Title: Identification of genomic signatures and multiple lineage markers from the second and third wave samples of COVID-19 in Western Rajasthan, India

Naveen Prakash Bokolia[1*], Ravisekhar Gadepalli[1**]

[1]Viral Research and Diagnostic Laboratory, Microbiology Department, All India Institute of Medical Sciences, Jodhpur (342001), India

**Corresponding Author: Phone: +91-9680009207

Email: gadepallir@aiimsjodhpur.edu.in

[1]Viral Research and Diagnostic Laboratory, Microbiology Department, All India Institute of Medical Sciences, Jodhpur (342001), India

*Co-corresponding Author: Phone: +91-7807448895

Email: naveenbokolia@gmail.com

ABSTRACT

Most of the mutations occurred in SARS-CoV-2 are either relatively neutral or swiftly purged. However, some mutations have altered the functional aspects in terms of infectivity and transmission, host-viral interactions, disease severity and immune or vaccine escape. There are emerging evidence that certain mutations are jeopardizing the immune based therapies. The present research report is focused on the identification of genomic signatures of SARS-CoV-2 variant that caused mortality during second and third wave of COVID-19 in Western Rajasthan, India. We identified that Delta clade of SARS-CoV-2 is the predominant cause of mortality during second wave and even third wave in Western Rajasthan, India. Importantly, this study also revealed the unique and common substitution mutations within the spike domain, those are present in mortality and survived persons during the second and third wave of COVID-19 in India. In addition, this study also revealed the multiple lineage markers (Delta and Omicron), that would update with insightful understanding in the clade development of SARS-CoV-2.

Keywords: SARS-CoV-2, Genomic signatures, Mutations, Spike domain, Multiple lineage markers, Recombination, Delta and Omicron Variant

Introduction:

Coronaviruses are single stranded positive sense RNA viruses belong to Coronaviradeae family that caused COVID-19 disease (Kim et al., 2020; Zhu et al., 2020). The ongoing COVID-19 pandemic poses one of the greatest global threats in modern history and has already caused severe social and economic costs. The development of efficient and rapid sequencing methods to reconstruct the genomic sequence of SARS-CoV-2, the etiological agent of COVID-19, has been fundamental for the design of diagnostic molecular tests and to devise effective measures and strategies to mitigate the diffusion of the pandemic (Aleem et al., 2021; Banho et al., 2021; Ignatieva et al., 2022). Whole genome sequencing data of SARS-CoV-2 has important role in the elucidation of lineage of virus (with evolution), tracing the pathogenicity, transmission and spread within a particular state or local region (Ignatieva et al., 2022; Kannan et al., 2021; Lu et al., 2020; Naveca et al., 2021; Tegally et al., 2020; Washington et al., 2021; Zhan et al., 2020).

According to one CDC report (April, 2022), Omicron variant infection occurs among the 10 persons who already have had infection with Delta variant within 90 days. Delta variant remains predominant followed by overtaken by Omicron variant (Roskosky et al., 2022). In India, similar scenario happened, where the Omicron associated infections risen during Jan-Feb 2022.

Rajasthan state shares 5.66% population of total Indian population, and Jodhpur district has second largest population (after Jaipur) with the estimate of 4,276,374. Thus the present study is designed to provide a broad overview of SARS-CoV-2 variant specification that is/was circulated among the survived persons and hospitalized patients who died during the hospitalization at tertiary care unit of All India Institute of Medical Sciences, Jodhpur, India. Although, during third wave the number

of deaths were comparatively very less in India (Arnaout and Arnaout, 2022; Jha et al., 2022). However, it is important to determine the prime cause of infection with morbidity and mortality in a specified region, and to update the information about previous and current SARS-CoV-2 variants in circulation.

Within this perspective we categorized the SARS-CoV-2 samples in four different categories. First category involves “second wave survived”, second category involves “second wave hospitalized and deceased”, third category involves “third wave survived” and fourth category involves “third wave hospitalized and deceased”. The comparative analysis (between different categories) of genomic signatures was done with respect to spike domain of SARS-CoV-2. The present study revealed that Delta is the causative agent of mortality in third wave of COVID-19 as well, and not the Omicron variant of SARS-CoV-2. In addition, the multiple lineages are detected from the second wave samples of COVID-19, those might be important in determining the genomic aspects of clade development of SARS-CoV-2. This study updates and defines the unique and common mutations in the RNA genome of SARS-CoV-2 variants.

Materials and Methods:

Categorization of SARS-CoV-2 positive samples used in this study:

This study was designed to identify SARS-CoV-2 variants and genomic signatures from the second and third wave of COVID-19 patients those were died during hospitalization at tertiary care unit of AIIMS Hospital, Jodhpur, Rajasthan. This study does not involve specific parameters those are based on gender, age, sex or associated disease complications. In the present investigation already stored samples during a particular COVID-19 wave (peak phase) was used, however only samples those fit in to the particular Ct value criteria (RNA integrity) were processed further for sequencing. We involved 50 SARS-CoV-2 positive samples those were collected from non-hospitalized individuals, 50 samples from hospitalized and died patients during second wave of COVID-19. Similarly, we involved 35 SARS-CoV-2 positive samples those were collected during third wave of COVID-19, and 16 samples from the patients those died during hospitalization at tertiary care unit of AIIMS Jodhpur, Rajasthan. We could involve only 16 samples from third wave mortality because of the lower cases of hospitalization and mortality during third wave. In addition, a parallel comparison was done with SARS-CoV-2 genome sequencing data obtained from the samples of non-hospitalized (survived) COVID-19 positive individuals.

Real Time RT-PCR of SARS-CoV-2 positive samples:

The stored SARS-CoV-2 positive samples (-80^{o}C storage) were taken and properly thawed at room temperature. The total RNA of samples was extracted by using QIAamp Viral RNA kit. In order to proceed for sequencing experiments, we determined the Ct value of extracted RNA samples by real time RT-PCR. The positive samples

those fall within the range of 20-25 Ct values were further proceeded for whole genome sequencing.

