COVID-19 subject sdrop1_molpath

2021-01-14

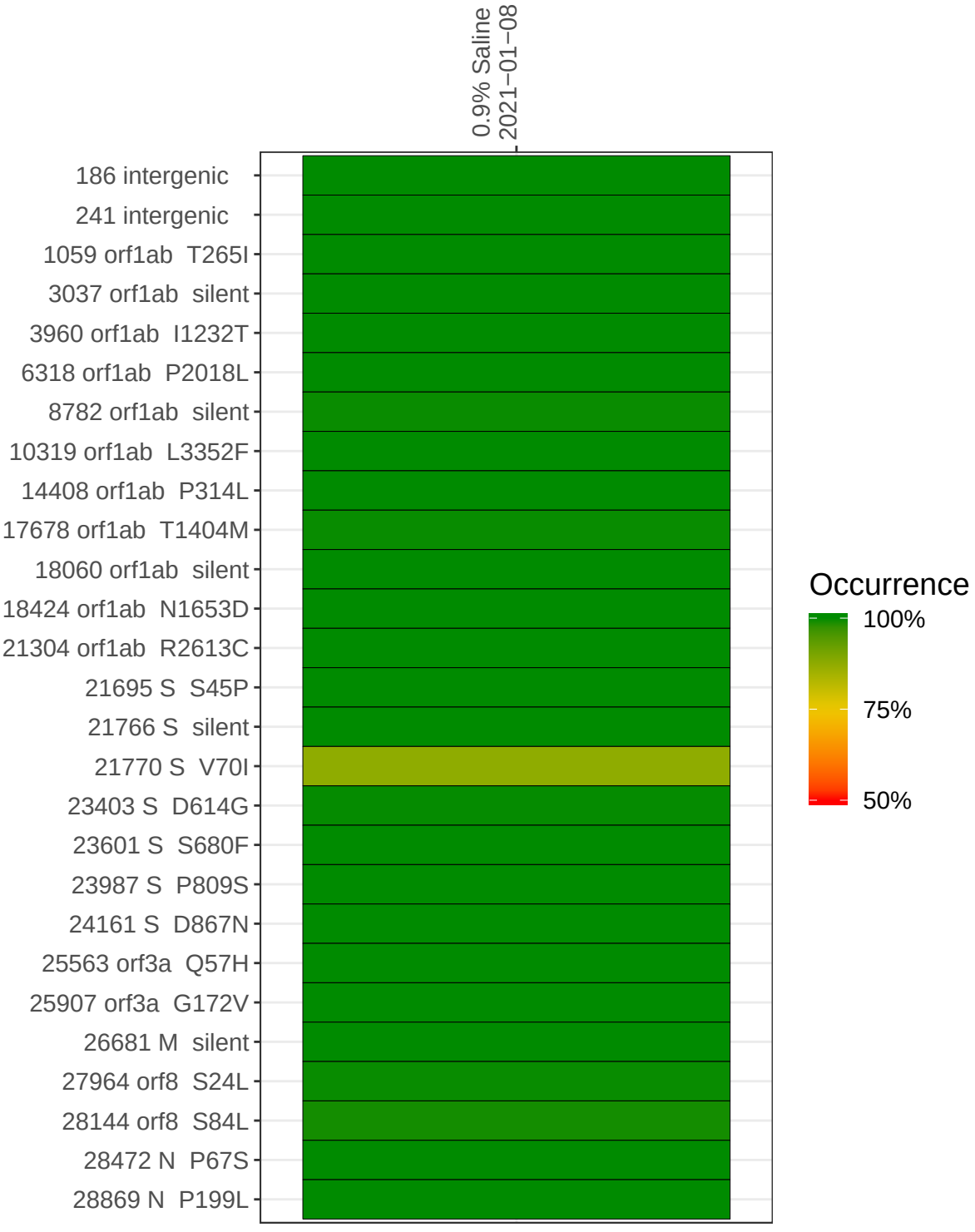
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

