

Manuscript Title: Phyloepidemiology and adaptive evolution of SARS-CoV2 during the first and second wave of COVID-19 in India

The name(s) of the author(s): Anuj Tewari, Rajesh Kumar

The affiliation(s) and address(es) of the author(s): Department of Veterinary Microbiology, College of Veterinary & Animal Sciences, Govind Ballabh Pant University of Agriculture & Technology, Pantnagar, Uttarakhand, India.

*Corresponding author email: anuj474@gmail.com

Figure S1

Figures showing amino acid substitution in NSP2 & NSP3 of SARS-CoV-2 under positive selective pressure

Figure S2

Figures showing amino acid substitution in NSP13 & S protein of SARS-CoV-2 under positive selective pressure

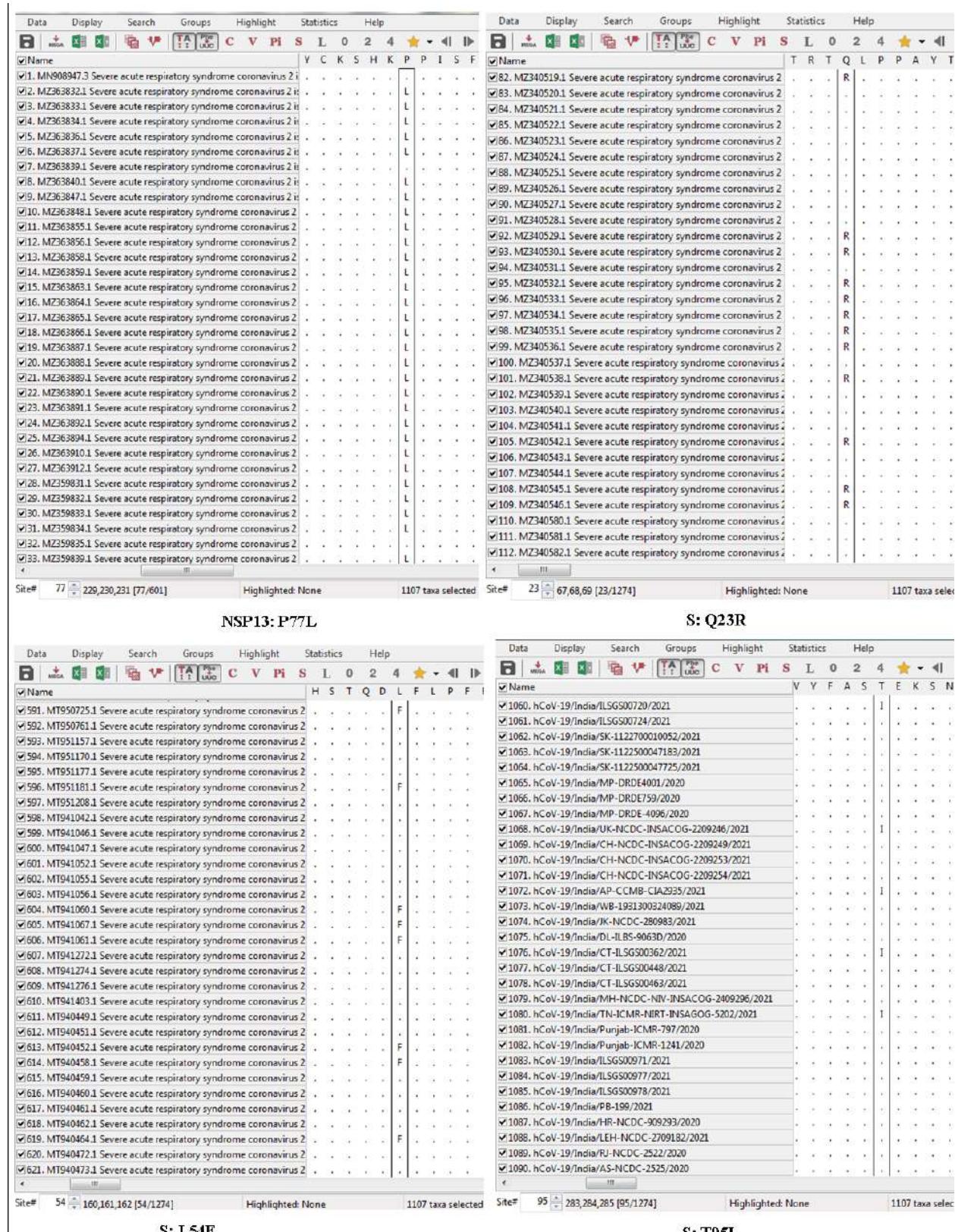


Figure S3

Figures showing amino acid substitution in S protein & ORF3a of SARS-CoV-2 under positive selective

S: G142D, G142S		S: E484D, E484Q, E484K	
Data	Display	Data	Display
Search	Groups	Search	Groups
Highlight	Statistics	Highlight	Statistics
Help		Help	
<input checked="" type="checkbox"/> Name	D P F L G V Y H K N	<input checked="" type="checkbox"/> Name	C N G V E G F N C
✓1070. hCoV-19/India/CH-NCDC-INSACOG-2209253/2021	- - - - -	✓974. hCoV-19/India/PY-SEQ-181 S38 R1 001/2021	- - - - -
✓1071. hCoV-19/India/CH-NCDC-INSACOG-2209254/2021	- - - - -	✓975. hCoV-19/India/PY-SEQ-183 S40 R1 001/2021	- - - - -
✓1072. hCoV-19/India/AP-CCMB-CIA2935/2021	- - - - -	✓976. hCoV-19/India/OR-ILSG501022/2021	- - - - -
✓1073. hCoV-19/India/WB-1931300324089/2021	- - - - -	✓977. hCoV-19/India/CT-ILSG501037/2021	- - - - -
✓1074. hCoV-19/India/JK-NCDC-28983/2021	- - - - -	✓978. hCoV-19/India/HR-THT1637/2021	- - - - -
✓1075. hCoV-19/India/DL-ILBS-9063D/2020	- - - - -	✓979. hCoV-19/India/un-SRI-HER-Chenna/2021	- - - - -
✓1076. hCoV-19/India/CT-ILSG500362/2021	- - - - -	✓980. hCoV-19/India/DL-NCDC-12099/2021	- - - - -
✓1077. hCoV-19/India/CT-ILSG500448/2021	- - - - -	✓981. hCoV-19/India/PB-IMT-P131/2020	- - - - -
✓1078. hCoV-19/India/CT-ILSG500463/2021	- - - - -	✓982. hCoV-19/India/MH-NIV-227244/2020	- - - - -
✓1079. hCoV-19/India/MH-NCDC-NIV-INSACOG-2409296/2021	- - - - -	✓983. hCoV-19/India/OR-ILSG500137/2020	- - - - -
