Open Access Research Journal of
Life Sciences
Journals home page: https://oarjpublication/journals/oarjls/
ISSN: 2783-025X (Online)

OARJ OPEN ACCESS RESEARCH JOURNALS

(RESEARCH ARTICLE)

Check for updates

Phylodynamic of hCoV-19 genomic variants: Overview of mutations SARS-CoV-2 on data hub

SIMÕES, R.S.Q [1, 2, *]

[1] *Department of Health and Medical Sciences, Santa Úrsula University, Fernando Ferrari, 75 – Botafogo, Rio de Janeiro, Brazil.*
[2] *PhD Tropical Medicine, Oswaldo Cruz Foundation, Avenida Brazil 4.365 – Manguinhos, Rio de Janeiro, Brazil.*

Open Access Research Journal of Life Sciences, 2026, 11(01), 034-050

Publication history: Received on 16 January 2026; revised on 04 March 2026; accepted on 07 March 2026

Article DOI: https://doi.org/10.53022/oarjls.2026.11.1.0011

Abstract

Coronavirus disease 2019 (COVID-19) is a highly infectious illness caused by the acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which started in China in December 2019 and rapidly spread across the world became the pandemic disease of the 21st century. A systematic review and meta-analysis to analyze the impact of the novel genomic variants more recently detected in countries where people have a higher risk of developing severe illness from COVID-19 were investigated using SARS-CoV-2 Data Hub, NCBI virus assembly. Four SARS-CoV-2 genomic variants have been selected and analyzed comparatively in different years. In addition, verifying the tracking of 19 hCoV-19 variants by accessible via GISAID EpiCoV database was used to analyze the relative variant genome frequency per region (exponentially smoothed alpha=0.3), phylodynamic of hCoV-19 variant and most recent sample collection per country SARS-CoV-2 hCov 19 variants to overview of mutations also registered. As results, geography dates using PANGO lineage about 20 tracking of 19 hCoV-19 genomic variants including number of shared genomic sequences, numbers of countries and first detected was investigated. So, the emergence and reemergence of viral diseases may be accompanied by the genomic and epidemiology surveillance to mitigate any risk of propagation of the novel possible variant viral. This study dataset was a backward-looking for circulating genomic variants and genetic engineering for the development of biotechnological platforms vaccines for SARS-CoV-2 candidates licensed or under development associated others respiratory viral infections.

Keywords: COVID-19; Genomic variants; SARS-CoV-2, Phylodynamic; Mutations

1 Introduction

Coronavirus disease 2019 (COVID-19) is a highly infectious illness caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) which started in China in December 2019 and quickly spread across the world became the pandemic disease of the 21st century. Novel genomic variants SARS-CoV-2 based on phylogeny and the viral taxonomy among several clusters had been documented in countries where people showed a higher risk of developing severe illness from COVID-19 [1]. SARS-CoV-2 is classified in *Nidovirales* order as member of *Coronaviridae* family, *Orthocoronavirinae* subfamily which virus are enveloped, positive-sense single-stranded RNA (ssRNA) coronavirus. In terms of viral genomics, this RNA virus are spherical particles of 100–160 nm in diameter with a genome of 27–32 kb presenting a protein open reading frame 1 (ORF-1) responsible by cleaved into 16 non-structural proteins (NSPs1–NSPs16) [1,2].

Retrospective epidemiology study had been described tracking of 19 hCoV-19 genomic variants and most recent sample collection per country SARS-CoV-2 hCov 19 variants [3]. Genetic recombination makes be possible genomic fragments and glycoproteins linked the cellular pathways cross geographic barriers to adapt into new species facilitated interspecies transmissibility adapting to a new host named as spillover phenomenon [4]. Moreover, biotechnological

* Corresponding author: Profa. Dra. Rachel Siqueira de Queiroz Simões

strategies adopted to SARS-CoV-2 vaccine development associated with the research about genomic variants had been reported demonstrating the high importance of monitoring up-to-date data on this approach worldwide [5].

The main objective of this study was carried out through the selection of four SARS-CoV-2 genomic variants to analyze comparatively among different years. In addition, the tracking of 19 hCoV-19 variants by accessible via GISAID EpiCoV database was used to analyze the relative variant genome frequency per region (exponentially smoothed alpha=0.3), phylodynamic of hCoV-19 variant and most recent sample collection per country SARS-CoV-2 hCov 19 variants to overview of mutations.

2 Material and Methods

A systematic review and meta-analysis to analyze the impact of the novel genomic variants (https://cov-lineages.org/) more recently detected in countries where people have a higher risk of developing severe illness from COVID-19 were investigated using SARS-CoV-2 Data Hub, NCBI virus assembly (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/). In addition, verifying the tracking of 19 hCoV-19 variants by accessible via GISAID EpiCoV database (https://gisaid.org/hcov19-variants/) was used to analyze the relative variant genome frequency per region (exponentially smoothed alpha=0.3), phylodynamic of hCoV-19 variant and most recent sample collection per country SARS-CoV-2 hCov 19 variants to overview of mutations also registered.

3 Results

Four SARS-CoV-2 genomic variants have been selected and analyzed comparatively in different years. As results, geography dates using PANGO lineage about 20 tracking of 19 hCoV-19 genomic variants including number of shared genomic sequences, numbers of countries and first detected was investigated. According to the SARS-CoV-2 data center, 14,732,459 nucleotides, 64,442,658 proteins, and 166,698 virus assemblies have been identified by the NCBI. However, the five main lineages in international growth are: NY.15; KP.1.1.3; MC.10; LB.1.3.1 and LP.8.1. Then, there are another five main growth lineages in the United States which are: PA.1; XEC.1; LB.1.3.1; MC.8.1 and MB.1.1.1. Some of these lineages have mutations and are best represented in table 1 and 2 below:

Table 1 International lineages and their mutations

International Lineages	Mutations
NY.15	3CLpro:S81F ORF3a:A103T
KP.1.1.3	nsp3:N760 nsp13:T366 S:K1086R
MC10	RdRp:T644 nsp13:D32
LB.1.3.1	nsp14:T311 ORF7a:T28I
LP.8.1	nsp2:Q383 nsp3:N760 nsp3:A1179V nsp3:T1465I nsp13:T366 S:F186L S:R190S S:V445G S:K1086R ORF3a:P178L

GenBank and SRA records were compared to the SARS-CoV reference sequence RefSeq NC_045512 (assembly NCBI: GCF_009858895.2) to identify alterations in the viruses' genomes. The mutations are used to assign Pango lineages to the sequence records.

Table 2 Identification of lineages, Gen Bank Accession, Geographic region and their total mutations

Lineages	Gen Bank	Geographic Region	Total Mutations	Mutations Spike Protein
NY.15	-	Denmark	115	44
NY.15	OZ309373.1	UK	144	68
NY.15	OZ309064.1	UK	142	68
NY.15	OZ308701.1	UK	138	68
NY.15	OZ309060.1	UK	143	68
NY.15	OZ309049.1	UK	143	68
NY.15	OZ308857.1	UK	132	66
NY.15	OZ309498.1	UK	137	65
NY.15	-	USA	118	45
NY.15	-	USA	118	44
NY.15	-	USA	119	48
NY.15	-	USA	119	48
NY.15	PX440188.1	USA	140	69
NY.15	-	USA	94	31
LP.8.1.mix	PX257845.1	USA	144	70
XFG.2	PX238522.1	USA	139	69
NY.15	-	USA	116	45
NY.15	-	USA	115	44
NY.15	-	USA	118	46
NY.15	-	USA	118	45
NY.15	-	USA	120	46
NY.15	-	USA	116	45
NY.15	-	USA	115	43
NY.15	-	USA	117	46
NY.15	-	USA	114	44
NY.15	-	USA	114	48
NY.15	-	USA	118	46

Lineages	Gen Bank	Geographic Region	Total Mutations	Mutations Spike Protein
KP.1.1.3	-	Spain	113	43
KP.1.1.3	-	Spain	113	43
XFC	-	Spain	116	42
KP.1.1.3	-	Spain	114	43
NY.16.1	-	Spain	109	43
NY.15	-	Denmark	115	44
NO.1.2	-	Denmark	121	46
XFC.1	-	Denmark	117	44
XFC.1	-	Netherlands	93	33
KP.1.1.3	-	Netherlands	92	30
KP.1.1.3	-	Netherlands	91	30
XFC.1	-	Netherlands	95	32
NO.1	-	Denmark	125	45
KP.1.1.3	-	Denmark	108	41
not assignable	-	Denmark	123	46
XFC	-	Denmark	119	45
KP.1.1.3	-	Slovakia	97	35
KP.1.1.3	-	Slovakia	98	27
not assignable	-	Denmark	125	50
KP.1.1.3	-	USA	98	31
NO.1	-	USA	93	31
KP.1.1.3	-	USA	89	30
NO.1.2	-	USA	94	31
PD.2.1	-	USA	97	31
KP.1.1.3	-	USA	87	30
KP.1.1.3	-	USA	87	30
KP.1.1.3	-	USA	116	46
XFG.2	PX238518.1	USA	135	71
XFG.2	PX238573.1	USA	136	71
NO.1	PX238568.1	USA	113	44
KP.1.1.3	PX238528.1	USA	112	47
KP.1.1.3	PX238539.1	USA	111	45
XFC	PX238515.1	USA	82	28
PQ.1	PX238547.1	USA	85	39
XFG.2	PX238522.1	USA	139	69
NY.1	PX236726.1	USA	141	72
NY.1		USA	115	48
NO.1.2	PX236744.1	USA	121	44
NO.1.2	-	USA	123	45
NO.1.2	PX236721.1	USA	149	72
no lineage	-	USA	83	30
NO.1	-	USA	95	31
no lineage	-	USA	97	30
NO.1	-	USA	94	31
NY.15	-	USA	94	31
NO.1	-	USA	92	32
KP.1.1.3	-	USA	94	32
PF.2.4	-	USA	93	31
KP.1.1.3	-	USA	87	29
PF.2.4	-	USA	89	29
KP.1.1.3	-	USA	90	30
KP.1.1.3	-	USA	93	29
PF.2.4	-	USA	93	30
NO.1	-	USA	91	31
KP.1.1.3	-	USA	90	31
NO.1	-	USA	93	30
KP.1.1.3	-	USA	116	45
XFC.1	-	USA	117	45
NO.1	-	USA	113	44
KP.1.1.3	-	USA	115	45
XFC.1	-	USA	117	47
NY.16.1	-	USA	116	47
PF.2	-	USA	115	46
XFC	-	USA	121	48
XFC.1	-	USA	115	45
NO.1	-	USA	115	45
KP.1.1.3	-	USA	119	46
KP.1.1.3	-	USA	118	48
NY.16.1	-	USA	117	47
NO.1.2	-	USA	108	36
KP.1.1.3	-	USA	110	42
NO.1.2	-	USA	92	28
NY.16.1	-	USA	95	32
NY.3.1	-	USA	97	31
NO.1.2	-	USA	104	33
NO.1.2	-	USA	98	31
not assignable	-	USA	67	20
NO.1	PX275797.1	USA	138	69
KP.1.1.3	-	USA	88	29
KP.1.1.3	-	USA	94	30
KP.1.1.3	-	USA	87	28
NO.1	-	USA	91	31
KP.1.1.3	-	USA	95	30
PF.2.4	-	USA	89	29
NO.1	-	USA	91	31
XFC.1	-	USA	82	29
NO.1	-	USA	92	31
NO.1	-	USA	94	31
KP.1.1.3	-	USA	89	29
NO.1	-	USA	87	29
NO.1	-	USA	92	31
KP.1.1.3	-	USA	89	29
PF.2	-	USA	87	29
NO.1	-	USA	92	31
PF.2.4	-	USA	87	29
no lineage	-	USA	90	31
NO.1	-	USA	91	31
NO.1	-	USA	97	31
NO.1	-	USA	92	32
NO.1	-	USA	94	31

