

Supplemental Online Content

Zhao LP, Cohen S, Zhao M, et al. Using haplotype-based artificial intelligence to evaluate SARS-CoV-2 novel variants and mutations. *JAMA Netw Open*. 2023;6(2):e230191. doi:10.1001/jamanetworkopen.2023.0191

eMethods.

eTable 1. A List of Known Variants Assigned by GISAID

eTable 2. Haplotype Frequencies Among Alpha Viruses

eTable 3. Haplotype Frequencies Among Beta Viruses

eTable 4. Haplotype Frequencies Among Delta Viruses

eTable 5. Haplotype Frequencies Among Epsilon Viruses

eTable 6. Haplotype Frequencies Among Eta Viruses

eTable 7. Haplotype Frequencies Among Gamma Viruses

eTable 8. Haplotype Frequencies Among GH/490R Viruses

eTable 9. Haplotype Frequencies Among Iota Viruses

eTable 10. Haplotype Frequencies Among Kappa Viruses

eTable 11. Haplotype Frequencies Among Lambda Viruses

eTable 12. Haplotype Frequencies Among Mu Viruses

eTable 13. Haplotype Frequencies Among Omicron Viruses

eTable 14. Haplotype Frequencies Among Theta Viruses

eTable 15. Haplotype Frequencies Among Zeta Viruses

eTable 16. Concordance Analysis in the Training Set

eTable 17. Concordance Analysis in the Validation Set

eTable 18. Concordance Analysis in the Prospective Set

eTable 19. Four Mixture Variants with Omicron-Delta

eTable 20. Four Mixture Variants with Omicron-Alpha

eTable 21. Four Mixture Variants with Omicron-Zeta

eTable 22. Four Mixture Variants with Omicron-Epsilon

eTable 23. Mixture Variant and Corresponding Lineages

eTable 24. Unidentifiable Omicron Viruses and Corresponding Lineages

eTable 25. New Mutations Among Variant-Unassigned and Unidentifiable Viruses

eFigure 1. Heatmap-Representation of Selected Polymutant Temporal Profile From January 1, 2020 to March 14, 2022 Within Every Variant

eFigure 2. Misclassification Errors by Haplotype-Based Variant Prediction (HVP), When the Prediction Probability Threshold Value is Set at 0.9 to 1

eFigure 3. Temporal Patterns of Sixteen Polymutants Identified From Variant-Unassigned 524 Viruses That Are Unpredictable by HAI, Excluding Those Core Polymutants of All Fourteen Variants

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods

Polymutants. GISAID is continuously receiving and archiving viral sequences globally, along with sample-specific metadata. Besides date and location of collecting viruses, the pertinent data to this project are extracted substituting amino acids (known as “AA.substitutions”), from aligning nucleotide sequence and translating nucleotides to amino acids by GISAID. Upon obtaining the text string on all substituting amino acids, we process all individual samples, and align all substitutions in a matrix form. In comparison with amino acids in the reference sequence, we convert the amino acid substitution matrix into a matrix of binary indicators (0, 1) for wildtype and mutant type, respectively. Note that most amino acids are monomorphic, some take one mutant type, and very few amino acids may have more than two substituting mutants. We call an amino acid a polymutant if it includes three or more mutating amino acids in the study population. While being conceptually straightforward, converting text strings of AA.substitutions into polymutant matrix and a matrix of binary polymutant indicators enables computations required by Statistical Learning Strategy (SLS).

Modeling Temporal Trends. Each polymutant has a specific temporal expansion pattern. Let binary indicator y_{jt} (=1 or 0) denote the presence or absence of mutant type, respectively, observed at time t for the j th polymutant. To model the non-linear temporal expansion, we applied a generalized additive model (GAM) to regress the mutant indicator over sample collection time via the following probability model,

$$\Pr(y_{jt} = 1 | t) = \frac{1}{1 + \exp[-\alpha - s_j(t)]}, \quad (1)$$

where α is a constant coefficient and $s_j(t)$ is a non-linear function of time t , and both are estimated by the restricted maximum likelihood method¹. Upon completing the estimation, the above function is used with the estimated coefficient and non-linear function to compute the probability of observing mutating amino acid at the time t , yielding a p -value that measures if the function $s_j(t)$ deviates from zero. Also produced is the maximum proportion as $P_{\max} = \max[\Pr(y_{jt} = 1 | t)]$. The function “gam” was used to fit the GAM (R packages MCV²). The smoothing parameter $k = 7$ was chosen.

Upon fitting the GAM, SLS can use the fitted values to compute locally averaged mutation percentage (LAMP) daily from the first to the last reporting day. The temporal pattern of LAMP can be used to describe the temporal expansion of the j th polymutant. By using p -value, SLS calls a polymutant with a significant trend, if p -value < 0.05, and with a substantial presence if the maximum proportion (P_{\max}) is greater than 10%.

Haplotype Analysis. SARS-CoV-2 is an RNA virus, i.e., a single strand, and thus multiple polymutants from the same virus share the same haplotype. Once two or more polymutants are identified, SLS converts polymutant matrix to a vector of polymutant haplotypes and computes their haplotype frequencies as part of the haplotype analysis. When confining the analysis to viruses of a specific variant, SLS can generate variant specific haplotype frequencies, denoted as $f(h | \text{variant})$.

Bayes' Prediction Probability. When predicting variant type based on viral sequences or polymutant haplotype (H), the haplotype-based artificial intelligence (HAI) relies on the following posterior probability:

$$\Pr(V | H = h) = \frac{f(H = h | \text{Variant} = V) p(\text{Variant} = V)}{\sum_V f(H = h | \text{Variant} = V) p(\text{Variant} = V)}, \quad (2)$$

in which the summation is over all possible variants, $f(H = h | \text{Variant} = V)$ is an empirically estimated haplotype frequency of a specific viral variant, and $p(\text{Variant} = V)$ is the proportion of the variant. Note that HAI treats variant-unassigned viruses as a separate class. On each virus, HAI computes an array of variant-specific prediction probabilities. If a prediction probability exceeds 0.99,

HAI predicts the corresponding variant. If the prediction probability for the variant-unassigned viruses exceeds 0.99, HAI predicts that the corresponding viruses do not have any known variant assignment, i.e., unpredictable (**UP**). Otherwise, HAI predicts that viruses include a mixture variant (**MV**) of two or more variant-specific haplotypes. Note that the posterior probability above (2) is strictly derived under the assumption that all variants are exclusive of each other, but, when dealing with recombinants, the quantity on the right hand side (2) is preferably interpreted as a risk score, in which the summation in the denominator serves as a normalizing factor so that the summation of all risk scores equals one.

Post-prediction modification. Among those MV predictions, their polymutant haplotypes include polymutants of two or more variants, some of which are recombinants. To tease out which MVs are recombinants, HAI utilizes a post-prediction modification. It extracts variant-specific polymutants. We call a MV recombinant if the polymutant includes mutating amino acids from two variants. For all other MVs, they will be re-assigned to be of specific variants.

Statistical Software. The statistical package R (version: R 4.2.1) and RStudio (Release 782775e, 2022-07-22) are used to implement all computational procedures in the HAI.

eTable 1. A List of Known Variants Assigned by GISAID

Fourteen variants established at GISAID and their basic annotations: WHO nomenclature, clade/lineages assigned by phylogenetic analysis, location(s) where the variant is first reported, and current designation by CDC

ID WHO	Clade/Lineages	First Detected in	Variant
1 Alpha	B.1.1.7+Q.*	United Kingdom	VOC
2 Beta	GH/501Y.V2 (B.1.351+B.1.351.2+B.1.351.3)	South Africa	VOC
3 Delta	GK (B.1.617.2+AY.*)	India	VOC
4 Epsilon	GH/452R.V1 (B.1.429+B.1.427)	USA/California	VOI
5 Eta	G/484K.V3 (B.1.525)	UK/Nigeria	VOI
6 Gamma	GR/501Y.V3 (P.1+P.1.*)	Brazil/Japan	VOC
7 GH/490R	(B.1.640+B.1.640.*)	Congo/France	VUM
8 Iota	GH/253G.V1 (B.1.526)	USA/New York	VOI
9 Kappa	G/452R.V3 (B.1.617.1)	India	VOI
10 Lambda	GR/452Q.V1 (C.37+C.37.1)	Peru	VOI
11 Mu	GH (B.1.621+B.1.621.1)	Colombia	VOI
12 Omicron	GRA (B.1.1.529+BA.*)	Botswana/South Africa/Hong Kong	VOC
13 Theta	GR/1092K.V1 (P.3)	Philippines	VOI
14 Zeta	GR/484K.V2 (P.2)	Brazil	VOI

eTable 2. Haplotype Frequencies Among Alpha Viruses

alpha. Haplotype frequencies of core haplotypes associated with alpha-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

eTable 3. Haplotype Frequencies Among Beta Viruses

Beta. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Count	Freq	Mut	T85	NSP2	S794	NSP3	K837	NSP3	N1778	NSP3	K90	NSP5	P323	NSP12	T588	NSP13	L18	Spike	D80	Spike	D215	Spike	K417	Spike	E484	Spike	N501	Spike	D614	Spike	A701	Spike	Q57	NS3	S171	NS3	P71	E	I121	NS8	T205	N	
1	3323	15.78%	15	I	S	N	N	R	L	T	L	A	G	N	K	Y	G	V	H	L	L	I	F
2	2770	13.16%	16	F					
3	1733	8.23%	16	.	L					
4	1703	8.09%	17	.	.	.	S	.	.	.	F						
5	1411	6.70%	17	I	L	.	.								
6	982	4.66%	17	F	L	.	.								
7	933	4.43%	16	H	.	F							
8	385	1.83%	13	E	N							
9	372	1.77%	16	L	.	.									
10	276	1.31%	18	.	.	S	.	.	.	F	L	.	.										
11	270	1.28%	16	I							
12	241	1.14%	14	F	.	.	E	N							
13	220	1.04%	13	D	D							
14	209	0.99%	17	L	L	.	.											
15	195	0.93%	14	.	K	.	.	I	.	.	K	.	.	.	K	.	.	.	S	.	L							
16	187	0.89%	15	I	.	.	K	.	.	K	.	.	.	S	.	L							
17	118	0.56%	15	.	S	.	.	F	.	.	E	N							
18	115	0.55%	14	D						
19	106	0.50%	14	L	E	N						
20	98	0.47%	15	I	.	.	E	N	L						
21	96	0.46%	14	S					
22	95	0.45%	15	F	S						
23	93	0.44%	17	.	.	.	H	.	F	L						
24	83	0.39%	14	E						
25	77	0.37%	14	K	.	.	.	K						
26	72	0.34%	14	.	.	.	H	.	F	.	E	N						
27	71	0.34%	14	T					
28	70	0.33%	15	F	.	K						
29	65	0.31%	10	K	D	D	.	.	.	A	.	S					
30	61	0.29%	15	K	.	.	.	F					
31	60	0.28%	12	K	E	N					
32	56	0.27%	15	.	.	.	F	.	E	N	L					
33	54	0.26%	14	E	N	L					
34	54	0.26%	10	D	D	K	E	N					
35	52	0.25%	14	K				
36	49	0.23%	14	Q					
37	48	0.23%	14	K	.	.	I	.	E	N	L					
38	42	0.20%	16	K	.	.	I	.	.	D	D	L					
39	42	0.20%	12	D	D	P				

