Genome characterization of Turkey Rotavirus G strains from the United States identifies potential recombination events with human Rotavirus B strains

Fangzhou Chen,1,2 Todd P. Knutson,1 Robert E. Porter,1 Max Ciarlet,3 Sunil Kumar Mor1 and Douglas G. Marthaler1,4,*

Abstract
Rotavirus G (RVG) strains have been detected in a variety of avian species, but RVG genomes have been published from only a single pigeon and two chicken strains. Two turkey RVG strains were identified and characterized, one in a hatchery with no reported health issues and the other in a hatchery with high embryo/poult mortality. The two turkey RVG strains shared only an 85.3% nucleotide sequence identity in the VP7 gene while the other genes possessed high nucleotide identity among them (96.3–99.9%). Low nucleotide percentage identities (31.6–87.3%) occurred among the pigeon and chicken RVG strains. Interestingly, potential recombination events were detected between our RVG strains and a human RVB strain, in the VP6 and NSP3 segments. The epidemiology of RVG in avian flocks and the pathogenicity of the two different RVG strains should be further investigated to understand the ecology and impact of RVG in commercial poultry flocks.

Rotaviruses (RVs) are causative agents of acute viral gastroenteritis and diarrhoea in multiple animal species [1]. Belonging to the family Reoviridae, RVs consist of a non-enveloped virion encasing 11 double-stranded RNA segments, which encode for 5–6 non-structural proteins (NSP1–NSP5/NSP6) and 6 viral proteins (VP1–4, VP6 and VP7) [2]. The outer capsid of the virion is composed of the VP7 and VP4 proteins while the middle and inner layers of the capsid are formed by VP6 and VP2 proteins, respectively. The VP1 protein is an RNA-dependent RNA polymerase that coordinates rotavirus RNA packing and genome replication [3]. The VP3 protein is a multi-functional protein associated with guanylyltransferase [4] and methyltransferase [5] activities. The NSP1 protein is an important innate host immune response regulator that can inhibit expression of type I interferon [6] and counteract NFκB [7]. The NSP2 and NSP5 proteins mediate the assembly of viroplasm [8, 9]. The NSP3 protein can regulate viral mRNA translation by simultaneous interaction with eukaryotic translation initiation factor eIF4G and the mRNA 3' end [10]. The NSP4 protein is a multi-functional viral viral enterotoxin [11].

Currently, RV species are classified by the inner capsid protein VP6, and strains within the same rotavirus species share >60% nucleotide identity [12]. Eight official RV species have been identified (RVA-RVH), and two tentative new RV species (RVI and RVJ) were identified in dogs [13] and bats [14], respectively. Due to their high prevalence and pathogenicity, RVA is well characterized and has been identified in both mammalian and avian species [15, 16]. RVB, RVC, RVE and RVH have detected in various mammalian species, while RVD, RVF and RVG have been exclusively identified in different avian species [1]. Rotavirus-like viruses were first identified in intestinal contents specimens from turkey flocks by immune electron microscopy and genome electropherotyping methods in the USA [17]. RVD was first identified in chicken faeces [18–20], and later identified in Europe and Asia [21]. RVF was identified in turkey faeces [17] and was detected in broiler chicks with running and stunting syndrome [22]. RVG was first identified in a chicken from Northern Ireland [18] and later detected in turkeys [23]. Subsequently, RVG has been identified in Brazil, Germany, Italy, the Netherlands and South Africa [24–26]. RVG was found in broiler chicks with running and stunting syndrome [22], but its role in the diseases remains unclear.

For RVA, a sequence-based classification system was proposed in 2008 [27]. Subsequently, the Rotavirus Classification Working Group (RCWG) was involved in RVA...
classification [28]. To understand evolution and reassembly between host species, hundreds of complete genome sequences of different RV species have been reported, but turkey RVG sequences are lacking. Since RVG has not been isolated in cell culture, molecular epidemiology is the main approach to characterizing RVG strains to understand the genetic diversity and evolution compared to RVA. In the present study, we identified and characterized the genome sequences of two US turkey RVG strains identified from a hatchery with no reported signs of disease (RVG/turkey-wt/Minnesota-1/USA/2016) and a hatchery with high embryo/poul mortality (RVG/turkey-wt/Minnesota-2/USA/2016).

