COVID-19 subject 307

2020-09-09

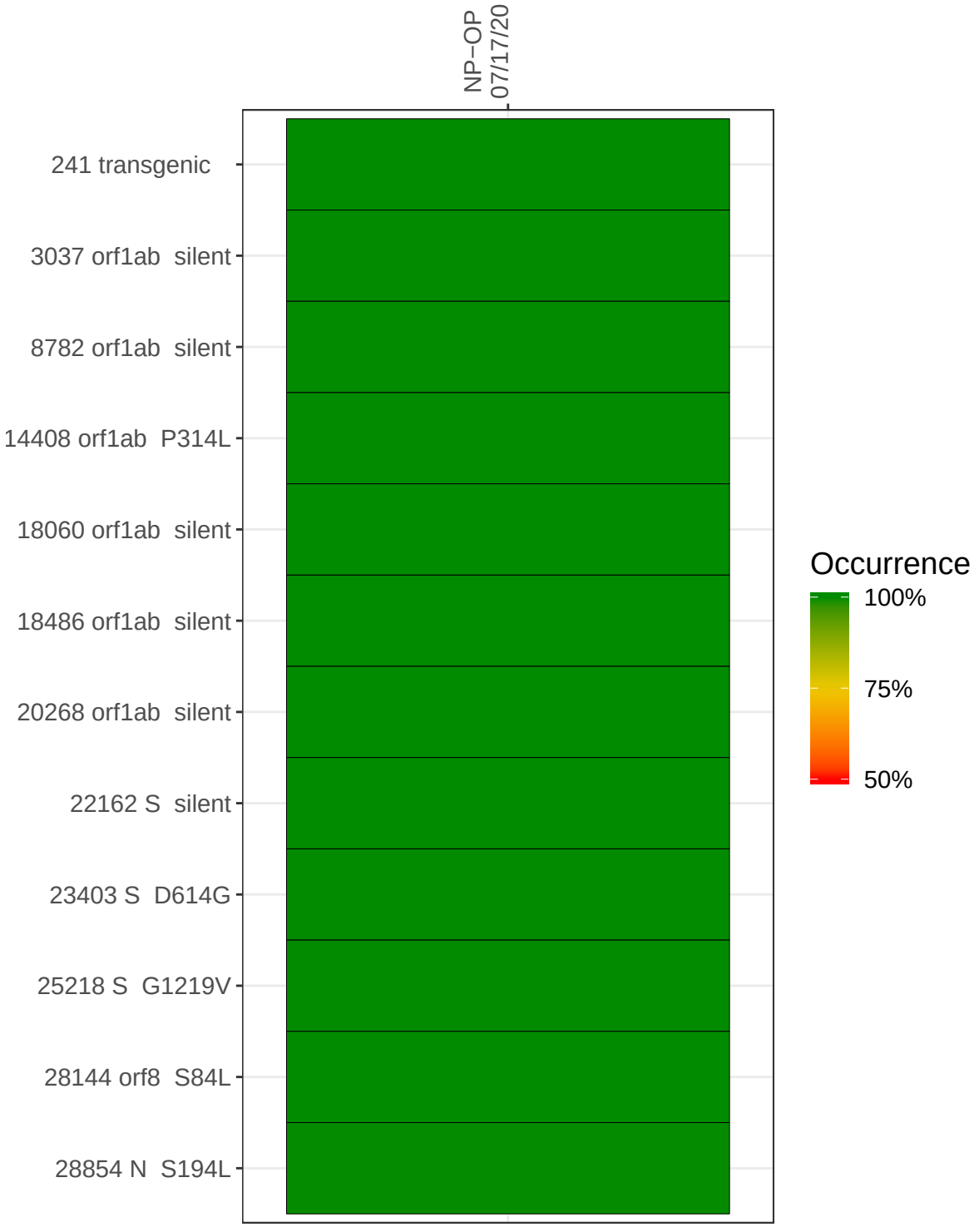
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

