COVID-19 subject 374

2021-01-08

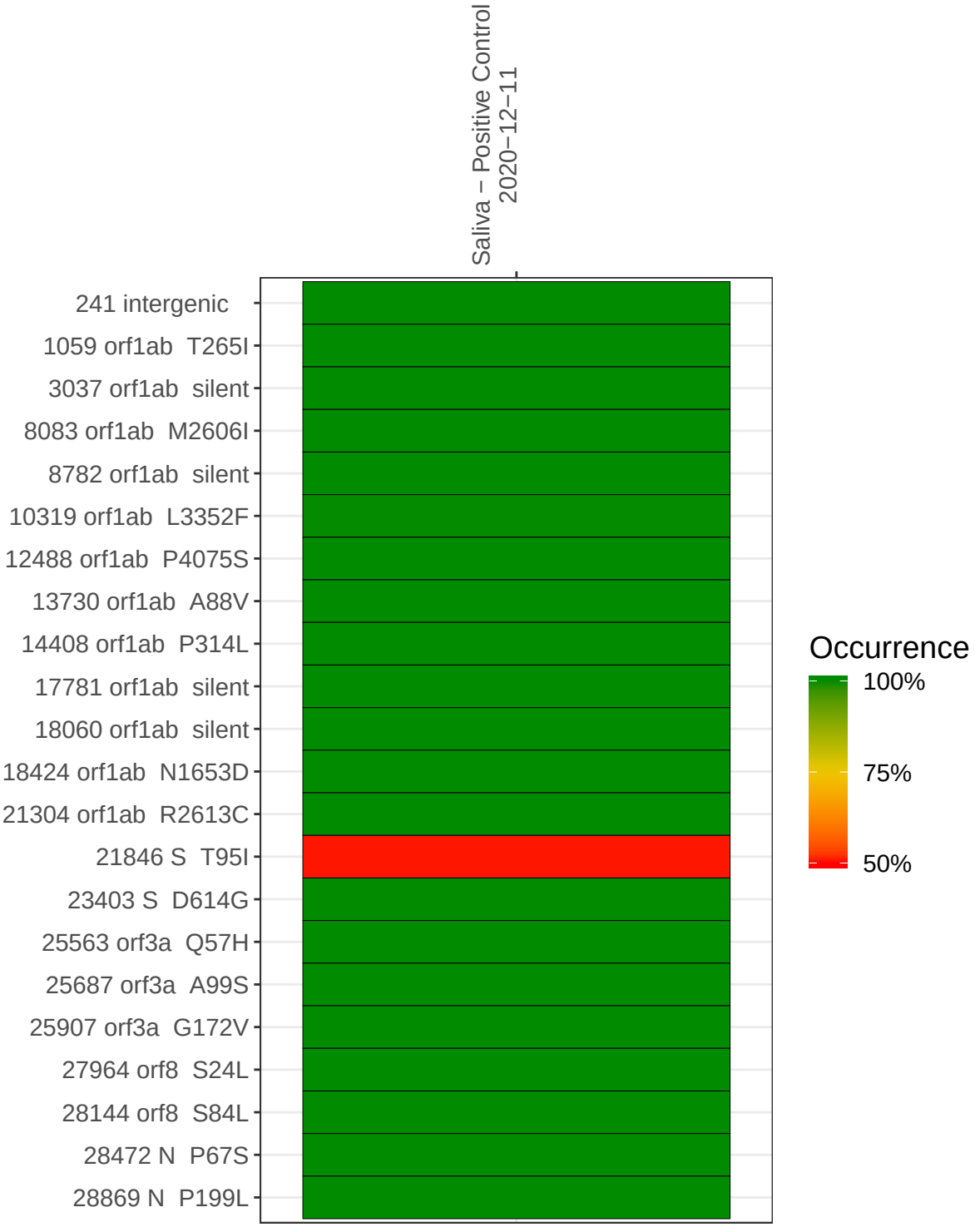
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

