COVID-19 subject SRR11783572

2020-09-29

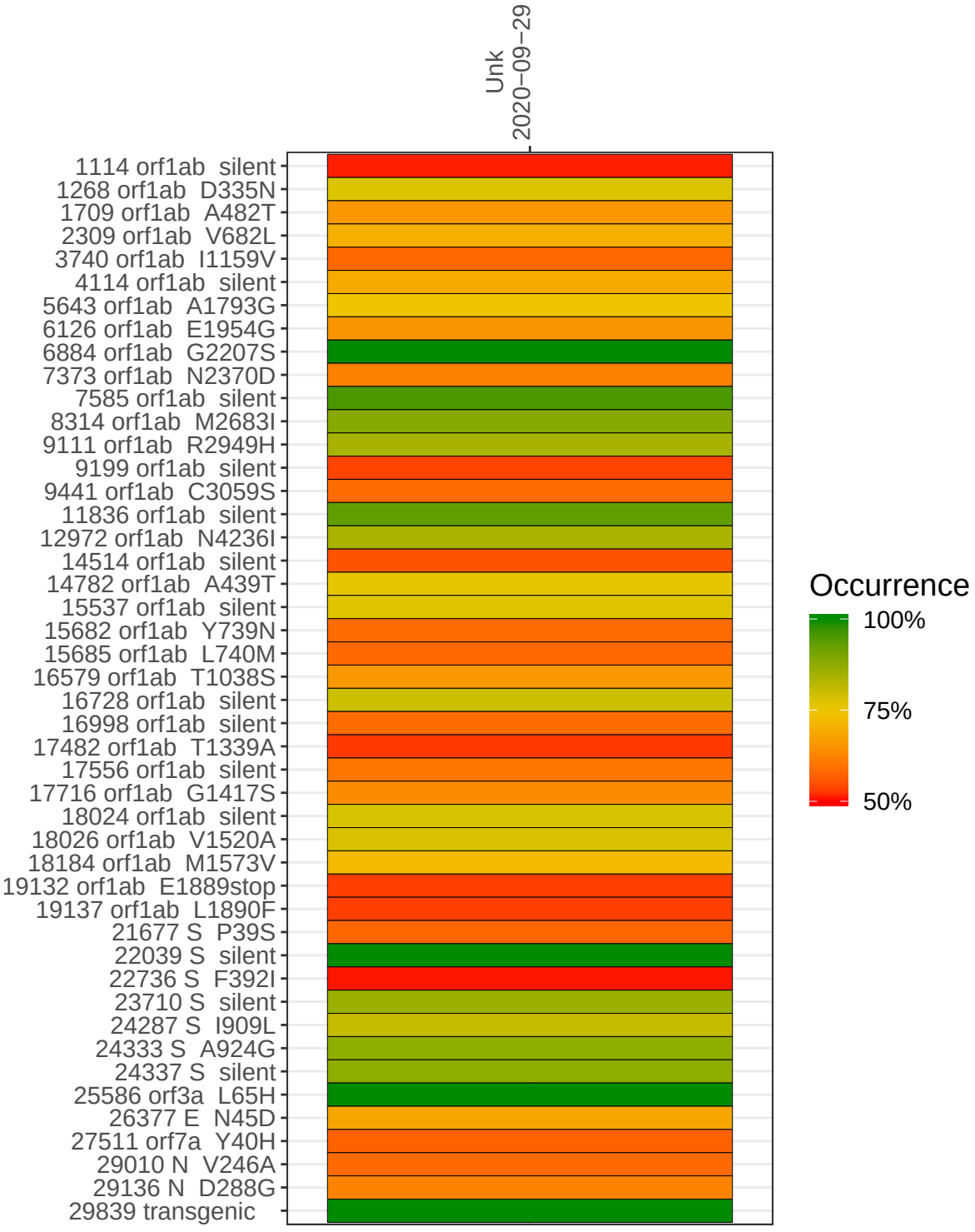
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

