Life Science Informatics Publications

Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences

Journal Home page http://www.rjlbpcs.com/

RJLBPCS

ISSN 2454-6348

Original Research Article **DOI: 10.26479/2020.0605.06**

INSILICO ANALYSIS OF MUTATIONAL VARIANTS OF SARS-CoV-2 RDRP PROTEIN

Deepak Kumar Jha[1], Niti Yashvardhini[2]*

1. Department of Zoology, P. C. Vigyan College, Chapra 841301, India.
2. Department of Microbiology, Patna Women's College, Patna 800 001, India.

ABSTRACT: SARS-CoV-2 (Severe Acute Respiratory syndrome) is an etiolating agent of COVID-19 disease, creating an alarming situation globally at this moment. Considering the increasing number of infected patients, the World Health Organization has declared public health emergency on 11th march 2020. Currently effective antiviral therapies have not been approved to prevent SARS-CoV-2 infections; therefore, we urgently need deep understanding of its origin, evolutionary aspects and molecular biology approach that are prerequisite to curb this pandemic disease. Since this virus is a member of genetically diverse RNA viruses having inherent capability to show high rates of mutation during circulations. Multiple mutations in SARS-CoV-2 genome drive viral evolution and hence enabling them to easily evade the pre-existing immunity of host and acquire drug resistance properties. RdRp (RNA dependent RNA polymerase) is a major component of the RNA virus because of its role in the replication /transcription process and lack of host cell homolog which also serves as the suitable target for the viral drug development. The aim of our study was to compare the SARS-CoV-2 RdRp sequence obtained from first Wuhan virus with those of Indian SARS-Cov-2 isolates. We focused mainly on those mutations which occurred in the RdRp protein and used for subsequent analysis. Altogether our study showed 13 mutations in the Indian SARS-CoV-2 RdRp protein (nsp12), a main replicase, followed by its phylogenetic analysis and homology modeling.

Keywords: SARS-CoV-2, Pandemic, COVID-19, RNA dependent RNA polymerase, Mutation**.**

Article History: Received: September 08, 2020; Revised: September 22, 2020; Accepted: October 02, 2020.

Corresponding Author: Dr. Niti Yashvardhini* Ph.D.

Department of Microbiology, Patna Women's College, Patna 800 001, India.

Email Address: nitiyashvardhini@gmail.com

1.INTRODUCTION

SARS-CoV-2, identified in December 2019, emerged as a pandemic of global concern and the causal agent of serious illness designated as COVID-19. As of 21 September, 2020, the total number of confirmed cases of COVID-19, was 30,949,804 including 959,116 deaths, reported by WHO [1,2]. SARS-CoV-2 is a positive single stranded, enveloped, genetically diverse RNA virus that belongs to Coronaviridae family and nidovirales order which is further divided into alpha, beta, gamma and delta coronaviruses genera respectively [3,4]. They are considered as largest RNA virus among all the RNA viruses having 30 Kb of genome size approximately. The genome of SARS-CoV-2 consist of 14 sequences of ORF, encoding 29 proteins which include three surfaces glycoprotein such as S (spike), E (envelope), M (membrane), and one nucleocapsid (N) protein which is essential for the assembly of complete virion particle. Nucleocapsid (N) protein plays an important role in the packaging of RNA into virion particles [5]. SARS-CoV-2 Possess ORF1ab gene, which constitutes two ORFs such as ORF1a and ORF1b. The position of these genes ranging from 251- 21541 nucleotides. The ORF1ab of the SARS-CoV-2 encodes 16 nsps (non-structural proteins) at the consensus region [6,7]. Coronavirus genome encodes RNA-dependent RNA polymerase (nsp 12) and helicase (nsp 13) protein respectively. RNA viruses showing high rates of mutation because of poor fidelity of its RNA dependent RNA polymerase, consequently these viruses exhibit overall genomic variations leading to antigenic variability [8]. Therefore, the objective of the present work focused on the recurrent mutational analysis of RdRp sequence and prediction of associated protein models as well as phylogeny of Indian SARS-CoV-2 isolates that might be acting as a therapeutic intervention and opens new avenues for combating SARS-CoV-2 infections.

