COVID-19 subject 263

2021-01-08

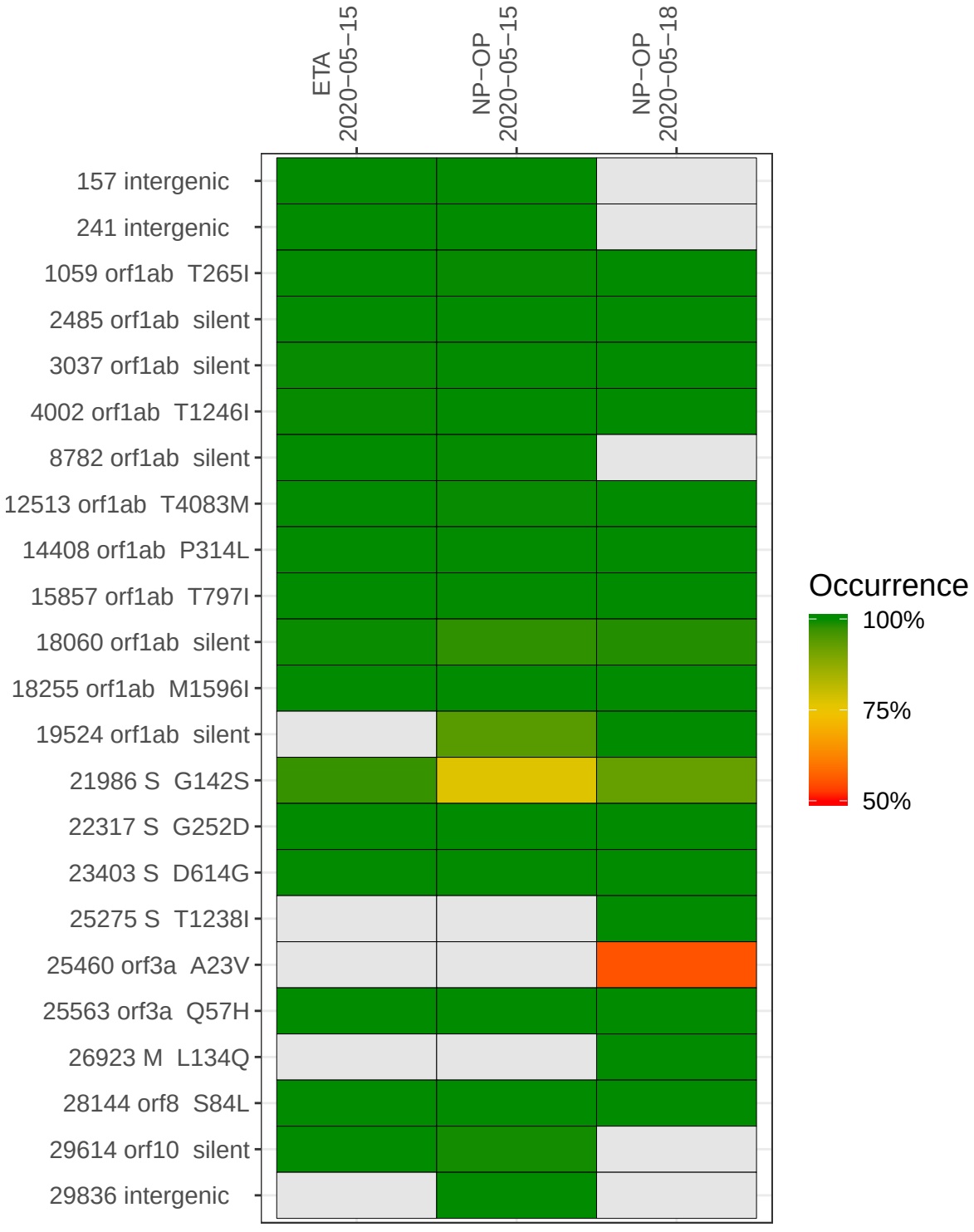
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

