COVID-19 subject 2749

2021-01-08

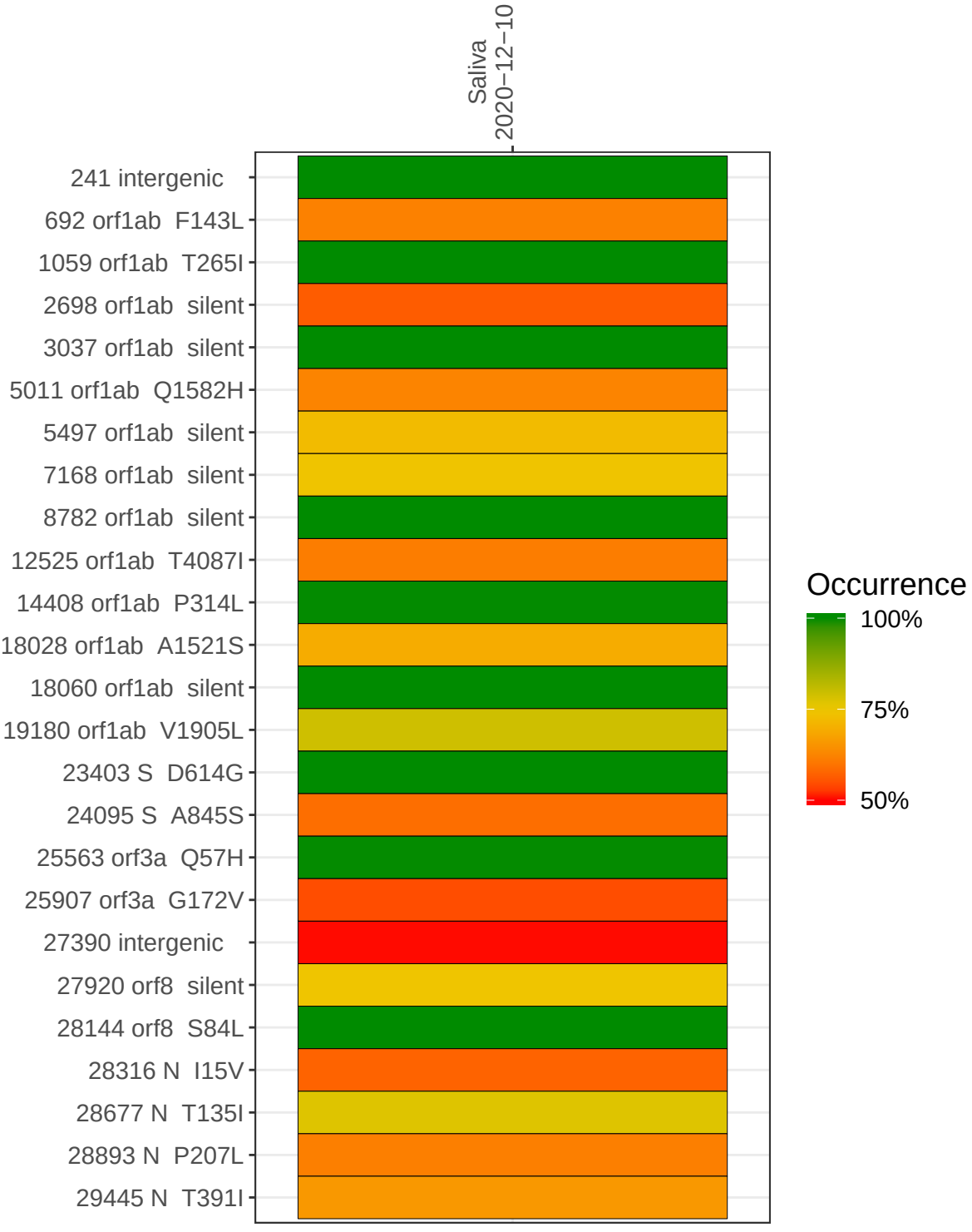
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

