(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 June 2005 (16.06.2005)

PCT

(10) International Publication Number
WO 2005/054473 A1

(51) International Patent Classification[7]: C12N 15/50, 15/05, 15/63, A61K 31/7088, 38/16, 39/215, A61P 31/14, 11/00, A61K 48/00

(21) International Application Number: PCT/CN2004/001419

(22) International Filing Date: 6 December 2004 (06.12.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/527,637 5 December 2003 (05.12.2003) US

(71) Applicant (*for all designated States except US*): **THE UNIVERSITY OF HONG KONG** [CN/CN]; G18, Eliot Hall, Pokfulam Road, Hong Kong (CN).

(72) Inventors: **CHYE, Meleen**; A2, Block 1, Tam Towers, 25 Sha Wan Drive, Pokfulam, Hong Kong (CN). **LI, Hongye**; South China Institute of Bontany, Chinese Academy of Sciences, Guangzhou 510650 (CN). **SATHISHKUMAR, Ramalingam**; Flat E, 24th Floor, Block 1, Elegant Garden, 409 Queen's Road West, Hong Kong (CN). **POON, Litman, Leo**; Flat H, 21/F, Block 2, Phase 3, Belvedere Gardens, Tsuen Wan, Hong Kong (CN). **PEIRIS, Sriyal, Malik, Joseph**; 19/F, Block 29, Baguio Villa, 550 Victoria Road, Pokfulam, Hong Kong (CN).

(74) Agent: **CHINA PATENT AGENT (H.K.) LTD.**; 22/F, Great Eagle Centre, 23 Harbour Road, Wanchai, Hong Kong (CN).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2005/054473 A1

(54) Title: GENETICALLY MODIFIED PLANTS COMPRISING SARS-CoV VIRAL NUCLEOTIDE SEQUENCES AND METHODS OF USE THEREOF FOR IMMUNIZATION AGAINST SARS

(57) Abstract: The invention relates to genetically modified plants and progeny thereof which constitute edible plant-derived mucosal vaccines and injectable plant-derived mucosal vaccines against Severe Acute Respiratory Syndrome (SARS). The invention relates to a recombinant vector that transforms specifically, but not limited to, the nuclei and/or plastids of tobacco, tomato and lettuce plants for antigen production. In specific embodiments, the plastid transformation vector expressing the nucleotide sequences are pCV1, pCV6 and pCV8, and their derivatives containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants. The present invention relates to methods of making the modified plants which comprises transformation of plants with vectors for nuclear expression and/or plastid expression of nucleotide sequences of the SARS-CoV virus, fragments, derivatives, analogs, or variants thereof. The present invention also relates to methods of immunization against SARS and methods of antibody detection using the SARS-CoV antigens generated by the plastid and/or nuclear vector(s) transformed plants.

GENETICALLY MODIFIED PLANTS COMPRISING SARS-CoV VIRAL NUCLEOTIDE SEQUENCES AND METHODS OF USE THEREOF FOR IMMUNIZATION AGAINST SARS

This application claims priority of U.S. Serial No. 60/527,637, filed December 5, 2003, the contents of which are hereby incorporated by reference into this application.

Throughout this application, certain publications are referenced. Full citations for these publications, as well as additional related references, may be found immediately preceding the claims. The disclosures of these publications are hereby incorporated by reference into this application in order to more fully describe the state of the art as of the date of the invention described and claimed herein.

BACKGROUND OF THE INVENTION

Recently, there has been an outbreak of atypical pneumonia in Guangdong province in mainland China. Between November 2002 and March 2003, there were 792 reported cases with 31 fatalities (WHO, Severe Acute Respiratory Syndrome (SARS), *Weekly Epidemiol. Rec.* (2003) Vol. 78, page 86). In response to this crisis, the Hospital Authority in Hong Kong has increased the surveillance on patients with severe atypical pneumonia. In the course of this investigation, a number of clusters of health care workers with the disease were identified. In addition, there were clusters of pneumonia incidents among persons in close contact with those infected. The disease was unusual in its severity and its progression in spite of the antibiotic treatment typical for the bacterial pathogens that are known to be commonly associated with atypical pneumonia. Some of the present inventors were one of the groups involved in the investigation of these patients. The disease was given the acronym Severe Acute Respiratory Syndrome ("SARS"), which is caused by a novel coronavirus (Peiris *et al.*, 2003, *Lancet* Vol. 61, pages 1319-25; Fouchier *et al.*, 2003, *Nature* Vol. 423, page 240) known as the SARS-CoV virus. The SARS-associated coronavirus is distinct from previously characterized members of the family *Coronaviridae* that cause respiratory and enteric diseases in animals including humans.

Plants play a critical role as nutrients for animals, including humans, and for the production of substances useful as pharmaceuticals, cosmetics and the like. The generation of transgenic plants by plant nuclear transformation has successfully produced mucosal vaccines against cholera, Norwalk virus, Hepatitis B, foot-and-mouth disease (Walmsley and Arntzen, 2003, *Curr. Opin. Biotech.* Vol. 14, pages 145-150). Viral antigens expressed in transgenic plants are effective in inducing mucosal and serum immune responses in animals, irrespective of parenteral or oral delivery (Daniell *et al.*, 2001, *Trends Plant Sci.* Vol. 6, pages 219-226).

Plastid transformation offers great benefits for recombinant DNA technology. In the plant cell, the mere presence of up to 10,000 more copies of the plastid genome as compared to a nuclear genome, ensures enhanced yield of the plastid-expressed foreign protein, or antigen in the case of vaccine production. Further, maternal inheritance of plastids would result in foreign gene containment due to the lack of pollen transmission (Daniell *et al.*, 2002, *Trends Plant Sci.* Vol. 7, pages 84-91). Human somatotropin accumulates to 7% total soluble protein, which is 300-fold greater than from nuclear-transformed tobacco (Staub *et al.*, 2000, *Plant Journal* Vol. 6, pages 547-553), while human serum albumin accumulates to 11.1% total soluble protein, which is 500-fold more than nuclear-transformed leaves (Fernandez-San Millan *et al.*, 2003, *Plant Biotech* Vol. 1, pages 71-79). The cholera toxin B subunit, the first plant-derived vaccine arising from plastid transformation, accumulates to 4.1% total soluble protein in tobacco leaves (Daniell *et al.*, 2001, *J. Mol. Biol.* Vol. 311, pages 1001-1009). In comparison, the yield of foreign protein from nuclear-transformed plants rarely exceeds 1% total soluble protein (Daniell *et al.*, 2001, *Trends Plant Sci.* Vol. 6, pages 219-226).

SUMMARY OF THE INVENTION

The present invention is based on the observation of the present inventors that genetically modified plants and progeny thereof expressing SARS-CoV viral antigens can be used as vaccines against SARS. When the genetically modified plants or progeny thereof are ingested by animals, or extracts from genetically modified plants or progeny thereof are injected into or ingested by animals, preferably human, antibodies are generated against the SARS-CoV viral antigens. An edible vaccine against the SARS-CoV virus, deliverable as fruits (*e.g.*, tomato), leaves (*e.g.*, lettuce), tubers (*e.g.*, potato), seeds (*e.g.*, rice or corn), flowers, stems or roots, would obliterate costly purification procedures required of recombinant vaccines from microorganisms. It has further advantages in easy storage, transport and administration by direct ingestion. Also, edible vaccines in the form of fruits, leaves, tubers, seeds, flowers, stems or roots can be grown and easily distributed in developing countries, cutting the costs of immunization programs, and omitting the need for refrigeration and delivery by injection which requires trained health personnel.

In accordance with the present invention, plant transformation vectors are engineered to provide SARS-CoV viral antigens for eliciting an immune response of an animal for the prevention and treatment of SARS. The antigen includes proteinaceous agent or molecules of the SARS-CoV virus, natural or artificial variants, or mutants thereof. The plant vectors comprising the SARS-CoV viral sequences may be engineered to provide one, two, three or more nucleotide sequences of the SARS-CoV virus. In accordance with the present invention, the antigenic sequences may be derived from the SARS-CoV virus, natural or artificial variants, or mutants thereof.

The present invention provides plant transformation vectors comprising SARS-CoV nucleotide sequences which encode fragments, derivatives, analogs, or variants of polypeptides of SARS-CoV virus. In a specific embodiment, the invention provides for transgenic and/or transplastomic tobacco plants. The present invention provides modified plants that comprise SARS-CoV antigens, including SARS-CoV viral polypeptides or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus. These activities include functional activities as well as immunogenic abilities. The present invention also provides a method of producing the modified plants which comprises transforming a plant with plastid and/or nuclear transformation vector comprising SARS-CoV viral sequences, fragments, derivatives, analogs, or variants thereof. In preferred embodiments, the present invention provides for modified plants that express less than the entire genome of the SARS-CoV virus. In preferred embodiments, the modified plants provide less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, less than 50%, less than 45%, less than 40%, less than 35%, less than 30%, less than 25%, less than 20%, less than 15%, less than 10%, less than 5%, or less than 1% of the entire genome of SARS-CoV virus. In preferred embodiments, the modified plants provide at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 polypeptides, fragments, derivatives, analogs and variants of the SARS-CoV virus.

In another specific embodiment, the invention provides a vaccine against SARS. The vaccine comprises the modified plant, its proteins, or extracts. The present invention also provides a method of immunization against SARS. The method comprises administering the modified plants to an animal so as to elicit the production of antibodies in the animal against the SARS-CoV viral antigens, which include SARS-CoV viral polypeptides, or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptide of the SARS-CoV virus.

In another specific embodiment, plant transformation vectors are engineered to provide SARS-CoV viral antigens for use as reagents in serological tests. One aspect of the invention relates to methods of detecting an antibody to a SARS-CoV viral antigen in a sample by incubating plant-derived SARS-CoV viral antigens with the sample and detecting the presence of antibodies bound to the plant-derived SARS-CoV viral antigens. The sample may be a biological fluid, such as blood, serum, plasma, saliva, urine, stool, sputum, nasopharyngeal aspirates, cells and tissues. Such plant-derived antigens would be cheaper alternative to those generated by infected cell lines for antibody detection.

In one embodiment, the invention relates to compositions comprising a SARS-CoV viral antigen produced by a transformed plant cell or plant. Methods of using the compositions to detect an antibody to a SARS-CoV viral antigen are also encompassed. In a specific embodiment, a method comprises contacting a sample with a composition comprising a SARS-CoV viral antigen produced by a transformed plant cell or plant, and detecting the presence of an antibody bound to the SARS-CoV viral antigen.

In one specific embodiment, the invention provides for a plastid transformation vector comprising the nucleotide sequences of the SARS-CoV virus encoding fragments, derivatives, analogs, or variants of polypeptides of SARS-CoV virus. In another embodiment, the invention provides for a plastid transformation vector which expresses SARS-CoV viral polypeptides, fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptide of the SARS-CoV virus. In a specific embodiment, the vector devoid of viral DNA is pMLVHisA. In a specific embodiment, the present invention provides a method of producing SARS-CoV viral antigens via the plastid transformation vector pCV1, pCV6 or pCV8, and its derivatives containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants.

In another specific embodiment, the invention provides for a nuclear transformation vector comprising the nucleotide sequences of the SARS-CoV virus, fragments, derivatives, analogs, or variants thereof. In another embodiment, the invention provides for a nuclear transformation vector which expresses SARS-CoV viral polypeptides, fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptide of the SARS-CoV virus. In a specific embodiment, the vector devoid of viral DNA is pSa7. In a specific embodiment, the present invention provides a method of producing SARS-CoV viral antigens via the nuclear transformation vector pCV2, pCV4 or pCV12, and its derivatives containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants.

In a specific embodiment, the present invention further provides transplastomic tobacco plants having therein a vector comprising a promoter derived from the rice plastid gene *psbA*, and a terminator. Plant cells containing a vector comprising a nucleotide sequence of the SARS-CoV virus is also an aspect of this invention. Plant parts of the modified plants, such as for example, fruits, leaves, tubers, seeds, flowers, stems or roots, which comprise cells expressing the SARS-CoV polypeptides, derivatives, analogs, or variants thereof are provided in the invention. The plant parts include parts that are separated from the whole plant or attached onto the whole plant. In a specific embodiment, the present invention further utilizes a selectable marker gene *aadA*, which specifies spectinomycin-resistance. In another embodiment, the present invention utilizes a start codon for recombinant $(His)_5$-

tagged protein and an *rbcL* terminator. In a preferred embodiment, the modified plant further comprises heterologous nucleotide sequences that express protease inhibitor protein.

In another specific embodiment, the invention provides for the construction of a nuclear transformation vector for expression of SARS-CoV antigens including polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus. In a specific embodiment, the present invention provides plant nuclear transformation vectors pCV2, pCV4 and pCV12.

In another specific embodiment, the invention provides for the construction of a plastid transformation vector for expression of SARS-CoV antigens including polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus. In a specific embodiment, the present invention provides plant plastid transformation vectors pCV1, pCV6 and pCV8.

In a specific embodiment, a nuclear transformation vector is used to express one or more SARS-CoV antigens, including polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus.

In a specific embodiment, a plastid transformation vector is used to express one or more SARS-CoV antigens, including polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus.

In a specific embodiment, a plastid transformation vector and a nuclear transformation vector are used to express one or more SARS-CoV antigens, including polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus.

The present invention provides a method of producing SARS-CoV viral antigens in plants, the antigens include polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus. The method comprises transforming a plant with a vector which comprises a promoter, operably associated with a coding sequence for one or more SARS-CoV viral polypeptides, and a terminator. Plant cells containing a vector which comprises one or more nucleic acid sequences encoding for the SARS-CoV viral antigens, including polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus, are also an aspect of this invention. Alternatively, the plant cells may contain one or more vectors of the present invention. The present invention provides plant parts, such as for example, fruits, leaves, tubers, seeds, flowers, stems, roots, and all other anatomical parts of the modified plant.

The present invention provides for a vaccine that elicits an immune response against SARS-CoV antigens, including SARS-CoV viral polypeptides, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus. In one specific embodiment, the invention provides a method of immunization against SARS. The method comprises ingesting an edible transformed plant parts comprising nucleic acid molecules of the SARS-CoV viral sequences encoding the SARS-CoV viral polypeptides. In another embodiment, the method comprises the consumption of the modified plant or intravenous injection or ingestion of an extract of the modified plant, which plant expresses SARS-CoV viral antigens, including SARS-CoV viral polypeptides, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus. In another specific embodiment, the present invention provides a method of immunization against SARS by feeding the modified plant of the present invention or injecting to or ingestion of an extract of the modified plant of the present invention by a rabbit, goat, cow, pig, sheep, horse, primate, civet, rodent, raccoon, raccoon dog, dog, ferret, ferret Badger, cat, avian, or any other species of animal, including human. The present invention further provides extracting antibodies from an immunized animal. The antibodies are immunospecific to SARS-CoV antigens, including viral polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A-1W show restriction sites of pCV1.

Figures 2A-2C show the nucleic acid sequence of pCV1 (SEQ ID NO:1).

Figures 3A-3B show restriction maps of pCV1 with specific restriction enzymes.

Figure 4 shows the pMLVHisA plastid transformation vector from which pCV1, pCV6 and pCV8 were derived. The flanking regions *rbcL* and *accD* are derived from the tobacco plastid genome for homologous recombination during plastid transformation. P_{psbA} represents the promoter for the expression of the inserted gene. T_{psbA} represents the terminator. P_{rrn} represents the promoter driving expression of the spectinomycin-resistance marker *aadA*. T_{rbcL} represents the *rbcL* terminator. Met represents the start codon for recombinant $(His)_5$-tagged protein. (A) shows the sequence (SEQ ID NO:2) and restriction sites of pCV8. (B) shows a restriction map of pCV8 with specific restriction enzymes. (C) shows a restriction map of pCV8 with specific restriction enzymes. (D) shows a restriction map of pCV8 with specific restriction enzymes.

Figure 5 shows the pSa7 nuclear transformation vector from which pCV2 and pCV4 were derived. RB and LB represent the right and left borders of T(transferred)-DNA for random insertion into the plant nuclear genome. NOS-Pro represents the nopaline synthase (NOS) promoter. NOS-ter represents the *NOS*-terminator. NPTII (Kan^R) represents neomycin phosphotransferase specifying kanamycin-resistance. CaMV35S-Pro represents the Cauliflower Mosaic Virus 35S promoter. SaPIN2a cDNA represents the cDNA encoding *Solanum americanum* proteinase inhibitor IIA.

Figures 6A-6I show the entire nucleic acid sequence of the SARS virus (SEQ ID NO:7).

Figures 7A-7LL show restriction sites of pCV8.

Figures 8A-8D show the nucleic acid sequence of pCV8 (SEQ ID NO:2).

Figure 9 shows a restriction map of pCV8 with specific restriction enzymes.

Figure 10 shows a restriction map of pCV8 with specific restriction enzymes.

Figure 11 shows a restriction map of pCV8 with specific restriction enzymes.

Figure 12 shows the pCV12 nuclear transformation vector for expression of a protein fusion consisting of the SARS-CoV S1 protein fused with green fluorescent protein (GFP). RB and LB represent the right and left borders of T(transferred)-DNA for random insertion into the plant nuclear genome. NOS-ter represents the *NOS*-terminator. CaMV35S-Pro represents the Cauliflower Mosaic Virus 35S promoter. S1 represents the SARS-CoV S1 protein and GFP, the green fluorescent protein.

Figure 13A-13B show the genomic DNA sequence of the M-gene (SEQ ID NO: 3), and the amino acid sequence of the M protein (SEQ ID NO:4), respectively.

Figures 14A-14B show the genomic DNA sequence of the S-gene (SEQ ID NO:5).

Figure 15 show the amino acid sequence of the S protein (SEQ ID NO:6).

Figures 16A-16NN show the deduced amino acid sequences obtained from SEQ ID NO:7 in three frames. An aster (*) indicates a stop codon which marks the end of a peptide.

Figures 17A-17NN show the deduced amino acid sequences obtained from the complement of SEQ ID NO:7 in three frames. An aster (*) indicates a stop codon which marks the end of a peptide.

Figures 18A-D show transient expression of S1:GFP in agroinfiltrated tobacco leaves. Representative tobacco leaf epidermal cells are shown by confocal microscopy 2 days following agroinfiltration (Yang Y. *et al.*, *In vivo* analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. *Plant J.* 2000; 22: 543-551) of *Agrobacterium tumefaciens* LBA4404 harboring plasmid pCV12 expressing S1:GFP fusion protein (A, C) or LBA4404 harboring pGDG expressing GFP alone (B, D). Bar represents 20μm.

Figures 19A-19XX show restriction sites of pCV2.

Figures 20A-20E show the nucleic acid sequence of pCV2 (SEQ ID NO:8).

Figures 21A-21NN show restriction sites of pCV6.

Figures 22A-22D show the nucleic acid sequence of pCV6 (SEQ ID NO:9).

Figures 23A-23I show plants obtained following transformation and their analysis. Figures 23A-23E show the regenerated plantlets obtained following *Agrobacterium*-mediated transformation of tobacco and lettuce using nuclear transformation vector pCV2 and analysis of these plantlets by PCR using primers 35S and NOS-ter followed by Southern blot analysis with a ^{32}P-radiolabeled S1 probe. Figures 23F-23I show the regenerated plantlets obtained following plastid transformation of tobacco using plastid transformation vector pCV1, PCR analysis of these plantlets using *S1* primers and Northern blot analysis with a ^{32}P-radiolabeled S1 probe. Figure 23A shows the plants regenerated from tobacco leaves after *Agrobacterium*-mediated transformation. Figure 23B shows the regenerated tobacco shoot used in the PCR analysis. Figure 23C shows the plants regenerated from lettuce cotyledons after *Agrobacterium*-mediated transformation. Figure 23D shows the regenerated lettuce shoot used in the PCR analysis. Figure 23E shows the presence of a 2.1-kb S1 hybridizing band (arrowed) in two independent tobacco lines (lanes 2 and 3), and two independent lettuce lines (lanes 4 and 5), and the absence of this band in wild-type tobacco (lane 6) and wild-type lettuce (lane 7). One other tobacco line tested negative (lane 1). Figure 23F and 23G show the plantlets regenerated from tobacco after particle bombardment. Figure 23H shows the specific 0.7-kb PCR band in regenerated tobacco (lane 2) which is absent in wild-type (lane 1). Figure 23I shows the presence of a 2.1-kb hybridizing *S1* mRNA band (arrowed) in a tobacco line (lane 2) which is absent in wild-type (lane 1).). Figure 23J shows western blot analysis of transplastomic tobacco expressing S1 using Ni-NTA conjugate in detection of His-tagged S1 protein. The arrow indicates the expected 73-kDa $(His)_5$-S1 band.

Figure 24 shows a Western blot analysis using antibodies against GFP show transient expression of S1:GFP in tobacco leaves following agroinfiltration. Total protein (200 μg) extracted from tobacco leaves, infiltrated with plasmid pGDG expressing GFP alone (lane 1) or plasmid pCV12 expressing S1:GFP fusion (lane 2), were separated on a 8% SDS-PAGE gel, blotted onto Hybond-C filters according to Sambrook *et al.* (1989. Molecular Cloning: A

Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor), and cross-reacted with antibodies against GFP (Clontech). Arrow indicates cross-reacting S1:GFP band (calculated size 99.1 kDa). M, molecular mass marker.

DETAILED DESCRIPTION OF THE INVENTION

DEFINITIONS OF TERMS:

As used herein, the term "modified plant or plant parts" refers to a plant or plant part, whether it is attached or detached from the whole plant. It also includes progeny of the modified plant or plant parts that are produced through sexual or asexual reproduction.

As used herein, the term "variant" refers either to a naturally occurring genetic mutant of SARS-CoV or a recombinantly prepared variation of the SARS-CoV virus, each of which contain one or more mutations in its genome compared to the SARS-CoV virus of a deposited virus, CCTCC-V200303, which sequence is shown in Figure 6A-6I (SEQ ID NO:7). The term "variant" may also refer to either a naturally occurring variation of a given peptide or a recombinantly prepared variation of a given peptide or protein in which one or more amino acid residues have been modified by amino acid substitution, addition, or deletion. A natural variant of SARS-CoV has a sequence that is different from the genomic sequence of the SARS-CoV virus due to one or more naturally occurred mutations, including, but not limited to, point mutations, rearrangements, insertions, deletions, etc., to the genomic sequence that may or may not result in a phenotypic change. Preferably, the variants include less than 25, less than 20, less than 15, less than 10, less than 5, less than 4, less than 3, or less than 2 amino acid substitutions, rearrangements, insertions, and/or deletions relative to the SARS-CoV virus.

In preferred embodiments, the variants have conservative amino acid substitutions that are made at one or more predicted non-essential amino acid residues (*i.e.*, amino acid residues which are not critical for the expression of the biological activities of the virus, *e.g.*, infectivity, replication ability, protein synthesis ability, assembling ability, and cytotoxic effect). A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a side chain with a similar charge. Families of amino acid residues having side chains with similar charges have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine), and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). In another embodiment, the variants have non-conservative amino acid substitution, *i.e.*, amino acid residues are replaced by an amino acid that does not have a side chain with a similar charge.

In a specific embodiment, the SARS-CoV sequences that may be used in the present invention include those deposited with GenBank® having accession nos. NC_004718, AY304495, AY304494, AY304493, AY304492, AY304491, AY304490, AY304489, AY304488, AY304487, AY304486, AY360146, AY278491, AY310120, AY278489, AY362699, AY362698, AY283798, AY283797, AY283796, AY283795, AY283794, AY268070, AY278741, AY340092, AY351680, AP006561, AP006560, AP006559, AP006558, AP006557, AY278554, AY348314, AY338175, AY338174, AY323977, AY322199, AY322198, AY322197, AH013000, AY322208, AY322207, AY322206, AY322205, AH012999, AY321118, AY323976, AY323975, AY323974, AY286320, AY290752, AY291315, AY307165, AY279354, AY278490, AY278487, AY297028, AY286402, AY274119, AY291451, AY271716, AY282752, AY278488, AY268049, AY269391, all of which are incorporated herein by reference in their entireties.

The SARS-CoV viral nucleotide sequences used in the present invention may be derived from a mutant SARS-CoV virus. Mutations can be introduced randomly along all or part of the coding sequence of the SARS-CoV virus or variants thereof, such as by saturation mutagenesis, and the resultant mutants can be screened for biological activity to identify mutants that retain activity. In preferred embodiments, the mutant polypeptides do not retain activity of the wild-type polypeptide. In specific embodiments, the mutant polypeptides do not retain the virulent activity of the SARS-CoV virus. Techniques for mutagenesis known in the art can also be used, including but not limited to, point-directed mutagenesis, chemical mutagenesis, *in vitro* site-directed mutagenesis, using, for example, the QuikChange® Site-Directed Mutagenesis Kit (Stratagene), etc. Non-limiting examples of such modifications include substitutions of amino acids to cysteines toward the formation of disulfide bonds; substitution of amino acids to tyrosine and subsequent chemical treatment of the polypeptide toward the formation of dityrosine bonds, as disclosed in detail herein; one or more amino acid substitutions and/or biological or chemical modification toward generating a binding pocket for a small molecule (substrate or inhibitor), and/or the introduction of side-chain specific tags (*e.g.*, to characterize molecular interactions or to capture protein-protein interaction partners). Biological modifications that are useful in the present invention comprises alkylation, phosphorylation, sulfation, oxidation or reduction, ADP-ribosylation, hydroxylation, glycosylation, glucosylphosphatidylinositol addition, ubiquitination, etc. Chemical modifications that are useful in the present invention comprise, *e.g.*, altering the charge of the recombinant virus. A positive or negative charge is chemically added to an amino acid residue where a charged amino acid residue is modified to an uncharged residue.

As used herein, the terms "antibody" and "antibodies" refer to monoclonal antibodies, bispecific antibodies, multispecific antibodies, human antibodies, humanized antibodies, chimeric antibodies, camelised antibodies, single domain antibodies, single-chain Fvs (scFv), single chain antibodies, Fab fragments, F(ab') fragments, disulfide-linked Fvs (sdFv), and anti-idiotypic (anti-Id) antibodies (including, *e.g.*, anti-Id antibodies to antibodies of the invention), and epitope-binding fragments of any of the above. In particular, antibodies include immunoglobulin molecules and immunologically active fragments of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site. Immunoglobulin molecules can be of any type (*e.g.*, IgG, IgE, IgM, IgD, IgA and IgY), class (*e.g.*, IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) or subclass.

As used herein, the term "an antibody that immunospecifically binds a polypeptide of the SARS-CoV virus" refers to an antibody that immunospecifically binds to the polypeptide encoded by SARS-CoV virus and does not non-specifically bind to other polypeptides. An antibody that immunospecifically binds to the polypeptide of the SARS-CoV virus does not cross-react with other antigens. Preferably, an antibody that immunospecifically binds to a polypeptide of the SARS-CoV virus does not cross-react with other antigens. An antibody that immunospecifically binds to the polypeptide of the SARS-CoV virus can be identified by, for example, immunoassays or other techniques known to those skilled in the art.

As used herein, the term "epitope" refers to a fragment of a SARS-CoV virus, polypeptide or protein having antigenic or immunogenic activity in an animal, preferably a mammal, and most preferably in a human. An epitope having immunogenic activity is a fragment of a polypeptide that elicits an antibody response in an animal. An epitope having antigenic activity is a fragment of a polypeptide or protein to which an antibody immunospecifically binds as determined by any method well known in the art, for example, by the immunoassays described herein. Antigenic epitopes need not necessarily be immunogenic.

As used herein, the term "antigenicity" refers to the ability of a substance (*e.g.*, foreign objects, microorganisms, drugs, antigens, proteins, peptides, polypeptides, nucleic acids, DNA, RNA, etc.) to trigger an immune response in a particular organism, tissue, and/or cell. Sometimes, the term "antigenic" is synonymous with the term "immunogenic".

As used herein, the term "immunogenicity" refers to the property of a substance (*e.g.*, foreign objects, microorganisms, drugs, antigens, proteins, peptides, polypeptides, nucleic acids, DNA, RNA, etc.) being able to evoke an immune response within an organism. Immunogenicity depends partly upon the size of the substance in question and partly upon how unlike the host molecules is the substance. Highly conserved proteins tend to have rather low immunogenicity.

As used herein, the term "hybridizes under stringent conditions" describes conditions for hybridization and washing under which nucleotide sequences having at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% identity to each other typically remain hybridized to each other. Such hybridization conditions are described in, for example but not limited to, Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6.; Basic Methods in Molecular Biology, Elsevier Science Publishing Co., Inc., N.Y. (1986), pp. 75-78, and 84-87; and Molecular Cloning, Cold Spring Harbor Laboratory, N.Y. (1982), pp. 387-389, and are well known to those skilled in the art. The conditions under which hybridization and/or washing can be carried out can range from 42-68°C and the washing buffer can comprise from 0.1X sodium chloride/sodium citrate (SSC), 0.5% SDS to 6X SSC, 0.5% SDS. Typically, hybridization can be carried out overnight at 65°C (high stringency conditions), 60°C (medium stringency conditions), or 55°C (low stringency conditions). The filters can be washed for 2x15 minutes with 0.1X SSC, 0.5% SDS at 65°C (high stringency washing). The filters were washed for 2x15 minutes with 0.1XSSC, 0.5% SDS at 63°C (medium stringency washing). For low stringency washing, the filters were washed at 60°C for 2x15 minutes at 2X SSC, 0.5% SDS. A preferred, non-limiting example of stringent hybridization conditions is hybridization in 6X SSC, 0.5% SDS at about 68°C followed by one or more washes in 2X SSC, 0.5% SDS at room temperature. Another preferred, non-limiting example of stringent hybridization conditions is hybridization in 6X SSC at about 45°C followed by one or more washes in 0.2X SSC, 0.1% SDS at about 50-65°C.

An "isolated" or "purified" antibody is substantially free of cellular material or other contaminating proteins from the biological fluid from which the antibody is derived. The language "substantially free of cellular material" includes preparations of an antibody in which the antibody is separated from cellular components of the cells from which it is isolated. Thus, an antibody that is substantially free of cellular material includes preparations of the antibody having less than about 30%, 20%, 10%, 5%, 2.5%, or 1%, (by dry weight) of contaminating protein. In a preferred embodiment of the present invention, the antibody is isolated or purified.

As used herein, the term "having a biological activity of the polypeptides of the SARS-CoV virus" refers to the characteristics of the polypeptides or proteins having a common biological activity similar or identical structural domain and/or having sufficient amino acid identity to SARS polypeptides or the polypeptides having the amino acid sequences as shown in Figures 13B, 15, 16A-16NN, and 17A-17NN, or a fragment thereof. Such common biological activities of the polypeptides of the invention include antigenicity and immunogenicity.

As used herein, the term "molecules having the similar activities as the polypeptide of the SARS-CoV virus" refers to polypeptides or proteins having similar biological activities and similar or identical structural domain and/or having sufficient amino acid identity to the polypeptides of the SARS-CoV virus. Such biological activities can include, but are not limited to, antigenicity, immunogenicity, cytotoxicity, hormonal activities, binding properties and affinities, pharmacological activities, stimulation or inhibition of growth proliferation and differentiation, induction of changes in cells, antiviral, antibacterial, antifungal and antiparasitic activities, etc. In preferred embodiments, polypeptides that are useful for the present invention may retain at least one, two, three, four, five, or more biological activities of the wild-type SARS-CoV virus (*e.g.*, infectivity, replication ability, protein synthesis ability, assembling ability, and cytotoxic effect).

As used herein, the term "portion" or "fragment" refers to a fragment of a nucleic acid molecule containing at least about 25, 30, 35, 40, 45, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, 1,500, 2,000, 2,500, 3,000, 3,500, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 15,000, 20,000, 25,000, or more contiguous nucleic acids in length of the relevant nucleic acid molecule and having at least one functional feature of the nucleic acid molecule (or the encoded protein has one functional feature of the protein encoded by the nucleic acid molecule); or a fragment of a protein or a polypeptide containing at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 450, 500, 600, 700, 800, 900, 1,000, 2,000, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, or more contiguous amino acid residues in length of the relevant protein or polypeptide and having at least one functional feature of the protein or polypeptide.

As used herein, the term "analog" in the context of proteinaceous agent (*e.g.*, proteins, polypeptides, peptides, and antibodies) refers to a proteinaceous agent that possesses a similar or identical function as a second proteinaceous agent but does not necessarily comprise a similar or identical amino acid sequence of the second proteinaceous agent, or possess a similar or identical structure of the second proteinaceous agent. In a specific embodiment, antibody analogs immunospecifically bind to the same epitope as the original antibodies from which the analogs were derived. In an alternative embodiment, antibody analogs immunospecifically bind to different epitopes than the original antibodies from which the analogs were derived. A proteinaceous agent that has a similar amino acid sequence refers to a second proteinaceous agent that satisfies at least one of the following: (a) a proteinaceous agent having an amino acid sequence that is at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% identical to the amino acid sequence of a second proteinaceous agent; (b) a proteinaceous agent encoded by a nucleotide sequence that hybridizes under stringent conditions to a nucleotide sequence encoding a second proteinaceous agent of at least 5 contiguous amino acid residues, at least 10 contiguous amino acid residues, at least 15 contiguous amino acid residues, at least 20 contiguous amino acid residues, at least 25 contiguous amino acid residues, at least 40 contiguous amino acid residues, at least 50 contiguous amino acid residues, at least 60 contiguous amino residues, at least 70 contiguous amino acid residues, at least 80 contiguous amino acid residues, at least 90 contiguous amino acid residues, at least 100 contiguous amino acid residues, at least 125 contiguous amino acid residues, or at least 150 contiguous amino acid residues; and (c) a proteinaceous agent encoded by a nucleotide sequence that is at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% identical to the nucleotide sequence encoding a second proteinaceous agent. A proteinaceous agent with similar structure to a second proteinaceous agent refers to a proteinaceous agent that has a similar secondary, tertiary or quaternary structure to the second proteinaceous agent. The structure of a proteinaceous agent can be determined by methods known to those skilled in the art, including but not limited to, peptide sequencing, X ray crystallography, nuclear magnetic resonance, circular dichroism, and crystallographic electron microscopy.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino acid or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % identity = number of identical overlapping positions/total number of positions x 100%). In one embodiment, the two sequences are the same length.

The determination of percent identity between two sequences can also be accomplished using a mathematical algorithm. A preferred, non limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul, 1990, *Proc. Natl. Acad. Sci. U.S.A.* Vol. 87, pages 2264-2268, modified as in Karlin and Altschul, 1993, *Proc. Natl. Acad. Sci. U.S.A.* Vol. 90, pages 5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul *et al.*, 1990, *J. Mol. Biol.* Vol. 215, page 403. BLAST nucleotide searches can be performed with the NBLAST nucleotide program parameters set, *e.g.*, for score=100, wordlength=12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the present invention. BLAST protein searches can be performed with the XBLAST program parameters set, *e.g.*, to score 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecule of the present invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, 1997, *Nucleic Acids Res.* Vol. 25, pages 3389-3402. Alternatively, PSI BLAST can be used to perform an iterated search which detects distant relationships between molecules (Id.). When utilizing BLAST, Gapped BLAST, and PSI Blast programs, the default parameters of the respective programs (*e.g.*, of XBLAST and

NBLAST) can be used (see, *e.g.*, the NCBI website). Another preferred, non limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, 1988, *CABIOS* Vol. 4, pages 11-17. Such an algorithm is incorporated in the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically only exact matches are counted.

As used herein, the term "analog" in the context of a non-proteinaceous analog refers to a second organic or inorganic molecule which possesses a similar or identical function as a first organic or inorganic molecule and is structurally similar to the first organic or inorganic molecule.

As used herein, the term "derivative" in the context of proteinaceous agent (*e.g.*, proteins, polypeptides, peptides, and antibodies) refers to a proteinaceous agent that comprises an amino acid sequence which has been altered by the introduction of amino acid residue substitutions, deletions, and/or additions. The term "derivative" as used herein also refers to a proteinaceous agent which has been modified, *i.e.*, by the covalent attachment of any type of molecule to the proteinaceous agent. For example, but not by way of limitation, an antibody may be modified, *e.g.*, by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. A derivative of a proteinaceous agent may be produced by chemical modifications using techniques known to those of skill in the art, including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic synthesis of tunicamycin, etc. Further, a derivative of a proteinaceous agent may contain one or more non-classical amino acids. A derivative of a proteinaceous agent possesses a similar or identical function as the proteinaceous agent from which it was derived.

As used herein, the term "derivative" in the context of a non-proteinaceous derivative refers to a second organic or inorganic molecule that is formed based upon the structure of a first organic or inorganic molecule. A derivative of an organic molecule includes, but is not limited to, a molecule modified, *e.g.*, by the addition or deletion of a hydroxyl, methyl, ethyl, carboxyl or amine group. An organic molecule may also be esterified, alkylated and/or phosphorylated.

As used herein, the terms "subject" and "patient" are used interchangeably. As used herein, the terms "subject" and "subjects" refer to an animal, preferably a mammal including a non-primate (*e.g.*, rabbits, goats, cows, pigs, sheep, horses, civets, rodents, raccoons, raccoon dogs, dogs, ferrets, ferret Badger, cats, and avian species) and a primate (*e.g.*, monkeys such as a cynomolgous monkey and humans), and more preferably a human.

As used herein, the terms "SARS" or "SARS related symptoms" include various clinical indications and classifications. As described herein, SARS include (1) asymptomatic or mild respiratory illness; (2) moderate respiratory illness; or (3) severe respiratory illness.

As used herein, the term "carrier" is a substance used to support or convey another substance such as a pigment, catalyst, or radioactive material.

As used herein, the term "vehicle" is a substance that facilitates the use of a drug, pigment, or other material mixed with it.

As used herein, the term "excipients" refers to inert substances which are commonly used as a diluent, vehicle, preservatives, binders, or stabilizing agent for drugs and includes, but not limited to, proteins (*e.g.*, serum albumin, etc.), amino acids (*e.g.*, aspartic acid, glutamic acid, lysine, arginine, glycine, histidine, etc.), fatty acids and phospholipids (*e.g.*, alkyl sulfonates, caprylate, etc.), surfactants (*e.g.*, SDS, polysorbate, nonionic surfactant, etc.), saccharides (*e.g.*, sucrose, maltose, trehalose, etc.) and polyols (*e.g.*, mannitol, sorbitol, etc.). Also see Remington's Pharmaceutical Sciences (by Joseph P. Remington, 18th ed., Mack Publishing Co., Easton, PA), which is hereby incorporated in its entirety.

As used herein, the term "operably associated" or "operably linked" refers to an association in which the regulatory regions (*e.g.*, promoter, enhancer) and the nucleic acid sequence to be expressed are covalently joined and positioned in such a way as to permit transcription, and under the appropriate condition, translation.

PREFERRED EMBODIMENTS AND EXPERIMENTAL DETAILS:

The present invention provides transgenic and transplastomic plants and their progeny that are generated by plastid and/or nuclear transformation vectors thereby producing a mucosal vaccine against SARS-CoV viral antigens. Viral antigens expressed in transgenic plants are effective in inducing mucosal and serum immune responses in animals, irrespective of parenteral or oral delivery. Viral antigens expressed in transgenic plants are also useful as reagents for antibody detection in serological tests. The present invention seeks to transform various types of plants with plant vectors comprising nucleotide sequences from the SARS-CoV virus including but not limited to nucleotide sequences of the SARS-CoV virus which encodes fragments, derivatives, analogs, or variants of polypeptides of SARS-CoV virus, or that it encodes an epitope or a proteinaceous molecule having similar activities as the polypeptides, fragments, derivatives, analogs, or variants thereof, of the SARS-CoV virus.

Analysis of the genome sequence of the SARS-associated coronavirus revealed that four of its eleven predicted open reading frames encode structural proteins including a spike glycoprotein (S) of 1255 deduced amino acids and a membrane glycoprotein (M) of 221 deduced amino acids; S and M presumably associate to form the viral envelope. The S glycoprotein, a type I membrane protein, functions in attachment to host receptors and has an N-terminal signal sequence (S1; amino acids 1-658) and a C-terminal transmembrane region followed by a cytoplasmic tail (Marra *et al.*, 2003; Rota *et al*, 2003). The N-terminal region of M also includes an uncleavable signal sequence and is exposed to the surface of the virus while its C-terminus resides within the viral membrane. While both S and M are believed to be targeted to the plasma membrane, M was also believed to be localized at the endoplasmic reticulum in plant cells. In preferred embodiments, the modified plant or plant parts or progeny of the modified plant comprise SARS-CoV viral sequences that encode the spike protein (S or S1) of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof. In a preferred embodiment, the modified plant or plant parts or progeny of the modified plant comprises nucleotide sequences that encode the N-terminal region (S1) of the spike protein (S) and/or the membrane protein (M) of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecule having the similar activities as the polypeptide of the SARS-CoV virus. In preferred embodiments, the plant or plant parts comprises the nucleotide sequence of SEQ ID NO:3, 5, or 7, or a fragment thereof, or that immunospecifically binds to the polypeptide having the nucleotide sequence of SEQ ID

NO:3, 5, or 7, or an analog, derivative, or fragment thereof, and/or polypeptides encoded by the nucleotide sequence of SEQ ID NO:3, 5, or 7, or a fragment thereof. The present invention also relates to a method of producing modified plants comprising one or more plastid and/or nuclear transformation vectors comprising one or more SARS-CoV nucleotide sequences which encode fragments, derivatives, analogs or variants of polypeptides of SARS-CoV virus. All varieties of plants may be used for the present invention. In a preferred embodiment, the plant is tobacco, lettuce, potato, tomato, banana, corn, rice, cereals, wheat, maize, barley, apple, pear, strawberry, carrot, sugar beets, yam, kiwifruit, or spinach.

CONSTRUCTION OF TRANSFORMATION VECTORS

The present invention relates to plant vectors comprising one or more nucleotide sequences of the SARS-CoV virus, or fragments, derivatives, analogs or variants thereof. In a preferred embodiment, the plant vectors are plastid and nuclear transformation vectors. In preferred embodiments, the nucleotide sequence is the M-gene and/or S-gene of the SARS-CoV virus. The present invention also relates to the construction of plastid and nuclear transformation vectors comprising one or more SARS-CoV viral nucleotide sequences, or fragments, derivatives, analogs or artificial or natural variants thereof. In specific embodiments, the invention provides plant transformation vectors that comprise one or more nucleotide sequences that are at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, 2,000, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, 10,000, 11,000, 12,000, 13,000, 14,000, 15,000, 16,000, 17,000, 18,000, 19,000, 20,000, 21,000, 22,000, 23,000, 24,000, 25,000, 26,000, 27,000, 28,000, 29,000, or more contiguous nucleotides of SEQ ID NO:7, or a complement thereof, or a nucleotide comprising a nucleotide sequence that hybridizes to the nucleotide sequences of the SARS-CoV virus under stringent conditions. In specific embodiments, the invention provides plant transformation vectors that comprise one or more nucleotide sequences that are at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, 300, 350, 400, 450, 500, 550, 600, 650, or more contiguous nucleotides of SEQ ID NO:3, or a nucleotide comprising a nucleotide sequence that hybridizes to the nucleotide sequences of the SARS-CoV virus under stringent conditions. In specific embodiments, the invention provides plant transformation vectors that comprise one or more nucleotide sequences that are at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, 2,000, 3,000, 3,500, or more contiguous nucleotides of SEQ ID NO:5, or a nucleotide comprising a nucleotide sequence that hybridizes to the nucleotide sequences of the SARS-CoV virus under stringent conditions. In specific embodiments, the invention provides plant transformation vectors that comprise one or more nucleotide sequences that encode polypeptides that are at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, or more contiguous amino acids of SEQ ID NO:4. In specific embodiments, the invention provides plant transformation vectors that comprise one or more nucleotide sequences that encode polypeptides that are at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 100, 150, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, or more contiguous amino acids of SEQ ID NO:6. In one embodiment, the plant vector comprises nucleotide sequences that encode amino acids 14-1195 of S. In preferred embodiments, the plant vector comprises nucleotide sequences that encode at least 20, 30, 40, 50, 60, 80, 100, 200, 300, 400, 500, or more contiguous amino acids in length of amino acids 14-1195 of S. In specific embodiments, the polypeptides are those shown in Figures 16A-16NN and 17A-17NN, or fragments, derivatives or analogs thereof. In specific embodiments, the invention provides plant transformation vectors that comprise one or more nucleotide sequences that encode any of the genes, or portions thereof, or variants, fragments, analogs, or derivatives, from the SARS-CoV virus. These genes include, but are not limited to, envelop protein (E protein), integral membrane protein (M protein), spike protein (S protein), nucleocapsid protein (N protein), hemagglutinin esterase (HE protein), and RNA-dependent RNA polymerase. Particularly useful are also those proteinaceous substances that are encoded by recombinant nucleic acid fragments of the SARS-CoV genome. Preferred are those that are within the open reading frames (ORFs), in particular, for eliciting SARS-CoV specific antibody or T cell responses, whether *in vivo* (*e.g.*, for protective or therapeutic purposes or for providing diagnostic antibodies) or *in vitro* (*e.g.*, by phage display technology or another technique useful for generating synthetic antibodies).

The viral vectors can be engineered to provide antigenic molecules, including nucleotide sequences, or polypeptides of the SARS-CoV virus, including recombinant and chimeric forms of the virus, or subunits of the virus, or fragments, derivatives, analogs, or variants thereof, or molecules having similar activities as the polypeptides of the SARS-CoV virus. The present invention further provides methods of preparing recombinant or chimeric forms of SARS-CoV antigen.

In another specific embodiment, the present invention provides methods for treating, ameliorating, managing, or preventing SARS by administering a vaccine preparations or antibodies of SARS-CoV virus alone or in combination with antivirals (*e.g.*, amantadine, rimantadine, zanamivir, abacavir, combivir, emtricitabine, tenofovir, trizivir, enfuvirtide, gancyclovir, acyclovir, ribavirin, penciclovir, oseltamivir, foscarnet zidovudine (AZT), didanosine (ddI), lamivudine (3TC), zalcitabine (ddC), stavudine (d4T), nevirapine, delavirdine, efavirenz, fosamprenavir, amprenavir, atazanavir, kaletra, indinavir, ritonavir, vidarabine, nelfinavir, saquinavir, relenza, tamiflu, pleconaril, interferons, etc.), steroids and corticosteroids such as prednisone, cortisone, fluticasone and glucocorticoid, antibiotics, analgesics, bronchodialaters, or other treatments for respiratory and/or viral infections. In another specific embodiment, the methods of the present invention provides the use of herbs, herbal extracts, Chinese medicine and other remedies in combination with the modified plants of the present invention for the prevention and treatment of SARS.

Furthermore, the present invention provides pharmaceutical compositions comprising anti-viral agents of the present invention and a pharmaceutically acceptable carrier, vehicle or excipient. In addition, the present invention provides pharmaceutical compositions comprising liposomally encapsulated plant extracts, preferably purified, or anti-viral agents of the present invention. The present invention also provides kits comprising the pharmaceutical compositions of the present invention.

The present invention encompasses the use of a nucleotide sequence that encodes a chimeric polypeptide of the SARS-CoV virus. In a specific embodiment, the chimeric polypeptide comprises amino acid sequences from two or more different strains of the SARS-CoV virus. In accordance with the present invention, the modified plants of the present invention further comprise a nucleotide sequence that is non-native to the viral genome. The nucleotide sequence that may be useful for the present invention includes a portion of the SARS-CoV viral sequence which further comprises a heterologous nucleotide sequence. A heterologous nucleotide sequence may be from a virus, bacteria, animal, or plant. In a preferred embodiment, the heterologous nucleotide sequence encodes a proteinase inhibitor. The heterologous nucleotide sequence renders the expressed SARS-CoV polypeptide more stable and reduces programmed cell death and increases shelf life of the modified plant of the present invention. In a specific embodiment, the heterologous nucleotide sequence (*e.g.*, *SaPIN2a* or *SaPIN2b*) renders the expressed SARS-CoV polypeptides more stable so that the modified plants that express the SARS-CoV polypeptide may be processed before ingestion, injection, or other methods of administration. *SaPIN2a* and *SaPIN2b*, which are proteinase inhibitor II genes isolated from *Solanum americanum*, a weed belonging to the Solanaceae family, encode serine proteinase inhibitor II proteins which confer insect resistance in transgenic plants (See U.S. Provisional Application No. 60/429,992 filed November 29, 2002; and U.S. Application No. 10/725,829, Attorney Docket No. 9661-043-999, filed December 1, 2003, each of which is incorporated by reference herein in its entirety). SaPIN2a and SaPIN2b are highly expressed in the phloem and have possible involvement in regulating proteolysis in the sieve elements (See, Xu *et al.*, 2001, *Plant Mol Biol.* Vol. 47, pages 727-738; and Xu *et al.*, 2003, *Planta* Vol. 218, pages 623-629, each of which is incorporated by reference herein in its entirety).

In certain embodiments, the present invention relates to vectors and nucleic acid molecules comprising the nucleotide sequence that encodes a chimeric polypeptide of the SARS-CoV virus. In a specific embodiment, a vector comprises a heterologous nucleotide sequence of SARS-CoV. In another embodiment, such heterologous nucleotide sequences have been added, inserted or substituted for native or non-native sequences. In accordance with the present invention, the nucleotide sequence of the SARS-CoV virus may be derived from different strains or variants of SARS-CoV virus.

A plant vector comprising chimeric SARS-CoV viral sequences may be of particular use for the generation of recombinant vaccines protecting against two or more viruses (Tao *et al.*, *J. Virol.* 72: 2955-2961; Durbin *et al.*, 2000, *J. Virol.* 74: 6821-6831; Skiadopoulos *et al.*, 1998, *J. Virol.* 72: 1762-1768 (1998); Teng *et al.*, 2000, *J. Virol.* 74: 9317-9321). For example, it can be envisaged that a plant vector comprising the SARS-CoV viral nucleotide sequences may express one or more peptides from variants of the SARS-CoV virus, and will protect a subject against infections by both the native SARS-CoV and the variant.

In accordance with the present invention, the plant vectors can be engineered to provide antigenic sequences which confer protection against infection by the SARS-CoV and natural variants thereof when ingested by a subject. The plant vectors may be engineered to provide one or more antigenic sequences of SARS-CoV virus. In accordance with the present invention, the antigenic sequences may be derived from the same virus, from different strains or variants of the same type of virus, or from different viruses.

The invention provides a host cell comprising a vector according to the invention. Plant plastid or nuclear transformation vector containing the nucleotide sequences of the SARS-CoV virus such as containing the full-length, portions or fragments of the SARS-CoV genome for the expression of SARS-CoV viral nucleic acids. These plant vectors may contain other sequences for the generation of chimeric SARS-CoV viral polypeptides which may contain mutations, deletions or insertions of the SARS-CoV polypeptides. Nucleotide sequences which encode SARS-CoV virus, fragments, derivatives, analogs, or variants of polypeptides of SARS-CoV virus include, but are not limited to, those deposited with GenBank® having accession nos. NC_004718, AY304495, AY304494, AY304493, AY304492, AY304491, AY304490, AY304489, AY304488, AY304487, AY304486, AY360146, AY278491, AY310120, AY278489, AY362699, AY362698, AY283798, AY283797, AY283796, AY283795, AY283794, AY268070, AY278741, AY340092, AY351680, AP006561, AP006560, AP006559, AP006558, AP006557, AY278554, AY348314, AY338175, AY338174, AY323977, AY322199, AY322198, AY322197, AH013000, AY322208, AY322207, AY322206, AY322205, AH012999, AY321118, AY323976, AY323975, AY323974, AY286320, AY290752, AY291315, AY307165, AY279354, AY278490, AY278487, AY297028, AY286402, AY274119, AY291451, AY271716, AY282752, AY278488, AY268049, AY269391, all of which are incorporated herein by reference in their entireties.

In one specific embodiment, plant cells may be transiently or stably expressing one or more nucleotide sequences of the SARS-CoV virus. Plants cells are modified by transfection (proteins or nucleic acid vectors), infection (viral vectors) or transduction (viral vectors). The modified plant may be used to modulate a subject's immune system by stimulating a humoral immune response, a cellular immune response or by stimulating tolerance to an antigen. As used herein, a subject means: humans, primates, rabbits, goats, cows, pigs, sheep, horses, civets, rodents, raccoons, raccoon dogs, dogs, ferrets, ferret Badger, cat, cats, avian species, or other non-human animals.

The nucleotide molecules encoding the SARS-CoV viral antigen proteins may be cloned by amplification. The term "amplified" refers to the process of making multiple copies of the nucleic acid from a single polynucleotide molecule. The amplification of polynucleotides can be carried out *in vitro* by biochemical processes known to those of skill in the art. The amplification agent may be any compound or system that will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, *Taq* polymerase, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, polymerase mutants, reverse transcriptase, ligase, and other enzymes, including heat-stable enzymes (*i.e.*, those enzymes that perform primer extension after being subjected to temperatures sufficiently elevated to cause denaturation). Suitable enzymes will facilitate combination of the nucleotides in the proper manner to form the primer extension products that are complementary to each mutant nucleotide strand. Generally, the synthesis will be initiated at the 3'-end of each primer and proceed in the 5'-direction along the template strand, until synthesis terminates, producing molecules of different lengths. There may be amplification agents, however, that initiate synthesis at the 5'-end and proceed in the other direction, using the same process as described above. In any event, the method of the invention is not to be limited to the embodiments of amplification described herein.

One method of *in vitro* amplification, which can be used according to this invention, is the polymerase chain reaction (PCR) described in U.S. Patent Nos. 4,683,202 and 4,683,195. The term "polymerase chain reaction" refers to a method for amplifying a DNA base sequence using a heat-stable DNA polymerase and two oligonucleotide primers, one complementary to the (+)-strand at one end of the sequence to be amplified and the other complementary to the (-)-strand at the other end. Because the newly synthesized DNA strands can subsequently serve as additional templates for the same primer sequences, successive rounds of primer annealing, strand elongation, and dissociation produce rapid and highly specific amplification of the desired sequence. The polymerase chain reaction is used to detect the presence of polynucleotides encoding cytokines in the sample. Many polymerase chain methods are known to those of skill in the art and may be used in the method of the invention. For example, DNA can be subjected to 30 to 35 cycles of amplification in a thermocycler as follows: 95°C for 30 sec, 52-60°C for 1 min, and 72°C for 1 min, with a final extension step of 72°C for 5 min. For another example, DNA can be subjected to 35 polymerase chain reaction cycles in a thermocycler at a denaturing temperature of 95°C for 30 sec, followed by varying annealing temperatures ranging from 54-58°C for 1 min, an extension step at 70°C for 1 min and a final extension step at 70°C.

The primers for use in amplifying the nucleotide sequences that encodes the proteinaceous molecules of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus may be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof so long as the primers are capable of hybridizing to the polynucleotides of interest. One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4,458,066. The exact length of primer will depend on many factors, including temperature, buffer, and nucleotide composition. The primer must prime the synthesis of extension products in the presence of the inducing agent for amplification.

Primers used according to the method of the invention are complementary to each strand of nucleotide sequence to be amplified. The term "complementary" means that the primers must hybridize with their respective strands under conditions, which allow the agent for polymerization to function. In other words, the primers that are complementary to the flanking sequences hybridize with the flanking sequences and permit amplification of the nucleotide sequence. Preferably, the 3' terminus of the primer that is extended has perfectly base paired complementarity with the complementary flanking strand. Primers and probes for the nucleotide sequence encoding the antigens, polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus can be developed using known methods combined with the present disclosure.

Those of ordinary skill in the art will know of various amplification methodologies that can also be utilized to increase the copy number of target nucleic acid. The polynucleotides that may be used for the present invention can be further evaluated, detected, cloned, sequenced, and the like, either in solution or after binding to a solid support, by any method usually applied to the detection of a specific nucleic acid sequence such as another polymerase chain reaction, oligomer restriction (Saiki *et al.*, 1985, *Bio/Technology* Vol. 3, pages 1008-1012), allele-specific oligonucleotide (ASO) probe analysis (Conner *et al.*, 1983, *Proc. Natl. Acad. Sci. USA* Vol. 80, page 278), oligonucleotide ligation assays (OLAs) (Landegren *et al.*, 1988, *Science* Vol. 241, page 1077), RNase Protection Assay (RPA) and the like. Molecular techniques for DNA analysis have been reviewed (Landegren *et al*, 1988, *Science* Vol. 242, pages 229-237). Following DNA amplification, the reaction product may be detected by

Southern blot analysis, without using radioactive probes. In such a process, for example, a small sample of DNA containing the polynucleotides obtained from the tissue or subject are amplified, and analyzed via a Southern blotting technique. The use of non-radioactive probes or labels is facilitated by the high level of the amplified signal.

The size of the primers used to amplify nucleotide sequences that encode antigens, polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus is at least 10, 15, 20, 25, 30 nucleotide in length. In particular, primers that amplify the M-gene or S-gene is most preferred. Preferably, the G:C ratio should be above 30%, 35%, 40%, 45%, 50%, 55%, 60% so as to prevent hair-pin structure on the primer. Furthermore, the amplicon should be sufficiently long enough to be detected by standard molecular biology methodologies. Preferably, the amplicon is at least 40, 60, 100, 200, 300, 400, 500, 600, 800, 1000, or more base pair in length.

The polynucleotides that may be used in the present invention include polynucleotides having the DNA sequences presented herein, and additionally include any nucleotide sequences encoding an epitope comprising contiguous and functional antigens, polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus encoding open reading frame (ORF) that hybridizes to a complement of the DNA sequences presented herein under highly stringent conditions. By way of example and not limitation, high stringency hybridization conditions can be defined as follows: The filter-bound DNA were hybridized in a solution containing 50% deionized formamide, 6X SSC, 5x Denhardt's, 1% SDS, 100 μg/ml denatured salmon sperm DNA at 42°C overnight (about 4-16 hours), and washing in 0.1X SSC, 0.1% SDS at 65°C (Ausubel F.M. *et al.*, eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York) and encodes a functionally equivalent gene product. For oligonucleotide probes, by way of example and not limitation, highly stringent conditions may refer, *e.g.*, to washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos).

Additionally contemplated polynucleotides that may be used in the present invention include any nucleotide sequences that hybridize under moderately stringent conditions to the complement of the DNA sequences that encode antigens, polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus. By way of example but not limitation, such moderately stringent conditions may include, *e.g.*, washing in 0.2X SSC/0.1% SDS at 42°C (Ausubel *et al.*, 1989, *supra*).

Additionally contemplated polynucleotides that may be used in the present invention include any nucleotide sequences that hybridize under low stringency conditions to the complement of the DNA sequences that encode antigens, polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus. By way of example and not limitation, procedures using such conditions of low stringency are described in Shilo and Weinberg, 1981, *Proc. Natl. Acad. Sci. USA* Vol. 78, pages 6789-6792. A variant may comprise one or more changes in the amino acid sequence of the protein, *e.g.*, by way of addition, substitution, or deletion of one or more amino acids, compared with the wild type protein. Any change should not abolish the ability of the protein to perform its function, though it may increase or decrease this ability depending on the nature of the changes. Preferably, the amino acid changes are conservative.

In various embodiments, antigen, polypeptide of the SARS-CoV virus or fragment, variant, analog, or derivative may be expressed as a fusion, or chimeric protein product (comprising the enzyme, fragment, analog, or derivative joined via a peptide bond to a heterologous protein sequence (of a different protein)). Such a chimeric gene product can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other by methods known in the art, in the proper coding frame, and expressing the chimeric product by methods commonly known in the art. Alternatively, such a chimeric product may be made by protein synthetic techniques, *e.g.*, by use of a peptide synthesizer. Preferably, the fragment, analog, and derivative of the enzyme in the fusion protein retains the ability to perform the enzyme's function.

For the construction of plant nuclear transformation vectors, the expression of the SARS-CoV nucleotide sequences, genes, fragments, derivatives, analogs or variants thereof may be driven by any of a number of regulatory elements. For example, viral promoters such as the 35S RNA and 19S RNA promoters of CaMV (Brisson *et al.*, 1984, *Nature* Vol. 310, pages

511-514), or the coat protein promoter of TMV (Takamatsu *et al.*, 1987, *EMBO J.* Vol. 6, pages 307-311) may be used; alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi *et al.*, 1984, *EMBO J.* Vol. 3, pages 1671-1680; Broglie *et al.*, 1984, *Science* Vol. 224, pages 838-843); or heat shock promoters, *e.g.*, soybean hsp17.5-E or hsp17.3-B (Gurley *et al.*, 1986, *Mol. Cell. Biol.* Vol. 6, pages 559-565) may be used. These constructs can be introduced into plant cells using Ti plasmids, Ri plasmids, plant virus vectors, direct DNA transformation, biolistics/particle bombardment, microinjection, electroporation, *etc.* For reviews of such techniques see, for example, Weissbach & Weissbach, 1988, Methods for Plant Molecular Biology, Academic Press, New York, Section VIII, pp. 421-463; and Grierson & Corey, 1988, Plant Molecular Biology, 2d Ed., Blackie, London, Ch. 7-9. As used herein, regulatory elements include but are not limited to inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Preferably the promoter is capable of directing expression in a particular tissue of the plant and/or at particular stages of development of the plant. The promoter may be heterologous or homologous to the plant. Preferably the promoter directs expression to the fruit, *e.g.*, tomato, the leaves, *e.g.*, lettuce, the endosperm of the plant seed or to the roots or tuber of the plant. A preferred promoter is the high molecular weight glutenin (HMWG) gene of wheat. Other suitable promoters will be known to the skilled man, such as the promoters of gliadin, branching enzyme, ADPG pyrophosphorylase, starch synthase and actin, for example.

A transformed plant with the ability to express SARS-CoV antigens including polypeptides of the SARS-CoV virus, or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus may be engineered by transforming a plant cell with a vector comprising a sequence encoding antigens, including polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus. In one embodiment, a plant promoter is operably associated with a sequence encoding the desired antigens. As used herein, the term "Operably associated" or "operably linked" refers to an association in which the regulatory regions (*e.g.*, promoter, enhancer) and the nucleic acid sequence to be expressed are covalently joined and positioned in such a way as to permit transcription, and under the appropriate condition, translation. In a preferred embodiment of the present invention, the associated promoter for nuclear transformation is a strong and non tissue- or developmental-specific plant promoter (*e.g.*, a promoter that strongly expresses in many or all plant tissue types). Examples of such strong, "constitutive" promoters include, but are not limited to, the CaMV 35S promoter (Odell *et al.*, 1985, *Nature* 313:810-812), the T-DNA mannopine synthetase promoter, and their various derivatives. In another preferred embodiment, an inducible or repressible promoter is used to express the SARS-CoV virus of interest in a plant, for example, a *tet* operator promoter as described in Weinmann *et al.*, 1994, *The Plant Journal* 5: 559-569; or a glucocorticoid-inducible promoter as described in McNellis *et al.*, 1998, *The Plant Journal* 14: 247-257; or an ethanol inducible promoter as described in Caddick *et al.*, 1998, *Nature Biotechnology* 16: 177-180. *See, also*, Gatz, 1995, *Methods In Cell Biology* 50: 411-424, which describes inducible and repressible gene expression systems for plants.

The promoters used for plastid transformation include strong and constitutive promoters in plastid expression including the *psbA* promoter (the *psbA* gene encodes the photosystem II 32kD protein) and the 16S rRNA operon (*rrn*) promoter, or modifications thereof of these promoters which have enhanced expression (Suzuki *et al.*, 2003, *Plant Cell* 15: 195-205; *see also* PCT publication no. WO 00/03012).

In one embodiment of the invention, SARS-CoV antigens will be localized in the apoplastic space from nuclear expression. The SARS-CoV antigens may be directed to the apoplastic space, when expressed in a plant, by expressing the antigens as fusion proteins together with a peptide that acts as a signal or transporter so that the antigen is localized in the apoplastic space of the transgenic plant. A variety of signal or transporter peptides can be used, for example, the PR1b signal sequence as described in Lund *et al.*, 1992, *Plant Molecular Biology* 18: 47-53; or the PR-1a, b and c signal sequences as described in Pfitzner *et al.*, 1987, *Nucleic Acids Research* 15: 4449-4465. A fusion protein comprising a signal or transporter peptide and an SARS-CoV antigen may be constructed by linking polynucleotides specific for each component to each other (*e.g.*, the polynucleotides are linked in frame) so that the desired fusion protein is made when the fusion polynucleotide is expressed in a transgenic plant. A skilled artisan would know how to construct a polynucleotide useful for expressing SARS-CoV antigens, including polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus in the apoplastic space of a transgenic plant.

In another embodiment of the present invention, it may be advantageous to engineer a plant with a vector comprising a sequence SARS-CoV antigens, polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus operably associated with a tissue- or developmental-specific promoter, such as, but not limited to, the tomato E8 fruit-specific promoter, the chalcone synthase (CHS) promoter, the patatin promoter.

In yet another embodiment of the present invention, it may be advantageous to transform a plant with a vector comprising a sequence encoding SARS-CoV antigens, polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus operably linked to a modified or artificial promoter. Typically, such promoters, constructed by recombining structural elements of different promoters, have unique expression patterns and/or levels not found in natural promoters. *See, e.g.,* Salina *et al.*, 1992, *Plant Cell* 4: 1485-1493, for examples of artificial promoters constructed from combining *cis*-regulatory elements with a promoter core.

In yet an additional embodiment of the present invention, the expression of SARS-CoV antigens, polypeptides of the SARS-CoV or fragments, derivatives, analogs, or variants thereof, or molecules having the similar activities as the polypeptides of the SARS-CoV virus may be engineered by increasing the copy number of the gene encoding the desired protein or polypeptide using techniques known in the art.

The present invention provides a vector capable of directing the expression of SARS-CoV viral nucleotide sequences, fragments, derivatives, analogs or variants thereof, in a genetically modified plant or progeny thereof including, for example, transgenic and transplastomic plants. The plant vector is constructed using general recombinant DNA and cloning techniques known in the art of biotechnology, *see, e.g.*, Sambrook *et al., supra*; Ausubel *et al., supra*. Such a polynucleotide construct typically comprises a polynucleotide sequence that encodes an engineered gene product and one or more regulatory polynucleotide sequence. Regulatory sequences useful for the polynucleotide construct of the invention include, but are not limited to, a promoter, an enhancer, an intron, a splice donor, a splice acceptor, a polyadenylation sequence, a RNA stability regulating sequence, or an element of any one of the above (*e.g.*, promoter elements including, but not limited to, a TATA box).

The regulatory elements useful for the present invention are capable of directing expression in a plant species in which expression of the nucleotide sequences that encode the SARS-CoV viral antigens is desired. In another preferred aspect, the regulatory elements are capable of directing expression in a cell type in which expression of the engineered gene product is desired in the plant species of interest.

Regulatory elements useful for the present invention are known to those of skill in the art, for example, promoter and enhancer elements of genes known to be expressed in the cell type and plant species of interest. A promoter useful for expression of the SARS-CoV viral antigens in a plant species of interest may also be isolated using routine experimentation, for example, by isolating a promoter region of a gene known to be expressed in the desired fashion. For example, one may screen a genomic library with a cDNA probe specific for the 5' end of a messenger RNA known to be expressed in the cell type of interest of the plant species of interest. Such a 5' end cDNA probe should preferably be only about 100 base pairs to about 300 base pairs so that the clones identified in the genomic library are likely to include the 5' end of the gene possibly including the promoter region of the gene for which the probe is specific. The promoter region typically includes about 1,000 to about 2,000 base pairs upstream of the transcription initiation site. Thus, a promoter useful for the expression of the engineered SARS-CoV nucleotide sequences, genes, fragments, derivatives, analogs or variants thereof of the present invention is a polynucleotide from about 2,000 base pairs upstream to about 50 base pairs downstream of the transcription initiation site of a gene known to be expressed in the cell type of interest in the plant species of interest, or is a portion of the polynucleotide.

In order to facilitate the proper processing of the SARS-CoV viral antigens, it may be necessary to include a nucleotide sequence that encodes a peptide sequence necessary for such processing. For example, a peptide sequence which is recognized by and functional in the transgenic host plant, for example, to facilitate the entry of the antigen into the endoplasmic reticulum may be necessary, *i.e.*, signal sequence.

<u>TRANSFORMATION OF PLANTS AND PLANT CELLS</u>

Plants and plant cells nuclei and plastids may be transformed using any method known in the art. In an embodiment of the present invention, *Agrobacterium* is employed to introduce the vector of the present invention in nuclear transformation. Such transformation preferably uses binary *Agrobacterium* T-DNA vectors (Bevan, 1984, *Nuc. Acid Res.* Vol. 12, pages 8711-8721) and the co-cultivation procedure (Horsch *et al.*, 1985, *Science* Vol. 227, pages 1229-1231). Generally, the *Agrobacterium* transformation system is used to engineer dicotyledonous plants (Bevan *et al.*, 1982, *Ann. Rev. Genet* Vol. 16, pages 357-384; Rogers *et al.*, 1986, *Methods Enzymol.* Vol. 118, pages 627-641). The *Agrobacterium* transformation system may also be used to transform, as well as transfer, DNA to monocotyledonous plants and plant cells (*see* Hernalsteen *et al.*, 1984, *EMBO J.* Vol. 3: pages 3039-3041 ; Hooykaas-Van Slogteren *et al.*, 1984, *Nature* Vol. 311, pages 763-764; Grimsley *et al.*, 1987, *Nature* Vol. 325, pages 1677-179; Boulton *et al.*, 1989, *Plant Mol. Biol.* Vol. 12, pages 31-40.; and Gould *et al.*, 1991, *Plant Physiol.* Vol. 95, pages 426-434).

In other embodiments, various alternative methods for introducing recombinant nucleic acid constructs into plants and plant cells may also be utilized. These other methods are particularly useful where the target is a monocotyledonous plant or plant cell. Alternative gene transfer and transformation methods include, but are not limited to, particle gun bombardment (biolistics), protoplast transformation through calcium-, polyethylene glycol (PEG)- or electroporation-mediated uptake of naked DNA (*see* Paszkowski *et al.*, 1984, *EMBO J.* Vol. 3, pages 2717-2722; Potrykus *et al.* 1985, *Molec. Gen. Genet.* Vol. 199, pages 169-177; Fromm *et al.*, 1985, *Proc. Nat. Acad. Sci. USA* Vol. 82, pages 5824-5828; and Shimamoto, 1989, *Nature* Vol. 338, pages 274-276) and electroporation of plant tissues (D'Halluin *et al.*, 1992, *Plant Cell* Vol. 4, pages 1495-1505). Additional methods for plant cell transformation include microinjection, silicon carbide mediated DNA uptake (Kaeppler *et al.*, 1990, *Plant Cell Reporter* Vol. 9, pages 415-418), and microprojectile bombardment (*see* Klein *et al.*, 1988, *Proc. Nat. Acad. Sci. USA* Vol. 85, pages 4305-4309; and Gordon-Kamm *et al.*, 1990, *Plant Cell* Vol. 2, pages 603-618). In any methods, selectable markers may be used, at least initially, in order to determine whether transformation has actually occurred. Useful selectable markers include enzymes which confer resistance to an antibiotic, such as gentamycin, hygromycin, kanamycin and the like. Alternatively, markers which provide a compound identifiable by a color change, such as GUS, or luminescence, such as luciferase, may be used.

The chimeric gene may also comprise a gene switch mechanism which determines under what conditions or when the coding sequence is to be expressed. The gene switch may be a chemically induced promoter or a temperature controlled promoter, for example.

In a specific embodiment, each plastid transformation construct (pCV1, pCV6 or pCV8, and its derivatives containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants) was introduced by particle gun bombardment into tobacco (Staub and Maliga, 1994, *Plant Journal* Vol. 6, pages 547-553) and tomato (Ruf *et al.*, 2001, *Nature Biotech* Vol. 19, pages 870-875). In a specific embodiment, production of transgenic tobacco and transgenic lettuce expressing S1 or M antigen by nuclear transformation was carried out using *Agrobacterium*-mediated transformation (Horsch *et al.*, 1985, *Science* Vol. 227, pages 1227-1231) with plasmids pCV2 [and its derivative containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants] and pCV4, respectively. In another specific embodiment, production of transgenic tomato expressing S1 or M antigen by nuclear transformation is carried out.

According to the present invention, a wide variety of plants and plant cell systems may be engineered for the desired physiological and agronomic characteristics described herein using the nucleic acid constructs of the present invention and the various transformation methods mentioned above. In preferred embodiments, target plants and plant cells for engineering include, but are not limited to, those monocotyledonous and dicotyledonous plants, such as crops including grain crops (*e.g.*, wheat, maize, rice, millet, barley), tobacco, fruit crops (*e.g.*, tomato, apple, pear, strawberry, orange), forage crops (*e.g.*, alfalfa), root vegetable crops (*e.g.*, carrot, potato, sugar beets, yam), leafy vegetable crops (*e.g.*, lettuce, spinach); flowering plants (*e.g.*, petunia, rose, chrysanthemum), conifers and pine trees (*e.g.*, pine fir, spruce); plants used in phytoremediation (*e.g.*, heavy metal accumulating plants); oil crops (*e.g.*, sunflower, rape seed); and plants used for experimental purposes (*e.g.*, *Arabidopsis*).

SCREENING OF TRANSFORMED PLANTS AND PLANT CELLS

According to the present invention, desired plants may be obtained by engineering one or more of the vectors expressing SARS-CoV antigens as described herein into a variety of plant cell types, including but not limited to, protoplasts, tissue culture cells, tissue and organ explants, pollens, embryos, as well as whole plants. In an embodiment of the present invention, the engineered plant material is selected or screened for transformants (those that have incorporated or integrated the introduced gene construct(s)) following the approaches and methods described below. An isolated transformant may then be regenerated into a plant and progeny thereof via sexual or asexual reproduction or growth. Alternatively, the engineered plant material may be regenerated into a plant before subjecting the derived plant to selection or screening for the marker gene traits. Procedures for regenerating plants from plant cells, tissues or organs, either before or after selecting or screening for marker gene(s), are well known to those skilled in the art.

A transformed plant cell, callus, tissue or plant may be identified and isolated by selecting or screening the engineered plant material for traits encoded by the marker genes present on the transforming DNA. For instance, selection may be performed by growing the engineered plant material on media containing inhibitory amount of the antibiotic or herbicide to which the transforming gene construct confers resistance. Further, transformed plants and plant cells may also be identified by screening for the activities of any visible marker genes (*e.g.*, the β-glucuronidase, luciferase, B or C1 genes) that may be present on the vector of the present invention. Such selection and screening methodologies are well known to those skilled in the art.

Physical and biochemical methods may also be used to identify plant or plant cell transformants containing the gene constructs of the present invention. These methods include but are not limited to: 1) Southern analysis or PCR amplification for detecting and determining the structure of the recombinant DNA insert; 2) Northern blot, S1 RNase protection, primer-extension or reverse transcriptase-PCR amplification for detecting and examining RNA transcripts of the gene constructs; 3) enzymatic assays for detecting enzyme activity, where such gene products are encoded by the gene construct; 4) protein gel electrophoresis (PAGE), Western blot techniques, immunoprecipitation, or enzyme-linked immunoassays, where the gene construct products are proteins. Additional techniques, such as *in situ* hybridization, enzyme staining, and immunostaining, also may be used to detect the presence or expression of the recombinant construct in specific plant organs and tissues. The methods for doing all these assays are well known to those skilled in the art.

In a specific embodiment, the selectable marker gene *aadA*, which specifies spectinomycin-resistance, is driven from the rice plastid 16S rRNA operon (*rrn*) promoter in plastid transformation. In a specific embodiment, the selectable marker gene *nptII*, which specifies kanamycin-resistance, is driven by the NOS (nopaline synthase) promoter in nuclear transformation.

The present invention relates to transgenic and transplastomic plants that express one or more epitopes of the SARS-CoV virus. A transgenic or transplastomic plant expressing a SARS-CoV antigen elicits the production of antibodies against SARS-CoV by the subject after ingestion of the modified plant or plant parts by a subject; or administration of plant extract to a subject using methods known in the art.

Examples of plants are monocots, dicots, crop plants (*i.e.*, any plant species grown for purposes of agriculture, food production for animals including humans, plants that are typically grown in groups of more than about 10 plants in order to harvest the entire plant or a part of the plant, *e.g.*, a fruit, a flower or a crop, *e.g.*, tobacco, grain, that the plants bear, etc.), trees (*i.e.*, fruit trees, trees grown for wood production, trees grown for decoration, etc.), flowers of any kind (*i.e.*, plants grown for purposes of decoration, for example, following their harvest), cactuses

.

Further examples of plants in which the SARS-CoV viral antigens may be expressed include Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Euphyllophytes, Spermatophyta, Magnoliophyta, Liliopsida, Commelinidae, Poales, Poaceae, *Oryza*, *Oryza sativa*, *Zea*, *Zea mays*, *Hordeum*, *Hordeum vulgare*, Triticum, *Triticum aestivum*, Eudicotyledons, Core eudicots, Asteridae, Euasterids, Rosidae, Eurosids II, Brassicales, Brassicaceae, *Arabidopsis*, Magnoliopsida, Solananae, Solanales, Solanaceae, *Solanum*, and *Nicotiana*.

Also included are, for example, crops of particular interest including Solanaceae, including processing and fresh market tomatoes, pepper and eggplant; leafy plants, including lettuce and spinach; *Brassicas*, including broccoli, brussels sprouts, calabrese, cale, cauliflower, red cabbage and white cabbage; cucurbits, including cucumber, melon, watermelon, zucchini and squash; large seeded plants, including peas, beans and sweetcorn; rooted plants, including carrots and onions; vegetatively propagated plants, including berries, grapes, banana, pineapple, kiwifruit, and rosaceous fruit and nut crops; and tropical crops, including mango and papaya.

Thus, the invention has use over a broad range of plants including, but not limited to, species from the genera *Anacardium, Arachis, Asparagus, Atropa, Avena, Brassica, Citrus, Citrullus, Capsicum, Carthamus, Cocos, Coffea, Cucumis, Cucurbita, Daucus, Elaeis, Fragaria, Glycine, Gossypium, Helianthus, Heterocallis, Hordeum, Hyoscyamus, Lactuca, Linum,* *Lolium, Lupinus, Lycopersicon, Malus, Manihot, Majorana, Medicago, Nicotiana, Olea, Oryza, Panieum, Panneserum, Persea, Phaseolus, Pistachia, Pisum, Pyrus, Prunus, Raphanus, Ricinus, Secale, Senecio, Sinapis, Solanum, Sorghum, Theobromus, Trigonella, Titicum, Vicia, Vitis, Vigna,* and *Zea.*

VACCINE FORMULATIONS AND ADMINISTRATION

The invention also provides vaccine formulations containing extracts of the modified plants of the present invention, which are suitable for administration to elicit a protective immune (humoral and/or cell mediated) response against certain antigens, *e.g.*, for the treatment and prevention of SARS.

Suitable preparations of such vaccines include injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection, may also be prepared. The preparation may also be emulsified, or the proteinaceous molecules encapsulated in liposomes. The active immunogenic ingredients are often mixed with carriers, vehicles or excipients which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, buffered saline, dextrose, glycerol, ethanol, sterile isotonic aqueous buffer or the like and combinations thereof. In addition, if desired, the vaccine preparation may also include minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, and/or adjuvants which enhance the effectiveness of the vaccine.

Examples of adjuvants which may be effective, include, but are not limited to: aluminum hydroxide, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine, N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine.

The effectiveness of an adjuvant may be determined by measuring the induction of anti-idiotype antibodies directed against the injected modified plant extract formulated with the particular adjuvant.

The composition can be a liquid solution, suspension, emulsion, tablet, pill, capsule, sustained release formulation, or powder. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate.

In another embodiment, the composition can be delivered in a vesicle, in particular a liposome (see Langer, 1990, *Science* 249:1527-1533; Treat *et al.*, in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, *ibid.*, pp. 317-327; *see* generally *ibid.*).

Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water-free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is administered by injection, an ampoule of sterile diluent can be provided so that the ingredients may be mixed prior to administration.

In a specific embodiment, the lyophilized modified immunoglobulin of the invention is provided in a first container; a second container comprises diluent consisting of an aqueous solution of 50% glycerin, 0.25% phenol, and an antiseptic (*e.g.*, 0.005% brilliant green).

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the vaccine formulations of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack or transdermal patch. The pack or dispenser device may be accompanied by instructions for administration. Composition comprising an extract of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

The subject to which the vaccine is administered is preferably a mammal, most preferably a human, but can also be a non-human animal including but not limited to rabbits, goats, cows, pigs, sheep, horses, civets, rodents, raccoons, raccoon dogs, dogs, ferrets, ferret Badger, cats, and avian species.

Many methods may be used to introduce the vaccine formulations of the invention; these include but are not limited to oral, intracerebral, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal routes, and via scarification (scratching through the top layers of skin, *e.g.*, using a bifurcated needle) or any other standard routes of immunization. In a specific embodiment, scarification is employed.

The precise dose of the extract of the modified plant to be employed in the formulation will also depend on the route of administration, and the nature of the patient, and should be decided according to the judgment of the practitioner and each patient's circumstances according to standard clinical techniques. An effective immunizing amount is that amount sufficient to produce an immune response to the SARS-CoV viral epitope in the host (*i.e.*, an anti-idiotype reaction) to which the vaccine preparation is administered. Effective doses may also be extrapolated from dose-response curves derived from animal model test systems.

The present invention provides methods of vaccination against SARS. The methods comprise administering whole modified plants or plant parts, extracts, to a subject for the production of antibodies against SARS-CoV viral antigens. In preferred embodiments, the vaccine is administered by ingestion of the modified plants or injection of the modified plant extracts. In certain embodiments, a first vaccine comprising a modified plant extract transformed by a vector for expression of a S antigen may be administered prior to, simultaneously with, or after administration of a second vaccine comprising a modified plant extract transformed by a vector for expression of a M antigen. The term "concurrently" is not limited to the administration of prophylactic or therapeutic composition at exactly the same time, but rather it is meant that the composition of the present invention and the other agent are administered to a mammal in a sequence and within a time interval such that the composition comprising the polynucleotides can act together with the other composition to provide an increased benefit than if they were administered otherwise. In various embodiments, the prophylactic or therapeutic compositions are administered less than 1 hour apart, at about 1 hour apart, at about 1 hour to about 2 hours apart, at about 2 hours to about 3 hours apart, at about 3 hours to about 4 hours apart, at about 4 hours to about 5 hours apart, at about 5 hours to about 6 hours apart, at about 6 hours to about 7 hours apart, at about 7 hours to about 8 hours apart, at about 8 hours to about 9 hours apart, at about 9 hours to about 10 hours apart, at about 10 hours to about 11 hours apart, at about 11 hours to about 12 hours apart, no more than 24 hours apart or no more than 48 hours apart. In preferred embodiments, two or more components are administered within the same patient visit.

In other embodiments, the prophylactic or therapeutic compositions are administered at about 30 minutes, at about 1 hour apart, at about 1 hour to about 2 hours apart, at about 2 hours to about 3 hours apart, at about 3 hours to about 4 hours apart, at about 4 hours to about 5 hours apart, at about 5 hours to about 6 hours apart, at about 6 hours to about 7 hours apart, at about 7 hours to about 8 hours apart, at about 8 hours to about 9 hours apart, at about 9 hours to about 10 hours apart, at about 10 hours to about 11 hours apart, at about 11 hours to about 12 hours apart, at about 1 to 2 days apart, at about 2 to 4 days apart, at about 4 to 6 days apart, at about 1 week part, at about 1 to 2 weeks apart, or more than 2 weeks apart. One skilled in the art would be able to determine such a time frame by determining the half life of the administered compositions.

In certain embodiments, the prophylactic or therapeutic compositions of the invention are cyclically administered to a subject. Cycling therapy involves the administration of a first composition for a period of time, followed by the administration of a second composition and/or third composition for a period of time and repeating this sequential administration. Cycling therapy can reduce the development of resistance to one or more of the therapies, avoid or reduce the side effects of one of the therapies, and/or improves the efficacy of the treatment.

In certain embodiments, prophylactic or therapeutic compositions are administered in a cycle of less than about 3 weeks, about once every two weeks, about once every 10 days or about once every week. One cycle can comprise the administration of a therapeutic or prophylactic composition by infusion over about 90 minutes every cycle, about 1 hour every cycle, about 45 minutes every cycle. Each cycle can comprise at least 1 week of rest, at least 2 weeks of rest, at least 3 weeks of rest. The number of cycles administered is from about 1 to about 12 cycles, more typically from about 2 to about 10 cycles, and more typically from about 2 to about 8 cycles.

In yet other embodiments, the therapeutic and prophylactic compositions of the invention are administered in metronomic dosing regiments, either by continuous infusion or frequent administration without extended rest periods. Such metronomic administration can involve dosing at constant intervals without rest periods. The dosing regimens encompass the chronic daily administration of relatively low doses for extended periods of time. In preferred embodiments, the use of lower doses can minimize toxic side effects and eliminate rest periods. In certain embodiments, the therapeutic and prophylactic compositions are delivered by chronic low-dose or continuous infusion ranging from about 24 hours to about 2 days, to about 1 week, to about 2 weeks, to about 3 weeks to about 1 month to about 2 months, to about 3 months, to about 4 months, to about 5 months, to about 6 months. The scheduling of such dose regimens can be optimalized by the skilled physician.

The dosage amounts and frequencies of administration provided herein are encompassed by the terms therapeutically effective and prophylactically effective. The dosage and frequency further will typically vary according to factors specific for each patient depending on the specific therapeutic or prophylactic composition administered, the severity and type of disease or disorder, the route of administration, as well as age, body weight, response, and the past medical history of the patient. Suitable regimens can be selected by one skilled in the art by considering such factors and by following, for example, dosages reported in the literature and recommended in the *Physician's Desk Reference* (56^{th} ed., 2002).

Various delivery systems are known and can be used to administer the therapeutic or prophylactic composition of the present invention, *e.g.*, encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the antibody or antibody fragment, receptor-mediated endocytosis (see, *e.g.*, Wu and Wu, *J. Biol. Chem.* 262:4429-4432 (1987)), construction of a nucleic acid as part of a retroviral or other vector, etc. Methods of administering a prophylactic or therapeutic composition of the invention include, but are not limited to, parenteral administration (*e.g.*, intradermal, intramuscular, intraperitoneal, intravenous and subcutaneous), epidural, and mucosal (*e.g.*, intranasal and oral routes). In a specific embodiment, prophylactic or therapeutic composition of the invention is administered intramuscularly, intravenously, or subcutaneously. The prophylactic or therapeutic composition may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (*e.g.*, oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local.

In a specific embodiment, it may be desirable to administer the prophylactic or therapeutic composition of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion, by injection, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers.

In yet another embodiment, the prophylactic or therapeutic composition can be delivered in a controlled release or sustained release system. In one embodiment, a pump may be used to achieve controlled or sustained release (see Langer, *supra*; Sefton, 1987, *CRC Crit. Ref. Biomed. Eng.* 14:20; Buchwald *et al.*, 1980, *Surgery* 88:507; Saudek *et al.*, 1989, *N. Engl. J. Med.* 321:574). In another embodiment, polymeric materials can be used to achieve controlled or sustained release of the therapeutic or prophylactic composition of the invention (see *e.g.*, Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974); Controlled Drug Bioavailability, Drug Product Design and Performance, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, 1983, *J., Macromol. Sci. Rev. Macromol. Chem.* 23:61; see also Levy *et al.*, 1985, *Science* 228:190; During *et al.*, 1989, *Ann. Neurol.* 25:351; Howard *et al.*, 1989, *J. Neurosurg.* 7 1:105); U.S. Patent No. 5,679,377; U.S. Patent No. 5,916,597; U.S. Patent No. 5,912,015; U.S. Patent No. 5,989,463; U.S. Patent No. 5,128,326; PCT Publication No. WO 99/15154; and PCT Publication No. WO 99/20253. Examples of polymers used in sustained release formulations include, but are not limited to, poly(2-hydroxy ethyl methacrylate), poly(methyl methacrylate), poly(acrylic acid), poly(ethylene-co-vinyl acetate), poly(methacrylic acid), polyglycolides (PLG), polyanhydrides, poly(N-vinyl pyrrolidone), poly(vinyl alcohol), polyacrylamide, poly(ethylene glycol), polylactides (PLA), poly(lactide-co-glycolides) (PLGA), and polyorthoesters. In a preferred embodiment, the polymer used in a sustained release formulation is inert, free of leachable impurities, stable on storage, sterile, and biodegradable. In yet another embodiment, a controlled or sustained release system can be placed in proximity of the prophylactic or therapeutic target, thus requiring only a fraction of the systemic dose (see, *e.g.*, Goodson, in Medical Applications of Controlled Release, supra, vol. 2, pp. 115-138 (1984)).

Controlled release systems are discussed in the review by Langer (1990, Science 249:1527-1533). Any technique known to one skilled in the art can be used to produce sustained release formulations comprising one or more therapeutic composition of the invention. See, *e.g.*, U.S. Patent No. 4,526,938, PCT publication WO 91/05548, PCT publication WO 96/20698,.Ning *et al.*, 1996, "Intratumoral Radioimmunotherapy of a Human Colon Cancer Xenograft Using a Sustained-Release Gel," Radiotherapy & Oncology 39:179-189, Song *et al.*, 1995, "Antibody Mediated Lung Targeting of Long-Circulating Emulsions," PDA Journal of Pharmaceutical Science & Technology 50:372-397, Cleek *et al.*, 1997, "Biodegradable Polymeric Carriers for a bFGF Antibody for Cardiovascular Application," Pro. Int'l. Symp. Control. Rel. Bioact. Mater. 24:853-854, and Lam *et al.*, 1997, "Microencapsulation of Recombinant Humanized Monoclonal Antibody for Local Delivery," Proc. Int'l. Symp. Control Rel. Bioact. Mater. 24:759-760, each of which is incorporated herein by reference in their entireties.

Antibodies may be isolated from a subject after administration of the vaccine of the present invention. Methods of production of antibodies against SARS comprise administering an effective amount of modified plant to a subject and isolating antibodies from a sample of the subject after a period of time. The sample may be a biological fluid, such as blood, serum, plasma, saliva, urine, stool, sputum, nasopharyngeal aspirates, cells and tissues.

DEMONSTRATION OF PROPHYLACTIC AND THERAPEUTIC UTILITY

The present invention provides methods for preventing, treating, ameliorating, and/or managing SARS by administration of the modified plants of the present invention to a subject in need thereof. Modified plants that are useful for the methods of the invention can be screened or assayed in a variety of ways for efficacy in treating or preventing SARS.

First, the immunopotency of a vaccine formulation containing the modified plant or extracts of the plant of the invention can be determined by monitoring the immune response of test animals following immunization with the vaccine. Generation of a humoral response may be taken as an indication of a generalized immune response, other components of which, particularly cell-mediated immunity, may be important for protection against a disease. Test animals may include mice, rabbits, chimpanzees and eventually human subjects. A vaccine made in this invention can be made to infect chimpanzees experimentally. However, since chimpanzees are a protected species, the antibody response to a vaccine of the invention can first be studied in a number of smaller, less expensive animals, with the goal of finding one or two best candidate modified plants that produce the best combinations of SARS-CoV viral epitopes to use be used in chimpanzee efficacy studies.

The safety of the vaccine can also be determined by observing or monitoring symptoms of subjects after administration of vaccine to a test subject. In a specific embodiment, the entire genome of SARS-CoV is used in the vaccine. In preferred embodiment, the plant express less than 10%, less than 20%, less than 25%, less than 30%, less than 35%, less than 40%, less than 45%, less than 50%, less than 55%, less than 60%, less than 65%, less than 70%, less than 80%, less than 90% of the SARS-CoV polypeptides.

The immune response of the test subjects can be analyzed by various approaches such as the reactivity of the resultant immune serum to antigens, as assayed by known techniques, *e.g.*, enzyme linked immunosorbent assay (ELISA), immunoblots, radioimmunoprecipitations, etc.; or protection from infection and/or attenuation of disease symptoms in immunized hosts.

As one example of suitable animal testing, the vaccine composition of the invention may be tested in rabbits for the ability to induce an immune response to the SARS viral antigens. For example, male specific-pathogen-free (SPF) young adult New Zealand White rabbits may be used. The test group of rabbits each receives an effective amount of the vaccine, via any route of administration, e.g., ingestion of modified plant or injection of modified plant extract. A control group of rabbits receives an injection in 1 mM Tris-HCl pH 9.0 of the vaccine containing the same kind of plant extract from an unmodified plant of the same type. Blood samples may be drawn from the rabbits every one or two weeks, and serum analyzed for antibodies specific for the viral antigen was directed using, *e.g.*, a radioimmunoassay (Abbott Laboratories). The presence of antibodies may be assayed using an ELISA. Because rabbits may give a variable response due to their outbred nature, it may also be useful to test the vaccines in mice.

In one embodiment, the invention relates to biological material collected from subjects to which the modified plants and/or plant extracts were administered, either by injection or ingestion. The biological material can be tested for the presence of SARS-CoV polypeptides or fragments thereof. Biological materials include, but are not limited to, blood, serum, plasma, saliva, urine, stool, sputum, nasopharyngeal aspirates, cells and tissues. The biological material can be collected 1 hour, 2 hours, 6 hours, 12 hours, 24 hours, 2 days, 3 days, or 1 week after administration of the modified plant and/or plant extracts of the present invention.

In addition, a modified plant of the invention may be tested by first administering the modified plant to a test subject, either animal or human, and then isolating the anti-anti-idiotypic antibodies (*i.e.*, the Ab3 antibodies) generated as part of the anti-idiotype response to the viral antigens. The isolated Ab3 may then be tested for the ability to bind the particular viral antigen (*e.g.*, by any immunoassays known in the art, for example, but not limited to, radioimmunoassays, ELISA, "sandwich" immunoassay, gel diffusion precipitin reactions, immunodiffusion assays, western blots, precipitation reactions, agglutination assays, complement fixation assays, immunofluorescence assays, protein A assays, immunoelectrophoresis assays, etc.)

In one aspect where the vaccine is directed against SARS-CoV viral antigen, the efficacy of the isolated Ab3 for treating SARS is screened by culturing SARS-CoV virus infected cells from a culture or a patient, contacting the cells with the Ab3 antibody to be tested, and comparing the proliferation or survival of the contacted cells with the proliferation or survival of cells not so contacted with the Ab3 antibody, wherein a lower level of proliferation or survival of the contacted cells indicates that the Ab3 antibody (which was elicited by immunization with the SARS-CoV viral antigen) is effective to treat the SARS in the patient. Many assays standard in the art can be used to assess such survival and/or growth; for example, cell proliferation can be assayed by measuring ^{3}H-thymidine incorporation, by direct cell count, by detecting changes in transcriptional activity of known genes such as proto-oncogenes (*e.g.*, *fos*, *myc*) or cell cycle markers; cell viability can be assessed by trypan blue staining, differentiation can be assessed visually based on changes in morphology.

The present invention also provides antibodies against the SARS-CoV virus subsequent to the administration of the modified plant of the invention to a subject in need thereof. An antibody that immunospecifically binds an antigen of the SARS-CoV virus may also be tested directly *in vivo*. To monitor the effect of an antibody of the invention, the level of the antigen is measured at suitable time intervals before, during, or after administration of the vaccine. Any change or absence of change in the amount of the antigen can be identified and correlated with the effect of the treatment on the subject.

In particular, in the case of SARS, the serum levels of a SARS-CoV antigen bears a direct relationship with severity of SARS. Generally, a decrease in the level of antigen is associated with efficacious treatment. The serum levels of SARS-CoV antigens in a treated subject can be measured every 1 hour, 2 hours, 3 hours, 6 hours, 12 hours, 24 hours, 2 days, 3 days, or 1 week after administration of the modified plant and/or plant extracts of the present invention to said subject.

In a preferred aspect, the approach that can be taken is to determine the levels of antigen at different time points and to compare these values with a baseline level. The baseline level can be either the level of the marker present in normal, disease free individuals; and/or the levels present prior to treatment, or during periods of stability. These levels can then be correlated with the disease course or treatment outcome.

The levels of antigen can be determined by any method well known in the art. For example, SARS-CoV viral antigen can be quantitated by known immunodiagnostic methods such as western blotting immunoprecipitation using any antibody against SARS-CoV antigen.

The strength of the immune response *in vivo* to the viral antigen may be determined by any method known in the art, for example, but not limited to, delayed hypersensitivity skin tests and assays of the activity of cytolytic T-lymphocytes *in vitro*.

Delayed hypersensitivity skin tests are of great value in the testing of the overall immunocompetence and cellular immunity to an antigen. Proper technique of skin testing requires that the antigens be stored sterile at 4°C, protected from light and reconstituted shortly before use. A 25- or 27-gauge need ensures intradermal, rather than subcutaneous, administration of antigen. Twenty-four and 48 hours after intradermal administration of the antigen, the largest dimensions of both erythema and induration are measured with a ruler. Hypoactivity to any given antigen or group of antigens is confirmed by testing with higher concentrations of antigen or, in ambiguous circumstances, by a repeat test with an intermediate test.

To test the activity of cytolytic T-lymphocytes, T-lymphocytes isolated from the immunized subject, *e.g.*, by the Ficoll-Hypaque centrifugation gradient technique, are re-stimulated with cells bearing the antigen against SARS was directed in 3 ml RPMI medium containing 10%

fetal calf serum. In some experiments, 33% secondary mixed lymphocyte culture supernatant or IL-2 is included in the culture medium as a source of T cell growth factors. In order to measure the primary response of cytolytic T-lymphocytes after immunization, the isolated T cells are cultured with or without the cells bearing the antigen. After six days, the cultures are tested for cytotoxity in a 4 hour ^{51}Cr-release assay. The spontaneous ^{51}Cr-release of the targets should reach a level less than 20% if immunization was effective (Heike *et al.*, *J. Immunotherapy* Vol. 15, pages 15-174).

The efficacy of the antibody can be assayed by administering the antibody to a subject (either a human subject or an animal model for the disease) and then monitoring either the levels of the SARS-CoV viral antigens or symptoms of the particular infectious disease. The levels of the SARS-CoV viral antigens may be determined by any method known in the art for assaying the levels of SARS-CoV viral antigens, *e.g.*, the viral titer, in the case of a virus, or bacterial levels (for example, by culturing of a sample from the patient), etc. A decrease in the levels of the viral antigens or an elimination, amelioration or reduction in the number of symptoms of SARS indicates that the antibody is effective. Symptoms of SARS include, but are not limited to, temperature of greater than 100.4°F (>38°C) (according to the CDC), chills, headache, malaise, body aches, cough, shortness of breath, difficulty breathing, and hypoxia. Radiographic evidence of pneumonia, respiratory distress syndrome, and autopsy findings are also useful to determine whether a subject is inflicted with, recovering from, or free of SARS.

EFFICACY OF THE PROPHYLACTIC AND THERAPEUTIC UTILITIES

Toxicity and efficacy of the prophylactic and/or therapeutic protocols of the instant invention can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Prophylactic and/or therapeutic agents that exhibit large therapeutic indices are preferred. While prophylactic and/or therapeutic agents that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such agents to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage of the prophylactic and/or therapeutic agents for use in humans. The dosage of such agents lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any agent used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC_{50} (*i.e.*, the concentration of the test compound that achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

EXAMPLES

The following examples illustrate the generation of transgenic or transplastomic tobacco plants that comprise the M and/or S1 proteins of the SARS-CoV virus.

PLASMID CONSTRUCTION

The backbone for construction of these plastid transformation vectors is derived from plasmid pVSR326 (Reddy *et al.*, 2002, *Mol Breeding* 9: 259-269; PCT international application no. PCT/EP00/12446, international publication no. WO 01/42441). Plasmid pVSR326 utilizes the tobacco plastid genome sequences spanning *rbcL-accD* to target the reporter gene encoding β-glucuronidase (*GUS*) into the chloroplast genome by homologous recombination (Reddy *et al.*, 2002). The promoter and terminator for *GUS* are derived from the rice plastid gene *psbA*; the *psbA* gene encoding the photosystem II 32kD protein. The selectable marker gene *aadA*, which specifies spectinomycin-resistance, is driven from the rice plastid 16S rRNA operon (*rrn*) promoter and *aadA* is located adjacent to *GUS* (Reddy *et al.*, 2002). In the pMLVHis vectors, other than unique restriction sites at this region, a start codon plus a $(His)_5$-tag is incorporated to facilitate the cloning and expression of $(His)_5$-tagged recombinant proteins.

Construction of pCV1, a plastid transformation vector for expression of the S1 antigen

The 2-kb *Spe*I-*Not*I fragment encoding S1 minus its start codon and signal peptide from pCRII-S1 was cloned into the *Nhe*I and *Not*I sites of pMLVHisA (Fig. 4). The plasmid pCRII-S1 was derived from pCRII (Invitrogen) and contains a PCR-amplified fragment of S1. In the resultant derivative, designated pCV1 (Fig. 1), amino acids 12-658 of S1 are fused in-frame to seventeen pMLVHisA-derived residues, including the 'ATG' start codon and a $(His)_5$-tag. Plasmid pMLVHisA is designed with a $(His)_5$-tag enabling recognition of the recombinant protein using antisera against the $(His)_5$-tag in western blot analysis. Also, the $(His)_5$-tag enables easy purification of the recombinant protein, if and when required, using Ni-NTA Agarose (Qiagen) affinity columns. The rice plastid *psbA* promoter drives expression of the recombinant protein and the terminator is also *psbA*-derived. The presence of flanking *rbcL* and *accD* sequences from the tobacco plastid genome enables homologous recombination to occur, resulting in the incorporation of *S1* (and *aadA*) into the plastid genome of the target plant.

Since plant chloroplast genome sequences are highly-conserved, the same construct was used for plastid transformation of tobacco and tomato, both belonging to the same family *Solanaceae*.

Construction of pCV2, a nuclear transformation vector for expression of the S1 antigen

The 2-kb *Bam*HI-*Xho*I fragment of *S1* encoding amino acids 1-658 from pCRII-S1 was cloned into the *Bam*HI and *Sal*I sites of pSa7 (Fig. 5), a plant nuclear transformation vector. In this case, the signal peptide of S1 (amino acids 1-13) is retained because previous reports suggest that the presence of endoplasmic reticulum-targeting signals on plant nuclear-expressed foreign proteins results in improved protein stability (Richter *et al.*, 2000, *Nature Biotech* 18: 1167-1171; Sojikul *et al.*, 2003, *Proc. Natl. Acad. Sci.* 100: 2209-2214). In the resultant plasmid, pCV2 (Fig. 19), *S1* is placed between the strong and constitutive Cauliflower Mosaic Virus 35S promoter and the nopaline synthase (*NOS*) terminator. This plasmid was introduced from *Escherichia coli* to *Agrobacterium tumefaciens* strain LBA4404 in the presence of *E. coli* helper strain HB101/pRK2013 by triparental mating (Horsch *et al.*, 1985, *Science* 227: 1227-1231).

Construction of pCV8, a plastid transformation vector for expression of the M antigen

The 0.7-kb *Nco*I-*Sac*I fragment encoding M from plasmid pCRII-M was cloned into the *Nco*I and *Sac*I sites of pMLVHisA (Fig. 4) to generate pCV7. The plasmid pCRII-M was derived from pCRII (Invitrogen) and contains a PCR-amplified fragment of M. Subsequently, plasmid pCV7 was digested with *Nco*I and a $(His)_6$-tag was introduced into this site using annealed oligomers ML527 (5'-CATGGCCGCGCGGGGTTCTCATCATCATCATCATCATGG-3'; SEQ ID NO:10) and ML528 (5'-CATGCCATGATGATGATGATGATGAGAACCCCGCGCGGC-3'; SEQ ID NO:11) to create pCV8 (Fig. 7). The resultant M expressed from pCV8 is a $(His)_6$-tagged protein.

Construction of pCV4, a nuclear transformation vector for expression of the M antigen

The 0.7-kb *Bam*HI fragment encoding M from plasmid pCRII-M was cloned into the *Bam*HI site of plant nuclear transformation vector pSa13. Plasmid pSa13 was derived from plasmid pSa7 (Fig. 5) by removal of a 0.58-kb *Sal*I fragment of *SaPIN2a* cDNA followed by religation.

Construction of pCV6, a plastid transformation vector for co-expression of the S1 and M antigens

A 1.2-kb partial *Bam*HI fragment consisting of the M-gene fused to the *psbA* terminator and the *rrn* promoter, in this order, from plasmid pCV3 was cloned into the *Bam*HI unique site of plasmid pCV1 to generate plastid transformation vector pCV6 (Fig. 21) that co-expresses the S1 and M antigens. Plasmid pCV3 was obtained by cloning a 0.7-kb *Eco*RV-*Sac*I fragment encoding M from pCRII-M in the *Stu*I-*Sac*I sites of pMLVHisA.

Codon usage optimization of SARS-CoV S1 gene by site directed mutagenesis

Plasmid pCRII-S1 was used as a mutagenesis template for generating derivatives of codon usage optimization. Mutagenesis was carried out by PCR with *PfuTurbo* DNA polymerase using the 'QuikChange Multi site-directed mutagenesis kit' (Stratagene), according to the manufacturer's specifications. Table 1 shows the oligonucleotides used for site-directed mutagenesis. The amplification procedure included denaturation at 95 °C for 1 min, followed by 30 cycles of denaturation (95 °C for 1 min), annealing (55 °C for 1 min) and extension (65 °C for 12 min). The amplified product was treated with *Dpn*I to remove template DNA and the mutated DNA was used in transformation of *E. coli* XL10-Gold cells. DNA sequencing analysis was performed to confirm each mutation. A total of 13 nucleotide changes (as summarized in Table 1) incorporated in an "optimized" S1 was then used for cloning in plastid and nuclear transformation vectors.

Table 1. Oligonucleotides used for site-directed mutagenesis of S1 for codon optimization

Affected Residue	Sequence of primer
	Forward primers:
R18	5'-GGTAGTGACCTTGACAGATGCACCACTTTTGAT-3'; SEQ ID NO:12
T75	5'-GGGTTTCATACTATTAATCATACTTTTGGCAACCCTGTCATAC-3'; SEQ ID NO:13
S113	5'-CCATGAACAACAAGTCACAGTCTGTGATTATTATTAACAATTCTACT-3'; SEQ ID NO:14
S169	5'-AGTACATATCTGATGCCTTTTCTCTTGATGTTTCAGAAAAGTC-3'; SEQ ID NO:15
L209	5'-CCTATAGATGTAGTTCGTGATCTTCCTTCTGGTTTTAACACTTTG-3'; SEQ ID NO:16
T247	5'-CAAGACATTTGGGGCACTTCAGCTGCAGCCTAT-3'; SEQ ID NO:17
A398	5'-GATGATGTAAGACAAATAGCTCCAGGACAAACTGG-3'; SEQ ID NO:18
P507	5'-TCTTTTGAACTTTTAAATGCACCTGCCACGGTTTGTGGACC-3'; SEQ ID NO:19
T509	5'-CTTTTAAATGCACCTGCCACTGTTTGTGGACCAAAATTATC-3'; SEQ ID NO:20
L597	5'-CTTCATCTGAAGTTGCTGTTCTTTATCAAGATGTTAACTGCAC-3'; SEQ ID NO:21
R620	5'-CAACTCACACCAGCTTGGAGAATATATTCTACTGGAAACAATG-3'; SEQ ID NO:22
	Reverse primers:
R18	5'-ATCAAAAGTGGTGCATCTGTCAAGGTCACTACC-3'; SEQ ID NO:23
R620	5'-CATTGTTTCCAGTAGAATATATTCTCCAAGCTGGTGTGAGTTG-3'; SEQ ID NO:24

*nucleotides in italics are mutated. The affected codons are underlined.

PLANT TRANSFORMATION

Plastid transformation of tobacco and tomato for expression of the S1 and/or M antigens

Each plastid transformation construct (pCV1, pCV6 or pCV8, and its derivatives containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants) was introduced by particle gun bombardment into tobacco (Staub and Maliga, 1994, *Plant Journal* 6: 547-553) and tomato (Ruf *et al.*, 2001, *Nature Biotech* 19: 870-875).

Plastid transformation of *Lycopersicon esculentum* cv. UC82B

Plastid transformation was carried out following Ruf *et al.,* 2001, by bombardment of young leaves with tungsten particles coated with the desired DNA (pCV1, pCV6 or pCV8, and its derivatives containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants).

Aseptical seeds were germinated on Murashige and Skoog (MS) medium (Murashige and Skoog, 1962, *Physiol. Plant.* 15: 473-497), supplemented with agar (5 g/liter) and sucrose (30 g/liter). For bombardment, young leaves from outgrowing axillary meristem were placed abaxial side facing up on RMOP medium containing MS salts, 6-benzylaminopurine (1 mg/liter), α-naphthalene acetic acid (0.1 mg/liter), thiamine (1 mg/liter), myo-inositol (100 mg/liter), agar (5 g/liter) at pH 5.7 and sucrose (30 g/liter). Bombardment was carried out using the particle delivery system, PDS 1000-He (Bio-Rad) and its accessories. Tungsten particles (M 17) were coated with desired DNA and leaves bombarded. After bombardment, leaves were cut into small pieces and placed on RMOP selection medium containing spectinomycin dihydrochloride (250 mg/liter). Greenish yellow calli were passaged for few more cycles in the selection medium to obtain homoplastomic plastid-containing calli. From this calli plant regeneration will be achieved.

Plastid transformation of *N. tabacum* cv. xanthi

Plastid transformation was carried out following Svab and Maliga (Svab *et al.*, 1993, *Proc. Natl. Acad. Acad. Sci. USA* 90: 913-917) by bombardment of leaves with tungsten particles coated with the desired DNA (pCV1, pCV6 or pCV8, and its derivatives containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants).

Aseptical seeds were germinated on Murashige and Skoog (MS) medium (Murashige and Skoog, 1962, *Physiol. Plant.* 15: 473-497), supplemented with agar (5 g/liter) and sucrose (30 g/liter). For bombardment, leaves were placed abaxial side facing up on RMOP medium containing MS salts, 6-benzylaminopurine (1 mg/liter), α-naphthalene acetic acid (0.1

mg/liter), thiamine (1 mg/liter), myo-inositol (100 mg/liter), agar (5 g/liter) at pH 5.7 and sucrose (30 g/liter). Bombardment was carried out using the particle delivery system, PDS 1000-He (Bio-Rad) and its accessories. Tungsten particles (M 17) were coated with desired DNA and leaves bombarded. After bombardment, leaves were cut into small pieces and placed on RMOP selection medium containing spectinomycin dihydrochloride (500 mg/liter). The regenerated plantlets obtained following plastid transformation of tobacco using vector pCV1 (Fig. 23F-23G) were passaged for five more cycles in selection medium to obtain homoplastomic plastid-containing plants. Results of PCR analysis (Figure 23H) using *S1* primers ML560 (5'-CAGAGAGAGTTTCCCGGATTC-3'; SEQ ID NO:25) and ML567 (5'-CAACCTATAGATGTAGTTCG-3'; SEQ ID NO:26) and of northern blot analysis (Fig. 23I) using an ^{32}P-radiolabeled S1 probe are shown. Figure 23J shows western blot analysis of transplastomic tobacco expressing a modified S1 (with nucleotide changes as stipulated in Table 1) using Ni-NTA conjugate in detection of a His-tagged S1 protein. The arrow indicates the expected 73-kDa $(His)_5$-S1 band.

Nuclear transformation of tobacco and lettuce for expression of S1 or M antigen

Production of transgenic tobacco and transgenic lettuce expressing S1 or M antigen by nuclear transformation was carried out using *Agrobacterium*-mediated transformation (Horsch *et al.*, 1985, *Science* 227: 1227-1231) with plasmids pCV2 [and its derivative containing an S1 with nucleotide (but not amino acid) changes to optimize codon usage for expression in plants] and pCV4, respectively.

Following *Agrobacterium*-mediated transformation, shoots of tobacco (Fig. 23A) and lettuce (Fig. 23C) were rooted and the regenerated shoots (Fig. 23, B, D) were used in PCR analysis using primers 35S and NOS-ter that are located within the CaMV 35S promoter and the *NOS*-terminator, respectively. Southern blot analysis using a ^{32}P-labeled S1 probe of these PCR-amplified DNA is shown in Fig. 23E. A 2.1-kb hybridizing band confirms the presence of S1-containing transgenic lines (Fig. 23E, lanes 2-5).

Nuclear transformation of tobacco for expression of S1:GFP antigen

Production of transgenic tobacco expressing a protein fusion consisting of the SARS-CoV S1 protein fused with green fluorescent protein (GFP) was carried out using *Agrobacterium*-mediated transformation (Horsch *et al. supra.*) with plasmid pCV12.

Two days after agroinfiltration, representative tobacco leaf epidermal cells were selected and observed by confocal microscopy (Yang Y. *et al.*, *In vivo* analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. *Plant J.* 2000; 22: 543-551) of *Agrobacterium tumefaciens* LBA4404 harboring plasmid pCV12 expressing S1:GFP fusion protein (Fig. 18, A, C) or LBA4404 harboring pGDG expressing GFP alone (Fig. 18, B, D). Bar represents 20μm.

Western blot analysis was performed using antibodies against GFP and showed transient expression of S1:GFP in the tobacco leaves following agroinfiltration. Total protein (200 μg) extracted from tobacco leaves, infiltrated with plasmid pGDG expressing GFP alone (Fig. 23, lane 1) or plasmid pCV12 expressing S1:GFP fusion (Fig. 23, lane 2), were separated on a 8% SDS-PAGE gel, blotted onto Hybond-C filters according to Sambrook *et al.* (1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor), and cross-reacted with antibodies against GFP (Clontech). The results are shown in Figure 24, wherein arrow indicates cross-reacting S1:GFP band (calculated size 99.1 kDa) and M is the molecular mass marker.

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

All publications, patents and patent applications mentioned in this specification are herein incorporated by reference into the specification to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference.

Citation or discussion of a reference herein shall not be construed as an admission that such is prior art to the present invention.

WHAT IS CLAIMED:

1. A vector comprising an isolated nucleic acid which nucleic acid sequence is set forth in SEQ ID NO:7, or a fragment, derivative or analog thereof.
2. The vector in accordance with claim 1, wherein the vector is a plastid transformation vector.
3. The vector in accordance with claim 2, wherein the vector is pCV1, pCV6, pCV8, or a derivative thereof containing an S1 with at least one nucleotide change.
4. The vector in accordance with claim 1, wherein the vector is a nuclear transformation vector.
5. The vector in accordance with claim 4, wherein the vector is pCV2, pCV4, pCV12, or a derivative thereof containing an S1 with at least one nucleotide change.
6. The vector in accordance with claim 1, wherein the isolated nucleic acid encodes the S protein of the SARS-CoV virus, the M protein of the SARS-CoV virus, or both.
7. The vector in accordance with claim 1, wherein the isolated nucleic acid comprises the nucleic acid sequence as set forth in SEQ ID NO:3, 5, or a fragment, derivative or analog thereof.
8. The vector in accordance with claim 1, wherein the isolated nucleic acid encodes the amino acid sequence as set forth in SEQ ID NO:4 or 6, or a fragment, derivative or analog thereof.
9. The vector in accordance with claim 1, wherein the isolated nucleic acid hybridizes under high stringency conditions to the nucleic acid sequence of SEQ ID NO:3, 5, or 7, or a complement thereof.
10. A plant cell comprising an isolated nucleic acid which nucleic acid sequence is set forth in SEQ ID NO:7, or a fragment, derivative or analog thereof.
11. The plant cell in accordance with claim 10, wherein the isolated nucleic acid encodes the S protein of the SARS-CoV virus, the M protein of the SARS-CoV virus, or both.
12. The plant cell in accordance with claim 10, wherein the isolated nucleic acid comprises the nucleic acid sequence as set forth in SEQ ID NO:3, 5, or a fragment, derivative or analog thereof.
13. The plant cell in accordance with claim 10, wherein the isolated nucleic acid encodes the amino acid sequence as set forth in SEQ ID NO:4 or 6, or a fragment, derivative or analog thereof.

14. The plant cell in accordance with claim 10, wherein the isolated nucleic acid hybridizes under high stringency conditions to the nucleic acid sequence of SEQ ID NO:3, 5, or 7, or a complement thereof.

15. The plant cell in accordance with claim 10, wherein the plant cell is isolated from a plant selected from the group consisting of tobacco, lettuce, tomato, potato, banana, corn, rice, cereals, wheat, maize, barley, apple, pear, strawberry, carrot, sugar beets, yam, kiwifruit, or spinach.

16. A plant cell comprising the vector in accordance with claim 1.

17. A method of immunization against SARS comprising administering to a subject a plant, wherein the plant comprises an isolated nucleic acid which nucleic acid sequence is set forth in SEQ ID NO:7, or a fragment, derivative, analog, or vector thereof.

18. The method in accordance to claim 17, wherein the isolated nucleic acid encodes the S protein of the SARS-CoV virus, the M protein of the SARS-CoV virus, or both.

19. The method in accordance to claim 17, wherein the isolated nucleic acid comprises the nucleic acid sequence as set forth in SEQ ID NO:3, 5, or a fragment, derivative or analog thereof.

20. The method in accordance to claim 17, wherein the isolated nucleic acid encodes the amino acid sequence as set forth in SEQ ID NO:4 or 6, or a fragment, derivative or analog thereof.

21. The method in accordance to claim 17, wherein the isolated nucleic acid hybridizes under high stringency conditions to the nucleic acid sequence of SEQ ID NO:3, 5, or 7, or a complement thereof.

22. The method in accordance to claim 17, where the administering to the subject is oral administration.

23. The method in accordance to claim 17, wherein subject is human.

24. A method of immunization against SARS comprising administering to a subject a vector in accordance with claim 1 and an acceptable pharmaceutical carrier.

25. The method in accordance to claim 24, wherein the subject is human.

26. A plant comprising the isolated nucleic acid which nucleic acid sequence is set forth in SEQ ID NO:7, or a fragment, derivative or analog thereof.

27. The plant in accordance with claim 26, which is selected from the group consisting of tobacco, lettuce, tomato, potato, banana, corn, rice, cereals, wheat, maize, barley, apple, pear, strawberry, carrot, sugar beets, yam, kiwifruit, or spinach.

28. A composition comprising a SARS-CoV viral antigen produced by the plant cell of claim 10.

29. A composition comprising a SARS-CoV viral antigen produced by the plant cell of claim16.

30. A composition comprising a SARS-CoV viral antigen produced by the plant of claim 26.

31. A method of detecting an antibody to a SARS-CoV viral antigen in a sample comprising:
   (a) contacting the sample with the composition of claim 28; and
   (b) detecting the presence of an antibody bound to the composition of claim 28, thereby detecting an antibody to a SARS-CoV viral antigen.

32. The method in accordance with claim 31, wherein the sample is a biological fluid.

33. The method in accordance with claim 32, wherein the biological fluid is blood, serum, plasma, saliva, urine, stool, sputum, nasopharyngeal aspirates, cells or tissues.

34. A method of detecting an antibody to a SARS-CoV viral antigen in a sample comprising:
   (a) contacting the sample with the composition of claim 29; and
   (b) detecting the presence of an antibody bound to the composition of claim 29, thereby detecting an antibody to a SARS-CoV viral antigen.

35. The method in accordance with claim 34, wherein the sample is a biological fluid.

36. The method in accordance with claim 35, wherein the biological fluid is blood, serum, plasma, saliva, urine, stool, sputum, nasopharyngeal aspirates, cells or tissues.

37. A method of detecting an antibody to a SARS-CoV viral antigen in a sample comprising:
   (a) contacting the sample with the composition of claim 30; and
   (b) detecting the presence of an antibody bound to the composition of claim 30, thereby detecting an antibody to a SARS-CoV viral antigen.

38. The method in accordance with claim 37, wherein the sample is a biological fluid.

39. The method in accordance with claim 38, wherein the biological fluid is blood, serum, plasma, saliva, urine, stool, sputum, nasopharyngeal aspirates, cells or tissues.

Figure 1-A

```
                                          HpyCH4III
                                       HpaII
                                        NciI
                                       ScrFI
                                      PfoI
                                    StyD4I
                                    BsmAI
                                    BsmBI
                                     AluI
                                    CviJI
                                Fnu4HI
                              HpyCH4V
                               NlaIII
                                 NspI
                                 TseI                                                   BstF5I
   BstUI                         MnlI              MaeIII                                NciI
    HhaI                       FatI                Tsp45I        SfaNI                  ScrFI       DrdI
  BstUI             HphI      HphI                 BbvI       AluI                     HpaII        HgaI
 HinP1I       HpyCH4III   Hpy188I                  BslI       CviJI          AciI    StyD4I   FokI  CviJI
   tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccg
 1 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 100
   agcgcgcaaagccactactgccacttttggagactgtgtacgtcgagggcctctgccagtgtcgaacagacattcgcctacggccctcgtctgttcgggc

a   S  R  V  S  V  M  T  V  K  T  S  D  T  C  S  S  R  R  R  S  Q  L  V  C  K  R  M  P  G  A  D  K  P  V -
b    R  A  F  R  *  *  R  *  K  P  L  T  H  A  A  P  G  D  G  H  S  L  S  V  S  G  C  R  E  Q  T  S  P  -
c     A  R  F  G  D  D  G  E  N  L  *  H  M  Q  L  P  E  T  V  T  A  C  L  *  A  D  A  G  S  R  Q  A  R -
```

```
                                                                                    AciI
                                                                              HpyF10VI
                                                                                BstAPI
                                                                                  MwoI
                                                          BspCNI                  NdeI
                                                          BseMII           Bme1580I
                                                           SfaNI           BsiHKAI
   BstUI                            MseI            Hpy188I       DdeI   Bsp1286I
    HhaI     FauI                 CviJI           TauI         RsaI        Hpy8I
   FauI   AciI                   Cac8I         Fnu4HI        Csp6I     HpyCH4V
  HinP1I  MspA1I      AciI      CviJI           AciI          TatI      ApaLI                  SfaNI
    tcagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaata
101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 200
    agtcccgcgcagtcgcccacaaccgcccacagccccgaccgaattgatacgccgtagtctcgtctaacatgactctcacgtggtatacgccacactttat

a      R  A  R  Q  R  V  L  A  G  V  G  A  G  L  T  M  R  H  Q  S  R  L  Y  *  E  C  T  I  C  G  V  K  Y -
b    S  G  R  V  S  G  C  W  R  V  S  G  L  A  *  L  C  G  I  R  A  D  C  T  E  S  A  P  Y  A  V  *  N  T -
c     Q  G  A  S  A  G  V  G  G  C  R  G  W  L  N  Y  A  A  S  E  Q  I  V  L  R  V  H  H  M  R  C  E  I  -
```

```
                               HpyF10VI
                                   MwoI
                                   BbvI
                                 BbeI                                                 CviJI
                                HaeII                                                HaeIII
                                 HhaI                                                 Cac8I
                                SfaNI                                               Sau96I
                                NlaIV                   HpyCH4III                     AciI
                                 SfoI                  HhaI                        MboII
                                BsaHI                  FspI                     HpyF10VI
                               HinP1I               HinP1I                         MwoI
                             HpyF10VI              Fnu4HI                         BsiEI
                                 NarI              CviJI                         BstKTI
                                 BanI                TseI                          PvuI
                                 KasI         HpyF10VI                             DpnI
  AciI                AciI       MwoI            BglI                              FauI                MnlI
                                                 MwoI                             MboI               EarI
    ccgcacagatgcgtaaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctat
201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 300
    ggcgtgtctacgcattcctcttttatggcgtagtccgcggtaagcggtaagtccgacgcgttgacaacccttcccgctagccacgcccggagaagcgata

a     R  T  D  A  *  G  E  N  T  A  S  G  A  I  R  H  S  G  C  A  T  V  G  K  G  D  R  C  G  P  L  R  Y  -
b      A  Q  M  R  K  E  K  I  P  H  Q  A  P  F  A  I  Q  A  A  Q  L  L  G  R  A  I  G  A  G  L  F  A  I -
c     P  H  R  C  V  R  R  K  Y  R  I  R  R  H  S  P  F  R  L  R  N  C  W  E  G  R  S  V  R  A  S  S  L  L-
```

Figure 1-B

```
                                                                                          TaiI
           Cac8I                                             BstNI                     Hpy99I|
          AluI |                                             ScrFI                 HpyCH4IV ||                 BsrI
         CviJI |                     HpyCH4V                BsaJI|             Hpy188III   | ||                CviJI|
        MspA1I |                    Fnu4HI |               PspGI||             MaeIII  |   | ||               HaeIII|
         PvuII |                   BstF5I| |              StyD4I||             Tsp45I  |   | ||             Hpy99I|| HindIII
        Cac8I | |BbvI                TseI| |FokI MseI  MaeIII  |||BmrI         BsrI |  |   | ||              EaeI||| TspRI|
              | | |  |                   || |   |    |       |    |||   |        | |  |   | ||                 ||||        ||
    tacgccagctggcgaaagggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgccaa
301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 400
    atgcggtcgaccgctttcccccctacacgacgttccgctaattcaacccattgcggtcccaaaagggtcagtgctgcaacattttgctgccggtcacggtt

a    Y  A  S  W  R  K  G  D  V  L  Q  G  D  *  V  G  *  R  Q  G  F  P  S  H  D  V  V  K  R  R  P  V  P  S -
b   T  P  A  G  E  R  G  M  C  C  K  A  I  K  L  G  N  A  R  V  F  P  V  T  T  L  *  N  D  G  Q  C  Q  -
c     R  Q  L  A  K  G  G  C  A  A  R  R  L  S  W  V  T  P  G  F  S  Q  S  R  R  C  K  T  T  A  S  A  K -


                                                    BstF5I         BslI
                                                     RsaI|   Tsp509I  |
   BceAI                                            Csp6I||   FokI   | |
  AluI |          BsmFI                             TatI||MseI |     | |               HpyCH4III
 CviJI |   Tsp509I  |                       BslI     ||||  |   |     | |          BceAI    |  CviJI
     | |        |   |                          |     ||||  |   |     | |            |      |    |
    gcttgaaagagataaattgaacaagtatggtcgtcccctgttgggatgtactattaaacctaaattggggttatctgctaaaaactacggtagagccgtt
401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 500
    cgaactttctctatttaacttgttcataccagcaggggacaaccctacatgataatttggatttaaccccaatagacgattttgatgccatctcggcaa

a      L  K  E  I  N  *  T  S  M  V  V  P  C  W  D  V  L  L  N  L  N  W  G  Y  L  L  K  T  T  V  E  P  F -
b   A  *  K  R  *  I  E  Q  V  W  S  S  P  V  G  M  Y  Y  *  T  *  I  G  V  I  C  *  K  L  R  *  S  R  L -
c    L  E  R  D  K  L  N  K  Y  G  R  P  L  L  G  C  T  I  K  P  K  L  G  L  S  A  K  N  Y  G  R  A  V -


                  Hpy8I                             MmeI
                TspDTI |                           Hpy8I|            PpiI       BstKTI
BbsI             AciI | |                          TaiI||           XcmI|        DpnI|
MboII      XmnI BstUI | |           PpiI  HpyCH4IV  |||         BsaXI ||       MboI ||          BsaXI
    |        |     |  | |              |         |  |||             | ||          | ||             |
    tatgaatgtcttcgcggtggacttgatttactaaagatgatgagaacgtgaactcacaaccatttatgcgttggagagatcgtttcttatttgtgccg
501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 600
    atacttacagaagcgccacctgaactaaatgatttctactactcttgcacttgagtgttggtaaatacgcaacctctctagcaaagaataaacacggc

a     M  N  V  F  A  V  D  L  I  L  L  K  M  M  R  T  *  T  H  N  H  L  C  V  G  E  I  V  S  Y  F  V  P  -
b      *  M  S  S  R  W  T  *  F  Y  *  R  *  *  E  R  E  L  T  T  I  Y  A  L  E  R  S  F  L  I  L  C  R -
c   Y  E  C  L  R  G  G  L  D  F  T  K  D  D  E  N  V  N  S  Q  P  F  M  R  W  R  D  R  F  L  F  C  A  E -


                                                                      NlaIII
                                                                    FatI |
                                                                   RsaI| |
                                                                 Csp6I|| |
                                                                 PstI ||| |          MboII
                                                          HpyCH4V  | ||| |        BstKTI :
                                                            SfcI | | ||| |         DpnI| :   HpyF10VI
      PsiI          CviJI              HphI       BspMI BsmI  | | | ||| |NspI    BclI|| :  AluI    |
        |               |                 |           |     |  | | | ||| |      MboI ||| : CviJI MwoI||
    aagcactttataaagcacaggctgaaacaggtgaaatcaaagggcattacttgaatgctactgcaggtacatgcgaagaaatgatcaaaagagctgtatt
601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 700
    ttcgtgaaatatttcgtgtccgactttgtccactttagtttcccgtaatgaacttacgatgacgtccatgtacgcttctttactagttttctcgacataa

a   K  H  F  I  K  H  R  L  K  Q  V  K  S  K  G  I  T  *  M  L  L  Q  V  H  A  K  K  *  S  K  E  L  Y  L -
b    S  T  L  *  S  T  G  *  N  R  *  N  Q  R  A  L  L  E  C  Y  C  R  Y  M  R  R  N  D  Q  K  S  C  I  -
c     A  L  Y  K  A  Q  A  E  T  G  E  I  K  G  H  Y  L  N  A  T  A  G  T  C  E  E  M  I  K  R  A  V  F -


                      BsiEI
                     BstKTI   NlaIII
                       PvuI    NsiI |                                                                Hpy8I
       BcgI            DpnI| BfrBI | |       HphI                    AluI                             AccI|
  Tsp509I |     Hpy188I   || FatI  | |    BcgI |HinfI                CviJI                            FokI ||
BfaI    | |        MboI   ||HpyCH4V | |  MseI  | | TfiI  AciI      BfaI |CviJI                     MboII ||
    |    | |           |  ||      | | |     |  | |    |     |          |  |   |                         | ||
    tgctagagaattgggcgttccgatcgtaatgcatgactacttaacgggggattcaccgcaaatactagcttggctcattattgccgagataatggtcta
701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 800
    acgatctcttaacccgcaaggctagcattacgtactgatgaattgccccctaagtggcgtttatgatcgaaccgagtaataacggctctattaccagat

a     L  E  N  W  A  F  R  S  *  C  M  T  T  *  R  G  D  S  P  Q  I  L  A  W  L  I  I  A  E  I  M  V  Y  -
b   C  *  R  I  G  R  S  D  R  N  A  *  L  L  N  G  G  I  H  R  K  Y  *  L  G  S  L  L  P  R  *  W  S  T -
c    A  R  E  L  G  V  P  I  V  M  H  D  Y  L  T  G  G  F  T  A  N  T  S  L  A  H  Y  C  R  D  N  G  L  -
```

Figure 1-C

```
AciI NlaIII NspI SphI BsrDI Cac8I HpyF10VI NsiI BstAPI MwoI BfrBI FatI HpyCH4V HpyCH4V HpyCH4III BstF5I
HinfI TfiI FatI NlaIII MboII
BciVI NciI ScrFI BslI HpaII StyD4I
TaiI BsaAI SnaBI HpyCH4IV MaeIII

    cttcttcacatccaccgtgcaatgcatgcggttattgatagacagaagaatcatggtatccacttccgggtattagcaaaagcgttacgtatgtctggtg
801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 900
    gaagaagtgtaggtggcacgttacgtacgccaataactatctgtcttcttagtaccataggtgaaggcccataatcgttttcgcaatgcatacagaccac

a    F  F  T  S  T  V  Q  C  M  R  L  L  I  D  R  R  I  M  V  S  T  S  G  Y  *  Q  K  R  Y  V  C  L  V  -
b     S  S  H  P  P  C  N  A  C  G  Y  *  *  T  E  E  S  W  Y  P  L  P  G  I  S  K  S  V  T  Y  V  W  W -
c    L  L  H  I  H  R  A  M  H  A  V  I  D  R  Q  K  N  H  G  I  H  F  R  V  L  A  K  A  L  R  M  S  G  G -

BstKTI DpnI BaeI MboI BaeI
HpyCH4III KpnI NlaIV RsaI Csp6I Acc65I BanI
BaeI Hpy8I BaeI BsmAI HphI CviJI

    gagatcatattcactctggtaccgtagtaggtaaacttgaaggtgaaagagacataactttgggctttgttgatttactgcgtgatgattttgttgaaca
901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1000
    ctctagtataagtgagaccatggcatcatccatttgaacttccactttctctgtattgaaacccgaaacaactaaatgacgcactactaaaacaacttgt

a    E  I  I  F  T  L  V  P  *  *  V  N  L  K  V  K  E  T  *  L  W  A  L  L  I  Y  C  V  M  I  L  L  N  K -
b     R  S  Y  S  L  W  Y  R  S  R  *  T  *  R  *  K  R  H  N  F  G  L  C  *  F  T  A  *  *  F  C  *  T  -
c     D  H  I  H  S  G  T  V  V  G  K  L  E  G  E  R  D  I  T  L  G  F  V  D  L  L  R  D  D  F  V  E  Q -

BstKTI TaqI DpnI MboI AciI BpuEI BstUI Hpy188III SmlI Hin4I
Hin4I Eco57I Eco57MI BstNI ScrFI PspGI SexAI StyD4I BsaI BsmAI
BsaJI BtgI Hpy188III CviJI MnlI HpyCH4IV TaiI NdeI Cac8I

     agatcgaagtcgcggtatttatttcactcaagattgggtctctttaccaggtgttctacccgtggcttcaggaggtattcacgtttggcatatgcctgct
1001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1100
     tctagcttcagcgccataaataaagtgagttctaacccagagaaatggtccacaagatgggcaccgaagtcctccataagtgcaaaccgtatacggacga

a      I  E  V  A  V  F  I  S  L  K  I  G  S  L  Y  Q  V  F  Y  P  W  L  Q  E  V  F  T  F  G  I  C  L  L -
b     R  S  K  S  R  Y  L  F  H  S  R  L  G  L  F  T  R  C  S  T  R  G  F  R  R  Y  S  R  L  A  Y  A  C  S -
c      D  R  S  R  G  I  Y  F  T  Q  D  W  V  S  L  P  G  V  L  P  V  A  S  G  G  I  H  V  W  H  M  P  A  -

MnlI HpyCH4III FokI SfcI RsaI Csp6I BstF5I HinfI TfiI TaqII TspGWI BstKTI DpnI BglII BstYI MboI Hpy188I
FokI BsaJI StyI BstF5I BceAI
NlaIV HpyF10VI BanI MwoI BstNI ScrFI HhaI PspGI StyD4I HinP1I
AluI CviJI TaqI AluI CviJI

     ctgaccgagatctttggggatgattccgtactacagttcggtggaggaactttaggacatccttggggtaatgcgccaggtgccgtagctaatcgagtag
1101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1200
     gactggctctagaaacccctactaaggcatgatgtcaagccacctccttgaaatcctgtaggaaccccattacgcggtccacggcatcgattagctcatc

a      *  P  R  S  L  G  M  I  P  Y  Y  S  S  V  E  E  L  *  D  I  L  G  V  M  R  Q  V  P  *  L  I  E  * -
b       D  R  D  L  W  G  *  F  R  T  T  V  R  W  R  N  F  R  T  S  L  G  *  C  A  R  C  R  S  *  S  S  S -
c     L  T  E  I  F  G  D  D  S  V  L  Q  F  G  G  G  T  L  G  H  P  W  G  N  A  P  G  A  V  A  N  R  V  A -
```

Figure 1-D

Restriction sites (1201–1300): XbaI, Hpy188III, BfaI, MwoI, HpyF10VI, FatI, NlaIII, NspI, CviJI, AluI, HpyCH4IV, BmgBI, TaiI, MboI, DpnI, BstKTI, TspDTI, DdeI, Bpu10I, Hpy188III, BspCNI, BseMII, Tsp509I, XmnI, MnlI, Hpy188III, NruI, BstUI, TspDTI, CviJI, MwoI, HpyF10VI, CacBI, HpyCH4V, MwoI, HpyF10VI, BbvI, NlaIV, StyD4I, CviJI, BbvI, ScrFI, NciI, HpaII, Bsp1286I, BanII, BfaI, TseI, CviJI, AluI, Fnu4HI

```
     ctctagaagcatgtgtaaaagctcgtaatgaaggacgtgatcttgctcaggaaggtaatgaaattattcgcgaggcttgcaaatggagcccggaactagc
1201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1300
     gagatcttcgtacacattttcgagcattacttcctgcactagaacgagtccttccattactttaataagcgctccgaacgtttacctcgggccttgatcg

a     L  *  K  H  V  *  K  L  V  M  K  D  V  I  L  L  R  K  V  M  K  L  F  A  R  L  A  N  G  A  R  N  *  L  -
b      S  R  S  M  C  K  S  S  *  *  R  T  *  S  C  S  G  R  *  *  N  Y  S  R  G  L  Q  M  E  P  G  T  S   -
c       L  E  A  C  V  K  A  R  N  E  G  R  D  L  A  Q  E  G  N  E  I  I  R  E  A  C  K  W  S  P  E  L  A  -
```

Restriction sites (1301–1400): TseI, Fnu4HI, MboI, DpnI, BstKTI, Tsp509I, MseI, HpyCH4V, Fnu4HI, TseI, Hpy8I, TspRI, HpyCH4IV, BtsI, BbvI, TaiI, HpyCH4III, AccI, Hpy8I, Tsp509I

```
     tgctgcttgtgaagtatggaaagagatcgtatttaattttgcagcagtggacgttttggataagtaaaaacagtagacattagcagataaattagcagga
1301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1400
     acgacgaacacttcatacctttctctagcataaattaaaacgtcgtcacctgcaaaacctattcatttttgtcatctgtaatcgtctatttaatcgtcct

a     L  L  V  K  Y  G  K  R  S  Y  L  I  L  Q  Q  W  T  F  W  I  S  K  N  S  R  H  *  Q  I  N  *  Q  E  -
b    C  C  L  *  S  M  E  R  D  R  I  *  F  C  S  S  G  R  F  G  *  V  K  T  V  D  I  S  R  *  I  S  R  K -
c     A  A  C  E  V  W  K  E  I  V  F  N  F  A  A  V  D  V  L  D  K  *  K  Q  *  T  L  A  D  K  L  A  G  -
```

Restriction sites (1401–1500): BpuEI, SmlI, Tsp509I, MseI, Tsp509I, Tsp509I, HpyCH4V, Tsp509I, MseI, Sau96I, HaeIII, CviJI, CviJI

```
     aataaagaaggataaggagaaagaactcaagtaattatccttcgttctcttaattgaattgcaattaaactcggcccaatcttttactaaaaggattgag
1401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1500
     ttatttcttcctattcctctttcttgagttcattaataggaagcaagagaattaacttaacgttaatttgagccgggttagaaaatgattttcctaactc

a      I  K  K  D  K  E  K  E  L  K  *  L  S  F  V  L  L  I  E  L  Q  L  N  S  A  Q  S  F  T  K  R  I  E  -
b       *  R  R  I  R  R  K  N  S  S  N  Y  P  S  F  S  *  L  N  C  N  *  T  R  P  N  L  L  L  K  G  L  S -
c     N  K  E  G  *  G  E  R  T  Q  V  I  I  L  R  S  L  N  *  I  A  I  K  L  G  P  I  F  Y  *  K  D  *  A -
```

Restriction sites (1501–1600): TfiI, HinfI, HpyCH4V, DdeI, BfaI, XbaI, Hpy188III, BfaI, PleI, MlyI

```
     ccgaatacaacaaagattctattgcatatattttgactaagtatatacttacctagatatacaagatttgaaatacaaaatctagaaaactaaatcaaaa
1501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1600
     ggcttatgttgtttctaagataacgtatataaaactgattcatatatgaatggatctatatgttctaaactttatgttttagatcttttgatttagtttt

a     P  N  T  T  K  I  L  L  H  I  F  *  L  S  I  Y  L  P  R  Y  T  R  F  E  I  Q  N  L  E  N  *  I  K  I -
b      R  I  Q  Q  R  F  Y  C  I  Y  F  D  *  V  Y  T  Y  L  D  I  Q  D  L  K  Y  K  I  *  K  T  K  S  K  -
c       E  Y  N  K  D  S  I  A  Y  I  L  T  K  Y  I  L  T  *  I  Y  K  I  *  N  T  K  S  R  K  L  N  Q  N  -
```

Figure 1-E

```
DdeI HinfI MboI DpnI BstKTI TaqI AlwI CviJI AluI Cac8I HpyCH4V FatI Cac8I SphI NspI NlaIII SfcI Cac8I HpyCH4V PstI SalI TaqI AccI Hpy8I HincII HpyCH4III Hpy188I TatI Csp6I RsaI

     tctaagactcaaatctttctattgttgtcttggatcatcgagcttgcatgcctgcagtcgactttcacagtttccattctgaaatgttctctgtactata
1601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1700
     agattctgagtttagaaagataacaacagaacctagtagctcgaacgtacggacgtcagctgaaagtgtcaaaggtaagactttacaagagacatgatat

a     * D S N L S I V V L D H R A C M P A V D F H S F H S E M F S V L * -
b    S K T Q I F L L L S W I I E L A C L Q S T F T V S I L K C S L Y Y N -
c     L R L K S F Y C C L G S S S L H A C S R L S Q F P F * N V L C T I -
```

```
HinfI TfiI Tsp509I Hpy8I MboII AccI Hpy188I EarI MnlI CviJI HpyCH4V

     ataaatagtaagtgaatcaacttactactaaaaaaattagtagacttcctcttggaatagaaatagcctatttctacatagggaaagtcgtgtgcaatg
1701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1800
     tatttatcattcacttagttgaatgatgatttttttaatcatctgaaggagaaccttatctttatcggataaagatgtatccctttcagcacacgttac

a     * I V S E S T Y Y * K N * * T S S S E * K * P I S T * G K S C A M -
b      K * * V N Q L T T K K I S R L P L R N R N S L F L H R E S R V Q * -
c    I N S K * I N L L L K K L V D F L P G I E I A Y F Y I G K V V C N E -
```

```
BsrDI HpyCH4V Cac8I TspDTI MnlI MaeIII Tsp509I FokI BstF5I Hpy188I SpeI BfaI BccI BsmFI StyD4I HpaII ScrFI NciI HinfI TaqI StyD4I XmaI StyD4I BsaJI AvaI ScrFI NciI HpaII SmaI ScrFI NciI MmeI PleI MlyI

     aaaaatgcaagcacgatttggggagaggtttttctctattgtaacaaggaataattatctactccatccgactagttccgggttcgagtcccgggcaac
1801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1900
     tttttacgttcgtgctaaacccctctccaaaaagagataacattgttccttattaatagatgaggtaggctgatcaaggcccaagctcagggcccgttg

a     K N A S T I W G E V F S L L * Q G I I I Y S I R L V P G S S P G Q P -
b      K M Q A R F G E R F F L Y C N K E * L S T P S D * F R V R V P G N -
c      K C K H D L G R G F F S I V T R N N Y L L H P T S S G F E S R A T -
```

```
NdeI BfaI BseMII BspCNI Hpy188I DdeI TspDTI Tsp509I ApoI Tsp509I ApoI

     ccatatggaaactagaaaggagcaatctgagttttgatttttcactcacttcatttacaaaatttttttggtttggtaaatttctgttgtatggatatacaa
1901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2000
     ggtatacctttgatctttcctcgttagactcaaaactaaaaagtgagtgaagtaaatgttttaaaaaaaccaaaccatttaaaacaacatacctatatgtt

a     I W K L E R S N L S F D F S L T S F T K F F G L V N F V V W I Y N -
b    P Y G N * K G A I * V L I F H S L H L Q N F L V W * I L L Y G Y T T -
c     H M E T R K E Q S E F * F F T H F I Y K I F W F G K F C C M D I Q -
```

Figure 1-F

```
HpyCH4III  CviJI  Cac8I  CviJI  HincII  Hpy8I  HpyCH4III  MseI  CviJI  BfaI  MseI  HpyCH4IV  AflIII  BsaAI  TaiI  HinfI

     ctgtcggggctggcttggttgacattggtatatagtctatattatactgttaaataacaagccttctattatctttctagttaatacgtgtgcttgggag
2001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2100
     gacagccccgaccgaaccaactgtaaccatatatcagatataatatgacaatttattgttcggaagataatagaaagatcaattatgcacacgaaccctc

a     C  R  G  W  L  G  *  H  W  Y  I  V  Y  I  I  L  L  N  N  K  P  S  I  I  F  L  V  N  T  C  A  W  E  -
b      V  G  A  G  L  V  D  I  G  I  *  S  I  L  Y  C  *  I  T  S  L  L  L  S  F  *  L  I  R  V  L  G  S  -
c    L  S  G  L  A  W  L  T  L  V  Y  S  L  Y  Y  T  V  K  *  Q  A  F  Y  Y  L  S  S  *  Y  V  C  L  G  V  -

MlyI  Tsp509I  HpyCH4V  PleI  BstKTI  DpnI  BglII  BstYI  MboI  BsaJI  BtgI  FatI  NcoI  StyI
AciI  BstUI  HhaI  BstUI  HinP1I  TauI  AciI  Fnu4HI  CviJI  HaeIII  NlaIII  EaeI  FauI
MslI  FatI  FauI  Sau96I  NlaIII  AvaII  BtgI  BsaJI  AciI  MspA1I  BstUI  AciI  SacII  Sau96I  PspOMI
Eco0109I  Sau96I  NlaIV  NlaIV  HaeIII  CviJI  Cac8I  Bsp1286I  Bme1580I  BanII  ApaI  AciI  BfaI  BslI  FauI
Tsp45I  MaeIII  AgeI  BsaWI  BsrFI

     tccttgcaatttgaataaaccaagatcttaccatggccgcgcggggttctcatcatcatcatcatggtccgcgggggcccgctagtggtagtgaccttga
2101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2200
     aggaacgttaaacttatttggttctagaatggtaccggcgcgccccaagagtagtagtagtagtaccaggcgcccccgggcgatcaccatcactggaact

a     S  L  Q  F  E  *  T  K  I  L  P  W  P  R  G  V  L  I  I  I  I  M  V  R  G  G  P  L  V  V  V  T  L  T  -
b      P  C  N  L  N  K  P  R  S  Y  H  G  R  A  G  F  S  S  S  S  S  W  S  A  G  A  R  *  W  *  *  P  *  -
c      L  A  I  *  I  N  Q  D  L  T  M  A  A  R  G  S  H  H  H  H  H  G  P  R  G  P  A  S  G  S  D  L  D  -

HpaII  ApaLI  HpyCH4V  Hpy8I  Bsp1286I  BsiHKAI  Bme1580I  CviJI  AluI  Tsp509I  TspDTI  MnlI
Hpy8I  Hpy188III  Tsp509I  ApoI  MboI  DpnI  BstKTI  TspDTI  Hpy188I

     ccggtgcaccactttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttactatcctgatgaaattttagatcagac
2201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2300
     ggccacgtggtgaaaaactactacaagttcgaggattaatgtgagttgtatgaagtagatactccccccaaatgataggactactttaaaatctagtctg

a      G  A  P  L  L  M  M  F  K  L  L  I  T  L  N  I  L  H  L  *  G  G  F  T  I  L  M  K  F  L  D  Q  T  -
b     P  V  H  H  F  *  *  C  S  S  S  *  L  H  S  T  Y  F  I  Y  E  G  G  L  L  S  *  *  N  F  *  I  R  H  -
c      R  C  T  T  F  D  D  V  Q  A  P  N  Y  T  Q  H  T  S  S  M  R  G  V  Y  Y  P  D  E  I  F  R  S  D  -

MseI  DdeI  Hpy188III  MboII  BspCNI  BseMII  MaeIII  TspDTI  MseI  AseI  HpyCH4IV  TaiI  BccI  MseI

     actctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactattaatcatacgtttggcaaccctgtcatacctttta
2301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2400
     tgagaaataaattgagtcctaaataaagaaggtaaaataagattacaatgtcccaaagtatgataattagtatgcaaaccgttgggacagtatggaaaat

a      L  F  I  *  L  R  I  Y  F  F  H  F  I  L  M  L  Q  G  F  I  L  L  I  I  R  L  A  T  L  S  Y  L  L  -
b       S  L  F  N  S  G  F  I  S  S  I  L  F  *  C  Y  R  V  S  Y  Y  *  S  Y  V  W  Q  P  C  H  T  F  *  -
c     T  L  Y  L  T  Q  D  L  F  L  P  F  Y  S  N  V  T  G  F  H  T  I  N  H  T  F  G  N  P  V  I  P  F  K  -
```

Figure 1-G

```
                                                                                            HpyCH4III
                    Fnu4HI                                                                    TspDTI
  BstF5I             TseI              BsaJI                          NlaIII   MaeIII
  BbvI       FokI              TspGWI  BtgI    BslI           FatI             Tsp45I                     MseI
     aggatggtatttatttgctgccacagagaaatcaaatgttgtccgtggttgggtttttggttctaccatgaacaacaagtcacagtcggtgattattat
2401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2500
     tcctaccataaataaacgacggtgtctctttagtttacaacaggcaccaacccaaaaaccaagatggtacttgttgttcagtgtcagccactaataata

a     R  M  V  F  I  L  L  P  Q  R  N  Q  M  L  S  V  V  G  F  L  V  L  P  *  T  T  S  H  S  R  *  L  L  L -
b      G  W  Y  L  F  C  C  H  R  E  I  K  C  C  P  W  L  G  F  W  F  Y  H  E  Q  Q  V  T  V  G  D  Y  Y  -
c       D  G  I  Y  F  A  A  T  E  K  S  N  V  V  R  G  W  V  F  G  S  T  M  N  N  K  S  Q  S  V  I  I  I -

                                                                                            Hpy8I
                                                                                           RsaI
                                                                                         Csp6I
                                                                                      NlaIII
                                                                                   BsaJI
                                                                                   BtgI
                             NlaIII                                                FatI
                              NspI       MaeIII                                    NcoI
  Tsp509I                  MaeIII        Tsp45I                                    StyI
  HphI                     FatI   Tsp509I
     taacaattctactaatgttgttatacgagcatgtaactttgaattgtgtgacaacccttctcttgctgtttctaaacccatgggtacacagacacatact
2501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2600
     attgttaagatgattacaacaatatgctcgtacattgaaacttaacacactgttgggaagaaacgacaaagatttgggtacccatgtgtctgtgtatga

a      T  I  L  L  M  L  L  Y  E  H  V  T  L  N  C  V  T  T  L  S  L  L  F  L  N  P  W  V  H  R  H  I  L -
b     *  Q  F  Y  *  C  C  Y  T  S  M  *  L  *  I  V  *  Q  P  F  L  C  C  F  *  T  H  G  Y  T  D  T  Y  Y -
c     N  N  S  T  N  V  V  I  R  A  C  N  F  E  L  C  D  N  P  F  F  A  V  S  K  P  M  G  T  Q  T  H  T  -

                 HpyCH4V
                  Hin4I
             Tsp509I          RsaI
              MseI           Csp6I
              NsiI           SfaNI
           BfrBI             TatI                                                         DraI
  TaqI  HpyCH4V           TaqI      Hpy188I      Hin4I      Hpy188I           Tsp509I MseI
     atgatattcgataatgcatttaattgcactttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgag
2601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2700
     tactataagctattacgtaaattaacgtgaaagctcatgtatagactacggaaaagcgaactacaaagtcttttcagtccattaaaatttgtgaatgctc

a     *  Y  S  I  M  H  L  I  A  L  S  S  T  Y  L  M  P  F  R  L  M  F  Q  K  S  Q  V  I  L  N  T  Y  E  -
b      D  I  R  *  C  I  *  L  H  F  R  V  H  I  *  C  L  F  A  *  C  F  R  K  V  R  *  F  *  T  L  T  R -
c     M  I  F  D  N  A  F  N  C  T  F  E  Y  I  S  D  A  F  S  L  D  V  S  E  K  S  G  N  F  K  H  L  R  E -

                                                                       BstKTI
                                                                        DpnI
   DraI                                                               MboI
   MseI BccI                      PsiI  CviJI      SfcI      Hpy188III                  MseI
     agtttgtgtttaaaaataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaa
2701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2800
     tcaaacacaaatttttatttctacccaaagagatacaaatattcccgatagttggatatctacatcaagcactagatggaagaccaaaattgtgaaactt

a     S  L  C  L  K  I  K  M  G  F  S  M  F  I  R  A  I  N  L  *  M  *  F  V  I  Y  L  L  V  L  T  L  *  N -
b      V  C  V  *  K  *  R  W  V  S  L  C  L  *  G  L  S  T  Y  R  C  S  S  *  S  T  F  W  F  *  H  F  E  -
c       F  V  F  K  N  K  D  G  F  L  Y  V  Y  K  G  Y  Q  P  I  D  V  V  R  D  L  P  S  G  F  N  T  L  K -
```

Figure 1-H

```
acctatttttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacgtcagctgca
2801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2900
tggataaaaattcaacggagaaccataattgtaatgtttaaaatctcggtaagaatgtcggaaaagtggacgagttctgtaaaccccgtgcagtcgacgt

a   L  F  L  S  C  L  L  V  L  T  L  Q  I  L  E  P  F  L  Q  P  F  H  L  L  K  T  F  G  A  R  Q  L  Q  -
b  T  Y  F  *  V  A  S  W  Y  *  H  Y  K  F  *  S  H  S  Y  S  L  F  T  C  S  R  H  L  G  H  V  S  C  S -
c  P  I  F  K  L  P  L  G  I  N  I  T  N  F  R  A  I  L  T  A  F  S  P  A  Q  D  I  W  G  T  S  A  A  -

gcctatttttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgttgattgttctcaaaatccacttg
2901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3000
cggataaaaacaaccgataaatttcggttgatgtaaatacgagttcatactacttttaccatgttagtgtctacgacaactaacaagagttttaggtgaac

a   P  I  L  L  A  I  *  S  Q  L  H  L  C  S  S  M  M  K  M  V  Q  S  Q  M  L  L  I  V  L  K  I  H  L  -
b    L  F  C  W  L  F  K  A  N  Y  I  Y  A  Q  V  *  *  K  W  Y  N  H  R  C  C  *  L  F  S  K  S  T  C  -
c  A  Y  F  V  G  Y  L  K  P  T  T  F  M  L  K  Y  D  E  N  G  T  I  T  D  A  V  D  C  S  Q  N  P  L  A -

ctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaatttcagggttgttccctcaggagatgttgtgagattccc
3001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3100
gacttgagtttacgagacaattctcgaaactctaactgtttccttaaatggtctggagattaaagtcccaacaagggagtcctctacaacactctaaggg

a  L  N  S  N  A  L  L  R  A  L  R  L  T  K  E  F  T  R  P  L  I  S  G  L  F  P  Q  E  M  L  *  D  S  L -
b   *  T  Q  M  L  C  *  E  L  *  D  *  Q  R  N  L  P  D  L  *  F  Q  G  C  S  L  R  R  C  C  E  I  P  -
c   E  L  K  C  S  V  K  S  F  E  I  D  K  G  I  Y  Q  T  S  N  F  R  V  V  P  S  G  D  V  V  R  F  P  -

taatattacaaacttgtgtccttttggagaggtttttaatgctactaaattcccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgct
3101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3200
attataatgtttgaacacaggaaaacctctccaaaaattacgatgatttaagggaagacagatacgtaccctctcttttttttaaagattaacacaacga

a   I  L  Q  T  C  V  L  L  E  R  F  L  M  L  L  N  S  L  L  S  M  H  G  R  E  K  K  F  L  I  V  L  L  -
b  *  Y  Y  K  L  V  S  F  W  R  G  F  *  C  Y  *  I  P  F  C  L  C  M  G  E  K  K  N  F  *  L  C  C  * -
c   N  I  T  N  L  C  P  F  G  E  V  F  N  A  T  K  F  P  S  V  Y  A  W  E  R  K  K  I  S  N  C  V  A  -

gattactctgtgctctacaactcaacatttttttcaaccttttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatg
3201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3300
ctaatgagacacgagatgttgagttgtaaaaaaagttggaaattcacgataccgcaaagacggtgattcaacttactagaaacgaagaggttacagatac

a   I  T  L  C  S  T  T  Q  H  F  F  Q  P  L  S  A  M  A  F  L  P  L  S  *  M  I  F  A  S  P  M  S  M  -
b    L  L  C  A  L  Q  L  N  I  F  F  N  L  *  V  L  W  R  F  C  H  *  V  E  *  S  L  L  L  Q  C  L  C  -
c  D  Y  S  V  L  Y  N  S  T  F  F  S  T  F  K  C  Y  G  V  S  A  T  K  L  N  D  L  C  F  S  N  V  Y  A -
```

Restriction sites (2801–2900): MseI, MnlI, MseI, ApoI, Tsp509I, CviJI, CviJI, BpuEI, HphI, Hpy188III, SmlI, AarI, BspMI, BbvI, HpyCH4IV, Bsp1286I, Bme1580I, BmgBI, TaiI, TseI, PvuII, MspA1I, CviJI, AluI, SfcI, Fnu4HI, TseI, HpyCH4V, Fnu4HI, PstI

Restriction sites (2901–3000): CviJI, BbvI, CviJI, DraI, MseI, CviJI, BpuEI, SmlI, Csp6I, SfaNI, RsaI, TspDTI, AlwNI

Restriction sites (3001–3100): MseI, AluI, CviJI, ApoI, Tsp509I, Tsp509I, MnlI, BsaXI, DdeI, Bsu36I, Hpy188III, MnlI, BspCNI, BseMII, TfiI, HinfI

Restriction sites (3101–3200): BsaXI, SspI, MnlI, MseI, ApoI, Tsp509I, HpyCH4V, FatI, BfrBI, NsiI, NlaIII, ApoI, Tsp509I, Tsp509I

Restriction sites (3201–3300): Bsp1286I, BsiHKAI, MseI, DdeI, MboI, DpnI, BstKTI, HpyCH4V

Figure 1-I

```
                                          BstNI
                                          ScrFI
                                          HaeII
                                          HhaI
                                          PspGI
                                          StyD4I                                      TspDTI
   HinfI                                  BsaXI                       Tsp509I  Tsp509I
    TfiI     BsaXI                        HinP1I         BsrI          PsiI  PsiI            FatI
      cagattcttttgtagtcaagggagatgatgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcat
 3301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3400
      gtctaagaaaacatcagttccctctactacattctgtttatcgcggtcctgtttgaccacaataacgactaatattaatatttaacggtctactaaagta

a      Q  I  L  L  *  S  R  E  M  M  *  D  K  *  R  Q  D  K  L  V  L  L  L  I  I  I  I  N  C  Q  M  I  S  W -
b       R  F  F  C  S  Q  G  R  *  C  K  T  N  S  A  R  T  N  W  C  Y  C  *  L  *  L  *  I  A  R  *  F  H  -
c        D  S  F  V  V  K  G  D  D  V  R  Q  I  A  P  G  Q  T  G  V  I  A  D  Y  N  Y  K  L  P  D  D  F  M -

                                                                                                 CviJI
                                                                                                 HaeIII
                                                                                               EcoO109I
                                                                                                 Sau96I
                                                                                           Bpu10I
                                                                                             DdeI
                                                                                             AluI
                                                                                             CviJI
                                                                                          Cac8I
                                                                  Tsp509I               HindIII
                         SfaNI                              BsrI                       NlaIII
 NlaIII                  BfaI                        Tsp509I PsiI  PsiI      DdeI  FatI
      gggttgtgtccttgcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggccc
 3401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3500
      cccaacacaggaacgaaccttatgatccttgtaactacgatgaagttgaccattaatattaatatttatatccatagaatctgtaccgttcgaatccggg

a       V  V  S  L  L  G  I  L  G  T  L  M  L  L  Q  L  V  I  I  I  I  N  I  G  I  L  D  M  A  S  L  G  P  -
b      G  L  C  P  C  L  E  Y  *  E  H  *  C  Y  F  N  W  *  L  *  L  *  I  *  V  S  *  T  W  Q  A  *  A  L -
c       G  C  V  L  A  W  N  T  R  N  I  D  A  T  S  T  G  N  Y  N  Y  K  Y  R  Y  L  R  H  G  K  L  R  P   -

                                             HpyCH4V              AarI     CviJI
                                          HpyF10VI                BspMI    HaeIII
                                            BstAPI               Tsp509I    MscI
  BsmAI                        BccI          MwoI                  MseI    EaeI  MseI
      tttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccattaaatgattatggttttt
 3501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3600
      aaactctctctgtatagattacacggaaagaggggactaccgtttggaacgtggggtggacgagaattaacaataaccggtaatttactaataccaaaaa

a       L  R  E  T  Y  L  M  C  L  S  P  L  M  A  N  L  A  P  H  L  L  L  I  V  I  G  H  *  M  I  M  V  F   -
b        *  E  R  H  I  *  C  A  F  L  P  *  W  Q  T  L  H  P  T  C  S  *  L  L  L  A  I  K  *  L  W  F  L -
c      F  E  R  D  I  S  N  V  P  F  S  P  D  G  K  P  C  T  P  P  A  L  N  C  Y  W  P  L  N  D  Y  G  F  Y -

                                                                     HpyCH4III
                                                                     BsaJI
                                                                      BtgI
                                                                     CviJI
                                                                    HaeIII
                                                                     EaeI
                                                                   HpaII
                                          RsaI                    BsrFI                Tsp509I
                                          ScaI               HpyCH4V             AvaII
                  CviJI                  Csp6I          DraI                     Hpy8I
            BsrI                         TatI           MseI                     Sau96I
      acaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaactttaaatgcaccggccacggtttgtggaccaaaattatccac
 3601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3700
      tgtggtgatgaccgtaaccgatggttggaatgtctcaacatcatgaaagaaaacttgaaatttacgtggccggtgccaaacacctggttttaataggtg

a      T  P  L  L  A  L  A  T  N  L  T  E  L  *  Y  F  L  L  N  F  *  M  H  R  P  R  F  V  D  Q  N  Y  P  L -
b       H  H  Y  W  H  W  L  P  T  L  Q  S  C  S  T  F  F  *  T  F  K  C  T  G  H  G  L  W  T  K  I  I  H   -
c        T  T  T  G  I  G  Y  Q  P  Y  R  V  V  V  L  S  F  E  L  L  N  A  P  A  T  V  C  G  P  K  L  S  T  -

                                                              HincII
                                                               HpaI
                                                   BsrI        Hpy8I
                            Tsp509I MlyI            RsaI        MboII
                     Tsp509I        MseI   BaeI    TspRI        MseI
  TspRI MseI    BsrI   TspRI    MseI   PleI HinfI    Csp6I   BsrI               BaeI            BceAI Tsp509I
      tgaccttattaagaaccagtgtgtcaattttaattttaatggactcactggtactggtgtgttaactccttcttcaaagagatttcaaccatttcaacaa
 3701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3800
      actggaataattcttggtcacacagttaaaattaaaattacctgagtgaccatgaccacacaattgaggaagaagtttctctaaagttggtaaagttgtt

a       T  L  L  R  T  S  V  S  I  L  I  L  M  D  S  L  V  L  V  C  *  L  L  L  Q  R  D  F  N  H  F  N  N  -
b      *  P  Y  *  E  P  V  C  Q  F  *  F  *  W  T  H  W  Y  W  C  V  N  S  F  F  K  E  I  S  T  I  S  T  I -
c       D  L  I  K  N  Q  C  V  N  F  N  F  N  G  L  T  G  T  G  V  L  T  P  S  S  K  R  F  Q  P  F  Q  Q  -
```

Figure 1-J

```
                                       BstKTI
                                        DpnI|
                                      BstYI ||
                                        MboI ||
                                  Hpy188III | ||
      CviJI                    TspRI      | | ||
     HaeIII          TspGWI  HinfI |  TaqI| | ||                     HphI
       EaeI |   Hpy188I   |   TfiI |AlwI || | ||           Hpy188I SspI |                         Tsp509I
         | |          |   |      | |    | || | ||                 |    | |                             |
     tttggccgtgatgtttctgatttcactgattccgttcgagatcctaaaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaa
3801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3900
     aaaccggcactacaaagactaaagtgactaaggcaagctctaggattttgtagactttataatctgtaaagtggaacgagaaaacccccacattcacatt

a      L  A  V  M  F  L  I  S  L  I  P  F  E  I  L  K  H  L  K  Y  *  T  F  H  L  A  L  L  G  V  *  V  *  -
b       W  P  *  C  F  *  F  H  *  F  R  S  R  S  *  N  I  *  N  I  R  H  F  T  L  L  F  W  G  C  K  C  N  -
c     F  G  R  D  V  S  D  F  T  D  S  V  R  D  P  K  T  S  E  I  L  D  I  S  P  C  S  F  G  G  V  S  V  I  -

                                                                                                BstKTI
                                                                                                 DpnI|
                                                                                               MboI ||
                                                                                          HpyCH4V | ||
                                                                                           NlaIII | ||
                                                                                       HpyF10VI | | ||
                                                                                         BstAPI|| | ||
          TspDTI                                                    HpyCH4V                      || | ||
          BstNI |                                   Eco57I HincII      :          Tsp509I        || | ||
          ScrFI |                                  Eco57MI   HpaI      :        SfcI     |       || | ||
          PspGI | |                              Hpy188III|  Hpy8I     :     TspDTI|     | MwoI|| | ||
        .StyD4I | |             Hpy188I               BsgI || MseI|    :   TspRI  ||     |FatI|| | ||
              | | |                   |                  | ||    ||    |       |  ||     |  |||| | ||
     ttacacctggaacaaatgcttcatctgaagttgctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacacc
3901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4000
     aatgtggaccttgtttacgaagtagacttcaacgacaagatatagttctacaattgacgtgactacaaagatgtcgttaagtacgtctagttgagtgtgg

a     L  H  L  E  Q  M  L  H  L  K  L  L  F  Y  I  K  M  L  T  A  L  M  F  L  Q  Q  F  M  Q  I  N  S  H  Q  -
b      Y  T  W  N  K  C  F  I  *  S  C  C  S  I  S  R  C  *  L  H  *  C  F  Y  S  N  S  C  R  S  T  H  T  -
c       T  P  G  T  N  A  S  S  E  V  A  V  L  Y  Q  D  V  N  C  T  D  V  S  T  A  I  H  A  D  Q  L  T  P  -

                                                                                  HincII
                                                                                   Hpy8I
                                                                                    AccI|
                                                          SmlI                      TaqI|
              HhaI                                      HinfI |             BlpI  NlaIII||
           HinP1I |                             Hpy188III   | |             DdeI    NspI||
         AluI   | |                      BpuEI     MlyI |   | |      CviJI BspCNI  AluI|     SalI||
        CviJI   | |                       BsrI :    PleI |   | |   Cac8I |BseMII|  CviJI|FatI  |||
            |   | |                          | :       | |   | |       | |     ||      ||   |  |||
     agcttggcgcatatattctactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacatt
4001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4100
     tcgaaccgcgtatataagatgacctttgttacataaggtctgagttcgtccgacagaatatcctcgactcgtacagctgtgaagaatactcacgctgtaa

a      L  G  A  Y  I  L  L  E  T  M  Y  S  R  L  K  Q  A  V  L  *  E  L  S  M  S  T  L  L  M  S  A  T  F  -
b     S  L  A  H  I  F  Y  W  K  Q  C  I  P  D  S  S  R  L  S  Y  R  S  *  A  C  R  H  F  L  *  V  R  H  S  -
c      A  W  R  I  Y  S  T  G  N  N  V  F  Q  T  Q  A  G  C  L  I  G  A  E  H  V  D  T  S  Y  E  C  D  I  -

                                                                                                 ClaI
                                                                                                 TaqI
                                                                                              TauI |
                                                                                            AciI | |
                                                                                           BsiEI | |
                                                                                          Fnu4HI | |
                                                                                          CviJI| | |
                                                                                         HaeIII| | |
                                                                                            TauI| | |
                                                                                         BsrI || | |  Tsp509I
                                                                          BsaBI          EaeI || | |    BanII |
                                                                        EcoRV |          EagI || | |  BsiHKAI |
                                                                       PstI | |       Fnu4HI || | | Bsp1286I |
                                                                   HpyCH4V| | |         NotI || | |     SacI |
                                                                 ApoI   | | | |        TspRI || | |    AluI| |
          Cac8I                                                 EcoRI   | | | |         AciI| || | |   CviJI| |
          AluI |                MaeIII                         Tsp509I  | | | |      BccI || || | | |EcoICRI| |
         CviJI |                BfaI |  HpyCH4III     BfaI    CviJI |  SfcI| | | |    BstXI| || || | | | SfaNI| |
             | |                   | |          |         |       | |     | | | : |         || || || : | |  || | |
     cctattggagctggcatttgtgctagttaccatacagTTAGACTAGAAagccgaattctgcagatatccatcacactggcggccgcatcgatgagctct
4101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4200
     ggataacctcgaccgtaaacacgatcaatggtatgtcAATCTGATCTTtcggcttaagacgtctataggtagtgtgaccgccggcgtagctactcgaga

a      L  L  E  L  A  F  V  L  V  T  I  Q  L  D  *  K  S  R  I  L  Q  I  S  I  T  L  A  A  A  S  M  S  S  -
b       Y  W  S  W  H  L  C  *  L  P  Y  S  *  T  R  K  A  E  F  C  R  Y  P  S  H  W  R  P  H  R  *  A  L  -
c     P  I  G  A  G  I  C  A  S  Y  H  T  V  R  L  E  K  P  N  S  A  D  I  H  H  T  G  G  R  I  D  E  L  *  -
```

