Emergence of Two Distinct Variants of SARS-CoV-2 and Explosive Second Wave of COVID-19: An Experience From A Tertiary Care Hospital, Pune, India

Shubham Shrivastava

Interactive Research School for Health Affairs

Suhas T Mhaske

Interactive Research School for Health Affairs

Meera S Modak

Bharati Vidyapeeth Deemed University Medical College

Rashmi G Virkar

Interactive Research School for Health Affairs

Shamburaje S Pisal

Interactive Research School for Health Affairs

Akhilesh Chandra Mishra

Interactive Research School for Health Affairs

Vidya A Arankalle (✉ varankalle@yahoo.com)

Interactive Research School for Health Affairs <https://orcid.org/0000-0002-9940-3437>

Research Article

Keywords: Variant, Kappa, Delta, COVID-19, RBD, VoC, genomes, co-circulation, SARS-CoV-2

Posted Date: September 3rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-824960/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Archives of Virology on January 9th, 2022. See the published version at <https://doi.org/10.1007/s00705-021-05320-7>.

Abstract

The emergence of novel variants of SARS-CoV-2 in several countries has been associated with the increased transmission/ reduced neutralization potential of antibodies against the Wuhan virus (wild type). Currently, India is recovering from an explosive second wave of COVID-19. This prospective study was conducted quarterly for one year (May 2020 to June 2021) at a tertiary care hospital, Pune city, western India. Receptor binding domain (RBD, n = 319) and full genome (n = 20) sequences from viral RNA positive nasopharyngeal swabs of COVID-19 patients representing first and second waves were used for analysis. We did not detect Brazil, South Africa and California variants. Till December, only wild type strain was prevalent. Concurrent with the upsurge of second wave in March-2021, 73% (33/45) of RBD sequences harboured L452R/E484Q mutations characteristic of Kappa variant. In April, co-circulation of Kappa (37%) and Delta (L452R/T478K, 59%) variants was recorded. During May and June, delta variant became the predominant circulating variant and coincided with a significant decline in the number of COVID-19 cases. Of the 20 full genome sequences, six isolates each exhibited signature mutations of Kappa and Delta variants respectively. With several states witnessing reduction in the number of COVID-19 cases, continuous monitoring of newer mutations and assessment of these mutations in influencing virus transmissibility and impact on vaccinated / previously exposed individuals would be necessary.

1. Introduction

The pandemic of COVID-19 disease caused by SARS-CoV-2 is in the second year and continues to be the global public health concern. Several countries are experiencing resurgence leading to second/third waves of the disease [1]. Several vaccines have been approved / approved for emergency use and are being used in different countries as per the availability and national policies [2]. However, the proportion of the global vaccinated population remains low.

The original SARS-CoV-2 strain from Wuhan (Wild type) was rapidly transmitted to a large number of countries through infected travellers followed by establishment of community transmission and further rapid spread. In India, the first COVID-19 case was reported on 29th January 2020 in a student returning from China [3]. At the peak in the month of September, 97,860 cases were recorded [4]. Subsequently, the number of cases reduced considerably with lowest cases in the months of January and February. The second wave started from the middle of March 2021 with highest number of cases (414,188) recorded on 6th May 2021. Currently, India is experiencing significant drop in the number of active infections to 41,831 as on 31st July 2021. The state of Maharashtra was the first to report increase in the number of cases during the second wave and remains the worst affected state, so far.

Though mutation is an ongoing process for RNA viruses, the identification of a variant harbouring a set of mutations in the spike protein accompanying enhanced transmissibility of SARS-CoV-2 in UK (UK variant, B.1.1.7, alpha) was alarming [5, 6]. Subsequently, additional variants of concern in terms of higher transmissibility, virulence and impact of vaccines currently in use were identified in countries from different continents. These include the South Africa variant (B.1.351, beta) [7], the Brazil variants (P.1 and P.2, gamma) [8, 9] California variants (B.1.429, epsilon and B.1.427) [10, 11] and the most recent Indian variant (B.1.617.2, delta) [12].

The first case in the state of Maharashtra was reported on 19th March 2020 from Pune, a metropolitan city with 50,57,709 population that has remained the major hotspot. To monitor SARS-CoV-2 strains over time on a quarterly basis, a single-centre study was undertaken at Pune. The changing variants over the two waves of the disease are reported here.

2. Materials And Methods

2.1 Clinical specimens

In the month of May'2020, 8 nasopharyngeal swab (NPS) specimens with confirmed COVID-19 were collected from a designated COVID treatment facility. At the beginning of the pandemic, diagnostic facilities were available only at a few national laboratories. From June first week onwards, BVDUMC, a tertiary care hospital at Pune catering to a population of 7.5 million (2011 census), was permitted to undertake real time PCR (RT-PCR) testing for COVID-19 diagnosis. Following the approval, NPS specimens collected from the suspected COVID-19 patients were tested for the presence of SARS-CoV-2 RNA. Depending on the availability, three different kits were used (Mylab, India; Covipath, Thermofisher, USA and Aargene, France) for RT-PCR. NPS specimens were subjected to RNA extraction (Qiagen, USA) and RT-PCR was performed according to the instructions provided by the manufacturers. SARS-CoV-2 positive NPS were stored in aliquots at -80°C till tested. Samples showing Ct values of < 20 in RT-PCR were selected for amplifying RBD region (669 nt) using PCR. The number of such samples analysed at different times were: May 2020 (n = 5), September 2020 (n = 10), December 2020 (n = 15), March 2021 (n = 45) including 20 samples collected in the 1st week and 25 samples in the 4th week of the month, Apr 2021 (n = 91), May 2021 (n = 132) and June 2021 (n = 21).