eTable 4. Haplotype Frequencies Among Delta Viruses

Delta. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Count	Freq	Mut	NSP3	NSP3	NSP3	NSP3	NSP3	NSP4	NSP4	NSP4	NSP4	NSP6	NSP6	NSP12	NSP12	NSP13	NSP14	Spike	NS3	S26	I82	V82	NS7a	NS7b	M	T120	NS7a	NS7b	N	R203	G215	D377										
1	482881	22.55%	27	S	P	L	S	A	A	I	A	V	L	S	L	V	R	T	D	G	A	K	G	R	N	L	T	A	I	I	G	M	C	Y									
2	276071	12.89%	29	V	I							
3	233724	10.92%	26	G							
4	205786	9.61%	28	I								
5	74239	3.47%	22	A	L	P	P	.	V	V	T	T	A	.	.	A	.	.	V	T	.	G	.	.	.									
6	64840	3.03%	27	I	G								
7	50506	2.36%	30	.	.	.	V	I	.	V									
8	37778	1.76%	28	.	.	.	V	I	G									
9	37472	1.75%	27	.	.	.	V	G									
10	33301	1.56%	21	A	L	P	P	.	V	V	T	T	A	.	.	A	.	.	G	V	T	.	G	.	.	.									
11	30783	1.44%	25	G	E								
12	15479	0.72%	21	A	L	P	P	.	V	V	T	T	A	.	.	A	T	.	G	.	.	.									
13	11681	0.55%	26	.	.	P								
14	10633	0.50%	25	G	.	.	.	D								
15	10548	0.49%	20	A	L	P	P	.	V	V	T	T	A	.	.	A	.	G	T	.	G										
16	10494	0.49%	26	E								
17	8398	0.39%	24	G	.	.	.	V	T							
18	7751	0.36%	24	G	.	.	D	.	.	D								
19	6618	0.31%	28	.	.	V	.	.	.	G	.	.	I								
20	6330	0.30%	25	F	V	T							
21	6227	0.29%	28	V	I							
22	6076	0.28%	27	.	.	V	I	V	T							
23	5657	0.26%	26	.	.	V	G	E							
24	5648	0.26%	23	G	E	G	E	.	.	V	T								
25	5640	0.26%	27	I	I	.	.	.	V							
26	5437	0.25%	27	I	E	I	E							
27	4911	0.23%	26	V						
28	4782	0.22%	28	.	P	.	V	.	.	.	I	I							
29	4736	0.22%	26	I	G	E	.	.	.	I	G	E							
30	4657	0.22%	26	.	.	P						
31	4533	0.21%	26	D						
32	4517	0.21%	25	G	G	D	.	.	.							
33	4437	0.21%	26	I	G	.	.	.	I	G	.	.	D							
34	4261	0.20%	28	.	.	V	.	.	.	I	E	.	.	.	I	E							
35	4184	0.20%	17	A	L	P	P	.	V	.	T	T	.	G	P	A	T	.	G								
36	4060	0.19%	27	.	P	.	V	.	.	G	.	I	.	.	I							
37	3911	0.18%	26	.	.	V	.	.	.	I	I	.	.	.	V	T	T							
38	3793	0.18%	19	A	L	P	P	.	V	V	T	T	A	.	A	.	G	E	T	.	G									
39	3778	0.18%	24	G	.	L	T	.	G	.	L	T								

eTable 5. Haplotype Frequencies Among Epsilon Viruses

Epsilon. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Count	Freq	Mut	NSP2	NSP9	NSP12	NSP13	S13	Spike	W152	Spike	D614	Spike	Q57	NS3	A65	NS8	V100	N	T205	N	M234	N	
				T85	I65	P323	P53	D260	NSP13	L452	D614	Q57	NS3	A65	NS8	V100	N	T205	N	M234	N			
1	11106	30.92%	10	I	V	L	P	Y	I	C	R	G	H	A	V	I	M		
2	7653	21.30%	12	L	.	I	
3	6205	17.27%	10	.	I	.	L	
4	2642	7.35%	11	.	I	.	L	V	
5	716	1.99%	11	I	
6	605	1.68%	9	.	.	P	
7	497	1.38%	9	S	
8	443	1.23%	10	.	I	.	L	.	W	.	.	V	
9	321	0.89%	9	.	I	.	L	.	S	
10	304	0.85%	11	.	.	P	L	.	I	
11	272	0.76%	11	S	L	.	I	
12	262	0.73%	9	.	I	
13	256	0.71%	9	.	I	P	L	
14	236	0.66%	9	W	
15	169	0.47%	9	L	
16	158	0.44%	11	W	.	.	L	.	I	
17	148	0.41%	10	.	I	P	L	V	
18	148	0.41%	8	.	I	P	
19	141	0.39%	11	L	.	.	L	.	I	
20	117	0.33%	11	V
21	108	0.30%	9	.	I	.	L	.	S	W	.	V
22	107	0.30%	9	T
23	98	0.27%	9	.	I	.	L	.	W
24	98	0.27%	8	S	W
25	82	0.23%	10	.	I	.	L	.	.	L	.	V
26	79	0.22%	11	.	I	L	.	I
27	78	0.22%	8	.	I	.	L	.	S	W
28	74	0.21%	8	D	S
29	71	0.20%	10	.	I	P	L	.	I
30	70	0.19%	10	.	I	.	L	.	S	.	.	V
31	64	0.18%	9	.	I	.	L	.	.	L
32	54	0.15%	9	.	I	.	L	.	W	L	.	V
33	52	0.14%	10	S	W	.	.	L	.	I
34	51	0.14%	9	D
35	46	0.13%	8	W	L
36	44	0.12%	8	.	I	.	.	S
37	41	0.11%	8	.	I	Q

eTable 6. Haplotype Frequencies Among Eta Viruses

Eta. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Core Haplotypes	Count	Freq	Mut	T1189	NSP3	K1771	NSP3	P323	NSP12	K160	NSP16	Q52	Spike	A67	Spike	E484	Spike	D614	Spike	Q677	Spike	F888	Spike	L21	E	I82	M	A12	N	N	I2U5
					I	K	F	K	R	V	K	G	H	L	F	T	G	I														
1	IKFKRVKGHLFTGI	2188	46.41%	12	I	K	F	K	R	V	K	G	H	L	F	T	G	I														
2	IKFRRVKGHLFTGI	536	11.37%	13	.	.	.	R														
3	IRFKRVKGHLFTGI	425	9.02%	13	.	R														
4	IKFKQVKGHLLTGI	338	7.17%	10	Q	L	.	.													
5	IKFKRVKGHLFTGI	204	4.33%	11	L	.	.													
6	IKFRRVKGHLFTAT	149	3.16%	11	.	.	.	R	A	T												
7	IKFKQVKGHLLTGI	90	1.91%	11	Q														
8	IKFKRAKGHLFTGI	61	1.29%	11	A														
9	IKFKRVKGHLFTGT	50	1.06%	11	T													
10	IKFKRVKGHLFTAI	49	1.04%	11	A	.												
11	IKFKRVEGHHLFTGI	35	0.74%	11	E														
12	IKFKQAKGHLFTGI	25	0.53%	10	Q	A														
13	IKFKRVKGHLFTVI	21	0.45%	12	V	.												
14	IKFKRVKGHFFTGI	20	0.42%	11	F	.	.	.														
15	TKFKRVKGHLFTGI	19	0.40%	11	T														
16	IKFRRVKGHLFTGT	17	0.36%	12	.	.	R	T														
17	IRFKQVKGHLLTGI	17	0.36%	11	.	R	.	.	Q	L	.	.	.															
18	IRFKRVKGHLFTGI	15	0.32%	12	.	R	L	.	.	.															
19	IKFKRVKGHLFTGT	14	0.30%	10	L	.	.	T															
20	IRFKRAKGHLFTGI	11	0.23%	12	.	R	.	.	A															
21	IRFKRVKGHLFTGT	11	0.23%	12	.	R	T														
22	IKFKQAEGHLFTGI	10	0.21%	9	.	.	.	Q	A	E															
23	IRFKQVKGHLLFTGI	10	0.21%	12	.	R	.	Q															
24	IKFKQVKGHLLTGT	9	0.19%	9	.	.	.	Q	L	.	.	.	T															
25	IKFKQVKGHLLTVI	9	0.19%	10	.	.	.	Q	L	.	V	.																
26	IKFKRAKGHLFTGI	9	0.19%	10	A	.	.	.	L	.	.	.																
27	IKFKRVEGHHLFTGT	8	0.17%	10	E	T															
28	TKFKRVKGHFFTGT	8	0.17%	9	T	F	.	.	A	.														
29	TKFKRVKGHLFTAI	8	0.17%	10	T	A	.														
30	IKFKQVKGHFFIGI	7	0.15%	9	.	.	.	Q	.	.	F	.	I	.	.	.																
31	IKFKRVEGHFFTGT	7	0.15%	9	E	.	F	.	T	.	.	.																
32	IRFKRVEGHHLFTGI	7	0.15%	12	.	R	.	.	E																
33	TKFKRVKGHLLTAT	7	0.15%	8	T	L	.	A	T																
34	IKFKRVKGHLLTVI	6	0.13%	11	L	.	V	.																
35	IKFKRAKGHLFTAI	5	0.11%	10	A	A	.	.																
36	IKFKRVEGHLLTGI	5	0.11%	10	E	.	.	L	.	.	.																	
37	IKFKRVKGHFTAI	5	0.11%	10	F	.	A	.	.																	
38	IKFKRVKGHFFTGT	5	0.11%	10	F	.	.	T	.	.																	
39	IKFKRVKGHLFTGV	5	0.11%	12	V	.	.	.																

eTable 7. Haplotype Frequencies Among Gamma Viruses

Gamma. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Freq	%	#	NSP3	NSP3	NSP4	NSP12	NSP13	L370	K977	S184	P323	E341	L18	T20	P26	D138	R190	K417	E484	N501	D614	H655	T1027	V1176	S253	E92	NS8				
				L	Q	S	L	D	F	N	S	Y	S	T	K	Y	I	F	P	R	K	R	K	R	N	P80	N	R203	N	G204	N	
1	36853	61.22%	21	L	Q	S	L	D	F	N	S	Y	S	T	K	Y	I	F	P	R	K	R	K	R	N	P80	N	R203	N	G204	N	
2	7567	12.57%	22	.	.	N
3	1626	2.70%	20	R	
4	1400	2.33%	19	R	G		
5	1375	2.28%	20	D		
6	1365	2.27%	20	K		
7	1005	1.67%	19	E	N		
8	324	0.54%	20	T			
9	263	0.44%	20	R	.			
10	259	0.43%	20	.	.	N	T	V			
11	246	0.41%	18	E	.	T	E			
12	230	0.38%	21	.	.	N	D			
13	220	0.37%	21	.	.	N	K			
14	205	0.34%	20	E			
15	203	0.34%	21	.	.	N	T				
16	181	0.30%	20	S			
17	173	0.29%	17	E	N	R	G	.				
18	166	0.28%	18	R	.	E	N				
19	157	0.26%	20	E				
20	137	0.23%	19	T	V				
21	135	0.22%	20	V				
22	133	0.22%	21	F			
23	124	0.21%	20	N			
24	122	0.20%	19	R	K			
25	100	0.17%	16	L	T	P	D	R			
26	91	0.15%	18	L	T	P			
27	90	0.15%	20	.	K			
28	90	0.15%	17	.	.	.	L	T	P	.	R			
29	83	0.14%	18	R	R	G	.					
30	72	0.12%	19	D	.	K			
31	72	0.12%	20	H			
32	68	0.11%	20	.	.	E			
33	67	0.11%	20	S	P	.	.				
34	66	0.11%	20			
35	64	0.11%	20	.	N	E	N				
36	62	0.10%	19	D	E				

