COVID-19 subject 235

2021-01-19

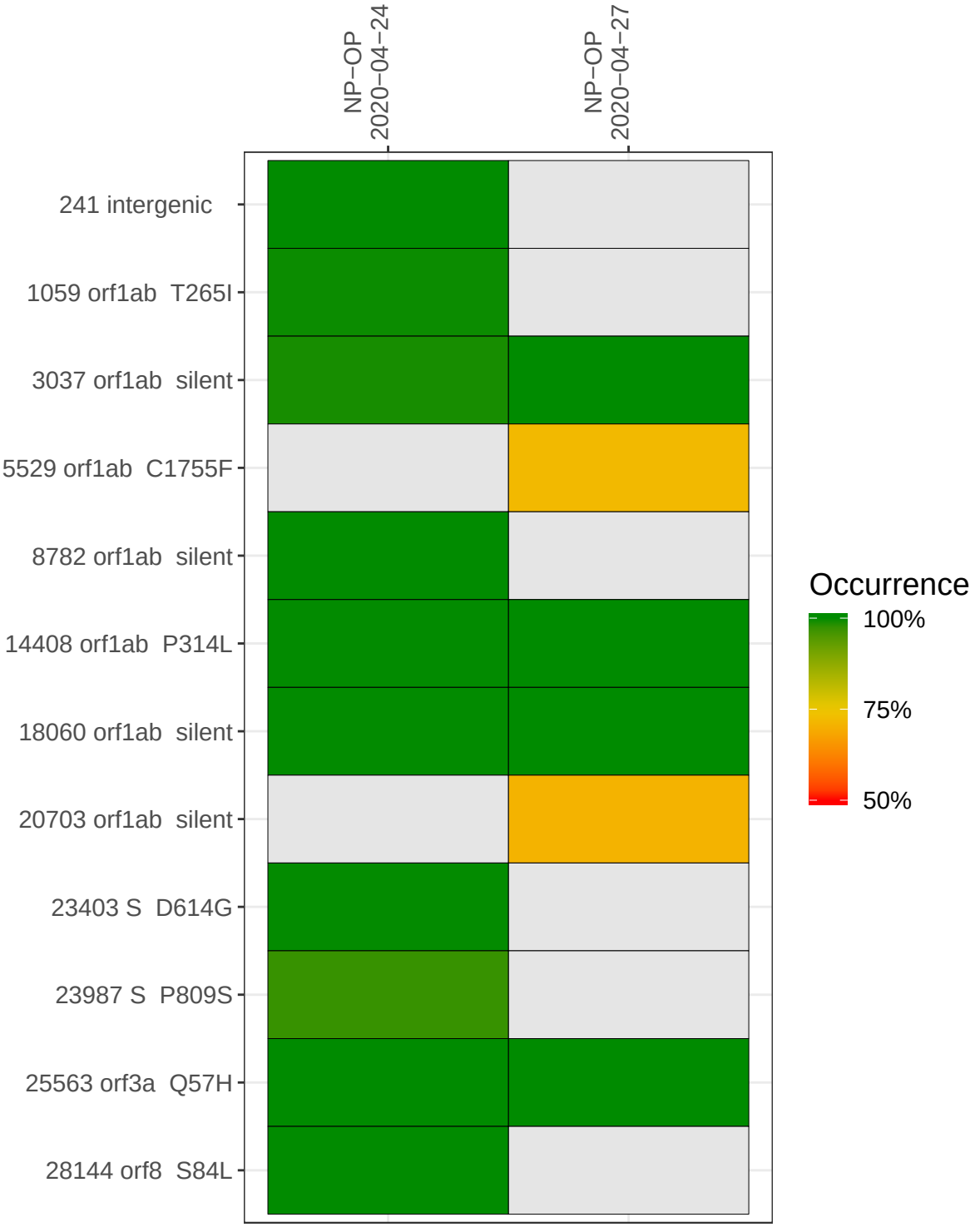
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

