Spectrum of mutations in the MECP2 gene in patients with infantile autism and Rett syndrome

CHING-WAN LAM, WAI-LAN YEUNG, CHUNG-HUNG KO, PRISCILLA M K POON, SUI-FAN TONG, KWOK-YIN CHAN, IVAN F M LO, LISA Y S CHAN, JOANNIE HUI, VIRGINIA WONG, CHI-PUI PANG, Y M DENNIS LO and TAI-FAI FOK

J. Med. Genet. 2000;37;41-
doi:10.1136/jmg.37.12.e41

Electronic letter

J Med Genet 2000;**37** (http://jmedgenet.com/cgi/content/full/37/12/*e*41)

Spectrum of mutations in the *MECP2* gene in patients with infantile autism and Rett syndrome

EDITOR—Rett syndrome (RTT, MIM 312750) is a progressive neurological disorder, occurring almost exclusively in females during their first two years of life . RTT is one of the most common causes of mental retardation in females, with an incidence of 1 in 10 000-15 000 female births. Patients with classical RTT appear to develop normally until 6-18 months of age, then gradually lose speech and purposeful hand use, and, eventually, develop microcephaly, seizures, autism, ataxia, hyperventilation, and stereotypic hand movements. After the initial regression, the clinical condition stabilises and patients usually survive into adulthood. Laboratory investigations have not shown any metabolic abnormalities in affected subjects.[1] RTT is included in the differential diagnosis of autistic disorder in girls.[1–3] Qualitative abnormalities in social and communicative development and stereotypic behaviour are typically present in RTT. Definitive diagnosis is often delayed until after the loss of purposeful hand movements and the relatively characteristic hyperventilation later in childhood,[1] and earlier diagnosis would be desirable.

The occurrence of a few familial cases with maternal inheritance suggests that RTT is an X linked dominant mutation with lethality in hemizygous males. Previous exclusion mapping studies using RTT families identified a locus in Xq28.[4–6] Using a systematic gene screening approach, Amir *et al*[7] and Wan *et al*[8] have identified mutations in the *MECP2* gene, which encodes methyl-CpG binding protein 2, as the cause of some RTT cases. Most of the mutations are de novo and occur at a CpG dinucleotide. The cytosine in the CpG dinucleotide is a frequent site for DNA methylation and deamination of methylated cytosine to thymine causes the transition.

The MeCP2 protein silences methylated chromatin by recruiting a histone deacetylase complex.[9] Unlike most other transcriptional repressor proteins, however, the binding site of MeCP2 occurs frequently in genomic DNA as it requires only a single methylated CpG base pair to bind. MeCP2 contains two functional domains, an 85 amino acid methyl-CpG binding domain (MBD) (residues 78-162), essential for its binding to 5-methylcytosine, and a 104 amino acid transcriptional repressor domain (TRD) (residues 207-310) that interacts with histone deacetylase and the transcription corepressor Sin3. It has been shown that interactions between this transcription repressor complex and chromatin bound MeCP2 leads to deacetylation of core histones, which in turn leads to changes in chromatin architecture and transcriptional repression.

To date, the mutational spectrum of *MECP2* in Chinese is not known. To define the role of *MECP2* mutations causing RTT in our population, we have screened *MECP2* mutations by denaturing high performance liquid chromatography (DHPLC). As RTT is included in the differential diagnosis of autistic disorders in girls, we have also screened *MECP2* mutations in a group of female patients with autism and mental retardation.

We screened genomic DNA from 13 sporadic RTT patients and 21 patients with autism and mental retardation by DHPLC and by direct DNA sequencing. All the subjects were unrelated females and were ethnic Chinese, with no family history of the disease. The clinical findings met the criteria of inclusion and exclusion for the diagnosis of RTT.[10] Patients with autism and mental retardation were obtained from a previous study.[11] The diagnosis of autism was based on clinical features and evaluated by diagnostic criteria from DSM-IV.[12] Most of them had onset of autistic features at less than 3 years of age. Informed consent was obtained from the patients or the parents.

