COVID-19 subject 10-1017

2020-12-22

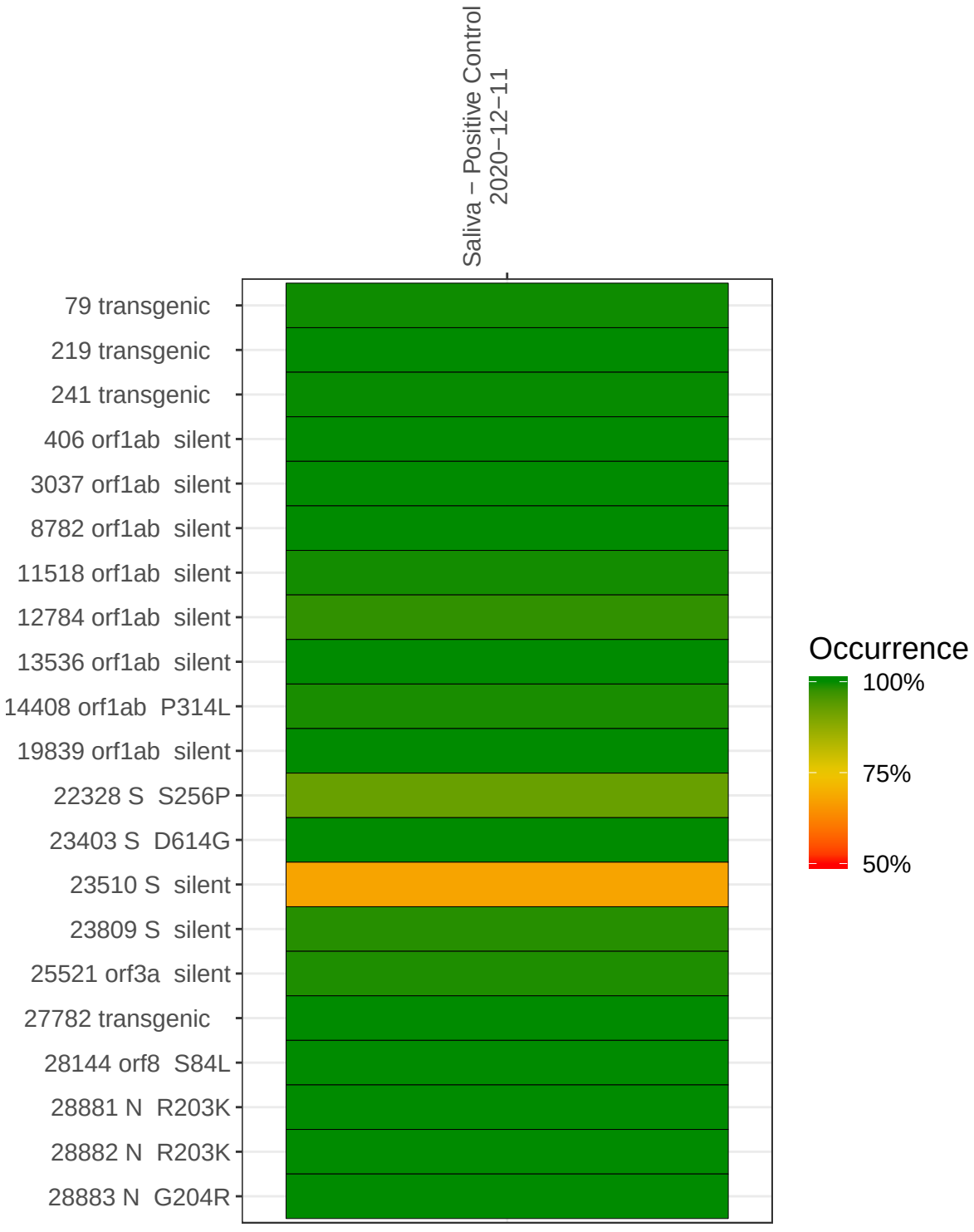
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

