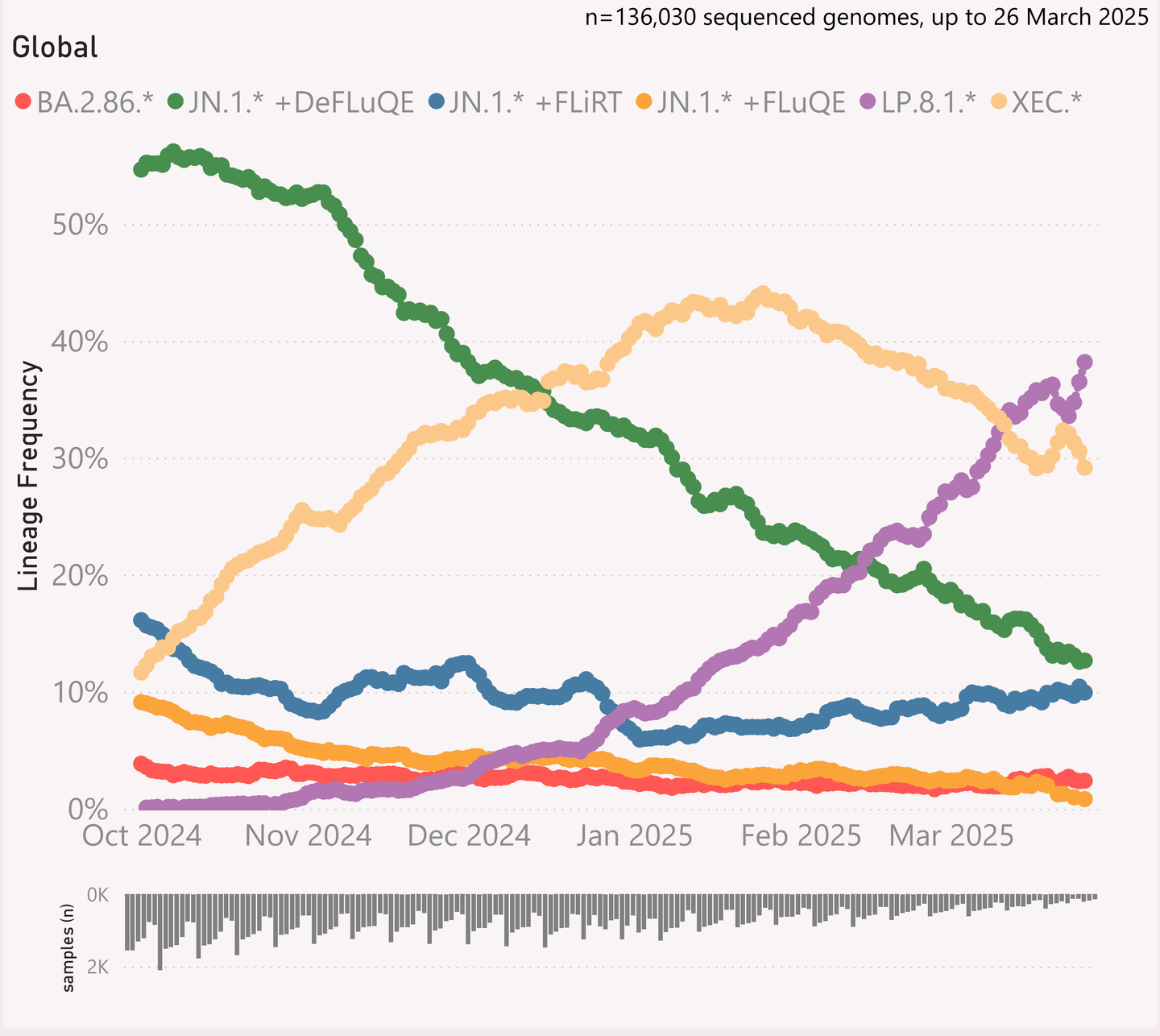
COVID-19 Genomic Sequencing analysis. Lineage Frequencies. Data from gisaid.org, with nextclade lineages. Source files and more info on github. Last Update: 2025-04-12 07:37 (UTC)

n=136,030 sequenced genomes, up to 26 March 2025



This page shows the frequency of the top 6 "L2" lineages, across recent months.