Figure 1-K

```
        MseI
        PacI                                              BfaI
  Tsp509I |                                               AluI|
     AseI||                                              CviJI|                                   Eco57I
     MseI||  |CviJI                Tsp509I               MboII|                                   Eco57MI
         ||  |                        |                     ||                                      |
     aattaattaaggcttttctgctaacatatagcaatttttgaagaaaggaaagctagaaataccaatatcttgctgaagcaagatattgggtatttcttt
4201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4300
     ttaattaattccgaaaagacgattgtatatcgttaaaaacttctttcctttcgatctttatgggttatagaacgacttcgttctataacccataaagaaa

a     N  *  L  R  L  F  C  *  H  I  A  I  F  E  E  R  K  A  R  N  T  Q  Y  L  A  E  A  R  Y  W  V  F  L  F -
b      I  N  *  G  F  S  A  N  I  *  Q  F  L  K  K  G  K  L  E  I  P  N  I  L  L  K  Q  D  I  G  Y  F  F  -
c       L  I  K  A  F  L  L  T  Y  S  N  F  *  R  K  E  S  *  K  Y  P  I  S  C  *  S  K  I  L  G  I  S  F -

                                        Hin4I                                                     ApoI
                              ApoI      HincII                                                    EcoRI
                              EcoRI     HpaI                     HpyCH4V                        Tsp509I
           HinfI            Tsp509I     Hpy8I                    BciVI|      Hpy188III          RsaI |
           TfiI           Hpy188I|      MseI| Hpy99I          TspDTI  ||Hin4I  BssSI |          Csp6I| |
             |                   ||         ||   |                 |  ||  |        | |              || |
     tttttttatttttgaatctttctattctgaattcagttaacgacgagatttagtatcctttcttgcactttcataactcgtgaaatgccgagttggtacga
4301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4400
     aaaaaaataaaaacttagaaagataagacttaagtcaattgctgctctaaatcataggaaagaacgtgaaagtattgagcactttacggctcaaccatgct

a      F  L  F  *  I  F  L  F  *  I  Q  L  T  T  R  F  S  I  L  S  C  T  F  I  T  R  E  M  P  S  W  Y  E  -
b     F  F  Y  F  E  S  F  Y  S  E  F  S  *  R  R  D  L  V  S  F  L  A  L  S  *  L  V  K  C  R  V  G  T  N -
c      F  F  I  L  N  L  S  I  L  N  S  V  N  D  E  I  *  Y  P  F  L  H  F  H  N  S  *  N  A  E  L  V  R   -

                                                                                          BsrDI
                                                                                        FauI |
                                                                  BstUI          CviJI  | |
                                                                   NruI         HaeIII  | |
                                                             Hpy188III|           MscI  | |
                              BslI                               HinfI||         EaeI |  | |        SfaNI
       Tsp509I            EcoNI |                                 TfiI||      TspDTI| |  | |  AciI    |
             |                 ||                                    | ||           || |  | |    |     |
     attcccccaatttgcgacctaccataggatttgttatgtaaataggtatatgttccttttccattatgaatcgcgattgtatggccaaccattgcgggtag
4401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4500
     taagggggttaaacgctggatggtatcctaaacaatacatttatccatatacaaggaaaggtaatacttagcgctaacataccggttggtaacgcccatc

a     F  P  Q  F  A  T  Y  H  R  I  C  Y  V  N  R  Y  M  F  L  S  I  M  N  R  D  C  M  A  N  H  C  G  *  -
b      S  P  N  L  R  P  T  I  G  F  V  M  *  I  G  I  C  S  F  P  L  *  I  A  I  V  W  P  T  I  A  G  R -
c     I  P  P  I  C  D  L  P  *  D  L  L  C  K  *  V  Y  Y  V  P  F  H  Y  E  S  R  L  Y  G  Q  P  L  R  V  E -

                   TaiI
                MaeIII|
            HpyCH4IV| |
              NlaIV || |
              AvaII| || |
             Sau96I| || |
               NciI|| || |
              ScrFI|| || |
               SmaI|| || |
              HpaII|| || |
               NciI|| || |
              ScrFI|| || |
              AvaI| || || |
             BsaJI| || || |
            StyD4I| || || |                                                          KpnI
              XmaI| || || |                                                        NlaIV |
            StyD4I|| || || |                                                        RsaI |
         BsmI     |||| || || |                      AluI                           Csp6I| |
         BfaI|    |||| || || |  BseRI              CviJI              BccI       Acc65I|| |        MnlI
             ||   |||| || || | |BsmFI |TspDTI    MnlI  |         Hin4I  |          BanI|| |    AciI  |
             ||   |||| || || |   | |     |          |   |              |  |             ||| |      |  |
     aatgctagatgcccgggaccacgttactattgtttctttctcctccttcatattgagcttttctattttgccaataaatgatggtacctccccccgcca
4501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4600
     ttacgatctacgggccctggtgcaatgataacaaagaaagaggaggaagtataactcgaaaagataaaacggttatttactaccatggaggggggcggt

a     N  A  R  C  P  G  P  R  Y  Y  C  F  F  L  L  L  H  I  E  L  F  Y  F  C  Q  *  M  M  V  P  P  P  A  T -
b      M  L  D  A  R  D  H  V  T  I  V  S  F  S  S  F  I  L  S  F  S  I  F  A  N  K  *  W  Y  L  P  P  P  -
c       C  *  M  P  G  T  T  L  L  L  F  L  S  P  P  S  Y  *  A  F  L  F  L  P  I  N  D  G  T  S  P  R  H -
```

Figure 1-L

```
BsiEI
BstKTI
Hin4I                                                                      CviJI              HaeII
PvuI                                                                     BstNI |              HhaI|
TaqI                                      TaiI             BpmI          ScrFI |              AfeI||
DpnI|                                  BmgBI |          Eco57MI          PspGI | |          HinP1I|||
FauI||        MnlI     CviJI       HpyCH4IV| |       CviJI   |           StyD4I | |   Hpy188I   ||||
MboI|||
cgatcgaacgggaatggataagaggcttgtgggattgacgtgatagggtagggttggctatactgctggtggcgaactccaggctaataatctgaagcgc
4601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4700
gctagcttgcccttacctattctccgaacaccctaactgcactatcccatcccaaccgatatgacgaccaccgcttgaggtccgattattagacttcgcg

a   I  E  R  E  W  I  R  G  L  W  D  *  R  D  R  V  G  L  A  I  L  L  V  A  N  S  R  L  I  I  *  S  A -
b  R  S  N  G  N  G  *  E  A  C  G  I  D  V  I  G  *  G  W  L  Y  C  W  W  R  T  P  G  *  *  S  E  A  L -
c   D  R  T  G  M  D  K  R  L  V  G  L  T  *  *  G  R  V  G  Y  T  A  G  G  E  L  Q  A  N  N  L  K  R  -

BstKTI, DpnI, NlaIV, BamHI, BstYI, Eco57I, Eco57MI, AlwI, MnlI, MboI, Hpy188III, BssSI, CviJI, AlwI,
HpyF10VI, AciI, HpyF10VI, MwoI, MwoI, HinfI, TaqI, MlyI, PleI, HgaI, Hpy188I, MnlI, BsaHI, TaqI,
HinP1I, HhaI, HaeII, Hpy188III, TaqI, BceAI, BccI
ttggaagttgtagggaggatcctatggctcgtgaagcggttatcgccgaagtatcgactcaactatcagaggtagttggcgtcatcgagcgccatctcga
4701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4800
aaccttcaacatccctcctaggataccgagcacttcgccaatagcggcttcatagctgagttgatagtctccatcaaccgcagtagctcgcggtagagct

a   W  K  L  *  G  G  S  Y  G  S  *  S  G  Y  R  R  S  I  D  S  T  I  R  G  S  W  R  H  R  A  P  S  R  -
b    G  S  C  R  E  D  P  M  A  R  E  A  V  I  A  E  V  S  T  Q  L  S  E  V  V  G  V  I  E  R  H  L  E -
c  L  E  V  V  G  R  I  L  W  L  V  K  R  L  S  P  K  Y  R  L  N  Y  Q  R  *  L  A  S  S  S  A  I  S  N -

HpyCH4IV, Hpy99I, TaiI, EaeI, Cac8I, HaeIII, CviJI, Csp6I, RsaI, Csp6I, RsaI, CviJI, NlaIV, AciI, BccI,
TspRI, BtsI, BceAI, AciI, Fnu4HI, BstF5I, TauI, HaeIII, CviJI, FokI, CviJI, HpyCH4III, DraIII, TspRI,
Eco57MI, Eco57I, MaeIII, Tsp45I, MaeIII, HpyCH4III, BstEII, HpyCH4III, CviJI, HphI
accgacgttgctggccgtacatttgtacggctccgcagtggatggcggcctgaagccacacagtgatattgatttgctggttacggtgaccgtaaggctt
4801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4900
tggctgcaacgaccggcatgtaaacatgccgaggcgtcacctaccgccggacttcggtgtgtcactataactaaacgaccaatgccactggcattccgaa

a   T  D  V  A  G  R  T  F  V  R  L  R  S  G  W  R  P  E  A  T  Q  *  Y  *  F  A  G  Y  G  D  R  K  A  * -
b   P  T  L  L  A  V  H  L  Y  G  S  A  V  D  G  G  L  K  P  H  S  D  I  D  L  L  V  T  V  T  V  R  L  -
c    R  R  C  W  P  Y  I  C  T  A  P  Q  W  M  A  A  *  S  H  T  V  I  L  I  C  W  L  R  *  P  *  G  L -

Hin4I, BstUI, AciI, Fnu4HI, TauI, TspDTI, Cac8I, CviJI, AluI, BclI, MboI, BstKTI, DpnI, Hin4I, CviJI,
StyD4I, PspGI, BsaJI, ScrFI, BstNI, HinfI, TfiI, AciI, HinP1I, Eco57MI, BstUI, BpmI, HhaI, SfcI, HphI,
MaeIII, Tsp45I, MslI, ApaLI
gatgaaacaacgcggcgagctttgatcaacgaccttttggaaacttcggcttcccctggagagagcgagattctccgcgctgtagaagtcaccattgttg
4901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5000
ctactttgttgcgccgctcgaaactagttgctggaaaacctttgaagccgaaggggacctctctcgctctaagaggcgcgacatcttcagtggtaacaac

a    *  N  N  A  A  S  F  D  Q  R  P  F  G  N  F  G  F  P  W  R  E  R  D  S  P  R  C  R  S  H  H  C  C -
b   D  E  T  T  R  R  A  L  I  N  D  L  L  E  T  S  A  S  P  G  E  S  E  I  L  R  A  V  E  V  T  I  V  V-
c    M  K  Q  R  G  E  L  *  S  T  T  F  W  K  L  R  L  P  L  E  R  A  R  F  S  A  L  *  K  S  P  L  L  -
```

Figure 1-M

```
Tsp509I HpyCH4V HpyF10VI BstUI HhaI HinP1I BlpI DdeI AluI CviJI MwoI
Hpy99I TspGWI Bme1580I BsiHKAI Bsp1286I Hpy8I HpyCH4V BsaJI BtgI
HhaI HinP1I Fnu4HI TseI BbvI BsrDI BspMI HpyCH4V MboII CviJI Cac8I CviJI TaqI

     tgcacgacgacatcattccgtggcgttatccagctaagcgcgaactgcaatttggagaatggcagcgcaatgacattcttgcaggtatcttcgagccagc
5001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5100
     acgtgctgctgtagtaaggcaccgcaataggtcgattcgcgcttgacgttaaacctcttaccgtcgcgttactgtaagaacgtccatagaagctcggtcg

a     A  R  R  H  H  S  V  A  L  S  S  *  A  R  T  A  I  W  R  M  A  A  Q  *  H  S  C  R  Y  L  R  A  S  -
b      H  D  D  I  I  P  W  R  Y  P  A  K  R  E  L  Q  F  G  E  W  Q  R  N  D  I  L  A  G  I  F  E  P  A  -
c     C  T  T  T  S  F  R  G  V  I  Q  L  S  A  N  C  N  L  E  N  G  S  A  M  T  F  L  Q  V  S  S  S  Q  P -

BsiEI BstKTI PvuI TaqI DpnI MboI BstKTI DpnI MboI CviJI
BsaJI StyI AvaII Sau96I AciI TauI Fnu4HI AciI MspA1I MnlI
Hpy188III NlaIV HpaII BsaWI BstKTI DpnI EciI AlwI MboI

     cacgatcgacattgatctggctatcttgctgacaaaagcaagagaacatagcgttgccttggtaggtccagcggcggaggaactctttgatccggttcct
5101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5200
     gtgctagctgtaactagaccgatagaacgactgttttcgttctcttgtatcgcaacggaaccatccaggtcgccgcctccttgagaaactaggccaagga

a    H  D  R  H  *  S  G  Y  L  A  D  K  S  K  R  T  *  R  C  L  G  R  S  S  G  G  G  T  L  *  S  G  S  * -
b     T  I  D  I  D  L  A  I  L  L  T  K  A  R  E  H  S  V  A  L  V  G  P  A  A  E  E  L  F  D  P  V  P  -
c     R  S  T  L  I  W  L  S  C  *  Q  K  Q  E  N  I  A  L  P  W  *  V  Q  R  R  R  N  S  L  I  R  F  L  -

BstKTI MnlI DpnI BstYI MboI AlwI HaeII HhaI HinP1I MseI TspDTI
TauI AciI Fnu4HI BmrI Cac8I BsrI CviJI BsmFI HpyCH4IV TaiI AciI

     gaacaggatctatttgaggcgctaaatgaaaccttaacgctatggaactcgccgcccgactgggctggcgatgagcgaaatgtagtgcttacgttgtccc
5201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5300
     cttgtcctagataaactccgcgatttactttggaattgcgataccttgagcggcgggctgacccgaccgctactcgctttacatcacgaatgcaacaggg

a      T  G  S  I  *  G  A  K  *  N  L  N  A  M  E  L  A  A  R  L  G  W  R  *  A  K  C  S  A  Y  V  V  P  -
b     E  Q  D  L  F  E  A  L  N  E  T  L  T  L  W  N  S  P  P  D  W  A  G  D  E  R  N  V  V  L  T  L  S  R -
c      N  R  I  Y  L  R  R  *  M  K  P  *  R  Y  G  T  R  R  P  T  G  L  A  M  S  E  M  *  C  L  R  C  P  -

HpaII BsrFI MaeIII HhaI HinP1I RsaI Csp6I FauI BslI BbvI HhaI BstUI HinP1I
Fnu4HI TseI BstF5I BsrI HpyF10VI BglI MwoI FokI
BsrI CviJI HaeIII Sau96I Cac8I NaeI HpaII BmrI BsrFI NgoMIV Cac8I HaeII HhaI HinP1I BsrDI BmrI CviJI

     gcatttggtacagcgcagtaaccggcaaaatcgcgccgaaggatgtcgctgccgactgggcaatggagcgcctgccggcccagtatcagcccgtcatact
5301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5400
     cgtaaaccatgtcgcgtcattggccgttttagcgcggcttcctacagcgacggctgacccgttacctcgcggacggccgggtcatagtcgggcagtatga

a     H  L  V  Q  R  S  N  R  Q  N  R  A  E  G  C  R  C  R  L  G  N  G  A  P  A  G  P  V  S  A  R  H  T  -
b      I  W  Y  S  A  V  T  G  K  I  A  P  K  D  V  A  A  D  W  A  M  E  R  L  P  A  Q  Y  Q  P  V  I  L -
c     A  F  G  T  A  Q  *  P  A  K  S  R  R  R  M  S  L  P  T  G  Q  W  S  A  C  R  P  S  I  S  P  S  Y  L-
```

Figure 1-N

```
BstKTI, DpnI, MboI, MnlI, HhaI, BstUI, Cac8I, HhaI, HinP1I, HpyF10VI, MwoI, BssHII, BstUI, HinP1I,
CviJI, HaeIII, MboII, MmeI, MboII, BstKTI, DpnI, MboI, CviJI, HpyF10VI, BfaI, AluI, CviJI, MwoI,
ApoI, Tsp509I, TaiI, BsaAI, DraIII, HpyCH4IV, MboII, Hpy8I, BcgI, DdeI, BstKTI, DpnI, MboI

     tgaagctagacaggcttatcttggacaagaagaagatcgcttggcctcgcgcgcagatcagttggaagaatttgtccactacgtgaaaggcgagatcact
5401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5500
     acttcgatctgtccgaatagaacctgttcttcttctagcgaaccggagcgcgcgtctagtcaaccttcttaaacaggtgatgcactttccgctctagtga

a     * S * T G L S W T R R R S L G L A R R S V G R I C P L R E R R D H * -
b     E A R Q A Y L G Q E E D R L A S R A D Q L E E F V H Y V K G E I T -
c     K L D R L I L D K K K I A W P R A Q I S W K N L S T T * K A R S L -
```

```
Hpy188III, TaqI, AvaI, SmlI, XhoI, BsmAI, DdeI, Tsp509I, BcgI, EarI, Hin4I, MboII, BstKTI, DpnI, MboI, MaeIII

     aaggtagttggcaaataactcgagactaagtggataaaattagatagaaaaaggtctaaataaaaaagaagagaaatagaaagatcaaaaatcagttac
5501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5600
     ttccatcaaccgtttattgagctctgattcacctattttaatctatctttttccagatttattttttcttctctttatctttctagtttttagtcaatg

a     G S W Q I T R D * V D K I R * K K G L N K K E E K * K D Q K S V T -
b     K V V G K * L E T K W I K L D R K K V * I K K K R N R K I K N Q L R -
c     R * L A N N S R L S G * N * I E K R S K * K R R E I E R S K I S Y -
```

```
Tsp509I, Hin4I, HpyCH4V, MboII, MboII, Tsp509I, MseI, Tsp509I, HpyCH4V, Hpy188I, BsmAI, BsmBI,
CviJI, DraI, MseI, DpnI, BsmAI, MboI

     gaaaatgcagtaattcttcttttttcttctaattgattgcaattaaactcgtctcaatctgaaaaaagattgagccgagtttaaatagattttgatacga
5601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5700
     cttttacgtcattaagaagaaaaaagaagattaactaacgttaatttgagcagagttagacttttttctaactcggctcaaatttatctaaaactatgct

a     K M Q * F F F F L L I D C N * T R L N L K K D * A E F K * I L I R -
b     K C S N S S F F F * L I A I K L V S I * K K I E P S L N R F * Y D -
c     E N A V I L L F S S N * L Q L N S S Q S E K R L S R V * I D F D T I -
```

```
NlaIII, Hpy188III, BspHI, BstKTI, FatI, HinfI, TfiI, Hpy188III, MnlI, MboII, BfaI, Hpy188III, XbaI,
TaqI, AlwI, AluI, CviJI, Bsu36I, DdeI, BstKTI, DpnI, MboI

     tcatgagacttgacaaatcgggattcctctattctatatatttagaagatataaaggtataatacaataaataaatacaatctagaagctcgatgatcct
5701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5800
     agtactctgaactgtttagccctaaggagataagatatataaatcttctatatttccatattatgttatttatttatgttagatcttcgagctactagga

a     S * D L T N R D S S I L Y I * K I * R Y N T I N K Y N L E A R * S L -
b     H E T * Q I G I P L F Y I F R R Y K G I I Q * I N T I * K L D D P -
c     M R L D K S G F L Y S I Y L E D I K V * Y N K * I Q S R S S M I L -
```

Figure 1-O

```
                                                              BstF5I
                                          HinfI               BsmFI|
                         SfcI             TfiI  CviJI   FokI   MnlI|   BccI TspGWI
     taggattggtatattcttttctatcctgtagtttgtagtttccctgaatcaagccaagtatcacacctctttctacccatcctgtatattgtcccctttg
5801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5900
     atcctaaccatataagaaaagataggacatcaaacatcaaagggacttagttcggttcatagtgtggagaaagatgggtaggacatataacaggggaaac

a     G  L  V  Y  S  F  L  S  C  S  L  *  F  P  *  I  K  P  S  I  T  P  L  S  T  H  P  V  Y  C  P  L  C  -
b    *  D  W  Y  I  L  F  Y  P  V  V  C  S  F  P  E  S  S  Q  V  S  H  L  F  L  P  I  L  Y  I  V  P  F  V -
c     R  I  G  I  F  F  S  I  L  *  F  V  V  S  L  N  Q  A  K  Y  H  T  S  F  Y  P  S  C  I  L  S  P  L   -

                                          ApoI
                    Tsp509I             Tsp509I                                                 Tsp509I
                      MseI|               MseI |
     ttccgtgttgaaatagaaccttaatttattacttatttttttattaaatttagatttgttagtgattagatattagtattagacgagatttacgaaac
5901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6000
     aaggcacaactttatcttggaattaaataatgaataaaaaaataatttaaatctaaacaatcactaatctataatcataatctgctctaaaatgctttg

a     S  V  L  K  *  N  L  N  L  L  L  I  F  L  L  N  F  R  F  V  S  D  *  I  L  V  L  D  E  I  L  R  N  -
b      P  C  *  N  R  T  L  I  Y  Y  L  F  F  Y  *  I  L  D  L  L  V  I  R  Y  *  Y  *  T  R  F  Y  E  T -
c    F  R  V  E  I  E  P  *  F  I  T  Y  F  F  I  K  F  *  I  C  *  *  L  D  I  S  I  R  R  D  F  T  K  Q -

                                              ApoI                          TauI
                                            Tsp509I                     AciI |
                  MnlI                      Hin4I |                   Fnu4HI |     Tsp509I
                Hin4I |            SfaNI   TaqI  | |                  CviJI| |     MseI  |   FalI
     aattatttttttatttctttataggagaggacaaatctctttttcgatgcgaatttgacacgacataggagaagccgccctttattaaaaattatatta
6001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6100
     ttaataaaaaaataaagaaatatcctctcctgtttagagaaaaagctacgcttaaactgtgctgtatcctcttcggcgggaaataatttttaatataat

a    N  Y  F  F  I  S  L  *  E  R  T  N  L  F  F  R  C  E  F. D  T  T  *  E  K  P  P  F  I  K  N  Y  I  I -
b     I  I  F  L  F  L  Y  R  R  G  Q  I  S  F  F  D  A  N  L  T  R  H  R  R  S  R  P  L  L  K  I  I  L  -
c      L  F  F  Y  F  F  I  G  E  D  K  S  L  F  S  M  R  I  *  H  D  I  G  E  A  A  L  Y  *  K  L  Y  Y -

                                                              Hpy188I
     DraI                           AseI                 HinfI  |
    MseI|              NlaIV        MseI        MmeI     TfiI   |                                MseI
     ttttaaataatataaaggggggttccaacatattaatatatagtgaagtgttcccccagattcagaactttttttcaatactcacaatccttattagttaa
6101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6200
     aaaatttatcatatttccccccaaggttgtataattatatatcacttcacaagggggtctaagtcttgaaaaaaagttatgagtgttaggaataatcaatt

a      L  N  N  I  K  G  V  P  T  Y  *  Y  I  V  K  C  S  P  R  F  R  T  F  F  Q  Y  S  Q  S  L  L  V  N  -
b    F  *  I  I  *  R  G  F  Q  H  I  N  I  *  *  S  V  P  P  D  S  E  L  F  F  N  T  H  N  P  Y  *  L  I -
c     F  K  *  Y  K  G  G  S  N  I  L  I  Y  S  E  V  F  P  Q  I  Q  N  F  F  S  I  L  T  I  L  I  S  *   -

                                                                 TspDTI                        Hin4I
       BfaI                 DdeI   Hpy188I               Tsp509I  Hin4I |            TspDTI     FatI |
     taatcctagtgattggatttctatgcttagtctgataggaaataagatattcaaataaataattttatagcgaatgactattcatctattgtattttcat
6201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6300
     attaggatcactaacctaaagatacgaatcagactatcctttattctataagtttatttattaaaatatcgcttactgataagtagataacataaaagta

a     N  P  S  D  W  I  S  M  L  S  L  I  G  N  K  I  F  K  *  I  I  L  *  R  M  T  I  H  L  L  Y  F  H  -
b      I  L  V  I  G  F  L  C  L  V  *  *  E  I  R  Y  S  N  K  *  F  Y  S  E  *  L  F  I  Y  C  I  F  M -
c    *  S  *  *  L  D  F  Y  A  *  S  D  R  K  *  D  I  Q  I  N  N  F  I  A  N  D  Y  S  S  I  V  F  S  C -

                                                                BstBI
                                                                 TaqI
                                                             AarI |
 HpyCH4V                              Tsp509I                BspMI|
  NlaIII                 BccI         MseI|TaqI      MseI    XmnI||              CviJI
     gcaaataggggggcaagaaaactctatggaaagatggtggtttaattcgatgttgtttaagaaggagttcgaacgcaggtgtgggctaaataaatcaatgg
6301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6400
     cgtttatccccccgttcttttgagatacctttctaccaccaaattaagctacaacaaattcttcctcaagcttgcgtccacacccgatttatttagttacc

a    A  N  R  G  Q  E  N  S  M  E  R  W  W  F  N  S  M  L  F  K  K  E  F  E  R  R  C  G  L  N  K  S  M  G -
b     Q  I  G  G  K  K  T  L  W  K  D  G  G  L  I  R  C  C  L  R  R  S  S  N  A  G  V  G  *  I  N  Q  W  -
c      K  *  G  A  R  K  L  Y  G  K  M  V  V  *  F  D  V  V  *  E  G  V  R  T  Q  V  W  A  K  *  I  N  G -
```

Figure 1-P

```
                                                                                  Tsp509I
                                   BsaBI      TspDTI                    Hpy188III    |
                                 BstKTI |    MmeI  |                     MaeIII      |
                                   DpnI| |TspDTI |  |                    Tsp45I      |          HpyCH4V
      AvaII                       BstYI || |MboII| |  |             MnlI HinfI   |      |          BsaXI |
      Sau96I               AlwI MboI || |TaqI|| |  |        BsaXI |   TfiI   |      | BfaI  | |
        |                   |    | || | ||| |  |          | |     |   |      |  |   | |
     gcagtcttggtcctattgaaaataccaatgaagatccaaatcgaaaagtgaaaaacattcatagttggaggaatcgtgacaattctagttgcagtaatgt
6401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6500
     cgtcagaaccaggataacttttatggttacttctaggtttagcttttcacttttttgtaagtatcaacctccttagcactgttaagatcaacgtcattaca

a      S  L  G  P  I  E  N  T  N  E  D  P  N  R  K  V  K  N  I  H  S  W  R  N  R  D  N  S  S  C  S  N  V -
b     A  V  L  V  L  L  K  I  P  M  K  I  Q  I  E  K  *  K  T  F  I  V  G  G  I  V  T  I  L  V  A  V  M  L -
c      Q  S  W  S  Y  *  K  Y  Q  *  R  S  K  S  K  S  E  K  H  S  *  L  E  E  S  *  Q  F  *  L  Q  *  C -

                          ApoI
                        Tsp509I
                       Hpy188I|
             MseI TspDTI    ||    Hpy188I                   BsmAI  HpyCH4III            BccI
               |     |      ||       |                        |        |                  |
     tgattatttattcggcgttaaagacattcggaatttcatctctgatgacactttttagttagtgataggaatggagacagttattccatctattttgat
6501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6600
     actaataaataagccgcaatttctgtaagccttaaagtagagactactgtgaaaaatcaatcactatccttacctctgtcaataaggtagataaaacta

a     D  Y  L  F  G  V  K  D  I  R  N  F  I  S  D  D  T  F  L  V  S  D  R  N  G  D  S  Y  S  I  Y  F  D  -
b      I  I  Y  S  A  L  K  T  F  G  I  S  S  L  M  T  L  F  *  L  V  I  G  M  E  T  V  I  P  S  I  L  I -
c     *  L  F  I  R  R  *  R  H  S  E  F  H  L  *  *  H  F  F  S  *  *  *  E  W  R  Q  L  F  H  L  F  *  Y -

                          BspCNI
                         BseMII|
                        BstKTI||
                         DpnI|||   Hpy188I    BfaI                          Tsp509I
                       MboI ||||     DdeI|Hpy8I |                     TaqI    TaqI | Hpy188I
                         |  ||||      ||    |  |                        |       | |     |
     attgaaaatcatattttgagattgacaacgatcattcttttctgagtgaactagaaagttctttttatagttatcgaaactcgaattatcggaataatg
6601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6700
     taacttttagtataaaactctaactgttgctagtaagaaaagactcacttgatctttcaagaaaaatatcaatagctttgagcttaatagccttattac

a     I  E  N  H  I  F  E  I  D  N  D  H  S  F  L  S  E  L  E  S  S  F  Y  S  Y  R  N  S  N  Y  R  N  N  G -
b      L  K  I  I  F  L  R  L  T  T  I  I  L  F  *  V  N  *  K  V  L  F  I  V  I  E  T  R  I  I  G  I  M  -
c     *  K  S  Y  F  *  D  *  Q  R  S  F  F  S  E  *  T  R  K  F  F  L  *  L  S  K  L  E  L  S  E  *  W -

                                NlaIII
                                  NspI
                                MmeI |
           BstKTI                                                  HpyCH4V
             DpnI|             AflIII ||                   AseI  Eco57I|     BsaBI
           BstYI ||  Tsp509I    FatI ||                   MseI Eco57MI|MboII  |
      AlwI MboI ||  MboII |     PciI ||                     |      ||   |     |
        |     | ||    | |       | | |                       |      ||   |     |
     gatttaggggcgaagatccctactataattcttacatgtatgatactcaatatagttggaataatcacattaatagttgcattgatagttatcttcagtc
6701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6800
     ctaaatccccgcttctagggatgatattaagaatgtacatactatgagttatatcaaccttattagtgtaattatcaacgtaactatcaatagaagtcag

a       F  R  G  E  D  P  Y  Y  N  S  Y  M  Y  D  T  Q  Y  S  W  N  N  H  I  N  S  C  I  D  S  Y  L  Q  S -
b      D  L  G  A  K  I  P  T  I  I  L  T  C  M  I  L  N  I  V  G  I  I  T  L  I  V  A  L  I  V  I  F  S  L -
c      I  *  G  R  R  S  L  L  *  F  L  H  V  *  Y  S  I  *  L  E  *  S  H  *  *  L  H  *  *  L  S  S  V -

                                               BceAI
                                            HpyCH4III
                                               MaeIII
                                   HpyCH4III     |            CviJI
                                      MaeIII     |           HaeIII                     HphI
 BsmAI                 PsiI     Tsp509I Tsp45I   |   HphI Sau96I |                 TaqI  |      MnlI
   |                    |             |    |     |     |      | |                   | |       |
     tcaaatctgtatagatacttccattataagtggtagtgagaattacggtgacagttacatttatagggccgtttgtggtggtgaaagtcgaaatagtagt
6801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6900
     agtttagacatatctatgaaggtaatattcaccatcactcttaatgccactgtcaatgtaaatatcccggcaaacaccaccactttcagctttatcatca

a      Q  I  C  I  D  T  S  I  I  S  G  S  E  N  Y  G  D  S  Y  I  Y  R  A  V  C  G  G  E  S  R  N  S  S  -
b       K  S  V  *  I  L  P  L  *  V  V  V  R  I  T  V  T  V  T  F  I  G  P  F  V  V  V  K  V  E  I  V  V -
c      S  N  L  Y  R  Y  F  H  Y  K  W  *  *  E  L  R  *  Q  L  H  L  *  G  R  L  W  W  *  K  S  K  *  *  * -
```

Figure 1-Q

```
                                                                                  MaeIII
                                                                                TaqI |
                                                                               AvaI| |
                                                                        Hpy188III| |
                                                                              SmlI| |
                                                                              XhoI| |
                                                                           BstKTI|| |
                                                                             DpnI||| |
              Hpy8I                               MseI                       MboI |||| |
          BsrI  |              HpyF10VI           BtsI|                       MnlI |||| |
       NlaIV |AccI|                MwoI|         TspRI|                        |||| |
          |  |  ||                   ||              ||                        |||||
     gaaaacgagggttccagtagacgaactcgcacgaagggcagtgatttaactataagagaaagttctaatgatctcgaggtaactcaaaaatacaggcatt
6901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7000
     cttttgctcccaaggtcatctgcttgagcgtgcttcccgtcactaaattgatattctctttcaagattactagagctccattgagtttttatgtccgtaa

a     E  N  E  G  S  S  R  R  T  R  T  K  G  S  D  L  T  I  R  E  S  S  N  D  L  E  V  T  Q  K  Y  R  H  L -
b      K  T  R  V  P  V  D  E  L  A  R  R  A  V  I  *  L  *  E  K  V  L  M  I  S  R  *  L  K  N  T  G  I  -
c       K  R  G  F  Q  *  T  N  S  H  E  G  Q  *  F  N  Y  K  R  K  F  *  *  S  R  G  N  S  K  I  Q  A  F -

                                              ApoI
                              Tsp509I  Tsp509I                              TspDTI
            Tsp509I              MseI |  PsiI |                      SspI  Hpy8I |        EcoRV
               |                    |  |   |   |                        |     |  |           |
     tgtgggttcaatgcgaaaattgttatggattaaattataagaaattttttgaaatcaaaaatgaatatttgtgaacaatgtggatatcatttgaaaatgag
7001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7100
     acacccaagttacgcttttaacaatacctaatttaatattctttaaaaaacttagtttttacttataaacacttgttacacctatagtaaacttttactc

a      W  V  Q  C  E  N  C  Y  G  L  N  Y  K  K  F  L  K  S  K  M  N  I  C  E  Q  C  G  Y  H  L  K  M  S -
b     C  G  F  N  A  K  I  V  M  D  *  I  I  R  N  F  *  N  Q  K  *  I  F  V  N  N  V  D  I  I  *  K  *  V-
c      V  G  S  M  R  K  L  L  W  I  K  L  *  E  I  F  E  I  K  N  E  Y  L  *  T  M  W  I  S  F  E  N  E  -

          ApoI
         EcoRI                                                    Tsp509I
       Tsp509I         NlaIII  AluI                          BsrBI  |                    HpaII
  Hpy188I  |          FatI  |  CviJI          Tsp509I        AciI | |                    CviJI|
       |   |            |   |    |                  |            | |  |                      ||
     tagttcagatagaattcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagccggaagcataaagt
7101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7200
     atcaagtctatcttaagcattagtaccagtatcgacaaaggacacactttaacaataggcgagtgttaaggtgtgttgtatgctcggccttcgtatttca

a      S  S  D  R  I  R  N  H  G  H  S  C  F  L  C  E  I  V  I  R  S  Q  F  H  T  T  Y  E  P  E  A  *  S -
b       V  Q  I  E  F  V  I  M  V  I  A  V  S  C  V  K  L  L  S  A  H  N  S  T  Q  H  T  S  R  K  H  K  V -
c     *  F  R  *  N  S  *  S  W  S  *  L  F  P  V  *  N  C  Y  P  L  T  I  P  H  N  I  R  A  G  S  I  K  C-

                                                                                                 AseI
                                                                                                 MseI
                NlaIV                                                 AciI                 HpyCH4V |
             HpyF10VI|                                                Cac8I               Fnu4HI | |
           BstNI    ||                                                TspRI                  AluI| | |
           ScrFI    ||                                          HpyF10VI |                  CviJI| | |
          BsaJI|    ||                                              MwoI||                 MspA1I| | |
          CviJI||   ||                                             BtsI |||                 PvuII| | |
          PspGI|||BanI||              AluI        Tsp509I       HhaI | |||    Hpy188III      TseI| | |
         StyD4I|||MwoI||              CviJI          AseI|   HinP1I |  | |||    BsrI |          Cac8I || | |
              |||    |||                |          MseI||        |  |  | |||    FauI |   BbvI    | || | |
     gtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcatta
7201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7300
     catttcggaccccacggattactcactcgattgagtgtaattaacgcaacgcgagtgacgggcgaaaggtcagccctttggacagcacggtcgacgtaat

a     V  K  P  G  V  P  N  E  *  A  N  S  H  *  L  R  C  A  H  C  P  L  S  S  R  E  T  C  R  A  S  C  I  N-
b      *  S  L  G  C  L  M  S  E  L  T  H  I  N  C  V  A  L  T  A  R  F  P  V  G  K  P  V  V  P  A  A  L  -
c       K  A  W  G  A  *  *  V  S  *  L  T  L  I  A  L  R  S  L  P  A  F  Q  S  G  N  L  S  C  Q  L  H  *  -
```

Figure 1-R

```
                AciI
               BstUI
                HhaI
                MnlI
               BslI
             TspDTI
              BstUI
             HinP1I                                                                   BbvI
            BsaXI                                                                     HhaI
           FauI                         HaeII                                     HinP1I           Cac8I
         CviJI                          HhaI                                     Fnu4HI             TauI
        HaeIII                       BsaXI                                        TseI          Fnu4HI
        EaeI                        HinP1I                                      TaqII            AciI
                                  HpyF10VI        EarI        MlyI     TspRI                   Fnu4HI
   HinfI                             MwoI         SapI        PleI  HinfI                       CviJI
   TfiI                  AciI       MboII         AciI     BbvI  MnlI                 BsiEI    TseI
     atgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcga
7301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7400
     tacttagccggttgcgcgcccctctccgccaaacgcataacccgcgagaaggcgaaggagcgagtgactgagcgacgcgagccagcaagccgacgccgct

a      E  S  A  N  A  R  G  E  A  V  C  V  L  G  A  L  P  L  P  R  S  L  T  R  C  A  R  S  F  G  C  G  E  -
b     M  N  R  P  T  R  G  E  R  R  F  A  Y  W  A  L  F  R  F  L  A  H  *  L  A  A  L  G  R  S  A  A  A  S -
c      *  I  G  Q  R  A  G  R  G  G  L  R  I  G  R  S  S  A  S  S  L  T  D  S  L  R  S  V  V  R  L  R  R  -

                                                                                                 NlaIV
                                                                                              BstNI
                                                                                              ScrFI
          AluI                                                                 NlaIII         CviJI
         CviJI                                                                  NspI         HaeIII
    HpyF10VI                                                                 AflIII       Cac8I    PspGI
  AciI                                                      HinfI             FatI       CviJI    StyD4I
 BsrBI   MwoI               AciI  HpyCH4III                  TfiI              PciI       HaeIII   BslI
     gcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccagga
7401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7500
     cgccatagtcgagtgagtttccgccattatgccaataggtgtcttagtcccctattgcgtcctttcttgtacactcgttttccggtcgttttccggtcct

a      R  Y  Q  L  T  Q  R  R  *  Y  G  Y  P  Q  N  Q  G  I  T  Q  E  R  T  C  E  Q  K  A  S  K  R  P  G  -
b       G  I  S  S  L  K  G  G  N  T  V  I  H  R  I  R  G  *  R  R  K  E  H  V  S  K  R  P  A  K  G  Q  E -
c     A  V  S  A  H  S  K  A  V  I  R  L  S  T  E  S  G  D  N  A  G  K  N  M  *  A  K  G  Q  Q  K  A  R  N -

                    Cac8I
                 HpyF10VI                                                           Hpy188I
              BstUI                                                                  HgaI
               TauI                                                               DrdI
             AciI                                                                 MnlI
           Fnu4HI                                                                SmlI
           CviJI                          AciI                                 Hpy99I
          HaeIII                         NlaIV                               TaqI
HpyCH4III BslI          MwoI       EciI  CviJI                BpuEI  SfaNI
     accgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgac
7501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7600
     tggcatttttccggcgcaacgaccgcaaaaaggtatccgaggcggggggactgctcgtagtgtttttagctgcgagttcagtctccaccgctttgggctg

a      T  V  K  R  P  R  C  W  R  F  S  I  G  S  A  P  L  T  S  I  T  K  I  D  A  Q  V  R  G  G  E  T  R  Q -
b       P  *  K  G  R  V  A  G  V  F  P  *  A  P  P  P  *  R  A  S  Q  K  S  T  L  K  S  E  V  A  K  P  D  -
c       R  K  K  A  A  L  L  A  F  F  H  R  L  R  P  P  D  E  H  H  K  N  R  R  S  S  Q  R  W  R  N  P  T  -

                                                                           EciI
                                                                          HpaII
                                                   MnlI                  BsaWI
                               BstNI              HhaI                    BslI
            BstNI              ScrFI            HinP1I                 BciVI
            ScrFI              BsaJI       BssSI                        TauI
           PspGI               PspGI  AluI                             AciI                AciI
          StyD4I              StyD4I  CviJI               Hpy188I   Fnu4HI         MmeI
     aggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctccct
7601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7700
     tcctgatatttctatggtccgcaaaggggggaccttcgagggagcacgcgagaggacaaggctgggacggcgaatggcctatggacaggcggaaagaggga

a        D  Y  K  D  T  R  R  F  P  L  E  A  P  S  C  A  L  L  F  R  P  C  R  L  P  D  T  C  P  P  F  S  L  -
b      R  T  I  K  I  P  G  V  S  P  W  K  L  P  R  A  L  S  C  S  D  P  A  A  Y  R  I  P  V  R  L  S  P  F -
c       G  L  *  R  Y  Q  A  F  P  P  G  S  S  L  V  R  S  P  V  P  T  L  P  L  T  G  Y  L  S  A  F  L  P  -
```

Figure 1-S

```
                                                                                            Bme1580I
                                                                                             BsiHKAI
                                                                                    CviJI Bsp1286I
                 HaeII                                                          AluI  |  Hpy8I  |
                  HhaI|         AluI                              BseMII        BseYI  |HpyCH4V |
Hpy188III   HinP1I ||        CviJI    SfcI      DdeI            BspCNI|         CviJI  |ApaLI | |        BbvI
     |           | ||          |        |          |                 ||             |   |  | | |          |
     tcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttc
7701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7800
     agcccttcgcaccgcgaaagagtatcgagtgcgacatccatagagtcaagccacatccagcaagcgaggttcgacccgacacacgtgcttggggggcaag

a      R  E  A  W  R  F  L  I  A  H  A  V  G  I  S  V  R  C  R  S  F  A  P  S  W  A  V  C  T  N  P  P  F  -
b       G  K  R  G  A  F  S  *  L  T  L  *  V  S  Q  F  G  V  G  R  S  L  Q  A  G  L  C  A  R  T  P  R  S  -
c     S  G  S  V  A  L  S  H  S  S  R  C  R  Y  L  S  S  V  *  V  V  R  S  K  L  G  C  V  H  E  P  P  V  Q  -

                                                                                                   BbvI
                                                                                                  BsrI|
                                                                                                TspRI|
                                                                                               BbvI ||
                                                                                             MaeIII ||
                                                                                            AlwNI | ||
                   HhaI                                                                  CviJI |  | ||
                 HinP1I |                                  HpaII                       Fnu4HI | | | ||
                Fnu4HI| |                                   MlyI                        TseI| | | | ||
               MspA1I|| |                                   NciI                       BsrI|| | | | ||
                  TseI|| |       MaeIII            HinfI    ScrFI                     Fnu4HI|| | | | ||
                 AciI ||| |      HpaII |     Hpy188III |     PleI|                     TspRI|| | | | ||
      CviJI BsiEI |||| |       BsaWI|     SmlI |    StyD4I||   BpuEI       MmeI      TseI||| | | | ||       MnlI
        |       | |||| |           || |          |   |        |||      |            |         ||||| | | | ||         |
     agcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattag
7801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7900
     tcgggctggcgacgcggaataggccattgatagcagaactcaggttgggccattctgtgctgaatagcggtgaccgtcgtcggtgaccattgtcctaatc

a     S  P  T  A  A  P  Y  P  V  T  I  V  L  S  P  T  R  *  D  T  T  Y  R  H  W  Q  Q  P  L  V  T  G  L  A  -
b      A  R  P  L  R  L  I  R  *  L  S  S  *  V  Q  P  G  K  T  R  L  I  A  T  G  S  S  H  W  *  Q  D  *  -
c       P  D  R  C  A  L  S  G  N  Y  R  L  E  S  N  P  V  R  H  D  L  S  P  L  A  A  A  T  G  N  R  I  S  -

                                                  CviJI        CviJI           HpyCH4III          HhaI
                 AciI   SfcI  Hpy188III         HaeIII     BslI  |   BfaI   BceAI |        HinP1I |      CviJI
                   |      |          .|              |         |   |     |       | |             | |          |
     cagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaag
7901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8000
     gtctcgctccatacatccgccacgatgtctcaagaacttcaccaccggattgatgccgatgtgatcttcctgtcataaaccatagacgcgagacgacttc

a      E  R  G  M  *  A  V  L  Q  S  S  *  S  G  G  L  T  T  A  T  L  E  G  Q  Y  L  V  S  A  L  C  *  S  -
b     Q  S  E  V  C  R  R  C  Y  R  V  L  E  V  V  A  *  L  R  L  H  *  K  D  S  I  W  Y  L  R  S  A  E  A  -
c      R  A  R  Y  V  G  G  A  T  E  F  L  K  W  W  P  N  Y  G  Y  T  R  R  T  V  F  G  I  C  A  L  L  K  -

                                       HpaII
                                     BstKTI|
                                      DpnI||                                                         BpuEI
                                      MboI|||                                                         HhaI
                               Hpy188III | |||                                                      BbvI |
                   Eco57I            AluI | | |||                                       Fnu4HI       BstUI |
  MaeIII         Eco57MI            AlwI | | |||       MspA1I                           TseI|      HinP1I |
   BsrI |      Hpy188I  |          CviJI | | |||        AciI |       AciI          Cac8I ||    HpyF10VI | |
      | |            |  |              | | | |||           | |          |        HpyCH4V | ||      MwoI| | |
     ccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgca
8001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8100
     ggtcaatggaagcctttttctcaaccatcgagaactaggccgtttgtttggtggcgaccatcgccaccaaaaaaacaaacgttcgtcgtctaatgcgcgt

a      Q  L  P  S  E  K  E  L  V  A  L  D  P  A  N  K  P  P  L  V  A  V  V  F  L  F  A  S  S  R  L  R  A  -
b       S  Y  L  R  K  K  S  W  *  L  L  I  R  Q  T  N  H  R  W  *  R  W  F  F  C  L  Q  A  A  D  Y  A  Q  -
c     P  V  T  F  G  K  R  V  G  S  S  *  S  G  K  Q  T  T  A  G  S  G  G  F  F  V  C  K  Q  Q  I  T  R  R  -
```

Figure 1-T

```
BstKTI
DpnI
BstYI
MboI
AlwI
AlwI
Hpy188III
SmlI
BstKTI                                                                                    MseI
DpnI                                                                                      TaiI
BstYI                                                                             HpyCH4IV
MboI        BstKTI                                                  BseMII                      NlaIII
            DpnI                                   DdeI             BspCNI          Hpy188III
            MboII                                  PpiI     TspRI                   BspHI
            MboI                       Hpy188I              HgaI                    PpiI FatI
     gaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatc
8101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8200
     ctttttttcctagagttcttctaggaaactagaaaagatgccccagactgcgagtcaccttgcttttgagtgcaattccctaaaaccagtactctaatag

a    E  K  K  D  L  K  K  I  L  *  S  F  L  R  G  L  T  L  S  G  T  K  T  H  V  K  G  F  W  S  *  D  Y  Q -
b     K  K  R  I  S  R  R  S  F  D  L  F  Y  G  V  *  R  S  V  E  R  K  L  T  L  R  D  F  G  H  E  I  I  -
c      K  K  G  S  Q  E  D  P  L  I  F  S  T  G  S  D  A  Q  W  N  E  N  S  R  *  G  I  L  V  M  R  L  S -

            BstKTI
      BstKTI     DpnI
        DpnI   BstYI
      BstYI     MboI        MseI
      MboI     BfaI  Tsp509I                                                     HpyCH4III
     HphI     AlwI     DraI               DraI                                    MaeIII               MnlI
MboII     AlwI         MseI               MseI TspDTI          Hpy8I   Hpy188I                         MseI
     aaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgctta
8201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8300
     tttttcctagaagtggatctaggaaaatttaatttttacttcaaaatttagttagatttcatatatactcatttgaaccagactgtcaatggttacgaat

a     K  G  S  S  P  R  S  F  *  I  K  N  E  V  L  N  Q  S  K  V  Y  M  S  K  L  G  L  T  V  T  N  A  *  -
b    K  K  D  L  H  L  D  P  F  K  L  K  M  K  F  *  I  N  L  K  Y  I  *  V  N  L  V  *  Q  L  P  M  L  N -
c     K  R  I  F  T  *  I  L  L  N  *  K  *  S  F  K  S  I  *  S  I  Y  E  *  T  W  S  D  S  Y  Q  C  L  -

                      BseMII
                      BspCNI
                     TspDTI
                     FokI
       NlaIV       BstKTI                 MlyI
       BanI        DpnI                   PleI              Hpy99I           MnlI
       TspRI  DdeI MboI       BstF5I  Hin4I     HinfI AhdI                 Hin4I         CviJI   BbvI
     atcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccat
8301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8400
     tagtcactccgtggatagagtcgctagacagataaagcaagtaggtatcaacggactgaggggcagcacatctattgatgctatgccctcccgaatggta

a     S  V  R  H  L  S  Q  R  S  V  Y  F  V  H  P  *  L  P  D  S  P  S  C  R  *  L  R  Y  G  R  A  Y  H  -
b      Q  *  G  T  Y  L  S  D  L  S  I  S  F  I  H  S  C  L  T  P  R  R  V  D  N  Y  D  T  G  G  L  T  I -
c    I  S  E  A  P  I  S  A  I  C  L  F  R  S  S  I  V  A  *  L  P  V  V  *  I  T  T  I  R  E  G  L  P  S-

                                                                                      HhaI
                                                                                    HinP1I
                         BpmI                                                     CviJI
       BsrI            Eco57MI                                                   HaeIII
      BccI           BstUI                                                    HpyF10VI
     NlaIV           AciI              Hpy188III
    CviJI   HpyCH4V  BsaI                NlaIV                             HpaII
   HaeIII    TspRI   BsmAI              CviJI                            CviJI Sau96I
   Sau96I  Fnu4HI    BsrDI            HpaII                            Cac8I   BglI                AvaII
   BmrI     TseI     Hin4I      HphI  BsrFI      Hin4I               CviJI     MwoI                Sau96I
     ctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtgg
8401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8500
     gaccggggtcacgacgttactatggcgctctgggtgcgagtggccgaggtctaaatagtcgttatttggtcggtcggccttcccggctcgcgtcttcacc

a    L  A  P  V  L  Q  *  Y  R  E  T  H  A  H  R  L  Q  I  Y  Q  Q  *  T  S  Q  P  E  G  P  S  A  E  V  V -
b     W  P  Q  C  C  N  D  T  A  R  P  T  L  T  G  S  R  F  I  S  N  K  P  A  S  R  K  G  R  A  Q  K  W  -
c      G  P  S  A  A  M  I  P  R  D  P  R  S  P  A  P  D  L  S  A  I  N  Q  P  A  G  R  A  E  R  R  S  G -
```

Figure 1-U

```
                                              TaiI
                                              AclI
                      Tsp509I    NciI         HpyCH4IV
        FokI          BccI       ScrFI   BfaI        HhaI
 HpyCH4V       BstF5I MnlI AseI  HpaII   AluI   MseI FspI
  EciI    AciI   BsrI   MseI   StyD4I  CviJI  BsrI   HinP1I      BsrDI
      tcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccatt
8501  ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8600
      aggacgttgaaataggcggaggtaggtcagataattaacaacggcccttcgatctcattcatcaagcggtcaattatcaaacgcgttgcaacaacggtaa

a      L  Q  L  Y  P  P  P  S  S  L  L  I  V  A  G  K  L  E  *  V  V  R  Q  L  I  V  C  A  T  L  L  P  L  -
b     S  C  N  F  I  R  L  H  P  V  Y  *  L  L  P  G  S  *  S  K  *  F  A  S  *  *  F  A  Q  R  C  C  H  C  -
c      P  A  T  L  S  A  S  I  Q  S  I  N  C  C  R  E  A  R  V  S  S  S  P  V  N  S  L  R  N  V  V  A  I  -

                                                                BstKTI
                                                                DpnI
                                        NlaIV                   NlaIII
                                   HpaII                        MboI
        MaeIII                     BsaWI      BstKTI      FatI
        SfaNI                      AluI       DpnI        AlwI          NlaIII     HpyCH4V
        Tsp45I       TspDTI        CviJI      MboI   MaeIII                FatI
 SfcI   MslI         Hpy99I   CviJI
      gctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgca
8601  ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8700
      cgatgtccgtagcaccacagtgcgagcagcaaaccataccgaagtaagtcgaggccaagggttgctagttccgctcaatgtactagggggtacaacacgt

a      L  Q  A  S  W  C  H  A  R  R  L  V  W  L  H  S  A  P  V  P  N  D  Q  G  E  L  H  D  P  P  C  C  A  -
b       Y  R  H  R  G  V  T  L  V  V  W  Y  G  F  I  Q  L  R  F  P  T  I  K  A  S  Y  M  I  P  H  V  V  Q  -
c     A  T  G  I  V  V  S  R  S  S  F  G  M  A  S  F  S  S  G  S  Q  R  S  R  R  V  T  *  S  P  M  L  C  K  -

                      Hpy188I
                      MnlI
                  BsiEI                                                        BbvI
                  BstKTI                                                       Tsp509I
                  PvuI                  TauI                                   TspRI
                  DpnI                  AciI                                   HpyCH4V
      AluI      Hpy188I                 Fnu4HI             MslI       BtsI
      CviJI                             CviJI    BtsI      NlaIII     Fnu4HI
  TaqII    AvaII                        HaeIII   TspRI  FatI          TseI
  AciI     Sau96I  MboI                 EaeI                                         FokI
      aaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttac
8701  ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8800
      tttttcgccaatcgaggaagccaggaggctagcaacagtcttcattcaaccggcgtcacaatagtgagtaccaataccgtcgtgacgtattaagagaatg

a      K  K  R  L  A  P  S  V  L  R  S  L  S  E  V  S  W  P  Q  C  Y  H  S  W  L  W  Q  H  C  I  I  L  L  L  -
b       K  S  G  *  L  L  R  S  S  D  R  C  Q  K  *  V  G  R  S  V  I  T  H  G  Y  G  S  T  A  *  F  S  Y  -
c        K  A  V  S  S  F  G  P  P  I  V  V  R  S  K  L  A  A  V  L  S  L  M  V  M  A  A  L  H  N  S  L  T  -

                                                                                      BcgI
                                                                                      BsaHI
                                                                                      HpaII
     BstF5I                                                                           NciI
     SfaNI                              RsaI                                          ScrFI
     NlaIII                             ScaI                           BsiEI          TaqII
   FatI                                 Csp6I   BspCNI            TauI                StyD4I
 HpyCH4III             MaeIII           TatI    BseMII  Hpy188I   Fnu4HI
 TspGWI       BccI     Tsp45I         BsrI      HphI    DdeI   BcgI   AciI       HgaI
      tgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcg
8801  ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8900
      acagtacggtaggcattctacgaaaagacactgaccactcatgagttggttcagtaagactcttatcacatacgccgctggctcaacgagaacgggccgc

a       S  C  H  P  *  D  A  F  L  *  L  V  S  T  Q  P  S  H  S  E  N  S  V  C  G  D  R  V  A  L  A  R  R  -
b      C  H  A  I  R  K  M  L  F  C  D  W  *  V  L  N  Q  V  I  L  R  I  V  Y  A  A  T  E  L  L  L  P  G  V  -
c       V  M  P  S  V  R  C  F  S  V  T  G  E  Y  S  T  K  S  F  *  E  *  C  M  R  R  P  S  C  S  C  P  A  -

                                                          BpuEI
                                                         TaiI                          AlwI
                                                      AclI                 BstKTI MspA1I
           HhaI                                    HpyCH4IV                  DpnI  AciI
        BstUI                                       XmnI                  BstYI   AlwI
        HinP1I               DraI   BsiHKAI                                 
        AciI                 MseI   Bsp1286I    MboII                SmlI  MboI   PpiI
      tcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgt
8901  ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9000
      agttatgccctattatggcgcggtgtatcgtcttgaaattttcacgagtagtaaccttttgcaagaagccccgcttttgagagttcctagaatggcgaca

a       Q  Y  G  I  I  P  R  H  I  A  E  L  *  K  C  S  S  L  E  N  V  L  R  G  E  N  S  Q  G  S  Y  R  C  -
b        N  T  G  *  Y  R  A  T  *  Q  N  F  K  S  A  H  H  W  K  T  F  F  G  A  K  T  L  K  D  L  T  A  V  -
c      S  I  R  D  N  T  A  P  H  S  R  T  L  K  V  L  I  I  G  K  R  S  S  G  R  K  L  S  R  I  L  P  L  L  -
```

Figure 1-V

```
                                   MboII
                              Bme1580I |
                               BsiHKAI |
                              Bsp1286I |
                                 Hpy8I |
                               HpyCH4V |
       BsrI                      ApaLI |
     BstKTI|                     Eco57I |
       DpnI|                   Eco57MI |   BstKTI
     BstYI |     MaeIII          BssSI |    DpnI|        SfaNI
     MboI  |   TaqI  |           PpiI  |   MboI ||TaqII  HphI |   TaqII               HphI
     tgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaa
9001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9100
     actctaggtcaagctacattgggtgagcacgtgggttgactagaagtcgtagaaaatgaaagtggtcgcaaagacccactcgtttttgtccttccgtttt

a     * D P V R C N P L V H P T D L Q H L L L S P A F L G E Q K Q E G K M -
b     E I Q F D V T H S C T Q L I F S I F Y F H Q R F W V S K N R K A K -
c     R S S S M * P T R A P N * S S A S F T F T S V S G * A K T G R Q N -

                                                                                              NlaIII
                                                                                               BciVI |
                                                                                               BsmAI |
      TauI                                                                                Hpy188III| |
      AciI |                                    MboII                                         BspHI|| |
    Fnu4HI |                        MslI        TspGWI          EarI       SspI                FatI|| |
     tgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatg
9101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9200
     acggcgttttttcccttattcccgctgtgcctttacaacttatgagtatgagaaggaaaaagttataataacttcgtaaatagtcccaataacagagtac

a      P Q K R E * G R H G N V E Y S Y S S F F N I I E A F I R V I V S * -
b     C R K K G N K G D T E M L N T H T L P F S I L L K H L S G L L S H E -
c     A A K K G I R A T R K C * I L I L F L F Q Y Y * S I Y Q G Y C L M -

                                                                                   DdeI
                                                 HhaI                            AatII|
                                               BstUI |                            TaiI|
                                              HinP1I |                            ZraI ||
     AciI                                      AciI ||                           BsaHI| ||
     BsrBI                                    NlaIV |||                      HpyCH4IV| ||
     agcggatacatatttgaatgtatttagaaaaataaacaaataggggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattatta
9201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9300
     tcgcctatgtataaacttacataaatcttttttatttgtttatccccaaggcgcgtgtaaaggggcttttcacggtggactgcagattctttggtaataat

a     A D T Y L N V F R K I N K * G F R A H F P E K C H L T S K K P L L -
b      R I H I * M Y L E K * T N R G S A H I S P K S A T * R L R N H Y Y -
c     S G Y I F E C I * K N K Q I G V P R T F P R K V P P D V * E T I I I -

                                 CviJI
                                HaeIII
                             Eco0109I|
       NlaIII                  Sau96I|
    Hpy188III |                BssSI ||
        BspHI||           Hpy188III  ||
         FatI|| MseI           MnlI| ||
     tcatgacattaacctataaaaataggcgtatcacgaggccctttcgtc
9301 ---------+---------+---------+---------+-------- 9348
     agtactgtaattggatatttttatccgcatagtgctccgggaaagcag

a     S * H * P I K I G V S R G P F V -
b      H D I N L * K * A Y H E A L S -
c      M T L T Y K N R R I T R P F R -
```

Figure 1-W

Enzymes that do cut and were not excluded:

AarI	AatII	AccI	Acc65I	AciI	AclI	AfeI	AflIII	AgeI	AhdI	AluI	AlwI	AlwNI
ApaI	ApaLI	ApoI	AseI	AvaI	AvaII	BaeI	BamHI	BanI	BanII	BbeI	BbsI	BbvI
BccI	BceAI	BcgI	BciVI	BclI	BfaI	BfrBI	BglI	BglII	BlpI	Bme1580I	BmgBI	BmrI
BpmI	Bpu10I	BpuEI	BsaI	BsaAI	BsaBI	BsaHI	BsaJI	BsaWI	BsaXI	BseMII	BseRI	BseYI
BsgI	BsiEI	BsiHKAI	BslI	BsmI	BsmAI	BsmBI	BsmFI	Bsp1286I	BspCNI	BspHI	BspMI	BsrI
BsrBI	BsrDI	BsrFI	BssHII	BssSI	BstAPI	BstBI	BstEII	BstF5I	BstKTI	BstNI	BstUI	BstXI
BstYI	Bsu36I	BtgI	BtsI	Cac8I	ClaI	Csp6I	CviJI	DdeI	DpnI	DraI	DraIII	DrdI
EaeI	EagI	EarI	EciI	Eco57I	EcoICRI	Eco57MI	EcoNI	EcoO109I	EcoRI	EcoRV	FalI	FatI
FauI	Fnu4HI	FokI	FspI	HaeII	HaeIII	HgaI	HhaI	Hin4I	HinP1I	HincII	HindIII	HinfI
HpaI	HpaII	HphI	Hpy8I	Hpy99I	Hpy188I	Hpy188III	HpyCH4III	HpyCH4IV	HpyCH4V	HpyF10VI	KasI	KpnI
MaeIII	MboI	MboII	MlyI	MmeI	MnlI	MscI	MseI	MslI	MspA1I	MwoI	NaeI	NarI
NciI	NcoI	NdeI	NgoMIV	NlaIII	NlaIV	NotI	NruI	NsiI	NspI	PacI	PciI	PfoI
PleI	PpiI	PsiI	PspGI	PspOMI	PstI	PvuI	PvuII	RsaI	SacI	SacII	SalI	SapI
Sau96I	ScaI	ScrFI	SexAI	SfaNI	SfcI	SfoI	SmaI	SmlI	SnaBI	SpeI	SphI	SspI
StyI	StyD4I	TaiI	TaqI	TaqII	TatI	TauI	TfiI	TseI	Tsp45I	Tsp509I	TspDTI	TspGWI
TspRI	XbaI	XcmI	XhoI	XmaI	XmnI	ZraI						

Enzymes that do not cut:

AflII	AleI	AloI	AscI	AsiSI	AvrII	BbvCI	BmtI	BplI	BsiWI	BspEI	BsrGI	BstZ17I
FseI	FspAI	MfeI	MluI	NheI	PflMI	PmeI	PmlI	PpuMI	PshAI	PsrI	RsrII	SanDI
SbfI	SfiI	SgrAI	SrfI	StuI	SwaI	Tth111I						

Enzymes excluded; MinCuts: 1 MaxCuts: 100000

NONE

Figure 2-A

pCV1 sequence

```
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttgaaaga gataaattga 420
acaagtatgg tcgtcccctg ttgggatgta ctattaaacc taaattgggg ttatctgcta 480
aaaactacgg tagagccgtt tatgaatgtc ttcgcggtgg acttgatttt actaaagatg 540
atgagaacgt gaactcacaa ccatttatgc gttggagaga tcgtttctta ttttgtgccg 600
aagcacttta taaagcacag gctgaaacag gtgaaatcaa agggcattac ttgaatgcta 660
ctgcaggtac atgcgaagaa atgatcaaaa gagctgtatt tgctagagaa ttgggcgttc 720
cgatcgtaat gcatgactac ttaacggggg gattcaccgc aaatactagc ttggctcatt 780
attgccgaga taatggtcta cttcttcaca tccaccgtgc aatgcatgcg gttattgata 840
gacagaagaa tcatggtatc cacttccggg tattagcaaa agcgttacgt atgtctggtg 900
gagatcatat tcactctggt accgtagtag gtaaacttga aggtgaaaga gacataactt 960
tgggctttgt tgatttactg cgtgatgatt ttgttgaaca agatcgaagt cgcggtattt 1020
atttcactca agattgggtc tctttaccag gtgttctacc cgtggcttca ggaggtattc 1080
acgtttggca tatgcctgct ctgaccgaga tctttgggga tgattccgta ctacagttcg 1140
gtggaggaac tttaggacat ccttggggta atgcgccagg tgccgtagct aatcgagtag 1200
ctctagaagc atgtgtaaaa gctcgtaatg aaggacgtga tcttgctcag gaaggtaatg 1260
aaattattcg cgaggcttgc aaatggagcc cggaactagc tgctgcttgt gaagtatgga 1320
aagagatcgt atttaatttt gcagcagtgg acgttttgga taagtaaaaa cagtagacat 1380
tagcagataa attagcagga aataaagaag gataaggaga aagaactcaa gtaattatcc 1440
ttcgttctct taattgaatt gcaattaaac tcggcccaat cttttactaa aaggattgag 1500
ccgaatacaa caaagattct attgcatata ttttgactaa gtatatactt acctagatat 1560
acaagatttg aaatacaaaa tctagaaaac taaatcaaaa tctaagactc aaatctttct 1620
attgttgtct tggatcatcg agcttgcatg cctgcagtcg actttcacag tttccattct 1680
gaaatgttct ctgtactata ataaatagta agtgaatcaa cttactacta aaaaaattag 1740
tagacttcct cttcggaata gaaatagcct atttctacat agggaaagtc gtgtgcaatg 1800
aaaaatgcaa gcacgatttg gggagaggtt ttttctctat tgtaacaagg aataattatc 1860
tactccatcc gactagttcc gggttcgagt cccgggcaac ccatatggaa actagaaagg 1920
agcaatctga gttttgattt ttcactcact tcatttacaa aatttttttgg tttggtaaat 1980
tttgttgtat ggatatacaa ctgtcggggc tggcttggtt gacattggta tatagtctat 2040
attatactgt taaataacaa gccttctatt atctttctag ttaatacgtg tgcttgggag 2100
tccttgcaat ttgaataaac caagatctta ccatggccgc gcggggttct catcatcatc 2160
atcatggtcc gcgggggccc gctagtggta gtgaccttga ccggtgcacc acttttgatg 2220
atgttcaagc tcctaattac actcaacata cttcatctat gaggggggtt tactatcctg 2280
atgaaatttt tagatcagac actctttatt taactcagga tttatttctt ccattttatt 2340
ctaatgttac agggtttcat actattaatc atacgtttgg caaccctgtc ataccttttta 2400
aggatggtat ttattttgct gccacagaga aatcaaatgt tgtccgtggt tgggtttttg 2460
gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct actaatgttg 2520
ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctgtt tctaaaccca 2580
tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact ttcgagtaca 2640
tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaaa cacttacgag 2700
agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggctat caacctatag 2760
atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt aagttgcctc 2820
ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct gctcaagaca 2880
tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact acatttatgc 2940
tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa aatccacttg 3000
ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaatttac cagacctcta 3060
atttcagggt tgttccctca ggagatgttg tgagattccc taatattaca aacttgtgtc 3120
cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg gagagaaaaa 3180
aaatttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt ttttcaacct 3240
ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc aatgtctatg 3300
cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgccagga caaactggtg 3360
```

Figure 2-B

ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc cttgcttgga 3420
atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat aggtatctta 3480
gacatggcaa gcttaggccc tttgagagag acatatctaa tgtgcctttc tcccctgatg 3540
gcaaaccttg caccccacct gctcttaatt gttattggcc attaaatgat tatggttttt 3600
acaccactac tggcattggc taccaacctt acagagttgt agtactttct tttgaacttt 3660
taaatgcacc ggccacggtt tgtggaccaa aattatccac tgaccttatt aagaaccagt 3720
gtgtcaattt taattttaat ggactcactg gtactggtgt gttaactcct tcttcaaaga 3780
gatttcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat tccgttcgag 3840
atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt gtaagtgtaa 3900
ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat gttaactgca 3960
ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc atatattcta 4020
ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag catgtcgaca 4080
cttcttatga gtgcgacatt cctattggag ctggcatttg tgctattacc atacagttag 4140
actagaaaag ggcgaattct gcagatatcc atcacactgg cggccgcatc gatgagctct 4200
aattaattaa ggcttttctg ctaacatata gcaatttttg aagaaaggaa agctagaaat 4260
acccaatatc ttgctgaagc aagatattgg gtatttcttt ttttttattt ttgaatcttt 4320
ctattctgaa ttcagttaac gacgagattt agtatccttt cttgcacttt cataactcgt 4380
gaaatgccga gttggtacga attcccccaa tttgcgacct accataggat ttgttatgta 4440
aataggtata tgttcctttc cattatgaat cgcgattgta tggccaacca ttgcgggtag 4500
aatgctagat gcccgggacc acgttactat tgtttctttc tcctccttca tattgagctt 4560
ttctattttt gccaataaat gatggtacct ccccccgcca cgatcgaacg ggaatggata 4620
agaggcttgt gggattgacg tgatagggta gggttggcta tactgctggt ggcgaactcc 4680
aggctaataa tctgaagcgc ttggaagttg tagggaggat cctatggctc gtgaagcggt 4740
tatcgccgaa gtatcgactc aactatcaga ggtagttggc gtcatcgagc gccatctcga 4800
accgacgttg ctggccgtac atttgtacgg ctccgcagtg gatggcggcc tgaagccaca 4860
cagtgatatt gatttgctgg ttacggtgac cgtaaggctt gatgaaacaa cgcggcgagc 4920
tttgatcaac gaccttttgg aaacttcggc ttcccctgga gagagcgaga ttctccgcgc 4980
tgtagaagtc accattgttg tgcacgacga catcattccg tggcgttatc cagctaagcg 5040
cgaactgcaa tttggagaat ggcagcgcaa tgacattctt gcaggtatct tcgagccagc 5100
cacgatcgac attgatctgg ctatcttgct gacaaaagca agagaacata gcgttgcctt 5160
ggtaggtcca gcggcggagg aactctttga tccggttcct gaacaggatc tatttgaggc 5220
gctaaatgaa accttaacgc tatggaactc gccgcccgac tgggctggcg atgagcgaaa 5280
tgtagtgctt acgttgtccc gcatttggta cagcgcagta accggcaaaa tcgcgccgaa 5340
ggatgtcgct gccgactggg caatggagcg cctgccggcc cagtatcagc ccgtcatact 5400
tgaagctaga caggcttatc ttggacaaga agaagatcgc ttggcctcgc gcgcagatca 5460
gttggaagaa tttgtccact acgtgaaagg cgagatcact aaggtagttg gcaaataact 5520
cgagactaag tggataaaat tagatagaaa aaaggtctaa ataaaaaaga agagaaatag 5580
aaagatcaaa aatcagttac gaaaatgcag taattcttct tttttcttct aattgattgc 5640
aattaaactc gtctcaatct gaaaaaagat tgagccgagt ttaaatagat tttgatacga 5700
tcatgagact tgacaaatcg ggattcctct attctatata tttagaagat ataaaggtat 5760
aatacaataa ataaatacaa tctagaagct cgatgatcct taggattggt atattctttt 5820
ctatcctgta gtttgtagtt tccctgaatc aagccaagta tcacacctct ttctacccat 5880
cctgtatatt gtcccctttg ttccgtgttg aaatagaacc ttaatttatt acttattttt 5940
ttattaaatt ttagatttgt tagtgattag atattagtat tagacgagat tttacgaaac 6000
aattattttt ttatttcttt ataggagagg acaaatctct tttttcgatg cgaatttgac 6060
acgacatagg agaagccgcc ctttattaaa aattatatta ttttaaataa tataaagggg 6120
gttccaacat attaatatat agtgaagtgt tcccccagat tcagaacttt ttttcaatac 6180
tcacaatcct tattagttaa taatcctagt gattggattt ctatgcttag tctgatagga 6240
aataagatat tcaaataaat aattttatag cgaatgacta ttcatctatt gtattttcat 6300
gcaaataggg ggcaagaaaa ctctatggaa agatggtggt ttaattcgat gttgtttaag 6360
aaggagttcg aacgcaggtg tgggctaaat aaatcaatgg gcagtcttgg tcctattgaa 6420
aataccaatg aagatccaaa tcgaaaagtg aaaaacattc atagttggag gaatcgtgac 6480
aattctagtt gcagtaatgt tgattattta ttcggcgtta aagacattcg gaatttcatc 6540
tctgatgaca ctttttagt tagtgatagg aatggagaca gttattccat ctattttgat 6600
attgaaaatc atatttttga gattgacaac gatcattctt ttctgagtga actagaaagt 6660
tctttttata gttatcgaaa ctcgaattat cggaataatg gatttagggg cgaagatccc 6720
tactataatt cttacatgta tgatactcaa tatagttgga ataatcacat taatagttgc 6780
attgatagtt atcttcagtc tcaaatctgt atagatactt ccattataag tggtagtgag 6840
aattacggtg acagttacat ttatagggcc gtttgtggtg gtgaaagtcg aaatagtagt 6900

Figure 2-C

```
gaaaacgagg gttccagtag acgaactcgc acgaagggca gtgatttaac tataagagaa 6960
agttctaatg atctcgaggt aactcaaaaa tacaggcatt tgtgggttca atgcgaaaat 7020
tgttatggat taaattataa gaaatttttg aaatcaaaaa tgaatatttg tgaacaatgt 7080
ggatatcatt tgaaaatgag tagttcagat agaattcgta atcatggtca tagctgtttc 7140
ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga agcataaagt 7200
gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg cgctcactgc 7260
ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg 7320
ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct 7380
cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca 7440
cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga 7500
accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgcccccct gacgagcatc 7560
acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg 7620
cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat 7680
acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt 7740
atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc 7800
agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg 7860
acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg 7920
gtgctacaga gttcttgaag tggtggccta actacggcta cactagaagg acagtatttg 7980
gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg 8040
gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca 8100
gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga 8160
acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga 8220
tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt 8280
ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt 8340
catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat 8400
ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag 8460
caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct 8520
ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt 8580
tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg 8640
cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca 8700
aaaaagcggt tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt 8760
tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat 8820
gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac 8880
cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa 8940
aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt 9000
tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt 9060
tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa 9120
gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt 9180
atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa 9240
taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtctaagaa accattatta 9300
tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcgtc 9348
```

Figure 3-A

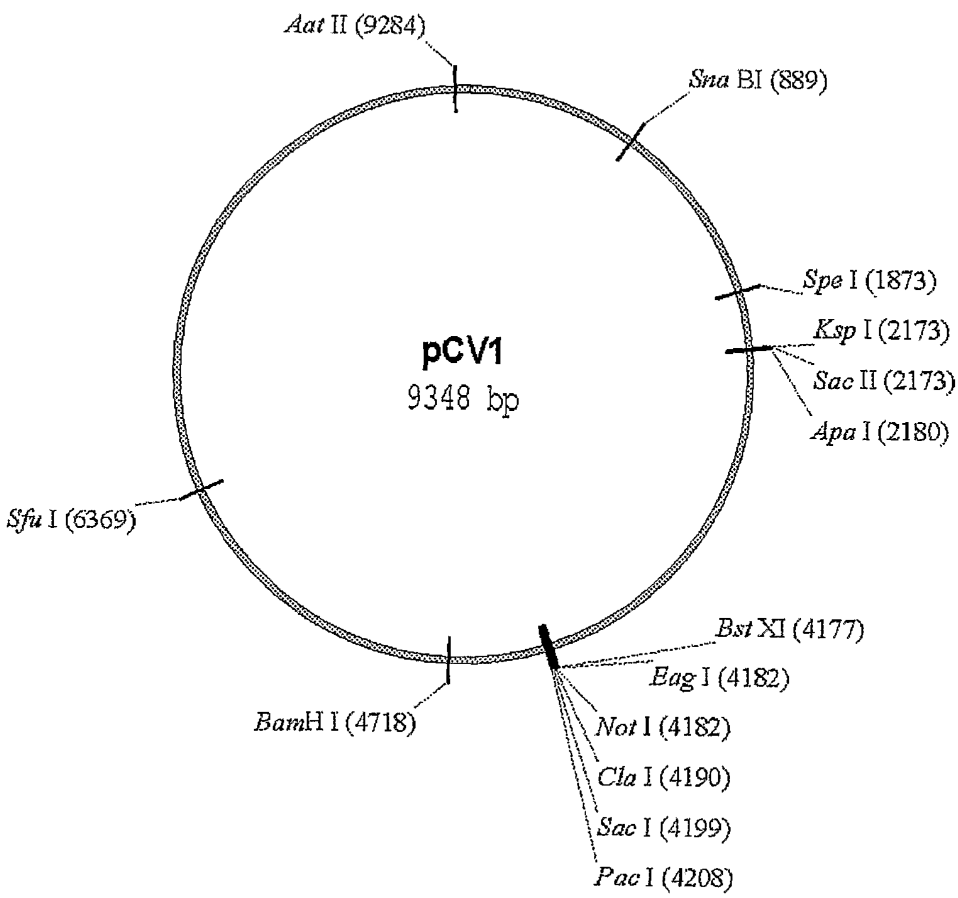

Aat II (9284)
Sna BI (889)
Spe I (1873)
Ksp I (2173)
Sac II (2173)
Apa I (2180)
pCV1
9348 bp
Sfu I (6369)
Bst XI (4177)
Eag I (4182)
Not I (4182)
Cla I (4190)
Sac I (4199)
Pac I (4208)
BamH I (4718)

Figure 3-B

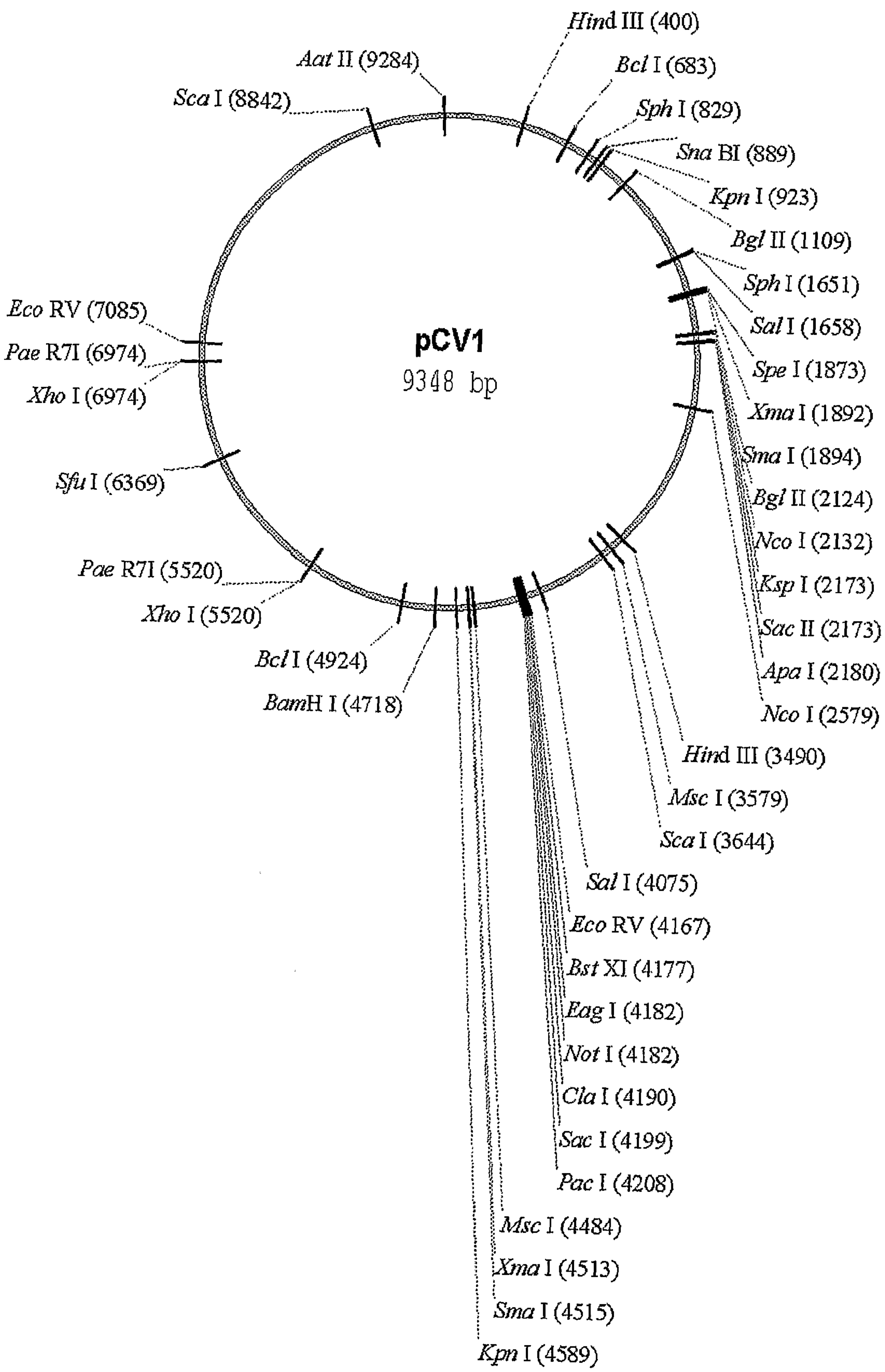

pCV1
9348 bp
Hind III (400)
Bcl I (683)
Sph I (829)
Sna BI (889)
Kpn I (923)
Bgl II (1109)
Sph I (1651)
Sal I (1658)
Spe I (1873)
Xma I (1892)
Sma I (1894)
Bgl II (2124)
Nco I (2132)
Ksp I (2173)
Sac II (2173)
Apa I (2180)
Nco I (2579)
Hind III (3490)
Msc I (3579)
Sca I (3644)
Sal I (4075)
Eco RV (4167)
Bst XI (4177)
Eag I (4182)
Not I (4182)
Cla I (4190)
Sac I (4199)
Pac I (4208)
Msc I (4484)
Xma I (4513)
Sma I (4515)
Kpn I (4589)
BamH I (4718)
Bcl I (4924)
Xho I (5520)
Pae R7I (5520)
Sfu I (6369)
Xho I (6974)
Pae R7I (6974)
Eco RV (7085)
Sca I (8842)
Aat II (9284)

Figure 3-C

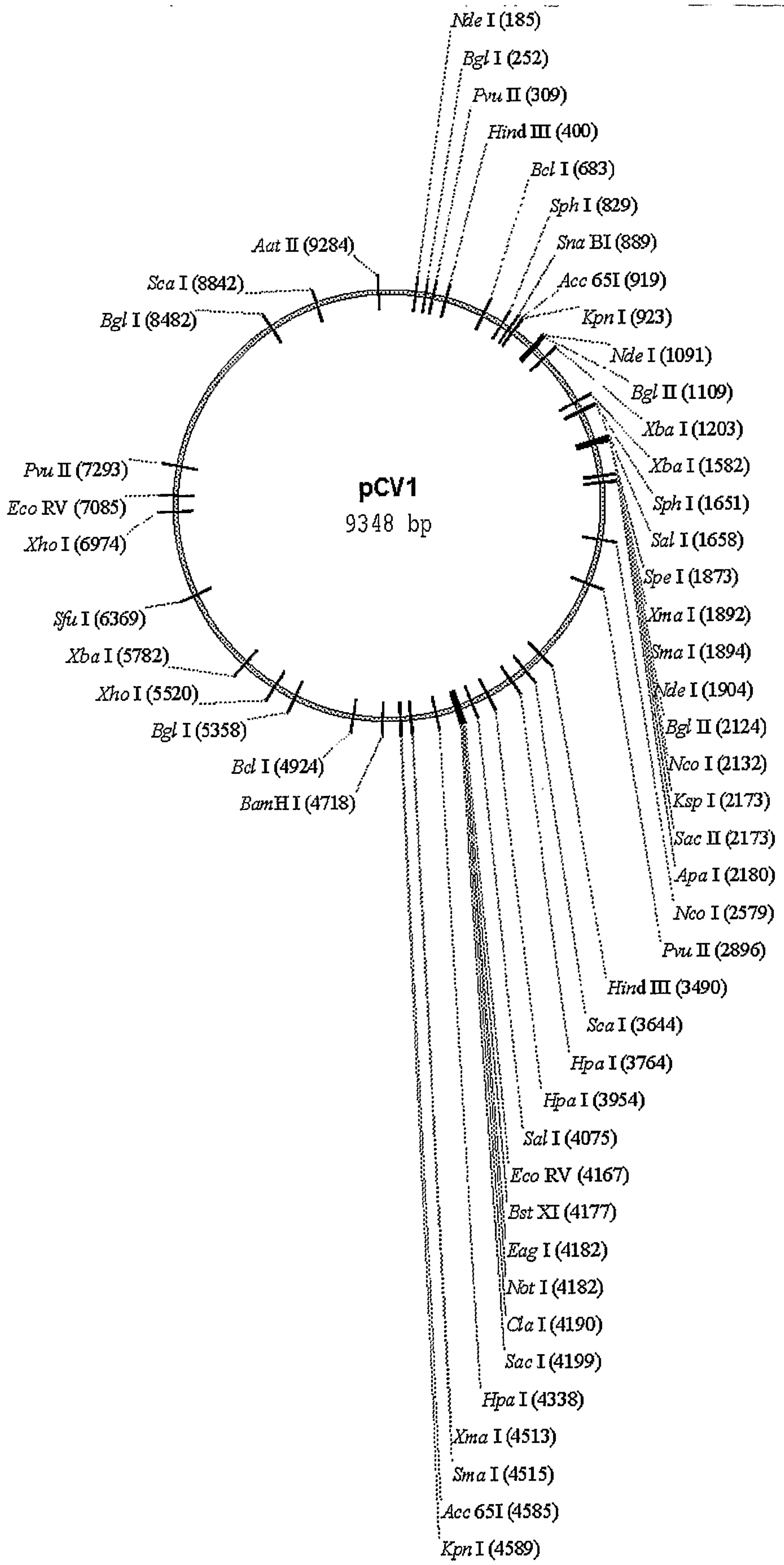

pCV1
9348 bp
Nde I (185)
Bgl I (252)
Pvu II (309)
Hind III (400)
Bcl I (683)
Sph I (829)
Sna BI (889)
Acc 65I (919)
Kpn I (923)
Nde I (1091)
Bgl II (1109)
Xba I (1203)
Xba I (1582)
Sph I (1651)
Sal I (1658)
Spe I (1873)
Xma I (1892)
Sma I (1894)
Nde I (1904)
Bgl II (2124)
Nco I (2132)
Ksp I (2173)
Sac II (2173)
Apa I (2180)
Nco I (2579)
Pvu II (2896)
Hind III (3490)
Sca I (3644)
Hpa I (3764)
Hpa I (3954)
Sal I (4075)
Eco RV (4167)
Bst XI (4177)
Eag I (4182)
Not I (4182)
Cla I (4190)
Sac I (4199)
Hpa I (4338)
Xma I (4513)
Sma I (4515)
Acc 65I (4585)
Kpn I (4589)
BamH I (4718)
Bcl I (4924)
Bgl I (5358)
Xho I (5520)
Xba I (5782)
Sfu I (6369)
Xho I (6974)
Eco RV (7085)
Pvu II (7293)
Bgl I (8482)
Sca I (8842)
Aat II (9284)

Figure 4

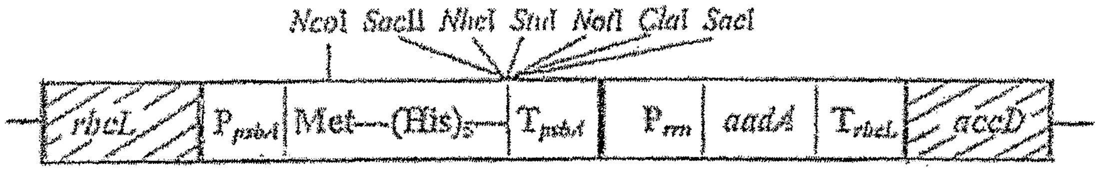

NcoI SacII NheI StuI NotI ClaI SacI
rbcL
P_{psbA}
Met---(His)$_5$---
T_{psbA}
P_{rrn}
aadA
T_{rbcL}
accD

Figure 5

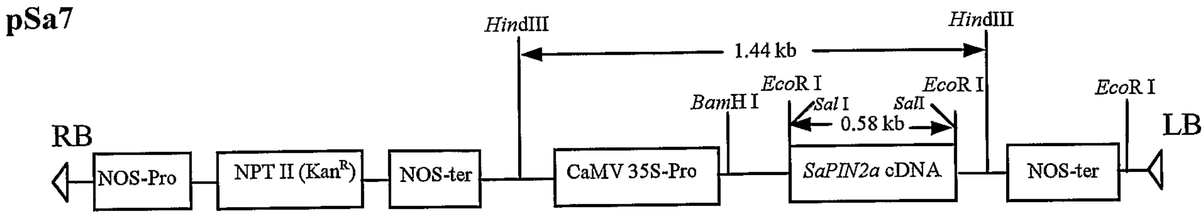

pSa7
HindIII
HindIII
1.44 kb
EcoR I
EcoR I
EcoR I
BamH I
Sal I
SalI
0.58 kb
RB
LB
NOS-Pro
NPT II (Kan^R)
NOS-ter
CaMV 35S-Pro
SaPIN2a cDNA
NOS-ter

Figure 6-A

1	ATATTAGGTT	TTTACCTACC	CAGGAAAAGC	CAACCAACCT	CGATCTCTTG	TAGATCTGTT
61	CTCTAAACGA	ACTTTAAAAT	CTGTGTAGCT	GTCGCTCGGC	TGCATGCCTA	GTGCACCTAC
121	GCAGTATAAA	CAATAATAAA	TTTTACTGTC	GTTGACAAGA	AACGAGTAAC	TCGTCCCTCT
181	TCTGCAGACT	GCTTACGGTT	TCGTCCGTGT	TGCAGTCGAT	CATCAGCATA	CCTAGGTTTC
241	GTCCGGGTGT	GACCGAAAGG	TAAGATGGAG	AGCCTTGTTC	TTGGTGTCAA	CGAGAAAACA
301	CACGTCCAAC	TCAGTTTGCC	TGTCCTTCAG	GTTAGAGACG	TGCTAGTGCG	TGGCTTCGGG
361	GACTCTGTGG	AAGAGGCCCT	ATCGGAGGCA	CGTGAACACC	TCAAAAATGG	CACTTGTGGT
421	CTAGTAGAGC	TGGAAAAAGG	CGTACTGCCC	CAGCTTGAAC	AGCCCTATGT	GTTCATTAAA
481	CGTTCTGATG	CCTTAAGCAC	CAATCACGGC	CACAAGGTCG	TTGAGCTGGT	TGCAGAAATG
541	GACGGCATTC	AGTACGGTCG	TAGCGGTATA	ACACTGGGAG	TACTCGTGCC	ACATGTGGGC
601	GAAACCCCAA	TTGCATACCG	CAATGTTCTT	CTTCGTAAGA	ACGGTAATAA	GGGAGCCGGT
661	GGTCATAGCT	ATGGCATCGA	TCTAAAGTCT	TATGACTTAG	GTGACGAGCT	TGGCACTGAT
721	CCCATTGAAG	ATTATGAACA	AAACTGGAAC	ACTAAGCATG	GCAGTGGTGC	ACTCCGTGAA
781	CTCACTCGTG	AGCTCAATGG	AGGTGCAGTC	ACTCGCTATG	TCGACAACAA	TTTCTGTGGC
641	CCAGATGGGT	ACCCTCTTGA	TTGCATCAAA	GATTTTCTCG	CACGCGCGGG	CAAGTCAATG
901	TGCACTCTTT	CCGAACAACT	TGATTACATC	GAGTCGAAGA	GAGGTGTCTA	CTGCTGCCGT
961	GACCATGAGC	ATGAAATTGC	CTGGTTCACT	GAGCGCTCTG	ATAAGAGCTA	CGAGCACCAG
1021	ACACCCTTCG	AAATTAAGAG	TGCCAAGAAA	TTTGACACTT	TCAAAGGGGA	ATGCCCAAAG
1081	TTTGTGTTTC	CTCTTAACTC	AAAAGTCAAA	GTCATTCAAC	CACGTGTTGA	AAAGAAAAAG
1141	ACTGAGGGTT	TCATGGGGCG	TATACGCTCT	GTGTACCCTG	TTGCATCTCC	ACAGGAGTGT
1201	AACAATATGC	ACTTGTCTAC	CTTGATGAAA	TGTAATCATT	GCGATGAAGT	TTCATGGCAG
1261	ACGTGCGACT	TTCTGAAAGC	CACTTGTGAA	CATTGTGGCA	CTGAAAATTT	AGTTATTGAA
1221	GGACCTACTA	CATGTGGGTA	CCTACCTACT	AATGCTGTAG	TGAAAATGCC	ATGTCCTGCC
1381	TGTCAAGACC	CAGAGATTGG	ACCTGAGCAT	AGTGTTGCAG	ATTATCACAA	CCACTCAAAC
1441	ATTGAAACTC	GACTCCGCAA	GGGAGGTAGG	ACTAGATGTT	TTGGAGGCTG	TGTGTTTGCC
1501	TATGTTGGCT	GCTATAATAA	GCGTGCCTAC	TGGGTTCCTC	GTGCTAGTGC	TGATATTGGC
1561	TCAGGCCATA	CTGGCATTAC	TGGTGACAAT	GTGGAGACCT	TGAATGAGGA	TCTCCTTGAG
1621	ATACTGAGTC	GTGAACGTGT	TAACATTAAC	ATTGTTGGCG	ATTTTCATTT	GAATGAAGAG
1681	GTTGCCATCA	TTTTGGCATC	TTTCTCTGCT	TCTACAAGTG	CCTTTATTGA	CACTATAAAG
1741	AGTCTTGATT	ACAAGTCTTT	CAAAACCATT	GTTGAGTCCT	GCGGTAACTA	TAAAGTTACC
1801	AAGGGAAAGC	CCGTAAAAGG	TGCTTGGAAC	ATTGGACAAC	AGAGATCAGT	TTTAACACCA
1661	CTGTGTGGTT	TTCCCTCACA	GGCTGCTGGT	GTTATCAGAT	CAATTTTTGC	GCGCACACTT
1921	GATGCAGCAA	ACCACTCAAT	TCCTGATTTG	CAAAGAGCAG	CTGTCACCAT	ACTTGATGGT
1981	ATTTCTGAAC	AGTCATTACG	TCTTGTCGAC	GCCATGGTTT	ATACTTCAGA	CCTGCTCACC
2041	AACAGTGTCA	TTATTATGGC	ATATGTAACT	GGTGGTCTTG	TACAACAGAC	TTCTCAGTGG
2101	TTGTCTAATC	TTTTGGGCAC	TACTGTTGAA	AAACTCAGGC	CTATCTTTGA	ATGGATTGAG
2161	GCGAAACTTA	GTGCAGGAGT	TGAATTTCTC	AAGGATGCTT	GGGAGATTCT	CAAATTTCTC
2221	ATTACAGGTG	TTTTTGACAT	CGTCAAGGGT	CAAATACAGG	TTGCTTCAGA	TAACATCAAG
2281	GATTGTGTAA	AATGCTTCAT	TGATGTTGTT	AACAAGGCAC	TCGAAATGTG	CATTGATCAA
2341	GTCACTATCG	CTGGCGCAAA	GTTGCGATCA	CTCAACTTAG	GTGAAGTCTT	CATCGCTCAA
2401	AGCAAGGGAC	TTTACCGTCA	GTGTATACGT	GGCAAGGAGC	AGCTGCAACT	ACTCATGCCT
2461	CTTAAGGCAC	CAAAAGAAGT	AACCTTTCTT	GAAGGTGATT	CACATGACAC	AGTACTTACC
2521	TCTGAGGAGG	TTGTTCTCAA	GAACGGTGAA	CTCGAAGCAC	TCGAGACGCC	CGTTGATAGC
2581	TTCACAAATG	GAGCTATCGT	CGGCACACCA	GTCTGTGTAA	ATGGCCTCAT	GCTCTTAGAG
2641	ATTAAGGACA	AAGAACAATA	CTGCGCATTG	TCTCCTGGTT	TACTGGCTAC	AAACAATGTC
2701	TTTCGCTTAA	AAGGGGGTGC	ACCAATTAAA	GGTGTAACCT	TTGGAGAAGA	TACTGTTTGG
2761	GAAGTTCAAG	GTTACAAGAA	TGTGAGAATC	ACATTTGAGC	TTGATGAACG	TGTTGACAAA
2821	GTGCTTAATG	AAAAGTGCTC	TGTCTACACT	GTTGAATCCG	GTACCGAAGT	TACTGAGTTT
2881	GCATGTGTTG	TAGCAGAGGC	TGTTGTGAAG	ACTTTACAAC	CAGTTTCTGA	TCTCCTTACC
2941	AACATGGGTA	TTGATCTTGA	TGAGTGGAGT	GTAGCTACAT	TCTACTTATT	TGATGATGCT
3001	GGTGAAGAAA	ACTTTTCATC	ACGTATGTAT	TGTTCCTTTT	ACCCTCCAGA	TGAGGAAGAA
3061	GAGGACGATG	CAGAGTGTGA	GGAAGAAGAA	ATTGATGAAA	CCTGTGAACA	TGAGTACGGT
3121	ACAGAGGATG	ATTATCAAGG	TCTCCCTCTG	GAATTTGGTG	CCTCAGCTGA	AACAGTTCGA
3181	GTTGAGGAAG	AAGAAGAGGA	AGACTGGCTG	GATGATACTA	CTGAGCAATC	AGAGATTGAG
3241	CCAGAACCAG	AACCTACACC	TGAAGAACCA	GTTAATCAGT	TTACTGGTTA	TTTAAAACTT
3301	ACTGACAATG	TTGCCATTAA	ATGTGTTGAC	ATCGTTAAGG	AGGCACAAAG	TGCTAATCCT

Figure 6-B

3361	ATGGTGATTG	TAAATGCTGC	TAACATACAC	CTGAAACATG	GTGGTGGTGT	AGCAGGTGCA
3421	CTCAACAAGG	CAACCAATGG	TGCCATGCAA	AAGGAGAGTG	ATGATTACAT	TAAGCTAAAT
3481	GGCCCTCTTA	CAGTAGGAGG	GTCTTGTTTG	CTTTCTGGAC	ATAATCTTGC	TAAGAAGTGT
3541	CTGCATGTTG	TTGGACCTAA	CCTAAATGCA	GGTGAGGACA	TCCAGCTTCT	TAAGGCAGCA
3601	TATGAAAATT	TCAATTCACA	GGACATCTTA	CTTGCACCAT	TGTTGTCAGC	AGGCATATTT
3661	GGTGCTAAAC	CACTTCAGTC	TTTACAAGTG	TGCGTGCAGA	CGGTTCGTAC	ACAGGTTTAT
3721	ATTGCAGTCA	ATGACAAAGC	TCTTTATGAG	CAGGTTGTCA	TGGATTATCT	TGATAACCTG
3781	AAGCCTAGAG	TGGAAGCACC	TAAACAAGAG	GAGCCACCAA	ACACAGAAGA	TTCCAAAACT
3841	GAGGAGAAAT	CTGTCGTACA	GAAGCCTGTC	GATGTGAAGC	CAAAAATTAA	GGCCTGCATT
3901	GATGAGGTTA	CCACAACACT	GGAAGAAACT	AAGTTTCTTA	CCAATAAGTT	ACTCTTGTTT
3961	GCTGATATCA	ATGGTAAGCT	TTACCATGAT	TCTCAGAACA	TGCTTAGAGG	TGAAGATATG
4021	TCTTTCCTTG	AGAAGGATGC	ACCTTACATG	GTAGGTGATG	TTATCACTAG	TGGTGATATC
4081	ACTTGTGTTG	TAATACCCTC	CAAAAAGGCT	GGTGGCACTA	CTGAGATGCT	CTCAAGAGCT
4141	TTGAAGAAAG	TGCCAGTTGA	TGAGTATATA	ACCACGTACC	CTGGACAAGG	ATGTGCTGGT
4201	TATACACTTG	AGGAAGCTAA	GACTGCTCTT	AAGAAATGCA	AATCTGCATT	TTATGTACTA
4261	CCTTCAGAAG	CACCTAATGC	TAAGGAAGAG	ATTCTAGGAA	CTGTATCCTG	GAATTTGAGA
4321	GAAATGCTTG	CTCATGCTGA	AGAGACAAGA	AAATTAATGC	CTATATGCAT	GGATGTTAGA
4381	GCCATAATGG	CAACCATCCA	ACGTAAGTAT	AAAGGAATTA	AAATTCAAGA	GGGCATCGTT
4441	GACTATGGTG	TCCGATTCTT	CTTTTATACT	AGTAAAGAGC	CTGTAGCTTC	TATTATTACG
4501	AAGCTGAACT	CTCTAAATGA	GCCGCTTGTC	ACAATGCCAA	TTGGTTATGT	GACACATGGT
4561	TTTAATCTTG	AAGAGGCTGC	GCGCTGTATG	CGTTCTCTTA	AAGCTCCTGC	CGTAGTGTCA
4621	GTATCATCAC	CAGATGCTGT	TACTACATAT	AATGGATACC	TCACTTCGTC	ATCAAAGACA
4681	TCTGAGGAGC	ACTTTGTAGA	AACAGTTTCT	TTGGCTGGCT	CTTACAGAGA	TTGGTCCTAT
4741	TCAGGACAGC	GTACAGAGTT	AGGTGTTGAA	TTTCTTAAGC	GTGGTGACAA	AATTGTGTAC
4801	CACACTCTGG	AGAGCCCCGT	CGAGTTTCAT	CTTGACGGTG	AGGTTCTTTC	ACTTGACAAA
4861	CTAAAGAGTC	TCTTATCCCT	GCGGGAGGTT	AAGACTATAA	AAGTGTTCAC	AACTGTGGAC
4921	AACACTAATC	TCCACACACA	GCTTGTGGAT	ATGTCTATGA	CATATGGACA	GCAGTTTGGT
4981	CCAACATACT	TGGATGGTGC	TGATGTTACA	AAAATTAAAC	CTCATGTAAA	TCATGAGGGT
5041	AAGACTTTCT	TTGTACTACC	TAGTGATGAC	ACACTACGTA	GTGAAGCTTT	CGAGTACTAC
5101	CATACTCTTG	ATGAGAGTTT	TCTTGGTAGG	TACATGTCTG	CTTTAAACCA	CACAAAGAAA
5161	TGGAAATTTC	CTCAAGTTGG	TGGTTTAACT	TCAATTAAAT	GGGCTGATAA	CAATTGTTAT
5221	TTGTCTAGTG	TTTTATTAGC	ACTTCAACAG	CTTGAAGTCA	AATTCAATGC	ACCAGCACTT
5281	CAAGAGGCTT	ATTATAGAGC	CCGTGCTGGT	GATGCTGCTA	ACTTTTGTGC	ACTCATACTC
5341	GCTTACAGTA	ATAAAACTGT	TGGCGAGCTT	GGTGATGTCA	GAGAAACTAT	GACCCATCTT
5401	CTACAGCATG	CTAATTTGGA	ATCTGCAAAG	CGAGTTCTTA	ATGTGGTGTG	TAAACATTGT
5461	GGTCAGAAAA	CTACTACCTT	AACGGGTGTA	GAAGCTGTGA	TGTATATGGG	TACTCTATCT
5521	TATGATAATC	TTAAGACAGG	TGTTTCCATT	CCATGTGTGT	GTGGTCGTGA	TGCTACACAA
5581	TATCTAGTAC	AACAAGAGTC	TTCTTTTGTT	ATGATGTCTG	CACCACCTGC	TGAGTATAAA
5641	TTACAGCAAG	GTACATTCTT	ATGTGCGAAT	GAGTACACTG	GTAACTATCA	GTGTGGTCAT
5701	TACACTCATA	TAACTGCTAA	GGAGACCCTC	TATCGTATTG	ACGGAGCTCA	CCTTACAAAG
5761	ATGTCAGAGT	ACAAAGGACC	AGTGACTGAT	GTTTTCTACA	AGGAAACATC	TTACACTACA
5821	ACCATCAAGC	CTGTGTCGTA	TAAACTCGAT	GGAGTTACTT	ACACAGAGAT	TGAACCAAAA
5881	TTGGATGGGT	ATTATAAAAA	GGATAATGCT	TACTATACAG	AGCAGCCTAT	AGACCTTGTA
5941	CCAACTCAAC	CATTACCAAA	TGCGAGTTTT	GATAATTTCA	AACTCACATG	TTCTAACACA
6001	AAATTTGCTG	ATGATTTAAA	TCAAATGACA	GGCTTCACAA	AGCCAGCTTC	ACGAGAGCTA
6061	TCTGTCACAT	TCTTCCCAGA	CTTGAATGGC	GATGTAGTGG	CTATTGACTA	TAGACACTAT
6121	TCAGCGAGTT	TCAAGAAAGG	TGCTAAATTA	CTGCATAAGC	CAATTGTTTG	GCACATTAAC
6181	CAGGCTACAA	CCAAGACAAC	GTTCAAACCA	AACACTTGGT	GTTTACGTTG	TCTTTGGAGT
6241	ACAAAGCCAG	TAGATACTTC	AAATTCATTT	GAAGTTCTGG	CAGTAGAAGA	CACACAAGGA
6301	ATGGACAATC	TTGCTTGTGA	AAGTCAACAA	CCCACCTCTG	AAGAAGTAGT	GGAAAATCCT
6361	ACCATACAGA	AGGAAGTCAT	AGAGTGTGAC	GTGAAAACTA	CCGAAGTTGT	AGGCAATGTC
6421	ATACTTAAAC	CATCAGATGA	AGGTGTTAAA	GTAACACAAG	AGTTAGGTCA	TGAGGATCTT
6481	ATGGCTGCTT	ATGTGGAAAA	CACAAGCATT	ACCATTAAGA	AACCTAATGA	GCTTTCACTA
6541	GCCTTAGGTT	TAAAAACAAT	TGCCACTCAT	GGTATTGCTG	CAATTAATAG	TGTTCCTTGG
6601	AGTAAAATTT	TGGCTTATGT	CAAACCATTC	TTAGGACAAG	CAGCAATTAC	AACATCAAAT
6661	TGCGCTAAGA	GATTAGCACA	ACGTGTGTTT	AACAATTATA	TGCCTTATGT	GTTTACATTA
6721	TTGTTCCAAT	TGTGTACTTT	TACTAAAAGT	ACCAATTCTA	GAATTAGAGC	TTCACTACCT
6781	ACAACTATTG	CTAAAAATAG	TGTTAAGAGT	GTTGCTAAAT	TATGTTTGGA	TGCCGGCATT

Figure 6-C

6841	AATTATGTGA	AGTCACCCAA	ATTTTCTAAA	TTGTTCACAA	TCGCTATGTG	GCTATTGTTG
6901	TTAAGTATTT	GCTTAGGTTC	TCTAATCTGT	GTAACTGCTG	CTTTTGGTGT	ACTCTTATCT
6961	AATTTTGGTG	CTCCTTCTTA	TTGTAATGGC	GTTAGAGAAT	TGTATCTTAA	TTCGTCTAAC
7021	GTTACTACTA	TGGATTTCTG	TGAAGGTTCT	TTTCCTTGCA	GCATTTGTTT	AAGTGGATTA
7081	GACTCCCTTG	ATTCTTATCC	AGCTCTTGAA	ACCATTCAGG	TGACGATTTC	ATCGTACAAG
7141	CTAGACTTGA	CAATTTTAGG	TCTGGCCGCT	GAGTGGGTTT	TGGCATATAT	GTTGTTCACA
7201	AAATTCTTTT	ATTTATTAGG	TCTTTCAGCT	ATAATGCAGG	TGTTCTTTGG	CTATTTTGCT
7261	AGTCATTTCA	TCAGCAATTC	TTGGCTCATG	TGGTTTATCA	TTAGTATTGT	ACAAATGGCA
7321	CCCGTTTCTG	CAATGGTTAG	GATGTACATC	TTCTTTGCTT	CTTTCTACTA	CATATGGAAG
7381	AGCTATGTTC	ATATCATGGA	TGGTTGCACC	TCTTCGACTT	GCATGATGTG	CTATAAGCGC
7441	AATCGTGCCA	CACGCGTTGA	GTGTACAACT	ATTGTTAATG	GCATGAAGAG	ATCTTTCTAT
7501	GTCTATGCAA	ATGGAGGCCG	TGGCTTCTGC	AAGACTCACA	ATTGGAATTG	TCTCAATTGT
7561	GACACATTTT	GCACTGGTAG	TACATTCATT	AGTGATGAAG	TTGCTCGTGA	TTTGTCACTC
7621	CAGTTTAAAA	GACCAATCAA	CCCTACTGAC	CAGTCATCGT	ATATTGTTGA	TAGTGTTGCT
7681	GTGAAAAATG	GCGCGCTTCA	CCTCTACTTT	GACAAGGCTG	GTCAAAAGAC	CTATGAGAGA
7741	CATCCGCTCT	CCCATTTTGT	CAATTTAGAC	AATTTGAGAG	CTAACAACAC	TAAAGGTTCA
7801	CTGCCTATTA	ATGTCATAGT	TTTTGATGGC	AAGTCCAAAT	GCGACGAGTC	TGCTTCTAAG
7861	TCTGCTTCTG	TGTACTACAG	TCAGCTGATG	TGCCAACCTA	TTCTGTTGCT	TGACCAAGCT
7921	CTTGTATCAA	ACGTTGGAGA	TAGTACTGAA	GTTTCCGTTA	AGATGTTTGA	TGCTTATGTC
7981	GACACCTTTT	CAGCAACTTT	TAGTGTTCCT	ATGGAAAAAC	TTAAGGCACT	TGTTGCTACA
8041	GCTCACAGCG	AGTTAGCAAA	GGGTGTAGCT	TTAGATGGTG	TCCTTTCTAC	ATTCGTGTCA
8101	GCTGCCCGAC	AAGGTGTTGT	TGATACCGAT	GTTGACACAA	AGGATGTTAT	TGAATGTCTC
8161	AAACTTTCAC	ATCACTCTGA	CTTAGAAGTG	ACAGGTGACA	GTTGTAACAA	TTTCATGCTC
8221	ACCTATAATA	AGGTTGAAAA	CATGACGCCC	AGAGATCTTG	GCGCATGTAT	TGACTGTAAT
8281	GCAAGGCATA	TCAATGCCCA	AGTAGCAAAA	AGTCACAATG	TTTCACTCAT	CTGGAATGTA
8341	AAAGACTACA	TGTCTTTATC	TGAACAGCTG	CGTAAACAAA	TTCGTACTGC	TGCCAAGAAG
8401	AACAACATAC	CTTTTACACT	AACTTGTGCT	ACAACTAGAC	AGGTTGTCAA	TGTCATAACT
8461	ACTAAAATCT	CACTCAAGGG	TGGTAAGATT	GTTAGTACTT	GTTTTAAACT	TATGCTTAAG
8521	GCCACATTAT	TGTGCGTTCT	TGCTGCATTG	GTTTGTTATA	TCGTTATGCC	AGTACATACA
8581	TTGTCAATCC	ATGATGGTTA	CACAAATGAA	ATCATTGGTT	ACAAAGCCAT	TCAGGATGGT
8641	GTCACTCGTG	ACATCATTTC	TACTGATGAT	TGTTTTGCAA	ATAAACATGC	TGGTTTTGAC
8701	GCATGGTTTA	GCCAGCGTGG	TGGTTCATAC	AAAAATGACA	AAAGCTGCCC	TGTAGTAGCT
8761	GCTATCATTA	CAAGAGAGAT	TGGTTTCATA	GTGCCTGGCT	TACCGGGTAC	TGTGCTGAGA
8821	GCAATCAATG	GTGACTTCTT	GCATTTTCTA	CCTCGTGTTT	TTAGTGCTGT	TGGCAACATT
8881	TGCTACACAC	CTTCCAAACT	CATTGAGTAT	AGTGATTTTG	CTACCTCTGC	TTGCGTTCTT
8941	GCTGCTGAGT	GTACAATTTT	TAAGGATGCT	ATGGGCAAAC	CTGTGCCATA	TTGTTATGAC
9001	ACTAATTTGC	TAGAGGGTTC	TATTTCTTAT	AGTGAGCTTC	GTCCAGACAC	TCGTTATGTG
9061	CTTATGGATG	GTTCCATCAT	ACAGTTTCCT	AACACTTACC	TGGAGGGTTC	TGTTAGAGTA
9121	GTAACAACTT	TTGATGCTGA	GTACTGTAGA	CATGGTACAT	GCGAAAGGTC	AGAAGTAGGT
9181	ATTTGCCTAT	CTACCAGTGG	TAGATGGGTT	CTTAATAATG	AGCATTACAG	AGCTCTATCA
9241	GGAGTTTTCT	GTGGTGTTGA	TGCGATGAAT	CTCATAGCTA	ACATCTTTAC	TCCTCTTGTG
9301	CAACCTGTGG	GTGCTTTAGA	TGTGTCTGCT	TCAGTAGTGG	CTGGTGGTAT	TATTGCCATA
9361	TTGGTGACTT	GTGCTGCCTA	CTACTTTATG	AAATTCAGAC	GTGTTTTTGG	TGAGTACAAC
9421	CATGTTGTTG	CTGCTAATGC	ACTTTTGTTT	TTGATGTCTT	TCACTATACT	CTGTCTGGTA
9481	CCAGCTTACA	GCTTTCTGCC	GGGAGTCTAC	TCAGTCTTTT	ACTTGTACTT	GACATTCTAT
9541	TTCACCAATG	ATGTTTCATT	CTTGGCTCAC	CTTCAATGGT	TTGCCATGTT	TTCTCCTATT
9601	GTGCCTTTTT	GGATAACAGC	AATCTATGTA	TTCTGTATTT	CTCTGAAGCA	CTGCCATTGG
9661	TTCTTTAACA	ACTATCTTAG	GAAAAGAGTC	ATGTTTAATG	GAGTTACATT	TAGTACCTTC
9721	GAGGAGGCTG	CTTTGTGTAC	CTTTTTGCTC	AACAAGGAAA	TGTACCTAAA	ATTGCGTAGC
9781	GAGACACTGT	TGCCACTTAC	ACAGTATAAC	AGGTATCTTG	CTCTATATAA	CAAGTACAAG
9841	TATTTCAGTG	GAGCCTTAGA	TACTACCAGC	TATCGTGAAG	CAGCTTGCTG	CCACTTAGCA
9901	AAGGCTCTAA	ATGACTTTAG	CAACTCAGGT	GCTGATGTTC	TCTACCAACC	ACCACAGACA
9961	TCAATCACTT	CTGCTGTTCT	GCAGAGTGGT	TTTAGGAAAA	TGGCATTCCC	GTCAGGCAAA
10021	GTTGAAGGGT	GCATGGTACA	AGTAACCTGT	GGAACTACAA	CTCTTAATGG	ATTGTGGTTG
10081	GATGACACAG	TATACTGTCC	AAGACATGTC	ATTTGCACAG	CAGAAGACAT	GCTTAATCCT
10141	AACTATGAAG	ATCTGCTCAT	TCGCAAATCC	AACCATAGCT	TTCTTGTTCA	GGCTGGCAAT
10201	GTTCAACTTC	GTGTTATTGG	CCATTCTATG	CAAAATTGTC	TGCTTAGGCT	TAAAGTTGAT
10261	ACTTCTAACC	CTAAGACACC	CAAGTATAAA	TTTGTCCGTA	TCCAACCTGG	TCAAACATTT

Figure 6-D

```
10321 TCAGTTCTAG CATGCTACAA TGGTTCACCA TCTGGTGTTT ATCAGTGTGC CATGAGACCT
10381 AATCATACCA TTAAAGGTTC TTTCCTTAAT GGATCATGTG GTAGTGTTGG TTTTAACATT
10441 GATTATGATT GCGTGTCTTT CTGCTATATG CATCATATGG AGCTTCCAAC AGGAGTACAC
10501 GCTGGTACTG ACTTAGAAGG TAAATTCTAT GGTCCATTTG TTGACAGACA AACTGCACAG
10561 GCTGCAGGTA CAGACACAAC CATAACATTA AATGTTTTGG CATGGCTGTA TGCTGCTGTT
10621 ATCAATGGTG ATAGGTGGTT TCTTAATAGA TTCACCACTA CTTTGAATGA CTTTAACCTT
10681 GTGGCAATGA AGTACAACTA TGAACCTTTG ACACAAGATC ATGTTGACAT ATTGGGACCT
10741 CTTTCTGCTC AAACAGGAAT TGCCGTCTTA GATATGTGTG CTGCTTTGAA AGAGCTGCTG
10801 CAGAATGGTA TGAATGGTCG TACTATCCTT GGTAGCACTA TTTTAGAAGA TGAGTTTACA
10861 CCATTTGATG TTGTTAGACA ATGCTCTGGT GTTACCTTCC AAGGTAAGTT CAAGAAAATT
10921 GTTAAGGGCA CTCATCATTG GATGCTTTTA ACTTTCTTGA CATCACTATT GATTCTTGTT
10981 CAAAGTACAC AGTGGTCACT GTTTTTCTTT GTTTACGAGA ATGCTTTCTT GCCATTTACT
11041 CTTGGTATTA TGGCAATTGC TGCATGTGCT ATGCTGCTTG TTAAGCATAA GCACGCATTC
11101 TTGTGCTTGT TTCTGTTACC TTCTCTTGCA ACAGTTGCTT ACTTTAATAT GGTCTACATG
11161 CCTGCTAGCT GGGTGATGCG TATCATGACA TGGCTTGAAT TGGCTGACAC TAGCTTGTCT
11221 GGTTATAGGC TTAAGGATTG TGTTATGTAT GCTTCAGCTT TAGTTTTGCT TATTCTCATG
11281 ACAGCTCGCA CTGTTTATGA TGATGCTGCT AGACGTGTTT GGACACTGAT GAATGTCATT
11341 ACACTTGTTT ACAAAGTCTA CTATGGTAAT GCTTTAGATC AAGCTATTTC CATGTGGGCC
11401 TTAGTTATTT CTGTAACCTC TAACTATTCT GGTGTCGTTA CGACTATCAT GTTTTTAGCT
11461 AGAGCTATAG TGTTTGTGTG TGTTGAGTAT TACCCATTGT TATTTATTAC TGGCAACACC
11521 TTACAGTGTA TCATGCTTGT TTATTGTTTC TTAGGCTATT GTTGCTGCTG CTACTTTGGC
11581 CTTTTCTGTT TACTCAACCG TTACTTCAGG CTTACTCTTG GTGTTTATGA CTACTTGGTC
11641 TCTACACAAG AATTTAGGTA TATGAACTCC CAGGGGCTTT TGCCTCCTAA GAGTAGTATT
11701 GATGCTTTCA AGCTTAACAT TAAGTTGTTG GGTATTGGAG GTAAACCATG TATCAAGGTT
11761 GCTACTGTAC AGTCTAAAAT GTCTGACGTA AAGTGCACAT CTGTGGTACT GCTCTCGGTT
11821 CTTCAACAAC TTAGAGTAGA GTCATCTTCT AAATTGTGGG CACAATGTGT ACAACTCCAC
11881 AATGATATTC TTCTTGCAAA AGACACAACT GAAGCTTTCG AGAAGATGGT TTCTCTTTTG
11941 TCTGTTTTGC TATCCATGCA GGGTGCTGTA GACATTAATA GGTTGTGCGA GGAAATGCTC
12001 GATAACCGTG CTACTCTTCA GGCTATTGCT TCAGAATTTA GTTCTTTACC ATCATATGCC
12061 GCTTATGCCA CTGCCCAGGA GGCCTATGAG CAGGCTGTAG CTAATGGTGA TTCTGAAGTC
12121 GTTCTCAAAA AGTTAAAGAA ATCTTTGAAT GTGGCTAAAT CTGAGTTTGA CCGTGATGCT
12181 GCCATGCAAC GCAAGTTGGA AAAGATGGCA GATCAGGCTA TGACCCAAAT GTACAAACAG
12241 GCAAGATCTG AGGACAAGAG GGCAAAAGTA ACTAGTGCTA TGCAAACAAT GCTCTTCACT
12301 ATGCTTAGGA AGCTTGATAA TGATGCACTT AACAACATTA TCAACAATGC GCGTGATGGT
12361 TGTGTTCCAC TCAACATCAT ACCATTGACT ACAGCAGCCA AACTCATGGT TGTTGTCCCT
12421 GATTATGGTA CCTACAAGAA CACTTGTGAT GGTAACACCT TTACATATGC ATCTGCACTC
12481 TGGGAAATCC AGCAAGTTGT TGATGCGGAT AGCAAGATTG TTCAACTTAG TGAAATTAAC
12541 ATGGACAATT CACCAAATTT GGCTTGGCCT CTTATTGTTA CAGCTCTAAG AGCCAACTCA
12601 GCTGTTAAAC TACAGAATAA TGAACTGAGT CCAGTAGCAC TACGACAGAT GTCCTGTGCG
12661 GCTGGTACCA CACAAACAGC TTGTACTGAT GACAATGCAC TTGCCTACTA TAACAATTCG
12721 AAGGGAGGTA GGTTTGTGCT GGCATTACTA TCAGACCACC AAGATCTCAA ATGGGCTAGA
12781 TTCCCTAAGA GTGATGGTAC AGGTACAATT TACACAGAAC TGGAACCACC TTGTAGGTTT
12641 GTTACAGACA CACCAAAAGG GCCTAAAGTG AAATACTTGT ACTTCATCAA AGGCTTAAAC
12901 AACCTAAATA GAGGTATGGT GCTGGGCAGT TTAGCTGCTA CAGTACGTCT TCAGGCTGGA
12961 AATGCTACAG AAGTACCTGC CAATTCAACT GTGCTTTCCT TCTGTGCTTT TGCAGTAGAC
13021 CCTGCTAAAG CATATAAGGA TTACCTAGCA AGTGGAGGAC AACCAATCAC CAACTGTGTG
13081 AAGATGTTGT GTACACACAC TGGTACAGGA CAGGCAATTA CTGTAACACC AGAAGCTAAC
13141 ATGGACCAAG AGTCCTTTGG TGGTGCTTCA TGTTGTCTGT ATTGTAGATG CCACATTGAC
13201 CATCCAAATC CTAAAGGATT CTGTGACTTG AAAGGTAAGT ACGTCCAAAT ACCTACCACT
13261 TGTGCTAATG ACCCAGTGGG TTTTACACTT AGAAACACAG TCTGTACCGT CTGCGGAATG
13321 TGGAAAGGTT ATGGCTGTAG TTGTGACCAA CTCCGCGAAC CCTTGATGCA GTCTGCGGAT
13381 GCATCAACGT TTTTAAACGG GTTTGCGGTG TAAGTGCAGC CCGTCTTACA CCGTGCGGCA
13441 CAGGCACTAG TACTGATGTC GTCTACAGGG CTTTTGATAT TTACAACGAA AAAAGTGCTG
13501 GTTTTGCAAA GTTCCTAAAA ACTAATTGCT GTCGCTTCCA GGAGAAGGAT GAGGAAGGCA
13561 ATTTATTAGA CTCTTACTTT GTAGTTAAGA GGCATACTAT GTCTAACTAC CAACATGAAG
13621 AGACTATTTA TAACTTGGTT AAAGATTGTC CAGCGGTTGC TGTCCATGAC TTTTTCAAGT
13681 TTAGAGTAGA TGGTGACATG GTACCACATA TATCACGTCA GCGTCTAACT AAATACACAA
13741 TGGCTGATTT AGTCTATGCT CTACGTCATT TTGATGAGGG TAATTGTGAT ACATTAAAAG
```

Figure 6-E

13801 AAATACTCGT CACATACAAT TGCTGTGATG ATGATTATTT CAATAAGAAG GATTGGTATG
13861 ACTTCGTAGA GAATCCTGAC ATCTTACGCG TATATGCTAA CTTAGGTGAG CGTGTACGCC
13921 AATCATTATT AAAGACTGTA CAATTCTGCG ATGCTATGCG TGATGCAGGC ATTGTAGGCG
13981 TACTGACATT AGATAATCAG GATCTTAATG GGAACTGGTA CGATTTCGGT GATTTCGTAC
14041 AAGTAGGACC AGGCTGCGGA GTTCCTATTG TGGATTCATA TTACTCATTG CTGATGCCCA
14101 TCCTCACTTT GACTAGGGCA TTGGCTGCTG AGTCCCATAT GGATGCTGAT CTCGCAAAAC
14161 CACTTATTAA GTGGGATTTG CTGAAATATG ATTTTACGGA AGAGAGACTT TGTCTCTTCG
14221 ACCGTTATTT TAAATATTGG GACCAGACAT ACCATCCCAA TTGTATTAAC TGTTTGGATG
14281 ATAGGTGTAT CCTTCATTGT GCAAACTTTA ATGTGTTATT TTCTACTGTG TTTCCACCTA
14341 CAAGTTTTGG ACCACTAGTA AGAAAAATAT TTGTAGATGG TGTTCCTTTT GTTGTTTCAA
14401 CTGGATACCA TTTTCGTGAG TTAGGAGTCG TACATAATCA GGATGTAAAC TTACATAGCT
14461 CGCGTCTCAG TTTCAAGGAA CTTTTAGTGT ATGCTGCTGA TCCAGCTATG CATGCAGCTT
14521 CTGGCAATTT ATTGCTAGAT AAACGCACTA CATGCTTTTC AGTAGCTGCA CTAACAAACA
14581 ATGTTGCTTT TCAAACTGTC AAACCCGGTA ATTTTAATAA AGACTTTTAT GACTTTGCTG
14641 TGTCTAAAGG TTTCTTTAAG GAAGGAAGTT CTGTTGAACT AAAACACTTC TTCTTTGCTC
14701 AGGATGGCAA CGCTGCTATC AGTGATTATG ACTATTATCG TTATAATCTG CCAACAATGT
14761 GTGATATCAG ACAACTCCTA TTCGTAGTTG AAGTTGTTGA TAAATACTTT GATTGTTACG
14821 ATGGTGGCTG TATTAATGCC AACCAAGTAA TCGTTAACAA TCTGGATAAA TCAGCTGGTT
14861 TCCCATTTAA TAAATGGGGT AAGGCTAGAC TTTATTATGA CTCAATGAGT TATGAGGATC
14941 AAGATGCACT TTTCGCGTAT ACTAAGCGTA ATGTCATCCC TACTATAACT CAAATGAATC
15001 TTAAGTATGC CATTAGTGCA AAGAATAGAG CTCGCACCGT AGCTGGTGTC TCTATCTGTA
15061 GTACTATGAC AAATAGACAG TTTCATCAGA AATTATTGAA GTCAATAGCC GCCACTAGAG
15121 GAGCTACTGT GGTAATTGGA ACAAGCAAGT TTTACGGTGG CTGGCATAAT ATGTTAAAAA
15181 CTGTTTACAG TGATGTAGAA ACTCCACACC TTATGGGTTG GGATTATCCA AAATGTGACA
15241 GAGCCATGCC TAACATGCTT AGGATAATGG CCTCTCTTGT TCTTGCTCGC AAACATAACA
15301 CTTGCTGTAA CTTATCACAC CGTTTCTACA GGTTAGCTAA CGAGTGTGCG CAAGTATTAA
15361 GTGAGATGGT CATGTGTGGC GGCTCACTAT ATGTTAAACC AGGTGGAACA TCATCCGGTG
15421 ATGCTACAAC TGCTTATGCT AATAGTGTCT TTAACATTTG TCAAGCTGTT ACAGCCAATG
15481 TAAATGCACT TCTTTCAACT GATGGTAATA AGATAGCTGA CAAGTATGTC CGCAATCTAC
15541 AACACAGGCT CTATGAGTGT CTCTATAGAA ATAGGGATGT TGATCATGAA TTCGTGGATG
15601 AGTTTTACGC TTACCTGCGT AAACATTTCT CCATGATGAT TCTTTCTGAT GATGCCGTTG
15661 TGTGCTATAA CAGTAACTAT GCGGCTCAAG GTTTAGTAGC TAGCATTAAG AACTTTAAGG
15721 CAGTTCTTTA TTATCAAAAT AATGTGTTCA TGTCTGAGGC AAAATGTTGG ACTGAGACTG
15781 ACCTTACTAA AGGACCTCAC GAATTTTGCT CACAGCATAC AATGCTAGTT AAACAAGGAG
15841 ATGATTACGT GTACCTGCCT TACCCAGATC CATCAAGAAT ATTAGGCGCA GGCTGTTTTG
15901 TCGATGATAT TGTCAAAACA GATGGTACAC TTATGATTGA AAGGTTCGTG TCACTGGCTA
15961 TTGATGCTTA CCCACTTACA AAACATCCTA ATCAGGAGTA TGCTGATGTC TTTCACTTGT
16021 ATTTACAATA CATTAGAAAG TTACATGATG AGCTTACTGG CCACATGTTG GACATGTATT
16081 CCGTAATGCT AACTAATGAT AACACCTCAC GGTACTGGGA ACCTGAGTTT TATGAGGCTA
16141 TGTACACACC ACATACAGTC TTGCAGGCTG TAGGTGCTTG TGTATTGTGC AATTCACAGA
16201 CTTCACTTCG TTGCGGTGCC TGTATTAGGA GACCATTCCT ATGTTGCAAG TGCTGCTATG
16261 ACCATGTCAT TTCAACATCA CACAAATTAG TGTTGTCTGT TAATCCCTAT GTTTGCAATG
16321 CCCCAGGTTG TGATGTCACT GATGTGACAC AACTGTATCT AGGAGGTATG AGCTATTATT
16381 GCAAGTCACA TAAGCCTCCC ATTAGTTTTC CATTATGTGC TAATGGTCAG GTTTTTGGTT
16441 TATACAAAAA CACATGTGTA GGCAGTGACA ATGTCACTGA CTTCAATGCG ATAGCAACAT
16501 GTGATTGGAC TAATGCTGGC GATTACATAC TTGCCAACAC TTGTACTGAG AGACTCAAGC
16561 TTTTCGCAGC AGAAACGCTC AAAGCCACTG AGGAAACATT TAAGCTGTCA TATGGTATTG
16621 CCACTGTACG CGAAGTACTC TCTGACAGAG AATTGCATCT TTCATGGGAG GTTGGAAAAC
16661 CTAGACCACC ATTGAACAGA AACTATGTCT TTACTGGTTA CCGTGTAACT AAAAATAGTA
16741 AAGTACAGAT TGGAGAGTAC ACCTTTGAAA AAGGTGACTA TGGTGATGCT GTTGTGTACA
16801 GAGGTACTAC GACATACAAG TTGAATGTTG GTGATTACTT TGTGTTGACA TCTCACACTG
16861 TAATGCCACT TAGTGCACCT ACTCTAGTGC CACAAGAGCA CTATGTGAGA ATTACTGGCT
16921 TGTACCCAAC ACTCAACATC TCAGATGAGT TTTCTAGCAA TGTTGCAAAT TATCAAAAGG
16981 TCGGCATGCA AAAGTACTCT ACACTCCAAG GACCACCTGG TACTGGTAAG AGTCATTTTG
17041 CCATCGGACT TGCTCTCTAT TACCCATCTG CTCGCATAGT GTATACGGCA TGCTCTCATG
17101 CAGCTGTTGA TGCCCTATGT GAAAAGGCAT TAAAATATTT GCCCATAGAT AAATGTAGTA
17161 GAATCATACC TGCGCGTGCG CGCGTAGAGT GTTTTGATAA ATTCAAAGTG AATTCAACAC
17221 TAGAACAGTA TGTTTTCTGC ACTGTAAATG CATTGCCAGA AACAACTGCT GACATTGTAG

Figure 6-F

17281	TCTTTGATGA	AATCTCTATG	GCTACTAATT	ATGACTTGAG	TGTTGTCAAT	GCTAGACTTC
17341	GTGCAAAACA	CTACGTCTAT	ATTGGCGATC	CTGCTCAATT	ACCAGCCCCC	CGCACATTGC
17401	TGACTAAAGG	CACACTAGAA	CCAGAATATT	TTAATTCAGT	GTGCAGACTT	ATGAAAACAA
17461	TAGGTCCAGA	CATGTTCCTT	GGAACTTGTC	GCCGTTGTCC	TGCTGAAATT	GTTGACACTG
17521	TGAGTGCTTT	AGTTTATGAC	AATAAGCTAA	AAGCACACAA	GGATAAGTCA	GCTCAATGCT
17561	TCAAAATGTT	CTACAAAGGT	GTTATTACAC	ATGATGTTTC	ATCTGCAATC	AACAGACCTC
17641	AAATAGGCGT	TGTAAGAGAA	TTTCTTACAC	GCAATCCTGC	TTGGAGAAAA	GCTGTTTTTA
17701	TCTCACCTTA	TAATTCACAG	AACGCTGTAG	CTTCAAAAAT	CTTAGGATTG	CCTACGCAGA
17761	CTGTTGATTC	ATCACAGGGT	TCTGAATATG	ACTATGTCAT	ATTCACACAA	ACTACTGAAA
17821	CAGCACACTC	TTGTAATGTC	AACCGCTTCA	ATGTGGCTAT	CACAAGGGCA	AAAATTGGCA
17881	TTTTGTGCAT	AATGTCTGAT	AGAGATCTTT	ATGACAAACT	GCAATTTACA	AGTCTAGAAA
17941	TACCACGTCG	CAATGTGGCT	ACATTACAAG	CAGAAAATGT	AACTGGACTT	TTTAAGGACT
18001	GTAGTAAGAT	CATTACTGGT	CTTCATCCTA	CACAGGCACC	TACACACCTC	AGCGTTGATA
18061	TAAAATTCAA	GACTGAAGGA	TTATGTGTTG	ACATACCAGG	CATACCAAAG	GACATGACCT
18121	ACCGTAGACT	CATCTCTATG	ATGGGTTTCA	AAATGAATTA	CCAAGTCAAT	GGTTACCCTA
18181	ATATGTTTAT	CACCCGCGAA	GAAGCTATTC	GTCACGTTCG	TGCGTGGATT	GGCTTTGATG
18241	TAGAGGGCTG	TCATGCAACT	AGAGATGCTG	TGGGTACTAA	CCTACCTCTC	CAGCTAGGAT
18301	TTTCTACAGG	TGTTAACTTA	GTAGCTGTAC	CGACTGGTTA	TGTTGACACT	GAAAATAACA
18361	CAGAATTCAC	CAGAGTTAAT	GCAAAACCTC	CACCAGGTGA	CCAGTTTAAA	CATCTTATAC
18421	CACTCATGTA	TAAAGGCTTG	CCCTGGAATG	TAGTGCGTAT	TAAGATAGTA	CAAATGCTCA
18481	GTGATACACT	GAAAGGATTG	TCAGACAGAG	TCGTGTTCGT	CCTTTGGGCG	CATGGCTTTG
18541	AGCTTACATC	AATGAAGTAC	TTTGTCAAGA	TTGGACCTGA	AAGAACGTGT	TGTCTGTGTG
18601	ACAAACGTGC	AACTTGCTTT	TCTACTTCAT	CAGATACTTA	TGCCTGCTGG	AATCATTCTG
18651	TGGGTTTTGA	CTATGTCTAT	AACCCATTTA	TGATTGATGT	TCAGCAGTGG	GGCTTTACGG
18721	GTAACCTTCA	GAGTAACCAT	GACCAACATT	GCCAGGTACA	TGGAAATGCA	CATGTGGCTA
18781	GTTGTGATGC	TATCATGACT	AGATGTTTAG	CAGTCCATGA	GTGCTTTGTT	AAGCGCGTTG
18841	ATTGGTCTGT	TGAATACCCT	ATTATAGGAG	ATGAACTGAG	GGTTAATTCT	GCTTGCAGAA
18901	AAGTACAACA	CATGGTTGTG	AAGTCTGCAT	TGCTTGCTGA	TAAGTTTCCA	GTTCTTCATG
18961	ACATTGGAAA	TCCAAAGGCT	ATCAAGTGTG	TGCCTCAGGC	TGAAGTAGAA	TGGAAGTTCT
19021	ACGATGCTCA	GCCATGTAGT	GACAAAGCTT	ACAAAATAGA	GGAACTCTTC	TATTCTTATG
19081	CTACACATCA	CGATAAATTC	ACTGATGGTG	TTTGTTTGTT	TTGGAATTGT	AACGTTGATC
19141	GTTACCCAGC	CAATGCAATT	GTGTGTAGGT	TTGACACAAG	AGTCTTGTCA	AACTTGAACT
19201	TACCAGGCTG	TGATGGTGGT	AGTTTGTATG	TGAATAAGCA	TGCATTCCAC	ACTCCAGCTT
19261	TCGATAAAAG	TGCATTTACT	AATTTAAAGC	AATTGCCTTT	CTTTTACTAT	TCTGATAGTC
19321	CTTGTGAGTC	TCATGGCAAA	CAAGTAGTGT	CGGATATTGA	TTATGTTCCA	CTCAAATCTG
19381	CTACGTGTAT	TACACGATGC	AATTTAGGTG	GTGCTGTTTG	CAGACACCAT	GCAAATGAGT
19441	ACCGACAGTA	CTTGGATGCA	TATAATATGA	TGATTTCTGC	TGGATTTAGC	CTATGGATTT
19501	ACAAACAATT	TGATACTTAT	AACCTGTGGA	ATACATTTAC	CAGGTTACAG	AGTTTAGAAA
19561	ATGTGGCTTA	TAATGTTGTT	AATAAAGGAC	ACTTTGATGG	ACACGCCGGC	GAAGCACCTG
19621	TTTCCATCAT	TAATAATGCT	GTTTACACAA	AGGTAGATGG	TATTGATGTG	GAGATCTTTG
19681	AAAATAAGAC	AACACTTCCT	GTTAATGTTG	CATTTGAGCT	TTGGGCTAAG	CGTAACATTA
19741	AACCAGTGCC	AGAGATTAAG	ATACTCAATA	ATTTGGGTGT	TGATATCGCT	GCTAATACTG
19801	TAATCTGGGA	CTACAAAAGA	GAAGCCCCAG	CACATGTATC	TACAATAGGT	GTCTGCACAA
19861	TGACTGACAT	TGCCAAGAAA	CCTACTGAGA	GTGCTTGTTC	TTCACTTACT	GTCTTGTTTG
19921	ATGGTAGAGT	GGAAGGACAG	GTAGACCTTT	TTAGAAACGC	CCGTAATGGT	GTTTTAATAA
19931	CAGAAGGTTC	AGTCAAAGGT	CTAACACCTT	CAAAGGGACC	AGCACAAGCT	AGCGTCAATG
20041	GAGTCACATT	AATTGGAGAA	TCAGTAAAAA	CACAGTTTAA	CTACTTTAAG	AAAGTAGACG
20101	GCATTATTCA	ACAGTTGCCT	GAAACCTACT	TTACTCAGAG	CAGAGACTTA	GAGGATTTTA
20161	AGCCCAGATC	ACAAATGGAA	ACTGACTTTC	TCGAGCTCGC	TATGGATGAA	TTCATACAGC
20221	GATATAAGCT	CGAGGGCTAT	GCCTTCGAAC	ACATCGTTTA	TGGAGATTTC	AGTCATGGAC
20281	AACTTGGCGG	TCTTCATTTA	ATGATAGGCT	TAGCCAAGCG	CTCACAAGAT	TCACCACTTA
20341	AATTAGAGGA	TTTTATCCCT	ATGGACAGCA	CAGTGAAAAA	TTACTTCATA	ACAGATGCGC
20401	AAACAGGTTC	ATCAAAATGT	GTGTGTTCTG	TGATTGATCT	TTTACTTGAT	GACTTTGTCG
20461	AGATAATAAA	GTCACAAGAT	TTGTCAGTGA	TTTCAAAAGT	GGTCAAGGTT	ACAATTGACT
20521	ATGCTGAAAT	TTCATTCATG	CTTTGGTGTA	AGGATGGACA	TGTTGAAACC	TTCTACCCAA
20581	AACTACAAGC	AAGTCAAGCG	TGGCAACCAG	GTGTTGCGAT	GCCTAACTTG	TACAAGATGC
20641	AAAGAATGCT	TCTTGAAAAG	TGTGACCTTC	AGAATTATGG	TGAAAATGCT	GTTATACCAA
20701	AAGGAATAAT	GATGAATGTC	GCAAAGTATA	CTCAACTGTG	TCAATACTTA	AATACACTTA

Figure 6-G

20761 CTTTAGCTGT ACCCTACAAC ATGAGAGTTA TTCACTTTGG TGCTGGCTCT GATAAAGGAG
20821 TTGCACCAGG TACAGCTGTG CTCAGACAAT GGTTGCCAAC TGGCACACTA CTTGTCGATT
20881 CAGATCTTAA TGACTTCGTC TCCGACGCAG ATTCTACTTT AATTGGAGAC TGTGCAACAG
20941 TACATACGGC TAATAAATGG GACCTTATTA TTAGCGATAT GTATGACCCT AGGACCAAAC
21001 ATGTGACAAA AGAGAATGAC TCTAAAGAAG GGTTTTTCAC TTATCTGTGT GGATTTATAA
21061 AGCAAAAACT AGCCCTGGGT GGTTCTATAG CTGTAAAGAT AACAGAGCAT TCTTGGAATG
21121 CTGACCTTTA CAAGCTTATG GGCCATTTCT CATGGTGGAC AGCTTTTGTT ACAAATGTAA
21181 ATGCATCATC ATCGGAAGCA TTTTTAATTG GGGCTAACTA TCTTGGCAAG CCGAAGGAAC
21241 AAATTGATGG CTATACCATG CATGCTAACT ACATTTTCTG GAGGAACACA AATCCTATCC
21301 AGTTGTCTTC CTATTCACTC TTTGACATGA GCAAATTTCC TCTTAAATTA AGAGGAACTG
21361 CTGTAATGTC TCTTAAGGAG AATCAAATCA ATGATATGAT TTATTCTCTT CTGGAAAAAG
21421 GTAGGCTTAT CATTAGAGAA AACAACAGAG TTGTGGTTTC AAGTGATATT CTTGTTAACA
21481 ACTAAACGAA CATGTTTATT TTCTTATTAT TTCTTACTCT CACTAGTGGT AGTGACCTTG
21541 ACCGGTGCAC CACTTTTGAT GATGTTCAAG CTCCTAATTA CACTCAACAT ACTTCATCTA
21601 TGAGGGGGGT TTACTATCCT GATGAAATTT TTAGATCAGA CACTCTTTAT TTAACTCAGG
21661 ATTTATTTCT TCCATTTTAT TCTAATGTTA CAGGGTTTCA TACTATTAAT CATACGTTTG
21721 GCAACCCTGT CATACCTTTT AAGGATGGTA TTTATTTTGC TGCCACAGAG AAATCAAATG
21781 TTGTCCGTGG TTGGGTTTTT GGTTCTACCA TGAACAACAA GTCACAGTCG GTGATTATTA
21841 TTAACAATTC TACTAATGTT GTTATACGAG CATGTAACTT TGAATTGTGT GACAACCCTT
21901 TCTTTGCTGT TTCTAAACCC ATGGGTACAC AGACACATAC TATGATATTC GATAATGCAT
21951 TTAATTGCAC TTTCGAGTAC ATATCTGATG CCTTTTCGCT TGATGTTTCA GAAAAGTCAG
22021 GTAATTTTAA ACACTTACGA GAGTTTGTGT TTAAAAATAA AGATGGGTTT CTCTATGTTT
22081 ATAAGGGCTA TCAACCTATA GATGTAGTTC GTGATCTACC TTCTGGTTTT AACACTTTGA
22141 AACCTATTTT TAAGTTGCCT CTTGGTATTA ACATTACAAA TTTTAGAGCC ATTCTTACAG
22201 CCTTTTCACC TGCTCAAGAC ATTTGGGGCA CGTCAGCTGC AGCCTATTTT GTTGGCTATT
22261 TAAAGCCAAC TACATTTATG CTCAAGTATG ATGAAAATGG TACAATCACA GATGCTGTTG
22321 ATTGTTCTCA AAATCCACTT GCTGAACTCA AATGCTCTGT TAAGAGCTTT GAGATTGACA
22381 AAGGAATTTA CCAGACCTCT AATTTCAGGG TTGTTCCCTC AGGAGATGTT GTGAGATTCC
22441 CTAATATTAC AAACTTGTGT CCTTTTGGAG AGGTTTTTAA TGCTACTAAA TTCCCTTCTG
22501 TCTATGCATG GGAGAGAAAA AAAATTTCTA ATTGTGTTGC TGATTACTCT GTGCTCTACA
22561 ACTCAACATT TTTTTCAACC TTTAAGTGCT ATGGCGTTTC TGCCACTAAG TTGAATGATC
22621 TTTGCTTCTC CAATGTCTAT GCAGATTCTT TTGTAGTCAA GGGAGATGAT GTAAGACAAA
22681 TAGCGCCAGG ACAAACTGGT GTTATTGCTG ATTATAATTA TAAATTGCCA GATGATTTCA
22741 TGGGTTGTGT CCTTGCTTGG AATACTAGGA ACATTGATGC TACTTCAACT GGTAATTATA
22801 ATTATAAATA TAGGTATCTT AGACATGGCA AGCTTAGGCC CTTTGAGAGA GACATATCTA
22861 ATGTGCCTTT CTCCCCTGAT GGCAAACCTT GCACCCCACC TGCTCTTAAT TGTTATTGGC
22921 CATTAAATGA TTATGGTTTT TACACCACTA CTGGCATTGG CTACCAACCT TACAGAGTTG
22981 TAGTACTTTC TTTTGAACTT TTAAATGCAC CGGCCACGGT TTGTGGACCA AAATTATCCA
23041 CTGACCTTAT TAAGAACCAG TGTGTCAATT TTAATTTTAA TGGACTCACT GGTACTGGTG
23101 TGTTAACTCC TTCTTCAAAG AGATTTCAAC CATTTCAACA ATTTGGCCGT GATGTTTCTG
23161 ATTTCACTGA TTCCGTTCGA GATCCTAAAA CATCTGAAAT ATTAGACATT TCACCTTGCT
23221 CTTTTGGGGG TGTAAGTGTA ATTACACCTG GAACAAATGC TTCATCTGAA GTTGCTGTTC
23281 TATATCAAGA TGTTAACTGC ACTGATGTTT CTACAGCAAT TCATGCAGAT CAACTCACAC
23341 CAGCTTGGCG CATATATTCT ACTGGAAACA ATGTATTCCA GACTCAAGCA GGCTGTCTTA
23401 TAGGAGCTGA GCATGTCGAC ACTTCTTATG AGTGCGACAT TCCTATTGGA GCTGGCATTT
23461 GTGCTAGTTA CCATACAGTT TCTTTATTAC GTAGTACTAG CCAAAAATCT ATTGTGGCTT
23521 ATACTATGTC TTTAGGTGCT GATAGTTCAA TTGCTTACTC TAATAACACC ATTGCTATAC
23581 CTACTAACTT TTCAATTAGC ATTACTACAG AAGTAATGCC TGTTTCTATG GCTAAAACCT
23641 CCGTAGATTG TAATATGTAC ATCTGCGGAG ATTCTACTGA ATGTGCTAAT TTGCTTCTCC
23701 AATATGGTAG CTTTTGCACA CAACTAAATC GTGCACTCTC AGGTATTGCT GCTGAACAGG
23761 ATCGCAACAC ACGTGAAGTG TTCGCTCAAG TCAAACAAAT GTACAAAACC CCAACTTTGA
23821 AATATTTTGG TGGTTTTAAT TTTTCACAAA TATTACCTGA CCCTCTAAAG CCAACTAAGA
23881 GGTCTTTTAT TGAGGACTTG CTCTTTAATA AGGTGACACT CGCTGATGCT GGCTTCATGA
23941 AGCAATATGG CGAATGCCTA GGTGATATTA ATGCTAGAGA TCTCATTTGT GCGCAGAAGT
23001 TCAATGGACT TACAGTGTTG CCACCTCTGC TCACTGATGA TATGATTGCT GCCTACACTG
24061 CTGCTCTAGT TAGTGGTACT GCCACTGCTG GATGGACATT TGGTGCTGGC GCTGCTCTTC
24121 AAATACCTTT TGCTATGCAA ATGGCATATA GGTTCAATGG CATTGGAGTT ACCCAAAATG
24181 TTCTCTATGA GAACCAAAAA CAAATCGCCA ACCAATTTAA CAAGGCGATT AGTCAAATTC

Figure 6-H

24241	AAGAATCACT	TACAACAACA	TCAACTGCAT	TGGGCAAGCT	GCAAGACGTT	GTTAACCAGA
24301	ATGCTCAAGC	ATTAAACACA	CTTGTTAAAC	AACTTAGCTC	TAATTTTGGT	GCAATTTCAA
24361	GTGTGCTAAA	TGATATCCTT	TCGCGACTTG	ATAAAGTCGA	GGCGGAGGTA	CAAATTGACA
24421	GGTTAATTAC	AGGCAGACTT	CAAAGCCTTC	AAACCTATGT	AACACAACAA	CTAATCAGGG
24481	CTGCTGAAAT	CAGGGCTTCT	GCTAATCTTG	CTGCTACTAA	AATGTCTGAG	TGTGTTCTTG
24541	GACAATCAAA	AAGAGTTGAC	TTTTGTGGAA	AGGGCTACCA	CCTTATGTCC	TTCCCACAAG
24601	CAGCCCCGCA	TGGTGTTGTC	TTCCTACATG	TCACGTATGT	GCCATCCCAG	GAGAGGAACT
24661	TCACCACAGC	GCCAGCAATT	TGTCATGAAG	GCAAAGCATA	CTTCCCTCGT	GAAGGTGTTT
24721	TTGTGTTTAA	TGGCACTTCT	TGGTTTATTA	CACAGAGGAA	CTTCTTTTCT	CCACAAATAA
24781	TTACTACAGA	CAATACATTT	GTCTCAGGAA	ATTGTGATGT	CGTTATTGGC	ATCATTAACA
24841	ACACAGTTTA	TGATCCTCTG	CAACCTGAGC	TTGACTCATT	CAAAGAAGAG	CTGGACAAGT
24901	ACTTCAAAAA	TCATACATCA	CCAGATGTTG	ATCTTGGCGA	CATTTCAGGC	ATTAACGCTT
24961	CTGTCGTCAA	CATTCAAAAA	GAAATTGACC	GCCTCAATGA	GGTCGCTAAA	AATTTAAATG
25021	AATCACTCAT	TGACCTTCAA	GAATTGGGAA	AATATGAGCA	ATATATTAAA	TGGCCTTGGT
25081	ATGTTTGGCT	CGGCTTCATT	GCTGGACTAA	TTGCCATCGT	CATGGTTACA	ATCTTGCTTT
25141	GTTGCATGAC	TAGTTGTTGC	AGTTGCCTCA	AGGGTGCATG	CTCTTGTGGT	TCTTGCTGCA
25201	AGTTTGATGA	GGATGACTCT	GAGCCAGTTC	TCAAGGGTGT	CAAATTACAT	TACACATAAA
25261	CGAACTTATG	GATTTGTTTA	TGAGATTTTT	TACTCTTGGA	TCAATTACTG	CACAGCCAGT
25321	AAAAATTGAC	AATGCTTCTC	CTGCAAGTAC	TGTTCATGCT	ACAGCAACGA	TACCGCTACA
253B1	AGCCTCACTC	CCTTTCGGAT	GGCTTGTTAT	TGGCGTTGCA	TTTCTTGCTG	TTTTTCAGAG
25441	CGCTACCAAA	ATAATTGCGC	TCAATAAAAG	ATGGCAGCTA	GCCCTTTATA	AGGGCTTCCA
25501	GTTCATTTGC	AATTTACTGC	TGCTATTTGT	TACCATCTAT	TCACATCTTT	TGCTTGTCGC
25561	TGCAGGTAAG	GAGGCGCAAT	TTTTGTACCT	CTATGCCTTG	ATATATTTTC	TACAATGCAT
25621	CAACGCATGT	AGAATTATTA	TGAGATGTTG	GCTTTGTTGG	AAGTGCAAAT	CCAAGAACCC
25661	ATTACTTTAT	GATGCCAACT	ACTTTGTTTG	CTGGCACACA	CATAACTATG	ACTACTGTAT
25681	ACCATATAAC	AGTGTCACAG	ATACAATTGT	CGTTACTGAA	GGiGACGGCA	TTTCAACACC
25741	ACCATATAAC	AGTGTCACAG	ATACAATTGT	CGTTACTGAA	GGiGACGGCA	TTTCAACACC
25801	AAAACTCAAA	GAAGACTACC	AAATTGGTGG	TTATTCTGAG	GATAGGCACT	CAGGTGTTAA
25861	AGACTATGTC	GTTGTACATG	GCTATTTCAC	CGAAGTTTAC	TACCAGCTTG	AGTCTACACA
25921	AATTACTACA	GACACTGGTA	TTGAAAATGC	TACATTCTTC	ATCTTTAACA	AGCTTGTTAA
25981	AGACCCACCG	AATGTGCAAA	TACACACAAT	CGACGGCTCT	TCAGGAGTTG	CTAATCCAGC
26041	AATGGATCCA	ATTTATGATG	AGCCGACGAC	GACTACTAGC	GTGCCTTTGT	AAGCACAAGA
26101	AAGTGAGTAC	GAACTTATGT	ACTCATTCGT	TTCGGAAGAA	ACAGGTACGT	TAATAGTTAA
26161	TAGCGTACTT	CTTTTTCTTG	CTTTCGTGGT	ATTCTTGCTA	GTCACACTAG	CCATCCTTAC
26221	TGCGCTTCGA	TTGTGTGCGT	ACTGCTGCAA	TATTGTTAAC	GTGAGTTTAG	TAAAACCAAC
26281	GGTTTACGTC	TACTCGCGTG	TTAAAAATCT	GAACTCTTCT	GAAGGAGTTC	CTGATCTTCT
26341	GGTCTAAACG	AACTAACTAT	TATTATTATT	CTGTTTGGAA	CTTTAACATT	GCTTATCATG
26401	GCAGACAACG	GTACTATTAC	CGTTGAGGAG	CTTAAACAAC	TCCTGGAACA	ATGGAACCTA
26461	GTAATAGGTT	TCCTATTCCT	AGCCTGGATT	ATGTTACTAC	AATTTGCCTA	TTCTAATCGG
26521	AACAGGTTTT	TGTACATAAT	AAAGCTTGTT	TTCCTCTGGC	TCTTGTGGCC	AGTAACACTT
26581	GCTTGTTTTG	TGCTTGCTGT	TGTCTACAGA	ATTAATTGGG	TGACTGGCGG	GATTGCGATT
26641	GCAATGGCTT	GTATTGTAGG	CTTGATGTGG	CTTAGCTACT	TCGTTGCTTC	CTTCAGGCTG
26701	TTTGCTCGTA	CCCGCTCAAT	GTGGTCATTC	AACCCAGAAA	CAAACATTCT	TCTCAATGTG
26761	CCTCTCCGGG	GGACAATTGT	GACCAGACCG	CTCATGGAAA	GTGAACTTGT	CATTGGTGCT
26821	GTGATCATTC	GTGGTCACTT	GCGAATGGCC	GGACACTCCC	TAGGGCGCTG	TGACATTAAG
26681	GACCTGCCAA	AAGAGATCAC	TGTGGCTACA	TCACGAACGC	TTTCTTATTA	CAAATTAGGA
26941	GCGTCGCAGC	GTGTAGGCAC	TGATTCAGGT	TTTGCTGCAT	ACAACCGCTA	CCGTATTGGA
27001	AACTATAAAT	TAAATACAGA	CCACGCCGGT	AGCAACGACA	ATATTGCTTT	GCTAGTACAG
27061	TAAGTGACAA	CAGATGTTTC	ATCTTGTTGA	CTTCCAGGTT	ACAATAGCAG	AGATATTGAT
27121	TATCATTATG	AGGACTTTCA	GGATTGCTAT	TTGGAATCTT	GACGTTATAA	TAAGTTCAAT
27181	AGTGAGACAA	TTATTTAAGC	CTCTAACTAA	GAAGAATTAT	TCGGAGTTAG	ATGATGAAGA
27241	ACCTATGGAG	TTAGATTATC	CATAAAACGA	ACATGAAAAT	TATTCTCTTC	CTGACATTGA
27301	TTGTATTTAC	ATCTTGCGAG	CTATATCACT	ATCAGGAGTG	TGTTAGAGGT	ACGACTGTAC
27361	TACTAAAAGA	ACCTTGCCCA	TCAGGAACAT	ACGAGGGCAA	TTCACCATTT	CACCCTCTTG
27421	CTGACAATAA	ATTTGCACTA	ACTTGCACTA	GCACACACTT	TGCTTTTGCT	TGTGCTGACG
27461	GTACTCGACA	TACCTATCAG	CTGCGTGCAA	GATCAGTTTC	ACCAAAACTT	TTCATCAGAC
27541	AAGAGGAGGT	TCAACAAGAG	CTCTACTCGC	CACTTTTTCT	CATTGTTGCT	GCTCTAGTAT
27601	TTTTAATACT	TTGCTTCACC	ATTAAGAGAA	AGACAGAATG	AATGAGCTCA	CTTTAATTGA
27661	CTTCTATTTG	TGCTTTTTAG	CCTTTCTGCT	ATTCCTTGTT	TTAATAATGC	TTATTATATT

Figure 6-I

```
27721 TTGGTTTTCA CTCGAAATCC AGGATCTAGA AGAACCTTGT ACCAAAGTCT AAACGAACAT
27781 GAAACTTCTC ATTGTTTTGA CTTGTATTTC TCTATGCAGT TGCATATGCA CTGTAGTACA
27B41 GCGCTGTGCA TCTAATAAAC CTCATGTGCT TGAAGATCCT TGTAAGGTAC AACACTAGGG
27901 GTAATACTTA TAGCACTGCT TGGCTTTGTG CTCTAGGAAA GGTTTTACCT TTTCATAGAT
27961 GGCACACTAT GGTTCAAACA TGCACACCTA ATGTTACTAT CAACTGTCAA GATCCAGCTG
28021 GTGGTGCGCT TATAGCTAGG TGTTGGTACC TTCATGAAGG TCACCAAACT GCTGCATTTA
26081 GAGACGTACT TGTTGTTTTA AATAAACGAA CAAATTAAAA TGTCTGATAA TGGACCCCAA
28141 TCAAACCAAC GTAGTGCCCC CCGCATTACA TTTGGTGGAC CCACAGATTC AACTGACAAT
28201 AACCAGAATG GAGGACGCAA TGGGGCAAGG CCAAAACAGC GCCGACCCCA AGGTTTACCC
28261 AATAATACTG CGTCTTGGTT CACAGCTCTC ACTCAGCATG GCAAGGAGGA ACTTAGATTC
28321 CCTCGAGGCC AGGGCGTTCC AATCAACACC AATAGTGGTC CAGATGACCA AATTGGCTAC
28381 TACCGAAGAG CTACCCGACG AGTTCGTGGT GGTGACGGCA AAATGAAAGA GCTCAGCCCC
28441 AGATGGTACT TCTATTACCT AGGAACTGGC CCAGAAGCTT CACTTCCCTA CGGCGCTAAC
28501 AAAGAAGGCA TCGTATGGGT TGCAACTGAG GGAGCCTTGA ATACACCCAA AGACCACATT
28561 GGCACCCGCA ATCCTAATAA CAATGCTGCC ACCGTGCTAC AACTTCCTCA AGGAACAACA
28621 TTGCCAAAAG GCTTCTACGC AGAGGGAAGC AGAGGCGGCA GTCAAGCCTC TTCTCGCTCC
28681 TCATCACGTA GTCGCGGTAA TTCAAGAAAT TCAACTCCTG GCAGCAGTAG GGGAAATTCT
28741 CCTGCTCGAA TGGCTAGCGG AGGTGGTGAA ACTGCCCTCG CGCTATTGCT GCTAGACAGA
28601 TTGAACCAGC TTGAGAGCAA AGTTTCTGGT AAAGGCCAAC AACAACAAGG CCAAACTGTC
28861 ACTAAGAAAT CTGCTGCTGA GGCATCTAAA AAGCCTCGCC AAAAACGTAC TGCCACAAAA
28921 CAGTACAACG TCACTCAAGC ATTTGGGAGA CGTGGTCCAG AACAAACCCA AGGAAATTTC
28981 GGGGACCAAG ACCTAATCAG ACAAGGAACT GATTACAAAC ATTGGCCGCA AATTGCACAA
29041 TTTGCTCCAA GTGCCTCTGC ATTCTTTGGA ATGTCACGCA TTGGCATGGA AGTCACACCT
29101 TCGGGAACAT GGCTGACTTA TCATGGAGCC ATTAAATTGG ATGACAAAGA TCCACAATTC
29161 AAAGACAACG TCATACTGCT GAACAAGCAC ATTGACGCAT ACAAAACATT CCCACCAACA
29221 GAGCCTAAAA AGGACAAAAA GAAAAAGACT GATGAAGCTC AGCCTTTGCC GCAGAGACAA
29281 AAGAAGCAGC CCACTGTGAC TCTTCTTCCT GCGGCTGACA TGGATGATTT CTCCAGACAA
29341 CTTCAAAATT CCATGAGTGG AGCTTCTGCT GATTCAACTC AGGCATAAAC ACTCATGATG
29401 ACCACACAAG GCAGATGGGC TATGTAAACG TTTTCGCAAT TCCGTTTACG ATACATAGTC
29461 TACTCTTGTG CAGAATGAAT TCTCGTAACT AAACAGCACA AGTAGGTTTA GTTAACTTTA
29521 ATCTCACATA GCAATCTTTA ATCAATGTGT AACATTAGGG AGGACTTGAA AGAGCCACCA
29581 CATTTTCATC GAGGCCACGC GGAGTACGAT CGAGGGTACA GTGAATAATG CTAGGGAGAG
29641 CTGCCTATAT GGAAGAGCCC TAATGTGTAA AATTAATTTT AGTAGTGCTA TCCCCATGTG
29701 ATTTTAATAG CTTCTTAGGA GAATGACAAA AAAAAAAAAA AA
```

Figure 7-A

pCV8

```
BstUI HhaI BstUI HinP1I HpyCH4III HphI HphI Hpy188I FatI MnlI TseI NspI NlaIII HpyCH4V Fnu4HI CviJI AluI BsmBI BsmAI StyD4I PfoI ScrFI NciI HpaII HpyCH4III BslI BbvI Tsp45I MaeIII CviJI AluI SfaNI AciI StyD4I HpaII ScrFI NciI BstF5I FokI CviJI HgaI DrdI

    tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccg
  1 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 100
    agcgcgcaaagccactactgccacttttggagactgtgtacgtcgagggcctctgccagtgtcgaacagacattcgcctacggccctcgtctgttcgggc

a    S  R  V  S  V  M  T  V  K  T  S  D  T  C  S  S  R  R  R  S  Q  L  V  C  K  R  M  P  G  A  D  K  P  V  -
b     R  A  F  R  *  *  R  *  K  P  L  T  H  A  A  P  G  D  G  H  S  L  S  V  S  G  C  R  E  Q  T  S  P  -
c      A  R  F  G  D  D  G  E  N  L  *  H  M  Q  L  P  E  T  V  T  A  C  L  *  A  D  A  G  S  R  Q  A  R  -

HinP1I FauI HhaI BstUI MspA1I AciI FauI AciI CviJI Cac8I CviJI MseI AciI Fnu4HI TauI Hpy188I SfaNI BseMII BspCNI TatI Csp6I RsaI DdeI ApaLI HpyCH4V Hpy8I Bsp1286I BsiHKAI Bme1580I NdeI MwoI BstAPI HpyF10VI AciI SfaNI

    tcagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaata
101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 200
    agtcccgcgcagtcgcccacaaccgcccacagccccgaccgaattgatacgccgtagtctcgtctaacatgactctcacgtggtatacgccacactttat

a     R  A  R  Q  R  V  L  A  G  V  G  A  G  L  T  M  R  H  Q  S  R  L  Y  *  E  C  T  I  C  G  V  K  Y  -
b    S  G  R  V  S  G  C  W  R  V  S  G  L  A  *  L  C  G  I  R  A  D  C  T  E  S  A  P  Y  A  V  *  N  T  -
c    Q  G  A  S  A  G  V  G  G  C  R  G  W  L  N  Y  A  A  S  E  Q  I  V  L  R  V  H  H  M  R  C  E  I  -
```

Figure 7-B

```
AciI  AciI  MwoI KasI BanI NarI HpyF10VI HinP1I BsaHI SfoI NlaIV SfaNI HhaI HaeII BbeI BbvI MwoI HpyF10VI
MwoI BglI HpyF10VI TseI CviJI Fnu4HI HinP1I FspI HhaI HpyCH4III
MboI FauI DpnI PvuI BstKTI BsiEI MwoI HpyF10VI MboII AciI Sau96I Cac8I HaeIII CviJI
EarI MnlI

    ccgcacagatgcgtaaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctat
201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 300
    ggcgtgtctacgcattcctcttttatggcgtagtccgcggtaagcggtaagtccgacgcgttgacaacccttcccgctagccacgcccggagaagcgata

a    R  T  D  A  *  G  E  N  T  A  S  G  A  I  R  H  S  G  C  A  T  V  G  K  G  D  R  C  G  P  L  R  Y  -
b     A  Q  M  R  K  E  K  I  P  H  Q  A  P  F  A  I  Q  A  A  Q  L  L  G  R  A  I  G  A  G  L  F  A  I  -
c   P  H  R  C  V  R  R  K  Y  R  I  R  R  H  S  P  F  R  L  R  N  C  W  E  G  R  S  V  R  A  S  S  L  L  -

Cac8I AluI CviJI MspA1I PvuII Cac8I  BbvI
HpyCH4V Fnu4HI BstF5I TseI  FokI MseI MaeIII
BstNI ScrFI BsaJI PspGI StyD4I  BmrI
Hpy188III MaeIII Tsp45I BsrI  HpyCH4IV Hpy99I TaiI
BsrI CviJI HaeIII Hpy99I EaeI  HindIII TspRI

    tacgccagctggcgaaagggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgccaa
301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 400
    atgcggtcgaccgctttccccctacacgacgttccgctaattcaacccattgcggtcccaaaagggtcagtgctgcaacattttgctgccggtcacggtt

a    Y  A  S  W  R  K  G  D  V  L  Q  G  D  *  V  G  *  R  Q  G  F  P  S  H  D  V  V  K  R  R  P  V  P  S  -
b     T  P  A  G  E  R  G  M  C  C  K  A  I  K  L  G  N  A  R  V  F  P  V  T  T  L  *  N  D  G  Q  C  Q  -
c      R  Q  L  A  K  G  G  C  A  A  R  R  L  S  W  V  T  P  G  F  S  Q  S  R  R  C  K  T  T  A  S  A  K  -
```

