Virus Research 334 (2023) 199171

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Genome analysis of the novel putative rotavirus species K

Reimar Johne [a,*], Simon H. Tausch [a], Rainer G. Ulrich [b], Katja Schilling-Loeffler [a]

[a] German Federal Institute for Risk Assessment, Berlin 10589, Germany
[b] Institute of Novel and Emerging Infectious Diseases, Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, and Partner Site Hamburg-Lübeck-Borstel-Riems, German Centre for Infection Research (DZIF), Greifswald-Insel, Riems 17493, Germany

ARTICLE INFO

Keywords:
Rotavirus
Genome sequence
Taxonomy
Phylogeny
Shrew

ABSTRACT

Rotaviruses are causative agents of diarrhea in humans and animals. Currently, the species rotavirus A-J (RVA-RVJ) and the putative species RVK and RVL are defined, mainly based on their genome sequence identities. RVK strains were first identified in 2019 in common shrews (*Sorex aranaeus*) in Germany; however, only short sequence fragments were available so far. Here, we analyzed the complete coding regions of strain RVK/shrew-wt/GER/KS14–0241/2013, which showed highest sequence identities with RVC. The amino acid sequence identity of VP6, which is used for rotavirus species definition, reached only 51% with other rotavirus reference strains thus confirming classification of RVK as a separate species. Phylogenetic analyses for the deduced amino acid sequences of all 11 virus proteins showed, that for most of them RVK and RVC formed a common branch within the RVA-like phylogenetic clade. Only the tree for the highly variable NSP4 showed a different branching; however, with very low bootstrap support. Comparison of partial nucleotide sequences of other RVK strains from common shrews of different regions in Germany indicated a high degree of sequence variability (61–97% identity) within the putative species. These RVK strains clustered separately from RVC genotype reference strains in phylogenetic trees indicating diversification of RVK independent from RVC. The results indicate that RVK represents a novel rotavirus species, which is most closely related to RVC.

1. Introduction

Rotaviruses are a major cause of diarrhea in humans and animals. Especially young individuals are affected, in which life-threatening severe diseases can occur. For humans, 128,500 deaths were estimated due to rotavirus infection worldwide for 2016 (Troeger et al., 2018). Rotavirus-induced disease is also very common in domestic animals (Otto et al., 2012, 2015). In addition, rotavirus infections have been identified in several wild animals (Althof et al., 2023; Simsek et al., 2021). Although distinct rotavirus types are commonly found in specific hosts, zoonotic transmission of rotaviruses between different animal species and between animals and humans have been repeatedly described (Martella et al., 2010; Díaz Alarcón et al., 2022).

The genome of rotaviruses consists of 11 segments of double-stranded RNA, which each encodes one of the structural proteins VP1-VP4, VP6-VP7, or the non-structural proteins NSP1-NSP5. The non-enveloped virus particle consists of VP2 forming the inner core, VP6 forming the intermediate shell, and VP4 and VP7 representing the outer capsid proteins. VP1 and VP3 as well as NSP1-NSP5 are involved in viral genome replication and interact with cellular proteins to enable virus propagation (Crawford et al., 2017).

Rotaviruses are grouped into the genus *Rotavirus* within the family *Sedoreoviridae* (Matthijnssens et al., 2022). The genus contains several rotavirus species, which are defined as "…unable to reassort their genome segments under normal circumstances, and each species may therefore represent a separate gene pool" (Matthijnssens et al., 2021). For demarcation of different rotavirus species, a cut-off value of 53% amino acid sequence identity for the VP6 was established (Matthijnssens et al., 2012). Until now, the species rotavirus A-D (RVA-RVD) and F-J (RVF-RVJ) have been designated by the International Committee on Taxonomy of Viruses (ICTV, 2020). Of these, RVA generally has the highest clinical importance in humans and several mammal species (Crawford et al., 2017). However, RVC is increasingly considered an important pathogen for domestic pigs (Theuns et al., 2016), as well as RVD for chicken (Otto et al., 2012).

Phylogenetically, the rotavirus species cluster into two major clades reflecting their evolutionary relationship. The RVA-like clade comprises RVA, RVC, RVD and RVF, whereas the RVB-like clade contains RVB and RVG-RVJ (Kindler et al., 2013; Johne et al., 2022). This phylogenetic relationship is also reflected by structural and functional differences

* Corresponding author.
E-mail address: reimar.johne@bfr.bund.de (R. Johne).

https://doi.org/10.1016/j.virusres.2023.199171
Received 9 May 2023; Received in revised form 7 July 2023; Accepted 8 July 2023
Available online 11 July 2023

between the viral proteins of the different clades as previously demonstrated for VP1 (Ogden et al., 2012). Furthermore, differences in the highly conserved terminal nucleotide sequences of the genome segments are evident between the representatives of the different clades. Whereas all rotaviruses show similar sequences at the 5′-termini, specific sequences are present at the 3′-termini of the RVA-like and RVB-like clades. The only exception is RVC, which uses a unique sequence at its 3′-termini (Matthijnssens et al., 2021).

In addition to the established rotavirus species, the tentative species rotavirus K (RVK) and L (RVL) have been identified in 2019 in common shrews (*Sorex araneus*) in Germany (Johne et al., 2019). For RVL, a complete genome sequence of a reference strain has been recently determined, which confirms its classification as a separate species and its affiliation to RVB-like clade (Johne et al., 2022). However, for RVK only short sequence fragments were available so far thus preventing analysis of its distinct phylogenetic relationship and final classification.

Here we describe the sequence analysis of complete coding regions of an RVK strain from a common shrew in Germany. Analysis of the deduced amino acid sequence of VP6 confirms the classification of this rotavirus into a separate species RVK. Sequence identity analyses and phylogenetic trees indicate that RVK has the closest relationship to RVC within the RVA-like clade. Analyses of partial sequences of other RVK strains assess the sequence variation within RVK, which points to diversification of this rotavirus species in common shrews.