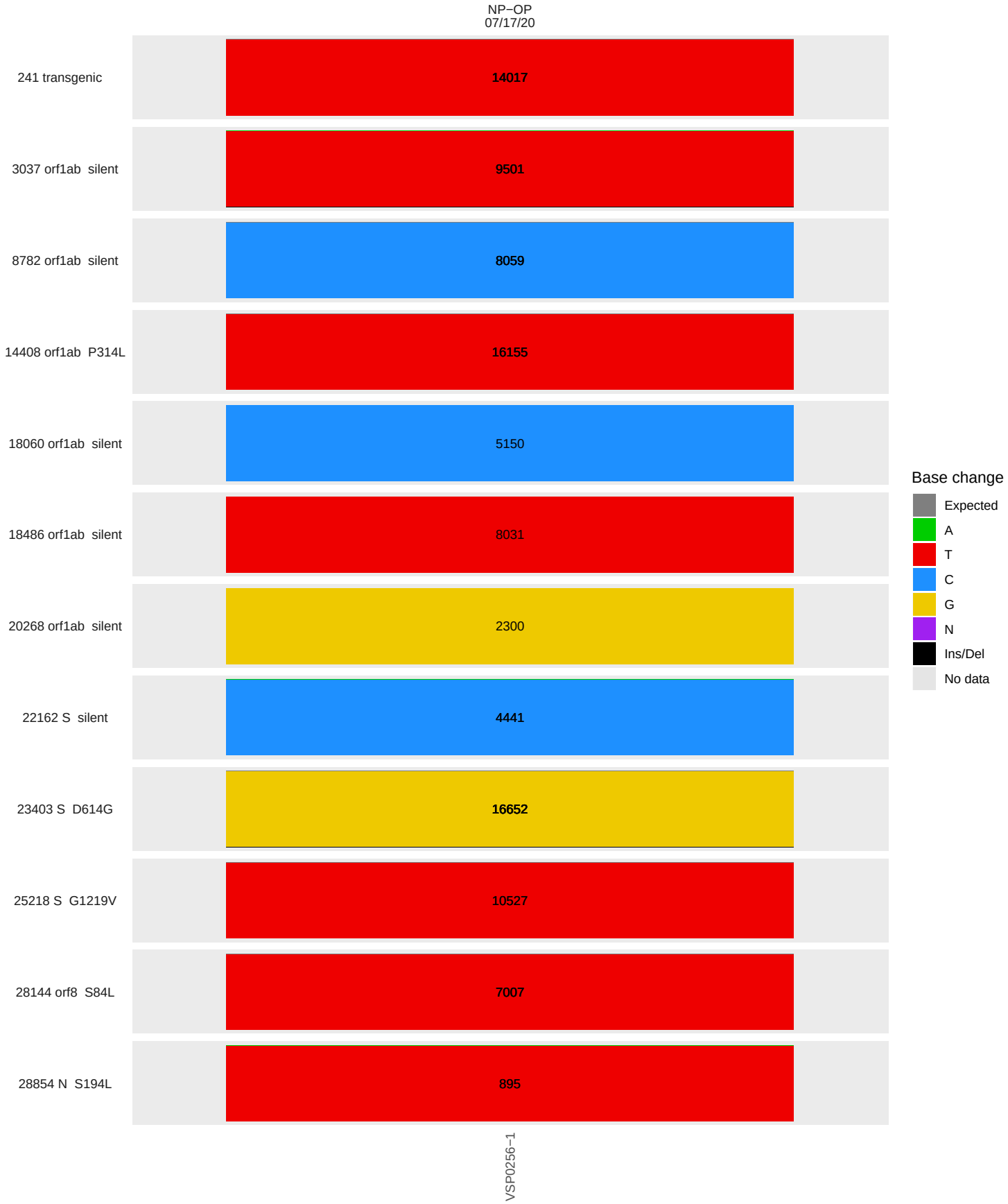
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0256-1	single experiment	313.5	NP-OP	07/17/20	22.60	99.1%	99.1%

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.

