COVID-19 subject 263

2020-11-30

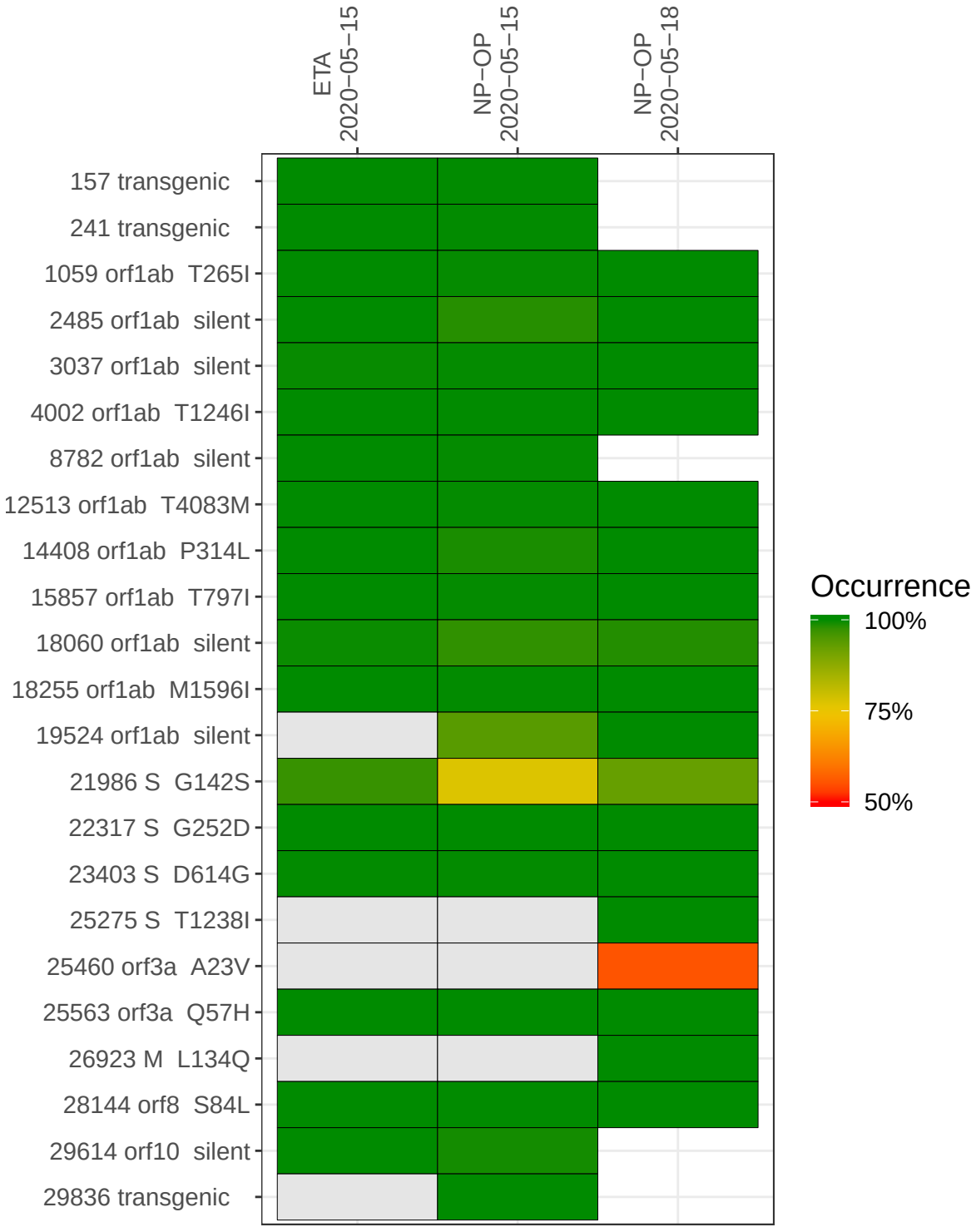
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