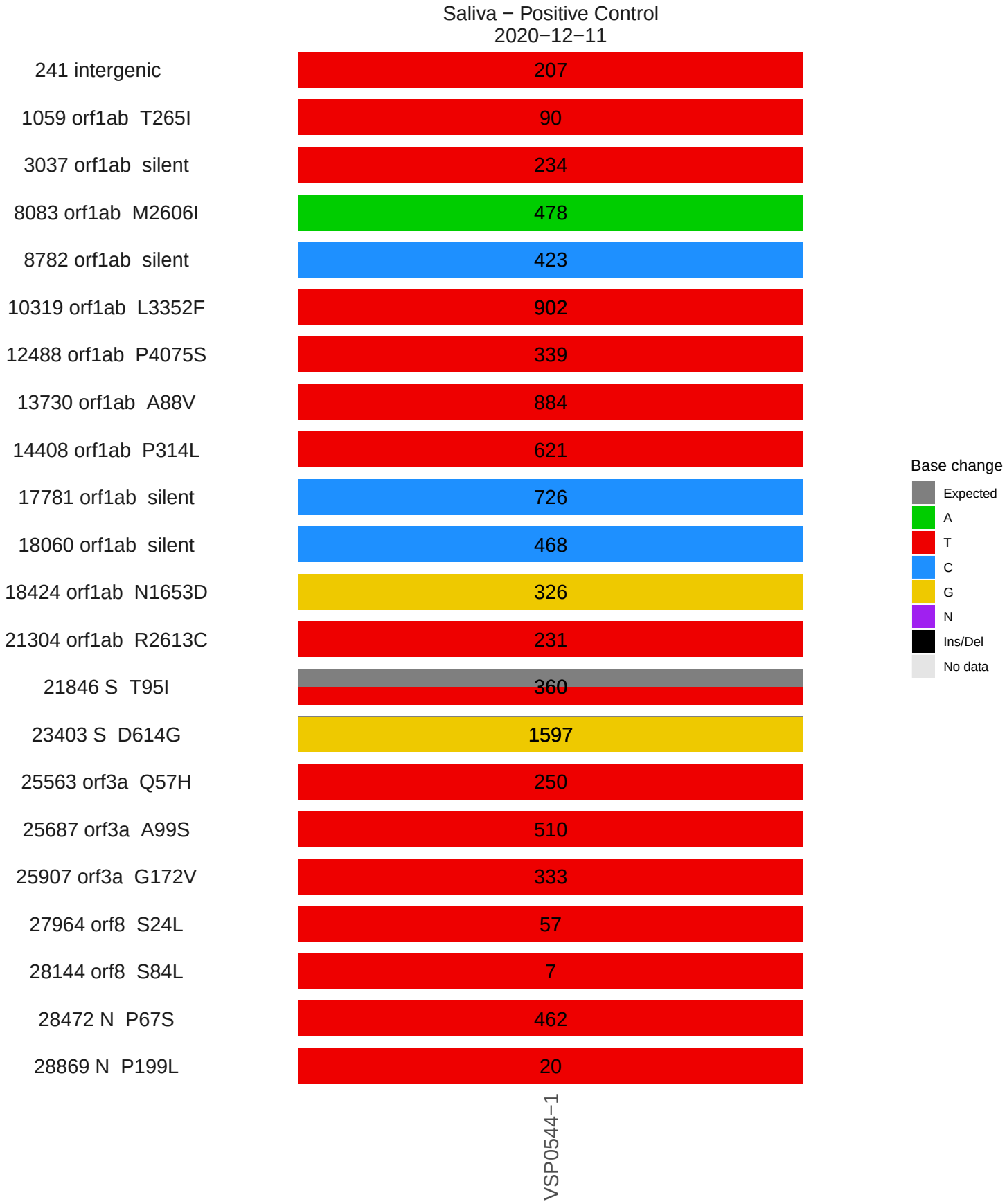
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0544-1	single experiment	NA	Saliva - Positive Control	2020-12-11	3.82	100.0%	79.1%

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.

