Identification of an avian group A rotavirus containing a novel VP4 gene with a close relationship to those of mammalian rotaviruses

Eva Trojnar,1,2 Jana Sachsenröder,1,2 Jochen Reetz,1 Peter H. Otto4 and Reimar Johne1

1Federal Institute for Risk Assessment, Max-Dohrn Straße 8-10, 10589 Berlin, Germany
2Free University Berlin, Takustraße 3, 14195 Berlin, Germany
3Institute for Molecular Biology and Bioinformatic, Charite, Arnimallee 22, 14195 Berlin, Germany
4Friedrich Loeffler Institute, Institute for Bacterial Infections and Zoonoses, Naumburger Straße 96a, 07743 Jena, Germany

INTRODUCTION

Group A rotaviruses (RVAs) are an important cause of diarrhoeal illness in humans, as well as in mammalian and avian animal species. Previous sequence analyses indicated that avian RVAs are related only distantly to mammalian RVAs. Here, the complete genomes of RVA strain 03V0002E10 from turkey (Meleagris gallopavo) and RVA strain 10V0112H5 from pheasant (Phasianus colchicus) were analysed using a combination of 454 deep sequencing and Sanger sequencing technologies. An adenine-rich insertion similar to that found in the chicken RVA strain 02V0002G3, but considerably shorter, was found in the 3' NCR of the NSP1 gene of the pheasant strain. Most genome segments of both strains were related closely to those of avian RVAs. The novel genotype N10 was assigned to the NSP2 gene of the pheasant RVA, which is related most closely to genotype N6 found in avian RVAs. However, this virus contains a VP4 gene of the novel genotype P37, which is related most closely to RVAs from pigs, dogs and humans. This strain either may represent an avian/mammalian rotavirus reassortant, or it carries an unusual avian rotavirus VP4 gene, thereby broadening the potential genetic and antigenic variability among RVAs.

Rotaviruses are classified as a genus within the family Reoviridae (Attoui et al., 2012). They are non-enveloped particles containing a genome of 11 segments of dsRNA with monocistronic coding capacity except for segment 11, which may encode two proteins. Based on antigenic and genome sequence properties, five rotavirus groups (A–E) and two tentative groups (F, G) can be distinguished, which also represent the taxonomically defined rotavirus species and tentative species, respectively (Attoui et al., 2012). In addition, the rotavirus NADRV has been described in adults in Asia (Yang et al., 2004). Recently, a classification system into the eight rotavirus species A–H (with NADRV designated rotavirus H) has been proposed, based on genetic data of genome segment 6 (Matthijnssens et al., 2012).

Among the rotavirus groups, RVAs have the highest clinical importance among humans and most mammalian species. The antigenic structures of RVAs eliciting neutralizing antibodies are the outer capsid proteins VP7 and VP4, which define the G- and P-types, respectively. Originally, G- and P-serotypes were defined based on antibody reactivity (Hoshino & Kapikian, 2000). Later on, sequence data of the VP7- and VP4-encoding genome segments were used for definition of G- and P-genotypes, leading to the present list of at least 27 different G-types and 35 different P-types in human and animal RVAs (Matthijnssens et al., 2011). Recently, a genotyping system involving all 11 genome segments has been developed (Matthijnssens et al., 2008).
The efficacy of the above-mentioned two human vaccines against the whole spectrum of RVA types is not yet known; however, their use has generally resulted in dramatic reductions in rotavirus disease (Usonis et al., 2012).

RVAs are distributed widely among animals, and diarrhoeic disease is caused mainly in young animals. Evidence of zoonotic transmission of rotavirus strains between animals (e.g. rabbits, dogs and cats) and humans arose from several epidemiological studies and genetic analyses of RVA strains (Martella et al., 2010). In addition, whole-genome analysis of RVA strains identified numerous reassortants, which contain an assembly of genome segments from human and mammalian animal RVA strains (Martella et al., 2010). It has been speculated that some of the animal strains transmitted to humans can adapt to humans and thereafter spread within the human population, as suggested for G9 and G12 strains (Matthijnssens et al., 2010).

In contrast to mammalian rotaviruses, the avian RVAs have been underinvestigated so far. However, they are distributed globally among chickens and turkeys and they are suspected to be involved in diarrhoea as well as in chronic diseases of poultry (Otto et al., 2012). Analysis of genome sequences of avian RVAs indicates that they are related only distantly to mammalian RVAs. Interspecies transmission and reassortment of avian RVAs have been described among avian hosts (Schumann et al., 2009). However, only two complete genome sequences of avian RVA strains, from chicken and pigeon, are available (Ito et al., 2001; Trojnar et al., 2009).

