

Distantly Related Rotaviruses in Common Shrews, Germany, 2004–2014

Appendix

Materials and Methods

Samples

Within a small mammal monitoring program, shrews were collected as bycatches during 2010–2013 at sites in 4 federal states (1). Additional shrews were collected between 2004 and 2014 in different regions of Germany (2). Collection of samples was done according to relevant legislation and by permission of the federal authorities (permits Regierungspräsidium Stuttgart 35–9185.82/0261, Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen 8.87–51.05.20.09.210, Landesamt für Landwirtschaft, Lebensmittelsicherheit und Fischerei Mecklenburg-Vorpommern 7221.3–030/09, Thüringer Landesamt für Lebensmittelsicherheit und Verbraucherschutz 22–2684–04–15–107/09) or were bycatches of the vole monitoring by local forestry institutions as part of their pest control measures. Trapping of the shrews was based on snap trapping, or the shrews were found dead in live traps.

The dissection followed a standard protocol including the collection of intestine samples. For this purpose, the frozen carcasses were thawed, intestinal contents (≈ 0.2 g) collected, transferred into fresh individual tubes and frozen again at -20°C until further analysis. In this investigation, 49 common shrews (*Sorex araneus*) were included.

Next-Generation Sequencing (NGS) Analysis of Pooled Samples

Initially, 2 pools were generated: 1 pool of 24 samples was generated by including 6 samples from each of the 4 monitoring sites. A second pool was generated by 25 samples from ongoing monitoring activities in 10 federal states of Germany. RNA was extracted from the pooled samples using the RNeasy Mini kit (QIAGEN, Hilden, Germany). Subsequently, the samples were processed as described (3). Briefly, double-stranded cDNA was generated using the cDNA-Synthesis System-Kit (Roche Life Science, Indianapolis, USA) and then transformed

into Ion Torrent compatible libraries. To this end, DNA was fragmented to a peak size of ≈ 500 bp with a M220 Focused-ultrasonicator (Covaris Inc., Brighton, UK). Then, fragment end-polishing and ligation of adapters (Ion Xpress Barcode Adapters; Life Technologies, Darmstadt, Germany) were performed with a GeneRead DNA Library L Core Kit (QIAGEN). After quality control (Agilent Bioanalyzer 2100 and High Sensitivity DNA Kit; both Agilent, Waldbronn, Germany) and quantification (KAPA Library Quantification Kit - Ion Torrent Personal Genome Machine (PGM) (Uni); Roche, Mannheim, Germany), libraries were pooled and sequenced with an Ion Torrent PGM in 400 bp run using HiQ reagents (Life Technologies, Darmstadt, Germany).

RT-PCR Screening for Specific Rotaviruses in Fecal Samples

A total of 46 individual samples of intestinal contents (3 of the initial 49 samples had been used up in other experiments) were diluted 1:5 in phosphate-buffered saline (PBS), and 100 μ l of the solution was subsequently subjected to nucleic acid extraction using the NucliSENS® easyMAG system (BioMerieux, Marcy l'Etoile, France) according to the manufacturer's instructions. Detection of rotaviruses was done by reverse transcription PCR (RT-PCR) using the One-Step RT-PCR kit (QIAGEN) with primers as described (Appendix Table 1) in a 2720 thermal cycler (Applied Biosystems). The thermal profile comprised 42°C for 30 min and 95°C for 15 min, followed by 40 cycles at 94°C for 30 s, 56°C for 30 s, and 74°C for 40 s, with a final incubation at 74°C for 5 min. PCR products were separated by electrophoresis on ethidium bromide-stained agarose gels.

NGS Analysis of Samples KS/12/0644 and KS/11/2281

A total of 100 μ l of each individual diluted fecal sample was clarified by centrifugation at 9,600 x g for 10 min. Thereafter, 70 μ l of the supernatant were mixed with 48.5 μ l nuclease-free water, 14 μ l 10x DNase buffer, 7 μ l TurboDNase (Invitrogen), and 0.5 μ l RNase A (QIAGEN). After incubation at 37°C for 1 h, RNA was extracted from the sample using the QIAamp viral RNA Mini Kit (QIAGEN) according to the manufacturer's instructions, but without adding carrier-RNA. The RNA was eluted in 60 μ l elution buffer (EB) and thereafter subjected to reverse transcription and random amplification using the WTA2 Whole Transcriptome Amplification Kit (Sigma, Deisenhofen, Germany). Briefly, 10 μ l of RNA preparation was mixed with 2.5 μ l Library Synthesis Solution and 4.1 μ l Nuclease-Free Water, heated at 95°C for 2 min, and immediately cooled to 18°C. Thereafter, 2.5 μ l Library Synthesis Buffer, 3.9 μ l

water, and 2.0 µl Library Synthesis Enzyme were added and incubated at 18°C for 10 min, 25°C for 10 min, 37°C for 30 min, 42°C for 10 min, and 72°C for 20 min, before cooling at 4°C. The following PCR was performed according to the manufacturer's instructions with 10 µl mixture from before in a total volume of 75 µl. The reaction mix was incubated at 94°C for 2 min and 17 cycles at 94°C for 30 sec and 70°C for 5 min each were performed. For NGS, the DNA libraries were generated from the PCR products with the TruSeq® Nano DNA Library Prep Kit (Illumina) and paired-end sequenced with 2 × 151 cycles using the NextSeq 500 sequencing system (Illumina).

The resulting 8,576,782 and 6,168,437 read pairs for samples KS/12/0644 and KS/11/2281 were subjected to trimming with fastp v0.14.1 with default parameters (4), yielding 6,498,311 and 4,878,954 read pairs, respectively. For a fast test for the presence of rotavirus reads with low similarity to known references, we classified rotavirus reads using RAMBO-K v1.21 (5). As a SPAdes v3.12.0 assembly (6) of the selected reads yielded only incomplete rotavirus genomes, we then assembled all trimmed reads using metaSPAdes (6). Resulting contigs were searched in the complete NCBI nr database using BLASTX v.2.6.0 with wordsize 4 (7). Contigs with hits to any rotavirus sequence were selected as putative rotavirus sequences and subjected to further manual analysis (see below).

PCR, FLAC, and RACE for Completion of Segment Sequences

First trials to complete the 5'- and 3'-ends of the genome segments were done by full-length amplification of cDNA (FLAC) as described (8). Further trials were done by rapid amplification of cDNA ends (RACE) using the 5' RACE System Kit (Invitrogen GmbH, Karlsruhe, Germany) as described (9). In both cases, specific primers delineated from NGS-derived sequences were used. In an additional set of experiments, specific RT-PCRs with these primers and those derived from an alignment of segment ends of other available rotavirus sequences were performed to complete the open reading frames for virus protein (VP) 1, VP6, and nonstructural protein 5 (NSP5). The RT-PCR was performed as described above for detection of rotaviruses; however, the cycling conditions were adapted to primer sequences and product lengths. Products with the expected lengths were purified using the QiaQuick DNA Purification Kit (QIAGEN) and subjected to dideoxy chain termination sequencing by a commercial supplier (Eurofins, Ebersberg, Germany).

Sequence Data Analysis

The NGS contigs were manually inspected using the SeqBuilder module of the DNASTAR software package (Lasergene, Madison, WI, USA) and flanking primer sequences were removed. Sequences were assembled from these NGS contigs and sequences from PCR, FLAC, or RACE products using the same software. The sequences were submitted to the GenBank database and accession numbers MN307962–MN207992 were assigned (also indicated on the branches of the phylogenetic trees). Sequences of the rotavirus type species and the rotavirus A genotype reference strains were used for alignments and phylogenetic comparisons. Nucleotide sequences were used for genotyping and assessment of phylogenetic relationship to other rotavirus A strains in the case of shrew rotavirus A sequences. As only very low nucleotide sequence similarities were evident for the species C-like and H-like rotaviruses of shrews compared with the reference strains, the deduced amino sequences were used, which showed a higher degree of sequence conservation. The phylogenetic trees were constructed using a neighbor-joining method implemented in the MEGALIGN module of the DNASTAR software package (Lasergene), and bootstrap analysis with 1000 trials and 111 random seeds was performed. Genotyping of rotavirus A genome segments was done as described (10). According to this protocol, partial sequences can only be assigned to an existing genotype if they have a minimum length of 500 nt and cover more than 50% of the open reading frame. In case they are only distantly related to the established genotypes, the complete sequence of the open reading frame has to be submitted to the Rotavirus Classification Working Group for genotype assignment, which was done for the VP1-, VP6- and NSP5-encoding genome segments of the shrew rotavirus A.