2. Material & methods

2.1. Samples

The investigated common shrew samples originated from a collection of 73 wild animals from different regions of Germany trapped between 2004 and 2016. Details of trapping and dissection of the animals have been described recently (Johne et al., 2019; Falkenhagen et al., 2022); additional 10 samples collected between 2012 and 2016 as bycatches within a small mammal monitoring program in different regions in Germany (Thuringia, Lower Saxony and Bavaria) were also included. Intestinal contents of the animals were diluted 1:10 in phosphate-buffered saline (PBS) and stored at −20°C. Thereafter, nucleic acids were isolated using the NucliSense platform on an E-MAG device (BioMerieux, Marcy I'Etoile, France), and subsequently stored at −80°C.

2.2. RVK detection by reverse transcription-polymerase chain reaction (RT-PCR)

Based on an alignment of available sequence fragments of RVK, a primer pair was designed for RT-PCR, targeting a region within the VP1 gene. Briefly, primers RVK-VP1s (5′-TAC ACT GMG ACT ATT YCC ACT ACA-3′) and RVK-VP1as (5′- ACR ACT GAW GTC TGT GAT TCT CC-3′) were used for amplification with the OneStep RT-PCR kit (Qiagen, Hilden, Germany). The cycling conditions were 42°C for 30min and 95°C for 15min followed by 40 cycles each including 95°C for 30s, 56°C for 30s and 74°C for 40s, with a final incubation at 74°C for 5min. RT-PCR products were separated on ethidium bromide-stained agarose gels and selected products were purified with the Monarch DNA Gel Extraction Kit (New England BioLabs, Ipswich, MA, USA). Thereafter, they were subjected to Sanger sequencing by a commercial company (Eurofins, Ebersberg, Germany) using the same primers, which were used for RT-PCR.

2.3. Next generation sequencing (NGS) and data analysis

For sample KS14–0241 a library was previously generated with the KAPA RNA HyperPrep Kit (Roche Diagnostics, Mannheim, Germany) and the KAPA Unique Dual-Indexed Adapter Kit for Illumina® platforms (Roche Diagnostics) (Johne et al., 2022). This previously generated and sequenced library was re-sequenced with 2×300 cycles on the MiSeq (Illumina, San Diego, CA, USA) as a single sample using the MiSeq Reagent Kit v3 (600) (Illumina) to increase sequencing depth and read length.

The generated raw sequences were trimmed using the fastp module (Chen et al., 2018) of the AQUAMIS pipeline (Deneke et al., 2021). Remaining reads were subjected to assembly using SKESA (Souvorov et al., 2018). All resulting contigs were screened for sequence similarities with rotavirus sequences using BLASTX (Altschul et al., 1997) against all rotavirus sequences available from the NCBI protein database (https://www.ncbi.nlm.nih.gov/protein/, accessed on 20 October 2021).

2.4. Conventional RT-PCR and rapid amplification of cDNA ends

Sequences of two gaps in the contigs of the VP2- and VP3-encoding sequences were generated by RT-PCR with specific primers delineated from the contig sequences, and determined by Sanger sequencing as described above. For sequencing of missing 5′- and 3′-termini of the genome segments, rapid amplification of cDNA ends (RACE) protocols were applied. Primer sequences for RACE were delineated from the NGS-derived contig sequences. First, the 5′ RACE-System (Invitrogen GmbH, Karlsruhe, Germany) was used. If no RACE product could be amplified by this kit, the 5′/3′ RACE Kit (Roche Diagnostics) was applied. Both kits were used according to the manufacturer's instructions. Products were separated on ethidium-bromide-stained agarose gels, purified with the Monarch DNA Gel Extraction Kit (New England BioLabs) and subjected to Sanger sequencing by a commercial company (Eurofins) using the RACE primers.

2.5. Sequence analysis

Contigs and sequences obtained from NGS and Sanger sequencing, respectively, were manually edited, assembled, translated into amino acid sequences and used for concatenation of sequences using the SeqBuilder Pro 17 module of the DNASTAR software (Lasergene, Madison, WI, USA). The final sequences determined for the 11 genome segments of strain RVK/shrew-wt/GER/KS14–0241/2013 and partial sequences of other RVK strains were deposited at GenBank database with the accession numbers OQ934016 - OQ934035. Alignments and calculations of sequence distances were performed using the ClustalW method with the MegAlign Pro 17 module of the DNASTAR software (Lasergene). The GenBank accession numbers of the reference strains for rotavirus species used in alignments are listed in Supplementary Data S1. Phylogenetic analyses were performed with MEGA X version 10.1.7 (Kumar et al., 2018). Sequences were aligned using ClustalW, and phylogenetic trees for amino acid sequences of each protein and the concatenated proteins were constructed by the Maximum Likelihood method with parameters: 1000 bootstrap replications, Jones-Taylor-Thornton method, uniform rates among sites, all sites used. For nucleotide sequences, phylogenetic trees were constructed by the Maximum Likelihood method with parameters: 1000 bootstrap replications, Tamura-Nei model, uniform rates among sites, all sites used.

