Brief Report

Computational analysis suggests putative intermediate animal hosts of the SARS-CoV-2

Peng Chu[1,2], Zheng Zhou[2], Zhichen Gao[1], Ruiqi Cai[1], Sijin Wu[3,*], Zhaolin Sun[2], Shuyuan Chen[4], Yongliang Yang[1,*]

[1]*Laboratoy of Innovative Drug Discovery, School of Bioengineering, Dalian University of Technology, Dalian, 116023, China*;
[2]*Department of Pharmacology, Dalian Medical University, 9 West Section, Lvshun South Road, Dalian 116044, China*;
[3]*College of Pharmacy, Ohio State University, USA*;
[4]*Bioinformatics Division*, *Faculty of Science, University of Melbourne, VIC 3052, Australia*.

Key Words: intermediate host; SARS-CoV-2; computational analysis; ACE2

*To whom correspondence may be addressed:

Yongliang Yang, E-mail: everbright99@foxmail.com **(lead contact)**

Sijin Wu, E-mail: Wu.3857@osu.edu

Abstract

The recent emerged SARS-CoV-2 may first transmit to intermediate animal host from bats before the spread to humans. The receptor recognition of ACE2 protein by SARS-CoVs or bat-originated coronaviruses is one of the most important determinant factors for the cross-species transmission and human-to-human transmission. To explore the hypothesis of possible intermediate animal host, we employed molecular dynamics simulation and free energy calculation to examine the binding of bat coronavirus with ACE2 proteins of 47 representing animal species collected from public databases.

Our results suggest that intermediate animal host may exist for the zoonotic transmission of SARS-CoV-2. Furthermore, we found that tree shrew and ferret may be two putative intermediate hosts for the zoonotic spread of SARS-CoV-2. Collectively, the continuous surveillance of pneumonia in human and suspicious animal hosts are crucial to control the zoonotic transmission events caused by SARS-CoV-2.

Introduction

An outbreak of infectious pneumonia caused by the novel coronavirus SARS-CoV-2 (also known as 2019-nCoV) has emerged as a pandemic resulting in over a million confirmed cases worldwide including more than sixty thousand deaths, and the number of infected cases is still rising. Noteworthy, SARS-CoV-2 is found to be 96% identical at the whole-genome level to a bat coronavirus (BatCov-RaTG13)[1]. Moreover, it has been confirmed that SARS-CoV-2 uses a surface glycoprotein, called the spike (S) protein to bind to the same host-cell entry receptor ACE2, as SARS-CoV. Importantly, the cross-species spillover of bat-adapted CoVs into an intermediate host has been accounted for the emergence of several zoonotic coronaviruses, including SARS-CoV and MERS-CoV. This intermediate host is identified as palm civets for SARS-CoV, while camels are thought to play the role of intermediate host for MERS-CoV. Analogously, as of March 2020, several facts implicate that another wild animal may act as an intermediate host between bats and humans for the spread of SARS-CoV-2[2]. However, the existence and identity of possible intermediate host for the SARS-CoV-2 has yet to be determined. Computational analysis of the viral receptors binding with the host receptors may help to quickly pin down possible intermediate animal hosts for SARS-CoV-2.

Results and Discussion

Although we don't rule out the possibility of a direct transmission from bats to human, the recent emerged SARS-CoV-2 may first spread to an intermediate animal from bats, similar as SARS-CoV and MERS-CoV (**Figure 1A**). To explore this hypothesis, we first examined the binding of four coronaviruses including SARS-CoV-2, SARS-CoV, RaTG13-CoV and Bat-CoV with human ACE2 from a computational approach. In brief, the RBD structures of RaTG13-CoV and Bat-CoV were first built by homology modeling based on the template structures[3] followed by molecular dynamics (MD) optimization. The complex structures of RaTG13-CoV and Bat-CoV with human ACE2 were predicted by protein docking followed by 20ns MD simulation. The binding free energies of four complex structures were calculated by MM/PBSA method[4]. Our results demonstrated that SARS-CoV-2 obtained the best binding energy (-904.76 kcal/mol) as compared to the other three CoVs (SARS-CoV: -785.71 kcal/mol; RaTG13-CoV: -682.61 kcal/mol; Bat-CoV: -284.76 kcal/mol, see **Figure 1B**). This is rather consistent with recent findings that SARS-CoV-2 binds much stronger with ACE2 than SARS-CoV[3]. Moreover, our results showed that two bat-originated CoVs can't bind efficiently with human ACE2 as compared with SARS-CoV-2 or SARS-CoV, implicating that an intermediate host may exist before the zoonotic transfer of coronaviruses. In another word, bat-originated coronaviruses may have to use an intermediate host as reservoir for adaptation or recombination before the efficient transmission to humans.

