Preventing future zoonosis: SARS-CoV-2 mutations enhance human-animal cross-transmission

JunJie Wee[1], and Guo-Wei Wei[1,2,3*]

[1]*Department of Mathematics, Michigan State University, East Lansing, MI, 48824, USA*
[2]*Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA*
[3]*Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA*

Introduction

The COVID-19 pandemic has driven substantial evolution of the SARS-CoV-2 virus, yielding subvariants that exhibit enhanced infectiousness in humans. However, this adaptive advantage may not universally extend to zoonotic transmission. In this work, we hypothesize that viral adaptations favoring animal hosts do not necessarily correlate with increased human infectivity. In addition, we consider the potential for gain-of-function mutations that could facilitate the virus's rapid evolution in humans following adaptation in animal hosts. Specifically, we identify the SARS-CoV-2 receptor-binding domain (RBD) mutations that enhance human–animal cross-transmission. To this end, we construct a multitask deep learning model, MT-TopLap trained on multiple deep mutational scanning datasets, to accurately predict the binding free energy changes upon mutation for the RBD to ACE2 of various species, including humans, cats, bats, deer, and hamsters. By analyzing these changes, we identified key RBD mutations such as Q498H in SARS-CoV-2 and R493K in the BA.2 variant that are likely to increase the potential for human–animal cross-transmission.

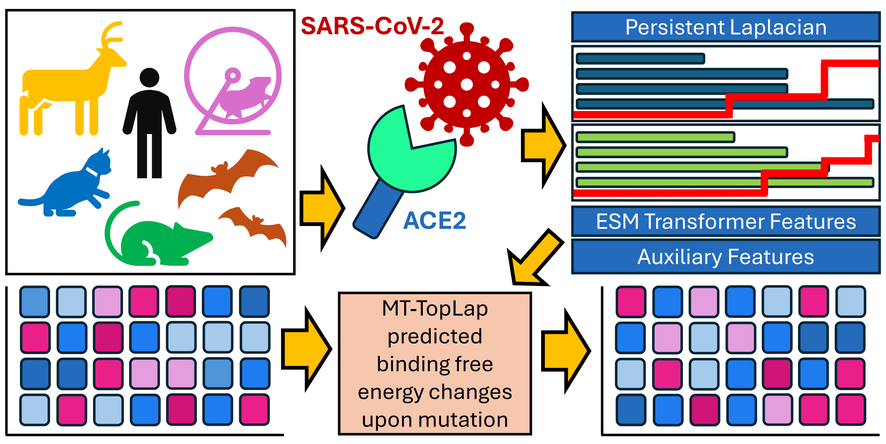

Figure: Illustration of the workflow for predicting deep mutational scanning of SARS-CoV-2 S protein RBD-ACE2 complexes. The input is a SARS-CoV-2 S protein RBD-ACE2 complex, and the output is the predicted DMS from MT-TopLap (the heatmap on the bottom right). By partitioning the SARS-CoV-2 RBD-ACE2 complex into different element/site-specific subsets, simplicial complexes are constructed using the Vietoris–Rips/Alpha complex for filtration. Harmonic and non-harmonic spectra are computed from the persistent Laplacians. The statistics of harmonic and non-harmonic spectra are used to generate feature vectors for deep learning.

Persistent Laplacians

Among all features used in MT-TopLap prediction, the application of Persistent Laplacians (PL) in characterizing proteins played an important role in the success of MT-TopLap in predicting protein-protein complex binding free energy changes upon mutation. PL is used to extract both harmonic and non-harmonic spectral properties essential for the characterization of protein-protein interactions, or in particular, a SARS-CoV-2 RBD-ACE2 complex. The harmonic spectra of PL can fully recover the Betti numbers of persistent homology (PH), a crucial tool in the early days of topological data analysis (TDA). The non-harmonic spectral properties capture the additional evolution of the homotopic shape of the PLs during the filtration, which is missing from PH.