Figure 7 C

```
                                                 BstF5I       BslI
  BceAI                                            RsaI|  Tsp509I |
  AluI |       BsmFI                              Csp6I||   FokI  | |         HpyCH4III
 CviJI |   Tsp509I |                       BslI   TatI|||MseI |   | |      BceAI     | CviJI
     | |         | |                          |       ||||   |  |   | |          |     |    |
     gcttgaaagagataaattgaacaagtatggtcgtcccctgttgggatgtactattaaacctaaattggggttatctgctaaaaactacggtagagccgtt
 401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 500
     cgaactttctctatttaacttgttcataccagcaggggacaaccctacatgataatttggatttaaccccaatagacgattttgatgccatctcggcaa

a     L  K  E  I  N  *  T  S  M  V  V  P  C  W  D  V  L  L  N  L  N  W  G  Y  L  L  K  T  T  V  E  P  F  -
b    A  *  K  R  *  I  E  Q  V  W  S  S  P  V  G  M  Y  Y  *  T  *  I  G  V  I  C  *  K  L  R  *  S  R  L -
c     L  E  R  D  K  L  N  K  Y  G  R  P  L  L  G  C  T  I  K  P  K  L  G  L  S  A  K  N  Y  G  R  A  V  -

                  Hpy8I                         MmeI
                TspDTI |                       Hpy8I|          PpiI      BstKTI
 BbsI            AciI | |                      TaiI||         XcmI|      DpnI|
 MboII    XmnI BstUI | |            PpiI  HpyCH4IV |||      BsaXI ||     MboI ||          BsaXI
     |       |     | | |               |        |  |||          | ||        | ||             |
     tatgaatgtcttcgcggtggacttgattttactaaagatgatgagaacgtgaactcacaaccatttatgcgttggagagatcgtttcttatttgtgccg
 501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 600
     atacttacagaagcgccacctgaactaaaatgatttctactactcttgcacttgagtgttggtaaatacgcaacctctctagcaaagaataaaacacggc

a     M  N  V  F  A  V  D  L  I  L  L  K  M  M  R  T  *  T  H  N  H  L  C  V  G  E  I  V  S  Y  F  V  P  -
b      *  M  S  S  R  W  T  *  F  Y  *  R  *  *  E  R  E  L  T  T  I  Y  A  L  E  R  S  F  L  I  L  C  R -
c    Y  E  C  L  R  G  G  L  D  F  T  K  D  D  E  N  V  N  S  Q  P  F  M  R  W  R  D  R  F  L  F  C  A  E -

                                                                  NlaIII
                                                                FatI |
                                                               RsaI| |
                                                              Csp6I|| |         MboII
                                                              PstI ||| |       BstKTI |
                                                         HpyCH4V | ||| |        DpnI| |      HpyF10VI
                                                           SfcI | | ||| |       BclI|| |  AluI       |
           PsiI      CviJI                  HphI       BspMI BsmI | | | |||NspI  MboI|| | CviJI  MwoI|
              |          |                     |           |    | | | | |||   |     ||| |    |      ||
     aagcactttataaagcacaggctgaaacaggtgaaatcaaagggcattacttgaatgctactgcaggtacatgcgaagaaatgatcaaaagagctgtatt
 601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 700
     ttcgtgaaatatttcgtgtccgactttgtccactttagtttcccgtaatgaacttacgatgacgtccatgtacgcttctttactagttttctcgacataa

a    K  H  F  I  K  H  R  L  K  Q  V  K  S  K  G  I  T  *  M  L  L  Q  V  H  A  K  K  *  S  K  E  L  Y  L -
b     S  T  L  *  S  T  G  *  N  R  *  N  Q  R  A  L  L  E  C  Y  C  R  Y  M  R  R  N  D  Q  K  S  C  I  -
c      A  L  Y  K  A  Q  A  E  T  G  E  I  K  G  H  Y  L  N  A  T  A  G  T  C  E  E  M  I  K  R  A  V  F  -
```

Figure 7-D

```
                          BsiEI
                          BstKTI    NlaIII                                                              Hpy8I
                            PvuI    NsiI |                                                              AccI|
     BcgI                   DpnI| BfrBI | |         HphI                  AluI                          FokI ||
  Tsp509I |         Hpy188I || FatI | |      BcgI |HinfI                 CviJI                         MboII ||
 BfaI   | |            MboI ||HpyCH4V | |    MseI | | TfiI  AciI      BfaI |CviJI                          | ||
   |    | |               | ||       | | |     |  | |    |     |         |  |     |                         | ||
   tgctagagaattgggcgttccgatcgtaatgcatgactacttaacgggggattcaccgcaaatactagcttggctcattattgccgagataatggtcta
701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 800
   acgatctcttaacccgcaaggctagcattacgtactgatgaattgccccctaagtggcgtttatgatcgaaccgagtaataacggctctattaccagat

a     L  E  N  W  A  F  R  S  *  C  M  T  T  *  R  G  D  S  P  Q  I  L  A  W  L  I  I  A  E  I  M  V  Y  -
b    C  *  R  I  G  R  S  D  R  N  A  *  L  L  N  G  G  I  H  R  K  Y  *  L  G  S  L  L  P  R  *  W  S  T -
c     A  R  E  L  G  V  P  I  V  M  H  D  Y  L  T  G  G  F  T  A  N  T  S  L  A  H  Y  C  R  D  N  G  L  -

                             AciI
                           NlaIII
                             NspI
                             SphI
                          BsrDI |
                           Cac8I |
                        HpyF10VI |
                            NsiI |
                         BstAPI| |
                           MwoI| |
                          BfrBI|| |
                           FatI|| |                                MboII                    TaiI
                        HpyCH4V|| |                                                 BciVI
                                                                  NlaIII |          NciI     BsaAI |
                  HpyCH4V   ||| |                                FatI  | |          ScrFI    SnaBI |
              HpyCH4III |   ||| |                             HinfI |  | |          BslI|  HpyCH4IV| |
       BstF5I        |  |   ||| |                              TfiI |  | |       HpaII|   MaeIII || |
           |         |  |   ||| |                                |  |  | |     StyD4I||      |   || |
   cttcttcacatccaccgtgcaatgcatgcggttattgatagacagaagaatcatggtatccacttccgggtattagcaaaagcgttacgtatgtctggtg
801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 900
   gaagaagtgtaggtggcacgttacgtacgccaataactatctgtcttcttagtaccataggtgaaggcccataatcgttttcgcaatgcatacagaccac

a     F  F  T  S  T  V  Q  C  M  R  L  L  I  D  R  R  I  M  V  S  T  S  G  Y  *  Q  K  R  Y  V  C  L  V  -
b      S  S  H  P  P  C  N  A  C  G  Y  *  *  T  E  E  S  W  Y  P  L  P  G  I  S  K  S  V  T  Y  V  W  W -
c    L  L  H  I  H  R  A  M  H  A  V  I  D  R  Q  K  N  H  G  I  H  F  R  V  L  A  K  A  L  R  M  S  G  G-
```

Figure 7-E

```
                    HpyCH4III
                         KpnI|
                       NlaIV ||
     BstKTI              RsaI ||
      DpnI|             Csp6I| ||
     BaeI ||          Acc65I|| ||      BaeI     BaeI
     MboI || BaeI       BanI|| ||   Hpy8I |     BsmAI      HphI      CviJI
        | ||  |           ||| ||       | |        |          |         |
    gagatcatattcactctggtaccgtagtaggtaaacttgaaggtgaaagagacataactttgggctttgttgatttactgcgtgatgatttgttgaaca
901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1000
    ctctagtataagtgagaccatggcatcatccatttgaacttccactttctctgtattgaaacccgaaacaactaaatgacgcactactaaaacaacttgt

a   E  I  I  F  T  L  V  P  *  *  V  N  L  K  V  K  E  T  *  L  W  A  L  L  I  Y  C  V  M  I  L  L  N  K -
b    R  S  Y  S  L  W  Y  R  S  R  *  T  *  R  *  K  R  H  N  F  G  L  C  *  F  T  A  *  *  F  C  *  T  -
c     D  H  I  H  S  G  T  V  V  G  K  L  E  G  E  R  D  I  T  L  G  F  V  D  L  L  R  D  D  F  V  E  Q -

                                                       Hin4I
                                                   Eco57I  |
                                                  Eco57MI  |
                                                 BstNI  |  |
                                                ScrFI   |  |
                                              PspGI |   |  |
                                              SexAI |   |  |
     BstKTI                                   StyD4I|   |  |      Hpy188III
       TaqI    AciI        Hpy188III           BsaI  |  |  |       CviJI  |           TaiI
      DpnI|    BpuEI             SmlI |        BsmAI |  |  |  |BsaJI  |   |    HpyCH4IV |  NdeI Cac8I
     MboI ||   BstUI      Hin4I |    |            |  |  |  |  |BtgI MnlI  |           |   |   |     |
        | ||      |           |    | |            |  |  |  |  |   |    |  |           |   |   |     |
     agatcgaagtcgcggtatttatttcactcaagattgggtctctttaccaggtgttctacccgtggcttcaggaggtattcacgtttggcatatgcctgct
1001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1100
     tctagcttcagcgccataaataaagtgagttctaacccagagaaatggtccacaagatgggcaccgaagtcctccataagtgcaaaccgtatacggacga

a     I  E  V  A  V  F  I  S  L  K  I  G  S  L  Y  Q  V  F  Y  P  W  L  Q  E  V  F  T  F  G  I  C  L  L -
b    R  S  K  S  R  Y  L  F  H  S  R  L  G  L  F  T  R  C  S  T  R  G  F  R  R  Y  S  R  L  A  Y  A  C  S -
c     D  R  S  R  G  I  Y  F  T  Q  D  W  V  S  L  P  G  V  L  P  V  A  S  G  G  I  H  V  W  H  M  P  A  -
```

Figure 7-F

```
Hpy188I MboI BstYI BglII DpnI BstKTI TspGWI TaqII TfiI HinfI BstF5I Csp6I RsaI SfcI FokI HpyCH4III MnlI FokI BstF5I StyI BsaJI BceAI HinP1I StyD4I PspGI HhaI ScrFI BstNI MwoI BanI HpyF10VI NlaIV CviJI AluI TaqI CviJI AluI

     ctgaccgagatctttggggatgattccgtactacagttcggtggaggaactttaggacatccttggggtaatgcgccaggtgccgtagctaatcgagtag
1101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1200
     gactggctctagaaacccctactaaggcatgatgtcaagccacctccttgaaatcctgtaggaaccccattacgcggtccacggcatcgattagctcatc

a     * P R S L G M I P Y Y S S V E E L * D I L G V M R Q V P * L I E * -
b     D R D L W G * F R T T V R W R N F R T S L G * C A R C R S * S S S -
c    L T E I F G D D S V L Q F G G G T L G H P W G N A P G A V A N R V A -
```

Figure 7-G

```
Hpy188III XbaI BfaI MwoI HpyF10VI FatI NlaIII NspI AluI CviJI
HpyCH4IV BmgBI TaiI MboI DpnI BstKTI TspDTI DdeI Bpu10I Hpy188III
BspCNI BseMII Tsp509I XmnI MnlI Hpy188III NruI BstUI TspDTI CviJI MwoI HpyF10VI Cac8I HpyCH4V
MwoI HpyF10VI BbvI NlaIV StyD4I CviJI BbvI ScrFI NciI HpaII Bsp1286I BanII
BfaI TseI CviJI AluI Fnu4HI

     ctctagaagcatgtgtaaaagctcgtaatgaaggacgtgatcttgctcaggaaggtaatgaaattattcgcgaggcttgcaaatggagcccggaactagc
1201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1300
     gagatcttcgtacacattttcgagcattacttcctgcactagaacgagtccttccattactttaataagcgctccgaacgtttacctcgggccttgatcg

a    L  *  K  H  V  *  K  L  V  M  K  D  V  I  L  L  R  K  V  M  K  L  F  A  R  L  A  N  G  A  R  N  *  L  -
b     S  R  S  M  C  K  S  S  *  *  R  T  *  S  C  S  G  R  *  *  N  Y  S  R  G  L  Q  M  E  P  G  T  S   -
c      L  E  A  C  V  K  A  R  N  E  G  R  D  L  A  Q  E  G  N  E  I  I  R  E  A  C  K  W  S  P  E  L  A  -
```

Figure 7-G

```
                                                  TaiI
                                                 BbvI|
                                 Fnu4HI      BtsI ||
                        BstKTI HpyCH4V| HpyCH4IV ||                   Hpy8I
   Fnu4HI                DpnI|Tsp509I  ||  TspRI ||                   AccI|
    TseI|               MboI || MseI| TseI| Hpy8I | ||          HpyCH4III ||       Tsp509I
     tgctgcttgtgaagtatggaaagagatcgtatttaattttgcagcagtggacgttttggataagtaaaaacagtagacattagcagataaattagcagga
1301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1400
     acgacgaacacttcatacctttctctagcataaattaaaacgtcgtcacctgcaaaacctattcatttttgtcatctgtaatcgtctatttaatcgtcct

a      L  L  V  K  Y  G  K  R  S  Y  L  I  L  Q  Q  W  T  F  W  I  S  K  N  S  R  H  *  Q  I  N  *  Q  E  -
b     C  C  L  *  S  M  E  R  D  R  I  *  F  C  S  S  G  R  F  G  *  V  K  T  V  D  I  S  R  *  I  S  R  K -
c      A  A  C  E  V  W  K  E  I  V  F  N  F  A  A  V  D  V  L  D  K  *  K  Q  *  T  L  A  D  K  L  A  G  -

                                                          MseI
                                                      Tsp509I |
                                                     HpyCH4V| |
                                                Tsp509I     || |   CviJI
                             Tsp509I       Tsp509I  |       || |   HaeIII
           BpuEI         SmlI   |            MseI|  |       || |  Sau96I|                         CviJI
     aataaagaaggataaggagaaagaactcaagtaattatccttcgttctcttaattgaattgcaattaaactcggcccaatcttttactaaaaggattgag
1401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1500
     ttatttcttcctattcctctttcttgagttcattaataggaagcaagagaattaacttaacgttaatttgagccgggttagaaaatgattttcctaactc

a      I  K  K  D  K  E  K  E  L  K  *  L  S  F  V  L  L  I  E  L  Q  L  N  S  A  Q  S  F  T  K  R  I  E  -
b      *  R  R  I  R  R  K  N  S  S  N  Y  P  S  F  S  *  L  N  C  N  *  T  R  P  N  L  L  L  K  G  L  S  -
c     N  K  E  G  *  G  E  R  T  Q  V  I  I  L  R  S  L  N  *  I  A  I  K  L  G  P  I  F  Y  *  K  D  *  A -

                                                                        BfaI
           HinfI                                                   Hpy188III                  MlyI
            TfiI  HpyCH4V         DdeI            BfaI                  XbaI|                  PleI
     ccgaatacaacaaagattctattgcatatattttgactaagtatatacttacctagatatacaagatttgaaatacaaaatctagaaaactaaatcaaaa
1501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1600
     ggcttatgttgtttctaagataacgtatataaaactgattcatatatgaatggatctatatgttctaaactttatgttttagatcttttgatttagtttt

a     P  N  T  T  K  I  L  L  H  I  F  *  L  S  I  Y  L  P  R  Y  T  R  F  E  I  Q  N  L  E  N  *  I  K  I -
b      R  I  Q  Q  R  F  Y  C  I  Y  F  D  *  V  Y  T  Y  L  D  I  Q  D  L  K  Y  K  I  *  K  T  K  S  K  -
c       E  Y  N  K  D  S  I  A  Y  I  L  T  K  Y  I  L  T  *  I  Y  K  I  *  N  T  K  S  R  K  L  N  Q  N  -
```

Figure 7-H

```
Enzymes (1601-1700): DdeI, HinfI, MboI, DpnI, BstKTI, TaqI, AlwI, CviJI, AluI, Cac8I, HpyCH4V, FatI, Cac8I, SphI, NspI, NlaIII, SfcI, Cac8I, HpyCH4V, PstI, SalI, TaqI, AccI, Hpy8I, HincII, HpyCH4III, Hpy188I, TatI, Csp6I, RsaI

     tctaagactcaaatctttctattgttgtcttggatcatcgagcttgcatgcctgcagtcgactttcacagtttccattctgaaatgttctctgtactata
1601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1700
     agattctgagtttagaaagataacaacagaacctagtagctcgaacgtacggacgtcagctgaaagtgtcaaaggtaagactttacaagagacatgatat

a     * D S N L S I V V L D H R A C M P A V D F H S F H S E M F S V L * -
b    S K T Q I F L L L S W I I E L A C L Q S T F T V S I L K C S L Y Y N -
c     L R L K S F Y C C L G S S S L H A C S R L S Q F P F * N V L C T I -

Enzymes (1701-1800): HinfI, TfiI, Tsp509I, AccI, MboII, Hpy8I, Hpy188I, EarI, MnlI, CviJI, HpyCH4V

     ataaatagtaagtgaatcaacttactactaaaaaaattagtagacttcctcttcggaatagaaatagcctatttctacatagggaaagtcgtgtgcaatg
1701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1800
     tatttatcattcacttagttgaatgatgatttttttaatcatctgaaggagaagccttatctttatcggataaagatgtatccctttcagcacacgttac

a     * I V S E S T Y Y * K N * * T S S S E * K * P I S T * G K S C A M -
b     K * * V N Q L T T K K I S R L P L R N R N S L F L H R E S R V Q * -
c    I N S K * I N L L L K K L V D F L F G I E I A Y F Y I G K V V C N E -
```

Figure 7-I

```
                                                                                                                          MlyI
                                                                                                                         PleI|
                                                                                                                       MmeI ||
                                                                                                                       NciI ||
                                                                                                                      ScrFI ||
                                                                                                                       SmaI ||
                                                                               NciI                                  HpaII| ||
                                                                              ScrFI                                   NciI| ||
                                                                             HpaII|                                  ScrFI| ||
                                                                            StyD4I||                                  AvaI|| ||
                                                                         BsmFI   |||                                 BsaJI|| ||
                                                                         BccI|   |||                                StyD4I|| ||
                      TspDTI                                             BfaI|   |||                                  XmaI|| ||
                   Cac8I   |                                             SpeI||  |||                               StyD4I||| ||
            HpyCH4V   |    |                      Tsp509I          Hpy188I |||  |||                            HinfI  |||| ||
        BsrDI     |   |    | MnlI          MaeIII  FokI|      BstF5I     | |||  |||                           TaqI |  |||| ||
            |     |   |    |    |                |       ||           |     | |||  |||                            | |  |||| ||
        aaaaatgcaagcacgatttggggagaggtttttctctattgtaacaaggaataattatctactccatccgactagttccgggttcgagtcccgggcaac
   1801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1900
        tttttacgttcgtgctaaacccctctccaaaaagagataacattgttccttattaatagatgaggtaggctgatcaaggcccaagctcagggcccgttg

a       K  N  A  S  T  I  W  G  E  V  F  S  L  L  *  Q  G  I  I  I  Y  S  I  R  L  V  P  G  S  S  P  G  Q  P -
b        K  M  Q  A  R  F  G  E  R  F  F  L  Y  C  N  K  E  *  L  S  T  P  S  D  *  F  R  V  R  V  P  G  N  -
c         K  C  K  H  D  L  G  R  G  F  F  S  I  V  T  R  N  N  Y  L  L  H  P  T  S  S  G  F  E  S  R  A  T  -

                     BspCNI
                    BseMII|   Hpy188I                               ApoI                    ApoI
          NdeI    BfaI    ||     DdeI|        TspDTI              Tsp509I                 Tsp509I
             |       |    ||        ||             |                    |                       |
        ccatatggaaactagaaaggagcaatctgagttttgatttttcactcacttcatttacaaaatttttggtttggtaaatttttgttgtatggatatacaa
   1901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2000
        ggtatacctttgatctttcctcgttagactcaaaactaaaaagtgagtgaagtaaatgttttaaaaaccaaaccatttaaaacaacatacctatatgtt

a         I  W  K  L  E  R  S  N  L  S  F  D  F  S  L  T  S  F  T  K  F  F  G  L  V  N  F  V  V  W  I  Y  N  -
b        P  Y  G  N  *  K  G  A  I  *  V  L  I  F  H  S  L  H  L  Q  N  F  L  V  W  *  I  L  L  Y  G  Y  T  T -
c         H  M  E  T  R  K  E  Q  S  E  F  *  F  F  T  H  F  I  Y  K  I  F  W  F  G  K  F  C  C  M  D  I  Q  -
```

Figure 7-J

```
                                                                                    TaiI
                                                                                 BsaAI |
                                                                               AflIII| |
                                                                             HpyCH4IV| |
                  CviJI                                     MseI              MseI   || |
                Cac8I | HincII                                               BfaI |   || |      HinfI
    HpyCH4III  CviJI | |  Hpy8I                  HpyCH4III |      CviJI          |    |   || |        |
         |       | | |    |                            | |          |            |    |   || |        |
         ctgtcggggctggcttggttgacattggtatatagtctatattatactgttaaataacaagccttctattatctttctagttaatacgtgtgcttgggag
    2001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2100
         gacagccccgaccgaaccaactgtaaccatatatcagatataatatgacaatttattgttcggaagataatagaaagatcaattatgcacacgaaccctc

a         C  R  G  W  L  G  *  H  W  Y  I  V  Y  I  I  L  L  N  N  K  P  S  I  I  F  L  V  N  T  C  A  W  E  -
b          V  G  A  G  L  V  D  I  G  I  *  S  I  L  Y  C  *  I  T  S  L  L  L  S  F  *  L  I  R  V  L  G  S  -
c         L  S  G  L  A  W  L  T  L  V  Y  S  L  Y  Y  T  V  K  *  Q  A  F  Y  Y  L  S  S  *  Y  V  C  L  G  V -
```

Figure 7-K

```
                                                                                    ApoI
                                                                                    EcoRI
                                                                                    Tsp509I
                                                                                    AlwI
                                                                                    PstI
                                                                                    HpyCH4V
                                                                                    HpyF10VI
                                                                                    SfcI
                                                                                    AlwNI
                                                                                    BstAPI
                                                                                    BstKTI
                                                                                    MwoI
                                                                                    DpnI
                                                                                    BstYI
                                                                                    MboI
                                                                                    FauI
                                                                                    BmtI
                                                                                    BslI
                                                                                    Cac8I
                                                                                    BfaI
                                                                                    NheI
                                                                                    AciI
                                                                                    ApaI
                                                                                    BanII
                                                                                    Bme1580I
                                                                                    Bsp1286I
                                                                                    Cac8I
                                                                                    CviJI
                                                                                    HaeIII
                                                     AciI                           NlaIV
                                                     BstUI                          NlaIV
                                                     HhaI                           Sau96I
                                                     BstUI                          EcoO109I
                                                     HinP1I                         PspOMI
                                                     TauI                           Sau96I
                                                     AciI                           SacII
                                                     Fnu4HI                         AciI
                                                     CviJI                          BstUI
                                                     HaeIII                         MspA1I
                                                     NlaIII                         AciI
                                                     EaeI                           BsaJI
                                                     FauI                           BtgI
                                                     BsaJI                          AvaII
                                   BstKTI                                           NlaIII
      MlyI                         DpnI  BtgI                                       Sau96I
      Tsp509I                      BglII FatI                                       FauI
      HpyCH4V                      BstYI NcoI                                       FatI
      PleI                         MboI  StyI                                       MslI

    tccttgcaatttgaataaaccaagatcttaccatggccgcgcggggttctcatcatcatcatcatggtccgcgggggcccgctagcaggatctgcagaat
2101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2200
    aggaacgttaaacttatttggttctagaatggtaccggcgcgccccaagagtagtagtagtagtaccaggcgcccccgggcgatcgtcctagacgtctta

a    S  L  Q  F  E  *  T  K  I  L  P  W  P  R  G  V  L  I  I  I  I  M  V  R  G  G  P  L  A  G  S  A  E  F  -
b     P  C  N  L  N  K  P  R  S  Y  H  G  R  A  G  F  S  S  S  S  S  W  S  A  G  A  R  *  Q  D  L  Q  N  -
c      L  A  I  *  I  N  Q  D  L  T  M  A  A  R  G  S  H  H  H  H  H  G  P  R  G  P  A  S  R  I  C  R  I  -
```

Figure 7-L

```
                          AlwI
                        NlaIII
                      BsaJI
                       BtgI
                       FatI
                       NcoI
                       StyI
                    BstKTI
                     DpnI
                    NlaIV
                   BamHI
                   BstYI                                                     BstNI                                   BstNI
                                                                             ScrFI                                   ScrFI
                                                                              PfoI                                  CviJI
                                   RsaI                       MseI          PspGI                                   PspGI
                CviJI           Csp6I HpyCH4III               AluI         StyD4I            BfaI                  StyD4I
                AlwI MboI    HpyCH4III   MnlI                CviJI          BseRI          NlaIV                  BfaI
      tcggcttGGATCCATGgcagacaacggtactattaccgttgaggagcttaaacaactcctggaacaatggaacctagtaataggtttcctattcctagcc
 2201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2300
      agccgaaCCTAGGTACcgtctgttgccatgataatggcaactcctcgaatttgttgaggaccttgttaccttggatcattatccaaaggataaggatcgg

a       G  L  D  P  W  Q  T  T  V  L  L  P  L  R  S  L  N  N  S  W  N  N  G  T  *  *  *  V  S  Y  S  *  P  -
b      S  A  W  I  H  G  R  Q  R  Y  Y  Y  R  *  G  A  *  T  T  P  G  T  M  E  P  S  N  R  F  P  I  P  S  L -
c       R  L  G  S  M  A  D  N  G  T  I  T  V  E  E  L  K  Q  L  L  E  Q  W  N  L  V  I  G  F  L  F  L  A  -

                                                                                     MaeIII
                                                                                     BsrI
                                                                                   CviJI
                                                                                  HaeIII
                                                                                    MscI
                                                                                   EaeI
                                                     RsaI                       HpyF10VI
                                                    Csp6I         AluI              MwoI
                                                   BsrGI         CviJI              MnlI
        MaeIII  Tsp509I              Hpy188I        TatI      HindIII        CviJI
      tggattatgttactacaatttgcctattctaatcggaacaggtttttgtacataataaagcttgttttcctctggctcttgtggccagtaacacttgctt
 2301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2400
      acctaatacaatgatgttaaacggataagattagccttgtccaaaaacatgtattatttcgaacaaaaggagaccgagaacaccggtcattgtgaacgaa

a       G  L  C  Y  Y  N  L  P  I  L  I  G  T  G  F  C  T  *  *  S  L  F  S  S  G  S  C  G  Q  *  H  L  L  -
b        D  Y  V  T  T  I  C  L  F  *  S  E  Q  V  F  V  H  N  K  A  C  F  P  L  A  L  V  A  S  N  T  C  L -
c      W  I  M  L  L  Q  F  A  Y  S  N  R  N  R  F  L  Y  I  I  K  L  V  F  L  W  L  L  W  P  V  T  L  A  C -
```

Figure 7-M

```
                                  Tsp509I
                                     AseI|
                                     MseI|
                             Tsp509I    ||
                              TaqII |   ||
                          Hpy8I |   |   ||    FauI             HphI
                           SfcI |   |   ||MaeIII|      BsrI     |                  BsrDI
                Cac8I      AccI||   |   ||Tsp45I|      AciI|    |       HpyCH4V CviJI|
                  |         ||  |   |   ||      ||      ||      |           |      ||
        gttttgtgcttgctgttgtctacagaattaattgggtgactggcgggattgcgattgcaatggcttgtattgtaggcttgatgtggcttagctacttcgt
   2401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2500
        caaaacacgaacgacaacagatgtcttaattaacccactgaccgccctaacgctaacgttaccgaacataacatccgaactacaccgaatcgatgaagca

a       V  L  C  L  L  L  S  T  E  L  I  G  *  L  A  G  L  R  L  Q  W  L  V  L  *  A  *  C  G  L  A  T  S  L -
b        F  C  A  C  C  C  L  Q  N  *  L  G  D  W  R  D  C  D  C  N  G  L  Y  C  R  L  D  V  A  *  L  L  R  -
c         F  V  L  A  V  V  Y  R  I  N  W  V  T  G  G  I  A  I  A  M  A  C  I  V  G  L  M  W  L  S  Y  F  V -

 (right-hand labels, positions ~2460-2500: CviJI | ; Eco57I, Eco57MI, AluI|, CviJI|, BlpI ||, DdeI ||, CviJI| ||)

                                                                                   BslI
                                                                                   NciI      MaeIII
                             BsrBI                                                ScrFI      Tsp45I
                             AciI |                                               BsaJI|   MfeI   |
                          RsaI  | |                                               HpaII|Tsp509I   |
            CviJI        Csp6I| | |  FauI                     MboII               StyD4I||MnlI|   |   BsmFI
              |             ||  | |    |                        |                   |||   |   |   |     |
        tgcttccttcaggctgtttgctcgtacccgctcaatgtggtcattcaacccagaaacaaacattcttctcaatgtgcctctccggggacaattgtgacc
   2501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2600
        acgaaggaagtccgacaaacgagcatgggcgagttacaccagtaagttgggtctttgtttgtaagaagagttacacggagaggcccctgttaacactgg
                 <-- ML523
a        L  P  S  G  C  L  L  V  P  A  Q  C  G  H  S  T  Q  K  Q  T  F  F  S  M  C  L  S  G  G  Q  L  *  P  -
b       C  F  L  Q  A  V  C  S  Y  P  L  N  V  V  I  Q  P  R  N  K  H  S  S  Q  C  A  S  P  G  D  N  C  D  Q -
c        A  S  F  R  L  F  A  R  T  R  S  M  W  S  F  N  P  E  T  N  I  L  L  N  V  P  L  R  G  T  I  V  T  -
```

Figure 7-N

```
                                                                                          MaeIII
                                                                                          Tsp45I
                                                                                          HaeII|
                                                                                          HhaI||
         NlaIII                                                                   HinP1I |||        AvaII
       FatI |                     BstKTI                     HpaII          BfaI     | |||       EcoO109I
    BsrBI |  |                     DpnI|                    CviJI|        AvrII|     | |||          PpuMI
     XcmI |  |                    BclI ||    MaeIII        HaeIII|        BsaJI|     | |||         Sau96I
     AciI | | |     Hpy8I         MboI ||    Tsp45I         EaeI ||        StyI|     | |||       MseI   |
          | | |  |     |              | ||       |              | ||           ||     | |||          |   |
     agaccgctcatggaaagtgaacttgtcattggtgctgtgatcattcgtggtcacttgcgaatggccggacactccctagggcgctgtgacattaaggacc
2601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2700
     tctggcgagtacctttcacttgaacagtaaccacgacactagtaagcaccagtgaacgcttaccggcctgtgagggatcccgcgacactgtaattcctgg

a     D  R  S  W  K  V  N  L  S  L  V  L  *  S  F  V  V  T  C  E  W  P  D  T  P  *  G  A  V  T  L  R  T  -
b      T  A  H  G  K  *  T  C  H  W  C  C  D  H  S  W  S  L  A  N  G  R  T  L  P  R  A  L  *  H  *  G  P  -
c     R  P  L  M  E  S  E  L  V  I  G  A  V  I  I  R  G  H  L  R  M  A  G  H  S  L  G  R  C  D  I  K  D  L -

                 Hpy188III
              CviJI  |
             TspRI|  |
        HpyCH4III ||  |
         BstKTI | ||  |                                                     TspRI
           DpnI| | ||  |                                                  HinfI |
          MboI || | ||  |                        Fnu4HI                    TfiI |          HpyCH4V
    BspMI  | || | ||  |                Tsp509I  Hpy99I|                  BbvI| |          Fnu4HI |
          |  | || | ||  |      XmnI      HgaI |    TseI|               BbvI || |            TseI| |
          |  | || | ||  |        |          | |        ||                 | || |              || |
     tgccaaaagagatcactgtggctacatcacgaacgctttcttattacaaattaggagcgtcgcagcgtgtaggcactgattcaggttttgctgcatacaa
2701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2800
     acggttttctctagtgacaccgatgtagtgcttgcgaaagaataatgtttaatcctcgcagcgtcgcacatccgtgactaagtccaaaacgacgtatgtt

a     C  Q  K  R  S  L  W  L  H  H  E  R  F  L  I  T  N  *  E  R  R  S  V  *  A  L  I  Q  V  L  L  H  T  T -
b      A  K  R  D  H  C  G  Y  I  T  N  A  F  L  L  Q  I  R  S  V  A  A  C  R  H  *  F  R  F  C  C  I  Q  -
c       P  K  E  I  T  V  A  T  S  R  T  L  S  Y  Y  K  L  G  A  S  Q  R  V  G  T  D  S  G  F  A  A  Y  N  -
```

Figure 7-O

```
                                                                                  ApoI
                                                                                  EcoRI
                                                                                  Tsp509I
                                                                              BcgI
                                                                          HpyCH4III          BceAI
                                                                             RsaI        ApoI
                                              HpaII                         Csp6I        EcoRI
   HpyCH4III              MseI                BcgI                          TatI         Tsp509I
 AciI                 Tsp509I                 BsrFI            SspI       BfaI         CviJI
     ccgctaccgtattggaaactataaattaaatacagaccacgccggtagcaacgacaatattgctttgctagtacagtaaGAATTCaagccgaattccagc
2801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2900
     ggcgatggcataacctttgatatttaatttatgtctggtgcggccatcgttgctgttataacgaaacgatcatgtcattCTTAAGttcggcttaaggtcg

a      A  T  V  L  E  T  I  N  *  I  Q  T  T  P  V  A  T  T  I  L  L  C  *  Y  S  K  N  S  S  R  I  P  A  -
b     P  L  P  Y  W  K  L  *  I  K  Y  R  P  R  R  *  Q  R  Q  Y  C  F  A  S  T  V` R  I  Q  A  E  F  Q  H -
c      R  Y  R  I  G  N  Y  K  L  N  T  D  H  A  G  S  N  D  N  I  A  L  L  V  Q  *  E  F  K  P  N  S  S  -

                                   MseI
                                   PacI
                              Tsp509I
                                 AseI
                                 MseI
              AlwI            Tsp509I
              BfaI            BanII
              SpeI           BsiHKAI
            Hin4I            Bsp1286I
          MaeIII                SacI
          BsiEI                AluI
         CviJI                 AlwI
        HaeIII                CviJI
          TauI               EcoICRI
         BsrI             Hpy188I
         EaeI            BstKTI
         EagI             DpnI
       Fnu4HI            NlaIV                                                      BfaI
        TspRI           BamHI                                                       AluI
         AciI           BstYI                 CviJI                                CviJI
   BstXI                 MboI                 Hin4I               Tsp509I          MboII
     acactggcggccgttactagtggatccgagctctaattaattaaggcttttctgctaacatatagcaattttgaagaaaggaaagctagaaatacccaa
2901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3000
     tgtgaccgccggcaatgatcacctaggctcgagattaattaattccgaaaagacgattgtatatcgttaaaacttctttcctttcgatctttatgggtt

a      H  W  R  P  L  L  V  D  P  S  S  N  *  L  R  L  F  C  *  H  I  A  I  F  E  E  R  K  A  R  N  T  Q  -
b       T  G  G  R  Y  *  W  I  R  A  L  I  N  *  G  F  S  A  N  I  *  Q  F  L  K  K  G  K  L  E  I  P  N -
c      T  L  A  A  V  T  S  G  S  E  L  *  L  I  K  A  F  L  L  T  Y  S  N  F  *  R  K  E  S  *  K  Y  P  I -
```

Figure 7-P

```
                                                                                   Hin4I
                                                                          ApoI   HincII
                                                                         EcoRI     HpaI                    HpyCH4V
                                                                       Tsp509I     Hpy8I                    BciVI|
                         Eco57I                 HinfI                  Hpy188I|    MseI|  Hpy99I      TspDTI   ||
                         Eco57MI                 TfiI                        ||        ||     |              |  ||
     tatcttgctgaagcaagatattgggtatttcttttttttttatttgaatctttctattctgaattcagttaacgacgagatttagtatcctttcttgca
3001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3100
     atagaacgacttcgttctataacccataaagaaaaaaaaaataaaacttagaaagataagacttaagtcaattgctgctctaaatcataggaaagaacgt

a    Y  L  A  E  A  R  Y  W  V  F  L  F  F  L  F  *  I  F  L  F  *  I  Q  L  T  T  R  F  S  I  L  S  C  T -
b     I  L  L  K  Q  D  I  G  Y  F  F  F  F  Y  F  E  S  F  Y  S  E  F  S  *  R  R  D  L  V  S  F  L  A  -
c      S  C  *  S  K  I  L  G  I  S  F  F  F  I  L  N  L  S  I  L  N  S  V  N  D  E  I  *  Y  P  F  L  H  -

                                   ApoI
                                  EcoRI
                                Tsp509I
         Hpy188III                 RsaI |                      BslI
   Hin4I  BssSI |                 Csp6I||   Tsp509I       EcoNI |
      |      | |                      ||  |       |            | |
     ctttcataactcgtgaaatgccgagttggtacgaattccccaatttgcgacctaccataggatttgttatgtaaataggtatatgttcctttccattat
3101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3200
     gaaagtattgagcactttacggctcaaccatgcttaaggggttaaacgctggatggtatcctaaacaatacatttatccatatacaaggaaaggtaata

a      F  I  T  R  E  M  P  S  W  Y  E  F  P  Q  F  A  T  Y  H  R  I  C  Y  V  N  R  Y  M  F  L  S  I  M  -
b     L  S  *  L  V  K  C  R  V  G  T  N  S  P  N  L  R  P  T  I  G  F  V  M  *  I  G  I  C  S  F  P  L  * -
c      F  H  N  S  *  N  A  E  L  V  R  I  P  P  I  C  D  L  P  *  D  L  L  C  K  *  V  Y  V  P  F  H  Y  -
```

Figure 7-Q

```
                                                         TaiI
                                                     MaeIII |
                                                 HpyCH4IV|  |
                                                  NlaIV  || |
                                                 AvaII|  || |
                                                Sau96I|  || |
                                                 NciI ||  || |
                                                ScrFI ||  || |
                                                 SmaI ||  || |
                                               HpaII| ||  || |
                                                NciI| ||  || |
                                               ScrFI| ||  || |
                    BsrDI                      AvaI|| ||  || |
                   FauI |                     BsaJI|| ||  || |
     BstUI     CviJI  | |                    StyD4I|| ||  || |
      NruI    HaeIII  | |                      XmaI|| ||  || |
 Hpy188III|      MscI | |                   StyD4I||| ||  || |                           AluI
   HinfI ||      EaeI | |        SfaNI   BsmI   |||| ||  || |  BseRI                     CviJI
    TfiI ||  TspDTI|  | |  AciI      |  BfaI|   |||| ||  || |BsmFI |TspDTI        MnlI    |
       | ||      || | | |    |       |     ||   |||| ||  || |    | |     |           |    |
       gaatcgcgattgtatggccaaccattgcgggtagaatgctagatgcccgggaccacgttactattgtttctttctcctccttcatattgagcttttctat
  3201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3300
       cttagcgctaacataccggttggtaacgcccatcttacgatctacgggccctggtgcaatgataacaaagaaagaggaggaagtataactcgaaaagata

a       N  R  D  C  M  A  N  H  C  G  *  N  A  R  C  P  G  P  R  Y  Y  C  F  F  L  L  L  H  I  E  L  F  Y  -
b        I  A  I  V  W  P  T  I  A  G  R  M  L  D  A  R  D  H  V  T  I  V  S  F  S  S  F  I  L  S  F  S  I  -
c       E  S  R  L  Y  G  Q  P  L  R  V  E  C  *  M  P  G  T  T  L  L  L  F  L  S  P  P  S  Y  *  A  F  L  F -

                                   BsiEI
                                  BstKTI
                                   Hin4I
                                    PvuI
                   KpnI             TaqI
                NlaIV |            DpnI|
                  RsaI |           FauI||
                Csp6I| |          MboI|||                                 TaiI                   BpmI
       BccI  Acc65I|| |    MnlI   ||||                                BmgBI |                 Eco57MI
 Hin4I   |     BanI|| |  AciI  |  ||||     MnlI      CviJI    HpyCH4IV|  |          CviJI      |
     |   |        ||| |     |  |  ||||        |          |          || |              |     |
       tttgccaataaatgatggtacctccccccgccacgatcgaacgggaatggataagaggcttgtgggattgacgtgatagggtagggttggctatactgc
  3301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3400
       aaaacggttatttactaccatggagggggcggtgctagcttgcccttacctattctccgaacaccctaactgcactatcccatcccaaccgatatgacg

a       F  C  Q  *  M  M  V  P  P  P  A  T  I  E  R  E  W  I  R  G  L  W  D  *  R  D  R  V  G  L  A  I  L  L -
b        F  A  N  K  *  W  Y  L  P  P  P  R  S  N  G  N  G  *  E  A  C  G  I  D  V  I  G  *  G  W  L  Y  C  -
c         L  P  I  N  D  G  T  S  P  R  H  D  R  T  G  M  D  K  R  L  V  G  L  T  *  *  G  R  V  G  Y  T  A  -
```

Figure 7-R

```
                                                       BstKTI
                                                        DpnI|
                                                       NlaIV|           HpyF10VI
                                                      BamHI ||      AciI |
                                                      BstYI ||  HpyF10VI |
       CviJI          HaeII                                 ||     MwoI| |
      BstNI |         HhaI|         Eco57I              |   || Hpy188III || |                  HinfI
      ScrFI |        AfeI||        Eco57MI              |   ||  BssSI |  || |                  TaqI |
    PspGI |  |      HinP1I|||       AlwI|               |   || CviJI| |  || |            MlyI   | |
   StyD4I |  |   Hpy188I  ||||   MnlI  ||MboI           |   || AlwI || |  || |   MwoI|    PleI   | |    MnlI
        |  |  |        |   ||||     |   ||    |   ||   | || |  || |       ||      |    | |      |
     tggtggcgaactccaggctaataatctgaagcgcttggaagttgtagggaggatcctatggctcgtgaagcggttatcgccgaagtatcgactcaactat
3401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3500
     accaccgcttgaggtccgattattagacttcgcgaaccttcaacatccctcctaggataccgagcacttcgccaatagcggcttcatagctgagttgata

a      V  A  N  S  R  L  I  I  *  S  A  W  K  L  *  G  G  S  Y  G  S  *  S  G  Y  R  R  S  I  D  S  T  I  -
b     W  W  R  T  P  G  *  *  S  E  A  L  G  S  C  R  E  D  P  M  A  R  E  A  V  I  A  E  V  S  T  Q  L  S -
c      G  G  E  L  Q  A  N  N  L  K  R  L  E  V  V  G  R  I  L  W  L  V  K  R  L  S  P  K  Y  R  L  N  Y   -

                                         RsaI
                                       Csp6I|
                                     CviJI ||
                                    HaeIII ||
                                   Cac8I | ||
                                 TaiI |  | ||
                               Hpy99I| |  | ||
                            HpyCH4IV || |  | ||                                        CviJI
                              BccI | || |  | ||                                         HaeIII
                            BceAI| | || |  | ||                                          TauI
                            TaqI || | || |  | ||                     BccI        BstF5I |
                      Hpy188III| || | || |  | ||                   AciI |        Fnu4HI |
                         HaeII || || | || |  | ||                NlaIV | |          AciI| |          DraIII
                          HhaI| || || | || |  | ||               CviJI| | |        BceAI || |       HpyCH4III
  HgaI                 HinP1I || || || | || |  | ||          RsaI   || | |        BtsI| || |  CviJI      |
Hpy188I       BsaHI   TaqI  | || || || | ||EaeI  | ||        Csp6I|   || | |      TspRI| || |  FokI |      |
     |            |      |   | || || || | ||  |   | ||           ||   || | |          || || |    |  |      |
     cagaggtagttggcgtcatcgagcgccatctcgaaccgacgttgctggccgtacatttgtacggctccgcagtggatggcggcctgaagccacacagtga
3501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3600
     gtctccatcaaccgcagtagctcgcggtagagcttggctgcaacgaccggcatgtaaacatgccgaggcgtcacctaccgccggacttcggtgtgtcact

a      R  G  S  W  R  H  R  A  P  S  R  T  D  V  A  G  R  T  F  V  R  L  R  S  G  W  R  P  E  A  T  Q  *   -
b       E  V  V  G  V  I  E  R  H  L  E  P  T  L  L  A  V  H  L  Y  G  S  A  V  D  G  G  L  K  P  H  S  D  -
c     Q  R  *  L  A  S  S  S  A  I  S  N  R  R  C  W  P  Y  I  C  T  A  P  Q  W  M  A  A  *  S  H  T  V  I -
```

Figure 7-S

```
                                                AluI
                                               CviJI
                                              Cac8I
                      HpyCH4III              TspDTI                                                  BstNI
                   BstEII                    TauI                                                    ScrFI
                HpyCH4III                 Fnu4HI     BstKTI                                        BsaJI
  Eco57I           MaeIII                  AciI      DpnI                                          PspGI
 Eco57MI           Tsp45I     CviJI        BstUI    BclI                                          StyD4I
TspRI       MaeIII            HphI     Hin4I        MboI                Hin4I       CviJI
     tattgatttgctggttacggtgaccgtaaggcttgatgaaacaacgcggcgagctttgatcaacgaccttttggaaacttcggcttcccctggagagagc
3601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3700
     ataactaaacgaccaatgccactggcattccgaactactttgttgcgccgctcgaaactagttgctggaaaacctttgaagccgaaggggacctctctcg

a     Y  *  F  A  G  Y  G  D  R  K  A  *  *  N  N  A  A  S  F  D  Q  R  P  F  G  N  F  G  F  P  W  R  E  R -
b     I  D  L  L  V  T  V  T  V  R  L  D  E  T  T  R  R  A  L  I  N  D  L  L  E  T  S  A  S  P  G  E  S  -
c      L  I  C  W  L  R  *  P  *  G  L  M  K  Q  R  G  E  L  *  S  T  T  F  W  K  L  R  L  P  L  E  R  A -

                                                                                   Tsp509I
                                         Hpy99I                                   HpyCH4V
            HphI                        TspGWI                                 HpyF10VI
            SfcI                     Bme1580I                                 BstUI
           HhaI                       BsiHKAI                                 HhaI
          BpmI                       Bsp1286I                              HinP1I
         BstUI                          Hpy8I                             BlpI
       Eco57MI                        HpyCH4V                             DdeI                      HinP1I
  HinfI  HinP1I    MaeIII            ApaLI              BsaJI             AluI                      Fnu4HI
  TfiI   AciI      Tsp45I           MslI                BtgI             CviJI       MwoI           TseI
     gagattctccgcgctgtagaagtcaccattgttgtgcacgacgacatcattccgtggcgttatccagctaagcgcgaactgcaatttggagaatggcagc
3701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3800
     ctctaagaggcgcgacatcttcagtggtaacaacacgtgctgctgtagtaaggcaccgcaataggtcgattcgcgcttgacgttaaacctcttaccgtcg

a      D  S  P  R  C  R  S  H  H  C  C  A  R  R  H  H  S  V  A  L  S  S  *  A  R  T  A  I  W  R  M  A  A -
b     E  I  L  R  A  V  E  V  T  I  V  V  H  D  D  I  I  P  W  R  Y  P  A  K  R  E  L  Q  F  G  E  W  Q  R -
c      R  F  S  A  L  *  K  S  P  L  L  C  T  T  T  S  F  R  G  V  I  Q  L  S  A  N  C  N  L  E  N  G  S  -
```

Figure 7-T

```
                                           BsiEI
                                           BstKTI
       BbvI                   CviJI         PvuI           CviJI
      BsrDI                  Cac8I          TaqI     BstKTI
     BspMI  HpyCH4V         CviJI          DpnI       DpnI                                          BsaJI    AvaII
  HhaI       MboII        TaqI            MboI       MboI                                            StyI    Sau96I
     gcaatgacattcttgcaggtatcttcgagccagccacgatcgacattgatctggctatcttgctgacaaaagcaagagaacatagcgttgccttggtagg
3801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3900
     cgttactgtaagaacgtccatagaagctcggtcggtgctagctgtaactagaccgatagaacgactgttttcgttctcttgtatcgcaacggaaccatcc

a     Q  *  H  S  C  R  Y  L  R  A  S  H  D  R  H  *  S  G  Y  L  A  D  K  S  K  R  T  *  R  C  L  G  R  -
b      N  D  I  L  A  G  I  F  E  P  A  T  I  D  I  D  L  A  I  L  L  T  K  A  R  E  H  S  V  A  L  V  G  -
c     A  M  T  F  L  Q  V  S  S  S  Q  P  R  S  T  L  I  W  L  S  C  *  Q  K  Q  E  N  I  A  L  P  W  *  V  -

                           Hpy188III
                            NlaIV
      AciI             HpaII
      TauI            BsaWI                 BstKTI
  Fnu4HI             BstKTI                  MnlI        HaeII
   AciI               DpnI                  DpnI         HhaI                            TauI          Cac8I
 MspA1I               EciI                BstYI       HinP1I                            AciI          BsrI
  MnlI          AlwI MboI                 MboI   AlwI                 MseI TspDTI     Fnu4HI        CviJI
     tccagcggcggaggaactctttgatccggttcctgaacaggatctatttgaggcgctaaatgaaaccttaacgctatggaactcgccgcccgactgggct
3901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4000
     aggtcgccgcctccttgagaaactaggccaaggacttgtcctagataaactccgcgatttactttggaattgcgataccttgagcggcgggctgacccga

a     S  S  G  G  G  T  L  *  S  G  S  *  T  G  S  I  *  G  A  K  *  N  L  N  A  M  E  L  A  A  R  L  G  W  -
b      P  A  A  E  E  L  F  D  P  V  P  E  Q  D  L  F  E  A  L  N  E  T  L  T  L  W  N  S  P  P  D  W  A  -
c       Q  R  R  R  N  S  L  I  R  F  L  N  R  I  Y  L  R  R  *  M  K  P  *  R  Y  G  T  R  R  P  T  G  L  -
```

Figure 7-U

```
                                                     HpaII
                                                    BsrFI|
                                              MaeIII    ||
                                                HhaI |  ||
                                           HinP1I |  |  ||
                                              RsaI | |  |  ||                    BbvI                      BsrI
                                             Csp6I| | |  |  ||                    HhaI          Fnu4HI   HpyF10VI |
                            TaiI            FauI || | |  |  ||             BstUI |             TseI|      BglI| |
   BmrI      BsmFI   HpyCH4IV | AciI BslI |  || | |  |  ||             HinP1I |          BstF5I|| FokI   MwoI| | BmrI
     |          |           |   |    |    |  || | |  |  ||                  | |               ||| |      || |    |
     ggcgatgagcgaaatgtagtgcttacgttgtcccgcatttggtacagcgcagtaaccggcaaaatcgcgccgaaggatgtcgctgccgactgggcaatgg
4001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4100
     ccgctactcgctttacatcacgaatgcaacagggcgtaaaccatgtcgcgtcattggccgttttagcgcggcttcctacagcgacggctgacccgttacc

a      R  *  A  K  C  S  A  Y  V  V  P  H  L  V  Q  R  S  N  R  Q  N  R  A  E  G  C  R  C  R  L  G  N  G  -
b     G  D  E  R  N  V  V  L  T  L  S  R  I  W  Y  S  A  V  T  G  K  I  A  P  K  D  V  A  A  D  W  A  M  E  -
c      A  M  S  E  M  *  C  L  R  C  P  A  F  G  T  A  Q  *  P  A  K  S  R  R  R  M  S  L  P  T  G  Q  W   -

                                                                                              BstKTI
                                                                                               DpnI|
                                                                                             MboI ||
                                                                                             MnlI ||
                                                                                           HhaI | ||
                                                                                        BstUI | | ||
                                                                                        Cac8I | | ||
                                                                                         HhaI | | ||
                                                                                       HinP1I | | ||
                                                                                     HpyF10VI | | ||
                                                                                      MwoI| | | ||
                                                                                    BssHII|| | | ||
                                                                                     BstUI|| | | ||
           BsrI                                                                      HinP1I|| | | ||
          CviJI |                                                                CviJI  ||| | | ||
        HaeIII |                                                                HaeIII  ||| | | ||
        Sau96I| |                                                                MboII  ||| | | ||
        Cac8I|| |                                                               MmeI |  ||| | | ||
         NaeI|| |                                         CviJI               MboII| |  ||| | | ||
        HpaII||| |                                    HpyF10VI |          BstKTI   || |  ||| | | ||
        BmrI|||| |                                    BfaI   | |           DpnI|   || |  ||| | | ||
       BsrFI|||| |                                    AluI|  | |          MboI ||   || |  ||| | | ||
      NgoMIV|||| |                                   CviJI| MwoI| |
      Cac8I ||||| |
     HaeII| ||||| |
     HhaI|| ||||| |
   HinP1I ||| ||||| |
   BsrDI| ||| ||||| |          CviJI
     || ||| ||||| |             |                 ||   || |           | ||       || |  ||| | | ||
     agcgcctgccggcccagtatcagcccgtcatacttgaagctagacaggcttatcttggacaagaagaagatcgcttggcctcgcgcgcagatcagttgga
4101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4200
     tcgcggacggccgggtcatagtcgggcagtatgaacttcgatctgtccgaatagaacctgttcttcttctagcgaaccggagcgcgcgtctagtcaacct

a     A  P  A  G  P  V  S  A  R  H  T  *  S  *  T  G  L  S  W  T  R  R  R  S  L  G  L  A  R  R  S  V  G   -
b      R  L  P  A  Q  Y  Q  P  V  I  L  E  A  R  Q  A  Y  L  G  Q  E  E  D  R  L  A  S  R  A  D  Q  L  E  -
c     S  A  C  R  P  S  I  S  P  S  Y  L  K  L  D  R  L  I  L  D  K  K  K  I  A  W  P  R  A  Q  I  S  W  K -
```

Figure 7-V

```
                          TaiI                              Hpy188III
                        BsaAI |                 BcgI           TaqI|
                       DraIII |                DdeI|           AvaI||
                    HpyCH4IV| |         BstKTI  ||            SmlI||
        ApoI      MboII    || |          DpnI|  ||            XhoI||            Tsp509I
     Tsp509I      Hpy8I|   || |          MboI || ||         BsmAI |||  DdeI     BcgI |                    EarI
           |          ||   || |             | ||  ||            |  |||    |        |  |                    |
       agaatttgtccactacgtgaaaggcgagatcactaaggtagttggcaaataactcgagactaagtggataaaattagatagaaaaaaggtctaaataaaa
  4201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4300
       tcttaaacaggtgatgcactttccgctctagtgattccatcaaccgtttattgagctctgattcacctattttaatctatctttttttccagatttatttt

a      R  I  C  P  L  R  E  R  R  D  H  *  G  S  W  Q  I  T  R  D  *  V  D  K  I  R  *  K  K  G  L  N  K  K -
b       E  F  V  H  Y  V  K  G  E  I  T  K  V  V  G  K  *  L  E  T  K  W  I  K  L  D  R  K  K  V  *  I  K  -
c        N  L  S  T  T  *  K  A  R  S  L  R  *  L  A  N  N  S  R  L  S  G  *  N  *  I  E  K  R  S  K  *  K -

                      BstKTI                  Tsp509I
                        DpnI|                  Hin4I |                                   MseI        Hpy188I
                       MboI ||               HpyCH4V |                            Tsp509I |          BsmAI |
          Hin4I MboII  | ||    MaeIII         MboII  | MboII         Tsp509I    HpyCH4V| |          BsmBI |
              |     |  | ||        |              |   |    |             |             || |              |  |
       aagaagagaaatagaaagatcaaaaatcagttacgaaaatgcagtaattcttctttttcttctaattgattgcaattaaactcgtctcaatctgaaaaa
  4301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4400
       ttcttctctttatctttctagtttttagtcaatgcttttacgtcattaagaagaaaaaagaagattaactaacgttaatttgagcagagttagactttttt

a       E  E  K  *  K  D  Q  K  S  V  T  K  M  Q  *  F  F  F  F  L  L  I  D  C  N  *  T  R  L  N  L  K  K  -
b      K  K  R  N  R  K  I  K  N  Q  L  R  K  C  S  N  S  S  F  F  F  *  L  I  A  I  K  L  V  S  I  *  K  K -
c       R  R  E  I  E  R  S  K  I  S  Y  E  N  A  V  I  L  L  F  S  S  N  *  L  Q  L  N  S  S  Q  S  E  K  -

                                      NlaIII
                                 Hpy188III |
                                    BspHI| |
                                   BstKTI| |
                                     FatI| |
                                    DpnI|| |               HinfI
                 DraI              BsmAI||| |                TfiI
          CviJI  MseI|              MboI|||| |   Hpy188III    |           MnlI                    MboII
              |      ||                 ||||| |           |    |              |                       |
       agattgagccgagtttaaatagatttgatacgatcatgagacttgacaaatcgggattcctctattctatatttagaagatataaaggtataataca
  4401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4500
       tctaactcggctcaaatttatctaaactatgctagtactctgaactgtttagccctaaggagataagatatataaatcttctatatttccatattatgt

a       D  *  A  E  F  K  *  I  L  I  R  S  *  D  L  T  N  R  D  S  S  I  L  Y  I  *  K  I  *  R  Y  N  T  -
b        I  E  P  S  L  N  R  F  *  Y  D  H  E  T  *  Q  I  G  I  P  L  F  Y  I  F  R  R  Y  K  G  I  I  Q -
c      R  L  S  R  V  *  I  D  F  D  T  I  M  R  L  D  K  S  G  F  L  Y  S  I  Y  L  E  D  I  K  V  *  Y  N -
```

Figure 7-W

```
                                        Bsu36I
                                Taql      DdeI
                       BfaI    AlwI| BstKTI |
                   Hpy188III   AluI|| DpnI| |                         HinfI
                        XbaI| CviJI||MboI || |                 SfcI    TfiI  CviJI    FokI
     ataaataaatacaatctagaagctcgatgatccttaggattggtatattcttttctatcctgtagtttgtagtttccctgaatcaagccaagtatcacac
4501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4600
     tatttatttatgttagatcttcgagctactaggaatcctaaccatataagaaaagataggacatcaaacatcaaagggacttagttcggttcatagtgtg

a    I  N  K  Y  N  L  E  A  R  *  S  L  G  L  V  Y  S  F  L  S  C  S  L  *  F  P  *  I  K  P  S  I  T  P -
b     *  I  N  T  I  *  K  L  D  D  P  *  D  W  Y  I  L  F  Y  P  V  V  C  S  F  P  E  S  S  Q  V  S  H  -
c      K  *  I  Q  S  R  S  S  M  I  L  R  I  G  I  F  F  S  I  L  *  F  V  V  S  L  N  Q  A  K  Y  H  T  -

         BstF5I                                                         ApoI
         BsmFI|                                  Tsp509I              Tsp509I
          MnlI|   BccI TspGWI                     MseI|                MseI |
     ctctttctacccatcctgtatattgtcccctttgttccgtgttgaaatagaaccttaatttattacttattttttattaaattttagatttgttagtga
4601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4700
     gagaaagatgggtaggacatataacaggggaaacaaggcacaactttatcttggaattaaataatgaataaaaaaataatttaaaatctaaacaatcact

a     L  S  T  H  P  V  Y  C  P  L  C  S  V  L  K  *  N  L  N  L  L  L  I  F  L  L  N  F  R  F  V  S  D  -
b    L  F  L  P  I  L  Y  I  V  P  F  V  P  C  *  N  R  T  L  I  Y  Y  L  F  F  Y  *  I  L  D  L  L  V  I -
c     S  F  Y  P  S  C  I  L  S  P  L  F  R  V  E  I  E  P  *  F  I  T  Y  F  F  I  K  F  *  I  C  *  *  -

                                                                                    ApoI
                                                                                  Tsp509I
                                                    MnlI                          Hin4I |
                                 Tsp509I           Hin4I |          SfaNI    TaqI    | |
     ttagatattagtattagacgagattttacgaaacaattatttttttatttctttataggaggaggacaaatctctttttcgatgcgaatttgacacgaca
4701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4800
     aatctataatcataatctgctctaaaatgctttgttaataaaaaaataaagaaatatcctctcctgtttagagaaaaagctacgcttaaactgtgctgt

a     *  I  L  V  L  D  E  I  L  R  N  N  Y  F  F  I  S  L  *  E  R  T  N  L  F  F  R  C  E  F  D  T  T  -
b      R  Y  *  Y  *  T  R  F  Y  E  T  I  I  F  L  F  L  Y  R  R  G  Q  I  S  F  F  D  A  N  L  T  R  H  -
c    L  D  I  S  I  R  R  D  F  T  K  Q  L  F  F  Y  F  F  I  G  E  D  K  S  L  F  S  M  R  I  *  H  D  I -
```

Figure 7-X

```
                 TauI
               AciI |
             Fnu4HI |     Tsp509I               DraI
             CviJI| |     MseI    |   FalI      MseI|             NlaIV             AseI
                  || |         |  |     |          ||                |               MseI              MmeI         HinfI  Hpy188I
                                                                                                                     TfiI  |
     taggagaagccgccctttattaaaaattatattatttaaataatataaagggggttccaacatattaatatatagtgaagtgttcccccagattcagaa
4801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4900
     atcctcttcggcgggaaataatttttaatataataaaatttattatatttcccccaaggttgtataattatatatcacttcacaagggggtctaagtctt

a    * E K P P F I K N Y I I L N N I K G V P T Y * Y I V K C S P R F R T -
b     R R S R P L L K I I L F * I I * R G F Q H I N I * * S V P P D S E  -
c      G E A A L Y * K L Y Y F K * Y K G G S N I L I Y S E V F P Q I Q N -

                                                                    Hpy188I
                                    MseI      BfaI              DdeI   |                         Tsp509I
                                       |         |                 |   |                               |
     cttttttcaatactcacaatccttattagttaataatcctagtgattggatttctatgcttagtctgataggaaataagatattcaaataaataatttt
4901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5000
     gaaaaaaagttatgagtgttaggaataatcaattattaggatcactaacctaaagatacgaatcagactatcctttattctataagtttatttattaaaa

a      F F Q Y S Q S L L V N N P S D W I S M L S L I G N K I F K * I I L -
b     L F F N T H N P Y * L I I L V I G F L C L V * * E I R Y S N K * F Y -
c     F F S I L T I L I S * * S * * L D F Y A * S D R K * D I Q I N N F  -

                                    HpyCH4V
                                     NlaIII                                                                          AarI
       TspDTI                       Hin4I |                                          Tsp509I                         BspMI
     Hin4I |            TspDTI      FatI | |                       BccI               MseI|TaqI          MseI        XmnI|
         |  |               |           | | |                         |                  || |              |            ||
     atagcgaatgactattcatctattgtattttcatgcaaatagggggcaagaaaactctatggaaagatggtggtttaattcgatgttgtttaagaaggag
5001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5100
     tatcgcttactgataagtagataacataaaagtacgtttatcccccgttcttttgagatacctttctaccaccaaattaagctacaacaaattcttcctc

a      * R M T I H L L Y F H A N R G Q E N S M E R W W F N S M L F K K E  -
b       S E * L F I Y C I F M Q I G G K K T L W K D G G L I R C C L R R S -
c     I A N D Y S S I V F S C K * G A R K L Y G K M V V * F D V V * E G V -
```

Figure 7-Y

```
                                                                                 BsaBI     TspDTI
                                                                               BstKTI |   MmeI  |
                                                                                DpnI| |TspDTI | |
      BstBI                                   AvaII                           BstYI || |MboII| |          MnlI
       TaqI          CviJI                    Sau96I                      AlwI MboI || |TaqI|| |        BsaXI |
        |              |                        |                           |    |  ||  |   ||| |  |          | |
        ttcgaacgcaggtgtgggctaaataaatcaatgggcagtcttggtcctattgaaaataccaatgaagatccaaatcgaaaagtgaaaaacattcatagtt
   5101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5200
        aagcttgcgtccacacccgatttatttagttacccgtcagaaccaggataacttttatggttacttctaggtttagcttttcacttttgtaagtatcaa

a       F  E  R  R  C  G  L  N  K  S  M  G  S  L  G  P  I  E  N  T  N  E  D  P  N  R  K  V  K  N  I  H  S  W -
b        S  N  A  G  V  G  *  I  N  Q  W  A  V  L  V  L  L  K  I  P  M  K  I  Q  I  E  K  *  K  T  F  I  V  -
c         R  T  Q  V  W  A  K  *  I  N  G  Q  S  W  S  Y  *  K  Y  Q  *  R  S  K  S  K  S  E  K  H  S  *  L  -

               Tsp509I
       Hpy188III  |
          MaeIII  |
          Tsp45I  |    HpyCH4V                                      ApoI
       HinfI  |   |    BsaXI |                                     Tsp509I
        TfiI  |   | BfaI  |  |                                    Hpy188I|
           |  |   |   |   |  |                         MseI TspDTI     ||       Hpy188I
           |  |   |   |   |  |                           |     |       ||          |
        ggaggaatcgtgacaattctagttgcagtaatgttgattatttattcggcgttaaagacattcggaatttcatctctgatgacactttttagttagtga
   5201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5300
        cctccttagcactgttaagatcaacgtcattacaactaataaataagccgcaatttctgtaagccttaaagtagagactactgtgaaaaaatcaatcact

a         R  N  R  D  N  S  S  C  S  N  V  D  Y  L  F  G  V  K  D  I  R  N  F  I  S  D  D  T  F  L  V  S  D -
b       G  G  I  V  T  I  L  V  A  V  M  L  I  I  Y  S  A  L  K  T  F  G  I  S  S  L  M  T  L  F  *  L  V  I -
c        E  E  S  *  Q  F  *  L  Q  *  C  *  L  F  I  R  R  *  R  H  S  E  F  H  L  *  *  H  F  F  S  *  *  -

                                                                                  BspCNI
                                                                                 BseMII|
                                                                                BstKTI||
                                                                                  DpnI|||  Hpy188I   BfaI
     BsmAI  HpyCH4III          BccI                                             MboI ||||   DdeI|Hpy8I |
        |        |               |                                                 | ||||     ||    |  |
        taggaatggagacagttattccatctattttgatattgaaaatcatatttttgagattgacaacgatcattcttttctgagtgaactagaaagttctttt
   5301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5400
        atccttacctctgtcaataaggtagataaaactataacttttagtataaaaactctaactgttgctagtaagaaaagactcacttgatctttcaagaaaa

a         R  N  G  D  S  Y  S  I  Y  F  D  I  E  N  H  I  F  E  I  D  N  D  H  S  F  L  S  E  L  E  S  S  F  -
b          G  M  E  T  V  I  P  S  I  L  L  I  L  K  I  I  F  L  R  L  T  T  I  I  L  F  *  V  N  *  K  V  L  F -
c       *  E  W  R  Q  L  F  H  L  F  *  Y  *  K  S  Y  F  *  D  *  Q  R  S  F  F  S  E  *  T  R  K  F  F  L -
```

Figure 7-Z

```
                                                               NlaIII
                                                                 NspI
                                         BstKTI                MmeI |
                                           DpnI|             AflIII | |
              Tsp509I                    BstYI ||  Tsp509I     FatI | |
         TaqI   TaqI | Hpy188I      AlwI MboI  ||  MboII |     PciI | |
          |      | |     |             |    |  ||     | |        | | |
     tatagttatcgaaactcgaattatcggaataatggatttaggggcgaagatccctactataattcttacatgtatgatactcaatatagttggaataatc
5401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5500
     atatcaatagctttgagcttaatagccttattacctaaatccccgcttctagggatgatattaagaatgtacatactatgagttatatcaaccttattag

a    Y  S  Y  R  N  S  N  Y  R  N  N  G  F  R  G  E  D  P  Y  Y  N  S  Y  M  Y  D  T  Q  Y  S  W  N  N  H -
b     I  V  I  E  T  R  I  I  G  I  M  D  L  G  A  K  I  P  T  I  I  L  T  C  M  I  L  N  I  V  G  I  I  -
c     *  L  S  K  L  E  L  S  E  *  W  I  *  G  R  R  S  L  L  *  F  L  H  V  *  Y  S  I  *  L  E  *  S  -

                                                                                 BceAI
                                                                              HpyCH4III
                                                                                MaeIII
         HpyCH4V                                                        HpyCH4III |
     AseI  Eco57I|    BsaBI                                               MaeIII  |
     MseI Eco57MI|MboII |           BsmAI               PsiI      Tsp509I Tsp45I  |  HphI Sau96I
      |          ||   | |             |                   |          |      |     |    |     |
     acattaatagttgcattgatagttatcttcagtctcaaatctgtatagatacttccattataagtggtagtgagaattacggtgacagttacatttatag
5501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5600
     tgtaattatcaacgtaactatcaatagaagtcagagtttagacatatctatgaaggtaatattcaccatcactcttaatgccactgtcaatgtaaatatc

a      I  N  S  C  I  D  S  Y  L  Q  S  Q  I  C  I  D  T  S  I  I  S  G  S  E  N  Y  G  D  S  Y  I  Y  R  -
b     T  L  I  V  A  L  I  V  I  F  S  L  K  S  V  *  I  L  P  L  *  V  V  V  R  I  T  V  T  V  T  F  I  G -
c     H  *  *  L  H  *  *  L  S  S  V  S  N  L  Y  R  Y  F  H  Y  K  W  *  *  E  L  R  *  Q  L  H  L  *  -

                                                 Hpy8I
   CviJI                 HphI                  BsrI  |           HpyF10VI    MseI
   HaeIII             TaqI |      MnlI      NlaIV |AccI|            MwoI| BtsI|
      |                 |  |        |          |  |  ||                || TspRI|
     ggccgtttgtggtggtgaaagtcgaaatagtagtgaaaacgagggttccagtagacgaactcgcacgaagggcagtgatttaactataagagaaagttct
5601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5700
     ccggcaaacaccaccactttcagctttatcatcacttttgctcccaaggtcatctgcttgagcgtgcttcccgtcactaaattgatattctctttcaaga

a     A  V  C  G  G  E  S  R  N  S  S  E  N  E  G  S  S  R  R  T  R  T  K  G  S  D  L  T  I  R  E  S  S  -
b      P  F  V  V  V  K  V  E  I  V  V  K  T  R  V  P  V  D  E  L  A  R  R  A  V  I  *  L  *  E  K  V  L -
c    G  R  L  W  W  *  K  S  K  *  *  *  K  R  G  F  Q  *  T  N  S  H  E  G  Q  *  F  N  Y  K  R  K  F  * -
```

Figure 7-AA

```
                 MaeIII
               TaqI  |
              AvaI|  |
         Hpy188III|  |
              SmlI|  |
              XhoI|  |
           BstKTI||  |
            DpnI|||  |
           MboI ||||  |                                                            ApoI
           MnlI ||||  |                              Tsp509I          Tsp509I   Tsp509I
              | ||||  |                              Tsp509I          MseI |  PsiI  |                 SspI
              | ||||  |                                 |             |    |   |    |                   |
     aatgatctcgaggtaactcaaaaatacaggcatttgtgggttcaatgcgaaaattgttatggattaaattataagaaattttgaaatcaaaaatgaata
5701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5800
     ttactagagctccattgagtttttatgtccgtaaacacccaagttacgcttttaacaatacctaatttaatattctttaaaactttagtttttacttat

a    N  D  L  E  V  T  Q  K  Y  R  H  L  W  V  Q  C  E  N  C  Y  G  L  N  Y  K  K  F  L  K  S  K  M  N  I -
b     M  I  S  R  *  L  K  N  T  G  I  C  G  F  N  A  K  I  V  M  D  *  I  I  R  N  F  *  N  Q  K  *  I  -
c      *  S  R  G  N  S  K  I  Q  A  F  V  G  S  M  R  K  L  L  W  I  K  L  *  E  I  F  E  I  K  N  E  Y -

                                                ApoI
                                                EcoRI
      TspDTI                                  Tsp509I          NlaIII   AluI                       Tsp509I
     Hpy8I |        EcoRV                Hpy188I  |            FatI  |  CviJI       Tsp509I     BsrBI  |
         | |            |                     |   |               |  |    |              |       AciI | |
     tttgtgaacaatgtggatatcatttgaaaatgagtagttcagatagaattcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcaca
5801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5900
     aaacacttgttacacctatagtaaacttttactcatcaagtctatcttaagcattagtaccagtatcgacaaaggacacactttaacaataggcgagtgt

a     C  E  Q  C  G  Y  H  L  K  M  S  S  S  D  R  I  R  N  H  G  H  S  C  F  L  C  E  I  V  I  R  S  Q -
b    F  V  N  N  V  D  I  I  *  K  *  V  V  Q  I  E  F  V  I  M  V  I  A  V  S  C  V  K  L  L  S  A  H  N -
c    L  *  T  M  W  I  S  F  E  N  E  *  F  R  *  N  S  *  S  W  S  *  L  F  P  V  *  N  C  Y  P  L  T  -

                                           NlaIV
                                       HpyF10VI|                                                     AciI
                                      BstNI   ||                                                     Cac8I
                                      ScrFI   ||                                                     TspRI
                                     BsaJI|   ||                                            HpyF10VI    |
                                    CviJI||   ||                                               MwoI|    |
             HpaII                  PspGI||BanI||      AluI            Tsp509I         BtsI  ||    |
             CviJI|                StyD4I||MwoI||      CviJI             AseI|   HhaI |   ||    |
                 ||                   ||| |   |||          |           MseI|   HinP1I |  |  ||    |
     attccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctt
5901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6000
     taaggtgtgttgtatgctcggccttcgtatttcacatttcggaccccacggattactcactcgattgagtgtaattaacgcaacgcgagtgacgggcgaa

a     F  H  T  T  Y  E  P  E  A  *  S  V  K  P  G  V  P  N  E  *  A  N  S  H  *  L  R  C  A  H  C  P  L  -
b      S  T  Q  H  T  S  R  K  H  K  V  *  S  L  G  C  L  M  S  E  L  T  H  I  N  C  V  A  L  T  A  R  F -
c    I  P  H  N  I  R  A  G  S  I  K  C  K  A  W  G  A  *  *  V  S  *  L  T  L  I  A  L  R  S  L  P  A  F -
```

Figure 7-BB

```
                                                 AciI
                                                BstUI
                                                 HhaI
                                                 MnlI
                                                 BslI|
                      AseI                     TspDTI|
                      MseI                     BstUI||
                HpyCH4V  |                    HinP1I||
                Fnu4HI | |                   BsaXI |||                        HaeII
                  AluI|| |                FauI   | |||                        HhaI|
                 CviJI|| |              CviJI|   | |||                      BsaXI ||
                MspA1I|| |             HaeIII|   | |||                     HinP1I ||
Hpy188III        PvuII|| |                EaeI ||   | |||                 HpyF10VI | ||   EarI      MlyI
 BsrI    |        TseI|| |HinfI          |  ||   | |||                      MwoI|  | ||   SapI      PleI
 FauI    |  BbvI Cac8I ||| | TfiI        |  ||   | |||          AciI       MboII|| | ||   AciI|     BbvI |
tccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcac
6001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6100
aggtcagccctttggacagcacggtcgacgtaattacttagccggttgcgcgcccctctccgccaaacgcataacccgcgagaaggcgaaggagcgagtg

a     S  S  R  E  T  C  R  A  S  C  I  N  E  S  A  N  A  R  G  E  A  V  C  V  L  G  A  L  P  L  P  R  S  L  -
b      P  V  G  K  P  V  V  P  A  A  L  M  N  R  P  T  R  G  E  R  R  F  A  Y  W  A  L  F  R  F  L  A  H  -
c       Q  S  G  N  L  S  C  Q  L  H  *  *  I  G  Q  R  A  G  R  G  G  L  R  I  G  R  S  S  A  S  S  L  T  -

                                             AluI
                                            CviJI
                                        HpyF10VI |
          BbvI                             AciI | |
          HhaI                            BsrBI | |
        HinP1I |                         Cac8I | | |
       Fnu4HI | |                         TauI| | | |
        TseI| | |                     Fnu4HI | | | | |
    TaqII || | |                         AciI|| | | | |
    TspRI || | |                     Fnu4HI || | | | |
  HinfI | || | |                     CviJI| || | | | |
  MnlI| | || | |  BsiEI               TseI| || | | |  MwoI| |    AciI HpyCH4III      HinfI
                                                                                    TfiI
tgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaag
6101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6200
actgagcgacgcgagccagcaagccgacgccgctcgccatagtcgagtgagtttccgccattatgccaataggtgtcttagtcccctattgcgtcctttc

a      T  R  C  A  R  S  F  G  C  G  E  R  Y  Q  L  T  Q  R  R  *  Y  G  Y  P  Q  N  Q  G  I  T  Q  E  R  -
b     *  L  A  A  L  G  R  S  A  A  A  S  G  I  S  S  L  K  G  G  N  T  V  I  H  R  I  R  G  *  R  R  K  E  -
c      D  S  L  R  S  V  V  R  L  R  R  A  V  S  A  H  S  K  A  V  I  R  L  S  T  E  S  G  D  N  A  G  K  -
```

Figure 7-CC

```
                                   HpyCH4III                 Cac8I
                                     NlaIV |              HpyF10VI
                                  BstNI  | |            BstUI     |
                                  ScrFI  | |             TauI     |
        NlaIII                   CviJI | | |            AciI |    |
          NspI                  HaeIII | | |          Fnu4HI |    |
     AflIII |          Cac8I    PspGI | | |          CviJI| |    |                AciI
       FatI |         CviJI |  StyD4I | | |        HaeIII| |    |              NlaIV |
       PciI |        HaeIII | BslI  | | | |    BslI    || |  MwoI|    EciI    CviJI| |           BpuEI SfaNI
          | |             | |    |  | | | |       |    || |     ||      |          || |               |     |
          aacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaa
     6201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6300
          ttgtacactcgttttccggtcgttttccggtccttggcatttttccggcgcaacgaccgcaaaaaggtatccgaggcggggggactgctcgtagtgtttt

a          T  C  E  Q  K  A  S  K  R  P  G  T  V  K  R  P  R  C  W  R  F  S  I  G  S  A  P  L  T  S  I  T  K  -
b           H  V  S  K  R  P  A  K  G  Q  E  P  *  K  G  R  V  A  G  V  F  P  *  A  P  P  P  *  R  A  S  Q  K  -
c          N  M  *  A  K  G  Q  Q  K  A  R  N  R  K  K  A  A  L  L  A  F  F  H  R  L  R  P  P  D  E  H  H  K  N -

                   Hpy188I
                    HgaI |                                                                      MnlI
                DrdI   | |                                   BstNI               BstNI        HhaI |
                MnlI   | |                                   ScrFI               ScrFI      HinP1I | |
               SmlI |  | |                                  PspGI |             BsaJI |    BssSI   | | |
           Hpy99I|  |  | |                                  StyD4I |            PspGI |  AluI  |   | | |
         TaqI   ||  |  | |                                          |           StyD4I | CviJI  |   | | |   Hpy188I
            |   ||  |  | |                                         ||               ||     |    |   | | |      |
          atcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccct
     6301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6400
          tagctgcgagttcagtctccaccgctttgggctgtcctgatatttctatggtccgcaaagggggaccttcgagggagcacgcgagaggacaaggctggga

a          I  D  A  Q  V  R  G  G  E  T  R  Q  D  Y  K  D  T  R  R  F  P  L  E  A  P  S  C  A  L  L  F  R  P  C -
b           S  T  L  K  S  E  V  A  K  P  D  R  T  I  K  I  P  G  V  S  P  W  K  L  P  R  A  L  S  C  S  D  P  -
c            R  R  S  S  Q  R  W  R  N  P  T  G  L  *  R  Y  Q  A  F  P  P  G  S  S  L  V  R  S  P  V  P  T  L  -
```

Figure 7-DD

```
                EciI
              HpaII|
             BsaWI||
              BslI||
          BciVI  |||
          TauI|  |||                                     HaeII
        AciI ||  |||      AciI                           HhaI|       AluI                           BseMII
      Fnu4HI ||  |||   MmeI  |         Hpy188III    HinP1I ||      CviJI    SfcI      DdeI      BspCNI|
           | ||  |||      |  |              |            | ||          |       |         |            ||
        gccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgc
   6401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6500
        cggcgaatggcctatggacaggcggaaagagggaagcccttcgcaccgcgaaagagtatcgagtgcgacatccatagagtcaagccacatccagcaagcg

a         R  L  P  D  T  C  P  P  F  S  L  R  E  A  W  R  F  L  I  A  H  A  V  G  I  S  V  R  C  R  S  F  A  -
b        A  A  Y  R  I  P  V  R  L  S  P  F  G  K  R  G  A  F  S  *  L  T  L  *  V  S  Q  F  G  V  G  R  S  L -
c         P  L  T  G  Y  L  S  A  F  L  P  S  G  S  V  A  L  S  H  S  S  R  C  R  Y  L  S  S  V  *  V  V  R  -

                                                     HhaI
                   Bme1580I                      HinP1I |                                     HpaII
                     BsiHKAI                    Fnu4HI | |                                      MlyI
               CviJI Bsp1286I                  MspA1I| | |                                      NciI
           AluI    |  Hpy8I |                    TseI| | |    MaeIII            HinfI          ScrFI
          BseYI    |HpyCH4V |           CviJI  AciI || | |   HpaII |    Hpy188III  |           PleI|
          CviJI    |ApaLI | |       BbvI     | BsiEI || | |  BsaWI| |          SmlI | StyD4I||     BpuEI
              |    |    | | |          |     |     | ||| |      || |            |  |      |||        |
        tccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttat
   6501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6600
        aggttcgacccgacacacgtgcttgggggcaagtcgggctggcgacgcggaataggccattgatagcagaactcaggttgggccattctgtgctgaata

a        P  S  W  A  V  C  T  N  P  P  F  S  P  T  A  A  P  Y  P  V  T  I  V  L  S  P  T  R  *  D  T  T  Y  -
b         Q  A  G  L  C  A  R  T  P  R  S  A  R  P  L  R  L  I  R  *  L  S  S  *  V  Q  P  G  K  T  R  L  I -
c        S  K  L  G  C  V  H  E  P  P  V  Q  P  D  R  C  A  L  S  G  N  Y  R  L  E  S  N  P  V  R  H  D  L  S -
```

Figure 7-EE

```
                         BbvI
                        BsrI|
                       TspRI|
                      BbvI ||
                   MaeIII  ||
                  AlwNI |  ||
               CviJI  | |  ||
            Fnu4HI |  | |  ||
             TseI| |  | |  ||
           BsrI || |  | |  ||
         Fnu4HI || |  | |  ||
          TspRI || |  | |  ||                                                      CviJI    CviJI
  MmeI     TseI| || |  | |  ||     MnlI              AciI  SfcI Hpy188III       HaeIII   BslI  |   BfaI
     |          || || |  | |  ||       |                   |     |          |              |       |   |     |
     cgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactag
6601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6700
     gcggtgaccgtcgtcggtgaccattgtcctaatcgtctcgctccatacatccgccacgatgtctcaagaacttcaccaccggattgatgccgatgtgatc

a    R  H  W  Q  Q  P  L  V  T  G  L  A  E  R  G  M  *  A  V  L  Q  S  S  *  S  G  G  L  T  T  A  T  L  E -
b     A  T  G  S  S  H  W  *  Q  D  *  Q  S  E  V  C  R  R  C  Y  R  V  L  E  V  V  A  *  L  R  L  H  *  -
c      P  L  A  A  A  T  G  N  R  I  S  R  A  R  Y  V  G  G  A  T  E  F  L  K  W  W  P  N  Y  G  Y  T  R -

                                                                      HpaII
                                                                    BstKTI|
                                                                      DpnI||
                                                                     MboI |||
                                                               Hpy188III | |||
                                       MaeIII      Eco57I          AluI | | |||
  HpyCH4III            HhaI            BsrI |     Eco57MI          AlwI | | |||                 MspA1I
     BceAI |         HinP1I |        CviJI| |   Hpy188I  |        CviJI | | | |||                 AciI |   AciI
         |  |             | |             || |         |  |            | | | |||                   |  |      |
     aaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggt
6701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6800
     ttcctgtcataaaccatagacgcgagacgacttcggtcaatggaagcctttttctcaaccatcgagaactaggccgtttgtttggtggcgaccatcgcca

a      G  Q  Y  L  V  S  A  L  C  *  S  Q  L  P  S  E  K  E  L  V  A  L  D  P  A  N  K  P  P  L  V  A  V -
b     K  D  S  I  W  Y  L  R  S  A  E  A  S  Y  L  R  K  K  S  W  *  L  L  I  R  Q  T  N  H  R  W  *  R  W -
c      R  T  V  F  G  I  C  A  L  L  K  P  V  T  F  G  K  R  V  G  S  S  *  S  G  K  Q  T  T  A  G  S  G  -
```

Figure 7-FF

```
                                                         BstKTI
                                                          DpnI|
                                                        BstYI ||
                                                         MboI ||
                                                      AlwI  | ||
                           BpuEI                    AlwI |  | ||
                            HhaI               Hpy188III |  | ||
                           BbvI |                   SmlI ||  | ||
            Fnu4HI        BstUI |               BstKTI| ||  | ||  BstKTI                                BseMII
              TseI|      HinP1I |                 DpnI|| ||  | ||   DpnI|                DdeI           BspCNI|
            Cac8I ||  HpyF10VI | |               BstYI ||| ||  | ||  MboII|                PpiI|    TspRI      ||
         HpyCH4V | ||      MwoI| | |               MboI ||| ||  | ||  MboI ||      Hpy188I    ||  HgaI |      ||
               | | ||         || | |                 | ||| ||  | ||     | ||           |     ||   |   |      ||
     ggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaa
6801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6900
     ccaaaaaaacaaacgttcgtcgtctaatgcgcgtctttttttcctagagttcttctaggaaactagaaaagatgccccagactgcgagtcaccttgcttt

a     V  F  L  F  A  S  S  R  L  R  A  E  K  K  D  L  K  K  I  L  *  S  F  L  R  G  L  T  L  S  G  T  K  -
b      F  F  C  L  Q  A  A  D  Y  A  Q  K  K  R  I  S  R  R  S  F  D  L  F  Y  G  V  *  R  S  V  E  R  K  -
c     G  F  F  V  C  K  Q  Q  I  T  R  R  K  K  G  S  Q  E  D  P  L  I  F  S  T  G  S  D  A  Q  W  N  E  N -

                                                      BstKTI
                                        BstKTI         DpnI|
                                          DpnI|      BstYI ||
                         NlaIII         BstYI ||      MboI ||        MseI
           MseI      Hpy188III |          MboI ||     BfaI | ||Tsp509I  |
           TaiI         BspHI| |        HphI | ||   AlwI | | ||  DraI|  |               DraI
   HpyCH4IV  |       PpiI FatI| |     MboII  | | ||AlwI| | | || MseI||  |               MseI|TspDTI
          |  |          |   || |          |  | | ||  || | | ||    |||  |                 ||    |
     actcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatata
6901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7000
     tgagtgcaattccctaaaaccagtactctaatagtttttcctagaagtggatctaggaaaatttaatttttacttcaaaatttagttagatttcatatat

a     T  H  V  K  G  F  W  S  *  D  Y  Q  K  G  S  S  P  R  S  F  *  I  K  N  E  V  L  N  Q  S  K  V  Y  M -
b      L  T  L  R  D  F  G  H  E  I  I  K  K  D  L  H  L  D  P  F  K  L  K  M  K  F  *  I  N  L  K  Y  I  -
c      S  R  *  G  I  L  V  M  R  L  S  K  R  I  F  T  *  I  L  L  N  *  K  *  S  F  K  S  I  *  S  I  Y  -
```

Figure 7-GG

```
                                                         BseMII
                                                        BspCNI|
                                                       TspDTI ||
                                                       FokI | ||
             HpyCH4III                   NlaIV       BstKTI| | ||                      MlyI
                MaeIII           MnlI    BanI |        DpnI|| | ||                      PleI
      Hpy8I  Hpy188I |           MseI |  TspRI |  DdeI MboI ||| | ||  BstF5I  Hin4I  | HinfI AhdI
          |       |  |              | |      | |     |    | ||| | ||     |       |  |     |    |
     tgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtc
7001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7100
     actcatttgaaccagactgtcaatggttacgaattagtcactccgtggatagagtcgctagacagataaagcaagtaggtatcaacggactgaggggcag

a      S  K  L  G  L  T  V  T  N  A  *  S  V  R  H  L  S  Q  R  S  V  Y  F  V  H  P  *  L  P  D  S  P  S  -
b     *  V  N  L  V  *  Q  L  P  M  L  N  Q  *  G  T  Y  L  S  D  L  S  I  S  F  I  H  S  C  L  T  P  R  R -
c     E  *  T  W  S  D  S  Y  Q  C  L  I  S  E  A  P  I  S  A  I  C  L  F  R  S  S  I  V  A  *  L  P  V   -

                                     BsrI                    BpmI
                                   BccI |                 Eco57MI
                                  NlaIV |              BstUI    |
                                 CviJI| |              AciI |   |            Hpy188III
                                HaeIII| |  HpyCH4V    BsaI | |   |              NlaIV |
                               Sau96I|| |   TspRI|   BsmAI | |   |             CviJI| |
               MnlI              BmrI ||| |Fnu4HI||   BsrDI | |   |           HpaII || |
Hpy99I       Hin4I |      CviJI  BbvI | ||| | TseI|||  Hin4I| | |  HphI     BsrFI| || |Hin4I
     |            | |          |     | | ||| |    ||||     || | |    |          || || |   |
     gtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataa
7101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7200
     cacatctattgatgctatgccctcccgaatggtagaccggggtcacgacgttactatggcgctctgggtgcgagtggccgaggtctaaatagtcgttatt

a     C  R  *  L  R  Y  G  R  A  Y  H  L  A  P  V  L  Q  *  Y  R  E  T  H  A  H  R  L  Q  I  Y  Q  Q  *   -
b      V  D  N  Y  D  T  G  G  L  T  I  W  P  Q  C  C  N  D  T  A  R  P  T  L  T  G  S  R  F  I  S  N  K  -
c     V  *  I  T  T  I  R  E  G  L  P  S  G  P  S  A  A  M  I  P  R  D  P  R  S  P  A  P  D  L  S  A  I  N -
```

Figure 7-HH

```
                        HhaI
                     HinP1I |
                 CviJI     | |
                HaeIII     | |
            HpyF10VI |     | |
          HpaII     | |     | |                  FokI                   Tsp509I      NciI
        CviJI|Sau96I |     | |              HpyCH4V |                BccI    |      ScrFI   BfaI
       Cac8I || BglI| |     | |       AvaII       | |        BstF5I  MnlI|AseI|    HpaII|  AluI|
      CviJI | || MwoI| |     | |       Sau96I EciI| |    AciI      |BsrI ||MseI|  StyD4I||  CviJI|
      accagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttc
 7201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7300
      tggtcggtcggccttcccggctcgcgtcttcaccaggacgttgaaataggcggaggtaggtcagataattaacaacggcccttcgatctcattcatcaag

a     T  S  Q  P  E  G  P  S  A  E  V  V  L  Q  L  Y  P  P  P  S  S  L  L  I  V  A  G  K  L  E  *  V  V  R -
b      P  A  S  R  K  G  R  A  Q  K  W  S  C  N  F  I  R  L  H  P  V  Y  *  L  L  P  G  S  *  S  K  *  F  -
c       Q  P  A  G  R  A  E  R  R  S  G  P  A  T  L  S  A  S  I  Q  S  I  N  C  C  R  E  A  R  V  S  S  S -

                   TaiI
                 AclI |                                                            NlaIV
              HpyCH4IV |                    MaeIII                          HpaII  |
                 HhaI | |                    SfaNI                          BsaWI| |
       MseI      FspI| | |                   Tsp45I          TspDTI        AluI || |      DpnI
   BsrI |      HinP1I|| | |  BsrDI  SfcI    MslI  |         Hpy99I |  CviJI  CviJI || |     MboI |
      gccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacga
 7301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7400
      cggtcaattatcaaacgcgttgcaacaacggtaacgatgtccgtagcaccacagtgcgagcagcaaaccataccgaagtaagtcgaggccaagggttgct

a      Q  L  I  V  C  A  T  L  L  P  L  L  Q  A  S  W  C  H  A  R  R  L  V  W  L  H  S  A  P  V  P  N  D -
b     A  S  *  *  F  A  Q  R  C  C  H  C  Y  R  H  R  G  V  T  L  V  V  W  Y  G  F  I  Q  L  R  F  P  T  I -
c      P  V  N  S  L  R  N  V  V  A  I  A  T  G  I  V  V  S  R  S  S  F  G  M  A  S  F  S  S  G  S  Q  R  -
```

Figure 7-II

```
                                                                               Hpy188I
                                                                            MnlI   |
                  BstKTI                                              BsiEI   |    |
                   DpnI|                                             BstKTI   |    |           TauI
                 NlaIII||                                              PvuI   |    |           AciI |
                   MboI||                     AluI                     DpnI|  |    |         Fnu4HI |
                FatI  |||        HpyCH4V      CviJI             Hpy188I ||  |    |          CviJI| |
              AlwI |  |||   NlaIII    |     TaqII  |   AvaII            | ||  |    |         HaeIII| |   BtsI
 BstKTI  MaeIII | |  |||FatI   |     |     AciI |  |  Sau96I     MboI  | ||  |    |           EaeI || |  TspRI
      |       | | |  ||||   |     |     |      |  |  |      |         | ||  |    |              | || |     |
      tcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcac
 7401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7500
      agttccgctcaatgtactagggggtacaacacgtttttcgccaatcgaggaagccaggaggctagcaacagtcttcattcaaccggcgtcacaatagtg

a      Q  G  E  L  H  D  P  P  C  C  A  K  K  R  L  A  P  S  V  L  R  S  L  S  E  V  S  W  P  Q  C  Y  H  -
b       K  A  S  Y  M  I  P  H  V  V  Q  K  S  G  *  L  L  R  S  S  D  R  C  Q  K  *  V  G  R  S  V  I  T  -
c      S  R  R  V  T  *  S  P  M  L  C  K  K  A  V  S  S  F  G  P  P  I  V  V  R  S  K  L  A  A  V  L  S  L-

                       BbvI                BstF5I
                    Tsp509I|                 SfaNI
                      TspRI||              NlaIII|                          RsaI
                  HpyCH4V |||                FatI ||                          ScaI
        MslI     BtsI  |  |||      HpyCH4III |   ||                         Csp6I|  BspCNI
    NlaIII|  Fnu4HI |  |  |||       TspGWI |   ||                MaeIII    TatI||  BseMII|  Hpy188I
 FatI   || TseI| |  |  |||  FokI      | |   ||       BccI       Tsp45I  BsrI |||  HphI ||   DdeI|    BcgI
    |   ||    || |  |  |||    |       | |   ||         |            |     | |||     | ||       ||       |
      tcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaata
 7501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7600
      agtaccaataccgtcgtgacgtattaagagaatgacagtacggtaggcattctacgaaaagacactgaccactcatgagttggttcagtaagactcttat

a      S  W  L  W  Q  H  C  I  I  L  L  L  S  C  H  P  *  D  A  F  L  *  L  V  S  T  Q  P  S  H  S  E  N  S-
b       H  G  Y  G  S  T  A  *  F  S  Y  C  H  A  I  R  K  M  L  F  C  D  W  *  V  L  N  Q  V  I  L  R  I  -
c        M  V  M  A  A  L  H  N  S  L  T  V  M  P  S  V  R  C  F  S  V  T  G  E  Y  S  T  K  S  F  *  E  *  -
```

Figure 7-JJ

```
                                BcgI
                             BsaHI |
                          HpaII  | |
                           NciI  | |                                                               BpuEI
                BsiEI     ScrFI  | |               HhaI                                          TaiI |
            TauI  |      TaqII|  | |             BstUI |                                        AclI | |
       Fnu4HI |   |     StyD4I||  | |             HinP1I |                DraI  BsiHKAI      HpyCH4IV | |
        AciI| |   |  HgaI   |||  | |             AciI | |                MseI|  Bsp1286I  MboII  XmnI | |
           || |   |    |    |||  | |                | | |                  ||       |        |     |   | |
     gtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttct
7601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7700
     cacatacgccgctggctcaacgagaacgggccgcagttatgccctattatggcgcggtgtatcgtcttgaaattttcacgagtagtaacettttgcaaga

a      V  C  G  D  R  V  A  L  A  R  R  Q  Y  G  I  I  P  R  H  I  A  E  L  *  K  C  S  S  L  E  N  V  L  -
b     V  Y  A  A  T  E  L  L  L  P  G  V  N  T  G  *  Y  R  A  T  *  Q  N  F  K  S  A  H  H  W  K  T  F  F  -
c      C  M  R  R  P  S  C  S  C  P  A  S  I  R  D  N  T  A  P  H  S  R  T  L  K  V  L  I  I  G  K  R  S   -

                                                                        MboII
                                                                 Bme1580I |
                                                                  BsiHKAI |
                                                                 Bsp1286I |
                                                                    Hpy8I | |
                                                                  HpyCH4V | |
                                  AlwI      BsrI                   ApaLI | | |
                         BstKTI MspA1I |  BstKTI|                  Eco57I | | |
                            DpnI|  AciI |    DpnI||               Eco57MI | | |     BstKTI
                         BstYI ||  AlwI | |BstYI ||     MaeIII      BssSI | | | |      DpnI|               SfaNI
                   SmlI  MboI ||  |PpiI| | | MboI |||  TaqI   |       PpiI | | | |  MboI ||TaqII  HphI   |   TaqII
                      |     | ||  |  || | |    | |||    |     |         |  | | | |    | ||  |      |     |     |
     tcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcacca
7701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7800
     agccccgcttttgagagttcctagaatggcgacaactctaggtcaagctacattgggtgagcacgtgggttgactagaagtcgtagaaaatgaaagtggt

a     R  G  E  N  S  Q  G  S  Y  R  C  *  D  P  V  R  C  N  P  L  V  H  P  T  D  L  Q  H  L  L  L  S  P   -
b      G  A  K  T  L  K  D  L  T  A  V  E  I  Q  F  D  V  T  H  S  C  T  Q  L  I  F  S  I  F  Y  F  H  Q  -
c     S  G  R  K  L  S  R  I  L  P  L  L  R  S  S  S  M  *  P  T  R  A  P  N  *  S  S  A  S  F  T  F  T  S -

                                       TauI
                                    AciI |                                 MboII
              HphI                Fnu4HI |                      MslI      TspGWI          EarI      SspI
                 |                     | |                         |           |              |         |
     gcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaata
7801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7900
     cgcaaagacccactcgtttttgtccttccgttttacggcgttttttcccttattcccgctgtgcctttacaacttatgagtatgagaaggaaaaagttat

a     A  F  L  G  E  Q  K  Q  E  G  K  M  P  Q  K  R  E  *  G  R  H  G  N  V  E  Y  S  Y  S  S  F  F  N  I -
b      R  F  W  V  S  K  N  R  K  A  K  C  R  K  K  G  N  K  G  D  T  E  M  L  N  T  H  T  L  P  F  S  I  -
c       V  S  G  *  A  K  T  G  R  Q  N  A  A  K  K  G  I  R  A  T  R  K  C  *  I  L  I  L  F  L  F  Q  Y  -
```

Figure 7-KK

```
                                            AciI
                                           BsrBI
                                         NlaIII |
                                          BciVI | |
                                          BsmAI | |
                                     Hpy188III| | |
                                        BspHI|| | |                                                    HhaI
                                         FatI|| | |                                                  BstUI |
                                             ||| | |                                               HinP1I |
                                                                                                   AciI | |
                                                                                                 NlaIV | | |
                                                                                                     | | | |
     ttattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaataggggttccgcgcacatttccccga
7901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8000
     aataacttcgtaaatagtcccaataacagagtactcgcctatgtataaacttacataaatctttttatttgtttatccccaaggcgcgtgtaaaggggct

a      I  E  A  F  I  R  V  I  V  S  *  A  D  T  Y  L  N  V  F  R  K  I  N  K  *  G  F  R  A  H  F  P  E  -
b     L  L  K  H  L  S  G  L  L  S  H  E  R  I  H  I  *  M  Y  L  E  K  *  T  N  R  G  S  A  H  I  S  P  K -
c      Y  *  S  I  Y  Q  G  Y  C  L  M  S  G  Y  I  F  E  C  I  *  K  N  K  Q  I  G  V  P  R  T  F  P  R  -

                                                                                     CviJI
                 DdeI                                                               HaeIII
               AatII|                                                            Eco0109I|
                TaiI|                    NlaIII                                    Sau96I|
               ZraI ||              Hpy188III |                                 BssSI    ||
              BsaHI| ||                 BspHI| |                             Hpy188III    ||
           HpyCH4IV| ||                  FatI| |MseI                           MnlI   |   ||
                   || ||                     || |   |                            |    |   ||
     aaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgtc
8001 ---------+---------+---------+---------+---------+---------+---------+---------+-- 8082
     tttcacggtggactgcagattctttggtaataatagtactgtaattggatatttttatccgcatagtgctccgggaaagcag

a      K  C  H  L  T  S  K  K  P  L  L  S  *  H  *  P  I  K  I  G  V  S  R  G  P  F  V  -
b       S  A  T  *  R  L  R  N  H  Y  Y  H  D  I  N  L  *  K  *  A  Y  H  E  A  L  S   -
c     K  V  P  P  D  V  *  E  T  I  I  I  M  T  L  T  Y  K  N  R  R  I  T  R  P  F  R  -
```

Figure 7-LL

Enzymes that do cut and were not excluded:

AarI	AatII	AccI	Acc65I	AciI	AclI	AfeI	AflIII	AhdI	AluI	AlwI	AlwNI	ApaI
ApaLI	ApoI	AseI	AvaI	AvaII	AvrII	BaeI	BamHI	BanI	BanII	BbeI	BbsI	BbvI
BccI	BceAI	BcgI	BciVI	BclI	BfaI	BfrBI	BglI	BglII	BlpI	Bme1580I	BmgBI	BmrI
BmtI	BpmI	Bpu10I	BpuEI	BsaI	BsaAI	BsaBI	BsaHI	BsaJI	BsaWI	BsaXI	BseMII	BseRI
BseYI	BsiEI	BsiHKAI	BslI	BsmI	BsmAI	BsmBI	BsmFI	Bsp1286I	BspCNI	BspHI	BspMI	BsrI
BsrBI	BsrDI	BsrFI	BsrGI	BssHII	BssSI	BstAPI	BstBI	BstEII	BstF5I	BstKTI	BstNI	BstUI
BstXI	BstYI	Bsu36I	BtgI	BtsI	Cac8I	Csp6I	CviJI	DdeI	DpnI	DraI	DraIII	DrdI
EaeI	EagI	EarI	EciI	Eco57I	EcoICRI	Eco57MI	EcoNI	EcoO109I	EcoRI	EcoRV	FalI	FatI
FauI	Fnu4HI	FokI	FspI	HaeII	HaeIII	HgaI	HhaI	Hin4I	HinP1I	HincII	HindIII	HinfI
HpaI	HpaII	HphI	Hpy8I	Hpy99I	Hpy188I	Hpy188III	HpyCH4III	HpyCH4IV	HpyCH4V	HpyF10VI	KasI	KpnI
MaeIII	MboI	MboII	MfeI	MlyI	MmeI	MnlI	MscI	MseI	MslI	MspA1I	MwoI	NaeI
NarI	NciI	NcoI	NdeI	NgoMIV	NheI	NlaIII	NlaIV	NruI	NsiI	NspI	PacI	PciI
PfoI	PleI	PpiI	PpuMI	PsiI	PspGI	PspOMI	PstI	PvuI	PvuII	RsaI	SacI	SacII
SalI	SapI	Sau96I	ScaI	ScrFI	SexAI	SfaNI	SfcI	SfoI	SmaI	SmlI	SnaBI	SpeI
SphI	SspI	StyI	StyD4I	TaiI	TaqI	TaqII	TatI	TauI	TfiI	TseI	Tsp45I	Tsp509I
TspDTI	TspGWI	TspRI	XbaI	XcmI	XhoI	XmaI	XmnI	ZraI				

Enzymes that do not cut:

AflII	AgeI	AleI	AloI	AscI	AsiSI	BbvCI	BplI	BsgI	BsiWI	BspEI	BstZ17I	ClaI
FseI	FspAI	MluI	NotI	PflMI	PmeI	PmlI	PshAI	PsrI	RsrII	SanDI	SbfI	SfiI
SgrAI	SrfI	StuI	SwaI	Tth111I								

Enzymes excluded; MinCuts: 1 MaxCuts: 100000

NONE

Figure 8-A

```
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttgaaaga gataaattga 420
acaagtatgg tcgtcccctg ttgggatgta ctattaaacc taaattgggg ttatctgcta 480
aaaactacgg tagagccgtt tatgaatgtc ttcgcggtgg acttgatttt actaaagatg 540
atgagaacgt gaactcacaa ccatttatgc gttggagaga tcgtttctta ttttgtgccg 600
aagcacttta taaagcacag gctgaaacag gtgaaatcaa agggcattac ttgaatgcta 660
ctgcaggtac atgcgaagaa atgatcaaaa gagctgtatt tgctagagaa ttgggcgttc 720
cgatcgtaat gcatgactac ttaacggggg gattcaccgc aaatactagc ttggctcatt 780
attgccgaga taatggtcta cttcttcaca tccaccgtgc aatgcatgcg gttattgata 840
gacagaagaa tcatggtatc cacttccggg tattagcaaa agcgttacgt atgtctggtg 900
gagatcatat tcactctggt accgtagtag gtaaacttga aggtgaaaga gacataactt 960
tgggctttgt tgatttactg cgtgatgatt ttgttgaaca agatcgaagt cgcggtattt 1020
atttcactca agattgggtc tctttaccag gtgttctacc cgtggcttca ggaggtattc 1080
acgtttggca tatgcctgct ctgaccgaga tctttgggga tgattccgta ctacagttcg 1140
gtggaggaac tttaggacat ccttggggta atgcgccagg tgccgtagct aatcgagtag 1200
ctctagaagc atgtgtaaaa gctcgtaatg aaggacgtga tcttgctcag gaaggtaatg 1260
aaattattcg cgaggcttgc aaatggagcc cggaactagc tgctgcttgt gaagtatgga 1320
aagagatcgt atttaatttt gcagcagtgg acgttttgga taagtaaaaa cagtagacat 1380
tagcagataa attagcagga aataaagaag gataaggaga aagaactcaa gtaattatcc 1440
ttcgttctct taattgaatt gcaattaaac tcggcccaat cttttactaa aaggattgag 1500
ccgaatacaa caaagattct attgcatata ttttgactaa gtatatactt acctagatat 1560
acaagatttg aaatacaaaa tctagaaaac taaatcaaaa tctaagactc aaatctttct 1620
attgttgtct tggatcatcg agcttgcatg cctgcagtcg actttcacag tttccattct 1680
gaaatgttct ctgtactata ataaatagta agtgaatcaa cttactacta aaaaaattag 1740
tagacttcct cttcggaata gaaatagcct atttctacat agggaaagtc gtgtgcaatg 1800
aaaaatgcaa gcacgatttg ggagaggtt ttttctctat tgtaacaagg aataattatc 1860
tactccatcc gactagttcc gggttcgagt cccgggcaac ccatatggaa actagaaagg 1920
agcaatctga gttttgattt ttcactcact tcatttacaa aattttttgg tttggtaaat 1980
tttgttgtat ggatatacaa ctgtcggggc tggcttggtt gacattggta tatagtctat 2040
attatactgt taaataacaa gccttctatt atctttctag ttaatacgtg tgcttgggag 2100
tccttgcaat ttgaataaac caagatctta ccatggccgc gcggggttct catcatcatc 2160
atcatcatgg catggcagac aacggtacta ttaccgttga ggagcttaaa caactcctgg 2220
aacaatggaa cctagtaata ggtttcctat tcctagcctg gattatgtta ctacaatttg 2280
cctattctaa tcggaacagg tttttgtaca taataaagct tgttttcctc tggctcttgt 2340
```

Figure 8-B

```
ggccagtaac acttgcttgt tttgtgcttg ctgttgtcta cagaattaat tgggtgactg 2400
gcgggattgc gattgcaatg gcttgtattg taggcttgat gtggcttagc tacttcgttg 2460
cttccttcag gctgtttgct cgtacccgct caatgtggtc attcaaccca gaaacaaaca 2520
ttcttctcaa tgtgcctctc cgggggacaa ttgtgaccag accgctcatg gaaagtgaac 2580
ttgtcattgg tgctgtgatc attcgtggtc acttgcgaat ggccggacac tccctagggc 2640
gctgtgacat taaggacctg ccaaaagaga tcactgtggc tacatcacga acgctttctt 2700
attacaaatt aggagcgtcg cagcgtgtag gcactgattc aggttttgct gcatacaacc 2760
gctaccgtat tggaaactat aaattaaata cagaccacgc cggtagcaac gacaatattg 2820
ctttgctagt acagtaagaa ttcaagccga attccagcac actggcggcc gttactagtg 2880
gatccgagct ctaattaatt aaggcttttc tgctaacata tagcaatttt tgaagaaagg 2940
aaagctagaa atacccaata tcttgctgaa gcaagatatt gggtatttct ttttttttta 3000
ttttgaatct ttctattctg aattcagtta acgacgagat ttagtatcct ttcttgcact 3060
ttcataactc gtgaaatgcc gagttggtac gaattccccc aatttgcgac ctaccatagg 3120
atttgttatg taaataggta tatgttcctt tccattatga atcgcgattg tatggccaac 3180
cattgcgggt agaatgctag atgcccggga ccacgttact attgtttctt tctcctcctt 3240
catattgagc ttttctattt ttgccaataa atgatggtac ctcccccgc cacgatcgaa 3300
cgggaatgga taagaggctt gtgggattga cgtgataggg tagggttggc tatactgctg 3360
gtggcgaact ccaggctaat aatctgaagc gcttggaagt tgtagggagg atcctatggc 3420
tcgtgaagcg gttatcgccg aagtatcgac tcaactatca gaggtagttg gcgtcatcga 3480
gcgccatctc gaaccgacgt tgctggccgt acatttgtac ggctccgcag tggatggcgg 3540
cctgaagcca cacagtgata ttgatttgct ggttacggtg accgtaaggc ttgatgaaac 3600
aacgcggcga gctttgatca acgacctttt ggaaacttcg gcttcccctg gagagagcga 3660
gattctccgc gctgtagaag tcaccattgt tgtgcacgac gacatcattc cgtggcgtta 3720
tccagctaag cgcgaactgc aatttggaga atggcagcgc aatgacattc ttgcaggtat 3780
cttcgagcca gccacgatcg acattgatct ggctatcttg ctgacaaaag caagagaaca 3840
tagcgttgcc ttggtaggtc cagcggcgga ggaactcttt gatccggttc ctgaacagga 3900
tctatttgag gcgctaaatg aaaccttaac gctatggaac tcgccgcccg actgggctgg 3960
cgatgagcga aatgtagtgc ttacgttgtc ccgcatttgg tacagcgcag taaccggcaa 4020
aatcgcgccg aaggatgtcg ctgccgactg ggcaatggag cgcctgccgg cccagtatca 4080
gcccgtcata cttgaagcta gacaggctta tcttggacaa gaagaagatc gcttggcctc 4140
gcgcgcagat cagttggaag aatttgtcca ctacgtgaaa ggcgagatca ctaaggtagt 4200
tggcaaataa ctcgagacta agtggataaa attagataga aaaaggtct aaataaaaaa 4260
gaagagaaat agaaagatca aaaatcagtt acgaaaatgc agtaattctt ctttttctt 4320
ctaattgatt gcaattaaac tcgtctcaat ctgaaaaaag attgagccga gtttaaatag 4380
attttgatac gatcatgaga cttgacaaat cgggattcct ctattctata tatttagaag 4440
atataaaggt ataatacaat aaataaatac aatctagaag ctcgatgatc cttaggattg 4500
gtatattctt ttctatcctg tagtttgtag tttccctgaa tcaagccaag tatcacacct 4560
ctttctaccc atcctgtata ttgtcccctt tgttccgtgt tgaaatagaa ccttaattta 4620
```

Figure 8-C

```
ttacttattt ttttattaaa ttttagattt gttagtgatt agatattagt attagacgag 4680
attttacgaa acaattattt ttttatttct ttataggaga ggacaaatct ctttttttcga 4740
tgcgaatttg acacgacata ggagaagccg ccctttatta aaaattatat tattttaaat 4800
aatataaagg gggttccaac atattaatat atagtgaagt gttcccccag attcagaact 4860
ttttttcaat actcacaatc cttattagtt aataatccta gtgattggat ttctatgctt 4920
agtctgatag gaaataagat attcaaataa ataattttat agcgaatgac tattcatcta 4980
ttgtattttc atgcaaatag ggggcaagaa aactctatgg aaagatggtg gtttaattcg 5040
atgttgttta agaaggagtt cgaacgcagg tgtgggctaa ataaatcaat gggcagtctt 5100
ggtcctattg aaaataccaa tgaagatcca aatcgaaaag tgaaaaacat tcatagttgg 5160
aggaatcgtg acaattctag ttgcagtaat gttgattatt tattcggcgt taaagacatt 5220
cggaatttca tctctgatga cactttttta gttagtgata ggaatggaga cagttattcc 5280
atctattttg atattgaaaa tcatattttt gagattgaca acgatcattc ttttctgagt 5340
gaactagaaa gttcttttta tagttatcga aactcgaatt atcggaataa tggatttagg 5400
ggcgaagatc cctactataa ttcttacatg tatgatactc aatatagttg gaataatcac 5460
attaatagtt gcattgatag ttatcttcag tctcaaatct gtatagatac ttccattata 5520
agtggtagtg agaattacgg tgacagttac atttataggg ccgtttgtgg tggtgaaagt 5580
cgaaatagta gtgaaaacga gggttccagt agacgaactc gcacgaaggg cagtgattta 5640
actataagag aaagttctaa tgatctcgag gtaactcaaa aatacaggca tttgtgggtt 5700
caatgcgaaa attgttatgg attaaattat aagaaatttt tgaaatcaaa aatgaatatt 5760
tgtgaacaat gtggatatca tttgaaaatg agtagttcag atagaattcg taatcatggt 5820
catagctgtt tcctgtgtga aattgttatc cgctcacaat tccacacaac atacgagccg 5880
gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag ctaactcaca ttaattgcgt 5940
tgcgctcact gcccgctttc cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg 6000
gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg 6060
actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa 6120
tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 6180
aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc 6240
ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 6300
aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 6360
cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct 6420
cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 6480
aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 6540
cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 6600
ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 6660
ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta 6720
gctcttgatc cggcaaacaa accaccgctg gtagcggtgg tttttttgtt tgcaagcagc 6780
agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 6840
acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 6900
```

Figure 8-D

```
tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 6960
agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 7020
gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg 7080
agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc 7140
cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 7200
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 7260
cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt 7320
cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc 7380
ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt 7440
tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 7500
catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt 7560
gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata 7620
gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga 7680
tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag 7740
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 7800
aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt 7860
attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga 7920
aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag 7980
aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc 8040
```

Figure 9

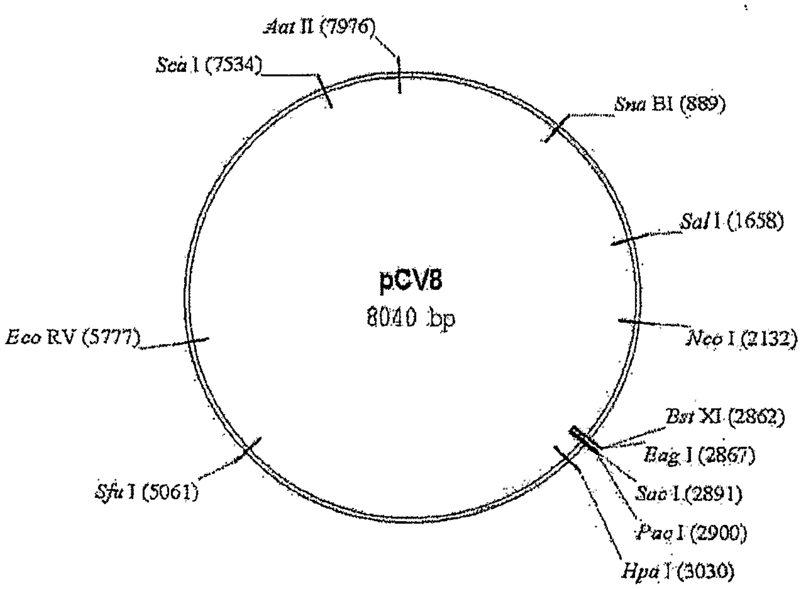

Aat II (7976)
Sca I (7534)
Sna BI (889)
Sal I (1658)
pCV8
8040 bp
Eco RV (5777)
Nco I (2132)
Bst XI (2862)
Eag I (2867)
Sfu I (5061)
Sac I (2891)
Pac I (2900)
Hpa I (3030)

Figure 10

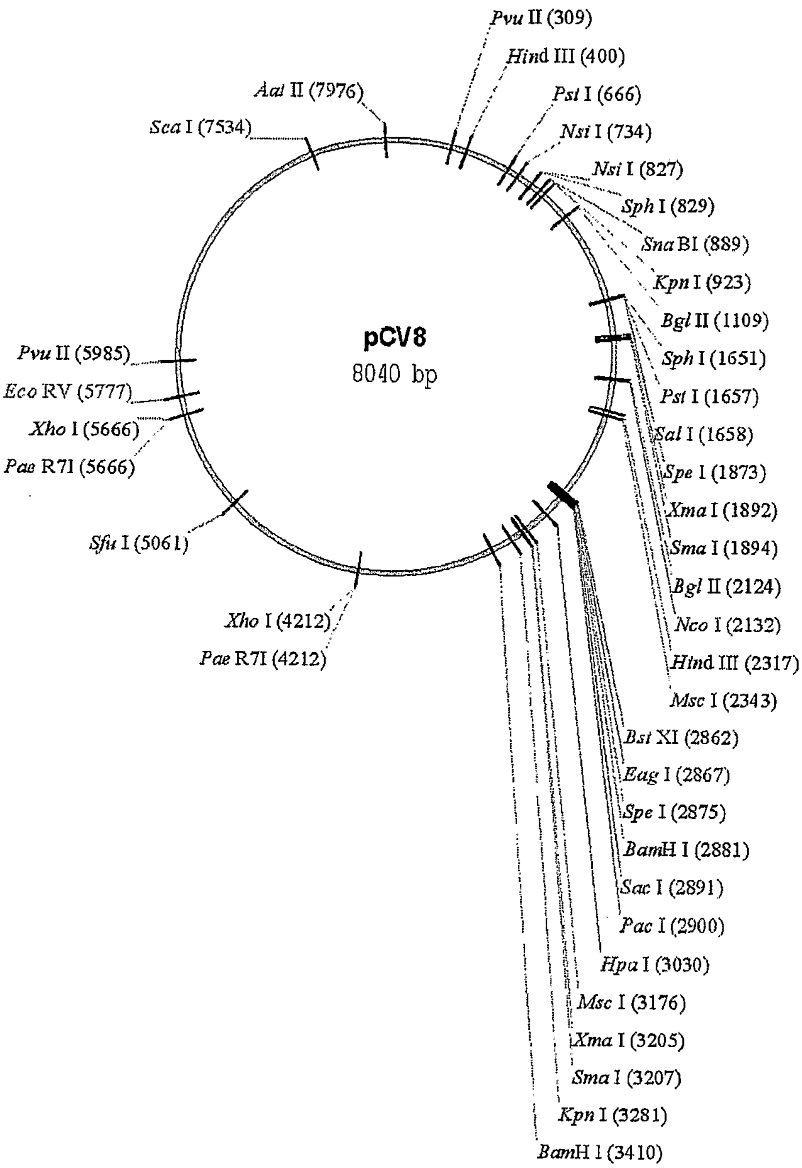
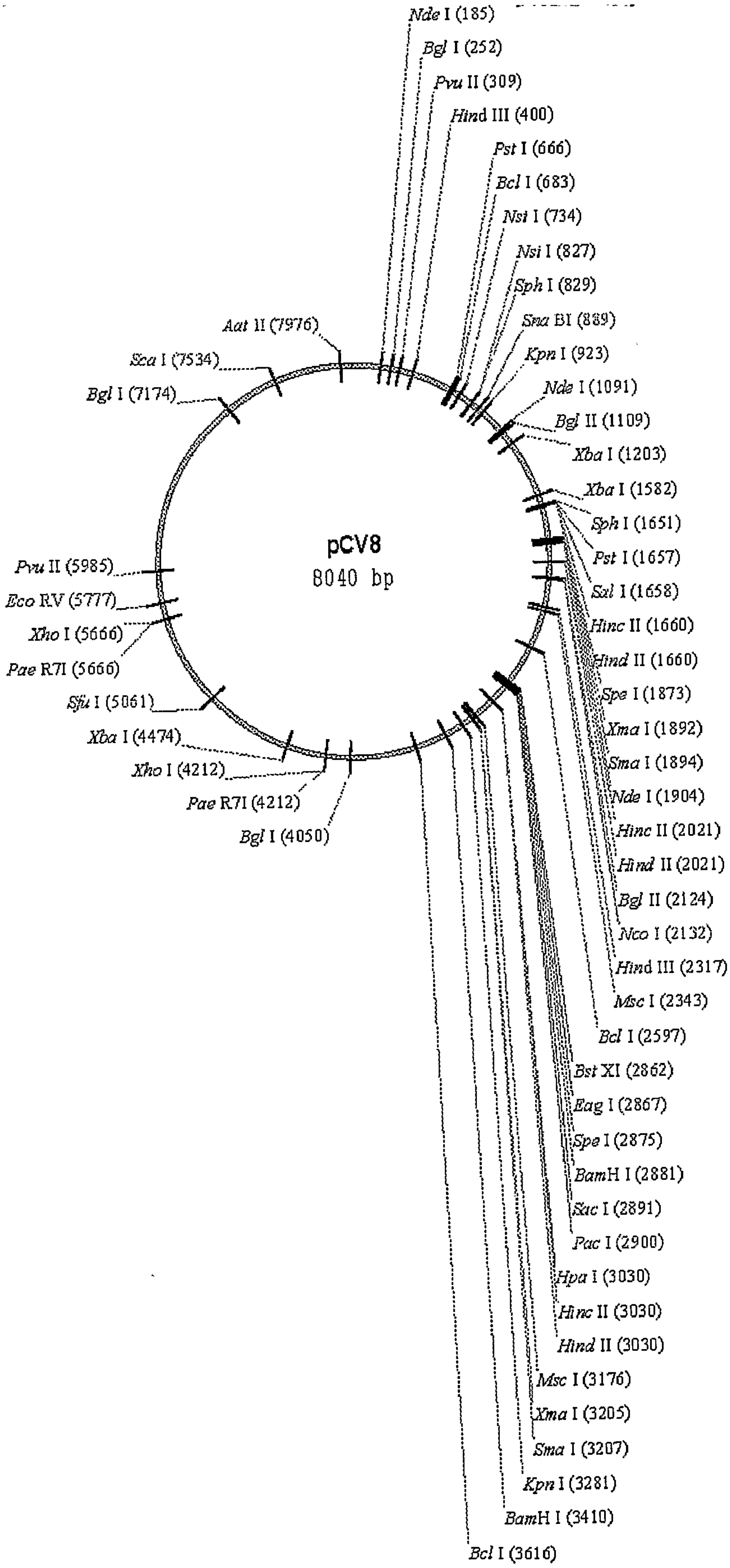

pCV8
8040 bp
Pvu II (309)
Hind III (400)
Pst I (666)
Nsi I (734)
Nsi I (827)
Sph I (829)
Sna BI (889)
Kpn I (923)
Bgl II (1109)
Sph I (1651)
Pst I (1657)
Sal I (1658)
Spe I (1873)
Xma I (1892)
Sma I (1894)
Bgl II (2124)
Nco I (2132)
Hind III (2317)
Msc I (2343)
Bst XI (2862)
Eag I (2867)
Spe I (2875)
BamH I (2881)
Sac I (2891)
Pac I (2900)
Hpa I (3030)
Msc I (3176)
Xma I (3205)
Sma I (3207)
Kpn I (3281)
BamH I (3410)
Xho I (4212)
Pae R7I (4212)
Sfu I (5061)
Pae R7I (5666)
Xho I (5666)
Eco RV (5777)
Pvu II (5985)
Sca I (7534)
Aat II (7976)

Figure 11


pCV8
8040 bp
Nde I (185)
Bgl I (252)
Pvu II (309)
Hind III (400)
Pst I (666)
Bcl I (683)
Nsi I (734)
Nsi I (827)
Sph I (829)
Sna BI (889)
Kpn I (923)
Nde I (1091)
Bgl II (1109)
Xba I (1203)
Xba I (1582)
Sph I (1651)
Pst I (1657)
Sal I (1658)
Hinc II (1660)
Hind II (1660)
Spe I (1873)
Xma I (1892)
Sma I (1894)
Nde I (1904)
Hinc II (2021)
Hind II (2021)
Bgl II (2124)
Nco I (2132)
Hind III (2317)
Msc I (2343)
Bcl I (2597)
Bst XI (2862)
Eag I (2867)
Spe I (2875)
BamH I (2881)
Sac I (2891)
Pac I (2900)
Hpa I (3030)
Hinc II (3030)
Hind II (3030)
Msc I (3176)
Xma I (3205)
Sma I (3207)
Kpn I (3281)
BamH I (3410)
Bcl I (3616)
Bgl I (4050)
Pae R7I (4212)
Xho I (4212)
Xba I (4474)
Sfu I (5061)
Pae R7I (5666)
Xho I (5666)
Eco RV (5777)
Pvu II (5985)
Bgl I (7174)
Sca I (7534)
Aat II (7976)

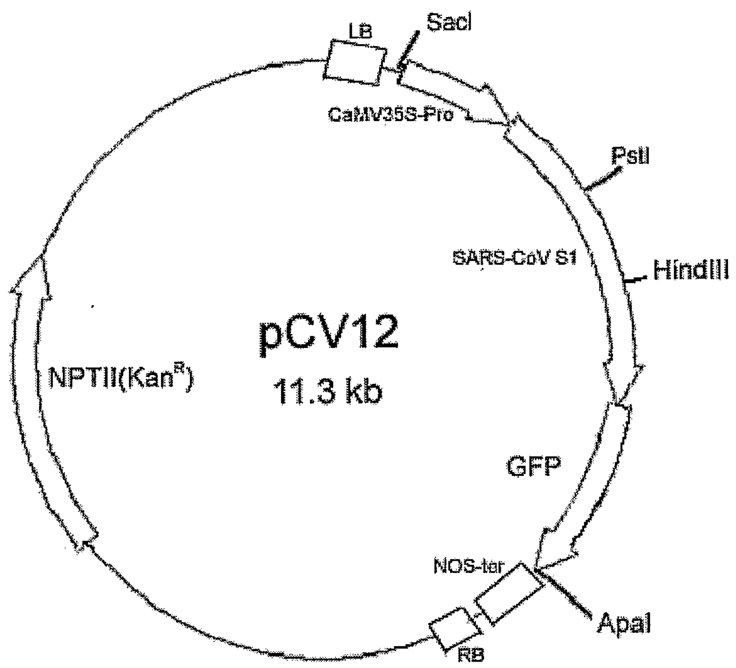

LB
SacI
CaMV35S-Pro
PstI
SARS-CoV S1
HindIII
pCV12
11.3 kb
NPTII(Kan^R)
GFP
NOS-ter
ApaI
RB

Figure 13-A

```
atggcagaca acggtactat taccgttgag gagcttaaac aactcctgga acaatggaac 60
ctagtaatag gtttcctatt cctagcctgg attatgttac tacaatttgc ctattctaat 120
cggaacaggt ttttgtacat aataaagctt gttttcctct ggctcttgtg gccagtaaca 180
cttgcttgtt ttgtgcttgc tgctgtctac agaattaatt gggtgactgg cgggattgcg 240
attgcaatgg cttgtattgt aggcttgatg tggcttagct acttcgttgc ttccttcagg 300
ctgtttgctc gtacccgctc aatgtggtca ttcaacccag aaacaaacat tcttctcaat 360
gtgcctctcc gggggacaat tgtgaccaga ccgctcatgg aaagtgaact tgtcattggt 420
gctgtgatca ttcgtggtca cttgcgaatg gccggacact ccctagggcg ctgtgacatt 480
aaggacctgc caaaagagat cactgtggct acatcacgaa cgctttctta ttacaaatta 540
ggagcgtcgc agcgtgtagg cactgattca ggttttgctg catacaaccg ctaccgtatt 600
ggaaactata aattaaatac agaccacgcc ggtagcaacg acaatattgc tttgctagta 660
cagtaa                                                            666
```

Figure 13-B

```
  1 MADNGTITVE ELKQLLEQWN LVIGFLFLAW IMLLQFAYSN RNRFLYIIKL VFLWLLWPVT
 61 LACFVLAAVY RINWVTGGIA IAMACIVGLM WLSYFVASFR LFARTRSMWS FNPETNILLN
121 VPLRGTIVTR PLMESELVIG AVIIRGHLRM AGHSLGRCDI KDLPKEITVA TSRTLSYYKL
181 GASQRVGTDS GFAAYNRYRI GNYKLNTDHA GSNDNIALLV Q
```

Figure 14-A

```
atgtttattt tcttattatt tcttactctc actagtggta gtgaccttga ccggtgcacc 60
acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gagggggggtt 120
tactatcctg atgaaatttt tagatcagac actctttatt taactcagga tttatttctt 180
ccattttatt ctaatgttac agggtttcat actattaatc atacgtttgg caaccctgtc 240
ataccttttta aggatggtat ttattttgct gccacagaga aatcaaatgt tgtccgtggt 300
tgggtttttg gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct 360
actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctgtt 420
tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaaa 540
cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggctat 600
caacctatag atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt 660
aagttgcctc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct 720
gctcaagaca tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact 780
acatttatgc tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa 840
aatccacttg ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaatttac 900
cagacctcta atttcagggt tgttccctca ggagatgttg tgagattccc taatattaca 960
aacttgtgtc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg 1020
gagagaaaaa aaatttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt 1080
ttttcaacct ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc 1140
aatgtctatg cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgccagga 1200
caaactggtg ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc 1260
cttgcttgga atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat 1320
aggtatctta gacatggcaa gcttaggccc tttgagagag acatatctaa tgtgcctttc 1380
tcccctgatg gcaaaccttg caccccacct gctcttaatt gttattggcc attaaatgat 1440
tatggttttt acaccactac tggcattggc taccaacctt acagagttgt agtactttct 1500
tttgaacttt taaatgcacc ggccacggtt tgtggaccaa aattatccac tgaccttatt 1560
aagaaccagt gtgtcaattt taattttaat ggactcactg gtactggtgt gttaactcct 1620
tcttcaaaga gatttcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat 1680
tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgcgc ttttgggggt 1740
gtaagtgtaa ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat 1800
gttaactgca ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc 1860
atatattcta ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag 1920
catgtcgaca cttcttatga gtgcgacatt cctattggag ctggcatttg tgctagttac 1980
catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct 2040
ttaggtgctg atagttcaat tgcttactct aataacacca ttgctatacc tactaacttt 2100
tcaattagca ttactacaga agtaatgcct gtttctatgg ctaaaacctc cgtagattgt 2160
aatatgtaca tctgcggaga ttctactgaa tgtgctaatt tgcttctcca atatggtagc 2220
ttttgcacac aactaaatcg tgcactctca ggtattgctg ctgaacagga tcgcaacaca 2280
```

Figure 14-B

```
cgtgaagtgt tcgctcaagt caaacaaatg tacaaaaccc caactttgaa atattttggt 2340
ggttttaatt tttcacaaat attacctgac cctctaaagc caactaagag gtcttttatt 2400
gaggacttgc tctttaataa ggtgacactc gctgatgctg gcttcatgaa gcaatatggc 2460
gaatgcctag gtgatattaa tgctagagat ctcatttgtg cgcagaagtt caatggactt 2520
acagtgttgc cacctctgct cactgatgat atgattgctg cctacactgc tgctctagtt 2580
agtggtactg ccactgctgg atggacattt ggtgctggcg ctgctcttca aataccttttt 2640
gctatgcaaa tggcatatag gttcaatggc attggagtta cccaaaatgt tctctatgag 2700
aaccaaaaac aaatcgccaa ccaatttaac aaggcgatta gtcaaattca agaatcactt 2760
acaacaacat caactgcatt gggcaagctg caagacgttg ttaaccagaa tgctcaagca 2820
ttaaacacac ttgttaaaca acttagctct aattttggtg caatttcaag tgtgctaaat 2880
gatatccttt cgcgacttga taaagtcgag gcggaggtac aaattgacag gttaattaca 2940
ggcagacttc aaagccttca aacctatgta acacaacaac taatcagggc tgctgaaatc 3000
agggcttctg ctaatcttgc tgctactaaa atgtctgagt gtgttcttgg acaatcaaaa 3060
agagttgact tttgtggaaa gggctaccac cttatgtcct tcccacaagc agccccgcat 3120
ggtgttgtct tcctacatgt cacgtatgtg ccatcccagg agaggaactt caccacagcg 3180
ccagcaattt gtcatgaagg caaagcatac ttccctcgtg aaggtgtttt tgtgtttaat 3240
ggcacttctt ggtttattac acagaggaac ttcttttctc cacaaataat tactacagac 3300
aatacatttg tctcaggaaa ttgtgatgtc gttattggca tcattaacaa cacagtttat 3360
gatcctctgc aacctgagct tgactcattc aaagaagagc tggacaagta cttcaaaaat 3420
catacatcac cagatgttga tcttggcgac atttcaggca ttaacgcttc tgtcgtcaac 3480
attcaaaaag aaattgaccg cctcaatgag gtcgctaaaa atttaaatga atcactcatt 3540
gaccttcaag aattgggaaa atatgagcaa tatattaaat ggccttggta tgtttggctc 3600
ggcttcattg ctggactaat tgccatcgtc atggttacaa tcttgctttg ttgcatgact 3660
agttgttgca gttgcctcaa gggtgcatgc tcttgtggtt cttgctgcaa gtttgatgag 3720
gatgactctg agccagttct caagggtgtc aaattacatt acacataa 3768
```

Figure 15

```
   1 MFIFLLFLTL TSGSDLDRCT TFDDVQAPNY TQHTSSMRGV YYPDEIFRSD TLYLTQDLFL
  61 PFYSNVTGFH TINHTFGNPV IPFKDGIYFA ATEKSNVVRG WVFGSTMNNK SQSVIIINNS
 121 TNVVIRACNF ELCDNPFFAV SKPMGTQTHT MIFDNAFNCT FEYISDAFSL DVSEKSGNFK
 181 HLREFVFKNK DGFLYVYKGY QPIDVVRDLP SGFNTLKPIF KLPLGINITN FRAILTAFSP
 241 AQDIWGTSAA AYFVGYLKPT TFMLKYDENG TITDAVDCSQ NPLAELKCSV KSFEIDKGIY
 301 QTSNFRVVPS GDVVRFPNIT NLCPFGEVFN ATKFPSVYAW ERKKISNCVA DYSVLYNSTF
 361 FSTFKCYGVS ATKLNDLCFS NVYADSFVVK GDDVRQIAPG QTGVIADYNY KLPDDFMGCV
 421 LAWNTRNIDA TSTGNYNYKY RYLRHGKLRP FERDISNVPF SPDGKPCTPP ALNCYWPLND
 481 YGFYTTTGIG YQPYRVVVLS FELLNAPATV CGPKLSTDLI KNQCVNFNFN GLTGTGVLTP
 541 SSKRFQPFQQ FGRDVSDFTD SVRDPKTSEI LDISPCSFGG VSVITPGTNA SSEVAVLYQD
 601 VNCTDVSTAI HADQLTPAWR IYSTGNNVFQ TQAGCLIGAE HVDTSYECDI PIGAGICASY
 661 HTVSLLRSTS QKSIVAYTMS LGADSSIAYS NNTIAIPTNF SISITTEVMP VSMAKTSVDC
 721 NMYICGDSTE CANLLLQYGS FCTQLNRALS GIAAEQDRNT REVFAQVKQM YKTPTLKYFG
 781 GFNFSQILPD PLKPTKRSFI EDLLFNKVTL ADAGFMKQYG ECLGDINARD LICAQKFNGL
 841 TVLPPLLTDD MIAAYTAALV SGTATAGWTF GAGAALQIPF AMQMAYRFNG IGVTQNVLYE
 901 NQKQIANQFN KAISQIQESL TTTSTALGKL QDVVNQNAQA LNTLVKQLSS NFGAISSVLN
 961 DILSRLDKVE AEVQIDRLIT GRLQSLQTYV TQQLIRAAEI RASANLAATK MSECVLGQSK
1021 RVDFCGKGYH LMSFPQAAPH GVVFLHVTYV PSQERNFTTA PAICHEGKAY FPREGVFVFN
1081 GTSWFITQRN FFSPQIITTD NTFVSGNCDV VIGIINNTVY DPLQPELDSF KEELDKYFKN
1141 HTSPDVDLGD ISGINASVVN IQKEIDRLNE VAKNLNESLI DLQELGKYEQ YIKWPWYVWL
1201 GFIAGLIAIV MVTILLCCMT SCCSCLKGAC SCGSCCKFDE DDSEPVLKGV KLHYT
```

Figure 16-A

```
  1 - ATATTAGGTTTTTACCTACCCAGGAAAAGCCAACCAACCTCGATCTCTTGTAGATCTGTT - 60
    - I  L  G  F  Y  L  P  R  K  S  Q  P  T  S  I  S  C  R  S  V
    -  Y  *  V  F  T  Y  P  G  K  A  N  Q  P  R  S  L  V  D  L  F
    -   I  R  F  L  P  T  Q  E  K  P  T  N  L  D  L  L  *  I  C  S
 61 - CTCTAAACGAACTTTAAAATCTGTGTAGCTGTCGCTCGGCTGCATGCCTAGTGCACCTAC - 120
    - L  *  T  N  F  K  I  C  V  A  V  A  R  L  H  A  *  C  T  Y
    -  S  K  R  T  L  K  S  V  *  L  S  L  G  C  M  P  S  A  P  T
    -   L  N  E  L  *  N  L  C  S  C  R  S  A  A  C  L  V  H  L  R
121 - GCAGTATAAACAATAATAAATTTTACTGTCGTTGACAAGAAACGAGTAACTCGTCCCTCT - 180
    - A  V  *  T  I  I  N  F  T  V  V  D  K  K  R  V  T  R  P  S
    -  Q  Y  K  Q  *  *  I  L  L  S  L  T  R  N  E  *  L  V  P  L
    -   S  I  N  N  N  K  F  Y  C  R  *  Q  E  T  S  N  S  S  L  F
181 - TCTGCAGACTGCTTACGGTTTCGTCCGTGTTGCAGTCGATCATCAGCATACCTAGGTTTC - 240
    - S  A  D  C  L  R  F  R  P  C  C  S  R  S  S  A  Y  L  G  F
    -  L  Q  T  A  Y  G  F  V  R  V  A  V  D  H  Q  H  T  *  V  S
    -   C  R  L  L  T  V  S  S  V  L  Q  S  I  I  S  I  P  R  F  R
241 - GTCCGGGTGTGACCGAAAGGTAAGATGGAGAGCCTTGTTCTTGGTGTCAACGAGAAAACA - 300
    - V  R  V  *  P  K  G  K  M  E  S  L  V  L  G  V  N  E  K  T
    -  S  G  C  D  R  K  V  R  W  R  A  L  F  L  V  S  T  R  K  H
    -   P  G  V  T  E  R  *  D  G  E  P  C  S  W  C  Q  R  E  N  T
301 - CACGTCCAACTCAGTTTGCCTGTCCTTCAGGTTAGAGACGTGCTAGTGCGTGGCTTCGGG - 360
    - H  V  Q  L  S  L  P  V  L  Q  V  R  D  V  L  V  R  G  F  G
    -  T  S  N  S  V  C  L  S  F  R  L  E  T  C  *  C  V  A  S  G
    -   R  P  T  Q  F  A  C  P  S  G  *  R  R  A  S  A  W  L  R  G
361 - GACTCTGTGGAAGAGGCCCTATCGGAGGCACGTGAACACCTCAAAAATGGCACTTGTGGT - 420
    - D  S  V  E  E  A  L  S  E  A  R  E  H  L  K  N  G  T  C  G
    -  T  L  W  K  R  P  Y  R  R  H  V  N  T  S  K  M  A  L  V  V
    -   L  C  G  R  G  P  I  G  G  T  *  T  P  Q  K  W  H  L  W  S
421 - CTAGTAGAGCTGGAAAAAGGCGTACTGCCCCAGCTTGAACAGCCCTATGTGTTCATTAAA - 480
    - L  V  E  L  E  K  G  V  L  P  Q  L  E  Q  P  Y  V  F  I  K
    -  *  *  S  W  K  K  A  Y  C  P  S  L  N  S  P  M  C  S  L  N
    -   S  R  A  G  K  R  R  T  A  P  A  *  T  A  L  C  V  H  *  T
481 - CGTTCTGATGCCTTAAGCACCAATCACGGCCACAAGGTCGTTGAGCTGGTTGCAGAAATG - 540
    - R  S  D  A  L  S  T  N  H  G  H  K  V  V  E  L  V  A  E  M
    -  V  L  M  P  *  A  P  I  T  A  T  R  S  L  S  W  L  Q  K  W
    -   F  *  C  L  K  H  Q  S  R  P  Q  G  R  *  A  G  C  R  N  G
541 - GACGGCATTCAGTACGGTCGTAGCGGTATAACACTGGGAGTACTCGTGCCACATGTGGGC - 600
    - D  G  I  Q  Y  G  R  S  G  I  T  L  G  V  L  V  P  H  V  G
    -  T  A  F  S  T  V  V  A  V  *  H  W  E  Y  S  C  H  M  W  A
    -   R  H  S  V  R  S  *  R  Y  N  T  G  S  T  R  A  T  C  G  R
601 - GAAACCCCAATTGCATACCGCAATGTTCTTCTTCGTAAGAACGGTAATAAGGGAGCCGGT - 660
    - E  T  P  I  A  Y  R  N  V  L  L  R  K  N  G  N  K  G  A  G
    -  K  P  Q  L  H  T  A  M  F  F  F  V  R  T  V  I  R  E  P  V
    -   N  P  N  C  I  P  Q  C  S  S  S  *  E  R  *  *  G  S  R  W
661 - GGTCATAGCTATGGCATCGATCTAAAGTCTTATGACTTAGGTGACGAGCTTGGCACTGAT - 720
    - G  H  S  Y  G  I  D  L  K  S  Y  D  L  G  D  E  L  G  T  D
    -  V  I  A  M  A  S  I  *  S  L  M  T  *  V  T  S  L  A  L  I
    -   S  *  L  W  H  R  S  K  V  L  *  L  R  *  R  A  W  H  *  S
```

Figure 16-B

```
721 - CCCATTGAAGATTATGAACAAAACTGGAACACTAAGCATGGCAGTGGTGCACTCCGTGAA - 780
    - P  I  E  D  Y  E  Q  N  W  N  T  K  H  G  S  G  A  L  R  E
    -  P  L  K  I  M  N  K  T  G  T  L  S  M  A  V  V  H  S  V  N
    -   H  *  R  L  *  T  K  L  E  H  *  A  W  Q  W  C  T  P  *  T
781 - CTCACTCGTGAGCTCAATGGAGGTGCAGTCACTCGCTATGTCGACAACAATTTCTGTGGC - 840
    - L  T  R  E  L  N  G  G  A  V  T  R  Y  V  D  N  N  F  C  G
    -  S  L  V  S  S  M  E  V  Q  S  L  A  M  S  T  T  I  S  V  A
    -   H  S  *  A  Q  W  R  C  S  H  S  L  C  R  Q  Q  F  L  W  P
841 - CCAGATGGGTACCCTCTTGATTGCATCAAAGATTTTCTCGCACGCGCGGGCAAGTCAATG - 900
    - P  D  G  Y  P  L  D  C  I  K  D  F  L  A  R  A  G  K  S  M
    -  Q  M  G  T  L  L  I  A  S  K  I  F  S  H  A  R  A  S  Q  C
    -   R  W  V  P  S  *  L  H  Q  R  F  S  R  T  R  G  Q  V  N  V
901 - TGCACTCTTTCCGAACAACTTGATTACATCGAGTCGAAGAGAGGTGTCTACTGCTGCCGT - 960
    - C  T  L  S  E  Q  L  D  Y  I  E  S  K  R  G  V  Y  C  C  R
    -  A  L  F  P  N  N  L  I  T  S  S  R  R  E  V  S  T  A  A  V
    -   H  S  F  R  T  T  *  L  H  R  V  E  E  R  C  L  L  L  P  *
961 - GACCATGAGCATGAAATTGCCTGGTTCACTGAGCGCTCTGATAAGAGCTACGAGCACCAG -
1020
    - D  H  E  H  E  I  A  W  F  T  E  R  S  D  K  S  Y  E  H  Q
    -  T  M  S  M  K  L  P  G  S  L  S  A  L  I  R  A  T  S  T  R
    -   P  *  A  *  N  C  L  V  H  *  A  L  *  *  E  L  R  A  P  D
1021 - ACACCCTTCGAAATTAAGAGTGCCAAGAAATTTGACACTTTCAAAGGGGAATGCCCAAAG -
1080
    - T  P  F  E  I  K  S  A  K  K  F  D  T  F  K  G  E  C  P  K
    -  H  P  S  K  L  R  V  P  R  N  L  T  L  S  K  G  N  A  Q  S
    -   T  L  R  N  *  E  C  Q  E  I  *  H  F  Q  R  G  M  P  K  V
1081 - TTTGTGTTTCCTCTTAACTCAAAAGTCAAAGTCATTCAACCACGTGTTGAAAAGAAAAAG -
1140
    - F  V  F  P  L  N  S  K  V  K  V  I  Q  P  R  V  E  K  K  K
    -  L  C  F  L  L  T  Q  K  S  K  S  F  N  H  V  L  K  R  K  R
    -   C  V  S  S  *  L  K  S  Q  S  H  S  T  T  C  *  K  E  K  D
1141 - ACTGAGGGTTTCATGGGGCGTATACGCTCTGTGTACCCTGTTGCATCTCCACAGGAGTGT -
1200
    - T  E  G  F  M  G  R  I  R  S  V  Y  P  V  A  S  P  Q  E  C
    -  L  R  V  S  W  G  V  Y  A  L  C  T  L  L  H  L  H  R  S  V
    -   *  G  F  H  G  A  Y  T  L  C  V  P  C  C  I  S  T  G  V  *
1201 - AACAATATGCACTTGTCTACCTTGATGAAATGTAATCATTGCGATGAAGTTTCATGGCAG -
1260
    - N  N  M  H  L  S  T  L  M  K  C  N  H  C  D  E  V  S  W  Q
    -  T  I  C  T  C  L  P  *  *  N  V  I  I  A  M  K  F  H  G  R
    -   Q  Y  A  L  V  Y  L  D  E  M  *  S  L  R  *  S  F  M  A  D
1261 - ACGTGCGACTTTCTGAAAGCCACTTGTGAACATTGTGGCACTGAAAATTTAGTTATTGAA -
1320
    - T  C  D  F  L  K  A  T  C  E  H  C  G  T  E  N  L  V  I  E
    -  R  A  T  F  *  K  P  L  V  N  I  V  A  L  K  I  *  L  L  K
    -   V  R  L  S  E  S  H  L  *  T  L  W  H  *  K  F  S  Y  *  R
1321 - GGACCTACTACATGTGGGTACCTACCTACTAATGCTGTAGTGAAAATGCCATGTCCTGCC -
1380
    - G  P  T  T  C  G  Y  L  P  T  N  A  V  V  K  M  P  C  P  A
    -  D  L  L  H  V  G  T  Y  L  L  M  L  *  *  K  C  H  V  L  P
    -   T  Y  Y  M  W  V  P  T  Y  *  C  C  S  E  N  A  M  S  C  L
1381 - TGTCAAGACCCAGAGATTGGACCTGAGCATAGTGTTGCAGATTATCACAACCACTCAAAC -
1440
    - C  Q  D  P  E  I  G  P  E  H  S  V  A  D  Y  H  N  H  S  N
    -  V  K  T  Q  R  L  D  L  S  I  V  L  Q  I  I  T  T  T  Q  T
    -   S  R  P  R  D  W  T  *  A  *  C  C  R  L  S  Q  P  L  K  H
1441 - ATTGAAACTCGACTCCGCAAGGGAGGTAGGACTAGATGTTTTGGAGGCTGTGTGTTTGCC -
1500
```

-I E T R L R K G G R T R C F G G C V F A
- L K L D S A R E V G L D V L E A V C L P
- * N S T P Q G R * D * M F W R L C V C L

Figure 16-C

```
 1501 - TATGTTGGCTGCTATAATAAGCGTGCCTACTGGGTTCCTCGTGCTAGTGCTGATATTGGC -
1560
      - Y  V  G  C  Y  N  K  R  A  Y  W  V  P  R  A  S  A  D  I  G
      -  M  L  A  A  I  I  S  V  P  T  G  F  L  V  L  V  L  I  L  A
      -   C  W  L  L  *  *  A  C  L  L  G  S  S  C  *  C  *  Y  W  L
 1561 - TCAGGCCATACTGGCATTACTGGTGACAATGTGGAGACCTTGAATGAGGATCTCCTTGAG -
1620
      - S  G  H  T  G  I  T  G  D  N  V  E  T  L  N  E  D  L  L  E
      -  Q  A  I  L  A  L  L  V  T  M  W  R  P  *  M  R  I  S  L  R
      -   R  P  Y  W  H  Y  W  *  Q  C  G  D  L  E  *  G  S  P  *  D
 1621 - ATACTGAGTCGTGAACGTGTTAACATTAACATTGTTGGCGATTTTCATTTGAATGAAGAG -
1680
      - I  L  S  R  E  R  V  N  I  N  I  V  G  D  F  H  L  N  E  E
      -  Y  *  V  V  N  V  L  T  L  T  L  L  A  I  F  I  *  M  K  R
      -   T  E  S  *  T  C  *  H  *  H  C  W  R  F  S  F  E  *  R  G
 1681 - GTTGCCATCATTTTGGCATCTTTCTCTGCTTCTACAAGTGCCTTTATTGACACTATAAAG -
1740
      - V  A  I  I  L  A  S  F  S  A  S  T  S  A  F  I  D  T  I  K
      -  L  P  S  F  W  H  L  S  L  L  L  Q  V  P  L  L  T  L  *  R
      -   C  H  H  F  G  I  F  L  C  F  Y  K  C  L  Y  *  H  Y  K  E
 1741 - AGTCTTGATTACAAGTCTTTCAAAACCATTGTTGAGTCCTGCGGTAACTATAAAGTTACC -
1800
      - S  L  D  Y  K  S  F  K  T  I  V  E  S  C  G  N  Y  K  V  T
      -  V  L  I  T  S  L  S  K  P  L  L  S  P  A  V  T  I  K  L  P
      -   S  *  L  Q  V  F  Q  N  H  C  *  V  L  R  *  L  *  S  Y  Q
 1801 - AAGGGAAAGCCCGTAAAAGGTGCTTGGAACATTGGACAACAGAGATCAGTTTTAACACCA -
1860
      - K  G  K  P  V  K  G  A  W  N  I  G  Q  Q  R  S  V  L  T  P
      -  R  E  S  P  *  K  V  L  G  T  L  D  N  R  D  Q  F  *  H  H
      -   G  K  A  R  K  R  C  L  E  H  W  T  T  E  I  S  F  N  T  T
 1861 - CTGTGTGGTTTTCCCTCACAGGCTGCTGGTGTTATCAGATCAATTTTTGCGCGCACACTT -
1920
      - L  C  G  F  P  S  Q  A  A  G  V  I  R  S  I  F  A  R  T  L
      -  C  V  V  F  P  H  R  L  L  V  L  S  D  Q  F  L  R  A  H  L
      -   V  W  F  S  L  T  G  C  W  C  Y  Q  I  N  F  C  A  H  T  *
 1921 - GATGCAGCAAACCACTCAATTCCTGATTTGCAAAGAGCAGCTGTCACCATACTTGATGGT -
1980
      - D  A  A  N  H  S  I  P  D  L  Q  R  A  A  V  T  I  L  D  G
      -  M  Q  Q  T  T  Q  F  L  I  C  K  E  Q  L  S  P  Y  L  M  V
      -   C  S  K  P  L  N  S  *  F  A  K  S  S  C  H  H  T  *  W  Y
 1981 - ATTTCTGAACAGTCATTACGTCTTGTCGACGCCATGGTTTATACTTCAGACCTGCTCACC -
2040
      - I  S  E  Q  S  L  R  L  V  D  A  M  V  Y  T  S  D  L  L  T
      -  F  L  N  S  H  Y  V  L  S  T  P  W  F  I  L  Q  T  C  S  P
      -   F  *  T  V  I  T  S  C  R  R  H  G  L  Y  F  R  P  A  H  Q
 2041 - AACAGTGTCATTATTATGGCATATGTAACTGGTGGTCTTGTACAACAGACTTCTCAGTGG -
2100
      - N  S  V  I  I  M  A  Y  V  T  G  G  L  V  Q  Q  T  S  Q  W
      -  T  V  S  L  L  W  H  M  *  L  V  V  L  Y  N  R  L  L  S  G
      -   Q  C  H  Y  Y  G  I  C  N  W  W  S  C  T  T  D  F  S  V  V
 2101 - TTGTCTAATCTTTTGGGCACTACTGTTGAAAAACTCAGGCCTATCTTTGAATGGATTGAG -
2160
      - L  S  N  L  L  G  T  T  V  E  K  L  R  P  I  F  E  W  I  E
      -  C  L  I  F  W  A  L  L  L  K  N  S  G  L  S  L  N  G  L  R
      -   V  *  S  F  G  H  Y  C  *  K  T  Q  A  Y  L  *  M  D  *  G
 2161 - GCGAAACTTAGTGCAGGAGTTGAATTTCTCAAGGATGCTTGGGAGATTCTCAAATTTCTC -
2220
      - A  K  L  S  A  G  V  E  F  L  K  D  A  W  E  I  L  K  F  L
```

```
     - R  N  L  V  Q  E  L  N  F  S  R  M  L  G  R  F  S  N  F  S
     -  E  T  *  C  R  S  *  I  S  Q  G  C  L  G  D  S  Q  I  S  H
2221 - ATTACAGGTGTTTTTGACATCGTCAAGGGTCAAATACAGGTTGCTTCAGATAACATCAAG -
2280
     - I  T  G  V  F  D  I  V  K  G  Q  I  Q  V  A  S  D  N  I  K
     - L  Q  V  F  L  T  S  S  R  V  K  Y  R  L  L  Q  I  T  S  R
     -  Y  R  C  F  *  H  R  Q  G  S  N  T  G  C  F  R  *  H  Q  G
```

Figure 16-D

```
2281 - GATTGTGTAAAATGCTTCATTGATGTTGTTAACAAGGCACTCGAAATGTGCATTGATCAA -
2340
     - D  C  V  K  C  F  I  D  V  V  N  K  A  L  E  M  C  I  D  Q
     -  I  V  *  N  A  S  L  M  L  L  T  R  H  S  K  C  A  L  I  K
     -   L  C  K  M  L  H  *  C  C  *  Q  G  T  R  N  V  H  *  S  S
2341 - GTCACTATCGCTGGCGCAAAGTTGCGATCACTCAACTTAGGTGAAGTCTTCATCGCTCAA -
2400
     - V  T  I  A  G  A  K  L  R  S  L  N  L  G  E  V  F  I  A  Q
     -  S  L  S  L  A  Q  S  C  D  H  S  T  *  V  K  S  S  S  L  K
     -   H  Y  R  W  R  K  V  A  I  T  Q  L  R  *  S  L  H  R  S  K
2401 - AGCAAGGGACTTTACCGTCAGTGTATACGTGGCAAGGAGCAGCTGCAACTACTCATGCCT -
2460
     - S  K  G  L  Y  R  Q  C  I  R  G  K  E  Q  L  Q  L  L  M  P
     -  A  R  D  F  T  V  S  V  Y  V  A  R  S  S  C  N  Y  S  C  L
     -   Q  G  T  L  P  S  V  Y  T  W  Q  G  A  A  A  T  T  H  A  S
2461 - CTTAAGGCACCAAAAGAAGTAACCTTTCTTGAAGGTGATTCACATGACACAGTACTTACC -
2520
     - L  K  A  P  K  E  V  T  F  L  E  G  D  S  H  D  T  V  L  T
     -  L  R  H  Q  K  K  *  P  F  L  K  V  I  H  M  T  Q  Y  L  P
     -   *  G  T  K  R  S  N  L  S  *  R  *  F  T  *  H  S  T  Y  L
2521 - TCTGAGGAGGTTGTTCTCAAGAACGGTGAACTCGAAGCACTCGAGACGCCCGTTGATAGC -
2580
     - S  E  E  V  V  L  K  N  G  E  L  E  A  L  E  T  P  V  D  S
     -  L  R  R  L  F  S  R  T  V  N  S  K  H  S  R  R  P  L  I  A
     -   *  G  G  C  S  Q  E  R  *  T  R  S  T  R  D  A  R  *  *  L
2581 - TTCACAAATGGAGCTATCGTCGGCACACCAGTCTGTGTAAATGGCCTCATGCTCTTAGAG -
2640
     - F  T  N  G  A  I  V  G  T  P  V  C  V  N  G  L  M  L  L  E
     -  S  Q  M  E  L  S  S  A  H  Q  S  V  *  M  A  S  C  S  *  R
     -   H  K  W  S  Y  R  R  H  T  S  L  C  K  W  P  H  A  L  R  D
2641 - ATTAAGGACAAAGAACAATACTGCGCATTGTCTCCTGGTTTACTGGCTACAAACAATGTC -
2700
     - I  K  D  K  E  Q  Y  C  A  L  S  P  G  L  L  A  T  N  N  V
     -  L  R  T  K  N  N  T  A  H  C  L  L  V  Y  W  L  Q  T  M  S
     -   *  G  Q  R  T  I  L  R  I  V  S  W  F  T  G  Y  K  Q  C  L
2701 - TTTCGCTTAAAAGGGGGTGCACCAATTAAAGGTGTAACCTTTGGAGAAGATACTGTTTGG -
2760
     - F  R  L  K  G  G  A  P  I  K  G  V  T  F  G  E  D  T  V  W
     -  F  A  *  K  G  V  H  Q  L  K  V  *  P  L  E  K  I  L  F  G
     -   S  L  K  R  G  C  T  N  *  R  C  N  L  W  R  R  Y  C  L  G
2761 - GAAGTTCAAGGTTACAAGAATGTGAGAATCACATTTGAGCTTGATGAACGTGTTGACAAA -
2820
     - E  V  Q  G  Y  K  N  V  R  I  T  F  E  L  D  E  R  V  D  K
     -  K  F  K  V  T  R  M  *  E  S  H  L  S  L  M  N  V  L  T  K
     -   S  S  R  L  Q  E  C  E  N  H  I  *  A  *  *  T  C  *  Q  S
2821 - GTGCTTAATGAAAAGTGCTCTGTCTACACTGTTGAATCCGGTACCGAAGTTACTGAGTTT -
2880
     - V  L  N  E  K  C  S  V  Y  T  V  E  S  G  T  E  V  T  E  F
     -  C  L  M  K  S  A  L  S  T  L  L  N  P  V  P  K  L  L  S  L
     -   A  *  *  K  V  L  C  L  H  C  *  I  R  Y  R  S  Y  *  V  C
2881 - GCATGTGTTGTAGCAGAGGCTGTTGTGAAGACTTTACAACCAGTTTCTGATCTCCTTACC -
2940
     - A  C  V  V  A  E  A  V  V  K  T  L  Q  P  V  S  D  L  L  T
     -  H  V  L  *  Q  R  L  L  *  R  L  Y  N  Q  F  L  I  S  L  P
     -   M  C  C  S  R  G  C  C  E  D  F  T  T  S  F  *  S  P  Y  Q
2941 - AACATGGGTATTGATCTTGATGAGTGGAGTGTAGCTACATTCTACTTATTTGATGATGCT -
3000
     - N  M  G  I  D  L  D  E  W  S  V  A  T  F  Y  L  F  D  D  A
     -  T  W  V  L  I  L  M  S  G  V  *  L  H  S  T  Y  L  M  M  L
```

```
     -  H  G  Y  *  S  *  *  V  E  C  S  Y  I  L  L  I  *  *  C  W
3001 - GGTGAAGAAAACTTTTCATCACGTATGTATTGTTCCTTTTACCCTCCAGATGAGGAAGAA -
3060
     - G  E  E  N  F  S  S  R  M  Y  C  S  F  Y  P  P  D  E  E  E
     -  V  K  K  T  F  H  H  V  C  I  V  P  F  T  L  Q  M  R  K  K
     -   *  R  K  L  F  I  T  Y  V  L  F  L  L  P  S  R  *  G  R  R
```

Figure 16-E

```
3061 - GAGGACGATGCAGAGTGTGAGGAAGAAGAAATTGATGAAACCTGTGAACATGAGTACGGT -
3120
       - E  D  D  A  E  C  E  E  E  E  I  D  E  T  C  E  H  E  Y  G
       -  R  T  M  Q  S  V  R  K  K  K  L  M  K  P  V  N  M  S  T  V
       -   G  R  C  R  V  *  G  R  R  N  *  *  N  L  *  T  *  V  R  Y
3121 - ACAGAGGATGATTATCAAGGTCTCCCTCTGGAATTTGGTGCCTCAGCTGAAACAGTTCGA -
3180
       - T  E  D  D  Y  Q  G  L  P  L  E  F  G  A  S  A  E  T  V  R
       -  Q  R  M  I  I  K  V  S  L  W  N  L  V  P  Q  L  K  Q  F  E
       -   R  G  *  L  S  R  S  P  S  G  I  W  C  L  S  *  N  S  S  S
3181 - GTTGAGGAAGAAGAAGAGGAAGACTGGCTGGATGATACTACTGAGCAATCAGAGATTGAG -
3240
       - V  E  E  E  E  E  E  D  W  L  D  D  T  T  E  Q  S  E  I  E
       -  L  R  K  K  K  R  K  T  G  W  M  I  L  L  S  N  Q  R  L  S
       -   *  G  R  R  R  G  R  L  A  G  *  Y  Y  *  A  I  R  D  *  A
3241 - CCAGAACCAGAACCTACACCTGAAGAACCAGTTAATCAGTTTACTGGTTATTTAAAACTT -
3300
       - P  E  P  E  P  T  P  E  E  P  V  N  Q  F  T  G  Y  L  K  L
       -  Q  N  Q  N  L  H  L  K  N  Q  L  I  S  L  L  V  I  *  N  L
       -   R  T  R  T  Y  T  *  R  T  S  *  S  V  Y  W  L  F  K  T  Y
3301 - ACTGACAATGTTGCCATTAAATGTGTTGACATCGTTAAGGAGGCACAAAGTGCTAATCCT -
3360
       - T  D  N  V  A  I  K  C  V  D  I  V  K  E  A  Q  S  A  N  P
       -  L  T  M  L  P  L  N  V  L  T  S  L  R  R  H  K  V  L  I  L
       -   *  Q  C  C  H  *  M  C  *  H  R  *  G  G  T  K  C  *  S  Y
3361 - ATGGTGATTGTAAATGCTGCTAACATACACCTGAAACATGGTGGTGGTGTAGCAGGTGCA -
3420
       - M  V  I  V  N  A  A  N  I  H  L  K  H  G  G  G  V  A  G  A
       -  W  *  L  *  M  L  L  T  Y  T  *  N  M  V  V  V  *  Q  V  H
       -   G  D  C  K  C  C  *  H  T  P  E  T  W  W  W  C  S  R  C  T
3421 - CTCAACAAGGCAACCAATGGTGCCATGCAAAAGGAGAGTGATGATTACATTAAGCTAAAT -
3480
       - L  N  K  A  T  N  G  A  M  Q  K  E  S  D  D  Y  I  K  L  N
       -  S  T  R  Q  P  M  V  P  C  K  R  R  V  M  I  T  L  S  *  M
       -   Q  Q  G  N  Q  W  C  H  A  K  G  E  *  *  L  H  *  A  K  W
3481 - GGCCCTCTTACAGTAGGAGGGTCTTGTTTGCTTTCTGGACATAATCTTGCTAAGAAGTGT -
3540
       - G  P  L  T  V  G  G  S  C  L  L  S  G  H  N  L  A  K  K  C
       -  A  L  L  Q  *  E  G  L  V  C  F  L  D  I  I  L  L  R  S  V
       -   P  S  Y  S  R  R  V  L  F  A  F  W  T  *  S  C  *  E  V  S
3541 - CTGCATGTTGTTGGACCTAACCTAAATGCAGGTGAGGACATCCAGCTTCTTAAGGCAGCA -
3600
       - L  H  V  V  G  P  N  L  N  A  G  E  D  I  Q  L  L  K  A  A
       -  C  M  L  L  D  L  T  *  M  Q  V  R  T  S  S  F  L  R  Q  H
       -   A  C  C  W  T  *  P  K  C  R  *  G  H  P  A  S  *  G  S  I
3601 - TATGAAAATTTCAATTCACAGGACATCTTACTTGCACCATTGTTGTCAGCAGGCATATTT -
3660
       - Y  E  N  F  N  S  Q  D  I  L  L  A  P  L  L  S  A  G  I  F
       -  M  K  I  S  I  H  R  T  S  Y  L  H  H  C  C  Q  Q  A  Y  L
       -   *  K  F  Q  F  T  G  H  L  T  C  T  I  V  V  S  R  H  I  W
3661 - GGTGCTAAACCACTTCAGTCTTTACAAGTGTGCGTGCAGACGGTTCGTACACAGGTTTAT -
3720
       - G  A  K  P  L  Q  S  L  Q  V  C  V  Q  T  V  R  T  Q  V  Y
       -  V  L  N  H  F  S  L  Y  K  C  A  C  R  R  F  V  H  R  F  I
       -   C  *  T  T  S  V  F  T  S  V  R  A  D  G  S  Y  T  G  L  Y
3721 - ATTGCAGTCAATGACAAAGCTCTTTATGAGCAGGTTGTCATGGATTATCTTGATAACCTG -
3780
       - I  A  V  N  D  K  A  L  Y  E  Q  V  V  M  D  Y  L  D  N  L
```

```
     - L  Q  S  M  T  K  L  F  M  S  R  L  S  W  I  I  L  I  T  *
     -  C  S  Q  *  Q  S  S  L  *  A  G  C  H  G  L  S  *  *  P  E
3781 - AAGCCTAGAGTGGAAGCACCTAAACAAGAGGAGCCACCAAACACAGAAGATTCCAAAACT -
3840
     - K  P  R  V  E  A  P  K  Q  E  E  P  P  N  T  E  D  S  K  T
     -  S  L  E  W  K  H  L  N  K  R  S  H  Q  T  Q  K  I  P  K  L
     -  A  *  S  G  S  T  *  T  R  G  A  T  K  H  R  R  F  Q  N  *
```

