SARS-CoV-2 sequencing report

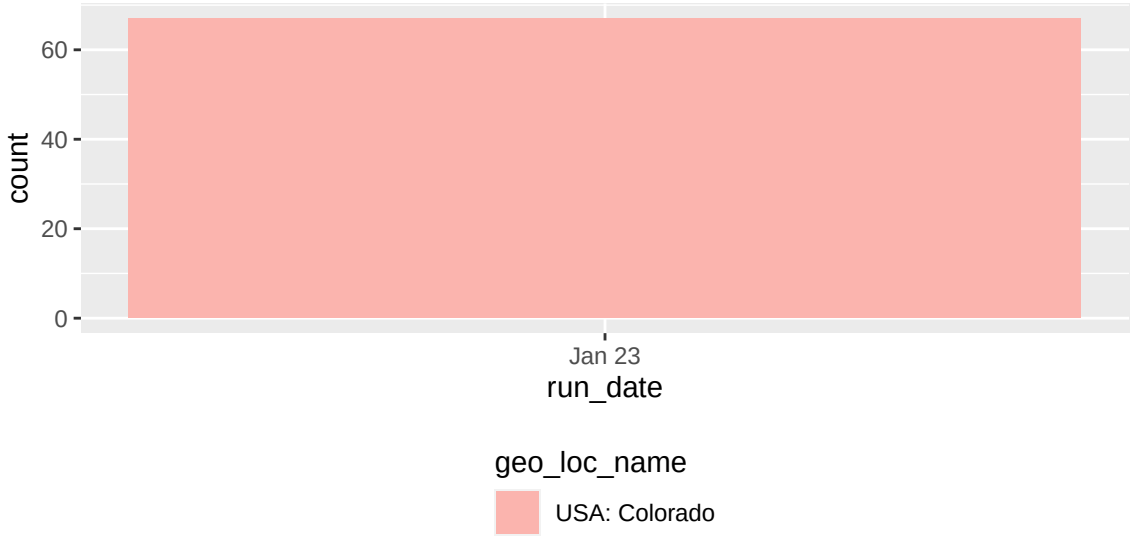
The Broad Institute Viral Genomics group, in partnership with the Genomics Platform and Data Sciences Platform, has been engaged in viral sequencing of COVID-19 patients since March 2020.

This report, generated on 2021-03-08, summarizes our sequencing activity for patient samples provided by **Colorado Mesa University**. It is current as of the epiweek ending **2021-01-23**.

Weekly summary

	week ending 2021-01-23	cumulative total
samples sequenced	67	67
genomes assembled	41	41
genomes submittable	30	30
Variants of Concern	0	0

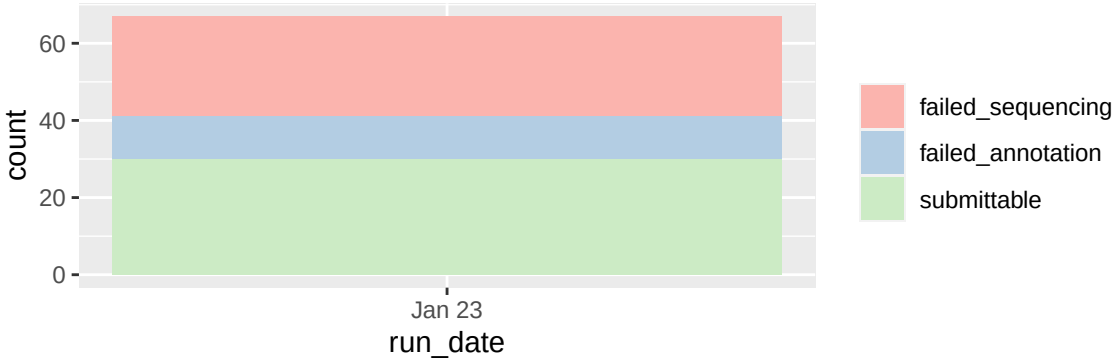
Sequencing activity over time



This describes sequencing attempts over time, broken down by geography.

