ZOOLOGICAL RESEARCH

# Mutations in spike protein and allele variations in ACE2 impact targeted therapy strategies against SARS-CoV-2

Chuan-Jun Shu[1,2,3,#], Xuan Huang[4,#], Hui-Hao Tang[2], Ding-Ding Mo[5], Jian-Wei Zhou[1,*], Cheng Deng[2,*]

[1] *Department of Molecular Cell Biology & Toxicology*, *Center for Global Health*, *School of Public Health*, *Nanjing Medical University*, *Nanjing*, *Jiangsu* 211166, *China*

[2] *Jiangsu Key Laboratory for Biodiversity and Biotechnology*, *College of Life Sciences*, *Nanjing Normal University*, *Nanjing*, *Jiangsu* 210023, *China*

[3] *Department of Bioinformatics*, *School of Biomedical Engineering and Informatics*, *Nanjing Medical University*, *Nanjing*, *Jiangsu* 211166, *China*

[4] *Reproductive Medical Center*, *Jinling Hospital Affiliated to Medical School of Nanjing University*, *Nanjing*, *Jiangsu* 210002, *China*

[5] *School of Chemical Biology and Biotechnology*, *Peking University Shenzhen Graduate School*, *Shenzhen*, *Guangdong* 518055, *China*

**ABSTRACT**

Coronavirus disease 2019 (COVID-19), which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has spread rapidly worldwide with high rates of transmission and substantial mortality. To date, however, no effective treatments or enough vaccines for COVID-19 are available. The roles of angiotensin converting enzyme 2 (ACE2) and spike protein in the treatment of COVID-19 are major areas of research. In this study, we explored the potential of ACE2 and spike protein as targets for the development of antiviral agents against SARS-CoV-2. We analyzed clinical data, genetic data, and receptor binding capability. Clinical data revealed that COVID-19 patients with comorbidities related to an abnormal renin-angiotensin system exhibited more early symptoms and poorer prognoses. However, the relationship between ACE2 expression and COVID-19 progression is still not clear. Furthermore, if ACE2 is not a good targetable protein, it would not be applicable across a wide range of populations. The spike-S1 receptor-binding domain that interacts with ACE2 showed various amino acid mutations based on sequence analysis. We identified two spike-S1 point mutations (V354F and V470A) by receptor-ligand docking and binding enzyme-linked immunosorbent assays. These variants enhanced the binding of the spike protein to ACE2 receptors and were potentially associated with increased infectivity. Importantly, the number of patients infected with the V354F and V470A mutants has



Received: 14 October 2020; Accepted: 15 March 2021; Online: 16 March 2021

Foundation items: This work was supported by the National Key Research and Development Program of China (2018YFD0900602), National Natural Science Foundation of China (31970388, 31701234), Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), Natural Science Foundation of the Jiangsu Higher Education Institutions (17KJB180006), Natural Science Foundation from Jiangsu Province (BK20160043, BK20151546, 15KJA180004 and BK20171035), and Jiangsu Distinguished Professor Funding

[#]Authors contributed equally to this work

*Corresponding authors, E-mail: jwzhou@njmu.edu.cn; dengcheng@njnu.edu.cn

DOI: 10.24272/j.issn.2095-8137.2020.301

increased with the development of the SARS-CoV-2 pandemic. These results suggest that ACE2 and spike-S1 are likely not ideal targets for the design of peptide drugs to treat COVID-19 in different populations.

**Keywords:** SARS-CoV-2; COVID-19; ACE2; Spike protein; Receptor-ligand docking; Drug therapy

## INTRODUCTION

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic has spread rapidly and widely since the end of 2019 (Holshue et al., 2020). The World Health Organization (WHO) declared coronavirus disease 2019 (COVID-19), which is caused by SARS-CoV-2, a global pandemic (https://www.who.int/) in March 2020. As of mid-January 2021, there have been 89 416 559 confirmed cases and 1 935 028 deaths, affecting approximately 223 countries. COVID-19 symptoms can range from mild, self-limited respiratory disease to severe progressive pneumonia, multiple organ failure, and death (Huang et al., 2020a).

To date, no specific therapeutic agents or sufficient vaccines that have been approved for COVID-19 therapy (Shi et al., 2020b). Several therapies, such as remdesivir, favipiravir, inactivated vaccine, chloroquine, hydroxychloroquine, convalescent plasma, and neutralizing antibodies from convalescent plasma, are under investigation, but the therapeutic efficacy of these drugs and treatments is not yet known, and clinical trial results remain controversial (Beigel et al., 2020; Chen et al., 2020; Gao et al., 2020; Li et al., 2020a; Shen et al., 2020; Shi et al., 2020b; Wang et al., 2020a). For example, several studies have shown that chloroquine, hydroxychloroquine, and convalescent plasma are potential therapeutic options for COVID-19 patients (Huang et al., 2020b; Shen et al., 2020; Wang et al., 2020b), whereas other studies have indicated that these three therapies do not significantly shorten the time to clinical improvement when compared with standard treatments alone (Geleris et al., 2020; Li et al., 2020a).

Peptide drugs are widely applied in medical treatment due to their relative safety, low production complexity, and remarkable curative action. Consequently, they have garnered increasing interest as antiviral agents (Boas et al., 2019). The design of peptide drugs usually begins with functional protein fragments (Mirochnik et al., 2009). Research has shown that the SARS-CoV-2 spike protein binds to receptor angiotensin converting enzyme 2 (ACE2) for viral entry (Wrapp et al., 2020). The receptor binding domain (RBD) that interacts with ACE2 is located at the S1 region, i.e., spike-S1 (Wang et al., 2020a; Wrapp et al., 2020). ACE2 and spike-S1 are considered attractive targets for the development of antiviral agents against human coronaviruses, such as SARS-CoV-2 and SARS-CoV. Therefore, peptide-based drugs could potentially be utilized for the treatment of SARS-CoV-2 infection by targeting viral proteins (such as spikes), mimicking ACE2, and disrupting spike-ACE2 interactions. In theory, a functional fragment of ACE2 and spike-S1 could be utilized against SARS-CoV-2 to decrease the receptor binding of SARS-CoV-2 for COVID-19 therapy.

However, amino acid substitutions in SARS-CoV-2 and ACE2 could be major challenges for the above strategy. Notably, amino acid substitutions in spike-S1 have the potential to generate SARS-CoV-2 variants with increased infectivity. Furthermore, comorbidities have significant indications regarding COVID-19 outcome (Zheng et al., 2020). Therefore, comorbidities should be considered for drugs designed to block SARS-CoV-2 entry into cells or clear SARS-CoV-2.

In this study, certain underlying diseases that appear to aggravate COVID-19 patient condition were explored based on clinical data. Analysis identified several problems associated with designing peptide drugs that target ACE2. Furthermore, similar to the allele frequency in spike interaction domains of ACE2, genomic analysis of SARS-CoV-2 and binding enzyme-linked immunosorbent assay (ELISA) of the receptor-ligand pair (ACE2-spike-S1 interaction) identified two spike-S1 point mutations with increased ACE2 binding and associated infectivity. These results suggest that mutations in ACE2 and spike-S1 may pose a challenge for the exploration of potential broad-spectrum peptide drugs for the treatment of COVID-19 in different populations.

## METHODS AND MATERIALS

### Data preparation

Clinical data, including age, sex, clinical features, computed tomography (CT) imaging, comorbidities, complications, treatment processes, and time information for the first 37 patients who died of COVID-19 in China, were downloaded from the public database of the Hubei CDC (Center for Disease Control) in China (http://www.hbcdc.cn/).

The three-dimensional (3D) structures of the ACE2 and spike protein of SARS-CoV-2 were downloaded from the Protein Data Bank (PDB ID: 6VSB) (Wrapp et al., 2020). The complex structure was determined by cryogenic electron microscopy (Cryo-EM). Coding variants in ACE2 and corresponding allele frequency differences between populations were obtained from the China Metabolic Analytics Project (ChinaMAP) (Cao et al., 2020b), Genome Aggregation Database (gnomAD) (Lek et al., 2016), 1000 Genomes Project (1KGP) (The 1000 Genomes Project Consortium, 2015), Trans-Omics for Precision Medicine (TOP Med) Program (Taliun et al., 2021), and United Kingdom 10K (UK10K) Project (The UK10K Consortium, 2015). Whole-genome sequences of SARS-CoV-2, SARS, bat-CoV, and pangolin-CoV were downloaded from the China National Center for Bioinformation (CNCB)/Beijing Institute of Genomics (BIG) database (https://bigd.big.ac.cn/ncov), Global Initiative of Sharing All Influenza Data (GISAID) EpiFlu™ database, and National Center for Biotechnology Information (NCBI) database. According to gene locations, the nucleotide

sequences of RBD regions of coronavirus and their corresponding amino acids were acquired via BioEdit v2.0 (Tippmann, 2004).