Appendix References

1. Fischer S, Mayer-Scholl A, Imholt C, Spierling NG, Heuser E, Schmidt S, et al. *Leptospira* genospecies and sequence type prevalence in small mammal populations in Germany. *Vector Borne Zoonotic Dis.* 2018;18:188–99. [PubMed https://doi.org/10.1089/vbz.2017.2140](https://doi.org/10.1089/vbz.2017.2140)
2. Mayer-Scholl A, Hammerl JA, Schmidt S, Ulrich RG, Pfeffer M, Woll D, et al. *Leptospira* spp. in rodents and shrews in Germany. *Int J Environ Res Public Health.* 2014;11:7562–74. [PubMed https://doi.org/10.3390/ijerph110807562](https://doi.org/10.3390/ijerph110807562)

3. Wylezich C, Papa A, Beer M, Höper D. A versatile sample processing workflow for metagenomic pathogen detection. *Sci Rep.* 2018;8:13108. [PubMed https://doi.org/10.1038/s41598-018-31496-1](https://doi.org/10.1038/s41598-018-31496-1)
4. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34:i884–90. [PubMed https://doi.org/10.1093/bioinformatics/bty560](https://doi.org/10.1093/bioinformatics/bty560)
5. Tausch SH, Renard BY, Nitsche A, Dabrowski PW. RAMBO-K: rapid and sensitive removal of background sequences from next generation sequencing data. *PLoS One.* 2015;10:e0137896. [PubMed https://doi.org/10.1371/journal.pone.0137896](https://doi.org/10.1371/journal.pone.0137896)
6. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 2017;27:824–34. [PubMed https://doi.org/10.1101/gr.213959.116](https://doi.org/10.1101/gr.213959.116)
7. Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA. Database indexing for production MegaBLAST searches. *Bioinformatics.* 2008;24:1757–64. [PubMed https://doi.org/10.1093/bioinformatics/btn322](https://doi.org/10.1093/bioinformatics/btn322)
8. Kindler E, Trojnar E, Heckel G, Otto PH, Johne R. Analysis of rotavirus species diversity and evolution including the newly determined full-length genome sequences of rotavirus F and G. *Infect Genet Evol.* 2013;14:58–67. [PubMed https://doi.org/10.1016/j.meegid.2012.11.015](https://doi.org/10.1016/j.meegid.2012.11.015)
9. Schielke A, Sachs K, Lierz M, Appel B, Jansen A, Johne R. Detection of hepatitis E virus in wild boars of rural and urban regions in Germany and whole genome characterization of an endemic strain. *Virology.* 2009;6:58. [PubMed https://doi.org/10.1186/1743-422X-6-58](https://doi.org/10.1186/1743-422X-6-58)
10. Matthijnssens J, Ciarlet M, Rahman M, Attoui H, Bányai K, Estes MK, et al. Recommendations for the classification of group A rotaviruses using all 11 genomic RNA segments. *Arch Virol.* 2008;153:1621–9. [PubMed https://doi.org/10.1007/s00705-008-0155-1](https://doi.org/10.1007/s00705-008-0155-1)

Appendix Table 1. Primers used for detection of RVA, RVC-like and RVH-like viruses in shrews*

Virus	Gene	Designation	Sequence (5'-3')	Product length, bp
RVA	VP1	S-RVA-s	CGTTTCACGTAGGATTCAGGAA	225
		S-RVA-as	CTCCATTTGACGCTGATGACAT	
RVC-like	VP1	S-RVC-like-s	AACTGTGTCAAATGACGTCACA	218
		S-RVC-like-as	GAAGTGTTCGATCGCACGATT	
RVH-like	VP3	RVH-like-s	CCAGACAACCTACTCATTCTGTTG	207
		RVH-like-as	TTGTGTGTGGTTCCTTTACCT	

*RVA, rotavirus A; RVC, rotavirus C; RVH, rotavirus H; VP, virus protein; bp, base pairs.

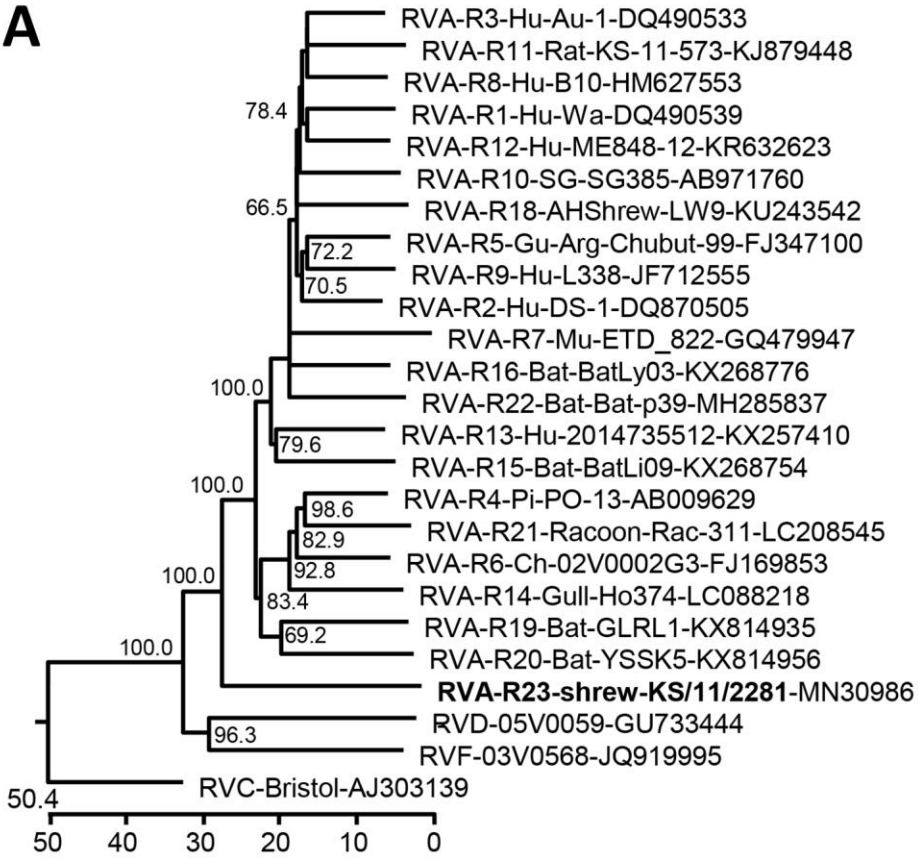
Appendix Table 2. Contig and read numbers and lengths generated by NGS of samples KS/12/0644 and KS/11/2281*

Category	KS/12/0644			KS/11/2281		
	RVA	RVC-like	RVH-like	RVA	RVC-like	RVH-like
Contig number	23	7	18	25	10	5
Contig lengths	200–1135	193–667	374–2584	164–821	636–3017	264–506
Read coverage	1.6–11.9	1.1–9.3	3.8–22.4	1.1–3.8	4.6–19.6	1.3–1.9
Total rotavirus read number		3597			2389	
Total read number†		8,576,782			6,168,437	

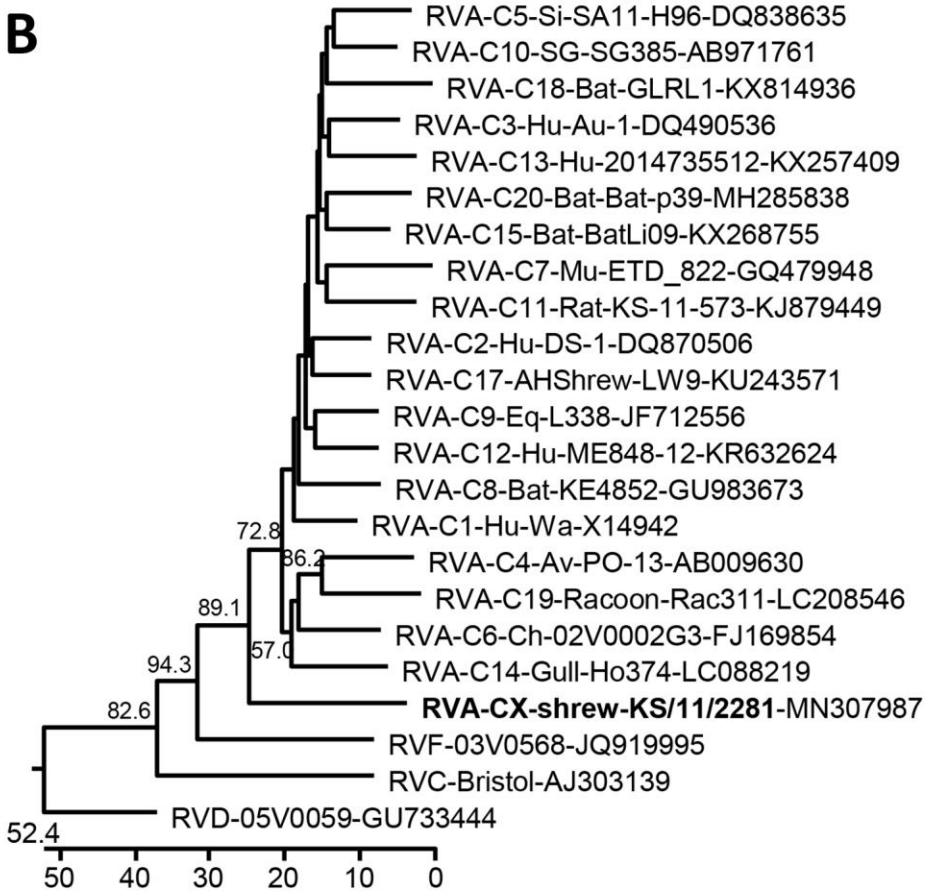
*NGS, next-generation sequencing; RVA, rotavirus A; RVC, rotavirus C; RVH, rotavirus H.

