Indonesian Journal of Biotechnology

VOLUME 30(2), 2025, 98-106 | RESEARCH ARTICLE

Mutations in RNA-dependent RNA polymerase could be major cause of high pandemic potential of SARS-CoV-2: An in silico study

Bhawna Sharma[1], Bennet Angel[1,*], Annette Angel[1], Vinod Joshi[1], Shareef Mohammed Buvvaji[1], Neha Singh[1], Khushbu Kumari[1], Aarya Chitransh[1], Poorna Khaneja[1], Devendra Kumar[2]

[1]Centre of Excellence in Virology & Immunology, Sharda University, Uttar Pradesh, India-201310
[2]Department of Respiratory Medicine, Sharda School of Medical Sciences & Research, Sharda Hospital, Greater Noida, U.P., India-201310
*Corresponding author: bennetangel@gmail.com

SUBMITTED 4 February 2025 **REVISED** 20 April 2025 **ACCEPTED** 2 May 2025

ABSTRACT Human coronaviruses (HCoVs) are responsible for mild common cold to severe pneumonia-like symptoms in infected individuals. The first HCoV was HCoV-229E, discovered in 1962 in the US, which causes moderate symptoms. Since then, HCoVs have evolved, leading to epidemics or the recent SARS-CoV-2 pandemic. The main objective of this study was to understand the modifications occurred and what led to the transition from mild to pandemic form. Of the viral proteins, the RNA-dependent RNA polymerase (RdRp) plays a crucial role in viral evolution, mutation, pathogenesis and transmission; this protein was therefore analyzed using in silico tools. We observed that RdRp has shown many mutations during its transition from mild to severe forms in HCoVs, which may have affected its enzymatic activity. The RdRp of HoV-229E and HCoV-NL63 showed 171 mutations, while SARS-CoV-2 showed the presence of 312. SARS-CoV-2 also showed a reduction in hydrophobic amino acid compared to the other HCoVs, consequently contributing to faster replication. Although mutations in the RdRp subdomains were found, yet five conserved regions was also presence among all the seven HCoVs; the finger and thumb subdomains had one conserved region, while the palm subdomains had three. Therefore, it can be inferred that on one hand the mutations reported in RdRp appeared to be the major cause of increased virulence leading to sporadic disease outbreaks, while on the other hand the presence of five conserved regions might prove to be potential targets for the development of new antiviral drugs.

KEYWORDS Evolution; Human coronaviruses; Mutations; NSP12; RdRp; SARS-CoV-2

1. Introduction

Human coronaviruses (HCoVs) were first discovered in the 1960s (Burrell et al. 2016). These belongs to the family Coronaviridae. Coronaviruses are among the largest RNA viruses with positive-sense and single-stranded RNA genome. It has the largest genome of about 27 to 30 kb among all RNA viruses (Liu et al. 2021). The Coronaviridae family has four genera viz. Alpha, Beta, Gamma, and Delta Coronaviruses. HCoV-229E and HCoV-NL63 belongs to the *Alphacoronavirus* while the other five (HCoV-OC43, SARS-CoV-1, HCoV-HKU1, MERS, and SARS-CoV-2) are *Betacoronaviruses* (Cicaloni et al. 2022; Miłek and Blicharz-Domańska 2018). Coronaviruses are common in birds and animals, however, human diseases are linked mainly to the Alpha and Beta coronaviruses (Cicaloni et al. 2022). HCoV-OC43 and HCoV-HKU1 were originated from rodents and the other human coronaviruses originated from bats (Tang et al. 2022). The severe form i.e. severe respiratory distress syndrome (SARS), was isolated between 2002–2003 (Guangdong, China) and it became epidemic (Ludwig and Zarbock 2020) leading to death of nearly 774 people (Haagmans and Osterhaus 2009). Later in December 2019, more severe form of coronavirus (i.e. SARS-CoV-2) appeared in China and became the cause of one of the worst pandemics of the world (Wei et al. 2019; Lamers and Haagmans 2022).

When we compare the genomic organization of all the 7 HCoVs, they show presence of two open reading frames (ORFs) i.e., ORF1a and ORF1b coding for the structural proteins (spike, membrane, envelope and nucleocapsid proteins) and non-structural proteins (NSP1 to NSP16), respectively (Zhao et al. 2020; Wong and Saier 2021). Some HCoVs also have hemagglutinin-esterase in their structure. Human coronaviruses such as HCoV-229E, HCoV-NL63, HCoV-OC43 and HCoV-HKU1 typically cause mild respiratory illnesses, such as common cold like symptoms, while SARS-CoV, MERS, and SARS-CoV-2 can result in severe or mild, both in humans and birds (www.medicalnewstoday.com).

Variations, possibly due to number of factors, (Supp. Figure 1) have been reported in the genomic organization of all seven HCoVs and we hypothesize that this might have affected the clinical outcome in various categories of patients. Therefore, we aimed to study one of the important proteins of HCoV that has undergone highest number of variations in all these 7 HCoVs. We thus focused on RdRp protein (RNA-dependent RNA polymerase), the enzyme involved in formation of replication complex/factory. It plays a key role in the intracellular replication of viral genome and formation of new virion copies. The structure of RdRp of SARS-CoV-2 has finger, palm and thumb subdomains giving it a right hand like shape. The subunits of NSP7 and NSP8 attaches to the thumb region, whereas NSP8 extra site attaches to the finger domain. NSP7 and NSP8 increases the binding of NSP12 with the primer RNA and also stimulates the function of the NSP12 protein (Sarma et al. 2022).

2. Materials and Methods

2.1. Genomic and proteomic studies on human coronaviruses

To understand the sequence and structural properties of RdRp among various HCoVs following methods were used i.e., the heat map generation of gene and protein sequences, structural studies using SWISS modelling, and finally observing the RdRp protein by its characteristics such as hydrophobicity, solubility, and isoelectric nature.

