International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.127 ISSN: 2581-4621

Higher binding affinity in mutant G614 SARS-CoV2 may explain the higher prevalence of Anosmia in European countries

Priyanka S[1], Harshita S.S[1], Geethamani P.R[1], Pooja M[1], Brinda. L[1], Manohar G M[2*]

[1]Department of PG studies in Biotechnology, Nrupathunga University, Bangalore -01.
[2]halliLabs, Ragihalli, Anekal Tq, Bangalore-83
Orcid ID : https://orcid.org/0009-0004-6806-7081

Abstract:
Anosmia is a condition where there is a loss of smell in people affected by covid 19. The prevalence of anosmia differs according to the geographical location. The prevalence of chemosensory dysfunction in Europeans was found to be three times higher than that in Asians. We observed that these geographical locations are also characterised by the occurrence of unique mutant strains G614 found in European countries and D614 wild-type strain of the SARS-CoV2 in Asian countries. Assuming that there may be a correlation between the mutation and the infection efficiency that in turn might affect the anosmia prevalence we analysed the binding affinity (with its receptor ACE2) and other features of the mutation. Using FoldX we built a mutant model G614 and found that the stability of the protein is not much affected. However, the difference in interaction energy with the receptor ACE2 showed a difference of around 10 kcal/mol. FoldX analyze complex analysis showed that the interaction energy is stronger in mutant strain hence high binding affinity. Pymol visualisation showed that the mutation was on the surface of the Spike protein. A phylogenetic tree of available ACE2 human sequences showed that Anosmia prevalence is not correlated with the variation in ACE2 sequences of different ethnic/geographical samples. Our study finds that the binding affinity of the mutant strain is high and therefore a possibility of explaining the correlation observed with the mutant strain and prevalence of anosmia.
Keywords: SARS-CoV2, MutationD614G, Spike-ACE2 Complex, Binding affinity, anosmia.

1. INTRODUCTION

In December 2019, the COVID-19 outbreak originated in Wuhan, China and rapidly spread across the world, causing a global pandemic (1). To date, there have been more than 210 million confirmed cases of COVID-19 and 3.9 million reported deaths as a result of the virus(2). Reduction of smell and taste is now recognized as one of the cardinal symptoms of COVID-19. The deficit appears to be most often transient, with a regaining of smell and taste after several days to weeks or months, but the anosmia differs from other virus-associated deficits in its sudden onset and its rapid recovery (3). In addition to the host response, variation in viral strain could contribute to disease severity and spread efficiency. At the level of the virus, there are differences between virus strains. The D614G mutation, especially, is responsible for the enhanced cell entry or binding of the SARS-CoV-2 spike protein to the ACE2 protein. D614G mutation in SARS-CoV-2, a non-synonymous mutation resulting in a replacement of aspartic acid with glycine at position 614 of the virus's spike protein [4]. Early in the pandemic, the D614 variant was predominant before it was rapidly replaced by the G614 variant [4,5]. At the host level, Angiotensin-converting enzyme 2 (ACE2), the entry protein to which the viral spike protein binds, has different variants, resulting in differing virus binding affinities and the frequency of such ACE2 variants is known to differ between ethnicities (5).
We observed a correlation in the geographical locations characterised by the occurrence of unique mutant strains G614 found in European countries and D614 wild-type strain of the SARS-CoV2 in

DOI:10.46647/ijetms.2023.v07i03.127 ISSN: 2581-4621

Asian countries (6) and also the prevalence of anosmia was more in European countries at 54.8% compared to Asian countries 17.7% (7).
We formulated a hypothesis " difference in the anosmia prevalence across geographical locations can be explained through the increased or decreased affinity of the Spike protein with the ACE2 receptor." We created a mutant Spike-ACE2 Complex G614 using FoldX software. Stability energy values and binding affinity energy values were studied. Binding affinity is increased in the mutant spike protein. There is a change in the electrical nature surrounding the mutation. Then mutation G614 increases the binding affinity thereby enabling the efficient entry of the virus. But whether this is connected with increased anosmia prevalence needs to be established.

2. MATERIALS AND METHODS

2.1 Retrieval of Sequences

The protein ACE2 Precursor isoform 1 (NP_001358344) sequence of humans was retrieved from protein database of NCBI (www.ncbi.nlm.nih.gov). Using the retrieved human sequence BLASTp was performed in a non-redundant database (blast.ncbi.nlm.nih.gov/Blast.cgi) to find other human sequences of different ethnicity.