**Genetic and phylogenetic analyses**

The allele frequencies of variants located in the spike-binding regions of ACE2 were summarized. The database variants that affect amino acid sequences for spike-binding regions of ACE2 were then identified. Synonymous (dS)/nonsynonymous substitution (dN) sites between coronaviruses were calculated by DnaSP and BioEdit (Rozas et al., 2017). The dN and dS values were also calculated using DnaSP (Rozas et al., 2017). The nucleotide/amino acid sequences of proteins and similarities between sequences were aligned using BioEdit. Unrooted tree topology based on multiple alignments of amino acids was established with the neighbor-joining method in MEGA 6.06 (Lewis et al., 1995). Branching consistency was tested using bootstrap analysis with 500 resamplings of the data in MEGA 6.06. We numbered the residues in spike-S1 with the deletion of the signal peptide. We also numbered S2 starting with 1. The main reason is that we wanted to unify residues between the sequence and corresponding structure.

**Structural analysis**

SWISS-Model was used to construct the structure of spike-S1 mutants based on the wild-type SARS-CoV-2 spike-S1 protein structure (Waterhouse et al., 2018). We used the ProtParam tool (https://web.expasy.org/protparam/) to analyze the physical and chemical parameters of proteins (Garg et al., 2016). The largest possible binding pocket of these proteins (ACE2, spike-S1 and its mutants) was predicted by Discovery Studio v3.0 (Gao & Huang, 2011). These predicted pockets were utilized to construct an initial coarse model of the spike-S1-ACE2 complexes. The complex structures were refined by Rosetta software (RosettaDock and FlexPepDock modules) (Rohl et al., 2004). The final structure was obtained based on energy scores. The interaction scores were calculated by Rosetta. High-quality 3D images of proteins were drawn by PyMOL (Ordog, 2008).

**Expression of spike-S1 protein and related mutants and ligand receptor-binding ELISA**

The SARS-CoV-2 spike-S1 and related mutants were cloned into the pcDNA3.1 vector and C-terminally fused to a flag-his-tag. Two point mutants (V354F and V470A) were constructed by site-directed mutagenesis, and the vectors were transfected into HEK293 cells. The expressed proteins were secreted into the medium due to the presence of their own signal peptides. Proteins were purified via Ni-NTA columns. The eluted protein was dialyzed against phosphate-buffered saline (PBS) for ligand-receptor binding assays.

The human ACE2 protein (Nouweisen Company, China) was immobilized onto a microtiter plate at 2 μg/mL overnight. The spike-S1 recombinant protein was added as a ligand at different concentrations (0 μg/mL to 8 μg/mL) and then incubated for 1.5 h at 37 °C to allow receptor-ligand interaction. Horseradish peroxidase (HRP) anti-FLAG tag antibody (Bioworld, China) (diluted 1:8 000) was added for 1 h after washing three times with PBS. After adding PBS washing buffer another three times, the signal was visualized and optical density at 450 nm (OD450) was recorded using a microtiter plate reader. Experiments were repeated independently at least three times. The results were analyzed using GraphPad Prism 5 (www.graphpad.com).

**Statistical analysis**

Data were analyzed using GraphPad Prism 8. Statistical significance ($P$<0.05) between the means of two groups was determined using Dunnett's $t$-test or chi-square test.

## RESULTS

**Most COVID-19 patients who died exhibited abnormal regulation of renin-angiotensin system**

ACE2, which plays a critical role in the regulation of the renin-angiotensin system (RAS), is a functional receptor for coronaviruses, including SARS-CoV, Middle East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV-2 (Zheng et al., 2020). To explore the therapeutic potential of ACE2 inhibitors for COVID-19 patients, clinical data on the first 37 COVID-19 patients who died in China were downloaded from publicly available data of the Hubei CDC in China. Clinical data consisted of age, sex, clinical features, CT imaging, comorbidities, complications, treatment processes, and time point information for disease progression. According to the comorbidity data, COVID-19 deaths were associated with hypertension, diabetes, cardiocerebrovascular diseases, atherosclerosis, liver cirrhosis, myxoma, Parkinson's disease, and chronic bronchitis (patients: 23/37, 62.16%). All COVID-19 patients that died usually had more than one comorbidity. Hence, the ratio of COVID-19 death patients with the top four comorbidities to all COVID-19 death patients was 19/37 (51.35%). The top four comorbidities were associated with an abnormal RAS (Figure 1A). This suggests that most COVID-19 patients who died had comorbidities, especially RAS-associated comorbidities.

To analyze the effects of abnormal RAS on COVID-19 patients, the patients were divided into three groups: i.e., G1 (COVID-19 patients with comorbidities associated with abnormal RAS, 19 patients), G2 (COVID-19 patients with comorbidities but normal RAS, four patients), and G3 (COVID-19 patients without comorbidities, 14 patients). The main treatments were oxygen inhalation, anti-infective therapy, and antiviral therapy. The main complications were respiratory failure and shock, with CT images also indicating ground glass opacity and infectious lesions in the lungs. These results indicated that the treatments, complications, and CT data in the three groups were similar (Supplementary Table S1). Therefore, the three groups were compared under four variables, i.e., age, days from symptoms to death, sex, and type of early symptoms. The average ages (mean±*SEM*

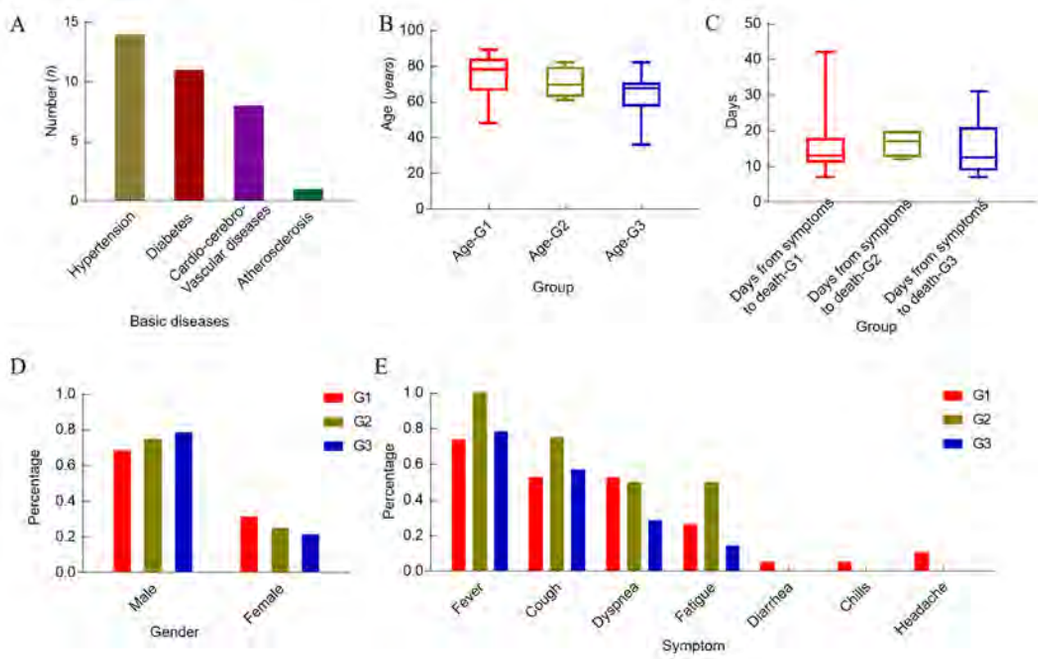


**Figure 1 Clinical data on COVID-19 patients who died**

A: Comorbidities in COVID-19 deaths. B: Age distribution in three groups, i.e., G1 (COVID-19 patients with comorbidities associated with abnormal RAS, 75.00±2.57 (mean±*SEM* (standard error of the mean))), G2 (COVID-19 patients with comorbidities but normal RAS, 70.50±4.44), and G3 (COVID-19 patients without comorbidities, 65.29±3.18). For age distribution, *P*-values between G1 and G2, G1 and G3, and G2 and G3 were 0.79, 0.80, and 0.70, respectively. C: Days of symptoms until death (time from early symptoms to death in G1 (15.47±1.71), G2 (16.50±2.06), and G3 (15.43±2.20) groups). For days of symptoms until death, *P*-values between G1 and G2, G1 and G3, and G2 and G3 were 0.36, 0.67, and 0.28, respectively. D: Sex ratio in three groups (G1-male (0.68), G2-male (0.75), G3-male (0.79); G1-female (0.32), G2-female (0.25), G3-female (0.21)). E: Early symptoms for patients in three groups (fever (G1: 0.74, G2: 1, G3: 0.79); cough (G1: 0.53, G2: 0.75, G3: 0.57); dyspnea (G1: 0.53, G2: 0.50, G3: 0.29); fatigue (G1: 0.26, G2: 0.50, G3: 0.14); diarrhea (G1: 0.05); chills (G1: 0.05); and headache (G1: 0.11).