eTable 8. Haplotype Frequencies Among GH/490R Viruses

GH/490R. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Freq	%	#	P129	NSP2	G221	NSP2	E272	NSP2	L1301	NSP3	A1537	NSP3	S386	NSP4	R401	NSP4	T492	NSP4	V149	NSP6	Q586	NSP13	R216	NSP16	P9	Spike	E96	Spike	R190	Spike	I210	Spike	Y449	Spike	F490	Spike	N501	Spike	D936	Spike	T32	NS3	A54	NS3	D22	N	E378
				L	G	G	F	L	H	F	S	H	A	F	H	I	A	Q	R	L	Q	S	T	N	R	Y	H	I	S	Y	Q																	
1	78	18.10%	20	L	G	G	F	L	H	F	S	H	A	F	H	I	A	Q	R	L	Q	S	T	N	R	Y	H	I	S	Y	Q																	
2	69	16.01%	19	A	.	.																
3	37	8.58%	21	.	S	H	H	Y	.	.	.	A	.	.															
4	35	8.12%	22	.	S	H	H	A	.	.															
5	22	5.10%	18	A	D	.																
6	17	3.94%	18	Y	.	.	A	.	.																
7	13	3.02%	19	Y																
8	12	2.78%	18	R	I																
9	10	2.32%	9	.	.	E	L	A	S	R	T	V	I	.	S	.	D	.	A	.	E	.																		
10	9	2.09%	15	E	R	I	A	D	.																			
11	7	1.62%	17	R	I	.	.	.	A																		
12	6	1.39%	20	V	.	.	.																		
13	5	1.16%	8	.	.	E	L	A	S	R	T	V	I	Y	S	.	D	.	A	.	E	.																			
14	5	1.16%	16	.	.	.	A	R	I	.	.	.	A																			
15	5	1.16%	19	S	Y																	
16	4	0.93%	8	.	.	E	L	A	S	R	T	V	R	I	.	S	.	D	.	A	.	E	.																			
17	4	0.93%	10	.	.	E	L	A	S	R	T	V	I	.	S	.	D	.	.	.	E	.	.																			
18	4	0.93%	18	.	.	.	A	A																			
19	4	0.93%	19	.	.	.	A																		
20	4	0.93%	18	R	A																			
21	4	0.93%	21	.	S	H	H	.	.	R	A																			
22	3	0.70%	9	.	.	E	L	A	S	R	T	V	R	I	.	S	.	D	.	.	E	.	.	.																			
23	3	0.70%	16	E	R	I	D																		
24	3	0.70%	19	S																		
25	3	0.70%	18	.	.	L	A	H	H	Y	.	.	.	A																			
26	3	0.70%	20	.	S	H	H	.	.	.	R	I	.	.	.	A																			
27	2	0.46%	5	.	.	E	L	A	S	R	T	V	.	.	.	E	R	I	.	S	.	D	T	A	D	E	.	.																				
28	2	0.46%	17	E	E	A	D																			
29	2	0.46%	19	D																			
30	2	0.46%	17	.	.	L	A	A																			
31	2	0.46%	17	.	S	H	H	.	.	.	R	I	Y	F	N	.	A																			
32	2	0.46%	17	P	A	D																			
33	1	0.23%	7	.	.	E	L	A	S	R	T	V	.	.	P	.	R	I	.	S	.	D	.	A	.	E	.	.	.																			
34	1	0.23%	15	.	.	E	L	Y	.	.	.	A	.	E																		
35	1	0.23%	14	.	.	.	A	E	R	I	A	D																			
36	1	0.23%	14	.	.	.	A	R	I	.	.	.	A	D	E																				
37	1	0.23%	17	.	.	A	R	I																			
38	1	0.23%	17	.	.	A	I	A	A	.	.	.																				
39	1	0.23%	18	.	.	A	Y	.	.	.	A	.	.	.	E	.	.																			
40	1	0.23%	17	.	.	A	Y	.	.	.	A	.	.	.	A	.	.	.																			

eTable 9. Haplotype Frequencies Among Iota Viruses

Iota. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Freq	%	#	NSP2	NSP4	NSP12	NSP13	Spike	NS3	NS3	NS7a	NS8	N	N	N	N											
				T85	L438	P323	Q88	L5	T95	D253	S477	E484	D614	A701	Q957	P42	Q57	L116	T11	P13	P199	S202	M234				
1	10264	47.21%	15	I	P	L	H	F	I	G	S	K	G	V	Q	L	H	L	I	P	L	S	I				
2	5329	24.51%	16	N	E	.	A	R	.	.	F	.	L	P	R	M					
3	1395	6.42%	15	N	E				
4	902	4.15%	14	E				
5	296	1.36%	15	N	E	.	A	R	.	.	L	P	R	M	.	.	.				
6	279	1.28%	14	A			
7	178	0.82%	14	L			
8	152	0.70%	15	N	E	.	A	.	.	F	.	L	P	R	M	.	.				
9	105	0.48%	14	P			
10	101	0.46%	14	.	.	.	Q			
11	87	0.40%	14	T			
12	84	0.39%	14	D			
13	83	0.38%	14	N	E	.	A			
14	77	0.35%	13	A	.	.	A	.	.	T			
15	74	0.34%	15	E	.	A	R	.	F	.	L	P	R	M			
16	72	0.33%	15	.	.	.	L	.	N	E	.	A	R	.	F	.	L	P	R	M			
17	55	0.25%	13	N	E	.	A	.	.	F	T	.	P	R	M			
18	53	0.24%	13	.	P	A			
19	49	0.23%	13	.	.	L	.	E		
20	46	0.21%	15	.	.	Q	.	.	N	E	.	A	R	.	F	.	L	P	R	M			
21	43	0.20%	15	N	E	.	A	R	P	.	F	.	L	P	R	M			
22	43	0.20%	16	F		
23	41	0.19%	14	.	.	L	.	N	E	.	A	R	.	.	L	P	R	M			
24	38	0.17%	14	E	.	A	R	.	.	L	P	R	M			
25	37	0.17%	13	P	.	M		
26	36	0.17%	16	L	
27	33	0.15%	15	D	N	E	.	A	R	.	F	.	L	P	R	M		
28	29	0.13%	14	L	
29	29	0.13%	14	.	P	
30	28	0.13%	15	.	.	.	N	E	.	A	R	.	F	T	L	P	R	M	
31	26	0.12%	13	E	.	A	
32	26	0.12%	13	.	Q	L	
33	24	0.11%	13	.	.	D	E	
34	24	0.11%	13	.	.	N	E	.	A	.	.	T	
35	24	0.11%	13	.	.	L	T

eTable 10. Haplotype Frequencies Among Kappa Viruses

Kappa. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Freq	%	#	NSP3	NSP6	NSP12	NSP13	NSP15	NSP15	NSP15	Spike	Spike	Spike	Spike	Spike	H1101	Spike	NS3	M	NS7a	NS8	N	N	N						
				T749	T77	P323	G206	M429	P65	K259	S261	I95	G142	E154	L452	E484	D614	P681	Q1071	H1101	S26	I82	V82	S69	D3	R203	D377			
1	679	17.36%	20	I	A	L	C	I	P	R	A	I	D	K	R	Q	G	R	H	H	L	S	A	S	D	M	Y			
2	228	5.83%	21	S			
3	165	4.22%	19	.	.	.	G	.	.	S	T	D	.	I	.	L	Y	.	.				
4	145	3.71%	18	.	.	.	G	.	.	S	T	D	.	I	.	L				
5	126	3.22%	18	G	Q				
6	96	2.45%	15	.	.	.	G	.	.	S	T	G	.	.	.	Q	D	.	I				
7	89	2.28%	19	E				
8	78	1.99%	18	G	E				
9	65	1.66%	19	G				
10	65	1.66%	17	.	.	.	G	.	.	S	T	D	.	I				
11	57	1.46%	17	T	G	E				
12	57	1.46%	17	.	.	.	G	.	.	T	I				
13	49	1.25%	16	.	.	.	G	.	.	S	T	G	.	.	.	Q	D	.	I	.	Y				
14	40	1.02%	15	.	.	.	G	.	.	T	G	E	I				
15	38	0.97%	14	.	.	.	G	.	.	S	T	G	E	.	.	Q	D	.	I				
16	37	0.95%	18	T	G				
17	34	0.87%	18	.	.	.	G	.	.	S	T	Y	.	I	.	L				
18	34	0.87%	16	.	.	.	G	.	.	S	T	G	E	.	.	D	.	I	.	L				
19	34	0.87%	18	.	.	.	G	.	.	S	T	G	.	.	.	D	.	I	.	L	Y				
20	33	0.84%	15	.	.	.	G	.	.	T	G	.	.	.	Q	.	I				
21	32	0.82%	16	.	.	.	G	.	.	S	T	G	.	.	.	D	.	I			
22	29	0.74%	18	.	.	.	G	.	.	S	T	.	E	.	.	D	.	I	.	L	Y			
23	26	0.66%	15	.	.	.	G	.	.	S	T	G	E	.	.	D	.	I			
24	23	0.59%	17	.	.	.	G	.	.	S	T	G	.	.	.	D	.	I	.	L			
25	23	0.59%	19	T		
26	22	0.56%	17	G	E	.	.	.	Q		
27	22	0.56%	20	S	.	.	E		
28	21	0.54%	14	.	.	.	G	.	.	T	G	E	.	.	Q	.	I			
29	21	0.54%	16	.	.	.	G	.	.	T	G	.	.	.	I		
30	21	0.54%	17	.	.	.	G	.	.	S	T	G	E	.	.	D	.	I	.	L	Y			
31	20	0.51%	16	.	.	.	G	.	.	S	T	G	E	.	.	D	.	I	.	Y		
32	18	0.46%	19	V	S	.	.	Q		
33	18	0.46%	19	.	.	.	S	.	.	G	E		
34	18	0.46%	20	.	.	.	S	.	.	G		
35	18	0.46%	20	T	.	.	S	
36	17	0.43%	19	V
37	17	0.43%	19	Q	
38	16	0.41%	17	.	.	.	G	.	.	S	T	.	E	.	.	D	.	I	.	L	
39	15	0.38%	16	T	G	E	.	.	Q	
40	15	0.38%	17	.	.	.	G	.	.	S	T	G	.	.	.	D	.	I	.	Y	

eTable 11. Haplotype Frequencies Among Lambda Viruses Lambda

Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Freq	%	#	NSP3	NSP3	NSP3	NSP3	NSP4	NSP4	NSP5	NSP12	Spike	Spike	Spike	Spike	Spike	Spike	NS3	N	N	N	N	N			
1	2885	56.57%	18	I	S	V	P	I	S	L	V	I	N	Q	S	G	N	A	L	A	K	R	C			
2	472	9.25%	17	D			
3	401	7.86%	17	G			
4	323	6.33%	20	S	.	P			
5	137	2.69%	16	G	T			
6	104	2.04%	19	D	.	.	.	S	.	P			
7	54	1.06%	19	P		
8	40	0.78%	16	G	.	.	D		
9	36	0.71%	15	G	T	D		
10	30	0.59%	17	P		
11	27	0.53%	15	D	R	G		
12	23	0.45%	16	.	P	D		
13	23	0.45%	18	D	P		
14	23	0.45%	16	R	G		
15	20	0.39%	18	S	
16	18	0.35%	15	G	.	G	T	
17	17	0.33%	17	R	
18	17	0.33%	17	T	
19	12	0.24%	17	.	P	
20	11	0.22%	17	F	
21	11	0.22%	18	F	
22	10	0.20%	14	G	.	G	T	D	
23	9	0.18%	16	L	F	
24	7	0.14%	14	.	P	D	L	F	
25	7	0.14%	19	G	S	.	P	
26	7	0.14%	17	G	
27	6	0.12%	15	.	P	L	F	
28	6	0.12%	17	D	S	
29	6	0.12%	14	D	R	G	G	
30	6	0.12%	19	S
31	6	0.12%	19	S

eTable 12. Haplotype Frequencies Among Mu Viruses Mu

Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Freq	%	#	NSP3	T237	T720	NSP3	N1329	NSP3	T189	NSP4	T492	NSP4	Q160	NSP6	P323	NSP12	Y521	NSP12	E261	NSP13	P419	NSP13	P491	NSP13	S261	NSP15	T95	Spike	Y144	Spike	Y145	Spike	R346	Spike	E484	Spike	N501	Spike	D614	Spike	P681	Spike	D950	Spike	I20	NS3	Q57	NS3	L106	NS3	V256	NS3	N257	NS3	V259	NS3	T11	NS8	P38	NS8	S67	NS8	T205	N
1	1406	17.96%	21	A	I	N	T	I	R	L	Y	E	S	P	S	I	S	N	K	K	Y	G	H	N	I	H	L	I	N	V	K	S	F	I																															
2	667	8.52%	27	.	.	D	I	S	L																														
3	561	7.16%	24	C	D	M																															
4	482	6.16%	22	F																															
5	405	5.17%	22	S																															
6	238	3.04%	19	Y	Y																															
7	224	2.86%	20	Y																															
8	210	2.68%	22	C																															
9	140	1.79%	26	.	.	D	I	.	.	.	S	L	H	L	.	S	.	.																																
10	125	1.60%	20	B	S																																
11	110	1.40%	21	C	B	S																																
12	96	1.23%	25	C	D	.	L	M																																
13	92	1.17%	20	S																															
14	81	1.03%	21	S	.	N	Y																															
15	75	0.96%	21	Y	F																														
16	64	0.82%	21	S	.	Y																															
17	62	0.79%	20	T	Y																															
18	59	0.75%	21	S	B	S																															
19	57	0.73%	23	C	D	B	M	S																														
20	49	0.63%	21	B	.	F	.	.	.	S																														
21	44	0.56%	23	C	D	.	.	Y	M																															
22	42	0.54%	23	C	D	M	S																															
23	42	0.54%	18	T	Y	Y																															
24	37	0.47%	19	E	N	F	.	.	S																															
25	37	0.47%	21	F	.	.	S																															
26	36	0.46%	21	C	.	.	Y																															
27	31	0.40%	19	D	.	.	S																															
28	26	0.33%	22	C	D	.	.	Y	Y	.	.	.	M																															
29	26	0.33%	19	S	T																															
30	25	0.32%	21	S	.	T	Y																														
31	24	0.31%	18	Y	Y	V																														
32	24	0.31%	21	S	S																															
33	22	0.28%	26	.	D	I	.	.	.	S	H	L																															
34	21	0.27%	23	C	D	.	.	T	Y	.	.	.	M																															
35	21	0.27%	20	D																														
36	20	0.26%	23	C	D																															
37	20	0.26%	20	V																														
38	20	0.26%	20	P																															
39	19	0.24%	25	.	D	I	.	.	.	S	L	.	Y	H																															