Next, we sought to assess and identify the possible intermediate animal hosts for bat coronaviruses before it spread to human. Briefly, we performed an exhaustive search and collected ACE2 sequences of 47 representing animals from the Pfam database and NCBI database. All the ACE2 structures of various animals were first built by homology modeling followed by MD optimization. Here, we choose

the RBD structure of bat-originated coronavirus RaTG13-CoV as a probe to examine the binding with ACE2 receptor because early study has found that the RaTG13 bat virus remains the closet to SARS-CoV-2 across the genome. Again, our assumption is that bat-originated coronaviruses may first jump to an intermediate animal host in which the coronavirus acquired mutation or recombination before the efficient transmission to human. Moreover, it has been demonstrated that bat-originated SARS-like coronaviruses can use ACE2 protein as the cell entry receptor[5]. In addition, the concept that an ACE2-utilizing bat coronavirus can infect intermediate host before the transmission to humans has been established[6]. Subsequently, the complex structures of RaTG13-CoV-RBD with 47 ACE2 receptors were predicted by protein docking followed by MD simulation and the binding free energies of all complex structures were computed by MM/PBSA method. Interestingly, our results revealed that *Tupaia chinensis* (tree shrew) and *Mustela furo* (ferret) obtained the top-two best binding free energies as compared to other species (ΔG: -678.22 kcal/mol for tree shrew; ΔG: -665.86 kcal/mol for ferret; **see Figure 2**). In addition, the binding of RaTG13-CoV-RBD with tree shrew and ferret is much stronger than that of pangolin (ΔG, -506.49 kcal/mol; see **Figure 2 and SI_dataset_1).** This suggest that pangolin may not be the intermediate host should SARS-CoV-2 was a nature selection by mutation or recombination from bat coronaviruses. The major limitation of the present study is the lack of additional datasets of ACE2 proteins for more animal species. In addition, our analysis is based on the hypothesis that bat-originated coronaviruses use ACE2 as a cell-entry receptor for the infection of intermediate animals and therefore we only focus on the interaction of CoV-RBDs with ACE2 receptors. Nevertheless, our study suggests that intermediate animals may exist and supports the call for surveillance of suspicious intermediate animal hosts to control the spread of SARS-CoV-2.

Conclusion

In summary, our computational analysis revealed that tree shrew and ferret are two suspicious mammals as intermediate animal host for the zoonotic spread of SARS-CoV-2. We want to remind the reader that our study is based on the hypothesis that SARS-CoV-2 may be a cross-species spillover of bat-originated CoVs into an intermediate host via ACE2 before the exposure to human. Moreover, we don't rule out the possibility of a direct transmission of coronaviruses from bats or other natural reservoir to human, which is warranted for further investigation. The continuous surveillance of pneumonia in humans and suspicious animals including tree shrew and ferret are critical to understand and control the zoonotic transmission events caused by SARS-CoV-2. Nevertheless, our study provides the medical and public health research communities with predictive insights which may facilitate to control and battle against the devastating SARS-CoV-2.

Acknowledgments

Dr. Y. Yang's laboratory was supported by the National Natural Science Foundation in China (Grant: 81874301) and the Fundamental Research Funds for Central University (Grant: DLUT15QY43).

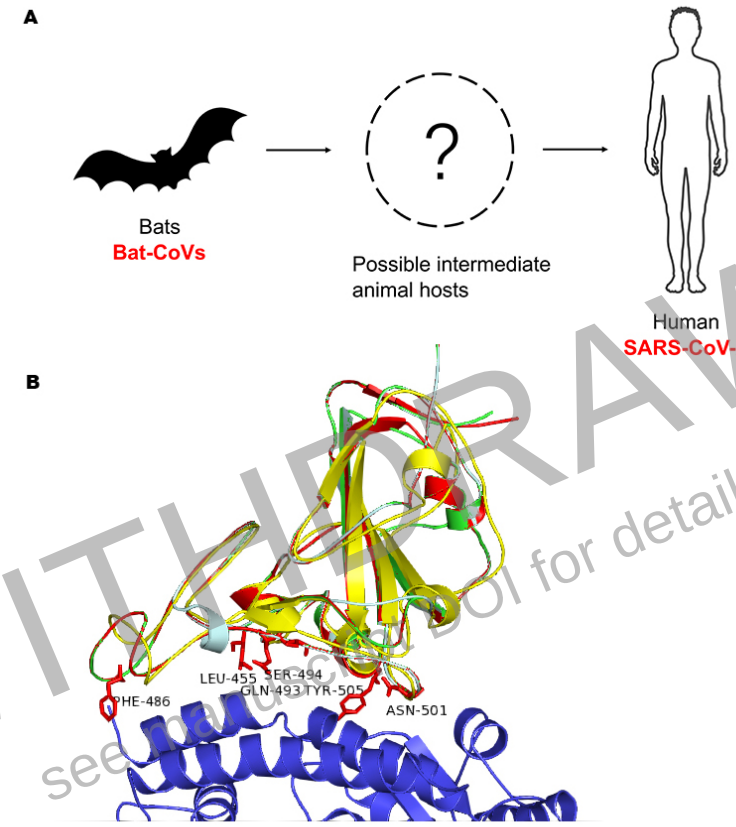
Conflict of interest statement

None of the authors declared conflict of interests.