2.2 Viral RNA extraction, PCR and sequencing

Viral RNA was extracted using QIAmp viral RNA mini kit (Qiagen CA, USA) and cDNA was synthesized using high-capacity reverse transcription kit using reverse primer (RBD-R). The PCR product was amplified using AmpliTaq polymerase, forward RBD-F (5' GCGGTCTAGAATGAATATTACAACTTGTGCCCT 3') and reverse RBD-R (5' TACATGCACCACTGTTAACTCGAGCCAT 3') primers. PCR products (~ 600 bp) were gel purified and sequenced from both the ends using Big Dye terminator cycle sequencing kit (Thermofisher Scientific, MA, USA). RBD sequences were confirmed by BLAST (www.ncbi.nlm.nih.gov/BLAST). The forward and reverse sequences were aligned and manually edited using Codon Code aligner v.9.0.1 software to obtain the consensus sequence. Partial RBD sequences were submitted to GenBank at www.ncbi.nlm.nih.gov (Supplementary Table 1).

2.3 Sequencing of SARS-CoV-2 full genomes

SARS-CoV-2 full genomes were sequenced from NPS specimens using Ion Torrent targeted next generation sequencing platform (Thermofisher Scientific, MA, USA). The analysis was performed on a Torrent Suite Server v5.14 using plugins specifically for the Ion Ampliseq SARS-CoV-2 panel including IRMA report to build a consensus sequence. The consensus sequences were assembled using Trinity assembler v 1.3.0.2 and contigs were annotated using NCBI database.

Complete genome sequences were submitted to GenBank at www.ncbi.nlm.nih.gov (accession number MT416725, MT416726, MW969752 to MW969758, MZ021503 to MZ021506, MZ574051 to MZ574056). One sequence was deposited in GISAID database (EPI_ISL_1710598).

2.4 Sequence analysis

The sequences obtained during the present study and other sequences retrieved from GenBank and GISAID were aligned against Wuhan-Hu-1 reference strain (NC045512) using MAFFT online alignment tool (v7.475, 2020). Pangolin v2.1.6 ([www.github.com/cov-lineages/pangolin](https://github.com/cov-lineages/pangolin)) was used to determine the lineages. The Nextclade web tool (v0.14.2) (<https://clades.nextstrain.org>) was used to compare the generated sequences to reference strains to assign the clades.

3. Results

3.1 Comparison of COVID-19 cases in Pune and diagnosed at Bharati hospital (March 2020-June 2021)

Figure 1A depicts month wise number of COVID-19 cases in Pune city as per the data provided by Pune Municipal Corporation [13]. From the first case detected on 19th March 2020, the number of cases increased steadily till June with a sharp rise in the month of July. At peak in September 2020, 49,918 cases were reported that was followed by sharp decline and low number of cases till February 2021. In March 2021, there was a sharp increase again that reached at its peak with 150,175 cases in the month of April. From 3rd week of April, a sharp decline was observed continuing through the months of May and June. During these two months, 8,590 cases were reported. Figure 1B presents the number of patients advised and with confirmed COVID-19 diagnosis at Bharati Vidyapeeth (Deemed to be University) Medical College following permission of the government to provide diagnostic services. Clearly, the patterns of number of COVID-19 cases detected at different times were identical when Pune city and Bharati hospital were compared.

Table 1
SARS-CoV-2 RBD mutant profiles during the first (2020) and second (2021) waves of COVID-19 at Pune, western India*

Collection Month (No. of samples sequenced)	Original (Wuhan)	UK Alpha (N501Y)					
		UK Alpha (N501Y)	Indian**		Others		
Kappa (L452R, E484Q)	Delta (L452R, T478K)	N440K	E484K	E484Q			
May 2020 (5)	5 (100%)	0	0	0	0	0	0
Sep 2020 (10)	10 (100%)	0	0	0	0	0	0
Dec 2020 (15)	13 (87%)	0	0	0	2 (13%)	0	0
Mar 2021 (45)	4 (8.9%)	0	33 (73.3%)	2 (4.4%)	2 (4.4%)	3 (6.7%)	1 (2.2%)
Apr 2021 (91)	1 (1.1%)	1 (1.1%)	34 (37.3%)	54 (59.3%)	1 (1.1%)	0	0
May 2021 (132)	0	0	10 (7.6%)	122 (92.4%)	0	0	0
June 2021 (21)	0	0	0	21 (100%)	0	0	0

*Mutants specific to Brazil/South Africa were not found. **Additionally in 2021, V382L mutation was recorded 14/289 (4.8%).

