Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Indian Journal of Medical Microbiology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Indian Journal of Medical Microbiology

journal homepage: www.journals.elsevier.com/indian-journal-of-medical-microbiology

Original Research Article

Genome characterization, phylogenomic assessment and spatio-temporal dynamics study of highly mutated BA variants from India

Poulomi Sarkar [a,b], Sarthak Banerjee [a,b], Saikat Chakrabarti [a,b], Partha Chakrabarti [a], Arun Bandyopadhyay [a,b], Arpita Ghosh Mitra [c], Soumen Saha [c], Aviral Roy [c], Siddik Sarkar [a,b,*]

[a] CSIR-Indian Institute of Chemical Biology (IICB), Kolkata, WB 700032, India
[b] IICB-Translational Research Unit of Excellence, Salt Lake, WB 700091, India
[c] MEDICA Super-specialty Hospital, Kolkata, India

ARTICLE INFO

Keywords:
BA.2/BA.2-like
Omicron
SARS-CoV-2
Whole genome sequencing (WGS)
Phylogenomic network

ABSTRACT

Purpose: The emergence of highly mutated and transmissible BA variants has caused an unprecedented surge in COVID-19 infections worldwide. Thorough analysis of its genome structure and phylogenomic evolutionary details will serve as scientific reference for future research.

Method: Here, we have analyzed the BA variants from India using whole-genome sequencing, spike protein mutation study, spatio-temporal surveillance, phylogenomic assessment and epitope mapping.

Results: The predominance of BA.2/BA.2-like was observed in India during COVID-19 third wave. Genome analysis and mutation study highlighted the existence of 2128 amino acid changes within BA as compared to NC_045512.2. Presence of 23 unknown mutation sites (spanning region 61–831) were observed among the Indian BA variants as compared to the global BA strains. Unassigned probable Omicron showed the highest number of mutations (370) followed by BA.1 (104), BA.2.3 (56), and BA.2 (27). Presence of mutations 'Q493R + Q498R + N501Y', and 'K417 N + E484A + N501Y' remained exclusive to BA.2 as well as unassigned probable Omicron. The time-tree and phylogenomic network assessed the evolutionary relationship of the BA variants. Existence of 424 segregating sites and 113 parsimony informative sites within BA genomes were observed through haplotype network analysis. Epitope mapping depicted the presence of unique antigenic sites within the receptor binding domain of the BA variants that could be exploited for robust vaccine development.

Conclusion: These findings provide important scientific insights about the nature, diversity, and evolution of Indian BA variants. The study further divulges in the avenues of therapeutic upgradation for better management and eventual eradication of COVID-19.

1. Introduction

Emergence of new SARS-CoV-2 'variants of concern' is challenging the efforts of global scientific and health community in eradicating the COVID-19 pandemic. Since these variants exhibit high adaptive features, the epidemic control of the disease becomes difficult [1]. During the last two years, several mutated strains of SARS-CoV-2 have emerged that exhibited increased immune responses within human body resulting in post COVID severe diseases [2]. The highly transmissible Delta variant brought about a sudden surge across the globe that peaked amongst all Variants of Concern (VoCs) [3,4]. The unprecedented surge of Delta (B.1.617.2) and Delta plus (AY.4/AY.x) in India could be linked to its unique mutations in spike protein amino acid sequences. Mutation D614G, present in all SARS CoV-2 variants, was reported to be one of the predominant single nucleotide polymorphisms of the SARS CoV-2 [5]. This mutation gives a certain replication advantage to the virus. Thus, D614G was associated to increased human transmission and infectivity events [6–8]. The spike protein mutation L452R which is found in B.1.617.1, B.1.617.2 and AY.4/AY.x variants is very important for the pathogenicity of the virus. It confers increased ACE2 binding affinity and decreased antibody binding capacity, thereby increasing host immune evasion by the virus [9–11]. Mutation P681R have been reported to facilitate S1-S2 cleavage at Furin cleavage site [9]. All these mutations facilitate viral replication, transmission and virulence within host cells. Some of the signature mutations of Delta variants are T19R, T478K and D950N which are present in N terminal domain, RBD and S2 region

* Corresponding author. CSIR-Indian Institute of Chemical Biology (IICB), Kolkata, WB-700032, India
E-mail address: siddik.sarkar@iicb.res.in (S. Sarkar).

https://doi.org/10.1016/j.ijmmb.2022.10.006
Received 9 June 2022; Received in revised form 8 September 2022; Accepted 19 October 2022
Available online xxxx
0255-0857/

Please cite this article as: Sarkar P et al., Genome characterization, phylogenomic assessment and spatio-temporal dynamics study of highly mutated BA variants from India, Indian Journal of Medical Microbiology, https://doi.org/10.1016/j.ijmmb.2022.10.006

respectively. Previous literature has highlighted that T19R and T478K (located in epitope binding region) impairs monoclonal antibody mediated neutralization while D950N affects spike protein dynamics thereby increasing virulence [12]. Following this, a highly transmissible BA/Omicron variant emerged in November 2021 in South Africa [13]. In January 2022 over 100,000 BA/omicron genomes were reported [14]. Moreover, this variant was associated with elevated immune escape. Several mutations in its genome conferred resistance to antibodies, thereby rendering the prevailing vaccines ineffective [13]. Previously, it was reported that the BA/Omicron had 37 mutations in its spike protein with 15 mutations alone in the Receptor Binding Domain (RBD) region [15]. Apart from the known mutations, this lineage has also undergone several mutations with time that have not been reported previously in any of the SARS-CoV-2 variants [16]. With increasing genome characterization for Omicron, presence of more than 50 mutations are found in the spike protein region of Omicron/BA variants. Greater than 19 mutations are solely present in the receptor binding domain of the spike protein. The most common mutations which are seen in the RBD region are G339D, R346K, K417N, S371L, S373P, S375F, N440K, Q498R, and N501Y [1]. The BA.1 subvariant of Omicron have deletions related to alpha at del 69-70 and delY144 which are absent in BA.2. Apart from shared mutations with other SARS-CoV-2 variants BA.1 and BA.2 have 30 and 25 unique mutations respectively in their spike protein amino acids sequence [18]. The delY144 mutation was associated with increased infection in several immunosuppressed COVID-19 patients indicating its involvement in immune escape [18]. Detailed study of various Omicron sub lineages also highlighted presence of R346K mutation in BA.1 that lead to its sub-classification to BA.1.1 (previously named as 'Mu' variant) [18].

