

COVID-19 subject 257

2020-09-14

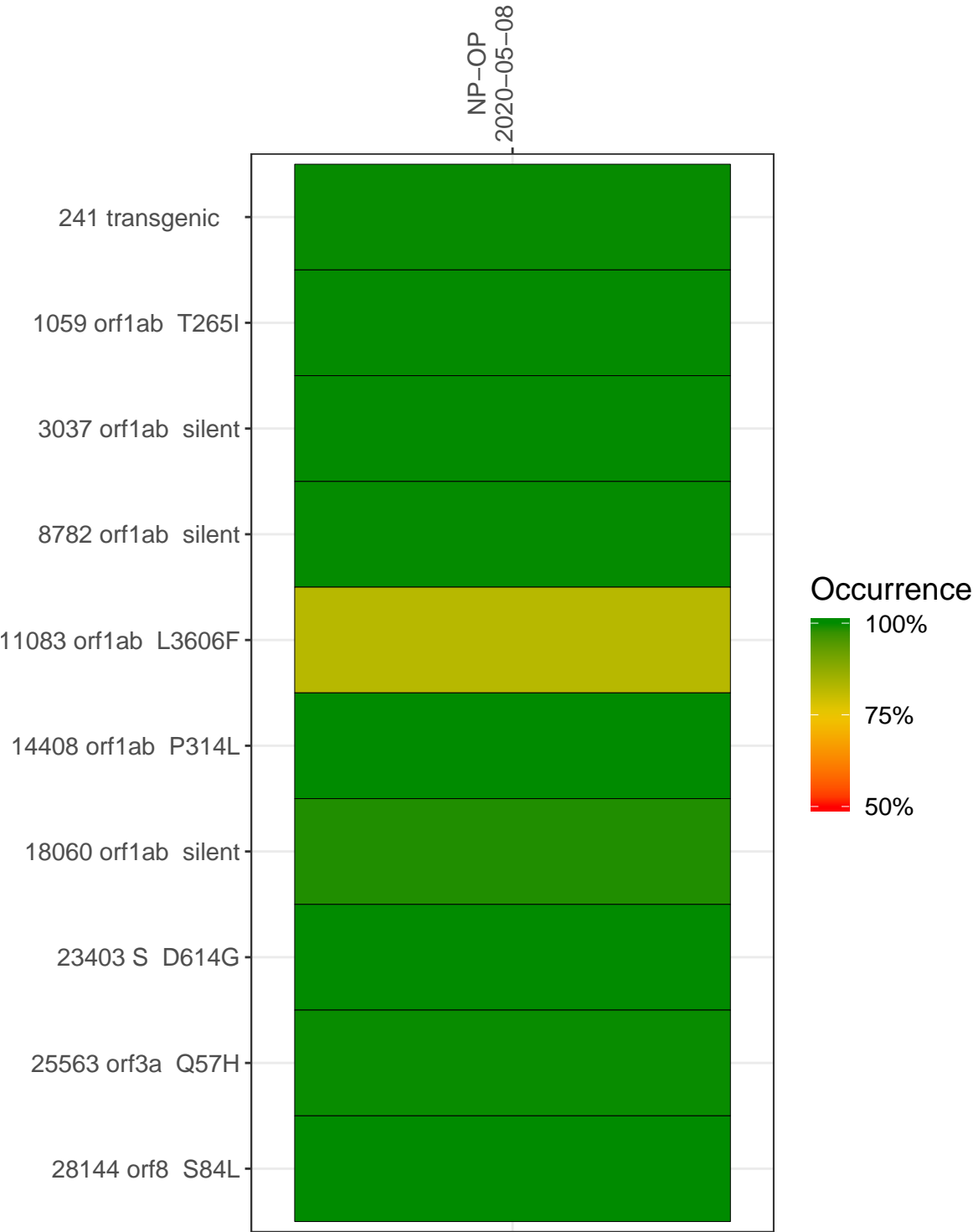
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0102-1	single experiment	51800	NP-OP	2020-05-08	29.93	99.8%	99.8%

