

COVID-19 subject 238

2020-11-30

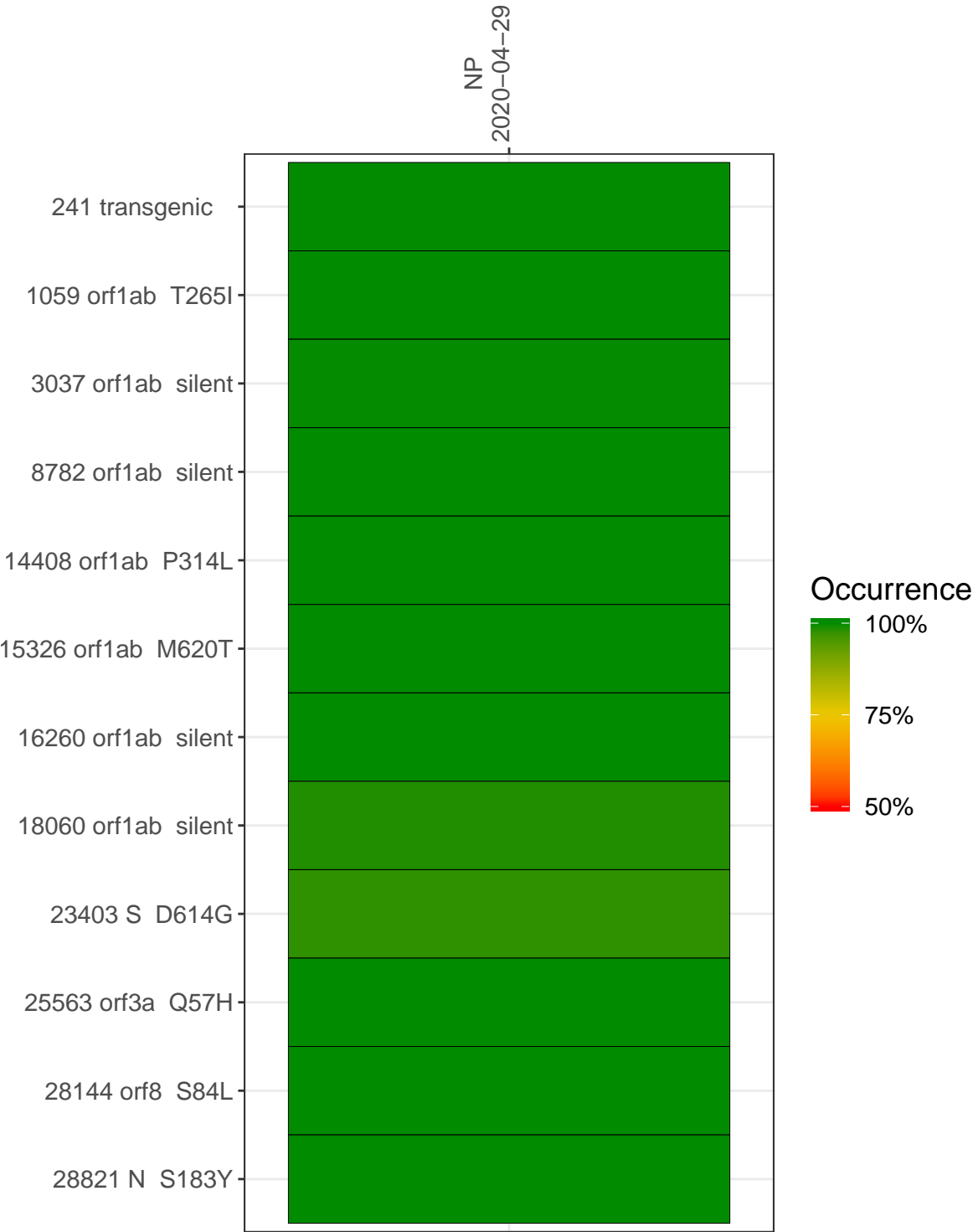
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0039	composite	NA	NP	2020-04-29	29.89	99.8%	99.8%
VSP0039-1m	single experiment	NA	NP	2020-04-29	29.67	99.2%	99.2%
VSP0039-2	single experiment	2465	NP	2020-04-29	29.89	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.



