Molecular characterization of a new species in the genus *Alphacoronavirus* associated with mink epizootic catarrhal gastroenteritis

Anastasia N. Vlasova,[1] Rebecca Halpin,[2] Shiliang Wang,[2] Elodie Ghedin,[2,3] David J. Spiro[2] and Linda J. Saif[1]

Correspondence
Linda J. Saif
saif.2@osu.edu

[1]Food Animal Health Research Program, Ohio Agricultural Research and Development Center, Ohio State University, 1680 Madison Avenue, Wooster, OH 44691, USA

[2]Viral Genomics Group, The J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, MD 20850, USA

[3]Department of Computational and Systems Biology, Center for Vaccine Research, University of Pittsburgh School of Medicine, 3501 Fifth Avenue, Pittsburgh, PA 15261, USA

A coronavirus (CoV) previously shown to be associated with catarrhal gastroenteritis in mink (*Mustela vison*) was identified by electron microscopy in mink faeces from two fur farms in Wisconsin and Minnesota in 1998. A pan-coronavirus and a genus-specific RT-PCR assay were used initially to demonstrate that the newly discovered mink CoVs (MCoVs) were members of the genus *Alphacoronavirus*. Subsequently, using a random RT-PCR approach, full-genomic sequences were generated that further confirmed that, phylogenetically, the MCoVs belonged to the genus *Alphacoronavirus*, with closest relatedness to the recently identified but only partially sequenced (fragments of the polymerase, and full-length spike, 3c, envelope, nucleoprotein, membrane, 3x and 7b genes) ferret enteric coronavirus (FRECV) and ferret systemic coronavirus (FRSCV). The molecular data presented in this study provide the first genetic evidence for a new coronavirus associated with epizootic catarrhal gastroenteritis outbreaks in mink and demonstrate that MCoVs possess high genomic variability and relatively low overall nucleotide sequence identities (91.7 %) between contemporary strains. Additionally, the new MCoVs appeared to be phylogenetically distant from human (229E and NL63) and other alphacoronaviruses and did not belong to the species *Alphacoronavirus 1*. It is proposed that, together with the partially sequenced FRECV and FRSCV, they comprise a new species within the genus *Alphacoronavirus*.

INTRODUCTION

Mink epizootic catarrhal gastroenteritis (ECG) was first described in 1975 (Larsen & Gorham, 1975), and later several million mink were reported to be affected in different countries (the USA, Canada, Scandinavia, PR China and the former USSR; Gorham *et al.*, 1990). The disease occurs seasonally and at greater frequency in mink of ⩾4 months. Together with high morbidity (approaching 100 %) and low mortality (<5 %), ECG in mink resembles that in ferrets (Gorham *et al.*, 1990; Wise *et al.*, 2006). Usually, infected mink become anorexic and develop mucoid diarrhoea within 2–6 days; however, coronavirus (CoV)-like particles have occasionally been demonstrated in faeces from clinically healthy mink

(Gorham *et al.*, 1990). Due to anorexia, infected mink lose body condition and pelt quality, which is of economic concern to mink producers (Gorham *et al.*, 1990). CoV was suggested and confirmed by electron microscopy to be an aetiological agent of ECG (Gorham *et al.*, 1990; Larsen & Gorham, 1975). Other enteric viruses such as rotavirus, parvovirus and calicivirus were suggested to enhance the severity of the ECG disease complex (Evermann *et al.*, 1983; Macartney *et al.*, 1988; Parrish *et al.*, 1988). Until now, CoV detected in ECG cases has not been isolated or sequenced for further characterization.

As described in the 2009 report of the International Committee on Taxonomy of Viruses (ICTV; http://www.ictvonline.org/virusTaxonomy.asp?version=2009), the family *Coronaviridae* now consists of two subfamilies – *Coronavirinae* and *Torovirinae*. Members of the subfamily *Coronavirinae* are enveloped viruses with a helical capsid and a

positive-sense non-segmented RNA (27–32 kb) genome (Spaan *et al.*, 1988; Tyrrell *et al.*, 1975). The RNA replication machinery possesses low fidelity, resulting in a high mutation rate and broad genomic diversity among the virus progeny, which are known as quasispecies (Domingo *et al.*, 1998). Like other members of the order *Nidovirales*, CoVs produce a set of 3′ nested transcripts with a common short leader sequence at the 5′ terminus (Cavanagh, 1997; Gorbalenya *et al.*, 2006; Spaan *et al.*, 1988). The subfamily *Coronovirinae* contains three genera: *Alphacoronavirus* (former CoV group 1), *Beta-coronavirus* (former group 2) and *Gammacoronavirus* (former group 3), with the species *Alphacoronavirus 1* corresponding to former subgroup 1a and other *Alphacoronavirus* species to former subgroup 1b (González *et al.*, 2003; ICTV 2009 report). The virions are pleomorphic and vary in size from 60 to 220 nm, with the surface spike (S) glycoprotein forming an exterior crown-like structure (Spaan *et al.*, 1988; Tyrrell *et al.*, 1975).

In 2002–2003, severe acute respiratory syndrome (SARS)-CoV emerged in the Guangdong province of China and later affected 29 countries, resulting in more than 8000 cases with at least 700 fatalities (Drosten *et al.*, 2003; Ksiazek *et al.*, 2003; Peiris *et al.*, 2004). SARS-CoV was shown to be of animal origin, with horseshoe bats as a potential natural reservoir (Lau *et al.*, 2005; Li *et al.*, 2005). Palm civets and raccoon dogs were suspected to be intermediate hosts (Guan *et al.*, 2003). It was demonstrated by full-genomic comparative analysis that SARS-like CoVs isolated from palm civets are under strong selective pressure and are genetically most closely related to SARS-CoV strains infecting humans early in the outbreaks (Song *et al.*, 2005). Palm civets are carnivores from the suborder Fissipedia together with raccoon dogs, dogs, cats, raccoons, hyenas, mongooses, bears, skunks, ferrets (*Mustela putorius*) and mink (*Mustela vison*) (Heller *et al.*, 2006). Cats, ferrets and palm civets have all been shown to be susceptible to experimental infection with SARS-CoV Urbani strain (Martina *et al.*, 2003; Wu *et al.*, 2005), and a mink lung cell line (Mv1Lu) was also permissive to SARS-CoV expressing a functional ACE2 receptor for viral entry (Gillim-Ross *et al.*, 2004; Heller *et al.*, 2006; Mossel *et al.*, 2005).

Here, a pan-coronavirus and a genus-specific RT-PCR assay were used to demonstrate that two newly discovered mink CoVs (MCoVs) are members of the genus *Alphacoronavirus*. Generation of full genomic sequences further confirmed that, phylogenetically, these MCoVs belonged to the genus *Alphacoronavirus*. According to the available sequence data [nucleoprotein (N) and S protein amino acid sequences], and together with previous studies (Pratelli *et al.*, 2003; Wise *et al.*, 2006; Wu *et al.*, 2005), our data demonstrated higher genetic diversity among CoVs from carnivores. Due to the crucial role they play in the food chain, carnivores harbouring CoVs may serve as virus reservoirs and contribute to the evolution and emergence of new CoV strains with zoonotic potential.

RESULTS

Identification of novel MCoVs and attempted virus isolation

CoV-like particles were first detected by electron microscopy (EM) in faeces of diarrhoeic mink clinically diagnosed with ECG in 1998. Using pan-coronavirus and alphacoronavirus-specific RT-PCR assays on eight mink faecal samples, we obtained products of the predicted sizes of 452 and 390 bp for the polymerase and N gene regions, respectively. After direct sequencing of the PCR products, a BLAST search showed the sequences to be authentic coronavirus sequences, with closest similarity to the recently identified ferret enteric coronavirus (FRECV) (Wise *et al.*, 2006), and to a lesser extent to transmissible gastroenteritis virus (TGEV), canine coronavirus (CCoV) and feline infectious peritonitis virus (FIPV). These initial findings provide the first genetic evidence that an enteric coronavirus is shed in the diarrhoeal faeces of mink, confirming a previous report suggesting CoV as an aetiological agent of ECG in mink (Gorham *et al.*, 1990). Despite the previous report of serological cross-reactivity between TGEV and MCoV (Have *et al.*, 1992), we were unable to detect CoV cross-reactive antigens in mink faeces using monoclonal or polyclonal antibodies to TGEV by ELISA or Western blotting. Our attempts to isolate CoV from mink faecal samples using a number of cell-culture types successful for other CoVs, including Vero E6, CrFK, ST, HRT-18, A59 and Ma-104 cells among others, were also unsuccessful. This failure was probably due to the absence of viable CoV after sample storage for 11 years; also, MCoVs may grow poorly in cell culture, as was observed previously for type I feline enteric CoV (Dye *et al.*, 2007).

