Transitions Omicron subvariant and amplicon dropouts

Update 2022/04/02

Bette Korber, Will Fischer, Hyejin Yoon, James Theiler

Contents:

Los Alamos NATIONAL LABORATORY

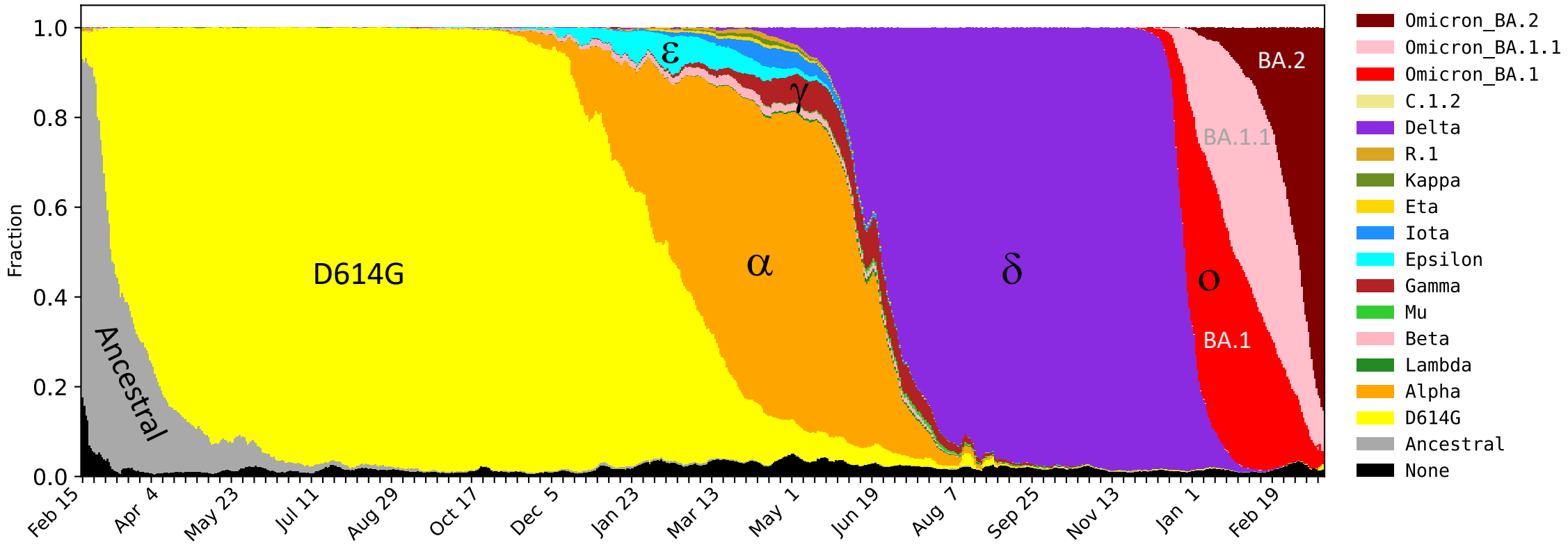
GISAID

LA-UR-21-28226

Transitions in major global lineages between: 2020/02/15 though 2022/03/30

Global: 9,234,832 sequences

For now, it is time to focus on BA.2



GISAID

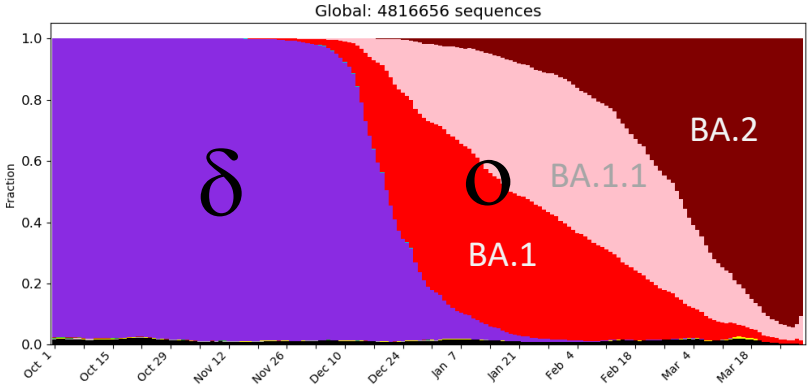
Los Alamos NATIONAL LABORATORY

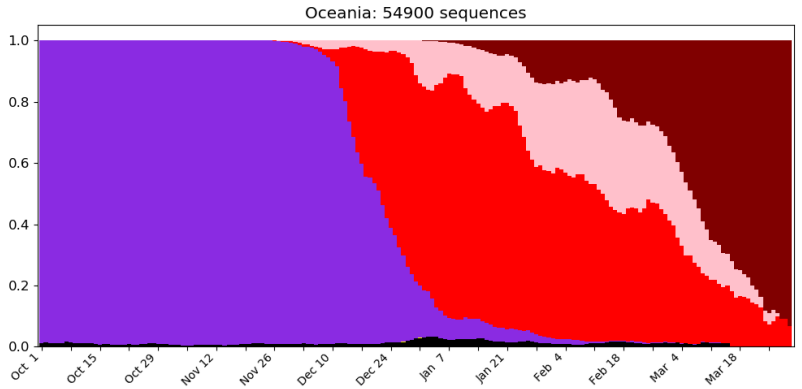
With thanks to those who share sequences through GISAID

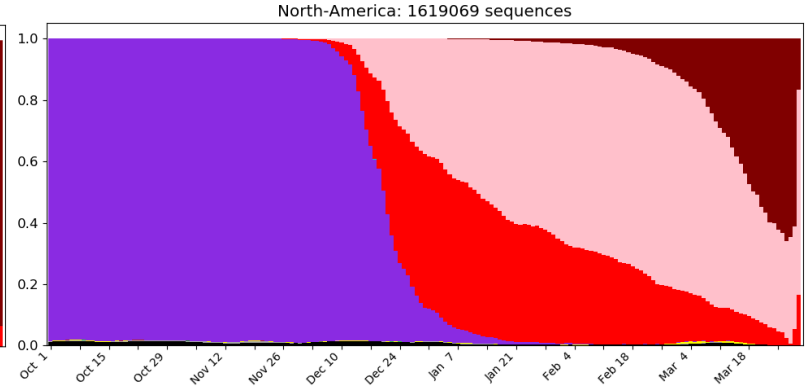
Omicron Transitions by continent: BA.2 is rapidly gaining ground, e Americas

