Prediction analysis of SARS-COV-2 entry in Livestock and Wild animals

Manas Ranjan Praharaj[1], Priyanka Garg[1], Raja Ishaq Nabi Khan[2], Shailesh Sharma[1], Manjit Panigrahi[2], B P Mishra[2], Bina Mishra[2], G Sai kumar[2], Ravi Kumar Gandham[1*], Raj Kumar Singh[2*], Subeer Majumdar[1*] and Trilochan Mohapatra[3]

[1] National Institute of Animal Biotechnology, Hyderabad, Telangana, India

[2] ICAR-Indian Veterinary Research Institute, Izatnagar, UP, India

[3] Indian Council of Agricultural Research, New Delhi, India

[*] Corresponding authors: Ravi Kumar Gandham: ravigandham@niab.org.in; R.K Singh: rks_virology@rediffmail.com; Subeer Majumdar: director@niab.org.in

Abstract

SARS-CoV-2 is a viral pathogen causing life threatening disease in human. Interaction between spike protein of SARS-CoV-2 and ACE2 receptor on the cells is a potential factor in the infectivity of a host. Using in-silico analysis, the protein and nucleotide sequences of ACE2 were initially compared across different species to identify key differences among them. This phylogeny and alignment comparison did not lead to any meaningful conclusion on viral entry facilitation in different hosts. The 6LZG - Structure of novel coronavirus spike receptor-binding domain complexed with its receptor - ACE2, was taken as a reference, to model the ACE2 receptor of various species and assess its comparative binding ability to the spike receptor-binding domain of SARS-CoV-2. Out of the several parameters estimated concerning binding of ACE2 with spike receptor-binding domain, a significant difference between the known infected and uninfected species was observed for Entropy side chain, Van der Waals, Solvation Polar, Solvation Hydrophobic and Interface Residues. However, these parameters did not specifically categorize the animals into infected or uninfected, for all the Orders (of animals). This clearly established the fact that no single parameter should be used to predict SARS-CoV-2 entry. The logistic regression model constructed upon taking all the parameters led to inclusion of parameters - Interaction energy, entropy sidechain and entropy mainchain for estimating the probability of viral entry in different species. In the mammalian class, most of the species of Carnivores, Artiodactyls, Perissodactyls, Pholidota, and Primates showed high probability of viral entry. However, among the primates, baboons have very low probability of viral entry. Among rodents, hamster was highly probable for viral entry with rats and mice having a very low probability. Rabbits have a medium probability of viral entry. In Birds, ducks have a very low probability, while chickens seemed to have medium probability and turkey showed the highest probability of viral entry. Although, viral entry alone does not determine infection in host(s), the predictions emerged out of this study may prompt us to closely follow certain species of animals for determining pathogenic insult by SARS-CoV-2 and for determining their ability to act as a carrier and/or disseminator.

Keywords: SARS-CoV-2, COVID-19, Livestock, ACE2, modeling

Introduction

Three large-scale disease outbreaks during the past two decades, *viz.*, Severe Acute Respiratory Syndrome (SARS), Middle East Respiratory Syndrome (MERS), and Swine Acute Diarrhea Syndrome (SADS) were caused by three zoonotic coronaviruses. SARS and MERS, which emerged in 2003 and 2012, respectively, caused a worldwide pandemic claiming 774 (8,000 SARS cases) and 866 (2,519 MERS cases) human lives, respectively[1], while SADS devastated livestock production by causing fatal diseases in pigs in 2017. The SARS and MERS viruses had several common factors in having originated from bats in China and being pathogenic to human or livestock[2-4]. Seventeen years after the first highly pathogenic human coronavirus, SARS-CoV-2 is devastating the world with 4,014,436 cases and 276,251 deaths (as on May 9, 2020)[5]. This outbreak was first identified in Wuhan City, Hubei Province, China, in December 2019 and notified by WHO on 5th January 2020. The disease has since been named as COVID-19 by WHO.

Coronaviruses (CoVs) are an enveloped, crown-like viral particles belonging to the subfamily Orthocoronavirinae in the family Coronaviridae and the order Nidovirales. They harbor a positive-sense, single-strand RNA (+ssRNA) genome of 27–32 kb in size. Two large overlapping polyproteins, ORF1a and ORF1b, that are processed into the viral polymerase (RdRp) and other nonstructural proteins involved in RNA synthesis or host response modulation, cover two thirds of the genome. The rest 1/3 of the genome encodes for four structural proteins (spike (S), envelope (E), membrane (M), and nucleocapsid (N)) and other accessory proteins. The four structural proteins and the ORF1a/ORF1b are relatively consistent among the CoVs, however, number and size of accessory proteins govern the length of the CoV genome[4]. This genome expansion is said to have facilitated acquisition of genes that encode accessory proteins, which are beneficial for CoVs to adapt to a specific host[6, 7]. Next generation sequencing has increased the detection and identification of new CoV species resulting in expansion of CoV subfamily. Currently, there are four genera (α-, β-, δ-, and γ-) with thirty-eight unique species in CoV subfamily (ICTV classification) including the three highly pathogenic CoVs, *viz.*, SARS-CoV-1, MERS-CoV, SARS-CoV-2 are β-CoVs[8].

Coronaviruses are notoriously promiscuous. Bats host thousands of these types, without succumbing to illness. The CoVs are known to infect mammals and birds, including dogs, chickens, cattle, pigs, cats, pangolins, and bats. These viruses have the potential to leap to new species and in this process mutate along the way to adapt to their new host(s). COVID -19, global crisis likely started with CoV infected horseshoe bat in China. The SARS-CoV-2 is spreading around the world in the hunt of entirely new reservoir hosts for re-infecting people in the future[9]. Recent reports of COVID-19 in a Pomeranian dog and a German shepherd in Hong Kong[10]; in a domestic cat in Belgium[11]; in five Malayan tigers and three lions at the Bronx Zoo in New York City[12] and in minks[13] make it all the more necessary to predict species that could be the most likely potential reservoir hosts in times to come.

Angiotensin-converting enzyme 2 (ACE2), an enzyme that physiologically counters RAAS activation functions as a receptor for both the SARS viruses (SARS-CoV-1 and SARS-CoV-2)[14-16]. ACE2 is found attached to the outer surface of cells in the lungs, arteries, heart, kidney, and intestines[17, 18]. The potential factor in the infectivity of a cell is the interaction between SARS viruses and the ACE2 receptor[19, 20]. By comparing the ACE2 sequence, several species that might be infected with SARS-CoV2 have been identified[21]. Recent studies, exposing cells/animals to the SARS-CoV2, revealed humans, horseshoe bats, civets, ferrets, cats and pigs could be infected with the virus and mice, dogs, pigs, chickens, and ducks could not be or poorly infected[16, 22]. Pigs, chickens, fruit bats, and ferrets are being exposed to SARS-CoV2 at Friedrich-Loeffler Institute and initial results suggest that Egyptian fruit bats and ferrets are susceptible, whereas pigs and chickens are not[23]. In this cause of predicting potential hosts, no studies on ACE2 sequence comparison among species along with homology modeling and prediction, to define its interaction with the spike protein of SARS-CoV-2 are available. Therefore, the present study is taken to identify viral entry in potential hosts through sequence comparison, homology modeling and prediction.

Materials and methods

Sequence analysis

In this study, 48 (mammalian, reptilian and avian species) ACE2 complete/partial protein and nucleotide sequences available on NCBI were analyzed (**Table 1**) to understand the possible difference(s) in the ACE2 sequences that may correlate with SARS-CoV-2 viral entry into the cell. Within the mammalian class, Orders - Artiodactyla, Perrisodactyla, Chiroptera, Rodentia, Carnivora, Lagomorpha, Primates, Pholidota and Proboscidea; within the Reptilian class, Orders - Testutides and Crocodile; and within the Avian class, Orders – Acciptriformes, Anseriformes and Galliformes, were considered in the study. These orders were considered keeping in view all the possible reservoir hosts/ laboratory animal models that can possibly be infected with the SARS-CoV-2. The within between group distances were calculated in Mega 6.0[24]. The Codon-based Z test of selection (strict-neutrality (dN=dS)) to evaluate synonymous and non-synonymous substitutions across the ACE2 sequences among the Orders was done.

Phylogenetic analysis

Phylogenetic analysis of the protein sequences was done using MEGA 6.0[24]. Initially, the sequence alignment was done using Clustal W[25]. The aligned sequences were analyzed for the best nucleotide substitution model on the basis of Bayesian information criterion scores using the **JModel**Test software v2.1.7[26]. The tree was constructed by the Neighbor-joining method with the best model obtained using 1000 bootstrap replicates.

Homology modeling

The Structure of novel coronavirus spike receptor-binding domain complexed with its receptor ACE2 which was determined through X-ray diffraction is available at PDB database with ID 6LZG[27]. This available ACE2 model from PDB databse is used for homology modeling using SWISS-MODEL[28]. SWISS-MODEL is a fully automatic homology modeling server for protein structure, which can be accessed through ExPASy web server.

Protein-protein docking

The spike receptor-binding domain of 6LZG was used in docking along with the homology modelled structures of ACE2 proteins of all the hosts, *i.e.*, ACE2 of 48 hosts as a receptor and spike receptor-binding domain of SARS-CoV-2 (from 6LZG) as a ligand for protein-protein docking. GRAMM-X docking server was used for protein-protein docking, which generated a docked complex[29]. Post-docking analysis was carried out using Chimera software[30], which is an extensible program for interactive visualization and analysis of molecular structures for use in structural biology. It provides the user with high quality 3D images, density maps, trajectories of small molecules and biological macromolecules, such as proteins. The homology modelled structure(s) of each species are compared with the human 6LZG to calculate the RMSD (root mean squared deviation). As most the deviation values could not be calculated with 6LZG model, the deviation(s) with respect to different human models 108a and 6M18[31] were calculated. A significant ($P < 0.05$) correlation in the deviation values calculated from 6LZG and 6M18 was observed. As most of the values could be calculated as deviations from 6M18, these values were used for further analysis along with the parameters below.

For the binding of the modelled structure of ACE2 and the spike receptor-binding domain, using FoldX software[32], several parameters (referred as spike binding properties of ACE2) – Interaction Energy, Backbone Hydrogen bond, Side chain Hydrogen bond, Van-der-Waals interaction, Electrostatic interaction, Solvation polar, Solvation hydrophobic and Entropy sidechain, entropy mainchain, torsional clash, backbone clash, helix dipole, disulfide, electrostatic kon, Interface Residues, Interface Residue Clashing and Interface Residues VdW Clashing were estimated.

Statistical analysis for prediction

Till date, clear-cut information of 17 species that are either infected or uninfected with SARS-CoV2 is available (**Supplementary table 1**). Initially, for each parameter (spike binding properties of ACE2), the difference between the infected and uninfected is tested using both Mann-Whitney non-parametric test was done using GraphPad Prism 7.00 (GraphPad Software, La Jolla, California, USA). For those parameters that were significant the difference between Order(s) and the infected/uninfected groups was

established using Mann-Whitney non-parametric test (Note: if a species is included in the infected/uninfected group, the same is not included in its Order on comparing the Order(s) with infected/uninfected group) (**Supplementary table 2** for more information). Later, a Logistic regression model was constructed on all the 18 parameters (17 from FoldX and RMSD w.r.t 6M18) estimated above. With 18 parameters, the minimum sample size required to derive statistics that represent each parameter, is 1000[33] *(n =100 + xi i.e here :- n = 100 + (100 + (50 × 18) = 1000, with a minimum of 50 events per parameter).* The data needed to be extrapolated to at least 1000. This needed us to take an assumption that the ACE2 structure and sequence is conserved within a species. For the species - *Homo sapiens*, we compared around 60 ACE2 sequences and found that all the compared sequences were completely identical. With this assumption that the spike binding properties of ACE2 within a species is conserved and because of the pandemic nature of the disease the data was extrapolated. All the parameters were included in the glm - logistic regression to construct the best model (based on R^2) for prediction. The goodness of fit was tested with Hosmer and Lemeshow goodness of fit test. The reduction in null deviance was tested with Chi-square test.