†number of all generated reads including rotavirus reads and non-rotavirus reads from the sample.

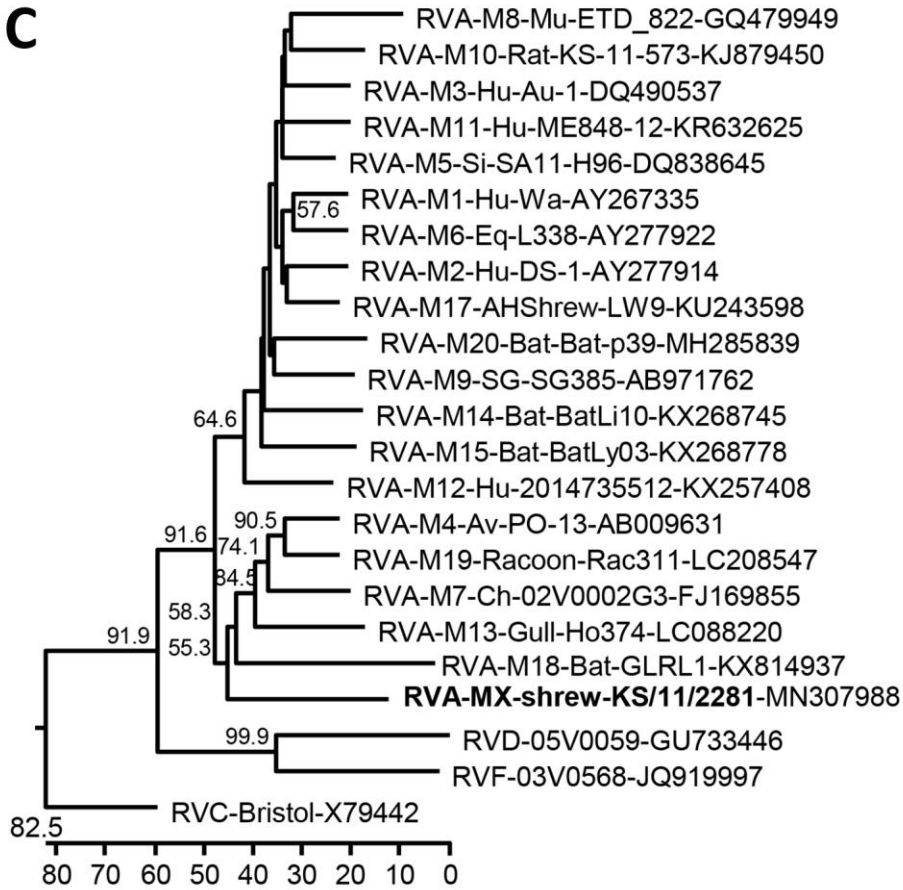
A

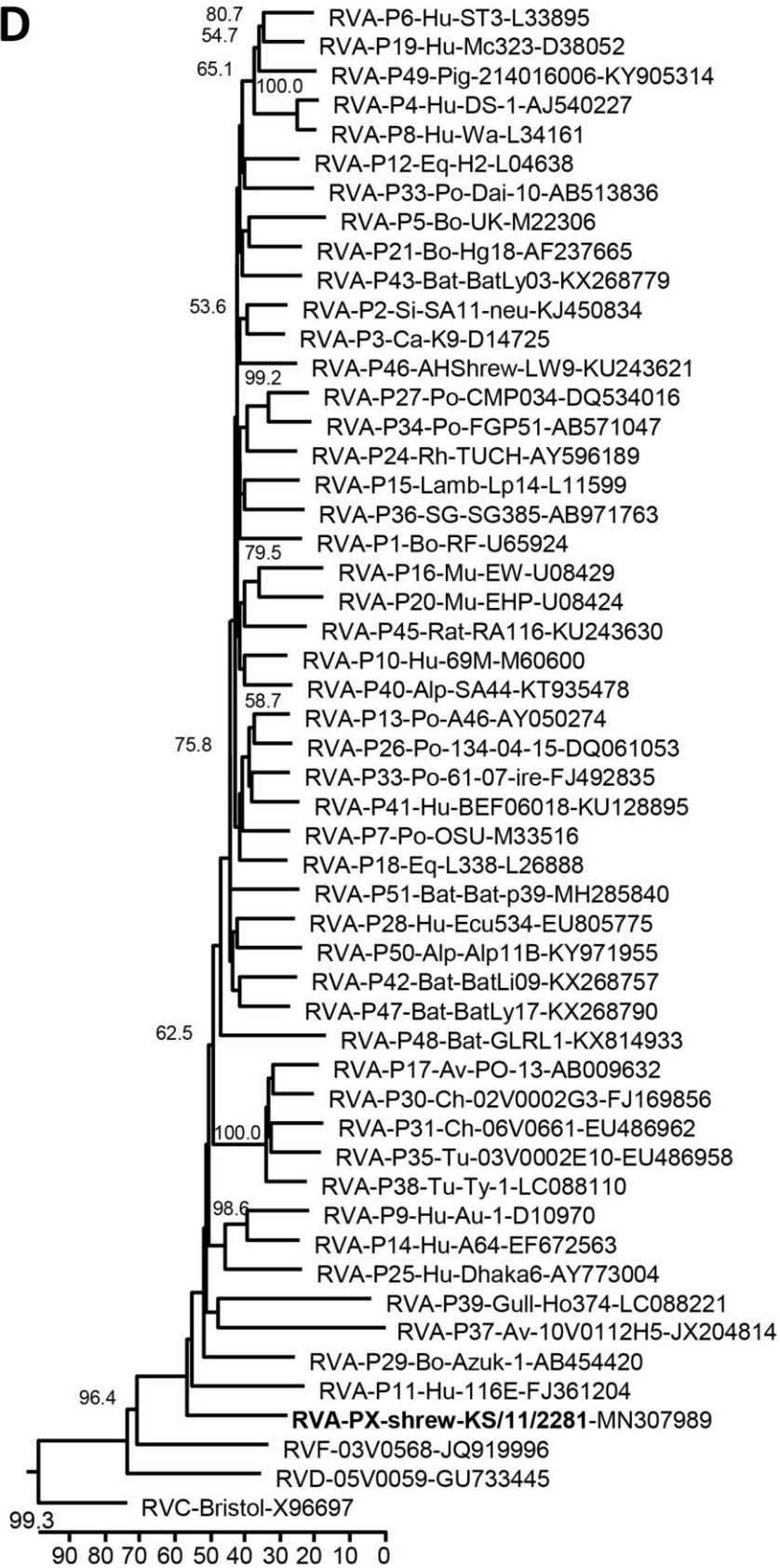


B