eTable 13. Haplotype Frequencies Among Omicron Viruses

Omicron. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Freq	%	#	NSP1	T24	K88	G489	NSP3	L1266	NSP3	A1892	NSP3	T24	N327	NSP4	I42	NSP4	T492	NSP4	P132	NSP5	I42	NSP5	T112	NSP12	I42	NSP13	T112	NSP14	I42	NSP15	T112	NSP16	I42	NSP17	T112	NSP18	I42	NSP19	T112	NSP20	I42	NSP21	T112	NSP22	I42	NSP23	T112	NSP24	I42	NSP25	T112	NSP26	I42	NSP27	T112	NSP28	I42	NSP29	T112	NSP30	I42	NSP31	T112	NSP32	I42	NSP33	T112	NSP34	I42	NSP35	T112	NSP36	I42	NSP37	T112	NSP38	I42	NSP39	T112	NSP40	I42	NSP41	T112	NSP42	I42	NSP43	T112	NSP44	I42	NSP45	T112	NSP46	I42	NSP47	T112	NSP48	I42	NSP49	T112	NSP50	I42	NSP51	T112	NSP52	I42	NSP53	T112	NSP54	I42	NSP55	T112	NSP56	I42	NSP57	T112	NSP58	I42	NSP59	T112	NSP60	I42	NSP61	T112	NSP62	I42	NSP63	T112	NSP64	I42	NSP65	T112	NSP66	I42	NSP67	T112	NSP68	I42	NSP69	T112	NSP70	I42	NSP71	T112	NSP72	I42	NSP73	T112	NSP74	I42	NSP75	T112	NSP76	I42	NSP77	T112	NSP78	I42	NSP79	T112	NSP80	I42	NSP81	T112	NSP82	I42	NSP83	T112	NSP84	I42	NSP85	T112	NSP86	I42	NSP87	T112	NSP88	I42	NSP89	T112	NSP90	I42	NSP91	T112	NSP92	I42	NSP93	T112	NSP94	I42	NSP95	T112	NSP96	I42	NSP97	T112	NSP98	I42	NSP99	T112	NSP100	I42	NSP101	T112	NSP102	I42	NSP103	T112	NSP104	I42	NSP105	T112	NSP106	I42	NSP107	T112	NSP108	I42	NSP109	T112	NSP110	I42	NSP111	T112	NSP112	I42	NSP113	T112	NSP114	I42	NSP115	T112	NSP116	I42	NSP117	T112	NSP118	I42	NSP119	T112	NSP120	I42	NSP121	T112	NSP122	I42	NSP123	T112	NSP124	I42	NSP125	T112	NSP126	I42	NSP127	T112	NSP128	I42	NSP129	T112	NSP130	I42	NSP131	T112	NSP132	I42	NSP133	T112	NSP134	I42	NSP135	T112	NSP136	I42	NSP137	T112	NSP138	I42	NSP139	T112	NSP140	I42	NSP141	T112	NSP142	I42	NSP143	T112	NSP144	I42	NSP145	T112	NSP146	I42	NSP147	T112	NSP148	I42	NSP149	T112	NSP150	I42	NSP151	T112	NSP152	I42	NSP153	T112	NSP154	I42	NSP155	T112	NSP156	I42	NSP157	T112	NSP158	I42	NSP159	T112	NSP160	I42	NSP161	T112	NSP162	I42	NSP163	T112	NSP164	I42	NSP165	T112	NSP166	I42	NSP167	T112	NSP168	I42	NSP169	T112	NSP170	I42	NSP171	T112	NSP172	I42	NSP173	T112	NSP174	I42	NSP175	T112	NSP176	I42	NSP177	T112	NSP178	I42	NSP179	T112	NSP180	I42	NSP181	T112	NSP182	I42	NSP183	T112	NSP184	I42	NSP185	T112	NSP186	I42	NSP187	T112	NSP188	I42	NSP189	T112	NSP190	I42	NSP191	T112	NSP192	I42	NSP193	T112	NSP194	I42	NSP195	T112	NSP196	I42	NSP197	T112	NSP198	I42	NSP199	T112	NSP200	I42	NSP201	T112	NSP202	I42	NSP203	T112	NSP204	I42	NSP205	T112	NSP206	I42	NSP207	T112	NSP208	I42	NSP209	T112	NSP210	I42	NSP211	T112	NSP212	I42	NSP213	T112	NSP214	I42	NSP215	T112	NSP216	I42	NSP217	T112	NSP218	I42	NSP219	T112	NSP220	I42	NSP221	T112	NSP222	I42	NSP223	T112	NSP224	I42	NSP225	T112	NSP226	I42	NSP227	T112	NSP228	I42	NSP229	T112	NSP230	I42	NSP231	T112	NSP232	I42	NSP233	T112	NSP234	I42	NSP235	T112	NSP236	I42	NSP237	T112	NSP238	I42	NSP239	T112	NSP240	I42	NSP241	T112	NSP242	I42	NSP243	T112	NSP244	I42	NSP245	T112	NSP246	I42	NSP247	T112	NSP248	I42	NSP249	T112	NSP250	I42	NSP251	T112	NSP252	I42	NSP253	T112	NSP254	I42	NSP255	T112	NSP256	I42	NSP257	T112	NSP258	I42	NSP259	T112	NSP260	I42	NSP261	T112	NSP262	I42	NSP263	T112	NSP264	I42	NSP265	T112	NSP266	I42	NSP267	T112	NSP268	I42	NSP269	T112	NSP270	I42	NSP271	T112	NSP272	I42	NSP273	T112	NSP274	I42	NSP275	T112	NSP276	I42	NSP277	T112	NSP278	I42	NSP279	T112	NSP280	I42	NSP281	T112	NSP282	I42	NSP283	T112	NSP284	I42	NSP285	T112	NSP286	I42	NSP287	T112	NSP288	I42	NSP289	T112	NSP290	I42	NSP291	T112	NSP292	I42	NSP293	T112	NSP294	I42	NSP295	T112	NSP296	I42	NSP297	T112	NSP298	I42	NSP299	T112	NSP300	I42	NSP301	T112	NSP302	I42	NSP303	T112	NSP304	I42	NSP305	T112	NSP306	I42	NSP307	T112	NSP308	I42	NSP309	T112	NSP310	I42	NSP311	T112	NSP312	I42	NSP313	T112	NSP314	I42	NSP315	T112	NSP316	I42	NSP317	T112	NSP318	I42	NSP319	T112	NSP320	I42	NSP321	T112	NSP322	I42	NSP323	T112	NSP324	I42	NSP325	T112	NSP326	I42	NSP327	T112	NSP328	I42	NSP329	T112	NSP330	I42	NSP331	T112	NSP332	I42	NSP333	T112	NSP334	I42	NSP335	T112	NSP336	I42	NSP337	T112	NSP338	I42	NSP339	T112	NSP340	I42	NSP341	T112	NSP342	I42	NSP343	T112	NSP344	I42	NSP345	T112	NSP346	I42	NSP347	T112	NSP348	I42	NSP349	T112	NSP350	I42	NSP351	T112	NSP352	I42	NSP353	T112	NSP354	I42	NSP355	T112	NSP356	I42	NSP357	T112	NSP358	I42	NSP359	T112	NSP360	I42	NSP361	T112	NSP362	I42	NSP363	T112	NSP364	I42	NSP365	T112	NSP366	I42	NSP367	T112	NSP368	I42	NSP369	T112	NSP370	I42	NSP371	T112	NSP372	I42	NSP373	T112	NSP374	I42	NSP375	T112	NSP376	I42	NSP377	T112	NSP378	I42	NSP379	T112	NSP380	I42	NSP381	T112	NSP382	I42	NSP383	T112	NSP384	I42	NSP385	T112	NSP386	I42	NSP387	T112	NSP388	I42	NSP389	T112	NSP390	I42	NSP391	T112	NSP392	I42	NSP393	T112	NSP394	I42	NSP395	T112	NSP396	I42	NSP397	T112	NSP398	I42	NSP399	T112	NSP400	I42	NSP401	T112	NSP402	I42	NSP403	T112	NSP404	I42	NSP405	T112	NSP406	I42	NSP407	T112	NSP408	I42	NSP409	T112	NSP410	I42	NSP411	T112	NSP412	I42	NSP413	T112	NSP414	I42	NSP415	T112	NSP416	I42	NSP417	T112	NSP418	I42	NSP419	T112	NSP420	I42	NSP421	T112	NSP422	I42	NSP423	T112	NSP424	I42	NSP425	T112	NSP426	I42	NSP427	T112	NSP428	I42	NSP429	T112	NSP430	I42	NSP431	T112	NSP432	I42	NSP433	T112	NSP434	I42	NSP435	T112	NSP436	I42	NSP437	T112	NSP438	I42	NSP439	T112	NSP440	I42	NSP441	T112	NSP442	I42	NSP443	T112	NSP444	I42	NSP445	T112	NSP446	I42	NSP447	T112	NSP448	I42	NSP449	T112	NSP450	I42	NSP451	T112	NSP452	I42	NSP453	T112	NSP454	I42	NSP455	T112	NSP456	I42	NSP457	T112	NSP458	I42	NSP459	T112	NSP460	I42	NSP461	T112	NSP462	I42	NSP463	T112	NSP464	I42	NSP465	T112	NSP466	I42	NSP467	T112	NSP468	I42	NSP469	T112	NSP470	I42	NSP471	T112	NSP472	I42	NSP473	T112	NSP474	I42	NSP475	T112	NSP476	I42	NSP477	T112	NSP478	I42	NSP479	T112	NSP480	I42	NSP481	T112	NSP482	I42	NSP483	T112	NSP484	I42	NSP485	T112	NSP486	I42	NSP487	T112	NSP488	I42	NSP489	T112	NSP490	I42	NSP491	T112	NSP492	I42	NSP493	T112	NSP494	I42	NSP495	T112	NSP496	I42	NSP497	T112	NSP498	I42	NSP499	T112	NSP500	I42	NSP501	T112	NSP502	I42	NSP503	T112	NSP504	I42	NSP505	T112	NSP506	I42	NSP507	T112	NSP508	I42	NSP509	T112	NSP510	I42	NSP511	T112	NSP512	I42	NSP513	T112	NSP514	I42	NSP515	T112	NSP516	I42	NSP517	T112	NSP518	I42	NSP519	T112	NSP520	I42	NSP521	T112	NSP522	I42	NSP523	T112	NSP524	I42	NSP525	T112	NSP526	I42	NSP527	T112	NSP528	I42	NSP529	T112	NSP530	I42	NSP531	T112	NSP532	I42	NSP533	T112	NSP534	I42	NSP535	T112	NSP536	I42	NSP537	T112	NSP538	I42	NSP539	T112	NSP540	I42	NSP541	T112	NSP542	I42	NSP543	T112	NSP544	I42	NSP545	T112	NSP546	I42	NSP547	T112	NSP548	I42	NSP549	T112	NSP550	I42	NSP551	T112	NSP552	I42	NSP553	T112	NSP554	I42	NSP555	T112	NSP556	I42	NSP557	T112	NSP558	I42	NSP559	T112	NSP560	I42	NSP561	T112	NSP562	I42	NSP563	T112	NSP564	I42	NSP565	T112	NSP566	I42	NSP567	T112	NSP568	I42	NSP569	T112	NSP570	I42	NSP571	T112	NSP572	I42	NSP573	T112	NSP574	I42	NSP575	T112	NSP576	I42	NSP577	T112	NSP578	I42	NSP579	T112	NSP580	I42	NSP581	T112	NSP582	I42	NSP583	T112	NSP584	I42	NSP585	T112	NSP586	I42	NSP587	T112	NSP588	I42	NSP589	T112	NSP590	I42	NSP591	T112	NSP592	I42	NSP593	T112	NSP594	I42	NSP595	T112	NSP596	I42	NSP597	T112	NSP598	I42	NSP599	T112	NSP600	I42	NSP601	T112	NSP602	I42	NSP603	T112	NSP604	I42	NSP605	T112	NSP606	I42	NSP607	T112	NSP608	I42	NSP609	T112	NSP610	I42	NSP611	T112	NSP612	I42	NSP613	T112	NSP614	I42	NSP615	T112	NSP616	I42	NSP617	T112	NSP618	I42	NSP619	T112	NSP620	I42	NSP621	T112	NSP622	I42	NSP623	T112	NSP624	I42	NSP625	T112	NSP626	I42	NSP627	T112	NSP628	I42	NSP629	T112	NSP630	I42	NSP631	T112	NSP632	I42	NSP633	T112	NSP634	I42	NSP635	T112	NSP636	I42	NSP637	T112	NSP638	I42	NSP639	T112	NSP640	I42	NSP641	T112	NSP642	I42	NSP643	T112	NSP644	I42	NSP645	T112	NSP646	I42	NSP647	T112	NSP648

