Genetic characterization of two novel megriviruses in geese

Fumin Wang,1 Te Liang,1 Ning Liu,1 Kang Ning,1 Kangzhen Yu2 and Dabing Zhang1,*

Abstract

Two goose megriviruses (W18 and HN56) were detected and sequenced. Both viruses possessed megrivirus-like genomic features, including unusually long genomes (9840 and 10101 nt). W18 and HN56 were very similar to duck megrivirus (DMV) in the P2 and P3 regions, but much less similar in the P1 and 2A1 regions. HN56 may be a potential recombinant virus, with a distinct P1 region possibly originating from an unknown picornavirus. W18 may represent a novel type of DMV, showing a P1 sequence identity of 67%. Similar levels (64–68%) of P1 sequence identity were also displayed by melegrivirus A with W18 and DMV, demonstrating an equal genetic separation of the capsid region among W18, DMV and melegrivirus A. For the 2A1 region, the divergence among W18, HN56 and DMV was remarkable, involving point mutations and a long insertion/deletion. The present work contributes to the understanding of unique features and phylogeny of megriviruses.

The genus Melegrivirus, one of the 31 official genera within the family Picornaviridae, was established in 2013 [1]. To date, the genome sequence has been determined for 11 megrivirus strains and 10 of the genomes are over 9 kb. Megivirus genome has a picornavirus-like layout, but also several unique features, such as a large insertion of 0.86–1.2 kb between VP1 and a recognizable 2A-like region containing H-box/NC motifs. The insertion sequence may encode an additional one to three distinct 2A polypeptides of unknown function [2–8]. Some megriviruses identified in chicken and turkey have long 3′UTRs which were predicted to contain a potential second ORF [2, 9].

The genus currently contains a single official species, namely Melegrivirus A (MeV-A) (www.picornaviridae.com/megivirus/megivirus.htm), with turkey hepatitis virus 1 (THV-1) strains 2993D and 0091.1 as prototypes [3]. Recently described chicken picornavirus 4 (ChPV4) 5C from Hong Kong, China, and turkey megrivirus B407 from Hungary may be members of the species MeV-A [2, 6]. THV-1-like partial 3′UTR sequences have been amplified from chicken and turkey in the USA [9].

Apart from the type species, there are at least three groups of potential novel megriviruses waiting to be classified. Mesiviruses, detected in pigeon fæces collected from Hong Kong, PR China (mesivirus-1 HK21) and Hungary (mesivirus-2 GALII5), share low amino acid (aa) identities of 41–43, 30–34 and 47 % in the P1, P2 and P3 regions, respectively, with MeV-A 0091.1. Therefore, it may be a novel species within the genus Melegrivirus [8]. Liao et al. [7] described the characterization of a novel megrivirus, named duck megrivirus (DMV), which was detected in Pekin ducks in China. DMV is distinct from previously reported megiviruses, but shares a higher level of aa identity with MeV-A in the capsid region (67–68%) compared with those in the nonstructural proteins (<52%) [7]. In 2014/2015, four strains of novel megriviruses were identified in chickens. They are chicken megrivirus (CMV) B21 and CHK-IV from Hungary [2], ChPV5 27C from Hong Kong [6] and chicken proventriculitis virus (CPV) 03 from South Korea [5]. Pairwise comparisons reveal that the four CMVs are very closely related to one another (>91% aa identity in the P1, P2 and P3 regions), except that ChPV5 27C shares a lower aa identity value with other three strains in the capsid region (74%). This suggests that the four CMV strains may belong to two genotypes of the same species. Interestingly, the chicken viruses exhibit discordant phylogenetic relationships in the structural and nonstructural proteins when compared with MeV-A (aa identity: P1, 36–37%; P2, 80–83%; P3, 96–97%) [2, 5, 6]. This phenomenon will have an impact on determination of the taxonomic relationship of CMV to MeV-A. Here, we report the genetic characterization of two novel megriviruses in geese, which we named goose megrivirus (GMV) provisionally.
In June 2014, a disease with sporadic death occurred in five flocks of geese on a farm located in Hunan Province, PR China. The farmer reported that the main gross pathological change was enteritis. Six fresh faecal samples and 12 tissues (e.g. intestine, liver and proventriculus) were sampled from the diseased geese and submitted to our lab for the purpose of diagnosis. In May 2015, a disease occurred again in two flocks of geese on the farm. The disease resulted in losses of approximately 20% in 15- to 30-day-old geese. Gross pathological changes were found chiefly in the liver, which was covered with small white necrotic foci. We collected 24 tissue samples including liver, spleen, kidney and intestine from dead geese after autopsy. Two intestine and two liver were selected from the 2015 samples for testing avian influenza virus, Newcastle disease virus and goose parvovirus, which were all negative. Subsequently, RNA was extracted from a mixture of aliquots of the four samples and subjected to random amplification [7, 10] (Supplementary Material). Following sequencing and sequence analysis, four megrivirus-like sequences were derived and assembled into two fragments of 809 and 293 bp. The two sequences were shown by BLASTx analysis in GenBank (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to encode partial VP0–VP3 and 3C–3D proteins of megrivirus, which exhibited the highest identities of 72 and 100% with DMV, respectively. The remainder of the genome was determined using the RNA extracted from a single intestinal sample (W18) by reverse transcription (RT) PCR and 5′ and 3′ RACE [11] with primers (Table S1, available in the online Supplementary Material) designed based on sequences of the two fragments and the DMV genome. Once the complete genomic sequence had been determined, it was confirmed by sequencing overlapping DNA fragments generated by RT-PCR with W18-specific primers (Table S2).