2. MATERIALS AND METHODS

2.1. Sequence Retrieval

The full length protein sequences of SARS-CoV-2 were downloaded from NCBI virus database submitted from India in the month of June, 2020. A total of 167 full length sequences were released from India and these were used for mutational analysis in this study. The first virus sequence reported from 'Wuhan wet sea food market area' (Accession number YP_009724389) [9] was used as a reference sequence.

2.2. Multiple Sequence Alignment

The full length protein sequences were aligned using CLUSTAL Omega online server, conducts alignment on the basis of HMM profiling [10]. Jalview an allignmet viewer was used to assess the aligned files and the similarities and differences of sequences were viewed. The mutations in the RdRp region were mainly documented for further analysis.

2.3. Phylogenetic analysis

To study the phylogenetic relationship among the mutated RdRp Indian SARS-CoV-2 isolates, MEGAX (Molecular Evolutionary Genetics Analysis) software was used. MEGAX created

Neighbor-joining phylogenetic tree with 1000 replicated bootstraps [11]. MEGAX first aligns sequences using MUSCLE programme and prepares phylogenetic tree.

2.4. Protein Modeling

The protein models were prepared for wild type as well as mutated RdRp protein sequences using Phyre software [12]. Phyre2 is user friendly structure predicting software which predicts models on the basis of results of homology modeling.

