VIRUS EVOLUTION

Virus Evolution, 2023, **9(2)**, 1–11

DOI: https://doi.org/10.1093/ve/vead051

Advance Access Publication 18 August 2023

Research Article

Newly identified lineages of porcine hemagglutinating encephalomyelitis virus exhibit respiratory phenotype

Wan-Ting He,[1,2] Dongyan Li,[1] Guy Baele,[2,‡,*] Jin Zhao,[1] Zhiwen Jiang,[1] Xiang Ji,[3] Michael Veit,[4] Marc A. Suchard,[5,§,*] Edward C. Holmes,[6,**,*] Philippe Lemey,[2,††,*] Maciej F. Boni,[7,†] and Shuo Su[1,†,‡‡,*]

[1]Jiangsu Engineering Laboratory of Animal Immunology, Institute of Immunology, College of Veterinary Medicine, Academy for Advanced Interdisciplinary Studies, Nanjing Agricultural University, Nanjing 210095, China, [2]Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory for Clinical and Epidemiological Virology, KU Leuven, Leuven 3000, Belgium, [3]Department of Mathematics, School of Science & Engineering, Tulane University, New Orleans, LA 70118, USA, [4]Institute for Virology, Center for Infection Medicine, Veterinary Faculty, Free University Berlin, Berlin 14163, Germany, [5]Department of Biostatistics, Fielding School of Public Health, and Departments of Biomathematics and Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA, [6]Sydney Institute for Infectious Diseases, School of Life and Environmental Sciences and School of Medical Sciences, The University of Sydney, Sydney, NSW 2006, Australia and [7]Center for Infectious Disease Dynamics, Department of Biology, Pennsylvania State University, University Park, PA 16802, USA

[†]These authors contributed equally.

[‡]https://orcid.org/0000-0002-1915-7732

[§]https://orcid.org/0000-0001-9818-479X

[**]https://orcid.org/0000-0001-9596-3552

[††]https://orcid.org/0000-0003-2826-5353

[‡‡]https://orcid.org/0000-0003-0187-1185

*Corresponding author: E-mail: shuosu@njau.edu.cn; ssh5658485@163.com

Abstract

Swine pathogens have a long history of zoonotic transmission to humans, occasionally leading to sustained outbreaks or pandemics. Through a retrospective epidemiological study of swine populations in China, we describe novel lineages of porcine hemagglutinating encephalomyelitis virus (PHEV) complex coronaviruses (CoVs) that cause exclusively respiratory symptoms with no signs of the neurological symptoms typically associated with classical PHEV infection. Through large-scale epidemiological surveillance, we show that these novel lineages have circulated in at least eight provinces in southeastern China. Phylogenetic and recombination analyses of twenty-four genomes identified two major viral lineages causing respiratory symptoms with extensive recombination within them, between them, and between classical PHEV and the novel respiratory variant PHEV (rvPHEV) lineages. Divergence times among the sampled lineages in the PHEV virus complex date back to 1886–1958 (mean estimate 1928), with the two major rvPHEV lineages separating approximately 20 years later. Many rvPHEV viruses show amino acid substitutions at the carbohydrate-binding site of hemagglutinin esterase (HE) and/or have lost the cysteine required for HE dimerization. This resembles the early adaptation of human CoVs, where HE lost its hemagglutination ability to adapt to growth in the human respiratory tract. Our study represents the first report of the evolutionary history of rvPHEV circulating in swine and highlights the importance of characterizing CoV diversity and recombination in swine to identify pathogens with outbreak potential that could threaten swine farming.

Keywords: novel porcine hemagglutinating encephalomyelitis virus; swine coronavirus; recombination.

1. Introduction

Coronaviruses (CoVs) belong to the subfamily *Orthocoronavirinae* (family *Coronaviridae*) and can infect a wide range of hosts—including humans, wild mammals, domestic mammals, and birds—resulting in diverse sets of symptoms including respiratory disease, intestinal to nervous system disorders, and hepatitis. Clinical outcomes following infection range from asymptomatic or mild to severe and lethal (Weiss and Navas-Martin 2005; Su et al. 2016; Cui, Li, and Shi 2019; Lu et al. 2023). *Orthocoronavirinae* are phylogenetically classified into four genera: *Alpha-*, *Beta-*, *Gamma-*, and *Deltacoronavirus* (https://talk.ictvonline.org/taxonomy/). CoVs exhibit rich genetic diversity and complex evolutionary histories due to relatively high rates of nucleotide substitution and recombination, as well as frequent cross-species transmission (Woo et al. 2009; Graham and Baric 2010). Host jumps of CoVs from animals to humans have resulted in several novel pathogen emergence events with high fatality rates, including severe acute respiratory syndrome (SARS), Middle East respiratory syndrome, and the recent pandemic of SARS-CoV-2 (Guan et al. 2003; Vijgen et al. 2005; Zhou et al. 2020), the largest singular infectious disease mortality event in more than 100 years. One important vehicle for the movement of viruses from wildlife

to humans is via swine populations that are in relatively frequent contact with humans and other livestock. For example, pigs have been known to transmit influenza virus, pseudorabies virus, and Nipah virus to humans causing severe clinical outcomes (Goh et al. 2000; Sun et al. 2020; Liu et al. 2021).

Six CoVs are currently known to infect pigs. Of these, four belong to the genus *Alphacoronavirus*: transmissible gastroenteritis CoV, porcine respiratory CoV (PRCV), porcine epidemic diarrhea virus (PEDV), and swine acute diarrhea syndrome CoV. In contrast, porcine hemagglutinating encephalomyelitis virus (PHEV) belongs to the genus *Betacoronavirus*, while porcine delta CoV (PDCoV) is a member of the genus *Deltacoronavirus*. PHEV is the only known neurotropic CoV of pigs, with clinical manifestations that include vomiting and wasting disease, as well as encephalomyelitis (Mora-Diaz et al. 2019). Although it was one of the first porcine CoVs to be identified and isolated, only a limited number of PHEV studies have been performed and only about a dozen genome sequences are available to date. PHEV is classified as a member of the subgenus *Embecovirus*, which also contains human, bovine, and equine CoVs associated with intestinal or respiratory disease (Vijgen et al. 2006).

Embecovirus genomes encode sixteen nonstructural proteins (nsp1–16) and the structural proteins hemagglutinin esterase (HE), spike (S), envelope (E), membrane (M), and nucleocapsid (N). The S protein is central for viral entry and hence the key determinant of host and tissue tropism (Hulswit, de Haan, and Bosch 2016). Like S from other members of the Embecoviruses, the PHEV S protein binds to acetylated sialic acids on surface glycoproteins, apparently its only cellular receptors (Huang et al. 2015; Matrosovich, Herrler, and Klenk 2015; Hulswit, de Haan, and Bosch 2016). The HE protein is also involved in attachment to these receptors and has acetylesterase and hence receptor-destroying activity (Klausegger et al. 1999). Another protein unique for the Embeco- and Merbecoviruses is NS2 (ORF2)—a phosphodiesterase that antagonizes activation of RNase L and hence degradation of the viral genome (Goldstein et al. 2017).

PHEV was first identified in Canada in 1957 and subsequently in Europe, the Americas, and Asia between the 1970s and the 2010s (Roe and Alexander 1958; Cartwright et al. 1969; Pensaert and Callebaut 1974; Hirano and Ono 1998; Quiroga et al. 2008; Rho et al. 2011; Mora-Diaz et al. 2020). The virus has been reported in Jilin province China, where it caused encephalitis in piglets (Gao et al. 2011; Dong et al. 2014). However, PHEV variants recently found in the USA (GenBank accessions from KY419103 to KY419112) seem to cause only respiratory disease, and hence, their clinical manifestations are different from earlier reports (Lorbach et al. 2017).

Here, we describe the identification—through metagenomic sequencing—of a novel respiratory variant PHEV (rvPHEV) in pigs with influenza-like symptoms in the absence of neurological signs or encephalomyelitis. Further monitoring of the prevalence and evolution of these emerging Betacoronaviruses is needed to prevent future outbreaks in swine and to monitor potential public health risk posed by virus recombination. Our study also suggests that the genetic variation of animal CoVs may have an important impact on clinical symptoms and tissue tropism, posing a challenge to clinical diagnosis.

2. Methods

2.1 Collection and processing of clinical samples

In April 2021, an outbreak of influenza-like illness in piglets was reported on a farm in Fujian province, Southeast China. To investigate the cause of this disease, we collected nasal swab samples from ten piglets with respiratory symptoms. Samples were kept cold and transported to the laboratory, where they were stored at −80°C before use. Sample collection from all animals was approved by the Institutional Animal Care and Use Committee of Nanjing Agricultural University, Nanjing, China (number SYXK2017-0007).

