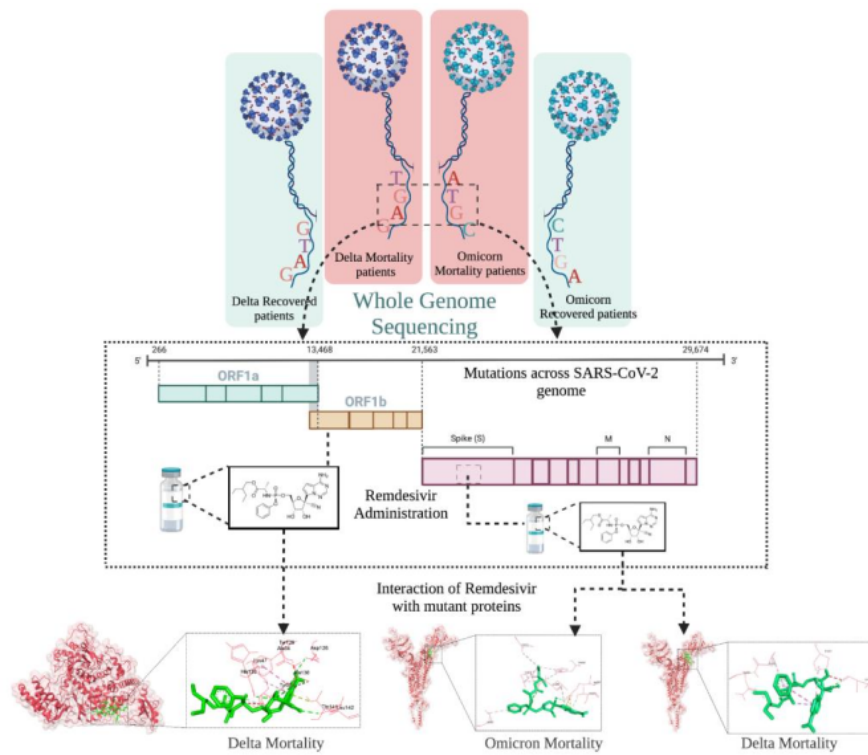# SARS-CoV-2 VOCs, Mutational diversity and clinical outcome: are they modulating drug efficacy by altered binding strength?

By Sheeba and Varsha

Abstract

The global COVID-19 pandemic continues due to the emerging Severe Acute Respiratory syndrome coronavirus 2 (SARS-CoV-2) variants of concern (VOC). Here, we performed comprehensive analysis of in-house sequenced SARS-CoV-2 genome mutations dynamics in the patients infected with the VOCs - Delta and Omicron, within Recovered and Mortality patients. Statistical analysis highlighted significant mutations - T4685A, N4992N, and G5063S in RdRp; T19R in NTD spike; K444N and N532H in RBD spike, associated with Delta mortality. Mutations, T19I in NTD spike, Q493R and N440K in the RBD spike were significantly associated with Omicron mortality. We performed molecular docking for possible effect of significant mutations on the binding of Remdesivir. We found that Remdesivir showed less binding efficacy with the mutant Spike protein of both Delta and Omicron mortality compared to recovered patients. This indicates that mortality associated mutations could have a modulatory effect on drug binding which could be associated with disease outcome.

Figure 1: Summary of the possible functional role of SARS-CoV-2 VOCs, mutations, and drug efficacy; with modulatory role in disease severity and clinical outcome. Distribution of SARS-CoV-2 infected individuals with VOCs - Delta and Omicron and segregation based on their disease outcome into *Recovered* and *Mortality*. Genome sequencing reveals the mutation spectrum of both VOCs across the different clinical outcome, and the modulatory potential of mutations vis-à-vis interaction to Remdesivir, a widely used drug during the different waves of COVID-19.

Introduction

The persistent upsurges of infectious cases engendered by the origin of SARS-CoV-2 and its novel variants of concern (VOC) resulting in more than 570 million cases worldwide have undoubtedly overburdened the public healthcare, medical infrastructure and posed economic challenges. The viral genome has been evolving since its advent and surveilling viral genome diversity during SARS-CoV-2 evolution has therefore been a high priority to restrain viral prevalence (Troyano-Hernáez et al. 2021) . Since the SARS-CoV-2 single stranded RNA virus exhibits the ability to acquire rapid mutations as it proliferates within a geographical region, only a subset of mutations are naturally selected favouring enhanced viral transmissibility, replication, and host immune evasion (McCormick et al. 2021). Consequently, several VOCs have emerged, challenging the containment measures of the pandemic in spite of practiced social measures and by reducing vaccines and available therapeutics efficacy (Barton et al. 2021).

Emergence of SARS-CoV-2 VOCs have led to natural selection of several mutations with distinct functional consequences that are being explored/elucidated/understood in greater detail through usage of model systems, genomic means and mechanism/s. Mutations, which are the drivers of potential evolutionary changes over time could have an effect on epidemiology, antigenicity and immune escape mechanisms thereby influencing the overall virus fitness (Tao et al. 2021; Ong et al. 2021). The first widely studied substitution, D614G, became dominant and is present in most of the variants now with a role in enhancing the viral replication (Korber et al. 2020). Another substitution, N501Y, convergently evolved in the Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), and Omicron (B.1.1.529). The N501Y has been reported to enhance the spike protein interaction with the human host ACE2 receptor and has been shown to influence vaccine efficacy (Escalera et al. 2022; Salehi-Vaziri et al. 2022). Numerous other studies have highlighted that Delta or B.1.617.2, had a higher proportion of infected patients who warranted ICU admission and unfortunate mortality due to the presence of nine mutations in the Spike protein of SARS-CoV-2 (Tian et al. 2021; Lin et al. 2021). Omicron, harbouring the highest number of mutations in the Spike protein compared to other existing VOCs, have been reported to cause a milder disease phenotype than the Delta variant (Saxena et al. 2021; Singhal 2022). Hence, VOCs have been

attributed to exacerbate the disease severity sub-phenotypes and differential mortality rates reported worldwide (https://covid19.who.int).

Since the beginning of the pandemic, multiple drugs such as antivirals, antibiotics, and antimalarials were predicted to be effective against SARS-CoV-2 (Drożdżal et al. 2021). Towards that, several clinical trial studies and computational drug interaction studies have been done and published worldwide (Ahmed et al. 2020; Ribaudo et al. 2020; Pirzada et al. 2021; Aouidate et al. 2020). For instance, drugs such as chloroquine and hydroxychloroquine, antimalarial drugs, which were available in the markets were widely utilised to treat COVID-19 in the beginning of the pandemic for quicker antiviral intervention. Later it was realized that these drugs do not effectively inhibit SARS-CoV-2 replication and reduce the severity (Hoffmann et al. 2020; Saghir et al. 2021). Notably, remdesivir, an adenosine analogue, which inhibits RdRp of SARS-CoV-2, was widely used in India during the Delta driven 2nd wave of the COVID-19. However, WHO trial which included 11266 adults showed that the drug remdesivir had little or no impact towards improved survival of SARS-CoV-2 infected patients (Beigel et al. 2020). On the other hand, Gurung et al found that the drugs, Paritaprevir and Emetine, could serve as an antiviral against SARS-CoV-2 RdRp through molecular docking studies (Bahadur Gurung et al. 2020). Nevertheless, it is important to note that mutations emerging in the drug interaction sites or in the target protein can potentially influence the drug binding efficiencies with role in the antiviral or vaccine development. Mohammed et.al showed that remdesivir binds more efficiently with RdRp mutant P323L than the wildtype by molecular simulation (Mohammad et al. 2021). However, it was reported that nirmatrelvir, a drug targeting main protease of SARS-CoV-2 showed similar binding potential with its variants as well as the wildtype (Ullrich et al. 2022).