Genomic DNA was extracted from peripheral blood samples using a QIAamp Blood Kit (Qiagen) according to the manufacturer's instructions. PCR amplification was conducted using primer pairs and conditions described elsewhere.[7 8] PCR products were purified by MicroSpin columns S-300 (Pharmacia Biotech) according to the manufacturer's instructions. PCR products shorter than those expected from the wild type sequence (in patients PWH24 and PMH65) were extracted from agarose gels using QIAquick Gel Extraction kit (Qiagen) according to the manufacturer's instructions. Direct sequencing of the PCR products was performed using the ABI PRISM dRhodamine Terminator Cycle Sequencing Ready Reaction Mix (PE Biosystems). Sequencing fragments were separated by capillary electrophoresis and detected via laser induced fluorescence on an ABI PRISM 310 Genetic Analyzer (PE Biosystems). Both strands were sequenced to confirm all the mutations detected. Sequencing results were compared with the reference human *MECP2* sequence (GenBank X99686).

Heteroduplex analysis was performed on a WAVE™ DHPLC instrument (Transgenomic). Analysis was performed at a temperature sufficient to partially denature (melt) the DNA heteroduplexes. The melted heteroduplexes are resolved from the corresponding homoduplexes by ion pair reversed phase liquid chromatography. The procedure is referred to as temperature modulated heteroduplex chromatography (TMHC). THMC relies upon the physical changes in DNA molecules induced by mismatched heteroduplex formation during reannealing of wild type and mutant DNA. Between 5 and 10 μl of crude PCR product was loaded on a DNASep column (Transgenomic) and was eluted from the column by an acetonitrile gradient in a 0.1 mol/l triethylammonium acetate (TEAA) buffer, pH 7.0, at a constant flow rate of 0.9 ml/minute. The standard buffers are prepared from concentrated TEAA to give A=0.1 mol/l TEAA, B=0.1 mol/l TEAA, and 25% acetonitrile. The gradient was created by mixing eluents A and B. The recommended gradient for mutation detection is a slope of 2% increase in buffer B per minute. Eluted DNA fragments were detected with ultraviolet absorption at wave length 260 nm. The WAVE utility software helps to determine the correct temperature for mutation scanning based on the sequence of the wild type DNA.

We used a methylation specific PCR assay developed at the human androgen receptor locus (*HUMARA*) on the X chromosome for X inactivation studies. The X inactivation pattern is defined as the ratio of the corrected peak area of a smaller allele to the corrected peak area of a larger allele.[13]

Among the 13 RTT patients, we identified one missense mutation, two nonsense mutations, one microdeletion, and

