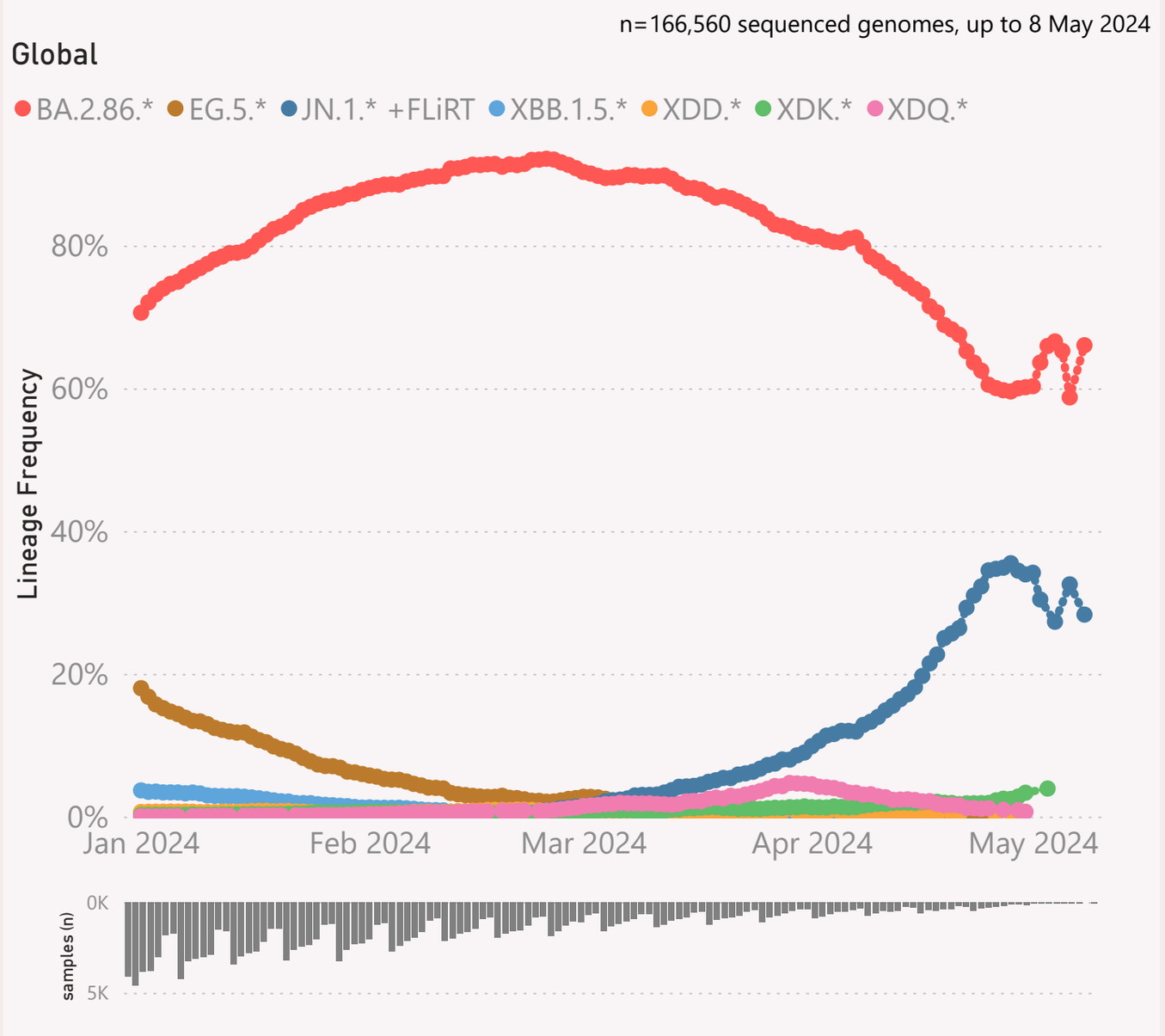
COVID-19 Genomic Sequencing analysis. Lineage Frequencies. Data from gisaid.org, with nextclade lineages. Source files and more info on github. Last Update: 2024-05-11 02:31 (UTC)



This page shows the frequency of the top 7 "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.