3.1 Genomic Variant FORMER VOI (BA.2.86+BA.2.86.*)

About 113 countries shared 25,683 GRA (BA.2.86+BA.2.86.*) genome sequences with unprecedented speed from sample collection to make these data publicly accessible via GISAID EpiCoV, in some cases within less than 24 hours. Moreover, from SARS-CoV-2 Data Hub, NCBI virus assembly (GCF_009858895.2) showed 9.112.993 nucleotide, and 47.593.360 protein registered at GenBank/RefSeq. SARS-CoV-2 outbreak Sequence Statistics resulted 9110990 collected from countries and 9108454 samples hosts. The graphs and images of the genomic variant VOI (BA.2.86+BA.2.86.*) can be better visualized by following figures 1 until 6.

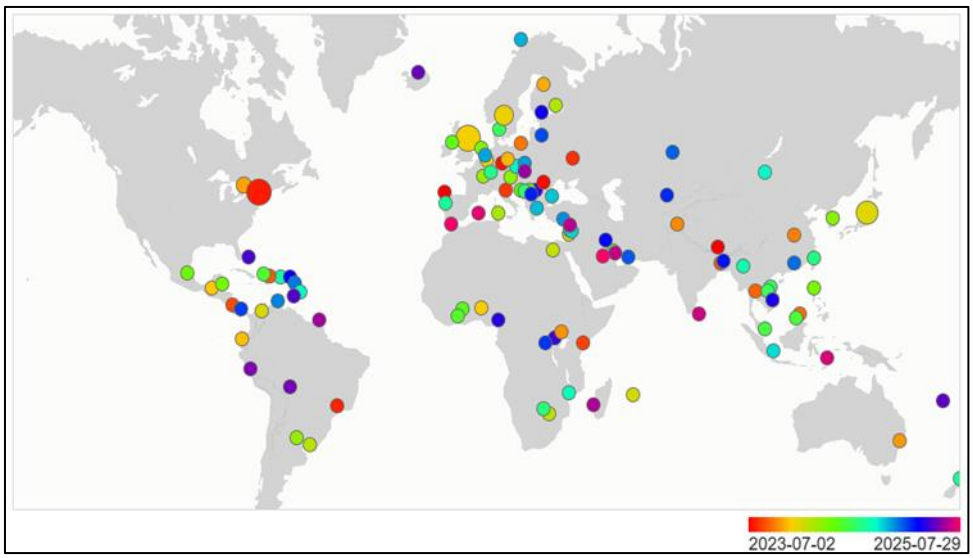

Figure 1 Map of tracked variant occurrence- BA.2.86 (excluding JN.1, JN.1.*). Accessed August 2025

As of 16 January 2026 - 1725UTC, 110 countries shared 24,993 (BA.2.86+BA.2.86.*) genome sequences with unprecedented speed from sample collection to make these data publicly accessible via GISAID EpiCoV, in some cases within less than 24 hours.

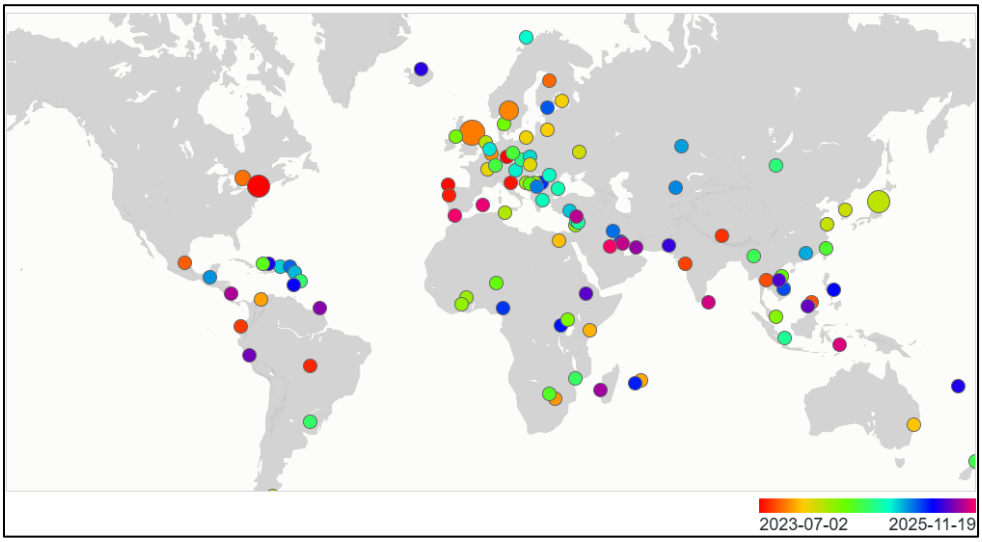

Figure 2 Map of tracked variant occurrence (BA.2.86+BA.2.86.*). Accessed January 2026

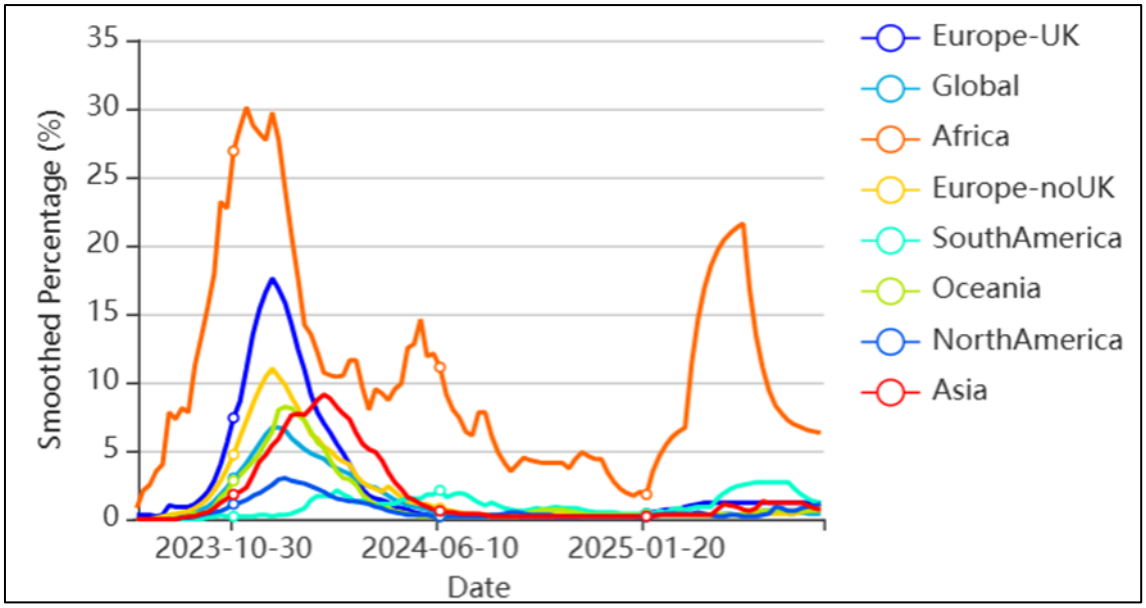

Figure 3 Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3) - BA.2.86 (excluding JN.1, JN.1.*)