3.2 RBD sequence analysis

In view of the role of RBD in the attachment of the virus to ACE2 receptor, neutralization of the virus by antibodies and as a major target for vaccine development, monitoring of RBD sequences over time was undertaken. Table 1 describes mutational analysis of the RBD region over a period of one year (May 2020 to Jun 2021). Till December 2020, none of the samples exhibited characteristic mutations (K417N/T, E484K, N501Y, T478K) documented for the variants of concern identified so far. 2/15 samples from December 2020 did show N440K mutation. In the beginning of March 2021, when the number of cases started showing an upward trend (Fig. 1B), we did not find N501Y mutation characteristic of the UK variant (shared by Brazil and SA variants) as well as K417N mutation in the Brazil/SA variants.

Strikingly, 70% of the NPS specimens sequenced during the beginning of March 2021 and 83% NPS samples sequenced at the end of March exhibited L452R mutation observed in the California variant. In addition, contrary to E484K mutation seen in the strains from Brazil and South Africa, Indian strains exhibited E484Q mutation, defined later as kappa variant. Overall, in the month of March, 33/45 (73%) samples harbour India-specific, L452R/E484Q mutations. Week-wise analysis revealed the same trend, with 70% and 76% of kappa-related cases in 1st and last week of Mar 2021. Simultaneous exponential rise in the number of COVID-19 cases (Fig. 1B) revealed a clear association of emergence of this mutant with the current second wave of the disease. Unfortunately, we did not collect samples during January or February and hence the possibility of earlier detection was missed. As specified in the Table 1, we did identify other unique mutations in a small proportion of NPS specimens.

Analysis of 91 RBD sequences from the samples collected during the month of April led to striking observations. At this time, 34 variants were L452R/E484Q (37%) while 54 sequences harboured L452R/T478K mutations (59%), characteristic of the delta variant. One sequence each exhibited N501Y mutation characteristics of UK strain, N440K mutation and no mutation. Overall, the L452R mutation increased from 0/30 till December to 123/136 (90%) by April ($p < 0.001$). Importantly, as compared to March (35/45), a significant rise in L452R mutants was seen in April (88/91, $p = 0.001$). Interestingly at the same time, T478K mutation increased drastically from 4% in March to 59% in April ($p < 0.001$). The increasing trend of the dominant mutation, L452R/T478K continued in the month of May and June however, we observed significant decline in the number of patients seeking COVID-19 diagnosis at the hospital (Table 1). Our results indicated that the dominant clade G virus strains seems to have been replaced by the variants of concern, L452R/T478K in Pune during the second wave of the disease.

3.3 Analyses based on SARS-CoV-2 full genome sequences from Pune, India

During 2020-21, we sequenced 20 full SARS-CoV-2 genomes from Pune (2020: May: 2, September: 2; 2021: March: 10, April: 2, May: 4). The four mutations, C241T, C3037T, C14408T, and A23403G were observed in all the 20 genomes from the clade "G" (named after the Spike D614G mutation). Of the 10 genomes from March 2021 selected for sequencing, four were wild type while 6 were Kappa variant (B.1.617.1) as per RBD analysis.

As shown in Table 2, two sequences each from May and September 2020 belonged to lineage B.1.1.306 and clade 20B. Of the 10 genomes from March 2021, 2 sequences (211295 – MW969753 and 210761 – MZ021503) belonged to the original prevalent lineage B.1.1.306, clade 20B while the other two sequences, 210922 (MW969752) and 210896 (EPI_ISL_1710598) categorized into separate lineages, B.1.1 (clade 20B) and B.1.36.29 (clade 20A) respectively, suggestive of simultaneous low level of circulation of the wild type virus. Remaining six sequences from March 2021 (210871-MZ021506, 210927-MZ021505, 210929-MZ021504, 211290-MW969754, 211294-MW969755 and 211406-MW969756) formed a distinct lineage B.1.617.1, Kappa variant (clade 21B). Two sequences from Apr 2021 (212095, 212098) and 4 sequences from May 2021 (213366, 213522, 213523 and 213570) formed a distinct clade 21A and belonged to B.1.617.2 lineage (Delta variant, VoC).

The complete genome sequence analysis of sample, 210896 (EPI_ISL_1710598) revealed in-frame stop codon in Orf3a protein at 261 amino acid position (26173 nucleotide position). The first strain with this amino acid change was reported in Germany (hCoV-19/Germany/NW-KRO-2355/2020) in March 2020 and recently in England in June 2021 (hCoV-19/England/HSLL-17FBAD3/2021). It has been reported in 35 countries so far (<https://www.gisaid.org/epiflu-applications/covsurver-mutations-app/>).

Spike protein

Next, we compared spike protein from the Indian variants with other known variants (Table 2). We found two clusters of Indian variants, B.1.617.1 (Kappa) and B.1.617.2 (Delta) and were distinct from the other variants of concern. With the observed D614G mutation, clade G continued to be the only clade circulating so far in Pune. Mutations specific to the recently emergent variants (UK, South Africa, Brazil and California) were not shared by the Indian variants, however, signature mutations of Delta variant were observed in all six sequences obtained in Apr and May months of year 2021.