When RNA extracted from a faecal sample (designated HN56) collected in 2014 was amplified with primers 2Cf and 2Cr (Table S3) designed over the conserved 2C region of picornaviruses, two fragments of 315 and 1538 bp were amplified. BLASTx analyses revealed that the 315 and 1538 bp sequences encoded VP3–VP1 and 2A–2C proteins of megrivirus, which showed maximum aa identities with megrivirus-1 (56%) and DMV (94%), respectively. Based on the sequences of the two fragments together with the DMV and W18 genomes, additional primers (Table S3) were designed and used to amplify the remainder of the HN56 genome. The initial genomic sequence was verified by sequencing overlapping DNA fragments generated by RT-PCR with HN56-specific primers (Table S4).

Sequence analysis using DNAMAN 5.2.2 (Lynnon) showed that the GMV W18 and HN56 genomes comprised 9840 and 10101 nt, respectively, excluding the poly(A) tail. The former was close to that (9883 nt) of sicinivirus Pf-CHK1/SiV, the largest picornavirus genome [12], whereas the latter was larger than that of Pf-CHK1/SiV, making it the largest among picornaviruses sequenced to date. The GC content of both W18 and HN56 was 46.2%, similar to those of the 0091.1 (46.2%), 2993D (46.1%) and B407 (45.6%) isolates of MeV-A, the GALI15 (46.8%) isolate of mesivirus and the LY isolate (45.3%) of DMV (Table 1). The W18 and HN56 genomes contained single ORFs of 8658 and 8919, encoding putative polyproteins of 2885 and 2972 aa, respectively. The ORFs of W18 and HN56 were preceded by 5′UTRs of 519 and 520 nt, and followed by 3′UTRs of 663 and 662 nt, respectively (Fig. 1). Unlike previously reported megriviruses (e.g. B407, B21 and CHK148) [2, 9], W18 and HN56 appeared to lack an additional ORF in the 3′UTR.

Based on sequence alignment with their closest relative DMV, the possible cleavage sites of the W18/HN56 polyproteins were predicted to be VP0|VP3 (Q 377/E 378) [A], VP3|VP1 (Q 625/626) [G], VP1|2A1 (Q 892) [N], 2A1|2A2 (E 950/1037) [A], 2A2|2A3 (E 1244/1331) [A], 2A3|2B (Q 1461/1548) [N], 2B|2C (E 1656/1743) [A], 2C|3A (Q 2001/2088) [A], 3A|3B (E 2180/2267) [A], 3B|3C (E 2209/2296) [G] and 3C|3D (E 2411/2498) [A]. Thus, the studied polyproteins may comprise three capsid proteins (VP0, VP3 and VP1) and nine nonstructural proteins (2A1–2C and 3A–3D) (Fig. 1). Using the CD-search [13], Rhv-like domains (pfam entry, cd00205) were detected in the P1 polyproteins of W18 (residues 151–337, 406–608 and 649–845) and HN56 (residues 184–342, 413–614 and 658–847). Examination revealed that the 2C, 3C and 3D proteins of W18/HN56 contained characteristic motifs conserved in picornaviruses, including motifs typical for helicase (G 1792/1879 xxGxGKS and D 1843/1938 DxxQ), cysteine protease (G 2368/2455 xCG) and RNA polymerase (K 2577/2664 DELR, D 2655/2742 xXXD, G 2796/2793 xXTSG, y 2746/2833 xGDD and E 2795/2883 LKR). Similar to other megriviruses [2, 3, 5–8], W18/HN56 also contained a long insertion of 1056/1317 nt between VP1 and a recognizable 2A region containing H-box/NC motifs [4]. The insertion sequence may encode an additional two distinct 2A polyproteins of unknown function, namely 2A1 (58/145 aa) and 2A2 (294 aa). The recognizable 2A H-box/NC was designated 2A3 (217 aa).