Sequencing, assembly and validation of MCoV genomic sequences

Full-length genome sequences were obtained for two MCoVs (WD1127 from Wisconsin and WD1133 from Minnesota) that originated from two independent ECG outbreaks on fur farms in the USA in 1998. Random RT-PCR and priming (Allander *et al.*, 2005; Djikeng *et al.*, 2008) were used to generate primary sequencing data. Gaps were then closed with unique primers designed on known sequences. The 5′ and 3′ ends of the genomes were defined using a 5′- and 3′-RACE system (Qiagen). Raw sequence reads were trimmed to remove amplicon primer-linker and low-quality sequences. Additional sequencing was performed to ensure fourfold sequence coverage across each genome; no polymorphisms were observed in the two MCoV genomes.

Overall genomic identities and phylogenetic analysis of nucleotide sequences

Comparative sequence analysis based on full genomic sequences confirmed that the MCoVs belonged to the

genus *Alphacoronavirus* and were phylogenetically distant from the recognized alphacoronavirus species and slightly closer to the species *Alphacoronavirus 1*, sharing 64.7–65.4 % and 55.8–57.3 % nucleotide identity with alphacoronavirus 1 and other (unclassified) alphacoronavirus representatives, respectively (Table 1), which is less than other members of the *Alphacoronavirus 1* species share with one another (>80 %). Based on full genomic sequence analysis of the two strains, we propose that these MCoVs be assigned to a new species, *Alphacoronavirus 2*, which should probably also include FRECV and FRSCV, which are closely related to MCoVs (based on available partial sequence information; Table 2) (Figs 1 and 2).

The full-length genomic identity between the two contemporary MCoV strains was relatively low (91.7 %; Table 1) compared with that for the same species of CoVs isolated from ruminants (>98 %; Alekseev *et al.*, 2008) or swine (>96 %; Zhang *et al.*, 2007). However, for CoVs isolated from carnivores (canines and felines), the percentage identity is more variable (Table 3).

Genomic organization of MCoVs

Analysis of the full-length genome sequences revealed that they possessed the genomic organization and structure of known alphacoronaviruses with comparable genome size and similar gene order, 5′-untranslated region and 3′–poly(A) tail (Fig. 3). Based on the partial genomic sequence data available, the closest relative was the recently identified FRECV (Wise *et al.*, 2006). The MCoV genome sizes were 28 915 nt for WD1133 and 28 941 nt for WD1127, with poly(A) tails varying in length between 46 and 88 residues.

The major genes encoding the structural and non-structural proteins were arranged as follows: ORF1a/1b, S, 3c, envelope (E), membrane (M) and N followed by the accessory genes (ORF7a, 3x and 7b) encoding non-structural proteins (nsps) (Fig. 3).

Two long ORFs overlapping by 42 nt were predicted in the MCoV genomes: ORF1a of 12 056 and 12 020 nt for WD1127 and WD1133, respectively, and ORF1b of 8033 nt. The nucleotide sequences in the ORF1a–ORF1b overlapping regions have been proposed to form a pseudoknot tertiary structure that allows ribosomal shift of the reading frame (Brierley *et al.*, 1987) between ORF1a and ORF1b. The slippery site for the ribosomal shift (UUUAAAC) is identical in all CoV genomes sequenced to date. We identified it at genomic positions 12 293–12 299 and 12 257–12 263 for WD1127 and WD1133, respectively.

We also identified in the genomes of the MCoVs the minimal conserved transcription regulatory sequence (TRS), CTAAAC, required for discontinuous synthesis of the nested set of subgenomic RNAs (Budzilowicz *et al.*, 1985; Lai & Cavanagh, 1997; Pasternak *et al.*, 2001; Sawicki & Sawicki, 1998; Snijder *et al.*, 2003; Spaan *et al.*, 1988). It was located upstream of the non-replicase genes (except for ORFs 3x-like and 7b) and, surprisingly, was in the middle of both S genes.

Identification and analysis of the three ORFs downstream of the N gene

Alphacoronaviruses are known to contain an additional ORF (ORF7a) downstream of the N gene (Herrewegh *et al.*, 1995; Vennema *et al.*, 1992b), encoding an accessory small

Table 1. Percentage nucleotide identities between MCoVs and selected CoVs based on full-length genomic sequences

Percentage nucleotide identities between MCoVs and other selected CoVs are highlighted in bold. BCoV, Bovine CoV; IBV, infectious bronchitis virus; MHV, mouse hepatitis virus; PEDV, porcine epidemic diarrhoea virus.

TGEV M6	FIPV 79-1146	PEDV	HCoV NL63	HCoV 229E	MCoV WD1127	MCoV WD1133	HCoV OC43	HCoV HKU1	BCoV Mebus	MHV A59	SARS-CoV Tor2	IBV Beaudette	
	85.4	55.9	56.6	55.4	**65.3**	**65.4**	43.9	42.3	43.9	41.5	43.4	42.6	**TGEV M6**
		55.4	56.4	55.3	**64.7**	**65.1**	43.7	42.2	43.7	41.4	42.9	42.4	**FIPV 79-1146**
			64.6	63.2	**56.2**	**56.1**	43.2	41.7	43.0	41.1	43.0	42.4	**PEDV**
				69.5	**57.3**	**57.3**	45.0	44.6	45.0	42.3	43.8	43.7	**HCoV NL63**
					56.1	**55.8**	43.8	44.9	43.8	41.5	43.3	43.4	**HCoV 229E**
						91.7	42.1	40.7	42.2	39.8	41.1	41.0	**MCoV WD1127**
							42.0	40.7	42.2	39.7	41.0	40.9	**MCoV WD1133**
								71.2	96.1	72.4	51.5	48.1	**HCoV OC43**
									70.6	69.5	49.1	47.5	**HCoV HKU1**
										73	51.9	48.3	**BCoV Mebus**
											48.9	44.7	**MHV A59**
												48	**SARS-CoV Tor2**
													IBV Beaudette

Table 2. Numbers of amino acids (a) and percentage amino acid identities (b) of structural and non-structural alphacoronavirus proteins

Percentage amino acid identities between the two MCoVs are highlighted in bold. NA, Protein was not identified for some CoVs or the sequence (or complete sequence) was not available.

|	MCoV (WD1127/WD1133)	TGEV	CCoV	FIPV	FRECV
(a)					
ORF1a	4018/4006	4018	NA	3956	NA
ORF1b	2678	2680	NA	2680	NA
Spike	1438/1429	1447	1452	1452	1449
ORF3c	247/69*	NA	244	237	247
E	82	82	82	82	82
M	268	262	263	262	263
N	376	382	380	377	374
ORF7a	98	78	101	101	NA
ORF3x/3a	73	82	71	NA	74
ORF7b	204	NA	213	206	184
(b)					
ORF1a	**94.2**	56.1/56.5	NA	56.0/56.5	NA
ORF1b	**97.9**	84.5/84.1	NA	84.6/84.3	NA
Spike	**86.3**	64.3/64.8	61.2/61.8	61.3/61.6	67.3/66.3
ORF3c	**86.5***	NA	54.4/51.5	52.7/49.4	64.0/60.3
E	**96.4**	61.0/61.0	61.0/61.0	56.1/56.1	82.9/80.5
M	**94.4**	69.7/69.7	68.6/69.0	71.2/69.2	81.2/81.2
N	**98.1**	58.6/58.1	56.8/56.5	55.0/55.0	76.2/76.7
ORF7a	**96.0**	40.8/42.1	49.5/48.5	47.5/46.5	NA
ORF3x/3a	**87.8**	8.2/11.0	10.0/14.3	NA	41.9/43.2
ORF7b	**94.0**	NA	41.9/41.9	38.1/39.1	46.7/45.7