(standard error of the mean)) for G1, G2, and G3 were 75.00±2.57, 70.50±4.44, and 65.29±3.18, respectively (Figure 1B). Days from symptoms to death (mean±*SEM*) for G1, G2, and G3 were 15.47±1.71, 16.50±2.06, and 15.43±2.20, respectively (Figure 1C). The sex ratios (male/female) for G1, G2, and G3 were 2.17, 3.00, and 3.67, respectively (Figure 1D). These data suggest that a higher proportion of men than women died from COVID-19, which may be due to the higher expression levels of ACE2 in males (Zhao et al., 2020). According to the average age of COVID-19 patients with death outcome as well as *P*- values in the three groups (p-G1 vs. G2: 0.79; p-G1 vs G3: 0.80; p-G2 vs G3: 0.70), most COVID-19 patients who died were older males. As shown in Figure 1E, COVID-19 patients with comorbidities associated with abnormal RAS had more early symptoms than those without comorbidities (G1: 7, G2: 4, G3: 4), including fever, cough, dyspnea, fatigue, diarrhea, chills, and headache. These results indicate that older males with abnormal RAS accounted for the largest proportion of COVID-19 deaths and COVID-19 patients who died often exhibited numerous early symptoms.

Generally, ACE inhibitors are utilized to treat comorbidities with abnormal RAS (Fang et al., 2020). ACE inhibitors can affect the presence of ACE2 by decreasing the receptor levels for usage by SARS-CoV-2 (Fang et al., 2020; Li et al., 2017). The high expression of ACE2 in humans with comorbidities and abnormal RAS may facilitate SARS-CoV-2 entry into target cells in the respiratory and gastrointestinal systems and prolong viral clearance time (Zheng et al., 2020). Hence, COVID-19 patients with comorbidities associated with abnormal RAS often showed many early symptoms and exhibited poor prognosis. However, the relationship between ACE2 expression and COVID-19 progression is not clear, as no significant differences were found between G1 and G3 in days from symptoms to death (*P*=0.67) and patients with abnormal RAS always showed high ACE2 expression. Several studies in mice have indicated that ACE2 plays a critical role in responding to injury in the lungs (Gu et al., 2016; Hamzelou, 2020; Imai et al., 2007). Inhibiting ACE2 function likely prevents injuries from healing. Therefore, ACE2 can act as a good target for drugs that block SARS-CoV-2 infection; however, we do not yet know whether ACE2 is safe or applicable across a wide range of populations. Peptides that prevent interaction between SARS-CoV-2 and its ACE2 receptor, but exert no effect on receptor enzyme activity, could potentially be utilized to treat COVID-19 patients.

**Spike-S1 as a potential target for COVID-19**

SARS-CoV-2 uses the S1 subunit in a densely glycosylated spike (spike-S1) protein to bind to the host cell receptor (ACE2) to gain entry into host cells (Wrapp et al., 2020). Hence, to explore potential broad-spectrum peptide drugs for treating COVID-19 in different populations, variants in potential targets, i.e. , ACE2 and spike-S1, should be analyzed. The cryo-EM complex structures of ACE2 and spike protein were downloaded from the PDB database (PDB ID: 6VSB) (Wrapp et al., 2020). Binding sites between ACE2 and spike-S1 were analyzed, as shown in Figure 2A (Supplementary Table S2). There were 14 binding sites in spike-S1 and 12 binding sites in ACE2. The 12 binding sites in ACE2 were distributed among three interaction regions: i.e., 24–41 residues, 79–83 residues, and 330–353 residues (Figure 2B). In addition, 19 amino acid variants in the spike-binding regions of ACE2 were identified based on the ChinaMAP (Cao et al., 2020b), gnomAD (Lek et al., 2016), 1KGP (The 1000 Genomes Project Consortium, 2015), TOP Med program (Taliun et al., 2021), and UK10K project (Consortium et al., 2015) databases (Figure 2C). Only one amino acid variant, K26R (rs ID (identity document): rs4646116), was distributed in the spike-binding regions of ACE2, but the variant site did not belong to the interaction sites for ACE2-spike-S1 (Figure 2A, C ). As shown in Figure 2D, the allele frequencies for different populations were different (Supplementary Table S3). The highest allele frequency for this variant was 0.01 (gnomAD (ASJ (Ashkenazi Jewish))). The differences in allele frequencies of ACE2 coding variants among different populations suggest that diverse genetic backgrounds may affect ACE2 function. These results indicate that, in different populations, ACE2 may exhibit different binding capabilities for spike-S1 in SARS-CoV-2. Therefore, a drug that could inhibit the binding of one ACE2 variant in SARS-CoV-2 may not work in other patients with different ACE2 variants.

To analyze amino acid variants in the spike-S1 RBD, 10 992 complete genomes of SARS-CoV-2 were downloaded from the GISAID EpiFlu™ and CNCB/BIG databases (https://bigd.big.ac.cn/ncov). There were 14 binding sites in spike-S1 distributed in one RBD, i.e. , 417–505 residues (Figure 2E). In total, 26 amino acid variants were identified in the spike-S1 RBD (Figure 2F). These amino acid variants were distributed among 66 SARS-CoV-2 viruses

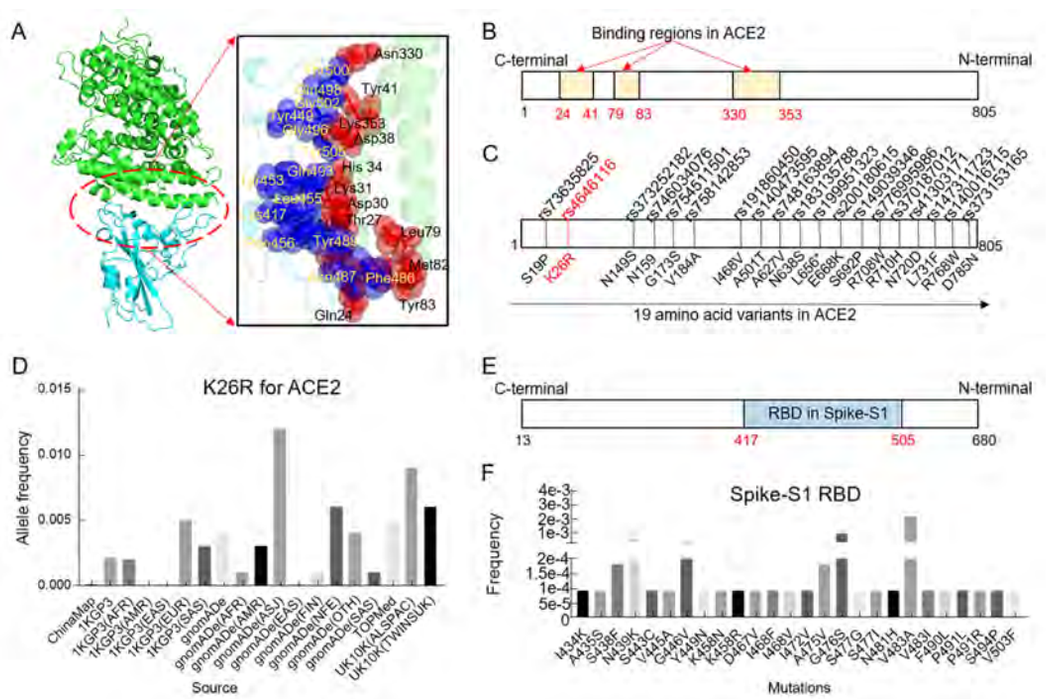


**Figure 2 Variants in binding regions of ACE2 and spike-S1**

A: Binding sites in ACE2 and spike-S1. Green and cyan schematic represents 3D structure of ACE2 and spike-S1, respectively. Yellow letters and blue spheres show spatial positions for ACE2-binding sites in spike-S1. Black letters and red spheres show spatial positions for spike-binding sites in ACE2. B: Spike-binding domain in ACE2. C: Nineteen amino acid substitutions in ACE2. D: Allele frequency for K26R in ACE2 for different populations. AFR, AMR, EAS, EUR, SAS, NFE, ALSPAC, TWINSUK, and OTH represent African/African American, Admixed American, East Asian, Europeans, South Asian, Non-Finnish European, Avon Longitudinal Study of Parents and Children, UK Adult Twin registry, and Other (population not assigned), respectively. E: ACE2-binding domain in spike-S1. F: Amino acid substitutions in ACE2-binding domain for spike-S1 and their corresponding frequencies.