In due course of time, various Omicron sub-lineages emerged across the country and they were found to have significant difference in the pattern of mutations within the spike protein. Earlier this year few studies claimed evolution of recombinant VoCs (denoted as 'Deltacron', 'Demicron'). According to World Health Organization such recombinant VoCs were considered to be result of laboratory contamination [17,18]. However, recent studies have provided evidence regarding presence of Deltacron having both the genetic mutations of Delta and Omicron [19–21]. Three clusters of Deltamicron or Deltacron was cultured and identified in Southern France [22]. Genetic recombination is a common occurrence in many human viruses. These recombination or mutations help in viral evolution through functional gene transfer which over time accumulates to give selective advantage to the virus in terms of altered pathogenicity, host receptor binding, zoonotic and anthropogenic transfers [18]. Recently multiple cases have been identified in different parts of the world where the patients were reported to be harboring co-infections with the Delta and Omicron variants [23,24]. The incidence of such cases in immunocompetent and epidemiologically unrelated patients, strengthens the understanding that whenever more than one highly transmissible variant is in circulation in the same region during a pandemic wave, there could be possibilities of co-infection [23]. However, it is not easy to find out if the infection by the two different lineages occurred simultaneously or sequentially [24]. However, co-infection by Delta and Omicron in an individual can be detected using PCR test for single nucleotide polymorphism (SNP) at the receptor binding domain of spike protein K417N and L452R and further confirmed by whole genome sequencing [18]. Literature report suggested that careful analysis of SNPs and allele discrimination plots can confirm presence of co-infections of Delta/Omicron in an individual [23]. Evolution and detection of any recombinant strain of Delta/Omicron was done using single molecule real time (SMRT) sequencing which showed presence of two break points (at nucleotide 22,036 and 22,193) in the 5′-region of the spike protein indicating it as the hotspot of recombination [23].

Despite several vaccine and diagnostic interventions the eradication of SARS-CoV-2 is still a challenge. Detailed analysis of novel SARS-CoV-2 variants become a critical tool for containing COVID-19 future surges. In the early phase of BA infections in India presence of BA.1, BA.1.1, and BA.2 was seen with BA.2 to be the most prevalent. This variant triggered the 'Omicron' wave in India during December 2021 to February 2022. However, a detailed analysis of the BA variant from India is still missing. Thus, the present study delves with genome characterization, mutation analysis, spatio-temporal dynamics study, epitope mapping and phylogenetic network assessment of BA variants from India during 'Omicron wave'.

2. Materials and methods

2.1. Data mining and plotting

All genome sequences and associated patient metadata of different BA/Omicron variants which were available from November 2021 to February 2022 from all over India were downloaded from the GISAID EpiCoV database (https://www.gisaid.org/) (Supplementary material 3). R/RStudio console (R-4.2.0) with different packages was used for creating the plots. The map of India was obtained from the URL: https://gadm.org/download_country.html selecting India as the country.

2.2. Genome sequencing and bioinformatics analysis

Apart from in silico analysis of different BA genomes from all over India, the genome characterization of BA variants from West Bengal was done using Whole Genome Sequencing (WGS). Sixty seven RNA samples were collected from West Bengal and subjected to whole-genome sequencing using the Illumina Miseq platform (Illumina Inc.). Library preparation was done using Illumina COVIDSeq RUO Kit. Miseq v3 flow cell was used for cluster generation and sequencing. Bioinformatic analysis was enabled by RSubread (release 3.15) and Illumina Basespace COVID Dragen Lineage. The Pangolin (version-4.0.6) and GISAID were used for the taxonomic identification. The mutation analysis was performed using the GISAID CoV server mutation app.

2.3. Time tree and phylogenomic network construction

BA sequences with collected during November 2021 to February 2022 were downloaded from GISAID EpiCoV database. High coverage sequences (n = 408, downloaded from GISAID along with in house whole genome sequences) were further screened and applied to MAFFT v7 web server (https://mafft.cbrc.jp/alignment/server/) for multiple alignment using NC_045512.2 as the reference genome. The aligned sequences were used for phylogenetic network construction using Population Analysis with Reticulate Trees (PopART v1.7) software [25]. The time tree was constructed using the MEGAX software with the Maximum Likelihood method [26].

2.4. Linear B cell epitope analysis

The linear B cell epitope analysis was done using the BepiPred 2.0 server [27]. The threshold value was set at 0.5 to achieve the most accurate prediction. The antigenicity of the predicted epitopes was calculated using Vaxijen 2.0 server [28] using 0.5 as the threshold antigenicity value.

2.5. Accession numbers

The annotated sequences were submitted to the GISAID EpiCov database with accession numbers: EPI_ISL_9781619 to EPI_ISL_9781691, EPI_ISL_8806019, and EPI_ISL_8806007.