*A mutation (deletion) in the WD1133 ORF3c sequence created a frame shift and resulted in a premature stop codon in 3c truncated nsp.

hydrophobic membrane-associated non-structural protein (Tung *et al.*, 1992) (Fig. 3). FIPV and CCoV are known to contain an additional ORF7b (Herrewegh *et al.*, 1995; Vennema *et al.*, 1992b), the product of which appears to be a secretory glycoprotein with no stable association with virions (Vennema *et al.*, 1992a). The FRECV genome has also been shown to contain ORF7b (Wise *et al.*, 2006) (Fig. 3). Additionally, the genome of this recently identified FRECV contained an additional ORF (in place of ORF7a) sharing 23.9 % identity with the 3x pseudogene of CCoV (Insavc-1 strain). TGEV has a counterpart to the CCoV pseudogene in a similar location (between the S and M genes) but with a 92 nt deletion (Horsburgh *et al.*, 1992; Wise *et al.*, 2006) (Fig. 3). We analysed the 3′ end of the MCoV genomes downstream of the N gene and identified three putative ORFs for both strains. The first was a gene corresponding to ORF7a (40.8–49.5 % nucleotide identity with TGEV, FIPV and CCoV ORF7a). The last gene shared 38.5–46.7 % nucleotide identity with ORF7b identified for FIPV, CCoV and FRECV. The short gene between ORF7a and ORF7b shared the highest identity with FRECV ORF3x-like gene (41.9–43.2 %) (Wise *et al.*, 2006), whereas identity with the 3x pseudogene of CCoV strain Insavc-1 was only

10–14.3 %, and with the TGEV ORF3 was even less (8.2–11 %). Thus, the MCoV genomes are organized into 10 ORFs comprising six major genes encoding structural and non-structural proteins or polyproteins (ORF1a, ORF1b, S, E, M and N) and four additional genes, ORF3c, ORF7a, ORF3x-like and ORF7b (Fig. 3).

Amino acid identities and differences in key residues of the putative CoV proteins

ORF1a/1b. Comparison of the predicted polypeptide sequences indicated the presence of two 2 aa deletions in WD1127 and one 16 aa deletion in WD1133, resulting in the ORF1a polypeptide being 12 aa longer in WD1127. Whereas deletions were found only in the highly variable ORF1a N-terminal part, numerous substitutions were scattered throughout the entire replicase gene complex. As has been observed previously for other CoVs, the MCoV ORF1b was more conserved than ORF1a between the two MCoV genomes and with the corresponding sequences of other CoVs. ORF1a amino acid identity between the two MCoVs was remarkably low – only 94.2 % for the two CoVs from the same year. The low amino acid identity with alphacoronavirus 1 (56.5 %) and other alphacoronaviruses
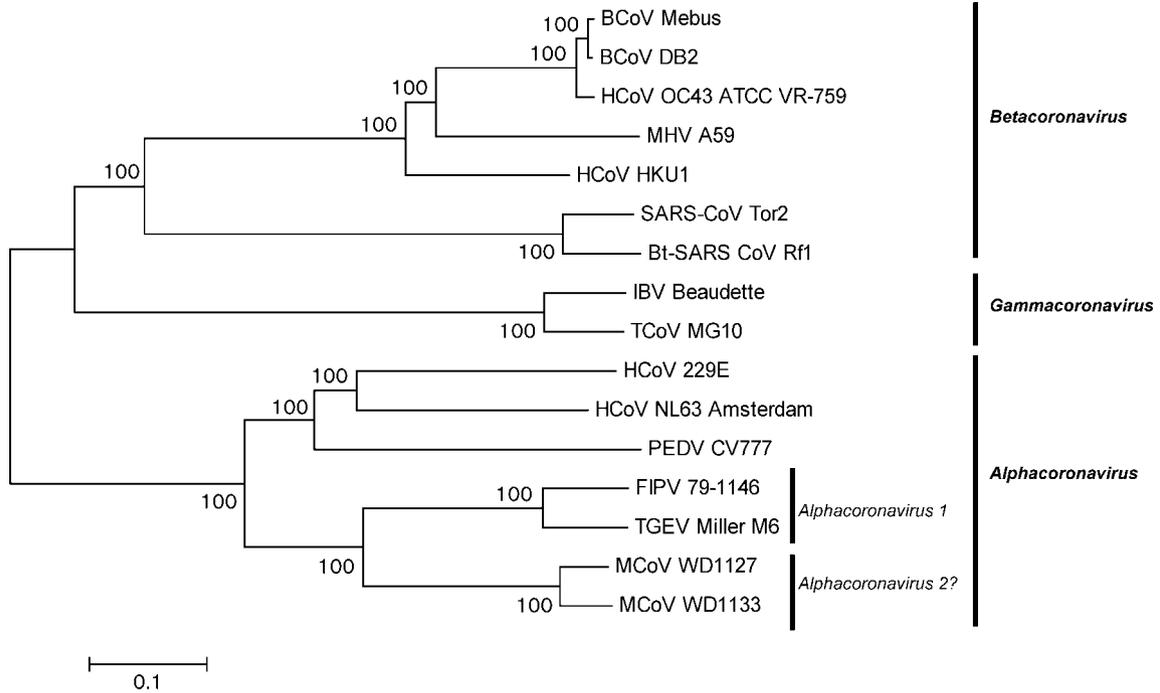
Fig. 1. Neighbour-joining tree of coronaviruses based on full genomic sequences. The tree was inferred using MEGA4. Bootstrap support values >95 % are indicated. Previously defined genera and species and a potential new species (*Alphacoronavirus 2*) are delineated by the bars on the right. The naming of these genera is as described in the 2009 report of the ICTV. Bar, number of nucleotide substitutions per site. TCoV, Turkey coronavirus; see text and tables for other abbreviations.
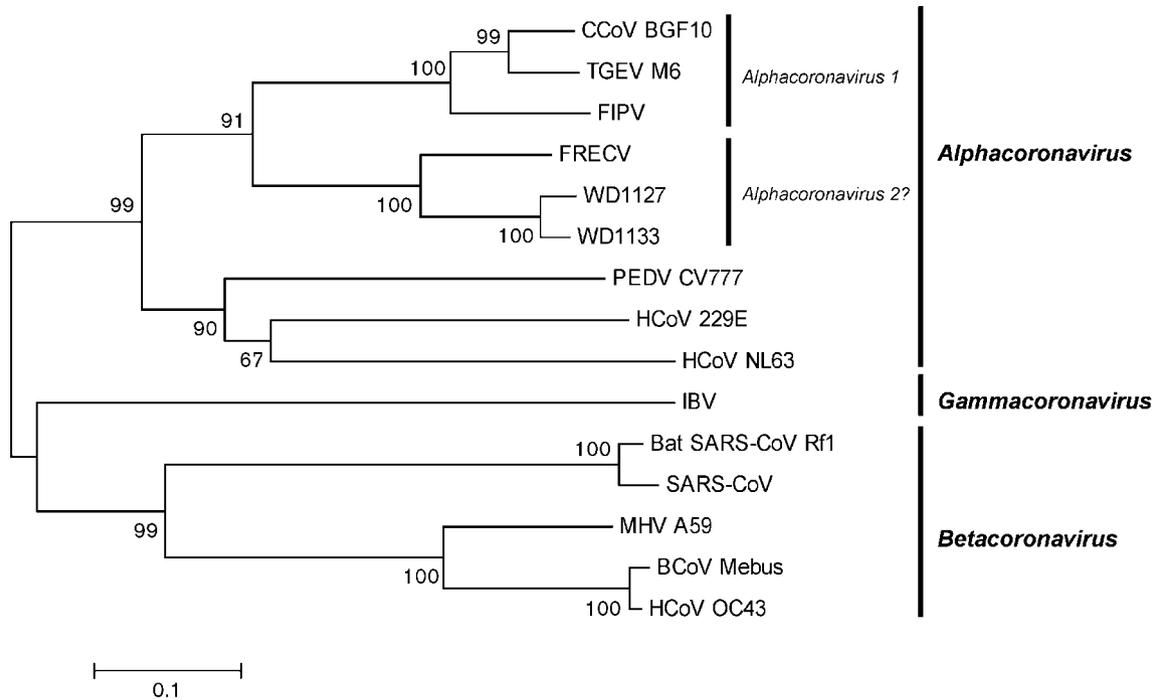
Fig. 2. Neighbour-joining tree of coronaviruses based on N gene sequences. The tree was inferred using MEGA4. Bootstrap support values ⩾90 % are indicated for every node except for that between alphacoronaviruses. Bar, number of nucleotide substitutions per site. See Fig. 1 for abbreviations.