Results and Discussion

Recognition of the receptor is an important determinant in identifying the host range and cross-species infection of viruses[34]. It has been established that ACE2 is the cellular receptor of SARS-CoV-2[16]. This study is targeted to predict viral entry in a host, *i.e.,* hosts that can be reservoir hosts (Artiodactyla, Perrisodactyla, Chiroptera, Carnivora, Lagomorpha, Primates, Pholidota, Proboscidea, Testutides, Crocodilia, Acciptriformes and Galliformes) and hosts that can be appropriate small animal laboratory models (Rodentia) of SARS-CoV-2 through sequence comparison and homology modeling of ACE2 and prediction

The protein and DNA sequence lengths of ACE2 varied in different hosts (**Table 1**). Among the sequences that were compared, the longest CDS was found in the Order - Chiroptera (*Myotis braditii* - 811 aa) and the smallest in the Order – Proboscidea (*Loxodonta africana* - 800 aa). *Homo sapiens* ACE2 is taken as a standard to compare

all the sequences because of the on-going pandemic nature of COVID-19 and the availability of its 3D structure - 6LZG[27]. The within group mean distance, the parameter indicative of variability of nucleotide sequences within the group was found to be minimum in Perrisodactyla followed by Primates and was maximum among the Galliformes followed by Chiroptera (**Table 2**). This indicates that within the group of primates, all the considered species are prone to be equally infected with SARS-CoV-2 as humans. Further, to establish the probability of SARS-CoV-2 entry into species of other Orders, the distance of all orders from Primates was assessed **(Table 3)**. This distance was found minimum for Perissodactyls followed by Carnivores and maximum for Galliformes followed by Anseriformes. This confirms with the recent reports of Chicken (Galliformes) and ducks (Anseriformes) not being infected with SARS-CoV-2[22], and tigers and lions being infected[12]. To decide a cut-off distance that can establish whether the species can be infected or not, the individual distance of each species from *Homo sapiens* was evaluated **(Supplementary Table 3)**. *Melaegris gallapova* (Turkey) is the species, which had the greatest distance from *Homo sapiens*. Recently, it was reported that SARS-CoV-2 does not infect pigs, chickens, ducks[22] and rats[35]. The minimum distance that corresponds to the species that is already established to be uninfected with the SARS-CoV-2 would be 0.187 of *Rattus norvegicus* (Rat). Considering this distance from *Homo sapiens* as a cut-off, would include all the carnivores, perissodactyls and few artiodactyls viz. Goat, buffalo, Bison and sheep, to be infected and excludes cattle (Artiodactyla), all the bats (Chiroptera) and birds (Galliformes, Anseriformes and Accipitriformes). Similar distance values were observed on evaluating the protein sequences as well (**Table-2 &Table-3**). These results do not lead to meaningful conclusions on viral entry in different species, thereby, making it inevitable to depend on other parameters like evaluating the spike-interacting domain of ACE2.

The spike interacting domain of the *Homo sapiens* ACE2 protein is defined in the UniProt ID Q9BYF1. The family and domains sections of the UniProt ID Q9BYF1 clearly marks the sequence location of the ACE2 - spike interacting domains as 30 - 41aa, 82 - 84 aa and 353 - 357 aa. The nucleotide sequence alignments at positions that correspond to the spike-binding domain of *Homo sapiens* ACE2 are 90-123 bp; 244-252

bp and 1058-1071 bp. This spike interacting ACE2 domain sequences at the nucleotide level and protein level (**Figure 1 and Figure 2**) were compared and evaluated. The alignment shows that the sequence is well conserved within the Orders, suggesting that the structure defined by the sequence is conserved within the Orders. The maximum variability *w.r.t.* the *Homo sapiens* sequence within these regions was observed for Galliformes, followed by Acciptriformes, Testidunes, Crocodilia and Chiroptera. The protein sequence alignment at 30-41aa, 82-84 aa and 353-357 also showed similar sequence conservation and variability (**Figure 2**). The Codon-based Test of Neutrality to understand the selection pressure on the ACE2 sequence in the process of evolution was done. The analysis showed that there was a significant negative selection between and within orders for the ACE2 sequence indicating that, though, there is a variation at the nucleotide level, the protein translation had synonymous substitutions dominating over the non-synonymous substitutions. This negative selection indicates that the structure of ACE2 is being conserved through the process of evolution.

The protein sequences that were aligned were further subjected to find the best substitution model for phylogenetic analysis. The best model on the basis of BIC was found to be JTT + G. The phylogenetic analysis clearly classified the sequences of the species into their Orders. All the sequences were clearly grouped into two clusters. The first cluster represented the Mammalian class and the second cluster was represented by two sub- clusters of Avian and Reptilian classes with high bootstrap values (**Figure 3**). Within the mammalian cluster, the artiodactyls were sub-clustered farthest to the primates and the rodents, lagomorphs and carnivores were found clustered close to the primates with reliable bootstrap values. This partially corroborates with the occurrence of SARS-CoV-2 infection in carnivores[22] since rats were found uninfected with SARS-CoV-2[35]. The Chiroptera sub-cluster had a sub-node constituting horseshoe bat (*Rhinolophus ferrumequinum*) and the fruit bats (*Pteropus Alecto and Rousettus aegyptiacus*). The COVID-19 outbreak in Wuhan in Dec 2019 was traced back to have a probable origin from horseshoe bat[16]. The virus strain RaTG13 isolated from this bat was found to have 96.2% sequence similarity with the human SARS-CoV-2. This suggests that the virus probably could enter the fruit-bat as well, since it clustered with horseshoe bat to a common sub-node. These results again leave us with no concrete

conclusions on viral entry in various hosts. Therefore, to assess the probability of viral entry in various species, homology modeling of ACE2 along with its interaction with coronavirus spike receptor-binding domain was analyzed for all the 48 hosts.

Homology modeling was done for all the ACE2 sequences based on the X-ray diffraction structure defined in 6LZG (PDB database). The models constructed were then studied for their interaction with the spike receptor binding domain defined in the same ID. It was observed that the modelled interaction of human ACE2 showed four hydrogen bonds between the ACE2 and Spike receptor binding domain. The hydrogen bonds between the ACE2 and Spike receptor binding domain varied for different species (Fig 4). In FoldX, several parameters were estimated for the binding of ACE2 with spike receptor binding domain. Logistic regression model was constructed on 17 species (known infected or uninfected) using these parameters. When each parameter was considered individually, significant difference between the infected and uninfected groups was observed for Entropy side chain, Van der Waals, Solvation Polar, Solvation Hydrophobic and Interface Residues (**Supplementary Table 4**). Each of the Order(s) was tested as a group for their possibility of infection by comparing them with the infected and uninfected groups all these significant parameters (**Figure 5, Figure 6 & Figure 7**). For the parameters - solvation hydrophobic and entropy side chain, artiodactyls were found significantly ($P<0.05$) different from the uninfected group and not significantly ($P<0.05$) different from the infected group (**Figure 5**). This indicates that the artiodactyls considered in the study can be infected. The testudines were significantly different from the infected and not significantly different from the uninfected groups for all the parameters (**Figure 6**). This suggests that the species considered under testudines may not be infected. However, analysis for the Order - Chiroptera revealed that this group is not significantly different from both the infected and uninfected groups (**Figure 7**) for all the five parameters, leaving no clue about the probability of infection in this group. This suggests that a single parameter at a time, as has been considered in recent reports[21], may not be considered and evaluated for estimating the probability of virus entry. Therefore, all the estimated parameters were considered in logistic regression to find the best possible independent variables that would influence the entry of the SARS-CoV-2.

On evaluating several models, we finally included a model with Interaction energy, entropy side chain and entropy main chain, as independent variables, with an R^2 of 0.807. Hosmer and Lemeshow goodness of fit test showed no significant difference between the model and the observed data ($p > 0.05$) indicating that the model constructed is a good fit. There was also a statistically significant reduction in null deviance on inclusion of these three parameters (**Supplementary Table 5**). The predicted probabilities are given in **Table 4**. Within the Order Artiodactyla, all species except *Sus scrofa* (Pig) had 99% probability of viral (SARS-CoV-2) entry using ACE2 as a receptor. It has been predicted that *Bos indicus* (Indian cattle) and *Bos taurus* (Exotic cattle) can act as intermediate hosts of SARS-CoV-2[36] and that pigs are not susceptible[22]. Also, Camels, which are reported to be infected with SARS-CoV[37] are equally capable of SARS-CoV-2 infection. Among the rodents, hamsters had the highest probability of viral entry[35]. It has been established that SARS-CoV-2 effectively infects hamster[38] and, rats and mice were found less probable[35]. All the Carnivores except *Lontra canadensis* (Otter) in the study had high probability of viral entry. Reports of SARS-CoV2 infection in cats[22], tigers and lions[12] substantiate our estimates obtained in the study. Rabbits had medium probability of viral entry showing some resemblance to the recent evidence of SARS-CoV-2 replication in rabbit cell lines[39]. In bats, the probability of viral entry was high in family Vespertilionidae. *Rhinolophus ferrumequinum* (horse-shoe bat) and *Phyllostomus discolor* (Pale spear-nosed bat) had lower probability of viral entry. The kidney cell line from the Rhinolophus genus was found infected with SARS-CoV but not with SARS-CoV-2[39]. However, probability of viral entry in chicken and ducks was found to be low. All the primates except baboon were predicted to have ~ 100% probability of viral entry as evident from the devasting nature of the disease in humans. Among the reptiles, both the testudines and crocodilia, showed low probability of viral entry. In the class Aves, *Anas platyrhynchos* (ducks) and *Haliaeetus albicilla* (eagles) showed the lowest probability followed by *Gallus gallus* (chicken). *Aquila chrysaetos chrysaetos* (Golden Eagle) and *Meleagris gallapova* (turkey) showed highest probability of viral entry. The 95% confidence intervals are narrow for most of the species indicating that the sample picked up randomly can have the probability of viral entry as mentioned in **Table 4.**

Most of the species considered in this study showed high probability of viral entry. However, viral entry is not the only factor that determines infection in COVID-19 as viral loads were found to be high in asymptomatic patients[40, 41]. The important factors that determine disease/infection(COVID-19) in host(s) are – Host defense potential, underlying health conditions, host behavior and number of contacts, Age, Atmospheric temperature, Population density, Airflow and ventilation and Humidity[42].