2.1.1 Heatmap generation of human coronaviruses RdRp

Protein and gene sequences of RdRp of all the 7 HCoVs were aligned using Clustal Omega in order to calculate pairwise percent identity. This was then converted to a distance matrix using 1-Percent Identity. This matrix was used, applying complete linkage clustering, to generate a heat map showing pairwise sequence divergence with a color gradient representing similarity. By understanding the metrics represented in the heat map and carefully analyzing the color gradients and clusters, the relationships and distinctive characteristics of the various coronaviruses can be identified (Babicki et al. 2016).

2.1.2 Sequence alignment of RdRp of all HCoVs

QIAGEN CLC Genomics Workbench 24.0.1 was used for sequence analysis and alignment. The software employs the "Needleman-Wunsch algorithm" by default for global alignment, with core parameters including the scoring scheme, gap penalties, substitution matrix, and alignment type. The RdRp sequences of seven HCoVs were aligned, and their genomics and proteomics analysis of the sequence were also done. The reference sequence numbers of RdRp protein included in the Multiple Sequence Alignment were retrieved from NCBI, i.e., AAP41036.1 for SARS-CoV-1, YP_009725307.1 for SARS-CoV-2, YP_009047223.1 for MERS, YP_009555260.1 for HCoV-OC43, YP_459941.1 for HKU1, NP_835352.1 for HCoV-229E, and YP_003766.2 for HCoV-NL63. The nucleotide sequence of the HCoVs was compared and checked for the mutation and similarity at the genomic level.

2.1.3 3D modelling

Further homology modeling was performed for obtaining the 3D models of RNA-dependent RNA polymerase (NSP12) for different HCoVs. The process involved selecting appropriate templates based on sequence similarity, optimizing sequence alignment, adjusting gap penalties and assessing the model quality.

The 8urb.1.A template was used for HCoV-229E and HCoV-NL63, with a Global Model Quality Estimate (GMQE) score of 0.87 for HCoV-NL63. HCoV-OC43 and HCoV-HKU1 were modeled using 7Krp.1.A, both achieving GMQE scores of 0.88. The 7uoe.1.A template was used for SARS-CoV-1 and MERS-CoV, resulting in GMQE scores of 0.91 and 0.89, respectively.

The FASTA sequences of the RdRp protein (NSP12) of 7 HCoVs are as follows i.e., HCoV-229E (NP_835352.1), HCoV-OC43 (YP_009555260.1), SARS-CoV-1 (YP_00992430.1), HCoV-NL63 (YP_003766.2), HCoV-HKU1 (YP_459941.1), MERS (YP_09047223.1), and SARS-CoV-2 (YP_009725307.1). The 3D structure of all proteins was modeled using Swiss modelling.

The SARS-CoV-2 RdRp (PDB ID: 6M71) was retrieved from the PDB databank. Due to the unavailability of other HCoVs RdRp 3D structures, we developed these using SWISS modeling (Protein Data Bank 1971). The core RdRp protein consists of three subdomains i.e., the finger (residue range L-366-A581, K621-G679), the palm (residue range T582-P620, T680-Q815), and the thumb subdomain (residue range H816-E920) (Vicenti et al. 2021). Using the PYMOL software (Schrödinger and DeLano 2020), the RdRp protein subdomains, finger, palm, and thumb were highlighted, and their 3D structure was generated.

2.1.4 Study of unique characteristics of RdRp protein of all HCoVs

The hydrophobicity plot of RdRp protein of coronaviruses was constructed and downloaded using the CLC Genomics Workbench 24.0.1 program (CLC QIAGEN, Germany). For calculating the hydrophobicity, the software used is a default Kyte and Doolittle scale. This scale is an effective approach for determining the hydrophobicity of protein sequences by assigning individual hydropathy scores to each amino acid. These scores are generally aggregated or averaged across the entire sequence to identify regions that are hydrophobic or hydrophilic. Regions with positive hydropathy scores are considered hydrophobic, while those with negative scores are classified as hydrophilic.

Solubility of RdRp protein of all HCoVs were obtained from the protein-sol software (Hebditch et al. 2017).

It is a user-friendly bioinformatics online server that uses protein sequence to calculate the predicted solubility compare with the solubility database. For calculating the solubility, the software employs a linear regression model incorporating sequence-derived features such as amino acid composition, hydropathy index (Kyte-Doolittle scale), isoelectric point, and net charge.

The isoelectric point (pI) of a protein is the pH at which the protein carries no net charge, and it is determined by the pKa values of the protein's ionizable groups. The net charge of the protein is calculated at different pH levels. Protein-Sol does not calculate the pI it instead uses external Bjellqvist method and Protparam or Biopython tool for calculation. Using this method isoelectric point (pI) of RdRp protein of all HCoVs were obtained from the protein-sol software (Hebditch et al. 2017).

2.2. Comparison of HCoVs sequences with similar viruses by Phylogenetic Tree construction

To understand the similarity among other viruses (other than coronaviruses), the phylogenetic analysis was also done. SARS-CoV-2 reference genome sequence (NC_45512.2) was taken for Pairwise genomic similarity analysis with viruses existing in the database, using the ViPTree v4.0 (Nishimura et al. 2017). This tool employs tBLASTx for sequence comparison and then translates nucleotide sequences into six reading frames (proteins) to identify conserved regions across viral genomes. The similarity score (SG) generated by tBLASTx ranges from 0 to 1, indicating the degree of genetic relation between viral sequences. Phylogenetic analysis was performed using the BIONJ (Bio Neighbor-Joining) algorithm, an improved version of the traditional Neighbor-Joining (NJ) method. BIONJ enhances tree accuracy by incorporating models that account for evolutionary rate variations among sequences. This approach provides a more precise representation of the evolutionary relationships within the dataset. The A,T,G,C nucleotide percentage was also calculated in all the HCoV genomes since it is known that Adenine and Thymine have higher mutation rate compared to cytosine and guanine (Braun et al. 2019).

