Sciencexpress Report

Isolation and Characterization of Viruses Related to the SARS Coronavirus from Animals in Southern China

Y. Guan,[1]*† B.J. Zheng,[1]† Y.Q. He,[2] X. L. Liu,[2] Z.X. Zhuang,[2] C.L. Cheung,[1] S.W. Luo,[1] P.H. Li,[1] L.J. Zhang,[1] Y.J. Guan,[1] K.M. Butt,[1] K.L. Wong,[1] K.W. Chan,[3] W. Lim,[4] K.F. Shortridge,[1] K.Y. Yuen,[1] J.S.M. Peiris,[1] L.L.M. Poon[1]

[1]Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Hong Kong SAR, P.R. China. [2]Center for Disease Control and Prevention, Shenzhen, Guangdong Province, P.R. China. [3]Department of Pathology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Hong Kong SAR, P.R. China. [4]Government Virus Unit, Department of Health, Hong Kong SAR, P.R. China.

*To whom correspondence should be addressed. E-mail: yguan@hkucc.hku.hk

†These authors contributed equally to this work.

A novel coronavirus (SCoV) is the etiological agent of the Severe Acute Respiratory Syndrome. SCoV-like viruses were isolated from Himalayan palm civets found in a live animal market in Guangdong, China. Evidence of virus infection was also detected in other animal, including a raccoon-dog, and in humans working at the same market. All the animal isolates retain a 29-nucleotide sequence, which is not found in most human isolates. The detection of SCoV-like viruses in small wild mammals in live retail market indicates a route of interspecies transmission, although the natural reservoir is not known.

Severe acute respiratory syndrome (SARS) is a recently emerged human disease associated with pneumonia (*1*). This disease was first recognized in Guangdong Province, China in November 2002. Subsequent to its introduction to Hong Kong in mid February 2003, the virus spread to more than 30 countries causing disease in over 7,900 patients across 5 continents (*2*). A novel coronavirus (SCoV) was identified as the etiological agent of SARS (*3*, *4*) and the virus causes a similar disease in cynomolgous macaques (*5*). Human SCoV appears to be an animal virus that crossed to humans relatively recently. Thus, identifying animals carrying the virus is of major scientific interest and public health importance. This prompted us to examine a range of domestic and wild mammals in Guangdong Province.

Since the early cases of SARS in Guangdong reportedly occurred in restaurant workers handling wild mammals as exotic food (*6*), our attention focused on wild animals recently captured and marketed for culinary purposes. We investigated a live animal retail market in Shenzhen. Animals were held, one per cage, in small wire cages. The animals sampled included seven wild, and one domestic animal species (Table 1). They originated from different regions of southern China and had been kept in separate storehouses before arrival to the market. The animals remained in the markets for a variable period of time and each stall holder had only a few animals of a given species. Animals from different stalls within the market were sampled. Nasal and fecal swabs were collected and stored in medium 199 with bovine serum albumin and antibiotics. Where possible, blood samples were collected for serology. Prior to sampling, all animals were examined by a veterinary surgeon and confirmed to be free of overt disease. Serum samples were also obtained, after informed consent, from animal (n = 35) and vegetable (n = 20) traders within the market. Sera (n = 60) submitted for routine laboratory tests from patients hospitalized for non-respiratory disease in Guangdong were anonymized and used for comparison.

Nasal and fecal swabs from 25 animals were tested for SCoV viral nucleic acid using RT-PCR for the N gene of the human SCoV. Swabs from 4 of 6 Himalayan palm civets were positive in the RT-PCR assay (Table 1). All specimens were inoculated on to FRhk-4 cells as previously described for virus isolation (*3*). Cytopathic effect was observed in cells inoculated with specimens from 4 Himalayan palm civets (*Paguma larvata*), two of whom were also RT-PCR positive in the original specimen. A virus was also detected by virus isolation and direct RT-PCR from the fecal swab of a raccoon-dog (*Nyctereutes procyonoides*). No virus was detectable in 6 other species sampled. Electron microscopy of one infected cell supernatant (SZ16) showed viral particles with a morphology compatible to a coronavirus (fig. S1). Sera from five animals had neutralizing antibody to the animal coronavirus; these were from three palm civets, a raccoon-dog and a Chinese ferret badger, respectively (Table 1).