3. Results

3.1. Screening for RVK and selection of a strain for NGS

Screening of 73 intestinal contents of common shrews from different regions in Germany using an RVK-VP1-specific RT-PCR identified RVK in eight (11%) of the samples. Sequencing of the PCR product by the Sanger method was successful for six of these samples, which confirmed the presence of RVK due to the highest sequence identities to RVK sequences by Nucleotide BLAST search (sequence analysis is presented in chapter 3.6). Out of these, samples KS/11/2281, KS/12/0644 and

KS14–0241 showed the strongest bands in the RVK-VP1-specific RT-PCR and the best quality of Sanger sequencing results. Of note, next generation sequencing data were already available for these three samples (Johne et al., 2019, 2022), by which RVA and RVK were originally identified in sample KS/11/2281, and RVL in samples KS/12/0644 and KS14–0241. A new analysis of these NGS data identified RVK-specific sequences in all of the three samples. Whereas these sequences were only fragmentary for samples KS/11/2281 and KS/12/0644, RVK-specific sequences with identities to all 11 rotavirus genome segments were identified in sample KS14–0241. However, even in the sequence data from this sample, only partial open reading frame sequences could be delineated. Therefore, the sample was selected for further NGS investigation with more sequencing depth and increased read length. Sample KS14–0241 originated from a female common shrew with a body weight of 6g, which was trapped near Stuttgart, southwest Germany, in June 2013, and the complete rotavirus strain designation is RVK/shrew-wt/GER/KS14–0241/2013.

3.2. Genome sequencing by NGS

NGS of sample KS14–0241 resulted in the generation of 49,112,464 paired reads, which resulted in a final set of 44,441,942 paired reads after trimming. By assembly, 26,135 contigs were generated, out of which 31 contigs with similarities to sequences of different rotavirus species were extracted via a BLASTX search. From these, 13 contigs showing the highest sequence similarities to RVK or RVC sequences available at NCBI Protein database were manually selected using BLASTX search. These contigs had a mean read coverage between 11x and 1,391x (average: 368x), and represented most parts of all eleven genome segments of RVK. Missing parts in the central region of the VP2- and VP3-encoding genome segments were amplified by RT-PCR using specific primers, and sequences were subsequently determined by the Sanger method. In addition, all 5′- and 3′-terminal sequences of the genome segments including parts of some open reading frames were missing. For amplification of the genome segment termini, RACE protocols were applied, leading to the completion of all open reading frames and extension into the non-coding sequence regions for most segments. However, the complete sequences of the non-coding segment termini could not be determined for all segments (see chapter 3.4).

3.3. Coding capacity of RVK and sequence identities with other rotaviruses

The RVK genome detected in sample KS14–0241 has 11 ORFs encoding the structural proteins VP1-VP4, VP6-VP7 and the non-structural proteins NSP1-NSP5. As shown in Table 1, the lengths of the corresponding amino acid sequences are in the range of that of other rotaviruses; only NSP4 with a length of 234 amino acid residues is slightly longer than that of the other rotaviruses (128–219 amino acid residues). Alignments of the amino acid sequences of all 11 proteins of rotavirus species reference strains and RVK are presented in Supplementary Data S2. The NSP4 alignment shows that an N-terminal extension and a short central insertion are responsible for the increased length of RVK-NSP4. The amino acid sequence identities of the RVK proteins and the other rotavirus species proteins are between 2% and 63%, with the lowest for NSP4 and the highest for VP2 (Table 2). Generally, RVK proteins show the highest amino acid sequence identities to RVC. Only NSP4 shows the highest identity to RVH; however, with only 13% identity. The highest amino acid sequence identity of VP6 is 51% with the RVC reference strain, which is below the cut-off value of 53% used for definition of rotavirus species.

3.4. Analysis of non-coding regions

The determined sequences of the non-coding regions of RVK are shown in Table 3. Usually, rotaviruses have conserved sequences at the 5′- and 3′-termini of all of their genome segments. It is evident from Table 3 that those conserved sequences could not be identified in the determined RVK genome segment sequences, suggesting that these sequences are incomplete. For the 5′-termini of the VP1- and VP7-encoding segments, similar sequences like those present in RVA, RVC and all other rotaviruses are evident (shown in bold face in Table 3); this might indicate, that the 5′-termini for these segments are complete. In contrast, no identity with terminal sequences of RVA, RVC or other rotaviruses could be identified for the 3′-ends.

Table 2
Identity of deduced amino acid sequences of RVK strain RVK/shrew-wt/GER/KS14–0241/2013 to the reference strains of the established rotavirus species RVA-RVJ and the putative species RVL.

	% amino acid sequence identity	Highest amino acid sequence identity to
VP1	21 – 59	RVC
VP2	11 - 63	RVC
VP3	11 - 49	RVC
VP4	9 - 34	RVC
VP6	9 - 51	RVC
VP7	11 - 46	RVC
NSP1 *	5 - 21	RVC
NSP2	11 - 49	RVC
NSP3	11 - 53	RVC
NSP4	2 - 13	RVH
NSP5	5 - 24	RVC

* for RVB, RVG and RVI, which encode two forms of NSP1 (NSP1–1 and NSP1–2), identities with NSP1–2 are shown.

Table 1
Amino acid sequence lengths of encoded proteins of rotavirus K strain RVK/shrew-wt/GER/KS14–0241/2013 (RVK, marked in bold face) compared to that of other rotavirus species reference strains.

Protein	Number of amino acid residues in proteins										
	RVA	RVB	RVC	RVD	RVF	RVG	RVH	RVI	RVJ	RVK	RVL
VP1	1088	1160	1090	1080	1086	1160	1167	1162	1168	**1086**	1168
VP2	881	934	884	914	904	991	973	983	986	**899**	981
VP3	835	764	693	686	694	768	719	701	715	**703**	715
VP4	776	750	744	778	738	772	823	777	826	**801**	852
VP6	397	391	395	399	396	391	396	395	395	**394**	396
VP7	326	249	332	317	295	247	258	270	245	**319**	245
NSP1*	495	321	394	574	547	324	395	390	401	**406**	326
NSP2	317	301	312	311	318	282	297	301	299	**313**	310
NSP3	315	347	402	371	370	303	262	273	331	**405**	425
NSP4	175	219	150	128	169	187	213	219	210	**234**	204
NSP5	198	170	212	195	218	181	176	157	165	**217**	186

* for RVB, RVG and RVI, which encode two forms of NSP1 (NSP1–1 and NSP1–2), the number of amino acid residues of NSP1–2 is shown.