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
2020-05-08

241 transgenic

1059 orf1ab T265I

3037 orf1ab silent

8782 orf1ab silent

11083 orf1ab L3606F

14408 orf1ab P314L

18060 orf1ab silent

23403 S D614G

25563 orf3a Q57H

28144 orf8 S84L

4091

3106

2563

2259

2707

4069

2553

6077

3763

10372

Base change

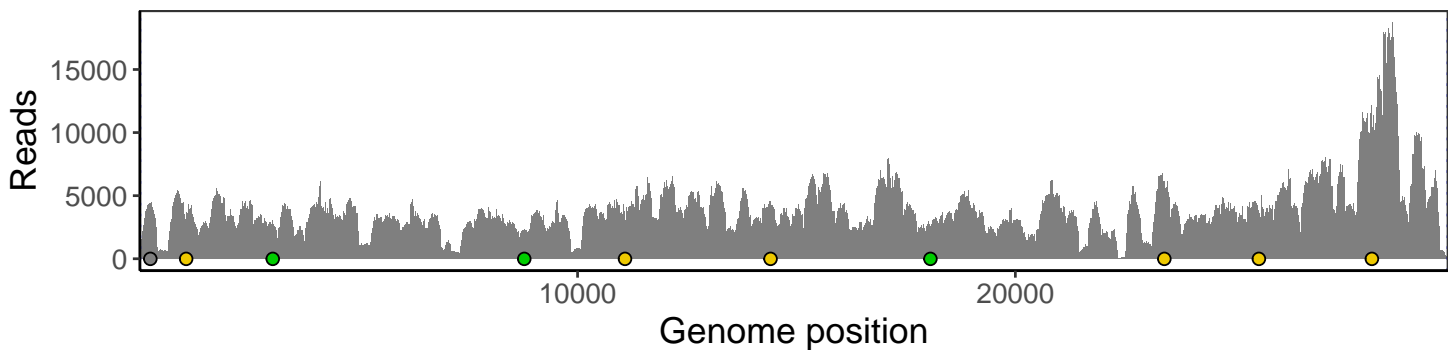
- Expected
- A
- T
- C
- G
- N
- Ins/Del
- No data

VSP0102-1

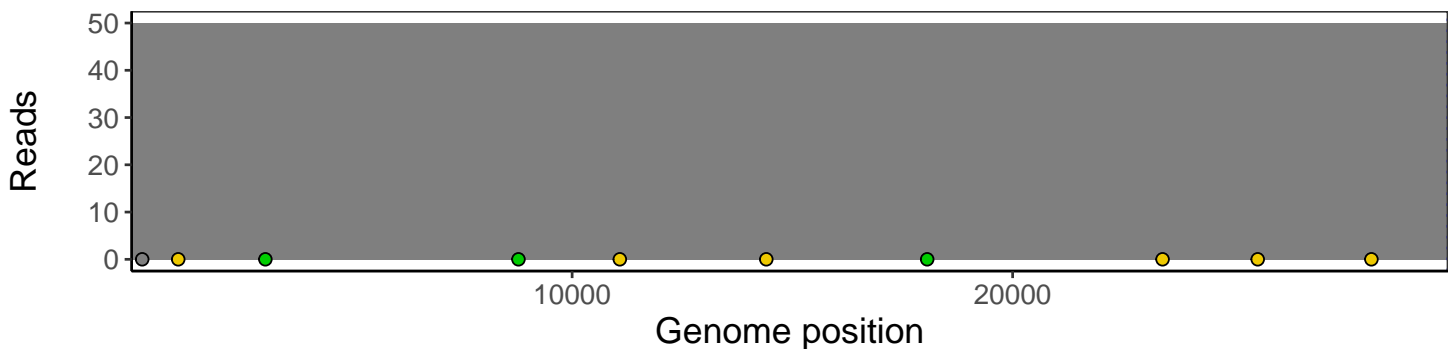
Analyses of individual experiments and composite results.

VSP0102-1 | 2020-05-08 | NP-OP | 257no-q | 51800 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 to 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.

