COVID-19 subject deWit_RM10_Vehicle

2020-10-02

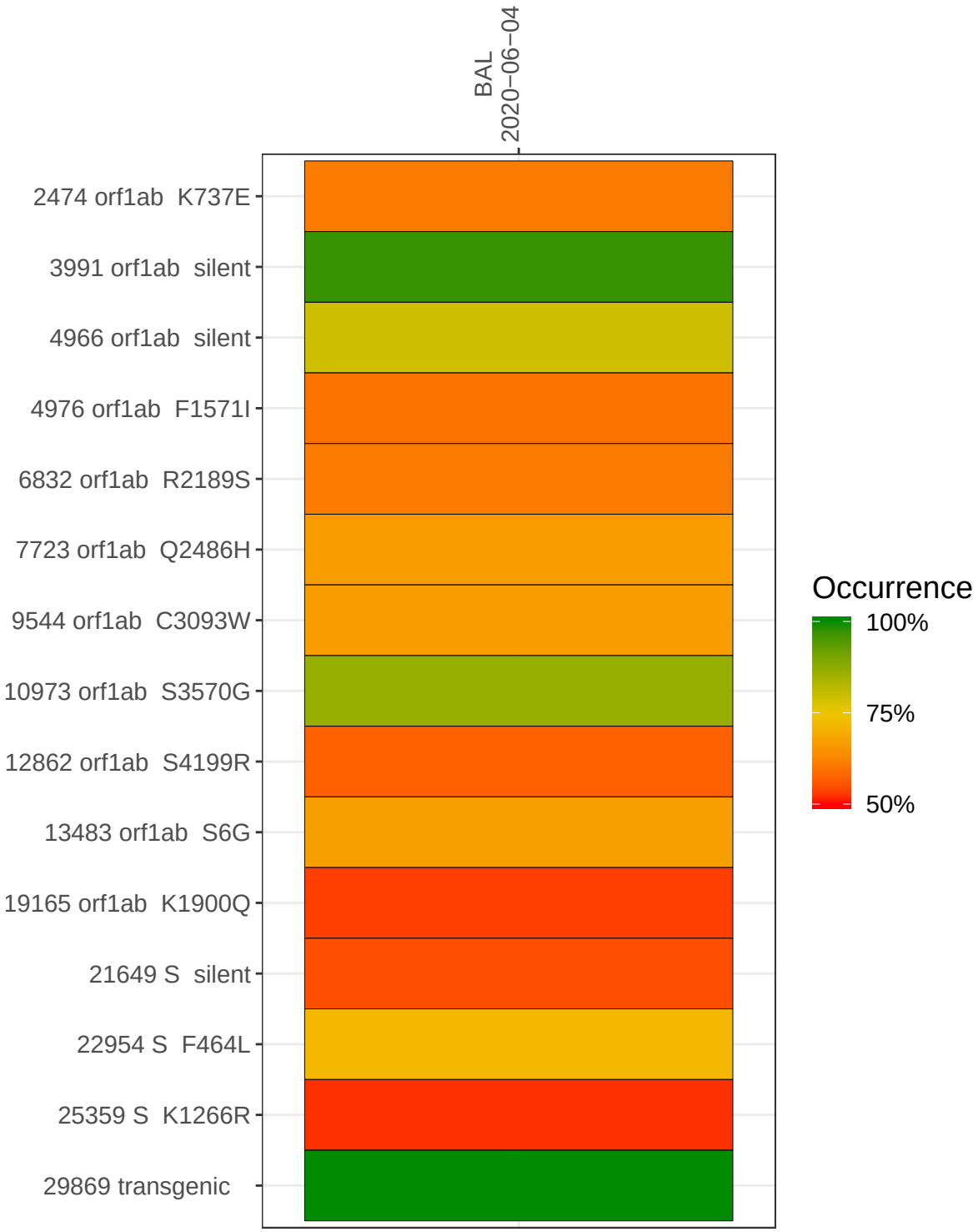
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

