

COVID-19 subject 227

2020-09-14

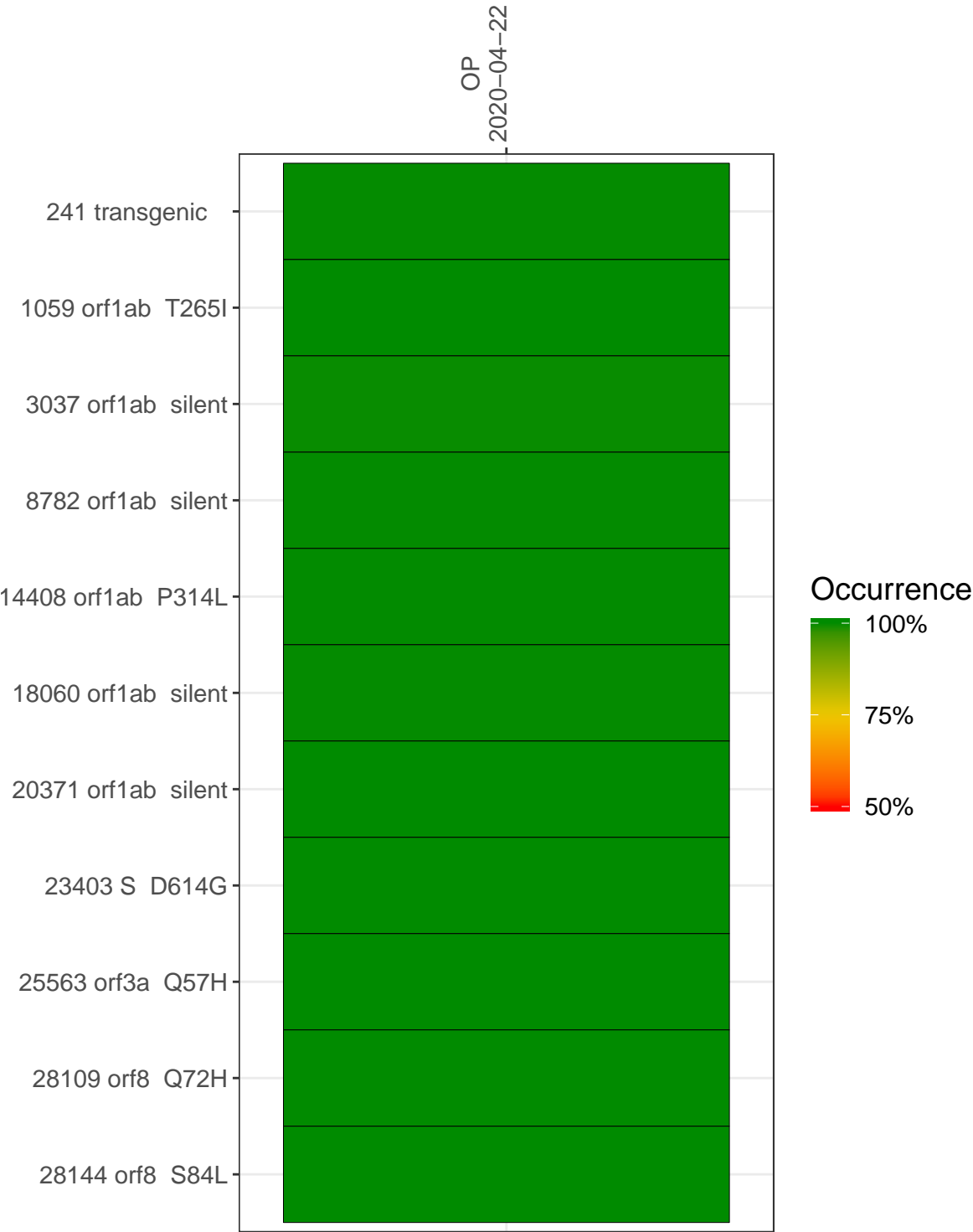
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0020	composite	NA	OP	2020-04-22	29.91	99.8%	99.8%
VSP0020-1m	single experiment	NA	OP	2020-04-22	20.45	97.4%	96.7%
VSP0020-2	single experiment	NA	OP	2020-04-22	29.91	99.8%	99.8%

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.