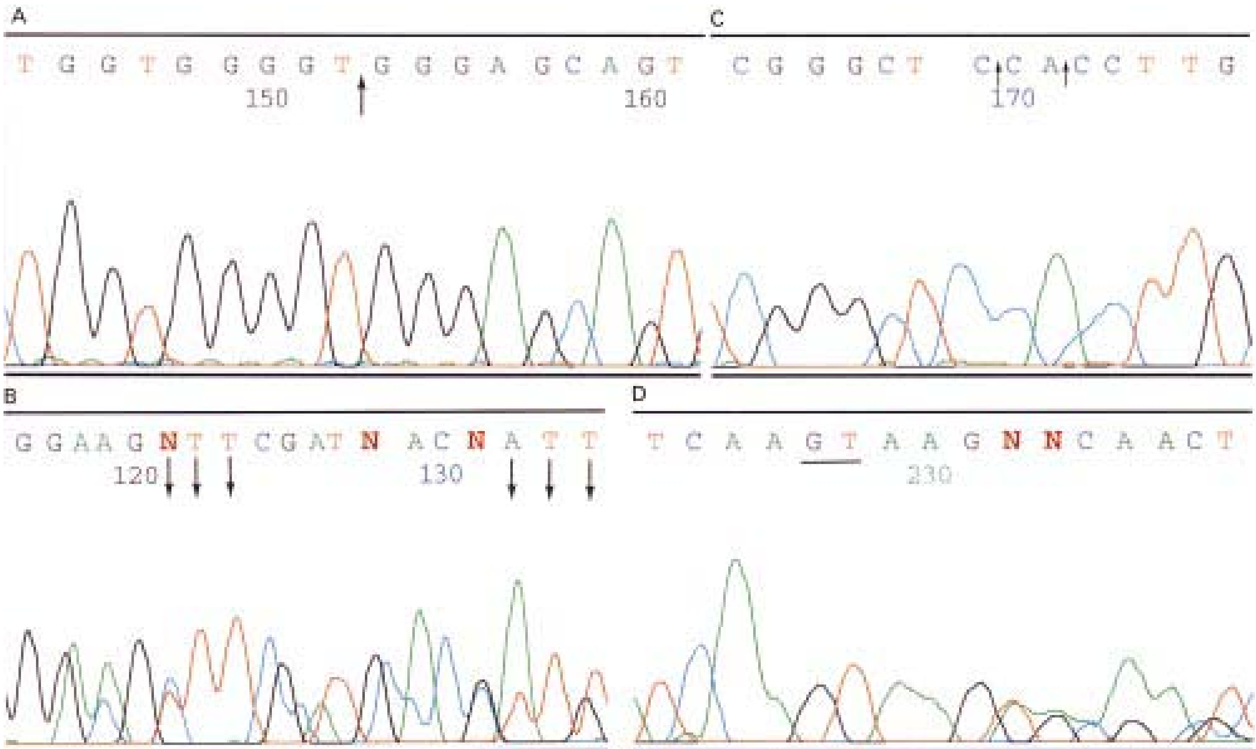

Figure 1 Analysis of the MECP2 gene. DNA sequence analysis of the four novel mutations found in patients PMH65 (A), PWH44 (B), PWH24 (C), and PWHA34 (D). Arrows in (A) and (C) indicate the deletion breakpoints. The 5' splice site of intron 2 is underlined in (D). The arrows in (B) indicate the positions of the sequence, 5'-TTT-3', in the sequences of the wild type and mutant alleles. The DNA sequences of (A) and (D) are shown in the sense direction. The DNA sequences of (B) and (C) are shown in the antisense direction.

two insertion/deletions (indels). Three of the mutations were novel (fig 1) and three of the mutations have previously been reported.[7 8 14] All are de novo mutations. None of these mutations were detected in 200 normal X chromosomes. Four of the six patients with *MECP2* mutations were heterozygous for the androgen receptor gene polymorphism and the XCI results are shown in table 1.

One missense mutation was detected. The mutation, 390C→T, is located in exon 2 and changes codon 106 from CGG to TGG. This mutation occurs at a CpG dinucleotide and changes the coded amino acid residue from arginine to tryptophan, that is, R106W, in patient CG1295. 390C→T, leading to the R106W substitution in the MBD, was previously found in affected half sisters but not in their common mother[7] and in an unrelated sporadic case.[8] The substituted arginine residue is conserved in MeCP2 from mammals to *Xenopus laevis*.[7]

The two nonsense mutations, which also occur at a CpG dinucleotide, that is, 576C→T (R168X) and 954 C→T (R294X) in patients CMC52 and PWH23, respectively, are located in exon 3. The 576C→T mutation changes codon 168 from CGA to TGA, changing the arginine codon to a stop codon. This mutation was found previously in five unrelated white subjects, one Japanese, and a Brazilian family with three affected sisters.[7 8] Our results confirmed that R168X is a frequent mutation causing RTT. Codon 168 is located between the MBD and the TRD. The putative truncated protein of 167 amino acids, lacking the nuclear localisation signal (NLS) within the TRD, is predicted to be located predominantly in the cytoplasm. The nonsense mutation in patient PWH23 is a C→T transition at nucleotide position 954, which converts a CGA to a TGA (R294X) that predicts truncation of the MeCP2 protein at residue 294 of 486. This mutation is located in the TRD. The truncated protein may cause abnormal folding or affect interactions with other proteins of the Sin3A/histone deacetylase silencing complex.

