COVID-19 subject 304

2021-01-10

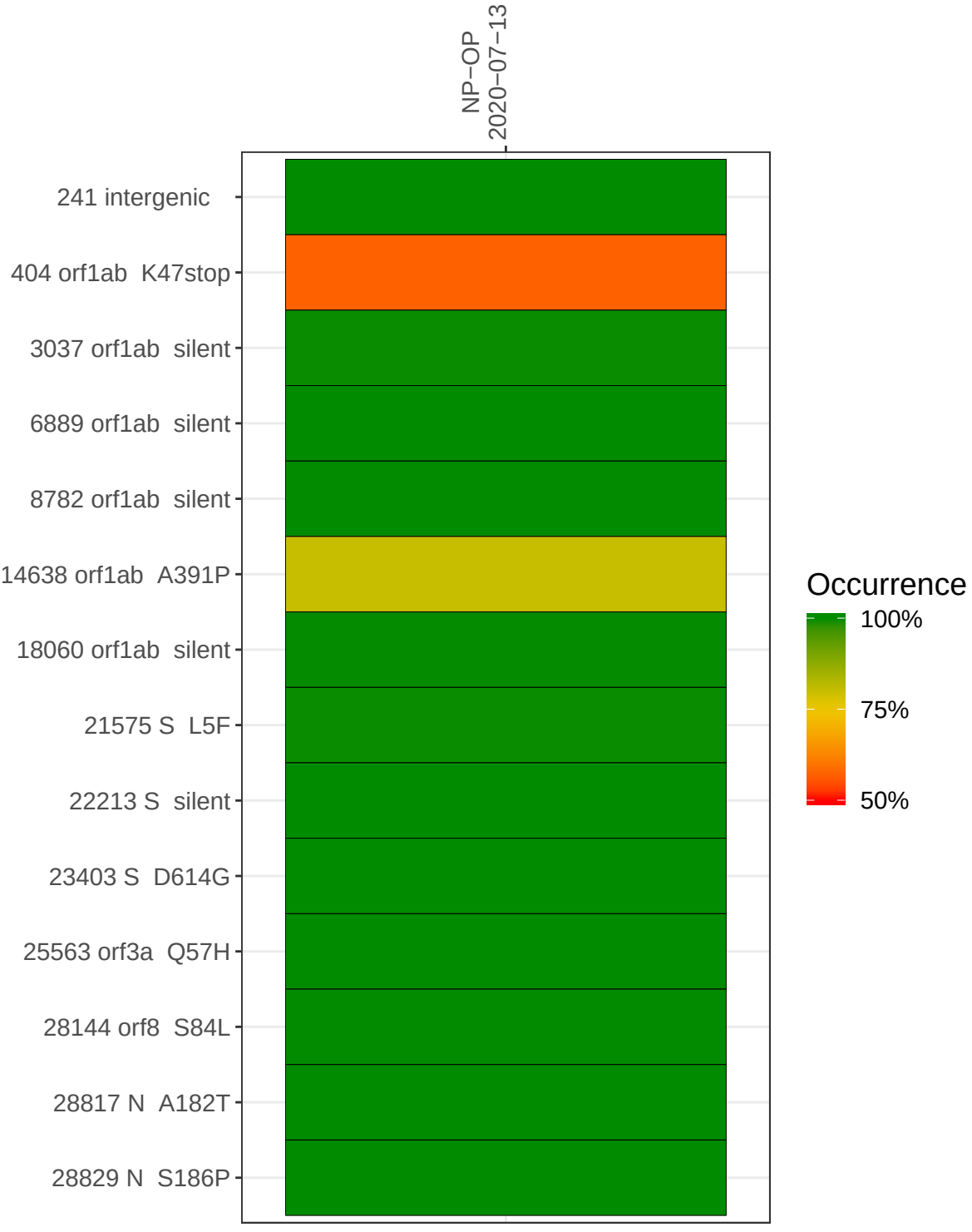
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

