COVID-19 subject SRR11783625

2020-09-29

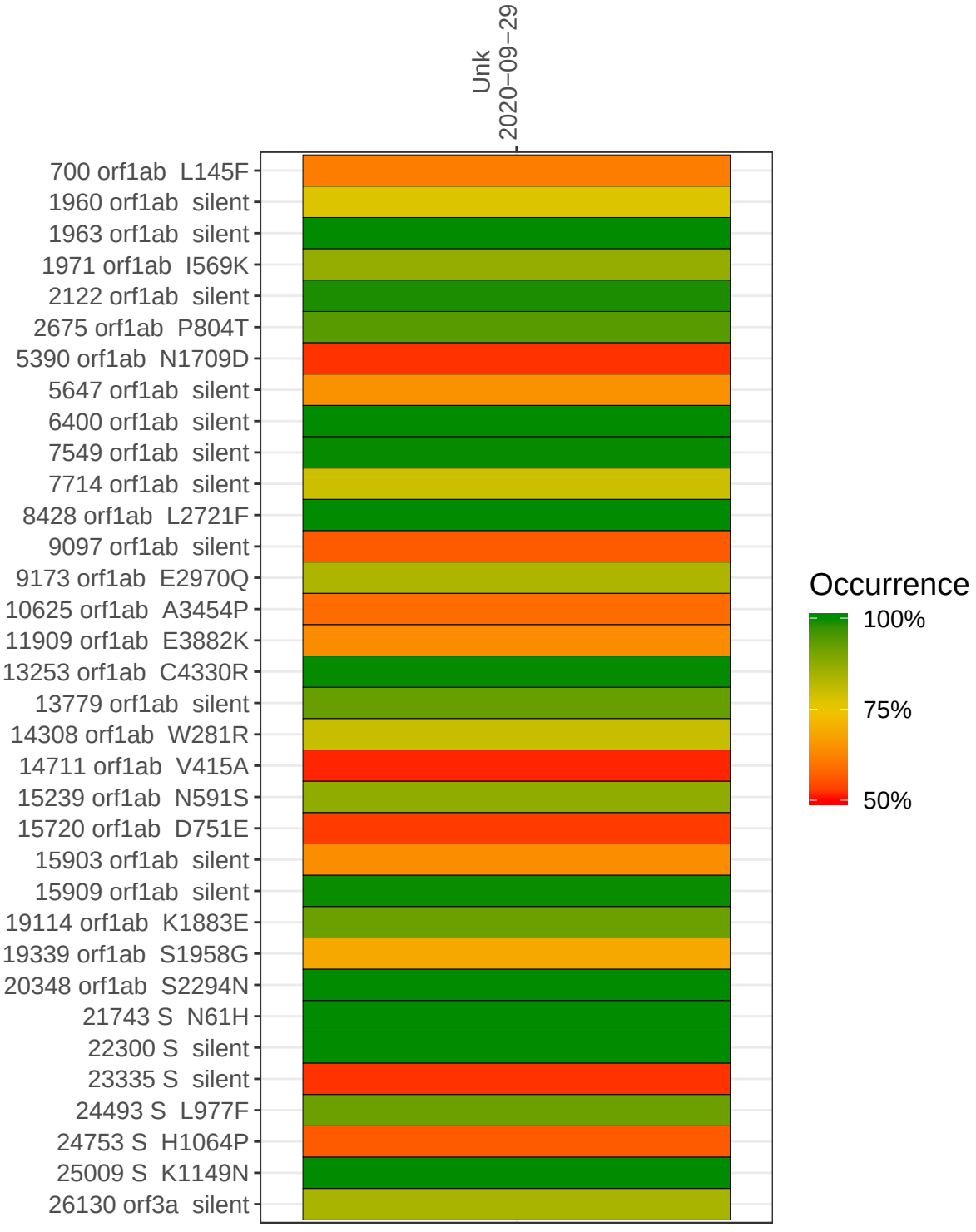
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

