

# Supplementary to “Characterisation of Clades of 18392 SARS-CoV-2 Genomes based on Signature SNPs unveils Continuous Evolution”

Nimisha Ghosh<sup>a,e</sup>, Indrajit Saha<sup>b,e,\*</sup>, Nikhil Sharma<sup>c,e</sup>, Suman Nandi<sup>d</sup>

<sup>a</sup>Department of Computer Science and Information Technology, Institute of Technical Education and Research, Siksha ‘O’ Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India

<sup>b</sup>Department of Computer Science and Engineering, National Institute of Technical Teachers’ Training and Research, Kolkata, West Bengal, India

<sup>c</sup>Department of Electronics and Communication Engineering, Jaypee Institute of Information Technology, Noida, Uttar Pradesh, India

<sup>d</sup>Department of Computer Science and Engineering, National Institute of Technical Teachers’ Training and Research, Kolkata, West Bengal, India

<sup>e</sup>Equally contributed

Table S1: Clade wise distribution of 15359 sequences for Global excluding India

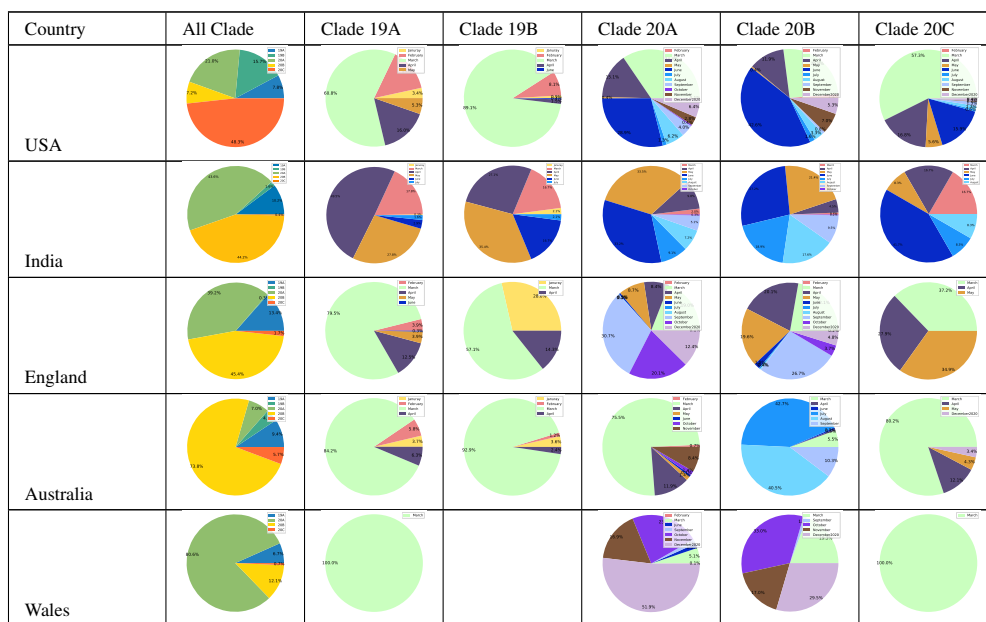
Country	19A	19B	20A	20B	20C	Country	19A	19B	20A	20B	20C
USA	263	530	707	244	1628	Thailand	2	16	1	3	4
England	336	7	981	1137	43	Northern Ireland	8	0	5	10	0
Australia	190	84	143	1499	116	Norway	8	0	6	1	5
Wales	97	0	1174	176	10	Austria	3	0	8	3	4
Scotland	109	14	514	212	22	Chile	0	8	5	1	1
Netherlands	141	8	241	126	41	Colombia	2	3	6	2	2
Belgium	100	1	185	226	27	Indonesia	12	0	3	0	0
China	394	90	23	12	13	Estonia	0	0	3	8	2
Iceland	95	16	152	91	66	Senegal	3	1	8	0	0
Portugal	33	11	128	189	8	Croatia	1	0	5	2	3
France	33	2	147	18	73	Georgia	4	1	4	1	1
Spain	17	129	97	19	4	Malaysia	10	1	0	0	0
New Zealand	43	10	63	77	56	Romania	0	0	5	6	0
Sweden	9	0	62	113	38	Ireland	3	0	2	5	0
Switzerland	19	0	71	48	25	Kenya	2	0	7	1	0
Italy	7	0	84	35	0	Latvia	5	0	2	2	0
Luxembourg	5	1	79	2	24	Nigeria	7	0	1	0	1
Denmark	2	0	32	9	66	Kuwait	5	0	1	1	0
Japan	71	4	7	24	0	Slovakia	1	0	4	1	0
Canada	7	25	31	16	23	Tunisia	0	0	5	1	0
Brazil	5	2	6	80	1	Bangladesh	1	0	0	3	0
Germany	23	2	5	14	25	Greece	0	1	0	3	0
Singapore	31	1	18	3	1	Qatar	4	0	0	0	0
Russia	0	0	9	41	2	Turkey	2	0	2	0	0
South Africa	1	0	11	5	34	Argentina	0	0	2	1	0
Kazakhstan	26	18	2	0	3	Belarus	2	0	1	0	0
Israel	1	0	8	31	0	Hungary	0	0	2	1	0
Poland	2	0	14	21	3	Saudi Arabia	1	0	2	0	0
Oman	16	0	6	16	1	Slovenia	2	0	1	0	0
Mexico	1	10	15	8	2	Pakistan	2	0	0	0	0
South Korea	17	19	0	0	0	Serbia	0	0	1	1	0
Peru	0	0	2	31	0	Cambodia	1	0	0	0	0
Czech Republic	0	0	9	20	3	Morocco	0	0	0	1	0
Vietnam	5	0	2	22	2	Nepal	1	0	0	0	0
Finland	3	0	13	4	6	Panama	0	0	1	0	0