Whole genome sequencing of SARS-CoV-2 on MiSeq system (Illumina platform):

The whole genome sequencing of SARS-CoV-2 was performed on MiSeq System (Illumina technology), that is installed at R'VRDL, Microbiology Department, AIIMS Jodhpur. The library preparation was done by COVIDSeq Assay Kit and the protocol was followed according to manufacturer's protocol instructions. The library was quantified on Qubit fluorimeter and 12 Pico molar concentration of final library was loaded for sequencing run. MiSeq Reagent Kit v3 (150-cycle) was used for sequencing run after final library loading. After the completion of sequencing run sequencing data was obtained in FASTQ format and subsequently used for analysis.

Analysis of sequencing data for clade verification, mutational profiling, genomic signatures and sharing the data at GISAID:

The initial analysis was performed by using the Illumina® DRAGEN COVID Lineage App. These mutations were listed in Table 2. Subsequently FASTA files were used for mutational analysis, and was done on CoVsurver mutations app-GISAID that delivers total number and types of mutations in spike domain of SARS-CoV-2. The clade analysis (assignment), verification, and phylogenetic tree was constructed by Nextclade v1.14.1 software. The sequencing data of all the samples relevant to this study has been shared on GISAID and accession IDs have been provided in supplementary data file (Khare et al., 2021).

Phylogenetic analysis of multiple lineage variants:

In this analysis, we involved six FASTA sequences of multiple lineage variants, two FASTA sequences of B.1.617.2 lineage, two FASTA sequences of AY.122 lineage and two FASTA sequences of B.1 lineage. The FASTA files of multiple lineage variants were analysed by using NGPhylogeny.fr tool. The analysis was done in "one-click'' mode(Lemoine et al., 2019). The resultant phylogenetic tree was exported, and presented as Figure 1.

Results:

B.1.617.2 and AY.122 lineages (Delta clade) remain persistent during third wave of COVID-19 in Western Rajasthan

In the present study report we investigated second wave (Delta wave), third wave and post third wave SARS-CoV-2 positive samples. Samples from respective wave was divided in to non-hospitalized and hospitalized (with mortality) category, and comparative variant analysis was done. We involved 16 SARS-CoV-2 positive samples (available) in the category of "hospitalized and died during third wave". Among them, in 14 samples we detected Delta and Delta plus variant that is categorized in B.1.617.2 and AY.122 lineages respectively. Remaining two patients were found to be infected with Kappa (B.1.617.1) and unassigned Omicron respectively. Sequencing output from the "third wave mortality patients" revealed that B.1.617.2 and AY.122 lineages of Delta clade remain the most persistent variants of SARS-CoV-2 during the third wave of COVID-19. In addition, study also revealed that

Delta variant of SARS-CoV-2 still remain the causative agent of death instead of Omicron variant that has been characterized during Jan-2022 to March-2022 (third Wave in India). Apart from mortality cases, the sequencing results from non-hospitalized COVID-19 individuals (third wave) revealed the Omicron variant of BA.1.1 and BA.2 lineages mainly (Table 3). The present research report indicates that during third wave of Omicron Clade, 14 SARS-CoV-2 patients died due to Delta variant instead of Omicron variant (Table 1).

Several studies have suggested that previously infected individuals with earlier circulating SARS-CoV-2 may have reduced protection from reinfection with the B.1.351 variant.

Identification of persistent or stable mutations within Delta clade during post COVID-19 phase:

The Variant of Concern (VOC), Delta, was first reported in India in December 2020, and was responsible for the deadly second wave in April 2021. It was then reported in USA in March 2021, and became the most dominant variant there, and had been detected throughout the world including UK. Studies reported that specific mutations in the spike domain makes the properties of virus deleterious in terms of infectivity, disease severity and pathogenicity. In this regard, there are several mutations those remain conserved during different waves of COVID-19 with respect to particular variant of SARS-CoV-2. These three categories were put in to the context of analysis because most of the samples belong to Delta clade only. It can be easily observed that the number of mutations decreased with the time in Delta clade. Even though the patients were died in third wave due to infection from the B.1.617.2 lineage and AY.122 lineage, however, the number of mutations in spike domain is significantly lower in comparison to second wave patient samples (including both mortality and survived). It indicates that the core mutations in spike domain remain invariable in a manner where the RNA biology of virus does not change in terms of infectivity and pathogenicity. These core mutations could be further helpful in the development of particular drug or vaccine, while considering the mutations in Delta clade in current scenario as well.

The comparative analysis between “second wave non-hospitalized”, “second wave Hospitalized (mortality)” and “third wave Hospitalized (mortality)” leads to identification of unique amino- acid substitution mutations within the spike domain of SARS-CoV-2. We characterized these mutations as “unique substitution mutations in spike domain” in the Delta clade, as these are unique in the group of mortality and survived patients of second wave and third wave in India.

Present study revealed that Delta variant of SARS-CoV-2 remained dominant in terms of pathogenicity and infectivity irrespective of ongoing emerging variants, specifically in Western Rajasthan, India.

Identification of multiple lineages from the second wave samples of COVID-19:

The present study involved sequencing data analysis of 100 SARS-CoV-2 positive samples from the second wave of COVID-19 in India (2021), those belonged to deceased and survived patients. Out of 100 samples, sequencing results from the

eight samples revealed the mixed mutations from the both Delta and Omicron variants (Table 4). These submissions are marked with "*ML*'' and belong to multiple lineages. The identified (listed) mutations are highlighted in bold in Table 4. These mutations were found to be present in the spike domain. Apparently, some of the Omicron specific mutations have been known/studied for the potential impact on SARS-CoV-2 pathogenesis. Within this context, Omicron specific mutations N679K and H655Y are present at the interface/junction of the S1 and S2 subunit, and both are associated with increased infectivity. In addition, deletion mutations at N-terminal domain (H69-V70) are associated with increased infectivity, antibody escape and diagnostic escape.

Importantly, mutations in receptor binding domain have also been found in one mixed lineage (EPI_ISL_14797957) and was characterized as BA.1 Pango lineage. This lineage bears some of the defining mutations of Delta variant in spike domain: L452R, P681R and T478K, however other major mutations like T19R and 156del and 157del mutations were not present. Same mixed lineage also bears Omicron specific substitutions: G339D, S373P, S375F those could be involved in increased infectivity, higher immune escape and transmissibility (Alkhatib et al., 2022; Cao et al., 2022; Liu et al., 2022; Long et al., 2020; Ou et al., 2022). Previous research reports mentioned about the G339D substitution mutation as rare mutation in SARS-CoV-2 genome database before the detection of Omicron, and this mutation is associated with significant increase in binding affinity with the ACE2 receptor (Alkhatib et al., 2022; Starr et al., 2020). Other than these, the significance of spike domain mutations like A67V and G142D in relation to the possible impact of SARS-CoV-2 pathogenesis has not been demonstrated.

