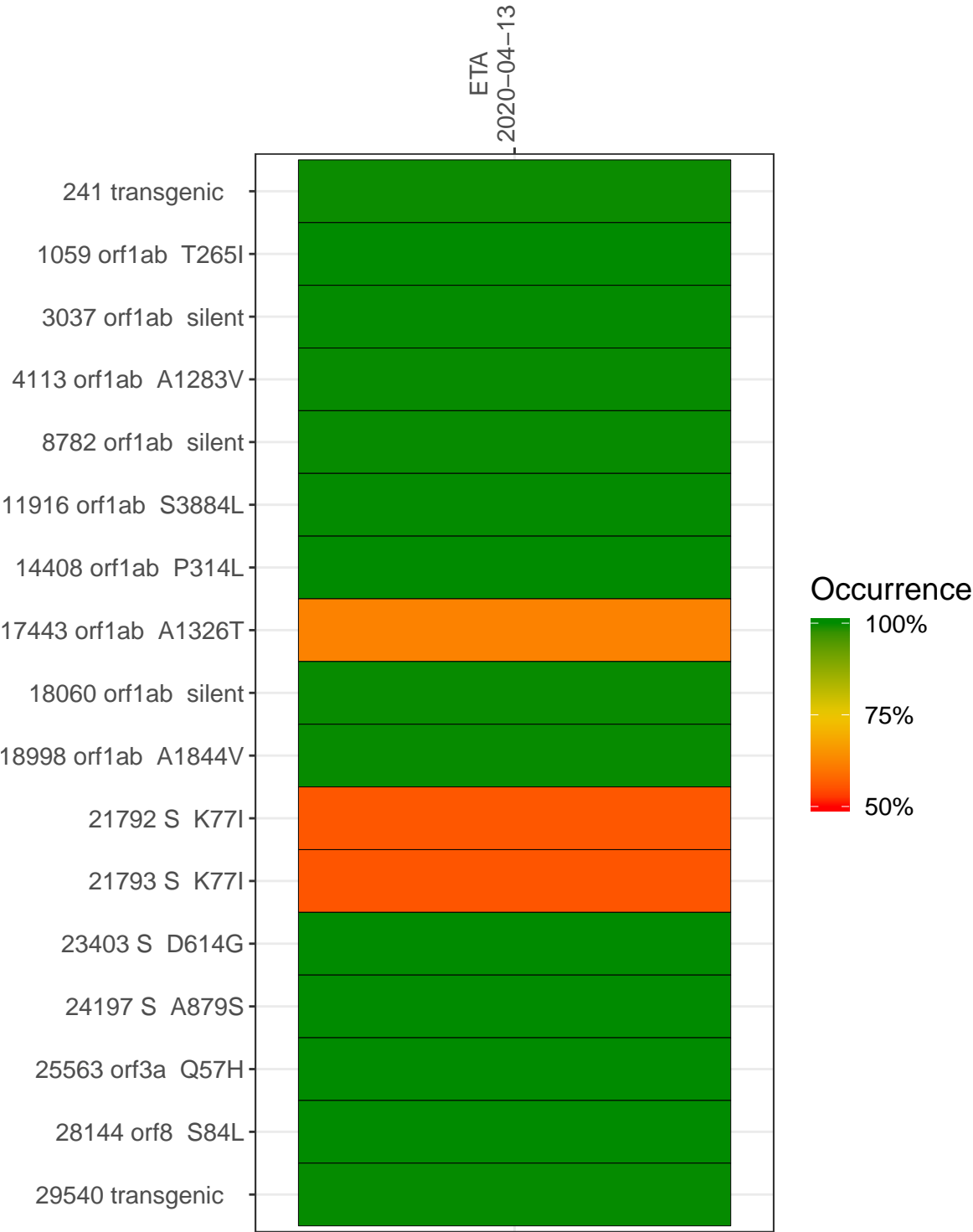# COVID-19 subject 211

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.
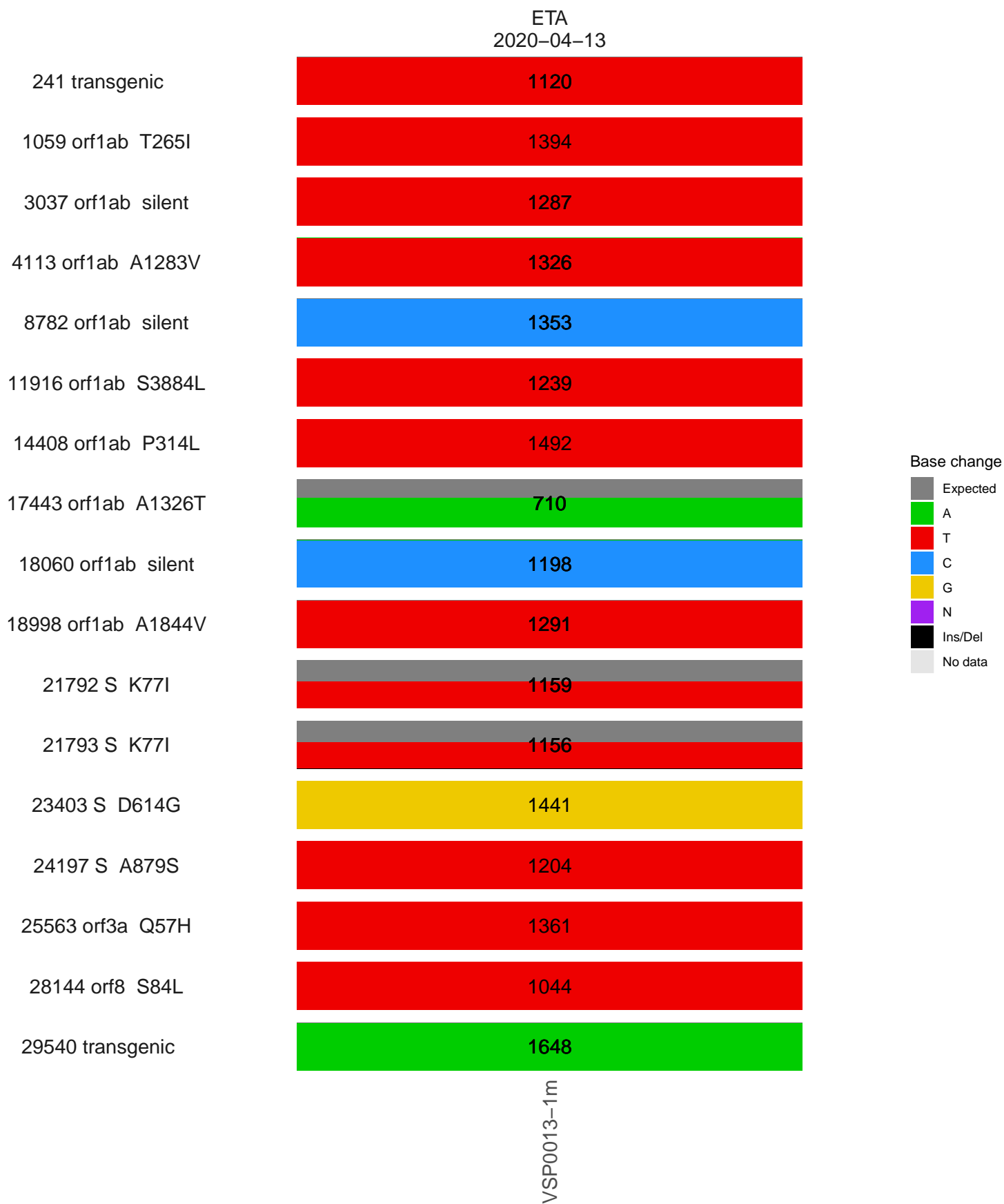Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0013-1m	single experiment	NA	ETA	2020-04-13	29.90	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.

