



COVID-19 genome surveillance: A geographical landscape and mutational mapping of SARS-CoV-2 variants in central India over two years

Krishna Khairnar^{a,b,*}, Siddharth Singh Tomar^{a,b}

^a Environmental Epidemiology & Pandemic Management (EE&PM), Council of Scientific and Industrial Research-National Environmental Engineering Research Institute (CSIR-NEERI), Nagpur, India

^b Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, UP, India

ARTICLE INFO

Keywords:

Pandemic
SARS-CoV-2
COVID-19
Genome surveillance
Next generation sequencing
Whole genome sequencing

ABSTRACT

Reading the viral genome through whole genome sequencing (WGS) enables the detection of changes in the viral genome. The rapid changes in the SARS-CoV-2 viral genome may cause immune escape leading to an increase in the pathogenicity or infectivity. Monitoring mutations through genomic surveillance helps understand the amino acid changes resulting from the mutation. These amino acid changes, especially in the spike glycoprotein, may have implications on the pathogenicity of the virus by rendering it immune-escape. The region of Vidarbha in Maharashtra represents 31.6 % of the state's total area. It holds 21.3 % of the total population. In total, 7457 SARS-CoV-2 positive samples belonging to 16 Indian States were included in the study, out of which 3002 samples passed the sequencing quality control criteria. The metadata of 7457 SARS-CoV-2 positive samples included in the study was sourced from the Integrated Health Information Platform (IHIP). The metadata of 3002 sequenced samples, including the FASTA sequence, was submitted to the Global Initiative on Sharing Avian Influenza Data (GISAID) and the Indian biological data centre (IBDC). This study identified 104 different SARS-CoV-2 pango-lineages classified into 19 clades. We have also analysed the mutation profiles of the variants found in the study, which showed eight mutations of interest, including L18F, K417N, K417T, L452R, S477N, N501Y, P681H, P681R, and mutation of concern E484K in the spike glycoprotein region. The study was from November 2020 to December 2022, making this study the most comprehensive genomic surveillance of SARS-CoV-2 conducted for the region.

1. Introduction

Viruses continue to mutate as they spread within the host population, develop into variants, and deviate genomically from the original virus. It is important to understand these changes in the viral genome and how they impact the aetiology of the virus. It is required to read the viral genome through whole genome sequencing (WGS) to detect changes in the viral genome. As the viral genome is comparatively smaller than the prokaryotic or eukaryotic genome, the frequency of mutations in the viral genome is comparatively high. The SARS-CoV-2 virus originated in December 2019 as a wild type, further diversifying into several variants of SARS-CoV-2. The rapid changes in the SARS-CoV-2 viral genome may cause immune escape leading to an increase in the pathogenicity or infectivity (Harvey et al., 2021a). Monitoring mutations through genomic surveillance allows the researchers to understand the amino

acid changes resulting from the mutation. These amino acid changes, especially in the spike glycoprotein, may have implications for the pathogenicity of the virus by making it escape the monoclonal antibody therapies and vaccine-induced immunity (Mengist et al., 2021). Hence, it is important to keep track of the changes in the viral genome to devise effective mitigation strategies. To track the emerging variants of SARS-CoV-2, healthcare authorities across the globe initiated genome surveillance programs. The government of India also initiated a nationwide SARS-CoV-2 genome surveillance network known as the Indian SARS-CoV-2 Consortium on Genomics (INSACOG), a forum established on 30 December 2020 under the Ministry of Health and Family Welfare to keep track of circulating SARS-CoV-2 variants in India (Genome, 2021).

The region of Vidarbha in Maharashtra represents 31.6 % of the state's total area. It holds 21.3 % of the total population. As per the 2011

* Corresponding author at: Environmental Epidemiology & Pandemic Management (EE&PM), Council of Scientific and Industrial Research-National Environmental Engineering Research Institute (CSIR-NEERI), Nagpur, India.

E-mail address: k_khairnar@neeri.res.in (K. Khairnar).

<https://doi.org/10.1016/j.virusres.2024.199365>

Received 28 March 2023; Received in revised form 21 March 2024; Accepted 22 March 2024

0168-1702/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

census data, the average population density of the region is 236 individuals per sq. km, with variations observed across different districts, ranging from 74 people per sq. km in Gadchiroli district to 470 people per sq. km in Nagpur district. It is also worth mentioning that the state of Maharashtra has emerged as a hotspot of SARS-CoV-2 infections since the pandemic's beginning. The origin of the virulent delta variant that drove the second wave of SARS-CoV-2 in India and around the globe was the district of Amravati in the Vidarbha region of Maharashtra (Kay and Pandya, 2021; Koshy, 2021). The largest city of the Vidarbha region is Nagpur, situated strategically at the centre of the north-south axis of India; Nagpur, an important urban centre of the region, also serves as a medical hub for the population living in surrounding rural and tribal areas. Nagpur is the third largest city in the Indian state of Maharashtra and the fourteenth largest city in India by population (Provisional Population Totals, 2011). Any outbreak in Nagpur city could quickly spread to other regions of India. Therefore, the findings of genome surveillance from the Vidarbha region hold great significance for understanding the dynamics of SARS-CoV-2 in India to control the spread of the SARS-CoV-2 virus in the population. This study was part of national SARS-CoV-2 genome surveillance efforts for the central region of India.

The Council of Scientific and Industrial Research-National Environmental Engineering Research Institute (CSIR-NEERI) Nagpur started the genomic surveillance of SARS-CoV-2 primarily for the Nagpur district of Maharashtra as gargle-based genome surveillance (Khairnar and Tomar, 2023). After becoming part of INSACOG, the genome surveillance mandate expanded to other districts of Maharashtra and other states of India like Tamil Nadu and the Union Territory of Jammu & Kashmir. In Total, 7457 SARS-CoV-2 positive samples belonging to 16 Indian States were included in the study, out of which 3002 samples passed the sequencing quality control criteria (Table 1). The duration of this study was from November 2020 to December 2022, which makes this study the most comprehensive genomic surveillance of SARS-CoV-2 conducted for the region.

2. Materials and method

2.1. Data collection

The metadata of 7457 SARS-CoV-2 positive samples included in the

Table 1
State-wise distribution of samples in the study.

S. No.	State/Union Territory	Samples Recieved	Samples Sequenced after applying selection criteria of <25 Ct value	% Sequenced
1	Tamil Nadu	1785	444	24.87
2	Maharashtra	5581	2516	45.08
3	Jammu & Kashmir	29	18	62.07
4	Kerala	17	8	47.06
5	Andhra Pradesh	3	1	33.33
6	Delhi	2	1	50.00
7	Haryana	1	1	100.00
8	Odisha	1	1	100.00
9	Uttar Pradesh	4	1	25.00
10	Bihar	1	0	0.00
11	Uttarakhand	1	0	0.00
12	Gujarat	10	0	0.00
13	West Bengal	4	0	0.00
14	Chattishgarh	2	2	100.00
15	Madhya Pradesh	8	6	75.00
16	Puduchery	8	3	37.50
Total		7457	3002	40.26

*Samples other than of Jammu & Kashmir, Tamil Nadu and Maharashtra are from visiting/interstate patients.

study was sourced from the Integrated Health Information Platform (IHIP) (IHIP, 2024). The hospitalisation and vaccination status information was obtained from the Indian Council of Medical Research (ICMR) COVID-19 data portal (ICMR, 2023). Within the vaccine information available group, 27.04 % of patients were vaccinated with Covaxin, 72.13 % were vaccinated with the Covishield vaccine, and 0.8 % with the Sputnik V vaccine. In the study population of 7457, the overall gender distribution was 52.9 % (3950) male, 47.02 % (3506) female and 0.01 % (1) transgender. In this study, the percentage distribution of the cases within the age groups was 0.36 % (0 to <2 years), 0.63 % (2 to <5 years), 5.8 % (5 to <15 years), 65.7 % (15 to <50 years), 18.5 % (50 to <65 years), and 8.97 % (\geq 65 years). The age distribution was according to the WHO Global Epidemiological Surveillance Standards for Influenza (WHO, 2013).

