COVID-19 subject 2752

2021-01-19

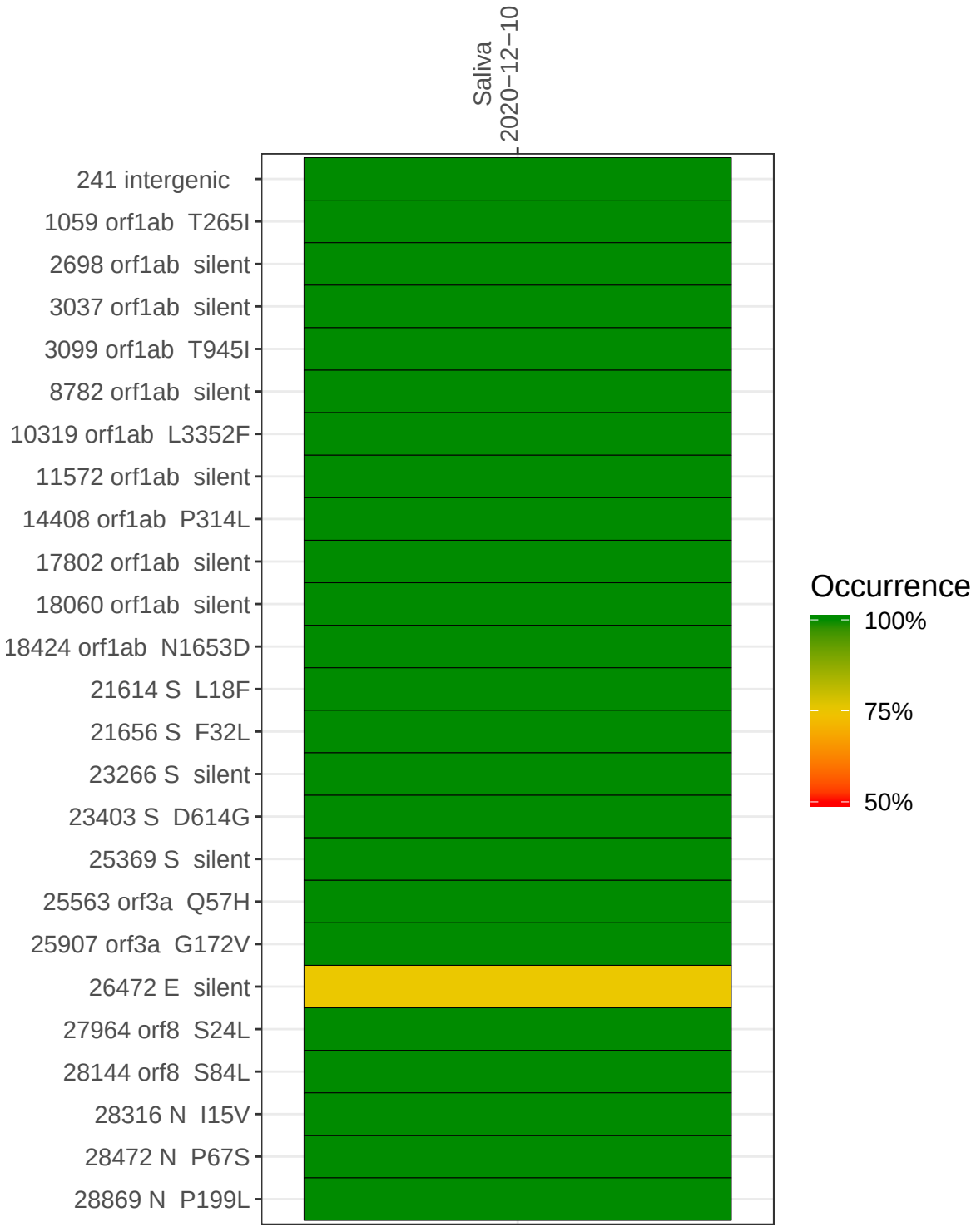
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