2.2 Pathogen detection and retrospective epidemiological survey

2.2.1 *Etiological investigation and metagenomic sequencing on samples from Fujian*

Based on the respiratory symptoms of piglets, polymerase chain reaction (PCR) diagnostics were performed for the following common clinical pathogens that may cause respiratory symptoms: swine influenza A virus, porcine respiratory reproductive syndrome virus, porcine circovirus 2, and PRCV. As routine virus testing failed to determine the etiological agent, we performed total RNA library sequencing using transcriptome technology to detect potential pathogens according to our previous method (He et al. 2022a).

2.2.2 *Retrospective epidemiological survey of rvPHEV in China*

After identifying the pathogen, a retrospective investigation of rvPHEV in China was conducted on samples collected between 2016 and 2021, including 771 samples from 2021, as well as other samples from the Chinese Piglet Metagenomic Project. All clinical samples were screened using primers of PHEV/rvPHEV, designed according to the available PHEV sequences in GenBank (www.ncbi.nlm.nih.gov), and contigs were annotated as rvPHEV. To obtain rvPHEV genomes and HE and Spike gene sequences, the sample RNA was extracted using the Trizol kit (Vazyme, China). The HiScript II 1st Strand cDNA Synthesis Kit (Vazyme, China) was used for cDNA synthesis. Subsequently, PCR was performed with the rvPHEV detection primers. Samples identified as positive were selected and further subjected to amplification reactions with Phanta® Max Super-Fidelity DNA Polymerase (Vazyme, China) using a set of primers for rvPHEV genome amplification designed based on reference genomes. Finally, purified PCR amplification products were sequenced by the Sanger dideoxy chain termination method, or a next-generation sequencing method was applied as described earlier. A total of twenty-four rvPHEV genomes and two additional variants with complete HE and spike genes were sequenced. Genbank accession numbers for these twenty-six samples are OQ798806–OQ798833.

2.3 Bioinformatic analysis

2.3.1 *Assembly, annotation, and abundance estimates*

For each library, trimmomatic was used with default parameters to remove adapters and quality trim the sequence reads (Bolger, Lohse, and Usadel 2014). The remaining reads were assembled *de novo* into contigs using Megahit and those less than 500 nt were filtered out (Li et al. 2015). All the assembled contigs were aligned to the National Center for Biotechnology Information (NCBI) nonredundant protein database (nr) using Diamond blastx (Buchfink, Reuter, and Drost 2021) with the *E*-value cut-off set to 1E-5. To estimate the relative abundance of each virus species in each library, quality-trimmed reads were first mapped to the swine and human genome sequences and then mapped to the SILVA ribosomal RNA database (www.arb-silva.de, version 132.1) using Bowtie2 (Langmead and Salzberg 2012; Quast et al. 2013) Unmapped reads were subsequently mapped to confirmed viral scaffolds using the 'end-to-end' setting, and the abundance of each virus species was estimated as the number of mapped reads in each library (He et al. 2022a).

2.3.2 Sequence collection and alignment

To describe the (potentially recombinant) origins and evolution of PHEV and rvPHEV, other viral reference genome sequences (from species *Betacoronavirus 1*) that are closely related to the PHEV complex were included in the analysis. Specifically, human CoV OC43 (NC_006213, AY903459), bovine CoV (NC_003045, DQ811784), and equine CoV (EF446615) were downloaded from the NCBI GenBank (https://www.ncbi.nlm.nih.gov/) (So et al. 2019). All available PHEV complete genomes, as well as HE and S gene sequences available before 24 September 2022, were similarly downloaded from GenBank. A total of fifteen reference PHEV whole-genome sequences were used in the analyses: one from Belgium, two from China, one from South Korea, and eleven from the USA, covering a time period from the 1970s to January 2021 (Table S1). There are three additional variants with two complete HE and two spike sequences that were used in subsequent analyses. Sequences were aligned using MAFFT v7.475 and inspected manually (Katoh and Standley 2013).

2.3.3 Recombination analysis

Recombination analysis was performed according to the procedures described in the study by Boni et al (2020). Briefly, the nonparametric mosaic detection program 3SEQ v1.8 (Lam, Ratmann, and Boni 2018) was run on the full alignment of forty-three Embecoviruses sequences. 3SEQ reports breakpoints as ranges, and all breakpoint ranges were combined into a single list of all possible breakpoints. This list was complemented to generate a list of non-breakpoints, which were then assembled into contiguous regions called breakpoint-free regions (BFRs). The BFRs were named A through G, sorted by length for BFRs longer than 1000 nt (Table S1).

To ensure that mosaic recombination signals were supported by phylogenetic evidence (Boni et al. 2010), we looked for evidence of additional recombination within and between the BFRs. BFRs A and B were further split into sub-segments due to phylogenetic recombination signals within these BFRs. Neighboring BFRs were combined when no phylogenetic incongruence signal could be detected between them (Table S1), and we refer to these concatenated BFRs as non-recombining regions (NRRs). Phylogenetic incongruence signals were assessed by building maximum likelihood (ML) trees using RAxML v8.2.8 (Stamatakis 2014) (100 bootstrap replicates) on neighboring segments and determining if there were sequences that jumped between sequence clades with bootstrap (BS) support $\geq$80 for membership in both clades.

To assess robustness of mosaic recombination signals, the full alignment of forty-three sequences was run through additional tests in RDP4 (Martin et al. 2015) (all with default settings) with Recombination Detection Program (RDP) (Martin and Rybicki 2000), GENECONV (Padidam, Sawyer, and Fauquet 1999), Bootscan (Martin et al. 2005), MaxChi (Maynard Smith 1992), Chimaera (Posada and Crandall 2001), and SisScan (Gibbs, Armstrong, and Gibbs 2000) all showing evidence of recombination in the alignment ($P < 0.001$). The mosaic signals tested by these methods (including 3SEQ) are in some cases very similar, and statistical agreement is expected; in other words, these are not tests of independent signals. PhylPro (Weiller 1998), a phylogenetic correlation approach, did not detect signs of recombination at the $P = 0.01$ level.

2.3.4 Phylogenetic analysis

ML phylogenetic trees were estimated for NRR5 (the longest NRR) and the S and HE genes using IQ-TREE v2.0.3 under the general time-reversible (GTR) substitution model with a discretized gamma distribution (Minh et al. 2020). The ML trees were used as input for the analysis of temporal signal (i.e. molecular clock structure) in TempEst (Rambaut et al. 2016). To estimate the divergence time of rvPHEV from traditional PHEV, BEAST v1.10.5 with BEAGLE v3 was used to perform time-measured phylogenetic reconstructions (Suchard et al. 2018; Ayres et al. 2019). To accommodate potentially different rates of evolution for specific lineages as well as random branch-specific rate variation, we employ a mixed effect molecular clock model in our Bayesian phylogenetic inference (Vrancken et al. 2014; Bletsa et al. 2019) in order to accommodate two outlier lineages that were likely evolving at independent rates (see Section 3). We further specified a GTR substitution model with a discretized gamma distribution to model among-site rate heterogeneity and a nonparametric coalescent skygrid model (Gill et al. 2013). A Markov chain Monte Carlo chain analysis was run for a total of 100 million iterations and sampled every 50,000th iteration. Convergence and proper statistical mixing were assessed using Tracer v1 7 (Rambaut et al. 2018). Upon removing burn-in, maximum clade credibility (MCC) trees were summarized using TreeAnnotator as part of the BEAST v1.10.5 package (Suchard et al. 2018).

3. Results

3.1 Epidemiological and clinical characterization of a novel swine respiratory Betacoronavirus in China

In April 2021, an outbreak of influenza-like illness was observed among piglets on a pig breeding farm in Fujian province in southeastern China (Fig. 1A). In addition to respiratory symptoms, some pigs also exhibited gastrointestinal symptoms. We collected nasal swabs from tens piglets, tested them for known viruses, and excluded any common respiratory pathogens. Through metatranscriptomic and Sanger sequencing, we identified a novel PHEV variant that encodes a phosphorylated structural protein characteristic for Embecovirus (Cox, Parker, and Babiuk 1991) (Fig. S1), but with a complete deletion of the NS2 gene (shown in Figs. 1B and 3B). This novel variant also differs clinically from classical PHEV, since it causes only respiratory symptoms, but no vomiting, wasting disease, encephalomyelitis, or other neurological symptoms. PHEV variants with a similar genome structure, which are also only associated with respiratory symptoms, have been previously detected in older pigs in the USA (Lorbach et al. 2017). Considering their atypical clinical symptoms and genome structure, we propose that these viruses are termed 'respiratory variant PHEV' or rvPHEV.