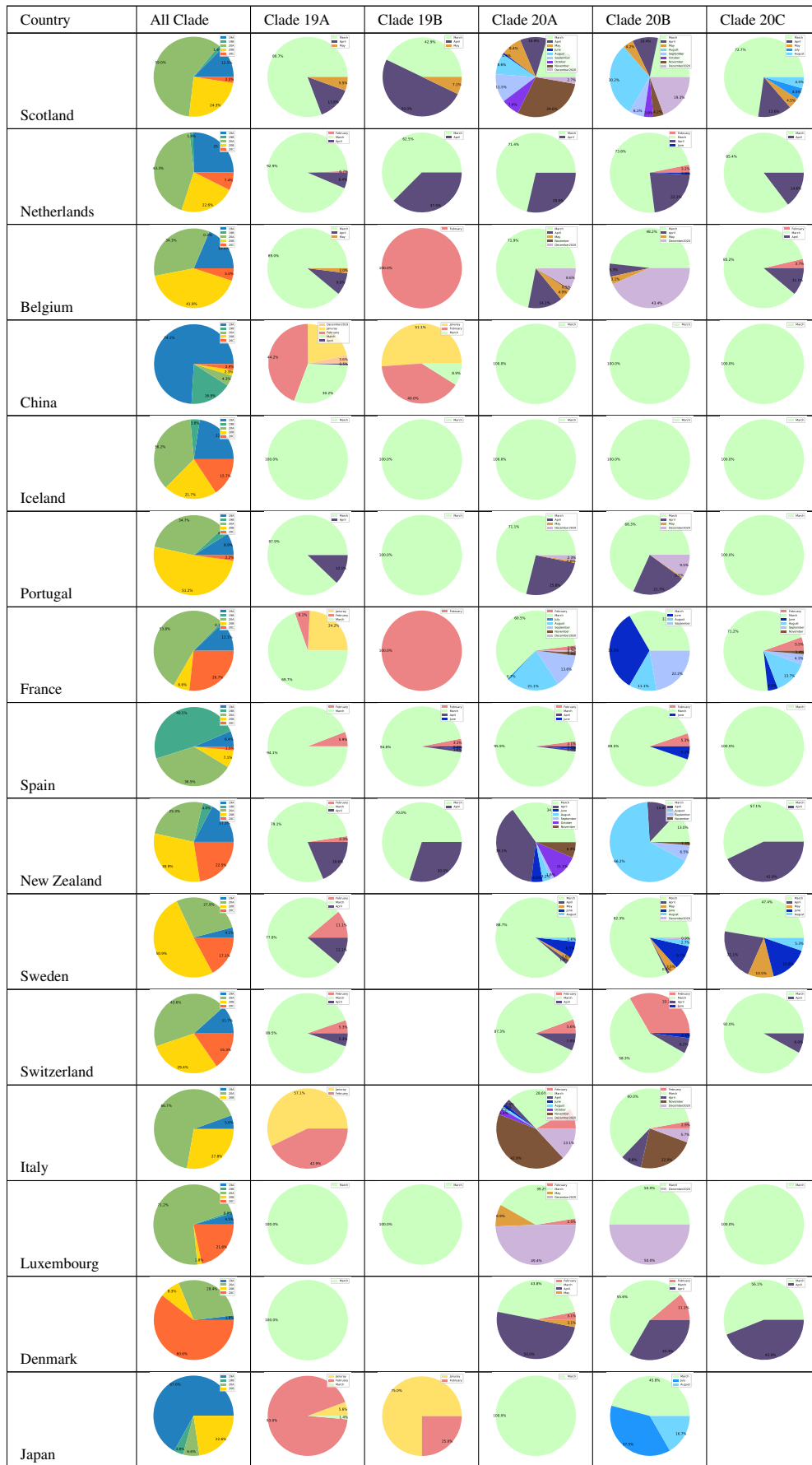
\*Corresponding author: indrajit@nittrkol.ac.in