To further validate the results from the neutralization test, a Western blot assay was used to detect anti-SCoV antibodies from these animal serum samples (Fig. 1). Positive signals were observed from samples SZ2, SZ3, SZ11 and SZ17 that were positive in the neutralization assay and from the positive control human serum. No positive signal was observed from those serum samples that were negative in the neutralization test. There was insufficient serum left over from the raccoon dog (SZ13) to be analyzed by this assay.

Sera from humans working in the market were tested for antibody to SZ16 virus by neutralization and indirect immunofluorescence assays. While 8 out of 20 (40%) of the wild animal traders and 3 of 15 (20%) of those who slaughter these animals had evidence of antibody, only 1 (5%) of 20 vegetable traders was seropositive. None of these workers reported SARS-like symptoms in the last six months. In comparison, none of 60 control sera from patients admitted to a Guangdong hospital for non-respiratory diseases was seropositive (Table 2).

Two of the virus isolates (SZ3 and SZ16) isolated from the nasal swabs of palm civets were completely sequenced and the amino acid sequence deduced. Two other viruses were partially sequenced, from the S gene to the 3' end of the virus

(GenBank accession number AY304486 to AY304489). Viral RNA sequences from these animal original swab samples were confirmed in an independent lab (*7*). The full-length genome sequences had 99.8% homology to the human SCoV indicating the human and animal SCoV-like viruses were closely related. Phylogenetic analysis of the S gene of both human and animal SCoV-like viruses indicated that the animal viruses are separated from the human virus cluster (Fig. 2 and fig. S2). However, the viruses SZ1, SZ3 and SZ16 from palm civets were phylogenetically distinct. The viruses SZ3 and SZ16 had 18 nucleotide differences between them over the 29,709 bp genome whereas the human SCoV isolated from 5 geographically separate sites (GZ50, CUHK-W1, Tor-2, HKU-39848 and Urbani) differed by only 14 nucleotides. On the other hand, animal virus SZ13 (raccoon-dog) and SZ16 (palm civet) were genetically almost identical, and transmission or contamination from one host to the other within the market cannot be excluded.

When the full genome of the animal (n = 2) and human (n = 5, see above) virus groups were compared the most striking difference was that these human viruses have a 29 nucleotide deletion (5′-CCTACTGGTTACCAACCTGAATGGAATAT -3′, residue 27869 to 27897) at 246 nucleotide upstream of the start codon of the N gene (Fig. 3). Of human SCoV sequences currently available in GenBank, there was only one (GZ01) with this additional 29 nucleotide sequence. In addition to that there were 43 to 57 nucleotide differences observed over the rest of the genome. Most of these differences were found in S gene coding region. The existence of the additional 29 nucleotide sequence in the animal viruses results in demolishing the open reading frames (ORFs) 10 and 11 (*8*) and merging these two ORFs into a new ORF encoding a putative protein of 122 amino acids (Fig. 3). This putative peptide has a high homology to the putative proteins encoded by ORF10 and ORF11. Since ORF11 does not have a typical transcription regulatory sequence for SCoV (*6*) the putative ORF11 reported by others may just be the direct result of the deletion of the 29 nucleotide sequence. BLAST search of this peptide yields no significant match to any other known peptide. Further investigation is required to elucidate the biological significant of this finding.

When the S-gene sequences of the four animal viruses were compared with 11 human SCoV viruses, 38 nucleotide polymorphisms were noted, and 26 of them being non-synonymous changes (Table 3). The S genes among the 4 animal viruses had 8 nucleotide differences while there were 20 nucleotide differences among 11 human viruses. Thus the animal viruses, though isolated from 1 market, are no less divergent than the human viruses isolated from Hong Kong, Guangdong, Canada and Vietnam. However, whereas 14 (70%) of the 20 polymorphisms among the human viruses were non-synonymous mutations, only 2 (25%) of the 8 nucleotide substitutions within the animal viruses were. An amino acid deletion (nucleotide positions: 21690-21692) was observed in two of the human viruses (GZ43 and GZ60). Of the 38 polymorphisms, there were 11 consistent nucleotide signatures that appeared to distinguish animal and human viruses. The observation that the human and animal viruses are phylogenetically distinct (Fig. 2) makes it highly unlikely that the isolation of SCoV-like viruses in these wild animals is due to the transmission of SCoV from human to animals.