In order to get insights on the evolution pattern of mixed lineages, we analysed the sequences of mixed lineages along with B.1, B.617.2 and AY.122 lineages. Since, the mixed lineage samples derived from the second wave of COVID-19, therefore, we purposefully involved first wave and second wave FASTA sequences during the analysis of phylogenetic tree. The mixed lineage indicates the possible recombination events between the GK and G clade variants (Figure 1). Subsequently, these events could have occurred with the RNA genome of different SARS-CoV-2 variants. The hierarchy pattern of mixed lineage SARS-CoV-2 clade development indicates that amino acid substitutions might have been started during the second wave (Delta wave) of COVID-19 in India that later on established as distinct Omicron clade.

Discussion:

The fourth VOC, Delta, was first reported in India in December 2020, and was responsible for the deadly second wave of COVID-19 (Jha et al., 2022). It was then reported in USA in March 2021, and became the most dominant variant there, and had spread throughout the world including UK. The lineage is comprised of three subtypes: B.1.617.1, B.1.617.2 and B.1.617.3, with diverse mutations in NTD and RBD of spike protein with higher immune evasion potentials. The B.1.617.2 and B.1.617.3 variants differ from each other mainly by two mutations: T478K and E484Q in the RBD. The Delta variant was described more transmissible than the viruses causing SARS, MERS and Ebola as well as influenza and common cold, and as contagious as chickenpox by CDC (https://www.yalemedicine.org/news/5-things-to-know-delta-

variant-covid). It was reported that there was the rise in number of cases (10 fold) and mortality (3 fold) in India, with the emergence of Delta variant, displacing alpha and kappa lineages (Dhar et al., 2021). A Scottish study demonstrated that the risk of hospitalisation was higher (double) in Delta infections as compared to Alpha, in non-vaccinated groups (Sheikh et al., 2021). These are combining perspectives that inform about the underlying reason of persistent impact of Delta clade that remains during third wave of COVID-19 as well. In present study, this impact was identified in terms of mortality in COVID-19 patients during the hospitalization at tertiary care unit in Western Rajasthan, India. Whereas Omicron variant was reported to be of lower risk in terms of serious disease, hospitalization and mortality in India and other countries (South Africa, England, Scotland, Denmark and Canada) (Espenhain et al., 2021; Singhal, 2022; Wolter et al., 2021).

In addition, present study identifies the unique amino acid substitution mutations in the mortality group of patients in Delta clade, that could be of further consideration and helpful for future vaccine development. For example, the unique amino acid substitution mutations can be used in a cumulative manner or in an independent manner, according to further docking or experimental outcome that results in increase in affinity or any significant observation. In addition, the reduction in number of mutations in spike domain is described among the group of “third wave mortality with hospitalization”. These markedly reduced number of mutations in spike domain further signifies that RNA biology of SARS-CoV-2 remains unchanged depending on the type and respective context of mutation, however a significant number of core mutations (that defines a particular Clade or lineage of variant) must be present in order to propagate at same pathogenicity and infectivity.

This study revealed that mixed lineage bearing the amino acid substitutions (as markers) from the both Delta and Omicron variants. The role of amino acid substitutions has been identified in the Omicron specific variant that modulates the potential impact of SARS-CoV-2 pathogenesis (Alkhatib et al., 2022; Ou et al., 2022). These findings indicates the possibility that during the infection cycle, the RNA genome of SARS-CoV-2 may have undergone through the spontaneous mutations and deletions events. In addition, the second probability could be of recombination events between the RNA genome of GK and G clade variants of SARS-CoV-2. The probability of recombination events between Alpha and Delta variants could be higher because the mixed lineages have the amino acids substitutions from the both Delta and Omicron variants. Since Omicron variant has more mutations then Alpha, Beta, Gamma and Delta, those promotes stronger immune escape and accelerated transmission (Xu et al., 2022). The accumulation of increased mutations signify the impact and possibility of recombination events. Some of the recent research reports have provided the evidence of possible intervariant and intravariant recombination of Delta and Omicron variants (Focosi and Maggi, 2022; Ou et al., 2022; VanInsberghe et al., 2021). One study provided the clear evidence of iterlineage recombination events between B.1.1.7 and B.1.177 lineages, where recombinant variant had spike gene from B.1.1.7 variant (Jackson et al., 2021). In addition, recent studies point out towards the human chronic infections and/ or animal reservoirs could be the contributing factor in the evolution of RNA genome of SARS-CoV-2 (Tegally et al.,

2022). This study signifies the identification of mixed lineage markers from the second wave of COVID-19 in India. These markers (amino acid substitutions) would be insightful and significant in understanding the evolution pattern of the SARS-CoV-2 RNA genome within the context of recombination events.

Conflict of Interest:

The authors declare that there are no conflicts of interest.

Data Availability:

All the sequencing data relevant to article, has been submitted on GISAID (Khare et al., 2021). The supplementary file of respective sequences (accession IDs) has been made available as supplementary data.

Acknowledgment:

We are very grateful and thankful for becoming the part of INSACOG. We are very grateful to NIV Pune, as a mentor institute of AIIMS Jodhpur, and providing the sequencing kits for the whole genome sequencing of SARS-CoV-2.

Aleem, A., Akbar Samad, A.B., Slenker, A.K., 2021. Emerging Variants of SARS-CoV-2 And Novel Therapeutics Against Coronavirus (COVID-19), StatPearls. © 2021, StatPearls Publishing LLC., Treasure Island FL.

Alkhatib, M., Salpini, R., Carioti, L., Ambrosio, F.A., D'Anna, S., Duca, L., Costa, G., Bellocchi, M.C., Piermatteo, L., Artese, A., Santoro, M.M., Alcaro, S., Svicher, V., Ceccherini-Silberstein, F., 2022. Update on SARS-CoV-2 Omicron Variant of Concern and Its Peculiar Mutational Profile. Microbiology Spectrum 10, e02732-02721.

Arnaout, R., Arnaout, R., 2022. Visualizing omicron: COVID-19 deaths vs. cases over time. PLOS ONE 17, e0265233.