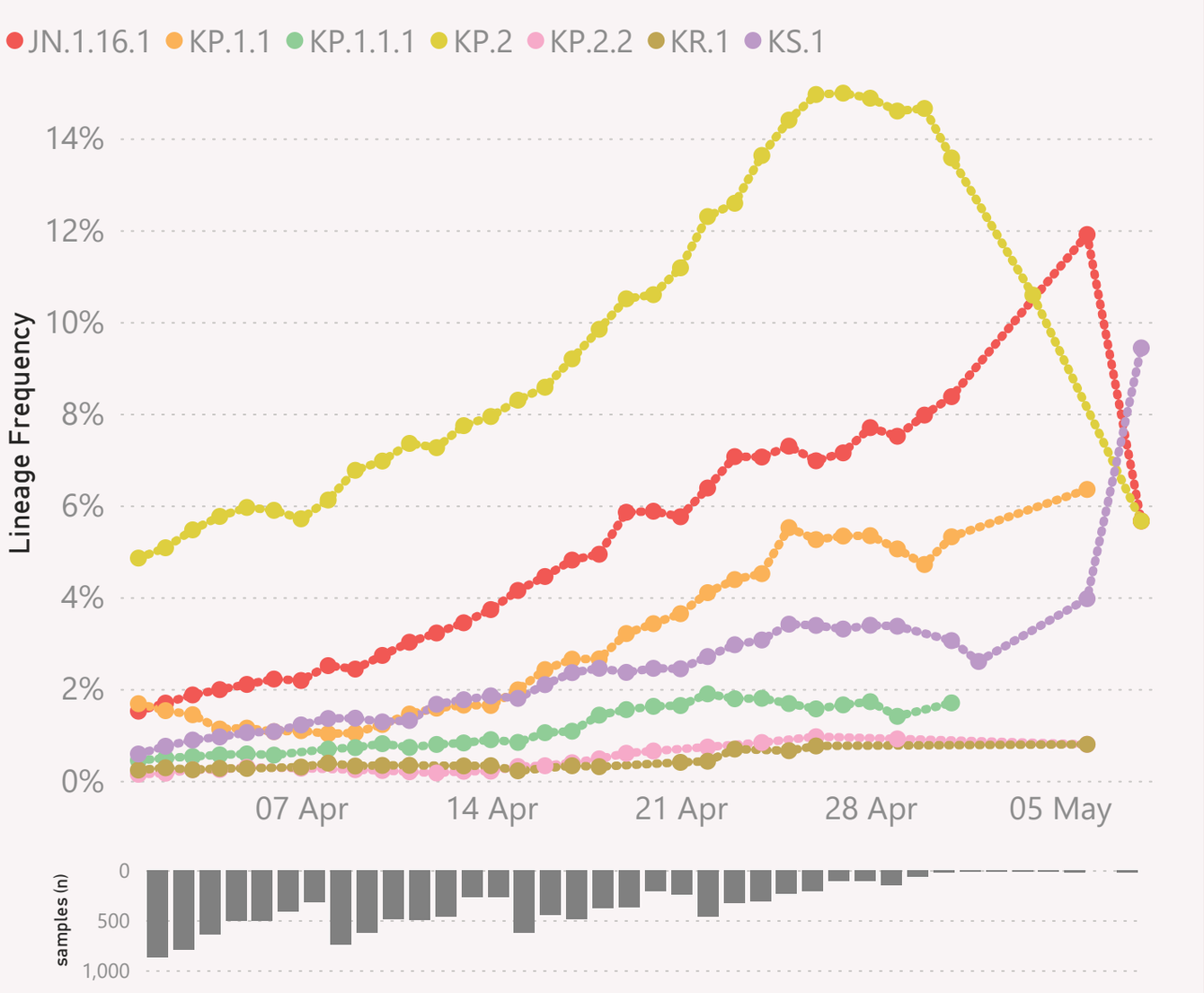
The detailed Lineage classifications are quite numerous and dynamic, so the "Lineage L2" groups give a simpler and more stable basis for analysis and comparison.

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.

n=12,159 sequenced genomes, up to 8 May 2024



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 "JN.1.* + FLiRT".