Table 3. Lowest and mean amino acid identities for the N protein of porcine, human, canine and feline alphacoronaviruses

Mean and lowest amino acid identities were defined for at least ten different strains of the same species from the same host. The N protein sequence was chosen as a conserved protein and because its sequence is available for the majority of strains in the genus *Alphacoronavirus*. The lowest and mean amino acid identities for the N protein of FIPV and CCoV are highlighted in bold. PRCV, Porcine respiratory coronavirus; see text and figures for other abbreviations.

Species	Amino acid identity (%)	
	Lowest	Mean
TGEV/PRCV	96.3	98.2
PEDV	95.7	97.5
HCoV NL63	99.4	99.7
HCoV 229E	96.4	98.7
FIPV	**89.6**	**92.0**
CCoV	**91.3**	**96.3**

(38.4–39.7 %) is insufficient to group MCoVs with either species. Whereas based on ORF1a and ORF1b amino acid sequence analysis, the MCoVs seemed to be more closely related to alphacoronavirus 1, all *Alphacoronavirus 1* species members share >80 % amino acid identity. The newly established (ICTV, 2009) species demarcation criterion within each genus has been defined as 90 % amino acid identity in seven conserved replicase domains (including nsp12 and nsp13). Whilst TGEV, FIPV and PRCV share $\geqslant$97 % amino acid identity in these regions, the MCoVs displayed a maximum of 88.3 % amino acid identity (range 73.5–88.3 %) in these regions with alphacoronavirus 1 and thus cannot be allocated to the same species. Other alphacoronaviruses also share a low amino acid identity in this region, similar to that observed for MCoVs. Consequently, it appears that MCoVs occupy an intermediate position within this genus and should be designated a new species.

S protein. We observed a low amino acid identity of 52.5 % in the N-terminal part of the MCoV S proteins (~270 aa), which probably represents a putative hypervariable region analogous to the S1 subunit of other CoVs (e.g. MHV, BCoV and IBV). The rest of the S proteins shared 93.7 % amino acid identity resulting in an overall amino acid identity of 86.3 % between the two MCoV strains (Table 2). In addition to multiple substitutions, there were six short deletions (1–4 aa) in WD1133 and one in WD1127 in the putative hypervariable region.

Genomic comparison with other alphacoronaviruses demonstrated 61.2–61.8 % overall amino acid identity between MCoV, FIPV and CCoV S proteins and 64.3–64.8 % amino acid identity between MCoV and TGEV S proteins (Table 2). An interesting observation was that the first 270 aa shared 46–57 % amino acid identity with
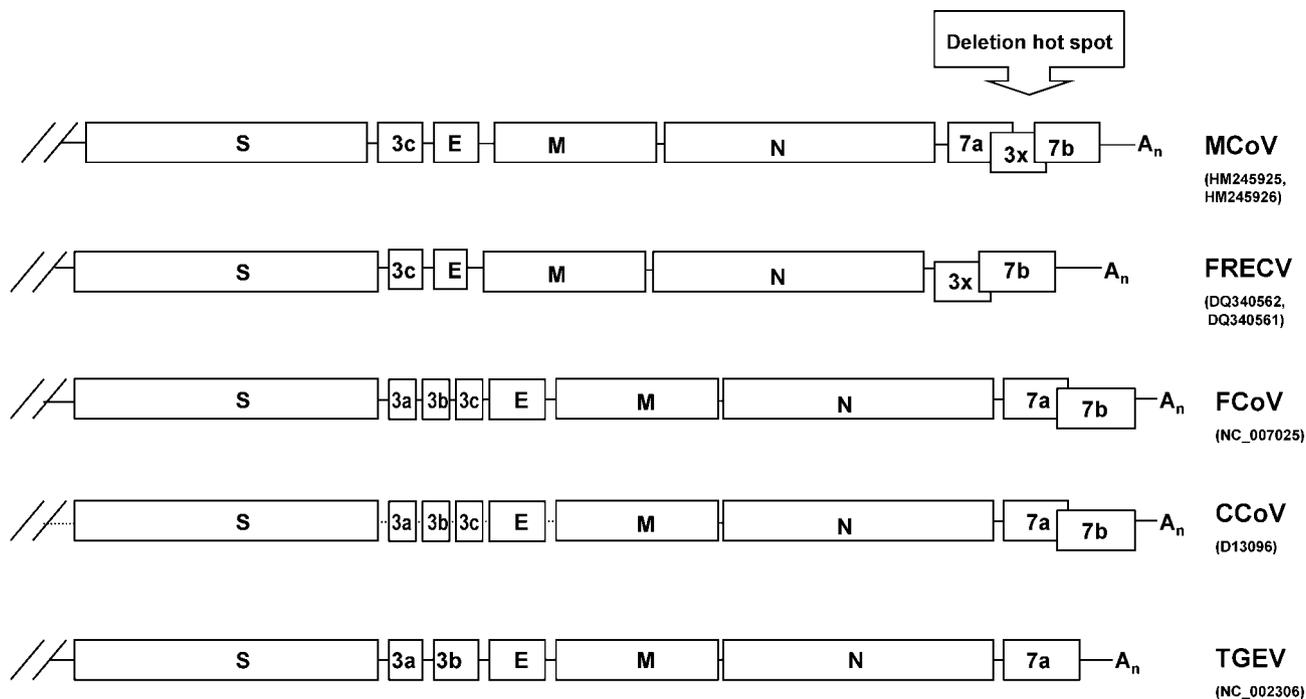
Fig. 3. Schematic diagram of the gene arrangements of the 3′-terminal region of the MCoV, FRECV, FCoV, CCoV and TGEV genomes. The ORFs coding for structural (S, E, M and N) and non-structural (3a/3b/3c, 3x, 7a and 7b) proteins are represented in boxes. Dotted lines represent genomic regions that have not yet been sequenced.

TGEV, but only 20 and 45 % amino acid identity with FIPV and CCoV, respectively. After the first 270 aa, we observed higher (68.2–69.2 %) amino acid identity with FIPV, CCoV and TGEV. Furthermore, we observed 66.3–67.3 % overall amino acid identity between MCoV and FRECV S proteins. For the putative hypervariable region (first 270 aa) of the MCoV and FRECV/ferret systemic coronavirus (FRSCV) S proteins, we observed a low amino acid identity of 34.1–43.7 %, with WD1127 sharing 39.6–43.7 % and WD1133 sharing 34.1–35.9 % amino acid identity. Apart from this, no extensive identity was observed between MCoV and any other CoV S proteins.

ORF3c. ORF3c is an accessory triple-spanning membrane protein analogous to SARS-CoV 3a (Oostra *et al.*, 2006). The predicted ORF3c protein for MCoV WD1133 was 178 aa shorter than for MCoV WD1127 due to a nonsense mutation resulting in a premature stop codon. The amino acid identity between WD1127 and WD1133 ORF3c was 86.5 %, or 73.9 % if considering the truncated WD1133 3c protein. Mutations in WD1133 ORF3c are interesting in view of previous findings on feline and ferret enteric CoVs that acquired high virulence (FIPV) or systemic tropism (FRSCV) and contained various ORF3c sequence alterations including minor and large deletions, insertions and mutations (Chang *et al.*, 2010; Wise *et al.*, 2010). However, we did not observe a higher amino acid identity between the WD1127 or WD1133 ORF3c and the ORF3c from FRECV or FRSCV (data not shown).

