<table>
<thead>
<tr>
<th>Title</th>
<th>The viruses of wild pigeon droppings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Phan, TG; Vo, NP; Boros, Á; Pankovics, P; Reuter, G; Li, OTW; Wang, C; Deng, X; Poon, LLM; Delwart, E</td>
</tr>
<tr>
<td>Citation</td>
<td>PLoS One, 2013, v. 8 n. 9, p. e72787</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2013</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/195732">http://hdl.handle.net/10722/195732</a></td>
</tr>
<tr>
<td>Rights</td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
The Viruses of Wild Pigeon Droppings

Tung Gia Phan1,2, Nguyen Phung Vo1,3, Ákos Boros4, Péter Pankovics4, Gábor Reuter4, Olive T. W. Li5, Chunling Wang5, Xutao Deng1, Leo L. M. Poon6, Eric Delwart1,2*

1 Blood Systems Research Institute, San Francisco, California, United States of America, 2 Department of Laboratory Medicine, University of California San Francisco, San Francisco, California, United States of America, 3 Pharmacology Department, School of Pharmacy, Ho Chi Minh City University of Medicine and Pharmacy, Ho Chi Minh, Vietnam, 4 Regional Laboratory of Virology, National Reference Laboratory of Gastroenteric Viruses, ANTISZ Regional Institute of State Public Health Service, Pécs, Hungary, 5 Stanford Genome Technology Center, Stanford, California, United States of America, 6 Centre of Influenza Research and School of Public Health, University of Hong Kong, Hong Kong SAR

Abstract

Birds are frequent sources of emerging human infectious diseases. Viral particles were enriched from the feces of 51 wild urban pigeons (Columba livia) from Hong Kong and Hungary, their nucleic acids randomly amplified and then sequenced. We identified sequences from known and novel species from the viral families Circoviridae, Paroviridae, Picornaviridae, Reoviridae, Adenovirus, Astroviridae, and Caliciviridae (listed in decreasing number of reads), as well as plant and insect viruses likely originating from consumed food. The near full genome of a new species of a proposed parovirus genus provisionally called Aviparvovirus contained an unusually long middle ORF showing weak similarity to an ORF of unknown function from a fowl adenovirus. Picornaviruses found in both Asia and Europe that are distantly related to the turkey megrivirus and contained a highly divergent 2A1 region were named mesiviruses. All eleven segments of a novel rotavirus subgroup related to a chicken rotavirus in group G were sequenced and phylogenetically analyzed. This study provides an initial assessment of the enteric virome in the droppings of pigeons, a feral urban species with frequent human contact.

Introduction

Many infectious diseases in humans are caused by pathogens originating from a wide variety of animals. More than 60% of emerging diseases are estimated to originate from wildlife [1–4]. Public awareness of zoonoses has recently increased because of their public health and economic impacts. Birds are recognized as frequent reservoirs for viruses that are of concern to humans; notably influenza A which is capable of infecting other mammals thereby facilitating genome segment reassortments and changes in tropism and transmission efficiency [5–8]. Sporadic human infections of the virulent H5N1 resulting from direct contact with infected poultry or wild birds have been reported in 15 countries, mainly in Asia [9–13], and H7N9 has recently emerged as a virus of concern. The prevalence of avian influenza viruses was 12% of oropharyngeal and 20% of cloacal swab specimens collected from urban pigeons in Slovakia [14]. H5N1 was found in a dead feral pigeon in Hong Kong [15] but is generally apathogenic in this host species and the overall risk of H5N1 transmission from pigeons to humans or chickens appears low [16,17]. West Nile virus (WNV) and Saint Louis encephalitis (SLE) virus, two arboviruses in the Flavivirus genus transmitted by mosquitoes bites, are disseminated by wild birds [18–23]. WNV-specific antibody and viremia was found in 25.7% and 11% of rock pigeons, respectively in the United States [24]. WNV was also isolated in pools of brains, kidneys, heart and spleen of feral pigeons and mappies [25]. Pigeons developed low levels of WNV viremia; insufficient to infect mosquitoes [24,26]. Avian paramyxoviruses, including Newcastle disease virus, are common domestic and wild bird pathogens [27–30]. Paramyxovirus type-1 can be found in pigeons worldwide [31–34] but the clinical signs vary depending on the immunity of the host and virulence of the specific isolates [35]. While human infection with Newcastle disease virus is rare, at least two outbreaks of conjunctivitis due to Newcastle disease virus have been reported in poultry workers [28,36,37]. Chicken anemia virus (CAV), until recently the only member of the gyrovirus genus, is highly contagious and causes severe anemia, hemorrhage and depletion of lymphoid tissue in chickens [38–40]. Related gyroviruses were recently characterized in human skin [41–44] indicating possible human tropism. Gyrovirus DNA was also detected in three blood samples of solid organ transplant patients and in one HIV-infected person [44] as well as in 0.85% of healthy French blood donations [45].