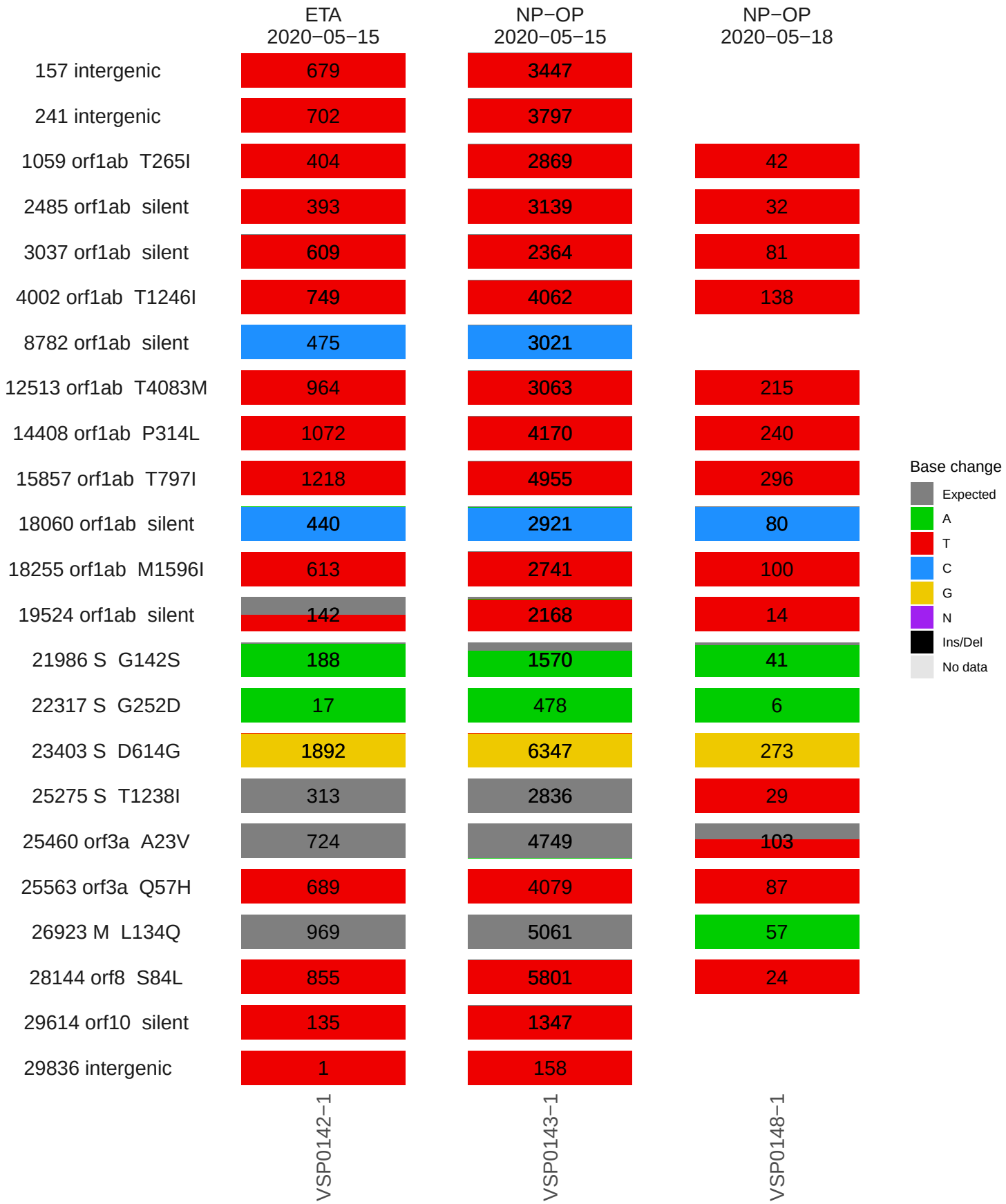
Table 1. Sample summary.

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	100.0%	99.6%
VSP0143-1	single experiment	172000	NP-OP	2020-05-15	29.87	100.0%	99.8%
VSP0148-1	single experiment	NA	NP-OP	2020-05-18	8.68	100.0%	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.