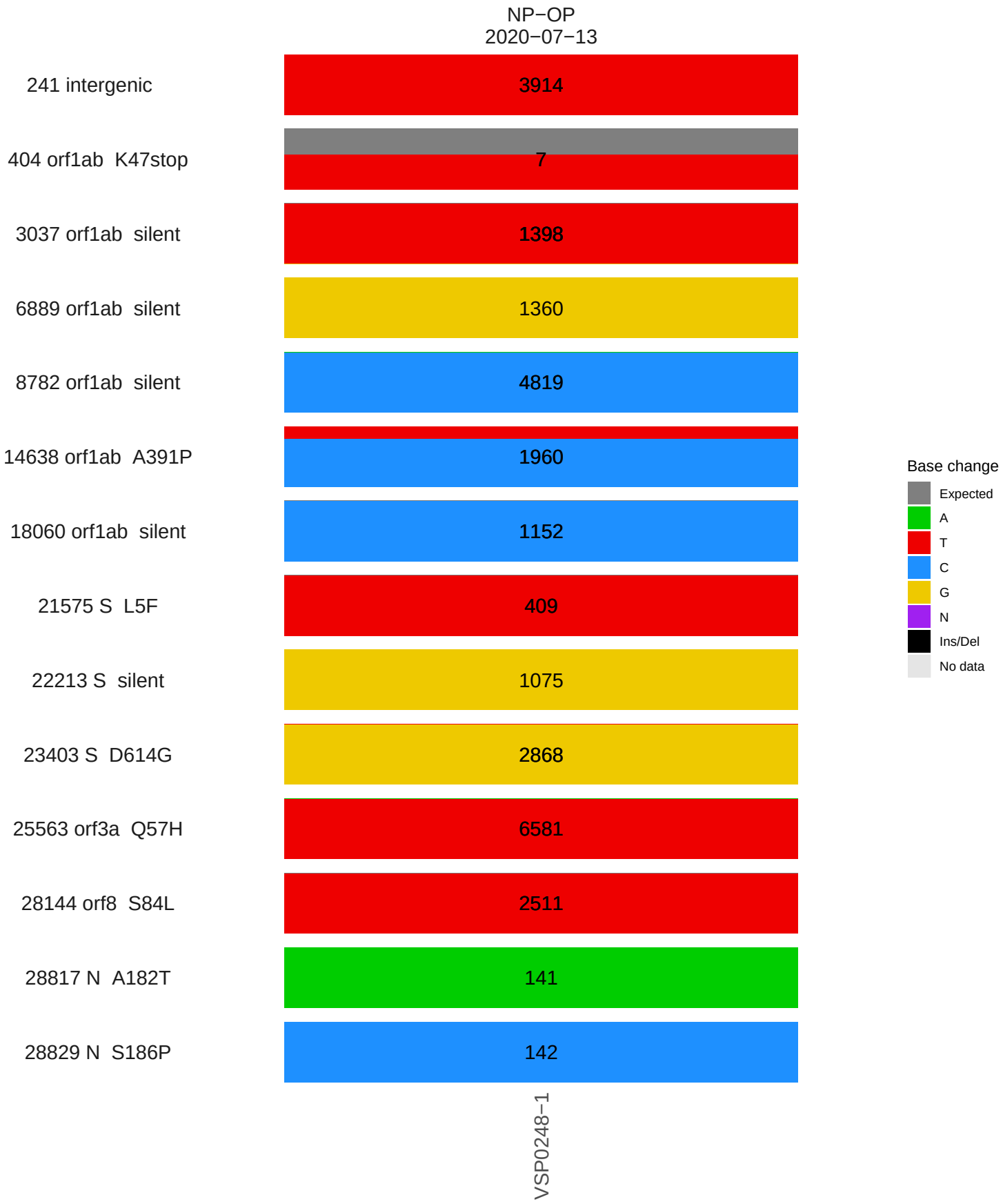
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0248-1	single experiment	NA	NP-OP	2020-07-13	8.38	88.4%	83.1%

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.