Using a retrospective epidemiological analysis from 2016 to 2021, we subsequently detected a total of ninety-two positive rvPHEV samples. When multiple positives were present in the same farm at the same time, genome sequences were identical for these samples (data not shown), and only one sequenced genome per farm was included in our study, resulting in a total of twenty-four novel whole-genome sequences. HE and S gene sequences were obtained from two additional samples, and these viruses were named rvPHEV1 through rvPHEV26. These rvPHEV were detected in several provinces in China, including Guangdong, Guangxi, Guizhou, Yunnan, Fujian, Zhejiang, Jiangxi, and Anhui. Of note, in rvPHEV 1, 2, 3, 5, 10, 13, and 23 the entire NS2 gene was deleted, and all the remaining rvPHEV showed premature termination codon ranging from nucleotides 27 to 111 (Figs. 1B and 3B). In all cases, the NS2 gene does not encode for a functional protein.

We chose rvPHEV13 (an NS2-deleted strain) to intranasally inoculate 15-day-old piglets for a preliminary evaluation of pathogenicity and clinical symptoms, as we identified no other

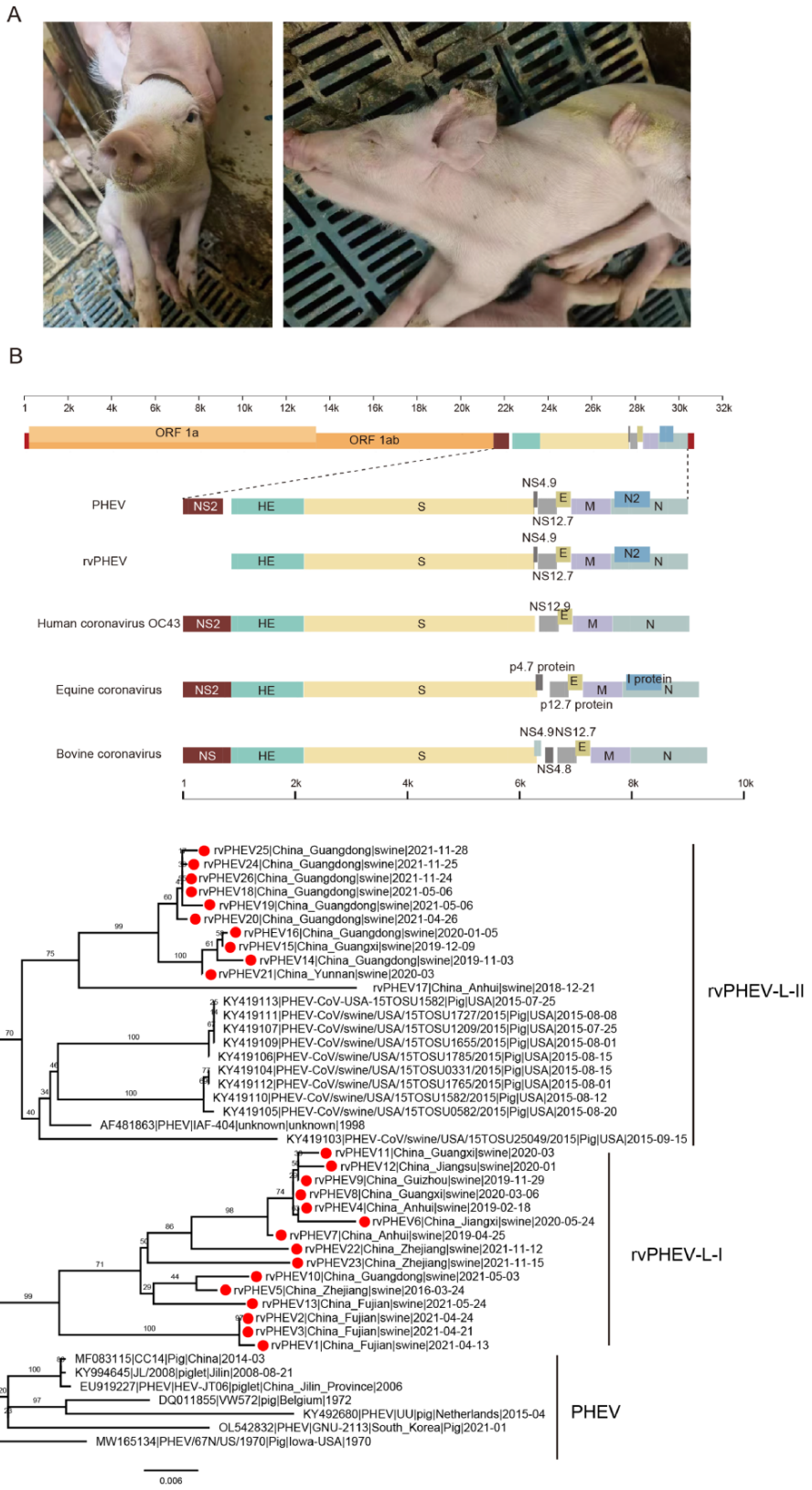

Figure 1. Signs and symptoms and genome characteristic of rvPHEV. (A) The clinical symptoms of rvPHEV-positive piglets are runny nose and lethargy. (B) Genome structure of *Embecovirus*, PHEV, rvPHEV, human CoV OC43, equine CoV, and bovine CoV. (C) ML phylogenetic tree of the S1 gene and rooted at classical PHEV Lineage.

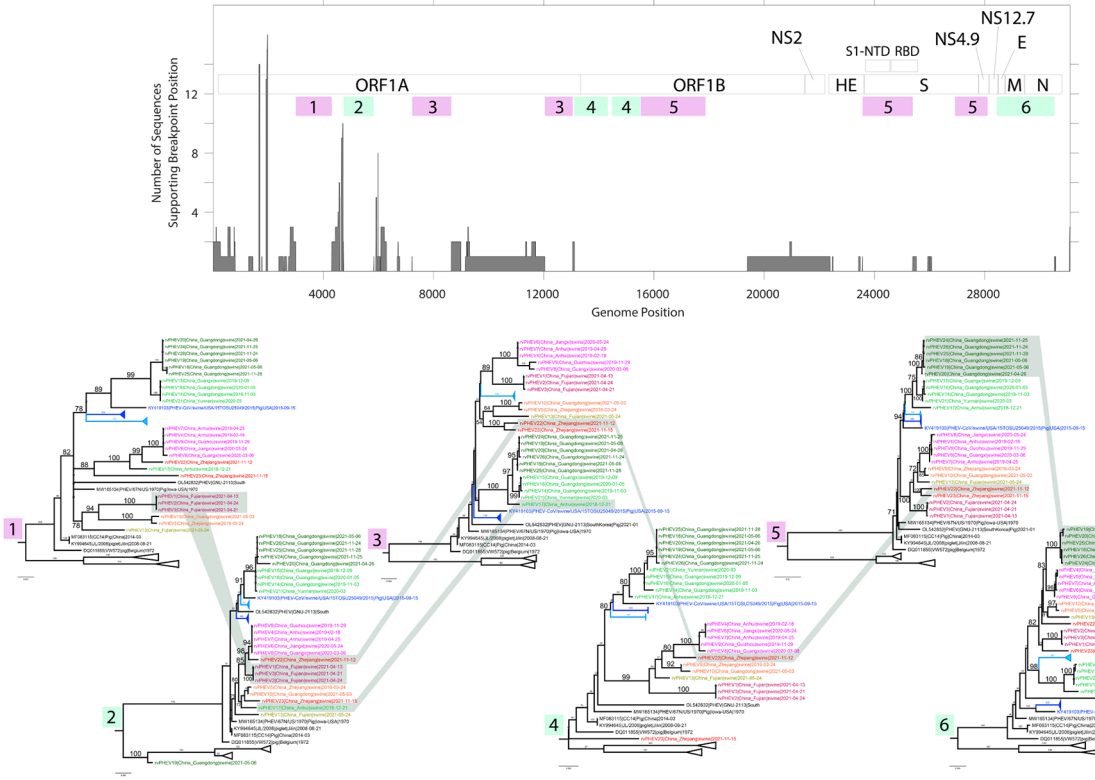

Figure 2. Recombination breakpoints and phylogenetic incongruence among NRRs. The top panel shows the level of breakpoint support at all possible 31,101 positions in the alignment of PHEVs and rvPHEVs used in this analysis. The y-axis shows the number of sequences (of forty-three) that support a breakpoint at a particular position when those sequences are evaluated as potential recombinants. The light gray boxes at the top of the plot outline the open reading frames in the virus genome, and the pink and green boxes show the positions of the six NRRs inferred for this analysis. Note that NRRs do not have to be contiguous. Phylogenetic trees for the six NRRs are shown below the plot. rvPHEV L-I sequences are shown in orange, red, and violet (rvPHEV1–10, 13, 22, and 23), rvPHEV L-II sequences are shown in light green and dark green, and their sibling clades of American PHEV sequences are shown in turquoise (KY419103), dark blue, and light blue, with two clades collapsed as triangles to reduce the size of the tree. Relevant bootstrap values ≥70 per cent are enlarged on all trees. The gray-shaded areas connecting the different trees show the phylogenetically incongruent positions of several example sequences and clades: rvPHEV1-3, rvPHEV17, rvPHEV22, and a L-II subclade sampled in Guangdong province in 2021.

suspected pathogenic viruses in this sample (cycle threshold of 21). After 24 h, the piglets exhibited rhinitis, but no neurological symptoms were found during the 5-day period post-inoculation. Clinical symptoms were mild, and the virus could only be detected in respiratory organs. There were no neurological symptoms or infection of the nerve tissue (Li et al., unpublished data).