NP
2020-04-29

	VSP0039-1m	VSP0039-2
241 transgenic	1044	1493
1059 orf1ab T265I	1092	1261
3037 orf1ab silent	411	1110
8782 orf1ab silent	684	1787
14408 orf1ab P314L	633	2758
15326 orf1ab M620T	1118	5090
16260 orf1ab silent	212	1134
18060 orf1ab silent	724	1454
23403 S D614G	1286	4735
25563 orf3a Q57H	1110	2126
28144 orf8 S84L	970	5533
28821 N S183Y	372	1654

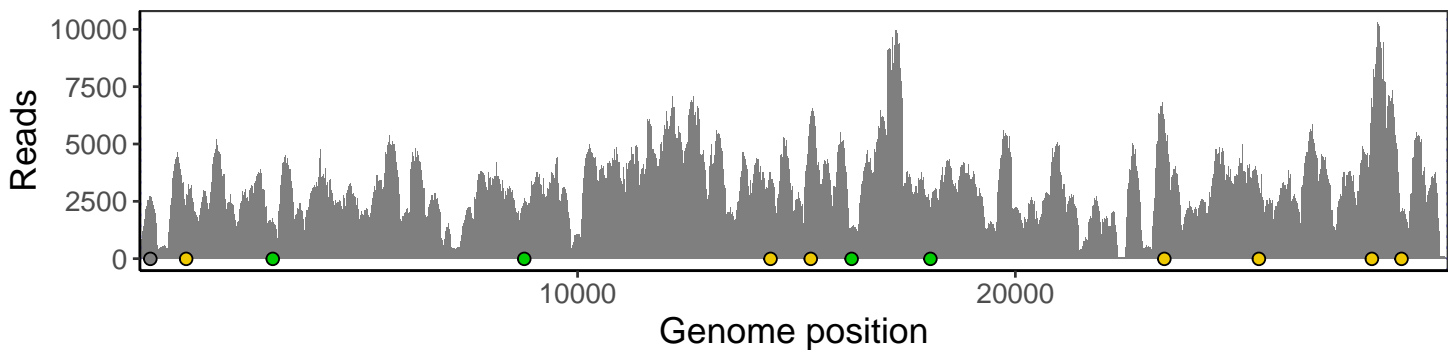
Base change



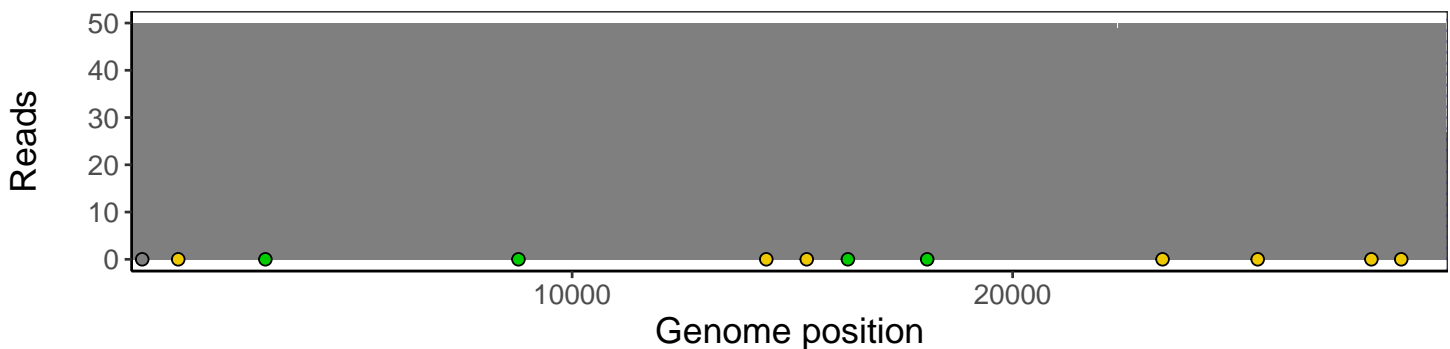
Analyses of individual experiments and composite results.

VSP0039 | 2020-04-29 | NP | 238n-tri | composite result

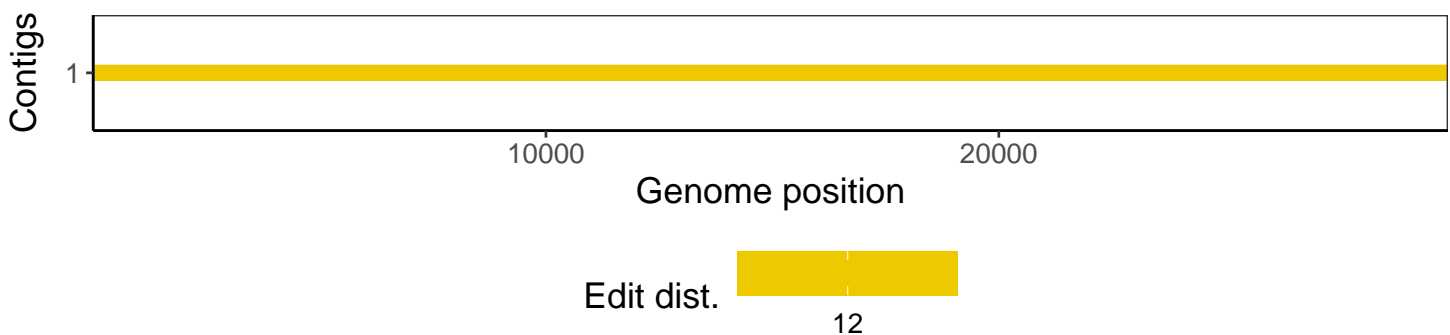
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



