The characterization of multiple novel paramyxoviruses highlights the diverse nature of the subfamily *Orthoparamyxovirinae*

Bert Vanmechelen,[1],[†] Sien Meurs,[1] Marie Horemans,[1] Arne Loosen,[1] Tibe Joly Maes,[1] Lies Laenen,[1] Valentijn Vergote,[1] Fara Raymond Koundouno,[2] N'Faly Magassouba,[3] Mandy Kader Konde,[4] Ibrahima Sory Condé,[4] Miles W. Carroll,[5] and Piet Maes[1],[*],[‡]

[1]KU Leuven, Department of Microbiology, Immunology and Transplantation, Laboratory of Clinical and Epidemiological Virology, Rega Institute for Medical Research, Herestraat 49/Box 1040, Leuven BE3000, Belgium, [2]Laboratoire des fièvres hémorragiques et virales de Guinée, Conakry Guinea—Direction préfectorale de la santé de Gueckedou, Conakry, Guinea, [3]Projet des Fièvres Hémorragiques en Guinée, Laboratoire de Recherche en Virologie, Conakry, Guinea, [4]Center of Excellence for Training, Research on Malaria and Disease Control in Guinea (CEFORPAG), Quartier Nongo, Ratoma J96C+GFW, Conakry, Guinea and [5]Pandemic Science Centre, Oxford University, Nuffield Department of Medicine, Wellcome Centre for Human Genetics, CCMP1st South, Roosevelt Dr, Headington, Oxford OX3 7BN, UK

[†]https://orcid.org/0000-0002-8350-4356
[‡]https://orcid.org/0000-0002-4571-5232
*Corresponding author: E-mail: piet.maes@kuleuven.be

Abstract

The subfamily *Orthoparamyxovirinae* is a group of single-stranded, negative-sense RNA viruses that contains many human, animal, and zoonotic pathogens. While there are currently only forty-two recognized species in this subfamily, recent research has revealed that much of its diversity remains to be characterized. Using a newly developed nested PCR-based screening assay, we report here the discovery of fifteen orthoparamyxoviruses in rodents and shrews from Belgium and Guinea, thirteen of which are believed to represent new species. Using a combination of nanopore and sanger sequencing, complete genomes could be determined for almost all these viruses, enabling a detailed evaluation of their genome characteristics. While most viruses are thought to belong to the rapidly expanding genus *Jeilongvirus*, we also identify novel members of the genera *Narmovirus*, *Henipavirus,* and *Morbillivirus*. Together with other recently discovered orthoparamyxoviruses, both henipaviruses and the morbillivirus discovered here appear to form distinct rodent-/shrew-borne clades within their respective genera, clustering separately from all currently classified viruses. In the case of the henipaviruses, a comparison of the different members of this clade revealed the presence of a secondary conserved open reading frame, encoding for a transmembrane protein, within the F gene, the biological relevance of which remains to be established. While the characteristics of the viruses described here shed further light on the complex evolutionary origin of paramyxoviruses, they also illustrate that the diversity of this group of viruses in terms of genome organization appears to be much larger than previously assumed.

Key words: Paramyxovirus; Orthoparamyxovirinae; screening; jeilongvirus; henipavirus; morbillivirus.

Introduction

In the past decade, the increased availability and reduced cost of high-throughput sequencing (HTS) platforms have resulted in the discovery of countless novel viruses (Greninger 2018; Zhang et al. 2019). A significant contribution to our understanding of viral diversity has been made by using large-scale metagenomics to expand the scope of putative hosts being studied. However, much of the virosphere also remains unexplored in previously established virus reservoirs, as demonstrated by the continuous discovery of new viruses in well-known virus hosts such as rodents and bats (Gibb et al. 2021). Although the pathological relevance and zoonotic potential of many of these newly discovered viruses is difficult to estimate solely from their genome sequences, valuable information can nonetheless be deduced by inferring their phylogenetic relationship with known viruses (Epstein and Anthony 2017). As illustrated by several recent large viral epidemics and pandemics (e.g. Ebola virus, Zika virus, Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)), the world is ill-prepared for adequately containing and managing virus outbreaks, especially if they are caused by unknown viruses (SARS-CoV-2) or viruses previously only reported in other parts of the world (Ebola virus, Zika virus) (Maxmen 2021). Through an improved understanding of the spread and diversity of animal-borne viruses, the risk of emergence of known or novel zoonoses locally or globally can be estimated more accurately, representing an important first step in improving global viral epidemic preparedness.

One virus family of particular interest for the potential emergence of novel human pathogens is the family *Paramyxoviridae*. Second only to the family *Rhabdoviridae*, paramyxoviruses form one of the largest known families of negative-sense, monopartite RNA viruses (Kuhn et al. 2021). All paramyxoviruses share a similar genome organization, having a 15–21 kb genome that encodes six major genes, although members of the genus *Jeilongvirus* and some other viruses can have one or more additional genes of which the function remains poorly understood (Rima et al. 2019). Currently, the family *Paramyxoviridae* includes eighty-six recognized species, spread across four distinct subfamilies, the largest being the subfamily *Orthoparamyxovirinae* (Kuhn et al. 2021). In addition to several important animal pathogens (e.g. canine distemper virus, fer-de-lance virus, rinderpest virus), this subfamily is also home to human pathogenic viruses such as measles virus, mumps virus and parainfluenza viruses, as well as zoonotic pathogens, like Hendra virus and Nipah virus, that can infect humans and other animals (Thibault et al. 2017; Rima et al. 2019). In recent years, many new orthoparamyxoviruses have been discovered and several of them are assumed to be pathogenic for animals or even humans. Especially the genus *Jeilongvirus* has seen a rapid expansion, with new members being discovered not only primarily in rodents and bats but also in carnivores and eulipotyphla (Noh et al. 2018; Vanmechelen et al. 2018, 2020, 2021; Wu et al. 2018, 2021; Sakaguchi et al. 2020; de Souza et al. 2021; Larsen et al. 2021; Lee et al. 2021). Furthermore, PCR-based screening studies have shown that much of the orthoparamyxovirus diversity remains to be properly characterized (Drexler et al. 2012; Conrardy et al. 2014; Wilkinson et al. 2014; Ghawar et al. 2017; Rizzo et al. 2017; Berto et al. 2018; Bourgarel et al. 2018; Markotter et al. 2019; Mendenhall et al. 2019; Onyuok et al. 2019; Prada et al. 2019; Hoarau et al. 2021; Kohl et al. 2021). These studies are often performed using the screening assays developed by Tong and colleagues, which are capable of detecting paramyxoviruses belonging to a diverse range of genera (Tong et al. 2008).

In this study, we screened ninety-five small mammals from Belgium and Guinea, belonging to a variety of rodent and shrew species, for the presence of paramyxoviruses. For this screening, we developed a custom nested reverse transcription PCR (RT-PCR) that was found to outperform the semi-nested assay of Tong et al. in terms of detectable diversity. Additionally, we attempted to obtain at least one complete genome sequence for each of the different viruses that were detected. In addition to Beilong virus (BeiV) and Pohorje myodes paramyxovirus 1 (PMPV-1), viruses belonging to thirteen novel putative species were detected, clustering within the genera *Jeilongvirus, Morbillivirus, Narmovirus,* and *Henipavirus*. For all but one of these novel viruses, a (near-)complete genome sequence could be determined.

Materials and methods
Sample collection
Cadavers of small mammals (rodents and shrews) that had been killed by cats or vehicles were collected in Belgium and stored at –20°C prior to transport to the laboratory. Following dissection, organs were treated with RNA*later* (Thermo Fisher Scientific, Waltham, MA, USA) at 4°C for 24 h, and subsequently stored at –80°C after the removal of excess RNA*later*. In addition, organs were also collected from rodents and shrews that had been trapped and killed near Guéckédou (Nzérékoré Region, Guinea) in the frame of routine pest control. Identification of the caught animals was done by amplifying and sequencing part of the mitochondrial *cytochrome b* gene using the primer set (CytB Uni fw 5′-TCATCMTGATGAAAYTTYGG-3′, CytB Uni rev 5′-ACTGGYTGDCCBCCRATTCA-3′) developed by Schlegel and colleagues (Schlegel et al. 2012).