Banho, C.A., Sacchetto, L., Campos, G.R.F., Bittar, C., Possebon, F.S., Ullmann, L.S., Marques, B.d.C., da Silva, G.C.D., Moraes, M.M., Parra, M.C.P., Negri, A.F., Boldrin, A.C., Barcelos, M.D., dos Santos, T.M.I.L., Milhim, B.H.G.A., da Rocha, L.C., Dourado, F.S., dos Santos, A.L., Ciconi, V.B., Patuto, C., Versiani, A.F., da Silva, R.A., Lobl, E.E.d.O., Hernandes, V.M., Zini, N., Pacca, C.C., Estofolete, C.F., Ferreira, H.L., Rahal, P., Araújo, J.P., Cohen, J.A., Kerr, C.C., Althouse, B.M., Vasilakis, N., Nogueira, M.L., 2021. Effects of SARS-CoV-2 P.1 introduction and the impact of COVID-19 vaccination on the epidemiological landscape of São José Do Rio Preto, Brazil. medRxiv, 2021.2007.2028.21261228.

Cao, Y., Wang, J., Jian, F., Xiao, T., Song, W., Yisimayi, A., Huang, W., Li, Q., Wang, P., An, R., 2022. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature 602, 657-663.

Dhar, M.S., Marwal, R., Vs, R., Ponnusamy, K., Jolly, B., Bhoyar, R.C., Sardana, V., Naushin, S., Rophina, M., Mellan, T.A., Mishra, S., Whittaker, C., Fatihi, S., Datta, M., Singh, P., Sharma, U., Ujjainiya, R., Bhatheja, N., Divakar, M.K., Singh, M.K., Imran, M., Senthivel, V., Maurya, R., Jha, N., Mehta, P., A, V., Sharma, P., Vr, A., Chaudhary, U., Soni, N., Thukral, L., Flaxman, S., Bhatt, S., Pandey, R., Dash, D., Faruq, M., Lall, H., Gogia, H., Madan, P., Kulkarni, S., Chauhan, H., Sengupta, S., Kabra, S., Gupta, R.K., Singh, S.K., Agrawal, A., Rakshit, P., Nandicoori, V., Tallapaka, K.B., Sowpati, D.T., Thangaraj, K., Bashyam, M.D., Dalal, A., Sivasubbu, S., Scaria, V., Parida, A., Raghav, S.K., Prasad, P., Sarin, A., Mayor, S., Ramakrishnan, U., Palakodeti, D., Seshasayee, A.S.N., Bhat, M., Shouche, Y., Pillai, A., Dikid, T., Das, S., Maitra, A., Chinnaswamy, S., Biswas, N.K., Desai, A.S., Pattabiraman, C., Manjunatha, M.V., Mani,

R.S., Arunachal Udupi, G., Abraham, P., Atul, P.V., Cherian, S.S., 2021. Genomic characterization and epidemiology of an emerging SARS-CoV-2 variant in Delhi, India. Science, eabj9932.

Espenhain, L., Funk, T., Overvad, M., Edslev, S.M., Fonager, J., Ingham, A.C., Rasmussen, M., Madsen, S.L., Espersen, C.H., Sieber, R.N., Stegger, M., Gunalan, V., Wilkowski, B., Larsen, N.B., Legarth, R., Cohen, A.S., Nielsen, F., Lam, J.U.H., Lavik, K.E., Karakis, M., Spiess, K., Marving, E., Nielsen, C., Wiid Svarrer, C., Bybjerg-Grauholm, J., Olsen, S.S., Jensen, A., Krause, T.G., Müller, L., 2021. Epidemiological characterisation of the first 785 SARS-CoV-2 Omicron variant cases in Denmark, December 2021. Eurosurveillance 26, 2101146.

Focosi, D., Maggi, F., 2022. Recombination in Coronaviruses, with a Focus on SARS-CoV-2. Viruses 14.

Ignatieva, A., Hein, J., Jenkins, P.A., 2022. Ongoing Recombination in SARS-CoV-2 Revealed through Genealogical Reconstruction. Molecular Biology and Evolution 39.

Jackson, B., Boni, M.F., Bull, M.J., Colleran, A., Colquhoun, R.M., Darby, A.C., Haldenby, S., Hill, V., Lucaci, A., McCrone, J.T., Nicholls, S.M., O'Toole, Á., Pacchiarini, N., Poplawski, R., Scher, E., Todd, F., Webster, H.J., Whitehead, M., Wierzbicki, C., Loman, N.J., Connor, T.R., Robertson, D.L., Pybus, O.G., Rambaut, A., 2021. Generation and transmission of interlineage recombinants in the SARS-CoV-2 pandemic. Cell 184, 5179-5188.e5178.

Jha, P., Deshmukh, Y., Tumbe, C., Suraweera, W., Bhowmick, A., Sharma, S., Novosad, P., Fu, S.H., Newcombe, L., Gelband, H., Brown, P., 2022. COVID mortality in India: National survey data and health facility deaths. Science 375, 667-671.

Kannan, S.R., Spratt, A.N., Cohen, A.R., Naqvi, S.H., Chand, H.S., Quinn, T.P., Lorson, C.L., Byrareddy, S.N., Singh, K., 2021. Evolutionary analysis of the Delta and Delta Plus variants of the SARS-CoV-2 viruses. Journal of autoimmunity 124, 102715.

Khare, S., Gurry, C., Freitas, L., Schultz, M.B., Bach, G., Diallo, A., Akite, N., Ho, J., Lee, R.T., Yeo, W., Curation Team, G.C., Maurer-Stroh, S., 2021. GISAID's Role in Pandemic Response. China CDC weekly 3, 1049-1051.

Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N., Chang, H., 2020. The Architecture of SARS-CoV-2 Transcriptome. Cell 181, 914-921 e910.

Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., Gascuel, O., 2019. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. Nucleic acids research 47, W260-W265.

Liu, L., Iketani, S., Guo, Y., Chan, J.F.-W., Wang, M., Liu, L., Luo, Y., Chu, H., Huang, Y., Nair, M.S., 2022. Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. Nature 602, 676-681.