Table 1. Species considered in this study

Order	Name (common name)	Accession number	Nucleotide length (bp)	Accession number	Amino acid length(aa)
Artiodactyla	*Bos indicus* **(Indian Cattle)**	XM_019956160.1	2436	XP_019811719.1	811
	Bos indicus x Bos taurus **(Indian crossbred Cattle)**	XM_027533926.1	2436	XP_027389727.1	811
	Bos taurus **(Exotic Cattle)**	XM_005228428.4	2436	XP_005228485.1	811
	Bubalus bubalis **(Buffalo)**	XM_006041540.2	2412	XP_006041602.1	803
	Bison bison bison **(American bison)**	XM_010834699.1	1294	XP_010833001.1	431
	Camelus bactrianus **(Double humped Camel)**	XM_010968001.1	2418	XP_010966303.1	805
	Camelus dromedaries **(Single humped camel)**	XM_010993415.2	2418	XP_010991717.1	805
	Capra hircus **(Goat)**	NM_001290107.1	2415	NP_001277036.1	804
	Ovis aries **(Sheep)**	XM_012106267.3	2415	XP_011961657.1	804
	Sus scrofa **(Pig)**	NM_001123070.1	2418	NP_001116542.1	805
Perissodactyla	*Equus asinus* **(Donkey)**	XM_014857647.1	2352	XP_014713133.1	783
	Equus caballus **(Horse)**	XM_001490191.5	2418	XP_001490241.1	805
Chiroptera	*Pteropus alecto* **(Black fruit bat)**	XM_006911647.1	2418	XP_006911709.1	805
	Rhinolophus ferrumequinum **(Greater horseshoe bat)**	AB297479.1	2418	BAH02663.1	805
	Myotis brandtii **(Brandt's bat)**	XM_014544294.1	2460	XP_014399780.1	819
	Eptesicus fuscus **(Big brown bat)**	XM_008154928.2	2436	XP_008153150.1	811
	Desmodus rotundus **(Common vampire bat)**	XM_024569930.1	2415	XP_024425698.1	804
	Phyllostomus discolor **(Pale spear-nosed bat)**	XM_028522516.1	2415	XP_028378317.1	804
	Rousettus aegyptiacus **(Egyptian fruit bat)**	XM_016118926.1	2418	XP_015974412.1	805
Pholidota	*Manis javanica* **(Sunda pangolin)**	XM_017650257.1	2418	XP_017505746.1	805
Carnivora	*Felis catus* **(Cat)**	XM_023248796.1	2424	XP_023104564.1	807
	Panthera tigris altaica **(Siberian Tiger)**	XM_007090080.2	2394	XP_007090142.1	797

	Mustela putorius furo **(Ferret)**	XM_004758885.2	2418	XP_004758942.1	805
	Canis lupus familiaris **(Dog)**	NM_001165260.1	2415	NP_001158732.1	804
	Vulpes vulpes **(Red Fox)**	XM_025986727.1	2415	XP_025842512.1	804
	Lontra canadensis **(North American river otter)**	XM_032880138.1	2418	XP_032736029.1	805
Rodentia	*Mus musculus* **(Mouse)**	NM_027286.4	2418	NP_081562.2	805
	Rattus norvegicus **(Rat)**	NM_001012006.1	2418	NP_001012006.1	805
	Cricetulus griseus **(Hamster)**	XM_027432806.1	2412	XP_027288607.1	805
Lagomorpha	*Oryctolagus cuniculus* **(Rabbit)**	XM_002719845.3	2418	XP_002719891.1	805
	Ochotona princeps **(American pika)**	XM_004597492.2	2427	XP_004597549.2	808
Primates	*Homo sapiens* **(Human)**	NM_001371415.1	2418	NP_001358344.1	805
	Pan troglodytes **(Chimpanzee)**	XM_016942979.1	2418	XP_016798468.1	805
	Papio anubis **(Baboon)**	XM_021933040.1	2418	XP_021788732.1	805
	Macaca nemestrina **(Southern pig-tailed monkey)**	XM_011735203.2	2418	XP_011733505.1	805
	Macaca mulatta **(Rhesus monkey)**	NM_001135696.1	2418	NP_001129168.1	805
	Macaca fascicularis **(Crab eating monkey)**	XM_005593037.2	2418	XP_005593094.1	805
Proboscidea	*Loxodonta Africana* **(African elephant)**	XM_023555192.1	2403	XP_023410960.1	800
Galliformes	*Gallus gallus* **(Chicken)**	XM_416822.5	2427	XP_416822.2	808
	Meleagris gallopavo **(Turkey)**	XM_019612009.2	2586	XP_019467554.1	861
Anseriformes	*Anas platyrhynchos* **(Mallard)**	XM_013094461.3	2418	XP_012949915.2	805
Accipitriformes	*Aquila chrysaetos chrysaetos* **(Golden Eagle)**	XM_029999165.1	2430	XP_029855025.1	809
	Haliaeetus albicilla **(White-tailed eagle)**	XM_009927339.1	1887	XP_009925641.1	629
Crocodilia	*Alligator sinensis* **(Chinese alligator)**	XM_025210843.1	2412	XP_025066628.1	803
	Crocodylus porosus **(Salt water alligator)**	XM_019529281.1	2412	XP_019384826.1	803
Testudines	*Pelodiscus sinensis* **(Chinese softshell turtle)**	XM_006122829.3	2427	XP_006122891.1	808
	Chelonia mydas **(Green sea turtle)**	XM_007070499.1	2436	XP_007070561.1	811

	Chrysemys picta bellii **(Painted turtle)**	XM_024108749.1	2487	XP_023964517.1	828

Table 2. With Mean group distance among the Orders

Order	Within Mean group distance (DNA)	Within Mean group distance (Protein)
Perrisodactyla	0.01	0.02
Primates	0.02	0.03
Accipitriformes	0.03	0.03
Crocodilia	0.04	0.05
Carnivora	0.07	0.10
Testudines	0.07	0.11
Artiodactyla	0.08	0.10
Rodentia	0.10	0.12
Lagomorpha	0.12	0.13
Chiroptera	0.14	0.23
Galliformes	0.21	0.28

Table 3. Between group distance (between Primates and other groups)

Order	Primates (DNA)	Primates (Protein)
Perrisodactyla	0.131	0.164
Carnivora	0.162	0.199
Pholidota	0.163	0.183
Lagomorpha	0.165	0.197
Rodentia	0.179	0.211
Chiroptera	0.181	0.249
Artiodactyla	0.186	0.241
Proboscidea	0.189	0.237
Testudines	0.516	0.573
Crocodilia	0.518	0.565
Accipitriformes	0.562	0.528
Anseriformes	0.594	0.587
Galliformes	0.605	0.653

Table 4. Probability of viral entry in different species

Class	Order	Family	Species (Common name)	Probability of Viral Entry (95% Confidence Interval)
Mammalia	Artiodactyla	Bovidae	*Bos indicus* **(Indian Cattle)**	9.979E-01(9.91E-01 – 9.99E-01)
			Bos taurus **(Exotic Cattle)**	9.978E-01(9.91E-01 – 9.99E-01)
			Bubalus bubalis **(Buffalo)**	9.954E-01(9.92E-01 – 1.00E+00)
			Bison bison bison **(American bison)**	1.00E+00(1.00E+00 – 1.00E+00)
			Bos indicus x *Bos taurus* **(Indian crossbred Cattle)**	9.979E-01(9.92E-01 – 1.00E+00)
		Camilidae	*Camelus bactrianus* **(Double humped Camel)**	9.989E-01(9.93E-01 – 1.00E+00)
			Camelus dromedaries **(Single humped camel)**	9.989E-01(9.93E-01 – 1.00E+00)
		Caprinae	*Capra hircus* **(Goat)**	9.998E-01(9.99E-01 – 1.00E+00)
			Ovis aries **(Sheep)**	9.998E-01(9.99E-01 – 1.00E+00)
		Suidae	*Sus scrofa* **(Pig)**	1.40E-02(2.44E-03 – 7.58E-02)
	Perrisodactyla	Equidae	*Equus asinus* **(Donkey)**	1.00E+00(1.00E+00 – 1.00E+00)
			Equus caballus **(Horse)**	4.647E-01(4.12E-02 – 9.46E-01)
	Carnivora	Mustelidae	*Mustela putorius furo* **(Ferret)**	9.95E-01(9.50E-01 – 1.00E+00)
			Lontra canadensis **(North American river otter)**	4.971E-07(5.74E-09 – 4.31E-05)
		Felidae	*Panthera tigris altaica* **(Siberian Tiger)**	9.57E-01(8.80E-01 – 9.86E-01)
		Canidae	*Vulpes Vulpes* **(Red Fox)**	9.889E-01(9.41E-01 – 9.98E-01)
			Canis lupus familiaris **(Dog)**	1.000E+00(9.99E-01 – 1.00E+00)
		Felidae	*Felis catus* **(Cat)**	1.000E+00(1.00E+00 – 1.00E+00)
	Chiroptera	Rhinolophidae	*Rhinolophus ferrumequinum* **(Greater horseshoe bat)**	9.269E-04(8.24E-05 – 1.03E-02)
		Phyllostomidae	*Desmodus rotundus* **(Common vampire bat)**	9.928E-01(9.47E-01 – 9.99E-01)
			Phyllostomus discolor **(Pale spear-nosed bat)**	7.237E-04(3.23E-05 – 1.60E-02)

		Vespertilionidae	*Eptesicus fuscus* **(Big brown bat)**	1.000E+00(1.00E+00 – 1.00E+00)
			Myotis brandtii **(Brandt's bat)**	9.998E-01(9.99E-01 – 1.00E+00)
		Pteropodidae	*Pteropus Alecto* **(Black fruit bat)**	2.650E-01(9.60E-03 – 9.31E-01)
			Rousettus aegyptiacus **(Egyptian fruit bat)**	4.83E-01(3.84E-01 – 5.84E-01)
	Rodentia	Cricetidae	*Cricetulus griseus* **(Hamster)**	8.92E-01(7.92E-01 – 9.47E-01)
		Muridae	*Mus musculus* **(Mouse)**	2.05E-04(4.56E-06 – 9.12E-03)
			Rattus norvegicus **(Rat)**	1.41E-03(1.91E-04 – 1.03E-02)
	Lagomorpha	Leporidae	*Oryctolagus cuniculus* **(Rabbit)**	6.760E-01(2.73E-01 – 9.20E-01)
		Ochotonidae	*Ochotona princeps* **(American pika)**	1.275E-01(2.94E-02 – 4.13E-01)
	Pholidota	Manidae	*Manis javanica* **(Sunda pangolin)**	1.000E+00(9.99E-01 – 1.00E+00)
	Primates	Hominidae	*Homo sapiens* **(Human)**	1.00E+00(9.98E-01 – 1.00E+00)
		Cercopithecidae	*Macaca fascicularis* **(Crab eating monkey)**	1.00E+00(1.00E+00 – 1.00E+00)
			Macaca mulatta **(Rhesus monkey)**	1.00E+00(9.99E-01 – 1.00E+00)
			Macaca nemestrina **(Southern pig-tailed monkey)**	1.00E+00(1.00E+00 – 1.00E+00)
		Hominidae	*Pan troglodytes* **(Chimpanzee)**	1.00E+00(9.98E-01 - 1.00E+00)
		Cercopithecidae	*Papio Anubis* **(Baboon)**	1.109E-09(4.36E-13 – 2.82E-06)
	Probosidae	Elephantidae	*Loxodonta Africana* **(African elephant)**	9.998E-01(9.98E-01 – 1.00E+00)
Reptiles	Testidunes	Cheloniidae	*Chelonia mydas* **(Green sea turtle)**	9.371E-03(9.11E-05 – 4.95E-01)
		Emydidae	*Chrysemys picta bellii* **(Painted turtle)**	3.781E-09(2.23E-13 – 6.41E-05)
		Trionychidae	*Pelodiscus sinensis* **(Chinese softshell turtle)**	3.851E-04(4.26E-06 – 3.37E-02)
	Crocodilia	Alligatoridae	*Alligator sinensis* **(Chinese alligator)**	6.27E-04(1.64E-06 – 1.94E-01)
		Crocodylidae	*Crocodylus porosus* **(Salt water alligator)**	2.223E-02(1.66E-04 – 7.57E-01)
Aves	Galliformes	Phasianidae	*Gallus gallus* **(Chicken)**	6.58E-01(5.81E-01 – 7.28E-01)
			Meleagris gallapova **(Turkey)**	1.00E+00(1.00E+00 – 1.00E+00)
	Anseriformes	Anatidae	*Anas platyrhynchus* **(Mallard)**	1.84E-10(1.11E-14 – 3.05E-06)
	Accipitriformes	Accipitridae	*Haliaeetus albicilla* **(White-tailed eagle)**	4.168E-01(1.61E-01 – 7.27E-01)
			Aquila chrysaetos chrysaetos **(Golden Eagle)**	9.999E-01(9.99E-01 – 1.00E+00)