E, M and N proteins. The N protein appeared to be the most conserved structural protein between the two MCoV strains with 98.1 % amino acid identity, whilst for the E and M proteins we observed 96.4 and 94.4 % amino acid identity, respectively. No deletions or insertions were observed for these proteins and only 8, 3 and 15 aa substitutions were detected for the N, E and M proteins, respectively. However, the MCoV N proteins differed more when compared with those of TGEV, FIPV and CCoV (55.0–58.6 % amino acid identity) than was observed for the E and M proteins (56.1–71.2 % amino acid identity) (Table 2).

When compared with other alphacoronavirus N proteins, we observed five amino acid deletions common for FRECV and MCoV N proteins (at residues 157–161, 226, 341–343, 375 and 384 of the TGEV N protein) and two insertions in common for the FRECV and MCoV N proteins: a 2 aa insertion between residues 14 and 15 of the TGEV N protein and a 1 aa insertion between residues 204 and 205 of the TGEV N protein. Additionally, we identified a unique 2 aa insertion between residues 359 and 360 of the TGEV N protein.

DISCUSSION

The entire genomes (~29 kb) of two MCoVs from independent outbreaks of ECG on mink farms in the USA were sequenced. To our knowledge, this is the first report of the full genomic sequencing of MCoVs. In view of the lack of sequence data for CoVs from carnivores in public databases, addition of the complete genome sequencing information for the MCoVs will aid in the characterization of animal CoV diversity and contribute to the establishment of new taxonomic units.

Our inability to isolate either of the two MCoVs in cell culture may have been be due to low sample quality or lack of viable CoV after sample storage for 11 years. Alternatively, MCoVs may grow poorly in cell culture, as has been observed previously for type I feline enteric CoV (Dye *et al.*, 2007). To date, no one has reported the successful propagation of MCoVs in cell culture using mink faecal samples. To address this issue, fresh mink faecal samples containing viable MCoVs or cell cultures of mink origin may be needed.

Phylogenetic analysis based on full-length genome sequences clearly demonstrated that the MCoVs belonged to the genus *Alphacoronavirus* (Fig. 1). Based on the limited sequence data available (excluding most of ORF1a/1b), the closest relatedness observed was to FRECV. However, the taxonomic position of FRECV among other alphacoronaviruses has not yet been clearly affirmed (Wise *et al.*, 2006) (Fig. 2). The MCoVs appeared to be more closely related to alphacoronavirus 1 than to other alphacoronaviruses (with pairwise nucleotide sequence identities of 64.7–65.4 and 55.8–57.3 %, respectively). However, MCoV amino acid identity to alphacoronavirus 1 in seven conserved replicase domains did not reach the newly established threshold of 90 % to be considered the same species. Thus, it has yet to be determined whether MCoVs alone or together with FRECV should form (as we propose) a new species (*Alphacoronavirus 2*) within the genus *Alphacoronavirus* (Fig. 1).

The lower genetic identity (91.7 %) between the two contemporary MCoV strains, together with previous sequencing data for canine and feline alphacoronaviruses (Table 3), demonstrate higher genomic diversity among CoVs isolated from carnivores compared with those isolated from herbivores and omnivores. This diversity may pose a greater potential for interspecies transmission and adaptation to new species, as was observed for SARS-CoV introduced by palm civets into the human population (Guan *et al.*, 2003; Kan *et al.*, 2005; Wang *et al.*, 2005).

In vitro experiments have demonstrated that frameshifting between ORF1a and ORF1b of CoV genomes occurs in approximately 20–30 % of translations (Ziebuhr, 2005), regulating the molar ratio of ORF1a and readthrough product ORF1ab. The more conserved ORF1b (compared with variable ORF1a) encodes core replicative enzymes (RNA-dependent RNA polymerase and helicase), whilst the main cysteine and accessory proteinases are encoded by ORF1a. Phylogenetic analysis of ORF1a and ORF1b polyproteins revealed the same relatedness of the MCoVs

within the genus – an intermediate position between *Alphacoronavirus 1* and other alphacoronavirus species or an early split off from *Alphacoronavirus 1* (data not shown).

Furthermore, comparative phylogenetic analysis of the S, E, M and N protein sequences and small accessory ORFs of the MCoVs demonstrated similar results to that based on the full genomic nucleotide sequences, further supporting classification of MCoV as an alphacoronavirus, with the highest levels of similarity to FRECV, followed by TGEV, CCoV and FIPV. Thus, the extent of genetic relatedness between MCoV and other alphacoronaviruses appears to be consistent, providing no evidence for recent recombination events or mosaic evolution of new CoVs from mink.

It has been suggested that several genes are associated with differences in pathogenicity, including the S gene and the accessory genes 3a, 3b, 3c, 7a and 7b (Kennedy *et al.*, 2001; Park *et al.*, 2008; Penzes *et al.*, 2001; Rottier *et al.*, 2005; Vennema *et al.*, 1998; Woods, 2001). Loss of 3c function was suggested previously to correlate with an increased FIPV virulence or acquisition of systemic tropism by FRSCV (Haijema *et al.*, 2004; Pedersen, 2009; Vennema *et al.*, 1998; Wise *et al.*, 2010). The FIPV and FRSCV strains carry mutations or large deletions inactivating the gene for 3c (Pedersen, 2009; Vennema *et al.*, 1998; Wise *et al.*, 2010), which was suggested to be strictly required for replication in gut tissues but dispensable for systemic replication. Considering a significant truncation in WD1133 ORF3c, it would be of interest to inoculate mink experimentally with MCoV WD1127 and WD1133 to see whether there are any differences in clinical signs or pathogenicity. In feline and canine alphacoronaviruses, the ORF3 region is represented by ORF3a, 3b and 3c, whereas in porcine alphacoronaviruses (TGEV and PRCV) ORF3c is missing and the ferret and MCoVs lack ORF3a and ORF3b. Interestingly, it was reported previously that, for most porcine alphacoronaviruses (TGEV, PRCV and PEDV), deletions in ORF3a and ORF3b correlate with attenuated virus phenotype (Izeta *et al.*, 1999; Park *et al.*, 2008; Penzes *et al.*, 2001; Woods, 2001; Zhang *et al.*, 2007), which is in contrast to the effect of mutations in the ORF3c region of alphacoronaviruses from carnivores. These data confirm the dynamic genetic aspects of the coronavirus ORF3 region and different functions associated with each component – 3a, 3b and 3c – that should be investigated in more detail.

The unique number and arrangement of additional small ORFs (7a, 3x-like and 7b) downstream of the N protein in MCoVs appear to be the same in both strains. Although all of these ORFs or their counterparts have been found in other CoVs (FRECV, TGEV, CCoV and FIPV) in various combinations and in different regions of the genomes, none was reported to be essential for virus replication *in vitro* (Herrewegh *et al.*, 1995; Vennema *et al.*, 1992a, b). In the CCoV and TGEV genomes, ORF3x/3 is located between the S and M genes; however, TGEV (Purdue-115 and FS772/70 strains) ORF3 contains a deletion of 92 nt (Horsburgh *et al.*, 1992). In FRECV, an ORF with 23.9 % identity to the 3x

pseudogene of CCoV Insavc-1 (Horsburgh *et al.*, 1992) was found in the genomic location of ORF7a, whilst the latter was missing (Wise *et al.*, 2006). Based on previous data, it was suggested that ORF3x is an evolutionarily redundant sequence that does not appear to encode a functional viral protein, and which could be the result of an insertional event in CCoV and FRECV (Horsburgh *et al.*, 1992; Wise *et al.*, 2006). ORF7a encoding a small hydrophobic membrane-associated protein (Tung *et al.*, 1992) was found in FIPV, TGEV, CCoV and MCoV at exactly the same position downstream of the N protein; however, it was missing in FRECV (Wise *et al.*, 2006). ORF7b, previously suggested to encode a secretory glycoprotein serving as a mediator for host immune responses (Herrewegh *et al.*, 1995), was present in all the aforementioned alphacoronaviruses except for TGEV (Herrewegh *et al.*, 1995; Vennema *et al.*, 1992b; Wise *et al.*, 2006). This comparative analysis confirmed that both of these accessory proteins are probably dispensable for CoV replication and possibly pathogenicity. The genomic region downstream of the N protein is known to be a 'deletion hot spot' (Collisson *et al.*, 1990; De Groot *et al.*, 1988; Horsburgh *et al.*, 1992) with frequent deletion and insertion events making the presence of all three ORFs coding for accessory proteins in the MCoV genomes very interesting. To explain these findings, more sequencing data should be generated for other MCoVs, both historical and recent strains.

