



Original Article  
Evolutionary Genetics

# ACE2 diversity in placental mammals reveals the evolutionary strategy of SARS-CoV-2

Bibiana S.O. Fam<sup>1</sup>, Pedro Vargas-Pinilla<sup>2</sup>, Carlos Eduardo G. Amorim<sup>3</sup>, Vinicius A. Sortica<sup>1</sup>  
and Maria Cátira Bortolini<sup>1</sup>

<sup>1</sup>Universidade Federal do Rio Grande do Sul, Instituto de Biociências, Departamento de Genética, Porto Alegre, RS, Brazil.

<sup>2</sup>Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, Departamento de Bioquímica e Imunologia, Ribeirão Preto, SP, Brazil.

<sup>3</sup>University of Lausanne, Department of Computational Biology, Lausanne, Switzerland.

## Abstract

The recent emergence of SARS-CoV-2 is responsible for the current pandemic of COVID-19, which uses the human membrane protein ACE2 as a gateway to host-cell infection. We performed a comparative genomic analysis of 70 ACE2 placental mammal orthologues to identify variations and contribute to the understanding of evolutionary dynamics behind this successful adaptation to infect humans. Our results reveal that 4% of the ACE2 sites are under positive selection, all located in the catalytic domain, suggesting possibly taxon-specific adaptations related to the ACE2 function, such as cardiovascular physiology. Considering all variable sites, we selected 30 of them located at the critical ACE2 binding sites to the SARS-CoV-like viruses for analysis in more detail. Our results reveal a relatively high diversity of ACE2 between placental mammal species, while showing no polymorphism within human populations, at least considering the 30 inter-species variable sites. A perfect scenario for natural selection favored this opportunistic new coronavirus in its trajectory of infecting humans. We suggest that SARS-CoV-2 became a specialist coronavirus for human hosts. Differences in the rate of infection and mortality could be related to the innate immune responses, other unknown genetic factors, as well as non-biological factors.

**Keywords:** ACE2, placental mammals, SARS-CoV-2, COVID-19, inter and intra-species diversity.

Received: April 06, 2020; Accepted: April 22, 2020.

## Introduction

The novel coronavirus SARS-CoV-2 (denominated before as 2019-nCoV) is a single-stranded RNA virus member of the Coronaviridae family. The estimates of the most recent common ancestor (MRCA) for the four coronavirus genera of this family range from 10,000 to millions of years (Wertheim *et al.*, 2013). The most commented member of the Coronaviridae family at the moment, SARS-CoV-2, has a very recent origin. Based on genome sequence from different SARS-CoV-2 strains and a yearly mutation rate of  $1.24 \times 10^{-3}$  per site, Li *et al.* (2020a) estimated that it originated on November 24 2019.

SARS-CoV-2 is responsible for the current pandemic of Corona Virus Disease 2019 (COVID-19) (Chan *et al.*, 2020; Coronaviridae Study Group, 2020; Huang *et al.*, 2020; Wang *et al.*, 2020). The COVID-19 was first reported in Wuhan (China) in December 2019 and spread worldwide,

with a high rate of transmission, thousands of infected patients and many deaths between December 2019 and April 2020, and these numbers continue to grow. On March 11, World Health Organization (WHO) characterized COVID-19 as a pandemic, the first one caused by a coronavirus (WHO, situation report, March 11, 2020). SARS-CoV-2 causes a severe respiratory syndrome in humans, being transmitted arguably through different routes (*i.e.*, fomites, airborne or fecal-oral) by an animal to human and human to humans (Chan *et al.*, 2020; Huang *et al.*, 2020; Li R *et al.*, 2020), similar to SARS-CoV (Severe Acute Respiratory Syndrome-related coronavirus) and MERS-CoV (Middle East Respiratory Syndrome-related coronavirus). SARS-CoV was initially transmitted to humans by the palm civet (*Paguma larvata*; Kan *et al.*, 2005) and MERS-CoV by the dromedary (*Camelus dromedarius*; Zaki *et al.*, 2012). SARS-CoV-2 is classified as *Beta coronavirus* subgenus *Sarbecovirus* with ~70% of similarity with SARS-CoV (Wu *et al.*, 2020). Genetic studies from different research groups found that SARS-CoV-2 is highly similar to the bat (*Rhinolophus affinis*) coronavirus (Guo *et al.*, 2020; Ji *et al.*, 2020; Li R *et al.*, 2020; Wu *et al.*, 2020; Zhou *et al.*, 2020). Other

Send correspondence to Maria Cátira Bortolini. Universidade Federal do Rio Grande do Sul, Instituto de Biociências, Departamento de Genética. Av. Bento Gonçalves, 9500. Prédio 43323, CEP: 91501-970, Porto Alegre, RS, Brazil. E-mail: maria.bortolini@ufrgs.br.

studies comparing genomes of different coronaviruses showed local genomic similarities between the human coronavirus and those from snakes, turtles, pangolins, and minks, suggesting these animals as possible intermediate-hosts (Guo *et al.*, 2020; Ji *et al.*, 2020; Li X *et al.*, 2020; Liu *et al.*, 2020; Wan *et al.*, 2020; Wu *et al.*, 2020). The most recent metagenomic study on the topic reinforces the idea that pangolins (*Manis javanica*) should be considered as the most plausible intermediate-host of SARS-CoV-2 related viruses, and that they acquired these viruses independently from bats or another animal hosts (Lam *et al.*, 2020). The authors found putative recombination signals between the pangolins coronaviruses, bat coronaviruses, and human SARS-CoV-2 (Lam *et al.*, 2020), which highlights the complexity of the origin of the new human coronavirus. Although there are uncertainties about the original and intermediate hosts, mutation and/or recombination among coronaviruses may have enabled cross-species infection (Letko *et al.*, 2020; Wan *et al.*, 2020).

Similar to SARS-CoV, SARS-CoV-2 utilizes the Angiotensin-Converting Enzyme 2 (ACE2) as target site to infect cells (Li *et al.*, 2003; Li *et al.*, 2005a,b; Chen *et al.*, 2020; Hoffmann *et al.*, 2020; Letko *et al.*, 2020; Li F *et al.*, 2020; Li R *et al.*, 2020; Luan *et al.*, 2020; Walls *et al.*, 2020; Wan *et al.*, 2020; Xu *et al.*, 2020). *ACE2* gene is located on chromosome X, and its product ACE2, a zinc metallopeptidase protein, comprises 805 amino acids containing a single catalytic domain. ACE2 is mainly expressed at vascular endothelium, myocardium, lungs, kidneys, and intestines (Turner *et al.*, 2004; McKinney *et al.*, 2014; Zhang *et al.*, 2020). It is a strict carboxypeptidase that hydrolyzes its substrate removing a single amino acid from their respective C-terminal. It is responsible for cleaving Angiotensin I and II into peptides Angiotensin 1-9 and Angiotensin 1-7, respectively, both are key elements connected with cardiovascular physiology, regulation of vascular tonus, blood pressure, electrolyte balance, and water intake (Millan *et al.*, 1990; Donoghue *et al.*, 2000; Oudit *et al.*, 2009; McKinney *et al.*, 2014).

