COVID-19 subject DOH1

2021-01-11

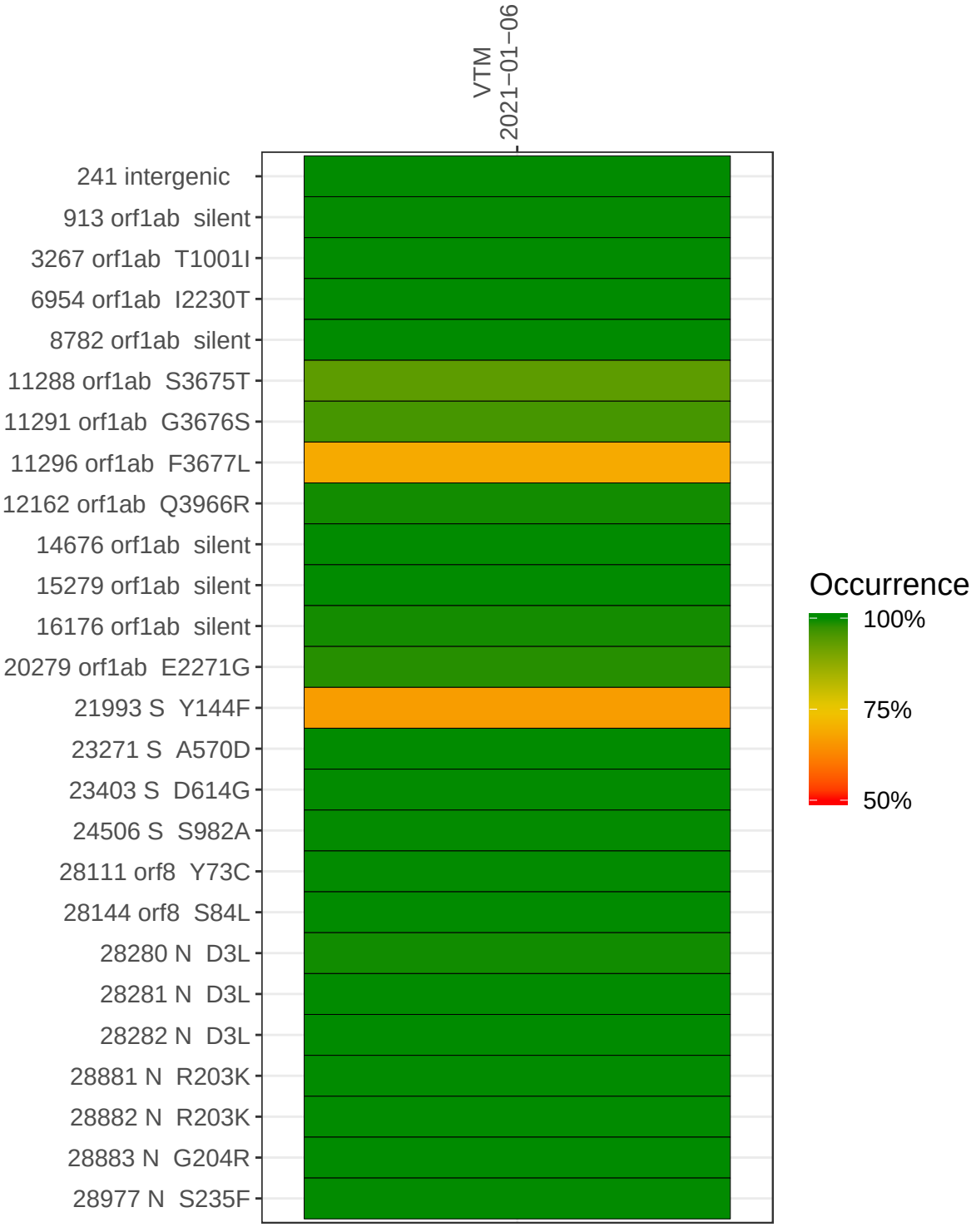
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