Table 3
Determined sequences for the non-coding regions of the RVK strain RVK/shrew-wt/GER/KS14–0241/2013 genome segments and comparison with terminal consensus sequences of RVA and RVC. The sequences of strands of positive polarity are shown in 5′–3′ direction. The nucleotides of RVK identical to the 5′-terminal consensus sequences of RVA and RVC are marked in bold face.

Rotavirus species / encoded gene	5'-Terminus	3'-Terminus
RVK		
VP1	**GGCAAAUUAA**GUACUGAGCAG	UAAAGAGCU
VP2	CGCGAACGUUGGUCUGGGUGACA	UAAGCUGAUUUAACAUUCCAG
VP3	CCUGUAGUACUUGUUGUAGUCUCGCC UCGCAAGGAGUUGGAGCU	UAGCACUGCCUAGCCUAACCGCGCUCUGUGGC
VP4	GCACCACU	UAGUGGUCG
VP6	CUCAAG	UAGUGAGACCACGAUGUACAAAGUUGUGGUUAA CAGAGCACUAGACAGCAUGUUUCGUACAGCCAA UCAUACCAGGGUAACCGGGCAGCACGGAAUGG GAGAACCAACCGCCGGCCAACCUGGAGCACUA UGAGUAACUGGC
VP7	**GGCUUUUUAA**AGGAUCGGUGAUUUCCG AUAGGAGUCCCAUCACGCAGUGUGAAGG	UAAUCCGAAGGACACCUUGAUGGGGUAACUC UGAC
NSP1	GAAAGGGAUCAAGGAAGAGACCAUC	UGACUUCCUCCCCCCUGACGUUA CCGCAAUGACGCAGCAAUCGCAAUG
NSP2	UCCGUCCCAGUCUGAUCAACUGGCGCC	UAAUCACUUAGACAGGAG
NSP3	GUCAAGAGGCAAA	UAAUACUAAUUGAUAUGCCACCUGGAGAAAAUA GUUACAAGUAGAUAGAUUUUGACGUGGAAAAU CUUAACACAAUGCGUGCAGACUGGAUGCCGGGU CACUGUACCCUCAGCCCUUGAAAACUGGAUUGU GGCGAUCCACGCAAUUGUGA
NSP4	GUUCGAGGAAGGAGGAGGCU	UAAUGUAAUAUAAUUCUAGGUGACC
NSP5	GCGGUG	UAGAUGUGCUGCCAGUCUAU UUUCCAUGGUGACAUGAGUUAG CGUCGAUGGGUGAC
RVA (consensus)	**GGC**($^{A}/_{U}$)$_7$	UGUGACC
RVC (consensus)	**GGC**($^{A}/_{U}$)$_7$	UGUGGCU

3.5. Phylogenetic analysis of deduced amino acid sequences

The phylogenetic relationship of RVK strain RVK/shrew-wt/GER/KS14–0241/2013 with the reference strains of other rotaviruses species was analyzed based on amino acid sequences of the proteins. The resulting phylogenetic trees are shown in Fig. 1, indicating that RVK generally clusters with RVC. In a larger context, RVK and RVC branch together with RVA, RVD and RVF in one clade (RVA-like clade), whereas the other rotaviruses form another clade (RVB-like clade). The branching is only different for NSP4, where RVK branches together with RVA, RVD, RVF and RVI, but RVC clusters within the RVB-like clade. However, bootstrap support is very low for this tree. An additional tree based on the concatenated amino acid sequences of all encoded proteins of rotavirus species reference strains confirms the grouping of RVK into the RVA-like clade and its closest relationship with RVC (Fig. 2).

3.6. Analysis of sequence variability between different RVK strains

Additional partial RVK-specific sequences were available from the NGS data of two other common shrew samples (KS/11/2281 and KS/12/0644) and from the RVK-specific PCR-product for VP1 of three additional samples. The nucleotide sequence identities between the corresponding regions of these samples ranged from 61% to 97%, indicating high sequence variability between different RVK strains. The distinct identity range for each segment fragment is shown in Supplementary Data S3. Phylogenetic trees based on these nucleotide sequence fragments, which are shown exemplarily for the VP1- and VP6-encoding fragments in Fig. 3, confirm the diversity within the putative RVK species. In these trees, sequences of genotype reference strains for RVC (Wang et al., 2021) have also been included. It is evident from the trees that RVC and RVK sequences cluster in well-separated branches, which contain each a variety of sub-branches reflecting the variability within each virus species.

4. Discussion

Rotaviruses display a high degree of genetic heterogeneity, which is reflected by the definition of the nine species RVA-RVD and RVF-RVJ (ICTV 2020), and the tentative species RVK and RVL (Johne et al., 2019). Recently, genome sequence identities have mostly been used for distinction of different virus species, and a sequence identity cut-off of 53% for the deduced amino acid sequence of VP6 has been established for rotavirus species definition (Matthijnssens et al., 2012). Here, we generated the first nearly complete genome sequence of an RVK strain, which could be used for its further characterization and classification. Based on its VP6 sequence, the highest amino acid sequence identity to other rotavirus species is 51% with RVC, which is below the cut-off value of 53%. Therefore, the classification of RVK as a separate rotavirus species is confirmed based on this criterion.