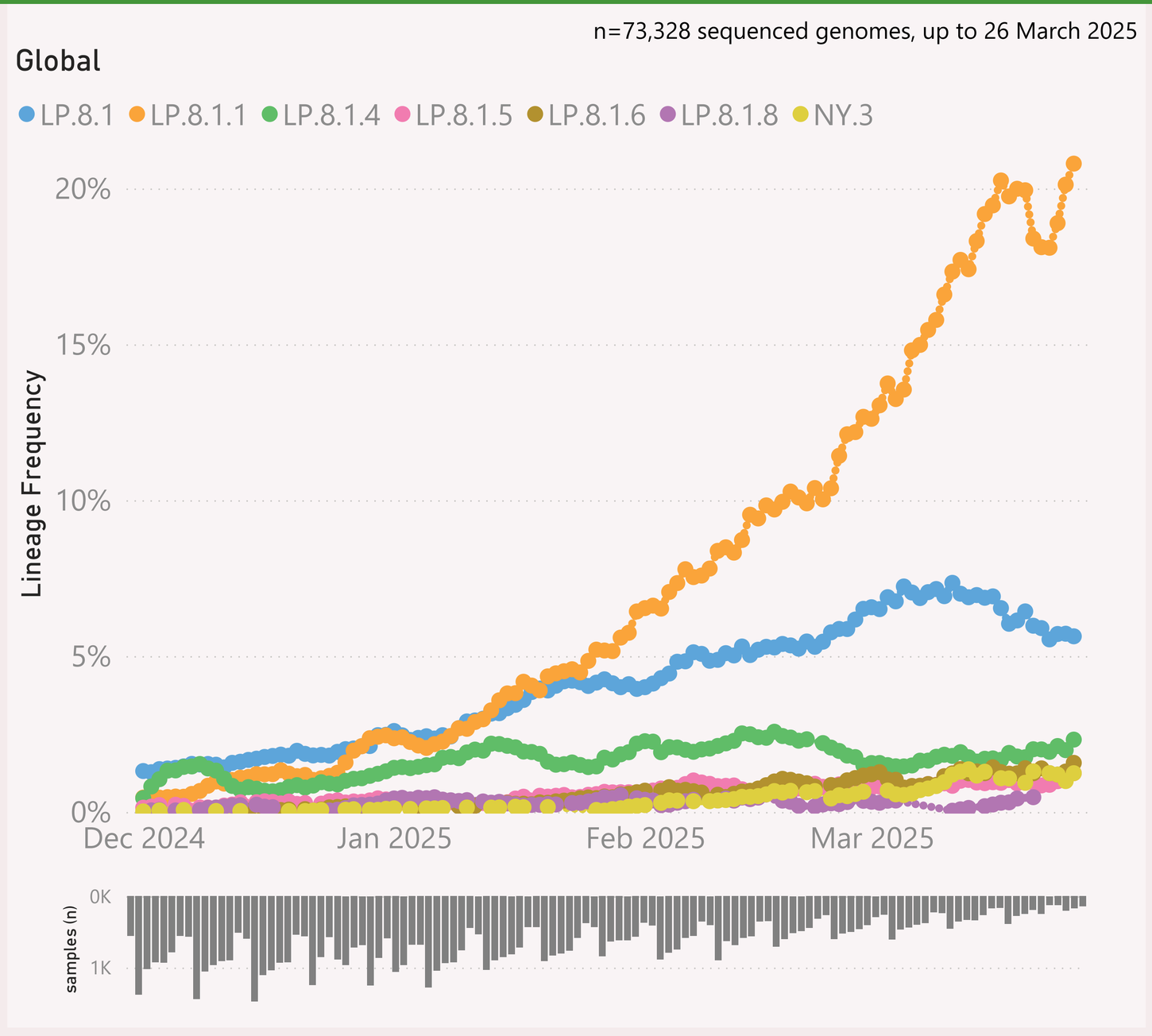
The detailed Lineage classifications are provided by Nextclade. I roll those up into "L2" groups, which roughly follow the WHO Variant definitions. For example, my "BA.2.86.*" group includes BA.2.86 and all it's descendants, e.g. the JN.* lineages.

The detailed Lineage classifications are quite numerous and dynamic, so the "Lineage L2" groups give a simpler and more stable basis for analysis and comparison.

The frequency shown at each point is based on the 7-day rolling average across all lineages.

The grey column chart across the bottom shows the volume of sequences available by date. As there can be long sample and data processing times, it is quite routine for recent dates to show lower sample sizes.

The frequency results calculated for the most recent dates might not be representative, due to those lower sample sizes.