2.6. Ethics approval

Ethical clearances were taken from the CSIR-Indian Institute of Chemical Biology (IICB/IRB/2020/6) and MEDICA Super-Specialty Hospital CREC/2020/JUL/1(a)) human ethics committees. The work

was executed under expert supervision following appropriate COVID-19 protocols.

3. Results and discussion

3.1. Spatio-temporal dynamics of Omicron variants in India

The first incidence of BA/Omicron variants in India was reported in late November 2021 as a part of Sentinel surveillance. However, the data was submitted to GISAID and made publicly available in February 2022. Hence, the official report of BA cases came to light in December 2021 from Karnataka from a male who had an international travel history.

Since our study is based on details of BA lineage up to February 2022 we have used the previous annotations for all BA variants. BA.2 was the most dominant type to circulate among Indian masses, followed by BA.1 and BA.1.1 (Fig. 1, Supplementary material 1). Maximum BA events were reported from December 2021 to January 2022 coinciding with the peak of the third wave of SARS-CoV-2 in India (Supplementary material 1). Transmission dynamics of BA.2 indicated that it outnumbered the other two variants during December 2021 (Supplementary material 1). Due to lag in sequence submission to GISAID from India undulations were observed in the graph. The initial infection rate for BA.2 was calculated to be 37.5 infections per day. Detailed analysis of the metadata revealed that in November 2021 first few cases of BA.1 were reported from Karnataka [3], Maharashtra [3], Delhi [2], Gujarat [2], and Punjab [1]. In the following month, cases rose and spread to 19 other Indian states with Maharashtra reporting the highest number (n = 688). In January 2022 BA.1 cases showed a slight dip with Maharashtra reporting the highest (452) cases. However, in February 2022 the overall BA.1 cases decreased considerably and 14 cases were reported from Telangana. The incidences of BA.1.1 were also seen from November 2021 with the first report from Gujarat. In December 2021 the transmission trends of BA.1.1 showed a similar pattern as BA.1, and Maharashtra was the worst-hit state with a maximum of 347 infections registered. Cases spiked in Maharashtra in January 2022 with 376 infections. But, in February Maharashtra showed a decline in overall caseload and Telangana reported the highest number of sequences. The transmission pattern of BA.2 followed a slightly different trend. The first two sequences were reported from Maharashtra in November 2021 which reached 130 in December 2021. During this time West Bengal reported maximum cases (1177 sequences reported). In January the infection rate of BA variants declined slowly and Maharashtra again showed the highest infection with 875 sequences. Infection rates in West Bengal declined considerably during January 2022. In February 2022 Telangana again showed the highest transmission with 278 cases. The variance in BA cases in different states/UTs could be attributed to the amount of genome sequencing done and made available in a public database(s). The more the number of cases reported, the higher the count seen. The Spatio-temporal dynamics of BA in India suggested that during the peak, Ladakh, Himachal Pradesh, Uttarakhand, Jharkhand, and Chhattisgarh showed considerably low cases possibly due to under-reporting of sequences. Due to the higher daily public mobility in Delhi, Karnataka, Maharashtra, West Bengal, and some other southern states of India, more community transmission happened and hence, higher cases were seen.

3.2. Genome analysis and mutation signatures

Whole-genome sequencing of 67 viral RNA samples (obtained from West Bengal) was done and on average > 500,000 sequences were retrieved. Individual samples with >97% identity were mapped against the reference genome (NC_045512.2) and a consensus sequence of size >29,000 bp was derived. Taxonomic classification of the samples in the GISAID EpiCoV database showed that they belonged to the GRA clade (BA/Omicron). As per Pango-lineage (updated in February 2022), 60 samples belonged to BA.2, two to BA.1 lineage, and one to BA.2.3. However, four samples could not be designated to any known BA lineage.