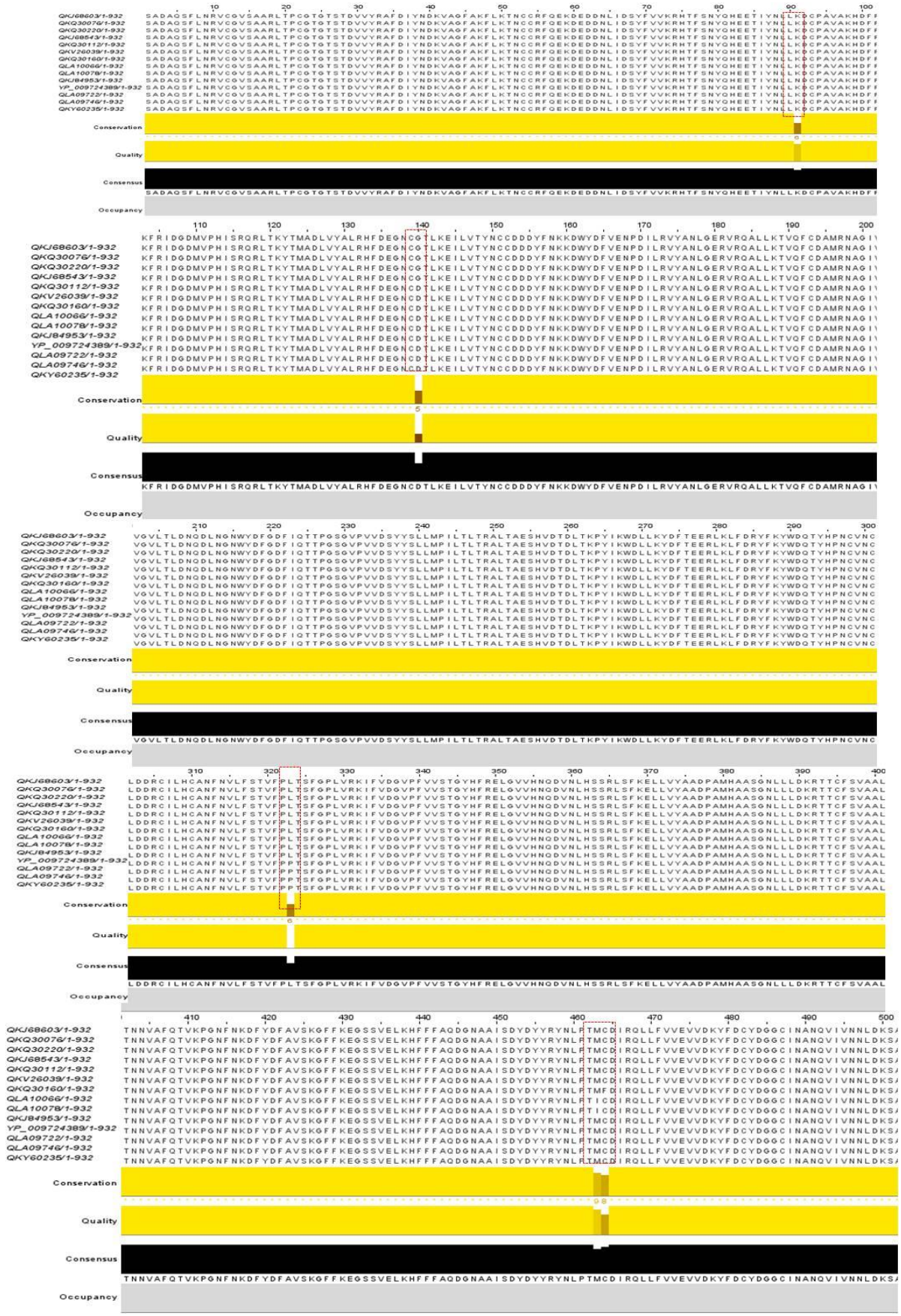
3. RESULTS AND DISCUSSION

3.1. Identification of mutation in RdRp proteins of Indian isolates

The SARS-CoV-2 protein sequences, 7096 amino acid long were retrieved from NCBI virus database hub. Out of 167 sequences released from India, 13 isolates showed mutation in the RdRp region as compared to Wuhan type SARS-CoV-2 RdRp after aligning the sequences using CUSTAL Omega platform (Figure 1). Out of these 13 isolates, one was a double mutant at sites D140G and V880I. Single point mutations were detected at sites K91N, M463I, C464F, L638F, V880I and T908I; other mutants showed variation at the same sites only (Table 1). Among the thirteen mutants only those which occurred at unique sites were further characterized to study the impact of mutation.

Table 1: Showing the details of RdRp sequences found mutated from the Wuhan SARS-CoV-2 sequence. Details include the accession number of the isolate position of mutation and the wild type and mutated amino acid sequence.

S. No.	Accession No.	Position of mutation	Wild sequence	Mutated sequence
1.	QKJ84953	91	K	N
2	QKJ68603	140	D	G
		880	V	I
3.	QLA10066	463	M	I
4.	QLA10078	463	M	I
5.	QKQ30160	464	C	F
6.	QLA09722	638	L	F
7.	QLA09746	638	L	F
8.	QKY60235	638	L	F
9.	QKQ30076	880	V	I
10.	QKQ30220	880	V	I
11.	QKJ68543	880	V	I
12.	QKQ30112	880	V	I
13.	QKV26039	908	T	I