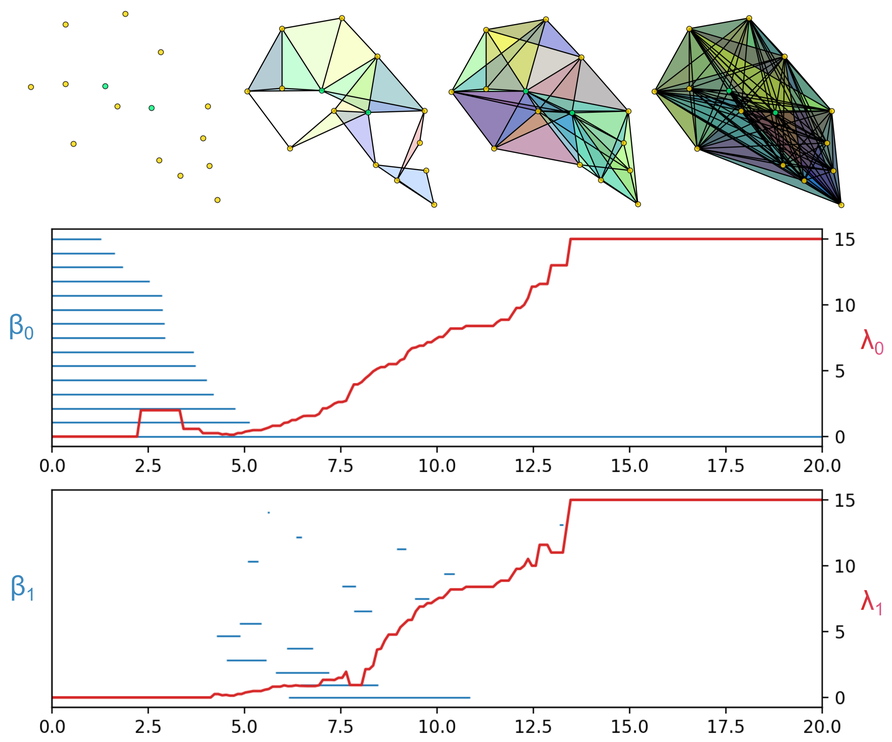

Figure: Illustration of harmonic spectra and non-harmonic spectra of persistent Laplacians from a mutation neighborhood of residue 498 in SARS-CoV-2 S protein PDBID: 6M0J.

- **Chain Group:** For a simplicial complex K, a k-th chain c_k is the formal sum of k-simplicies in K, i.e. $c_k = \sum_i \alpha_i \sigma_i^k$. Let kth chain group C_k be the free Abelian group generated by oriented k-simplices.
- **Bounday operator:** A boundary operator $\partial_k : C_k \to C_{k-1}$ defined on a k-th chain c_k is

$$\partial_k c_k = \sum_{i=0}^{k} \alpha_i \partial_k \sigma_i^k. \tag{1}$$

- **Topological Laplacian**

$$\Delta_k := \partial_{k+1}\partial_{k+1}^* + \partial_k^* \partial_k. \tag{2}$$

- **Combinatorial Laplacian matrix:**

$$\mathbf{L}_k = \mathbf{B}_{k+1}\mathbf{B}_{k+1}^\top + \mathbf{B}_k^\top \mathbf{B}_k. \tag{3}$$

- **Filtration process:**

$$\varnothing = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_m = K. \tag{4}$$

- **Persistent combinatorial Laplacian:** Based on the filtration process, a sequence of combinatorial Laplacian matrices is generated:

$$\mathbf{L}_k^0, \mathbf{L}_k^1, \mathbf{L}_k^2, \mathbf{L}_k^3 \cdots, \mathbf{L}_k^m \tag{5}$$

where $\mathbf{L}_k^t = \mathbf{L}_k(K_t)$ for $0 \le t \le m$.

Reference

Wee, J., Chen, J., & Wei, G. W. (2024). Preventing future zoonosis: SARS-CoV-2 mutations enhance human–animal cross-transmission. Computers in Biology and Medicine, 182, 109101.

BA.2 RBD mutations enhance human-hamster cross-transmission

The predictions indicate that the BA.2 RBD is more adaptive to humans than to hamsters, with highest BFE change for RBD-human ACE2 is R493V with 2.64 kcal/mol, while for RBD-hamster ACE2, it is R493K at 0.655 kcal/mol. However, RBD-human ACE2 and RBD-hamster ACE2 share a highly positive mutation R493K, with both mutations showing BFE changes above 0.5 kcal/mol, potentially enhancing human-hamster cross-transmission.