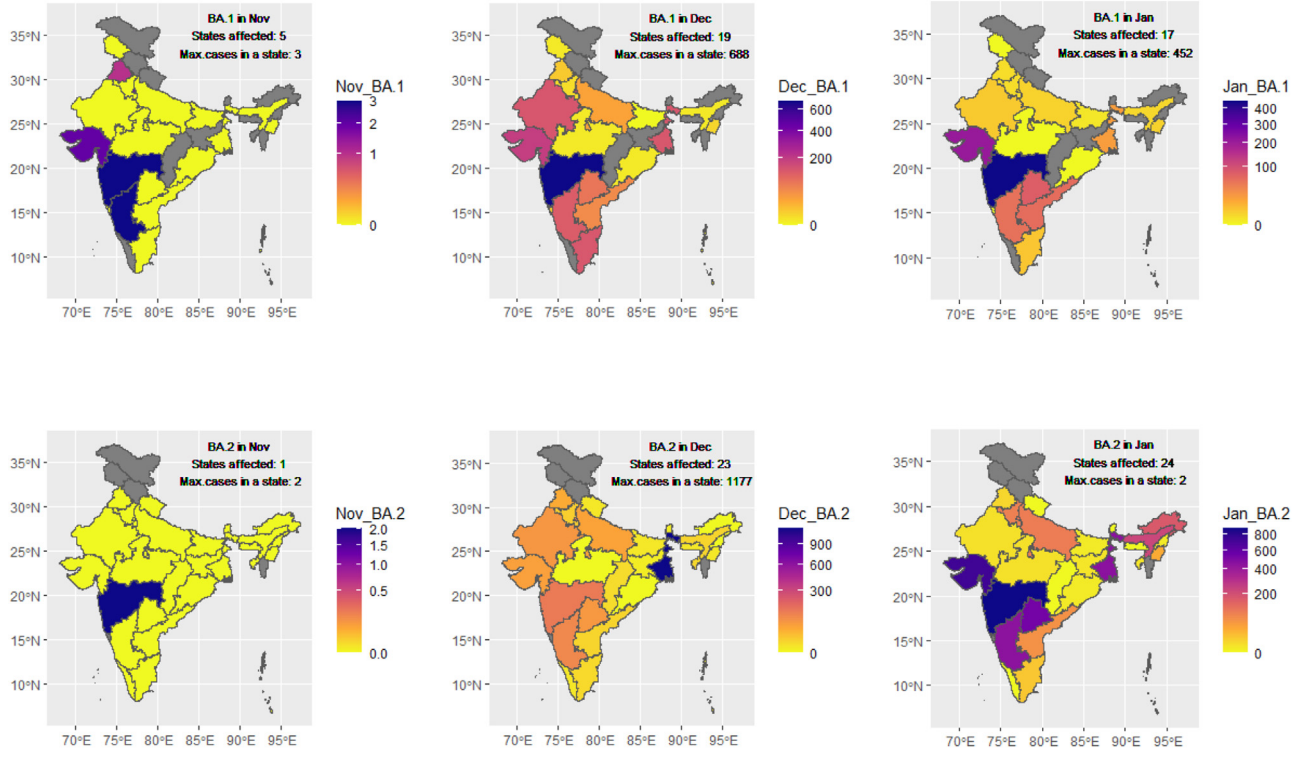

Fig. 1. Spatio-temporal incidences of different variants of BA/Omicron in Indian states and union territories (UTs). The dynamic incidences of Delta-surge in Indian states and UTs were shown in indicated months. Note: The geographical/administrative boundaries are used only for the purpose of data representation. The spatial polygonal coordinates or boundaries of state might differ.

Hence, they got classified as 'unassigned probable Omicron'.

Spike protein amino acid mutation analysis was performed for all the 67 samples against the reference strain NC_045512.2. The presence of 2128 mutation events and four patterns were distinct (Supplementary material 1). They were mostly concentrated within the receptor binding domain (RBD) (19–900 amino acids) of the spike protein (Fig. 2). The highest number of total amino acid changes were observed for unassigned probable Omicron (1605), followed by BA.1 (237), BA.2.3 (218), and BA.2 (68). The number of mutations within the spike protein of unassigned probable Omicron, BA.1, BA.2.3, and BA.2 were 370, 104, 56, and 27 respectively. Mutations at 'H655Y + N679K + P681H' were common to all BA variants. The overlapping mutations of BA with other SARS CoV-2 variants included H69de, T478K, N501Y, and D614G. These mutations play role in robust hACE2 and RBD binding as well as cell-to-cell fusion which induces the production of more nucleated cells eventually leading to an increase in transmissibility or infectivity of these variants [29]. Infectivity of the virus can be defined as the process of horizontal transmission of viral particles within different hosts. The spike protein mutations which alter the hACE2 and RBD binding affinity determines the horizontal host transmission or infectivity of the virus. Apart from this the spike protein mutations which cause acceleration in Furin cleavage which in turn increases host cell attachment of the virus also elevate the infectivity [30]. Triple mutation 'Q493R + Q498R + N501Y', and 'K417N + E484A + N501Y' remained exclusive to BA.2 as well as unassigned probable Omicron. Simulation studies and experimental findings have shown that these triple mutations enhance the binding affinity of hACE2 with RBD [29,31]. Signature mutations of BA.2 were seen at T19I and N440K. Similarly, mutations at D405N, S477N, and Y505H remained inclusive between BA.2 and BA.2.3. Literature-based data revealed that 'H655Y + N679K + P681H' mutations accelerate Furin cleavage and fusion of the virus with the host cell membrane, thereby, elevating the infectivity rate [32,33]. Triple mutation 'K417N + E484A + N501Y' in Omicron helps in immune evasion by neutralizing antibodies [34]. Mutations at D405N, K417N, N440K, S477N, T478K, E484A, Q493R, Q498R, N501Y, and Y505H at the RBD site also participate in escaping immunity. Moreover mutations such as N440K, T478K, E484A, Q493R, and Q498R which were present in Omicron/BA enhances the stability of the hACE2-RBD complex. Thus, all these mutations increases the infectivity of Omicron. The binding free energies (entropy and enthalpy) determines the stability of binding of spike protein RBD with the human ACE2 receptor. More the binding more stable the complex will be and more will be the viral transmissibility/infectivity [35]. The patient metadata corroborated well with this immune evasion, as the trend showed the susceptibility of vaccinated people to BA/Omicron. The mutations were not only present within the spike protein but many substitutions were also present in the ORF1ab, envelope protein (E), nucleocapsid protein (N) that participates in viral replication and transcription [29].

Apart from this, the presence of 23 unknown mutation sites were found. Unassigned probable Omicron showed the presence of deletions/substitutions at N61X, T63X, N74X, L242X, A243X, L244X, N709X, and A831X. They might be associated with the removal of a potential *N*-glycosylation site, thereby, affecting antigenicity as well as host specificity of the variant [36,37]. Similarly, in the case of BA.1 unique changes at N165X, T167X, N282X, T284X, N450X, L452X, and Y453X, R454X, and H519X were observed. These mutations also help in antigenic drift [36]. BA.1 also harboured two mutations at N317X and L517X. These changes affect the virulence and host range of the virus [38]. The BA.2.3 variant contained mutations at A435X, D428X, T430X, and I434X. They could be linked to host change and antibody recognition [39]. The Omicron variants have gone considerable number of mutations in their spike protein region. Mutation at T95I which was present in the genome of the Omicron variants participate in the antigenic drift, thereby affecting the antibody binding [40]. Again our analysis showed presence of K417N, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y. and Y505H mutations in the antigenic site 1 of the virus. Presence of these mutations in the antigenic site of the viral spike protein is associated with weakening the effect of several antibodies [41,42]. Studies by several researchers have depicted the inactivation of various antibodies by BA/Omicron viral particles [15,40,44].