✓1080. hCoV-19/India/TN-ICMR-NIRT-INSACOG-G-5202/2021	- - - - -	✓984. hCoV-19/India/KA-NIMH-SEQ-217/2021	- - - - -
✓1081. hCoV-19/India/Punjab-ICMR-797/2020	- - - - -	✓985. hCoV-19/India/KA-NIMH-SEQ-228/2021	- - - - -
✓1082. hCoV-19/India/Punjab-ICMR-1241/2020	- - - - -	✓986. hCoV-19/India/MH-NCCS-87405/2021	- - - - -
✓1083. hCoV-19/India/ILSG500971/2021	- - - - -	✓987. hCoV-19/India/G-ICMR-NIV-INSACOG-GSEQ-565/2020	- - - - -
✓1084. hCoV-19/India/ILSG500977/2021	- - - - -	✓988. hCoV-19/India/TN-SEQ152 S18 R1 001/2021	- - - - -
✓1085. hCoV-19/India/ILSG500978/2021	- - - - -	✓989. hCoV-19/India/TN-SEQ154 S20 R1 001/2021	- - - - -
✓1086. hCoV-19/India/PB-199/2021	- - - - -	✓990. hCoV-19/India/HR-NCDC-INSACOG-1309166/2021	- - - - -
✓1087. hCoV-19/India/HR-NCDC-909293/2020	- - - - -	✓991. hCoV-19/India/HR-NCDC-INSACOG-1309175/2021	- - - - -
✓1088. hCoV-19/India/LEH-NCDC-2709182/2021	- - - - -	✓992. hCoV-19/India/HR-NCDC-INSACOG-1309180/2021	- - - - -
✓1089. hCoV-19/India/IL-NCDC-2522/2020	- - - - -	✓993. hCoV-19/India/JK-NCDC-INSACOG-1309124/2021	- - - - -
✓1090. hCoV-19/India/AS-NCDC-2525/2020	- - - - -	✓994. hCoV-19/India/JK-NCDC-INSACOG-1309128/2021	- - - - -
✓1091. hCoV-19/India/BR-ILSG500281/2021	- - - - -	✓995. hCoV-19/India/un-NCVTC-P50/2020	- - - - -
✓1092. hCoV-19/India/BR-ILSG500284/2021	- - - - -	✓996. hCoV-19/India/un-NCVTC-P70/2020	- - - - -
✓1093. hCoV-19/India/OR-ILSG500294/2021	- - - - -	✓997. hCoV-19/India/CT-ILSG501071/2021	- - - - -
✓1094. hCoV-19/India/BR-ILSG500298/2020	- - - - -	✓998. hCoV-19/India/KA-NIMH-SEQ-964/2021	- - - - -
✓1095. hCoV-19/India/OR-ILSG500333/2021	- - - - -	✓999. hCoV-19/India/TN-C134 S185 R1 001/2021	- - - - -
✓1096. hCoV-19/India/ILSG500928/2021	- - - - -	✓1000. hCoV-19/India/AP-CCMB-CIA2385/2021	- - - - -
✓1097. hCoV-19/India/ILSG500942/2021	- - - - -	✓1001. hCoV-19/India/AP-CCMB-CIA2383/2021	- - - - -
✓1098. hCoV-19/India/ILSG500948/2021	- - - - -	✓1002. hCoV-19/India/AP-CCMB-CIA2782/2021	- - - - -
✓1099. hCoV-19/India/ILSG500979/2021	Y	✓1003. hCoV-19/India/JK-NCDC-1109254/2020	- - - - -
✓1100. hCoV-19/India/ILSG500996/2021	- - - - -	✓1004. hCoV-19/India/AR-1223700015970/2021	- - - - -