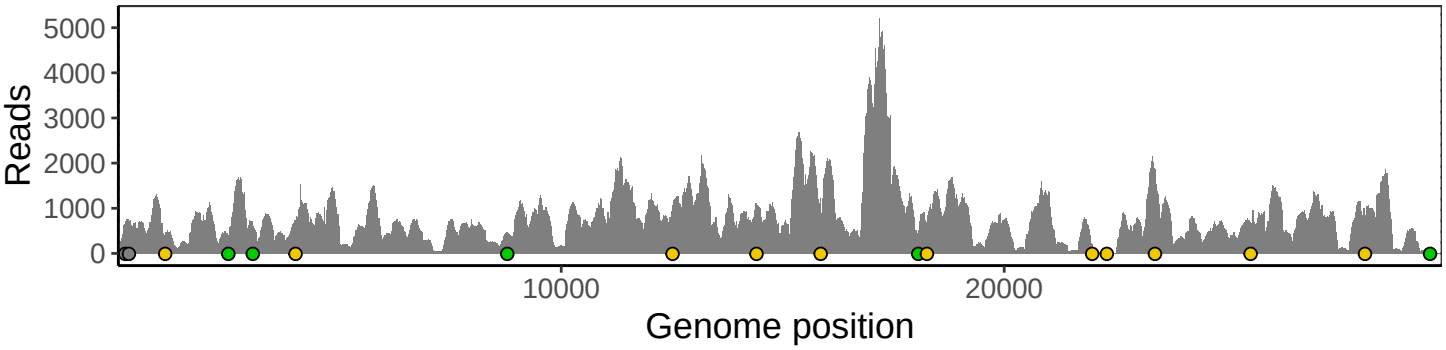



ETA
2020−05−15
NP−OP
2020−05−15
NP−OP
2020−05−18
157 intergenic
679
3447
241 intergenic
702
3797
1059 orf1ab T265I
404
2869
42
2485 orf1ab silent
393
3139
32
3037 orf1ab silent
609
2364
81
4002 orf1ab T1246I
749
4062
138
8782 orf1ab silent
475
3021
12513 orf1ab T4083M
964
3063
215
14408 orf1ab P314L
1072
4170
240
15857 orf1ab T797I
1218
4955
296
18060 orf1ab silent
440
2921
80
18255 orf1ab M1596I
613
2741
100
19524 orf1ab silent
142
2168
14
21986 S G142S
188
1570
41
22317 S G252D
17
478
6
23403 S D614G
1892
6347
273
25275 S T1238I
313
2836
29
25460 orf3a A23V
724
4749
103
25563 orf3a Q57H
689
4079
87
26923 M L134Q
969
5061
57
28144 orf8 S84L
855
5801
24
29614 orf10 silent
135
1347
29836 intergenic
1
158
VSP0142−1
VSP0143−1
VSP0148−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

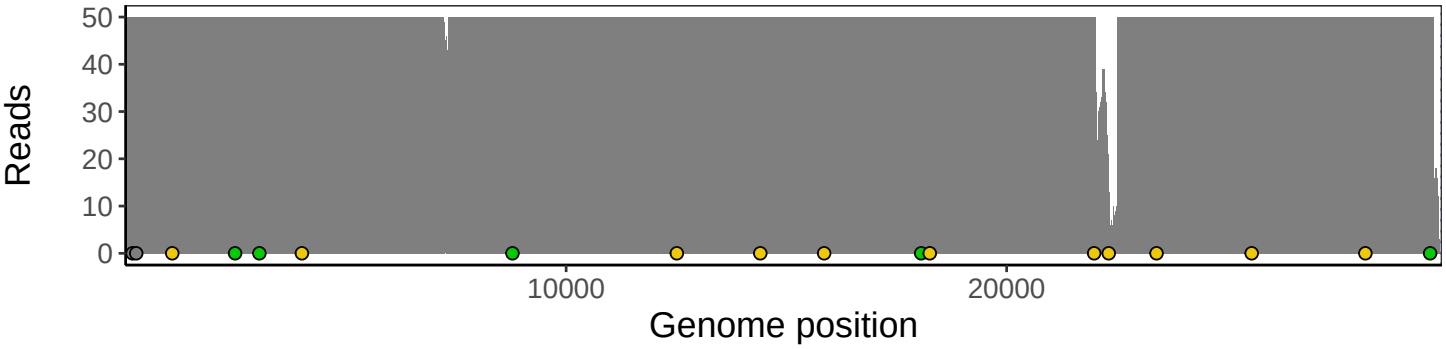
Analyses of individual experiments and composite results.

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

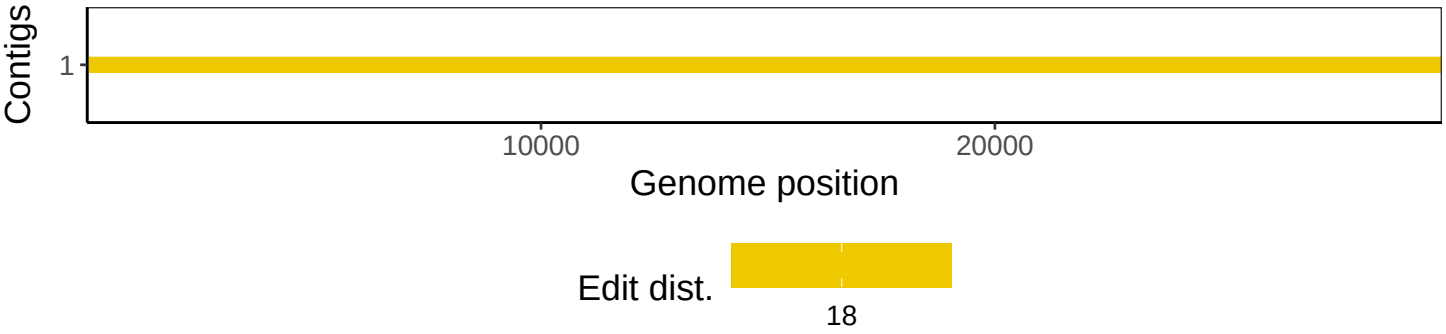
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