ACE2 is a cell-surface non-raft protein with little intracellular localization. When binding to coronaviruses, the protein internalizes down-regulating activity from the cell surface (Hamming *et al.*, 2007). After SARS-CoV-like viruses and ACE2 attach, molecular cascade signal events result in entry of the coronavirus into the host cell (Simmons *et al.*, 2013), showing a successful evolutionary infection strategy. Kuba *et al.* (2005) showed that ACE2 down regulation mediated by SARS-CoV contributes to the acute lung injury, the most important clinical implication of human respiratory diseases caused by a coronavirus.

The success of a zoonotic spillover (transmission of a pathogen from a vertebrate animal to a human) is a complex process. It depends on several cultural, ecological, and climatic conditions, as well as features of viruses and hosts (Olival *et al.*, 2017). The efficiency in recognizing the host binding receptor is a crucial step, which defines the preference of the virus for a given species, tissue or cell type. The interaction of SARS-CoV-2 spike (S) glycoprotein, through

its receptor-binding domain (RBD), happens via an optimized binding to human ACE2 domains (Andersen *et al.*, 2020; Hoffman *et al.*, 2020; Walls *et al.*, 2020). According to recent studies, five amino acid changes associated with natural selection in the critical binding sites (L455, F486, Q493, S494, N501) located at the SARS-CoV-2 S glycoprotein would be responsible for this high tropism with human ACE2 (Andersen *et al.*, 2020; Wan *et al.*, 2020). Furthermore, S glycoprotein of SARS-CoV-2 contains a cleavage site for furin proteases at the junction of subunits S1 and S2 (Coutard *et al.*, 2020). During viral infection, S glycoprotein needs to be cleaved by host-cell proteases into S1 (which contains RBD) and S2 subunits. After this cleavage, the exposition of the S2 mediates fusion of the viral and host-cell membranes (Yan *et al.*, 2020). Furin proteases are abundant in the respiratory tract, and this characteristic of the SARS-CoV-2 has been considered to be essential for successful infection of SARS-CoV-2 in human cells (Lam *et al.*, 2020). In contrast, the highly related bat and pangolin SARS-CoV-2-like coronaviruses do not have the furin cleavage site, which in principle could be a barrier to zoonotic coronavirus spillover. Moreover, human-infecting coronaviruses such as HCoV-OC43, MERS-CoV, and HKU1 also have been demonstrated to be cleaved at an S1/S2 cleavage site by furin proteases (Coutard *et al.*, 2020). However, in SARS-CoV, furin-mediated cleavage of the S glycoprotein appears not to occur naturally (Simmons *et al.*, 2004) and the introduction of a functional furin cleavage site in the S1/S2 junction of SARS-CoV S glycoprotein resulted in a dramatic enhancement of cell-cell fusion (Follis and Nunberg, 2004).

Changes in key amino acids (AA) related to an efficient interaction between SARS-CoV-like S glycoprotein and ACE2 are crucial to cross-species infection, multi-host infections, as well as differences in susceptibility to disease and its symptoms in animal species. The evolutionary phenomenon can be evoked to explain at least part of these events and conditions. To contribute to a better understanding of this complex process, we performed a comparative genomic analysis of 70 ACE2 orthologues, representative of placental mammal species, including pangolin, civet, bat, mink, and five New World monkey (NWM) species (Parvorder Platyrrhini). Furthermore, we describe the genetic variation of key binding sites across human populations. Our immediate goal is to reveal the taxon-specific variability in ACE2 amino acid sequence, its possible influence in cross-species SARS-CoV-2 infection, potential hosts, and other related topics that can be important and very timely considering the world outbreak of COVID-19.

## Material and Methods

### Sequences

Coding sequences for *ACE2* gene for 70 placental mammal species were retrieved through BLAST from GenBank (www.ncbi.nlm.nih.gov accessed at 20/02/2020). For these analyses we considered 16 domestic and 54 wild species including representatives of orders Carnivora (N =

9), Ungulata (N = 11), Chiroptera (N = 10), Primates (N = 22), Rodentia (N = 9), and others (Table S1). Sequences were aligned according to the MUSCLE algorithm (Edgar, 2004) in AliView software (Larsson, 2014).

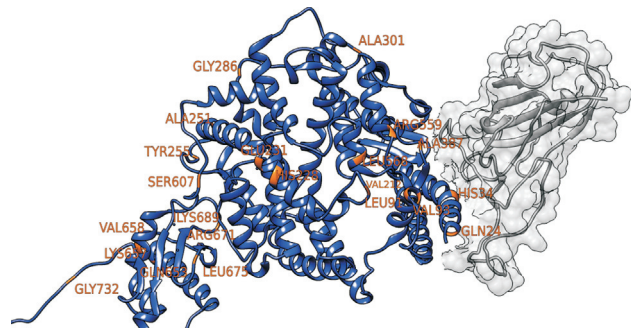
To investigate the variation within *Homo sapiens*, we used the Ensembl database (www.ensembl.org; Hunt. *et al.*, 2018) and the UNIPROT database (www.uniprot.org, The Uniprot Consortium, 2019). Both platforms used data from the 1000 Genome Project.

## Evolutionary Analysis

The consensus phylogeny of species, based on neutral molecular markers, was obtained from the Timetree server (www.timetree.org; Kumar *et al.*, 2017). This phylogeny is used as a reference by specialized literature following Meredith *et al.* (2011). To better understand evolutionary patterns that are acting on *ACE2*, we performed a Site Model (NsSite) test with Codeml package in PAML 4.9 software (Yang, 2007), which allows for inter-specific phylogenetic comparison of substitution rates in protein-coding genes. This test allows for  $\omega$  variation ( $\omega = dN/dS$ , being dN the rate of non-synonymous substitutions and dS the rate of synonymous substitutions) across sites. It fits neutral ( $\omega \approx 1$ ), selective constraint ( $\omega < 1$ ), or positive selection ( $\omega > 1$ ) models to the observed levels of variation. To determine which evolutionary model best fits the analyzed data, we performed Likelihood Ratio Tests (Yang, 1998). For these, we compared the neutral model (M1) against the model that allows positive selection (M2a) with  $df = 2$  (Yang, 2007). Bayes Empirical Bayes (BEB) was used to infer sites with a high probability of being under positive selection (Yang *et al.*, 2005). Furthermore, we classify amino acid variation found in placental mammals *ACE2* coding sequences into four classes of chemical similarity accordingly Grantham score (Grantham, 1974): conservative (GS =  $\leq 50$ ), moderately conservative (GS = 51–100), moderately radical (GS = 101–150), and radical ( $\geq 151$ ).