OP
2020-04-22

241 transgenic	202	1908
1059 orf1ab T265I	330	869
3037 orf1ab silent	215	857
8782 orf1ab silent	625	747
14408 orf1ab P314L	484	1586
18060 orf1ab silent	229	834
20371 orf1ab silent	257	313
23403 S D614G	1103	2305
25563 orf3a Q57H	626	2637
28109 orf8 Q72H	1026	2793
28144 orf8 S84L	904	2464

Base change

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

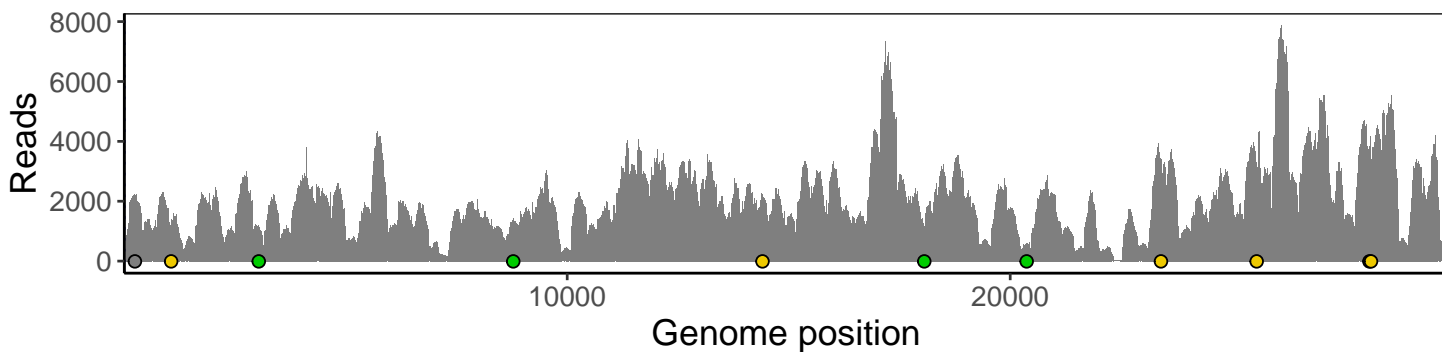
VSP0020-1m

VSP0020-2

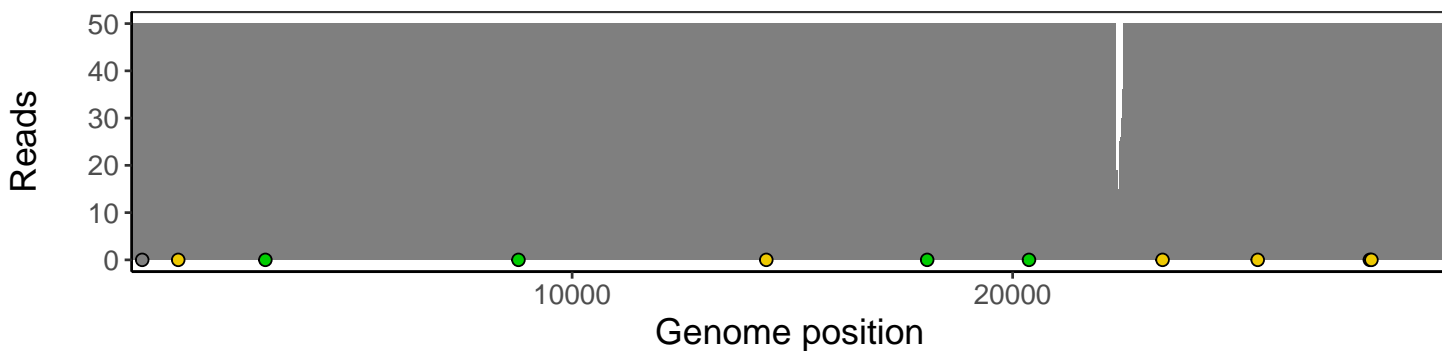
Analyses of individual experiments and composite results.

VSP0020 | 2020-04-22 | OP | 227o | composite result

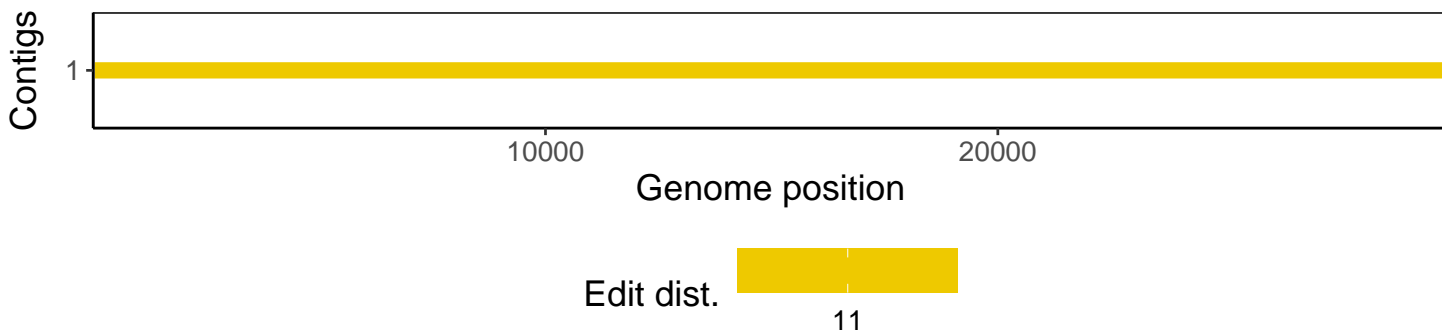
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