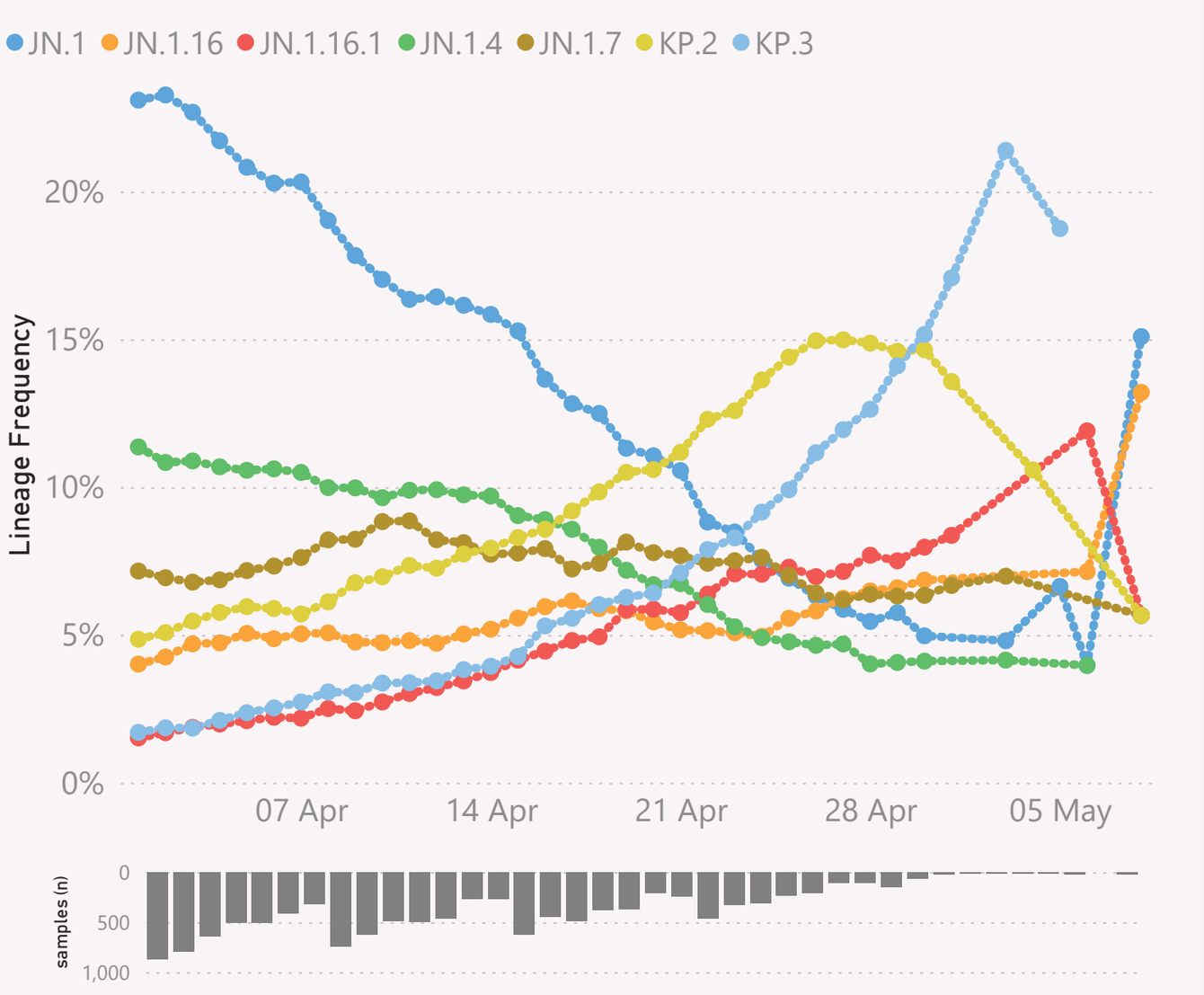
The Lineage classifications are provided by Nextclade. The colour assignments are random.

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.