## Results

Our results reveal that the M2 model (NsSite;  $\omega > 1$ ) fits the placental mammal *ACE2* data significantly better than the neutral model ( $\omega_2 = 2.58680$ ,  $p < 0.001$ ; Table S2). Considering the M2 model, 4% of the *ACE2* sites are under positive selection, a value high enough to infer a non-neutral molecular evolution for this gene as whole, at least considering placental mammals. Note however, that this evolutionary scenario is not homogeneous across the different domains of the *ACE2* protein. A parallel analysis, Bayes Empirical Bayes (BEB), highlights 22 sites (24, 34, 91, 93, 212, 228, 231, 251, 255, 286, 301, 387, 559, 568, 607, 653, 657, 658, 671, 675, 689, and 732) (Table S3; Figure 1) with a high probability ( $> 95\%$ ) of being under positive selection, all of them being in the active sites at the extracellular portion of the protein (residues 18-740), where the *ACE2* catalytic domain is located. Of the remaining *ACE2* sites, 69% are under functional constraint due to the negative selection



**Figure 1** - Structural analysis of sites under positive selection in placental mammals (orange) of human *ACE2* (blue) binding to SARS-CoV2 spike glycoprotein (gray). The Protein Data Bank (PDB) ID is 6M17 (Yan *et al.*, 2020).

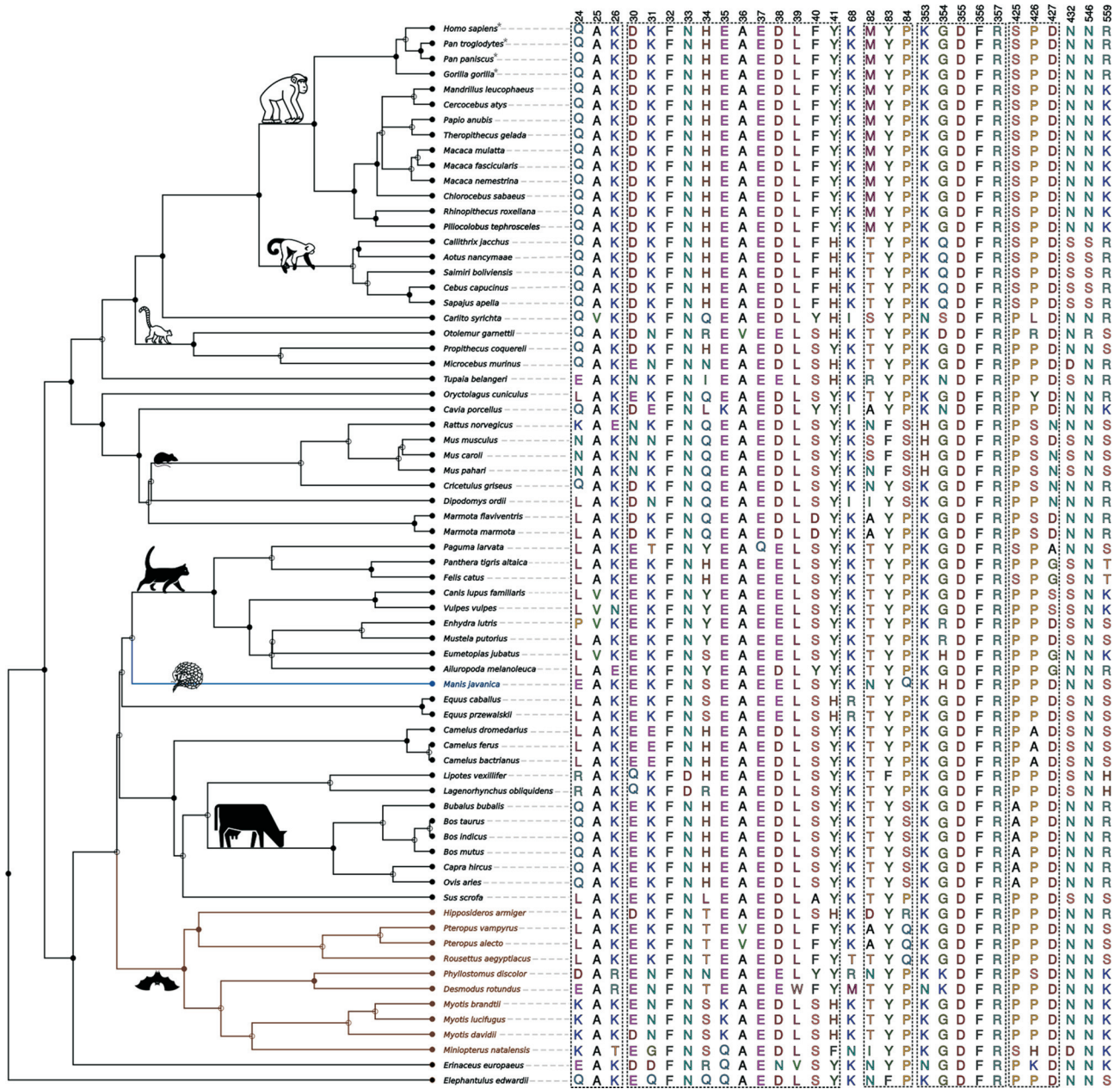
( $\omega_0 = 0.10125$ ), while 27% are with selective relaxation ( $\omega_1 = 1$ ). For instance, the signal peptide (1-17), transmembrane domain (741 -761), and cytoplasmatic domain (761-805; see  $\omega_3$  column of the Table S4) are more conserved, showing little variation across species, none of them with positive selection signals (Tables S2 and S4). We note that *ACE2* is a critical element in the cascade of regulation of blood pressure and cardiovascular condition, among other functions, and that these AA changes with a significant signal of positive selection could be taxon-specific, related to particular adaptive phenotypes found in some of these diverse placental mammals.

After that first evolutionary analysis, we examined all variable sites found considering the 70 orthologues. However, we choose only the 30 variable sites with information that they are functional and/or involved in the interaction between host-*ACE2* and SARS-CoV-like viruses (Towler *et al.*, 2004; Li F *et al.*, 2005; Li R *et al.*, 2020; Luan *et al.*, 2020; Wan *et al.*, 2020). Figure 2 shows the AAs observed in 30 chosen variable sites across the *ACE2* orthologues of 70 selected placental mammal species. Three of them (24, 34, and 559) are target sites for positive selection, according to BEB (Table S3). The topology of Figure 2 recovers the phylogenetic relationship between the major taxonomic groups investigated. For example, *Homo sapiens* and all other great apes (family Hominoidea) share the same amino acids in the 30 positions, and only one divergence from Cercopithecidae species at the positive selection site 559 (conservative AA change, R559K; GS = 26) (Figure 2).