References

[1] COVID-19, MERS & SARS. https://www.niaid.nih.gov/diseases-conditions/covid-19 last accessed).
[2] Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N Engl J Med 2003;348:1967-76.
[3] Zhou P, Fan H, Lan T, Yang XL, Shi WF, Zhang W, et al. Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. Nature 2018;556:255-8.
[4] Fan Y, Zhao K, Shi ZL, Zhou P. Bat Coronaviruses in China. Viruses 2019;11.
[5] COVID-19 CORONAVIRUS PANDEMIC. https://www.worldometers.info/coronavirus/ last accessed).
[6] Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, et al. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc Natl Acad Sci U S A 2014;111:E3900-9.
[7] Forni D, Cagliani R, Clerici M, Sironi M. Molecular Evolution of Human Coronavirus Genomes. Trends Microbiol 2017;25:35-48.
[8] Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 2012;367:1814-20.
[9] Scientists hunt for the next potential coronavirus animal host. https://www.nationalgeographic.com/animals/2020/03/coronavirus-animal-reservoir-research/ last accessed).
[10] Hong Kong: Blood tests confirm that pomerarian caught COVID-19 after its owner. https://www.theweek.in/news/world/2020/03/26/hong-kong-blood-tests-confirm-that-pomerarian-caught-covid-19-after-its-owner.html last accessed).
[11] A cat appears to have caught the coronavirus, but it's complicated. https://www.sciencenews.org/article/cats-animals-pets-coronavirus-covid19 last accessed).
[12] Four tigers, three lions test Covid-19 positive at Bronx Zoo. https://timesofindia.indiatimes.com/world/us/four-tigers-three-lions-test-covid-19-positive-at-bronx-zoo/articleshow/75319387.cms last accessed).
[13] Mink found to have coronavirus on two Dutch farms: ministry. https://in.reuters.com/article/health-coronavirus-netherlands-mink/mink-found-to-have-coronavirus-on-two-dutch-farms-ministry-idINKCN2280K2 last accessed).
[14] Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 2003;426:450-4.
[15] Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 2020;181:271-80 e8.
[16] Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020;579:270-3.
[17] Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. J Pathol 2004;203:631-7.

[18] Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, et al. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. Circ Res 2000;87:E1-9.
[19] Li F, Li W, Farzan M, Harrison SC. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 2005;309:1864-8.
[20] Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 2020;367:1260-3.
[21] Qiu Y, Zhao YB, Wang Q, Li JY, Zhou ZJ, Liao CH, et al. Predicting the angiotensin converting enzyme 2 (ACE2) utilizing capability as the receptor of SARS-CoV-2. Microbes Infect 2020.
[22] Shi J, Wen Z, Zhong G, Yang H, Wang C, Huang B, et al. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. Science 2020.
[23] Novel Coronavirus SARS-CoV-2: Fruit bats and ferrets are susceptible, pigs and chickens are not. https://www.fli.de/en/press/press-releases/press-singleview/novel-coronavirus-sars-cov-2-fruit-bats-and-ferrets-are-susceptible-pigs-and-chickens-are-not/ last accessed).
[24] Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 2013;30:2725-9.
[25] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673-80.
[26] Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 2012;9:772.
[27] Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. Cell 2020.
[28] Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46:W296-W303.
[29] Tovchigrechko A, Vakser IA. GRAMM-X public web server for protein-protein docking. Nucleic Acids Res 2006;34:W310-4.
[30] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605-12.
[31] Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 2020;367:1444-8.
[32] Strokach A, Corbi-Verge C, Kim PM. Predicting changes in protein stability caused by mutation using sequence-and structure-based methods in a CAGI5 blind challenge. Human Mutation 2019;40:1414-23.
[33] Bujang MA, Sa'at N, Sidik T, Joo LC. Sample Size Guidelines for Logistic Regression from Observational Studies with Large Population: Emphasis on the Accuracy Between Statistics and Parameters Based on Real Life Clinical Data. Malays J Med Sci 2018;25:122-30.
[34] Li F. Receptor recognition and cross-species infections of SARS coronavirus. Antiviral Res 2013;100:246-54.
[35] Zhang H, Penninger JM, Li Y, Zhong N, Slutsky AS. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. Intensive Care Med 2020;46:586-90.

[36] Luan J, Jin X, Lu Y, Zhang L. SARS-CoV-2 spike protein favors ACE2 from Bovidae and Cricetidae. J Med Virol 2020.
[37] Gong SR, Bao LL. The battle against SARS and MERS coronaviruses: Reservoirs and Animal Models. Animal Model Exp Med 2018;1:125-33.
[38] Lau SY, Wang P, Mok BW, Zhang AJ, Chu H, Lee AC, et al. Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. Emerg Microbes Infect 2020:1-15.
[39] Chu H, Chan JF-W, Yuen TT-T, Shuai H, Yuan S, Wang Y, et al. Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. The Lancet Microbe 2020.
[40] Rabi FA, Al Zoubi MS, Kasasbeh GA, Salameh DM, Al-Nasser AD. SARS-CoV-2 and Coronavirus Disease 2019: What We Know So Far. Pathogens 2020;9.
[41] Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. N Engl J Med 2020;382:1177-9.
[42] Lakshmi Priyadarsini S, Suresh M. Factors influencing the epidemiological characteristics of pandemic COVID 19: A TISM approach. International Journal of Healthcare Management 2020:1-10.

Abbreviations

SARS: Severe Acute Respiratory Syndrome
MERS: Middle East Respiratory Syndrome
SADS: Swine Acute Diarrhea Syndrome
SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2
COVID-19: Coronavirus disease 2019
ACE2: Angiotensin-converting enzyme 2
WHO: World Health Organization
ICTV: International Committee on Taxonomy of Viruses
PDB: Protein Data Bank
RMSD: Root-mean-square deviation
CDS: Coding Sequence

Author's contributions

MRP performed sequence alignment and phylogeny of nucleotide and amino acid and drafted the manuscript. PG and SS performed protein modelling and docking and estimated the different parameters from FoldX. RINK retrieved the amino acid and nucleotide sequences and edited the manuscripts. MP, GSK and BM edited and proofread the manuscript. RKG did complete statistical analysis and manuscript development. TM, SM, RKS, RKG and BPM conceptualized and planned the entire study.

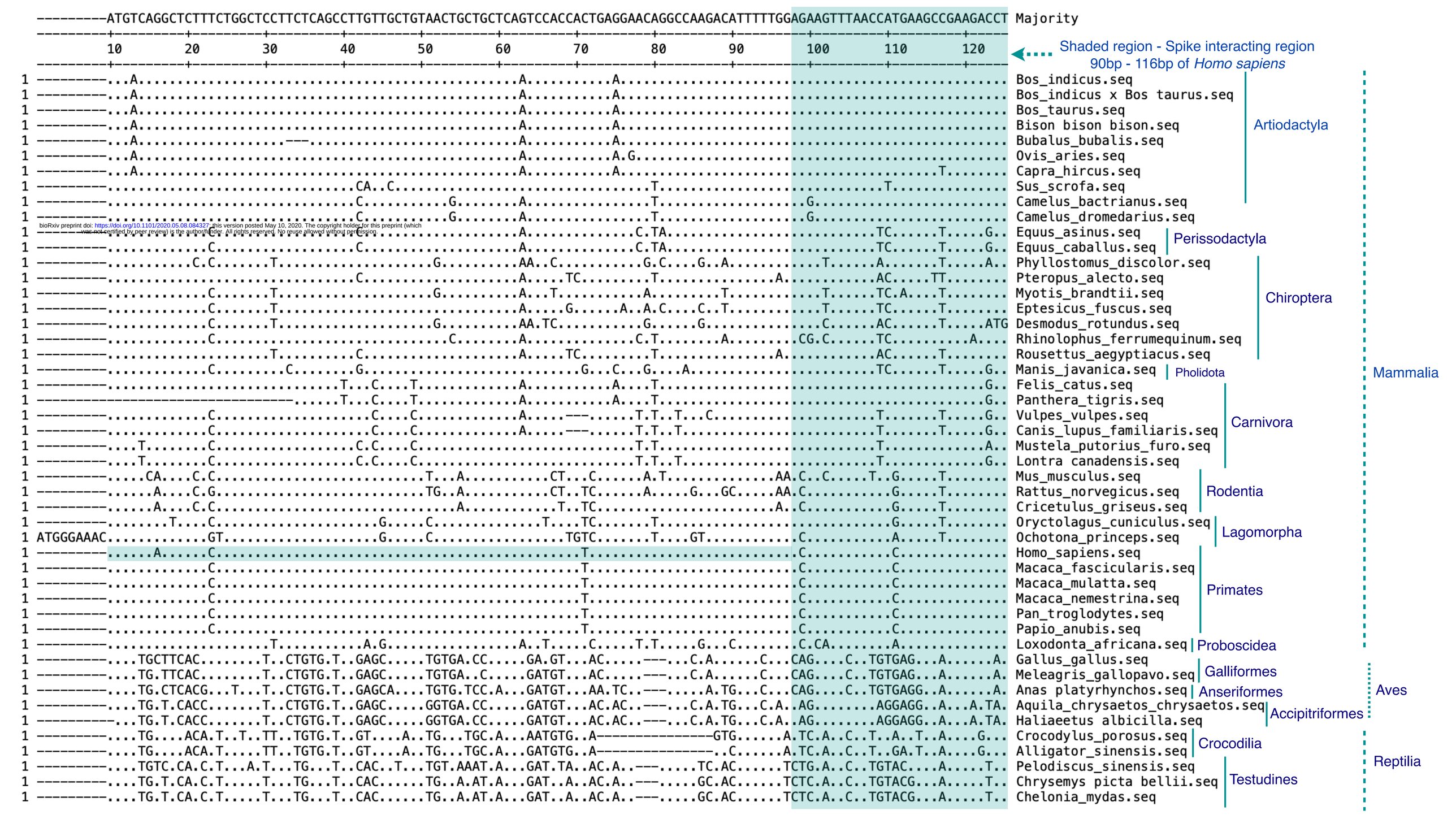
Competing interests

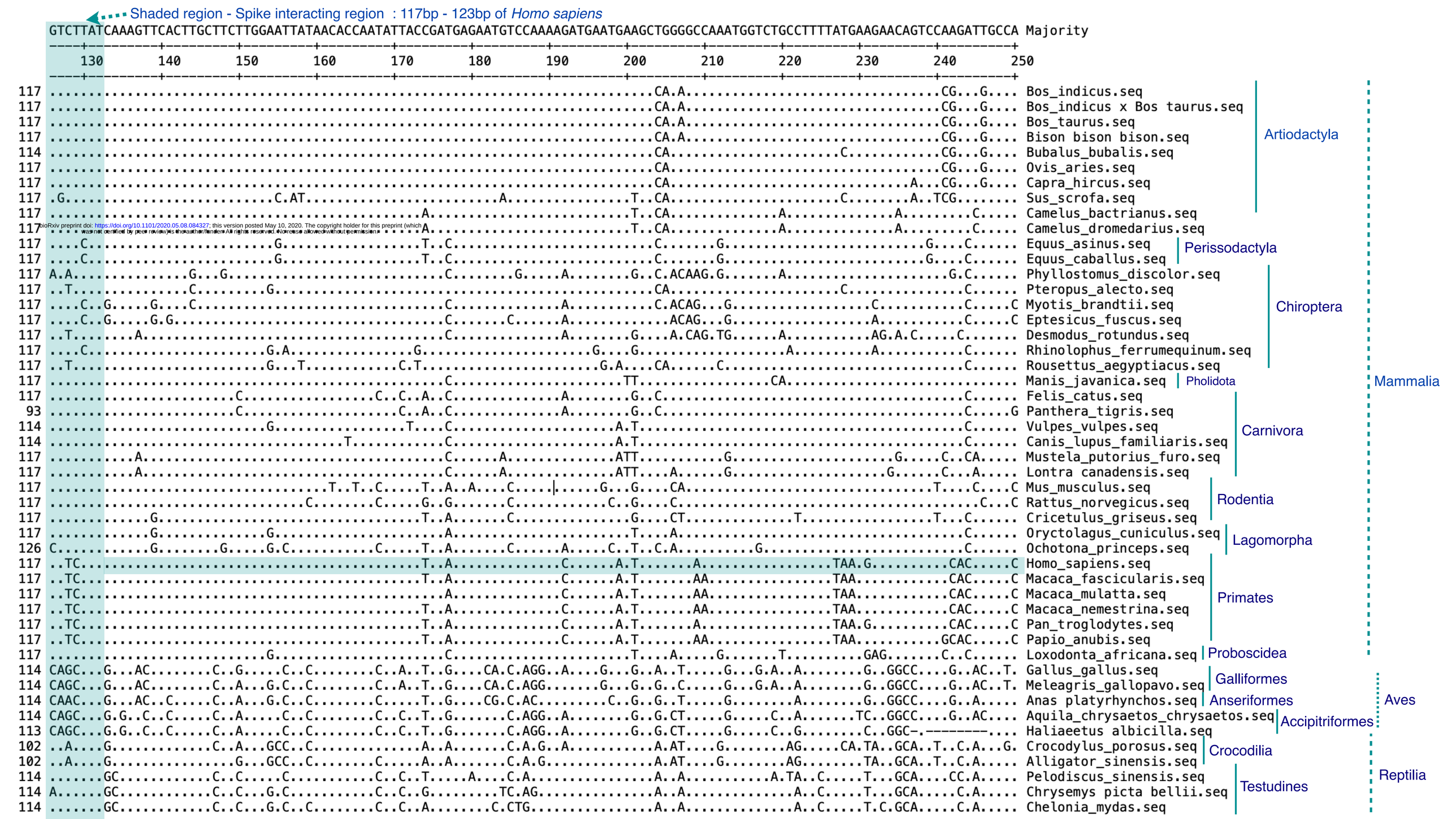
The author has declared no competing interests.