Pigeons are therefore natural reservoirs for pathogens that have caused emerging and re-emerging diseases in humans. In order to better understand the viruses shed by pigeons to which humans are frequently exposed, we genetically characterized the viral community in droppings from wild pigeons in Hong Kong and Hungary following an unbiased amplification method and deep sequencing.


Editor: Dhanasekaran Vijaykrishna, Duke-NUS Graduate Medical School, Singapore

Received May 9, 2013; Accepted July 12, 2013; Published September 4, 2013

Copyright: © 2013 Phan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors acknowledge NHLBI R01HL083254 and support from BSRI to ED, EMPERIE grant EU FP7 223498 to L.LMP. Grants from the Hungarian Scientific Research Fund (OTKA, K83013 and PD108482) and support from the János Bolyai Research Scholarship of the Hungarian Academy of Sciences to GR. AB was supported by the National Excellence Program TÁMOP 4.2.4. A/2-11-1-2012-0001 subsidized by the European Union and co-financed by the European Social Fund. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: delwarte@medicine.ucsf.edu
Genome Acquisition of Novel Viruses

Sequences showing significant but divergent BLASTx hits to parvoviruses, picornaviruses or rotaviruses in the same sample pool were linked together using RT-PCR or PCR and primers based on the initial short sequence reads or contigs. 5’ and 3’ rapid amplification of cDNA end (5’- and 3’ RACE) was used to acquire the 5’ and 3’ extremities of viral genome. Phylogenetic analyses were performed using novel virus sequences, their best mammalian viral hits belonged to the Parvoviridae family. The subfamily Parvovirinae currently consists of five genera: Parvovirus, Rotavirus, Adenovirus, Astrovirus, and Calicivirus. The Picornaviridae-related sequences showed best hits to three different clades, Circovirus (1824 reads), Calicivirus (20 reads) and Gyrovirus (2 reads). 77% of the circovirus hits showed best matches to pigeon circovirus of which 108 sequences had >90% aa-identity to pigeon circovirus. 74% of the Parvovirinae sequences showed best hits to AAV (Adeno-associated virus). A majority of the picornavirus sequences (85%; 759/889) showed best hits to Turkey hepatitis virus [52]. The adenovirus sequences showed best hits to three difference genera Mastadenovirus (3 reads), Adenovirinae (77 reads) and Sadenovirus (296 reads).

The pigeon fecal sample from Hungary was analyzed using 6,818 unique reads from 454 pyrosequencing and showed viral hits to picornaviruses,astroviruses, and tobamoviruses.

We further characterized a subset of divergent mammalian viruses by acquiring their nearly complete or full genomes and phylogenetically comparing them to their closest viral relatives.

Proposed New Avian Parvovirus Genus

Parvoviruses are small non-enveloped viruses with linear single-stranded DNA genomes of approximately 5-kb. The family Parvoviridae is classified into two subfamilies, Parvovirinae and Densovirinae, which infect vertebrates and invertebrates, respectively [53–57]. The subfamily Parvovirinae currently consists of five genera

Viral Metagenomic Overview

The distributions of the viral hits among the ten pools of five fecal samples from Hong Kong are shown in Table 1 based on BLASTx E scores <10^{-5}. A range of best hits (BLASTx E scores <10^{-5}) to different viral families infecting plants (Geminiviridae, Nanoviridae, Caulimoviridae, Hypoviridae, Umbraviridae, Mynovirus, Togaviridae, Tombusviridae) and insects (Iflaviruses, Tenuiviruses, Chronic bee paralysis virus, Laem Singh virus) were detected (Table 1). The mammalian viral hits belonged to the Circoviridae, Parvovirinae subfamily, Picornaviridae, Rotavirus, Adenovirus, Astrovirus, and Calicivirus. The Circoviridae-related sequences showed best hits to three different clades, Circovirus (1824 reads), Calicivirus (20 reads) and Gyrovirus (2 reads). 77% of the circovirus hits showed best matches to pigeon circovirus of which 108 sequences had >90% aa-identity to pigeon circovirus. 74% of the Parvovirinae sequences showed best hits to AAV (Adeno-associated virus). A majority of the picornavirus sequences (85%; 759/889) showed best hits to Turkey hepatitis virus [52]. The adenovirus sequences showed best hits to three difference genera Mastadenovirus (3 reads), Adenovirinae (77 reads) and Sadenovirus (296 reads).