3.2 Recombination and evolutionary history of PHEV and rvPHEV

The history of PHEV and novel rvPHEV lineage is highly recombinant with mosaic recombination signals abundant and easily detectable ($P < 10^{-160}$, 3SEQ Lam, Ratmann, and Boni 2018). Figure 2A shows the number of sequences (out of the complete alignment of forty-three) supporting inferred breakpoint positions throughout the genome. A substantial portion of the recombination events involved one of the outgroup sequences, equine CoV NC99 (EF446615.1), and further recombination events involving this sequence were ignored as these would only show evidence for deep ancestral recombination. Extracting all contiguous regions longer than 1000 nt where no breakpoints were inferred gives seven BFRs labeled A to G (Table S1) that can be further broken down or combined depending on the mosaic and phylogenetic signals present within and between these BFRs. Mosaic signals were detected within BFR A ($P < 4 \times 10^{-6}$) and BFR B ($P < 3 \times 10^{-5}$). Based on additional phylogenetic incongruence detected within these segments, both were split into two sub-segments each. Based on the presence/absence of phylogenetic incongruence signals between neighboring BFRs, some of these regions were combined (concatenated) into the six NRRs shown in Table S1. The longest of these, NRR5, is a concatenated region of 5349 nt in length.

Based on these six NRRs, rvPHEV lineages I and II can be defined for the Embecovirus subclade that contains both PHEV and rvPHEV sequences. Using NRR5, we define Lineage I (L-I) as the clade consisting of rvPHEV sequences 1 through 10, 13, 22, and 23 (orange, plum, pink, and red sequences in Fig. 2B). Similarly, Lineage II (L-II) is the clade defined by rvPHEV sequences 14–21 and 24–26 (light and dark green). The within-lineage pairwise nucleotide diversity is 0.017 for L-I and 0.017 for L-II. The between-lineage nucleotide pairwise distance is 0.031.

Phylogenetic relationships among individual sequences and between L-I and L-II show substantial phylogenetic incongruence among the NRRs, supporting both recent and ancestral recombination events (Fig. 2). Four individual sequences—rvPHEV 17, 19, 22, and 23—show clear recombinant histories, some with multiple breakpoints and potentially multiple past recombination events. rvPHEV 17 appears to be L-II in the majority of its genome (all bootstrap values ≥80 for NRR3–6) but has clear ancestry in different subclades of L-I in the 3- to 6-kb region that includes NRR1 and NRR2 (bootstrap values ≥80). rvPHEV 22, sampled in Zhejiang province in 2021, is a likely recombinant of the rvPHEV4–9 clade (several provinces, 2019–20) and rvPHEV 23 (also sampled in

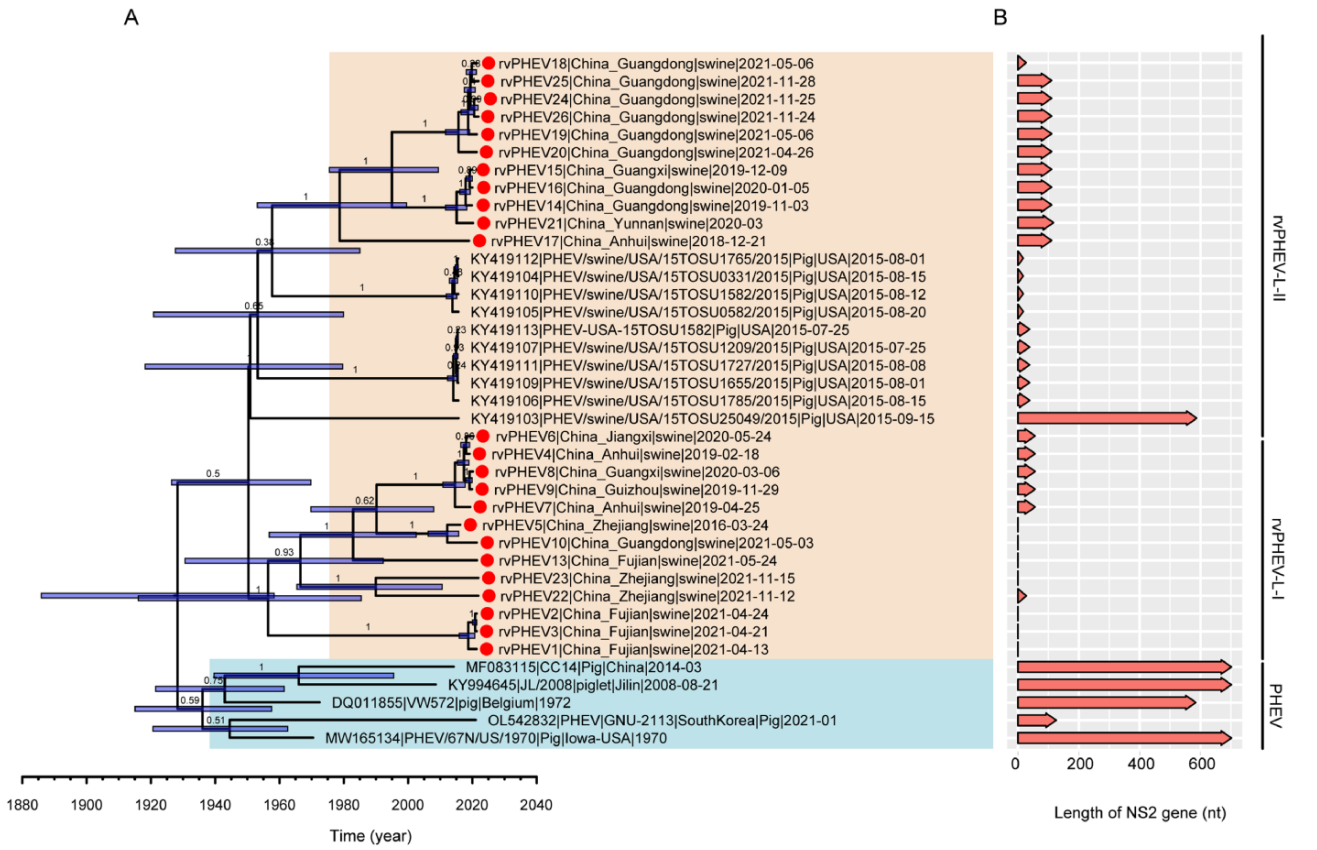

Figure 3. The time-calibrated MCC tree inferred for NRR5. (A) The MCC tree was constructed using Bayesian divergence time estimation in BEAST v1.10.5 using a mixed effects molecular clock model. The numbers on each branch indicate the posterior support value, and the blue bars show the 95 per cent HPD interval of the height of each node. Strains with red dots are those newly sequenced in this study. (B) The length of the NS2 gene of each strain is shown next to each tree tip label.

Zhejiang in 2021), with bootstrap values ≥80 supporting a complex multi-breakpoint recombinant history in NRR1 through NRR5.

Interestingly, respiratory variants 19 and 23 show evidence of past recombination with unsampled parental sequences. rvPHEV 23 appears as a distant outgroup to the entire PHEV complex in NRR4, and rvPHEV 19 is the result of a likely interspecies recombination event with the NRR2 genomic region (nucleotides 4723–5828 in ORF1A) having ancestry in the Embecovirus lineage containing human CoV OC43 and bovine CoV (100 percent bootstrap value). At least one recombination event between L-I and L-II can be readily seen in the NRR phylogenies, with a breakpoint after the end of the spike reading frame. A subclade of rvPHEV sampled in Guangdong province during April–November 2021 is a recombinant of older L-II viruses (rvPHEV14-16, 21; BS = 75) and L-I viruses closely related to rvPHEV4–9 (BS = 83). Probable inter-Lineage recombination is also seen in the American PHEV strains (see NRRs 3 and 6 in Fig. 2).