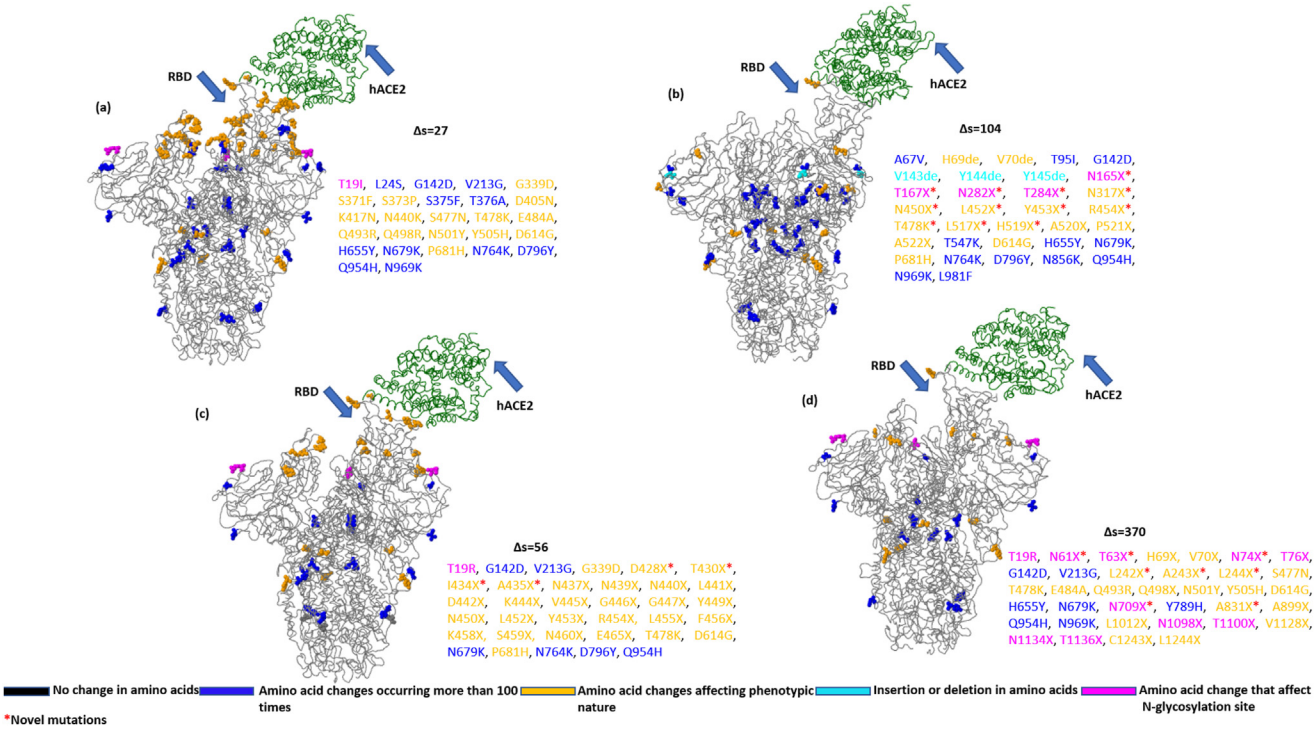

Fig. 2. Mutation profile of the spike protein of different BA variants in India. The variants depicted are (a) BA.2, (b) BA.1, (c) BA.2.3, and (d) Unassigned probable Omicron. Δs denotes the number of amino acid changes within the spike protein.

3.3. Phylogenomic network and temporal evolution of BA variants

To infer the evolutionary relationship of the different BA variants a parsimony based Transitive Consistency Score (TCS) network was constructed (Fig. 3). The alignment of 409 whole genome sequences of BA and their further phylogenomic analysis showed presence of 424 segregating sites. Among these sites 113 parsimony informative sites were seen. The negative Tajima's D statistics value ($D = -2.6548$, $p\ (D \geq -2.6548) = 0.999922$) indicated the importance of these parsimony informative sites in viral evolution. It further highlighted the presence of high number of low frequency polymorphisms which is directly linked to population size expansion [43]. The phylogenetic network comprised of two major networks interconnected through edges. The hatch marks depict the number of mutations present. The pie charts at the centre represent the haplotypes. The more number of similar haplotypes bigger is the circle. The colour codes are linked to the different geographical regions of India. The in house sequences did not show close proximity to the reference genome. The overall structure of the network suggests that due to acquisition of several number of mutations the BA strains are highly diverse than the reference genome NC_045512.2. Some of the sequences shared close proximity amongst each other and they are listed in Supplementary material 2.

The temporal evolution of BA/Omicron (based on spike protein mutations) with respect to the other variants was studied using a time-constrained Maximum Likelihood (ML) tree developed using the Jones-Taylor-Thronton model (Supplementary material 1). The spike protein sequence of the Wuhan-Hu-1 (NC_045512.2) strain was used as a reference outgroup. The nodes denote the different variants and the branch length corresponds to the time of evolution. The analysis revealed two clades, one comprising of Omicron variants and the other containing Delta and Delta-like. BA.1 and BA.1.1 shared relatedness amongst each other while they differed from BA.2 which was placed in a separate branch. On the other hand, Delta (B.1.617.2) and Delta-like (AY.4) were closely related and are placed in the same clade. The topology of the tree further predicted the time of evolution of the spike protein. From the tree topology, it can be inferred that Delta, Delta-like, and Omicron variants diverged from NC_045512.2. The substitutions accumulated within spike protein of Delta and Omicron and they evolved as new variants in early and late 2021 respectively. It can be further predicted that BA is not a sub-lineage of Delta or Delta-like, rather they have a common ancestor which is NC_045512.2.

