COVID-19 subject 272

2020-10-23

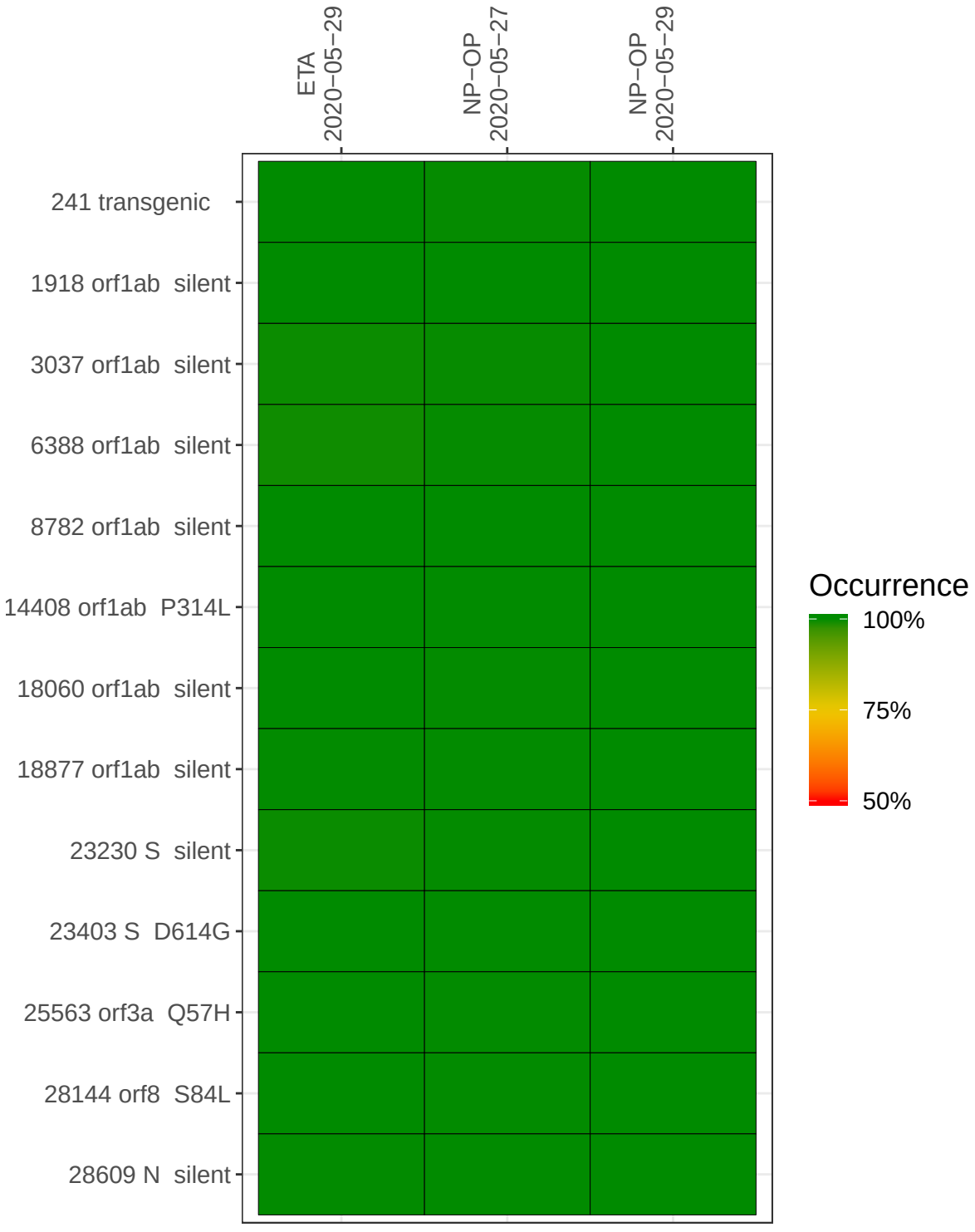
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