The global endeavour to track the evolutionary trajectory of SARS-CoV-2 throughout the pandemic and screen for efficacious drugs against newly emerging viral variants is crucial. In this context, we aimed to comprehensively analyze the mutational spectrum of in-house sequenced SARS-CoV-2 VOCs affected individuals with two distinct clinical outcomes of *Recovered vs Mortality*, to identify viral genomic variability/ies, possibly affecting the drug efficacy. To elucidate and understand this aspect, we investigated the molecular docking characteristics with different SARS-CoV-2 target proteins including Nsp12, Nsp13 and Spike protein. Findings

indicate the functional role of mutations within the SARS-CoV-2 genome through sub-optimal drug binding compromising the efficacy.

Methods

Patient Categorisation and samples segregation

A total of 249 Nasopharyngeal RNA samples were obtained from the hospital admitted SARS-CoV-2 infected patients for this study. All the patient samples were given anonymous identities at CSIR-IGIB. Further, the patients were classified into four groups based on the disease/clinical outcomes obtained from the clinical data (Recovered and Mortality). This was further overlaid with the SARS-CoV-2 genome sequencing confirmed variants of concern, Delta and Omicron, leading to the infection (**Supplementary file 1**).

Whole Genome Sequencing

The whole genome sequencing of 249 SARS-CoV-2 genome sequences was performed utilizing both Oxford Nanopore technology (ONT) and Illumina (Nextseq 2000 and MiSeq) platforms.

For Nanopore based sequencing: Out of 249 samples, 32 samples were sequenced through Rapid Barcoding kit protocol (SQK-RBK110.96) available for the ONT platform. In summary, 50ng of the extracted viral RNA from the nasopharyngeal samples was used to reverse transcribe the RNA into complementary DNA using LunaScript RT SuperMix (New England Biolab. Cat No. E3010L). The single stranded cDNA was further utilized to amplify the viral genome using rapid barcoding primers (IDT Product number: 10007184) and Q5 High-Fidelity 2X master mix (New England Biolabs, Cat. No. M0494S). The amplicons were then ligated to the unique barcode sequences and purified using SPRI beads. The prepared library was ligated to the adapter sequences and loaded on MinION Mk1B/ Mk1C for sequencing (**Figure 1A**).

Illumina sequencing: The remaining 217 samples were sequenced using Illumina COVIDSeq (Cat. No.20043675 and reference guide: 1000000126053 v04). The cDNA was synthesized from the extracted nasopharyngeal RNA and was further used to amplify SARS-CoV-2 genome through two separate multiplex PCR reactions. This was followed by Tagmentation, a process to fragment and ligate amplicons with adapter sequences. The tagged amplicons then underwent a bead based

purification where they were further amplified in a second round of PCR with addition of index adapters and sequences essential for sequencing cluster generation. The indexed samples were pooled and purified using Illumina Tune beads. The final library was quantified using Qubit dsDNA HS Assay kit (Cat. No. Q32854). The prepared library was denatured and diluted to prepare a loading concentration of 11pM and 1nM, accordingly using the MiSeq and Nextseq 2000 System Denature and Dilute Libraries Guide respectively (Illumina, Document no. 15039740 v10 & Document # protocol 1000000126053 v04). The denatured library was then sequenced using MiSeq reagent kit v3 or NextSeq 1000/2000 P2 flow cell.

Sequencing Data Analysis

The ONT minION raw fast5 files were analyzed using the ARTIC end-to-end pipeline upto variant calling. These raw files were basecalled and demultiplexed through the Guppy basecaller that functions using the algorithms of ONT with a phred score cut-off >7 (Nanopore Community). Consequently, all the reads having a score below 7 were eliminated to filter out the low-quality reads which would compromise mutation calling confidence. The demultiplexed fast5 files were normalized according to the average size of the amplicon (1200bp) and were aligned to the reference SARS-CoV-2 genome (MN908947.3) using Minimap2 v2.17 (Li, 2018). Following their alignment to the reference genome, the fast5 files were indexed through Nanopolish (Loman et al., 2015) for variant calling from the minimap output files. Bcftools v1.8 was used to create consensus fasta with normalized minimap output (Danecek et al. 2021).

For the fastq files generated in Illumina sequencing data, a FASTAQC was performed to examine the phred quality score with a cut-off >20 (Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data). The low quality reads were filtered out and the adapter sequences were trimmed through the Trim Galore tool (Babraham Bioinformatics - TrimGalore). In order to separate any human read contamination, the sequences were aligned using HISAT2 algorithm on human data build hg38 (Kim et al., 2019). The consensus fasta files were then generated using the BEDTools and variant calling was performed with high phred quality score reads (Quinlan and Hall, 2010). The genomic coverage and depth for all the samples are available in **Supplementary file 2**.
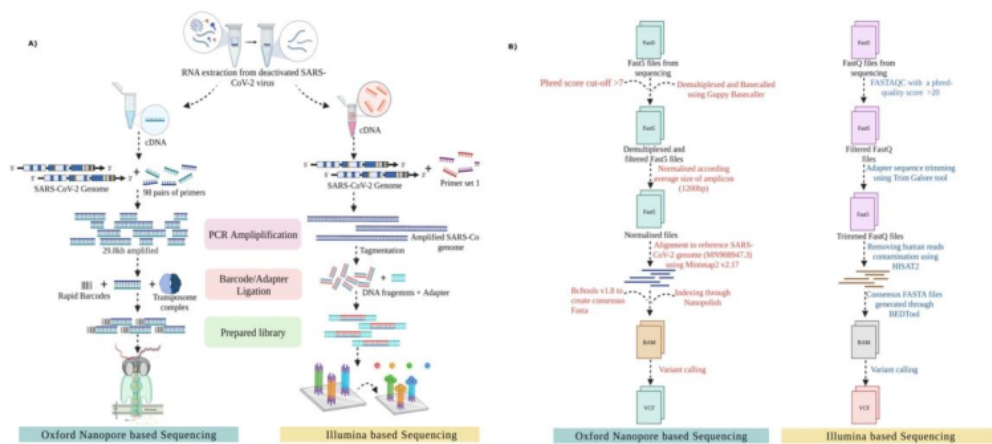
Figure 1: Methodology used in our study. (A) A graphical representation of the sequencing workflow from nasopharyngeal RNA extraction to library preparation, followed by sequencing on different platforms – ONT and Illumina. (B) The sequencing analysis workflow for both ONT and Illumina platforms highlighting the data QC, read annotation and mutation calling.

Mutation Analysis

The genomic diversity of 249 SARS-CoV-2 samples was analyzed by calculating the frequency of mutations vis-a-vis samples within the mortality and recovered patients across the VOCs – Delta and Omicron. Since two different sequencing platforms were utilized, in order to deduce the heterozygous calls, a haplotype based variant caller freebayes was used for Illumina sequencing data (Garrison and Marth, 2012). To further detect the genetic variation from the bam file, a filter of depth 20 was applied. Two public databases were used to obtain lineage defining mutations for Delta and Omicron variant (https://covariants.org/shared-mutations, https://github.com/cov-lineages/constellations).

Statistical Analysis

The statistical analysis was performed using R and Python programming languages. The Fisher exact test of independence for evaluating non-random association between the two categorical variables was performed to differentiate the mutation profile for all the four clinical groups by considering the Delta mortality and recovered as one group and the Omicron mortality and recovered as another group. The p-values were calculated from two-tailed tests considering 0.05

as the significance level. To quantify the strength of association between mutation and the clinical groups, phi-correlation coefficient was calculated.

Phylogenetic Analysis

The multiple sequence alignment of 249 SARS-CoV-2 genome sequences with the Wuhan reference strain (NC_045512.2) was performed through MAFFT (v7.475) (Katoh et al. 2002). The aligned sequences were manually trimmed and a phylogenetic tree was constructed through IQ-tree (Nguyen et al. 2015). The lineage stratification was done using PANGOLIN (https://cov-lineages.org/resources/pangolin.html). The phylogenetic analysis was then visualized using the FIGTREE tool (http://tree.bio.ed.ac.uk/software/figtree/).