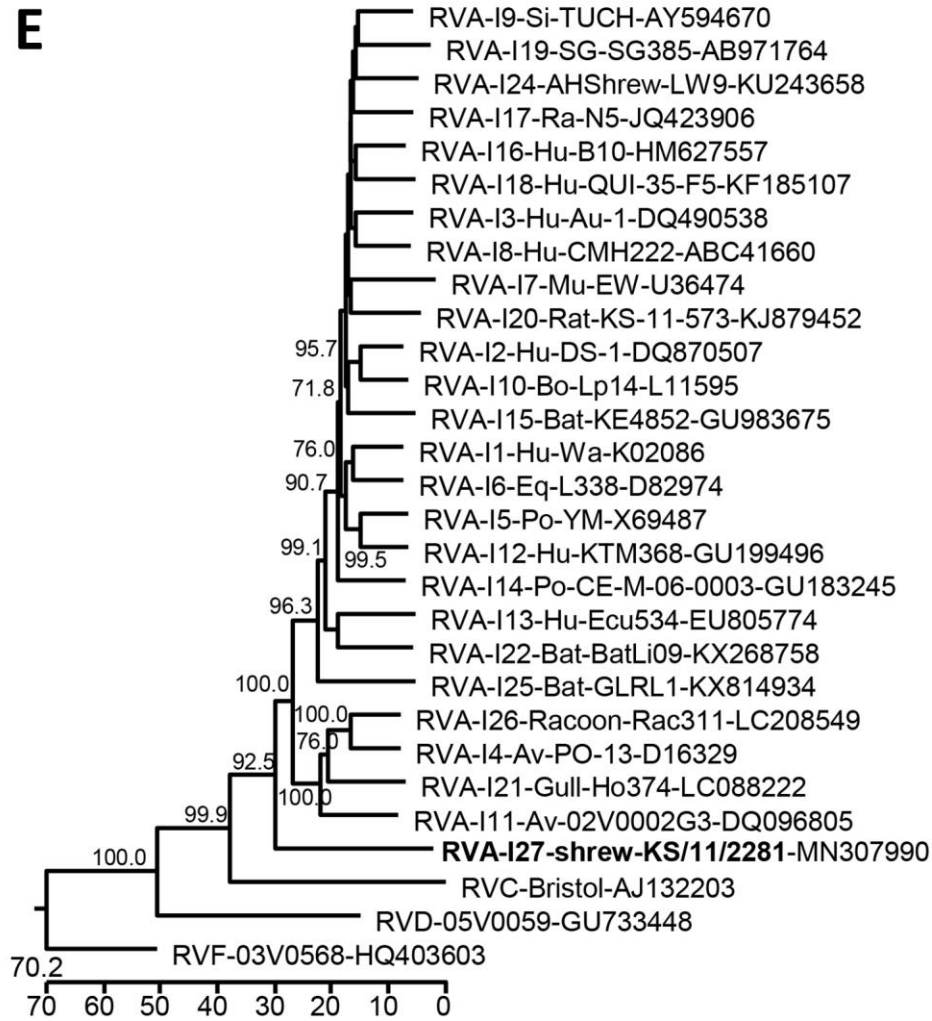


C

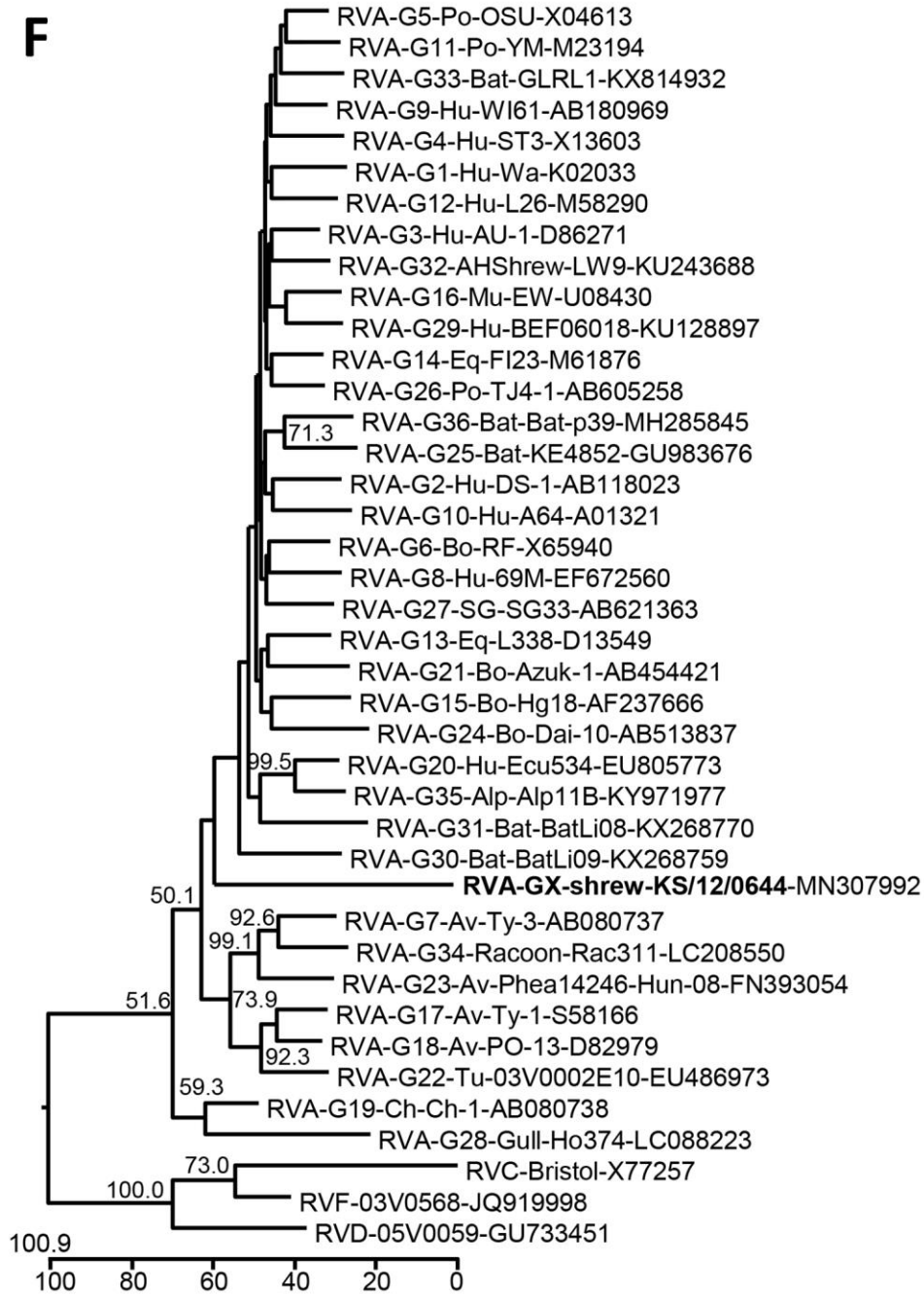


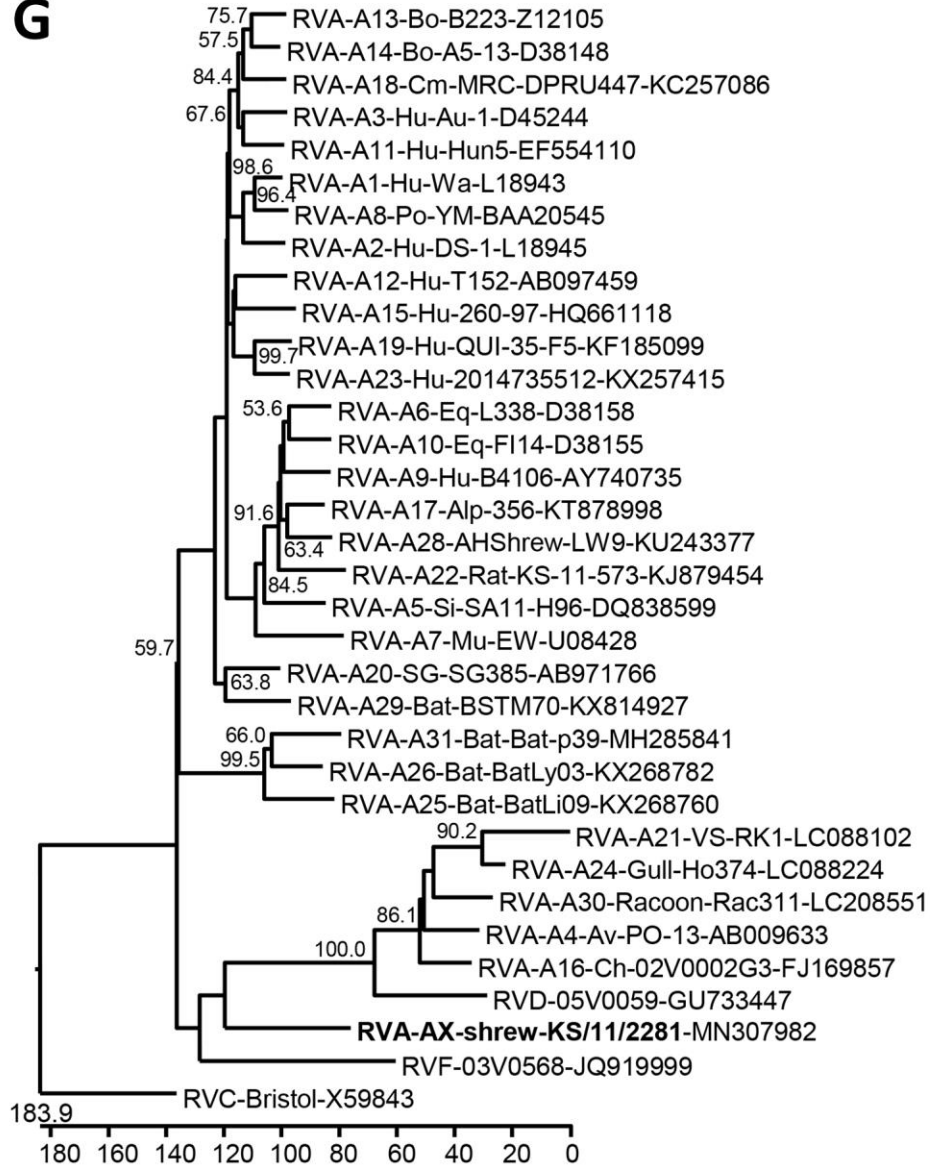
D

E

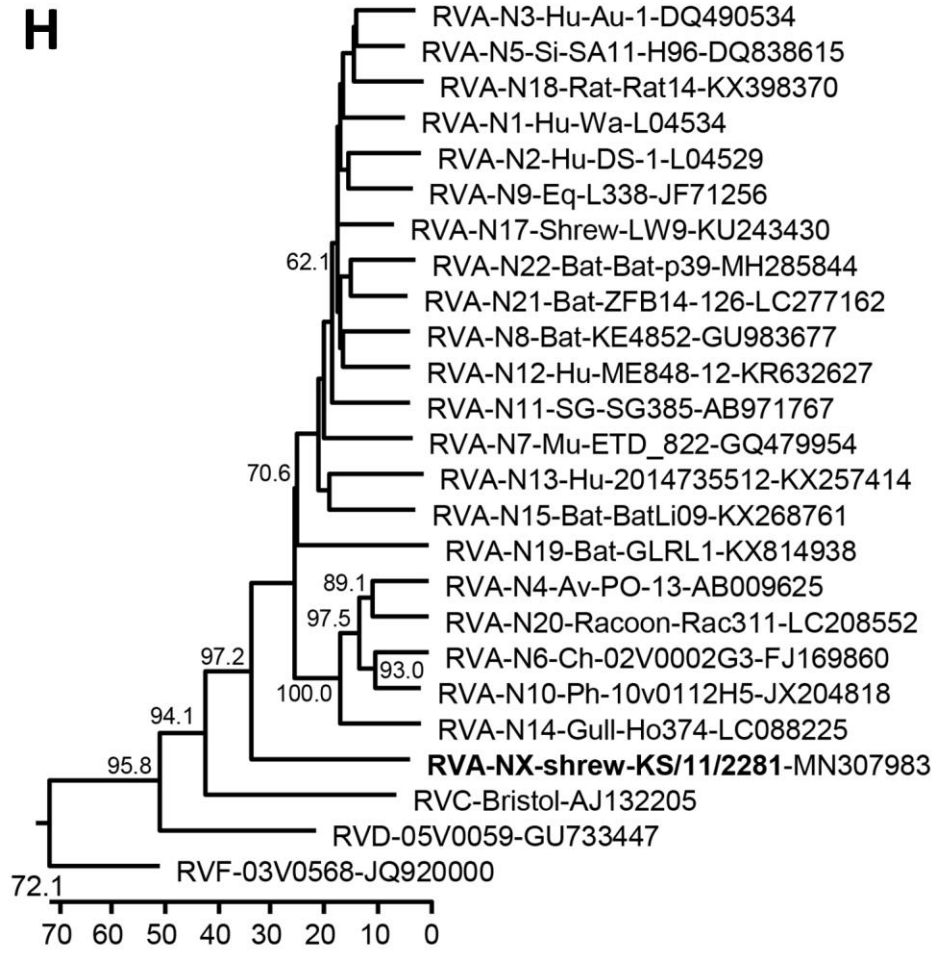