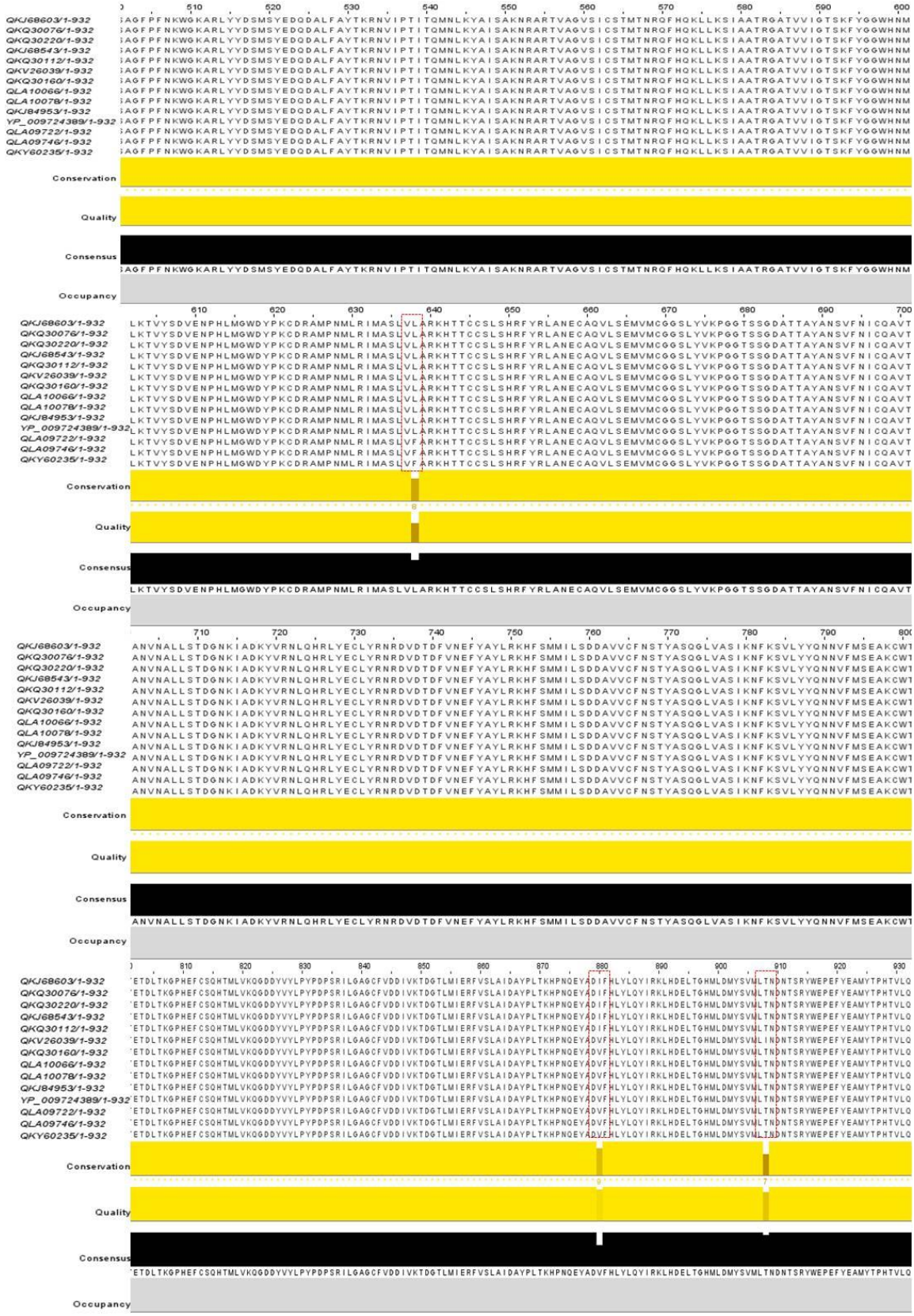


Figure 1: Multiple sequence alignment of RNA dependent RNA polymerase protein of Indian SARS-CoV-2 with that of Wuhan SARS-CoV-2. The mutated regions are marked with red boxes.

3.2. Phylogenetic analysis

To explore the phylogenetic relationships amongst the mutant and wild type SARS-CoV-2 isolates, a neighbor- joining phylogenetic tree was constructed using MEGAX. The analysis revealed the occurrence of Indian isolates on different clusters of the tree, showing the multi-faced nature of the virus (Figure 2). These mutations help these viruses to adapt new geographical regions.