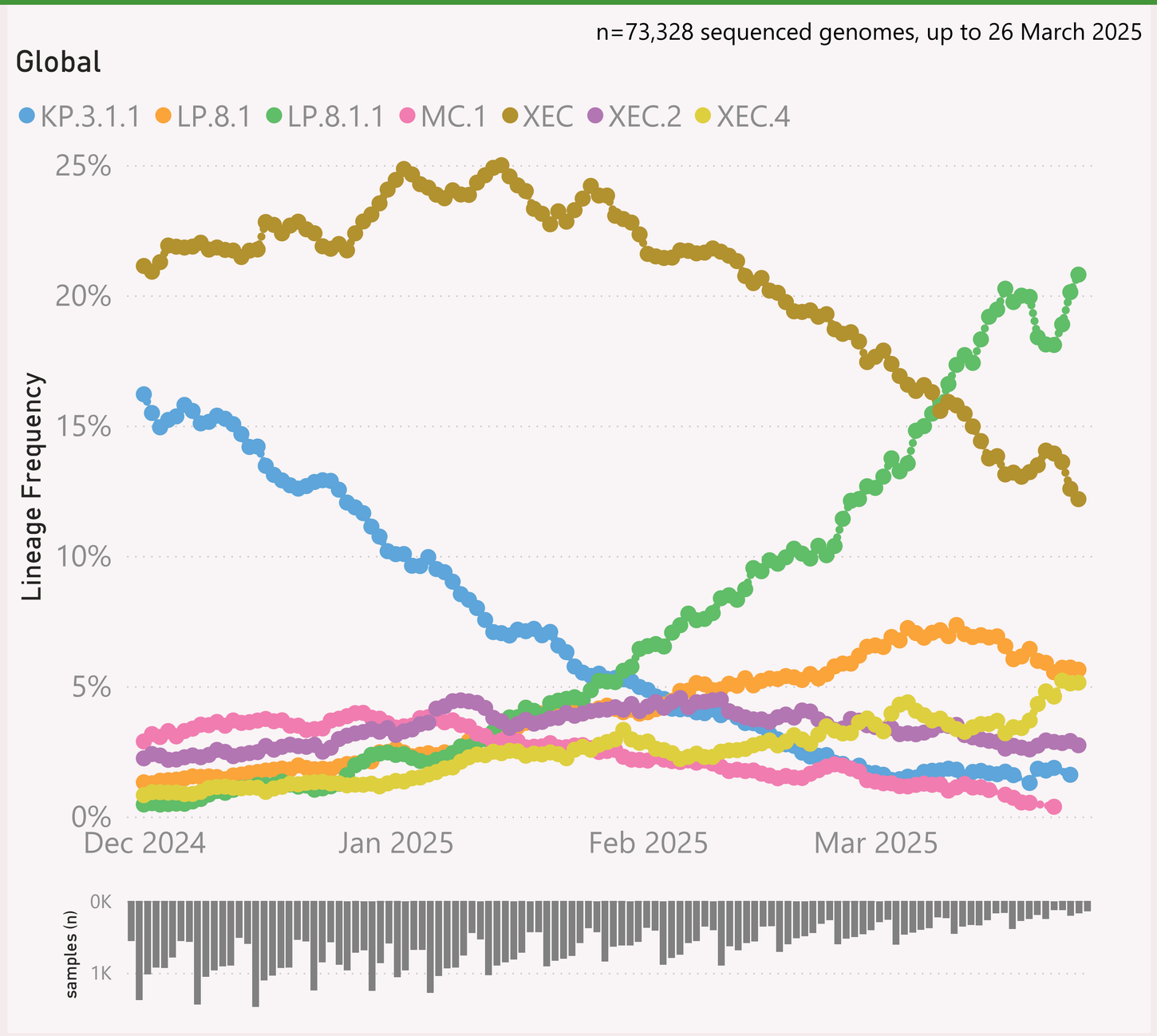
This page shows the frequency of the top 7 lineages, across recent months. The lineages are filtered for a "Lineage L2" group of interest, currently LP.8.1.*.

The Lineage classifications are provided by Nextclade. The colour assignments are random.

The frequency shown at each point is based on the 7-day rolling average across all lineages.

The grey column chart across the bottom shows the volume of sequences available by date. As there can be long sample and data processing times, it is quite routine for recent dates to show lower sample sizes.

The frequency results calculated for the most recent dates might not be representative, due to those lower sample sizes.