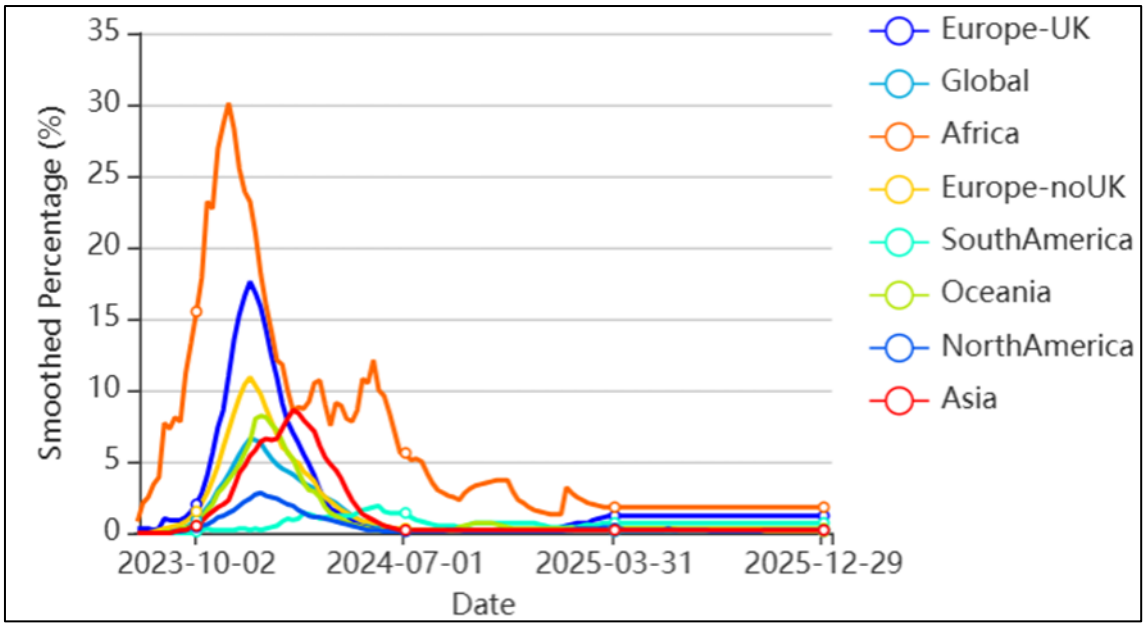

Figure 4 Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3) - BA.2.86 (excluding JN.1, JN.1.*). Accessed January 2026

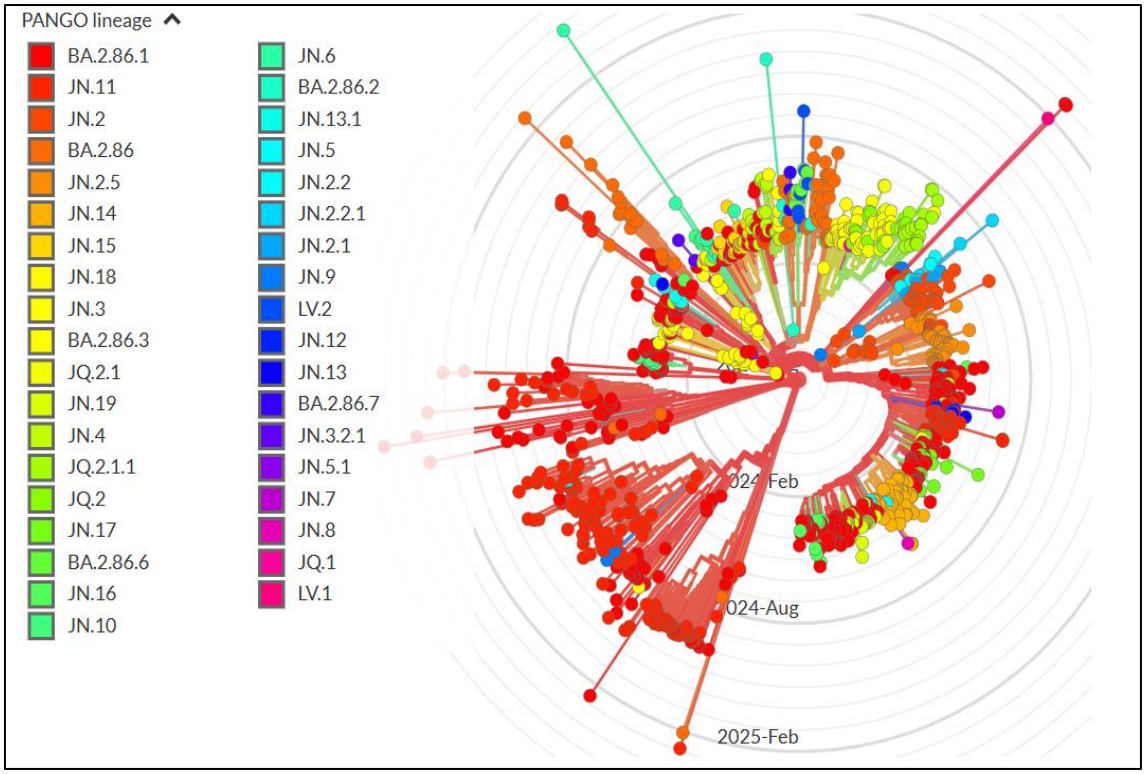

Figure 5 Phylodynamic of hCoV-19 variant BA.2.86 (excluding JN.1, JN.1.*) across the Globe – As od 9 August 2025 – 215UTC, 114 countries shared 26,435 (BA.2.86) genome sequences. Accessed August 2025

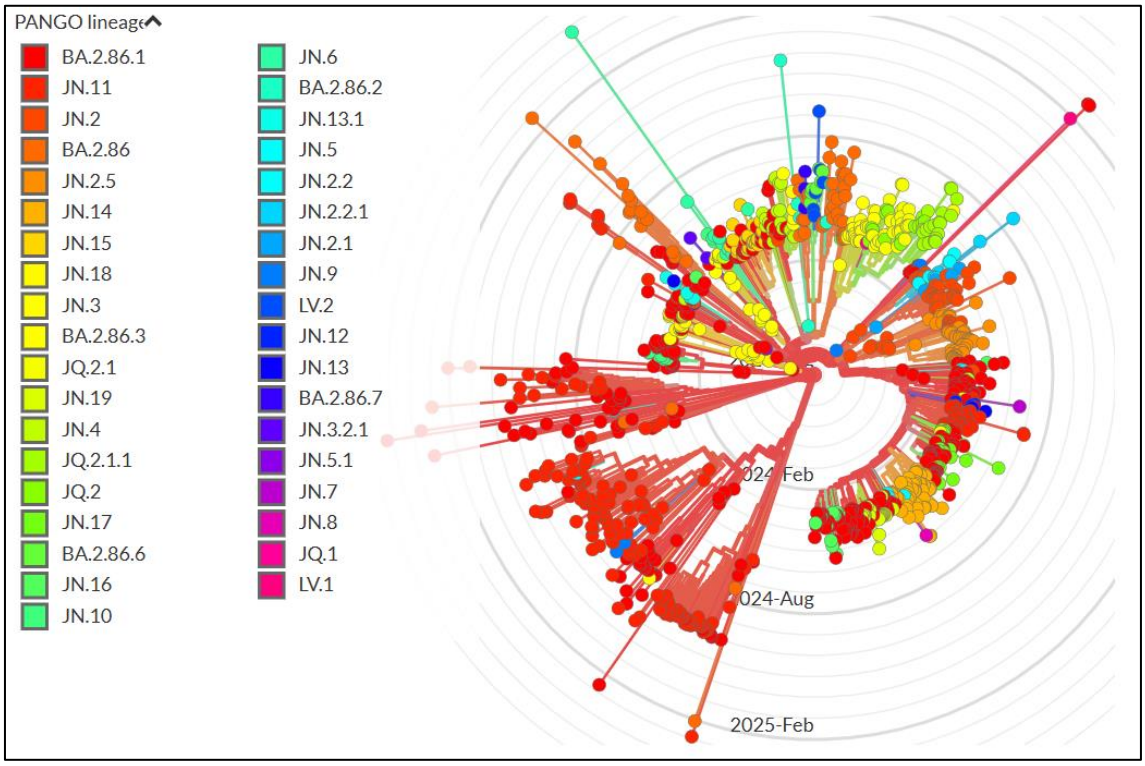

Figure 6 Phylodynamic of hCoV-19 variant BA.2.86 (excluding JN.1, JN.1.*) across the Globe showing 1.044 of 1.044 genomes collected between Sep 2023 and May 2025, last updated 2025-05-26. Accessed January 2026

3.2 Genomic Variant VOI JN.1+JN.1.* excluding current VUMs

As of 16 January 2026 – 1725 UTC, 155 countries shared 403,435 (JN.1+JN.1.* excluding current VUMs) genome sequences with unprecedented speed from sample collection to making these data publicly accessible via GISAID EpiCoV, in some cases within less than 24 hours. The graphs and images of the genomic variant VOI JN.1+JN.1.* excluding current VUMs can be better visualized by following figures 7 until 12.