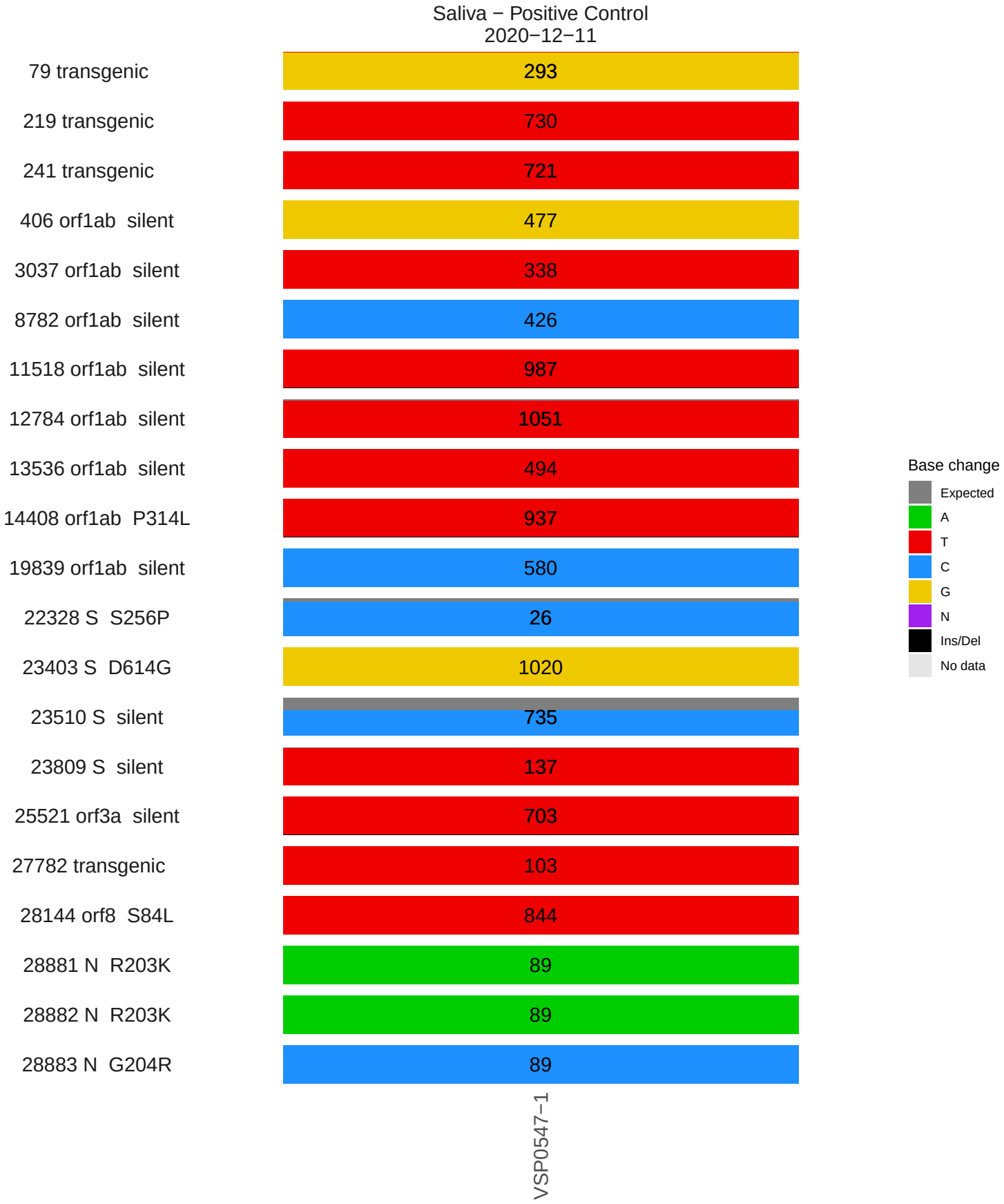
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0547-1	single experiment	NA	Saliva - Positive Control	2020-12-11	29.85	99.9%	99.6%

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.