The pigeon fecal sample from Hungary was analyzed using 6,818 unique reads from 454 pyrosequencing and showed viral hits to picornaviruses,astroviruses, and tobamoviruses. We further characterized a subset of divergent mammalian viruses by acquiring their nearly complete or full genomes and phylogenetically comparing them to their closest viral relatives.

Proposed New Avian Parvovirus Genus

Parvoviruses are small non-enveloped viruses with linear single-stranded DNA genomes of approximately 5-kb. The family Parvoviridae is classified into two subfamilies, Parvovirinae and Densovirinae, which infect vertebrates and invertebrates, respectively [53–57]. The subfamily Parvovirinae currently consists of five genera

node are given if >70%. The generated phylogenetic trees were visualized using the program MEGA version 5. Resulting trees were examined for consistency with published phylogenetic trees. Sequence identity was measured using BioEdit [50]. The sequence distance comparison was calculated by using SSE [51]. Conserved amino acid analyses were performed over the alignments produced by CLUSTAL X (version 2.0.3). The hypothetical cleavage map of the picornavirus polyprotein was derived from alignments with other closest picornaviruses and NetPicoRNA prediction. Putative ORFs in the genome were predicted by NCBI ORF finder. The secondary structure of 5’ and 3’ UTR was predicted using the Mfold program.

Bioinformatics Analysis

We received 36.5 million 250-bp paired-end reads generated on the Illumina Mi-Seq platform. Short reads were debarcoded using vendor software from Illumina. A total of 6,818 reads from 454 pyrosequencing were trimmed of their primer sequence and adjacent eight nucleotides corresponding to the randomized part of the primer. In-house analysis pipelines were developed to process both datasets. Adaptors were trimmed by in-house code. The cleaned reads were de novo assembled using SOAPdenovo2 (38). The assembled contigs, along with singlets, were aligned to an in-house viral protein database using NCBI BLASTx. The significant hits to viral sequences were then aligned to an in-house non-virus-non-redundant (NVNR) universal protein database using BLASTx. Hits with more significant adjusted E-value to NVNR than to virus were removed.

Results

Viral Metagenomic Overview

The distributions of the viral hits among the ten pools of five fecal samples from Hong Kong are shown in Table 1 based on BLASTx E scores <10^{-5}. A range of best hits (BLASTx E scores <10^{-5}) to different viral families infecting plants (Geminiviridae, Nanoviridae, Caulimoviridae, Hypoviridae, Umbraviridae, Mynovirus, Togaviridae, Tombusviridae) and insects (Iflaviruses, Tenuiviruses, Chronic bee paralysis virus, Laem Singh virus) were detected (Table 1). The mammalian viral hits belonged to the Circoviridae, Parvovirinae subfamily, Picornaviridae, Rotavirus, Adenovirus, Astrovirus, and Calicivirus. The Circoviridae-related sequences showed best hits to three different clades, Circovirus (1824 reads), Calicivirus (20 reads) and Gyrovirus (2 reads). 77% of the circovirus hits showed best matches to pigeon circovirus of which 108 sequences had >90% aa-identity to pigeon circovirus. 74% of the Parvovirinae sequences showed best hits to AAV (Adeno-associated virus). A majority of the picornavirus sequences (85%; 759/889) showed best hits to Turkey hepatitis virus [52]. The adenovirus sequences showed best hits to three difference genera Mastadenovirus (3 reads), Adenovirinae (77 reads) and Sadenovirus (296 reads).

The pigeon fecal sample from Hungary was analyzed using 6,818 unique reads from 454 pyrosequencing and showed viral hits to picornaviruses,astroviruses, and tobamoviruses. We further characterized a subset of divergent mammalian viruses by acquiring their nearly complete or full genomes and phylogenetically comparing them to their closest viral relatives.

Proposed New Avian Parvovirus Genus

Parvoviruses are small non-enveloped viruses with linear single-stranded DNA genomes of approximately 5-kb. The family Parvoviridae is classified into two subfamilies, Parvovirinae and Densovirinae, which infect vertebrates and invertebrates, respectively [53–57]. The subfamily Parvovirinae currently consists of five genera
Table 1. Distribution of sequence reads to different viral species/families in 10 sample pools from Hong Kong and the single sample from Hungary.

