COVID-19 subject 161

2020-12-22

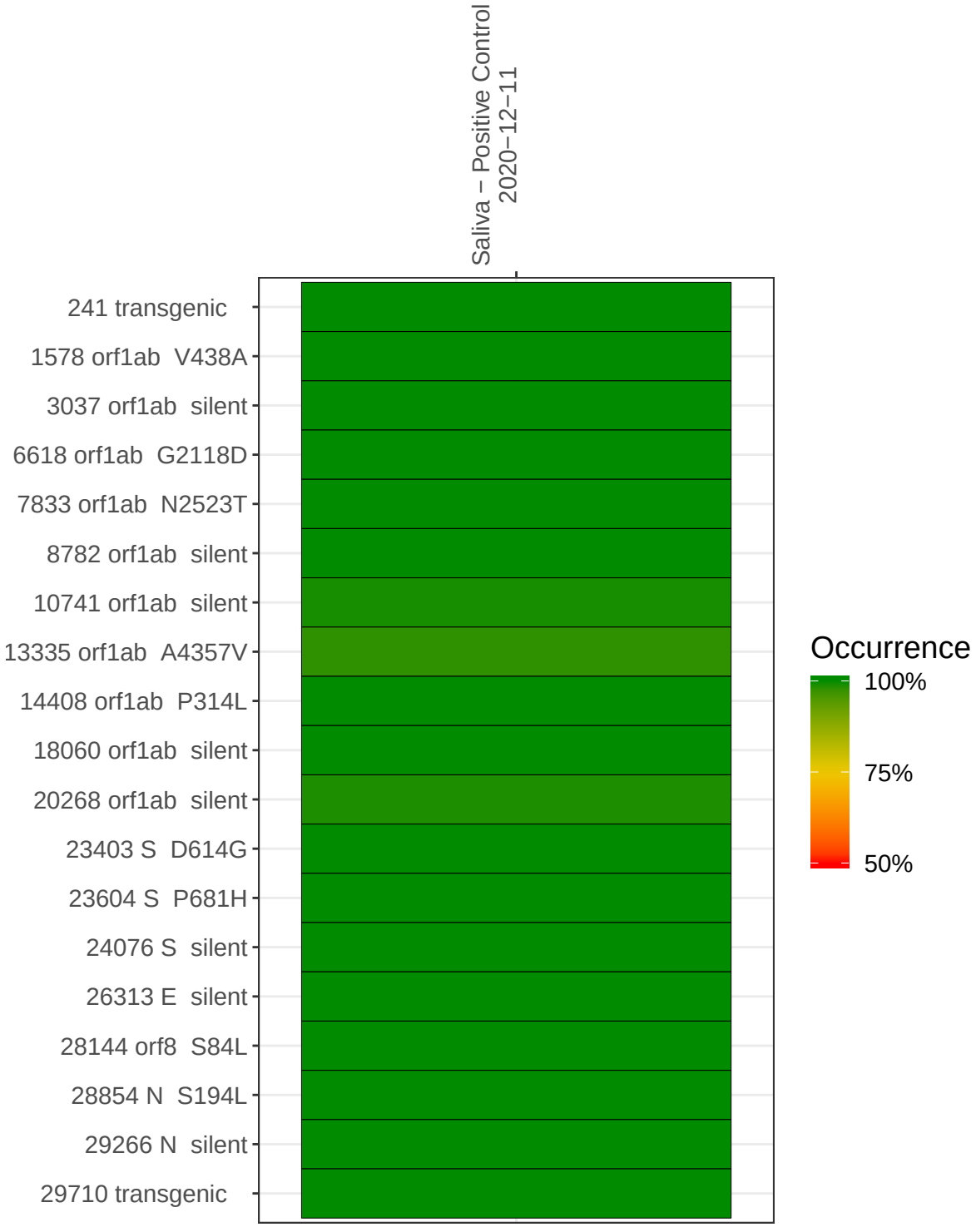
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