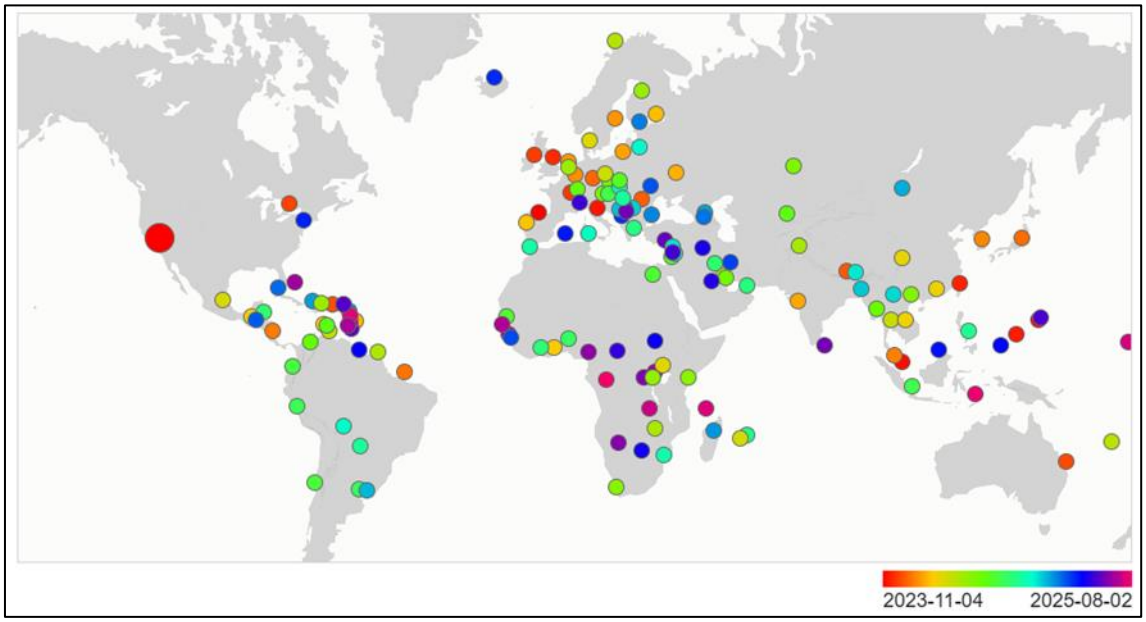

Figure 7 Map of tracked variant occurrence - JN.1+JN.1.* excluding current VUMs (accessed August 2025)

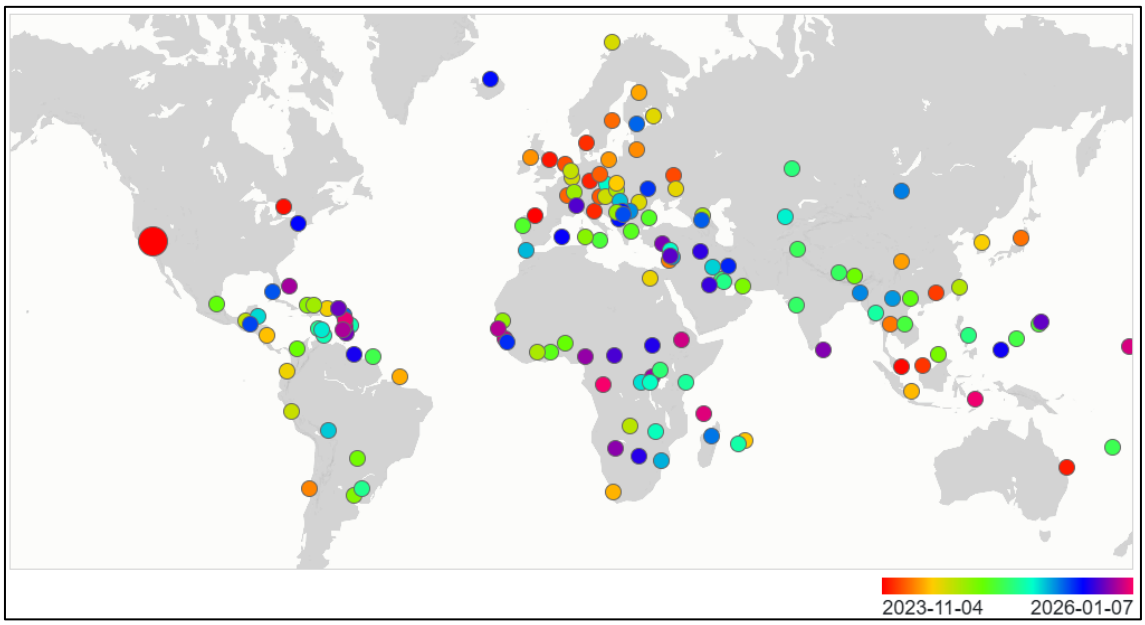

Figure 8 Map of tracked variant occurrence - JN.1+JN.1.* excluding current VUMs (accessed January 2026)

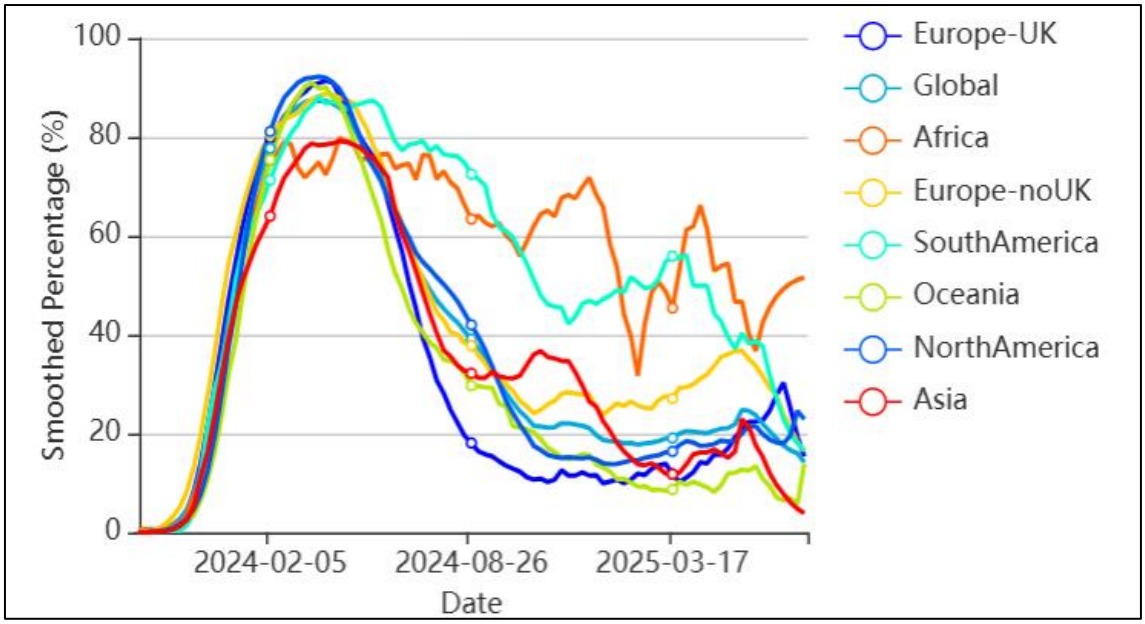

Figure 9 Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3) accessed August 2025

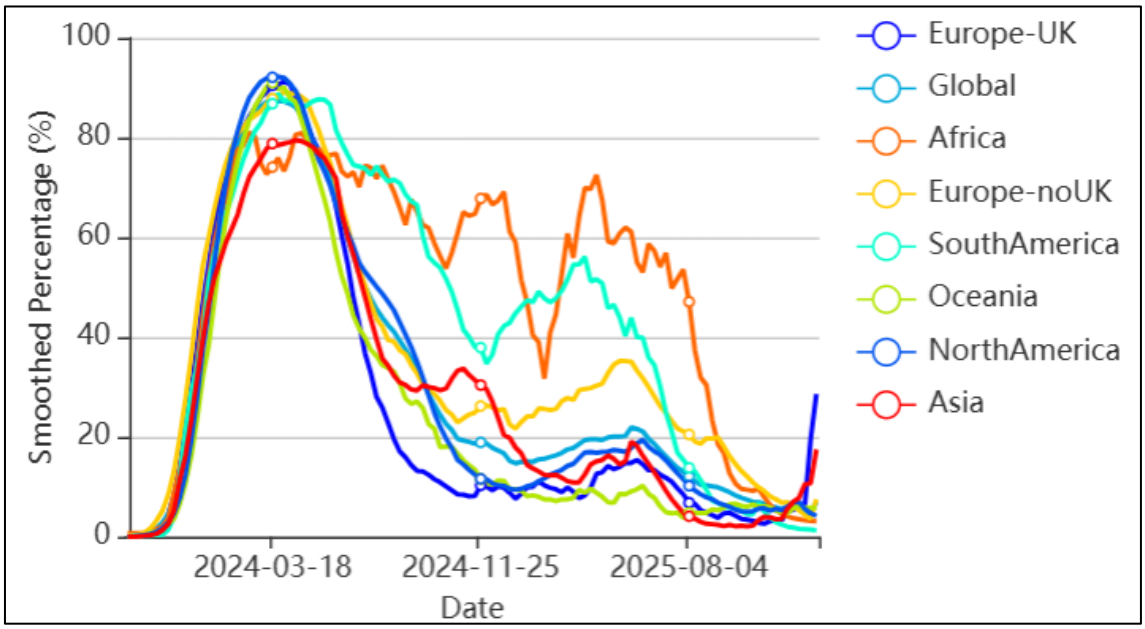

Figure10 Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3) accessed January 2026

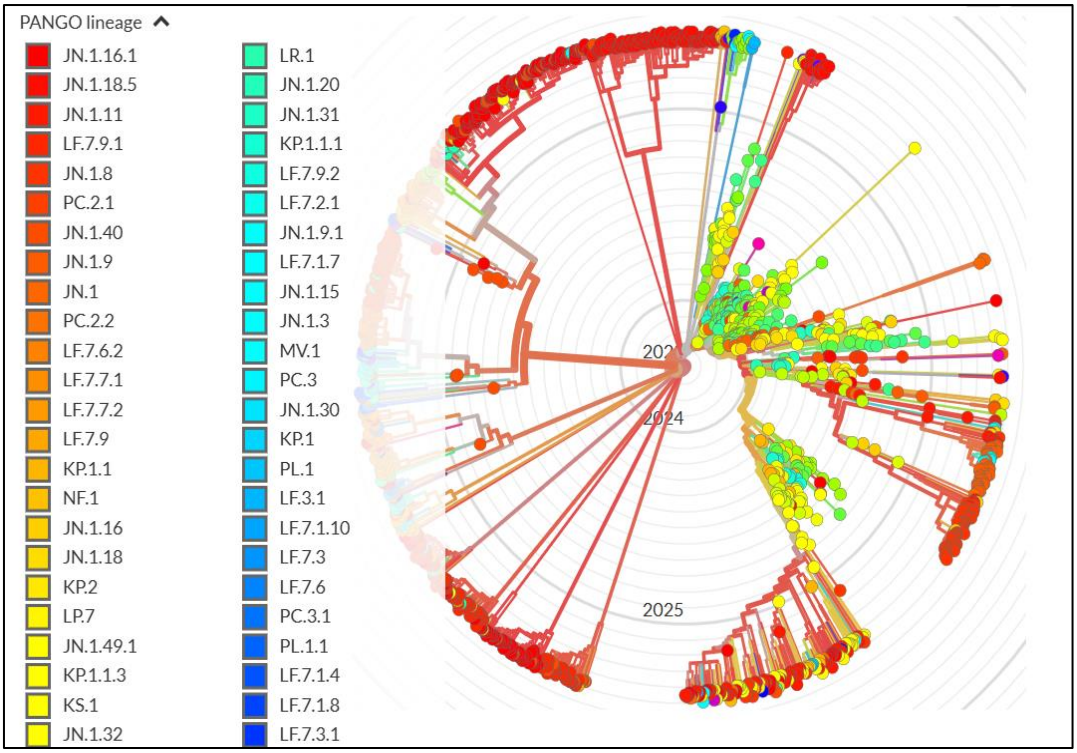

Figure 11 Phylodynamic of hCoV-19 variant JN.1+JN.1.* across the Globe -showing 1.399 genomes, As of 10 August 2025 - 152 countries shared 400,808 (JN.1+JN.1.* excluding current VUMs) genome sequences (accessed August 2025)