Our findings suggest that the markets provide a venue for the animal SCoV-like viruses to amplify and transmit to new hosts, including humans and this is critically important from the point of view of public health. It is not, however, clear whether any one or more of these animals are the natural reservoir in the wild. It is conceivable that civets, raccoon-dog and ferret badgers were all infected from another, as yet unknown animal source, which is in fact the true reservoir in nature. However, because of the culinary practices of southern China, these market animals may be intermediate hosts that increase the opportunity for transmission of infection to humans. Further extensive surveillance on animals will help to better understand the animal reservoir in nature and the inter-species transmission events that led to the origin of the SARS outbreak.

References and Notes

1. WHO (http://www.who.int/csr/sars/en/)
2. WHO: Cumulative Number of Reported Probable Cases of Severe Acute Respiratory Syndrome (SARS) (http://www.who.int/csr/sars/country/2003_05_20/en/)
3. J.S.M. Peiris et al., *Lancet*, **361**, 1319 (2003).
4. T.G. Ksiazek et al., *N. Engl. J. Med.* **348**, 1953 (2003).
5. R.A. Fouchier et al., *Nature* **423**, 240 (2003).
6. N.S. Zhong et al., *Lancet* in press.
7. K. Holmes, unpublished data.
8. M.A. Marra et al., *Science*, **300**, 1399 (2003).
9. China species information system (http://www.chinabiodiversity.com).
10. S. Kumar et al., *Bioinformatics* **17**, 1244 (2001).
11. M Kimura, *J. Mol. Evol.* **16**, 111 (1980).

12. We thank the Department of Health and Department of Agriculture of Shenzhen Government for facilitating the study. We gratefully acknowledge the encouragement and support of Prof. L.C. Tsui, Vice-Chancellor, The University of Hong Kong. We thank X.Y. Zhao from Department of the Microbiology, The University of Hong Kong for the excellent technical assistance. We also thank C.C. Hon and Dr. F.C. Leung from the Department of Zoology, The University of Hong Kong, and Dr. Richard Webby from St. Jude Children's Research Hospital (Memphis, TN, U.S.A.) for the assistance in the phylogenetic analysis. We thank Dr. K.V. Holmes's lab from the Department of Microbiology, University of Colorado Health Sciences Center (Denver, U.S.A) to validate the animal viral sequences. Supported by research funding from Public Health Research (Grant A195357), the National Institute of Allergy and Infectious Diseases, USA, the Wellcome Trust (067072/D/02/Z) and SARS research funds from The University of Hong Kong.

Supporting Online Material

www.sciencemag.org/cgi/content/full/[ms. no.]/DC1
Materials and Methods
Figs. S1 and S2
References and Notes

22 May 2003; accepted 26 August 2003
Published online 4 September 2003;
10.1126/science.1087139
Include this information when citing this paper.

Fig. 1. Detection of antibodies against recombinant nucleocapsid protein of SCoV in animal sera by Western Blot Assay. Recombinant nuleocapsid protein (NP, 49.6 kDa) was used as an antigen to detect anti-SCoV antibodies in animal

sera. Protein A-HRP was used as a secondary antibody and reactive bands were visualized by the ECL western blotting system. Human convalescent serum sample from a SARS patient (Human) was used as a positive control. Blots reacted with animal (SZ2, SZ3, SZ11, SZ17, SZ7, SZ16 or SZ19) or human sera are indicated. Results from the neutralization test for anti-SCoV antibodies in these serum samples are also shown.