2.2 Sequence comparison by sequence alignment

Sequential comparison of human protein ACE2 sequences of different geographical origins was performed by aligning the sequences in the Molecular evolutionary analysis tool (Phylogenetic and molecular evolutionary analyses) were conducted using MEGA version 11 (8).

2.3 Structure visualizations

UCSF Chimera X, UCSF Chimera and Pymol were used for Structural visualizations.

2.4 Retrieval of the protein

Crystal structure of SARS-CoV-2 spike receptor-binding domain bound with ACE2- 6M0J was retrieved from the Protein data bank (https://www.rcsb.org)

2.5 Fold X

The mutations of protein 6M0J were studied in Fold X at three levels, Repair, Build model and Analyze Complex. The following commands were used for Repair PDB, Sequence Detail, BuildModel and AnalyseComplex.
FoldX --command=RepairPDB --pdb=Rp. pdb,
FoldX --command=SequenceDetail --pdb=SD.pdb
FoldX --command=BuildModel --pdb=BM.pdb --mutant-file=individual_list.txt
FoldX--command=AnalyseComplex--pdb=AC.pdb--analyseComplexChains=A,B --complexWithDNA=true

2.6 Phylogenetic tree

Maximum likelihood tree with the Test of phylogeny Bootstrap method, 100 replications; Nucleotide substitution method, general time reversible model was employed for constructing phylogenetic trees using the multiple sequence alignment of human models of different ethnicity in MEGA Software.

3. RESULTS

Anosmia is a condition where there is loss of smell in people affected by covid 19. It is also considered as a symptom of infection. The prevalence of anosmia differs according to their geographical locations. A lower percentage of patients presenting with anosmia was found in Asian countries compared to European countries (9). The prevalence of chemosensory dysfunction in Europeans was found to be three times higher than that in Asians. The prevalence of anosmia among Asians was 17.7%, as compared to 54.8% among Europeans (10). D614 wild-type strain of the SARS-CoV2 having Aspartic acid located at 614 position is found in Asian countries while G614 mutant strain with Glycine at 614 position is found in European countries (11).
We hypothesised that the mutation G614 may confer a higher binding affinity for the spike protein to the human ACE2 receptor than the wild-type D614 and a higher binding affinity might result in a higher prevalence of anosmia (12).We wanted to explain the difference in the anosmia prevalence

DOI:10.46647/ijetms.2023.v07i03.127 ISSN: 2581-4621

across geographical locations through the increased or decreased affinity of the Spike protein with the receptor.

In order to find whether there is any change in the stability between the wild type (D614) Spike ACE2 complex (PDB ID:6M0J) was retrieved from PDB and the mutant counterpart protein complex (G614) was created using FoldX. We created a mutant protein containing Glycine at the location in place of Aspartic acid in the wild type. Stability energy values and binding affinity energy values were obtained (see below).

Before any protein can be imported for further analysis in FoldX the protein has to be repaired using the Repair pdb command.

Repair PDB: Identifies those residues which have bad torsion angles, VanderWaals' clashes, or total energy, and repairs them. 6M0J protein retrieved from Protein Data Bank was repaired using the repair command [FoldX --command=RepairPDB --pdb=RP .pdb] of FoldX.

Sequence list of the repaired pdb (6M0J) was obtained. Because sometimes in the Foldx program, the numbering of the amino acid sequence is different from what is allotted in the PDB structure. Therefore the numbers of locations of the amino acids in the sequence and the numbers that the Foldx can be mismatched. Using the sequence list command(FoldX --command=SequenceDetail --pdb=SD.pdb), Foldx numbers the amino acids and these numbers can be used for further analysis. Using the Foldx tool we mutated the amino acid Aspartic acid to Glycine to know the difference in energy levels.

Build model - The build model introduces mutations and optimises the structure of new protein variants. The energy function of FoldX is only able to calculate the energy difference in an accurate manner between the wild type and the mutated one. The build model command was used to create a mutant of G614 from the wild-type spike protein D614. The command used was FoldX --command=BuildModel --pdb=BM.pdb --mutant-file=individual_list.txt

Free energy changes happening between the wild type and the mutated protein can be predicted which gives an idea about the stability of the protein (Whether mutation confers more stability or less stability).The lower the Free energy, the higher the stability.