3.3 Time-scaled evolutionary history of rvPHEV

We focused on the longest NRR5 to infer a time-calibrated evolutionary history using Bayesian phylogenetic inference. NRR5 encompasses part of the ORF1b gene, the whole N-terminal domain, almost the entire receptor-binding domain of the S1 subunit, and a part of S2 subunit of the spike protein. We first assessed the temporal signal of NRR5 using TempEst, which yielded a coefficient of determination (R^2) of 0.36, indicating some correlation between the dates of the tips and the divergence from the root (Fig. S2). However, two lineages stood out in showing either lower or higher divergence than expected based on the root-to-tip regression. The lineage potentially evolving at a lower evolutionary rate consisted of two Chinese classical PHEV strains (JL/2008 [KY994645] and CC14 [MF083115]). The lineage potentially evolving at a higher rate consisted of three Chinese rvPHEV strains from Fujian (rvPHEV1-3). In both cases, the monophyletic clustering of the genomes involved suggests that standard explanations for outlier behavior such as errors in the sampling date of specific sequences are unlikely. To accommodate their potentially different evolutionary rates, as well as other sources of branch-specific rate variation, we employed a mixed effects molecular clock model for Bayesian divergence time estimation (Bletsa et al. 2019). This model identified a considerably lower rate for the PHEV lineage but not a considerably higher rate for the rvPHEV lineage, so we eventually only included a fixed effect for the former in the mixed effects clock. The result of this Bayesian reconstruction is represented as a MCC tree in Fig. 3A, indicating that the time to the most recent common ancestor (tMRCA) of PHEV and rvPHEV was estimated to be between 1886 and 1958 (95 percent highest posterior density (HPD) interval), with a mean date of 1928 (Fig. 3A). rvPHEV was estimated to have diverged from PHEV between 1926 and 1970 (mean of 1950). We can conclude that rvPHEV has diverged into two independent lineages, rvPHEV L-I (1916–85, mean of 1956) and rvPHEV L-II (1918–80, mean of 1951); both have a very high posterior support close to 1. The mean evolutionary rate of PHEV and rvPHEV excluding the two outlier sequences was estimated to be 2.70×10^{-4} nucleotide

Table 1. Prediction of proprotein convertase cleavage sites.

Virus	Sequence	Furin cleavage score
PHEV	TALRSRR/SFTT	0.688
rvPHEV-L-II	TALRSRR/SFTT	0.688
rvPHEV-L-II[a]	T**S**LRSRR/S**L**TT	0.742
rvPHEV1-3	TA**R**RSRR/SFTT	0.786
Other rvPHEV-L-I	TALRSRR/S**V**TT	0.798

[a]Strains: KY419106, KY419107, KY419109, KY419111, and KY419113. Amino acid mutations are indicated in bold.

substitutions per site per year (with 95 per cent HPD interval being $1.48–4.04 \times 10^{-4}$ subs/site/year).

The variants of the L-I lineage have only been identified in some provinces of China Fujian ($n = 4$), Guangdong ($n = 1$), Zhejiang ($n = 3$), Anhui ($n = 2$), Guangxi ($n = 1$), Guangdong ($n = 1$), and Guizhou ($n = 1$). Based on the available data, the L-II samples found that Guangdong ($n = 8$), Yunan ($n = 1$), Guangxi ($n = 1$), and Anhui ($n = 1$) share a common ancestor with those viruses found in the USA. Since the emergence of rvPHEV, its effective population size has increased, reaching a peak around 2007 and then declining sharply from 2007 to 2015. We also inferred time-measured evolutionary histories based on the HE and S genes. Both the structure of the phylogenetic tree and the estimated tMRCA are similar to the results obtained for NRR5 (Fig. S3).

3.4 Amino acid substitutions in the HE gene of swine respiratory Betacoronaviruses likely affect its receptor-binding activity

The alignment of classic PHEV and rvPHEV S gene sequences revealed few unique changes and only one of potential functional relevance. It involves changes at the cleavage site between S1 and S2, some of which are predicted to lead to a moderate increase in the furin cleavage score compared with classic PHEV stains (Table 1). In contrast, the HE gene of rvPHEV strains contains several unique amino acid changes not present in classical PHEV strains (Fig. 4A). Those that likely affect the hemagglutinating activity are highlighted in the 3D structure of closely related bovine HE (Zeng et al. 2008) (Fig. 4A and B). The carbohydrate-binding site (CBS) of HE is composed of two hydrophobic pockets (Fig. 4C, cyan sticks) and hydrophilic amino acids that form hydrogen bonds with various atoms of sialic acid, either directly (Fig. 4C, green sticks) or via a water molecule (Fig. 4C, blue sticks). The latter residues are substituted not only in all variants of the rvPHEV-L-I lineage, in rvPHEV 26, but also in several PHEV strains by amino acids, which are not able to form hydrogen bonds (Fig. 4D–G, red sticks). Many rvPHEV strains contain additional changes either in the hydrophobic pockets, which restrict access to the binding site, or in the amino acids forming hydrogen bonds that likely have important functional effects. rvPHEV strains 1–3 contain two other unique changes in the C-terminal part of HE, located a few amino acids upstream of the transmembrane region. The highly conserved Pro 383 is replaced by a Ser, and Cys385 by a Phe. Cys385 forms the single disulfide bond between two monomers, and hence, this change likely affects the formation of the HE dimer and its functionality (Phillips, Gomez-Navarro, and Miller 2020).

4. Discussion

Due to the coronavirus disease (COVID)-19 pandemic, there has been increased research effort directed toward CoV diversity and evolution. Despite this, the swine betacoronavirus PHEV remains poorly studied. Infection with classical PHEV variants often results in encephalomyelitis, leading to vomiting, inability to eat, and neurological symptoms such as abnormal gait and tremors (Cartwright et al. 1969; Quiroga et al. 2008; Li et al. 2016; Shi et al. 2018). Our data provide an enhanced understanding of the diversity and circulation of an endemic novel PHEV-like respiratory variant (rvPHEV) and of the evolutionary relationships between classic PHEV and rvPHEV. A PHEV variant in the USA that caused influenza-like illness in older pigs has been previously described (Lorbach et al. 2017), and another PHEV variant was recently identified in South Korea by sequencing samples from pigs with diarrheal symptoms (Kim et al. 2022). Our study shows, for the first time, that there is substantial diversity in PHEV-like variants in China, with some clades clearly distant from classical PHEV (Vijgen et al. 2006; Lorbach et al. 2017). Based on the differences in the genome structure—especially in the NS2 gene (Figs. 1B and 3B)—and clinical symptoms (Fig. 1A), we propose that this expanded group of PHEV variants is termed rvPHEV. In addition, we provide evidence that rvPHEV is a respiratory pathogen and characterize the genomes of clinically relevant strains circulating in swine herds in the south of China. Both natural and laboratory-infected pigs (unpublished data) show similar clinical symptoms and relatively high virus titers. Considering that the virus has not been isolated, and few fecal samples from diarrheal pigs were positive for rvPHEV, detailed animal experiments with virus isolates are urgently needed to further evaluate the tissue tropism, pathogenicity, and clinical symptoms in swine.

Retrospective epidemiology suggests that rvPHEV has been circulating in at least eight provinces in China. Phylogenetic analysis shows that rvPHEVs can be divided into two lineages, rvPHEV-L-I and rvPHEV-L-II. Of note, there is a possible interspecies recombination event detected for rvPHEV19, on NRR2 (located in the ORF1a protein, 4734–5828nt), with this region being closer to BCoV and HCoV OC43 than to classical PHEV or rvPHEV. Genomic similarity patterns across rvPHEV23 are also consistent with recombination involving unsampled parental viruses, suggesting that the diversity of unsampled Embecoviruses (potentially outside swine populations) may be relatively large. Interspecies recombination among Betacoronaviruses has been previously detected among camel, rabbit, and rodent CoVs (So et al. 2019). There are several recombination breakpoints at the start and end of the HE gene (Fig. 2), although our phylogenetic evidence does not directly implicate HE as a focal point of past recombination as HE appears to reside in the middle of NRR5. The BFR overlapping the HE gene was not analyzed in this study as its length (932nt) fell below the 1000-nt threshold set for evaluating BFRs. It has been suggested that the HE gene, as a unique structural protein that is expressed only in the subgenus *Embecovirus*, is believed to have been acquired from influenza C virus through heterologous recombination (Luytjes et al. 1988). We also detected inter- and intra-lineage recombination among classical PHEV, rvPHEV-L-I, and rvPHEV-L-II. This is consistent with frequent coinfections that can only occur in high-prevalence settings, indicating that a lack of surveillance for CoVs in pigs may be the reason that the full diversity of swine Embecoviruses has not been described. Since swines are raised and traded for meat globally, they deserves attention and continuous monitoring. This is illustrated by the swine-origin human influenza virus pandemic of 2009 (Newman et al. 2008; Shinde et al. 2009; Smith et al. 2009) and swine-origin PDCoV that independently infected Haitian children several times in 2014–15 (Lednicky et al. 2021).