The intestinal caecum samples from 3-day-old poults from two Minnesota hatcheries (MN-1=hatchery with no disease and MN-2=hatchery with high embryo/poul mortality and lesions of white, caseous exudate within dilated caeca) were submitted to the Veterinary Diagnostic Laboratory of the University of Minnesota. The caeca of two poults from each hatchery were pooled and tested for routine bacterial culture and conventional multiplex PCR for turkey enteric viral pathogens, including RVA, astrovirus type 2 and reovirus. Both pools were negative for RVA, astrovirus and reovirus by RT-PCR. The sample from MN-1 was negative on bacterial culture while MN-2 was positive for Salmonella. The two pools were individually analysed by Next Generation

Fig. 1. Similarity plot of different segments of avian RVG strains and human RVB strain. (a) Similarity plot of the complete VP6 gene of RVG strains and human RVB strain: RVB/Human-wt/BAN/Bang117/2002. (b) Similarity plot of the complete NSP3 gene of RVG strains and human RVB strain: RVB/Human-wt/JAP/MMR-B1/2007. The study strain, RVG/turkey-wt/USA/Minnesota-1/2016/VP6, was set as query strain in both analyses. The vertical and horizontal axes indicate the nucleotide sequence identities and nucleotide position (bp) in the graph, respectively.
Sequencing (NGS) on Illumina MiSeq (2×250 bases). Rotavirus G was the only viral pathogen identified in the samples, using a metagenomics pipeline [29]. The pool from MN-1 contained 281,073 (12.6%); 281,073/2,233,562) RVG reads while the pool from MN-2 contained 10,393 (0.3%); 10,393/3,080,286) RVG reads. The RVG reads were de novo assembled using bowtie2 V2.2.4 [30], resulting in the 11 segments of RVG. The genome sequences of RVG/turkey-wt/Minnesota-1/USA/2016 and RVG/turkey-wt/Minnesota-2/USA/2016 were deposited in GenBank under the accession nos KY689676–KY689686 and MF120214–MF120224, respectively.

Blast was performed on the RVG gene segments, which indicated high nucleotide identity to human RVB VP6 and NSP3 gene segments. Thus, RDP4 [31] was used to investigate potential recombination events between turkey RVG strains and human RVB strains. The sub-programmes GEN-ECONV, MaxChi and Chimaera in RDP4 detected potential recombination in the VP6 gene segment, where the turkey RVG strains harboured a similar region in the 110–230 nucleotides to the human RVB/Human-wt/BAN/Bang117/2002 (GenBank accession no. GU391305.1) Fig. 1(a), which was visualized using a similarity plot with a two-parameter (Kimura) distance model and a sliding window of 60 base pairs and step size of 10 base pairs. With RDP4, GEN-ECONV, BootScan, SiScan and TOPAL sub-programmes detected potential recombination between the turkey RVG strains, which harboured a similar region in the 130–270 nucleotides to the human RVB/Human-wt/JAP/MMR-B1/2007 (GenBank accession no. GU370059.1) Fig. 1(b). While recombination within a RV species is rare, recombination with different RV species is unlikely to occur [32]. This identification of potential recombination could be an artefact related to protein function, which is necessary for viral persistence. Blastp analysis of the 11 RVG proteins indicated low amino acid identity (<60%) to RVB proteins, supporting an evolutionary or biological relationship between RVG and RVB.

The RVG gene segments were aligned with publicly available RVG sequences in GenBank, and nucleotide and amino acid sequences identities were calculated with Lasergene package MegAlign software v7.1.0 (DNASTAR, Inc., Madison, WI, USA). The RVG VP6 sequences from MN-1 and MN-2 shared 66.2–79.1% of their nucleotide identities with the reported three RVG strains, confirming the identification of turkey RVG [12]. The length of the open reading frame (ORF) along with nucleotide and amino acid identities of the MN-1 gene segments were compared to the corresponding frames of RVG/turkey-wt/Minnesota-2/USA/2016, chicken and pigeon RVG strains (Table 1). When comparing MN-1 and MN-2, the RVG ORFs of VP7, VP4, VP6, VP1-VP3, NSP1, NSP2, NSP4 and NSP5 were the same length, except for the NSP3 protein, which was 63 aa longer (Table 1).