RNA extraction
RNA was extracted from kidney tissue with or without viral enrichment procedures. Homogenization was performed by shredding 20 mg tissue in a tube filled with 135-μl sterile PBS (Thermo Fisher Scientific) and five zirconium oxide beads (2.8 mm, Thermo Fisher Scientific), using a Bertin homogenizer for 2 min at 4.000 rpm. For standard extractions, the homogenate was immediately subjected to RNA extraction using the RNeasy mini kit (Qiagen, Venlo, The Netherlands) according to the manufacturer's instructions. For viral enrichment, 20X homemade enzyme buffer (1-M Tris, 100-mM $CaCl_2$, and 30-mM $MgCl_2$, pH 8), 1-μl microccocal nuclease (New England Biolabs, Ipswich, MA, USA) and 2 μl benzonase (Merck-Millipore, Burlington, MA, USA) were added to the homogenate. After a 2-h incubation at 37°C, EDTA was added to a final concentration of 10 mM to stop enzymatic digestion. The homogenate was then centrifuged at 17.000 g for 3 min, after which the resulting supernatant was filtered through a 0.8-μM filter by centrifuging at 17.000 g for 1 min. The resulting filtrate was then subjected to RNA extraction using the Viral RNA mini kit (Qiagen), according to the manufacturer's instructions.

Screening PCR
Nested RT-PCRs were performed using the OneStep RT-PCR kit (Qiagen). In each outer reaction, generating a 251-bp fragment, 2-μl RNA extract was mixed with 10-nmol dNTPs, 5X reaction buffer, 30-pmol forward primer (SCR_14F: 5′-ATGATGAARGGNCATGC-3′), 30-pmol reverse primer (SCR_26R: 5′-GCYTTRTCYTTCATRTACAT-3′), and 1 μl of the supplied enzyme mix, in a total volume of 25 μl. Cycling conditions were as follows: 30 min at 50°C and 15 min at 95°C, followed by 40 cycles of 30 s at 94°C, 30 s at 47°C and 1 min at 72°C. A final elongation step was performed at 72°C for 10 min. For the inner reaction, generating a 235-bp fragment, 2 μl of the outer reaction product was used. All reaction and cycling conditions were identical to the ones used for the outer reaction, with the exception of the used primer set (SCR_16F: 5′-AARGGNCATGCHHTNTTCTG-3′, SCR_25R: 5′-TTCATRTACATDGTNAGRTC-3′) and the omission of the reverse transcription step at 50°C.

Generation of cDNA libraries
RNA extracts from paramyxovirus-positive samples were converted into cDNA libraries using the SISPA method previously described by Greniger et al. (Greninger et al. 2015). Briefly, 4-μl RNA was reverse transcribed by adding 1 μl sol-PrimerA (40 pmol/μl, 5′-GTTTCCCACTGGAGGATA-N_9-3′) and incubating at 65°C for 5 min and at 25°C for another 5 min. Next, 5 μl SuperScript IV master mix was added, consisting of 5× SSIV buffer, 10-nmol dNTPs, 50-nmol DTT, 20U RNaseOUT (Thermo Fisher Scientific) and 100 U SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific). The resulting mix was incubated at 23°C for 10 min, followed by 10 min at 50°C. For second strand synthesis, 5 μl Sequenase mix was added, containing 1-μl Sequenase buffer, 3.85-μl water, and 0.15-μl Sequenase Version 2.0 DNA Polymerase (Thermo Fisher Scientific). Following an 8-min incubation step at 37°C, a secondary Sequenase mix, containing 0.45-μl dilution buffer and 0.15-μl enzyme, was added, and the 8-min incubation at 37°C was repeated. For amplification, 5 μl of the cDNA was mixed with LongAmp Taq 2X Master Mix (New England Biolabs) and 100 pmol Sol-PrimerB (5′-GTTTCCCACTGGAGGATA-3′) in a total volume of 50 μl. Cycling

conditions were as follows: 2 min at 94°C, followed by 30 cycles of 30 s at 94°C, 45 s at 50°C and 1 min at 65°C, and a final 10-min elongation step at 65°C. Clean-up was performed using 0.4X AMPure XP beads (Beckman Coulter, Brea, CA, USA) according to the manufacturer's instructions.

Nanopore sequencing

Preparation of sequencing libraries was done using the SQK-LSK109 kit (Oxford Nanopore Technologies) according to the manufacturer's protocol. When sequencing more than one sample per flow cell, the EXP-NBD104 and EXP-NDB114 kits were used for barcoding of individual samples. For each sample, 200-ng cDNA was used as input and ∼120 fmol library was loaded onto the R.9.4.1 flow cell. Sequencing was performed on a MinION Mk1B or GridION. Basecalling (and demultiplexing) was performed using Guppy v3 and above. Porechop (github.com/rrwick/Porechop) was used to remove sol-Primer sequences added during cDNA generation. To identify reads corresponding to paramyxovirus sequences, a tblastx search of all reads against a custom BLAST database containing a representative sequence of each known orthoparamyxovirus species was done.

Genome completion

If sufficient in number (>100), the subset of paramyxovirus reads was assembled into contigs using Canu v2.0 (Koren et al. 2017). These contigs were polished using Medaka v0.9+ to achieve high quality sequences. To close remaining gaps in the sequence and to confirm the sequence of regions with low coverage (<20X), primer sets were designed based on the Medaka-polished contigs and used for the RT-PCR amplification of regions of interest using the OneStep RT-PCR kit (Qiagen) according to the manufacturer's instructions. When insufficient reads were available, primers were designed based directly on the sequence of the individual nanopore reads. The resulting amplicons were purified using PureIT ExoZap PCR CleanUp (Ampliqon, Odense, Denmark) and sent to Macrogen Europe (Amsterdam, The Netherlands) for Sanger sequencing. The obtained chromatograms were corrected using Chromas v2.7 and, if applicable, joined with the Medaka contigs using Seqman v7.0.0. Genome ends were determined using the Roche 5'/3'race kit, 2nd generation (Hoffmann-La Roche, Basel, Switzerland) as described previously (Vanmechelen et al. 2020).

Phylogenetic analysis

The NCBI Open Reading Frame Finder (ncbi.nlm.nih.gov/orffinder) was used to identify the open reading frames (ORFs) present in the newly sequenced genomes. For each of the six major proteins (N, P, M, F, RBP, and L), separate alignments were made using MAFFT v7.489, incorporating a representative sequence of each (putative) paramyxovirus species for which a (near-)complete genome sequence is publicly available (Katoh and Standley 2013). A list of all used accession numbers can be found in Supplementary Table S1. The resulting alignments were trimmed with trimAl v1.4.rev15, removing all columns with gaps in >10 per cent of all sequences, after which the trimmed alignments were concatenated (Capella-Gutierrez, Silla-Martinez, and Gabaldon 2009). Following model selection using IQ-TREE, phylogenetic trees were inferred using BEAST v1.10.4, employing an LG + G4 + I amino acid substitution model (Le and Gascuel 2008; Nguyen et al. 2015; Suchard et al. 2018). The BeiV tree was made similarly but based on the complete nucleotide sequences of all available BeiV genomes, employing a GTR + G4 + I nucleotide substitution model

(Lanave et al. 1984). Analyses were run until adequate sample sizes were obtained (ESS > 200) and maximum clade credibility trees were summarized from the posterior tree distribution using TreeAnnotator with a burn-in of 10 per cent. FigTree v1.4.3 was used to visualize the resulting trees.

Attempts at virus isolation

Tissue homogenate was made by shredding 30–50 mg kidney tissue in in a tube filled with 200 µl sterile PBS (Thermo Fisher Scientific) and five zirconium oxide beads (2.8 mm, Thermo Fisher Scientific), using a Bertin homogenizer for 2 min at 4.000 rpm. The homogenate was diluted 1/5 in Opti-MEM (Thermo Fisher Scientific) and centrifuged at 17.000 g for 2 min to pellet cell debris. Vero E6 or BHK21J cells, grown to 80 per cent confluency in 6-well plates, were then overlayed with 250 µl diluted homogenate for 90 min. Next, the homogenate was removed and cells were washed once with 1 ml sterile PBS before adding 2-ml DMEM supplemented with 2.5 per cent Fetal bovine serum (Biowest), 200-mM L-glutamine, 1 per cent Penicillin-Streptomycin, 1 per cent Gentamicin and 0.2 per cent Fungizone (all Thermo Fisher Scientific). Cells were inspected daily for cytopathogenic effects and passaged weekly three times. Supernatans from each passage was subjected to RNA extraction using the Viral RNA mini kit (Qiagen) and the resulting RNA was screened for the presence of paramyxovirus using the nested RT-PCR described above.

Results

Host determination and paramyxovirus screening

To investigate the presence of previously undiscovered paramyxoviruses in small mammals, we collected ninety-five rodents and shrews from sampling sites in Guinea ($n = 71$) and Belgium ($n = 24$) (Table 1). Using Sanger sequencing of the mitochondrial *Cytochrome b* gene, eighteen different hosts could be identified. We initially screened the seventy-one Guinean animals for the presence of paramyxoviruses using an in-house developed nested RT-PCR. Twenty-two samples were found to be positive (31 per cent). To validate whether there were any false negatives, we also screened these seventy-one animals using the 'PAR' and 'RES-MOR-HEN' primer sets previously developed by Tong et al. (Tong et al. 2008). However, these assays showed a reduced performance compared to our newly developed primer set, detecting only six and nine positives, respectively. Both assays also failed to detect any additional positive samples (Supplementary Table S2, Supplementary Fig. S1). Therefore, the Belgian samples were only screened using our novel assay. An additional ten positives were found (41.7 per cent), bringing the total to thirty-two positive samples, spread across twelve of the eighteen different hosts. Based on the sequences obtained by Sanger sequencing the screening amplicons, these thirty-two positives were thought to belong to fifteen separate paramyxovirus species. Viruses that were found in multiple animals were only detected in members of the same host species, and only in reddish-white Mastomys (*Mastomys erythroleucus*) and field voles (*Microtus agrestis*) did we detect viruses belonging to more than one species (Wilson and Reeder 2005).