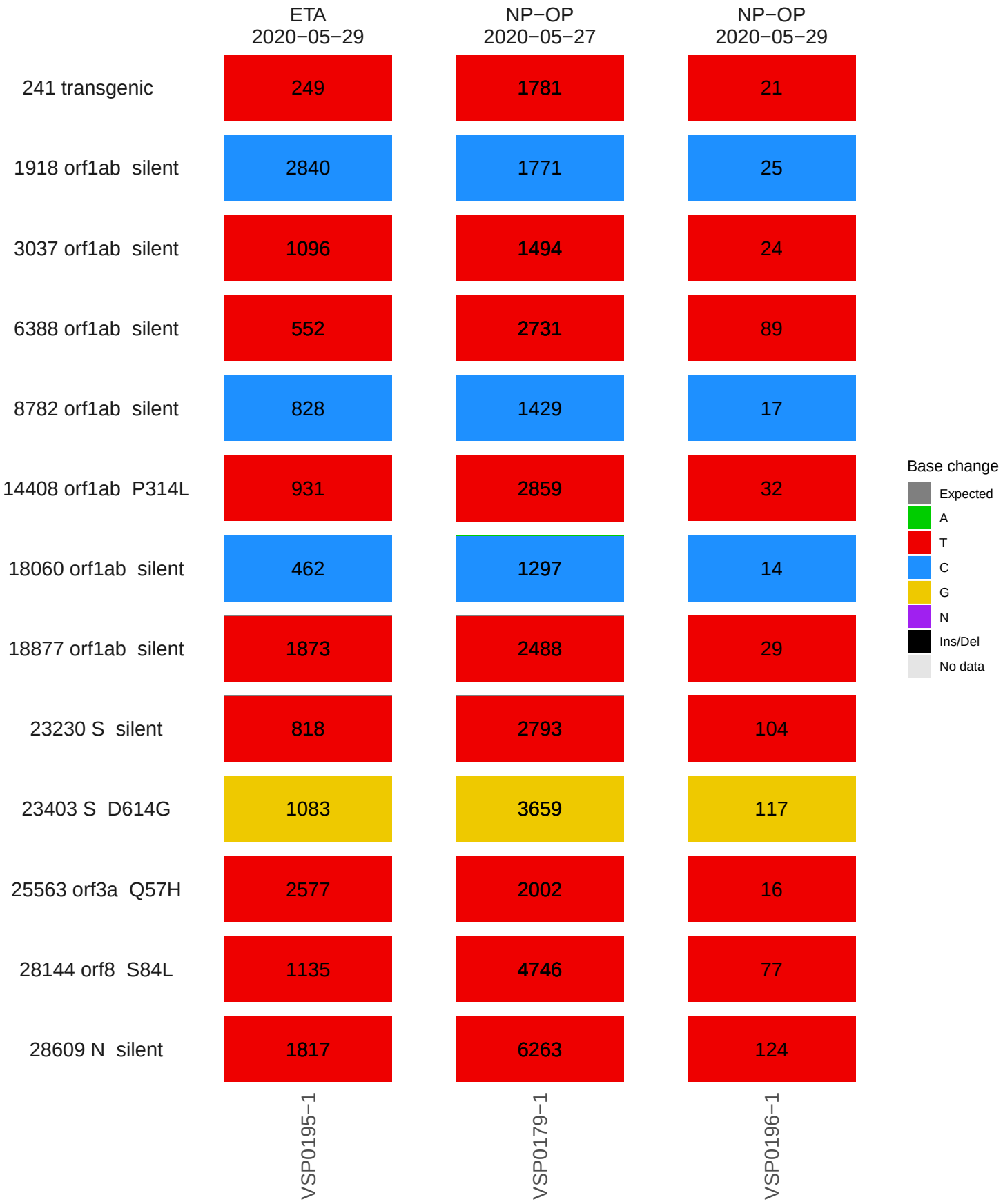
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0179-1	single experiment	14500	NP-OP	2020-05-27	29.41	99.8%	99.8%
VSP0195-1	single experiment	1530000	ETA	2020-05-29	25.63	99.7%	98.7%
VSP0196-1	single experiment	14300	NP-OP	2020-05-29	12.50	99.1%	97.7%
VSP0260-1	single experiment	NA	Stool	2020-06-17	NA	NA	NA

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.

