

Cell Reports, Volume 39

Supplemental information

**Emerging viruses: Cross-species transmission
of coronaviruses, filoviruses, henipaviruses,
and rotaviruses from bats**

Jin Tian, Jiumeng Sun, Dongyan Li, Ningning Wang, Lifang Wang, Chang Zhang, Xiaorong Meng, Xiang Ji, Marc A. Suchard, Xu Zhang, Alexander Lai, Shuo Su, and Michael Veit

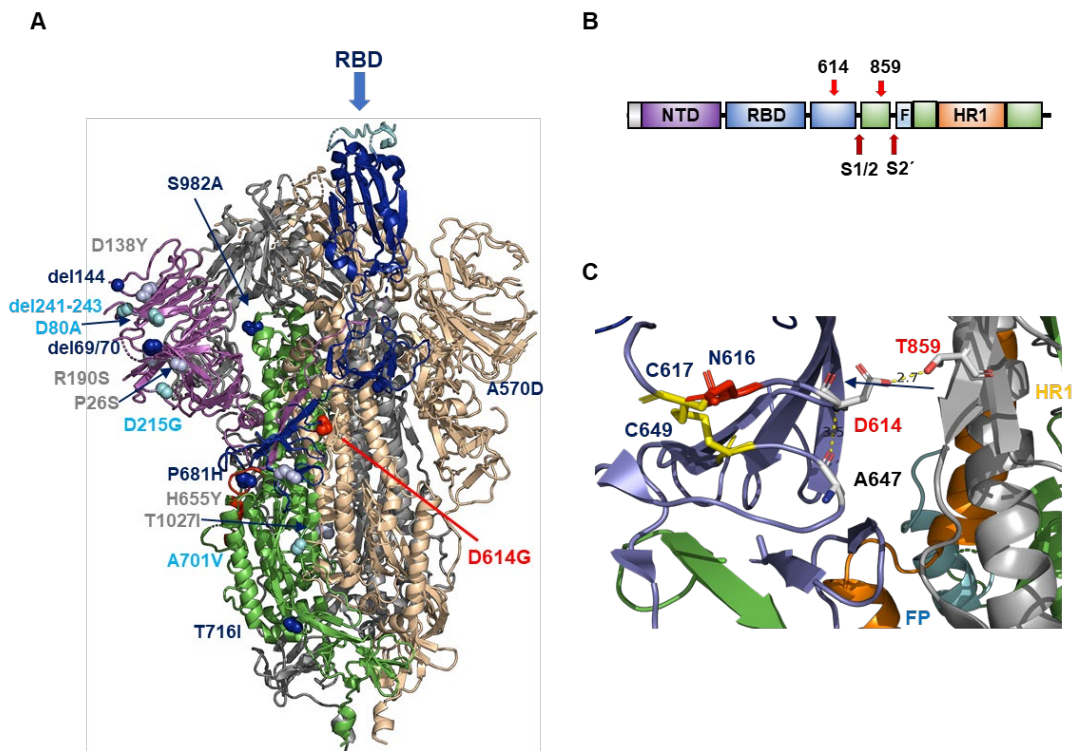


Figure S1. Amino acid substitutions in the spike protein. (A) Cryo-EM structure of the trimeric S highlighting the location of amino acid substitutions in SARS-CoV-2 variants. The N-terminal domain (NTD) of the S1 subunit is in magenta, the remaining part of S1 including the receptor binding domain (RBD) is in blue and the S2 subunit is in green. Residue Asp 614, which was substituted early in the pandemic by a Gly is shown as a red sphere. The substitutions in the variant B.1.1.7 are shown as blue spheres, in the variant B.1.35 as cyan spheres and in the variant P.1 as grey spheres. The RBD of the colored monomer is in the up conformation, the other two in the down conformation. The amino acids substituted in the RBD are shown in Fig 3b. The figure was created from Pdb file 6VSB. (B) Primary structure of S showing the location of positions 614 and 859. The S1 subunit is colored as in a. The S2 subunit is in green except for the heptad repeat region 1 (HR1), which is in orange and the fusion peptide F, which is in blue. The proteolytic cleavage sites S1/2 and S2' are indicated by red arrows. (C) Hydrogen bond between Asp614 and Thr859. Asp614 forms the apex of a loop that contains a used N-glycosylation site, Asn 616. It is covalently connected to another loop by a disulphide-linkage between Cys 617 and Cys 649 and via interaction of the main chain atoms of Asp614 with Ala647. Both loops are at the interface between two monomers of the trimeric spike protein. The carboxyl group of Asp 614 forms a hydrogen bond with the -OH group of Thr 859, which is part of a short β -strand in the S2 subunit between the fusion peptide and the heptad region. Protein domains are colored as in B. The figure was created from pdb file 6VSB.