The indel in patient PWH44, 824delCins11, is located in the TRD. The mutation involves a deletion of the last nucleotide at codon 250 followed by insertion of an 11 bp sequence, that is, 5'-TCAGGAAGCTT-3' and causes a shift of the reading frame. This shift creates a stop codon TGA at codon 261, that is, P261X. A truncated protein of 260 amino acids results and the TRD domain is disrupted. The indel in patient PWH24 is 1118del131insTG. This indel starts at codon 348, changing the codon from GAG to GTG. This changes the amino acid at position 348 from glutamic acid to valine, that is, E348V, but does not change the reading frame. Forty three codons, from codons 349 to 391, are deleted, that is, S349-P391del43,[15] leaving the C-terminal of the protein from amino acid residues 392 to 486 intact. However, this deletion eliminates both the poly-His and poly-Pro domains and a truncated protein of 443 amino acids results. The microdeletion, 1231del41, in patient PMH65, involves a deletion of 41 bases starting from the second nucleotide of codon 386 to the last nucleotide of codon 399. This mutation causes the deletion of the poly-Pro domain (codons 384 to 393). The deleted region is flanked by a direct repeat of four cytosine bases. This small deletion may be caused by replication slippage errors resulting from looping out of the template strand during DNA replication. This mutation shifts the reading frame at codon 386 and creates a stop codon TGA at codon 389. A truncated protein of 388 amino acids, P389X, results.

We identified three novel *MECP2* polymorphisms in the sequence analysis of the RTT and autistic patients. The single nucleotide polymorphism (SNP) IVS2+22C→G in patients CMC51 and PWH55 was located in intron 2 (data not shown). We found this SNP in normal males, indicating that IVS2+22C→G is a neutral polymorphism. Another SNP, 676C→T, changes codon 201 from GCG to GTG in patient QMH58, changing the amino acid from alanine to valine, that is, A201V (data not shown). The SNP is de novo as it is not found in either parent. The

Table 1 Mutations of the MECP2 gene in patients with RTT and infantile autism

Patient	*Diagnosis*	*Mutation type*	*Exon/IVS*	*Domain affected*	*Nucleotide change*	*Amino acid change*	*CpG hotspot*	*X chromosome inactivation*	*Reference*
CG1295	RTT	Missense	2	MBD	390C→T	R106W	+	66:34	7, 8
CMC52	RTT	Nonsense	3	TRD	576C→T	R168X	+	79:21	7, 8
PWH23	RTT	Nonsense	3	TRD	954C→T	R294X	+	Not informative	14
PMH65	RTT	Deletion	3	Poly-Pro domain	1231del41	P389X	–	Not informative	This report
PWH44	RTT	Indel	3	TRD	824delCins11	P261X	–	73:27	This report
PWH24	RTT	Indel	3	Poly-Pro & His domains	1118del131insTG	[E348V, S349-P391del43]	–	50:50	This report
PWHA34	Infantile autism	Splicing	IVS2	Donor splice junction	IVS+2delTAAG	Unknown	–	Not informative	This report

Table 2 SpliceView analysis of the donor splice junction of intron 2 of the MECP2 gene

Donor splice junction	*Exon sequence*	*Intron sequence*	*Score*
IVS2+1	CAA	GTAAGT	84
IVS2+5	TAA	GTAAGA	81
IVS2+77	AAT	GTATGT	83
IVS+131	CAG	GTGTGC	84
IVS+1 (mutant)	CAA	GTAAGA	82

amino acid is not conserved and the SNP has been found in normal males, indicating that 676C→T is a neutral polymorphism. In patient PWHA5, we found a SNP changing the ninth base of the 3'-UTR from G to A (data not shown). However, this nucleotide is not conserved and the nucleotide at the analogous position in mouse is adenine. This SNP represents a rare polymorphism.

In one of the patients with infantile autism, PWHA34, we found a mutation, IVS2+2delTAAG, in the 5' splicing site of intron 2, causing a deletion of four bases TAAG from the second base of the intron. This mutation was not found in her parents or 200 normal X chromosomes. The mutation was probably caused by mispairing of a direct repeat of 5'-taag-3' in the sequence 5'-gtaagTaaggagcaactcctatct-3'. The mutation retains a GT dinucleotide, that is, 5'-gTaaggagcaactcctatct-3', but the sequence of the splicing site will change from IVS+6 position onwards. Using SpliceView (http://www.itba.mi.cnr.it/webgene/), the mutant splice site has a lower consensus value (score 82) than the wild type splice site (score 84). Two downstream splice sites which have higher consensus values (IVS+77 with a score of 83 and IVS+131 with a score of 84) may act as the new 5' splice sites (table 2). This is predicted to cause aberrant splicing with partial intron 2 retention and premature termination. Unfortunately, mRNA was not available to evaluate the predicted result.

