

COVID-19 subject 428

2021-01-11

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	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0514-1	single experiment	NA	Saliva	2020-12-07	1.08	69.5%	64.3%

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
2020-12-07

241 intergenic
404 orf1ab K47stop
1578 orf1ab V438A
3176 orf1ab P971S
8782 orf1ab silent
10741 orf1ab silent
20268 orf1ab silent
23403 S D614G
24076 S silent
27739 orf7a L116F
28144 orf8 S84L
28854 N S194L

6207
6
1374
1367
6085
223
1494
9803
2255
2287
10520
1377

Base change

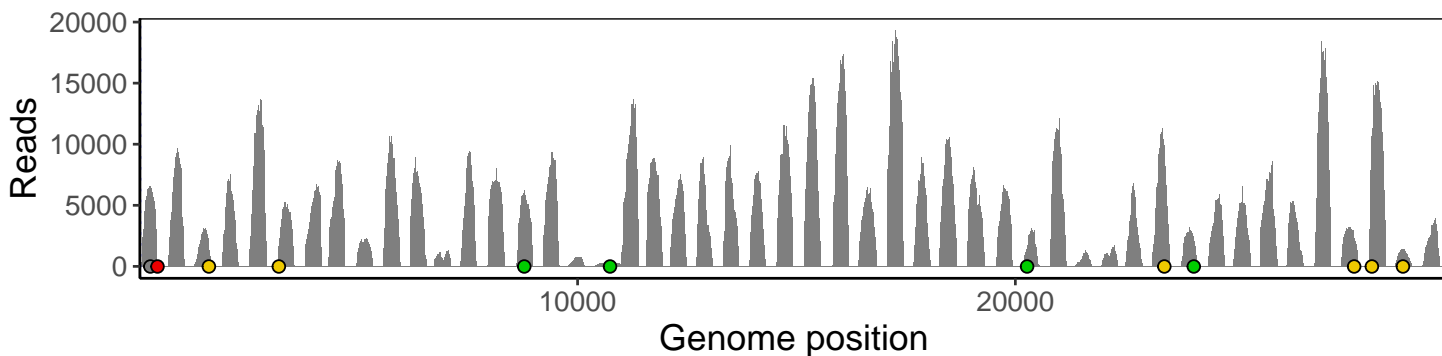
- Expected
- A
- T
- C
- G
- N
- Ins/Del
- No data

VSP0514-1

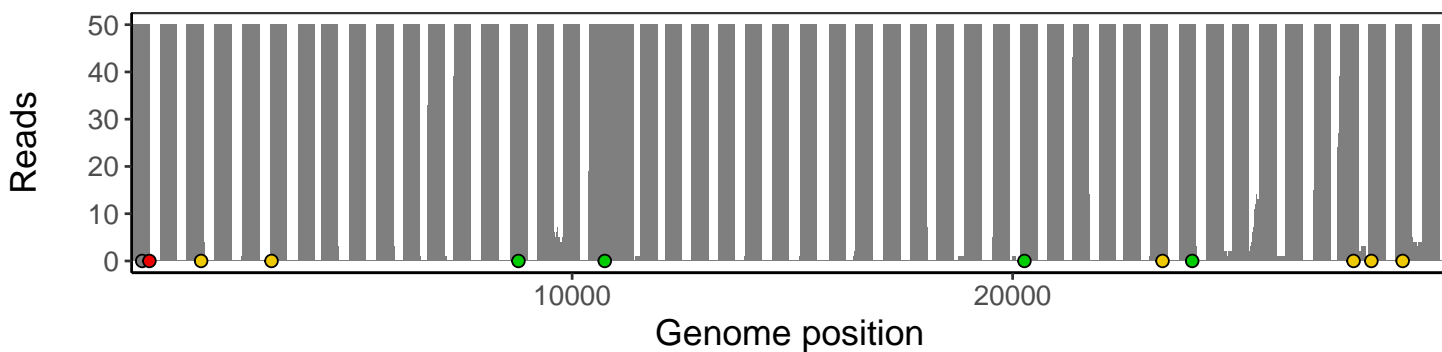
Analyses of individual experiments and composite results.

VSP0514-1 | 2020-12-07 | Saliva | 428s | 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.