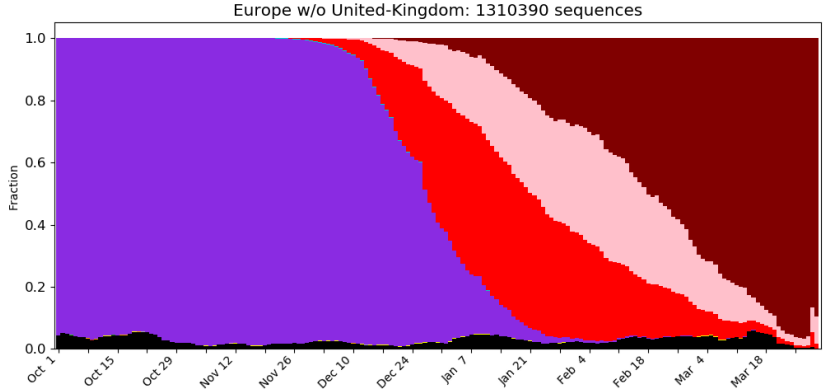
2021/10/01 – 2022/04/01

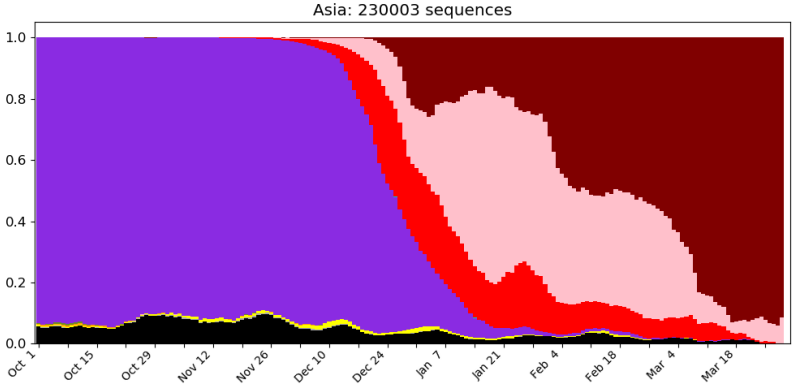
GLOBAL

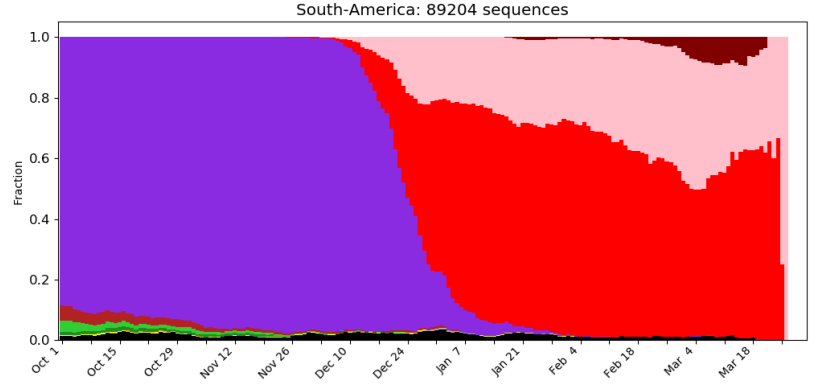


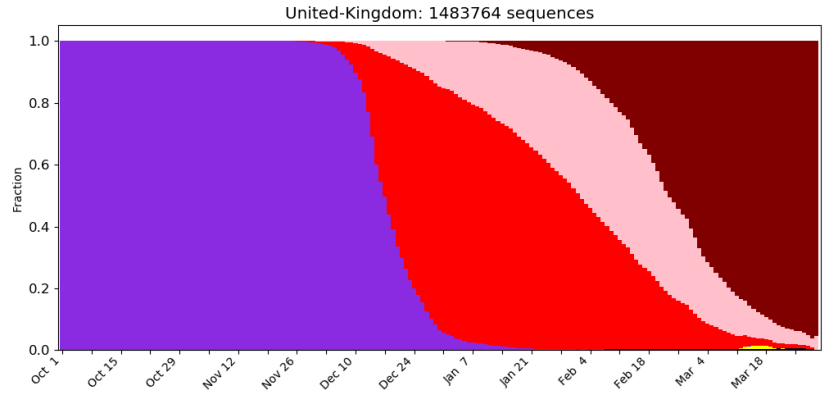


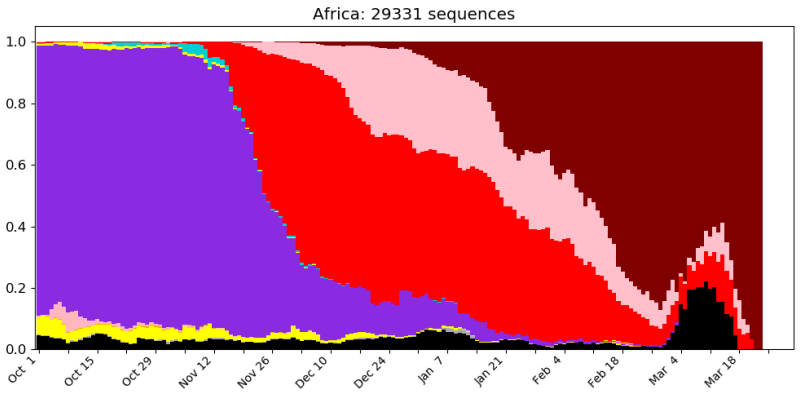












The transition to from BA.1 to BA.2 is essentially complete in most of the world, but began expanding later in the Americas. BA.2 is now >50% of the samples in N. America, and approaching 10% in S. American. The “nones” were often chimeric ancestral forms.

BA.2 Spike variants included in our current VOI spread sheet and alignments

	BA.2 variants	Variants relative to BA.2	Number of sequences that exactly match this pattern, on April 2, 2022, Limited to full QC'd data	Number of sequences that contain this pattern, on April 2, 2022, limited to full QC'd data	Typical example of EPI_ISL number for this pattern of substitutions
A	BA.2_Omicron_baseline	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K	267193	302831	EPI_ISL_11089722
B	BA.2_+212SGR	T19I,L24S,P25-,P26-,A27-,G142D,**+212SGR**,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K	537	578	EPI_ISL_11028691
C	BA.2_E484V	T19I,L24S,P25-,P26-,A27-,G142.,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,**E484V**,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K	13	15	EPI_ISL_11262104
D	BA.2_S704L	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,**S704L**,N764K,D796Y,Q954H,N969K]	875	1017	EPI_ISL_10858864
E	BA.2_S787H	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,**Q787H**,D796Y,Q954H,N969K	717	759	EPI_ISL_11152597
F	BA.2_A846G	T19I,L24-,P25-,P26-,A27S,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,**A846G**,Q954H,N969K	810	907	EPI_ISL_10096403
G	BA.2_P1162L	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K,**P1162L**	1115	1157	EPI_ISL_10966824
H	BA.2_P1162S	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K,**P1162S**	245	262	EPI_ISL_9850727
I	BA.2_I1221T	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K,**I1221T**	740	891	EPI_ISL_10915570
J	BA.2_M1229I	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K,**M1229I**	602	610	EPI_ISL_10200123
	Chimera or reversions to ancestral;				
K	BA.2_L24_P25_P26_A27_	T19I, G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K	14312	319168	EPI_ISL_8707013
L	BA.2_N440_	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N, S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K	4779	308168	EPI_ISL_10006638
M	BA.2_R408_	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N, K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K	1116	304120	EPI_ISL_9772288
N	BA.2xBA.1	T19I,L24S,P25-,P26-,A27-,A67V,H69-,V70-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K	55	98	EPI_ISL_10991177
O	BA.2xBA.1.11_	T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,R346K,S371L,S373P,S375F,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K	53	70	EPI_ISL_11159397
P	BA.1xBA.2	A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F	81	216	EPI_ISL_10443647

Row A shows the Omicron BA.2 baseline consensus and most common form.