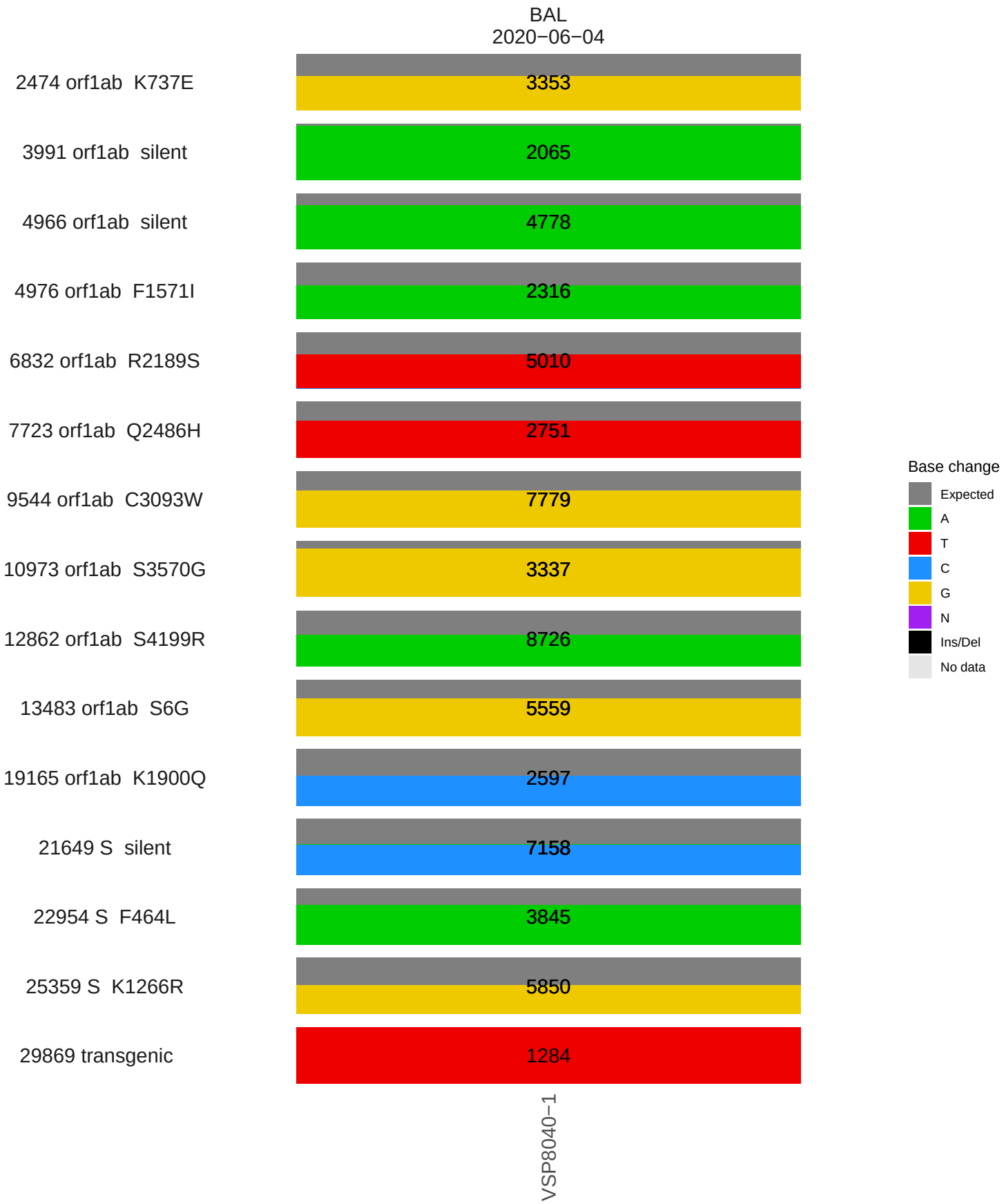
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP8040-1	single experiment	NA	BAL	2020-06-04	4.04	99.7%	99.3%

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.