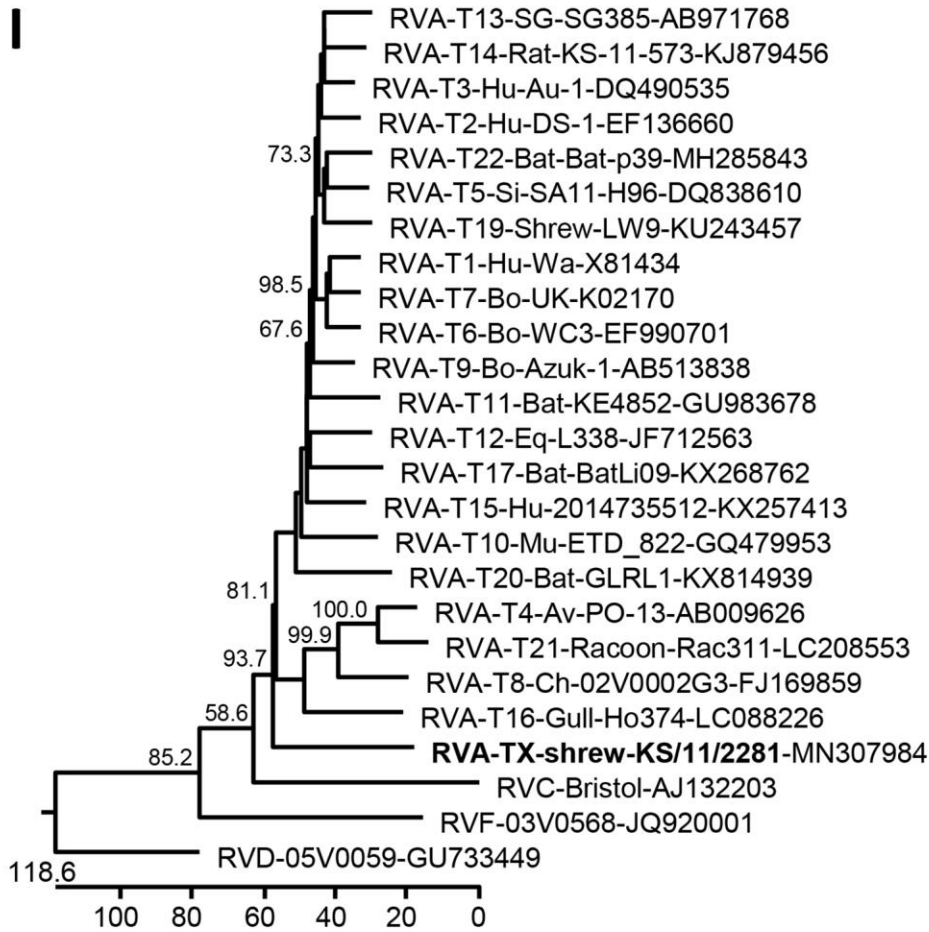
F



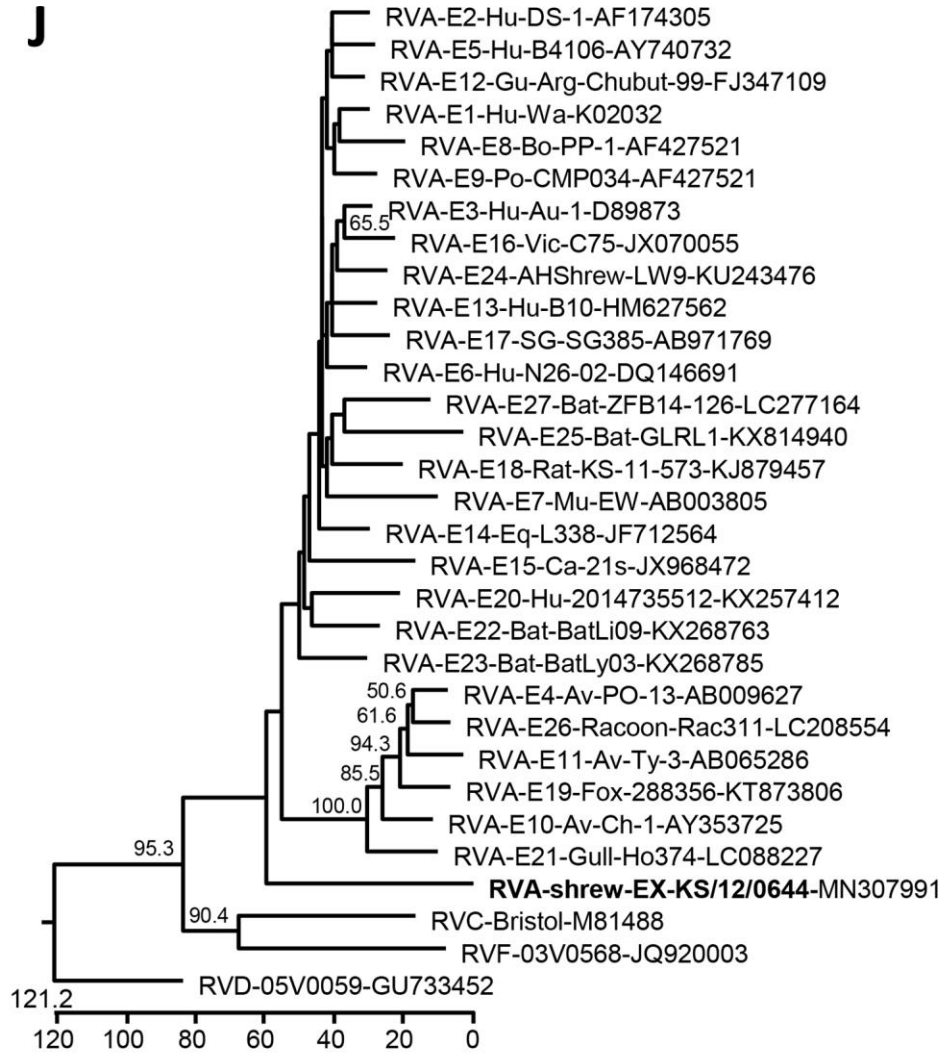
G

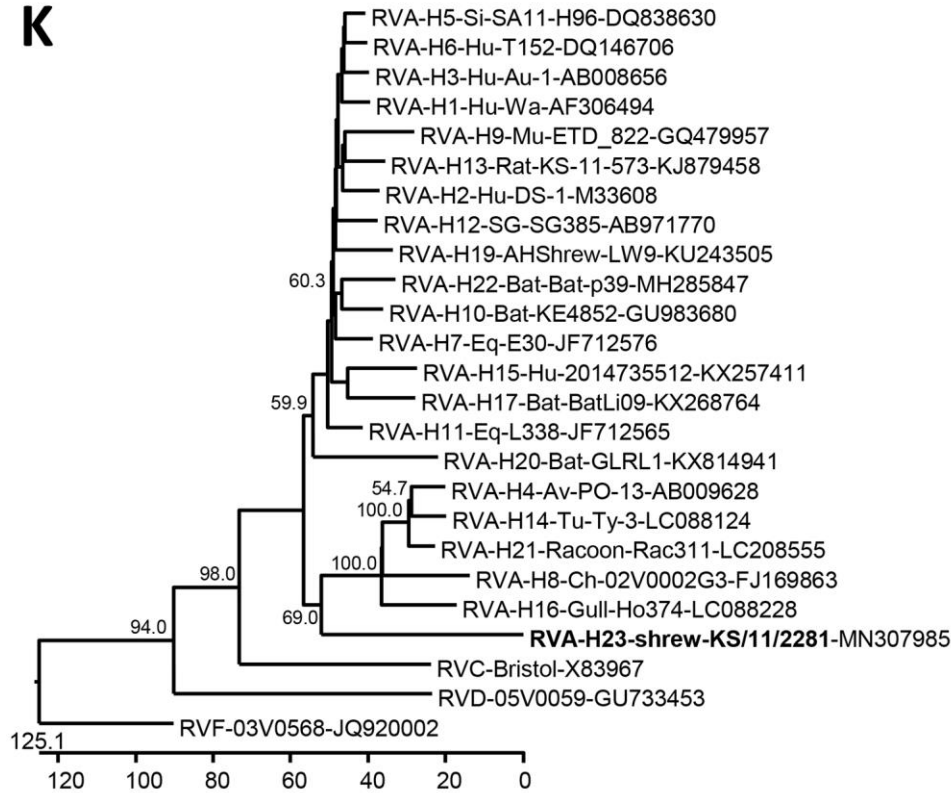
H



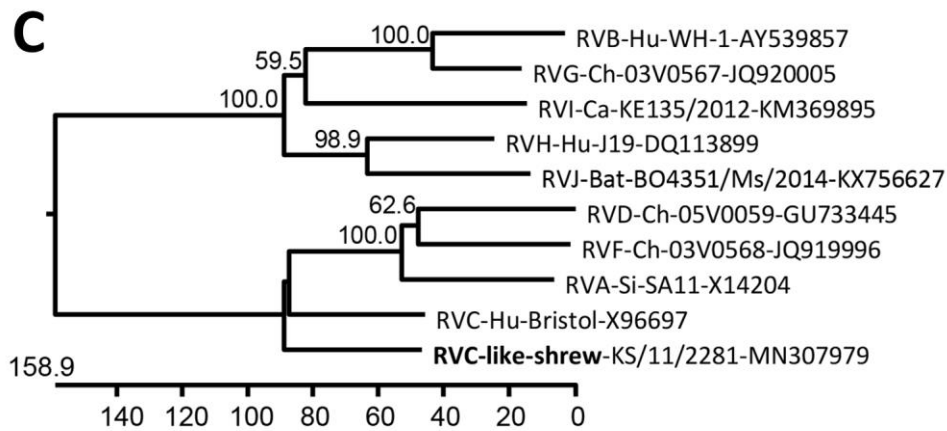
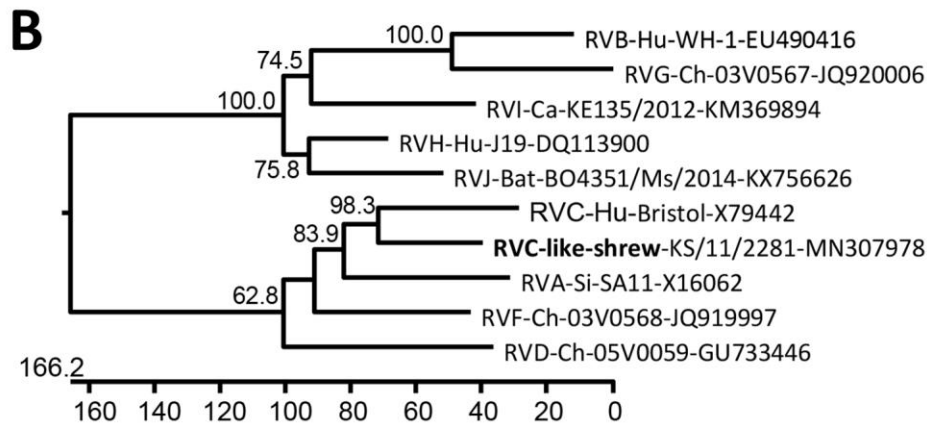
