

# COVID-19 subject 211-TCE

*2020-10-23*

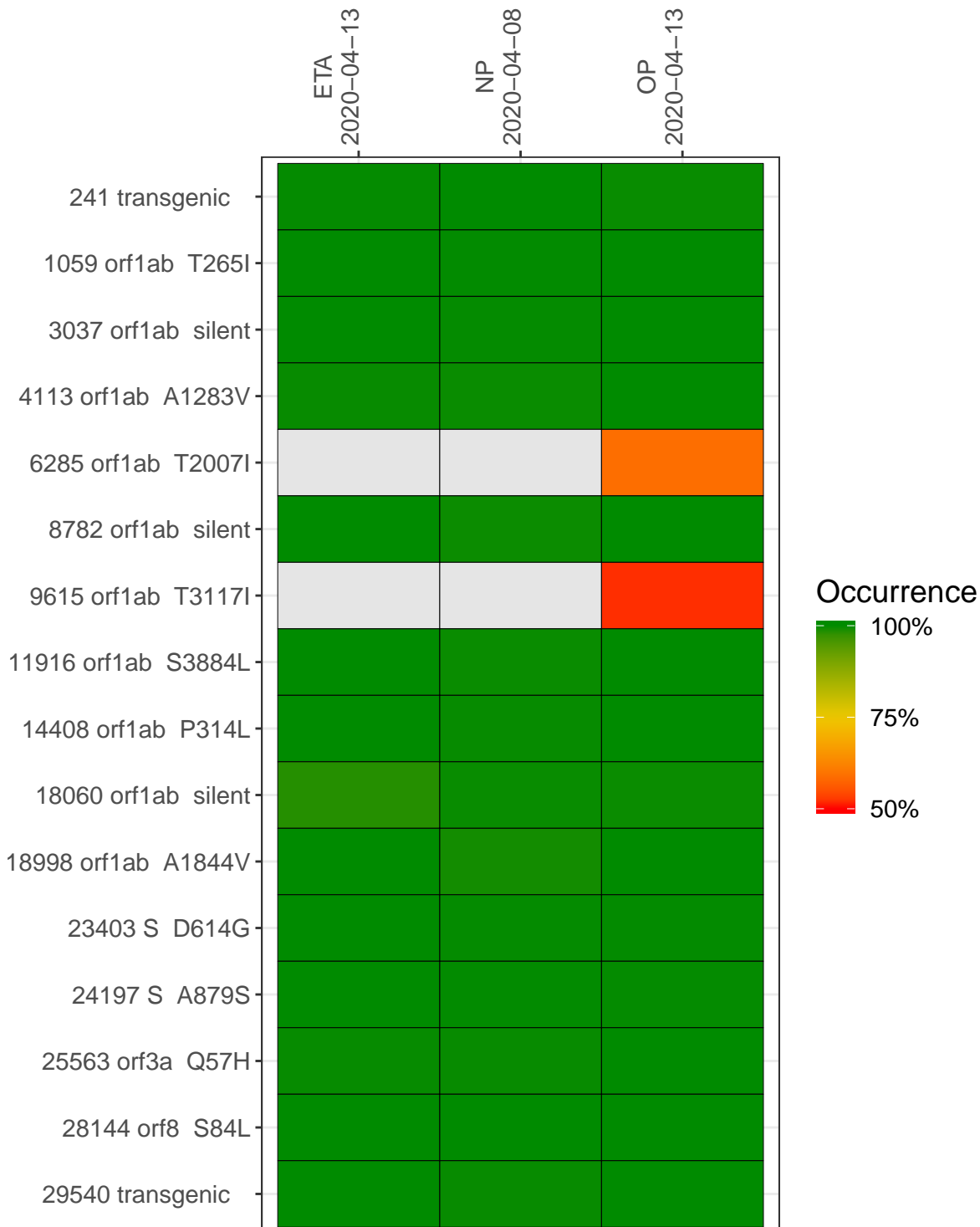
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage ( $\geq 5$ reads)
VSP0184-1m	single experiment	NA	NP	2020-04-08	29.86	99.9%	99.8%
VSP0185-1m	single experiment	NA	OP	2020-04-13	29.88	99.9%	99.8%
VSP0186-1m	single experiment	NA	ETA	2020-04-13	29.80	99.9%	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.