The TRS (CTAAAC) upstream of the non-replicase genes in the CoV genome is hypothesized to direct discontinuous synthesis of negative-sense subgenomic mRNAs serving as templates for subgenomic mRNA synthesis (Pasternak *et al.*, 2001; Sawicki & Sawicki, 1998). The conserved TRS was found in FRECV upstream of the 3x-like ORF and in CCoV Insavc-1 upstream of ORF7a but not upstream of ORF7b, leading to the assumption that ORF7a/7b (3x-like/7b) are probably being expressed from polycistronic mRNAs (Horsburgh *et al.*, 1992; Wise *et al.*, 2006), as polycistronic CoV mRNAs have been identified previously (Liu *et al.*, 1991; Liu & Inglis, 1992). Therefore, the MCoVs 3′-terminal genes (7a, 3x-like and 7b) can probably also be expressed from polycistronic mRNA.

In conclusion, our study provides the first genomic evidence and molecular confirmation for a CoV in mink. This CoV was previously suggested to be an aetiological agent of mink ECG and can now be classified as a potential new species (*Alphacoronavirus 2*) of the diverse genus *Alphacoronavirus*. Whether the new *Alphacoronavirus 2* species proposed in this manuscript includes mink and ferret CoVs or only MCoV will be better defined after the full genomic sequence for FRECV/FRSCV is available.

METHODS

History of mink faecal samples. Eight mink faecal samples were submitted to our laboratory in 1998. The samples from diarrhoeal animals in two fur farms in Minnesota and Wisconsin were designated WD1126–WD1133. All faecal samples were shown to be

positive for CoV by electron microscopy and in our pan-coronavirus and alphacoronavirus-specific reverse RT-PCR assays.

EM. EM was carried out as described previously (Saif *et al.*, 1991). Briefly, the faecal samples were diluted as 20 % suspensions in minimal essential medium containing 1 % antibiotics and 1 % non-essential amino acids (Gibco) and clarified by centrifugation. The supernatants were filtered sequentially through 0.8, 0.45 and 0.2 μm syringe filters to remove bacteria and other contamination. The CoVs were pelleted by ultracentrifugation (35 000 *g* for 30 min). The pellets were then mixed with filtered distilled water and an equal volume of 3 % phosphotungstic acid (pH 7.0) and placed on Formvar carbon-coated grids. Specimens were evaluated in an electron microscope (Philips 201; Norelco).

RT-PCR. Total RNA was extracted from clarified and filtered faecal samples using an RNeasy Mini kit (Qiagen) according to the manufacturer's instructions. A one-step RT-PCR assay was performed as described previously (Hasoksuz *et al.*, 2007). The primers used in RT-PCR were designed from the published sequence of the polymerase and N genes of the CoV strains. The following primer pairs were designed or modified and used for genome detection of MCoVs: pan-coronavirus universal primers IN-2deg (5′-GGGDTG-GGAYTAYCCHAARTGYGA-3′, forward) and IN-4deg (5′-TARCA-VACAACISYRTCRTCA-3′, reverse), targeting a 452 bp fragment of the polymerase gene (modified from Ksiazek *et al.*, 2003); and alphacoronavirus-specific primers Gr1F (5′-GADGGWGTYKTCT-GGGTTGC-3′, forward) and Gr1R1 (5′-GTTYTCTTCCAGGTGT-GTTTG-3′, reverse) capable of detecting of all alphacoronaviruses, targeting a 390 bp fragment of the nucleoprotein gene. Samples WD1126 and WD1133 were used for full-genome sequencing using primers Gr1F and Gr1R1.

Sequencing. RNA was extracted as described above. A random RT-PCR protocol (Djikeng *et al.*, 2008) was used to generate assemblies; gaps were dosed with unique primers designed based on initial sequence data. Primers were designed every 500 bp along the genome. An M13 sequence tag was added to the 5′ end of each primer to be used for sequencing (5′-TGTAAAACGACGGCCAGT-3′ for forward primers and 5′-CAGGAAACAGCTATGACC-3′ for reverse primers). Primer sequences are available from the authors on request. RT-PCRs were performed with 50–200 ng CoV RNA using a Qiagen OneStep RT-PCR kit (Qiagen) according to the manufacturer's instructions. Duplicate reactions were analysed for quality control purposes by agarose gel electrophoresis. Amplicons were prepared for sequencing by incubation at 37 °C for 60 min with 0.5 U shrimp alkaline phosphatase (USB) and 1 U exonuclease I (USB) to inactivate remaining dNTPs and to digest the single-stranded primers. The enzymes were inactivated by incubation at 72 °C for 15 min.

Sequencing reactions were performed on a standard high-throughput sequencing system using BigDye Terminator chemistry (Applied Biosystems) with 2 μl template cDNA. Each amplicon was sequenced from each end using M13 primers described above. Sequencing reactions were analysed on a 3730xl ABI sequencer (Applied Biosystems).

Sequencing reads were downloaded, trimmed to remove the amplicon primer-linker sequence as well as low-quality sequence, assembled using Minimus, part of the open-source AMOS (Pop *et al.*, 2004) project (http://amos.sourceforge.net) and edited using AutoEditor (Gajer *et al.*, 2004), as well as by manual curation using CloE (Closure Editor, http://cloe.sourceforge.net). To close gaps between assembled contigs, strain-specific primers were designed, RT-PCRs were performed and amplicons were sequenced as described above. Additional primer design, cDNA synthesis and sequencing were performed to ensure greater than fourfold sequence coverage along the CoV genomes.

Table 4. GenBank accession numbers for sequences used for phylogenetic trees construction based on the complete genome and the N gene

Virus	GenBank no.
Alphacoronaviruses	
HCoV 229E	NC_002645
HCoV NL63	NC_005831
PEDV CV777	NC_003436
FIPV 79-1146	NC_007025
CCoV BGF10	AY342160*
FRECV	DQ340562*
TGEV M6	DQ811785
Betacoronaviruses	
HCoV OC43 ATCC VR-759	AY391777
HCoV HKU1	NC_006577
MHV A59	NC_001846
BCoV Mebus	U00735
BCoV DB2	DQ811784
SARS-CoV Tor2	NC_004718
Bat SARS-CoV Rf1	NC_009695
Gammacoronaviruses	
IBV Beaudette	NC_001451
TCoV MG10	NC_010800

*The complete genome sequence was not analysed.

All apparent polymorphisms were checked against reference data, and ambiguities were analysed by RT-PCR and cloning. Each assembly was analysed using Viral Genome ORF Reader (VIGOR) (Wang *et al.*, 2010), a program designed to predict viral protein sequences. VIGOR checked segment length, alignments with reference sequences and fidelity of reading frames, correlated amino acid mutations with nucleotide polymorphisms and detected potential sequence errors.

Sequence analyses. The CoV genome references downloaded from GenBank and used in the phylogenetic analyses are listed in Table 4. Sequence alignment and phylogenetic analysis were performed using the CLUSTAL W method of the Lasergene Biocomputing Software (DNASTAR) and MEGA4. The MCoV sequences were compared with the human and animal CoV strains in GenBank. The deduced amino acid sequences were then assembled and analysed using the MEGALIGN module of the Lasergene Biocomputing Software.