3. Results and Discussion

3.1. Genomic and proteomic studies on human coronaviruses

3.1.1 Heatmap analysis of amino acid and nucleotide sequence of RdRp protein in human coronaviruses

Figure 1a-b indicates the obtained heatmap. The diagonal cells (from the top-left to the bottom-right) were mainly all blue, which indicated that 100% similarity was found among HCoVs. While the HCoV with more intense blue color may belong to the same evolutionary lineages, HCoV with light magenta and yellow color represent less similarity to each other.

HCoV-229E and HCoV-NL63 were similar as both belong to the *Alphacoronavirus* genus. HCoV-OC43 and HCoV-HKU1 shared similarities, they are included in *Betacoronaviruses* and both are originated from rodents. SARS-CoV-1 and SARS-CoV-2 were closely related, both belong to the same genus and came from the same origin i.e. bats. MERS-CoV, however, was more distinct from other human coronaviruses, as it belongs to a unique evolutionary lineage and originated in bats before crossing into humans through dromedary camels as shown in Figure 1.

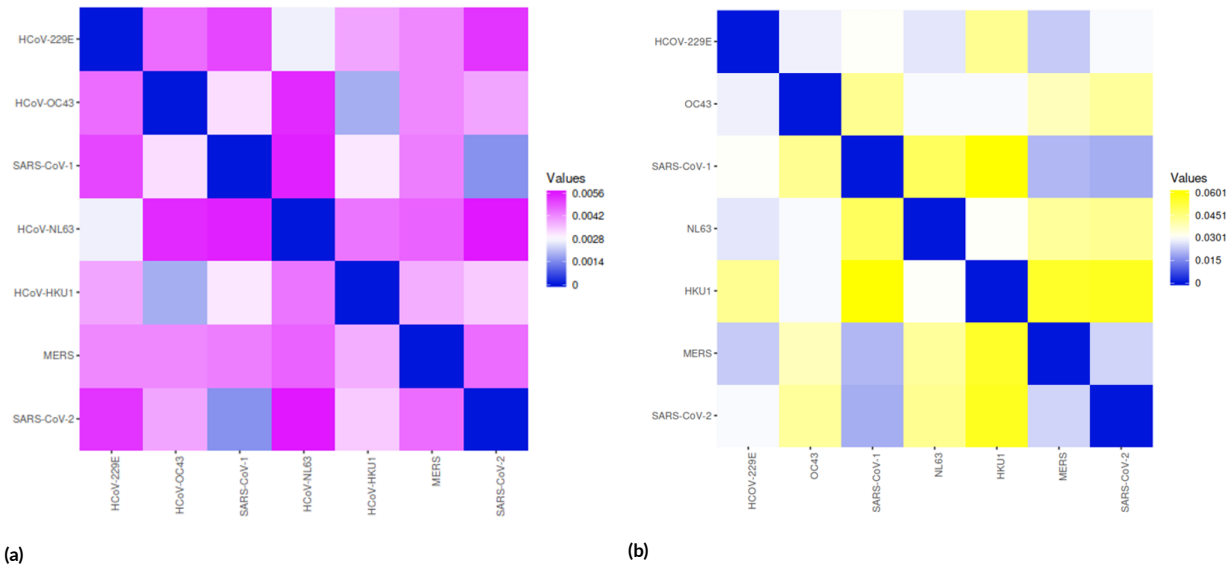

FIGURE 1 Heatmap represents the RdRp proteomic (a) and genomic (b) similarity among human coronaviruses http://www.heatmapper.ca/pairwise/.

3.1.2 Sequence alignment of RdRp of all HCoVs

During sequence alignment of RdRp proteins of HCoVs, it was found that the RdRp core protein has five conserved regions. Of which the palm subdomain had three conserved regions while the thumb subdomain had only one conserved region (in all the 7 HCoVs). The conserved region for finger subdomain was 621KCDRA625, palm subdomain was 614LMGWDYP620, 751KHFSMMILSDD761, 808GPHEFCSQ815), and thumb subdomain was 860VSLAIDAYPL869 (Figure 2 and Supp. Figure 2). The similarity of all 7 RdRp proteins can be seen as much as 49% in the finger, 52% in palm and 42.3% in thumb subdomains. The percentage of similarity of nucleotide and amino acid sequences of the RdRp showed in Supp. Table 1.