Figure 16-F

```
3841 - GAGGAGAAATCTGTCGTACAGAAGCCTGTCGATGTGAAGCCAAAAATTAAGGCCTGCATT -
3900
       - E  E  K  S  V  V  Q  K  P  V  D  V  K  P  K  I  K  A  C  I
       -  R  R  N  L  S  Y  R  S  L  S  M  *  S  Q  K  L  R  P  A  L
       -   G  E  I  C  R  T  E  A  C  R  C  E  A  K  N  *  G  L  H  *
3901 - GATGAGGTTACCACAACACTGGAAGAAACTAAGTTTCTTACCAATAAGTTACTCTTGTTT -
3960
       - D  E  V  T  T  T  L  E  E  T  K  F  L  T  N  K  L  L  L  F
       -  M  R  L  P  Q  H  W  K  K  L  S  F  L  P  I  S  Y  S  C  L
       -   *  G  Y  H  N  T  G  R  N  *  V  S  Y  Q  *  V  T  L  V  C
3961 - GCTGATATCAATGGTAAGCTTTACCATGATTCTCAGAACATGCTTAGAGGTGAAGATATG -
4020
       - A  D  I  N  G  K  L  Y  H  D  S  Q  N  M  L  R  G  E  D  M
       -  L  I  S  M  V  S  F  T  M  I  L  R  T  C  L  E  V  K  I  C
       -   *  Y  Q  W  *  A  L  P  *  F  S  E  H  A  *  R  *  R  Y  V
4021 - TCTTTCCTTGAGAAGGATGCACCTTACATGGTAGGTGATGTTATCACTAGTGGTGATATC -
4080
       - S  F  L  E  K  D  A  P  Y  M  V  G  D  V  I  T  S  G  D  I
       -  L  S  L  R  R  M  H  L  T  W  *  V  M  L  S  L  V  V  I  S
       -   F  P  *  E  G  C  T  L  H  G  R  *  C  Y  H  *  W  *  Y  H
4081 - ACTTGTGTTGTAATACCCTCCAAAAAGGCTGGTGGCACTACTGAGATGCTCTCAAGAGCT -
4140
       - T  C  V  V  I  P  S  K  K  A  G  G  T  T  E  M  L  S  R  A
       -  L  V  L  *  Y  P  P  K  R  L  V  A  L  L  R  C  S  Q  E  L
       -   L  C  C  N  T  L  Q  K  G  W  W  H  Y  *  D  A  L  K  S  F
4141 - TTGAAGAAAGTGCCAGTTGATGAGTATATAACCACGTACCCTGGACAAGGATGTGCTGGT -
4200
       - L  K  K  V  P  V  D  E  Y  I  T  T  Y  P  G  Q  G  C  A  G
       -  *  R  K  C  Q  L  M  S  I  *  P  R  T  L  D  K  D  V  L  V
       -   E  E  S  A  S  *  *  V  Y  N  H  V  P  W  T  R  M  C  W  L
4201 - TATACACTTGAGGAAGCTAAGACTGCTCTTAAGAAATGCAAATCTGCATTTTATGTACTA -
4260
       - Y  T  L  E  E  A  K  T  A  L  K  K  C  K  S  A  F  Y  V  L
       -  I  H  L  R  K  L  R  L  L  L  R  N  A  N  L  H  F  M  Y  Y
       -   Y  T  *  G  S  *  D  C  S  *  E  M  Q  I  C  I  L  C  T  T
4261 - CCTTCAGAAGCACCTAATGCTAAGGAAGAGATTCTAGGAACTGTATCCTGGAATTTGAGA -
4320
       - P  S  E  A  P  N  A  K  E  E  I  L  G  T  V  S  W  N  L  R
       -  L  Q  K  H  L  M  L  R  K  R  F  *  E  L  Y  P  G  I  *  E
       -   F  R  S  T  *  C  *  G  R  D  S  R  N  C  I  L  E  F  E  R
4321 - GAAATGCTTGCTCATGCTGAAGAGACAAGAAAATTAATGCCTATATGCATGGATGTTAGA -
4380
       - E  M  L  A  H  A  E  E  T  R  K  L  M  P  I  C  M  D  V  R
       -  K  C  L  L  M  L  K  R  Q  E  N  *  C  L  Y  A  W  M  L  E
       -   N  A  C  S  C  *  R  D  K  K  I  N  A  Y  M  H  G  C  *  S
4381 - GCCATAATGGCAACCATCCAACGTAAGTATAAAGGAATTAAAATTCAAGAGGGCATCGTT -
4440
       - A  I  M  A  T  I  Q  R  K  Y  K  G  I  K  I  Q  E  G  I  V
       -  P  *  W  Q  P  S  N  V  S  I  K  E  L  K  F  K  R  A  S  L
       -   H  N  G  N  H  P  T  *  V  *  R  N  *  N  S  R  G  H  R  *
4441 - GACTATGGTGTCCGATTCTTCTTTTATACTAGTAAAGAGCCTGTAGCTTCTATTATTACG -
4500
       - D  Y  G  V  R  F  F  F  Y  T  S  K  E  P  V  A  S  I  I  T
       -  T  M  V  S  D  S  S  F  I  L  V  K  S  L  *  L  L  L  L  R
       -   L  W  C  P  I  L  L  L  Y  *  *  R  A  C  S  F  Y  Y  Y  E
4501 - AAGCTGAACTCTCTAAATGAGCCGCTTGTCACAATGCCAATTGGTTATGTGACACATGGT -
4560
```

-K L N S L N E P L V T M P I G Y V T H G
- S * T L * M S R L S Q C Q L V M * H M V
- A E L S K * A A C H N A N W L C D T W F

Figure 16-G

```
4561 - TTTAATCTTGAAGAGGCTGCGCGCTGTATGCGTTCTCTTAAAGCTCCTGCCGTAGTGTCA -
4620
      - F  N  L  E  E  A  A  R  C  M  R  S  L  K  A  P  A  V  V  S
      -  L  I  L  K  R  L  R  A  V  C  V  L  L  K  L  L  P  *  C  Q
      -   *  S  *  R  G  C  A  L  Y  A  F  S  *  S  S  C  R  S  V  S
4621 - GTATCATCACCAGATGCTGTTACTACATATAATGGATACCTCACTTCGTCATCAAAGACA -
4680
      - V  S  S  P  D  A  V  T  T  Y  N  G  Y  L  T  S  S  S  K  T
      -  Y  H  H  Q  M  L  L  L  H  I  M  D  T  S  L  R  H  Q  R  H
      -   I  I  T  R  C  C  Y  Y  I  *  W  I  P  H  F  V  I  K  D  I
4681 - TCTGAGGAGCACTTTGTAGAAACAGTTTCTTTGGCTGGCTCTTACAGAGATTGGTCCTAT -
4740
      - S  E  E  H  F  V  E  T  V  S  L  A  G  S  Y  R  D  W  S  Y
      -  L  R  S  T  L  *  K  Q  F  L  W  L  A  L  T  E  I  G  P  I
      -   *  G  A  L  C  R  N  S  F  F  G  W  L  L  Q  R  L  V  L  F
4741 - TCAGGACAGCGTACAGAGTTAGGTGTTGAATTTCTTAAGCGTGGTGACAAAATTGTGTAC -
4800
      - S  G  Q  R  T  E  L  G  V  E  F  L  K  R  G  D  K  I  V  Y
      -  Q  D  S  V  Q  S  *  V  L  N  F  L  S  V  V  T  K  L  C  T
      -   R  T  A  Y  R  V  R  C  *  I  S  *  A  W  *  Q  N  C  V  P
4801 - CACACTCTGGAGAGCCCCGTCGAGTTTCATCTTGACGGTGAGGTTCTTTCACTTGACAAA -
4860
      - H  T  L  E  S  P  V  E  F  H  L  D  G  E  V  L  S  L  D  K
      -  T  L  W  R  A  P  S  S  F  I  L  T  V  R  F  F  H  L  T  N
      -   H  S  G  E  P  R  R  V  S  S  *  R  *  G  S  F  T  *  Q  T
4861 - CTAAAGAGTCTCTTATCCCTGCGGGAGGTTAAGACTATAAAAGTGTTCACAACTGTGGAC -
4920
      - L  K  S  L  L  S  L  R  E  V  K  T  I  K  V  F  T  T  V  D
      -  *  R  V  S  Y  P  C  G  R  L  R  L  *  K  C  S  Q  L  W  T
      -   K  E  S  L  I  P  A  G  G  *  D  Y  K  S  V  H  N  C  G  Q
4921 - AACACTAATCTCCACACACAGCTTGTGGATATGTCTATGACATATGGACAGCAGTTTGGT -
4980
      - N  T  N  L  H  T  Q  L  V  D  M  S  M  T  Y  G  Q  Q  F  G
      -  T  L  I  S  T  H  S  L  W  I  C  L  *  H  M  D  S  S  L  V
      -   H  *  S  P  H  T  A  C  G  Y  V  Y  D  I  W  T  A  V  W  S
4981 - CCAACATACTTGGATGGTGCTGATGTTACAAAAATTAAACCTCATGTAAATCATGAGGGT -
5040
      - P  T  Y  L  D  G  A  D  V  T  K  I  K  P  H  V  N  H  E  G
      -  Q  H  T  W  M  V  L  M  L  Q  K  L  N  L  M  *  I  M  R  V
      -   N  I  L  G  W  C  *  C  Y  K  N  *  T  S  C  K  S  *  G  *
5041 - AAGACTTTCTTTGTACTACCTAGTGATGACACACTACGTAGTGAAGCTTTCGAGTACTAC -
5100
      - K  T  F  F  V  L  P  S  D  D  T  L  R  S  E  A  F  E  Y  Y
      -  R  L  S  L  Y  Y  L  V  M  T  H  Y  V  V  K  L  S  S  T  T
      -   D  F  L  C  T  T  *  *  *  H  T  T  *  *  S  F  R  V  L  P
5101 - CATACTCTTGATGAGAGTTTTCTTGGTAGGTACATGTCTGCTTTAAACCACACAAAGAAA -
5160
      - H  T  L  D  E  S  F  L  G  R  Y  M  S  A  L  N  H  T  K  K
      -  I  L  L  M  R  V  F  L  V  G  T  C  L  L  *  T  T  Q  R  N
      -   Y  S  *  *  E  F  S  W  *  V  H  V  C  F  K  P  H  K  E  M
5161 - TGGAAATTTCCTCAAGTTGGTGGTTTAACTTCAATTAAATGGGCTGATAACAATTGTTAT -
5220
      - W  K  F  P  Q  V  G  G  L  T  S  I  K  W  A  D  N  N  C  Y
      -  G  N  F  L  K  L  V  V  *  L  Q  L  N  G  L  I  T  I  V  I
      -   E  I  S  S  S  W  W  F  N  F  N  *  M  G  *  *  Q  L  L  F
5221 - TTGTCTAGTGTTTTATTAGCACTTCAACAGCTTGAAGTCAAATTCAATGCACCAGCACTT -
5280
      - L  S  S  V  L  L  A  L  Q  Q  L  E  V  K  F  N  A  P  A  L
```

```
     - C  L  V  F  Y  *  H  F  N  S  L  K  S  N  S  M  H  Q  H  F
     -  V  *  C  F  I  S  T  S  T  A  *  S  Q  I  Q  C  T  S  T  S
5281 - CAAGAGGCTTATTATAGAGCCCGTGCTGGTGATGCTGCTAACTTTTGTGCACTCATACTC -
5340
     - Q  E  A  Y  Y  R  A  R  A  G  D  A  A  N  F  C  A  L  I  L
     -  K  R  L  I  I  E  P  V  L  V  M  L  L  T  F  V  H  S  Y  S
     -  R  G  L  L  *  S  P  C  W  *  C  C  *  L  L  C  T  H  T  R
```

Figure 16-H

```
 5341 - GCTTACAGTAATAAAACTGTTGGCGAGCTTGGTGATGTCAGAGAAACTATGACCCATCTT -
5400
      -A  Y  S  N  K  T  V  G  E  L  G  D  V  R  E  T  M  T  H  L
      - L  T  V  I  K  L  L  A  S  L  V  M  S  E  K  L  *  P  I  F
      -  L  Q  *  *  N  C  W  R  A  W  *  C  Q  R  N  Y  D  P  S  S
 5401 - CTACAGCATGCTAATTTGGAATCTGCAAAGCGAGTTCTTAATGTGGTGTGTAAACATTGT -
5460
      -L  Q  H  A  N  L  E  S  A  K  R  V  L  N  V  V  C  K  H  C
      - Y  S  M  L  I  W  N  L  Q  S  E  F  L  M  W  C  V  N  I  V
      -  T  A  C  *  F  G  I  C  K  A  S  S  *  C  G  V  *  T  L  W
 5461 - GGTCAGAAAACTACTACCTTAACGGGTGTAGAAGCTGTGATGTATATGGGTACTCTATCT -
5520
      -G  Q  K  T  T  T  L  T  G  V  E  A  V  M  Y  M  G  T  L  S
      - V  R  K  L  L  P  *  R  V  *  K  L  *  C  I  W  V  L  Y  L
      -  S  E  N  Y  Y  L  N  G  C  R  S  C  D  V  Y  G  Y  S  I  L
 5521 - TATGATAATCTTAAGACAGGTGTTTCCATTCCATGTGTGTGTGGTCGTGATGCTACACAA -
5580
      -Y  D  N  L  K  T  G  V  S  I  P  C  V  C  G  R  D  A  T  Q
      - M  I  I  L  R  Q  V  F  P  F  H  V  C  V  V  V  M  L  H  N
      -  *  *  S  *  D  R  C  F  H  S  M  C  V  W  S  *  C  Y  T  I
 5581 - TATCTAGTACAACAAGAGTCTTCTTTTGTTATGATGTCTGCACCACCTGCTGAGTATAAA -
5640
      -Y  L  V  Q  Q  E  S  S  F  V  M  M  S  A  P  P  A  E  Y  K
      - I  *  Y  N  K  S  L  L  L  L  *  C  L  H  H  L  L  S  I  N
      -  S  S  T  T  R  V  F  F  C  Y  D  V  C  T  T  C  *  V  *  I
 5641 - TTACAGCAAGGTACATTCTTATGTGCGAATGAGTACACTGGTAACTATCAGTGTGGTCAT -
5700
      -L  Q  Q  G  T  F  L  C  A  N  E  Y  T  G  N  Y  Q  C  G  H
      - Y  S  K  V  H  S  Y  V  R  M  S  T  L  V  T  I  S  V  V  I
      -  T  A  R  Y  I  L  M  C  E  *  V  H  W  *  L  S  V  W  S  L
 5701 - TACACTCATATAACTGCTAAGGAGACCCTCTATCGTATTGACGGAGCTCACCTTACAAAG -
5760
      -Y  T  H  I  T  A  K  E  T  L  Y  R  I  D  G  A  H  L  T  K
      - T  L  I  *  L  L  R  R  P  S  I  V  L  T  E  L  T  L  Q  R
      -  H  S  Y  N  C  *  G  D  P  L  S  Y  *  R  S  S  P  Y  K  D
 5761 - ATGTCAGAGTACAAAGGACCAGTGACTGATGTTTTCTACAAGGAAACATCTTACACTACA -
5820
      -M  S  E  Y  K  G  P  V  T  D  V  F  Y  K  E  T  S  Y  T  T
      - C  Q  S  T  K  D  Q  *  L  M  F  S  T  R  K  H  L  T  L  Q
      -  V  R  V  Q  R  T  S  D  *  C  F  L  Q  G  N  I  L  H  Y  N
 5821 - ACCATCAAGCCTGTGTCGTATAAACTCGATGGAGTTACTTACACAGAGATTGAACCAAAA -
5880
      -T  I  K  P  V  S  Y  K  L  D  G  V  T  Y  T  E  I  E  P  K
      - P  S  S  L  C  R  I  N  S  M  E  L  L  T  Q  R  L  N  Q  N
      -  H  Q  A  C  V  V  *  T  R  W  S  Y  L  H  R  D  *  T  K  I
 5881 - TTGGATGGGTATTATAAAAAGGATAATGCTTACTATACAGAGCAGCCTATAGACCTTGTA -
5940
      -L  D  G  Y  Y  K  K  D  N  A  Y  Y  T  E  Q  P  I  D  L  V
      - W  M  G  I  I  K  R  I  M  L  T  I  Q  S  S  L  *  T  L  Y
      -  G  W  V  L  *  K  G  *  C  L  L  Y  R  A  A  Y  R  P  C  T
 5941 - CCAACTCAACCATTACCAAATGCGAGTTTTGATAATTTCAAACTCACATGTTCTAACACA -
6000
      -P  T  Q  P  L  P  N  A  S  F  D  N  F  K  L  T  C  S  N  T
      - Q  L  N  H  Y  Q  M  R  V  L  I  I  S  N  S  H  V  L  T  Q
      -  N  S  T  I  T  K  C  E  F  *  *  F  Q  T  H  M  F  *  H  K
 6001 - AAATTTGCTGATGATTTAAATCAAATGACAGGCTTCACAAAGCCAGCTTCACGAGAGCTA -
6060
      -K  F  A  D  D  L  N  Q  M  T  G  F  T  K  P  A  S  R  E  L
```

```
     - N  L  L  M  I  *  I  K  *  Q  A  S  Q  S  Q  L  H  E  S  Y
     -  I  C  *  *  F  K  S  N  D  R  L  H  K  A  S  F  T  R  A  I
6061 - TCTGTCACATTCTTCCCAGACTTGAATGGCGATGTAGTGGCTATTGACTATAGACACTAT -
6120
     - S  V  T  F  F  P  D  L  N  G  D  V  V  A  I  D  Y  R  H  Y
     -  L  S  H  S  S  Q  T  *  M  A  M  *  W  L  L  T  I  D  T  I
     -   C  H  I  L  P  R  L  E  W  R  C  S  G  Y  *  L  *  T  L  F
```

Figure 16-I

```
6121 - TCAGCGAGTTTCAAGAAAGGTGCTAAATTACTGCATAAGCCAATTGTTTGGCACATTAAC -
6180
       - S  A  S  F  K  K  G  A  K  L  L  H  K  P  I  V  W  H  I  N
       -  Q  R  V  S  R  K  V  L  N  Y  C  I  S  Q  L  F  G  T  L  T
       -   S  E  F  Q  E  R  C  *  I  T  A  *  A  N  C  L  A  H  *  P
6181 - CAGGCTACAACCAAGACAACGTTCAAACCAAACACTTGGTGTTTACGTTGTCTTTGGAGT -
6240
       - Q  A  T  T  K  T  T  F  K  P  N  T  W  C  L  R  C  L  W  S
       -  R  L  Q  P  R  Q  R  S  N  Q  T  L  G  V  Y  V  V  F  G  V
       -   G  Y  N  Q  D  N  V  Q  T  K  H  L  V  F  T  L  S  L  E  Y
6241 - ACAAAGCCAGTAGATACTTCAAATTCATTTGAAGTTCTGGCAGTAGAAGACACACAAGGA -
6300
       - T  K  P  V  D  T  S  N  S  F  E  V  L  A  V  E  D  T  Q  G
       -  Q  S  Q  *  I  L  Q  I  H  L  K  F  W  Q  *  K  T  H  K  E
       -   K  A  S  R  Y  F  K  F  I  *  S  S  G  S  R  R  H  T  R  N
6301 - ATGGACAATCTTGCTTGTGAAAGTCAACAACCCACCTCTGAAGAAGTAGTGGAAAATCCT -
6360
       - M  D  N  L  A  C  E  S  Q  Q  P  T  S  E  E  V  V  E  N  P
       -  W  T  I  L  L  V  K  V  N  N  P  P  L  K  K  *  W  K  I  L
       -   G  Q  S  C  L  *  K  S  T  T  H  L  *  R  S  S  G  K  S  Y
6361 - ACCATACAGAAGGAAGTCATAGAGTGTGACGTGAAAACTACCGAAGTTGTAGGCAATGTC -
6420
       - T  I  Q  K  E  V  I  E  C  D  V  K  T  T  E  V  V  G  N  V
       -  P  Y  R  R  K  S  *  S  V  T  *  K  L  P  K  L  *  A  M  S
       -   H  T  E  G  S  H  R  V  *  R  E  N  Y  R  S  C  R  Q  C  H
6421 - ATACTTAAACCATCAGATGAAGGTGTTAAAGTAACACAAGAGTTAGGTCATGAGGATCTT -
6480
       - I  L  K  P  S  D  E  G  V  K  V  T  Q  E  L  G  H  E  D  L
       -  Y  L  N  H  Q  M  K  V  L  K  *  H  K  S  *  V  M  R  I  L
       -   T  *  T  I  R  *  R  C  *  S  N  T  R  V  R  S  *  G  S  Y
6481 - ATGGCTGCTTATGTGGAAAACACAAGCATTACCATTAAGAAACCTAATGAGCTTTCACTA -
6540
       - M  A  A  Y  V  E  N  T  S  I  T  I  K  K  P  N  E  L  S  L
       -  W  L  L  M  W  K  T  Q  A  L  P  L  R  N  L  M  S  F  H  *
       -   G  C  L  C  G  K  H  K  H  Y  H  *  E  T  *  *  A  F  T  S
6541 - GCCTTAGGTTTAAAAACAATTGCCACTCATGGTATTGCTGCAATTAATAGTGTTCCTTGG -
6600
       - A  L  G  L  K  T  I  A  T  H  G  I  A  A  I  N  S  V  P  W
       -  P  *  V  *  K  Q  L  P  L  M  V  L  L  Q  L  I  V  F  L  G
       -   L  R  F  K  N  N  C  H  S  W  Y  C  C  N  *  *  C  S  L  E
6601 - AGTAAAATTTTGGCTTATGTCAAACCATTCTTAGGACAAGCAGCAATTACAACATCAAAT -
6660
       - S  K  I  L  A  Y  V  K  P  F  L  G  Q  A  A  I  T  T  S  N
       -  V  K  F  W  L  M  S  N  H  S  *  D  K  Q  Q  L  Q  H  Q  I
       -   *  N  F  G  L  C  Q  T  I  L  R  T  S  S  N  Y  N  I  K  L
6661 - TGCGCTAAGAGATTAGCACAACGTGTGTTTAACAATTATATGCCTTATGTGTTTACATTA -
6720
       - C  A  K  R  L  A  Q  R  V  F  N  N  Y  M  P  Y  V  F  T  L
       -  A  L  R  D  *  H  N  V  C  L  T  I  I  C  L  M  C  L  H  Y
       -   R  *  E  I  S  T  T  C  V  *  Q  L  Y  A  L  C  V  Y  I  I
6721 - TTGTTCCAATTGTGTACTTTTACTAAAAGTACCAATTCTAGAATTAGAGCTTCACTACCT -
6780
       - L  F  Q  L  C  T  F  T  K  S  T  N  S  R  I  R  A  S  L  P
       -  C  S  N  C  V  L  L  L  K  V  P  I  L  E  L  E  L  H  Y  L
       -   V  P  I  V  Y  F  Y  *  K  Y  Q  F  *  N  *  S  F  T  T  Y
6781 - ACAACTATTGCTAAAAATAGTGTTAAGAGTGTTGCTAAATTATGTTTGGATGCCGGCATT -
6840
       - T  T  I  A  K  N  S  V  K  S  V  A  K  L  C  L  D  A  G  I
```

```
     - Q  L  L  L  K  I  V  L  R  V  L  L  N  Y  V  W  M  P  A  L
     -  N  Y  C  *  K  *  C  *  E  C  C  *  I  M  F  G  C  R  H  *
6841 - AATTATGTGAAGTCACCCAAATTTTCTAAATTGTTCACAATCGCTATGTGGCTATTGTTG - 6900
     - N  Y  V  K  S  P  K  F  S  K  L  F  T  I  A  M  W  L  L  L
     -  I  M  *  S  H  P  N  F  L  N  C  S  Q  S  L  C  G  Y  C  C
     -  L  C  E  V  T  Q  I  F  *  I  V  H  N  R  Y  V  A  I  V  V
```

Figure 16-J

```
6901 - TTAAGTATTTGCTTAGGTTCTCTAATCTGTGTAACTGCTGCTTTTGGTGTACTCTTATCT -
6960
       -L  S  I  C  L  G  S  L  I  C  V  T  A  A  F  G  V  L  L  S
       - *  V  F  A  *  V  L  *  S  V  *  L  L  L  L  V  Y  S  Y  L
       -  K  Y  L  L  R  F  S  N  L  C  N  C  C  F  W  C  T  L  I  *
6961 - AATTTTGGTGCTCCTTCTTATTGTAATGGCGTTAGAGAATTGTATCTTAATTCGTCTAAC -
7020
       -N  F  G  A  P  S  Y  C  N  G  V  R  E  L  Y  L  N  S  S  N
       - I  L  V  L  L  L  I  V  M  A  L  E  N  C  I  L  I  R  L  T
       -  F  W  C  S  F  L  L  *  W  R  *  R  I  V  S  *  F  V  *  R
7021 - GTTACTACTATGGATTTCTGTGAAGGTTCTTTTCCTTGCAGCATTTGTTTAAGTGGATTA -
7080
       -V  T  T  M  D  F  C  E  G  S  F  P  C  S  I  C  L  S  G  L
       - L  L  L  W  I  S  V  K  V  L  F  L  A  A  F  V  *  V  D  *
       -  Y  Y  Y  G  F  L  *  R  F  F  S  L  Q  H  L  F  K  W  I  R
7081 - GACTCCCTTGATTCTTATCCAGCTCTTGAAACCATTCAGGTGACGATTTCATCGTACAAG -
7140
       -D  S  L  D  S  Y  P  A  L  E  T  I  Q  V  T  I  S  S  Y  K
       - T  P  L  I  L  I  Q  L  L  K  P  F  R  *  R  F  H  R  T  S
       -  L  P  *  F  L  S  S  S  *  N  H  S  G  D  D  F  I  V  Q  A
7141 - CTAGACTTGACAATTTTAGGTCTGGCCGCTGAGTGGGTTTTGGCATATATGTTGTTCACA -
7200
       -L  D  L  T  I  L  G  L  A  A  E  W  V  L  A  Y  M  L  F  T
       - *  T  *  Q  F  *  V  W  P  L  S  G  F  W  H  I  C  C  S  Q
       -  R  L  D  N  F  R  S  G  R  *  V  G  F  G  I  Y  V  V  H  K
7201 - AAATTCTTTTATTTATTAGGTCTTTCAGCTATAATGCAGGTGTTCTTTGGCTATTTTGCT -
7260
       -K  F  F  Y  L  L  G  L  S  A  I  M  Q  V  F  F  G  Y  F  A
       - N  S  F  I  Y  *  V  F  Q  L  *  C  R  C  S  L  A  I  L  L
       -  I  L  L  F  I  R  S  F  S  Y  N  A  G  V  L  W  L  F  C  *
7261 - AGTCATTTCATCAGCAATTCTTGGCTCATGTGGTTTATCATTAGTATTGTACAAATGGCA -
7320
       -S  H  F  I  S  N  S  W  L  M  W  F  I  I  S  I  V  Q  M  A
       - V  I  S  S  A  I  L  G  S  C  G  L  S  L  V  L  Y  K  W  H
       -  S  F  H  Q  Q  F  L  A  H  V  V  Y  H  *  Y  C  T  N  G  T
7321 - CCCGTTTCTGCAATGGTTAGGATGTACATCTTCTTTGCTTCTTTCTACTACATATGGAAG -
7380
       -P  V  S  A  M  V  R  M  Y  I  F  F  A  S  F  Y  Y  I  W  K
       - P  F  L  Q  W  L  G  C  T  S  S  L  L  L  S  T  T  Y  G  R
       -  R  F  C  N  G  *  D  V  H  L  L  C  F  F  L  L  H  M  E  E
7381 - AGCTATGTTCATATCATGGATGGTTGCACCTCTTCGACTTGCATGATGTGCTATAAGCGC -
7440
       -S  Y  V  H  I  M  D  G  C  T  S  S  T  C  M  M  C  Y  K  R
       - A  M  F  I  S  W  M  V  A  P  L  R  L  A  *  C  A  I  S  A
       -  L  C  S  Y  H  G  W  L  H  L  F  D  L  H  D  V  L  *  A  Q
7441 - AATCGTGCCACACGCGTTGAGTGTACAACTATTGTTAATGGCATGAAGAGATCTTTCTAT -
7500
       -N  R  A  T  R  V  E  C  T  T  I  V  N  G  M  K  R  S  F  Y
       - I  V  P  H  A  L  S  V  Q  L  L  L  M  A  *  R  D  L  S  M
       -  S  C  H  T  R  *  V  Y  N  Y  C  *  W  H  E  E  I  F  L  C
7501 - GTCTATGCAAATGGAGGCCGTGGCTTCTGCAAGACTCACAATTGGAATTGTCTCAATTGT -
7560
       -V  Y  A  N  G  G  R  G  F  C  K  T  H  N  W  N  C  L  N  C
       - S  M  Q  M  E  A  V  A  S  A  R  L  T  I  G  I  V  S  I  V
       -  L  C  K  W  R  P  W  L  L  Q  D  S  Q  L  E  L  S  Q  L  *
7561 - GACACATTTTGCACTGGTAGTACATTCATTAGTGATGAAGTTGCTCGTGATTTGTCACTC -
7620
       -D  T  F  C  T  G  S  T  F  I  S  D  E  V  A  R  D  L  S  L
```

- T H F A L V V H S L V M K L L V I C H S
- H I L H W * Y I H * * * S C S * F V T P
7621 - CAGTTTAAAAGACCAATCAACCCTACTGACCAGTCATCGTATATTGTTGATAGTGTTGCT - 7680
- Q F K R P I N P T D Q S S Y I V D S V A
- S L K D Q S T L L T S H R I L L I V L L
- V * K T N Q P Y * P V I V Y C * * C C C

Figure 16-K

```
7681 - GTGAAAAATGGCGCGCTTCACCTCTACTTTGACAAGGCTGGTCAAAAGACCTATGAGAGA -
7740
     - V  K  N  G  A  L  H  L  Y  F  D  K  A  G  Q  K  T  Y  E  R
     -  *  K  M  A  R  F  T  S  T  L  T  R  L  V  K  R  P  M  R  D
     -   E  K  W  R  A  S  P  L  L  *  Q  G  W  S  K  D  L  *  E  T
7741 - CATCCGCTCTCCCATTTTGTCAATTTAGACAATTTGAGAGCTAACAACACTAAAGGTTCA -
7800
     - H  P  L  S  H  F  V  N  L  D  N  L  R  A  N  N  T  K  G  S
     -  I  R  S  P  I  L  S  I  *  T  I  *  E  L  T  T  L  K  V  H
     -   S  A  L  P  F  C  Q  F  R  Q  F  E  S  *  Q  H  *  R  F  T
7801 - CTGCCTATTAATGTCATAGTTTTTGATGGCAAGTCCAAATGCGACGAGTCTGCTTCTAAG -
7860
     - L  P  I  N  V  I  V  F  D  G  K  S  K  C  D  E  S  A  S  K
     -  C  L  L  M  S  *  F  L  M  A  S  P  N  A  T  S  L  L  L  S
     -   A  Y  *  C  H  S  F  *  W  Q  V  Q  M  R  R  V  C  F  *  V
7861 - TCTGCTTCTGTGTACTACAGTCAGCTGATGTGCCAACCTATTCTGTTGCTTGACCAAGCT -
7920
     - S  A  S  V  Y  Y  S  Q  L  M  C  Q  P  I  L  L  L  D  Q  A
     -  L  L  L  C  T  T  V  S  *  C  A  N  L  F  C  C  L  T  K  L
     -   C  F  C  V  L  Q  S  A  D  V  P  T  Y  S  V  A  *  P  S  S
7921 - CTTGTATCAAACGTTGGAGATAGTACTGAAGTTTCCGTTAAGATGTTTGATGCTTATGTC -
7980
     - L  V  S  N  V  G  D  S  T  E  V  S  V  K  M  F  D  A  Y  V
     -  L  Y  Q  T  L  E  I  V  L  K  F  P  L  R  C  L  M  L  M  S
     -   C  I  K  R  W  R  *  Y  *  S  F  R  *  D  V  *  C  L  C  R
7981 - GACACCTTTTCAGCAACTTTTAGTGTTCCTATGGAAAAACTTAAGGCACTTGTTGCTACA -
8040
     - D  T  F  S  A  T  F  S  V  P  M  E  K  L  K  A  L  V  A  T
     -  T  P  F  Q  Q  L  L  V  F  L  W  K  N  L  R  H  L  L  L  Q
     -   H  L  F  S  N  F  *  C  S  Y  G  K  T  *  G  T  C  C  Y  S
8041 - GCTCACAGCGAGTTAGCAAAGGGTGTAGCTTTAGATGGTGTCCTTTCTACATTCGTGTCA -
8100
     - A  H  S  E  L  A  K  G  V  A  L  D  G  V  L  S  T  F  V  S
     -  L  T  A  S  *  Q  R  V  *  L  *  M  V  S  F  L  H  S  C  Q
     -   S  Q  R  V  S  K  G  C  S  F  R  W  C  P  F  Y  I  R  V  S
8101 - GCTGCCCGACAAGGTGTTGTTGATACCGATGTTGACACAAAGGATGTTATTGAATGTCTC -
8160
     - A  A  R  Q  G  V  V  D  T  D  V  D  T  K  D  V  I  E  C  L
     -  L  P  D  K  V  L  L  I  P  M  L  T  Q  R  M  L  L  N  V  S
     -   C  P  T  R  C  C  *  Y  R  C  *  H  K  G  C  Y  *  M  S  Q
8161 - AAACTTTCACATCACTCTGACTTAGAAGTGACAGGTGACAGTTGTAACAATTTCATGCTC -
8220
     - K  L  S  H  H  S  D  L  E  V  T  G  D  S  C  N  N  F  M  L
     -  N  F  H  I  T  L  T  *  K  *  Q  V  T  V  V  T  I  S  C  S
     -   T  F  T  S  L  *  L  R  S  D  R  *  Q  L  *  Q  F  H  A  H
8221 - ACCTATAATAAGGTTGAAAACATGACGCCCAGAGATCTTGGCGCATGTATTGACTGTAAT -
8280
     - T  Y  N  K  V  E  N  M  T  P  R  D  L  G  A  C  I  D  C  N
     -  P  I  I  R  L  K  T  *  R  P  E  I  L  A  H  V  L  T  V  M
     -   L  *  *  G  *  K  H  D  A  Q  R  S  W  R  M  Y  *  L  *  C
8281 - GCAAGGCATATCAATGCCCAAGTAGCAAAAAGTCACAATGTTTCACTCATCTGGAATGTA -
8340
     - A  R  H  I  N  A  Q  V  A  K  S  H  N  V  S  L  I  W  N  V
     -  Q  G  I  S  M  P  K  *  Q  K  V  T  M  F  H  S  S  G  M  *
     -   K  A  Y  Q  C  P  S  S  K  K  S  Q  C  F  T  H  L  E  C  K
8341 - AAAGACTACATGTCTTTATCTGAACAGCTGCGTAAACAAATTCGTACTGCTGCCAAGAAG -
8400
     - K  D  Y  M  S  L  S  E  Q  L  R  K  Q  I  R  T  A  A  K  K
```

```
     - K  T  T  C  L  Y  L  N  S  C  V  N  K  F  V  L  L  P  R  R
     -  R  L  H  V  F  I  *  T  A  A  *  T  N  S  Y  C  C  Q  E  E
8401 - AACAACATACCTTTTACACTAACTTGTGCTACAACTAGACAGGTTGTCAATGTCATAACT -
8460
     - N  N  I  P  F  T  L  T  C  A  T  T  R  Q  V  V  N  V  I  T
     -  T  T  Y  L  L  H  *  L  V  L  Q  L  D  R  L  S  M  S  *  L
     -  Q  H  T  F  Y  T  N  L  C  Y  N  *  T  G  C  Q  C  H  N  Y
```

Figure 16-L

```
8461 - ACTAAAATCTCACTCAAGGGTGGTAAGATTGTTAGTACTTGTTTTAAACTTATGCTTAAG -
8520
     - T  K  I  S  L  K  G  G  K  I  V  S  T  C  F  K  L  M  L  K
     -  L  K  S  H  S  R  V  V  R  L  L  V  L  V  L  N  L  C  L  R
     -   *  N  L  T  Q  G  W  *  D  C  *  Y  L  F  *  T  Y  A  *  G
8521 - GCCACATTATTGTGCGTTCTTGCTGCATTGGTTTGTTATATCGTTATGCCAGTACATACA -
8580
     - A  T  L  L  C  V  L  A  A  L  V  C  Y  I  V  M  P  V  H  T
     -  P  H  Y  C  A  F  L  L  H  W  F  V  I  S  L  C  Q  Y  I  H
     -   H  I  I  V  R  S  C  C  I  G  L  L  Y  R  Y  A  S  T  Y  I
8581 - TTGTCAATCCATGATGGTTACACAAATGAAATCATTGGTTACAAAGCCATTCAGGATGGT -
8640
     - L  S  I  H  D  G  Y  T  N  E  I  I  G  Y  K  A  I  Q  D  G
     -  C  Q  S  M  M  V  T  Q  M  K  S  L  V  T  K  P  F  R  M  V
     -   V  N  P  *  W  L  H  K  *  N  H  W  L  Q  S  H  S  G  W  C
8641 - GTCACTCGTGACATCATTTCTACTGATGATTGTTTTGCAAATAAACATGCTGGTTTTGAC -
8700
     - V  T  R  D  I  I  S  T  D  D  C  F  A  N  K  H  A  G  F  D
     -  S  L  V  T  S  F  L  L  M  I  V  L  Q  I  N  M  L  V  L  T
     -   H  S  *  H  H  F  Y  *  *  L  F  C  K  *  T  C  W  F  *  R
8701 - GCATGGTTTAGCCAGCGTGGTGGTTCATACAAAAATGACAAAAGCTGCCCTGTAGTAGCT -
8760
     - A  W  F  S  Q  R  G  G  S  Y  K  N  D  K  S  C  P  V  V  A
     -  H  G  L  A  S  V  V  V  H  T  K  M  T  K  A  A  L  *  *  L
     -   M  V  *  P  A  W  W  F  I  Q  K  *  Q  K  L  P  C  S  S  C
8761 - GCTATCATTACAAGAGAGATTGGTTTCATAGTGCCTGGCTTACCGGGTACTGTGCTGAGA -
8820
     - A  I  I  T  R  E  I  G  F  I  V  P  G  L  P  G  T  V  L  R
     -  L  S  L  Q  E  R  L  V  S  *  C  L  A  Y  R  V  L  C  *  E
     -   Y  H  Y  K  R  D  W  F  H  S  A  W  L  T  G  Y  C  A  E  S
8821 - GCAATCAATGGTGACTTCTTGCATTTTCTACCTCGTGTTTTTAGTGCTGTTGGCAACATT -
8880
     - A  I  N  G  D  F  L  H  F  L  P  R  V  F  S  A  V  G  N  I
     -  Q  S  M  V  T  S  C  I  F  Y  L  V  F  L  V  L  L  A  T  F
     -   N  Q  W  *  L  L  A  F  S  T  S  C  F  *  C  C  W  Q  H  L
8881 - TGCTACACACCTTCCAAACTCATTGAGTATAGTGATTTTGCTACCTCTGCTTGCGTTCTT -
8940
     - C  Y  T  P  S  K  L  I  E  Y  S  D  F  A  T  S  A  C  V  L
     -  A  T  H  L  P  N  S  L  S  I  V  I  L  L  P  L  L  A  F  L
     -   L  H  T  F  Q  T  H  *  V  *  *  F  C  Y  L  C  L  R  S  C
8941 - GCTGCTGAGTGTACAATTTTTAAGGATGCTATGGGCAAACCTGTGCCATATTGTTATGAC -
9000
     - A  A  E  C  T  I  F  K  D  A  M  G  K  P  V  P  Y  C  Y  D
     -  L  L  S  V  Q  F  L  R  M  L  W  A  N  L  C  H  I  V  M  T
     -   C  *  V  Y  N  F  *  G  C  Y  G  Q  T  C  A  I  L  L  *  H
9001 - ACTAATTTGCTAGAGGGTTCTATTTCTTATAGTGAGCTTCGTCCAGACACTCGTTATGTG -
9060
     - T  N  L  L  E  G  S  I  S  Y  S  E  L  R  P  D  T  R  Y  V
     -  L  I  C  *  R  V  L  F  L  I  V  S  F  V  Q  T  L  V  M  C
     -   *  F  A  R  G  F  Y  F  L  *  *  A  S  S  R  H  S  L  C  A
9061 - CTTATGGATGGTTCCATCATACAGTTTCCTAACACTTACCTGGAGGGTTCTGTTAGAGTA -
9120
     - L  M  D  G  S  I  I  Q  F  P  N  T  Y  L  E  G  S  V  R  V
     -  L  W  M  V  P  S  Y  S  F  L  T  L  T  W  R  V  L  L  E  *
     -   Y  G  W  F  H  H  T  V  S  *  H  L  P  G  G  F  C  *  S  S
9121 - GTAACAACTTTTGATGCTGAGTACTGTAGACATGGTACATGCGAAAGGTCAGAAGTAGGT -
9180
     - V  T  T  F  D  A  E  Y  C  R  H  G  T  C  E  R  S  E  V  G
     -  *  Q  L  L  M  L  S  T  V  D  M  V  H  A  K  G  Q  K  *  V
```

```
     -  N  N  F  *  C  *  V  L  *  T  W  Y  M  R  K  V  R  S  R  Y
9181 - ATTTGCCTATCTACCAGTGGTAGATGGGTTCTTAATAATGAGCATTACAGAGCTCTATCA -
9240
     - I  C  L  S  T  S  G  R  W  V  L  N  N  E  H  Y  R  A  L  S
     -  F  A  Y  L  P  V  V  D  G  F  L  I  M  S  I  T  E  L  Y  Q
     -  L  P  I  Y  Q  W  *  M  G  S  *  *  *  A  L  Q  S  S  I  R
```

Figure 16-M

```
 9241 - GGAGTTTTCTGTGGTGTTGATGCGATGAATCTCATAGCTAACATCTTTACTCCTCTTGTG -
9300
       -G  V  F  C  G  V  D  A  M  N  L  I  A  N  I  F  T  P  L  V
       - E  F  S  V  V  L  M  R  *  I  S  *  L  T  S  L  L  L  L  C
       -  S  F  L  W  C  *  C  D  E  S  H  S  *  H  L  Y  S  S  C  A
 9301 - CAACCTGTGGGTGCTTTAGATGTGTCTGCTTCAGTAGTGGCTGGTGGTATTATTGCCATA -
9360
       -Q  P  V  G  A  L  D  V  S  A  S  V  V  A  G  G  I  I  A  I
       - N  L  W  V  L  *  M  C  L  L  Q  *  W  L  V  V  L  L  P  Y
       -  T  C  G  C  F  R  C  V  C  F  S  S  G  W  W  Y  Y  C  H  I
 9361 - TTGGTGACTTGTGCTGCCTACTACTTTATGAAATTCAGACGTGTTTTTGGTGAGTACAAC -
9420
       -L  V  T  C  A  A  Y  Y  F  M  K  F  R  R  V  F  G  E  Y  N
       - W  *  L  V  L  P  T  T  L  *  N  S  D  V  F  L  V  S  T  T
       -  G  D  L  C  C  L  L  L  Y  E  I  Q  T  C  F  W  *  V  Q  P
 9421 - CATGTTGTTGCTGCTAATGCACTTTTGTTTTTGATGTCTTTCACTATACTCTGTCTGGTA -
9480
       -H  V  V  A  A  N  A  L  L  F  L  M  S  F  T  I  L  C  L  V
       - M  L  L  L  L  M  H  F  C  F  *  C  L  S  L  Y  S  V  W  Y
       -  C  C  C  C  *  C  T  F  V  F  D  V  F  H  Y  T  L  S  G  T
 9481 - CCAGCTTACAGCTTTCTGCCGGGAGTCTACTCAGTCTTTTACTTGTACTTGACATTCTAT -
9540
       -P  A  Y  S  F  L  P  G  V  Y  S  V  F  Y  L  Y  L  T  F  Y
       - Q  L  T  A  F  C  R  E  S  T  Q  S  F  T  C  T  *  H  S  I
       -  S  L  Q  L  S  A  G  S  L  L  S  L  L  L  V  L  D  I  L  F
 9541 - TTCACCAATGATGTTTCATTCTTGGCTCACCTTCAATGGTTTGCCATGTTTTCTCCTATT -
9600
       -F  T  N  D  V  S  F  L  A  H  L  Q  W  F  A  M  F  S  P  I
       - S  P  M  M  F  H  S  W  L  T  F  N  G  L  P  C  F  L  L  L
       -  H  Q  *  C  F  I  L  G  S  P  S  M  V  C  H  V  F  S  Y  C
 9601 - GTGCCTTTTTGGATAACAGCAATCTATGTATTCTGTATTTCTCTGAAGCACTGCCATTGG -
9660
       -V  P  F  W  I  T  A  I  Y  V  F  C  I  S  L  K  H  C  H  W
       - C  L  F  G  *  Q  Q  S  M  Y  S  V  F  L  *  S  T  A  I  G
       -  A  F  L  D  N  S  N  L  C  I  L  Y  F  S  E  A  L  P  L  V
 9661 - TTCTTTAACAACTATCTTAGGAAAAGAGTCATGTTTAATGGAGTTACATTTAGTACCTTC -
9720
       -F  F  N  N  Y  L  R  K  R  V  M  F  N  G  V  T  F  S  T  F
       - S  L  T  T  I  L  G  K  E  S  C  L  M  E  L  H  L  V  P  S
       -  L  *  Q  L  S  *  E  K  S  H  V  *  W  S  Y  I  *  Y  L  R
 9721 - GAGGAGGCTGCTTTGTGTACCTTTTTGCTCAACAAGGAAATGTACCTAAAATTGCGTAGC -
9780
       -E  E  A  A  L  C  T  F  L  L  N  K  E  M  Y  L  K  L  R  S
       - R  R  L  L  C  V  P  F  C  S  T  R  K  C  T  *  N  C  V  A
       -  G  G  C  F  V  Y  L  F  A  Q  Q  G  N  V  P  K  I  A  *  R
 9781 - GAGACACTGTTGCCACTTACACAGTATAACAGGTATCTTGCTCTATATAACAAGTACAAG -
9840
       -E  T  L  L  P  L  T  Q  Y  N  R  Y  L  A  L  Y  N  K  Y  K
       - R  H  C  C  H  L  H  S  I  T  G  I  L  L  Y  I  T  S  T  S
       -  D  T  V  A  T  Y  T  V  *  Q  V  S  C  S  I  *  Q  V  Q  V
 9841 - TATTTCAGTGGAGCCTTAGATACTACCAGCTATCGTGAAGCAGCTTGCTGCCACTTAGCA -
9900
       -Y  F  S  G  A  L  D  T  T  S  Y  R  E  A  A  C  C  H  L  A
       - I  S  V  E  P  *  I  L  P  A  I  V  K  Q  L  A  A  T  *  Q
       -  F  Q  W  S  L  R  Y  Y  Q  L  S  *  S  S  L  L  P  L  S  K
 9901 - AAGGCTCTAAATGACTTTAGCAACTCAGGTGCTGATGTTCTCTACCAACCACCACAGACA -
```

```
9960
     -K  A  L  N  D  F  S  N  S  G  A  D  V  L  Y  Q  P  P  Q  T
     - R  L  *  M  T  L  A  T  Q  V  L  M  F  S  T  N  H  H  R  H
     -  G  S  K  *  L  *  Q  L  R  C  *  C  S  L  P  T  T  T  D  I
```

Figure 16-N

```
 9961 - TCAATCACTTCTGCTGTTCTGCAGAGTGGTTTTAGGAAAATGGCATTCCCGTCAGGCAAA -
10020
      -S  I  T  S  A  V  L  Q  S  G  F  R  K  M  A  F  P  S  G  K
      - Q  S  L  L  L  F  C  R  V  V  L  G  K  W  H  S  R  Q  A  K
      -  N  H  F  C  C  S  A  E  W  F  *  E  N  G  I  P  V  R  Q  S
10021 - GTTGAAGGGTGCATGGTACAAGTAACCTGTGGAACTACAACTCTTAATGGATTGTGGTTG -
10080
      -V  E  G  C  M  V  Q  V  T  C  G  T  T  T  L  N  G  L  W  L
      - L  K  G  A  W  Y  K  *  P  V  E  L  Q  L  L  M  D  C  G  W
      -  *  R  V  H  G  T  S  N  L  W  N  Y  N  S  *  W  I  V  V  G
10081 - GATGACACAGTATACTGTCCAAGACATGTCATTTGCACAGCAGAAGACATGCTTAATCCT -
10140
      -D  D  T  V  Y  C  P  R  H  V  I  C  T  A  E  D  M  L  N  P
      - M  T  Q  Y  T  V  Q  D  M  S  F  A  Q  Q  K  T  C  L  I  L
      -  *  H  S  I  L  S  K  T  C  H  L  H  S  R  R  H  A  *  S  *
10141 - AACTATGAAGATCTGCTCATTCGCAAATCCAACCATAGCTTTCTTGTTCAGGCTGGCAAT -
10200
      -N  Y  E  D  L  L  I  R  K  S  N  H  S  F  L  V  Q  A  G  N
      - T  M  K  I  C  S  F  A  N  P  T  I  A  F  L  F  R  L  A  M
      -  L  *  R  S  A  H  S  Q  I  Q  P  *  L  S  C  S  G  W  Q  C
10201 - GTTCAACTTCGTGTTATTGGCCATTCTATGCAAAATTGTCTGCTTAGGCTTAAAGTTGAT -
10260
      -V  Q  L  R  V  I  G  H  S  M  Q  N  C  L  L  R  L  K  V  D
      - F  N  F  V  L  L  A  I  L  C  K  I  V  C  L  G  L  K  L  I
      -  S  T  S  C  Y  W  P  F  Y  A  K  L  S  A  *  A  *  S  *  Y
10261 - ACTTCTAACCCTAAGACACCCAAGTATAAATTTGTCCGTATCCAACCTGGTCAAACATTT -
10320
      -T  S  N  P  K  T  P  K  Y  K  F  V  R  I  Q  P  G  Q  T  F
      - L  L  T  L  R  H  P  S  I  N  L  S  V  S  N  L  V  K  H  F
      -  F  *  P  *  D  T  Q  V  *  I  C  P  Y  P  T  W  S  N  I  F
10321 - TCAGTTCTAGCATGCTACAATGGTTCACCATCTGGTGTTTATCAGTGTGCCATGAGACCT -
10380
      -S  V  L  A  C  Y  N  G  S  P  S  G  V  Y  Q  C  A  M  R  P
      - Q  F  *  H  A  T  M  V  H  H  L  V  F  I  S  V  P  *  D  L
      -  S  S  S  M  L  Q  W  F  T  I  W  C  L  S  V  C  H  E  T  *
10381 - AATCATACCATTAAAGGTTCTTTCCTTAATGGATCATGTGGTAGTGTTGGTTTTAACATT -
10440
      -N  H  T  I  K  G  S  F  L  N  G  S  C  G  S  V  G  F  N  I
      - I  I  P  L  K  V  L  S  L  M  D  H  V  V  V  V  L  V  L  T  L
      -  S  Y  H  *  R  F  F  P  *  W  I  M  W  *  C  W  F  *  H  *
10441 - GATTATGATTGCGTGTCTTTCTGCTATATGCATCATATGGAGCTTCCAACAGGAGTACAC -
10500
      -D  Y  D  C  V  S  F  C  Y  M  H  H  M  E  L  P  T  G  V  H
      - I  M  I  A  C  L  S  A  I  C  I  I  W  S  F  Q  Q  E  Y  T
      -  L  *  L  R  V  F  L  L  Y  A  S  Y  G  A  S  N  R  S  T  R
```

```
10501 - GCTGGTACTGACTTAGAAGGTAAATTCTATGGTCCATTTGTTGACAGACAAACTGCACAG -
10560
      -A  G  T  D  L  E  G  K  F  Y  G  P  F  V  D  R  Q  T  A  Q
      - L  V  L  T  *  K  V  N  S  M  V  H  L  L  T  D  K  L  H  R
      -  W  Y  *  L  R  R  *  I  L  W  S  I  C  *  Q  T  N  C  T  G
10561 - GCTGCAGGTACAGACACAACCATAACATTAAATGTTTTGGCATGGCTGTATGCTGCTGTT -
10620
      -A  A  G  T  D  T  T  I  T  L  N  V  L  A  W  L  Y  A  A  V
      - L  Q  V  Q  T  Q  P  *  H  *  M  F  W  H  G  C  M  L  L  L
      -  C  R  Y  R  H  N  H  N  I  K  C  F  G  M  A  V  C  C  C  Y
10621 - ATCAATGGTGATAGGTGGTTTCTTAATAGATTCACCACTACTTTGAATGACTTTAACCTT -
10680
      -I  N  G  D  R  W  F  L  N  R  F  T  T  T  L  N  D  F  N  L
      - S  M  V  I  G  G  F  L  I  D  S  P  L  L  *  M  T  L  T  L
      -  Q  W  *  *  V  V  S  *  *  I  H  H  Y  F  E  *  L  *  P  C
```

Figure 16-O

```
10681 - GTGGCAATGAAGTACAACTATGAACCTTTGACACAAGATCATGTTGACATATTGGGACCT -
10740
      -V  A  M  K  Y  N  Y  E  P  L  T  Q  D  H  V  D  I  L  G  P
      - W  Q  *  S  T  T  M  N  L  *  H  K  I  M  L  T  Y  W  D  L
      -  G  N  E  V  Q  L  *  T  F  D  T  R  S  C  *  H  I  G  T  S
10741 - CTTTCTGCTCAAACAGGAATTGCCGTCTTAGATATGTGTGCTGCTTTGAAAGAGCTGCTG -
10800
      -L  S  A  Q  T  G  I  A  V  L  D  M  C  A  A  L  K  E  L  L
      - F  L  L  K  Q  E  L  P  S  *  I  C  V  L  L  *  K  S  C  C
      -  F  C  S  N  R  N  C  R  L  R  Y  V  C  C  F  E  R  A  A  A
10801 - CAGAATGGTATGAATGGTCGTACTATCCTTGGTAGCACTATTTTAGAAGATGAGTTTACA -
10860
      -Q  N  G  M  N  G  R  T  I  L  G  S  T  I  L  E  D  E  F  T
      - R  M  V  *  M  V  V  L  S  L  V  A  L  F  *  K  M  S  L  H
      -  E  W  Y  E  W  S  Y  Y  P  W  *  H  Y  F  R  R  *  V  Y  T
10861 - CCATTTGATGTTGTTAGACAATGCTCTGGTGTTACCTTCCAAGGTAAGTTCAAGAAAATT -
10920
      -P  F  D  V  V  R  Q  C  S  G  V  T  F  Q  G  K  F  K  K  I
      - H  L  M  L  L  D  N  A  L  V  L  P  S  K  V  S  S  R  K  L
      -  I  *  C  C  *  T  M  L  W  C  Y  L  P  R  *  V  Q  E  N  C
10921 - GTTAAGGGCACTCATCATTGGATGCTTTTAACTTTCTTGACATCACTATTGATTCTTGTT -
10980
      -V  K  G  T  H  H  W  M  L  L  T  F  L  T  S  L  L  I  L  V
      - L  R  A  L  I  I  G  C  F  *  L  S  *  H  H  Y  *  F  L  F
      -  *  G  H  S  S  L  D  A  F  N  F  L  D  I  T  I  D  S  C  S
10981 - CAAAGTACACAGTGGTCACTGTTTTTCTTTGTTTACGAGAATGCTTTCTTGCCATTTACT -
11040
      -Q  S  T  Q  W  S  L  F  F  F  V  Y  E  N  A  F  L  P  F  T
      - K  V  H  S  G  H  C  F  S  L  F  T  R  M  L  S  C  H  L  L
      -  K  Y  T  V  V  T  V  F  L  C  L  R  E  C  F  L  A  I  Y  S
11041 - CTTGGTATTATGGCAATTGCTGCATGTGCTATGCTGCTTGTTAAGCATAAGCACGCATTC -
```

```
11100
       -L  G  I  M  A  I  A  A  C  A  M  L  L  V  K  H  K  H  A  F
       - L  V  L  W  Q  L  L  H  V  L  C  C  L  L  S  I  S  T  H  S
       -  W  Y  Y  G  N  C  C  M  C  Y  A  A  C  *  A  *  A  R  I  L
 11101 - TTGTGCTTGTTTCTGTTACCTTCTCTTGCAACAGTTGCTTACTTTAATATGGTCTACATG -
11160
       -L  C  L  F  L  L  P  S  L  A  T  V  A  Y  F  N  M  V  Y  M
       - C  A  C  F  C  Y  L  L  L  Q  Q  L  L  T  L  I  W  S  T  C
       -  V  L  V  S  V  T  F  S  C  N  S  C  L  L  *  Y  G  L  H  A
 11161 - CCTGCTAGCTGGGTGATGCGTATCATGACATGGCTTGAATTGGCTGACACTAGCTTGTCT -
11220
       -P  A  S  W  V  M  R  I  M  T  W  L  E  L  A  D  T  S  L  S
       - L  L  A  G  *  C  V  S  *  H  G  L  N  W  L  T  L  A  C  L
       -  C  *  L  G  D  A  Y  H  D  M  A  *  I  G  *  H  *  L  V  W
 11221 - GGTTATAGGCTTAAGGATTGTGTTATGTATGCTTCAGCTTTAGTTTTGCTTATTCTCATG -
11280
       -G  Y  R  L  K  D  C  V  M  Y  A  S  A  L  V  L  L  I  L  M
       - V  I  G  L  R  I  V  L  C  M  L  Q  L  *  F  C  L  F  S  *
       -  L  *  A  *  G  L  C  Y  V  C  F  S  F  S  F  A  Y  S  H  D
 11281 - ACAGCTCGCACTGTTTATGATGATGCTGCTAGACGTGTTTGGACACTGATGAATGTCATT -
11340
       -T  A  R  T  V  Y  D  D  A  A  R  R  V  W  T  L  M  N  V  I
       - Q  L  A  L  F  M  M  M  L  L  D  V  F  G  H  *  *  M  S  L
       -  S  S  H  C  L  *  *  C  C  *  T  C  L  D  T  D  E  C  H  Y
 11341 - ACACTTGTTTACAAAGTCTACTATGGTAATGCTTTAGATCAAGCTATTTCCATGTGGGCC -
11400
       -T  L  V  Y  K  V  Y  Y  G  N  A  L  D  Q  A  I  S  M  W  A
       - H  L  F  T  K  S  T  M  V  M  L  *  I  K  L  F  P  C  G  P
       -  T  C  L  Q  S  L  L  W  *  C  F  R  S  S  Y  F  H  V  G  L
```

Figure 16-P

```
 11401 - TTAGTTATTTCTGTAACCTCTAACTATTCTGGTGTCGTTACGACTATCATGTTTTTAGCT -
11460
       -L  V  I  S  V  T  S  N  Y  S  G  V  V  T  T  I  M  F  L  A
       - *  L  F  L  *  P  L  T  I  L  V  S  L  R  L  S  C  F  *  L
       -  S  Y  F  C  N  L  *  L  F  W  C  R  Y  D  Y  H  V  F  S  *
 11461 - AGAGCTATAGTGTTTGTGTGTGTTGAGTATTACCCATTGTTATTTATTACTGGCAACACC -
11520
       -R  A  I  V  F  V  C  V  E  Y  Y  P  L  L  F  I  T  G  N  T
       - E  L  *  C  L  C  V  L  S  I  T  H  C  Y  L  L  L  A  T  P
       -  S  Y  S  V  C  V  C  *  V  L  P  I  V  I  Y  Y  W  Q  H  L
 11521 - TTACAGTGTATCATGCTTGTTTATTGTTTCTTAGGCTATTGTTGCTGCTGCTACTTTGGC -
11580
       -L  Q  C  I  M  L  V  Y  C  F  L  G  Y  C  C  C  C  Y  F  G
       - Y  S  V  S  C  L  F  I  V  S  *  A  I  V  A  A  A  T  L  A
       -  T  V  Y  H  A  C  L  L  F  L  R  L  L  L  L  L  L  L  W  P
 11581 - CTTTTCTGTTTACTCAACCGTTACTTCAGGCTTACTCTTGGTGTTTATGACTACTTGGTC -
11640
       -L  F  C  L  L  N  R  Y  F  R  L  T  L  G  V  Y  D  Y  L  V
       - F  S  V  Y  S  T  V  T  S  G  L  L  L  V  F  M  T  T  W  S
       -  F  L  F  T  Q  P  L  L  Q  A  Y  S  W  C  L  *  L  L  G  L
```

```
11641 - TCTACACAAGAATTTAGGTATATGAACTCCCAGGGGCTTTTGCCTCCTAAGAGTAGTATT -
11700
      - S  T  Q  E  F  R  Y  M  N  S  Q  G  L  L  P  P  K  S  S  I
      -  L  H  K  N  L  G  I  *  T  P  R  G  F  C  L  L  R  V  V  L
      -   Y  T  R  I  *  V  Y  E  L  P  G  A  F  A  S  *  E  *  Y  *
11701 - GATGCTTTCAAGCTTAACATTAAGTTGTTGGGTATTGGAGGTAAACCATGTATCAAGGTT -
11760
      - D  A  F  K  L  N  I  K  L  L  G  I  G  G  K  P  C  I  K  V
      -  M  L  S  S  L  T  L  S  C  W  V  L  E  V  N  H  V  S  R  L
      -   C  F  Q  A  *  H  *  V  V  G  Y  W  R  *  T  M  Y  Q  G  C
11761 - GCTACTGTACAGTCTAAAATGTCTGACGTAAAGTGCACATCTGTGGTACTGCTCTCGGTT -
11820
      - A  T  V  Q  S  K  M  S  D  V  K  C  T  S  V  V  L  L  S  V
      -  L  L  Y  S  L  K  C  L  T  *  S  A  H  L  W  Y  C  S  R  F
      -   Y  C  T  V  *  N  V  *  R  K  V  H  I  C  G  T  A  L  G  S
11821 - CTTCAACAACTTAGAGTAGAGTCATCTTCTAAATTGTGGGCACAATGTGTACAACTCCAC -
11880
      - L  Q  Q  L  R  V  E  S  S  S  K  L  W  A  Q  C  V  Q  L  H
      -  F  N  N  L  E  *  S  H  L  L  N  C  G  H  N  V  Y  N  S  T
      -   S  T  T  *  S  R  V  I  F  *  I  V  G  T  M  C  T  T  P  Q
11881 - AATGATATTCTTCTTGCAAAAGACACAACTGAAGCTTTCGAGAAGATGGTTTCTCTTTTG -
11940
      - N  D  I  L  L  A  K  D  T  T  E  A  F  E  K  M  V  S  L  L
      -  M  I  F  F  L  Q  K  T  Q  L  K  L  S  R  R  W  F  L  F  C
      -   *  Y  S  S  C  K  R  H  N  *  S  F  R  E  D  G  F  S  F  V
11941 - TCTGTTTTGCTATCCATGCAGGGTGCTGTAGACATTAATAGGTTGTGCGAGGAAATGCTC -
12000
      - S  V  L  L  S  M  Q  G  A  V  D  I  N  R  L  C  E  E  M  L
      -  L  F  C  Y  P  C  R  V  L  *  T  L  I  G  C  A  R  K  C  S
      -   C  F  A  I  H  A  G  C  C  R  H  *  *  V  V  R  G  N  A  R
12001 - GATAACCGTGCTACTCTTCAGGCTATTGCTTCAGAATTTAGTTCTTTACCATCATATGCC -
12060
      - D  N  R  A  T  L  Q  A  I  A  S  E  F  S  S  L  P  S  Y  A
      -  I  T  V  L  L  F  R  L  L  L  Q  N  L  V  L  Y  H  H  M  P
      -   *  P  C  Y  S  S  G  Y  C  F  R  I  *  F  F  T  I  I  C  R
12061 - GCTTATGCCACTGCCCAGGAGGCCTATGAGCAGGCTGTAGCTAATGGTGATTCTGAAGTC -
12120
      - A  Y  A  T  A  Q  E  A  Y  E  Q  A  V  A  N  G  D  S  E  V
      -  L  M  P  L  P  R  R  P  M  S  R  L  *  L  M  V  I  L  K  S
      -   L  C  H  C  P  G  G  L  *  A  G  C  S  *  W  *  F  *  S  R
```

Figure 16-Q

```
12121 - GTTCTCAAAAAGTTAAAGAAATCTTTGAATGTGGCTAAATCTGAGTTTGACCGTGATGCT -
12180
      - V  L  K  K  L  K  K  S  L  N  V  A  K  S  E  F  D  R  D  A
      -  F  S  K  S  *  R  N  L  *  M  W  L  N  L  S  L  T  V  M  L
      -   S  Q  K  V  K  E  I  F  E  C  G  *  I  *  V  *  P  *  C  C
12181 - GCCATGCAACGCAAGTTGGAAAAGATGGCAGATCAGGCTATGACCCAAATGTACAAACAG -
12240
      - A  M  Q  R  K  L  E  K  M  A  D  Q  A  M  T  Q  M  Y  K  Q
      -  P  C  N  A  S  W  K  R  W  Q  I  R  L  *  P  K  C  T  N  R
```

```
      - H A T Q V G K D G R S G Y D P N V Q T G
12241 - GCAAGATCTGAGGACAAGAGGGCAAAAGTAACTAGTGCTATGCAAACAATGCTCTTCACT -
12300
      -A R S E D K R A K V T S A M Q T M L F T
      - Q D L R T R G Q K * L V L C K Q C S S L
      -  K I * G Q E G K S N * C Y A N N A L H Y
12301 - ATGCTTAGGAAGCTTGATAATGATGCACTTAACAACATTATCAACAATGCGCGTGATGGT -
12360
      -M L R K L D N D A L N N I I N N A R D G
      - C L G S L I M M H L T T L S T M R V M V
      -  A * E A * * * C T * Q H Y Q Q C A * W L
12361 - TGTGTTCCACTCAACATCATACCATTGACTACAGCAGCCAAACTCATGGTTGTTGTCCCT -
12420
      -C V P L N I I P L T T A A K L M V V V P
      - V F H S T S Y H * L Q Q P N S W L L S L
      -  C S T Q H H T I D Y S S Q T H G C C P *
12421 - GATTATGGTACCTACAAGAACACTTGTGATGGTAACACCTTTACATATGCATCTGCACTC -
12480
      -D Y G T Y K N T C D G N T F T Y A S A L
      - I M V P T R T L V M V T P L H M H L H S
      -  L W Y L Q E H L * W * H L Y I C I C T L
12481 - TGGGAAATCCAGCAAGTTGTTGATGCGGATAGCAAGATTGTTCAACTTAGTGAAATTAAC -
12540
      -W E I Q Q V V D A D S K I V Q L S E I N
      - G K S S K L L M R I A R L F N L V K L T
      -  G N P A S C * C G * Q D C S T * * N * H
12541 - ATGGACAATTCACCAAATTTGGCTTGGCCTCTTATTGTTACAGCTCTAAGAGCCAACTCA -
12600
      -M D N S P N L A W P L I V T A L R A N S
      - W T I H Q I W L G L L L L Q L * E P T Q
      -  G Q F T K F G L A S Y C Y S S K S Q L S
12601 - GCTGTTAAACTACAGAATAATGAACTGAGTCCAGTAGCACTACGACAGATGTCCTGTGCG -
12660
      -A V K L Q N N E L S P V A L R Q M S C A
      - L L N Y R I M N * V Q * H Y D R C P V R
      -  C * T T E * * T E S S S T T T D V L C G
12661 - GCTGGTACCACACAAACAGCTTGTACTGATGACAATGCACTTGCCTACTATAACAATTCG -
12720
      -A G T T Q T A C T D D N A L A Y Y N N S
      - L V P H K Q L V L M T M H L P T I T I R
      -  W Y H T N S L Y * * Q C T C L L * Q F E
12721 - AAGGGAGGTAGGTTTGTGCTGGCATTACTATCAGACCACCAAGATCTCAAATGGGCTAGA -
12780
      -K G G R F V L A L L S D H Q D L K W A R
      - R E V G L C W H Y Y Q T T K I S N G L D
      -  G R * V C A G I T I R P P R S Q M G * I
12781 - TTCCCTAAGAGTGATGGTACAGGTACAATTTACACAGAACTGGAACCACCTTGTAGGTTT -
12840
      -F P K S D G T G T I Y T E L E P P C R F
      - S L R V M V Q V Q F T Q N W N H L V G L
      -  P * E * W Y R Y N L H R T G T T L * V C
12841 - GTTACAGACACACCAAAAGGGCCTAAAGTGAAATACTTGTACTTCATCAAAGGCTTAAAC -
12900
      -V T D T P K G P K V K Y L Y F I K G L N
      - L Q T H Q K G L K * N T C T S S K A * T
      -  Y R H T K R A * S E I L V L H Q R L K Q
```

Figure 16-R

```
12901 - AACCTAAATAGAGGTATGGTGCTGGGCAGTTTAGCTGCTACAGTACGTCTTCAGGCTGGA -
12960
      - N  L  N  R  G  M  V  L  G  S  L  A  A  T  V  R  L  Q  A  G
      -  T  *  I  E  V  W  C  W  A  V  *  L  L  Q  Y  V  F  R  L  E
      -   P  K  *  R  Y  G  A  G  Q  F  S  C  Y  S  T  S  S  G  W  K
12961 - AATGCTACAGAAGTACCTGCCAATTCAACTGTGCTTTCCTTCTGTGCTTTTGCAGTAGAC -
13020
      - N  A  T  E  V  P  A  N  S  T  V  L  S  F  C  A  F  A  V  D
      -  M  L  Q  K  Y  L  P  I  Q  L  C  F  P  S  V  L  L  Q  *  T
      -   C  Y  R  S  T  C  Q  F  N  C  A  F  L  L  C  F  C  S  R  P
13021 - CCTGCTAAAGCATATAAGGATTACCTAGCAAGTGGAGGACAACCAATCACCAACTGTGTG -
13080
      - P  A  K  A  Y  K  D  Y  L  A  S  G  G  Q  P  I  T  N  C  V
      -  L  L  K  H  I  R  I  T  *  Q  V  E  D  N  Q  S  P  T  V  *
      -   C  *  S  I  *  G  L  P  S  K  W  R  T  T  N  H  Q  L  C  E
13081 - AAGATGTTGTGTACACACACTGGTACAGGACAGGCAATTACTGTAACACCAGAAGCTAAC -
13140
      - K  M  L  C  T  H  T  G  T  G  Q  A  I  T  V  T  P  E  A  N
      -  R  C  C  V  H  T  L  V  Q  D  R  Q  L  L  *  H  Q  K  L  T
      -   D  V  V  Y  T  H  W  Y  R  T  G  N  Y  C  N  T  R  S  *  H
13141 - ATGGACCAAGAGTCCTTTGGTGGTGCTTCATGTTGTCTGTATTGTAGATGCCACATTGAC -
13200
      - M  D  Q  E  S  F  G  G  A  S  C  C  L  Y  C  R  C  H  I  D
      -  W  T  K  S  P  L  V  V  L  H  V  V  C  I  V  D  A  T  L  T
      -   G  P  R  V  L  W  W  C  F  M  L  S  V  L  *  M  P  H  *  P
13201 - CATCCAAATCCTAAAGGATTCTGTGACTTGAAAGGTAAGTACGTCCAAATACCTACCACT -
13260
      - H  P  N  P  K  G  F  C  D  L  K  G  K  Y  V  Q  I  P  T  T
      -  I  Q  I  L  K  D  S  V  T  *  K  V  S  T  S  K  Y  L  P  L
      -   S  K  S  *  R  I  L  *  L  E  R  *  V  R  P  N  T  Y  H  L
13261 - TGTGCTAATGACCCAGTGGGTTTTACACTTAGAAACACAGTCTGTACCGTCTGCGGAATG -
13320
      - C  A  N  D  P  V  G  F  T  L  R  N  T  V  C  T  V  C  G  M
      -  V  L  M  T  Q  W  V  L  H  L  E  T  Q  S  V  P  S  A  E  C
      -   C  *  *  P  S  G  F  Y  T  *  K  H  S  L  Y  R  L  R  N  V
13321 - TGGAAAGGTTATGGCTGTAGTTGTGACCAACTCCGCGAACCCTTGATGCAGTCTGCGGAT -
13380
      - W  K  G  Y  G  C  S  C  D  Q  L  R  E  P  L  M  Q  S  A  D
      -  G  K  V  M  A  V  V  V  T  N  S  A  N  P  *  C  S  L  R  M
      -   E  R  L  W  L  *  L  *  P  T  P  R  T  L  D  A  V  C  G  C
13381 - GCATCAACGTTTTTAAACGGGTTTGCGGTGTAAGTGCAGCCCGTCTTACACCGTGCGGCA -
13440
      - A  S  T  F  L  N  G  F  A  V  *  V  Q  P  V  L  H  R  A  A
      -  H  Q  R  F  *  T  G  L  R  C  K  C  S  P  S  Y  T  V  R  H
      -   I  N  V  F  K  R  V  C  G  V  S  A  A  R  L  T  P  C  G  T
13441 - CAGGCACTAGTACTGATGTCGTCTACAGGGCTTTTGATATTTACAACGAAAAAAGTGCTG -
13500
      - Q  A  L  V  L  M  S  S  T  G  L  L  I  F  T  T  K  K  V  L
      -  R  H  *  Y  *  C  R  L  Q  G  F  *  Y  L  Q  R  K  K  C  W
      -   G  T  S  T  D  V  V  Y  R  A  F  D  I  Y  N  E  K  S  A  G
13501 - GTTTTGCAAAGTTCCTAAAAACTAATTGCTGTCGCTTCCAGGAGAAGGATGAGGAAGGCA -
13560
      - V  L  Q  S  S  *  K  L  I  A  V  A  S  R  R  R  M  R  K  A
      -  F  C  K  V  P  K  N  *  L  L  S  L  P  G  E  G  *  G  R  Q
      -   F  A  K  F  L  K  T  N  C  C  R  F  Q  E  K  D  E  E  G  N
13561 - ATTTATTAGACTCTTACTTTGTAGTTAAGAGGCATACTATGTCTAACTACCAACATGAAG -
13620
```

```
      - I  Y  *  T  L  T  L  *  L  R  G  I  L  C  L  T  T  N  M  K
      -  F  I  R  L  L  L  C  S  *  E  A  Y  Y  V  *  L  P  T  *  R
      -   L  L  D  S  Y  F  V  V  K  R  H  T  M  S  N  Y  Q  H  E  E
13621 - AGACTATTTATAACTTGGTTAAAGATTGTCCAGCGGTTGCTGTCCATGACTTTTTCAAGT -
13680
      - R  L  F  I  T  W  L  K  I  V  Q  R  L  L  S  M  T  F  S  S
      -  D  Y  L  *  L  G  *  R  L  S  S  G  C  C  P  *  L  F  Q  V
      -   T  I  Y  N  L  V  K  D  C  P  A  V  A  V  H  D  F  F  K  F
```

Figure 16-S

```
13681 - TTAGAGTAGATGGTGACATGGTACCACATATATCACGTCAGCGTCTAACTAAATACACAA -
13740
      -L  E  *  M  V  T  W  Y  H  I  Y  H  V  S  V  *  L  N  T  Q
      - *  S  R  W  *  H  G  T  T  Y  I  T  S  A  S  N  *  I  H  N
      -  R  V  D  G  D  M  V  P  H  I  S  R  Q  R  L  T  K  Y  T  M
13741 - TGGCTGATTTAGTCTATGCTCTACGTCATTTTGATGAGGGTAATTGTGATACATTAAAAG -
13800
      -W  L  I  *  S  M  L  Y  V  I  L  M  R  V  I  V  I  H  *  K
      - G  *  F  S  L  C  S  T  S  F  *  *  G  *  L  *  Y  I  K  R
      -  A  D  L  V  Y  A  L  R  H  F  D  E  G  N  C  D  T  L  K  E
13801 - AAATACTCGTCACATACAATTGCTGTGATGATGATTATTTCAATAAGAAGGATTGGTATG -
13860
      -K  Y  S  S  H  T  I  A  V  M  M  I  I  S  I  R  R  I  G  M
      - N  T  R  H  I  Q  L  L  *  *  *  L  F  Q  *  E  G  L  V  *
      -  I  L  V  T  Y  N  C  C  D  D  D  Y  F  N  K  K  D  W  Y  D
13861 - ACTTCGTAGAGAATCCTGACATCTTACGCGTATATGCTAACTTAGGTGAGCGTGTACGCC -
13920
      -T  S  *  R  I  L  T  S  Y  A  Y  M  L  T  *  V  S  V  Y  A
      - L  R  R  E  S  *  H  L  T  R  I  C  *  L  R  *  A  C  T  P
      -  F  V  E  N  P  D  I  L  R  V  Y  A  N  L  G  E  R  V  R  Q
13921 - AATCATTATTAAAGACTGTACAATTCTGCGATGCTATGCGTGATGCAGGCATTGTAGGCG -
13980
      -N  H  Y  *  R  L  Y  N  S  A  M  L  C  V  M  Q  A  L  *  A
      - I  I  I  K  D  C  T  I  L  R  C  Y  A  *  C  R  H  C  R  R
      -  S  L  L  K  T  V  Q  F  C  D  A  M  R  D  A  G  I  V  G  V
13981 - TACTGACATTAGATAATCAGGATCTTAATGGGAACTGGTACGATTTCGGTGATTTCGTAC -
14040
      -Y  *  H  *  I  I  R  I  L  M  G  T  G  T  I  S  V  I  S  Y
      - T  D  I  R  *  S  G  S  *  W  E  L  V  R  F  R  *  F  R  T
      -  L  T  L  D  N  Q  D  L  N  G  N  W  Y  D  F  G  D  F  V  Q
14041 - AAGTAGCACCAGGCTGCGGAGTTCCTATTGTGGATTCATATTACTCATTGCTGATGCCCA -
14100
      -K  *  H  Q  A  A  E  F  L  L  W  I  H  I  T  H  C  *  C  P
      - S  S  T  R  L  R  S  S  Y  C  G  F  I  L  L  I  A  D  A  H
      -  V  A  P  G  C  G  V  P  I  V  D  S  Y  Y  S  L  L  M  P  I
14101 - TCCTCACTTTGACTAGGGCATTGGCTGCTGAGTCCCATATGGATGCTGATCTCGCAAAAC -
14160
      -S  S  L  *  L  G  H  W  L  L  S  P  I  W  M  L  I  S  Q  N
      - P  H  F  D  *  G  I  G  C  *  V  P  Y  G  C  *  S  R  K  T
      -  L  T  L  T  R  A  L  A  A  E  S  H  M  D  A  D  L  A  K  P
14161 - CACTTATTAAGTGGGATTTGCTGAAATATGATTTTACGGAAGAGAGACTTTGTCTCTTCG -
14220
      -H  L  L  S  G  I  C  *  N  M  I  L  R  K  R  D  F  V  S  S
      - T  Y  *  V  G  F  A  E  I  *  F  Y  G  R  E  T  L  S  L  R
      -  L  I  K  W  D  L  L  K  Y  D  F  T  E  E  R  L  C  L  F  D
14221 - ACCGTTATTTTAAATATTGGGACCAGACATACCATCCCAATTGTATTAACTGTTTGGATG -
14280
      -T  V  I  L  N  I  G  T  R  H  T  I  P  I  V  L  T  V  W  M
      - P  L  F  *  I  L  G  P  D  I  P  S  Q  L  Y  *  L  F  G  *
      -  R  Y  F  K  Y  W  D  Q  T  Y  H  P  N  C  I  N  C  L  D  D
14281 - ATAGGTGTATCCTTCATTGTGCAAACTTTAATGTGTTATTTTCTACTGTGTTTCCACCTA -
14340
      -I  G  V  S  F  I  V  Q  T  L  M  C  Y  F  L  L  C  F  H  L
      - *  V  Y  P  S  L  C  K  L  *  C  V  I  F  Y  C  V  S  T  Y
      -  R  C  I  L  H  C  A  N  F  N  V  L  F  S  T  V  F  P  P  T
14341 - CAAGTTTTGGACCACTAGTAAGAAAAATATTTGTAGATGGTGTTCCTTTTGTTGTTTCAA -
14400
```

```
      - Q  V  L  D  H  *  *  E  K  Y  L  *  M  V  F  L  L  L  F  Q
      -  K  F  W  T  T  S  K  K  N  I  C  R  W  C  S  F  C  C  F  N
      -   S  F  G  P  L  V  R  K  I  F  V  D  G  V  P  F  V  V  S  T
14401 - CTGGATACCATTTTCGTGAGTTAGGAGTCGTACATAATCAGGATGTAAACTTACATAGCT - 14460
      - L  D  T  I  F  V  S  *  E  S  Y  I  I  R  M  *  T  Y  I  A
      -  W  I  P  F  S  *  V  R  S  R  T  *  S  G  C  K  L  T  *  L
      -   G  Y  H  F  R  E  L  G  V  V  H  N  Q  D  V  N  L  H  S  S
```

Figure 16-T

```
14461 - CGCGTCTCAGTTTCAAGGAACTTTTAGTGTATGCTGCTGATCCAGCTATGCATGCAGCTT -
14520
        -R  V  S  V  S  R  N  F  *  C  M  L  L  I  Q  L  C  M  Q  L
        - A  S  Q  F  Q  G  T  F  S  V  C  C  *  S  S  Y  A  C  S  F
        -  R  L  S  F  K  E  L  L  V  Y  A  A  D  P  A  M  H  A  A  S
14521 - CTGGCAATTTATTGCTAGATAAACGCACTACATGCTTTTCAGTAGCTGCACTAACAAACA -
14580
        -L  A  I  Y  C  *  I  N  A  L  H  A  F  Q  *  L  H  *  Q  T
        - W  Q  F  I  A  R  *  T  H  Y  M  L  F  S  S  C  T  N  K  Q
        -  G  N  L  L  L  D  K  R  T  T  C  F  S  V  A  A  L  T  N  N
14581 - ATGTTGCTTTTCAAACTGTCAAACCCGGTAATTTTAATAAAGACTTTTATGACTTTGCTG -
14640
        -M  L  L  F  K  L  S  N  P  V  I  L  I  K  T  F  M  T  L  L
        - C  C  F  S  N  C  Q  T  R  *  F  *  *  R  L  L  *  L  C  C
        -  V  A  F  Q  T  V  K  P  G  N  F  N  K  D  F  Y  D  F  A  V
14641 - TGTCTAAAGGTTTCTTTAAGGAAGGAAGTTCTGTTGAACTAAAACACTTCTTCTTTGCTC -
14700
        -C  L  K  V  S  L  R  K  E  V  L  L  N  *  N  T  S  S  L  L
        - V  *  R  F  L  *  G  R  K  F  C  *  T  K  T  L  L  L  C  S
        -  S  K  G  F  F  K  E  G  S  S  V  E  L  K  H  F  F  F  A  Q
14701 - AGGATGGCAACGCTGCTATCAGTGATTATGACTATTATCGTTATAATCTGCCAACAATGT -
14760
        -R  M  A  T  L  L  S  V  I  M  T  I  I  V  I  I  C  Q  Q  C
        - G  W  Q  R  C  Y  Q  *  L  *  L  L  S  L  *  S  A  N  N  V
        -  D  G  N  A  A  I  S  D  Y  D  Y  Y  R  Y  N  L  P  T  M  C
14761 - GTGATATCAGACAACTCCTATTCGTAGTTGAAGTTGTTGATAAATACTTTGATTGTTACG -
14820
        -V  I  S  D  N  S  Y  S  *  L  K  L  L  I  N  T  L  I  V  T
        - *  Y  Q  T  T  P  I  R  S  *  S  C  *  *  I  L  *  L  L  R
        -  D  I  R  Q  L  L  F  V  V  E  V  V  D  K  Y  F  D  C  Y  D
14821 - ATGGTGGCTGTATTAATGCCAACCAAGTAATCGTTAACAATCTGGATAAATCAGCTGGTT -
14880
        -M  V  A  V  L  M  P  T  K  *  S  L  T  I  W  I  N  Q  L  V
        - W  W  L  Y  *  C  Q  P  S  N  R  *  Q  S  G  *  I  S  W  F
        -  G  G  C  I  N  A  N  Q  V  I  V  N  N  L  D  K  S  A  G  F
14881 - TCCCATTTAATAAATGGGGTAAGGCTAGACTTTATTATGACTCAATGAGTTATGAGGATC -
14940
        -S  H  L  I  N  G  V  R  L  D  F  I  M  T  Q  *  V  M  R  I
        - P  I  *  *  M  G  *  G  *  T  L  L  *  L  N  E  L  *  G  S
        -  P  F  N  K  W  G  K  A  R  L  Y  Y  D  S  M  S  Y  E  D  Q
14941 - AAGATGCACTTTTCGCGTATACTAAGCGTAATGTCATCCCTACTATAACTCAAATGAATC -
15000
        -K  M  H  F  S  R  I  L  S  V  M  S  S  L  L  *  L  K  *  I
        - R  C  T  F  R  V  Y  *  A  *  C  H  P  Y  Y  N  S  N  E  S
        -  D  A  L  F  A  Y  T  K  R  N  V  I  P  T  I  T  Q  M  N  L
15001 - TTAAGTATGCCATTAGTGCAAAGAATAGAGCTCGCACCGTAGCTGGTGTCTCTATCTGTA -
15060
        -L  S  M  P  L  V  Q  R  I  E  L  A  P  *  L  V  S  L  S  V
        - *  V  C  H  *  C  K  E  *  S  S  H  R  S  W  C  L  Y  L  *
        -  K  Y  A  I  S  A  K  N  R  A  R  T  V  A  G  V  S  I  C  S
15061 - GTACTATGACAAATAGACAGTTTCATCAGAAATTATTGAAGTCAATAGCCGCCACTAGAG -
15120
        -V  L  *  Q  I  D  S  F  I  R  N  Y  *  S  Q  *  P  P  L  E
        - Y  Y  D  K  *  T  V  S  S  E  I  I  E  V  N  S  R  H  *  R
        -  T  M  T  N  R  Q  F  H  Q  K  L  L  K  S  I  A  A  T  R  G
15121 - GAGCTACTGTGGTAATTGGAACAAGCAAGTTTTACGGTGGCTGGCATAATATGTTAAAAA -
```

15180
```
-E  L  L  W  *  L  E  Q  A  S  F  T  V  A  G  I  I  C  *  K
-  S  Y  C  G  N  W  N  K  Q  V  L  R  W  L  A  *  Y  V  K  N
-   A  T  V  V  I  G  T  S  K  F  Y  G  G  W  H  N  M  L  K  T
```

Figure 16-U

```
15181 - CTGTTTACAGTGATGTAGAAACTCCACACCTTATGGGTTGGGATTATCCAAAATGTGACA -
15240
      - L  F  T  V  M  *  K  L  H  T  L  W  V  G  I  I  Q  N  V  T
      -  C  L  Q  *  C  R  N  S  T  P  Y  G  L  G  L  S  K  M  *  Q
      -   V  Y  S  D  V  E  T  P  H  L  M  G  W  D  Y  P  K  C  D  R
15241 - GAGCCATGCCTAACATGCTTAGGATAATGGCCTCTCTTGTTCTTGCTCGCAAACATAACA -
15300
      - E  P  C  L  T  C  L  G  *  W  P  L  L  F  L  L  A  N  I  T
      -  S  H  A  *  H  A  *  D  N  G  L  S  C  S  C  S  Q  T  *  H
      -   A  M  P  N  M  L  R  I  M  A  S  L  V  L  A  R  K  H  N  T
15301 - CTTGCTGTAACTTATCACACCGTTTCTACAGGTTAGCTAACGAGTGTGCGCAAGTATTAA -
15360
      - L  A  V  T  Y  H  T  V  S  T  G  *  L  T  S  V  R  K  Y  *
      -  L  L  *  L  I  T  P  F  L  Q  V  S  *  R  V  C  A  S  I  K
      -   C  C  N  L  S  H  R  F  Y  R  L  A  N  E  C  A  Q  V  L  S
15361 - GTGAGATGGTCATGTGTGGCGGCTCACTATATGTTAAACCAGGTGGAACATCATCCGGTG -
15420
      - V  R  W  S  C  V  A  A  H  Y  M  L  N  Q  V  E  H  H  P  V
      -  *  D  G  H  V  W  R  L  T  I  C  *  T  R  W  N  I  I  R  *
      -   E  M  V  M  C  G  G  S  L  Y  V  K  P  G  G  T  S  S  G  D
15421 - ATGCTACAACTGCTTATGCTAATAGTGTCTTTAACATTTGTCAAGCTGTTACAGCCAATG -
15480
      - M  L  Q  L  L  M  L  I  V  S  L  T  F  V  K  L  L  Q  P  M
      -  C  Y  N  C  L  C  *  *  C  L  *  H  L  S  S  C  Y  S  Q  C
      -   A  T  T  A  Y  A  N  S  V  F  N  I  C  Q  A  V  T  A  N  V
15481 - TAAATGCACTTCTTTCAACTGATGGTAATAAGATAGCTGACAAGTATGTCCGCAATCTAC -
15540
      - *  M  H  F  F  Q  L  M  V  I  R  *  L  T  S  M  S  A  I  Y
      -  K  C  T  S  F  N  *  W  *  *  D  S  *  Q  V  C  P  Q  S  T
      -   N  A  L  L  S  T  D  G  N  K  I  A  D  K  Y  V  R  N  L  Q
15541 - AACACAGGCTCTATGAGTGTCTCTATAGAAATAGGGATGTTGATCATGAATTCGTGGATG -
15600
      - N  T  G  S  M  S  V  S  I  E  I  G  M  L  I  M  N  S  W  M
      -  T  Q  A  L  *  V  S  L  *  K  *  G  C  *  S  *  I  R  G  *
      -   H  R  L  Y  E  C  L  Y  R  N  R  D  V  D  H  E  F  V  D  E
15601 - AGTTTACGCTTACCTGCGTAAACATTTCTCCATGATGATTCTTTCTGATGATGCCGTTG -
15660
      - S  F  T  L  T  C  V  N  I  S  P  *  *  F  F  L  M  M  P  L
      -  V  L  R  L  P  A  *  T  F  L  H  D  D  S  F  *  *  C  R  C
      -   F  Y  A  Y  L  R  K  H  F  S  M  M  I  L  S  D  D  A  V  V
15661 - TGTGCTATAACAGTAACTATGCGGCTCAAGGTTTAGTAGCTAGCATTAAGAACTTTAAGG -
15720
      - C  A  I  T  V  T  M  R  L  K  V  *  *  L  A  L  R  T  L  R
      -  V  L  *  Q  *  L  C  G  S  R  F  S  S  *  H  *  E  L  *  G
      -   C  Y  N  S  N  Y  A  A  Q  G  L  V  A  S  I  K  N  F  K  A
15721 - CAGTTCTTTATTATCAAAATAATGTGTTCATGTCTGAGGCAAAATGTTGGACTGAGACTG -
15780
      - Q  F  F  I  I  K  I  M  C  S  C  L  R  Q  N  V  G  L  R  L
      -  S  S  L  L  S  K  *  C  V  H  V  *  G  K  M  L  D  *  D  *
      -   V  L  Y  Y  Q  N  N  V  F  M  S  E  A  K  C  W  T  E  T  D
15781 - ACCTTACTAAAGGACCTCACGAATTTTGCTCACAGCATACAATGCTAGTTAAACAAGGAG -
15840
      - T  L  L  K  D  L  T  N  F  A  H  S  I  Q  C  *  L  N  K  E
      -  P  Y  *  R  T  S  R  I  L  L  T  A  Y  N  A  S  *  T  R  R
      -   L  T  K  G  P  H  E  F  C  S  Q  H  T  M  L  V  K  Q  G  D
15841 - ATGATTACGTGTACCTGCCTTACCCAGATCCATCAAGAATATTAGGCGCAGGCTGTTTTG -
15900
```

```
      - M  I  T  C  T  C  L  T  Q  I  H  Q  E  Y  *  A  Q  A  V  L
      -  *  L  R  V  P  A  L  P  R  S  I  K  N  I  R  R  R  L  F  C
      -   D  Y  V  Y  L  P  Y  P  D  P  S  R  I  L  G  A  G  C  F  V
15901 - TCGATGATATTGTCAAAACAGATGGTACACTTATGATTGAAAGGTTCGTGTCACTGGCTA -
15960
      - S  M  I  L  S  K  Q  M  V  H  L  *  L  K  G  S  C  H  W  L
      -  R  *  Y  C  Q  N  R  W  Y  T  Y  D  *  K  V  R  V  T  G  Y
      -   D  D  I  V  K  T  D  G  T  L  M  I  E  R  F  V  S  L  A  I
```

Figure 16-V

```
15961 - TTGATGCTTACCCACTTACAAAACATCCTAATCAGGAGTATGCTGATGTCTTTCACTTGT -
16020
      -L  M  L  T  H  L  Q  N  I  L  I  R  S  M  L  M  S  F  T  C
      -  *  C  L  P  T  Y  K  T  S  *  S  G  V  C  *  C  L  S  L  V
      -   D  A  Y  P  L  T  K  H  P  N  Q  E  Y  A  D  V  F  H  L  Y
16021 - ATTTACAATACATTAGAAAGTTACATGATGAGCTTACTGGCCACATGTTGGACATGTATT -
16080
      -I  Y  N  T  L  E  S  Y  M  M  S  L  L  A  T  C  W  T  C  I
      -  F  T  I  H  *  K  V  T  *  *  A  Y  W  P  H  V  G  H  V  F
      -   L  Q  Y  I  R  K  L  H  D  E  L  T  G  H  M  L  D  M  Y  S
16081 - CCGTAATGCTAACTAATGATAACACCTCACGGTACTGGGAACCTGAGTTTTATGAGGCTA -
16140
      -P  *  C  *  L  M  I  T  P  H  G  T  G  N  L  S  F  M  R  L
      -  R  N  A  N  *  *  *  H  L  T  V  L  G  T  *  V  L  *  G  Y
      -   V  M  L  T  N  D  N  T  S  R  Y  W  E  P  E  F  Y  E  A  M
16141 - TGTACACACCACATACAGTCTTGCAGGCTGTAGGTGCTTGTGTATTGTGCAATTCACAGA -
16200
      -C  T  H  H  I  Q  S  C  R  L  *  V  L  V  Y  C  A  I  H  R
      -  V  H  T  T  Y  S  L  A  G  C  R  C  L  C  I  V  Q  F  T  D
      -   Y  T  P  H  T  V  L  Q  A  V  G  A  C  V  L  C  N  S  Q  T
16201 - CTTCACTTCGTTGCGGTGCCTGTATTAGGAGACCATTCCTATGTTGCAAGTGCTGCTATG -
16260
      -L  H  F  V  A  V  P  V  L  G  D  H  S  Y  V  A  S  A  A  M
      -  F  T  S  L  R  C  L  Y  *  E  T  I  P  M  L  Q  V  L  L  *
      -   S  L  R  C  G  A  C  I  R  R  P  F  L  C  C  K  C  C  Y  D
16261 - ACCATGTCATTTCAACATCACACAAATTAGTGTTGTCTGTTAATCCCTATGTTTGCAATG -
16320
      -T  M  S  F  Q  H  H  T  N  *  C  C  L  L  I  P  M  F  A  M
      -  P  C  H  F  N  I  T  Q  I  S  V  V  C  *  S  L  C  L  Q  C
      -   H  V  I  S  T  S  H  K  L  V  L  S  V  N  P  Y  V  C  N  A
16321 - CCCCAGGTTGTGATGTCACTGATGTGACACAACTGTATCTAGGAGGTATGAGCTATTATT -
16380
      -P  Q  V  V  M  S  L  M  *  H  N  C  I  *  E  V  *  A  I  I
      -  P  R  L  *  C  H  *  C  D  T  T  V  S  R  R  Y  E  L  L  L
      -   P  G  C  D  V  T  D  V  T  Q  L  Y  L  G  G  M  S  Y  Y  C
16381 - GCAAGTCACATAAGCCTCCCATTAGTTTCCATTATGTGCTAATGGTCAGGTTTTTGGTT -
16440
      -A  S  H  I  S  L  P  L  V  F  H  Y  V  L  M  V  R  F  L  V
      -  Q  V  T  *  A  S  H  *  F  S  I  M  C  *  W  S  G  F  W  F
      -   K  S  H  K  P  P  I  S  F  P  L  C  A  N  G  Q  V  F  G  L
16441 - TATACAAAAACACATGTGTAGGCAGTGACAATGTCACTGACTTCAATGCGATAGCAACAT -
16500
      -Y  T  K  T  H  V  *  A  V  T  M  S  L  T  S  M  R  *  Q  H
      -  I  Q  K  H  M  C  R  Q  *  Q  C  H  *  L  Q  C  D  S  N  M
      -   Y  K  N  T  C  V  G  S  D  N  V  T  D  F  N  A  I  A  T  C
16501 - GTGATTGGACTAATGCTGGCGATTACATACTTGCCAACACTTGTACTGAGAGACTCAAGC -
16560
      -V  I  G  L  M  L  A  I  T  Y  L  P  T  L  V  L  R  D  S  S
      -  *  L  D  *  C  W  R  L  H  T  C  Q  H  L  Y  *  E  T  Q  A
      -   D  W  T  N  A  G  D  Y  I  L  A  N  T  C  T  E  R  L  K  L
16561 - TTTTCGCAGCAGAAACGCTCAAAGCCACTGAGGAAACATTTAAGCTGTCATATGGTATTG -
16620
      -F  S  Q  Q  K  R  S  K  P  L  R  K  H  L  S  C  H  M  V  L
      -  F  R  S  R  N  A  Q  S  H  *  G  N  I  *  A  V  I  W  Y  C
      -   F  A  A  E  T  L  K  A  T  E  E  T  F  K  L  S  Y  G  I  A
16621 - CCACTGTACGCGAAGTACTCTCTGACAGAGAATTGCATCTTTCATGGGAGGTTGGAAAAC -
```

16680
-P L Y A K Y S L T E N C I F H G R L E N
- H C T R S T L * Q R I A S F M G G W K T
- T V R E V L S D R E L H L S W E V G K P

Figure 16-W

```
16681 - CTAGACCACCATTGAACAGAAACTATGTCTTTACTGGTTACCGTGTAACTAAAAATAGTA -
16740
      - L  D  H  H  *  T  E  T  M  S  L  L  V  T  V  *  L  K  I  V
      -  *  T  T  I  E  Q  K  L  C  L  Y  W  L  P  C  N  *  K  *  *
      -   R  P  P  L  N  R  N  Y  V  F  T  G  Y  R  V  T  K  N  S  K
16741 - AAGTACAGATTGGAGAGTACACCTTTGAAAAAGGTGACTATGGTGATGCTGTTGTGTACA -
16800
      - K  Y  R  L  E  S  T  P  L  K  K  V  T  M  V  M  L  L  C  T
      -  S  T  D  W  R  V  H  L  *  K  R  *  L  W  *  C  C  C  V  Q
      -   V  Q  I  G  E  Y  T  F  E  K  G  D  Y  G  D  A  V  V  Y  R
16801 - GAGGTACTACGACATACAAGTTGAATGTTGGTGATTACTTTGTGTTGACATCTCACACTG -
16860
      - E  V  L  R  H  T  S  *  M  L  V  I  T  L  C  *  H  L  T  L
      -  R  Y  Y  D  I  Q  V  E  C  W  *  L  L  C  V  D  I  S  H  C
      -   G  T  T  T  Y  K  L  N  V  G  D  Y  F  V  L  T  S  H  T  V
16861 - TAATGCCACTTAGTGCACCTACTCTAGTGCCACAAGAGCACTATGTGAGAATTACTGGCT -
16920
      - *  C  H  L  V  H  L  L  *  C  H  K  S  T  M  *  E  L  L  A
      -  N  A  T  *  C  T  Y  S  S  A  T  R  A  L  C  E  N  Y  W  L
      -   M  P  L  S  A  P  T  L  V  P  Q  E  H  Y  V  R  I  T  G  L
16921 - TGTACCCAACACTCAACATCTCAGATGAGTTTTCTAGCAATGTTGCAAATTATCAAAAGG -
16980
      - C  T  Q  H  S  T  S  Q  M  S  F  L  A  M  L  Q  I  I  K  R
      -  V  P  N  T  Q  H  L  R  *  V  F  *  Q  C  C  K  L  S  K  G
      -   Y  P  T  L  N  I  S  D  E  F  S  S  N  V  A  N  Y  Q  K  V
16981 - TCGGCATGCAAAAGTACTCTACACTCCAAGGACCACCTGGTACTGGTAAGAGTCATTTTG -
17040
      - S  A  C  K  S  T  L  H  S  K  D  H  L  V  L  V  R  V  I  L
      -  R  H  A  K  V  L  Y  T  P  R  T  T  W  Y  W  *  E  S  F  C
      -   G  M  Q  K  Y  S  T  L  Q  G  P  P  G  T  G  K  S  H  F  A
17041 - CCATCGGACTTGCTCTCTATTACCCATCTGCTCGCATAGTGTATACGGCATGCTCTCATG -
17100
      - P  S  D  L  L  S  I  T  H  L  L  A  *  C  I  R  H  A  L  M
      -  H  R  T  C  S  L  L  P  I  C  S  H  S  V  Y  G  M  L  S  C
      -   I  G  L  A  L  Y  Y  P  S  A  R  I  V  Y  T  A  C  S  H  A
17101 - CAGCTGTTGATGCCCTATGTGAAAAGGCATTAAAATATTTGCCCATAGATAAATGTAGTA -
17160
      - Q  L  L  M  P  Y  V  K  R  H  *  N  I  C  P  *  I  N  V  V
      -  S  C  *  C  P  M  *  K  G  I  K  I  F  A  H  R  *  M  *  *
      -   A  V  D  A  L  C  E  K  A  L  K  Y  L  P  I  D  K  C  S  R
17161 - GAATCATACCTGCGCGTGCGCGCGTAGAGTGTTTTGATAAATTCAAAGTGAATTCAACAC -
17220
      - E  S  Y  L  R  V  R  A  *  S  V  L  I  N  S  K  *  I  Q  H
      -  N  H  T  C  A  C  A  R  R  V  F  *  *  I  Q  S  E  F  N  T
      -   I  I  P  A  R  A  R  V  E  C  F  D  K  F  K  V  N  S  T  L
17221 - TAGAACAGTATGTTTTCTGCACTGTAAATGCATTGCCAGAAACAACTGCTGACATTGTAG -
17280
      - *  N  S  M  F  S  A  L  *  M  H  C  Q  K  Q  L  L  T  L  *
      -  R  T  V  C  F  L  H  C  K  C  I  A  R  N  N  C  *  H  C  S
      -   E  Q  Y  V  F  C  T  V  N  A  L  P  E  T  T  A  D  I  V  V
17281 - TCTTTGATGAAATCTCTATGGCTACTAATTATGACTTGAGTGTTGTCAATGCTAGACTTC -
17340
      - S  L  M  K  S  L  W  L  L  I  M  T  *  V  L  S  M  L  D  F
      -  L  *  *  N  L  Y  G  Y  *  L  *  L  E  C  C  Q  C  *  T  S
      -   F  D  E  I  S  M  A  T  N  Y  D  L  S  V  V  N  A  R  L  R
```

17341 - GTGCAAAACACTACGTCTATATTGGCGATCCTGCTCAATTACCAGCCCCCCGCACATTGC - 17400
- V Q N T T S I L A I L L N Y Q P P A H C
- C K T L R L Y W R S C S I T S P P H I A
- A K H Y V V Y I G D P A Q L P A P R T L L
17401 - TGACTAAAGGCACACTAGAACCAGAATATTTTAATTCAGTGTGCAGACTTATGAAAACAA - 17460
- * L K A H * N Q N I L L I Q C A D L * K Q
- D * R H T R T R I F * F S V Q T Y E N N
- T K G T L E P E Y F N S V C R L M K T I

Figure 16-X

```
17461 - TAGGTCCAGACATGTTCCTTGGAACTTGTCGCCGTTGTCCTGCTGAAATTGTTGACACTG -
17520
      - * V Q T C S L E L V A V V L L K L L T L
      - R S R H V P W N L S P L S C * N C * H C
      -  G P D M F L G T C R R C P A E I V D T V
17521 - TGAGTGCTTTAGTTTATGACAATAAGCTAAAAGCACACAAGGATAAGTCAGCTCAATGCT -
17580
      - * V L * F M T I S * K H T R I S Q L N A
      - E C F S L * Q * A K S T Q G * V S S M L
      -  S A L V Y D N K L K A H K D K S A Q C F
17581 - TCAAAATGTTCTACAAAGGTGTTATTACACATGATGTTTCATCTGCAATCAACAGACCTC -
17640
      - S K C S T K V L L H M M F H L Q S T D L
      - Q N V L Q R C Y Y T * C F I C N Q Q T S
      -  K M F Y K G V I T H D V S S A I N R P Q
17641 - AAATAGGCGTTGTAAGAGAATTTCTTACACGCAATCCTGCTTGGAGAAAAGCTGTTTTTA -
17700
      - K * A L * E N F L H A I L L G E K L F L
      - N R R C K R I S Y T Q S C L E K S C F Y
      -  I G V V R E F L T R N P A W R K A V F I
17701 - TCTCACCTTATAATTCACAGAACGCTGTAGCTTCAAAAATCTTAGGATTGCCTACGCAGA -
17760
      - S H L I I H R T L * L Q K S * D C L R R
      - L T L * F T E R C S F K N L R I A Y A D
      -  S P Y N S Q N A V A S K I L G L P T Q T
17761 - CTGTTGATTCATCACAGGGTTCTGAATATGACTATGTCATATTCACACAAACTACTGAAA -
17820
      - L L I H H R V L N M T M S Y S H K L L K
      - C * F I T G F * I * L C H I H T N Y * N
      -  V D S S Q G S E Y D Y V I F T Q T T E T
17821 - CAGCACACTCTTGTAATGTCAACCGCTTCAATGTGGCTATCACAAGGGCAAAAATTGGCA -
17880
      - Q H T L V M S T A S M W L S Q G Q K L A
      - S T L L * C Q P L Q C G Y H K G K N W H
      -  A H S C N V N R F N V A I T R A K I G I
17881 - TTTTGTGCATAATGTCTGATAGAGATCTTTATGACAAACTGCAATTTACAAGTCTAGAAA -
17940
      - F C A * C L I E I F M T N C N L Q V * K
      - F V H N V * * R S L * Q T A I Y K S R N
      -  L C I M S D R D L Y D K L Q F T S L E I
17941 - TACCACGTCGCAATGTGGCTACATTACAAGCAGAAAATGTAACTGGACTTTTTAAGGACT -
18000
      - Y H V A M W L H Y K Q K M * L D F L R T
      - T T S Q C G Y I T S R K C N W T F * G L
      -  P R R N V A T L Q A E N V T G L F K D C
18001 - GTAGTAAGATCATTACTGGTCTTCATCCTACACAGGCACCTACACACCTCAGCGTTGATA -
18060
      - V V R S L L V F I L H R H L H T S A L I
      - * * D H Y W S S S Y T G T Y T P Q R * Y
      -  S K I I T G L H P T Q A P T H L S V D I
18061 - TAAAATTCAAGACTGAAGGATTATGTGTTGACATACCAGGCATACCAAAGGACATGACCT -
18120
      - * N S R L K D Y V L T Y Q A Y Q R T * P
      - K I Q D * R I M C * H T R H T K G H D L
      -  K F K T E G L C V D I P G I P K D M T Y
18121 - ACCGTAGACTCATCTCTATGATGGGTTTCAAAATGAATTACCAAGTCAATGGTTACCCTA -
```

18180
- T V D S S L * W V S K * I T K S M V T L
- P * T H L Y D G F Q N E L P S Q W L P *
- R R L I S M M G F K M N Y Q V N G Y P N

Figure 16-Y

```
                    18181 -
ATATGTTTATCACCCGCGAAGAAGCTATTCGTCACGTTCGTGCGTGGATTGGCTTTGATG - 18240
       - I  C  L  S  P  A  K  K  L  F  V  T  F  V  R  G  L  A  L  M
       -  Y  V  Y  H  P  R  R  S  Y  S  S  R  S  C  V  D  W  L  *  C
       -   M  F  I  T  R  E  E  A  I  R  H  V  R  A  W  I  G  F  D  V
18241 - TAGAGGGCTGTCATGCAACTAGAGATGCTGTGGGTACTAACCTACCTCTCCAGCTAGGAT -
18300
       - *  R  A  V  M  Q  L  E  M  L  W  V  L  T  Y  L  S  S  *  D
       -  R  G  L  S  C  N  *  R  C  C  G  Y  *  P  T  S  P  A  R  I
       -   E  G  C  H  A  T  R  D  A  V  G  T  N  L  P  L  Q  L  G  F
18301 - TTTCTACAGGTGTTAACTTAGTAGCTGTACCGACTGGTTATGTTGACACTGAAAATAACA -
18360
       - F  L  Q  V  L  T  *  *  L  Y  R  L  V  M  L  T  L  K  I  T
       -  F  Y  R  C  *  L  S  S  C  T  D  W  L  C  *  H  *  K  *  H
       -   S  T  G  V  N  L  V  A  V  P  T  G  Y  V  D  T  E  N  N  T
18361 - CAGAATTCACCAGAGTTAATGCAAAACCTCCACCAGGTGACCAGTTTAAACATCTTATAC -
18420
       - Q  N  S  P  E  L  M  Q  N  L  H  Q  V  T  S  L  N  I  L  Y
       -  R  I  H  Q  S  *  C  K  T  S  T  R  *  P  V  *  T  S  Y  T
       -   E  F  T  R  V  N  A  K  P  P  P  G  D  Q  F  K  H  L  I  P
18421 - CACTCATGTATAAAGGCTTGCCCTGGAATGTAGTGCGTATTAAGATAGTACAAATGCTCA -
18480
       - H  S  C  I  K  A  C  P  G  M  *  C  V  L  R  *  Y  K  C  S
       -  T  H  V  *  R  L  A  L  E  C  S  A  Y  *  D  S  T  N  A  Q
       -   L  M  Y  K  G  L  P  W  N  V  V  R  I  K  I  V  Q  M  L  S
18481 - GTGATACACTGAAAGGATTGTCAGACAGAGTCGTGTTCGTCCTTTGGGCGCATGGCTTTG -
18540
       - V  I  H  *  K  D  C  Q  T  E  S  C  S  S  F  G  R  M  A  L
       -  *  Y  T  E  R  I  V  R  Q  S  R  V  R  P  L  G  A  W  L  *
       -   D  T  L  K  G  L  S  D  R  V  V  F  V  L  W  A  H  G  F  E
18541 - AGCTTACATCAATGAAGTACTTTGTCAAGATTGGACCTGAAAGAACGTGTTGTCTGTGTG -
18600
       - S  L  H  Q  *  S  T  L  S  R  L  D  L  K  E  R  V  V  C  V
       -  A  Y  I  N  E  V  L  C  Q  D  W  T  *  K  N  V  L  S  V  *
       -   L  T  S  M  K  Y  F  V  K  I  G  P  E  R  T  C  C  L  C  D
18601 - ACAAACGTGCAACTTGCTTTTCTACTTCATCAGATACTTATGCCTGCTGGAATCATTCTG -
18660
       - T  N  V  Q  L  A  F  L  L  H  Q  I  L  M  P  A  G  I  I  L
       -  Q  T  C  N  L  L  F  Y  F  I  R  Y  L  C  L  L  E  S  F  C
       -   K  R  A  T  C  F  S  T  S  S  D  T  Y  A  C  W  N  H  S  V
18661 - TGGGTTTTGACTATGTCTATAACCCATTTATGATTGATGTTCAGCAGTGGGGCTTTACGG -
18720
       - W  V  L  T  M  S  I  T  H  L  *  L  M  F  S  S  G  A  L  R
       -  G  F  *  L  C  L  *  P  I  Y  D  *  C  S  A  V  G  L  Y  G
       -   G  F  D  Y  V  Y  N  P  F  M  I  D  V  Q  Q  W  G  F  T  G
18721 - GTAACCTTCAGAGTAACCATGACCAACATTGCCAGGTACATGGAAATGCACATGTGGCTA -
18780
       - V  T  F  R  V  T  M  T  N  I  A  R  Y  M  E  M  H  M  W  L
       -  *  P  S  E  *  P  *  P  T  L  P  G  T  W  K  C  T  C  G  *
       -   N  L  Q  S  N  H  D  Q  H  C  Q  V  H  G  N  A  H  V  A  S
18781 - GTTGTGATGCTATCATGACTAGATGTTTAGCAGTCCATGAGTGCTTTGTTAAGCGCGTTG -
18840
       - V  V  M  L  S  *  L  D  V  *  Q  S  M  S  A  L  L  S  A  L
       -  L  *  C  Y  H  D  *  M  F  S  S  P  *  V  L  C  *  A  R  *
       -   C  D  A  I  M  T  R  C  L  A  V  H  E  C  F  V  K  R  V  D
18841 - ATTGGTCTGTTGAATACCCTATTATAGGAGATGAACTGAGGGTTAATTCTGCTTGCAGAA -
```

```
18900
      - I  G  L  L  N  T  L  L  *  E  M  N  *  G  L  I  L  L  A  E
      -  L  V  C  *  I  P  Y  Y  R  R  *  T  E  G  *  F  C  L  Q  K
      -   W  S  V  E  Y  P  I  I  G  D  E  L  R  V  N  S  A  C  R  K
18901 - AAGTACAACACATGGTTGTGAAGTCTGCATTGCTTGCTGATAAGTTTCCAGTTCTTCATG -
18960
      - K  Y  N  T  W  L  *  S  L  H  C  L  L  I  S  F  Q  F  F  M
      -  S  T  T  H  G  C  E  V  C  I  A  C  *  *  V  S  S  S  S  *
      -   V  Q  H  M  V  V  K  S  A  L  L  A  D  K  F  P  V  L  H  D
```

Figure 16-Z

```
18961 - ACATTGGAAATCCAAAGGCTATCAAGTGTGTGCCTCAGGCTGAAGTAGAATGGAAGTTCT -
19020
      - T  L  E  I  Q  R  L  S  S  V  C  L  R  L  K  *  N  G  S  S
      -  H  W  K  S  K  G  Y  Q  V  C  A  S  G  *  S  R  M  E  V  L
      -   I  G  N  P  K  A  I  K  C  V  P  Q  A  E  V  E  W  K  F  Y
19021 - ACGATGCTCAGCCATGTAGTGACAAAGCTTACAAAATAGAGGAACTCTTCTATTCTTATG -
19080
      - T  M  L  S  H  V  V  T  K  L  T  K  *  R  N  S  S  I  L  M
      -  R  C  S  A  M  *  *  Q  S  L  Q  N  R  G  T  L  L  F  L  C
      -   D  A  Q  P  C  S  D  K  A  Y  K  I  E  E  L  F  Y  S  Y  A
19081 - CTACACATCACGATAAATTCACTGATGGTGTTTGTTTGTTTTGGAATTGTAACGTTGATC -
19140
      - L  H  I  T  I  N  S  L  M  V  F  V  C  F  G  I  V  T  L  I
      -  Y  T  S  R  *  I  H  *  W  C  L  F  V  L  E  L  *  R  *  S
      -   T  H  H  D  K  F  T  D  G  V  C  L  F  W  N  C  N  V  D  R
19141 - GTTACCCAGCCAATGCAATTGTGTGTAGGTTTGACACAAGAGTCTTGTCAAACTTGAACT -
19200
      - V  T  Q  P  M  Q  L  C  V  G  L  T  Q  E  S  C  Q  T  *  T
      -  L  P  S  Q  C  N  C  V  *  V  *  H  K  S  L  V  K  L  E  L
      -   Y  P  A  N  A  I  V  C  R  F  D  T  R  V  L  S  N  L  N  L
19201 - TACCAGGCTGTGATGGTGGTAGTTTGTATGTGAATAAGCATGCATTCCACACTCCAGCTT -
19260
      - Y  Q  A  V  M  V  V  V  C  M  *  I  S  M  H  S  T  L  Q  L
      -  T  R  L  *  W  W  *  F  V  C  E  *  A  C  I  P  H  S  S  F
      -   P  G  C  D  G  G  S  L  Y  V  N  K  H  A  F  H  T  P  A  F
19261 - TCGATAAAAGTGCATTTACTAATTTAAAGCAATTGCCTTTCTTTTACTATTCTGATAGTC -
19320
      - S  I  K  V  H  L  L  I  *  S  N  C  L  S  F  T  I  L  I  V
      -  R  *  K  C  I  Y  *  F  K  A  I  A  F  L  L  L  F  *  *  S
      -   D  K  S  A  F  T  N  L  K  Q  L  P  F  F  Y  Y  S  D  S  P
19321 - CTTGTGAGTCTCATGGCAAACAAGTAGTGTCGGATATTGATTATGTTCCACTCAAATCTG -
19380
      - L  V  S  L  M  A  N  K  *  C  R  I  L  I  M  F  H  S  N  L
      -  L  *  V  S  W  Q  T  S  S  V  G  Y  *  L  C  S  T  Q  I  C
      -   C  E  S  H  G  K  Q  V  V  S  D  I  D  Y  V  P  L  K  S  A
19381 - CTACGTGTATTACACGATGCAATTTAGGTGGTGCTGTTTGCAGACACCATGCAAATGAGT -
19440
      - L  R  V  L  H  D  A  I  *  V  V  L  F  A  D  T  M  Q  M  S
      -  Y  V  Y  Y  T  M  Q  F  R  W  C  C  L  Q  T  P  C  K  *  V
      -   T  C  I  T  R  C  N  L  G  G  A  V  C  R  H  H  A  N  E  Y
19441 - ACCGACAGTACTTGGATGCATATAATATGATGATTTCTGCTGGATTTAGCCTATGGATTT -
19500
      - T  D  S  T  W  M  H  I  I  *  *  F  L  L  D  L  A  Y  G  F
      -  P  T  V  L  G  C  I  *  Y  D  D  F  C  W  I  *  P  M  D  L
      -   R  Q  Y  L  D  A  Y  N  M  M  I  S  A  G  F  S  L  W  I  Y
19501 - ACAAACAATTTGATACTTATAACCTGTGGAATACATTTACCAGGTTACAGAGTTTAGAAA -
19560
      - T  N  N  L  I  L  I  T  C  G  I  H  L  P  G  Y  R  V  *  K
      -  Q  T  I  *  Y  L  *  P  V  E  Y  I  Y  Q  V  T  E  F  R  K
      -   K  Q  F  D  T  Y  N  L  W  N  T  F  T  R  L  Q  S  L  E  N
19561 - ATGTGGCTTATAATGTTGTTAATAAAGGACACTTTGATGGACACGCCGGCGAAGCACCTG -
19620
      - M  W  L  I  M  L  L  I  K  D  T  L  M  D  T  P  A  K  H  L
      -  C  G  L  *  C  C  *  *  R  T  L  *  W  T  R  R  R  S  T  C
      -   V  A  Y  N  V  V  N  K  G  H  F  D  G  H  A  G  E  A  P  V
19621 - TTTCCATCATTAATAATGCTGTTTACACAAAGGTAGATGGTATTGATGTGGAGATCTTTG -
```

19680
- F P S L I M L F T Q R * M V L M W R S L
- F H H * * C C L H K G R W Y * C G D L *
- S I I N N A V Y T K V D G I D V E I F E

Figure 16-AA

19681 - AAAATAAGACAACACTTCCTGTTAATGTTGCATTTGAGCTTTGGGCTAAGCGTAACATTA - 19740
- K I R Q H F L L M L H L S F G L S V T L
- K * D N T S C * C C I * A L G * A * H *
- N K T T L P V N V A F E L W A K R N I K
19741 - AACCAGTGCCAGAGATTAAGATACTCAATAATTTGGGTGTTGATATCGCTGCTAATACTG - 19800
- N Q C Q R L R Y S I I W V L I S L L I L
- T S A R D * D T Q * F G C * Y R C * Y C
- P V P E I K I L N N L G V D I A A N T V
19801 - TAATCTGGGACTACAAAAGAGAAGCCCCAGCACATGTATCTACAATAGGTGTCTGCACAA - 19860
- * S G T T K E K P Q H M Y L Q * V S A Q
- N L G L Q K R S P S T C I Y N R C L H N
- I W D Y K R E A P A H V S T I G V C T M
19861 - TGACTGACATTGCCAAGAAACCTACTGAGAGTGCTTGTTCTTCACTTACTGTCTTGTTTG - 19920
- * L T L P R N L L R V L V L H L L S C L
- D * H C Q E T Y * E C L F F T Y C L V *
- T D I A K K P T E S A C S S L T V L F D
19921 - ATGGTAGAGTGGAAGGACAGGTAGACCTTTTTAGAAACGCCCGTAATGGTGTTTTAATAA - 19980
- M V E W K D R * T F L E T P V M V F * *
- W * S G R T G R P F * K R P * W C F N N
- G R V E G Q V D L F R N A R N G V L I T
19981 - CAGAAGGTTCAGTCAAAGGTCTAACACCTTCAAAGGGACCAGCACAAGCTAGCGTCAATG - 20040
- Q K V Q S K V * H L Q R D Q H K L A S M
- R R F S Q R S N T F K G T S T S * R Q W
- E G S V K G L T P S K G P A Q A S V N G
20041 - GAGTCACATTAATTGGAGAATCAGTAAAAACACAGTTTAACTACTTTAAGAAAGTAGACG - 20100
- E S H * L E N Q * K H S L T T L R K * T
- S H I N W R I S K N T V * L L * E S R R
- V T L I G E S V K T Q F N Y F K K V D G
20101 - GCATTATTCAACAGTTGCCTGAAACCTACTTTACTCAGAGCAGAGACTTAGAGGATTTTA - 20160
- A L F N S C L K P T L L R A E T * R I L
- H Y S T V A * N L L Y S E Q R L R G F *
- I I Q Q L P E T Y F T Q S R D L E D F K
20161 - AGCCCAGATCACAAATGGAAACTGACTTTCTCGAGCTCGCTATGGATGAATTCATACAGC - 20220
- S P D H K W K L T F S S S L W M N S Y S
- A Q I T N G N * L S R A R Y G * I H T A
- P R S Q M E T D F L E L A M D E F I Q R
20221 - GATATAAGCTCGAGGGCTATGCCTTCGAACACATCGTTTATGGAGATTTCAGTCATGGAC -

```
20280
      -D  I  S  S  R  A  M  P  S  N  T  S  F  M  E  I  S  V  M  D
      - I  *  A  R  G  L  C  L  R  T  H  R  L  W  R  F  Q  S  W  T
      -  Y  K  L  E  G  Y  A  F  E  H  I  V  Y  G  D  F  S  H  G  Q
20281 - AACTTGGCGGTCTTCATTTAATGATAGGCTTAGCCAAGCGCTCACAAGATTCACCACTTA -
20340
      -N  L  A  V  F  I  *  *  *  A  *  P  S  A  H  K  I  H  H  L
      - T  W  R  S  S  F  N  D  R  L  S  Q  A  L  T  R  F  T  T  *
      -  L  G  G  L  H  L  M  I  G  L  A  K  R  S  Q  D  S  P  L  K
20341 - AATTAGAGGATTTTATCCCTATGGACAGCACAGTGAAAAATTACTTCATAACAGATGCGC -
20400
      -N  *  R  I  L  S  L  W  T  A  Q  *  K  I  T  S  *  Q  M  R
      - I  R  G  F  Y  P  Y  G  Q  H  S  E  K  L  L  H  N  R  C  A
      -  L  E  D  F  I  P  M  D  S  T  V  K  N  Y  F  I  T  D  A  Q
20401 - AAACAGGTTCATCAAAATGTGTGTGTTCTGTGATTGATCTTTTACTTGATGACTTTGTCG -
20460
      -K  Q  V  H  Q  N  V  C  V  L  *  L  I  F  Y  L  M  T  L  S
      - N  R  F  I  K  M  C  V  F  C  D  *  S  F  T  *  *  L  C  R
      -  T  G  S  S  K  C  V  C  S  V  I  D  L  L  L  D  D  F  V  E
```

Figure 16-BB

```
20461 - AGATAATAAAGTCACAAGATTTGTCAGTGATTTCAAAAGTGGTCAAGGTTACAATTGACT -
20520
      - R  *  *  S  H  K  I  C  Q  *  F  Q  K  W  S  R  L  Q  L  T
      -  D  N  K  V  T  R  F  V  S  D  F  K  S  G  Q  G  Y  N  *  L
      -   I  I  K  S  Q  D  L  S  V  I  S  K  V  V  K  V  T  I  D  Y
20521 - ATGCTGAAATTTCATTCATGCTTTGGTGTAAGGATGGACATGTTGAAACCTTCTACCCAA -
20580
      - M  L  K  F  H  S  C  F  G  V  R  M  D  M  L  K  P  S  T  Q
      -  C  *  N  F  I  H  A  L  V  *  G  W  T  C  *  N  L  L  P  K
      -   A  E  I  S  F  M  L  W  C  K  D  G  H  V  E  T  F  Y  P  K
20581 - AACTACAAGCAAGTCAAGCGTGGCAACCAGGTGTTGCGATGCCTAACTTGTACAAGATGC -
20640
      - N  Y  K  Q  V  K  R  G  N  Q  V  L  R  C  L  T  C  T  R  C
      -  T  T  S  K  S  S  V  A  T  R  C  C  D  A  *  L  V  Q  D  A
      -   L  Q  A  S  Q  A  W  Q  P  G  V  A  M  P  N  L  Y  K  M  Q
20641 - AAAGAATGCTTCTTGAAAAGTGTGACCTTCAGAATTATGGTGAAAATGCTGTTATACCAA -
20700
      - K  E  C  F  L  K  S  V  T  F  R  I  M  V  K  M  L  L  Y  Q
      -  K  N  A  S  *  K  V  *  P  S  E  L  W  *  K  C  C  Y  T  K
      -   R  M  L  L  E  K  C  D  L  Q  N  Y  G  E  N  A  V  I  P  K
20701 - AAGGAATAATGATGAATGTCGCAAAGTATACTCAACTGTGTCAATACTTAAATACACTTA -
20760
      - K  E  *  *  *  M  S  Q  S  I  L  N  C  V  N  T  *  I  H  L
      -  R  N  N  D  E  C  R  K  V  Y  S  T  V  S  I  L  K  Y  T  Y
      -   G  I  M  M  N  V  A  K  Y  T  Q  L  C  Q  Y  L  N  T  L  T
20761 - CTTTAGCTGTACCCTACAACATGAGAGTTATTCACTTTGGTGCTGGCTCTGATAAAGGAG -
20820
      - L  *  L  Y  P  T  T  *  E  L  F  T  L  V  L  A  L  I  K  E
      -  F  S  C  T  L  Q  H  E  S  Y  S  L  W  C  W  L  *  *  R  S
      -   L  A  V  P  Y  N  M  R  V  I  H  F  G  A  G  S  D  K  G  V
20821 - TTGCACCAGGTACAGCTGTGCTCAGACAATGGTTGCCAACTGGCACACTACTTGTCGATT -
20880
      - L  H  Q  V  Q  L  C  S  D  N  G  C  Q  L  A  H  Y  L  S  I
      -  C  T  R  Y  S  C  A  Q  T  M  V  A  N  W  H  T  T  C  R  F
      -   A  P  G  T  A  V  L  R  Q  W  L  P  T  G  T  L  L  V  D  S
20881 - CAGATCTTAATGACTTCGTCTCCGACGCAGATTCTACTTTAATTGGAGACTGTGCAACAG -
20940
      - Q  I  L  M  T  S  S  P  T  Q  I  L  L  *  L  E  T  V  Q  Q
      -  R  S  *  *  L  R  L  R  R  R  F  Y  F  N  W  R  L  C  N  S
      -   D  L  N  D  F  V  S  D  A  D  S  T  L  I  G  D  C  A  T  V
20941 - TACATACGGCTAATAAATGGGACCTTATTATTAGCGATATGTATGACCCTAGGACCAAAC -
21000
      - Y  I  R  L  I  N  G  T  L  L  L  A  I  C  M  T  L  G  P  N
      -  T  Y  G  *  *  M  G  P  Y  Y  *  R  Y  V  *  P  *  D  Q  T
      -   H  T  A  N  K  W  D  L  I  I  S  D  M  Y  D  P  R  T  K  H
21001 - ATGTGACAAAAGAGAATGACTCTAAAGAAGGGTTTTTCACTTATCTGTGTGGATTTATAA -
21060
      - M  *  Q  K  R  M  T  L  K  K  G  F  S  L  I  C  V  D  L  *
      -  C  D  K  R  E  *  L  *  R  R  V  F  H  L  S  V  W  I  Y  K
      -   V  T  K  E  N  D  S  K  E  G  F  F  T  Y  L  C  G  F  I  K
21061 - AGCAAAAACTAGCCCTGGGTGGTTCTATAGCTGTAAAGATAACAGAGCATTCTTGGAATG -
21120
      - S  K  N  *  P  W  V  V  L  *  L  *  R  *  Q  S  I  L  G  M
      -  A  K  T  S  P  G  W  F  Y  S  C  K  D  N  R  A  F  L  E  C
      -   Q  K  L  A  L  G  G  S  I  A  V  K  I  T  E  H  S  W  N  A
21121 - CTGACCTTTACAAGCTTATGGGCCATTTCTCATGGTGGACAGCTTTTGTTACAAATGTAA -
```

```
21180
      - L  T  F  T  S  L  W  A  I  S  H  G  G  Q  L  L  L  Q  M  *
      -  *  P  L  Q  A  Y  G  P  F  L  M  V  D  S  F  C  Y  K  C  K
      -   D  L  Y  K  L  M  G  H  F  S  W  W  T  A  F  V  T  N  V  N
```

Figure 16-CC

```
21181 - ATGCATCATCATCGGAAGCATTTTTAATTGGGGCTAACTATCTTGGCAAGCCGAAGGAAC -
21240
      - M  H  H  H  R  K  H  F  *  L  G  L  T  I  L  A  S  R  R  N
      -  C  I  I  I  G  S  I  F  N  W  G  *  L  S  W  Q  A  E  G  T
      -   A  S  S  S  E  A  F  L  I  G  A  N  Y  L  G  K  P  K  E  Q
21241 - AAATTGATGGCTATACCATGCATGCTAACTACATTTTCTGGAGGAACACAAATCCTATCC -
21300
      - K  L  M  A  I  P  C  M  L  T  T  F  S  G  G  T  Q  I  L  S
      -  N  *  W  L  Y  H  A  C  *  L  H  F  L  E  E  H  K  S  Y  P
      -   I  D  G  Y  T  M  H  A  N  Y  I  F  W  R  N  T  N  P  I  Q
21301 - AGTTGTCTTCCTATTCACTCTTTGACATGAGCAAATTTCCTCTTAAATTAAGAGGAACTG -
21360
      - S  C  L  P  I  H  S  L  T  *  A  N  F  L  L  N  *  E  E  L
      -  V  V  F  L  F  T  L  *  H  E  Q  I  S  S  *  I  K  R  N  C
      -   L  S  S  Y  S  L  F  D  M  S  K  F  P  L  K  L  R  G  T  A
21361 - CTGTAATGTCTCTTAAGGAGAATCAAATCAATGATATGATTTATTCTCTTCTGGAAAAAG -
21420
      - L  *  C  L  L  R  R  I  K  S  M  I  *  F  I  L  F  W  K  K
      -  C  N  V  S  *  G  E  S  N  Q  *  Y  D  L  F  S  S  G  K  R
      -   V  M  S  L  K  E  N  Q  I  N  D  M  I  Y  S  L  L  E  K  G
21421 - GTAGGCTTATCATTAGAGAAAACAACAGAGTTGTGGTTTCAAGTGATATTCTTGTTAACA -
21480
      - V  G  L  S  L  E  K  T  T  E  L  W  F  Q  V  I  F  L  L  T
      -  *  A  Y  H  *  R  K  Q  Q  S  C  G  F  K  *  Y  S  C  *  Q
      -   R  L  I  I  R  E  N  N  R  V  V  V  S  S  D  I  L  V  N  N
21481 - ACTAAACGAACATGTTTATTTTCTTATTATTTCTTACTCTCACTAGTGGTAGTGACCTTG -
21540
      - T  K  R  T  C  L  F  S  Y  Y  F  L  L  S  L  V  V  V  T  L
      -  L  N  E  H  V  Y  F  L  I  I  S  Y  S  H  *  W  *  *  P  *
      -   *  T  N  M  F  I  F  L  L  F  L  T  L  T  S  G  S  D  L  D
21541 - ACCGGTGCACCACTTTTGATGATGTTCAAGCTCCTAATTACACTCAACATACTTCATCTA -
21600
      - T  G  A  P  L  L  M  M  F  K  L  L  I  T  L  N  I  L  H  L
      -  P  V  H  H  F  *  *  C  S  S  S  *  L  H  S  T  Y  F  I  Y
      -   R  C  T  T  F  D  D  V  Q  A  P  N  Y  T  Q  H  T  S  S  M
21601 - TGAGGGGGGTTTACTATCCTGATGAAATTTTTAGATCAGACACTCTTTATTTAACTCAGG -
21660
      - *  G  G  F  T  I  L  M  K  F  L  D  Q  T  L  F  I  *  L  R
      -  E  G  G  L  L  S  *  *  N  F  *  I  R  H  S  L  F  N  S  G
      -   R  G  V  Y  Y  P  D  E  I  F  R  S  D  T  L  Y  L  T  Q  D
21661 - ATTTATTTCTTCCATTTTATTCTAATGTTACAGGGTTTCATACTATTAATCATACGTTTG -
21720
      - I  Y  F  F  H  F  I  L  M  L  Q  G  F  I  L  L  I  I  R  L
      -  F  I  S  S  I  L  F  *  C  Y  R  V  S  Y  Y  *  S  Y  V  W
      -   L  F  L  P  F  Y  S  N  V  T  G  F  H  T  I  N  H  T  F  G
21721 - GCAACCCTGTCATACCTTTTAAGGATGGTATTTATTTTGCTGCCACAGAGAAATCAAATG -
21780
      - A  T  L  S  Y  L  L  R  M  V  F  I  L  L  P  Q  R  N  Q  M
      -  Q  P  C  H  T  F  *  G  W  Y  L  F  C  C  H  R  E  I  K  C
      -   N  P  V  I  P  F  K  D  G  I  Y  F  A  A  T  E  K  S  N  V
21781 - TTGTCCGTGGTTGGGTTTTTGGTTCTACCATGAACAACAAGTCACAGTCGGTGATTATTA -
21840
      - L  S  V  V  G  F  L  V  L  P  *  T  T  S  H  S  R  *  L  L
      -  C  P  W  L  G  F  W  F  Y  H  E  Q  Q  V  T  V  G  D  Y  Y
      -   V  R  G  W  V  F  G  S  T  M  N  N  K  S  Q  S  V  I  I  I
```

```
21841 - TTAACAATTCTACTAATGTTGTTATACGAGCATGTAACTTTGAATTGTGTGACAACCCTT -
21900
      - L  T  I  L  L  M  L  L  Y  E  H  V  T  L  N  C  V  T  T  L
      -  *  Q  F  Y  *  C  C  Y  T  S  M  *  L  *  I  V  *  Q  P  F
      -   N  N  S  T  N  V  V  I  R  A  C  N  F  E  L  C  D  N  P  F
```

Figure 16-DD

```
21901 - TCTTTGCTGTTTCTAAACCCATGGGTACACAGACACATACTATGATATTCGATAATGCAT -
21960
      -S  L  L  F  L  N  P  W  V  H  R  H  I  L  *  Y  S  I  M  H
      - L  C  C  F  *  T  H  G  Y  T  D  T  Y  Y  D  I  R  *  C  I
      -  F  A  V  S  K  P  M  G  T  Q  T  H  T  M  I  F  D  N  A  F
21961 - TTAATTGCACTTTCGAGTACATATCTGATGCCTTTTCGCTTGATGTTTCAGAAAAGTCAG -
22020
      -L  I  A  L  S  S  T  Y  L  M  P  F  R  L  M  F  Q  K  S  Q
      - *  L  H  F  R  V  H  I  *  C  L  F  A  *  C  F  R  K  V  R
      -  N  C  T  F  E  Y  I  S  D  A  F  S  L  D  V  S  E  K  S  G
22021 - GTAATTTTAAACACTTACGAGAGTTTGTGTTTAAAAATAAAGATGGGTTTCTCTATGTTT -
22080
      -V  I  L  N  T  Y  E  S  L  C  L  K  I  K  M  G  F  S  M  F
      - *  F  *  T  L  T  R  V  C  V  *  K  *  R  W  V  S  L  C  L
      -  N  F  K  H  L  R  E  F  V  F  K  N  K  D  G  F  L  Y  V  Y
22081 - ATAAGGGCTATCAACCTATAGATGTAGTTCGTGATCTACCTTCTGGTTTTAACACTTTGA -
22140
      -I  R  A  I  N  L  *  M  *  F  V  I  Y  L  L  V  L  T  L  *
      - *  G  L  S  T  Y  R  C  S  S  *  S  T  F  W  F  *  H  F  E
      -  K  G  Y  Q  P  I  D  V  V  R  D  L  P  S  G  F  N  T  L  K
22141 - AACCTATTTTTAAGTTGCCTCTTGGTATTAACATTACAAATTTTAGAGCCATTCTTACAG -
22200
      -N  L  F  L  S  C  L  L  V  L  T  L  Q  I  L  E  P  F  L  Q
      - T  Y  F  *  V  A  S  W  Y  *  H  Y  K  F  *  S  H  S  Y  S
      -  P  I  F  K  L  P  L  G  I  N  I  T  N  F  R  A  I  L  T  A
22201 - CCTTTTCACCTGCTCAAGACATTTGGGGCACGTCAGCTGCAGCCTATTTTGTTGGCTATT -
22260
      -P  F  H  L  L  K  T  F  G  A  R  Q  L  Q  P  I  L  L  A  I
      - L  F  T  C  S  R  H  L  G  H  V  S  C  S  L  F  C  W  L  F
      -  F  S  P  A  Q  D  I  W  G  T  S  A  A  A  Y  F  V  G  Y  L
22261 - TAAAGCCAACTACATTTATGCTCAAGTATGATGAAAATGGTACAATCACAGATGCTGTTG -
22320
      -*  S  Q  L  H  L  C  S  S  M  M  K  M  V  Q  S  Q  M  L  L
      - K  A  N  Y  I  Y  A  Q  V  *  *  K  W  Y  N  H  R  C  C  *
      -  K  P  T  T  F  M  L  K  Y  D  E  N  G  T  I  T  D  A  V  D
22321 - ATTGTTCTCAAAATCCACTTGCTGAACTCAAATGCTCTGTTAAGAGCTTTGAGATTGACA -
22380
      -I  V  L  K  I  H  L  L  N  S  N  A  L  L  R  A  L  R  L  T
      - L  F  S  K  S  T  C  *  T  Q  M  L  C  *  E  L  *  D  *  Q
      -  C  S  Q  N  P  L  A  E  L  K  C  S  V  K  S  F  E  I  D  K
22381 - AAGGAATTTACCAGACCTCTAATTTCAGGGTTGTTCCCTCAGGAGATGTTGTGAGATTCC -
22440
      -K  E  F  T  R  P  L  I  S  G  L  F  P  Q  E  M  L  *  D  S
      - R  N  L  P  D  L  *  F  Q  G  C  S  L  R  R  C  C  E  I  P
      -  G  I  Y  Q  T  S  N  F  R  V  V  P  S  G  D  V  V  R  F  P
22441 - CTAATATTACAAACTTGTGTCCTTTTGGAGAGGTTTTTAATGCTACTAAATTCCCTTCTG -
22500
      -L  I  L  Q  T  C  V  L  L  E  R  F  L  M  L  L  N  S  L  L
      - *  Y  Y  K  L  V  S  F  W  R  G  F  *  C  Y  *  I  P  F  C
      -  N  I  T  N  L  C  P  F  G  E  V  F  N  A  T  K  F  P  S  V
22501 - TCTATGCATGGGAGAGAAAAAAAATTTCTAATTGTGTTGCTGATTACTCTGTGCTCTACA -
22560
      -S  M  H  G  R  E  K  K  F  L  I  V  L  L  I  T  L  C  S  T
      - L  C  M  G  E  K  K  N  F  *  L  C  C  *  L  L  C  A  L  Q
      -  Y  A  W  E  R  K  K  I  S  N  C  V  A  D  Y  S  V  L  Y  N
```

```
22561 - ACTCAACATTTTTTTCAACCTTTAAGTGCTATGGCGTTTCTGCCACTAAGTTGAATGATC -
22620
      - T  Q  H  F  F  Q  P  L  S  A  M  A  F  L  P  L  S  *  M  I
      -  L  N  I  F  F  N  L  *  V  L  W  R  F  C  H  *  V  E  *  S
      -   S  T  F  F  S  T  F  K  C  Y  G  V  S  A  T  K  L  N  D  L
```

Figure 16-EE

```
22621 - TTTGCTTCTCCAATGTCTATGCAGATTCTTTTGTAGTCAAGGGAGATGATGTAAGACAAA -
22680
      - F  A  S  P  M  S  M  Q  I  L  L  *  S  R  E  M  M  *  D  K
      -  L  L  L  Q  C  L  C  R  F  F  C  S  Q  G  R  *  C  K  T  N
      -   C  F  S  N  V  Y  A  D  S  F  V  V  K  G  D  D  V  R  Q  I
22681 - TAGCGCCAGGACAAACTGGTGTTATTGCTGATTATAATTATAAATTGCCAGATGATTTCA -
22740
      - *  R  Q  D  K  L  V  L  L  L  I  I  I  I  N  C  Q  M  I  S
      -  S  A  R  T  N  W  C  Y  C  *  L  *  L  *  I  A  R  *  F  H
      -   A  P  G  Q  T  G  V  I  A  D  Y  N  Y  K  L  P  D  D  F  M
22741 - TGGGTTGTGTCCTTGCTTGGAATACTAGGAACATTGATGCTACTTCAACTGGTAATTATA -
22800
      - W  V  V  S  L  L  G  I  L  G  T  L  M  L  L  Q  L  V  I  I
      -  G  L  C  P  C  L  E  Y  *  E  H  *  C  Y  F  N  W  *  L  *
      -   G  C  V  L  A  W  N  T  R  N  I  D  A  T  S  T  G  N  Y  N
22801 - ATTATAAATATAGGTATCTTAGACATGGCAAGCTTAGGCCCTTTGAGAGAGACATATCTA -
22860
      - I  I  N  I  G  I  L  D  M  A  S  L  G  P  L  R  E  T  Y  L
      -  L  *  I  *  V  S  *  T  W  Q  A  *  A  L  *  E  R  H  I  *
      -   Y  K  Y  R  Y  L  R  H  G  K  L  R  P  F  E  R  D  I  S  N
22861 - ATGTGCCTTTCTCCCCTGATGGCAAACCTTGCACCCCACCTGCTCTTAATTGTTATTGGC -
22920
      - M  C  L  S  P  L  M  A  N  L  A  P  H  L  L  L  I  V  I  G
      -  C  A  F  L  P  *  W  Q  T  L  H  P  T  C  S  *  L  L  L  A
      -   V  P  F  S  P  D  G  K  P  C  T  P  P  A  L  N  C  Y  W  P
22921 - CATTAAATGATTATGGTTTTTACACCACTACTGGCATTGGCTACCAACCTTACAGAGTTG -
22980
      - H  *  M  I  M  V  F  T  P  L  L  A  L  A  T  N  L  T  E  L
      -  I  K  *  L  W  F  L  H  H  Y  W  H  W  L  P  T  L  Q  S  C
      -   L  N  D  Y  G  F  Y  T  T  T  G  I  G  Y  Q  P  Y  R  V  V
22981 - TAGTACTTTCTTTTGAACTTTTAAATGCACCGGCCACGGTTTGTGGACCAAAATTATCCA -
23040
      - *  Y  F  L  L  N  F  *  M  H  R  P  R  F  V  D  Q  N  Y  P
      -  S  T  F  F  *  T  F  K  C  T  G  H  G  L  W  T  K  I  I  H
      -   V  L  S  F  E  L  L  N  A  P  A  T  V  C  G  P  K  L  S  T
23041 - CTGACCTTATTAAGAACCAGTGTGTCAATTTTAATTTTAATGGACTCACTGGTACTGGTG -
23100
      - L  T  L  L  R  T  S  V  S  I  L  I  L  M  D  S  L  V  L  V
      -  *  P  Y  *  E  P  V  C  Q  F  *  F  *  W  T  H  W  Y  W  C
      -   D  L  I  K  N  Q  C  V  N  F  N  F  N  G  L  T  G  T  G  V
23101 - TGTTAACTCCTTCTTCAAAGAGATTTCAACCATTTCAACAATTTGGCCGTGATGTTTCTG -
23160
      - C  *  L  L  L  Q  R  D  F  N  H  F  N  N  L  A  V  M  F  L
      -  V  N  S  F  F  K  E  I  S  T  I  S  T  I  W  P  *  C  F  *
      -   L  T  P  S  S  K  R  F  Q  P  F  Q  Q  F  G  R  D  V  S  D
23161 - ATTTCACTGATTCCGTTCGAGATCCTAAAACATCTGAAATATTAGACATTTCACCTTGCT -
23220
      - I  S  L  I  P  F  E  I  L  K  H  L  K  Y  *  T  F  H  L  A
      -  F  H  *  F  R  S  R  S  *  N  I  *  N  I  R  H  F  T  L  L
      -   F  T  D  S  V  R  D  P  K  T  S  E  I  L  D  I  S  P  C  S
23221 - CTTTTGGGGGTGTAAGTGTAATTACACCTGGAACAAATGCTTCATCTGAAGTTGCTGTTC -
23280
      - L  L  G  V  *  V  *  L  H  L  E  Q  M  L  H  L  K  L  L  F
      -  F  W  G  C  K  C  N  Y  T  W  N  K  C  F  I  *  S  C  C  S
      -   F  G  G  V  S  V  I  T  P  G  T  N  A  S  S  E  V  A  V  L
23281 - TATATCAAGATGTTAACTGCACTGATGTTTCTACAGCAATTCATGCAGATCAACTCACAC -
```

23340
- Y I K M L T A L M F L Q Q F M Q I N S H
- I S R C * L H * C F Y S N S C R S T H T
- Y Q D V N C T D V S T A I H A D Q L T P

Figure 16-FF

```
23341 - CAGCTTGGCGCATATATTCTACTGGAAACAATGTATTCCAGACTCAAGCAGGCTGTCTTA -
23400
      - Q  L  G  A  Y  I  L  L  E  T  M  Y  S  R  L  K  Q  A  V  L
      -  S  L  A  H  I  F  Y  W  K  Q  C  I  P  D  S  S  R  L  S  Y
      -   A  W  R  I  Y  S  T  G  N  N  V  F  Q  T  Q  A  G  C  L  I
23401 - TAGGAGCTGAGCATGTCGACACTTCTTATGAGTGCGACATTCCTATTGGAGCTGGCATTT -
23460
      - *  E  L  S  M  S  T  L  L  M  S  A  T  F  L  L  E  L  A  F
      -  R  S  *  A  C  R  H  F  L  *  V  R  H  S  Y  W  S  W  H  L
      -   G  A  E  H  V  D  T  S  Y  E  C  D  I  P  I  G  A  G  I  C
23461 - GTGCTAGTTACCATACAGTTTCTTTATTACGTAGTACTAGCCAAAAATCTATTGTGGCTT -
23520
      - V  L  V  T  I  Q  F  L  Y  Y  V  V  L  A  K  N  L  L  W  L
      -  C  *  L  P  Y  S  F  F  I  T  *  Y  *  P  K  I  Y  C  G  L
      -   A  S  Y  H  T  V  S  L  L  R  S  T  S  Q  K  S  I  V  A  Y
23521 - ATACTATGTCTTTAGGTGCTGATAGTTCAATTGCTTACTCTAATAACACCATTGCTATAC -
23580
      - I  L  C  L  *  V  L  I  V  Q  L  L  T  L  I  T  P  L  L  Y
      -  Y  Y  V  F  R  C  *  *  F  N  C  L  L  *  *  H  H  C  Y  T
      -   T  M  S  L  G  A  D  S  S  I  A  Y  S  N  N  T  I  A  I  P
23581 - CTACTAACTTTTCAATTAGCATTACTACAGAAGTAATGCCTGTTTCTATGGCTAAAACCT -
23640
      - L  L  T  F  Q  L  A  L  L  Q  K  *  C  L  F  L  W  L  K  P
      -  Y  *  L  F  N  *  H  Y  Y  R  S  N  A  C  F  Y  G  *  N  L
      -   T  N  F  S  I  S  I  T  T  E  V  M  P  V  S  M  A  K  T  S
23641 - CCGTAGATTGTAATATGTACATCTGCGGAGATTCTACTGAATGTGCTAATTTGCTTCTCC -
23700
      - P  *  I  V  I  C  T  S  A  E  I  L  L  N  V  L  I  C  F  S
      -  R  R  L  *  Y  V  H  L  R  R  F  Y  *  M  C  *  F  A  S  P
      -   V  D  C  N  M  Y  I  C  G  D  S  T  E  C  A  N  L  L  L  Q
23701 - AATATGGTAGCTTTTGCACACAACTAAATCGTGCACTCTCAGGTATTGCTGCTGAACAGG -
23760
      - N  M  V  A  F  A  H  N  *  I  V  H  S  Q  V  L  L  L  N  R
      -  I  W  *  L  L  H  T  T  K  S  C  T  L  R  Y  C  C  *  T  G
      -   Y  G  S  F  C  T  Q  L  N  R  A  L  S  G  I  A  A  E  Q  D
23761 - ATCGCAACACACGTGAAGTGTTCGCTCAAGTCAAACAAATGTACAAAACCCCAACTTTGA -
23820
      - I  A  T  H  V  K  C  S  L  K  S  N  K  C  T  K  P  Q  L  *
      -  S  Q  H  T  *  S  V  R  S  S  Q  T  N  V  Q  N  P  N  F  E
      -   R  N  T  R  E  V  F  A  Q  V  K  Q  M  Y  K  T  P  T  L  K
23821 - AATATTTTGGTGGTTTTAATTTTTCACAAATATTACCTGACCCTCTAAAGCCAACTAAGA -
23880
      - N  I  L  V  V  L  I  F  H  K  Y  Y  L  T  L  *  S  Q  L  R
      -  I  F  W  W  F  *  F  F  T  N  I  T  *  P  S  K  A  N  *  E
      -   Y  F  G  G  F  N  F  S  Q  I  L  P  D  P  L  K  P  T  K  R
23881 - GGTCTTTTATTGAGGACTTGCTCTTTAATAAGGTGACACTCGCTGATGCTGGCTTCATGA -
23940
      - G  L  L  L  R  T  C  S  L  I  R  *  H  S  L  M  L  A  S  *
      -  V  F  Y  *  G  L  A  L  *  *  G  D  T  R  *  C  W  L  H  E
      -   S  F  I  E  D  L  L  F  N  K  V  T  L  A  D  A  G  F  M  K
23941 - AGCAATATGGCGAATGCCTAGGTGATATTAATGCTAGAGATCTCATTTGTGCGCAGAAGT -
24000
      - S  N  M  A  N  A  *  V  I  L  M  L  E  I  S  F  V  R  R  S
      -  A  I  W  R  M  P  R  *  Y  *  C  *  R  S  H  L  C  A  E  V
      -   Q  Y  G  E  C  L  G  D  I  N  A  R  D  L  I  C  A  Q  K  F
24001 - TCAATGGACTTACAGTGTTGCCACCTCTGCTCACTGATGATATGATTGCTGCCTACACTG -
```

24060

- S M D L Q C C H L C S L M I * L L P T L
- Q W T Y S V A T S A H * * Y D C C L H C
- N G L T V L P P L L T D D M I A A Y T A

Figure 16-GG

```
 24061 - CTGCTCTAGTTAGTGGTACTGCCACTGCTGGATGGACATTTGGTGCTGGCGCTGCTCTTC -
24120
       - L  L  *  L  V  V  L  P  L  L  D  G  H  L  V  L  A  L  L  F
       -  C  S  S  *  W  Y  C  H  C  W  M  D  I  W  C  W  R  C  S  S
       -   A  L  V  S  G  T  A  T  A  G  W  T  F  G  A  G  A  A  L  Q
 24121 - AAATACCTTTTGCTATGCAAATGGCATATAGGTTCAATGGCATTGGAGTTACCCAAAATG -
24180
       - K  Y  L  L  L  C  K  W  H  I  G  S  M  A  L  E  L  P  K  M
       -  N  T  F  C  Y  A  N  G  I  *  V  Q  W  H  W  S  Y  P  K  C
       -   I  P  F  A  M  Q  M  A  Y  R  F  N  G  I  G  V  T  Q  N  V
 24181 - TTCTCTATGAGAACCAAAAACAAATCGCCAACCAATTTAACAAGGCGATTAGTCAAATTC -
24240
       - F  S  M  R  T  K  N  K  S  P  T  N  L  T  R  R  L  V  K  F
       -  S  L  *  E  P  K  T  N  R  Q  P  I  *  Q  G  D  *  S  N  S
       -   L  Y  E  N  Q  K  Q  I  A  N  Q  F  N  K  A  I  S  Q  I  Q
 24241 - AAGAATCACTTACAACAACATCAACTGCATTGGGCAAGCTGCAAGACGTTGTTAACCAGA -
24300
       - K  N  H  L  Q  Q  H  Q  L  H  W  A  S  C  K  T  L  L  T  R
       -  R  I  T  Y  N  N  I  N  C  I  G  Q  A  A  R  R  C  *  P  E
       -   E  S  L  T  T  T  S  T  A  L  G  K  L  Q  D  V  V  N  Q  N
 24301 - ATGCTCAAGCATTAAACACACTTGTTAAACAACTTAGCTCTAATTTTGGTGCAATTTCAA -
24360
       - M  L  K  H  *  T  H  L  L  N  N  L  A  L  I  L  V  Q  F  Q
       -  C  S  S  I  K  H  T  C  *  T  T  *  L  *  F  W  C  N  F  K
       -   A  Q  A  L  N  T  L  V  K  Q  L  S  S  N  F  G  A  I  S  S
 24361 - GTGTGCTAAATGATATCCTTTCGCGACTTGATAAAGTCGAGGCGGAGGTACAAATTGACA -
24420
       - V  C  *  M  I  S  F  R  D  L  I  K  S  R  R  R  Y  K  L  T
       -  C  A  K  *  Y  P  F  A  T  *  *  S  R  G  G  G  T  N  *  Q
       -   V  L  N  D  I  L  S  R  L  D  K  V  E  A  E  V  Q  I  D  R
 24421 - GGTTAATTACAGGCAGACTTCAAAGCCTTCAAACCTATGTAACACAACAACTAATCAGGG -
24480
       - G  *  L  Q  A  D  F  K  A  F  K  P  M  *  H  N  N  *  S  G
       -  V  N  Y  R  Q  T  S  K  P  S  N  L  C  N  T  T  T  N  Q  G
       -   L  I  T  G  R  L  Q  S  L  Q  T  Y  V  T  Q  Q  L  I  R  A
 24481 - CTGCTGAAATCAGGGCTTCTGCTAATCTTGCTGCTACTAAAATGTCTGAGTGTGTTCTTG -
24540
       - L  L  K  S  G  L  L  L  I  L  L  L  L  K  C  L  S  V  F  L
       -  C  *  N  Q  G  F  C  *  S  C  C  Y  *  N  V  *  V  C  S  W
       -   A  E  I  R  A  S  A  N  L  A  A  T  K  M  S  E  C  V  L  G
 24541 - GACAATCAAAAAGAGTTGACTTTTGTGGAAAGGGCTACCACCTTATGTCCTTCCCACAAG -
24600
       - D  N  Q  K  E  L  T  F  V  E  R  A  T  T  L  C  P  S  H  K
       -  T  I  K  K  S  *  L  L  W  K  G  L  P  P  Y  V  L  P  T  S
       -   Q  S  K  R  V  D  F  C  G  K  G  Y  H  L  M  S  F  P  Q  A
 24601 - CAGCCCCGCATGGTGTTGTCTTCCTACATGTCACGTATGTGCCATCCCAGGAGAGGAACT -
24660
       - Q  P  R  M  V  L  S  S  Y  M  S  R  M  C  H  P  R  R  G  T
       -  S  P  A  W  C  C  L  P  T  C  H  V  C  A  I  P  G  E  E  L
       -   A  P  H  G  V  V  F  L  H  V  T  Y  V  P  S  Q  E  R  N  F
 24661 - TCACCACAGCGCCAGCAATTTGTCATGAAGGCAAAGCATACTTCCCTCGTGAAGGTGTTT -
24720
       - S  P  Q  R  Q  Q  F  V  M  K  A  K  H  T  S  L  V  K  V  F
       -  H  H  S  A  S  N  L  S  *  R  Q  S  I  L  P  S  *  R  C  F
       -   T  T  A  P  A  I  C  H  E  G  K  A  Y  F  P  R  E  G  V  F
 24721 - TTGTGTTTAATGGCACTTCTTGGTTTATTACACAGAGGAACTTCTTTTCTCCACAAATAA -
24780
```

```
      - L  C  L  M  A  L  L  G  L  L  H  R  G  T  S  F  L  H  K  *
      -  C  V  *  W  H  F  L  V  Y  Y  T  E  E  L  L  F  S  T  N  N
      -   V  F  N  G  T  S  W  F  I  T  Q  R  N  F  F  S  P  Q  I  I
24781 - TTACTACAGACAATACATTTGTCTCAGGAAATTGTGATGTCGTTATTGGCATCATTAACA -
24840
      - L  L  Q  T  I  H  L  S  Q  E  I  V  M  S  L  L  A  S  L  T
      -  Y  Y  R  Q  Y  I  C  L  R  K  L  *  C  R  Y  W  H  H  *  Q
      -   T  T  D  N  T  F  V  S  G  N  C  D  V  V  I  G  I  I  N  N
```

Figure 16-HH

```
 24841 - ACACAGTTTATGATCCTCTGCAACCTGAGCTTGACTCATTCAAAGAAGAGCTGGACAAGT -
 24900
       - T  Q  F  M  I  L  C  N  L  S  L  T  H  S  K  K  S  W  T  S
       -  H  S  L  *  S  S  A  T  *  A  *  L  I  Q  R  R  A  G  Q  V
       -   T  V  Y  D  P  L  Q  P  E  L  D  S  F  K  E  E  L  D  K  Y
 24901 - ACTTCAAAAATCATACATCACCAGATGTTGATCTTGGCGACATTTCAGGCATTAACGCTT -
 24960
       - T  S  K  I  I  H  H  Q  M  L  I  L  A  T  F  Q  A  L  T  L
       -  L  Q  K  S  Y  I  T  R  C  *  S  W  R  H  F  R  H  *  R  F
       -   F  K  N  H  T  S  P  D  V  D  L  G  D  I  S  G  I  N  A  S
 24961 - CTGTCGTCAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTCGCTAAAAATTTAAATG -
 25020
       - L  S  S  T  F  K  K  K  L  T  A  S  M  R  S  L  K  I  *  M
       -  C  R  Q  H  S  K  R  N  *  P  P  Q  *  G  R  *  K  F  K  *
       -   V  V  N  I  Q  K  E  I  D  R  L  N  E  V  A  K  N  L  N  E
 25021 - AATCACTCATTGACCTTCAAGAATTGGGAAAATATGAGCAATATATTAAATGGCCTTGGT -
 25080
       - N  H  S  L  T  F  K  N  W  E  N  M  S  N  I  L  N  G  L  G
       -  I  T  H  *  P  S  R  I  G  K  I  *  A  I  Y  *  M  A  L  V
       -   S  L  I  D  L  Q  E  L  G  K  Y  E  Q  Y  I  K  W  P  W  Y
 25081 - ATGTTTGGCTCGGCTTCATTGCTGGACTAATTGCCATCGTCATGGTTACAATCTTGCTTT -
 25140
       - M  F  G  S  A  S  L  L  D  *  L  P  S  S  W  L  Q  S  C  F
       -  C  L  A  R  L  H  C  W  T  N  C  H  R  H  G  Y  N  L  A  L
       -   V  W  L  G  F  I  A  G  L  I  A  I  V  M  V  T  I  L  L  C
 25141 - GTTGCATGACTAGTTGTTGCAGTTGCCTCAAGGGTGCATGCTCTTGTGGTTCTTGCTGCA -
 25200
       - V  A  *  L  V  V  A  V  A  S  R  V  H  A  L  V  V  L  A  A
       -  L  H  D  *  L  L  Q  L  P  Q  G  C  M  L  L  W  F  L  L  Q
       -   C  M  T  S  C  C  S  C  L  K  G  A  C  S  C  G  S  C  C  K
 25201 - AGTTTGATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTCAAATTACATTACACATAAA -
 25260
       - S  L  M  R  M  T  L  S  Q  F  S  R  V  S  N  Y  I  T  H  K
       -  V  *  *  G  *  L  *  A  S  S  Q  G  C  Q  I  T  L  H  I  N
       -   F  D  E  D  D  S  E  P  V  L  K  G  V  K  L  H  Y  T  *  T
 25261 - CGAACTTATGGATTTGTTTATGAGATTTTTTACTCTTGGATCAATTACTGCACAGCCAGT -
 25320
       - R  T  Y  G  F  V  Y  E  I  F  Y  S  W  I  N  Y  C  T  A  S
       -  E  L  M  D  L  F  M  R  F  F  T  L  G  S  I  T  A  Q  P  V
       -   N  L  W  I  C  L  *  D  F  L  L  L  D  Q  L  L  H  S  Q  *
 25321 - AAAAATTGACAATGCTTCTCCTGCAAGTACTGTTCATGCTACAGCAACGATACCGCTACA -
 25380
       - K  N  *  Q  C  F  S  C  K  Y  C  S  C  Y  S  N  D  T  A  T
       -  K  I  D  N  A  S  P  A  S  T  V  H  A  T  A  T  I  P  L  Q
       -   K  L  T  M  L  L  L  Q  V  L  F  M  L  Q  Q  R  Y  R  Y  K
 25381 - AGCCTCACTCCCTTTCGGATGGCTTGTTATTGGCGTTGCATTTCTTGCTGTTTTTCAGAG -
 25440
       - S  L  T  P  F  R  M  A  C  Y  W  R  C  I  S  C  C  F  S  E
       -  A  S  L  P  F  G  W  L  V  I  G  V  A  F  L  A  V  F  Q  S
       -   P  H  S  L  S  D  G  L  L  L  A  L  H  F  L  L  F  F  R  A
 25441 - CGCTACCAAAATAATTGCGCTCAATAAAAGATGGCAGCTAGCCCTTTATAAGGGCTTCCA -
 25500
       - R  Y  Q  N  N  C  A  Q  *  K  M  A  A  S  P  L  *  G  L  P
       -  A  T  K  I  I  A  L  N  K  R  W  Q  L  A  L  Y  K  G  F  Q
       -   L  P  K  *  L  R  S  I  K  D  G  S  *  P  F  I  R  A  S  S
 25501 - GTTCATTTGCAATTTACTGCTGCTATTTGTTACCATCTATTCACATCTTTTGCTTGTCGC -
```

25560
-V H L Q F T A A I C Y H L F T S F A C R
- F I C N L L L L F V T I Y S H L L L V A
- S F A I Y C C Y L L P S I H I F C L S L

Figure 16-II

```
25561 - TGCAGGTAAGGAGGCGCAATTTTTGTACCTCTATGCCTTGATATATTTTCTACAATGCAT -
25620
      - C  R  *  G  G  A  I  F  V  P  L  C  L  D  I  F  S  T  M  H
      -  A  G  K  E  A  Q  F  L  Y  L  Y  A  L  I  Y  F  L  Q  C  I
      -   Q  V  R  R  R  N  F  C  T  S  M  P  *  Y  I  F  Y  N  A  S
25621 - CAACGCATGTAGAATTATTATGAGATGTTGGCTTTGTTGGAAGTGCAAATCCAAGAACCC -
25680
      - Q  R  M  *  N  Y  Y  E  M  L  A  L  L  E  V  Q  I  Q  E  P
      -  N  A  C  R  I  I  M  R  C  W  L  C  W  K  C  K  S  K  N  P
      -   T  H  V  E  L  L  *  D  V  G  F  V  G  S  A  N  P  R  T  H
25681 - ATTACTTTATGATGCCAACTACTTTGTTTGCTGGCACACACATAACTATGACTACTGTAT -
25740
      - I  T  L  *  C  Q  L  L  C  L  L  A  H  T  *  L  *  L  L  Y
      -  L  L  Y  D  A  N  Y  F  V  C  W  H  T  H  N  Y  D  Y  C  I
      -   Y  F  M  M  P  T  T  L  F  A  G  T  H  I  T  M  T  T  V  Y
25741 - ACCATATAACAGTGTCACAGATACAATTGTCGTTACTGAAGGTGACGGCATTTCAACACC -
25800
      - T  I  *  Q  C  H  R  Y  N  C  R  Y  *  R  *  R  H  F  N  T
      -  P  Y  N  S  V  T  D  T  I  V  V  T  E  G  D  G  I  S  T  P
      -   H  I  T  V  S  Q  I  Q  L  S  L  L  K  V  T  A  F  Q  H  Q
25801 - AAAACTCAAAGAAGACTACCAAATTGGTGGTTATTCTGAGGATAGGCACTCAGGTGTTAA -
25860
      - K  T  Q  R  R  L  P  N  W  W  L  F  *  G  *  A  L  R  C  *
      -  K  L  K  E  D  Y  Q  I  G  G  Y  S  E  D  R  H  S  G  V  K
      -   N  S  K  K  T  T  K  L  V  V  I  L  R  I  G  T  Q  V  L  K
25861 - AGACTATGTCGTTGTACATGGCTATTTCACCGAAGTTTACTACCAGCTTGAGTCTACACA -
25920
      - R  L  C  R  C  T  W  L  F  H  R  S  L  L  P  A  *  V  Y  T
      -  D  Y  V  V  V  H  G  Y  F  T  E  V  Y  Y  Q  L  E  S  T  Q
      -   T  M  S  L  Y  M  A  I  S  P  K  F  T  T  S  L  S  L  H  K
25921 - AATTACTACAGACACTGGTATTGAAAATGCTACATTCTTCATCTTTAACAAGCTTGTTAA -
25980
      - N  Y  Y  R  H  W  Y  *  K  C  Y  I  L  H  L  *  Q  A  C  *
      -  I  T  T  D  T  G  I  E  N  A  T  F  F  I  F  N  K  L  V  K
      -   L  L  Q  T  L  V  L  K  M  L  H  S  S  S  L  T  S  L  L  K
25981 - AGACCCACCGAATGTGCAAATACACACAATCGACGGCTCTTCAGGAGTTGCTAATCCAGC -
26040
      - R  P  T  E  C  A  N  T  H  N  R  R  L  F  R  S  C  *  S  S
      -  D  P  P  N  V  Q  I  H  T  I  D  G  S  S  G  V  A  N  P  A
      -   T  H  R  M  C  K  Y  T  Q  S  T  A  L  Q  E  L  L  I  Q  Q
26041 - AATGGATCCAATTTATGATGAGCCGACGACGACTACTAGCGTGCCTTTGTAAGCACAAGA -
26100
      - N  G  S  N  L  *  *  A  D  D  D  Y  *  R  A  F  V  S  T  R
      -  M  D  P  I  Y  D  E  P  T  T  T  T  S  V  P  L  *  A  Q  E
      -   W  I  Q  F  M  M  S  R  R  R  L  L  A  C  L  C  K  H  K  K
26101 - AAGTGAGTACGAACTTATGTACTCATTCGTTTCGGAAGAAACAGGTACGTTAATAGTTAA -
26160
      - K  *  V  R  T  Y  V  L  I  R  F  G  R  N  R  Y  V  N  S  *
      -  S  E  Y  E  L  M  Y  S  F  V  S  E  E  T  G  T  L  I  V  N
      -   V  S  T  N  L  C  T  H  S  F  R  K  K  Q  V  R  *  *  L  I
26161 - TAGCGTACTTCTTTTTCTTGCTTTCGTGGTATTCTTGCTAGTCACACTAGCCATCCTTAC -
26220
      - *  R  T  S  F  S  C  F  R  G  I  L  A  S  H  T  S  H  P  Y
      -  S  V  L  L  F  L  A  F  V  V  F  L  L  V  T  L  A  I  L  T
      -   A  Y  F  F  F  L  L  S  W  Y  S  C  *  S  H  *  P  S  L  L
26221 - TGCGCTTCGATTGTGTGCGTACTGCTGCAATATTGTTAACGTGAGTTTAGTAAAACCAAC -
26280
```

```
      - C  A  S  I  V  C  V  L  L  Q  Y  C  *  R  E  F  S  K  T  N
      -  A  L  R  L  C  A  Y  C  C  N  I  V  N  V  S  L  V  K  P  T
      -   R  F  D  C  V  R  T  A  A  I  L  L  T  *  V  *  *  N  Q  R
26281 - GGTTTACGTCTACTCGCGTGTTAAAAATCTGAACTCTTCTGAAGGAGTTCCTGATCTTCT -
26340
      - G  L  R  L  L  A  C  *  K  S  E  L  F  *  R  S  S  *  S  S
      -  V  Y  V  Y  S  R  V  K  N  L  N  S  S  E  G  V  P  D  L  L
      -   F  T  S  T  R  V  L  K  I  *  T  L  L  K  E  F  L  I  F  W
```