Long, S.W., Olsen, R.J., Christensen, P.A., Bernard, D.W., Davis, J.J., Shukla, M., Nguyen, M., Saavedra, M.O., Yerramilli, P., Pruitt, L., 2020. Molecular architecture of early dissemination and massive second wave of the SARS-CoV-2 virus in a major metropolitan area. MBio 11, e02707-02720.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L., Chen, J., Meng, Y., Wang, J., Lin, Y., Yuan, J., Xie, Z., Ma, J., Liu, W.J., Wang, D., Xu, W., Holmes, E.C., Gao, G.F., Wu, G., Chen, W., Shi, W., Tan, W., 2020. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet (London, England) 395, 565-574.

Naveca, F.G., Nascimento, V., Souza, V., de Lima Corado, A., Nascimento, F., Silva, G., Mejía, M., Brandão, M.J., Costa, Á., Duarte, D., Pessoa, K., Jesus, M., Gonçalves, L., Fernandes, C., Mattos, T., Abdalla, L., Santos, J.H., Martins, A., Chui, F.M., Val, F.F., Melo, G.C.d., Simão, M.X., Sampaio, V.d.S., Mourão, M.P., Lacerda, M.V., Batista, É.L.R., Magalhães, A.L.Á., Dábilla, N., Pereira, L.C.G., Vinhal, F., Miyajima, F., Dias, F.B.S., dos Santos, E.R., Coêlho, D., Ferraz, M., Lins, R., Wallau, G.L., Delatorre, E., Gräf, T., Siqueira, M.M., Resende, P.C., Bello, G., 2021. Spread of Gamma (P.1) sub-lineages carrying Spike mutations close to the furin cleavage site and deletions in the N-terminal domain drives ongoing transmission of SARS-CoV-2 in Amazonas, Brazil. medRxiv, 2021.2009.2012.21263453.

Ou, J., Lan, W., Wu, X., Zhao, T., Duan, B., Yang, P., Ren, Y., Quan, L., Zhao, W., Seto, D., Chodosh, J., Luo, Z., Wu, J., Zhang, Q., 2022. Tracking SARS-CoV-2 Omicron diverse spike gene mutations identifies multiple inter-variant recombination events. Signal Transduction and Targeted Therapy 7, 138.

385 Roskosky, M., Borah, B.F., DeJonge, P.M., Donovan, C.V., Blevins, L.Z., Lafferty, A.G., Pringle, J.C., Kelso,
386 P., Temte, J.L., Temte, E., Barlow, S., Goss, M., Uzicanin, A., Bateman, A., Florek, K., Kawakami, V.,
387 Lewis, J., Loughran, J., Pogosjans, S., Kay, M., Duchin, J., Lunn, S., Schnitzler, H., Arora, S., Tate, J.,
388 Ricaldi, J., Kirking, H., 2022. Notes from the Field: SARS-CoV-2 Omicron Variant Infection in 10 Persons
389 Within 90 Days of Previous SARS-CoV-2 Delta Variant Infection - Four States, October 2021-January
390 2022. MMWR. Morbidity and mortality weekly report 71, 524-526.
391 Sheikh, A., McMenamin, J., Taylor, B., Robertson, C., 2021. SARS-CoV-2 Delta VOC in Scotland:
392 demographics, risk of hospital admission, and vaccine effectiveness. The Lancet 397, 2461-2462.
393 Singhal, T., 2022. The Emergence of Omicron: Challenging Times Are Here Again! Indian Journal of
394 Pediatrics 89, 490-496.
395 Starr, T.N., Greaney, A.J., Hilton, S.K., Ellis, D., Crawford, K.H., Dingens, A.S., Navarro, M.J., Bowen, J.E.,
396 Tortorici, M.A., Walls, A.C., 2020. Deep mutational scanning of SARS-CoV-2 receptor binding domain
397 reveals constraints on folding and ACE2 binding. Cell 182, 1295-1310. e1220.
398 Tegally, H., Moir, M., Everatt, J., Giovanetti, M., Scheepers, C., Wilkinson, E., Subramoney, K., Makatini,
399 Z., Moyo, S., Amoako, D.G., Baxter, C., Althaus, C.L., Anyaneji, U.J., Kekana, D., Viana, R., Giandhari, J.,
400 Lessells, R.J., Maponga, T., Maruapula, D., Choga, W., Matshaba, M., Mbulawa, M.B., Msomi, N.,
401 Bester, A.P., Claassen, M., Doolabh, D., Mudau, I., Mbhele, N., Engelbrecht, S., Goedhals, D., Hardie,
402 D., Hsiao, N.-Y., Iranzadeh, A., Ismail, A., Joseph, R., Maharaj, A., Mahlangu, B., Mahlakwane, K., Davis,
403 A., Marais, G., Mlisana, K., Mnguni, A., Mohale, T., Motsatsi, G., Mwangi, P., Ntuli, N., Nyaga, M.,
404 Olubayo, L., Radibe, B., Ramphal, Y., Ramphal, U., Strasheim, W., Tebeila, N., van Wyk, S., Wilson, S.,
405 Lucaci, A.G., Weaver, S., Maharaj, A., Pillay, Y., Davids, M., Mendes, A., Mayaphi, S., Naidoo, Y., Pillay,
406 S., Sanko, T.J., San, J.E., Scott, L., Singh, L., Magini, N.A., Smith-Lawrence, P., Stevens, W., Dor, G.,
407 Tshiabuila, D., Wolter, N., Preiser, W., Treurnicht, F.K., Venter, M., Chiloane, G., McIntyre, C., O'Toole,
408 A., Ruis, C., Peacock, T.P., Roemer, C., Kosakovsky Pond, S.L., Williamson, C., Pybus, O.G., Bhiman, J.N.,
409 Glass, A., Martin, D.P., Jackson, B., Rambaut, A., Laguda-Akingba, O., Gaseitsiwe, S., von Gottberg, A.,
410 de Oliveira, T., consortium, N.-S., 2022. Emergence of SARS-CoV-2 Omicron lineages BA.4 and BA.5 in
411 South Africa. Nature Medicine 28, 1785-1790.
412 Tegally, H., Wilkinson, E., Giovanetti, M., Iranzadeh, A., Fonseca, V., Giandhari, J., Doolabh, D., Pillay,
413 S., San, E.J., Msomi, N., Mlisana, K., von Gottberg, A., Walaza, S., Allam, M., Ismail, A., Mohale, T.,
414 Glass, A.J., Engelbrecht, S., Van Zyl, G., Preiser, W., Petruccione, F., Sigal, A., Hardie, D., Marais, G.,
415 Hsiao, M., Korsman, S., Davies, M.-A., Tyers, L., Mudau, I., York, D., Maslo, C., Goedhals, D., Abrahams,
416 S., Laguda-Akingba, O., Alisoltani-Dehkordi, A., Godzik, A., Wibmer, C.K., Sewell, B.T., Lourenço, J.,
417 Alcantara, L.C.J., Pond, S.L.K., Weaver, S., Martin, D., Lessells, R.J., Bhiman, J.N., Williamson, C., de
418 Oliveira, T., 2020. Emergence and rapid spread of a new severe acute respiratory syndrome-related
419 coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. medRxiv,
420 2020.2012.2021.20248640.
421 VanInsberghe, D., Neish, A.S., Lowen, A.C., Koelle, K., 2021. Recombinant SARS-CoV-2 genomes are
422 currently circulating at low levels. bioRxiv.
423 Washington, N.L., Gangavarapu, K., Zeller, M., Bolze, A., Cirulli, E.T., Barrett, K.M.S., Larsen, B.B.,
424 Anderson, C., White, S., Cassens, T., Jacobs, S., Levan, G., Nguyen, J., Ramirez, J.M., Rivera-Garcia, C.,
425 Sandoval, E., Wang, X., Wong, D., Spencer, E., Robles-Sikisaka, R., Kurzban, E., Hughes, L.D., Deng, X.,
426 Wang, C., Servellita, V., Valentine, H., De Hoff, P., Seaver, P., Sathe, S., Gietzen, K., Sickler, B., Antico,
427 J., Hoon, K., Liu, J., Harding, A., Bakhtar, O., Basler, T., Austin, B., Isaksson, M., Febbo, P.G., Becker, D.,
428 Laurent, M., McDonald, E., Yeo, G.W., Knight, R., Laurent, L.C., de Feo, E., Worobey, M., Chiu, C.,
429 Suchard, M.A., Lu, J.T., Lee, W., Andersen, K.G., 2021. Genomic epidemiology identifies emergence
430 and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. medRxiv,
431 2021.2002.2006.21251159.
432 Wolter, N., Jassat, W., Walaza, S., Welch, R., Moultrie, H., Groome, M., Amoako, D.G., Everatt, J.,
433 Bhiman, J.N., Scheepers, C., Tebeila, N., Chiwandire, N., du Plessis, M., Govender, N., Ismail, A., Glass,
434 A., Mlisana, K., Stevens, W., Treurnicht, F.K., Makatini, Z., Hsiao, N.-y., Parboosing, R., Wadula, J.,