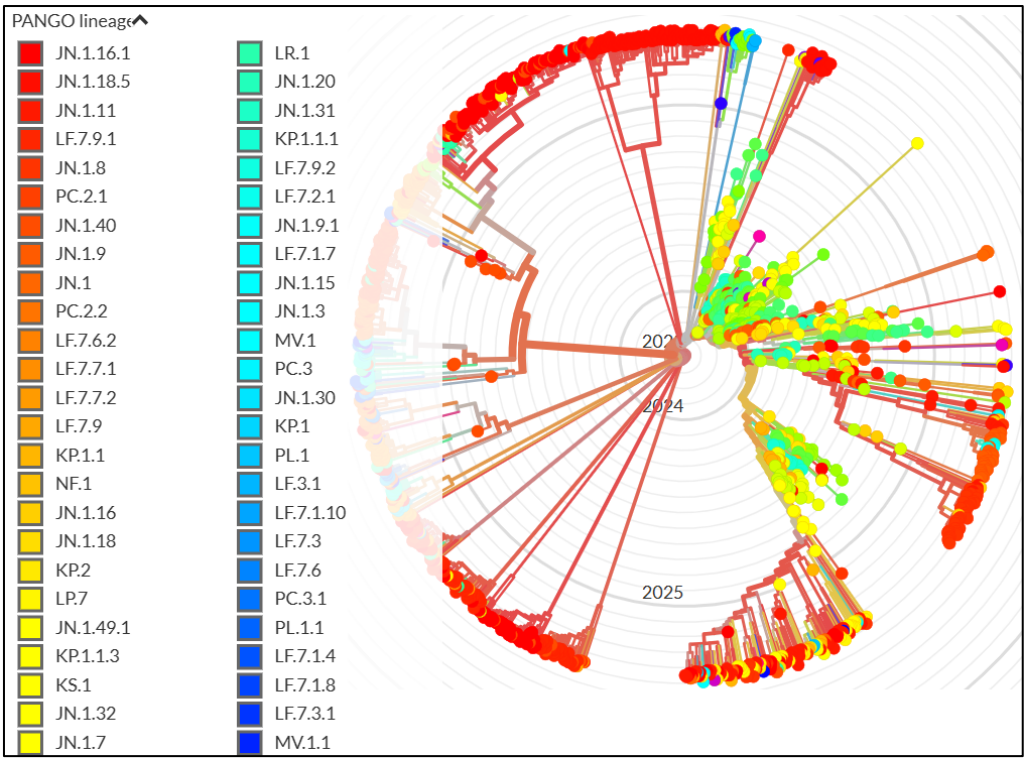

Figure 12 Phylodynamic of hCoV-19 variant JN.1+JN.1.* across the Globe -showing 1.951 of 1.951 genomes collected between Jan 2023 and Jul 2024, last updated 2025-05-26 (accessed January 2026)

3.3 Genomic Variant FORMER VUM (BA.2.75+BA.2.75.*)

As of 16 January 2026 – 1725 UTC, 140 countries shared 129,413 (BA.2.75+BA.2.75.*) genome sequences with unprecedented speed from sample collection to making these data publicly accessible via GISAID EpiCoV, in some cases within less than 24 hours. The graphs and images of the genomic variant VUM (BA.2.75+BA.2.75.*) can be better visualized by following figures 13 until 18.

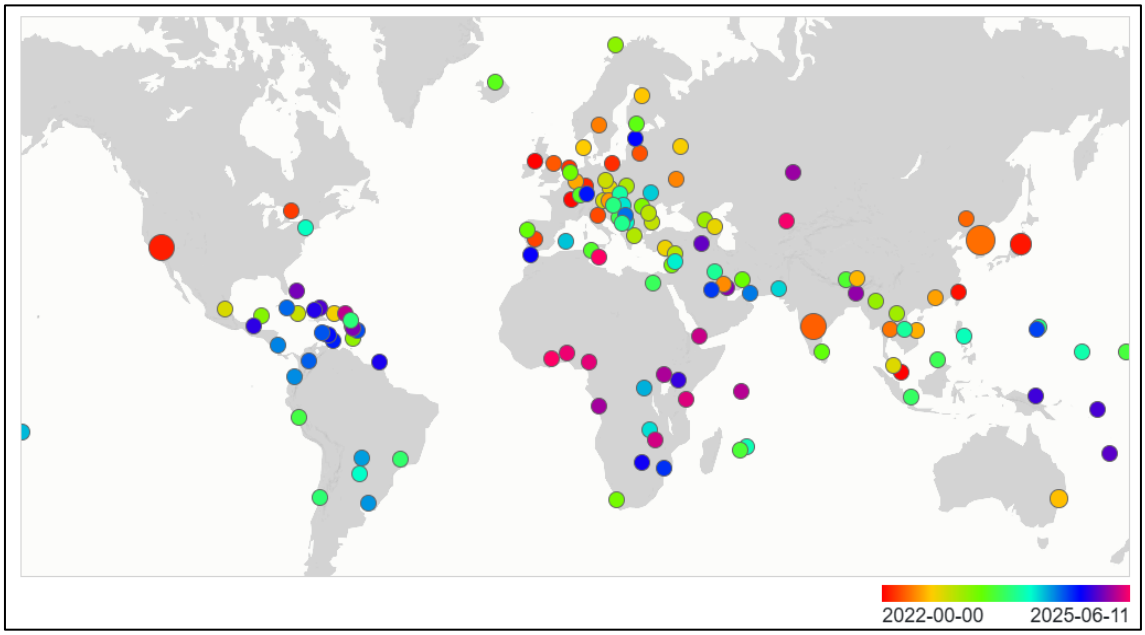

Figure 13 Map of tracked variant occurrence - (BA.2.75+BA.2.75.*). Accessed August 2025

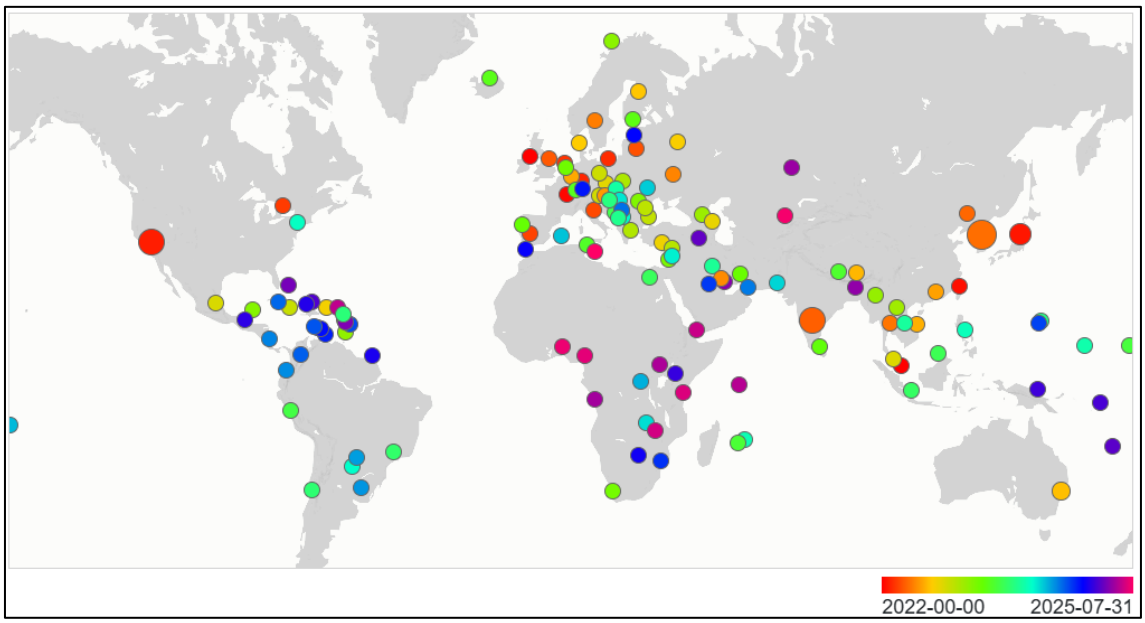

Figure 14 Map of tracked variant occurrence - (BA.2.75+BA.2.75.*). Accessed January 2026