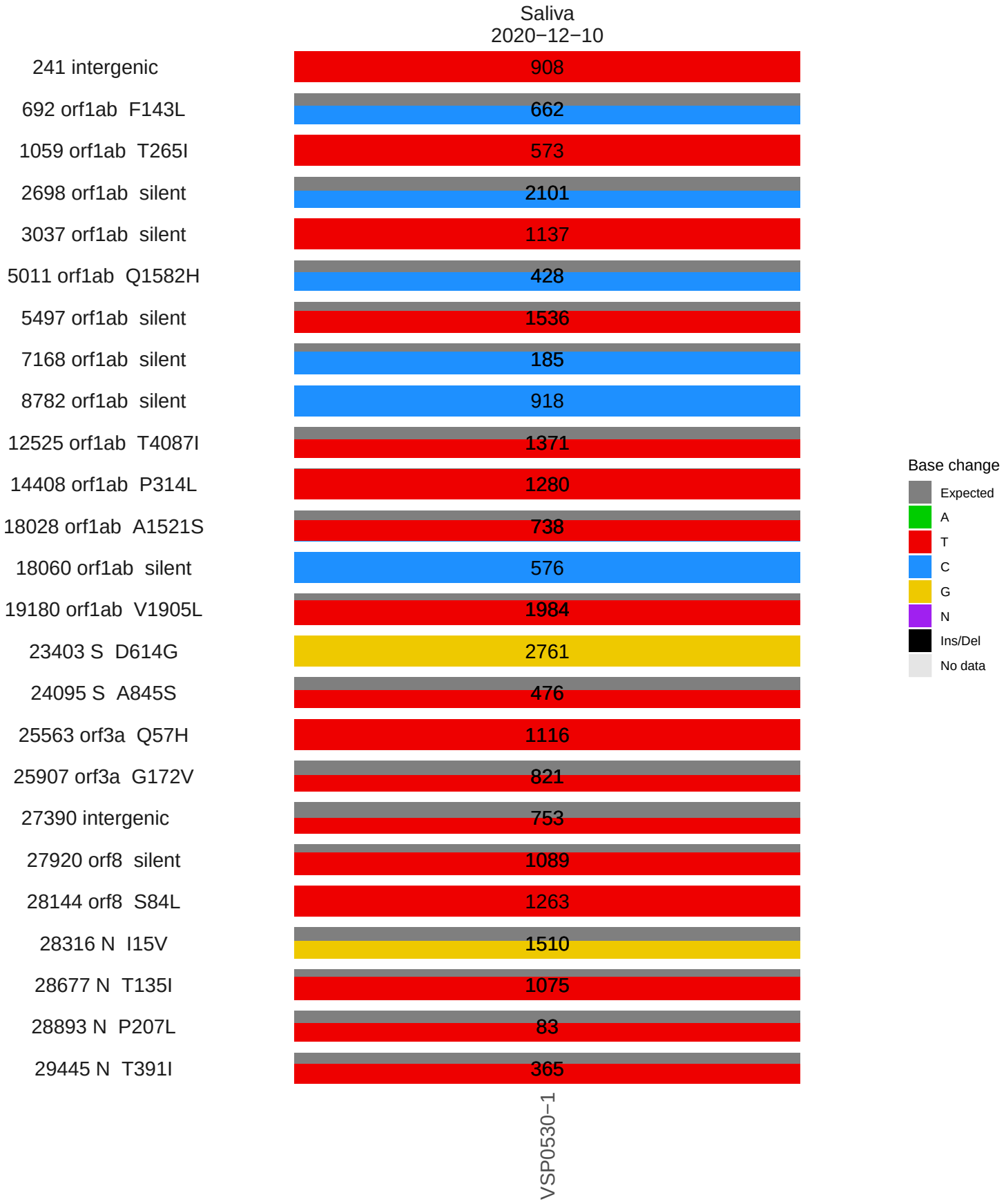
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0530-1	single experiment	NA	Saliva	2020-12-10	21.64	100.0%	99.2%

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.