Genome completion

To validate the hypothesis that the differences between the fifteen sequence groups were sufficient to warrant their classification

Table 1. Overview of collected animals.

Host family	Host species	Common host name	Paramyxovirus positivity rate (Positives/total)	Country
Muridae	*Mastomys erythroleucus*	Reddish-white Mastomys	11/42	Guinea
	Mastomys natalensis	Natal Mastomys	2/9	Guinea
	Uranomys ruddi	Rudd's Bristle-furred Rat	4/6	Guinea
	Lophuromys sikapusi	Rusty-bellied Brush-furred Rat	2/3	Guinea
	Praomys rostratus	West African Praomys	2/2	Guinea
	Dasymys rufulus	West African Dasymys	0/2	Guinea
	Hybomys planifrons	Liberian Forest Hybomys	0/1	Guinea
	Hylomyscus alleni	Allen's Hylomyscus	0/1	Guinea
	Mus setulosus	Peters's Mouse	0/1	Guinea
	Mus minutoides	Southern African Pygmy Mouse	0/1	Guinea
	Rattus norvegicus	Brown Rat	1/4	Belgium
	Apodemus sylvaticus	Long-tailed Field Mouse	1/3	Belgium
Soricidae	*Crocidura grandiceps*	Large-headed Shrew	1/3	Guinea
	Crocidura russula	Greater White-toothed Shrew	3/8	Belgium
	Sorex minutus	Eurasian Pygmy Shrew	0/1	Belgium
Cricetidae	*Microtus arvalis*	Common Vole	2/5	Belgium
	Microtus agrestis	Field Vole	2/2	Belgium
	Myodes glareolus	Bank Vole	1/1	Belgium

in separate species, we attempted to determine the complete genome sequence for a representative strain from each group. Nanopore sequencing of randomly amplified cDNA was used to obtain more sequence information for each virus, after which targeted RT-PCRs spanning remaining gaps or poorly covered regions of the genome were used to determine the rest of the genome sequence. Genomes were completed by Sanger sequencing of the genome ends following 5'/3' RACE. Through this approach, (coding-)complete genomes could be obtained for thirteen of the fifteen different viruses. For a fourteenth sequence, the 3' end of the genome and one gap covering part of the G and L ORFs could not be determined due to the low integrity of the available RNA. Lastly, repeated efforts to obtain additional sequence information for the fifteenth virus, which was found in three Guinean reddish-white Mastomys samples, only managed to identify two small fragments of the N and L ORFs, presumable due to a very low viral load in these three samples. An overview of the organization of the different obtained genomes and some of their characteristics are shown in Fig. 1 and Table 2. The naming of the different viruses is based on the genus to which they belong, the host in which they were detected, the location in which these hosts were found or a combination of these characteristics.

Excluding the three sequences for which part of the genome sequence remains to be determined, the genomes sequenced here are all compliant with the 'rule-of-six', which states that paramyxovirus genomes must have a length that is an exact multiple of six to enable genome replication (Kolakofsky et al. 2005). However, their exact lengths vary greatly, representing both the smallest (14,994 nucleotides) and largest (20,454 nucleotides) known orthoparamyxoviruses. The diverse nature of the genomes characterized in this study is further illustrated by their scattered phylogenetic clustering throughout the entire subfamily *Orthoparamyxovirinae*, with the exception of a single clade that contains the genus *Respirovirus* and the non-mammalian genera *Ferlavirus* and *Aquaparamyxovirus* (Fig. 2). While most viruses we found seem to belong to the genus *Jeilongvirus*, we also detected three novel narmoviruses, two henipaviruses and one morbillivirus. Unfortunately, attempts to rescue infectious virus from animal tissue using Vero E6 and BHK21J cells were unsuccessful.

Characterization of novel Jeilongviruses

Nine of the fifteen paramyxoviruses detected here are thought to belong to the genus *Jeilongvirus*. This includes memaer virus, for which insufficient sequence data could be obtained to do a meaningful phylogenetic analysis. Based on blastn comparisons, this virus appears to be most similar to the here reported memana virus, to ruloma virus and to Longquan Niviventer niviventer jeilongvirus 1 (Table 3), although sufficiently dissimilar, especially given the conserved nature of the detected regions, to consider memaer virus part of a separate viral species. Conversely, two of the jeilongviruses detected in this study, Pohorje myodes paramyxovirus 1 (PMPV-1) strain BE/Ninove/Mg/1/2019 and Beilong virus (BeiV) strain BE/Ninove/Rn/1/2019, appear to fall within the already recognized species *Myodes jeilongvirus* and *Beilong jeilongvirus*, respectively. While the latter is intriguingly similar to all known BeiV genomes, phylogenetically clustering amidst them and sharing 94.50–96.14 per cent nucleotide identity (Fig. 3), the PMPV-1 sequence is more dissimilar from the only other known sequence of PMPV-1, sharing only 84.25 per cent identity on nucleotide level. Interestingly, the variation between these two sequences appears to be roughly equal throughout the coding regions of the genome, with a notable exception being the C-terminal domain of the G gene, which codes for the receptor-binding protein (RBP; Fig. 4). While the first half and the last quarter of this protein display a comparable degree of conservation as the rest of the genome, the third quarter bears little resemblance to the other available PMPV-1 sequence and appears to be even less conserved than the non-coding regions of the genome. A similar pattern, albeit less pronounced, can also be observed when comparing the BeiV sequence presented here with all known BeiV genomes.

Previously, we showed that the glycoproteins of jeilongviruses share certain unique features not found in those of other paramyxoviruses (Vanmechelen et al. 2018). In addition to their exceptional and highly variable length, the C-terminal regions of jeilongvirus RBPs are also characterized by a remarkably high fraction of proline, serine, and threonine, despite their lack of sequence conservation at the nucleotide level. However, since our initial analysis, many new jeilongviruses have been discovered.
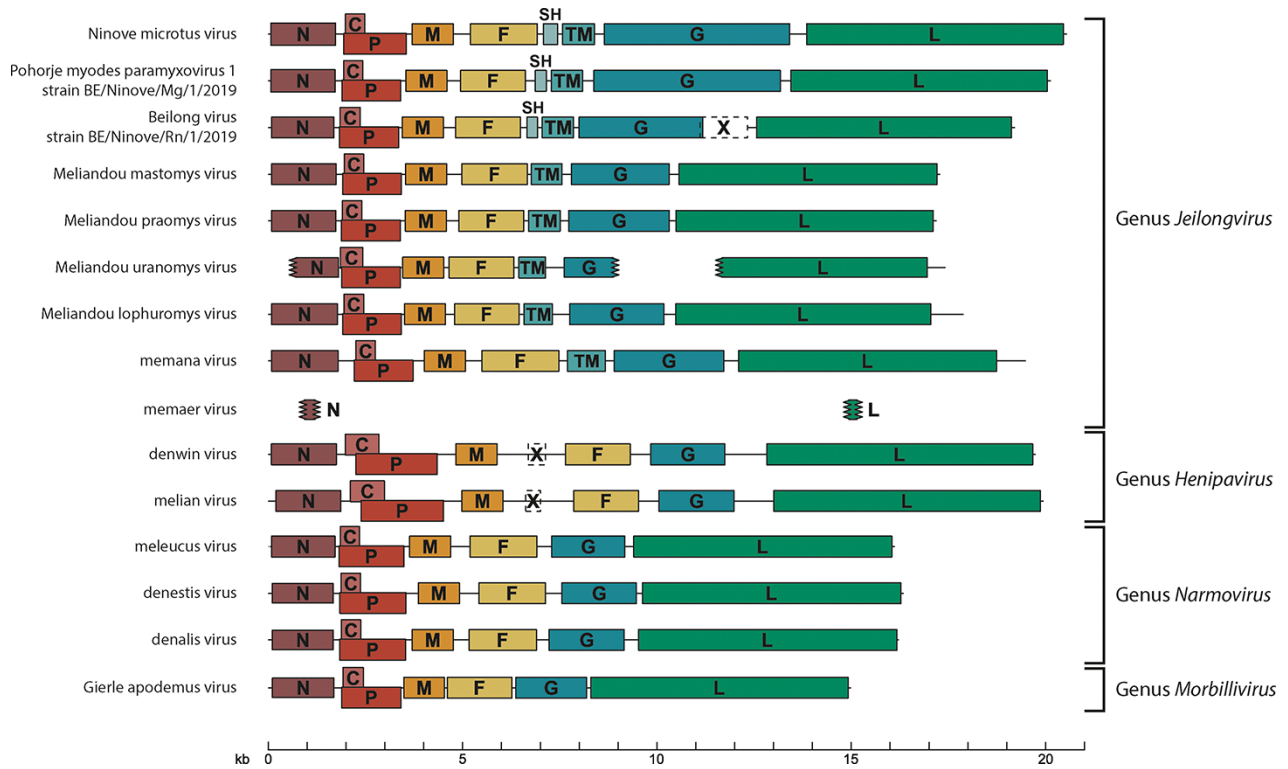
Figure 1. Genome organization of novel paramyxoviruses. The fifteen viruses described in this study have differing genome layouts, which seem to correlate with their presumptive classification in four separate genera. All genomes and ORFs they encode are drawn according to scale. 'X' ORFs indicate ORFs of significant length (>100 amino acids) of which it is unclear if they are biologically active.