The mutations in the amino acids of RdRp of all 7

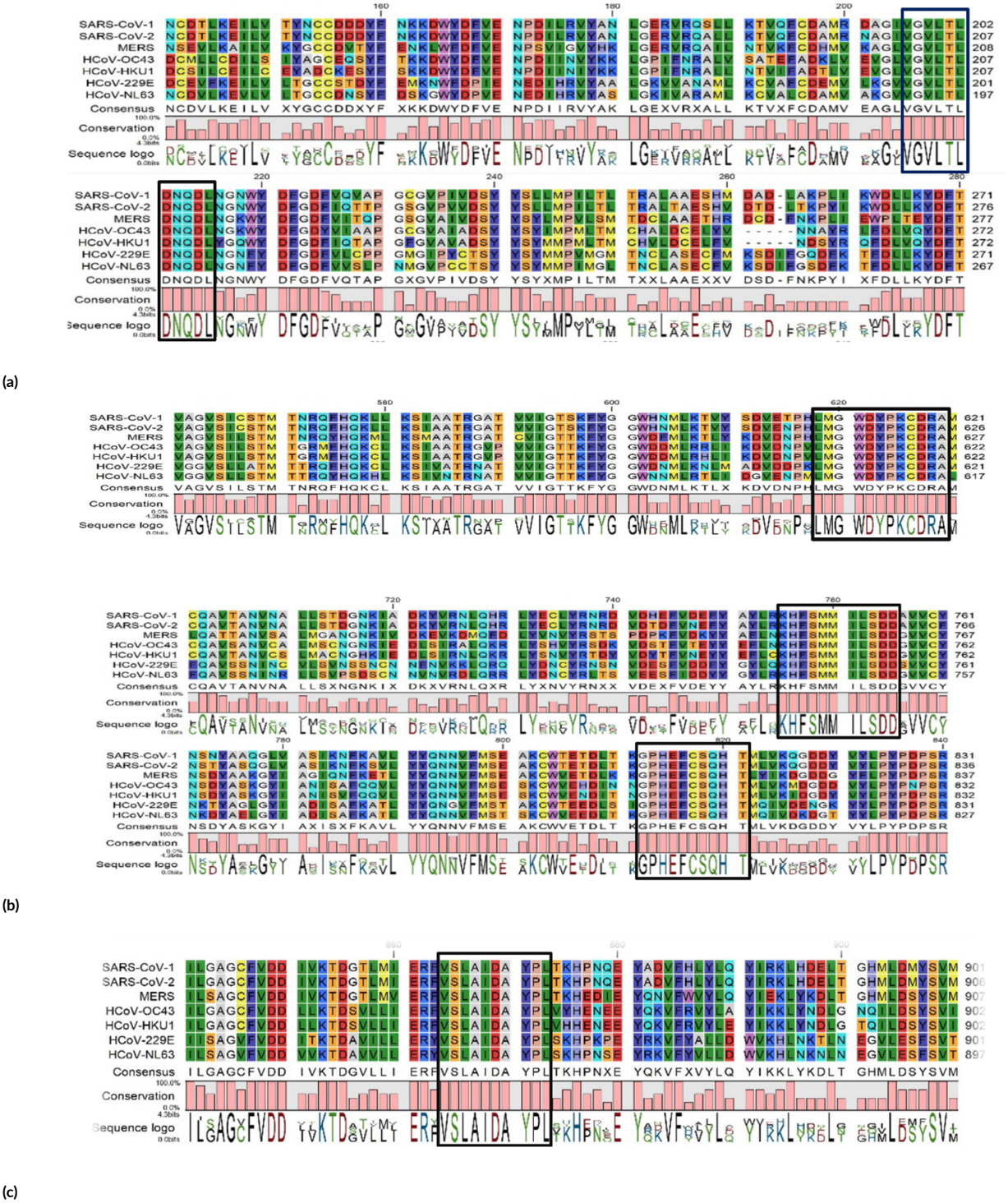

FIGURE 2 RdRp protein sequence alignment with five conserved regions marked in the boxes using the CLC Genomics workbench 24.0.1.

human coronaviruses were checked. Figure 3 and Supp. Figure 3 shows the major mutations with change and no change in chemical properties in the amino acids. The individual mutations in RdRp protein of HCoVs have been compiled in Supp. Table 2.

3.1.3 3D modelling

Supp. Figure 4 highlights the finger, palm and thumb subdomains residue range w.r.t SARS-CoV-2 RdRp (PDB ID: 6M71) structure using the PyMOL software.

3.1.4 Study of unique characteristics of RdRp Protein of all HCoVs

It was observed that SARS-CoV-2 showed decrease number in hydrophobic amino acids, which might contribute to greater solubility, flexibility and potentially a more efficient RdRp. It could enhance replication fidelity and speed, which could be the reason of increased transmissibility as compared to the other HCoVs (Table 1 and Supp. Figure 5). The 3D surface representation is shown in Supp. Figure 6.

Several coronaviruses i.e., HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1 are less pathogenic. They showed moderate solubility, which indicate stability but they had slower replication. MERS-CoV have lower solubility and SARS-CoV-2 have the lowest solubility supporting more efficient RNA binding and replication. The solubility of all HCoVs shown in the Supp. Figure 7.

All RdRp proteins showed a net positive charge at acidic pH (low pH) and shift to a net negative charge at basic pH (high pH). The isoelectric point of seasonal coronaviruses i.e. HCoV-229E, HCoV-OC43, HCoV-NL63 and HCoV-HKU1, ranges from 6.1 to 6.3, which was slightly lower compared to the one belong to other coronaviruses