ACKNOWLEDGEMENTS

REFERENCES

Alekseev, K. P., Vlasova, A. N., Jung, K., Hasoksuz, M., Zhang, X., Halpin, R., Wang, S., Ghedin, E., Spiro, D. & Saif, L. J. (2008). Bovine-like coronaviruses isolated from four species of captive wild ruminants are homologous to bovine coronaviruses, based on complete genomic sequences. *J Virol* **82**, 12422–12431.

Allander, T., Tammi, M. T., Eriksson, M., Bjerkner, A., Tiveljung-Lindell, A. & Andersson, B. (2005). Cloning of a human parvovirus by molecular screening of respiratory tract samples. *Proc Natl Acad Sci U S A* **102**, 12891–12896.

Brierley, I., Boursnell, M. E., Binns, M. M., Bilimoria, B., Blok, V. C., Brown, T. D. & Inglis, S. C. (1987). An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J* **6**, 3779–3785.

Budzilowicz, C. J., Wilczynski, S. P. & Weiss, S. R. (1985). Three intergenic regions of coronavirus mouse hepatitis virus strain A59 genome RNA contain a common nucleotide sequence that is homologous to the 3′ end of the viral mRNA leader sequence. *J Virol* **53**, 834–840.

Cavanagh, D. (1997). *Nidovirales*: a new order comprising *Coronaviridae* and *Arteriviridae*. *Arch Virol* **142**, 629–633.

Chang, H.-W., de Groot, R. J., Egberink, H. F. & Rottier, P. J. (2010). Feline infectious peritonitis: insights into feline coronavirus pathobiogenesis and epidemiology based on genetic analysis of the viral 3c gene. *J Gen Virol* **91**, 415–420.

Collisson, E. W., Williams, A. K., Vonder Haar, R., Li, W. & Sneed, L. W. (1990). Sequence comparisons of the 3′ end of the genomes of five strains of avian infectious bronchitis virus. *Adv Exp Med Biol* **276**, 373–377.

De Groot, R. J., Andeweg, A. C., Horzinek, M. C. & Spaan, W. J. (1988). Sequence analysis of the 3′-end of the feline coronavirus FIPV 79-1146 genome: comparison with the genome of porcine coronavirus TGEV reveals large insertions. *Virology* **167**, 370–376.

Djikeng, A., Halpin, R., Kuzmickas, R., Depasse, J., Feldblyum, J., Sengamalay, N., Afonso, C., Zhang, X., Anderson, N. G. & other authors (2008). Viral genome sequencing by random priming methods. *BMC Genomics* **9**, 5.

Domingo, E., Baranowski, E., Ruiz-Jarabo, C. M., Martín-Hernández, A. M., Sáiz, J. C. & Escarmís, C. (1998). Quasispecies structure and persistence of RNA viruses. *Emerg Infect Dis* **4**, 521–527.

Drosten, C., Preiser, W., Günther, S., Schmitz, H. & Doerr, H. W. (2003). Severe acute respiratory syndrome: identification of the etiological agent. *Trends Mol Med* **9**, 325–327.

Dye, C., Temperton, N. & Siddell, S. G. (2007). Type I feline coronavirus spike glycoprotein fails to recognize aminopeptidase N as a functional receptor on feline cell lines. *J Gen Virol* **88**, 1753–1760.

Evermann, J. F., Smith, A. W., Skilling, D. E. & McKeirnan, A. J. (1983). Ultrastructure of newly recognized caliciviruses of the dog and mink. *Arch Virol* **76**, 257–261.

Gajer, P., Schatz, M. & Salzberg, S. L. (2004). Automated correction of genome sequence errors. *Nucleic Acids Res* **32**, 562–569.

Gillim-Ross, L., Taylor, J., Scholl, D. R., Ridenour, J., Masters, P. S. & Wentworth, D. E. (2004). Discovery of novel human and animal cells infected by the severe acute respiratory syndrome coronavirus by replication-specific multiplex reverse transcription-PCR. *J Clin Microbiol* **42**, 3196–3206.

González, J. M., Gomez-Puertas, P., Cavanagh, D., Gorbalenya, A. E. & Enjuanes, L. (2003). A comparative sequence analysis to revise the current taxonomy of the family *Coronaviridae*. *Arch Virol* **148**, 2207–2235.

Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. & Snijder, E. J. (2006). *Nidovirales*: evolving the largest RNA virus genome. *Virus Res* **117**, 17–37.

Gorham, J. R., Evermann, J. F., Ward, A., Pearson, R., Shen, D., Hartsough, G. R. & Leathers, C. (1990). Detection of coronavirus-like particles from mink with epizootic catarrhal gastroenteritis. *Can J Vet Res* **54**, 383–384.

Guan, Y., Zheng, B. J., He, Y. Q., Liu, X. L., Zhuang, Z. X., Cheung, C. L., Luo, S. W., Li, P. H., Zhang, L. J. & other authors (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* **302**, 276–278.

Haijema, B. J., Volders, H. & Rottier, P. J. (2004). Live, attenuated coronavirus vaccines through the directed deletion of group-specific genes provide protection against feline infectious peritonitis. *J Virol* **78**, 3863–3871.

Hasoksuz, M., Alekseev, K., Vlasova, A., Zhang, X., Spiro, D., Halpin, R., Wang, S., Ghedin, E. & Saif, L. J. (2007). Biologic, antigenic, and full-length genomic characterization of a bovine-like coronavirus isolated from a giraffe. *J Virol* **81**, 4981–4990.

Have, P., Moving, V., Svansson, V., Uttenthal, A. & Bloch, B. (1992). Coronavirus infection in mink (*Mustela vison*). Serological evidence of infection with a coronavirus related to transmissible gastroenteritis virus and porcine epidemic diarrhea virus. *Vet Microbiol* **31**, 1–10.

Heller, L. K., Gillim-Ross, L., Olivieri, E. R. & Wentworth, D. E. (2006). *Mustela vison* ACE2 functions as a receptor for SARS-coronavirus. *Adv Exp Med Biol* **581**, 507–510.

Herrewegh, A. A., Vennema, H., Horzinek, M. C., Rottier, P. J. & de Groot, R. J. (1995). The molecular genetics of feline coronaviruses: comparative sequence analysis of the ORF7a/7b transcription unit of different biotypes. *Virology* **212**, 622–631.

Horsburgh, B. C., Brierley, I. & Brown, T. D. (1992). Analysis of a 9.6 kb sequence from the 3′ end of canine coronavirus genomic RNA. *J Gen Virol* **73**, 2849–2862.

Izeta, A., Smerdou, C., Alonso, S., Penzes, Z., Mendez, A., Plana-Durán, J. & Enjuanes, L. (1999). Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. *J Virol* **73**, 1535–1545.

Kan, B., Wang, M., Jing, H., Xu, H., Jiang, X., Yan, M., Liang, W., Zheng, H., Wan, K. & other authors (2005). Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. *J Virol* **79**, 11892–11900.

Kennedy, M., Boedeker, N., Gibbs, P. & Kania, S. (2001). Deletions in the 7a ORF of feline coronavirus associated with an epidemic of feline infectious peritonitis. *Vet Microbiol* **81**, 227–234.

Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J. A. & other authors (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**, 1953–1966.

Lai, M. M. & Cavanagh, D. (1997). The molecular biology of coronaviruses. *Adv Virus Res* **48**, 1–100.

Larsen, A. E. & Gorham, J. R. (1975). A new mink enteritis: an initial report. *Vet Med Small Anim Clin* **70**, 291–292.

Lau, S. K., Woo, P. C., Li, K. S., Huang, Y., Tsoi, H.-W., Wong, B. H., Wong, S. S., Leung, S.-Y., Chan, K.-H. & Yuen, K.-Y. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc Natl Acad Sci U S A* **102**, 14040–14045.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J. H., Wang, H., Crameri, G., Hu, Z. & other authors (2005). Bats are natural reservoirs of SARS-like coronaviruses. *Science* **310**, 676–679.