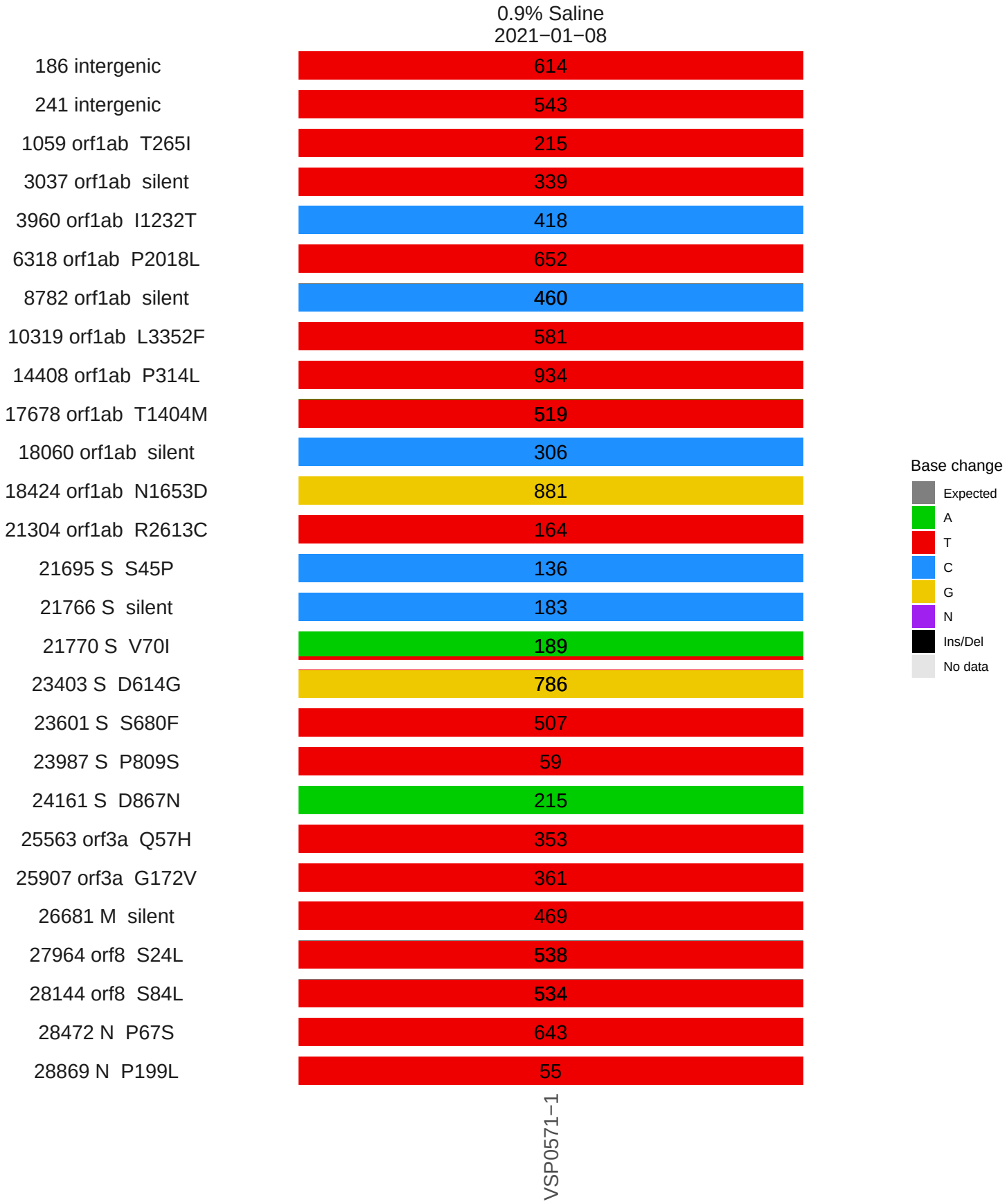
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0571-1	single experiment	NA	0.9% Saline	2021-01-08	29.82	99.9%	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.