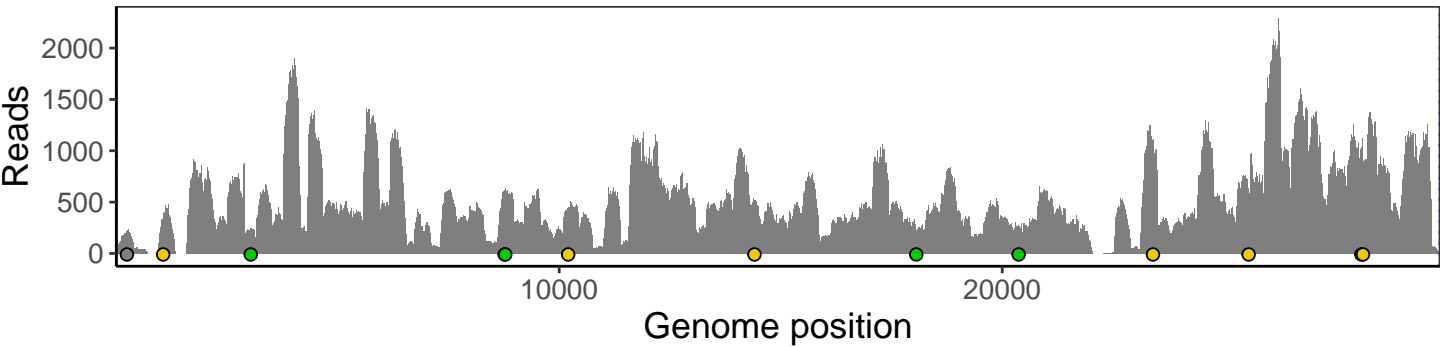
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



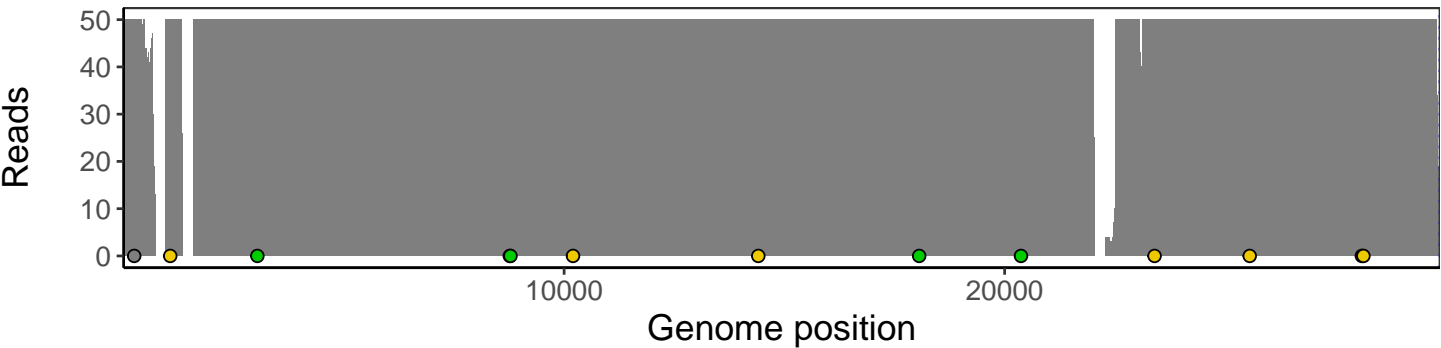
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