eTable 14. Haplotype Frequencies Among Theta Viruses

Theta. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

eTable 15. Haplotype Frequencies Among Zeta Viruses

Zeta. Haplotype frequencies of core haplotypes associated with variant-specific polymutants, for those haplotypes with 0.1% haplotype frequency.

ID	Freq	%	#	NSP5											
				L205	L71	P323	E484	D614	V1176	A119	R203	G204	M234	N	
1	2550	0.8736	10	V	F	L	K	G	F	S	K	R	I		
2	85	0.0291	9	L		
3	80	0.0274	9	.	.	.	E		
4	21	0.0072	9	V		
5	18	0.0062	9	A	.	.	.		
6	17	0.0058	8	.	.	P	E		
7	17	0.0058	9	.	.	P		
8	14	0.0048	8	.	.	P	.	.	V		
9	12	0.0041	7	R	G	M	.		
10	10	0.0034	8	R	G	.			
11	10	0.0034	9	R	.	.			
12	9	0.0031	9	M	.			
13	8	0.0027	8	.	.	.	E	M	.		
14	8	0.0027	10	L	.			
15	4	0.0014	8	A	.	M	.			
16	4	0.0014	8	V	.	.	M	.		
17	3	0.001	6	L	.	E	D	V		
18	3	0.001	8	A	R	.	.	.		

eTable 16. Concordance Analysis in the Training Set

Result from the concordance analysis of GISAID assigned variant (columns) with haplotype-based variant assignment (rows) in the training data set with 4,393,998 viruses (50% of all viruses at GISAID, downloaded on March 14, 2022).

Prediction ¹	Alpha	Beta	Delta	Epsilon	Eta	Gamma	GH/490R	Iota	Kappa	Lambda	Mu	Omicron	Theta	Zeta	UA ⁴	Sub-total
Alpha	573,599			4		3									2,497	576,103
Beta		8,021													90	8,111
Delta	1	1	2,120,806						6			2			1,486	2,122,302
Epsilon	1		9	35,883					1		3	1			145,822	181,720
Eta	1		2		4,695										33	4,731
Gamma			1			59,471									95	59,567
GH/490R						422										422
Iota	1	1					18,589								68	18,659
Kappa	1		2					3,796							48	3,847
Lambda					1				5,077						6	5,084
Mu				1					5,634						38	5,673
Omicron			2						1,487,867						704	1,488,573
Theta										148					1	149
Zeta					4				1				1	2,918	24,546	27,470
MV²	385	12,940	19,917	11	16	658	8	3,140	43	9	2,193	112,137	181	1	7,633	159,272
UP³	446	97	1,702	53	3	63	1	12	66	13	2	1,333	1	360,335	364,127	
Sub-total	574,435	21,060	2,142,445	35,948	4,714	60,200	431	21,742	3,911	5,100	7,832	1,601,340	330	2,920	543,402	5,025,810

¹Posterior probability threshold=0.99; ²MV for mixture of variants; ³UP for unpredictable variant; ⁴UC for unclassified variant

eTable 17. Concordance Analysis in the Validation Set

Result from the concordance analysis of GISAID assigned variant (columns) with haplotype-based variant assignment (rows) in the training data set with 4,393,998 viruses (50% of all viruses at GISAID, downloaded on March 14, 2022).

Prediction ¹	Alpha	Beta	Delta	Epsilon	Eta	Gamma	GH/490R	Iota	Kappa	Lambda	Mu	Omicron	Theta	Zeta	UA ⁴	Sub-total
Alpha	1,148,060	1	4			6			1						5,218	1,153,290
Beta		15,492													145	15,637
Delta	4		4,246,693			1			11			5			3,184	4,249,898
Epsilon	2	8	24	71,899		1			3			7	3		298,020	369,967
Eta	1	1	5		9,398										92	9,497
Gamma			1			118,898									193	119,092
GH/490R						8										8
Iota		2					37,951								143	38,096
Kappa			40					7,558							113	7,711
Lambda					2				10,058						16	10,076
Mu				1						7,348					60	7,409
Omicron			2						1		3,003,288				1,600	3,004,891
Theta												414		1	415	
Zeta	3	5	3			12		4			2		5,886	53,014	58,929	
MV²	732	26,350	35,585	56	54	1,399	879	5,315	90	12	8,518	194,988	211	1	18,028	292,218
UP³	825	122	3,248	84	3	86		14	98	24	5	3,232		1	706,744	714,486
Sub-total	1,149,627	41,981	4,285,605	72,040	9,455	120,405	887	43,287	7,758	10,095	15,878	3,201,518	625	5,888	1,086,571	10,051,620

¹Posterior probability threshold=0.99; ²MV for mixture of variants; ³UP for unpredictable variant; ⁴UA for unassigned variant

eTable 18. Concordance Analysis in the Prospective Set

Result from the concordance analysis of GISAID assigned variant (columns) with haplotype-based variant assignment (rows) in the prospective data set with 343,431 viruses, collected after March 14, 2022 and downloaded on May 18, 2022.

Prediction¹	Alpha	Delta	Epsilon	Eta	Lambda	Omicron	Zeta	UA⁴	Sub-total
Alpha	2					4			6
Delta		171				3			174
Epsilon			0			2		3	5
Eta				0				1	1
Lambda					1				1
Omicron						319,812		567	320,379
Zeta							0	1	1
MV²		5				22,072		30	22,107
UP³		4				1,699		524	2,227
Sub-total	2	180		1	343,592		1,126	344,901	

eTable 19. Four Mixture Variants with Omicron-Delta

Four recombinants of Omicron with A) Delta, B) Alpha, C) Zeta and D) Epsilon variants, in which all observed recombinants include polymutants in their respective variant-specific core haplotypes and indicator of 1 and 2 corresponded to the variant (1) recombinant with Omicron (2)

eTable 20. Four Mixture Variants with Omicron-Alpha

Four recombinants of Omicron with A) Delta, B) Alpha, C) Zeta and D) Epsilon variants, in which all observed recombinants include polymutants in their respective variant-specific core haplotypes and indicator of 1 and 2 corresponded to the variant (1) recombinant with Omicron (2)

eTable 21. Four Mixture Variants with Omicron-Zeta

Four recombinants of Omicron with A) Delta, B) Alpha, C) Zeta and D) Epsilon variants, in which all observed recombinants include polymutants in their respective variant-specific core haplotypes and indicator of 1 and 2 corresponded to the variant (1) recombinant with Omicron (2)

Epsilon-Omicron	Freq	NSP1	NSP2	NSP3	NSP4	NSP5	NSP6	NSP7	NSP8	NSP9	NSP10	NSP11	NSP12	NSP13
Indicator		S135	T85	K38	G489	L1266	N38	A1892	N392	A142	T527	L264	N384	N394
1	96 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . H I I . E T L K R . R													
2	94 R I . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R . R													
3	88 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R R													
4	8 R I . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R . R													
5	8 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R . R													
6	7 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R R													
7	6 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N . . N K A R . R Y H . Y K H K Y . H K . H I I . E T L K R . R													
8	6 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R R													
9	6 R . I . S . . F I F I H . . C V I I S . . D . G . . N S N K . . N K A R . R Y H . Y K H K Y . H K . H I I . E T L K R . R													
10	6 R . I . S . . F I F I H . V C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R . R													
11	5 R I . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . . . Y K H K Y . H K . . I I . E T L K R . R													
12	5 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . . . Y K H K Y . H K . . I I . E T L K R R													
13	4 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . H I I . Z T L K R . R													
14	4 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . . I I . E T . K R R													
15	4 . I . R . I T . . . I H V . . V . . V I D I . D K L P F . . N K S . N K A R S R Y H K Y K H K Y K H K F . . I G E T L K R . .													
16	3 R I . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S . K . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R . R													
17	3 R I . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S . . N K A R . R Y H . Y K H K Y . H K . . I I . E T L K R . R													
18	3 R I . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S N Y K H K Y . H K . . I I . E T L K R . R													
19	3 R . I . S . . F I F I H . . C V I I . . D . G D . F P F A N S N K . . N K A R . R Y H . Y K H K Y . H K . H I I . E T L K R . R													
20	3 R . I . S . . F I F I H . . C V I I S . . D . G D . F P F A N S . . N K A R . R Y H . Y K H K Y . H K . H I I . E T L K R . R													

eTable 22. Four Mixture Variants with Omicron-Epsilon
Geo- and temporal distribution of Epsilon-Omicron recombinants in the world.

Collection Date	Africa/Mauritius	Africa/South Africa	Asia/Israel	Asia/Japan	Asia/Malaysia	Asia/Singapore	Asia/South Korea	Asia/Thailand	Europe/Austria	Europe/Belgium	Europe/Croatia	Europe/Czech Republic	Europe/Denmark	Europe/Estonia	Europe/Finland	Europe/France	Europe/Germany	Europe/Ireland	Europe/Italy	Europe/Lithuania	Europe/Luxembourg	Europe/Netherlands	Europe/Norway	Europe/Poland	Europe/Portugal	Europe/Slovakia	Europe/Slovenia	Europe/Spain	Europe/Sweden	Europe/Switzerland	Europe/Turkey	Europe/United Kingdom	North America/Canada	North America/Costa Rica	North America/USA	Oceania/Australia	Oceania/New Zealand	South America/Brazil	South America/Chile	South America/Colombia	Total
2022-03-15								1																										6							
2022-03-16	1										1																						8								
2022-03-17	1										1							1	1		2			1	1								8								
2022-03-18											1	1								1	1		3	1	2							10									
2022-03-19											2	1											1		1	1						6									
2022-03-20																																1									
2022-03-21												2									1		3	1	1						8										
2022-03-22											1	1	1	4							1		2	3							15										
2022-03-23		1																		2			3	2							8										
2022-03-24		1																1	2				5	3	1						13										
2022-03-25												1											1		1						3										
2022-03-26	3	3																													6										
2022-03-27	1										1												1								3										
2022-03-28											1	1	1	1						1	1		1	1		1				9											
2022-03-29	2										1			2						1			1	3						10											
2022-03-30											1		1							2										5											
2022-03-31											1	1		3						2			1							8											
2022-04-01											1											3		1		1				6											
2022-04-02	1	1										1										3	1							7											
2022-04-03												1								1		2	2						6												
2022-04-04	1										2	1											1		1		1			6											
2022-04-05											1			1						1		1	1	1	1	1	1	1	8												
2022-04-06	1										3									1	1	7	2						15												
2022-04-07	1										1			1	1	1				1	3	1	1						11												
2022-04-08	2										1									1		4								8											
2022-04-09												1									4	2	3						10												
2022-04-10		1									1										3		1							9											
2022-04-11		2									1	6	5	3		1	2	1				7	1	2	1					31											
2022-04-12											2	2	3		1	1			1			5	1	1					17												
2022-04-13	1	1									2		2							1		3	1	1	1				14												
2022-04-14												3		6					1	1		7	1	4					23												
2022-04-15												4										1	1	4						10											
2022-04-16	1										1	1									2	1	3	1						15											
2022-04-17								1				1								1		3	2							8											
2022-04-18	1											4									9	1	3	2						20											
2022-04-19	1										8	1	3			1	1				7	1	3	1	1				28												
2022-04-20								1				1								2	3	1	1						10												
2022-04-21	1										2	3	1	1						1		3	1	5					18												
2022-04-22	2	1									1	2	4							2		8	5						26												
2022-04-23	1										1	1	2	1							5	4							15												
2022-04-24											1		3									1	1	7						14											
2022-04-25											1	2	6				1				2	2	7						23												
2022-04-26											2	1	3		1	1				1		9	14						32												
2022-04-27	1										1		1								6	2	4						16												
2022-04-28												5								1		8	5						19												
2022-04-29												1									5	1	8						15												
2022-04-30																					1		3	1						5											
2022-05-01		1		1																	1	1	2							6											
2022-05-02		2	1									1									1		3							8											
2022-05-03												1								3		4	3						11												
2022-05-04											1		1	2							2									6											
2022-05-05												1																		1											
2022-05-06												1										2								3											
2022-05-08												1																		1											
2022-05-09												1																		1											
total	1	3	13	6	4	1	4	5	1	13	6	12	33	1	2	19	71	7	7	1	4	14	6	4	4	2	2	18	13	1	1	158	31	1	124	8	3	1	3	1	609

eTable 23. Mixture Variant and Corresponding Lineages
 Relationship of mixture variants with their assigned lineages