By comparison to the amino acid sequences of other rotavirus proteins, the highest sequence identities of RVK were generally with RVC. Also, in phylogenetic trees, RVK amino acid sequences generally clustered together with RVC sequences, indicating a close phylogenetic relationship of both viruses. RVC has been frequently detected in diseased humans (Zhao et al., 2022), but it is increasingly identified as an important pathogen in pigs (Theuns et al., 2016) and with lower detection rates also in other animal species (Otto et al., 2015). The relationship of RVK with RVC was evident for all virus proteins, with the exception of NSP4, which showed very low sequence identities and a different branching in phylogenetic trees. However, NSP4 sequences are known to be highly variable between different rotavirus species (Johne et al., 2022), and the low bootstrap support in the respective phylogenetic tree reflects the difficulties in reconstruction of phylogenetic relationship for this protein. Despite the high amino acid sequence diversity, it has been shown that NSP4 of RVA and RVC are able to exert similar functions as an enterotoxin (Sasaki et al., 2001).

In a broader evolutionary view, the rotaviruses can be separated into the two major phylogenetic RVA- and RVB-like clades (Kindler et al., 2013; Johne et al., 2022). Based on our phylogenetic trees, RVK can be classified into the RVA-like clade. Beside the phylogenetic clustering,

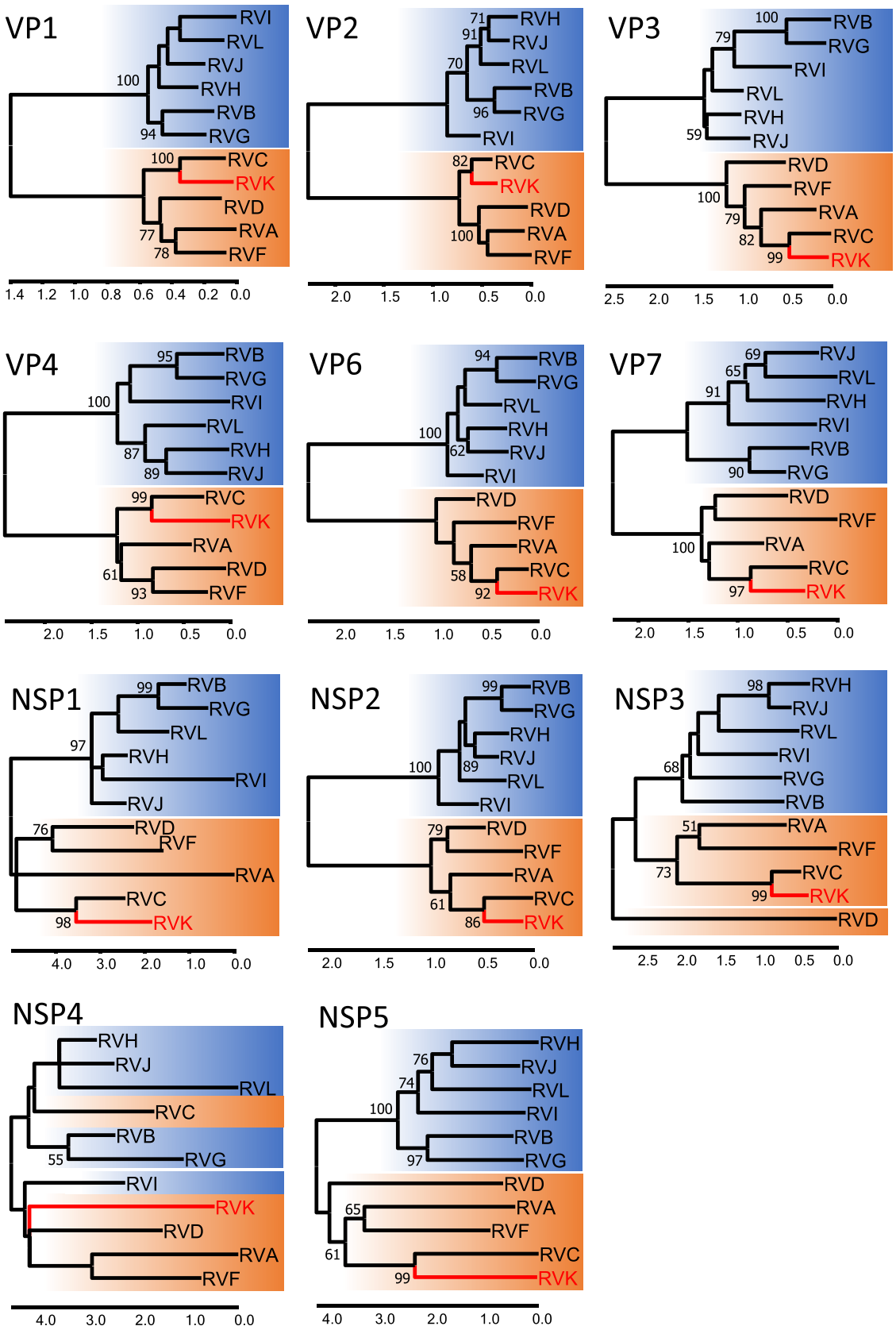

Fig. 1. Phylogenetic relationship of RVK proteins VP1-VP4, VP6-VP7 and NSP1-NSP5 with that of other rotavirus species. The complete deduced amino acid sequences of the encoded proteins of RVK strain RVK/shrew-wt/GER/KS14–0241/2013 were compared with those of reference strains of the other rotavirus species (as specified in Supplementary Data S1) by the Maximum Likelihood method using MEGA X. RVK is marked in red and bootstrap values > 50% are shown. Scaled in amino acid substitutions per site. Rotavirus species belonging to the evolutionary RVA-like clade are shaded in orange, whereas those of the RVB-like clade are shaded in blue. For RVB, RVG and RVI, which encode two forms of NSP1 (NSP1–1 and NSP1–2), only NSP1–2 was used in the NSP1 tree.