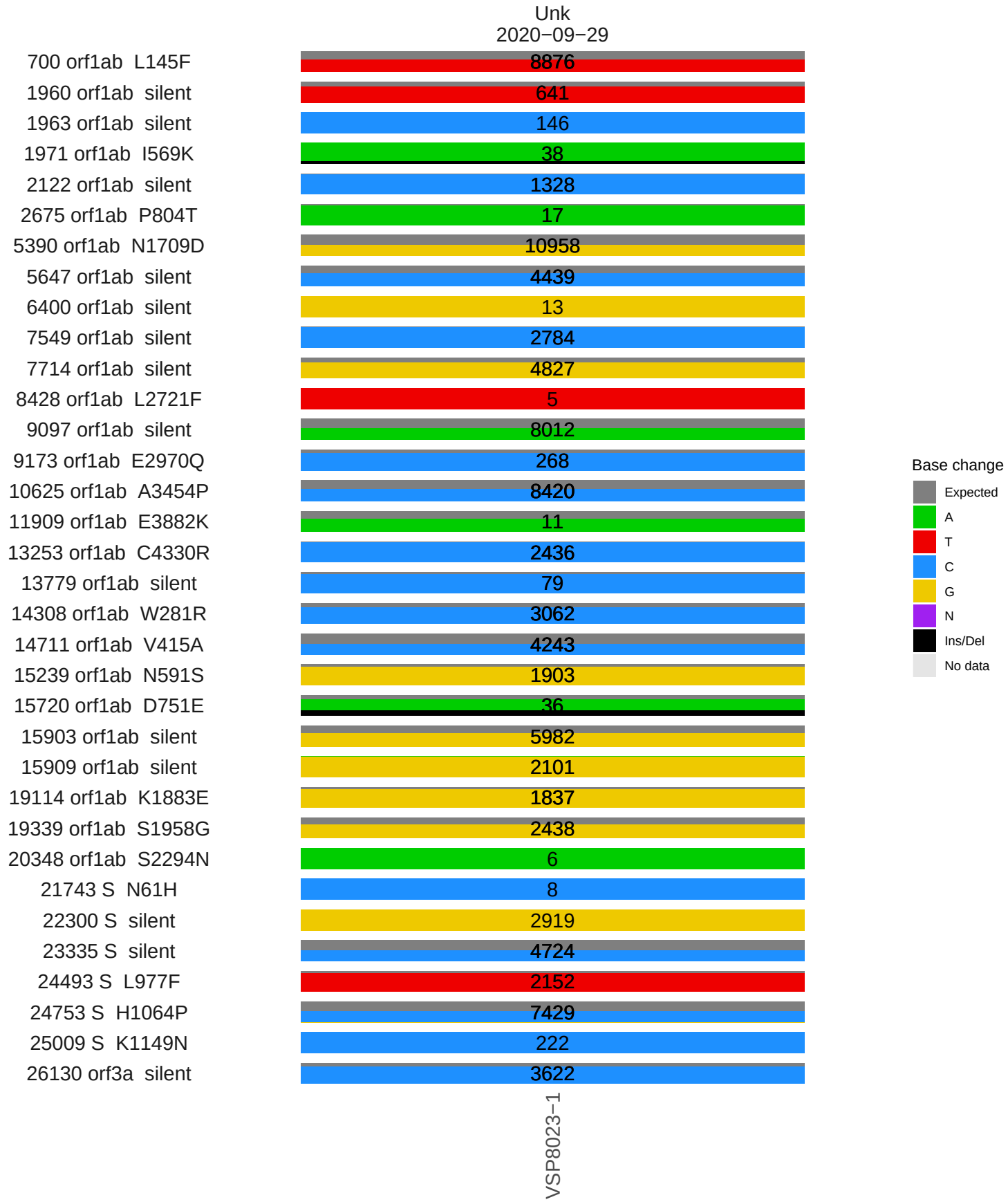
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP8023-1	single experiment	NA	Unk	2020-09-29	2.52	96.0%	95.4%

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.