Rows B-J are common mutations on a BA.2 backbone, with the exception of row C, “BA.2_E484V” which was not common but still interesting. Position 484 is a key position for RBD antibody escape, is usually E484A in Omicron, and the E484V was common in both BA.1 and BA.1.1. 484V is the most common mutation in this position in BA.2, though still very rare.

Rows K-M are reversions to ancestral forms. Row K is currently the most common BA.2 Spike sublineage variant, and a 9-base deletion which impacts Spike amino acids 24-27 is reverted to the ancestral. Rows N-P show some of the more common, though still rare, forms of BA.2/BA.1 or BA.1.1 chimera.

Recombination and chimeric sequences

I. Recombination is an evolutionary mechanism used by coronaviruses

- It is possible that recombination between two distinct SARS-CoV-2 variants (e.g. Omicron and Delta) may occur within a host that is naturally co-infected. Thus, a natural recombinant form with selective advantage may be able expand in SARS-CoV-2.

II. There are two ways apparent recombination can arise in the laboratory

- If two variants are present in a sample, recombination can occur during PCR amplification, giving rise to recombinant sequences generated in vitro.
- If the dominant variant in a sample has a primer mismatch, a rare variant or low-level contaminant in the sample may be preferentially amplified, giving rise to a chimeric sequence that is an apparent recombinant.

III. It is important be aware of chimeric sequences in subsequent analyses.

- Chimeric sequences can impact conclusions based on phylogenetic analyses. Sometimes they will yield particularly long branches within a clade, sometimes they will form distinctive branches between the two clades representing the lineages from which the parents were derived. Either artifact can impact conclusions such as timing the origins of a lineage and tree-based estimates of positive selection.
- Common chimeric sequence forms could result in noisy estimates of mutational frequencies.
- If a natural recombination event was confirmed (for example, its sequence was confirmed and/or a recombinant lineage began to be transmitted and resampled in multiple geographic regions) it would be an interesting event in its own right, but could also impact phylogenetic analyses that assume no recombination.

Many chimeric sequences are evident among Omicron variants. These include (i) chimeric stretches of either ancestral or Delta sequence in an Omicron backbone, (ii) Omicron BA.1 in a BA.2 backbone, or (iii) BA.1 chimeric stretches in Delta backbones. While many these chimeric sequences are likely to be *in vitro* artifacts, some reflect an actual *in vivo* recombination event

Amplicon Dropout and SARS-CoV-2 sequences

As new variants arise among SARS-CoV-2 sequences, the mutations they carry can disrupt primer interactions and leave some parts of the full genome sequences unresolved; this situation will persist until new primers with superior performance can be adopted and incorporated into sequencing protocols. These adjustments will naturally have different paces in different laboratories.

The unresolved sections are generally filled in with N's.

To enable reporting sequences in the critical weeks just as a new highly transmissible variant is detected and found to be expanding, in the earliest days of a new expansion many sequences will have long stretches of N's. As people adapt their primer sets and incorporate sequencing strategies that better capture these regions, more complete forms of the variants become increasingly available over time.

The next three slides (8-10) illustrate how this scenario has played out with Omicron sequences.

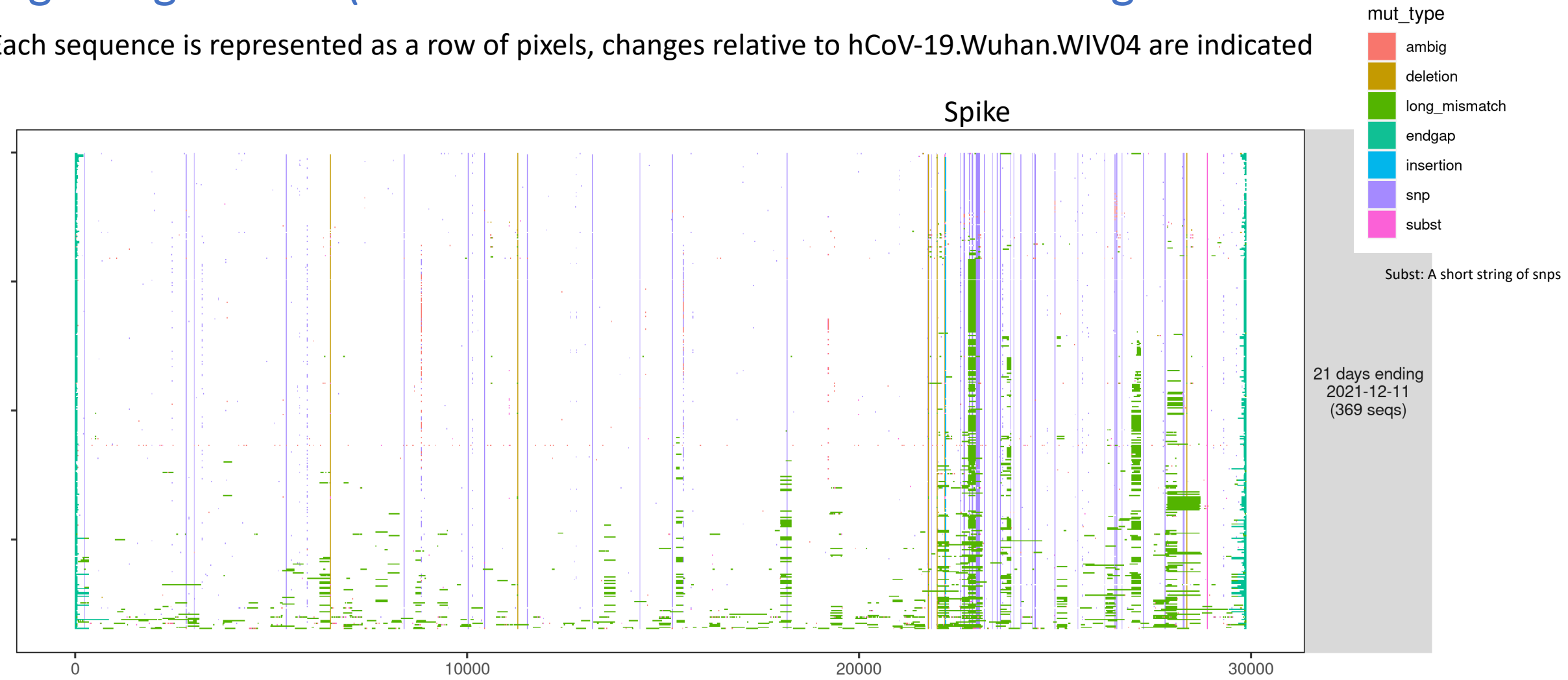
Omicron carried so many mutations relative to earlier variants that the issue of amplicon dropout was exacerbated. However, this problem is not unique to omicron, and most variants presented similar initial sequencing challenges as they first arose in the past.