To obtain more information about the genetic variability of avian RVAs, the complete genomes of two other RVA strains isolated from turkey and pheasant have been sequenced here and compared with those of known RVAs.

RESULTS

Genome sequence analysis

RVA strains 03V0002E10 and 10V0112H5 were isolated from diarrhoeic turkey and pheasant, respectively, and subjected to genome sequencing. Using the 454 FLX sequencing method, about 97.9% of the turkey RVA genome and 95.9% of the pheasant RVA genome could be sequenced with two-to 153-fold sequence coverage. Missing sequences were exclusively located at the termini of the segments and in an adenine-rich sequence of the NSP1 gene (see below), and were subsequently amplified by classical RT-PCR (13 fragments for the turkey strain and 10 fragments for the pheasant strain). In parallel, the complete VP4 gene of the pheasant strain was amplified by RT-PCR and analysed by classical Sanger sequencing. However, no differences were evident when this sequence was compared with the corresponding fragment derived from 454 FLX sequencing, with the exception of 68 nt missing at the 5’ terminus in the 454 FLX sequence.

The genome size of the turkey strain 03V0002E10 is 18 863 nt; that of the pheasant strain 10V0112H5 is 18 938 nt. For both viruses, the RVA genes encoding the structural proteins VP1–4, VP6–7 and non-structural proteins NSP1–NSP5/6 could be identified. The 3’-terminal NCR of the NSP1 genes had a remarkably different length. As shown in Fig. 1, the NSP1 gene of the pheasant RVA contains a 124 nt long adenine-rich insertion in this region, which resembles that found in chicken RVA strain 02V0002G3 (Trojnar et al., 2009), but is considerably shorter. The turkey RVA, as well as the pigeon strain PO-13, do not show a corresponding insertion. Additionally, the VP6 gene of the turkey strain has a short 17 nt tandem duplication of nt 1288–1304, which is located in the 3’ NCR.

Sequence comparison to other rotavirus strains, genotyping and phylogenetic analysis

Comparison of the new sequences revealed, for most of the genome segments, the highest percentages of nucleotide sequence identities to those of avian RVAs (Table 1). However, for the VP4-encoding genome segment of the pheasant RVA strain 10V0112H5, the highest percentages of sequence identity were found to cogent genes of RVAs from dogs, pigs and humans. Accordingly, the first 100 hits
Table 1. Range of nucleotide sequence identities of the genome segments of turkey RVA 03V0002E10 and pheasant RVA 10V0112H5 to established RVA reference genotypes

<table>
<thead>
<tr>
<th>Gene-encoding segment</th>
<th>03V0002E10 (turkey)</th>
<th>10V0112H5 (pheasant)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Identity (%)</td>
<td>Highest identity to genotype (host)</td>
</tr>
<tr>
<td>VP1</td>
<td>68.0–89.8</td>
<td>R4 (pigeon)</td>
</tr>
<tr>
<td>VP2</td>
<td>67.2–85.9</td>
<td>C4 (pigeon)</td>
</tr>
<tr>
<td>VP3</td>
<td>59.3–90.3</td>
<td>M4 (pigeon)</td>
</tr>
<tr>
<td>VP4</td>
<td>60.0–99.7</td>
<td>P[35] (turkey)</td>
</tr>
<tr>
<td>VP6</td>
<td>64.9–93.6</td>
<td>I4 (pigeon)</td>
</tr>
<tr>
<td>VP7</td>
<td>61.1–99.6</td>
<td>G22 (turkey)</td>
</tr>
<tr>
<td>NSP1</td>
<td>33.1–87.3</td>
<td>A16 (chicken)</td>
</tr>
<tr>
<td>NSP2</td>
<td>61.2–93.9</td>
<td>N4 (pigeon)</td>
</tr>
<tr>
<td>NSP3</td>
<td>53.0–94.5</td>
<td>T4 (pigeon)</td>
</tr>
<tr>
<td>NSP4</td>
<td>44.4–97.4</td>
<td>E11 (turkey)</td>
</tr>
<tr>
<td>NSP5</td>
<td>59.4–93.6</td>
<td>H4 (pigeon)</td>
</tr>
</tbody>
</table>

The genotype assignment is mirrored by the branching of phylogenetic trees set up for the nucleotide sequences of each genome segment (Fig. S1, available in JGV Online). For the VP4 gene, the pheasant RVA strain branches closest to human RVA strains (Fig. 2), thus generally confirming its assignment to mammalian RVA sequences, as indicated previously by sequence identity calculations and BLASTN searches.