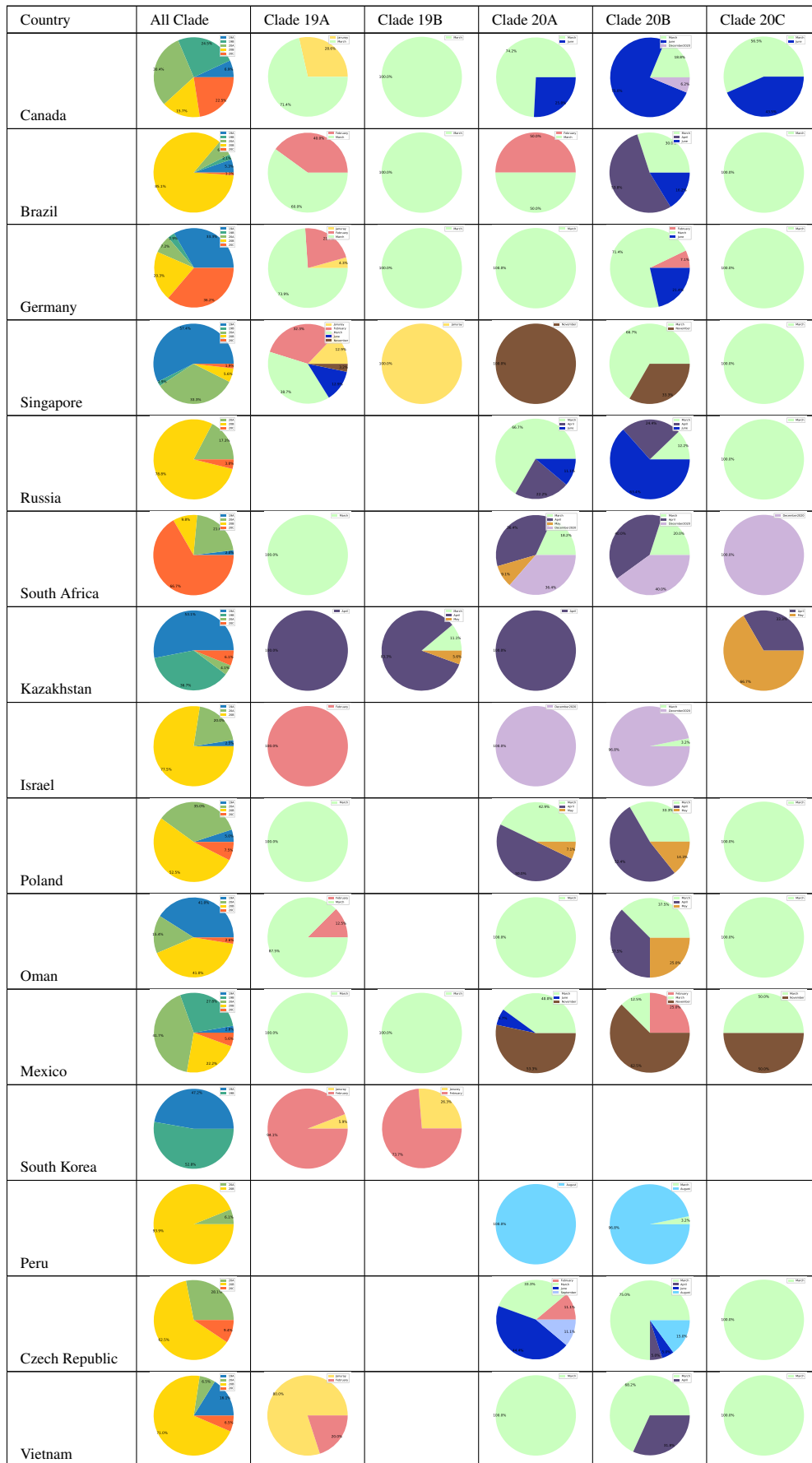
Table S2: Clade wise distribution of 3033 Indian sequences

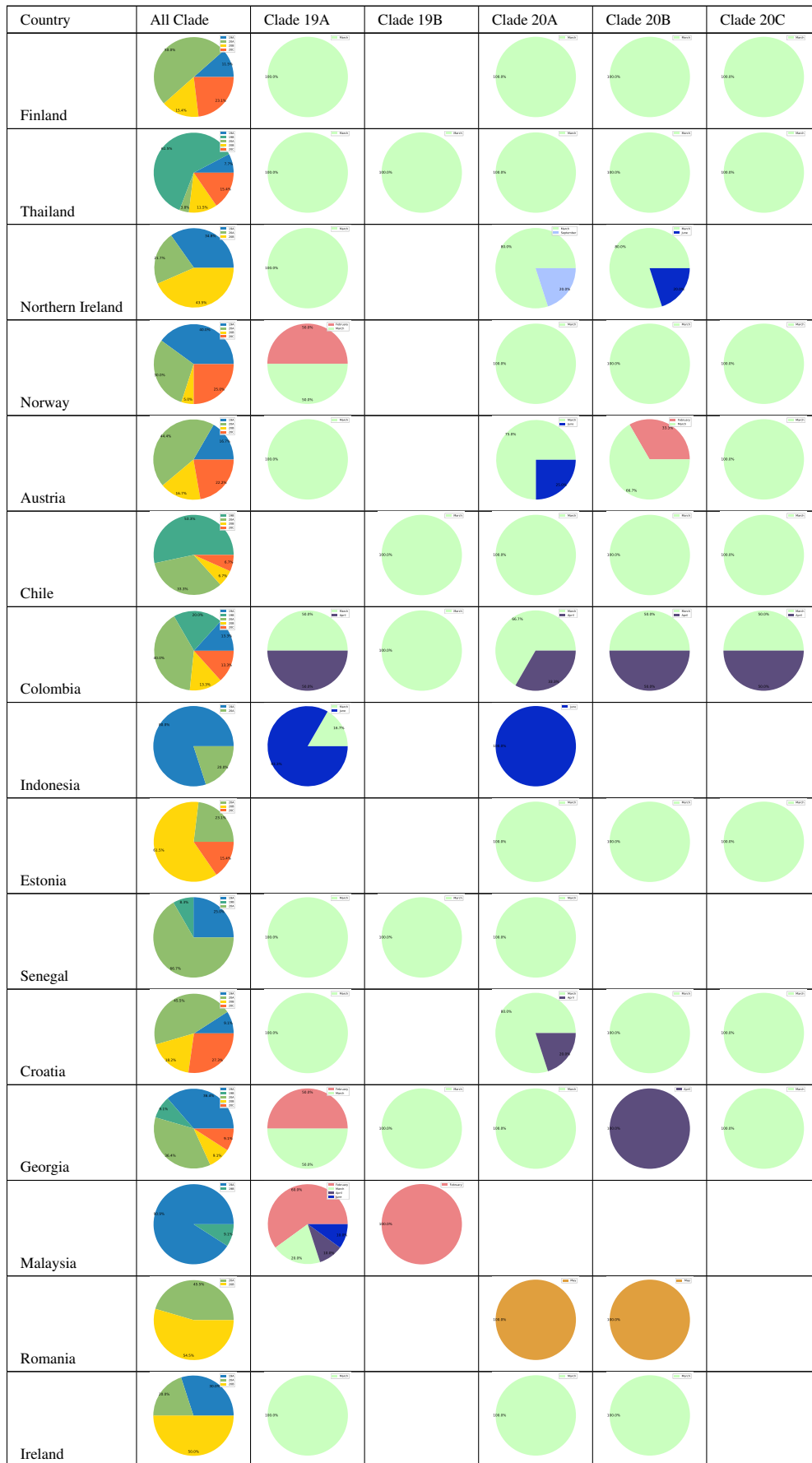
State	19A	19B	20A	20B	20C
Maharashtra	39	8	289	808	0
Gujarat	16	12	559	21	3
Telangana	94	0	59	311	2
West Bengal	9	14	154	15	0
Delhi	55	1	79	19	2
Karnataka	25	2	25	51	0
Odisha	6	10	28	45	4
Haryana	15	0	44	29	1
Uttarakhand	2	1	40	25	0
Madhya Pradesh	10	0	25	1	0
Tamil Nadu	15	0	1	15	0
Uttar Pradesh	4	0	16	1	0
Rajasthan	4	0	2	0	0
Punjab	4	0	1	0	0
Ladakh	5	0	0	0	0
Bihar	2	0	0	0	0
Assam	2	0	0	0	0
Andhra Pradesh	1	0	0	1	0
Jammu and Kashmir	1	0	0	0	0
Total	309	48	1322	1342	12