The wild-type D614 energy was -146.048 kcal /mol and that of the mutant G614 was found to be -146.04 kcal/mol.The energy change between the wild type and mutant type is very minimal (-0.008kcal/mol) , we can assume that the stability of the protein is not very much affected by the mutation.

Analyze Complex -Determines the interaction energy between two molecules or two groups of molecules. The way it operates is by unfolding the selected targets and determining the stability of the remaining molecules and then subtracting the sum of the individual energies from the global energy.

Analyze complexes were used to find out the difference between the affinity of the two complexes - 1. Wild type with ACE2 and 2. Mutant with ACE2.

The interaction energy of wild type(D614) is -6.57206 kcal/mol and mutant type (G614) -16.8897 kcal/mol.

After the interaction between wild-type strain D614 and mutant-type strain G614, the energy is lowered (-16.889 kcal/mol) compared to the wild type (-6.57206 kcal/mol) . So the interaction of the G614 spike protein of mutant type with the ACE2 has more negative interaction energy. Hence, the interaction is stronger (Since it leads to the lowest energy level).

We also checked whether a mutation is present on the surface using a Pymol tool (13). Mutant protein, G614 created using the Foldx tool, was used for analysis . We hypothesised that if the mutated residue is on the surface, then the stability is less affected. Wild type strain has Aspartic acid which has a longer side chain , whereas mutant type strain containing Glycine does not have a side chain at all. These amino acids are present on the surface of protein since chains do not make contact with others. If the mutation is present on the surface but still may be involved in binding the receptor (ACE2). So even though the mutation all by itself is not affecting the stability of the protein (from our Build

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.127 ISSN: 2581-4621

model results) it may be affecting the binding of the ACE2 protein and therefore indirectly affecting the Anosmia prevalence).

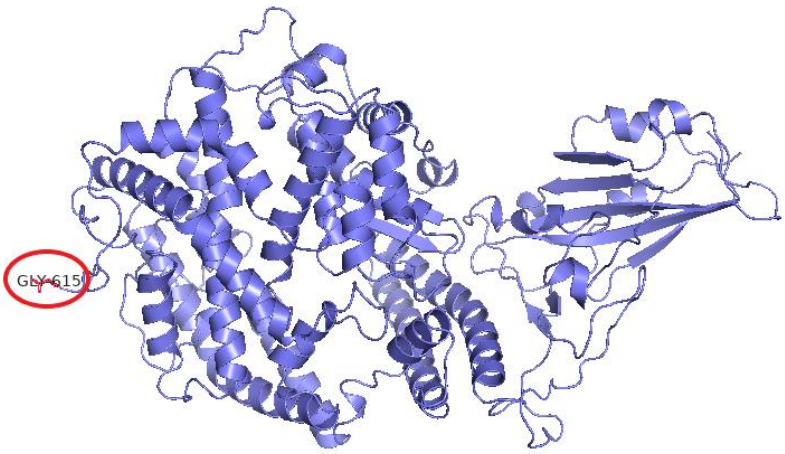
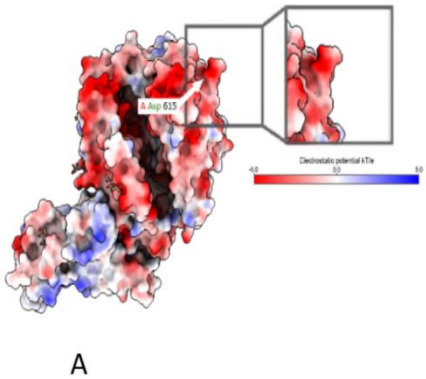


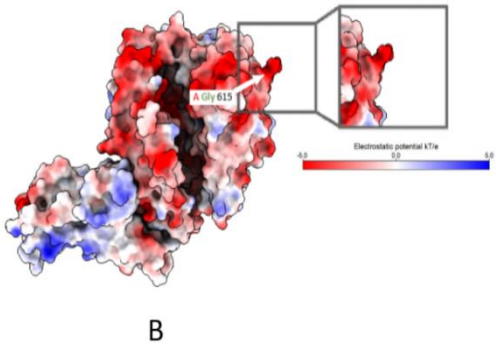
Fig -1 : Spike ACE2 complex (6M0J) showing G614 mutation present on the surface shown in red colour inside red circle

We wanted to see whether the mutation is creating any changes in the electrostatic nature of the environment. Electrostatic potential at the region (APBS electrostatics) was studied using the ChimeraX tool (14)for both wild and mutant types.