Figure 16-JJ

```
26341 - GGTCTAAACGAACTAACTATTATTATTATTCTGTTTGGAACTTTAACATTGCTTATCATG -
26400
      - G  L  N  E  L  T  I  I  I  I  L  F  G  T  L  T  L  L  I  M
      -  V  *  T  N  *  L  L  L  L  F  C  L  E  L  *  H  C  L  S  W
      -   S  K  R  T  N  Y  Y  Y  Y  S  V  W  N  F  N  I  A  Y  H  G
26401 - GCAGACAACGGTACTATTACCGTTGAGGAGCTTAAACAACTCCTGGAACAATGGAACCTA -
26460
      - A  D  N  G  T  I  T  V  E  E  L  K  Q  L  L  E  Q  W  N  L
      -  Q  T  T  V  L  L  P  L  R  S  L  N  N  S  W  N  N  G  T  *
      -   R  Q  R  Y  Y  Y  R  *  G  A  *  T  T  P  G  T  M  E  P  S
26461 - GTAATAGGTTTCCTATTCCTAGCCTGGATTATGTTACTACAATTTGCCTATTCTAATCGG -
26520
      - V  I  G  F  L  F  L  A  W  I  M  L  L  Q  F  A  Y  S  N  R
      -  *  *  V  S  Y  S  *  P  G  L  C  Y  Y  N  L  P  I  L  I  G
      -   N  R  F  P  I  P  S  L  D  Y  V  T  T  I  C  L  F  *  S  E
26521 - AACAGGTTTTTGTACATAATAAAGCTTGTTTTCCTCTGGCTCTTGTGGCCAGTAACACTT -
26580
      - N  R  F  L  Y  I  I  K  L  V  F  L  W  L  L  W  P  V  T  L
      -  T  G  F  C  T  *  *  S  L  F  S  S  G  S  C  G  Q  *  H  L
      -   Q  V  F  V  H  N  K  A  C  F  P  L  A  L  V  A  S  N  T  C
26581 - GCTTGTTTTGTGCTTGCTGTTGTCTACAGAATTAATTGGGTGACTGGCGGGATTGCGATT -
26640
      - A  C  F  V  L  A  V  V  Y  R  I  N  W  V  T  G  G  I  A  I
      -  L  V  L  C  L  L  L  S  T  E  L  I  G  *  L  A  G  L  R  L
      -   L  F  C  A  C  C  C  L  Q  N  *  L  G  D  W  R  D  C  D  C
26641 - GCAATGGCTTGTATTGTAGGCTTGATGTGGCTTAGCTACTTCGTTGCTTCCTTCAGGCTG -
26700
      - A  M  A  C  I  V  G  L  M  W  L  S  Y  F  V  A  S  F  R  L
      -  Q  W  L  V  L  *  A  *  C  G  L  A  T  S  L  L  P  S  G  C
      -   N  G  L  Y  C  R  L  D  V  A  *  L  L  R  C  F  L  Q  A  V
26701 - TTTGCTCGTACCCGCTCAATGTGGTCATTCAACCCAGAAACAAACATTCTTCTCAATGTG -
26760
      - F  A  R  T  R  S  M  W  S  F  N  P  E  T  N  I  L  L  N  V
      -  L  L  V  P  A  Q  C  G  H  S  T  Q  K  Q  T  F  F  S  M  C
      -   C  S  Y  P  L  N  V  V  I  Q  P  R  N  K  H  S  S  Q  C  A
26761 - CCTCTCCGGGGGACAATTGTGACCAGACCGCTCATGGAAAGTGAACTTGTCATTGGTGCT -
26820
      - P  L  R  G  T  I  V  T  R  P  L  M  E  S  E  L  V  I  G  A
      -  L  S  G  G  Q  L  *  P  D  R  S  W  K  V  N  L  S  L  V  L
      -   S  P  G  D  N  C  D  Q  T  A  H  G  K  *  T  C  H  W  C  C
26821 - GTGATCATTCGTGGTCACTTGCGAATGGCCGGACACTCCCTAGGGCGCTGTGACATTAAG -
26880
      - V  I  I  R  G  H  L  R  M  A  G  H  S  L  G  R  C  D  I  K
      -  *  S  F  V  V  T  C  E  W  P  D  T  P  *  G  A  V  T  L  R
      -   D  H  S  W  S  L  A  N  G  R  T  L  P  R  A  L  *  H  *  G
26881 - GACCTGCCAAAAGAGATCACTGTGGCTACATCACGAACGCTTTCTTATTACAAATTAGGA -
26940
      - D  L  P  K  E  I  T  V  A  T  S  R  T  L  S  Y  Y  K  L  G
      -  T  C  Q  K  R  S  L  W  L  H  H  E  R  F  L  I  T  N  *  E
      -   P  A  K  R  D  H  C  G  Y  I  T  N  A  F  L  L  Q  I  R  S
26941 - GCGTCGCAGCGTGTAGGCACTGATTCAGGTTTTGCTGCATACAACCGCTACCGTATTGGA -
27000
      - A  S  Q  R  V  G  T  D  S  G  F  A  A  Y  N  R  Y  R  I  G
      -  R  R  S  V  *  A  L  I  Q  V  L  L  H  T  T  A  T  V  L  E
      -   V  A  A  C  R  H  *  F  R  F  C  C  I  Q  P  L  P  Y  W  K
27001 - AACTATAAATTAAATACAGACCACGCCGGTAGCAACGACAATATTGCTTTGCTAGTACAG -
27060
```

```
-N  Y  K  L  N  T  D  H  A  G  S  N  D  N  I  A  L  L  V  Q
-  T  I  N  *  I  Q  T  T  P  V  A  T  T  I  L  L  C  *  Y  S
-   L  *  I  K  Y  R  P  R  R  *  Q  R  Q  Y  C  F  A  S  T  V
```

Figure 16-KK

```
27061 - TAAGTGACAACAGATGTTTCATCTTGTTGACTTCCAGGTTACAATAGCAGAGATATTGAT -
27120
      -* V  T  T  D  V  S  S  C  *  L  P  G  Y  N  S  R  D  I  D
      - K  *  Q  Q  M  F  H  L  V  D  F  Q  V  T  I  A  E  I  L  I
      -  S  D  N  R  C  F  I  L  L  T  S  R  L  Q  *  Q  R  Y  *  L
27121 - TATCATTATGAGGACTTTCAGGATTGCTATTTGGAATCTTGACGTTATAATAAGTTCAAT -
27180
      -Y  H  Y  E  D  F  Q  D  C  Y  L  E  S  *  R  Y  N  K  F  N
      - I  I  M  R  T  F  R  I  A  I  W  N  L  D  V  I  I  S  S  I
      -  S  L  *  G  L  S  G  L  L  F  G  I  L  T  L  *  *  V  Q  *
27181 - AGTGAGACAATTATTTAAGCCTCTAACTAAGAAGAATTATTCGGAGTTAGATGATGAAGA -
27240
      -S  E  T  I  I  *  A  S  N  *  E  E  L  F  G  V  R  *  *  R
      - V  R  Q  L  F  K  P  L  T  K  K  N  Y  S  E  L  D  D  E  E
      -  *  D  N  Y  L  S  L  *  L  R  R  I  I  R  S  *  M  M  K  N
27241 - ACCTATGGAGTTAGATTATCCATAAAACGAACATGAAAATTATTCTCTTCCTGACATTGA -
27300
      -T  Y  G  V  R  L  S  I  K  R  T  *  K  L  F  S  S  *  H  *
      - P  M  E  L  D  Y  P  *  N  E  H  E  N  Y  S  L  P  D  I  D
      -  L  W  S  *  I  I  H  K  T  N  M  K  I  I  L  F  L  T  L  I
27301 - TTGTATTTACATCTTGCGAGCTATATCACTATCAGGAGTGTGTTAGAGGTACGACTGTAC -
27360
      -L  Y  L  H  L  A  S  Y  I  T  I  R  S  V  L  E  V  R  L  Y
      - C  I  Y  I  L  R  A  I  S  L  S  G  V  C  *  R  Y  D  C  T
      -  V  F  T  S  C  E  L  Y  H  Y  Q  E  C  V  R  G  T  T  V  L
27361 - TACTAAAAGAACCTTGCCCATCAGGAACATACGAGGGCAATTCACCATTTCACCCTCTTG -
27420
      -Y  *  K  N  L  A  H  Q  E  H  T  R  A  I  H  H  F  T  L  L
      - T  K  R  T  L  P  I  R  N  I  R  G  Q  F  T  I  S  P  S  C
      -  L  K  E  P  C  P  S  G  T  Y  E  G  N  S  P  F  H  P  L  A
27421 - CTGACAATAAATTTGCACTAACTTGCACTAGCACACACTTTGCTTTTGCTTGTGCTGACG -
27480
      -L  T  I  N  L  H  *  L  A  L  A  H  T  L  L  L  L  V  L  T
      - *  Q  *  I  C  T  N  L  H  *  H  T  L  C  F  C  L  C  *  R
      -  D  N  K  F  A  L  T  C  T  S  T  H  F  A  F  A  C  A  D  G
27481 - GTACTCGACATACCTATCAGCTGCGTGCAAGATCAGTTTCACCAAAACTTTTCATCAGAC -
27540
      -V  L  D  I  P  I  S  C  V  Q  D  Q  F  H  Q  N  F  S  S  D
      - Y  S  T  Y  L  S  A  A  C  K  I  S  F  T  K  T  F  H  Q  T
      -  T  R  H  T  Y  Q  L  R  A  R  S  V  S  P  K  L  F  I  R  Q
27541 - AAGAGGAGGTTCAACAAGAGCTCTACTCGCCACTTTTTCTCATTGTTGCTGCTCTAGTAT -
27600
      -K  R  R  F  N  K  S  S  T  R  H  F  F  S  L  L  L  L  *  Y
      - R  G  G  S  T  R  A  L  L  A  T  F  S  H  C  C  C  S  S  I
      -  E  E  V  Q  Q  E  L  Y  S  P  L  F  L  I  V  A  A  L  V  F
27601 - TTTTAATACTTTGCTTCACCATTAAGAGAAAGACAGAATGAATGAGCTCACTTTAATTGA -
27660
      -F  *  Y  F  A  S  P  L  R  E  R  Q  N  E  *  A  H  F  N  *
      - F  N  T  L  L  H  H  *  E  K  D  R  M  N  E  L  T  L  I  D
      -  L  I  L  C  F  T  I  K  R  K  T  E  *  M  S  S  L  *  L  T
27661 - CTTCTATTTGTGCTTTTTAGCCTTTCTGCTATTCCTTGTTTTAATAATGCTTATTATATT -
27720
      -L  L  F  V  L  F  S  L  S  A  I  P  C  F  N  N  A  Y  Y  I
      - F  Y  L  C  F  L  A  F  L  L  F  L  V  L  I  M  L  I  I  F
      -  S  I  C  A  F  *  P  F  C  Y  S  L  F  *  *  C  L  L  Y  F
```

```
27721 - TTGGTTTTCACTCGAAATCCAGGATCTAGAAGAACCTTGTACCAAAGTCTAAACGAACAT -
27780
      -L  V  F  T  R  N  P  G  S  R  R  T  L  Y  Q  S  L  N  E  H
      - W  F  S  L  E  I  Q  D  L  E  E  P  C  T  K  V  *  T  N  M
      -  G  F  H  S  K  S  R  I  *  K  N  L  V  P  K  S  K  R  T  *
```

Figure 16-LL

```
27781 - GAAACTTCTCATTGTTTTGACTTGTATTTCTCTATGCAGTTGCATATGCACTGTAGTACA -
27840
      -E  T  S  H  C  F  D  L  Y  F  S  M  Q  L  H  M  H  C  S  T
      - K  L  L  I  V  L  T  C  I  S  L  C  S  C  I  C  T  V  V  Q
      -  N  F  S  L  F  *  L  V  F  L  Y  A  V  A  Y  A  L  *  Y  S
27841 - GCGCTGTGCATCTAATAAACCTCATGTGCTTGAAGATCCTTGTAAGGTACAACACTAGGG -
27900
      -A  L  C  I  *  *  T  S  C  A  *  R  S  L  *  G  T  T  L  G
      - R  C  A  S  N  K  P  H  V  L  E  D  P  C  K  V  Q  H  *  G
      -  A  V  H  L  I  N  L  M  C  L  K  I  L  V  R  Y  N  T  R  G
27901 - GTAATACTTATAGCACTGCTTGGCTTTGTGCTCTAGGAAAGGTTTTACCTTTTCATAGAT -
27960
      -V  I  L  I  A  L  L  G  F  V  L  *  E  R  F  Y  L  F  I  D
      - *  Y  L  *  H  C  L  A  L  C  S  R  K  G  F  T  F  S  *  M
      -  N  T  Y  S  T  A  W  L  C  A  L  G  K  V  L  P  F  H  R  W
27961 - GGCACACTATGGTTCAAACATGCACACCTAATGTTACTATCAACTGTCAAGATCCAGCTG -
28020
      -G  T  L  W  F  K  H  A  H  L  M  L  L  S  T  V  K  I  Q  L
      - A  H  Y  G  S  N  M  H  T  *  C  Y  Y  Q  L  S  R  S  S  W
      -  H  T  M  V  Q  T  C  T  P  N  V  T  I  N  C  Q  D  P  A  G
28021 - GTGGTGCGCTTATAGCTAGGTGTTGGTACCTTCATGAAGGTCACCAAACTGCTGCATTTA -
28080
      -V  V  R  L  *  L  G  V  G  T  F  M  K  V  T  K  L  L  H  L
      - W  C  A  Y  S  *  V  L  V  P  S  *  R  S  P  N  C  C  I  *
      -  G  A  L  I  A  R  C  W  Y  L  H  E  G  H  Q  T  A  A  F  R
28081 - GAGACGTACTTGTTGTTTTAAATAAACGAACAAATTAAAATGTCTGATAATGGACCCCAA -
28140
      -E  T  Y  L  L  F  *  I  N  E  Q  I  K  M  S  D  N  G  P  Q
      - R  R  T  C  C  F  K  *  T  N  K  L  K  C  L  I  M  D  P  N
      -  D  V  L  V  V  L  N  K  R  T  N  *  N  V  *  *  W  T  P  I
28141 - TCAAACCAACGTAGTGCCCCCCGCATTACATTTGGTGGACCCACAGATTCAACTGACAAT -
28200
      -S  N  Q  R  S  A  P  R  I  T  F  G  G  P  T  D  S  T  D  N
      - Q  T  N  V  V  P  P  A  L  H  L  V  D  P  Q  I  Q  L  T  I
      -  K  P  T  *  C  P  P  H  Y  I  W  W  T  H  R  F  N  *  Q  *
28201 - AACCAGAATGGAGGACGCAATGGGGCAAGGCCAAAACAGCGCCGACCCCAAGGTTTACCC -
28260
      -N  Q  N  G  G  R  N  G  A  R  P  K  Q  R  R  P  Q  G  L  P
      - T  R  M  E  D  A  M  G  Q  G  Q  N  S  A  D  P  K  V  Y  P
      -  P  E  W  R  T  Q  W  G  K  A  K  T  A  P  T  P  R  F  T  Q
28261 - AATAATACTGCGTCTTGGTTCACAGCTCTCACTCAGCATGGCAAGGAGGAACTTAGATTC -
28320
      -N  N  T  A  S  W  F  T  A  L  T  Q  H  G  K  E  E  L  R  F
      - I  I  L  R  L  G  S  Q  L  S  L  S  M  A  R  R  N  L  D  S
      -  *  Y  C  V  L  V  H  S  S  H  S  A  W  Q  G  G  T  *  I  P
28321 - CCTCGAGGCCAGGGCGTTCCAATCAACACCAATAGTGGTCCAGATGACCAAATTGGCTAC -
28380
      -P  R  G  Q  G  V  P  I  N  T  N  S  G  P  D  D  Q  I  G  Y
      - L  E  A  R  A  F  Q  S  T  P  I  V  V  Q  M  T  K  L  A  T
      -  S  R  P  G  R  S  N  Q  H  Q  *  W  S  R  *  P  N  W  L  L
28381 - TACCGAAGAGCTACCCGACGAGTTCGTGGTGGTGACGGCAAAATGAAAGAGCTCAGCCCC -
28440
      -Y  R  R  A  T  R  R  V  R  G  G  D  G  K  M  K  E  L  S  P
      - T  E  E  L  P  D  E  F  V  V  V  T  A  K  *  K  S  S  A  P
      -  P  K  S  Y  P  T  S  S  W  W  *  R  Q  N  E  R  A  Q  P  Q
```

28441 - AGATGGTACTTCTATTACCTAGGAACTGGCCCAGAAGCTTCACTTCCCTACGGCGCTAAC - 28500

- R W Y F Y Y L G T G P E A S L P Y G A N
- D G T S I T * E L A Q K L H F P T A L T
- M V L L L P R N W P R S F T S L R R * Q

Figure 16-MM

```
 28501 - AAAGAAGGCATCGTATGGGTTGCAACTGAGGGAGCCTTGAATACACCCAAAGACCACATT -
28560
       - K  E  G  I  V  W  V  A  T  E  G  A  L  N  T  P  K  D  H  I
       -  K  K  A  S  Y  G  L  Q  L  R  E  P  *  I  H  P  K  T  T  L
       -   R  R  H  R  M  G  C  N  *  G  S  L  E  Y  T  Q  R  P  H  W
 28561 - GGCACCCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACTTCCTCAAGGAACAACA -
28620
       - G  T  R  N  P  N  N  N  A  A  T  V  L  Q  L  P  Q  G  T  T
       -  A  P  A  I  L  I  T  M  L  P  P  C  Y  N  F  L  K  E  Q  H
       -   H  P  Q  S  *  *  Q  C  C  H  R  A  T  T  S  S  R  N  N  I
 28621 - TTGCCAAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAGTCAAGCCTCTTCTCGCTCC -
28680
       - L  P  K  G  F  Y  A  E  G  S  R  G  G  S  Q  A  S  S  R  S
       -  C  Q  K  A  S  T  Q  R  E  A  E  A  A  V  K  P  L  L  A  P
       -   A  K  R  L  L  R  R  G  K  Q  R  R  Q  S  S  L  F  S  L  L
 28681 - TCATCACGTAGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAAATTCT -
28740
       - S  S  R  S  R  G  N  S  R  N  S  T  P  G  S  S  R  G  N  S
       -  H  H  V  V  A  V  I  Q  E  I  Q  L  L  A  A  V  G  E  I  L
       -   I  T  *  S  R  *  F  K  K  F  N  S  W  Q  Q  *  G  K  F  S
 28741 - CCTGCTCGAATGGCTAGCGGAGGTGGTGAAACTGCCCTCGCGCTATTGCTGCTAGACAGA -
28800
       - P  A  R  M  A  S  G  G  G  E  T  A  L  A  L  L  L  L  D  R
       -  L  L  E  W  L  A  E  V  V  K  L  P  S  R  Y  C  C  *  T  D
       -   C  S  N  G  *  R  R  W  *  N  C  P  R  A  I  A  A  R  Q  I
 28801 - TTGAACCAGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAACAACAAGGCCAAACTGTC -
28860
       - L  N  Q  L  E  S  K  V  S  G  K  G  Q  Q  Q  Q  G  Q  T  V
       -  *  T  S  L  R  A  K  F  L  V  K  A  N  N  N  K  A  K  L  S
       -   E  P  A  *  E  Q  S  F  W  *  R  P  T  T  T  R  P  N  C  H
 28861 - ACTAAGAAATCTGCTGCTGAGGCATCTAAAAAGCCTCGCCAAAAACGTACTGCCACAAAA -
28920
       - T  K  K  S  A  A  E  A  S  K  K  P  R  Q  K  R  T  A  T  K
       -  L  R  N  L  L  L  R  H  L  K  S  L  A  K  N  V  L  P  Q  N
       -   *  E  I  C  C  *  G  I  *  K  A  S  P  K  T  Y  C  H  K  T
 28921 - CAGTACAACGTCACTCAAGCATTTGGGAGACGTGGTCCAGAACAAACCCAAGGAAATTTC -
28980
       - Q  Y  N  V  T  Q  A  F  G  R  R  G  P  E  Q  T  Q  G  N  F
       -  S  T  T  S  L  K  H  L  G  D  V  V  Q  N  K  P  K  E  I  S
       -   V  Q  R  H  S  S  I  W  E  T  W  S  R  T  N  P  R  K  F  R
 28981 - GGGGACCAAGACCTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAA -
29040
       - G  D  Q  D  L  I  R  Q  G  T  D  Y  K  H  W  P  Q  I  A  Q
       -  G  T  K  T  *  S  D  K  E  L  I  T  N  I  G  R  K  L  H  N
       -   G  P  R  P  N  Q  T  R  N  *  L  Q  T  L  A  A  N  C  T  I
 29041 - TTTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCACACCT -
29100
       - F  A  P  S  A  S  A  F  F  G  M  S  R  I  G  M  E  V  T  P
       -  L  L  Q  V  P  L  H  S  L  E  C  H  A  L  A  W  K  S  H  L
       -   C  S  K  C  L  C  I  L  W  N  V  T  H  W  H  G  S  H  T  F
 29101 - TCGGGAACATGGCTGACTTATCATGGAGCCATTAAATTGGATGACAAAGATCCACAATTC -
29160
       - S  G  T  W  L  T  Y  H  G  A  I  K  L  D  D  K  D  P  Q  F
       -  R  E  H  G  *  L  I  M  E  P  L  N  W  M  T  K  I  H  N  S
       -   G  N  M  A  D  L  S  W  S  H  *  I  G  *  Q  R  S  T  I  Q
 29161 - AAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCCCACCAACA -
29220
```

```
      - K  D  N  V  I  L  L  N  K  H  I  D  A  Y  K  T  F  P  P  T
      -  K  T  T  S  Y  C  *  T  S  T  L  T  H  T  K  H  S  H  Q  Q
      -   R  Q  R  H  T  A  E  Q  A  H  *  R  I  Q  N  I  P  T  N  R
29221 - GAGCCTAAAAAGGACAAAAAGAAAAAGACTGATGAAGCTCAGCCTTTGCCGCAGAGACAA -
29280
      - E  P  K  K  D  K  K  K  K  T  D  E  A  Q  P  L  P  Q  R  Q
      -  S  L  K  R  T  K  R  K  R  L  M  K  L  S  L  C  R  R  D  K
      -   A  *  K  G  Q  K  E  K  D  *  *  S  S  A  F  A  A  E  T  K
```

Figure 16-NN

```
29281 - AAGAAGCAGCCCACTGTGACTCTTCTTCCTGCGGCTGACATGGATGATTTCTCCAGACAA -
29340
      -K  K  Q  P  T  V  T  L  L  P  A  A  D  M  D  D  F  S  R  Q
      -  R  S  S  P  L  *  L  F  F  L  R  L  T  W  M  I  S  P  D  N
      -   E  A  A  H  C  D  S  S  S  C  G  *  H  G  *  F  L  Q  T  T
29341 - CTTCAAAATTCCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATG -
29400
      -L  Q  N  S  M  S  G  A  S  A  D  S  T  Q  A  *  T  L  M  M
      -  F  K  I  P  *  V  E  L  L  L  I  Q  L  R  H  K  H  S  *  *
      -   S  K  F  H  E  W  S  F  C  *  F  N  S  G  I  N  T  H  D  D
29401 - ACCACACAAGGCAGATGGGCTATGTAAACGTTTTCGCAATTCCGTTTACGATACATAGTC -
29460
      -T  T  Q  G  R  W  A  M  *  T  F  S  Q  F  R  L  R  Y  I  V
      -  P  H  K  A  D  G  L  C  K  R  F  R  N  S  V  Y  D  T  *  S
      -   H  T  R  Q  M  G  Y  V  N  V  F  A  I  P  F  T  I  H  S  L
29461 - TACTCTTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTTAACTTTA -
29520
      -Y  S  C  A  E  *  I  L  V  T  K  Q  H  K  *  V  *  L  T  L
      -  T  L  V  Q  N  E  F  S  *  L  N  S  T  S  R  F  S  *  L  *
      -   L  L  C  R  M  N  S  R  N  *  T  A  Q  V  G  L  V  N  F  N
29521 - ATCTCACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAAGAGCCACCA -
29580
      -I  S  H  S  N  L  *  S  M  C  N  I  R  E  D  L  K  E  P  P
      -  S  H  I  A  I  F  N  Q  C  V  T  L  G  R  T  *  K  S  H  H
      -   L  T  *  Q  S  L  I  N  V  *  H  *  G  G  L  E  R  A  T  T
29581 - CATTTTCATCGAGGCCACGCGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAG -
29640
      -H  F  H  R  G  H  A  E  Y  D  R  G  Y  S  E  *  C  *  G  E
      -  I  F  I  E  A  T  R  S  T  I  E  G  T  V  N  N  A  R  E  S
      -   F  S  S  R  P  R  G  V  R  S  R  V  Q  *  I  M  L  G  R  A
29641 - CTGCCTATATGGAAGAGCCCTAATGTGTAAAATTAATTTTAGTAGTGCTATCCCCATGTG -
29700
      -L  P  I  W  K  S  P  N  V  *  N  *  F  *  *  C  Y  P  H  V
      -  C  L  Y  G  R  A  L  M  C  K  I  N  F  S  S  A  I  P  M  *
      -   A  Y  M  E  E  P  *  C  V  K  L  I  L  V  V  L  S  P  C  D
29701 - ATTTTAATAGCTTCTTAGGAGAATGACAAAAAAAAAAAAAAA - 29742
      -I  L  I  A  S  *  E  N  D  K  K  K  K  K  X
      -  F  *  *  L  L  R  R  M  T  K  K  K  K  X
      -   F  N  S  F  L  G  E  *  Q  K  K  K  K  X
```

Figure 17-A

1 - TTTTTTTTTTTTTTTGTCATTCTCCTAAGAAGCTATTAAAATCACATGGGGATAGCACTA - 60
- F F F F F V I L L R S Y * N H M G I A L
- F F F F L S F S * E A I K I T W G * H Y
- F F F F C H S P K K L L K S H G D S T T
61 - CTAAAATTAATTTTACACATTAGGGCTCTTCCATATAGGCAGCTCTCCCTAGCATTATTC - 120
- L K L I L H I R A L P Y R Q L S L A L F
- * N * F Y T L G L F H I G S S P * H Y S
- K I N F T H * G S S I * A A L P S I I H
121 - ACTGTACCCTCGATCGTACTCCGCGTGGCCTCGATGAAAATGTGGTGGCTCTTTCAAGTC - 180
- T V P S I V L R V A S M K M W W L F Q V
- L Y P R S Y S A W P R * K C G G S F K S
- C T L D R T P R G L D E N V V A L S S P
181 - CTCCCTAATGTTACACATTGATTAAAGATTGCTATGTGAGATTAAAGTTAACTAAACCTA - 240
- L P N V T H * L K I A M * D * S * L N L
- S L M L H I D * R L L C E I K V N * T Y
- P * C Y T L I K D C Y V R L K L T K P T
241 - CTTGTGCTGTTTAGTTACGAGAATTCATTCTGCACAAGAGTAGACTATGTATCGTAAACG - 300
- L V L F S Y E N S F C T R V D Y V S * T
- L C C L V T R I H S A Q E * T M Y R K R
- C A V * L R E F I L H K S R L C I V N G
301 - GAATTGCGAAAACGTTTACATAGCCCATCTGCCTTGTGTGGTCATCATGAGTGTTTATGC - 360
- E L R K R L H S P S A L C G H H E C L C
- N C E N V Y I A H L P C V V I M S V Y A
- I A K T F T * P I C L V W S S * V F M P
361 - CTGAGTTGAATCAGCAGAAGCTCCACTCATGGAATTTTGAAGTTGTCTGGAGAAATCATC - 420
- L S * I S R S S T H G I L K L S G E I I
- * V E S A E A P L M E F * S C L E K S S
- E L N Q Q K L H S W N F E V V W R N H P
421 - CATGTCAGCCGCAGGAAGAAGAGTCACAGTGGGCTGCTTCTTTTGTCTCTGCGGCAAAGG - 480
- H V S R R K K S H S G L L L L S L R Q R
- M S A A G R R V T V G C F F C L C G K G
- C Q P Q E E E S Q W A A S F V S A A K A
481 - CTGAGCTTCATCAGTCTTTTTCTTTTTGTCCTTTTTAGGCTCTGTTGGTGGGAATGTTTT - 540
- L S F I S L F L F V L F R L C W W E C F
- * A S S V F F F L S F L G S V G G N V L
- E L H Q S F S F C P F * A L L V G M F C
541 - GTATGCGTCAATGTGCTTGTTCAGCAGTATGACGTTGTCTTTGAATTGTGGATCTTTGTC - 600
- V C V N V L V Q Q Y D V V F E L W I F V
- Y A S M C L F S S M T L S L N C G S L S
- M R Q C A C S A V * R C L * I V D L C H
601 - ATCCAATTTAATGGCTCCATGATAAGTCAGCCATGTTCCCGAAGGTGTGACTTCCATGCC - 660
- I Q F N G S M I S Q P C S R R C D F H A
- S N L M A P * * V S H V P E G V T S M P
- P I * W L H D K S A M F P K V * L P C Q
661 - AATGCGTGACATTCCAAAGAATGCAGAGGCACTTGGAGCAAATTGTGCAATTTGCGGCCA - 720
- N A * H S K E C R G T W S K L C N L R P
- M R D I P K N A E A L G A N C A I C G Q
- C V T F Q R M Q R H L E Q I V Q F A A N

Figure 17-B

```
 721 - ATGTTTGTAATCAGTTCCTTGTCTGATTAGGTCTTGGTCCCCGAAATTTCCTTGGGTTTG - 780
     -M  F  V  I  S  S  L  S  D  *  V  L  V  P  E  I  S  L  G  L
     - C  L  *  S  V  P  C  L  I  R  S  W  S  P  K  F  P  W  V  C
     -  V  C  N  Q  F  L  V  *  L  G  L  G  P  R  N  F  L  G  F  V
 781 - TTCTGGACCACGTCTCCCAAATGCTTGAGTGACGTTGTACTGTTTTGTGGCAGTACGTTT - 840
     -F  W  T  T  S  P  K  C  L  S  D  V  V  L  F  C  G  S  T  F
     - S  G  P  R  L  P  N  A  *  V  T  L  Y  C  F  V  A  V  R  F
     -  L  D  H  V  S  Q  M  L  E  *  R  C  T  V  L  W  Q  Y  V  F
 841 - TTGGCGAGGCTTTTTAGATGCCTCAGCAGCAGATTTCTTAGTGACAGTTTGGCCTTGTTG - 900
     -L  A  R  L  F  R  C  L  S  S  R  F  L  S  D  S  L  A  L  L
     - W  R  G  F  L  D  A  S  A  A  D  F  L  V  T  V  W  P  C  C
     -  G  E  A  F  *  M  P  Q  Q  Q  I  S  *  *  Q  F  G  L  V  V
 901 - TTGTTGGCCTTTACCAGAAACTTTGCTCTCAAGCTGGTTCAATCTGTCTAGCAGCAATAG - 960
     -L  L  A  F  T  R  N  F  A  L  K  L  V  Q  S  V  *  Q  Q  *
     - C  W  P  L  P  E  T  L  L  S  S  W  F  N  L  S  S  S  N  S
     -  V  G  L  Y  Q  K  L  C  S  Q  A  G  S  I  C  L  A  A  I  A
 961 - CGCGAGGGCAGTTTCACCACCTCCGCTAGCCATTCGAGCAGGAGAATTTCCCCTACTGCT -
1020
     -R  E  G  S  F  T  T  S  A  S  H  S  S  R  R  I  S  P  T  A
     - A  R  A  V  S  P  P  P  L  A  I  R  A  G  E  F  P  L  L  L
     -  R  G  Q  F  H  H  L  R  *  P  F  E  Q  E  N  F  P  Y  C  C
1021 - GCCAGGAGTTGAATTTCTTGAATTACCGCGACTACGTGATGAGGAGCGAGAAGAGGCTTG -
1080
     -A  R  S  *  I  S  *  I  T  A  T  T  *  *  G  A  R  R  G  L
     - P  G  V  E  F  L  E  L  P  R  L  R  D  E  E  R  E  E  A  *
     -  Q  E  L  N  F  L  N  Y  R  D  Y  V  M  R  S  E  K  R  L  D
1081 - ACTGCCGCCTCTGCTTCCCTCTGCGTAGAAGCCTTTTGGCAATGTTGTTCCTTGAGGAAG -
1140
     -T  A  A  S  A  S  L  C  V  E  A  F  W  Q  C  C  S  L  R  K
     - L  P  P  L  L  P  S  A  *  K  P  F  G  N  V  V  P  *  G  S
     -  C  R  L  C  F  P  L  R  R  S  L  L  A  M  L  F  L  E  E  V
1141 - TTGTAGCACGGTGGCAGCATTGTTATTAGGATTGCGGGTGCCAATGTGGTCTTTGGGTGT -
1200
     -L  *  H  G  G  S  I  V  I  R  I  A  G  A  N  V  V  F  G  C
     - C  S  T  V  A  A  L  L  L  G  L  R  V  P  M  W  S  L  G  V
     -  V  A  R  W  Q  H  C  Y  *  D  C  G  C  Q  C  G  L  W  V  Y
1201 - ATTCAAGGCTCCCTCAGTTGCAACCCATACGATGCCTTCTTTGTTAGCGCCGTAGGGAAG -
1260
     -I  Q  G  S  L  S  C  N  P  Y  D  A  F  F  V  S  A  V  G  K
     - F  K  A  P  S  V  A  T  H  T  M  P  S  L  L  A  P  *  G  S
     -  S  R  L  P  Q  L  Q  P  I  R  C  L  L  C  *  R  R  R  E  V
1261 - TGAAGCTTCTGGGCCAGTTCCTAGGTAATAGAAGTACCATCTGGGGCTGAGCTCTTTCAT -
1320
     -*  S  F  W  A  S  S  *  V  I  E  V  P  S  G  A  E  L  F  H
     - E  A  S  G  P  V  P  R  *  *  K  Y  H  L  G  L  S  S  F  I
     -  K  L  L  G  Q  F  L  G  N  R  S  T  I  W  G  *  A  L  S  F
1321 - TTTGCCGTCACCACCACGAACTCGTCGGGTAGCTCTTCGGTAGTAGCCAATTTGGTCATC -
1380
     -F  A  V  T  T  T  N  S  S  G  S  S  S  V  V  A  N  L  V  I
     - L  P  S  P  P  R  T  R  R  V  A  L  R  *  *  P  I  W  S  S
     -  C  R  H  H  H  E  L  V  G  *  L  F  G  S  S  Q  F  G  H  L
1381 - TGGACCACTATTGGTGTTGATTGGAACGCCCTGGCCTCGAGGGAATCTAAGTTCCTCCTT -
1440
     -W  T  T  I  G  V  D  W  N  A  L  A  S  R  E  S  K  F  L  L
     - G  P  L  L  V  L  I  G  T  P  W  P  R  G  N  L  S  S  S  L
     -  D  H  Y  W  C  *  L  E  R  P  G  L  E  G  I  *  V  P  P  C
1441 - GCCATGCTGAGTGAGAGCTGTGAACCAAGACGCAGTATTATTGGGTAAACCTTGGGGTCG -
1500
```

```
-A  M  L  S  E  S  C  E  P  R  R  S  I  I  G  *  T  L  G  S
- P  C  *  V  R  A  V  N  Q  D  A  V  L  L  G  K  P  W  G  R
-  H  A  E  *  E  L  *  T  K  T  Q  Y  Y  W  V  N  L  G  V  G
```

Figure 17-C

```
 1501 - GCGCTGTTTTGGCCTTGCCCCATTGCGTCCTCCATTCTGGTTATTGTCAGTTGAATCTGT -
1560
       -A  L  F  W  P  C  P  I  A  S  S  I  L  V  I  V  S  *  I  C
       - R  C  F  G  L  A  P  L  R  P  P  F  W  L  L  S  V  E  S  V
       -  A  V  L  A  L  P  H  C  V  L  H  S  G  Y  C  Q  L  N  L  W
 1561 - GGGTCCACCAAATGTAATGCGGGGGGCACTACGTTGGTTTGATTGGGGTCCATTATCAGA -
1620
       -G  S  T  K  C  N  A  G  G  T  T  L  V  *  L  G  S  I  I  R
       - G  P  P  N  V  M  R  G  A  L  R  W  F  D  W  G  P  L  S  D
       -  V  H  Q  M  *  C  G  G  H  Y  V  G  L  I  G  V  H  Y  Q  T
 1621 - CATTTTAATTTGTTCGTTTATTTAAAACAACAAGTACGTCTCTAAATGCAGCAGTTTGGT -
1680
       -H  F  N  L  F  V  Y  L  K  Q  Q  V  R  L  *  M  Q  Q  F  G
       - I  L  I  C  S  F  I  *  N  N  K  Y  V  S  K  C  S  S  L  V
       -  F  *  F  V  R  L  F  K  T  T  S  T  S  L  N  A  A  V  W  *
 1681 - GACCTTCATGAAGGTACCAACACCTAGCTATAAGCGCACCACCAGCTGGATCTTGACAGT -
1740
       -D  L  H  E  G  T  N  T  *  L  *  A  H  H  Q  L  D  L  D  S
       - T  F  M  K  V  P  T  P  S  Y  K  R  T  T  S  W  I  L  T  V
       -  P  S  *  R  Y  Q  H  L  A  I  S  A  P  P  A  G  S  *  Q  L
 1741 - TGATAGTAACATTAGGTGTGCATGTTTGAACCATAGTGTGCCATCTATGAAAAGGTAAAA -
1800
       -*  *  *  H  *  V  C  M  F  E  P  *  C  A  I  Y  E  K  V  K
       - D  S  N  I  R  C  A  C  L  N  H  S  V  P  S  M  K  R  *  N
       -  I  V  T  L  G  V  H  V  *  T  I  V  C  H  L  *  K  G  K  T
 1801 - CCTTTCCTAGAGCACAAAGCCAAGCAGTGCTATAAGTATTACCCCTAGTGTTGTACCTTA -
1860
       -P  F  L  E  H  K  A  K  Q  C  Y  K  Y  Y  P  *  C  C  T  L
       - L  S  *  S  T  K  P  S  S  A  I  S  I  T  P  S  V  V  P  Y
       -  F  P  R  A  Q  S  Q  A  V  L  *  V  L  P  L  V  L  Y  L  T
 1861 - CAAGGATCTTCAAGCACATGAGGTTTATTAGATGCACAGCGCTGTACTACAGTGCATATG -
1920
       -Q  G  S  S  S  T  *  G  L  L  D  A  Q  R  C  T  T  V  H  M
       - K  D  L  Q  A  H  E  V  Y  *  M  H  S  A  V  L  Q  C  I  C
       -  R  I  F  K  H  M  R  F  I  R  C  T  A  L  Y  Y  S  A  Y  A
 1921 - CAACTGCATAGAGAAATACAAGTCAAAACAATGAGAAGTTTCATGTTCGTTTAGACTTTG -
1980
       -Q  L  H  R  E  I  Q  V  K  T  M  R  S  F  M  F  V  *  T  L
       - N  C  I  E  K  Y  K  S  K  Q  *  E  V  S  C  S  F  R  L  W
       -  T  A  *  R  N  T  S  Q  N  N  E  K  F  H  V  R  L  D  F  G
 1981 - GTACAAGGTTCTTCTAGATCCTGGATTTCGAGTGAAAACCAAAATATAATAAGCATTATT -
2040
       -V  Q  G  S  S  R  S  W  I  S  S  E  N  Q  N  I  I  S  I  I
       - Y  K  V  L  L  D  P  G  F  R  V  K  T  K  I  *  *  A  L  L
       -  T  R  F  F  *  I  L  D  F  E  *  K  P  K  Y  N  K  H  Y  *
 2041 - AAAACAAGGAATAGCAGAAAGGCTAAAAAGCACAAATAGAAGTCAATTAAAGTGAGCTCA -
2100
       -K  T  R  N  S  R  K  A  K  K  H  K  *  K  S  I  K  V  S  S
```

```
     - K  Q  G  I  A  E  R  L  K  S  T  N  R  S  Q  L  K  *  A  H
     -  N  K  E  *  Q  K  G  *  K  A  Q  I  E  V  N  *  S  E  L  I
2101 - TTCATTCTGTCTTTCTCTTAATGGTGAAGCAAAGTATTAAAAATACTAGAGCAGCAACAA -
2160
     - F  I  L  S  F  S  *  W  *  S  K  V  L  K  I  L  E  Q  Q  Q
     -  S  F  C  L  S  L  N  G  E  A  K  Y  *  K  Y  *  S  S  N  N
     -  H  S  V  F  L  L  M  V  K  Q  S  I  K  N  T  R  A  A  T  M
2161 - TGAGAAAAAGTGGCGAGTAGAGCTCTTGTTGAACCTCCTCTTGTCTGATGAAAAGTTTTG -
2220
     - *  E  K  V  A  S  R  A  L  V  E  P  P  L  V  *  *  K  V  L
     -  E  K  K  W  R  V  E  L  L  L  N  L  L  L  S  D  E  K  F  W
     -  R  K  S  G  E  *  S  S  C  *  T  S  S  C  L  M  K  S  F  G
```

Figure 17-D

```
2221 - GTGAAACTGATCTTGCACGCAGCTGATAGGTATGTCGAGTACCGTCAGCACAAGCAAAAG -
2280
     - V  K  L  I  L  H  A  A  D  R  Y  V  E  Y  R  Q  H  K  Q  K
     -  *  N  *  S  C  T  Q  L  I  G  M  S  S  T  V  S  T  S  K  S
     -  E  T  D  L  A  R  S  *  *  V  C  R  V  P  S  A  Q  A  K  A
2281 - CAAAGTGTGTGCTAGTGCAAGTTAGTGCAAATTTATTGTCAGCAAGAGGGTGAAATGGTG -
2340
     - Q  S  V  C  *  C  K  L  V  Q  I  Y  C  Q  Q  E  G  E  M  V
     -  K  V  C  A  S  A  S  *  C  K  F  I  V  S  K  R  V  K  W  *
     -  K  C  V  L  V  Q  V  S  A  N  L  L  S  A  R  G  *  N  G  E
2341 - AATTGCCCTCGTATGTTCCTGATGGGCAAGGTTCTTTTAGTAGTACAGTCGTACCTCTAA -
2400
     - N  C  P  R  M  F  L  M  G  K  V  L  L  V  V  Q  S  Y  L  *
     -  I  A  L  V  C  S  *  W  A  R  F  F  *  *  Y  S  R  T  S  N
     -  L  P  S  Y  V  P  D  G  Q  G  S  F  S  S  T  V  V  P  L  T
2401 - CACACTCCTGATAGTGATATAGCTCGCAAGATGTAAATACAATCAATGTCAGGAAGAGAA -
2460
     - H  T  P  D  S  D  I  A  R  K  M  *  I  Q  S  M  S  G  R  E
     -  T  L  L  I  V  I  *  L  A  R  C  K  Y  N  Q  C  Q  E  E  N
     -  H  S  *  *  *  Y  S  S  Q  D  V  N  T  I  N  V  R  K  R  I
2461 - TAATTTTCATGTTCGTTTTATGGATAATCTAACTCCATAGGTTCTTCATCATCTAACTCC -
2520
     - *  F  S  C  S  F  Y  G  *  S  N  S  I  G  S  S  S  S  N  S
     -  N  F  H  V  R  F  M  D  N  L  T  P  *  V  L  H  H  L  T  P
     -  I  F  M  F  V  L  W  I  I  *  L  H  R  F  F  I  I  *  L  R
2521 - GAATAATTCTTCTTAGTTAGAGGCTTAAATAATTGTCTCACTATTGAACTTATTATAACG -
2580
     - E  *  F  F  L  V  R  G  L  N  N  C  L  T  I  E  L  I  I  T
     -  N  N  S  S  *  L  E  A  *  I  I  V  S  L  L  N  L  L  *  R
     -  I  I  L  L  S  *  R  L  K  *  L  S  H  Y  *  T  Y  Y  N  V
2581 - TCAAGATTCCAAATAGCAATCCTGAAAGTCCTCATAATGATAATCAATATCTCTGCTATT -
2640
     - S  R  F  Q  I  A  I  L  K  V  L  I  M  I  I  N  I  S  A  I
     -  Q  D  S  K  *  Q  S  *  K  S  S  *  *  *  S  I  S  L  L  L
     -  K  I  P  N  S  N  P  E  S  P  H  N  D  N  Q  Y  L  C  Y  C
2641 - GTAACCTGGAAGTCAACAAGATGAAACATCTGTTGTCACTTACTGTACTAGCAAAGCAAT -
2700
```

```
     - V  T  W  K  S  T  R  *  N  I  C  C  H  L  L  Y  *  Q  S  N
     -  *  P  G  S  Q  Q  D  E  T  S  V  V  T  Y  C  T  S  K  A  I
     -   N  L  E  V  N  K  M  K  H  L  L  S  L  T  V  L  A  K  Q  Y
2701 - ATTGTCGTTGCTACCGGCGTGGTCTGTATTTAATTTATAGTTTCCAATACGGTAGCGGTT -
2760
     - I  V  V  A  T  G  V  V  C  I  *  F  I  V  S  N  T  V  A  V
     -  L  S  L  L  P  A  W  S  V  F  N  L  *  F  P  I  R  *  R  L
     -   C  R  C  Y  R  R  G  L  Y  L  I  Y  S  F  Q  Y  G  S  G  C
2761 - GTATGCAGCAAAACCTGAATCAGTGCCTACACGCTGCGACGCTCCTAATTTGTAATAAGA -
2820
     - V  C  S  K  T  *  I  S  A  Y  T  L  R  R  S  *  F  V  I  R
     -  Y  A  A  K  P  E  S  V  P  T  R  C  D  A  P  N  L  *  *  E
     -   M  Q  Q  N  L  N  Q  C  L  H  A  A  T  L  L  I  C  N  K  K
2821 - AAGCGTTCGTGATGTAGCCACAGTGATCTCTTTTGGCAGGTCCTTAATGTCACAGCGCCC -
2880
     - K  R  S  *  C  S  H  S  D  L  F  W  Q  V  L  N  V  T  A  P
     -  S  V  R  D  V  A  T  V  I  S  F  G  R  S  L  M  S  Q  R  P
     -   A  F  V  M  *  P  Q  *  S  L  L  A  G  P  *  C  H  S  A  L
2881 - TAGGGAGTGTCCGGCCATTCGCAAGTGACCACGAATGATCACAGCACCAATGACAAGTTC -
2940
     - *  G  V  S  G  H  S  Q  V  T  T  N  D  H  S  T  N  D  K  F
     -  R  E  C  P  A  I  R  K  *  P  R  M  I  T  A  P  M  T  S  S
     -   G  S  V  R  P  F  A  S  D  H  E  *  S  Q  H  Q  *  Q  V  H
2941 - ACTTTCCATGAGCGGTCTGGTCACAATTGTCCCCCGGAGAGGCACATTGAGAAGAATGTT -
3000
     - T  F  H  E  R  S  G  H  N  C  P  P  E  R  H  I  E  K  N  V
     -  L  S  M  S  G  L  V  T  I  V  P  R  R  G  T  L  R  R  M  F
     -   F  P  *  A  V  W  S  Q  L  S  P  G  E  A  H  *  E  E  C  L
```

Figure 17-E

```
3001 - TGTTTCTGGGTTGAATGACCACATTGAGCGGGTACGAGCAAACAGCCTGAAGGAAGCAAC -
3060
       - C  F  W  V  E  *  P  H  *  A  G  T  S  K  Q  P  E  G  S  N
       -  V  S  G  L  N  D  H  I  E  R  V  R  A  N  S  L  K  E  A  T
       -   F  L  G  *  M  T  T  L  S  G  Y  E  Q  T  A  *  R  K  Q  R
3061 - GAAGTAGCTAAGCCACATCAAGCCTACAATACAAGCCATTGCAATCGCAATCCCGCCAGT -
3120
       - E  V  A  K  P  H  Q  A  Y  N  T  S  H  C  N  R  N  P  A  S
       -  K  *  L  S  H  I  K  P  T  I  Q  A  I  A  I  A  I  P  P  V
       -   S  S  *  A  T  S  S  L  Q  Y  K  P  L  Q  S  Q  S  R  Q  S
3121 - CACCCAATTAATTCTGTAGACAACAGCAAGCACAAAACAAGCAAGTGTTACTGGCCACAA -
3180
       - H  P  I  N  S  V  D  N  S  K  H  K  T  S  K  C  Y  W  P  Q
       -  T  Q  L  I  L  *  T  T  A  S  T  K  Q  A  S  V  T  G  H  K
       -   P  N  *  F  C  R  Q  Q  Q  A  Q  N  K  Q  V  L  L  A  T  R
3181 - GAGCCAGAGGAAAACAAGCTTTATTATGTACAAAAACCTGTTCCGATTAGAATAGGCAAA -
3240
       - E  P  E  E  N  K  L  Y  Y  V  Q  K  P  V  P  I  R  I  G  K
       -  S  Q  R  K  T  S  F  I  M  Y  K  N  L  F  R  L  E  *  A  N
       -   A  R  G  K  Q  A  L  L  C  T  K  T  C  S  D  *  N  R  Q  I
3241 - TTGTAGTAACATAATCCAGGCTAGGAATAGGAAACCTATTACTAGGTTCCATTGTTCCAG -
3300
       - L  *  *  H  N  P  G  *  E  *  E  T  Y  Y  *  V  P  L  F  Q
       -  C  S  N  I  I  Q  A  R  N  R  K  P  I  T  R  F  H  C  S  R
       -   V  V  T  *  S  R  L  G  I  G  N  L  L  L  G  S  I  V  P  G
3301 - GAGTTGTTTAAGCTCCTCAACGGTAATAGTACCGTTGTCTGCCATGATAAGCAATGTTAA -
3360
       - E  L  F  K  L  L  N  G  N  S  T  V  V  C  H  D  K  Q  C  *
       -  S  C  L  S  S  S  T  V  I  V  P  L  S  A  M  I  S  N  V  K
       -   V  V  *  A  P  Q  R  *  *  Y  R  C  L  P  *  *  A  M  L  K
3361 - AGTTCCAAACAGAATAATAATAATAGTTAGTTCGTTTAGACCAGAAGATCAGGAACTCCT -
3420
       - S  S  K  Q  N  N  N  N  S  *  F  V  *  T  R  R  S  G  T  P
       -  V  P  N  R  I  I  I  I  V  S  S  F  R  P  E  D  Q  E  L  L
       -   F  Q  T  E  *  *  *  *  L  V  R  L  D  Q  K  I  R  N  S  F
3421 - TCAGAAGAGTTCAGATTTTTAACACGCGAGTAGACGTAAACCGTTGGTTTACTAAACTC -
3480
       - S  E  E  F  R  F  L  T  R  E  *  T  *  T  V  G  F  T  K  L
       -  Q  K  S  S  D  F  *  H  A  S  R  R  K  P  L  V  L  L  N  S
       -   R  R  V  Q  I  F  N  T  R  V  D  V  N  R  W  F  Y  *  T  H
3481 - ACGTTAACAATATTGCAGCAGTACGCACACAATCGAAGCGCAGTAAGGATGGCTAGTGTG -
3540
       - T  L  T  I  L  Q  Q  Y  A  H  N  R  S  A  V  R  M  A  S  V
       -  R  *  Q  Y  C  S  S  T  H  T  I  E  A  Q  *  G  W  L  V  *
       -   V  N  N  I  A  A  V  R  T  Q  S  K  R  S  K  D  G  *  C  D
3541 - ACTAGCAAGAATACCACGAAAGCAAGAAAAAGAAGTACGCTATTAACTATTAACGTACCT -
3600
       - T  S  K  N  T  T  K  A  R  K  R  S  T  L  L  T  I  N  V  P
       -  L  A  R  I  P  R  K  Q  E  K  E  V  R  Y  *  L  L  T  Y  L
       -   *  Q  E  Y  H  E  S  K  K  K  K  Y  A  I  N  Y  *  R  T  C
3601 - GTTTCTTCCGAAACGAATGAGTACATAAGTTCGTACTCACTTTCTTGTGCTTACAAAGGC -
3660
       - V  S  S  E  T  N  E  Y  I  S  S  Y  S  L  S  C  A  Y  K  G
       -  F  L  P  K  R  M  S  T  *  V  R  T  H  F  L  V  L  T  K  A
       -   F  F  R  N  E  *  V  H  K  F  V  L  T  F  L  C  L  Q  R  H
3661 - ACGCTAGTAGTCGTCGTCGGCTCATCATAAATTGGATCCATTGCTGGATTAGCAACTCCT -
```

3720

- T L V V V V G S S * I G S I A G L A T P
- R * * S S S A H H K L D P L L D * Q L L
- A S S R R R L I I N W I H C W I S N S *

Figure 17-F

```
3721 - GAAGAGCCGTCGATTGTGTGTATTTGCACATTCGGTGGGTCTTTAACAAGCTTGTTAAAG -
3780
     -E  E  P  S  I  V  C  I  C  T  F  G  G  S  L  T  S  L  L  K
     -  K  S  R  R  L  C  V  F  A  H  S  V  G  L  *  Q  A  C  *  R
     -   R  A  V  D  C  V  Y  L  H  I  R  W  V  F  N  K  L  V  K  D
3781 - ATGAAGAATGTAGCATTTTCAATACCAGTGTCTGTAGTAATTTGTGTAGACTCAAGCTGG -
3840
     -M  K  N  V  A  F  S  I  P  V  S  V  V  I  C  V  D  S  S  W
     -  *  R  M  *  H  F  Q  Y  Q  C  L  *  *  F  V  *  T  Q  A  G
     -   E  E  C  S  I  F  N  T  S  V  C  S  N  L  C  R  L  K  L  V
3841 - TAGTAAACTTCGGTGAAATAGCCATGTACAACGACATAGTCTTTAACACCTGAGTGCCTA -
3900
     -*  *  T  S  V  K  *  P  C  T  T  T  *  S  L  T  P  E  C  L
     -  S  K  L  R  *  N  S  H  V  Q  R  H  S  L  *  H  L  S  A  Y
     -   V  N  F  G  E  I  A  M  Y  N  D  I  V  F  N  T  *  V  P  I
3901 - TCCTCAGAATAACCACCAATTTGGTAGTCTTCTTTGAGTTTTGGTGTTGAAATGCCGTCA -
3960
     -S  S  E  *  P  P  I  W  *  S  S  L  S  F  G  V  E  M  P  S
     -  P  Q  N  N  H  Q  F  G  S  L  L  *  V  L  V  L  K  C  R  H
     -   L  R  I  T  T  N  L  V  V  F  F  E  F  W  C  *  N  A  V  T
3961 - CCTTCAGTAACGACAATTGTATCTGTGACACTGTTATATGGTATACAGTAGTCATAGTTA -
4020
     -P  S  V  T  T  I  V  S  V  T  L  L  Y  G  I  Q  *  S  *  L
     -  L  Q  *  R  Q  L  Y  L  *  H  C  Y  M  V  Y  S  S  H  S  Y
     -   F  S  N  D  N  C  I  C  D  T  V  I  W  Y  T  V  V  I  V  M
4021 - TGTGTGTGCCAGCAAACAAAGTAGTTGGCATCATAAAGTAATGGGTTCTTGGATTTGCAC -
4080
     -C  V  C  Q  Q  T  K  *  L  A  S  *  S  N  G  F  L  D  L  H
     -  V  C  A  S  K  Q  S  S  W  H  H  K  V  M  G  S  W  I  C  T
     -   C  V  P  A  N  K  V  V  G  I  I  K  *  W  V  L  G  F  A  L
4081 - TTCCAACAAAGCCAACATCTCATAATAATTCTACATGCGTTGATGCATTGTAGAAAATAT -
4140
     -F  Q  Q  S  Q  H  L  I  I  I  L  H  A  L  M  H  C  R  K  Y
     -  S  N  K  A  N  I  S  *  *  F  Y  M  R  *  C  I  V  E  N  I
     -   P  T  K  P  T  S  H  N  N  S  T  C  V  D  A  L  *  K  I  Y
4141 - ATCAAGGCATAGAGGTACAAAAATTGCGCCTCCTTACCTGCAGCGACAAGCAAAAGATGT -
4200
     -I  K  A  *  R  Y  K  N  C  A  S  L  P  A  A  T  S  K  R  C
     -  S  R  H  R  G  T  K  I  A  P  P  Y  L  Q  R  Q  A  K  D  V
     -   Q  G  I  E  V  Q  K  L  R  L  L  T  C  S  D  K  Q  K  M  *
4201 - GAATAGATGGTAACAAATAGCAGCAGTAAATTGCAAATGAACTGGAAGCCCTTATAAAGG -
4260
     -E  *  M  V  T  N  S  S  S  K  L  Q  M  N  W  K  P  L  *  R
     -  N  R  W  *  Q  I  A  A  V  N  C  K  *  T  G  S  P  Y  K  G
     -   I  D  G  N  K  *  Q  Q  *  I  A  N  E  L  E  A  L  I  K  G
4261 - GCTAGCTGCCATCTTTTATTGAGCGCAATTATTTTGGTAGCGCTCTGAAAAACAGCAAGA -
4320
     -A  S  C  H  L  L  L  S  A  I  I  L  V  A  L  *  K  T  A  R
     -  L  A  A  I  F  Y  *  A  Q  L  F  W  *  R  S  E  K  Q  Q  E
     -   *  L  P  S  F  I  E  R  N  Y  F  G  S  A  L  K  N  S  K  K
4321 - AATGCAACGCCAATAACAAGCCATCCGAAAGGGAGTGAGGCTTGTAGCGGTATCGTTGCT -
4380
     -N  A  T  P  I  T  S  H  P  K  G  S  E  A  C  S  G  I  V  A
     -  M  Q  R  Q  *  Q  A  I  R  K  G  V  R  L  V  A  V  S  L  L
     -   C  N  A  N  N  K  P  S  E  R  E  *  G  L  *  R  Y  R  C  C
4381 - GTAGCATGAACAGTACTTGCAGGAGAAGCATTGTCAATTTTTACTGGCTGTGCAGTAATT -
```

```
4440
     -V  A  *  T  V  L  A  G  E  A  L  S  I  F  T  G  C  A  V  I
     - *  H  E  Q  Y  L  Q  E  K  H  C  Q  F  L  L  A  V  Q  *  L
     -  S  M  N  S  T  C  R  R  S  I  V  N  F  Y  W  L  C  S  N  *
 4441 - GATCCAAGAGTAAAAAATCTCATAAACAAATCCATAAGTTCGTTTATGTGTAATGTAATT -
4500
     -D  P  R  V  K  N  L  I  N  K  S  I  S  S  F  M  C  N  V  I
     - I  Q  E  *  K  I  S  *  T  N  P  *  V  R  L  C  V  M  *  F
     -  S  K  S  K  K  S  H  K  Q  I  H  K  F  V  Y  V  *  C  N  L
```

Figure 17-G

```
4501 - TGACACCCTTGAGAACTGGCTCAGAGTCATCCTCATCAAACTTGCAGCAAGAACCACAAG -
4560
     - * H P * E L A Q S H P H Q T C S K N H K
     - D T L E N W L R V I L I K L A A R T T R
     -  T P L R T G S E S S S S N L Q Q E P Q E
4561 - AGCATGCACCCTTGAGGCAACTGCAACAACTAGTCATGCAACAAAGCAAGATTGTAACCA -
4620
     - S M H P * G N C N N * S C N K A R L * P
     - A C T L E A T A T T S H A T K Q D C N H
     -  H A P L R Q L Q Q L V M Q Q S K I V T M
4621 - TGACGATGGCAATTAGTCCAGCAATGAAGCCGAGCCAAACATACCAAGGCCATTTAATAT -
4680
     - * R W Q L V Q Q * S R A K H T K A I * Y
     - D D G N * S S N E A E P N I P R P F N I
     -  T M A I S P A M K P S Q T Y Q G H L I Y
4681 - ATTGCTCATATTTTCCCAATTCTTGAAGGTCAATGAGTGATTCATTTAAATTTTTAGCGA -
4740
     - I A H I F P I L E G Q * V I H L N F * R
     - L L I F S Q F L K V N E * F I * I F S D
     -  C S Y F P N S * R S M S D S F K F L A T
4741 - CCTCATTGAGGCGGTCAATTTCTTTTTGAATGTTGACGACAGAAGCGTTAATGCCTGAAA -
4800
     - P H * G G Q F L F E C * R Q K R * C L K
     - L I E A V N F F L N V D D R S V N A * N
     -  S L R R S I S F * M L T T E A L M P E M
4801 - TGTCGCCAAGATCAACATCTGGTGATGTATGATTTTTGAAGTACTTGTCCAGCTCTTCTT -
4860
     - C R Q D Q H L V M Y D F * S T C P A L L
     - V A K I N I W * C M I F E V L V Q L F F
     -  S P R S T S G D V * F L K Y L S S S S L
4861 - TGAATGAGTCAAGCTCAGGTTGCAGAGGATCATAAACTGTGTTGTTAATGATGCCAATAA -
4920
     - * M S Q A Q V A E D H K L C C * * C Q *
     - E * V K L R L Q R I I N C V V N D A N N
     -  N E S S S G C R G S * T V L L M M P I T
4921 - CGACATCACAATTTCCTGAGACAAATGTATTGTCTGTAGTAATTATTTGTGGAGAAAAGA -
4980
     - R H H N F L R Q M Y C L * * L F V E K R
     - D I T I S * D K C I V C S N Y L W R K E
     -  T S Q F P E T N V L S V V I I C G E K K
4981 - AGTTCCTCTGTGTAATAAACCAAGAAGTGCCATTAAACACAAAAACACCTTCACGAGGGA -
5040
     - S S S V * * T K K C H * T Q K H L H E G
     - V P L C N K P R S A I K H K N T F T R E
     -  F L C V I N Q E V P L N T K T P S R G K
5041 - AGTATGCTTTGCCTTCATGACAAATTGCTGGCGCTGTGGTGAAGTTCCTCTCCTGGGATG -
5100
     - S M L C L H D K L L A L W * S S S P G M
     - V C F A F M T N C W R C G E V P L L G W
     -  Y A L P S * Q I A G A V V K F L S W D G
5101 - GCACATACGTGACATGTAGGAAGACAACACCATGCGGGGCTGCTTGTGGGAAGGACATAA -
5160
     - A H T * H V G R Q H H A G L L V G R T *
     - H I R D M * E D N T M R G C L W E G H K
     -  T Y V T C R K T T P C G A A C G K D I R
5161 - GGTGGTAGCCCTTTCCACAAAAGTCAACTCTTTTTGATTGTCCAAGAACACACTCAGACA -
```

5220
- G G S P F H K S Q L F L I V Q E H T Q T
- V V A L S T K V N S F * L S K N T L R H
- W * P F P Q K S T L F D C P R T H S D I

Figure 17-H

```
 5221 - TTTTAGTAGCAGCAAGATTAGCAGAAGCCCTGATTTCAGCAGCCCTGATTAGTTGTTGTG -
5280
      - F  *  *  Q  Q  D  *  Q  K  P  *  F  Q  Q  P  *  L  V  V  V
      -  F  S  S  S  K  I  S  R  S  P  D  F  S  S  P  D  *  L  L  C
      -   L  V  A  A  R  L  A  E  A  L  I  S  A  A  L  I  S  C  C  V
 5281 - TTACATAGGTTTGAAGGCTTTGAAGTCTGCCTGTAATTAACCTGTCAATTTGTACCTCCG -
5340
      - L  H  R  F  E  G  F  E  V  C  L  *  L  T  C  Q  F  V  P  P
      -  Y  I  G  L  K  A  L  K  S  A  C  N  *  P  V  N  L  Y  L  R
      -   T  *  V  *  R  L  *  S  L  P  V  I  N  L  S  I  C  T  S  A
 5341 - CCTCGACTTTATCAAGTCGCGAAAGGATATCATTTAGCACACTTGAAATTGCACCAAAAT -
5400
      - P  R  L  Y  Q  V  A  K  G  Y  H  L  A  H  L  K  L  H  Q  N
      -  L  D  F  I  K  S  R  K  D  I  I  *  H  T  *  N  C  T  K  I
      -   S  T  L  S  S  R  E  R  I  S  F  S  T  L  E  I  A  P  K  L
 5401 - TAGAGCTAAGTTGTTTAACAAGTGTGTTTAATGCTTGAGCATTCTGGTTAACAACGTCTT -
5460
      - *  S  *  V  V  *  Q  V  C  L  M  L  E  H  S  G  *  Q  R  L
      -  R  A  K  L  F  N  K  C  V  *  C  L  S  I  L  V  N  N  V  L
      -   E  L  S  C  L  T  S  V  F  N  A  *  A  F  W  L  T  T  S  C
 5461 - GCAGCTTGCCCAATGCAGTTGATGTTGTTGTAAGTGATTCTTGAATTTGACTAATCGCCT -
5520
      - A  A  C  P  M  Q  L  M  L  L  *  V  I  L  E  F  D  *  S  P
      -  Q  L  A  Q  C  S  *  C  C  C  K  *  F  L  N  L  T  N  R  L
      -   S  L  P  N  A  V  D  V  V  V  S  D  S  *  I  *  L  I  A  L
 5521 - TGTTAAATTGGTTGGCGATTTGTTTTTGGTTCTCATAGAGAACATTTTGGGTAACTCCAA -
5580
      - C  *  I  G  W  R  F  V  F  G  S  H  R  E  H  F  G  *  L  Q
      -  V  K  L  V  G  D  L  F  L  V  L  I  E  N  I  L  G  N  S  N
      -   L  N  W  L  A  I  C  F  W  F  S  *  R  T  F  W  V  T  P  M
 5581 - TGCCATTGAACCTATATGCCATTTGCATAGCAAAAGGTATTTGAAGAGCAGCGCCAGCAC -
5640
      - C  H  *  T  Y  M  P  F  A  *  Q  K  V  F  E  E  Q  R  Q  H
      -  A  I  E  P  I  C  H  L  H  S  K  R  Y  L  K  S  S  A  S  T
      -   P  L  N  L  Y  A  I  C  I  A  K  G  I  *  R  A  A  P  A  P
 5641 - CAAATGTCCATCCAGCAGTGGCAGTACCACTAACTAGAGCAGCAGTGTAGGCAGCAATCA -
5700
      - Q  M  S  I  Q  Q  W  Q  Y  H  *  L  E  Q  Q  C  R  Q  Q  S
      -  K  C  P  S  S  S  G  S  T  T  N  *  S  S  S  V  G  S  N  H
      -   N  V  H  P  A  V  A  V  P  L  T  R  A  A  V  *  A  A  I  I
 5701 - TATCATCAGTGAGCAGAGGTGGCAACACTGTAAGTCCATTGAACTTCTGCGCACAAATGA -
5760
      - Y  H  Q  *  A  E  V  A  T  L  *  V  H  *  T  S  A  H  K  *
      -  I  I  S  E  Q  R  W  Q  H  C  K  S  I  E  L  L  R  T  N  E
      -   S  S  V  S  R  G  G  N  T  V  S  P  L  N  F  C  A  Q  M  R
 5761 - GATCTCTAGCATTAATATCACCTAGGCATTCGCCATATTGCTTCATGAAGCCAGCATCAG -
5820
      - D  L  *  H  *  Y  H  L  G  I  R  H  I  A  S  *  S  Q  H  Q
      -  I  S  S  I  N  I  T  *  A  F  A  I  L  L  H  E  A  S  I  S
      -   S  L  A  L  I  S  P  R  H  S  P  Y  C  F  M  K  P  A  S  A
 5821 - CGAGTGTCACCTTATTAAAGAGCAAGTCCTCAATAAAAGACCTCTTAGTTGGCTTTAGAG -
5880
      - R  V  S  P  Y  *  R  A  S  P  Q  *  K  T  S  *  L  A  L  E
      -  E  C  H  L  I  K  E  Q  V  L  N  K  R  P  L  S  W  L  *  R
      -   S  V  T  L  L  K  S  K  S  S  I  K  D  L  L  V  G  F  R  G
 5881 - GGTCAGGTAATATTTGTGAAAAATTAAAACCACCAAAATATTTCAAAGTTGGGGTTTTGT -
```

```
5940
      - G  Q  V  I  F  V  K  N  *  N  H  Q  N  I  S  K  L  G  F  C
      -  V  R  *  Y  L  *  K  I  K  T  T  K  I  F  Q  S  W  G  F  V
      -   S  G  N  I  C  E  K  L  K  P  P  K  Y  F  K  V  G  V  L  Y
 5941 - ACATTTGTTTGACTTGAGCGAACACTTCACGTGTGTTGCGATCCTGTTCAGCAGCAATAC -
6000
      - T  F  V  *  L  E  R  T  L  H  V  C  C  D  P  V  Q  Q  Q  Y
      -  H  L  F  D  L  S  E  H  F  T  C  V  A  I  L  F  S  S  N  T
      -   I  C  L  T  *  A  N  T  S  R  V  L  R  S  C  S  A  A  I  P
```

Figure 17-I

```
6001 - CTGAGAGTGCACGATTTAGTTGTGTGCAAAAGCTACCATATTGGAGAAGCAAATTAGCAC - 6060
     - L  R  V  H  D  L  V  V  C  K  S  Y  H  I  G  E  A  N  *  H
     -  *  E  C  T  I  *  L  C  A  K  A  T  I  L  E  K  Q  I  S  T
     -   E  S  A  R  F  S  C  V  Q  K  L  P  Y  W  R  S  K  L  A  H
6061 - ATTCAGTAGAATCTCCGCAGATGTACATATTACAATCTACGGAGGTTTTAGCCATAGAAA - 6120
     - I  Q  *  N  L  R  R  C  T  Y  Y  N  L  R  R  F  *  P  *  K
     -  F  S  R  I  S  A  D  V  H  I  T  I  Y  G  G  F  S  H  R  N
     -   S  V  E  S  P  Q  M  Y  I  L  Q  S  T  E  V  L  A  I  E  T
6121 - CAGGCATTACTTCTGTAGTAATGCTAATTGAAAAGTTAGTAGGTATAGCAATGGTGTTAT - 6180
     - Q  A  L  L  L  *  *  C  *  L  K  S  *  *  V  *  Q  W  C  Y
     -  R  H  Y  F  C  S  N  A  N  *  K  V  S  R  Y  S  N  G  V  I
     -   G  I  T  S  V  V  M  L  I  E  K  L  V  G  I  A  M  V  L  L
6181 - TAGAGTAAGCAATTGAACTATCAGCACCTAAAGACATAGTATAAGCCACAATAGATTTTT - 6240
     - *  S  K  Q  L  N  Y  Q  H  L  K  T  *  Y  K  P  Q  *  I  F
     -  R  V  S  N  *  T  I  S  T  *  R  H  S  I  S  H  N  R  F  L
     -   E  *  A  I  E  L  S  A  P  K  D  I  V  *  A  T  I  D  F  W
6241 - GGCTAGTACTACGTAATAAAGAAACTGTATGGTAACTAGCACAAATGCCAGCTCCAATAG - 6300
     - G  *  Y  Y  V  I  K  K  L  Y  G  N  *  H  K  C  Q  L  Q  *
     -  A  S  T  T  *  *  R  N  C  M  V  T  S  T  N  A  S  S  N  R
     -   L  V  L  R  N  K  E  T  V  W  *  L  A  Q  M  P  A  P  I  G
6301 - GAATGTCGCACTCATAAGAAGTGTCGACATGCTCAGCTCCTATAAGACAGCCTGCTTGAG - 6360
     - E  C  R  T  H  K  K  C  R  H  A  Q  L  L  *  D  S  L  L  E
     -  N  V  A  L  I  R  S  V  D  M  L  S  S  Y  K  T  A  C  L  S
     -   M  S  H  S  *  E  V  S  T  C  S  A  P  I  R  Q  P  A  *  V
6361 - TCTGGAATACATTGTTTCCAGTAGAATATATGCGCCAAGCTGGTGTGAGTTGATCTGCAT - 6420
     - S  G  I  H  C  F  Q  *  N  I  C  A  K  L  V  *  V  D  L  H
     -  L  E  Y  I  V  S  S  R  I  Y  A  P  S  W  C  E  L  I  C  M
     -   W  N  T  L  F  P  V  E  Y  M  R  Q  A  G  V  S  *  S  A  *
6421 - GAATTGCTGTAGAAACATCAGTGCAGTTAACATCTTGATATAGAACAGCAACTTCAGATG - 6480
     - E  L  L  *  K  H  Q  C  S  *  H  L  D  I  E  Q  Q  L  Q  M
     -  N  C  C  R  N  I  S  A  V  N  I  L  I  *  N  S  N  F  R  *
     -   I  A  V  E  T  S  V  Q  L  T  S  *  Y  R  T  A  T  S  D  E
6481 - AAGCATTTGTTCCAGGTGTAATTACACTTACACCCCCAAAAGAGCAAGGTGAAATGTCTA - 6540
     - K  H  L  F  Q  V  *  L  H  L  H  P  Q  K  S  K  V  K  C  L
     -  S  I  C  S  R  C  N  Y  T  Y  T  P  K  R  A  R  *  N  V  *
     -   A  F  V  P  G  V  I  T  L  T  P  P  K  E  Q  G  E  M  S  N
6541 - ATATTTCAGATGTTTTAGGATCTCGAACGGAATCAGTGAAATCAGAAACATCACGGCCAA - 6600
     - I  F  Q  M  F  *  D  L  E  R  N  Q  *  N  Q  K  H  H  G  Q
     -  Y  F  R  C  F  R  I  S  N  G  I  S  E  I  R  N  I  T  A  K
     -   I  S  D  V  L  G  S  R  T  E  S  V  K  S  E  T  S  R  P  N
6601 - ATTGTTGAAATGGTTGAAATCTCTTTGAAGAAGGAGTTAACACACCAGTACCAGTGAGTC - 6660
     - I  V  E  M  V  E  I  S  L  K  K  E  L  T  H  Q  Y  Q  *  V
     -  L  L  K  W  L  K  S  L  *  R  R  S  *  H  T  S  T  S  E  S
     -   C  *  N  G  *  N  L  F  E  E  G  V  N  T  P  V  P  V  S  P
6661 - CATTAAAATTAAAATTGACACACTGGTTCTTAATAAGGTCAGTGGATAATTTTGGTCCAC -
```

6720

```
-H * N * N * H T G S * * G Q W I I L V H
- I K I K I D T L V L N K V S G * F W S T
-  L K L K L T H W F L I R S V D N F G P Q
```

Figure 17-J

```
6721 - AAACCGTGGCCGGTGCATTTAAAAGTTCAAAAGAAAGTACTACAACTCTGTAAGGTTGGT -
6780
     - K  P  W  P  V  H  L  K  V  Q  K  K  V  L  Q  L  C  K  V  G
     -  N  R  G  R  C  I  *  K  F  K  R  K  Y  Y  N  S  V  R  L  V
     -   T  V  A  G  A  F  K  S  S  K  E  S  T  T  T  L  *  G  W  *
6781 - AGCCAATGCCAGTAGTGGTGTAAAAACCATAATCATTTAATGGCCAATAACAATTAAGAG -
6840
     - S  Q  C  Q  *  W  C  K  N  H  N  H  L  M  A  N  N  N  *  E
     -  A  N  A  S  S  G  V  K  T  I  I  I  *  W  P  I  T  I  K  S
     -   P  M  P  V  V  V  *  K  P  *  S  F  N  G  Q  *  Q  L  R  A
6841 - CAGGTGGGGTGCAAGGTTTGCCATCAGGGGAGAAAGGCACATTAGATATGTCTCTCTCAA -
6900
     - Q  V  G  C  K  V  C  H  Q  G  R  K  A  H  *  I  C  L  S  Q
     -  R  W  G  A  R  F  A  I  R  G  E  R  H  I  R  Y  V  S  L  K
     -   G  G  V  Q  G  L  P  S  G  E  K  G  T  L  D  M  S  L  S  K
6901 - AGGGCCTAAGCTTGCCATGTCTAAGATACCTATATTTATAATTATAATTACCAGTTGAAG -
6960
     - R  A  *  A  C  H  V  *  D  T  Y  I  Y  N  Y  N  Y  Q  L  K
     -  G  P  K  L  A  M  S  K  I  P  I  F  I  I  I  I  T  S  *  S
     -   G  L  S  L  P  C  L  R  Y  L  Y  L  *  L  *  L  P  V  E  V
6961 - TAGCATCAATGTTCCTAGTATTCCAAGCAAGGACACAACCCATGAAATCATCTGGCAATT -
7020
     - *  H  Q  C  S  *  Y  S  K  Q  G  H  N  P  *  N  H  L  A  I
     -  S  I  N  V  P  S  I  P  S  K  D  T  T  H  E  I  I  W  Q  F
     -   A  S  M  F  L  V  F  Q  A  R  T  Q  P  M  K  S  S  G  N  L
7021 - TATAATTATAATCAGCAATAACACCAGTTTGTCCTGGCGCTATTTGTCTTACATCATCTC -
7080
     - Y  N  Y  N  Q  Q  *  H  Q  F  V  L  A  L  F  V  L  H  H  L
     -  I  I  I  I  S  N  N  T  S  L  S  W  R  Y  L  S  Y  I  I  S
     -   *  L  *  S  A  I  T  P  V  C  P  G  A  I  C  L  T  S  S  P
7081 - CCTTGACTACAAAAGAATCTGCATAGACATTGGAGAAGCAAAGATCATTCAACTTAGTGG -
7140
     - P  *  L  Q  K  N  L  H  R  H  W  R  S  K  D  H  S  T  *  W
     -  L  D  Y  K  R  I  C  I  D  I  G  E  A  K  I  I  Q  L  S  G
     -   L  T  T  K  E  S  A  *  T  L  E  K  Q  R  S  F  N  L  V  A
7141 - CAGAAACGCCATAGCACTTAAAGGTTGAAAAAAATGTTGAGTTGTAGAGCACAGAGTAAT -
7200
     - Q  K  R  H  S  T  *  R  L  K  K  M  L  S  C  R  A  Q  S  N
     -  R  N  A  I  A  L  K  G  *  K  K  C  *  V  V  E  H  R  V  I
     -   E  T  P  *  H  L  K  V  E  K  N  V  E  L  *  S  T  E  *  S
7201 - CAGCAACACAATTAGAAATTTTTTTCTCTCCCATGCATAGACAGAAGGGAATTTAGTAG -
7260
     - Q  Q  H  N  *  K  F  F  F  S  P  M  H  R  Q  K  G  I  *  *
     -  S  N  T  I  R  N  F  F  S  L  P  C  I  D  R  R  E  F  S  S
     -   A  T  Q  L  E  I  F  F  L  S  H  A  *  T  E  G  N  L  V  A
7261 - CATTAAAAACCTCTCCAAAAGGACACAAGTTTGTAATATTAGGGAATCTCACAACATCTC -
7320
     - H  *  K  P  L  Q  K  D  T  S  L  *  Y  *  G  I  S  Q  H  L
     -  I  K  N  L  S  K  R  T  Q  V  C  N  I  R  E  S  H  N  I  S
     -   L  K  T  S  P  K  G  H  K  F  V  I  L  G  N  L  T  T  S  P
7321 - CTGAGGGAACAACCCTGAAATTAGAGGTCTGGTAAATTCCTTTGTCAATCTCAAAGCTCT -
7380
     - L  R  E  Q  P  *  N  *  R  S  G  K  F  L  C  Q  S  Q  S  S
     -  *  G  N  N  P  E  I  R  G  L  V  N  S  F  V  N  L  K  A  L
     -   E  G  T  T  L  K  L  E  V  W  *  I  P  L  S  I  S  K  L  L
7381 - TAACAGAGCATTTGAGTTCAGCAAGTGGATTTTGAGAACAATCAACAGCATCTGTGATTG -
7440
     - *  Q  S  I  *  V  Q  Q  V  D  F  E  N  N  Q  Q  H  L  *  L
```

```
     - N  R  A  F  E  F  S  K  W  I  L  R  T  I  N  S  I  C  D  C
     -  T  E  H  L  S  S  A  S  G  F  *  E  Q  S  T  A  S  V  I  V
7441 - TACCATTTTCATCATACTTGAGCATAAATGTAGTTGGCTTTAAATAGCCAACAAAATAGG -
7500
     - Y  H  F  H  H  T  *  A  *  M  *  L  A  L  N  S  Q  Q  N  R
     - T  I  F  I  I  L  E  H  K  C  S  W  L  *  I  A  N  K  I  G
     -  P  F  S  S  Y  L  S  I  N  V  V  G  F  K  *  P  T  K  *  A
```

Figure 17-K

```
7501 - CTGCAGCTGACGTGCCCCAAATGTCTTGAGCAGGTGAAAAGGCTGTAAGAATGGCTCTAA -
7560
     - L  Q  L  T  C  P  K  C  L  E  Q  V  K  R  L  *  E  W  L  *
     -  C  S  *  R  A  P  N  V  L  S  R  *  K  G  C  K  N  G  S  K
     -   A  A  D  V  P  Q  M  S  *  A  G  E  K  A  V  R  M  A  L  K
7561 - AATTTGTAATGTTAATACCAAGAGGCAACTTAAAAATAGGTTTCAAAGTGTTAAAACCAG -
7620
     - N  L  *  C  *  Y  Q  E  A  T  *  K  *  V  S  K  C  *  N  Q
     -  I  C  N  V  N  T  K  R  Q  L  K  N  R  F  Q  S  V  K  T  R
     -   F  V  M  L  I  P  R  G  N  L  K  I  G  F  K  V  L  K  P  E
7621 - AAGGTAGATCACGAACTACATCTATAGGTTGATAGCCCTTATAAACATAGAGAAACCCAT -
7680
     - K  V  D  H  E  L  H  L  *  V  D  S  P  Y  K  H  R  E  T  H
     -  R  *  I  T  N  Y  I  Y  R  L  I  A  L  I  N  I  E  K  P  I
     -   G  R  S  R  T  T  S  I  G  *  *  P  L  *  T  *  R  N  P  S
7681 - CTTTATTTTTAAACACAAACTCTCGTAAGTGTTTAAAATTACCTGACTTTTCTGAAACAT -
7740
     - L  Y  F  *  T  Q  T  L  V  S  V  *  N  Y  L  T  F  L  K  H
     -  F  I  F  K  H  K  L  S  *  V  F  K  I  T  *  L  F  *  N  I
     -   L  F  L  N  T  N  S  R  K  C  L  K  L  P  D  F  S  E  T  S
7741 - CAAGCGAAAAGGCATCAGATATGTACTCGAAAGTGCAATTAAATGCATTATCGAATATCA -
7800
     - Q  A  K  R  H  Q  I  C  T  R  K  C  N  *  M  H  Y  R  I  S
     -  K  R  K  G  I  R  Y  V  L  E  S  A  I  K  C  I  I  E  Y  H
     -   S  E  K  A  S  D  M  Y  S  K  V  Q  L  N  A  L  S  N  I  I
7801 - TAGTATGTGTCTGTGTACCCATGGGTTTAGAAACAGCAAAGAAAGGGTTGTCACACAATT -
7860
     - *  Y  V  S  V  Y  P  W  V  *  K  Q  Q  R  K  G  C  H  T  I
     -  S  M  C  L  C  T  H  G  F  R  N  S  K  E  R  V  V  T  Q  F
     -   V  C  V  C  V  P  M  G  L  E  T  A  K  K  G  L  S  H  N  S
7861 - CAAAGTTACATGCTCGTATAACAACATTAGTAGAATTGTTAATAATAATCACCGACTGTG -
7920
     - Q  S  Y  M  L  V  *  Q  H  *  *  N  C  *  *  *  S  P  T  V
     -  K  V  T  C  S  Y  N  N  I  S  R  I  V  N  N  N  H  R  L  *
     -   K  L  H  A  R  I  T  T  L  V  E  L  L  I  I  I  T  D  C  D
7921 - ACTTGTTGTTCATGGTAGAACCAAAAACCCAACCACGGACAACATTTGATTTCTCTGTGG -
7980
     - T  C  C  S  W  *  N  Q  K  P  N  H  G  Q  H  L  I  S  L  W
     -  L  V  V  H  G  R  T  K  N  P  T  T  D  N  I  *  F  L  C  G
     -   L  L  F  M  V  E  P  K  T  Q  P  R  T  T  F  D  F  S  V  A
7981 - CAGCAAAATAAATACCATCCTTAAAAGGTATGACAGGGTTGCCAAACGTATGATTAATAG -
8040
     - Q  Q  N  K  Y  H  P  *  K  V  *  Q  G  C  Q  T  Y  D  *  *
     -  S  K  I  N  T  I  L  K  R  Y  D  R  V  A  K  R  M  I  N  S
     -   A  K  *  I  P  S  L  K  G  M  T  G  L  P  N  V  *  L  I  V
8041 - TATGAAACCCTGTAACATTAGAATAAAATGGAAGAAATAAATCCTGAGTTAAATAAAGAG -
8100
     - Y  E  T  L  *  H  *  N  K  M  E  E  I  N  P  E  L  N  K  E
     -  M  K  P  C  N  I  R  I  K  W  K  K  *  I  L  S  *  I  K  S
     -   *  N  P  V  T  L  E  *  N  G  R  N  K  S  *  V  K  *  R  V
8101 - TGTCTGATCTAAAAATTTCATCAGGATAGTAAACCCCCCTCATAGATGAAGTATGTTGAG -
8160
     - C  L  I  *  K  F  H  Q  D  S  K  P  P  S  *  M  K  Y  V  E
     -  V  *  S  K  N  F  I  R  I  V  N  P  P  H  R  *  S  M  L  S
     -   S  D  L  K  I  S  S  G  *  *  T  P  L  I  D  E  V  C  *  V
8161 - TGTAATTAGGAGCTTGAACATCATCAAAAGTGGTGCACCGGTCAAGGTCACTACCACTAG -
```

8220
- C N * E L E H H Q K W C T G Q G H Y H *
- V I R S L N I I K S G A P V K V T T T S
- * L G A * T S S K V V H R S R S L P L V

Figure 17-L

```
8221 - TGAGAGTAAGAAATAATAAGAAAATAAACATGTTCGTTTAGTTGTTAACAAGAATATCAC -
8280
     -* E * E I I R K * T C S F S C * Q E Y H
     - E S K K * * E N K H V R L V V N K N I T
     -  R V R N N K K I N M F V * L L T R I S L
8281 - TTGAAACCACAACTCTGTTGTTTTCTCTAATGATAAGCCTACCTTTTTCCAGAAGAGAAT -
8340
     -L K P Q L C C F L * * * A Y L F P E E N
     - * N H N S V V F S N D K P T F F Q K R I
     -  E T T T L L F S L M I S L P F S R R E *
8341 - AAATCATATCATTGATTTGATTCTCCTTAAGAGACATTACAGCAGTTCCTCTTAATTTAA -
8400
     -K S Y H * F D S P * E T L Q Q F L L I *
     - N H I I D L I L L K R H Y S S S S * F K
     -  I I S L I * F S L R D I T A V P L N L R
8401 - GAGGAAATTTGCTCATGTCAAAGAGTGAATAGGAAGACAACTGGATAGGATTTGTGTTCC -
8460
     -E E I C S C Q R V N R K T T G * D L C S
     - R K F A H V K E * I G R Q L D R I C V P
     -  G N L L M S K S E * E D N W I G F V F L
8461 - TCCAGAAAATGTAGTTAGCATGCATGGTATAGCCATCAATTTGTTCCTTCGGCTTGCCAA -
8520
     -S R K C S * H A W Y S H Q F V P S A C Q
     - P E N V V S M H G I A I N L F L R L A K
     -  Q K M * L A C M V * P S I C S F G L P R
8521 - GATAGTTAGCCCCAATTAAAAATGCTTCCGATGATGATGCATTTACATTTGTAACAAAAG -
8580
     -D S * P Q L K M L P M M M H L H L * Q K
     - I V S P N * K C F R * * C I Y I C N K S
     -  * L A P I K N A S D D D A F T F V T K A
8581 - CTGTCCACCATGAGAAATGGCCCATAAGCTTGTAAAGGTCAGCATTCCAAGAATGCTCTG -
8640
     -L S T M R N G P * A C K G Q H S K N A L
     - C P P * E M A H K L V K V S I P R M L C
     -  V H H E K W P I S L * R S A F Q E C S V
8641 - TTATCTTTACAGCTATAGAACCACCCAGGGCTAGTTTTTGCTTTATAAATCCACACAGAT -
8700
     -L S L Q L * N H P G L V F A L * I H T D
     - Y L Y S Y R T T Q G * F L L Y K S T Q I
     -  I F T A I E P P R A S F C F I N P H R *
8701 - AAGTGAAAAACCCTTCTTTAGAGTCATTCTCTTTTGTCACATGTTTGGTCCTAGGGTCAT -
8760
     -K * K T L L * S H S L L S H V W S * G H
     - S E K P F F R V I L F C H M F G P R V I
     -  V K N P S L E S F S F V T C L V L G S Y
8761 - ACATATCGCTAATAATAAGGTCCCATTTATTAGCCGTATGTACTGTTGCACAGTCTCCAA -
8820
     -T Y R * * * G P I Y * P Y V L L H S L Q
     - H I A N N K V P F I S R M Y C C T V S N
     -  I S L I I R S H L L A V C T V A Q S P I
8821 - TTAAAGTAGAATCTGCGTCGGAGACGAAGTCATTAAGATCTGAATCGACAAGTAGTGTGC -
8880
     -L K * N L R R R R S H * D L N R Q V V C
     - * S R I C V G D E V I K I * I D K * C A
     -  K V E S A S E T K S L R S E S T S S V P
8881 - CAGTTGGCAACCATTGTCTGAGCACAGCTGTACCTGGTGCAACTCCTTTATCAGAGCCAG -
8940
```

```
      - Q  L  A  T  I  V  *  A  Q  L  Y  L  V  Q  L  L  Y  Q  S  Q
      -  S  W  Q  P  L  S  E  H  S  C  T  W  C  N  S  F  I  R  A  S
      -   V  G  N  H  C  L  S  T  A  V  P  G  A  T  P  L  S  E  P  A
 8941 - CACCAAAGTGAATAACTCTCATGTTGTAGGGTACAGCTAAAGTAAGTGTATTTAAGTATT -
9000
      - H  Q  S  E  *  L  S  C  C  R  V  Q  L  K  *  V  Y  L  S  I
      -  T  K  V  N  N  S  H  V  V  G  Y  S  *  S  K  C  I  *  V  L
      -   P  K  *  I  T  L  M  L  *  G  T  A  K  V  S  V  F  K  Y  *
```

Figure 17-M

```
 9001 - GACACAGTTGAGTATACTTTGCGACATTCATCATTATTCCTTTTGGTATAACAGCATTTT -
9060
      -D  T  V  E  Y  T  L  R  H  S  S  L  F  L  L  V  *  Q  H  F
      - T  Q  L  S  I  L  C  D  I  H  H  Y  S  F  W  Y  N  S  I  F
      -  H  S  *  V  Y  F  A  T  F  I  I  I  P  F  G  I  T  A  F  S
 9061 - CACCATAATTCTGAAGGTCACACTTTTCAAGAAGCATTCTTTGCATCTTGTACAAGTTAG -
9120
      -H  H  N  S  E  G  H  T  F  Q  E  A  F  F  A  S  C  T  S  *
      - T  I  I  L  K  V  T  L  F  K  K  H  S  L  H  L  V  Q  V  R
      -  P  *  F  *  R  S  H  F  S  R  S  I  L  C  I  L  Y  K  L  G
 9121 - GCATCGCAACACCTGGTTGCCACGCTTGACTTGCTTGTAGTTTTGGGTAGAAGGTTTCAA -
9180
      -A  S  Q  H  L  V  A  T  L  D  L  L  V  V  L  G  R  R  F  Q
      - H  R  N  T  W  L  P  R  L  T  C  L  *  F  W  V  E  G  F  N
      -  I  A  T  P  G  C  H  A  *  L  A  C  S  F  G  *  K  V  S  T
 9181 - CATGTCCATCCTTACACCAAAGCATGAATGAAATTTCAGCATAGTCAATTGTAACCTTGA -
9240
      -H  V  H  P  Y  T  K  A  *  M  K  F  Q  H  S  Q  L  *  P  *
      - M  S  I  L  T  P  K  H  E  *  N  F  S  I  V  N  C  N  L  D
      -  C  P  S  L  H  Q  S  M  N  E  I  S  A  *  S  I  V  T  L  T
 9241 - CCACTTTTGAAATCACTGACAAATCTTGTGACTTTATTATCTCGACAAAGTCATCAAGTA -
9300
      -P  L  L  K  S  L  T  N  L  V  T  L  L  S  R  Q  S  H  Q  V
      - H  F  *  N  H  *  Q  I  L  *  L  Y  Y  L  D  K  V  I  K  *
      -  T  F  E  I  T  D  K  S  C  D  F  I  I  S  T  K  S  S  S  K
 9301 - AAAGATCAATCACAGAACACACACATTTGATGAACCTGTTTGCGCATCTGTTATGAAGT -
9360
      -K  D  Q  S  Q  N  T  H  I  L  M  N  L  F  A  H  L  L  *  S
      - K  I  N  H  R  T  H  T  F  *  *  T  C  L  R  I  C  Y  E  V
      -  R  S  I  T  E  H  T  H  F  D  E  P  V  C  A  S  V  M  K  *
 9361 - AATTTTTCACTGTGCTGTCCATAGGGATAAAATCCTCTAATTTAAGTGGTGAATCTTGTG -
9420
      -N  F  S  L  C  C  P  *  G  *  N  P  L  I  *  V  V  N  L  V
      - I  F  H  C  A  V  H  R  D  K  I  L  *  F  K  W  *  I  L  *
      -  F  F  T  V  L  S  I  G  I  K  S  S  N  L  S  G  E  S  C  E
 9421 - AGCGCTTGGCTAAGCCTATCATTAAATGAAGACCGCCAAGTTGTCCATGACTGAAATCTC -
9480
      -S  A  W  L  S  L  S  L  N  E  D  R  Q  V  V  H  D  *  N  L
      - A  L  G  *  A  Y  H  *  M  K  T  A  K  L  S  M  T  E  I  S
      -  R  L  A  K  P  I  I  K  *  R  P  P  S  C  P  *  L  K  S  P
 9481 - CATAAACGATGTGTTCGAAGGCATAGCCCTCGAGCTTATATCGCTGTATGAATTCATCCA -
9540
      -H  K  R  C  V  R  R  H  S  P  R  A  Y  I  A  V  *  I  H  P
      - I  N  D  V  F  E  G  I  A  L  E  L  I  S  L  Y  E  F  I  H
      -  *  T  M  C  S  K  A  *  P  S  S  L  Y  R  C  M  N  S  S  I
 9541 - TAGCGAGCTCGAGAAAGTCAGTTTCCATTTGTGATCTGGGCTTAAAATCCTCTAAGTCTC -
9600
      -*  R  A  R  E  S  Q  F  P  F  V  I  W  A  *  N  P  L  S  L
      - S  E  L  E  K  V  S  F  H  L  *  S  G  L  K  I  L  *  V  S
      -  A  S  S  R  K  S  V  S  I  C  D  L  G  L  K  S  S  K  S  L
 9601 - TGCTCTGAGTAAAGTAGGTTTCAGGCAACTGTTGAATAATGCCGTCTACTTTCTTAAAGT -
9660
      -C  S  E  *  S  R  F  Q  A  T  V  E  *  C  R  L  L  S  *  S
      - A  L  S  K  V  G  F  R  Q  L  L  N  N  A  V  Y  F  L  K  V
      -  L  *  V  K  *  V  S  G  N  C  *  I  M  P  S  T  F  L  K  *
 9661 - AGTTAAACTGTGTTTTTACTGATTCTCCAATTAATGTGACTCCATTGACGCTAGCTTGTG -
9720
```

```
     - S  *  T  V  F  L  L  I  L  Q  L  M  *  L  H  *  R  *  L  V
     -  V  K  L  C  F  Y  *  F  S  N  *  C  D  S  I  D  A  S  L  C
     -   L  N  C  V  F  T  D  S  P  I  N  V  T  P  L  T  L  A  C  A
9721 - CTGGTCCCTTTGAAGGTGTTAGACCTTTGACTGAACCTTCTGTTATTAAAACACCATTAC -
9780
     - L  V  P  L  K  V  L  D  L  *  L  N  L  L  L  L  K  H  H  Y
     -  W  S  L  *  R  C  *  T  F  D  *  T  F  C  Y  *  N  T  I  T
     -   G  P  F  E  G  V  R  P  L  T  E  P  S  V  I  K  T  P  L  R
```

Figure 17-N

```
 9781 - GGGCGTTTCTAAAAAGGTCTACCTGTCCTTCCACTCTACCATCAAACAAGACAGTAAGTG -
 9840
      - G  R  F  *  K  G  L  P  V  L  P  L  Y  H  Q  T  R  Q  *  V
      -  G  V  S  K  K  V  Y  L  S  F  H  S  T  I  K  Q  D  S  K  *
      -   A  F  L  K  R  S  T  C  P  S  T  L  P  S  N  K  T  V  S  E
 9841 - AAGAACAAGCACTCTCAGTAGGTTTCTTGGCAATGTCAGTCATTGTGCAGACACCTATTG -
 9900
      - K  N  K  H  S  Q  *  V  S  W  Q  C  Q  S  L  C  R  H  L  L
      -  R  T  S  T  L  S  R  F  L  G  N  V  S  H  C  A  D  T  Y  C
      -   E  Q  A  L  S  V  G  F  L  A  M  S  V  I  V  Q  T  P  I  V
 9901 - TAGATACATGTGCTGGGGCTTCTCTTTTGTAGTCCCAGATTACAGTATTAGCAGCGATAT -
 9960
      - *  I  H  V  L  G  L  L  F  C  S  P  R  L  Q  Y  *  Q  R  Y
      -  R  Y  M  C  W  G  F  S  F  V  V  P  D  Y  S  I  S  S  D  I
      -   D  T  C  A  G  A  S  L  L  *  S  Q  I  T  V  L  A  A  I  S
 9961 - CAACACCCAAATTATTGAGTATCTTAATCTCTGGCACTGGTTTAATGTTACGCTTAGCCC -
10020
      - Q  H  P  N  Y  *  V  S  *  S  L  A  L  V  *  C  Y  A  *  P
      -  N  T  Q  I  I  E  Y  L  N  L  W  H  W  F  N  V  T  L  S  P
      -   T  P  K  L  L  S  I  L  I  S  G  T  G  L  M  L  R  L  A  Q
10021 - AAAGCTCAAATGCAACATTAACAGGAAGTGTTGTCTTATTTTCAAAGATCTCCACATCAA -
10080
      - K  A  Q  M  Q  H  *  Q  E  V  L  S  Y  F  Q  R  S  P  H  Q
      -  K  L  K  C  N  I  N  R  K  C  C  L  I  F  K  D  L  H  I  N
      -   S  S  N  A  T  L  T  G  S  V  V  L  F  S  K  I  S  T  S  I
10081 - TACCATCTACCTTTGTGTAAACAGCATTATTAATGATGGAAACAGGTGCTTCGCCGGCGT -
10140
      - Y  H  L  P  L  C  K  Q  H  Y  *  *  W  K  Q  V  L  R  R  R
      -  T  I  Y  L  C  V  N  S  I  I  N  D  G  N  R  C  F  A  G  V
      -   P  S  T  F  V  *  T  A  L  L  M  M  E  T  G  A  S  P  A  C
10141 - GTCCATCAAAGTGTCCTTTATTAACAACATTATAAGCCACATTTTCTAAACTCTGTAACC -
10200
      - V  H  Q  S  V  L  Y  *  Q  H  Y  K  P  H  F  L  N  S  V  T
      -  S  I  K  V  S  F  I  N  N  I  I  S  H  I  F  *  T  L  *  P
      -   P  S  K  C  P  L  L  T  T  L  *  A  T  F  S  K  L  C  N  L
10201 - TGGTAAATGTATTCCACAGGTTATAAGTATCAAATTGTTTGTAAATCCATAGGCTAAATC -
10260
      - W  *  M  Y  S  T  G  Y  K  Y  Q  I  V  C  K  S  I  G  *  I
      -  G  K  C  I  P  Q  V  I  S  I  K  L  F  V  N  P  *  A  K  S
      -   V  N  V  F  H  R  L  *  V  S  N  C  L  *  I  H  R  L  N  P
10261 - CAGCAGAAATCATCATATTATATGCATCCAAGTACTGTCGGTACTCATTTGCATGGTGTC -
10320
      - Q  Q  K  S  S  Y  Y  M  H  P  S  T  V  G  T  H  L  H  G  V
      -  S  R  N  H  H  I  I  C  I  Q  V  L  S  V  L  I  C  M  V  S
      -   A  E  I  I  I  L  Y  A  S  K  Y  C  R  Y  S  F  A  W  C  L
10321 - TGCAAACAGCACCACCTAAATTGCATCGTGTAATACACGTAGCAGATTTGAGTGGAACAT -
10380
      - C  K  Q  H  H  L  N  C  I  V  *  Y  T  *  Q  I  *  V  E  H
      -  A  N  S  T  T  *  I  A  S  C  N  T  R  S  R  F  E  W  N  I
      -   Q  T  A  P  P  K  L  H  R  V  I  H  V  A  D  L  S  G  T  *
10381 - AATCAATATCCGACACTACTTGTTTGCCATGAGACTCACAAGGACTATCAGAATAGTAAA -
10440
      - N  Q  Y  P  T  L  L  V  C  H  E  T  H  K  D  Y  Q  N  S  K
      -  I  N  I  R  H  Y  L  F  A  M  R  L  T  R  T  I  R  I  V  K
      -   S  I  S  D  T  T  C  L  P  *  D  S  Q  G  L  S  E  *  *  K
10441 - AGAAAGGCAATTGCTTTAAATTAGTAAATGCACTTTTATCGAAAGCTGGAGTGTGGAATG -
10500
```

```
      -R  K  A  I  A  L  N  *  *  M  H  F  Y  R  K  L  E  C  G  M
      -  E  R  Q  L  L  *  I  S  K  C  T  F  I  E  S  W  S  V  E  C
      -   K  G  N  C  F  K  L  V  N  A  L  L  S  K  A  G  V  W  N  A
10501 - CATGCTTATTCACATACAAACTACCACCATCACAGCCTGGTAAGTTCAAGTTTGACAAGA -
10560
      -H  A  Y  S  H  T  N  Y  H  H  H  S  L  V  S  S  S  L  T  R
      -  M  L  I  H  I  Q  T  T  T  I  T  A  W  *  V  Q  V  *  Q  D
      -   C  L  F  T  Y  K  L  P  P  S  Q  P  G  K  F  K  F  D  K  T
```

Figure 17-O

```
10561 - CTCTTGTGTCAAACCTACACACAATTGCATTGGCTGGGTAACGATCAACGTTACAATTCC -
10620
      - L  L  C  Q  T  Y  T  Q  L  H  W  L  G  N  D  Q  R  Y  N  S
      -  S  C  V  K  P  T  H  N  C  I  G  W  V  T  I  N  V  T  I  P
      -   L  V  S  N  L  H  T  I  A  L  A  G  *  R  S  T  L  Q  F  Q
10621 - AAAACAAACAAACACCATCAGTGAATTTATCGTGATGTGTAGCATAAGAATAGAAGAGTT -
10680
      - K  T  N  K  H  H  Q  *  I  Y  R  D  V  *  H  K  N  R  R  V
      -  K  Q  T  N  T  I  S  E  F  I  V  M  C  S  I  R  I  E  E  F
      -   N  K  Q  T  P  S  V  N  L  S  *  C  V  A  *  E  *  K  S  S
10681 - CCTCTATTTTGTAAGCTTTGTCACTACATGGCTGAGCATCGTAGAACTTCCATTCTACTT -
10740
      - P  L  F  C  K  L  C  H  Y  M  A  E  H  R  R  T  S  I  L  L
      -  L  Y  F  V  S  F  V  T  T  W  L  S  I  V  E  L  P  F  Y  F
      -   S  I  L  *  A  L  S  L  H  G  *  A  S  *  N  F  H  S  T  S
10741 - CAGCCTGAGGCACACACTTGATAGCCTTTGGATTTCCAATGTCATGAAGAACTGGAAACT -
10800
      - Q  P  E  A  H  T  *  *  P  L  D  F  Q  C  H  E  E  L  E  T
      -  S  L  R  H  T  L  D  S  L  W  I  S  N  V  M  K  N  W  K  L
      -   A  *  G  T  H  L  I  A  F  G  F  P  M  S  *  R  T  G  N  L
10801 - TATCAGCAAGCAATGCAGACTTCACAACCATGTGTTGTACTTTTCTGCAAGCAGAATTAA -
10860
      - Y  Q  Q  A  M  Q  T  S  Q  P  C  V  V  L  F  C  K  Q  N  *
      -  I  S  K  Q  C  R  L  H  N  H  V  L  Y  F  S  A  S  R  I  N
      -   S  A  S  N  A  D  F  T  T  M  C  C  T  F  L  Q  A  E  L  T
10861 - CCCTCAGTTCATCTCCTATAATAGGGTATTCAACAGACCAATCAACGCGCTTAACAAAGC -
10920
      - P  S  V  H  L  L  *  *  G  I  Q  Q  T  N  Q  R  A  *  Q  S
      -  P  Q  F  I  S  Y  N  R  V  F  N  R  P  I  N  A  L  N  K  A
      -   L  S  S  S  P  I  I  G  Y  S  T  D  Q  S  T  R  L  T  K  H
10921 - ACTCATGGACTGCTAAACATCTAGTCATGATAGCATCACAACTAGCCACATGTGCATTTC -
10980
      - T  H  G  L  L  N  I  *  S  *  *  H  H  N  *  P  H  V  H  F
      -  L  M  D  C  *  T  S  S  H  D  S  I  T  T  S  H  M  C  I  S
      -   S  W  T  A  K  H  L  V  M  I  A  S  Q  L  A  T  C  A  F  P
10981 - CATGTACCTGGCAATGTTGGTCATGGTTACTCTGAAGGTTACCCGTAAAGCCCCACTGCT -
11040
      - H  V  P  G  N  V  G  H  G  Y  S  E  G  Y  P  *  S  P  T  A
      -  M  Y  L  A  M  L  V  M  V  T  L  K  V  T  R  K  A  P  L  L
      -   C  T  W  Q  C  W  S  W  L  L  *  R  L  P  V  K  P  H  C  *
11041 - GAACATCAATCATAAATGGGTTATAGACATAGTCAAAACCCACAGAATGATTCCAGCAGG -
11100
      - E  H  Q  S  *  M  G  Y  R  H  S  Q  N  P  Q  N  D  S  S  R
      -  N  I  N  H  K  W  V  I  D  I  V  K  T  H  R  M  I  P  A  G
      -   T  S  I  I  N  G  L  *  T  *  S  K  P  T  E  *  F  Q  Q  A
11101 - CATAAGTATCTGATGAAGTAGAAAAGCAAGTTGCACGTTTGTCACACAGACAACACGTTC -
11160
      - H  K  Y  L  M  K  *  K  S  K  L  H  V  C  H  T  D  N  T  F
      -  I  S  I  *  *  S  R  K  A  S  C  T  F  V  T  Q  T  T  R  S
      -   *  V  S  D  E  V  E  K  Q  V  A  R  L  S  H  R  Q  H  V  L
11161 - TTTCAGGTCCAATCTTGACAAAGTACTTCATTGATGTAAGCTCAAAGCCATGCGCCCAAA -
11220
      - F  Q  V  Q  S  *  Q  S  T  S  L  M  *  A  Q  S  H  A  P  K
      -  F  R  S  N  L  D  K  V  L  H  *  C  K  L  K  A  M  R  P  K
      -   S  G  P  I  L  T  K  Y  F  I  D  V  S  S  K  P  C  A  Q  R
11221 - GGACGAACACGACTCTGTCTGACAATCCTTTCAGTGTATCACTGAGCATTTGTACTATCT -
```

11280
- G R T R L C L T I L S V Y H * A F V L S
- D E H D S V * Q S F Q C I T E H L Y Y L
- T N T T L S D N P F S V S L S I C T I L

Figure 17-P

```
11281 - TAATACGCACTACATTCCAGGGCAAGCCTTTATACATGAGTGGTATAAGATGTTTAAACT -
11340
      - * Y A L H S R A S L Y T * V V * D V * T
      - N T H Y I P G Q A F I H E W Y K M F K L
      -  I R T T F Q G K P L Y M S G I R C L N W
11341 - GGTCACCTGGTGGAGGTTTTGCATTAACTCTGGTGAATTCTGTGTTATTTCAGTGTCAA -
11400
      - G H L V E V L H * L W * I L C Y F Q C Q
      - V T W W R F C I N S G E F C V I F S V N
      -  S P G G G F A L T L V N S V L F S V S T
11401 - CATAACCAGTCGGTACAGCTACTAAGTTAACACCTGTAGAAAATCCTAGCTGGAGAGGTA -
11460
      - H N Q S V Q L L S * H L * K I L A G E V
      - I T S R Y S Y * V N T C R K S * L E R *
      -  * P V G T A T K L T P V E N P S W R G R
11461 - GGTTAGTACCCACAGCATCTCTAGTTGCATGACAGCCCTCTACATCAAAGCCAATCCACG -
11520
      - G * Y P Q H L * L H D S P L H Q S Q S T
      - V S T H S I S S C M T A L Y I K A N P R
      -  L V P T A S L V A * Q P S T S K P I H A
11521 - CACGAACGTGACGAATAGCTTCTTCGCGGGTGATAAACATATTAGGGTAACCATTGACTT -
11580
      - H E R D E * L L R G * * T Y * G N H * L
      - T N V T N S F F A G D K H I R V T I D L
      -  R T * R I A S S R V I N I L G * P L T W
11581 - GGTAATTCATTTTGAAACCCATCATAGAGATGAGTCTACGGTAGGTCATGTCCTTTGGTA -
11640
      - G N S F * N P S * R * V Y G R S C P L V
      - V I H F E T H H R D E S T V G H V L W Y
      -  * F I L K P I I E M S L R * V M S F G M
11641 - TGCCTGGTATGTCAACACATAATCCTTCAGTCTTGAATTTTATATCAACGCTGAGGTGTG -
11700
      - C L V C Q H I I L Q S * I L Y Q R * G V
      - A W Y V N T * S F S L E F Y I N A E V C
      -  P G M S T H N P S V L N F I S T L R C V
11701 - TAGGTGCCTGTGTAGGATGAAGACCAGTAATGATCTTACTACAGTCCTTAAAAAGTCCAG -
11760
      - * V P V * D E D Q * * S Y Y S P * K V Q
      - R C L C R M K T S N D L T T V L K K S S
      -  G A C V G * R P V M I L L Q S L K S P V
11761 - TTACATTTTCTGCTTGTAATGTAGCCACATTGCGACGTGGTATTTCTAGACTTGTAAATT -
11820
      - L H F L L V M * P H C D V V F L D L * I
      - Y I F C L * C S H I A T W Y F * T C K L
      -  T F S A C N V A T L R R G I S R L V N C
11821 - GCAGTTTGTCATAAAGATCTCTATCAGACATTATGCACAAAATGCCAATTTTTGCCCTTG -
11880
      - A V C H K D L Y Q T L C T K C Q F L P L
      - Q F V I K I S I R H Y A Q N A N F C P C
      -  S L S * R S L S D I M H K M P I F A L V
11881 - TGATAGCCACATTGAAGCGGTTGACATTACAAGAGTGTGCTGTTTCAGTAGTTTGTGTGA -
11940
      - * * P H * S G * H Y K S V L F Q * F V *
      - D S H I E A V D I T R V C C F S S L C E
      -  I A T L K R L T L Q E C A V S V V C V N
11941 - ATATGACATAGTCATATTCAGAACCCTGTGATGAATCAACAGTCTGCGTAGGCAATCCTA -
```

```
12000
      - I  *  H  S  H  I  Q  N  P  V  M  N  Q  Q  S  A  *  A  I  L
      -  Y  D  I  V  I  F  R  T  L  *  *  I  N  S  L  R  R  Q  S  *
      -  M  T  *  S  Y  S  E  P  C  D  E  S  T  V  C  V  G  N  P  K
```

Figure 17-Q

```
12001 - AGATTTTTGAAGCTACAGCGTTCTGTGAATTATAAGGTGAGATAAAAACAGCTTTTCTCC -
12060
      -R  F  L  K  L  Q  R  S  V  N  Y  K  V  R  *  K  Q  L  F  S
      - D  F  *  S  Y  S  V  L  *  I  I  R  *  D  K  N  S  F  S  P
      -  I  F  E  A  T  A  F  C  E  L  *  G  E  I  K  T  A  F  L  Q
12061 - AAGCAGGATTGCGTGTAAGAAATTCTCTTACAACGCCTATTTGAGGTCTGTTGATTGCAG -
12120
      -K  Q  D  C  V  *  E  I  L  L  Q  R  L  F  E  V  C  *  L  Q
      - S  R  I  A  C  K  K  F  S  Y  N  A  Y  L  R  S  V  D  C  R
      -  A  G  L  R  V  R  N  S  L  T  T  P  I  *  G  L  L  I  A  D
12121 - ATGAAACATCATGTGTAATAACACCTTTGTAGAACATTTTGAAGCATTGAGCTGACTTAT -
12180
      -M  K  H  H  V  *  *  H  L  C  R  T  F  *  S  I  E  L  T  Y
      - *  N  I  M  C  N  N  T  F  V  E  H  F  E  A  L  S  *  L  I
      -  E  T  S  C  V  I  T  P  L  *  N  I  L  K  H  *  A  D  L  S
12181 - CCTTGTGTGCTTTTAGCTTATTGTCATAAACTAAAGCACTCACAGTGTCAACAATTTCAG -
12240
      -P  C  V  L  L  A  Y  C  H  K  L  K  H  S  Q  C  Q  Q  F  Q
      - L  V  C  F  *  L  I  V  I  N  *  S  T  H  S  V  N  N  F  S
      -  L  C  A  F  S  L  L  S  *  T  K  A  L  T  V  S  T  I  S  A
12241 - CAGGACAACGGCGACAAGTTCCAAGGAACATGTCTGGACCTATTGTTTTCATAAGTCTGC -
12300
      -Q  D  N  G  D  K  F  Q  G  T  C  L  D  L  L  F  S  *  V  C
      - R  T  T  A  T  S  S  K  E  H  V  W  T  Y  C  F  H  K  S  A
      -  G  Q  R  R  Q  V  P  R  N  M  S  G  P  I  V  F  I  S  L  H
12301 - ACACTGAATTAAAATATTCTGGTTCTAGTGTGCCTTTAGTCAGCAATGTGCGGGGGGCTG -
12360
      -T  L  N  *  N  I  L  V  L  V  C  L  *  S  A  M  C  G  G  L
      - H  *  I  K  I  F  W  F  *  C  A  F  S  Q  Q  C  A  G  G  W
      -  T  E  L  K  Y  S  G  S  S  V  P  L  V  S  N  V  R  G  A  G
12361 - GTAATTGAGCAGGATCGCCAATATAGACGTAGTGTTTTGCACGAAGTCTAGCATTGACAA -
12420
      -V  I  E  Q  D  R  Q  Y  R  R  S  V  L  H  E  V  *  H  *  Q
      - *  L  S  R  I  A  N  I  D  V  V  F  C  T  K  S  S  I  D  N
      -  N  *  A  G  S  P  I  *  T  *  C  F  A  R  S  L  A  L  T  T
12421 - CACTCAAGTCATAATTAGTAGCCATAGAGATTTCATCAAAGACTACAATGTCAGCAGTTG -
12480
      -H  S  S  H  N  *  *  P  *  R  F  H  Q  R  L  Q  C  Q  Q  L
      - T  Q  V  I  I  S  S  H  R  D  F  I  K  D  Y  N  V  S  S  C
      -  L  K  S  *  L  V  A  I  E  I  S  S  K  T  T  M  S  A  V  V
12481 - TTTCTGGCAATGCATTTACAGTGCAGAAAACATACTGTTCTAGTGTTGAATTCACTTTGA -
12540
      -F  L  A  M  H  L  Q  C  R  K  H  T  V  L  V  L  N  S  L  *
      - F  W  Q  C  I  Y  S  A  E  N  I  L  F  *  C  *  I  H  F  E
      -  S  G  N  A  F  T  V  Q  K  T  Y  C  S  S  V  E  F  T  L  N
12541 - ATTTATCAAAACACTCTACGCGCGCACGCGCAGGTATGATTCTACTACATTTATCTATGG -
12600
      -I  Y  Q  N  T  L  R  A  H  A  Q  V  *  F  Y  Y  I  Y  L  W
      - F  I  K  T  L  Y  A  R  T  R  R  Y  D  S  T  T  F  I  Y  G
      -  L  S  K  H  S  T  R  A  R  A  G  M  I  L  L  H  L  S  M  G
12601 - GCAAATATTTTAATGCCTTTTCACATAGGGCATCAACAGCTGCATGAGAGCATGCCGTAT -
12660
      -A  N  I  L  M  P  F  H  I  G  H  Q  Q  L  H  E  S  M  P  Y
      - Q  I  F  *  C  L  F  T  *  G  I  N  S  C  M  R  A  C  R  I
      -  K  Y  F  N  A  F  S  H  R  A  S  T  A  A  *  E  H  A  V  Y
12661 - ACACTATGCGAGCAGATGGGTAATAGAGAGCAAGTCCGATGGCAAAATGACTCTTACCAG -
```

12720

- T L C E Q M G N R E Q V R W Q N D S Y Q
- H Y A S R W V I E S K S D G K M T L T S
- T M R A D G * * R A S P M A K * L L P V

Figure 17-R

```
12721 - TACCAGGTGGTCCTTGGAGTGTAGAGTACTTTTGCATGCCGACCTTTTGATAATTTGCAA -
12780
      -Y  Q  V  V  L  G  V  *  S  T  F  A  C  R  P  F  D  N  L  Q
      - T  R  W  S  L  E  C  R  V  L  L  H  A  D  L  L  I  I  C  N
      -  P  G  G  P  W  S  V  E  Y  F  C  M  P  T  F  *  *  F  A  T
12781 - CATTGCTAGAAAACTCATCTGAGATGTTGAGTGTTGGGTACAAGCCAGTAATTCTCACAT -
12840
      -H  C  *  K  T  H  L  R  C  *  V  L  G  T  S  Q  *  F  S  H
      - I  A  R  K  L  I  *  D  V  E  C  W  V  Q  A  S  N  S  H  I
      -  L  L  E  N  S  S  E  M  L  S  V  G  Y  K  P  V  I  L  T  *
12841 - AGTGCTCTTGTGGCACTAGAGTAGGTGCACTAAGTGGCATTACAGTGTGAGATGTCAACA -
12900
      -S  A  L  V  A  L  E  *  V  H  *  V  A  L  Q  C  E  M  S  T
      - V  L  L  W  H  *  S  R  C  T  K  W  H  Y  S  V  R  C  Q  H
      -  C  S  C  G  T  R  V  G  A  L  S  G  I  T  V  *  D  V  N  T
12901 - CAAAGTAATCACCAACATTCAACTTGTATGTCGTAGTACCTCTGTACACAACAGCATCAC -
12960
      -Q  S  N  H  Q  H  S  T  C  M  S  *  Y  L  C  T  Q  Q  H  H
      - K  V  I  T  N  I  Q  L  V  C  R  S  T  S  V  H  N  S  I  T
      -  K  *  S  P  T  F  N  L  Y  V  V  V  P  L  Y  T  T  A  S  P
12961 - CATAGTCACCTTTTTCAAAGGTGTACTCTCCAATCTGTACTTTACTATTTTTAGTTACAC -
13020
      -H  S  H  L  F  Q  R  C  T  L  Q  S  V  L  Y  Y  F  *  L  H
      - I  V  T  F  F  K  G  V  L  S  N  L  Y  F  T  I  F  S  Y  T
      -  *  S  P  F  S  K  V  Y  S  P  I  C  T  L  L  F  L  V  T  R
13021 - GGTAACCAGTAAAGACATAGTTTCTGTTCAATGGTGGTCTAGGTTTTCCAACCTCCCATG -
13080
      -G  N  Q  *  R  H  S  F  C  S  M  V  V  *  V  F  Q  P  P  M
      - V  T  S  K  D  I  V  S  V  Q  W  W  S  R  F  S  N  L  P  *
      -  *  P  V  K  T  *  F  L  F  N  G  G  L  G  F  P  T  S  H  E
13081 - AAAGATGCAATTCTCTGTCAGAGAGTACTTCGCGTACAGTGGCAATACCATATGACAGCT -
13140
      -K  D  A  I  L  C  Q  R  V  L  R  V  Q  W  Q  Y  H  M  T  A
      - K  M  Q  F  S  V  R  E  Y  F  A  Y  S  G  N  T  I  *  Q  L
      -  R  C  N  S  L  S  E  S  T  S  R  T  V  A  I  P  Y  D  S  L
13141 - TAAATGTTTCCTCAGTGGCTTTGAGCGTTTCTGCTGCGAAAAGCTTGAGTCTCTCAGTAC -
13200
      -*  M  F  P  Q  W  L  *  A  F  L  L  R  K  A  *  V  S  Q  Y
      - K  C  F  L  S  G  F  E  R  F  C  C  E  K  L  E  S  L  S  T
      -  N  V  S  S  V  A  L  S  V  S  A  A  K  S  L  S  L  S  V  Q
13201 - AAGTGTTGGCAAGTATGTAATCGCCAGCATTAGTCCAATCACATGTTGCTATCGCATTGA -
13260
      -K  C  W  Q  V  C  N  R  Q  H  *  S  N  H  M  L  L  S  H  *
      - S  V  G  K  Y  V  I  A  S  I  S  P  I  T  C  C  Y  R  I  E
      -  V  L  A  S  M  *  S  P  A  L  V  Q  S  H  V  A  I  A  L  K
13261 - AGTCAGTGACATTGTCACTGCCTACACATGTGTTTTTGTATAAACCAAAAACCTGACCAT -
13320
      -S  Q  *  H  C  H  C  L  H  M  C  F  C  I  N  Q  K  P  D  H
      - V  S  D  I  V  T  A  Y  T  C  V  F  V  *  T  K  N  L  T  I
      -  S  V  T  L  S  L  P  T  H  V  F  L  Y  K  P  K  T  *  P  L
13321 - TAGCACATAATGGAAAACTAATGGGAGGCTTATGTGACTTGCAATAATAGCTCATACCTC -
13380
      -*  H  I  M  E  N  *  W  E  A  Y  V  T  C  N  N  S  S  Y  L
      - S  T  *  W  K  T  N  G  R  L  M  *  L  A  I  I  A  H  T  S
      -  A  H  N  G  K  L  M  G  G  L  C  D  L  Q  *  *  L  I  P  P
13381 - CTAGATACAGTTGTGTCACATCAGTGACATCACAACCTGGGGCATTGCAAACATAGGGAT -
```

13440

```
-L D T V V S H Q * H H N L G H C K H R D
- * I Q L C H I S D I T T W G I A N I G I
-  R Y S C V T S V T S Q P G A L Q T * G L
```

Figure 17-S

```
13441 - TAACAGACAACACTAATTTGTGTGATGTTGAAATGACATGGTCATAGCAGCACTTGCAAC -
13500
      -* Q T T L I C V M L K * H G H S S T C N
      - N R Q H * F V * C * N D M V I A A L A T
      -  T D N T N L C D V E M T W S * Q H L Q H
13501 - ATAGGAATGGTCTCCTAATACAGGCACCGCAACGAAGTGAAGTCTGTGAATTGCACAATA -
13560
      - I G M V S * Y R H R N E V K S V N C T I
      -  * E W S P N T G T A T K * S L * I A Q Y
      -  R N G L L I Q A P Q R S E V C E L H N T
13561 - CACAAGCACCTACAGCCTGCAAGACTGTATGTGGTGTGTACATAGCCTCATAAAACTCAG -
13620
      - H K H L Q P A R L Y V V C T * P H K T Q
      -  T S T Y S L Q D C M W C V H S L I K L R
      -  Q A P T A C K T V C G V Y I A S * N S G
13621 - GTTCCCAGTACCGTGAGGTGTTATCATTAGTTAGCATTACGGAATACATGTCCAACATGT -
13680
      - V P S T V R C Y H * L A L R N T C P T C
      -  F P V P * G V I I S * H Y G I H V Q H V
      -  S Q Y R E V L S L V S I T E Y M S N M W
13681 - GGCCAGTAAGCTCATCATGTAACTTTCTAATGTATTGTAAATACAAGTGAAAGACATCAG -
13740
      - G Q * A H H V T F * C I V N T S E R H Q
      -  A S K L I M * L S N V L * I Q V K D I S
      -  P V S S S C N F L M Y C K Y K * K T S A
13741 - CATACTCCTGATTAGGATGTTTTGTAAGTGGGTAAGCATCAATAGCCAGTGACACGAACC -
13800
      - H T P D * D V L * V G K H Q * P V T R T
      -  I L L I R M F C K W V S I N S Q * H E P
      -  Y S * L G C F V S G * A S I A S D T N L
13801 - TTTCAATCATAAGTGTACCATCTGTTTTGACAATATCATCGACAAAACAGCCTGCGCCTA -
13860
      - F Q S * V Y H L F * Q Y H R Q N S L R L
      -  F N H K C T I C F D N I I D K T A C A *
      -  S I I S V P S V L T I S S T K Q P A P N
13861 - ATATTCTTGATGGATCTGGGTAAGGCAGGTACACGTAATCATCTCCTTGTTTAACTAGCA -
13920
      - I F L M D L G K A G T R N H L L V * L A
      -  Y S * W I W V R Q V H V I I S L F N * H
      -  I L D G S G * G R Y T * S S P C L T S I
13921 - TTGTATGCTGTGAGCAAAATTCGTGAGGTCCTTTAGTAAGGTCAGTCTCAGTCCAACATT -
13980
      - L Y A V S K I R E V L * * G Q S Q S N I
      -  C M L * A K F V R S F S K V S L S P T F
      -  V C C E Q N S * G P L V R S V S V Q H F
13981 - TTGCCTCAGACATGAACACATTATTTTGATAATAAAGAACTGCCTTAAAGTTCTTAATGC -
14040
      - L P Q T * T H Y F D N K E L P * S S * C
      -  C L R H E H I I L I I K N C L K V L N A
      -  A S D M N T L F * * * R T A L K F L M L
14041 - TAGCTACTAAACCTTGAGCCGCATAGTTACTGTTATAGCACACAACGGCATCATCAGAAA -
14100
      - * L L N L E P H S Y C Y S T Q R H H Q K
      -  S Y * T L S R I V T V I A H N G I I R K
      -  A T K P * A A * L L L * H T T A S S E R
14101 - GAATCATCATGGAGAAATGTTTACGCAGGTAAGCGTAAAACTCATCCACGAATTCATGAT -
14160
```

```
      - E  S  S  W  R  N  V  Y  A  G  K  R  K  T  H  P  R  I  H  D
      -  N  H  H  G  E  M  F  T  Q  V  S  V  K  L  I  H  E  F  M  I
      -   I  I  M  E  K  C  L  R  R  *  A  *  N  S  S  T  N  S  *  S
14161 - CAACATCCCTATTTCTATAGAGACACTCATAGAGCCTGTGTTGTAGATTGCGGACATACT -
14220
      - Q  H  P  Y  F  Y  R  D  T  H  R  A  C  V  V  D  C  G  H  T
      -  N  I  P  I  S  I  E  T  L  I  E  P  V  L  *  I  A  D  I  L
      -   T  S  L  F  L  *  R  H  S  *  S  L  C  C  R  L  R  T  Y  L
```

Figure 17-T

```
14221 - TGTCAGCTATCTTATTACCATCAGTTGAAAGAAGTGCATTTACATTGGCTGTAACAGCTT -
14280
      - C  Q  L  S  Y  Y  H  Q  L  K  E  V  H  L  H  W  L  *  Q  L
      -  V  S  Y  L  I  T  I  S  *  K  K  C  I  Y  I  G  C  N  S  L
      -   S  A  I  L  L  P  S  V  E  R  S  A  F  T  L  A  V  T  A  *
14281 - GACAAATGTTAAAGACACTATTAGCATAAGCAGTTGTAGCATCACCGGATGATGTTCCAC -
14340
      - D  K  C  *  R  H  Y  *  H  K  Q  L  *  H  H  R  M  M  F  H
      -  T  N  V  K  D  T  I  S  I  S  S  C  S  I  T  G  *  C  S  T
      -   Q  M  L  K  T  L  L  A  *  A  V  V  A  S  P  D  D  V  P  P
14341 - CTGGTTTAACATATAGTGAGCCGCCACACATGACCATCTCACTTAATACTTGCGCACACT -
14400
      - L  V  *  H  I  V  S  R  H  T  *  P  S  H  L  I  L  A  H  T
      -  W  F  N  I  *  *  A  A  T  H  D  H  L  T  *  Y  L  R  T  L
      -   G  L  T  Y  S  E  P  P  H  M  T  I  S  L  N  T  C  A  H  S
14401 - CGTTAGCTAACCTGTAGAAACGGTGTGATAAGTTACAGCAAGTGTTATGTTTGCGAGCAA -
14460
      - R  *  L  T  C  R  N  G  V  I  S  Y  S  K  C  Y  V  C  E  Q
      -  V  S  *  P  V  E  T  V  *  *  V  T  A  S  V  M  F  A  S  K
      -   L  A  N  L  *  K  R  C  D  K  L  Q  Q  V  L  C  L  R  A  R
14461 - GAACAAGAGAGGCCATTATCCTAAGCATGTTAGGCATGGCTCTGTCACATTTTGGATAAT -
14520
      - E  Q  E  R  P  L  S  *  A  C  *  A  W  L  C  H  I  L  D  N
      -  N  K  R  G  H  Y  P  K  H  V  R  H  G  S  V  T  F  W  I  I
      -   T  R  E  A  I  I  L  S  M  L  G  M  A  L  S  H  F  G  *  S
14521 - CCCAACCCATAAGGTGTGGAGTTTCTACATCACTGTAAACAGTTTTTAACATATTATGCC -
14580
      - P  N  P  *  G  V  E  F  L  H  H  C  K  Q  F  L  T  Y  Y  A
      -  P  T  H  K  V  W  S  F  Y  I  T  V  N  S  F  *  H  I  M  P
      -   Q  P  I  R  C  G  V  S  T  S  L  *  T  V  F  N  I  L  C  Q
14581 - AGCCACCGTAAAACTTGCTTGTTCCAATTACCACAGTAGCTCCTCTAGTGGCGGCTATTG -
14640
      - S  H  R  K  T  C  L  F  Q  L  P  Q  *  L  L  *  W  R  L  L
      -  A  T  V  K  L  A  C  S  N  Y  H  S  S  S  S  S  G  G  Y  *
      -   P  P  *  N  L  L  V  P  I  T  T  V  A  P  L  V  A  A  I  D
14641 - ACTTCAATAATTTCTGATGAAACTGTCTATTTGTCATAGTACTACAGATAGAGACACCAG -
14700
      - T  S  I  I  S  D  E  T  V  Y  L  S  *  Y  Y  R  *  R  H  Q
      -  L  Q  *  F  L  M  K  L  S  I  C  H  S  T  T  D  R  D  T  S
      -   F  N  N  F  *  *  N  C  L  F  V  I  V  L  Q  I  E  T  P  A
14701 - CTACGGTGCGAGCTCTATTCTTTGCACTAATGGCATACTTAAGATTCATTTGAGTTATAG -
14760
      - L  R  C  E  L  Y  S  L  H  *  W  H  T  *  D  S  F  E  L  *
      -  Y  G  A  S  S  I  L  C  T  N  G  I  L  K  I  H  L  S  Y  S
      -   T  V  R  A  L  F  F  A  L  M  A  Y  L  R  F  I  *  V  I  V
14761 - TAGGGATGACATTACGCTTAGTATACGCGAAAAGTGCATCTTGATCCTCATAACTCATTG -
14820
      - *  G  *  H  Y  A  *  Y  T  R  K  V  H  L  D  P  H  N  S  L
      -  R  D  D  I  T  L  S  I  R  E  K  C  I  L  I  L  I  T  H  *
      -   G  M  T  L  R  L  V  Y  A  K  S  A  S  *  S  S  *  L  I  E
14821 - AGTCATAATAAAGTCTAGCCTTACCCCATTTATTAAATGGGAAACCAGCTGATTTATCCA -
14880
      - S  H  N  K  V  *  P  Y  P  I  Y  *  M  G  N  Q  L  I  Y  P
      -  V  I  I  K  S  S  L  T  P  F  I  K  W  E  T  S  *  F  I  Q
      -   S  *  *  S  L  A  L  P  H  L  L  N  G  K  P  A  D  L  S  R
14881 - GATTGTTAACGATTACTTGGTTGGCATTAATACAGCCACCATCGTAACAATCAAAGTATT -
```

14940 .
-D C * R L L G W H * Y S H H R N N Q S I
- I V N D Y L V G I N T A T I V T I K V F
- L L T I T W L A L I Q P P S * Q S K Y L

Figure 17-U

```
14941 - TATCAACAACTTCAACTACGAATAGGAGTTGTCTGATATCACACATTGTTGGCAGATTAT -
15000
      - Y  Q  Q  L  Q  L  R  I  G  V  V  *  Y  H  T  L  L  A  D  Y
      -  I  N  N  F  N  Y  E  *  E  L  S  D  I  T  H  C  W  Q  I  I
      -   S  T  T  S  T  T  N  R  S  C  L  I  S  H  I  V  G  R  L  *
15001 - AACGATAATAGTCATAATCACTGATAGCAGCGTTGCCATCCTGAGCAAAGAAGAAGTGTT -
15060
      - N  D  N  S  H  N  H  *  *  Q  R  C  H  P  E  Q  R  R  S  V
      -  T  I  I  V  I  I  T  D  S  S  V  A  I  L  S  K  E  E  V  F
      -   R  *  *  S  *  S  L  I  A  A  L  P  S  *  A  K  K  K  C  F
15061 - TTAGTTCAACAGAACTTCCTTCCTTAAAGAAACCTTTAGACACAGCAAAGTCATAAAAGT -
15120
      - L  V  Q  Q  N  F  L  P  *  R  N  L  *  T  Q  Q  S  H  K  S
      -  *  F  N  R  T  S  F  L  K  E  T  F  R  H  S  K  V  I  K  V
      -   S  S  T  E  L  P  S  L  K  K  P  L  D  T  A  K  S  *  K  S
15121 - CTTTATTAAAATTACCGGGTTTGACAGTTTGAAAAGCAACATTGTTTGTTAGTGCAGCTA -
15180
      - L  Y  *  N  Y  R  V  *  Q  F  E  K  Q  H  C  L  L  V  Q  L
      -  F  I  K  I  T  G  F  D  S  L  K  S  N  I  V  C  *  C  S  Y
      -   L  L  K  L  P  G  L  T  V  *  K  A  T  L  F  V  S  A  A  T
15181 - CTGAAAAGCATGTAGTGCGTTTATCTAGCAATAAATTGCCAGAAGCTGCATGCATAGCTG -
15240
      - L  K  S  M  *  C  V  Y  L  A  I  N  C  Q  K  L  H  A  *  L
      -  *  K  A  C  S  A  F  I  *  Q  *  I  A  R  S  C  M  H  S  W
      -   E  K  H  V  V  R  L  S  S  N  K  L  P  E  A  A  C  I  A  G
15241 - GATCAGCAGCATACACTAAAAGTTCCTTGAAACTGAGACGCGAGCTATGTAAGTTTACAT -
15300
      - D  Q  Q  H  T  L  K  V  P  *  N  *  D  A  S  Y  V  S  L  H
      -  I  S  S  I  H  *  K  F  L  E  T  E  T  R  A  M  *  V  Y  I
      -   S  A  A  Y  T  K  S  S  L  K  L  R  R  E  L  C  K  F  T  S
15301 - CCTGATTATGTACGACTCCTAACTCACGAAAATGGTATCCAGTTGAAACAACAAAAGGAA -
15360
      - P  D  Y  V  R  L  L  T  H  E  N  G  I  Q  L  K  Q  Q  K  E
      -  L  I  M  Y  D  S  *  L  T  K  M  V  S  S  *  N  N  K  R  N
      -   *  L  C  T  T  P  N  S  R  K  W  Y  P  V  E  T  T  K  G  T
15361 - CACCATCTACAAATATTTTTCTTACTAGTGGTCCAAAACTTGTAGGTGGAAACACAGTAG -
15420
      - H  H  L  Q  I  F  F  L  L  V  V  Q  N  L  *  V  E  T  Q  *
      -  T  I  Y  K  Y  F  S  Y  *  W  S  K  T  C  R  W  K  H  S  R
      -   P  S  T  N  I  F  L  T  S  G  P  K  L  V  G  G  N  T  V  E
15421 - AAAATAACACATTAAAGTTTGCACAATGAAGGATACACCTATCATCCAAACAGTTAATAC -
15480
      - K  I  T  H  *  S  L  H  N  E  G  Y  T  Y  H  P  N  S  *  Y
      -  K  *  H  I  K  V  C  T  M  K  D  T  P  I  I  Q  T  V  N  T
      -   N  N  T  L  K  F  A  Q  *  R  I  H  L  S  S  K  Q  L  I  Q
15481 - AATTGGGATGGTATGTCTGGTCCCAATATTTAAAATAACGGTCGAAGAGACAAAGTCTCT -
15540
      - N  W  D  G  M  S  G  P  N  I  *  N  N  G  R  R  D  K  V  S
      -  I  G  M  V  C  L  V  P  I  F  K  I  T  V  E  E  T  K  S  L
      -   L  G  W  Y  V  W  S  Q  Y  L  K  *  R  S  K  R  Q  S  L  S
15541 - CTTCCGTAAAATCATATTTCAGCAAATCCCACTTAATAAGTGGTTTTGCGAGATCAGCAT -
15600
      - L  P  *  N  H  I  S  A  N  P  T  *  *  V  V  L  R  D  Q  H
      -  F  R  K  I  I  F  Q  Q  I  P  L  N  K  W  F  C  E  I  S  I
      -   S  V  K  S  Y  F  S  K  S  H  L  I  S  G  F  A  R  S  A  S
15601 - CCATATGGGACTCAGCAGCCAATGCCCTAGTCAAAGTGAGGATGGGCATCAGCAATGAGT -
```

15660
- P Y G T Q Q P M P * S K * G W A S A M S
- H M G L S S Q C P S Q S E D G H Q Q * V
- I W D S A A N A L V K V R M G I S N E *

Figure 17-V

```
15661 - AATATGAATCCACAATAGGAACTCCGCAGCCTGGTGCTACTTGTACGAAATCACCGAAAT -
15720
      - N  M  N  P  Q  *  E  L  R  S  L  V  L  L  V  R  N  H  R  N
      -  I  *  I  H  N  R  N  S  A  A  W  C  Y  L  Y  E  I  T  E  I
      -   Y  E  S  T  I  G  T  P  Q  P  G  A  T  C  T  K  S  P  K  S
15721 - CGTACCAGTTCCCATTAAGATCCTGATTATCTAATGTCAGTACGCCTACAATGCCTGCAT -
15780
      - R  T  S  S  H  *  D  P  D  Y  L  M  S  V  R  L  Q  C  L  H
      -  V  P  V  P  I  K  I  L  I  I  *  C  Q  Y  A  Y  N  A  C  I
      -   Y  Q  F  P  L  R  S  *  L  S  N  V  S  T  P  T  M  P  A  S
15781 - CACGCATAGCATCGCAGAATTGTACAGTCTTTAATAATGATTGGCGTACACGCTCACCTA -
15840
      - H  A  *  H  R  R  I  V  Q  S  L  I  M  I  G  V  H  A  H  L
      -  T  H  S  I  A  E  L  Y  S  L  *  *  *  L  A  Y  T  L  T  *
      -   R  I  A  S  Q  N  C  T  V  F  N  N  D  W  R  T  R  S  P  K
15841 - AGTTAGCATATACGCGTAAGATGTCAGGATTCTCTACGAAGTCATACCAATCCTTCTTAT -
15900
      - S  *  H  I  R  V  R  C  Q  D  S  L  R  S  H  T  N  P  S  Y
      -  V  S  I  Y  A  *  D  V  R  I  L  Y  E  V  I  P  I  L  L  I
      -   L  A  Y  T  R  K  M  S  G  F  S  T  K  S  Y  Q  S  F  L  L
15901 - TGAAATAATCATCATCACAGCAATTGTATGTGACGAGTATTTCTTTTAATGTATCACAAT -
15960
      - *  N  N  H  H  H  S  N  C  M  *  R  V  F  L  L  M  Y  H  N
      -  E  I  I  I  I  T  A  I  V  C  D  E  Y  F  F  *  C  I  T  I
      -   K  *  S  S  S  Q  Q  L  Y  V  T  S  I  S  F  N  V  S  Q  L
15961 - TACCCTCATCAAAATGACGTAGAGCATAGACTAAATCAGCCATTGTGTATTTAGTTAGAC -
16020
      - Y  P  H  Q  N  D  V  E  H  R  L  N  Q  P  L  C  I  *  L  D
      -  T  L  I  K  M  T  *  S  I  D  *  I  S  H  C  V  F  S  *  T
      -   P  S  S  K  *  R  R  A  *  T  K  S  A  I  V  Y  L  V  R  R
16021 - GCTGACGTGATATATGTGGTACCATGTCACCATCTACTCTAAACTTGAAAAAGTCATGGA -
16080
      - A  D  V  I  Y  V  V  P  C  H  H  L  L  *  T  *  K  S  H  G
      -  L  T  *  Y  M  W  Y  H  V  T  I  Y  S  K  L  E  K  V  M  D
      -   *  R  D  I  C  G  T  M  S  P  S  T  L  N  L  K  K  S  W  T
16081 - CAGCAACCGCTGGACAATCTTTAACCAAGTTATAAATAGTCTCTTCATGTTGGTAGTTAG -
16140
      - Q  Q  P  L  D  N  L  *  P  S  Y  K  *  S  L  H  V  G  S  *
      -  S  N  R  W  T  I  F  N  Q  V  I  N  S  L  F  M  L  V  V  R
      -   A  T  A  G  Q  S  L  T  K  L  *  I  V  S  S  C  W  *  L  D
16141 - ACATAGTATGCCTCTTAACTACAAAGTAAGAGTCTAATAAATTGCCTTCCTCATCCTTCT -
16200
      - T  *  Y  A  S  *  L  Q  S  K  S  L  I  N  C  L  P  H  P  S
      -  H  S  M  P  L  N  Y  K  V  R  V  *  *  I  A  F  L  I  L  L
      -   I  V  C  L  L  T  T  K  *  E  S  N  K  L  P  S  S  S  F  S
16201 - CCTGGAAGCGACAGCAATTAGTTTTTAGGAACTTTGCAAAACCAGCACTTTTTTCGTTGT -
16260
      - P  G  S  D  S  N  *  F  L  G  T  L  Q  N  Q  H  F  F  R  C
      -  L  E  A  T  A  I  S  F  *  E  L  C  K  T  S  T  F  F  V  V
      -   W  K  R  Q  Q  L  V  F  R  N  F  A  K  P  A  L  F  S  L  *
16261 - AAATATCAAAAGCCCTGTAGACGACATCAGTACTAGTGCCTGTGCCGCACGGTGTAAGAC -
16320
      - K  Y  Q  K  P  C  R  R  H  Q  Y  *  C  L  C  R  T  V  *  D
      -  N  I  K  S  P  V  D  D  I  S  T  S  A  C  A  A  R  C  K  T
      -   I  S  K  A  L  *  T  T  S  V  L  V  P  V  P  H  G  V  R  R
16321 - GGGCTGCACTTACACCGCAAACCCGTTTAAAAACGTTGATGCATCCGCAGACTGCATCAA -
```

16380

- G L H L H R K P V * K R * C I R R L H Q
- G C T Y T A N P F K N V D A S A D C I K
- A A L T P Q T R L K T L M H P Q T A S R

Figure 17-W

```
16381 - GGGTTCGCGGAGTTGGTCACAACTACAGCCATAACCTTTCCACATTCCGCAGACGGTACA -
16440
      - G  F  A  E  L  V  T  T  T  A  I  T  F  P  H  S  A  D  G  T
      -  G  S  R  S  W  S  Q  L  Q  P  *  P  F  H  I  P  Q  T  V  Q
      -   V  R  G  V  G  H  N  Y  S  H  N  L  S  T  F  R  R  R  Y  R
16441 - GACTGTGTTTCTAAGTGTAAAACCCACTGGGTCATTAGCACAAGTGGTAGGTATTTGGAC -
16500
      - D  C  V  S  K  C  K  T  H  W  V  I  S  T  S  G  R  Y  L  D
      -  T  V  F  L  S  V  K  P  T  G  S  L  A  Q  V  V  G  I  W  T
      -   L  C  F  *  V  *  N  P  L  G  H  *  H  K  W  *  V  F  G  R
16501 - GTACTTACCTTTCAAGTCACAGAATCCTTTAGGATTTGGATGGTCAATGTGGCATCTACA -
16560
      - V  L  T  F  Q  V  T  E  S  F  R  I  W  M  V  N  V  A  S  T
      -  Y  L  P  F  K  S  Q  N  P  L  G  F  G  W  S  M  W  H  L  Q
      -   T  Y  L  S  S  H  R  I  L  *  D  L  D  G  Q  C  G  I  Y  N
16561 - ATACAGACAACATGAAGCACCACCAAAGGACTCTTGGTCCATGTTAGCTTCTGGTGTTAC -
16620
      - I  Q  T  T  *  S  T  T  K  G  L  L  V  H  V  S  F  W  C  Y
      -  Y  R  Q  H  E  A  P  P  K  D  S  W  S  M  L  A  S  G  V  T
      -   T  D  N  M  K  H  H  Q  R  T  L  G  P  C  *  L  L  V  L  Q
16621 - AGTAATTGCCTGTCCTGTACCAGTGTGTGTACACAACATCTTCACACAGTTGGTGATTGG -
16680
      - S  N  C  L  S  C  T  S  V  C  T  Q  H  L  H  T  V  G  D  W
      -  V  I  A  C  P  V  P  V  C  V  H  N  I  F  T  Q  L  V  I  G
      -   *  L  P  V  L  Y  Q  C  V  Y  T  T  S  S  H  S  W  *  L  V
16681 - TTGTCCTCCACTTGCTAGGTAATCCTTATATGCTTTAGCAGGGTCTACTGCAAAAGCACA -
16740
      - L  S  S  T  C  *  V  I  L  I  C  F  S  R  V  Y  C  K  S  T
      -  C  P  P  L  A  R  *  S  L  Y  A  L  A  G  S  T  A  K  A  Q
      -   V  L  H  L  L  G  N  P  Y  M  L  *  Q  G  L  L  Q  K  H  R
16741 - GAAGGAAAGCACAGTTGAATTGGCAGGTACTTCTGTAGCATTTCCAGCCTGAAGACGTAC -
16800
      - E  G  K  H  S  *  I  G  R  Y  F  C  S  I  S  S  L  K  T  Y
      -  K  E  S  T  V  E  L  A  G  T  S  V  A  F  P  A  *  R  R  T
      -   R  K  A  Q  L  N  W  Q  V  L  L  *  H  F  Q  P  E  D  V  L
16801 - TGTAGCAGCTAAACTGCCCAGCACCATACCTCTATTTAGGTTGTTTAAGCCTTTGATGAA -
16860
      - C  S  S  *  T  A  Q  H  H  T  S  I  *  V  V  *  A  F  D  E
      -  V  A  A  K  L  P  S  T  I  P  L  F  R  L  F  K  P  L  M  K
      -   *  Q  L  N  C  P  A  P  Y  L  Y  L  G  C  L  S  L  *  *  S
16861 - GTACAAGTATTTCACTTTAGGCCCTTTTGGTGTGTCTGTAACAAACCTACAAGGTGGTTC -
16920
      - V  Q  V  F  H  F  R  P  F  W  C  V  C  N  K  P  T  R  W  F
      -  Y  K  Y  F  T  L  G  P  F  G  V  S  V  T  N  L  Q  G  G  S
      -   T  S  I  S  L  *  A  L  L  V  C  L  *  Q  T  Y  K  V  V  P
16921 - CAGTTCTGTGTAAATTGTACCTGTACCATCACTCTTAGGGAATCTAGCCCATTTGAGATC -
16980
      - Q  F  C  V  N  C  T  C  T  I  T  L  R  E  S  S  P  F  E  I
      -  S  S  V  *  I  V  P  V  P  S  L  L  G  N  L  A  H  L  R  S
      -   V  L  C  K  L  Y  L  Y  H  H  S  *  G  I  *  P  I  *  D  L
16981 - TTGGTGGTCTGATAGTAATGCCAGCACAAACCTACCTCCCTTCGAATTGTTATAGTAGGC -
17040
      - L  V  V  *  *  *  C  Q  H  K  P  T  S  L  R  I  V  I  V  G
      -  W  W  S  D  S  N  A  S  T  N  L  P  P  F  E  L  L  *  *  A
      -   G  G  L  I  V  M  P  A  Q  T  Y  L  P  S  N  C  Y  S  R  Q
17041 - AAGTGCATTGTCATCAGTACAAGCTGTTTGTGTGGTACCAGCCGCACAGGACATCTGTCG -
```

17100
-K C I V I S T S C L C G T S R T G H L S
- S A L S S V Q A V C V V P A A Q D I C R
- V H C H Q Y K L F V W Y Q P H R T S V V

Figure 17-X

```
 17101 - TAGTGCTACTGGACTCAGTTCATTATTCTGTAGTTTAACAGCTGAGTTGGCTCTTAGAGC -
17160
       - * C Y W T Q F I I L * F N S * V G S * S
       - S A T G L S S L F C S L T A E L A L R A
       -  V L L D S V H Y S V V * Q L S W L L E L
 17161 - TGTAACAATAAGAGGCCAAGCCAAATTTGGTGAATTGTCCATGTTAATTTCACTAAGTTG -
17220
       - C N N K R P S Q I W * I V H V N F T K L
       -  V T I R G Q A K F G E L S M L I S L S *
       -   * Q * E A K P N L V N C P C * F H * V E
 17221 - AACAATCTTGCTATCCGCATCAACAACTTGCTGGATTTCCCAGAGTGCAGATGCATATGT -
17280
       - N N L A I R I N N L L D F P E C R C I C
       -  T I L L S A S T T C W I S Q S A D A Y V
       -   Q S C Y P H Q Q L A G F P R V Q M H M *
 17281 - AAAGGTGTTACCATCACAAGTGTTCTTGTAGGTACCATAATCAGGGACAACAACCATGAG -
17340
       - K G V T I T S V L V G T I I R D N N H E
       -  K V L P S Q V F L * V P * S G T T T M S
       -   R C Y H H K C S C R Y H N Q G Q Q P * V
 17341 - TTTGGCTGCTGTAGTCAATGGTATGATGTTGAGTGGAACACAACCATCACGCGCATTGTT -
17400
       - F G C C S Q W Y D V E W N T T I T R I V
       -  L A A V V N G M M L S G T Q P S R A L L
       -   W L L * S M V * C * V E H N H H A H C *
 17401 - GATAATGTTGTTAAGTGCATCATTATCAAGCTTCCTAAGCATAGTGAAGAGCATTGTTTG -
17460
       - D N V V K C I I I K L P K H S E E H C L
       -  I M L L S A S L S S F L S I V K S I V C
       -   * C C * V H H Y Q A S * A * * R A L F A
 17461 - CATAGCACTAGTTACTTTTGCCCTCTTGTCCTCAGATCTTGCCTGTTTGTACATTTGGGT -
17520
       - H S T S Y F C P L V L R S C L F V H L G
       -  I A L V T F A L L S S D L A C L Y I W V
       -   * H * L L L P S C P Q I L P V C T F G S
 17521 - CATAGCCTGATCTGCCATCTTTTCCAACTTGCGTTGCATGGCAGCATCACGGTCAAACTC -
17580
       - H S L I C H L F Q L A L H G S I T V K L
       -  I A * S A I F S N L R C M A A S R S N S
       -   * P D L P S F P T C V A W Q H H G Q T Q
 17581 - AGATTTAGCCACATTCAAAGATTTCTTTAACTTTTTGAGAACGACTTCAGAATCACCATT -
17640
       - R F S H I Q R F L * L F E N D F R I T I
       -  D L A T F K D F F N F L R T T S E S P L
       -   I * P H S K I S L T F * E R L Q N H H *
 17641 - AGCTACAGCCTGCTCATAGGCCTCCTGGGCAGTGGCATAAGCGGCATATGATGGTAAAGA -
17700
       - S Y S L L I G L L G S G I S G I * W * R
       -  A T A C S * A S W A V A * A A Y D G K E
       -   L Q P A H R P P G Q W H K R H M M V K N
 17701 - ACTAAATTCTGAAGCAATAGCCTGAAGAGTAGCACGGTTATCGAGCATTTCCTCGCACAA -
17760
       - T K F * S N S L K S S T V I E H F L A Q
       -  L N S E A I A * R V A R L S S I S S H N
       -   * I L K Q * P E E * H G Y R A F P R T T
 17761 - CCTATTAATGTCTACAGCACCCTGCATGGATAGCAAAACAGACAAAAGAGAAACCATCTT -
17820
```

```
      - P  I  N  V  Y  S  T  L  H  G  *  Q  N  R  Q  K  R  N  H  L
      -  L  L  M  S  T  A  P  C  M  D  S  K  T  D  K  R  E  T  I  F
      -   Y  *  C  L  Q  H  P  A  W  I  A  K  Q  T  K  E  K  P  S  S
17821 - CTCGAAAGCTTCAGTTGTGTCTTTTGCAAGAAGAATATCATTGTGGAGTTGTACACATTG -
17880
      - L  E  S  F  S  C  V  F  C  K  K  N  I  I  V  E  L  Y  T  L
      -  S  K  A  S  V  V  S  F  A  R  R  I  S  L  W  S  C  T  H  C
      -   R  K  L  Q  L  C  L  L  Q  E  E  Y  H  C  G  V  V  H  I  V
```

Figure 17-Y

```
17881 - TGCCCACAATTTAGAAGATGACTCTACTCTAAGTTGTTGAAGAACCGAGAGCAGTACCAC -
17940
      -C  P  Q  F  R  R  *  L  Y  S  K  L  L  K  N  R  E  Q  Y  H
      - A  H  N  L  E  D  D  S  T  L  S  C  *  R  T  E  S  S  T  T
      -  P  T  I  *  K  M  T  L  L  *  V  V  E  E  P  R  A  V  P  Q
17941 - AGATGTGCACTTTACGTCAGACATTTTAGACTGTACAGTAGCAACCTTGATACATGGTTT -
18000
      -R  C  A  L  Y  V  R  H  F  R  L  Y  S  S  N  L  D  T  W  F
      - D  V  H  F  T  S  D  I  L  D  C  T  V  A  T  L  I  H  G  L
      -  M  C  T  L  R  Q  T  F  *  T  V  Q  *  Q  P  *  Y  M  V  Y
18001 - ACCTCCAATACCCAACAACTTAATGTTAAGCTTGAAAGCATCAATACTACTCTTAGGAGG -
18060
      -T  S  N  T  Q  Q  L  N  V  K  L  E  S  I  N  T  T  L  R  R
      - P  P  I  P  N  N  L  M  L  S  L  K  A  S  I  L  L  L  G  G
      -  L  Q  Y  P  T  T  *  C  *  A  *  K  H  Q  Y  Y  S  *  E  A
18061 - CAAAAGCCCCTGGGAGTTCATATACCTAAATTCTTGTGTAGAGACCAAGTAGTCATAAAC -
18120
      -Q  K  P  L  G  V  H  I  P  K  F  L  C  R  D  Q  V  V  I  N
      - K  S  P  W  E  F  I  Y  L  N  S  C  V  E  T  K  *  S  *  T
      -  K  A  P  G  S  S  Y  T  *  I  L  V  *  R  P  S  S  H  K  H
18121 - ACCAAGAGTAAGCCTGAAGTAACGGTTGAGTAAACAGAAAAGGCCAAAGTAGCAGCAGCA -
18180
      -T  K  S  K  P  E  V  T  V  E  *  T  E  K  A  K  V  A  A  A
      - P  R  V  S  L  K  *  R  L  S  K  Q  K  R  P  K  *  Q  Q  Q
      -  Q  E  *  A  *  S  N  G  *  V  N  R  K  G  Q  S  S  S  S  N
18181 - ACAATAGCCTAAGAAACAATAAACAAGCATGATACACTGTAAGGTGTTGCCAGTAATAAA -
18240
      -T  I  A  *  E  T  I  N  K  H  D  T  L  *  G  V  A  S  N  K
      - Q  *  P  K  K  Q  *  T  S  M  I  H  C  K  V  L  P  V  I  N
      -  N  S  L  R  N  N  K  Q  A  *  Y  T  V  R  C  C  Q  *  *  I
18241 - TAACAATGGGTAATACTCAACACACACAAACACTATAGCTCTAGCTAAAAACATGATAGT -
18300
      -*  Q  W  V  I  L  N  T  H  K  H  Y  S  S  S  *  K  H  D  S
      - N  N  G  *  Y  S  T  H  T  N  T  I  A  L  A  K  N  M  I  V
      -  T  M  G  N  T  Q  H  T  Q  T  L  *  L  *  L  K  T  *  *  S
18301 - CGTAACGACACCAGAATAGTTAGAGGTTACAGAAATAACTAAGGCCCACATGGAAATAGC -
18360
      -R  N  D  T  R  I  V  R  G  Y  R  N  N  *  G  P  H  G  N  S
      - V  T  T  P  E  *  L  E  V  T  E  I  T  K  A  H  M  E  I  A
      -  *  R  H  Q  N  S  *  R  L  Q  K  *  L  R  P  T  W  K  *  L
18361 - TTGATCTAAAGCATTACCATAGTAGACTTTGTAAACAAGTGTAATGACATTCATCAGTGT -
18420
      -L  I  *  S  I  T  I  V  D  F  V  N  K  C  N  D  I  H  Q  C
      - *  S  K  A  L  P  *  *  T  L  *  T  S  V  M  T  F  I  S  V
      -  D  L  K  H  Y  H  S  R  L  C  K  Q  V  *  *  H  S  S  V  S
18421 - CCAAACACGTCTAGCAGCATCATCATAAACAGTGCGAGCTGTCATGAGAATAAGCAAAAC -
18480
      -P  N  T  S  S  S  I  I  I  N  S  A  S  C  H  E  N  K  Q  N
      - Q  T  R  L  A  A  S  S  *  T  V  R  A  V  M  R  I  S  K  T
      -  K  H  V  *  Q  H  H  H  K  Q  C  E  L  S  *  E  *  A  K  L
18481 - TAAAGCTGAAGCATACATAACACAATCCTTAAGCCTATAACCAGACAAGCTAGTGTCAGC -
18540
      -*  S  *  S  I  H  N  T  I  L  K  P  I  T  R  Q  A  S  V  S
      - K  A  E  A  Y  I  T  Q  S  L  S  L  *  P  D  K  L  V  S  A
      -  K  L  K  H  T  *  H  N  P  *  A  Y  N  Q  T  S  *  C  Q  P
18541 - CAATTCAAGCCATGTCATGATACGCATCACCCAGCTAGCAGGCATGTAGACCATATTAAA -
18600
```

```
      -Q  F  K  P  C  H  D  T  H  H  P  A  S  R  H  V  D  H  I  K
      - N  S  S  H  V  M  I  R  I  T  Q  L  A  G  M  *  T  I  L  K
      -  I  Q  A  M  S  *  Y  A  S  P  S  *  Q  A  C  R  P  Y  *  S
18601 - GTAAGCAACTGTTGCAAGAGAAGGTAACAGAAACAAGCACAAGAATGCGTGCTTATGCTT -
18660
      -V  S  N  C  C  K  R  R  *  Q  K  Q  A  Q  E  C  V  L  M  L
      - *  A  T  V  A  R  E  G  N  R  N  K  H  K  N  A  C  L  C  L
      -  K  Q  L  L  Q  E  K  V  T  E  T  S  T  R  M  R  A  Y  A  *
```

Figure 17-Z

```
18661 - AACAAGCAGCATAGCACATGCAGCAATTGCCATAATACCAAGAGTAAATGGCAAGAAAGC -
18720
      - N  K  Q  H  S  T  C  S  N  C  H  N  T  K  S  K  W  Q  E  S
      -  T  S  S  I  A  H  A  A  I  A  I  I  P  R  V  N  G  K  K  A
      -   Q  A  A  *  H  M  Q  Q  L  P  *  Y  Q  E  *  M  A  R  K  H
18721 - ATTCTCGTAAACAAAGAAAAACAGTGACCACTGTGTACTTTGAACAAGAATCAATAGTGA -
18780
      - I  L  V  N  K  E  K  Q  *  P  L  C  T  L  N  K  N  Q  *  *
      -  F  S  *  T  K  K  N  S  D  H  C  V  L  *  T  R  I  N  S  D
      -   S  R  K  Q  R  K  T  V  T  T  V  Y  F  E  Q  E  S  I  V  M
18781 - TGTCAAGAAAGTTAAAAGCATCCAATGATGAGTGCCCTTAACAATTTTCTTGAACTTACC -
18840
      - C  Q  E  S  *  K  H  P  M  M  S  A  L  N  N  F  L  E  L  T
      -  V  K  K  V  K  S  I  Q  *  *  V  P  L  T  I  F  L  N  L  P
      -   S  R  K  L  K  A  S  N  D  E  C  P  *  Q  F  S  *  T  Y  L
18841 - TTGGAAGGTAACACCAGAGCATTGTCTAACAACATCAAATGGTGTAAACTCATCTTCTAA -
18900
      - L  E  G  N  T  R  A  L  S  N  N  I  K  W  C  K  L  I  F  *
      -  W  K  V  T  P  E  H  C  L  T  T  S  N  G  V  N  S  S  S  K
      -   G  R  *  H  Q  S  I  V  *  Q  H  Q  M  V  *  T  H  L  L  K
18901 - AATAGTGCTACCAAGGATAGTACGACCATTCATACCATTCTGCAGCAGCTCTTTCAAAGC -
18960
      - N  S  A  T  K  D  S  T  T  I  H  T  I  L  Q  Q  L  F  Q  S
      -  I  V  L  P  R  I  V  R  P  F  I  P  F  C  S  S  S  F  K  A
      -   *  C  Y  Q  G  *  Y  D  H  S  Y  H  S  A  A  A  L  S  K  Q
18961 - AGCACACATATCTAAGACGGCAATTCCTGTTTGAGCAGAAAGAGGTCCCAATATGTCAAC -
19020
      - S  T  H  I  *  D  G  N  S  C  L  S  R  K  R  S  Q  Y  V  N
      -  A  H  I  S  K  T  A  I  P  V  *  A  E  R  G  P  N  M  S  T
      -   H  T  Y  L  R  R  Q  F  L  F  E  Q  K  E  V  P  I  C  Q  H
19021 - ATGATCTTGTGTCAAAGGTTCATAGTTGTACTTCATTGCCACAAGGTTAAAGTCATTCAA -
19080
      - M  I  L  C  Q  R  F  I  V  V  L  H  C  H  K  V  K  V  I  Q
      -  *  S  C  V  K  G  S  *  L  Y  F  I  A  T  R  L  K  S  F  K
      -   D  L  V  S  K  V  H  S  C  T  S  L  P  Q  G  *  S  H  S  K
19081 - AGTAGTGGTGAATCTATTAAGAAACCACCTATCACCATTGATAACAGCAGCATACAGCCA -
19140
      - S  S  G  E  S  I  K  K  P  P  I  T  I  D  N  S  S  I  Q  P
      -  V  V  V  N  L  L  R  N  H  L  S  P  L  I  T  A  A  Y  S  H
      -   *  W  *  I  Y  *  E  T  T  Y  H  H  *  *  Q  Q  H  T  A  M
19141 - TGCCAAAACATTTAATGTTATGGTTGTGTCTGTACCTGCAGCCTGTGCAGTTTGTCTGTC -
19200
      - C  Q  N  I  *  C  Y  G  C  V  C  T  C  S  L  C  S  L  S  V
      -  A  K  T  F  N  V  M  V  V  S  V  P  A  A  C  A  V  C  L  S
      -   P  K  H  L  M  L  W  L  C  L  Y  L  Q  P  V  Q  F  V  C  Q
19201 - AACAAATGGACCATAGAATTTACCTTCTAAGTCAGTACCAGCGTGTACTCCTGTTGGAAG -
19260
      - N  K  W  T  I  E  F  T  F  *  V  S  T  S  V  Y  S  C  W  K
      -  T  N  G  P  *  N  L  P  S  K  S  V  P  A  C  T  P  V  G  S
      -   Q  M  D  H  R  I  Y  L  L  S  Q  Y  Q  R  V  L  L  L  E  A
19261 - CTCCATATGATGCATATAGCAGAAAGACACGCAATCATAATCAATGTTAAAACCAACACT -
19320
      - L  H  M  M  H  I  A  E  R  H  A  I  I  I  N  V  K  T  N  T
      -  S  I  *  C  I  *  Q  K  D  T  Q  S  *  S  M  L  K  P  T  L
      -   P  Y  D  A  Y  S  R  K  T  R  N  H  N  Q  C  *  N  Q  H  Y
19321 - ACCACATGATCCATTAAGGAAAGAACCTTTAATGGTATGATTAGGTCTCATGGCACACTG -
```

19380

```
-T T * S I K E R T F N G M I R S H G T L
- P H D P L R K E P L M V * L G L M A H *
-  H M I H * G K N L * W Y D * V S W H T D
```