We found six *MECP2* mutations in 13 patients with classical RTT. Four of the six mutations (R168X, P261X, R294X, P389X) lead to premature termination of translation. The 1118del131insTG mutation leads to a truncated protein of 443 amino acids [E348V;S349-P391del43]. Three patients with *MECP2* mutations have moderately skewed XCI patterns. This is consistent with the fact that RTT patients as a group have a higher frequency of moderate skewing (65-80%) of XCI in lymphocytes, when compared with normal controls.[16] Like previous studies, we found that several (three out of six) of the mutations causing RTT are C→T transitions occurring at CpG dinucleotides. These mutations are probably the result of methylation deamination of the CpG dinucleotide. In addition, we found several direct repeats from codon 350 to 411. Within these 186 nucleotides, there are five simple direct repeats of four cytosine bases, two simple direct repeats of five cytosine bases, and two simple direct repeats of six cytosine bases. In addition, there are two direct repeats of three AGC and two direct repeats of three CAC. Together, there are 78 nucleotides located in a direct repeat sequence, accounting for about 42% of the sequence (fig 2). This part of the gene might be more vulnerable to rearrangement mutations.

The MeCP2 protein has one poly-Ala domain (residues 277 to 283: (5×Ala)-Glu-Ala), one poly-His domain (residues 366 to 372: 7×His), and one poly-Pro domain (residues 384 to 393: Pro-Pro-Leu-(5×Pro)-Glu-Pro). Although the functions of these three domains in the protein are unclear, they are all evolutionarily conserved from mammals to *Xenopus laevis*. We found two mutations that disrupt one or two of these domains. Interestingly, the mutation in patient PWH24 disrupts both the poly-His and poly-Pro domains without altering the reading frame

AGC AAG GAG AGC AG**C CCC** AAG GGG CGC **AGC AGC AGC** GCC TCC TCA
346 Ser Lys Glu Ser Ser Pro Lys Gly Arg Ser Ser Ser Ala Ser Ser
360

CCC CCC AAG AAG GAG **CAC CAC CAC** CAT **CAC CAC CAC** TCA GAG T**CC**
361 Pro Pro Lys Lys Glu His His His His His His His Ser Glu Ser
375

CCA AAG G**CC CCC** GTG CCA CTG CTC CCA **CCC C**TG **CCC CC**A CCT CCA
376 Pro Lys Ala Pro Val Pro Leu Leu Pro Pro Leu Pro Pro Pro Pro
390

CCT GAG CCC GAG AGC TCC GAG GA**C CCC** ACC AG**C CCC CC**T GAG **CCC**
391 Pro Glu Pro Glu Ser Ser Glu Asp Pro Thr Ser Pro Pro Glu Pro
405

CAG GAC TTG **AGC AGC AGC** GTC TGC AAA GAG GAG AAG ATG CCC AGA
406 Gln Asp Leu Ser Ser Ser Val Cys Lys Glu Glu Lys Met Pro Arg
420

Figure 2 Simple direct repeats in exon 3 of the MECP2 gene. The direct repeat sequences are shown in bold. The poly-His and poly-Pro domains are underlined.

and the rest of the C-terminal. Together, these results suggest that these domains are important for the normal function of the protein and that disruption of these domains might alter the conformation of the protein.

We identified a mutation in one of the 21 patients with infantile autism. The mutation involved the 5' splice site of intron 2. The affected patient, PWHA34, presented to us at 4 years of age with a mental age close to 2 years and significant difficulties in social interaction and communication. Her spoken language has not developed, but she did not show a regression phase in her clinical course. There is no evidence of seizures, kyphoscoliosis, stereotypic hand movements, or microcephaly. Unfortunately, we are unable to re-evaluate the clinical diagnosis because the patient has already been lost to follow up. Further investigations will be required to determine whether this mutation interferes with gene product function.

To date, only three *MECP2* mutations have been identified in 17 RTT families.[7 8 14] Thus, 14 RTT families do not have mutations in either the coding region or the intron/exon boundaries of *MECP2* to account for the disorder. The presence of abnormalities in the untranslated regions of the *MECP2* mRNA and genetic regulatory elements have yet to be explored, but it is also possible that another tightly linked locus may be present on chromosome Xq28. Until now, the diagnosis of RTT has relied solely on clinical observations. The discovery of *MECP2* as an RTT associated gene will enable the development of a test for earlier diagnosis using DNA based methods. Mutational analysis at the DNA level will increasingly contribute to diagnosis of RTT, particularly in atypical cases.