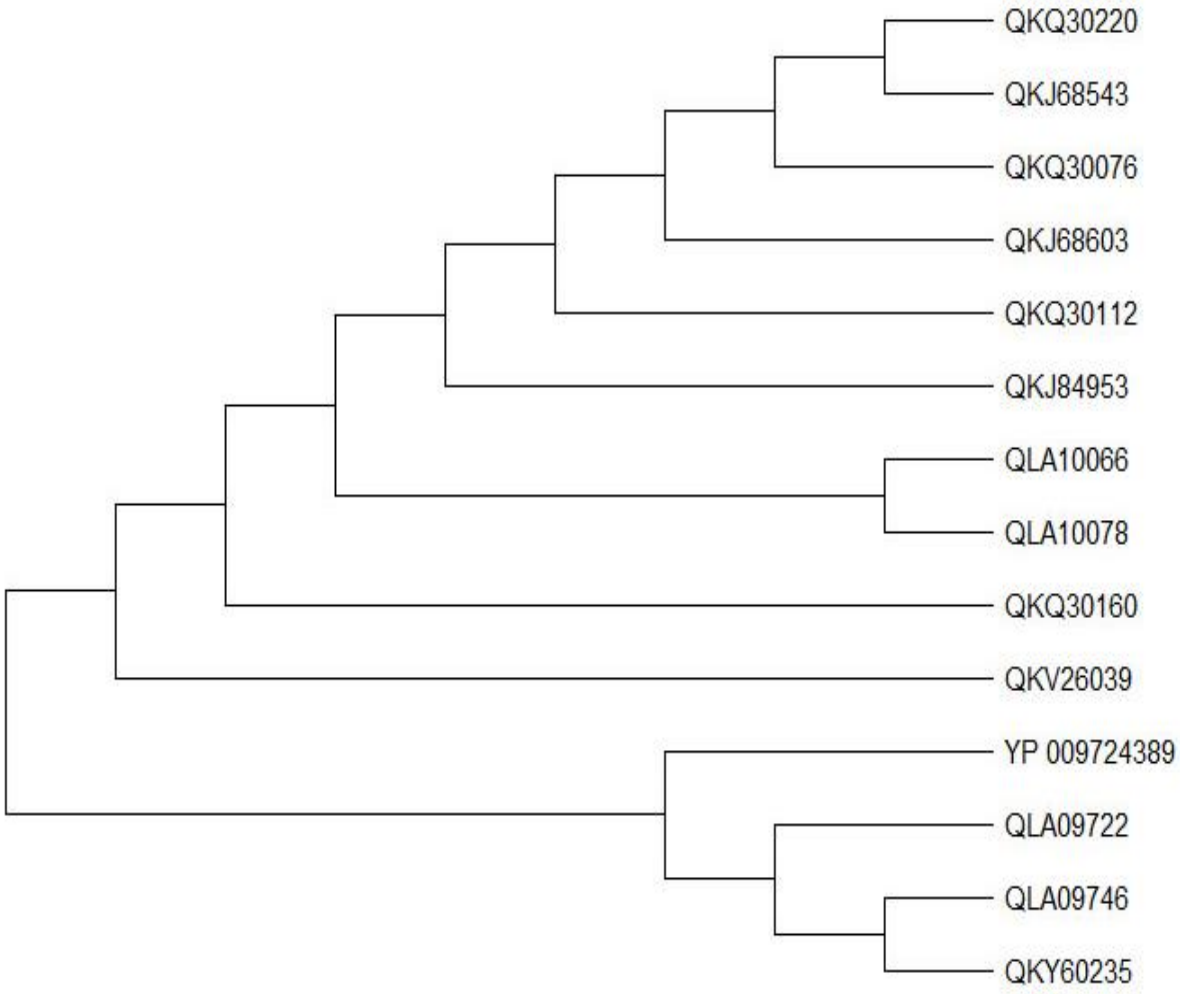


Figure 2: Phylogenetic tree of Indian SARS-CoV-2 and Wuhan SARS-CoV-2 isolates with reference to RdRp protein.

3.3. Homology modeling

We constructed 3D models of RdRp protein of SARS-CoV-2 using homology/ analogy modeling engine Phyre2. Protein models were built using Phyre2 online software to study the impact of mutation on the tertiary structure of RdRp proteins, as shown in figure 3. The models were constructed with nearly 90% accuracy in Phyre2. The difference occurring in the protein structure upon mutation in the RdRp protein is shown in the figure.