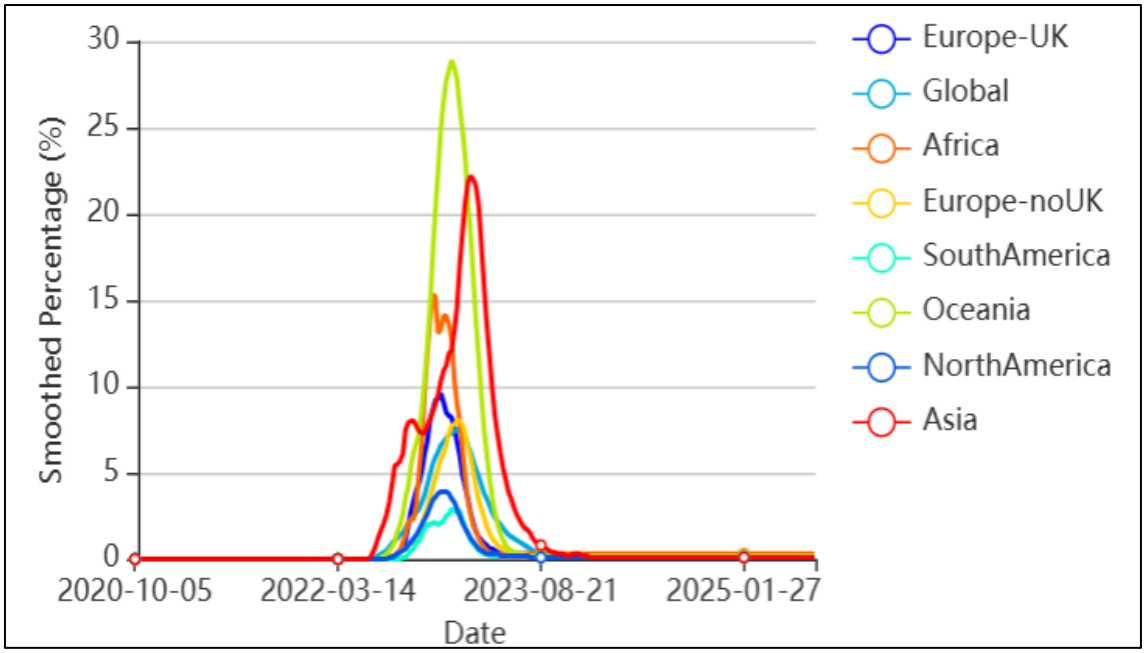

Figure 15 Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3) - (BA.2.75+BA.2.75.*). Accessed August 2025

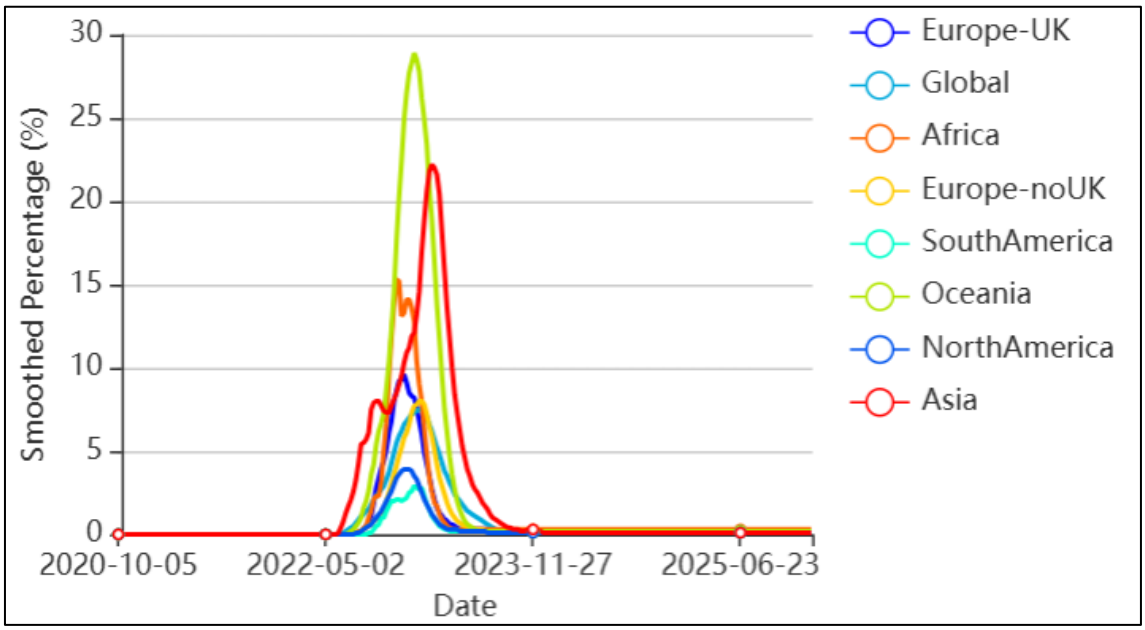

Figure 16 Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3). Accessed January 2026

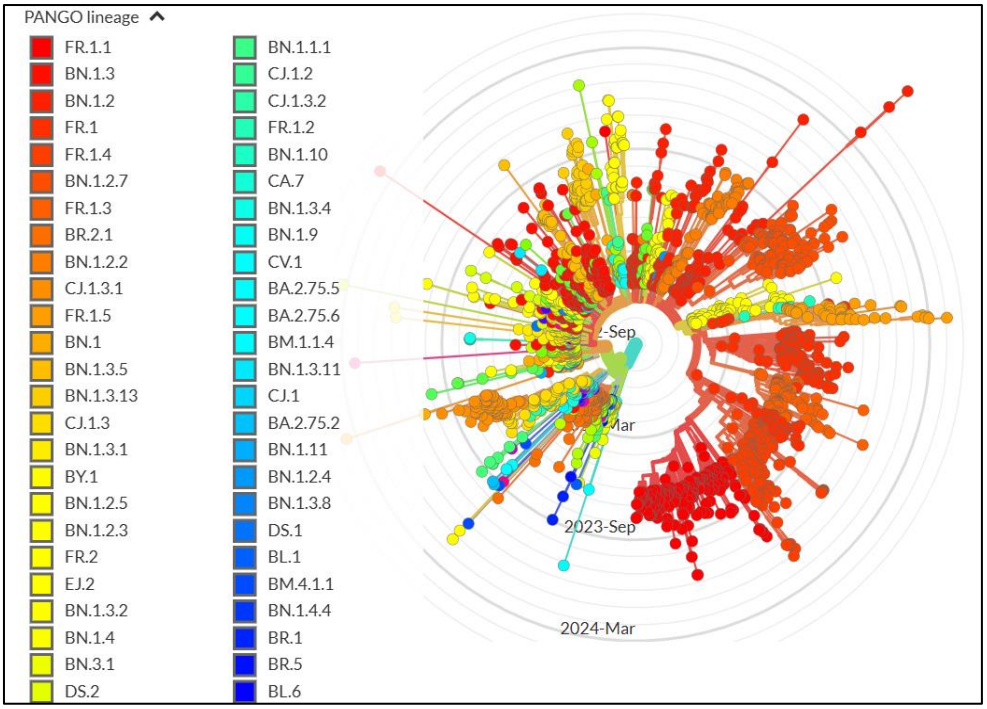

Figure 17 Phylodynamic of hCoV-19 variant (BA.2.75+BA.2.75.*) across the Globe – As of 10 August 2025 – 0013UTC,140 countries shared 129,287 (BA.2.75+BA.2.75.*) genome sequences. Accessed August 2025

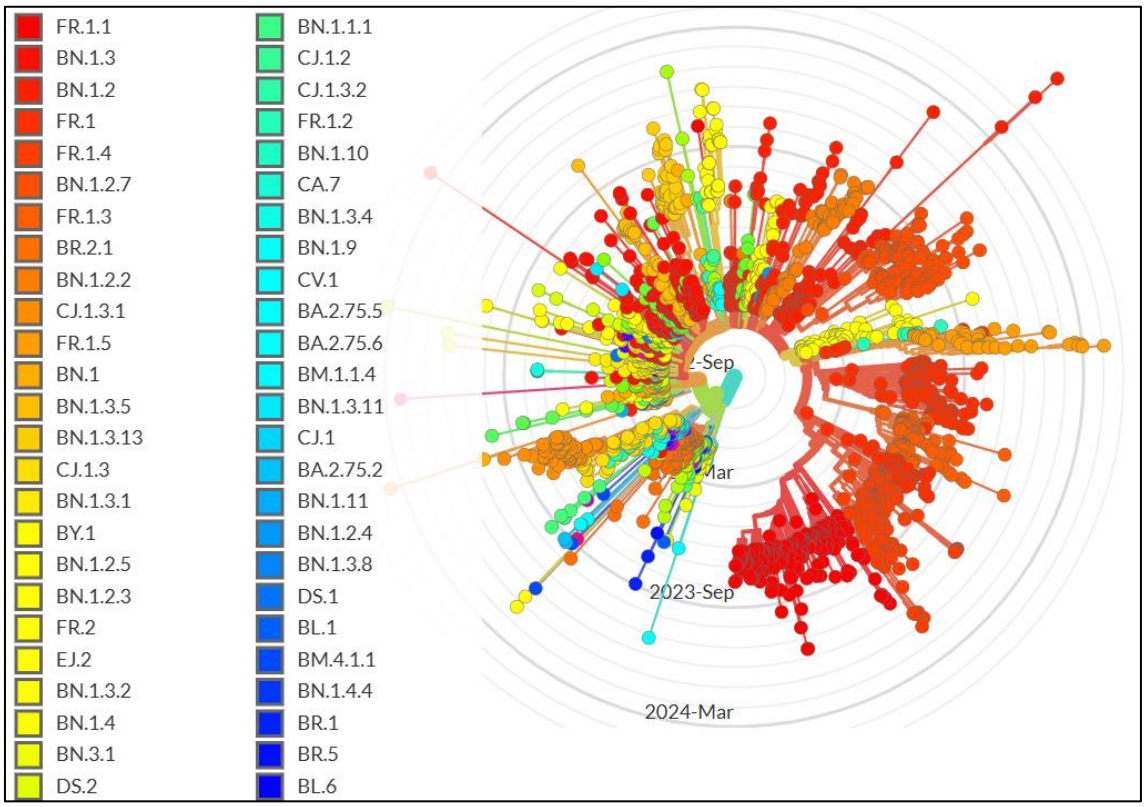

Figure 18 Phylodynamic of hCoV-19 variant (BA.2.75+BA.2.75.*) across the Globe showing 1.951 of 1.951 genomes collected between Jan 2023 and Jul 2024, last updated 2025-05-26. Accessed February 2026

3.4 Genomic Variant FORMER VOI (XBB.1.5+XBB.1.5.*)

As of 16 January 2026 - 1725UTC, 154 countries shared 408,097 (XBB.1.5+XBB.1.5.*) genome sequences with unprecedented speed from sample collection to making these data publicly accessible via GISAID EpiCoV, in some cases within less than 24 hours. The graphs and images of the genomic variant VOI (XBB.1.5+XBB.1.5.*) can be better visualized by following figures 19 until 24.