Our phylogenetic analyses demonstrated that the two lineages of rvPHEV have a common ancestor that dates back to 1950

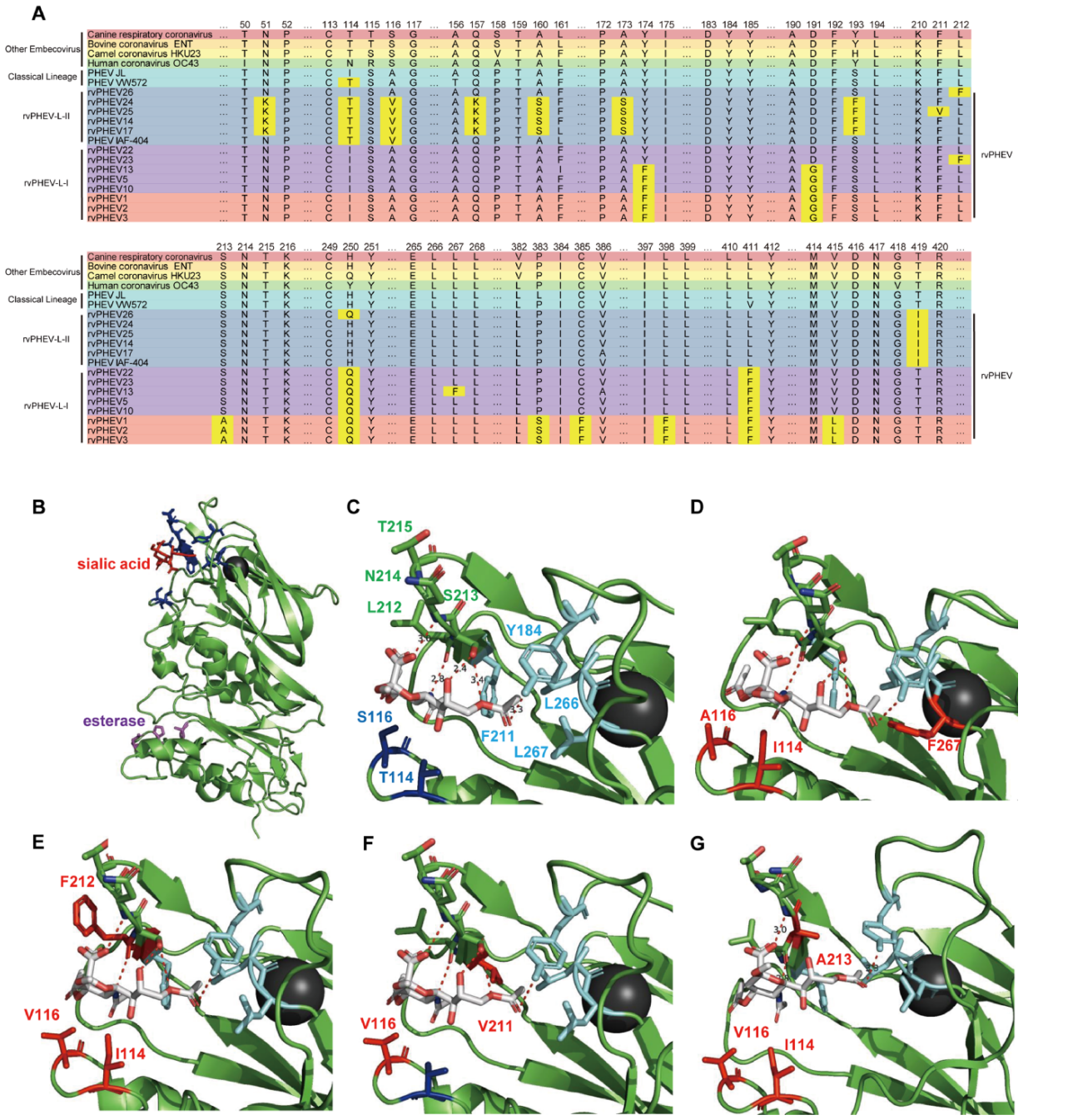

Figure 4. The structure of bovine HE (PDB: 3CLA) showing the amino acid substitutions in rvPHEV strains. (A) Unique amino acid substitutions in the HE protein of rvPHEV-L-I and rvPHEV-L-II, aligned with other Embecoviruses, similar canine CoVs, bovine CoVs, human CoV OC43, camel CoV, and classical PHEV. (B) The structure of an HE monomer from bovine CoV with bound 9–0 acetyl sialic acid analog (red stick). The amino acids constituting the sialic acid binding site are shown as blue sticks and the catalytic triad of the esterase domain as magenta sticks. The black sphere is a potassium ion. (C) Magnification of the 9–0 acetyl sialic acid binding site. The site is composed of two hydrophobic pockets, one formed by Tyr 184, Leu 266, Leu 267, and Phe 211, which accommodates the 9-O-acetyl group and the second by Phe 211, which binds the 5N-acetyl group of the receptor (cyan sticks). In addition, residues Ser 213, Asn 214, and Thr 215 form hydrogen bonds (dotted lines) with various atoms of sialic acid (green sticks). Furthermore, the hydroxy-amino acids Thr 114 and Ser 116 are connected via a water molecule (not depicted) to 0–7 of the 9-O-acetylated sialic acid analog (blue sticks). (D–G) Amino acid changes present in rvPHEV strains, but not in any PHEV strain are shown as red sticks. (D) The change Leu267Phe in rvPHEV13 narrows the hydrophobic pocket for the 9-O-acetyl group of the ligand. (E) The Leu212Phe substitution in rvPHEV 23 + 26 narrows the space for the C1 group of the ligand. (F) The change Phe211Val in rvPHEV 18, 19, 20, and 25 affects the hydrophobic pocket for the 5N-acetyl group of the receptor. The very similar change Phe211Ala in HE of rvPHEV abrogates receptor binding. (G) The change Ser213Ala in in rvPHEV 1–3 destroys the hydrogen bond with 0-8 of 9–0 acetyl sialic acid. Figures were created using PyMol 2.1.1 from pdb-file 3CLA using the mutagenesis wizard.

(posterior mean estimate) and had formed two independently circulating lineages in piglets by the early to mid-1950s. The mean evolutionary rate of PHEV and rvPHEV (2.70×10^{-4} expected substitutions per site per year) was lower than that reported previously for other porcine CoVs such as PEDV (1.93×10^{-3} expected substitutions per site per year) and PDCoV (1.67×10^{-3} expected substitutions per site per year) (He et al. 2020, 2022b), across 60 years of evolution in pigs. One striking feature of all rvPHEV viruses is the deletion of the NS2 gene, which exhibits >50 per cent amino acid identity to NS2 of murine hepatitis virus (MHV). NS2 of MHV antagonizes the Type I interferon antiviral response by cleaving 2',5'-oligoadenylate, the product of the interferon inducible 2',5'-oligoadenylate synthetase (Cervantes-Barragan et al. 2007). This prevents activation of the cellular RNase L and hence degradation of viral RNA. Although NS2 is non-essential for CoV replication in cell culture, a role for NS2 in viral pathogenicity is plausible, since the deletion of MHV NS2 significantly attenuates the virus when it is inoculated into mice. Although the function of NS2 of MHV is cell type specific (Roth-Cross, Bender, and Weiss 2008; Roth-Cross et al. 2009), it can be assumed that the deletion of the NS2 gene in rvPHEV strains contributes to the observed attenuation of these viruses. Many rvPHEVs have unique substitutions at the CBS of HE that are likely to diminish its function (Fig. 4). One of the substitutions (Phe211Val) has been experimentally shown to abolish the receptor-binding activity of bovine HE (Lang et al. 2020). Furthermore, rvPHEVs 1–3 exhibit a substitution of cysteine 385 that forms the only intermolecular disulfide bond between HE monomers. This exchange likely destabilizes the HE dimer and might even compromise its incorporation into virus particles since oligomerization and folding of glycoproteins are intimately linked and misfolded proteins are degraded by the quality control system of the endoplasmic reticulum (Ellgaard and Helenius 2003).

These changes in HE of rvPHEV resemble substitutions described in HCoVs OC43 and coronavirus HKU-1 after their cross-species transfer from cattle to humans. HE gradually lost its lectin function as a consequence of adaptation for growth in the human respiratory tract (Bakkers et al. 2017; Hurdiss et al. 2020; Lang et al. 2020). Although the cross-species transmission of CoVs from cattle to humans and the change in tissue tropism described here for rvPHEV are very different events, in both cases, viruses encountered a new glycan environment (Bosch et al. 2013). Since the receptor-binding and receptor-destroying activities of a virus are usually highly adapted to a particular tissue of a particular host (Wasik, Barnard, and Parrish 2016), they are temporarily unbalanced when a virus encounters a new glycan environment. We propose that the various mutations near the sialic acid binding site in HE of rvPHEV may be indicative of a viral population attempting to restore the functional balance between receptor binding and receptor elution.

In sum, we report here the emergence of new variants (named rvPHEV) of a porcine CoV. While these viruses currently constitute only a mild respiratory pathogen in pigs, the future impact of these novel swine viral variants, on livestock farming and human society in general, deserves further study. Additional regional surveillance studies should be conducted to monitor emerging genetic variants of rvPHEV, such that prevention and control measures can be implemented to detect range expansion.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Acknowledgements

W.-T.H. is the co-corresponding author (hewt@cpu.edu.cn) and now works at the School of Pharmacy, China Pharmaceutical University. We thank the Bioinformatics Center, Nanjing Agricultural University, for providing the high-performance computing platform.