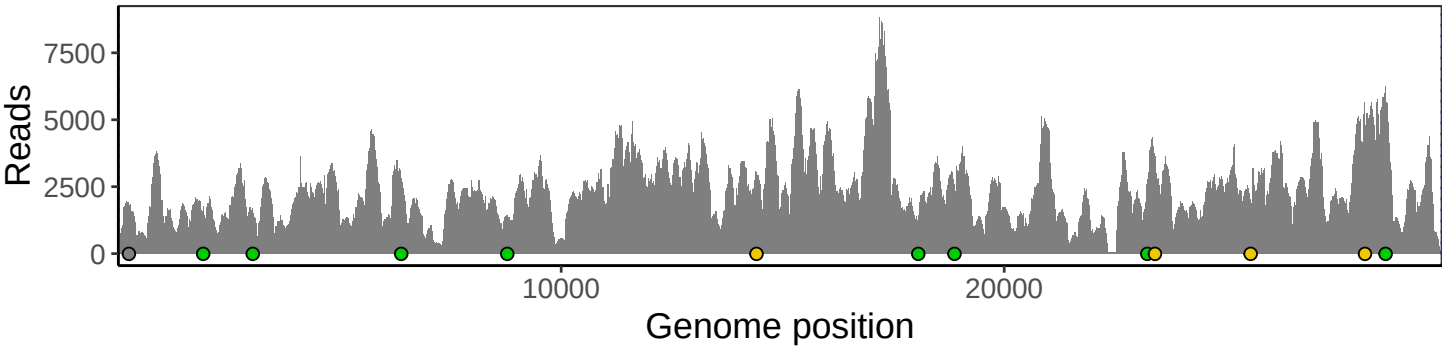



ETA
2020−05−29
NP−OP
2020−05−27
NP−OP
2020−05−29
241 transgenic
249
1781
21
1918 orf1ab silent
2840
1771
25
3037 orf1ab silent
1096
1494
24
6388 orf1ab silent
552
2731
89
8782 orf1ab silent
828
1429
17
14408 orf1ab P314L
931
2859
32
18060 orf1ab silent
462
1297
14
18877 orf1ab silent
1873
2488
29
23230 S silent
818
2793
104
23403 S D614G
1083
3659
117
25563 orf3a Q57H
2577
2002
16
28144 orf8 S84L
1135
4746
77
28609 N silent
1817
6263
124
VSP0195−1
VSP0179−1
VSP0196−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

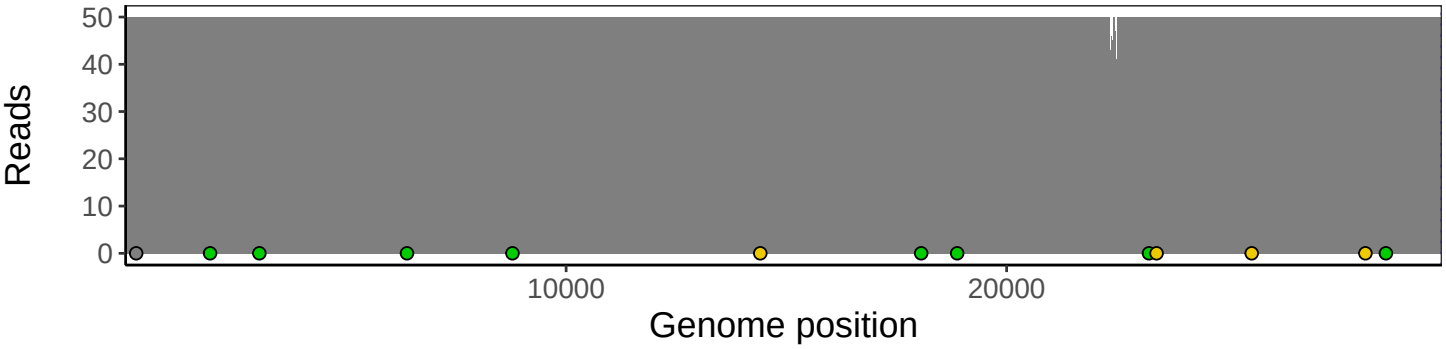
Analyses of individual experiments and composite results.