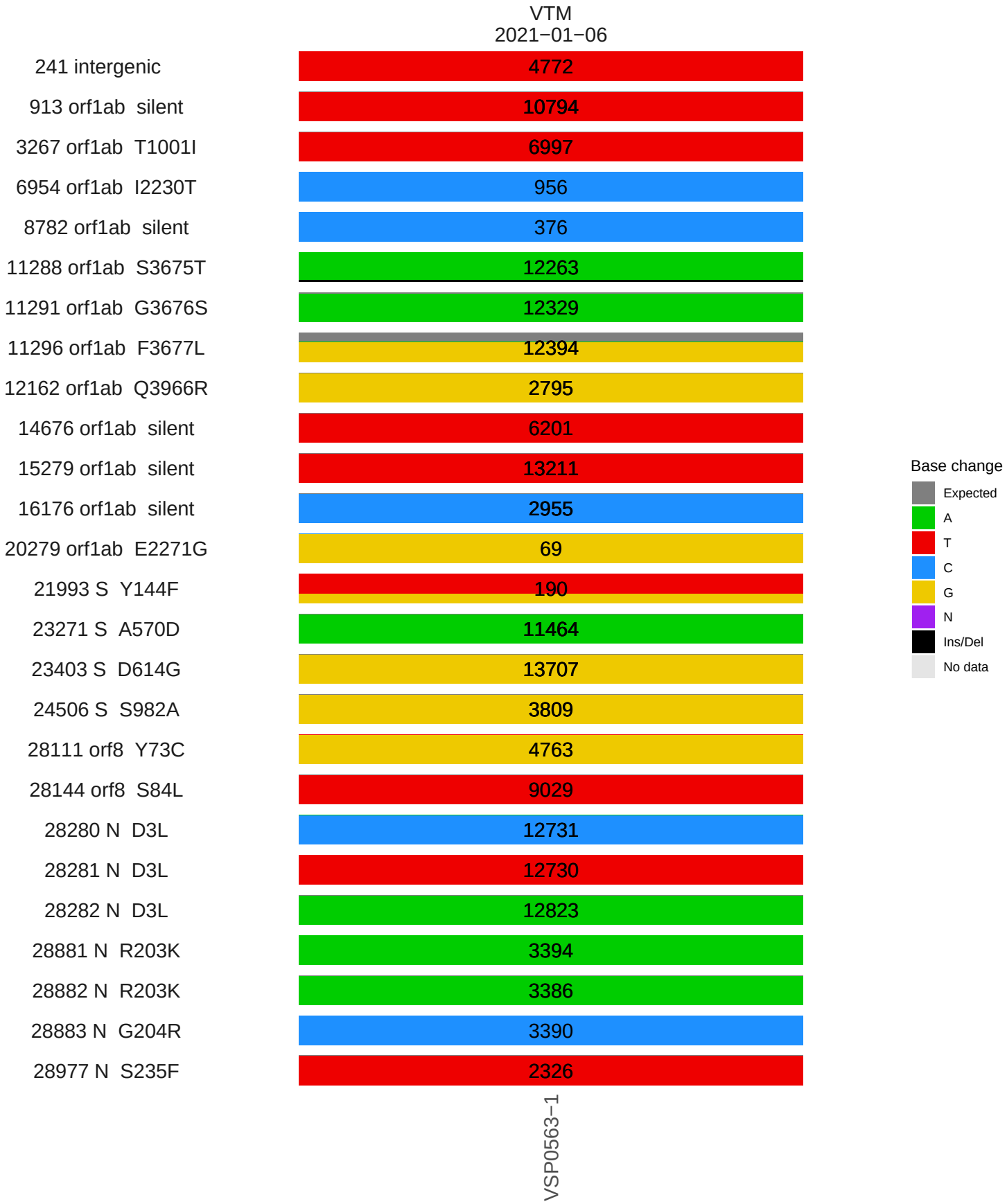
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0563-1	single experiment	NA	VTM	2021-01-06	1.17	65.6%	62.3%

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.