Implications of Amplicon Dropout for interpreting sequences:

1) For experimentalist scientists ordering reagents to explore the immunological and virological impact of mutational patterns carried by new variants of interest and concern, care must be taken to make sure that the sequence they are using actually reflects the common circulating form of the variant. Some bioinformatics sites have treated "N"s as ancestral states, and this can lead to an under-representation of amino acid changes in common amplicon dropout regions in newly emerging lineages. We have taken *great care* to avoid this problem in the variant representative sequences included in these LANL VOI/VOC representative sequence sets.
2) Amplicon dropout can result in artificial chimeric sequences, where a rare variant in a sample (e.g. a low level co-circulating form in vivo or low-level laboratory contaminant) is successfully amplified by primers that miss the targeted sequence. Such "chimeras" will include stretches of a distinct form embedded within the sequence. Examples of stretches of ancestral or Delta sequences within Omicron sequences can be readily identified, as the Omicron Spike is so distinctive; some examples are included in our fasta files and spread sheet, and shown on slides

The first 3 weeks of Omicron sequencing and amplicon dropouts: long strings of “N” (sometimes occasional bases intermingled with them)

Each sequence is represented as a row of pixels, changes relative to hCoV-19.Wuhan.WIV04 are indicated



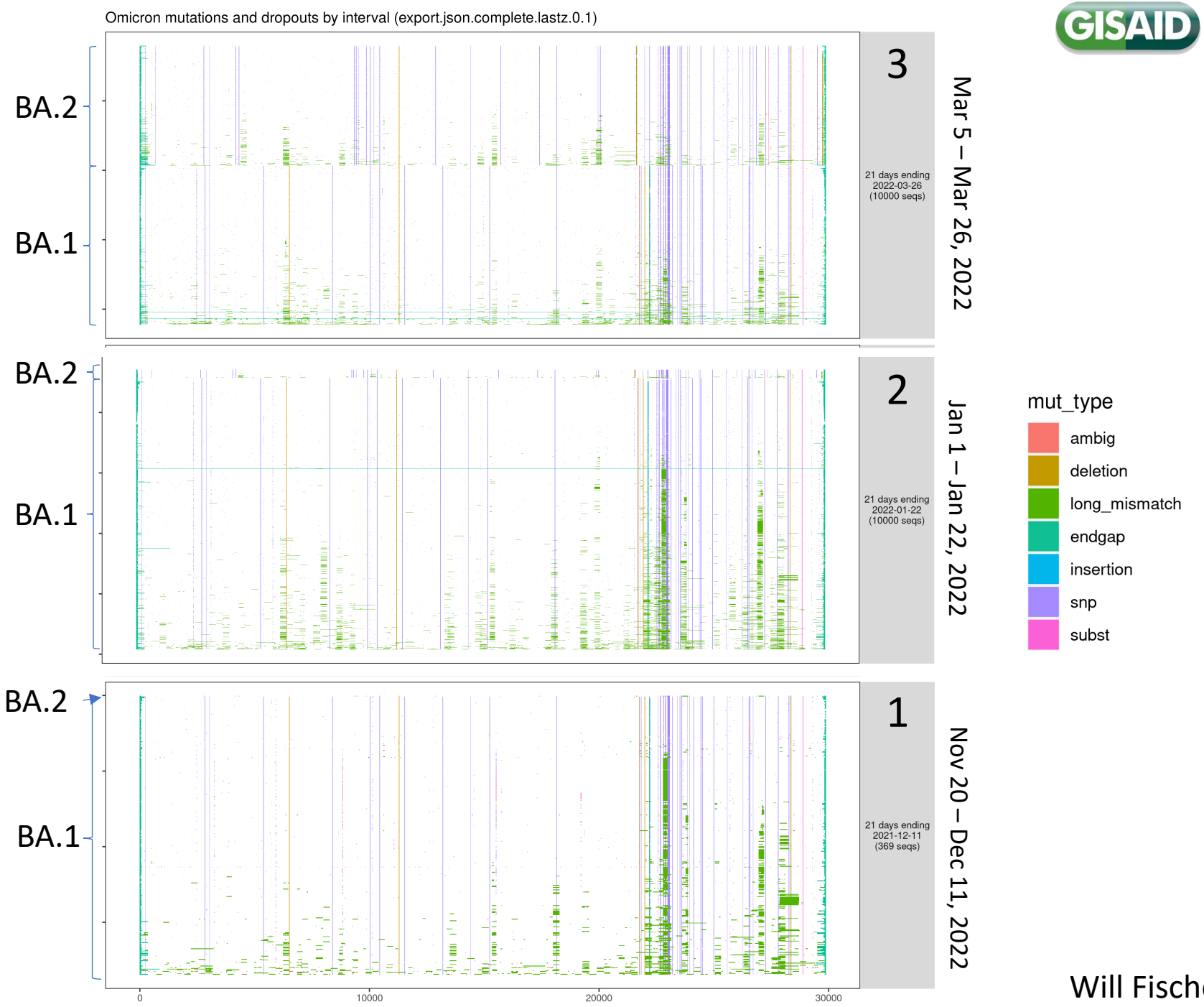
Green indicates stretches of N’s in each sequence (long_mismatch), the common vertical patterns are amplicon dropouts
N’s can more common than base calls in the RBD, and it is critical for bioinformatics groups not to assume these are ancestral.

Amplicon Dropout Regions in Omicron sequences over time, full genome sequences grouped by 3 weekperiods, and subsampled to 10,000 sequences.

Even the earliest complete sequences from November accurately captured the pattern of SNPS/indels common to the consensus forms of the Omicron lineage (also true of Beta, Delta, Gamma, Mu…).

“Green towers” diminish over time, as people improve sequencing strategies.

Of note: BA.2 was there from the earliest days, but had a delayed expansion.