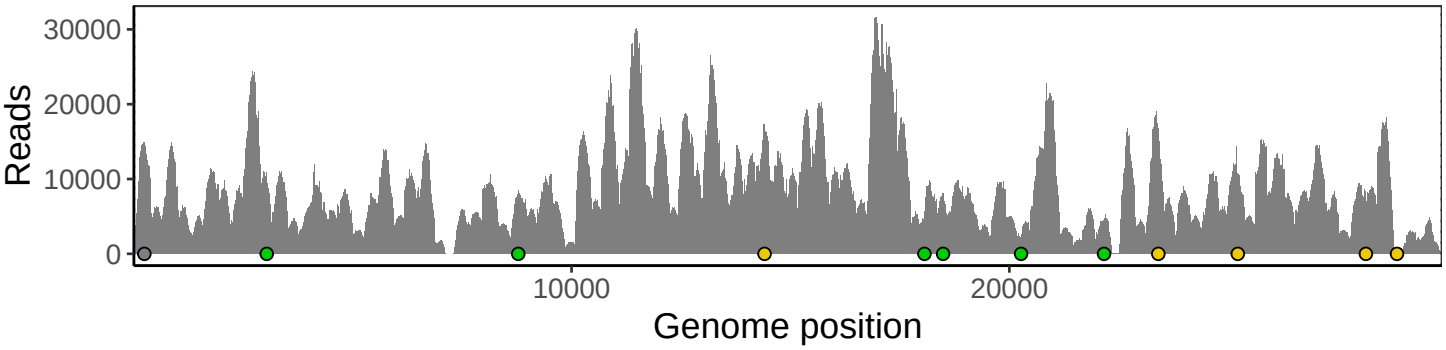



NP−OP
07/17/20
241 transgenic
14017
3037 orf1ab silent
9501
8782 orf1ab silent
8059
14408 orf1ab P314L
16155
18060 orf1ab silent
5150
18486 orf1ab silent
8031
20268 orf1ab silent
2300
22162 S silent
4441
23403 S D614G
16652
25218 S G1219V
10527
28144 orf8 S84L
7007
28854 N S194L
895
Base change
Expected
A
T
C
G
N
Ins/Del
No data
VSP0256−1

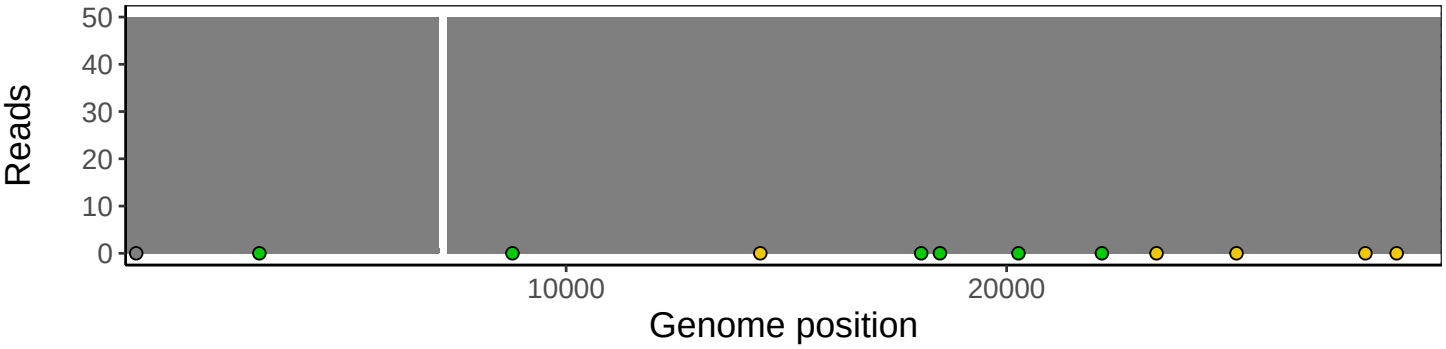
Analyses of individual experiments and composite results.

VSP0256-1 | 07/17/20 | NP-OP | 307no-q | 313.5 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 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.