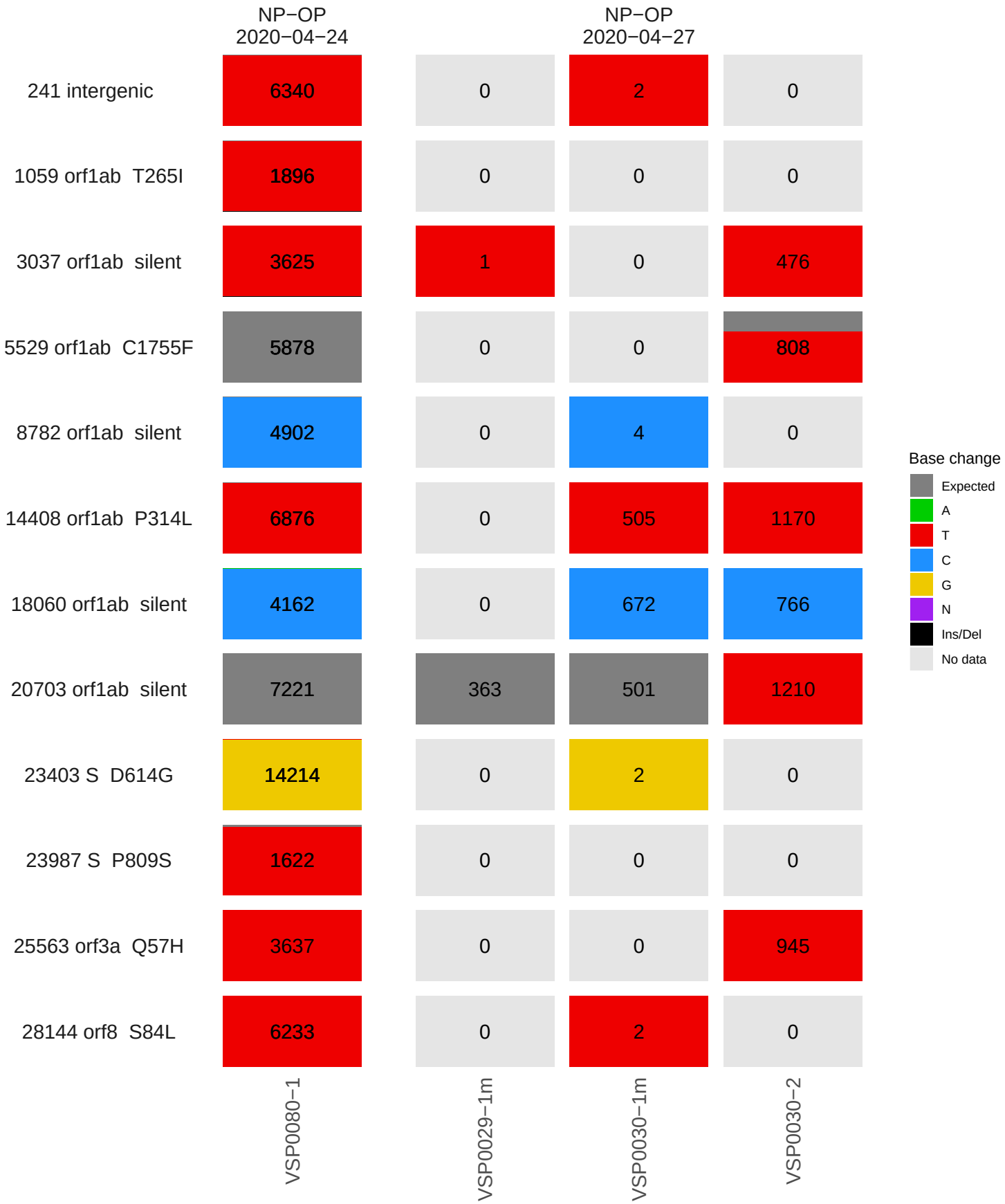
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0030	composite	NA	NP-OP	2020-04-27	2.47	76.7%	57.1%
VSP0029-1m	single experiment	NA	NP-OP	2020-04-27	0.50	17.5%	12.5%
VSP0030-1m	single experiment	NA	NP-OP	2020-04-27	1.05	51.1%	23.4%
VSP0030-2	single experiment	NA	NP-OP	2020-04-27	1.96	55.0%	53.0%
VSP0080-1	single experiment	176000	NP-OP	2020-04-24	29.87	99.9%	99.9%

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 or 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.

