COVID-19 subject 95-267

2021-01-19

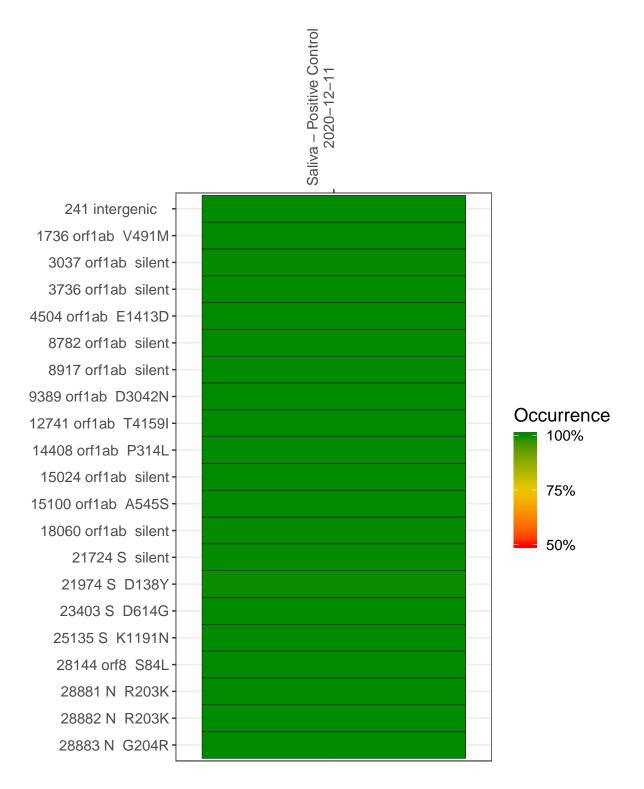
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

Experiment	Туре	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0545-1	single experiment	NA	Saliva - Positive Control	2020-12-11	22.41	98.9%	98.0%

Variants shared across samples

The heat map below shows how variants (reference genome USA-WA1-2020) are shared across subject samples where the percent variance is colored. Variants are called if a variant position is covered by 5 or more reads, the alternative base is found in > 50% of read pairs and the variant yields a PHRED score > 20. Gray tiles denote positions where the variant was not the major variant or no variants were found. The relative base compositions of each experiment used to calculate tiles are shown in the following plot where the total number of position reads are shown atop of each plot.



Saliva – Positive Control 2020–12–11

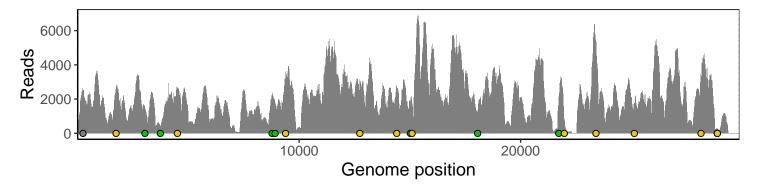
241 intergenic	2471
1736 orf1ab V491M	2717
3037 orf1ab silent	1436
3736 orf1ab silent	451
4504 orf1ab E1413D	2625
8782 orf1ab silent	2255
8917 orf1ab silent	1621
9389 orf1ab D3042N	3180
12741 orf1ab T4159I	3127
14408 orf1ab P314L	2586
15024 orf1ab silent	1910
15100 orf1ab A545S	1748
18060 orf1ab silent	1290
21724 S silent	2133
21974 S D138Y	786
23403 S D614G	5473
25135 S K1191N	1658
28144 orf8 S84L	3284
28881 N R203K	242
28882 N R203K	242
28883 N G204R	242
	SP0545-1



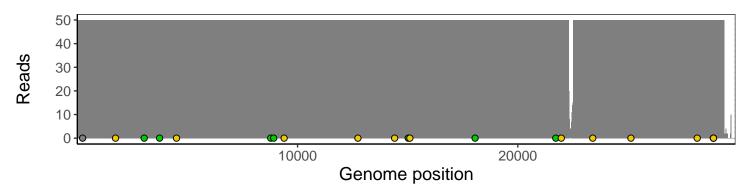
Analyses of individual experiments and composite results.

VSP0545-1 | 2020-12-11 | Saliva - Positive Control | 95-267 | genomes | single experiment

The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