	ETA 2020-04-13	NP 2020-04-08	OP 2020-04-13	
241 transgenic	944	1074	990	
1059 orf1ab T265I	640	2450	2642	
3037 orf1ab silent	591	1039	1690	
4113 orf1ab A1283V	1237	1826	1753	
6285 orf1ab T2007I	920	1752	129	
8782 orf1ab silent	435	730	120	
9615 orf1ab T3117I	321	61	660	
11916 orf1ab S3884L	568	483	1283	
14408 orf1ab P314L	517	704	1631	
18060 orf1ab silent	567	538	1633	
18998 orf1ab A1844V	665	1419	106	
23403 S D614G	2056	1946	1532	
24197 S A879S	538	2260	2322	
25563 orf3a Q57H	645	2387	2480	
28144 orf8 S84L	558	1249	386	
29540 transgenic	351	586	419	
	VSP0186-1m	VSP0184-1m	VSP0185-1m	

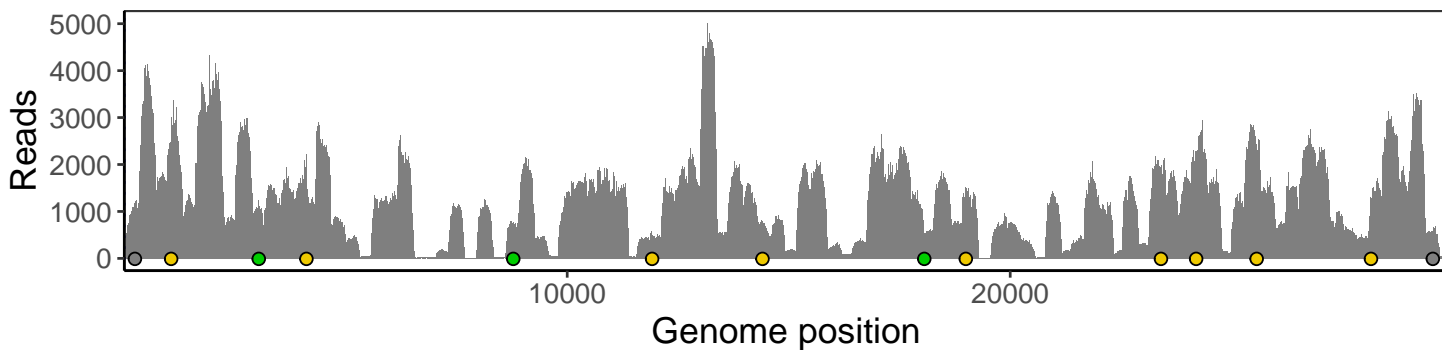
Base change

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

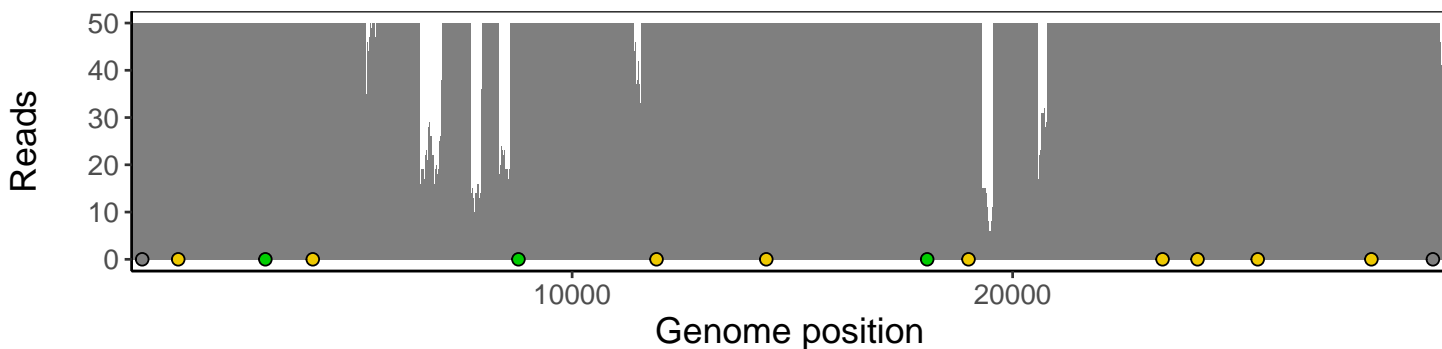
## Analyses of individual experiments and composite results.