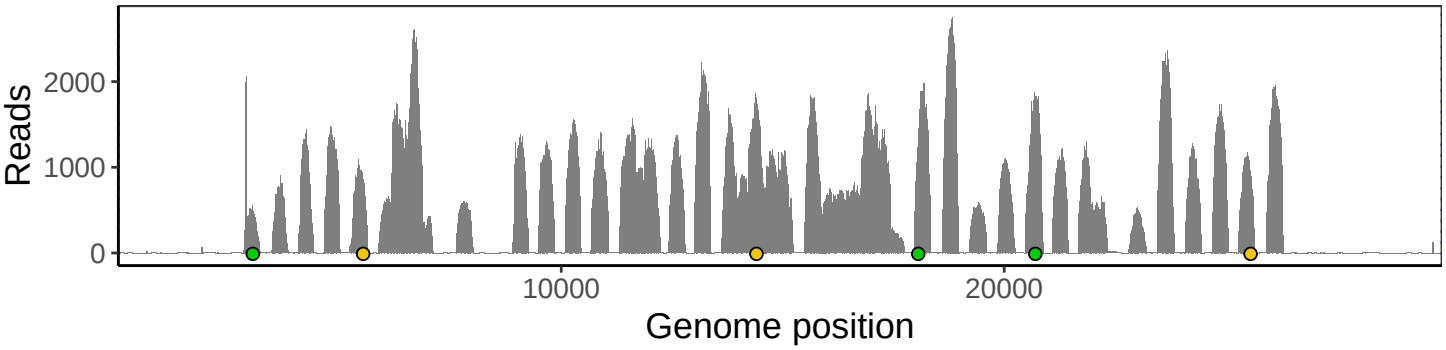



NP−OP
2020−04−24
NP−OP
2020−04−27
241 intergenic
1059 orf1ab T265I
3037 orf1ab silent
5529 orf1ab C1755F
8782 orf1ab silent
14408 orf1ab P314L
18060 orf1ab silent
20703 orf1ab silent
23403 S D614G
23987 S P809S
25563 orf3a Q57H
28144 orf8 S84L
6340
0
2
0
1896
0
0
0
3625
1
0
476
5878
0
0
808
4902
0
4
0
6876
0
505
1170
4162
0
672
766
7221
363
501
1210
14214
0
2
0
1622
0
0
0
3637
0
0
945
6233
0
2
0
VSP0080−1
VSP0029−1m
VSP0030−1m
VSP0030−2
Base change
Expected
A
T
C
G
N
Ins/Del
No data

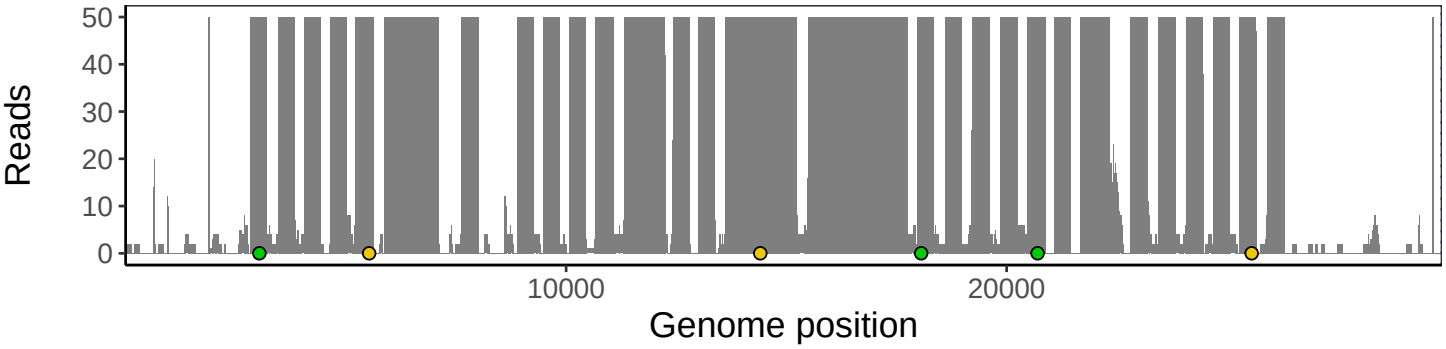
Analyses of individual experiments and composite results.