Fig. 2. Phylogenetic analysis of nucleotide acid sequence of spike gene of SCoV-like viruses. Nucleotide sequences of representative SCoV S genes (S gene coding region 21477 to 25244, 3768 bp) were analyzed. The phylogenetic tree was constructed by neighbor-joining method with bootstrap analysis (1000 replicates) using MEGA 2 (*10*). Number at the notes indicates bootstrap values in percentage. The scale bar shows genetic distance estimated using Kimura's two parameter substitution model (*11*). In addition to viruses sequenced in current study, the other sequences used in the analysis could be found in GenBank with accession number: from AY304490 to AY304495, AY278741, AY278554, AY278491, AY274119, and AY278489.

Fig. 3. A 29 nucleotide deletion in the human SCoV genome. (**A**) Genetic organization of SCoV-like viruses found in humans and animals. ORFs 1a and 1b, encoding the nonstructural polyproteins, and those encoding the S, E, M, and N structural proteins are indicated (green boxes). (**B**) Expanded view of the SCoV genomic sequence (27700 nt to 28200 nt, based on AY278554 numbering). ORFs for putative proteins and for N in human isolates are indicated as brown and green boxes, respectively (*8*). An extra 29 nucleotide sequence is present down-stream of the nucleotide of 27868 of the animal SCoV (based on AY278554 numbering). The presence of this 29 nt sequence in animal isolates results in fusing the ORFs 10 and 11 (upper panel) into a new ORF (lower panel; ORF10'; light blue box). (**C**) Protein sequence alignment of ORF10 and 11 from human isolates and ORF 10' from animal isolates.

Table 1. Animal species tested for coronavirus detection

Sample number	Animal type (Species)	Virus detection				Neutralizing antibody titer to SZ16
		RT-PCR detection		Isolation		
		Nasal	Fecal	Nasal	Fecal	
SZ1	HPC	+*	+			ND
SZ2	HPC	+	+			40
SZ3	HPC	+	+	+*		40
SZ4	HB					<20
SZ5	B					<20
SZ6	DC					ND
SZ7	DC					<20
SZ8	CH					ND
SZ9	CH					<20
SZ10	CM					<20
SZ11	CFB					160
SZ12	CFB					<20
SZ13	RD		+		+*	≥640
SZ14	CM					<20
SZ15	B					<20
SZ16	HPC	+	+	+*	+	<20
SZ17	HPC			+		≥640
SZ18	B					<20
SZ19	CH					<20
SZ20	CH					<20
SZ21	DC					<20
SZ22	DC					<20
SZ23	HB					ND
SZ24	HB					ND
SZ25	HPC			+		ND

Abbreviation of animal species: HPC, Himalayan palm civet (*Paguma larvata*); HB, Hog-badger (*Arctonyx collaris*); RD, raccoon-dog (*Nyctereutes procyonoides*); B, Beaver (*Castor fiber*); CM, Chinese muntjac (*Muntiacus reevesi*); DC, Domestic cat (*Felis catus*); CH, Chinese Hare (*Lepus sinensis*); CFB, Chinese Ferret-Badger (*Melogale moschata*) (*9*).
+ positive by RT-PCR or virus isolation, * the PCR product or virus isolates sequenced in the study. ND, not done.

Table 2. Prevalence of antibody to animal SCoV SZ16 in humans

Occupation	Sample numbers	Antibody positive (%)
Wild animal trader	20	8 (40)
Slaughterer of animals	15	3 (20)
Vegetable trader	20	1 (5)
Control*	60	0 (0)

