

# SARS-CoV-2 Associated Mutations, Strain Variations and Mortality During the First Epidemic Wave in India: A Study.

Neha Sharma

University Institute of Biotechnology, Chandigarh University, Gharuan, Mohali, Punjab, 140413.

\* Corresponding author : n.sharma137@gmail.com

## Abstract

India is no different to global health havoc created by SARS-CoV-2 outbreak, although the appearance of different waves and the magnitude of losses varied. The study of the first wave on timeline is of utmost importance due to the alarming emergency and lack of preparedness to deal with the new virus which entered the gene pool. Besides, population density and RNA virus mutation rate might have led to more strain variants. In the present study, comparative analysis of thirty-nine complete or nearly-complete genomes of SARS-CoV-2 during the first wave of COVID-19 in India was made and a number of mutations were observed. These observations further supported the evidence for genetic basis of rapid evolution of the virus.

**Keywords:** Coronavirus, First wave, India, Mutation, SARS-CoV-2, Variants.

The magnitude of toll on the health and lives of millions of people across the globe by RNA coronavirus SARS-CoV2 is massive and distressing<sup>1,2</sup>. The RNA viruses mutate at much higher rates (of the order of  $10^{-6}$  to  $10^{-4}$  substitutions per nucleotide per infection cell) than their host<sup>3,4</sup>. Since the case of its first incidence to notice in Wuhan, the capital of Hubei Province in the People's Republic of China and thereafter, the inexplicable mortality and fleeting

rate of infection intrigued the scientists to draw the correlation between the viral genomic variations and its spread<sup>5,6</sup>. Consequently, it led to a marathon of scientific efforts in the direction to deduce its structure and find vaccine.

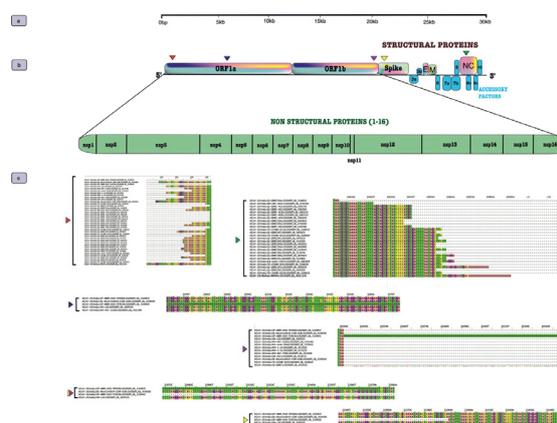


Figure 1: Polycistronic genome of SARS-CoV-2 (a) SARS-CoV-2 genome representing individual ORFs. (b) 16 non-structural proteins (NSPs) are embedded in polyprotein 1ab (PP1ab) while structural proteins include S (spike), E (envelope), M (membrane), NC (nucleocapsid) (c) pairwise nucleotide sequence alignment of the strains showing mutations in NSPS, S and NC parts of the SARS-CoV-2 genome. Coloured arrows in part b and c depict mutations and are correlated with respect to position and strains, respectively.

Table 1: Mutations found in the genome of SARS-CoV-2 strains isolated from thirty-nine patients during the first wave of COVID-19 in India. The number in the parenthesis indicated the location of amino acid in its protein.

Month	STRAIN			PATIENT DETAILS					MUTATIONS		
	Number	Lineage (Clade)	Location, India	Gender	Travel History	Age	Status	Sample collection date	Nucleotide*	Ge-nomic region*	Missense*
January	1.	hCoV-19/India/MH-1-27/2020	Maharashtra	Female	Wuhan, China	20	Recovered	2020-01-27	Gap of 3 nucleotides	NSP2 NSP3 NSP12 Spike	I(491)T P(1326)S A(406)V Y(145)del, R(408)I
	2.	hCoV-19/India/MH-1-31/2020	Maharashtra	Male	China	23	Recovered	2020-01-31	-	NSP2 NSP3 NSP13 Spike NS8	I(296)V P(1261)L T(214)I A(930)V L(84)S
March	1.	hCoV-19/India/UN-c32/2020	Unknown	Female	Italy	37	Unknown	2020-03-03	-	NSP12 Spike N	P(323)L D(614)G R(203)K, G(204)R
	2.	hCoV-19/India/UN-1621/2020	Unknown	Female	Indian citizen sampled at Iran	55	Unknown	2020-03-12	-	NSP2 NSP4 NSP6 Spike	R(27)C, V(198)I M(33)I L(37)F V(622)I