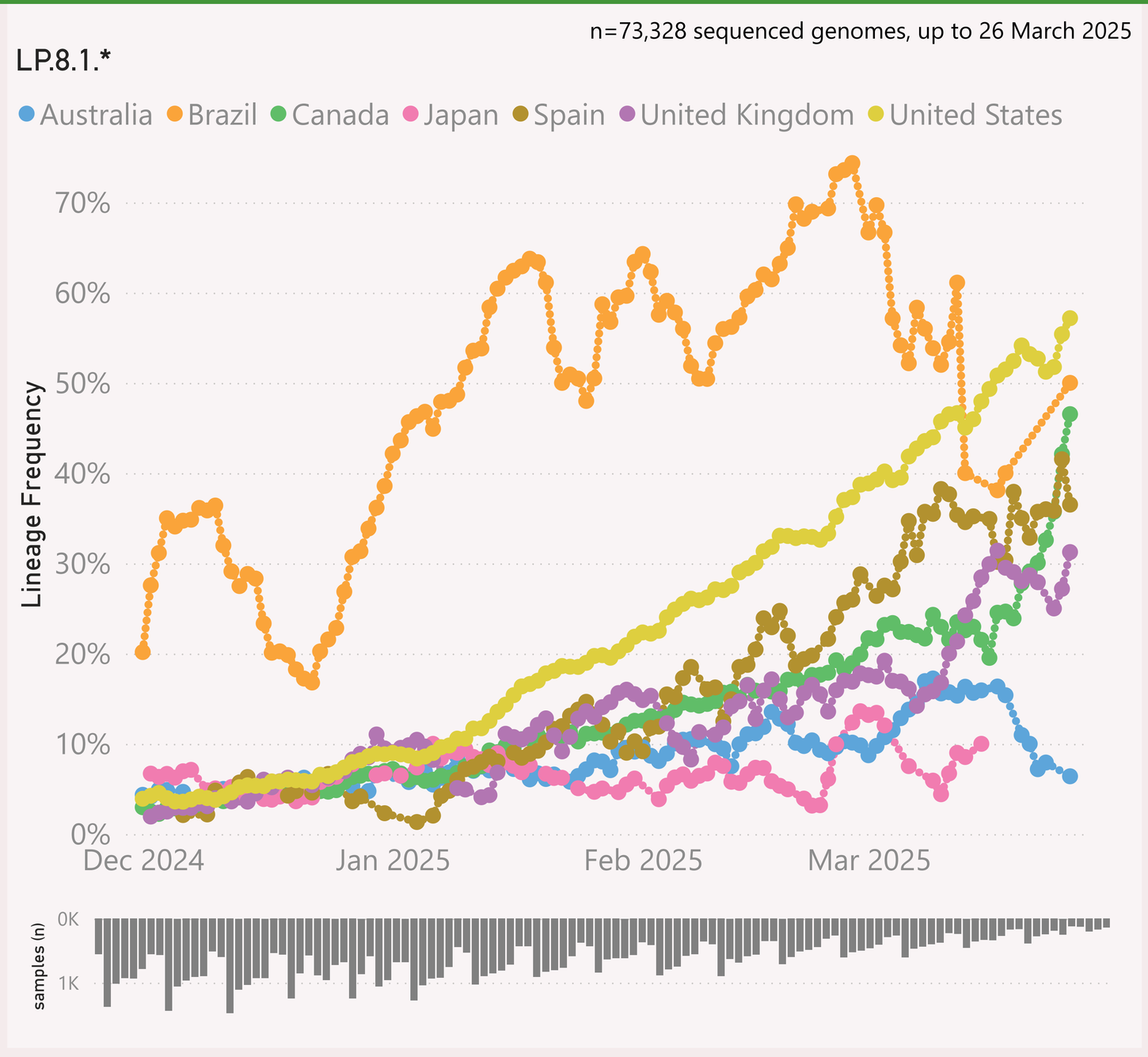
This page shows the frequency of the top 7 lineages, across recent months.

The Lineage classifications are provided by Nextclade. The colour assignments are random.

The frequency shown at each point is based on the 7-day rolling average across all lineages.

The grey column chart across the bottom shows the volume of sequences available by date. As there can be long sample and data processing times, it is quite routine for recent dates to show lower sample sizes.

The frequency results calculated for the most recent dates might not be representative, due to those lower sample sizes.



This page shows the frequency of a selected "Lineage L2" group of interest, for the 7 countries reporting the most samples over recent months.

