Short Communication

Inter- and intra-variant genetic heterogeneity of human coronavirus OC43 strains in France

Astrid Vabret,[1] Julia Dina,[1] Thomas Mourez,[1] Stéphanie Gouarin,[1] Joëlle Petitjean,[1] Sylvie van der Werf[2] and François Freymuth[1]

[1]Laboratory of Virology, University Hospital of Caen, Avenue Georges Clemenceau, 14033 Caen cedex, France

[2]Laboratoire de Génétique Moléculaire des Virus Respiratoires, Institut Pasteur, 25–28 rue du Docteur Roux, 75015 Paris, France

Correspondence
Astrid Vabret
vabret-a@chu-caen.fr

Received 23 March 2006
Accepted 30 June 2006

Human coronavirus OC43 (HCoV-OC43) causes acute, self-limited respiratory infections. A close relationship between bovine coronaviruses (BCoVs) and HCoV-OC43 has recently been demonstrated. This study includes seven clinical, non-cell culture-adapted, contemporary HCoV-OC43 strains detected in France in 2003. By using RT-PCR and clonal sequencing of the S1 gene of HCoV-OC43, the inter-variant heterogeneity of the HCoV-OC43 circulating strains was studied and the intra-variant diversity was assessed by investigation of a quasispecies cloud. This paper brings to the forefront a high genetic diversity of circulating HCoV-OC43 variants. Genetically different groups are defined among the variants described in this study. One of these variants holds characteristics of an outlier and presents a deletion of 12 nt, also found in BCoV strains. Moreover, the presence of HCoV-OC43 as a quasispecies cloud *in vivo* during an acute respiratory-tract illness was discovered. It has also been revealed that quasispecies-cloud sizes are similar for the two viral populations tested.

Coronaviruses (CoVs) occupy an important position in virology, being not only the cause of the severe acute respiratory sundrome (SARS) outbreak – the first human emergent infectious disease of the 21st century – but also a significant pathogen involved in many worldwide respiratory-tract infections. CoVs have been found in many mammalian and avian species, causing acute, chronic or persistent infections. They are enveloped viruses with a linear, non-segmented, positive-sense, single-stranded RNA genome of 27–31 kb in length (Cavanagh, 1997). Among the structural viral proteins, the spike (S) protein is cleaved into two subunits, S1 and S2. This protein plays an important role in the attachment of the virus to cell-surface receptors and induces the fusion of viral and cellular membranes. It could be implicated in the variation in host range and in the determination of tropism. In view of the published data, analysis of the S1 gene should be an optimal choice for revealing genetic diversity of coronavirus variants (Cavanagh, 1995; Gallagher & Buchmeier, 2001). Five types of human coronavirus (HCoV) have been described to date: HCoVs OC43, 229E, NL63, the recently described HCoV-HKU1 and SARS-CoV. HCoVs are known to cause acute respiratory infections and could be involved in enteric and neurological disorders (Zhang *et al.*, 1994; Arbour *et al.*, 2000). This study concerns HCoV-OC43 only. Vijgen *et al.* (2005a) submitted to GenBank (accession no. AY391777) the complete genome sequence of the HCoV-OC43 prototype strain (ATCC VR-759) isolated in 1967. These authors demonstrated a high rate of similarity with bovine coronaviruses (BCoVs) and postulated that HCoV-OC43 and BCoVs diverged from each other around the 1890s (Vijgen *et al.*, 2005a). They were then able to demonstrate the circulation of distinct HCoV-OC43 variants and provided evidence for the genetic variability of HCoV-OC43 strains (Vijgen *et al.*, 2005b). During this period, Jeong *et al.* (2005) analysed the S gene of some contemporary BCoV strains in Korea and showed the same genetic variability in BCoVs.

The current study includes seven clinical, non-cell culture-adapted, contemporary HCoV-OC43 strains. Our aim has been twofold. We first studied the inter-patient heterogeneity of the HCoV-OC43 circulating variants and then assessed the intra-patient diversity by investigation of a quasispecies cloud. The HCoV-OC43 S1 gene was amplified directly from seven respiratory specimens by reverse transcription followed by two rounds of 30-cycle PCR using increased-fidelity polymerase (Expand High Fidelity PCR system; Roche). The respiratory specimens were collected in seven children, aged from 3 to 36 months and admitted for upper or lower respiratory-tract illnesses to the University Hospital of Caen in 2003. Hereafter, these variants will be referred to as Caen11 THS, Caen14 BEL,

The GenBank/EMBL/DDBJ accession numbers for the nucleotide sequence data reported in this paper are DQ355400–DQ355408.