S: P681H, P681R		ORF3a: S26L	
Data	Display	Data	Display
Search	Groups	Search	Groups
Highlight	Statistics	Highlight	Statistics
Help		Help	
<input checked="" type="checkbox"/> Name	T Q N T N S P R R A S	<input checked="" type="checkbox"/> M11: Sequence Data Explorer (ORF3a-final.fas)	K D A T P S D F V
✓971. hCoV-19/India/PY-SEQ165 S34 R1 001/2021	- - - - -	✓1. M2363832.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓972. hCoV-19/India/TN-SEQ-168 S31 R1 001/2021	- - - - -	✓2. M2363832.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓973. hCoV-19/India/TN-SEQ-179 S31 R1 001/2021	- - - - -	✓3. M2363833.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓974. hCoV-19/India/PY-SEQ-181 S38 R1 001/2021	- - - - -	✓4. M2363834.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓975. hCoV-19/India/PY-SEQ-183 S40 R1 001/2021	- - - - -	✓5. M2363836.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓976. hCoV-19/India/OR-ILSG501022/2021	- - - - -	✓6. M2363837.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓977. hCoV-19/India/CT-ILSG501037/2021	- - - - -	✓7. M2363839.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓978. hCoV-19/India/HR-THT1637/2021	- - - - -	✓8. M2363840.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓979. hCoV-19/India/un-SRI-HER-Chenna/2021	- - - - -	✓9. M2363847.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓980. hCoV-19/India/DL-NCDC-12099/2021	- - - - -	✓10. M2363848.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓981. hCoV-19/India/PB-IMT-P131/2020	- - - - -	✓11. M2363855.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓982. hCoV-19/India/MH-NIV-227244/2020	- - - - -	✓12. M2363856.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓983. hCoV-19/India/OR-ILSG500137/2020	- - - - -	✓13. M2363858.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓984. hCoV-19/India/KA-NIMH-SEQ-217/2021	- - - - -	✓14. M2363859.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓985. hCoV-19/India/KA-NIMH-SEQ-228/2021	- - - - -	✓15. M2363863.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓986. hCoV-19/India/MH-NCCS-87405/2021	- - - - -	✓16. M2363864.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓987. hCoV-19/India/G-ICMR-NIV-INSACOG-GSEQ-565/2020	- - - - -	✓17. M2363865.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓988. hCoV-19/India/TN-SEQ152 S18 R1 001/2021	- - - - -	✓18. M2363866.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓989. hCoV-19/India/TN-SEQ154 S20 R1 001/2021	- - - - -	✓19. M2363887.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓990. hCoV-19/India/HR-NCDC-INSACOG-1309166/2021	- - - - -	✓20. M2363888.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓991. hCoV-19/India/HR-NCDC-INSACOG-1309175/2021	- - - - -	✓21. M2363899.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓992. hCoV-19/India/HR-NCDC-INSACOG-1309180/2021	- - - - -	✓22. M2363890.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓993. hCoV-19/India/JK-NCDC-INSACOG-1309124/2021	- - - - -	✓23. M2363891.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓994. hCoV-19/India/JK-NCDC-INSACOG-1309128/2021	- - - - -	✓24. M2363892.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓995. hCoV-19/India/un-NCVTC-P50/2020	- - - - -	✓25. M2363894.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓996. hCoV-19/India/un-NCVTC-P70/2020	- - - - -	✓26. M2363910.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓997. hCoV-19/India/CT-ILSG501071/2021	- - - - -	✓27. M2363912.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓998. hCoV-19/India/KA-NIMH-SEQ-964/2021	- - - - -	✓28. M2359831.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓999. hCoV-19/India/TN-C134 S185 R1 001/2021	- - - - -	✓29. M2359832.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓1000. hCoV-19/India/AP-CCMB-CIA2385/2021	- - - - -	✓30. M2359833.1 Severe acute respiratory syndrome coronavirus 2	- - - - -
✓1001. hCoV-19/India/AP-CCMB-CIA2383/2021	- - - - -	✓31. M2359834.1 Severe acute respiratory syndrome coronavirus 2	- - - - -