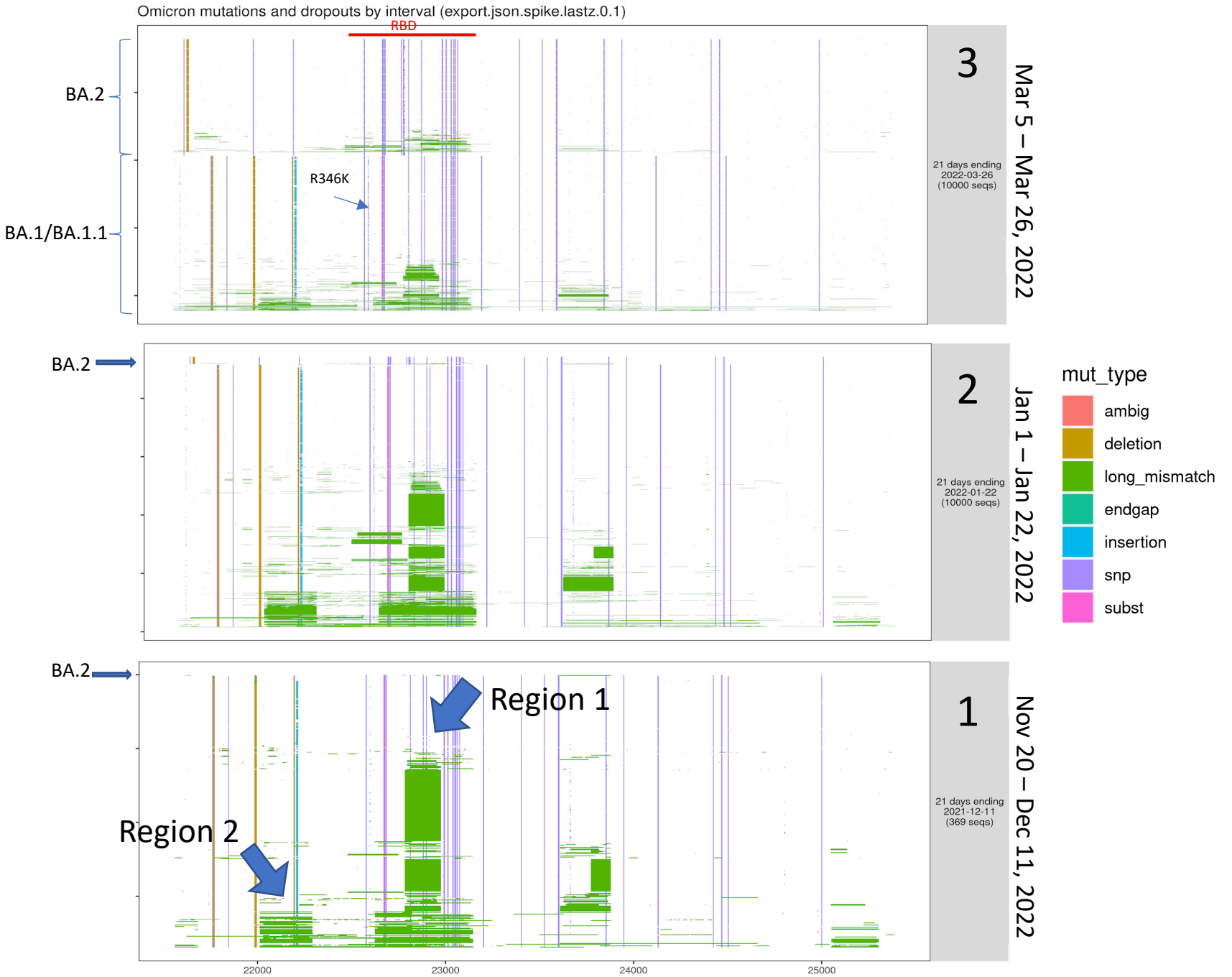


Will Fischer

Amplicon drop out in a Spike gene expansion

Note the frequent Amplicon dropout regions in the RBD, and how the coverage frequency improves over time.

The drop out regions are also often associated with a chimeric stretches (detailed on the next slide). The chimeras are sequences with a strings of either ancestral base calls, or delta mutations.



SARS-CoV-2 *spike* gene region expansion, positions relative to the full genome

GISAID

Los Alamos NATIONAL LABORATORY

BA.1 Chimeric sequences found >50 times among the 243,069 BA.1 sequences – 60 days prior to the date show below:

Counts: 2022-02-14

Region 2

Region 1

Counts: 2022-03-29

Count	perc	HD	Mutation string relative to the ancestral form:	
134941	55%	0	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	**BA.1**
*5918	2.4%	3	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I, G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	1728
*1738	0.7%	3	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F, S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	653
*1313	0.5%	5	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-, G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	223
828	0.3%	4	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F, S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H, D796Y,N856K,Q954H,N969K,L981F]	
567	0.2%	20	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
395	0.2%	7	[A67V, T95I,G142D,V143-,Y144-,Y145-, G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
330	0.1%	2	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-, +214EPE,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
298	0.1%	6	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-, G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,**A701V**,N764K,D796Y,N856K,Q954H,N969K,L981F]	
226	0.1%	9	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F, N440K,G446S, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
172	0.1%	8	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F,K417N,N440K,G446S, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
143	0.1%	5	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
95	0.0%	11	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
75	0.0%	19	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-, G339D, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
68	0.0%	8	[A67V, T95I,G142D,V143-,Y144-,Y145-, G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,**A701V**,N764K,D796Y,N856K,Q954H,N969K,L981F]	

BA.1 chimeric forms found >10 times carrying **Delta signature mutations.** We lowered the threshold for inclusion to 10 as Delta and BA.1 were cocirculating, so Delta chimeras have a greater potential to be biologically interesting.

*196	0.1%	4	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F, **L452R**,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	1
*152	0.1%	13	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-, **L452R**,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	0
*12	0.0%	19	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE, **L452R**, **T478K**, **D614G**, **P681R**,N764K,D796Y,N856K,Q954H,N969K,L981F]	0
*14	0.0%	23	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,**E156G,F157-,R158-**, **L452R**, T478K, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	0
20	0.0%	14	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,**E156G,F157-,R158-**, G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	0
*29	0.0%	20	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-, **L452R**, **T478K**, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	0

BA.1.1 Chimeric sequences among 211797 BA.1.1 sequences

167320	79 %	0	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
3342	1.6%	3	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
1602	0.8%	4	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H, D796Y,N856K,Q954H,N969K,L981F]	
356	0.2%	5	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-, G339D,R346K,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
328	0.2%	9	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, N440K,G446S, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
280	0.1%	8	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F,K417N,N440K,G446S,T547K,D614G, H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
227	0.1%	5	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
129	0.1%	8	[A67V,H69-,V70-,T95I, N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H, D796Y,N856K,Q954H,N969K,L981F]	
104	0.0%	11	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
94	0.0%	6	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
90	0.0%	7	[A67V, T95I,G142D,V143-,Y144-,Y145-, G339D,R346K,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
85	0.0%	4	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K, N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
78	0.0%	7	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H, D796Y,N856K,Q954H,N969K,L981F]	
71	0.0%	6	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
70	0.0%	5	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, S477N,T478K, Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H, D796Y,N856K,Q954H,N969K,L981F]	
64	0.0%	5	[T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,V736G,N764K,D796Y,N856K,Q954H,N969K,L981F]	
58	0.0%	7	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, N440K,G446S,S477N,T478K, T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	
56	0.0%	4	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, S477N,T478K, Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	

BA.1.1 forms found > 10 times carrying **Delta signature mutations**:

140	0.1%	4	[A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,R346K,S371L,S373P,S375F, **L452R**,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	1

These were initially identified in the sequences sampled 60 days prior 2/14/2022; all are diminishing or disappearing in the last 60 days. This possibly is because sequencing issues are getting fixed, or because BA.2 now dominates the pandemic, or both. Region 1 and 2 are frequent amplicon drop out regions, noted on the previous slide.

Omicron fragments in Delta backbones 60 days prior to 2/14: All were rare

Delta-lineage baseline Spike mutations:

T19R,T95I,G142D,E156-,F157-,R158G,L452R,T478K,D614G,P681R,D950N (Note: A27S, found in several chimera's listed below is commonly found with the Delta lineage.)

Omicron BA.1 lineage Spike baseline mutations:

A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F

These are the forms of Delta with stretches of Omicron sequences that were found more than one time within a distinct Delta Pango lineage. Purple are stretches of Delta, red of Omicron.