(Supplementary Table S4). Although variant site Y449N belonged to the interaction sites for ACE2-spike-S1, the frequency of the spike-S1 variant was 9.10E-05. The highest frequency for spike-S1 variants was <2.09E-03 (Figure 2F, Supplementary Table S4). The corresponding dN, dS, and dN/dS (ω) values of the genome and spike-S1 sequences for SARS-CoV-2 spike-S1 RBD variants were calculated by DnaSP6 (Rozas et al., 2017). Compared to the first genome sequence of SARS-CoV-2 (GenBank ID: MN908947.3), the average dN, dS, and dN/dS (ω) values (mean±*SEM*) for the genome sequences of SARS-CoV-2 spike-S1 RBD variants were 3.08E-04±3.99E-5, 4.91E-4±4.36E-5, and 0.64±5.87E-2, respectively (Supplementary Table S5). Compared to spike-S1 of the reference SARS-CoV-2, the average dN, dS, and dN/dS (ω) values (mean±*SEM*) for spike-S1 sequences of the SARS-CoV-2 spike-S1 RBD variants were 1.13E-3±1.08E-4, 3.31E-4±1.48E-4, and 0.46±0.14, respectively (Supplementary Table S5). According to these values, there were no positive or negative amino acid substitutions. These results indicate the presence of many amino acid substitutions in the RBD of spike-S1, but these substitutions are probably random events. Therefore, spike-S1 could be a potential target for exploring broad-spectrum peptide drugs for COVID-19 if spike-S1 mutants have the same binding capability as ACE2.

**Enhanced receptor-ligand binding capability in two spike-S1 mutants**

Initially, we compared 154 complete genomes of SARS-CoV-2 downloaded from the GISAID EpiFlu™ (access date 22 February 2020) and CNCB/BIG databases (https://bigd.big.ac.cn/ncov, access date 22 February 2020) to explore variations in SARS-CoV-2 in the early stages of the epidemic. The aligned genome size of SARS-CoV-2 was 29 674 bp and was comprised of 10 protein-coding sequence (CDS) regions (Figure 3A). The SARS-CoV-2-Wuhan01 virus (GenBank ID: MN908947.3) was compared to bat-CoV-RaTG13 (GISAID ID: EPI_ISL_402131), pangolin-CoV-Guangxi (GISAID ID: EPI_ISL_410538), and pangolin-CoV-Guangdong (GISAID ID: EPI_ISL_410721) genomes to assess the potential origin of SARS-CoV-2 before it infected humans. Compared with these three coronaviruses, the whole-genome sequence identity of SARS-CoV-2-Wuhan01 was 0.85 to 0.96 (bat-CoV-RaTG13: 0.96, pangolin-CoV-Guangxi: 0.85, pangolin-CoV-Guangdong: 0.90), which is higher than that between SARS-CoV-2 and SARS (~0.80). In addition, the dN/dS (ω) values between SARS-CoV-2 and other coronaviruses were: bat-CoV-RaTG13: 0.80E-02/0.17 (0.04); pangolin-CoV-Guangxi: 5.50E-2/0.72 (0.08), pangolin-CoV-Guangdong: 2.5E-2/0.47 (0.08); and SARS:0.11/0.93 (0.12). Compared to SARS-CoV-2-Wuhan01, there were 360/755, 1 215/3 003, and 804/2 033 synonymous/nonsynonymous substitution sites for bat-CoV-RaTG13, pangolin-CoV-Guangxi, and pangolin-CoV-Guangdong, respectively. As shown in Figure 3B, nonsynonymous substitution sites were mainly distributed in the ORF1ab and spike-S1 regions. These results indicate that the two regions had higher substitution rates than other SARS-CoV-2 regions in different hosts (Figure 3B).

Furthermore, to explore the evolution of SARS-CoV-2, we further examined genomic variations of SARS-CoV-2 in humans. The substitution sites in five proteins, i.e. , 3-chymotrypsin-like protease (Mpro), papain-like protease, helicase, RNA-dependent RNA polymerase, and spike-S1, were calculated as they may be potential drug targets for the development of antiviral agents or vaccines against SARS-CoV-2, SARS, and MERS. Subsequently, we identified 138 substitution sites (synonymous sites: 52, nonsynonymous sites: 86) in the 10 CDS regions, and all substitutions could be classified as single nucleotide polymorphisms (SNPs). The substitution sites were also mainly distributed in the ORF1ab and spike-S1 regions (Figure 3C). However, ORF8a, ORF3a, ORF10, and spike-S1 had higher substitution rates in SARS-CoV-2 (Figure 3D). With a nonsynonymous/synonymous site ratio >1, our results indicate that SARS-CoV-2 underwent adaptive evolution after human host infection.

The nonsynonymous substitutions could result in a SARS-CoV-2 variant with increased infectivity. To address potential increased SARS-CoV-2 infectivity, amino acid mutations caused by nonsynonymous substitutions in SARS-CoV-2-Wuhan01 versus SARS-CoV-2 mutants were analyzed. The corresponding amino acids in the bat-CoV-RaTG13, pangolin-CoV, and SARS viruses were then used as controls. As shown in Figure 3E, most amino acid substitutions from the SARS-CoV-2 viruses corresponded to conserved residues in the bat-CoV-RaTG13, pangolin-CoV, and SARS viruses, except for five amino acid substitutions in spike-S1 of SARS-CoV-2. Furthermore, the dN/dS (ω) values for the SARS-CoV-2 mutants were always below 1 (Table 1). These results indicate that most amino acid substitutions in SARS-CoV-2 are likely random events. However, the 10 amino acid substitutions recurred in SARS-CoV-2, i.e. , I789V of papain-like protease, H36Y and V354F of spike-S1, G251V of ORF3a, D209H of membrane glycoprotein, V62L, L84S, and P85S of ORF8, and S194L, S202N, and P344S of nucleocapsid phosphoprotein (Figure 3E). The spike-S1 protein binds to the human cellular receptor ACE2 and is responsible for viral infection. Hence, attention should be paid to amino acid substitutions occurring in the RBD region of spike-S1 or substitutions occurring independently more than three times in the early stages of the epidemic. In addition, epidemiological and evolutionary data for H36Y, N341D, D351Y, and V354F of spike-S1, G251V of ORF3a, V62L and L84S of ORF8, and S194L and S202N of nucleocapsid phosphoprotein were analyzed. The amino acid substitution of ORF8, i.e., L84S, had the highest substitution rate (value=36) among the 10 mutant residues (Figure 3E). The dN/dS (ω) value for the SARS-CoV-2 mutant-L84S was 3.6E-3/0.00 (NA) compared to the wild-type ORF8. However, this site is a reverse mutation compared to its ancestor (Figure 3E). We selected the whole genome of these 36 SARS-CoV-2 mutant viruses and their potential ancestors (bat-CoV and pangolin-CoV viruses) to construct a phylogenetic tree (neighbor-joining, bootstrap=500) (Figure 3F). As shown in Figure 3F, we found that the SARS-CoV-2 mutant with L84S in ORF8

