CONTACT: Yugmee Yogesh Gidiya yugmee13@gmail.com

Molecular Docking – Study Of Interaction between SARS COV-2 protein and Ligand (ACE II) for Vaccine Development with Phylogenetic Analysis and Homology Modelling

Yugmee Gidiya, Dr. Rahna. K. Rathnan
JSS Private School , Dubai

Abstract

The outbreak of a pneumonia of unknown origin in Wuhan, Hubei Province in late December 2019, caused by the SARS-CoV 2 has caused deaths world-wide. The novel coronavirus SARS-CoV-2 entry into host cells is mediated by its spike glycoprotein (S-glycoprotein), and the human angiotensin converting enzyme 2 (hACE2)has been identified as a cellular receptor. In this review, we also summarize the current knowledge about the epidemiology, phylogenesis, and molecular diagnostics of SARS-CoV-2. Phylogenetic analysis is essential to understand viral evolution, and for vaccine development strategies and therapies. Highly sensitive and specific diagnostic assays are key to case identification, contact tracing, identification of the animal source, and implementation of control measures.

Keywords: Docking; Ligand; Receptor; Vaccines; Coronavirus; SARS-CoV-2, ACE II , phylogenetic analysis, genetic sequence, multiple alignment, genome, mutation

Objectives and Introduction

- To study the evolution of SARS-CoV-2, as an emerging coronavirus by phylogenetic analysis and homology modelling.
- To study the interactions between the S Glycoprotein and the Human ACE II receptor in order to provide reliable alternatives to the existing vaccines.

(1)

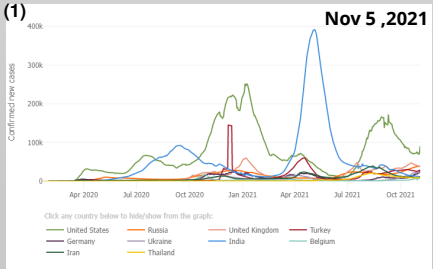


Fig 1: Outbreak evolution for the current most affected countries

Methodology

MULTIPLE SEQUNECE ALIGNMENT AND PHYLOGENETIC ANALYSIS

- Identification of DNA sequences of 10 different strains of SARS-CoV-2 in FASTA format from NCBI BLAST.
- Execute Multiple Sequence Alignment using CLUSTALw web service as the alignment tool.
- Derive pairwise and multiple sequence alignment.
- Construct a phylogenetic tree for the various strains – determine their evolutionary relationship with each other.
- This research follows the methodology of secondary data analysis or archival study. It involved the use of existing data, such as medical journals and related research for the purposes of prior study and insight into the thesis.

HOMOLOGY MODELLING

- SWISS-MODEL is a structural bioinformatics web-server dedicated to modelling of protein 3D structures.
- Retrieve the nucleotide sequence for the surface glycoprotein from NCBI Protein in FASTA format
- Plugging the sequence into the SWISS Model tool to obtain homology model of the protein.

MOLECULAR DOCKING

- Docking is a computational technique aimed at the prediction of the most favorable ligand–target spatial configuration and an estimate of the corresponding complex free energy.
- Identification and preparation of optimal ligand and receptor molecule configurations.
- Visualization of structures using docking software and studying interactions between the selected molecules.
- Quantitative evaluation of the free energy of binding between ACE2 and the active receptor binding domain of the SARS-CoV-2 spike protein for providing the thermodynamics of virus–receptor recognition.

Results & Inference

(2)

```
CLUSTAL 2.1 Multiple Sequence Alignments


Sequence type explicitly set to DNA
Sequence format is Pearson
Sequence 1: NC_045512.2 29903 bp
Sequence 2: MT671817.1  29901 bp
Sequence 3: MZ043010.1  29898 bp
Sequence 4: MZ824654.1  29903 bp
Sequence 5: MT039890.1   1470 bp
Sequence 6: MT066175.1   1470 bp
Sequence 7: MT066176.1   1470 bp
Sequence 8: LC528232.1   1470 bp
Sequence 9: LC528233.1   1470 bp
Sequence 10: MT007544.1   1470 bp
Start of Pairwise alignments
Aligning...
```

Fig 2: Results for Multiple Sequence Alignment. SARS-CoV-2 is a slow mutating virus , with only few base pair changes in different strains.

(3)

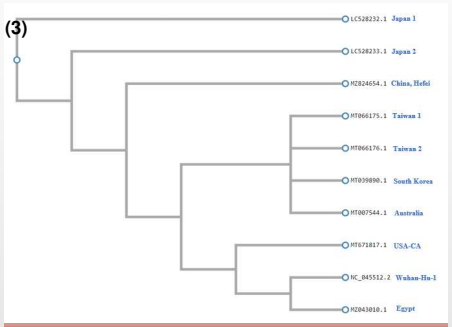
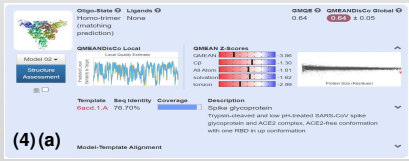


Fig 3: Phylogenetic Tree(Cladogram) showing different SARS CoV2 Strains . A mosaic pattern of transmission is observed indicating that it represents the result of travel of infected humans among different countries



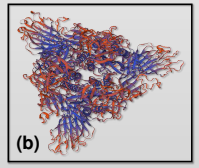


Fig 4(b: Cartoon Model of (S) Surface glycoprotein: SARS-CoV-2. It allows medical scientists to develop antibodies targeting the protein at it binding site to the host cell proteins.

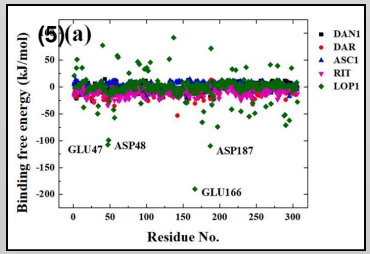
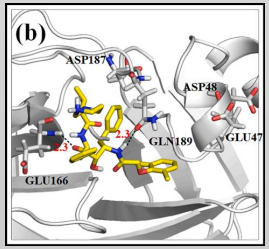


Fig 5(a): The bind free energy between all protein residues and selected drugs.
5(b): Interaction details between negatively charged residues and positively charged lopinavir (LOP1 system). The data in red is the interaction distance (Å).



Conclusion

The results can help clinicians prescribe therapies that are known to target certain genetic variants or reduce the person's risk of developing the disease or condition. Genomic analysis has paved the path to identifying potential medicines and cures for the pandemic while the vaccines continue to undergo trials in several regions of the world.

References

https://coronavirus.jhu.edu/data/new-cases

LIU, J., ZHAI, Y., LIANG, L., ZHU, D., ZHAO, Q. AND QIU, Y.
Molecular modeling evaluation of the binding effect of five protease inhibitors to COVID-19 main protease
In-text: (Liu et al., 2021)
Your Bibliography: Liu, J., Zhai, Y., Liang, L., Zhu, D., Zhao, Q. and Qiu, Y., 2021. Molecular modeling evaluation of the binding effect of five protease inhibitors to COVID-19 main protease. *Chemical Physics*, [online] 542, p.111080. Available at: <https://doi.org/10.1016/j.chemphys.2020.111080.>.