1.	hCoV-19/India/ GJ-GBRC5/2020	B.1 (GH)	Ahmed- abad	Female	-	66	Re- leased, Alive	2020- 04-13	-	NSP2 NSP3 NSP12 NSP13 Spike NS3 NS7a NS8 N	T(85)I, I(295) X, I(296) X, L(297) X, A(298) X, S(299) X, F(300) X, S(301)X S(371) T, C(417) X, V(418) X, G(1217) E, V(1434) X, L(1435) X, M(1436) X, S(1437) X, N(1438) X, L(1439) X, G(1440) X, M(1441) X, P(1442) X, S(1443)X P(323)I P(504)I D(614)G Q(57)H I(107)X, V(108) X, F(109) X, I(110) X, T(111) X, L(112) X, C(113) X E(110)stop S(33)I
2.	hCoV-19/ India/MH- NIV-11683/2020	B.1 (G)	Maharash- tra	Female	-	12	Mild	2020- 04-19	-	NSP5 NSP12 NSP16 Spike	A(116)V P(323)I G(39)C T(22)I, D(614) G
3.	hCoV-19/	B.6 (O)	Maharash-	Male	-	25	Mild	2020-	-	NSP3	S(1197)



3.	hCoV-19/India/KA-nimh-15835/2020	B.6 (O)	Karnataka	Female	-	15	Asymptomatic	2020-05-10	-	NSP3 NSP6 NSP12 Spike NS3 N	T(1198)K L(37)F A(97)V K(77)M, D(1084) Y L(52)I P(13)L
4.	hCoV-19/India/UP-NBRI-N38-TDT660/2020	B.1.1.32 (GR)	Uttar Pradesh	Female	-	86	Suspected coronavirus infection	2020-05-14	-	NSP2 NSP3 NSP12 Spike N	E(66)X, A(192)V S(848)X, I(849) X, K(850) X, W(851) X, A(852) X, D(853) X, N(854) X, N(855) X, C(856) X, A(994)D P(323)I D(614)G R(203)K, G(204) R

However, the situation was more tensed in India during the first wave of its outbreak due to lack of unpreparedness in all facets of medical and health care artillery<sup>7</sup>. The population density in India is multiple times higher than those of developed countries making the precautionary step of social distancing more challenging to follow and implement. Moreover, significance of face masks was more taxing with the illiterate section of the society mostly the daily wage workers<sup>8,9</sup>. These factors led to the multitude of strain mutations of the RNA virus and consequentially challenged the available medical facilities, which led to heavy loss of lives. To work strategically in the direction to find a vaccine amid this health chaos, it was imperative to deduce the structure of the RNA virus and the strain mutations accompanying its infection rate. The marathon began in this direction and a number of authorized laboratories actively participated over the country<sup>10</sup>.

In the present study an effort to evaluate the correlation of virus evolution and infection rate, SARS-CoV-2 genomes (complete or nearly complete) isolated from thirty-nine patients infected at different time period during the first wave of COVID-19 in India, were collected from Global Initiative on Sharing All Influenza Data (GISAID) [<https://www.gisaid.org/>]. These sequences were analysed on the basis of different parameters viz., detail(s) of patient and type(s) of mutation (Table 1). These sequences were also multiple aligned using the reference strain, hCoV-19/Wuhan/WIV04/2019 (Accession number: EPI\_ISL\_402124) and were found with missense mutations, nucleotide gaps, intermittent stop codons, long stretch of N nucleotides and many other alterations (Figure 1c). Most of the mutations were found in the non-structural proteins (NSPs), Spike (S) polyproteins and nucleocapsid (NC) proteins (Table 1; Figure 1b,c). The known role of position equivalent to most of these mutation(s) are antibody recognition sites, ligand binding and viral oligomerization interfaces<sup>11,12</sup>. To date, there is lack of information on the viral protein

However, the situation was more tensed in India during the first wave of its outbreak due to lack of unpreparedness in all facets of medical and health care artillery<sup>7</sup>. The population density in India is multiple times higher than those of developed countries making the precautionary step of social distancing more challenging to follow and implement. Moreover, significance of face masks was more taxing with the illiterate section of the society mostly the daily wage workers<sup>8,9</sup>. These factors led to the multitude of strain mutations of the RNA virus and consequentially challenged the available medical facilities, which led to heavy loss of lives. To work strategically in the direction to find a vaccine amid this health chaos, it was imperative to deduce the structure of the RNA virus and the strain mutations accompanying its infection rate. The marathon began in this direction and a number of authorized laboratories actively participated over the country<sup>10</sup>.