Lineage	Omicron-Delta	Omicron-Epsilon	Omicron-Alpha	Omicron-Zeta	Alpha-Epsilon	Delta-Kappa	Delta-Zeta	total
AY.25					1			1
AY.38						1		1
AY.4.14				1				1
AY.4.15					1			1
AY.45						1		1
BA.1	1	1	3	1				6
BA.1.1		30	2	2				34
BA.1.1.1		6						6
BA.1.1.14		1						1
BA.1.1.15		1						1
BA.1.1.2		5						5
BA.1.14		3						3
BA.1.15		1						1
BA.1.17		1						1
BA.1.17.2		1						1
BA.1.18		2						2
BA.1.19		1						1
BA.2	2	354	15	5	6			382
BA.2.1		5						5
BA.2.10		5		1	1			7
BA.2.12		4						4
BA.2.15		2						2
BA.2.2		1						1
BA.2.20		3						3
BA.2.21		1						1
BA.2.22		3						3
BA.2.23		4						4
BA.2.24		1						1
BA.2.26		1						1
BA.2.3		88		3				91
BA.2.3.4		9						9
BA.2.4		1						1
BA.2.7		1						1
BA.2.8		1						1
BA.2.9		71	5					76
BA.4		1						1
BA.5			1					1
total	3	609	25	10	10	3	2	662

eTable 24. Unidentifiable Omicron Viruses and Corresponding Lineages

HAI-unpredictable Omicron viruses and their assigned lineages

	O1	O2	O3	O4	total
BA.1	22	5	55	3	85
BA.1.1	3		54	1	58
BA.1.1.1			4		4
BA.1.1.13			1		1
BA.1.1.18			1		1
BA.1.15			2		2
BA.1.15.2			1		1
BA.1.17			5		5
BA.1.20			1		1
BA.1.9			1		1
BA.2	510	253		323	1086
BA.2.10	9	5			14
BA.2.12	23	13			36
BA.2.12.1	18	6			24
BA.2.18	2				2
BA.2.19		1			1
BA.2.23	35	9		3	47
BA.2.3	18	19		1	38
BA.2.3.2		3			3
BA.2.3.3	1				1
BA.2.31	4	2			6
BA.2.32	1				1
BA.2.5	1	1			2
BA.2.6	2				2
BA.2.7	2				2
BA.2.8		2			2
BA.2.9	27	16		9	52
BA.3	1	1			2
BA.3.1	4				4
BA.4	1				1
BA.5	3				3
Unassigned	131	62	9		202
XE	8	1			9
total	826	399	134	340	1699

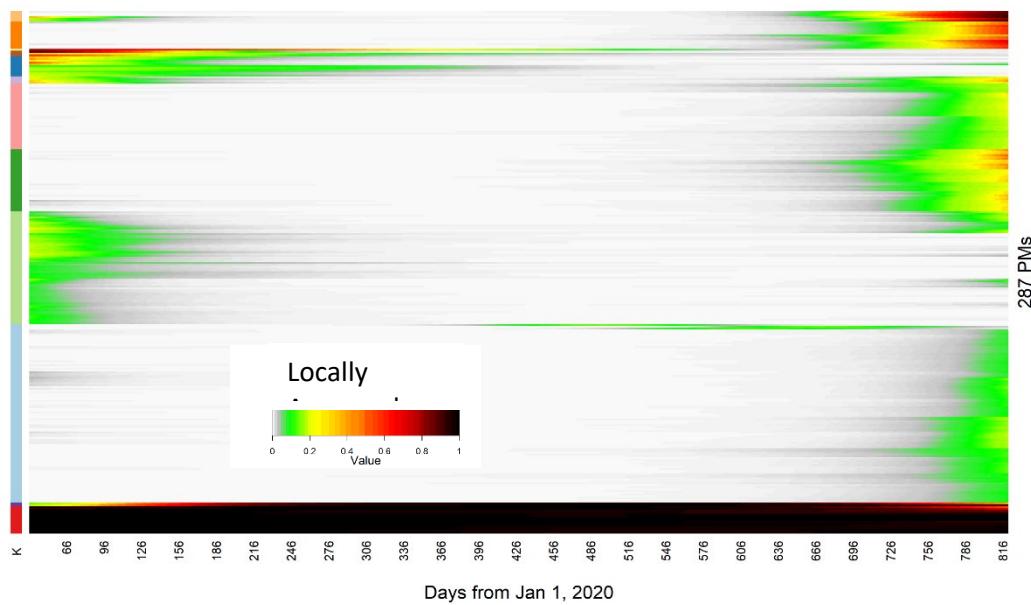
eTable 25. New Mutations Among Variant-Unassigned and Unidentifiable Viruses

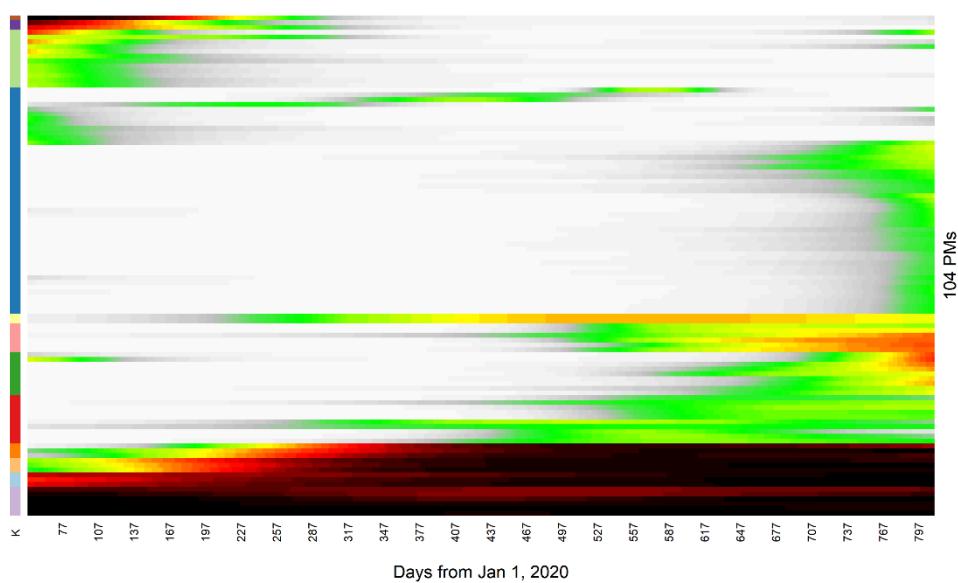
Ab initio mutations (excluding all core polymutants of 14 known variants) have been found to significant temporal trend

Polymutant	cluster	p-vlaue	LAMP	LAMP-max
1 NSP1_F143	O1	3.43E-02	2.25E-02	7.62E-02
2 NSP1_K141	O1	3.43E-02	2.25E-02	7.62E-02
3 NSP1_S142	O1	3.43E-02	2.25E-02	7.62E-02
4 Spike_A684	O1	1.94E-02	2.00E-02	9.05E-02
5 Spike_I68	O1	4.10E-02	2.09E-02	1.22E-01
6 NSP2_F356	O2	1.04E-02	2.34E-02	1.09E-01
7 NSP6_L105	O2	2.76E-04	8.18E-02	2.09E-01
8 NS3_H78	O3	3.56E-02	9.53E-02	1.67E-01
9 Spike_L24	O4	3.12E-05	3.30E-01	5.33E-01
10 Spike_P25	O4	1.66E-05	3.32E-01	5.41E-01
11 NSP6_F108	O5	2.09E-06	4.42E-01	6.02E-01
12 NSP6_G107	O5	4.32E-05	5.20E-01	6.70E-01
13 NSP6_S106	O5	6.73E-05	5.14E-01	6.61E-01
14 N_E31	O6	1.99E-02	6.52E-01	7.15E-01
15 N_R32	O6	1.70E-02	6.39E-01	7.00E-01
16 N_S33	O6	1.03E-02	6.49E-01	7.18E-01

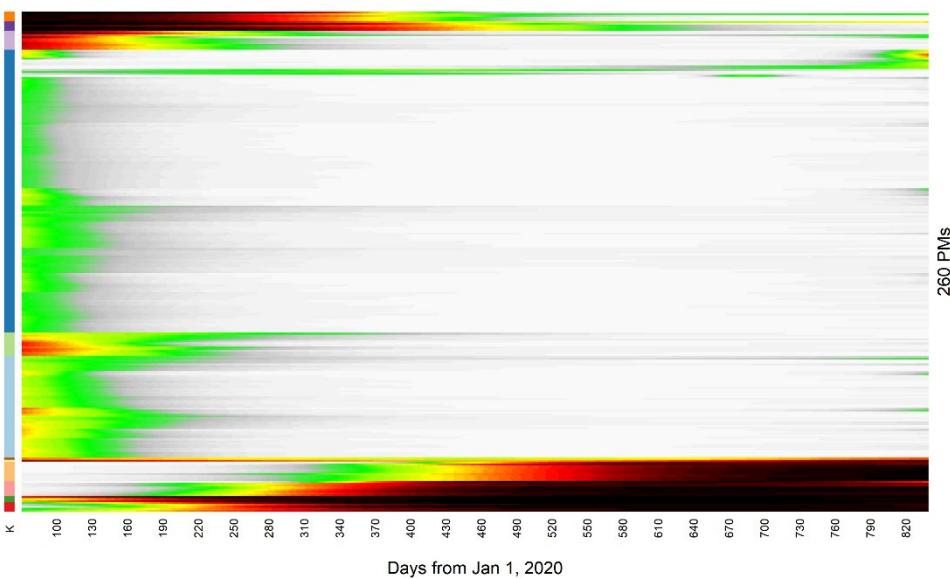
eFigure 1. Heatmap-representation of selected polymutant temporal profile from January 1, 2020 to March 14, 2022 within every variant: A) Alpha, B) Beta, C) Delta, D) Epsilon, E) Eta, F) Gamma, G) GH/490R, H) Iota, I) Kappa, J) Lambda, K) Mu L) Omicron, M) Theta, N) Zeta and O) variant-unassigned viruses. In each profile, estimated locally-averaged mutation percentages (LAMP) take value from 0 to 1, and are color-coded (the legend), and each row corresponds to a polymutant while each column corresponds to collection date.