The turkey RVG strains from the healthy and infected hatcheries (MN-1 and MN-2, respectively) shared high amino acid sequences identities in the remaining gene segments; 85.3% in VP7, 97.9% in VP4, 100% in VP6, 99.8% in VP1, 98.7% in VP2, 99.1% in VP3, 99.7% in NSP1, 99.0% in NSP2, 100% in NSP5, 100% in NSP4 and 100% in NSP5 (Table 1). Since RVG was identified in both hatcheries, RVG could be a non-pathogenic virus in the turkey, and the clinical disease in MN-2 could be attributed to Salmonella infection. Conversely, RVG might be associated with clinical diseases due to strain variability in the VP7, since MN-1 and MN-2 were only 85% identical. These hypotheses would have to be verified by further clinical experiments.

The nucleotide sequence identities of the VP7 and VP4 genes of the turkey strains to other RVG strains were 53.9–61.6 and 48.0–48.6%, respectively, suggesting a novel G and H gene on the genome segments.

<table>
<thead>
<tr>
<th>Table 1. Nucleotide and amino acid sequence identities (%) and lengths of coding regions of RVG/turkey-wt/USA/Minnesota-1/2016 compared to the four other RVG strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of the nucleotide and/or amino acid coding region is in parentheses.</td>
</tr>
<tr>
<td>Nucleotide</td>
</tr>
<tr>
<td>VP7 (250, 249)</td>
</tr>
<tr>
<td>VP4 (2367, 788)</td>
</tr>
<tr>
<td>VP6 (1176, 391)</td>
</tr>
<tr>
<td>VP1 (3483, 1160)</td>
</tr>
<tr>
<td>VP2 (2967, 988)</td>
</tr>
<tr>
<td>VP3 (2307, 768)</td>
</tr>
<tr>
<td>VP1 (975, 324)</td>
</tr>
<tr>
<td>NSP2 (903, 300)</td>
</tr>
<tr>
<td>NSP3 (1098, 365)</td>
</tr>
<tr>
<td>NSP4 (564, 187)</td>
</tr>
<tr>
<td>NSP5 (546, 181)</td>
</tr>
</tbody>
</table>
P genotype in turkeys. However, at present it is difficult to classify RVG strains due to the limited number of complete RVG genome sequences. The turkey VP4 gene segment had relatively low (~37%) amino acid identity to the other RVG strains (Table 1). Since VP4 is used for viral attachment, the low amino acid identity could be related to host restriction factors. Within the NSP1 gene, the pigeon strain HK18/2011 shared the lowest nucleotide identity (45.3%) with the study strain MN-1 while the nucleotide identity was higher with the chicken strains (~80.0%) (Table 1). The NSP3 protein can simultaneously interact with the eukaryotic translation initiation factor eIF4G and the 3' end of the mRNA to regulate cellular protein synthesis in RVA infections [10, 33], and the NSP3 of the turkey RVG strains was 192 nucleotides longer than that of the other RVG strains. The biological implications of the greater length in the NSP3 protein of turkey RVG strains are unknown. The turkey NSP4 gene segment shared a 31.6% nucleotide identity with that of the chicken strain Ger/03V2567 but had higher identity with the other RVG gene segments. In addition, the turkey RVG strains shared a 40.9% nucleotide identity with Ger/03V2567 in NSP5, but also had a low nucleotide identity (59.3%) with the pigeon strain (Table 1).

Phylogenetic analyses were implemented to evaluate the genetic relationship of our RVG strains with heterologous RVG chicken and pigeon using the maximum likelihood algorithm, with the GTR nucleotide substitution model (bootstrap analysis with 1000 replicates) (Fig. 2). In the VP6, VP1-VP3 and NSP1 phylogenetic trees, MN-1 and MN-2 share a common ancestor with the chicken RVG strains (Fig. 2c, d–f and g, respectively). In the VP7 and NSP5 phylogenetic trees, MN-1 and MN-2 share a common ancestor with the chicken RVG strains (Fig. 2g and h, respectively). Scale bars indicate nucleotide substitutions per site.