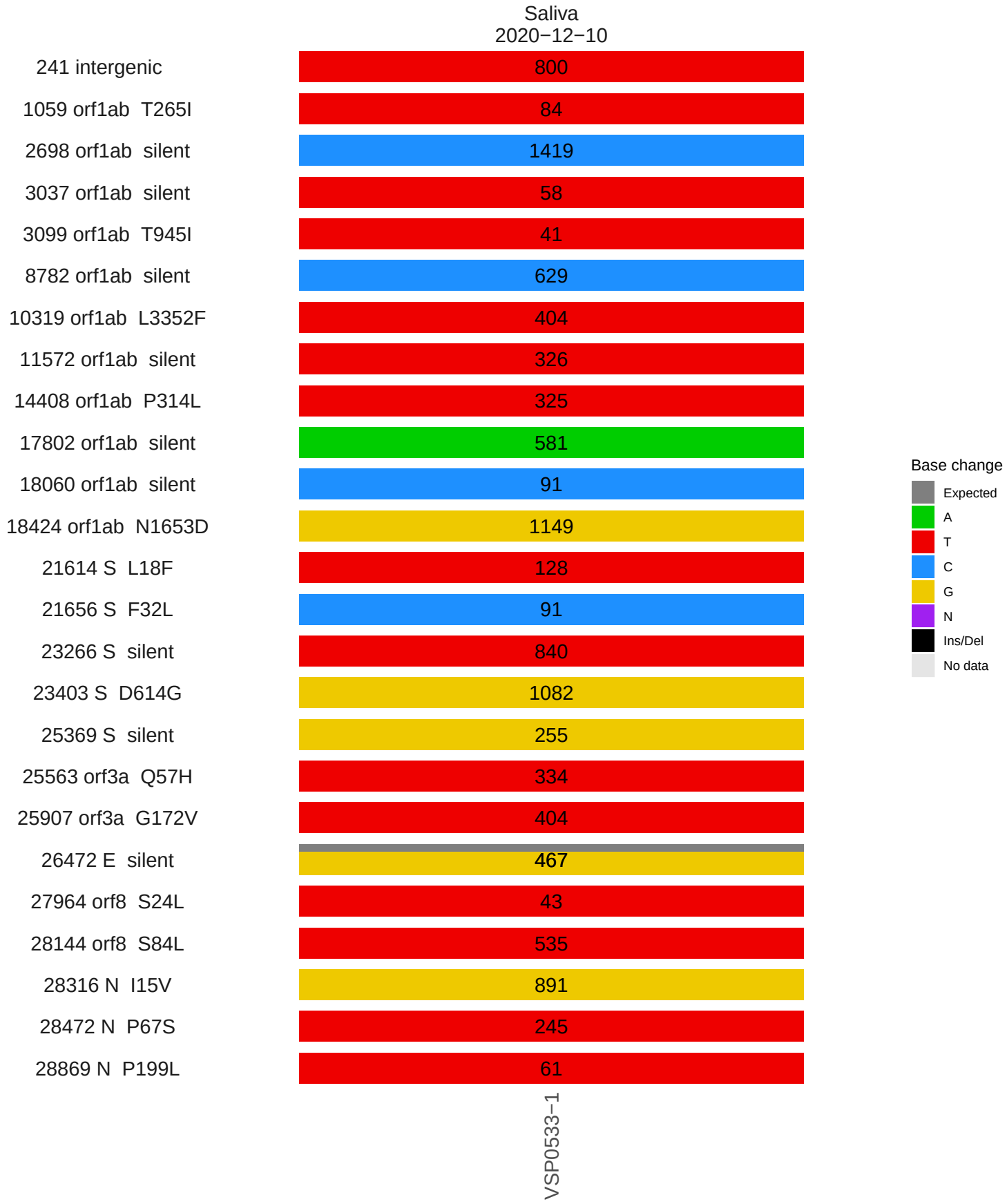
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0533-1	single experiment	NA	Saliva	2020-12-10	19.25	99.4%	97.1%

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 or 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.

