COVID-19 subject 2753

2021-01-19

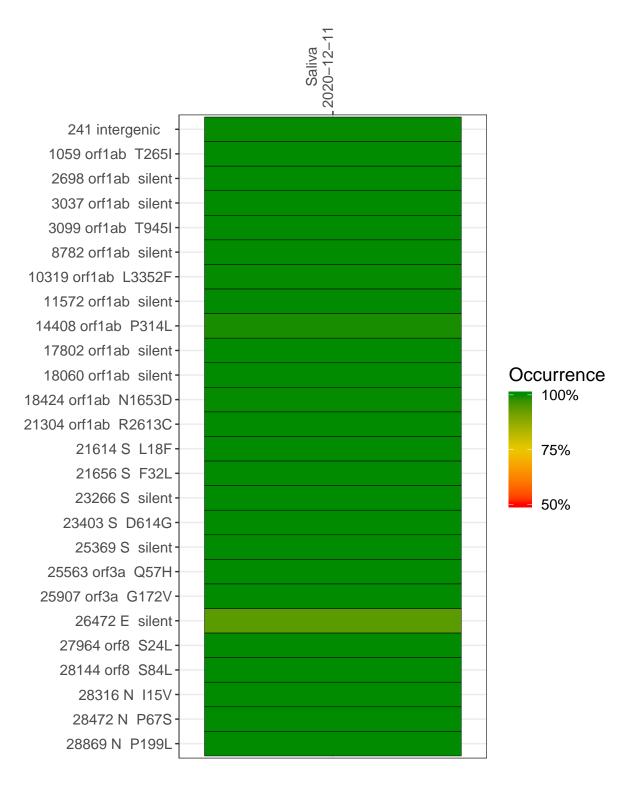
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

Experiment	Туре	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0534-1	single experiment	NA	Saliva	2020-12-11	18.02	99.8%	99.6%

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 2020–12–11

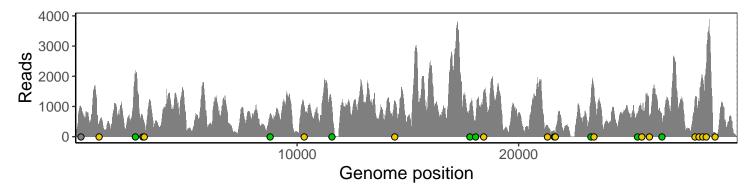
	2020-12-11
241 intergenic	972
1059 orf1ab T265I	422
2698 orf1ab silent	2050
3037 orf1ab silent	614
3099 orf1ab T945I	461
8782 orf1ab silent	752
10319 orf1ab L3352F	1034
11572 orf1ab silent	982
14408 orf1ab P314L	1146
17802 orf1ab silent	1045
18060 orf1ab silent	573
18424 orf1ab N1653D	1517
21304 orf1ab R2613C	376
21614 S L18F	173
21656 S F32L	120
23266 S silent	1273
23403 S D614G	1661
25369 S silent	601
25563 orf3a Q57H	893
25907 orf3a G172V	634
26472 E silent	675
27964 orf8 S24L	1216
28144 orf8 S84L	1137
28316 N I15V	1901
28472 N P67S	2239
28869 N P199L	131
	1 - 4



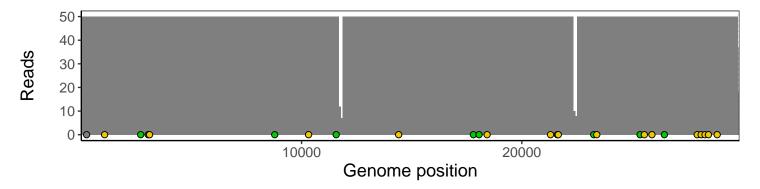
Analyses of individual experiments and composite results.

VSP0534-1 | 2020-12-11 | Saliva | 2753 | 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.