Sequencing performance over time

Samples sequenced weekly



This describes the total number of patient samples sequenced in this data set, plotted by the date of the sequencing run in our lab.

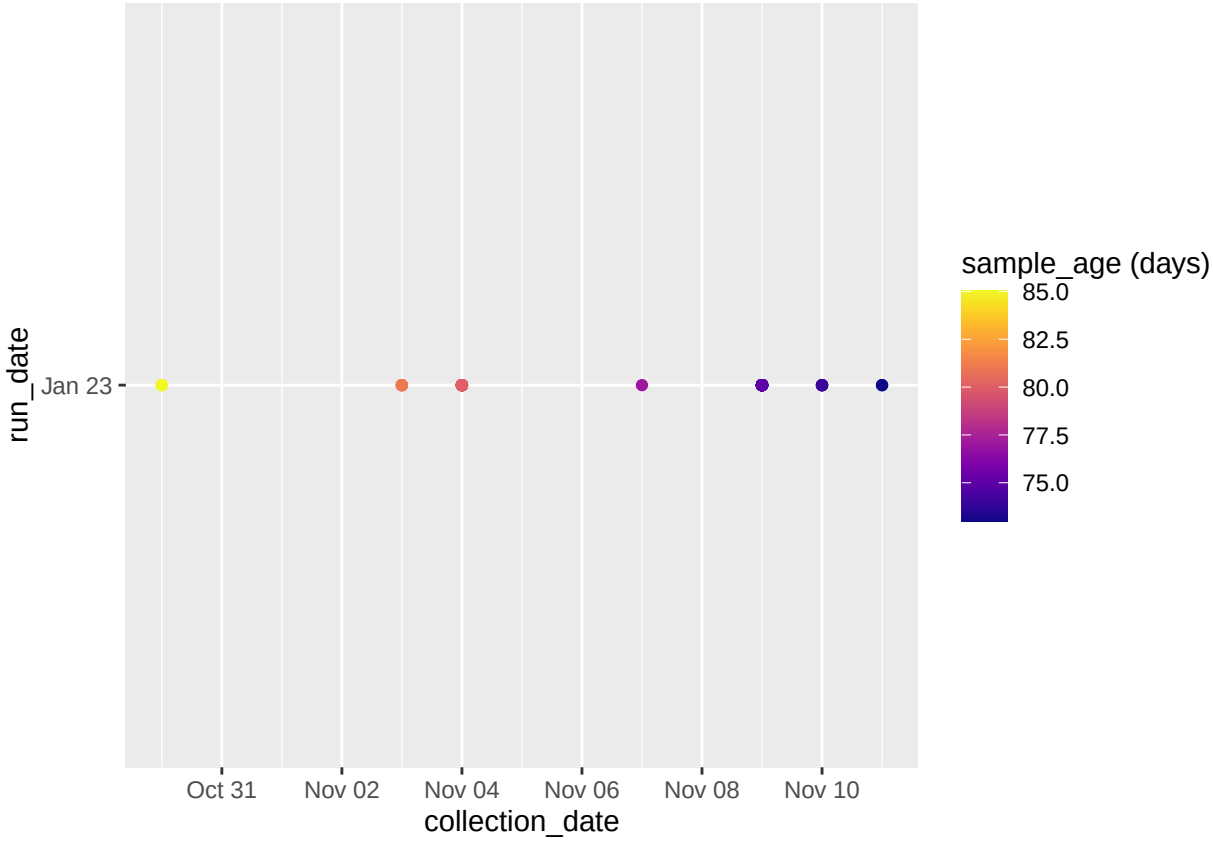
“Submittable” genomes pass all QC checks and are quickly released to public genome repositories. “Failed sequencing” are samples that failed to produce at least 24000 unambiguous base pairs of viral genome. Raw data from these samples are submitted to NCBI’s SRA database, but the genomes are not used for any analyses. “Failed annotation” are samples that produced a sufficiently complete genome, but did not pass NCBI’s VADR quality checks.