Table S3: Pie charts to represent clade wise evolution of 18392 SARS-CoV-2 genomes with respect to each month for 71 countries











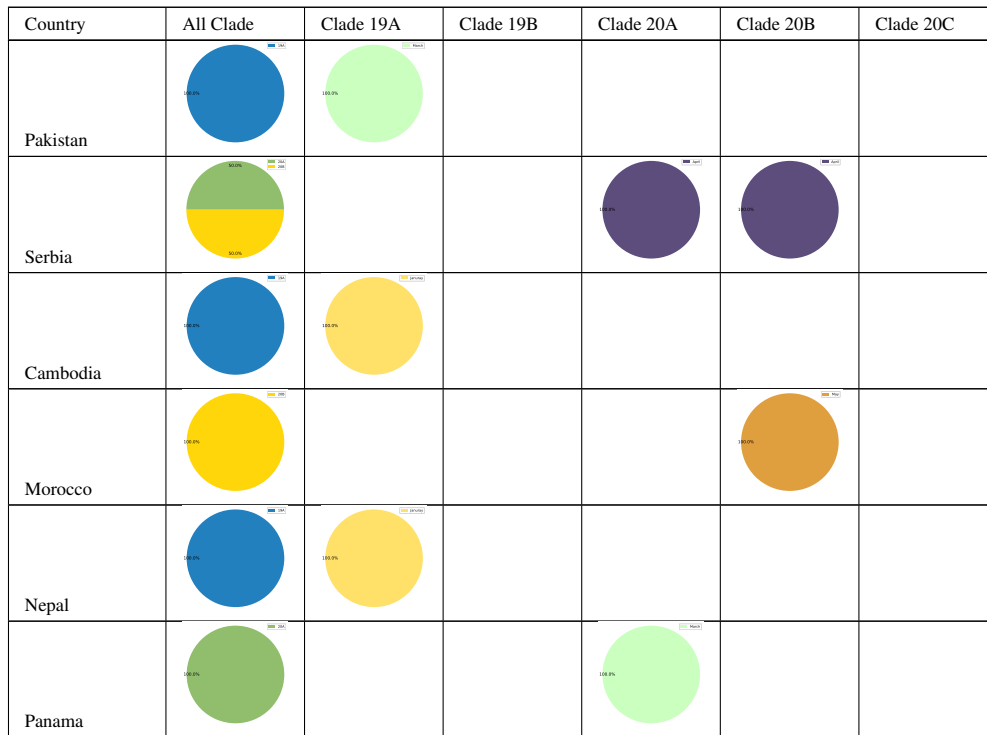


Table S4: Pie charts to represent clade wise evolution of 3033 Indian SARS-CoV-2 genomes with respect to each month for each state