DISCUSSION

Our knowledge of avian rotaviruses, their genetic variability and ability for interspecies transmission is very limited. As only two avian RVA genome sequences, from chicken and pigeon, were available (Ito et al., 2001; Trojnar et al., 2009), the sequences of two additional complete genomes of avian RVA isolates from turkey and pheasant were determined. A deep sequencing 454 FLX method was applied in order to enable rapid and reliable sequencing of the 11 genome segments. Although this method enabled sequencing of >95% of the genomes, short missing sequences were identified at the termini of many segments and in an adenine-rich region. Therefore, this method is suitable for most applications of rotavirus genome.

Table 2. Genotype constellations of the avian RVA strains completely sequenced so far

<table>
<thead>
<tr>
<th></th>
<th>VP7</th>
<th>VP4</th>
<th>VP6</th>
<th>VP1</th>
<th>VP2</th>
<th>VP3</th>
<th>NSP1</th>
<th>NSP2</th>
<th>NSP3</th>
<th>NSP4</th>
<th>NSP5</th>
</tr>
</thead>
<tbody>
<tr>
<td>02V0002G3 (chicken)</td>
<td>G19</td>
<td>P[30]</td>
<td>I1</td>
<td>R6</td>
<td>C6</td>
<td>M7</td>
<td>A16</td>
<td>N6</td>
<td>T8</td>
<td>E10</td>
<td>H8</td>
</tr>
<tr>
<td>PO-13 (pigeon)</td>
<td>G18</td>
<td>P[17]</td>
<td>I4</td>
<td>R4</td>
<td>C4</td>
<td>M4</td>
<td>A4</td>
<td>N4</td>
<td>T4</td>
<td>E4</td>
<td>H4</td>
</tr>
<tr>
<td>03V0002E10 (turkey)</td>
<td>G22</td>
<td>P[35]</td>
<td>I4</td>
<td>R4</td>
<td>C4</td>
<td>M4</td>
<td>A16</td>
<td>N4</td>
<td>T4</td>
<td>E11</td>
<td>H4</td>
</tr>
<tr>
<td>10V0112H5 (pheasant)</td>
<td>G23</td>
<td>P[37]</td>
<td>I4</td>
<td>R4</td>
<td>C4</td>
<td>M4</td>
<td>A16</td>
<td>N10</td>
<td>T4</td>
<td>E4</td>
<td>H4</td>
</tr>
<tr>
<td>Dhaka6 (human)</td>
<td>G11</td>
<td>P[25]</td>
<td>I1</td>
<td>R1</td>
<td>C1</td>
<td>M1</td>
<td>A1</td>
<td>N1</td>
<td>T1</td>
<td>E1</td>
<td>H1</td>
</tr>
</tbody>
</table>
sequencing, e.g. complete genotyping. However, further method development is necessary if the complete sequence, including terminal sequences of the genome segments, is required.

Whole-genome analysis of the two avian rotaviruses revealed a close relationship with other avian RVA strains, which may indicate an adaptation of these rotaviruses to avian hosts. Among the avian RVAs, considerable genetic

---

**Fig. 2.** Phylogenetic relationships of rotaviruses 03V0002E10 and 10V0112H5 derived from turkey and pheasant, respectively, to other RVA strains, based on the VP4-encoding genome segment. The tree was constructed using nucleotide sequences of the reference sequences of established genotypes (Matthijnssens et al., 2011). RVA strain GUB71 from wild boar, which showed the highest sequence identity to the pheasant RVA strain 10V0112H5 by BLASTN search, is underlined. Rotavirus D, strain 05V0049 from chicken, was used as an outgroup sequence. The tree was constructed using the neighbour-joining method implemented in the MEGALIGN module of the DNASTAR software package (Lasergene). Bootstrap values $>50\%$ are indicated. Strains are labelled with species, genotype, host abbreviation (Hu, human; Po, porcine; Lamb, lamb; Rh, rhesus monkey; Bo, bovine; Mu, murine; Eq, equine; Si, simian; Ca, canine; Ph, pheasant; Tu, turkey; Po, pigeon; Ch, chicken); strain designation and GenBank accession no. The mammalian and avian clades are indicated. The turkey strain (filled arrow) and the pheasant strain (open arrow) are marked.
variability was evident, and different genotype constellations were determined for the four completely sequenced avian strains. As the pheasant and turkey strain show a close relationship to the pigeon strain PO-13 for seven of the 11 segments, but a close relationship to the chicken strain 02V0002G3 for the NSP1-encoding segment, reassortment events among avian RVA strains, as first described by Schumann et al. (2009), were confirmed. Another kind of sequence variation was found due to insertions into the NCRs of some of the genome segments. Insertions, deletions and rearrangements in genome segments of rotaviruses have been reported previously (Giambigi et al., 1994; Matthijnssens et al., 2006; Desselberger, 1996). A large adenine-rich insertion has also been found in the NSP1 gene of the chicken rotavirus 02V0002G3 (Trojnar et al., 2009), resembling that of the pheasant RVA except for the difference in length. The function of this insertion is not clear, although polyadenylation may enhance the translation of RNA through binding of the poly(A)-binding protein, which interacts with translation initiation factors (Deo et al., 1999).