3.4. Prediction of linear B-cell epitopes and their antigenicity

To elucidate the possibility of advanced vaccine development prediction of epitopes targeting the spike protein of the virus was done. Linear B cell epitope prediction was targeted. Representative full-length consensus spike protein amino acid sequences of BA.2, BA.1, BA.1.1, and NC_045512.2 were used for the analysis. The test revealed the presence of multiple epitope sites within the spike protein of the variants with low ($\leq$0.1) to high ($\geq$1.0) antigenicity values. Amino acid sequences which showed moderate to high ($\geq$0.5) antigenic properties were selected (Table 1). These epitope regions were present between 13 and 720 amino acids for BA whereas for reference strain this region extended till 1270. BA.2 showed the highest number of epitopes with strong antigenicity as compared to BA.1.1. Surprisingly, BA.1 showed only one site with high antigenicity. The structure prediction showed the epitopes to be mostly coiled in nature. The surface accessibility prediction by NetsurfP ascertained them to remain exposed or embedded on the surface of the spike monomer. Vaccines played a quintessential role in restricting the severity of disease caused by SARS-CoV-2 variants. However, in the case of Omicron, the first incidences from Botswana, Africa as well as India showed infection in fully vaccinated individuals [44]. Three-dimensional structures of Antibody-RDB complexes highlighted that Omicron is twice as

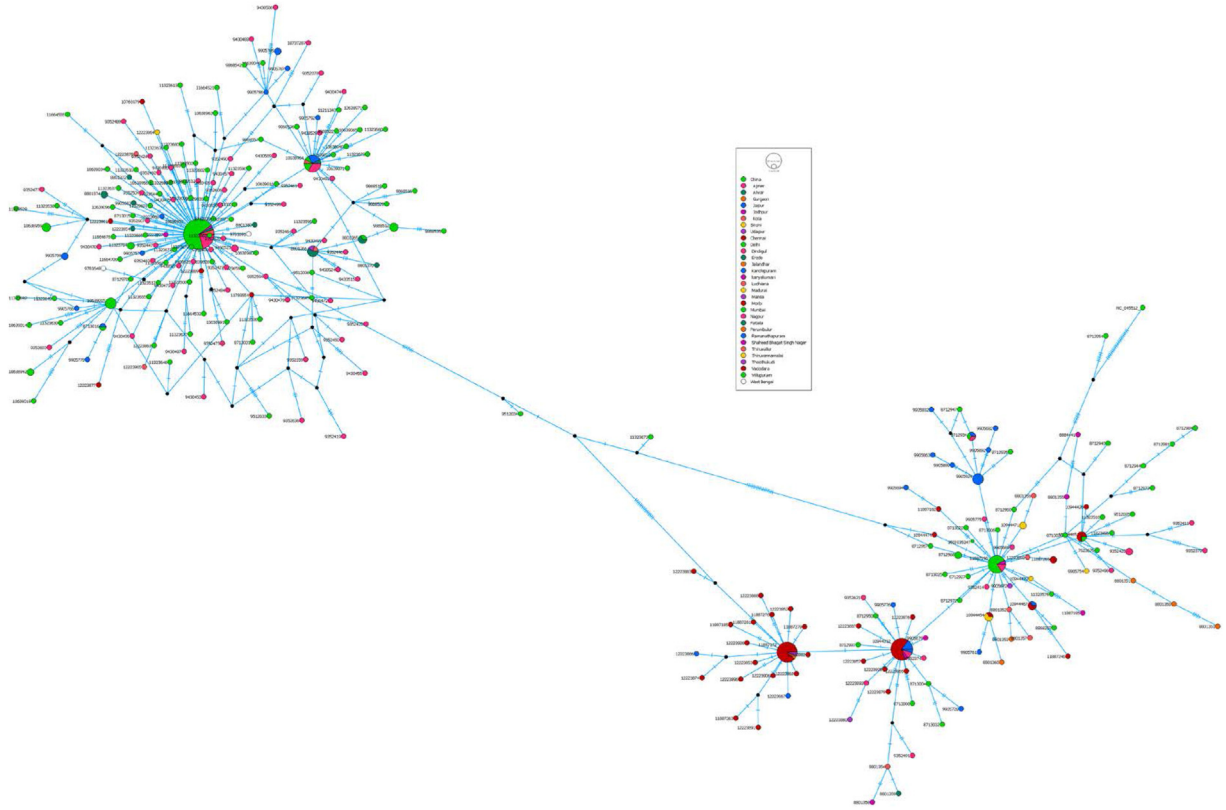

Fig. 3. Phylogenomic network of the BA variants. The lines depict the edges and the hatch marks represent the number of mutations accumulated. The pie charts indicate the number of haplotypes present. More the number of haplotypes bigger is the circle. The colour-codes demarcate the different states of India. The details of the accession number are provided in Supplementary material 2.

Table 1
Predicted linear B-cell epitopes and their antigenicity.