VSP0030 | 2020-04-27 | NP-OP | 235-qia | composite result

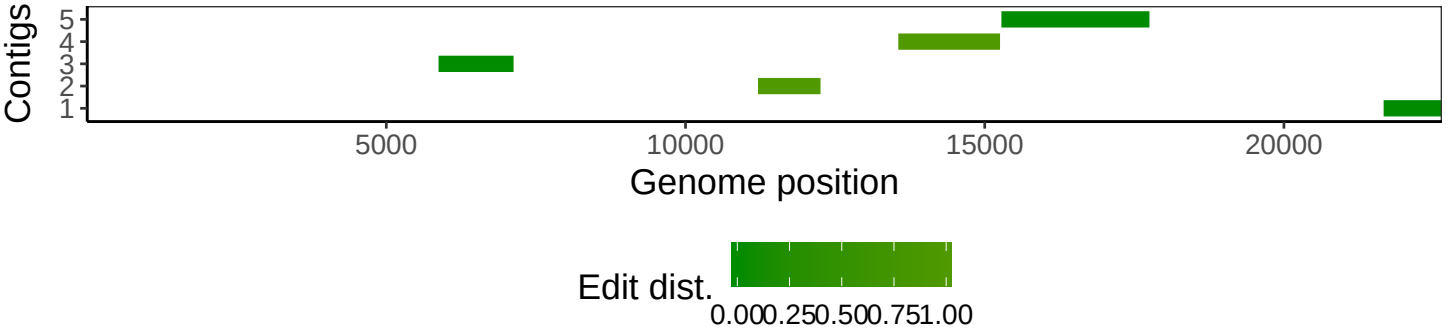
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