2.2. Materials

The SARS-CoV-2 viral RNA extraction was performed for gargle-based genome surveillance using an RNA release buffer containing Tris-EDTA and Proteinase K. The SARS-CoV-2 viral RNA extraction for OPS-NPS-based genome surveillance was performed using Insta NX automated nucleic acid extractor. Applied Biosystems QuantStudio 3 and StepOnePlus qRT-PCR machines were used for qRT-PCR, cDNA synthesis, PCR tiling, and rapid barcoding. Qubit 4 Fluorometer by Thermo Fisher Scientific was used to quantify nucleic acids for quality control (QC) of DNA libraries for Oxford Nanopore Technology (ONT)-based WGS. SARS-CoV-2 WGS was done on an Mk1C 6.3.9 and Mk1B ONT MinION sequencing platform.

2.3. Sample collection and RNA extraction

Out of 7457 SARS-CoV-2 positive samples included in the study, 707 were collected by saline gargle collection, and 6750 samples were collected using OPS-NPS VTM swabs from November 2020 to December 2022. The SARS-CoV-2 samples were collected using a saline gargle, and one-step RNA isolation from gargle samples was performed per the method developed by Khairnar et al. (Khairnar and Tomar, 2023). A video documentary of SOP on CSIR-NEERI's saline gargle-based SARS-CoV-2 molecular testing is available (Director CSIR-NEERI, 2021b). The SARS-CoV-2 samples collected using a nasopharyngeal-orpharyngeal (NP/OP) swab were suspended in a viral transport medium (VTM). The RNA from VTM swab samples was extracted using an automated RNA extractor using MBIN013 Insta NX Viral RNA Purification Kit according to the manufacturer's protocol (HIMEDIA) (Himedia, 2022). The isolated RNA samples were used immediately or stored at -80 °C till further use.

2.4. RT-PCR for detecting SARS-CoV-2

qRT-PCR was performed to quantitatively detect SARS-CoV-2 in samples using MBPCR255 Hi-PCR COVID-19 Triplex Probe PCR Kit (HIMEDIA) (Hi-PCR®, 2023) and/or Meril SARS-CoV-2 kit (RT PCR KIT, 2023) The primer-probe sets in the kits specifically detect viral RNA from SARS-CoV-2. The samples with a Ct value of \leq 38 for SARS-CoV-2 genes were considered positive for SARS-CoV-2 by RT-PCR. The sample's cycle threshold (Ct) value in RT-PCR exhibits inverse relation with the viral load. The study with OPS-NPS-based genome surveillance showed sequencing success, with samples exhibiting relatively lower Ct values. However, the saline gargle-based genome surveillance study showed sequencing success even with samples exhibiting relatively higher Ct values. Gargle-based genome surveillance enabled SARS-CoV-2 WGS of positive samples even with a lower viral load, exhibiting Ct values as high as >35 (Fig. 1).

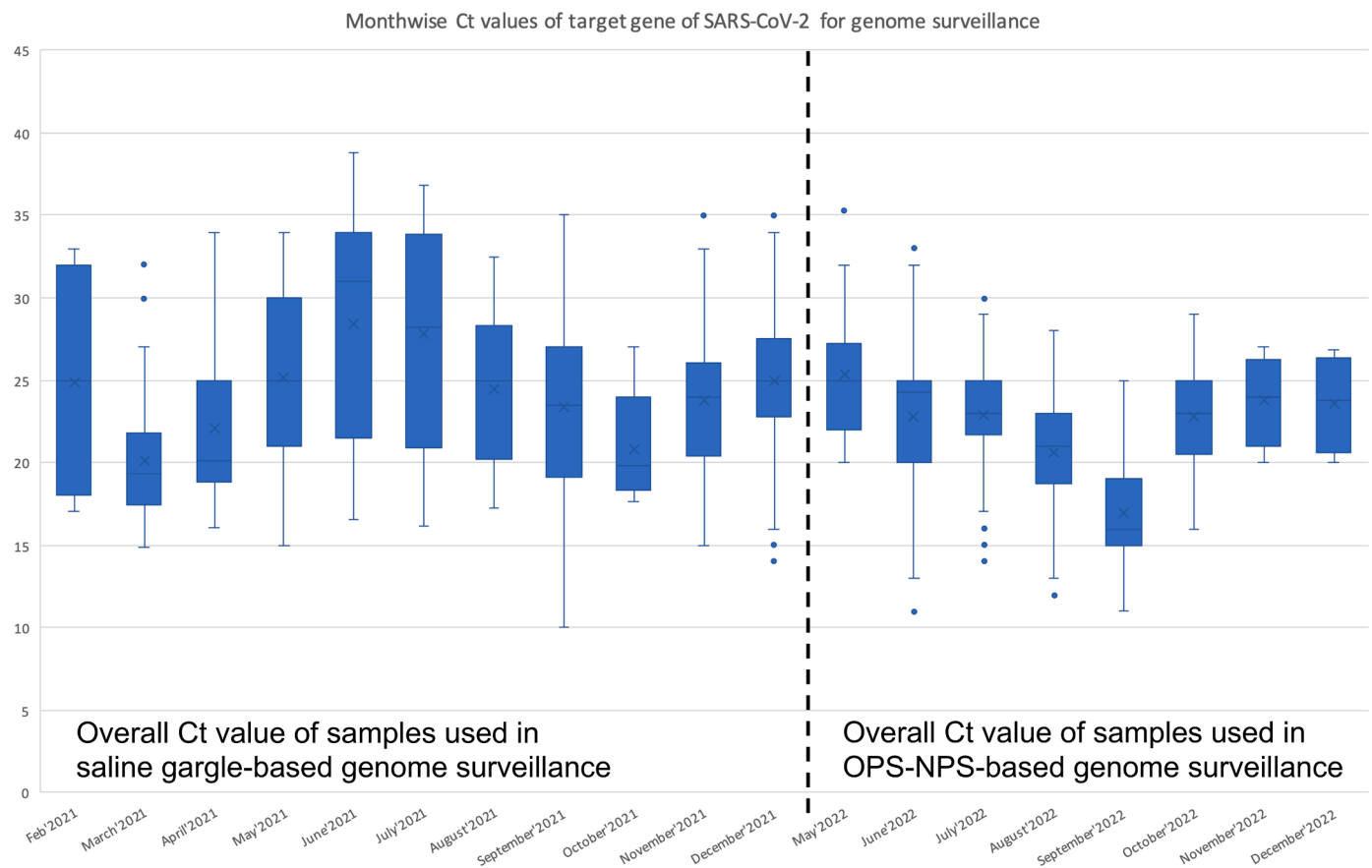


Fig. 1. Monthwise distribution of Ct value for SARS-CoV-2 samples used in saline gargle and OPS-NPS-based genome surveillance.

2.5. SARS-CoV-2 whole genome sequencing

Out of 7457 samples, only 3002 samples qualified for SARS-CoV-2 WGS according to the Ct value selection criteria of GISRS (WHO, 2022). The cDNA synthesis was done using the TAKARA Prime Script RT reagent kit (RR037A) (Takara bio inc, 2024). PCR tiling was done according to the midnight protocol of ONT (Oxford Nanopore Technologies, 2021). After pooling the samples, sequencing libraries were prepared by using ONT's Rapid Barcoding Kit-96 (SQK-RBK-110.96) The prepared cDNA libraries were cleaned by using SPRI magnetic beads and quantified using Qubit fluorometer followed by rapid adapter ligation and sequenced using MinION Mk1C or MinION Mk1B.

2.6. Bioinformatic analysis

2.6.1. Sequencing data processing

The live base-calling was performed using the Guppyv22.10.7 base-calling algorithm integrated into the MinION Mk1C sequencer (Wick et al., 2019). The processed FASTQ reads from the MinION sequencer were analysed using a graphical user interface bioinformatics platform COMMANDER developed by Genotypic Technology Pvt. Ltd (Commander, 2023). The FASTA sequence and the variant call performed by the COMMANDER pipeline were further confirmed by the web-based Pangolin COVID-19 Lineage Assigner (O'Toole et al., 2021). The sequencing quality control criteria of >70 % horizontal genome covered and >100X depth coverage were applied. The metadata of 3002 samples, including the FASTA sequence, was submitted to the GISAID (Gisaid.org, 2022) and the Indian biological data centre (IBDC) (INSA-COG - INDA-CA, 2023).