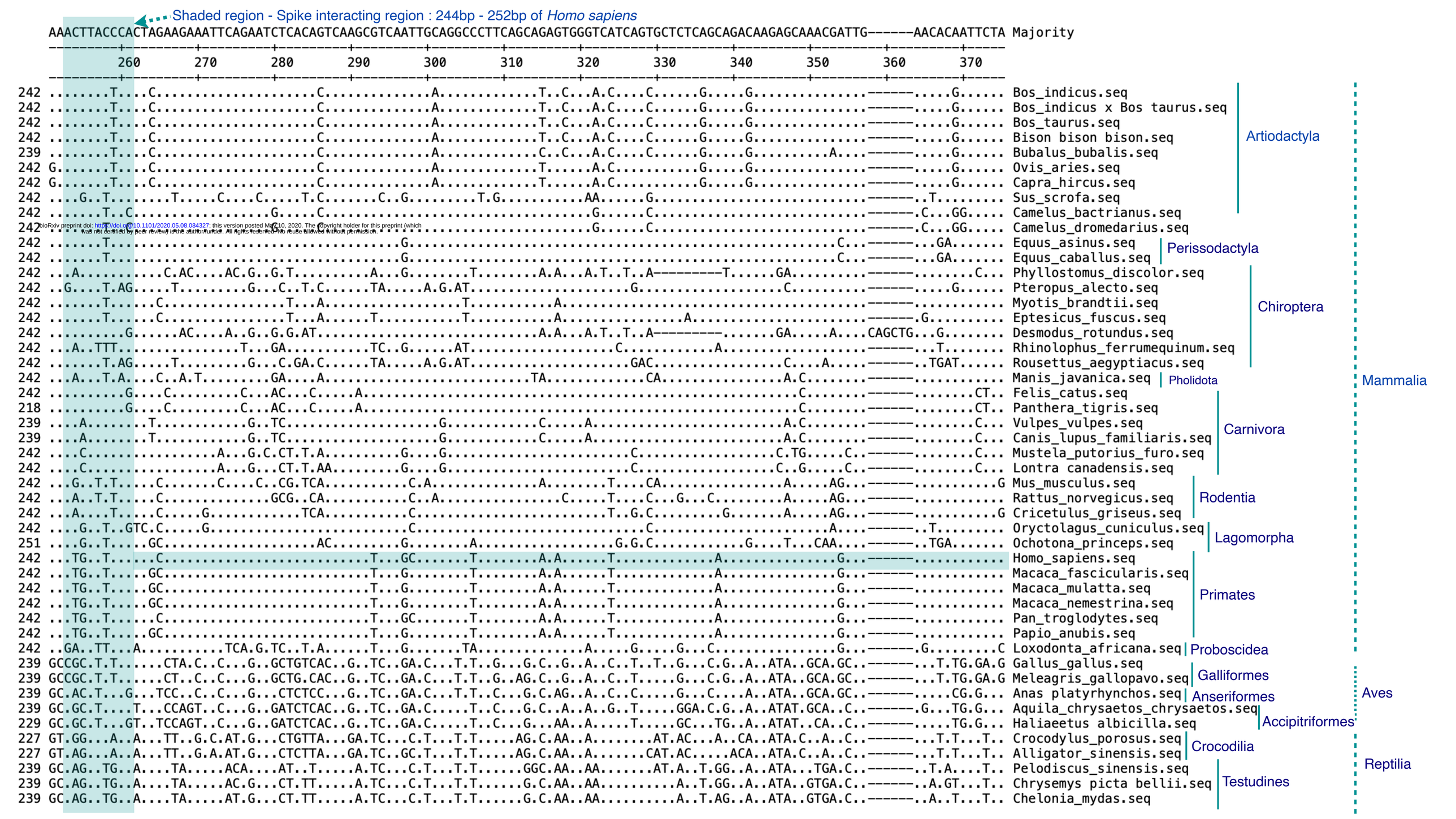
Acknowledgments

We are grateful to Director NIAB and Director IVRI for the support.

Figure 1. Nucleotide sequence alignment of the CDS region of ACE2. The shaded regions show the spike interacting domains.







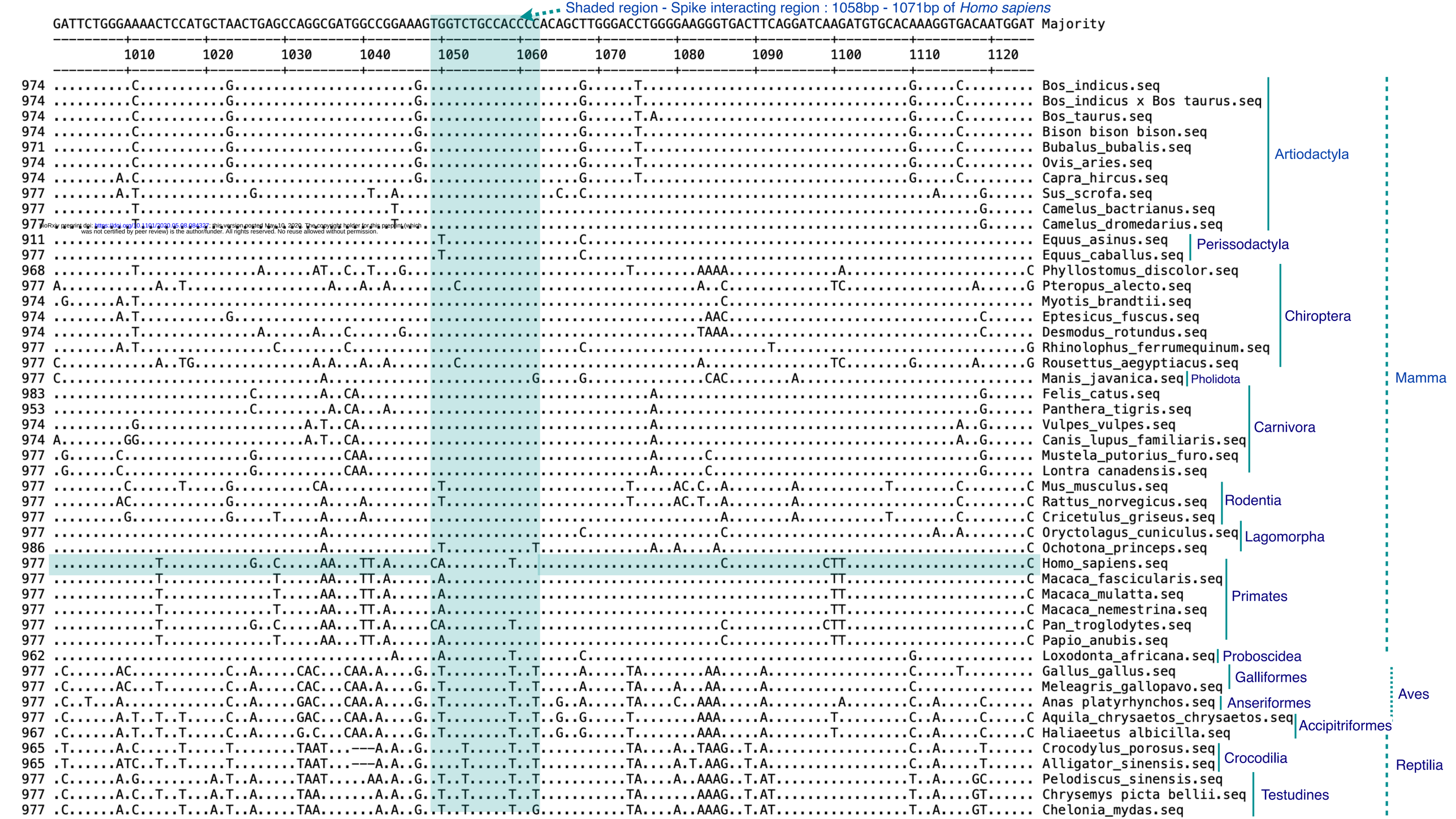
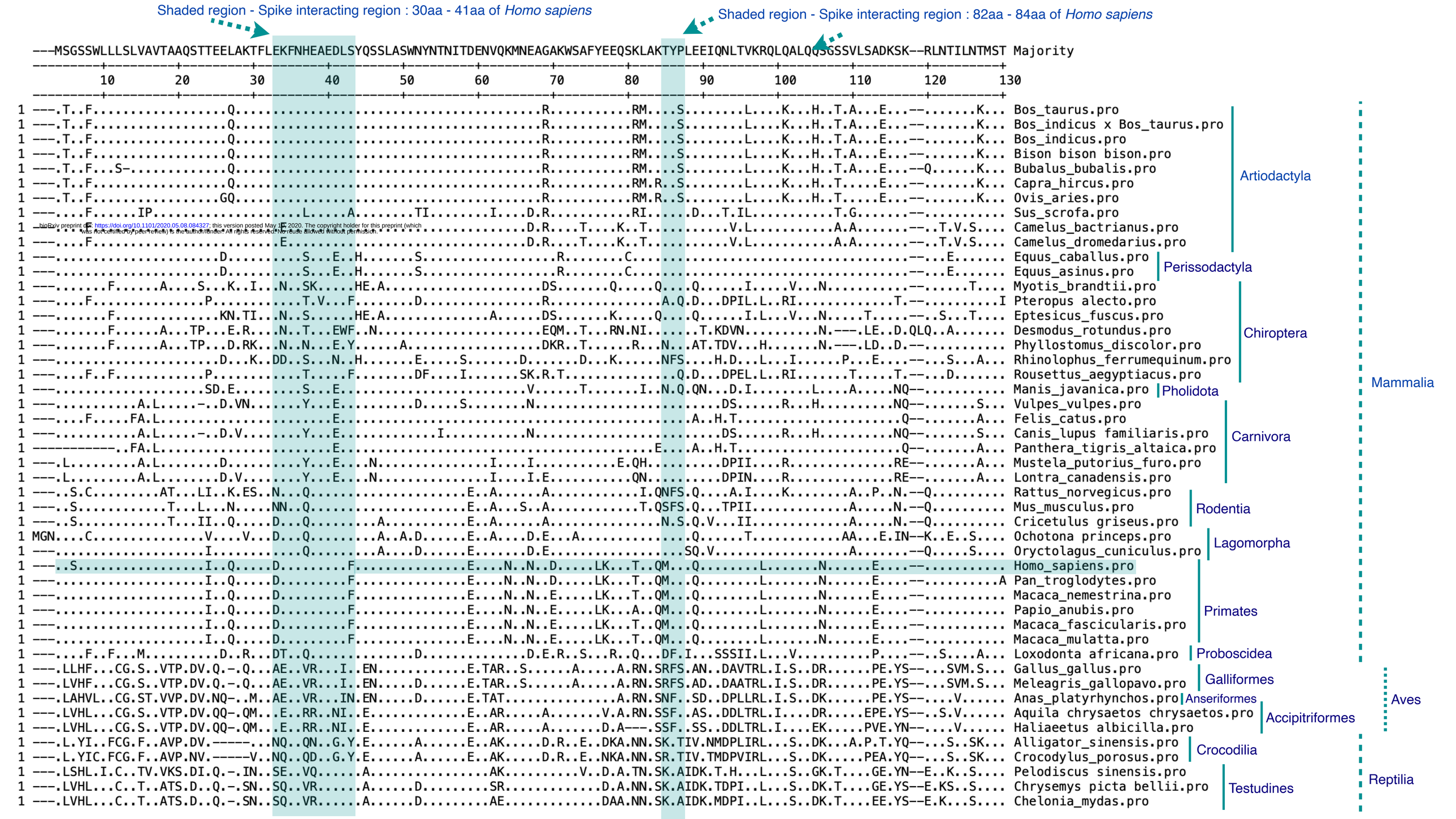