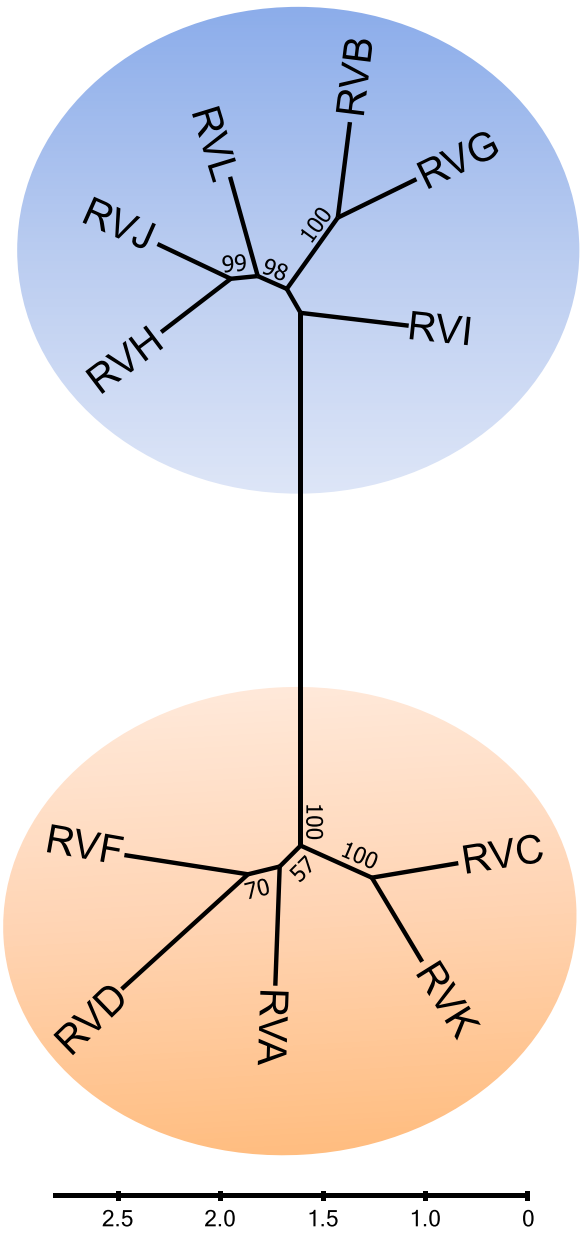

Fig. 2. Phylogenetic relationship of rotavirus species based on the concatenated amino acid sequences of VP1-VP4, VP6-VP7 and NSP1-NSP5. The concatenated amino acid sequences of proteins of RVK strain RVK/shrew-wt/GER/KS14–0241/2013 were compared with that of reference strains of the other rotavirus species (as specified in Supplementary Data S1) by the Maximum Likelihood method using MEGA X, and a radiated tree was constructed. Bootstrap values > 50% are shown. Scaled in amino acid substitutions per site. Rotavirus species belonging to the evolutionary RVA-like clade are shaded in orange, whereas those of the RVB-like clade are shaded in blue.

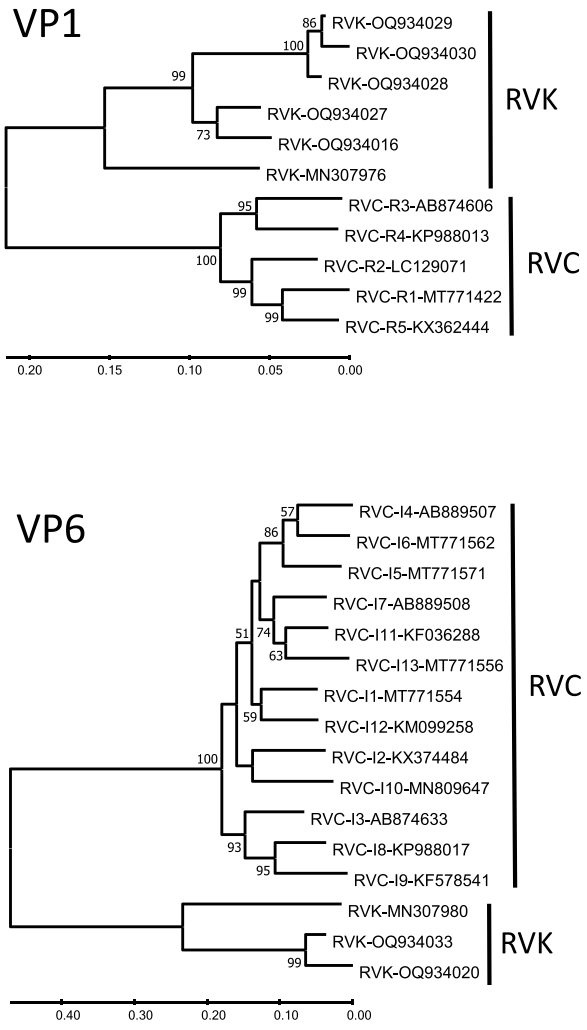

Fig. 3. Phylogenetic relationship of different RVK strains and RVC genotype reference strains. Overlapping partial nucleotide sequences (174 nucleotides for VP1, above, and 153 nucleotides for VP6, below) were compared by the Maximum Likelihood method using MEGA X. The virus species (RVK or RVC), genotype (for RVC only) and the GenBank accession numbers are shown at the branches. Bootstrap values > 50% are shown. Scaled in nucleotide substitutions per site.

both clades exert different structural properties and distinct mechanisms of RNA recognition as demonstrated by sequence analyses of VP1 (Ogden et al., 2012). These findings indicate that both clades separated early during evolution of rotaviruses, leading to two largely different groups of rotavirus species at present.

The analysis of the non-coding regions of the RVK genome was hampered by the incompleteness of the generated sequence data. Difficulties in sequencing of 5′- and 3′-termini of segmented RNA virus genomes are not uncommon, especially if only low virus genome amounts are present in the sample (Misu et al., 2023). From the generated data, the 5′-termini of the VP1 and VP7 genes seem to be complete, showing the consensus sequence GGC($^{A}/_{U}$)$_7$, which is common to RVA and RVC and similar to that found also in all other rotavirus species. However, no consensus sequence was identified in the generated sequences for the 3′-termini, most likely due to missing terminal sequences. RVA and most other representatives of RVA-like clade show the 3′-terminal sequence UGUGACC, but RVC has the sequence UGUGGCU (Johne et al., 2022). None of these terminal consensus sequences are completely found in the RVK sequence data determined in our study. As RVK clusters in the RVA-like clade, but is most closely related to RVC, it would be interesting to identify the complete 3′-terminal sequences in future studies.