n=12,159 sequenced genomes, up to 8 May 2024



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

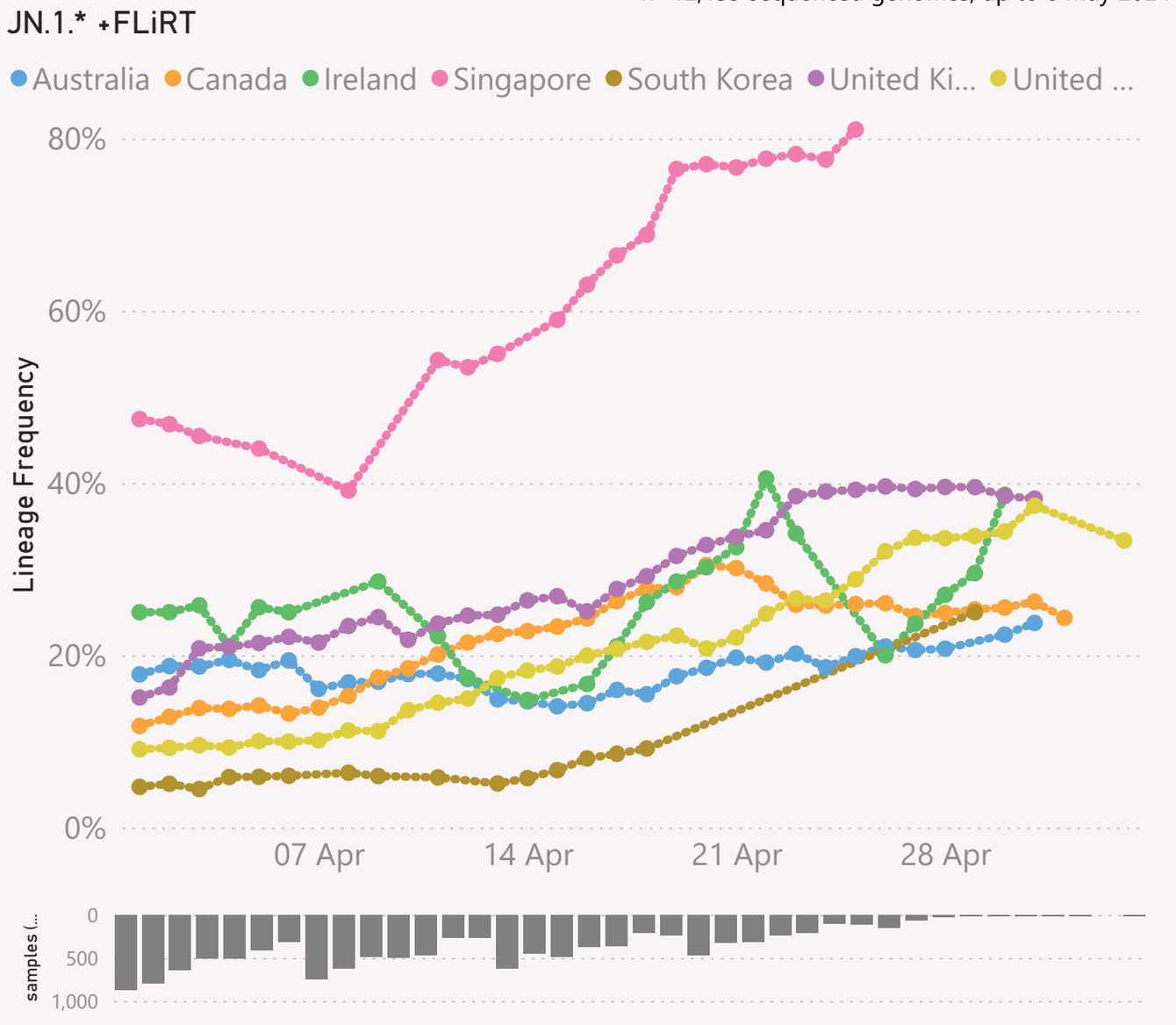
The Lineage classifications are provided by Nextclade. The colour assignments are random.

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.

n=12,159 sequenced genomes, up to 8 May 2024



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 "BA.2.86.*" group includes BA.2.86 and all it's descendants, e.g. the JN.* lineages.

