COVID-19 subject SRR11783612

2020-09-29

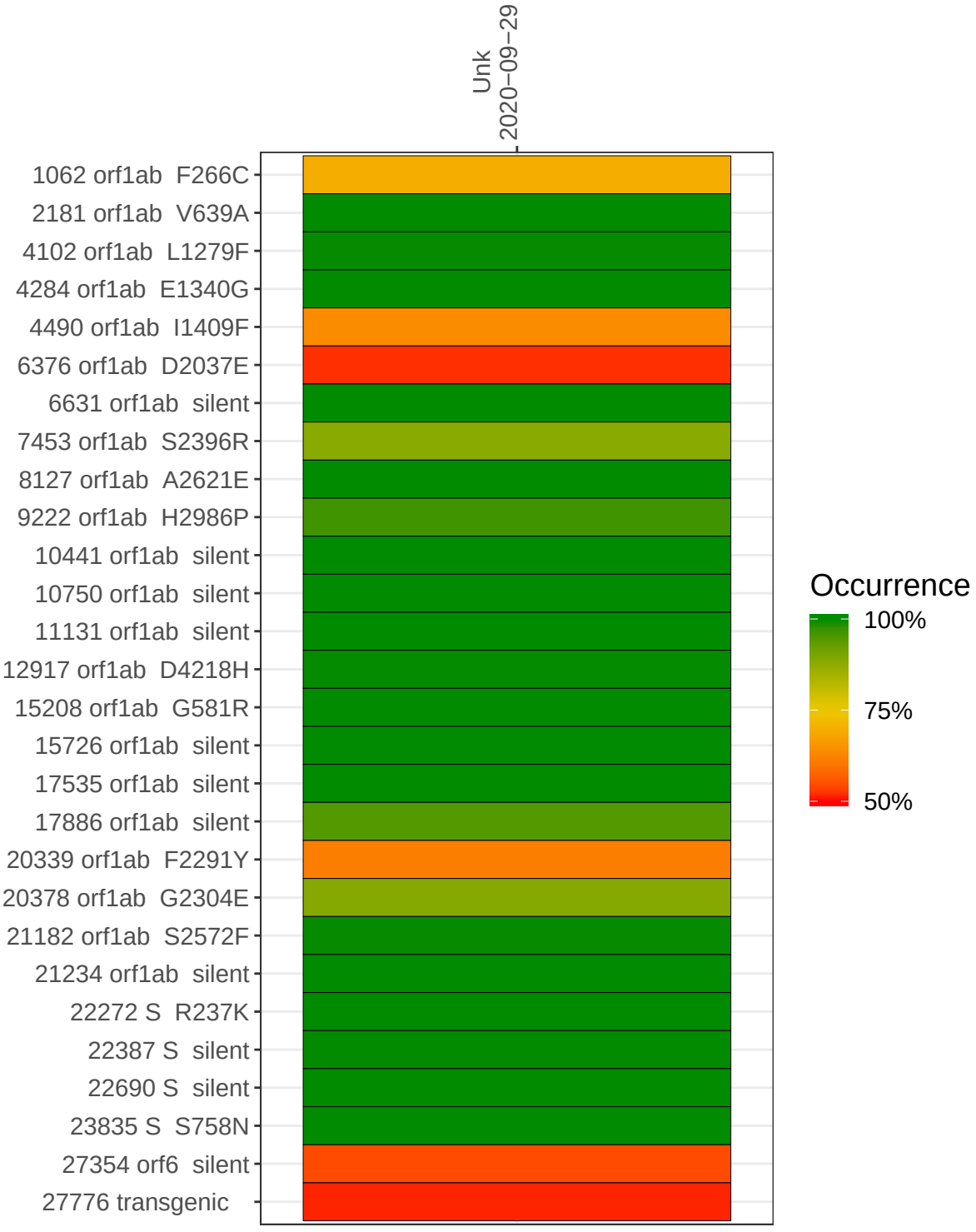
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