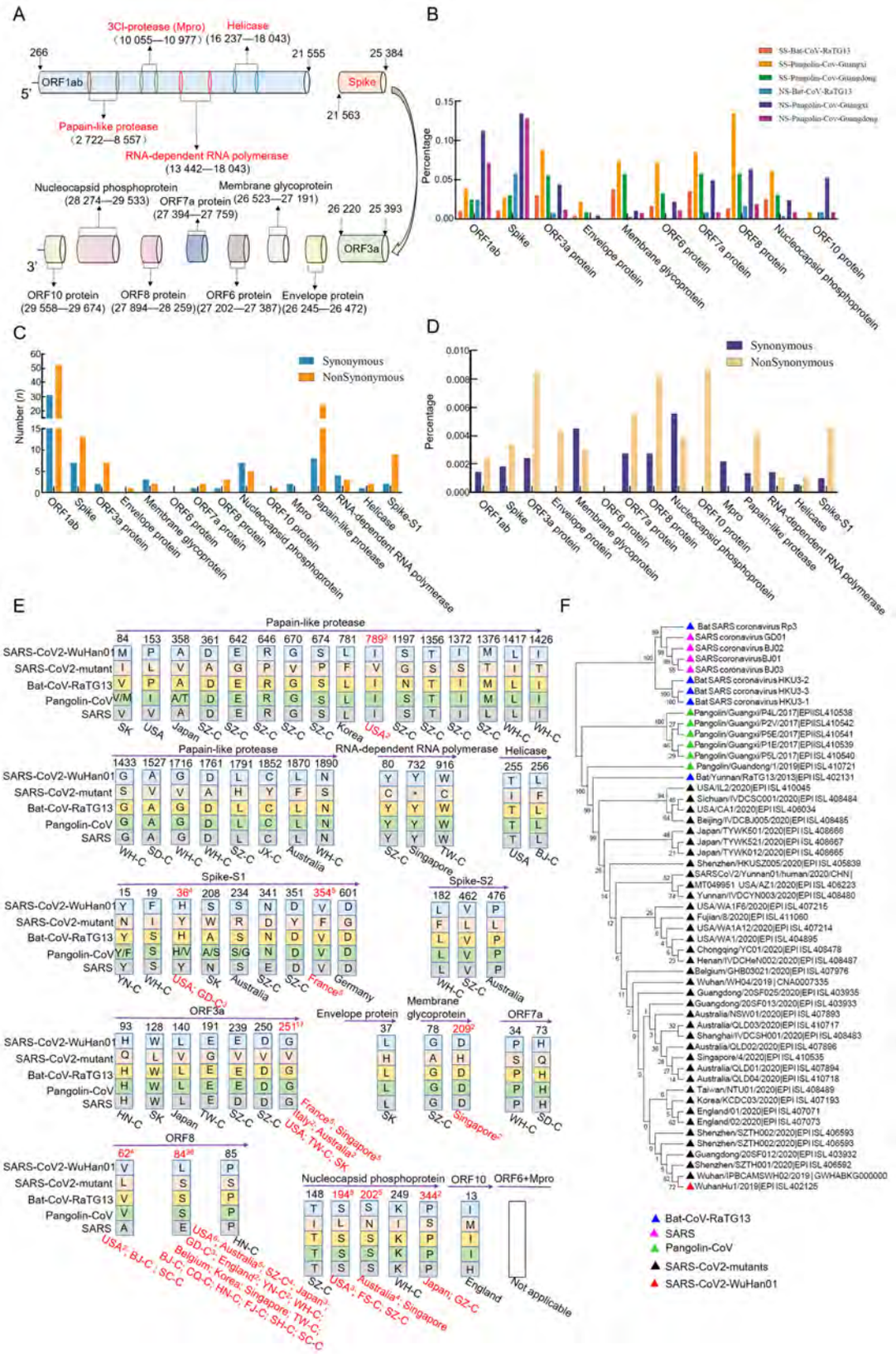
A
3Cl-protease (Mpro) (10 055—10 977)
Helicase (16 237—18 043)
266
21 555
25 384
ORF1ab
Spike
5'
21 563
Papain-like protease (2 722—8 557)
RNA-dependent RNA polymerase (13 442—18 043)
Nucleocapsid phosphoprotein (28 274—29 533)
ORF7a protein (27 394—27 759)
Membrane glycoprotein (26 523—27 191)
26 220
25 393
ORF3a
3'
ORF10 protein (29 558—29 674)
ORF8 protein (27 894—28 259)
ORF6 protein (27 202—27 387)
Envelope protein (26 245—26 472)
B
SS-Bat-CoV-RaTG13
SS-Pangolin-Cov-Guangxi
SS-Pangolin-Cov-Guangdong
NS-Bat-CoV-RaTG13
NS-Pangolin-Cov-Guangxi
NS-Pangolin-Cov-Guangdong
Percentage
ORF1ab
Spike
ORF3a protein
Envelope protein
Membrane glycoprotein
ORF6 protein
ORF7a protein
ORF8 protein
Nucleocapsid phosphoprotein
ORF10 protein
C
Synonymous
NonSynonymous
Number (n)
Mpro
Papain-like protease
RNA-dependent RNA polymerase
Helicase
Spike-S1
D
Synonymous
NonSynonymous
Percentage
E
Papain-like protease
SARS-CoV2-WuHan01
SARS-CoV2-mutant
Bat-CoV-RaTG13
Pangolin-CoV
SARS
RNA-dependent RNA polymerase
Helicase
Spike-S1
Spike-S2
ORF3a
Envelope protein
Membrane glycoprotein
ORF7a
ORF8
Nucleocapsid phosphoprotein
ORF10
ORF6+Mpro
F
Bat SARS coronavirus Rp3
SARS coronavirus GD01
SARS coronavirus BJ02
SARScoronavirusBJ01
SARS coronavirus BJ03
Bat SARS coronavirus HKU3-2
Bat SARS coronavirus HKU3-3
Bat SARS coronavirus HKU3-1
Pangolin/Guangxi/P4L/2017|EPI|SL410538
Pangolin/Guangxi/P2V/2017|EPI|SL410542
Pangolin/Guangxi/P5E/2017|EPI|SL410541
Pangolin/Guangxi/P1E/2017|EPI|SL410539
Pangolin/Guangxi/P5L/2017|EPI|SL410540
Pangolin/Guandong/1/2019|EPI|SL410721
Bat/Yunnan/RaTG13/2013|EPI|SL402131
USA/IL2/2020|EPI|SL410045
Sichuan/IVDCSC001/2020|EPI|SL408484
USA/CA1/2020|EPI|SL406034
Beijing/IVDCBJ005/2020|EPI|SL408485
Japan/TYWK501/2020|EPI|SL408666
Japan/TYWK521/2020|EPI|SL408667
Japan/TYWK012/2020|EPI|SL408665
Shenzhen/HKUSZ005/2020|EPI|SL405839
SARSCoV2/Yunnan01/human/2020/CHN|
MT049951 USA/AZ1/2020|EPI|SL406223
Yunnan/IVDCYN003/2020|EPI|SL408480
USA/WA1F6/2020|EPI|SL407215
Fujian/8/2020|EPI|SL411060
USA/WA1A12/2020|EPI|SL407214
USA/WA1/2020|EPI|SL404895
Chongqing/YC01/2020|EPI|SL408478
Henan/IVDCHeN002/2020|EPI|SL408487
Belgium/GHB03021/2020|EPI|SL407976
Wuhan/WH04/2019 | CNA0007335
Guangdong/20SF025/2020|EPI|SL403935
Guangdong/20SF013/2020|EPI|SL403933
Australia/NSW01/2020|EPI|SL407893
Australia/QLD03/2020|EPI|SL410717
Shanghai/IVDCSH001/2020|EPI|SL408483
Australia/QLD02/2020|EPI|SL407896
Singapore/4/2020|EPI|SL410535
Australia/QLD01/2020|EPI|SL407894
Australia/QLD04/2020|EPI|SL410718
Taiwan/NTU01/2020|EPI|SL408489
Korea/KCDC03/2020|EPI|SL407193
England/01/2020|EPI|SL407071
England/02/2020|EPI|SL407073
Shenzhen/SZTH002/2020|EPI|SL406593
Shenzhen/SZTH002/2020|EPI|SL406593
Guangdong/20SF012/2020|EPI|SL403932
Shenzhen/SZTH001/2020|EPI|SL406592
Wuhan/IPBCAMS/WH02/2019 | GWHABKG000000
WuhanHu1/2019|EPI|SL402125
Bat-CoV-RaTG13
SARS
Pangolin-CoV
SARS-CoV2-mutants
SARS-CoV2-WuHan01

**Figure 3 Summary information for 86 nonsynonymous sites in 10 CDS regions of SARS-CoV-2**

A: Genomic organization of SARS-CoV-2. Red letters represent proteins that are potential drug targets. B: Compared to SARS-CoV-2-Wuhan01, percentages of synonymous/nonsynonymous substitution sites in 10 CDS regions for bat-CoV-RaTG13, pangolin-CoV-Guangxi, and pangolin-CoV-Guangdong. C, D: Show number and percentage of synonymous/nonsynonymous substitution sites in 10 CDS regions and five potential drug targets for SARS-CoV-2, respectively. E: Amino acid substitutions in SARS-CoV-2 mutants and corresponding residues in bat-CoV-RaTG13, pangolin-CoV, and SARS viruses. Red indicates that amino acid substitution rates are greater than 2. F: Phylogenetic tree of SARS-CoV-2 mutants (L84S in ORF8).