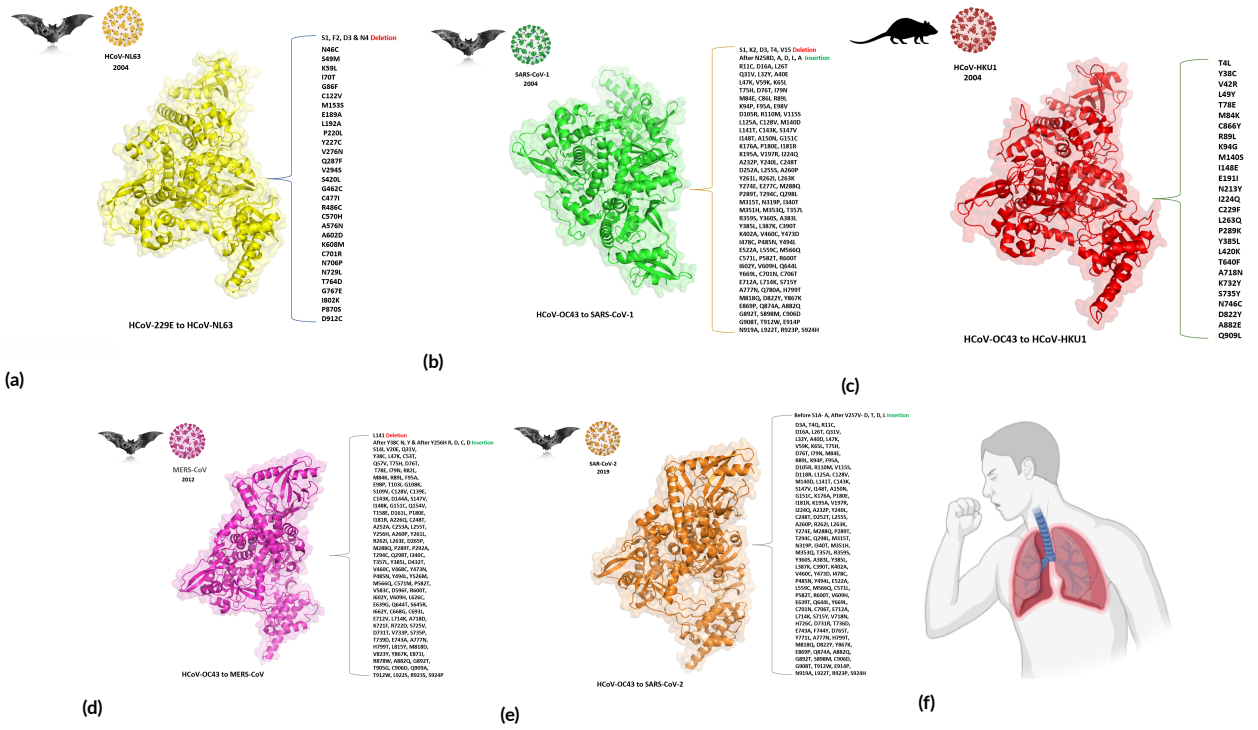

FIGURE 3 Major mutations (with no change in the chemical properties) in HCoVs.

TABLE 1 Percentage distribution of Hydrophobic amino acids in RdRp of all 7 HCoVs.

RdRp of Human Coronavirus	% of Hydrophobic amino acids observed in RdRp						
	Glycine	Proline	Phenylalanine	Alanine	Isoleucine	Leucine	Valine
229E	5.70%	2.90%	6.60%	6.00%	4.10%	8.20%	7.60%
OC43	5.20%	2.90%	5.70%	6.20%	4.80%	9.40%	8.6%.
SARS-CoV-1	5.00%	3.20%	60%	7.20%	3.50%	9.00%	8.20%
NL63	5.70%	2.80%	6.10%	5.60%	4.10%	9.20%	8.10%
HKU1	5.40%	2.80%	6.10%	5.60%	4.10%	9.20%	8.10%
MERS	5.60%	3.00%	6.10%	6.90%	3.80%	7.90%	8.10%
SARS-CoV-2	4.80%	3.20%	6.10%	6.90%	3.50%	8.90%	7.90%

i.e. SARS-CoV-1, MERS-CoV and SARS-Cov-2. They had pI of around 6.4. A slightly higher pI in SARS-CoV-2 might affect the interactions with RNA, cofactors, or therapeutic agents under physiological pH. Isoelectric point (pI) with predicted scaled solubility is shown in Table 2 and Supp. Figure 8.

3.2. Comparison of HCoVs sequences with similar viruses by Phylogenetic Tree construction

In order to analyze the relationship among the SARS-CoV-2 genome, other coronaviruses and other existing viruses available in the database, an evolutionary-based analysis was conducted. Viral phylogeny was constructed using the reference of SARS-CoV-2 whole genome sequence via the tBLAST. It showed the presence of 224 virus families and 200 host group (Figure 4). The circular cladogram results showed the global genome similarity relationship with seven HCoVs that are highlighted with red stars (Figure 4).

The phylogenetic analysis of HCoVs enhances our understanding of its origin and transmission. The phylogenetic tree results showed a close relationship between NL63 and 229E, which share a common reservoir bat and both belong to the family of *Alphacoronaviruses*. The branches of OC43 and HKU1 are *Betacoronaviruses* that originate from the same reservoir rodent. MERS shows a separate branch in the phylogenetic tree, and it was transmitted from camels to humans. And the pandemic caused by SARS-CoV-2 in 2019 shows a very close relationship with SARS-CoV-1. They both originated from bats (Supp. Figure 9).

When the individual nucleotides was checked , it was found that the percentage of adenine (A) nucleotide

TABLE 2 Isoelectric point (pI) of all human coronaviruses.

Human Coronaviruses	Isoelectric point (pI)	Predicted scaled solubility
229E	6.14	0.291
OC43	6.21	0.283
SARS-CoV-1	6.4	0.229
NL63	6.37	0.289
HKU1	6.22	0.287
MERS	6.41	0.25
SARS-CoV-2	6.44	0.233