<table>
<thead>
<tr>
<th>No. of reads</th>
<th>Sample ID</th>
<th>Mammalian</th>
<th>Insect</th>
<th>Plant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Circoviridae</td>
<td>Paroviridae</td>
<td>Picornavirus</td>
</tr>
<tr>
<td>22438</td>
<td>1</td>
<td>75</td>
<td>41</td>
<td>29</td>
</tr>
<tr>
<td>39193</td>
<td>2*</td>
<td>378</td>
<td>105</td>
<td>63</td>
</tr>
<tr>
<td>23643</td>
<td>3</td>
<td>81</td>
<td>89</td>
<td>162</td>
</tr>
<tr>
<td>19787</td>
<td>4*</td>
<td>98</td>
<td>307</td>
<td>133</td>
</tr>
<tr>
<td>90660</td>
<td>5*</td>
<td>121</td>
<td>203</td>
<td>205</td>
</tr>
<tr>
<td>8794</td>
<td>6</td>
<td>425</td>
<td>281</td>
<td>53</td>
</tr>
<tr>
<td>60389</td>
<td>7</td>
<td>100</td>
<td>41</td>
<td>50</td>
</tr>
<tr>
<td>42138</td>
<td>8</td>
<td>153</td>
<td>129</td>
<td>24</td>
</tr>
<tr>
<td>51485</td>
<td>9</td>
<td>154</td>
<td>198</td>
<td>133</td>
</tr>
<tr>
<td>75010</td>
<td>10</td>
<td>261</td>
<td>291</td>
<td>20</td>
</tr>
<tr>
<td>6818 GALII.5*</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>440355 Total</td>
<td></td>
<td>1846</td>
<td>1685</td>
<td>889</td>
</tr>
</tbody>
</table>

*Sample contained the initial novel parvovirus, rotavirus or picornavirus sequence reads that were further analyzed in the study.

SmV A: Sclerophthora macrospora virus A; CBPV: Chronic bee paralysis virus; SsHADV-1: Sclerotinia sclerotiorum hypovirulence-associated DNA virus 1; LSV: Laem Singh virus.

doi:10.1371/journal.pone.0072787.t001
Figure 1. Pigeon parvovirus genome and phylogeny. A. Genome organization of pigeon parvovirus. The alignment of the large central ORF of pigeon parvovirus and ORF78 of fowl adenovirus is shown. B. Phylogenetic analyses of NS and VP proteins of pigeon parvovirus and related parvoviruses. The scale indicated amino acid substitutions per position. ICTV approved and proposed Parvovirinae genera are shown with those containing avian parvoviruses labeled in bold font. GenBank accession numbers used are included in Table S3. doi:10.1371/journal.pone.0072787.g001
Bocavirus, Erythrovirus, Dependovirus, Amdovirus and Parvovirus; however, recently discovered genomes are expected to greatly increase that number [57,58]. Parvoviruses are widespread pathogens that can cause a range of diseases in a variety of mammals as well as birds and reptiles [57].

One hundred and five out of a total of 39,193 reads from one specimen pool (Table 1 pair 2) were related to parvovirus proteins (BLASTx E scores: $7 \times 10^{-8}$ to $5 \times 10^{-11}$). Gaps between short sequence reads were filled by PCR and then directly Sanger-sequenced. The nearly complete genome (5492-bp) of pigeon parvovirus, including a partial 5' UTR (252-bp), complete NS (631-aa), complete VP (696-aa) and a partial 3' UTR (54-bp) was acquired (GenBank KC876004). As expected the genome contained two large open reading frames (ORFs) with the left and right ORFs encoding non-structural (NS) and viral capsid protein (VP), respectively (Figure 1A). Distinct from its closest genetic relatives (chicken and turkey parvoviruses), pigeon parvovirus A contained an unusually long middle ORF (477-aa) (Figure 1A). The carboxy termini of this ORF (135-aa; position 343–477) showed best although weak similarity (BLASTx to NR E score: 8.9 $\times 10^{-2}$) to a segment of a 221-aa ORF43 predicted encoded protein from the large dsDNA fowl adenovirus genome [59], sharing 28% amino acid identity (Figure 1A). A smaller middle ORF (155-aa) was also detected, which did not show

---

**Figure 2. Mesivirus genome and phylogeny.** A. Genome organization of *Mesivirus* (*Picornaviridae*) and its sequence distance to the closest genetic relative, the turkey hepatitis virus (HQ189775) belonging to genus *Megrivirus*. B. Phylogenetic analyses of P3 regions of *Mesivirus-1* (KC876003) from Hong Kong and *Mesivirus-2* (KC811837) from Hungary and other representative picornaviruses. ICTV approved and proposed *Picornaviridae* genera are shown with those containing avian picornaviruses labeled in bold font. GenBank accession numbers used are included in Table S4.

```
doi:10.1371/journal.pone.0072787.g002
```
Table 2. Amino acid sequence identities (%) of eleven proteins of a novel pigeon rotavirus to representatives of other rotavirus species (A-H) belonging to the genus Rotavirus.