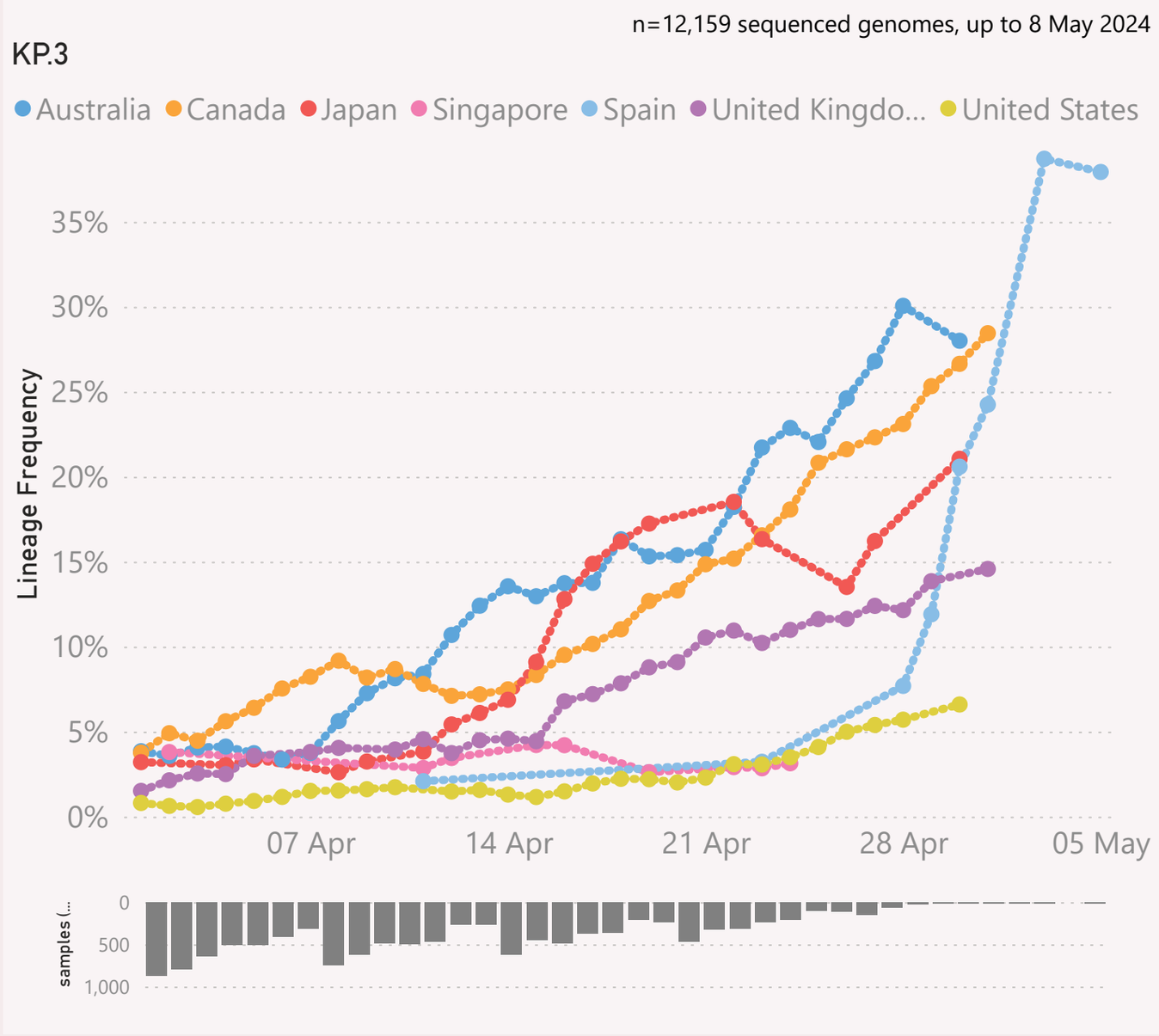
The detailed Lineage classifications are quite numerous and dynamic, so the "Lineage L2" groups give a simpler and more stable basis for analysis and comparison.

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

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.

COVID-19 Genomic Sequencing analysis - selected Lineage by Country. Data from gisaid.org, with nextclade lineages Source files and info on github. Last Update: 2024-05-11 02:31 (UTC)



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

The Lineage classifications are provided by Nextclade.