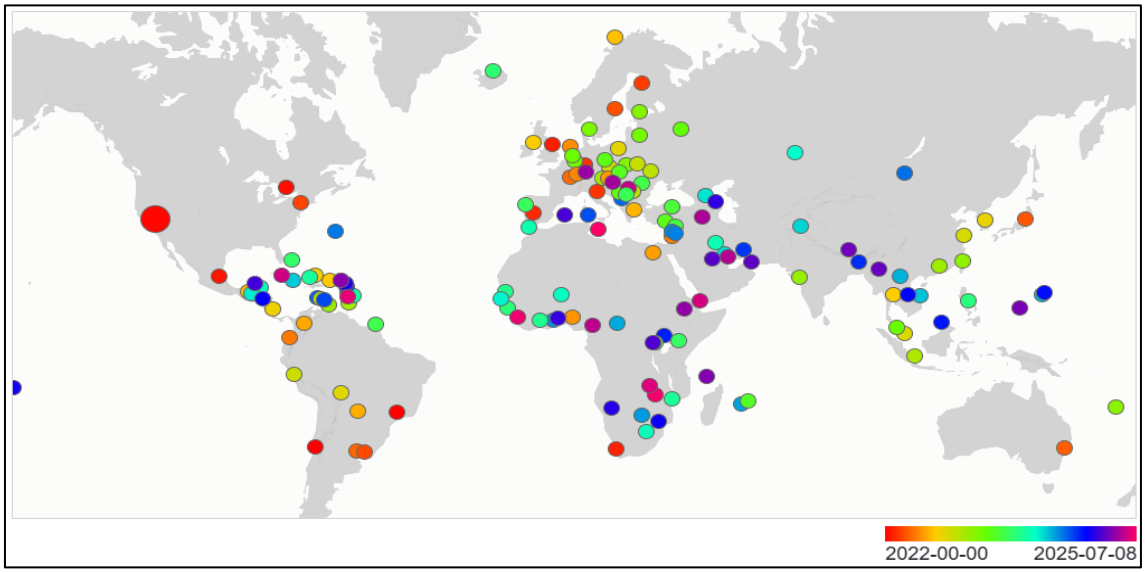

Figure 19 Map of tracked variant occurrence - (XBB.1.5+XBB.1.5.*)

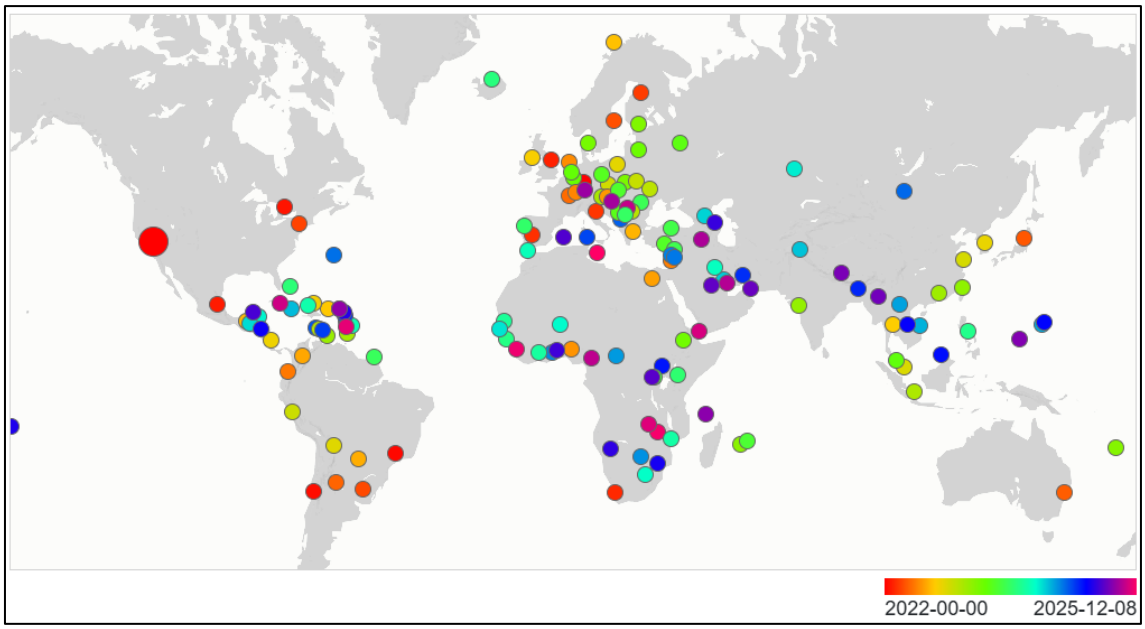

Figure 20 Map of tracked variant occurrence - (XBB.1.5+XBB.1.5.*). Accessed January 2026

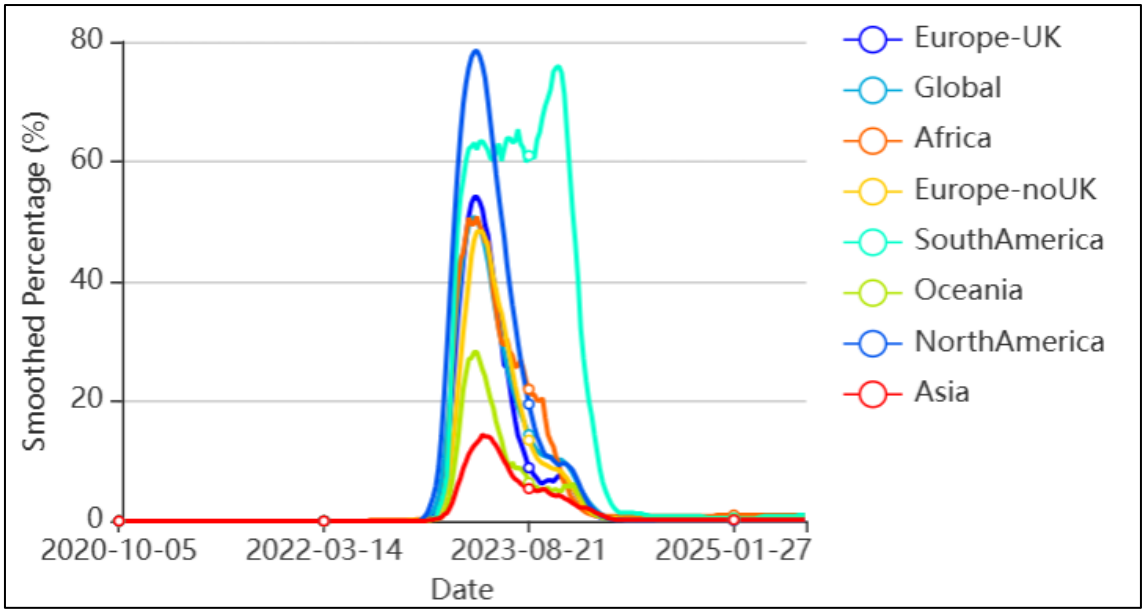

Figure 21 Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3) - (XBB.1.5+XBB.1.5.*). Accessed August 2025

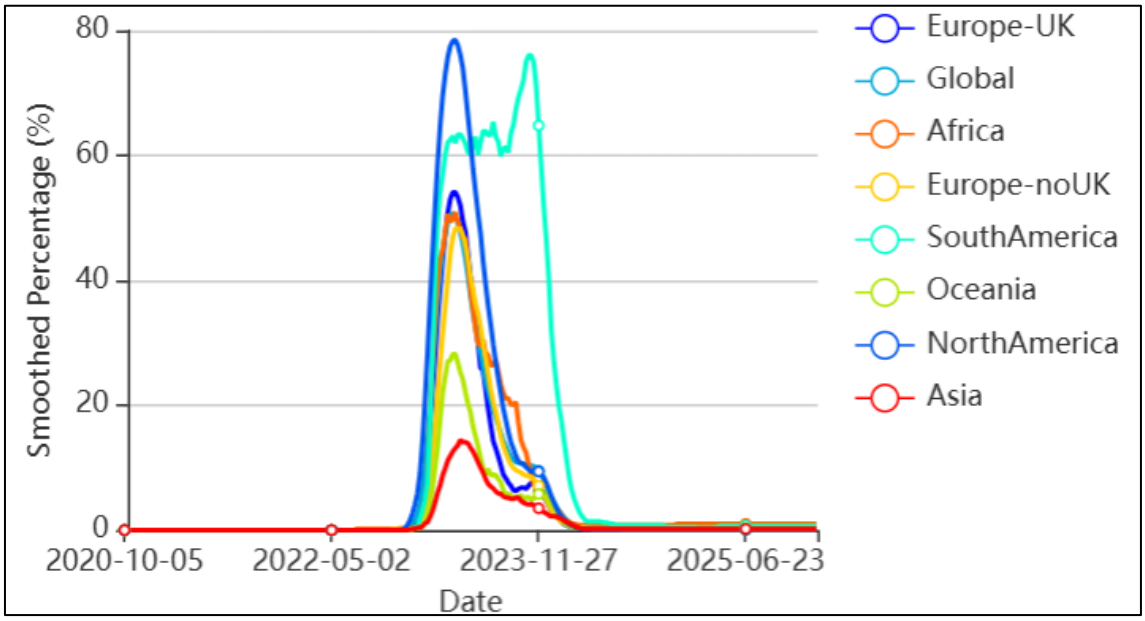

Figure 22 Relative Variant Genome Frequency per Region (exponentially smoothed alpha=0.3). Accessed January 2026