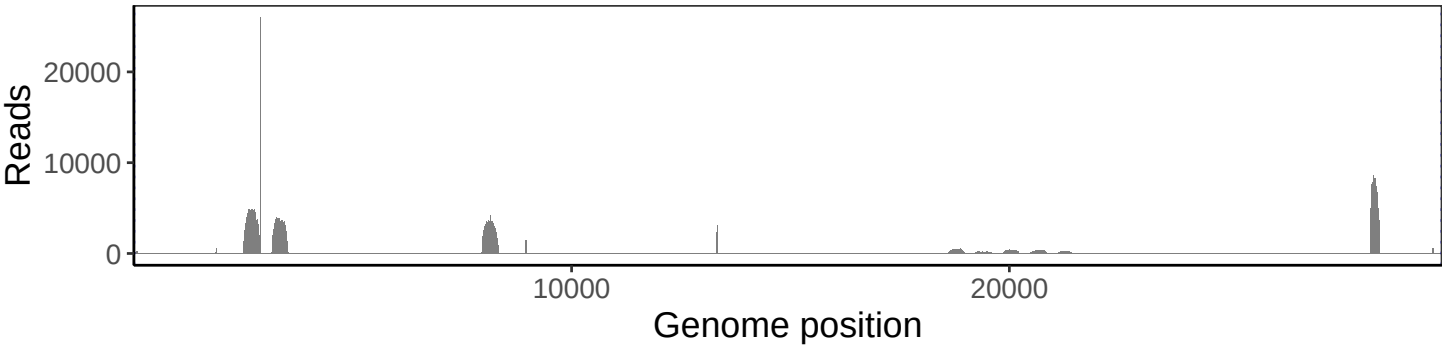


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

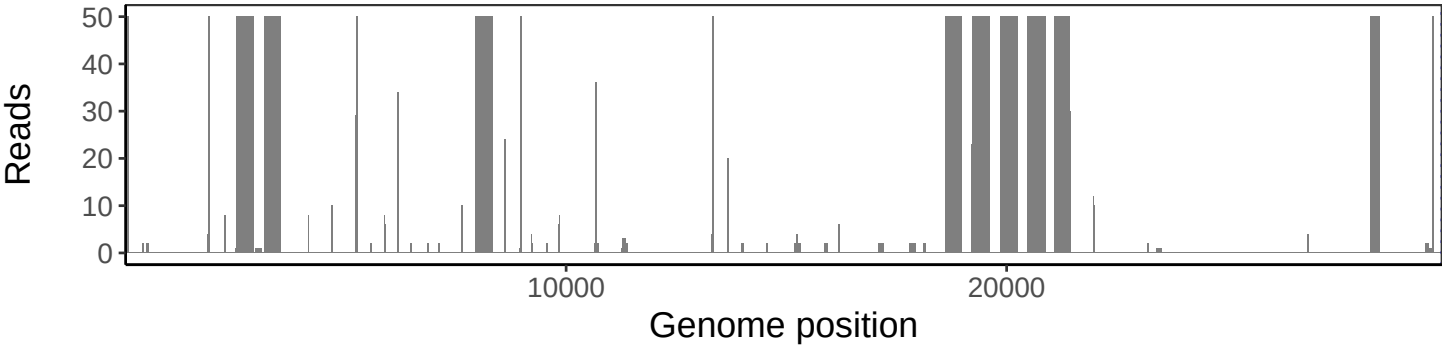


VSP0029-1m | 2020-04-27 | NP-OP | 235-tri | genomes | single experiment

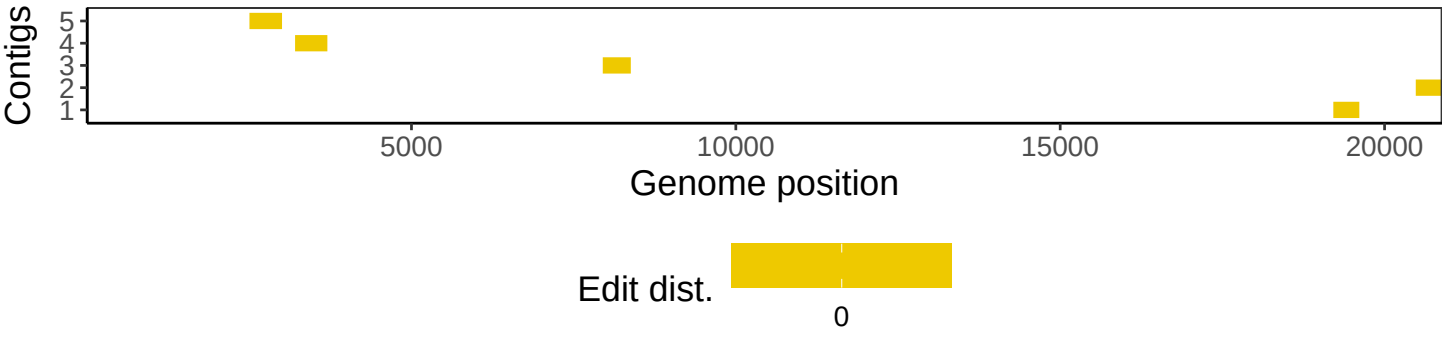
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