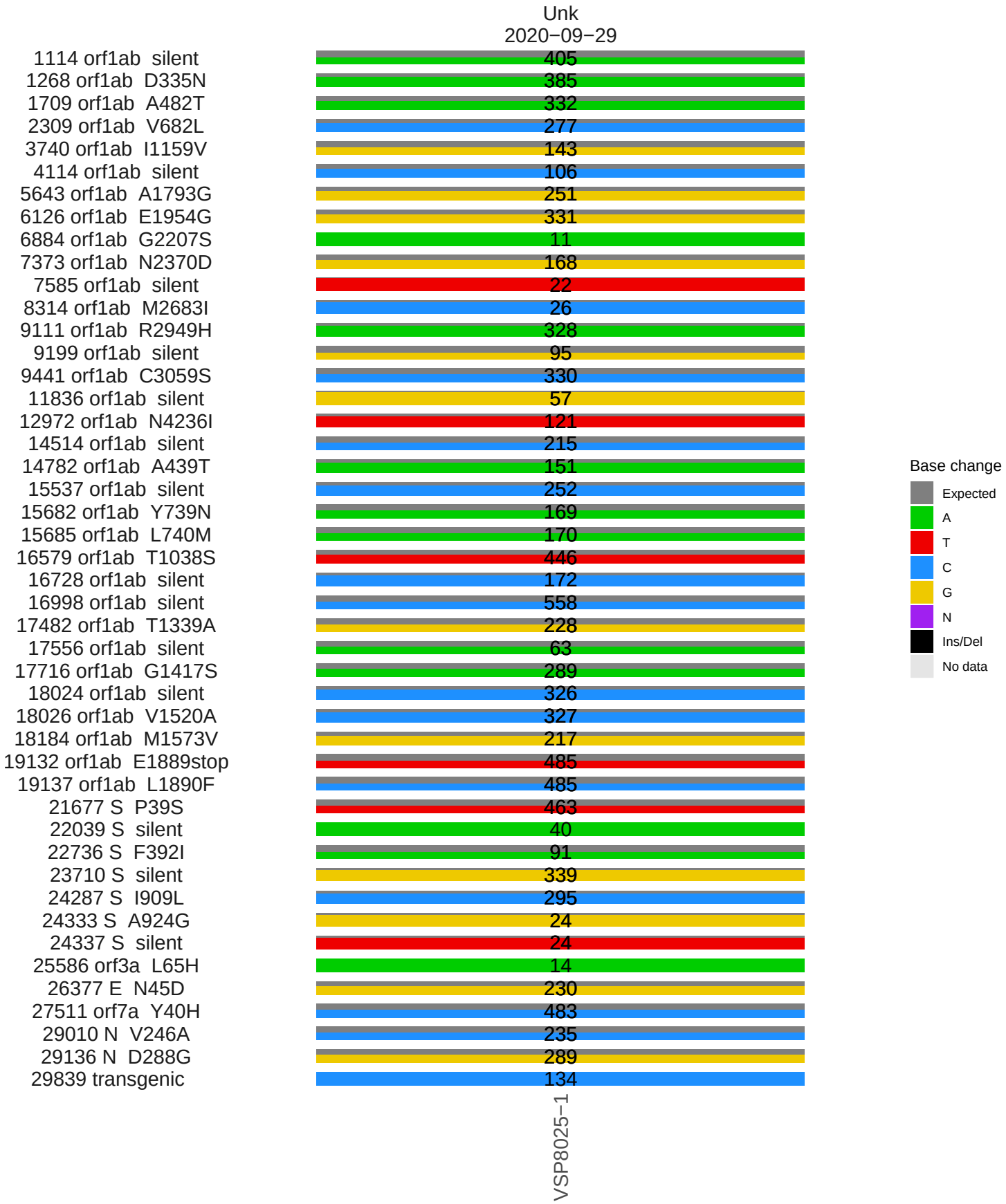
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP8025-1	single experiment	NA	Unk	2020-09-29	2.72	96.0%	94.5%

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.