To investigate the relationship of GMVs to other megriviruses, nucleotide identities for UTRs and aa identities for the P1, P2 and P3 regions were deduced by CLUSTAL W (http://www.genome.jp/tools/clustalw/) (Table 1). Based on alignments of the P1, 2C and 3CD aa sequences, phylogenetic trees (Fig. 2) were constructed by MEGAl6 [14]. W18 and HN56 were very closely related to each other and to DMV in P2 (>91% identity), P3 (>96%) and UTRs (>88%). In the P1 region, however, W18 and HN56 were highly divergent from each other and from other megriviruses. Of the GMVs, HN56 shared low levels of capsid sequence identity with DMV (45%) and W18 (50%), exhibiting remarkable sequence heterogeneity between P1 and other genomic regions (5′UTR and P2–P3–3′UTR). Interestingly, the closest relative of the HN56 capsid sequence was that of megrivirus-1 although the identity was still low (50%). The different topology patterns revealed by the phylogenetic trees of different regions for HN56 suggest that a recombination event may have occurred between DMV-like virus and an unknown, mesivirus-associated ancestral.
picornavirus with a distinct P1 region. The level of divergence (67% identity) between W18 and DMV was comparable to that which separates different types in a picornavirus species [15]. Thus, W18 could be regarded as a novel type of DMV. Interestingly, similar levels of capsid sequence identity were also displayed by W18 and MeV-A (64–65%) and by DMV and MeV-A (67–68%) (Table 1) [7], reflecting an equal genetic separation among W18, DMV and MeV-A. This may suggest that their P1 regions came from a common ancestor virus. Notably, both W18 and DMV displayed much lower sequence identities in the P2 (<37%) and P3 (<53%) regions with MeV-A (Table 1) [7], indicating that W18 and DMV are highly divergent from MeV-A in the nonstructural regions. Together these data suggest that the structural and nonstructural regions of megriviruses may evolve independently, potentially generating recombinant viruses with distinct phylogenetic positions.

Detailed domain comparisons further confirmed the above observations and showed several novel findings. First, the levels of diversity in the VP0, VP3 and VP1 proteins between W18 and DMV were comparable to those between different picornavirus types. In VP3, however, W18 shared higher identity (72–74%) with MeV-A than with DMV (70%) (Table S5). Second, the levels of VP0, VP3 and VP1 identities between W18/HN56 and other megriviruses. Values are nucleotide identities for 5’ and 3’ UTRs and aa identities for P1, P2 and P3 proteins. §Identity between W18 and HN56.

Table 1. Genomic features of members in the genus Megrivirus and identities between the studied virus and other megriviruses