pressure

Figure S4

Figures showing amino acid substitution in ORF9 region of SARS-CoV-2 under positive selective pressure

Data Display Search Groups Highlight Statistics Help											
C V P I S L O 2 4 ★ ◀ ▶											
Name Q R N A P R I T F G G											
Name R S R N S S R N S T I											
<input checked="" type="checkbox"/> 178. MZ336031.1 Severe acute respiratory syndrome coronavirus 2 L											
<input checked="" type="checkbox"/> 179. MZ336032.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 180. MZ336033.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 181. MZ325260.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 182. MZ325261.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 183. MZ325301.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 184. MZ325302.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 185. MZ325343.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 186. MZ325344.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 187. MZ325346.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 188. MZ325378.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 189. MZ325379.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 190. MZ325380.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 191. MZ325393.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 192. MZ325421.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 193. MZ325844.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 194. MZ325845.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 195. MZ317890.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 196. MZ317891.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 197. MZ317892.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 198. MZ317894.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 199. MZ317905.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 200. MZ317906.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 201. MZ317907.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 202. MZ317908.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 203. MZ317915.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 204. MZ317916.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 205. MZ317918.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 206. MZ317919.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 207. MZ317920.1 Severe acute respiratory syndrome coronavirus 2											
<input checked="" type="checkbox"/> 208. MZ317921.1 Severe acute respiratory syndrome coronavirus 2											
Site# 13	37,38,39 [13/419]	Highlighted: None	1107 taxa selected								
Site# 194	580,581,582 [194/419]	Highlighted: None	1107 taxa selected								

ORF9: P13L, P13S

ORF9: S194L

Figure S5

Phylogenetic analysis of SARS-CoV-2 genomes using iTOL (<https://itol.embl.de/>). The tree initially got divided into three clades.