Noteworthy, our purpose here is not to report each variation observed in the 70 orthologues (Figure 2), but to point out some thought-provoking findings to support the discussion and the main conclusion of the present study. For this, we built Figure S1, a simplified version of the data from Figure 2, where only sites and species commented in the next paragraphs are shown.

Position 24, which is one of the three sites with high probability of being under positive selection, has eight different amino acids segregating across taxon (Figures 2 and S1). Li W *et al.* (2005) showed that the amino acid change from Glutamine to Lysine at position 24 (Q24K), together





**Figure 2** - Crucial variable sites in placental mammalian ACE2 protein in the interaction with SARS-CoV-2. In color branches potential sources of SARS-CoV-2; in blue, Pangolin (*Manis javanica*) branch, and in brown bat (Chiroptera) branches.\* Species whose susceptibility to infection by SARS-CoV-2 may be similar to human. The full nodes in the tree are those in which we have a taxonomic classification defined as order, family, gender, etc., while those with an empty node, this information is not well supported. Animal icons available in Noun project (<https://thenounproject.com/>).

with a change from Lysine to Glutamate at position 26 (K26E), slightly inhibits interactions with SARS-CoV S glycoprotein (Li W *et al.*, 2005). *Rattus norvegicus* has this combination (24K and 26E), while mouse species (*Mus musculus*, *Mus caroli*, and *Mus pahari*) have an Asparagine (N) at the position 24. *Rattus norvegicus* ACE2 also has the combination of Asparagine (N), Phenylalanine (F), and Serine (S) at positions 82-84. According to Li W *et al.* (2005), this combination in rats (*R. norvegicus*) inhibits the interaction between ACE2 and SARS-CoV S glycoprotein. In addition, a Histidine (H) at positions 353 of ACE2 in the murine

species also could inhibit or reduce the efficiency of the interaction according to Wan *et al.* (2020). Our study shows that murine species' ACE2 have only 56% of identity with the human ACE2 regarding these 30 binding sites. This reinforces the idea that they cannot naturally bind to S glycoprotein of SARS-CoV-2. On the other hand, Glutamine (Q) and Lysine (K) at 24 and 26 ACE2 positions are found in humans and all primates, as well as in *Cavia porcellus* (Guinea-pig), a rodent used very often as an experimental animal. However, there are differences in other important binding sites, resulting in the identity between Guinea-pig

ACE2 and the human ACE2, regarding these 30 binding sites, being just 70%.

Cat (*Felis catus*) and dog (*Canis lupus familiaris*) are of special interest due to their close relationship with humans. Their ACE2 orthologues present at 77% and 73% of identity with the human ACE2 respectively, considering the 30 variable sites. Both species present identity with other species belonging to the Carnivora order, such as Leucine (L) at position 24. However, they also show differences from each other at positions 25, 34, 427, and 559, two of them are a target for positive selection (34 and 559; Table S3). In other words, 50% of the differences between the ACE2 binding sites of cat and dog are taxon-specific variants that may be connected with specific functions of the ACE2 protein in each of these two domesticated species.

Interestingly, the five NWm and Prosimian species show different patterns than other primates at position 82-84. They present a moderately conserved change at position 82 (M82T; GS = 81), which also represents a change in polarity. We note that, due to the great number of species that present the amino acid Threonine in that position, according to the principle of parsimony, Threonine is probably the ancestral allele, meaning that a non-synonymous mutation promoted the AA alteration Threonine (T) > Methionine (M) in the ancestral branches of the Catarrhini clade. However, the ancestor of the Platyrrhini maintained the ancestral allele T. At position 41, we also found the same amino acid in NWm and Prosimian species. The moderately conservative modification Y41H (GS = 83) represents changes in charge, once Histidine (H) is an acid AA. Another mutation at this same site (Y41A; GS = 112) implies a moderately radical change and polarity difference, since Tyrosine (Y) is a polar AA, while Alanine (A) is a non-polar AA. This type of alteration could potentially prevent viral interaction, at least with SARS-CoV according Li *et al.* (2005b). Based on the distribution of the AAs at position 41 of ACE2, it seems that both NWm and Prosimian clades shared the derived allele (H), also seen in the bat genera *Hipposideros* and *Myotis*, as well as in *Equus*, suggesting molecular convergence.

NWm ACE2 is also different from other Primates species due to two changes in the potential glycosylation sites at positions 432 and 546 (Luan *et al.*, 2020), where Arginine (N) was replaced by Serine (S). The glycosylation was suggested by the presence of electron density at all Arginine-linked site of the ACE2 N-terminal domain, including 432 and 546 and the loss of glycosylation sites may have functional implications (Towler *et al.*, 2004). Studies with MERS-CoV and with its host-cell entry receptor, DPP4, indicated that mutations that knock out glycosylation sites present in mouse DPP4 blocked the binding with MERS-CoV RBD (Peck *et al.*, 2017). The last authors concluded that glycosylation in DPP4 orthologue sites are a substantial barrier to MERS-CoV infection, particularly when combined with taxon-specific changes (Peck *et al.*, 2017).

We also observed a Serine (S) at the positively selected site 559, which reduces or inhibits the interaction of SARS-CoV S glycoprotein with ACE2 (Li W *et al.*, 2005). A Serine

in that position is found not only in rodents but also in some species of carnivores that were already considered as possible original or intermediate hosts, such as the pangolin (*Manis javanica*). Pangolin ACE2 has ~85% of similarity with human ACE2 (Lam *et al.*, 2020), but only 66% considering just these 30 sites investigated here. In addition, a Serine at position 559 is also observed in the two species of the Mustelidae family, ferret (*Mustela putorius*) and sea otter (*Enhydra lutris*).

Finally, we also leveraged intra-specific variation of the ACE2 gene using data from human populations available in the 1000 Genomes Project (accessed through Ensembl and UNIPROT platforms). This dataset revealed 660 missense variants (all with minor allele frequency of 0.005 or less), 260 synonymous variants, 14,352 intron variants, 67 5'UTR variants, and 268 3'UTR variants in the ACE2 gene in humans. Many of these have been investigated in case-control or genome-wide association studies for cardiovascular diseases and other correlated conditions (see for instance Ji *et al.*, 2017). Regarding the 30 ACE2 binding sites selected to present study (Figure 2), only one of them, at position 26, has two alleles segregating (K and R; Minor Allele Frequency or MAF = 0.002; rs4646116). Although we do not have population data for the other species investigated here, we observed that there is no relevant polymorphism regarding the 30 sites that promote the interaction between host-ACE2 and SARS-CoV-like viruses in the most studied species of all (*Homo sapiens*). This finding reinforces the idea that the AAs presented in Figure 2 are taxon-specific, *i.e.*, characteristic of a species or a taxonomic group, where intra-specific variation is absent or negligible. Also, the AA conservation within *Homo sapiens* in these 30 sites indicates an instigating evolutionary constraint whose reasons need to be further studied. A search in the genome of the Neanderthal (*Homo neanderthalensis*) and in the specimen of Denisova indicates that both present the same AAs as *Homo sapiens* in these 30 sites.