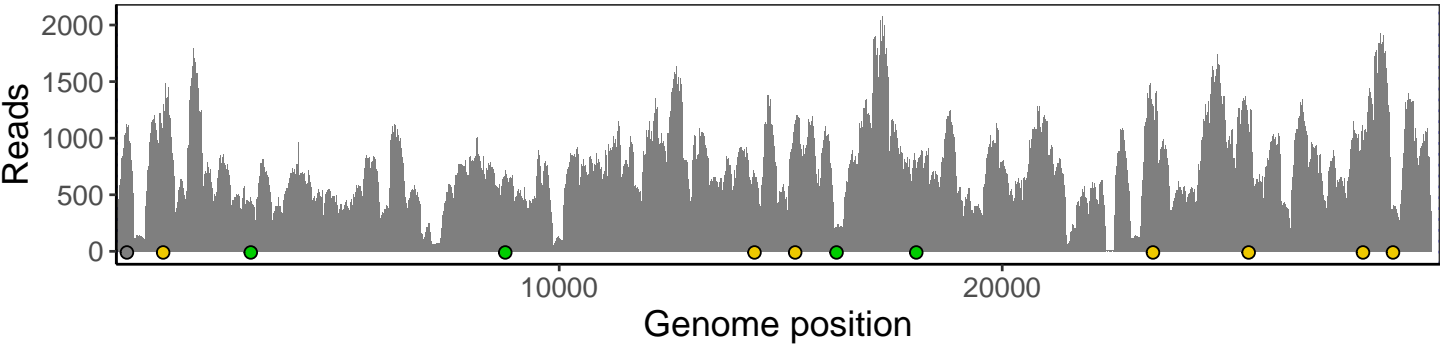
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



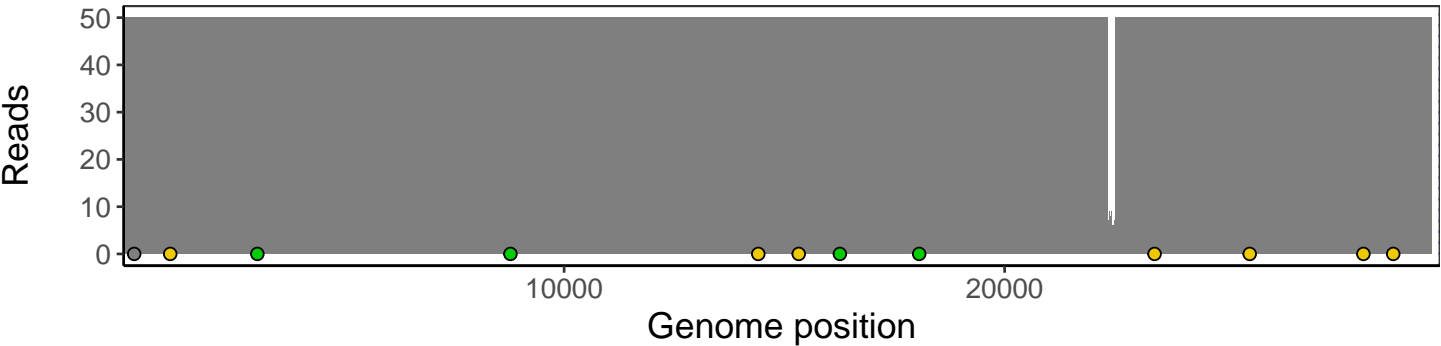
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