435 Hussey, H., Davies, M.-A., Boulle, A., von Gottberg, A., Cohen, C., 2021. Early assessment of the clinical
436 severity of the SARS-CoV-2 Omicron variant in South Africa. medRxiv, 2021.2012.2021.21268116.
437 Xu, A., Hong, B., Lou, F., Wang, S., Li, W., Shafqat, A., An, X., Zhao, Y., Song, L., Tong, Y., Fan, H., 2022.
438 Sub-lineages of the SARS-CoV-2 Omicron variants: Characteristics and prevention. MedComm 3, e172.
439 Zhan, X.-Y., Zhang, Y., Zhou, X., Huang, K., Qian, Y., Leng, Y., Yan, L., Huang, B., He, Y., 2020. Molecular
440 Evolution of SARS-CoV-2 Structural Genes: Evidence of Positive Selection in Spike Glycoprotein.
441 bioRxiv, 2020.2006.2025.170688.
442 Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan,
443 F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G.F., Tan, W., 2020. A Novel Coronavirus from Patients with
444 Pneumonia in China, 2019. The New England journal of medicine 382, 727-733.

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

Figure Legend:

Figure 1:

The presented phylogenetic tree indicates the possibility of evolution of mixed lineages due to intervariant recombination events between the RNA genome of SARS-CoV-2. The analysis also includes FASTA sequences of first wave and second wave variants. The GISAID accession IDs of first wave variants are EPI_ISL_14822459 and EPI_ISL_14822474. The GISAID accession IDs of second wave variants (lineage B.1.617.2 and AY.122) are EPI_ISL_14822481, EPI_ISL_14822482, EPI_ISL_14822497 and EPI_ISL_14822498 respectively. Remaining six sequences belong to multiple lineages. The phylogenetic tree indicates that RNA genome of B.1.617.2 lineage and B.1 lineage undergoes recombination events, subsequently multiple recombination events could have occurred that leads to development of SARS-CoV-2 clade.

TABLES:

	Different category of SARS-CoV-2 samples used in the present investigation			
Type of SARS-CoV-2 lineage	**Second Wave Non-Hospitalized (May 2021)**	**Second Wave Hospitalized Patients with mortality (May-June 2021)**	**Third Wave Non-Hospitalized Individuals (March-May 2022)**	**Third Wave Hospitalized Patients with mortality (March-May 2022)**
	Total number of samples used and respective numbers belong to particular lineage			
	50	**50**	**35**	**16**
B.1.1	1	3	Not Detected	Not Detected
B.1.617.1	1	4	Not Detected	1
B.1.617.2	21	21	Not Detected	7
AY.122	20	14	Not Detected	7
AY.38	4	4	Not Detected	Not Detected
AY.46	Not Detected	1	Not Detected	

BA.2	Not Detected	Not Detected	9	Not Detected
BA.1.1 BA.1.17 BA.1.18 (Omicron: BA.1-like)	Not Detected	Not Detected	26	Not Detected
Unassigned (Omicron)	3	3	Not Detected	1
Unassigned (Delta)	Not Detected	Not Detected	Not Detected	Not Detected

478

479 **Table 1: Table represent the number of samples belong to particular lineage of SARS-CoV-2, during the second and third**
480 **wave of COVID-19.**

481

Lineage	**Total Representative mutations in spike domain of SARS-CoV-2 during second wave (associated with survived COVID-19 positive individuals)**	**Total Representative mutations in spike domain of SARS-CoV-2 during second wave (associated with hospitalization and mortality)**	**Total representative mutations in spike domain of SARS-CoV-2 during third wave (associated with hospitalization and mortality)**
AY.122 Lineage	**T19R E156G F157del R158del** G446V **L452R T478K D614G P681R D950N**	**T19R** R21I A67V H69del V70del T95I **G142D** V143del Y144del Y145del **E156G F157del R158del** T307I **L452R T478K D614G** Q675H **P681R**(674) **D950N** N969K	**T19R** ins108AV **G142D E156G F157del R158del** N185D **L452R T478K D614G P681R D950N**