Caen15 VAL, Caen17 EYM, Caen BUT, Caen21 VUO and Caen VAC. Two of the variants – Caen BUT and Caen VAC – were used for the study of intra-species heterogeneity. Two laboratory strains propagated into human rectal tumour-cell strains (HRT18) were used as control, one being HCoV-OC43 ATCC number VR-759 (referred to as Caen7 OC43 Labo) and the other a BCoV referred to as Caen6 BCV. No information about space and time of sampling was available for this strain. The outer and inner primers used for the S1 gene amplification were designed from the sequence published by Vijgen *et al.* (2005a) (GenBank accession no. AY391777) as follows: OC897 (nt 23235–23255), 5′-CAATGCCAGGCAGTCTGATA-3′; OC4193 (nt 26505–26525), 5′-AGCAGTGGAGGCAACACTTT-3′; OC1111 (nt 23449–23469), 5′-TACCCCTATGGCAGATGTCC-3′; and OC4000 (nt 26312–26332), 5′-CAGGGGAAAAATTGATGTCG-3′. The second-round PCR generated an amplified product of 2883 bp, including the initiation codon ATG and the proteolytic-cleavage site of the HCoV-OC43 S protein (nt 23449–26332). Amplified S1 gene products were cloned into the PCR-XL-TOPO vector (Invitrogen). Inter-variant diversity was evaluated by analysing one clone per variant and laboratory strains used as control, whilst intra-variant diversity was evaluated by analysing 19 and 20 clones of the Caen VAC and Caen BUT variants, respectively. The DNA templates were sequenced on both strands. The nucleotide sequence data reported in this paper have been deposited in GenBank under accession numbers DQ355400–DQ355408. To access inter-patient diversity, a multiple nucleotide sequence alignment was prepared by using the BioEdit software package (Hall, 1999) and CLUSTAL X version 1.83 (Thompson *et al.*, 1997). This alignment included S gene sequence data of different HCoV-OC43 and BCoVs available in GenBank: prototype BCoV LY-138 (GenBank accession no. AF058942), BCoV L9 avirulent strain (M64667), BCoV Mebus (U00735) and BCQ.3994 (AF339836); contemporary Korean BCoVs KWD1–4 (AY935637–AY935640); prototype HCoV-OC43 ATCC VR-759 (AY391777), HCoV-OC43 sequenced by Künkel & Herrler (1993) (S62886) and contemporary Belgium strains from 2003 and 2004 (BE03 and BE04) described by Vijgen *et al.* (2005a) (AY903454–AY903460). CLUSTAL X version 1.83 was used to conduct phylogenetic analyses. A neighbour-joining phylogenetic tree was constructed by using HCoV-HKU1 as an outgroup and evaluated with 1000 bootstrap pseudoreplicates (Fig. 1). Two phylogenetic clusters, containing HCoV-OC43 strains and BCoVs, respectively, were determined. In the HCoV-OC43 branch, three clusters may be identified. The first cluster contains the laboratory-adapted cell-culture strains (Caen7 Labo, OC43 Paris, ATCC VR-759 and GenBank S62886). The second cluster contains two subgroups in which both of the contemporary Belgium HCoV-OC43 strains (2003 and 2004) have been placed. Interestingly, among our six 2003 HCoV-OC43 isolates found in this branch, three cluster with the 2003 Belgium HCoV-OC43 isolates (Caen11 THS, Caen17 EYM and Caen21 VUO),