Figure 2. Protein sequence alignment of ACE2. The shaded regions show the spike interacting domains.


Shaded region - Spike interacting region : 30aa - 41aa of *Homo sapiens*
Shaded region - Spike interacting region : 82aa - 84aa of *Homo sapiens*
---MSGSSWLLLSLVAVTAAQSTTEELAKTFLEKFNHEAEDLSYQSSLASWNYNTNITDENVQKMNEAGAKWSAFYEEQSKLAKTYPLEEIQNLTVKRQLQALQQSGSSVLSADKSK--RLNTILNTMST Majority
10 20 30 40 50 60 70 80 90 100 110 120 130
Bos_taurus.pro
Bos_indicus x Bos_taurus.pro
Bos_indicus.pro
Bison bison bison.pro
Bubalus_bubalis.pro
Capra_hircus.pro
Ovis_aries.pro
Sus_scrofa.pro
Camelus_bactrianus.pro
Camelus_dromedarius.pro
Equus_caballus.pro
Equus_asinus.pro
Myotis_brandtii.pro
Pteropus alecto.pro
Eptesicus_fuscus.pro
Desmodus_rotundus.pro
Phyllostomus_discolor.pro
Rhinolophus_ferrumequinum.pro
Rousettus_aegyptiacus.pro
Manis_javanica.pro
Vulpes_vulpes.pro
Felis_catus.pro
Canis_lupus familiaris.pro
Panthera_tigris_altaica.pro
Mustela_putorius_furo.pro
Lontra_canadensis.pro
Rattus_norvegicus.pro
Mus_musculus.pro
Cricetulus griseus.pro
Ochotona princeps.pro
Oryctolagus_cuniculus.pro
Homo_sapiens.pro
Pan_troglodytes.pro
Macaca_nemestrina.pro
Papio_anubis.pro
Macaca_fascicularis.pro
Macaca_mulatta.pro
Loxodonta africana.pro
Gallus_gallus.pro
Meleagris_gallopavo.pro
Anas_platyrhynchos.pro
Aquila chrysaetos chrysaetos.pro
Haliaeetus albicilla.pro
Alligator_sinensis.pro
Crocodylus_porosus.pro
Pelodiscus sinensis.pro
Chrysemys picta bellii.pro
Chelonia_mydas.pro
Artiodactyla
Perissodactyla
Chiroptera
Pholidota
Carnivora
Rodentia
Lagomorpha
Primates
Proboscidea
Galliformes
Anseriformes
Accipitriformes
Crocodilia
Testudines
Mammalia
Aves
Reptilia
bioRxiv preprint doi: https://doi.org/10.1101/2020.05.08.084327; this version posted May 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

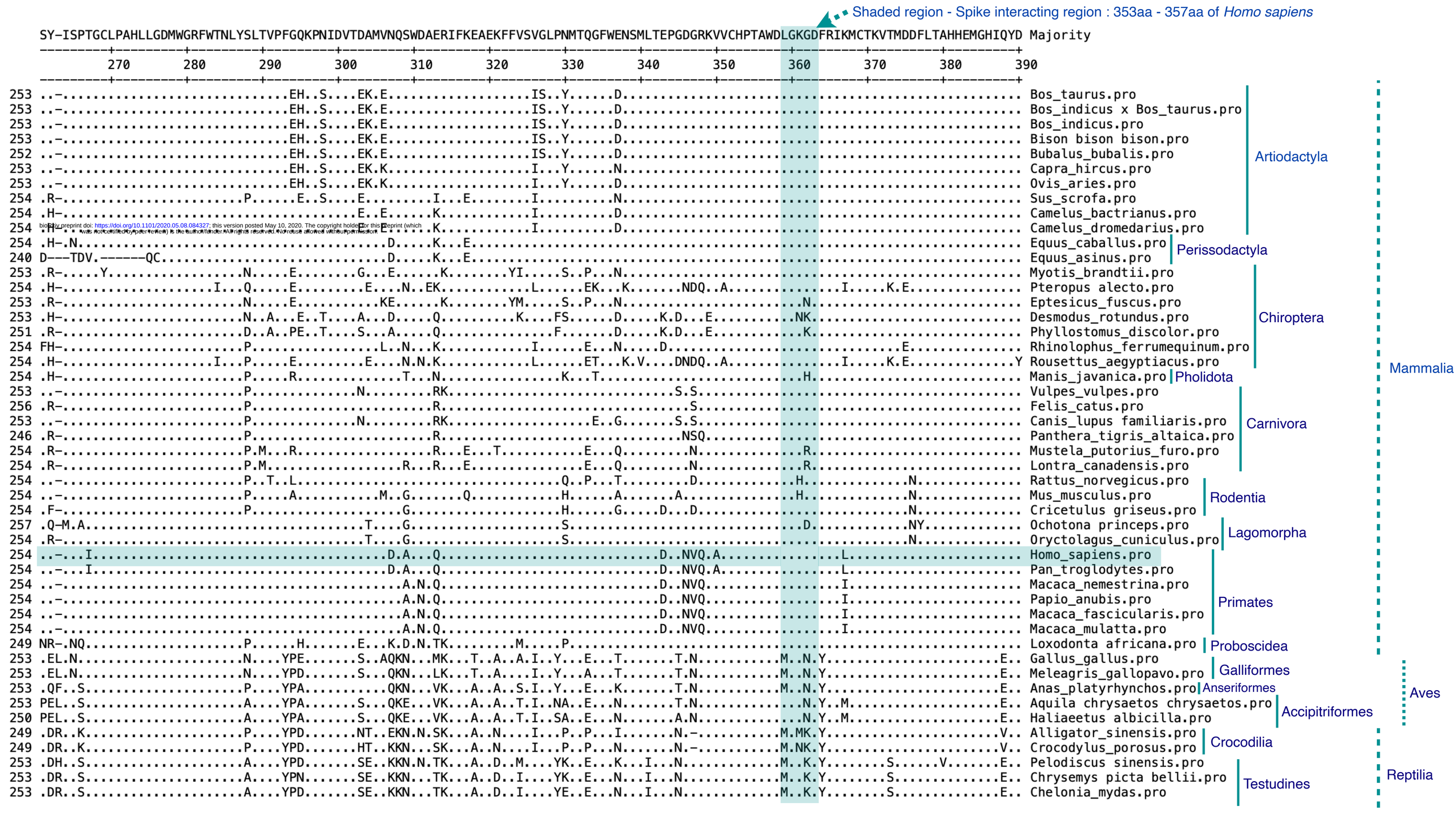


Figure 3. Phylogenetic analysis of ACE2 protein sequences. The tree was constructed using neighbor joining method in MEGA 6.0. The bootstrap values are given at each node.

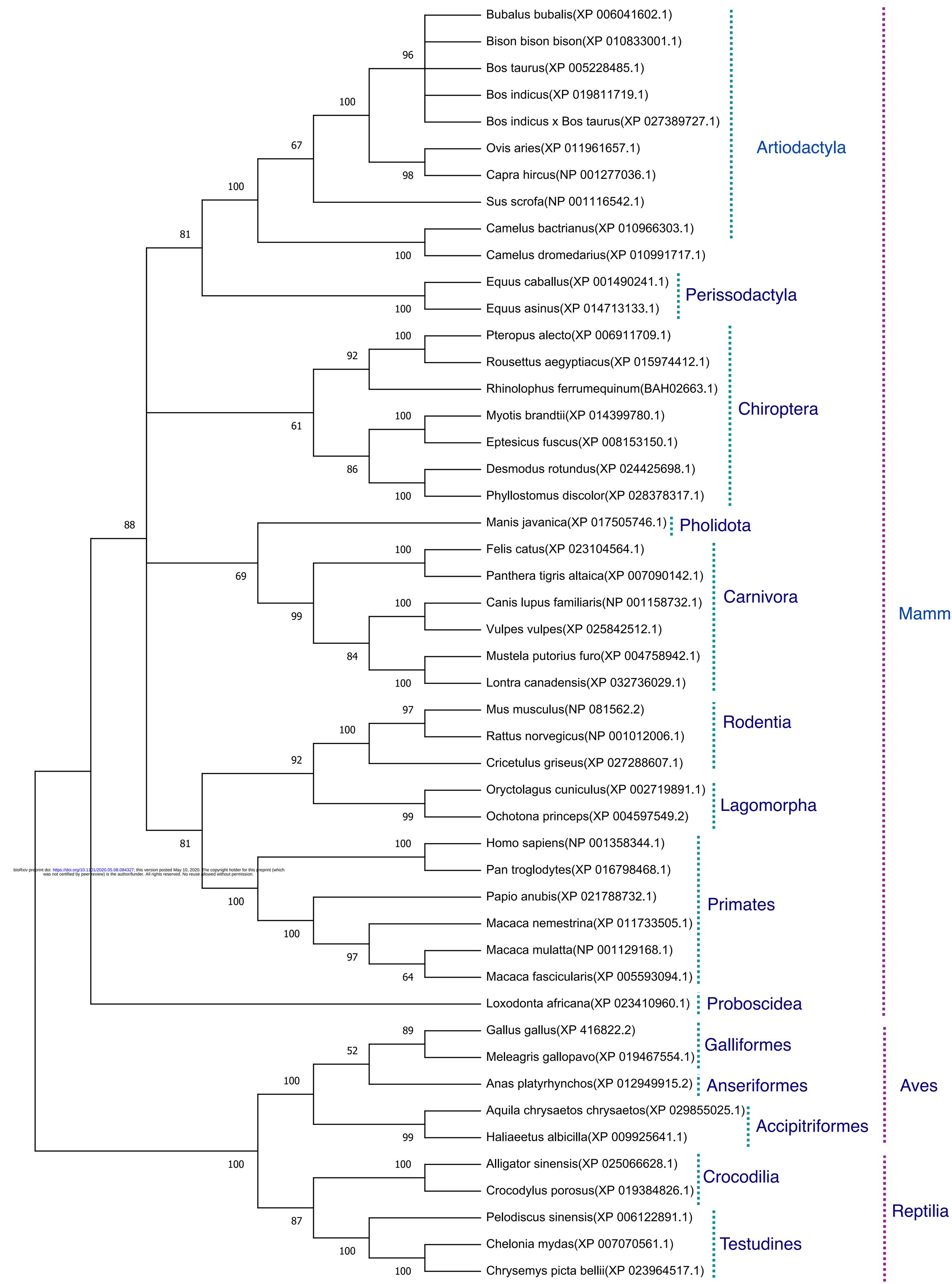


Figure 4. Representative protein modelled structures showing the interaction between ACE2 of (A) Human (B) Cat (C) Donkey (D) Exotic cattle (E) Chinese alligator & (F) Greater horseshoe bat, and spike receptor binding domain of SARS-CoV-2.