Figure S6: Phylogenetic analysis of SARS-CoV-2 genome using MEGA 11. The phylogenetic clusters are unlabelled lines. Sequences which didn't cluster with any of the sequences have been labeled and shown in the Phylogenetic Tree

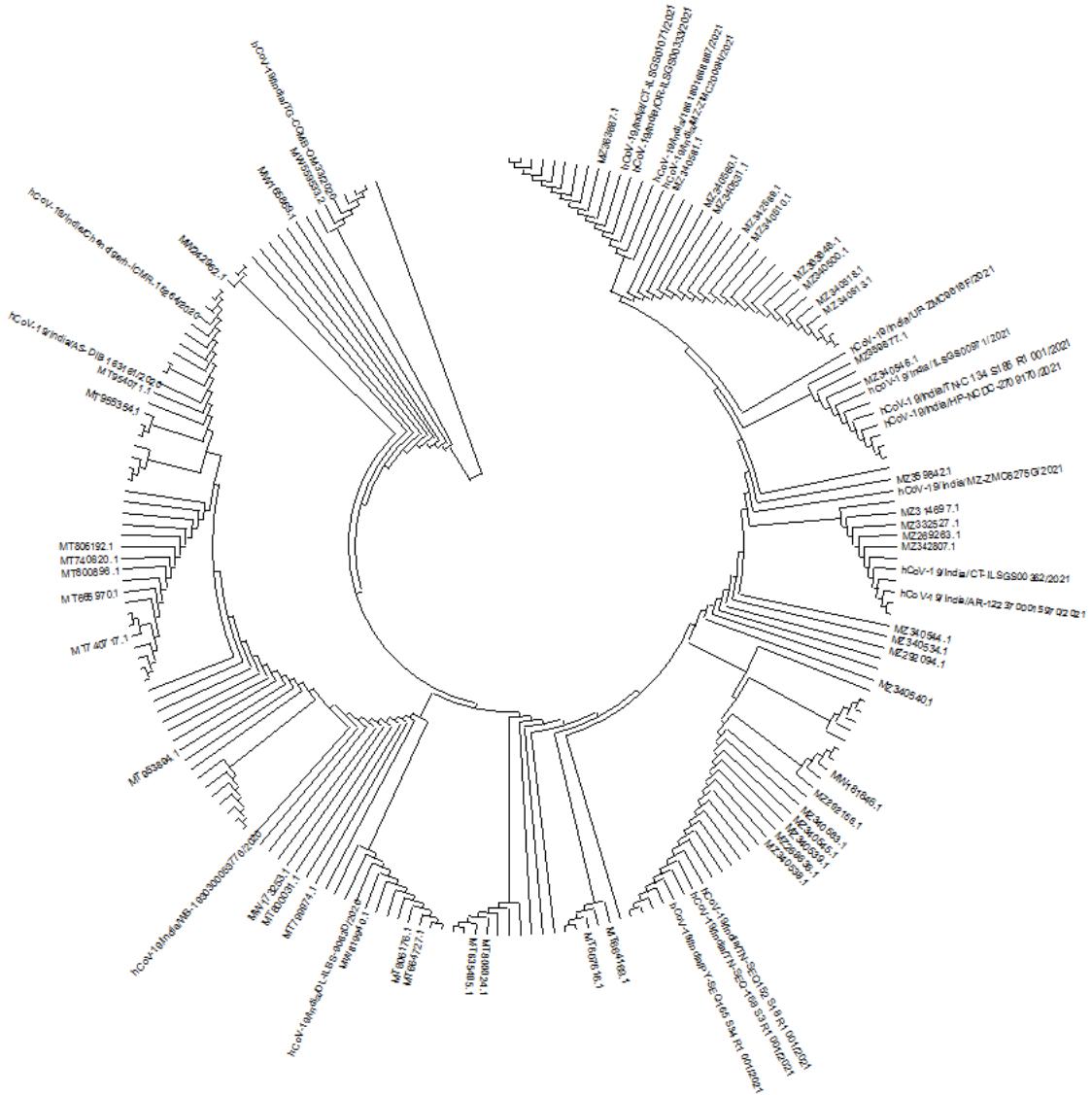


Table S1

The purifying and diversifying selection in the genome of SARS-CoV-2 as assessed by FEL method

Protein	Site showing purifying selection	Site showing diversifying selection
NSP1	16, 90, 159, 172	33
NSP2	36, 75, 94, 107, 190, 260, 335, 357, 384, 470, 475, 563, 595, 598, 614, 629	381, 488, 489
NSP3	21, 39, 68, 85, 106, 111, 236, 246, 255, 297, 327, 352, 512, 527, 550, 551, 586, 629, 660, 700, 701, 705, 744, 763, 819, 875, 878, 938, 940, 951, 955, 963, 976, 978, 982, 991, 1003, 1055, 1089, 1092, 1113, 1159, 1193, 1216, 1225, 1325, 1342, 1671, 1703, 1813, 1820, 1823, 1857, 1915	183, 822, 1682
NSP4	15, 67, 78, 121, 177, 201, 212, 240, 400	Nil
NSP5	14, 81, 137, 301	Nil
NSP6	16, 28, 119, 120, 214, 215, 279, 285	37
NSP7	15, 23, 32, 60	Nil
NSP8	14, 146	Nil
NSP9	11, 112	Nil
NSP10	130	Nil
NSP11	Nil	Nil
NSP12	92, 103, 144, 302, 513, 626, 753	513
NSP14	17, 143, 183, 280, 338, 424, 425, 459, 495, 514	Nil