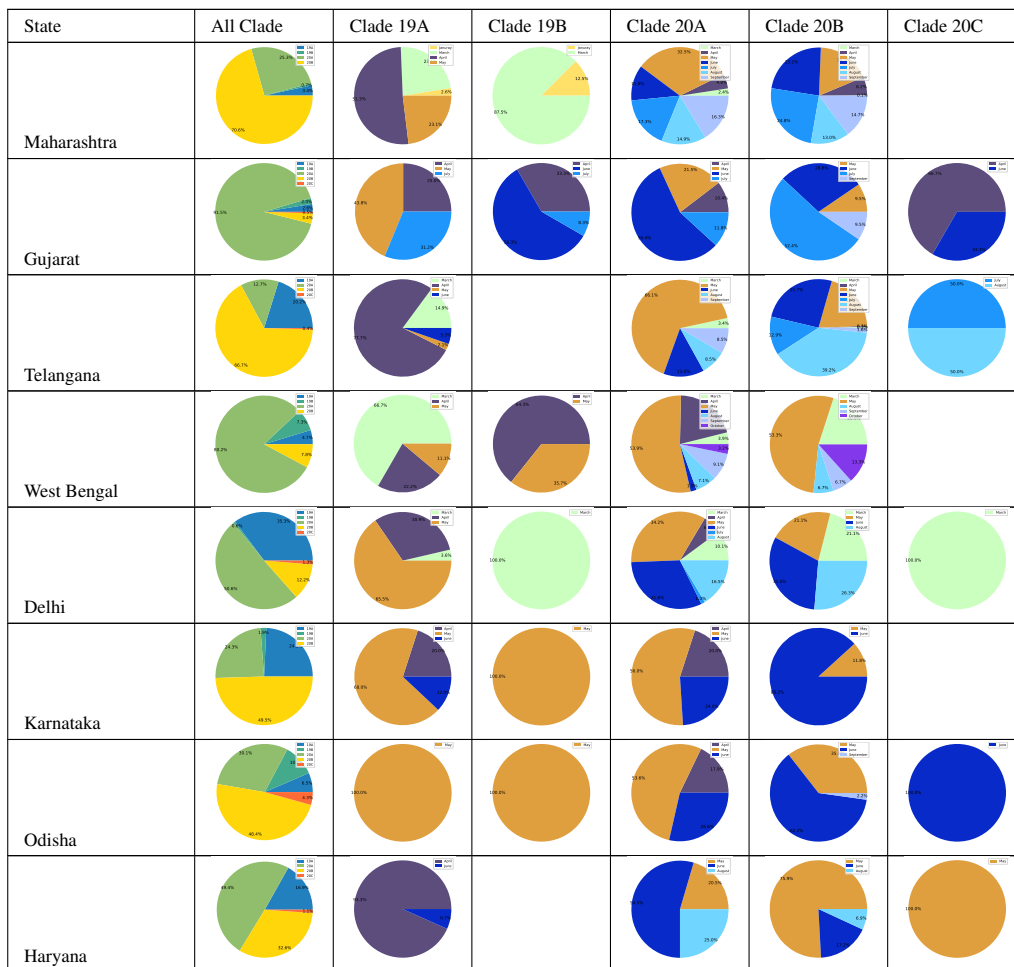
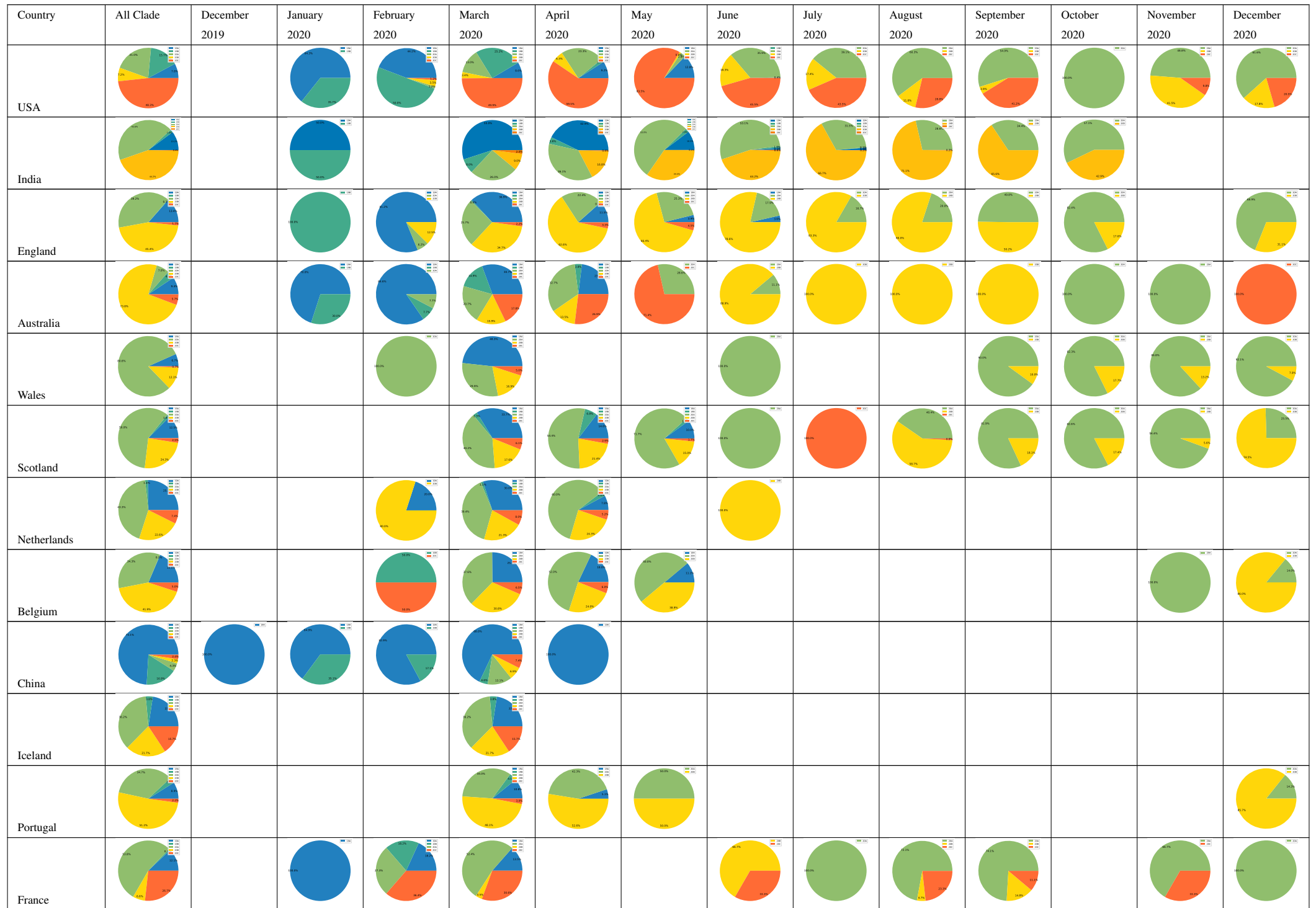