A)

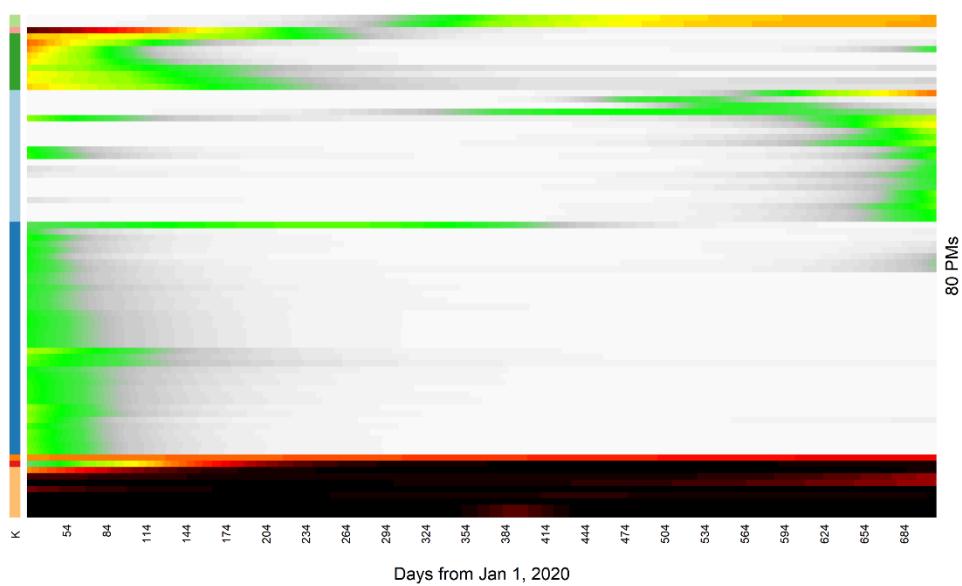


B) Beta

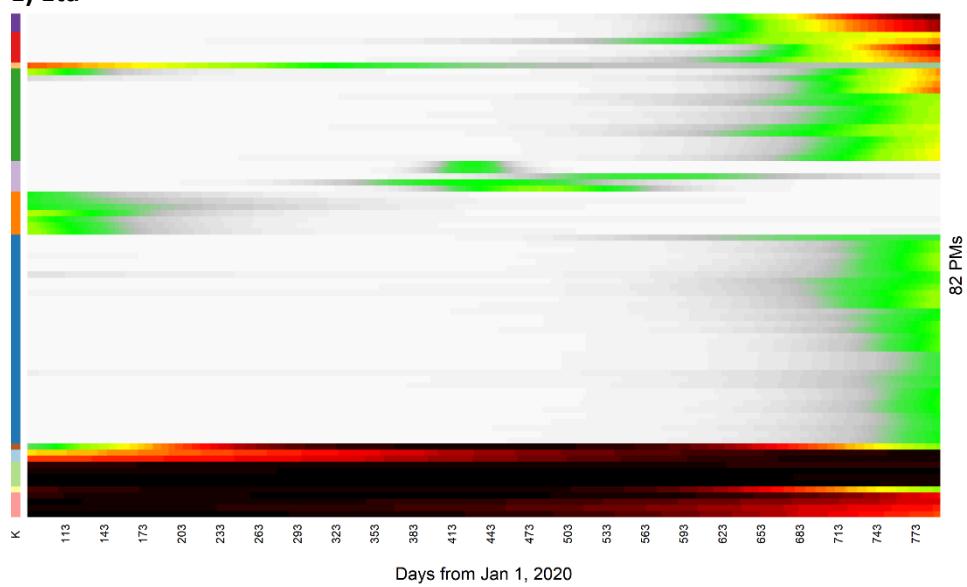
C) Delta



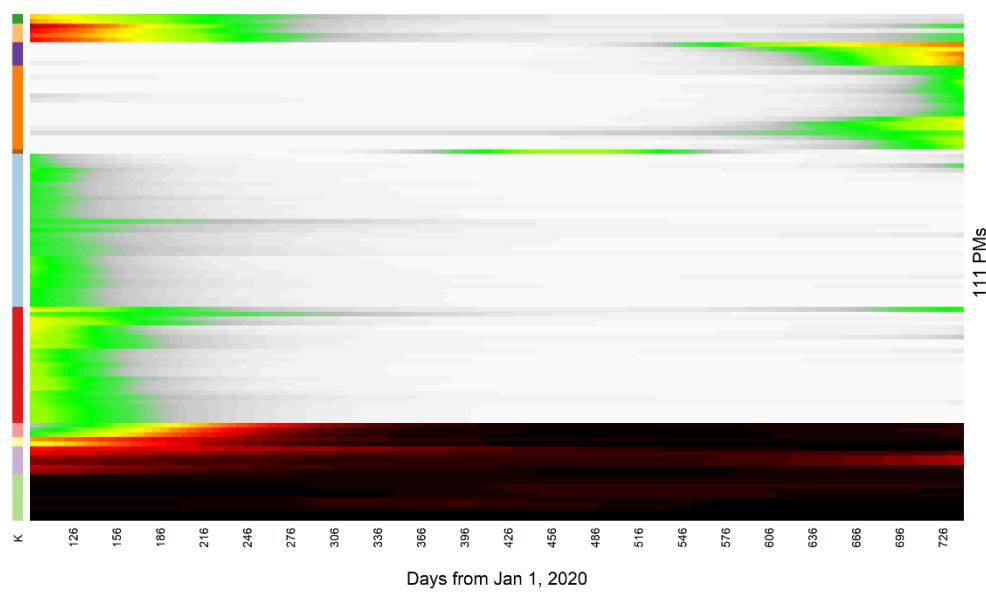
D) Epsilon



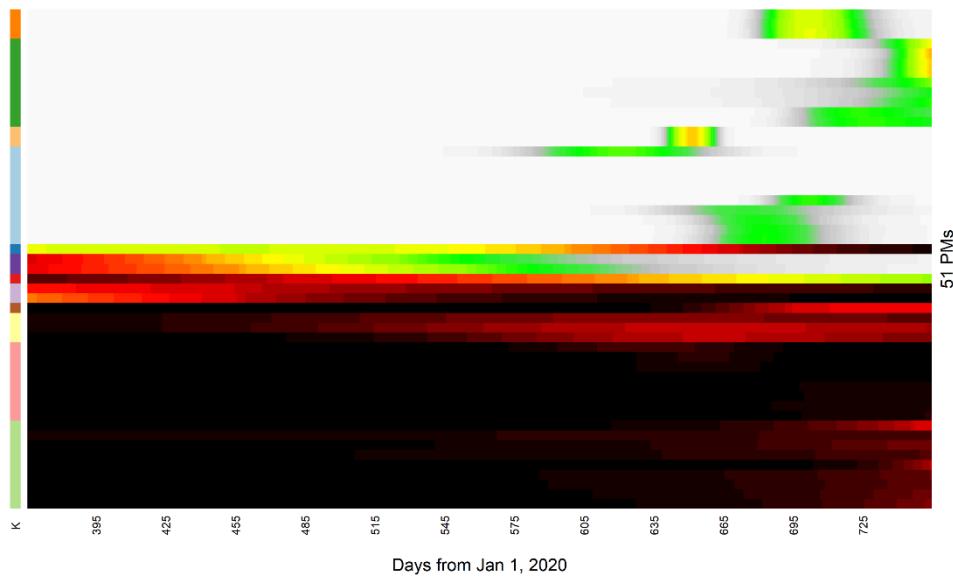
E) Eta



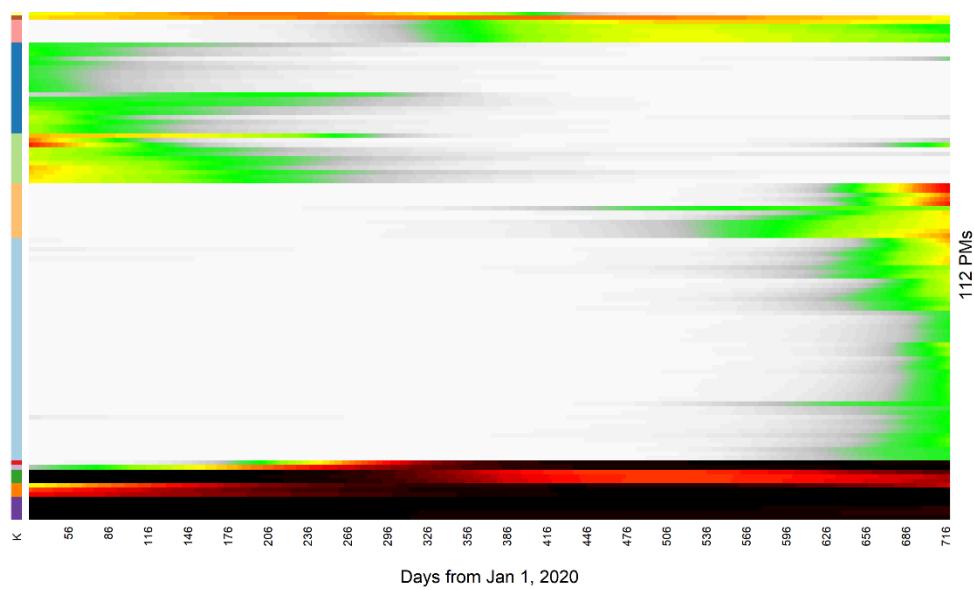
F) Gamma



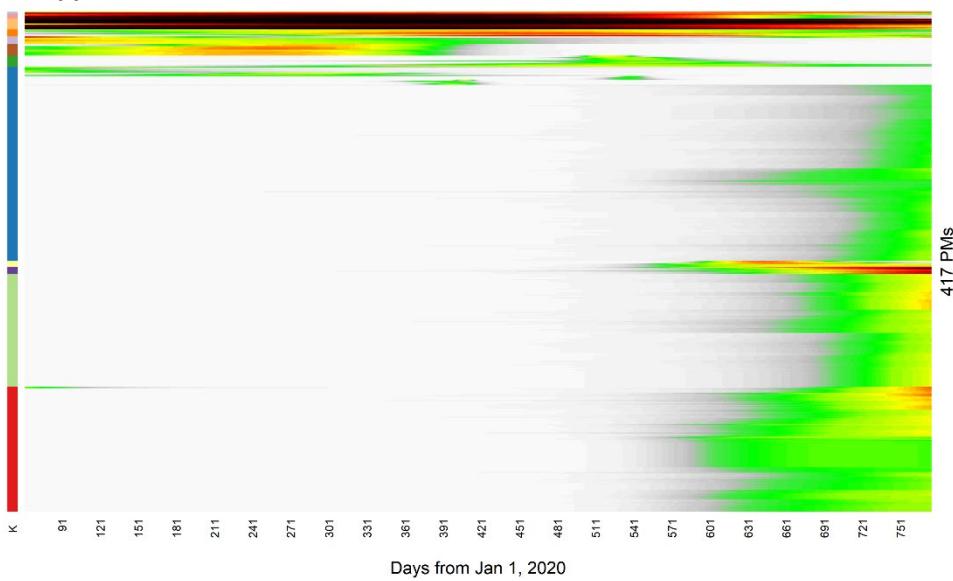
G) GH/450R



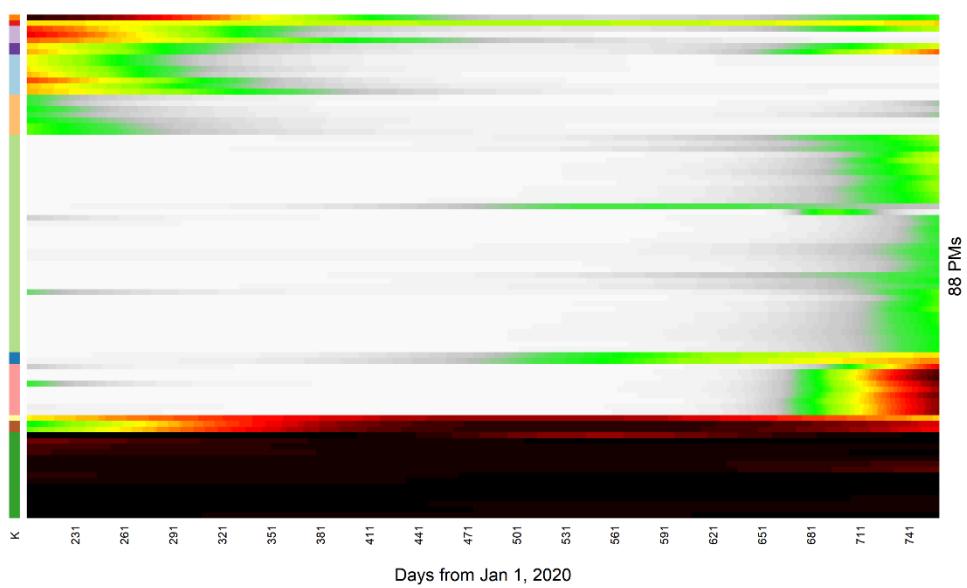
H) Iota



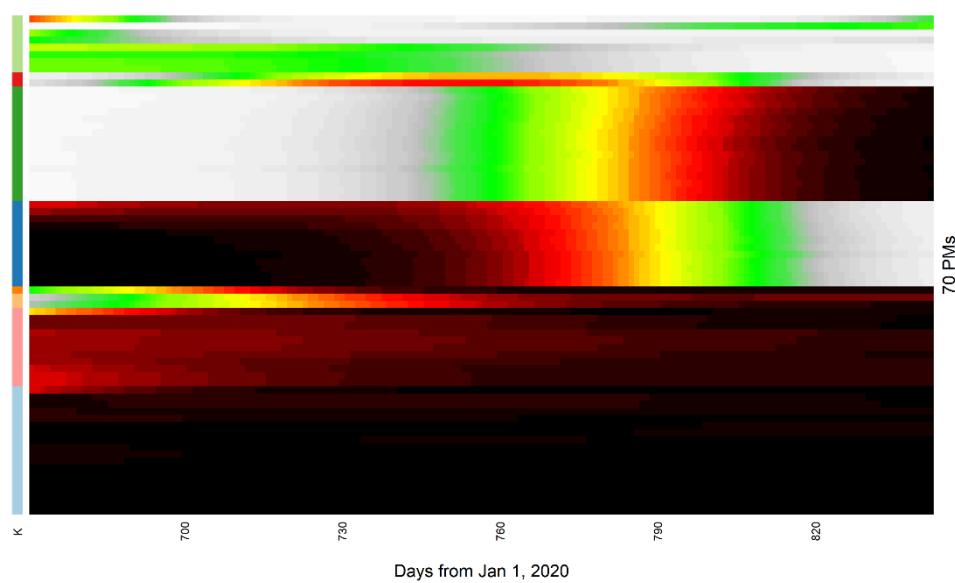
I) Kappa



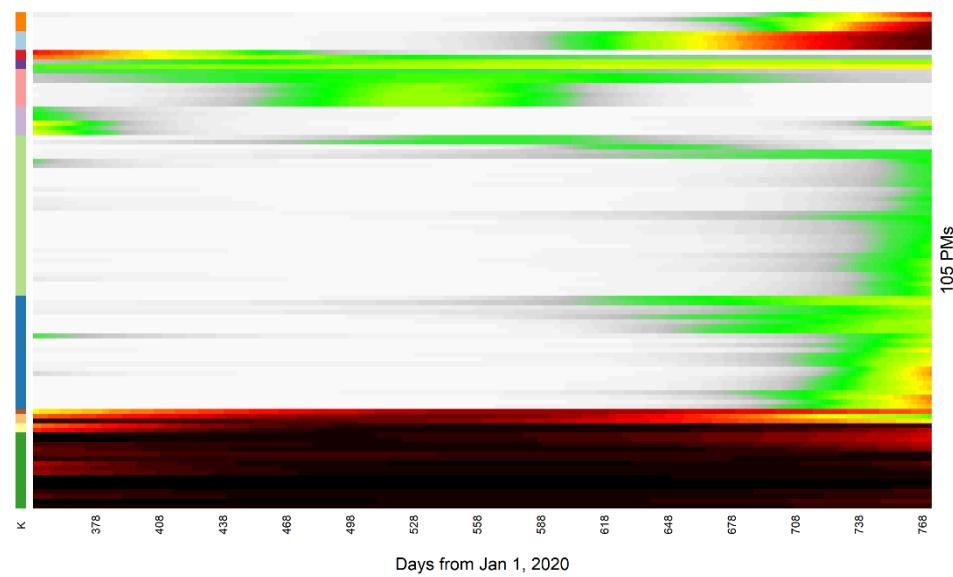
J) Lambda



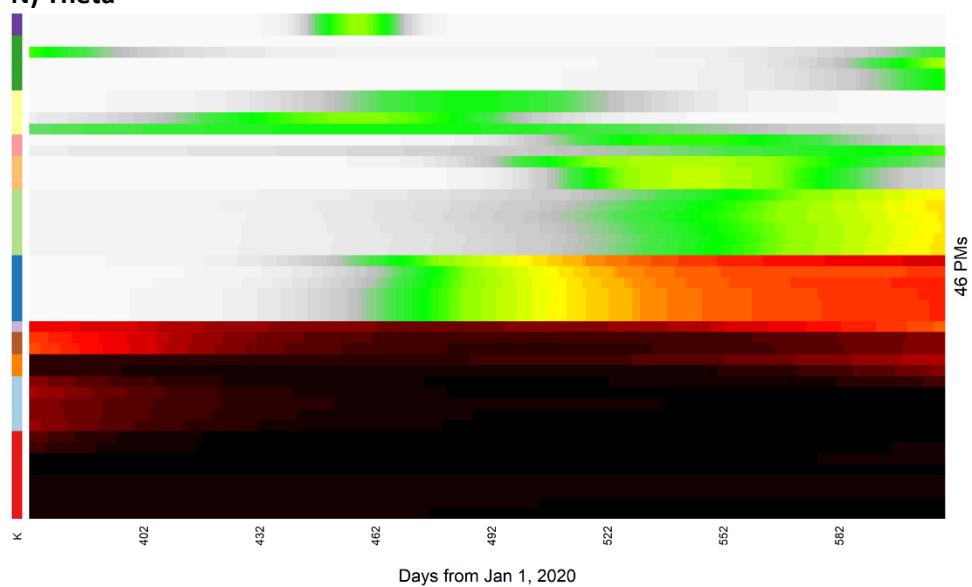