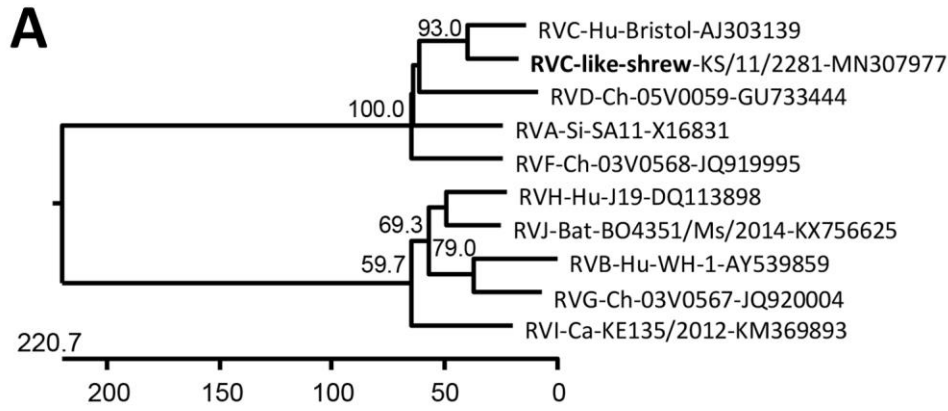


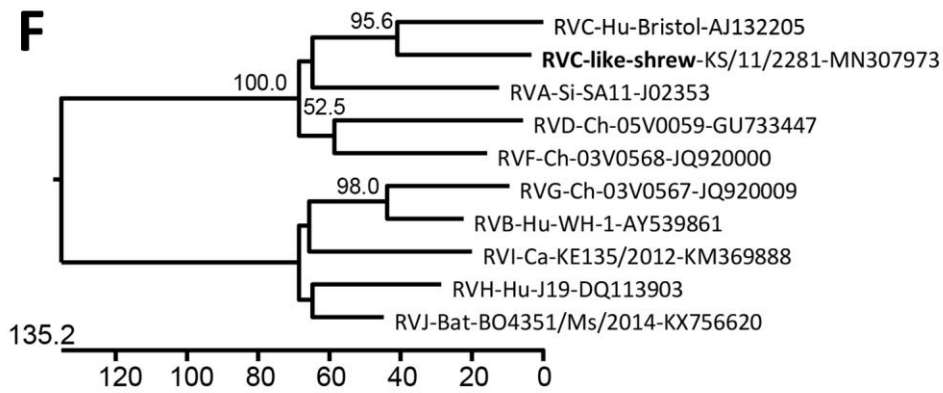
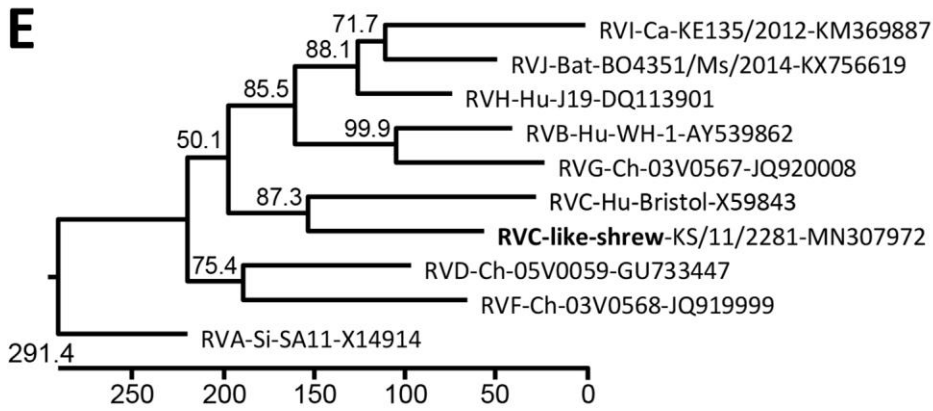
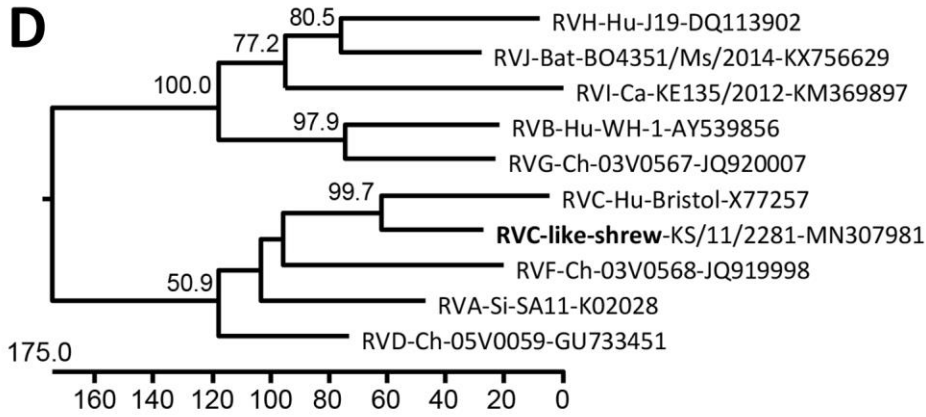
J