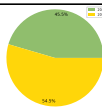
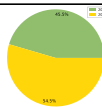
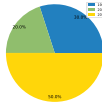
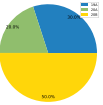
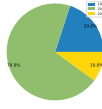



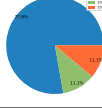
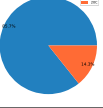
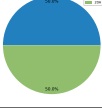

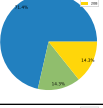
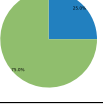
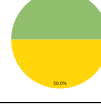
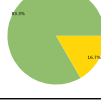

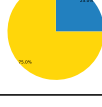
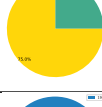
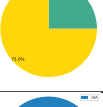
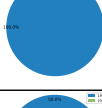
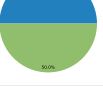
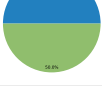
Table S5: Pie charts to represent month wise evolution of 18392 SARS-CoV-2 genomes with respect to each clade for 71 countries









Country	All Clade	December 2019	January 2020	February 2020	March 2020	April 2020	May 2020	June 2020	July 2020	August 2020	September 2020	October 2020	November 2020	December 2020
Romania														
Ireland														
Kenya														
Latvia														
Nigeria														
Kuwait														
Slovakia														
Tunisia														
Bangladesh														
Greece														
Qatar														
Turkey														

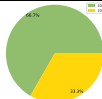
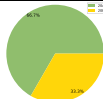
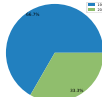
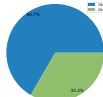
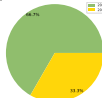
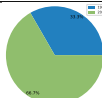
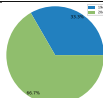
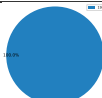
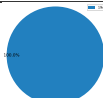
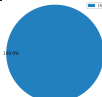
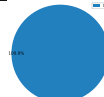
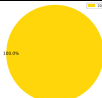
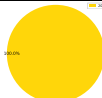
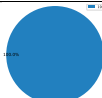
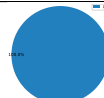
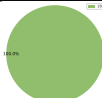
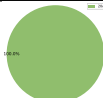
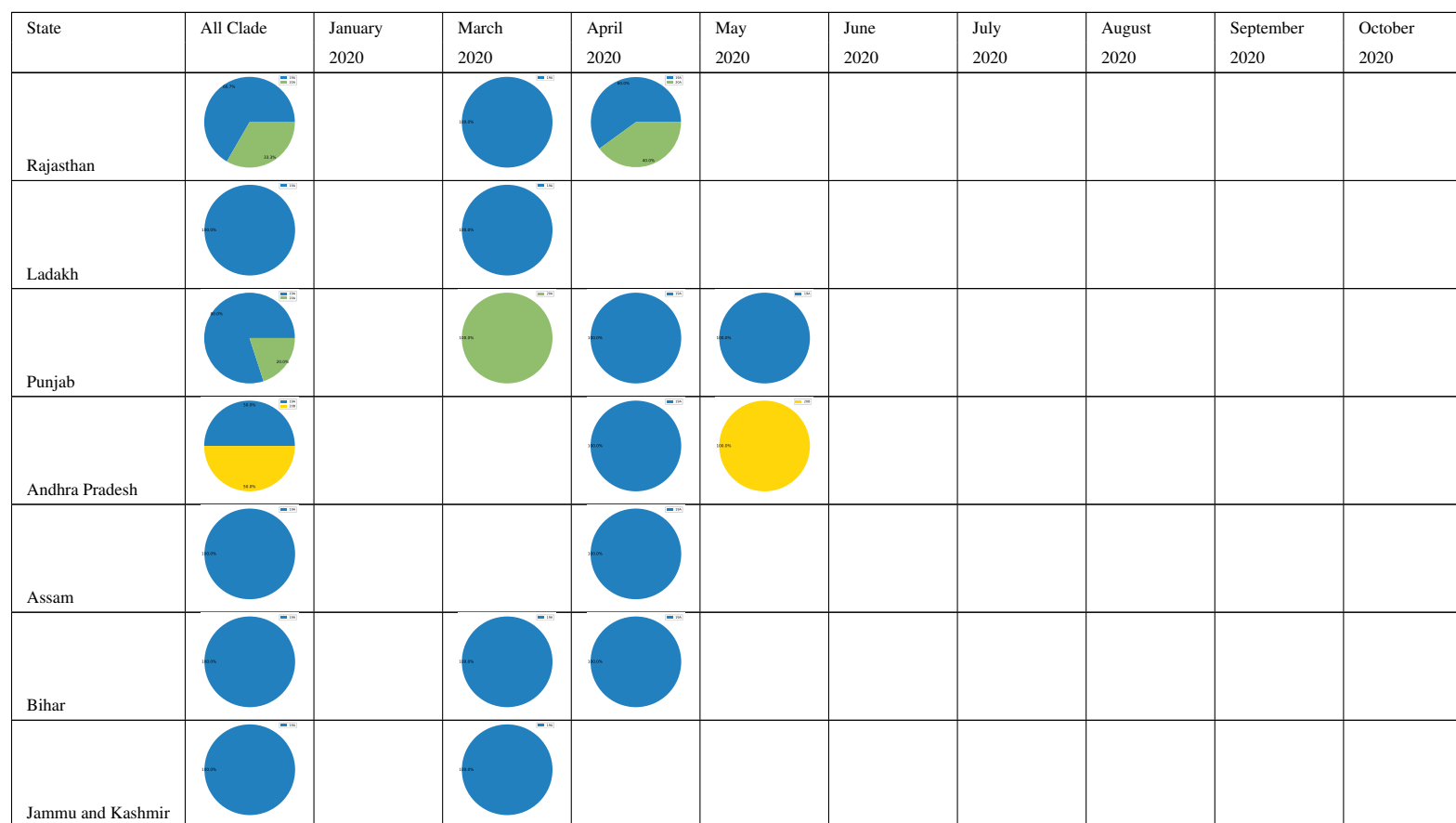
Country	All Clade	December 2019	January 2020	February 2020	March 2020	April 2020	May 2020	June 2020	July 2020	August 2020	September 2020	October 2020	November 2020	December 2020
Argentina														
Belarus														
Hungary														
Saudi Arabia														
Pakistan														
Cambodia														
Morocco														
Nepal														
Panama														