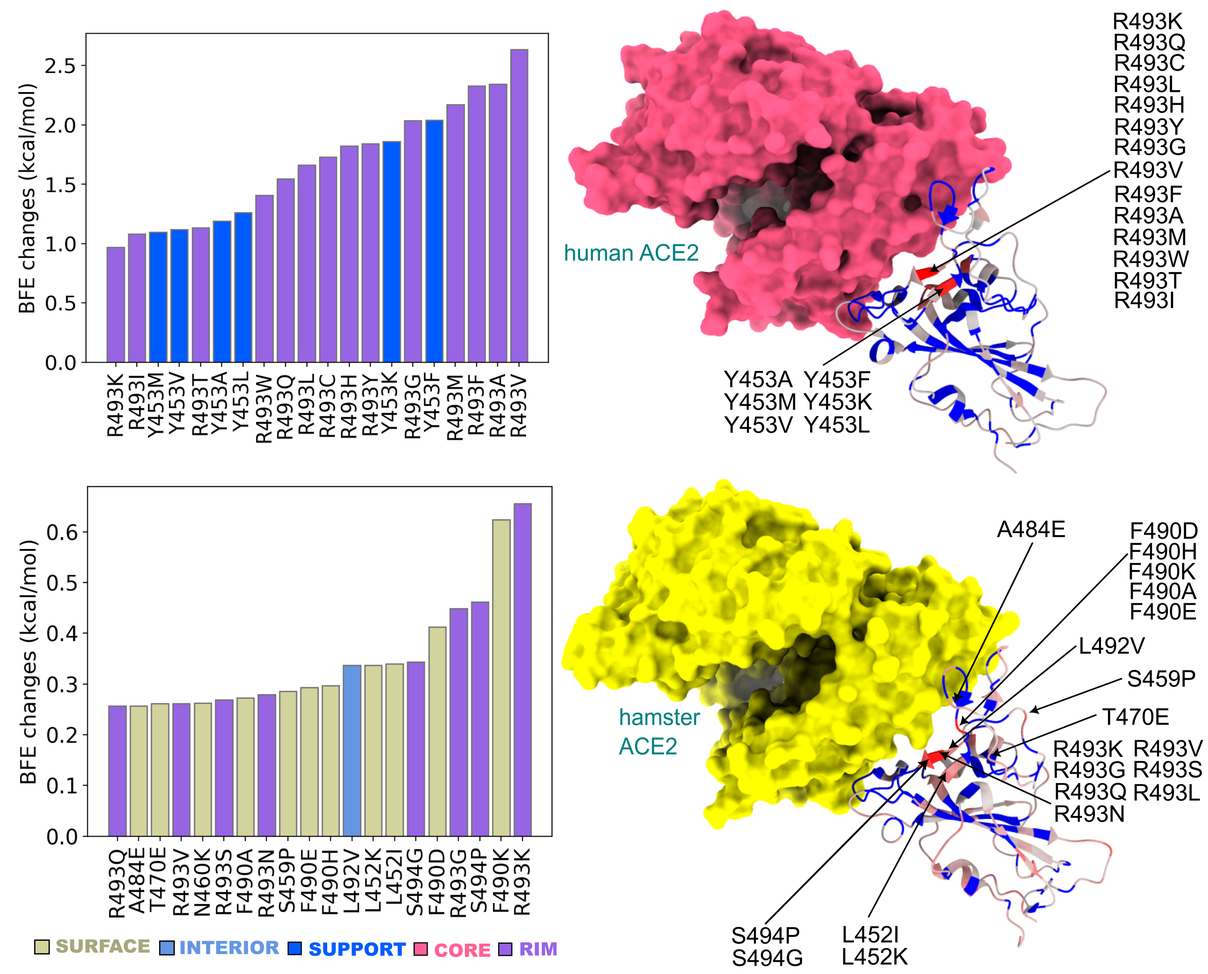

Figure: Top 20 binding free energy (BFE) changes induced by BA.2 S protein RBD mutations that are more than 0.5 kcal/mol for RBD-human ACE2 complex (top), RBD-hamster ACE2 complex (bottom).

Analysis of RBM by structural regions

In our analysis, we categorize receptor binding motif (RBM) residues into surface, interior regions and three binding interface regions: support, rim, and core. These regions significantly influence binding free energy (BFE) changes upon mutation. The same RBM residue from the SARS-CoV-2 S protein may be categorized differently based on its interaction with ACE2 in various RBD-ACE2 complexes due to variations in relative accessible surface area (rASA) calculations. This dynamic classification of RBD residues highlights the adaptability of the virus when interacting with ACE2 from different species.