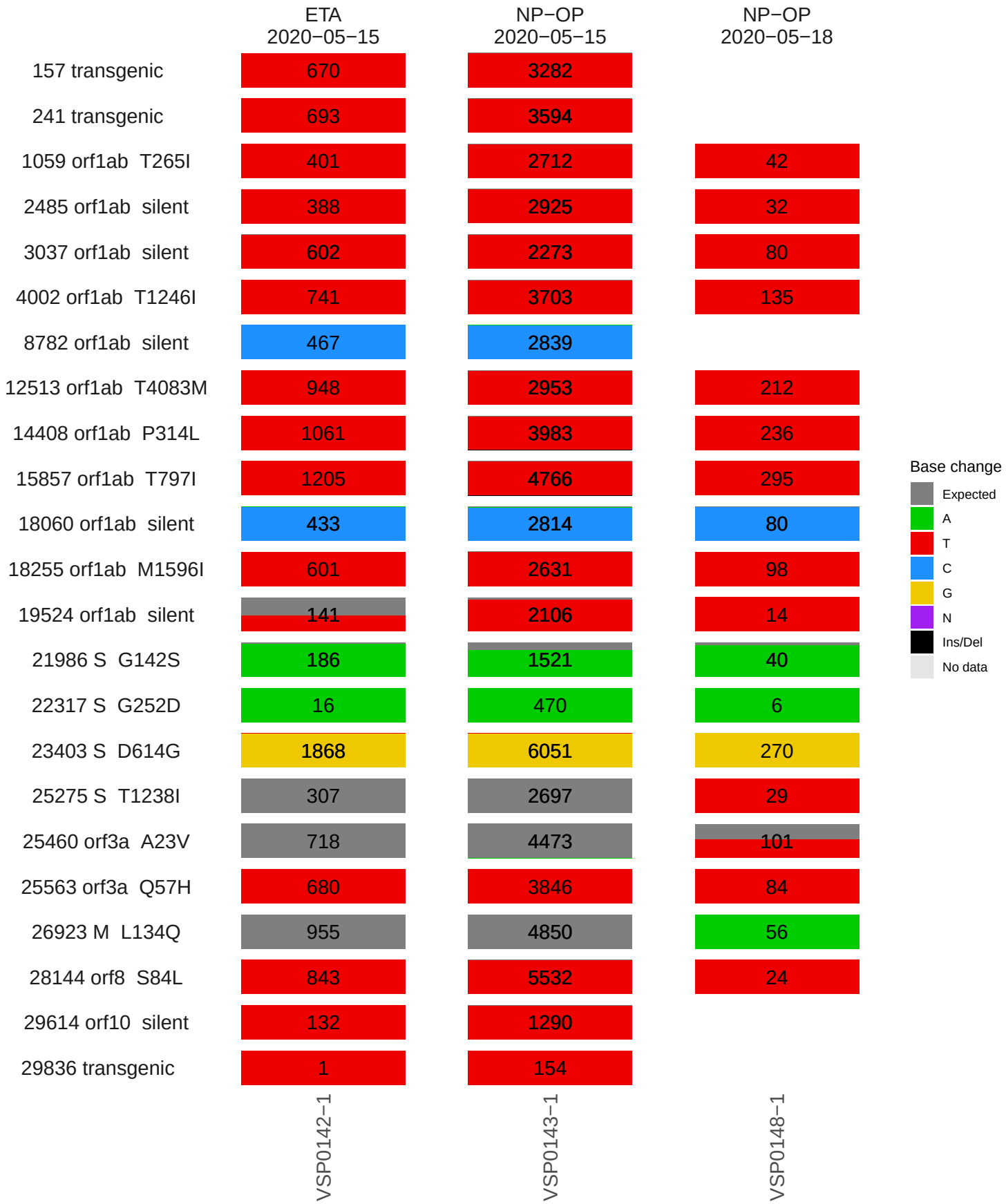
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0142-1	single experiment	NA	ETA	2020-05-15	29.81	99.7%	99.6%
VSP0143-1	single experiment	172000	NP-OP	2020-05-15	29.85	99.8%	99.8%
VSP0148-1	single experiment	NA	NP-OP	2020-05-18	8.68	90.7%	89.5%

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.