Table S6: Pie charts to represent month wise evolution of 3033 Indian SARS-CoV-2 genomes with respect to each clade for each state





Experiment Results and Files	Web Link
Analysed Nextstrain results for 15359 Global Excluding India SARS-CoV-2 Genomes	<a href="http://www.nitttrkol.ac.in/indrajit/projects/COVID-Evolution-SignatureSNPs-18K/downloads/supplementary/Analysis-of-Global-Excluding-India-15k-Genomes.xlsx">http://www.nitttrkol.ac.in/indrajit/projects/ COVID-Evolution-SignatureSNPs-18K/downloads/supplementary/Analysis-of-Global-Excluding-India-15k-Genomes.xlsx</a>
Analysed Nextstrain results for 3033 Indian SARS-CoV-2 Genomes	<a href="http://www.nitttrkol.ac.in/indrajit/projects/COVID-Evolution-SignatureSNPs-18K/downloads/supplementary/Analysis-of-Indian-3k-Genomes.xlsx">http://www.nitttrkol.ac.in/indrajit/projects/COVID-Evolution-SignatureSNPs-18K/downloads/supplementary/Analysis-of-Indian-3k-Genomes.xlsx</a>
Results of clade wise evolution of SARS-CoV-2 genomes with respect to each month for 71 countries	<a href="http://www.nitttrkol.ac.in/indrajit/projects/COVID-Evolution-SignatureSNPs-18K/downloads/supplementary/Global-Including-India-Country-Clade-Month-18392.xlsx">http://www.nitttrkol.ac.in/indrajit/projects/COVID-Evolution-SignatureSNPs-18K/downloads/supplementary/Global-Including-India-Country-Clade-Month-18392.xlsx</a>
Results of month wise evolution of SARS-CoV-2 genomes with respect to each clade for 71 countries	<a href="http://www.nitttrkol.ac.in/indrajit/projects/COVID-Evolution-SignatureSNPs-18K/downloads/supplementary/Global-Including-India-Country-Month-Clade-18392.xlsx">http://www.nitttrkol.ac.in/indrajit/projects/COVID-Evolution-SignatureSNPs-18K/downloads/supplementary/Global-Including-India-Country-Month-Clade-18392.xlsx</a>

Table S7: Various results of clade specific Signature SNPs for SARS-CoV-2 Genomes