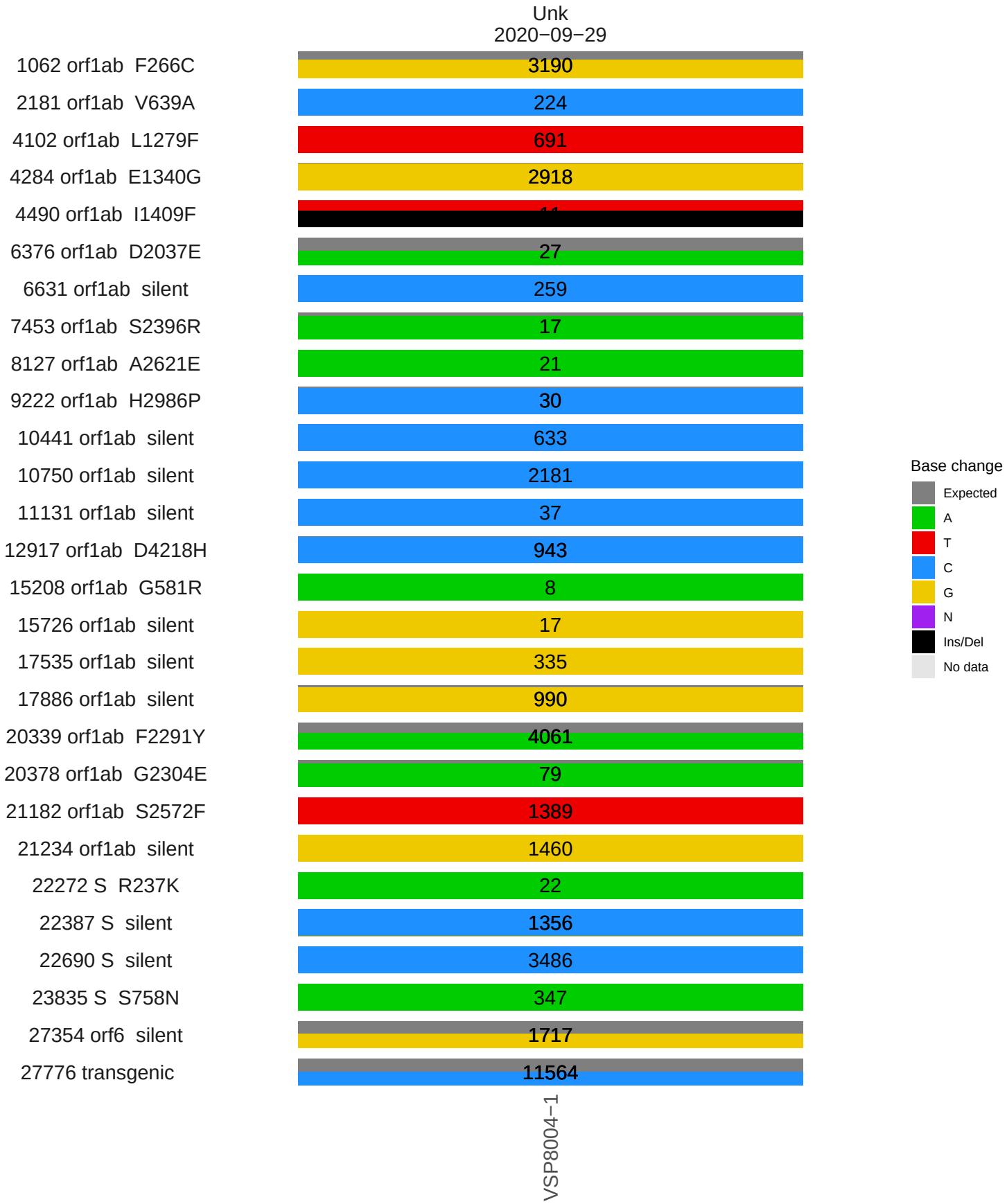
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP8004-1	single experiment	NA	Unk	2020-09-29	0.55	40.3%	37.6%

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.