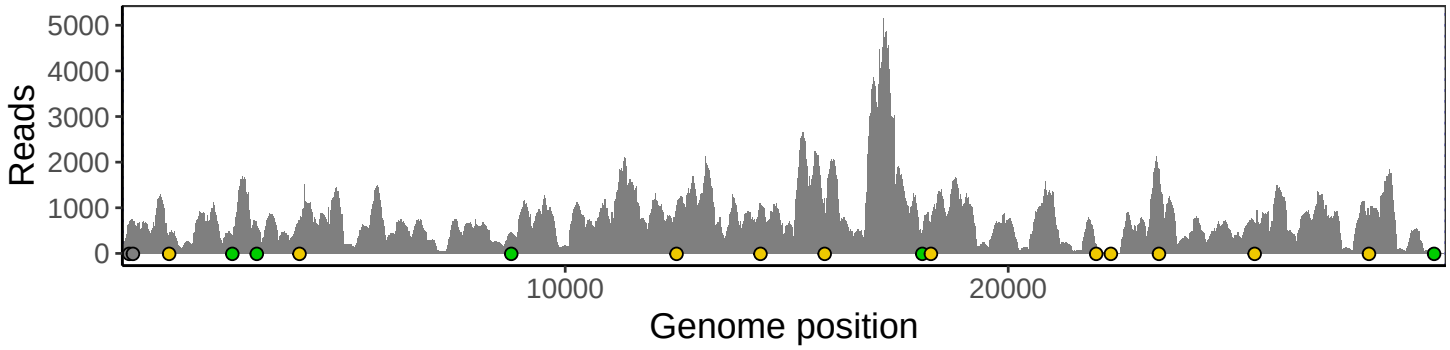



ETA
2020−05−15
NP−OP
2020−05−15
NP−OP
2020−05−18
157 transgenic
241 transgenic
1059 orf1ab T265I
2485 orf1ab silent
3037 orf1ab silent
4002 orf1ab T1246I
8782 orf1ab silent
12513 orf1ab T4083M
14408 orf1ab P314L
15857 orf1ab T797I
18060 orf1ab silent
18255 orf1ab M1596I
19524 orf1ab silent
21986 S G142S
22317 S G252D
23403 S D614G
25275 S T1238I
25460 orf3a A23V
25563 orf3a Q57H
26923 M L134Q
28144 orf8 S84L
29614 orf10 silent
29836 transgenic
670
693
401
388
602
741
467
948
1061
1205
433
601
141
186
16
1868
307
718
680
955
843
132
1
3282
3594
2712
2925
2273
3703
2839
2953
3983
4766
2814
2631
2106
1521
470
6051
2697
4473
3846
4850
5532
1290
154
42
32
80
135
212
236
295
80
98
14
40
6
270
29
101
84
56
24
Base change
Expected
A
T
C
G
N
Ins/Del
No data
VSP0142−1
VSP0143−1
VSP0148−1