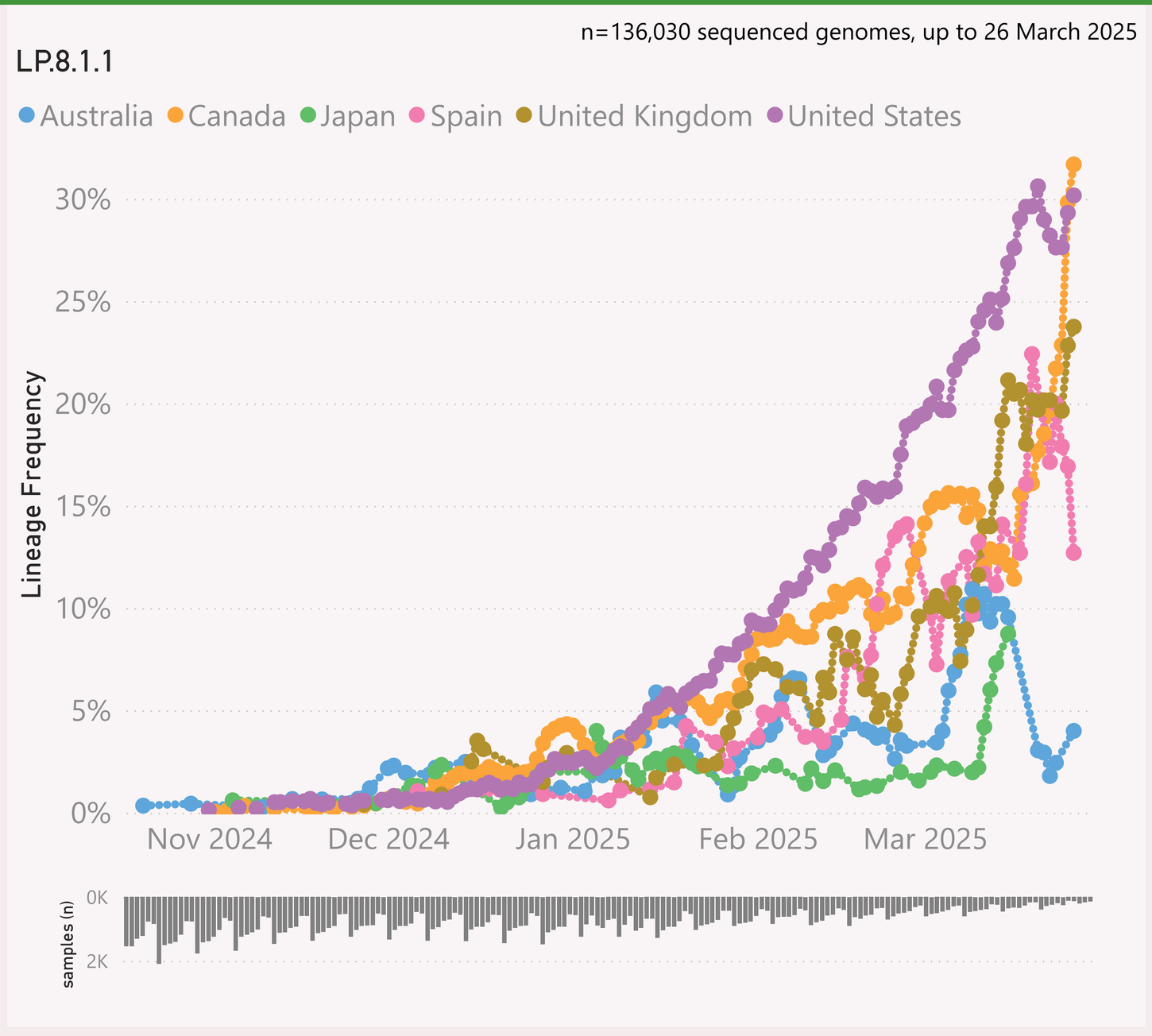
The detailed Lineage classifications are provided by Nextclade. I roll those up into "L2" groups, which roughly follow the WHO Variant definitions. For example, my "JN.1.* +FLiRT" group includes the descendants of JN.1.* with the mutations: F456L & R346T.

The detailed Lineage classifications are quite numerous and dynamic, so the "Lineage L2" groups give a simpler and more stable basis for analysis and comparison.

The frequency shown at each point is based on the 7-day rolling average across all lineages, for that country.

The grey column chart across the bottom shows the volume of sequences available by date. As there can be long sample and data processing times, it is quite routine for recent dates to show lower sample sizes.

The frequency results calculated for the most recent dates might not be representative, due to those lower sample sizes.