As more RVK-positive common shrew samples were identified in our sample collection, an initial analysis of RVK diversity could be done here. By this, a considerable sequence variability was identified. This argues against a single spillover event from another host, but indicates diversification of RVK within common shrews. For the closest relative of RVK, which is RVC, a similar high genetic diversity of strains has been recently identified (Wang et al., 2021), which is also reflected by our phylogenetic analysis of partial nucleotide sequences. However, despite the close relationship of RVK and RVC, strains of RVK branch separately from those of RVC indicating a distinct evolutionary history of both virus

species.

In our sample collection, only low amounts of the RVK genome were detected in the samples. This finding is similar to that described for RVI in cats and dogs, which mostly occurs in low amounts (Mihalov-Kovács et al., 2015). Despite these low amounts, an association of RVI infections with diarrhea has been recently described (Phan et al., 2017). In addition, we found that those samples showing the highest RVK genome amounts were also co-infected with other rotaviruses, e.g. sample KS/11/2281 with RVA and samples KS/12/0644 and KS14–0241 with RVL. Co-infections are not uncommon for rotaviruses and have been frequently described (Otto et al., 2012, 2015). Future studies should clarify the effect of co-infections on RVK replication and vice versa, and more generally the clinical significance of RVK infections.

5. Conclusions

The nearly complete genome of an RVK strain was generated here, which confirms its classification as a separate rotavirus species. RVK is most closely related to RVC and both virus species cluster within the phylogenetic RVA-like clade of rotaviruses. The non-coding genome segment termini of the RVK genome could not be completely sequenced so far, but should be analyzed in future to identify the conserved terminal sequences usually present in each rotavirus species. A considerable sequence variation was found between different RVK strains, which indicates diversification of this rotavirus species within common shrews as a host species for RVK. Further studies are needed to elucidate the pathogenicity, host specificity and the zoonotic potential of this novel rotavirus species in future.

CRediT authorship contribution statement

Reimar Johne: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Writing – original draft, Project administration, Funding acquisition. **Simon H. Tausch:** Methodology, Software, Data curation, Writing – review & editing. **Rainer G. Ulrich:** Conceptualization, Methodology, Writing – review & editing. **Katja Schilling-Loeffler:** Methodology, Investigation, Data curation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We like to thank Jörg Thiel, Florian Krüger, Marion Saathoff and Britta Oltmann for providing the shrews collected as by-catches during rodent monitoring, Stephan Drewes, René Ryll, Sandra Jagdmann, Dörte Kaufmann, Maysaa Dafalla, Madlen Walther, Arne Auste, Elfi Schlohsarczyk, Kathrin Jeske, Florian Binder for dissection of animals and Viola Haring and Patrick Slowikowski for sample archive management. We would also like to thank Stefanie Prosetzky and Anja Schlosser for their excellent technical assistance in sample analysis. This study was funded by the Deutsche Forschungsgemeinschaft (DFG), Germany (grant number JO 369/5–2), and an intramural grant (No. 1322–713) of the German Federal Institute for Risk Assessment.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.virusres.2023.199171.