<table>
<thead>
<tr>
<th>Pigeon</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural proteins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VP1</td>
<td>20</td>
<td>59</td>
<td>20</td>
<td>19</td>
<td>22</td>
<td>91</td>
<td>58</td>
</tr>
<tr>
<td>VP2</td>
<td>12</td>
<td>53</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>94</td>
<td>45</td>
</tr>
<tr>
<td>VP3</td>
<td>9</td>
<td>45</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>84</td>
<td>25</td>
</tr>
<tr>
<td>VP4</td>
<td>10</td>
<td>32</td>
<td>11</td>
<td>9</td>
<td>11</td>
<td>35</td>
<td>20</td>
</tr>
<tr>
<td>VP6</td>
<td>12</td>
<td>45</td>
<td>14</td>
<td>11</td>
<td>14</td>
<td>67</td>
<td>41</td>
</tr>
<tr>
<td>VP7</td>
<td>12</td>
<td>42</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>51</td>
<td>17</td>
</tr>
<tr>
<td>Non-structural proteins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSP1*</td>
<td>4</td>
<td>23</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>NSP2</td>
<td>10</td>
<td>60</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td>92</td>
<td>46</td>
</tr>
<tr>
<td>NSP3</td>
<td>10</td>
<td>23</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>64</td>
<td>15</td>
</tr>
<tr>
<td>NSP4</td>
<td>14</td>
<td>16</td>
<td>10</td>
<td>6</td>
<td>9</td>
<td>68</td>
<td>6</td>
</tr>
<tr>
<td>NSP5</td>
<td>13</td>
<td>34</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>50</td>
<td>21</td>
</tr>
</tbody>
</table>

GenBank numbers of these viruses are available in Table S5.
*A major peptide NSP1–2 used for identity calculation.
doi:10.1371/journal.pone.0072787.t002

Pigeon Picornavirus

Picornaviruses are small, non-enveloped, single-stranded RNA viruses whose prototype is poliovirus. The family Picornaviridae currently consists of seventeen genera (http://www.picornaviridae.com) that infect a wide range of hosts including mammals and bird species. Many other recently characterized picornavirus genomes have increased the number of potential Picornaviridae genera [63–66]. The genome of a typical picornavirus ranges from 7 Kb to 9 Kb and, with one exception, contains a single long ORF coding for a polyprotein [67]. Picornaviruses have been reported in fecal and respiratory specimens from numerous vertebrates including humans, bats, rodents, pigs, fish and reptiles [63,68–70].

In this study we characterize two complete genomes of a novel picornavirus from pigeon fecal samples collected from Hong Kong (strain HK-21, KC876003) and Hungary (strain GALII-5/2011/HUN, KC811837) with a 9,072/9,192 nucleotide (nt) long genome [excluding the poly(A) tail] (Figure 2A). The P1, P2 and P3 regions of GALII-5/2011/HUN showed 74%, 92% and 97% amino acid identity to strain HK-21 indicating a presence of two genotypes of the same picornavirus species. Polyprotein alignments revealed that strain GALII-5/2011/HUN had an alternative upstream start codon MREY instead of the putative start codon of MATF seen in strain HK-21 (Figure S2). A stop codon was found between MREY and MATF in strain HK-21 but none was found in strain GALII-5/2011/HUN. The start codon in the upstream MREY in GALII-5/2011/HUN was also located in a weaker Kozak consensus sequence [GGGGATGCGG] which plays a major role in the initiation of the translation process than MATF [GGAGATGGCG] (Figure S2). We therefore provisionally named these picoenaviruses as Mesivirus-1 (strain HK-21) and Mesivirus-2 (strain GALII-5/2011/HUN) for Mesivirus sister-clade virus.