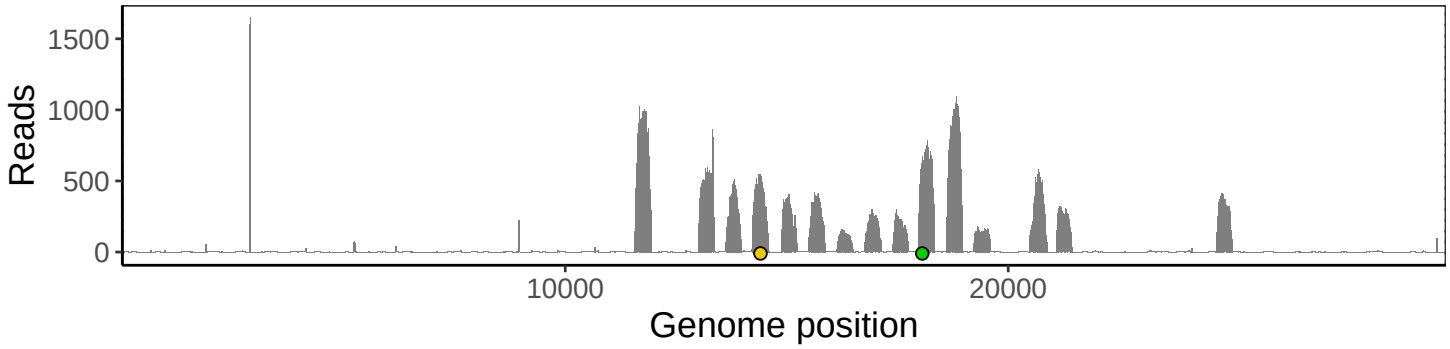


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

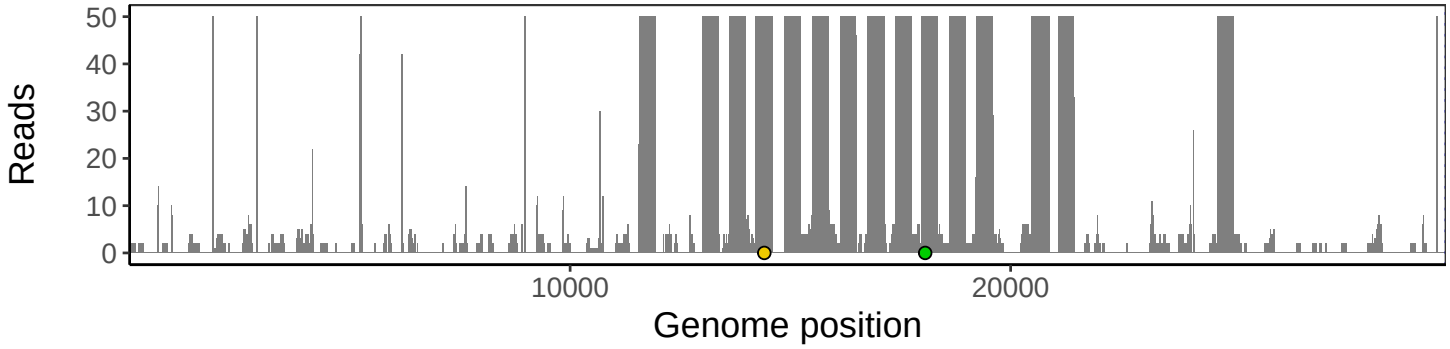


VSP0030-1m | 2020-04-27 | NP-OP | 235-qia | genomes | single experiment

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.



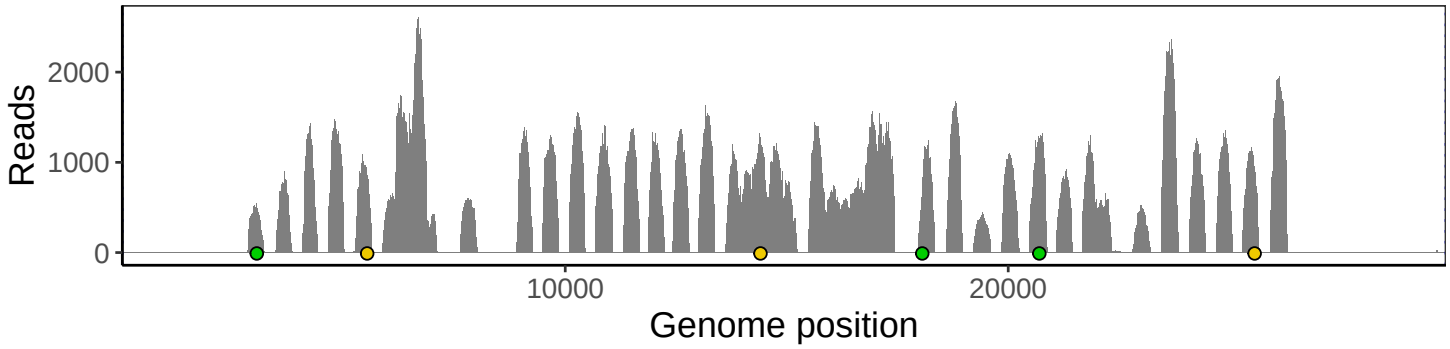
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