Table 2. Overview of detected paramyxoviruses.

| Genbank accession	Virus	Abbreviation	Host species	Common host name	Country	Complete (Y/N)	Putative genus
OK623353	melian virus	MeliV	*Crocidura grandiceps*	Large-headed Shrew	Guinea	Y	*Henipavirus*
OK623354	denwin virus	DewV	*Crocidura russula*	Greater White-toothed Shrew	Belgium	Y	*Henipavirus*
OK623355	Ninove microtus virus	NiMiV	*Microtus agrestis*	Field Vole	Belgium	Y	*Jeilongvirus*
OK623356	Gierle apodemus virus	GapV	*Apodemus sylvaticus*	Long-tailed Field Mouse	Belgium	Y	*Morbillivirus*
OK623357	Pohorje myodes paramyxovirus 1 strain BE/Ninove/Mg/1/2019	PMPV-1	*Myodes glareolus*	Bank Vole	Belgium	Y	*Jeilongvirus*
OK623358	Beilong virus strain BE/Ninove/Rn/1/2019	BeiV	*Rattus norvegicus*	Brown Rat	Belgium	Y	*Jeilongvirus*
OK623359	denestis virus	DeneV	*Microtus agrestis*	Field Vole	Belgium	Y	*Narmovirus*
OK623360	denalis virus	DenaV	*Microtus arvalis*	Common Vole	Belgium	Y	*Narmovirus*
OK623361	Meliandou mastomys virus	MeMV	*Mastomys erythroleucus*	Reddish-white Mastomys	Guinea	Y	*Jeilongvirus*
OK623362	Meliandou praomys virus	MePV	*Praomys rostratus*	West African Praomys	Guinea	Y	*Jeilongvirus*
OK623363	memana virus	MemaV	*Mastomys natalensis*	Natal Mastomys	Guinea	Y	*Jeilongvirus*
OK623364	Meliandou lophuromys virus	MeLoV	*Lophuromys sikapusi*	Rusty-bellied Brush-furred Rat	Guinea	Y	*Jeilongvirus*
OK623365	meleucus virus	meleV	*Mastomys erythroleucus*	Reddish-white Mastomys	Guinea	N[a]	*Narmovirus*
OK623366	Meliandou uranomys virus	MeUV	*Uranomys ruddi*	Rudd's Bristle-furred Rat	Guinea	N	*Jeilongvirus*
OK623367—OK623368	memaer virus	MmrV	*Mastomys erythroleucus*	Reddish-white Mastomys	Guinea	N	*Jeilongvirus*

[a]Coding-complete.
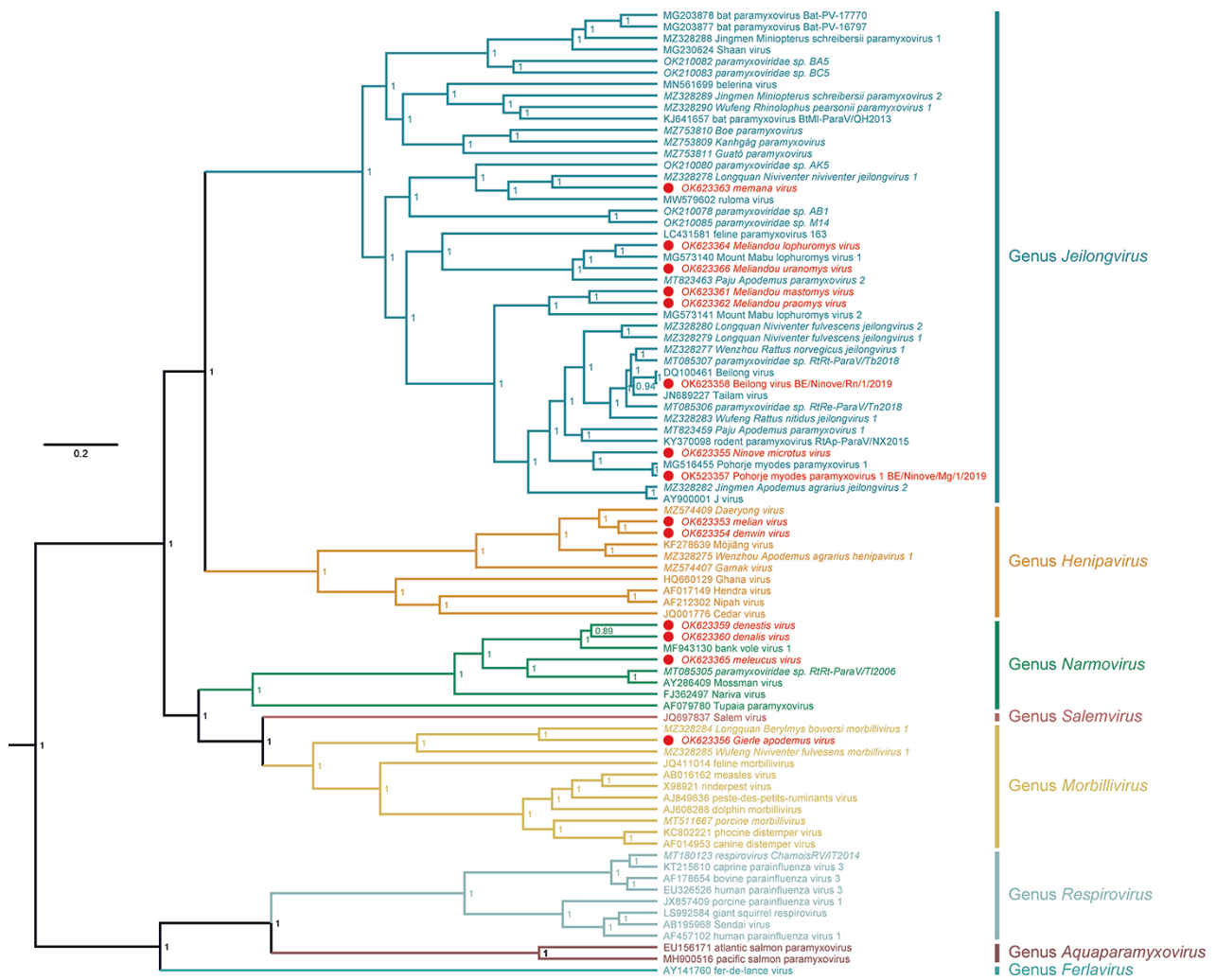
Figure 2. The subfamily *Orthoparamyxovirinae*. Bayesian phylogenetic tree of all eighty-four currently known members of the subfamily *Orthoparamyxovirinae* for which (near-)complete genomes are available. The tree was made based on a concatenation of individual alignments of the N, P, M, F, G, and L ORFs, covering 4,422 sites after trimming. Missing sites in partial alignments were marked as such ('?') prior to phylogenetic inference. Different genera within the subfamily are marked in different colors, according to the current ICTV classification. Viruses thought to belong to yet to be established species have been marked in italics. New sequences presented here are marked in red and are preceded by a red dot. Values at the nodes indicate posterior support. Phylogenetic tree drawn according to scale, with the scale bar denoting the number of substitutions/site. Used accession numbers are shown in the tree. Memaer virus was not included due to a lack of available sequence information.

Table 3. Blastn comparison of memaer virus.