Molecular Docking

Fasta sequence of proteins were taken from Uniprot [ID: P0DTD1 (ORF1ab), P0DTC2 (Spike glycoprotein)], and modelled using SWISS-MODEL (Waterhouse et al. 2018). The modelled protein was verified using PROCHECK, wherein the Ramachandran plot revealed 70% of the amino acid residues to be present in the most favoured region, 23.4% in the additional allowed region, 5% in generously allowed regions and only 1.5% in the disallowed region. Also, the RMSD value of modelled protein structure with respect to the protein crystal structure available in Uniport was observed to be 1.4 (**Supplementary file 3**).

The structure for remdesivir was downloaded from NCBI-PUBCHEM database (cid:121304016) in SDF format. SDF format of ligand was converted into PDB format using pymol (ref: The PyMOL Molecular Graphics System, Version 2.5.2 Schrödinger, LLC). Docking has been done through Autodock vina 4.2 (Morris et al. 2009). PDBQT formats of protein and ligand were then prepared using Autodock Tools v1.5.7 (Morris et al. 2009). Subsequently, free ions, water, and ligand were removed, kollman charges and polar hydrogens were introduced to the protein structure prior to docking. Visualizations of docked structures were done using PyMOL v2.5.2 and ChimeraX (Pettersen et al. 2021). Further interaction studies have been done using discovery studio (Pawar and Rohane 2021).

Remdesivir, a prodrug with broad spectrum antiviral activity has been widely utilized against SARS-CoV-2 infections (Khan et al. 2021). An in-vitro study by Wang et al have demonstrated

its antiviral mechanism as a nucleotide analogue in VERO-E6 cell lines (Wang et al. 2020). Moreover, clinical trials have shown the improvement of clinical outcomes in patients presenting moderate to severe COVID-19 upon Remdesivir administration (Gottlieb et al. 2022). These findings prompted us to consider and evaluate the binding strength of remdesivir to different viral mutant proteins in variable clinical outcomes.

Molecular dynamics simulations:

MD simulation was performed to verify the docking results. CHARMM36 force-field (da Silva et al. 2022) was used where ligands have been parameterized using CGENEFF server. The protein-ligand complexes of all the subgroups (Delta Recovered, Delta Mortality, Omicron Recovered, Omicron Mortality) along with the wildtype prepared from Autodock-vina 4.2 were subjected to 100ns MD simulations on the Nvidia-AVX2_256 GPU using the Gromacs-2020.2 CUDA module (Lindahl et al. 2020). The protein was prepared using the TIP3 water model with a cubic box of 1 Å of buffer distance. To reduce the bad contacts and clashes, energy minimization was done using the steepest descent method, for 50000 steps. Equilibrium was attained in two steps for all the complexes, with the first being 100 ps of NVT [number of particles (N), system volume (V) and temperature (T)] equilibration and the second one at 100 ps of NPT [number of particles (N), system volume (V) and pressure (P)] equilibration. A Berendsen thermostat and a Parrinello-Rahman barostat was used to maintain the system's temperature (300°C) and pressure respectively. (Berendsen et al. 1984) (Parrinello 1981). Moreover, the LINCS approach was used to address the system's long-range interaction (Hess et al. 1997). Further downstream analysis of simulation results was performed using xmgrace 5.1.2 (Turner, P. J. "XMGRACE, Version 5.1.19." Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR 2 (2005)).

Results

Mutation Profile of SARS-CoV-2 VOCs within distinct Clinical Outcomes

In alignment with our study theme, segregation based on clinical outcome and SARS-CoV-2 whole genome sequencing data within our cohort of 249 patients, were categorized into four groups - Delta recovered (67), Delta mortality (69), Omicron recovered (96) and Omicron mortality (17), respectively. The rationale was to evaluate the association of mutations with different clinical outcomes presented by the COVID-19 patients and its possible impact on drug efficacy. Mutation spectrum analysis across our samples yielded a total of 697 mutations across the cohort, of which 418 mutations were observed in the Delta subclinical groups (recovered and mortality) and 279 mutations within the Omicron subgroups (recovered and mortality), respectively (**Supplementary file 4**). This evinces towards a higher number of mutations exhibited by the Delta variant (3.07/sample) relative to that of the Omicron (2.46/sample). **Figure 2B** represents the cohort frequency of mutations and their corresponding genomic orientation across the viral genome for all the four subgroups. Delving further, we performed SARS-CoV-2 gene normalization to observe the relative abundance of mutations with respect to the gene length across the viral genome (**Figure 2D**). This revealed an overrepresentation of mutations in the ORF8 region in the Delta and within the 3'UTR region in the Omicron. However, the least number of normalised mutations were observed in the Membrane gene within the Delta and ORF1ab gene in the Omicron.
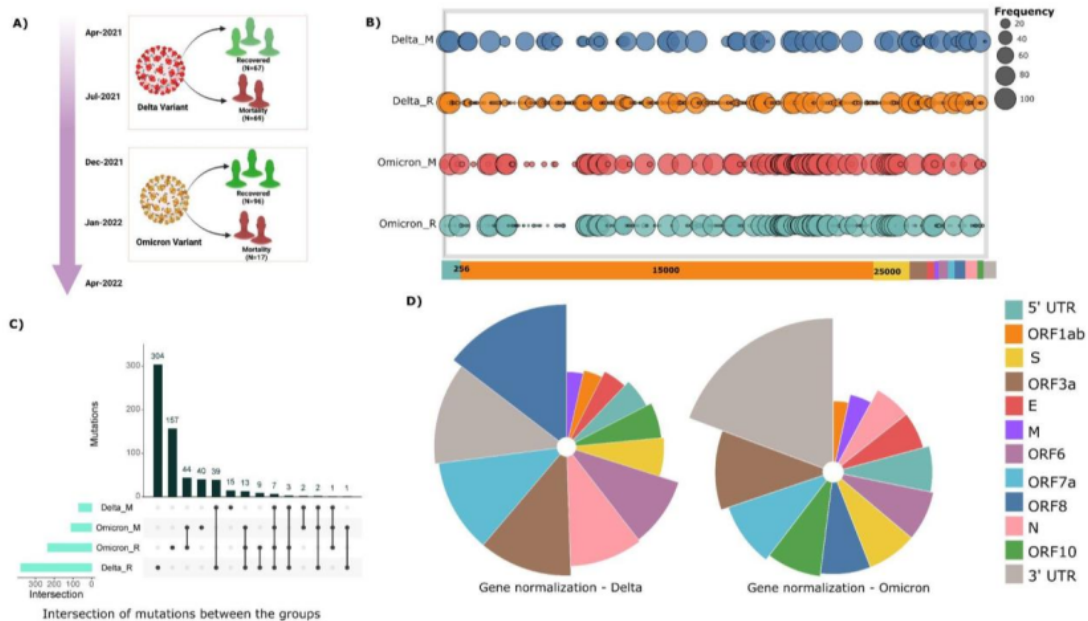
Figure 2: Distribution of mutation frequency across the viral genome in our cohort of 249 COVID-19 patients. (A) Sample distribution for patients infected with Delta and Omicron variant, categorised based on their clinical outcomes - *recovered and mortality*. (B) Mutational landscape for all the four subgroups, depicting the cohort frequency of mutations according to its genome coordinates. (C) Intersection of mutations between the subgroups – Delta mortality, Omicron mortality, Omicron recovered and Delta recovered. (D) Visual representation of gene normalization for both the VOCs, wherein the number of mutations were normalized with respect to the length of the gene.