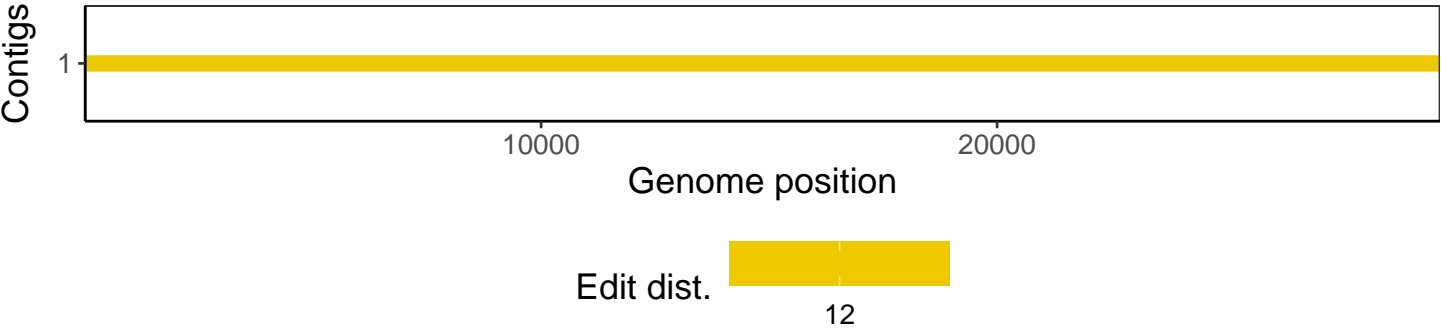
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



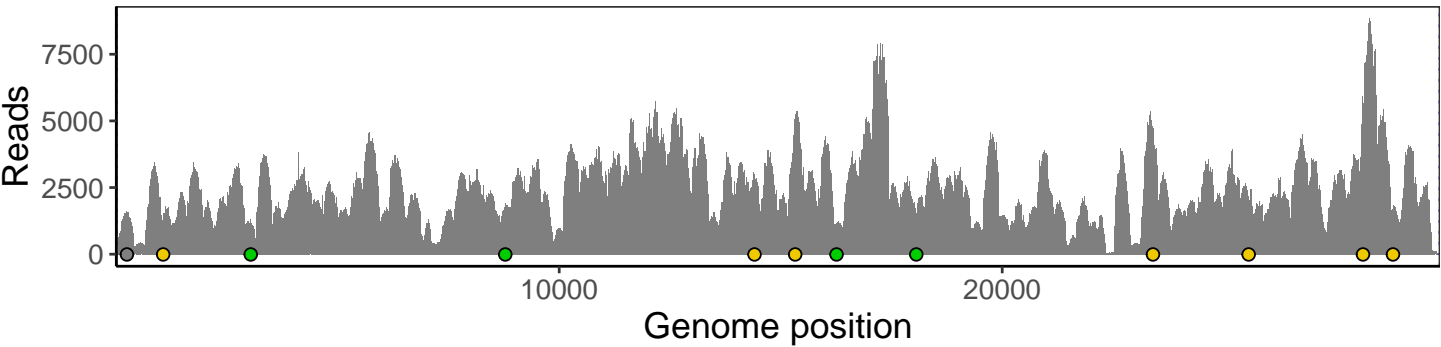
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



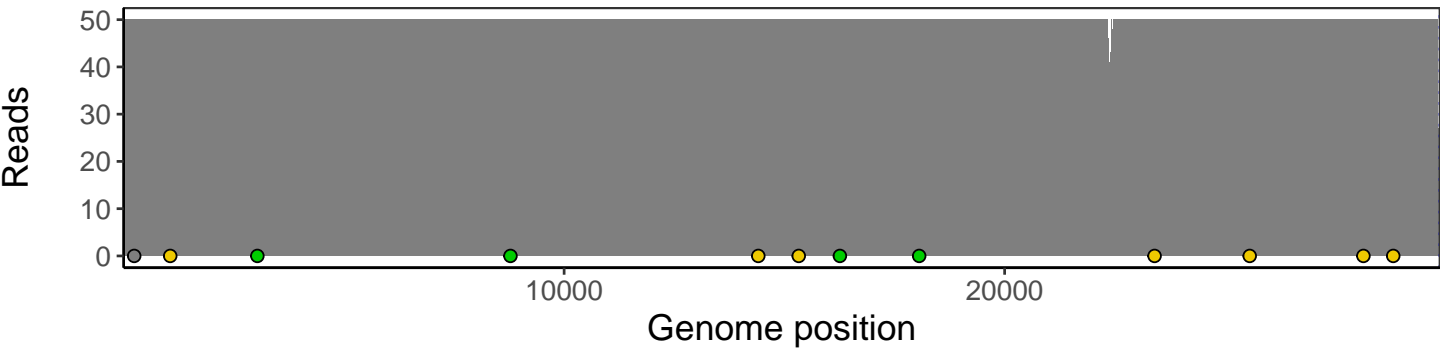
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