Two sequences from 2021 (211295 – MW969753 and 210761 – MZ021503) belonging to the original prevalent lineage B.1.1.306 had four characteristic mutations, L18F, A27S, E484K and Q675H in the spike protein that was missing in all the four sequences from 2020 suggesting that the SARS-CoV-2 is continuously evolving. Two genomes, 210922 (MW969752) and 210896 (EPI_ISL_1710598) had unique mutations at V143F, Q677H and at N440K respectively (Table 2). All the six Indian variants of B.1.617.1 lineage exhibited 5 unique mutations at G142D, L452R, E484Q, P681R, and Q1071H. Additionally, E154K mutation was recorded in 5/6 and V382L, D1153Y mutations were recorded in 4/6 Kappa variants. For Delta variant, all the 6 sequences belonged to B.1.617.2 lineage with 6 unique mutations at T19R, G142D, L452R, T478K, P681R, D950N and deletion at F157 and R158 aa positions with substitution of E156G. Notably, P681R mutation was adjacent to the furin cleavage site (682–685). *In-vitro* experiments conducted with site-directed mutations have shown that this mutation leads to increase in the fusion activity of SARS-CoV-2 spike protein [14]. Importantly, in the UK variant B.1.1.7, P681H was recorded (Table 2). The acquisition of A222V and K417N in Delta variant (AY.2 lineage) was associated with the upsurge in COVID-19 cases in Europe and USA and these amino acid changes were missing in the Delta variants of this study.

Pangolin Lineages	Sequence ID*	N-terminal domain in S1 subunit														RBD				
		AA position														AA position				
		18	19	20	26	27	70	95	138	142	143	154	171	190	222	382	417	433	440	
B	NC-045512, Wuhan	L	T	T	P	A	V	T	D	G	V	E	V	R	A	V	K	V	N	
B.1.1.306	8003-IRSHA (MT416725)																			
	8004-IRSHA (MT416726)																			
B.1.1.306	CD208550 (MW969758)																			
	CD208560 (MW969757)																	I		
B.1.1	CD210922 (MW969752)										F									
B.1.36.29	CD210896 (EPI_ISL_1710598)																		K	
B.1.1.306	CD211295 (MW969753)	F				S							I							
B.1.1.306	CD210761 (MZ021503)	F				S														
B.1.617.1	CD210871 (MZ021506)							I		D										
B.1.617.1	CD210927 (MZ021505)									D		K				L				
B.1.617.1	CD210929 (MZ021504)									D		K				L				
B.1.617.1	CD211290 (MW969754)									D		K				L				
B.1.617.1	CD211406 (MW969756)									D		K				L				
B.1.617.1	CD211294 (MW969755)									D		K								
B.1.617.2	CD212098 (MZ574054)		R							D					V					
B.1.617.2	CD213366 (MZ574051)		R							D										
B.1.617.2	CD212095 (MZ574052)		R							D										
B.1.617.2	CD213570 (MZ574053)		R							D										
B.1.617.2	CD213522 (MZ574055)		R							D										
B.1.617.2	CD213523 (MZ574056)		R							D										
B.1.617.2	EPI_ISL_2880930 (India/HR)		R												V					
B.1.617.2	EPI_ISL_2272481 (India/DL)		R							D					V					
B.1.617.2	EPI_ISL_2036277 (India/WB)		R							D										
B.1.617.2	EPI_ISL_2897823 (India/MH)		R																	
B.1.617.2	EPI_ISL_2956019 (Sweden)		R							D					V					

*Sequence IDs in bold denote sequences from India; Black: wild type and red: variant viruses

Pangolin Lineages	Sequence ID*	N-terminal domain in S1 subunit														RBD											
		AA position														AA position											
		18	19	20	26	27	70	95	138	142	143	154	171	190	222	382	417	433	440								
B.1.617.2	EPI_ISL_2958767 (France)	R																									
AY.2	EPI_ISL_2923711 (USA/CA-CDPH)	R																		F		V		N			
AY.2	EPI_ISL_2545667 (USA/HI-TAMC)	R																		F		D		V		N	
AY.2	EPI_ISL_2929274 (USA/CA-CDC)	R																		F		D		V		N	
AY.2	EPI_ISL_2928214 (USA/CA-OC)	R																		F		V		N			
B.1.1.7 (UK)	EPI_ISL_601443																										
B.1.1.7 (L18F)	EPI_ISL_720875	F																									
B.1.1.7 (F490S)	EPI_ISL_736026																										
B.1.1.7 (S494P)	EPI_ISL_741039																										
B.1.1.7 (E484K)	EPI_ISL_782148																										
B.1.351 (SA)	EPI_ISL_1012924																			N							
P.1 (Brazil)	EPI_ISL_792681	F		N	S				Y					S			T										
P.1 (Brazil)	EPI_ISL_875566	F		N	S				Y					S			T										
P.1 (Brazil)	EPI_ISL_875567	F		N	S				Y					S			T										
P.1 (Brazil)	EPI_ISL_875568	F		N	S				Y					S			T										
B.1.427 (California, L452R)	EPI_ISL_1620465																										
B.1.429	EPI_ISL_824555																										
B.1.525 (E484K)	EPI_ISL_1615794																										
B.1.2	EPI_ISL_824741																										
B.1.1.432	EPI_ISL_913915																										
*Sequence IDs in bold denote sequences from India; Black: wild type and red: variant viruses																											

Table 3
Mutational analyses in nucleocapsid protein with reference to Wuhan strain (accession number, NC045512).