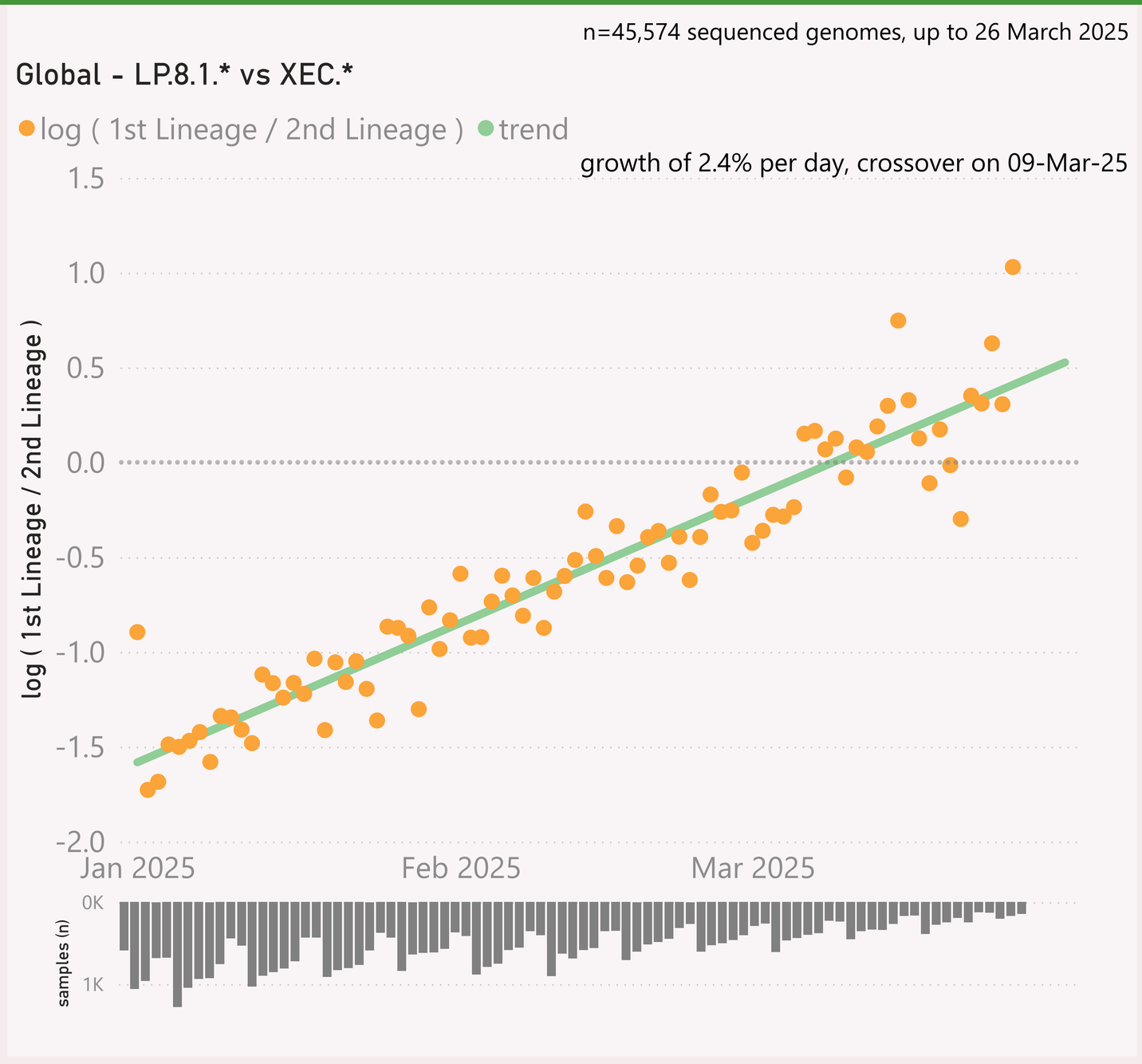
This page shows the frequency of a selected Lineage of interest, for the 6 countries reporting the most samples over recent months.

The Lineage classifications are provided by Nextclade.

The frequency shown at each point is based on the 7-day rolling average across all lineages, for that country.

The grey column chart across the bottom shows the volume of sequences available by date. As there can be long sample and data processing times, it is quite routine for recent dates to show lower sample sizes.

The frequency results calculated for the most recent dates might not be representative, due to those lower sample sizes.