The mesivirus 5’ UTR contained a highly conserved, 20 nucleotides long motif and apical “8”-like structure 407GGAGGTTTGCTGAATATTGGAACGTCCAC [457] [with unknown function] that were also seen in avian origin picornaviruses, including turkey hepatitis virus (genus Megavirus), duck hepatitis A virus-1 (genus Aviparvovirus), quail picornavirus (unassigned genus) and pigeon picornavirus B (unassigned genus) [52,71,72]. All these picornaviruses had type IV-like IRES [http://www.picornaviridae.com], however, we could not draw the potential secondary structure of mesivirus IRES. The hypothetical cleavage map of the polyproteins of mesiverses were derived from alignments with other known picornaviruses and described from alignments with other known picornaviruses and NetPicoRNA prediction [73]. Mesiverses did not have a putative L protein preceding the capsid region. The P1 regions were 807/809-aa (Mesivirus-1/Mesivirus-2) in length containing rhf-like superfamily domains (Figure 2A) revealed by Conserved Domain Database (CDD) search [74]. The conserved motif GXXXT/S for myristylation was also absent. Similar to the megrivirus, there was no potential VP2/VP4 cleavage site in mesiverses. The P1 polyproteins of mesivirus-1 and -2 could be theoretically cleaved at VP0/VP3 (Q389–390 A/Y/T), VP3/VP1 (Q557/558 D/G) and VP1/2A1 (Q807/809 D/E). BLASTx showed that the P1 region had the highest aa-identity of 43%-41% to the megrivirus (strain 0091.1, HQ189775) (Table S2). The P2 polyproteins (1019/1021-aa) encoded non-structural proteins of mesivirus-1 and -2 cleaved at 2A1/2A2 (G1065/1059 R), 2A2/2B (Q1290/1284 D/A), 2B/2C (E1508/1491 D/A) and 2C/3A (E1626/1630 D/A) (Table S2). Similar to the megrivirus Turkey hepatitis virus, both mesivirus genomes encoded an unusual 2A1 (208/290 aa) at the N-terminal end of the P2 polyprotein. When mesivirus and mesivirus polyproteins where aligned their 2A1 regions showed the weakest homology among their genomes with only the central part of 2A1 (28%) showing 31% identity to that of the megrivirus 2A1 strain 0091.1 (accession number HQ189775) (Figure S1A). The 195-aa-long 2A1 proteins of both mesiverses contained an HBox/NC domain (H1127/1131 WG and N1188/1192 CT) that was also observed in the...
199-aa-long 2A2 of megriviruses [75]. Characteristic features of picornaviruses, including the RNA-helicase superfamily and NTPase motifs, were identified in the mesivirus 2C proteins. The P3 of megriviruses were 881-aa in length encoding proteins 3A, 3B (VPg, small genome-linked protein), 3Cpro (protease) and 3Dpol (RNA-dependent RNA polymerase). The P3 polypeptides were cleaved at 3A/3B (E2008/2012 Q A), 3B/3C (E2036/2040 Q G) and 3C/3D (Q2233/2237 Q G) (Table S2). The 3Cpro contained the H-D-C catalytic triad and conserved GXCG [G2190/2194 FCG] motif known as an active site of protease [76]. The 3Dpol contained conserved RdRp motifs K2298/2302 ELR, Y2367/2371 GDD and E2485/2489 LKR. It also had a conserved motif G2527/2531 GMPSG that was also seen in other picornaviruses [77].

Phylogenetic analysis of P1–P3 also showed that mesiviruses were most closely related to megrivirus (Figure 2B and Figure S1B) and in pair-wise comparison mesivirus-1/mesivirus-2 had identities of 43%/41%, 34%/30% and 47%/47% to turkey megrivirus (strain 0091.1) at the amino acid level for the P1, P2 and P3 regions respectively (Table S2). According to the ICTV, the member of a picornavirus genus should share ≥40%.
50% aa-identities in P1, P2 and P3 regions respectively (http://www.picornaviridae.com). Although the identities of the P2 and P3 regions qualified mesiviruses as belonging to a novel genus, its P1 region shared more than the 40% aa-identity threshold to megrivirus. Mesiviruses might therefore be considered as a new picornavirus species in the genus *Megrivirus* or given their distinct 2A1 and high P2 and P3 divergence to megrivirus as members of a separate genus.