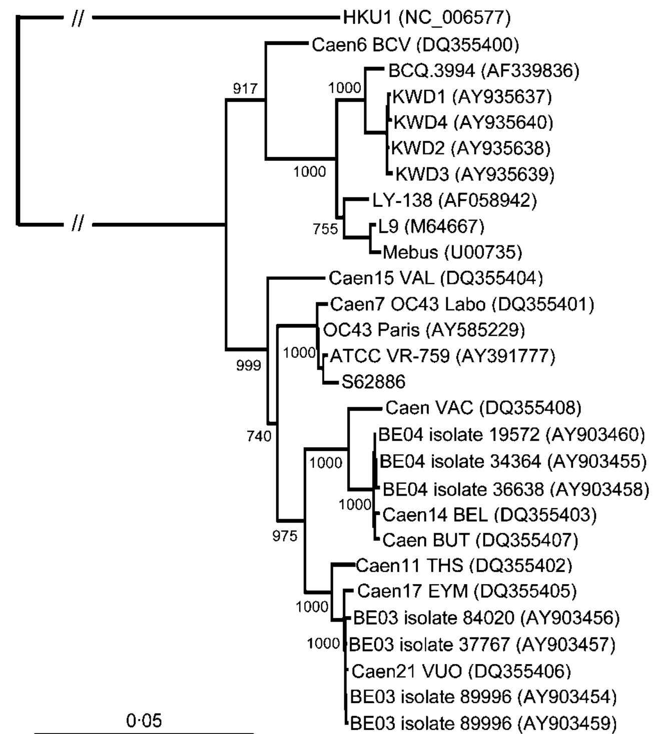

Fig. 1. Neighbour-joining phylogenetic tree of the S1 gene nucleotide sequence data of group 2 coronaviruses: prototype BCoVs LY-138, L9 avirulent strain, Mebus and BCQ.3994; laboratory BCoV strain Caen6 BCV; contemporary Korea BCoVs KWD1–4; HCoV-OC43 ATCC-VR759 and Caen7 OC43 Labo; contemporary Belgium strains from 2003 and 2004 (BE03 and BE04); and our seven variants (Caen15 VAL, Caen11 THS, Caen17 EYM, Caen21 VUO, Caen14 BEL, Caen VAC and Caen BUT). For Caen VAC and Caen BUT, the consensus sequence obtained from the different clones studied has been used. Bar, 0·05 substitutions per site.

whilst the three others (Caen VAC, Caen BUT and Caen14 BEL) cluster with the 2004 Belgium isolates. A parsimony tree has also been deduced by using a heuristic algorithm with PAUP version 4.0b (Swofford, 2003) and shows nearly the same distribution of BCoV and HCoV strains (tree not shown). These results confirm the existence of several genetically distinct HCoV-OC43 variants with different possible temporal- and geographical-circulation patterns and reveal that some HCoV-OC43 variants found in Belgium in 2004 were already circulating in France 1 year before, i.e. in 2003. The variant Caen15 VAL holds characteristics of an outlier and presents a deletion of 12 nt (nt 457–468), also found in all BCoV strains (results not shown). This variant was sampled from a 19-month-old child suffering from acute respiratory-tract illness without presenting any distinctive clinical or epidemiological features. In the BCoV branch, the cell culture-adapted prototype strains and contemporary isolates were distributed into two clusters and several subclusters according to the sampling date (from 1965 to 2003). Caen6 BCV also

holds characteristics of an outlier. Unfortunately, no sampling data were available for this strain used as a control. In order to verify whether bovine-to-human interspecies-transmission events have occurred and thereby resulted in the circulation of new variants, it will be necessary to compare more strains of BCoVs and HCoV-OC43 sampled from the same area without any cell-culture amplification. One of the features of the Belgium contemporary HCoV-OC43 strain was an amino acid change in the last position of the proteolytic-cleavage site of the S protein, resulting in a RRSRR motif identical to that of BCoVs (Vijgen *et al.*, 2005b). The amino acid sequence RRSRR at the predicted cleavage site was identified in our seven contemporary variants and in Caen6 BCV, whilst Caen7 OC43 Labo contained the sequence RRSRG at the predicted cleavage site of the S protein. Cleavage of the coronavirus S protein into the subunits S1 and S2 was not required for virus–cell fusion. Some coronaviruses produce virions with up to 100 % cleaved S protein, whereas no instance of cleaved S protein has been observed in others. The extent of S cleavage depends on the type of coronavirus and the type of host cell studied (Künkel & Herrler, 1993; Cavanagh, 1995). It is not yet possible to say whether this amino acid change at the proteolytic-cleavage site is related to increased or decreased viral infectivity.