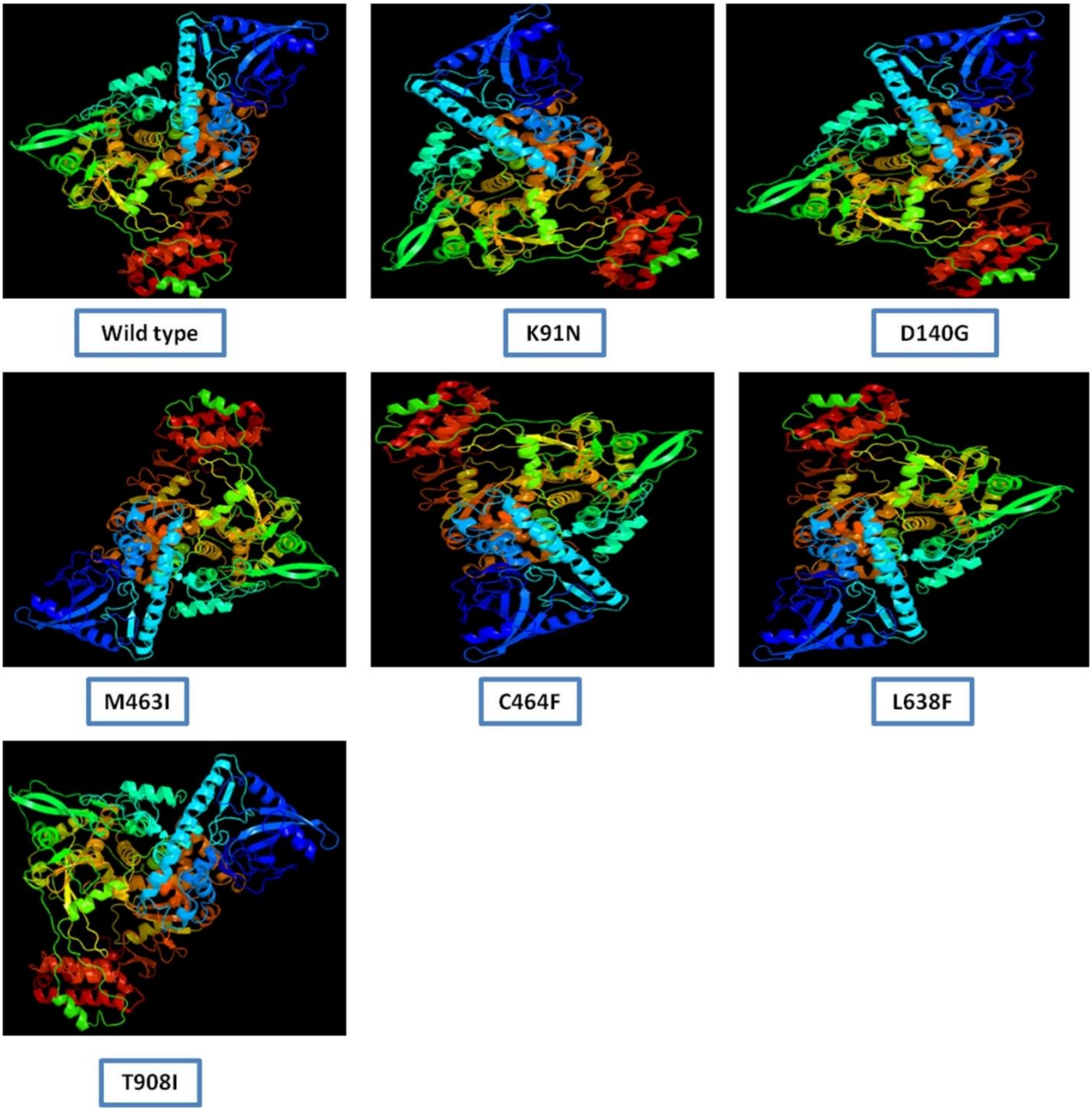


Figure 3: Homology modeling of mutant and wild type SARS-CoV-2 RdRp protein.

Recurrent mutation facilitates viral evolution as well as genome variability and therefore, enabling viruses to escape the pre-existing immunity of host and quickly acquire drug resistance. SARS-CoV-2 broke out from sea food market, Wuhan, China, further began to rapidly spread all over the globe [13-14]. Several factors are directly linked with the transmissibility of SARS-CoV-2, like population density, geographical distribution, health care system as well as environmental conditions [15-19]. Notably, mutations in the sequence of RdRp of Coronavirus affecting the fidelity rate of viral replication that subsequently affects the viral load (severity) and virulence properties of this virus. However, comprehensive research works are urgently required to assess the possible impact of mutations in the fidelity of RdRp protein [20]**.** Similarly, viruses having mutant RdRp sequences might be showing resistant properties towards anti viral therapeutic agents such as remdesivir which

is a most potent and commonly used repurposing drug for SARS-CoV-2 currently. These implications strongly favors RdRp mutations a promising vaccine candidate and drug target for pharmacological as well as epidemiological surveillances.