Funding

S.S., W.-T.H., D.L., J.Z., and Z.J. were financially supported by the National Key Research and Development Program of China (grant number 2021YFD1801101) and Six Talent Peaks Project of Jiangsu Province of China (grant number NY-045). P.L. acknowledges the support of the Wellcome Trust (Collaborators Award 206298/Z/17/Z – ARTIC network), the European Research Council (grant agreement number 725422 – ReservoirDOCS), the NIH grant R01AI153044 and the Research Foundation – Flanders ('Fonds voor Wetenschappelijk Onderzoek – Vlaanderen', G0D5117N and G051322N). G.B. acknowledges support from the Internal Funds KU Leuven (grant number C14/18/094) and the Research Foundation – Flanders ('Fonds voor Wetenschappelijk Onderzoek – Vlaanderen', G0E1420N, G098321N). M.F.B. was funded by National Institutes of Health grant NIAID R01AI153355 and Bill and Melinda Gates Foundation grant INV-005517. Experimental work in the laboratory of M.V. was funded by the German Research Foundation. X.J. acknowledges support from the NVIDIA academic hardware grant program. E.C.H. was funded by an Australian Research Council Australian Laureate Fellowship (FL170100022) and AIR@InnoHK administered by the Innovation and Technology Commission, Hong Kong Special Administrative Region, China.

Conflict of interest: None declared.

Significance

Cross-species transmission and genetic recombination have frequently been observed in CoVs among animals. Here, we report for the first time two new lineages of a respiratory variant of PHEV (rvPHEV) circulating in China. This recently discovered clinical phenotype shows exclusively respiratory symptoms in swine and none of the neurological symptoms associated with classical PHEV. These rvPHEVs show extensive recombination within the rvPHEV/PHEV complex of viruses. In addition, in both lineages, potentially adaptive mutations appear in the HE protein that may be associated with changes in tissue tropism. Further studies of these and other CoVs in swine will be critical for determining if these viruses pose a risk to animal farming.

References

Ayres, D. L. et al. (2019) 'BEAGLE 3: Improved Performance, Scaling, and Usability for a High-Performance Computing Library for Statistical Phylogenetics', *Systematic Biology*, 68: 1052–61.

Bakkers, M. J. et al. (2017) 'Betacoronavirus Adaptation to Humans Involved Progressive Loss of Hemagglutinin-Esterase Lectin Activity', *Cell Host & Microbe*, 21: 356–66.

Bletsa, M. et al. (2019) 'Divergence Dating Using Mixed Effects Clock Modelling: An Application to HIV-1', *Virus Evolution*, 5: vez036.

Bolger, A. M., Lohse, M., and Usadel, B. (2014) 'Trimmomatic: A Flexible Trimmer for Illumina Sequence Data', *Bioinformatics*, 30: 2114–20.

Boni, M. F. et al. (2010) 'Guidelines for Identifying Homologous Recombination Events in Influenza A Virus', *PLoS One*, 5: e10434.

Boni, M. F. et al. (2020) 'Evolutionary Origins of the SARS-CoV-2 Sarbecovirus Lineage Responsible for the COVID-19 Pandemic', *Nat Microbiol*, 5: 1408–17.

Bosch, A. A. et al. (2013) 'Viral and Bacterial Interactions in the Upper Respiratory Tract', *PLoS Pathogens*, 9: e1003057.

Buchfink, B., Reuter, K., and Drost, H. G. (2021) 'Sensitive Protein Alignments at Tree-of-life Scale Using DIAMOND', *Nature Methods*, 18: 366–8.

Cartwright, S. F. et al. (1969) 'Vomiting and Wasting Disease of Piglets', *Veterinary Record*, 84: 175–6.

Cervantes-Barragan, L. et al. (2007) 'Control of Coronavirus Infection through Plasmacytoid Dendritic-cell-derived Type I Interferon', *Blood*, 109: 1131–7.

Cox, G. J., Parker, M. D., and Babiuk, L. A. (1991) 'Bovine Coronavirus Nonstructural Protein Ns2 Is a Phosphoprotein', *Virology*, 185: 509–12.

Cui, J., Li, F., and Shi, Z. L. (2019) 'Origin and Evolution of Pathogenic Coronaviruses', *Nature Reviews. Microbiology*, 17: 181–92.

Dong, B. et al. (2014) 'Identification and Genetic Characterization of Porcine Hemagglutinating Encephalomyelitis Virus from Domestic Piglets in China', *Archives of Virology*, 159: 2329–37.

Ellgaard, L., and Helenius, A. (2003) 'Quality Control in the Endoplasmic Reticulum', *Nature Reviews. Molecular Cell Biology*, 4: 181–91.

Gao, W. et al. (2011) 'Vomiting and Wasting Disease Associated with Hemagglutinating Encephalomyelitis Viruses Infection in Piglets in Jilin, China', *Virology Journal*, 8: 130.

Gibbs, M. J., Armstrong, J. S., and Gibbs, A. J. (2000) 'Sister-Scanning: A Monte Carlo Procedure for Assessing Signals in Recombinant Sequences', *Bioinformatics*, 16: 573–82.

Gill, M. S. et al. (2013) 'Improving Bayesian Population Dynamics Inference: A Coalescent-Based Model for Multiple Loci', *Molecular Biology and Evolution*, 30: 713–24.

Goh, K. J. et al. (2000) 'Clinical Features of Nipah Virus Encephalitis among Pig Farmers in Malaysia', *New England Journal of Medicine*, 342: 1229–35.

Goldstein, S. A. et al. (2017) 'Lineage A Betacoronavirus NS2 Proteins and the Homologous Torovirus Berne Pp1a Carboxy-Terminal Domain Are Phosphodiesterases that Antagonize Activation of RNase L', *Journal of Virology*, 91: 10–128.

Graham, R. L., and Baric, R. S. (2010) 'Recombination, Reservoirs, and the Modular Spike: Mechanisms of Coronavirus Cross-Species Transmission', *Journal of Virology*, 84: 3134–46.

Guan, Y. et al. (2003) 'Isolation and Characterization of Viruses Related to the SARS Coronavirus from Animals in Southern China', *Science*, 302: 276–8.

He, W. T. et al. (2020) 'Genomic Epidemiology, Evolution, and Transmission Dynamics of Porcine Deltacoronavirus', *Molecular Biology and Evolution*, 37: 2641–54.

——— et al. (2022a) 'Virome Characterization of Game Animals in China Reveals a Spectrum of Emerging Pathogens', *Cell*, 185: 1117–1129.e1118.

——— et al. (2022b) 'Phylogeography Reveals Association between Swine Trade and the Spread of Porcine Epidemic Diarrhea Virus in China and across the World', *Molecular Biology and Evolution*, 39: msab364.

Hirano, N., and Ono, K. (1998) 'A Serological Survey of Human Coronavirus in Pigs of the Tohoku District of Japan', *Advances in Experimental Medicine and Biology*, 440: 491–4.

Huang, X. et al. (2015) 'Human Coronavirus HKU1 Spike Protein Uses O-acetylated Sialic Acid as an Attachment Receptor Determinant and Employs Hemagglutinin-Esterase Protein as a Receptor-Destroying Enzyme', *Journal of Virology*, 89: 7202–13.

Hulswit, R. J., de Haan, C. A., and Bosch, B. J. (2016) 'Coronavirus Spike Protein and Tropism Changes', *Advances in Virus Research*, 96: 29–57.

Hurdiss, D. L. et al. (2020) 'Cryo-EM Structure of coronavirus-HKU1 Haemagglutinin Esterase Reveals Architectural Changes Arising from Prolonged Circulation in Humans', *Nature Communications*, 11: 4646.

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.

Kim, Y. et al. (2022) 'Complete Genome Sequence of a Novel Porcine Hemagglutinating Encephalomyelitis Virus Strain Identified in South Korea', *Archives of Virology*, 167: 1381–5.

Klausegger, A. et al. (1999) 'Identification of a Coronavirus Hemagglutinin-Esterase with a Substrate Specificity Different from Those of Influenza C Virus and Bovine Coronavirus', *Journal of Virology*, 73: 3737–43.

Lam, H. M., Ratmann, O., and Boni, M. F. (2018) 'Improved Algorithmic Complexity for the 3SEQ Recombination Detection Algorithm', *Molecular Biology and Evolution*, 35: 247–51.

Lang, Y. et al. (2020) 'Coronavirus Hemagglutinin-Esterase and Spike Proteins Coevolve for Functional Balance and Optimal Virion Avidity', *Proceedings of the National Academy of Sciences of the United States of America*, 117: 25759–70.