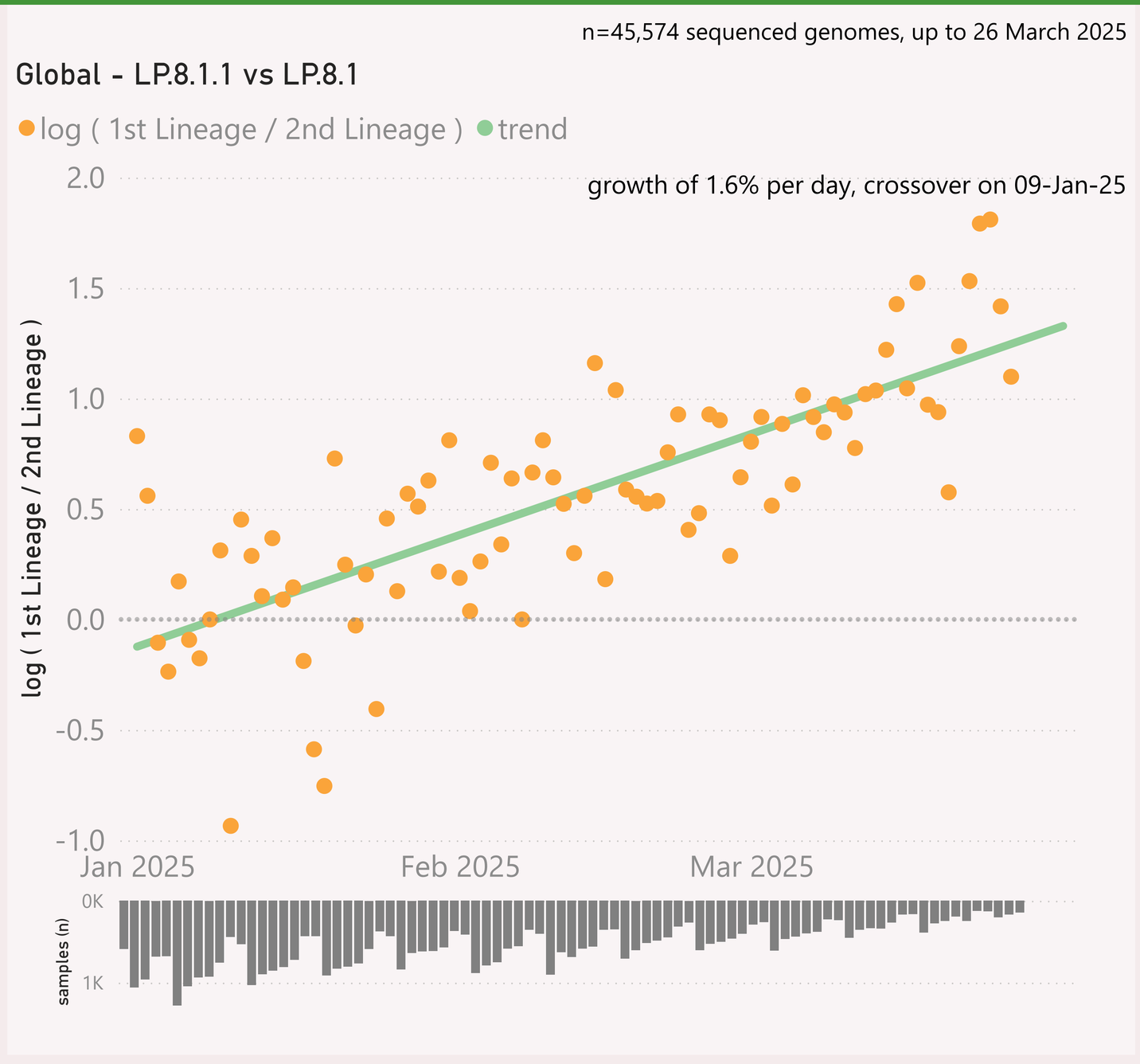


This page compares the relative frequency of 2 selected "Lineage L2" groups, over recent months. A challenging Lineage L2 is selected first, and compared to the incumbent.

The trend is shown as a green line and expressed as a daily growth % advantage. If the green line crosses over the 0.0 line, the date when that occurred or is predicted to occur will be shown. At that point the challenging Lineage L2 is considered to have "crossed over" or taken over dominance from the incumbent Lineage L2.

The Lineage classifications are provided by Nextclade. I add the "Lineage L2" groups, typically following common variant groupings, but occasionally being "creative".

The grey column chart across the bottom shows the volume of sequences available by date. As there can be long sample and data processing times, it is quite routine for recent dates to show lower sample sizes.



This page compares the relative frequency of 2 selected Lineages, over recent months. A challenging Lineage is selected first, and compared to the incumbent.

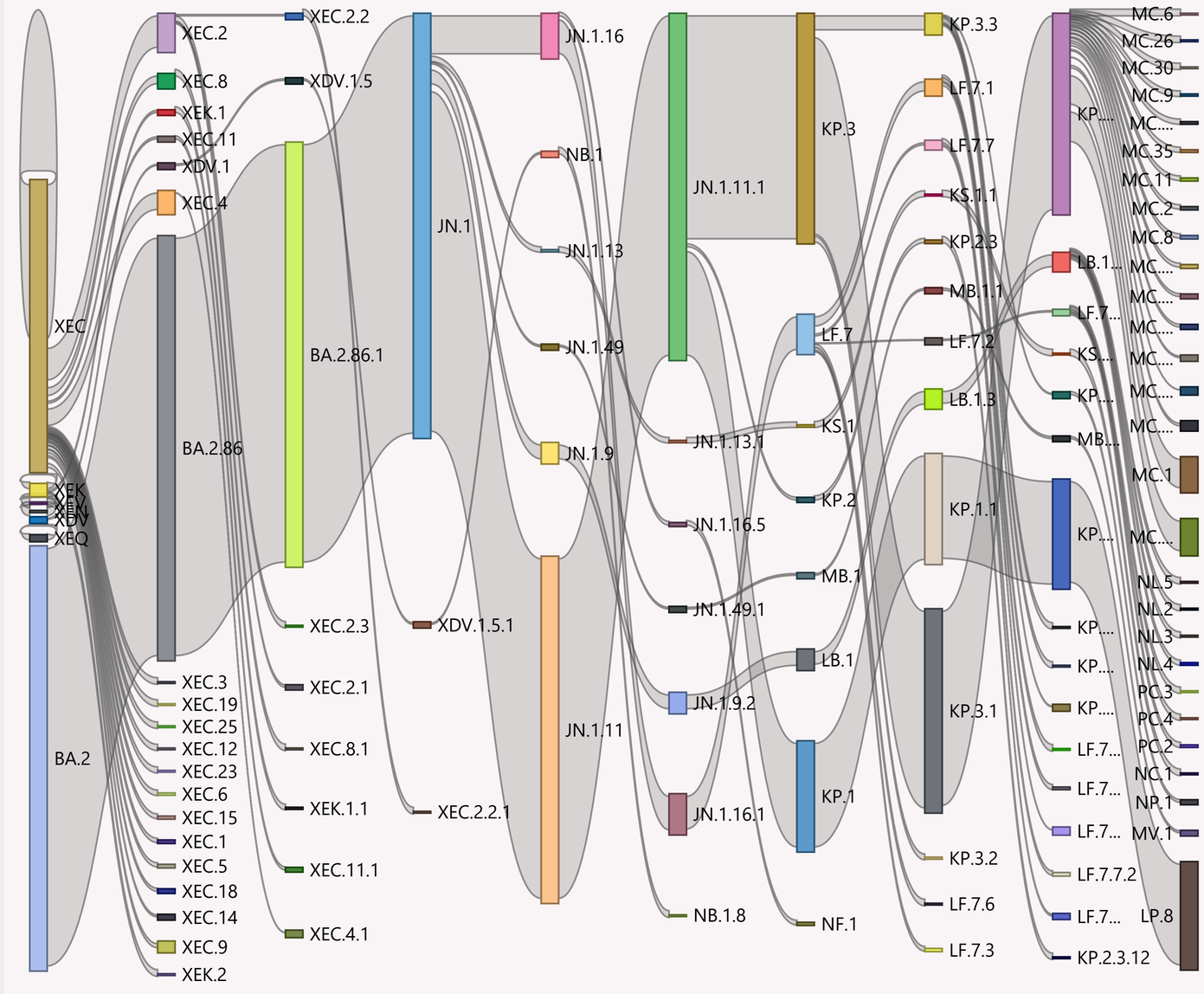
The trend is shown as a green line and expressed as a daily growth % advantage. If the green line crosses over the 0.0 line, the date when that occurred or is predicted to occur will be shown. At that point the challenging Lineage is considered to have "crossed over" or taken over dominance from the incumbent Lineage