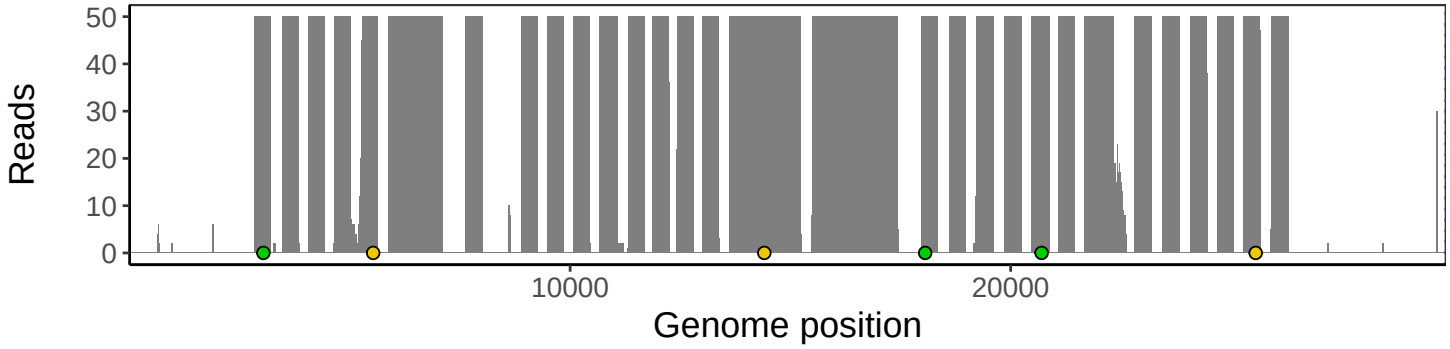
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

VSP0030-2 | 2020-04-27 | NP-OP | 235-qia | genomes | single experiment

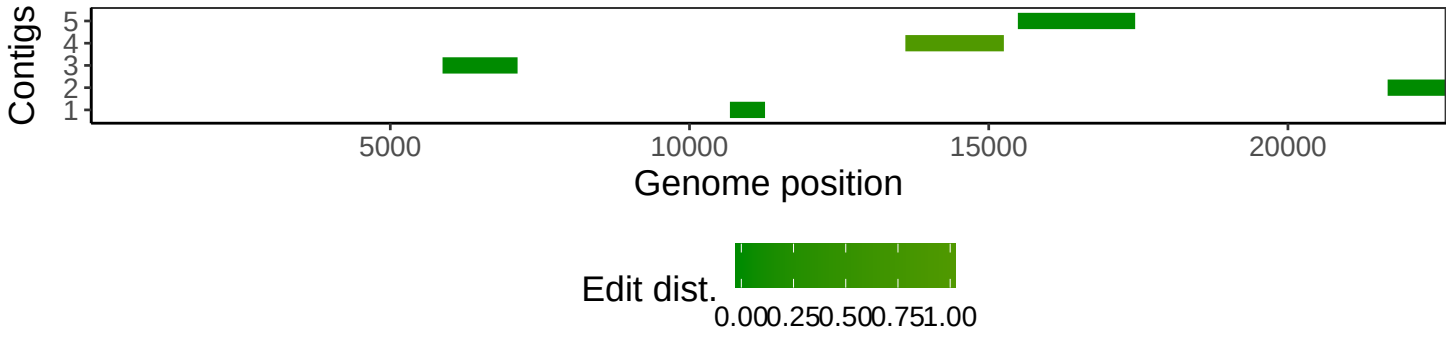
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