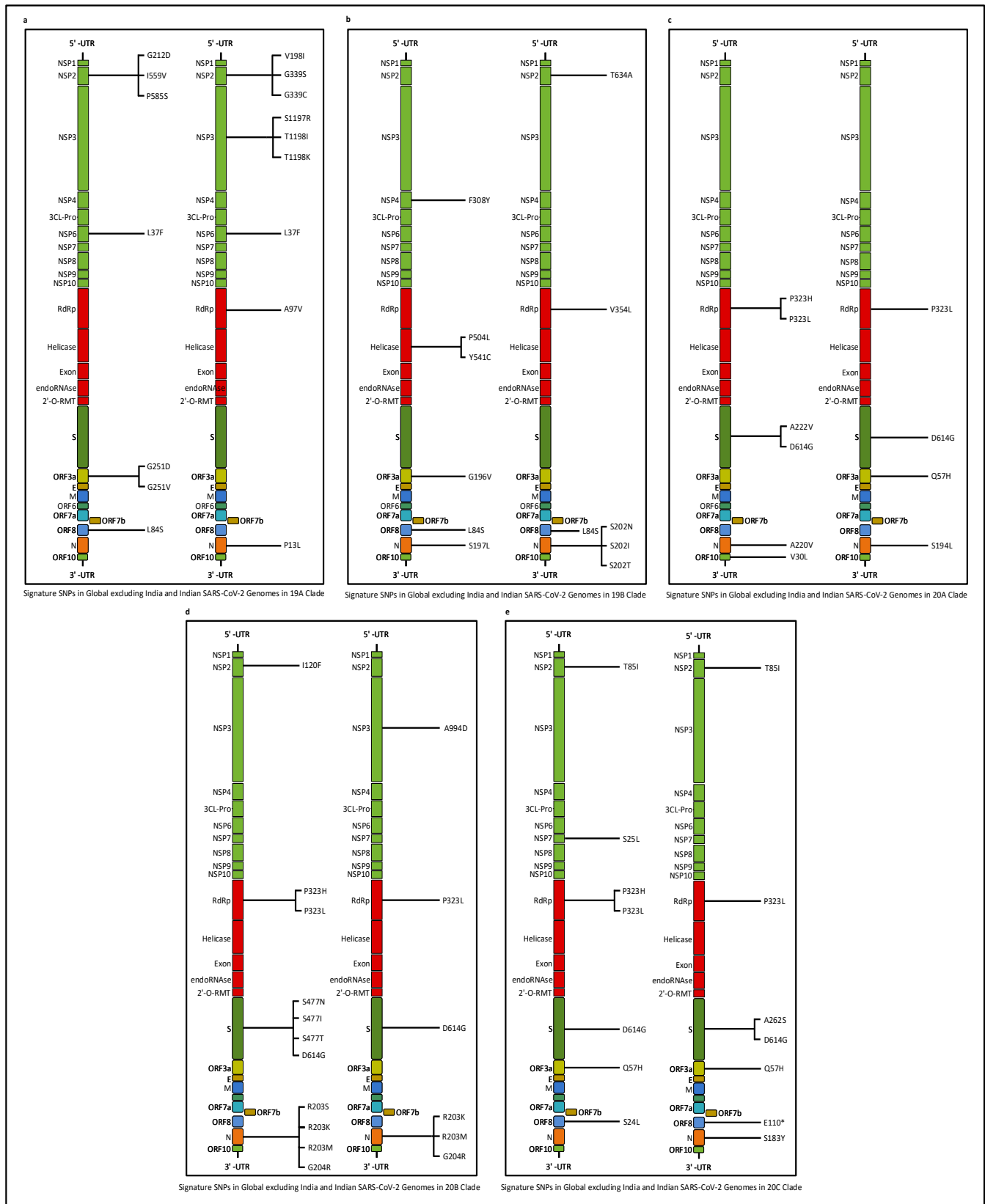


Figure S3: Amino acid changes in the proteins for the non-synonymous signature SNPs of Global excluding India and Indian SARS-CoV-2 genomes in 5 clades

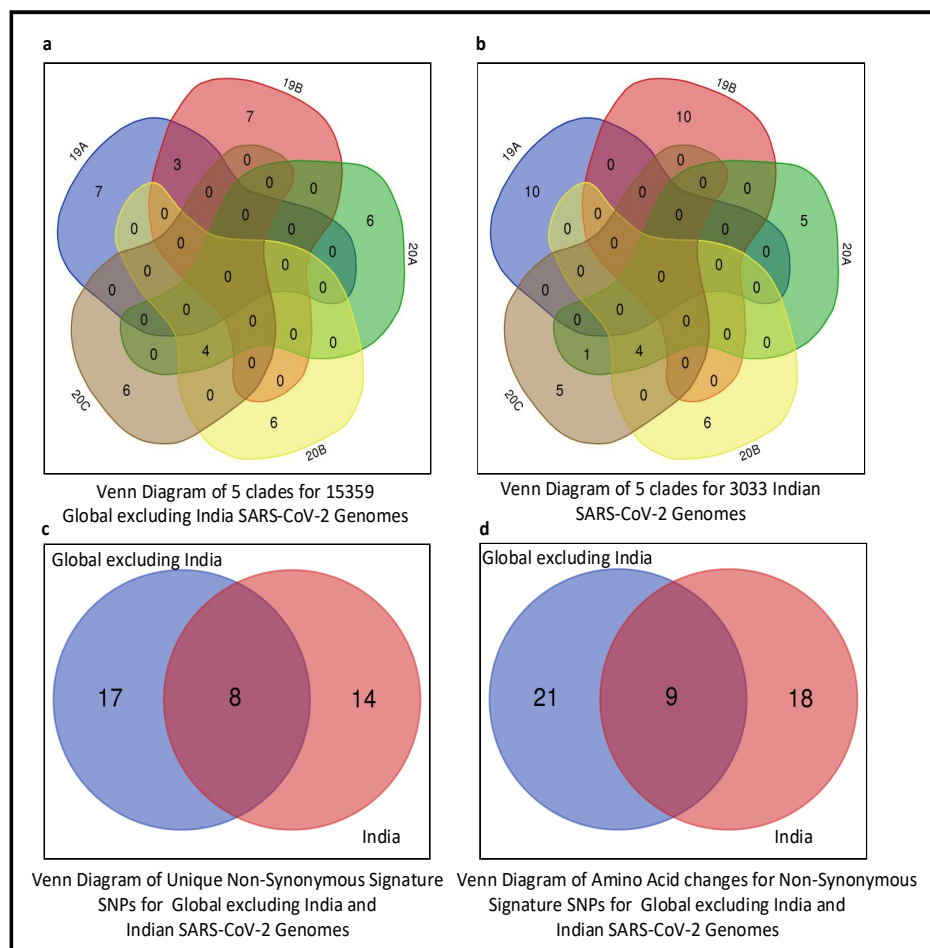


Figure S4: Venn diagrams of Global excluding India and Indian genomes to represent common non-synonymous signature SNPs and corresponding common amino acid changes