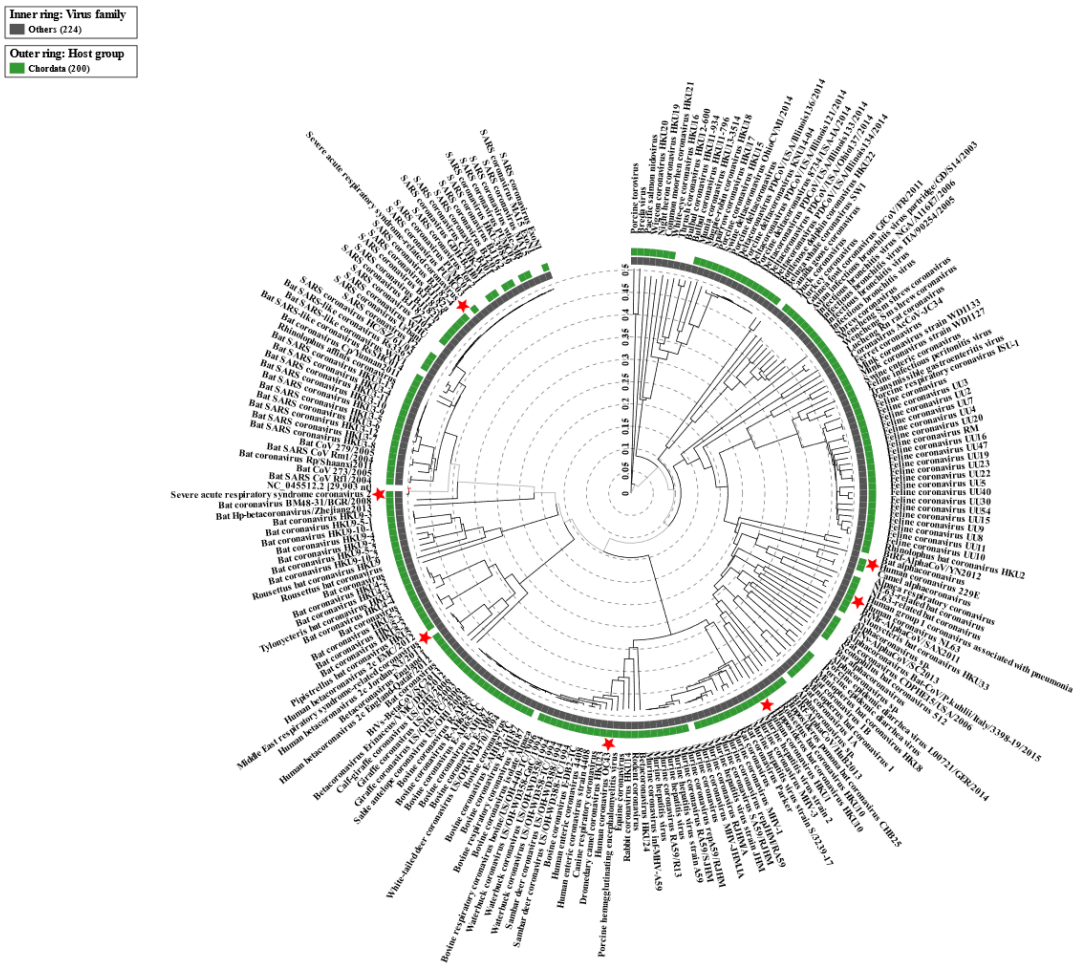

FIGURE 4 Viral tree of ssRNA represented in the circular view. The tree was constructed by the ViPTree. The cladogram tree includes 224 virus families (Inner ring) and 200 host groups (outer ring). For reference, Human coronaviruses were highlighted in red stars.

in SARS-CoV-2 was higher, followed by SARS-CoV-1. HCoV-HKU1 contained the highest amount of thymine (T) nucleotide, i.e. 40.10%, followed by HCoV-NL63, i.e. 39.21% (Figure 5).

3.3. Discussion

Coronaviruses were first identified in animals in the 1920s and later in humans in the 1960s. Till date, seven HCoVs have been discovered and some of which are seasonal and cause epidemics while some caused pandemics. Within the genome, the RdRp is a key enzyme involved in viral replication. When the sequence alignment of RdRp of all 7 HCoVs was done, significant mutations were observed; 171 mutations in *Alphacoronaviruses* (HCoV-229E and HCoV-NL63), and varying degrees of mutations in *Betacoronaviruses*, including 283 mutations, 4 insertions & 5 deletions in SARS-CoV-1, 103 in HCoV-HKU1, 293 mutations, with 6 insertions and 1 deletion in MERS-CoV and 312 mutations with 5 insertions in SARS-CoV-2 (the highest) all compared to HCoV-OC43. These mutations may not only lead to structural diversity but may also affect replication efficiency and the viral evolution.

The core RdRp of viruses have three subdomains, i.e., finger (nucleoside triphosphate entry), palm (polymerization), thumb (template binding) subdomains resembling a right-handed cup. The palm subdomain is the most conserved and catalytically active region, while the thumb subdomain is the most variable (Wolf et al. 2018; Venkataraman et al. 2018). The RdRp subdomains exhibit sequence similarity rates of 49% in the finger, 52% in the palm, and 42.3% in the thumb region. Among the five conserved regions in the core RdRp, the palm subdomain contains the most conserved regions, while the thumb subdomain is the most variable and has one conserved region. Unlike eukaryotic DNA polymerases, which possess proofreading mechanisms to correct replication errors, viral polymerases such as RdRp generally lack this function. This absence may contribute to higher mutation rates and drives the evolution observed in human coronaviruses HCoVs. The palm subdomain which has three conserved regions (residues ranges from 582-620, 680-815) could also be considered for a drug target.

The evaluation of the hydrophobicity, solubility, and isoelectric nature of these 7 RdRps revealed some important observations. RdRp of less pathogenic coronaviruses (e.g. 229E, OC43, NL63, HKU1) had balanced hydrophobicity, moderate solubility and a pI of 6.2–6.3, supporting stable but slower replication which contributes to mild infections. RdRp of MERS-CoV showed strong hydrophobicity and low solubility, potentially leading to protein aggregation or membrane interactions, which may contribute to its higher pathogenicity. RdRp of SARS-CoV-2 showed distinct hydrophobic peaks, lowest solubility and highest pI (∼6.4). This indicates stronger binding with the RNA, efficient replication and greater interaction with cofactors further contributing to rapid replication and high transmission potential among all the 7HCoVs.