Species	24	27	28	31	34	37	38	41	42	45	79	82	83	84	90	91	92	322	323	324	325	329	330	353	354	355	357	Virus binding
Human	Q	T	F	K	H	E	D	Y	Q	L	L	M	Y	P	N	L	T	N	M	T	Q	E	N	K	G	D	R	+
Ferret	L	T	F	K	Y	E	E	Y	Q	L	H	T	Y	P	D	P	I	N	M	T	E	Q	N	K	R	D	R	+
Raccoon	L	T	F	K	Y	E	E	Y	Q	L	L	T	Y	P	D	S	T	N	M	T	Q	E	N	K	G	D	R	+
Civet	L	T	F	T	Y	Q	E	Y	Q	V	L	T	Y	P	D	A	K	N	M	T	Q	E	N	K	G	D	R	+
Mouse	N	T	F	N	Q	E	D	Y	Q	L	T	S	F	S	T	P	I	H	M	T	Q	A	N	H	G	D	R	+/-
Rat	K	S	F	K	Q	E	D	Y	Q	L	I	N	F	S	N	A	T	Q	M	T	P	T	N	H	G	D	R	-
Chinese horseshoe bat-YN	E	M	F	K	T	E	D	H	Q	L	L	N	Y	S	N	V	T	N	M	T	E	N	N	K	G	D	R	?
Chinese horseshoe bat-GX	E	I	F	K	T	E	D	H	Q	L	L	N	Y	P	N	V	T	N	M	T	E	N	N	K	G	D	R	?
Chinese horseshoe bat-HB	R	T	F	E	S	E	N	Y	Q	L	L	N	Y	P	N	V	T	N	M	T	E	N	N	K	G	D	R	?
Chicken	E	T	F	E	V	E	D	Y	E	L	N	R	F	S	D	A	V	N	M	T	E	T	N	K	N	D	R	-

Table S1. Comparison of the ACE2 residues interfacing with SARS-CoV receptor-binding domain (RBD) of different species. Red numbers: contact to SARS-CoV-2 S and to SARS-CoV S. Blue numbers: contact to SARS-CoV S only. Amino acids differences to human ACE2 are indicated as red characters. Highlighted in grey: N-glycosylation sites. Accession numbers: human (*Homo sapiens*): BAB40370.1; Ferret (*Mustela putorius furo*): BAE53380.1; Raccoon (*Nyctereutes procyonoides*): ABW16956.1; Civet (*Paguma larvata*): AAX63775.1; Mouse (*Mus musculus*): ABN80106.1; Rat (*Rattus norvegicus*): AAW78017.1; Chinese horseshoe bat-YN (*Rhinolophus sinicus*-YN): AGZ48803.1; Chinese horseshoe bat-GX (*Rhinolophus sinicus*-GX): ADN93472.1; Chinese horseshoe bat-HB (*Rhinolophus sinicus*-HB): ADN93475.1; Chicken (*Gallus gallus*): XP_416822.2. -YN indicates an individual from Yunnan; -GX indicates an individual from Guangxi; -HB indicates an individual from Hubei. Amino acids differences to human ACE2 are indicated as red characters, some of the amino acids do not directly contact the S-protein, but are present in a region that encode a N-glycosylation site in ACE2 proteins from other animals.