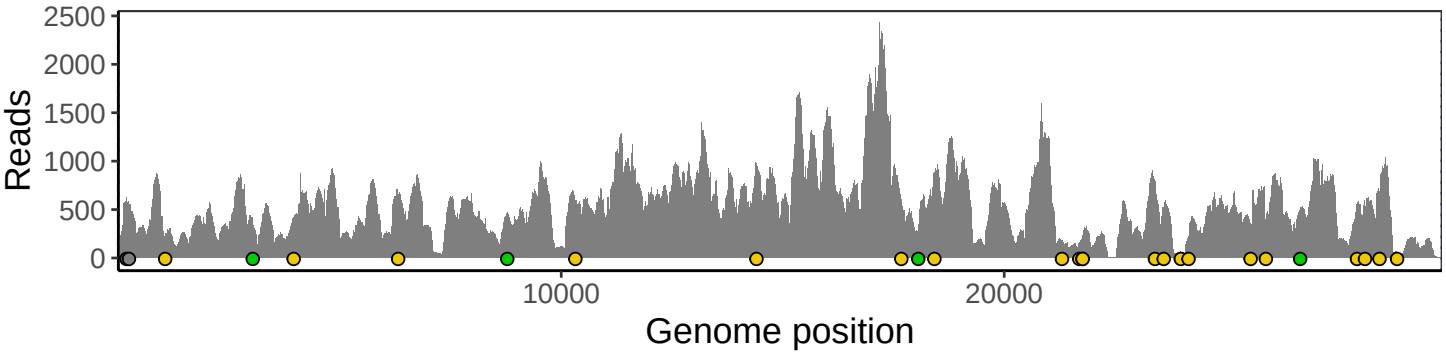



0.9% Saline
2021−01−08
186 intergenic
241 intergenic
1059 orf1ab T265I
3037 orf1ab silent
3960 orf1ab I1232T
6318 orf1ab P2018L
8782 orf1ab silent
10319 orf1ab L3352F
14408 orf1ab P314L
17678 orf1ab T1404M
18060 orf1ab silent
18424 orf1ab N1653D
21304 orf1ab R2613C
21695 S S45P
21766 S silent
21770 S V70I
23403 S D614G
23601 S S680F
23987 S P809S
24161 S D867N
25563 orf3a Q57H
25907 orf3a G172V
26681 M silent
27964 orf8 S24L
28144 orf8 S84L
28472 N P67S
28869 N P199L
614
543
215
339
418
652
460
581
934
519
306
881
164
136
183
189
786
507
59
215
353
361
469
538
534
643
55
VSP0571−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

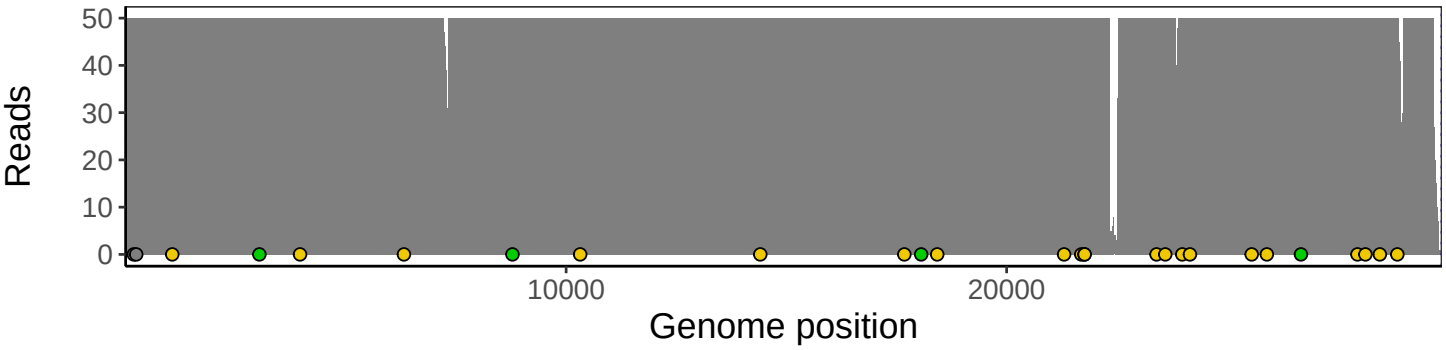
Analyses of individual experiments and composite results.

VSP0571-1 | 2021-01-08 | 0.9% Saline | sdrop1_molpath | 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.