<table>
<thead>
<tr>
<th>Virus*</th>
<th>Abbreviation</th>
<th>Strain†</th>
<th>GenBank accession no.</th>
<th>Genomic features</th>
<th>Identity (§%)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Size (nt)</td>
<td>GC content (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melegrivirus A</td>
<td>MeV-A</td>
<td>2993D</td>
<td>HM751199</td>
<td>9 075</td>
<td>46.1</td>
</tr>
<tr>
<td></td>
<td>0091.1</td>
<td>HQ189775</td>
<td>9 040</td>
<td>46.2</td>
<td>55/56</td>
</tr>
<tr>
<td></td>
<td>B407</td>
<td>KP961188</td>
<td>9 739</td>
<td>45.6</td>
<td>56/54</td>
</tr>
<tr>
<td></td>
<td>ChPV4/5C</td>
<td>KP979335</td>
<td>9 564</td>
<td>44.7</td>
<td>57/58</td>
</tr>
<tr>
<td>Mesivirus</td>
<td>Mesivirus-1</td>
<td>HK21</td>
<td>KC876003</td>
<td>9 072</td>
<td>47.5</td>
</tr>
<tr>
<td></td>
<td>Mesivirus-2</td>
<td>GALI15</td>
<td>KC811837</td>
<td>9 192</td>
<td>46.8</td>
</tr>
<tr>
<td>Chicken megrivirus</td>
<td>CMV</td>
<td>B21</td>
<td>KP961186</td>
<td>9 560</td>
<td>44.8</td>
</tr>
<tr>
<td></td>
<td>CHK-IV</td>
<td>KP961187</td>
<td>9 566</td>
<td>44.9</td>
<td>56/57</td>
</tr>
<tr>
<td></td>
<td>CPV/03</td>
<td>KJ690629</td>
<td>8 851</td>
<td>44.1</td>
<td>59/58</td>
</tr>
<tr>
<td></td>
<td>ChPV5/27C</td>
<td>KP979336</td>
<td>9 591</td>
<td>43.8</td>
<td>56/55</td>
</tr>
<tr>
<td>Duck megrivirus</td>
<td>DMV</td>
<td>LY</td>
<td>KC663628</td>
<td>9 700</td>
<td>45.3</td>
</tr>
<tr>
<td>Goose megrivirus</td>
<td>GMV</td>
<td>W18</td>
<td>KY369299</td>
<td>9 840</td>
<td>46.2</td>
</tr>
<tr>
<td></td>
<td>HN56</td>
<td>KY369300</td>
<td>10 101</td>
<td>46.2</td>
<td>–</td>
</tr>
</tbody>
</table>

* Italic, official species. Roman, unassigned species.
† ChPV4/5C, originally named chicken picornavirus 4 (strain 5C) [6], shared high aa identities with MeV-A and was regarded as a member of the species Melegrivirus A in the present study. CPV/03 and ChPV5/27C, originally named chicken proventriculitis virus (strain 03) and chicken picornavirus 5 (strain 27C), respectively [5, 6], shared high aa identities with CMV B21 and CHK-IV and were regarded as members of unassigned species ‘Chicken megrivirus’.
‡ Identity between W18/HN56 and other megriviruses. Values are nucleotide identities for 5’ and 3’ UTRs and aa identities for P1, P2 and P3 proteins. §Identity between W18 and HN56.

Table 1. Genomic features of members in the genus Megrivirus and identities between the studied virus and other megriviruses

* Italic, official species. Roman, unassigned species.
† ChPV4/5C, originally named chicken picornavirus 4 (strain 5C) [6], shared high aa identities with MeV-A and was regarded as a member of the species Melegrivirus A in the present study. CPV/03 and ChPV5/27C, originally named chicken proventriculitis virus (strain 03) and chicken picornavirus 5 (strain 27C), respectively [5, 6], shared high aa identities with CMV B21 and CHK-IV and were regarded as members of unassigned species ‘Chicken megrivirus’.
‡ Identity between W18/HN56 and other megriviruses. Values are nucleotide identities for 5’ and 3’ UTRs and aa identities for P1, P2 and P3 proteins.
§ Identity between W18 and HN56.

picornavirus with a distinct P1 region. The level of divergence (67% identity) between W18 and DMV was comparable to that which separates different types in a picornavirus species [15]. Thus, W18 could be regarded as a novel type of DMV. Interestingly, similar levels of capsid sequence identity were also displayed by W18 and MeV-A (64–65%) and by DMV and MeV-A (67–68%) (Table 1) [7], reflecting an equal genetic separation among W18, DMV and MeV-A. This may suggest that their P1 regions came from a common ancestor virus. Notably, both W18 and DMV displayed much lower sequence identities in the P2 (<37%) and P3 (<53%) regions with MeV-A (Table 1) [7], indicating that W18 and DMV are highly divergent from MeV-A in the nonstructural regions. Together these data suggest that the structural and nonstructural regions of megriviruses may evolve independently, potentially generating recombinant viruses with distinct phylogenetic positions.

Detailed domain comparisons further confirmed the above observations and showed several novel findings. First, the levels of diversity in the VP0, VP3 and VP1 proteins between W18 and DMV were comparable to those between different picornavirus types. In VP3, however, W18 shared higher identity (72–74%) with MeV-A than with DMV (70%) (Table S5). Second, the levels of VP0, VP3 and VP1 proteins between W18 and DMV were comparable to those between different picornavirus types. In VP3, however, W18 shared higher identity (72–74%) with MeV-A than with DMV (70%) (Table S5). Second, the levels of VP0, VP3 and VP1 proteins between W18 and DMV were comparable to those between different picornavirus types. In VP3, however, W18 shared higher identity (72–74%) with MeV-A than with DMV (70%) (Table S5).