**Table 1 Molecular divergence among SARS-CoV-2 and its mutants**

| | dN/ds(ω)-H | dN/ds(ω)-L | dN/ds(ω)>>1 mutant |
|---|---|---|---|
| Mpro | 0/4.70E-3 (0) | 0/4.70E-3 (0) | N/A |
| Papain-like protease | 2.00E-3/3.00E-3 (0.67) | 2.00E-4/7.00E-4 (0.29) | N/A |
| RNA-dependent RNA polymerase | 5.00E-4/0 (N/A) | 0/1.60E-3 (0) | N/A |
| Helicase | 7.00E-4/0 (N/A) | 0/2.30E-3 (0) | N/A |
| Spike-S1 | 7.00E-4/0 (N/A) | 1.30E-3/2.1E-3(0.62) | N/A |
| Spike-S2 | 7.00E-4/0 (N/A) | 0/2.50E-3 (0) | N/A |
| ORF3a | 3.20E-3/0 (N/A) | 0/5.10E-3 (0) | N/A |
| Envelope protein | 6.10E-3/0 (0) | N/A | N/A |
| Membrane glycoprotein | 2.00E-3/6.10E-3(0.33) | 0/1.23E-2 (0) | N/A |
| ORF6 protein | N/A | N/A | N/A |
| ORF7a protein | 3.70E-03/0 (N/A) | N/A | N/A |
| ORF8 protein | 7.20E-3/0 (N/A) | 0/1.22E-2 (0) | N/A |
| Nucleocapsid phosphoprotein | 1.00E-3/0 (N/A) | 0/3.40E-3 (0) | N/A |
| ORF10 protein | 1.15E-2/0 (N/A) | N/A | N/A |

H and L represent highest dN/dS (ω) values for SARS mutants. N/A: Not available.

had a closer phylogenetic relationship to its ancestors compared to that of SARS-CoV-2-Wuhan01 (Figure 3F). Hence, the SARS-CoV-2 mutant L84S was more likely be an ancestral variant than a virus with increased infectivity. Previous research has indicated that the SARS-CoV-2 mutant L84S has reduced infectivity (Tang et al., 2020). Therefore, we focused on SARS-CoV-2 without L84S in ORF8.

The amino acid substitution of ORF3a, i.e., G251V, had the second highest substitution rate (value=17) among the 10 mutant residues (Figure 3E). These SARS-CoV-2 mutants were mainly isolated from patients in France and Singapore. Interestingly, five patients from France with 100% identical protein sequences (Supplementary Table S6) shared the same V354F mutation, and this amino acid substitution was in the RBD region of spike-S1 (V354F). In addition, the SARS-CoV-2 mutant-V354F had no L84S mutant in ORF8. The dN/dS (ω) value for SARS-CoV-2 mutant-V354F was 7.0E-4/0.00 (NA) compared to the wild-type spike-S1. These results suggest that the SARS-CoV-2 mutant with the V354F-spike-S1 mutation potentially exhibits increased infectivity.

Since spike-S1 is a key factor for viruses to enter human cells, we explored whether it might contribute to the increased infectivity of SARS-CoV-2 mutants based on receptor-ligand (ACE2-spike-S1 mutants) docking. Amino acid substitution in the RBD region of spike-S1 (Figure 4A, B) could theoretically affect the infectivity of SARS-CoV-2. Therefore, three mutants found in spike-S1 (N341D, D351Y, and V354F) were utilized to investigate the docking of spike-S1 to ACE2. Based on sequence, physical, and chemical parameters, i.e., theoretical pI (isoelectric point), GRAVY (grand average of hydropathicity), and instability index, were analyzed using the ProtParam tool (https://web.expasy.org/protparam/) for wild-type and mutants. As shown in Figure 4C, the three amino acid substitutions affected the physical and chemical parameters of the spike-S1 protein, with the V354F mutant having the smallest impact among the three mutants. We further tested whether these three mutants could result in different interaction models between ACE2 and spike-S1. With the 3D structure of ACE2 and spike-S1 downloaded from the PDB database (PDB ID: 6ACJ and 6VSB), the spike-S1 mutant structures were obtained using SWISS-MODEL based on a homologous modeling algorithm (Song et al., 2018; Wrapp et al., 2020). Complexes of the ACE2-spike-S1 mutants were obtained by Discovery studio and Rosetta package software. As shown in Figure 4D, V354F had the lowest energy and interaction score in the ACE2-spike-S1 complexes based on structural comparison and protein-protein interaction analysis (Figure 4D, Supplementary Table S6). These results further suggest that spike-S1 with the V354F amino acid substitution had a higher affinity for ACE2 than wild-type spike-S1, which potentially led to increased infectivity. Therefore, given its potential increase in infectivity, SARS-CoV-2 mutant-V354F deserves close epidemic surveillance. Such action could reduce the worldwide threat of a potential explosion in SARS-CoV-2 (Shu et al., 2020) and increase the success rate of therapeutic drug design.

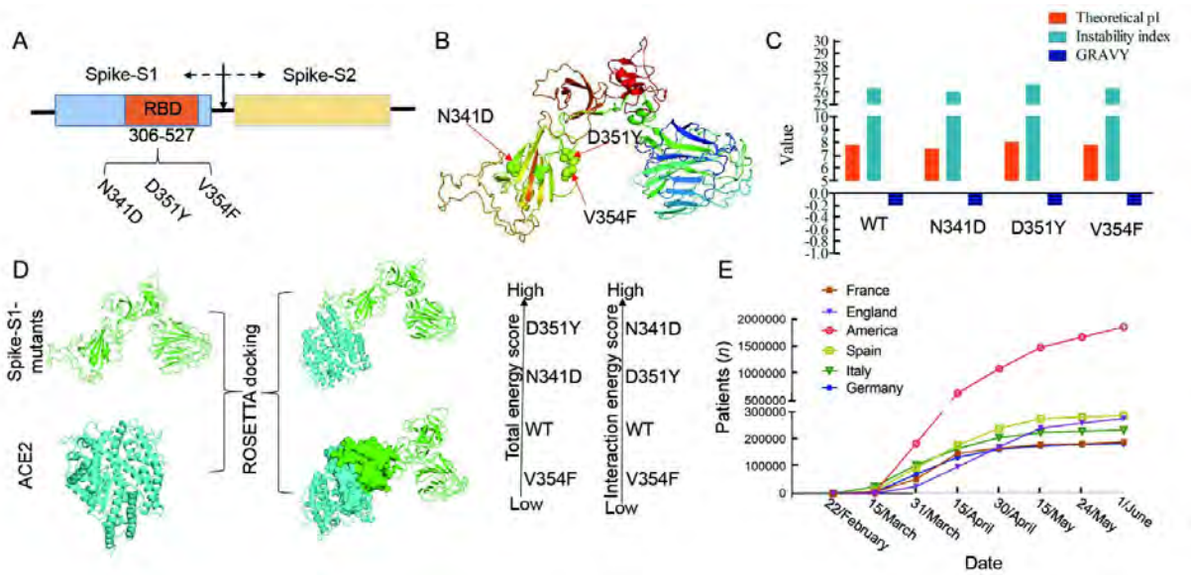


**Figure 4 Structural pharmacological analysis among three spike-S1 protein mutants**

A: RMD domain in spike-S1. B: Distribution of three spike mutations, i.e., N341D, D351Y, and V354F, in 3D structure. C: Physical and chemical parameters for spike-S1 and its mutants. D: Energy and interaction scores for complex structures of ACE2-spike-S1 and its mutants. E: Number of COVID-19 patients in Europe and USA.

To further explore the epidemic impact of the V354F mutant, spike-S1 sequences were downloaded from the GISAID EpiFlu$^{TM}$ and CNCB/BIG databases at three different time points (22 February, 25 March, and 10 April 2020). As shown in Figure 5A and B , an increasing number of amino acid substitutions were seen in the RBD of spike-S1 with the spread of SARS-CoV-2, but most amino acid substitutions were random events, except for V354F, G463S, and V470A. Furthermore, the detected number of patients with these three mutants, i.e. , V354F, G463S, and V470A, increased. By 10 April 2020, based on 16 491 complete genomes of SARS-CoV-2, the number of mutant viruses with V354F, G463S, and V470A of spike-S1 reached 20, 19, and 30, respectively (Figure 5C). According to interaction energy analysis, SARS-CoV-2 strains with the spike-S1 V354F or V470A substitutions were likely novel strains with increased infectivity (Figures 4D, 5A).

To examine the binding capacity of spike-S1 mutants to ACE2, we constructed expression plasmids for two mutants, V354F and V470A (additional variant subsequently reported in the USA in Figure 5A–C) and purified sufficient protein for receptor-ligand binding assay. Both the V354F and V470A variants exhibited increased binding to the human ACE2 receptor compared to the wild-type, consistent with our computer simulation (Figures 4D, 5A). As shown in Figure 5D, both the V354F and V470A mutants had significantly lower $EC_{50}$ (concentration for 50% of maximal effect) values ($EC_{50}$=1.32 μg/mL for V354F and $EC_{50}$=1.24 μg/mL for V470A) than the wild-type ($EC_{50}$=1.86 μg/mL), demonstrating that the V354F and V470A mutants had higher binding affinity and potentially increased infectivity. Therefore, the SARS-CoV-2 mutants (V354F and V470A) with potentially increased infectivity should be under close epidemic surveillance to avoid potential additional waves of SARS-CoV-2 infection worldwide.