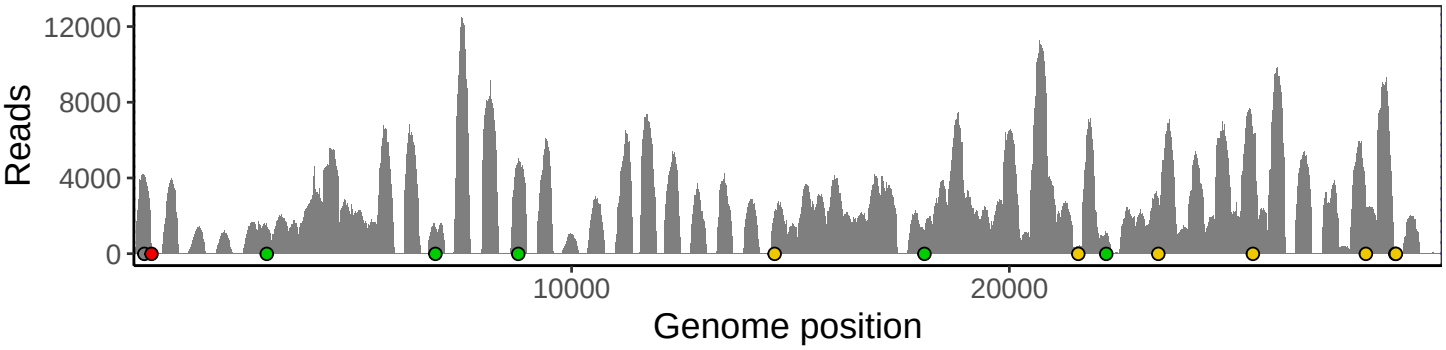



NP−OP
2020−07−13
241 intergenic
3914
404 orf1ab K47stop
7
3037 orf1ab silent
1398
6889 orf1ab silent
1360
8782 orf1ab silent
4819
14638 orf1ab A391P
1960
18060 orf1ab silent
1152
21575 S L5F
409
22213 S silent
1075
23403 S D614G
2868
25563 orf3a Q57H
6581
28144 orf8 S84L
2511
28817 N A182T
141
28829 N S186P
142
VSP0248−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

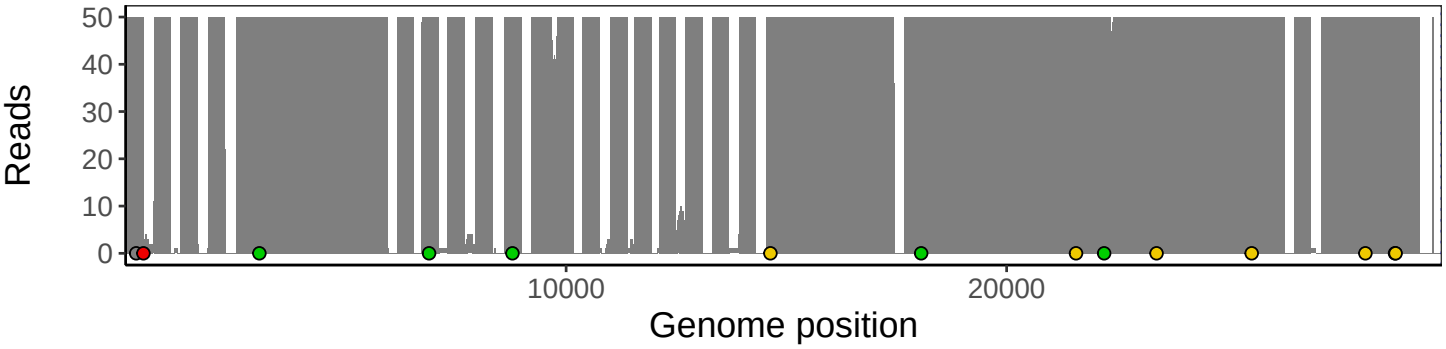
Analyses of individual experiments and composite results.

VSP0248-1 | 2020-07-13 | NP-OP | 304no-q | NA 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.