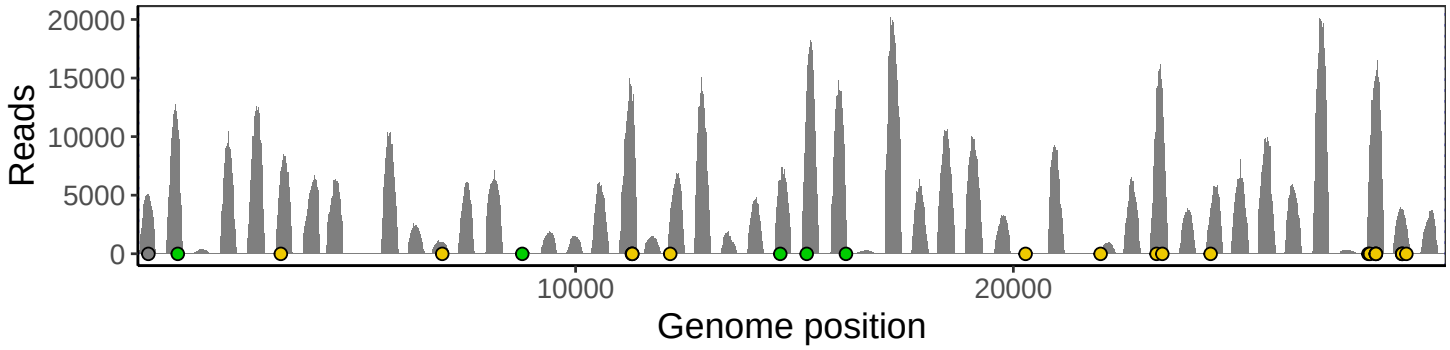



VTM
2021−01−06
241 intergenic
913 orf1ab silent
3267 orf1ab T1001I
6954 orf1ab I2230T
8782 orf1ab silent
11288 orf1ab S3675T
11291 orf1ab G3676S
11296 orf1ab F3677L
12162 orf1ab Q3966R
14676 orf1ab silent
15279 orf1ab silent
16176 orf1ab silent
20279 orf1ab E2271G
21993 S Y144F
23271 S A570D
23403 S D614G
24506 S S982A
28111 orf8 Y73C
28144 orf8 S84L
28280 N D3L
28281 N D3L
28282 N D3L
28881 N R203K
28882 N R203K
28883 N G204R
28977 N S235F
4772
10794
6997
956
376
12263
12329
12394
2795
6201
13211
2955
69
190
11464
13707
3809
4763
9029
12731
12730
12823
3394
3386
3390
2326
VSP0563−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

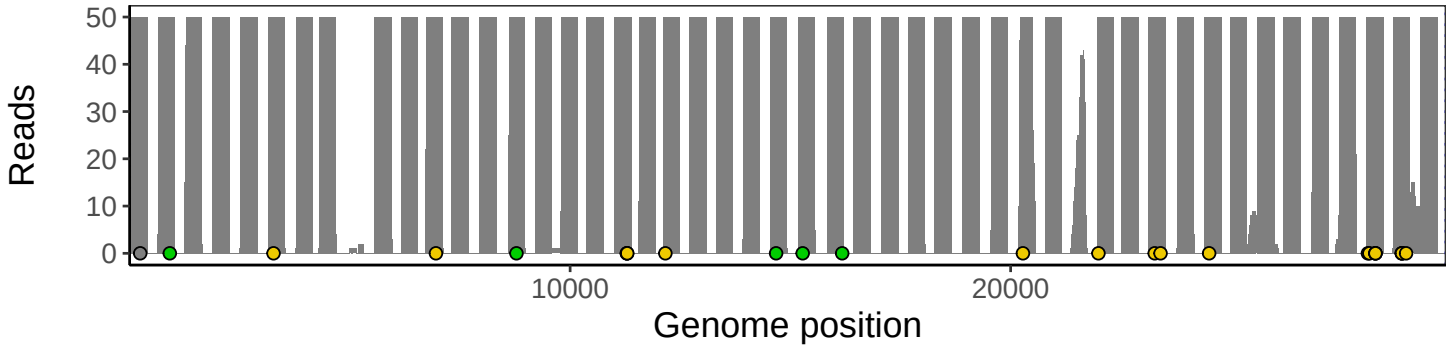
Analyses of individual experiments and composite results.

VSP0563-1 | 2021-01-06 | VTM | DOH1 | 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.