2.6.2. Mutation mapping

The mutation mapping of SARS-CoV-2 variants was done using the Outbreak.info web server (Outbreak.info, 2023). Outbreak.info server calculates the percent prevalence of mutations as a ratio of the number of sequences carrying a given set of mutations on a given day at a specific place (or all locations) (x) divided by the total number of sequences on that day in that location (n). A 95% Confidence interval was calculated using Jeffrey's Interval: 2.5 quantiles of $(x + 0.5, n - x + 0.5)$ to 97.5 quantiles of $(x + 0.5, n - x + 0.5)$. GISAID assigns PANGO lineage for each unique sequence. The Pango nomenclature system's classification is subject to changes as new lineages are uploaded to the GISAID database. While fundamental sequencing data remains the same in these cases, reporting for particular lineages may vary if analysed during a different period. The criteria of mutations with >75 % prevalence in at least one lineage were applied.

3. Results

The two-year-long study used Oxford nanopore technologies (ONT) based next-generation sequencing (NGS) platform. This study identified 104 different SARS-CoV-2 pango-lineages classified into 19 clades (Fig. 2). The detailed metadata of 3002 samples has been uploaded on the GISAID portal, and the accession numbers of the published sequences are provided in the supplementary file.

The study revealed a diverse variant profile of 104 pango-lineages, which included 12 major variants with >1 % prevalence among the total variants like B.1.1.529 (Omicron), (16.89 %), BA.2 (33.91 %), BA.2.75 (8.59 %), AY.112 (6.70 %), BA.2.38 (3.76 %), B.1.617.2 (2.76 %), BA.2.12 (2.76 %), B.1.617.1 (2.47 %), BA.2.10 (1.53 %), BA.5 (1.47 %), BA.2.76 (1.47 %), and AY.127 (1.07 %). As per the variant prevalence data retrieved from the open-sourced covlineages.org (Lineages.

3.2. Epidemiologically significant and emerging SARS-CoV-2 lineages

This study observed some epidemiologically significant variants from August 2022 to December 2022. The variants BM.1.1 (BA.2.75.3.1.1) and BM.1.1.3 (BA.2.75.3.1.3) belonging to the 22D clade were observed in September 2022; the BM.1.1.3 variant was majorly reported from India (20.1 %), USA (14.8 %), and United Kingdom (8.4 %) (Lineages.org. Cov, 2023). Another important group of variants belonged to the clade 22F: XBB, XBB.1, and XBB.2.

3.3. Phylogenetic analysis of major variants

Phylogenetic analysis of 12 major variants was performed to understand the phylogenetic relationships between the variants. FASTA sequences of the sequenced samples with >90 % horizontal coverage were selected. Multiple sequence alignment was performed using Molecular Evolution Genetic Analysis (MEGA) software version 11.0.13 (Tamura et al., 2021). The phylogenetic tree using aligned sequences was constructed by the Interactive Tree of Life (ITOL v5) web server (Letunic and Bork, 2021). A clear divergence in the clades of Delta and Omicron lineages was observed (Fig. 3)

3.4. Hospitalisation and death cases

The study has also evaluated the hospitalisation, reinfection, and death rates among Vaccinated and Unvaccinated patients. The vaccination data of 1521 patients were available, of which 1158 were unvaccinated, and 363 were vaccinated.

In the unvaccinated group, 106 patients were hospitalised, 1029 were home-isolated, 13 were reinfection cases, and ten died of SARS-CoV-2 infection. In the vaccinated patient group, 39 patients were hospitalised, 320 were home-isolated, 3 were reinfection cases, and one patient died due to the SARS-CoV-2 infection, as summarised in (Table 2a).

A chi-square test of independence was performed to examine the relationship between vaccination status and case outcomes. The chi-square statistic is calculated as 2.3005. The p-value is 0.5124. The result is not significant at $p < 0.05$ (Table 2b)

3.5. Variants prevailing in the vaccinated population

Among the vaccinated patients, the vaccine type information was available for 122 patients, out of which 88 were vaccinated with Oxford-AstraZeneca vaccine (ChAdOx1), also known as Covishield, manufactured under licence by Serum Institute of India (World Health Organization: WHO, 2022), and 33 were vaccinated with indigenously developed Covaxin (BBV152) (World Health Organization: WHO, 2022a). One patient was vaccinated with the Russian Sputnik V vaccine (Logunov et al., 2021). The major SARS-CoV-2 variant prevailing among vaccinated patients who received Covishield or Covaxin was BA.2, followed by other variants of Omicron (Fig. 4). The chi-square statistic is 1.1067. The p-value is 0.8932. The result is not significant at $p < 0.05$. means that no statistically significant association exists between the

Table 2a

The number of cases reported in the study, categorized by hospitalization status and vaccination status.

Status	Hospitalized	Not Hospitalized	Reinfection cases	Death cases
Unvaccinated	106	1029	13	10
Vaccinated	39	320	3	1
Total	145	1349	16	11

variants reported in vaccinated cases and the vaccine administered.

3.6. Mutation profile of the variants found in the study

An increase in infectivity and immune escape of the SARS-CoV-2 virus results from mutations in the SARS-CoV-2 genome. We have analysed the mutation profiles of major variants in our study using the Outbreak.info web server (Outbreak.info, 2023) (Fig. 5a). The analysis showed seven mutations of interest (MOI) which include K417N, K417T, L452R, S477N, N501Y, P681H, P681R and one mutation of concern (MOC) E484K in the spike glycoprotein region (Fig. 5a). The K417N mutation was found to be >60 % MOI prevalence among the GISAID sequences of B.1.1.529, BA.2, BA.2.10, BA.2.12, BA.2.38, BA.2.76, BA.2.75, and BA.5; while the Mutation K417T appeared with >60 % MOI prevalence in BA.2.38. L452R mutation was observed among the AY.112, B.1.617.2, and AY.127 variants of delta lineage with >90 % MOI prevalence. Interestingly, the L452R mutation was also noticed in the BA.5 variant of the Omicron lineage, recognised as a delta-like mutation (Islam et al., 2022). Mutations S477N, N501Y and P681H showed 65–90 % MOI prevalence in B.1.1.529, BA.2, BA.2.10, BA.2.12, BA.2.38, BA.2.76, BA.2.75, and BA.5. However mutation P681R appeared in the variants of delta lineages AY.112, B.1.617.2, and AY.127 with >90 % MOI prevalence. The E484K mutation showed <0.1 % MOC prevalence.

We have also analysed the mutation profiles of some epidemiologically significant variants found in the later phase of the study using the data from the Outbreak.info web server. The analysis showed eight MOI, including L18F, K417N, K417T, L452R, S477N, N501Y, P681H, P681R, and MOC E484K in the spike glycoprotein region (Fig. 5b). The L18F mutation showed <0.1 % MOI prevalence in BA.2.75.2, BA.5.1.1, BA.5.2, BA.5.2.1, XBB.1, and XBB.2. The mutation profile showed that all the variants isolated in the study retained the MOI K417N with 70–90 % MOI prevalence. The K417T mutation showed <0.1 % MOI prevalence in BA.2.75.21, BA.2.75.2, BA.2.75.6, BA.2.75.7, BL.1, BM.1.1, BF.3, BA.2.12.1, BA.4 and BQ.1.1. The L452R mutation showed 70–95 % MOI prevalence among the BR.1, BA.5.1.1, BA.5.1.3, BA.5.2, BA.5.2.1, BF.3, BA.4, and BQ.1.1. The mutation profile showed that all the variants isolated in the study retained the MOI S477N and N501Y with 80–90 % MOI prevalence.