## Discussion

Species of the Coronaviridae family can infect a wide variety of animals and humans, causing severe respiratory, enteric, hepatic, and neurological diseases (Weiss and Leibowitz, 2011). In humans, coronavirus has caused mainly respiratory tract infections and, for decades, were considered with little clinical and epidemiological importance until SARS-CoV and MERS-CoV outbreaks (Kuiken *et al.*, 2003; Zaki *et al.*, 2012). Today, SARS-Cov-2 is responsible for the COVID-19, a pandemic that has challenged governments and peoples of every country in the world. COVID-19 already killed more persons than diseases caused by SARS-CoV and MERS-CoV combined, emphasizing the importance of this pathology as a highly relevant public health concern (WHO; <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>).

SARS-CoV-like viruses use ACE2 as their gateway to the host cell. Figure 2 illustrates the diversity of 70 ACE2 orthologues, considering 30 potential binding sites to the S



glycoproteins of the SARS-CoV-like viruses. This fact opens up a wide range of possibilities for this type of coronavirus, whose rate of evolution is high, due to its high rates of mutation and recombination (Li *et al.*, 2020a). These phenomena that generate variability in the RNA viruses explain why they change host more frequently and jump between distantly related species more often than other pathogens (Woolhouse *et al.*, 2005a,b; Davies and Pedersen, 2008; Longdon *et al.*, 2014). These facts would explain a predicted origin of SARS-CoV-2 from original and/or intermediate phylogenetically distant hosts (for instance, bat and pangolin; Guo *et al.*, 2020; Ji *et al.*, 2020; Lam *et al.*, 2020; Li X *et al.*, 2020; Wu *et al.*, 2020; Zhou *et al.*, 2020).

It was also known that specific pathogen mutations are often required to enhance pathogen's fitness in a new host (Longdon *et al.*, 2014). This is exactly what happened for the emergence of SARS-CoV-2. Recent investigations suggest the jump across species that resulted in the emergence of SARS-CoV-2 in humans was mainly due to five amino acid changes in critical S glycoprotein binding sites (Wan *et al.*, 2020; Andersen *et al.*, 2020). In addition to that, the S glycoprotein of SARS-CoV-2 contains a cleavage site for furin proteases at the junction of subunits S1 and S2 as other human coronavirus (HCoV-OC43, MERS-CoV, and HKU1), but differently than what occurs with bat and pangolin SARS-CoV-2-like coronavirus, and SARS-CoV (Coutard *et al.*, 2020). These events are associated with the high and efficient tropism with human ACE2 (Andersen *et al.*, 2020; Lam *et al.*, 2020).

As opposed to the key differences observed across the 70 orthologues in mammals, no relevant polymorphism was identified in these 30 ACE2 binding sites considering human populations. In that sense, *Homo sapiens* is a "homogeneous" primate species, at least in regards to the 30 ACE2 important binding sites. In addition to that, *Homo sapiens* has a large population (census size of ~7.5 billion individuals) currently living in virtually all habitable places on Earth, including regions with very high demographic density. Based on these observations, we can suggest that the high and efficient tropism is extended to all human populations, in addition to demographic and cultural conditions linked to particular human populations that may facilitate contagion and dispersion of this opportunistic coronavirus.

Hosts and pathogens are always in an evolutionary arms race. Several investigations reveal that coronavirus evolved to recognize different host-receptors, while others explain the host's strategies to defend themselves. For example, Peng *et al.* (2017) showed that a hepatitis coronavirus (MHV) infection drives the evolution of a mouse receptor (mCEACAM1a). In this case, this event represented an answer to the intense selective pressure from a lethal infection where MHV uses mCEACAM1a as its host-receptor and changes in critical amino acid residues have led to the emergence of a derived-allele (mCEACAM1b) that is much more deficient as MHV receptor (Peng *et al.*, 2017). Consequently, mice that are homozygous for this derived-allele are

highly resistant to death from MHV infections (Peng *et al.*, 2017).

The first scientific report on the use of ACE2 as a host-cell entry portal by a coronavirus involved SARS-CoV (Li *et al.*, 2003). If this happened before (some coronavirus use the ACE2 protein as a gateway to the host-cell), either with humans or other species, we currently do not know. Therefore, and considering only what we know about the history of SARS-CoV, it is a short timeframe for us to expect any evolutionary "reaction" from the host with a relatively long life cycle, at least regarding human ACE2. In other words, we do not expect to see for the human ACE2 (in relation to both SARS-CoV and SARS-CoV-2) what was reported in the study of Peng *et al.* (2017) for the mouse receptor and MHV. As a consequence of the very recent evolution of SARS-CoV-2 Spike glycoprotein and gain of a cleavage site for furin proteases, in addition to the homogeneity of critical ACE2 binding sites across human populations and other demographic and cultural factors already mentioned above, SARS-CoV-2 obtained a significant edge in the evolutionary arms race, and become a highly infectious and pathogenic coronavirus for humans.

However, it is very well known that subsequent to cell entry, pathogens can be blocked by neutralizing host-antibodies, which appears to be occurring in SARS-CoV-2 infections in humans (Conti *et al.*, 2020). According to the WHO the majority of humans infected by SARS-CoV-2 have mild symptoms or are asymptomatic, despite the fact that they have the same AAs in the ACE2 binding sites as reported here. The mechanisms of how antibodies neutralize a coronavirus infection have already been published (Hoffmann *et al.*, 2004; Reguera *et al.*, 2012; Li G *et al.*, 2020), illustrating that an innate agile and efficient host-immune system (like those in young individuals and/or in individuals who do not develop an immunological over-response) is the first weapon to block the infection of opportunistic coronaviruses. Eventual population variability in other genetic systems, which may also be involved in this heterogeneity of COVID-19 symptoms and outcomes, cannot be ruled out. We must also remember that the ACE2 gene has many polymorphisms outside the regions coding for key binding sites, as mentioned in the last paragraph of the results section. In this way, we cannot rule out that some of these SNPs may also be involved in differential susceptibility for infection, as well as to the differences in the symptoms and outcomes observed across human populations.

On the other hand, the differences in these 30 important ACE2 binding sites of interaction with SARS-CoV-like viruses observed between species and taxonomic groups is striking, probably impacting SARS-Cov-2's ability to bind with ACE2 in some of these species. SARS-CoV-2 capacity to infect other 42 mammalian species was recently predicted based on changes at positions 31, 35, 38, 82, and 353 in the ACE2 orthologues when compared with the human ACE2 (Luan *et al.*, 2020). These authors suggested that 76% of the 42 mammal species investigated, including pets (dog, *Canis lupus familiaris*, and cat, *Felis catus*), horse (*Equus caballus*),

lus), cattle (*Bos taurus*), and sheep (*Ovis aries*), may bind to S protein of SARS-CoV-2. The idea of a “generalist” coronavirus, capable of infecting other species of mammals (pig, ferret, cat, monkey) had also been argued by other authors (Wan *et al.*, 2020), despite of the differences in their correspondent ACE2 orthologues.