Gene fragment	Virus	Accession number	Max score	Total score	Query cover	E value	% Identity
N (OK623367)	memana virus	OK623363	235	235	99%	6e-65	78.57%
	Longquan Niviventer niviventer jeilongvirus 1	MZ328278	191	191	91%	6e-44	76.54%
	ruloma virus	MW579602	159	159	97%	3e-34	72.92%
L (OK623368)	memana virus	OK623363	204	204	98%	3e-55	80.93%
	ruloma virus	MW579602	154	154	98%	3e-33	75.81%
	Longquan Niviventer niviventer jeilongvirus 1	MZ328278	95.1	95.1	98%	3e-15	70.37%

Redoing the analysis discussed in (Vanmechelen et al. 2018), using all currently available jeilongvirus G genes, shows that the genus *Jeilongvirus* can be divided into five clades based on the length and sequence composition of the RBP (Fig. 5). Members of clade I show no elongation and have RBPs that are comparable to those of other orthoparamyxoviruses. Members of clades II, III, and IV show a modest increase in size and their extension is often but not necessarily characterized by P/T/S-enrichment. Lastly, members of clade V have exceptionally large RBPs, 2–3 times the average length of an orthoparamyxovirus RBP, and their proteins are all characterized by P/T/S-enrichment. Interestingly, several members of this clade have acquired stop codons in their G genes, resulting in the split of the original G ORF into two smaller ORFs, 'G' and 'X'. It remains to be determined if these 'X' ORFs, which fall within the G gene, are expressed, either independently or as part of the RBP. Additionally, while many seem to follow this same pattern, P/T/S-enrichment appears less ubiquitous than previously assumed (Fig. 5). Interestingly, the peak in P/T/S-enrichment appears to
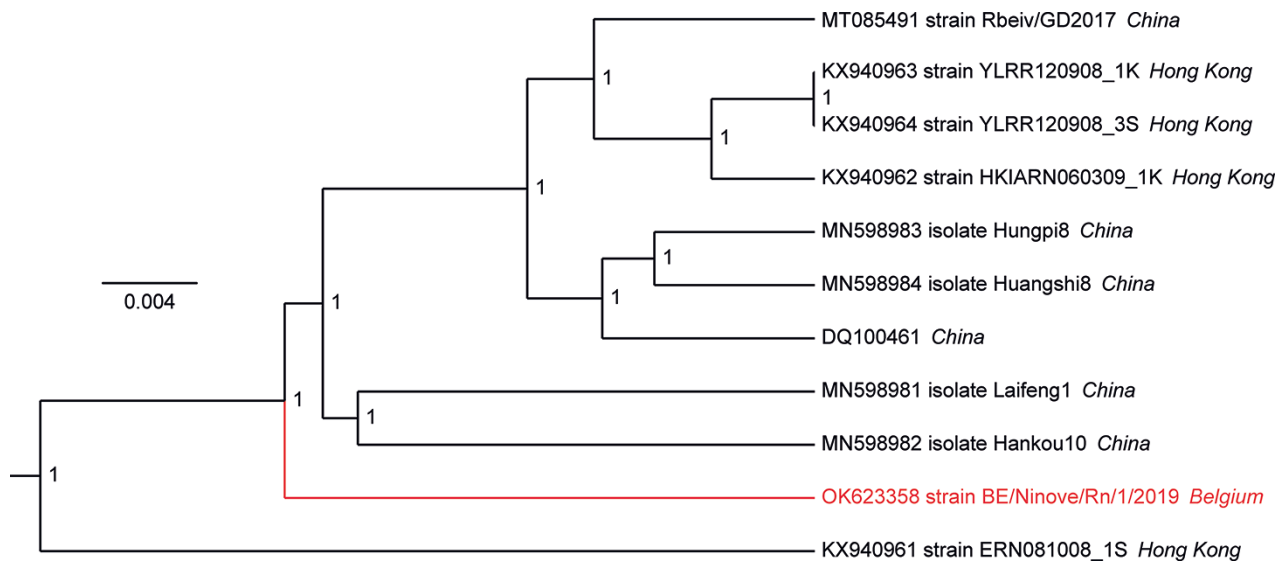
Figure 3. Phylogenetic tree of all (near-)complete BeiV genomes. Bayesian phylogenetic tree based on a nucleotide alignment of all available BeiV genome sequences. Despite its different geographical origin, the strain identified here (from Belgium, marked in red) clusters among the other known BeiV sequences from China and Hong Kong. Values at the nodes indicate posterior support. Phylogenetic tree drawn according to scale, with the scale bar denoting the number of substitutions/site. Used accession numbers and countries of origin are shown in the tree.

coincide with the least conserved region at the nucleotide level, as illustrated by comparing Figs 4 and 5 for the novel strains of PMPV-1 and BeiV reported here.

A new hypothetical protein characterizes a separate henipavirus clade

While most novel viruses described here were found in rodents, we also detected two new viruses in shrews: denwin virus in a greater white-toothed shrew (*Crocidura russula*) and melian virus in a large-headed forest shrew (*Crocidure grandiceps*) (Wilson and Reeder 2005). Both these viruses cluster within the genus *Henipavirus*, together with the recently discovered Gamak virus and Daeryong virus, who were also detected in shrews, and the rodent-borne Mòjiāng virus and Wenzhou Apodemus agrarius henipavirus 1 (Fig. 2). Together, these viruses seem to form a distinct lineage within the genus *Henipavirus*, separate from all bat-borne henipaviruses. Interestingly, all the viruses within this clade, including denwin virus and melian virus, contain a small ORF ('X') between the M and F ORFs that is not seen in the genomes of other paramyxoviruses. This ORF seems to be conserved both at the nucleotide and amino acid level, coding for a small transmembrane protein (Fig. 6). However, this ORF does not have its own gene start and stop signals but instead falls within the boundaries of the F gene. Whether this ORF is expressed as part of a multicistronic mRNA or through another mechanism remains to be established.

Discovery of other paramyxoviruses

Besides jeilongviruses and henipaviruses, we also discovered three new narmoviruses and one morbillivirus. Unlike the abovementioned viruses, these four viruses have a more traditional paramyxovirus genome organization, characterized by only six genes (N-P/V/C-M-F-G-L) (Fig. 1). In accordance with the hosts in which they were found, denestis virus (*Microtus agrestis*) and denalis virus (*Microtus arvalis* (Common vole)) cluster closely together, near Bank vole virus 1, while meleucus virus is more closely related to Mossman virus and the yet unnamed

'Paramyxoviridae sp. isolate RtRt-ParaV/Tl2006' (Fig. 2) (Wilson and Reeder 2005). Similar to the abovementioned situation in the genus *Henipavirus*, the clustering of apodemus morbillivirus, alongside the recently discovered Wufeng Niviventer fulvescens morbillivirus 1 and Longquan Berylmys bowersi morbillivirus 1, appears to indicate the existence of a distinct rodent-borne morbillivirus clade, forming a distant sister clade to all non-rodent morbilliviruses.