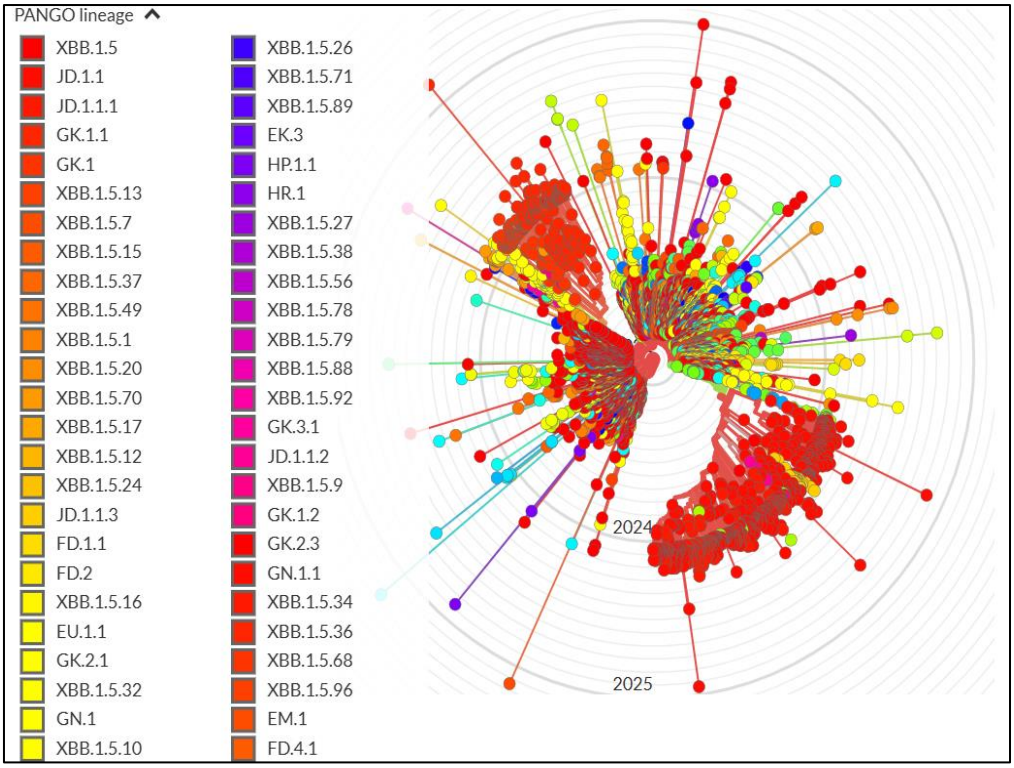

Figure 23 Phylodynamics of hCoV-19 variant XBB.1.5+XBB.1.5.* across the Globe. Accessed August 2025

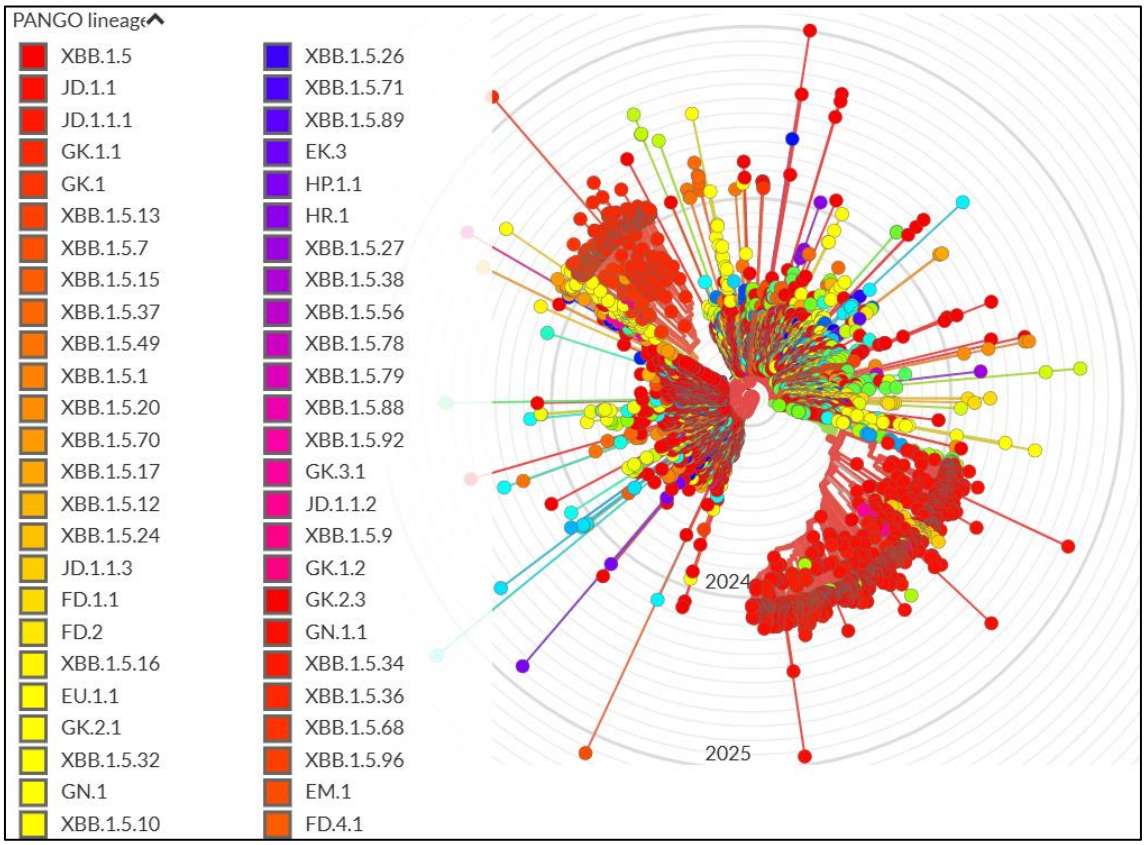

Figure 24 Phylodynamics of hCoV-19 variant XBB.1.5+XBB.1.5.* across the Globe showing 3.881 of 3.881 genomes collected between Jan 2023 and Mar 2025, last updated 2025-05-26. Accessed January 2026

4 Discussions

The authors suggest follow-up with new molecular research associated platforms and databases using bio impressions tools to clarify some genetics rearrangements. Preparing a rapid response to re-emerging viral outbreaks requires aligned genomic and epidemiological surveillance to mitigate potential risks of reintroduction of new pathogens. Furthermore, concern about the emergence of new viral lineages raises a red flag for a one health approach with constant-scale molecular surveillance. Several molecular studies with approach systematic review and meta-analysis during SARS-CoV-2 pandemic have described the impact of genome sequences shared in many countries including patients with comorbidities in COVID-19 [3]. Most recent reported occurrences in different countries on the European, Asian, American, Oceania, Africa continent according to tracking of variants (VOC Delta GK (B.1.617.2 + AY)) first detected in India and cases reported in different countries according to tracking of variants (VUM GH/490R (B.1.640 + B.1.640*)) first detected in Congo/France had been documented for each country, strain name, collection date and GISAID accession into novel lineages [5].

5 Conclusions

The authors conclude from the results obtained in this study and according to SARS-CoV-2 Data Hub, NCBI virus assembly GCF_009858895.2, with nucleotide accession NC_045512.2 have been reported 9.191.725 nucleotides, 48.172.028 proteins by NCBI taxonomy updates to virus classification (https://www.ncbi.nlm.nih.gov/sars-cov-2/, accessed on 17 February 2026) has been contributing to the elucidation of the evolutionary dynamics of the virus. So, the emergence and reemergence of viral diseases may be accompanied by the genomic and epidemiology surveillance to mitigate any risk of propagation of the novel possible variant viral. This study dataset was a backward-looking for circulating genomic variants and genetic engineering for the development of biotechnological platforms vaccines for SARS-CoV-2 candidates licensed or under development associated others respiratory viral infections. Advanced genomic tools such as Next-Generation Sequencing (NGS) combined with awareness of three-dimensional cell cultures as viral organoids mimicking a pathophysiological environment may be key to the mechanism of new pharmaceutical molecules as well as to the development of new vaccines. These modern techniques require manipulation in sterile environments free from microbiological contamination, such as the use of isolators and restricted access barrier systems in aseptic processing.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare no conflict of interest.

References

[1] Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species severe acute respiratory syndrome related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. 2020, 5, 536–544. https://doi.org/10.1038/s41564-020-0695-z

[2] Woo et al., ICTV Virus Taxonomy Profile: Coronaviridae 2023, Journal of General Virology (2023) 104, 001843. https://doi.org/10.1099/jgv.0.001843

[3] Simões, R.S.Q et al., 2025. Novel Genomic Variants Sars-Cov-2 and blood glucose in the post-pandemic period: Retrospective epidemiology study. International Journal of Science and Technology Research Archive, 2025, 08(01), 059-068. DOI: 10.53771/ijstra.2025.8.1.0022

[4] Simões, R.S.Q. Animal and Human Coronavirus Disease. COJ Technical & Scientific Research. 2(5). COJTS.000547. 2020. DOI. 10.31031/COJTS.2020.02.000547

[5] Simões, R.S.Q.; et al. Classical and Next-Generation Vaccine Platforms to SARS-CoV-2: Biotechnological Strategies and Genomic Variants. Int. J. Environ. Res. Public Health 2022, 19, 2392. https://doi.org/10.3390/ijerph19042392