Tabular view by CDC epiweek

epiweek ending	samples sequenced	genomes assembled	genomes submittable
2021-01-23	67	41	30

Timeliness of surveillance

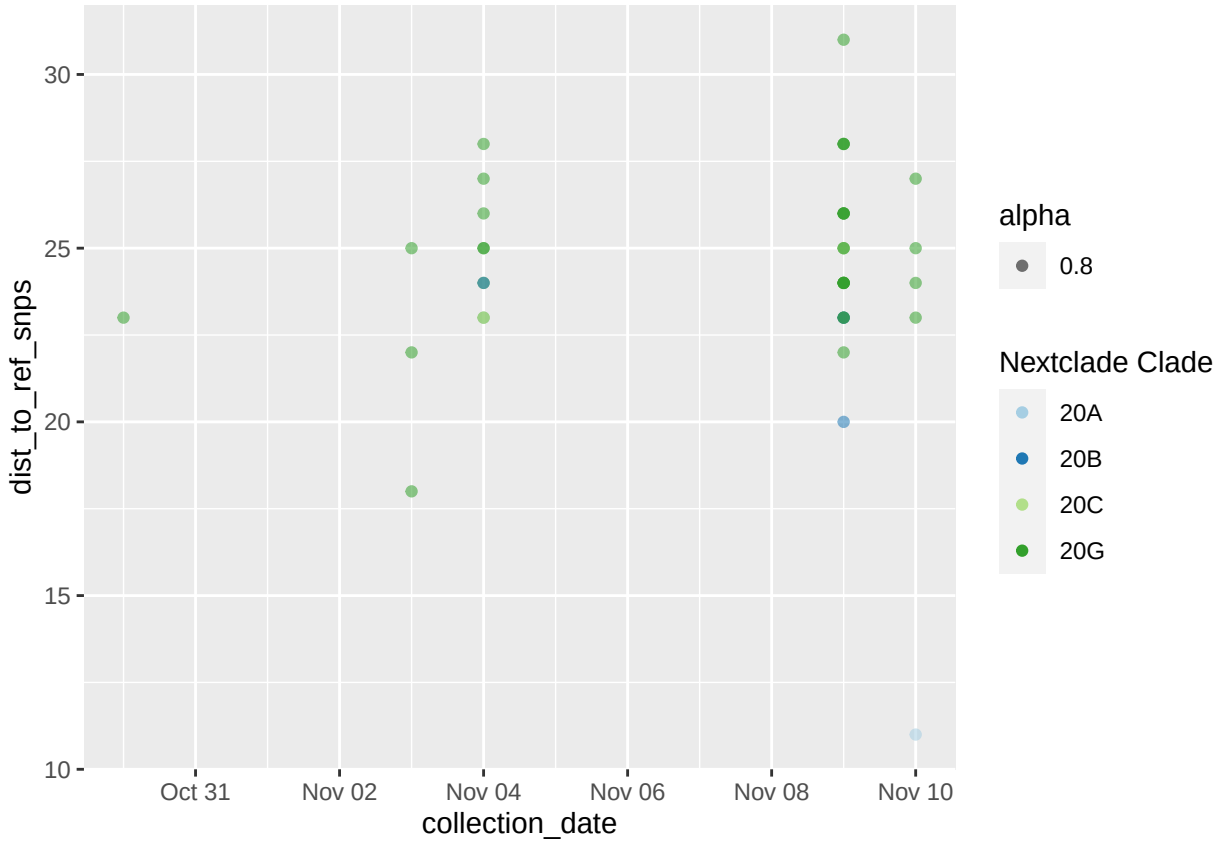
Sequencing date vs collection date



This plot describes the “timeliness” of the sequencing run for the purpose of real-time surveillance of circulating lineages and variants of interest. Note that this plot likely includes many samples that were sequenced for non-surveilance purposes.