Phylogenetic Analysis

Comparative genomic analysis revealed distribution of the samples across the lineages of B.1.617.2 (43.37%) and BA.2 (36.14%). Several B.1.617.2 (Delta) sub-lineages including AY.112, AY.120, AY.122, AY.127, AY.46.5, AY.98 were together categorized as AY.* lineage (11.24%). Moreover, sub-lineages of BA.2 (Omicron), as evident in **Figure 3** were BA.2.10, BA.2.10.1, BA.2.2, BA.2.12, BA.2.23, and BA.2.3.4 (6.42%). We further observed the presence of an Omicron lineage B.1.1.529 (1.2%) followed by the sub-lineages XQ (0.8%), XT (0.4%) and BA.1.1.18 (0.4%). Subsequently, stratification of lineages based on the clinical outcome revealed the overrepresentation of B.1.617.2 across Delta recovered (58%) and mortality patients (100%) and BA.2 dominance for the Omicron recovered (79.16%) and mortality (82.35%) patients. Importantly, all the Delta mortality patients belonged to the lineage B.1.617.2.
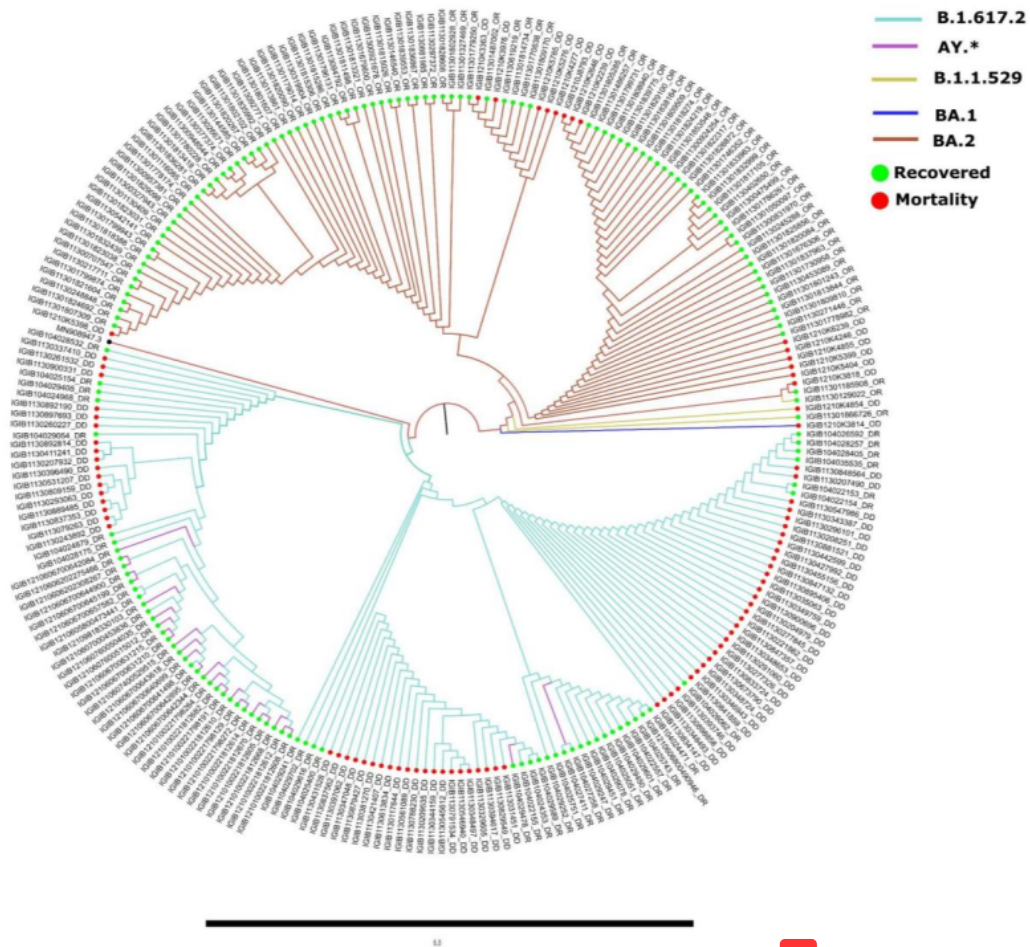
Figure 3: Phylogenetic analysis of 249 SARS-CoV-2 genome sequences in comparison to the wild type strain. The distribution of the lineages and sub-lineages across in-house sequenced Delta and Omicron VOCs is representative of the community prevalence in India during their respective waves.

Statistical & Correlation Analysis reveals mutations with potential significance

To further comprehend the significance of mutations associated with clinical outcomes across the VOCs, Fisher's exact test was performed considering a p-value of <0.05. We observed 50 statistically significant mutations in the Delta clinical subgroups and 11 significant mutations in

the Omicron (available as **supplementary file 4**). Phi-Coefficient correlation analysis of these significant mutations revealed 20 mutations to be associated with the Delta recovered patients and 30 mutations were found to be associated with the Delta mortality patients. The same analysis highlighted three mutations for their association with Omicron recovered and eight mutations with the Omicron mortality patients. The distribution of significant mutations in different regions of SARS-CoV-2 genome including ORF1ab, Spike, ORF7a, ORF8 and Nucleocapsid protein has been elucidated in **Figure 4**.
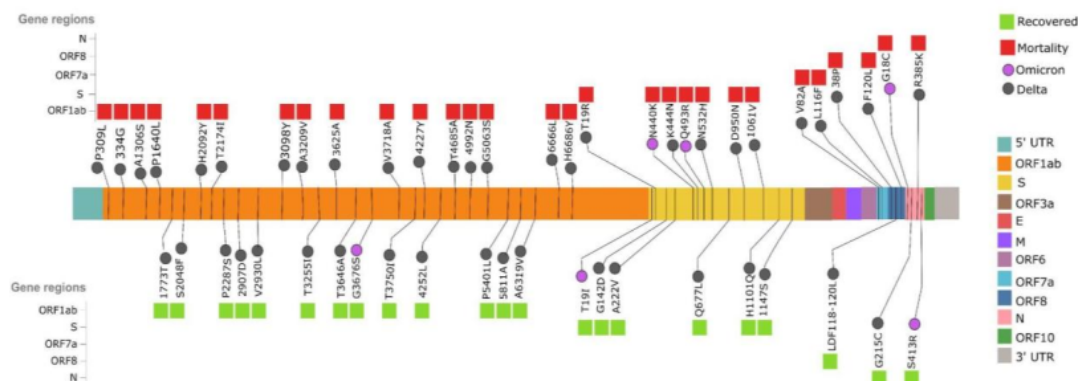
Figure 4: Statistically significant mutations within the recovered and mortality patients of Delta and Omicron variants. The pan genomes distribution of the significant mutations across VOCs observed for the recovered (green) and mortality patients during the waves of higher SARS-CoV-2 infectivity (red).

Intriguingly, most of the significant mutations (n=28) in the Delta patients were seen in the ORF1ab region and notably, three mutations T4685A (p=4.61E-23), N4992N (p=0.0006) and G5063S (p=0.0005) were observed in the Nsp12/RdRp region of the Delta mortality patients, whereas no mutations in RdRp were seen in the Delta Recovered. This could possibly emphasize the critical role of RdRp in the viral replication, transcription and its subsequent contribution in elevating viral load thereby influencing the progression of disease severity. Contrarily, two mutations, P5401L and A5811A in the Nsp13/Helicase region were found to be associated (p-value=4.30E-06 and 0.0012) within the Delta recovered patients.

Furthermore, among the significant mutations observed in the spike region for the Delta infected patients, six mutations were associated with mortality and five mutations with the recovered, respectively. It is important to mention that one of the mutations observed in the NTD region of the Spike protein, T19R in the Delta mortality patients (*p*-value= 9.93E-13) have been previously predicted to reduce the inhibitory effect of monoclonal antibodies. Interestingly, two significant mutations, K444N and N532H (*p*-value=0.0006, 0.0006) in the mortality patients were observed to be present in the RBD region of spike protein which could possibly impact the Spike-RBD interaction with the ACE2 receptor.