4. Discussion

In this study, the variants of the Omicron lineage have dominated the variant distribution post-December 2021; the omicron variants reported

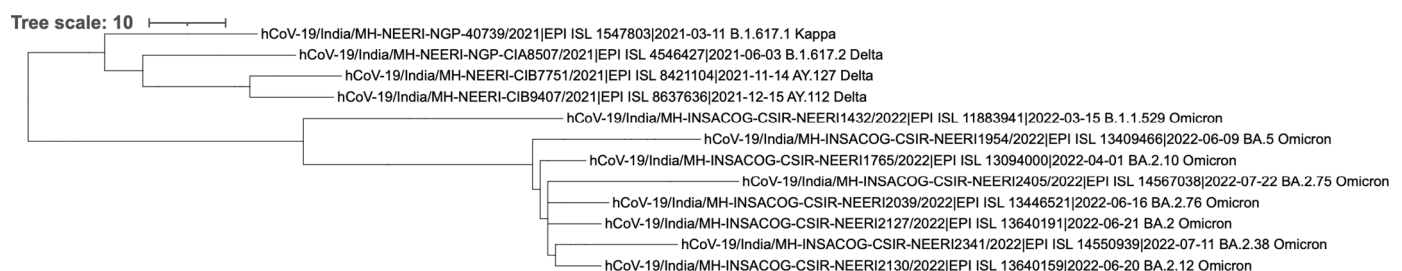


Fig. 3. Phylogenetic analysis of Major variants (>1 % Prevalence) found in the study.

Table 2b

Represents the observed frequency, followed by the (expected frequency) and the [chi-square value for each cell].

Status	Hospitalized	Not Hospitalized	Reinfection cases	Death cases	Row Totals
Unvaccinated	106 (110.39) [0.17]	1029 (1027.05) [0.00]	13 (12.18) [0.06]	10 (8.37) [0.32]	1158
Vaccinated	39 (34.61) [0.56]	320 (321.95) [0.01]	3 (3.82) [0.18]	1 (2.63) [1.01]	363
Column Totals	145	1349	16	11	1521 (Grand Total)

Variants Found in Vaccinated Cases

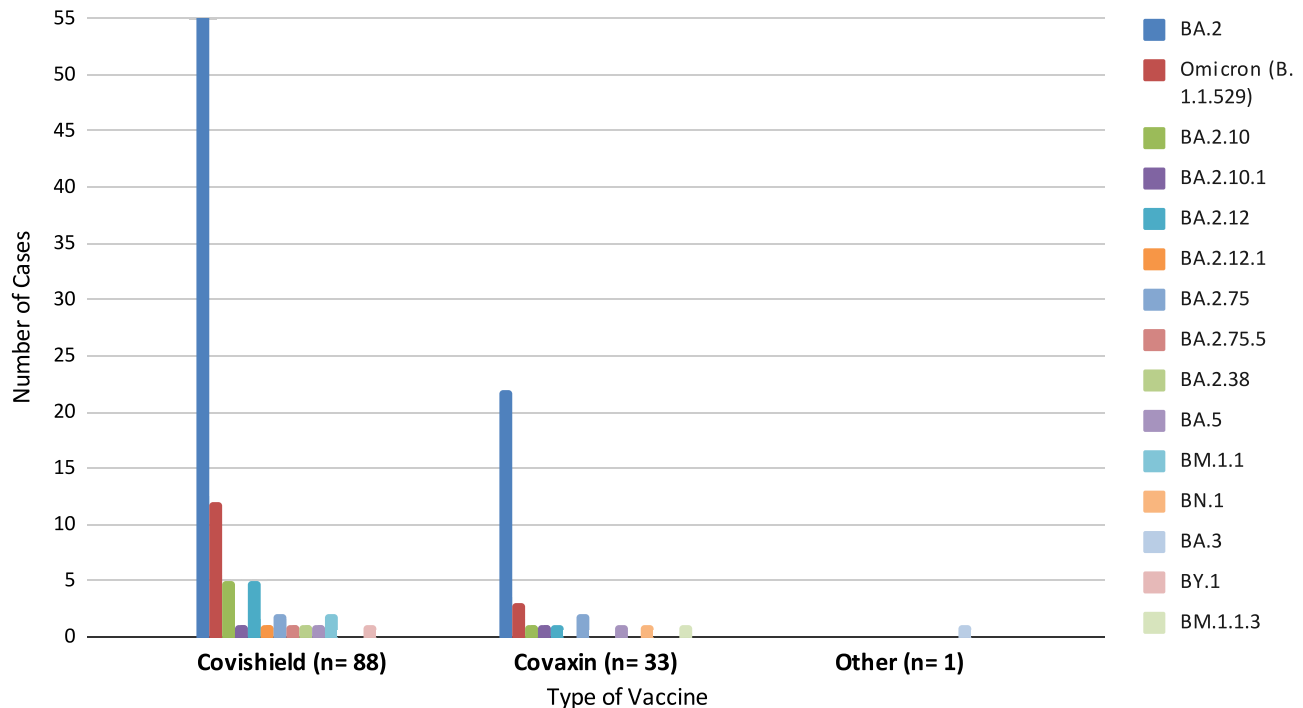


Fig. 4. Summary of SARS-CoV-2 variants prevailed among vaccinated cases who received Covishield, Covaxin and Sputnik V.

in this study belong to 10 different clades, as shown in Fig. 2. The Omicron variant was first detected in Botswana and South Africa on 24 November 2021 (Mallapaty, 2022) and was declared as a variant of concern (VOC) by the world health organisation (WHO) on 26 November 2021. Within one month of first reporting in Botswana, the B.1.1.529 (Omicron) variant was reported in our study from a sample collected on 21 December 2021. The sample was a 22-year-old female from Nagpur city in Maharashtra without international travel history. The prevalence of variants belonging to the Omicron lineage continued to proliferate and dominated the variant distribution afterwards. Other Omicron sub-variants, such as BA.1, BA.2, BA.3, BA.4, and BA.5, evolved from B.1.1.529 (Shrestha et al., 2022). Two sub-variants of BA.5 have evolved, BQ.1 and BQ.1.1. The Omicron variant is reported to be an immune escape variant as it has caused multiple cases of infection in vaccinated individuals (Cao et al., 2024). A study by Chakraborty et al. (2022) showed that the mutations D614G, E484A, N501Y, K417N, Y505H, and G496S improved S-glycoprotein accessibility to engage with the ACE2 receptor, which may contribute to enhancing the infectivity of Omicron variant.

The month-wise dynamics of SARS-CoV-2 over two years clearly showed the complete transition of SARS-CoV-2 from delta to omicron variant in January of 2022. The detailed month-wise distribution and the relative percentage share are elaborated in Fig. 6a and b. The trend of the SARS-CoV-2 wave pattern observed during the pandemic in Maharashtra showed consonance with the Indian SARS-CoV-2 wave pattern. The comparative analysis of SARS-CoV-2 cases of major variants in India and Maharashtra from November 2020 to December 2022 is

shown in Fig. 7. The Pearson correlation coefficient between the cases in Maharashtra and India for SARS-CoV-2 variants was analysed, which showed a weak positive linear relationship for BA.5 (0.402), medium positive linear relationship for B.1.1.529 Omicron (0.725); and strong positive linear relationship for BA.2 (0.958), BA.2.75 (0.948), AY.112 (0.982), BA.2.38 (0.914), B.1.617.2 (0.874), BA.2.12 (0.885), B.1.617.1 (0.930), BA.2.10 (0.923), BA.2.76 (0.896), and AY.127 (0.880).

The trend of the SARS-CoV-2 wave pattern observed during the pandemic in Maharashtra showed consonance with the Indian SARS-CoV-2 wave pattern.

The XBB and its related variants were responsible for increased SARS-CoV-2 cases in the USA, Singapore, and Europe (Callaway, 2023). XBB.1.5 has become the most common variant in the USA since December 2022. Since February 2023, the XBB.1.16 has been India's most dominating SARS-CoV-2 variant, as per the GISAID database (Gisaid.Org 2022). XBB variants show high immune evasion and increased hACE2 binding due to an F486S amino acid substitution in the spike glycoprotein (Salsekar et al., 2023).

Hospitalisation rate, reinfection, and death rate among Vaccinated and Unvaccinated patients implied no significant association between case outcomes and vaccination status for the sample studies. The observed association could plausibly occur due to random variation. The result emphasises the need for caution in making firm conclusions and leaves scope for further investigation.