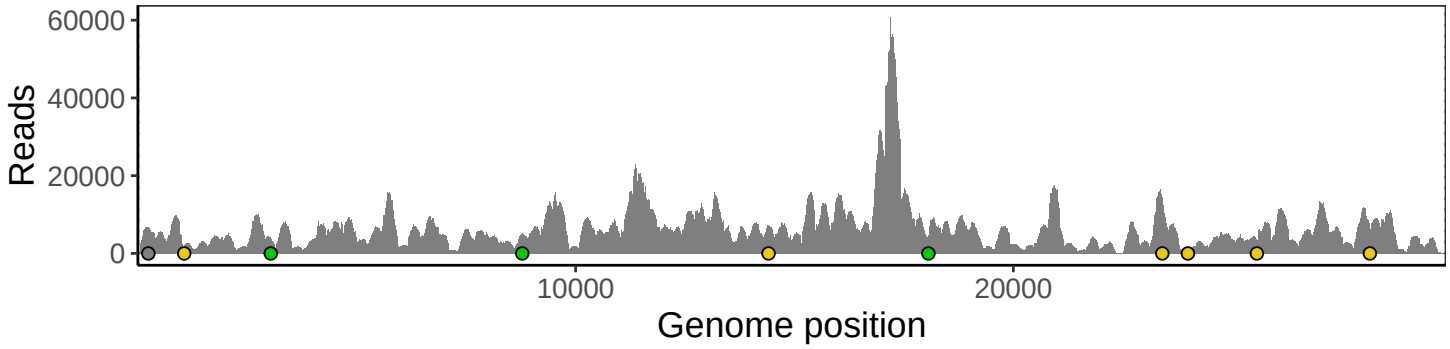


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

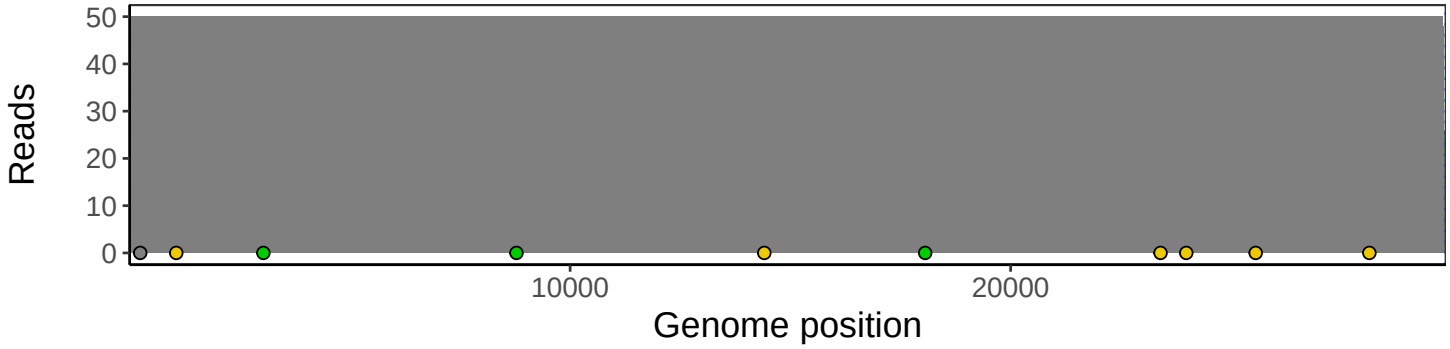
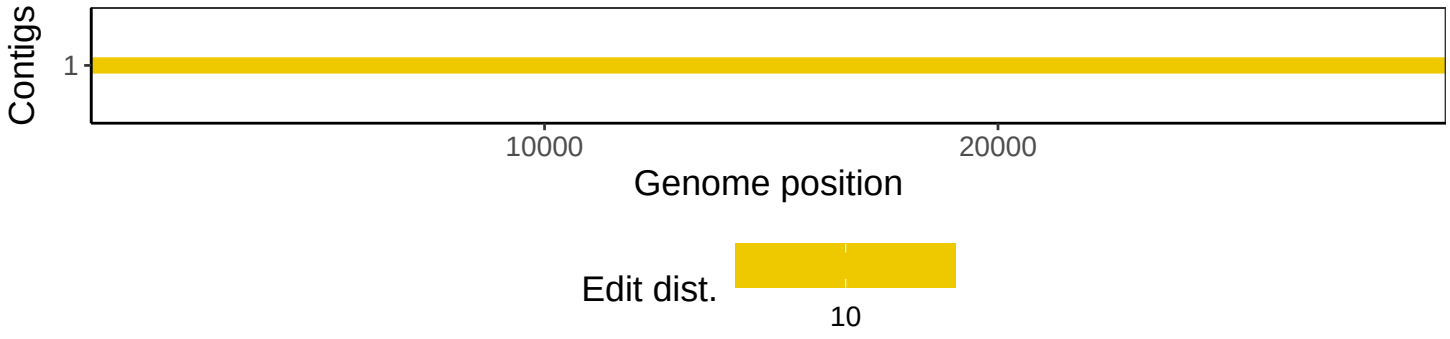


VSP0080-1 | 2020-04-24 | NP-OP | 235no-t | 176000 genomes | single experiment

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.