B.1.617.2 Lineage	**T19R A67V H69del V70del T95I** E132D E156G **F157del R158del A222V** G446V **L452R T478K** E583D **D614G** H655Y N679K P681H **P681R** A845S N856K **D950N** Q954H **N969K** L981F	**T19R** T29A **A67V H69del V70del K77T T95I** S98F G142D W152L **E156G F157del R158del V213G A222V L452R T478K** Q613H **D614G P681R** N764K D796Y **D950N N969K** L981F A1070S	**T19R A67V H69del V70del K77T T95I E156G F157del R158del V213G A222V L452R T478K D614G P681R D950N N969K**

482

483 **Table 2:** The table represents the total number of mutations present in spike domain within each categorized set of samples.

484

	Total Representative mutations in spike domain of SARS-CoV-2 during third wave (COVID-19 positive) associated with BA.2.38 Variant	Total Representative mutations in spike domain of SARS-CoV-2 during third wave (COVID-19 positive) associated with BA.2 Variant
Mutations in Spike domain	**T19I L24del P25del P26del A27S G142D** V213G **G339D S371F S373P S375F T376A D405N R408S K417N** K417T Q498R N501Y Y505H N556K **D614G** V642G **H655Y N679K P681H N764K D796Y Q954H N969K**	**T19I L24del P25del P26del A27S L84I G142D** N185D V213G Y248N **G339D** R346T **S371F S373P S375F T376A D405N R408S K417N** N440K G446S L452M N460K S477N T478K E484A Q493R Q498R N501Y Y505H **D614G H655Y** T676A **N679K P681H N764K D796Y Q954H N969K**

485

486 **Table 3:** Table presents the total number of substitution mutations present in a particular lineage during third wave of COVID-19.

487

Virus Name	Accession ID	Clade	Pango Lineage	Amino Acid Substitutions:	Variant
hCoV-19/India/RJ-INSACOG-AIIMSJ-193799/2022	EPI_ISL_14806594	GK	AY.122 (Pango v.4.1.2 PLEARN-v1.12), Delta (B.1.617.2-like) (Scorpio)	**Spike A67V**, Spike D614G, ***Spike D950N***, **Spike H69del**, ***Spike L452R***, Spike **N969K**, Spike P681R, ***Spike T19R***, **Spike T95I**, **Spike T478K**, **Spike V70del**, M I82T, N D63G, N D377Y, NS3 D238Y, NS3 S26L, NS3 S60Y, NS6 T10I, NS7b T40I, NSP2 K81N, NSP3 A488S, NSP4 T492I, NSP4 V167L, NSP12 G671S, NSP12 P323L, NSP13 P77L, NSP14 A394V	VOC Delta GK (B.1.617.2+AY.*) first detected in India
hCoV-19/India/RJ-INSACOG-AIIMSJ-135613/2021	EPI_ISL_14806565	GRA	B.1.1 (Pango v.4.1.2 PUSHER-v1.12), Probable Omicron (Unassigned) (Scorpio)	Spike D614G, **Spike D796Y**, **Spike N764K**, Spike N856K, **Spike N969K**, ***Spike P681R***, **Spike Q954H**, M I82T, N A254T, N D63G, N E31del, N G204R, N P13L, N R32del, N R203K, N S33del, NS3 S26L, NSP1 A79T, NSP3 K38R, NSP4 T492I, NSP5 P132H, NSP6 G107del, NSP6 L105del, NSP6 S106del, NSP6 T181I, NSP6 V149A, NSP12 P323L	

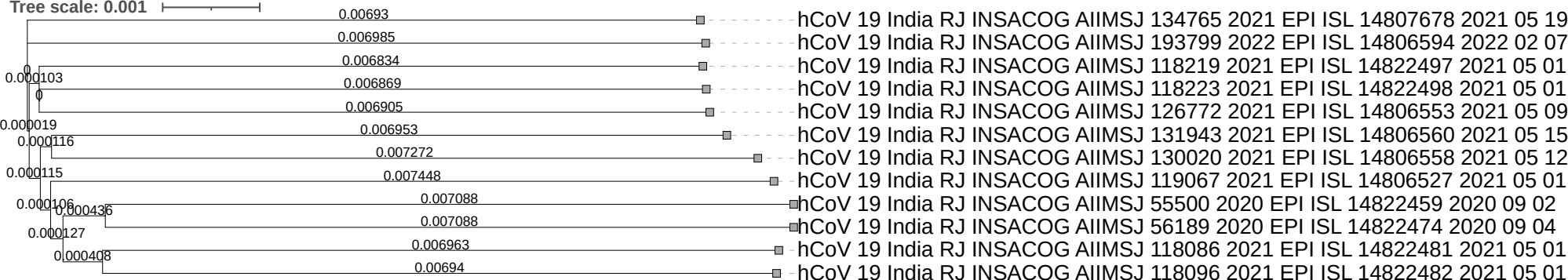
hCoV-19/India/RJ-INSACOG-AIIMSJ-134765/2021	EPI_ISL_14807678	GK	AY.122 (Pango v.4.1.2 PLEARN-v1.12), Delta (B.1.617.2-like) (Scorpio)	**Spike A67V**, Spike D614G, ***Spike D950N***, Spike E156G, Spike F157del, **Spike H69del**, ***Spike L452R***, Spike **N969K**, Spike P681R, Spike R158del, ***Spike T19R***, **Spike T95I**, **Spike T478K**, **Spike V70del**, M I82T, N D63G, N D377Y, N M210V, N R209A, NS3 D238Y, NS3 S26L, NS3 T34M, NS7b T40I, NSP2 K81N, NSP3 A488S, NSP3 P1228L, NSP3 P1469S, NSP3 V477I, NSP4 T492I, NSP4 V167L, NSP12 G671S, NSP12 P323L, NSP13 P77L	VOC Delta GK (B.1.617.2+AY.*) first detected in India
hCoV-19/India/RJ-INSACOG-AIIMSJ-131943/2021	EPI_ISL_14806560	GK	AY.122 (Pango v.4.1.2 PLEARN-v1.12), Delta (B.1.617.2-like) (Scorpio)	Spike D614G, Spike **D796Y**, Spike D950N, **Spike G142D**, Spike **N764K**, Spike **N969K**, Spike P681R, ***Spike T19R***, **Spike T478K**, M I82T, N D63G, N P13L, N R203M, NS3 D238Y, NS3 S26L, NS7b T40I, NSP2 K81N, NSP3 A488S, NSP3 K38R, NSP4 T492I, NSP4 V167L, NSP12 P323L, NSP13 P77L	VOC Delta GK (B.1.617.2+AY.*) first detected in India
hCoV-19/India/RJ-INSACOG-AIIMSJ-130020/2021	EPI_ISL_14806558	GRA	AY.122 (Pango v.4.1.2 PLEARN-v1.12)	**Spike A67V**, Spike D614G, **Spike D796Y**, Spike D950N, Spike H69del, Spike L981F, **Spike N679K**, Spike N764K, Spike **N969K**, **Spike P681H**, **Spike T95I**, **Spike T547K**, **Spike V70del**, M I82T, N G204R, N I131T, N P13L, N R203K, NS3 S26L, NS3 T14I, NS7b T40I, NSP2 K81N, NSP3 A488S,	VOC Delta GK (B.1.617.2+AY.*) first detected in India