Furthermore, animals that are easily accessible to humans should be monitored to prevent cross-infection, which could result in a novel coronavirus with increased infectivity or a new recombination coronavirus (Shi et al., 2020a). These findings highlight the challenge of designing drugs for COVID-19 therapy. As the coronavirus infects human cells via ACE2, corresponding sequences of 16 mammals were downloaded from the NCBI database. Their sequence similarities and phylogenetic relationships are shown in Figure 6A and B, respectively. The sequence logo for the 16 ACE2 sequences is presented in Figure 6C. The key amino acids of human ACE2 for interaction with SARS-CoV-2 spike-S1 were not conserved in the 15 other species. Subsequently, the interaction energy scores between SARS-CoV-2 and ACE2 of the different species were analyzed. We found that *Felis catus*, *Paguma larvata*, *Pan troglodytes*, *Pongo abelii*, and *Sus scrofa* may be more susceptible to SARS-CoV-2 infection, consistent with other studies (Figure 6D) (Shi et al., 2020a). Hence, animals with different ACE2 sequences or key interaction amino acids in ACE2 could affect the infectivity of SARS-CoV-2. Some animals may be infected with SARS-CoV-2 because they have the same interaction residues as human ACE2. The aligned interaction sites in ACE2 were 24, 30, 34, 38, 83, and 353 (Figures 2A, 6C). However, previous study has indicated that there is no direct evidence to support the existence of ACE2 variants that are resistant to coronavirus S-protein binding in different populations (Cao et al., 2020a). Therefore, we should focus our attention on

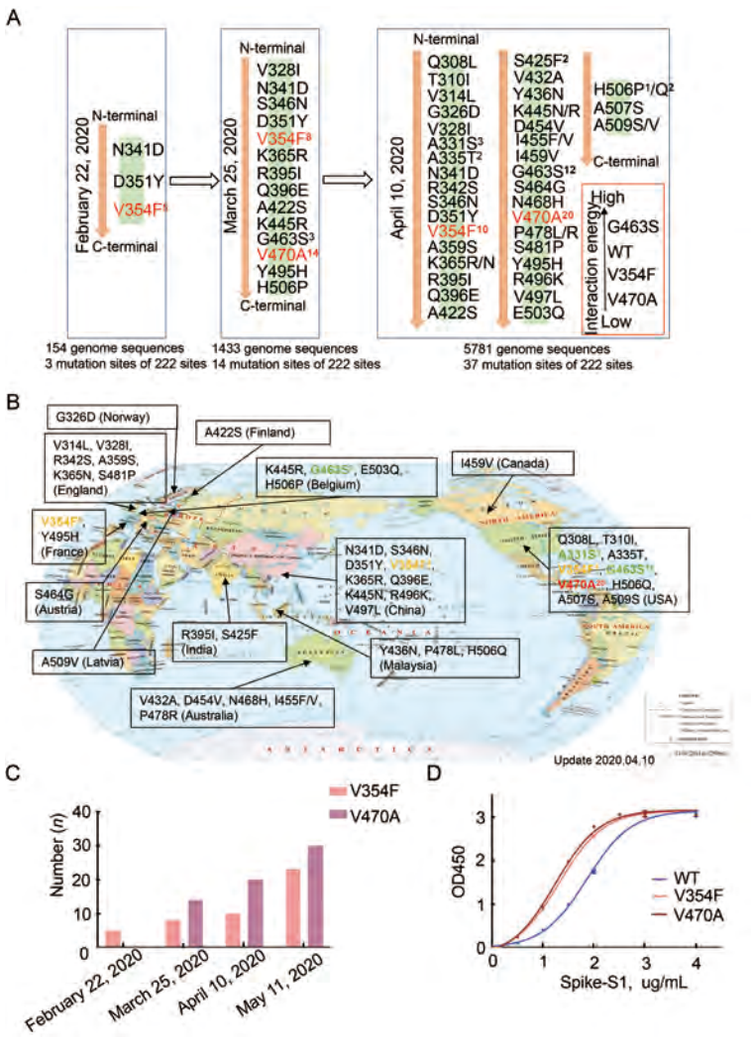


**Figure 5 Epidemic situation for spike-S1 mutants and receptor-ligand binding of spike-S1 mutants**

A: Number of amino acid mutants in spike-S1 at different time points. B: Distribution of spike-S1 mutants worldwide. C: Number of V354F and V470A mutants at different times. D: Optical density at 450 nm (OD450) for wild-type spike-S1 and its mutants, i.e., V354F and V470A mutants.

SARS-CoV-2 mutants, especially mutants with amino acid substitutions in the RBD region of spike-S1. For example, the V354F and V470A mutants have potentially increased infectivity. These mutants may increase the difficulty of developing effective COVID-19 therapies.

## DISCUSSION AND CONCLUSIONS

Based on clinical data obtained from COVID-19 patients, we found that patients with abnormal RAS often exhibited more early symptoms and poorer outcomes after standard treatment. ACE2, which is a functional receptor for SARS-CoV-2, plays a critical role in the process of RAS. However, the relationship between ACE2 expression and COVID-19 progression is not clear. In different populations, ACE2 demonstrates different binding capabilities for spike-S1 in SARS-CoV-2. Furthermore, the amino acid substitutions present in the RBD of spike-S1 are likely random events. Hence, compared to ACE2, the spike-S1 RBD acts as a good target for broad-spectrum peptide drugs that could block SARS-CoV-2 infection in different populations. Although SARS-CoV-2 genomic variations are still low, some nonsynonymous substitutions have led to SARS-CoV-2 variants with increased infectivity in humans. Here, we found that SARS-CoV-2 with V354F or V470A substitutions of spike-S1 showed increased infectivity. Both V354F and V470A are also known to increase the entry of pseudoviruses into cells expressing ACE2 (Li et al., 2020b). These findings increase the challenge of designing peptide drugs for COVID-19 therapy and represent potentially dangerous virus strains. Although the clinical COVID-19 data and number of mutants tested for infectivity are limited in this study, the challenges for the treatment of COVID-19 based on existing analyses are clear. As mutations or allele variations result in different target domains, ACE2 and spike-S1 do not appear to be ideal

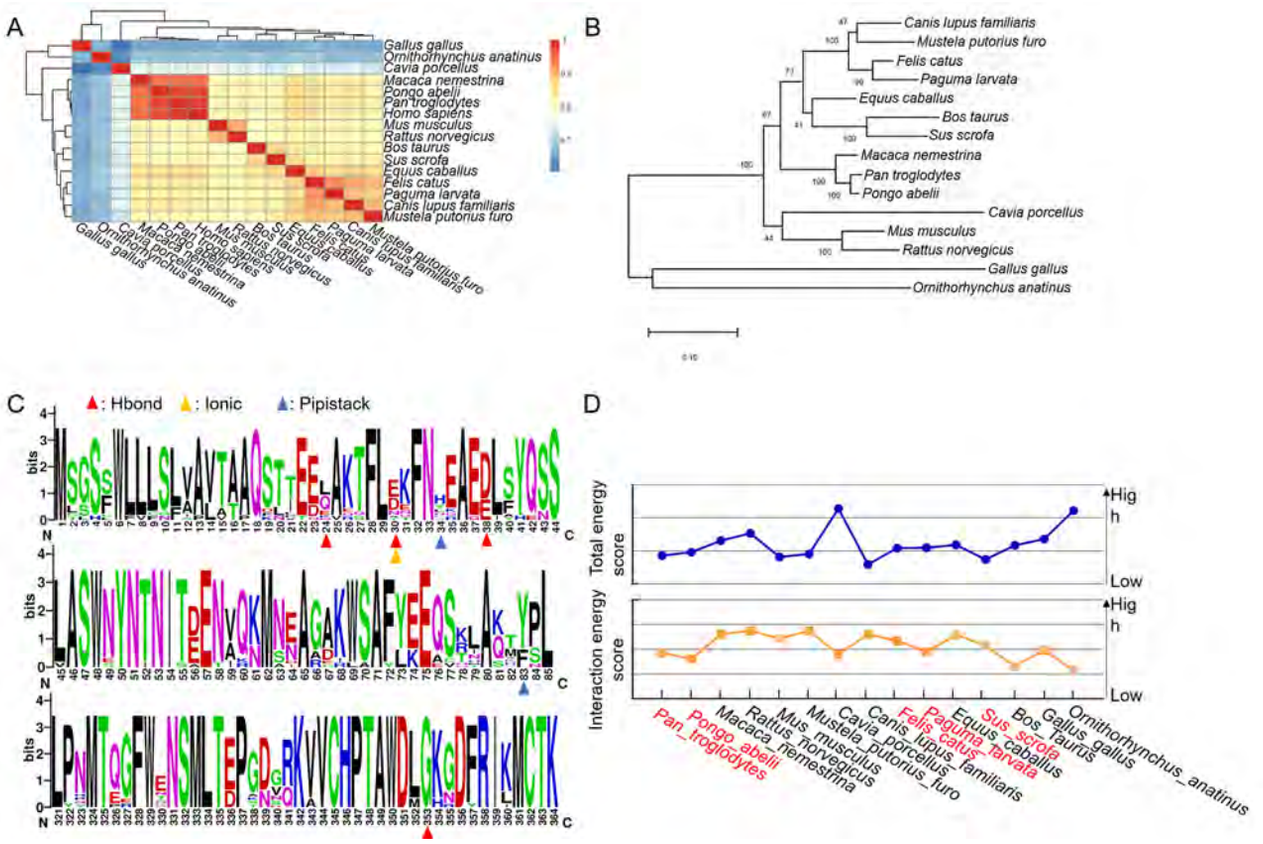


**Figure 6 Different ACE2 sequences have different binding energies to spike-S1**

A: Sequence similarity between different ACE2 sequences. B: Phylogenetic tree for different ACE2 sequences. C: Sequence logo for ACE2. Red, yellow, and blue triangles represent ACE2 residues that could potentially form a hydrogen bond with the spike-S1 protein of SARS-CoV-2-Wuhan01 and an ionic or PiPistack interaction with the corresponding residues in spike-S1. D: Binding interaction energy scores for ACE2 and spike-S1.

targets for the design of small-molecule drugs to treat COVID-19 across all populations.

## SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

## COMPETING INTERESTS

The authors declare that they have no competing interests.

## AUTHORS' CONTRIBUTIONS

C.J.S. and X.H. designed and carried out the study and drafted the manuscript. C.J.S., C.D. and H.H.T. performed and analyzed the experiments. C.J.S., C.D., D.D.M., J.W.Z. conceived the study and were the main writers of the manuscript. All authors read and approved the final version of the manuscript.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the GISAID EpiFlu™ (www.gisaid.org) and CNCB/BIG (https://bigd.big.ac.cn/ncov) databases.

## REFERENCES

Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, et al. 2020. Remdesivir for the treatment of Covid-19-preliminary report. *New England Journal of Medicine*, **383**(10): 992−994.

Boas LCPV, Campos ML, Berlanda RLA, De Carvalho Neves N, Franco OL. 2019. Antiviral peptides as promising therapeutic drugs. *Cellular and Molecular Life Sciences*, **76**(18): 3525−3542.

Cao YN, Li L, Feng ZM, Wan SQ, Huang PD, Sun XH, et al. 2020a. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discovery*, **6**(1): 11.

Cao YN, Li L, Xu M, Feng ZM, Sun XH, Lu JL, et al. 2020b. The ChinaMAP analytics of deep whole genome sequences in 10, 588 individuals. *Cell Research*, **30**(9): 717−731.

Chen C, Zhang Y, Huang JY, Yin P, Cheng ZS, Wu JY, et al. 2020. Favipiravir versus Arbidol for COVID-19: a randomized clinical trial. *medRxiv*, doi: 10.1101/2020.03.17.20037432.

Fang L, Karakiulakis G, Roth M. 2020. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection?. *The Lancet Respiratory Medicine*, **8**(4): e21.

Gao Q, Bao LL, Mao HY, Wang L, Xu KW, Yang MN, et al. 2020. Development of an inactivated vaccine candidate for SARS-CoV-2.

*Science*, **369**(6499): 77−81.

Gao YD, Huang JF. 2011. An extension strategy of discovery studio 2.0 for non-bonded interaction energy automatic calculation at the residue level. *Zoological Research*, **32**(3): 262−266.

Garg VK, Avashthi H, Tiwari A. 2016. MFPPI–Multi FASTA ProtParam interface. *Bioinformation*, **12**(2): 74−77.

Geleris J, Sun Y, Platt J, Zucker J, Baldwin MR, Hripcsak G, et al. 2020. Observational Study of Hydroxychloroquine in Hospitalized Patients with Covid-19. *New England Journal of Medicine*, **382**(25): 2411−2418.

Gu HJ, Xie ZD, Li TL, Zhang SG, Lai CC, Zhu P, et al. 2016. Angiotensin-converting enzyme 2 inhibits lung injury induced by respiratory syncytial virus. *Scientific Reports*, **6**(1): 19840.

Hamzelou J. 2020. Does a cell protein explain covid-19 severity?. *New Scientist*, **246**(3277): 9.

Holshue ML, Debolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. 2020. First Case of 2019 Novel Coronavirus in the United States. *New England Journal of Medicine*, **382**(10): 929−936.

Huang CL, Wang YM, Li XW, Ren LL, Zhao JP, Hu Y, et al. 2020a. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, **395**(10223): 497−506.

Huang MX, Li M, Xiao F, Pang PF, Liang JB, Tang TT, et al. 2020b. Preliminary evidence from a multicenter prospective observational study of the safety and efficacy of chloroquine for the treatment of COVID-19. *National Science Review*, **7**(9): 1428−1436.

Imai Y, Kuba K, Penninger JM. 2007. Angiotensin-converting enzyme 2 in acute respiratory distress syndrome. *Cellular and Molecular Life Sciences*, **64**(15): 2006−2012.

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, **536**(7616): 285−291.

Lewis PO, Kumar S, Tamura K, Nei M. 1995. MEGA: Molecular evolutionary genetics analysis, version 1.02. *Systematic Biology*, **44**(4): 576−577.

Li L, Zhang W, Hu Y, Tong XL, Zheng SG, Yang JT, et al. 2020a. Effect of convalescent plasma therapy on time to clinical improvement in patients with severe and life-threatening COVID-19: a randomized clinical trial. *JAMA*, **324**(5): 460−470.

Li QQ, Wu JJ, Nie JH, Zhang L, Hao H, Liu S, et al. 2020b. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell*, **182**(5): 1284−1294.

Li XC, Zhang J, Zhuo JL. 2017. The vasoprotective axes of the renin-angiotensin system: Physiological relevance and therapeutic implications in cardiovascular, hypertensive and kidney diseases. *Pharmacological Research*, **125**: 21−38.

Mirochnik Y, Aurora A, Schulze-Hoepfner FT, Deabes A, Shifrin V, Beckmann R, et al. 2009. Short pigment epithelial-derived factor-derived peptide inhibits angiogenesis and tumor growth. *Clinical Cancer Research*, **15**(5): 1655−1663.

Ordog R. 2008. PyDeT, a PyMOL plug-in for visualizing geometric concepts around proteins. *Bioinformation*, **2**(8): 346−347.

Rohl CA, Strauss CEM, Misura KMS, Baker D. 2004. Protein structure prediction using Rosetta. *Methods in Enzymology*, **383**: 66−93.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, **34**(12): 3299−3302.

Shen CG, Wang ZQ, Zhao F, Yang Y, Li JX, Yuan J, et al. 2020. Treatment of 5 critically Ill patients with COVID-19 with convalescent plasma. *JAMA*, **323**(16): 1582−1589.

Shi JZ, Wen ZY, Zhong GX, Yang HL, Wang C, Huang BY, et al. 2020a. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science*, **368**(6494): 1016−1020.

Shi R, Shan C, Duan XM, Chen ZH, Liu PP, Song JW, et al. 2020b. A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature*, **584**(7819): 120−124.

Shu CJ, Xiao K, Sun X. 2020. Structural basis for the high conductivity of microbial pili as potential nanowires. *Journal of Nanoscience and Nanotechnology*, **20**(1): 64−80.

Song WF, Gui M, Wang XQ, Xiang Y. 2018. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathogens*, **14**(8): e1007236.

Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, et al. 2021. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. *Nature*, **590**(7845): 290−299.

Tang XL, Wu CC, Li X, Song YH, Yao XM, Wu XK, et al. 2020. On the origin and continuing evolution of SARS-CoV-2. *National Science Review*, **7**(6): 1012−1023.

The 1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. *Nature*, **526**(7571): 68−74.

The UK10K Consortium. 2015. The UK10K project identifies rare variants in health and disease. *Nature*, **526**(7571): 82−90.

Tippmann HF. 2004. Analysis for free: comparing programs for sequence analysis. *Briefings in Bioinformatics*, **5**(1): 82−87.

Wang H, Zhang YT, Huang BY, Deng W, Quan YR, Wang WL, et al. 2020a. Development of an inactivated vaccine candidate, BBIBP-CorV, with potent protection against SARS-CoV-2. *Cell*, **182**(3): 713−721.

Wang ML, Cao RY, Zhang LK, Yang XL, Liu J, Xu MY, et al. 2020b. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Research*, **30**(3): 269−271.

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*, **46**(W1): W296−W303.

Wrapp D, Wang NS, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, **367**(6483): 1260−1263.

Zhao Y, Zhao ZX, Wang YJ, Zhou YQ, Ma Y, Zuo W. 2020. Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov. *bioRxiv*, doi: 10.1101/2020.01.26.919985.

Zheng YY, Ma YT, Zhang JY, Xie X. 2020. COVID-19 and the cardiovascular system. *Nature Reviews Cardiology*, **17**(5): 259−260.