Here, we have expanded the comparative analysis to more species and sites of interaction between ACE2 orthologues and SARS-CoV-like viruses. In sum, our results show that ACE2 murine species have only 56% of identity with human ACE2, and have also several critical taxon-specific changes that explain the barrier to infection with SARS-CoV-2. For the other species and taxon investigated here, however, the ability to block (or not) the infection to SARS-CoV-2 via this receptor is not easy to predict. At the moment, preliminary results indicate that a macaque (*Macaca mulatta*) inoculated with SARS-CoV-2 recapitulates moderate respiratory disease observed in the majority of human cases (Munster *et al.*, 2020). Note that *Macaca mulatta*, similar to other species of the family Cercopithecidae, has only one change (R559K) relative to human and other species of the family Hominidae (*Pan troglodytes*, *Pan paniscus* and *Gorilla gorilla*; Figure 2). We also found a set of AA changes in the ACE2 of NWM species, which can characterize different patterns of binding to SARS-CoV-2, making them potentially less susceptible to the infection than humans and great apes. Earlier, one *in vivo* study (Greenough *et al.*, 2005) demonstrated SARS-CoV infection outcome in *Callithrix sp.*, however, the pattern of virus replication and the potential risk for infection in nature for this and other NWM species has never been demonstrated. Therefore, *Callithrix sp.*, although being a good model for MERS-CoV infection (McAuliffe *et al.*, 2004; Van Doremalen and Munster, 2015), it has not been established as an animal model to SARS-CoV infection. Note that MERS-CoV uses another host-cell entry receptor, DPP4, rather than ACE2.

When more *in vitro* and *in vivo* functional studies are performed, we will know what is the minimum AA combination in the ACE2 binding sites and in other critical regions of the receptor that is indispensable for an active connection with SARS-CoV-2 S glycoprotein. Until then, the predictions are speculative, but they can bring subsidies for future studies and strategies to contain the infection and its harmful consequences. Here, we suggest that there are not many ACE2 orthologues with tropism for SARS-CoV-2, precisely because of its ability, shaped by natural selection, to use *Homo sapiens* ACE2 as an efficient and successful gateway to the host-cell. To our knowledge, there is only one preliminary *in vivo* study to test the susceptibility of animals to SARS-CoV-2 (Shi *et al.*, 2020). These authors intra-nasally inoculated ferrets, cats, dogs, pigs, chickens, and ducks with a dosage of SARS-CoV-2, and euthanized them some days post inoculation. They reported that ferrets and cats are highly susceptible to SARS-CoV-2, dogs have low susceptibility, and livestock including pigs, chickens, and ducks are not susceptible to the SARS-CoV-2. On the other hand, there are no scientific record up to date of natural infection by

SARS-CoV-2 in animals living in anthropogenic environments (only rare press reports that described that cats, as well as tigers and lions from the Bronx Zoo in New York City were infected), an observation that challenges previous predictions that this coronavirus would infect a wide range of species (Luan *et al.*, 2020; Wan *et al.*, 2020). The absence of a veterinary epidemiological emergency during the current human COVID-19 pandemic, at least considering pets, also challenges preliminary results, which indicated a high susceptibility of cats to SARS-CoV-2 (Shi *et al.*, 2020). If there is such a susceptibility to SARS-CoV-2 infection to these and other animals, it appears that it is much more restricted than that observed for humans.

It is also known that viruses that take advantage of receptors with conserved AA sequence have a broad host range (Woolhouse *et al.*, 2012). It is worth recalling that ACE2 has conserved the AA sequence in crucial binding sites within *Homo sapiens*, but not between placental mammal species investigated here. In other words, the high mutation rate SARS-CoV-2 may provide alternatives to infect one of several species, but when a high tropism with the host cell is found, it can be predicted that the binding region has little (if any) variation within the target species.

Our findings reveal also that an essential part of this taxon-specific ACE2 diversity can even be attributed to the positive natural selection, since it is a membrane protein related to the water intake, blood pressure control, and other cardiovascular conditions. Besides that, a common trade-off of adapting to novel hosts is that the performance on the original host is reduced (Longdon *et al.*, 2014). In other words, SARS-CoV-2 is a coronavirus with high performance to infect *Homo sapiens* individuals, from any population. In contrast, it could not infect, at least with the same efficiency, other species, as well as jump and re-infect the original or intermediate phylogenetically distant hosts (bats and/or pangolins). New mutations and/or recombination can turn SARS-CoV-2 into a derived-coronavirus capable of breaking down species barriers again, feeding the endless cycle of the arms race between host and pathogen.

On the other hand, closely related species may have similar levels of susceptibility, regardless of their distance from the pathogen’s natural host (Longdon *et al.*, 2014). Thus, due to the homology that ACE2 orthologues of the great apes have with the human ACE2 regarding the 30 binding sites investigated here, perhaps, only they are naturally and equally susceptible to develop symptoms of COVID-19, a concern that has already been demonstrated by members of the “Great Ape Health Consortium” (Gillespie and Leendertz, 2020).

## Conclusion

Our comparative analysis of the ACE2 coding region from 70 placental mammal species, revealed many variable sites. Of these, we selected 30 sites with records to be relevant for interaction with SARS-CoV-like S glycoproteins. For three of them, there is a significant signal of the action of the Darwinian (positive) selection. We observed extraordi-

nary conservation within *Homo sapiens* and relatively high variation among placental mammal species. Our data indicate that SARS-CoV-2 has similar potential to infect any human population, already corroborated by the pandemic character of COVID-19. We can suggest that any difference in the rate of infection and mortality of this pathogen/disease among humans is related to other factors such as comorbidities of the infected individuals, cultural practices, demographic density, capacity of the country's health systems, suppression or mitigation measures, in addition to differences in the SARS-CoV-2 strains, as well as the effective innate immune response from infected individuals. The involvement of other genetic pathways, in addition to the ACE2 protein and the immune system, cannot be ruled out either, but only the progress of the studies will be able to identify them, as well as their eventual role in the COVID-19 human pandemic.

In conclusion, SARS-CoV-2 is highly specialized for infecting humans of all populations, *i.e.* it is a “human-population generalist coronavirus”, at least considering its gateway to the cell, the ACE2. In contrast, it is not a “generalist” coronavirus with the ability to infect, naturally and easily, a range of other species, including pets.

## Acknowledgments

This research was financially supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES - Brasil, Finance code 001, the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Conflict of Interest

The authors declare that there is no conflict of interest that could be understood as prejudicial to the present research.