Discussion

Many novel orthoparamyxoviruses have been discovered in recent years, especially thanks to the increased availability of HTS platforms However, despite the advancements that have been made, the cost per sample remains too high to allow the unfiltered sequencing of large sample collections (Epstein and Anthony 2017). In studies focusing on one or more particular pathogens, targeted methods that allow the selection of a specific subset of samples therefore remain of interest to maximize the return on investment of HTS experiments. Conversely, too stringent selection methods might result in interesting samples erroneously being excluded from further analysis. Therefore, we sought to make a PCR-based screening approach that would maximize the number and diversity of orthoparamyxoviruses that could be detected using a singular assay. By opting for a nested PCR approach that targets a small region of the L gene using four degenerate primers, we were able to detect a wide range of orthoparamyxoviruses, covering each major clade within the subfamily besides the one containing the genus *Respirovirus* and related non-mammalian viruses. Using our new assay, we found >33 per cent of the animals in this study to be positive for paramyxoviruses. Even though this study had a limited sample size and the possibility of false negatives due to insufficient viral load or a too divergent sequence cannot be excluded, this percentage seems to fall in line with what has previously been reported in paramyxovirus screening studies of small mammals (Sasaki et al. 2014; Wilkinson et al. 2014; Larsen et al. 2021). However, few previous studies have reported a similarly broad diversity as
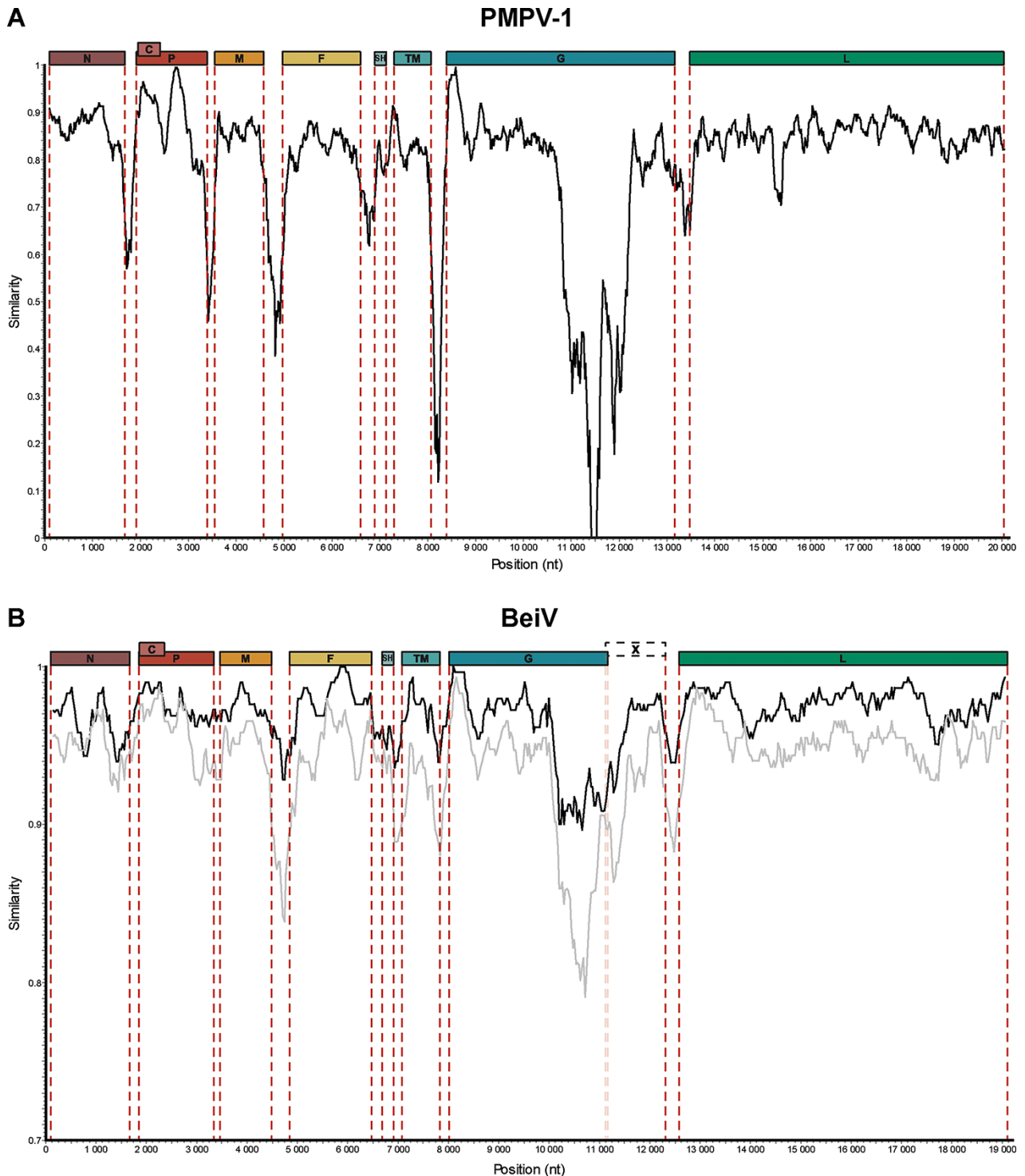
Figure 4. Similarity comparison of the new PMPV-1 and BeiV strains. Panel A shows a comparison of the two known PMPV-1 strains, highlighting a comparable degree of conservation all throughout the coding regions of the genome, with a notable exception in the middle of the G gene. Panel B shows a comparable but significantly weaker pattern when comparing the new BeiV strain described here with all (black line) known BeiV genomes or only the most divergent one (KX940961; grey line). Gene borders are marked by dashed red lines, with faint lines indicating the transition from the 'G' to the 'X' ORF within the G gene of BeiV. Figure made using SimPlot v3.5.1, employing a window size of 200 (A)/300 (B) bp and a step size of 20 (A)/30 (B) bp.

observed here, which might be attributable to the new assay used in this study, as most previous studies were performed using the semi-nested assays developed by Tong and colleagues. The design of their primers is based on the principle of CODEHOP, in which a consensus 5′-sequence is linked to a degenerate 3′-end. The 3′-end maximizes the diversity of sequences that can be detected while the 5′-region ensures a more efficient PCR amplification. However, despite the perfect matching of the 3′-degenerate cores of these primers with the new sequences described here, the sensitivity of these assays appeared insufficient to detect most of the
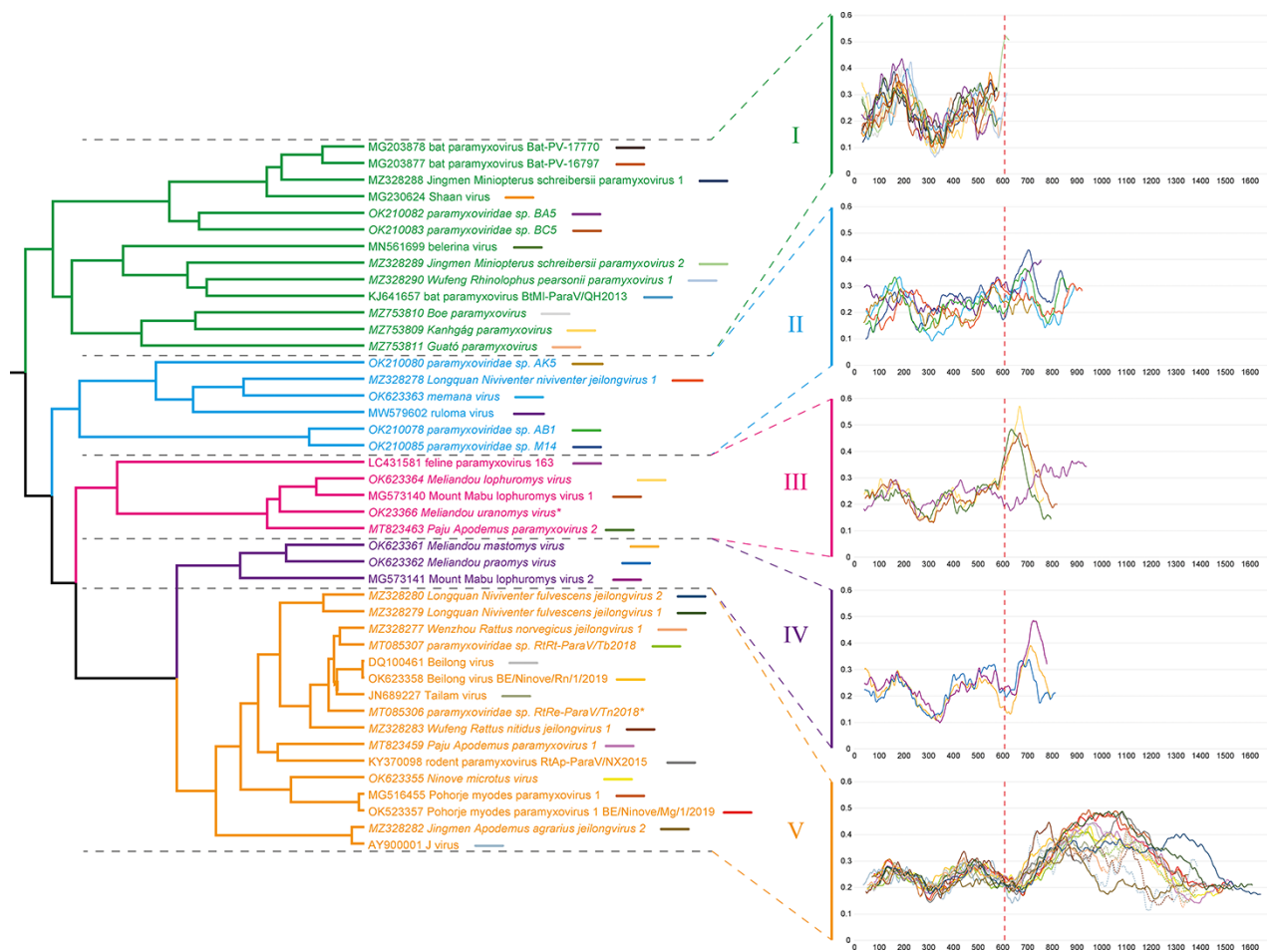
Figure 5. Different RBP organizations in the genus *Jeilongvirus*. Close-up of the tree shown in Fig. 2, depicting only the genus *Jeilongvirus*. Next to the tree are graphs indicating the length of the different jeilongvirus RBPs and their cumulative fraction of proline, serine and threonine, calculated using a sliding window corresponding to 10 per cent of the length of each protein. Based on their phylogenetic relationship and the organization of their RBP, jeilongviruses can be divided into five clades. Members of clade I have RBPs that are comparable in length to those of non-jeilongvirus orthoparamyxoviruses (average length: 607 aa, dashed red line). Conversely, the RBPs of members of clade V are up to 150 per cent larger and their C-terminal extensions are characterized by a very high fraction of proline, serine and threonine. In some members of this clade, the 'G' ORF has acquired a stop codon somewhere within this region, resulting in a secondary 'X' ORF within the G gene. It is unclear if this 'X' ORF is expressed, either by itself or as part of the RBP (dashed lines). Clades II, III and IV are characterized by a more modest expansion of the RBP (up to 50 per cent) that is in some cases characterized by proline-, serine- and threonine-enrichment. Sequences marked by a '*' are not included in the graphs because their G genes have not been fully sequenced. Viruses thought to belong to yet to be established species have been marked in italics.

paramyxoviruses that could be detected in our dataset using our novel assay. This seems to suggest that previous studies might have underestimated the true diversity of paramyxoviruses and that more sensitive assays are needed to better characterize the prevalence and diversity of these viruses.