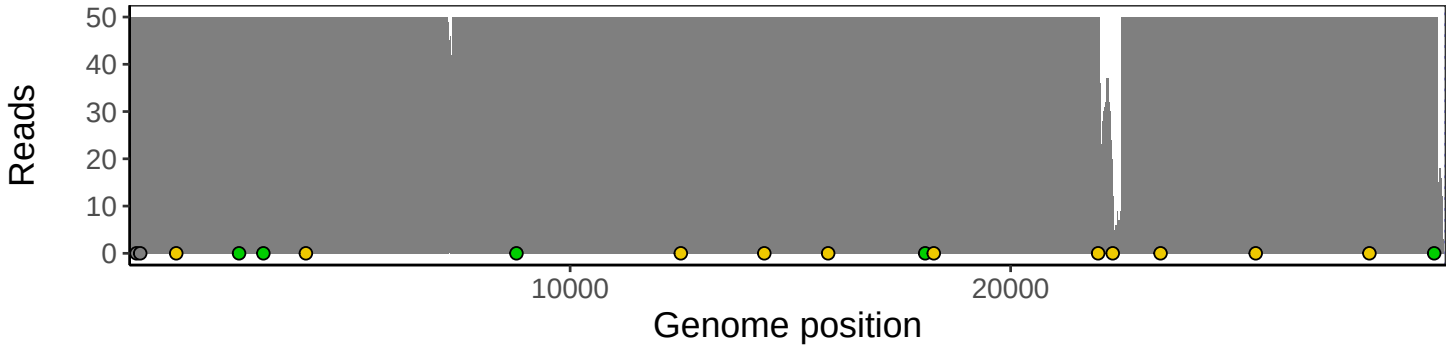
Analyses of individual experiments and composite results.

VSP0142-1 | 2020-05-15 | ETA | 263e-q1 | genomes | single experiment

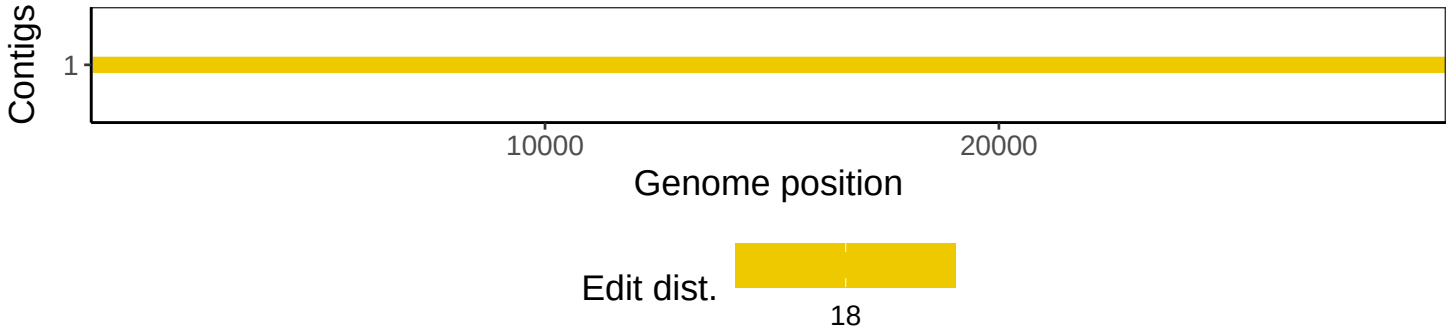
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



Excerpt from plot above focusing on reads coverage from 0 to 50 NT.