Author Contributions

Y.Y conceived the project, designed the experiments, and wrote the manuscript. S.W. and C.L. participated in the design of the experiments. P.C, and Z.Z., designed and performed most of the simulations, whereas Z.G., Q.C. and S.C. performed bioinformatics analysis. Q.L. and Z.S. participated in the coordination of the experiments. Y.Y. provided funds. S.C. helped in the writing of the manuscript.

Figures



CoVs	Binding free energy of CoV-RBDs with hACE2 (kcal/mol)	Sequence identity AA (%)
SARS-CoV-2	-904.76	100
SARS-CoV	-785.71	81
RaTG13-CoV	-682.61	97
Bat-CoV	-284.76	76

Figure 1. Another animal may act as an intermediate host between bats and human for the transmission of SARS-CoV-2. **A). The possible transmission route of SARS-CoV-2; B). The overlay and the binding free energies of four CoV-RBDs with human ACE2 computed by MM/PBSA method**. The RBD of SARS-CoV-2 was depicted in red cartoon, the RBD of SARS-CoV was depicted in green cartoon, the RaTG13-CoV was depicted in yellow cartoon and the Bat-CoV was depicted in cyan cartoon. Human ACE2 was displayed in blue cartoon. Six hotspot residues including Phe486, Leu455, Ser494, Gln493, Tyr505 and Asn501 in the RBD of SARS-CoV-2 were displayed in stick model. The values of sequence identify was computed in the amino acid level by the BLAST tools.

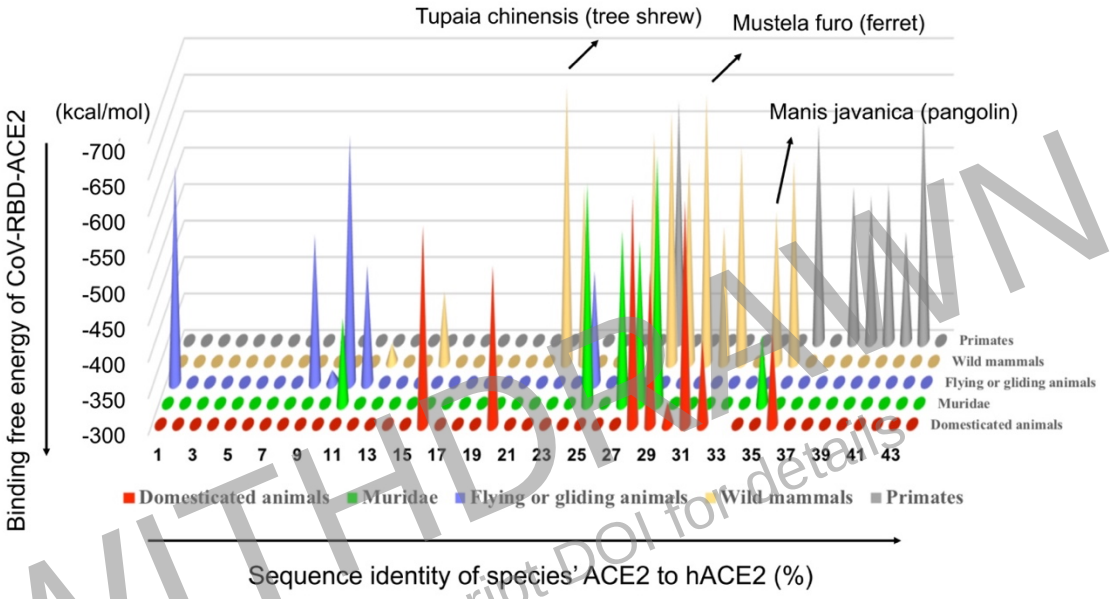


Figure 2. Three-dimensional conical plot of binding free energies of RaTG13-CoV-RBD complexed with ACE2 proteins of 47 representing animal species. The top two best binding free energies for tree shrew (-675.71 kcal/mol) and ferret (-663.57 kcal/mol) were highlighted with pangolin as a comparison (-504.76 kcal/mol). The horizonal axis represents the sequence identity of ACE2 protein of animal species with human ACE2 in amino acid level computed by BLAST tool. The vertical axis represents five animal groups classified by domesticated animals (red color), muridae (green color), flying or gliding animals (blue color), other wild mammals (sand color) and primates (grey color). The Z axis represents the binding free energies of RaTG13-CoV-RBD complexed with ACE2 proteins computed by MM/PBSA method.

References

1. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 2020;579(7798):270-273.
2. Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet.* 2020;395(10224):565-574.
3. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science.* 2020.
4. Chen F, Liu H, Sun H, et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. *Phys Chem Chem Phys.* 2016;18(32):22129-22139.
5. Ge XY, Li JL, Yang XL, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature.* 2013;503(7477):535-538.
6. Demogines A, Farzan M, Sawyer SL. Evidence for ACE2-utilizing coronaviruses (CoVs) related to severe acute respiratory syndrome CoV in bats. *J Virol.* 2012;86(11):6350-6353.