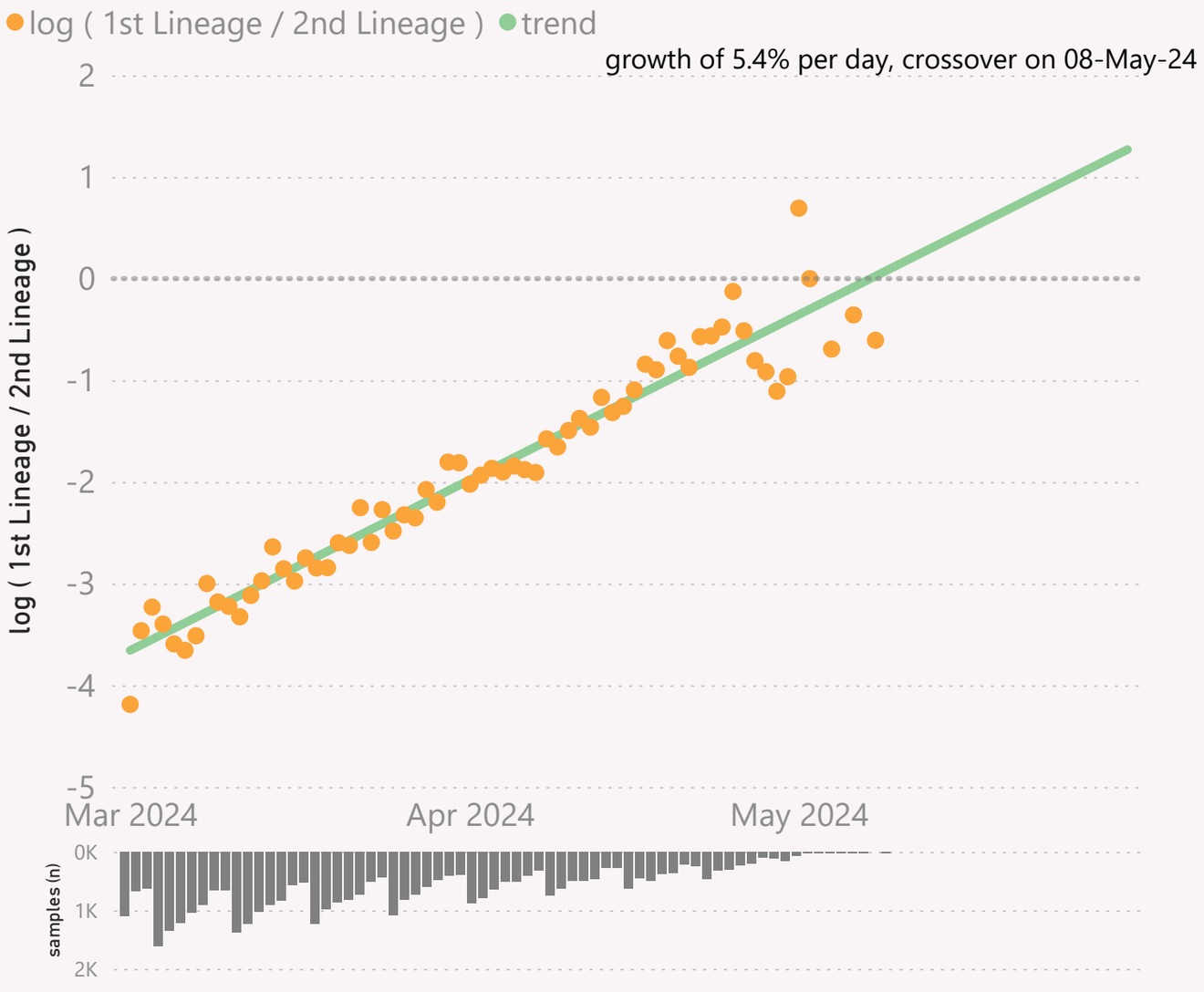
The frequency shown at each point is based on the 7-day rolling average across all lineages, for that state.

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.

n=38,434 sequenced genomes, up to 8 May 2024

Global - JN.1.* +FLiRT vs BA.2.86.*



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.