VSP0184-1m | 2020-04-08 | NP | 5 | genomes | single experiment

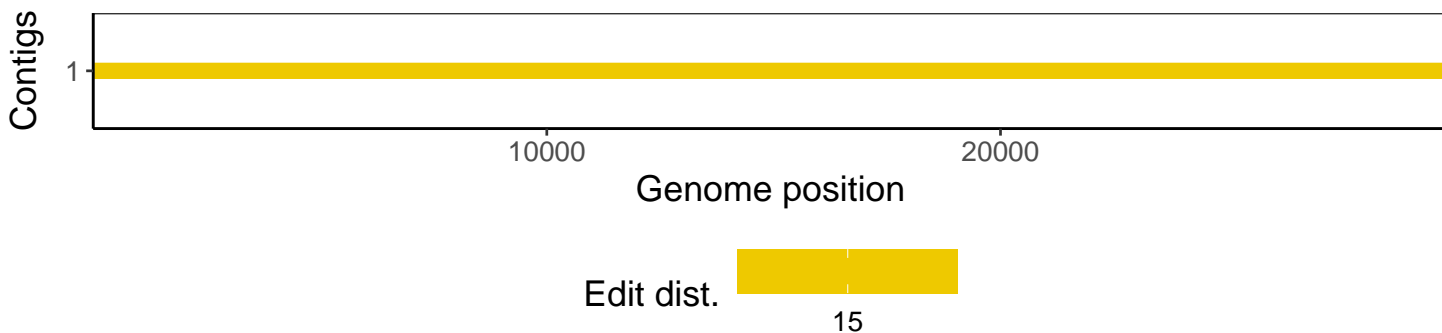
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



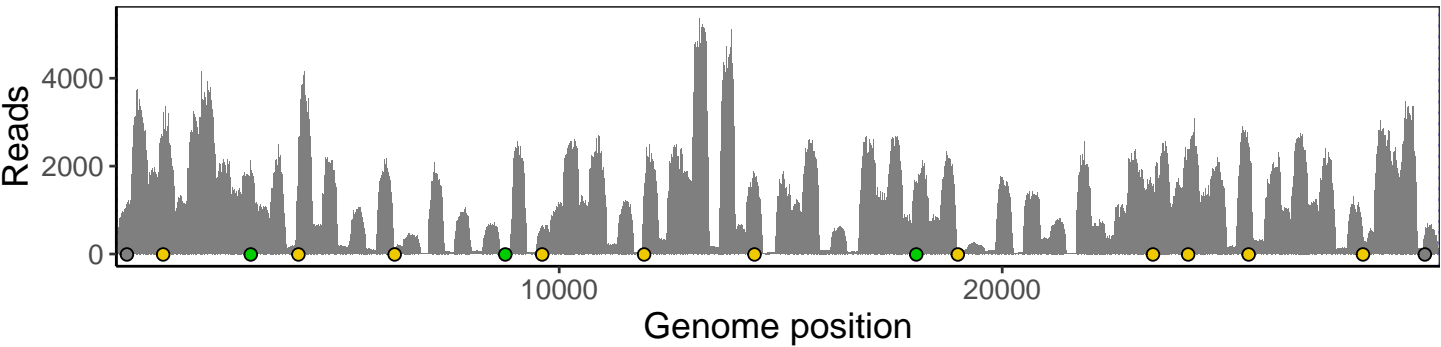
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



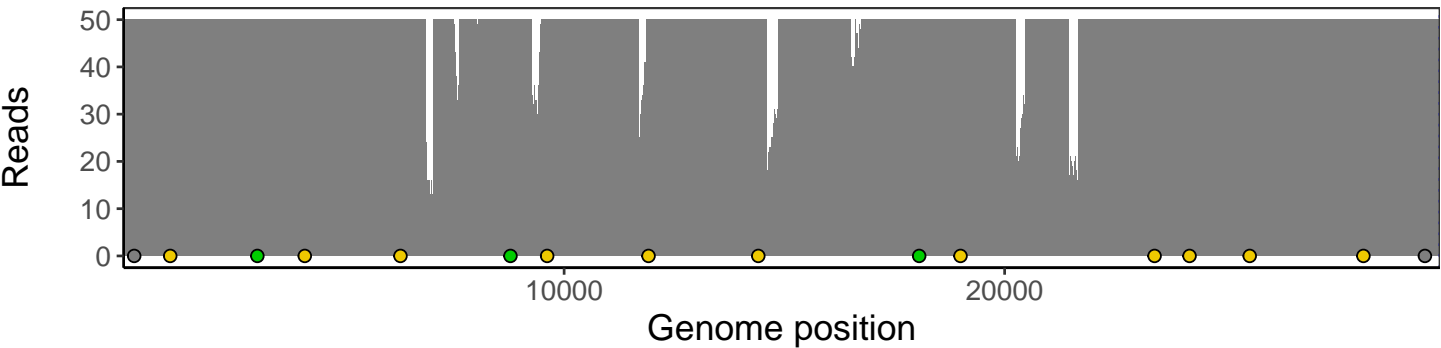
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