The nucleotide sequences of 19 and 20 clones derived from Caen VAC and Caen BUT have also been studied. For HCoV-OC43 Caen VAC, a total of 47 substitutions (45 transitions and two transversions), of which 31 were non-silent and 16 silent, were found. Some substitutions were recurrent or present within more than one clone. Two substitutions were present within two clones (a C at nt 732 in VAC4 and -13; a G at nt 2560 in VAC13 and -14), whilst two other substitutions were present within three clones (an A at nt 706 in VAC7, -11 and -19; a C at nt 1030 in VACU3, -8 and -10). The nucleotide sequence of three clones from the 19 studied clones – VAC5, -10 and -18 – proved identical and there was a total of 17 variants (Table 1). For HCoV-OC43 Caen BUT, we found a total of 31 substitutions (30 transitions and one transversion), of which 18 were non-silent and 13 silent. Two substitutions were present within two clones (a C at nt 599 in BUT3 and -18; a C at nt 832 in BUT9 and -11) and one was present within three clones (a C at nt 639 in BUT5, -6 and -15). The nucleotide sequences of the four clones – BUT2, -13, -14 and -19 – on one side and the two clones – BUT7 and -16 – on the other were identical, and there was a total of 16 variants (Table 2). All of these changes appeared to be distributed throughout the S1 gene; no hot spots or clustering in the location of the mutations were noticed. No in-frame stop codon was found in the analysed clones. Therefore, for each Caen BUT and Caen VAC variant, several clones were identical and could represent a major form. However, most of the clones represented minority or unique forms. Such a heterogeneous population structure containing phylogenetically non-identical but related variants is commonly termed quasispecies. This concept is nevertheless characterized by a

Table 1. Nucleotide variations observed in 19 clones of the Caen VAC variant

Only differences were scored. Dots indicate sequence identity.

Clone	Position																																							
	36	59	126	182	294	359	466	500	505	515	582	643	699	706	732	751	858	869	976	1028	1030	1254	1407	1440	1529	1607	1709	1957	1969	1973	2006	2087	2259	2261	2381	2382	2410	2525	2541	2560	2625
VAC5	T	A	T	T	A	A	G	C	T	A	T	A	A	G	T	C	T	A	T	A	T	A	A	T	A	T	G	T	T	A	A	A	T	A	G	T	T	A	T	A	T
VAC10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VAC18	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VAC4	.	.	.	.	.	.	.	.	.	.	C	.	.	.	C	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VAC13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.
VAC14	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.
VAC7	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VAC11	C	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	G	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.
VAC19	.	G	C	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VACU3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	A	.	.	.	.	.	.	.	A	.	.	.	.	.	.
VACU8	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	G	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.
VACU10	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	T	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	C	G	.	.	C	.	.	.	.
VACU4	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	.	.
VAC6	.	.	.	.	.	.	.	.	.	G	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.	.	.	.	.	.
VAC12	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
VAC16	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
VAC17	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.
VACU5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.
VACU7	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Table 2. Nucleotide variations observed in 20 clones of the Caen BUT variant

Only differences were scored. Dots indicate sequence identity.

Clone	Position																							
	119	302	409	441	453	471	599	639	672	832	968	977	990	1016	1204	1465	1964	2007	2191	2300	2499	2503	2579	2580
BUT2	C	T	G	A	T	T	T	T	A	T	T	A	A	T	A	G	A	T	T	C	A	T	T	T
BUT13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BUT14	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BUT19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BUT7	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.
BUT16	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.
BUT5	T	.	A	.	.	.	.	C	.	.	.	.	.	.	.	A	.	.	C	.	G	.	.	.
BUT6	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.
BUT15	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BUT9	.	C	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BUT11	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BUT3	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BUT18	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.
BUT1	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
BUT4	.	.	.	.	.	.	.	.	.	.	.	C	G	.	.	.	.	C	.	.	.	.	.	.
BUT8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.
BUT12	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.
BUT17	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.
BUT20	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
BUT24	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.