A

Fig 2.A Wild type showing negative electrostatic potential



B

Fig 2.B Mutant type showing negative electrostatic potential

Red colour indicates a negative charge and the blue colour indicates positive charge; There is a small change in the magnitude of the negative charge from wild to mutant type at 614 positions (where aspartic acid is mutated to glycine) which was indicated by a change in colour from light to dark red.

Multiple sequence alignment

Multiple sequence alignment was performed for the ACE2 sequences to find out whether any variation in sequences across different ethnicities is correlated with the prevalence of Anosmia.

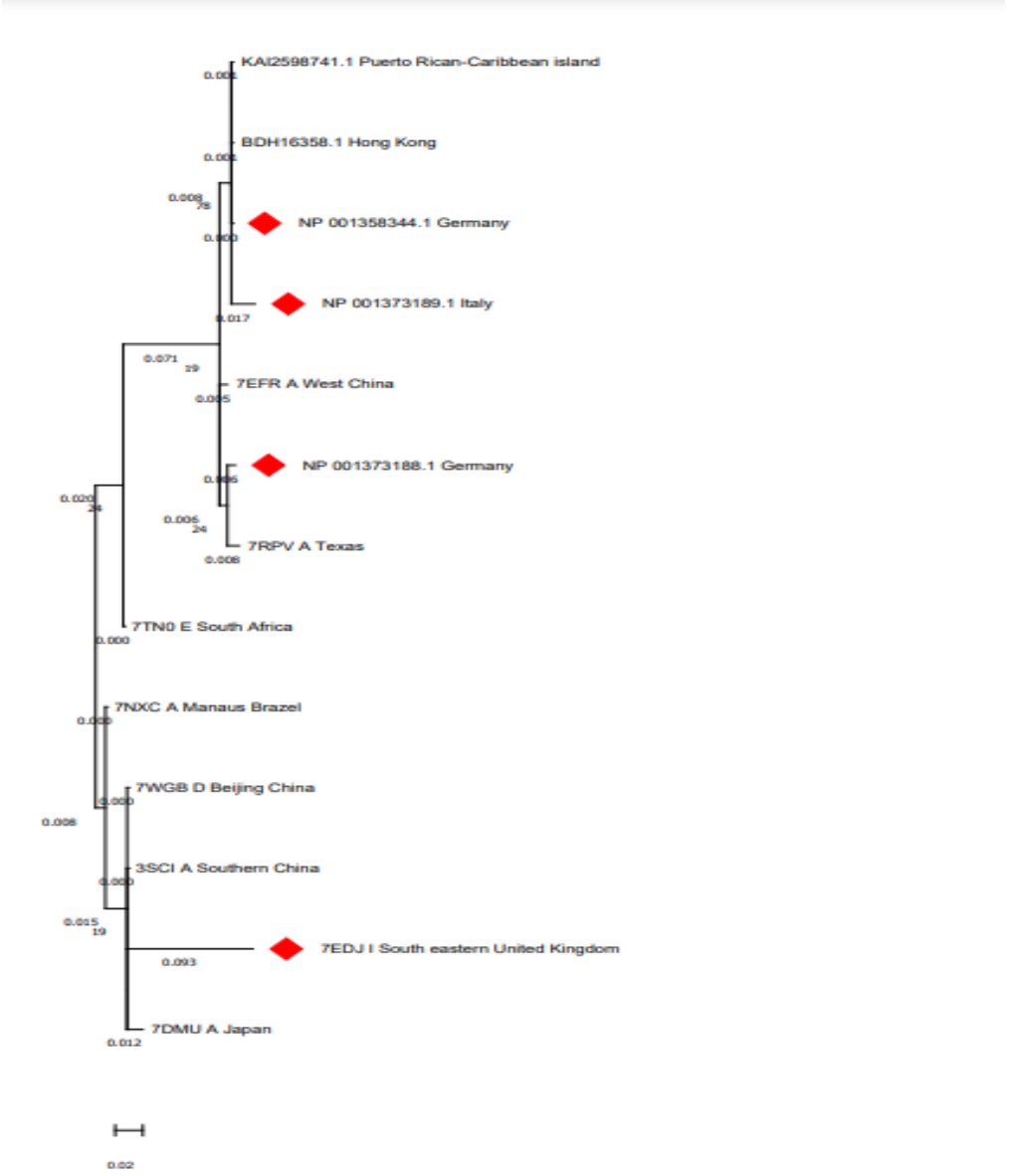
The protein ACE2 Precursor isoform 1 (NP_001358344) was retrieved from the NCBI database. Blastp was performed for the above sequence, by applying parameters such as percentage and query coverage 70-100 % and E-valve was set to 0.00. Blast results produced 41 sequences out of which 13 protein sequences were taken considering different ethnicities around the world. Namely,