L) Omicron



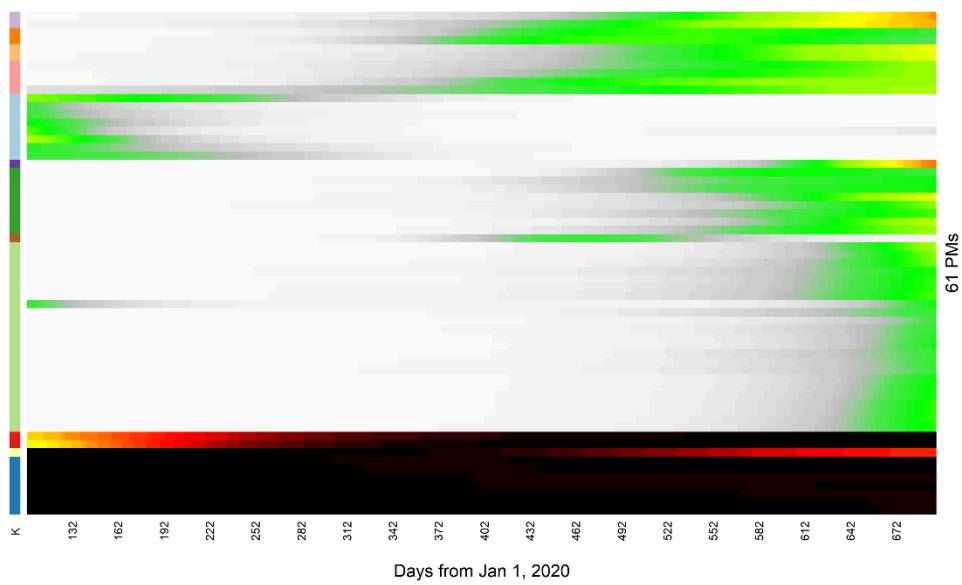
M) Mu



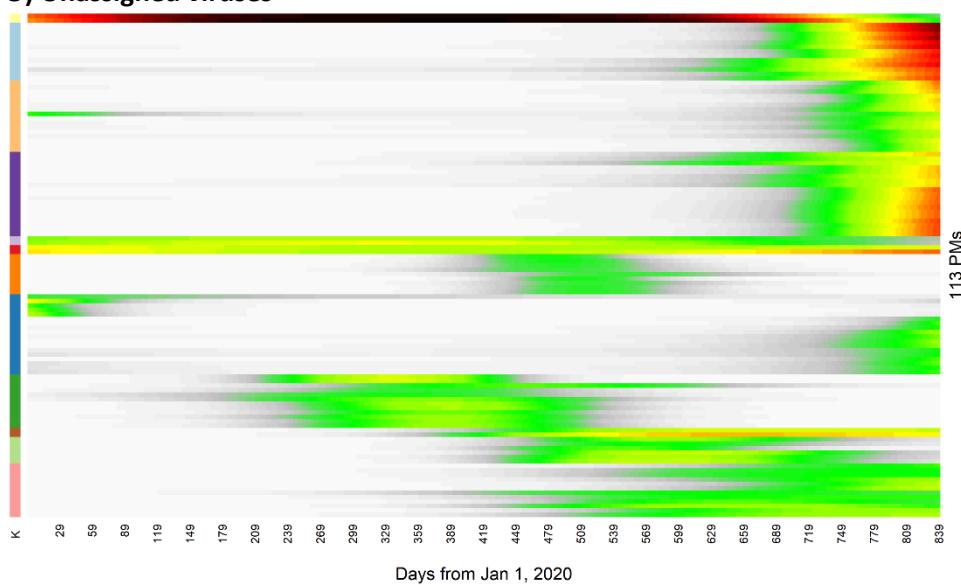
N) Theta



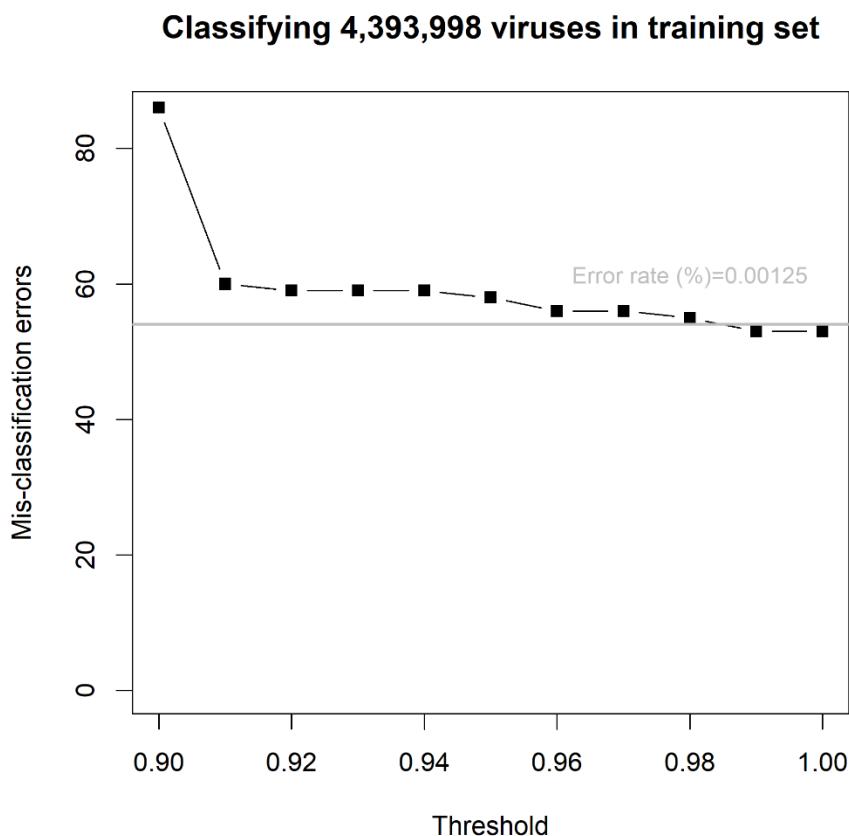
N) Zeta



O) Unassigned viruses



eFigure 2. Misclassification errors by haplotype-based variant prediction (HVP), when the prediction probability threshold value is set at 0.9 to 1



eFigure 3. Temporal patterns of sixteen polymutants identified from variant-unassigned 524 viruses that are unpredictable by HAI, excluding those core polymutants of all fourteen variants.