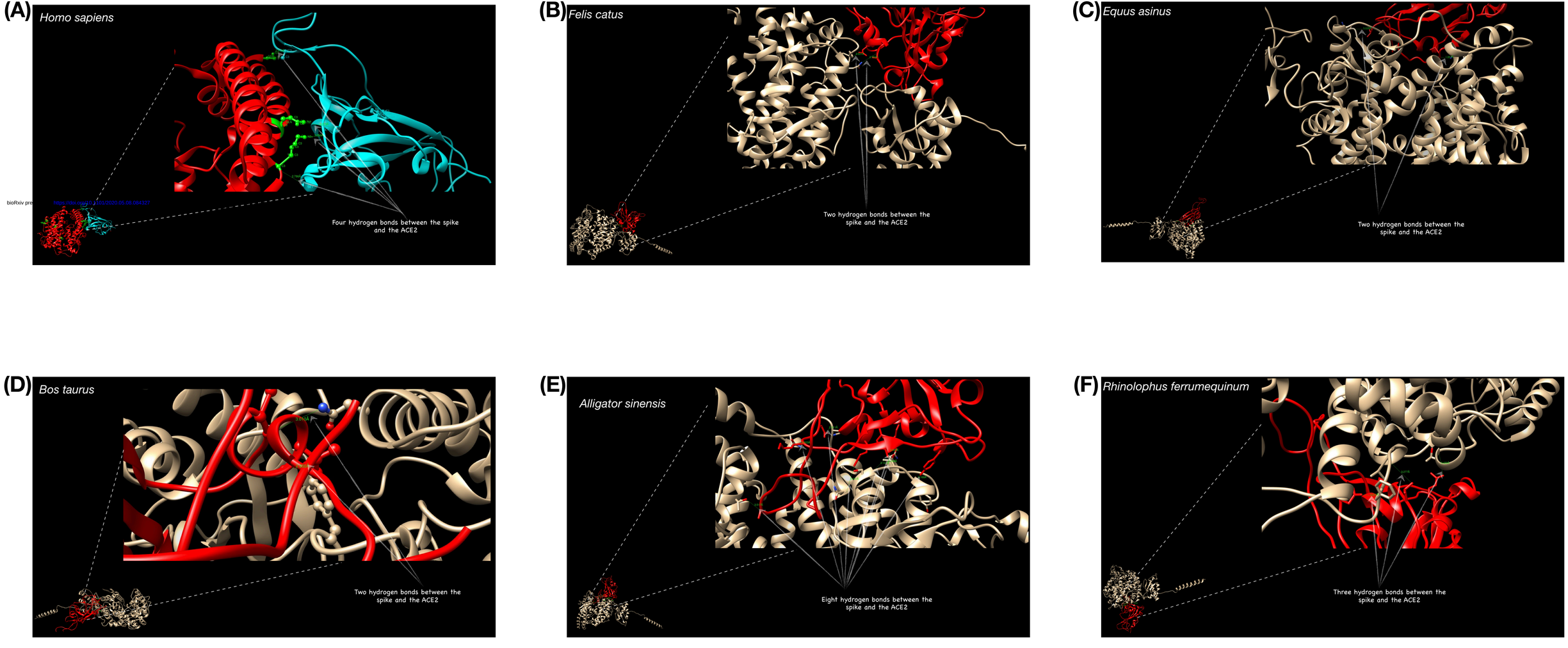
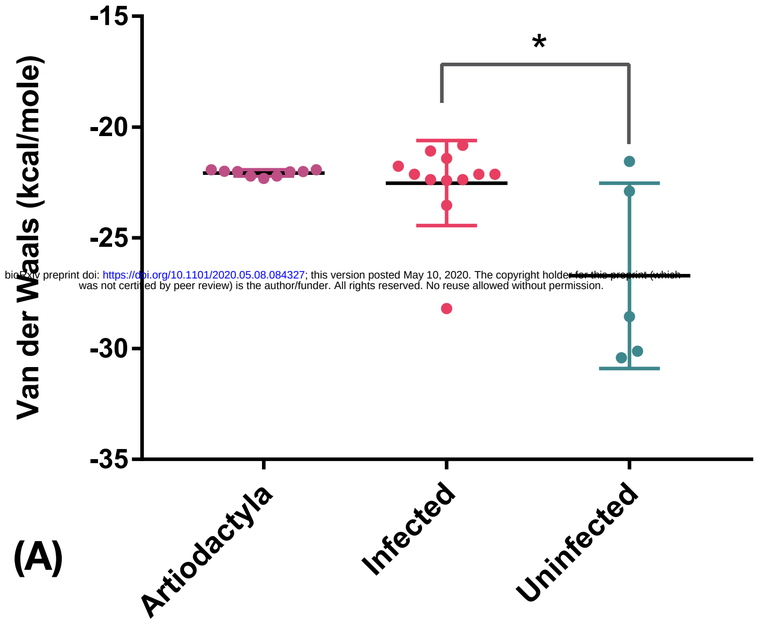
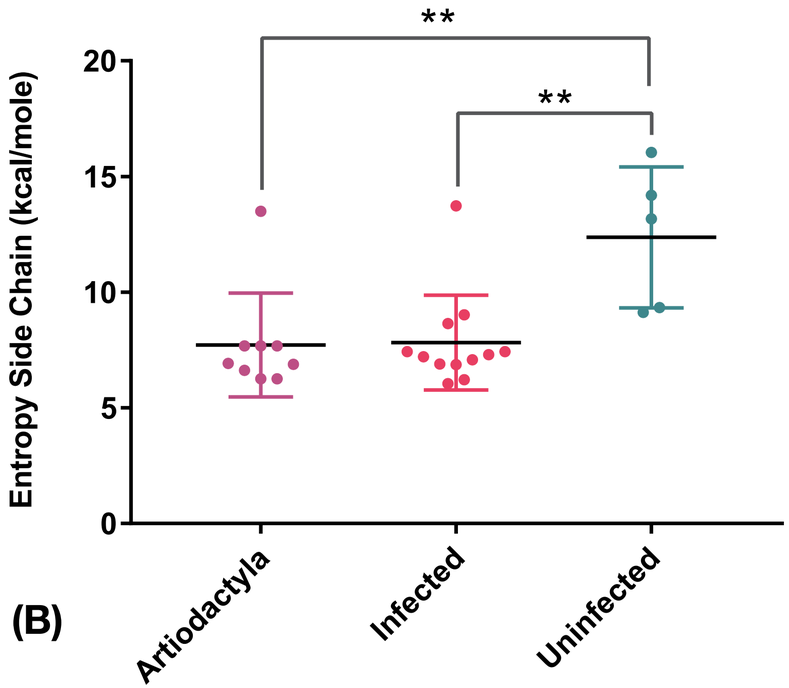
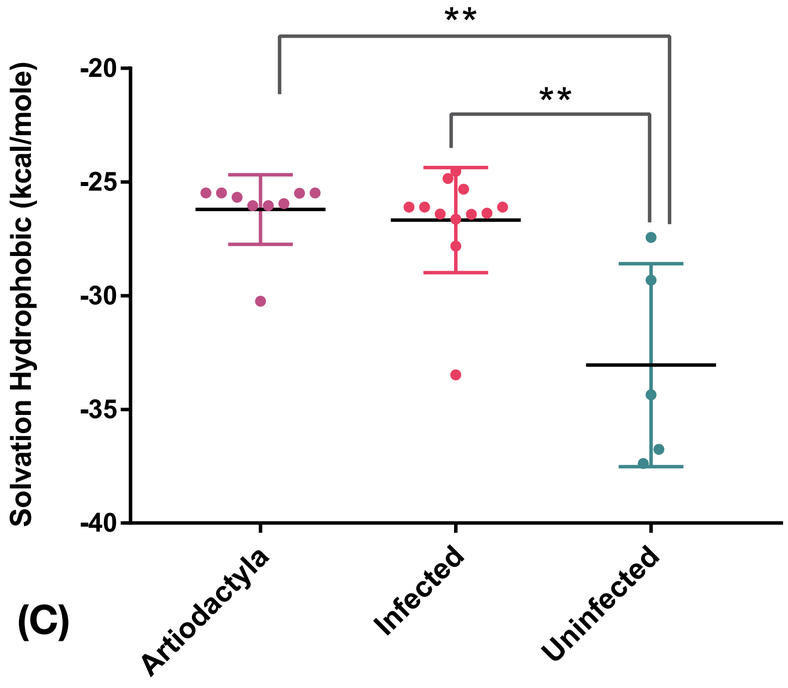
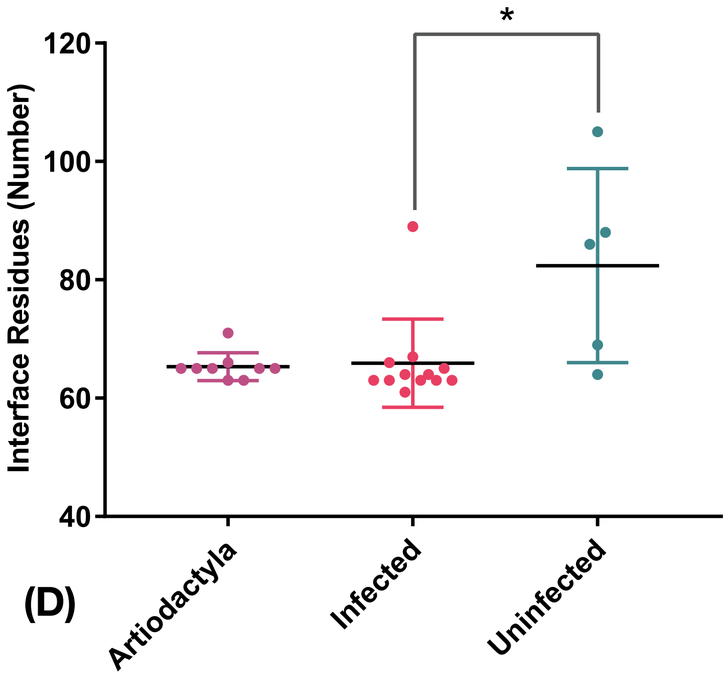


Figure 5. Scatterplot showing the comparison of Artiodactyls with infected and uninfected groups for the all five significant parameters (A). Van der Waals – No Significant difference on comparison of Artiodactyls with infected and uninfected groups. (B). Entropy side chain - Significant difference on comparison of Artiodactyls with uninfected group and no significant difference from infected group. (C). Solvation hydrophobic - Significant difference on comparison of Artiodactyls with uninfected and no significant difference from infected. (D). Interface residues - No Significant difference on comparison of Artiodactyls with infected and uninfected groups. (E). Solvation polar - No Significant difference on comparison of Artiodactyls with infected and uninfected groups. ** Significance at $P < 0.01$; * Significance at $P < 0.05$ after Mann-Whitney test on comparing two groups at a time.