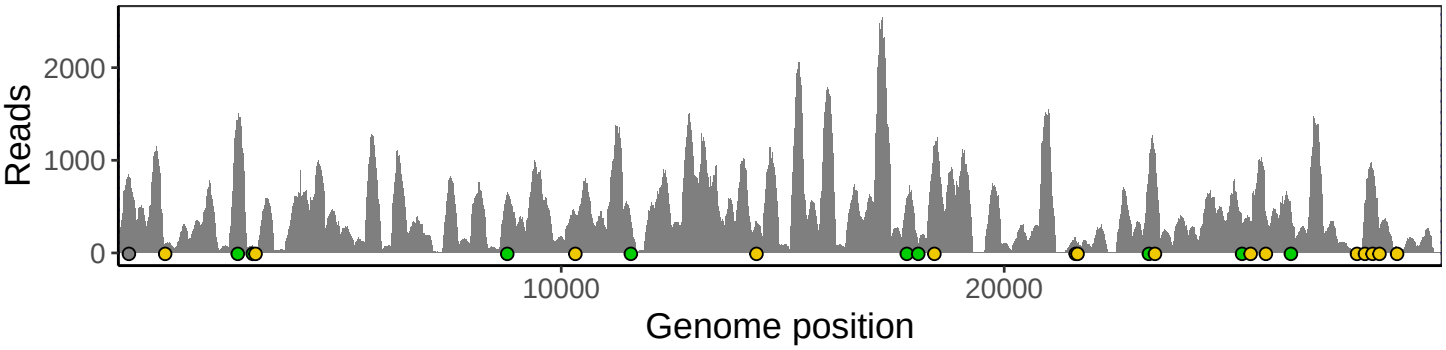



Saliva
2020−12−10
241 intergenic
800
1059 orf1ab T265I
84
2698 orf1ab silent
1419
3037 orf1ab silent
58
3099 orf1ab T945I
41
8782 orf1ab silent
629
10319 orf1ab L3352F
404
11572 orf1ab silent
326
14408 orf1ab P314L
325
17802 orf1ab silent
581
18060 orf1ab silent
91
18424 orf1ab N1653D
1149
21614 S L18F
128
21656 S F32L
91
23266 S silent
840
23403 S D614G
1082
25369 S silent
255
25563 orf3a Q57H
334
25907 orf3a G172V
404
26472 E silent
467
27964 orf8 S24L
43
28144 orf8 S84L
535
28316 N I15V
891
28472 N P67S
245
28869 N P199L
61
VSP0533−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

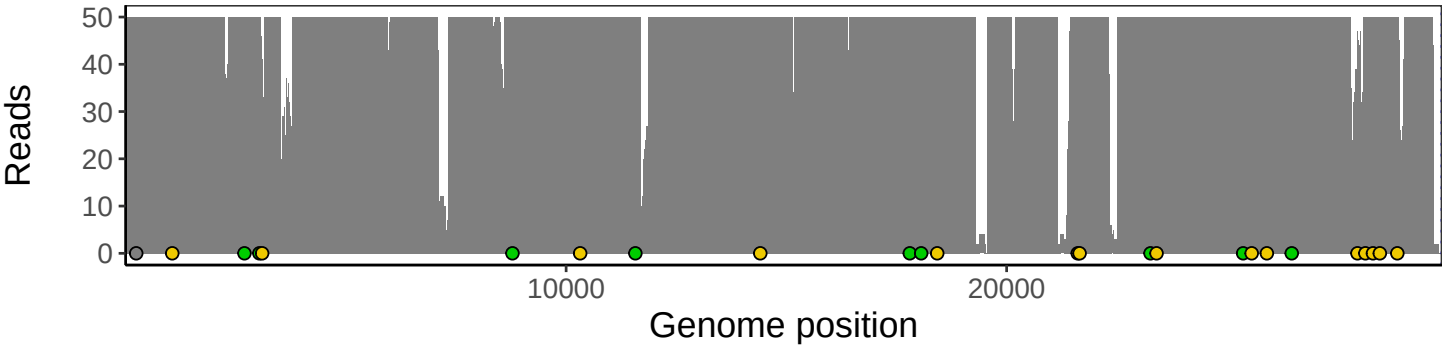
Analyses of individual experiments and composite results.

VSP0533-1 | 2020-12-10 | Saliva | 2752 | 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.