## Author Contributions

BSOF contribute to conceptualization, formal analysis, investigation, and writing-original draft. PVP contributes to conceptualization, formal analysis, investigation, and writing-original draft. CEGA contributes to investigation, and writing-review & editing. VAS contribute to conceptualization, investigation, writing-original draft, and writing-review & editing. MCB contributed to supervision (Lead), conceptualization, investigation, writing-original draft, and writing-review & editing.

## References

Andersen KG, Rambaut A, Lipkin WI, Holmes EC and Garry RF (2020) The proximal origin of SARS-CoV-2. *Nat Med* 26:450-452.

Chan JFW, Yuan S, Kok KH, To KKW, Chu H, Yang J, Xing F, Liu J, Yip CCY, Poon RWS *et al.* (2020) A familial cluster of

pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 395:514-523.

Chen Y, Guo Y, Pan Y and Zhao ZJ (2020) Structure analysis of the receptor binding of 2019-nCoV. *Biochem Biophys Res Commun* 525:135-140.

Conti P, Gallenga CE, Tetè G, Caraffa A, Ronconi G, Younes A, Toniato E, Ross R and Kritas SK (2020) How to reduce the likelihood of coronavirus-19 (CoV-19 or SARS-CoV-2) infection and lung inflammation mediated by IL-1. *J Biol Regul Homeost Agents* (in press).

Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020) The species *Severe acute respiratory syndrome-related coronavirus*: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* 5:536-544.

Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG and Decroly E (2020) The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res* 176:104742.

Davies TJ and Pedersen AB (2008) Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proc Biol Sci* 275:1695-1701.

Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, Donovan M, Woolf B, Robison K, Jeyaseelan R *et al.* (2000) A novel angiotensin-converting enzyme-related Carboxypeptidase (ACE2) converts angiotensin I to angiotensin1-9. *Circ Res* 87:E1-9.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792-1797.

Follis KE, York J and Nunberg JH (2006) Furin cleavage of the SARS coronavirus spike glycoprotein enhances cell-cell fusion but does not affect virion entry. *Virology* 350:358-369.

Gillespie TR and Leendertz FH (2020) COVID-19: protect great apes during human pandemics. *Nature* 579:497-497.

Grantham R (1974) Amino acid difference formula to help explain protein evolution. *Science* 185:862-864.

Greenough TC, Carville A, Coderre J, Somasundaran M, Sullivan JL, Luzuriaga K and Mansfield K (2005) Pneumonitis and multi-organ system disease in common marmosets (*Callithrix jacchus*) infected with the Severe Acute Respiratory Syndrome-Associated Coronavirus. *Am J Pathol* 167:455-463.

Guo Q, Li M, Wang C, Wang P, Fang Z, Tan J, Wu S, Xiao Y and Zhu H (2020) Host and infectivity prediction of Wuhan 2019 novel coronavirus using deep learning algorithm. *bioRxiv*:914044.

Hamming I, Cooper ME, Haagmans BL, Hooper NM, Korstanje R, Osterhaus AD, Timens W, Turner AJ, Navis G and van Goor H (2007) The emerging role of ACE2 in physiology and disease. *J Pathol* 212:1-11.

Hofmann H, Hattermann K, Marzi A, Gramberg T, Geier M, Krumbiegel M, Kuate S, Uberla K, Niedrig M and Pöhlmann S (2004) S protein of severe acute respiratory syndrome-associated coronavirus mediates entry into hepatoma cell lines and is targeted by neutralizing antibodies in infected patients. *J Virol* 78:6134-6142.

Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A *et al.* (2020) SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181:271-280.e8

Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X *et al.* (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395:497-506.



- Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Shepard D, Parton A, Armean IM, Trevanion SJ, Flicek P *et al.* (2018) Ensembl variation resources. Database 2018:12.
- Ji LD, Li JY, Yao BB, Cai XB, Shen QJ and Xu J (2017) Are genetic polymorphisms in the renin-angiotensin-aldosterone system associated with essential hypertension? Evidence from genome-wide association studies. *J Hum Hypertens* 31:695–698.
- Ji W, Wang W, Zhao X, Zai J and Li X (2020) Cross-species transmission of the newly identified coronavirus 2019-nCoV. *J Med Virol* 92:433–440.
- Kan B, Wang M, Jing H, Xu H, Jiang X, Yan M, Liang W, Zheng H, Wan K, Liu Q *et al.* (2005) Molecular evolution analysis and geographic investigation of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms. *J Virol* 79:11892–11900.
- Kuba K, Imai Y, Rao S, Gao H, Guo F, Guan B, Huan Y, Yang P, Zhang Y, Deng W *et al.* (2005) A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat Med* 11:875–879.
- Kuiken T, Fouchier RA, Schutten M, Rimmelzwaan GF, van Amerongen G, van Riel D, Laman JD, de Jong T, van Doornum G, Lim W *et al.* (2003) Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* 362:263–270.
- Kumar S, Stecher G, Suleski M and Hedges SB (2017) TimeTree: A resource for timelines, timetrees, and divergence times. *Mol Biol Evol* 34:1812–1819.
- Lam TT, Shum MH, Zhu HC, Tong YG, Ni XB, Liao YS, Wei W, Cheung WY, Li WJ, Li LF *et al.* (2020) Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature* (in press).
- Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30:3276–3278.
- Letko M, Marzi A and Munster V (2020) Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* 5:562–569.
- Li F, Li W, Farzan M and Harrison SC (2005) Structure of SARS Coronavirus spike receptor-binding domain complexed with receptor. *Science* 309:1864–1868.
- Li G, Fan Y, Lai Y, Han T, Li Z, Zhou P, Pan P, Wang W, Hu D, Liu X *et al.* (2020) Coronavirus infections and immune responses. *J Med Virol* 92:424–432.
- Li R, Qiao S and Zhang G (2020) Analysis of angiotensin-converting enzyme 2 (ACE2) from different species sheds some light on cross-species receptor usage of a novel coronavirus 2019-nCoV. *J Infect* 80:469–496.
- Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC *et al.* (2003) Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426:450–454.
- Li W, Zhang C, Sui J, Kuhn JH, Moore MJ, Luo S, Wong SK, Huang IC, Xu K, Vasilieva N *et al.* (2005) Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J* 24:1634–1643.
- Li X, Zai J, Zhao Q, Nie Q, Li Y, Foley BT and Chaillon A (2020) Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J Med Virol* (in press)
- Liu Z, Xiao X, Wei X, Li J, Yang J, Tan H, Zhu J, Zhang Q, Wu J and Liu L (2020) Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *J Med Virol* (in press).
- Longdon B, Brockhurst MA, Russell CA, Welch JJ and Jiggins FM (2014) The evolution and genetics of virus host shifts. *PLoS Pathog* 10:e1004395.
- Luan J, Lu Y, Jin X and Zhang L (2020) Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. *Biochem Biophys Res Commun* 526:165–169.
- McAuliffe J, Vogel L, Roberts A, Fahle G, Fischer S, Shieh WJ, Butler E, Zaki S, St. Claire M, Murphy B *et al.* (2004) Replication of SARS coronavirus administered into the respiratory tract of African Green, rhesus and cynomolgus monkeys. *Virology* 330:8–15.
- McKinney CA, Fattah C, Loughrey CM, Milligan G and Nicklin SA (2014) Angiotensin-(1-7) and angiotensin-(1-9): function in cardiac and vascular remodelling. *Clin Sci* 126:815–827.
- Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simao TLL, Stadler T *et al.* (2011) Impacts of the cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334:521–524.
- Millan MA, Jacobowitz DM, Catt KJ and Aguilera G (1990) Distribution of angiotensin II receptors in the brain of nonhuman primates. *Peptides* 11:243–253.
- Munster VJ, Koopmans M, van Doremalen N, van Riel D and de Wit E (2020) A novel Coronavirus emerging in China — key questions for impact assessment. *N Engl J Med* 382:692–694.
- Olivall KJ, Hosseini PR, Zambrana-Torrel C, Ross N, Bogich TL and Daszak P (2017) Host and viral traits predict zoonotic spillover from mammals. *Nature* 546:646–650.
- Oudit GY, Kassiri Z, Jiang C, Liu PP, Poutanen SM, Penninger JM and Butany J (2009) SARS-coronavirus modulation of myocardial ACE2 expression and inflammation in patients with SARS. *Eur J Clin Invest* 39:618–625.
- Peck KM, Scobey T, Swanstrom J, Jensen KL, Burch CL, Baric RS and Heise MT (2017) Permissivity of dipeptidyl peptidase 4 orthologs to Middle East Respiratory Syndrome Coronavirus is governed by glycosylation and other complex determinants. *J Virol* 91:e00534-17.
- Peng G, Yang Y, Pasquarella JR, Xu L, Qian Z, Holmes KV and Li F (2017) Structural and molecular evidence suggesting coronavirus-driven evolution of mouse receptor. *J Biol Chem* 292:2174–2181.
- Reguera J, Santiago C, Mudgal G, Ordoño D, Enjuanes L and Casanovas JM (2012) Structural bases of coronavirus attachment to host Aminopeptidase N and its inhibition by neutralizing antibodies. *PLoS Pathog* 8:e1002859.
- Simmons G, Reeves JD, Rennekamp AJ, Amberg SM, Piefer AJ and Bates P (2004) Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. *Proc Natl Acad Sci USA* 101:4240–4245.
- Simmons G, Zmora P, Gierer S, Heurich A and Pöhlmann S (2013) Proteolytic activation of the SARS-coronavirus spike protein: Cutting enzymes at the cutting edge of antiviral research. *Antiviral Res* 100:605–614.
- Shi J, Wen Z, Zhong G, Yang H, Wang C, Huang B, Liu R, He X, Shuai L, Sun Z *et al.* (2020) Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science* eabb7015.
- Towler P, Staker B, Prasad SG, Menon S, Tang J, Parsons T, Ryan D, Fisher M, Williams D, Dales NA *et al.* (2004) ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J Biol Chem* 279:17996–18007.