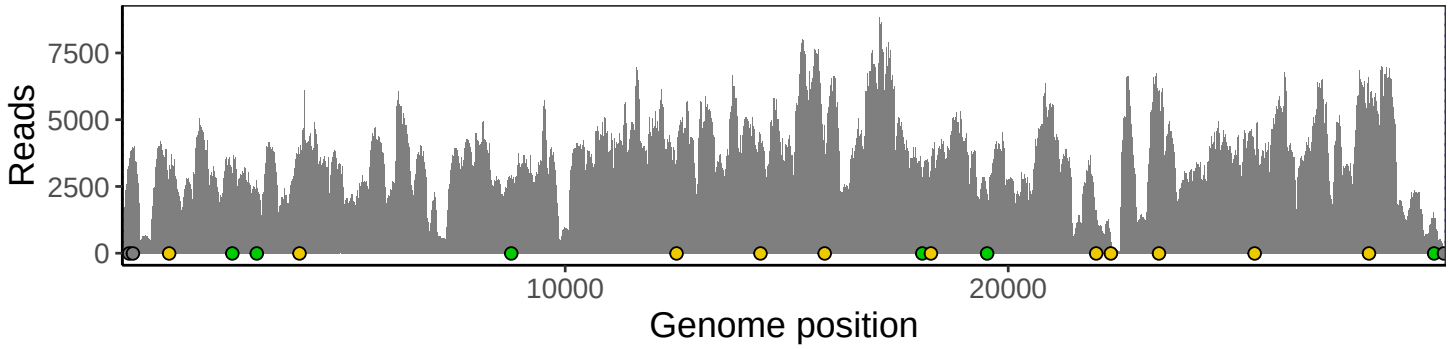


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

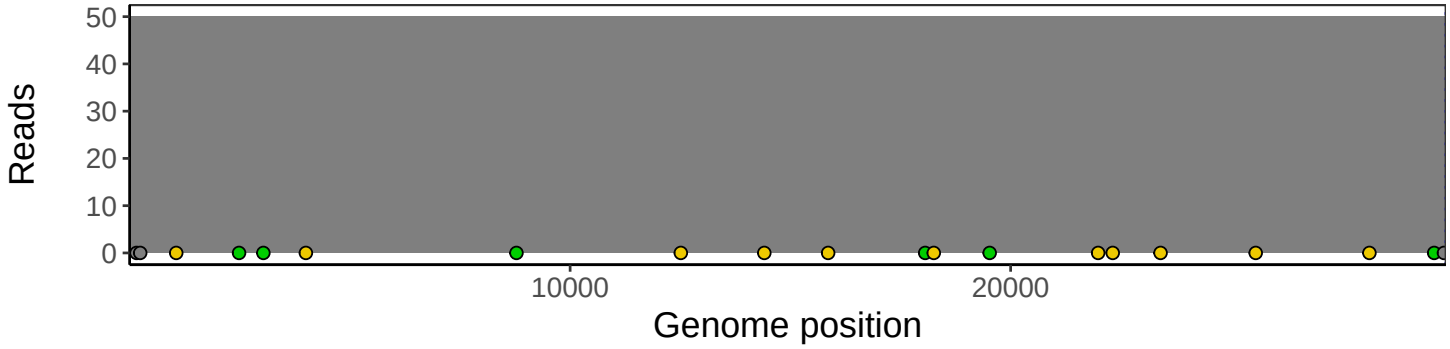


VSP0143-1 | 2020-05-15 | NP-OP | 263no-q1 | 172000 genomes | single experiment

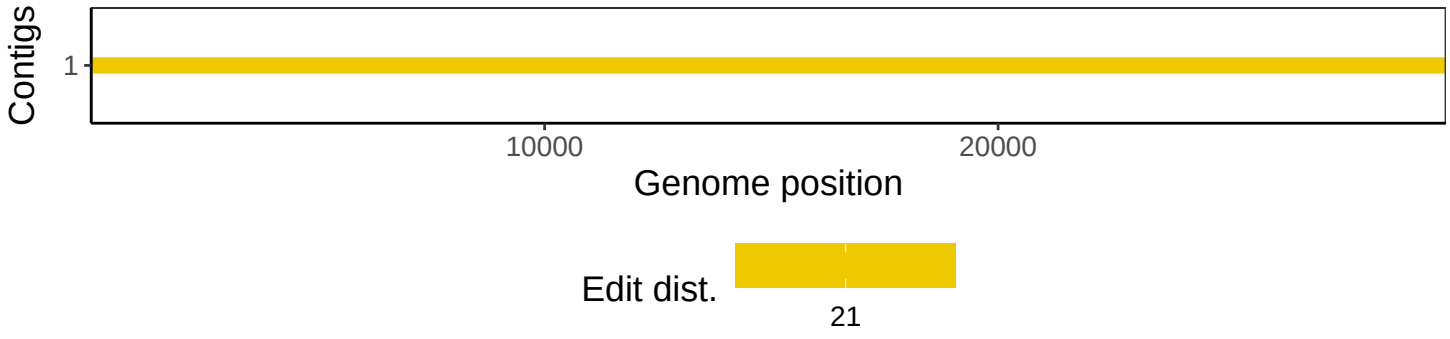
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



Excerpt from plot above focusing on reads coverage from 0 to 50 NT.