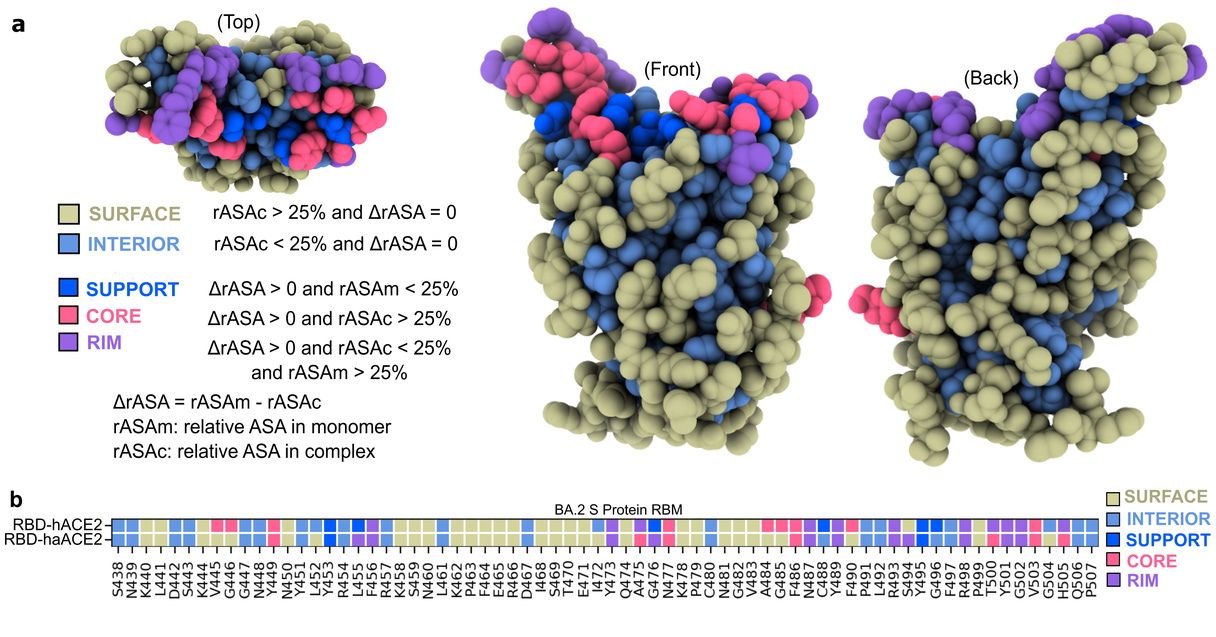

Figure: (a) SARS-CoV-2 S protein RBD residues colored by structural regions. (b) Categorization of BA.2 S protein RBM by structural regions for BA.2 RBD-human ACE2 (RBD-hACE2) and BA.2 RBD-hamster ACE2 (RBD-haACE2).

Conclusion

The hypothesis that viral adaptation favoring animal hosts does not necessarily correlate with increased human infectivity is supported by the distinct evolutionary pressures experienced by the virus in different hosts. Adaptations that enhance viral binding and replication in animal hosts may not directly translate to increased infectivity in humans due to the unique structural and biochemical differences in human ACE2 receptors compared to those of other species. For instance, mutations that increase binding affinity to animal ACE2 may not confer the same advantage for binding to human ACE2, as these mutations might disrupt the precise molecular interactions required for efficient human infection. Therefore, while a virus may adapt to an animal host environment, these adaptations might not necessarily facilitate or enhance transmission to humans, highlighting the complexity of zoonotic transmission pathways and the multifaceted nature of viral evolution across different species. This complexity underscores the importance of evaluating viral mutations within the context of each specific host environment to accurately predict their potential impact on human infectivity.

Acknowledgement

This work was supported in part by NIH grants R01GM126189, R01AI164266, and R01AI146210, NSF grants DMS-2052983, DMS-1761320, and IIS-1900473, NASA grant 80NSSC21M0023, MSU Foundation, Bristol-Myers Squibb 65109, and Pfizer.

https://junjiewee.github.io/ https://users.math.msu.edu/users/weig/ weejunji@msu.edu weig@msu.edu