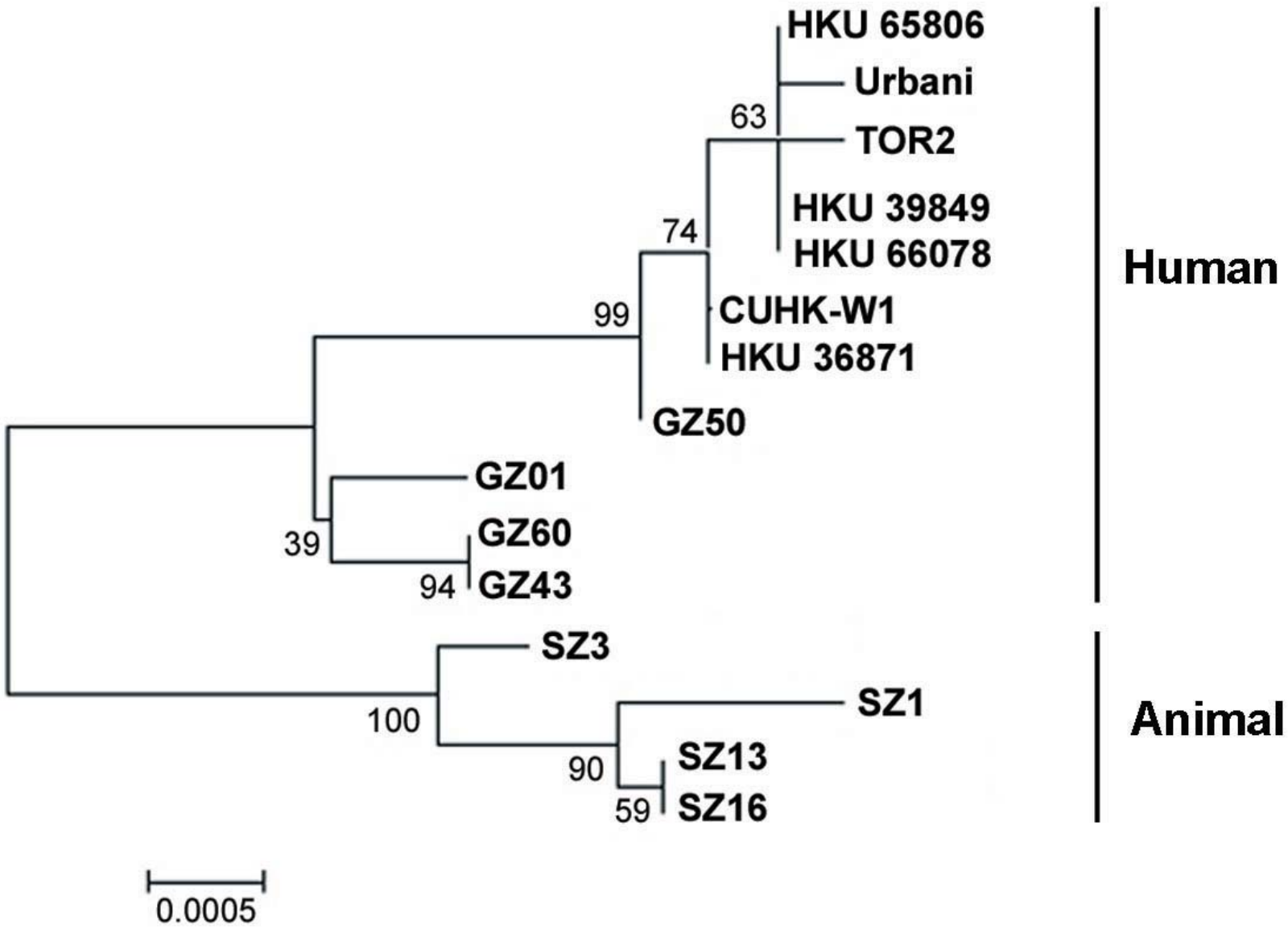
*Anonymized serum specimens from patients hospitalized for non-respiratory diseases in Guangdong.