Upon examining the mutations present in the Omicron subgroups, we observed two mutations, Q493R and N440K in the RBD region of Spike protein to be strongly associated (*p*-value=1.55E-12 and 0.0029) with the mortality. Spike mutations, particularly in the RBD region have been attributed to confer immune escape properties to the virus and thereby assisting in the progression of the disease severity. At the same time, T19I in the NTD region of Spike protein was found to be associated (*p*-value=0.0045) with the recovered patients. Also, other significant mutations, such as G18C and R413S in the Nucleocapsid gene were found to be associated with mortality (*p*-value= 0.0214) and recovered (*p*-value= 0.0214) patients, respectively. A comprehensive literature search for all the significant mutations was performed to elucidate the functional consequence of these mutations in altering the viral genome characteristics (**supplementary file 5**).

Interactions of Remdesivir with SARS-CoV-2 Nsp12, Nsp13 protein & Spike protein

To explore and understand the modulatory potential of the observed significant mutations towards available therapeutics, molecular docking was performed to assess the interaction strength and affinity of drugs with the mutated SARS-CoV-2 proteins. We investigated the interaction of the drug Remdesivir with the viral proteins, Nsp12, Nsp13 and Spike. The significant mutations observed in the Delta and Omicron patients were considered for the 3D homology modelling of the respective proteins. **Table 1** enumerates the binding affinity for each ligand-protein complex, mutations in the active sites, interacting residues present at the active site of protein and the number of hydrogen bonds involved in the interaction of remdesivir with the predicted region of protein in each group.

Table 1: Docking results of SARS-CoV-2 Nsp12, Nsp13 and Spike mutant proteins against Remdesivir.

Drug	Protein	Patient Group	Mutations	Mutation in Active site	Binding Affinity	Interacting Residues	Number of H-Bonds
Remdesivir	Nsp12 Wildtype	Wildtype			-8.1	Lys47, Tyr129, Ala130, Asp135, Ser709, Lys780, Asn781	6
	Nsp12_ mutation	Delta Mortality	T4685A, N4992N, G5063S	None	-8.2	Leu142, His133, Ala130, Thr141, Tyr32, Asp126, Ala34, Tyr129, Lys47	6
	Nsp13_ Wildtype	Wildtype			-8.7	Pro406, Pro408, Leu412, Phe422, Tyr515, Asn516, Thr532, Ala553, His554, Arg560	8
	Nsp13_D R	Delta Recovered	P5401L, A5811A	None	-8.5	Arg409, His554, Ala553, Leu412, Leu417	3

Remdesivir	S protein Wildtype	S protein Wildtype			-6.3	Asp40, Lys206, Tyr38, Ala222	5
	S protein_ DR	Delta Recovered	G142D, A222V, Q677V, H1101Q, 1147S	None	-7.3	Val367,Ser371, Asn343,Leu44 1,Phe374,Trp4 36	10
	S protein _ DM	Delta Mortality	T19R, K444N, N532H, D950N, 1061V	K444N, N532H	-5.8	Asn331,Gln563 ,Pro521,Pro330	5
	S protien_ OR	Omicron Recovered	T19I	None	-7.1	Phe565,Asp428 ,Thr430,Phe51 5,Phe464,Glu5 16,Leu517,Pro 426	6
	Sprotien_ OM	Omicron Mortality	N440K, Q493R	N440K, Q493R	-6.4	Arg567,Thr573 ,Leu546,Val38 2,Ala522,Leu5 17,Gly545,Phe 565	3

Interaction study by docking revealed that the remdesivir drug binds to Nsp12 wildtype and Nsp12 mutant in the Delta mortality patients with nearly the same binding efficacy. Consequently, we observed six hydrogen bonds in the active site of Nsp12 wildtype and mutant proteins, four hydrophobic bonds in the active site of the wildtype and three hydrophobic bonds in the active site of mutant Nsp12 protein interacting with the remdesivir, respectively. Notably, hydrophobic interactions in the wildtype Nsp12 were formed through three amino acids - Ala130, Tyr129 and Lys780 and the mutant Nsp12 hydrophobic interactions involved the amino acids Tyr129, Ala130, and Lys47 (Table1). The appearance of a salt bridge during docking can indicate Nsp12 as a preferential binding site for Remdesivir. On the other hand, docking of remdesivir with Nsp13

wildtype and Nsp13 mutant in the delta recovered patients revealed that the efficiency of binding is more within the wildtype than the mutant protein. We also observed more hydrogen bonds and hydrophobic bonds in the drug interaction with the wildtype as compared with the mutant Nsp13 protein (**supplementary file 6**). Though remdesivir is widely studied to interact with Nsp12, Nsp14 and Nsp5/3CL Protease, we observed a considerable binding affinity of Remdesivir with the Spike protein.
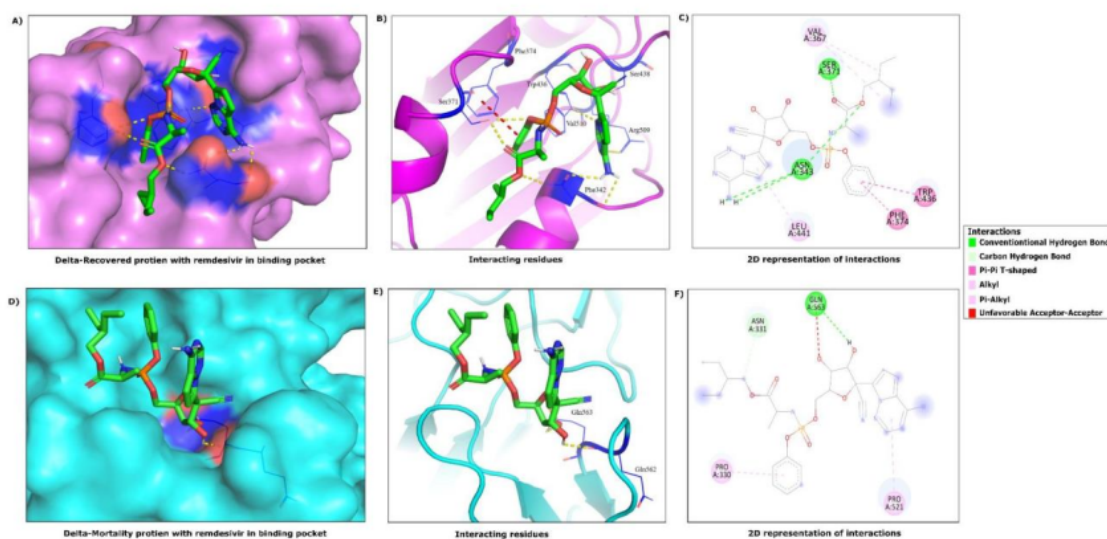
Figure 5: 3D and 2D view of interactions observed between Remdesivir and viral mutant Spike proteins. (A) and (C) are showing 3D & 2D view of interactions between Remdesivir and RBD- Spike region in the Delta Recovered patients. (B) and (D) highlight 3D & 2D view of interactions between Remdesivir and RBD-Spike region in the Delta Mortality patients.