Hydrophobicity helps a protein remain stable and physiologically active by enabling it to have a smaller surface area and fewer unwanted interactions with water. In the study done by Matyášek et al. (2021) , they analysed that mutation in amino acids gives the exceptional trend

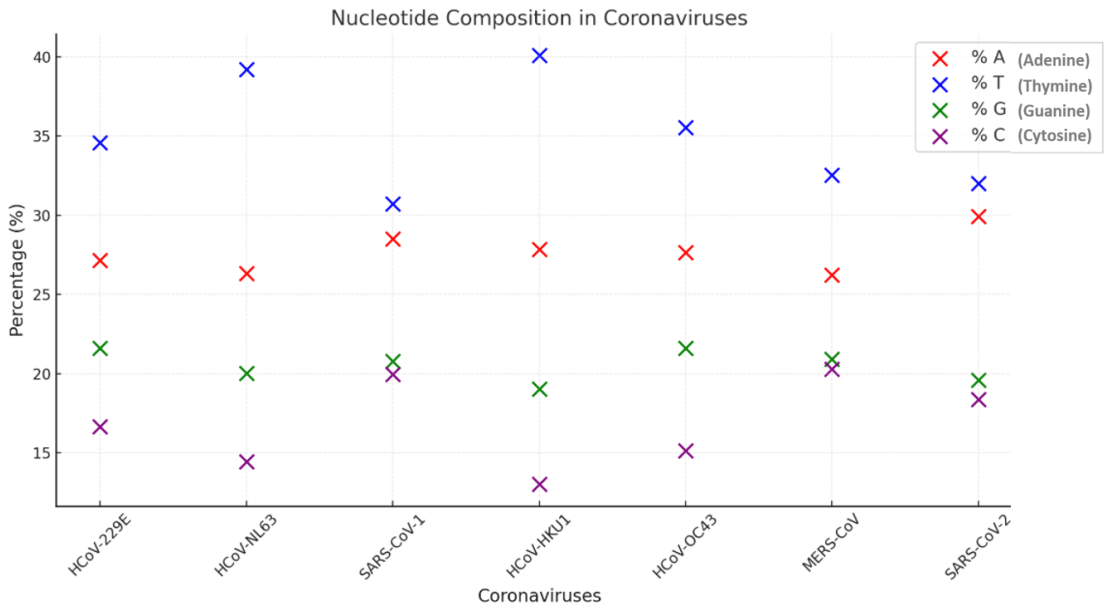

FIGURE 5 Percentage of nucleotides contents in HCoVs.

of increasing and decreasing hydrophobic amino acids in a broad range of SARS-CoV-2. High number of substitutions in amino acid leads to increase in hydrophobicity of the protein. Hydrophobic amino acids have the potential to enhance viral replication by altering the protein folding and interact inside RNA polymerase complexes. This may assist in faster replication. In Hepatitis C Virus (HCV), the RNA-dependent RNA polymerase (NS5B) contains a C-terminal hydrophobic domain essential for anchoring the polymerase to intracellular membranes and helps in facilitating effective replication complex formation (Lee et al. 2006). Another study done by Smertina et al. (2019) on calicivirus RNA-dependent RNA polymerase have indicated hydrophobic motifs that interact with intracellular membranes, thus promoting the formation of replication complexes. Mutations disrupting these hydrophobic interactions have been shown to impair viral replication, underscoring the importance of hydrophobic amino acids in the replication process. The occurrence of higher percentage of adenine (A) nucleotide in RdRp of SARS-CoV-2 and SARS-CoV-1 predicts their increased mutability, facilitating faster evolution but potentially reducing stability.

Mutations in the RdRp protein can also lead to an increase in the virus replication rate which is directly proportional to the pathogenicity. The rapid rate of replication and mutation facilitates the emergence of novel antigenic variations and the virus's ability to adapt to the host cell. A mutation can also cause change in specificity of enzyme, which in turn can affect the synthesis of enzymatic viral proteins (Wang et al. 2020). In our previous study, we reported the drug targets against RdRp and repurposed those drugs to develop a new drug molecule to inhibit its enzyme activity (Sharma et al. 2023).

Limitations of this study is that this is purely an in silico study. Further *in vitro* and *in vivo* studies may be required to validate our findings for an in-depth understanding of pathogenicity of HCoVs. Based on the present findings, RdRp emerged as the major enzymatic proteins of SARS-CoV-2, which can be responsible for severe pathogenicity based on the increased number of mutations and decreased hydrophobicity. Future in silico studies could be required to understand the mechanism in neutralizing RdRp protein and especially in infected lung cell lines for suggesting an effective anti-viral molecule.

Present study reported the transition of mutation trend of RdRp viral proteins in mild (HCOV229E; HCOVNL63) and severe (SARS-CoV-2) coronaviruses. In addition, the hydrophobicity trend of coronaviruses has also been studied. It is suggested that the study can be focused on RdRp viral protein and also focused on the event of outbreaks of coronaviruses for developing the RdRp-based effective drug target by appropriate anti-viral molecule.

4. Conclusions

The study of mutation pattern of RdRp protein in different HCoVs showed that mutation in this proteins can lead to enhanced viral replication, increased pathogenicity and immune evasion by HCoVs. Present study of mutation analysis represents a useful approach to predict the pandemic potential of HCoVs.

Acknowledgments

We are thankful to Indian Council of Medical Research, India (2021-9858) for providing financial support.

Authors' contributions

VJ, BA, AA perceived the concept. BS, SMB, PK, NS conducted the study and developed the design of this experiment. KK, DK, AC set up the facilities for the experimental investigation and oversaw it. BS, SMB, PK, NS, BA, KK, AC, DK performed the data analysis and literature review. VJ, BA, AA, BS, SMB did the interpretation of the findings and performed the initial drafting. KK, NS, PK, AC, DK reviewed and updated the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare no competing interest.

References

Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS. 2016. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 44(1):W147–53. doi:10.1093/NAR/GKW419.