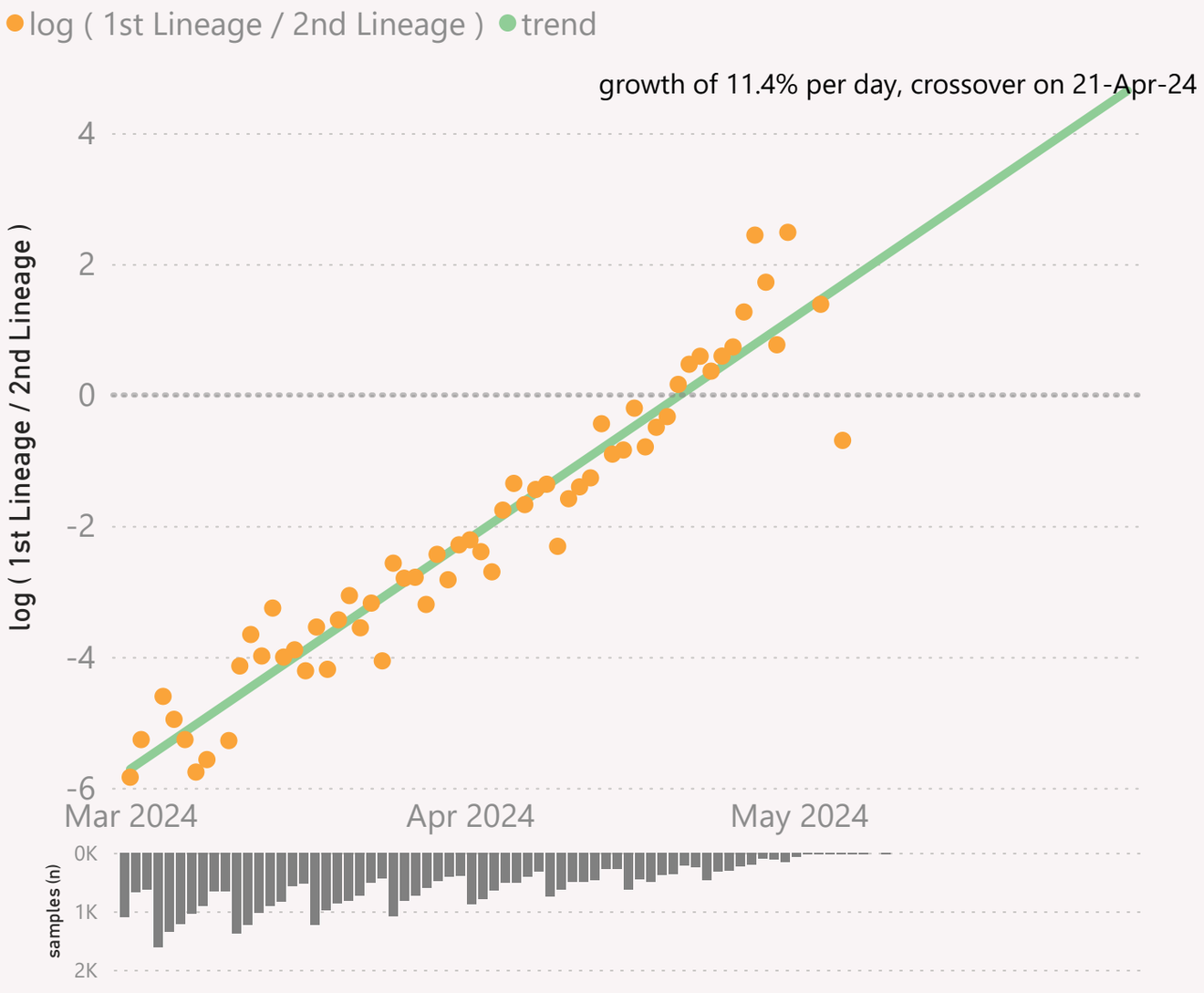
The trend is shown as a green line and expressed as a daily growth % advantage. If the green line crosses over the 0.0 line, the date when that occurred or is predicted to occur will be shown. At that point the challenging Lineage L2 is considered to have "crossed over" or taken over dominance from the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