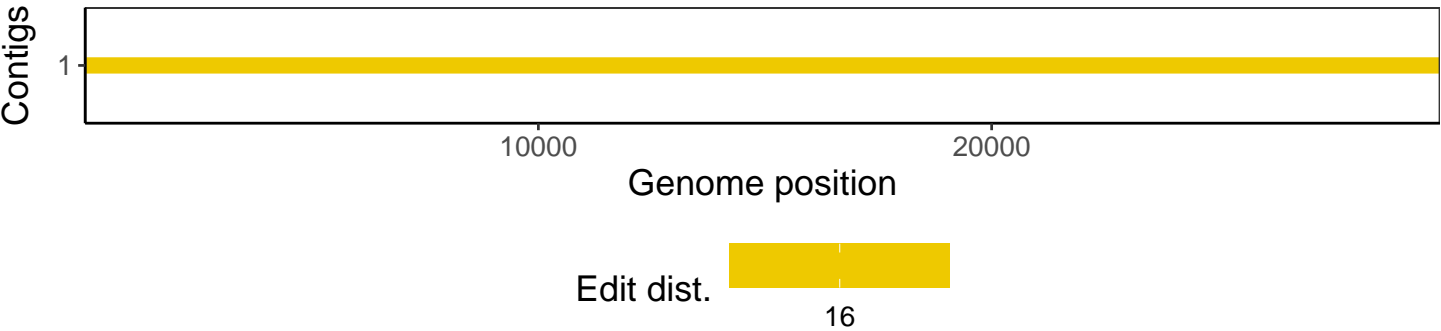
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



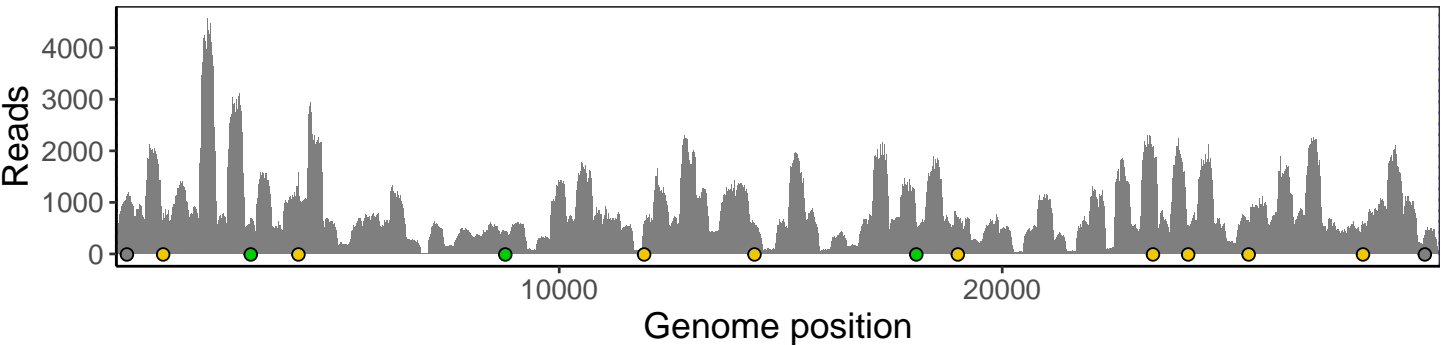
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



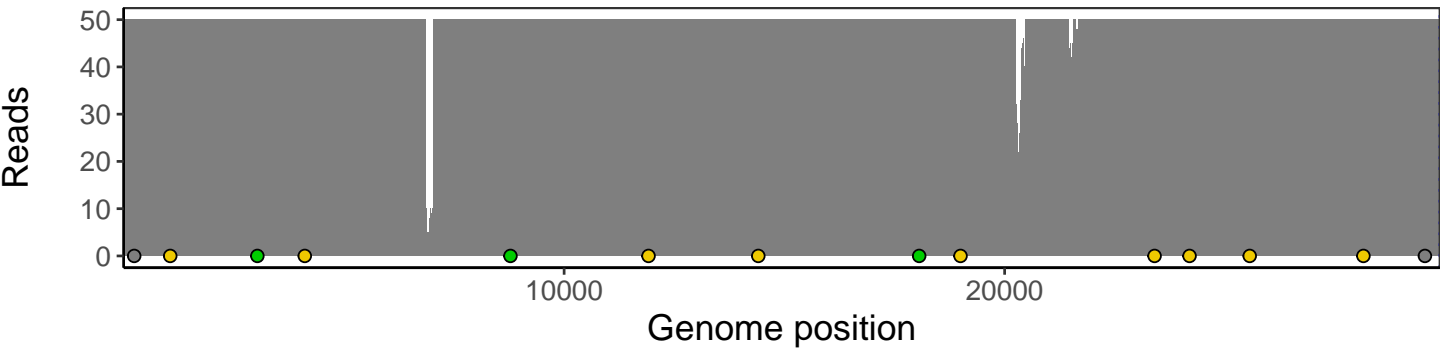
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