Counts: 2022-02-14

Counts: 2022-03-29

Pango designation	N_Pango	N	perc	HD	[Spike mutation strings]	
AY.103	4244	2	0.0%	7	[T19R,G142D,E156G,F157-,R158-,N211-,L212I,+214EPE,L452R,T478K,Q613H,D614G,P681R,S691F,D950N]	0
AY.4	6536	2	0.0%	29	[T19R,A27S,T95I,G142D,E156G,F157-,R158-,N211-,L212I,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	3
B.1.617.2	2013	2	0.1%	4	[T19R,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	1
AY.126	1382	2	0.1%	6	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,I850L,D950N]	0
B.1.617.2	2013	2	0.1%	12	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,E554Q,D614G,H655Y,N679K,P681H,N856K,D950N,Q954H]	0
AY.100	1350	2	0.1%	6	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	
AY.39	818	2	0.2%	6	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	
AY.25	1152	2	0.2%	8	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	
AY.122	5732	6	0.1%	7	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	This exact from of Spike 25 times in 6 AY Delta Pango sublineages 1
AY.3	1434	6	0.4%	7	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	
AY.4	6536	7	0.1%	6	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	
AY.103	4244	11	0.3%	7	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	
AY.1	15	2	13.3%	19	[T19R,T95I,G142D,E156G,F157-,R158-,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,D950N]	0
AY.98.1	791	3	0.4%	7	[T19R,A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,P251L,L452R,T478K,D614G,P681R,D950N]	0
AY.4	6536	4	0.1%	32	[T19R,A27S,T95I,G142D,E156G,F157-,R158-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]	23
B.1.617.2	2013	5	0.2%	3	[T19R,T95I,G142D,V143-,Y144-,Y145-,E156G,F157-,R158-,L452R,T478K,D614G,P681R,D950N]	1

Only one was increasing in the most recent 60 day period: it went from being found 4 times to 24 times, so we have added it to our VOI list, though rare:

		N_Pango	N	%
Feb 14:	AY.4	6536	4	0.1%
March 29:	AY.4	148	23	15.5%

[T19R,A27S,T95I,G142D,E156G,F157-,R158-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F]

There were 162 *additional* distinct chimeric sequences with stretches of Omicron in Delta backgrounds that were only found once in a given Pango lineage set, in addition to the set of 17 that were found multiple times shown above.

There were no Omicron BA.2 specific mutations found in Delta backgrounds found more than 1x. (3/27/2022 update)

T19I,L24-,P25-,P26-,A27S,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K

3/27/2022: BA.2 Chimeric sequences found over 50 times among BA.2 sequences :
One is found strikingly often in the last 60 days.

```
BA.2 307400 248977 81.0%    [T19I,L24S,P25-,P26-,A27-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 11595 3.8%      [T19I,                  G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 97 0.0%         [T19I,                  G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,      S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 85 0.0% 5  [L5F,T19I,                  G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 84 0.0% 5       [T19I,W64L,             G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 75 0.0% 3       [T19I,L24S,P25-,P26-,A27-,      V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,           S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 72 0.0% 5       [T19I,                  G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K,C1236F
```

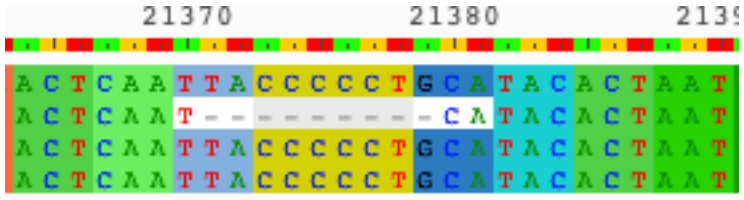
BA.1

A67V,H69-,V70-,T95I,G142D,V143-,Y144-,Y145-,N211-,L212I,+214EPE,G339D,S371L,S373P,S375F,K417N,N440K,G446S,S477N,T478K,E484A,Q493R,G496S,Q498R,N501Y,Y505H,T547K,D614G,H655Y,N679K,P681H,N764K,D796Y,N856K,Q954H,N969K,L981F

BA.2 forms found over 40 times carrying **BA.1 signature mutations**:

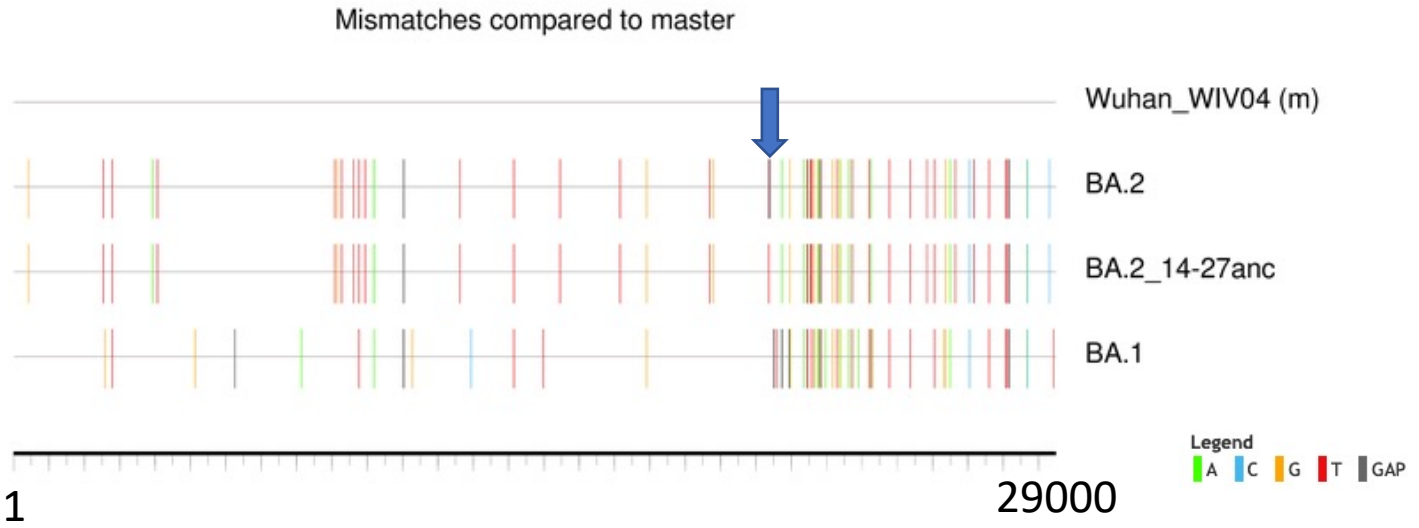
```
BA.2 307400 248977 81.0%   [T19I,L24S,P25-,P26-,A27-,G142D,           V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 197 0.1% 2     [T19I,L24S,P25-,P26-,A27-,G142D,N211-,L212I,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 53 0.0% 6      [T19I,                  G142D,N211-,L212I,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
BA.2 307400 41 0.0% 3      [T19I,L24S,P25-,P26-,A27-,A67V,H69-,V70-,G142D,V213G,G339D,S371F,S373P,S375F,T376A,D405N,R408S,K417N,N440K,S477N,T478K,E484A,Q493R,Q498R,N501Y,Y505H,D614G,H655Y,N679K,P681H,N764K,D796Y,Q954H,N969K]
```

In the entire genome, it differs from the BA.2 backbone only in terms of this single 9 base deletion -> ancestral:



The 9-base deletion that defines the L24S,P25-,P26-,A27-change in Spike in BA.2 is restored

Artifact, recombinant, insertion?



Chimeric/Recombinant Sequence Screening

I. GISAID notices: The GISAID development team has added a screening feature to alert submitters when they have submitted a likely chimeric sequence that carries mutational signatures characteristic of two distinct variants (*e.g.* Delta and Omicron).

Such sequences will be tagged 'Under investigation', and the submitter will be notified and asked to please review the sequence and confirm its accuracy. Upon confirmation the tag will be removed.