Mutational analyses in nucleocapsid protein with reference to Wuhan strain (accession number, NC045512).													
Domains of nucleocapsid protein		N-terminal domain					SR-rich		Linker region				
Amino acid positions		2	3	63	80	119	142	194	203	204	236	377	
Pangolin Lineages	Sequence ID*												
B	NC-045512 (Wuhan)	S	D	D	P	A	P	S	R	G	S	D	
B.1.1.306	8003-IRSHA (MT416725)								K	R			
	8004-IRSHA (MT416726)								K	R			
B.1.1.306	208550 (MW969758)								K	R			
	208560 (MW969757)								K	R			
B.1.1	CD210922 (MW969752)								K	R			
B.1.36.29	CD210896 (EPI_ISL_1710598)	P							L				
B.1.1.306	CD211295 (MW969753)	Y						S	K	R			
B.1.1.306	CD210761 (MZ021503)	Y			T					K	R		
B.1.617.1	CD210871 (MZ021506)								M				Y
B.1.617.1	CD210927 (MZ021505)								M				Y
B.1.617.1	CD210929 (MZ021504)								M				Y
B.1.617.1	CD211290 (MW969754)								M				Y
B.1.617.1	CD211406 (MW969756)								M				Y
B.1.617.1	CD211294 (MW969755)								M				Y
B.1.617.2	CD212098 (MZ574054)				G				M				Y
B.1.617.2	CD213366 (MZ574051)				G				M				Y
B.1.617.2	CD212095 (MZ574052)				G				M				Y
B.1.617.2	CD213570 (MZ574053)				G				M				Y
B.1.617.2	CD213522 (MZ574055)				G				M				Y
B.1.617.2	CD213523 (MZ574056)				G				M				Y
B.1.617.2	EPI_ISL_2880930 (India/HR)				G				M				Y
B.1.617.2	EPI_ISL_2272481 (India/DL)				G				M				Y
B.1.617.2	EPI_ISL_2036277 (India/WB)				G				M				Y
B.1.617.2	EPI_ISL_2897823 (India/MH)				G				M				Y
B.1.617.2	EPI_ISL_2956019 (Sweden)				G				M				Y
B.1.617.2	EPI_ISL_2958767 (France)				G				M				Y
AY.2	EPI_ISL_2923711 (USA/CA-CDPH)				G				M				Y
AY.2	EPI_ISL_2545667 (USA/HI-TAMC)				G				M				Y
AY.2	EPI_ISL_2929274 (USA/CA-CDC)				G				M				Y
AY.2	EPI_ISL_2928214 (USA/CA-OC)				G				M				Y
B.1.1.7 (UK)	EPI_ISL_601443	L							K	R	F		
B.1.1.7 (L18F)	EPI_ISL_720875	L							K	R	F		
B.1.1.7 (F490S)	EPI_ISL_736026	L							K	R	F		
B.1.1.7 (S494P)	EPI_ISL_741039	L							K	R	F		
B.1.1.7 (E484K)	EPI_ISL_782148	L							K	R	F		
B.1.351 (SA)	EPI_ISL_1012924												
P.1 (Brazil)	EPI_ISL_792681				R				K	R			

*Sequence IDs in bold denote sequences from India; Black: wild type and red: variant viruses

Domains of nucleocapsid protein		N-terminal domain	SR-rich	Linker region
P.1 (Brazil)	EPI_ISL_875566	R	K	R
P.1 (Brazil)	EPI_ISL_875567	R	K	R
P.1 (Brazil)	EPI_ISL_875568	R	K	R
B.1.427 (California, L452R)	EPI_ISL_1620465			
B.1.429	EPI_ISL_824555			
B.1.525 (E484K)	EPI_ISL_1615794			
B.1.2	EPI_ISL_824741			
B.1.1.432	EPI_ISL_913915		K	R
*Sequence IDs in bold denote sequences from India; Black: wild type and red: variant viruses				

Nucleoprotein

The Indian variants were distinct in this protein as well (Table 3). The Indian variants did not share the D3L, S236F and P80R mutations seen in the UK and Brazil variants respectively. The two Indian sequences (21295, 210761) had tyrosine (Y) instead of Leucine (L) at 3rd amino acid position as observed in the UK strain. The R203M and D377Y mutation in all the twelve Indian variants. At 203, the change is from hydrophilic Arginine to hydrophobic Methionine while at 377 position, Aspartic acid is changed to Tyrosine. The biological significance of these mutations needs to be evaluated. Remaining Indian sequences continued to harbour R203K and G204R mutations that was present since May 2020. One Indian sequence (210896_EPI_ISL_1710598) was unique with no amino acid substitution at 203 and 204 positions. The six Indian variant exhibited unique mutation at D63G amino acid position as observed in Delta variant.