Intriguingly, both the mortality group patients demonstrated a decrease in binding affinity and number of H-bonds in comparison to that of the recovered patients for the Spike protein. Notably, the presence of two mutations in the RBD region of Spike protein can possibly influence the binding characteristics, thereby resulting in a weaker interaction between Remdesivir and spike protein across both the mortality group patients. An increase in binding affinity and number of H-bonds (**Figure 5 & 6**) revealed by the recovered patients, when compared to the wildtype and

mortality patients, is indicative of the possibility of these mutations promoting a stronger interaction between Remdesivir and the SARS-CoV-2 Spike protein.
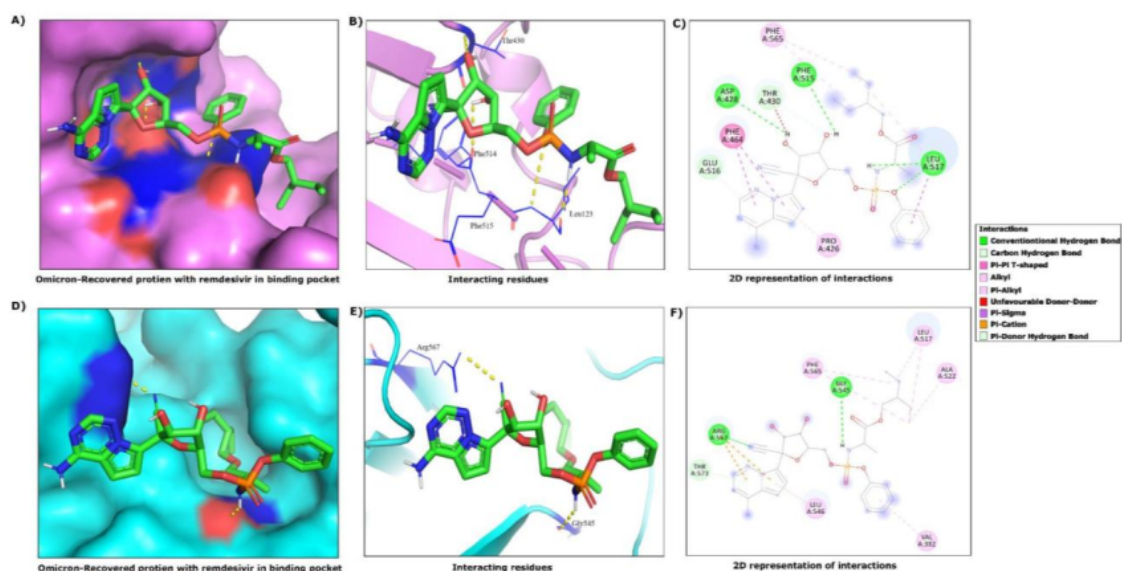

Figure 6: 3D and 2D view of interactions observed between Remdesivir and viral mutant **proteins.** (A) and (C) highlights 3D & 2D view of interactions between Remdesivir and Spike-RBD mutant in the Omicron Recovered. (B) and (D) emphasizes 3D & 2D view of interactions between Remdesivir and Spike-RBD mutant in Omicron Mortality patients.

Molecular Dynamics Simulation of Remdesivir and SARS-CoV-2 Spike proteins

Based on the findings of molecular docking, MD simulation was performed to verify the difference in the stability of the predicted Remdesivir-Spike complex across all the four sub-group. The structural dynamics of both the ligand and ligand-protein complex were evaluated by measuring their respective RMSD and RMSF values. Delving further, we observed the average RMSD values for Remdesivir-spike complex to be 1.54nm and convergence of the RMSD curve after 85ns for the Delta mortality group. Although the average RMSD value of the same complex was found to be 1.3nm for the Delta recovered group along with the curve stabilization after 85ns of MD simulation (**Figure 7A**). The variability in the average RMSD values between the clinically

distinct groups can possibly suggest a more stabilised and enclosed accommodation of Remdesivir within the predicted binding pockets of the RBD-spike region in the Delta recovered group. This was further substantiated by the ligand RMSD trajectories observed in **Figure 7 (B)**, wherein Remdesivir confers a ligand shift around 30ns and fluctuates randomly from the active binding pocket until 100ns in the Delta mortality group. On the other hand, the ligand RMSD curve of the Delta recovered group revealed a stable accommodation of remdesivir within the active pockets of the protein upto 45ns. Moreover, we observed greater residual fluctuations in the Delta mortality group as compared to the Delta recovered group (**Figure 7C**).
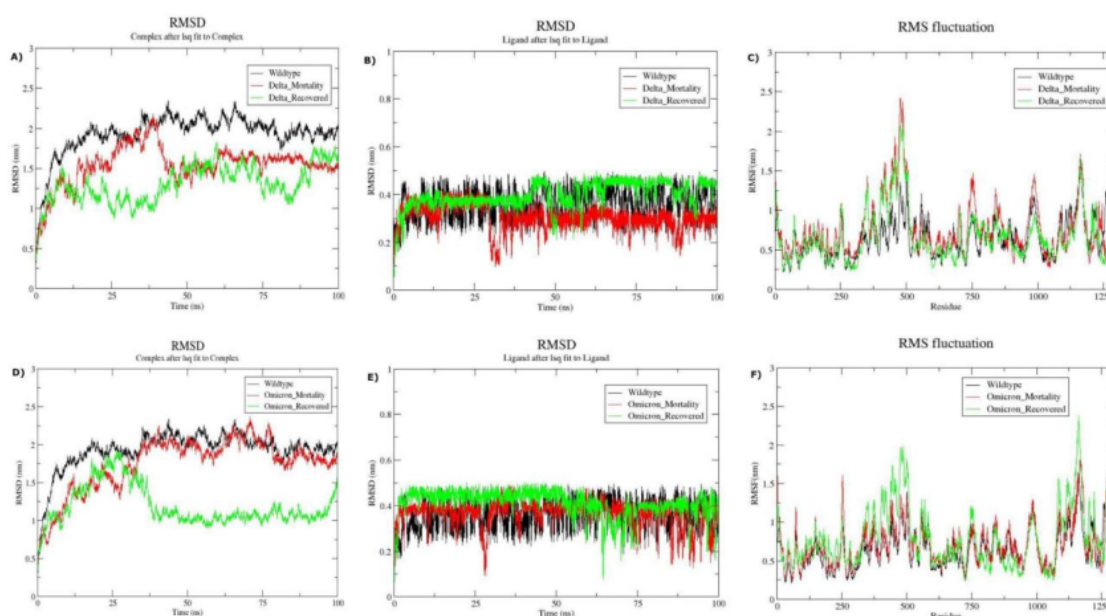
Figure 7: Structural Dynamics of Remdesivir and SARS-CoV-2 Spike protein: (A) RMSD plot of Remdesivir-Spike complex as a function of time for Delta groups (B) RMSD plot of Remdesivir for Delta groups (C) Root Mean Square Fluctuation vs residue for Delta groups.

(D) RMSD plot of Remdesivir-Spike complex as a function of time for Omicron groups (E) RMSD plot of Remdesivir for Omicron groups (F) Root Mean Square Fluctuation vs residue for Omicron groups.

Further investigation of the structural dynamics for the Omicron mortality group revealed a constant fluctuation in the RMSD values of the Remdesivir-spike complex until the completion of MD simulation **(Figure 7D).** The average RMSD for the complex was found to be 1.74nm. However, we observed the RMSD trajectory of the Remdesivir-Spike complex to be stable around 50ns to 75ns for the Omicron recovered group, contributing to an average RMSD of 1.18nm. Interestingly, the ligand RMSD curve of the Omicron mortality group exhibited a similar pattern of ligand shift from the binding site of the protein around 30ns as that of the Delta mortality group and endured the random fluctuations upto 100ns of the simulation **(Figure 7E)**. We further observed the stabilization of the ligand RMSD curve from 5ns to 50ns for the Omicron recovered group. Although a comparatively stable interaction of Remdesivir within the binding pockets of the Spike protein was observed for the Omicron recovered group than the mortality group, the Root Mean Square Fluctuations per residue were observed to be higher in the recovered group **(Figure 7F)**.

