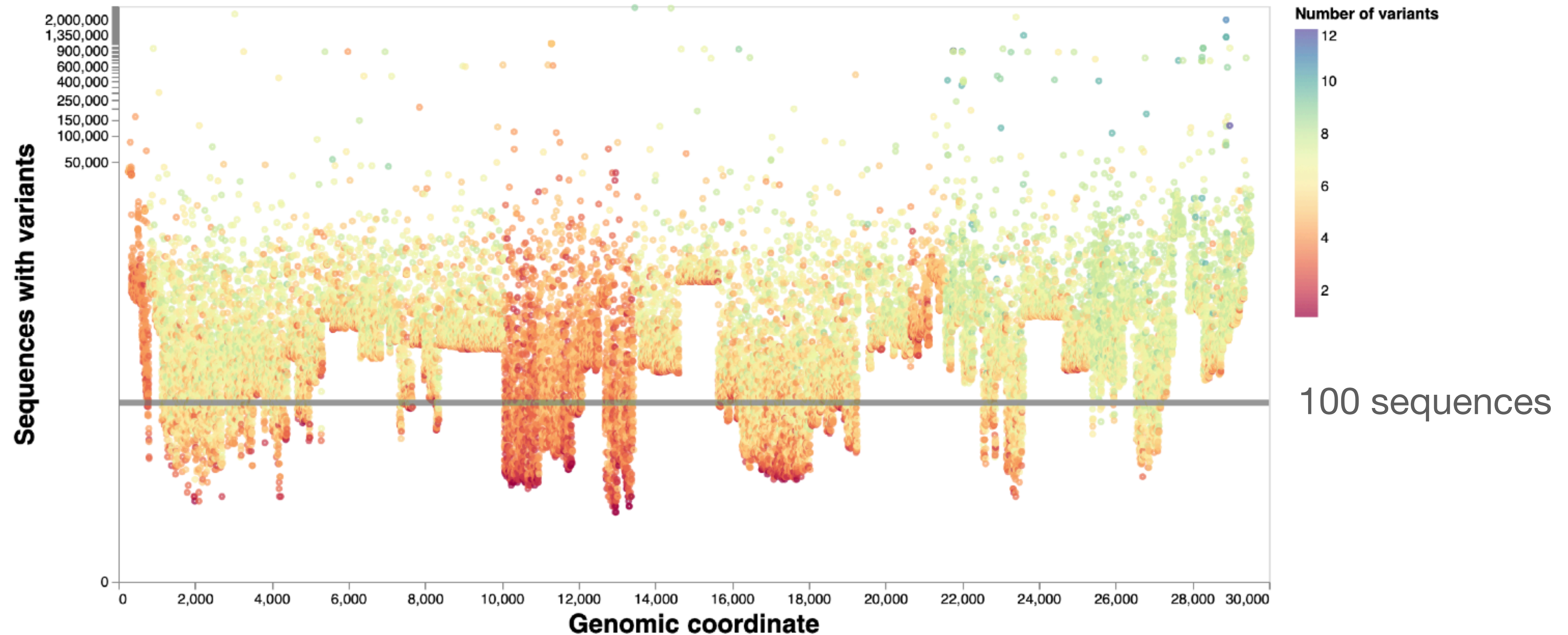


How much genomic variation is there in SARS-CoV-2?

- There are >4M genomes in GISAID and >1.3M NGS datasets in SRA for SARS-CoV-2
- At this point, nearly every genomic position has sequences with multiple allelic variants (AVs)



Which AVs are interesting and non-obvious?

- Most variation exists at low ($<1\%$) frequency.
- Variants more frequent than 5% are rare (~ 100) and can be exhaustively tested.