CHING-WAN LAM*
WAI-LAN YEUNG†
CHUNG-HUNG KO‡
PRISCILLA M K POON*
SUI-FAN TONG*
KWOK-YIN CHAN§
IVAN F M LO¶
LISA Y S CHAN*
JOANNIE HUI†
VIRGINIA WONG**
CHI-PUI PANG††
Y M DENNIS LO*
TAI-FAI FOK†

Department of Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China
†*Department of Paediatrics, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China*
‡*Department of Paediatrics, Caritas Medical Center, Hong Kong, China*
§*Department of Paediatrics, Princess Margaret Hospital, Hong Kong, China*
¶*Clinical Genetic Service, Department of Health, Hong Kong, China*
***Department of Paediatrics, Queen Mary Hospital, Hong Kong, China*
††*Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong Eye Hospital, Hong Kong, China*

Correspondence to: Dr Lam, ching-wanlam@cuhk.edu.hk

1 Hagberg B. Rett syndrome: clinical peculiarities and biological mysteries. *Acta Paediatr Scand* 1995;**84**:971-6.
2 Gillberg C. The borderland of autism and Rett syndrome: five case histories to highlight diagnostic difficulties. *J Autism Dev Disord* 1989;**19**:545-59.
3 Olsson B, Rett A. Autism and Rett syndrome: behavioral investigations and differential diagnosis. *Dev Med Child Neurol* 1987;**29**:429-41.
4 Webb T, Clarke A, Hanefeld F, Pereira JL, Rosenbloom L, Woods CG. Linkage analysis in Rett syndrome families suggests that there may be a critical region at Xq28. *J Med Genet* 1998;**35**:997-1003.
5 Xiang F, Zhang Z, Clarke A, Joseluiz P, Sakkubai N, Sarojini B, Delozier-Blanchet CD, Hansmann I, Edstrom L, Anvret M. Chromosome mapping of Rett syndrome: a likely candidate region on the telomere of Xq. *J Med Genet* 1998;**35**:297-300.
6 Sirianni N, Naidu S, Pereira J, Pillotto RF, Hoffman EP. Rett syndrome: confirmation of X-linked dominant inheritance, and localization of the gene to Xq28. *Am J Hum Genet* 1998;**63**:1552-8.
7 Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999;**23**:185-8.
8 Wan M, Lee SS, Zhang X, Houwink-Manville I, Song HR, Amir RE, Budden S, Naidu S, Pereira JL, Lo IF, Zoghbi HY, Schanen NC, Francke U. Rett syndrome and beyond: recurrent spontaneous and familial MECP2 mutations at CpG hotspots. *Am J Hum Genet* 1999;**65**:1520-9.
9 Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 1998;**393**:386-9.
10 The Rett Syndrome Diagnostic Criteria Work Group. Diagnostic criteria for Rett syndrome. *Ann Neurol* 1988;**23**:425-8.
11 Poon PM, Chen QL, Lai KY, Wong CK, Pang CP. CGG repeat interruptions in the FMR1 gene in patients with infantile autism. *Clin Chem Lab Med* 1998;**36**:649-53.
12 American Psychiatric Association. *Diagnostic and statistical manual of mental retardation disorders*. 4th ed. Washington, DC: APA, 1994.
13 Kubota T, Nonoyama S, Tonoki H, Masuno M, Imaizumi K, Kojima M, Wakui K, Shimadzu M, Fukushima Y. A new assay for the analysis of X-chromosome inactivation based on methylation-specific PCR. *Hum Genet* 1999;**104**:49-55.
14 Xiang F, Buervenich S, Nicolao P, Bailey ME, Zhang Z, Anvret M. Mutation screening in Rett syndrome patients. *J Med Genet* 2000;**37**:250-5.
15 Dunnen JT, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat* 2000;**15**:7-12.
16 Krepischi AC, Kok F, Otto PG. X chromosome-inactivation patterns in patients with Rett syndrome. *Hum Genet* 1998;**102**:319-21.