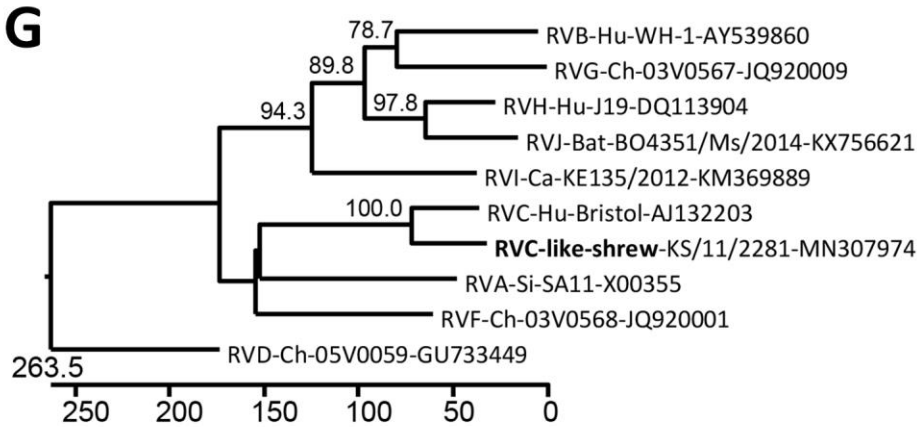


K

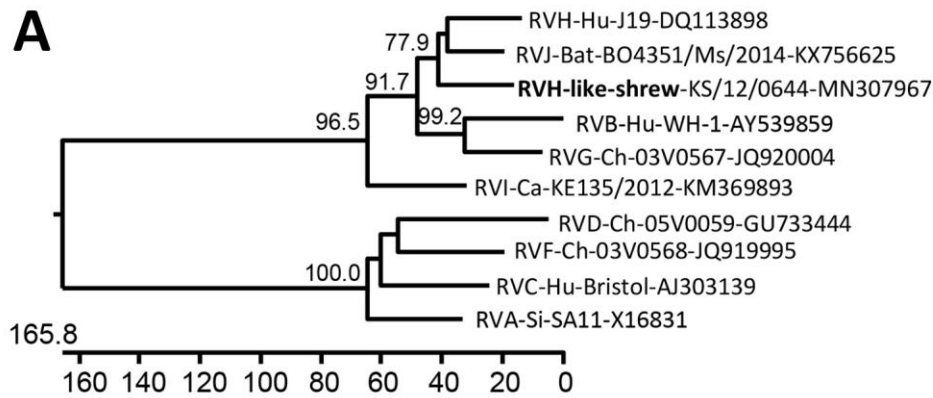
Appendix Figure 1. Phylogenetic relationship of the shrew rotavirus A (RVA) with genotype reference strains based on complete or partial nucleotide sequences of the genome segments encoding virus protein (VP) 1 (A), VP2 (B), VP3 (C), VP4 (D), VP6 (E), VP7 (F), nonstructural protein (NSP) 1 (G), NSP2 (H), NSP3 (I), NSP4 (J), and NSP5 (K). The rotavirus genotype, host (for abbreviations see below), strain designation and GenBank accession numbers are indicated at the branches of the tree. Rotavirus C (RVC), D (RVD), and F (RVF) strains are included as outgroups. The shrew virus from this study is given in bold letters. The tree was constructed using a neighbor-joining method implemented in the MEGALIGN module of DNASTAR software package (Lasergene) and a bootstrap analysis with 1000 trials and 111 random seeds. Bootstrap values of >50% are shown. The trees are scaled in nucleotide substitutions per 100 residues. Alp, alpaca; Av, avian; Bo, bovine; Ca, canine; Ch, chicken; Cm, camel; Eq, equine; Gu, guanaco; Hu, human; Mu, murine; Ph, pheasant; Pi, pigeon; Po, porcine; Ra, rabbit; Rh, rhesus monkey; RVA, rotavirus A; RVC, rotavirus C; RVD, rotavirus D; RVF, rotavirus F; SG, sugar glider; AHShrew, Asian house shrew; shrew, common shrew from this study; Si, simian; Tu, turkey; Vic, vicuna.

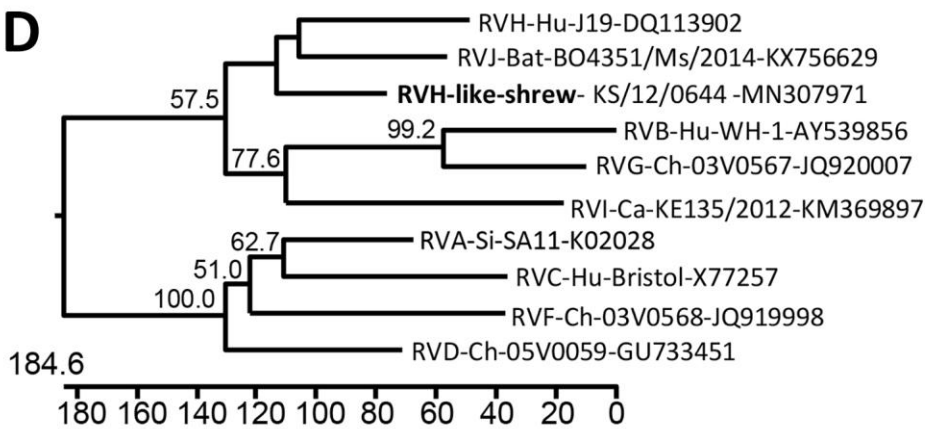
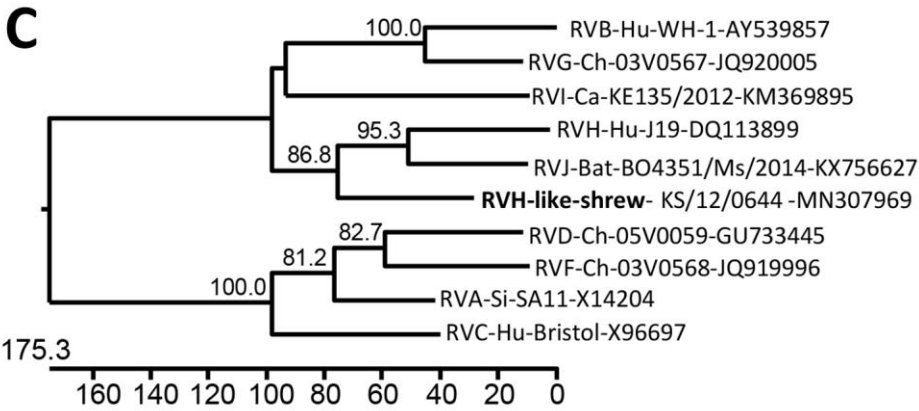
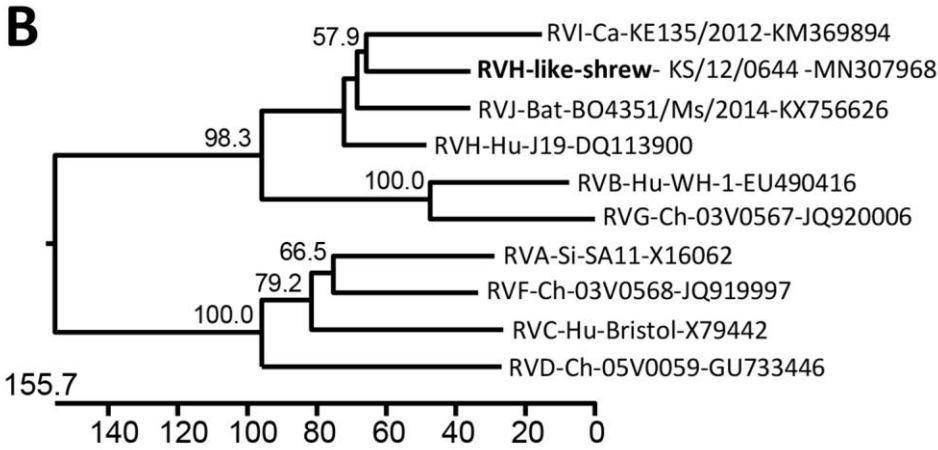