ETA
2020−04−13

Position	Value	Base
241 transgenic	1120	T
1059 orf1ab T265I	1394	T
3037 orf1ab silent	1287	T
4113 orf1ab A1283V	1326	T
8782 orf1ab silent	1353	C
11916 orf1ab S3884L	1239	T
14408 orf1ab P314L	1492	T
17443 orf1ab A1326T	710	A (Expected)
18060 orf1ab silent	1198	C
18998 orf1ab A1844V	1291	T
21792 S K77I	1159	T (Expected)
21793 S K77I	1156	T (Expected)
23403 S D614G	1441	G
24197 S A879S	1204	T
25563 orf3a Q57H	1361	T
28144 orf8 S84L	1044	T
29540 transgenic	1648	A

VSP0013−1m

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

Analyses of individual experiments and composite results.

VSP0013-1m | 2020-04-13 | ETA | 211-tri | 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.
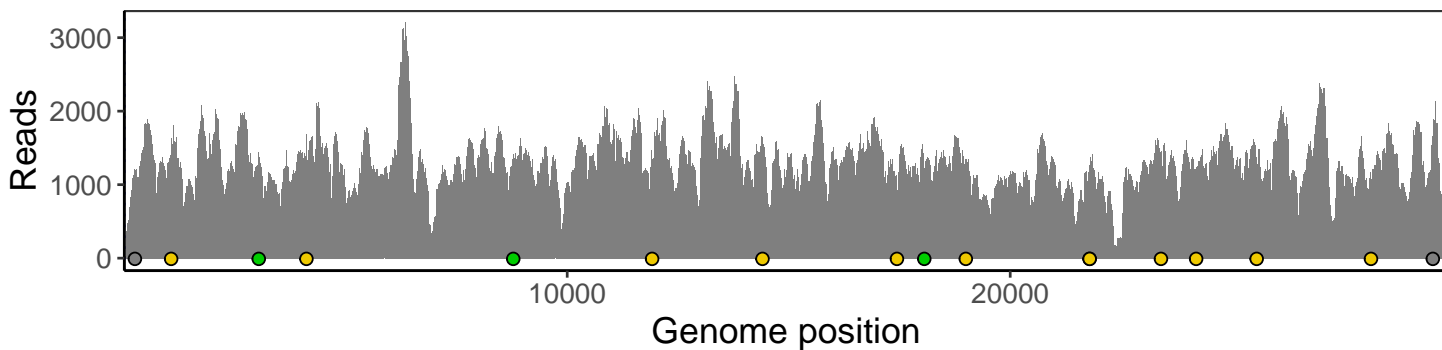
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.
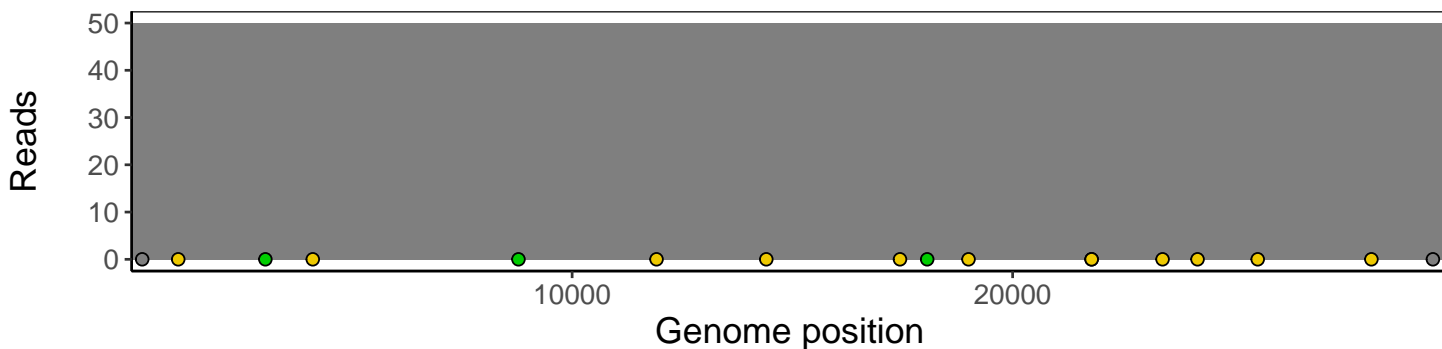
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

4