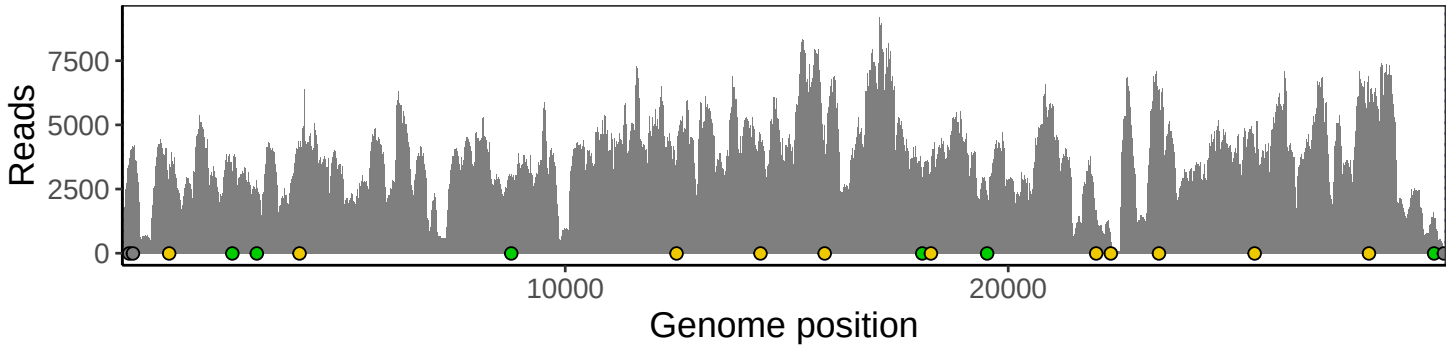


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

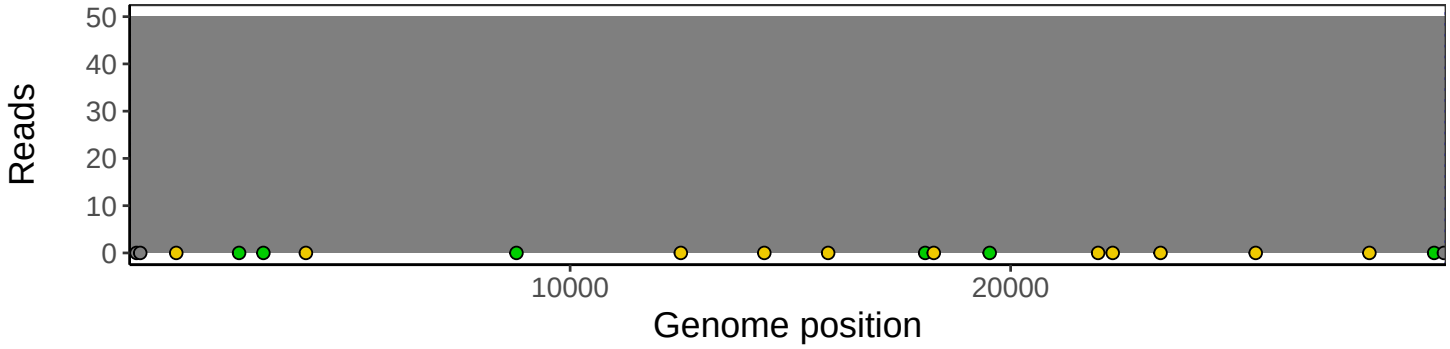


VSP0143-1 | 2020-05-15 | NP-OP | 263no-q1 | 172000 genomes | single experiment

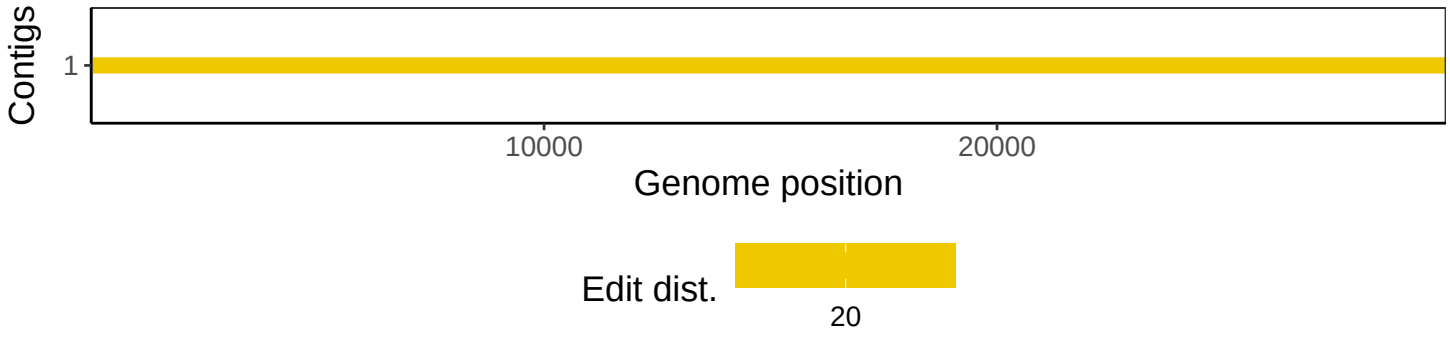
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