Most interesting, the VP4-encoding genome segment of the pheasant RVA represents a novel genotype, which is related most closely to that of mammalian RVAs. Although its distinct origin cannot be traced, it may be speculated that it results from reassortment with an as-yet-unknown mammalian RVA strain. A prerequisite for creation of reassortants is a simultaneous replication of avian and mammalian rotavirus strains in a doubly infected cell in vitro or in a host. Recently, mammalian rotaviruses have been detected in chickens and turkeys (Wani et al., 2003; Asano et al., 2011) and an avian rotavirus has been identified previously in cattle (Bru¨ssow et al., 1992). Experimental infection of mice with the pigeon rotavirus PO-13 resulted in infection and diarrhoea (Mori et al., 2001). In addition, a reassortant turkey RVA strain containing a VP4 gene of a simian RVA has been previously created by experimental co-infection of cell cultures (Kool et al., 1992). Although all these studies strongly suggested the theoretical possibility of avian/mammalian rotavirus reassortment, analysis of the pheasant strain provides the first suggestive evidence for such reassortment under field conditions. However, given the fact that only very few avian RVA sequences have been determined so far and that the [P37] genotype has not yet been described in mammalian RVAs, this genotype may also represent an unusual avian VP4 genotype. More avian and mammalian RVA strains need to be analysed in future to answer this question.

VP4 is an outer capsid protein of the rotavirus particle, which is a target for rotavirus-specific antibodies. In addition, VP4 has been implicated in other biological functions, e.g. host species restriction, plaque phenotype, virulence, sialic-acid dependence and, more recently, binding to histo-blood antigens (Estes & Kapikian, 2007; Hu et al., 2012). Therefore, by exchange of the VP4-encoding genome segment or by extensive mutations within the gene, changes in the antigenicity, as well as biological properties, of the virus could be expected. If avian and mammalian RVAs can exchange genetic material by reassortment or the genetic variability among the avian genes is unusually high, the variability of RVAs may therefore be broader than expected. Further studies should focus on the analysis of more rotavirus field isolates, especially those that are difficult to type with routine methods, for the presence of avian rotavirus genes in mammalian RVAs and vice versa.

### METHODS

**Isolation of rotaviruses from turkey and pheasant.** RVA strain 03V0002E10 was isolated in 2003 from the intestinal content of a turkey (Melesaggeris gallopavo) with diarrhoea by inoculation onto MA104 cell cultures as described previously (Schumann et al., 2009). From this strain, the gene regions encoding VP4, VP6, VP7 and NSP5 were already known (Schumann et al., 2009). Using the same method, RVA strain 10V0112H5 was isolated in 2010 from a pheasant (Phasianus colchicus) with diarrhoea, which was kept in an aviary in Germany.