4. CONCLUSION

Present findings therefore clearly suggest that transmissibility of SARS-CoV-2 is greatly associated with the emergence of several novel mutations that spread at many new locations of the genome of this virus. The reports also provide crucial insight for functional validation to better understand the molecular means of differential disease severity and to develop specific control strategies to prevent COVID-19 infections.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The author confirms that the data supporting the findings of this research are available within the article.

FUNDING

None

ACKNOWLEDGEMENT

We thankfully acknowledge technical assistance received from Mr. Dinesh Kumar Dinkar.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interests.

REFERENCES

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet. 2020;395: 497–506.
2. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N. Engl. J. Med. 2020;382: 1199–1207.
3. Woo PCY, Wang M, Lau SKP, Xu H, Poon RWS, Guo R, Wong BHL, Gao K, Tsoi HW, Huang Y, et al. Comparative analysis of twelve genomes of three novel group 2c and group 2d coronaviruses reveals unique group and subgroup features. J. Virol. 2007; 81(4):1574–1585.
4. Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV).

Nucleic Acids Res. 2018;46(D1):D708–D717.

5. McBride R, Fielding BC. The role of severe acute respiratory syndrome (SARS) coronavirus accessory proteins in virus pathogenesis. Viruses 2012;4:2902–23.
6. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, O'Meara MJ, et al. 2020. A SARS-CoV-2-human protein-protein interaction map reveals drug targets and potential drug repurposing, bioRxiv 2020.
7. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. Nature 2020; 579(7798):265–269.
8. Sackman AM, McGee LW, Morrison AJ, Pierce J, Anisman J, Hamilton H, Sanderbeck S, Newman C, Rokyta DR. Mutation-driven parallel evolution during viral adaptation. Mol. Biol. Evol. 2017; 34(12):3243–3253.
9. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, Cao L, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science 2020; 368 (6492):779-782.
10. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nuc. Acids Res. 2019; 47(W1):W636–W641.
11. Sudhir K, Glen S, Michael L, Christina K, Koichiro T. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Bio and Evol. 2018; 35(6) :1547-1549.
12. Kelley L, Mezulis S, Yates C, Wass M, Sternberg M. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols 2015; 10 (6):845-858.
13. Domingo E. Viruses at the edge of adaptation. Virology. 2000.
14. Domingo E, Holland JJ. RNA virus mutations and ftness for survival. Annu Rev Microbiol. 1997.
15. Wang M, Jiang A, Gong L, Luo L, Guo W, Li C, Zheng J, et al. Temperature significant change COVID-19 transmission in 429 cities, medRxiv, 2020.
16. Lim SY, Osuna C, Lakritz J, Chen E, Yoon G, Taylor R, et al. Galidesivir, a direct-acting antiviral drug, Abrogates Viremia in Rhesus Macaques challenged with zika virus. Open Forum Infect Dis. 2017.
17. Agostini MK, Andre EL, Sims AC, Graham RL, Sheahan TP, Lu X, et al. Coronavirus susceptibility to the antiviral remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease. MBio. 2018.
18. Morgenstern B, Michaelis M, Baer PC, Doerr HW, Cinatl JJr. Ribavirin and interferon-β synergistically inhibit SARS-associated coronavirus replication in animal and human cell lines. Biochem. Biophys. Res. Commun. 2005.
19. Delang L, Froeyen M, Herdewijn P, Neyts J. Identification of a novel resistance mutation for benzimidazole inhibitors of the HCV RNA-dependent RNA polymerase. Antivir. Res. 2012; 93(1);30–38.
20. Duffy S. Why are RNA virus mutation rates so damn high? PLOS Biol 2018;16(8):e3000003.