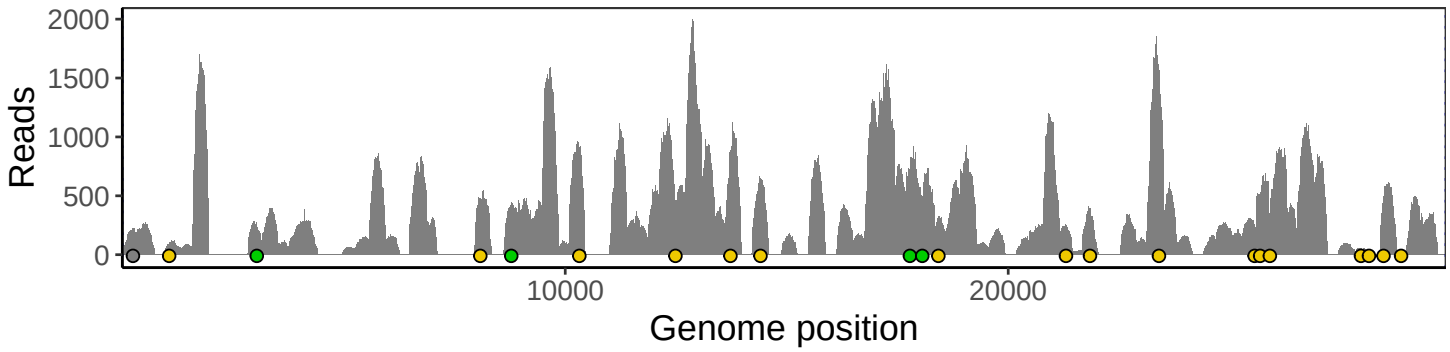



Saliva − Positive Control
2020−12−11
241 intergenic
207
1059 orf1ab T265I
90
3037 orf1ab silent
234
8083 orf1ab M2606I
478
8782 orf1ab silent
423
10319 orf1ab L3352F
902
12488 orf1ab P4075S
339
13730 orf1ab A88V
884
14408 orf1ab P314L
621
17781 orf1ab silent
726
18060 orf1ab silent
468
18424 orf1ab N1653D
326
21304 orf1ab R2613C
231
21846 S T95I
360
23403 S D614G
1597
25563 orf3a Q57H
250
25687 orf3a A99S
510
25907 orf3a G172V
333
27964 orf8 S24L
57
28144 orf8 S84L
7
28472 N P67S
462
28869 N P199L
20
VSP0544−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

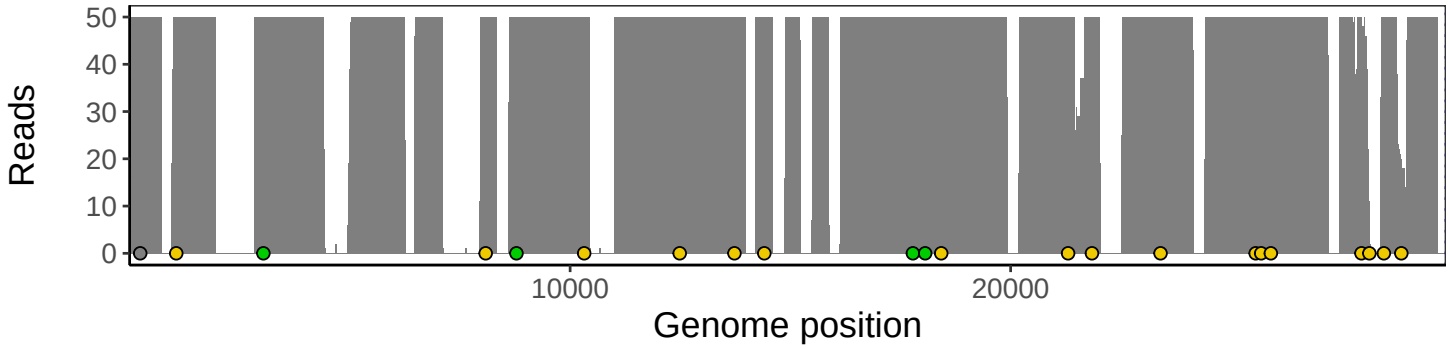
Analyses of individual experiments and composite results.

VSP0544-1 | 2020-12-11 | Saliva - Positive Control | 374 | 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.