Species	24	26	27	28	30	31	34	35	37	38	40	41	42	43	78	79	81	82	83	84	90	91	92	95	322	323	324	330	353	354	355	357	393	Virus binding
Human	Q	K	T	F	D	K	H	E	E	D	F	Y	Q	S	T	L	Q	M	Y	P	N	L	T	L	N	M	T	N	K	G	D	R	R	+
Dog	L	K	T	F	E	K	Y	E	E	E	S	Y	Q	S	K	L	K	T	Y	P	D	S	T	R	N	M	T	N	K	G	D	R	R	+
Cat	L	K	T	F	E	K	H	E	E	E	S	Y	Q	S	K	L	K	T	Y	P	N	T	T	R	N	M	T	N	K	G	D	R	R	+
Ferret	L	K	T	F	E	K	Y	E	E	E	S	Y	Q	S	Q	H	K	T	Y	R	D	P	I	R	N	M	T	N	K	R	D	R	R	+
Pig	L	K	T	F	E	K	L	E	E	D	A	Y	Q	S	R	I	K	T	Y	P	T	L	I	R	N	M	T	N	K	G	D	R	R	+
Hamster	Q	K	T	F	D	K	Q	E	E	D	S	Y	Q	S	K	L	K	N	Y	S	N	L	T	R	Y	M	T	N	K	G	D	R	R	+
Raccoon dog	L	N	T	F	E	K	Y	E	E	E	S	Y	Q	S	K	L	K	T	Y	P	D	S	T	R	N	M	T	N	K	G	D	R	R	+
Civet	L	K	T	F	E	T	Y	E	Q	E	S	Y	Q	S	K	L	Q	T	Y	P	D	A	K	R	N	M	T	N	K	G	D	R	R	+
Chinese horseshoe bat-YN	E	K	M	F	D	K	T	K	E	D	S	H	Q	S	K	L	K	N	Y	S	N	V	T	L	N	M	T	N	K	G	D	R	R	+
Chinese horseshoe bat-GX	E	K	I	F	D	K	T	K	E	D	S	H	Q	S	K	L	K	N	Y	P	N	V	T	L	N	M	T	N	K	G	D	R	R	+
Chinese horseshoe bat-HB	R	K	T	F	D	E	S	E	E	N	S	Y	Q	S	K	L	K	N	Y	P	N	V	T	L	N	M	T	N	K	G	D	R	R	+
Horseshoe bat	R	K	I	F	D	N	R	E	E	E	S	Y	Q	S	K	L	K	N	Y	P	T	V	P	L	N	M	T	N	K	G	D	R	R	+
Egyptian fruit bat	L	K	T	F	E	K	T	E	E	D	F	Y	Q	S	K	L	K	T	Y	Q	D	P	E	L	N	M	T	K	K	G	D	R	R	+
Pangolin	E	K	T	F	E	K	S	E	E	E	S	Y	Q	S	K	I	K	N	Y	Q	N	D	T	R	K	M	T	N	K	H	D	R	R	+
Mouse	N	K	T	F	N	N	Q	E	E	D	S	Y	Q	S	K	T	Q	S	F	S	T	P	I	R	H	M	T	N	H	G	D	R	R	-
Rat	K	E	S	F	N	K	Q	E	E	D	S	Y	Q	S	K	I	Q	N	F	S	N	A	T	R	Q	M	T	N	H	G	D	R	R	-
Chicken	E	Q	T	F	A	E	V	R	E	D	S	Y	E	N	R	N	S	R	F	S	D	A	V	L	N	M	T	N	K	N	D	R	R	-
Marmoset	Q	K	T	F	D	K	H	E	E	D	F	H	E	N	K	L	Q	T	Y	P	N	L	T	L	N	M	T	N	K	Q	D	R	R	-
Koala	R	K	E	F	E	T	K	E	E	E	S	Y	Q	S	N	I	R	T	F	P	D	P	Q	L	N	M	T	N	K	G	D	R	R	-