Furthermore, the study results indicate the immune escape properties of the Omicron variant. WHO, in its Weekly epidemiological update on COVID-19 issued on 14 December 2021, indicated that after the 15th

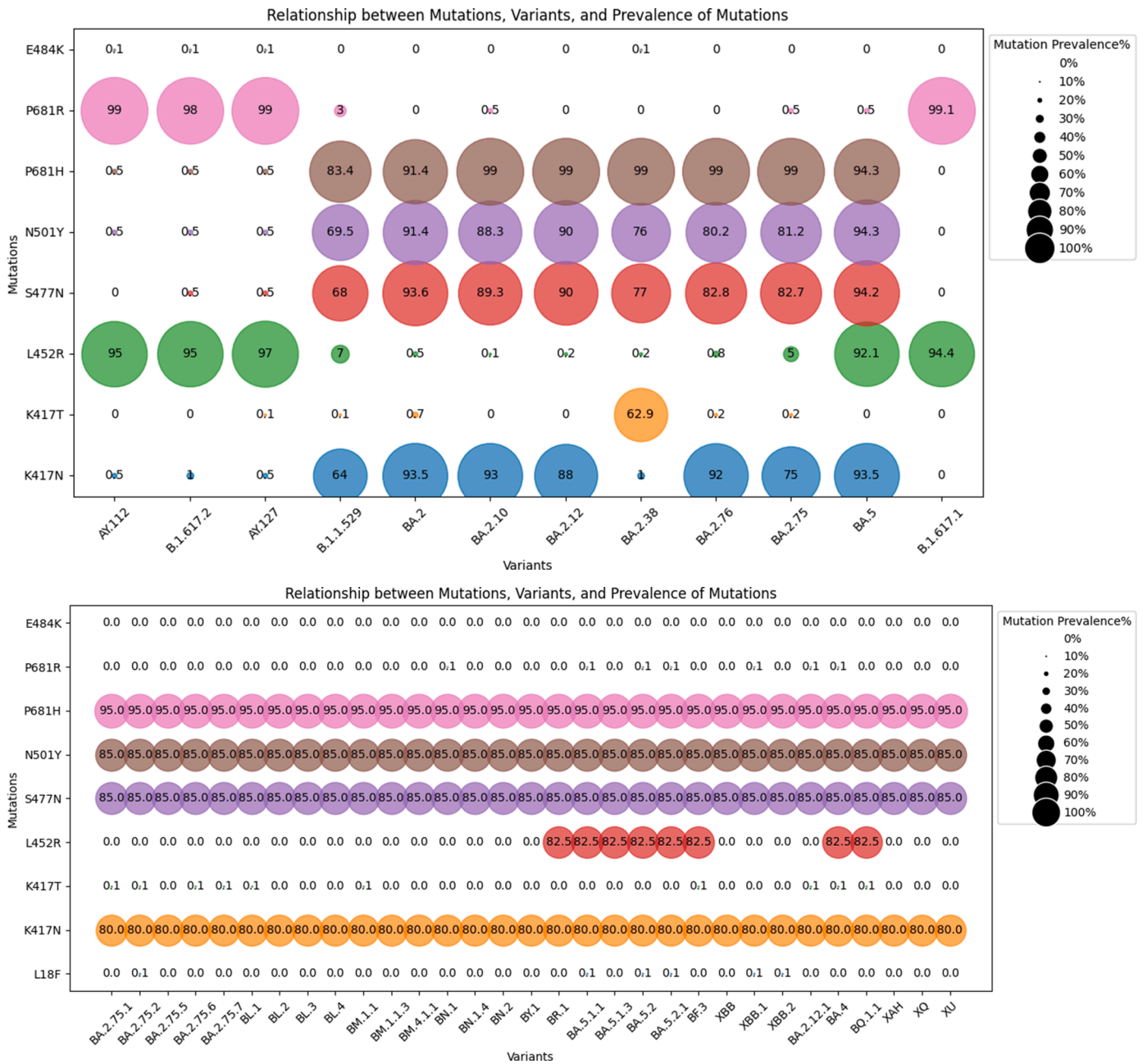


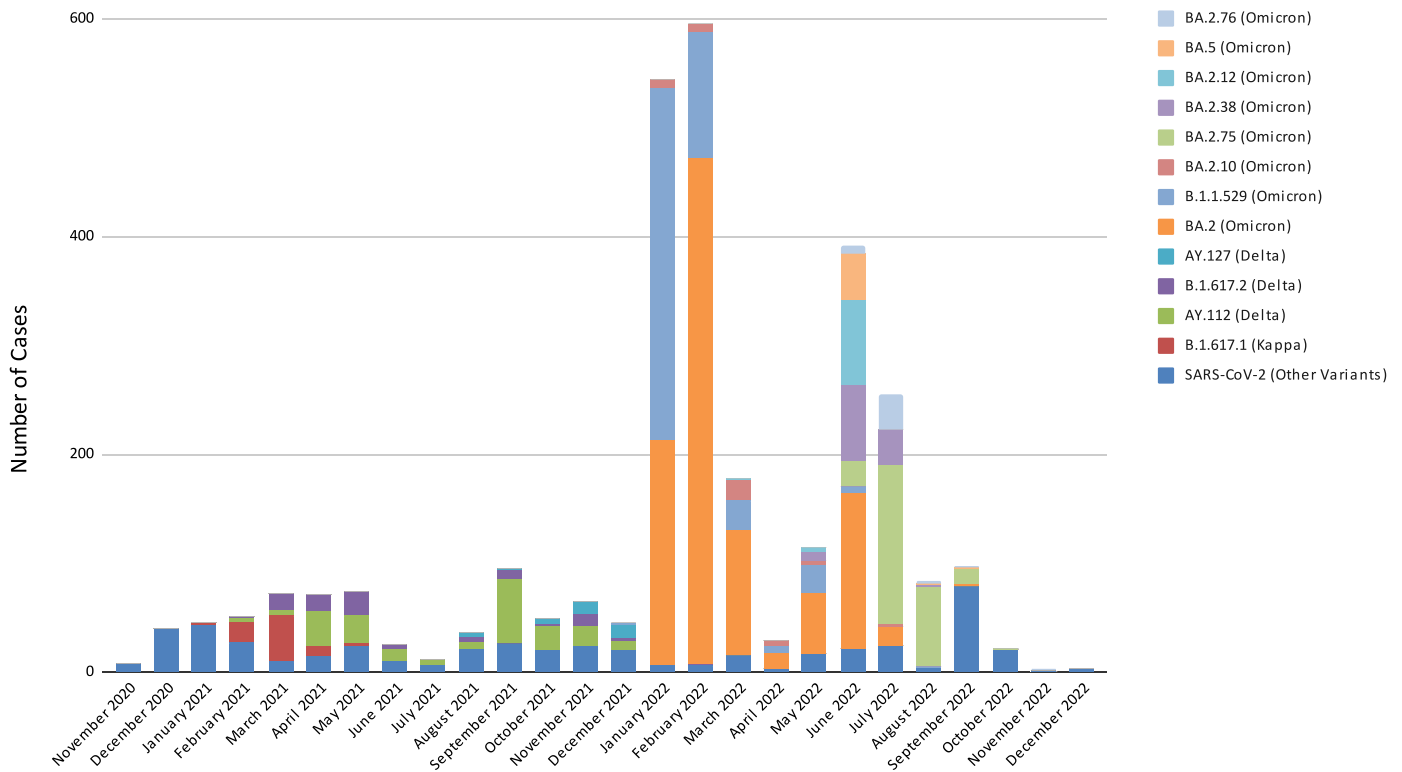
Fig. 5. (a): Relationship between Percentage prevalence of Mutation and Major variants (>1 % Prevalence) detected throughout the study. (b): Relationship between Percentage prevalence of Mutation and epidemiologically significant variants (<1 % Prevalence) detected in the later phase of the study.

week, the original two-dose regimen of the ChAdOx1 vaccine is ineffective against symptomatic illness caused by the Omicron variant. Andrews et al. (2022) also showed that the ChAdOx1 vaccine showed only 17.8 % effectiveness against the Omicron variant of SARS-CoV-2 after 15–19 weeks post-second dose (Andrews et al., 2022). These findings may explain the sudden rise in the Omicron cases in India during December 2021. The vaccine effectiveness may have waned off and rendered ineffective in protecting the population against the omicron variant after 15–19 weeks post-vaccination. However, this study has not observed any significant association between the type of vaccine administered and the variants found among vaccinated cases.