Braun T, Bordería AV, Barbezange C, Vignuzzi M, Louzoun Y. 2019. Long-Term context-dependent genetic adaptation of the viral genetic cloud. Bioinformatics 35(11):1907–1915. doi:10.1093/bioinformatics/bty891.

Burrell CJ, Howard CR, Murphy FA. 2016. Coronavirus - an overview | ScienceDirect Topics Chapter 31. Fenner White's Med. Virol. p. 437–446.

Cicaloni V, Costanti F, Pasqui A, Bianchini M, Niccolai N, Bongini P. 2022. A Bioinformatics Approach to Investigate Structural and Non-Structural Proteins in Human Coronaviruses. Front. Genet. 13:891418. doi:10.3389/fgene.2022.891418.

Haagmans BL, Osterhaus ADME. 2009. SARS. Academic Press. p. 671–683.

Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J. 2017. Protein-Sol: A web tool for predicting protein solubility from sequence. Bioinformatics 33(19):3098–3100. doi:10.1093/bioinformatics/btx345.

Lamers MM, Haagmans BL. 2022. SARS-CoV-2 pathogenesis. Nat. Rev. Microbiol. 20(5):270–284. doi:10.1038/s41579-022-00713-0.

Lee H, Liu Y, Mejia E, Paul AV, Wimmer E. 2006. The C-Terminal Hydrophobic Domain of Hepatitis C Virus

RNA Polymerase NS5B Can Be Replaced with a Heterologous Domain of Poliovirus Protein 3A. J. Virol. 80(22):11343–11354. doi:10.1128/jvi.02072-05.

Liu DX, Liang JQ, Fung TS. 2021. Human Coronavirus-229E, -OC43, -NL63, and -HKU1 (Coronaviridae). Ref. Mod. Life Sci. 1-5:428–440. doi:10.1016/B978-0-12-809633-8.21501-X.

Ludwig S, Zarbock A. 2020. Coronaviruses and SARS-CoV-2: A Brief Overview. Anesth. Analg. 131(1):93–96. doi:10.1213/ANE.0000000000004845.

Matyášek R, Řehůřková K, Marošiová KB, Kovařík A. 2021. Mutational asymmetries in the SARS-Cov-2 genome may lead to increased hydrophobicity of virus proteins. Genes (Basel). 12(6):826. doi:10.3390/genes12060826.

Miłek J, Blicharz-Domańska K. 2018. Coronaviruses in avian species-review with focus on epidemiology and diagnosis in wild birds. J. Vet. Res. 62(3):249. doi:10.2478/jvetres-2018-0035.

Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. 2017. ViPTree: The viral proteomic tree server. Bioinformatics 33(15):2379–2380. doi:10.1093/bioinformatics/btx157.

Protein Data Bank. 1971. Protein Data Bank.

Sarma H, Jamir E, Sastry GN. 2022. Protein-protein interaction of RdRp with its co-factor NSP8 and NSP7 to decipher the interface hotspot residues for drug targeting: A comparison between SARS-CoV-2 and SARS-CoV. J. Mol. Struct. 1257:132602. doi:10.1016/j.molstruc.2022.132602.

Schrödinger L, DeLano W. 2020. PyMOL. URL http://www.pymol.org/pymol.

Sharma B, Angel B, Naik VU, Angel A, Joshi V, Shareef B, Singh N, Khan AS, Khaneja P, Barthwal S, Joshi R, Peer NM, Yadav K, Tomar K, Singh SP. 2023. Repurposed drug molecules targeting NSP12 protein of SARS-CoV-2: An in-silico study. AIMS Mol. Sci. 10(4):322–342. doi:10.3934/molsci.2023019.

Smertina E, Urakova N, Strive T, Frese M. 2019. Calicivirus RNA-dependent RNA polymerases: Evolution, structure, protein dynamics, and function. Front. Microbiol. 10(JUN):1280. doi:10.3389/fmicb.2019.01280.

Tang G, Liu Z, Chen D. 2022. Human coronaviruses: Origin, host and receptor. J. Clin. Virol. 155:105246. doi:10.1016/j.jcv.2022.105246.

Venkataraman S, Prasad BV, Selvarajan R. 2018. RNA dependent RNA polymerases: Insights from structure, function and evolution. Viruses 10(2):76. doi:10.3390/v10020076.

Vicenti I, Zazzi M, Saladini F. 2021. SARS-CoV-2 RNA-dependent RNA polymerase as a therapeutic target for COVID-19. Expert Opin. Ther. Pat. 31(4):325–337. doi:10.1080/13543776.2021.1880568.

Wang H, Pipes L, Nielsen R. 2020. Synonymous mutations and the molecular evolution of SARS-CoV-2 origins. Virus Evol. 7(1):veaa098. doi:10.1093/ve/veaa098.

Wei M, Yang N, Wang F, Zhao G, Gao H, Li Y. 2019. Epidemiology of Coronavirus Disease 2019 (COVID-19) Caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Disaster Med. Public Health Prep. 14(6):796–804. doi:10.1017/dmp.2020.155.

Wolf YI, Kazlauskas D, Iranzo J, Lucía-Sanz A, Kuhn JH, Krupovic M, Dolja VV, Koonin EV. 2018. Origins and evolution of the global RNA virome. MBio 9(6):e02329–18. doi:10.1128/mBio.02329-18.

Wong NA, Saier MH. 2021. The SARS-coronavirus infection cycle: A survey of viral membrane proteins, their functional interactions and pathogenesis. Int. J. Mol. Sci. 22(3):1308. doi:10.3390/ijms22031308.

Zhao X, Ding Y, Du J, Fan Y. 2020. 2020 update on human coronaviruses: One health, one world. Med. Nov. Technol. Devices 8:100043. doi:10.1016/j.medntd.2020.100043.