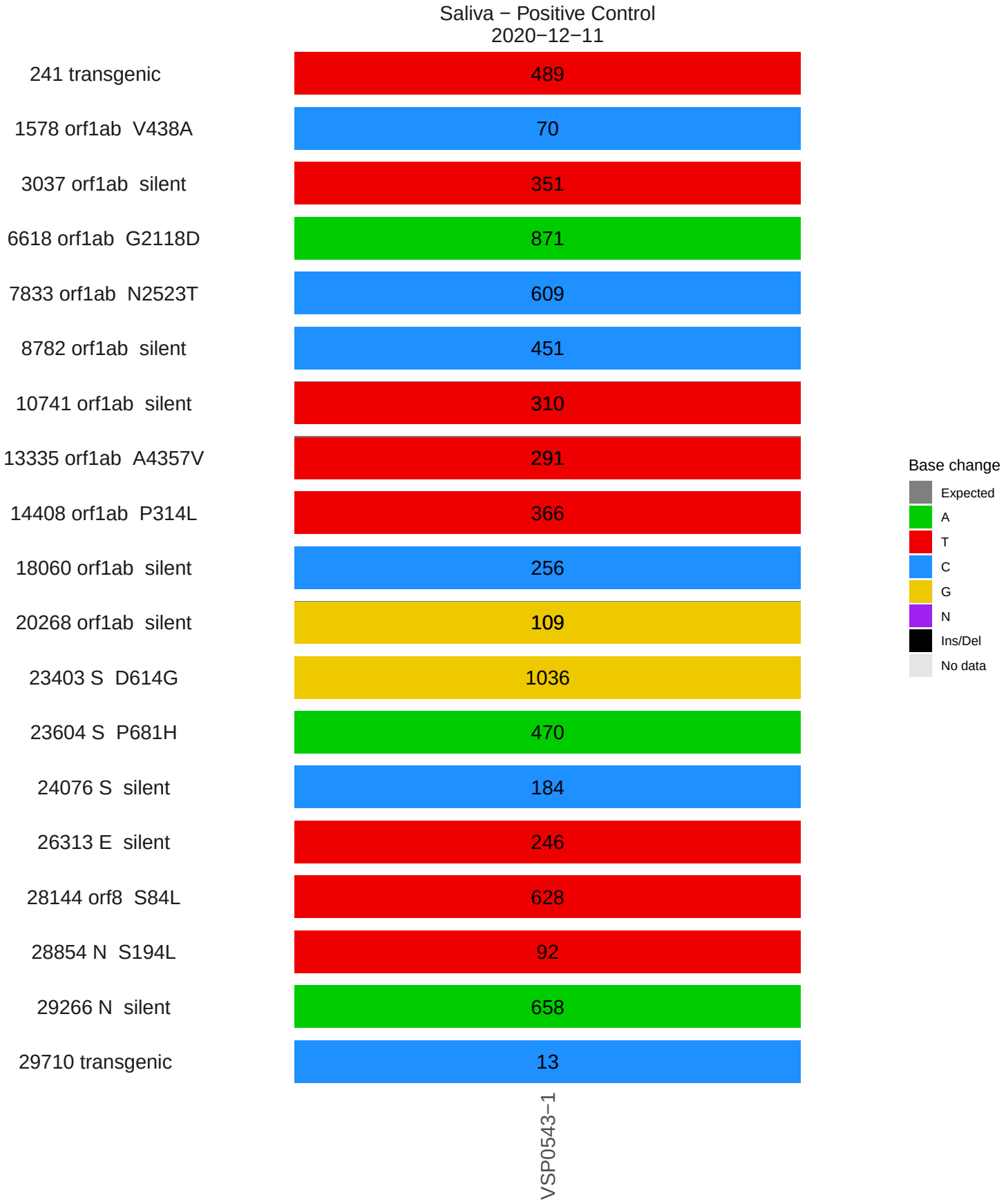
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0543-1	single experiment	NA	Saliva - Positive Control	2020-12-11	29.81	99.8%	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.