References

Althof, N., Trojnar, E., Johne, R., 2023. Rotaviruses in wild ungulates from Germany, 2019-2022 Microorganisms 11 (3), 566.
Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic. Acids Res. 25 (17), 3389–3402.
Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34 (17), i884–i890.
Crawford, S.E., Ramani, S., Tate, J.E., Parashar, U.D., Svensson, L., Hagbom, M., Franco, M.A., Greenberg, H.B., O'Ryan, M., Kang, G., Desselberger, U., Estes, M.K., 2017. Rotavirus infection. Nature Rev. Dis. Primers 3, 17083.
Deneke, C., Brendebach, H., Uelze, L., Borowiak, M., Malorny, B., Tausch, S.H., 2021. Species-specific quality control, assembly and contamination detection in microbial isolate sequences with AQUAMIS. Genes 12 (5), 644. Basel.
Díaz Alarcón, R.G., Liotta, D.J., Miño, S., 2022. Zoonotic RVA: state of the art and distribution in the animal world. Viruses 14 (11), 2554.
Falkenhagen, A., Tausch, S.H., Labutin, A., Grützke, J., Heckel, G., Ulrich, R.G., Johne, R., 2022. Genetic and biological characteristics of species A rotaviruses detected in common shrews suggest a distinct evolutionary trajectory. Virus Evol. 8 (1), veac004.
International Committee on Taxonomy of Viruses: Virus taxonomy: 2020 Release. Available online: https://talk.ictvonline.org/taxonomy/(accessed on 19 March 2023).
Johne, R., Tausch, S.H., Grützke, J., Falkenhagen, A., Patzina-Mehling, C., Beer, M., Höper, D., Ulrich, R.G., 2019. Distantly related rotaviruses in common shrews, Germany, 2004-2014 Emerg. Inf. Dis. 25 (12), 2310–2314.
Johne, R., Schilling-Loeffler, K., Ulrich, R.G., Tausch, S.H., 2022. Whole genome sequence analysis of a prototype strain of the novel putative rotavirus species L. Viruses 14 (3), 462.
Kindler, E., Trojnar, E., Heckel, G., Otto, P.H., Johne, R., 2013. Analysis of rotavirus species diversity and evolution including the newly determined full-length genome sequences of rotavirus F and G. Infect. Genet. Evol. 14, 58–67.
Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35 (6), 1547–1549.
Martella, V., Bányai, K., Matthijnssens, J., Buonavoglia, C., Ciarlet, M., 2010. Zoonotic aspects of rotaviruses. Vet. Microbiol. 140 (3–4), 246–255.
Matthijnssens, J., Otto, P.H., Ciarlet, M., Desselberger, U., Van Ranst, M., Johne, R., 2012. VP6-sequence-based cutoff values as a criterion for rotavirus species demarcation. Arch. Virol. 157 (6), 1177–1182.
J. Matthijnssens, H. Attoui, K. Bányai, C.P.D. Brussaard, P. Pranav Danthi, M. del Vas, T. S. Dermody, R. Duncan, Q. Fang, R. Johne, et al., 2021. Reoviridae. virus taxonomy: 2020 release. International committee on taxonomy of viruses. Available online: https://talk.ictvonline.org/ictv-reports/ictv_online_report/dsrna-viruses/w/reoviridae (accessed on 19 March 2023).
Matthijnssens, J., Attoui, H., Bányai, K., Brussaard, C.P.D., Danthi, P., Del Vas, M., Dermody, T.S., Duncan, R., Fäng, Q., Johne, R., Mertens, P.P.C., Mohd Jaafar, F., Patton, J.T., Sasaya, T., Suzuki, N., Wei, T., 2022. ICTV virus taxonomy profile: sedoreoviridae, 2022 J. Gen. Virol. 103 (10). https://doi.org/10.1099/jgv.0.001782.
Mihalov-Kovács, E., Gellért, Á., Marton, S., Farkas, S.L., Fehér, E., Oldal, M., Jakab, F., Martella, V., Bányai, K., 2015. Candidate new rotavirus species in sheltered dogs. Hungary. Emerg. Inf. Dis. 21 (4), 660–663.
Misu, M., Yoshikawa, T., Sugimoto, S., Takamatsu, Y., Kurosu, T., Ouji, Y., Yoshikawa, M., Shimojima, M., Ebihara, H., Saijo, M., 2023. Rapid whole genome sequencing methods for RNA viruses. Front. Microbiol. 14, 1137086.
Ogden, K.M., Johne, R., Patton, J.T., 2012. Rotavirus RNA polymerases resolve into two phylogenetically distinct classes that differ in their mechanism of template recognition. Virology 431 (1–2), 50–57.
Otto, P.H., Rosenhain, S., Elschner, M.C., Hotzel, H., Machnowska, P., Trojnar, E., Hoffmann, K., Johne, R., 2015. Detection of rotavirus species A, B and C in domestic mammalian animals with diarrhoea and genotyping of bovine species A rotavirus strains. Vet. Microbiol. 179 (3–4), 168–176.
Otto, P.H., Ahmed, M.U., Hotzel, H., Machnowska, P., Reetz, J., Roth, B., Trojnar, E., Johne, R., 2012. Detection of avian rotaviruses of groups A, D, F and G in diseased chickens and turkeys from Europe and Bangladesh. Vet. Microbiol. 156 (1–2), 8–15.
Phan, T.G., Leutenegger, C.M., Chan, R., Delwart, E., 2017. Rotavirus I in feces of a cat with diarrhea. Virus Genes 53 (3), 487–490.
Sasaki, S., Horie, Y., Nakagomi, T., Oseto, M., Nakagomi, O., 2001. Group C rotavirus NSP4 induces diarrhea in neonatal mice. Arch. Virol. 146 (4), 801–806.
Simsek, C., Corman, V.M., Everling, H.U., Lukashev, A.N., Rasche, A., Maganga, G.D., Binger, T., Jansen, D., Beller, L., Deboutte, W., Gloza-Rausch, F., Seebens-Hoyer, A., Yordanov, S., Sylverken, A., Oppong, S., Sarkodie, Y.A., Vallo, P., Leroy, E.M., Bourgarel, M., Yinda, K.C., Matthijnssens, J., 2021. At least seven distinct rotavirus genotype constellations in bats with evidence of reassortment and zoonotic transmissions. MBio 12 (1), e02755, 20.
Souvorov, A., Agarwala, R., Lipman, D.J., 2018. SKESA: strategic k-mer extension for scrupulous assemblies. Genome Biol. 19 (1), 153.

Theuns, S., Vyt, P., Desmarets, L.M.B., Roukaerts, I.D.M., Heylen, E., Zeller, M., Matthijnssens, J., Nauwynck, H.J., 2016. Presence and characterization of pig group A and C rotaviruses in feces of Belgian diarrheic suckling piglets. Virus Res. 213, 172–183.

Troeger, C., Khalil, I.A., Rao, P.C., Cao, S., Blacker, B.F., Ahmed, T., Armah, G., Bines, J. E., Brewer, T.G., Colombara, D.V., Kang, G., Kirkpatrick, B.D., Kirkwood, C.D., Mwenda, J.M., Parashar, U.D., Petri Jr, W.A., Riddle, M.S., Steele, A.D., Thompson, R.L., Walson, J.L., Reiner Jr, R.C., 2018. Rotavirus vaccination and the global burden of rotavirus diarrhea among children younger than 5 years. JAMA Pediatr. 172 (10), 958–965.

Wang, Y., Porter, E.P., Lu, N., Zhu, C., Noll, L.W., Hamill, V., Brown, S.J., Palinski, R.M., Bai, J., 2021. Whole-genome classification of rotavirus C and genetic diversity of porcine strains in the USA. J. Gen. Virol. 102 (5) https://doi.org/10.1099/jgv.0.001598.

Zhao, S., Jin, X., Zang, L., Liu, Z., Wen, X., Ran, X., 2022. Global infection rate of rotavirus C during 1980-2022 and analysis of critical factors in the host range restriction of virus VP4. Viruses 14 (12), 2826.