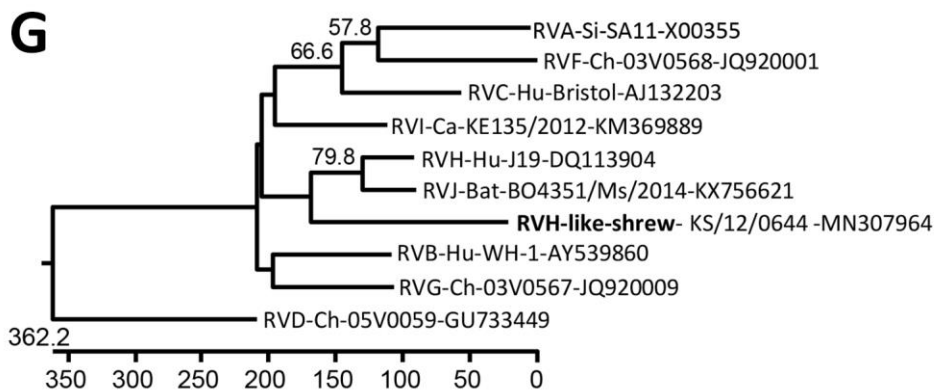
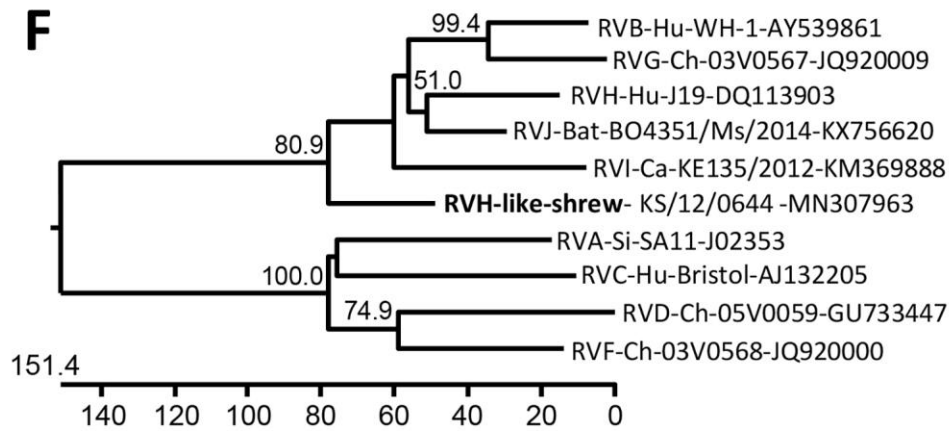
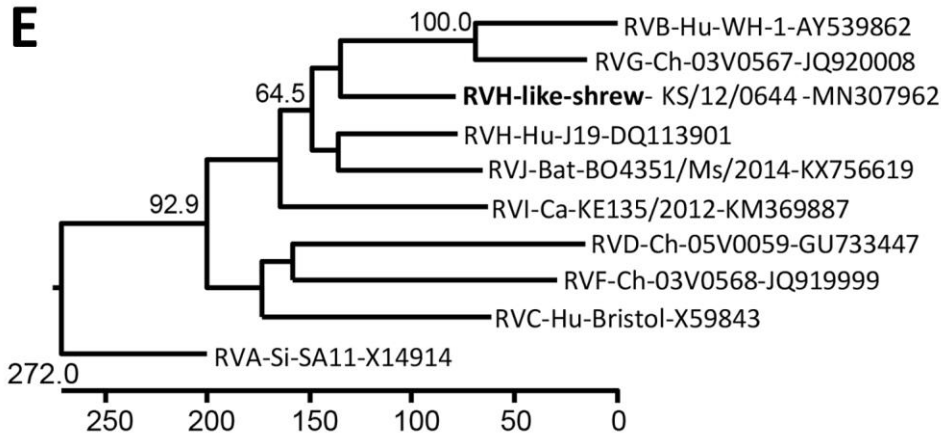




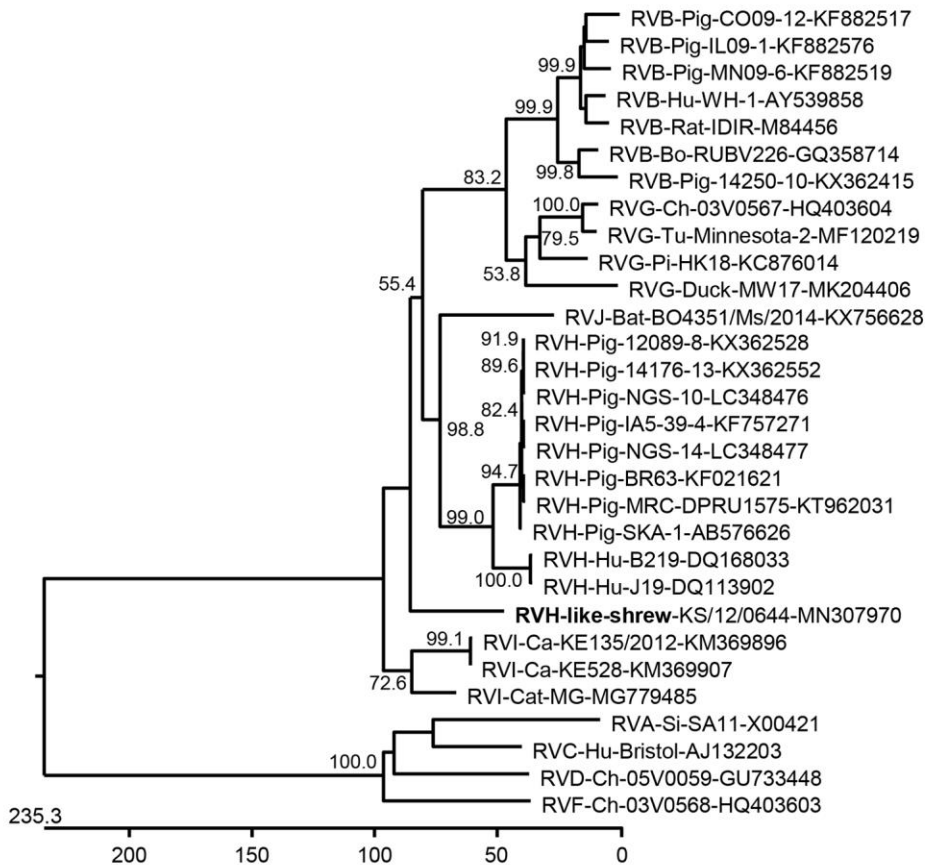
Appendix Figure 2. Phylogenetic relationship of the shrew species C-like rotavirus with the rotavirus A-J type species based on the deduced amino acid sequences of partial virus protein (VP) 2 (A), VP3 (B), VP4 (C), VP7 (D), nonstructural protein (NSP) 1 (E), NSP2 (F), and NSP3 (G). The rotavirus species (RVA-RVJ), host (for abbreviations see below), strain designation, and GenBank accession numbers are indicated at the branches of the tree. The shrew virus is given in bold letters. The tree was constructed using a neighbor-joining method implemented in the MEGALIGN module of DNASTAR software package (Lasergene) and a bootstrap analysis with 1000 trials and 111 random seeds. Bootstrap values of >50% are shown. The trees are scaled in amino acid substitutions per 100 residues. aa, amino acid residues; Ca, canine; Ch, chicken; Hu, human; RVA, rotavirus A; RVB, rotavirus B; RVC, rotavirus C; RVD, rotavirus D; RVF, rotavirus F; RVG, rotavirus G; RVH, rotavirus H; RVI, rotavirus I; RVJ, rotavirus J; Si, simian.







Appendix Figure 3. Phylogenetic relationship of the shrew species H-like rotavirus with the rotavirus A-J type species based on the deduced amino acid sequences of partial or complete virus protein (VP) 2 (A), VP3 (B), VP4 (C), VP7 (D), nonstructural protein (NSP) 1 (E), NSP2 (F), and NSP3 (G). The rotavirus species (RVA-RVJ), host (for abbreviations see below), strain designation, and GenBank accession numbers are indicated at the branches of the tree. The shrew virus is given in bold letters. The tree was constructed using a neighbor-joining method implemented in the MEGALIGN module of DNASTAR software package (Lasergene) and a bootstrap analysis with 1000 trials and 111 random seeds. Bootstrap values of >50% are shown. The trees are scaled in amino acid substitutions per 100 residues. aa, amino acid residues; Ca, canine; Ch, chicken; Hu, human; RVA, rotavirus A; RVB, rotavirus B; RVC, rotavirus C; RVD, rotavirus D; RVF, rotavirus F; RVG, rotavirus G; RVH, rotavirus H; RVI, rotavirus I; RVJ, rotavirus J; Si, simian.



Appendix Figure 4. Phylogenetic relationship of the shrew species H-like rotavirus within the rotavirus B/G/H/I/J cluster. Type species of rotavirus A, C, D, and F are included as outgroup. The complete deduced amino acid sequences of VP6 were used for construction of a phylogenetic tree using a neighbor-joining method implemented in the MEGALIGN module of DNASTAR software package (Lasergene) and a bootstrap analysis with 1000 trials and 111 random seeds. Bootstrap values of >50%

are shown. The tree is scaled in amino acid substitutions per 100 residues. The rotavirus species (RVA-RVJ), host (for abbreviations see below), strain designation, and GenBank accession numbers are indicated at the branches of the tree. The shrew virus is given in bold letters. Bo, bovine; Ca, canine; Ch, chicken; Hu, human; Pi, pigeon; RVA, rotavirus A; RVB, rotavirus B; RVC, rotavirus C; RVD, rotavirus D; RVF, rotavirus F; RVG, rotavirus G; RVH, rotavirus H; RVI, rotavirus I; RVJ, rotavirus J; Si, simian; Tu, turkey.