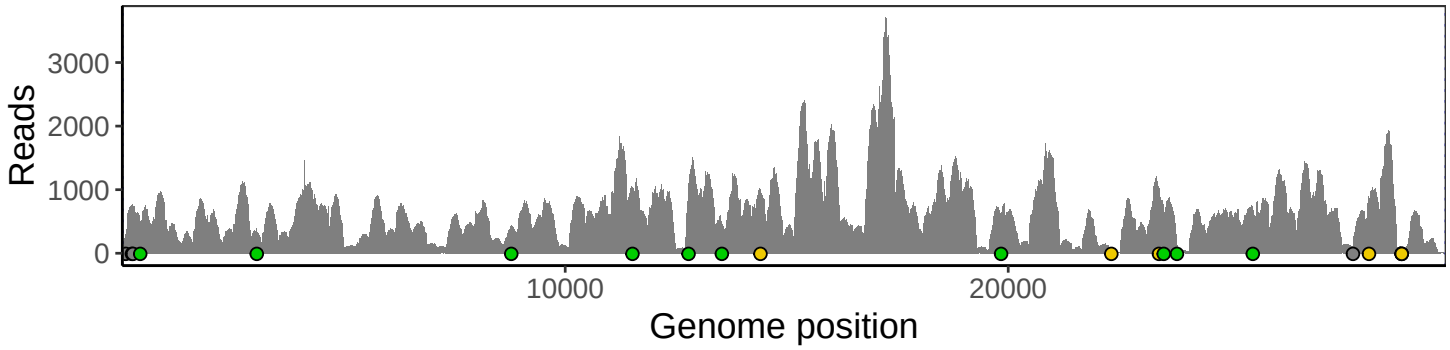



Saliva − Positive Control
2020−12−11
79 transgenic
219 transgenic
241 transgenic
406 orf1ab silent
3037 orf1ab silent
8782 orf1ab silent
11518 orf1ab silent
12784 orf1ab silent
13536 orf1ab silent
14408 orf1ab P314L
19839 orf1ab silent
22328 S S256P
23403 S D614G
23510 S silent
23809 S silent
25521 orf3a silent
27782 transgenic
28144 orf8 S84L
28881 N R203K
28882 N R203K
28883 N G204R
293
730
721
477
338
426
987
1051
494
937
580
26
1020
735
137
703
103
844
89
89
89
VSP0547−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

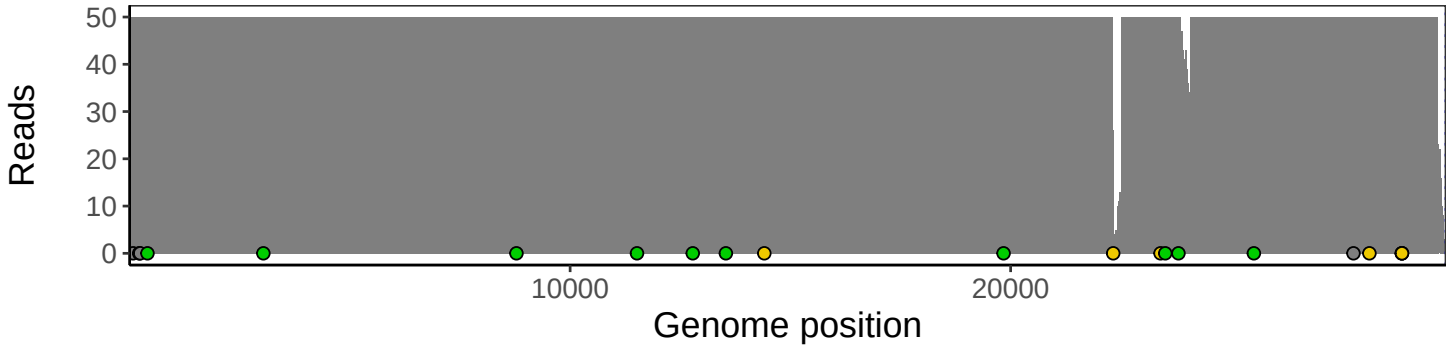
Analyses of individual experiments and composite results.

VSP0547-1 | 2020-12-11 | Saliva - Positive Control | 10-1017 | 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.