**Pigeon Rotavirus**

Rotaviruses consist of at least eight groups or species (A through H) with multiple P and G genotypes which together comprise a genus in the family *Reoviridae*. Rotavirus has a non-enveloped,
The six structural proteins (VP1–VP7) showed best hits by BLASTx to group G rotavirus, first described in chicken feces from Northern Ireland [81]. While VP1–VP3 shared 54%–94% aa-identities with group G rotavirus, other structural proteins showed lower aa-identities, ranging from 33% to 67% (Table 2). Classification of rotavirus is based mainly on the inner capsid protein VP6 and according to ICTV definition the members of the same rotavirus species should share >60% nt-identity in VP6 region [82]. The sequence (1176-n) of the VP6-encoding genome segment shared 68% nt-identity to group G rotavirus (Figure 3A). Direct observation of VP6 alignment between HK18 and its closest relatives revealed numerous mismatches in the group-specific antigenic regions (Figure S3), indicating that HK18 represents a candidate prototype for a novel subgroup within group G. Two structural proteins VP4 (the protease-cleaved polyprotein) and VP7 (the glycoprotein) on the outermost surface of rotavirus are targets for neutralizing antibodies and have also been used to classify rotaviruses into their P and G genotypes, respectively [83]. The nucleotide sequence alignments of HK18 VP4 and VP7 showed low identity to other rotaviruses (Figure 4A). In addition, HK18 and group G rotavirus shared only 35% for VP4 and 51% for VP7 at the amino acid level, suggesting that HK18 should be considered as novel P and G genotypes, tentatively proposed as G2P [2] within group G rotavirus. The VP1 protein of HK18 contained three conserved RNA-dependent RNA polymerase motifs 558AEKII560, 568LSN571 and 608 KKI RYLGVASEK704T573 and 683 LG VASGEK704T573, which has unknown function [84]. Similar to group G rotavirus, the VP3 protein of HK18 did not possess the NTP-binding sequence motif present in group C rotavirus [85]. However, an alignment of VP3 sequences from different rotavirus groups showed that the VP3 protein of HK18 shared a high degree of conservation in ALYXLSN [ALYXLSN] motif, which has unknown function [86]. In order to phylogenetically classify pigeon rotavirus HK18, its six structural proteins were aligned to other rotavirus representatives. While the VP1–VP3, VP6 and VP7 of HK18 were clustered with those of group G rotavirus (Figure 3B and Figure S4), its VP4 was located on a branch diverging closer to the root (Figure 4B).

The pair-wise amino acid sequence analysis demonstrated that nonstructural proteins NSP2–NSP5 shared top identities to group G rotavirus rather than other groups, ranging from 50% to 92% at the amino acid level (Table 2). However, NSP1 had a very low aa-identity, less than 31% to group G rotavirus. Like group G rotavirus, the NSP1 gene of HK18 had two ORFs encoding two minor and major peptides labeled as NSP1–1 (104-aa) and NSP1–2 (310-aa), respectively. While NSP1–1 had the best identity of 30% to group G rotavirus, NSP1–2 shared the closest match to group B rotavirus with 23% identity versus 19% to group G. The NSP1 zinc-binding domain could not be identified in HK18. The NSP2 was found to have a conserved sequence HGXGHXXRVV and histidine triad (His-X-His-X-His-XX) located at the RNA binding domain [87]. Phylogenetic analysis of NSP2–NSP5 showed that the group G rotavirus was the closest relative to HK18 (Figure S4) in all fragments except for NSP1 where it appeared basal to both groups G and B. A similar observation was also made for avian group A rotaviruses where their NSP1 could not be grouped into taxonomic species [88].

**Discussion**

Metagenomics has been used to analyze viral nucleic acids in feces collected from humans and a growing range of animals including primates, horse, bats, rodents, pigs, dogs, and turkeys [42,58,65,69,89-92]. We investigated viral sequences in 51 wild pigeon droppings followed by viral sequence similarity searches revealing sequences closely or distantly related to viral genomes in GenBank. The eukaryotic viruses detected included insect and plant viruses likely reflecting the pigeon diet as well as viruses already known to infect pigeons such as pigeon circovirus and adenoviruses. Also detected and characterized were novel genomes in viral families known to infect vertebrates.

Circovirus-like inclusion bodies were initially identified in the bursa of pigeons [93,94] and the molecular epidemiology of pigeon circovirus worldwide has confirmed its importance as a pathogen associated with a wide range of illnesses in pigeon populations, including weight loss, respiratory distress and diarrhea [94–96]; possibly as a result of induced immunodeficiency aggravating the pathogenicity of co-infections such as those of adenoviruses [97]. Both pigeon circovirus and diverse adenovirus sequences were identified here.