- Turner AJ, Hiscox JA and Hooper NM (2004) ACE2: from vasopeptidase to SARS virus receptor. *Trends Pharmacol Sci* 25:291–294.
- UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47:D506–D515.
- Van Doremalen N and Munster VJ (2015) Animal models of Middle East respiratory syndrome coronavirus infection. *Antiviral Res* 122:28–38.
- Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT and Veesler D (2020) Structure, function, and antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181:281–292.e6.
- Wan Y, Shang J, Graham R, Baric RS and Li F (2020) Receptor recognition by the novel Coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS Coronavirus. *J Virol* 94:e00127–20.
- Wang C, Horby PW, Hayden FG and Gao GF (2020) A novel coronavirus outbreak of global health concern. *Lancet* 395:470–473.
- Weiss SR and Leibowitz JL (2011) Coronavirus Pathogenesis. In: Maramorosch K, Shatkin AJ and Murphy FA (eds) *Advances in Virus Research*. Academic Press, Burlington, pp 85–164.
- Wertheim JO, Chu DKW, Peiris JSM, Kosakovsky Pond SL and Poon LLM (2013) A case for the ancient origin of Coronaviruses. *J Virol* 87:7039–7045.
- Woolhouse M, Scott F, Hudson Z, Howey R and Chase-Topping M (2012) Human viruses: discovery and emergence. *Philos Trans R Soc Lond B Biol Sci* 367:2864–2871.
- Woolhouse MEJ and Gowtage-Sequeria S (2005a) Host range and emerging and reemerging pathogens. *Emerg Infect Dis* 11:1842–1847.
- Woolhouse MEJ, Haydon DT and Antia R (2005b) Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol Evol* 20:238–244.
- Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, Meng J, Zhu Z, Zhang Z, Wang J *et al.* (2020) Genome composition and divergence of the novel Coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* 27:325–328.
- Xu X, Chen P, Wang J, Feng J, Zhou H, Li X, Zhong W and Hao P (2020) Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci China Life Sci* 63:457–460.
- Yan R, Zhang Y, Li Y, Xia L, Guo Y and Zhou Q (2020) Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* 367:1444–1448.
- Yang Z (1998) Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol* 15:568–573.
- Yang Z (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591.
- Yang Z, Wong WSW and Nielsen R (2005) Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22:1107–1118.
- Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME and Fouchier RAM (2012) Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* 367:1814–1820.
- Zhang H, Penninger JM, Li Y, Zhong N and Slutsky AS (2020) Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. *Intensive Care Med* 46:586–590.
- Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL *et al.* (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579:270–273.

## Internet Resources

- Ensembl, [www.ensembl.org](http://www.ensembl.org) (accessed 2 March 2020)
- GenBank, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov) (accessed 2 March 2020)
- Timetree, [www.timetree.org](http://www.timetree.org) (accessed 3 March 2020)
- Uniprot, [www.uniprot.org](http://www.uniprot.org) (accessed 5 March 2020)
- WHO - World Health Organization, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> (accessed 15 March 2020).

## Supplementary Material

- Table S1 - Placental mammal species analyzed.
- Table S2 - Evolutionary rates estimated considering 70 ACE2 placental mammal orthologues.
- Table S3 - Bayesian Empirical Bayes (BEB) analysis indicating sites with high probability (> 95%) to be under positive selection
- Table S4 - Bayes Empirical Bayes (BEB). General results. Signal peptide (1-17), extracellular domain (residues 18-740), transmembrane domain (741 -761), and cytoplasmic domain (761-805). See  $\omega$  column.
- Figure S1 - Highlighted variable sites in placental mammalian ACE2 protein in the interaction with SARS-CoV-2.

*Associate Editor: Carlos F.M. Menck*

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.