Variant name	Predicted epitope sequence	Region	Antigenicity[a]
BA.2	SNVTWFHAIHVSGTNGTKRFDNPVL	60–84	0.5298
	GKQGNFKNLREF	181–192	0.5402
	INLXRDLPQGFSA	210–222	1.7522
	VEKGIYQTSNFRVQPTES	308–325	0.7311
	VSPTKLNDLCFTNVYADS	382–399	1.3513
	TGXIADYNYKLPDDFT	415–430	0.9994
	KKSTNLVKNKCVNFNFNGLTGTG	528–550	1.1386
	SYQTQTKSXRRARSVASQSIIAYTMSLGAENSVAYSNNS	673–711	0.5263
BA.1	YQTQTKSHRRARSVASQSIIAYTMSLGAENSVAYSNNSIA	674–612	0.5848
BA.1.1	SQCVNLXTRTQLPPAYTNSFTRGV	13–36	0.7419
	FELLHAPATVCGPKKSTNLVKNKCVNFNFNGL	515–548	0.6509
	SYQTQTKSHRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSV	673–736	0.7582
NC_045512.2	QCVNLTTRTQLPPAYTNSFTRGV	14–36	0.7515
	FSNVTWFHAIHVSGTNGTKRFDN	59–81	0.6767
	DLEGKQGNFKNLRE	178–191	0.9256
	NSASFSTFKCYGVSPTKLNDLCFTNV	370–395	1.3609
	GDEVRQIAPGQTGKIADYNYK	404–424	1.3212
	AYTMSLGAENSVAYSN	694–709	0.6003
	SCCKFDEDDSEPVLKGVKL	1252–1270	0.6085

[a] Antigenicity values >0.5 are represented in the table.

likely to evade vaccine immunity as compared to Delta [29]. Hence, currently available vaccines may prove ineffective in providing immunity against Omicron. In this regard, there is a pressing need of designing more effective vaccines against Omicron and other SARS-CoV-2 variants.

4. Conclusions

BA variants with their highest number of mutations are more virulent among other SARS-CoV-2 strains. Thorough genomic characterization, epidemiological surveillance, phylogenetic inference and epitope prediction of BA will help in designing preventive measures against more deadly future variants of the SARS-CoV-2. This study will be beneficial for upgrading healthcare infrastructure as well as provide strong scientific reference for monitoring and prevention of future SARS-CoV-2 surges.

Funding

Council of Scientific and Industrial Research India and CSIR-Indian Institute of Chemical Biology supported the work through project MLP-132.

Ethics approval

Ethical clearances for the present study were taken from the CSIR-Indian Institute of Chemical Biology (IICB/IRB/2020/6) and MEDICA Super-specialty Hospital (CREC/2020/JUL/1(a)) human ethics committees. The work was executed under expert supervision following appropriate COVID-19 protocols.

Declaration of competing interest

All the authors declare that no competing interests exist.

Acknowledgment

The authors gratefully acknowledge the Council of Scientific & Industrial Research (CSIR), Ministry of Science and Technology, INDIA for providing structural and financial support through project MLP-132. The authors further thank the GISAID team for enabling data download and mutation analysis. Indian SARS-CoV-2 consortium (INSACOG) is duly acknowledged for providing SARS-CoV-2 RNA samples for this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmmb.2022.10.006.

References

[1] Cedro-Tanda A, Gómez-Romero L, de Anda-Jauregui G, Garnica-López D, Alfaro-Mora Y, Sánchez-Xochipa S, et al. Early genomic, epidemiological, and clinical description of the SARS-CoV-2 omicron variant in Mexico City. Viruses 2022;14(3):545.
[2] Saha P, Bose S, Srivastava AK, Chaudhary AA, Lall R, Prasad S. Jeopardy of COVID-19: rechecking the perks of phytotherapeutic interventions. Molecules 2021;26(22):6783.
[3] Dhar MS, Marwal R, Vs R, Ponnusamy K, Jolly B, Bhoyar RC, et al. Genomic characterization and epidemiology of an emerging SARS-CoV-2 variant in Delhi, India. Science 2021;374(6570):995–9.
[4] Sarkar P, Banerjee S, Saha SA, Mitra P, Sarkar S. Genome surveillance of SARS-CoV-2 variants and their role in pathogenesis focusing on second wave of COVID-19 in India. medRxiv; 2022.
[5] Yurkovetskiy L, Wang X, Pascal KE, Tomkins-Tinch C, Nyalile TP, Wang Y, et al. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. Cell 2020;183(3):739–51. e8.
[6] Furuyama TN, Antoneli F, Carvalho IM, Briones MR, Janini LM. Temporal data series of COVID-19 epidemics in the USA, Asia and Europe suggests a selective sweep of SARS-CoV-2 Spike D614G variant. 2020. arXiv preprint arXiv:200611609.
[7] Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 2020;182(4):812–27. e19.
[8] Volz E, Hill V, McCrone JT, Price A, Jorgensen D, O'Toole Á, et al. Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. Cell 2021;184(1):64–75. e11.
[9] Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, et al. SARS-CoV-2 spike mutations, L452R, T478K, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. Microorganisms 2021;9(7):1542. https://doi.org/10.3390/microorganisms9071542.
[10] Starr TN, Greaney AJ, Dingens AS, Bloom JD. Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. Cell Rep Med 2021;2(4):100255.
[11] Verghese M, Jiang B, Iwai N, Mar M, Sahoo MK, Yamamoto F, et al. A SARS-CoV-2 variant with L452R and E484Q neutralization resistance mutations. J Clin Microbiol 2021;59(7):e00741-21.
[12] Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, et al. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. Nature 2021;596(7871):276–80.
[13] CDC COVID-19 response Team. SARS-CoV-2 B. 1.1. 529 (Omicron) variant—United States, december 1–8, 2021. MMWR (Morb Mortal Wkly Rep) 2021;70(50):1731.
[14] Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. Nature 2022:1–10.
[15] Cameroni E, Bowen JE, Rosen LE, Saliba C, Zepeda SK, Culap K, et al. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. Nature 2022;602(7898):664–70.