![Fig. 1. Genome organization of goose megriviruses W18 and HN56. Numbers indicate lengths in nucleotides (for UTRs) and amino acids (for each protein). The predicted cleavage sites are shown above the gene box of W18. The cleavage site VP0/VP3 of HN56 was different from that of W18 and is indicated independently.](image-url)
aa sequence identity shared by HN56 with W18 and DMV were generally low, ranging from 40 to 52 %. By contrast, HN56 shared a higher level of identity (61–63 \%) with mesi-viruses in VP3 (Table S6). Third, the 2A1 sequence diversity among duck and GMVs was remarkable. Compared with W18 and DMV, HN56 contained an insertion of 87 aa at the 3' portion of 2A1 (Fig. S1), suggesting the evolution of duck and GMVs may involve a large insertion/deletion of a 261 nt sequence. In the 174 nt 2A1 region (excluding the insertion/deletion sequence), W18 and HN56 shared 65–67 \% nt identity and 62 \% aa identity with DMV, suggesting that GMVs and DMV have been separated for a long time and that the difference between GMVs and DMV is most likely the result of point mutations accumulated over time. W18 and HN56 shared 81 \% nt identity and 91 \% aa identity with each other and with DMV, indicating that most mutations are synonymous.

The region containing internal ribosome entry site (IRES) for W18 and HN56 shared identities of 91 and 89 \% with that of DMV, respectively, suggesting that they may possess a variant-type IVB IRES with domains II and III and a long helix III4 ending with the ‘8’-like 20 nt-long conserved structure at the top of domain III (Fig. S2), like those of DMV and other megriviruses [2, 6–8].

To understand the prevalence of megrivirus infections in goose flocks on the farm, the field samples collected in 2014 and 2015 were tested using RT-PCR methods with two pairs of primers specific to the VP1 genes of W18 and HN56, respectively (Supplementary Material). Of the 42 samples, 23 (54.8 \%) were positive using W18-specific RT-PCR, and 12 (28.6 \%) were positive using HN56-specific RT-PCR. In total, 35 (83.3 \%) samples tested positive for megriviruses.

For virus isolation, 0.3 ml filtrate produced from the field sample was inoculated into yolk sac of five 6-day-old specific pathogen-free embryonated chicken eggs. Attempts to isolate the HN56 virus were unsuccessful. However, the W18 was readily propagated in the yolk sac of chicken eggs. All embryos inoculated with the W18 sample died within 2–5 days postinoculation. On subsequent four passages, 100 \% mortality of embryos was also observed, with death occurring between 2 and 10 days postinoculation. Dead embryos exhibited gross lesions similar to those caused by THV-1 [16]. Using the W18-specific RT-PCR method, yolk fluids harvested from each passage were all positive for the W18 megrivirus. The full-length genome sequence of the fifth passage virus was determined, which was shown to have only three nucleotide changes (A927G, G6626A and C8694T) compared with the W18 virus that was present in the field sample. Unfortunately, the disease observed in the field failed to be repeated by inoculation of the fifth passage virus into 10-day-old goslings.

Taken together, although the causative agent of the goose disease remains to be clarified, we have identified two novel GMVs. This suggests that geese may be an important reservoir for diverse picornaviruses. The GMVs possess numerous unique genome features, such as an unusually long genome, a large insertion of 1056–1317 nt between VP1 and a recognizable 2A-like region containing H-box/NC motifs, the variant-type IVB IRES and heterogeneous sequence relationships between structural and nonstructural regions. Our data, in conjunction with previous evidences [2, 5–7], support the view that sequence heterogeneity between structural and nonstructural genome regions may be a general phenomenon in some members of the genus *Megrivirus*, which could be regarded as a unique genome feature of...
these viruses. The present work contributes to the understanding of unique features and phylogeny of megriviruses, which may be of help in the classification of megriviruses.

Funding information
This study was funded by the China Agriculture Research System (CARS-43).

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
All procedures involving animals were approved by the Animal Welfare and Ethical Censor Committee at China Agricultural University, and Beijing Administration Committee of Laboratory Animals under the leadership of Beijing Association for Science and Technology (Approval ID SYXK [Jing] 2015-0028).

References


Five reasons to publish your next article with a Microbiology Society journal
1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.