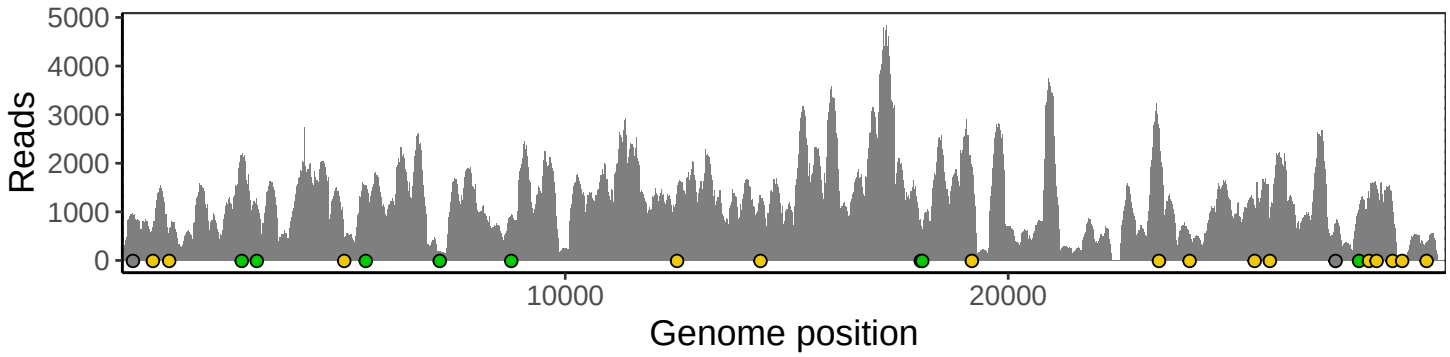



Saliva
2020−12−10
241 intergenic
908
692 orf1ab F143L
662
1059 orf1ab T265I
573
2698 orf1ab silent
2101
3037 orf1ab silent
1137
5011 orf1ab Q1582H
428
5497 orf1ab silent
1536
7168 orf1ab silent
185
8782 orf1ab silent
918
12525 orf1ab T4087I
1371
14408 orf1ab P314L
1280
18028 orf1ab A1521S
738
18060 orf1ab silent
576
19180 orf1ab V1905L
1984
23403 S D614G
2761
24095 S A845S
476
25563 orf3a Q57H
1116
25907 orf3a G172V
821
27390 intergenic
753
27920 orf8 silent
1089
28144 orf8 S84L
1263
28316 N I15V
1510
28677 N T135I
1075
28893 N P207L
83
29445 N T391I
365
VSP0530−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

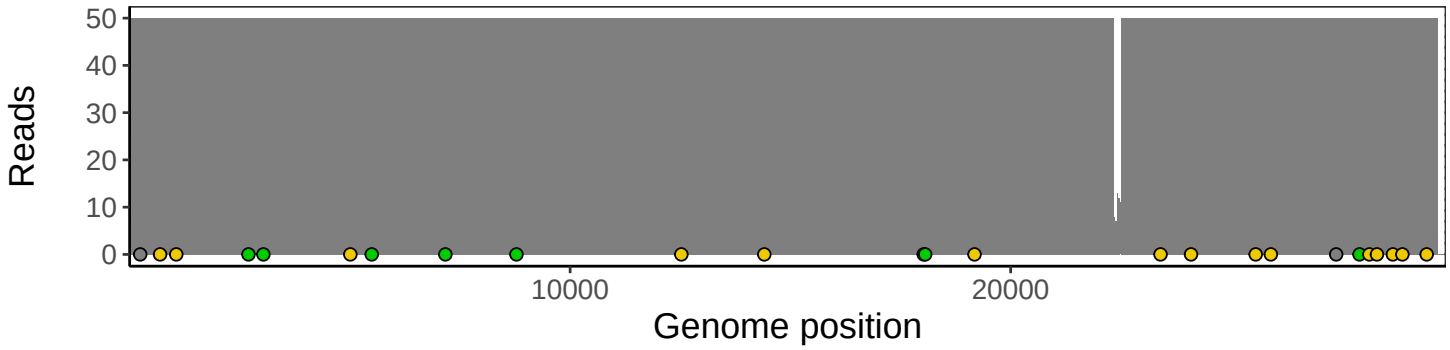
Analyses of individual experiments and composite results.

VSP0530-1 | 2020-12-10 | Saliva | 2749 | 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.