1) Germany (NP_001358344.1) 2) PuertoRican-uk(KAI 2598741.1)

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.127 ISSN: 2581-4621

3) HongKong-china(BDH16358.1) 4)Germany (NP_001373188.1)
5) Southeastern- UK (7EDJ_I) 6) Brazil (7NXC_A)
7) Southern china (3SCI_A) 8) South Africa (7TN0_E)
9) West china (7EFR_A) 10) Beijing-China (7WGB_D)
11) Japan (7DMU_A) 12) Texas(7RPV_A) 13)Italy (NP_001373189.1)

These sequences were used to construct Maximum likelihood phylogenetic trees with bootstrap replications set to 100. The General Reverse Transcriptase model (rtREV) model (15) method and anosmia prevalence were mapped on the same.



Red diamonds represent sequences from European countries and the rest are from Asian, African and American countries.
Each clade has a different ACE2 (Eg ; China and other countries). ACE2 is present as clusters (mixed up).

DOI:10.46647/ijetms.2023.v07i03.127 ISSN: 2581-4621

From the phylogenetic tree, we can infer that Anosmia prevalence is not correlated with the variation in ACE2 sequences of different ethnic/regional samples. So the responsibility for the degree of Anosmia prevalence could be a function of the mutations in the Spike protein.

DISCUSSION:

Anosmia varies according to the geographical locations and it has been found that Europeans are more affected than Asians. In Asians we can see the D614 strain and in Europeans G614. We hypothesized that the prevalence of anosmia might be dependent on a particular mutant in the same geographical locations. In our research work, we have found that the stability of the complex is not affected by mutation
(- 146.04kcal/mol) from the wild type (-146.048 kcal/mol). But there is a substantial difference in the binding affinity between the Spike and the receptor when the spike protein is mutated into G614
(-16.8897kcal/mol) from D614(-6.57206kcal/mol). It could be possible that the change in the binding affinity could be causing the difference in anosmia prevalence. The mutation is however found on the surface of the protein and there is a slight change in the electrical nature of the motif surrounding the mutation (more negative). An emerging field of interest and a major hypothesis is that differences in the prevalence of anosmia may be caused by variations in the binding affinity of the ACE2 receptor for the virus and therefore may regulate infectivity and spreading of the virus. We also checked whether variations in ACE2 in different locations can explain the difference in the prevalence of Anosmia. But we did not find any evidence to that purpose in our phylogenetic tree constructed using the available ACE2 sequences. However, an exhaustive collection of ACE2 from various different parts of the world can give a clearer picture. Differences between populations in this aspect need to be verified by future studies, but if confirmed, they would have considerable implications for defining which populations are most vulnerable to COVID-19 infection and how best and most effectively we can manage the pandemic by a customized approach.
We have concentrated on one particular mutation which is the defining mutation of the strains found in their respective geographical locations. However, there will be a number of other mutations that differ between the wild type (D614) and the mutant strain(G614) chosen. A comprehensive analysis in the same way could define the possible correlation between the prevalence of anosmia and other mutations.

CONCLUSION: Our study finds that the binding affinity of the mutant strain is high. However, more work needs to be done to establish all the downstream effects that could lead to a difference in Anosmia prevalence. Such studies can shed more light on other mutations and possible linkages to specific covid symptoms.