VSP0179-1 | 2020-05-27 | NP-OP | 272no-q | 14500 genomes | single experiment

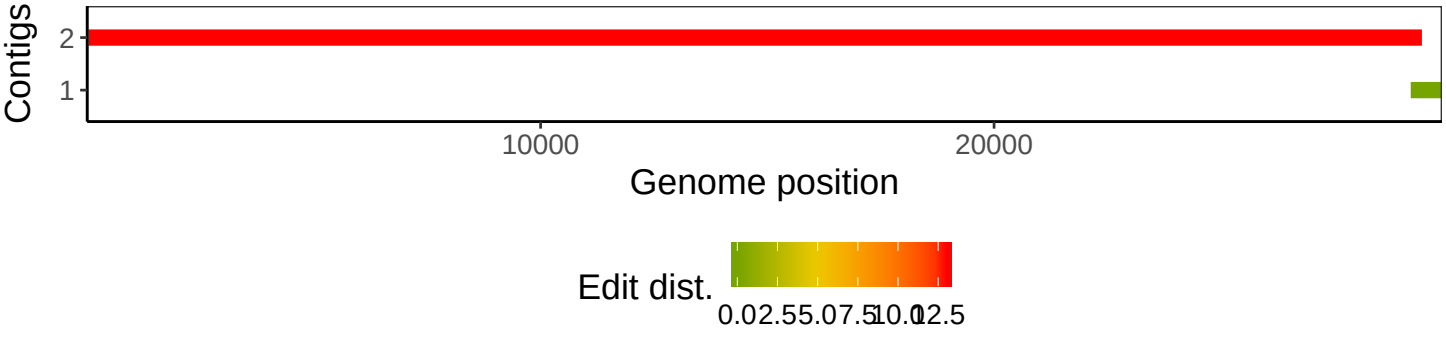
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