Figure 17-AA

```
19381 - ATAAACACCAGATGGTGAACCATTGTAGCATGCTAGAACTGAAAATGTTTGACCAGGTTG -
19440
      -I  N  T  R  W  *  T  I  V  A  C  *  N  *  K  C  L  T  R  L
      - *  T  P  D  G  E  P  L  *  H  A  R  T  E  N  V  *  P  G  W
      -  K  H  Q  M  V  N  H  C  S  M  L  E  L  K  M  F  D  Q  V  G
19441 - GATACGGACAAATTTATACTTGGGTGTCTTAGGGTTAGAAGTATCAACTTTAAGCCTAAG -
19500
      -D  T  D  K  F  I  L  G  C  L  R  V  R  S  I  N  F  K  P  K
      - I  R  T  N  L  Y  L  G  V  L  G  L  E  V  S  T  L  S  L  S
      -  Y  G  Q  I  Y  T  W  V  S  *  G  *  K  Y  Q  L  *  A  *  A
19501 - CAGACAATTTTGCATAGAATGGCCAATAACACGAAGTTGAACATTGCCAGCCTGAACAAG -
19560
      -Q  T  I  L  H  R  M  A  N  N  T  K  L  N  I  A  S  L  N  K
      - R  Q  F  C  I  E  W  P  I  T  R  S  *  T  L  P  A  *  T  R
      -  D  N  F  A  *  N  G  Q  *  H  E  V  E  H  C  Q  P  E  Q  E
19561 - AAAGCTATGGTTGGATTTGCGAATGAGCAGATCTTCATAGTTAGGATTAAGCATGTCTTC -
19620
      -K  A  M  V  G  F  A  N  E  Q  I  F  I  V  R  I  K  H  V  F
      - K  L  W  L  D  L  R  M  S  R  S  S  *  L  G  L  S  M  S  S
      -  S  Y  G  W  I  C  E  *  A  D  L  H  S  *  D  *  A  C  L  L
19621 - TGCTGTGCAAATGACATGTCTTGGACAGTATACTGTGTCATCCAACCACAATCCATTAAG -
19680
      -C  C  A  N  D  M  S  W  T  V  Y  C  V  I  Q  P  Q  S  I  K
      - A  V  Q  M  T  C  L  G  Q  Y  T  V  S  S  N  H  N  P  L  R
      -  L  C  K  *  H  V  L  D  S  I  L  C  H  P  T  T  I  H  *  E
19681 - AGTTGTAGTTCCACAGGTTACTTGTACCATGCACCCTTCAACTTTGCCTGACGGGAATGC -
19740
      -S  C  S  S  T  G  Y  L  Y  H  A  P  F  N  F  A  *  R  E  C
      - V  V  V  P  Q  V  T  C  T  M  H  P  S  T  L  P  D  G  N  A
      -  L  *  F  H  R  L  L  V  P  C  T  L  Q  L  C  L  T  G  M  P
19741 - CATTTTCCTAAAACCACTCTGCAGAACAGCAGAAGTGATTGATGTCTGTGGTGGTTGGTA -
19800
      -H  F  P  K  T  T  L  Q  N  S  R  S  D  *  C  L  W  W  L  V
      - I  F  L  K  P  L  C  R  T  A  E  V  I  D  V  C  G  G  W  *
      -  F  S  *  N  H  S  A  E  Q  Q  K  *  L  M  S  V  V  V  G  R
19801 - GAGAACATCAGCACCTGAGTTGCTAAAGTCATTTAGAGCCTTTGCTAAGTGGCAGCAAGC -
19860
      -E  N  I  S  T  *  V  A  K  V  I  *  S  L  C  *  V  A  A  S
      - R  T  S  A  P  E  L  L  K  S  F  R  A  F  A  K  W  Q  Q  A
      -  E  H  Q  H  L  S  C  *  S  H  L  E  P  L  L  S  G  S  K  L
19861 - TGCTTCACGATAGCTGGTAGTATCTAAGGCTCCACTGAAATACTTGTACTTGTTATATAG -
19920
      -C  F  T  I  A  G  S  I  *  G  S  T  E  I  L  V  L  V  I  *
      - A  S  R  *  L  V  V  S  K  A  P  L  K  Y  L  Y  L  L  Y  R
      -  L  H  D  S  W  *  Y  L  R  L  H  *  N  T  C  T  C  Y  I  E
19921 - AGCAAGATACCTGTTATACTGTGTAAGTGGCAACAGTGTCTCGCTACGCAATTTTAGGTA -
19980
      -S  K  I  P  V  I  L  C  K  W  Q  Q  C  L  A  T  Q  F  *  V
      - A  R  Y  L  L  Y  C  V  S  G  N  S  V  S  L  R  N  F  R  Y
      -  Q  D  T  C  Y  T  V  *  V  A  T  V  S  R  Y  A  I  L  G  T
19981 - CATTTCCTTGTTGAGCAAAAAGGTACACAAAGCAGCCTCCTCGAAGGTACTAAATGTAAC -
20040
      -H  F  L  V  E  Q  K  G  T  Q  S  S  L  L  E  G  T  K  C  N
      - I  S  L  L  S  K  K  V  H  K  A  A  S  S  K  V  L  N  V  T
      -  F  P  C  *  A  K  R  Y  T  K  Q  P  P  R  R  Y  *  M  *  L
20041 - TCCATTAAACATGACTCTTTTCCTAAGATAGTTGTTAAAGAACCAATGGCAGTGCTTCAG -
```

20100

- S I K H D S F P K I V V K E P M A V L Q
- P L N M T L F L R * L L K N Q W Q C F R
- H * T * L F S * D S C * R T N G S A S E

Figure 17-BB

```
20101 - AGAAATACAGAATACATAGATTGCTGTTATCCAAAAAGGCACAATAGGAGAAAACATGGC -
20160
      - R  N  T  E  Y  I  D  C  C  Y  P  K  R  H  N  R  R  K  H  G
      -  E  I  Q  N  T  *  I  A  V  I  Q  K  G  T  I  G  E  N  M  A
      -   K  Y  R  I  H  R  L  L  L  S  K  K  A  Q  *  E  K  T  W  Q
20161 - AAACCATTGAAGGTGAGCCAAGAATGAAACATCATTGGTGAAATAGAATGTCAAGTACAA -
20220
      - K  P  L  K  V  S  Q  E  *  N  I  I  G  E  I  E  C  Q  V  Q
      -  N  H  *  R  *  A  K  N  E  T  S  L  V  K  *  N  V  K  Y  K
      -   T  I  E  G  E  P  R  M  K  H  H  W  *  N  R  M  S  S  T  S
20221 - GTAAAAGACTGAGTAGACTCCCGGCAGAAAGCTGTAAGCTGGTACCAGACAGAGTATAGT -
20280
      - V  K  D  *  V  D  S  R  Q  K  A  V  S  W  Y  Q  T  E  Y  S
      -  *  K  T  E  *  T  P  G  R  K  L  *  A  G  T  R  Q  S  I  V
      -   K  R  L  S  R  L  P  A  E  S  C  K  L  V  P  D  R  V  *  *
20281 - GAAAGACATCAAAAACAAAAGTGCATTAGCAGCAACAACATGGTTGTACTCACCAAAAAC -
20340
      - E  R  H  Q  K  Q  K  C  I  S  S  N  N  M  V  V  L  T  K  N
      -  K  D  I  K  N  K  S  A  L  A  A  T  T  W  L  Y  S  P  K  T
      -   K  T  S  K  T  K  V  H  *  Q  Q  Q  H  G  C  T  H  Q  K  H
20341 - ACGTCTGAATTTCATAAAGTAGTAGGCAGCACAAGTCACCAATATGGCAATAATACCACC -
20400
      - T  S  E  F  H  K  V  V  G  S  T  S  H  Q  Y  G  N  N  T  T
      -  R  L  N  F  I  K  *  *  A  A  Q  V  T  N  M  A  I  I  P  P
      -   V  *  I  S  *  S  S  R  Q  H  K  S  P  I  W  Q  *  Y  H  Q
20401 - AGCCACTACTGAAGCAGACACATCTAAAGCACCCACAGGTTGCACAAGAGGAGTAAAGAT -
20460
      - S  H  Y  *  S  R  H  I  *  S  T  H  R  L  H  K  R  S  K  D
      -  A  T  T  E  A  D  T  S  K  A  P  T  G  C  T  R  G  V  K  M
      -   P  L  L  K  Q  T  H  L  K  H  P  Q  V  A  Q  E  E  *  R  C
20461 - GTTAGCTATGAGATTCATCGCATCAACACCACAGAAAACTCCTGATAGAGCTCTGTAATG -
20520
      - V  S  Y  E  I  H  R  I  N  T  T  E  N  S  *  *  S  S  V  M
      -  L  A  M  R  F  I  A  S  T  P  Q  K  T  P  D  R  A  L  *  C
      -   *  L  *  D  S  S  H  Q  H  H  R  K  L  L  I  E  L  C  N  A
20521 - CTCATTATTAAGAACCCATCTACCACTGGTAGATAGGCAAATACCTACTTCTGACCTTTC -
20580
      - L  I  I  K  N  P  S  T  T  G  R  *  A  N  T  Y  F  *  P  F
      -  S  L  L  R  T  H  L  P  L  V  D  R  Q  I  P  T  S  D  L  S
      -   H  Y  *  E  P  I  Y  H  W  *  I  G  K  Y  L  L  L  T  F  R
20581 - GCATGTACCATGTCTACAGTACTCAGCATCAAAAGTTGTTACTACTCTAACAGAACCCTC -
20640
      - A  C  T  M  S  T  V  L  S  I  K  S  C  Y  Y  S  N  R  T  L
      -  H  V  P  C  L  Q  Y  S  A  S  K  V  V  T  T  L  T  E  P  S
      -   M  Y  H  V  Y  S  T  Q  H  Q  K  L  L  L  L  *  Q  N  P  P
20641 - CAGGTAAGTGTTAGGAAACTGTATGATGGAACCATCCATAAGCACATAACGAGTGTCTGG -
20700
      - Q  V  S  V  R  K  L  Y  D  G  T  I  H  K  H  I  T  S  V  W
      -  R  *  V  L  G  N  C  M  M  E  P  S  I  S  T  *  R  V  S  G
      -   G  K  C  *  E  T  V  *  W  N  H  P  *  A  H  N  E  C  L  D
20701 - ACGAAGCTCACTATAAGAAATAGAACCCTCTAGCAAATTAGTGTCATAACAATATGGCAC -
20760
      - T  K  L  T  I  R  N  R  T  L  *  Q  I  S  V  I  T  I  W  H
      -  R  S  S  L  *  E  I  E  P  S  S  K  L  V  S  *  Q  Y  G  T
      -   E  A  H  Y  K  K  *  N  P  L  A  N  *  C  H  N  N  M  A  Q
```

```
20761 - AGGTTTGCCCATAGCATCCTTAAAAATTGTACACTCAGCAGCAAGAACGCAAGCAGAGGT -
20820
      -R  F  A  H  S  I  L  K  N  C  T  L  S  S  K  N  A  S  R  G
      - G  L  P  I  A  S  L  K  I  V  H  S  A  A  R  T  Q  A  E  V
      -  V  C  P  *  H  P  *  K  L  Y  T  Q  Q  Q  E  R  K  Q  R  *
```

Figure 17-CC

```
20821 - AGCAAAATCACTATACTCAATGAGTTTGGAAGGTGTGTAGCAAATGTTGCCAACAGCACT -
20880
      - S  K  I  T  I  L  N  E  F  G  R  C  V  A  N  V  A  N  S  T
      -  A  K  S  L  Y  S  M  S  L  E  G  V  *  Q  M  L  P  T  A  L
      -   Q  N  H  Y  T  Q  *  V  W  K  V  C  S  K  C  C  Q  Q  H  *
20881 - AAAAACACGAGGTAGAAAATGCAAGAAGTCACCATTGATTGCTCTCAGCACAGTACCCGG -
20940
      - K  N  T  R  *  K  M  Q  E  V  T  I  D  C  S  Q  H  S  T  R
      -  K  T  R  G  R  K  C  K  K  S  P  L  I  A  L  S  T  V  P  G
      -   K  H  E  V  E  N  A  R  S  H  H  *  L  L  S  A  Q  Y  P  V
20941 - TAAGCCAGGCACTATGAAACCAATCTCTCTTGTAATGATAGCAGCTACTACAGGGCAGCT -
21000
      - *  A  R  H  Y  E  T  N  L  S  C  N  D  S  S  Y  Y  R  A  A
      -  K  P  G  T  M  K  P  I  S  L  V  M  I  A  A  T  T  G  Q  L
      -   S  Q  A  L  *  N  Q  S  L  L  *  *  *  Q  L  L  Q  G  S  F
21001 - TTTGTCATTTTTGTATGAACCACCACGCTGGCTAAACCATGCGTCAAAACCAGCATGTTT -
21060
      - F  V  I  F  V  *  T  T  T  L  A  K  P  C  V  K  T  S  M  F
      -  L  S  F  L  Y  E  P  P  R  W  L  N  H  A  S  K  P  A  C  L
      -   C  H  F  C  M  N  H  H  A  G  *  T  M  R  Q  N  Q  H  V  Y
21061 - ATTTGCAAAACAATCATCAGTAGAAATGATGTCACGAGTGACACCATCCTGAATGGCTTT -
21120
      - I  C  K  T  I  I  S  R  N  D  V  T  S  D  T  I  L  N  G  F
      -  F  A  K  Q  S  S  V  E  M  M  S  R  V  T  P  S  *  M  A  L
      -   L  Q  N  N  H  Q  *  K  *  C  H  E  *  H  H  P  E  W  L  C
21121 - GTAACCAATGATTTCATTTGTGTAACCATCATGGATTGACAATGTATGTACTGGCATAAC -
21180
      - V  T  N  D  F  I  C  V  T  I  M  D  *  Q  C  M  Y  W  H  N
      -  *  P  M  I  S  F  V  *  P  S  W  I  D  N  V  C  T  G  I  T
      -   N  Q  *  F  H  L  C  N  H  H  G  L  T  M  Y  V  L  A  *  R
21181 - GATATAACAAACCAATGCAGCAAGAACGCACAATAATGTGGCCTTAAGCATAAGTTTAAA -
21240
      - D  I  T  N  Q  C  S  K  N  A  Q  *  C  G  L  K  H  K  F  K
      -  I  *  Q  T  N  A  A  R  T  H  N  N  V  A  L  S  I  S  L  K
      -   Y  N  K  P  M  Q  Q  E  R  T  I  M  W  P  *  A  *  V  *  N
21241 - ACAAGTACTAACAATCTTACCACCCTTGAGTGAGATTTTAGTAGTTATGACATTGACAAC -
21300
      - T  S  T  N  N  L  T  T  L  E  *  D  F  S  S  Y  D  I  D  N
      -  Q  V  L  T  I  L  P  P  L  S  E  I  L  V  V  M  T  L  T  T
      -   K  Y  *  Q  S  Y  H  P  *  V  R  F  *  *  L  *  H  *  Q  P
21301 - CTGTCTAGTTGTAGCACAAGTTAGTGTAAAAGGTATGTTGTTCTTCTTGGCAGCAGTACG -
21360
      - L  S  S  C  S  T  S  *  C  K  R  Y  V  V  L  L  G  S  S  T
      -  C  L  V  V  A  Q  V  S  V  K  G  M  L  F  F  L  A  A  V  R
      -   V  *  L  *  H  K  L  V  *  K  V  C  C  S  S  W  Q  Q  Y  E
21361 - AATTTGTTTACGCAGCTGTTCAGATAAAGACATGTAGTCTTTTACATTCCAGATGAGTGA -
21420
      - N  L  F  T  Q  L  F  R  *  R  H  V  V  F  Y  I  P  D  E  *
      -  I  C  L  R  S  C  S  D  K  D  M  *  S  F  T  F  Q  M  S  E
      -   F  V  Y  A  A  V  Q  I  K  T  C  S  L  L  H  S  R  *  V  K
21421 - AACATTGTGACTTTTTGCTACTTGGGCATTGATATGCCTTGCATTACAGTCAATACATGC -
21480
      - N  I  V  T  F  C  Y  L  G  I  D  M  P  C  I  T  V  N  T  C
      -  T  L  *  L  F  A  T  W  A  L  I  C  L  A  L  Q  S  I  H  A
      -   H  C  D  F  L  L  L  G  H  *  Y  A  L  H  Y  S  Q  Y  M  R
21481 - GCCAAGATCTCTGGGCGTCATGTTTTCAACCTTATTATAGGTGAGCATGAAATTGTTACA -
```

21540
- A K I S G R H V F N L I I G E H E I V T
- P R S L G V M F S T L L * V S M K L L Q
- Q D L W A S C F Q P Y Y R * A * N C Y N

Figure 17-DD

```
21541 - ACTGTCACCTGTCACTTCTAAGTCAGAGTGATGTGAAAGTTTGAGACATTCAATAACATC -
21600
      -T  V  T  C  H  F  *  V  R  V  M  *  K  F  E  T  F  N  N  I
      - L  S  P  V  T  S  K  S  E  *  C  E  S  L  R  H  S  I  T  S
      -  C  H  L  S  L  L  S  Q  S  D  V  K  V  *  D  I  Q  *  H  P
21601 - CTTTGTGTCAACATCGGTATCAACAACACCTTGTCGGGCAGCTGACACGAATGTAGAAAG -
21660
      -L  C  V  N  I  G  I  N  N  T  L  S  G  S  *  H  E  C  R  K
      - F  V  S  T  S  V  S  T  T  P  C  R  A  A  D  T  N  V  E  R
      -  L  C  Q  H  R  Y  Q  Q  H  L  V  G  Q  L  T  R  M  *  K  G
21661 - GACACCATCTAAAGCTACACCCTTTGCTAACTCGCTGTGAGCTGTAGCAACAAGTGCCTT -
21720
      -D  T  I  *  S  Y  T  L  C  *  L  A  V  S  C  S  N  K  C  L
      - T  P  S  K  A  T  P  F  A  N  S  L  *  A  V  A  T  S  A  L
      -  H  H  L  K  L  H  P  L  L  T  R  C  E  L  *  Q  Q  V  P  *
21721 - AAGTTTTTCCATAGGAACACTAAAAGTTGCTGAAAAGGTGTCGACATAAGCATCAAACAT -
21780
      -K  F  F  H  R  N  T  K  S  C  *  K  G  V  D  I  S  I  K  H
      - S  F  S  I  G  T  L  K  V  A  E  K  V  S  T  *  A  S  N  I
      -  V  F  P  *  E  H  *  K  L  L  K  R  C  R  H  K  H  Q  T  S
21781 - CTTAACGGAAACTTCAGTACTATCTCCAACGTTTGATACAAGAGCTTGGTCAAGCAACAG -
21840
      -L  N  G  N  F  S  T  I  S  N  V  *  Y  K  S  L  V  K  Q  Q
      - L  T  E  T  S  V  L  S  P  T  F  D  T  R  A  W  S  S  N  R
      -  *  R  K  L  Q  Y  Y  L  Q  R  L  I  Q  E  L  G  Q  A  T  E
21841 - AATAGGTTGGCACATCAGCTGACTGTAGTACACAGAAGCAGACTTAGAAGCAGACTCGTC -
21900
      -N  R  L  A  H  Q  L  T  V  V  H  R  S  R  L  R  S  R  L  V
      - I  G  W  H  I  S  *  L  *  Y  T  E  A  D  L  E  A  D  S  S
      -  *  V  G  T  S  A  D  C  S  T  Q  K  Q  T  *  K  Q  T  R  R
21901 - GCATTTGGACTTGCCATCAAAAACTATGACATTAATAGGCAGTGAACCTTTAGTGTTGTT -
21960
      -A  F  G  L  A  I  K  N  Y  D  I  N  R  Q  *  T  F  S  V  V
      - H  L  D  L  P  S  K  T  M  T  L  I  G  S  E  P  L  V  L  L
      -  I  W  T  C  H  Q  K  L  *  H  *  *  A  V  N  L  *  C  C  *
21961 - AGCTCTCAAATTGTCTAAATTGACAAAATGGGAGAGCGGATGTCTCTCATAGGTCTTTTG -
22020
      -S  S  Q  I  V  *  I  D  K  M  G  E  R  M  S  L  I  G  L  L
      - A  L  K  L  S  K  L  T  K  W  E  S  G  C  L  S  *  V  F  *
      -  L  S  N  C  L  N  *  Q  N  G  R  A  D  V  S  H  R  S  F  D
22021 - ACCAGCCTTGTCAAAGTAGAGGTGAAGCGCGCCATTTTTCACAGCAACACTATCAACAAT -
22080
      -T  S  L  V  K  V  E  V  K  R  A  I  F  H  S  N  T  I  N  N
      - P  A  L  S  K  *  R  *  S  A  P  F  F  T  A  T  L  S  T  I
      -  Q  P  C  Q  S  R  G  E  A  R  H  F  S  Q  Q  H  Y  Q  Q  Y
22081 - ATACGATGACTGGTCAGTAGGGTTGATTGGTCTTTTAAACTGGAGTGACAAATCACGAGC -
22140
      -I  R  *  L  V  S  R  V  D  W  S  F  K  L  E  *  Q  I  T  S
      - Y  D  D  W  S  V  G  L  I  G  L  L  N  W  S  D  K  S  R  A
      -  T  M  T  G  Q  *  G  *  L  V  F  *  T  G  V  T  N  H  E  Q
22141 - AACTTCATCACTAATGAATGTACTACCAGTGCAAAATGTGTCACAATTGAGACAATTCCA -
22200
      -N  F  I  T  N  E  C  T  T  S  A  K  C  V  T  I  E  T  I  P
      - T  S  S  L  M  N  V  L  P  V  Q  N  V  S  Q  L  R  Q  F  Q
      -  L  H  H  *  *  M  Y  Y  Q  C  K  M  C  H  N  *  D  N  S  N
22201 - ATTGTGAGTCTTGCAGAAGCCACGGCCTCCATTTGCATAGACATAGAAAGATCTCTTCAT -
22260
```

```
      -I  V  S  L  A  E  A  T  A  S  I  C  I  D  I  E  R  S  L  H
      - L  *  V  L  Q  K  P  R  P  P  F  A  *  T  *  K  D  L  F  M
      -  C  E  S  C  R  S  H  G  L  H  L  H  R  H  R  K  I  S  S  C
22261 - GCCATTAACAATAGTTGTACACTCAACGCGTGTGGCACGATTGCGCTTATAGCACATCAT -
22320
      -A  I  N  N  S  C  T  L  N  A  C  G  T  I  A  L  I  A  H  H
      - P  L  T  I  V  V  H  S  T  R  V  A  R  L  R  L  *  H  I  M
      -  H  *  Q  *  L  Y  T  Q  R  V  W  H  D  C  A  Y  S  T  S  C
```

Figure 17-EE

```
22321 - GCAAGTCGAAGAGGTGCAACCATCCATGATATGAACATAGCTCTTCCATATGTAGTAGAA -
22380
      - A  S  R  R  G  A  T  I  H  D  M  N  I  A  L  P  Y  V  V  E
      -  Q  V  E  E  V  Q  P  S  M  I  *  T  *  L  F  H  M  *  *  K
      -   K  S  K  R  C  N  H  P  *  Y  E  H  S  S  S  I  C  S  R  K
22381 - AGAAGCAAAGAAGATGTACATCCTAACCATTGCAGAAACGGGTGCCATTTGTACAATACT -
22440
      - R  S  K  E  D  V  H  P  N  H  C  R  N  G  C  H  L  Y  N  T
      -  E  A  K  K  M  Y  I  L  T  I  A  E  T  G  A  I  C  T  I  L
      -   K  Q  R  R  C  T  S  *  P  L  Q  K  R  V  P  F  V  Q  Y  *
22441 - AATGATAAACCACATGAGCCAAGAATTGCTGATGAAATGACTAGCAAAATAGCCAAAGAA -
22500
      - N  D  K  P  H  E  P  R  I  A  D  E  M  T  S  K  I  A  K  E
      -  M  I  N  H  M  S  Q  E  L  L  M  K  *  L  A  K  *  P  K  N
      -   *  *  T  T  *  A  K  N  C  *  *  N  D  *  Q  N  S  Q  R  T
22501 - CACCTGCATTATAGCTGAAAGACCTAATAAATAAAAGAATTTTGTGAACAACATATATGC -
22560
      - H  L  H  Y  S  *  K  T  *  *  I  K  E  F  C  E  Q  H  I  C
      -  T  C  I  I  A  E  R  P  N  K  *  K  N  F  V  N  N  I  Y  A
      -   P  A  L  *  L  K  D  L  I  N  K  R  I  L  *  T  T  Y  M  P
22561 - CAAAACCCACTCAGCGGCCAGACCTAAAATTGTCAAGTCTAGCTTGTACGATGAAATCGT -
22620
      - Q  N  P  L  S  G  Q  T  *  N  C  Q  V  *  L  V  R  *  N  R
      -  K  T  H  S  A  A  R  P  K  I  V  K  S  S  L  Y  D  E  I  V
      -   K  P  T  Q  R  P  D  L  K  L  S  S  L  A  C  T  M  K  S  S
22621 - CACCTGAATGGTTTCAAGAGCTGGATAAGAATCAAGGGAGTCTAATCCACTTAAACAAAT -
22680
      - H  L  N  G  F  K  S  W  I  R  I  K  G  V  *  S  T  *  T  N
      -  T  *  M  V  S  R  A  G  *  E  S  R  E  S  N  P  L  K  Q  M
      -   P  E  W  F  Q  E  L  D  K  N  Q  G  S  L  I  H  L  N  K  C
22681 - GCTGCAAGGAAAAGAACCTTCACAGAAATCCATAGTAGTAACGTTAGACGAATTAAGATA -
22740
      - A  A  R  K  R  T  F  T  E  I  H  S  S  N  V  R  R  I  K  I
      -  L  Q  G  K  E  P  S  Q  K  S  I  V  V  T  L  D  E  L  R  Y
      -   C  K  E  K  N  L  H  R  N  P  *  *  *  R  *  T  N  *  D  T
22741 - CAATTCTCTAACGCCATTACAATAAGAAGGAGCACCAAAATTAGATAAGAGTACACCAAA -
22800
      - Q  F  S  N  A  I  T  I  R  R  S  T  K  I  R  *  E  Y  T  K
      -  N  S  L  T  P  L  Q  *  E  G  A  P  K  L  D  K  S  T  P  K
      -   I  L  *  R  H  Y  N  K  K  E  H  Q  N  *  I  R  V  H  Q  K
22801 - AGCAGCAGTTACACAGATTAGAGAACCTAAGCAAATACTTAACAACAATAGCCACATAGC -
22860
      - S  S  S  Y  T  D  *  R  T  *  A  N  T  *  Q  Q  *  P  H  S
      -  A  A  V  T  Q  I  R  E  P  K  Q  I  L  N  N  N  S  H  I  A
      -   Q  Q  L  H  R  L  E  N  L  S  K  Y  L  T  T  I  A  T  *  R
22861 - GATTGTGAACAATTTAGAAAATTTGGGTGACTTCACATAATTAATGCCGGCATCCAAACA -
22920
      - D  C  E  Q  F  R  K  F  G  *  L  H  I  I  N  A  G  I  Q  T
      -  I  V  N  N  L  E  N  L  G  D  F  T  *  L  M  P  A  S  K  H
      -   L  *  T  I  *  K  I  W  V  T  S  H  N  *  C  R  H  P  N  I
22921 - TAATTTAGCAACACTCTTAACACTATTTTTAGCAATAGTTGTAGGTAGTGAAGCTCTAAT -
22980
      - *  F  S  N  T  L  N  T  I  F  S  N  S  C  R  *  *  S  S  N
      -  N  L  A  T  L  L  T  L  F  L  A  I  V  V  G  S  E  A  L  I
      -   I  *  Q  H  S  *  H  Y  F  *  Q  *  L  *  V  V  K  L  *  F
22981 - TCTAGAATTGGTACTTTTAGTAAAAGTACACAATTGGAACAATAATGTAAACACATAAGG -
```

23040
-S R I G T F S K S T Q L E Q * C K H I R
- L E L V L L V K V H N W N N N V N T * G
- * N W Y F * * K Y T I G T I M * T H K A

Figure 17-FF

```
23041 - CATATAATTGTTAAACACACGTTGTGCTAATCTCTTAGCGCAATTTGATGTTGTAATTGC -
23100
      - H  I  I  V  K  H  T  L  C  *  S  L  S  A  I  *  C  C  N  C
      -  I  *  L  L  N  T  R  C  A  N  L  L  A  Q  F  D  V  V  I  A
      -   Y  N  C  *  T  H  V  V  L  I  S  *  R  N  L  M  L  *  L  L
23101 - TGCTTGTCCTAAGAATGGTTTGACATAAGCCAAAATTTTACTCCAAGGAACACTATTAAT -
23160
      - C  L  S  *  E  W  F  D  I  S  Q  N  F  T  P  R  N  T  I  N
      -  A  C  P  K  N  G  L  T  *  A  K  I  L  L  Q  G  T  L  L  I
      -   L  V  L  R  M  V  *  H  K  P  K  F  Y  S  K  E  H  Y  *  L
23161 - TGCAGCAATACCATGAGTGGCAATTGTTTTAAACCTAAGGCTAGTGAAAGCTCATTAGG -
23220
      - C  S  N  T  M  S  G  N  C  F  *  T  *  G  *  *  K  L  I  R
      -  A  A  I  P  *  V  A  I  V  F  K  P  K  A  S  E  S  S  L  G
      -   Q  Q  Y  H  E  W  Q  L  F  L  N  L  R  L  V  K  A  H  *  V
23221 - TTTCTTAATGGTAATGCTTGTGTTTTCCACATAAGCAGCCATAAGATCCTCATGACCTAA -
23280
      - F  L  N  G  N  A  C  V  F  H  I  S  S  H  K  I  L  M  T  *
      -  F  L  M  V  M  L  V  F  S  T  *  A  A  I  R  S  S  *  P  N
      -   S  *  W  *  C  L  C  F  P  H  K  Q  P  *  D  P  H  D  L  T
23281 - CTCTTGTGTTACTTTAACACCTTCATCTGATGGTTTAAGTATGACATTGCCTACAACTTC -
23340
      - L  L  C  Y  F  N  T  F  I  *  W  F  K  Y  D  I  A  Y  N  F
      -  S  C  V  T  L  T  P  S  S  D  G  L  S  M  T  L  P  T  T  S
      -   L  V  L  L  *  H  L  H  L  M  V  *  V  *  H  C  L  Q  L  R
23341 - GGTAGTTTTCACGTCACACTCTATGACTTCCTTCTGTATGGTAGGATTTTCCACTACTTC -
23400
      - G  S  F  H  V  T  L  Y  D  F  L  L  Y  G  R  I  F  H  Y  F
      -  V  V  F  T  S  H  S  M  T  S  F  C  M  V  G  F  S  T  T  S
      -   *  F  S  R  H  T  L  *  L  P  S  V  W  *  D  F  P  L  L  L
23401 - TTCAGAGGTGGGTTGTTGACTTTCACAAGCAAGATTGTCCATTCCTTGTGTGTCTTCTAC -
23460
      - F  R  G  G  L  L  T  F  T  S  K  I  V  H  S  L  C  V  F  Y
      -  S  E  V  G  C  *  L  S  Q  A  R  L  S  I  P  C  V  S  S  T
      -   Q  R  W  V  V  D  F  H  K  Q  D  C  P  F  L  V  C  L  L  L
23461 - TGCCAGAACTTCAAATGAATTTGAAGTATCTACTGGCTTTGTACTCCAAAGACAACGTAA -
23520
      - C  Q  N  F  K  *  I  *  S  I  Y  W  L  C  T  P  K  T  T  *
      -  A  R  T  S  N  E  F  E  V  S  T  G  F  V  L  Q  R  Q  R  K
      -   P  E  L  Q  M  N  L  K  Y  L  L  A  L  Y  S  K  D  N  V  N
23521 - ACACCAAGTGTTTGGTTTGAACGTTGTCTTGGTTGTAGCCTGGTTAATGTGCCAAACAAT -
23580
      - T  P  S  V  W  F  E  R  C  L  G  C  S  L  V  N  V  P  N  N
      -  H  Q  V  F  G  L  N  V  V  L  V  V  A  W  L  M  C  Q  T  I
      -   T  K  C  L  V  *  T  L  S  W  L  *  P  G  *  C  A  K  Q  L
23581 - TGGCTTATGCAGTAATTTAGCACCTTTCTTGAAACTCGCTGAATAGTGTCTATAGTCAAT -
23640
      - W  L  M  Q  *  F  S  T  F  L  E  T  R  *  I  V  S  I  V  N
      -  G  L  C  S  N  L  A  P  F  L  K  L  A  E  *  C  L  *  S  I
      -   A  Y  A  V  I  *  H  L  S  *  N  S  L  N  S  V  Y  S  Q  *
23641 - AGCCACTACATCGCCATTCAAGTCTGGGAAGAATGTGACAGATAGCTCTCGTGAAGCTGG -
23700
      - S  H  Y  I  A  I  Q  V  W  E  E  C  D  R  *  L  S  *  S  W
      -  A  T  T  S  P  F  K  S  G  K  N  V  T  D  S  S  R  E  A  G
      -   P  L  H  R  H  S  S  L  G  R  M  *  Q  I  A  L  V  K  L  A
23701 - CTTTGTGAAGCCTGTCATTTGATTTAAATCATCAGCAAATTTTGTGTTAGAACATGTGAG -
23760
```

```
      - L  C  E  A  C  H  L  I  *  I  I  S  K  F  C  V  R  T  C  E
      -  F  V  K  P  V  I  *  F  K  S  S  A  N  F  V  L  E  H  V  S
      -   L  *  S  L  S  F  D  L  N  H  Q  Q  I  L  C  *  N  M  *  V
23761 - TTTGAAATTATCAAAACTCGCATTTGGTAATGGTTGAGTTGGTACAAGGTCTATAGGCTG -
23820
      - F  E  I  I  K  T  R  I  W  *  W  L  S  W  Y  K  V  Y  R  L
      -  L  K  L  S  K  L  A  F  G  N  G  *  V  G  T  R  S  I  G  C
      -   *  N  Y  Q  N  S  H  L  V  M  V  E  L  V  Q  G  L  *  A  A
```

Figure 17-GG

```
23821 - CTCTGTATAGTAAGCATTATCCTTTTTATAATACCCATCCAATTTTGGTTCAATCTCTGT -
23880
      -L  C  I  V  S  I  I  L  F  I  I  P  I  Q  F  W  F  N  L  C
      - S  V  *  *  A  L  S  F  L  *  Y  P  S  N  F  G  S  I  S  V
      -  L  Y  S  K  H  Y  P  F  Y  N  T  H  P  I  L  V  Q  S  L  C
23881 - GTAAGTAACTCCATCGAGTTTATACGACACAGGCTTGATGGTTGTAGTGTAAGATGTTTC -
23940
      -V  S  N  S  I  E  F  I  R  H  R  L  D  G  C  S  V  R  C  F
      - *  V  T  P  S  S  L  Y  D  T  G  L  M  V  V  V  *  D  V  S
      -  K  *  L  H  R  V  Y  T  T  Q  A  *  W  L  *  C  K  M  F  P
23941 - CTTGTAGAAAACATCAGTCACTGGTCCTTTGTACTCTGACATCTTTGTAAGGTGAGCTCC -
24000
      -L  V  E  N  I  S  H  W  S  F  V  L  *  H  L  C  K  V  S  S
      - L  *  K  T  S  V  T  G  P  L  Y  S  D  I  F  V  R  *  A  P
      -  C  R  K  H  Q  S  L  V  L  C  T  L  T  S  L  *  G  E  L  R
24001 - GTCAATACGATAGAGGGTCTCCTTAGCAGTTATATGAGTGTAATGACCACACTGATAGTT -
24060
      -V  N  T  I  E  G  L  L  S  S  Y  M  S  V  M  T  T  L  I  V
      - S  I  R  *  R  V  S  L  A  V  I  *  V  *  *  P  H  *  *  L
      -  Q  Y  D  R  G  S  P  *  Q  L  Y  E  C  N  D  H  T  D  S  Y
24061 - ACCAGTGTACTCATTCGCACATAAGAATGTACCTTGCTGTAATTTATACTCAGCAGGTGG -
24120
      -T  S  V  L  I  R  T  *  E  C  T  L  L  *  F  I  L  S  R  W
      - P  V  Y  S  F  A  H  K  N  V  P  C  C  N  L  Y  S  A  G  G
      -  Q  C  T  H  S  H  I  R  M  Y  L  A  V  I  Y  T  Q  Q  V  V
24121 - TGCAGACATCATAACAAAAGAAGACTCTTGTTGTACTAGATATTGTGTAGCATCACGACC -
24180
      -C  R  H  H  N  K  R  R  L  L  L  Y  *  I  L  C  S  I  T  T
      - A  D  I  I  T  K  E  D  S  C  C  T  R  Y  C  V  A  S  R  P
      -  Q  T  S  *  Q  K  K  T  L  V  V  L  D  I  V  *  H  H  D  H
24181 - ACACACACATGGAATGGAAACACCTGTCTTAAGATTATCATAAGATAGAGTACCCATATA -
24240
      -T  H  T  W  N  G  N  T  C  L  K  I  I  I  R  *  S  T  H  I
      - H  T  H  G  M  E  T  P  V  L  R  L  S  *  D  R  V  P  I  Y
      -  T  H  M  E  W  K  H  L  S  *  D  Y  H  K  I  E  Y  P  Y  T
24241 - CATCACAGCTTCTACACCCGTTAAGGTAGTAGTTTTCTGACCACAATGTTTACACACCAC -
24300
      -H  H  S  F  Y  T  R  *  G  S  S  F  L  T  T  M  F  T  H  H
      - I  T  A  S  T  P  V  K  V  V  V  F  *  P  Q  C  L  H  T  T
      -  S  Q  L  L  H  P  L  R  *  *  F  S  D  H  N  V  Y  T  P  H
24301 - ATTAAGAACTCGCTTTGCAGATTCCAAATTAGCATGCTGTAGAAGATGGGTCATAGTTTC -
24360
      -I  K  N  S  L  C  R  F  Q  I  S  M  L  *  K  M  G  H  S  F
      - L  R  T  R  F  A  D  S  K  L  A  C  C  R  R  W  V  I  V  S
      -  *  E  L  A  L  Q  I  P  N  *  H  A  V  E  D  G  S  *  F  L
24361 - TCTGACATCACCAAGCTCGCCAACAGTTTTATTACTGTAAGCGAGTATGAGTGCACAAAA -
24420
      -S  D  I  T  K  L  A  N  S  F  I  T  V  S  E  Y  E  C  T  K
      - L  T  S  P  S  S  P  T  V  L  L  L  *  A  S  M  S  A  Q  K
      -  *  H  H  Q  A  R  Q  Q  F  Y  Y  C  K  R  V  *  V  H  K  S
24421 - GTTAGCAGCATCACCAGCACGGGCTCTATAATAAGCCTCTTGAAGTGCTGGTGCATTGAA -
24480
      -V  S  S  I  T  S  T  G  S  I  I  S  L  L  K  C  W  C  I  E
      - L  A  A  S  P  A  R  A  L  *  *  A  S  *  S  A  G  A  L  N
      -  *  Q  H  H  Q  H  G  L  Y  N  K  P  L  E  V  L  V  H  *  I
24481 - TTTGACTTCAAGCTGTTGAAGTGCTAATAAAACACTAGACAAATAACAATTGTTATCAGC -
```

24540
-F D F K L L K C * * N T R Q I T I V I S
- L T S S C * S A N K T L D K * Q L L S A
- * L Q A V E V L I K H * T N N N C Y Q P

Figure 17-HH

```
24541 - CCATTTAATTGAAGTTAAACCACCAACTTGAGGAAATTTCCATTTCTTTGTGTGGTTTAA -
24600
      - P  F  N  *  S  *  T  T  N  L  R  K  F  P  F  L  C  V  V  *
      -  H  L  I  E  V  K  P  P  T  *  G  N  F  H  F  F  V  W  F  K
      -   I  *  L  K  L  N  H  Q  L  E  E  I  S  I  S  L  C  G  L  K
24601 - AGCAGACATGTACCTACCAAGAAAACTCTCATCAAGAGTATGGTAGTACTCGAAAGCTTC -
24660
      - S  R  H  V  P  T  K  K  T  L  I  K  S  M  V  V  L  E  S  F
      -  A  D  M  Y  L  P  R  K  L  S  S  R  V  W  *  Y  S  K  A  S
      -   Q  T  C  T  Y  Q  E  N  S  H  Q  E  Y  G  S  T  R  K  L  H
24661 - ACTACGTAGTGTGTCATCACTAGGTAGTACAAAGAAAGTCTTACCCTCATGATTTACATG -
24720
      - T  T  *  C  V  I  T  R  *  Y  K  E  S  L  T  L  M  I  Y  M
      -  L  R  S  V  S  S  L  G  S  T  K  K  V  L  P  S  *  F  T  *
      -   Y  V  V  C  H  H  *  V  V  Q  R  K  S  Y  P  H  D  L  H  E
24721 - AGGTTTAATTTTTGTAACATCAGCACCATCCAAGTATGTTGGACCAAACTGCTGTCCATA -
24780
      - R  F  N  F  C  N  I  S  T  I  Q  V  C  W  T  K  L  L  S  I
      -  G  L  I  F  V  T  S  A  P  S  K  Y  V  G  P  N  C  C  P  Y
      -   V  *  F  L  *  H  Q  H  H  P  S  M  L  D  Q  T  A  V  H  M
24781 - TGTCATAGACATATCCACAAGCTGTGTGTGGAGATTAGTGTTGTCCACAGTTGTGAACAC -
24840
      - C  H  R  H  I  H  K  L  C  V  E  I  S  V  V  H  S  C  E  H
      -  V  I  D  I  S  T  S  C  V  W  R  L  V  L  S  T  V  V  N  T
      -   S  *  T  Y  P  Q  A  V  C  G  D  *  C  C  P  Q  L  *  T  L
24841 - TTTTATAGTCTTAACCTCCCGCAGGGATAAGAGACTCTTTAGTTTGTCAAGTGAAAGAAC -
24900
      - F  Y  S  L  N  L  P  Q  G  *  E  T  L  *  F  V  K  *  K  N
      -  F  I  V  L  T  S  R  R  D  K  R  L  F  S  L  S  S  E  R  T
      -   L  *  S  *  P  P  A  G  I  R  D  S  L  V  C  Q  V  K  E  P
24901 - CTCACCGTCAAGATGAAACTCGACGGGGCTCTCCAGAGTGTGGTACACAATTTTGTCACC -
24960
      - L  T  V  K  M  K  L  D  G  A  L  Q  S  V  V  H  N  F  V  T
      -  S  P  S  R  *  N  S  T  G  L  S  R  V  W  Y  T  I  L  S  P
      -   H  R  Q  D  E  T  R  R  G  S  P  E  C  G  T  Q  F  C  H  H
24961 - ACGCTTAAGAAATTCAACACCTAACTCTGTACGCTGTCCTGAATAGGACCAATCTCTGTA -
25020
      - T  L  K  K  F  N  T  *  L  C  T  L  S  *  I  G  P  I  S  V
      -  R  L  R  N  S  T  P  N  S  V  R  C  P  E  *  D  Q  S  L  *
      -   A  *  E  I  Q  H  L  T  L  Y  A  V  L  N  R  T  N  L  C  K
25021 - AGAGCCAGCCAAAGAAACTGTTTCTACAAAGTGCTCCTCAGATGTCTTTGATGACGAAGT -
25080
      - R  A  S  Q  R  N  C  F  Y  K  V  L  L  R  C  L  *  *  R  S
      -  E  P  A  K  E  T  V  S  T  K  C  S  S  D  V  F  D  D  E  V
      -   S  Q  P  K  K  L  F  L  Q  S  A  P  Q  M  S  L  M  T  K  *
25081 - GAGGTATCCATTATATGTAGTAACAGCATCTGGTGATGATACTGACACTACGGCAGGAGC -
25140
      - E  V  S  I  I  C  S  N  S  I  W  *  *  Y  *  H  Y  G  R  S
      -  R  Y  P  L  Y  V  V  T  A  S  G  D  D  T  D  T  T  A  G  A
      -   G  I  H  Y  M  *  *  Q  H  L  V  M  I  L  T  L  R  Q  E  L
25141 - TTTAAGAGAACGCATACAGCGCGCAGCCTCTTCAAGATTAAAACCATGTGTCACATAACC -
25200
      - F  K  R  T  H  T  A  R  S  L  F  K  I  K  T  M  C  H  I  T
      -  L  R  E  R  I  Q  R  A  A  S  S  R  L  K  P  C  V  T  *  P
      -   *  E  N  A  Y  S  A  Q  P  L  Q  D  *  N  H  V  S  H  N  Q
```

25201 - AATTGGCATTGTGACAAGCGGCTCATTTAGAGAGTTCAGCTTCGTAATAATAGAAGCTAC - 25260

- N W H C D K R L I * R V Q L R N N R S Y
- I G I V T S G S F R E F S F V I I E A T
- L A L * Q A A H L E S S A S * * * K L Q

Figure 17-II

```
25261 - AGGCTCTTTACTAGTATAAAAGAAGAATCGGACACCATAGTCAACGATGCCCTCTTGAAT -
25320
      -R  L  F  T  S  I  K  E  E  S  D  T  I  V  N  D  A  L  L  N
      - G  S  L  L  V  *  K  K  N  R  T  P  *  S  T  M  P  S  *  I
      -  A  L  Y  *  Y  K  R  R  I  G  H  H  S  Q  R  C  P  L  E  F
25321 - TTTAATTCCTTTATACTTACGTTGGATGGTTGCCATTATGGCTCTAACATCCATGCATAT -
25380
      -F  N  S  F  I  L  T  L  D  G  C  H  Y  G  S  N  I  H  A  Y
      - L  I  P  L  Y  L  R  W  M  V  A  I  M  A  L  T  S  M  H  I
      -  *  F  L  Y  T  Y  V  G  W  L  P  L  W  L  *  H  P  C  I  *
25381 - AGGCATTAATTTTCTTGTCTCTTCAGCATGAGCAAGCATTTCTCTCAAATTCCAGGATAC -
25440
      -R  H  *  F  S  C  L  F  S  M  S  K  H  F  S  Q  I  P  G  Y
      - G  I  N  F  L  V  S  S  A  *  A  S  I  S  L  K  F  Q  D  T
      -  A  L  I  F  L  S  L  Q  H  E  Q  A  F  L  S  N  S  R  I  Q
25441 - AGTTCCTAGAATCTCTTCCTTAGCATTAGGTGCTTCTGAAGGTAGTACATAAAATGCAGA -
25500
      -S  S  *  N  L  F  L  S  I  R  C  F  *  R  *  Y  I  K  C  R
      - V  P  R  I  S  S  L  A  L  G  A  S  E  G  S  T  *  N  A  D
      -  F  L  E  S  L  P  *  H  *  V  L  L  K  V  V  H  K  M  Q  I
25501 - TTTGCATTTCTTAAGAGCAGTCTTAGCTTCCTCAAGTGTATAACCAGCACATCCTTGTCC -
25560
      -F  A  F  L  K  S  S  L  S  F  L  K  C  I  T  S  T  S  L  S
      - L  H  F  L  R  A  V  L  A  S  S  S  V  *  P  A  H  P  C  P
      -  C  I  S  *  E  Q  S  *  L  P  Q  V  Y  N  Q  H  I  L  V  Q
25561 - AGGGTACGTGGTTATATACTCATCAACTGGCACTTTCTTCAAAGCTCTTGAGAGCATCTC -
25620
      -R  V  R  G  Y  I  L  I  N  W  H  F  L  Q  S  S  *  E  H  L
      - G  Y  V  V  I  Y  S  S  T  G  T  F  F  K  A  L  E  S  I  S
      -  G  T  W  L  Y  T  H  Q  L  A  L  S  S  K  L  L  R  A  S  Q
25621 - AGTAGTGCCACCAGCCTTTTTGGAGGGTATTACAACACAAGTGATATCACCACTAGTGAT -
25680
      -S  S  A  T  S  L  F  G  G  Y  Y  N  T  S  D  I  T  T  S  D
      - V  V  P  P  A  F  L  E  G  I  T  T  Q  V  I  S  P  L  V  I
      -  *  C  H  Q  P  F  W  R  V  L  Q  H  K  *  Y  H  H  *  *  *
25681 - AACATCACCTACCATGTAAGGTGCATCCTTCTCAAGGAAAGACATATCTTCACCTCTAAG -
25740
      -N  I  T  Y  H  V  R  C  I  L  L  K  E  R  H  I  F  T  S  K
      - T  S  P  T  M  *  G  A  S  F  S  R  K  D  I  S  S  P  L  S
      -  H  H  L  P  C  K  V  H  P  S  Q  G  K  T  Y  L  H  L  *  A
25741 - CATGTTCTGAGAATCATGGTAAAGCTTACCATTGATATCAGCAAACAAGAGTAACTTATT -
25800
      -H  V  L  R  I  M  V  K  L  T  I  D  I  S  K  Q  E  *  L  I
      - M  F  *  E  S  W  *  S  L  P  L  I  S  A  N  K  S  N  L  L
      -  C  S  E  N  H  G  K  A  Y  H  *  Y  Q  Q  T  R  V  T  Y  W
25801 - GGTAAGAAACTTAGTTTCTTCCAGTGTTGTGGTAACCTCATCAATGCAGGCCTTAATTTT -
25860
      -G  K  K  L  S  F  F  Q  C  C  G  N  L  I  N  A  G  L  N  F
      - V  R  N  L  V  S  S  S  V  V  V  T  S  S  M  Q  A  L  I  F
      -  *  E  T  *  F  L  P  V  L  W  *  P  H  Q  C  R  P  *  F  L
25861 - TGGCTTCACATCGACAGGCTTCTGTACGACAGATTTCTCCTCAGTTTTGGAATCTTCTGT -
25920
      -W  L  H  I  D  R  L  L  Y  D  R  F  L  L  S  F  G  I  F  C
      - G  F  T  S  T  G  F  C  T  T  D  F  S  S  V  L  E  S  S  V
      -  A  S  H  R  Q  A  S  V  R  Q  I  S  P  Q  F  W  N  L  L  C
25921 - GTTTGGTGGCTCCTCTTGTTTAGGTGCTTCCACTCTAGGCTTCAGGTTATCAAGATAATC -
25980
```

```
      -V  W  W  L  L  L  F  R  C  F  H  S  R  L  Q  V  I  K  I  I
      -  F  G  G  S  S  C  L  G  A  S  T  L  G  F  R  L  S  R  *  S
      -   L  V  A  P  L  V  *  V  L  P  L  *  A  S  G  Y  Q  D  N  P
25981 - CATGACAACCTGCTCATAAAGAGCTTTGTCATTGACTGCAATATAAACCTGTGTACGAAC -
26040
      -H  D  N  L  L  I  K  S  F  V  I  D  C  N  I  N  L  C  T  N
      -  M  T  T  C  S  *  R  A  L  S  L  T  A  I  *  T  C  V  R  T
      -   *  Q  P  A  H  K  E  L  C  H  *  L  Q  Y  K  P  V  Y  E  P
```

Figure 17-JJ

```
26041 - CGTCTGCACGCACACTTGTAAAGACTGAAGTGGTTTAGCACCAAATATGCCTGCTGACAA -
26100
      - R  L  H  A  H  L  *  R  L  K  W  F  S  T  K  Y  A  C  *  Q
      -  V  C  T  H  T  C  K  D  *  S  G  L  A  P  N  M  P  A  D  N
      -   S  A  R  T  L  V  K  T  E  V  V  *  H  Q  I  C  L  L  T  T
26101 - CAATGGTGCAAGTAAGATGTCCTGTGAATTGAAATTTTCATATGCTGCCTTAAGAAGCTG -
26160
      - Q  W  C  K  *  D  V  L  *  I  E  I  F  I  C  C  L  K  K  L
      -  N  G  A  S  K  M  S  C  E  L  K  F  S  Y  A  A  L  R  S  W
      -   M  V  Q  V  R  C  P  V  N  *  N  F  H  M  L  P  *  E  A  G
26161 - GATGTCCTCACCTGCATTTAGGTTAGGTCCAACAACATGCAGACACTTCTTAGCAAGATT -
26220
      - D  V  L  T  C  I  *  V  R  S  N  N  M  Q  T  L  L  S  K  I
      -  M  S  S  P  A  F  R  L  G  P  T  T  C  R  H  F  L  A  R  L
      -   C  P  H  L  H  L  G  *  V  Q  Q  H  A  D  T  S  *  Q  D  Y
26221 - ATGTCCAGAAAGCAAACAAGACCCTCCTACTGTAAGAGGGCCATTTAGCTTAATGTAATC -
26280
      - M  S  R  K  Q  T  R  P  S  Y  C  K  R  A  I  *  L  N  V  I
      -  C  P  E  S  K  Q  D  P  P  T  V  R  G  P  F  S  L  M  *  S
      -   V  Q  K  A  N  K  T  L  L  L  *  E  G  H  L  A  *  C  N  H
26281 - ATCACTCTCCTTTTGCATGGCACCATTGGTTGCCTTGTTGAGTGCACCTGCTACACCACC -
26340
      - I  T  L  L  L  H  G  T  I  G  C  L  V  E  C  T  C  Y  T  T
      -  S  L  S  F  C  M  A  P  L  V  A  L  L  S  A  P  A  T  P  P
      -   H  S  P  F  A  W  H  H  W  L  P  C  *  V  H  L  L  H  H  H
26341 - ACCATGTTTCAGGTGTATGTTAGCAGCATTTACAATCACCATAGGATTAGCACTTTGTGC -
26400
      - T  M  F  Q  V  Y  V  S  S  I  Y  N  H  H  R  I  S  T  L  C
      -  P  C  F  R  C  M  L  A  A  F  T  I  T  I  G  L  A  L  C  A
      -   H  V  S  G  V  C  *  Q  H  L  Q  S  P  *  D  *  H  F  V  P
26401 - CTCCTTAACGATGTCAACACATTTAATGGCAACATTGTCAGTAAGTTTTAAATAACCAGT -
26460
      - L  L  N  D  V  N  T  F  N  G  N  I  V  S  K  F  *  I  T  S
      -  S  L  T  M  S  T  H  L  M  A  T  L  S  V  S  F  K  *  P  V
      -   P  *  R  C  Q  H  I  *  W  Q  H  C  Q  *  V  L  N  N  Q  *
26461 - AAACTGATTAACTGGTTCTTCAGGTGTAGGTTCTGGTTCTGGCTCAATCTCTGATTGCTC -
26520
      - K  L  I  N  W  F  F  R  C  R  F  W  F  W  L  N  L  *  L  L
      -  N  *  L  T  G  S  S  G  V  G  S  G  S  G  S  I  S  D  C  S
      -   T  D  *  L  V  L  Q  V  *  V  L  V  L  A  Q  S  L  I  A  Q
26521 - AGTAGTATCATCCAGCCAGTCTTCCTCTTCTTCTTCCTCAACTCGAACTGTTTCAGCTGA -
26580
      - S  S  I  I  Q  P  V  F  L  F  F  F  L  N  S  N  C  F  S  *
      -  V  V  S  S  S  Q  S  S  S  S  S  S  S  S  T  R  T  V  S  A  E
      -   *  Y  H  P  A  S  L  P  L  L  L  P  Q  L  E  L  F  Q  L  R
26581 - GGCACCAAATTCCAGAGGGAGACCTTGATAATCATCCTCTGTACCGTACTCATGTTCACA -
26640
      - G  T  K  F  Q  R  E  T  L  I  I  I  L  C  T  V  L  M  F  T
      -  A  P  N  S  R  G  R  P  *  *  S  S  S  V  P  Y  S  C  S  Q
      -   H  Q  I  P  E  G  D  L  D  N  H  P  L  Y  R  T  H  V  H  R
26641 - GGTTTCATCAATTTCTTCTTCCTCACACTCTGCATCGTCCTCTTCTTCCTCATCTGGAGG -
26700
      - G  F  I  N  F  F  F  L  T  L  C  I  V  L  F  F  L  I  W  R
      -  V  S  S  I  S  S  S  S  H  S  A  S  S  S  S  S  S  S  G  G
      -   F  H  Q  F  L  L  P  H  T  L  H  R  P  L  L  P  H  L  E  G
```

26701 - GTAAAAGGAACAATACATACGTGATGAAAAGTTTTCTTCACCAGCATCATCAAATAAGTA - 26760

-V K G T I H T * * K V F F T S I I K * V
- * K E Q Y I R D E K F S S P A S S N K *
- K R N N T Y V M K S F L H Q H H Q I S R

Figure 17-KK

```
 26761 - GAATGTAGCTACACTCCACTCATCAAGATCAATACCCATGTTGGTAAGGAGATCAGAAAC -
26820
       - E  C  S  Y  T  P  L  I  K  I  N  T  H  V  G  K  E  I  R  N
       -  N  V  A  T  L  H  S  S  R  S  I  P  M  L  V  R  R  S  E  T
       -   M  *  L  H  S  T  H  Q  D  Q  Y  P  C  W  *  G  D  Q  K  L
 26821 - TGGTTGTAAAGTCTTCACAACAGCCTCTGCTACAACACATGCAAACTCAGTAACTTCGGT -
26880
       - W  L  *  S  L  H  N  S  L  C  Y  N  T  C  K  L  S  N  F  G
       -  G  C  K  V  F  T  T  A  S  A  T  T  H  A  N  S  V  T  S  V
       -   V  V  K  S  S  Q  Q  P  L  L  Q  H  M  Q  T  Q  *  L  R  Y
 26881 - ACCGGATTCAACAGTGTAGACAGAGCACTTTTCATTAAGCACTTTGTCAACACGTTCATC -
26940
       - T  G  F  N  S  V  D  R  A  L  F  I  K  H  F  V  N  T  F  I
       -  P  D  S  T  V  *  T  E  H  F  S  L  S  T  L  S  T  R  S  S
       -   R  I  Q  Q  C  R  Q  S  T  F  H  *  A  L  C  Q  H  V  H  Q
 26941 - AAGCTCAAATGTGATTCTCACATTCTTGTAACCTTGAACTTCCCAAACAGTATCTTCTCC -
27000
       - K  L  K  C  D  S  H  I  L  V  T  L  N  F  P  N  S  I  F  S
       -  S  S  N  V  I  L  T  F  L  *  P  *  T  S  Q  T  V  S  S  P
       -   A  Q  M  *  F  S  H  S  C  N  L  E  L  P  K  Q  Y  L  L  Q
 27001 - AAAGGTTACACCTTTAATTGGTGCACCCCCTTTTAAGCGAAAGACATTGTTTGTAGCCAG -
27060
       - K  G  Y  T  F  N  W  C  T  P  F  *  A  K  D  I  V  C  S  Q
       -  K  V  T  P  L  I  G  A  P  P  F  K  R  K  T  L  F  V  A  S
       -   R  L  H  L  *  L  V  H  P  L  L  S  E  R  H  C  L  *  P  V
 27061 - TAAACCAGGAGACAATGCGCAGTATTGTTCTTTGTCCTTAATCTCTAAGAGCATGAGGCC -
27120
       - *  T  R  R  Q  C  A  V  L  F  F  V  L  N  L  *  E  H  E  A
       -  K  P  G  D  N  A  Q  Y  C  S  L  S  L  I  S  K  S  M  R  P
       -   N  Q  E  T  M  R  S  I  V  L  C  P  *  S  L  R  A  *  G  H
 27121 - ATTTACACAGACTGGTGTGCCGACGATAGCTCCATTTGTGAAGCTATCAACGGGCGTCTC -
27180
       - I  Y  T  D  W  C  A  D  D  S  S  I  C  E  A  I  N  G  R  L
       -  F  T  Q  T  G  V  P  T  I  A  P  F  V  K  L  S  T  G  V  S
       -   L  H  R  L  V  C  R  R  *  L  H  L  *  S  Y  Q  R  A  S  R
 27181 - GAGTGCTTCGAGTTCACCGTTCTTGAGAACAACCTCCTCAGAGGTAAGTACTGTGTCATG -
27240
       - E  C  F  E  F  T  V  L  E  N  N  L  L  R  G  K  Y  C  V  M
       -  S  A  S  S  S  P  F  L  R  T  T  S  S  E  V  S  T  V  S  C
       -   V  L  R  V  H  R  S  *  E  Q  P  P  Q  R  *  V  L  C  H  V
 27241 - TGAATCACCTTCAAGAAAGGTTACTTCTTTTGGTGCCTTAAGAGGCATGAGTAGTTGCAG -
27300
       - *  I  T  F  K  K  G  Y  F  F  W  C  L  K  R  H  E  *  L  Q
       -  E  S  P  S  R  K  V  T  S  F  G  A  L  R  G  M  S  S  C  S
       -   N  H  L  Q  E  R  L  L  L  L  V  P  *  E  A  *  V  V  A  A
 27301 - CTGCTCCTTGCCACGTATACACTGACGGTAAAGTCCCTTGCTTTGAGCGATGAAGACTTC -
27360
       - L  L  L  A  T  Y  T  L  T  V  K  S  L  A  L  S  D  E  D  F
       -  C  S  L  P  R  I  H  *  R  *  S  P  L  L  *  A  M  K  T  S
       -   A  P  C  H  V  Y  T  D  G  K  V  P  C  F  E  R  *  R  L  H
 27361 - ACCTAAGTTGAGTGATCGCAACTTTGCGCCAGCGATAGTGACTTGATCAATGCACATTTC -
27420
       - T  *  V  E  *  S  Q  L  C  A  S  D  S  D  L  I  N  A  H  F
       -  P  K  L  S  D  R  N  F  A  P  A  I  V  T  *  S  M  H  I  S
       -   L  S  *  V  I  A  T  L  R  Q  R  *  *  L  D  Q  C  T  F  R
 27421 - GAGTGCCTTGTTAACAACATCAATGAAGCATTTTACACAATCCTTGATGTTATCTGAAGC -
27480
```

```
      - E  C  L  V  N  N  I  N  E  A  F  Y  T  I  L  D  V  I  *  S
      -  S  A  L  L  T  T  S  M  K  H  F  T  Q  S  L  M  L  S  E  A
      -   V  P  C  *  Q  H  Q  *  S  I  L  H  N  P  *  C  Y  L  K  Q
27481 - AACCTGTATTTGACCCTTGACGATGTCAAAAACACCTGTAATGAGAAATTTGAGAATCTC -
27540
      - N  L  Y  L  T  L  D  D  V  K  N  T  C  N  E  K  F  E  N  L
      -  T  C  I  *  P  L  T  M  S  K  T  P  V  M  R  N  L  R  I  S
      -   P  V  F  D  P  *  R  C  Q  K  H  L  *  *  E  I  *  E  S  P
```

Figure 17-LL

```
27541 - CCAAGCATCCTTGAGAAATTCAACTCCTGCACTAAGTTTCGCCTCAATCCATTCAAAGAT -
27600
      -P  S  I  L  E  K  F  N  S  C  T  K  F  R  L  N  P  F  K  D
      - Q  A  S  L  R  N  S  T  P  A  L  S  F  A  S  I  H  S  K  I
      -  K  H  P  *  E  I  Q  L  L  H  *  V  S  P  Q  S  I  Q  R  *
27601 - AGGCCTGAGTTTTTCAACAGTAGTGCCCAAAAGATTAGACAACCACTGAGAAGTCTGTTG -
27660
      -R  P  E  F  F  N  S  S  A  Q  K  I  R  Q  P  L  R  S  L  L
      - G  L  S  F  S  T  V  V  P  K  R  L  D  N  H  *  E  V  C  C
      -  A  *  V  F  Q  Q  *  C  P  K  D  *  T  T  T  E  K  S  V  V
27661 - TACAAGACCACCAGTTACATATGCCATAATAATGACACTGTTGGTGAGCAGGTCTGAAGT -
27720
      -Y  K  T  T  S  Y  I  C  H  N  N  D  T  V  G  E  Q  V  *  S
      - T  R  P  P  V  T  Y  A  I  I  M  T  L  L  V  S  R  S  E  V
      -  Q  D  H  Q  L  H  M  P  *  *  *  H  C  W  *  A  G  L  K  Y
27721 - ATAAACCATGGCGTCGACAAGACGTAATGACTGTTCAGAAATACCATCAAGTATGGTGAC -
27780
      -I  N  H  G  V  D  K  T  *  *  L  F  R  N  T  I  K  Y  G  D
      - *  T  M  A  S  T  R  R  N  D  C  S  E  I  P  S  S  M  V  T
      -  K  P  W  R  R  Q  D  V  M  T  V  Q  K  Y  H  Q  V  W  *  Q
27781 - AGCTGCTCTTTGCAAATCAGGAATTGAGTGGTTTGCTGCATCAAGTGTGCGCGCAAAAAT -
27840
      -S  C  S  L  Q  I  R  N  *  V  V  C  C  I  K  C  A  R  K  N
      - A  A  L  C  K  S  G  I  E  W  F  A  A  S  S  V  R  A  K  I
      -  L  L  F  A  N  Q  E  L  S  G  L  L  H  Q  V  C  A  Q  K  L
27841 - TGATCTGATAACACCAGCAGCCTGTGAGGGAAAACCACACAGTGGTGTTAAAACTGATCT -
27900
      -*  S  D  N  T  S  S  L  *  G  K  T  T  Q  W  C  *  N  *  S
      - D  L  I  T  P  A  A  C  E  G  K  P  H  S  G  V  K  T  D  L
      -  I  *  *  H  Q  Q  P  V  R  E  N  H  T  V  V  L  K  L  I  S
27901 - CTGTTGTCCAATGTTCCAAGCACCTTTTACGGGCTTTCCCTTGGTAACTTTATAGTTACC -
27960
      -L  L  S  N  V  P  S  T  F  Y  G  L  S  L  G  N  F  I  V  T
      - C  C  P  M  F  Q  A  P  F  T  G  F  P  L  V  T  L  *  L  P
      -  V  V  Q  C  S  K  H  L  L  R  A  F  P  W  *  L  Y  S  Y  R
27961 - GCAGGACTCAACAATGGTTTTGAAAGACTTGTAATCAAGACTCTTTATAGTGTCAATAAA -
28020
      -A  G  L  N  N  G  F  E  R  L  V  I  K  T  L  Y  S  V  N  K
      - Q  D  S  T  M  V  L  K  D  L  *  S  R  L  F  I  V  S  I  K
      -  R  T  Q  Q  W  F  *  K  T  C  N  Q  D  S  L  *  C  Q  *  R
28021 - GGCACTTGTAGAAGCAGAGAAAGATGCCAAAATGATGGCAACCTCTTCATTCAAATGAAA -
28080
      -G  T  C  R  S  R  E  R  C  Q  N  D  G  N  L  F  I  Q  M  K
      - A  L  V  E  A  E  K  D  A  K  M  M  A  T  S  S  F  K  *  K
      -  H  L  *  K  Q  R  K  M  P  K  *  W  Q  P  L  H  S  N  E  N
28081 - ATCGCCAACAATGTTAATGTTAACACGTTCACGACTCAGTATCTCAAGGAGATCCTCATT -
28140
      -I  A  N  N  V  N  V  N  T  F  T  T  Q  Y  L  K  E  I  L  I
      - S  P  T  M  L  M  L  T  R  S  R  L  S  I  S  R  R  S  S  F
      -  R  Q  Q  C  *  C  *  H  V  H  D  S  V  S  Q  G  D  P  H  S
28141 - CAAGGTCTCCACATTGTCACCAGTAATGCCAGTATGGCCTGAGCCAATATCAGCACTAGC -
28200
      -Q  G  L  H  I  V  T  S  N  A  S  M  A  *  A  N  I  S  T  S
      - K  V  S  T  L  S  P  V  M  P  V  W  P  E  P  I  S  A  L  A
      -  R  S  P  H  C  H  Q  *  C  Q  Y  G  L  S  Q  Y  Q  H  *  H
28201 - ACGAGGAACCCAGTAGGCACGCTTATTATAGCAGCCAACATAGGCAAACACACAGCCTCC -
```

28260

- T R N P V G T L I I A A N I G K H T A S
- R G T Q * A R L L * Q P T * A N T Q P P
- E E P S R H A Y Y S S Q H R Q T H S L Q

Figure 17-MM

```
 28261 - AAAACATCTAGTCCTACCTCCCTTGCGGAGTCGAGTTTCAATGTTTGAGTGGTTGTGATA -
28320
       -K  T  S  S  P  T  S  L  A  E  S  S  F  N  V  *  V  V  V  I
       -  K  H  L  V  L  P  P  L  R  S  R  V  S  M  F  E  W  L  *  *
       -   N  I  *  S  Y  L  P  C  G  V  E  F  Q  C  L  S  G  C  D  N
 28321 - ATCTGCAACACTATGCTCAGGTCCAATCTCTGGGTCTTGACAGGCAGGACATGGCATTTT -
28380
       -I  C  N  T  M  L  R  S  N  L  W  V  L  T  G  R  T  W  H  F
       -  S  A  T  L  C  S  G  P  I  S  G  S  *  Q  A  G  H  G  I  F
       -   L  Q  H  Y  A  Q  V  Q  S  L  G  L  D  R  Q  D  M  A  F  S
 28381 - CACTACAGCATTAGTAGGTAGGTACCCACATGTAGTAGGTCCTTCAATAACTAAATTTTC -
28440
       -H  Y  S  I  S  R  *  V  P  T  C  S  R  S  F  N  N  *  I  F
       -  T  T  A  L  V  G  R  Y  P  H  V  V  G  P  S  I  T  K  F  S
       -   L  Q  H  *  *  V  G  T  H  M  *  *  V  L  Q  *  L  N  F  Q
 28441 - AGTGCCACAATGTTCACAAGTGGCTTTCAGAAAGTCGCACGTCTGCCATGAAACTTCATC -
28500
       -S  A  T  M  F  T  S  G  F  Q  K  V  A  R  L  P  *  N  F  I
       -  V  P  Q  C  S  Q  V  A  F  R  K  S  H  V  C  H  E  T  S  S
       -   C  H  N  V  H  K  W  L  S  E  S  R  T  S  A  M  K  L  H  R
 28501 - GCAATGATTACATTTCATCAAGGTAGACAAGTGCATATTGTTACACTCCTGTGGAGATGC -
28560
       -A  M  I  T  F  H  Q  G  R  Q  V  H  I  V  T  L  L  W  R  C
       -  Q  *  L  H  F  I  K  V  D  K  C  I  L  L  H  S  C  G  D  A
       -   N  D  Y  I  S  S  R  *  T  S  A  Y  C  Y  T  P  V  E  M  Q
 28561 - AACAGGGTACACAGAGCGTATACGCCCCATGAAACCCTCAGTCTTTTTCTTTTCAACACG -
28620
       -N  R  V  H  R  A  Y  T  P  H  E  T  L  S  L  F  L  F  N  T
       -  T  G  Y  T  E  R  I  R  P  M  K  P  S  V  F  F  F  S  T  R
       -   Q  G  T  Q  S  V  Y  A  P  *  N  P  Q  S  F  S  F  Q  H  V
 28621 - TGGTTGAATGACTTTGACTTTTGAGTTAAGAGGAAACACAAACTTTGGGCATTCCCCTTT -
28680
       -W  L  N  D  F  D  F  *  V  K  R  K  H  K  L  W  A  F  P  F
       -  G  *  M  T  L  T  F  E  L  R  G  N  T  N  F  G  H  S  P  L
       -   V  E  *  L  *  L  L  S  *  E  E  T  Q  T  L  G  I  P  L  *
 28681 - GAAAGTGTCAAATTTCTTGGCACTCTTAATTTCGAAGGGTGTCTGGTGCTCGTAGCTCTT -
28740
       -E  S  V  K  F  L  G  T  L  N  F  E  G  C  L  V  L  V  A  L
       -  K  V  S  N  F  L  A  L  L  I  S  K  G  V  W  C  S  *  L  L
       -   K  C  Q  I  S  W  H  S  *  F  R  R  V  S  G  A  R  S  S  Y
 28741 - ATCAGAGCGCTCAGTGAACCAGGCAATTTCATGCTCATGGTCACGGCAGCAGTAGACACC -
28800
       -I  R  A  L  S  E  P  G  N  F  M  L  M  V  T  A  A  V  D  T
       -  S  E  R  S  V  N  Q  A  I  S  C  S  W  S  R  Q  Q  *  T  P
       -   Q  S  A  Q  *  T  R  Q  F  H  A  H  G  H  G  S  S  R  H  L
 28801 - TCTCTTCGACTCGATGTAATCAAGTTGTTCGGAAAGAGTGCACATTGACTTGCCCGCGCG -
28860
       -S  L  R  L  D  V  I  K  L  F  G  K  S  A  H  *  L  A  R  A
       -  L  F  D  S  M  *  S  S  C  S  E  R  V  H  I  D  L  P  A  R
       -   S  S  T  R  C  N  Q  V  V  R  K  E  C  T  L  T  C  P  R  V
 28861 - TGCGAGAAAATCTTTGATGCAATCAAGAGGGTACCCATCTGGGCCACAGAAATTGTTGTC -
28920
       -C  E  K  I  F  D  A  I  K  R  V  P  I  W  A  T  E  I  V  V
       -  A  R  K  S  L  M  Q  S  R  G  Y  P  S  G  P  Q  K  L  L  S
       -   R  E  N  L  *  C  N  Q  E  G  T  H  L  G  H  R  N  C  C  R
 28921 - GACATAGCGAGTGACTGCACCTCCATTGAGCTCACGAGTGAGTTCACGGAGTGCACCACT -
```

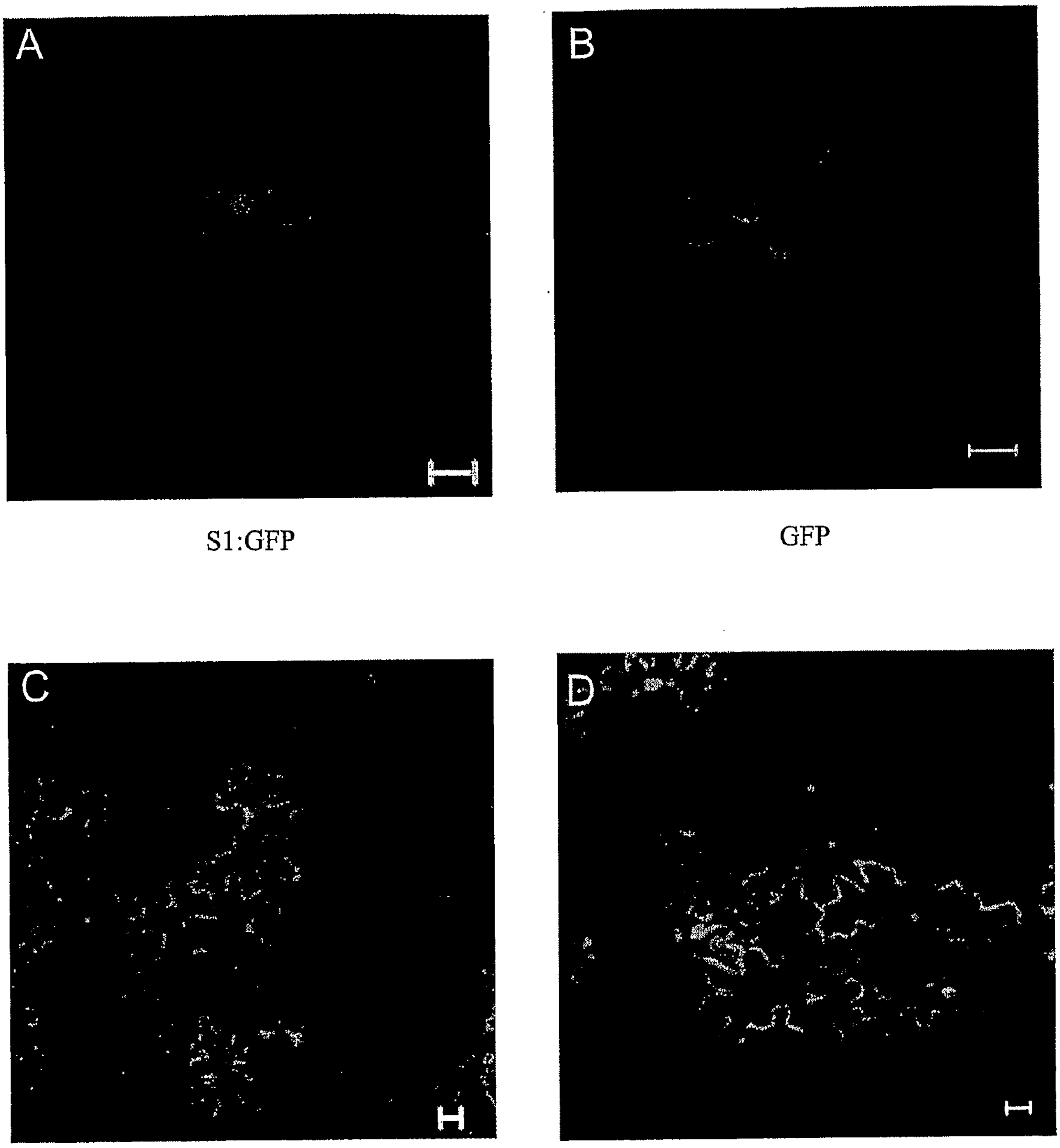
```
28980
      -D  I  A  S  D  C  T  S  I  E  L  T  S  E  F  T  E  C  T  T
      - T  *  R  V  T  A  P  P  L  S  S  R  V  S  S  R  S  A  P  L
      -  H  S  E  *  L  H  L  H  *  A  H  E  *  V  H  G  V  H  H  C
28981 - GCCATGCTTAGTGTTCCAGTTTTGTTCATAATCTTCAATGGGATCAGTGCCAAGCTCGTC -
29040
      -A  M  L  S  V  P  V  L  F  I  I  F  N  G  I  S  A  K  L  V
      - P  C  L  V  F  Q  F  C  S  *  S  S  M  G  S  V  P  S  S  S
      -  H  A  *  C  S  S  F  V  H  N  L  Q  W  D  Q  C  Q  A  R  H
```

Figure 17-NN

```
29041 - ACCTAAGTCATAAGACTTTAGATCGATGCCATAGCTATGACCACCGGCTCCCTTATTACC -
29100
      - T  *  V  I  R  L  *  I  D  A  I  A  M  T  T  G  S  L  I  T
      -  P  K  S  *  D  F  R  S  M  P  *  L  *  P  P  A  P  L  L  P
      -   L  S  H  K  T  L  D  R  C  H  S  Y  D  H  R  L  P  Y  Y  R
29101 - GTTCTTACGAAGAAGAACATTGCGGTATGCAATTGGGGTTTCGCCCACATGTGGCACGAG -
29160
      - V  L  T  K  K  N  I  A  V  C  N  W  G  F  A  H  M  W  H  E
      -  F  L  R  R  R  T  L  R  Y  A  I  G  V  S  P  T  C  G  T  S
      -   S  Y  E  E  E  H  C  G  M  Q  L  G  F  R  P  H  V  A  R  V
29161 - TACTCCCAGTGTTATACCGCTACGACCGTACTGAATGCCGTCCATTTCTGCAACCAGCTC -
29220
      - Y  S  Q  C  Y  T  A  T  T  V  L  N  A  V  H  F  C  N  Q  L
      -  T  P  S  V  I  P  L  R  P  Y  *  M  P  S  I  S  A  T  S  S
      -   L  P  V  L  Y  R  Y  D  R  T  E  C  R  P  F  L  Q  P  A  Q
29221 - AACGACCTTGTGGCCGTGATTGGTGCTTAAGGCATCAGAACGTTTAATGAACACATAGGG -
29280
      - N  D  L  V  A  V  I  G  A  *  G  I  R  T  F  N  E  H  I  G
      -  T  T  L  W  P  *  L  V  L  K  A  S  E  R  L  M  N  T  *  G
      -   R  P  C  G  R  D  W  C  L  R  H  Q  N  V  *  *  T  H  R  A
29281 - CTGTTCAAGCTGGGGCAGTACGCCTTTTTCCAGCTCTACTAGACCACAAGTGCCATTTTT -
29340
      - L  F  K  L  G  Q  Y  A  F  F  Q  L  Y  *  T  T  S  A  I  F
      -  C  S  S  W  G  S  T  P  F  S  S  S  T  R  P  Q  V  P  F  L
      -   V  Q  A  G  A  V  R  L  F  P  A  L  L  D  H  K  C  H  F  *
29341 - GAGGTGTTCACGTGCCTCCGATAGGGCCTCTTCCACAGAGTCCCCGAAGCCACGCACTAG -
29400
      - E  V  F  T  C  L  R  *  G  L  F  H  R  V  P  E  A  T  H  *
      -  R  C  S  R  A  S  D  R  A  S  S  T  E  S  P  K  P  R  T  S
      -   G  V  H  V  P  P  I  G  P  L  P  Q  S  P  R  S  H  A  L  A
29401 - CACGTCTCTAACCTGAAGGACAGGCAAACTGAGTTGGACGTGTGTTTTCTCGTTGACACC -
29460
      - H  V  S  N  L  K  D  R  Q  T  E  L  D  V  C  F  L  V  D  T
      -  T  S  L  T  *  R  T  G  K  L  S  W  T  C  V  F  S  L  T  P
      -   R  L  *  P  E  G  Q  A  N  *  V  G  R  V  F  S  R  *  H  Q
29461 - AAGAACAAGGCTCTCCATCTTACCTTTCGGTCACACCCGGACGAAACCTAGGTATGCTGA -
29520
      - K  N  K  A  L  H  L  T  F  R  S  H  P  D  E  T  *  V  C  *
      -  R  T  R  L  S  I  L  P  F  G  H  T  R  T  K  P  R  Y  A  D
      -   E  Q  G  S  P  S  Y  L  S  V  T  P  G  R  N  L  G  M  L  M
29521 - TGATCGACTGCAACACGGACGAAACCGTAAGCAGTCTGCAGAAGAGGGACGAGTTACTCG -
29580
      - *  S  T  A  T  R  T  K  P  *  A  V  C  R  R  G  T  S  Y  S
      -  D  R  L  Q  H  G  R  N  R  K  Q  S  A  E  E  G  R  V  T  R
      -   I  D  C  N  T  D  E  T  V  S  S  L  Q  K  R  D  E  L  L  V
29581 - TTTCTTGTCAACGACAGTAAAATTTATTATTGTTTATACTGCGTAGGTGCACTAGGCATG -
29640
      - F  L  V  N  D  S  K  I  Y  Y  C  L  Y  C  V  G  A  L  G  M
      -  F  L  S  T  T  V  K  F  I  I  V  Y  T  A  *  V  H  *  A  C
      -   S  C  Q  R  Q  *  N  L  L  L  F  I  L  R  R  C  T  R  H  A
29641 - CAGCCGAGCGACAGCTACACAGATTTTAAAGTTCGTTTAGAGAACAGATCTACAAGAGAT -
29700
      - Q  P  S  D  S  Y  T  D  F  K  V  R  L  E  N  R  S  T  R  D
      -  S  R  A  T  A  T  Q  I  L  K  F  V  *  R  T  D  L  Q  E  I
      -   A  E  R  Q  L  H  R  F  *  S  S  F  R  E  Q  I  Y  K  R  S
```

```
29701 - CGAGGTTGGTTGGCTTTTCCTGGGTAGGTAAAAACCTAATAT - 29742
      - R  G  W  L  A  F  P  G  *  V  K  T  *  Y  X
      -  E  V  G  W  L  F  L  G  R  *  K  P  N  X
      -   R  L  V  G  F  S  W  V  G  K  N  L  I  X
```

Figure 18



S1:GFP GFP

S1:GFP GFP

Figure 19-A

```
                                                                                              TaiI
                                                                                          BmgBI |
                                                              MboII                   HpyCH4IV| |
                 HaeII                                    RsaI  BstUI |                  NlaIII  || |
                 HhaI|                                  Csp6I|  AciI | |       EarI      FalI  | || |
               HinP1I ||                                TatI||  AluI | | |      SapI     FatI| | || |
          Hpy99I   | ||                              HphI ||| CviJI | | |   Hpy188III|    HhaI| | || |
           TaqII   | ||                       Hpy99I  |  ||| HphI |  | | |     BccI ||  HinP1I || | || |
              |    | ||                            |  |  |||  |    |  | | |      |  ||       | || | || |
           tgagcgtcgcaaaggcgctcggtcttgccttgctcgtcggtgatgtacttcaccagctccgcgaagtcgctcttcttgatggagcgcatggggacgtg
ct
         1 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 100
           actcgcagcgtttccgcgagccagaacggaacgagcagccactacatgaagtggtcgaggcgcttcagcgagaagaactacctcgcgtacccctgcac
ga

a          * A S Q R R S V L P C S S V M Y F T S S A K S L F L M E R M G T C
L -
b           E R R K G A R S C L A R R * C T S P A P R S R S S * W S A W G R A
-
c           S V A K A L G L A L L V G D V L H Q L R E V A L L D G A H G D V
L -
```

```
                                    CviJI
                                     TauI
                                 Fnu4HI |
                                   AciI| |
                             HpyF10VI || |
                          BsiEI    | || |
                         CviJI|    | || |
                        HaeIII|    | || |
                          EaeI ||  | || |
                          EagI ||  | || |
                        HpaII| ||  | || |                                                  BccI
                         NciI| ||  | || |                                               NlaIII|
                        ScrFI| ||  | || |                                               EciI ||
                      BsaJI || ||  | || |                                          Hpy188III ||
                     StyD4I | || ||  | || |                                 HpyF10VI      |  ||
                   HhaI  | || ||  | || |                                    MwoI|     |  ||
               BstUI |   | || ||  | || |                                    MnlI ||    |  ||
              HinP1I |   | || ||  | || |                                 AvaII | ||    |  ||
                   |  |  | || ||  | || |          CviJI              Sau96I | ||    |  ||
HpyF10VI
           BceAI | |  | || ||  | || |         NlaIII|       AvaI AciI | | ||    BspHI| ||       AluI
MboII|
          BsmFI | | | | || || MwoI| || |      FatI  ||    BsaJI|BslI| | | ||      FatI| ||      CviJI
MwoI|
               | | | | || ||  || || |           |   ||       ||   || | | ||        || ||          |
||
           tggcaatcacgcgcacccccccggccgttttagcggctaaaaaagtcatggctctgccctcgggcggaccacgcccatcatgaccttgccaagctcgtc
ct
       101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 200
           accgttagtgcgcgtgggggggccggcaaaatcgccgatttttttcagtaccgagacgggagcccgcctggtgcgggtagtactggaacggttcgagcag
ga

a           A I T R T P R P F * R L K K S W L C P R A D H A H H D L A K L V
L -
b          W Q S R A P P G R F S G * K S H G S A L G R T T P I M T L P S S S
C -
c           G N H A H P P A V L A A K K V M A L P S G G P R P S * P C Q A R P
-
```

```
                                                                                         HpyF10V
I
                                                                                            BglI
|
                                                             AciI
BstNI|
                                                             BslI
MwoI|
                                                            BstUI
ScrFI|
                                                             HhaI                         BsaJI
||
                                                          HinP1I |                        PspGI
||
                                                        HpyF10VI |                        StyD4I
||
```

```
                     BstKTI                                         FauI    | |                                 TauI|
||                    DpnI|    HpyF10VI                             HhaI    | |                                AciI ||
||                   EarI||       BglI|    BstKTI                 BstUI |   | |                               Fnu4HI ||
||                   MboI|||      MwoI|     HphI                 HinP1I |   | |                               CviJI| ||
||                  TaqI||||     MnlI ||    DpnI| AlwI            AciI | |  | |           CviJI              HaeIII| ||
||             MboII   |||||  Cac8I | ||  MboI ||BceAI       SfaNI | | |MwoI| |    Hpy99I   |    HphI   Cac8I || ||
||                 |   |||||      | | ||     | ||    |          | | | |   || |          |   |        |      | || ||
||            gcttctcttcgatcttcgccagcagggcgaggatcgtggcatcaccgaaccgcgccgtgcgcgggtcgtcggtgagccagagtttcagcaggccgccc
ag
         201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 300
              cgaagagaagctagaagcggtcgtcccgctcctagcaccgtagtggcttggcgcggcacgcgcccagcagccactcggtctcaaagtcgtccggcggg
tc

a               L  L  F  D  L  R  Q  Q  G  E  D  R  G  I  T  E  P  R  R  A  R  V  V  G  E  P  E  F  Q  Q  A  A  Q
-
b                F  S  S  I  F  A  S  R  A  R  I  V  A  S  P  N  R  A  V  R  G  S  S  V  S  Q  S  F  S  R  P  P
R  -
c               A  S  L  R  S  S  P  A  G  R  G  S  W  H  H  R  T  A  P  C  A  G  R  R  *  A  R  V  S  A  G  R  P
G -
```

Figure 19-B

```
                                                BsiHKAI
                                               Bsp1286I
                                                 TaiI  |
       HpyF10VI                               BmgBI |  |
           MwoI|                           HpyCH4IV| |  |
          SfaNI||                              AciI || |  |
         BstNI |||                            BstUI || |  |
         ScrFI |||                             BslI| || |  |
        BsaJI | |||                          Cac8I|| || |  |
        PspGI | |||                          AluI ||| || |  |
       StyD4I | |||                        CviJI ||| || |  |
        CviJI| | |||                     Cac8I | ||| || |  |
      HaeIII| | |||                     CviJI | | ||| || |  |
         TauI| | |||                   HaeIII | | | ||| || |  |
      Sau96I|| | |||                   Cac8I | | | | ||| || |  |
      Fnu4HI||| | |||                 Sau96I | | | | ||| || |  |
    HpyF10VI|||| | |||                  AciI | | | | | ||| || |  |
        AciI||||| | ||| HpyF10VI | | | | | | | ||| || |  |          Hpy99I
BceAI
        BglI|||| | |||    MwoI| | | | | | | ||| || |  |          BsaHI |
CviJI|
        MwoI|||| | |||FauI   || | | | | | | ||| || |  |  Hpy8I     | |  HgaI
HaeIII|
            ||||| | |||   |  || | | | | | | ||| || |  |        |    | |     |
||
          gcggcccaggtcgccattgatgcgggccagctcgcggacgtgctcatagtccacgacgcccgtgatttgtagccctggccgacggccagcaggtagg
cc
      301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 400
          cgccgggtccagcggtaactacgcccggtcgagcgcctgcacgagtatcaggtgctgcgggcactaaaacatcgggaccggctgccggtcgtccatcc
gg

a         A  A  Q  V  A  I  D  A  G  Q  L  A  D  V  L  I  V  H  D  A  R  D  F  V  A  L  A  D  G  Q  Q  V  G
R -
b          R  P  R  S  P  L  M  R  A  S  S  R  T  C  S  *  S  T  T  P  V  I  L  *  P  W  P  T  A  S  R  *  A
-
c           G  P  G  R  H  *  C  G  P  A  R  G  R  A  H  S  P  R  R  P  *  F  C  S  P  G  R  R  P  A  G  R
P  -
```

(Right-hand enzyme stack, positions 301–400:)

```
                Cac8I
             CviJI |
            HaeIII |
           Hpy99I| |
            EaeI|| |
      BspMI   ||| |
     CviJI |   ||| |
    HaeIII |   ||| |
      EaeI | |  ||| |
    BstNI| | |  ||| |
    ScrFI| | |  ||| |
    BsaJI || | |  ||| |
    PspGI || | |  ||| |
   StyD4I || | |  ||| |
   CviJI| || | |  ||| |
          || || | |  ||| |
```

```
                           TauI
                         AciI |
                       Fnu4HI |
                     HpyF10VI |
                         MwoI| |
                         TauI| |
                       AciI || |
                     Fnu4HI || |
                       TauI| || |
                     AciI || || |
                    BsiEI || || |
                   Fnu4HI || || |
                    CviJI| || || |
                   HaeIII| || || |
                   Cac8I || || || |
                    EaeI || || || |
                    EagI || || || |
                    NaeI || || || |
FokI
               HpaII| || || || |                                                              CviJI
!
              BsrFI|| || || || |                                  BbvI                    TspDTI |
|
             NgoMIV|| || || || |                              Hpy8I |                  BstNI | |
|
             NlaIII|| || || || |               Hpy188III       RsaI| |         Fnu4HI    ScrFI | |
|
             FatI  ||| || || || |               EarI    |    Csp6I|| |         CviJI| PspGI | | |
|
          CviJI |  ||| || || || |  MboII   MnlI SapI     |    TatI||| |          TseI| StyD4I | | |
|
               | |  ||| || || || |      |     |   |     |     |||| |             ||     || | |
|
          gacaggctcatgccggccgccgccgccttttcctcaatcgctcttcgttcgtctggaaggcagtacaccttgataggtgggctgcccttcctggttgg
ct
      401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 500
          ctgtccgagtacggccggcggcggcggaaaaggagttagcgagaagcaagcagaccttccgtcatgtggaactatccacccgacgggaaggaccaacc
ga

a          Q  A  H  A  G  R  R  R  L  F  L  N  R  S  S  F  V  W  K  A  V  H  L  D  R  W  A  A  L  P  G  W
L  -
```

b D R L M P A A A A F S S I A L R S S G R Q Y T L I G G L P F L V G
L -
c T G S C R P P P P P F P Q S L F V R L E G S T P * * V G C P S W L A
-

Figure 19-C

```
                                                        HpyF10VI
                                                            MwoI|
                                                         CviJI ||
                                                       Cac8I | ||
                                         AciI      CviJI | | ||
                                       Cac8I |     HaeIII | | ||
                                        NaeI |     Cac8I | | | ||
                                      HpaII| |      EaeI | | | ||
                                     BsrFI|| |      NaeI | | | ||
                                    NgoMIV||| |    HpaII|| | | | ||                              BstNI
                 BccI               SgrAI||| |    BsrFI|| | | | ||           HinfI           ScrFI
      BstF5I   Cac8I           MnlI  |||| |    CviJI|| | | | ||            TfiI           PspGI |
      CviJI|AciI    |     MaeIII  |  |||| |  |NgoMIV|| | | | ||        MnlI   |      BsiHKAI| | |
           ||  |    |          |  |  |||| |  |     ||| | | | ||           |   |      Bsp1286I|| | |
                                                                                                StyD4I |
                                                                                                 AciI | |
        tggtttcatcagccatccgcttgccctcatctgttacgccggcggtagccggccagcctcgcagagcaggattcccgttgagcaccgccaggtgcgaa
ta
   501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 600
        accaaagtagtcggtaggcgaacgggagtagacaatgcggccgccatcggccggtcggagcgtctcgtcctaagggcaactcgtggcggtccacgctt
at

a        G  F  I  S  H  P  L  A  L  I  C  Y  A  G  G  S  R  P  A  S  Q  S  R  I  P  V  E  H  R  Q  V  R  I
-
b         V  S  S  A  I  R  L  P  S  S  V  T  P  A  V  A  G  Q  P  R  R  A  G  F  P  L  S  T  A  R  C  E
*  -
c        W  F  H  Q  P  S  A  C  P  H  L  L  R  R  R  *  P  A  S  L  A  E  Q  D  S  R  *  A  P  P  G  A  N
K -


                                 CviJI
                                HaeIII
                                  HphI|
                                Sau96I||
                           FauI    |||
                           AciI|   |||
                           BslI|   |||                              BciVI
                          BstUI|   |||                             BsaHI |
                           BslI||  |||                        CviJI    | |
                         Cac8I|||  |||                    HpaII  |     | |
                 BsmFI   BsrBI ||||  |||                     NciI  |     | |
          TspRI     |    AciI | ||||  |||                   ScrFI  |     | |
         PpiI |     |    FauI| | ||||  |||                StyD4I |  |     | |
    HpyCH4III |     |   MboII| | ||||  |||PpiI      BceAI     |  |  |     | |       BsaJI       Hpy8I
            |  |    |      || | ||||  |||  |        MmeI|     |  |  |     | |   HgaI  StyI       AccI|
        agggacagtgaagaaggaacacccgctcgcgggtgggcctacttcacctatcctgcccggctgacgccgttggatacaccaaggaaagtctacacgaa
cc
   601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 700
        tccctgtcacttcttccttgtgggcgagcgcccacccggatgaagtggataggacgggccgactgcggcaacctatgtggttcctttcagatgtgctt
gg

a       R  D  S  E  E  G  T  P  A  R  G  W  A  Y  F  T  Y  P  A  R  L  T  P  L  D  T  P  R  K  V  Y  T  N
P -
b        G  T  V  K  K  E  H  P  L  A  G  G  P  T  S  P  I  L  P  G  *  R  R  W  I  H  Q  G  K  S  T  R  T
-
c         G  Q  *  R  R  N  T  R  S  R  V  G  L  L  H  L  S  C  P  A  D  A  V  G  Y  T  K  E  S  L  H  E
P  -


                                                                                         AciI
                                                                                      MspA1I      HaeII
                                                                                    Fnu4HI |       HhaI|
                                                                                BslI  HpyCH4V| |      BbvI||
                            BccI      BstF5I    FokI                            BslI|    TseI| |   HinP1I|||
                              |          |        |                               ||        || |        ||||
        ctttggcaaaatcctgtatatcgtgcgaaaaaggatggatataccgaaaaaatcgctataatgaccccgaagcagggttatgcagcggaaaagcgcca
cg
   701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 800
        gaaaccgttttaggacatatagcacgcttttttcctacctatatggctttttttagcgatattactggggcttcgtcccaatacgtcgccttttcgcggt
gc

a        L  A  K  S  C  I  S  C  E  K  G  W  I  Y  R  K  N  R  Y  N  D  P  E  A  G  L  C  S  G  K  A  P
R  -
b       L  W  Q  N  P  V  Y  R  A  K  K  D  G  Y  T  E  K  I  A  I  M  T  P  K  Q  G  Y  A  A  E  K  R  H
A -
c        F  G  K  I  L  Y  I  V  R  K  R  M  D  I  P  K  K  S  L  *  *  P  R  S  R  V  M  Q  R  K  S  A  T
-
```

Figure 19-D

```
                                          TauI
                                      Fnu4HI |
                                        AciI||
                                       BciVI||
                                       BslI|||
                                    EciI |||
                               HpaII    |  |||                        BssSI             BstNI
                               BsaWI|   |  |||                        HhaI|             ScrFI
                          MmeI    ||    |  |||                      HinP1I ||          BsaJI|
  Hpy188III       AciI     |     ||    |  |||  Hpy188I              MnlI| ||          PspGI||        BstNI
         |          |       |     ||    |  ||| |     |                 || ||          StyD4I||       ScrFI
                                                                                     AluI |||       PspGI |
                                                                                     CviJI |||      StyD4I |
     cttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctt
 801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 900
     gaagggcttccctctttccgcctgtccataggccattcgccgtcccagccttgtcctctcgcgtgctccctcgaaggtccccctttgcggaccatagaaa

a     F  P  K  G  E  R  R  T  G  I  R  *  A  A  G  S  E  Q  E  S  A  R  G  S  F  Q  G  E  T  P  G  I  F  -
b      S  R  R  E  K  G  G  Q  V  S  G  K  R  Q  G  R  N  R  R  A  H  E  G  A  S  R  G  K  R  L  V  S  L  -
c     L  P  E  G  R  K  A  D  R  Y  P  V  S  G  R  V  G  T  G  E  R  T  R  E  L  P  G  G  N  A  W  Y  L  Y  -

                                                                                              CviJI
                                                                                             HaeIII
                                                                                               TauI
                                                                                           Fnu4HI |
                             DrdI                                                            AciI| |
                            MnlI|                                                           BstUI| |
                         SmlI ||Hpy99I                        CviJI                   HpyF10VI |  | |
                 Hpy188I   |  ||SfaNI|                        NlaIV|                    MwoI|  | |
                 HgaI   |   | ||TaqI||       BpuEI          AciI ||          EciI Cac8I   ||  | |
                    |   |   | ||  |||          |              |  ||           |     |     ||  | |
     atagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctt
 901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1000
     tatcaggacagcccaaagcggtggagactgaactcgcagctaaaaacactacgagcagtccccccgcctcggataccttttttgcggtcgttgcgccggaa

a     I  V  L  S  G  F  A  T  S  D  L  S  V  D  F  C  D  A  R  Q  G  G  G  A  Y  G  K  T  P  A  T  R  P  F  -
b     *  S  C  R  V  S  P  P  L  T  *  A  S  I  F  V  M  L  V  R  G  A  E  P  M  E  K  R  Q  Q  R  G  L  -
c      S  P  V  G  F  R  H  L  *  L  E  R  R  F  L  *  C  S  S  G  G  R  S  L  W  K  N  A  S  N  A  A  F  -

             CviJI
            HaeIII
          BstNI  |
          ScrFI  |
        PspGI |  |
       StyD4I |  |                NlaIII
       NlaIV| |  |      CviJI        NspI
 HpyCH4III || |  |     HaeIII   AflIII |
     BslI | || |  |     Cac8I |   FatI  |
          | || |  |      BslI | |   PciI  |            HinfI                                     AluI
          | || |  |       | | |     |   |             TfiI   HpyCH4III   AciI                CviJI
     |    |  ||  |  |        ||  |       |   |              |          |        |                  |
     tttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagct
1001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1100
     aaatgccaaggaccggaaaacgaccggaaaacgagtgtacaagaaaggacgcaataggggactaagacacctattggcataatggcggaaactcactcga

a      Y  G  S  W  P  F  A  G  L  L  L  T  C  S  F  L  R  Y  P  L  I  L  W  I  T  V  L  P  P  L  S  E  L  -
b     F  T  V  P  G  L  L  L  A  F  C  S  H  V  L  S  C  V  I  P  *  F  C  G  *  P  Y  Y  R  L  *  V  S  *  -
c     L  R  F  L  A  F  C  W  P  F  A  H  M  F  F  P  A  L  S  P  D  S  V  D  N  R  I  T  A  F  E  *  A  -
```

Figure 19-E

```
                                                                                                BsiEI
                                                                                               CviJI|
                                                                                              HaeIII|
                                                                                                TauI|
                                                                                              EaeI ||
                                                                                              EagI ||
                                                                                            Fnu4HI ||
                                                                                              MnlI ||
                                                                                           AciI| ||
                                                                                          BstUI| ||
                                                                                           HhaI| ||
                                                                                      HinP1I || ||
                                                                                   CviJI    | || ||
                                                                                  HaeIII    | || ||
                                                                               HpyF10VI |   | || ||
                                                                                   BglI||   | || ||
                          TspRI                                                    MwoI||   | || ||
               CviJI       MlyI|                                                BceAI |||   | || ||
             Fnu4HI |      PleI||                                            TauI |   |||   | || ||
               TauI||      BbvI|||                                           MnlI||   |||   | || ||
               TseI||     MnlI ||||                                          AciI|||  |||   | || ||
            AciI |||     TaqII | ||||                                      Fnu4HI|||  |||   | || ||
          Fnu4HI  |||    HinfI | | ||||                                    CviJI||||  |||   | || ||
        Cac8I  | |||  Fnu4HI | | | ||||                                   HaeIII||||  |||   | || ||
      BsrBI | | |||     HhaI| | | | ||||                                 MboII |||||  |||   | || ||
     AciI | | | |||     TseI|| | | | ||||      AciI       HaeII       | |||||  |||   | || ||
 HpyF10VI | | | | |||   HinP1I || | | | ||||    EarI |      HhaI|       | |||||  |||   | || ||
   MwoI|| | | | | |||   BbvI | || | | | ||||    SapI |  HinP1I ||       | |||||  |||   | || ||
       gataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgccagaaggccgccagagaggccgagcgcggccg
tg
  1101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 1200
       ctatggcgagcggcgtcggcttgctggctcgcgtcgctcagtcactcgctccttcgccttctcgcggtcttccggcggtctctccggctcgcgccggc
ac

a        I  P  L  A  A  A  E  R  P  S  A  A  S  Q  *  A  R  K  R  K  S  A  R  R  P  P  E  R  P  S  A  A  V
-
b         Y  R  S  P  Q  P  N  D  R  A  Q  R  V  S  E  R  G  S  G  R  A  P  E  G  R  Q  R  G  R  A  R  P
*  -
c       D  T  A  R  R  S  R  T  T  E  R  S  E  S  V  S  E  E  A  E  E  R  Q  K  A  A  R  E  A  E  R  G  R
E -

                                         Fnu4HI                                                  BsaXI
                                         CviJI|                                                 CviJI|
                                          TseI|             AciI                               NlaIII||
                                          HgaI||            BstUI                              FokI |||
                                         Cac8I|||       HpyF10VI |                            FatI | |||
                               CviJI    AciI ||||           MwoI| |       MnlI            Hpy8I | | |||
                        NlaIII  FauI    BslI ||||        Hpy188I|| |      HgaI|            AccI| | | |||
   CviJI      BfaI HgaI FatI   |BbvI |  TspDTI ||||      BsaHI   ||| |BsaXI||            BstF5I || | | |||
       aggcttggacgctagggcagggcatgaaaaagcccgtagcgggctgctacgggcgtctgacgcggtggaaagggggaggggatgttgtctacatggct
ct
  1201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 1300
       tccgaacctgcgatcccgtcccgtactttttcgggcatcgcccgacgatgcccgcagactgcgccacctttcccccctcccctacaacagatgtaccga
ga

a       R  L  G  R  *  G  R  A  *  K  S  P  *  R  A  A  T  G  V  *  R  G  G  K  G  E  G  M  L  S  T  W  L
C -
b        G  L  D  A  R  A  G  H  E  K  A  R  S  G  L  L  R  A  S  D  A  V  E  R  G  R  G  C  C  L  H  G  S
-
c         A  W  T  L  G  Q  G  M  K  K  P  V  A  G  C  Y  G  R  L  T  R  W  K  G  G  G  D  V  V  Y  M  A
L  -
```

Figure 19-F

```
                                                TaqII
                                              MaeIII |
                                              Tsp45I |
                                          BstKTI   | |
                                            BbvI|  | |
                                            DpnI|  | |
                                            HphI|  | |
                                          BclI ||  | |
                                          MboI ||  | |
                                Hpy188III |  | ||  | |
                                   AvaII  |  | ||  | |
                                  Sau96I  |  | ||  | |
                                    AciI |  |  | ||  | |
                                 MspA1I |  |  | ||  | |
                                  Hin4I| |  |  | ||  | |
                                Fnu4HI|| |  |  | ||  | |                    Hpy188III
                              HpyF10VI|| |  |  | ||  | |                      TaiI |
                                  MwoI|| |  |  | ||  | |                  AclI  | |
                                  TseI|| |  |  | ||  | |             HpyCH4IV | | |
                             HpaII  ||| |  |  | ||  | |          Hin4I      |  | |
                         HhaI  |  |||| |  |  | ||  | |          AvaII |     |  | |
     SfcI           HinP1I |  |  |||| |  |  | ||  | |        Sau96I | |     |  | |      CviJI        MnlI
      |                  |  |  |  |||| |  |  | ||  | |        BslI |  | |     |  | |        |            |
      gctgtagtgagtgggttgcgctccggcagcggtcctgatcaatcgtcaccctttctcggtccttcaacgttcctgacaacgagcctcctttcgccaatc
 1301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1400
      cgacatcactcacccaacgcgaggccgtcgccaggactagttagcagtgggaaagagccaggaagttgcaaggactgttgctcggaggaaaagcggttag

a       C  S  E  W  V  A  L  R  Q  R  S  *  S  I  V  T  L  S  R  S  F  N  V  P  D  N  E  P  P  F  R  Q    S -
b      A  V  V  S  G  L  R  S  G  S  G  P  D  Q  S  S  P  F  L  G  P  S  T  F  L  T  T  S  L  L  F  A  N   P -
c       L  *  *  V  G  C  A  P  A  A  V  L  I  N  R  H  P  F  S  V  L  Q  R  S  *  Q  R  A  S  F  S  P  I  -


                                                                                   AciI
                                                                                  Cac8I
                                                     HgaI                        CviJI |                     CviJI
                                                   CviJI|                      HaeIII |                     NlaIV|
                                                 HpaII ||                        TauI |                  AciI ||
                                                BsrFI| ||                      Sau96I| |                 BsrBI ||
                                               AvaII || ||                    Fnu4HI|| |             HpyF10VI |  ||
                                               RsrII || ||                      AciI||| |                MwoI| |  ||
                                              Sau96I || ||                     BstUI||| |               TauI || |  ||
                          AloI            HpaII | || ||              HpyF10VI |||| |          Fnu4HI |  || |  ||
                         BbvI|  TaqI  Hpy188III | | || ||         BsaHI     | |||| |        HpyF10VI |  || |  ||
  BsmFI          HinfI   || MlyI|      BsaWI| | | || ||  Hpy99I  |          | |||| |            AciI| |  || |  ||
  TaqI|    BstUI |  ||HgaI|  Fnu4HI  || | | || ||  TaqI |  |          | |||| |          MspA1I| |  || |  ||
  HphI||BccI AciI | |  ||PleI|  TseI|BspEI| | | || ||AloI| |  |      MwoI | |||| |            MwoI| |  || |  ||
   |||   |    | | |  ||  |||    ||   || | | || ||  || |  |         || |||| |           FauI||| |  || |  ||
      catcgacaatcaccgcgagtccctgctcgaacgctgcgtccggaccggcttcgtcgaaggcgtctatcgcggcccgcaacagcggcgagagcggagcctg
 1401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1500
      gtagctgttagtggcgctcagggacgagcttgcgacgcaggcctggccgaagcagcttccgcagatagcgccgggcgttgtcgccgctctcgcctcggac

a      I  D  N  H  R  E  S  L  L  E  R  C  V  R  T  G  F  V  E  G  V  Y  R  G  P  Q  Q  R  R  E  R  S  L  -
b       S  T  I  T  A  S  P  C  S  N  A  A  S  G  P  A  S  S  K  A  S  I  A  A  R  N  S  G  E  S  G  A   C -
c      H  R  Q  S  P  R  V  P  A  R  T  L  R  P  D  R  L  R  R  R  R  L  S  R  P  A  T  A  A  R  A  E  P   V -
```

Figure 19-G

```
                                  HpyF10VI
                                Cac8I     |
                                 NaeI     |
                               HpaII|     |
                              BsrFI||     |
                             NgoMIV||     |
                            Cac8I |||     |
                         HpyF10VI |||     |
                            MwoI| |||     |
                           HhaI|| |||     |
                        BstUI ||| |||     |
                       HinP1I ||| |||     |        Cac8I
                         TauI ||| |||     |       CviJI |
                        BcgI| ||| |||     |      HaeIII |
                       AciI|| ||| |||     |     Cac8I | |
                     Fnu4HI|| ||| |||     |      NaeI | |
Tth111I
                       TauI||| ||| |||    |     HpaII| | |               HpyCH4III
HpyF10VI|
                       AciI |||| ||| |||  |    BseRI|| | |              NlaIV      |                        HhaI
||
                     Fnu4HI |||| ||| |||  |    BsrFI|| | | HpyF10VI  CviJI|      |       AluI         HgaI |
||
                     NlaIV | |||| ||| |||  |  NgoMIV|| | |     MwoI| HaeIII|      |      CviJI        HinP1I |
||
                   BanI | | |||| ||| |||  | SfaNI ||| | |  BcgI ||  MnlI||       |BceAI    |          BsmFI | |
||
             HpyCH4III | | |||| ||| |||MwoI|BpuEI| ||| | |  SmlI ||Sau96I||       |TspRI    |  BsaBI     | | |
MwoI||
                  | | | |||| ||| |||   ||    || ||| | |     | ||     |||        |     |     |      |     | | |
|||
          ttcaacggtgccgccgcgctcgccggcatcgctgtcgccggcctgctcctcaagcacggccccaacagtgaagtagctgattgtcatcagcgcattga
cg
   1501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 1600
          aagttgccacggcggcgcgagcggccgtagcgacagcggccggacgaggagttcgtgccggggttgtcacttcatcgactaacagtagtcgcgtaact
gc

a         F  N  G  A  A  A  L  A  G  I  A  V  A  G  L  L  L  K  H  G  P  N  S  E  V  A  D  C  H  Q  R  I  D
G -
b          S  T  V  P  P  R  S  P  A  S  L  S  P  A  C  S  S  S  T  A  P  T  V  K  *  L  I  V  I  S  A  L  T
-
c           Q  R  C  R  R  A  R  R  H  R  C  R  R  P  A  P  Q  A  R  P  Q  Q  *  S  S  *  L  S  S  A  H  *
R  -

                                                                                                      Bst
UI
                                                                                                      Cac
8I
                                                                                                       Hh
aI
                                                                                                     HinP
1I
                                                                                                   BssHII
|
                                                                                                   BstF5I
|
                                                                                                    BstUI
|
                                                                                                    Cac8I
|
                                                                                                     HhaI
|
                                                                                                   HinP1I
|
                                                                                                BssHII |
|
                                                                                                HinP1I |
|
                                                                                            NlaIII   | |
|
                                                         BsiEI                               FatI |  | |
|
                                                        CviJI|                          Cac8I |   |  | |
|
                                                       HaeIII|                           NaeI |   |  | |
|
                                                      Hpy99I||                          HpaII| |  |  | |
|
                BceAI                           BstUI    |||                           BsrFI|| |  |  | |
|
              BsiEI |                            HhaI    |||                          NgoMIV|| |  |  | |
|
```

```
                CviJI|  |                    HinP1I |     |||                 SfaNI|||  |    |    | |
               HaeIII|  |                    Fnu4HI | |   |||                 Cac8I||||  |    |    | |
                 EaeI ||  |                    AluI| | |   |||               BstUI |||||  |    |    | |
                 EagI ||  |                   CviJI| | |   |||              BsiEI | |||||  |    |    | |
             HpyF10VI ||  |                BceAI   || | |   |||          HpaII | | |||||  |    |    | |
                HpaII| ||  |                HgaI |  || | |   |||            NciI | | |||||  |    |    | |
                 MwoI| ||  |           HpyF10VI| |  || | |   |||           ScrFI | | |||||  |    |    | |
                 NciI| ||  |              BslI   || |  || | |   |||          HhaI |  | | |||||  |    |    | |
                ScrFI| ||  |              BbvI|   || |  || | |   |||       StyD4I |  | | |||||  |    |    | |
              BsaJI || ||  |          EcoNI||   || |  || | |   |||       HinP1I | |  | | |||||  |    |    | |
             StyD4I || ||  |  MnlI   |||MnlI|| |  || | |EaeI|||         BccI | | |  | | |||||  |    |    | |
          BsaHI  | || ||  | AciI  |FauI||MwoI|| |TseI| | |EagI|||       AciI |  | | |  | | |||||  |    |    | |
              |  | || ||  |   |  |  |||  |||  |  || | |  ||||     | |  | | |  | | |||||  |    |    | |
       gcgtccccggccgaaaaacccgcctcgcagaggaagcgaagctgcgcgtcggccgtttccatctgcggtgcgcccggtcgcgtgccggcatggatgcg
cg
  1601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 1700
       cgcaggggccggctttttgggcggagcgtctccttcgcttcgacgcgcagccggcaaaggtagacgccacgcgggccagcgcacggccgtacctacgc
gc

a        V  P  G  R  K  T  R  L  A  E  E  A  K  L  R  V  G  R  F  H  L  R  C  A  R  S  R  A  G  M  D  A
R  -
b       A  S  P  A  E  K  P  A  S  Q  R  K  R  S  C  A  S  A  V  S  I  C  G  A  P  G  R  V  P  A  W  M  R
A -
c        R  P  R  P  K  N  P  P  R  R  G  S  E  A  A  R  R  P  F  P  S  A  V  R  P  V  A  C  R  H  G  C  A
-
```

Figure 19-H

```
                                                          Hpy188I
                                                         BstKTI |
                                                          DpnI| |
                                                        Eco57I| |
                                                       Eco57MI| |
                      Cac8I                              Hin4I| |
                   HpyF10VI                             MboI || |
                    AlwNI|                        Hpy188III | || |
                   BstAPI|                            BsmI  | | || |
                    HaeII|                       Cac8I |    | | || |
                     MwoI|
                    HhaI||
HinfI
                    BbvI |||          AciI | |   | | || |                                      BsmI
|                  HinP1I |||       Fnu4HI | | |  | | || |                                   EciI|
|             BccI  HpyF10VI |||      AluI| | | |  | | || |                             HpyF10VI ||
|          AciI |   MwoI| |||       CviJI| | | |  | | || |    BsrI                     BstAPI| ||
|        BstUI |  Fnu4HI|| |||      TseI| | | |  | | || |   HaeII        CviJI  NlaIV     || ||
|      FokI | |  TseI||| ||| BbvI   || | | |  | | || |   HhaI|       Hpy99I|  BanI |     || ||
||TfiI HhaI | | | Cac8I |||| ||| FauI  || | | |  | | || |  HinP1I ||   Hpy99I ||Hin4I| | MwoI|
|         | | | |   | |||| |||    |   || | | |  | | || |       | ||      | ||   || |    || ||
   cgccatcgcggtaggcgagcagcgcctgcctgaagctgcgggcattcccgatcagaaatgagcgccagtcgtcgtcggctctcggcaccgaatgcgta
tg
  1701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 1800
   gcggtagcgccatccgctcgtcgcggacggacttcgacgcccgtaagggctagtctttactcgcggtcagcagcagccgagagccgtggcttacgcat
ac

a      A  I  A  V  G  E  Q  R  L  P  E  A  A  G  I  P  D  Q  K  *  A  P  V  V  V  G  S  R  H  R  M  R  M
-
b       P  S  R  *  A  S  S  A  C  L  K  L  R  A  F  P  I  R  N  E  R  Q  S  S  S  A  L  G  T  E  C  V
* -
c      R  H  R  G  R  R  A  A  P  A  *  S  C  G  H  S  R  S  E  M  S  A  S  R  R  R  L  S  A  P  N  A  Y
D -

                                                                     Fnu4HI
                                                                   HpyF10VI
                                                                      CviJI|
                                                                       MwoI|
                                                                       TseI|
                        BsrI                                        Eco57I||
                      CviJI|                  AciI                 Eco57MI||
                     HaeIII|                 Cac8I                   Cac8I|||
                      EaeI ||             HpyF10VI                    NaeI|||
                    HgaI | ||                HaeII|                  HaeII||||
                  CviJI| | ||                 MwoI|                  HpaII||||
               NlaIII|| | ||                 HhaI||                 BsrFI|||||
            HpyF10VI||| | ||             HinP1I |||                  HhaI|||||
               BglI|||| | ||            Fnu4HI | |||                NgoMIV|||||
               MwoI|||| | ||              TseI| | |||             HinP1I ||||||
               FatI ||||| | ||         Hpy99I || | |||Hpy188III    HpyF10VI | ||||||
           Cac8I | ||||| | ||         TaqI | || | |||   FauI |     BbvI   | | ||||||      HpyCH4III
BsrI
|        AciI  | | ||||| | ||      TspRI| | || | ||| BbvI | |   BsrI|MwoI| | ||||||    EciI   |AciI
|            |  | | ||||| | ||         || | || | ||    | | |      ||  || | ||||||       |       | |
   attctccgccagcatggcttcggccagtgcgtcgagcagcgcccgcttgttcctgaagtgccagtaaagcgccggctgctgaaccccccaaccgttccg
cc
  1801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 1900
   taagaggcggtcgtaccgaagccggtcacgcagctcgtcgcgggcgaacaaggacttcacggtcatttcgcggccgacgacttggggggttggcaaggc
gg

a      I  L  R  Q  H  G  F  G  Q  C  V  E  Q  R  P  L  V  P  E  V  P  V  K  R  R  L  L  N  P  Q  P  F  R
Q -
b       F  S  A  S  M  A  S  A  S  A  S  S  S  A  R  L  F  L  K  C  Q  *  S  A  G  C  *  T  P  N  R  S  A
-
c      S  P  P  A  W  L  R  P  V  R  R  A  A  P  A  C  S  *  S  A  S  K  A  P  A  A  E  P  P  T  V  P
P -

                                                                   HpyCH4V
                                                                  Fnu4HI |
                                                 TauI             CviJI| |
                                              Fnu4HI |             TseI| |
```

```
                                                     BstNI   | |            TspGWI ||  |
                                                     ScrFI   | |          TspRI  | ||  |
                                                    BsaJI|   | |         AlwI |  | ||  |
                                                   PspGI||   | |   HpyCH4III| |  | ||  |
                  Hpy8I                           StyD4I||   | |  BstKTI  || |  | ||  |
                  AccI|                          AvaII |||   | |  BbvI|   || |  | ||  |          NlaIII
            HpyCH4III ||                        Sau96I |||   | |  DpnI|   || |  | ||  |       FatI  |
           Hpy188I  | ||                        MnlI | |||AciI| |MboI | || || |  | ||  |     Hin4I|  |
                 |  | ||                            | | |||   || |    | || || |  | ||  |        ||  |
        agtttgcgtgtcgtcagaccgtctacgccgacctcgttcaacaggtccagggcggcacggatcactgtattcggctgcaacttgtcatgcttgacac
tt
   1901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 2000
        tcaaacgcacagcagtctggcagatgcggctggagcaagttgtccaggtcccgccgtgcctagtgacataagccgacgttgaaacagtacgaactgtg
aa

a         F  A  C  R  Q  T  V  Y  A  D  L  V  Q  Q  V  Q  G  G  T  D  H  C  I  R  L  Q  L  C  H  A  *  H
F  -
b        S  L  R  V  V  R  P  S  T  P  T  S  F  N  R  S  R  A  A  R  I  T  V  F  G  C  N  F  V  M  L  D  T
L -
c         V  C  V  S  S  D  R  L  R  R  P  R  S  T  G  P  G  R  H  G  S  L  Y  S  A  A  T  L  S  C  L  T  L
-
```

Figure 19-I

```
                                                                              HpaII
                                                   BstUI            AciI  Hpy188III
                                                    HhaI          MspA1I     BsaWI
                                                  BstUI          MnlI        BspEI          TaiI
                                                 HinP1I        AvaII        AvaII         BmgBI
                                            HinfI             Hpy188I      Sau96I     HpyCH4IV
                                             TfiI             Sau96I        MnlI      CviJI
         TspRI   Hin4I   Hpy8I   Hin4I   TspRI    AciI        Hin4I     CviJI       HaeIII
      tatcactgataaacataatatgtccaccaacttatcagtgataaagaatccgcgcgttcaatcggaccagcggaggctggtccggaggccagacgtga
aa
  2001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 2100
      atagtgactatttgtattatacaggtggttgaatagtcactatttcttaggcgcgcaagttagcctggtcgcctccgaccaggcctccggtctgcact
tt

a      I  T  D  K  H  N  M  S  T  N  L  S  V  I  K  N  P  R  V  Q  S  D  Q  R  R  L  V  R  R  P  D  V  K
-
b       S  L  I  N  I  I  C  P  P  T  Y  Q  *  *  R  I  R  A  F  N  R  T  S  G  G  W  S  G  G  Q  T  *
N  -
c      Y  H  *  *  T  *  Y  V  H  Q  L  I  S  D  K  E  S  A  R  S  I  G  P  A  E  A  G  P  E  A  R  R  E
T -

                                           Hpy99I                                                BstKTI
                                          Tth111I                                                  DpnI
                                        HpyF10VI                                                   MnlI
                                            MwoI                                                  MboI
                                           TaqI                                                  BstUI
                                          HhaI                                                    HhaI
                                       BstUI                                                   HinP1I
                                     HinP1I                                                  HpyF10VI
                                     TspRI                                           CviJI
                              HpyCH4III                                             HaeIII
                              BsiHKAI                                             Sau96I
                             Bsp1286I                                               NciI
                        Hpy188I                                                     ScrFI
                  Tsp509I                                                          HpaII
                 BspCNI                                                          StyD4I
                 BstKTI                                  BbvI                  Fnu4HI
                BseMII                                   SfaNI              HpyF10VI
                 DpnI                                 CviJI          HpaII MwoI
                MboI        DdeI                 HgaI  HaeIII        BsrFI  TseI           MwoI
      cccaacataccccctgatcgtaattctgagcactgtcgcgctcgacgctgtcggcatcggcctgattatgccggtgctgccgggcctcctgcgcgatct
gg
  2101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 2200
      gggttgtatggggactagcattaagactcgtgacagcgcgagctgcgacagccgtagccggactaatacggccacgacggcccggaggacgcgctaga
cc

a      P  N  I  P  L  I  V  I  L  S  T  V  A  L  D  A  V  G  I  G  L  I  M  P  V  L  P  G  L  L  R  D  L
V -
b       P  T  Y  P  *  S  *  F  *  A  L  S  R  S  T  L  S  A  S  A  *  L  C  R  C  C  R  A  S  C  A  I  W
-
c       Q  H  T  P  D  R  N  S  E  H  C  R  A  R  R  C  R  H  R  P  D  Y  A  G  A  A  G  P  P  A  R  S
G  -

                                                                                                 BstKTI
                                                                                                   DpnI
                 AciI                                                                          Hpy188I
              AatII                                                                        HhaI
               TaiI                                                                      BstUI
             Hpy99I                                                                      Cac8I
            MaeIII                                                                        HhaI
            Tsp45I                                                                       HinP1I
              ZraI                                                                     HpyF10VI
             BsaHI                                  HpyF10VI                             MwoI
          HpyCH4IV                                      MwoI                MmeI       BssHII
          HphI                                      HaeII                  HhaI        HinP1I
       TaqI                                          HhaI                 FspI
   Hpy8I                           BsmI          HinP1I         Tsp509I  HinP1I     BslI
                                                 Cac8I       HpyCH4V    Cac8I     BseYI              MboI
      ttcactcgaacgacgtcaccgcccactatggcattctgctggcgctgtatgcgttggtgcaatttgcctgcgcacctgtgctgggcgcgctgtcggat
cg
  2201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 2300
      aagtgagcttgctgcagtggcgggtgataccgtaagacgaccgcgacatacgcaaccacgttaaacggacgcgtggacacgacccgcgcgacagccta
gc

a      H  S  N  D  V  T  A  H  Y  G  I  L  L  L  A  L  Y  A  L  V  Q  F  A  C  A  P  V  L  G  A  L  S  D
R  -
```

b F T R T T S P P T M A F C W R C M R W C N L P A H L C W A R C R I
V -
c S L E R R H R P L W H S A G A V C V G A I C L R T C A G R A V G S
-

Figure 19-J

```
                                                  BslI
                                               BstKTI|
                                                 DpnI||
                                              BglII |||
                                              BstYI |||
                                               MboI |||
                                             BbeI | |||
                                            HaeII | |||
                                            HhaI| | |||
                                          NlaIV|| | |||
                                           SfoI|| | |||
                                         BsaHI||| | |||
                                        HinP1I||| | |||
                                          NarI||| | |||
                                         BanI|||| | |||
                                         KasI|||| | |||
                                       Cac8I||||| | |||
                                        NaeI||||| | |||
              HpyF10VI                 HpaII|||||| | |||
          CviJI     |                  BsrFI|||||| | |||
         HaeIII     |                  CviJI|||||| | |||
           TauI     |                 HaeIII|||||| | |||
          EaeI |    |                NgoMIV||||||| | |||                HpyCH4V
        Fnu4HI |    |                Cac8I |||||||| | |||                 NlaIII
          AciI||    |                 EaeI |||||||| | |||                   NspI
          TauI||    |              BsmAI | |||||||| | |||                   SphI
      Fnu4HI ||     |              BsmBI | |||||||| | |||                Cac8I |
        AciI| ||    |          HpyF10VI | | |||||||| | |||                FatI | |
   AlwI  || ||  |    MwoI|     MwoI |  | | |||||||| | ||| NlaIV     BslI  | | |      MslI
   |     || ||  |       ||        || |  | |||||||| | |||    |          |   | | |       |
   tttcgggcggcggccaatcttgctcgtctcgctggccggcgccagatctggggaaccctgtggttggcatgcacatacaaatggacgaacggataaac
ct
  2301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 2400
   aaagcccgccgccggttagaacgagcagagcgaccggccgcggtctagaccccttgggacaccaaccgtacgtgtatgtttacctgcttgcctatttg
ga

a     F  G  R  R  P  I  L  L  V  S  L  A  G  A  R  S  G  E  P  C  G  W  H  A  H  T  N  G  R  T  D  K  P
-
b      S  G  G  G  Q  S  C  S  S  R  W  P  A  P  D  L  G  N  P  V  V  G  M  H  I  Q  M  D  E  R  I  N
L  -
c     F  R  A  A  A  N  L  A  R  L  A  G  R  R  Q  I  W  G  T  L  W  L  A  C  T  Y  K  W  T  N  G  *  T
F  -

                                                                                               FauI
                                                                                          DraI |
                                                                                          PmeI |
             DraI                                                   AciI                 MseI| |
    TspGWI   MseI|  Hpy188I                            DdeI  Hpy8I   |     FauI         TspRI  || |
      |         ||     |                                 |      |    |       |             |   || |
   tttcacgccctttttaaatatccgattattctaataaacgctcttttctcttaggtttacccgccaatatatcctgtcaaacactgatagtttaaactg
aa
  2401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 2500
   aaagtgcgggaaaatttataggctaataagattatttgcgagaaaagagaatccaaatgggcggttatataggacagtttgtgactatcaaatttgac
tt

a     F  H  A  L  L  N  I  R  L  F  *  *  T  L  F  S  L  R  F  T  R  Q  Y  I  L  S  N  T  D  S  L  N  *
R  -
b      F  T  P  F  *  I  S  D  Y  S  N  K  R  S  F  L  L  G  L  P  A  N  I  S  C  Q  T  L  I  V  *  T  E
-
c       S  R  P  F  K  Y  P  I  I  L  I  N  A  L  F  S  *  V  Y  P  P  I  Y  P  V  K  H  *  *  F  K  L
K  -

                                 MseI
                          Tsp509I |
                         AciI  |  |
                        BsrBI  |  |
                     NlaIII |  |  |
                Hpy188III | |  |  |                                         CviJI
                   BspHI| | |  |  |                                   HpyF10VI |
                  BstKTI| | |  |  |                                      AciI| |
                    FatI| | |  |  |                MlyI                 BstUI| |
                    DpnI|| | |  |  |               PleI|          HpyF10VI |  | |
                   BclI ||| | |  |  |               TaiI|              MwoI| |  | |
               Hpy188I ||| | |  |  |      HpyCH4IV  ||             FauI|| |  | |
                  MboI ||| | |  |  |      MaeIII  |  ||           BceAI ||| |  | |                  TaiI
                Eco57I| ||| | |  |  |     Tsp45I  |  ||            FauI| ||| |  HgaI| |HpyCH4IV    |
   AciI        Eco57MI| ||| | |  |  |     HinfI|  |  ||       AciI  || ||| |  MwoI| |  BsmFI|      |
     |              || ||| | |  |  |          ||  |  ||          |   || ||| |    || |      ||      |
   ggcgggaaacgacaatctgatcatgagcggagaattaagggagtcacgttatgacccccgccgatgacgcgggacaagccgttttacgtttggaactg
ac
  2501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 2600
```

```
        ccgccctttgctgttagactagtactcgcctcttaattccctcagtgcaatactgggggcggctactgcgccctgttcggcaaaatgcaaaccttgac
tg

a        R E T T I * S * A E N * G S H V M T P A D D A G Q A V L R L E L
T -
b       G G K R Q S D H E R R I K G V T L * P P P M T R D K P F Y V W N *
Q -
c        A G N D N L I M S G E L R E S R Y D P R R * R G T S R F T F G T D
-
```

Figure 19-K

```
Hpy188III
BseMII|
BspCNI||
SacII |||
AciI| |||
BstUI| |||
MspA1I| |||
TauI| |||
AciI || |||
BsaJI || |||
BtgI || |||
Fnu4HI || |||
CviJI| || |||
TaiI DdeI || || |||
AclI | CviJI | || || |||
HpyCH4IV |
Bsu36I
AciI | | NlaIV|FauI || || ||| MseI CviJI|
DdeI
| | | || | || || ||| | ||
|
```

```
Hpy188I
TaiI |
HpyCH4IV | |
BpmI | | |
Eco57MI | | |
BlpI | | | |
DdeI | | | | BstUI
AluI| | | | | HhaI
HinP1I |
| | | | | ||
```

```
    agaaccgcaacgttgaaggagccactcagccgcgggtttctggagtttaatgagctaagcacatacgtcagaaaccattattgcgcgttcaaaagtcg
cc
2601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 2700
    tcttggcgttgcaacttcctcggtgagtcggcgcccaaagacctcaaattactcgattcgtgtatgcagtctttggtaataacgcgcaagttttcagc
gg

a     E  P  Q  R  *  R  S  H  S  A  A  G  F  W  S  L  M  S  *  A  H  T  S  E  T  I  I  A  R  S  K  V  A
-
b      N  R  N  V  E  G  A  T  Q  P  R  V  S  G  V  *  *  A  K  H  I  R  Q  K  P  L  L  R  V  Q  K  S
P -
c    R  T  A  T  L  K  E  P  L  S  R  G  F  L  E  F  N  E  L  S  T  Y  V  R  N  H  Y  C  A  F  K  S  R
L -
```

```
BmtI
Cac8I |
BfaI| |
AluI|| |
MaeIII CviJI|| |
Tsp45I NheI|| | SspI
| ||| | |
```

```
ApoI
Tsp509I
TaiI |
TspRI | |
HpyCH4IV| | |
|| | | BsaJI
MnlI BciVI
Tsp509I HinfI
| |
MlyI
PleI|
BsmAI|| BsgI
||| |
```

```
    taaggtcactatcagctagcaaatatttcttgtcaaaaatgctccactgacgttccataaattcccctcggtatccaattagagtctcatattcactc
tc
2701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 2800
    attccagtgatagtcgatcgtttataaagaacagtttttacgaggtgactgcaaggtatttaaggggagccataggttaatctcagagtataagtgag
ag

a     *  G  H  Y  Q  L  A  N  I  S  C  Q  K  C  S  T  D  V  P  *  I  P  L  G  I  Q  L  E  S  H  I  H  S
Q -
b      K  V  T  I  S  *  Q  I  F  L  V  K  N  A  P  L  T  F  H  K  F  P  S  V  S  N  *  S  L  I  F  T  L
-
c       R  S  L  S  A  S  K  Y  F  L  S  K  M  L  H  *  R  S  I  N  S  P  R  Y  P  I  R  V  S  Y  S  L
S -
```

```
BstKTI
AlwI|
DpnI|
MboI ||
BsaBI| ||
Hpy188III|| ||
BstKTI||| ||
DpnI|||| ||
BstYI ||||| ||
MboI ||||| ||
HpaII | ||||| ||
BsaWI| | ||||| ||
HpyCH4V || | ||||| ||
CviJI
| || | ||||| ||
|
```

```
NlaIII
FatI |
AlwI | |BccI
| | | |
Cac8I
HpyCH4V |
BspMI | |
| | |
```

```
TauI
AciI |
BsiEI |
Fnu4HI |
CviJI| |
HaeIII| |
EaeI || |
EagI || |
HpaII| || |
TaqII|| || |MnlI CviJI
||| || | | |
```

```
    aatccaaataatctgcaccggatctggatcgtttcgcatgattgaacaagatggattgcacgcaggttctccggccgcttgggtggagaggctattcg
gc
2801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 2900
    ttaggtttattagacgtggcctagacctagcaaagcgtactaacttgttctacctaacgtgcgtccaagaggccggcgaacccacctctccgataagc
cg

a     S  K  *  S  A  P  D  L  D  R  F  A  *  L  N  K  M  D  C  T  Q  V  L  R  P  L  G  W  R  G  Y  S
A -
b    N  P  N  N  L  H  R  I  W  I  V  S  H  D  *  T  R  W  I  A  R  R  F  S  G  R  L  G  G  E  A  I  R
L -
```

C I Q I I C T G S G S F R M I E Q D G L H A G S P A A W V E R L F G

Figure 19-L

```
                                                                             HpaII
                                                                              NciI
                                                                             ScrFI
                                                                             BbeI|
                                                                            HaeII|
                                                                            HhaI||
                                               TauI                        StyD4I||
                                             AciI |                        NlaIV|||
                                           Fnu4HI |                        SfoI|||
                                       Hpy188I  | |               HhaI  |  BsaHI||||
                 BmrI     Fnu4HI          |     | |           HinP1I |  BsaHI||||
                BbvI |    CviJI|          |     | |         HpyF10VI |HinP1I||||
BanI
 |           Bme1580I| |    TseI|         |     | |     CviJI      | |  NarI||||                    HpaII
 |           Bsp1286I| |   BceAI||        |     | |  HpaII |       | |  BanI|||||                  BsaWI|
 |              BsrI || |  SfaNI|||       |     | |  BslI| |  MwoI| |  KasI|||||    Hpy188III   DrdI  ||
 |                  | || |     ||||       |     | |     || |      || |  ||||||           |          |  ||
        tatgactgggcacaacagacaatcggctgctctgatgccgccgtgttccggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacctgtc
cg
  2901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 3000
        atactgacccgtgttgtctgttagccgacgagactacggcggcacaaggccgacagtcgcgtccccgcgggccaagaaaaacagttctggctggacag
gc

a        M  T  G  H  N  R  Q  S  A  A  L  M  P  P  C  S  G  C  Q  R  R  G  A  R  F  F  L  S  R  P  T  C  P
-
b         *  L  G  T  T  D  N  R  L  L  *  C  R  R  V  P  A  V  S  A  G  A  P  G  S  F  C  Q  D  R  P  V
R  -
c       Y  D  W  A  Q  Q  T  I  G  C  S  D  A  A  V  F  R  L  S  A  Q  G  R  P  V  L  F  V  K  T  D  L  S
G -
```

```
                                                                                          FauI
                                                                                      MaeIII |
                                                                                      Tsp45I |
                                                                                        TaiI| |
                                                      CviJI                           Hpy99I|| |
                                                     HaeIII                          Tth111I|| |
                                                       MscI                        HpyCH4IV ||| |
                                                     Cac8I |                          BbvI| ||| |
                                                      EaeI |                       BsiHKAI || ||| |
                                                    CviJI| |                      Bsp1286I || ||| |
                                               HpyF10VI | | |                  AluI    | || ||| |
                                             CviJI     | | | |                CviJI    | || ||| |
                                              TauI     | | | |               MspA1I    | || ||| |
                                           Fnu4HI |    | | | |                PvuII    | || ||| |
                                             AciI||    | | | |             Fnu4HI |    | || ||| |
                                            BstUI||    | | | |               HhaI||    | || ||| |
                                             HhaI||    | | | |               TseI||    | || ||| |
   Bme1580I          PstI              HinP1I |||    | | | |               FspI||    | || ||| |
   Bsp1286I       HpyCH4V |          Fnu4HI | |||    | | | |             HinP1I|||    | || ||| |         BbvI
    TaqII |          MnlI |           TseI| | |||  MwoI| | | |          HpyF10VI ||||    | || ||| |         AciI|
    NlaIV| |         SfcI | |TspDTI | | | |||  |BbvI || | | |  Hpy99I   MwoI| ||||   TaqI | || ||| |       TspRI||
         || |           | | |     | | | | |||  |   | || | | |    |        || ||||      |  | || ||| |         |||
        gtgccctgaatgaactgcaggacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagcg
gg
  3001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 3100
        cacgggacttacttgacgtcctgctccgtcgcgccgatagcaccgaccggtgctgcccgcaaggaacgcgtcgacacgagctgcaacagtgacttcgc
cc

a       V  P  *  M  N  C  R  T  R  Q  R  G  Y  R  G  W  P  R  R  A  F  L  A  Q  L  C  S  T  L  S  L  K  R
E -
b        C  P  E  *  T  A  G  R  G  S  A  A  I  V  A  G  H  D  G  R  S  L  R  S  C  A  R  R  C  H  *  S  G
-
c         A  L  N  E  L  Q  D  E  A  A  R  L  S  W  L  A  T  T  G  V  P  C  A  A  V  L  D  V  V  T  E  A
G  -
```

```
                                                                                          HpyCH4V
          Eco57I            NciI                                                     BccI  |
          Eco57MI           ScrFI                                                  CviJI  |
AciI
             BsrI|          BsaJI|    BstKTI                                      BciVI|   |
TauI
          Fnu4HI|           Hin4I|    DpnI|                                     NlaIII|   |
BsrDI|
          CviJI||           HpaII|  BstYI || AlwI          Hin4I              FatI
||Hin4I|Fnu4HI||
           TseI||BsmFI    StyD4I||   MboI ||HphI|        Hin4I |          SfaNI |    ||BbvI||
AciI|||
                |||    |         |||     | ||   ||            | |              | |     ||     |||
||||
```

```
         aagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgc
gg
   3101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 3200
         ttccctgaccgacgataacccgcttcacggccccgtcctagaggacagtagagtggaacgaggacggctctttcataggtagtaccgactacgttacg
cc

a         G  T  G  C  Y  W  A  K  C  R  G  R  I  S  C  H  L  T  L  L  L  P  R  K  Y  P  S  W  L  M  Q  C
G  -
b        K  G  L  A  A  I  G  R  S  A  G  A  G  S  P  V  I  S  P  C  S  C  R  E  S  I  H  H  G  *  C  N  A
A -
c         R  D  W  L  L  L  G  E  V  P  G  Q  D  L  L  S  S  H  L  A  P  A  E  K  V  S  I  M  A  D  A  M  R
-
```

Figure 19-M

```
              AlwI
           HpyF10VI
             MwoI|
         HpyCH4V  ||
        Fnu4HI |  ||
      HpyF10VI |  ||
        CviJI| |  ||         CviJI
Hpy188III
        MwoI| |  ||     HpaII  |
BstKTI|
        TauI| |  ||   BstKTI|  |
DpnI||
        TseI| |  ||    DpnI||  |          TaqI
|||
     Fnu4HI || |  || MboI  |||  |         BspMI|
|||
           | || |  ||    | |||  |           ||
|||
        cggctgcatacgcttgatccggctacctgcccattcgaccaccaagcgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcga
tc
   3201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 3300
        gccgacgtatgcgaactaggccgatggacgggtaagctggtggttcgctttgtagcgtagctcgctcgtgcatgagcctaccttcggccagaacagct
ag

a        G  C  I  R  L  I  R  L  P  A  H  S  T  T  K  R  N  I  A  S  S  E  H  V  L  G  W  K  P  V  L  S  I
-
b         A  A  Y  A  *  S  G  Y  L  P  I  R  P  P  S  E  T  S  H  R  A  S  T  Y  S  D  G  S  R  S  C  R
S  -
c        R  L  H  T  L  D  P  A  T  C  P  F  D  H  Q  A  K  H  R  I  E  R  A  R  T  R  M  E  A  G  L  V  D
Q -
```

Right-hand enzyme labels, upper block: Hpy188I, BccI, RsaI, TaiI, Csp6I, BsaAI, BsiHKAI, Bsp1286I, HpyCH4IV, Cac8I, TaqI, SfaNI, BstF5I, HpaII, BsrFI, CviJI, FokI, MboI, TaqI

```
                                                                                  NlaIII
                                                                                    NspI
                                                                                    SphI
                                                                               Cac8I |
                                                                              FatI | |
                                                                              HhaI | |
                                                                           BstUI | | |
                                                                           Cac8I | | |
                                                                            HhaI | | |
                                                                         HinP1I | | |
                                                                          BssHII | | |
                                                                          HinP1I | | |
                                                                        HpyF10VI | | |
                                                                         SmlI
                                                                      CviJI|
                                                                   BstNI |
                                                                   ScrFI ||
                                                                  PspGI  ||
                                                                 StyD4I  ||
                                                        HpyCH4III
                                                            BpuEI
                                                         CviJI
                                                       Cac8I
                                                       HhaI
                                                      BstUI
                                                    HinP1I
                                                     BanII
                                                  Bsp1286I
                                                     Cac8I                                                BsaJI
       EarI                                          CviJI                                                 BtgI
       SapI                                         SfaNI                                                  FatI
    Hpy188III|                                   HpyF10VI                                                  NcoI
      BstF5I||                                      MboII                                                  StyI
      BstKTI||                                       MwoI                                           SfaNI |
        DpnI|||                                                                                   Hpy99I | |
        FalI||||                                                                                  BceAI| | |
       MboI|||||  FokI                                                     MwoI|             Hpy188III | | |
           ||||||   |                                                                        BstKTI  || | |
                                                                                     Hpy99I MboI  DpnI|MaeIII
                                                                                                   ||Tsp45I
        aggatgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgatgatctcgtcgtgacc
ca
   3301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 3400
        tcctactagacctgcttctcgtagtccccgagcgcggtcggcttgacaagcggtccgagttccgcgcgtacgggctgccgctactagagcagcactgg
gt

a       R  M  I  W  T  K  S  I  R  G  S  R  Q  P  N  C  S  P  G  S  R  R  A  C  P  T  A  M  I  S  S  *  P
M -
b       G  *  S  G  R  R  A  S  G  A  R  A  S  R  T  V  R  Q  A  Q  G  A  H  A  R  R  R  *  S  R  R  D  P
-
c        D  D  L  D  E  E  H  Q  G  L  A  P  A  E  L  F  A  R  L  K  A  R  M  P  D  G  D  D  L  V  V  T
H  -
```

Figure 19-N

```
                                                                              BseYI
                                                                              CviJI
                                                                            Cac8I |
                                                         HinfI               NaeI |
                                                          TfiI              HpaII||
                                                  Hpy188III |               BsrFI|||
                                                   TspDTI | |               CviJI|||     Hpy188III   EciI
                                                     TauI|| |             HaeIII|||    AciI     |    |
                 Cac8I                               AciI||| |           NgoMIV|||  AvaII |      |    |
           Cac8I   |                               Fnu4HI||| |          EaeI ||||   RsrII |      |    |
       HpyF10VI    |                               CviJI|||| |   HpyCH4III | ||||  Sau96I |      |    |
    MslI      |    |              NlaIII           HaeIII||| |    TaqII  | ||||   AciI |  |  BslI|    |
 NlaIII|  MwoI|    |          FatI   |             EaeI |||| |  TaqI  |  | ||||      |  |  |    ||    |
      ||      ||   |             |   |                | |||| |    |   |  | ||||      |  |  |    ||    |
       tggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttctggattcatcgactgtggccggctgggtgtggcggaccgctatcaggaca
ta
  3401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 3500
       accgctacggacgaacggcttatagtaccacctttaccggcgaaaagacctaagtagctgacaccggccgacccacaccgcctggcgatagtcctgt
at

a       A  M  P  A  C  R  I  S  W  W  K  M  A  A  F  L  D  S  S  T  V  A  G  W  V  W  R  T  A  I  R  T
*  -
b      W  R  C  L  L  A  E  Y  H  G  G  K  W  P  L  F  W  I  H  R  L  W  P  A  G  C  G  G  P  L  S  G  H
S -
c       G  D  A  C  L  P  N  I  M  V  E  N  G  R  F  S  G  F  I  D  C  G  R  L  G  V  A  D  R  Y  Q  D  I
-

                                              CviJI
                                           Eco57I |
                                          Eco57MI |                                HinfI
                                       TauI     | |                                 TfiI
                                     MboII|     | |                         Hpy188III |
                                   Fnu4HI||     | |                            BsrBI  | |       HhaI
                  EarI       AluI       |||     | |                   HpyCH4III  TauI | | |  HinP1I |
     CviJI        SapI      CviJI AciI|||     | |   AciI   BssSI   MnlI  AciI  | | |Fnu4HI| |
       |            |          |      ||||     | |    |       |       |   Fnu4HI | | |  TseI| | |
       gcgttggctacccgtgatattgctgaagagcttggcggcgaatgggctgaccgcttcctcgtgctttacggtatcgccgctcccgattcgcagcgcat
cg
  3501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 3600
       cgcaaccgatgggcactataacgacttctcgaaccgccgcttacccgactggcgaaggagcacgaaatgccatagcggcgagggctaagcgtcgcgta
gc

a       R  W  L  P  V  I  L  L  K  S  L  A  A  N  G  L  T  A  S  S  C  F  T  V  S  P  L  P  I  R  S  A  S
-
b        V  G  Y  P  *  Y  C  *  R  A  W  R  R  M  G  *  P  L  P  R  A  L  R  Y  R  R  S  R  F  A  A  H
R -
c       A  L  A  T  R  D  I  A  E  E  L  G  G  E  W  A  D  R  F  L  V  L  Y  G  I  A  A  P  D  S  Q  R  I
A -

                                     HinfI
                                  AciI |
                                 BsrBI |                                        BspMI
                                 MlyI| |                                        BssSI
                                 PleI| |                                   Hpy188III
                    BspCNI  Hpy188I|| |                              HpyF10VI   |
                   BseMII|     DdeI||| |          BsmFI           TaqII      |   |        HinfI
    SfaNI          MboII ||  FauI |||| |          BstBI|       Hpy99I| HgaI|   |         TfiI
    BbvI | Hpy188III |   ||    |  |||| |          TaqI| BplI   BsaHI|| MwoI|   |BccI TaqI |
       | |         | |   ||    |  |||| |             ||   |        |||    ||   |  |    |  |
      ccttctatcgccttcttgacgagttcttctgagcgggactctggggttcgaaatgaccgaccaagcgacgcccaacctgccatcacgagatttcgatt
cc
  3601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 3700
      ggaagatagcggaagaactgctcaagaagactcgccctgagaccccaagctttactggctggttcgctgcgggttggacggtagtgctctaaagctaa
gg

a      P  S  I  A  F  L  T  S  S  S  E  R  D  S  G  V  R  N  D  R  P  S  D  A  Q  P  A  I  T  R  F  R  F
H -
b       L  L  S  P  S  *  R  V  L  L  S  G  T  L  G  F  E  M  T  D  Q  A  T  P  N  L  P  S  R  D  F  D  S
-
c        F  Y  R  L  L  D  E  F  F  *  A  G  L  W  G  S  K  *  P  T  K  R  R  P  T  C  H  H  E  I  S  I
P -
```

Figure 19-O

```
                                                                                            MboII
                                                                                           NlaIII
                                                                                             AlwI|
                                                                                          FatI  ||
                                                                                      BstKTI |  ||
                                                                                       DpnI| |  ||
                                                                                     BstYI || |  ||
                                                                                     MboI || |  ||
                                                                                   MnlI | || |  ||
                                                                                  BslI| | || |  ||
                                                                                 AciI|| | || |  ||
                                                                                BstUI|| | || |  ||
                                                                                HhaI|| | || |  ||
                                                                               FokI ||| | || |  ||
                                                                             HinP1I ||| | || |  ||
                                                                      FauI     |||| | || |  ||
                                                                   BstF5I |    |||| | || |  ||
                                                                   BstKTI |    |||| | || |  ||
                                                                    DpnI| |    |||| | || |  ||
                                                                   MboI || |    |||| | || |  ||
                                                                 BsmFI| || |    |||| | || |  ||
                                                               HgaI || || |    |||| | || |  ||
                                                              AlwI| || || |    |||| | || |  ||
                                                             CviJI|| || || |    |||| | || |  ||
                                                            Cac8I ||| || || |    |||| | || |  ||
                                                             NaeI ||| || || |    |||| | || |  ||
                                                           HpaII| ||| || || |    |||| | || |  ||
                                                            BpmI|| ||| || || |    |||| | || |  ||
                                                           BsrFI|| ||| || || |    |||| | || |  ||
                                                         Eco57MI|| ||| || || |    |||| | || |  ||
                                                          NgoMIV|| ||| || || |    |||| | || |  ||
                                                          BsaHI ||| ||| || || |    |||| | || |  ||
                        HinfI                    NciI  | ||| ||| || || |    |||| | || |  ||
          TauI           TfiI                   ScrFI  | ||| ||| || || |    |||| | || |  ||
          AciI |      Hpy188I|                 HpaII|  | ||| ||| || || |    |||| | || |  ||
        Fnu4HI |      TspDTI ||                PfoI||  | ||| ||| || || |    |||| | || |  ||
BsaJI
       AciI | |       CviJI | ||          StyD4I||  | ||| ||| || || |    |||| | || |  ||
BtgI
          | | |           | | ||             |||  | ||| ||| || || |    |||| | || |  ||
|
        accgccgccttctatgaaaggttgggcttcggaatcgttttccgggacgccggctggatgatcctccagcgcggggatctcatgctggagttcttcgc
cc
   3701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 3800
        tggcggcggaagatactttccaacccgaagccttagcaaaaggccctgcggccgacctactaggaggtcgcgcccctagagtacgacctcaagaagcg
gg

a         R  R  L  L  *  K  V  G  L  R  N  R  F  P  G  R  R  L  D  D  P  P  A  R  G  S  H  A  G  V  L  R
p  -
b        T  A  A  F  Y  E  R  L  G  F  G  I  V  F  R  D  A  G  W  M  I  L  Q  R  G  D  L  M  L  E  F  F  A
H -
c         P  P  P  S  M  K  G  W  A  S  E  S  F  S  G  T  P  A  G  *  S  S  S  A  G  I  S  C  W  S  S  S  P
-
```

```
                                                                                  Bpu10I
                                                                                    DdeI
                                                                              Hpy188III|
                                                                                 BstKTI||
Sau96I
          BstKTI                                                                   DpnI|||
PspOMI|
           BpmI|            BsiEI                           BsiEI                 MboI ||||
Sau96I|
           DpnI|          AciI   |                         CviJI|          BspCNI | ||||           BstKTI
||
        Eco57MI|    HpyF10VI |   |                        HaeIII|         BseMII| | ||||            DpnI|
||
         BstYI || AciI     | |   |  NlaIV                   EaeI ||          AlwI|| | ||||           HgaI||
||
          MboI || AlwI   MwoI| |TaqI BanI | EcoRV            EagI ||         BceAI||| | ||||          MboI|||
||
             | ||     |    || |   |    | |    |                 | ||          |||| | ||||            ||||
||
        acgggatctctgcggaacaggcggtcgaaggtgccgatatcattacgacagcaacggccgacaagcacaacgccacgatcctgagcgacaatatgatc
gg
   3801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 3900
        tgccctagagacgccttgtccgccagcttccacggctatagtaatgctgtcgttgccggctgttcgtgttgcggtgctaggactcgctgttatactag
cc
```

a R D L C G T G G R R C R Y H Y D S N G R Q A Q R H D P E R Q Y D R
-
b G I S A E Q A V E G A D I I T T A T A D K H N A T I L S D N M I
G -
c T G S L R N R R S K V P I S L R Q Q R P T S T T P R S * A T I * S
G -

Figure 19-P

```
                                          SfaNI
                                      BstNI |
                                      ScrFI |
                 HgaI               BsaJI | |
                Hpy8I               PspGI | |
             BsaHI  |              StyD4I | |
            ApaI |  |             BceAI | | |
           BanII |  |           HpyF10VI| | |
        Bme1580I |  |              MwoI | | |
        Bsp1286I |  |             TauI || | |
           HpaII |  |         Fnu4HI | || | |                     EcoRV
            NciI |  |       HpyF10VI | || | |                  TaqII |
           ScrFI |  |           AciI|| || | |               BstUI | |
         CviJI | |  |           MwoI|| || | |              AciI | | |
        HaeIII | |  |         Hpy99I||| || | |            BbvI| | | | Fnu4HI
          NlaIV| |  |   BsaHI      ||| || | |      HpyCH4V || | | |  TseI| Hpy188I          AciI
        StyD4I | |  |     |        ||| || | |            | || | | |     || |                 |
       gcccggcgtccacatcaacggcgtcggcggcgactgcccaggcaagaccgagatgcaccgcgatatcttgctgcgttcggatatttcgtggagttcc
cg
  3901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 4000
       cgggccgcaggtgtagttgccgcagccgccgctgacgggtccgttctggctctacgtggcgctatagaacgacgcaagcctataaaagcacctcaagg
gc

a       A  R  R  P  H  Q  R  R  R  R  R  L  P  R  Q  D  R  D  A  P  R  Y  L  A  A  F  G  Y  F  R  G  V  P
A -
b        P  G  V  H  I  N  G  V  G  G  D  C  P  G  K  T  E  M  H  R  D  I  L  L  R  S  D  I  F  V  E  F  P
-
c         P  A  S  T  S  T  A  S  A  A  T  A  Q  A  R  P  R  C  T  A  I  S  C  C  V  R  I  F  S  W  S  S
R  -

            AlwI
           HpaII          BsiEI
            NciI         BstKTI
           ScrFI  BstF5I  FokI                                                   HpyF10VI
        StyD4I |  BstKTI  PvuI                  MseI                       HpaII      |
         Hin4I|| DpnI|  DpnI|                  AflII|    HinfI             BslI|      |
         FauI||| MboI ||MboI ||       Hin4I      SmlI|     TfiI           BsrFI| MwoI|   BsaBI  Tsp509I
            ||| |   | ||   | ||          |          ||        |               || ||       |        |
       ccacagacccggatgatccccgatcgttcaaacatttggcaataaagtttcttaagattgaatcctgttgccggtcttgcgatgattatcatataatt
tc
  4001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 4100
       ggtgtctgggcctactaggggctagcaagtttgtaaaccgttatttcaaagaattctaacttaggacaacggccagaacgctactaatagtatattaa
ag

a       T  D  P  D  D  P  R  S  F  K  H  L  A  I  K  F  L  K  I  E  S  C  C  R  S  C  D  D  Y  H  I  I
S  -
b      P  Q  T  R  M  I  P  D  R  S  N  I  W  Q  *  S  F  L  R  L  N  P  V  A  G  L  A  M  I  I  I  *  F
L -
c       H  R  P  G  *  S  P  I  V  Q  T  F  G  N  K  V  S  *  D  *  I  L  L  P  V  L  R  *  L  S  Y  N  F
-

                                      TaiI
                                HpyCH4IV |
                                 NlaIII| |
                                   NsiI|| |
                                BfrBI ||| |
                                 FatI ||| |
                              HpyCH4V ||| |
                           NlaIII | ||| |
                        AflIII  | | ||| |
                          FatI  | | ||| |
                          PciI  | | ||| |
                      MseI  |   | | ||| |                                                FauI
        MseI  Tsp509I  |    |   | | ||| |                                             MlyI |
        TaiI NlaIII |  |    |   | | ||| |                                             PleI| |
   HpyCH4IV |   NspI |  |    |   | | ||| |                                         Tsp509I| |
 Tsp509I  | |FatI  | |  |    |NspI| | ||| |BccI     BsmFI                   HinfI AciI  || |     MseI
       |  | |  |   | |  |    |   | | ||| |  |          |                         |    |  || |        |
       tgttgaattacgttaagcatgtaataattaacatgtaatgcatgacgttatttatgagatgggtttttatgattagagtcccgcaattatacatttaa
ta
  4101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 4200
       acaacttaatgcaattcgtacattattaattgtacattacgtactgcaataaatactctacccaaaaatactaatctcagggcgttaatatgtaaatt
at

a       V  E  L  R  *  A  C  N  N  *  H  V  M  H  D  V  I  Y  E  M  G  F  Y  D  *  S  P  A  I  I  H  L  I
-
b        L  N  Y  V  K  H  V  I  I  N  M  *  C  M  T  L  F  M  R  W  V  F  M  I  R  V  P  Q  L  Y  I  *
Y  -
c       C  *  I  T  L  S  M  *  *  L  T  C  N  A  *  R  Y  L  *  D  G  F  L  *  L  E  S  R  N  Y  T  F  N
T -
```

Figure 19-Q

```
                                         AciI
CviJI
                                        BstUI
TauI
                    HhaI                 HhaI                                            Fnu4HI
|
                   Hin4I              BstUI |                          CviJI               AciI|
|
                  BstUI |             Cac8I |                        HaeIII                TauI|
|
                  Cac8I |              HhaI |                       Sau96I |             Fnu4HI ||
|
                   HhaI |            HinP1I |                      BstKTI | |              AciI| ||
|
                 HinP1I |            BssHII | |         MaeIII     DpnI| | |              Hin4I| ||
|
                BssHII | |   Tsp509I  BstUI | |        Hin4I|    MboI || | |            Cac8I || ||
|
   BstUI        HinP1I | | BfaI     | HinP1I | |   Hin4I    || BfaI | || | |         MnlI   | || ||
|
       |              | | |  |      |       | | |        |   ||    | | || | |         |    | || ||
|
       cgcgatagaaaacaaaatatagcgcgcaaactaggataaattatcgcgcgcggtgtcatctatgttactagatcgggcctcctgtcaatgctggcggc
gg
  4201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 4300
       gcgctatcttttgttttatatcgcgcgtttgatcctatttaatagcgcgcgccacagtagatacaatgatctagcccggaggacagttacgaccgccg
cc

a      R  D  R  K  Q  N  I  A  R  K  L  G  *  I  I  A  R  G  V  I  Y  V  T  R  S  G  L  L  S  M  L  A  A
A -
b       A  I  E  N  K  I  *  R  A  N  *  D  K  L  S  R  A  V  S  S  M  L  L  D  R  A  S  C  Q  C  W  R  R
-
c        R  *  K  T  K  Y  S  A  Q  T  R  I  N  Y  R  A  R  C  H  L  C  Y  *  I  G  P  P  V  N  A  G  G
G  -

                                                                      Hpy188I
                                                                         DdeI|
                                                                        MnlI||
                                                                     CviJI |||
                                                                      TauI |||
                                                                   Fnu4HI | |||
                                                                      MnlI | |||
                                                                     AciI| | |||
                                                                 BspCNI || | |||
                                                                BseMII| || | |||
                                                          Hpy188I    || || | |||
                                                       MnlI     |    || || | |||
                                                      AciI|     |    || || | |||
                                                  BspCNI ||     |    || || | |||
                                                 BseMII| ||     |    || || | |||
                                           Hpy188I    || ||     |    || || | |||
                                              DdeI|   || ||     |    || || | |||
                                          CviJI  ||   || ||     |    || || | |||
                                         MnlI |  ||   || ||     |    || || | |||
                                   BspCNI  |  |  ||   || ||     |    || || | |||
                                  BseMII|  |  |  ||   || ||     |    || || | |||
                            Hpy188I    ||  |  |  ||   || ||     |    || || | |||
                               DdeI|   ||  |  |  ||   || ||     |    || || | |||
                           CviJI  ||   ||  |  |  ||   || ||     |    || || | |||
                            TauI  ||   ||  |  |  ||   || ||     |    || || | |||
                       Fnu4HI |   ||   ||  |  |  ||   || ||     |    || || | |||
                         MnlI |   ||   ||  |  |  ||   || ||     |    || || | |||                HpaII
                        AciI| |   ||   ||  |  |  ||   || ||     |    || || | |||                BsaWI|
                    BspCNI || |   ||   ||  |  |  ||   || ||     |    || || | |||               NlaIV||
                   BseMII| || |   ||   ||  |  |  ||   || || DdeI|    || || | |||        AciI   |||        CviJI
                        || || |   ||   ||  |  |  ||   || ||    ||    || || | |||         |     |||            |
       ctctggtggtggttctggtggcggctctgagggtggtggctctgagggtggcggttctgagggtggcggctctgagggaggcggttccggtggtggct
ct
  4301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 4400
       gagaccaccaccaagaccaccgccgagactcccaccaccgagactcccaccgccaagactcccaccgccgagactccctccgccaaggccaccaccga
ga

a        L  V  V  V  L  V  A  A  L  R  V  V  A  L  R  V  A  V  L  R  V  A  A  L  R  E  A  V  P  V  V  A
L  -
b       L  W  W  W  F  W  W  R  L  *  G  W  W  L  *  G  W  R  F  *  G  W  R  L  *  G  R  R  F  R  W  W  L
W -
c       S  G  G  G  S  G  G  G  S  E  G  G  G  S  E  G  G  G  S  E  G  G  G  S  E  G  G  G  S  G  G  G  S
-
```

```
                                                                                       Hpy188I
                                                                                  HpyCH4III  |
                                                                                     TspDTI  |
                                                                                  SfcI    |  |
                                                                                 HhaI|    |  |
      HpaII                                                                     BstUI ||    |  |
     BsaWI|              BccI                                                  HinP1I ||    |  |
    NlaIV||            HphI |         TspDTI          CviJI               TaqII     | ||    |  |         HgaI
        |||                 | |              |               |                    |     | ||    |  |            |
        ggttccggtgattttgattatgaaaagatggcaaacgctaataagggggctatgaccgaaaatgccgatgaaaacgcgctacagtctgacgctaaagg
ca
   4401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 4500
        ccaaggccactaaaactaatacttttctaccgtttgcgattattcccccgatactggcttttacggctacttttgcgcgatgtcagactgcgatttcc
gt

a         V  P  V  I  L  I  M  K  R  W  Q  T  L  I  R  G  L  *  P  K  M  P  M  K  T  R  Y  S  L  T  L  K  A
-
b          F  R  *  F  *  L  *  K  D  G  K  R  *  *  G  G  Y  D  R  K  C  R  *  K  R  A  T  V  *  R  *  R
Q  -
c         G  S  G  D  F  D  Y  E  K  M  A  N  A  N  K  G  A  M  T  E  N  A  D  E  N  A  L  Q  S  D  A  K  G
K -
```

Figure 19-R

```
                                                                            CviJI
                                          ClaI                            HaeIII
                                          TaqI                             HphI|
                                      TspDTI |                          HpaII ||
                                        BccI | |                       TaiI | ||
                                   Fnu4HI |  | |              HpyCH4IV    | | ||
        HinfI          FalI        TseI|  |  | |            MaeIII     |  | | ||
         TfiI          BbvI|HpyCH4III ||  |  | |            Tsp45I     |  | | ||                     BsrI
            |              ||          || |  | |                 |     |  | | ||                        |
        aacttgattctgtcgctactgattacggtgctgctatcgatggtttcattggtgacgtttccggccttgctaatggtaatggtgctactggtgatttt
gc
   4501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 4600
        ttgaactaagacagcgatgactaatgccacgacgatagctaccaaagtaaccactgcaaaggccggaacgattaccattaccacgatgaccactaaaa
cg

a       N  L  I  L  S  L  L  I  T  V  L  L  S  M  V  S  L  V  T  F  P  A  L  L  M  V  M  V  L  L  V  I  L
L -
b        T  *  F  C  R  Y  *  L  R  C  C  Y  R  W  F  H  W  *  R  F  R  P  C  *  W  *  W  C  Y  W  *  F  C
-
c         L  D  S  V  A  T  D  Y  G  A  A  I  D  G  F  I  G  D  V  S  G  L  A  N  G  N  G  A  T  G  D  F
A  -

                                                       HphI
        Tsp509I                                   Tsp509I|             XmnI
          BpuEI|                                     HphI ||         Tsp509I|
        CviJI ||                          HpyCH4III|     ||     TspGWI     ||         SspI                   BsaXI
        HphI| ||              SmlI  MaeIII      |     ||     MseI  |     ||    TspDTI  |              MnlI  |
      Cac8I|| ||             CviJI|  Tsp45I     |     ||   HphI |  |     ||   BsaXI |  |             BslI|MnlI
         ||| ||                  ||       |     ||    ||      |  |  |    ||       | |  |                || |
        tggctctaattcccaaatggctcaagtcggtgacggtgataattcacctttaatgaataatttccgtcaatatttaccttccctccctcaatcggttg
aa
   4601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 4700
        accgagattaagggtttaccgagttcagccactgccactattaagtggaaattacttattaaaggcagttataaatggaagggagggagttagccaac
tt

a         A  L  I  P  K  W  L  K  S  V  T  V  I  I  H  L  *  *  I  I  S  V  N  I  Y  L  P  S  L  N  R  L
N  -
b       W  L  *  F  P  N  G  S  S  R  *  R  *  *  F  T  F  N  E  *  F  P  S  I  F  T  F  P  P  S  I  G  *
M -
c        G  S  N  S  Q  M  A  Q  V  G  D  G  D  N  S  P  L  M  N  N  F  R  Q  Y  L  P  S  L  P  Q  S  V  E
-

                                                          HinfI
                                                           TfiI
                                                       CviJI |
                                                      HaeIII |
                                                      BsaXI| |
                                                       EaeI|| |
                                                      FauI||| |             Cac8I
                                                      BslI|||| |              AluI |
                                                   TspDTI||||| |             CviJI |
                        HpyF10VI                   BstUI|||||| |            MspA1I |
                            MwoI|                   HhaI|||||| |             PvuII |
                        BsaXI ||                    MnlI||||||| |          Fnu4HI | |
                      CviJI | ||                   BstUI|||||||| |        HpyCH4V| | |
AciI
                     HaeIII | ||               HinP1I|||||||| |  AseI       || | |       Hpy188III      BsrI
|
                    Sau96I| | ||   AciI        AciI ||||||||| |  MseI TseI| | |    BbvI           | FauI  |
|
|                         || | ||    |          | |||||||| |     |      || | |      |          |   |  |
        tgtcgcccttttgtctttggcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaa
ag
   4701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 4800
        acagcgggaaaacagaaaccgggttatgcgtttggcggagaggggcgcgcaaccggctaagtaattacgtcgaccgtgctgtccaaagggctgacctt
tc

a         V  A  L  L  S  L  A  Q  Y  A  N  R  L  S  P  R  V  G  R  F  I  N  A  A  G  T  T  G  F  P  T  G  K
-
b          S  P  F  C  L  W  P  N  T  Q  T  A  S  P  R  A  L  A  D  S  L  M  Q  L  A  R  Q  V  S  R  L  E
S  -
c       C  R  P  F  V  F  G  P  I  R  K  P  P  L  P  A  R  W  P  I  H  *  C  S  W  H  D  R  F  P  D  W  K
A -

                                                            CviJI
                       AseI                            HpyF10VI |
                       MseI                               BstNI| |
                Tsp509I |                                  MwoI| |
             HhaI     | |                                 ScrFI| |
        BtsI |        | |                                BsaJI|| |
```

```
                HinPlI |      | |                      PspGI ||  |
              HpyF10VI |      | |                     StyD4I ||  |
                 TspRI |      | |          AluI        NlaIV |  ||  |               CviJI
        Cac8I    MwoI| |      | |          CviJI       BanI |  |  ||  |             HpaII  |              Tsp509I
          |          || |      | |            |             |  |  |  ||  |               |  |                   |
          cgggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaatt
gt
    4801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 4900
          gcccgtcactcgcgttgcgttaattacactcaatcgagtgagtaatccgtggggtccgaaatgtgaaatacgaaggccgagcatacaacacaccttaa
ca

a         R  A  V  S  A  T  Q  L  M  *  V  S  S  L  I  R  H  P  R  L  Y  T  L  C  F  R  L  V  C  C  V  E  L
* -
b          G  Q  *  A  Q  R  N  *  C  E  L  A  H  S  L  G  T  P  G  F  T  L  Y  A  S  G  S  Y  V  V  W  N  C
-
c           G  S  E  R  N  A  I  N  V  S  *  L  T  H  *  A  P  Q  A  L  H  F  M  L  P  A  R  M  L  C  G  I
V  -
```