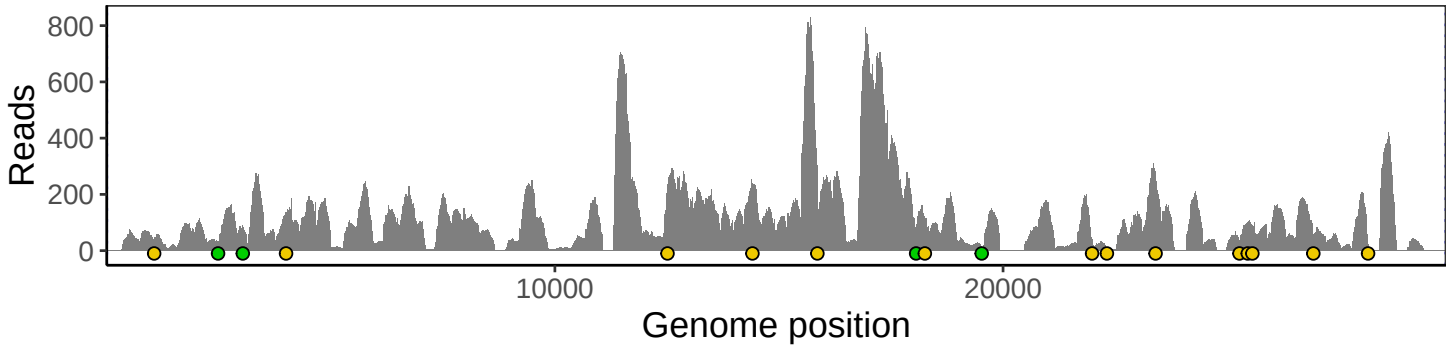


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

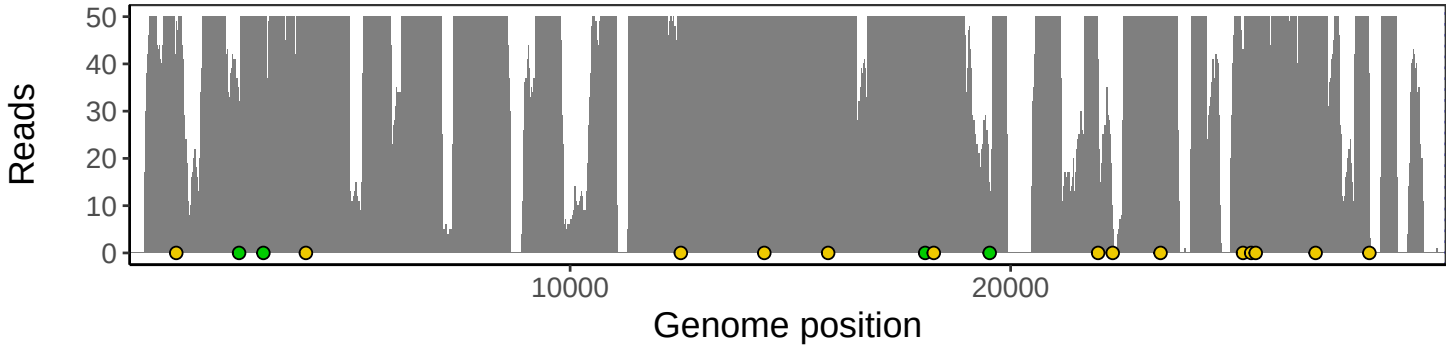


VSP0148-1 | 2020-05-18 | NP-OP | 263no-q2 | genomes | single experiment

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.