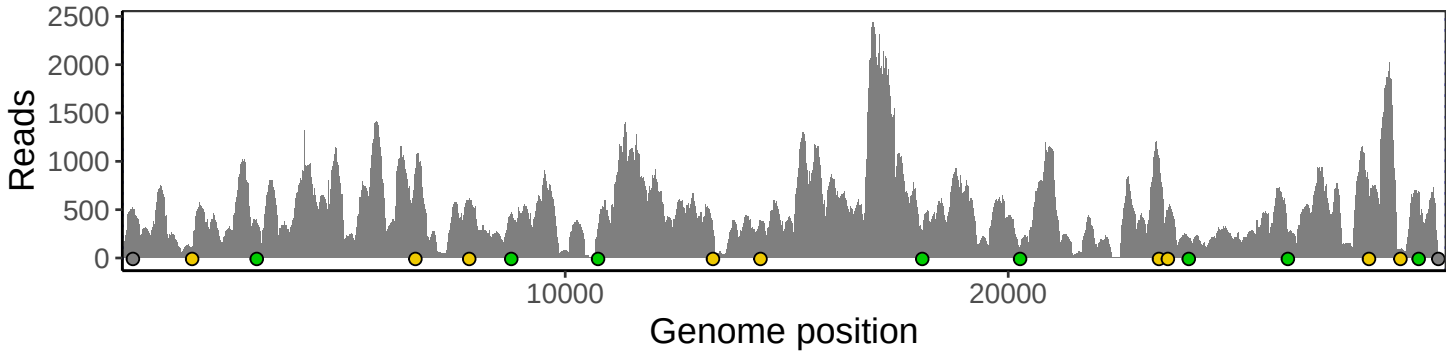



Saliva − Positive Control
2020−12−11
241 transgenic
489
1578 orf1ab V438A
70
3037 orf1ab silent
351
6618 orf1ab G2118D
871
7833 orf1ab N2523T
609
8782 orf1ab silent
451
10741 orf1ab silent
310
13335 orf1ab A4357V
291
14408 orf1ab P314L
366
18060 orf1ab silent
256
20268 orf1ab silent
109
23403 S D614G
1036
23604 S P681H
470
24076 S silent
184
26313 E silent
246
28144 orf8 S84L
628
28854 N S194L
92
29266 N silent
658
29710 transgenic
13
VSP0543−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

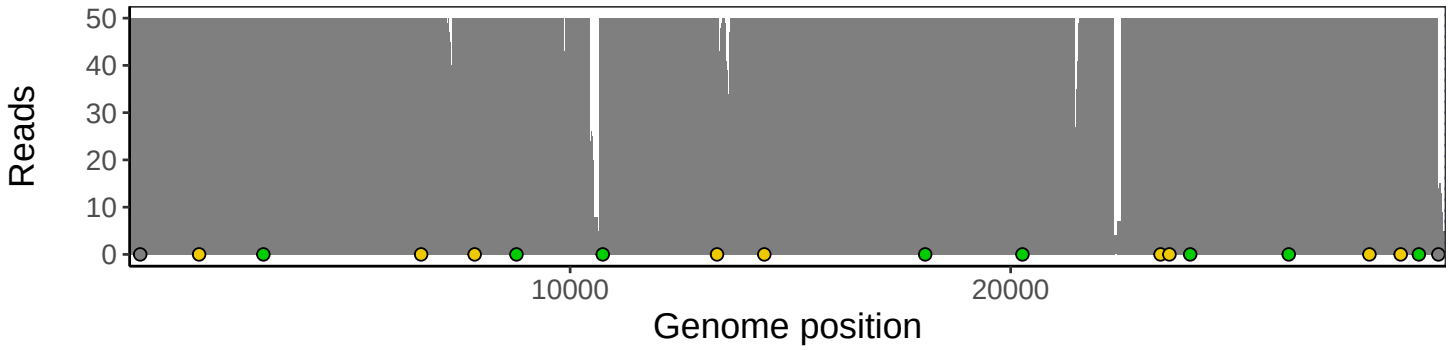
Analyses of individual experiments and composite results.

VSP0543-1 | 2020-12-11 | Saliva - Positive Control | 161 | 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.