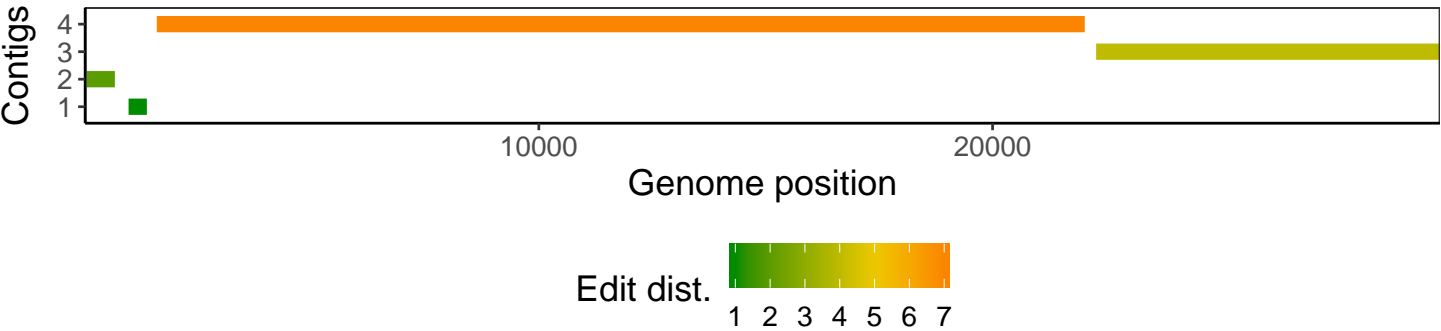
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



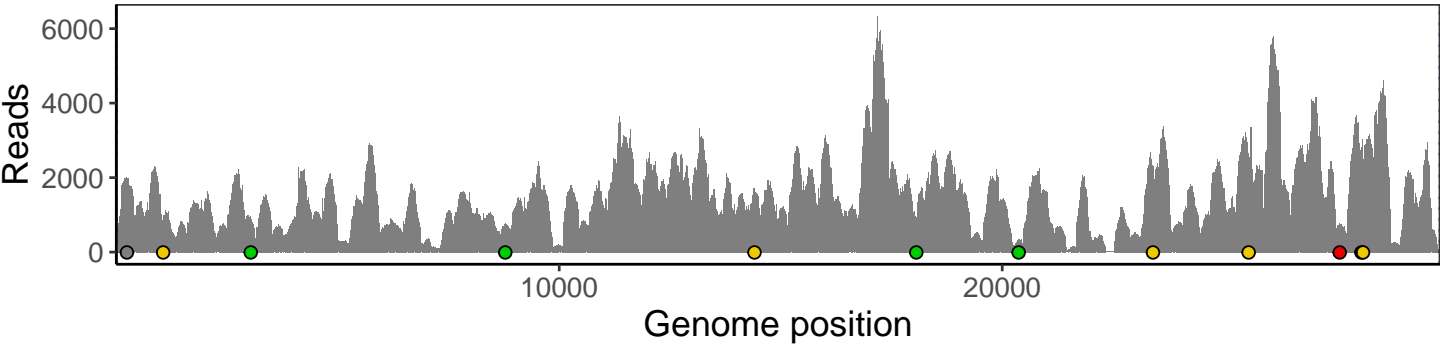
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



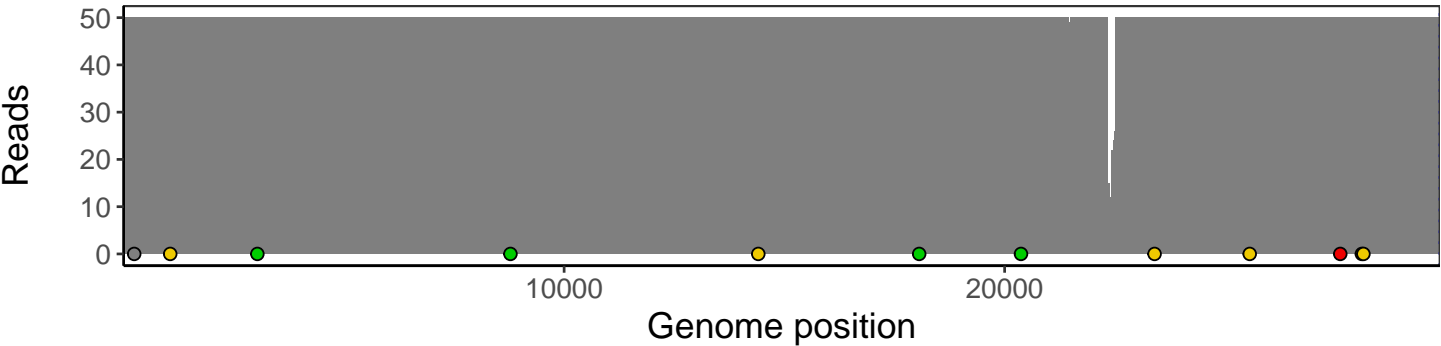
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