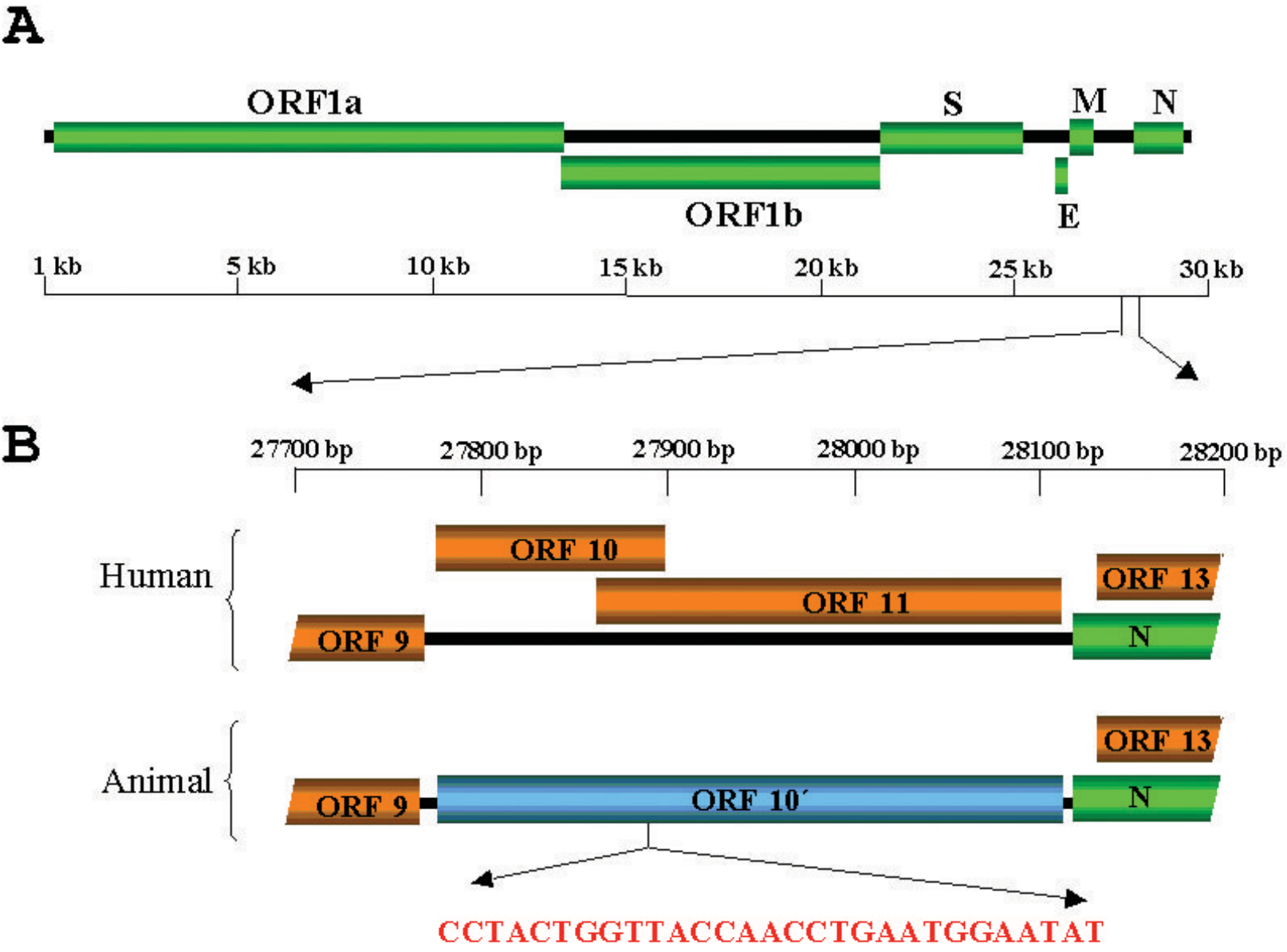
Table 3. Nucleotide sequence variation of the S gene of animal and human SCoV