Other genomic regions:

We compared the other genome regions as well. All the six kappa variant-specific mutations included: Nsp3-T749I, Nsp6-T77A (Orf1a), Nsp12-P323L, Nsp13-M429I, Nsp15-K259R (Orf1b), S26L (Orf3a), I33T (Orf6) and V82A (Orf7a). The other proteins (E, M, Orf7b, Orf8 and Orf10) in Kappa variant remained unchanged. All the six Indian delta variant-specific mutations included: G210T (5'UTR) Nsp2-P129L, Nsp3-P822L, Nsp4-A446V, Nsp6-V149A (Orf1a), Nsp12-P323L, Nsp12-G671S, Nsp13-P77L (Orf1b), S26L (Orf3a), I82T (M), V82A, T120I (Orf7a) and deletions at D (119), F (120) positions in Orf8 region. The other proteins (E, Orf6, Orf7b and Orf10) remained unchanged in the delta variant.

4. Discussion

We report here the emergence of two variants of SARS-CoV-2 associated with the progression of the second wave of COVID-19 in Pune, India. Though our study is restricted to a single hospital, it does reflect situation in the city. The variants of concern (VoC) identified as, UK, Brazil, and SA were not detected during the first wave of disease in the year 2020. This observation indicated that replacement of Clade L of original Wuhan strain with Clade G (spike D614G mutation) as the predominant circulating strain could be the reason for the spread of SARS-CoV-2 during the first wave of disease in India. All the Pune strains exhibited this mutation and belonged to the clade G. Notably, The D614G mutation was shown to enhance rate of SARS-CoV-2 infection in the upper airways [15, 16].

The second wave in India was first noticed in the state of Maharashtra to which Pune city belongs. Emergence of the second wave was characterized by the appearance of the kappa variant in the first week of the month of March 2021 (Table 1). Within weeks, the virus evolved further, replacing E484Q mutation with T478K named as the delta variant. L452R mutation was maintained. Subsequently, both variants co-circulated followed by the dominance of the delta variant. At this time, the cases did show a sharp decline.

Being a large country with distinct geographic and socioeconomic differences, the second wave of COVID-19 was recorded in different states at different time points after the state of Maharashtra. As a result, as against the early peak in Pune and state of Maharashtra, the country peak was of the second wave was recorded in the month of May [13, 17]. It was noticed that the UK variant was predominant in some states in north India, Punjab and Delhi [18].). At the time of submission of this manuscript, similar observation on co-circulation of novel Kappa and Delta variants was reported from Maharashtra [12]. According to WHO, kappa and delta variants accounted for 21% and 7% of sequenced samples from India in late April [19]. However, in a recent study from Houston, USA [20] all the known mutants in variable proportions were detectable from single centre.

Outside India, United Kingdom reported upsurge in largest number of COVID-19 cases of delta variant prompting WHO to declare delta variant as variant of concern on 6th May 2021. In 3rd week of June, rise in 79% of COVID-19 cases in UK were due to new infections by delta variant. Around the same time, delta variant accounted for more than 20% of cases in USA and reported to have a higher growth rate than the UK variant [19, 21]. According to the technical briefings by Public health England, although the delta variant is now dominant strain in UK, the hospital admissions and deaths are not rising at the same rate as infectivity rate [22, 23].

In the kappa variant with L452R and E484Q mutations in the RBD region, substitution of hydrophobic leucine (L) with hydrophilic arginine (R) is similar to the California variant that was shown to correlate with increased transmissibility and decreased antibody neutralization [24]. At 484 position, as against lysine in the SA and Brazil variants, the kappa variant exhibited glutamine, though E484K was present in two different sequences without any change at 452. The

E484K variant was found to be of serious concern and a potential immune escape because of no/partial effect of some of the therapeutic monoclonal antibodies and due to decreased *in-vitro* virus neutralization by the serum samples derived from the patients infected with the original SARS-CoV-2. Here, an elegant study by Greaney et al (2021) is noteworthy [25]. By employing deep-mutational scanning method, the effect of even single mutations in the RBD region on binding / neutralization of well characterized plasma samples from COVID-19 patients was determined. Of note, mutations at site E484 were most striking in terms of reduction in the neutralization titers. More importantly, mutations to K, Q or P led to 10-60-fold decrease in the neutralizing antibody titers.

As far as delta variant is concerned, T478 is located near the interface with ACE2. Moderate resistance to neutralization by two monoclonal antibodies as well as convalescent serum from two patients was documented with T478I mutation [26]. An increased affinity of T478K mutant to ACE2 in *in-vitro* experiments has been shown [27]. This is clearly reflected by the delta variant associated increase in COVID-19 cases in many countries. It is pertinent to note that the recent increase in COVID-19 cases in New York, USA was observed to be due to the introduction of B.1.243 lineage with a P681H and T478K mutations [28]. Immunoreactivity of the variants currently in circulation in Pune with antibodies induced by the wild type viral infection and vaccines would respectively decide degree of reinfections / infections among previously infected COVID-19 patients and vaccine recipients. In a vaccine breakthrough infection in healthcare workers across three centres in Delhi, India, B.1.617.2 had a 3.2-fold higher transmissibility than B.1.1.7 [29]. In line with the same observations, Delta variant caused higher rate of re-infections when compared to Alpha strain in UK [23].