In this study, the mutation profile of the SARS-CoV-2 variants from Aug’2022 to Dec’2022 showed that all the variants isolated in the study retained the MOI P681H with 90–100 % MOI prevalence. The P681R mutation showed <0.1 % MOI prevalence in BN.1, BA.5.1.1, BA.5.2, BA.5.2.1, XBB.1, BA.2.12.1 and BA.4. The MOC E484K consistently

remain undetected among the variants found in the later part of the study. L18F substitution is one of the most prevalent mutations sequenced to date and is present in both the B.1.351 and P.1 lineages (Harvey et al., 2021b). L18F is associated with escaping N-terminal domain binding antibodies (McCallum et al., 2021). K417N mutation had reduced neutralisation by Casirivimab (REGN10933) and Etesevimab monoclonal antibodies (Tada et al., 2022). K417T mutations showed stronger binding with the hACE2 receptor contributing to increased pathogenicity of the variants possessing this mutation (Khan et al., 2021). By December 2020 and February 2021, L452R occurred independently in multiple lineages of delta worldwide, indicating that this amino acid substitution is most likely the consequence of viral adaptability due to increased immunity in the population (Cherian et al., 2021). The L452R mutation in RBD reduces the neutralising action of RBD-specific mAbs such as regdanvimab (CT-P59), etesevimab (LY-CoV016), and bamlanivimab (LY-CoV555) (McCallum et al., 2021).

DISTRIBUTION OF SARS-CoV-2 MAJOR VARIANTS FROM NOVEMBER 2020 TO DECEMBER 2022



PERCENTAGE SHARE OF SARS-CoV-2 MAJOR VARIANTS FROM NOV 2020-DEC 2022

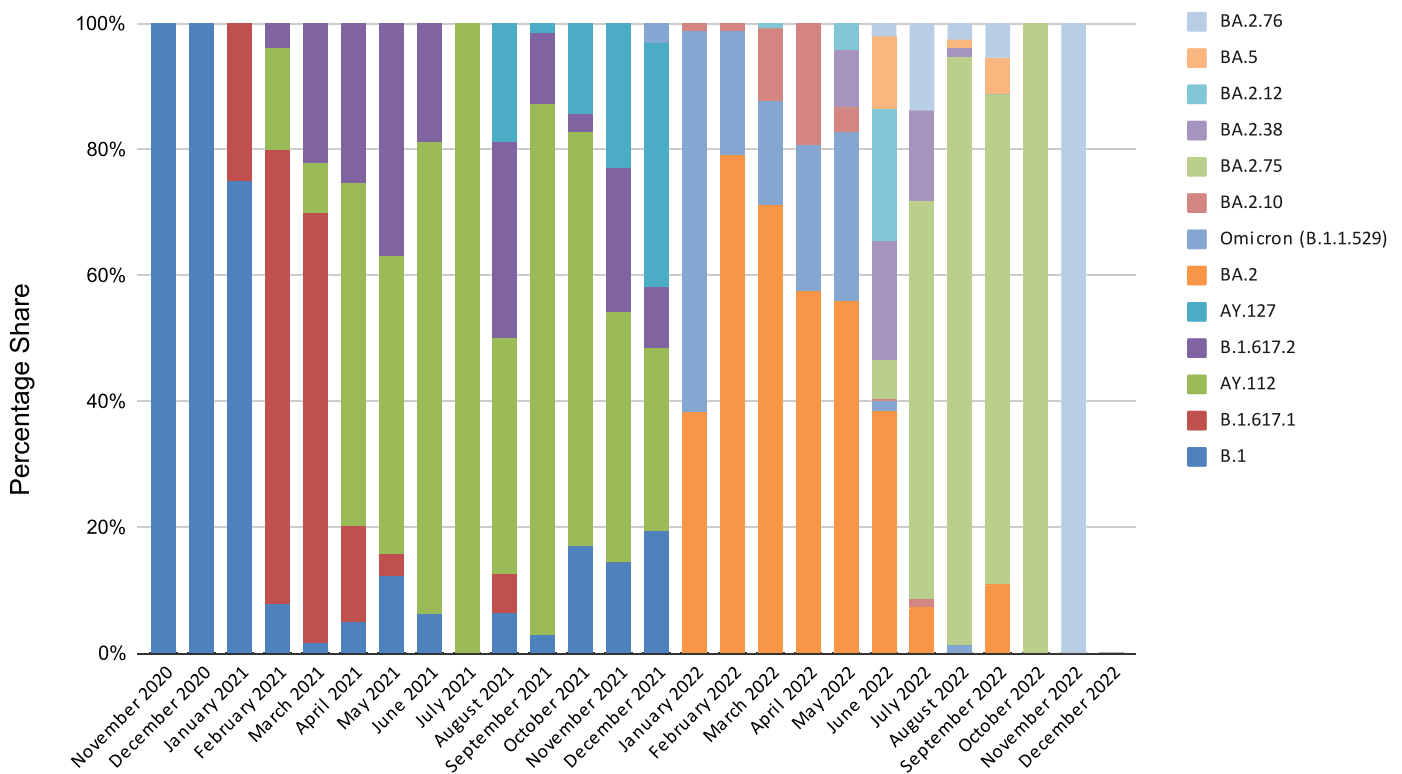


Fig. 6. (a): Monthly distribution of pango-lineages found in the SARS-CoV-2 genome surveillance. (b): Percentage-wise monthly distribution of major pango-lineages found in the SARS-CoV-2 genome surveillance.

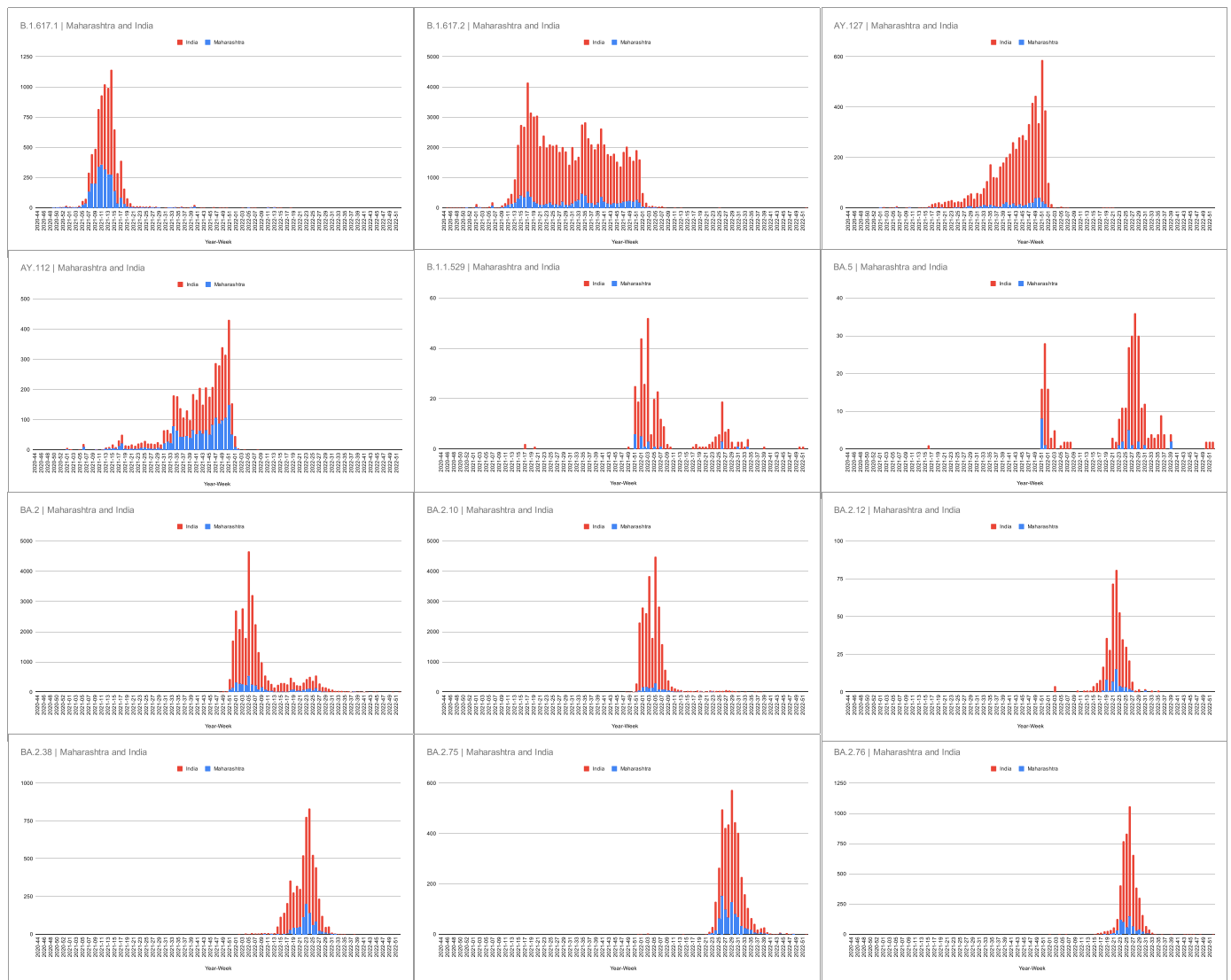


Fig. 7. Comparison of cases of major variants in India and Maharashtra from November 2020 to December 2022.