Liu, D. X. & Inglis, S. C. (1992). Identification of two new polypeptides encoded by mRNA5 of the coronavirus infectious bronchitis virus. *Virology* **186**, 342–347.

Liu, D. X., Cavanagh, D., Green, P. & Inglis, S. C. (1991). A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. *Virology* **184**, 531–544.

Macartney, L., Parrish, C. R., Binn, L. N. & Carmichael, L. E. (1988). Characterization of minute virus of canines (MVC) and its pathogenicity for pups. *Cornell Vet* **78**, 131–145.

Martina, B. E., Haagmans, B. L., Kuiken, T., Fouchier, R. A., Rimmelzwaan, G. F., Van Amerongen, G., Peiris, J. S., Lim, W. & Osterhaus, A. D. (2003). Virology: SARS virus infection of cats and ferrets. *Nature* **425**, 915.

Mossel, E. C., Huang, C., Narayanan, K., Makino, S., Tesh, R. B. & Peters, C. J. (2005). Exogenous ACE2 expression allows refractory cell lines to support severe acute respiratory syndrome coronavirus replication. *J Virol* **79**, 3846–3850.

Oostra, M., de Haan, C. A., de Groot, R. J. & Rottier, P. J. (2006). Glycosylation of the severe acute respiratory syndrome coronavirus triple-spanning membrane proteins 3a and M. *J Virol* **80**, 2326–2336.

Park, S.-J., Moon, H.-J., Luo, Y., Kim, H.-K., Kim, E.-M., Yang, J.-S., Song, D.-S., Kang, B.-K., Lee, C.-S. & Park, B.-K. (2008). Cloning and further sequence analysis of the ORF3 gene of wild- and attenuated-type porcine epidemic diarrhea viruses. *Virus Genes* **36**, 95–104.

Parrish, C. R., Have, P., Foreyt, W. J., Evermann, J. F., Senda, M. & Carmichael, L. E. (1988). The global spread and replacement of canine parvovirus strains. *J Gen Virol* **69**, 1111–1116.

Pasternak, A. O., van den Born, E., Spaan, W. J. & Snijder, E. J. (2001). Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. *EMBO J* **20**, 7220–7228.

Pedersen, N. C. (2009). A review of feline infectious peritonitis virus infection: 1963-2008. *J Feline Med Surg* **11**, 225–258.

Peiris, J. S., Guan, Y. & Yuen, K. Y. (2004). Severe acute respiratory syndrome. *Nat Med* **10** (Suppl), S88–S97.

Penzes, Z., Gonzalez, J. M., Calvo, E., Izeta, A., Smerdou, C., Méndez, A., Sanchez, C. M., Sola, I., Almazan, F. & Enjuanes, L. (2001). Complete genome sequence of transmissible gastroenteritis coronavirus PUR46-MAD clone and evolution of the Purdue virus cluster. *Virus Genes* **23**, 105–118.

Pop, M., Phillippy, A., Delcher, A. L. & Salzberg, S. L. (2004). Comparative genome assembly. *Brief Bioinform* **5**, 237–248.

Pratelli, A., Martella, V., Decaro, N., Tinelli, A., Camero, M., Cirone, F., Elia, G., Cavalli, A., Corrente, M. & other authors (2003). Genetic diversity of a canine coronavirus detected in pups with diarrhoea in Italy. *J Virol Methods* **110**, 9–17.

Rottier, P. J., Nakamura, K., Schellen, P., Volders, H. & Haijema, B. J. (2005). Acquisition of macrophage tropism during the pathogenesis of feline infectious peritonitis is determined by mutations in the feline coronavirus spike protein. *J Virol* **79**, 14122–14130.

Saif, L. J., Brock, K. V., Redman, D. R. & Kohler, E. M. (1991). Winter dysentery in dairy herds: electron microscopic and serological evidence for an association with coronavirus infection. *Vet Rec* **128**, 447–449.

Sawicki, S. G. & Sawicki, D. L. (1998). A new model for coronavirus transcription. *Adv Exp Med Biol* **440**, 215–219.

Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L., Guan, Y., Rozanov, M., Spaan, W. J. & Gorbalenya, A. E. (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J Mol Biol* **331**, 991–1004.

Song, H.-D., Tu, C.-C., Zhang, G.-W., Wang, S.-Y., Zheng, K., Lei, L.-C., Chen, Q.-X., Gao, Y. -W., Zhou, H.-Q. & other authors (2005). Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proc Natl Acad Sci U S A* **102**, 2430–2435.

Spaan, W., Cavanagh, D. & Horzinek, M. C. (1988). Coronaviruses: structure and genome expression. *J Gen Virol* **69**, 2939–2952.

Tung, F. Y., Abraham, S., Sethna, M., Hung, S. L., Sethna, P., Hogue, B. G. & Brian, D. A. (1992). The 9-kDa hydrophobic protein encoded at the 3′ end of the porcine transmissible gastroenteritis coronavirus genome is membrane-associated. *Virology* **186**, 676–683.

Tyrrell, D. A., Almeida, J. D., Cunningham, C. H., Dowdle, W. R., Hofstad, M. S., McIntosh, K., Tajima, M., Zakstelskaya, L. Y., Easterday, B. C. & other authors (1975). *Coronaviridae. Intervirology* **5**, 76–82.

Vennema, H., Heijnen, L., Rottier, P. J., Horzinek, M. C. & Spaan, W. J. (1992a). A novel glycoprotein of feline infectious peritonitis coronavirus contains a KDEL-like endoplasmic reticulum retention signal. *J Virol* **66**, 4951–4956.

Vennema, H., Rossen, J. W., Wesseling, J., Horzinek, M. C. & Rottier, P. J. (1992b). Genomic organization and expression of the 3′ end of the canine and feline enteric coronaviruses. *Virology* **191**, 134–140.

Vennema, H., Poland, A., Foley, J. & Pedersen, N. C. (1998). Feline infectious peritonitis viruses arise by mutation from endemic feline enteric coronaviruses. *Virology* **243**, 150–157.

Wang, M., Yan, M., Xu, H., Liang, W., Kan, B., Zheng, B., Chen, H., Zheng, H., Xu, Y. & other authors (2005). SARS-CoV infection in a restaurant from palm civet. *Emerg Infect Dis* **11**, 1860–1865.

Wang, S., Sundaram, J. P. & Spiro, D. (2010). VIGOR, an annotation program for small viral genomes. *BMC Bioinformatics* **11**, 451.

Wise, A. G., Kiupel, M. & Maes, R. K. (2006). Molecular characterization of a novel coronavirus associated with epizootic catarrhal enteritis (ECE) in ferrets. *Virology* **349**, 164–174.

Wise, A. G., Kiupel, M., Garner, M. M., Clark, A. K. & Maes, R. K. (2010). Comparative sequence analysis of the distal one-third of the genomes of a systemic and an enteric ferret coronavirus. *Virus Res* **149**, 42–50.

Woods, R. D. (2001). Efficacy of a transmissible gastroenteritis coronavirus with an altered ORF-3 gene. *Can J Vet Res* **65**, 28–32.

Wu, D., Tu, C., Xin, C., Xuan, H., Meng, Q., Liu, Y., Yu, Y., Guan, Y., Jiang, Y. & other authors (2005). Civets are equally susceptible to experimental infection by two different severe acute respiratory syndrome coronavirus isolates. *J Virol* **79**, 2620–2625.

Zhang, X., Hasoksuz, M., Spiro, D., Halpin, R., Wang, S., Stollar, S., Janies, D., Hadya, N., Tang, Y. & other authors (2007). Complete genomic sequences, a key residue in the spike protein and deletions in nonstructural protein 3b of US strains of the virulent and attenuated coronaviruses, transmissible gastroenteritis virus and porcine respiratory coronavirus. *Virology* **358**, 424–435.

Ziebuhr, J. (2005). The coronavirus replicase. *Curr Top Microbiol Immunol* **287**, 57–94.