Table S2. Comparison of the ACE2 residues interfacing with SARS-CoV-2 receptor-binding domain (RBD) of different species. Red numbers: contact to the spike of both SARS-CoV S and SARS-CoV-2. Green numbers: contact to SARS-CoV-2 S only. Black numbers: variable amino acids close to contacting amino acids. Amino acids differences to human ACE2 are indicated as red characters. Highlighted in grey: N-glycosylation sites. Residues 90 and 322 contain carbohydrates in human ACE2. ACE2 proteins from two other New World monkeys (*Tufted capuchin*, XM_032285963.1 and *Squirrel monkey*, XM_010336623.1), which also do not serve as SARS-CoV-2 receptor have identical residues compared to marmoset. Accession numbers: Human: BAB40370.1;

Dog (*Canis lupus familiaris*): ACT66277.1; Cat (*Felis catus*): MT269671; Ferret (*Mustela putorius furo*): BAE53380.1; Pig (*Sus scrofa*): NP_001116542.1; Hamster (*Mesocricetus auratus*): XP_005074266.1; Raccoon dog (*Nyctereutes procyonoides*): ABW16956.1; Civet (*Paguma larvata*): AAX63775.1; Chinese horseshoe bat-YN (*Rhinolophus sinicus*-YN): AGZ48803.1; Chinese horseshoe bat-GX (*Rhinolophus sinicus*-GX): ADN93472.1; Chinese horseshoe bat-HB (*Rhinolophus sinicus*-HB): ADN93475.1; Horseshoe bat (*Rhinolophus affinis*): QMQ39222.1; Egyptian fruit bat (*Rousettus aegyptiacus*): XM_016118926.1; Pangolin (*Manis javanica*): XP_017505746.1; Mouse (*Mus musculus*): ABN80106.1; Rat (*Rattus norvegicus*): AAW78017.1; Chicken (*Gallus gallus*): XP_416822.2; Marmoset (*Callithrix jacchus*): XM_008988993.3; Koala (*Phascolarctos cinereus*): XM_021007494.1. -YN indicates an individual from Yunnan; -GX indicates an individual from Guangxi; -HB indicates an individual from Hubei.