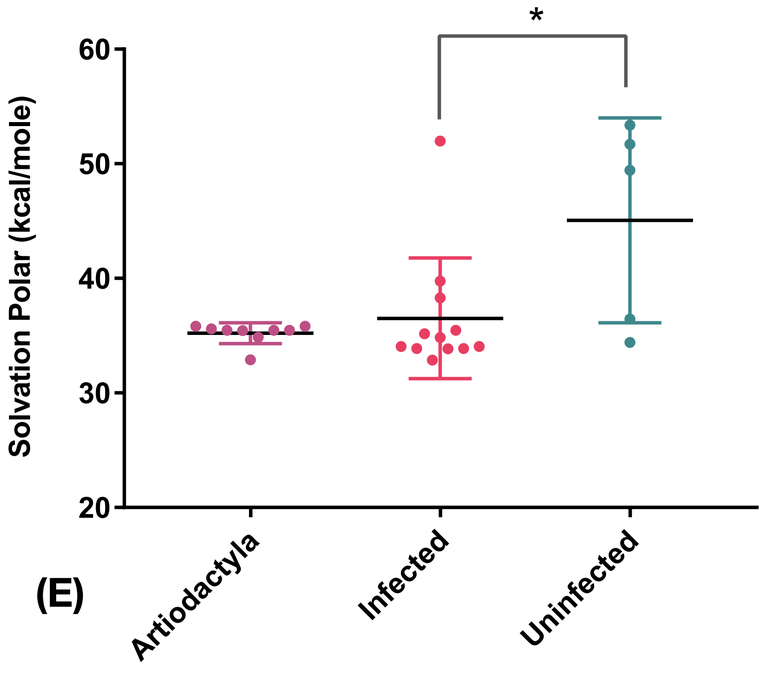
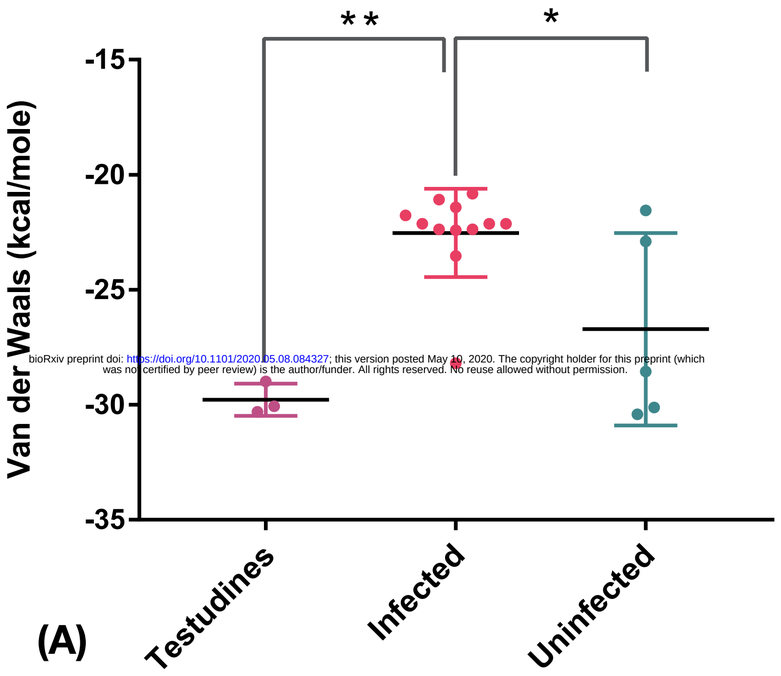
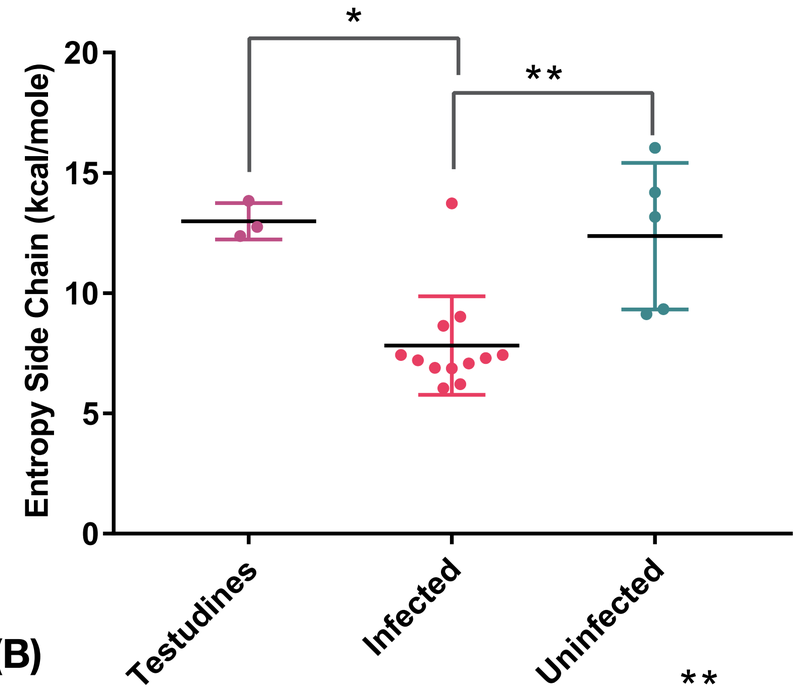
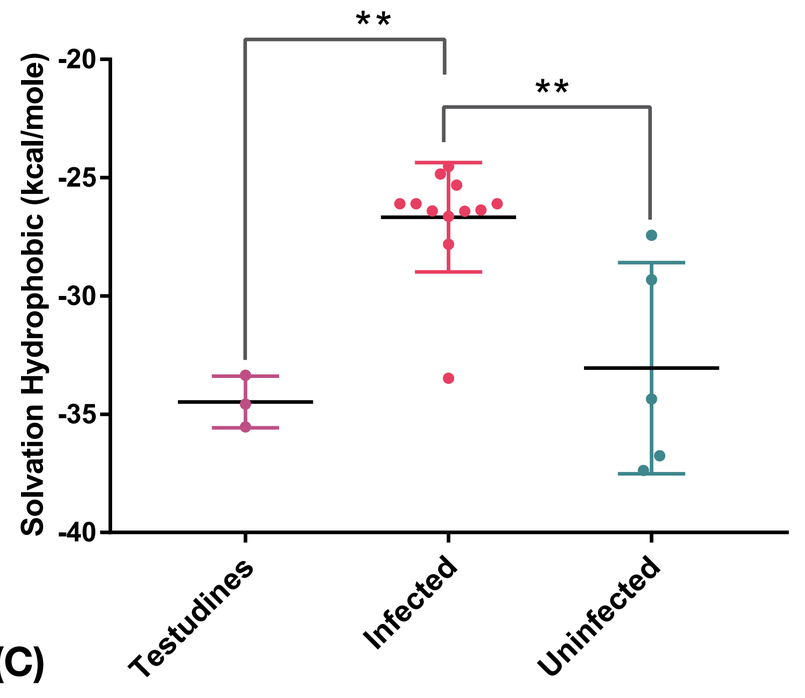
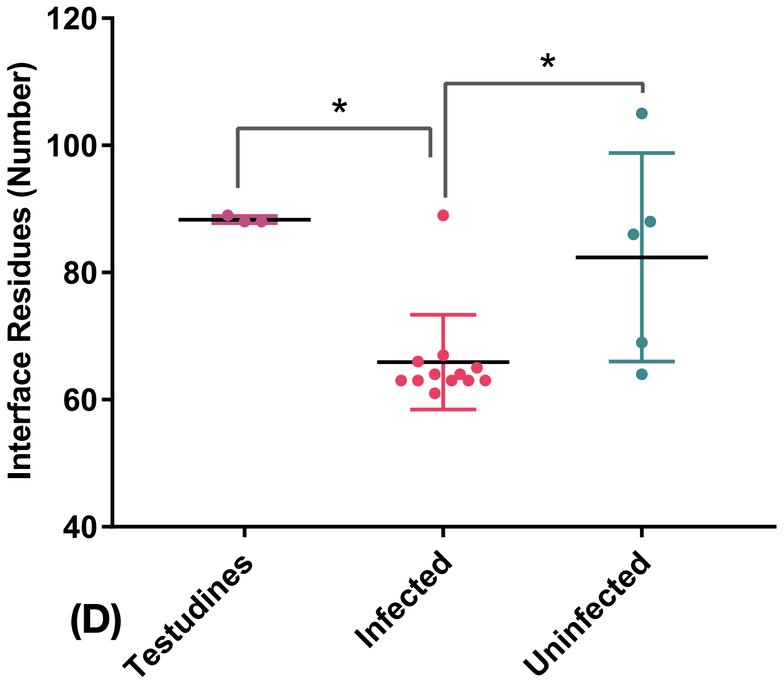


Figure 6. Scatterplot showing the comparison of Testudines with infected and uninfected groups for the all five significant parameters (A). Van der Waals – Significant difference on comparison of Testudines with infected and no significant difference from uninfected. (B). Entropy side chain - Significant difference on comparison of Testudines with infected and no significant difference from uninfected. (C). Solvation hydrophobic - Significant difference on comparison of Testudines with infected and no significant difference from uninfected. (D). Interface residues- Significant difference on comparison of Testudines with infected and no significant difference from uninfected. (E). Solvation polar - Significant difference on comparison of Testudines with infected and no significant difference from uninfected. ** Significance at $P < 0.01$; * Significance at $P < 0.05$ after Mann-Whitney test on comparing two groups at a time.









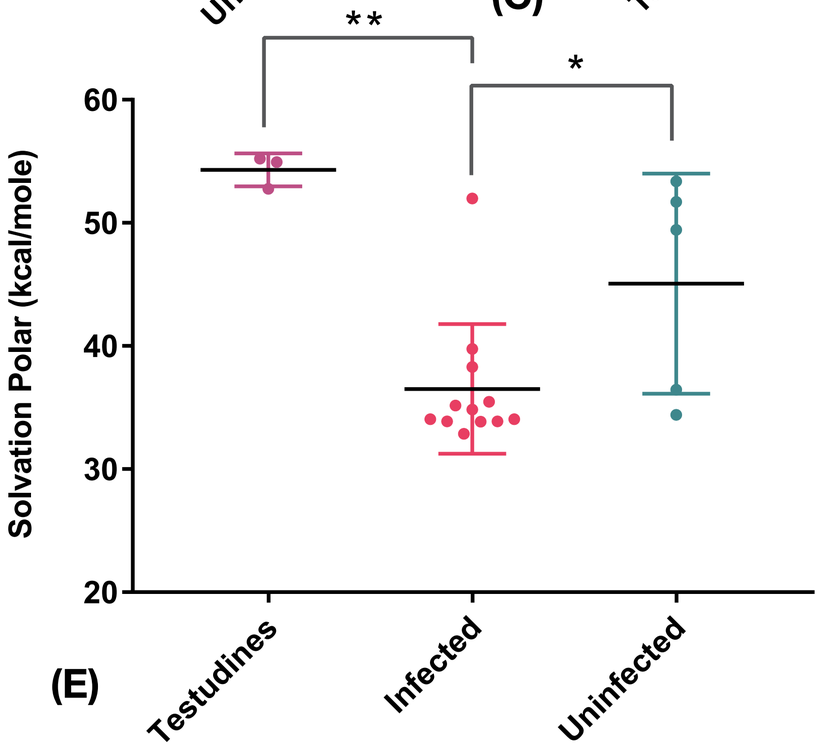


Figure 7. Scatterplot showing the comparison of Chiroptera with infected and uninfected groups for the all five significant parameters (A). Van der Waals – No Significant difference on comparison of Chiroptera with infected and uninfected groups. (B). Entropy side chain - No Significant difference on comparison of Chiroptera with infected and uninfected groups. (C). Solvation hydrophobic - No Significant difference on comparison of Chiroptera with infected and uninfected groups. (D). Interface residues- No Significant difference on comparison of Chiroptera with infected and uninfected groups. (E). Solvation polar - No Significant difference on comparison of Chiroptera with infected and uninfected groups. ** Significance at $P < 0.01$; * Significance at $P < 0.05$ after Mann-Whitney test on comparing two groups at a time.