Nine of the fifteen viruses we detected belong to the genus *Jeilongvirus*. Prior to the establishment of this genus in 2019, only six putative jeilongviruses genome sequences were known, although some studies had already hinted at the existence of many more viruses remaining to be characterized (Drexler et al. 2012; Wu et al. 2016; Amarasinghe et al. 2019). However, since 2019, this genus has known the strongest expansion of all paramyxovirus genera. The latest taxonomy release of the International Committee for the Taxonomy of Viruses (ICTV) already recognizes fifteen species in this genus, and many viruses belonging to additional putative species have been described and/or published on GenBank in the last 3 years and are awaiting classification (International Committee on Taxonomy of Viruses (ICTV) 2022). Consequently, jeilongviruses now represent half of all

known orthoparamyxoviruses (Figs 2, 5). Interestingly, this expansion has also revealed a remarkable diversity in genome organization within members of this genus. Aside from some minor exceptions, paramyxovirus genomes contain six genes (N-P/V/C-M-F-G-L) and members of the same genus typically have comparable genome lengths (Rima et al. 2019). Jeilongvirus genomes, however, contain one or two additional genes and can vary in length from <16 kb to >20.5 kb. The first of these extra genes is commonly named 'TM' because it encodes a protein with a transmembrane region. A protein like this is present in the genome of all jeilongviruses and, in the case of J virus, it has been shown to interact with the F protein and the RBP, and to promote cell-to-cell fusion (Li et al. 2015). However, even though all Jeilongviruses encode a TM protein, it remains to be established whether these proteins exert similar functions, as the sequences of TM proteins from different clades appear to be phylogenetically unrelated, differing strongly in terms of length and sequence composition. As noted previously (see (Vanmechelen et al. 2018)), a similar observation can be made for the second additional jeilongvirus
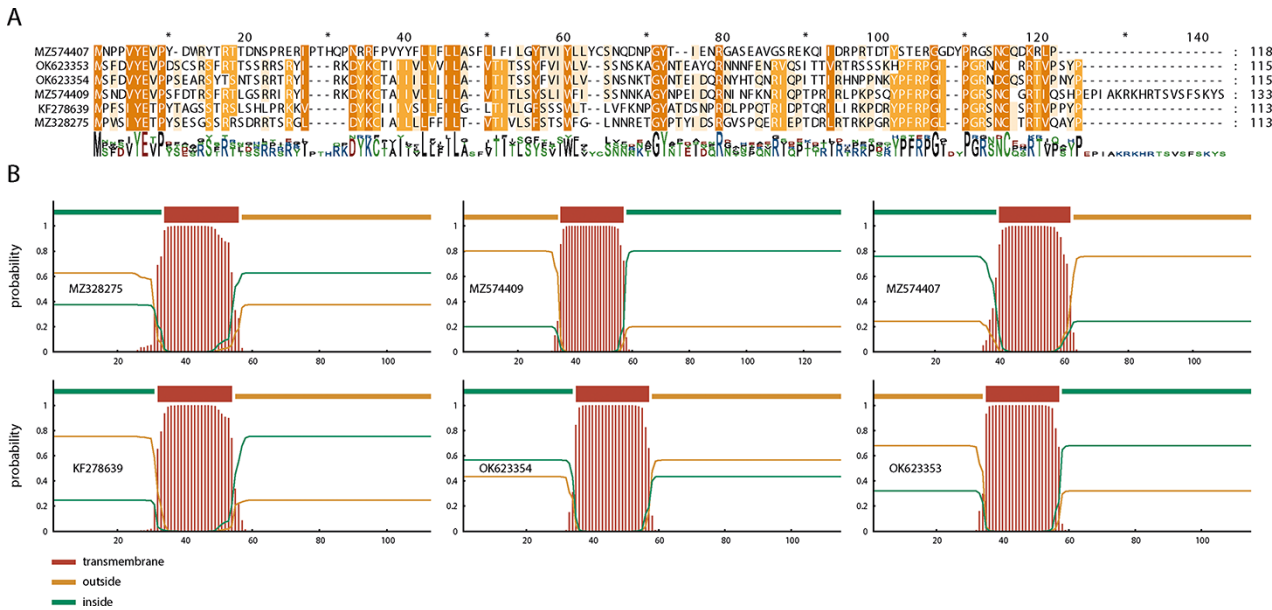
Figure 6. Conservation of a secondary ORF within the henipavirus F gene. (A) All members of the rodent-/shrew-borne clade of henipaviruses have a secondary ORF in the F gene, preceding the actual F ORF. Both the length and sequence composition of this 'X' ORF show a certain degree of conservation, hinting at the biological relevance of this ORF. (B) All these 'X' ORFs are predicted to encode a transmembrane protein. Plots made using TMHMM—2.0 (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0).

gene, SH, which is present only in members of certain clades of the genus. Future research will be needed to determine whether the TM and SH proteins of different clades are functional homologues of each other and to determine the precise function of these proteins.

Like the genus *Jeilongvirus*, the genus *Narmovirus* was only established in 2019 and little is known about the pathological relevance of these viruses for animal and human health. Conversely, the genera *Morbillivirus* and *Henipavirus* were established many years ago and many of their members are known to be important human, animal or zoonotic pathogens (Pfeffermann et al. 2018; Weatherman, Feldmann, and de Wit 2018). The new viruses described here, in combination with other novel genomes published by other groups in the last year, illustrate the existence of a separate clade of rodent- and/or shrew-borne viruses that diverged early on from the rest of the genus within both the genus *Henipavirus* and the genus *Morbillivirus*. A question that remains to be resolved is whether the members of these separate clades are equally pathogenic as their sister clade counterparts and, especially in the case of the henipaviruses, if they have the same zoonotic potential. A possible argument against this hypothesis might be that Mòjiāng virus, the first rodent-borne henipavirus to be discovered, has been shown to be incapable of using the same host-entry pathways as other henipaviruses due to differences in its RBP (Rissanen et al. 2017; Cheliout Da Silva et al. 2021). Interestingly, the availability of additional sequences of related viruses has made it apparent that members of this rodent-/shrew-borne henipavirus clade contain a conserved ORF coding for a transmembrane protein nested within the F gene. Although the mechanism through which this ORF is expressed remains to be established, as well as the function of the resulting protein, its existence highlights that more complex genome organizations are not a unique feature of jeilongviruses but seem to be the result of independent host-specific adaptation events that have occurred in multiple lineages within the subfamily *Orthoparamyxovirinae*.

It will be of interest to determine whether this transmembrane protein is a functional homologue to any of the transmembrane proteins found in jeilongviruses and if similar proteins can also be found in other paramyxovirus lineages that remain to be discovered.

Data availability

The genome sequences (accession numbers OK623353-OK623368) and cytochrome B gene sequences (accession numbers ON419583-ON419677) generated in this study have been submitted to GenBank.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Acknowledgements

Conflict of interest: The authors declare no competing interests.

Author contributions

B.V. and P.M. designed the study. B.V., S.M., M.H., A.L., and T.J.M. performed the experimental work. L.L., V.V., M.W.C., F.R.K., I.S.C., M.K.K., and P.M. ensured sample collection. M.W.C. and P.M. provided reagents and materials. B.V. wrote the main manuscript text and prepared the figures. All authors reviewed the manuscript and approved the final version.

References

Amarasinghe, G. K. et al. (2019) 'Taxonomy of the Order Mononegavirales: Update 2019', *Archives of Virology*, 164: 1967–80.

Berto, A. et al. (2018) 'Detection of Potentially Novel Paramyxovirus and Coronavirus Viral RNA in Bats and Rats in the Mekong Delta Region of Southern Viet Nam', *Zoonoses and Public Health*, 65: 30–42.

Bourgarel, M. et al. (2018) 'Circulation of Alphacoronavirus, Betacoronavirus and Paramyxovirus in Hipposideros Bat Species in Zimbabwe', *Infection, Genetics and Evolution*, 58: 253–7.

Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. (2009) 'trimAl: A Tool for Automated Alignment Trimming in Large-scale Phylogenetic Analyses', *Bioinformatics*, 25: 1972–3.

Cheliout Da Silva, S. et al. (2021) 'Functional Analysis of the Fusion and Attachment Glycoproteins of Mojiang Henipavirus', *Viruses*, 13: 517.