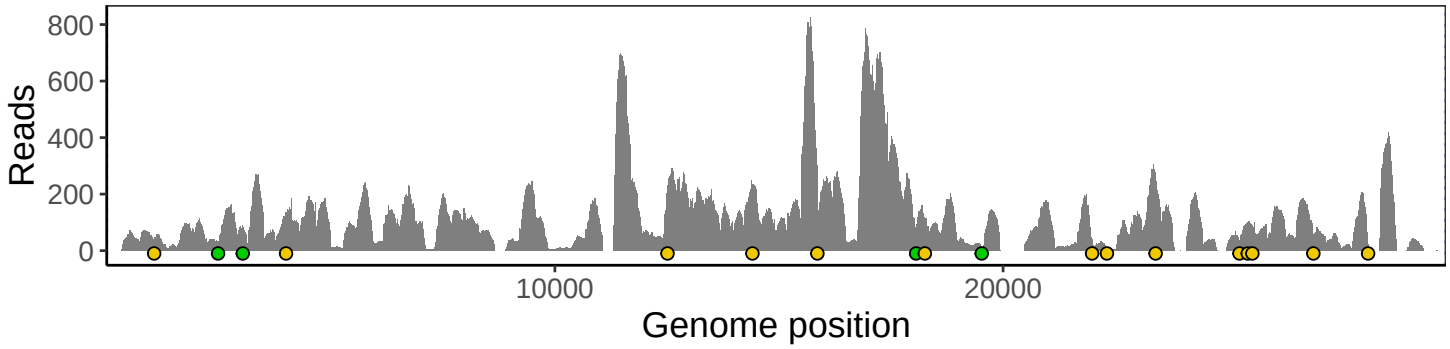


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

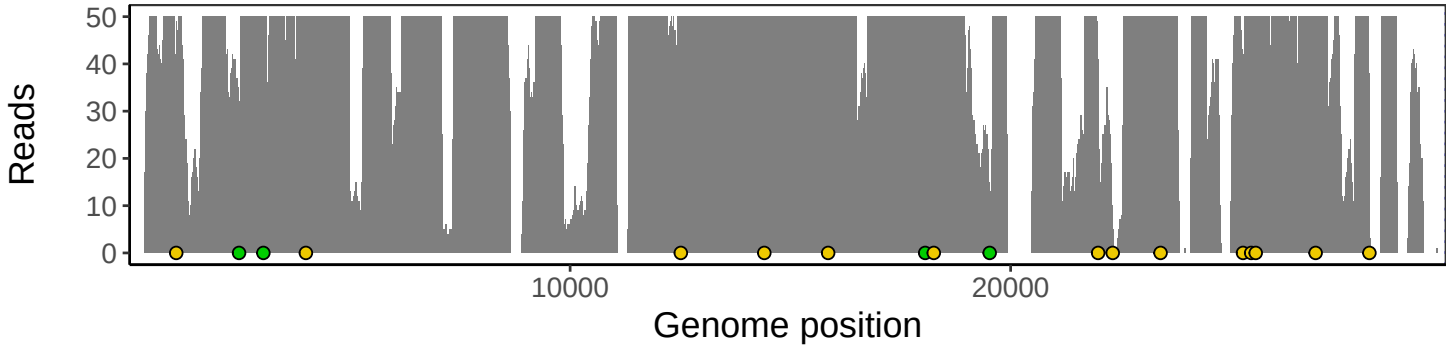


VSP0148-1 | 2020-05-18 | NP-OP | 263no-q2 | genomes | single experiment

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.