**Genome sequencing using the 454 FLX method.** RVAs were grown in MA104 cells and the supernatants were subjected to CsCl density ultracentrifugation to concentrate and purified the virus particles. The samples were loaded on a stepwise CsCl density gradient with density layers of 1.7, 1.5, 1.35 and 1.2 g ml−1, and centrifuged at 65 000 g for 14 h at 10 °C (Sachsenroder et al., 2012). An opalescent virus-containing double band (containing double- and triple-layered particles; Patton et al., 2000) was collected using a syringe. To eliminate free DNA present in the virus concentrate, 500 μl of the purified virus suspension was treated with 50 units DNase I (2000 U mg−1, bovine pancreas, grade II; Roche Diagnostics) for 45 min at 37 °C, followed by heat inactivation for 10 min at 65 °C. Thereafter, RNA was extracted from 200 μl using an RTP Pathogen kit (Invitrek). In order to enable sequencing of the genome-segment ends, a DNA linker was ligated as described previously (Trojnar et al., 2010; Maan et al., 2007). Briefly, 16 μl RNA was mixed with 2 μl oligonucleotide SP9 (Trojnar et al., 2010), 1 μl BSA, 3 μl T4 RNA-Ligase buffer (New England Biolabs) and 2 μl T4 RNA Ligase (10 U; New England Biolabs) in a total volume of 30 μl. After incubation overnight at 17 °C, the RNA was purified using MobiSpin S-400 columns (MobiTec). The purified RNA (75 ng per reaction as quantified by UV spectrometry) was denatured at 95 °C for 2 min and randomly primed for cDNA synthesis using a TransPlex Complete Whole Transcriptom Amplification (WTA) kit (WTA2; Sigma-Aldrich) according to the protocol recommended by the supplier. Further steps were performed essentially as described by Sachsenroder et al. (2012). Aliquots of 75 μl each were removed from the WTA reaction at different cycle numbers, and purified as well as size-selected using MobiSpin S-400 columns. A total of 1 mg DNA was applied to deep sequencing on a 1/16 plate of the GS-FLX sequencer 454 Titanium [GS Titanium SV emPCR kit (Lib-L) v2; GS Titanium PicoTiterPlate kit 70675; GS Titanium Sequencing kit XL70tt; Life Sciences, Roche] according to the manufacturer’s protocol. Primary sequence analysis was applied to raw sequence reads, which were subjected to primer/adaptor trimming using SeqMan (DNASTAR; Lasergene) and selection for a minimum length of 50 nt. All primary reads were subjected to de novo contig assembly using 454 Newbler Assembler software (http://www.my454.com/), with criteria of 90% minimum overlap identity and a minimum overlap length of 40 nt. In addition to de novo sequence assembly, the reads and contigs were aligned to reference genomes using the read-mapping software CLC Genomics Workbench (http://www.clcbio.com/index.php?id=1326) and Geneious (http://www.geneious.com).
Genome-sequence completion using RT-PCR and Sanger sequencing. Missing sequences and those showing low sequence coverage after 454 FLX sequencing were amplified by classical RT-PCR followed by DNA sequencing. Purified RNA after ligation of the linker (as described above) was reverse-transcribed using a Qiagen LongRange 2Step RT-PCR kit (Qiagen) and subjected to PCR using primer 5-15-1 (binding site on the linker; Trojnar et al., 2010) as well as gene-specific primers constructed on the basis of the determined sequences. The complete VP4-encoding segment of strain 10V0112H5 was amplified in three overlapping fragments using the primer sequences. The complete VP4-encoding segment of strain 10V0112H5 was amplified in three overlapping fragments using the primer combinations 5-15-1 and 5-15-1, respectively. All PCR products were purified from agarose gels using a QIAquick gel extraction kit (Qiagen). Both strands of the PCR products were sequenced directly in an ABI 3730 DNA Analyzer (Applied Biosystems) using the respective PCR primers.

**Sequence analysis.** Complete genome segment sequences were assembled using the SeqBuilder module of the DNAstar software package (Lasergene) and submitted to GenBank with accession numbers JX204822–JX204832 (turkey strain 03V0000E10, segments 1–11) and JX204811–JX204821 (pheasant strain10V0112H5, segments 1–11). Sequence alignments and construction of phylogenetic trees were performed using the CLUSTAL W method with the IUB residue weight table (Thompson et al., 1994) and the neighbour-joining method as implemented in the MEGALIGN module of the above-mentioned software package. Bootstrap analysis of phylogenetic trees was performed with 1000 trials and 111 random seeds. In addition to phylogenetic analysis, assignment of genotypes to the RVA sequences was performed using the online tool RotaC (Maes et al., 2009).

**ACKNOWLEDGEMENTS**

We wish to thank I. N. Clarke (University of Southampton, UK) for reviewing the English language in the manuscript. We thank AniCon Labor GmbH (W. Bielenberg, Höttinghausen, Germany) for support with 454 FLX sequencing in part of the pheasant rotavirus strain. This investigation was supported in part by the Deutsche Forschungsgemeinschaft, projects JO369/4-1 (E. T., R. J.) and SFB852 (J. S., S. T., R. J.).

**REFERENCES**