II. Pango lineages: Lineages are currently being assigned to identify some recombinant/chimeric sequences, currently they are being designated with by lineages names starting with "X". Only a subset have a breakpoint in Spike.

https://github.com/cov-lineages/pango-designation/releases

III. Natural recombinants: Some recombinant/chimeric sequences have been shown to be likely biologically valid recombinants, though these have been rarely sampled to date:

Colson et al. medRxiv: Culture and identification of a "Deltamicron" SARS-CoV-2 in a three cases cluster in southern France.
https://www.medrxiv.org/content/10.1101/2022.03.03.22271812v1

Lacek et al., bioRxiv: Identification of a Novel SARS-CoV-2 Delta-Omicron Recombinant Virus in the United States
https://doi.org/10.1101/2022.03.19.484981

NOTE: It is interest to resolve the origin of a recombinant even if only rarely sampled and found in a small local cluster. If such a lineage does not expand, however, it would be of limited interest for immunological and virological experiments.

IV. Sequencing artifacts: Even chimeras that result from sequencing artifacts are important to identify, as they can complicate interpretation of analyses bases on phylogenetic inference and can lead to erroneous conclusions. Chimeras and natural recombinants are present in some of the global data, and so should be carefully accounted for in phylogenetic analyses.

March April 2, BA.2 with ancestral amino acids between Spike 24-27:

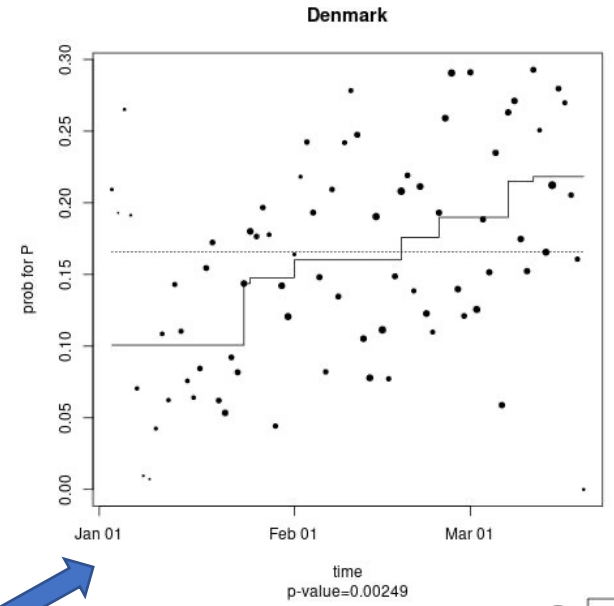
This is the most common BA.2 variant, but only consistently increasing in the places where it is most common and where BA.2 was first established, Denmark and India:

Position: SPIKE 24-27 LPPA
Assumption: Test amino acid form is increasing over time
Correlated variant: Do not consider. Include all sequences
Range of dates: 2022-01-03 - 2022-04-02
Pango lineage designation in GISAID (version: 2022-02-28): BA.2
Hosts: Human

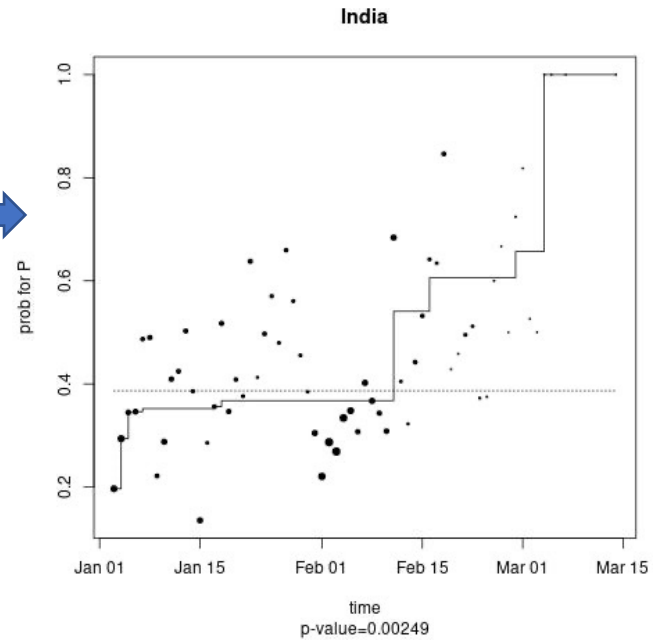
All countries where BA.2 with Spike 24-27 ancestral was sequenced >10 times – represented by "#P" column.

Country level

	# P	# Others	Total	P/Total (%)	# days	Time window	p-val
Australia	17	1438	1455	1.17	74	75	0.89303
Denmark	16925	85109	102034	16.59	76	75	0.00249
France	45	6875	6920	0.65	68	70	0.801
Germany	58	14754	14812	0.39	65	64	0.35821
India	3985	6319	10304	38.67	64	70	0.00249
Indonesia	22	844	866	2.54	70	70	0.84328
Ireland	58	445	503	11.53	45	55	0.25373
Italy	74	863	937	7.90	59	69	0.00249
Mauritius	13	51	64	20.31	28	40	0.16667
Romania	66	250	316	20.89	35	58	0.64179
Sweden	22	5106	5128	0.43	71	70	0.99751
Switzerland	295	3458	3753	7.86	65	70	0.93284
USA	48	8080	8128	0.59	75	80	0.73134
United-Kingdom	50	194294	194344	0.03	77	76	0.97015



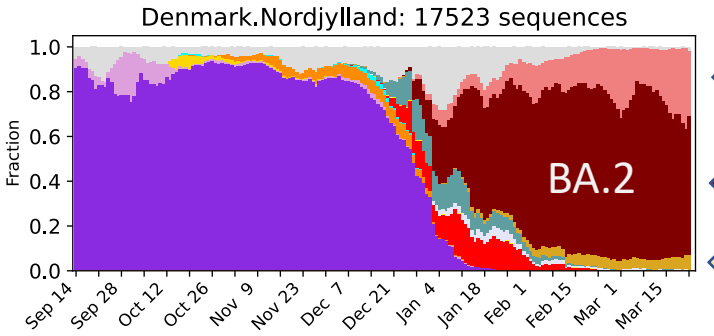
These plots show the increase in BA.2 variants with the Spike 24-27 ancestral form, relative to other BA.2 sequences, over time.



GISAID

Los Alamos NATIONAL LABORATORY

Denmark – The BA.2 with ancestral is slowly increasing in each region sampled.