In the present study an effort to evaluate the correlation of virus evolution and infection rate, SARS-CoV-2 genomes (complete or nearly complete) isolated from thirty-nine patients infected at different time period during the first wave of COVID-19 in India, were collected from Global Initiative on Sharing All Influenza Data (GISAID) [<https://www.gisaid.org/>]. These sequences were analysed on the basis of different parameters viz., detail(s) of patient and type(s) of mutation (Table 1). These sequences were also multiple aligned using the reference strain, hCoV-19/Wuhan/WIV04/2019 (Accession number: EPI\_ISL\_402124) and were found with missense mutations, nucleotide gaps, intermittent stop codons, long stretch of N nucleotides and many other alterations (Figure 1c). Most of the mutations were found in the non-structural proteins (NSPs), Spike (S) polyproteins and nucleocapsid (NC) proteins (Table 1; Figure 1b,c). The known role of position equivalent to most of these mutation(s) are antibody recognition sites, ligand binding and viral oligomerization interfaces<sup>11,12</sup>. To date, there is lack of information on the viral protein

conformational changes and localization of amino acids involved. Such studies will be very significant and will play a key role in undermining the life threatening ability of the virus during the preceding waves of its spread in India.

#### Conflicts of interest

The authors declare no conflicts of interest/competing interests.

#### Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

#### Acknowledgements

The author gratefully acknowledge the Chandigarh University for providing facilities. The author is also thankful to the anonymous reviewers for the valuable suggestions.

#### References

1. Phan, T., Genetic diversity and evolution of SARS-CoV-2. *Infect. Genet. Evol.*, 2020, 81, 104260.
2. Zheng, J., SARS-CoV-2: an emerging coronavirus that causes a global threat. *Int. J. Biol. Sci.*, 2020, 16(10), 1678.
3. Sanjuán, R. and Domingo-Calap, P., Mechanisms of viral mutation. *Cell. Mol. Life. Sci.*, 2016, 73(23), 4433-4448.
4. Peck, K. M. and Luring, A. S., Complexities of viral mutation rates. *J. Virol.*, 2018, 92(14), e01031-17.
5. Luring, A. S. and Andino, R., (2010) Quasispecies theory and the behavior of RNA viruses. *PLoS. Pathog.*, 2010, 6(7), e1001005.
6. Sironi, M., Hasnain, S. E., Phan, T., Luciani, F., Shaw, M. A., Sallum, M. A., Mirhashemi, M. E., Morand, S. and González-Candelas,

- F., SARS-CoV-2 and COVID-19: A genetic, epidemiological, and evolutionary perspective. *Infect. Genet. Evol.*, 2020, 84, 104384.
7. Lajawala, R., Sonavane, S. and De Sousa, A., Various Facets of Anxiety during the Pandemic in India-Critical Issues, 2021, [Online First], IntechOpen, DOI: 10.5772/intechopen.97797. Available from: <https://www.intechopen.com/online-first/various-facets-of-anxiety-during-the-pandemic-in-india-critical-issues>.
  8. Desai, D., Urban densities and the Covid-19 pandemic: Upending the sustainability myth of global megacities. *ORF. Occas. Pap.*, 2020, 244(4), 1-4.
  9. Ghosh, A., Nundy, S. and Mallick, T. K., How India is dealing with COVID-19 pandemic. *Sens. Int.*, 2020, 1, 100021.
  10. Alai, S., Gujar, N., Joshi, M., Gautam, M. and Gairola, S., Pan-India novel coronavirus SARS-CoV-2 genomics and global diversity analysis in spike protein. *Heliyon*, 2021, 7(3), e06564.
  11. Fung, T.S. and Liu, D.X., Human coronavirus: host-pathogen interaction. *Ann. Rev. Microbiol.*, 2019, 73, 529-557.
  12. Yu, F., Du, L., Ojcius, D. M., Pan, C. and Jiang, S., Measures for diagnosing and treating infections by a novel coronavirus responsible for a pneumonia outbreak originating in Wuhan, China. *Microbes. Infect.*, 2020, 22(2), 74-79.