dynamic evolution under selective pressure, such as the immunological response, and is described mainly in chronic or persistent infections (Domingo *et al.*, 1996). Our patients were sampled during the acute period of illness, a few days after infection by HCoV-OC43, at a time when we might assume that the viral population had not yet reached equilibrium with the emergence of a major variant. The description of viral quasispecies has often been based on small amplicons ($<0{\cdot}5$ kb) (Smith *et al.*, 1997). In our case, we chose to analyse larger genomic segments to increase the analytical power and reveal a greater complexity of viral swarm. As a result, the direct sequencing of PCR products, obtained by limiting dilution and commonly recommended to address the issue of misincorporation and artefactual sources of variation, was not possible (Smith *et al.*, 1997; Arias *et al.*, 2001). The possibility that the sequence variation was due to *in vitro* artefacts needs to be addressed. Fortunately, some substitutions were present within more than one clone and are likely to represent segregating polymorphism veritably present within the viral population. The recurrent detection of several mutations within the viral populations constitutes a strong argument for the existence of different variants within the viral population infecting one patient. Our findings concur with the description of quasispecies populations in acute infections due to *Hepatitis A virus*, *Hepatitis E virus* and *Dengue virus* (Wang *et al.*, 2002; Sánchez *et al.*, 2003; Grandadam *et al.*, 2004). To date, no other study has been carried out on acute HCoV-OC43 infection. Genetic diversity allows viral populations to evolve in an ever-changing environment with selective pressure and can have an important biological impact (Vignuzzi *et al.*, 2006). Some minority variants can infect different organs, adapt to them and therefore persist. Our results correspond with the observations of Arbour *et al.* (2000), who have detected HCoV RNA in many human brain specimens and considered this phenomenon as a neuroinvasion by human respiratory coronaviruses. They suggested that, given the fact that most human beings have been in contact with coronaviruses as respiratory pathogens during their childhood, the presence of HCoV RNA in brain samples correlates with a persistent infection within the central nervous system (Arbour *et al.*, 2000).

Acknowledgements

We thank Paul Denis and Francis Delpeyroux for their help.

References

Arbour, N., Ray, D., Newcombe, J. & Talbot, P. J. (2000). Neuroinvasion by human respiratory coronaviruses. *J Virol* **74**, 8913–8921.

Arias, A., Lázaro, E., Escarmís, C. & Domingo, E. (2001). Molecular intermediates of fitness gain of an RNA virus: characterization of a mutant spectrum by biological and molecular cloning. *J Gen Virol* **82**, 1049–1060.

Cavanagh, D. (1995). The coronavirus surface glycoprotein. In *The Coronaviridae*, pp. 73–113, Edited by S. G. Siddell. New York: Plenum.

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

Domingo, E., Escarmís, C., Sevilla, N., Moya, A., Elena, S. F., Quer, J., Novella, I. S. & Holland, J. J. (1996). Basic concepts in RNA virus evolution. *FASEB J* **10**, 859–864.

Gallagher, T. M. & Buchmeier, M. J. (2001). Coronavirus spike proteins in viral entry and pathogenesis. *Virology* **279**, 371–374.

Grandadam, M., Tebbal, S., Caron, M., Siriwardana, M., Larouze, B., Koeck, J. L., Buisson, Y., Enouf, V. & Nicand, E. (2004). Evidence for hepatitis E virus quasispecies. *J Gen Virol* **85**, 3189–3194.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95–98.

Jeong, J.-H., Kim, G.-Y., Yoon, S.-S. & 9 other authors (2005). Molecular analysis of S gene of spike glycoprotein of winter dysentery bovine coronavirus circulated in Korea during 2002–2003. *Virus Res* **108**, 207–212.

Künkel, F. & Herrler, G. (1993). Structural and functional analysis of the surface protein of human coronavirus OC43. *Virology* **195**, 195–202.

Sánchez, G., Bosch, A., Gómez-Mariano, G., Domingo, E. & Pintó, R. M. (2003). Evidence for quasispecies distributions in the human hepatitis A virus genome. *Virology* **315**, 34–42.

Smith, D. B., McAllister, J., Casino, C. & Simmonds, P. (1997). Virus 'quasispecies': making a mountain out of a molehill? *J Gen Virol* **78**, 1511–1519.

Swofford, D. L. (2003). PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sunderland, MA: Sinauer Associates.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.

Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348.

Vijgen, L., Keyaerts, E., Moës, E., Thoelen, I., Wollants, E., Lerney, P., Vandamme, A.-M. & Van Ranst, M. (2005a). Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J Virol* **79**, 1595–1604.

Vijgen, L., Keyaerts, E., Lerney, P., Moës, E., Li, S., Vandamme, A.-M. & Van Ranst, M. (2005b). Circulation of genetically distinct contemporary human coronavirus OC43 strains. *Virology* **337**, 85–92.

Wang, W.-K., Lin, S.-R., Lee, C.-M., King, C.-C. & Chang, S.-C. (2002). Dengue type 3 virus in plasma is a population of closely related genomes: quasispecies. *J Virol* **76**, 4662–4665.

Zhang, X. M., Herbst, W., Kousoulas, K. G. & Storz, J. (1994). Biological and genetic characterization of a hemagglutinating coronavirus isolated from a diarrhoeic child. *J Med Virol* **44**, 152–161.