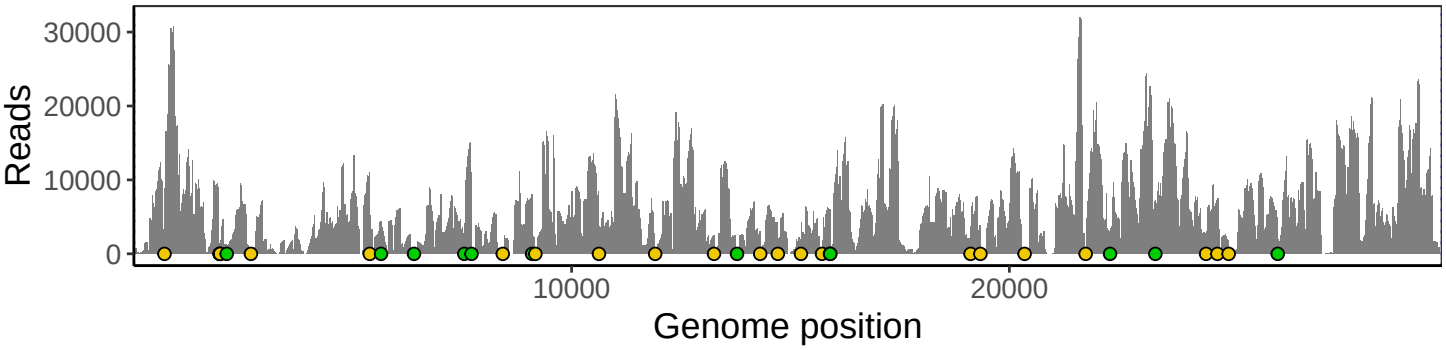



Unk
2020−09−29
700 orf1ab L145F
1960 orf1ab silent
1963 orf1ab silent
1971 orf1ab I569K
2122 orf1ab silent
2675 orf1ab P804T
5390 orf1ab N1709D
5647 orf1ab silent
6400 orf1ab silent
7549 orf1ab silent
7714 orf1ab silent
8428 orf1ab L2721F
9097 orf1ab silent
9173 orf1ab E2970Q
10625 orf1ab A3454P
11909 orf1ab E3882K
13253 orf1ab C4330R
13779 orf1ab silent
14308 orf1ab W281R
14711 orf1ab V415A
15239 orf1ab N591S
15720 orf1ab D751E
15903 orf1ab silent
15909 orf1ab silent
19114 orf1ab K1883E
19339 orf1ab S1958G
20348 orf1ab S2294N
21743 S N61H
22300 S silent
23335 S silent
24493 S L977F
24753 S H1064P
25009 S K1149N
26130 orf3a silent
8876
641
146
38
1328
17
10958
4439
13
2784
4827
5
8012
268
8420
11
2436
79
3062
4243
1903
36
5982
2101
1837
2438
6
8
2919
4724
2152
7429
222
3622
VSP8023−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

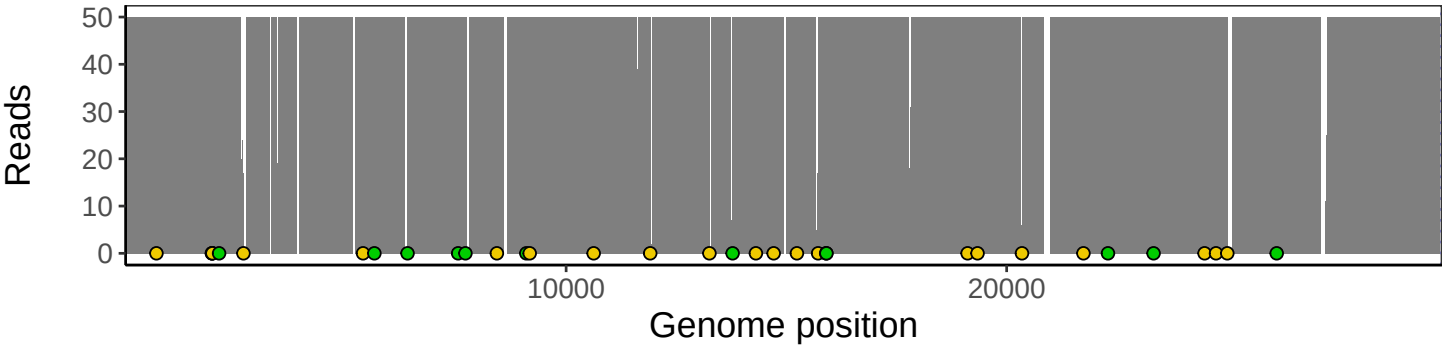
Analyses of individual experiments and composite results.

VSP8023-1 | 2020-09-29 | Unk | SRR11783625 | 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.