The S477N mutation is reported for a 37-fold increase in the RBD/ACE2 binding, thereby increasing the infectivity of variants having this mutation (Zahradník et al., 2021). *In vitro* studies showed that the mutation N501Y significantly increased ACE2 affinity imparted by a single RBD mutation (Starr et al., 2020). Studies using pseudoviruses with N501Y RBD mutation revealed that the neutralising activity of plasma from vaccinated individuals decreased significantly (Xie et al., 2024). The mutation P681H is reported to confer resistance against innate immune response (Lista et al., 2022). According to a study, the P681R mutation might be connected to the increased frequency of non-synonymous mutations in SARS-CoV-2 variants found in Maharashtra during February 2021 (Cherian et al., 2021). The mutation P681R in the furin cleavage site led to increased basicity of the poly-basic stretch that might improve the transmissibility of the SARS-CoV-2 virus (Lista et al., 2022). The mutation E484K was discovered to be one of the most prevalent mutations to evade monoclonal antibodies and natural class 2 polyclonal antibody response (Weisblum et al., 2020; Wang et al., 2021). E484K is among the most common RBD mutations reported *in-vitro* studies and sequences on the GISAD database.

Author declarations

I confirm all relevant ethical guidelines have been followed and any

necessary IRB and/or ethics committee approvals have been obtained. The Institutional Ethics Committee (IEC) of the Council of Scientific and Industrial Research-National Environmental Engineering Research Institute (CSIR-NEERI), Nagpur, India, approved the research. I confirm that all necessary patient/participant consent has been obtained and the appropriate institutional forms have been archived, and that any patient/participant/sample identifiers included were not known to anyone (e.g., hospital staff, patients or participants themselves) outside the research group so cannot be used to identify individuals.

Supplementary file 1: Line listing of GISAID submission with accession number for 3002 SARS-CoV-2 genome sequences from the study

Supplementary file 2: Weekly line listing of the Pearson correlation coefficient between the cases in Maharashtra and India for SARS-CoV-2 variants.

CRediT authorship contribution statement

Krishna Khairnar: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Siddharth Singh Tomar:** Data curation, Formal analysis, Software, Validation, Visualization, Writing – original draft.

Declaration of competing interest

The authors declare no conflicts of interest.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported and funded by CSIR-NEERI, Nagpur through project MLP-158 from March 2020 to March 2022 followed by project MLP-205 from March 2022 to April 2023. The authors would like to acknowledge the valuable contribution of all university students and support staff from Nagpur Municipal Corporation who were involved in this project.

It is certified that the manuscript has been checked for plagiarism by the institute knowledge resource centre through iThenticate (anti-plagiarism software) KRC No. CSIR-NEERI/KRC/2023/MARCH/EVC/2.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.virusres.2024.199365](https://doi.org/10.1016/j.virusres.2024.199365).