Evolutionary Clock

Genetic distance root-to-tip vs sample collection date



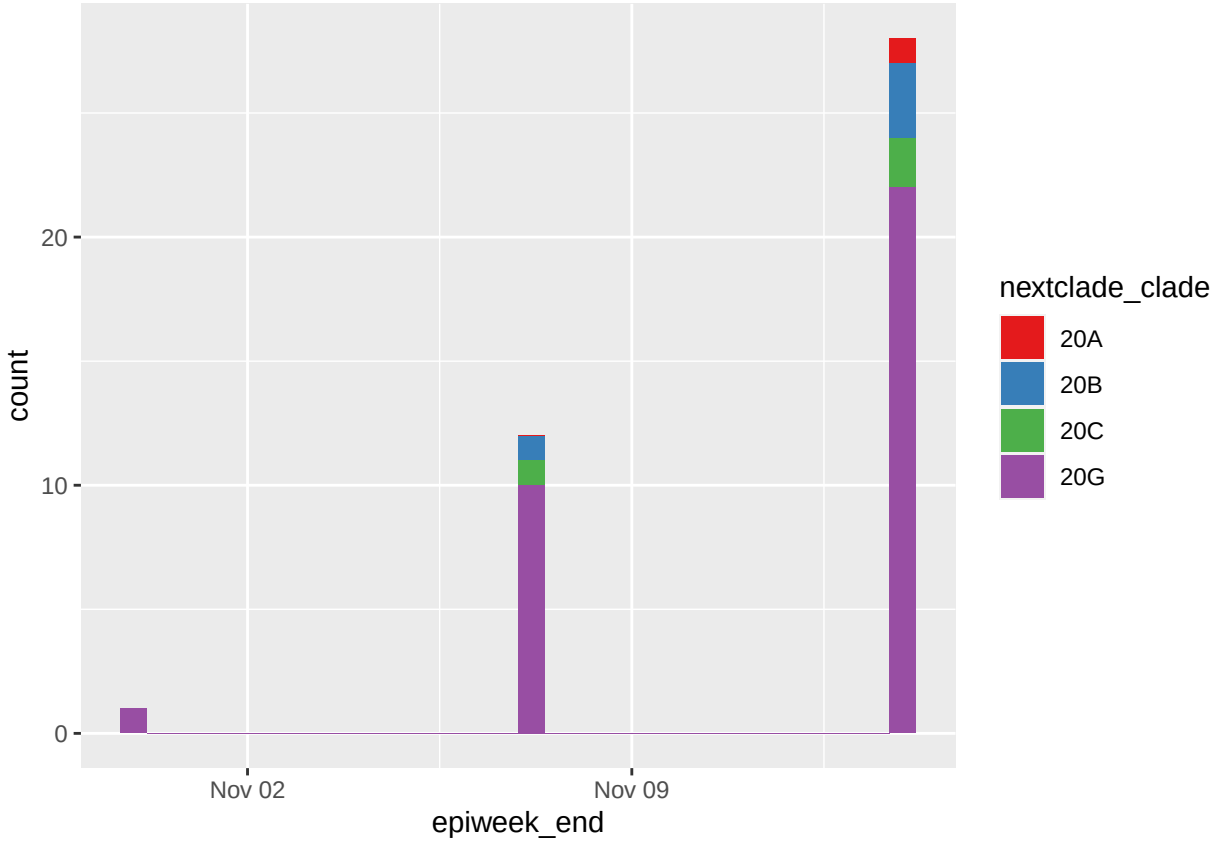
A “root-to-tip plot” plots the genetic distance of each sample from Wuhan Hu-1 against the date it was collected. It is generally somewhat linear. Outliers on this plot may be indicative of laboratory or metadata errors, or of evolutionarily unusual lineages (such as B.1.1.7).

Phylogenetic Clades and Variants

Reportable Variants of Concern (VoCs) by CDC epiweek of sample collection

Collection epiweek ending

Nextclade classifications vs sample collection date



This shows the breakdown of major phylogenetic clades over time, using the Nextclade naming system. Variants of Concern (VoCs) are highlighted as specially named Nextclade clades. Nextclade clade 20I/501Y.V1 corresponds to PANGO lineage B.1.1.7, 20H/501Y.V2 corresponds to B.1.351, and 20J/501Y.V3 corresponds to P.1.