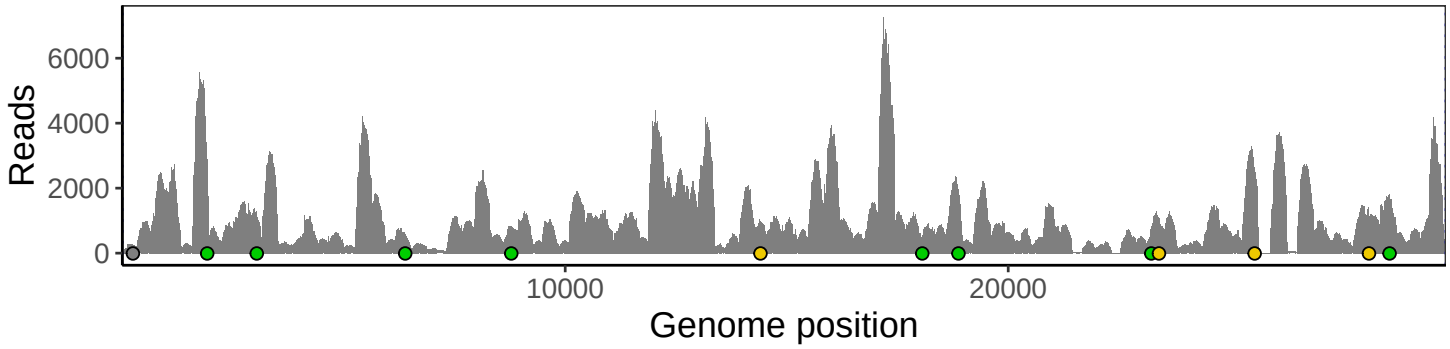


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

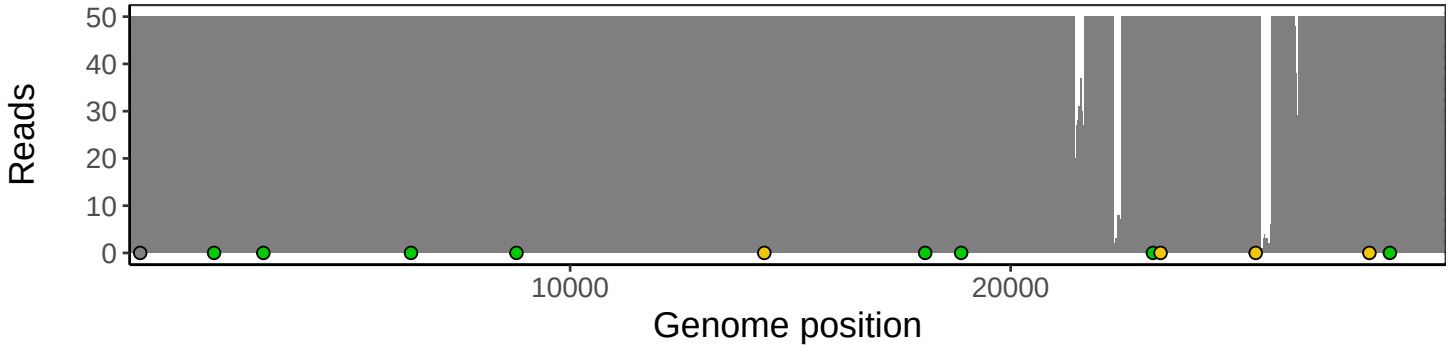


VSP0195-1 | 2020-05-29 | ETA | 272e-q | 1530000 genomes | single experiment

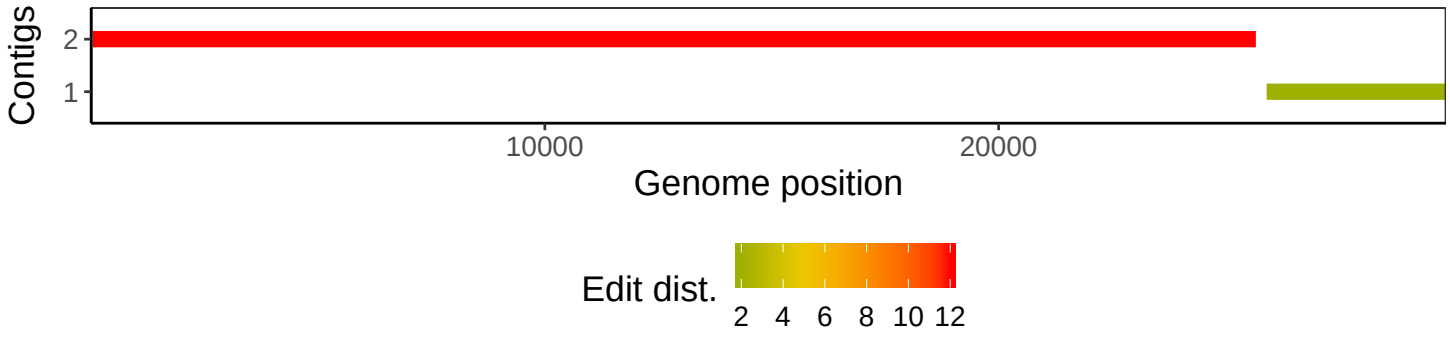
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