Conrardy, C. et al. (2014) 'Molecular Detection of Adenoviruses, Rhabdoviruses, and Paramyxoviruses in Bats from Kenya', *The American Journal of Tropical Medicine and Hygiene*, 91: 258–66.

de Souza, W. M. et al. (2021) 'Paramyxoviruses from Neotropical Bats Suggest a Novel Genus and Nephrotropism', *Infection, Genetics and Evolution*, 95: 105041.

Drexler, J. F. et al. (2012) 'Bats Host Major Mammalian Paramyxoviruses', *Nature Communications*, 3: 796.

Epstein, J. H., and Anthony, S. J. (2017) 'Viral Discovery as a Tool for Pandemic Preparedness', *Revue Scientifique Et Technique de l'OIE*, 36: 499–512.

Ghawar, W. et al. (2017) 'Insight into the Global Evolution of Rodentia Associated Morbilli-related Paramyxoviruses', *Scientific Reports*, 7: 1974.

Gibb, R. et al. (2021) 'Mammal Virus Diversity Estimates are Unstable Due to Accelerating Discovery Effort', *bioRxiv*. 2021.2008.2010.455791.

Greninger, A. L. (2018) 'A Decade of RNA Virus Metagenomics Is (Not) Enough', *Virus Research*, 244: 218–29.

——— et al. (2015) 'Rapid Metagenomic Identification of Viral Pathogens in Clinical Samples by Real-time Nanopore Sequencing Analysis', *Genome Medicine*, 7: 99.

Hoarau, A. O. G. et al. (2021) 'Investigation of Astrovirus, Coronavirus and Paramyxovirus Co-infections in Bats in the Western Indian Ocean', *Virology Journal*, 18: 205.

International Committee on Taxonomy of Viruses (ICTV). (2022), *Virus Taxonomy: 2021 Release*. <https://talk.ictvonline.org/taxonomy/> accessed 05 May 2022.

Katoh, K., and Standley, D. M. (2013) 'MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability', *Molecular Biology and Evolution*, 30: 772–80.

Kohl, C. et al. (2021) 'The Virome of German Bats: Comparing Virus Discovery Approaches', *Scientific Reports*, 11: 7430.

Kolakofsky, D. et al. (2005) 'Paramyxovirus mRNA Editing, the "rule of six" and Error Catastrophe: A Hypothesis', *Journal of General Virology*, 86: 1869–77.

Koren, S. et al. (2017) 'Canu: Scalable and Accurate Long-read Assembly via Adaptive K-mer Weighting and Repeat Separation', *Genome Research*, 27: 722–36.

Kuhn, J. H. et al. (2021) '2021 Taxonomic Update of Phylum Negarnaviricota (Riboviria: Orthornavirae), Including the Large Orders Bunyavirales and Mononegavirales', *Archives of Virology*, 166: 3513–3566.

Lanave, C. et al. (1984) 'A New Method for Calculating Evolutionary Substitution Rates', *Journal of Molecular Evolution*, 20: 86–93.

Larsen, B. B. et al. (2021) 'Evolution and Diversity of Bat and Rodent Paramyxoviruses from North America', *Journal of Virology*, JVI0109821.

Le, S. Q., and Gascuel, O. (2008) 'An Improved General Amino Acid Replacement Matrix', *Molecular Biology and Evolution*, 25: 1307–20.

Lee, S. H. et al. (2021) 'Novel Paju Apodemus Paramyxovirus 1 and 2, Harbored by Apodemus Agrarius in the Republic of Korea', *Virology*, 562: 40–9.

Li, Z. et al. (2015) 'Type II Integral Membrane Protein, TM of J Paramyxovirus Promotes Cell-to-cell Fusion', *Proceedings of the National Academy of Sciences of the United States of America*, 112: 12504–9.

Markotter, W. et al. (2019) 'Paramyxo- and Coronaviruses in Rwandan Bats', *Tropical Medicine and Infectious Disease*, 4: 99.

Maxmen, A. (2021) 'Has COVID Taught Us Anything about Pandemic Preparedness?', *Nature*, 596: 332–5.

Mendenhall, I. H. et al. (2019) 'Diversity and Evolution of Viral Pathogen Community in Cave Nectar Bats (Eonycteris Spelaea)', *Viruses*, 11: 250.

Nguyen, L. T. et al. (2015) 'IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-likelihood Phylogenies', *Molecular Biology and Evolution*, 32: 268–74.

Noh, J. Y. et al. (2018) 'Isolation and Characterization of Novel Bat Paramyxovirus B16-40 Potentially Belonging to the Proposed Genus Shaanvirus', *Scientific Reports*, 8: 12533.

Onyuok, S. O. et al. (2019) 'Molecular Detection and Genetic Characterization of Novel RNA Viruses in Wild and Synanthropic Rodents and Shrews in Kenya', *Frontiers in Microbiology*, 10: 2696.

Pfeffermann, K. et al. (2018) 'Morbillivirus Pathogenesis and Virus-Host Interactions', *Advances in Virus Research*, 100: 75–98.

Prada, D. et al. (2019) 'Viral Diversity of Microbats within the South West Botanical Province of Western Australia', *Viruses*, 11: 1157.

Rima, B. et al. (2019) 'ICTV Virus Taxonomy Profile: Paramyxoviridae', *Journal of General Virology*, 100: 1593–4.

Rissanen, I. et al. (2017) 'Idiosyncratic Mojiang Virus Attachment Glycoprotein Directs a Host-cell Entry Pathway Distinct from Genetically Related Henipaviruses', *Nature Communications*, 8: 16060.

Rizzo, F. et al. (2017) 'Coronavirus and Paramyxovirus in Bats from Northwest Italy', *BMC Veterinary Research*, 13: 396.

Sakaguchi, S. et al. (2020) 'Molecular Characterization of Feline Paramyxovirus in Japanese Cat Populations', *Archives of Virology*, 165: 413–8.

Sasaki, M. et al. (2014) 'Molecular Epidemiology of Paramyxoviruses in Zambian Wild Rodents and Shrews', *Journal of General Virology*, 95: 325–30.

Schlegel, M. et al. (2012) 'Molecular Identification of Small Mammal Species Using Novel Cytochrome B Gene-derived Degenerated Primers', *Biochemical Genetics*, 50: 440–7.

Suchard, M. A. et al. (2018) 'Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10', *Virus Evolution*, 4: vey016.

Thibault, P. A. et al. (2017) 'Zoonotic Potential of Emerging Paramyxoviruses: Knowns and Unknowns', *Advances in Virus Research*, 98: 1–55.

Tong, S. et al. (2008) 'Sensitive and Broadly Reactive Reverse transcription-PCR Assays to Detect Novel Paramyxoviruses', *Journal of Clinical Microbiology*, 46: 2652–8.

Vanmechelen, B. et al. (2018) 'Discovery and Genome Characterization of Three New Jeilongviruses, a Lineage of Paramyxoviruses Characterized by Their Unique Membrane Proteins', *BMC Genomics*, 19: 617.

——— et al. (2021) 'Genome Sequence of Ruloma Virus, a Novel Paramyxovirus Clustering Basally to Members of the Genus Jeilongvirus', *Microbiology Resource Announcements*, 10: e00325-21.

———— et al. (2020) 'Common Occurrence of Belerina Virus, a Novel Paramyxovirus Found in Belgian Hedgehogs', *Scientific Reports*, 10: 19341.

Weatherman, S., Feldmann, H., and de Wit, E. (2018) 'Transmission of Henipaviruses', *Current Opinion in Virology*, 28: 7–11.

Wilkinson, D. A. et al. (2014) 'Highly Diverse Morbillivirus-related Paramyxoviruses in Wild Fauna of the Southwestern Indian Ocean Islands: Evidence of Exchange between Introduced and Endemic Small Mammals', *Journal of Virology*, 88: 8268–77.

Wilson, D. E., and Reeder, D. A. M. (2005) *Mammal Species of the World: A Taxonomic and Geographic Reference*. Johns Hopkins University Press, Baltimore.

Wu, Z. et al. (2021) 'Decoding the RNA Viromes in Rodent Lungs Provides New Insight into the Origin and Evolutionary Patterns of Rodent-borne Pathogens in Mainland Southeast Asia', *Microbiome*, 9: 18.

———— et al. (2018) 'Comparative Analysis of Rodent and Small Mammal Viromes to Better Understand the Wildlife Origin of Emerging Infectious Diseases', *Microbiome*, 6: 178.

———— et al. (2016) 'Deciphering the Bat Virome Catalog to Better Understand the Ecological Diversity of Bat Viruses and the Bat Origin of Emerging Infectious Diseases', *The ISME Journal*, 10: 609–20.

Zhang, Y. Z. et al. (2019) 'Expanding the RNA Virosphere by Unbiased Metagenomics', *Annual Review of Virology*, 6: 119–39.