Discussion

Several vaccines and drugs are being used globally towards pandemic management with differential efficacy with hierarchical layers of modulation. This is in combination with the vaccine development for the below 18 age groups with specific success towards the same as well. Hence, global efforts are towards developing, exploring, repurposing several antiviral drugs to subside the clinical severity, especially continuously evolving and emerging SARS-CoV-2 variants of concern. Towards that, repurposing of available drugs, such as remdesivir and dexamethasone, has been shown to be efficacious in managing the disease severity symptoms (Kelleni et. al. 2021). At the same, we have observed differential efficacy of the drug in a subset of the SARS-CoV-2 infected patients. What can be the reason? Is it the different VOCs or mutations or genomic region-specific mutations? However, our understanding about the association of mutations with different clinical outcomes and the potency of these drugs is restricted (Stevens et al. 2022). To evaluate and elucidate this aspect, we stratified SARS-CoV-2 infected patients based on the clinical outcome and assessed the binding interactions of Remdesivir with different mutant viral proteins.

Extensive studies, computational and experimental, involving the screening of drugs with patients presenting variable clinical symptoms, is important to optimize treatment and identify other available treatment options as well.

In the present study, through genomic analysis of 249 SARS-CoV-2 genome sequences, we sought to investigate the differences in the association of mutations manifested by SARS-CoV-2 VOCs infected individuals with the disease severity leading to differential clinical outcomes. All the mutations observed within our cohort were analysed for their functional relevance and their respective viral genome coordinates. This revealed the distribution of significant mutations in the ORF1ab (28), Spike (11), ORF8 (3), Intergenic (3), Orf7a (2), Nucleocapsid (2) and only one mutation in the 3'UTR region across the viral genomes in the Delta infected patients. Contrarily, we observed a greater number of significant mutations in the Spike (5) followed by mutations in the Nucleocapsid (2), 3'UTR (2), 5'UTR (1) and the ORF1ab (1) genes in the Omicron patients. Furthermore, we found the co-presence of five mutations (P309L in NSP2, P1640L in NSP3, A3209V in NSP4, V3718A in NSP6, and R385K in the N gene) in the Delta mortality patients, which has been recently reported to be prevalent in the symptomatic patients than the asymptomatic ones (Banerjee et al. 2022). On the other hand, three co-occurring mutations G142D in the Spike, P2287S in the Nsp3 and T3255I in the Nsp4 region of the ORF1ab was observed in the Delta recovered patients which has been previously reported for their correlation with milder infectious cases of SARS-CoV-2 (Nilgiriwala et al. 2022).

Mutations in the Spike RBD region are crucial for increased infectivity and antibody resistance for SARS-CoV-2 emerging variants, and thereby promoting viral pathogenesis (Barton et al. 2021). Keeping a constant track of these RBD mutations along with non-RBD mutations is imperative for the development of potential vaccines and therapies as well as monitoring the efficacy of the ones currently in use (Almehdi et al. 2021). Celik et. al. evaluated SARS-CoV-2 variants, including Iota, Mu, Delta plus, Kappa, Lambda and C.1.2 for their changes in the interacting residues between spike-RBD region and hACE2 receptor, and highlighted the increase in the stability of the interacting complex as compared to the wild-type leading to enhanced virulence (Celik et al. 2022). Furthermore, Shafiq et. al. demonstrated that one of the recent SARS-CoV-2 variant in Tanzania, A.30 strain, manifesting the mutations R346K, T478K and E484K in the Spike-RBD

region aids the virus with immune escape properties (Shafiq et al. 2022). The Omicron variant, harbours a large number of mutations in the RBD region of the spike protein, including K417N, G446S, Q493R, and Q498R, leading to more stable interactions with the hACE2 protein (Khan et al. 2022). Hence, these hotspot mutations in the Omicron have been considered as a potential target for drug design and repurposing to combat the viral spread (Khan et al. 2022).

Interestingly, one of the significant mutations, K444N in the RBD region of the spike protein, observed in the Delta mortality patients was identified by Weisblum et. al. for conferring resistance to neutralizing antibodies against the spike protein (Weisblum et al. 2020). Other significant mutations in the Spike NTD region, such as T19R and D950N, observed in the Delta mortality have been attributed as a potential target for anti-NTD neutralizing antibodies (Planas et al. 2021). Furthermore, the Omicron mortality patients also manifested mutations in the RBD region of the Spike protein, including Q943R and N440K, which has been reported to enhance the binding affinity towards hACE2 receptor and escape from the antibody neutralization, consequently increasing chances of reinfection (Kullappan et al. 2021; Focosi et al. 2021).

To evaluate the propensity of these associated mutations towards functional modulation in SARS-CoV-2 pathogenesis, we computationally evaluated the docking parameters across both the Delta and Omicron sub-group of patients for different viral proteins. Since remdesivir has been previously shown to have high binding affinity to RdRp, we observed that RdRp mutant (T4685A, N4992N and G5063S) in the Delta mortality exhibited nearly same binding affinity for Remdesivir as that of the wildtype (Byléhn et al. 2021; Khan et al. 2021). This possibly suggest that the RdRp mutants identified in the study does not affect remdesivir binding and does not modulate efficacy between recovered and the mortality clinical outcome. Nsp13 mutants (P5401L and A5811A) conferred decreased binding affinity towards remdesivir, indicating a weaker interaction between the protein and the ligand. Notably, the mutations K444N and N532H in the Delta mortality, and Q493R and N440K in the Omicron mortality within the RBD region of Spike protein resulted in a lower binding affinity for remdesivir than their respective recovered patients (Table1). Consequently, these mutations observed across both VOCs hold the potential to ameliorate the disease severity by reducing the binding efficacy to remdesivir.

Considering Molecular Dynamics Simulation as an effective technique to validate the stability profiles of the protein-ligand complexes presented in our study, we performed 100ns simulation for all the clinical groups. Intriguingly, the Remdesivir-Spike protein complex of the Delta and Omicron mortality groups exhibited a higher average RMSD as compared to their respective recovered groups. This can thereby signify the relative instability of the Remdesivir-Spike protein complex of the Delta and Omicron mortality groups. The dynamic behaviour of Remdesivir observed as ligand shift from the binding pockets of the Spike protein in both the mortality groups further evinces towards inadequate accommodation of Remdesivir within the Spike protein. However, further in-vitro validation studies are required to support the role of these mutant proteins in disease transmission and pathogenicity.

At present, the availability of several antiviral drugs, immune modulators and monoclonal antibody treatments have comparatively reduced the current disease burden (Favalli et al. 2020; Kumar et al. 2021). The global vaccination data highlight a substantial number of individuals to be fully vaccinated worldwide, further aiding in reduction of morbidity and mortality (Ritchie et al. 2020). Nevertheless, robust genomic surveillance can highlight emerging SARS-CoV-2 variants and help identify mutations manifesting characteristics affecting the public health and clinical interventions.

Conclusion

The present study was initiated to, i) comprehensively investigate the viral genome heterogeneity represented by SARS-CoV-2 infected individuals with different VOCs, ii) identify significant mutations potentially modulating the different clinical outcomes, and iii) explore the functional aspect of such mutations for their propensity to impact the efficacy of available and future therapeutics. Our findings reveal significant mutations, T4685A, N4992N, and G5063S in the RdRp; T19R in the NTD spike; and K444N and N532H in the RBD spike region to be associated with the Delta mortality patients. At the same time, we observed the mutations, T19I in the NTD Spike, Q493R and N440K in the RBD spike, G18C and R413S in the Nucleocapsid gene to be significantly associated with the Omicron mortality patients. Subsequent docking studies of Remdesivir with mutant viral proteins including Nsp12, Nsp13 and Spike RBD region

demonstrated a reduction in binding strength of Remdesivir to mutant Spike proteins for both the mortality groups. The docking results were further verified through MD simulation. In combination, these findings suggest the possible role of these mutations in reduced drug efficacy and potential role in progression of disease severity.

SARS-CoV-2 VOCs, Mutational diversity and clinical outcome: are they modulating drug efficacy by altered binding strength?