ACKNOWLEDGEMENT:
We deeply appreciate the help and mentoring from **Mr Srikanth Lingappa,** a Research Scholar from the University of Bristol, UK and also the resource person from halliLabs for regular inputs for this research.

REFERENCES

1. Veronese S.; Sbarbati A. Chemosensory Systems in COVID-19: Evolution of Scientific Research. ACS Chem. Neurosci. 2021, 12, 813–824. 10.1021/acschemneuro.0c00788.
2. Najafloo, Raziyeh et al. “Mechanism of Anosmia Caused by Symptoms of COVID-19 and Emerging Treatments.” ACS chemical neuroscience vol. 12,20 (2021): 3795-3805. doi:10.1021/acschemneuro.1c00477
3. Printza, Athanasia et al. “Smell and Taste Loss Recovery Time in COVID-19 Patients and Disease Severity.” Journal of clinical medicine vol. 10,5 966. 2 Mar. 2021, doi:10.3390/jcm10050966

DOI:10.46647/ijetms.2023.v07i03.127 ISSN: 2581-4621

4. Korber B, Fischer WM, Gnanakaran S, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell. 2020;182(4):812–827. doi: 10.1016/j.cell.2020.06.043.
5. Butowt R, Bilinska K, von Bartheld CS. Chemosensory dysfunction in COVID-19: integration of genetic and epidemiological data points to D614G spike protein variant as a contributing factor. ACS Chem Neurosci. 2020;11(20):3180–3184. doi: 10.1021/acschemneuro.0c00596.
6. Zhang L, Jackson CB, Mou H, Ojha A, Rangarajan ES, Izard T, Farzan M, Choe H. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. bioRxiv [Preprint]. 2020 Jun 12:2020.06.12.148726. doi: 10.1101/2020.06.12.148726
7. von Bartheld, Christopher S et al. "Prevalence of Chemosensory Dysfunction in COVID-19 Patients: A Systematic Review and Meta-analysis Reveals Significant Ethnic Differences." medRxiv : the preprint server for health sciences 2020.06.15.20132134. 17 Jun.
8. Koichiro Tamura, Glen Stecher, and Sudhir Kumar (2021) MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Molecular Biology and Evolution* 38:3022-3027
9. Kumar AA, Lee SWY, Lock C and Keong NC (2021) Geographical Variations in Host Predisposition to COVID-19 Related Anosmia, Ageusia, and Neurological Syndromes. *Front. Med.* 8:661359. doi: 10.3389/fmed.2021.661359
10. Lizhou Zhang,Cody B Jackson, Huihui Mou,Amrita Ojha,Erumbi S Rangarajan,Tina Izard,Michael Farzan,and Hyeryun Choe.The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity.Version 1. bioRxiv. Preprint. 2020 Jun 12. doi: 10.1101/2020.06.12.148726
11. Elisa Frullanti, Chiara Fallerini, ,Ottavia Spiga,Anna Maria Pinto.The D614G Virus Mutation Enhances Anosmia in COVID-19 Patients: Evidence from a Systematic Review and Meta-analysis of Studies from South Asia.ACS Chem Neurosci 2021 Oct 6; 12(19): 3535–3549. Published online 2021 Sep 17. doi: 10.1021/acschemneuro.1c00542
12. von Bartheld CS, Hagen MM, Butowt R. The D614G Virus Mutation Enhances Anosmia in COVID-19 Patients: Evidence from a Systematic Review and Meta-analysis of Studies from South Asia. ACS Chem Neurosci. 2021 Oct 6;12(19):3535-3549. doi: 10.1021/acschemneuro.1c00542.
13. Schrödinger, L., & DeLano, W. (2020). PyMOL. Retrieved from http://www.pymol.org/pymol.
14. Pettersen, Eric F et al. "UCSF ChimeraX: Structure visualization for researchers, educators, and developers." Protein science : a publication of the Protein Society vol. 30,1 (2021): 70-82. doi:10.1002/pro.3943.
15. Dimmic MW, Rest JS, Mindell DP, Goldstein RA. rtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny. J Mol Evol. 2002 Jul;55(1):65-73. doi: 10.1007/s00239-001-2304-y.

SUPPLEMENTARY SECTION

https://docs.google.com/document/d/1usH9SWJBh0s0p1xD2Ca9NDjVf0AoNwu55Z9iVohlU4g/edit?usp=sharing