Interestingly, two Indian strains (211295-MW969753, 210761-MZ021503) had combination of two substitutions, L18F and E484K which may have significant role in disease transmission that needs further evaluation. L18F substitution was dominant in England before the emergence of UK strain. It was shown that the percentage of genomes with spike L18F mutation increased from 35% in September to 43% in October and 52% in November 2020 suggesting the beneficial role of this mutation in virus transmission [30].

Apart from the mutations of concern described above, V382L (9/55, 16.4%) and N440K (4/55, 7.3%) substitutions were recorded in the 2021 sequences (Table 1). The prevalence of RBD variant with N440K was estimated to be 2.1% in India with ~ 34% in Andhra Pradesh, a southern Indian state [31]. N440K mutation was also reported in COVID-19 reinfection cases from north India [32]. N440K mutation showed enhanced binding affinity to the human ACE2 receptor resulted in immune escape from a panel of neutralizing monoclonal antibodies [33, 34].

Protease cleavage site (675–692) in the spike protein is most prone to mutation. Both kappa and delta variants exhibited mutation at P681 position located adjacent to the furin cleavage site (682–685) that was shown to facilitate efficient SARS-CoV-2 transmission and infection [14, 35]. The rapid global spread of UK variant with P681H, P681R mutations is well known. Though we did recognize substitutions in the other regions of the genome, the significance remains unclear.

The Delta variant is continuously evolving and a new variant – AY.1 (delta plus, B.1.617.2.1) carrying an additional K417N mutation is emerging in India and UK. This particular mutation was first identified in B.1.351, beta variant in Brazil. Acquisition of spike K417N and A222V mutations in Delta, resulted in the emergence of AY.2 and being reported to be associated with upsurge in new COVID-19 infections in the United States.

In summary, the second wave of COVID-19 in Pune in March 2021 was associated with the first emergence of kappa variant followed immediately by the delta variant that almost completely replaced the kappa variant. Though the cases have declined significantly since May, molecular surveillance needs to be continued. Monitoring of the emergence of newer mutations and assessment of the association with virus transmission vaccinated/previously exposed individuals remains of utmost importance.

Declarations

Funding

This work was supported by Department of Biotechnology (DBT) – BIRAC (grant number: BT/NBM0095/02/18). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Acknowledgements

We thank Ms Abha Thite, Ms Meghana Walke and Ms Shraddha Hatangadi for excellent technical assistance, Mr Rahul Patil for statistical analysis, Mr Rahul Kadu and Dinesh Nande for BSL-3 laboratory support. Timely help from the staff of Microbiology Department, Bharati Vidyapeeth (Deemed to be University) Medical College and Hospital is greatly acknowledged.

Author's contributions

VAA: Concept, study design and supervision; SS, STM, RGV, and SSP performed the experiments, data analysis and interpretation; MM: patient's recruitment and clinical management; SS and VAA: writing of the manuscript; ACM: review the results and manuscript. All authors read and approved the final manuscript.

Institutional Review Board Statement

This study was approved by the "human ethics committee" of Bharati Vidyapeeth Deemed to be University Medical College (BVDUMC), Pune. The study was conducted according to the guidelines of the Declaration of Helsinki.

Data Availability

All data generated or analyzed during this study are included in this article. Full list of sequence identifiers is clearly mentioned in the text.

Conflicts of interests

None of the authors have any conflict of interest.