7   Margot Paco-Chipana, Camilo Febres-Molina, Jorge Alberto Aguilar-Pineda, Badhin Gómez. "Novel In Silico Insights into Rv1417 and Rv2617c as Potential Protein Targets: The Importance of the Medium on the Structural Interactions with Exported Repetitive Protein (Erp) of Mycobacterium tuberculosis", Polymers, 2022
Crossref

25 words — < 1%

8   wjgnet.com
Internet

25 words — < 1%

9   Chad A. Brambley, Austin L. Bolay, Henry Salvo, Amanda L. Jansch et al. " Structural Characterization of sp. KT-1 PahZ1-Catalyzed Biodegradation of Thermally Synthesized Poly(aspartic acid) ", ACS Sustainable Chemistry & Engineering, 2020
Crossref

24 words — < 1%

10   punchout.mesoscale.com
Internet

24 words — < 1%

11   Priyanka Mehta, Varsha Ravi, Priti Devi, Ranjeet Maurya et al. "Mutational dynamics across VOCs in International travellers and Community transmission underscores importance of Spike-ACE2 interaction", Microbiological Research, 2022
Crossref

23 words — < 1%

12   Ranjeet Maurya, Pallavi Mishra, Aparna Swaminathan, Varsha Ravi et al. "SARS-CoV-2 Mutations and COVID-19 Clinical Outcome: Mutation Global Frequency Dynamics and Structural Modulation Hold the Key", Frontiers in Cellular and Infection Microbiology, 2022
Crossref

20 words — < 1%

13   Pulleri Kandi Harsha, Chitra Pattabiraman, Anson Kunjumon George, Madhusudhan J. et al. "The

19 words — < 1%

role of SARS-CoV-2 genomic surveillance and innovative analytical platforms for informing public health preparedness in Bengaluru, India.", Cold Spring Harbor Laboratory, 2022
Crossref Posted Content

14   Tamalika Ka, Utkarsh Narsaria, Srijita Basak, Debashrito De, Filippo Castiglion, David M. Mueller, Anurag P. Srivas. "A Candidate multi-epitope vaccine against SARS-CoV-2", Research Square Platform LLC, 2020
Crossref Posted Content

19 words — < 1%

15   ircmj.com
Internet

18 words — < 1%

16   Akshay Kanakan, Priyanka Mehta, Priti Devi, Sheeba Saifi et al. "Clinico-Genomic Analysis Reiterates Mild Symptoms Post-vaccination Breakthrough: Should We Focus on Low-Frequency Mutations?", Frontiers in Microbiology, 2022
Crossref

16 words — < 1%

17   Manoj Kumar, Sneha Gulati, Asgar Hussain Ansari, Rhythm Phutela et al. "RAY: CRISPR diagnostic for rapid and accurate detection of SARS-CoV2 variants on a paper strip", Cold Spring Harbor Laboratory, 2021
Crossref Posted Content

16 words — < 1%

18   neuro.unboundmedicine.com
Internet

15 words — < 1%

19   Ruei-Min Lu, Kang-Hao Liang, Hsiao-Ling Chiang, Fu-Fei Hsu et al. "Broadly neutralizing antibodies against Omicron variants of SARS-CoV-2 derived from mRNA-lipid nanoparticle-immunized mice", Cold Spring Harbor Laboratory, 2022
Crossref Posted Content

14 words — < 1%

20 www.mdpi.com
Internet   14 words — < 1%

21 Szendrey, Mark Graham. "Effects of Covid-19   10 words — < 1%
    Drugs Chloroquine, Hydroxychloroquine,
    Azithromycin and Remdesivir on hErg Expression and
    Function", Queen's University (Canada), 2021
    ProQuest

22 www.medrxiv.org
Internet   10 words — < 1%

23 ANNU NÄKKI, TAPIO VIDEMAN, URHO M. KUJALA,   9 words — < 1%
    MATTI SUHONEN et al. " Candidate Gene
    Association Study of Magnetic Resonance Imaging-based Hip
    Osteoarthritis (OA): Evidence for Gene as a Common
    Predisposing Factor for Hip OA and Lumbar Disc Degeneration
    ", The Journal of Rheumatology, 2011
    Crossref

24 Abhisek Kumar Behera, Sushmita Basu, Sarah S.   9 words — < 1%
    Cherian. "Molecular mechanism of the enhanced
    viral fitness contributed by secondary mutations in the
    hemagglutinin protein of oseltamivir resistant H1N1 influenza
    viruses: Modeling studies of antibody and receptor binding",
    Gene, 2015
    Crossref

25 Darshan Contractor, Christoph Globisch, Shiv   9 words — < 1%
    Swaroop, Alok Jain. "Structural basis of Omicron
    immune evasion: A comparative computational study",
    Computers in Biology and Medicine, 2022
    Crossref

26 E Sarvikivi, M Salminen, C Savolainen-Kopra, N 9 words — < 1%
    Ikonen et al. "COVID-19 cases in spectators
    returning to Finland from UEFA Euro 2020 matches in Saint

Petersburg", Epidemiology and Infection, 2022
Crossref

27 Krebs, Bruna Baumgarten, and Joelma Freire De Mesquita. "Amyotrophic Lateral Sclerosis Type 20 - In Silico Analysis and Molecular Dynamics Simulation of hnRNPA1", PLoS ONE, 2016.
Crossref
9 words — < 1%

28 www.nature.com
Internet
9 words — < 1%

29 Chiara Zagni, Andrea Citarella, Mahjoub Oussama, Antonio Rescifina et al. "Hydroxamic Acid-Based Histone Deacetylase (HDAC) Inhibitors Bearing a Pyrazole Scaffold and a Cinnamoyl Linker", International Journal of Molecular Sciences, 2019
Crossref
8 words — < 1%

30 Pandey, Bharati, Sonam Grover, Chetna Tyagi, Sukriti Goyal, Salma Jamal, Aditi Singh, Jagdeep Kaur, and Abhinav Grover. "Molecular principles behind pyrazinamide resistance due to mutations in panD gene in Mycobacterium tuberculosis", Gene, 2016.
Crossref
8 words — < 1%

31 www.frontiersin.org
Internet
8 words — < 1%

32 Somayye Taghvaei, Farzaneh Sabouni, Zarrin Minuchehr. "Identification of Natural Products as SENP2 Inhibitors for Targeted Therapy in Heart Failure", Frontiers in Pharmacology, 2022
Crossref
7 words — < 1%

33 Sun-Ah Cho, Katie Jo Rohn-Glowacki, Yazun B. Jarrar, Myeongjin Yi, Woo-Young Kim, Jae-Gook
7 words — < 1%

Shin, Su-Jun Lee. "Analysis of genetic polymorphism and biochemical characterization of a functionally decreased variant in prostacyclin synthase gene (CYP8A1) in humans", Archives of Biochemistry and Biophysics, 2015
Crossref

34 Nivya Mendon, Rayees Ganie, Shubham Kesarwani, Drisya Dileep, Sarika Sasi, Prakash Lama, Anchal Chandra, Minhajuddin Sirajuddin. "Peptide derived nanobody inhibits entry of SARS-CoV-2 variants", Cold Spring Harbor Laboratory, 2022   6 words — < 1%
Crossref Posted Content

35 Pedro F.N. Souza, Felipe P. Mesquita, Jackson L. Amaral, Patrícia G.C. Landim et al. "The spike glycoprotein of SARS-CoV-2: A review of how mutations of spike glycoproteins have driven the emergence of variants with high transmissibility and immune escape", International Journal of Biological Macromolecules, 2022   6 words — < 1%
Crossref

36 Waleemas Jairak, Ekkapat Chamsai, Kitikhun Udom, Kamonpan Charoenkul et al. "SARS-CoV-2 delta variant infection in domestic dogs and cats, Thailand", Scientific Reports, 2022   6 words — < 1%
Crossref

37 s3.amazonaws.com
Internet   6 words — < 1%

EXCLUDE QUOTES   OFF   EXCLUDE SOURCES   OFF
EXCLUDE BIBLIOGRAPHY   ON   EXCLUDE MATCHES   OFF