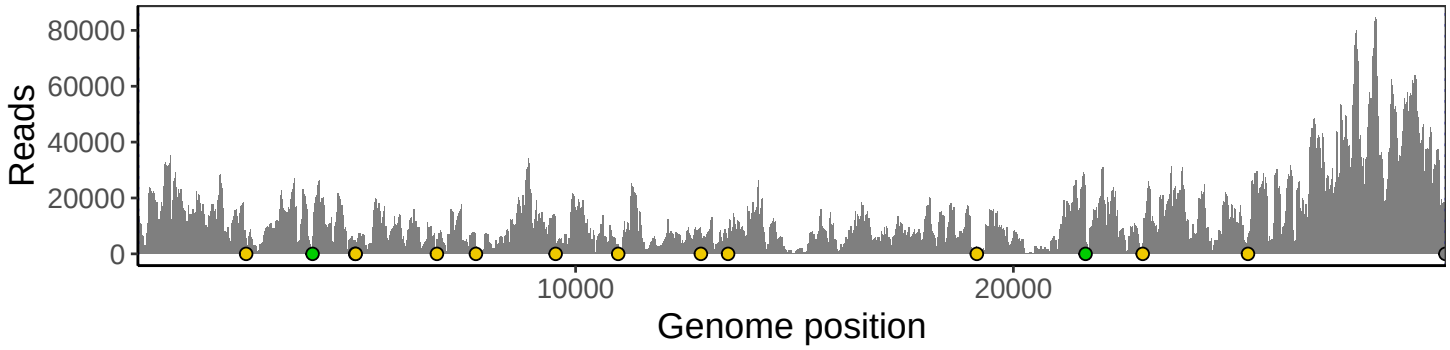



BAL
2020−06−04
2474 orf1ab K737E
3353
3991 orf1ab silent
2065
4966 orf1ab silent
4778
4976 orf1ab F1571I
2316
6832 orf1ab R2189S
5010
7723 orf1ab Q2486H
2751
9544 orf1ab C3093W
7779
10973 orf1ab S3570G
3337
12862 orf1ab S4199R
8726
13483 orf1ab S6G
5559
19165 orf1ab K1900Q
2597
21649 S silent
7158
22954 S F464L
3845
25359 S K1266R
5850
29869 transgenic
1284
VSP8040−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

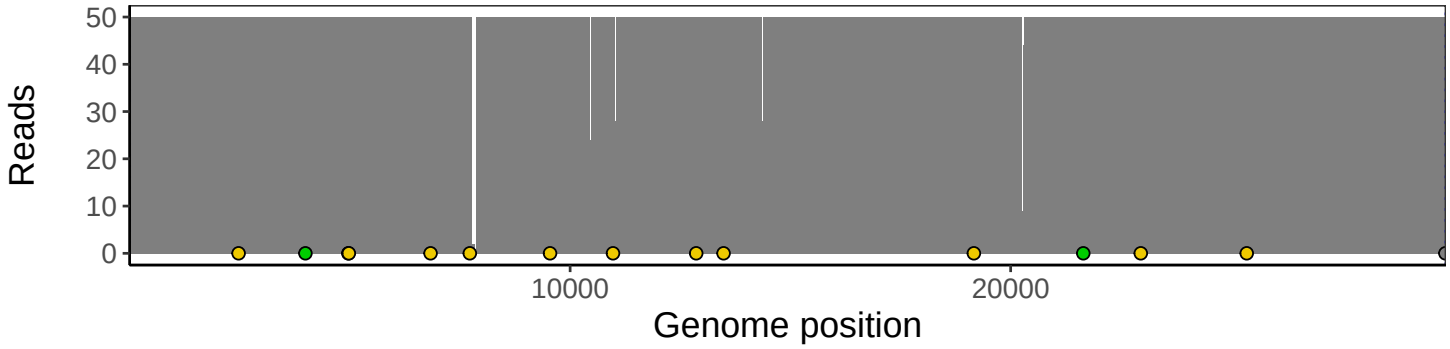
Analyses of individual experiments and composite results.

VSP8040-1 | 2020-06-04 | BAL | SRR11783589 | 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.