Species	229	267	269	286	288	289	291	294	295	298	317	322	334	335	336	341	344	346	392	Virus binding
<i>H. sapiens</i>	N	K	F	Q	T	A	A	L	I	H	R	Y	S	G	R	V	Q	I	K	+
<i>C. dromedarius</i>	N	K	F	Q	V	P	A	L	I	H	R	Y	T	G	R	V	Q	I	R	+
<i>S. scrofa</i>	N	K	F	Q	V	P	A	L	I	H	R	Y	T	G	R	V	Q	I	S	+
<i>M. auratus</i>	N	K	F	Q	T	A	E	L	T	H	R	Y	N	L	T	L	Q	V	K	-
<i>M. musculus</i>	N	K	F	Q	P	A	A	A	R	H	R	Y	N	L	T	S	Q	V	K	-
<i>R. ferrumequinum</i>	N	K	F	Q	V	A	A	L	I	H	R	Y	N	D	R	V	Q	I	E	++
<i>R. aegyptiacus</i>	N	K	F	Q	V	A	A	L	I	H	R	Y	D	G	S	V	Q	M	E	++
<i>H. armiger</i>	N	K	F	Q	V	A	A	L	I	H	R	Y	N	G	R	V	Q	I	E	++
<i>S. bilineata</i>	N	K	F	Q	P	A	A	L	I	H	R	Y	S	G	R	V	Q	I	Q	++
<i>A. jamaicensis</i>	N	K	F	Q	T	A	A	L	I	H	R	Y	S	G	R	V	Q	I	Q	+
<i>P. vampyrus</i>	N	K	F	Q	V	A	A	L	T	H	R	Y	N	G	S	V	Q	M	K	+
<i>E. fuscus</i>	N	K	Y	Q	T	A	A	L	I	H	R	Y	-	P	R	V	Q	I	E	+
<i>C. perspicillata</i>	N	K	F	Q	T	A	A	L	T	H	R	Y	S	G	R	V	Q	I	Q	-/+
<i>P. abramus</i>	D	K	Y	Q	T	A	A	L	I	H	R	Y	Y	S	T	V	Q	I	D	-/+
<i>M. davidii</i>	N	K	Y	Q	V	A	A	L	T	H	R	Y	T	P	K	V	Q	I	P	-/+
<i>P. alecto</i>	N	K	F	Q	V	A	A	L	T	H	R	Y	N	G	S	V	Q	L	K	-/+
<i>A. planirostris</i>	N	K	F	Q	T	A	A	L	I	H	R	Y	S	G	R	V	Q	I	Q	--/+
<i>M. lucifugus</i>	N	K	Y	Q	I	A	A	L	I	H	R	Y	T	P	R	V	Q	I	E	--/+
<i>M. brandtii</i>	N	K	Y	Q	V	A	A	L	I	H	R	Y	T	P	I	V	Q	I	K	--/+
<i>E. buettikoferi</i>	N	K	F	Q	I	A	A	L	I	H	R	Y	N	G	S	V	Q	M	E	--/+
<i>D. rotundus</i>	N	K	F	Q	T	A	A	L	T	H	Q	Y	S	G	R	V	Q	I	Q	--/+

Table S3. Comparison of the DPP4 residues interfacing with MERS-CoV RBD of different species. Amino acids differences to human DPP4 are indicated as red characters. Highlighted in grey: N-glycosylation sites. Plus and minus indicates whether MERS binds to this receptor with higher (++), the same (+) less (-) or much lower affinity compared to the human receptor. Note that only *Myotis s.*, *Pipistrellus s.* and *Eptesicus s.* belong to the *Vespertilionidae* family. Accession numbers: *H. sapiens* (*Homo sapiens*): NM_001379604.1; *C. dromedarius* (*Camelus dromedarius*): XM_031451547.1; *S. scrofa* (*Sus scrofa*): NM_214257.1; *M. auratus* (*Mesocricetus auratus*): NM_001310571.1; *M. musculus* (*Mus musculus*): NM_010074.3; *R. ferrumequinum* (*Rhinolophus*

ferrumequinum): MH299899.1; *R. aegyptiacus* (*Rousettus aegyptiacus*): MH299900.1; *H. armiger* (*Hipposideros armiger*): XM_019636841.1; *S. bilineata* (*Saccopteryx bilineata*): MH299901.1; *A. jamaicensis* (*Artibeus jamaicensis*): KF574262.1; *P. vampyrus* (*Pteropus vampyrus*): XM_023529782.1; *E. fuscus* (*Eptesicus fuscus*): XM_028141092.1; *C. perspicillata* (*Carollia perspicillata*): MH299896.1; *P. abramus* (*Pipistrellus abramus*): MH345672.1; *M. davidii* (*Myotis davidii*): NW_006293001.1;; *P. alecto* (*Pteropus alecto*): NW_006442484.1; *A. planirostris* (*Artibeus planirostris*): MH299895.1; *M. lucifugus* (*Myotis lucifugus*): XM_014445986.2; *M. brandtii* (*Myotis brandtii*): XM_005859372.2; *E. buettikoferi* (*Epomops buettikoferi*): MH299897.1; *D. rotundus* (*Desmodus rotundus*): XM_024579835.1.