[16] Wolter N, Jassat W, Walaza S, Welch R, Moultrie H, Groome M, et al. Early assessment of the clinical severity of the SARS-CoV-2 Omicron variant in South Africa. Medrxiv; 2021.
[17] Kreier F. Deltacron: the story of the variant that wasn't. Nature 2022;602(7895):19.
[18] Ou J, Lan W, Wu X, Zhao T, Duan B, Yang P, et al. Tracking SARS-CoV-2 Omicron diverse spike gene mutations identifies multiple inter-variant recombination events. Signal Transduct Targeted Ther 2022;7(1):1–9.
[19] Moisan A, Mastrovito B, De Oliveira F, Martel M, Hedin H, Leoz M, et al. Evidence of transmission and circulation of Deltacron XD recombinant SARS-CoV-2 in Northwest France. Clin Infect Dis 2022. https://doi.org/10.1093/cid/ciac360.
[20] Karbalaei M, Keikha M. Deltacron is a recombinant variant of SARS-CoV-2 but not a laboratory mistake. Ann Med Surg (Lond) 2022;79:104032. https://doi.org/10.1016/j.amsu.2022.104032.
[21] Farheen S, Araf Y, Tang YD, Zheng C. The Deltacron conundrum: its origin and potential health risks. J Med Virol 2022;94:5096–102. https://doi.org/10.1002/jmv.27990.
[22] Colson P, Fournier PE, Delerce J, Million M, Bedotto M, Houhamdi L, et al. Culture and identification of a "Deltamicron" SARS-CoV-2 in a three cases cluster in southern France. J Med Virol 2022;94(8):3739–49. https://doi.org/10.1002/jmv.27789.
[23] Combes P, Bisseux M, Bal A, Marin P, Latour J, Archimbaud C, et al. Evidence of co-infections during Delta and Omicron SARS-CoV-2 variants co-circulation through prospective screening and sequencing. Clin Microbiol Infect 2022;28(11):1503.e5–8. https://doi.org/10.1016/j.cmi.2022.06.030.
[24] Rockett RJ, Draper J, Gall M, Sim EM, Arnott A, Agius JE, et al. Co-infection with SARS-CoV-2 Omicron and Delta variants revealed by genomic surveillance. Nat Commun 2022;13(1):1–7.
[25] Leigh JW, Bryant D. POPART: full-feature software for haplotype network construction. Methods Ecol Evol 2015;6(9):1110–6.
[26] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 2018;35(6):1547.
[27] Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res 2017;45(W1):W24–9. https://doi.org/10.1093/nar/gkx346.
[28] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinf 2007;8(1):1–7.
[29] Tian D, Sun Y, Xu H, Ye Q. The emergence and epidemic characteristics of the highly mutated SARS-CoV-2 Omicron variant. J Med Virol 2022;94:1383–2376. https://doi.org/10.1002/jmv.27643.
[30] Nguyen HL, Thai NQ, Nguyen PH, Li MS. SARS-CoV-2 Omicron variant binds to human cells more strongly than the wild type: evidence from molecular dynamics simulation. J Phys Chem B 2022;126(25):4669–78.
[31] Lupala CS, Ye Y, Chen H, Su X-D, Liu H. Mutations on RBD of SARS-CoV-2 Omicron variant result in stronger binding to human ACE2 receptor. Biochem Biophys Res Commun 2022;590:34–41.
[32] Sabir DK. Analysis of SARS-COV2 spike protein variants among Iraqi isolates. Gene Rep 2022;26:101420.
[33] Mertens J, Coppens J, Loens K, Le Mercier M, Xavier BB, Lammens C, et al. Monitoring the SARS-CoV-2 pandemic: screening algorithm with single nucleotide polymorphism detection for the rapid identification of established and emerging variants. Clin Microbiol Infect 2022;28(1):124–9.
[34] Alenquer M, Ferreira F, Lousa D, Valério M, Medina-Lopes M, Bergman M-L, et al. Signatures in SARS-CoV-2 spike protein conferring escape to neutralizing antibodies. PLoS Pathog 2021;17(8):e1009772.
[35] Han Y, Wang Z, Wei Z, Schapiro I, Li J. Binding affinity and mechanisms of SARS-CoV-2 variants. Comput Struct Biotechnol J 2021;19:4184–91.
[36] Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, Oosthuysen B, et al. SARS-CoV-2 501Y. V2 escapes neutralization by South African COVID-19 donor plasma. Nat Med 2021;27(4):622–5.
[37] Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S, et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. Cell 2020;182(5):1284–94. e9.
[38] Starr TN, Greaney AJ, Hilton SK, Ellis D, Crawford KH, Dingens AS, et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. Cell 2020;182(5):1295–310. e20.
[39] Greaney AJ, Starr TN, Gilchuk P, Zost SJ, Binshtein E, Loes AN, et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. Cell Host Microbe 2021;29(1):44–57. e9.
[40] VanBlargan LA, Errico JM, Halfmann PJ, Zost SJ, Crowe JE, Purcell LA, et al. An infectious SARS-CoV-2 B. 1.1. 529 Omicron virus escapes neutralization by therapeutic monoclonal antibodies. Nat Med 2022;28(3):490–5.
[41] Syed AM, Ciling A, Taha TY, Chen IP, Khalid MM, Sreekumar B, et al. Omicron mutations enhance infectivity and reduce antibody neutralization of SARS-CoV-2 virus-like particles. Proc Natl Acad Sci USA 2022;119(31):e2200592119.
[42] McCallum M, Czudnochowski N, Rosen LE, Zepeda SK, Bowen JE, Walls AC, et al. Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement. Science 2022;375(6583):864–8.
[43] Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 1989;123(3):585–95.
[44] Chen J, Wang R, Gilby NB, Wei G-W. Omicron variant (B. 1.1. 529): infectivity, vaccine breakthrough, and antibody resistance. J Chem Inf Model 2022;62(2):412–22. https://doi.org/10.1021/acs.jcim.1c01451.