Figure 19-S

```
                                                                          NlaIV
                                                                       AvaII |
                                                                   Eco0109I |
                                                                      PpuMI |
                                                                     Sau96I |
                                                                       PstI| |
                                                                       SbfI| |
                                                                HpyCH4V || |
                                                                  Cac8I | || |
                                                                   SfcI | || |
                                                              NlaIII | | || |
                                                                NspI | | || |
                                                                SphI | | || |
                                                           Cac8I | | | || |
                                                          FatI | | | | || |
                                                      HpyCH4V | | | | || |
                                                        BspMI| | | | || |
                                                       Cac8I|| | | | || |
                                                    HpyF10VI|| | | | || |
                                                       BsmFI||| | | | || |
                                                        MwoI||| | | | || |
                                                        AluI||| | | | || |
    AciI                           AluI      NlaIII      CviJI|||| | | | || |                Tsp509I
    BsrBI Tsp509I                  CviJI   FatI  |    HindIII |||| | | | || |       CviJI    |Hpy188I
       |       |                       |      |  |          | |||| | | | || |          |       |   |
    gagcggataacaatttcacacaggaaacagctatgaccatgattacgccaagcttgcatgcctgcaggtccccagattagccttttcaatttcagaaa
ga
4901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 5000
    ctcgcctattgttaaagtgtgtcctttgtcgatactggtactaatgcggttcgaacgtacggacgtccaggggtctaatcggaaaagttaaagtcttt
ct

a      A  D  N  N  F  T  Q  E  T  A  M  T  M  I  T  P  S  L  H  A  C  R  S  P  D  *  P  F  Q  F  Q  K
E  -
b    E  R  I  T  I  S  H  R  K  Q  L  *  P  *  L  R  Q  A  C  M  P  A  G  P  Q  I  S  L  F  N  F  R  K
N -
c    S  G  *  Q  F  H  T  G  N  S  Y  D  H  D  Y  A  K  L  A  C  L  Q  V  P  R  L  A  F  S  I  S  E  R
-

                                                             AvaI
                                                         BpmI |
                                                      Eco57MI |
                                 Fnu4HI    Hpy188III  BstKTI | |                 BstNI
                               HpyF10VI       BbvI |    DpnI| | |                 ScrFI
      BccI      MnlI       BspMI  MwoI|         BsaI |   Hin4I|| | |                PfoI |
    BsmI |    BslI |       CviJI  TseI|        BsmAI |   MboI||| | |              PspGI |
       |  |       |  |          |   ||            |  |    ||||| | |             StyD4I |        Hin4I        SfaNI
    atgctaacccacagatggttagagaggcttacgcagcaggtctcatcaagacgatctacccgagcaataatctccaggaaatcaaataccttcccaag
aa
5001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 5100
    tacgattgggtgtctaccaatctctccgaatgcgtcgtccagagtagttctgctagatgggctcgttattagaggtcctttagtttatggaagggttc
tt

a     C  *  P  T  D  G  *  R  G  L  R  S  R  S  H  Q  D  D  L  P  E  Q  *  S  P  G  N  Q  I  P  S  Q  E
-
b     A  N  P  Q  M  V  R  E  A  Y  A  A  G  L  I  K  T  I  Y  P  S  N  N  L  Q  E  I  K  Y  L  P  K
K  -
c    M  L  T  H  R  W  L  E  R  L  T  Q  Q  V  S  S  R  R  S  T  R  A  I  I  S  R  K  S  N  T  F  P  R
R -

                                                                                    RsaI
                                                                                    ScaI
                                                                                 Csp6I|
                                                                                 TatI||
                                                                         Hpy188I  |||
                                                                        BstKTI |  |||
                                                                         DpnI| |  |||
                                                                        MboI || |  |||
          Hpy188III                                                            
            HinfI |          Hpy188III      BpuEI       Hpy188III | || |  |||              HinfI
MseI HpyCH4V  TfiI |       HpyCH4V   |SfaNI   |          SmlI | | || |  |||     BsrI      TfiI
   |      |      |  |             |  |  |     |                | | | || |  |||      |          |
    ggttaaagatgcagtcaaaagattcaggactaactgcatcaagaacacagagaaagatatatttctcaagatcagaagtactattccagtatggacga
tt
5101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 5200
    ccaatttctacgtcagttttctaagtcctgattgacgtagttcttgtgtctctttctatataaagagttctagtcttcatgataaggtcatacctgct
aa

a    G  *  R  C  S  Q  K  I  Q  D  *  L  H  Q  E  H  R  E  R  Y  I  S  Q  D  Q  K  Y  Y  S  S  M  D  D
S -
b     V  K  D  A  V  K  R  F  R  T  N  C  I  K  N  T  E  K  D  I  F  L  K  I  R  S  T  I  P  V  W  T  I
-
```

C L K M Q S K D S G L T A S R T Q R K I Y F S R S E V L F Q Y G R
F -

Figure 19-T

```
                                                                                HinfI
                                                                              NlaIII|
                                                                           BsaJI  ||
                                                                            BtgI  ||
                                                                            FatI  ||      MnlI
                                                                            NcoI  ||   MlyI |
                                          MlyI                    TspRI     StyI  ||  HinfI| |
       Cac8I        BsaJI                 PleI|              HinfI |   CviJI|  ||   PleI| |
       CviJI |      StyI   FalI     HinfI BsmAI||              TfiI |  HaeIII|  ||   TfiI| |
caaggcttgcttcacaaaccaaggcaagtaatagagattggagtctctaaaaaggtagttcccactgaatcaaaggccatggagtcaaagattcaaatag
5201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5300
gttccgaacgaagtgtttggttccgttcattatctctaacctcagagatttttccatcaagggtgacttagtttccggtacctcagtttctaagtttatc

a    R  L  A  S  Q  T  K  A  S  N  R  D  W  S  L  *  K  G  S  S  H  *  I  K  G  H  G  V  K  D  S  N  R -
b   Q  G  L  L  H  K  P  R  Q  V  I  E  I  G  V  S  K  K  V  V  P  T  E  S  K  A  M  E  S  K  I  Q  I  E -
c   K  A  C  F  T  N  Q  G  K  *  *  R  L  E  S  L  K  R  *  F  P  L  N  Q  R  P  W  S  Q  R  F  K  * -

                                                                                        BsiHKAI
                                                                                        Bsp1286I
                                              HinfI                                     FalI |
                                              PsrI|                              NlaIII  | |
    BceAI                                     MlyI||                         FatI  |  |  |
    AvaII|                  HpyCH4III         PleI|||                    HincII |  |  |  |
  EcoO109I|                 BsrI   |         BsmAI||||                    Hpy8I |  |  |  |
    PpuMI|             TspDTI |    |          MlyI |||||               MboII | |  |  |  |
    Sau96I|           PsrI |  |    |    HinfI PleI |||||    MboII  XmnI  | | |  |  |  |
aggacctaacagaactcgccgtaaagactggcgaacagttcatacagagtctcttacgactcaatgacaagaagaaaatcttcgtcaacatggtggagca
5301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5400
tcctggattgtcttgagcggcatttctgaccgcttgtcaagtatgtctcagagaatgctgagttactgttcttcttttagaagcagttgtaccacctcgt

a    G  P  N  R  T  R  R  K  D  W  R  T  V  H  T  E  S  L  T  T  Q  *  Q  E  E  N  L  R  Q  H  G  G  A -
b     D  L  T  E  L  A  V  K  T  G  E  Q  F  I  Q  S  L  L  R  L  N  D  K  K  K  I  F  V  N  M  V  E  H -
c   R  T  *  Q  N  S  P  *  R  L  A  N  S  S  Y  R  V  S  Y  D  S  M  T  R  R  K  S  S  S  T  W  W  S  T -

                                                  MfeI
                                                Tsp509I
                                                 BsmAI|
                                                 MboII|
                                    BsmAI        Hin4I||                  HpaII
      Hpy8I                        Hpy188I      BseMII|||              Hpy188III
      AccI|                        DdeI |       BspCNI||||               BsaWI|
   DrdI ||    Hin4I    HpyCH4III    |  |        BbsI |||||               BseRI|
    |   ||      |         |    |    |  |          |  |||||               BspEI|    BsaJI
cgacacacttgtctactccaaaaatatcaaagatacagtctcagaagaccaaagggcaattgagacttttcaacaaagggtaatatccggaaacctcctc
5401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5500
gctgtgtgaacagatgaggtttttatagtttctatgtcagagtcttctggtttcccgttaactctgaaaagttgtttcccattataggcctttggaggag

a    R  H  T  C  L  L  Q  K  Y  Q  R  Y  S  L  R  R  P  K  G  N  *  D  F  S  T  K  G  N  I  R  K  P  P  R -
b     D  T  L  V  Y  S  K  N  I  K  D  T  V  S  E  D  Q  R  A  I  E  T  F  Q  Q  R  V  I  S  G  N  L  L -
c      T  H  L  S  T  P  K  I  S  K  I  Q  S  Q  K  T  K  G  Q  L  R  L  F  N  K  G  *  Y  P  E  T  S  S -

        MnlI
       BsrDI|
       MnlI ||        MaeIII                                              BccI
    HinfI | ||   AluI    |                                           HpyF10VI |            SfaNI
```

```
                                                                          NlaIV                 MwoI| |            CviJI |
        TfiI | ||       CviJI       |                                                                                   
   Hpy188I|  | ||BseYI      |Tsp45I                          MboII        CviJI|            BsrDI || |            HaeIII |
          || | ||     |      |      |                            |             ||                 | || |                | |
          ggattccattgcccagctatctgtcactttattgtgaagatagtggaaaaggaaggtggctcctacaaatgccatcattgcgataaaggaaaggccat
cg
     5501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 5600
          cctaaggtaacgggtcgatagacagtgaaataacacttctatcaccttttccttccaccgaggatgtttacggtagtaacgctatttcctttccggta
gc

a          I  P  L  P  S  Y  L  S  L  Y  C  E  D  S  G  K  G  R  W  L  L  Q  M  P  S  L  R  *  R  K  G  H
R   -
b         G  F  H  C  P  A  I  C  H  F  I  V  K  I  V  E  K  E  G  G  S  Y  K  C  H  H  C  D  K  G  K  A  I
V -
c          D  S  I  A  Q  L  S  V  T  L  L  *  R  *  W  K  R  K  V  A  P  T  N  A  I  I  A  I  K  E  R  P  S
-
```

Figure 19-U

```
                                                                                       TaiI
                                                                                  BmgBI |
                        BccI                                                 HpyCH4IV| |
                      NlaIV|                                                  MboII || |
                      TspRI|                                                 BbsI | || |
                    AvaII ||                                                MboII | || |
                   Sau96I ||                                               BbsI | | || |
                   PshAI | ||         NlaIV                               TaiI | | | || |
            HpyCH4III| | ||        AvaII |                     BseRI      |    | | | || |
          MboII     || | ||       Sau96I |    BssSI      SfaNI|HpyCH4IV   |    | | | || |
  BccI BsmFI    | MnlI|| | ||     BslI | |  MnlI |     TaqII  ||  XmnI    |    | | | || |      Hin4I
   |     |      |    ||| | ||       |  | |   |   |        |   ||    |     |    | | | || |        |
   ttgaagatgcctctgccgacagtggtcccaaagatggacccccacccacgaggagcatcgtggaaaaagaagacgttccaaccacgtcttcaaagcaa
gt
5601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 5700
   aacttctacggagacggctgtcaccagggtttctacctgggggtgggtgctcctcgtagcaccttttttcttctgcaaggttggtgcagaagtttcgtt
ca

a     *  R  C  L  C  R  Q  W  S  Q  R  W  T  P  T  H  E  E  H  R  G  K  R  R  R  S  N  H  V  F  K  A  S
-
b    E  D  A  S  A  D  S  G  P  K  D  G  P  P  P  T  R  S  I  V  E  K  E  D  V  P  T  T  S  S  K  Q
V  -
c   L  K  M  P  L  P  T  V  V  P  K  M  D  P  H  P  R  G  A  S  W  K  K  K  T  F  Q  P  R  L  Q  S  K
W -


                           Hin4I
                          TaiI |
                        Hin4I| |
                       TspRI|| |                                              TspDTI
                                          FokI        Hin4I                    MnlI|
  MmeI       EcoRV    HpyCH4IV||| |  BstF5I  HgaI   Hin4I  |              TspDTI  ||Hin4I
   |           |             |||| |     |       |      |    |                |    ||   |
   ggattgatgtgatatctccactgacgtaagggatgacgcacaatcccactatccttcgcaagacccttcctctatataaggaagttcatttcatttgg
ag
5701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 5800
   cctaactacactatagaggtgactgcattccctactgcgtgttagggtgataggaagcgttctgggaaggagatatattccttcaagtaaagtaaacc
tc

a   G  L  M  *  Y  L  H  *  R  K  G  *  R  T  I  P  L  S  F  A  R  P  F  L  Y  I  R  K  F  I  S  F  G
E -
b    D  *  C  D  I  S  T  D  V  R  D  D  A  Q  S  H  Y  P  S  Q  D  P  S  S  I  *  G  S  S  F  H  L  E
-
c     I  D  V  I  S  P  L  T  *  G  M  T  H  N  P  T  I  L  R  K  T  L  P  L  Y  K  E  V  H  F  I  W
R  -


                           AlwI
                         NlaIII                                                  Bme1580I
                       FatI  |                                                    BsiHKAI
                     BsmFI|  |                                                   Bsp1286I
             AlwI  BstKTI||  |                                                      Hpy8I |
             BfaI   DpnI||   |                                                    HpyCH4V |
        Hpy188III  NlaIV||   |                                                   ApaLI | |
             XbaI|BamHI|||   |                                                  HpaII || |
  MlyI HinfI    ||BstYI|||   |                                                   AgeI| || |
  PleI  MnlI    || MboI|||   |                           BfaI   MaeIII         BsaWI|| || |
    |     |     ||     |||   |                           SpeI|   Tsp45I        BsrFI|| || |
                                                           ||      |              || || |
   agaacacgggggactctagaggatccatgtttatttcttattatttcttactctcactagtggtagtgaccttgaccggtgcaccacttttgatgat
gt
5801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 5900
   tcttgtgccccctgagatctcctaggtacaaataaaagaataataaagaatgagagtgatcaccatcactggaactggccacgtggtgaaaactacta
ca

a    N  T  G  D  S  R  G  S  M  F  I  F  L  L  F  L  T  L  T  S  G  S  D  L  D  R  C  T  T  F  D  D
V  -
b   R  T  R  G  T  L  E  D  P  C  L  F  S  Y  Y  F  L  L  S  L  V  V  V  T  L  T  G  A  P  L  L  M  M
F -
c    E  H  G  G  L  *  R  I  H  V  Y  F  L  I  I  S  Y  S  H  *  W  *  *  P  *  P  V  H  H  F  *  *  C
-


                                                                   Hpy188I
                                                                  TspDTI|                   MboII
     Tsp509I                                                     BstKTI||          Hpy188III  |
   AluI  |                                   Hpy188III  ApoI      DpnI|||            DdeI  |  |
  CviJI  |   TspDTI        MnlI            Hpy8I     |Tsp509I    MboI ||||         MseI |  |  |
     |   |     |             |                |      |     |        | ||||            |  |  |  |
   tcaagctcctaattacactcaacatacttcatctatgaggggggtttactatcctgatgaaattttagatcagacactctttatttaactcaggatt
ta
5901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 6000
   agttcgaggattaatgtgagttgtatgaagtagatactccccccaaatgataggactactttaaaatctagtctgtgagaaataaattgagtcctaa
at
```

a Q A P N Y T Q H T S S M R G V Y Y P D E I F R S D T L Y L T Q D L
-
b K L L I T L N I L H L * G G F T I L M K F L D Q T L F I * L R I
Y -
c S S S * L H S T Y F I Y E G G L L S * * N F * I R H S L F N S G F
I -

Figure 19-V

```
                                                    TaiI
                                              HpyCH4IV |
      BseMII               TspDTI         AseI      |  |                      MseI     BstF5I         Fnu4HI
     BspCNI|              MaeIII|         MseI      |  |                      BccI|    BbvI  |    FokI  TseI|
          ||                    ||           |      |  |                         ||       |  |       |  ||
       tttcttccatttattctaatgttacagggtttcatactattaatcatacgtttggcaaccctgtcatacctttaaggatggtatttattttgctgc
ca
  6001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 6100
       aaagaaggtaaaataagattacaatgtcccaaagtatgataattagtatgcaaaccgttgggacagtatggaaattcctaccataaataaaacgacg
gt

a       F  L  P  F  Y  S  N  V  T  G  F  H  T  I  N  H  T  F  G  N  P  V  I  P  F  K  D  G  I  Y  F  A  A
T -
b        F  F  H  F  I  L  M  L  Q  G  F  I  L  L  I  I  R  L  A  T  L  S  Y  L  L  R  M  V  F  I  L  L  P
-
c         S  S  I  L  F  *  C  Y  R  V  S  Y  Y  *  S  Y  V  W  Q  P  C  H  T  F  *  G  W  Y  L  F  C  C
H  -

                                                        HpyCH4III
                                                          TspDTI|
                         BsaJI                  NlaIII  MaeIII ||       Tsp509I
            TspGWI       BtgI  BslI        FatI  |  Tsp45I     ||        HphI |
                 |          |     |           |  |       |     ||       MseI| |
       cagagaaatcaaatgttgtccgtggttgggtttttggttctaccatgaacaacaagtcacagtcggtgattattattaacaattctactaatgttgtt
at
  6101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 6200
       gtctctttagtttacaacaggcaccaacccaaaaaccaagatggtacttgttgttcagtgtcagccactaataataattgttaagatgattacaacaa
ta

a         E  K  S  N  V  V  R  G  W  V  F  G  S  T  M  N  N  K  S  Q  S  V  I  I  I  N  N  S  T  N  V  V
I  -
b       Q  R  N  Q  M  L  S  V  V  G  F  L  V  L  P  *  T  T  S  H  S  R  *  L  L  L  T  I  L  L  M  L  L
Y -
c        R  E  I  K  C  C  P  W  L  G  F  W  F  Y  H  E  Q  Q  V  T  V  G  D  Y  Y  *  Q  F  Y  *  C  C  Y
-

                                                            Hpy8I
                                                            RsaI|
                                                          Csp6I||
                                                       NlaIII |||
                                                     BsaJI  | |||                                Tsp509I
          NlaIII                                     BtgI   | |||                                  MseI|
            NspI          MaeIII                     FatI   | |||                                  NsiI ||
         MaeIII|          Tsp45I                     NcoI   | |||                           BfrBI |  ||
          FatI ||    Tsp509I   |                     StyI   | |||                   TaqI HpyCH4V |  ||
             | ||          |   |                        |   | |||                      |      |  |  ||
       acgagcatgtaactttgaattgtgtgacaacccttctttgctgtttctaaacccatgggtacacagacacatactatgatattcgataatgcattta
at
  6201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 6300
       tgctcgtacattgaaacttaacacactgttgggaaagaaacgacaaagatttgggtacccatgtgtctgtgtatgatactataagctattacgtaaat
ta

a        R  A  C  N  F  E  L  C  D  N  P  F  F  A  V  S  K  P  M  G  T  Q  T  H  T  M  I  F  D  N  A  F  N
-
b         E  H  V  T  L  N  C  V  T  T  L  S  L  L  F  L  N  P  W  V  H  R  H  I  L  *  Y  S  I  M  H  L
I  -
c       T  S  M  *  L  *  I  V  *  Q  P  F  L  C  C  F  *  T  H  G  Y  T  D  T  Y  Y  D  I  R  *  C  I  *
L -

                RsaI
              Csp6I|
              SfaNI|
  HpyCH4V     TatI||                                                DraI                         DraI
    Hin4I|  TaqI  |||  Hpy188I        Hin4I      Hpy188I     Tsp509I MseI|                     MseI|BccI
        ||     |  |||        |            |            |           |   ||                         ||   |
       tgcactttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaattttaaacacttacgagagtttgtgtttaaaaataaaga
tg
  6301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 6400
       acgtgaaagctcatgtatagactacggaaaagcgaactacaaagtcttttcagtccattaaaatttgtgaatgctctcaaacacaaatttttatttct
ac

a       C  T  F  E  Y  I  S  D  A  F  S  L  D  V  S  E  K  S  G  N  F  K  H  L  R  E  F  V  F  K  N  K  D
G -
b        A  L  S  S  T  Y  L  M  P  F  R  L  M  F  Q  K  S  Q  V  I  L  N  T  Y  E  S  L  C  L  K  I  K  M
-
c         H  F  R  V  H  I  *  C  L  F  A  *  C  F  R  K  V  R  *  F  *  T  L  T  R  V  C  V  *  K  *  R
W  -

                                                BstKTI
                                                  DpnI|
                                                MboI ||
```

```
                    PsiI CviJI      SfcI      Hpy188III | ||         MseI                 MseI
                       |     |         |              |  | ||          |                    |
        ggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaaacctatttttaagttgcctctt
gg
   6401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 6500
        ccaaagagatacaaatattcccgatagttggatatctacatcaagcactagatggaagaccaaaattgtgaaactttggataaaaattcaacggagaa
cc

a         F  L  Y  V  Y  K  G  Y  Q  P  I  D  V  V  R  D  L  P  S  G  F  N  T  L  K  P  I  F  K  L  P  L
G  -
b       G  F  S  M  F  I  R  A  I  N  L  *  M  *  F  V  I  Y  L  L  V  L  T  L  *  N  L  F  L  S  C  L  L
V -
c        V  S  L  C  L' *  G  L  S  T  Y  R  C  S  S  *  S  T  F  W  F  *  H  F  E  T  Y  F  *  V  A  S  W
-
```

Figure 19-W

```
                                                                          CviJI
                                                                          PstI|
                                                                        Fnu4HI||
                                                                       HpyCH4V|||
                                                                          TseI|||
                                                                     Fnu4HI ||||
                                                                       SfcI ||||
                                                                       AluI| ||||
                                                                      CviJI| ||||
                                                                     MspA1I| ||||
                                                                      PvuII| ||||
                                                                       TseI| ||||
                                                                   TaiI  || ||||
                                                                 BmgBI | || ||||
                                                  AarI     Bme1580I| | || ||||
                                      CviJI       BspMI    Bsp1286I| | || ||||
CviJI
       MnlI      ApoI                 BpuEI |     Hpy188III |   HpyCH4IV| | || ||||        CviJI   DraI
 |
       MseI    Tsp509I       CviJI     HphI |     SmlI | |   BbvI      || | || ||||        BbvI  | MseI|
 |
         |        |        |          | |        | | |      |         || | || ||||          |     |   ||
 |
        tattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacgtcagctgcagcctatttgttggctatttaa
ag
   6501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 6600
        ataattgtaatgtttaaaatctcggtaagaatgtcggaaaagtggacgagttctgtaaaccccgtgcagtcgacgtcggataaaacaaccgataaatt
tc

a        I  N  I  T  N  F  R  A  I  L  T  A  F  S  P  A  Q  D  I  W  G  T  S  A  A  A  Y  F  V  G  Y  L  K
-
b         L  T  L  Q  I  L  E  P  F  L  Q  P  F  H  L  L  K  T  F  G  A  R  Q  L  Q  P  I  L  L  A  I  *
S  -
c        Y  *  H  Y  K  F  *  S  H  S  Y  S  L  F  T  C  S  R  H  L  G  H  V  S  C  S  L  F  C  W  L  F  K
A -


                                           TspDTI
                                        RsaI   |
                                       SfaNI   |
      BpuEI           SmlI            Csp6I|   |    AlwNI                                           MseI
        |               |                 ||   |      |                                               |
        ccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgttgattgttctcaaaatccacttgctgaactcaaatgctctgttaa
ga
   6601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 6700
        ggttgatgtaaatacgagttcatactacttttaccatgttagtgtctacgacaactaacaagagttttaggtgaacgacttgagtttacgagacaatt
ct

a       P  T  T  F  M  L  K  Y  D  E  N  G  T  I  T  D  A  V  D  C  S  Q  N  P  L  A  E  L  K  C  S  V  K
S -
b        Q  L  H  L  C  S  S  M  M  K  M  V  Q  S  Q  M  L  L  I  V  L  K  I  H  L  L  N  S  N  A  L  L  R
-
c         N  Y  I  Y  A  Q  V  *  *  K  W  Y  N  H  R  C  C  *  L  F  S  K  S  T  C  *  T  Q  M  L  C  *
E  -


                                                                    HinfI
                                                  Hpy188III          TfiI
                                                  Bsu36I |     BseMII   |
   AluI          ApoI                               DdeI | |    BspCNI|   |        SspI
   CviJI        Tsp509I          Tsp509I  MnlI    BsaXI | |   MnlI   ||   |      BsaXI|                MnlI
     |             |               |        |        |   | |     |   ||   |         ||                   |
        gctttgagattgacaaaggaatttaccagacctctaatttcagggttgttccctcaggagatgttgtgagattccctaatattacaaacttgtgtcct
tt
   6701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 6800
        cgaaactctaactgtttccttaaatggtctggagattaaagtcccaacaagggagtcctctacaacactctaagggattataatgtttgaacacagga
aa

a         F  E  I  D  K  G  I  Y  Q  T  S  N  F  R  V  V  P  S  G  D  V  V  R  F  P  N  I  T  N  L  C  P
F  -
b        A  L  R  L  T  K  E  F  T  R  P  L  I  S  G  L  F  P  Q  E  M  L  *  D  S  L  I  L  Q  T  C  V  L
L -
c         L  *  D  *  Q  R  N  L  P  D  L  *  F  Q  G  C  S  L  R  R  C  C  E  I  P  *  Y  Y  K  L  V  S  F
-


                                          NlaIII
                                           NsiI |
                                        BfrBI | |          Tsp509I
                   ApoI                  FatI | |     ApoI     |                    BsiHKAI
          MseI   Tsp509I             HpyCH4V  | |   Tsp509I    |                    Bsp1286I
            |       |                     | | |        |       |                        |
```

```
        tggagaggttttaatgctactaaattcccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgctgattactctgtgctctacaact
ca
   6801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 6900
        acctctccaaaaattacgatgatttaagggaagacagatacgtaccctctcttttttttaaagattaacacaacgactaatgagacacgagatgttga
gt

a        G  E  V  F  N  A  T  K  F  P  S  V  Y  A  W  E  R  K  K  I  S  N  C  V  A  D  Y  S  V  L  Y  N  S
-
b         E  R  F  L  M  L  L  N  S  L  L  S  M  H  G  R  E  K  K  F  L  I  V  L  L  I  T  L  C  S  T  T
Q  -
c       W  R  G  F  *  C  Y  *  I  P  F  C  L  C  M  G  E  K  K  N  F  *  L  C  C  *  L  L  C  A  L  Q  L
N -
```

Figure 19-X

```
                   MseI                 DdeI      BstKTI                     HinfI
                                                  DpnI|                      TfiI
                                                  MboI ||              HpyCH4V  |       BsaXI
acatttttttcaacctttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatgcagattcttttgtagtcaaggg ag
6901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------- 7000
tgtaaaaaaagttggaaattcacgataccgcaaagacggtgattcaacttactagaaacgaagaggttacagatacgtctaagaaaacatcagttccc tc

a   T  F  F  S  T  F  K  C  Y  G  V  S  A  T  K  L  N  D  L  C  F  S  N  V  Y  A  D  S  F  V  V  K  G  D -
b    H  F  F  Q  P  L  S  A  M  A  F  L  P  L  S  *  M  I  F  A  S  P  M  S  M  Q  I  L  L  *  S  R  E  -
c     I  F  F  N  L  *  V  L  W  R  F  C  H  *  V  E  *  S  L  L  L  Q  C  L  C  R  F  F  C  S  Q  G  R -

                 BstNI
                 ScrFI
                 HaeII
                 HhaI
                 PspGI
                 StyD4I                              TspDTI
                 BsaXI                     Tsp509I
                 HinP1I                  Tsp509I                    NlaIII
BfaI                        BsrI          PsiI  PsiI             FatI
atgatgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcatgggttgtgtccttgcttggaat ac
7001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------- 7100
tactacattctgtttatcgcggtcctgtttgaccacaataacgactaatattaatatttaacggtctactaaagtacccaacacaggaacgaacctta tg

a    D  V  R  Q  I  A  P  G  Q  T  G  V  I  A  D  Y  N  Y  K  L  P  D  D  F  M  G  C  V  L  A  W  N  T -
b   M  M  *  D  K  *  R  Q  D  K  L  V  L  L  L  I  I  I  I  N  C  Q  M  I  S  W  V  V  S  L  L  G  I  L -
c    *  C  K  T  N  S  A  R  T  N  W  C  Y  C  *  L  *  L  *  I  A  R  *  F  H  G  L  C  P  C  L  E  Y  -

                                                                     CviJI
                                                                     HaeIII
                                                                   EcoO109I
                                                                     Sau96I
                                                                Bpu10I
                                                                  DdeI
                                                                  AluI
                                                                  CviJI
                                                                Cac8I
                                                              HindIII
                           Tsp509I                          NlaIII
                        BsrI                                              BsmAI
SfaNI                Tsp509I PsiI  PsiI              DdeI FatI
taggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggccctttgagagagacatatctaatg tg
7101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------- 7200
atccttgtaactacgatgaagttgaccattaatattaatatttatatccatagaatctgtaccgttcgaatccgggaaactctctctgtatagattac ac

a   R  N  I  D  A  T  S  T  G  N  Y  N  Y  K  Y  R  Y  L  R  H  G  K  L  R  P  F  E  R  D  I  S  N  V  -
b    G  T  L  M  L  L  Q  L  V  I  I  I  I  N  I  G  I  L  D  M  A  S  L  G  P  L  R  E  T  Y  L  M  C -
c   *  E  H  *  C  Y  F  N  W  *  L  *  L  *  I  *  V  S  *  T  W  Q  A  *  A  L  *  E  R  H  I  *  C  A -

                  HpyCH4V        AarI      CviJI
                HpyF10VI         BspMI     HaeIII
                 BstAPI         Tsp509I    MscI                                    CviJI
       BccI       MwoI           MseI      EaeI  MseI                          BsrI
cctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccattaaatgattatggttttacaccactactggcattggcta cc
7201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------- 7300
ggaaagaggggactaccgtttggaacgtggggtggacgagaattaacaataaccggtaatttactaataccaaaaatgtggtgatgaccgtaaccgat gg
```

a P F S P D G K P C T P P A L N C Y W P L N D Y G F Y T T T G I G Y
Q -
b L S P L M A N L A P H L L L I V I G H * M I M V F T P L L A L A T
-
c F L P * W Q T L H P T C S * L L L A I K * L W F L H H Y W H W L
P -

Figure 19-Y

```
                                                  HpyCH4III
                                                   BsaJI |
                                                     BtgI |
                                                  CviJI| |
                                                 HaeIII| |
                                                  EaeI || |
                                                HpaII| || |
                          RsaI               BsrFI|| || |      Tsp509I
                          ScaI           HpyCH4V ||| || | AvaII    |
                        Csp6I|             DraI  | ||| || | Hpy8I    |
                        TatI||             MseI| | ||| || | Sau96I   |      TspRI MseI   BsrI
TspRI
                          |||               ||  | |||| || | .  |     |        |    |       |
|
          aaccttacagagttgtagtactttcttttgaactttaaatgcaccggccacggtttgtggaccaaaattatccactgaccttattaagaaccagtgt
gt
    7301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 7400
          ttggaatgtctcaacatcatgaaagaaaacttgaaaatttacgtggccggtgccaaacacctggttttaataggtgactggaataattcttggtcaca
ca

a           P  Y  R  V  V  V  L  S  F  E  L  L  N  A  P  A  T  V  C  G  P  K  L  S  T  D  L  I  K  N  Q  C
V  -
b         N  L  T  E  L  *  Y  F  L  L  N  F  *  M  H  R  P  R  F  V  D  Q  N  Y  P  L  T  L  L  R  T  S  V
S -
c          T  L  Q  S  C  S  T  F  F  *  T  F  K  C  T  G  H  G  L  W  T  K  I  I  H  *  P  Y  *  E  P  V  C
-

                                        HincII
                                          HpaI
                                 BsrI    Hpy8I
                  MlyI           RsaI    MboII                                      CviJI
         Tsp509I MseI  BaeI      TspRI   MseI|                                      HaeIII          TspGWI
   Tsp509I MseI| PleI HinfI      Csp6I|  BsrI ||            BaeI          BceAI Tsp509I EaeI |       Hpy188I |
         |     ||    |    |          ||     |  ||              |              |       |    | |           |  |
          caatttaattttaatggactcactggtactggtgtgttaactcctcttcaaagagatttcaaccatttcaacaatttggccgtgatgtttctgatt
tc
    7401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 7500
          gttaaaattaaaattacctgagtgaccatgaccacacaattgaggaagaagtttctctaaagttggtaaagttgttaaaccggcactacaaagactaa
ag

a          N  F  N  F  N  G  L  T  G  T  G  V  L  T  P  S  S  K  R  F  Q  P  F  Q  Q  F  G  R  D  V  S  D  F
-
b          I  L  I  L  M  D  S  L  V  L  V  C  *  L  L  L  Q  R  D  F  N  H  F  N  N  L  A  V  M  F  L  I
S  -
c          Q  F  *  F  *  W  T  H  W  Y  W  C  V  N  S  F  F  K  E  I  S  T  I  S  T  I  W  P  *  C  F  *  F
H -

                  BstKTI
                   DpnI|
                 BstYI ||
                  MboI ||
           Hpy188III | ||
          TspRI      | | ||                                                             TspDTI
                                                                                      BstNI  |
                                                                                      ScrFI  |
        HinfI |  TaqI| | ||                                                           PspGI | |
         TfiI |AlwI || | ||        Hpy188I SspI   HphI                        Tsp509I StyD4I | |
            |  |   | |:| | ||            |   |   |                                |     |  |  |
          actgattccgttcgagatcctaaaacatctgaaatattagacatttcaccttgctcttttgggggtgtaagtgtaattacacctggaacaaatgcttc
at
    7501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 7600
          tgactaaggcaagctctaggattttgtagactttataatctgtaaagtggaacgagaaaacccccacattcacattaatgtggaccttgtttacgaag
ta

a          T  D  S  V  R  D  P  K  T  S  E  I  L  D  I  S  P  C  S  F  G  G  V  S  V  I  T  P  G  T  N  A  S
S -
b          L  I  P  F  E  I  L  K  H  L  K  Y  *  T  F  H  L  A  L  L  G  V  *  V  *  L  H  L  E  Q  M  L  H
-
c           *  F  R  S  R  S  *  N  I  *  N  I  R  H  F  T  L  L  F  W  G  C  K  C  N  Y  T  W  N  K  C  F
I  -

                                                                BstKTI
                                                                 DpnI|
                                                             MboI ||
                                                         HpyCH4V | ||
                                                          NlaIII | ||
                                                      HpyF10VI | | ||
                                      HpyCH4V             BstAPI| | | ||
                     Eco57I HincII      |          Tsp509I     || | | ||                 HhaI
                    Eco57MI  HpaI       |       SfcI     |     || | | ||              HinP1I |
                  Hpy188III| Hpy8I      |    TspDTI|     | MwoI| | | ||              AluI  | |
   Hpy188I            BsgI || MseI|     | TspRI  ||     |FatI|| | | ||             CviJI  | |
```

```
           |              |  ||    ||   |      |  ||    |  ||| | | ||           |  | |
           ctgaagttgctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacaccagcttggcgcatatattctact
gg
     7601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 7700
           gacttcaacgacaagatatagttctacaattgacgtgactacaaagatgtcgttaagtacgtctagttgagtgtggtcgaaccgcgtatataagatga
cc

a           E  V  A  V  L  Y  Q  D  V  N  C  T  D  V  S  T  A  I  H  A  D  Q  L  T  P  A  W  R  I  Y  S  T
G  -
b          L  K  L  L  F  Y  I  K  M  L  T  A  L  M  F  L  Q  Q  F  M  Q  I  N  S  H  Q  L  G  A  Y  I  L  L
E -
c           *  S  C  C  S  I  S  R  C  *  L  H  *  C  F  Y  S  N  S  C  R  S  T  H  T  S  L  A  H  I  F  Y  W
-
```

Figure 19-Z

```
                                                   HincII
                                                   Hpy8I
                                                   AccI|
                       SmlI                        TaqI|
                   HinfI |               BlpI  NlaIII||
        Hpy188III   |    |               DdeI   NspI||                          Cac8I
  BpuEI   MlyI    | |    |    CviJI BspCNI AluI|    SalI||                          AluI |
 BsrI |   PleI    | |    |   Cac8I |BseMII CviJI|FatI |||                          CviJI |
 aaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacatcctattggagctggcatttgtgct
7701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7800
 tttgttacataaggtctgagttcgtccgacagaatatcctcgactcgtacagctgtgaagaatactcacgctgtaaggataacctcgaccgtaaacacga

a    N  N  V  F  Q  T  Q  A  G  C  L  I  G  A  E  H  V  D  T  S  Y  E  C  D  I  P  I  G  A  G  I  C  A  -
b     T  M  Y  S  R  L  K  Q  A  V  L  *  E  L  S  M  S  T  L  L  M  S  A  T  F  L  L  E  L  A  F  V  L  -
c   K  Q  C  I  P  D  S  S  R  L  S  Y  R  S  *  A  C  R  H  F  L  *  V  R  H  S  Y  W  S  W  H  L  C  Y  -
```

```
ApoI, Tsp509I, TaqI, AluI, CviJI, RsaI, Csp6I, ApaI, BanII, Bme1580I, Bsp1286I, HpaII, NciI, ScrFI,
CviJI, HaeIII, NlaIV, StyD4I, NlaIV, Sau96I, Eco0109I, MnlI, PspOMI, Sau96I, BslI
MnlI, TaqI, BsrBI, TauI, AciI, BsiEI, Fnu4HI, CviJI, HaeIII, TauI, BsrI, EaeI, EagI, Fnu4HI, NotI,
TspRI, AciI, BccI, BstXI, TaqI, AvaI, SmlI, XhoI
BsaBI, EcoRV, PstI, HpyCH4V, HpyF10VI, SfcI, MwoI, ApoI, EcoRI¹, Tsp509I
HpyCH4III   BfaI
 attaccatacagTTAGACTAGAAAAGggCGAATTCTGCAGATATCCATCACACTGGCGGCCGCtcgacctcgaggggggggcccggtacagctcgaattc
7801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7900
 taatggtatgtcAATCTGATCTTTTCccGCTTAAGACGTCTATAGGTAGTGTGACCGCCGGCGagctggagctccccccccgggccatgtcgagcttaag

a    I  T  I  Q  L  D  *  K  R  A  N  S  A  D  I  H  H  T  G  G  R  S  T  S  R  G  G  P  V  Q  L  E  F  p -
b     L  P  Y  S  *  T  R  K  G  R  I  L  Q  I  S  I  T  L  A  A  A  R  P  R  G  G  A  R  Y  S  S  N  F  -
c      Y  H  T  V  R  L  E  K  G  E  F  C  R  Y  P  S  H  W  R  P  L  D  L  E  G  G  P  G  T  A  R  I  S  -
```

```
      BsiEI
      BstKTI
       PvuI                                         HpyF10VI                                    MseI
      DpnI|              MseI                  HpaII |                                         TaiI
     MboI ||            AflII|     HinfI       BslI  |                                HpyCH4IV |
!FatI                    SmlI|     TfiI      BsrFI|  MwoI|  BsaBI  Tsp509I  Tsp509I      |
 cccgatcgttcaaacatttggcaataaagtttcttaagattgaatcctgttgccggtcttgcgatgattatcatataatttctgttgaattacgttaagc
7901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8000
 gggctagcaagtttgtaaaccgttatttcaaagaattctaacttaggacaacggccagaacgctactaatagtatattaaagacaacttaatgcaattcg

a    R  S  F  K  H  L  A  I  K  F  L  K  I  E  S  C  C  R  S  C  D  D  Y  H  I  I  S  V  E  L  R  *  A -
b   P  D  R  S  N  I  W  Q  *  S  F  L  R  L  N  P  V  A  G  L  A  M  I  I  I  *  F  L  L  N  Y  V  K  H -
c    P  I  V  Q  T  F  G  N  K  V  S  *  D  *  I  L  L  P  V  L  R  *  L  S  Y  N  F  C  *  I  T  L  S  -
```

Figure 19-AA

```
                                   TaiI
                           HpyCH4IV |
                            NlaIII| |
                             NsiI|| |
                          BfrBI | || |
                           FatI | || |
                        HpyCH4V | || |
                     NlaIII   | | || |
                 AflIII  |    | | || |
                  FatI   |    | | || |                                          FauI
                  PciI   |    | | || |                                     MlyI |
               MseI |    |    | | || |                                    PleI| |
          Tsp509I | |    |    | | || |                                 Tsp509I| |
        NlaIII  | | |    |    | | || |                                      || |
         NspI   | | |NspI|    | | || |BccI      BsmFI         HinfI AciI   || |   MseI   BstUI
            |   | | |    |    | | || |   |          |             |    |   || |      |      |
         atgtaataattaacatgtaatgcatgacgttatttatgagatgggtttttatgattagagtcccgcaattatacatttaatacgcgatagaaaacaaa
at
    8001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 8100
         tacattattaattgtacattacgtactgcaataaatactctacccaaaaatactaatctcagggcgttaatatgtaaattatgcgctatcttttgttt
ta

a          C  N  N  *  H  V  M  H  D  V  I  Y  E  M  G  F  Y  D  *  S  P  A  I  I  H  L  I  R  D  R  K  Q  N
-
b           V  I  I  N  M  *  C  M  T  L  F  M  R  W  V  F  M  I  R  V  P  Q  L  Y  I  *  Y  A  I  E  N  K
I -
c          M  *  *  L  T  C  N  A  *  R  Y  L  *  D  G  F  L  *  L  E  S  R  N  Y  T  F  N  T  R  *  K  T  K
Y -
```

```
                                             AciI
                                            BstUI
                                             HhaI                         ApoI
           HhaI                             BstUI |                       EcoRI[2]
           Hin4I                            Cac8I |                     Tsp509I
          BstUI |                            HhaI |               Hpy188III  |                Hpy99I            Hpy99I
          Cac8I |                          HinP1I |                  BstKTI| |         BsrI      |       Hpy188III|
           HhaI |                         BssHII| |                  BceAI|| |        TspRI      |          MaeIII|
         HinP1I |                        BstUI  | |                  DpnI|| |        CviJI|      |          Tsp45I|
       BssHII | |              Tsp509I   HinP1I | |   MaeIII        MboI ||| |       HaeIII|      |            TaiI||
       HinP1I | | BfaI               |   |    | |  Hin4I   |   BfaI  | ||| |        EaeI ||      |HpyCH4IV   |||      BsrI
            | | |   |                |   |    | |     |    |     |  | ||| |           | ||      |     |     |||         |
         atagcgcgcaaactaggataaattatcgcgcgcggtgtcatctatgttactagatcgggaattcactggccgtcgttttacaacgtcgtgactgggaa
aa
    8101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 8200
         tatcgcgcgtttgatcctatttaatagcgcgcgccacagtagatacaatgatctagcccttaagtgaccggcagcaaaatgttgcagcactgaccctt
tt

a         I  A  R  K  L  G  *  I  I  A  R  G  V  I  Y  V  T  R  S  G  I  H  W  P  S  F  Y  N  V  V  T  G  K
T -
b          *  R  A  N  *  D  K  L  S  R  A  V  S  S  M  L  L  D  R  E  F  T  G  R  R  F  T  T  S  *  L  G  K
-
c           S  A  Q  T  R  I  N  Y  R  A  R  C  H  L  C  Y  *  I  G  N  S  L  A  V  V  L  Q  R  R  D  W  E
N  -
```

```
                                                                                                BsiEI
                                                                                               BstKTI
                                                                                             HpyF10VI
                                                                                                 PvuI
                                                                                                DpnI|
                                                                                                MwoI|
                                                                                               FauI||
                                                                                               MboI|||
        MaeIII                                                                              MboII ||||
        BstNI |                                                  Cac8I                     AciI | ||||
        ScrFI |                                 BstF5I            AluI |                   Cac8I | ||||
         BmrI | |                          Fnu4HI  |             CviJI |                  CviJI| | ||||
        BsaJI | |                     FokI HpyCH4V|  |          MspA1I |                 HaeIII| | ||||
        PspGI | |                                                PvuII |      MnlI              
HinP1I
    StyD4I |  |          MseI  |      TseI|   |    BbvI     Cac8I || |     EarI|    Sau96I| |   | ||||            HpyCH4III
|
            | | |          |  |         ||    |       |          | | |        ||        || |   | ||||                    |
|
         ccctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagt
tg
    8201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 8300
         gggaccgcaatgggttgaattagcggaacgtcgtgtagggggaaagcggtcgaccgcattatcgcttctccgggcgtggctagcgggaagggttgtca
ac

a          L  A  L  P  N  L  I  A  L  Q  H  I  P  L  S  P  A  G  V  I  A  K  R  P  A  P  I  A  L  P  N  S
C -
b         P  W  R  Y  P  T  *  S  P  C  S  T  S  P  F  R  Q  L  A  *  *  R  R  G  P  H  R  S  P  F  P  T  V
A -
```

c P G V T Q L N R L A A H P P F A S W R N S E E A R T D R P S Q Q L

Figure 19-BB

```
                              HpyF10VI
                          BsrBI  |
                           AciI |  |
                          Cac8I |  |
                          BbeI| |  |
                         HaeII| |  |
                         HhaI|| |  |
                       NlaIV||| |  |
                        SfoI||| |  |
                       BbvI|||| |  |
                      BsaHI|||| |  |
                     HinP1I|||| |  |
                       NarI|||| |  |
                      BanI||||| |  |                           CviJI
                      KasI||||| |  |                         Cac8I |
                 HpyF10VI|||||| |  |                          NaeI |
              CviJI     ||||||| |  |                         HpaII||
         Fnu4HI |       ||||||| |  |                        BsrFI|||
          HhaI| |       ||||||| |  |                       NgoMIV||||
          TseI| |  BglI ||||||| |  MwoI|               HpyF10VI ||||              AluI             BanII
          FspI|| | MwoI ||||||| | FauI||               MwoI|  ||||               CviJI        Bsp1286I
                                  MboII||              TaiI|  ||||                              NlaIV|
                                                   HpyCH4IV  ||  ||||                            CviJI||
          cgcagcctgaatggcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctccctta
     8301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8400
          gcgtcggacttaccgcgggcgaggaaagcgaaagaagggaaggaaagagcggtgcaagcggccgaaaggggcagttcgagatttagcccccgagggaaat

a          A  A  *  M  A  P  A  P  F  A  F  F  P  S  F  L  A  T  F  A  G  F  P  R  Q  A  L  N  R  G  L  P  L  -
b           Q  P  E  W  R  P  L  L  S  L  S  S  L  P  F  S  P  R  S  P  A  F  P  V  K  L  *  I  G  G  S  L  *  -
c          R  S  L  N  G  A  R  S  F  R  F  L  P  F  L  S  R  H  V  R  R  L  S  P  S  S  S  K  S  G  A  P  F  R  -

                                                                                   CviJI
                                                                                  HaeIII
                                                                               Sau96I |
                                                                             DraIII | |
                                                                               TaiI|  | |
                                                                             BsaAI ||  | |
                                                   TaqII                      HphI ||  | |
         Hpy188I            NlaIV          BceAI |                        HpyCH4IV| ||  | |          HpyCH4III
         NlaIV |            BanI |TaqI      MnlI  |          BccI          Hpy8I || ||  | |       BccI     MmeI|
          gggttccgatttagtgctttacggcacctcgaccccaaaaaacttgatttgggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgcc
     8401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8500
          cccaaggctaaatcacgaaatgccgtggagctggggtttttgaactaaacccactaccaagtgcatcacccggtagcgggactatctgccaaaaagcgg

a          G  F  R  F  S  A  L  R  H  L  D  P  K  K  L  D  L  G  D  G  S  R  S  G  P  S  P  *  *  T  V  F  R  P  -
b           G  S  D  L  V  L  Y  G  T  S  T  P  K  N  L  I  W  V  M  V  H  V  V  G  H  R  P  D  R  R  F  F  A  -
c          V  P  I  *  C  F  T  A  P  R  P  Q  K  T  *  F  G  *  W  F  T  *  W  A  I  A  L  I  D  G  F  S  P  -

                        MlyI
                        TaiI
                       PleI|
                 HpyCH4IV ||
                   Hpy8I | ||
              HinfI    |  | ||
               DrdI|   |  | ||                AloI
              TaiI ||   |  | ||                PpiI
       HpyCH4IV | ||   |  | ||               HinfI|
         AloI |  | ||   |  | ||              BsaXI||                               CviJI
        BsaXI |  | ||   |  | ||    MlyI        ||                           BslI   |
         PpiI |  | ||   |  | ||    PleI        ||                          AvaI|   |
                                  MseI      Hpy8I||           BsrI              BslI|   |              PsiI
          ctttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagggat
     8501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8600
          gaaactgcaacctcaggtgcaagaaattatcacctgagaacaaggtttgaccttgttgtgagttgggatagagcccgataagaaaactaaatattcccta

a           L  T  L  E  S  T  F  F  N  S  G  L  L  F  Q  T  G  T  T  L  N  P  I  S  G  Y  S  F  D  L  *  G  I  -
b          L  *  R  W  S  P  R  S  L  I  V  D  S  C  S  K  L  E  Q  H  S  T  L  S  R  A  I  L  L  I  Y  K  G  F  -
```

C F D V G V H V L * * W T L V P N W N N T Q P Y L G L F F * F I R D
-

Figure 19-CC

```
                                                                                 HpyCH4V                    BseMII
                                                                               Fnu4HI |                    BspCNI|
                                                                                 TseI||                     AciI ||
                                                                               Cac8I |||                  BstNI  | ||
                                                                            AciI  | |||                  ScrFI  |  ||
                                                                        HpyF10VI  | |||                 CviJI |  |  ||
                                                                           MwoI|  | |||                HaeIII |  |  ||
                                                                         AvaII ||  | |||                 PspGI |  |  ||
                                                                         Hpy8I ||  | |||                StyD4I |  |  ||
                        NlaIV                        BseYI              Sau96I ||  | |||              Sau96I | |  |  ||
                     Hpy188I |           BccI       Cac8I |             BbvI  | ||  | |||             DdeI   | | |  |  ||
                           | |             |            | |                |  | ||  | |||               |    | | |  |  ||
      tttgccgatttcggaaccaccatcaaacaggatttcgcctgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggccaggcggtgaagg
 8601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8700
      aaacggctaaagccttggtggtagtttgtcctaaaagcggacgaccccgtttggtcgcacctggcgaacgacgttgagagagtcccggtccgccacttcc

a       L  P  I  S  E  P  P  S  N  R  I  F  A  C  W  G  K  P  A  W  T  A  C  C  N  S  L  R  A  R  R  *  R  -
b        C  R  F  R  N  H  H  Q  T  G  F  S  P  A  G  A  N  Q  R  G  P  L  A  A  T  L  S  G  P  G  G  E  G  -
c       F  A  D  F  G  T  T  I  K  Q  D  F  R  L  L  G  Q  T  S  V  D  R  L  L  Q  L  S  Q  G  Q  A  V  K  G  -

          AluI
         CviJI
        MspA1I                  BsrI                   RsaI              AciI
         PvuII                 TspRI                  Csp6I|             TaiI|               HgaI
       HphI |             BsmAI   |         BmrI      TatI||  HpyCH4IV   ||              MseI |
            | |           BsmBI   |         HphI|      BsrI |||MseI   |   ||      BsrDI   | |     DdeI  Tsp509I
            | |               |   |            ||         ||||     |   |  ||          |   | |       |       |
      gcaatcagctgttgcccgtctcactggtgaaaagaaaaaccaccccagtacattaaaaacgtccgcaatgtgttattaagttgtctaagcgtcaatttgt
 8701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8800
      cgttagtcgacaacgggcagagtgaccacttttcttttggtggggtcatgtaatttttgcaggcgttacacaataattcaacagattcgcagttaaaca

a      A  I  S  C  C  P  S  H  W  *  K  E  K  P  P  Q  Y  I  K  N  V  R  N  V  L  L  S  C  L  S  V  N  L  F  -
b       Q  S  A  V  A  R  L  T  G  E  K  K  N  H  P  S  T  L  K  T  S  A  M  C  Y  *  V  V  *  A  S  I  C  -
c        N  Q  L  L  P  V  S  L  V  K  R  K  T  T  P  V  H  *  K  R  P  Q  C  V  I  K  L  S  K  R  Q  F  V  -

                                                                                                      NciI
                                                         AluI                                        ScrFI
                                                        CviJI                                        BccI|
                                                      Fnu4HI |                                      HpaII|
                                                        TseI||                                      BbvI||
                                CviJI               BsiEI ||                        CviJI          PfoI||
                               Cac8I |     AluI    HpaII ||      BbvI            Fnu4HI |         StyD4I||
   Hpy8I                      CviJI | |    CviJI   BsrFI| ||  HphI  |      TaqI     TseI| |         HgaI |||
       |                          | | |      |         || ||    |   |        |          || |           | |||
      ttacaccacaatatatcctgccaccagccagccaacagctccccgaccggcagctcggcacaaaatcaccactcgatacaggcagcccatcagtccggga
 8801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8900
      aatgtggtgttatataggacggtggtcggtcggttgtcgaggggctggccgtcgagccgtgttttagtggtgagctatgtccgtcgggtagtcaggccct

a       T  P  Q  Y  I  L  P  P  A  S  Q  Q  L  P  D  R  Q  L  G  T  K  S  P  L  D  T  G  S  P  S  V  R  D  -
b      L  H  H  N  I  S  C  H  Q  P  A  N  S  S  P  T  G  S  S  A  Q  N  H  H  S  I  Q  A  A  H  Q  S  G  T  -
c       Y  T  T  I  Y  P  A  T  S  Q  P  T  A  P  R  P  A  A  R  H  K  I  T  T  R  Y  R  Q  P  I  S  P  G  -

                                                                                            BceAI
                                                  NlaIII                                     NciI
                                                 MaeIII|                                    ScrFI
                                                  FatI ||                                  HpaII|
                                                 SfaNI ||                                 StyD4I||
                                        HpyF10VI   |   ||                                Fnu4HI |||
  BsaHI                               TauI    |    |   ||                                  AluI| |||
 BceAI|    BsmFI                   Fnu4HI |   |    |   ||                                 CviJI| |||
   FauI|    AciI | CviJI             AciI| |BstAPI  |   ||                  BbvI    DdeI   TseI|| |||
Tth111I|MspA1I |BceAI|      BslI     || |  MwoI  |   ||         Hpy188I  |      MboII    || |||
      ||     | |    ||         |     || |     ||   |   ||               |  |          |     || |||
```

```
         cggcgtcagcgggagagccgttgtaaggcggcagactttgctcatgttaccgatgctattcggaagaacggcaactaagctgccgggtttgaaacacg
ga
    8901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 9000
         gccgcagtcgccctctcggcaacattccgccgtctgaaacgagtacaatggctacgataagccttcttgccgttgattcgacggcccaaactttgtgc
ct

a         G  V  S  G  R  A  V  V  R  R  Q  T  L  L  M  L  P  M  L  F  G  R  T  A  T  K  L  P  G  L  K  H  G
-
b          A  S  A  G  E  P  L  *  G  G  R  L  C  S  C  Y  R  C  Y  S  E  E  R  Q  L  S  C  R  V  *  N  T
D  -
c         R  R  Q  R  E  S  R  C  K  A  A  D  F  A  H  V  T  D  A  I  R  K  N  G  N  *  A  A  G  F  E  T  R
M -
```

Figure 19-DD

```
Enzymes (9001-9100): MboI DpnI MnlI BstKTI BstF5I BstUI AciI TspGWI FokI FatI NspI NlaIII MaeIII BbvI TseI Hin4I Fnu4HI MboI BclI DpnI BstKTI AlwI Hin4I MboI BstYI MnlI DpnI BstKTI Hpy188I Tsp509I TspDTI CviJI

     tgatctcgcggagggtagcatgttgattgtaacgatgacagagcgttgctgcctgtgatcaaatatcatctccctcgcagagatccgaattatcagcc
tt
9001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 9100
     actagagcgcctcccatcgtacaactaacattgctactgtctcgcaacgacggacactagtttatagtagagggagcgtctctaggcttaatagtcgg
aa

a     * S R G G * H V D C N D D R A L L P V I K Y H L P R R D P N Y Q P
S -
b      D L A E G S M L I V T M T E R C C L * S N I I S L A E I R I I S L
-
c      I S R R V A C * L * R * Q S V A A C D Q I S S P S Q R S E L S A
F -

Enzymes (9101-9200): HinP1I Hin4I MseI Tsp45I MaeIII HpyCH4III CviJI AhdI TaqI MboI DpnI BstKTI SmlI Hpy188III Hin4I BpuEI BseMII BspCNI MnlI CviJI Fnu4HI AciI TauI MspA1I DdeI BseMII Bpu10I BbvCI BspCNI MwoI HpyF10VI CviJI AluI DdeI

     cttattcatttctcgcttaaccgtgacaggctgtcgatcttgagaactatgccgacataataggaaatcgctggataaagccgctgaggaagctgagt
gg
9101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 9200
     gaataagtaaagagcgaattggcactgtccgacagctagaactcttgatacggctgtattatcctttagcgacctatttcggcgactccttcgactca
cc

a      Y S F L A * P * Q A V D L E N Y A D I I G N R W I K P L R K L S
G -
b     L I H F S L N R D R L S I L R T M P T * * E I A G * S R * G S * V
A -
c      L F I S R L T V T G C R S * E L C R H N R K S L D K A A E E A E W
-

Enzymes (9201-9300): HhaI HaeII Hpy8I Hin4I HpyCH4IV AclI TaiI Hpy8I HincII EcoRV HphI BsrDI Hin4I Csp6I RsaI AloI Sau96I SanDI PpuMI Eco0109I AvaII NlaIV NlaIV Hpy188I BsmFI SfaNI FokI HpyCH4III BslI HaeIII CviJI AloI

     cgctatttctttagaagtgaacgttgacgatatcaactcccctatccattgctcaccgaatggtacaggtcggggacccgaagttccgactgtcggcc
tg
9201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 9300
     gcgataaagaaatcttcacttgcaactgctatagttgaggggataggtaacgagtggcttaccatgtccagcccctgggcttcaaggctgacagccgg
ac

a      A I S L E V N V D D I N S P I H C S P N G T G R G P E V P T V G L
-
b      L F L * K * T L T I S T P L S I A H R M V Q V G D P K F R L S A
* -
c     R Y F F R S E R * R Y Q L P Y P L L T E W Y R S G T R S S D C R P
D -
```

Figure 19-EE

```
BstKTI
TaqI
DpnI
MboI
CviJI
SfaNI
Hin4I
HpyF10VI
MmeI
HpaII
MwoI
NciI
ScrFI
BsaJI
StyD4I
NsiI
BfrBI
BstF5I
HpyCH4V
BslI
BslI
BstKTI
BspCNI
DpnI
BseMII
BglII
BstYI
MboI
BsrI
CviJI
Hin4I
HpyF10VI
DdeI MwoI
CviJI BmrI
HpyCH4III
SfcI
NlaIV
HpaII
NciI
ScrFI
BsaJI
StyD4I
BstKTI
DpnI
MlyI
BstYI
MboI
PleI
BstUI
NruI
Hpy188III
AlwI
HinfI
TaqI
        atgcatccccggctgatcgaccccagatctggggctgagaaagcccagtaaggaaacaactgtaggttcgagtcgcgagatcccccggaaccaaagga
ag
   9301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 9400
        tacgtaggggccgactagctggggtctagaccccgactctttcgggtcattcctttgttgacatccaagctcagcgctctagggggccttggtttcct
tc

a        M  H  P  R  L  I  D  P  R  S  G  A  E  K  A  Q  *  G  N  N  C  R  F  E  S  R  D  P  P  E  P  K  E
V -
b         C  I  P  G  *  S  T  P  D  L  G  L  R  K  P  S  K  E  T  T  V  G  S  S  R  E  I  P  R  N  Q  R  K
-
c          A  S  P  A  D  R  P  Q  I  W  G  *  E  S  P  V  R  K  Q  L  *  V  R  V  A  R  S  P  G  T  K  G
S  -

CviJI
HaeIII
BstKTI
DpnI
FauI
Hpy188I
BsrBI
MseI AciI MboI
CviJI
HaeIII
AloI
BstNI
ScrFI
PspGI
StyD4I
CviJI
BstKTI
DpnI
MboI
AciI
SfaNI
SfcI
NlaIV Hpy188III
AloI
AlwI
Hin4I
AluI
CviJI
EciI
BsrI
        taggttaaacccgctccgatcaggccgagccacgccaggccgagaacattggttcctgtaggcatcgggattggcggatcaaacactaaagctactgg
aa
   9401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 9500
        atccaatttgggcgaggctagtccggctcggtgcggtccggctcttgtaaccaaggacatccgtagccctaaccgcctagtttgtgatttcgatgacc
tt

a         G  *  T  R  S  D  Q  A  E  P  R  Q  A  E  N  I  G  S  C  R  H  R  D  W  R  I  K  H  *  S  Y  W
N  -
b        *  V  K  P  A  P  I  R  P  S  H  A  R  P  R  T  L  V  P  V  G  I  G  I  G  G  S  N  T  K  A  T  G
T -
c         R  L  N  P  L  R  S  G  R  A  T  P  G  R  E  H  W  F  L  *  A  S  G  L  A  D  Q  T  L  K  L  L  E
-

AciI
BstNI
ScrFI
PspGI
StyD4I
Hin4I
BsrI
MnlI
TauI
AciI
BsiEI
Fnu4HI
CviJI
HaeIII
EaeI
EagI
HpaII
MnlI
HphI
MnlI
FauI
AciI
BslI
HpyCH4III
NlaIV
Hpy188I
NlaIII
BsaJI
BtgI
FatI
NcoI
StyI
        cgagcagaagtcctccggccgccagttgccaggcggtaaaggtgagcagaggcacgggaggttgccacttgcgggtcagcacggttccgaacgccatg
ga
   9501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 9600
        gctcgtcttcaggaggccggcggtcaacggtccgccatttccactcgtctccgtgccctccaacggtgaacgcccagtcgtgccaaggcttgcggtac
ct
```

a - b E - c K -

E Q K S S G R Q L P G G K G E Q R H G R L P L A G Q H G S E R H G

S R S P P A A S C Q A V K V S R G T G G C H L R V S T V P N A M

R A E V L R P P V A R R * R * A E A R E V A T C G S A R F R T P W

Figure 19-FF

```
Fnu4HI MspA1I TseI AciI Cac8I HpyF10VI MwoI CviJI FauI HaeIII Sau96I BstNI ScrFI BslI
Fnu4HI HaeII HhaI TseI AfeI AlwI HinP1I BfaI BstKTI DpnI BstYI MboI
AciI Cac8I HpyF10VI
BstNI ScrFI PspGI HgaI BsaJI PspGI StyD4I BbvI AciI Hpy99I HaeII BbvI BsaHI HincII HhaI AciI StyD4I AciI FauI Hpy8I HinP1I MwoI FauI CviJI

     aaccgcccccgccaggcccgctgcgacgccgacaggatctagcgctgcgtttggtgtcaacaccaacagcgccacgcccgcagttccgcaaatagccc cc
9601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9700
     ttggcgggggcggtccgggcgacgctgcggctgtcctagatcgcgacgcaaaccacagttgtggttgtcgcggtgcgggcgtcaaggcgtttatcggg gg

a     N  R  P  R  Q  A  R  C  D  A  D  R  I  *  R  C  V  W  C  Q  H  Q  Q  R  H  A  R  S  S  A  N  S  P  Q -
b      T  A  P  A  R  P  A  A  T  P  T  G  S  S  A  A  F  G  V  N  T  N  S  A  T  P  A  V  P  Q  I  A  P -
c       P  P  P  P  G  P  L  R  R  R  Q  D  L  A  L  R  L  V  S  T  P  T  A  P  R  P  Q  F  R  K  *  P  P -

BstUI TauI AciI Fnu4HI Hpy99I HpyCH4III HaeII HhaI HinP1I HpyCH4V Fnu4HI AlwNI CviJI TauI TseI BccI Fnu4HI AciI MspA1I TspDTI
TauI Fnu4HI AciI BsrBI
HpaII NciI ScrFI StyD4I AciI
AciI AvaII Sau96I BccI CviJI BfaI BbvI BsgI

     aggaccgccatcaatcgtatcgggctacctagcagagcggcagagatgaacacgaccatcagcggctgcacagcgcctaccgtcgccgcgaccccgcc cg
9701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9800
     tcctggcggtagttagcatagcccgatggatcgtctcgccgtctctacttgtgctggtagtcgccgacgtgtcgcggatggcagcggcgctggggcgg gc

a      D  R  H  Q  S  Y  R  A  T  *  Q  S  G  R  D  E  H  D  H  Q  R  L  H  S  A  Y  R  R  R  D  P  A  R -
b     R  T  A  I  N  R  I  G  L  P  S  R  A  A  E  M  N  T  T  I  S  G  C  T  A  P  T  V  A  A  T  P  P  G -
c      G  P  P  S  I  V  S  G  Y  L  A  E  R  Q  R  *  T  R  P  S  A  A  A  Q  R  L  P  S  P  R  P  R  P -
```

Figure 19-GG

```
                                                                                    HinfI
                                                                                     TfiI
                                                                                   SfaNI|
                                                                                   MmeI||
                                                                               TspDTI |||
                                                                               MboII| |||
                                                                              FokI || |||
                                                                           BstF5I| || |||
                                                                            HhaI|| || |||
          AciI                                                   FokI      BsaBI||| || |||
          BslI|                                                 SfaNI |    FspI|||| || |||
         Cac8I||                                                BsaJI| |   FspAI||||| || |||
     HpyF10VI|||      BpmI       Hpy188III                     HhaI|| |  HinP1I|||||| || |||
        BglI||||    Eco57MI        AluI |                   HinP1I ||| |   MmeI||||||| || |||
        MwoI|||||Hpy8I    |        CviJI |              MseI      | ||| |    |||||||| || |||       HinfI
        FauI||||||AccI|   |       TaqII | |          SspI |MnlI   | ||| | BstF5I ||||| || |||XcmI    TfiI
         ||||| ||     |             | |  |              | |   |   | ||| |     |  |||| || ||| |        |
         gcaggcggtagaccgaaataaacaacaagctccagaatagcgaaatattaagtgcgccgaggatgaagatgcgcatccaccagattcccgttggaatc
tg
   9801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 9900
         cgtccgccatctggctttatttgttgttcgaggtcttatcgctttataattcacgcggctcctacttctacgcgtaggtggtctaagggcaaccttag
ac

a         Q  A  V  D  R  N  K  Q  Q  A  P  E  *  R  N  I  K  C  A  E  D  E  D  A  H  P  P  D  S  R  W  N  L
-
b          R  R  *  T  E  I  N  N  K  L  Q  N  S  E  I  L  S  A  P  R  M  K  M  R  I  H  Q  I  P  V  G  I
C  -
c         A  G  G  R  P  K  *  T  T  S  S  R  I  A  K  Y  *  V  R  R  G  *  R  C  A  S  T  R  F  P  L  E  S
V -
```

```
                                                    FauI
                                                 Fnu4HI|
                                                   TseI||
                                              Fnu4HI |||
                                                TseI| |||
                                               AciI|| |||
                                              Cac8I|| |||
                                           HpyF10VI|| |||
                                               MwoI||| |||
                                        HpyF10VI|| || |||
                                            MwoI|| || |||
                                           FauI||| || |||
                BssSI                    Cac8I|||| || |||                                        BanII
              Hpy188III                   NaeI|||| || |||                                      Bsp1286I
              BstKTI  |                  HpaII||||| || |||       BbvI                           NlaIV|
               DpnI|  |                  BsrFI||||| || |||   BbvI  |      CviJI          HpyF10VI CviJI||
               MboI||  |               NgoMIV|||||| || |||  HpaII|  |      HaeIII        MnlI|   |||          HpaII
   Hpy188I   || ||  |               AciI |||||| || |||  BsrFI||  |BsaJI   |  MnlI   MwoI||   |||     SfaNI  |
        |   | ||  |                 |  |||  ||| || |||  ||  |   |     |    |    |       ||   |||        |  |
         tcggacgatcatcacgagcaataaacccgccggcaacgcccgcagcagcataccggcgacccctcggcctcgctgttcgggctccacgaaaacgccgg
ac
   9901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 10000
         agcctgctagtagtgctcgttatttgggcggccgttgcgggcgtcgtcgtatggccgctggggagccggagcgacaagcccgaggtgcttttgcggcc
tg

a        S  D  D  H  H  E  Q  *  T  R  R  Q  R  P  Q  Q  H  T  G  D  P  S  A  S  L  F  G  L  H  E  N  A  G
Q -
b         R  T  I  I  T  S  N  K  P  A  G  N  A  R  S  S  I  P  A  T  P  R  P  R  C  S  G  S  T  K  T  P  D
-
c          G  R  S  S  R  A  I  N  P  P  A  T  P  A  A  A  Y  R  R  P  L  G  L  A  V  R  A  P  R  K  R  R
T  -
```

```
                                                                                                BsmI
                                                                                               BsaJI|
                                                                                                BtgI|
                                                                                        HpyF10VI  ||
                                                                                            MwoI| ||
                                                                                        BbeI  |  ||
                                                                                       HaeII  |  ||
                                                                                        HhaI| |  ||
                                                                                      NlaIV|| |  ||
                                                                                       SfoI||| |  ||
                              CviJI                                                   BsaHI|||| |  ||
                             HaeIII                                                  HinP1I||||| |  ||
          HpyF10VI            NlaIV|                MboII  BstKTI                      NarI|||||| |  ||
       HhaI      |          Sau96I||              CviJI|   DpnI|                       BanI||||||| |  ||
     HgaI |  MwoI| BsaJI        |||            BceAI ||  TaqII|       Hpy99I            KasI|||||||| |  ||
   HinP1I |BceAI|| StyI         |||      MnlI  BbsI | ||  MboI ||        TaqI |           ||||||||| |  ||
        |  |   ||   |           |||        |     |  | ||    |  ||          | |          ||||      ||  ||
         agatgcgccttgtgagcgtccttggggccgtcctcctgtttgaagaccgacagcccaatgatctcgccgtcgatgtaggcgccgaatgccacggcatc
tc
  10001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 10100
         tctacgcggaacactcgcaggaaccccggcaggaggacaaacttctggctgtcgggttactagagcggcagctacatccgcggcttacggtgccgtag
ag
```

a M R L V S V L G A V L L F E D R Q P N D L A V D V G A E C H G I
S -
b R C A L * A S L G P S S C L K T D S P M I S P S M * A P N A T A S
R -
c D A P C E R P W G R P P V * R P T A Q * S R R R C R R R M P R H L
-

Figure 19-HH

```
                                                                                              BstKTI
                                                                                               DpnI|
                                                                                              BsgI||
                                                                                             BstYI|||
                                                                                           Hpy188I|||
                                                                                              MboI|||
           MnlI                                           NlaIV                     AluI      BslI ||||
          CviJI|                                        AvaII                      CviJI      BpmI ||||
         BseRI ||                                       Sau96I |              MboII           Eco57MI ||||
        NlaIII ||                                      Hpy8I | |          Hpy188III|          CviJI | ||||
       BsaJI | ||                                      MnlI |  |           TspGWI||          HaeIII | ||||
       BtgI  | ||                             BsiHKAI |    |  |          Eco57I  |||           Sau96I | | ||||
  BceAI     FatI | ||                          Bsp1286I |   |  |        Eco57MI   |||
HpyCH4III|  NcoI | ||                       BssSI      |   |  |
 SfaNI  ||  StyI | ||
gcaaccgttcagcgaacgcctccatgggcttttctcctcgtgctcgtaaacggacccgaacatctctggagctttcttcagggccgacaatcggatc
tc
10101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 10200
cgttggcaagtcgcttgcggaggtacccgaaaaagaggagcacgagcatttgcctgggcttgtagagacctcgaaagaagtcccggctgttagcctag
ag

a     Q  P  F  S  E  R  L  H  G  L  F  L  L  V  L  V  N  G  P  E  H  L  W  S  F  L  Q  G  R  Q  S  D  L
-
b      N  R  S  A  N  A  S  M  G  F  F  S  S  C  S  *  T  D  P  N  I  S  G  A  F  F  R  A  D  N  R  I
S -
c    A  T  V  Q  R  T  P  P  W  A  F  S  P  R  A  R  K  R  T  R  T  S  L  E  L  S  S  G  P  T  I  G  S
R -

                 BsrBI
                  TauI
                 AciI |
                 BsiEI |
                Fnu4HI |
                 CviJI| |                MseI
               HaeIII| |          CviJI |
                                Hpy188I |  |
HhaI
|              Hpy99I|| |          HinfI  | | |                                                 BstUI
|               EaeI||| |        Hpy99I   | | |                                                 Cac8I
|               EagI||| |          TfiI   | | |                                                  HhaI
|               TaiI|||| |         TaqI | | | |                  MseI                            HinP1I
|             BceAI ||||| |       Hin4I| | | | |                Hin4I|                         BssHII |
|  AlwI       BmgBI ||||| |      BspCNI|| | | | |          Tsp509I  ||                        HinP1I |
|  AciI|     HpyCH4IV| ||||| |   BseMII||| | | | |     MfeI    |   ||          BceAI        HgaI | |
|  BstUI|   HpyCH4V || ||||| |  CviJI|||| |DdeI| | |Tsp509I   |   ||       MnlI   |  BsmFI | | |
       ||       | || ||||| |     ||||| |  || | |   |    |   ||             |   |    | | | |
gcggaaatcctgcacgtcggccgctccaagccgtcgaatctgagccttaatcacaattgtcaatttaatcctctgtttatcggcagttcgtagagcg
cg
10201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 10300
cgcctttaggacgtgcagccggcgaggttcggcagcttagactcggaattagtgttaacagttaaaattaggagacaaatagccgtcaagcatctcgc
gc

a     A  E  I  L  H  V  G  R  S  K  P  S  N  L  S  L  N  H  N  C  Q  F  *  S  S  V  Y  R  Q  F  V  E  R
A -
b      R  K  S  C  T  S  A  A  P  S  R  R  I  *  A  L  I  T  I  V  N  F  N  P  L  F  I  G  S  S  *  S  A
-
c      G  N  P  A  R  R  P  L  Q  A  V  E  S  E  P  *  S  Q  L  S  I  L  I  L  C  L  S  A  V  R  R  A
R -

                                                                                    Fnu4HI
                                                                                   HpyF10VI
                                                                                     CviJI|
                                                                                      MwoI|
                                                        Hpy188III                     TseI|
                                                        BcgI    |                   Cac8I ||
                   BcgI                              AciI  |    |                   HaeII| ||
                 BspCNI |                        Bme1580I  |    |                    HhaI|| ||
               HpyF10VI |                        Bsp1286I  |    |                   AfeI||| ||
                BseMII| |                           BtsI   |    |                HinP1I|||| ||
                  MwoI| |                          Cac8I   |    |                  XcmI||||| ||
                 AvaI || |                       HpyF10VI  |    |          HpyF10VI|||||| ||         HpaII
           Hpy188III || |                           TspRI  |    |               BbvI |||||| ||        BslI|
     HpyF10VI   | || |  DdeI      Hpy99I            BcgI|  |    |              BsrI|MwoI||||||| ||      CviJI|
      MwoI|     | || |BcgI |HgaI   TaqI |           MwoI|  |FauI |                   || ||||||| ||      BseYI  ||
         ||     | || |  |  |   |       ||             ||   |  | |                  ||  ||||||| ||             ||
ccgtgcgtcccgagcgatactgagcgaagcaagtgcgtcgagcagtgcccgcttgttcctgaaatgccagtaaagcgctggctgctgaacccccagcc
gg
```

```
   10301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 10400
         ggcacgcagggctcgctatgactcgcttcgttcacgcagctcgtcacgggcgaacaaggactttacggtcatttcgcgaccgacgacttgggggtcgg
cc

a          V  R  P  E  R  Y  *  A  K  Q  V  R  R  A  V  P  A  C  S  *  N  A  S  K  A  L  A  A  E  P  P  A
G  -
b        P  C  V  P  S  D  T  E  R  S  K  C  V  E  Q  C  P  L  V  P  E  M  P  V  K  R  W  L  L  N  P  Q  P
E -
c         R  A  S  R  A  I  L  S  E  A  S  A  S  S  S  A  R  L  F  L  K  C  Q  *  S  A  G  C  *  T  P  S  R
-
```

Figure 19-II

```
                                                                          MboII
                                                                     Fnu4HI   |
                                                                    MspA1I|   |
                                                                      TauI|   |
                                                                      TseI|   |
                                                                     AciI ||  |
                                          BstNI                    Fnu4HI ||  |
                                          ScrFI                     CviJI| || |
                                         BsrDI |                   HaeIII|| || |                      HpyF10VI
              BfaI                       PspGI |       BstNI      BstNI || || |                          MwoI|
           CviJI |                       SexAI |       ScrFI      ScrFI || || |                     CviJI ||
          HaeIII |                      StyD4I |      BsaJI |    PspGI| || || |                    Cac8I | ||
        Eco0109I||                   HpyCH4V | |      PspGI |   StyD4I| || || |                    EarI| | ||
          Sau96I||         HpyCH4V         | | |     StyD4I |   BbvI  | || || |              MnlI  | | | ||
               ||  |          |            | | |         |  |     |  | | || || |                |    | | | ||
     aactgaccccacaaggccctagcgtttgcaatgcaccaggtcatcattgacccaggcgtgttccaccaggccgctgcctcgcaactcttcgcaggctt
cg
10401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 10500
     ttgactggggtgttccgggatcgcaaacgttacgtggtccagtagtaactgggtccgcacaaggtggtccggcgacggagcgttgagaagcgtccgaa
gc

a     T  D  P  T  R  P  *  R  L  Q  C  T  R  S  S  L  T  Q  A  C  S  T  R  P  L  P  R  N  S  S  Q  A  S
-
b      L  T  P  Q  G  P  S  V  C  N  A  P  G  H  H  *  P  R  R  V  P  P  G  R  C  L  A  T  L  R  R  L
R  -
c     N  *  P  H  K  A  L  A  F  A  M  H  Q  V  I  I  D  P  G  V  F  H  Q  A  A  A  S  Q  L  F  A  G  F
A -


                                                                                                   FokI
                                                                                                  AluI|
                                                                                                 BceAI|
                                                                                                 CviJI|
                                                                                               Cac8I ||
                                                                                          BpuEI   | ||
                                                                                       HpyF10VI   | ||
                                                                                           MwoI|  | ||
                                           AciI                                           HpaII | | ||
                                     HpyF10VI |                                            NciI|| | ||
                                         MwoI||                                           ScrFI|| | ||
                                       NlaIII||                                        StyD4I ||| | ||
                                      FatI  ||                                         NlaIV| ||| | ||
                                  MnlI  |   ||                                        CviJI|| ||| | ||
                                 AciI|  |   ||                                   HpyF10VI ||| ||| | ||
                              BstKTI||  |   ||                    EciI              MwoI| ||| ||| | ||
         HhaI                   DpnI|||  |   ||                    SmlI              RsaI|| ||| ||| | ||
       BspMI|           AleI   Hpy188I |||| |   ||                 AluI|            Csp6I||| ||| ||| | ||
       BstUI||          MslI  AlwI  | |||| |   ||                CviJI| AciI        |||| ||| ||| | ||
      HinP1I||         AciI| HinfI  | |||| |   ||                FauI||BsrBI       |||| ||| ||| | ||
       MboII||    FauI BstUI| TfiI MboI |||| |   ||               ||| |           |||| ||| ||| | ||
       Cac8I |||    |     ||    |   | |||| |   ||               ||| |           |||| ||| ||| | ||
     ccgacctgctcgcgccacttcttcacgcgggtggaatccgatccgcacatgaggcggaaggtttccagcttgagcgggtacggctcccggtgcgagct
ga
10501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 10600
     ggctggacgagcgcggtgaagaagtgcgcccaccttaggctaggcgtgtactccgccttccaaaggtcgaactcgcccatgccgagggccacgctcga
ct

a     P  T  C  S  R  H  F  F  T  R  V  E  S  D  P  H  M  R  R  K  V  S  S  L  S  G  Y  G  S  R  C  E  L
K -
b      R  P  A  R  A  T  S  S  R  G  W  N  P  I  R  T  *  G  G  R  F  P  A  *  A  G  T  A  P  G  A  S  *
-
c       D  L  L  A  P  L  L  H  A  G  G  I  R  S  A  H  E  A  E  G  F  Q  L  E  R  V  R  L  P  V  R  A
E  -


                                       RsaI
                                     Csp6I|
        BstF5I       CviJI          AciI ||
      BceAI |       HaeIII        Cac8I | ||        ApoI
       TaqI |      Hpy99I |       AluI | | ||      Tsp509I
  TspGWI |  |      Sau96I |Hpy99I CviJI | | ||      NdeI |         TspDTI  Cac8I          Hpy99I
      |  |  |          | |   |      |  | | ||        |   |            |       |               |
     aatagtcgaacatccgtcgggccgtcggcgacagcttgcggtacttctcccatatgaatttcgtgtagtggtcgccagcaaacagcacgacgatttcc
tc
10601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 10700
     ttatcagcttgtaggcagcccggcagccgctgtcgaacgccatgaagagggtatacttaaagcacatcaccagcggtcgtttgtcgtgctgctaaagg
ag

a      *  S  N  I  R  R  A  V  G  D  S  L  R  Y  F  S  H  M  N  F  V  *  W  S  P  A  N  S  T  T  I  S
S  -
b     N  S  R  T  S  V  G  P  S  A  T  A  C  G  T  S  P  I  *  I  S  C  S  G  R  Q  Q  T  A  R  R  F  P
R -
c       I  V  E  H  P  S  G  R  R  R  Q  L  A  V  L  L  P  Y  E  F  R  V  V  V  A  S  K  Q  H  D  D  F  L
-
```

Figure 19-JJ

```
                 BstNI
                 ScrFI
                AlwNI|
               PspGI||
              StyD4I||
             AvaII |||                                                      Bme1580I
          Eco0109I |||                                                      Bsp1286I
             PpuMI |||                  BstNI                               NlaIV  |
            Sau96I |||                  ScrFI                              BanI |  |
       Hpy188III | |||                   PfoI |                           BsgI | | |
BmgBI                                   PspGI |                          BstNI | | |
          MnlI | |||                   StyD4I |                          ScrFI | | |
HpyCH4IV|                                                                 BslI| | |
       BstKTI| | |||                    AvaII | |          Fnu4HI        PspGI|| | | |
BceAI||                                                                 StyD4I|| | | |
         DpnI|| | |||                HpyCH4III | |           TseI|       BbvI ||| | | |
Hpy188I|||                                                              MmeI ||| | | |
      Hpy99I||| | |||                   Sau96I | |         Fnu4HI ||   HinfI| ||| | | |      BsiEI
||||                           TaiI     BsaJI  | | | AciI  HpyCH4V| ||   TfiI| ||| | | | AciI  |
       MboI|||| | ||| HpyCH4IV |        BsmFI  | | | BstUI AciI   || ||  BbvI || ||| | | |BstUI |
||||
      TaqI||||| | ||| BslI     | |      BtgI   | | | BslI| HgaI TseI| ||
||||
         |||||| | |||    |     | |        |    | | |   ||    |    || ||    | || ||| | | |    |   |
||||
   gtcgatcaggacctggcaacgggacgttttcttgccacggtccaggacgcggaagcggtgcagcagcgacaccgattccaggtgcccaacgcggtcgg
ac
10701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 10800
   cagctagtcctggaccgttgccctgcaaaagaacggtgccaggtcctgcgccttcgccacgtcgtcgctgtggctaaggtccacgggttgcgccagcc
tg

a     S  I  R  T  W  Q  R  D  V  F  L  P  R  S  R  T  R  K  R  C  S  S  D  T  D  S  R  C  P  T  R  S  D
-
b      R  S  G  P  G  N  G  T  F  S  C  H  G  P  G  R  G  S  G  A  A  A  T  P  I  P  G  A  Q  R  G  R
T  -
c     V  D  Q  D  L  A  T  G  R  F  L  A  T  V  Q  D  A  E  A  V  Q  Q  R  H  R  F  Q  V  P  N  A  V  G
R -
```

```
                                                                                        BstNI
                                                                                        ScrFI
                                                                                       AlwNI|
                                                                                        BslI|
                                                                                       PflMI|
                                                                                      PspGI||
                                                                                     StyD4I||
                                                                                     AvaII |||
                                                                                  Eco0109I |||
                                                                                     PpuMI |||
                                                                                    Sau96I |||
                                                                                  BstNI |  |||
                                      BsmI                                       ScrFI  |  |||
                              HpyF10VI |                   CviJI      BstKTI    BsaJI | | |||
                        SfcI  BstUI  | |        CviJI      HaeIII |    TaqI     PspGI | | |||
          CviJI    Hpy99I |    HhaI  | |        HaeIII       EaeI |    DpnI|    StyD4I | | |||     AluI
        TaiI |      BccI | |HinP1I |MwoI| |BsaJI   | MnlI   HpaII| |  MboI ||   CviJI| | | |||     CviJI
           |  |         |  | |     | |   || |     |    |      |      BsrFI||  |   | ||     || | | |||       |
   gtgaagcccatcgccgtcgcctgtaggcgcgacaggcattcctcggccttcgtgtaataccggccattgatcgaccagcccaggtcctggcaaagctc
gt
10801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 10900
   cacttcgggtagcggcagcggacatccgcgctgtccgtaaggagccggaagcacattatggccggtaactagctggtcgggtccaggaccgtttcgag
ca

a     V  K  P  I  A  V  A  C  R  R  D  R  H  S  S  A  F  V  *  Y  R  P  L  I  D  Q  P  R  S  W  Q  S  S
*  -
b      *  S  P  S  P  S  P  V  G  A  T  G  I  P  R  P  S  C  N  T  G  H  *  S  T  S  P  G  P  G  K  A  R
-
c       E  A  H  R  R  R  L  *  A  R  Q  A  F  L  G  L  R  V  I  P  A  I  D  R  P  A  Q  V  L  A  K  L
V  -
```

Figure 19-KK

```
                                                                                                   Bb
vI                                                                                               HpaI
I|                                                                                              BsrFI
||                                                                                             Hpy99I
||                                                                                             SgrAI
||                                                                                            BsaHI
|||                                                                                       FauI  |
|||                                                                                       TaqI  |
|||                                                                                    AluI |   |
|||                                                                                   CviJI |   |
|||                                                                                Fnu4HI | |   |
|||                  HphI                                                            TseI| | |   |
|||              Cac8I |                                                            AciI || | |   |
|||            CviJI | |                BcgI                                       Cac8I || | |   |
|||          BstKTI | | |              BbvI |           AarI                     CviJI | || | |   |
|||            DpnI| | | |             RsaI |          BspMI                    HaeIII | || | |   |
|||     TaiI   BcgI|| | | |      HhaI  Csp6I| |       Fnu4HI   |                Hpy99I| | || | |   |
||| HpyCH4IV | MboI||| | | |    HinP1I |BstUI || |       TseI|   |BsrI  MmeI   Sau96I| | || | |   |
|||     | |    |||| | | |        | |    | || |          ||   |    |     |       || | || | |   |
    agaacgtgaaggtgatcggctcgccgatagggggtgcgcttcgcgtactccaacacctgctgccacaccagttcgtcatcgtcggcccgcagctcgacgcc
10901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 11000
    tcttgcacttccactagccgagcggctatccccacgcgaagcgcatgaggttgtggacgacggtgtggtcaagcagtagcagccgggcgtcgagctgcgg

a     N  V  K  V  I  G  S  P  I  G  V  R  F  A  Y  S  N  T  C  C  H  T  S  S  S  S  S  A  R  S  S  T  P  -
b    R  T  *  R  *  S  A  R  R  *  G  C  A  S  R  T  P  T  P  A  A  T  P  V  R  H  R  R  P  A  A  R  R  R  -
c     E  R  E  G  D  R  L  A  D  R  G  A  L  R  V  L  Q  H  L  L  P  H  Q  F  V  I  V  G  P  Q  L  D  A  -

                                                              AciI
                                                              BbvI
                                                              BslI
                                                             BstUI
                                                              HhaI
                                                           BstUI |
                                                          HinP1I |
                                                        HpyF10VI |
                                                           MwoI| |
                                                       HaeII  || |
                                                        FauI| || |
                                                        HhaI| || |
                    TaiI                             HinP1I || || |
                   HphI|          TaiI
                 BmgBI||        BmgBI |
             HpyCH4IV|||
AciI
           BstKTI     |||| HpyCH4IV| |       Fnu4HI | ||  || |              BstUI
BsrBI
|      HgaI    DpnI|  |||| HincII || |     HpyCH4V| | ||  || |             HhaI    BceAI    HphI
|   MboII MboI ||     |||| Hpy8I || |         TseI| | ||  || |MnlI      HinP1I |  Hpy8I|FauI  |
|       |    | ||     ||||    | || |           || | ||  || |   |             | |     ||   |   |
ag  ggtgtaggtgatcttcacgtccttgttgacgtggaaaatgaccttgtttgcagcgcctcgcgcgggattttcttgttgcgcgtggtgaacagggcag
11001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 11100
tc  ccacatccactagaagtgcaggaacaactgcaccttttactggaacaaaacgtcgcggagcgcgccctaaaagaacaacgcgcaccacttgtcccgtc

a     V  *  V  I  F  T  S  L  L  T  W  K  M  T  L  F  C  S  A  S  R  G  I  F  L  L  R  V  V  N  R  A  E  -
b    C  R  *  S  S  R  P  C  *  R  G  K  *  P  C  F  A  A  P  R  A  G  F  S  C  C  A  W  *  T  G  Q  S  -
```

```
c        G  V  G  D  L  H  V  L  V  D  V  E  N  D  L  V  L  Q  R  L  A  R  D  F  L  V  A  R  G  E  Q  G  R
A -

                                                 HhaI
                                              HinP1I |
                                           BsaJI   | |
                                            BtgI   | |
       CviJI                              CviJI|   | |                       BbvI
      HaeIII                  SfaNI      HaeIII|   | |              HpyCH4V |         Fnu4HI
      Cac8I |               Cac8I |        EaeI ||   | |        BceAI  Fnu4HI  | |         TseI|
   HpyF10VI |            HpyF10VI |       HpaII| ||   | |   BbvI    |   AluI| | |   BstKTI   ||
     Sau96I |                MwoI||      SfaNI| ||   | |   TaqI    |  CviJI| | |    DpnI|   ||
       MwoI||                    ||           || ||   | |     |     |   TseI| | |    MboI || ||
         cgggccgtgtcgtttggcatcgctcgcatcgtgtccggccacggcgcaatatcgaacaaggaaagctgcatttccttgatctgctgcttcgtgtgttt
ca
   11101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 11200
         gcccggcacagcaaaccgtagcgagcgtagcacaggccggtgccgcgttatagcttgttcctttcgacgtaaaggaactagacgacgaagcacacaaa
gt

a        R  A  V  S  F  G  I  A  R  I  V  S  G  H  G  A  I  S  N  K  E  S  C  I  S  L  I  C  C  F  V  C  F
S -
b         G  P  C  R  L  A  S  L  A  S  C  P  A  T  A  Q  Y  R  T  R  K  A  A  F  P  *  S  A  A  S  C  V  S
-
c          G  R  V  V  W  H  R  S  H  R  V  R  P  R  R  N  I  E  Q  G  K  L  H  F  L  D  L  L  L  R  V  F
Q  -
```

Figure 19-LL

```
                              CviJI
                             HaeIII
                         HpyF10VI |               AvaII          MnlI
                       Cac8I    | |            EcoO109I       AciI |
                     CviJI |    | |               PpuMI    Cac8I | |
                    HaeIII |    | |              Sau96I     NaeI | |
                      TauI |    | |              BstNI |   HpaII|  | |
                   Fnu4HI| |BglI| |              ScrFI |  BsrFI||  | |
                    AciI| | |MwoI| |            PspGI | | NgoMIV|| | |                        BccI MnlI
TaqI
                   BstUI| | |SfiI| |        MnlI StyD4I | |  SgrAI|| | |                BstUI|TaqI|BsaBI
|
                      || | |   || |            |       | | |      ||| | |                   ||    ||    |
|
         gcaacgcggcctgcttggcctcgctgacctgttttgccaggtcctcgccggcggtttttcgcttcttggtcgtcatagttcctcgcgtgtcgatggtc
at
   11201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 11300
         cgttgcgccggacgaaccggagcgactggacaaaacggtccaggagcggccgccaaaaagcgaagaaccagcagtatcaaggagcgcacagctaccag
ta

a          N  A  A  C  L  A  S  L  T  C  F  A  R  S  S  P  A  V  F  R  F  L  V  V  I  V  P  R  V  S  M  V
I -
b         A  T  R  P  A  W  P  R  *  P  V  L  P  G  P  R  R  R  F  F  A  S  W  S  S  *  F  L  A  C  R  W  S
S -
c         Q  R  G  L  L  G  L  A  D  L  F  C  Q  V  L  A  G  G  F  S  L  L  G  R  H  S  S  S  R  V  D  G  H
-


                                                                          HpyF10VI
                                                                         Cac8I    |
                                                                        BceAI|    |
                                                                         AciI||   |
                                                                        BstUI||   |
                                                                         HhaI||   |
                                                                     HinP1I |||   |
                                                                   HpyF10VI |||   |
                                                                      FauI | |||   |
                                                                   BsiEI| | |||   |
                                                                   CviJI|| | |||   |
                                                                  HaeIII|| | |||   |
                                          MnlI                      TauI|| | |||   |
                                   Hpy188III|                       EaeI||| | |||   |
                                       BsmAI ||                      EagI||| | |||   |                   Cac8I
                                       BsmBI ||                    Fnu4HI||| | |||   |              HpyCH4V |
                                   BspMI |  .||                   BsaJI | ||| | |||   |             BsrI    | |
                                  TauI | |   ||                                        |          CviJI|   | |
                                 AciI | | |   ||                                       |          NlaIV||   | |
                              Fnu4HI | | |   ||
BstUI
                   HpyF10VI | | | |     ||    BstUI     BtgI AciI| ||| | ||| | HpyF10VI||| | |
HinP1I
                      MwoI| | | | |TaqI|| Hpy99I HgaI | BccI| ||MwoI| ||MwoI| MwoI|||| | | BceAI
|
                         || | | | | |||      |      | |  || |||  || |||  ||     |||||    | |      |
|
         cgacttcgccaaacctgccgcctcctgttcgagacgacgcgaacgctccacggcggccgatggcgcgggcagggcagggggagccagttgcacgctgt
cg
   11301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 11400
         gctgaagcggtttggacggcggaggacaagctctgctgcgcttgcgaggtgccgccggctaccgcgcccgtcccgtccccctcggtcaacgtgcgaca
gc

a          D  F  A  K  P  A  A  S  C  S  R  R  R  E  R  S  T  A  A  D  G  A  G  R  A  G  G  A  S  C  T  L  S
-
b         T  S  P  N  L  P  P  P  V  R  D  D  A  N  A  P  R  R  P  M  A  R  A  G  Q  G  E  P  V  A  R  C
R -
c         R  L  R  Q  T  C  R  L  L  F  E  T  T  R  T  L  H  G  G  R  W  R  G  Q  G  R  G  S  Q  L  H  A  V
A -
```

Figure 19-MM

```
                                                                                    FokI
                                                                                  Cac8I|
                                                                                 BccI ||
                                                                                CviJI ||
                                                                                 TauI ||
                                                                             Fnu4HI | ||
                                                                              AciI| | ||
                                                                        HpyCH4III || | ||
                                                                           MslI | || | ||
                                                                        NlaIII| | || | ||
                                                                       FatI  || | || | ||
                                                                    Cac8I |  || | || | ||
                                                                   HhaI | |  || | || | ||
                                                                 HinP1I | | |  || | || | ||
          CviJI                                                 AciI  | | | |  || | || | ||
         HaeIII                                                BstUI  | | | |  || | || | ||
           EaeI |                  AvaII                     TspGWI | | | | |  || | || | ||
     BstKTI  | |                  Sau96I        Hpy99I         FauI | | | | | |  || | || | ||
       DpnI| | |          Cac8I       |      BccI     |        BsrI| | | | | | |  || | || | ||
     MboI || | |          AluI |      |    CviJI|     |            || | | | | | |  || | || | ||
    TaqI| || | |        CviJI  |      |  TaqI  ||    |            || | | | | | |  || | || | ||      BstF5I
    HhaI || || | |             |      |     |  ||     |            || | | | | | |  || | || | ||           |
   cgctcgatcttggccgtagcttgctggaccatcgagccgacggactggaaggtttcgcggggcgcacgcatgacggtgcggcttgcgatggtttcggc
at
 11401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 11500
   gcgagctagaaccggcatcgaacgacctggtagctcggctgcctgaccttccaaagcgccccgcgtgcgtactgccacgccgaacgctaccaaagccg
ta

a        R  S  I  L  A  V  A  C  W  T  I  E  P  T  D  W  K  V  S  R  G  A  R  M  T  V  R  L  A  M  V  S  A
S -
b          A  R  S  W  P  *  L  A  G  P  S  S  R  R  T  G  R  F  R  G  A  H  A  *  R  C  G  L  R  W  F  R  H
-
c         L  D  L  G  R  S  L  L  D  H  R  A  D  G  L  E  G  F  A  G  R  T  H  D  G  A  A  C  D  G  F  G
I  -


                          BstKTI
                            DpnI|
                            FauI||
                          Hpy99I||                                HinfI
       HgaI                 MboI|||                                TfiI
       AciI|                EciI||||                               HphI|
     SfaNI|                 TaqI||||                            Hpy188I|
  HpyF10VI ||                                                     TaiI ||
BsaJI                                                 HpyCH4IV  | ||                 MnlI
   MwoI| ||             BstUI  |||||          HpaII              | | ||           AciI  |MlyI
HpaII
   BsaJI || ||MnlI AciI |  |||||         BsaWI|TspDTI |  | ||       FauI  BslI|  |PleI HinfI
StyD4I|
||
   cctcggcggaaaaccccgcgtcgatcagttcttgcctgtatgccttccggtcaaacgtccgattcattcaccctccttgcgggattgccccgactcac
gc
 11501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 11600
   ggagccgccttttggggcgcagctagtcaagaacggacatacggaaggccagtttgcaggctaagtaagtgggaggaacgccctaacggggctgagtg
cg

a          S  A  E  N  P  A  S  I  S  S  C  L  Y  A  F  R  S  N  V  R  F  I  H  P  P  C  G  I  A  P  T  H
A  -
b        P  R  R  K  T  P  R  R  S  V  L  A  C  M  P  S  G  Q  T  S  D  S  F  T  L  L  A  G  L  P  R  L  T
P -
c         L  G  G  K  P  R  V  D  Q  F  L  P  V  C  L  P  V  K  R  P  I  H  S  P  S  L  R  D  C  P  D  S  R
-


                                       BsaJI                                                        BcgI
                                        StyI                                                   HpyCH4III
                                      NlaIV |                                                    Hpy8I |
                                       FauI||                                                    AccI| |
                                       BanI|||                                                 TspDTI| |
                                     BstNI ||||                                              BceAI  || |
                   Hpy188III          ScrFI||||                                            NlaIV|   || |
           Bme1580I   |              PspGI ||||                                            AvaII||  || |
CviJI
       NciI  Bsp1286I  |             StyD4I | ||| |                  TaqII            Sau96I ||  ||
|HaeIII
    ScrFI     BsrDI |  |             AciI | | ||| |Hpy188III    BcgI    BsmFI |     BsrDI  | ||  ||  |EaeI
|
|
   cggggcaatgtgcccttattcctgatttgacccgcctggtgccttggtgtccagataatccaccttatcggcaatgaagtcggtcccgtagaccgtct
gg
 11601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 11700
   gccccgttacacgggaataaggactaaactgggcggaccacggaaccacaggtctattaggtggaatagccgttacttcagccagggcatctggcaga
cc
```

a - G A M C P Y S * F D P P G A L V S R * S T L S A M K S V P * T V W
b G - G Q C A L I P D L T R L V P W C P D N P P Y R Q * S R S R R P S
c A - R G N V P L F L I * P A W C L G V Q I I H L I G N E V G P V D R L

Figure 19-NN

```
                                                                                    SfaNI
                                                                                  CviJI |
                                                                               MnlI  |  |
                                                                            DdeI  |  |  |
                                                                         CviJI |  |  |  |
                                                                        HaeIII |  |  |  |
                                                                     BsaJI  |  |  |  |  |
                                                                   BspCNI|  |  |  |  |  |
                                                                    CviJI|  |  |  |  |  |
                                  HinfI                            HaeIII|  |  |  |  |  |
                                   TfiI                          BseMII||  |  |  |  |  |
         RsaI            Hpy188I|                                Sau96I|||  |  |  |  |  |
        Csp6I|         BsgI    ||    HpyCH4V          BceAI    BsaJI  ||||  |  |  |  |  |
            ||           |     ||       |               |      BtgI   ||||  |  |  |  |  |
    ccgtccttctcgtacttggtattccgaatcttgccctgcacgaataccagcgaccccttgcccaaatacttgccgtgggcctcggcctgagagccaaa
ac
  11701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 11800
    ggcaggaagagcatgaaccataaggcttagaacgggacgtgcttatggtcgctggggaacgggtttatgaacggcacccggagccggactctcggttt
tg

a       P  S  F  S  Y  L  V  F  R  I  L  P  C  T  N  T  S  D  P  L  P  K  Y  L  P  W  A  S  A  *  E  P  K
H -
b        R  P  S  R  T  W  Y  S  E  S  C  P  A  R  I  P  A  T  P  C  P  N  T  C  R  G  P  R  P  E  S  Q  N
-
c         V  L  L  V  L  G  I  P  N  L  A  L  H  E  Y  Q  R  P  L  A  Q  I  L  A  V  G  L  G  L  R  A  K
T  -

                                                                     TspDTI
                                         Cac8I                        RsaI|
                                          NaeI                       Csp6I||
                           HhaI         HpaII|        HhaI           FokI||
                       HinP1I |        BsrFI||      SfaNI|    BfaI    |||       BstF5I
           AciI         MboII |      NgoMIV||     HinP1I||    BslI    |||Tsp509I  |BsrI          SspI
             |              | |            |||          |||      |    |||    |    |  |              |
    acttgatgcggaagaagtcggtgcgctcctgcttgtcgccggcatcgttgcgccacatctaggtactaaaacaattcatccagtaaaatataatattt
ta
  11801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 11900
    tgaactacgccttcttcagccacgcgaggacgaacagcggccgtagcaacgcggtgtagatccatgattttgttaagtaggtcattttatattataaa
at

a        L  M  R  K  K  S  V  R  S  C  L  S  P  A  S  L  R  H  I  *  V  L  K  Q  F  I  Q  *  N  I  I  F
Y  -
b       T  *  C  G  R  S  R  C  A  P  A  C  R  R  H  R  C  A  T  S  R  Y  *  N  N  S  S  S  K  I  *  Y  F
I  -
c        L  D  A  E  E  V  G  A  L  L  L  V  A  G  I  V  A  P  H  L  G  T  K  T  I  H  P  V  K  Y  N  I  L
-

                                                                              BsiEI
                            BsrI                                              HpaII
                       BstKTI |                                              BsaWI|
                         DpnI| |                 HpyCH4III                   MnlI ||
                        BmrI|| |                MboII   |                 BstKTI| ||
                        MboI|| |            TaqI  |     |                   TaqI| ||
                    CviJI | || |           AluI|  |     |                  DpnI|| ||
                   AlwI | | || |          CviJI|  |     |          EcoRV   MboI|| ||          HgaI
                      | | | || |              ||  |     |            |        |||| ||            |
    ttttctcccaatcaggcttgatccccagtaagtcaaaaaatagctcgacatactgttcttccccgatatcctccctgatcgaccggacgcagaaggca
at
  11901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 12000
    aaaagagggttagtccgaactaggggtcattcagttttttatcgagctgtatgacaagaaggggctataggagggactagctggcctgcgtcttccgt
ta

a        F  L  P  I  R  L  D  P  Q  *  V  K  K  *  L  D  I  L  F  F  P  D  I  L  P  D  R  P  D  A  E  G  N
-
b         F  S  Q  S  G  L  I  P  S  K  S  K  N  S  S  T  Y  C  S  S  P  I  S  S  L  I  D  R  T  Q  K  A
M  -
c        F  S  P  N  Q  A  *  S  P  V  S  Q  K  I  A  R  H  T  V  L  P  R  Y  P  P  *  S  T  G  R  R  R  Q
C  -

                                                                                 BstNI
                                                                                 ScrFI
                                                                                BsmAI|
                           TauI       BstKTI                                   BsaJI||
   EciI                    AciI |        DpnI|                          MslI      PspGI||   BsaJI
  BsrDI|        AciI     Fnu4HI |        MboI ||    CviJI           BccI    |   BceAI StyD4I||   BtgI
      ||          |          |  |           | ||      |               |    |       |     |||      |
    gtcataccacttgtccgccctgccgcttctcccaagatcaataaagccacttactttgccatctttcacaaagatgttgctgtctcccaggtcgccgt
gg
  12001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 12100
```

```
       cagtatggtgaacaggcgggacggcgaagagggttctagttatttcggtgaatgaaacggtagaaagtgtttctacaacgacagagggtccagcggca
cc

a      V  I  P  L  V  R  P  A  A  S  P  K  I  N  K  A  T  Y  F  A  I  F  H  K  D  V  A  V  S  Q  V  A  V
G -
b       S  Y  H  L  S  A  L  P  L  L  P  R  S  I  K  P  L  T  L  P  S  F  T  K  M  L  L  S  P  R  S  P  W
-
c        H  T  T  C  P  P  C  R  F  S  Q  D  Q  *  S  H  L  L  C  H  L  S  Q  R  C  C  C  L  P  G  R  R
G  -
```

Figure 19-OO

```
                                                                                          MboII
                                                                                       BbsI    |
                                                                                      MboII    |
                                                                                     AlwI |    |
                                                                                    DraI| |    |
                                                                                   MseI|| |    |
                                                                              BstKTI  ||| |    |
                                                                                DpnI| ||| |    |
                                                                              BstYI || ||| |    |
                                                                               MboI || ||| |    |
                                                                             AciI | || ||| |    |
                                                                            BstUI | || ||| |    |
                                                                             HhaI | || ||| |    |
                                                                          BstUI | | || ||| |    |
                       MnlI                                              HinPlI | | || ||| |    |
                      CviJI|          DraI                              Cac8I | | | || ||| |    |
                      EarI ||         MseI|                             AluI | | | | || ||| |    |
       MboII     TspGWI | ||      FalI  ||                            CviJI | | | | | || ||| |    |        BmrI  BsrI
        |         |   | ||         |   ||                            | | | | | | || ||| |    |         |     |
       gaaaagacaagttcctcttcgggcttttccgtctttaaaaaatcatacagctcgcgcggatctttaaatggagtgtcttcttcccagttttcgcaatc
ca
  12101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 12200
       cttttctgttcaaggagaagcccgaaaaggcagaaatttttagtatgtcgagcgcgcctagaaatttacctcacagaagaagggtcaaaagcgttag
gt

a        K  D  K  F  L  F  G  L  F  R  L  *  K  I  I  Q  L  A  R  I  F  K  W  S  V  F  F  P  V  F  A  I
H  -
b       E  K  T  S  S  S  S  G  F  S  V  F  K  K  S  Y  S  S  R  G  S  L  N  G  V  S  S  S  Q  F  S  Q  S
T -
c        K  R  Q  V  P  L  R  A  F  P  S  L  K  N  H  T  A  R  A  D  L  *  M  E  C  L  L  P  S  F  R  N  P
-

                                                          AluI
                                                         CviJI
                                                    HpyF10VI |
                                               CviJI       | |
           BstKTI                                TauI      | |
            DpnI|                              Fnu4HI |    | |
          MboI ||                           BlpI    | |    | |                        TaqI
       CviJI | ||                           DdeI    | |  MwoI| |                  BsmFI |
       HaeIII | ||                         CviJI|AciI| |  DdeI|| |             BccI | |
       EaeI | | ||              Tsp509I   CviJI|AciI| | DdeI|| |         Hpy188I | | |        SfaNI
CviJI

|        | |  | ||              |        ||   || |    ||| |            |  | | |          |
       catcggccagatcgttattcagtaagtaatccaattcggctaagcggctgtctaagctattcgtatagggacaatccgatatgtcgatggagtgaaag
ag
  12201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 12300
       gtagccggtctagcaataagtcattcattaggttaagccgattcgccgacagattcgataagcatatccctgttaggctatacagctacctcactttc
tc

a        I  G  Q  I  V  I  Q  *  V  I  Q  F  G  *  A  A  V  *  A  I  R  I  G  T  I  R  Y  V  D  G  V  K  E
-
b         S  A  R  S  L  F  S  K  *  S  N  S  A  K  R  L  S  K  L  F  V  *  G  Q  S  D  M  S  M  E  *  K
S  -
c        H  R  P  D  R  Y  S  V  S  N  P  I  R  L  S  G  C  L  S  Y  S  Y  R  D  N  P  I  C  R  W  S  E  R
A -

                                                                                              MnlI
                                                                                           NlaIII
                                                                                     Hpy188III |
                                                                                        BspHI| |
                                                                                         FatI| |
                                                                                         BpmI|| |
                                                                                    Eco57MI|| |
                                                                       BsaHI        BccI | ||| |
                                       MboII                         HpyF10VI    CviJI | ||| |
                                      TspDTI                          MwoI|     HaeIII | ||| |
      HpyCH4V              TaqI       CviJI |                   EarI    ||       HgaI| | ||| |
     BcgI | AciI          AluI |      BcgI | |                 Hpy188I BslI ||       || | ||| |
      |   |  |          CviJI | |    TspDTI| | |      MboII         |    |   ||       || | ||| |
       cctgatgcactccgcatacagctcgataatcttttcagggctttgttcatcttcatactcttccgagcaaaggacgccatcggcctcactcatgagca
ga
  12301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 12400
       ggactacgtgaggcgtatgtcgagctattagaaaagtcccgaaacaagtagaagtatgagaaggctcgtttcctgcggtagccggagtgagtactcgt
ct

a        P  D  A  L  R  I  Q  L  D  N  L  F  R  A  L  F  I  F  I  L  F  R  A  K  D  A  I  G  L  T  H  E  Q
I -
b         L  M  H  S  A  Y  S  S  I  I  F  S  G  L  C  S  S  S  Y  S  S  E  Q  R  T  P  S  A  S  L  M  S  R
-
c         *  C  T  P  H  T  A  R  *  S  F  Q  G  F  V  H  L  H  T  L  P  S  K  G  R  H  R  P  H  S  *  A
D  -
```

Figure 19-PP

```
                                    AvaII                          AluI
                   BccI           Eco0109I                        CviJI
                 NlaIII|            PpuMI                        BsgI |                             SfaNI
               FatI   ||           Sau96I                     Fnu4HI |         CviJI          NlaIII |               BsmFI        XcmI
      BceAI  CviJI  |  ||         HpyCH4V  |                   TseI| |        BbvI  |          FatI  ||                 |           |
  ttgctccagccatcatgccgttcaaagtgcaggacctttggaacaggcagctttccttccagccatagcatcatgtcctttttcccgttccacatcata
gg
12401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 12500
  aacgaggtcggtagtacggcaagtttcacgtcctggaaaccttgtccgtcgaaaggaaggtcggtatcgtagtacaggaaaaagggcaaggtgtagtat
cc

a     A  P  A  I  M  P  F  K  V  Q  D  L  W  N  R  Q  L  S  F  Q  P  *  H  H  V  L  F  P  F  H  I  I
G -
b     L  L  Q  P  S  C  R  S  K  C  R  T  F  G  T  G  S  F  P  S  S  H  S  I  M  S  F  S  R  S  T  S  *
V -
c     C  S  S  H  H  A  V  Q  S  A  G  P  L  E  Q  A  A  F  L  P  A  I  A  S  C  P  F  P  V  P  H  H  R
-

                CviJI
    NlaIV    HpaII |              TspDTI                                    BsmAI
   AvaII |   BsrFI|  |              DraI |                        AluI     Bpu10I |
  Sau96I | TspGWI ||  |             MseI| |                       CviJI     DdeI  | TspGWI
  tggtccctttataccggctgtccgtcatttttaaatataggttttcattttctcccaccagcttatataccttagcaggagacattccttccgtatct
tt
12501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 12600
  accagggaaatatggccgacaggcagtaaaaatttatatccaaaagtaaaagagggtggtcgaatatatggaatcgtcctctgtaaggaaggcataga
aa

a     G  P  F  I  P  A  V  R  H  F  *  I  *  V  F  I  F  S  H  Q  L  I  Y  L  S  R  R  H  S  F  R  I  F
-
b     V  P  L  Y  R  L  S  V  I  F  K  Y  R  F  S  F  S  P  T  S  L  Y  T  L  A  G  D  I  P  S  V  S
F -
c     W  S  L  Y  T  G  C  P  S  F  L  N  I  G  F  H  F  L  P  P  A  Y  I  P  *  Q  E  T  F  L  P  Y  L
L -

                BstKTI
    AciI         DpnI|
   MspA1I        MboI ||            HpaII                                        HpyCH4III
  Fnu4HI |       TaqI| ||           BsaWI|                                        MnlI |  DraI
   TseI| |       BbvI|| ||   Tsp509I    ||                  HphI  CviJI.          SfcI | | MseI|
  tacgcagcggtatttttcgatcagtttttcaattccggtgatattctcattttagccatttattatttccttcctcttttctacagtatttaaagat
ac
12601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 12700
  atgcgtcgccataaaaagctagtcaaaaaagttaaggccactataagagtaaaatcggtaaataataaaggaaggagaaaagatgtcataaatttcta
tg

a     Y  A  A  V  F  F  D  Q  F  F  Q  F  R  *  Y  S  H  F  S  H  L  L  F  P  S  S  F  L  Q  Y  L  K  I
P -
b     T  Q  R  Y  F  S  I  S  F  F  N  S  G  D  I  L  I  L  A  I  Y  Y  F  L  P  L  F  Y  S  I  *  R  Y
-
c     R  S  G  I  F  R  S  V  F  S  I  P  V  I  F  S  F  *  P  F  I  I  S  F  L  F  S  T  V  F  K  D
T -

                                           BsmI
                                         HpyCH4V
  Tsp509I                               TspRI |
    AluI |                     HpyCH4III    | |                                    AluI
   CviJI | PsiI          Tsp509I        |   | |                MseI               CviJI
  cccaagaagctaattataacaagacgaactccaattcactgttccttgcattctaaaaccttaaataccagaaaacagctttttcaaagttgttttca
aa
12701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 12800
  gggttcttcgattaatattgttctgcttgaggttaagtgacaaggaacgtaagattttggaatttatggtcttttgtcgaaaaagtttcaacaaaagt
tt

a     Q  E  A  N  Y  N  K  T  N  S  N  S  L  F  L  A  F  *  N  L  K  Y  Q  K  T  A  F  S  K  L  F  S
K -
b     P  K  K  L  I  I  T  R  R  T  P  I  H  C  S  L  H  S  K  T  L  N  T  R  K  Q  L  F  Q  S  C  F  Q
S -
c     P  R  S  *  L  *  Q  D  E  L  Q  F  T  V  P  C  I  L  K  P  *  I  P  E  N  S  F  F  K  V  V  F  K
-

                                     BstXI
                                    MslI |
                                   BbvI| |
                   CviJI          SfaNI| |
                  NlaIV|        Tsp509I|| |                                     MnlI
```

```
                        Hpy99I ||        TaqII   ||| |       Fnu4HI                                     BpmI |
                   TaqI      | ||      TspGWI |  ||| |        TseI| HphI                  BccI  Eco57MI |
                      |      | ||            | |  ||| |           || |                       |       ||
          gttggcgtataacatagtatcgacggagccgattttgaaaccacaattatgggtgatgctgccaacttactgatttagtgtatgatggtgtttttgag
gt
   12801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 12900
          caaccgcatattgtatcatagctgcctcggctaaaactttggtgttaatacccactacgacggttgaatgactaaatcacatactaccacaaaaactc
ca

a          L  A  Y  N  I  V  S  T  E  P  I  L  K  P  Q  L  W  V  M  L  P  T  Y  *  F  S  V  *  W  C  F  *  G
-
b           W  R  I  T  *  Y  R  R  S  R  F  *  N  H  N  Y  G  *  C  C  Q  L  T  D  L  V  Y  D  G  V  F  E
V  -
c          V  G  V  *  H  S  I  D  G  A  D  F  E  T  T  I  M  G  D  A  A  N  L  L  I  *  C  M  M  V  F  L  R
C -
```

Figure 19-QQ

```
gctccagtggcttctgtgtctatcagctgtccctcctgttcagctactgacggggtggtgcgtaacggcaaaagcaccgccggacatcagcgctatct
12901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 13000
cgaggtcaccgaagacacagatagtcgacagggaggacaagtcgatgactgccccaccacgcattgccgttttcgtggcggcctgtagtcgcgataga

a   A P V A S V S I S C P S C S A T D G V V R N G K S T A G H Q R Y L
b   L Q W L L C L S A V P P V Q L L T G W C V T A K A P P D I S A I S
c   S S G F C V Y Q L S L L F S Y * R G G A * R Q K H R R T S A L S

gctctcactgccgtaaaacatggcaactgcagttcacttacaccgcttctcaacccggtacgcaccagaaaatcattgatatggccatgaatggcgtt
13001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 13100
cgagagtgacggcattttgtaccgttgacgtcaagtgaatgtggcgaagagttgggccatgcgtggtcttttagtaactataccggtacttaccgcaa

a   S H C R K T W Q L Q F T Y T A S Q P G T H Q K I I D M A M N G V
b   A L T A V K H G N C S S L T P L L N P V R T R K S L I W P * M A L
c   L S L P * N M A T A V H L H R F S T R Y A P E N H * Y G H E W R W

atgccgggcaacagcccgcattatgggcgttggcctcaacacgattttacgtcacttaaaaaactcaggccgcagtcggtaacctcgcgcatacagcc
13101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 13200
tacggcccgttgtcgggcgtaatacccgcaaccggagttgtgctaaaatgcagtgaatttttgagtccggcgtcagccattggagcgcgtatgtcgg

a   C R A T A R I M G V G L N T I L R H L K N S G R S R * P R A Y S R
b   A G Q Q P A L W A L A S T R F Y V T * K T Q A A V G N L A H T A
c   M P G N S P H Y G R W P Q H D F T S L K K L R P Q S V T S R I Q P
```

Restriction sites, 12901–13000: HpyF10VI, MwoI, BsmFI, TspRI, CviJI, HpyF10VI, MwoI, BsrI, BsiHKAI, Bsp1286I, AluI, CviJI, MspA1I, PvuII, AlwNI, AluI, CviJI, MnlI, MaeIII, BceAI, HpaII, AciI, BceAI, HaeII, HhaI, AfeI, HinP1I

Restriction sites, 13001–13100: MwoI, TspRI, HpyF10VI, MwoI, BtsI, NlaIII, FatI, Hpy8I, PstI, HpyCH4V, SfcI, AciI, RsaI, Csp6I, HpaII, NciI, ScrFI, StyD4I, SfaNI, NlaIII, FatI, CviJI, HaeIII, MslI, MscI, MmeI, EaeI

Restriction sites, 13101–13200: BstF5I, NciI, ScrFI, HpaII, StyD4I, TspDTI, HpyF10VI, AciI, Cac8I, CviJI, FokI, HpyF10VI, MwoI, FauI, BslI, Hin4I, CviJI, HaeIII, TaiI, MaeIII, Tsp45I, HpyCH4IV, MnlI, MseI, BseMII, BstEII, MaeIII, BspCNI, TauI, AciI, Fnu4HI, CviJI, HaeIII, Hin4I, DdeI, NciI, ScrFI, HpaII, CviJI, StyD4I, HpyF10VI, HhaI, BstUI, MnlI, HinP1I, MwoI

Next section (continued): DrdI, AatII, CviJI, Cac8I

```
               TaiI |                                                            HaeII| |
              BtsI | |                                                           HhaI|| |
             TspRI | |                                                           AfeI||| |
BsmAI
              ZraI | |                                                       HinP1I|||| |
HpaII|
            BsaHI| | |     AciI                                              Cac8I ||||| |
Hpy188III|
          HpyCH4IV| | |   BstUI                  CviJI                       HhaI | ||||| |      BstUI
BsaWI||
        MaeIII   || | |   HhaI                  TspRI|                   BstUI | | ||||| |  AflIII |
BspEI||
        Tsp45I   || | | HinP1I |        HpyCH4III   ||        CviJI  HinP1I | | ||||| |     MluI |HpyCH4III
|||
               |  || | |     | |             |   ||          |      | | | ||||| |       | |         |
|||
      gcagtgacgtcatcgtctgcgcggaaatggacgaacagtggggctatgtcggggctaaatcgcgccagcgctggctgttttacgcgtatgacagtctc
cg
  13201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 13300
      cgtcactgcagtagcagacgcgcctttacctgcttgtcaccccgatacagccccgatttagcgcggtcgcgaccgacaaaatgcgcatactgtcagag
gc

a      A  V  T  S  S  S  A  R  K  W  T  N  S  G  A  M  S  G  L  N  R  A  S  A  G  C  F  T  R  M  T  V  S
G -
b       Q  *  R  H  R  L  R  G  N  G  R  T  V  G  L  C  R  G  *  I  A  P  A  L  A  V  L  R  V  *  Q  S  P
-
c        S  D  V  I  V  C  A  E  M  D  E  Q  W  G  Y  V  G  A  K  S  R  Q  R  W  L  F  Y  A  Y  D  S  L
R  -
```

Figure 19-RR

```
                                  TaiI
                               BsaAI |
                           HpyCH4IV| |
                              HhaI || |
                      BbsI   FspI| || |                   BseYI              MaeIII
BstF5I                                                  Hpy99I              Tsp45I
      HpyCH4III|HinP1I|| || |                                           Cac8I |        TaiI
|                                       HgaI    | BsaHI     CviJI |  |     BmgBI |        HphI
|        BslI  ||MboII||| || |  Hpy8I  HphI  |  | HgaI     HphI  |  |  HpyCH4IV| |       BccI |
|           |  ||   |||| || |     |       |  |  |    |         || |       || |            | |
      gaagacggttgttgcgcacgtattcggtgaacgcactatggcgacgctggggcgtcttatgagcctgctgtcaccctttgacgtggtgatatggatga
cg
 13301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 13400
      cttctgccaacaacgcgtgcataagccacttgcgtgataccgctgcgaccccgcagaatactcggacgacagtgggaaactgcaccactatacctact
gc
a       R  R  L  L  R  T  Y  S  V  N  A  L  W  R  R  W  G  V  L  *  A  C  C  H  P  L  T  W  *  Y  G  *
R -
b      E  D  G  C  C  A  R  I  R  *  T  H  Y  G  D  A  G  A  S  Y  E  P  A  V  T  L  *  R  G  D  M  D  D
G -
c       K  T  V  V  A  H  V  F  G  E  R  T  M  A  T  L  G  R  L  M  S  L  L  S  P  F  D  V  V  I  W  M  T
-

                                                                                                   Fnu4
HI
                                                                                            Hpy188I
|                                                                                       HinfI   |
|                                                                                        TfiI   |
|                                                                                       MnlI|   |
|         FokI                                                                     BspCNI  ||   |
|       TspGWI                                                                    BseMII| ||   |
|       MspA1I|                                                                  TauI  || ||   |
|        TauI|                             Eco57I                           Fnu4HI |  || ||   |
|        AciI ||                          Eco57MI                            AciI| |  || ||   |
|      Fnu4HI ||                            TaiI |                           BbvI| |  || ||   |
|       CviJI| ||                         BsaAI | |                          BsrBI| |  || ||   |
|      HaeIII| ||                     HpyCH4IV| | |                    HpyF10VI || |  || ||   |
|       Cac8I || ||                    HpyCH4V || | |                 Tsp509I  | || |  || ||   |
|        EaeI || ||          TspDTI  Fnu4HI | || | |             Fnu4HI  |    | || |  || ||   |
|      BstF5I| || ||      BbvI     |  AluI| | || | |               TseI|  |    | || |  || ||   |
|        FokI| || || HinfI  BsgI| FauI| CviJI| | || | |         HpyF10VI ||  |    | || |  || ||   |
|      CviJI|| || ||  TfiI AciI||BslI|| TseI| | || | | Cac8I    MwoI| ||  |MwoI| || |  ||
||DdeI|TseI|
         ||| || ||    |    |||   |||    || | || | |   |      || ||  |   || || |  || ||  ||
||     gatggctggccgctgtatgaatcccgcctgaagggaaagctgcacgtaatcagcaagcgatatacgcagcgaattgagcggcataacctgaatctgag
gc
 13401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 13500
       ctaccgaccggcgacatacttagggcggacttccctttcgacgtgcattagtcgttcgctatatgcgtcgcttaactcgccgtattggacttagactc
cg
a        M  A  G  R  C  M  N  P  A  *  R  E  S  C  T  *  S  A  S  D  I  R  S  E  L  S  G  I  T  *  I  *  G
-
b        W  L  A  A  V  *  I  P  P  E  G  K  A  A  R  N  Q  Q  A  I  Y  A  A  N  *  A  A  *  P  E  S  E
A -
c       D  G  W  P  L  Y  E  S  R  L  K  G  K  L  H  V  I  S  K  R  Y  T  Q  R  I  E  R  H  N  L  N  L  R
Q -

                                                            Tth111I
                                                           NlaIII |
                                                          FatI  | |
      BstNI                                            HpyCH4V  | |
```

```
  ScrFI                                TspGWI               Fnu4HI |    |   |
   AlwNI|   BseYI                 BsmFI     |                 AluI|  |    |   |
  PspGI||   CviJI            BceAI     |    |                CviJI|  |    |   |                           MmeI
 StyD4I|| BbvI   |      PshAI     |    |    | BbvI             TseI|  |    |   |                   Hpy188I    |
       |||    |   |          |     |    |    |   |               ||  |    |   |                       |      |
       agcacctggcacggctgggacggaagtcgctgtcgttctcaaaatcggtggagctgcatgacaaagtcatcgggcattatctgaacataaaacactat
ca
 13501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------
-+ 13600
       tcgtggaccgtgccgaccctgccttcagcgacagcaagagttttagccacctcgacgtactgtttcagtagcccgtaatagacttgtattttgtgata
gt

a       S  T  W  H  G  W  D  G  S  R  C  R  S  Q  N  R  W  S  C  M  T  K  S  S  G  I  I  *  T  *  N  T  I
N -
b        A  P  G  T  A  G  T  E  V  A  V  V  L  K  I  G  G  A  A  *  Q  S  H  R  A  L  S  E  H  K  T  L  S
-
c         H  L  A  R  L  G  R  K  S  L  S  F  S  K  S  V  E  L  H  D  K  V  I  G  H  Y  L  N  I  K  H  Y
Q  -
```

Figure 19-SS

```
                                                  BstNI
                                                  ScrFI
              Tsp509I                            BsaJI|                 HpyCH4V
                MlyI |       ApoI      AluI     PspGI||                  NlaIII
        HinfI   PleI||    Tsp509I     CviJI    StyD4I||                FatI   |
            |       |||          |         |         |||                  |   |
      ataagttggagtcattacccaattatgatagaatttacaagctataaggttattgtcctgggtttcaagcattagtccatgcaagtttttatgctttg
cc
 13601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 13700
      tattcaacctcagtaatgggttaatactatcttaaatgttcgatattccaataacaggacccaaagttcgtaatcaggtacgttcaaaaatacgaaac
gg

a      K  L  E  S  L  P  N  Y  D  R  I  Y  K  L  *  G  Y  C  P  G  F  Q  A  L  V  H  A  S  F  Y  A  L
P  -
b     I  S  W  S  H  Y  P  I  M  I  E  F  T  S  Y  K  V  I  V  L  G  F  K  H  *  S  M  Q  V  F  M  L  C
P -
c      *  V  G  V  I  T  Q  L  *  *  N  L  Q  A  I  R  L  L  S  W  V  S  S  I  S  P  C  K  F  L  C  F  A
-

                        HpyF10VI
                    Fnu4HI     |
                      HhaI|    |
                      TseI|    |
                    BstUI ||   |
                    Cac8I ||   |
                     HhaI ||   |
                   HinP1I ||   |
                  BssHII | ||  |                                           BceAI
        SfcI    BbvI    HinP1I | || MwoI|                    MboII EcoRV  Hin4I |
          |        |         | | ||   ||                          |     |       ||
      cattctatagatatattgataagcgcgctgcctatgccttgcccctgaaatccttacatacggcgatatcttctatataaaagatatattatcttat
ca
 13701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 13800
      gtaagatatctatataactattcgcgcgacggatacggaacgggggactttaggaatgtatgccgctatagaagatatattttctatataatagaata
gt

a       I  L  *  I  Y  *  *  A  R  C  L  C  L  A  P  *  N  P  Y  I  R  R  Y  L  L  Y  K  R  Y  I  I  L  S
-
b        F  Y  R  Y  I  D  K  R  A  A  Y  A  L  P  P  E  I  L  T  Y  G  D  I  F  Y  I  K  D  I  L  S  Y
Q  -
c      H  S  I  D  I  L  I  S  A  L  P  M  P  C  P  L  K  S  L  H  T  A  I  S  S  I  *  K  I  Y  Y  L  I
S -

                                            MnlI
                                         EarI |
                         BstF5I   BstF5I |    |
                         MboII |    MnlI||    |                    TspDTI     HaeII
              BseRI     TspDTI |   MboII||    |        MnlI  AluI  DraI |      HhaI|
       Hin4I   FokI       FokI |  MnlI ||||   |      EarI |  CviJI MseI| |HinP1I ||     Tsp509I
           |       |        || |     | ||||   |         | |     |    || |     | ||           |
      gtattgtcaatatattcaaggcaatctgcctcctcatcctcttcatcctcttcgtcttggtagcttttaaatatggcgcttcatagagtaattctgt
aa
 13801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 13900
      cataacagttatataagttccgttagacggaggagtaggagaagtaggagaagcagaaccatcgaaaaatttataccgcgaagtatctcattaagaca
tt

a      V  L  S  I  Y  S  R  Q  S  A  S  S  S  S  S  S  S  S  S  W  *  L  F  K  Y  G  A  S  *  S  N  S  V
K -
b       Y  C  Q  Y  I  Q  G  N  L  P  P  H  P  L  H  P  L  R  L  G  S  F  L  N  M  A  L  H  R  V  I  L  *
-
c       I  V  N  I  F  K  A  I  C  L  L  I  L  F  I  L  F  V  L  V  A  F  *  I  W  R  F  I  E  *  F  C
K  -

                                                                                                  BstUI
                                                                                               AciI |
                                                                                           NlaIV  | |
        TspDTI                                                                            BanI |  | |
      Tsp509I|                                                                        BsiHKAI| |  | |
       AvaII ||                        HphI                    BseYI                 Bsp1286I| |  | |
     Sau96I  ||            BsaJI      MnlI |                 MnlI |             BssSI  BcgI| | |  | |
          |  ||                |         | |                    | |                 |     || | |  | |
      aggtccaattctcgttttcatacctcggtataatcttacctatcacctcaaatggttcgctgggtttatcgcacccccgaacacgagcacggcacccg
cg
 13901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14000
      tccaggttaagagcaaaagtatggagccatattagaatggatagtggagtttaccaagcgacccaaatagcgtgggggcttgtgctcgtgccgtgggc
gc

a       V  Q  F  S  F  S  Y  L  G  I  I  L  P  I  T  S  N  G  S  L  G  L  S  H  P  R  T  R  A  R  H  P
R  -
```

b R S N S R F H T S V * S Y L S P Q M V R W V Y R T P E H E H G T R
D -
c G P I L V F I P R Y N L T Y H L K W F A G F I A P P N T S T A P A
-

Figure 19-TT

```
AciI
                                                                                                          Fnu4
HI
                                                    FauI
CviJI|
                                                  NlaIII
HaeIII|
                                                FatI    |                                           HpyF10VI
||
                                            TspGWI |    |                                              BglI|
||
                                              AciI| |   |                                             BstNI|
||
                                          HpyF10VI| |   |                                              MwoI|
||
                                              MwoI|| |  |                                             ScrFI|
||
                                           NlaIV||| |   |                  BsiEI                   BsaJI ||
||
                                           CviJI|||| |  |                 CviJI|                   PspGI ||
||
                                          HaeIII|||| |  |                HaeIII|                  StyD4I ||
||
                                         Sau96I||||| |  |               Hpy99I||               TauI   | ||
||
                                          Cac8I|||||| | |                EaeI|||            AciI |    | ||
||
                        BcgI               NaeI|||||| | |                EagI|||           Fnu4HI |    | ||
||
                    BsaJI |              HpaII||||||| | |            HpyF10VI|||          CviJI| |    | ||
||
                     BsmI |             BsrFI|||||||| | |                BglI||||        HaeIII| |    | ||
||
     BceAI   HpyF10VI  | |            NgoMIV|||||||| | |           BsmI   |||||         Cac8I || |    | ||
||
    FauI |      MwoI|StyI |Tsp509I    ||||||||| |   |    TspDTI |MwoI||||         BceAI | || |    | ||
||
       | |        ||   | |      |     ||||||||| |   |         | |   |||||            | | || |    | ||
||
     accactatgccaagaatgcccaaggtaaaaattgccggccccgccatgaagtccgtgaatgccccgacggccgaagtgaagggcaggccgccacccag
gc
  14001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14100
     tggtgatacggttcttacgggttccatttttaacggccggggcggtacttcaggcacttacggggctgccggcttcacttcccgtccggcggtgggtc
cg

a      P  L  C  Q  E  C  P  R  *  K  L  P  A  P  P  *  S  P  *  M  P  R  R  P  K  *  R  A  G  R  H  P  G
-
b       H  Y  A  K  N  A  Q  G  K  N  C  R  P  R  H  E  V  R  E  C  P  D  G  R  S  E  G  Q  A  A  T  Q
A  -
c      T  T  M  P  R  M  P  K  V  K  I  A  G  P  A  M  K  S  V  N  A  P  T  A  E  V  K  G  R  P  P  P  R
P -


                        BstNI
                        ScrFI
                        BslI|
                       PspGI||
                       SexAI||
                      StyD4I||
                      NlaIV|||
                     BanI ||||
                    MnlI| ||||
                   HpaII|| ||||                                                             BstF5I
                   NciI|| ||||                                                         BstKTI   |
                  ScrFI|| ||||                                                          DpnI|   |
                  TspRI|| ||||                                        HgaI              MboI ||  |
                StyD4I ||| ||||                                      MslI |          CviJI  | ||  |
             HpyF10VI | ||| ||||                                    TaiI |          FokI |  | ||  |
                 MwoI| | ||| ||||                               BmgBI| |         AvaI|  |  | ||  |
                BtsI || | ||| ||||                               AarI| |         HhaI| |  |  | ||  |
              TaqII| || | ||| ||||                              BspMI| |        Hin4I| |  |  | ||  |
              TauI || || | ||| ||||                          HpyCH4IV| |        BstUI|| |  |  | ||  |
             AciI | || || | ||| ||||                             TauI || |       HinP1I|| |  |  | ||  |
           Fnu4HI | || || | ||| ||||                       Fnu4HI | || |        Hpy99I|| |  |  | ||  |
            TauI| | || || | ||| ||||                     HpyF10VI | || |     BsaHI  | || |  |  | ||  |
                || | || || | ||| ||||                          AciI| | || |     NciI |  | || |  |  | ||  |
                || | || || | ||| ||||                         MwoI| | || |    ScrFI |  | || |  |  | ||  |
                || | || || | ||| ||||                     AlwNI || | || |    HpaII| |  | || |  |  | ||  |
                || | || || | ||| |||| SfaNI   TaqI  Cac8I  |  || | || |   |StyD4I|| |  | || |  |  | ||  |
                || | || || | ||| ||||   |       |      |   |  || | || |   ||| | |  | || |  |  | ||  |
     cgccgccctcactgcccggcacctggtcgctgaatgtcgatgccagcacctgcggcacgtcaatgcttccgggcgtcgcgctcgggctgatcgcccat
cc
  14101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14200
     gcggcgggagtgacgggccgtggaccagcgacttacagctacggtcgtggacgccgtgcagttacgaaggcccgcagcgcgagcccgactagcgggta
gg
```

a
p - R R P H C P A P G R * M S M P A P A A R Q C F R A S R S G * S P I
b
- A A L T A R H L V A E C R C Q H L R H V N A S G R R A R A D R P S
c
p - P P S L P G T W S L N V D A S T C G T S M L P G V A L G L I A H

Figure 19-UU

```
                                                                                              BbvI
                                                   BceAI                                      MnlI
                                           Fnu4HI  |                                        BslI |
                                            HaeII  |                                       BccI | |
                                         HpyF10VI  |                        BsaJI        BstNI| | |
                                             HhaI| |                        BsiEI        ScrFI| | |
                                             MwoI| |                       CviJI|        AlwNI|| | |
                         HpaII               TseI| |                      HaeIII|        BsaJI|| | |
                          NciI               AfeI|| |                       TauI|        BsaJI||| | |
                         ScrFI             HinP1I||| |                     EaeI ||       PspGI||| | |
                         BslI|            Cac8I |||| |                     EagI ||      StyD4I||| | |
                      BstKTI||    HpyF10VI  |  |||| |                    Fnu4HI ||      CviJI |||| | |
                      StyD4I||      BstAPI  |  |||| |                     AciI| ||   Fnu4HI | |||| | |
                        DpnI|||     BbvI |  |  |||| |                    BstUI| ||     HhaI| | |||| | |
        BccI      Hin4I ||||     BsrDI|  || |  |||| |  BslI     CviJI    HphI|| ||     TseI|| | |||| | |
     MaeIII |AlwI MboI |||| BslI  ||MwoI| |  |||| |  MnlI    HaeIII    MnlI|| || HinP1I || | |||| | |
BstF5I
|   cgttactgccccgatcccggcaatggcaaggactgccagcgctgccattttggggtgaggccgttcgcggccgagggcgcagcccctgggggatg
gg
 14201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14300
    gcaatgacggggctagggccgttaccgttcctgacggtcgcgacggtaaaaccccactccggcaagcgccggctccccgcgtcggggaccccctac
cc

a      L  L  P  R  S  R  Q  W  Q  G  L  P  A  L  P  F  L  G  *  G  R  S  R  P  R  G  A  A  P  G  G  M
G  -
b     R  Y  C  P  D  P  G  N  G  K  D  C  Q  R  C  H  F  W  G  E  A  V  R  G  R  G  A  Q  P  L  G  G  W
E -
c      V  T  A  P  I  P  A  M  A  R  T  A  S  A  A  I  F  G  V  R  P  F  A  A  E  G  R  S  P  W  G  D  G
-


                                                                                         BstNI
                                                                                         ScrFI
                                                                                       BsaJI |
                                                                                       PspGI |
                                                                                      StyD4I |
                                                                                      CviJI| |
                                                                                   Fnu4HI || |
                   NciI                                                              HhaI| || |
                  ScrFI                                                              TseI| || |
                 HpaII|                                                  MaeIII     HinP1I || || |
      BstUI      CviJI||                                                 Tsp45I   HpyF10VI || || |
       FokI     HaeIII||                                                 AciI |      MwoI|| || || |
       FauI|     StyD4I||                                               BstUI |  HpyF10VI||| || || |
       AciI||    Cac8I |||                              Bme1580I          HhaI |    HhaI  |||| || || |
      Cac8I||     MnlI |||                              Bsp1286I      HinP1I  | BstUI  |||| || || |     BbvI
     CviJI |||   Sau96I |||                              NlaIV|   BslI Cac8I | |HinP1I |MwoI|||| || || |MseI |
     HaeIII|||    AciI |  |||  Hpy188III                  BanI ||          |  | |       |   |||| || || |     |
     Sau96I||| |  FauI|  |||       TaqI|
    aggcccgcgttagcgggccgggagggttcgagaaggggggcaccccccttcggcgtgcgcggtcacgcgcacagggcgcagccctggttaaaaacaa
gg
 14301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14400
    tccgggcgcaatcgcccggccctcccaagctcttccccccgtggggggaagccgcacgcgccagtgcgcgtgtcccgcgtcgggaccaatttttgtt
cc

a     G  P  R  *  R  A  G  R  V  R  E  G  G  A  P  P  F  G  V  R  G  H  A  H  R  A  Q  P  W  L  K  T  R
-
b      A  R  V  S  G  P  G  G  F  E  K  G  G  H  P  P  S  A  C  A  V  T  R  T  G  R  S  P  G  *  K  Q
G  -
c     R  P  A  L  A  G  R  E  G  S  R  R  G  G  T  P  L  R  R  A  R  S  R  A  Q  G  A  A  L  V  K  N  K
V -

                                        CviJI
                 DraI                   HaeIII
                 MseI|                   EaeI |                                   EciI
             BspMI ||                        |                                       |
    PsiI SspI    |  ||       MseI        AciI  | |   BslI AciI      HpyCH4V  |        Hin4I   Hpy8I
       |    |    |  ||         |            |  | |      |   |               |  |            |       |
    tttataaatattggtttaaaagcaggttaaaagacaggttagcggtggccgaaaaacgggcggaaacccttgcaaatgctggatttctgcctgtgga
ca
 14401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14500
    aaatatttataaccaaattttcgtccaattttctgtccaatcgccaccggctttttgcccgcctttgggaacgtttacgacctaaagacggacacct
gt

a     F  I  N  I  G  L  K  A  G  *  K  T  G  *  R  W  P  K  N  G  R  K  P  L  Q  M  L  D  F  L  P  V  D
S -
```

b L * I L V * K Q V K R Q V S G G R K T G G N P C K C W I F C L W T
-
c Y K Y W F K S R L K D R L A V A E K R A E T L A N A G F S A C G
Q -

Figure 19-VV

```
                                                                                AlwI
                                                                              HhaI|
                                                                            BstUI ||
                                                                          HinP1I  ||
                                                                        BstKTI |  ||
                              HhaI                                       DpnI| |  ||
                          HinP1I |            MnlI                      MboI || |  ||                 MnlI  BstUI  |    HhaI
     CviJI            MnlI Hin4I | |       BpuEI |             SmlI     MnlI |  || |  ||              BpuEI|  HinP1I  |SmlI
       |                 |    |  | |           | |               |        |   |  || |  ||                 ||       |  |    |
       gcccctcaaatgtcaataggtgcgcccctcatctgtcagcactctgcccctcaagtgtcaaggatcgcgcccctcatctgtcagtagtcgcgcccctc
aa
   14501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14600
       cggggagtttacagttatccacgcggggagtagacagtcgtgagacggggagttcacagttcctagcgcggggagtagacagtcatcagcgcggggag
tt

a         P  S  N  V  N  R  C  A  P  H  L  S  A  L  C  P  S  S  V  K  D  R  A  P  H  L  S  V  V  A  P  L
K -
b       A  P  Q  M  S  I  G  A  P  L  I  C  Q  H  S  A  P  Q  V  S  R  I  A  P  L  I  C  Q  *  S  R  P  S
S -
c        P  L  K  C  Q  *  V  R  P  S  S  V  S  T  L  P  L  K  C  Q  G  S  R  P  S  S  V  S  S  R  A  P  Q
-

                                                                                                     Al
uI
                                                                                                    Cvi
JI
                                                                                                  Cac8I
|
                                                                                                CviJI |
|
                                                                                              HaeIII |
|
                         CviJI                                                                  MscI |
|
                       BstNI |                                                           Cac8I | |
|
                       ScrFI |                                                            EaeI | |
|
                     BsaJI | |                                                        HpyF10VI | |
|
         Bme1580I  PspGI | |                                                              MwoI| | |
|
   MnlI AciI Bsp1286I StyD4I | | Hpy8I  XcmI         BstUI                     MnlI      CviJI|| | |
|
      |   |        |      | | |     |      |              |                       |          ||| | |
       gtgtcaataccgcagggcacttatccccaggcttgtccacatcatctgtgggaaactcgcgtaaaatcaggcgttttcgccgatttgcgaggctggcc
ag
   14601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14700
       cacagttatggcgtcccgtgaataggggtccgaacaggtgtagtagacaccctttgagcgcattttagtccgcaaaagcggctaaacgctccgaccgg
tc

a        C  Q  Y  R  R  A  L  I  P  R  L  V  H  I  I  C  G  K  L  A  *  N  Q  A  F  S  P  I  C  E  A  G  Q
-
b         V  N  T  A  G  H  L  S  P  G  L  S  T  S  S  V  G  N  S  R  K  I  R  R  F  R  R  F  A  R  L  A
S -
c       V  S  I  P  Q  G  T  Y  P  Q  A  C  P  H  H  L  W  E  T  R  V  K  S  G  V  F  A  D  L  R  G  W  P
A -

              BsiEI                                  BpuEI
             CviJI|                                   NciI
            HaeIII|                                  ScrFI
            Cac8I ||                                HpaII|
             EaeI ||                                HhaI||
             EagI ||                              StyD4I||
             NaeI ||                              BstUI |||          SmlI
           HpaII| ||                             HinP1I |||          HphI|
           BsrFI|| ||                              TauI |||          MlyI|
          Hpy99I|| ||                            AciI | |||          PleI||               AciI
          NgoMIV|| ||                          Fnu4HI | |||        NlaIV|||              TaiI|
            TaiI ||| ||              Cac8I                           CviJI||||      HpyCH4IV  ||
         BmgBI | ||| ||            CviJI |      HincII | | |||       HaeIII||||       HincII |  ||
     HpyCH4IV| | ||| ||          TaqI  | |      Hpy8I  | | |||       Sau96I|||||     Hpy8I  |  ||                MnlI
             || |  ||| ||          |   | |      MnlI   | | |||HinfI     |||||EciI MnlI |  ||             MnlI |
             || |  ||| ||          |   | |        |    | | |||  |       |||||  |    |  |  ||                |  |
       ctccacgtcgccggccgaaatcgagcctgcccctcatctgtcaacgccgcgccgggtgagtcggcccctcaagtgtcaacgtccgcccctcatctgtc
ag
   14701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14800
```

gaggtgcagcggccggctttagctcggacggggagtagacagttgcggcgcggcccactcagccggggagttcacagttgcaggcggggagtagacag
tc

a L H V A G R N R A C P S S V N A A P G E S A P Q V S T S A P H L S
V -
b S T S P A E I E P A P H L S T P R R V S R P L K C Q R P P L I C Q
-
c P R R R P K S S L P L I C Q R R A G * V G P S S V N V R P S S V
S -

Figure 19-WW

```
                                            SacII
                                             AciI|
                                            BstUI|
                                           MspA1I|
                                             TauI|
                                          AciI ||
                                         BsaJI ||
                                         BsiEI ||
                                          BtgI ||
                                        Fnu4HI ||
                                        CviJI| ||
                                       HaeIII| ||
                                         TauI| ||
                                       EaeI || ||
                                       EagI || ||
                                     Fnu4HI || ||
                                       NotI || ||                                            CviJI
                                      AciI| || ||                                           HaeIII
                                   Cac8I || || ||                                         Sau96I |
                                    NaeI || || ||                                      HpyCH4V | |
                                  HpaII| || || ||                                    BceAI   | | |
                                  BciVI|| || || ||                                  BstUI |  | | |
        CviJI                     BsrFI|| || || ||          Hpy99I                  HhaI |  | | |
       HaeIII        BstUI       NgoMIV|| || || ||        HpyF10VI|      HinP1I  |  |  | | |
      Sau96I |        AciI |                              MwoI||HpyF10VI |  |  |  | | |
EaeI                                                      TaqI |||  MwoI| |  |  |  | | |
      TspRI |     MnlI | |       SgrAI|| || || ||   BsmAI  CviJI  |  |||BceAI || | | |  | | |
EagI
|          | |       | | |         ||| || || ||       |     |  |  |||    | || | | |  | | |
           tgagggccaagttttccgcgaggtatccacaacgccggcggccgcggtgtctcgcacacggcttcgacggcgtttctggcgcgtttgcagggccatag
ac
   14801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+--------
-+ 14900
           actcccggttcaaaaggcgctccataggtgttgcggccgccggcgccacagagcgtgtgccgaagctgccgcaaagaccgcgcaaacgtcccggtatc
tg

a            R  A  K  F  S  A  R  Y  P  Q  R  R  R  P  R  C  L  A  H  G  F  D  G  V  S  G  A  F  A  G  P  *
T -
b           *  G  P  S  F  P  R  G  I  H  N  A  G  G  R  G  V  S  H  T  A  S  T  A  F  L  A  R  L  Q  G  H  R
R -
c            E  G  Q  V  F  R  E  V  S  T  T  P  A  A  A  V  S  R  T  R  L  R  R  R  F  W  R  V  C  R  A  I  D
-

                HpyF10VI
                    MwoI|
                 TauI  ||
             BceAI |   ||
            Fnu4HI |   ||
              AciI||   ||
            MspA1I||   ||
         HpyF10VI ||   ||
             MnlI ||   ||
            MwoI| ||   ||
           BseYI|| ||   ||
           CviJI||| ||   ||
          Cac8I |||| ||   ||
          TauI | |||| ||   ||
        AciI | | |||| ||   ||          HpaII
       BsiEI | | |||| ||   ||           NciI
      Fnu4HI | | |||| ||   ||          ScrFI
      CviJI| | | |||| ||   ||        CviJI |
     HaeIII| | | |||| ||   ||       StyD4I |
          || | | |||| ||   ||            | |
          ggccgccagcccagcggcgagggcaaccagcccgg
   14901 ---------+---------+---------+----- 14935
          ccggcggtcgggtcgccgctcccgttggtcgggcc

a          A  A  S  P  A  A  R  A  T  S  P   -
b           P  P  A  Q  R  R  G  Q  P  A  R  -
c         G  R  Q  P  S  G  E  G  N  Q  P    -
```

Figure 19-XX

Enzymes that do cut and were not excluded:

AarI AatII AccI AciI AclI AfeI AflII AflIII AgeI AhdI AleI
AloI AluI
AlwI AlwNI ApaI ApaLI ApoI AseI AvaI AvaII BaeI BamHI BanI
BanII BbeI
BbsI BbvI BbvCI BccI BceAI BcgI BciVI BclI BfaI BfrBI BglI
BglII BlpI
Bme1580I BmgBI BmrI BmtI BplI BpmI Bpu10I BpuEI BsaI BsaAI BsaBI
BsaHI BsaJI
BsaWI BsaXI BseMII BseRI BseYI BsgI BsiEI BsiHKAI BslI BsmI BsmAI
BsmBI BsmFI
Bsp1286I BspCNI BspEI BspHI BspMI BsrI BsrBI BsrDI BsrFI BssHII BssSI
BstAPI BstBI
BstEII BstF5I BstKTI BstNI BstUI BstXI BstYI Bsu36I BtgI BtsI Cac8I
ClaI Csp6I
CviJI DdeI DpnI DraI DraIII DrdI EaeI EagI EarI EciI Eco57I
Eco57MI EcoNI
EcoO109I EcoRI EcoRV FalI FatI FauI Fnu4HI FokI FspI FspAI HaeII
HaeIII HgaI
HhaI Hin4I HinP1I HincII HindIII HinfI HpaI HpaII HphI Hpy8I Hpy99I
Hpy188I Hpy188III
HpyCH4III HpyCH4IV HpyCH4V HpyF10VI KasI MaeIII MboI MboII MfeI MluI MlyI
MmeI MnlI
MscI MseI MslI MspA1I MwoI NaeI NarI NciI NcoI NdeI NgoMIV
NheI NlaIII
NlaIV NotI NruI NsiI NspI PciI PflMI PfoI PleI PmeI PpiI
PpuMI PshAI
PsiI PspGI PspOMI PsrI PstI PvuI PvuII RsaI RsrII SacII SalI
SanDI SapI
Sau96I SbfI ScaI ScrFI SexAI SfaNI SfcI SfiI SfoI SgrAI SmlI
SpeI SphI
SspI StyI StyD4I TaiI TaqI TaqII TatI TauI TfiI TseI Tsp45I
Tsp509I TspDTI
TspGWI TspRI Tth111I XbaI XcmI XhoI XmnI ZraI

Enzymes that do not cut:

Acc65I AscI AsiSI AvrII BsiWI BsrGI BstZ17I EcoICRI FseI KpnI PacI
PmlI SacI
SmaI SnaBI SrfI StuI SwaI XmaI

Enzymes excluded; MinCuts: 1 MaxCuts: 100000

NONE

Figure 20-A

```
tgagcgtcgc aaaggcgctc ggtcttgcct tgctcgtcgg tgatgtactt caccagctcc 60
gcgaagtcgc tcttcttgat ggagcgcatg ggacgtgct tggcaatcac gcgcaccccc 120
cggccgtttt agcggctaaa aaagtcatgg ctctgccctc gggcggacca cgcccatcat 180
gaccttgcca agctcgtcct gcttctcttc gatcttcgcc agcagggcga ggatcgtggc 240
atcaccgaac cgcgccgtgc gcgggtcgtc ggtgagccag agtttcagca ggccgcccag 300
gcggcccagg tcgccattga tgcgggccag ctcgcggacg tgctcatagt ccacgacgcc 360
cgtgattttg tagccctggc cgacggccag caggtaggcc gacaggctca tgccggccgc 420
cgccgccttt tcctcaatcg ctcttcgttc gtctggaagg cagtacacct tgataggtgg 480
gctgcccttc ctggttggct tggtttcatc agccatccgc ttgccctcat ctgttacgcc 540
ggcggtagcc ggccagcctc gcagagcagg attcccgttg agcaccgcca ggtgcgaata 600
agggacagtg aagaaggaac acccgctcgc gggtgggcct acttcaccta tcctgcccgg 660
ctgacgccgt tggatacacc aaggaaagtc tacacgaacc ctttggcaaa atcctgtata 720
tcgtgcgaaa aaggatggat ataccgaaaa aatcgctata atgaccccga agcagggtta 780
tgcagcggaa aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg 840
gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt 900
atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag 960
gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt 1020
gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta 1080
ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt 1140
cagtgagcga ggaagcggaa gagcgccaga aggccgccag agaggccgag cgcggccgtg 1200
aggcttggac gctagggcag ggcatgaaaa agcccgtagc gggctgctac gggcgtctga 1260
cgcggtggaa agggggaggg gatgttgtct acatggctct gctgtagtga gtgggttgcg 1320
ctccggcagc ggtcctgatc aatcgtcacc ctttctcggt ccttcaacgt tcctgacaac 1380
gagcctcctt ttcgccaatc catcgacaat caccgcgagt ccctgctcga acgctgcgtc 1440
cggaccggct tcgtcgaagg cgtctatcgc ggcccgcaac agcggcgaga gcggagcctg 1500
ttcaacggtg ccgccgcgct cgccggcatc gctgtcgccg gcctgctcct caagcacggc 1560
cccaacagtg aagtagctga ttgtcatcag cgcattgacg gcgtccccgg ccgaaaaacc 1620
cgcctcgcag aggaagcgaa gctgcgcgtc ggccgtttcc atctgcggtg cgcccggtcg 1680
cgtgccggca tggatgcgcg cgccatcgcg gtaggcgagc agcgcctgcc tgaagctgcg 1740
ggcattcccg atcagaaatg agcgccagtc gtcgtcggct ctcggcaccg aatgcgtatg 1800
attctccgcc agcatggctt cggccagtgc gtcgagcagc gcccgcttgt tcctgaagtg 1860
ccagtaaagc gccggctgct gaacccccaa ccgttccgcc agtttgcgtg tcgtcagacc 1920
gtctacgccg acctcgttca acaggtccag ggcggcacgg atcactgtat tcggctgcaa 1980
ctttgtcatg cttgacactt tatcactgat aaacataata tgtccaccaa cttatcagtg 2040
ataaagaatc cgcgcgttca atcggaccag cggaggctgg tccggaggcc agacgtgaaa 2100
cccaacatac ccctgatcgt aattctgagc actgtcgcgc tcgacgctgt cggcatcggc 2160
ctgattatgc cggtgctgcc gggcctcctg cgcgatctgg ttcactcgaa cgacgtcacc 2220
gcccactatg gcattctgct ggcgctgtat gcgttggtgc aatttgcctg cgcacctgtg 2280
ctgggcgcgc tgtcggatcg tttcgggcgg cggccaatct tgctcgtctc gctggccggc 2340
gccagatctg gggaaccctg tggttggcat gcacatacaa atggacgaac ggataaacct 2400
tttcacgccc ttttaaatat ccgattattc taataaacgc tcttttctct taggtttacc 2460
cgccaatata tcctgtcaaa cactgatagt ttaaactgaa ggcgggaaac gacaatctga 2520
tcatgagcgg agaattaagg gagtcacgtt atgacccccg ccgatgacgc gggacaagcc 2580
gttttacgtt tggaactgac agaaccgcaa cgttgaagga gccactcagc cgcgggtttc 2640
tggagtttaa tgagctaagc acatacgtca gaaaccatta ttgcgcgttc aaaagtcgcc 2700
taaggtcact atcagctagc aaatatttct tgtcaaaaat gctccactga cgttccataa 2760
attcccctcg gtatccaatt agagtctcat attcactctc aatccaaata atctgcaccg 2820
gatctggatc gtttcgcatg attgaacaag atggattgca cgcaggttct ccggccgctt 2880
gggtggagag gctattcggc tatgactggg cacaacagac aatcggctgc tctgatgccg 2940
ccgtgttccg gctgtcagcg caggggcgcc cggttctttt tgtcaagacc gacctgtccg 3000
gtgccctgaa tgaactgcag gacgaggcag cgcggctatc gtggctggcc acgacgggcg 3060
```

Figure 20-B

```
ttccttgcgc agctgtgctc gacgttgtca ctgaagcggg aagggactgg ctgctattgg 3120
gcgaagtgcc ggggcaggat ctcctgtcat ctcaccttgc tcctgccgag aaagtatcca 3180
tcatggctga tgcaatgcgg cggctgcata cgcttgatcc ggctacctgc ccattcgacc 3240
accaagcgaa acatcgcatc gagcgagcac gtactcggat ggaagccggt cttgtcgatc 3300
aggatgatct ggacgaagag catcaggggc tcgcgccagc cgaactgttc gccaggctca 3360
aggcgcgcat gcccgacggc gatgatctcg tcgtgaccca tggcgatgcc tgcttgccga 3420
atatcatggt ggaaaatggc cgcttttctg gattcatcga ctgtggccgg ctgggtgtgg 3480
cggaccgcta tcaggacata gcgttggcta cccgtgatat tgctgaagag cttggcggcg 3540
aatgggctga ccgcttcctc gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg 3600
ccttctatcg ccttcttgac gagttcttct gagcgggact ctggggttcg aaatgaccga 3660
ccaagcgacg cccaacctgc catcacgaga tttcgattcc accgccgcct tctatgaaag 3720
gttgggcttc ggaatcgttt tccgggacgc cggctggatg atcctccagc gcggggatct 3780
catgctggag ttcttcgccc acgggatctc tgcggaacag gcggtcgaag gtgccgatat 3840
cattacgaca gcaacggccg acaagcacaa cgccacgatc ctgagcgaca atatgatcgg 3900
gcccggcgtc cacatcaacg gcgtcggcgg cgactgccca ggcaagaccg agatgcaccg 3960
cgatatcttg ctgcgttcgg atattttcgt ggagttcccg ccacagaccc ggatgatccc 4020
cgatcgttca aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc 4080
gatgattatc atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg 4140
catgacgtta tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata 4200
cgcgatagaa aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc 4260
tatgttacta gatcgggcct cctgtcaatg ctggcggcgg ctctggtggt ggttctggtg 4320
gcggctctga gggtggtggc tctgagggtg gcggttctga gggtggcggc tctgagggag 4380
gcggttccgg tggtggctct ggttccggtg attttgatta tgaaaagatg gcaaacgcta 4440
ataagggggc tatgaccgaa aatgccgatg aaaacgcgct acagtctgac gctaaaggca 4500
aacttgattc tgtcgctact gattacggtg ctgctatcga tggtttcatt ggtgacgttt 4560
ccggccttgc taatggtaat ggtgctactg gtgattttgc tggctctaat tcccaaatgg 4620
ctcaagtcgg tgacggtgat aattcacctt taatgaataa tttccgtcaa tatttacctt 4680
ccctccctca atcggttgaa tgtcgccctt ttgtctttgg cccaatacgc aaaccgcctc 4740
tccccgcgcg ttggccgatt cattaatgca gctggcacga caggtttccc gactggaaag 4800
cgggcagtga gcgcaacgca attaatgtga gttagctcac tcattaggca ccccaggctt 4860
tacactttat gcttccggct cgtatgttgt gtggaattgt gagcggataa caatttcaca 4920
caggaaacag ctatgaccat gattacgcca agcttgcatg cctgcaggtc cccagattag 4980
ccttttcaat ttcagaaaga atgctaaccc acagatggtt agagaggctt acgcagcagg 5040
tctcatcaag acgatctacc cgagcaataa tctccaggaa atcaaatacc ttcccaagaa 5100
ggttaaagat gcagtcaaaa gattcaggac taactgcatc aagaacacag agaaagatat 5160
atttctcaag atcagaagta ctattccagt atggacgatt caaggcttgc ttcacaaacc 5220
aaggcaagta atagagattg gagtctctaa aaaggtagtt cccactgaat caaaggccat 5280
ggagtcaaag attcaaatag aggacctaac agaactcgcc gtaaagactg gcgaacagtt 5340
catacagagt ctcttacgac tcaatgacaa gaagaaaatc ttcgtcaaca tggtggagca 5400
cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat 5460
tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gcccagctat 5520
ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg 5580
cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggtccca aagatggacc 5640
cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 5700
ggattgatgt gatatctcca ctgacgtaag ggatgacgca caatcccact atccttcgca 5760
agacccttcc tctatataag gaagttcatt tcatttggag agaacacggg ggactctaga 5820
ggatccatgt ttattttctt attatttctt actctcacta gtggtagtga ccttgaccgg 5880
tgcaccactt ttgatgatgt tcaagctcct aattacactc aacatacttc atctatgagg 5940
ggggtttact atcctgatga aatttttaga tcagacactc tttatttaac tcaggattta 6000
tttcttccat tttattctaa tgttacaggg tttcatacta ttaatcatac gtttggcaac 6060
cctgtcatac cttttaagga tggtatttat tttgctgcca cagagaaatc aaatgttgtc 6120
cgtggttggg tttttggttc taccatgaac aacaagtcac agtcggtgat tattattaac 6180
aattctacta atgttgttat acgagcatgt aactttgaat tgtgtgacaa ccctttcttt 6240
```

Figure 20-C

```
gctgtttcta aacccatggg tacacagaca catactatga tattcgataa tgcatttaat 6300
tgcactttcg agtacatatc tgatgccttt tcgcttgatg tttcagaaaa gtcaggtaat 6360
tttaaacact tacgagagtt tgtgtttaaa aataaagatg ggtttctcta tgtttataag 6420
ggctatcaac ctatagatgt agttcgtgat ctaccttctg gttttaacac tttgaaacct 6480
atttttaagt tgcctcttgg tattaacatt acaaatttta gagccattct tacagccttt 6540
tcacctgctc aagacatttg ggcacgtca gctgcagcct attttgttgg ctatttaaag 6600
ccaactacat ttatgctcaa gtatgatgaa aatggtacaa tcacagatgc tgttgattgt 6660
tctcaaaatc cacttgctga actcaaatgc tctgttaaga gctttgagat tgacaaagga 6720
atttaccaga cctctaattt cagggttgtt ccctcaggag atgttgtgag attccctaat 6780
attacaaact tgtgtccttt tggagaggtt tttaatgcta ctaaattccc ttctgtctat 6840
gcatgggaga gaaaaaaaat ttctaattgt gttgctgatt actctgtgct ctacaactca 6900
acattttttt caacctttaa gtgctatggc gtttctgcca ctaagttgaa tgatctttgc 6960
ttctccaatg tctatgcaga ttcttttgta gtcaagggag atgatgtaag acaaatagcg 7020
ccaggacaaa ctggtgttat tgctgattat aattataaat tgccagatga tttcatgggt 7080
tgtgtccttg cttggaatac taggaacatt gatgctactt caactggtaa ttataattat 7140
aaatataggt atcttagaca tggcaagctt aggccctttg agagagacat atctaatgtg 7200
cctttctccc ctgatggcaa accttgcacc ccacctgctc ttaattgtta ttggccatta 7260
aatgattatg gtttttacac cactactggc attggctacc aaccttacag agttgtagta 7320
ctttcttttg aacttttaaa tgcaccggcc acggtttgtg gaccaaaatt atccactgac 7380
cttattaaga accagtgtgt caattttaat tttaatggac tcactggtac tggtgtgtta 7440
actccttctt caaagagatt tcaaccattt caacaatttg gccgtgatgt ttctgatttc 7500
actgattccg ttcgagatcc taaaacatct gaaatattag acatttcacc ttgctctttt 7560
gggggtgtaa gtgtaattac acctggaaca aatgcttcat ctgaagttgc tgttctatat 7620
caagatgtta actgcactga tgtttctaca gcaattcatg cagatcaact cacaccagct 7680
tggcgcatat attctactgg aaacaatgta ttccagactc aagcaggctg tcttatagga 7740
gctgagcatg tcgacacttc ttatgagtgc gacattccta ttggagctgg catttgtgct 7800
attaccatac agttagacta gaaaagggcg aattctgcag atatccatca cactggcggc 7860
cgctcgacct cgaggggggg cccggtacag ctcgaatttc cccgatcgtt caaacatttg 7920
gcaataaagt ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt 7980
tctgttgaat tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag 8040
atgggttttt atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat 8100
atagcgcgca aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcggga 8160
attcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcgtt acccaactta 8220
atcgccttgc agcacatccc cctttcgcca gctggcgtaa tagcgaagag gcccgcaccg 8280
atcgcccttc ccaacagttg cgcagcctga atggcgcccg ctcctttcgc tttcttccct 8340
tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctcccttta 8400
gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgattt gggtgatggt 8460
tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg 8520
ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat ctcgggctat 8580
tcttttgatt tataagggat tttgccgatt tcggaaccac catcaaacag gattttcgcc 8640
tgctggggca aaccagcgtg gaccgcttgc tgcaactctc tcagggccag gcggtgaagg 8700
gcaatcagct gttgcccgtc tcactggtga aaagaaaaac caccccagta cattaaaaac 8760
gtccgcaatg tgttattaag ttgtctaagc gtcaatttgt ttacaccaca atatatcctg 8820
ccaccagcca gccaacagct ccccgaccgg cagctcggca caaaatcacc actcgataca 8880
ggcagcccat cagtccggga cggcgtcagc gggagagccg ttgtaaggcg gcagactttg 8940
ctcatgttac cgatgctatt cggaagaacg gcaactaagc tgccgggttt gaaacacgga 9000
tgatctcgcg gagggtagca tgttgattgt aacgatgaca gagcgttgct gcctgtgatc 9060
aaatatcatc tccctcgcag agatccgaat tatcagcctt cttattcatt tctcgcttaa 9120
ccgtgacagg ctgtcgatct tgagaactat gccgacataa taggaaatcg ctggataaag 9180
ccgctgagga agctgagtgg cgctatttct ttagaagtga acgttgacga tatcaactcc 9240
cctatccatt gctcaccgaa tggtacaggt cggggacccg aagttccgac tgtcggcctg 9300
atgcatcccc ggctgatcga ccccagatct ggggctgaga aagcccagta aggaaacaac 9360
```