BA.2, ancestral at S 24-27

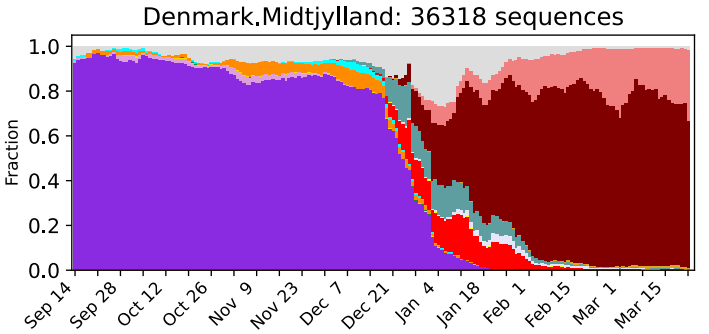
BA.2

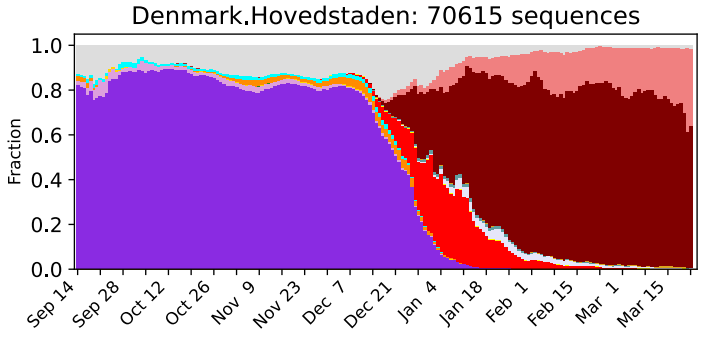
BA.2, insertion +212SGR

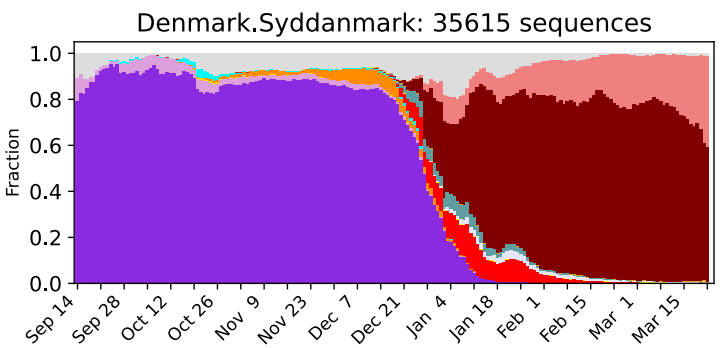
other
Omicron_BA.2+P25_,P26_
Omicron_BA.2
Omicron_BA.2+212SGR
Omicron+A701V
Omicron_BA.1.1
Omicron_BA.1.1+214GAR
Omicron_BA.1
Omicron_BA.1+214GAR
Delta+FurinRelated
Delta+Y145H,A222V
Delta+S112L
Delta+V1264L
Delta+T95_,A222V
Delta
B.1.640.2
B.1.640.1
Gamma+FurinRelated
Gamma
Mu
Lambda
Beta
Alpha

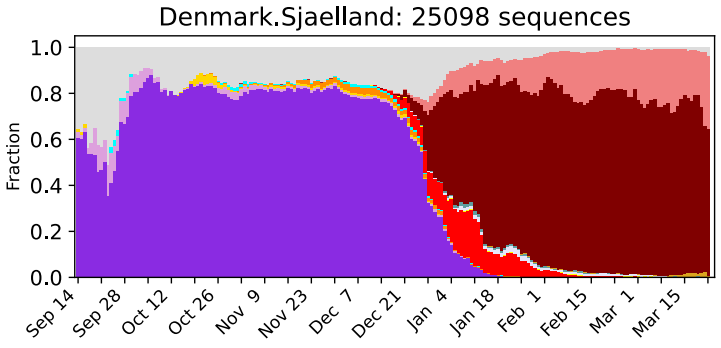
GISAID

Los Alamos NATIONAL LABORATORY





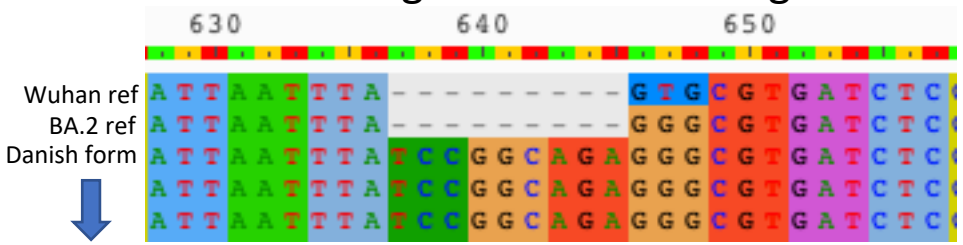




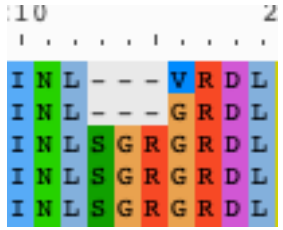
Insertion S 212 +SGR in BA.2 is found most often in Denmark, and though found at a modest levels it is increasing locally.

It is currently found 1,127 times in GISAID, and is now sampled throughout Europe, and in Australia, South Africa, and Israel

Codon aligned nucleotide region



Translation +212SGR



LANL Alignment, amino acid changes relative to BA.2

Note: GISAID currently usually lists this insertion as Spike +213GRG. As it arose is a BA.2 backbone, this would align codon GGG with the TCC codon in the insertion, and this is evolutionarily less likely.

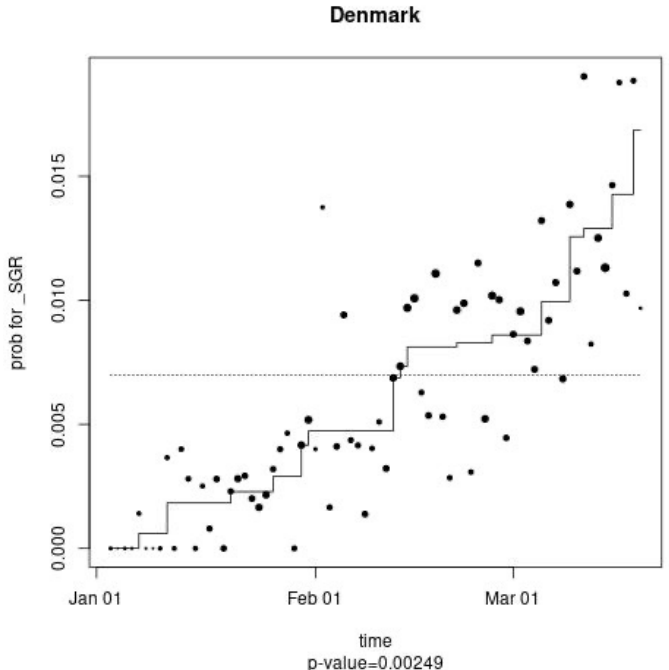
```
TTAGGG --------- CGTGAT      NLG --- RD    BA.2
TTATTC GGCAGAGGG CGTGAT      NLS GRG RD
```

This insertion is interesting as Omicron is sampling indel variants in this region, and other variants have also carried three amino acid insertions in this region:

BA.1: +214 EPE	Global transition
BA.1: loss of 214 insertion	May be a sequencing artifact
BA.1: +214 EPE change to +214 GAR	~5000 times
BA.2: L212F V213I +213GGG	Rare, India
B.1.214.2 +214 TDR	-- Belgium, France ~ a thousand
A.2.5. +214 AAG (+ D215Y)	-- US, central America, a few thousand

Country level

	# _SGR	# Others	Total	_SGR/Total (%)	# days	Time window	p-val
Denmark	712	101322	102034	0.70	76	75	0.00249



With thanks to the sequencing teams in Denmark for sharing their data through GISAID.