Virus	Nucleotide residue																																				
	21622	21690	21691	21692	21700	21706	22130	22157	22192	22205	22207	22258	22407	22502	22507	22555	22646	22913	22936	22978	23003	23148	23205	23295	23470	23578	23617	23703	23737	23808	23856	24156	24197	24368	24551	24857	24963	25012
SZ3	C	A	T	T	C	A	T	A	T	T	C	A	G	G	G	C	A	A	G	T	G	T	T	C	C	T	C	G	T	G	C	G	C	G	C	T	G	T
SZ16	C	A	T	T	C	A	T	A	T	C	C	A	G	G	G	C	G	A	G	T	G	A	T	C	C	T	C	G	T	G	C	G	C	T	C	T	G	T
SZ1	C	A	T	T	C	A	T	A	T	T	C	A	G	G	G	C	G	A	G	T	T	A	T	C	C	T	T	G	T	G	C	G	T	T	T	T	G	T
SZ13	C	A	T	T	C	A	T	A	T	T	C	A	G	G	G	C	G	A	G	T	G	A	T	C	C	T	C	G	T	G	C	G	C	T	C	T	G	T
GZ01	C	A	T	T	C	A	C	C	T	C	C	C	A	G	G	T	G	T	C	A	G	T	T	T	T	C	C	A	C	G	T	A	C	G	C	T	A	T
GZ43	C	--	--	--	G	A	T	C	T	C	C	C	A	G	G	T	G	T	C	T	G	T	T	T	C	C	C	A	C	G	C	A	C	G	C	T	A	C
GZ60	C	--	--	--	G	A	T	C	T	C	C	C	A	G	G	T	G	T	C	T	G	T	T	T	C	C	C	A	C	G	C	A	C	G	C	T	A	C
GZ50	T	A	T	T	C	A	T	C	C	C	C	C	G	A	A	T	G	T	C	T	G	T	T	T	T	C	C	A	C	G	C	A	C	G	T	T	A	T
CUHK-W1	C	A	T	T	C	A	T	C	C	C	C	C	G	A	A	T	G	T	C	T	G	T	T	T	T	C	C	A	C	T	C	A	C	G	T	T	A	T
HKU-36871	C	A	T	T	C	A	T	C	C	C	C	C	G	A	A	T	G	T	C	T	G	T	T	T	T	C	C	A	C	T	C	A	C	G	T	T	A	T
HKU-39848	C	A	T	T	C	G	T	C	C	C	T	C	G	A	A	T	G	T	C	T	G	T	T	T	T	C	C	A	C	T	C	A	C	G	T	T	A	T
HKU-66078	C	A	T	T	C	G	T	C	C	C	T	C	G	A	A	T	G	T	C	T	G	T	T	T	T	C	C	A	C	T	C	A	C	G	T	T	A	T
HKU-65806	C	A	T	T	C	G	T	C	C	C	T	C	G	A	A	T	G	T	C	T	G	T	T	T	T	C	C	A	C	T	C	A	C	G	T	T	A	T
Urbani	C	A	T	T	C	G	T	C	C	C	T	C	G	A	A	T	G	T	C	T	G	T	T	T	T	C	C	A	C	T	C	A	C	G	T	C	A	T
Tor2	C	A	T	T	C	G	T	C	C	C	T	C	G	A	A	T	G	T	C	T	G	T	G	T	T	C	C	A	C	T	C	A	C	G	T	T	A	T

Note: The nucleotide residues are based on AY278554 numbering. Non-silent mutations are highlighted in red. "--" indicates a nucleotide deletion.

SZ2
SZ3
SZ11
SZ17
SZ7
SZ16
SZ19
Human
250kDa
NP
50kDa
30kDa
25kDa
10kDa
Presence of neutralizing antibody : + + + + − − − +


HKU 65806
Urbani
63
TOR2
HKU 39849
74
HKU 66078
Human
99
CUHK-W1
HKU 36871
GZ50
GZ01
39
GZ60
94
GZ43
SZ3
SZ1
100
Animal
90
SZ13
59
SZ16
0.0005



C

```
          1                                                   50
ORF10     MKLLIVLTCI SLCSCICTVV QRCASNKPHV LEDPCKVQH~ ~~~~~~~~~~
ORF11     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~MC LKILVRYNTR
ORF10'    MKLLIVLTCI SLCSCICTVV QRCASNKPHV LEDPCPTGYQ PEWNIRYNTR

          51                                                  100
ORF10     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
ORF11     GNTYSTAWLC ALGKVLPFHR WHTMVQTCTP NVTINCQDPA GGALIARCWY
ORF10'    GNTYSTAWLC ALGKVLPFHR WHTMVQTCTP NVTINCQDPA GGALIARCWY

          101                  122
ORF10     ~~~~~~~~~~ ~~~~~~~~~~ ~~
ORF11     LHEGHQTAAF RDVLVVLNKR TN
ORF10'    LHEGHQTAAF RDVLVVLNKR TN
```