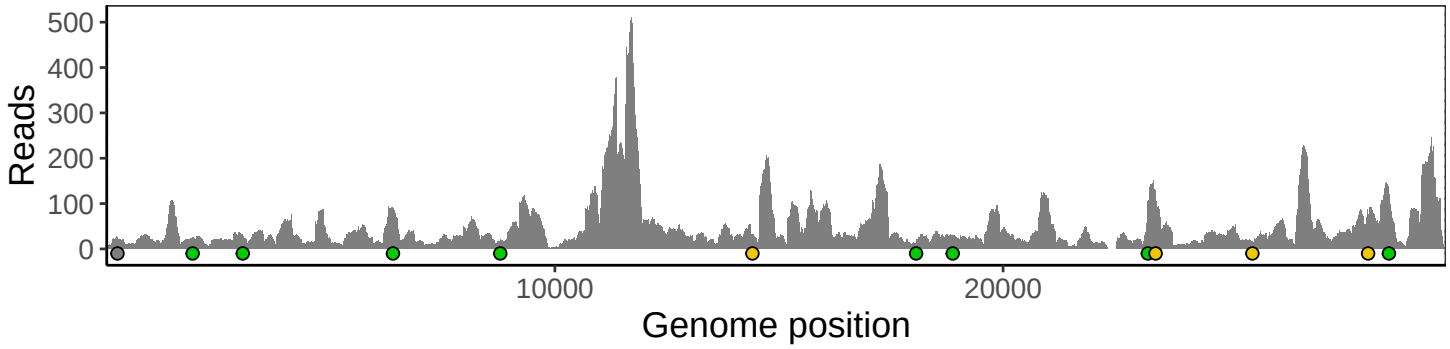


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

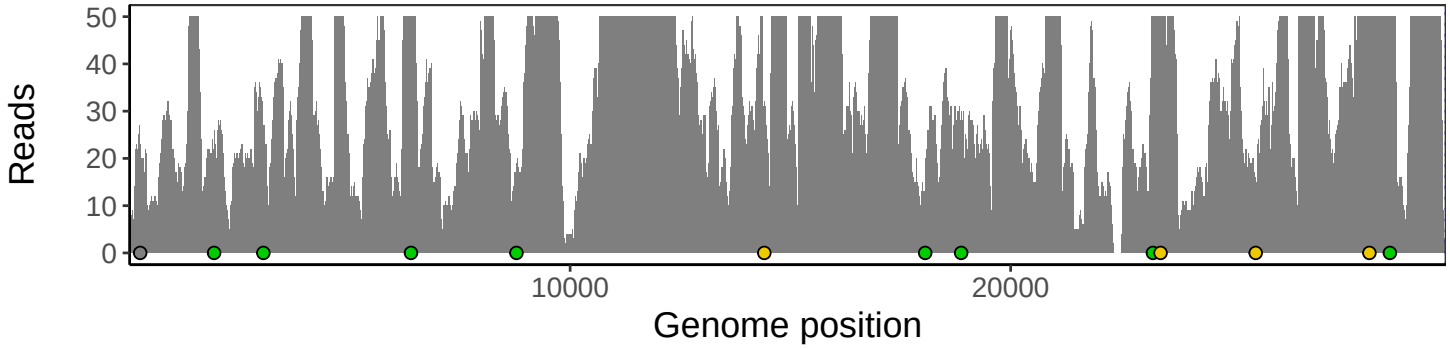


VSP0196-1 | 2020-05-29 | NP-OP | 272no-q | 14300 genomes | single experiment

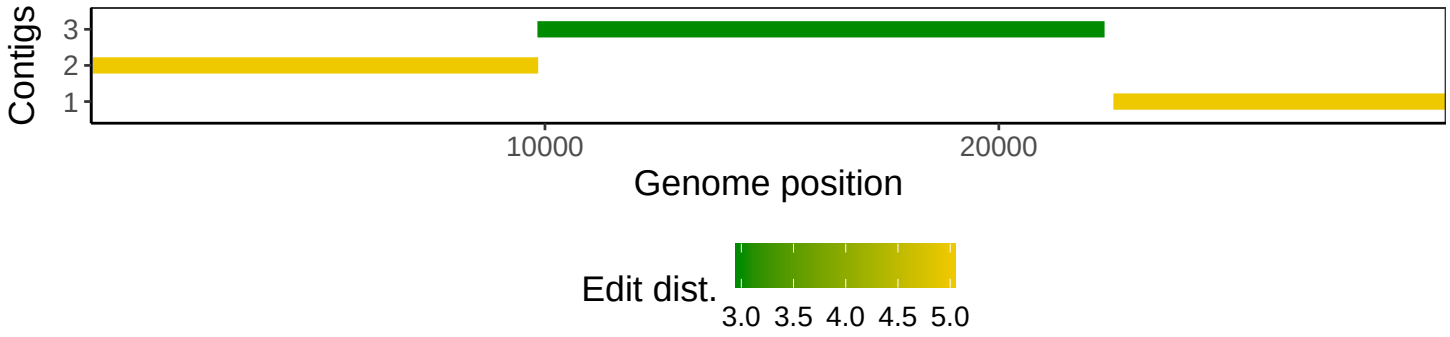
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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.



No pileup data available.

No contig data available.