References

- Andrews, N., Stowe, J., Kirsebom, F.C.M., Toffa, S.E., Rickeard, T., Gallagher, E., Gower, C., Kall, M., Groves, N., O'Connell, A., Simons, D.S., Blomquist, P., Zaidi, A. A., Nash, S., Aziz, N.A., Thelwall, S., Dabrera, G., Myers, R.M., Amirthalingam, G., Bernal, J.L., 2022. Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant. *N. Engl. J. Med.* 386 (16), 1532–1546. <https://doi.org/10.1056/nejmoa2119451>.
- Callaway, E. (2023, January 9). Coronavirus variant XBB.1.5 rises in the United States - is it a global threat? *Nature News*. Retrieved March 22, 2023, from <https://www.nature.com/articles/d41586-023-00014-3>.
- Cao, Y.R., Wang, J., Jian, F., Xiao, T., et al., 2024. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* 602 (7898), 657–663. <https://doi.org/10.1038/s41586-021-04385-3>.
- Chakraborty, C., Bhattacharya, M., Sharma, A., Mallik, B., 2022. Omicron (B.1.1.529) - A newly mutated variant: mapped location and probable properties of its mutations with an emphasis on S-glycoprotein. *Int. J. Biol. Macromol.* 219, 980–997. <https://doi.org/10.1016/j.ijbiomac.2022.07.254>.
- Cherian, S., Potdar, V., Jadhav, S., et al., 2021. SARS-CoV-2 spike mutations, L452R, T478K, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *Microorganisms* 9 (7), 1542. <https://doi.org/10.3390/microorganisms9071542>.
- Commander, A command line free GUI-based sequencing data analysis tool developed by Genotypic Technology, Bangalore India <https://www.genotypic.co.in/commander/>. Accessed on March 22, 2023.
- Director CSIR-NEERI. Saline Gargle RT PCR Innovation by CSIR NEERI. YouTube. Published online August 9, 2021. <https://www.youtube.com/watch?v=10PZFIF4GpQ>.
- Genome Sequencing by INSACOG shows variants of concern and a Novel variant in India. (2021, March 21). pib.gov.in. Retrieved March 13, 2023, from <https://pib.gov.in/PressReleaseframePage.aspx?PRID=1707177>.
- Gisaid.org. 2022. <https://gisaid.org/submission-tracker-global/> (accessed March 22, 2023).
- Harvey, W.T., Carabelli, A.M., Jackson, B., et al., 2021b. SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* 19, 409–424. <https://doi.org/10.1038/s41579-021-00573-0>.
- Harvey, W.T., Carabelli, A.M., Jackson, B., et al., 2021a. SARS-CoV-2 variants, spike mutations, and immune escape. *Nat. Rev. Microbiol.* 19, 409–424. <https://doi.org/10.1038/s41579-021-00573-0>.
- Himedia. Insta NX® Viral RNA purification kit.; 2022. <https://www.himedialabs.com/in/mbin013-insta-nx-viral-rna-purification-kit.html>.
- Hi-PCR® COVID-19 triplex probe PCR kit. HiMedia leading BioSciences company. Retrieved March 22, 2023, from <https://www.himedialabs.com/in/hi-pcr-covid-19-triplex-probe-pcr-kit.html>.
- ICMR. ICMR public dashboard @ influenza.icmr.org.in https://influenza.icmr.org.in/public_dashboard/ Accessed on March 22, 2023.
- IHIP. Integrated health information platform integrated disease surveillance programme Ministry of Health and Family Welfare, 2024 Government of India @ ihip.nhp.gov.in. <https://ihip.nhp.gov.in/idsp/#1/>.
- INSACOG - INDA-CA. INDA. Retrieved March 22, 2023, from <https://inda.rcb.ac.in/insacog/indexpage>.
- Islam, M.R., Shahriar, M., Bhuiyan, M.A., 2022. The latest Omicron BA.4 and BA.5 lineages are frowning toward COVID-19 preventive measures: a threat to global public health. *Health Sci. Rep.* 5 (6), e884. <https://doi.org/10.1002/hsr2.884>. Oct 13.
- Kay, C., Pandya, D., 2021, December 29. How Delta variant spread in India: Deadly errors, inaction, Covid crisis. *Bloomberg.com*. <https://www.bloomberg.com/news/features/2021-12-29/how-delta-variant-spread-in-india-deadly-errors-inaction-covid-crisis>.
- Khairnar K, Tomar S. Non-invasive SARS-CoV-2 genome surveillance and its utility in resource-poor settings during the Delta wave of the COVID-19 pandemic. *medRxiv (Cold Spring Harbor Laboratory)*. Published online February 18, 2023. [10.1101/2023.02.16.23286031](https://doi.org/10.1101/2023.02.16.23286031).
- Khan, A., Zia, T., Suleman, M., Khan, T., Ali, S.S., Abbasi, A.A., Mohammad, A., Wei, D., 2021. Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: an insight from structural data. *J. Cell Physiol.* 236, 7045–7057. <https://doi.org/10.1002/jcp.30367>.
- Koshy, J., 2021, October, 3. The story of how a mystifying novel coronavirus variant, Delta, has India and the globe in its grip. *The Hindu*. <https://www.thehindu.com/sci-tech/health/the-story-of-how-a-mystifying-novel-coronavirus-variant-delta-has-india-and-the-globe-in-its-grip/article36772942.ece>.
- Leticun, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49 (W1) <https://doi.org/10.1093/nar/gkab301>. VolumeIssue2 JulyPages W293–W296.
- Lineages.org. Cov. Retrieved March 22, 2023, from <https://cov-lineages.org/>.
- Lista, M.J., Winstone, H., Wilson, H.D., Dyer, A., Pickering, S., Galao, R.P., De Lorenzo, G., Cowton, V.M., Furnon, W., Suarez, N., Orton, R., Palmirani, M., Patel, A. H., Snell, L., Nebbia, G., Swanson, C., Neil, S.J.D., 2022. The P681H mutation in the spike glycoprotein of the alpha variant of SARS-CoV-2 escapes IFITM restriction and is necessary for Type I interferon resistance. *J. Virol.* 96 (23), e0125022 <https://doi.org/10.1128/jvi.01250-22>.
- Logunov, D.Y., Dolzhikova, I.V., Shcheplyakov, D.V., Tukhvatulin, A.I., Zubkova, O.V., Dzharullaeva, A.S., Kovyrshina, A.V., Lubenets, N.L., Grousova, D.M., Erokhova, A. S., Botnikov, A.G., Izhaeva, F.M., Popova, O., Ozharovskaya, T.A., Esmagambetov, I. B., Favorskaya, I.A., Zrelkin, D.I., Voronina, D.V., Sherbini, D.N., Gintsburg, A.L., 2021. Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. *The Lancet* 397, 671–681. [https://doi.org/10.1016/s0140-6736\(21\)00234-8](https://doi.org/10.1016/s0140-6736(21)00234-8).
- Mallapaty, S., 2022. Where did Omicron come from? Three key theories. *Nature* 602 (7895), 26–28. <https://doi.org/10.1038/d41586-022-00215-2>.
- McCallum, M., De Marco, A., Lempp, F.A., Tortorici, M.A., Pinto, D., Walls, A.C., Beltramo, M., et al., 2021. N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* 184 (9), 2332–2347. <https://doi.org/10.1016/j.cell.2021.03.028> e16Accessed March 21, 2023.
- Mengist, H.M., Kombe Kombe, A.J., Mekonnen, D., Abebaw, A., Getachew, M., Jin, T., 2021. Mutations of SARS-CoV-2 spike protein: implications on immune evasion and vaccine-induced immunity. *Semin. Immunol.* 55, 101533 <https://doi.org/10.1016/j.smim.2021.101533>.
- O'Toole, A., Scher, E., Underwood, A., et al., 2021. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evolution* 7 (2). <https://doi.org/10.1093/ve/veab064>.
- Outbreak.info SARS-COV-2 data explorer. outbreak.info. Retrieved March 22, 2023, from <https://outbreak.info/>.
- Oxford Nanopore Technologies. PCR tiling of SARS-CoV-2 virus with rapid barcoding and midnight RT PCR expansion (SQK-RBK110.96 and EXPMRT001). Vol 3.; 2021. <https://nanoporetech.com/resource-centre/knowledge> Accessed on March 22, 2023.
- Provisional Population Totals, Census of India 2011 cities having population 1 Lakh and above provisional population totals, Census of India 2011 Cities Having Population 1 Lakh and Above.; 2011. <https://www.census2011.co.in/city.php>.
- RT PCR KIT: COVID-19 one-step rt-PCR kit. Meril Life. Retrieved March 22, 2023, from <https://www.merillife.com/assets/pdfs/medical-devices/covid-19-one-step-rt-pc-r-kit-1608096936.pdf.pdf>.
- Salsekar, L., Rahangdale, S., Tamboli, E., & Khairnar, K. (2023). Unique amino acid substitution in RBD region of SARS-CoV-2 omicron XAY.2. [10.1101/2023.01.09.523246](https://doi.org/10.1101/2023.01.09.523246).
- Shrestha, L.B., Foster, C., Rawlinson, W., Tedla, N., Bull, R.A., 2022. Evolution of the sars-cov-2 omicron variants BA.1 to Ba.5: implications for immune escape and transmission. *Rev. Med. Virol.* 32 (5) <https://doi.org/10.1002/rmv.2381>.
- Starr, T.N., et al., 2020. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* 182. <https://doi.org/10.1016/j.cell.2020.08.012>, 1295–1310.e1220.
- Takara bio inc. RR037A for research use PrimeScript™ RT reagent kit (Perfect Real Time) product manual. 2024 https://www.takarabio.com/documents/UserManual/RR037A_e_v2008Da.pdf.
- Tada, T., Zhou, H., Dcosta, B.M., et al., 2022. Increased resistance of SARS-CoV-2 Omicron variant to neutralization by vaccine-elicited and therapeutic antibodies. *EBioMedicine* 78, 103944. <https://doi.org/10.1016/j.ebiom.2022.103944>.
- Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027.
- Wang, Z., Schmidt, F., Weisblum, Y., et al., 2021. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* 592, 616–622. <https://doi.org/10.1038/s41586-021-03324-6>.
- Weisblum, Y., Schmidt, F., Zhang, F., DaSilva, J., Poston, D., et al., 2020. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife* 9, e61312. <https://doi.org/10.7554/eLife.61312>.

- WHO. Global epidemiological surveillance standards for Influenza; 2013. <https://www.who.int/publications/i/item/9789241506601> Accessed on March 22, 2023.
- WHO. End-to-end integration of SARS-CoV-2 and influenza sentinel surveillance. 2022 https://www.who.int/publications/i/item/WHO-2019-nCoV-Integrated_sentinel_surveillance-2022.1.
- Wick, R.R., Judd, L.M., Holt, K.E., 2019. Performance of neural network base-calling tools for Oxford Nanopore sequencing. *Genome Biol.* 1–10. <https://doi.org/10.1186/s13059-019-1727-y>.
- World Health Organization: WHO. (2022, June 10). The Bharat Biotech BBV152 COVAXIN vaccine against COVID-19: what you need to know. <https://www.who.int/news-room/feature-stories/detail/the-bharat-biotech-bbv152-covaxin-vaccine-against-covid-19-what-you-need-to-know>.
- World Health Organization: WHO. The Oxford/AstraZeneca (ChAdOx1-S [recombinant] vaccine) COVID-19 vaccine: what you need to know. Published June 13, 2022. <https://www.who.int/news-room/feature-stories/detail/the-oxford-astrazeneca-covid-19-vaccine-what-you-need-to-know>.
- Xie X, Zou J, Fontes-Garfias CR, et al. Neutralization of N501Y mutant SARS-CoV-2 by BNT162b2 vaccine-elicited sera. *bioRxiv (Cold Spring Harbor Laboratory)*. Published online January 7, 2021. [10.1101/2021.01.07.425740](https://doi.org/10.1101/2021.01.07.425740)Xie.
- Zahradník, J., Tuekprakhon, A., Ginn, H., Duyvesteyn, H., Bahar, M.W., Khan, S., Avinoam, O., Zhou, D., Nutalai, R., Supasa, P., Wang, B., Dejnirattisai, W., Liu, C., Djokaite, A., Temperton, N., Mongkolsapaya, J., Fry, E.E., Ren, J., Screaton, G., ... Stuart, D. (2021). Omicron-B.1.1.529 leads to widespread escape from neutralizing antibody responses. *bioRxiv (Cold Spring Harbor Laboratory)*. <https://doi.org/10.1101/2021.12.03.471045>.