The Lineage classifications are provided by Nextclade.

The grey column chart across the bottom shows the volume of sequences available by date. As there can be long sample and data processing times, it is quite routine for recent dates to show lower sample sizes.

n=73,476 sequenced genomes, up to 26 March 2025

Global



This page shows the hierarchy of the significant Lineages, over recent months.

The hierarchy can be read from left to right, starting with the earliest/highest Lineages being broken down into more detailed child Lineages.

The vertical height of each bar segment represents the relative volume of all the samples of that specific Lineage, as well as all it's descendants.

The full picture is typically quite busy, so insignificant Lineages (with few samples, or at the extreme top or bottom of the hierarchy) are not shown.

The Lineage classifications are provided by Nextclade.

COVID-19 Genomic Sequencing analysis - Data Currency. Data from gisaid.org. Source files and info on github.

Last Update: 2025-04-12 07:37 (UTC)

Data Submitted in the last 8 weeks

Country	# Samples Sequenced	Latest Collection date	by Collection date	Latest Submission date	by Submission date
United States	18,089	26/03/2025		29/03/2025	
Canada	3,972	26/03/2025		29/03/2025	
Japan	2,565	26/03/2025		29/03/2025	
Australia	1,829	26/03/2025		29/03/2025	
Brazil	1,637	26/03/2025		29/03/2025	
Spain	1,314	26/03/2025		29/03/2025	
United Kingdom	1,092	26/03/2025		29/03/2025	
France	625	26/03/2025		29/03/2025	
Chile	566	09/03/2025		17/03/2025	
Germany	482	20/03/2025		29/03/2025	
Singapore	426	26/03/2025		29/03/2025	
Russia	422	19/03/2025		29/03/2025	
Slovenia	336	24/03/2025		29/03/2025	
Denmark	326	24/03/2025		29/03/2025	
Greece	326	01/03/2025		21/03/2025	
South Korea	322	20/03/2025		29/03/2025	
New Zealand	320	25/03/2025		29/03/2025	
Italy	284	25/03/2025		29/03/2025	
Peru	267	10/01/2025		29/03/2025	
China	259	25/03/2025		29/03/2025	
Luxembourg	252	14/02/2025		29/03/2025	
Sweden	245	19/03/2025		29/03/2025	
Ghana	211	19/12/2024		05/03/2025	
Austria	184	27/01/2025		24/02/2025	
Norway	175	21/03/2025		29/03/2025	
Ireland	156	24/03/2025		28/03/2025	
Netherlands	153	25/03/2025		29/03/2025	
South Africa	135	08/03/2025		28/03/2025	
Total	**38,781**	**26/03/2025**		**29/03/2025**	

This page shows the volume and currency/timeliness of the genomic sequencing data shared via GISAID, over the last 8 weeks, for the countries sharing the most samples.

Each sample shared comes with a Collection date - when the PCR test for that sample was collected. The GISAID system also records a Submission date for each sample, which is typically the date that sample was uploaded.

The latest date of each type is shown, along with "sparkline"-style mini charts to give a flavour for the spread of recent data by Collection date and by Submission date.