NSP15	53, 85, 192, 214, 247, 335	Nil
NSP16	97, 112, 122, 141, 191, 199, 204	126, 298
ORF3a	36, 78, 129, 252	26, 57, 93, 155, 172, 223
M gene	71, 93, 135	82
ORF6	7, 61	Nil
ORF7a	33	81
ORF7b	10, 36	Nil
ORF8	14, 31, 48, 49, 72	Nil
N gene	35, 36, 39, 61, 110, 157, 173, 234, 268, 287, 298, 302, 320, 321, 337	13, 18, 63, 194, 214, 215, 235, 251, 379
ORF 10	28	Nil
Spike:	32, 47, 61, 66, 81, 87, 111, 134, 146, 151, 206, 213, 227, 237, 238, 256, 294, 302, 321, 341, 342, 355, 372, 380, 386, 394, 433, 475, 518, 534, 555, 629, 636, 642, 679, 682, 693, 707, 789, 808, 820, 827, 852, 860, 881, 888, 905, 907, 912, 942, 948, 971, 995, 996, 1045, 1053, 1060, 1091, 1110, 1128, 1148, 1178, 1216, 1238, 1254, 1273	5, 23, 54, 95, 142, 452, 501, 681, 879, 950, 1118

Table S2

The purifying and diversifying selection in the genome of SARS-CoV-2 as assessed by MEME method

Protein	Site showing Episodic Positive/diversifying selection
NSP2	381
NSP3	183, 882, 1682
NSP6	37
NSP 12 ploymerase	571
NSP 15	414
NSP 16	126, 298
ORF 3a	26, 57, 93, 155, 172, 223
M gene	82
Orf 7a	81-0.08, 1.82
Orf 8	73-0.01 0.719
Orf 9	3, 13, 63, 194, 204, 235, 251
Spike	23, 54, 95, 142, 152, 452, 501, 627, 681, 879, 943, 1118