COVID-19 subject 3157

2021-01-10

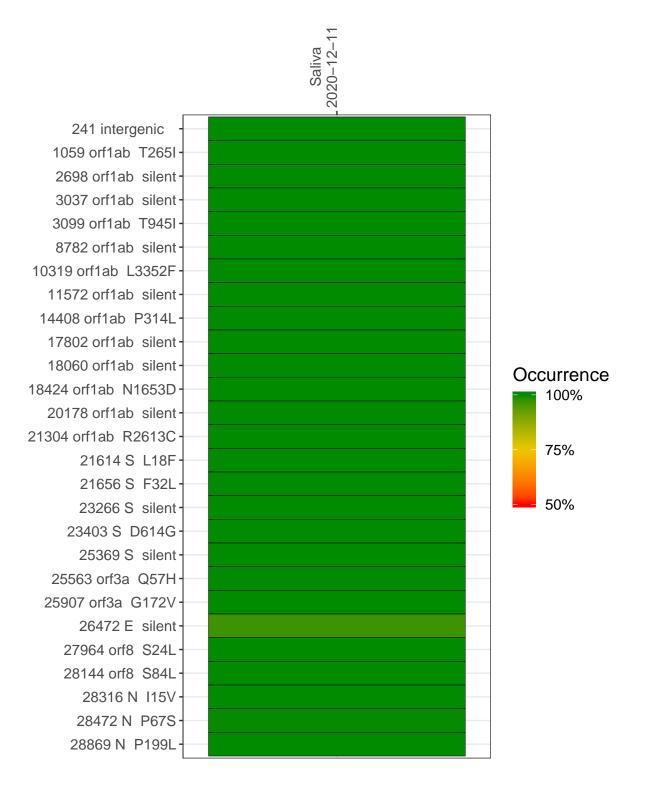
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

Experiment	Туре	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0540-1	single experiment	NA	Saliva	2020-12-11	17.99	99.7%	98.9%

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 for 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

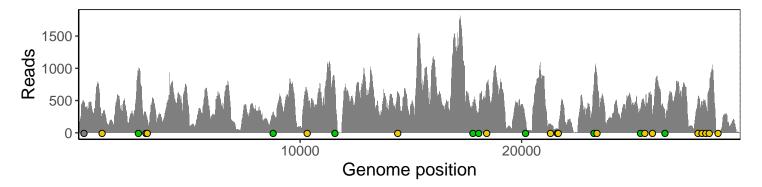
	2020-12-11
241 intergenic	471
1059 orf1ab T265I	262
2698 orf1ab silent	953
3037 orf1ab silent	287
3099 orf1ab T945I	228
8782 orf1ab silent	330
10319 orf1ab L3352F	647
11572 orf1ab silent	696
14408 orf1ab P314L	625
17802 orf1ab silent	464
18060 orf1ab silent	328
18424 orf1ab N1653D	769
20178 orf1ab silent	208
21304 orf1ab R2613C	124
21614 S L18F	58
21656 S F32L	40
23266 S silent	739
23403 S D614G	903
25369 S silent	308
25563 orf3a Q57H	442
25907 orf3a G172V	263
26472 E silent	331
27964 orf8 S24L	519
28144 orf8 S84L	423
28316 N I15V	444
28472 N P67S	651
28869 N P199L	23
	10-1



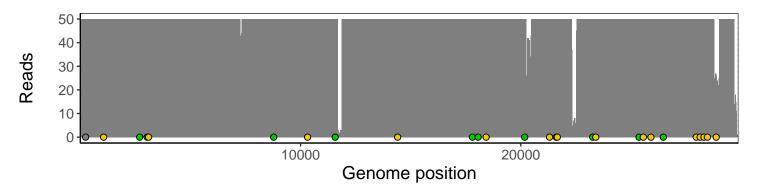
Analyses of individual experiments and composite results.

VSP0540-1 | 2020-12-11 | Saliva | 3157 | 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.