References

1. Geneva: World Health Organization. WHO coronavirus disease (COVID-19) dashboard (2020) <https://covid19.who.int/>. Accessed 23 Apr 2021
2. World Health Organization. Status of COVID-19 vaccines within WHO EUL/PQ evaluation process (2020) https://extranet.who.int/pqweb/sites/default/files/documents/Status_COVID_VAX_16Feb2021.pdf. Accessed 23 Apr 2021
3. Andrews MA et al (2020) First confirmed case of COVID-19 infection in India: A case report. *Indian J Med Res* 151(5):490–492
4. <https://> Accessed on 23 Apr 2021
5. Leung K, Shum MHH, Leung GM, Lam TT, Wu JT (2021) Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020. *Euro Surveill* 26(1):2002106
6. Volz E et al (2020) Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data. *medRxiv*. <https://doi.org/10.1101/2020.12.30.20249034>
7. Tegally H et al (2020) Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. *medRxiv* <https://doi.org/10.1101/2020.12.21.20248640>
8. Sabino EC et al (2021) Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* 397(10273):452–455
9. Claro IM et al (2020) Local transmission of SARS-CoV-2 lineage B.1.1.7, Brazil, December 2020. *Emerg Infect Dis* 27(3):970–972
10. Zhang W, Davis BD, Chen SS, Martinez JMC, Plummer JT, Vail E (2021) Emergence of a Novel SARS-CoV-2 Variant in Southern California. *JAMA* 325(13):1324–1326
11. Tchesnokova V et al (2021) Acquisition of the L452R mutation in the ACE2-binding interface of Spike protein triggers recent massive expansion of SARS-CoV-2 variants. *bioRxiv* <https://doi.org/10.1101/2021.02.22.432189>
12. Cherian S et al (2021) Convergent evolution of SARS-CoV-2 spike mutations, L452R, E484Q and P681R in the second wave of COVID-19 in Maharashtra, India. *bioRxiv* <https://doi.org/10.1101/2021.04.22.440932>
13. <http://www.cessi.in/coronavirus/pune>. Accessed on 09 August 2021
14. Hoffmann M, Kleine-Weber H, Pohlmann S (2020) A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol Cell* 78(4):779–784.e5
15. Plante JA et al (2021) Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* 592(7852):116–121
16. Hou YJ et al (2020) SARS-CoV-2 D614G Variant Exhibits Enhanced Replication ex vivo and Earlier Transmission in vivo. *bioRxiv* <https://doi.org/10.1101/2020.09.28.317685>
17. Boehm E et al (2021) Novel SARS-CoV-2 variants: the pandemics within the pandemic. *Clin Microbiol Infect* 27(8):1109–1117
18. Vaidyanathan G (2021) Coronavirus variants are spreading in India - what scientists know so far. *Nature* 593(7859):321–322
19. World Health Organization. Weekly epidemiological update on COVID-19-11 May 2021. <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19-11-may-2021>
20. Long LS et al (2021) Sequence Analysis of 20,453 Severe Acute Respiratory Syndrome Coronavirus 2 Genomes from the Houston Metropolitan Area Identifies the Emergence and Widespread Distribution of Multiple Isolates of All Major Variants of Concern. *Am J Pathol* S0002-9440(21):00108–00105
21. Bolze A et al (2021) Rapid displacement of SARS-CoV-2 variant B.1.1.7 by B.1.617.2 and P.1 in the United States. *medRxiv* <https://doi.org/10.1101/2021.06.20.21259195>
22. Adrian O'Dowd (2021) Covid-19: Cases of delta variant rise by 79%, but rate of growth slows. *BMJ* 373:n1596
23. SARS-CoV-2 variants of concern and variants under investigation (publishing.service.gov.uk). Technical briefing 16
24. Deng X et al (2021) Transmission, infectivity, and antibody neutralization of an emerging SARS-CoV-2 variant in California carrying a L452R spike protein mutation. *medRxiv* <https://doi.org/10.1101/2021.03.07.21252647>
25. Greaney AJ et al (2021) Comprehensive mapping of mutations to the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human serum antibodies. *bioRxiv* doi. <https://doi.org/10.1101/2020.12.31.425021>
26. Liu Z et al (2021) Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe* 29(3):477–488.e4
27. Zahradnik J et al (2021) SARS-CoV-2 RBD in vitro evolution follows contagious mutation spread, yet generates an able infection inhibitor. *bioRxiv* <https://doi.org/10.1101/2021.01.06.425392>
28. Lasek-Nesselquist E, Pata J, Schneider E, St. George K (2021) A tale of three SARS-CoV-2 variants with independently acquired P681H mutations in New York State. *medRxiv* doi: <https://doi.org/10.1101/2021.03.10.21253285>
29. Mlcochova P et al. SARS-CoV-2 B.1.617.2 Delta variant emergence and vaccine breakthrough. *Research Square*
30. Grabowski F, Preibisch P, Giziński S, Kocharczyk M, Lipniacki T (2021) SARS-CoV-2 Variant of Concern 202012/01 Has about Twofold Replicative Advantage and Acquires Concerning Mutations. *Viruses* 13(3):392
31. Jolly B et al (2020) Genetic epidemiology of variants associated with immune escape from global SARS-CoV-2 genomes. *bioRxiv* <https://doi.org/10.1101/2020.12.24.424332>

32. Gupta V et al (2020) Asymptomatic reinfection in two healthcare workers from India with genetically distinct SARS-CoV-2. Clin Infect Dis Sep 23:ciaa1451

33. Thomson EC et al (2020) The circulating SARS-CoV-2 spike variant N439K maintains fitness while evading antibody-mediated immunity. bioRxiv <https://doi.org/10.1101/2020.11.04.355842>

34. Weisblum Y et al (2020) Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. Elife 9:e61312. doi:10.7554/eLife.61312

35. Peacock TP et al (2020) The furin cleavage site of SARS-CoV-2 spike protein is a key determinant for transmission due to enhanced replication in airway cells. bioRxiv. doi:<https://doi.org/10.1101/2020.09.30.318311>

Figures

Figure 1

(A). Number of confirmed COVID-19 cases in Pune city from March 2020 to June 2021 and (B) Number of NPS specimens from suspected COVID-19 cases tested at the tertiary care hospital and numbers scored positive for SARS-CoV-2 RNA by RT-PCR (June 2020 to June 2021).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1Variantmanuscript.docx](#)