Langmead, B., and Salzberg, S. L. (2012) 'Fast Gapped-Read Alignment with Bowtie 2', *Nature Methods*, 9: 357–9.

Lednicky, J. A. et al. (2021) 'Independent Infections of Porcine Deltacoronavirus among Haitian Children', *Nature*, 600: 133–7.

Li, D. et al. (2015) 'MEGAHIT: An Ultra-Fast Single-Node Solution for Large and Complex Metagenomics Assembly via Succinct de Bruijn Graph', *Bioinformatics*, 31: 1674–6.

Li, Z. et al. (2016) 'The Evidence of Porcine Hemagglutinating Encephalomyelitis Virus Induced Nonsuppurative Encephalitis as the Cause of Death in Piglets', *PeerJournal*, 4: e2443.

Liu, Q. et al. (2021) 'A Novel Human Acute Encephalitis Caused by Pseudorabies Virus Variant Strain', *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, 73: e3690–70.

Lorbach, J. N. et al. (2017) 'Porcine Hemagglutinating Encephalomyelitis Virus and Respiratory Disease in Exhibition Swine, Michigan, USA, 2015', *Emerging Infectious Diseases*, 23: 1168–71.

Lu, M. et al. (2023) 'Zoonotic Risk Assessment among Farmed Mammals', *Cell*, 186: 2040–2040.e2041.

Luytjes, W. et al. (1988) 'Sequence of Mouse Hepatitis Virus A59 mRNA 2: Indications for RNA Recombination between Coronaviruses and Influenza C Virus', *Virology*, 166: 415–22.

Martin, D., and Rybicki, E. (2000) 'RDP: Detection of Recombination amongst Aligned Sequences', *Bioinformatics*, 16: 562–3.

Martin, D. P. et al. (2015) 'RDP4: Detection and Analysis of Recombination Patterns in Virus Genomes', *Virus Evolution*, 1: vev003.

——— et al. (2005) 'A Modified Bootscan Algorithm for Automated Identification of Recombinant Sequences and Recombination Breakpoints', *AIDS Research and Human Retroviruses*, 21: 98–102.

Matrosovich, M., Herrler, G., and Klenk, H. D. (2015) 'Sialic Acid Receptors of Viruses', *Topics in Current Chemistry*, 367: 1–28.

Maynard Smith, J. (1992) 'Analyzing the Mosaic Structure of Genes', *Journal of Molecular Evolution*, 34: 126–9.

Minh, B. Q. et al. (2020) 'IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era', *Molecular Biology and Evolution*, 37: 1530–4.

Mora-Diaz, J. C. et al. (2019) 'Porcine Hemagglutinating Encephalomyelitis Virus: A Review', *Frontiers in Veterinary Science*, 6: 53.

——— et al. (2020) 'Detecting and Monitoring Porcine Hemagglutinating Encephalomyelitis Virus, an Underresearched Betacoronavirus', *mSphere*, 5: 10–128.

Newman, A. P. et al. (2008) 'Human Case of Swine Influenza A (H1N1) Triple Reassortant Virus Infection, Wisconsin', *Emerging Infectious Diseases*, 14: 1470–2.

Padidam, M., Sawyer, S., and Fauquet, C. M. (1999) 'Possible Emergence of New Geminiviruses by Frequent Recombination', *Virology*, 265: 218–25.

Pensaert, M. B., and Callebaut, P. E. (1974) 'Characteristics of a Coronavirus Causing Vomition and Wasting in Pigs', *Archiv Fur Die Gesamte Virusforschung*, 44: 35–50.

Phillips, B. P., Gomez-Navarro, N., and Miller, E. A. (2020) 'Protein Quality Control in the Endoplasmic Reticulum', *Current Opinion in Cell Biology*, 65: 96–102.

Posada, D., and Crandall, K. A. (2001) 'Evaluation of Methods for Detecting Recombination from DNA Sequences: Computer Simulations', *Proceedings of the National Academy of Sciences of the United States of America*, 98: 13757–62.

Quast, C. et al. (2013) 'The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools', *Nucleic Acids Research*, 41: D590–596.

Quiroga, M. A. et al. (2008) 'Hemagglutinating Encephalomyelitis Coronavirus Infection in Pigs, Argentina', *Emerging Infectious Diseases*, 14: 484–6.

Rambaut, A. et al. (2018) 'Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7', *Systematic Biology*, 67: 901–4.

——— et al. (2016) 'Exploring the Temporal Structure of Heterochronous Sequences Using TempEst (Formerly Path-O-Gen)', *Virus Evolution*, 2: vew007.

Rho, S. et al. (2011) 'Detection and Genetic Analysis of Porcine Hemagglutinating Encephalomyelitis Virus in South Korea', *Virus Genes*, 42: 90–6.

Roe, C. K., and Alexander, T. J. (1958) 'A Disease of Nursing Pigs Previously Unreported in Ontario', *Canadian Journal of Comparative Medicine and Veterinary Science*, 22: 305–7.

Roth-Cross, J. K. et al. (2009) 'Organ-Specific Attenuation of Murine Hepatitis Virus Strain A59 by Replacement of Catalytic Residues in the Putative Viral Cyclic Phosphodiesterase Ns2', *Journal of Virology*, 83: 3743–53.

Roth-Cross, J. K., Bender, S. J., and Weiss, S. R. (2008) 'Murine Coronavirus Mouse Hepatitis Virus Is Recognized by MDA5 and Induces Type I Interferon in Brain Macrophages/Microglia', *Journal of Virology*, 82: 9829–38.

Shi, J. et al. (2018) 'Genomic Characterization and Pathogenicity of a Porcine Hemagglutinating Encephalomyelitis Virus Strain Isolated in China', *Virus Genes*, 54: 672–83.

Shinde, V. et al. (2009) 'Triple-reassortant Swine Influenza A (H1) in Humans in the United States, 2005-2009', *New England Journal of Medicine*, 360: 2616–25.

Smith, G. J. et al. (2009) 'Origins and Evolutionary Genomics of the 2009 Swine-Origin H1N1 Influenza A Epidemic', *Nature*, 459: 1122–5.

So, R. T. Y. et al. (2019) 'Diversity of Dromedary Camel Coronavirus HKU23 in African Camels Revealed Multiple Recombination Events among Closely Related Betacoronaviruses of the Subgenus Embecovirus', *Journal of Virology*, 93: 10–128.

Stamatakis, A. (2014) 'RAxML Version 8: A Tool for Phylogenetic Analysis and Post-analysis of Large Phylogenies', *Bioinformatics*, 30: 1312–3.

Su, S. et al. (2016) 'Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses', *Trends in Microbiology*, 24: 490–502.

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

Sun, H. et al. (2020) 'Prevalent Eurasian Avian-like H1N1 Swine Influenza Virus with 2009 Pandemic Viral Genes Facilitating Human Infection', *Proceedings of the National Academy of Sciences of the United States of America*, 117: 17204–10.

Vijgen, L. et al. (2005) 'Complete Genomic Sequence of Human Coronavirus OC43: Molecular Clock Analysis Suggests a Relatively Recent Zoonotic Coronavirus Transmission Event', *Journal of Virology*, 79: 1595–604.

Leen, V. et al. (2006) 'Evolutionary History of the Closely Related Group 2 Coronaviruses: Porcine Hemagglutinating Encephalomyelitis Virus, Bovine Coronavirus, and Human Coronavirus OC43', *Journal of Virology*, 80: 7270–4 .

Vrancken, B. et al. (2014) 'The Genealogical Population Dynamics of HIV-1 in a Large Transmission Chain: Bridging within and among Host Evolutionary Rates', *PLoS Computational Biology*, 10: e1003505.

Wasik, B. R., Barnard, K. N., and Parrish, C. R. (2016) 'Effects of Sialic Acid Modifications on Virus Binding and Infection', *Trends in Microbiology*, 24: 991–1001.

Weiller, G. F. (1998) 'Phylogenetic Profiles: A Graphical Method for Detecting Genetic Recombinations in Homologous Sequences', *Molecular Biology and Evolution*, 15: 326–35.

Weiss, S. R., and Navas-Martin, S. (2005) 'Coronavirus Pathogenesis and the Emerging Pathogen Severe Acute Respiratory Syndrome Coronavirus', *Microbiology and Molecular Biology Reviews*, 69: 635–64.

Woo, P. C. et al. (2009) 'Coronavirus Diversity, Phylogeny and Interspecies Jumping', *Experimental Biology and Medicine*, 234: 1117–27.

Zeng, Q. et al. (2008) 'Structure of Coronavirus Hemagglutinin-Esterase Offers Insight into Corona and Influenza Virus Evolution', *Proceedings of the National Academy of Sciences of the United States of America*, 105: 9065–9.

Zhou, P. et al. (2020) 'A Pneumonia Outbreak Associated with A New Coronavirus of Probable Bat Origin', *Nature*, 579: 270–3.