**Fig. 2.** Maximum likelihood phylogenetic trees of different segments of the study strains with the cognate gene of the other RVG strains. Phylogenetic trees based on VP7 (a), VP4 (b), VP6 (c), VP1 (d), VP2 (e), VP3 (f), NSP1 (g), NSP2 (h), NSP3 (i), NSP4 (j) and NSP5 (k). Bootstrap values are represented at major nodes. Black solid circles and solid squares indicate the different segments of our RVG strains: RVG/turkey-wt/USA/Minnesota-1/2016 and RVG/turkey-wt/USA/Minnesota-2/2016, respectively. Scale bars indicate nucleotide substitutions per site.
ancestor with the chicken RVG strain MRC-DPRU1679 (Fig. 2a and k, respectively). In the NSP3 phylogenetic tree, MN-1 and MN-2 have a closer genetic relationship with the pigeon RVG strain CHN/HK18/2011 than the chicken strains (Fig. 2i). The RVG phylogenetic trees indicate a wide gap in evolutionary information based on the long branch lengths in the different gene segments.

Previously identified rotavirus motifs were present in both US turkey RVG strains. In the VP1 protein, the three conserved RNA-dependent RNA polymerase motifs (35) AEKII LYT[4D/S][4W][4D]573, 630 LI[4R]YHG[4V][4S][4G][4K][4T][4K][4G][4N][4S][4F] NVALI644 (different residues are highlighted by underlining) and 683 MRVDGD[4D][4D][4V]VT693 were identified, with the second motif representing a potential metal-binding motif [34]. Another conserved motif, 456 ALYSLSN462, was observed in the VP3 protein but its function remains unknown [35]. The motif 744 KX[D/N]G747 in the VP3 protein is assumed to be associated with guanylyltransferase activity [36]. In the RNA-binding domain of NSP2 protein, the conserved amino acid sequence 229 HGXGHRXV237 and histidine triad 229 H-X-His-X-His-XX235 were observed; these facilitate the binding of nucleoside triphosphate [37].

In contrast to the avian RVA and RVD strains, knowledge of RVF and RVG is relatively scarce. RVG was first identified in broiler chicks with running and stunning syndrome [22] while its role in these diseases remains to be clarified [25]. RVG has consistently been reported from South Africa, Germany and Italy [24, 26] in association with diarrhea, growth depression and enteric diseases. While data on RVG are accumulating, the diversity of avian RVG, its infection status in avian flocks and its role in avian disease require further investigation. RV strains were previously reported in 2–3-day-old poults [17, 38]. In a retrospective study on pathogen detection in cases of poult enteritis syndrome, the study concluded that Salmonella and enteric viruses, specifically RVs, occupy a niche in the intestines of 10-day-old poults and may lead to poult enteritis syndrome at a very young age [38]. However, the RV species observed in these studies are unknown.

In conclusion, complete turkey RVG genomes of two strains were generated. Although RVG strains were identified in both a hatchery with no known health issues (MN-1) and one with high embryo/poult mortality (MN-2), the strains shared only 85.3 % nucleotide sequence identity and 88.4 % amino acid sequence identity in the VP7 gene segment. The turkey RVG strains shared some conserved motifs with other RVG strains. However, the RVG nucleotide and amino acid sequence identities ranged between 31.6–87.3 and 11.0–83.0 %, respectively, indicating high genetic diversity within the RVG species. The sections of high nucleotide identity noted in the VP6 and NSP3 gene segments suggest potential recombination events between turkey RVG and human RVB strains. The relationship between RVG and clinical disease cannot be confirmed at this time; clinical experiments are needed to characterize these newly identified turkey RVG strains. The diversity and classification of avian RVGs, the role of RVG in the avian disease and the inter-relation of human RVB and turkey RVG should be further investigated.

**Funding information**
This study was supported by the University of Minnesota Veterinary Diagnostic Laboratory. Fangzhou Chen was supported by the China Agriculture Research System (No. CARS-36) and China Scholarship Council No. 201606760012.

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

**Reference**
17. Theil KW, Reynolds DL, Saif YM. Comparison of immune electron microscopy and genome electrophorotyping techniques for...

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.