				NSP4 T147I, NSP4 T492I, NSP4 V167L, NSP5 P132H, NSP12 P323L, NSP13 P77L	
hCoV-19/India/RJ-INSACOG-AIIMSJ-126772/2021	EPI_ISL_14806553	GK	AY.122 (Pango v.4.1.2 PLEARN-v1.12), Delta (B.1.617.2-like) (Scorpio)	**Spike A67V**, Spike D614G, ***Spike D950N***, Spike G142D, **Spike H69del**, ***Spike L452R***, Spike P681R, **Spike T95I**, **Spike T478K**, **Spike V70del**, **Spike V143del**, Spike Y144del, Spike Y145del, M I82T, N D63G, N D377Y, N G215C, N R203M, N S37P, NS3 D238Y, NS3 S26L, NS7a T120I, NS7a V82A, NS7b T40I, NSP2 K81N, NSP3 A488S, NSP3 P1228L, NSP3 S1682F, NSP4 T492I, NSP4 V167L, NSP6 T77A, NSP12 G671S, NSP12 P323L, NSP13 P77L, NSP14 A394V	VOC Delta GK (B.1.617.2+AY.*) first detected in India

hCoV-19/India/RJ-INSACOG-AIIMSJ-119067/2021	EPI_ISL_14806527	G	B.1.617.2 (Pango v.4.1.2 PLEARN-v1.12), Delta (B.1.617.2-like) (Scorpio)	**Spike A67V**, Spike D614G, **Spike H69del**, **Spike H655Y**, **Spike N679K**, **Spike P681H**, **Spike Q954H**, ***Spike T19R***, **Spike T95I**, **Spike V70del**, M I82T, N A254S, N D377Y, N G215C, N R203M, NS3 S26L, NS7b T40I, NSP3 A1883V, NSP3 A1892T, NSP3 K38R, NSP3 P1228L, NSP4 T492I, NSP6 T77A, NSP12 P323L, NSP13 P77L, NSP14 A394V	VOC Delta GK (B.1.617.2+AY.*) first detected in India
hCoV-19/India/RJ-INSACOG-AIIMSJ-119196/2021	EPI_ISL_14797957	GRA	BA.1 (Pango v.4.1.2 PLEARN-v1.12)	Spike D405N, Spike D614G, **Spike D796Y**, ***Spike D950N***, **Spike G339D**, Spike K417N, ***Spike L452R***, Spike L981F, **Spike N764K**, **Spike N969K**, ***Spike P681R***, Spike R408S, Spike S371F, **Spike S373P**, **Spike S375F**, Spike T376A, ***Spike T478K***, **Spike T547K**, **Spike V213G**, E T9I, M I82T, N D63G, N E31del, N G204R, N P13L, N R32del, N R203K, N S33del, NS3 S26L, NSP1 L92F, NSP3 K38R, NSP4 T492I, NSP5 P132H, NSP6 G107del, NSP6 L105del, NSP6 S106del, NSP6 T181I, NSP6 V149A, NSP12 P323L	

Table 4: Table presents all of the mutations those are present in mixed lineages and unassigned (probable) Omicron variant during third wave of COVID-19 in India. Defining mutations of Delta clade are presented in bold only, whereas defining mutations of Omicron clade are highlighted with bold and italics.

492


Tree scale: 0.001
0.00693
0.006985
0.006834
0.0000103
0.006869
0
0.006905
0.0000019
0.006953
0.0000116
0.007272
0.0000115
0.007448
0.0000106
0.007088
0.000436
0.007088
0.000127
0.006963
0.000408
0.00694
hCoV 19 India RJ INSACOG AIIMSJ 134765 2021 EPI ISL 14807678 2021 05 19
hCoV 19 India RJ INSACOG AIIMSJ 193799 2022 EPI ISL 14806594 2022 02 07
hCoV 19 India RJ INSACOG AIIMSJ 118219 2021 EPI ISL 14822497 2021 05 01
hCoV 19 India RJ INSACOG AIIMSJ 118223 2021 EPI ISL 14822498 2021 05 01
hCoV 19 India RJ INSACOG AIIMSJ 126772 2021 EPI ISL 14806553 2021 05 09
hCoV 19 India RJ INSACOG AIIMSJ 131943 2021 EPI ISL 14806560 2021 05 15
hCoV 19 India RJ INSACOG AIIMSJ 130020 2021 EPI ISL 14806558 2021 05 12
hCoV 19 India RJ INSACOG AIIMSJ 119067 2021 EPI ISL 14806527 2021 05 01
hCoV 19 India RJ INSACOG AIIMSJ 55500 2020 EPI ISL 14822459 2020 09 02
hCoV 19 India RJ INSACOG AIIMSJ 56189 2020 EPI ISL 14822474 2020 09 04
hCoV 19 India RJ INSACOG AIIMSJ 118086 2021 EPI ISL 14822481 2021 05 01
hCoV 19 India RJ INSACOG AIIMSJ 118096 2021 EPI ISL 14822482 2021 05 01