Figure 20-D

tgtaggttcg agtcgcgaga tccccccggaa ccaaaggaag taggttaaac ccgctccgat 9420
caggccgagc cacgccaggc cgagaacatt ggttcctgta ggcatcggga ttggcggatc 9480
aaacactaaa gctactggaa cgagcagaag tcctccggcc gccagttgcc aggcggtaaa 9540
ggtgagcaga ggcacgggag gttgccactt gcgggtcagc acggttccga acgccatgga 9600
aaccgccccc gccaggcccg ctgcgacgcc gacaggatct agcgctgcgt ttggtgtcaa 9660
caccaacagc gccacgcccg cagttccgca aatagccccc aggaccgcca tcaatcgtat 9720
cgggctacct agcagagcgg cagagatgaa cacgaccatc agcggctgca cagcgcctac 9780
cgtcgccgcg accccgcccg gcaggcggta gaccgaaata aacaacaagc tccagaatag 9840
cgaaatatta agtgcgccga ggatgaagat gcgcatccac cagattcccg ttggaatctg 9900
tcggacgatc atcacgagca ataaacccgc cggcaacgcc cgcagcagca taccggcgac 9960
ccctcggcct cgctgttcgg gctccacgaa aacgccggac agatgcgcct tgtgagcgtc 10020
cttggggccg tcctcctgtt tgaagaccga cagcccaatg atctcgccgt cgatgtaggc 10080
gccgaatgcc acggcatctc gcaaccgttc agcgaacgcc tccatgggct tttctcctc 10140
gtgctcgtaa acggacccga acatctctgg agctttcttc agggccgaca atcggatctc 10200
gcggaaatcc tgcacgtcgg ccgctccaag ccgtcgaatc tgagccttaa tcacaattgt 10260
caattttaat cctctgttta tcggcagttc gtagagcgcg ccgtgcgtcc cgagcgatac 10320
tgagcgaagc aagtgcgtcg agcagtgccc gcttgttcct gaaatgccag taaagcgctg 10380
gctgctgaac cccagccgg aactgacccc acaaggccct agcgtttgca atgcaccagg 10440
tcatcattga cccaggcgtg ttccaccagg ccgctgcctc gcaactcttc gcaggcttcg 10500
ccgacctgct cgcgccactt cttcacgcgg gtggaatccg atccgcacat gaggcggaag 10560
gtttccagct tgagcgggta cggctcccgg tgcgagctga aatagtcgaa catccgtcgg 10620
gccgtcggcg acagcttgcg gtacttctcc catatgaatt tcgtgtagtg gtcgccagca 10680
aacagcacga cgatttcctc gtcgatcagg acctggcaac gggacgtttt cttgccacgg 10740
tccaggacgc ggaagcggtg cagcagcgac accgattcca ggtgcccaac gcggtcggac 10800
gtgaagccca tcgccgtcgc ctgtaggcgc gacaggcatt cctcggcctt cgtgtaatac 10860
cggccattga tcgaccagcc caggtcctgg caaagctcgt agaacgtgaa ggtgatcggc 10920
tcgccgatag gggtgcgctt cgcgtactcc aacacctgct gccacaccag ttcgtcatcg 10980
tcggccccgca gctcgacgcc ggtgtaggtg atcttcacgt ccttgttgac gtggaaaatg 11040
accttgtttt gcagcgcctc gcgcgggatt ttcttgttgc gcgtggtgaa cagggcagag 11100
cgggccgtgt cgtttggcat cgctcgcatc gtgtccggcc acggcgcaat atcgaacaag 11160
gaaagctgca tttccttgat ctgctgcttc gtgtgtttca gcaacgcggc ctgcttggcc 11220
tcgctgacct gttttgccag gtcctcgccg gcggtttttc gcttcttggt cgtcatagtt 11280
cctcgcgtgt cgatggtcat cgacttcgcc aaacctgccg cctcctgttc gagacgacgc 11340
gaacgctcca cggcggccga tggcgcgggc agggcagggg gagccagttg cacgctgtcg 11400
cgctcgatct tggccgtagc ttgctggacc atcgagccga cggactggaa ggtttcgcgg 11460
ggcgcacgca tgacggtgcg gcttgcgatg gtttcggcat cctcggcgga aaaccccgcg 11520
tcgatcagtt cttgcctgta tgccttccgg tcaaacgtcc gattcattca ccctccttgc 11580
gggattgccc cgactcacgc cggggcaatg tgcccttatt cctgatttga cccgcctggt 11640
gccttggtgt ccagataatc caccttatcg gcaatgaagt cggtcccgta gaccgtctgg 11700
ccgtccttct cgtacttggt attccgaatc ttgccctgca cgaataccag cgacccttg 11760
cccaaatact tgccgtgggc ctcggcctga gagccaaaac acttgatgcg gaagaagtcg 11820
gtgcgctcct gcttgtcgcc ggcatcgttg cgccacatct aggtactaaa acaattcatc 11880
cagtaaaata taatatttta ttttctccca atcaggcttg atccccagta agtcaaaaaa 11940
tagctcgaca tactgttctt ccccgatatc ctccctgatc gaccggacgc agaaggcaat 12000
gtcataccac ttgtccgccc tgccgcttct cccaagatca ataaagccac ttactttgcc 12060
atctttcaca aagatgttgc tgtctcccag gtcgccgtgg gaaaagacaa gttcctcttc 12120
gggcttttcc gtctttaaaa aatcatacag ctcgcgcgga tctttaaatg gagtgtcttc 12180
ttcccagttt tcgcaatcca catcggccag atcgttattc agtaagtaat ccaattcggc 12240
taagcggctg tctaagctat tcgtataggg acaatccgat atgtcgatgg agtgaaagag 12300
cctgatgcac tccgcataca gctcgataat cttttcaggg ctttgttcat cttcatactc 12360
ttccgagcaa aggacgccat cggcctcact catgagcaga ttgctccagc catcatgccg 12420

Figure 20-E

```
ttcaaagtgc aggacctttg gaacaggcag ctttccttcc agccatagca tcatgtcctt 12480
ttcccgttcc acatcatagg tggtcccttt ataccggctg tccgtcattt ttaaatatag 12540
gttttcattt tctcccacca gcttatatac cttagcagga gacattcctt ccgtatcttt 12600
tacgcagcgg tatttttcga tcagtttttt caattccggt gatattctca ttttagccat 12660
ttattatttc cttcctcttt tctacagtat ttaaagatac cccaagaagc taattataac 12720
aagacgaact ccaattcact gttccttgca ttctaaaacc ttaaatacca gaaaacagct 12780
ttttcaaagt tgttttcaaa gttggcgtat aacatagtat cgacggagcc gattttgaaa 12840
ccacaattat gggtgatgct gccaacttac tgatttagtg tatgatggtg tttttgaggt 12900
gctccagtgg cttctgtgtc tatcagctgt ccctcctgtt cagctactga cggggtggtg 12960
cgtaacggca aaagcaccgc cggacatcag cgctatctct gctctcactg ccgtaaaaca 13020
tggcaactgc agttcactta caccgcttct caacccggta cgcaccagaa aatcattgat 13080
atggccatga atggcgttgg atgccgggca acagcccgca ttatgggcgt tggcctcaac 13140
acgattttac gtcacttaaa aaactcaggc cgcagtcggt aacctcgcgc atacagccgg 13200
gcagtgacgt catcgtctgc gcggaaatgg acgaacagtg gggctatgtc ggggctaaat 13260
cgcgccagcg ctggctgttt tacgcgtatg acagtctccg gaagacggtt gttgcgcacg 13320
tattcggtga acgcactatg gcgacgctgg ggcgtcttat gagcctgctg tcaccctttg 13380
acgtggtgat atggatgacg gatggctggc cgctgtatga atcccgcctg aagggaaagc 13440
tgcacgtaat cagcaagcga tatacgcagc gaattgagcg gcataacctg aatctgaggc 13500
agcacctggc acggctggga cggaagtcgc tgtcgttctc aaaatcggtg gagctgcatg 13560
acaaagtcat cgggcattat ctgaacataa aacactatca ataagttgga gtcattaccc 13620
aattatgata gaatttacaa gctataaggt tattgtcctg ggtttcaagc attagtccat 13680
gcaagttttt atgctttgcc cattctatag atatattgat aagcgcgctg cctatgcctt 13740
gccccctgaa atccttacat acggcgatat cttctatata aaagatatat tatcttatca 13800
gtattgtcaa tatattcaag gcaatctgcc tcctcatcct cttcatcctc ttcgtcttgg 13860
tagcttttta aatatggcgc ttcatagagt aattctgtaa aggtccaatt ctcgttttca 13920
tacctcggta taatcttacc tatcacctca aatggttcgc tgggtttatc gcacccccga 13980
acacgagcac ggcacccgcg accactatgc caagaatgcc caaggtaaaa attgccggcc 14040
ccgccatgaa gtccgtgaat gccccgacgg ccgaagtgaa gggcaggccg ccacccaggc 14100
cgccgccctc actgcccggc acctggtcgc tgaatgtcga tgccagcacc tgcggcacgt 14160
caatgcttcc gggcgtcgcg ctcgggctga tcgcccatcc cgttactgcc ccgatcccgg 14220
caatggcaag gactgccagc gctgccattt ttggggtgag gccgttcgcg gccgaggggc 14280
gcagcccctg gggggatggg aggcccgcgt tagcgggccg ggagggttcg agaagggggg 14340
gcaccccccct tcggcgtgcg cggtcacgcg cacagggcgc agccctggtt aaaaacaagg 14400
tttataaata ttggtttaaa agcaggttaa aagacaggtt agcggtggcc gaaaaacggg 14460
cggaaaccct tgcaaatgct ggattttctg cctgtggaca gcccctcaaa tgtcaatagg 14520
tgcgcccctc atctgtcagc actctgcccc tcaagtgtca aggatcgcgc ccctcatctg 14580
tcagtagtcg cgcccctcaa gtgtcaatac cgcagggcac ttatccccag gcttgtccac 14640
atcatctgtg ggaaactcgc gtaaaatcag gcgttttcgc cgatttgcga ggctggccag 14700
ctccacgtcg ccggccgaaa tcgagcctgc ccctcatctg tcaacgccgc gccgggtgag 14760
tcggcccctc aagtgtcaac gtccgcccct catctgtcag tgagggccaa gttttccgcg 14820
aggtatccac aacgccggcg gccgcggtgt ctcgcacacg gcttcgacgg cgtttctggc 14880
gcgtttgcag ggccatagac ggccgccagc ccagcggcga gggcaaccag cccgg      14935
```

Figure 21-A

```
                                              HpyCH4III
                                            HpaII
                                             NciI
                                            ScrFI
                                            PfoI
                                           StyD4I
                                           BsmAI
                                           BsmBI
                                            AluI
                                           CviJI
                                        Fnu4HI
                                       HpyCH4V
                                       NlaIII
                                          NspI
                                          TseI
                                         MnlI                              BstF5I
       BstUI                             FatI                                NciI
        HhaI                                        MaeIII                  ScrFI
      BstUI                 HphI      HphI          Tsp45I     SfaNI       HpaII        DrdI
     HinP1I      HpyCH4III       Hpy188I            BbvI   AluI           StyD4I        HgaI
                                                    BslI   CviJI     AciI          FokI CviJI
      tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccg
    1 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 100
      agcgcgcaaagccactactgccacttttggagactgtgtacgtcgagggcctctgccagtgtcgaacagacattcgcctacggccctcgtctgttcgggc

a      S  R  V  S  V  M  T  V  K  T  S  D  T  C  S  S  R  R  R  S  Q  L  V  C  K  R  M  P  G  A  D  K  P  V -
b       R  A  F  R  *  *  R  *  K  P  L  T  H  A  A  P  G  D  G  H  S  L  S  V  S  G  C  R  E  Q  T  S  P  -
c        A  R  F  G  D  D  G  E  N  L  *  H  M  Q  L  P  E  T  V  T  A  C  L  *  A  D  A  G  S  R  Q  A  R  -

                                                                                            AciI
                                                                                       HpyF10VI
                                                                                        BstAPI
                                                                                          MwoI
                                                            BspCNI                        NdeI
                                                           BseMII           Bme1580I
                                                            SfaNI           BsiHKAI
        BstUI                       MseI            Hpy188I         DdeI    Bsp1286I
         HhaI     FauI             CviJI         TauI              RsaI      Hpy8I
        FauI   AciI                Cac8I    Fnu4HI               Csp6I    HpyCH4V
     HinP1I  MspA1I     AciI      CviJI      AciI                TatI    ApaLI                     SfaNI
      tcagggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaata
  101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 200
      agtcccgcgcagtcgcccacaaccgcccacagccccgaccgaattgatacgccgtagtctcgtctaacatgactctcacgtggtatacgccacactttat

a       R  A  R  Q  R  V  L  A  G  V  G  A  G  L  T  M  R  H  Q  S  R  L  Y  *  E  C  T  I  C  G  V  K  Y  -
b      S  G  R  V  S  G  C  W  R  V  S  G  L  A  *  L  C  G  I  R  A  D  C  T  E  S  A  P  Y  A  V  *  N  T -
c       Q  G  A  S  A  G  V  G  G  C  R  G  W  L  N  Y  A  A  S  E  Q  I  V  L  R  V  H  H  M  R  C  E  I  -
```

Figure 21-B

```
AciI  AciI  HpyF10VI MwoI BbvI BbeI HaeII HhaI SfaNI NlaIV SfoI BsaHI HinP1I HpyF10VI NarI BanI KasI MwoI
HpyF10VI BglI MwoI  Fnu4HI CviJI TseI HinP1I FspI HhaI HpyCH4III
MboI FauI DpnI PvuI BstKTI BsiEI MwoI HpyF10VI MboII AciI Sau96I Cac8I HaeIII CviJI  EarI MnlI

    ccgcacagatgcgtaaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctat
201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 300
    ggcgtgtctacgcattcctcttttatggcgtagtccgcggtaagcggtaagtccgacgcgttgacaacccttcccgctagccacgcccggagaagcgata

a    R  T  D  A  *  G  E  N  T  A  S  G  A  I  R  H  S  G  C  A  T  V  G  K  G  D  R  C  G  P  L  R  Y  -
b     A  Q  M  R  K  E  K  I  P  H  Q  A  P  F  A  I  Q  A  A  Q  L  L  G  R  A  I  G  A  G  L  F  A  I  -
c    P  H  R  C  V  R  R  K  Y  R  I  R  R  H  S  F  F  R  L  R  N  C  W  E  G  R  S  V  R  A  S  S  L  L -

Cac8I AluI CviJI MspA1I PvuII Cac8I  BbvI  HpyCH4V Fnu4HI BstF5I TseI  FokI  MseI  MaeIII
BstNI ScrFI BsaJI PspGI StyD4I  BmrI  Hpy188III MaeIII Tsp45I BsrI  HpyCH4IV Hpy99I TaiI
BsrI CviJI HaeIII Hpy99I EaeI  HindIII TspRI

    tacgccagctggcgaaagggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgccaa
301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 400
    atgcggtcgaccgctttccccctacacgacgttccgctaattcaacccattgcggtcccaaaagggtcagtgctgcaacattttgctgccggtcacggtt

a    Y  A  S  W  R  K  G  D  V  L  Q  G  D  *  V  G  *  R  Q  G  F  P  S  H  D  V  V  K  R  R  P  V  P  S -
b     T  P  A  G  E  R  G  M  C  C  K  A  I  K  L  G  N  A  R  V  F  P  V  T  T  L  *  N  D  G  Q  C  Q  -
c      R  Q  L  A  K  G  G  C  A  A  R  R  L  S  W  V  T  P  G  F  S  Q  S  R  R  C  K  T  T  A  S  A  K  -

BceAI AluI CviJI  Tsp509I  BsmFI  BslI  TatI Csp6I RsaI BstF5I  MseI  FokI  Tsp509I  BslI
BceAI  HpyCH4III  CviJI

    gcttgaaagagataaattgaacaagtatggtcgtcccctgttgggatgtactattaaacctaaattggggttatctgctaaaaactacggtagagccgtt
401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 500
    cgaactttctctatttaacttgttcataccagcaggggacaaccctacatgataatttggatttaaccccaatagacgatttttgatgccatctcggcaa

a     L  K  E  I  N  *  T  S  M  V  V  P  C  W  D  V  L  L  N  L  N  W  G  Y  L  L  K  T  T  V  E  P  F  -
b    A  *  K  R  *  I  E  Q  V  W  S  S  P  V  G  M  Y  Y  *  T  *  I  G  V  I  C  *  K  L  R  *  S  R  L -
c     L  E  R  D  K  L  N  K  Y  G  R  P  L  L  G  C  T  I  K  P  K  L  G  L  S  A  K  N  Y  G  R  A  V  -
```

Figure 21-C

```
                       Hpy8I                          MmeI
                     TspDTI |                       Hpy8I|          PpiI      BstKTI
   BbsI             AciI | |                        TaiI||          XcmI|      DpnI|
  MboII    XmnI BstUI | |           PpiI  HpyCH4IV  |||         BsaXI ||     MboI ||           BsaXI
     |        |    |  | |              |         |  |||              | ||        | ||               |
     tatgaatgtcttcgcggtggacttgattttactaaagatgatgagaacgtgaactcacaaccatttatgcgttggagagatcgtttcttattttgtgccg
 501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 600
     atacttacagaagcgccacctgaactaaaatgatttctactactcttgcacttgagtgttggtaaatacgcaacctctctagcaaagaataaaacacggc

a     M  N  V  F  A  V  D  L  I  L  L  K  M  M  R  T  *  T  H  N  H  L  C  V  G  E  I  V  S  Y  F  V  P  -
b      *  M  S  S  R  W  T  *  F  Y  *  R  *  *  E  R  E  L  T  T  I  Y  A  L  E  R  S  F  L  I  L  C  R  -
c     Y  E  C  L  R  G  G  L  D  F  T  K  D  D  E  N  V  N  S  Q  P  F  M  R  W  R  D  R  F  L  F  C  A  E  -

                                                                                 NlaIII
                                                                              FatI |
                                                                             RsaI| |
                                                                           Csp6I|| |
                                                                            PstI ||| |        MboII
                                                                      HpyCH4V | ||| |      BstKTI |
                                                                        SfcI | | ||| |       DpnI| |     HpyF10VI
          PsiI        CviJI                 HphI           BspMI BsmI | | | ||| |NspI  BclI || | AluI       |
             |            |                    |               |    | | | | ||| |    MboI || | CviJI MwoI|
     aagcactttataaagcacaggctgaaacaggtgaaatcaaagggcattacttgaatgctactgcaggtacatgcgaagaaatgatcaaaagagctgtatt
 601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 700
     ttcgtgaaatatttcgtgtccgactttgtccactttagtttcccgtaatgaacttacgatgacgtccatgtacgcttctttactagttttctcgacataa

a     K  H  F  I  K  H  R  L  K  Q  V  K  S  K  G  I  T  *  M  L  L  Q  V  H  A  K  K  *  S  K  E  L  Y  L  -
b      S  T  L  *  S  T  G  *  N  R  *  N  Q  R  A  L  L  E  C  Y  C  R  Y  M  R  R  N  D  Q  K  S  C  I  -
c       A  L  Y  K  A  Q  A  E  T  G  E  I  K  G  H  Y  L  N  A  T  A  G  T  C  E  E  M  L  K  R  A  V  F  -

                           BsiEI
                          BstKTI      NlaIII
                            PvuI      NsiI |                                                        Hpy8I
          BcgI             DpnI|    BfrBI | |        HphI              AluI                          AccI|
     Tsp509I |      Hpy188I || |      FatI | |     BcgI |HinfI         CviJI                        FokI ||
    BfaI   | |         MboI ||HpyCH4V | |  MseI | | TfiI  AciI   BfaI |CviJI                   MboII ||
       |   | |            | ||       | | |     | | |    |     |       | |    |                      | ||
     tgctagagaattgggcgttccgatcgtaatgcatgactacttaacgggggattcaccgcaaatactagcttggctcattattgccgagataatggtcta
 701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 800
     acgatctcttaacccgcaaggctagcattacgtactgatgaattgccccctaagtggcgtttatgatcgaaccgagtaataacggctctattaccagat

a      L  E  N  W  A  F  R  S  *  C  M  T  T  *  R  G  D  S  P  Q  I  L  A  W  L  I  I  A  E  I  M  V  Y  -
b     C  *  R  I  G  R  S  D  R  N  A  *  L  L  N  G  G  I  H  R  K  Y  *  L  G  S  L  L  P  R  *  W  S  T  -
c      A  R  E  L  G  V  P  I  V  M  H  D  Y  L  T  G  G  F  T  A  N  T  S  L  A  H  Y  C  R  D  N  G  L  -
```

Figure 21-D

```
                                         AciI
                                       NlaIII
                                         NspI
                                         SphI
                                      BsrDI
                                       Cac8I
                                    HpyF10VI
                                        NsiI
                                     BstAPI
                                       MwoI
                                     BfrBI                                          BciVI
                                      FatI                              MboII        NciI                   TaiI
                                  HpyCH4V                            NlaIII         ScrFI               BsaAI
                             HpyCH4V                                FatI            BslI                SnaBI
                         HpyCH4III                               HinfI             HpaII            HpyCH4IV
                     BstF5I                                      TfiI            StyD4I             MaeIII
       cttcttcacatccaccgtgcaatgcatgcggttattgatagacagaagaatcatggtatccacttccgggtattagcaaaagcgttacgtatgtctggtg
   801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 900
       gaagaagtgtaggtggcacgttacgtacgccaataactatctgtcttcttagtaccataggtgaaggcccataatcgttttcgcaatgcatacagaccac

a        F  F  T  S  T  V  Q  C  M  R  L  L  I  D  R  R  I  M  V  S  T  S  G  Y  *  Q  K  R  Y  V  C  L  V  -
b         S  S  H  P  P  C  N  A  C  G  Y  *  *  T  E  E  S  W  Y  P  L  P  G  I  S  K  S  V  T  Y  V  W  W  -
c        L  L  H  I  H  R  A  M  H  A  V  I  D  R  Q  K  N  H  G  I  H  F  R  V  L  A  K  A  L  R  M  S  G  G -

                           HpyCH4III
                                KpnI
                              NlaIV
                               RsaI
       BstKTI
        DpnI                 Csp6I
      BaeI                 Acc65I          BaeI      BaeI
      MboI    BaeI           BanI       Hpy8I      BsmAI        HphI        CviJI
       gagatcatattcactctggtaccgtagtaggtaaacttgaaggtgaaagagacataactttgggctttgttgatttactgcgtgatgattttgttgaaca
   901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1000
       ctctagtataagtgagaccatggcatcatccatttgaacttccactttctctgtattgaaacccgaaacaactaaatgacgcactactaaaacaacttgt

a        E  I  I  F  T  L  V  P  *  *  V  N  L  K  V  K  E  T  *  L  W  A  L  L  I  Y  C  V  M  I  L  L  N  K -
b         R  S  Y  S  L  W  Y  R  S  R  *  T  *  R  *  K  R  H  N  F  G  L  C  *  F  T  A  *  *  F  C  *  T  -
c          D  H  I  H  S  G  T  V  V  G  K  L  E  G  E  R  D  I  T  L  G  F  V  D  L  L  R  D  D  F  V  E  Q  -
```

Figure 21-E

```
                                          Hin4I
                                      Eco57I
                                     Eco57MI
                                    BstNI
                                    ScrFI
                                  PspGI
                                  SexAI
   BstKTI                              Hpy188III
     TaqI     AciI       Hpy188III       StyD4I          CviJI
     DpnI     BpuEI        SmlI          BsaI         BsaJI               TaiI
   MboI       BstUI   Hin4I              BsmAI        BtgI  MnlI   HpyCH4IV    NdeI  Cac8I
     agatcgaagtcgcggtatttatttcactcaagattgggtctctttaccaggtgttctacccgtggcttcaggaggtattcacgtttggcatatgcctgct
1001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1100
     tctagcttcagcgccataaataaagtgagttctaacccagagaaatggtccacaagatgggcaccgaagtcctccataagtgcaaaccgtatacggacga

a      I  E  V  A  V  F  I  S  L  K  I  G  S  L  Y  Q  V  F  Y  P  W  L  Q  E  V  F  T  F  G  I  C  L  L  -
b     R  S  K  S  R  Y  L  F  H  S  R  L  G  L  F  T  R  C  S  T  R  G  F  R  R  Y  S  R  L  A  Y  A  C  S -
c      D  R  S  R  G  I  Y  F  T  Q  D  W  V  S  L  P  G  V  L  P  V  A  S  G  G  I  H  V  W  H  M  P  A   -

                     MnlI
                  HpyCH4III
                    FokI
                    SfcI
                  RsaI
                 Csp6I                                                 NlaIV
             BstF5I                                                  HpyF10VI
            HinfI                                                      BanI
             TfiI                                                      MwoI
           TaqII                                                     BstNI
        TspGWI                                                       ScrFI
      BstKTI                                                        HhaI
       DpnI                                                         PspGI
      BglII                                          BsaJI         StyD4I
      BstYI                                           StyI       HinP1I       AluI          AluI
Hpy188I MboI                             FokI     BstF5I   BceAI                CviJI TaqI  CviJI
     ctgaccgagatctttggggatgattccgtactacagttcggtggaggaactttaggacatccttggggtaatgcgccaggtgccgtagctaatcgagtag
1101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1200
     gactggctctagaaacccctactaaggcatgatgtcaagccacctccttgaaatcctgtaggaaccccattacgcggtccacggcatcgattagctcatc

a     *  P  R  S  L  G  M  I  P  Y  Y  S  S  V  E  E  L  *  D  I  L  G  V  M  R  Q  V  P  *  L  I  E  *  -
b      D  R  D  L  W  G  *  F  R  T  T  V  R  W  R  N  F  R  T  S  L  G  *  C  A  R  C  R  S  *  S  S  S -
c     L  T  E  I  F  G  D  D  S  V  L  Q  F  G  G  G  T  L  G  H  P  W  G  N  A  P  G  A  V  A  N  R  V  A-
```

Figure 21-F

```
BanII
Bsp1286I
HpaII
NciI
ScrFI
BbvI
CviJI
StyD4I
NlaIV
BbvI
HpyF10VI
HpyCH4V
Cac8I
HpyF10VI
MwoI
CviJI
TspDTI
BstUI
NruI
Hpy188III
MnlI
XmnI
Hpy188III
Bpu10I
DdeI
TspDTI
BstKTI
DpnI
MboI
TaiI
BmgBI
HpyCH4IV
Tsp509I
BseMII
BspCNI
MwoI
NlaIII
FatI
HpyF10VI
MwoI
BfaI
Hpy188III
XbaI
NspI
AluI
CviJI
Fnu4HI
AluI
CviJI
TseI
BfaI

     ctctagaagcatgtgtaaaagctcgtaatgaaggacgtgatcttgctcaggaaggtaatgaaattattcgcgaggcttgcaaatggagcccggaactagc
1201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1300
     gagatcttcgtacacattttcgagcattacttcctgcactagaacgagtccttccattactttaataagcgctccgaacgtttacctcgggccttgatcg

a    L * K H V * K L V M K D V I L L R K V M K L F A R L A N G A R N * L -
b     S R S M C K S S * * R T * S C S G R * * N Y S R G L Q M E P G T S  -
c      L E A C V K A R N E G R D L A Q E G N E I I R E A C K W S P E L A -

TaiI
BbvI
Fnu4HI
BtsI
BstKTI
HpyCH4V
HpyCH4IV
Hpy8I
Fnu4HI
DpnI
Tsp509I
TspRI
AccI
TseI
MboI
MseI
TseI
Hpy8I
HpyCH4III
Tsp509I

     tgctgcttgtgaagtatggaaagagatcgtatttaattttgcagcagtggacgttttggataagtaaaaacagtagacattagcagataaattagcagga
1301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1400
     acgacgaacacttcatacctttctctagcataaattaaaacgtcgtcacctgcaaaacctattcatttttgtcatctgtaatcgtctatttaatcgtcct

a     L L V K Y G K R S Y L I L Q Q W T F W I S K N S R H * Q I N * Q E -
b    C C L * S M E R D R I * F C S S G R F G * V K T V D I S R * I S R K -
c     A A C E V W K E I V F N F A A V D V L D K * K Q * T L A D K L A G  -
```

Figure 21-G

```
                                                                  MseI
                                                              Tsp509I   |
                                                             HpyCH4V|   |
                                                          Tsp509I   ||  |   CviJI
                  Tsp509I                       Tsp509I       |     ||  |   HaeIII
       BpuEI      SmlI   |                       MseI|        |     ||  |  Sau96I|                   CviJI
           |         |    |                          ||       |     ||  |       ||                       |
     aataaagaaggataaggagaaagaactcaagtaattatccttcgttctcttaattgaattgcaattaaactcggcccaatcttttactaaaggattgag
1401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1500
     ttatttcttcctattcctctttcttgagttcattaataggaagcaagagaattaacttaacgttaatttgagccgggttagaaaatgattttcctaactc

a     I  K  K  D  K  E  K  E  L  K  *  L  S  F  V  L  L  I  E  L  Q  L  N  S  A  Q  S  F  T  K  R  I  E  -
b      *  R  R  I  R  R  K  N  S  S  N  Y  P  S  F  S  *  L  N  C  N  *  T  R  P  N  L  L  L  K  G  L  S  -
c     N  K  E  G  *  G  E  R  T  Q  V  I  I  L  R  S  L  N  *  I  A  I  K  L  G  P  I  F  Y  *  K  D  *  A -

                                                                                  BfaI
           HinfI                                                              Hpy188III              MlyI
            TfiI  HpyCH4V          DdeI               BfaI                       XbaI|              PleI
               |        |              |                  |                          ||                  |
     ccgaatacaacaaagattctattgcatatattttgactaagtatatacttacctagatatacaagatttgaaatacaaaatctagaaaactaaatcaaaa
1501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1600
     ggcttatgttgtttctaagataacgtatataaaactgattcatatatgaatggatctatatgttctaaactttatgttttagatcttttgatttagtttt

a     P  N  T  T  K  I  L  L  H  I  F  *  L  S  I  Y  L  P  R  Y  T  R  F  E  I  Q  N  L  E  N  *  I  K  I -
b      R  I  Q  Q  R  F  Y  C  I  Y  F  D  *  V  Y  T  Y  L  D  I  Q  D  L  K  Y  K  I  *  K  T  K  S  K  -
c       E  Y  N  K  D  S  I  A  Y  I  L  T  K  Y  I  L  T  *  I  Y  K  I  *  N  T  K  S  R  K  L  N  Q  N  -
```

Figure 21-H

```
HincII, Hpy8I, AccI, TaqI, SalI, PstI, HpyCH4V, Cac8I, SfcI, NlaIII, NspI, SphI, Cac8I, FatI, HpyCH4V, Cac8I, AluI, CviJI, AlwI, TaqI, BstKTI, DpnI, MboI, HinfI, DdeI, HpyCH4III, Hpy188I, RsaI, Csp6I, TatI

     tctaagactcaaatctttctattgttgtcttggatcatcgagcttgcatgcctgcagtcgactttcacagtttccattctgaaatgttctctgtactata
1601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1700
     agattctgagtttagaaagataacaacagaacctagtagctcgaacgtacggacgtcagctgaaagtgtcaaaggtaagactttacaagagacatgatat

a     * D S N L S I V V L D H R A C M P A V D F H S F H S E M F S V L * -
b    S K T Q I F L L L S W I I E L A C L Q S T F T V S I L K C S L Y Y N -
c     L R L K S F Y C C L G S S S L H A C S R L S Q F P F * N V L C T I -

Hpy8I, MboII, AccI, MnlI, EarI, Hpy188I, HinfI, TfiI, Tsp509I, CviJI, HpyCH4V

     ataaatagtaagtgaatcaacttactactaaaaaaattagtagacttcctcttcggaatagaaatagcctatttctacatagggaaagtcgtgtgcaatg
1701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1800
     tatttatcattcacttagttgaatgatgatttttttaatcatctgaaggagaagccttatctttatcggataaagatgtatccctttcagcacacgttac

a     * I V S E S T Y Y * K N * * T S S S E * K * P I S T * G K S C A M -
b      K * * V N Q L T T K K I S R L P L R N R N S L F L H R E S R V Q * -
c    I N S K * I N L L L K K L V D F L F G I E I A Y F Y I G K V V C N E -
```

Figure 21-I

```
                                                                                                MlyI
                                                                                               PleI|
                                                                                             MmeI ||
                                                                                             NciI ||
                                                                                            ScrFI ||
                                                                                             SmaI ||
                                                                   NciI              HpaII| ||
                                                                  ScrFI               NciI| ||
                                                                 HpaII|              ScrFI| ||
                                                               StyD4I||               AvaI|| ||
                                                            BsmFI   |||              BsaJI|| ||
                                                             BccI|  |||             StyD4I|| ||
                                                             BfaI|  |||               XmaI|| ||
                       TspDTI                                SpeI|| |||             StyD4I||| ||
                  Cac8I  |                                    |||  |||
            HpyCH4V |    |                      Tsp509I        Hpy188I ||| |||     HinfI |||| ||
        BsrDI   | |   |  MnlI          MaeIII     FokI|     BstF5I   | ||| |||    TaqI | |||| ||
          |     | |   |   |              |         ||        |     | ||| |||     | |  |||| ||
          aaaaatgcaagcacgatttggggagaggtttttctctattgtaacaaggaataattatctactccatccgactagttccgggttcgagtcccgggcaac
     1801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 1900
          tttttacgttcgtgctaaacccctctccaaaaagagataacattgttccttattaatagatgaggtaggctgatcaaggcccaagctcagggcccgttg

a          K  N  A  S  T  I  W  G  E  V  F  S  L  L  *  Q  G  I  I  I  Y  S  I  R  L  V  P  G  S  S  P  G  Q  P -
b           K  M  Q  A  R  F  G  E  R  F  F  L  Y  C  N  K  E  *  L  S  T  P  S  D  *  F  R  V  R  V  P  G  N  -
c            K  C  K  H  D  L  G  R  G  F  F  S  I  V  T  R  N  N  Y  L  L  H  P  T  S  S  G  F  E  S  R  A  T -

                      BspCNI
                     BseMII|  Hpy188I
        NdeI    BfaI     ||    DdeI|       TspDTI                ApoI               ApoI
          |       |      ||     ||           |                 Tsp509I            Tsp509I
          ccatatggaaactagaaaggagcaatctgagtttgattttcactcacttcatttacaaaatttttggtttggtaaatttttgttgtatggatatacaa
     1901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2000
          ggtatacctttgatctttcctcgttagactcaaaactaaaagtgagtgaagtaaatgttttaaaaaccaaaccatttaaaacaacatacctatatgtt

a            I  W  K  L  E  R  S  N  L  S  F  D  F  S  L  T  S  F  T  K  F  F  G  L  V  N  F  V  V  W  I  Y  N -
b          P  Y  G  N  *  K  G  A  I  *  V  L  I  F  H  S  L  H  L  Q  N  F  L  V  W  *  I  L  L  Y  G  Y  T  T -
c           H  M  E  T  R  K  E  Q  S  E  F  *  F  F  T  H  F  I  Y  K  I  F  W  F  G  K  F  C  C  M  D  I  Q  -

                                                                                           TaiI
                                                                                        BsaAI |
                                                                                      AflIII| |
                CviJI                                                               HpyCH4IV| |
               Cac8I | HincII                           MseI                         MseI   || |
   HpyCH4III  CviJI | |  Hpy8I               HpyCH4III |         CviJI        BfaI   |    || |       HinfI
          |      | | |    |                          | |           |            |     |    || |         |
          ctgtcggggctggcttggttgacattggtatatagtctatattatactgttaaataacaagccttctattatctttctagttaatacgtgtgcttgggag
     2001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2100
          gacagccccgaccgaaccaactgtaaccatatatcagatataatatgacaatttattgttcggaagataatagaaagatcaattatgcacacgaaccctc

a           C  R  G  W  L  G  *  H  W  Y  I  V  Y  I  I  L  L  N  N  K  P  S  I  I  F  L  V  N  T  C  A  W  E  -
b            V  G  A  G  L  V  D  I  G  I  *  S  I  L  Y  C  *  I  T  S  L  L  L  S  F  *  L  I  R  V  L  G  S -
c          L  S  G  L  A  W  L  T  L  V  Y  S  L  Y  Y  T  V  K  *  Q  A  F  Y  Y  L  S  S  *  Y  V  C  L  G  V -
```

Figure 21-J

```
                                    MaeIII
                                    Tsp45I
                                    FauI
                                    BslI
                                    BfaI
                                    AciI
                                    ApaI
                                    BanII
                                    Bme1580I
                                    Bsp1286I
                                    Cac8I
                                    CviJI
                                    HaeIII
                                    NlaIV
                                    NlaIV
               AciI                 Sau96I
               BstUI                EcoO109I
               HhaI                 PspOMI
               BstUI                Sau96I
               HinP1I               SacII
               TauI                 AciI
               AciI                 BstUI
               Fnu4HI               MspA1I
               CviJI                AciI
               HaeIII               BsaJI
               NlaIII               BtgI
               BaeI                 AvaII
               FauI                 NlaIII
               BsaJI                Sau96I
         BstKTI                     FauI
 MlyI    DpnI  BtgI                 FatI
 Tsp509I BglII FatI                 MslI                 AgeI
 HpyCH4V BstYI NcoI                                      BsaWI
 PleI    MboI  StyI                                      BsrFI
     tccttgcaatttgaataaaccaagatcttaccatggccgcgcggggttctcatcatcatcatcatggtccgcgggggcccgctagtggtagtgaccttga
2101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2200
     aggaacgttaaacttatttggttctagaatggtaccggcgcgccccaagagtagtagtagtagtaccaggcgcccccgggcgatcaccatcactggaact

a    S  L  Q  F  E  *  T  K  I  L  P  W  P  R  G  V  L  I  I  I  I  M  V  R  G  G  P  L  V  V  V  T  L  T -
b     P  C  N  L  N  K  P  R  S  Y  H  G  R  A  G  F  S  S  S  S  S  W  S  A  G  A  R  *  W  *  *  P  *  -
c      L  A  I  *  I  N  Q  D  L  T  M  A  A  R  G  S  H  H  H  H  H  G  P  R  G  P  A  S  G  S  D  L  D -

 Bme1580I
 BsiHKAI
 Bsp1286I                                                           Hpy188I
 Hpy8I                                                              TspDTI
 HpyCH4V         Tsp509I                                            BstKTI
 ApaLI           AluI                       Hpy188III  ApoI         DpnI
HpaII            CviJI  TspDTI      MnlI    Hpy8I      Tsp509I      MboI
     ccggtgcaccacttttgatgatgttcaagctcctaattacactcaacatacttcatctatgaggggggtttactatcctgatgaaattttagatcagac
2201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2300
     ggccacgtggtgaaaactactacaagttcgaggattaatgtgagttgtatgaagtagatactccccccaaatgataggactactttaaaatctagtctg

a      G  A  P  L  L  M  M  F  K  L  L  I  T  L  N  I  L  H  L  *  G  G  F  T  I  L  M  K  F  L  D  Q  T -
b     P  V  H  H  F  *  *  C  S  S  S  *  L  H  S  T  Y  F  I  Y  E  G  G  L  L  S  *  *  N  F  *  I  R  H -
c     R  C  T  T  F  D  D  V  Q  A  P  N  Y  T  Q  H  T  S  S  M  R  G  V  Y  Y  P  D  E  I  F  R  S  D  -
```

Figure 21-K

```
                  MboII                                                        TaiI
           Hpy188III |                                              HpyCH4IV  |
              DdeI | |     BseMII            TspDTI              AseI       | |                    MseI
          MseI  | | |    BspCNI|            MaeIII|              MseI       | |                    BccI|
             |  | | |         ||                 ||                 |       | |                      ||
     actctttatttaactcaggatttatttcttccattttattctaatgttacagggtttcatactattaatcatacgtttggcaaccctgtcatacctttta
2301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2400
     tgagaaataaattgagtcctaaataaagaaggtaaaataagattacaattgtcccaaagtatgataattagtatgcaaaccgttgggacagtatggaaaat

a     L  F  I  *  L  R  I  Y  F  F  H  F  I  L  M  L  Q  G  F  I  L  L  I  I  R  L  A  T  L  S  Y  L  L  -
b      S  L  F  N  S  G  F  I  S  S  I  L  F  *  C  Y  R  V  S  Y  Y  *  S  Y  V  W  Q  P  C  H  T  F  *  -
c    T  L  Y  L  T  Q  D  L  F  L  P  F  Y  S  N  V  T  G  F  H  T  I  N  H  T  F  G  N  P  V  I  P  F  K -

                                                                                  HpyCH4III
             Fnu4HI                                                               TspDTI|
   BstF5I    TseI|                   BsaJI                      NlaIII  MaeIII      ||
   BbvI |   FokI ||        TspGWI     BtgI  BslI               FatI |  Tsp45I       ||                MseI
      | |      | ||             |        |     |                  | |       |       ||                   |
     aggatggtatttattttgctgccacagagaaatcaaatgttgtccgtggttgggtttttggttctaccatgaacaacaagtcacagtcggtgattattat
2401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2500
     tcctaccataaataaaacgacggtgtctctttagtttacaacaggcaccaacccaaaaaccaagatggtacttgttgttcagtgtcagccactaataata

a    R  M  V  F  I  L  L  P  Q  R  N  Q  M  L  S  V  V  G  F  L  V  L  P  *  T  T  S  H  S  R  *  L  L  L -
b     G  W  Y  L  F  C  C  H  R  E  I  K  C  C  P  W  L  G  F  W  F  Y  H  E  Q  Q  V  T  V  G  D  Y  Y  -
c      D  G  I  Y  F  A  A  T  E  K  S  N  V  V  R  G  W  V  F  G  S  T  M  N  N  K  S  Q  S  V  I  I  I  -

                                                                                     Hpy8I
                                                                                     RsaI|
                                                                                   Csp6I||
                                                                                 NlaIII|||
                                                                               BsaJI | |||
                                 NlaIII                                         BtgI | |||
                                   NspI     MaeIII                              FatI | |||
   Tsp509I                       MaeIII|    Tsp45I                              NcoI | |||
   HphI |                         FatI || Tsp509I  |                            StyI | |||
      | |                            | ||       |  |                               | | |||
     taacaattctactaatgttgttatacgagcatgtaactttgaattgtgtgacaaccctttctttgctgtttctaaacccatgggtacacagacacatact
2501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2600
     attgttaagatgattacaacaatatgctcgtacattgaaacttaacacactgttgggaaagaaacgacaaagatttgggtacccatgtgtctgtgtatga

a      T  I  L  L  M  L  L  Y  E  H  V  T  L  N  C  V  T  T  L  S  L  L  F  L  N  P  W  V  H  R  H  I  L  -
b    *  Q  F  Y  *  C  C  Y  T  S  M  *  L  *  I  V  *  Q  P  F  L  C  C  F  *  T  H  G  Y  T  D  T  Y  Y -
c     N  N  S  T  N  V  V  I  R  A  C  N  F  E  L  C  D  N  P  F  F  A  V  S  K  P  M  G  T  Q  T  H  T  -
```

Figure 21-L

```
                                       HpyCH4V
                  Hin4I
               Tsp509I        RsaI
                  MseI       Csp6I
                  NsiI       SfaNI
               BfrBI          TatI                                             DraI
     TaqI HpyCH4V      TaqI     Hpy188I      Hin4I     Hpy188I     Tsp509I MseI
     atgatattcgataatgcatttaattgcactttcgagtacatatctgatgccttttcgcttgatgtttcagaaaagtcaggtaatttaaacacttacgag
2601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2700
     tactataagctattacgtaaattaacgtgaaagctcatgtatagactacggaaaagcgaactacaaagtcttttcagtccattaaatttgtgaatgctc

a     * Y S I M H L I A L S S T Y L M P F R L M F Q K S Q V I L N T Y E  -
b      D I R * C I * L H F R V H I * C L F A * C F R K V R * F * T L T R -
c     M I F D N A F N C T F E Y I S D A F S L D V S E K S G N F K H L R E -

                                                              BstKTI
                                                               DpnI
       DraI                                                   MboI
      MseI BccI                    PsiI CviJI    SfcI    Hpy188III         MseI
     agtttgtgtttaaaaataaagatgggtttctctatgtttataagggctatcaacctatagatgtagttcgtgatctaccttctggttttaacactttgaa
2701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2800
     tcaaacacaaatttttatttctacccaaagagatacaaatattcccgatagttggatatctacatcaagcactagatggaagaccaaaattgtgaaactt

a    S L C L K I K M G F S M F I R A I N L * M * F V I Y L L V L T L * N -
b     V C V * K * R W V S L C L * G L S T Y R C S S * S T F W F * H F E  -
c      F V F K N K D G F L Y V Y K G Y Q P I D V V R D L P S G F N T L K -

                                                                                         PstI
                                                                                       Fnu4HI
                                                                                      HpyCH4V
                                                                                         TseI
                                                                                       Fnu4HI
                                                                                         SfcI
                                                                                        AluI
                                                                                       CviJI
                                                                                      MspA1I
                                                                                       PvuII
                                                                                        TseI
                                                                                      TaiI
                                                                                    BmgBI
                                                        AarI                    Bme1580I
                                         CviJI          BspMI                   Bsp1286I
                          MnlI   ApoI    BpuEI      Hpy188III                  HpyCH4IV
         MseI             MseI Tsp509I   CviJI   HphI    SmlI   BbvI
     acctattttaagttgcctcttggtattaacattacaaattttagagccattcttacagccttttcacctgctcaagacatttggggcacgtcagctgca
2801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 2900
     tggataaaaattcaacggagaaccataattgtaatgtttaaaatctcggtaagaatgtcggaaaagtggacgagttctgtaaaccccgtgcagtcgacgt

a      L F L S C L L V L T L Q I L E P F L Q P F H L L K T F G A R Q L Q -
b     T Y F * V A S W Y * H Y K F * S H S Y S L F T C S R H L G H V S C S-
c      P I F K L P L G I N I T N F R A I L T A F S P A Q D I W G T S A A -
```

Figure 21-M

```
                          BpuEI                                            TspDTI
                         CviJI|                                        RsaI   |
                  CviJI  DraI ||                                      SfaNI   |
 CviJI       BbvI   | MseI|   ||               SmlI                  Csp6I|   |    AlwNI
   |           |    |    ||   ||                |                       ||    |      |
      gcctattttgttggctatttaaagccaactacatttatgctcaagtatgatgaaaatggtacaatcacagatgctgttgattgttctcaaaatccacttg
 2901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3000
      cggataaaacaaccgataaatttcggttgatgtaaatacgagttcatactacttttaccatgttagtgtctacgacaactaacaagagttttaggtgaac

a       P  I  L  L  A  I  *  S  Q  L  H  L  C  S  S  M  M  K  M  V  Q  S  Q  M  L  L  I  V  L  K  I  H  L  -
b        L  F  C  W  L  F  K  A  N  Y  I  Y  A  Q  V  *  *  K  W  Y  N  H  R  C  C  *  L  F  S  K  S  T  C  -
c      A  Y  F  V  G  Y  L  K  P  T  T  F  M  L  K  Y  D  E  N  G  T  I  T  D  A  V  D  C  S  Q  N  P  L  A -

                                                                                                 HinfI
                                                                          Hpy188III               TfiI
                                                                           Bsu36I |         BseMII |
                   AluI                 ApoI                                DdeI  |        BspCNI| |
              MseI CviJI               Tsp509I          Tsp509I  MnlI   BsaXI  |  |     MnlI   || |
                 |    |                   |                   |      |        |  |  |        |  || |
      ctgaactcaaatgctctgttaagagctttgagattgacaaaggaatttaccagacctctaatttcagggttgttccctcaggagatgttgtgagattccc
 3001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3100
      gacttgagtttacgagacaattctcgaaactctaactgtttccttaaatggtctggagattaaagtcccaacaagggagtcctctacaacactctaaggg

a       L  N  S  N  A  L  L  R  A  L  R  L  T  K  E  F  T  R  P  L  I  S  G  L  F  P  Q  E  M  L  *  D  S  L -
b        *  T  Q  M  L  C  *  E  L  *  D  *  Q  R  N  L  P  D  L  *  F  Q  G  C  S  L  R  R  C  C  E  I  P  -
c         E  L  K  C  S  V  K  S  F  E  I  D  K  G  I  Y  Q  T  S  N  F  R  V  V  P  S  G  D  V  V  R  F  P  -

                                                                    NlaIII
                                                                     NsiI |
                                                                  BfrBI | |
   SspI                                          ApoI                 FatI| |                   Tsp509I
  BsaXI|               MnlI           MseI     Tsp509I            HpyCH4V | |      ApoI               |
      ||                 |              |         |                     | | |    Tsp509I              |
      taatattacaaacttgtgtccttttggagaggtttttaatgctactaaattcccttctgtctatgcatgggagagaaaaaaaatttctaattgtgttgct
 3101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3200
      attataatgtttgaacacaggaaaacctctccaaaaattacgatgatttaagggaagacagatacgtaccctctcttttttttaaagattaacacaacga

a        I  L  Q  T  C  V  L  L  E  R  F  L  M  L  L  N  S  L  L  S  M  H  G  R  E  K  K  F  L  I  V  L  L -
b       *  Y  Y  K  L  V  S  F  W  R  G  F  *  C  Y  *  I  P  F  C  L  C  M  G  E  K  K  N  F  *  L  C  C  * -
c        N  I  T  N  L  C  P  F  G  E  V  F  N  A  T  K  F  P  S  V  Y  A  W  E  R  K  K  I  S  N  C  V  A  -

                                                                          BstKTI
           BsiHKAI                                                          DpnI|
          Bsp1286I                     MseI                 DdeI       MboI ||                HpyCH4V
                 |                        |                    |          | ||                      |
      gattactctgtgctctacaactcaacatttttttcaaccttttaagtgctatggcgtttctgccactaagttgaatgatctttgcttctccaatgtctatg
 3201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3300
      ctaatgagacacgagatgttgagttgtaaaaaaagttggaaattcacgataccgcaaagacggtgattcaacttactagaaacgaagaggttacagatac

a        I  T  L  C  S  T  T  Q  H  F  F  Q  P  L  S  A  M  A  F  L  P  L  S  *  M  I  F  A  S  P  M  S  M  -
b         L  L  C  A  L  Q  L  N  I  F  F  N  L  *  V  L  W  R  F  C  H  *  V  E  *  S  L  L  L  Q  C  L  C -
c       D  Y  S  V  L  Y  N  S  T  F  F  S  T  F  K  C  Y  G  V  S  A  T  K  L  N  D  L  C  F  S  N  V  Y  A -
```

Figure 21-N

```
                                    BstNI
                                    ScrFI
                                   HaeII|
                                   HhaI||
                                  PspGI||                                      TspDTI
                                  StyD4I||                           Tsp509I      |
   HinfI                          BsaXI|||                   Tsp509I      |       |
    TfiI     BsaXI               HinP1I||||          BsrI     PsiI| PsiI  |       |          FatI
     |        |                       |||||            |        ||    | |     |          |
     cagattcttttgtagtcaagggagatgatgtaagacaaatagcgccaggacaaactggtgttattgctgattataattataaattgccagatgatttcat
3301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3400
     gtctaagaaaacatcagttccctctactacattctgtttatcgcggtcctgtttgaccacaataacgactaatattaatatttaacggtctactaaagta

a    Q  I  L  L  *  S  R  E  M  M  *  D  K  *  R  Q  D  K  L  V  L  L  L  I  I  I  I  N  C  Q  M  I  S  W -
b     R  F  F  C  S  Q  G  R  *  C  K  T  N  S  A  R  T  N  W  C  Y  C  *  L  *  L  *  I  A  R  *  F  H  -
c      D  S  F  V  V  K  G  D  D  V  R  Q  I  A  P  G  Q  T  G  V  I  A  D  Y  N  Y  K  L  P  D  D  F  M  -

                                                                                                CviJI
                                                                                               HaeIII
                                                                                             EcoO109I|
                                                                                               Sau96I|
                                                                                           Bpu10I   ||
                                                                                             DdeI   ||
                                                                                             AluI|  ||
                                                                                            CviJI|  ||
                                                                                         Cac8I ||   ||
                                                    Tsp509I                           HindIII ||   ||
                         SfaNI                  BsrI     |                           NlaIII | ||   ||
NlaIII                   BfaI|              Tsp509I PsiI| PsiI              DdeI FatI   |  | ||   ||
     |                      ||                   |    ||   |                   |    |   |  | ||   ||
     gggttgtgtccttgcttggaatactaggaacattgatgctacttcaactggtaattataattataaatataggtatcttagacatggcaagcttaggccc
3401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3500
     cccaacacaggaacgaaccttatgatccttgtaactacgatgaagttgaccattaatattaatatttatatccatagaatctgtaccgttcgaatccggg

a     V  V  S  L  L  G  I  L  G  T  L  M  L  L  Q  L  V  I  I  I  I  N  I  G  I  L  D  M  A  S  L  G  P  -
b    G  L  C  P  C  L  E  Y  *  E  H  *  C  Y  F  N  W  *  L  *  L  *  I  *  V  S  *  T  W  Q  A  *  A  L -
c     G  C  V  L  A  W  N  T  R  N  I  D  A  T  S  T  G  N  Y  N  Y  K  Y  R  Y  L  R  H  G  K  L  R  P  -

                                         HpyCH4V                AarI       CviJI
                                       HpyF10VI |               BspMI      HaeIII
                                          BstAPI| |            Tsp509I       MscI
 BsmAI                         BccI          MwoI| |             MseI|      EaeI |MseI
     |                           |              || |               ||          |  |   |
     tttgagagagacatatctaatgtgcctttctcccctgatggcaaaccttgcaccccacctgctcttaattgttattggccattaaatgattatggttttt
3501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3600
     aaactctctctgtatagattacacggaaagaggggactaccgtttggaacgtggggtggacgagaattaacaataaccggtaatttactaataccaaaaa

a     L  R  E  T  Y  L  M  C  L  S  P  L  M  A  N  L  A  P  H  L  L  L  I  V  I  G  H  *  M  I  M  V  F  -
b     *  E  R  H  I  *  C  A  F  L  P  *  W  Q  T  L  H  P  T  C  S  *  L  L  L  A  I  K  *  L  W  F  L -
c    F  E  R  D  I  S  N  V  P  F  S  P  D  G  K  P  C  T  P  P  A  L  N  C  Y  W  P  L  N  D  Y  G  F  Y -
```

Figure 21-O

```
                                                                       HpyCH4III
                                                                       BsaJI
                                                                       BtgI
                                                                       CviJI
                                                                       HaeIII
                                                                       EaeI
                                                                       HpaII
                         RsaI                                          BsrFI           Tsp509I
                         ScaI                                  HpyCH4V              AvaII
           CviJI         Csp6I                           DraI                       Hpy8I
      BsrI               TatI                            MseI                       Sau96I
     acaccactactggcattggctaccaaccttacagagttgtagtactttcttttgaacttttaaatgcaccggccacggtttgtggaccaaaattatccac
3601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3700
     tgtggtgatgaccgtaaccgatggttggaatgtctcaacatcatgaaagaaaacttgaaaatttacgtggccggtgccaaacacctggttttaataggtg

a     T  P  L  L  A  L  A  T  N  L  T  E  L  *  Y  F  L  L  N  F  *  M  H  R  P  R  F  V  D  Q  N  Y  P  L -
b      H  H  Y  W  H  W  L  P  T  L  Q  S  C  S  T  F  F  *  T  F  K  C  T  G  H  G  L  W  T  K  I  I  H  -
c       T  T  T  G  I  G  Y  Q  P  Y  R  V  V  V  L  S  F  E  L  L  N  A  P  A  T  V  C  G  P  K  L  S  T  -

                                                                  HincII
                                                                  HpaI
                                                        BsrI      Hpy8I
                         Tsp509I MlyI                   RsaI      MboII
                  Tsp509I        MseI  BaeI             TspRI     MseI
TspRI MseI   BsrI TspRI   MseI   PleI  HinfI            Csp6I BsrI              BaeI          BceAI Tsp509I
     tgaccttattaagaaccagtgtgtcaatttaatttaatggactcactggtactggtgtgttaactccttcttcaaagagatttcaaccatttcaacaa
3701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3800
     actggaataattcttggtcacacagttaaattaaattacctgagtgaccatgaccacacaattgaggaagaagtttctctaaagttggtaaagttgtt

a      T  L  L  R  T  S  V  S  I  L  I  L  M  D  S  L  V  L  V  C  *  L  L  L  Q  R  D  F  N  H  F  N  N  -
b     *  P  Y  *  E  P  V  C  Q  F  *  F  *  W  T  H  W  Y  W  C  V  N  S  F  F  K  E  I  S  T  I  S  T  I -
c      D  L  I  K  N  Q  C  V  N  F  N  F  N  G  L  T  G  T  G  V  L  T  P  S  S  K  R  F  Q  P  F  Q  Q  -

                                        BstKII
                                        DpnI
                                        BstYI
                                        MboI
                                  Hpy188III
      CviJI                    TspRI
      HaeIII             TspGWI HinfI   TaqI                         HphI
      EaeI        Hpy188I      TfiI   AlwI            Hpy188I SspI                             Tsp509I
     tttggccgtgatgtttctgatttcactgattccgttcgagatcctaaaacatctgaaatattagacatttcaccttgctcttttggggtgtaagtgtaa
3801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 3900
     aaaccggcactacaaagactaaagtgactaaggcaagctctaggattttgtagactttataatctgtaaagtggaacgagaaaaccccacattcacatt

a      L  A  V  M  F  L  I  S  L  I  P  F  E  I  L  K  H  L  K  Y  *  T  F  H  L  A  L  L  G  V  *  V  *  -
b       W  P  *  C  F  *  F  H  *  F  R  S  R  S  *  N  I  *  N  I  R  H  F  T  L  L  F  W  G  C  K  C  N  -
c     F  G  R  D  V  S  D  F  T  D  S  V  R  D  P  K  T  S  E  I  L  D  I  S  P  C  S  F  G  G  V  S  V  I -
```

Figure 21-P

```
                                                                                                      BstKTI
                                                                                                        DpnI|
                                                                                                      MboI ||
                                                                                                 HpyCH4V | ||
                                                                                                  NlaIII | ||
                                                                                               HpyF10VI | | ||
        TspDTI                                                     HpyCH4V                         BstAPI| | | ||
      BstNI |                                    Eco57I HincII         |               Tsp509I     || | | ||
      ScrFI |                                   Eco57MI   HpaI         |          SfcI      |      || | | ||
     PspGI| |                                Hpy188III|  Hpy8I         |       TspDTI|      | MwoI| | | ||
    StyD4I| |                  Hpy188I           BsgI || MseI|         |    TspRI   ||      |FatI|| | | ||
         || |                     |                  | ||    ||        |        |   ||      | |||  | | ||
      ttacacctggaacaaatgcttcatctgaagttgctgttctatatcaagatgttaactgcactgatgtttctacagcaattcatgcagatcaactcacacc
 3901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4000
      aatgtggaccttgtttacgaagtagacttcaacgacaagatatagttctacaattgacgtgactacaaagatgtcgttaagtacgtctagttgagtgtgg

a     L  H  L  E  Q  M  L  H  L  K  L  L  F  Y  I  K  M  L  T  A  L  M  F  L  Q  Q  F  M  Q  I  N  S  H  Q -
b      Y  T  W  N  K  C  F  I  *  S  C  C  S  I  S  R  C  *  L  H  *  C  F  Y  S  N  S  C  R  S  T  H  T  -
c       T  P  G  T  N  A  S  S  E  V  A  V  L  Y  Q  D  V  N  C  T  D  V  S  T  A  I  H  A  D  Q  L  T  P  -

                                                                                    HincII
                                                                                     Hpy8I
                                                                                     AccI|
                                                  SmlI                               TaqI|
         HhaI                                   HinfI |                  BlpI NlaIII||
      HinP1I |                            Hpy188III | |                  DdeI   NspI||
    AluI   | |                    BpuEI    MlyI  |  | |    CviJI BspCNI  AluI|  SalI2|
   CviJI   | |                    BsrI |   PleI  |  | |  Cac8I |BseMII|  CviJI|FatI |||
       |   | |                       | |     |   |  | |     | |     ||      || |   |||
      agcttggcgcatatattctactggaaacaatgtattccagactcaagcaggctgtcttataggagctgagcatgtcgacacttcttatgagtgcgacatt
 4001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4100
      tcgaaccgcgtatataagatgacctttgttacataaggtctgagttcgtccgacagaatatcctcgactcgtacagctgtgaagaatactcacgctgtaa

a      L  G  A  Y  I  L  L  E  T  M  Y  S  R  L  K  Q  A  V  L  *  E  L  S  M  S  T  L  L  M  S  A  T  F  -
b     S  L  A  H  I  F  Y  W  K  Q  C  I  P  D  S  S  R  L  S  Y  R  S  *  A  C  R  H  F  L  *  V  R  H  S -
c       A  W  R  I  Y  S  T  G  N  N  V  F  Q  T  Q  A  G  C  L  I  G  A  E  H  V  D  T  S  Y  E  C  D  I  -
```

Figure 21-Q

```
                                                                                               ClaI
                                                                                               TaqI
                                                                                            TauI
                                                                                           AciI
                                                                                          BsiEI
                                                                                         Fnu4HI
                                                                                         CviJI
                                                                                        HaeIII
                                                                            BsaBI        TauI
                                                                          EcoRV          BsrI            Tsp509I
                                                                        PstI             EaeI             BanII
                                                                   HpyCH4V               EagI            BsiHKAI
                                                                 HpyF10VI               Fnu4HI           Bsp1286I
                                                                     SfcI                NotI               SacI
                                                                     MwoI                TspRI              AluI
      Cac8I                                                      ApoI                    AciI              CviJI
       AluI                                                     EcoRI               BccI                 EcoICRI
      CviJI                         HpyCH4III      BfaI       Tsp509I           BstXI                      SfaNI
      cctattggagctggcatttgtgctattaccatacagTTAGACTAGAAaagggcgaattctgcagatatccatcacactggcggccgcatcgatgagctct
4101  ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4200
      ggataacctcgaccgtaaacacgataatggtatgtcAATCTGATCTTtcccgcttaagacgtctataggtagtgtgaccgccggcgtagctactcgaga

a      L  L  E  L  A  F  V  L  L  P  Y  S  *  T  R  K  G  R  I  L  Q  I  S  I  T  L  A  A  A  S  M  S  S  -
b       Y  W  S  W  H  L  C  Y  Y  H  T  V  R  L  E  K  G  E  F  C  R  Y  P  S  H  W  R  P  H  R  *  A  L  -
c      P  I  G  A  G  I  C  A  I  T  I  Q  L  D  *  K  R  A  N  S  A  D  I  H  H  T  G  G  R  I  D  E  L  *  -

       MseI
       PacI
  Tsp509I
     AseI
     MseI  CviJI               Tsp509I                 BfaI
                                                       AluI
                                                      CviJI
                                                      MboII                                   Eco57I
                                                                                              Eco57MI
      aattaattaaggcttttctgctaacatatagcaatttttgaagaaaggaaagctagaaatacccaatatcttgctgaagcaagatattgggtatttcttt
4201  ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4300
      ttaattaattccgaaaagacgattgtatatcgttaaaaacttctttcctttcgatctttatgggttatagaacgacttcgttctataacccataaagaaa

a      N  *  L  R  L  F  C  *  H  I  A  I  F  E  E  R  K  A  R  N  T  Q  Y  L  A  E  A  R  Y  W  V  F  L  F  -
b       I  N  *  G  F  S  A  N  I  *  Q  F  L  K  K  G  K  L  E  I  P  N  I  L  L  K  Q  D  I  G  Y  F  F  -
c        L  I  K  A  F  L  L  T  Y  S  N  F  *  R  K  E  S  *  K  Y  P  I  S  C  *  S  K  I  L  G  I  S  F  -

                                       Hin4I
                             ApoI     HincII                                                              ApoI
                            EcoRI      HpaI                          HpyCH4V                             EcoRI
          HinfI            Tsp509I     Hpy8I                          BciVI      Hpy188III              Tsp509I
           TfiI           Hpy188I      MseI  Hpy99I          TspDTI       Hin4I  BssSI                   RsaI
                                                                                                        Csp6I
      tttttttattttgaatctttctattctgaattcagttaacgacgagatttagtatcctttcttgcactttcataactcgtgaaatgccgagttggtacga
4301  ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4400
      aaaaaaataaaacttagaaagataagacttaagtcaattgctgctctaaatcataggaaagaacgtgaaagtattgagcactttacggctcaaccatgct

a       F  L  F  *  I  F  L  F  *  I  Q  L  T  T  R  F  S  I  L  S  C  T  F  I  T  R  E  M  P  S  W  Y  E  -
b      F  F  Y  F  E  S  F  Y  S  E  F  S  *  R  R  D  L  V  S  F  L  A  L  S  *  L  V  K  C  R  V  G  T  N  -
c       F  F  I  L  N  L  S  I  L  N  S  V  N  D  E  I  *  Y  P  F  L  H  F  H  N  S  *  N  A  E  L  V  R  -
```

Figure 21-R

```
                                                                              BsrDI
                                                                         FauI |
                                                            BstUI   CviJI | |
                                                             NruI  HaeIII | |
                                                       Hpy188III|     MscI | |
                  BslI                                    HinfI ||    EaeI | | |        SfaNI
   Tsp509I      EcoNI |                                    TfiI ||  TspDTI| | | | AciI   |
        |           | |                                       |  ||      || | | |   |   |
     attcccccaatttgcgacctaccataggatttgttatgtaaataggtatatgttcctttccattatgaatcgcgattgtatggccaaccattgcgggtag
4401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4500
     taagggggttaaacgctggatggtatcctaaacaatacatttatccatatacaaggaaaggtaatacttagcgctaacataccggttggtaacgcccatc

a     F  P  Q  F  A  T  Y  H  R  I  C  Y  V  N  R  Y  M  F  L  S  I  M  N  R  D  C  M  A  N  H  C  G  *  -
b      S  P  N  L  R  P  T  I  G  F  V  M  *  I  G  I  C  S  F  P  L  *  I  A  I  V  W  P  T  I  A  G  R  -
c     I  P  P  I  C  D  L  P  *  D  L  L  C  K  *  V  Y  V  P  F  H  Y  E  S  R  L  Y  G  Q  P  L  R  V  E -

                 TaiI
             MaeIII |
          HpyCH4IV| |
           NlaIV  || |
          AvaII|  || |
         Sau96I|  || |
           NciI || || |
          ScrFI || || |
           SmaI || || |
         HpaII| || || |
          NciI| || || |
         ScrFI| || || |
          AvaI|| || || |
         BsaJI|| || || |
        StyD4I|| || || |
          XmaI|| || || |
       StyD4I||| || || |                                                                  KpnI
                                                                                        NlaIV |
                                                                                          RsaI |
                                                                      AluI               Csp6I| |
    BsmI  |||| || || |  BseRI                                        CviJI      BccI  Acc65I|| |      MnlI
    BfaI| |||| || || |BsmFI |TspDTI                         MnlI       |    Hin4I  |    BanI|| |  AciI  |
       || |||| || || |    | |    |                            |       |        | |        ||| |     |  |
     aatgctagatgcccgggaccacgttactattgtttctttctcctccttcatattgagcttttctattttgccaataaatgatggtacctcccccccgcca
4501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4600
     ttacgatctacgggccctggtgcaatgataacaaagaaagaggaggaagtataactcgaaaagataaaacggttatttactaccatggagggggggcggt

a     N  A  R  C  P  G  P  R  Y  Y  C  F  F  L  L  L  H  I  E  L  F  Y  F  C  Q  *  M  M  V  P  P  P  A  T -
b      M  L  D  A  R  D  H  V  T  I  V  S  F  S  S  F  I  L  S  F  S  I  F  A  N  K  *  W  Y  L  P  P  P  -
c     C  *  M  P  G  T  T  L  L  L  F  L  S  P  P  S  Y  *  A  F  L  F  L  P  I  N  D  G  T  S  P  R  H  -
```

Figure 21-S

```
         BsiEI
        BstKTI
         Hin4I
          PvuI
          TagI                                                                          CviJI          HaeII
          DpnI|                                                                  BstNI  |           HhaI|
         FauI||                          TaiI               BpmI                 ScrFI  |           AfeI||
         MboI|||     MnlI     CviJI    BmgBI |           Eco57MI               PspGI |  |         HinP1I|||
          ||||        |         |    HpyCH4IV| |        CviJI  |               StyD4I | |   Hpy188I  ||||
          ||||        |         |           || |            |  |                    | | |      |     ||||
     cgatcgaacgggaatggataagaggcttgtgggattgacgtgatagggtagggttggctatactgctggtggcgaactccaggctaataatctgaagcgc
4601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4700
     gctagcttgcccttacctattctccgaacaccctaactgcactatcccatcccaaccgatatgacgaccaccgcttgaggtccgattattagacttcgcg

a     I  E  R  E  W  I  R  G  L  W  D  *  R  D  R  V  G  L  A  I  L  L  V  A  N  S  R  L  I  I  *  S  A  -
b    R  S  N  G  N  G  *  E  A  C  G  I  D  V  I  G  *  G  W  L  Y  C  W  W  R  T  P  G  *  *  S  E  A  L -
c     D  R  T  G  M  D  K  R  L  V  G  L  T  *  *  G  R  V  G  Y  T  A  G  G  E  L  Q  A  N  N  L  K  R  -

                        AlwI
                      NlaIII
                   BsaJI  |
                    BtgI  |
                    FatI  |
                    NcoI  |
                    StyI  |
                 BstKTI|  |
                   DpnI||  |
                  NlaIV||  |
                BamHI1|||  |
                 BstYI |||  |
            Eco57I   | |||  |                                                    BstNI
           Eco57MI   | |||  |                                                    ScrFI
              AlwI|   | |||  |         RsaI                     MseI              PfoI |
          MnlI   ||MboI |||  | HpyCH4III Csp6I|HpyCH4III        AluI |            PspGI |            BfaI
            |    ||    | |||  |        |||  MnlI |           CviJI |            StyD4I |      NlaIV  |
            |    ||    | |||  |        |||    |  |              | |            BseRI| |         |   |
     ttggaagttgtagggaGGATCCATGgcagacaacggtactattaccgttgaggagcttaaacaactcctggaacaatggaacctagtaataggtttccta
4701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4800
     aaccttcaacatccctCCTAGGTACcgtctgttgccatgataatggcaactcctcgaatttgttgaggaccttgttaccttggatcattatccaaaggat

a     W  K  L  *  G  G  S  M  A  D  N  G  T  I  T  V  E  E  L  K  Q  L  L  E  Q  W  N  L  V  I  G  F  L  -
b      G  S  C  R  E  D  P  W  Q  T  T  V  L  L  P  L  R  S  L  N  N  S  W  N  N  G  T  *  *  *  V  S  Y -
c     L  E  V  V  G  R  I  H  G  R  Q  R  Y  Y  Y  R  *  G  A  *  T  T  P  G  T  M  E  P  S  N  R  F  P  I -
```

Figure 21-T

```
                                                                                         MaeIII
                                                                                          BsrI |
                                                                                        CviJI| |
                                                                                       HaeIII| |
                                                                                         MscI| |
                                                                                        EaeI | |
      BstNI                                                                          HpyF10VI| |
      ScrFI                                                                              MwoI| |
     CviJI |                                   RsaI                                      MnlI| |
     PspGI |                                  Csp6I|      AluI                               | |
    StyD4I |                                  BsrGI|     CviJI                               | |
    BfaI | | MaeIII Tsp509I          Hpy188I   TatI|| HindIII |               CviJI |  ||   || |
ttcctagcctggattatgttactacaatttgcctattctaatcggaacaggtttttgtacataataaagcttgttttcctctggctcttgtggccagtaa
4801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 4900
aaggatcggacctaatacaatgatgttaaacggataagattagccttgtccaaaaacatgtattatttcgaacaaaaggagaccgagaacaccggtcatt

a    F  L  A  W  I  M  L  L  Q  F  A  Y  S  N  R  N  R  F  L  Y  I  I  K  L  V  F  L  W  L  L  W  P  V  T -
b     S  *  P  G  L  C  Y  Y  N  L  P  I  L  I  G  T  G  F  C  T  *  *  S  L  F  S  S  G  S  C  G  Q  *  -
c      P  S  L  D  Y  V  T  T  I  C  L  F  *  S  E  Q  V  F  V  H  N  K  A  C  P  P  L  A  L  V  A  S  N -

                         Tsp509I
                           AseI|
                           MseI|
                    Tsp509I ||                                                                   AluI
                     TaqII | ||                                                                 CviJI
                  Hpy8I | | ||  FauI         HphI                                               BlpI |
                   SfcI | | ||MaeIII|    BsrI  |                 BsrDI                          DdeI |
        Cac8I      AccI|  | | ||Tsp45I|   AciI|  |      HpyCH4V CviJI|        CviJI          CviJI| |
cacttgcttgttttgtgcttgctgttgtctacagaattaattgggtgactggcgggattgcgattgcaatggcttgtattgtaggcttgatgtggcttag
4901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5000
gtgaacgaacaaaacacgaacgacaacagatgtcttaattaacccactgaccgccctaacgctaacgttaccgaacataacatccgaactacaccgaatc

a     L  A  C  F  V  L  A  V  V  Y  R  I  N  W  V  T  G  G  I  A  I  A  M  A  C  I  V  G  L  M  W  L  S  -
b    H  L  L  V  L  C  L  L  L  S  T  E  L  I  G  *  L  A  G  L  R  L  Q  W  L  V  L  *  A  *  C  G  L  A -
c     T  C  L  F  C  A  C  C  C  L  Q  N  *  L  G  D  W  R  D  C  D  C  N  G  L  Y  C  R  L  D  V  A  *  -

                                                                                         BslI
                                                                                         NciI
                                  BsrBI                                                 ScrFI
                                 AciI |                                               BsaJI|  MfeI
Eco57I                          RsaI | |                                              HpaII|Tsp509I
Eco57MI            CviJI       Csp6I| | |  FauI                  MboII               StyD4I||MnlI |
ctacttcgttgcttccttcaggctgtttgctcgtacccgctcaatgtggtcattcaacccagaaacaaacattcttctcaatgtgcctctccgggggaca
5001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5100
gatgaagcaacgaaggaagtccgacaaacgagcatgggcgagttacaccagtaagttgggtctttgtttgtaagaagagttacacggagaggccccctgt

a     Y  F  V  A  S  F  R  L  F  A  R  T  R  S  M  W  S  F  N  P  E  T  N  I  L  L  N  V  P  L  R  G  T  -
b      T  S  L  L  P  S  G  C  L  L  V  P  A  Q  C  G  H  S  T  Q  K  Q  T  F  F  S  M  C  L  S  G  G  Q -
c    L  L  R  C  F  L  Q  A  V  C  S  Y  P  L  N  V  V  I  Q  P  R  N  K  H  S  S  Q  C  A  S  P  G  D  N -
```

Figure 21-U

```
MaeIII, Tsp45I, BsmFI, AciI, XcmI, BsrBI, FatI, NlaIII, Hpy8I, BstKTI, DpnI, BclI, MboI, MaeIII, Tsp45I, HpaII, CviJI, HaeIII, EaeI, Bfal, AvrII, BsaJI, StyI, HinP1I, HhaI, HaeII, Tsp45I, MaeIII

     attgtgaccagaccgctcatggaaagtgaacttgtcattggtgctgtgatcattcgtggtcacttgcgaatggccggacactccctagggcgctgtgaca
5101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5200
     taacactggtctggcgagtacctttcacttgaacagtaaccacgacactagtaagcaccagtgaacgcttaccggcctgtgagggatcccgcgacactgt

a    I  V  T  R  P  L  M  E  S  E  L  V  I  G  A  V  I  I  R  G  H  L  R  M  A  G  H  S  L  G  R  C  D  I  -
b     L  *  P  D  R  S  W  K  V  N  L  S  L  V  L  *  S  F  V  V  T  C  E  W  P  D  T  P  *  G  A  V  T  -
c      C  D  Q  T  A  H  G  K  *  T  C  H  W  C  C  D  H  S  W  S  L  A  N  G  R  T  L  P  R  A  L  *  H  -

MseI, Sau96I, PpuMI, Eco0109I, AvaII, BspMI, MboI, DpnI, BstKTI, HpyCH4III, TspRI, CviJI, Hpy188III, XmnI, Tsp509I, HgaI, Fnu4HI, Hpy99I, TseI, BbvI, BbvI, TfiI, HinfI, TspRI, Fnu4HI, TseI

     ttaggacctgccaaaagagatcactgtggctacatcacgaacgctttcttattacaaattaggagcgtcgcagcgtgtaggcactgattcaggtttgc
5201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5300
     aattcctggacggttttctctagtgacaccgatgtagtgcttgcgaaagaataatgtttaatcctcgcagcgtcgcacatccgtgactaagtccaaaacg

a     K  D  L  P  K  E  I  T  V  A  T  S  R  T  L  S  Y  Y  K  L  G  A  S  Q  R  V  G  T  D  S  G  F  A  -
b    L  R  T  C  Q  K  R  S  L  W  L  H  H  E  R  F  L  I  T  N  *  E  R  R  S  V  *  A  L  I  Q  V  L  L -
c     *  G  P  A  K  R  D  H  C  G  Y  I  T  N  A  F  L  L  Q  I  R  S  V  A  A  C  R  H  *  F  R  F  C  -

HpyCH4V, AciI, HpyCH4III, Tsp509I, MseI, BsrFI, BcgI, HpaII, SspI, BfaI, TatI, Csp6I, RsaI, HpyCH4III, BcgI, Tsp509I, EcoRI, ApoI, CviJI, Tsp509I, EcoRI, ApoI

     tgcatacaaccgctaccgtattggaaactataaattaaatacagaccacgccggtagcaacgacaatattgctttgctagtacagtaaGAATTCaagccg
5301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5400
     acgtatgttggcgatggcataacctttgatatttaatttatgtctggtgcggccatcgttgctgttataacgaaacgatcatgtcattCTTAAGttcggc

a     A  Y  N  R  Y  R  I  G  N  Y  K  L  N  T  D  H  A  G  S  N  D  N  I  A  L  L  V  Q  *  E  F  K  P  -
b      H  T  T  A  T  V  L  E  T  I  N  *  I  Q  T  T  P  V  A  T  T  I  L  L  C  *  Y  S  K  N  S  S  R -
c    C  I  Q  P  L  P  Y  W  K  L  *  I  K  Y  R  P  R  R  *  Q  R  Q  Y  C  F  A  S  T  V  R  I  Q  A  E -
```

Figure 21-V

```
                                                 MseI
                                                 PacI
                                            Tsp509I |
                                               AseI||
                                               MseI||
                    AlwI               Tsp509I  |||
                    BfaI                 BanII | |||
                    SpeI|              BsiHKAI | |||
                  Hin4I ||            Bsp1286I | |||
                MaeIII | ||               SacI | |||
                 BsiEI| | ||             AluI | | |||
               CviJI|| | ||             AlwI | | |||
              HaeIII|| | ||            CviJI | | |||
                TauI|| | ||          EcoICRI | | |||
              BsrI ||| | ||        Hpy188I | | | |||
              EaeI ||| | ||       BstKTI | | | | |||
              EagI ||| | ||        DpnI| | | | | |||
            Fnu4HI ||| | ||       NlaIV| | | | | |||
             TspRI ||| | ||BamHI2|| | | | | | |||                                     BfaI
              AciI| ||| | ||BstYI || | | | | | |||CviJI                                 AluI|
BceAI   BstXI   || ||| | || MboI || | | | | | |||Hin4I        Tsp509I              CviJI|
  |       |     || ||| | ||  |  || | | | | | |||  |              |                 MboII|
aattccagcacactggcggccgttactagtggatccgagctctaattaattaaggcttttctgctaacatatagcaattttgaagaaaggaaagctaga
5401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5500
ttaaggtcgtgtgaccgccggcaatgatcacctaggctcgagattaattaattccgaaaagacgattgtatatcgttaaaacttctttcctttcgatct

a    N  S  S  T  L  A  A  V  T  S  G  S  E  L  *  L  I  K  A  F  L  L  T  Y  S  N  F  *  R  K  E  S  *  K -
b     I  P  A  H  W  R  P  L  L  V  D  P  S  S  N  *  L  R  L  F  C  *  H  I  A  I  F  E  E  R  K  A  R  -
c      F  Q  H  T  G  G  R  Y  *  W  I  R  A  L  I  N  *  G  F  S  A  N  I  *  Q  F  L  K  K  G  K  L  E  -

                                                                       Hin4I
                                                           ApoI       HincII
                                                          EcoRI         HpaI
                         Eco57I             HinfI       Tsp509I        Hpy8I
                        Eco57MI              TfiI       Hpy188I|        MseI| Hpy99I
                           |                  |             ||            ||     |
aatacccaatatcttgctgaagcaagatattgggtatttcttttttttttatttgaatctttctattctgaattcagttaacgacgagatttagtatcc
5501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5600
ttatgggttatagaacgacttcgttctataacccataaagaaaaaaaaaataaacttagaaagataagacttaagtcaattgctgctctaaatcatagg

a     Y  P  I  S  C  *  S  K  I  L  G  I  S  F  F  F  I  L  N  L  S  I  L  N  S  V  N  D  E  I  *  Y  P -
b    N  T  Q  Y  L  A  E  A  R  Y  W  V  F  L  F  F  L  F  *  I  F  L  F  *  I  Q  L  T  T  R  F  S  I  L -
c     I  P  N  I  L  L  K  Q  D  I  G  Y  F  F  F  F  Y  F  E  S  F  Y  S  E  F  S  *  R  R  D  L  V  S  -
```

Figure 21-W

```
                                       ApoI
                                       EcoRI
     HpyCH4V                           Tsp509I
      BciVI|      Hpy188III            RsaI |                        BslI
  TspDTI  ||Hin4I  BssSI |           Csp6I| |  Tsp509I         EcoNI |
      |   ||    |     | |                 || |        |               | |
     tttcttgcactttcataactcgtgaaatgccgagttggtacgaattcccccaatttgcgacctaccataggatttgttatgtaaataggtatatgttcct
5601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5700
     aaagaacgtgaaagtattgagcactttacggctcaaccatgcttaagggggttaaacgctggatggtatcctaaacaatacatttatccatatacaagga

a     F  L  H  F  H  N  S  *  N  A  E  L  V  R  I  P  P  I  C  D  L  P  *  D  L  L  C  K  *  V  Y  V  P  -
b      S  C  T  F  I  T  R  E  M  P  S  W  Y  E  F  P  Q  F  A  T  Y  H  R  I  C  Y  V  N  R  Y  M  F  L  -
c     F  L  A  L  S  *  L  V  K  C  R  V  G  T  N  S  P  N  L  R  P  T  I  G  F  V  M  *  I  G  I  C  S  F -

                                                                     TaiI
                                                                 MaeIII |
                                                              HpyCH4IV| |
                                                              NlaIV  || |
                                                             AvaII|  || |
                                                            Sau96I|  || |
                                                             NciI || || |
                                                            ScrFI || || |
                                                             SmaI || || |
                                                           HpaII| || || |
                                                            NciI| || || |
                                   BsrDI                   ScrFI| || || |
                                  FauI |                   AvaI|| || || |
              BstUI       CviJI  | |                     BsaJI|| || || |
               NruI      HaeIII  | |                    StyD4I|| || || |
        Hpy188III|         MscI  | |                      XmaI|| || || |
          HinfI ||        EaeI | | |      SfaNI     BsmI   StyD4I||| || || | BseRI                   AluI
           TfiI ||     TspDTI| | | |  AciI    |     BfaI|   |||| || || |BsmFI |TspDTI       MnlI  CviJI
              | ||          || | | |     |    |        ||   |||| || || |   |  |    |             |   |
     ttccattatgaatcgcgattgtatggccaaccattgcgggtagaatgctagatgcccgggaccacgttactattgtttctttctcctccttcatattgag
5701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5800
     aaggtaatacttagcgctaacataccggttggtaacgcccatcttacgatctacgggccctggtgcaatgataacaaagaaagaggaggaagtataactc

a     F  H  Y  E  S  R  L  Y  G  Q  P  L  R  V  E  C  *  M  P  G  T  T  L  L  L  F  L  S  P  P  S  Y  *  A -
b      S  I  M  N  R  D  C  M  A  N  H  C  G  *  N  A  R  C  P  G  P  R  Y  Y  C  F  F  L  L  L  H  I  E  -
c       P  L  *  I  A  I  V  W  P  T  I  A  G  R  M  L  D  A  R  D  H  V  T  I  V  S  F  S  S  F  I  L  S  -
```

Figure 21-X

```
                                         BsiEI
                                         BstKTI
                                         Hin4I
                                         PvuI
                     KpnI                TaqI
                   NlaIV                 DpnI
                    RsaI                 FauI
                   Csp6I                 MboI                                              TaiI
        BccI     Acc65I              MnlI                                               BmgBI
   Hin4I           BanI       AciI                   MnlI        CviJI        HpyCH4IV                  CviJI
     ctttctattttgccaataaatgatggtacctccccccgccacgatcgaacgggaatggataagaggcttgtgggattgacgtgatagggtagggttgg
5801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 5900
     gaaaagataaaaacggttatttactaccatggaggggggcggtgctagcttgcccttacctattctccgaacaccctaactgcactatcccatcccaacc

a     F  L  F  L  P  I  N  D  G  T  S  P  R  H  D  R  T  G  M  D  K  R  L  V  G  L  T  *  *  G  R  V  G  -
b    L  F  Y  F  C  Q  *  M  M  V  P  P  P  A  T  I  E  R  E  W  I  R  G  L  W  D  *  R  D  R  V  G  L  A -
c     F  S  I  F  A  N  K  *  W  Y  L  P  P  P  R  S  N  G  N  G  *  E  A  C  G  I  D  V  I  G  *  G  W   -

                                                            BstKTI
                                                            DpnI
                                                            NlaIV                    HpyF10VI
                                                            BamHI3                AciI
                                                            BstYI            HpyF10VI
                      CviJI              HaeII                                 MwoI
                    BstNI                HhaI        Eco57I          Hpy188III
                    ScrFI                AfeI        Eco57MI          BssSI                           HinfI
   BpmI            PspGI               HinP1I        AlwI             CviJI                    MlyI  TaqI
   Eco57MI         StyD4I          Hpy188I      MnlI       MboI        AlwI               MwoI  PleI
     ctatactgctggtggcgaactccaggctaataatctgaagcgcttggaagttgtagggaggatcctatggctcgtgaagcggttatcgccgaagtatcga
5901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6000
     gatatgacgaccaccgcttgaggtccgattattagacttcgcgaaccttcaacatccctcctaggataccgagcacttcgccaatagcggcttcatagct

a     Y  T  A  G  G  E  L  Q  A  N  N  L  K  R  L  E  V  V  G  R  I  L  W  L  V  K  R  L  S  P  K  Y  R   -
b      I  L  L  V  A  N  S  R  L  I  I  *  S  A  W  K  L  *  G  G  S  Y  G  S  *  S  G  Y  R  R  S  I  D  -
c    L  Y  C  W  W  R  T  P  G  *  *  S  E  A  L  G  S  C  R  E  D  P  M  A  R  E  A  V  I  A  E  V  S  T -
```

Figure 21-Y

```
MnlI, Hpy188I, HgaI, BsaHI, TaqI, HinP1I, HhaI, HaeII, Hpy188III, TaqI, BceAI, BccI, HpyCH4IV, Hpy99I, TaiI, Cac8I, HaeIII, CviJI, EaeI, Csp6I, RsaI, Csp6I, RsaI, CviJI, NlaIV, AciI, BccI, TspRI, BtsI, BceAI, AciI, Fnu4HI, BstF5I, TauI, HaeIII, CviJI, FokI, CviJI

     ctcaactatcagaggtagttggcgtcatcgagcgccatctcgaaccgacgttgctggccgtacatttgtacggctccgcagtggatggcggcctgaagcc
6001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6100
     gagttgatagtctccatcaaccgcagtagctcgcggtagagcttggctgcaacgaccggcatgtaaacatgccgaggcgtcacctaccgccggacttcgg

a     L  N  Y  Q  R  *  L  A  S  S  S  A  I  S  N  R  R  C  W  P  Y  I  C  T  A  P  Q  W  M  A  A  *  S  H  -
b      S  T  I  R  G  S  W  R  H  R  A  P  S  R  T  D  V  A  G  R  T  F  V  R  L  R  S  G  W  R  P  E  A   -
c       Q  L  S  E  V  V  G  V  I  E  R  H  L  E  P  T  L  L  A  V  H  L  Y  G  S  A  V  D  G  G  L  K  P  -

HpyCH4III, DraIII, TspRI, Eco57MI, Eco57I, MaeIII, Tsp45I, MaeIII, HpyCH4III, BstEII, HpyCH4III, CviJI, HphI, Hin4I, BstUI, AciI, Fnu4HI, TauI, TspDTI, Cac8I, CviJI, AluI, MboI, BclI, DpnI, BstKTI, Hin4I, CviJI, StyD4I, PspGI, BsaJI, ScrFI, BstNI

     acacagtgatattgatttgctggttacggtgaccgtaaggcttgatgaaacaacgcggcgagctttgatcaacgaccttttggaaacttcggcttccccct
6101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6200
     tgtgtcactataactaaacgaccaatgccactggcattccgaactactttgttgcgccgctcgaaactagttgctggaaaacctttgaagccgaaggggga

a      T  V  I  L  I  C  W  L  R  *  P  *  G  L  M  K  Q  R  G  E  L  *  S  T  T  F  W  K  L  R  L  P  L  -
b     T  Q  *  Y  *  F  A  G  Y  G  D  R  K  A  *  *  N  N  A  A  S  F  D  Q  R  P  F  G  N  F  G  F  P  W  -
c     H  S  D  I  D  L  L  V  T  V  T  V  R  L  D  E  T  T  R  R  A  L  I  N  D  L  L  E  T  S  A  S  P  -
```

Figure 21-Z

```
                                                                                  Tsp509I
                                        Hpy99I                                   HpyCH4V
                 HphI                  TspGWI                                 HpyF10VI
                 SfcI            Bme1580I                                   BstUI
                HhaI             BsiHKAI                                    HhaI
               BpmI              Bsp1286I                                HinP1I
              BstUI               Hpy8I                                 BlpI
             Eco57MI             HpyCH4V                                DdeI
     HinfI   HinP1I   MaeIII     ApaLI               BsaJI             AluI
     TfiI    AciI     Tsp45I    MslI                 BtgI              CviJI      MwoI
     ggagagagcgagattctccgcgctgtagaagtcaccattgttgtgcacgacgacatcattccgtggcgttatccagctaagcgcgaactgcaatttggag
6201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6300
     cctctctcgctctaagaggcgcgacatcttcagtggtaacaacacgtgctgctgtagtaaggcaccgcaataggtcgattcgcgcttgacgttaaacctc

a     E  R  A  R  F  S  A  L  *  K  S  P  L  L  C  T  T  T  S  F  R  G  V  I  Q  L  S  A  N  C  N  L  E  -
b      R  E  R  D  S  P  R  C  R  S  H  H  C  C  A  R  R  H  H  S  V  A  L  S  S  *  A  R  T  A  I  W  R  -
c     G  E  S  E  I  L  R  A  V  E  V  T  I  V  V  H  D  D  I  I  P  W  R  Y  P  A  K  R  E  L  Q  F  G  E -

                                                BsiEI
                                               BstKTI
        HhaI              HpyCH4V       CviJI   PvuI              CviJI
      HinP1I     BbvI                  Cac8I    TaqI     BstKTI
     Fnu4HI     BsrDI                 CviJI    DpnI      DpnI                                     BsaJI
      TseI     BspMI   MboII        TaqI      MboI       MboI                                      StyI
     aatggcagcgcaatgacattcttgcaggtatcttcgagccagccacgatcgacattgatctggctatcttgctgacaaaagcaagagaacatagcgttgc
6301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6400
     ttaccgtcgcgttactgtaagaacgtccatagaagctcggtcggtgctagctgtaactagaccgatagaacgactgttttcgttctcttgtatcgcaacg

a     N  G  S  A  M  T  F  L  Q  V  S  S  S  Q  P  R  S  T  L  I  W  L  S  C  *  Q  K  Q  E  N  I  A  L  P -
b      M  A  A  Q  *  H  S  C  R  Y  L  R  A  S  H  D  R  H  *  S  G  Y  L  A  D  K  S  K  R  T  *  R  C  -
c      W  Q  R  N  D  I  L  A  G  I  F  E  P  A  T  I  D  I  D  L  A  I  L  L  T  K  A  R  E  H  S  V  A  -

                                  Hpy188III
                 AciI              NlaIV
                 TauI            HpaII
               Fnu4HI           BsaWI               BstKTI
                AciI           BstKTI                MnlI          HaeII
              MspA1I             DpnI               DpnI           HhaI                                   TauI
       AvaII                    EciI              BstYI         HinP1I                                   AciI
      Sau96I  MnlI        AlwI  MboI              MboI    AlwI                    MseI  TspDTI           Fnu4HI
     cttggtaggtccagcggcggaggaactctttgatccggttcctgaacaggatctatttgaggcgctaaatgaaaccttaacgctatggaactcgccgccc
6401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6500
     gaaccatccaggtcgccgcctccttgagaaactaggccaaggacttgtcctagataaactccgcgatttactttggaattgcgataccttgagcggcggg

a      W  *  V  Q  R  R  R  N  S  L  I  R  F  L  N  R  I  Y  L  R  R  *  M  K  P  *  R  Y  G  T  R  R  P  -
b     L  G  R  S  S  G  G  G  T  L  *  S  G  S  *  T  G  S  I  *  G  A  K  *  N  L  N  A  M  E  L  A  A  R -
c      L  V  G  P  A  A  E  E  L  F  D  P  V  P  E  Q  D  L  F  E  A  L  N  E  T  L  T  L  W  N  S  P  P  -
```

Figure 21-AA

```
                                                        HpaII
                                                       BsrFI|
                                                  MaeIII   ||
                                                  HhaI |   ||
                                              HinP1I | |   ||
         BmrI                               RsaI  | | |   ||          BbvI
      Cac8I |                             Csp6I|  | | |   ||          HhaI        Fnu4HI   BglI
      BsrI | |                  TaiI      FauI ||  | | |   ||        BstUI |         TseI|    MwoI
     CviJI | |      BsmFI HpyCH4IV | AciI BslI | ||  | | |   ||       HinP1I |      BstF5I|| FokI |
         | | |          |         | |    |    | | ||  | | |   ||            | |           |||    | |
     gactgggctggcgatgagcgaaatgtagtgcttacgttgtcccgcatttggtacagcgcagtaaccggcaaaatcgcgccgaaggatgtcgctgccgact
6501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6600
     ctgacccgaccgctactcgctttacatcacgaatgcaacagggcgtaaaccatgtcgcgtcattggccgttttagcgcggcttcctacagcgacggctga

a     T  G  L  A  M  S  E  M  *  C  L  R  C  P  A  F  G  T  A  Q  *  P  A  K  S  R  R  R  M  S  L  P  T  -
b      L  G  W  R  *  A  K  C  S  A  Y  V  V  P  H  L  V  Q  R  S  N  R  Q  N  R  A  E  G  C  R  C  R  L  -
c     D  W  A  G  D  E  R  N  V  V  L  T  L  S  R  I  W  Y  S  A  V  T  G  K  I  A  P  K  D  V  A  A  D  W -

                                                                                                   DpnI
                                                                                                 MboI |
                                                                                                 MnlI |
                                                                                               HhaI | |
                                                                                             BstUI | | |
                      BsrI                                                                   Cac8I | | |
                    CviJI |                                                                   HhaI | | |
                   HaeIII |                                                                 HinP1I | | |
                   Sau96I| |                                                              HpyF10VI | | |
                    Cac8I|| |                                                                 MwoI| | | |
                     NaeI|| |                                                              BssHII|| | | |
                    HpaII||| |                                                               BstUI|| | | |
                     BmrI||| |                                                             HinP1I|| | | |
                   BsrFI|||| |                                                         CviJI   ||| | | |
                  NgoMIV|||| |                                                        HaeIII   ||| | | |
                  Cac8I ||||| |                                      CviJI              MboII   ||| | | |
                  HaeII| ||||| |                               HpyF10VI |               MmeI |   ||| | | |
                  HhaI|| ||||| |                                  BfaI | |             MboII| |   ||| | | |
              HinP1I ||| ||||| |                                 AluI| | |          BstKTI || |   ||| | | |
         BsrI  BsrDI| ||| ||||| |                              CviJI| MwoI| |         DpnI| || |   ||| | | |
  HpyF10VI | BmrI || ||| ||||| |       CviJI                    || || |          MboI || || |   ||| | | |
         | |    | || ||| ||||| |           |                    || || |            | || || |   ||| | | |
     gggcaatggagcgcctgccggcccagtatcagcccgtcatacttgaagctagacaggcttatcttggacaagaagaagatcgcttggcctcgcgcgcaga
6601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6700
     cccgttacctcgcggacggccgggtcatagtcgggcagtatgaacttcgatctgtccgaatagaacctgttcttcttctagcgaaccggagcgcgcgtct

a     G  Q  W  S  A  C  R  P  S  I  S  P  S  Y  L  K  L  D  R  L  I  L  D  K  K  K  I  A  W  P  R  A  Q  I -
b      G  N  G  A  P  A  G  P  V  S  A  R  H  T  *  S  *  T  G  L  S  W  T  R  R  R  S  L  G  L  A  R  R  -
c       A  M  E  R  L  P  A  Q  Y  Q  P  V  I  L  E  A  R  Q  A  Y  L  G  Q  E  E  D  R  L  A  S  R  A  D -
```

Figure 21-BB

```
                                   TaiI                             Hpy188III
                                BsaAI |             BcgI                 TaqI|
                               DraIII |             DdeI|                AvaI||
                             HpyCH4IV||        BstKTI  ||                SmlI||
              ApoI        MboII  || |           DpnI|  ||               XhoI||              Tsp509I
    BstKTI  Tsp509I     Hpy8I|   || |        MboI ||   ||         BsmAI  |||   DdeI        BcgI  |
         |        |         ||   || |           | ||   ||             |  |||      |          |   |
         tcagttggaagaatttgtccactacgtgaaaggcgagatcactaaggtagttggcaaataactcgagactaagtggataaaattagatagaaaaaaggtc
    6701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6800
         agtcaaccttcttaaacaggtgatgcactttccgctctagtgattccatcaaccgtttattgagctctgattcacctattttaatctatcttttttccag

a          S  W  K  N  L  S  T  T  *  K  A  R  S  L  R  *  L  A  N  N  S  R  L  S  G  *  N  *  I  E  K  R  S  -
b         S  V  G  R  I  C  P  L  R  E  R  R  D  H  *  G  S  W  Q  I  T  R  D  *  V  D  K  I  R  *  K  K  G  L -
c         Q  L  E  E  F  V  H  Y  V  K  G  E  I  T  K  V  V  G  K  *  L  E  T  K  W  I  K  L  D  R  K  K  V  -

                                   BstKTI                   Tsp509I
                                    DpnI|                 Hin4I  |                                  MseI
                                  MboI ||               HpyCH4V  |                          Tsp509I  |         BsmAI
              EarI       Hin4I MboII | ||   MaeIII        MboII  |  MboII        Tsp509I HpyCH4V|   |         BsmBI
                 |           |     | | ||        |            |  |     |              |        ||   |             |
         taaataaaaaagaagagaaatagaaagatcaaaaatcagttacgaaaatgcagtaattcttcttttttcttctaattgattgcaattaaactcgtctcaa
    6801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 6900
         atttatttttttcttctctttatctttctagtttttagtcaatgcttttacgtcattaagaagaaaaaagaagattaactaacgttaatttgagcagagtt

a         K  *  K  R  R  E  I  E  R  S  K  I  S  Y  E  N  A  V  I  L  L  F  S  S  N  *  L  Q  L  N  S  S  Q  -
b          N  K  K  E  E  K  *  K  D  Q  K  S  V  T  K  M  Q  *  F  F  F  F  L  L  I  D  C  N  *  T  R  L  N  -
c         *  I  K  K  K  R  N  R  K  I  K  N  Q  L  R  K  C  S  N  S  S  F  F  F  *  L  I  A  I  K  L  V  S  I -

                                                NlaIII
                                           Hpy188III |
                                               BspHI||
                                              BstKTI||
                                                FatI||
                                                DpnI||              HinfI
                              DraI            BsmAI|||               TfiI
      Hpy188I        CviJI    MseI|            MboI||||      Hpy188III  |              MnlI              MboII
            |            |       ||               |||||  |           |  |                 |                  |
         tctgaaaaaagattgagccgagtttaaatagatttgatacgatcatgagacttgacaaatcgggattcctctattctatatttagaagatataaagg
    6901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7000
         agacttttttctaactcggctcaaatttatctaaactatgctagtactctgaactgtttagccctaaggagataagatatataaatcttctatatttcc

a         S  E  K  R  L  S  R  V  *  I  D  F  D  T  I  M  R  L  D  K  S  G  F  L  Y  S  I  Y  L  E  D  I  K  V -
b          L  K  K  D  *  A  E  F  K  *  I  L  I  R  S  *  D  L  T  N  R  D  S  S  I  L  Y  I  *  K  I  *  R  -
c           *  K  K  I  E  P  S  L  N  R  F  *  Y  D  H  E  T  *  Q  I  G  I  P  L  F  Y  I  F  R  R  Y  K  G  -
```

Figure 21-CC

```
                                        Bsu36I
                                   TaqI     DdeI
                           BfaI    AlwI| BstKTI |
                      Hpy188III   AluI|| DpnI| |
                           XbaI| CviJI||MboI || |                               SfcI                  HinfI
                               ||    |||    | || |                                 |                  TfiI  CviJI
                                                                                                          |      |
      tataatacaataaataaatacaatctagaagctcgatgatccttaggattggtatattcttttctatcctgtagtttgtagtttccctgaatcaagccaa
 7001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7100
      atattatgttatttatttatgttagatcttcgagctactaggaatcctaaccatataagaaaagataggacatcaaacatcaaagggacttagttcggtt

a      *  Y  N  K  *  I  Q  S  R  S  S  M  I  L  R  I  G  I  F  F  S  I  L  *  F  V  V  S  L  N  Q  A  K  -
b     Y  N  T  I  N  K  Y  N  L  E  A  R  *  S  L  G  L  V  Y  S  F  L  S  C  S  L  *  F  P  *  I  K  P  S -
c      I  I  Q  *  I  N  T  I  *  K  L  D  D  P  *  D  W  Y  I  L  F  Y  P  V  V  C  S  F  P  E  S  S  Q  -

                     BstF5I
                     BsmFI|                                          ApoI
        FokI          MnlI|     BccI TspGWI              Tsp509I     Tsp509I
           |             ||        |     |                  MseI|     MseI |
                                                                ||         | |
      gtatcacacctctttctacccatcctgtatattgtcccctttgttccgtgttgaaatagaaccttaatttattacttatttttttattaaattttagatt
 7101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7200
      catagtgtggagaaagatgggtaggacatataacaggggaaacaaggcacaactttatcttggaattaaataatgaataaaaaaataatttaaaatctaa

a     Y  H  T  S  F  Y  P  S  C  I  L  S  P  L  F  R  V  E  I  E  P  *  F  I  T  Y  F  F  I  K  F  *  I  -
b      I  T  P  L  S  T  H  P  V  Y  C  P  L  C  S  V  L  K  *  N  L  N  L  L  L  I  F  L  L  N  F  R  F -
c     V  S  H  L  F  L  P  I  L  Y  I  V  P  F  V  P  C  *  N  R  T  L  I  Y  Y  L  F  F  Y  *  I  L  D  L -

                                                                                                  ApoI
                                                                                               Tsp509I
                                                             MnlI                              Hin4I |
                                           Tsp509I          Hin4I |        SfaNI      TaqI        | |
                                                 |              | |            |         |        | |
      tgttagtgattagatattagtattagacgagattttacgaaacaattatttttttatttctttataggagaggacaaatctcttttttcgatgcgaattt
 7201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7300
      acaatcactaatctataatcataatctgctctaaaatgctttgttaataaaaaaataaagaaatatcctctcctgtttagagaaaaaagctacgcttaaa

a     C  *  *  L  D  I  S  I  R  R  D  F  T  K  Q  L  F  F  Y  F  F  I  G  E  D  K  S  L  F  S  M  R  I  * -
b      V  S  D  *  I  L  V  L  D  E  I  L  R  N  N  Y  F  F  I  S  L  *  E  R  T  N  L  F  F  R  C  E  F  -
c       L  V  I  R  Y  *  Y  *  T  R  F  Y  E  T  I  I  F  L  F  L  Y  R  R  G  Q  I  S  F  F  D  A  N  L -

                   TauI
                  AciI |
                Fnu4HI |    Tsp509I         DraI                     AseI
                CviJI| |    MseI  |   FalI  MseI|         NlaIV      MseI               MmeI
                    || |       |  |      |     ||             |         |                  |
      gacacgacataggagaagccgcccttattaaaaattatattattttaaataatataaaggggggttccaacatattaatatatagtgaagtgttcccccca
 7301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7400
      ctgtgctgtatcctcttcggcgggaaataatttttaatataataaaatttattatatttccccccaaggttgtataattatatatcacttcacaaggggggt

a      H  D  I  G  E  A  A  L  Y  *  K  L  Y  Y  F  K  *  Y  K  G  G  S  N  I  L  I  Y  S  E  V  F  P  Q  -
b     D  T  T  *  E  K  P  P  F  I  K  N  Y  I  I  L  N  N  I  K  G  V  P  T  Y  *  Y  I  V  K  C  S  P  R -
c      T  R  H  R  R  S  R  P  L  L  K  I  I  L  F  *  I  I  *  R  G  F  Q  H  I  N  I  *  *  S  V  P  P  -
```

Figure 21-DD

```
  Hpy188I
 HinfI  |                                                                  Hpy188I
  TfiI  |                                   MseI     BfaI          DdeI       |
     |  |                                      |        |             |       |
     gattcagaacttttttcaatactcacaatccttattagttaataatcctagtgattggatttctatgcttagtctgataggaaataagatattcaaata
7401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7500
     ctaagtcttgaaaaaagttatgagtgttaggaataatcaattattaggatcactaacctaaagatacgaatcagactatcctttattctataagtttat

a     I  Q  N  F  F  S  I  L  T  I  L  I  S  *  *  S  *  *  L  D  F  Y  A  *  S  D  R  K  *  D  I  Q  I  -
b      F  R  T  F  F  Q  Y  S  Q  S  L  L  V  N  N  P  S  D  W  I  S  M  L  S  L  I  G  N  K  I  F  K  *  -
c     D  S  E  L  F  F  N  T  H  N  P  Y  *  L  I  I  L  V  I  G  F  L  C  L  V  *  *  E  I  R  Y  S  N  K -

                                          HpyCH4V
                                           NlaIII
            TspDTI                        Hin4I |                                Tsp509I
Tsp509I  Hin4I  |         TspDTI          FatI | |                   BccI          MseI|TaqI      MseI
      |      |  |              |             | | |                      |              || |          |
     aataatttatagcgaatgactattcatctattgtattttcatgcaaataggggcaagaaaactctatggaaagatggtggtttaattcgatgttgttt
7501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7600
     ttattaaatatcgcttactgataagtagataacataaaagtacgtttatccccgttcttttgagatacctttctaccaccaaattaagctacaacaaa

a    N  N  F  I  A  N  D  Y  S  S  I  V  F  S  C  K  *  G  A  R  K  L  Y  G  K  M  V  V  *  F  D  V  V  * -
b     I  I  L  *  R  M  T  I  H  L  L  Y  F  H  A  N  R  G  Q  E  N  S  M  E  R  W  W  F  N  S  M  L  F  -
c      *  F  Y  S  E  *  L  F  I  Y  C  I  F  M  Q  I  G  G  K  K  T  L  W  K  D  G  G  L  I  R  C  C  L  -

         BstBI                                                               BsaBI    TspDTI
          TaqI                                                            BstKTI |    MmeI |
        AarI |                                                              DpnI| |TspDTI | |
       BspMI |                                         AvaII             BstYI || |MboII| | |
       XmnI| |              CviJI                      Sau96I       AlwI MboI || |TaqI|| | |
          || |                  |                          |           |    | || |  ||| | |
     aagaaggagttcgaacgcaggtgtgggctaaataaatcaatgggcagtcttggtcctattgaaaataccaatgaagatccaaatcgaaaagtgaaaaaca
7601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7700
     ttcttcctcaagcttgcgtccacacccgatttatttagttacccgtcagaaccaggataacttttatggttacttctaggtttagcttttcactttttgt

a     E  G  V  R  T  Q  V  W  A  K  *  I  N  G  Q  S  W  S  Y  *  K  Y  Q  *  R  S  K  S  K  S  E  K  H  -
b    K  K  E  F  E  R  R  C  G  L  N  K  S  M  G  S  L  G  P  I  E  N  T  N  E  D  P  N  R  K  V  K  N  I -
c     R  R  S  S  N  A  G  V  G  *  I  N  Q  W  A  V  L  V  L  L  K  I  P  M  K  I  Q  I  E  K  *  K  T  -

                    Tsp509I
          Hpy188III       |
             MaeIII       |                                                 ApoI
             Tsp45I |     |   HpyCH4V                                    Tsp509I
   MnlI     HinfI   |     |    BsaXI |                                  Hpy188I|
 BsaXI |     TfiI   |     |  BfaI  | |                        MseI TspDTI    ||     Hpy188I
     | |        |   |     |     |  | |                           |      |    ||           |
     ttcatagttggaggaatcgtgacaattctagttgcagtaatgttgattatttattcggcgttaaagacattcggaatttcatctctgatgacactttttt
7701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7800
     aagtatcaacctccttagcactgttaagatcaacgtcattacaactaataaataagccgcaatttctgtaagccttaaagtagagactactgtgaaaaaa

a     S  *  L  E  E  S  *  Q  F  *  L  Q  *  C  *  L  F  I  R  R  *  R  H  S  E  F  H  L  *  *  H  F  F  -
b      H  S  W  R  N  R  D  N  S  S  C  S  N  V  D  Y  L  F  G  V  K  D  I  R  N  F  I  S  D  D  T  F  L -
c     F  I  V  G  G  I  V  T  I  L  V  A  V  M  L  I  I  Y  S  A  L  K  T  F  G  I  S  S  L  M  T  L  F  * -
```

Figure 21-EE

```
                                                                    BspCNI
                                                                   BseMII|
                                                                  BstKTI||
                                                                   DpnI|||  Hpy188I    BfaI
        BsmAI  HpyCH4III          BccI                            MboI ||||   DdeI|Hpy8I  |
     agttagtgataggaatggagacagttattccatctattttgatattgaaaatcatattttgagattgacaacgatcattcttttctgagtgaactagaa
7801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 7900
     tcaatcactatccttacctctgtcaataaggtagataaaactataacttttagtataaaaactctaactgttgctagtaagaaaagactcacttgatctt

a    S  *  *  *  E  W  R  Q  L  F  H  L  F  *  Y  *  K  S  Y  F  *  D  *  Q  R  S  F  F  S  E  *  T  R  K -
b     V  S  D  R  N  G  D  S  Y  S  I  Y  F  D  I  E  N  H  I  F  E  I  D  N  D  H  S  F  L  S  E  L  E  -
c      L  V  I  G  M  E  T  V  I  P  S  I  L  I  L  K  I  I  F  L  R  L  T  T  I  I  L  F  *  V  N  *  K -

                                                                          NlaIII
                                                                            NspI
                                                         BstKTI           MmeI|
                                                           DpnI|      AflIII||
              Tsp509I                                    BstYI||  Tsp509I FatI||
          TaqI  TaqI|Hpy188I                       AlwI MboI||  MboII|  PciI||
     agttcttttatagttatcgaaactcgaattatcggaataatggatttaggggcgaagatccctactataattcttacatgtatgatactcaatatagtt
7901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8000
     tcaagaaaaatatcaatagctttgagcttaatagccttattacctaaatccccgcttctagggatgatattaagaatgtacatactatgagttatatcaa

a      F  F  L  *  L  S  K  L  E  L  S  E  *  W  I  *  G  R  R  S  L  L  *  F  L  H  V  *  Y  S  I  *  L -
b    S  S  F  Y  S  Y  R  N  S  N  Y  R  N  N  G  F  R  G  E  D  P  Y  Y  N  S  Y  M  Y  D  T  Q  Y  S  W -
c     V  L  F  I  V  I  E  T  R  I  I  G  I  M  D  L  G  A  K  I  P  T  I  I  L  T  C  M  I  L  N  I  V  -

                                                                                               BceAI
                                                                                           HpyCH4III
                                                                                              MaeIII
             HpyCH4V                                                               HpyCH4III  |
         AseI Eco57I|    BsaBI                                                        MaeIII  |
         MseI Eco57MI|MboII  |          BsmAI                  PsiI        Tsp509I Tsp45I     |
     ggaataatcacattaatagttgcattgatagttatcttcagtctcaaatctgtatagatacttccattataagtggtagtgagaattacggtgacagtta
8001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8100
     ccttattagtgtaattatcaacgtaactatcaatagaagtcagagtttagacatatctatgaaggtaatattcaccatcactcttaatgccactgtcaat

a     E  *  S  H  *  *  L  H  *  *  L  S  S  V  S  N  L  Y  R  Y  F  H  Y  K  W  *  *  E  L  R  *  Q  L  -
b      N  N  H  I  N  S  C  I  D  S  Y  L  Q  S  Q  I  C  I  D  T  S  I  I  S  G  S  E  N  Y  G  D  S  Y -
c    G  I  I  T  L  I  V  A  L  I  V  I  F  S  L  K  S  V  *  I  L  P  L  *  V  V  V  R  I  T  V  T  V  T -

          CviJI
         HaeIII                    HphI                      Hpy8I                         MseI
     HphI Sau96I|               TaqI  |     MnlI         BsrI   |          HpyF10VI        BtsI|
        |     ||                  |   |       |      NlaIV |AccI|             MwoI|       TspRI|
     catttatagggccgtttgtggtggtgaaagtcgaaatagtagtgaaaacgagggttccagtagacgaactcgcacgaagggcagtgatttaactataaga
8101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8200
     gtaaatatcccggcaaacaccaccactttcagctttatcatcacttttgctcccaaggtcatctgcttgagcgtgcttcccgtcactaaattgatattct

a    H  L  *  G  R  L  W  W  *  K  S  K  *  *  *  K  R  G  F  Q  *  T  N  S  H  E  G  Q  *  F  N  Y  K  R -
b     I  Y  R  A  V  C  G  G  E  S  R  N  S  S  E  N  E  G  S  S  R  R  T  R  T  K  G  S  D  L  T  I  R  -
c      F  I  G  P  F  V  V  V  K  V  E  I  V  V  K  T  R  V  P  V  D  E  L  A  R  R  A  V  I  *  L  *  E -
```

Figure 21-FF

```
                         MaeIII
                       TaqI    |
                      AvaI|    |
                 Hpy188III|    |
                      SmlI|    |
                      XhoI|    |
                   BstKTI||    |
                     DpnI|||   |
                    MboI ||||  |                                                          ApoI
                    MnlI ||||  |                         Tsp509I              Tsp509I  Tsp509I
                         |||||  |                          |                  MseI | PsiI  |
       gaaagttctaatgatctcgaggtaactcaaaaatacaggcatttgtgggttcaatgcgaaaattgttatggattaaattataagaattttgaaatcaa
  8201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8300
       ctttcaagattactagagctccattgagtttttatgtccgtaaacacccaagttacgcttttaacaatacctaatttaatattcttaaaaactttagtt

a        K  F  *  *  S  R  G  N  S  K  I  Q  A  F  V  G  S  M  R  K  L  L  W  I  K  L  *  E  I  F  E  I  K  -
b       E  S  S  N  D  L  E  V  T  Q  K  Y  R  H  L  W  V  Q  C  E  N  C  Y  G  L  N  Y  K  K  F  L  K  S  K -
c        K  V  L  M  I  S  R  *  L  K  N  T  G  I  C  G  F  N  A  K  I  V  M  D  *  I  I  R  N  F  *  N  Q   -

                                                        ApoI
                                                        EcoRI
              TspDTI                                   Tsp509I      NlaIII   AluI
        SspI  Hpy8I |     EcoRV                Hpy188I   |          FatI  |  CviJI        Tsp509I
          |     |   |       |                       |     |              |  |    |             |
       aaatgaatatttgtgaacaatgtggatatcatttgaaaatgagtagttcagatagaattcgtaatcatggtcatagctgtttcctgtgtgaaattgttat
  8301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8400
       tttacttataaacacttgttacacctatagtaaacttttactcatcaagtctatcttaagcattagtaccagtatcgacaaaggacacactttaacaata

a       N  E  Y  L  *  T  M  W  I  S  F  E  N  E  *  F  R  *  N  S  *  S  W  S  *  L  F  P  V  *  N  C  Y   -
b        M  N  I  C  E  Q  C  G  Y  H  L  K  M  S  S  S  D  R  I  R  N  H  G  H  S  C  F  L  C  E  I  V  I -
c       K  *  I  F  V  N  N  V  D  I  I  *  K  *  V  V  Q  I  E  F  V  I  M  V  I  A  V  S  C  V  K  L  L  S -

                                                           NlaIV
                                                       HpyF10VI|
                                                     BstNI     ||
                                                     ScrFI     ||                                HpyF10VI
                                                    BsaJI|     ||                                   MwoI|
         Tsp509I                                    CviJI||    ||                     Tsp509I       BtsI ||
      BsrBI  |                  HpaII               PspGI||BanI||          AluI       AseI|       HhaI | ||
      AciI |  |                 CviJI|              StyD4I||MwoI||          CviJI      MseI|   HinP1I | | ||
         | |  |                     ||                   |||   |||              |          ||       | | | ||
       ccgctcacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcac
  8401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8500
       ggcgagtgttaaggtgtgttgtatgctcggccttcgtatttcacatttcggaccccacggattactcactcgattgagtgtaattaacgcaacgcgagtg

a       P  L  T  I  P  H  N  I  R  A  G  S  I  K  C  K  A  W  G  A  *  *  V  S  *  L  T  L  I  A  L  R  S  L -
b        R  S  Q  F  H  T  T  Y  E  P  E  A  *  S  V  K  P  G  V  P  N  E  *  A  N  S  H  *  L  R  C  A  H   -
c         A  H  N  S  T  Q  H  T  S  R  K  H  K  V  *  S  L  G  C  L  M  S  E  L  T  H  I  N  C  V  A  L  T  -
```

Figure 21-GG

```
                                                          AciI
                                                          BstUI
                                                          HhaI
                                                          MnlI
                                                          BslI
                                          AseI            TspDTI
                                          MseI            BstUI
                                   HpyCH4V                HinP1I
                                   Fnu4HI                 BsaXI
                                   AluI              FauI                          HaeII
                                   CviJI             CviJI                         HhaI
                                   MspA1I            HaeIII                        BsaXI
   AciI   Hpy188III                PvuII             EaeI                          HinP1I
   Cac8I  BsrI                     TseI       HinfI                               HpyF10VI   EarI
   TspRI  FauI        BbvI         Cac8I      TfiI                  AciI           MwoI       SapI
                                                                                   MboII      AciI
     tgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttc
8501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8600
     acgggcgaaaggtcagccctttggacagcacggtcgacgtaattacttagccggttgcgcgcccctctccgccaaacgcataacccgcgagaaggcgaag

a     P  A  F  Q  S  G  N  L  S  C  Q  L  H  *  *  I  G  Q  R  A  G  R  G  G  L  R  I  G  R  S  S  A  S  -
b    C  P  L  S  S  R  E  T  C  R  A  S  C  I  N  E  S  A  N  A  R  G  E  A  V  C  V  L  G  A  L  P  L  P -
c     A  R  F  P  V  G  K  P  V  V  P  A  A  L  M  N  R  P  T  R  G  E  R  R  F  A  Y  W  A  L  F  R  F  -

                                           AluI
                                           CviJI
                                        HpyF10VI
                 BbvI                  AciI
                 HhaI                  BsrBI
              HinP1I                  Cac8I
             Fnu4HI                  TauI
             TseI                  Fnu4HI
          TaqII                    AciI
   MlyI   TspRI                  Fnu4HI
   PleI HinfI                    CviJI                                                 HinfI
  BbvI  MnlI            BsiEI    TseI            MwoI          AciI  HpyCH4III         TfiI
     ctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataac
8601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8700
     gagcgagtgactgagcgacgcgagccagcaagccgacgccgctcgccatagtcgagtgagtttccgccattatgccaataggtgtcttagtcccctattg

a    S  L  T  D  S  L  R  S  V  V  R  L  R  R  A  V  S  A  H  S  K  A  V  I  R  L  S  T  E  S  G  D  N  -
b     R  S  L  T  R  C  A  R  S  F  G  C  G  E  R  Y  Q  L  T  Q  R  R  *  Y  G  Y  P  Q  N  Q  G  I  T -
c    L  A  H  *  L  A  A  L  G  R  S  A  A  A  S  G  I  S  S  L  K  G  G  N  T  V  I  H  R  I  R  G  *  R -
```

Figure 21-HH

```
                                        HpyCH4III
                                          NlaIV |          Cac8I
                                      BstNI  |  |    HpyF10VI
                                      ScrFI  |  |     BstUI |
                                      CviJI |  |  |    TauI |
                                      HaeIII |  |  |   AciI | |
          NlaIII                       PspGI |  |  |  Fnu4HI | |
            NspI                      StyD4I |  |  |  CviJI| | |               AciI
     AflIII  |        Cac8I               |  |  |  |  HaeIII| | |              NlaIV |
       FatI  |        CviJI |            |  |  |  |     |  | | |               CviJI| |
       FciI  |       HaeIII | BslI       |  |  |  | BslI || |  MwoI|   EciI    || |
        |    |          |  |   |        |  |  |  |   |  || |    ||     |      || |
     gcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagc
8701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8800
     cgtcctttcttgtacactcgttttccggtcgttttccggtccttggcatttttccggcgcaacgaccgcaaaaaggtatccgaggcggggggactgctcg

a    A  G  K  N  M  *  A  K  G  Q  Q  K  A  R  N  R  K  K  A  A  L  L  A  F  F  H  R  L  R  P  P  D  E  H -
b     Q  E  R  T  C  E  Q  K  A  S  K  R  P  G  T  V  K  R  P  R  C  W  R  F  S  I  G  S  A  P  L  T  S  -
c      R  K  E  H  V  S  K  R  P  A  K  G  Q  E  P  *  K  G  R  V  A  G  V  F  P  *  A  P  P  P  *  R  A -

                 Hpy188I
                 HgaI |
              DrdI |  |                                                                         MnlI
              MnlI |  |                                                     BstNI        HhaI |
             SmlI |  |  |                                   BstNI           ScrFI     HinP1I | |
           Hpy99I| |  |  |                                  ScrFI          BsaJI |  BssSI  |  | |
        TaqI    || |  |  |                                  PspGI |         PspGI | AluI  |    | | |
 BpuEI SfaNI |  || |  |  |                                 StyD4I |        StyD4I | CviJI  |   | | |
   |     |  |  || |  |  |                                     | |            | |    |     |    | | |
     atcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgt
8801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 8900
     tagtgtttttagctgcgagttcagtctccaccgctttgggctgtcctgatatttctatggtccgcaaagggggaccttcgagggagcacgcgagaggaca

a     H  K  N  R  R  S  S  Q  R  W  R  N  P  T  G  L  *  R  Y  Q  A  F  P  P  G  S  S  L  V  R  S  P  V  -
b    I  T  K  I  D  A  Q  V  R  G  G  E  T  R  Q  D  Y  K  D  T  R  R  F  P  L  E  A  P  S  C  A  L  L  F -
c     S  Q  K  S  T  L  K  S  E  V  A  K  P  D  R  T  I  K  I  P  G  V  S  P  W  K  L  P  R  A  L  S  C  -

                EciI
               HpaII|
              BsaWI||
               BslI||
             BciVI |||
              TauI| |||
            AciI || |||                                          HaeII
 Hpy188I  Fnu4HI || |||   MmeI    AciI        Hpy188III           HhaI|     AluI                              BseMII
   |          |  ||  |||     |     |                |        HinP1I || CviJI   SfcI      DdeI          BspCNI|
   |          |  ||  |||     |     |                |            | ||   |       |         |                ||
     tccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtag
8901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9000
     aggctgggacggcgaatggcctatggacaggcggaaagagggaagcccttcgcaccgcgaaagagtatcgagtgcgacatccatagagtcaagccacatc

a     P  T  L  P  L  T  G  Y  L  S  A  F  L  P  S  G  S  V  A  L  S  H  S  S  R  C  R  Y  L  S  S  V  *  -
b      R  P  C  R  L  P  D  T  C  P  P  F  S  L  R  E  A  W  R  F  L  I  A  H  A  V  G  I  S  V  R  C  R -
c    S  D  P  A  A  Y  R  I  P  V  R  L  S  P  F  G  K  R  G  A  F  S  *  L  T  L  *  V  S  Q  F  G  V  G -
```

Figure 21-II

```
                                              HhaI
              Bme1580I                    HinP1I                               HpaII
                 BsiHKAI                 Fnu4HI                                MlyI
          CviJI Bsp1286I                MspA1I                                 NciI
      AluI       Hpy8I                    TseI     MaeIII          HinfI       ScrFI
      BseYI     HpyCH4V          CviJI  AciI      HpaII    Hpy188III           PleI
      CviJI     ApaLI      BbvI        BsiEI      BsaWI          SmlI    StyD4I       BpuEI
     gtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagac
9001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9100
     cagcaagcgaggttcgacccgacacacgtgcttggggggcaagtcgggctggcgacgcggaataggccattgatagcagaactcaggttgggccattctg

a    V  V  R  S  K  L  G  C  V  H  E  P  P  V  Q  P  D  R  C  A  L  S  G  N  Y  R  L  E  S  N  P  V  R  H -
b     S  F  A  P  S  W  A  V  C  T  N  P  P  F  S  P  T  A  A  P  Y  P  V  T  I  V  L  S  P  T  R  *  D  -
c      R  S  L  Q  A  G  L  C  A  R  T  P  R  S  A  R  P  L  R  L  I  R  *  L  S  S  *  V  Q  P  G  K  T  -

                BbvI
                BsrI
               TspRI
               BbvI
             MaeIII
            AlwNI
          CviJI
        Fnu4HI
         TseI
        BsrI
       Fnu4HI
        TspRI                                                                  CviJI     CviJI
  MmeI   TseI                         MnlI              AciI  SfcI Hpy188III   HaeIII    BslI
     acgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacgg
9101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9200
     tgctgaatagcggtgaccgtcgtcggtgaccattgtcctaatcgtctcgctccatacatccgccacgatgtctcaagaacttcaccaccggattgatgcc

a     D  L  S  P  L  A  A  A  T  G  N  R  I  S  R  A  R  Y  V  G  G  A  T  E  F  L  K  W  W  P  N  Y  G -
b    T  T  Y  R  H  W  Q  Q  P  L  V  T  G  L  A  E  R  G  M  *  A  V  L  Q  S  S  *  S  G  G  L  T  T  A -
c     R  L  I  A  T  G  S  S  H  W  *  Q  D  *  Q  S  E  V  C  R  R  C  Y  R  V  L  E  V  V  A  *  L  R  -

                                                                         HpaII
                                                                       BstKTI
                                                                         DpnI
                                                                        MboI
                                                                  Hpy188III
                                          MaeIII    Eco57I            AluI
       HpyCH4III           HhaI           BsrI     Eco57MI            AlwI                   MspA1I
   BfaI  BceAI           HinP1I          CviJI   Hpy188I             CviJI                   AciI
     ctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgct
9201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9300
     gatgtgatcttcctgtcataaaccatagacgcgagacgacttcggtcaatggaagcctttttctcaaccatcgagaactaggccgtttgtttggtggcga

a     Y  T  R  R  T  V  F  G  I  C  A  L  L  K  P  V  T  F  G  K  R  V  G  S  S  *  S  G  K  Q  T  T  A  -
b      T  L  E  G  Q  Y  L  V  S  A  L  C  *  S  Q  L  P  S  E  K  E  L  V  A  L  D  P  A  N  K  P  P  L -
c     L  H  *  K  D  S  I  W  Y  L  R  S  A  E  A  S  Y  L  R  K  K  S  W  *  L  L  I  R  Q  T  N  H  R  W -
```

Figure 21-JJ

```
                                                                BstKTI
                                                                 DpnI|
                                                               BstYI ||
                                                                MboI ||
                                                             AlwI  | ||
                                          BpuEI            AlwI |  | ||
                                           HhaI       Hpy188III |  | ||
                                          BbvI |           SmlI | | ||
                               Fnu4HI     BstUI |         BstKTI| | | ||  BstKTI
                                 TseI|    HinP1I |         DpnI|| | | ||   DpnI|                DdeI
                               Cac8I ||  HpyF10VI ||      BstYI ||| | | ||  MboII|                PpiI|
        AciI                HpyCH4V | ||   MwoI| | |       MboI ||| | | || MboI ||     Hpy188I   || HgaI
          |                         | | ||       || | |        | ||| | | || | ||          |    || |
        ggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatctttgatcttttctacggggtctgacgctcagt
   9301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9400
        ccatcgccaccaaaaaaacaaacgttcgtcgtctaatgcgcgtcttttttcctagagttcttctaggaaactagaaaagatgccccagactgcgagtca

a        G  S  G  G  F  F  V  C  K  Q  Q  I  T  R  R  K  K  G  S  Q  E  D  P  L  I  F  S  T  G  S  D  A  Q  W -
b         V  A  V  V  F  L  F  A  S  S  R  L  R  A  E  K  K  D  L  K  K  I  L  *  S  F  L  R  G  L  T  L  S  -
c          *  R  W  F  F  C  L  Q  A  A  D  Y  A  Q  K  K  R  I  S  R  R  S  F  D  L  F  Y  G  V  *  R  S  V -

                                                                  BstKTI
                   MseI                                 BstKTI     DpnI|
                   TaiI                                  DpnI|   BstYI ||
             HpyCH4IV |          NlaIII                BstYI ||   MboI ||       MseI
           BseMII  |  |      Hpy188III |                MboI ||   BfaI | ||Tsp509I |
          BspCNI|  |  |         BspHI| |             HphI | || AlwI | | || DraI| |               DraI
        TspRI || |  |      PpiI  FatI| |         MboII | | ||AlwI| | | || MseI|| |          MseI|TspDTI
          |   ||  |  |         |    || |             |  | | ||  || | | ||    ||| |              ||   |
        ggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatcta
   9401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9500
        ccttgcttttgagtgcaattccctaaaaccagtactctaatagtttttcctagaagtggatctaggaaaatttaatttttacttcaaaatttagttagat

a         N  E  N  S  R  *  G  I  L  V  M  R  L  S  K  R  I  F  T  *  I  L  L  N  *  K  *  S  F  K  S  I  *  -
b        G  T  K  T  H  V  K  G  F  W  S  *  D  Y  Q  K  G  S  S  P  R  S  F  *  I  K  N  E  V  L  N  Q  S  K -
c         E  R  K  L  T  L  R  D  F  G  H  E  I  I  K  K  D  L  H  L  D  P  F  K  L  K  M  K  F  *  I  N  L  -

                                                                       BseMII
                                                                      BspCNI|
                                                                    TspDTI ||
                                                                     FokI | ||
                      HpyCH4III                    NlaIV         BstKTI| | ||                 MlyI
                        MaeIII           MnlI      BanI |          DpnI|| | ||                 PleI
              Hpy8I  Hpy188I |          MseI |     TspRI |    DdeI MboI ||| | ||  BstF5I  Hin4I  | HinfI
                  |       |  |            |  |         | |      |     | |||  | ||     |      |   |    |
        aagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctga
   9501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9600
        ttcatatatactcatttgaaccagactgtcaatggttacgaattagtcactccgtggatagagtcgctagacagataaagcaagtaggtatcaacggact

a         S  I  Y  E  *  T  W  S  D  S  Y  Q  C  L  I  S  E  A  P  I  S  A  I  C  L  F  R  S  S  I  V  A  *  -
b          V  Y  M  S  K  L  G  L  T  V  T  N  A  *  S  V  R  H  L  S  Q  R  S  V  Y  F  V  H  P  *  L  P  D -
c        K  Y  I  *  V  N  L  V  *  Q  L  P  M  L  N  Q  *  G  T  Y  L  S  D  L  S  I  S  F  I  H  S  C  L  T -
```

Figure 21-KK

```
                                              BsrI                                       BpmI
                                             BccI |                                    Eco57MI
                                            NlaIV |                                 BstUI   |
                                           CviJI| |                                AciI |   |                    Hpy188III
                                         HaeIII|  | HpyCH4V     BsaI | |   |                     NlaIV |
                                        Sau96I||  |  TspRI|    BsmAI | |   |                    CviJI| |
            Hpy99I           MnlI       BmrI |||  |Fnu4HI||    BsrDI | |   |                  HpaII  || |
        AhdI   |            Hin4I |.      CviJI   BbvI | |||  | TseI|||  Hin4I| | | HphI          BsrFI|  || |Hin4I
        ctccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttat
   9601 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9700
        gaggggcagcacatctattgatgctatgccctcccgaatggtagaccggggtcacgacgttactatggcgctctgggtgcgagtggccgaggtctaaata

a       L  P  V  V  *  I  T  T  I  R  E  G  L  P  S  G  P  S  A  A  M  I  P  R  D  P  R  S  P  A  P  D  L  S -
b        S  P  S  C  R  *  L  R  Y  G  R  A  Y  H  L  A  P  V  L  Q  *  Y  R  E  T  H  A  H  R  L  Q  I  Y  -
c         P  R  R  V  D  N  Y  D  T  G  G  L  T  I  W  P  Q  C  C  N  D  T  A  R  P  T  L  T  G  S  R  F  I -

                              HhaI
                           HinP1I |
                         CviJI  | |
                        HaeIII  | |
                     HpyF10VI | | |
                   HpaII    | | | |                 FokI                  Tsp509I        NciI
                  CviJI|Sau96I | | |       HpyCH4V   |                  BccI    |       ScrFI   BfaI
                Cac8I || BglI| | | |    AvaII      |  |       BstF5I    MnlI|AseI|      HpaII|  AluI|
              CviJI | || MwoI| | | |   Sau96I EciI|  |    AciI    |BsrI  ||MseI|   StyD4I||  CviJI|
        cagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagt
   9701 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9800
        gtcgttatttggtcggtcggccttcccggctcgcgtcttcaccaggacgttgaaataggcggaggtaggtcagataattaacaacggcccttcgatctca

a         A  I  N  Q  P  A  G  R  A  E  R  R  S  G  P  A  T  L  S  A  S  I  Q  S  I  N  C  C  R  E  A  R  V -
b       Q  Q  *  T  S  Q  P  E  G  P  S  A  E  V  V  L  Q  L  Y  P  P  P  S  S  L  L  I  V  A  G  K  L  E  * -
c        S  N  K  P  A  S  R  K  G  R  A  Q  K  W  S  C  N  F  I  R  L  H  P  V  Y  *  L  L  P  G  S  *  S  -

                              TaiI
                           AclI |                                                                  NlaIV
                       HpyCH4IV |                     MaeIII                                    HpaII |
                          HhaI| |                      SfaNI                                    BsaWI| |
                 MseI     FspI| |                     Tsp45I              TspDTI                AluI || |
             BsrI  |    HinP1I|| |  BsrDI  SfcI      MslI   |           Hpy99I |    CviJI      CviJI || |
        aagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggt
   9801 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 9900
        ttcatcaagcggtcaattatcaaacgcgttgcaacaacggtaacgatgtccgtagcaccacagtgcgagcagcaaaccataccgaagtaagtcgaggcca

a        S  S  S  P  V  N  S  L  R  N  V  V  A  I  A  T  G  I  V  V  S  R  S  S  F  G  M  A  S  F  S  S  G  -
b         V  V  R  Q  L  I  V  C  A  T  L  L  P  L  L  Q  A  S  W  C  H  A  R  R  L  V  W  L  H  S  A  P  V -
c       K  *  F  A  S  *  *  F  A  Q  R  C  C  H  C  Y  R  H  R  G  V  T  L  V  V  W  Y  G  F  I  Q  L  R  F -
```

Figure 21-LL

```
BstKTI DpnI MboI MaeIII FatI AlwI BstKTI DpnI NlaIII MboI FatI NlaIII HpyCH4V
AciI TaqII AluI CviJI AvaII Sau96I Hpy188I MboI DpnI PvuI BstKTI BsiEI MnlI Hpy188I
EaeI HaeIII CviJI Fnu4HI AciI TauI

     tcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcag
9901 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 10000
     agggttgctagttccgctcaatgtactagggggtacaacacgttttttcgccaatcgaggaagccaggaggctagcaacagtcttcattcaaccggcgtc

a     S  Q  R  S  R  R  V  T  *  S  P  M  L  C  K  K  A  V  S  S  F  G  P  P  I  V  V  R  S  K  L  A  A  V -
b      P  N  D  Q  G  E  L  H  D  P  P  C  C  A  K  K  R  L  A  P  S  V  L  R  S  L  S  E  V  S  W  P  Q  -
c       P  T  I  K  A  S  Y  M  I  P  H  V  V  Q  K  S  G  *  L  L  R  S  S  D  R  C  Q  K  *  V  G  R  S -

BtsI TspRI FatI NlaIII MslI Fnu4HI TseI BtsI HpyCH4V TspRI Tsp509I BbvI FokI TspGWI HpyCH4III
FatI NlaIII SfaNI BstF5I BccI MaeIII Tsp45I BsrI TatI Csp6I ScaI RsaI HphI BseMII BspCNI

      tgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcatt
10001 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 10100
      acaatagtgagtaccaataccgtcgtgacgtattaagagaatgacagtacggtaggcattctacgaaaagacactgaccactcatgagttggttcagtaa

a      L  S  L  M  V  M  A  A  L  H  N  S  L  T  V  M  P  S  V  R  C  F  S  V  T  G  E  Y  S  T  K  S  F -
b     C  Y  H  S  W  L  W  Q  H  C  I  I  L  L  L  S  C  H  P  *  D  A  F  L  *  L  V  S  T  Q  P  S  H  S -
c      V  I  T  H  G  Y  G  S  T  A  *  F  S  Y  C  H  A  I  R  K  M  L  F  C  D  W  *  V  L  N  Q  V  I  -

Hpy188I DdeI BcgI Fnu4HI AciI TauI BsiEI HgaI StyD4I TaqII ScrFI NciI HpaII BsaHI BcgI
AciI HinP1I BstUI HhaI DraI MseI BsiHKAI Bsp1286I MboII

      ctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattgga
10101 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 10200
      gactcttatcacatacgccgctggctcaacgagaacgggccgcagttatgccctattatggcgcggtgtatcgtcttgaaattttcacgagtagtaacct

a      *  E  *  C  M  R  R  P  S  C  S  C  P  A  S  I  R  D  N  T  A  P  H  S  R  T  L  K  V  L  I  I  G  -
b       E  N  S  V  C  G  D  R  V  A  L  A  R  R  Q  Y  G  I  I  P  R  H  I  A  E  L  *  K  C  S  S  L  E -
c     L  R  I  V  Y  A  A  T  E  L  L  L  P  G  V  N  T  G  *  Y  R  A  T  *  Q  N  F  K  S  A  H  H  W  K -
```

Figure 21-MM

```
                                                                                    MboII
                                                                                Bme1580I |
                                                                                 BsiHKAI |
                                                                                Bsp1286I |
                                                                                   Hpy8I | |
                                                                                 HpyCH4V | |
        BpuEI                                   AlwI      BsrI                   ApaLI | | |
      TaiI |                        BstKTI MspA1I |    BstKTI|                  Eco57I | | |
     AclI | |                        DpnI| AciI | |     DpnI||                 Eco57MI | | | |   BstKTI
  HpyCH4IV | |                     BstYI || AlwI | |  |BstYI |||   MaeIII      BssSI | | | |    DpnI|          SfaNI
      XmnI | |                  SmlI MboI ||PpiI| |  | MboI |||  TaqI  |      PpiI | | | | |  MboI ||TaqII  HphI |
         | | |                     |   | ||   || |  |    | |||    |   |        |  | | | | |     | ||   |     |  |
         aaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttta
   10201 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 10300
         tttgcaagaagccccgcttttgagagttcctagaatggcgacaactctaggtcaagctacattgggtgagcacgtgggttgactagaagtcgtagaaat

a        K  R  S  S  G  R  K  L  S  R  I  L  P  L  L  R  S  S  S  M  *  P  T  R  A  P  N  *  S  S  A  S  F  T -
b         N  V  L  R  G  E  N  S  Q  G  S  Y  R  C  *  D  P  V  R  C  N  P  L  V  H  P  T  D  L  Q  H  L  L  -
c          T  F  F  G  A  K  T  L  K  D  L  T  A  V  E  I  Q  F  D  V  T  H  S  C  T  Q  L  I  F  S  I  F  Y -

                                                TauI
                                            AciI |
            TaqII                 HphI     Fnu4HI |                        MslI         MboII
              |                     |          | |                           |          TspGWI           EarI
              |                     |          | |                           |             |               |
         ctttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcct
   10301 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 10400
         gaaagtggtcgcaaagacccactcgtttttgtccttccgttttacggcgttttttcccttattcccgctgtgcctttacaacttatgagtatgagaagga

a          F  T  S  V  S  G  *  A  K  T  G  R  Q  N  A  A  K  K  G  I  R  A  T  R  K  C  *  I  L  I  L  F  L  -
b         L  S  P  A  F  L  G  E  Q  K  Q  E  G  K  M  P  Q  K  R  E  *  G  R  H  G  N  V  E  Y  S  Y  S  S  F -
c          F  H  Q  R  F  W  V  S  K  N  R  K  A  K  C  R  K  K  G  N  K  G  D  T  E  M  L  N  T  H  T  L  P  -

                                             AciI
                                            BsrBI
                                          NlaIII |
                                          BciVI | |
                                          BsmAI | |                                                    HhaI
                                       Hpy188III| | |                                                BstUI |
                                          BspHI|| | |                                               HinP1I |
            SspI                           FatI|| | |                                                AciI | |
              |                                ||| | |                                             NlaIV | | |
              |                                ||| | |                                                 | | | |
         ttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaataggggttccgcgcaca
   10401 ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+ 10500
         aaaagttataataacttcgtaaatagtcccaataacagagtactcgcctatgtataaacttacataaatctttttatttgtttatccccaaggcgcgtgt

a         F  Q  Y  Y  *  S  I  Y  Q  G  Y  C  L  M  S  G  Y  I  F  E  C  I  *  K  N  K  Q  I  G  V  P  R  T  -
b          F  N  I  I  E  A  F  I  R  V  I  V  S  *  A  D  T  Y  L  N  V  F  R  K  I  N  K  *  G  F  R  A  H -
c         F  S  I  L  L  K  H  L  S  G  L  L  S  H  E  R  I  H  I  *  M  Y  L  E  K  *  T  N  R  G  S  A  H  I -
```

Figure 21-NN

```
                                                                       CviJI
                   DdeI                                               HaeIII
                 AatII|                                            EcoO109I|
                  TaiI|                NlaIII                        Sau96I|
                 ZraI ||          Hpy188III  |                    BssSI   ||
               BsaHI| ||             BspHI|  |                Hpy188III   ||
            HpyCH4IV| ||              FatI|  |MseI                MnlI |  ||
                    || ||                 ||  |  |                   |  |  ||
       tttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgtc
 10501 ---------+---------+---------+---------+---------+---------+---------+---------+---------+- 10591
       aaaggggcttttcacggtggactgcagattctttggtaataatagtactgtaattggatatttttatccgcatagtgctccgggaaagcag

a      F  P  R  K  V  P  P  D  V  *  E  T  I  I  I  M  T  L  T  Y  K  N  R  R  I  T  R  P  F  R   -
b       F  P  E  K  C  H  L  T  S  K  K  P  L  L  S  *  H  *  P  I  K  I  G  V  S  R  G  P  F  V  -
c        S  P  K  S  A  T  *  R  L  R  N  H  Y  Y  H  D  I  N  L  *  K  *  A  Y  H  E  A  L  S    -
```

Enzymes that do cut and were not excluded:

AarI	AatII	AccI	Acc65I	AciI	AclI	AfeI	AflIII	AgeI	AhdI	AluI	AlwI	AlwNI
ApaI	ApaLI	ApoI	AseI	AvaI	AvaII	AvrII	BaeI	BamHI	BanI	BanII	BbeI	BbsI
BbvI	BccI	BceAI	BcgI	BciVI	BclI	BfaI	BfrBI	BglI	BglII	BlpI	Bme1580I	BmgBI
BmrI	BpmI	Bpu10I	BpuEI	BsaI	BsaAI	BsaBI	BsaHI	BsaJI	BsaWI	BsaXI	BseMII	BseRI
BseYI	BsgI	BsiEI	BsiHKAI	BslI	BsmI	BsmAI	BsmBI	BsmFI	Bsp1286I	BspCNI	BspHI	BspMI
BsrI	BsrBI	BsrDI	BsrFI	BsrGI	BssHII	BssSI	BstAPI	BstBI	BstEII	BstF5I	BstKTI	BstNI
BstUI	BstXI	BstYI	Bsu36I	BtgI	BtsI	Cac8I	ClaI	Csp6I	CviJI	DdeI	DpnI	DraI
DraIII	DrdI	EaeI	EagI	EarI	EciI	Eco57I	EcoICRI	Eco57MI	EcoNI	EcoO109I	EcoRI	EcoRV
FalI	FatI	FauI	Fnu4HI	FokI	FspI	HaeII	HaeIII	HgaI	HhaI	Hin4I	HinP1I	HincII
HindIII	HinfI	HpaI	HpaII	HphI	Hpy8I	Hpy99I	Hpy188I	Hpy188III	HpyCH4III	HpyCH4IV	HpyCH4V	HpyF10VI
KasI	KpnI	MaeIII	MboI	MboII	MfeI	MlyI	MmeI	MnlI	MscI	MseI	MslI	MspA1I
MwoI	NaeI	NarI	NciI	NcoI	NdeI	NgoMIV	NlaIII	NlaIV	NotI	NruI	NsiI	NspI
PacI	PciI	PfoI	PleI	PpiI	PpuMI	PsiI	PspGI	PspOMI	PstI	PvuI	PvuII	RsaI
SacI	SacII	SalI	SapI	Sau96I	ScaI	ScrFI	SexAI	SfaNI	SfcI	SfoI	SmaI	SmlI
SnaBI	SpeI	SphI	SspI	StyI	StyD4I	TaiI	TaqI	TaqII	TatI	TauI	TfiI	TseI
Tsp45I	Tsp509I	TspDTI	TspGWI	TspRI	XbaI	XcmI	XhoI	XmaI	XmnI	ZraI		

Enzymes that do not cut:

AflII	AleI	AloI	AscI	AsiSI	BbvCI	BmtI	BplI	BsiWI	BspEI	BstZ17I	FseI	FspAI
MluI	NheI	PflMI	PmeI	PmlI	PshAI	PsrI	RsrII	SanDI	SbfI	SfiI	SgrAI	SrfI
StuI	SwaI	Tth111I										

Enzymes excluded; MinCuts: 1 MaxCuts: 100000

Figure 22-A

pCV6 sequence

```
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttgaaaga gataaattga 420
acaagtatgg tcgtcccctg ttgggatgta ctattaaacc taaattgggg ttatctgcta 480
aaaactacgg tagagccgtt tatgaatgtc ttcgcggtgg acttgatttt actaaagatg 540
atgagaacgt gaactcacaa ccatttatgc gttggagaga tcgtttctta ttttgtgccg 600
aagcacttta taaagcacag gctgaaacag gtgaaatcaa agggcattac ttgaatgcta 660
ctgcaggtac atgcgaagaa atgatcaaaa gagctgtatt tgctagagaa ttgggcgttc 720
cgatcgtaat gcatgactac ttaacggggg gattcaccgc aaatactagc ttggctcatt 780
attgccgaga taatggtcta cttcttcaca tccaccgtgc aatgcatgcg gttattgata 840
gacagaagaa tcatggtatc cacttccggg tattagcaaa agcgttacgt atgtctggtg 900
gagatcatat tcactctggt accgtagtag gtaaacttga aggtgaaaga gacataactt 960
tgggctttgt tgatttactg cgtgatgatt ttgttgaaca agatcgaagt cgcggtattt 1020
atttcactca agattgggtc tctttaccag gtgttctacc cgtggcttca ggaggtattc 1080
acgtttggca tatgcctgct ctgaccgaga tctttgggga tgattccgta ctacagttcg 1140
gtggaggaac tttaggacat ccttggggta atgcgccagg tgccgtagct aatcgagtag 1200
ctctagaagc atgtgtaaaa gctcgtaatg aaggacgtga tcttgctcag gaaggtaatg 1260
aaattattcg cgaggcttgc aaatggagcc cggaactagc tgctgcttgt gaagtatgga 1320
aagagatcgt atttaatttt gcagcagtgg acgttttgga taagtaaaaa cagtagacat 1380
tagcagataa attagcagga aataaagaag gataaggaga aagaactcaa gtaattatcc 1440
ttcgttctct taattgaatt gcaattaaac tcggcccaat cttttactaa aaggattgag 1500
ccgaatacaa caaagattct attgcatata ttttgactaa gtatatactt acctagatat 1560
acaagatttg aaatacaaaa tctagaaaac taaatcaaaa tctaagactc aaatctttct 1620
attgttgtct tggatcatcg agcttgcatg cctgcagtcg actttcacag tttccattct 1680
gaaatgttct ctgtactata ataaatagta agtgaatcaa cttactacta aaaaaattag 1740
tagacttcct cttcggaata gaaatagcct atttctacat agggaaagtc gtgtgcaatg 1800
aaaaatgcaa gcacgatttg ggagaggtt ttttctctat tgtaacaagg aataattatc 1860
tactccatcc gactagttcc gggttcgagt cccgggcaac ccatatggaa actagaaagg 1920
agcaatctga gttttgattt ttcactcact tcatttacaa aattttttgg tttggtaaat 1980
tttgttgtat ggatatacaa ctgtcggggc tggcttggtt gacattggta tatagtctat 2040
attatactgt taaataacaa gccttctatt atctttctag ttaatacgtg tgcttgggag 2100
tccttgcaat ttgaataaac caagatctta ccatggccgc gcggggttct catcatcatc 2160
atcatggtcc gcgggggccc gctagtggta gtgaccttga ccggtgcacc acttttgatg 2220
atgttcaagc tcctaattac actcaacata cttcatctat gaggggggtt tactatcctg 2280
atgaaatttt tagatcagac actctttatt taactcagga tttatttctt ccatttttatt 2340
ctaatgttac agggtttcat actattaatc atacgtttgg caaccctgtc ataccttttta 2400
aggatggtat ttatttgct gccacagaga aatcaaatgt tgtccgtggt tgggtttttg 2460
gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct actaatgttg 2520
ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctgtt tctaaaccca 2580
tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact ttcgagtaca 2640
tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaaa cacttacgag 2700
agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggctat caacctatag 2760
atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt aagttgcctc 2820
ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct gctcaagaca 2880
```

Figure 22-B

tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact acatttatgc 2940
tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa aatccacttg 3000
ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaatttac cagacctcta 3060
atttcagggt tgttccctca ggagatgttg tgagattccc taatattaca aacttgtgtc 3120
cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg gagagaaaaa 3180
aaatttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt ttttcaacct 3240
ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc aatgtctatg 3300
cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgccagga caaactggtg 3360
ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc cttgcttgga 3420
atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat aggtatctta 3480
gacatggcaa gcttaggccc tttgagagag acatatctaa tgtgcctttc tcccctgatg 3540
gcaaaccttg caccccacct gctcttaatt gttattggcc attaaatgat tatggttttt 3600
acaccactac tggcattggc taccaacctt acagagttgt agtactttct tttgaacttt 3660
taaatgcacc ggccacggtt tgtggaccaa aattatccac tgaccttatt aagaaccagt 3720
gtgtcaattt taattttaat ggactcactg gtactggtgt gttaactcct tcttcaaaga 3780
gatttcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat tccgttcgag 3840
atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt gtaagtgtaa 3900
ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat gttaactgca 3960
ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc atatattcta 4020
ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag catgtcgaca 4080
cttcttatga gtgcgacatt cctattggag ctggcatttg tgctattacc atacagttag 4140
actagaaaag ggcgaattct gcagatatcc atcacactgg cggccgcatc gatgagctct 4200
aattaattaa ggcttttctg ctaacatata gcaatttttg aagaaaggaa agctagaaat 4260
acccaatatc ttgctgaagc aagatattgg gtatttcttt tttttttatt ttgaatcttt 4320
ctattctgaa ttcagttaac gacgagattt agtatccttt cttgcacttt cataactcgt 4380
gaaatgccga gttggtacga attcccccaa tttgcgacct accataggat ttgttatgta 4440
aataggtata tgttcctttc cattatgaat cgcgattgta tggccaacca ttgcgggtag 4500
aatgctagat gcccgggacc acgttactat tgtttctttc tcctccttca tattgagctt 4560
ttctattttt gccaataaat gatggtacct ccccccgcca cgatcgaacg ggaatggata 4620
agaggcttgt gggattgacg tgatagggta gggttggcta tactgctggt ggcgaactcc 4680
aggctaataa tctgaagcgc ttggaagttg tagggaggat ccatggcaga caacggtact 4740
attaccgttg aggagcttaa acaactcctg gaacaatgga acctagtaat aggtttccta 4800
ttcctagcct ggattatgtt actacaattt gcctattcta atcggaacag gtttttgtac 4860
ataataaagc ttgttttcct ctggctcttg tggccagtaa cacttgcttg ttttgtgctt 4920
gctgttgtct acagaattaa ttgggtgact ggcgggattg cgattgcaat ggcttgtatt 4980
gtaggcttga tgtggcttag ctacttcgtt gcttccttca ggctgtttgc tcgtacccgc 5040
tcaatgtggt cattcaaccc agaaacaaac attcttctca atgtgcctct ccgggggaca 5100
attgtgacca gaccgctcat ggaaagtgaa cttgtcattg gtgctgtgat cattcgtggt 5160
cacttgcgaa tggccggaca ctccctaggg cgctgtgaca ttaaggacct gccaaaagag 5220
atcactgtgg ctacatcacg aacgctttct tattacaaat taggagcgtc gcagcgtgta 5280
ggcactgatt caggttttgc tgcatacaac cgctaccgta ttggaaacta taaattaaat 5340
acagaccacg ccggtagcaa cgacaatatt gctttgctag tacagtaaga attcaagccg 5400
aattccagca cactggcggc cgttactagt ggatccgagc tctaattaat taaggctttt 5460
ctgctaacat atagcaattt ttgaagaaag gaaagctaga aatacccaat atcttgctga 5520
agcaagatat tgggtatttc tttttttttt attttgaatc tttctattct gaattcagtt 5580
aacgacgaga tttagtatcc tttcttgcac tttcataact cgtgaaatgc cgagttggta 5640
cgaattcccc caatttgcga cctaccatag gatttgttat gtaaataggt atatgttcct 5700
ttccattatg aatcgcgatt gtatggccaa ccattgcggg tagaatgcta gatgcccggg 5760
accacgttac tattgtttct ttctcctcct tcatattgag cttttctatt tttgccaata 5820
aatgatggta cctccccccg ccacgatcga acgggaatgg ataagaggct tgtgggattg 5880
acgtgatagg gtagggttgg ctatactgct ggtggcgaac tccaggctaa taatctgaag 5940
cgcttggaag ttgtagggag gatcctatgg ctcgtgaagc ggttatcgcc gaagtatcga 6000
ctcaactatc agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg ttgctggccg 6060

Figure 22-C

```
tacatttgta cggctccgca gtggatggcg gcctgaagcc acacagtgat attgatttgc 6120
tggttacggt gaccgtaagg cttgatgaaa caacgcggcg agctttgatc aacgaccttt 6180
tggaaacttc ggcttcccct ggagagagcg agattctccg cgctgtagaa gtcaccattg 6240
ttgtgcacga cgacatcatt ccgtggcgtt atccagctaa gcgcgaactg caatttggag 6300
aatggcagcg caatgacatt cttgcaggta tcttcgagcc agccacgatc gacattgatc 6360
tggctatctt gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg 6420
aggaactctt tgatccggtt cctgaacagg atctatttga ggcgctaaat gaaaccttaa 6480
cgctatggaa ctcgccgccc gactgggctg gcgatgagcg aaatgtagtg cttacgttgt 6540
cccgcatttg gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact 6600
gggcaatgga gcgcctgccg gcccagtatc agcccgtcat acttgaagct agacaggctt 6660
atcttggaca agaagaagat cgcttggcct cgcgcgcaga tcagttggaa gaatttgtcc 6720
actacgtgaa aggcgagatc actaaggtag ttggcaaata actcgagact aagtggataa 6780
aattagatag aaaaaaggtc taaataaaaa agaagagaaa tagaaagatc aaaaatcagt 6840
tacgaaaatg cagtaattct tctttttct tctaattgat tgcaattaaa ctcgtctcaa 6900
tctgaaaaaa gattgagccg agtttaaata gattttgata cgatcatgag acttgacaaa 6960
tcgggattcc tctattctat atatttagaa gatataaagg tataatacaa taaataaata 7020
caatctagaa gctcgatgat ccttaggatt ggtatattct tttctatcct gtagtttgta 7080
gtttccctga atcaagccaa gtatcacacc tctttctacc catcctgtat attgtcccct 7140
ttgttccgtg ttgaaataga accttaattt attacttatt tttttattaa attttagatt 7200
tgttagtgat tagatattag tattagacga gattttacga aacaattatt tttttatttc 7260
tttataggag aggacaaatc tcttttttcg atgcgaattt gacacgacat aggagaagcc 7320
gccctttatt aaaaattata ttattttaaa taatataaag gggggttccaa catattaata 7380
tatagtgaag tgttcccca gattcagaac tttttttcaa tactcacaat ccttattagt 7440
taataatcct agtgattgga tttctatgct tagtctgata ggaaataaga tattcaaata 7500
aataatttta tagcgaatga ctattcatct attgtatttt catgcaaata gggggcaaga 7560
aaactctatg gaaagatggt ggtttaattc gatgttgttt aagaaggagt tcgaacgcag 7620
gtgtgggcta aataaatcaa tgggcagtct tggtcctatt gaaaatacca atgaagatcc 7680
aaatcgaaaa gtgaaaaaca ttcatagttg gaggaatcgt gacaattcta gttgcagtaa 7740
tgttgattat ttattcggcg ttaaagacat tcggaatttc atctctgatg acactttttt 7800
agttagtgat aggaatggag acagttattc catctatttt gatattgaaa atcatatttt 7860
tgagattgac aacgatcatt cttttctgag tgaactagaa agttcttttt atagttatcg 7920
aaactcgaat tatcggaata atggatttag ggcgaagat ccctactata attcttacat 7980
gtatgatact caatatagtt ggaataatca cattaatagt tgcattgata gttatcttca 8040
gtctcaaatc tgtatagata cttccattat aagtggtagt gagaattacg gtgacagtta 8100
catttatagg gccgtttgtg gtggtgaaag tcgaaatagt agtgaaaacg agggttccag 8160
tagacgaact cgcacgaagg gcagtgattt aactataaga gaaagttcta atgatctcga 8220
ggtaactcaa aaatacaggc atttgtgggt tcaatgcgaa aattgttatg gattaaatta 8280
taagaaattt ttgaaatcaa aaatgaatat ttgtgaacaa tgtggatatc atttgaaaat 8340
gagtagttca gatagaattc gtaatcatgg tcatagctgt ttcctgtgtg aaattgttat 8400
ccgctcacaa ttccacacaa catacgagcc ggaagcataa agtgtaaagc ctggggtgcc 8460
taatgagtga gctaactcac attaattgcg ttgcgctcac tgcccgcttt ccagtcggga 8520
aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 8580
attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 8640
cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 8700
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 8760
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 8820
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 8880
tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 8940
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 9000
gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc 9060
ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 9120
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 9180
aagtggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg 9240
```

Figure 22-D

aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 9300
ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 9360
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 9420
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 9480
tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc 9540
ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcctga 9600
ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca 9660
atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc 9720
ggaagggccg agcgcagaag tggtcctgca actttatccg cctccatcca gtctattaat 9780
tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc 9840
attgctacag gcatcgtggt gtcacgctcg tcgtttggta tggcttcatt cagctccggt 9900
tcccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc 9960
ttcggtcctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg 10020
gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgcttttc tgtgactggt 10080
gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg 10140
gcgtcaatac gggataatac cgcgccacat agcagaactt taaaagtgct catcattgga 10200
aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg 10260
taacccactc gtgcacccaa ctgatcttca gcatctttta ctttcaccag cgtttctggg 10320
tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa taagggcgac acggaaatgt 10380
tgaatactca tactcttcct ttttcaatat tattgaagca tttatcaggg ttattgtctc 10440
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca 10500
tttccccgaa aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat 10560
aaaaataggc gtatcacgag gccctttcgt c 10591

Figure 23

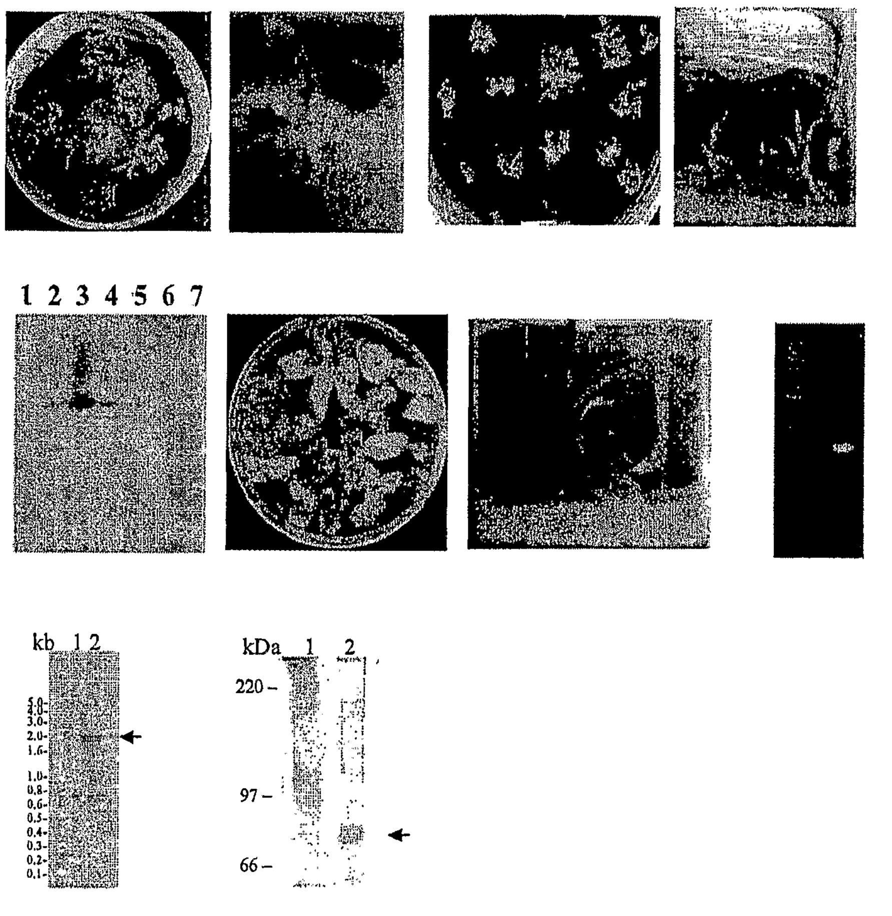
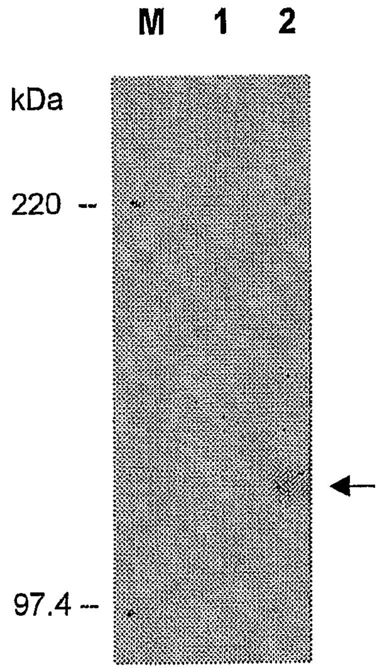


Figure 24



INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2004/001419

A. CLASSIFICATION OF SUBJECT MATTER

C12N15/50,15/05,15/63,A61K31/7088,38/16,39/215,A61P31/14,11/00,A61K 48/00
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N15, A61K, A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPOQUE ,CPRS ,CNKI, BA
SARS ,COV, SEVERE ACUTE RESPIRATORY SYNDROME ,CORONAVIRUS ,SPIKE GLYCOPROTEIN ,MEMBRANE GLYPROTEIN ,PLANT ,TRANSFORMATION ,MODIFY

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN, A ,1450173, 22 Oct 2003,see the whole document	1-2,4,10,14-16,28-30
A	DATABASE NCBI ONLINE ,ACCESSION NUMBER :AY 274119	1-2,4,10,14-16,28-30
A	CN, A ,1449826 ,22 Oct 2003,see the whole document	3,5-9,11-14,28-30
PA	CN,A,1548452, 24 ,Nov 2004,see the whole document	3,5-9,11-14,28-30

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
01,M ar 2005(01,03 2005)	31 · MAR 2005 (3 1 · 0 3 · 2 0 0 5)

Name and mailing address of the ISA/	Authorized officer
6,Xitucheng Road, Jimen Bridge, Haidian District, Beijing,100088,P.R.China	Wang Huimei
Facsimile No. (86-10)62019451	Telephone No. (86-10)62085294

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2004/001419

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:17-25,26-27and 31-39
   because they relate to subject matter not required to be searched by this Authority, namely:
   Claims 17-25 and 31-39 relate to the methods for the treatment and for the diagnosis of diseases, respectively. And claims 26-27 relate to the plant varieties. So the claims above are found unsearchable according to PCT rule 39.

2. ☐ Claims Nos.:
   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant,this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant.Consequently ,this international search report is restricted to the invention first mentioned in the claims;it is covered by claims Nos.:

Remark on protest

☐ The acditional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2004/001419

Patent document Cited in search report	Publication data	Patent family members	Publication data
CN 1450173 A	22-10-2003	NONE	--------
CN 1449826 A	22-10-2003	NONE	--------
CN 1548452 A	24-11-2004	NONE	----------