COVID-19 subject 2748

2021-01-08

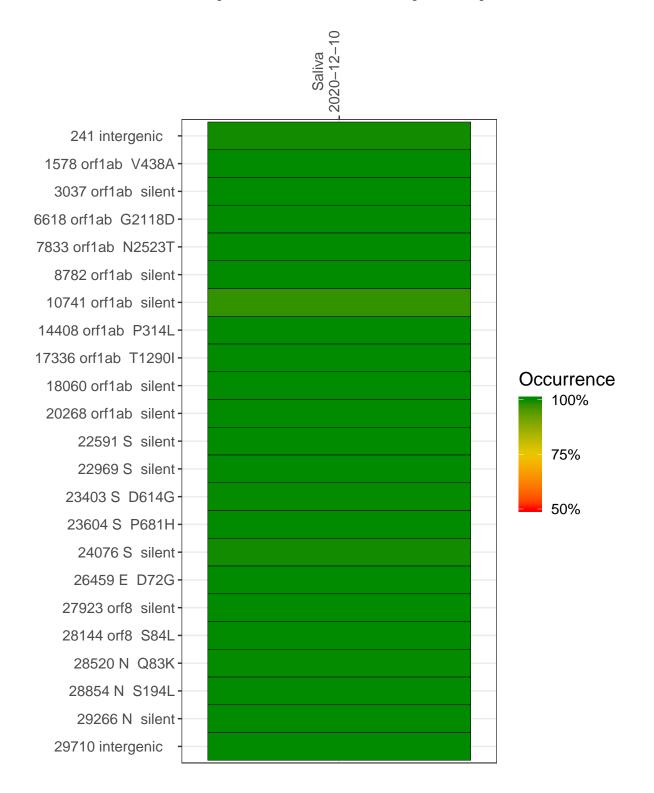
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

Experiment	Туре	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0529-1	single experiment	NA	Saliva	2020-12-10	29.82	100.0%	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 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 2020–12–10

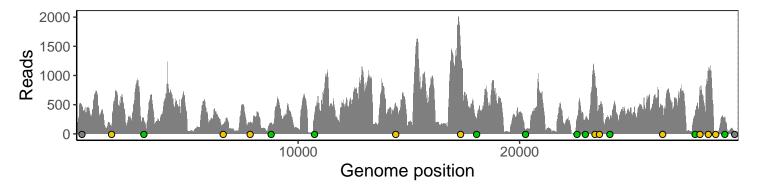
	2020-12-10
241 intergenic	453
1578 orf1ab V438A	82
3037 orf1ab silent	234
6618 orf1ab G2118D	151
7833 orf1ab N2523T	208
8782 orf1ab silent	196
10741 orf1ab silent	335
14408 orf1ab P314L	486
17336 orf1ab T1290I	1436
18060 orf1ab silent	255
20268 orf1ab silent	65
22591 S silent	256
22969 S silent	253
23403 S D614G	1002
23604 S P681H	471
24076 S silent	227
26459 E D72G	445
27923 orf8 silent	388
28144 orf8 S84L	738
28520 N Q83K	1024
28854 N S194L	46
29266 N silent	284
29710 intergenic	12
	VSP0529-1



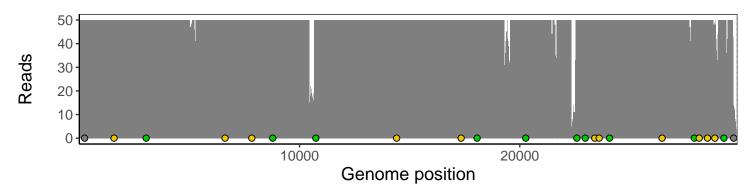
Analyses of individual experiments and composite results.

VSP0529-1 | 2020-12-10 | Saliva | 2748 | 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.