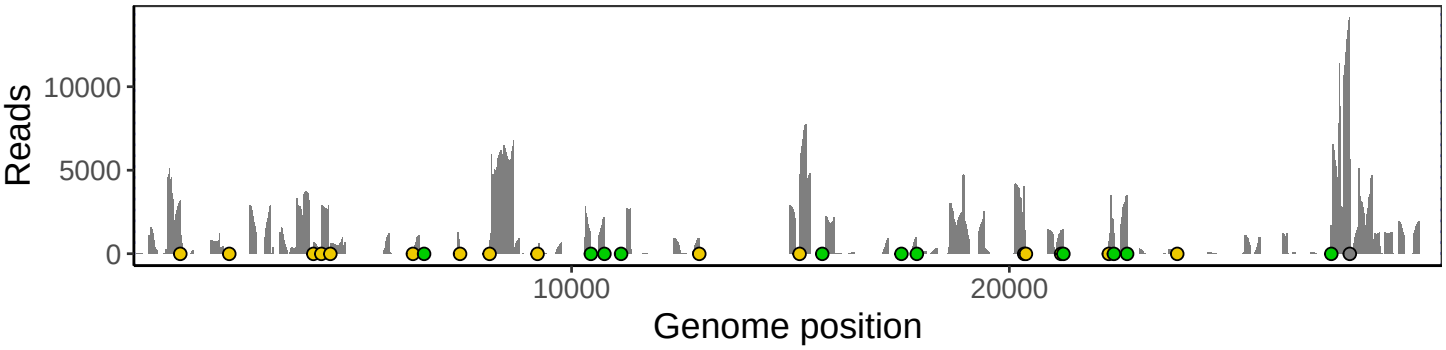



Unk
2020−09−29
1062 orf1ab F266C
2181 orf1ab V639A
4102 orf1ab L1279F
4284 orf1ab E1340G
4490 orf1ab I1409F
6376 orf1ab D2037E
6631 orf1ab silent
7453 orf1ab S2396R
8127 orf1ab A2621E
9222 orf1ab H2986P
10441 orf1ab silent
10750 orf1ab silent
11131 orf1ab silent
12917 orf1ab D4218H
15208 orf1ab G581R
15726 orf1ab silent
17535 orf1ab silent
17886 orf1ab silent
20339 orf1ab F2291Y
20378 orf1ab G2304E
21182 orf1ab S2572F
21234 orf1ab silent
22272 S R237K
22387 S silent
22690 S silent
23835 S S758N
27354 orf6 silent
27776 transgenic
3190
224
691
2918
11
27
259
17
21
30
633
2181
37
943
8
17
335
990
4061
79
1389
1460
22
1356
3486
347
1717
11564
VSP8004−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

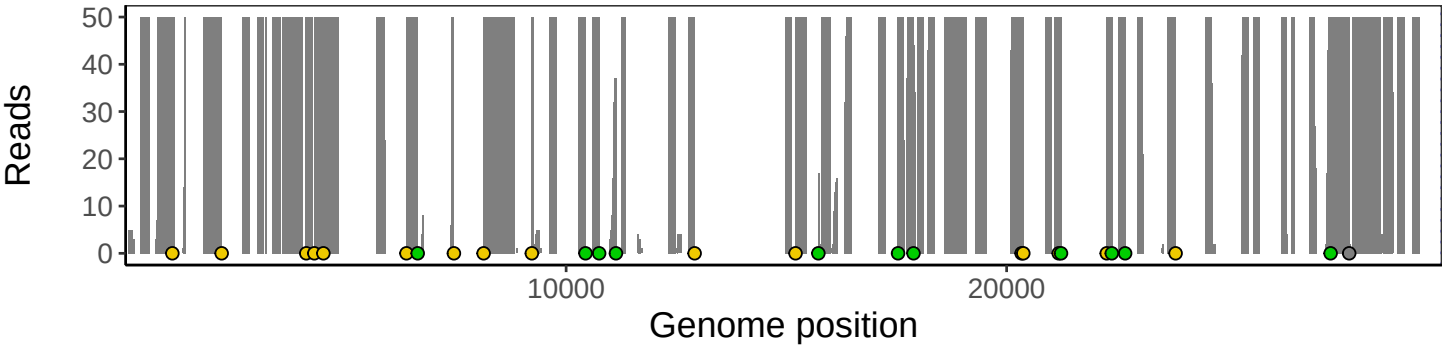
Analyses of individual experiments and composite results.

VSP8004-1 | 2020-09-29 | Unk | SRR11783612 | 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.