Species	420	421	423	424	425	426	428	501	502	503	504	505	506	528	Virus binding
Human	Y	Q	Y	P	S	G	D	D	D	F	F	V	Y	L	+
Chimpanzee	Y	Q	Y	P	S	G	D	D	D	F	F	V	Y	L	+
Gorilla	Y	Q	Y	P	S	G	D	D	D	F	F	V	Y	L	+
Dog	Y	Q	Y	P	S	G	D	D	D	F	Y	V	Y	L	+
Ferret	Y	Q	Y	P	S	G	D	D	D	F	F	V	Y	L	+
Pig	Y	H	Y	P	A	G	D	D	F	F	F	V	Y	L	+
Rousettus species	Y	Q	Y	P	A	G	D	D	D	F	F	V	Y	L	+
Pteropus species	Y	Q	Y	P	T	E	D	D	D	F	Y	V	Y	L	+
<i>E. helvum</i>	Y	Q	Y	P	A	G	D	D	F	F	F	T	Y	L	-
<i>E. crypturus</i>	Y	Q	Y	P	A	G	D	D	D	F	F	V	Y	L	+

Table S4. Comparison of the NPC1 residues interfacing with EBOV GP of different species. Blue numbers: most important amino acids. Amino acids differences to human NPC1 are indicated as red characters. Accession numbers: human (*Homo sapiens*): XM_005258278.5; Chimpanzee (*Pan troglodytes*): GABC01005095.1; Gorilla (*Gorilla gorilla gorilla*): XM_019014085.1; dog (*Canis lupus familiaris*): AF315034.1; Ferret (*Mustela putorius furo*): XM_004742838.2; Pig (*Sus scrofa*): AF169635.1; Rousettus species: *R. aegyptiacus* LC462995.; *R. leschenaultia* : LC462996.1, which have identical amino acids; Pteropus species: *P. vampyrus* XM_023530841.1; *P. dasymallus yayeyamae*: LC462999.1; *P. Alecto*: XP_006920089.1, which have identical amino acids; *E. helvum* (*Eidolon helvum*): LC462993.1; *E. crypturus* (*Epomophorus crypturus*): LC462994.1.

Species	106	117	118	119	120	121	122	123	124	125	126	127	128	129	130	133	Virus binding
Human	K	F	Q	E	F	S	P	N	L	W	G	L	E	F	Q	K	++
Horse	R	F	Q	E	F	S	P	N	L	W	G	L	E	F	Q	R	+
Pig	R	F	Q	E	F	S	P	N	L	W	G	L	E	F	Q	R	+
Mouse	R	F	Q	E	F	S	P	N	L	W	G	L	E	F	Q	K	-
<i>Pteropus</i> species	R	F	Q	E	F	S	P	N	L	W	G	L	E	F	Q	K	+
<i>E. fuscus</i>	R	F	Q	E	F	S	P	N	L	W	G	L	E	F	K	K	+
<i>M. natelensis</i>	R	F	Q	E	F	S	P	N	L	W	G	L	E	F	E	K	+

Table S5. Comparison of the ephrin-b2 residues of different species interfacing with Nipah virus G-protein. Amino acids differences to human ephrin-b2 are indicated as red characters. Accession numbers: Human (*Homo sapiens*): NM-004093.4; Horse (*Equus caballus*): NC-009160.3; Pig (*Sus scrofa*): NC-010453.5; Mouse (*Mus musculus*): BC057009.1. Pteropus specie indicates the identical sequences of *P. alecto*: NM-001290170.1 and *P. vampyrus*: NW-011888864.1. The following bat species also contain identical amino acids at the spike interface: *Desmodus-rotundus*: NW-020093757.1; *Hipposideros-armiger*: NW-017731751.1; *Rhinolophus-ferrumequinum*: NC-046287.1; *Myotis-lucifugus*: NW-005871081.1; *Myotis-brandtii*: NW-005371371.1; *Myotis-davidii*: NW-006287546.1; *Phyllostomus-discolor*: XM-028526285.1. Only *E. fuscus* (*Eptesicus-fuscus*): NW-007370674.1; and *M. natelensis* (*Miniopterus-natelensis*): NW-015504348.1 contains one amino acid substitution.