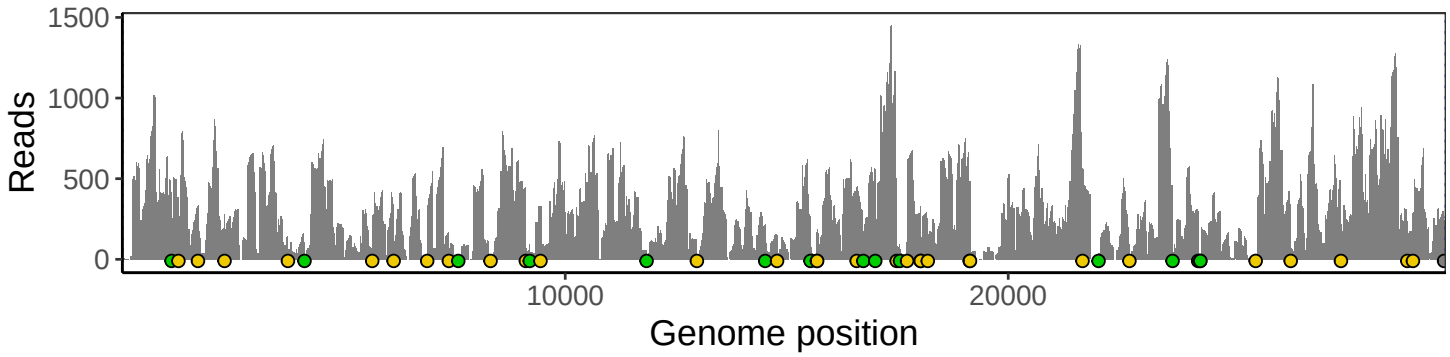



Unk
2020−09−29
1114 orf1ab silent
1268 orf1ab D335N
1709 orf1ab A482T
2309 orf1ab V682L
3740 orf1ab I1159V
4114 orf1ab silent
5643 orf1ab A1793G
6126 orf1ab E1954G
6884 orf1ab G2207S
7373 orf1ab N2370D
7585 orf1ab silent
8314 orf1ab M2683I
9111 orf1ab R2949H
9199 orf1ab silent
9441 orf1ab C3059S
11836 orf1ab silent
12972 orf1ab N4236I
14514 orf1ab silent
14782 orf1ab A439T
15537 orf1ab silent
15682 orf1ab Y739N
15685 orf1ab L740M
16579 orf1ab T1038S
16728 orf1ab silent
16998 orf1ab silent
17482 orf1ab T1339A
17556 orf1ab silent
17716 orf1ab G1417S
18024 orf1ab silent
18026 orf1ab V1520A
18184 orf1ab M1573V
19132 orf1ab E1889stop
19137 orf1ab L1890F
21677 S P39S
22039 S silent
22736 S F392I
23710 S silent
24287 S I909L
24333 S A924G
24337 S silent
25586 orf3a L65H
26377 E N45D
27511 orf7a Y40H
29010 N V246A
29136 N D288G
29839 transgenic
405
385
332
277
143
106
251
331
11
168
22
26
328
95
330
57
121
215
151
252
169
170
446
172
558
228
63
289
326
327
217
485
485
463
40
91
339
295
24
24
14
230
483
235
289
134
VSP8025−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

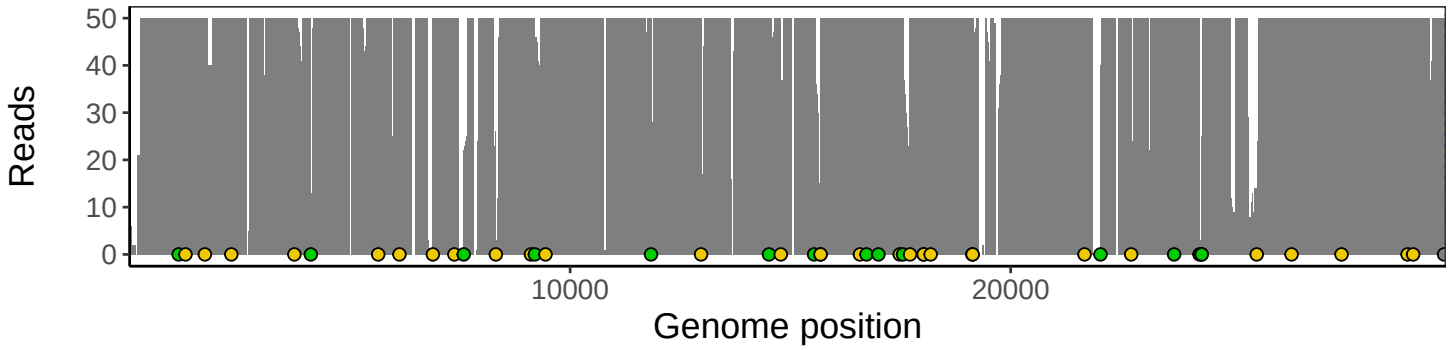
Analyses of individual experiments and composite results.

VSP8025-1 | 2020-09-29 | Unk | SRR11783572 | 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.