Parvoviruses cause a variety of mild to severe symptoms in birds [98–101]. The *Parvoviridae* subfamily has recently undergone a large expansion in the number of known genera and species [58,62,102,103]. Here a novel parvovirus was characterized in pigeon feces that was distantly related to chicken and turkey parvoviruses found in feces of farm animals with signs of enteric diseases [62,98]. The pigeon parvovirus genome, the first from host species, possessed an unusually long middle ORF that encoded a protein showing similarity to that encoded by an ORF of unknown function in the fowl adenovirus genome. This observation might reflect past lateral gene exchange between avian DNA viruses replicating in the nucleus. The pigeon, chicken and turkey parvoviruses cluster together phylogenetically, but have not yet been assigned to a *Parvovirinae* genus. We therefore propose a new genus provisionally called *Aviparvovirus* containing at least two species, pigeon parvovirus and the closely related turkey/chicken parvoviruses. Two related picornaviruses (81% nucleotide similarity) from pigeons in Hong Kong and Hungary were also characterized and provisionally named *Mesivirus*-1 and -2 reflecting a wide geographic distribution of this viral clade. Based on the phylogenetic analysis of the most conserved region (P3) many of the bird picornaviruses clustered together (*Gallivirus*, *Orthotogavirus*, *Paratogivirus*, *Mogievirus*, and *Mesivirus*) in a supported clade that includes only the mammalian infecting *Kobavirus* and *Salivirus* genera. It is therefore possible to speculate on the past existence of a strictly avian clade a member of which adapted to mammals later resulting in the *Salivirus/Kobavirus* clade.

Rotaviruses were first described as a causative agent of gastroenteritis in humans in 1973 by using electron microscopy to examine biopsies of duodenal mucosa from children with acute non-bacterial gastroenteritis [104]. The first pigeon rotavirus was isolated from feces in 1983 and belongs to group A [105]. We describe here the complete coding regions of all eleven segments of a group G rotavirus, the first from pigeons, distinct from chicken rotavirus [83]. The addition of a second pigeon rotavirus genome to the viral database also shows that the diversity of known avian rotaviruses is likely to continue to expand with multiple rotavirus species infecting the same host species.

Animal species living in or around human habitation, including some bats and rodents species and pigeons, are recognized as the reservoirs of multiple zoonotic pathogens. Because of increasing
contact between such peridomestic animals and humans a better understanding of these animals' virome can inform future studies of their cross-species transmissions potential. Monitoring viral exchanges between pigeons and highly exposed humans using nucleic acids or serological assays for avian viruses will be facilitated by improved knowledge of their virome.

Supporting Information

Figure S1 Alignment of putative 2A1 sequences of Mesivirus-1 and Megavirus. B Phylogenetic trees of P1 and P2 regions of Mesivirus and other picornavirus genera in the family Picornaviridae. (PDF)

Figure S2 Alignment of N-terminal amino acid (upper) and nucleotide (lower) sequences containing start and stop codons of two pigeon Mesiviruses. (PDF)

Figure S3 Alignment of VP6 proteins of the novel pigeon rotavirus and other closely-related rotavirus species. (PDF)

Figure S4 Phylogenetic trees of structural (shaded box) and non-structural proteins of the novel pigeon rotavirus and other rotavirus species (A-H). (PDF)

Table S1 Pairwise amino acid sequence identities (%) between NS and VP regions of the novel pigeon parvovirus, turkey parvovirus and representatives of parvovirus genera. GenBank numbers of these viruses are available in Table S3. (PDF)

Table S2 Coding potential/putative proteins of the genome of Mesivirus-1 and comparison of amino acid sequence identity (%) of its P1–3 and other closely-related picornaviruses in the Picornaviridae family. (PDF)

Table S3 Representative members in the subfamily Parovirinae for the phylogenetic trees in Figure 1 and their GenBank numbers. *Paroviruses were used for pairwise calculation in Table S1. (PDF)

Table S4 Representative members in the family Picornaviridae for the phylogenetic tree in Figure 2 and their GenBank numbers. (PDF)

Table S5 Representatives and their GenBank numbers in the genus Rotavirus for phylogenetic trees in Figures 2, 4, S4. *Rotaviruses were used for pairwise calculation in Table 2. (PDF)

Author Contributions

Conceived and designed the experiments: TGP AB GR LLMP ED. Performed the experiments: TGP NPV AB GR. Analyzed the data: TGP NPV AB GR CW XD ED. Contributed reagents/materials/analysis tools: TGP NPV AB GR ED. Conceived and designed the experiments: TGP AB GR LLMP ED. Wrote the paper: TGP AB GR ED.

References