n=38,434 sequenced genomes, up to 8 May 2024

Global - KP.3 vs JN.1



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

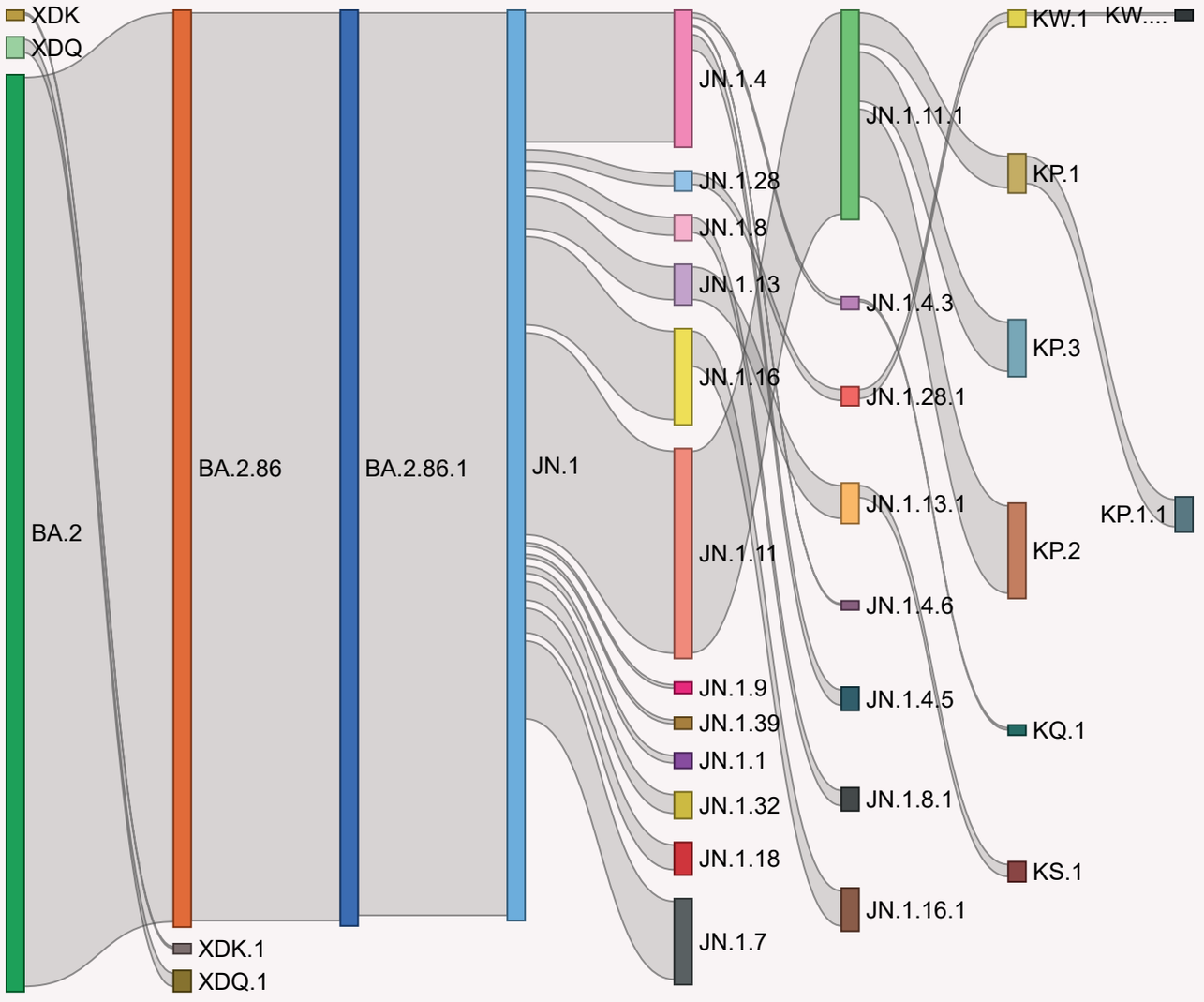
The Lineage classifications are provided by Nextclade.

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.

COVID-19 Genomic Sequencing sankey diagram. Data from gisaid.org, with nextclade lineages. Source files and more info on github. Last Update: 2024-05-11 02:31 (UTC)

n=12,159 sequenced genomes, up to 8 May 2024

Global



Show most significant labels

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: 2024-05-11 02:31 (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	24,378	4/05/2024		9/05/2024	
Japan	11,385	1/05/2024		9/05/2024	
Canada	5,778	2/05/2024		9/05/2024	
China	5,391	24/04/2024		8/05/2024	
South Korea	4,514	29/04/2024		9/05/2024	
United Kingdom	3,794	1/05/2024		9/05/2024	
Australia	2,302	1/05/2024		8/05/2024	
Brazil	1,862	15/04/2024		7/05/2024	
Spain	1,614	6/05/2024		9/05/2024	
Mexico	1,525	3/04/2024		7/05/2024	
Chile	1,043	17/04/2024		25/04/2024	
Denmark	1,019	22/04/2024		9/05/2024	
Russia	995	23/04/2024		2/05/2024	
India	831	11/04/2024		23/04/2024	
Lithuania	752	14/03/2024		25/04/2024	
Singapore	611	25/04/2024		6/05/2024	
Turkey	604	3/03/2024		28/03/2024	
New Zealand	586	2/04/2024		22/04/2024	
France	505	22/04/2024		7/05/2024	
Puerto Rico	488	19/04/2024		1/05/2024	
Ukraine	441	9/04/2024		3/05/2024	
Germany	432	19/04/2024		7/05/2024	
Netherlands	417	21/04/2024		6/05/2024	
Thailand	401	25/04/2024		9/05/2024	
Ireland	363	30/04/2024		9/05/2024	
Taiwan	353	8/05/2024		9/05/2024	
Italy	334	27/04/2024		9/05/2024	
Malaysia	300	18/03/2024		7/05/2024	
Total	**77,335**	**8/05/2024**		**9/05/2024**	

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.