COVID-19 subject 269

2020-09-14

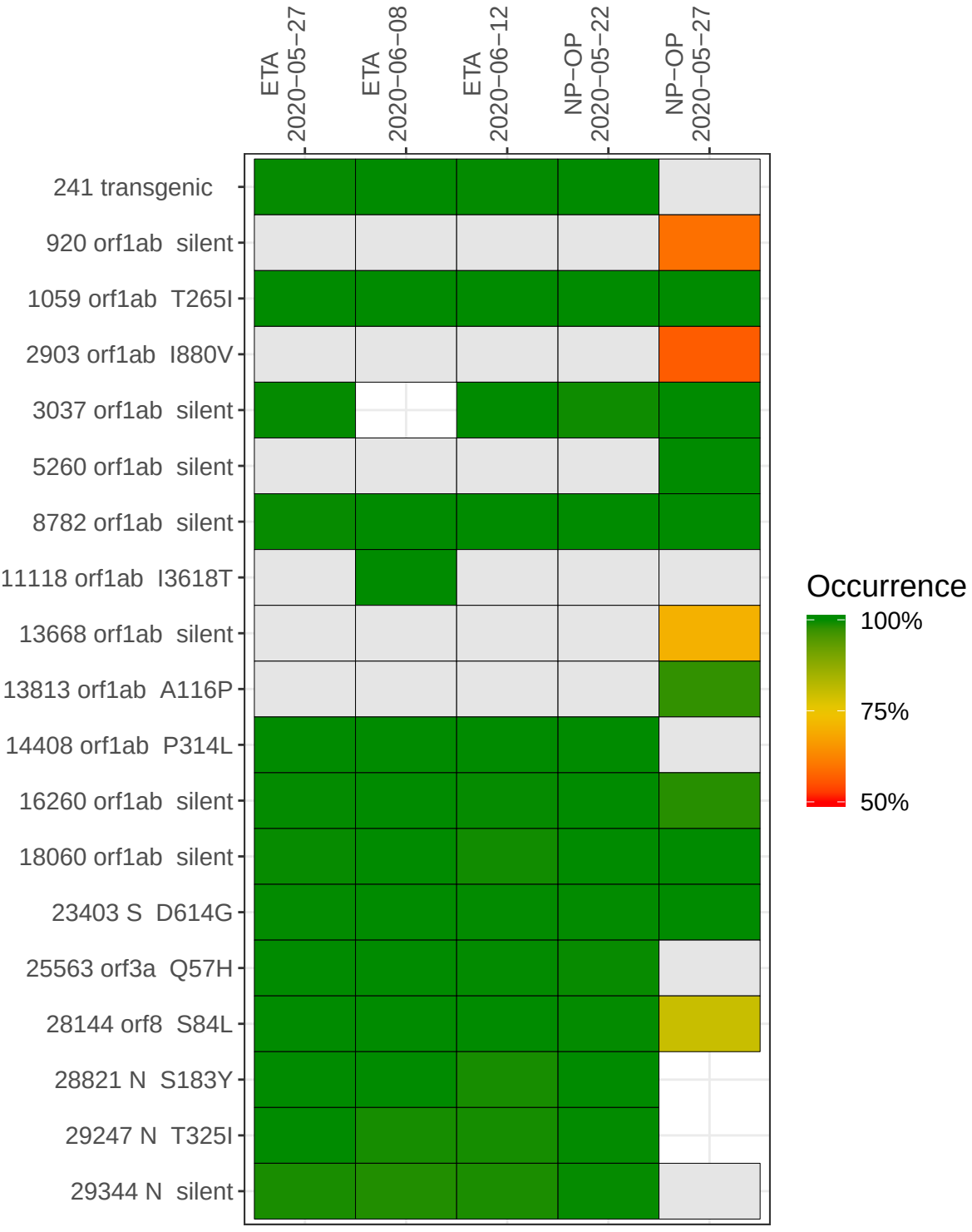
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

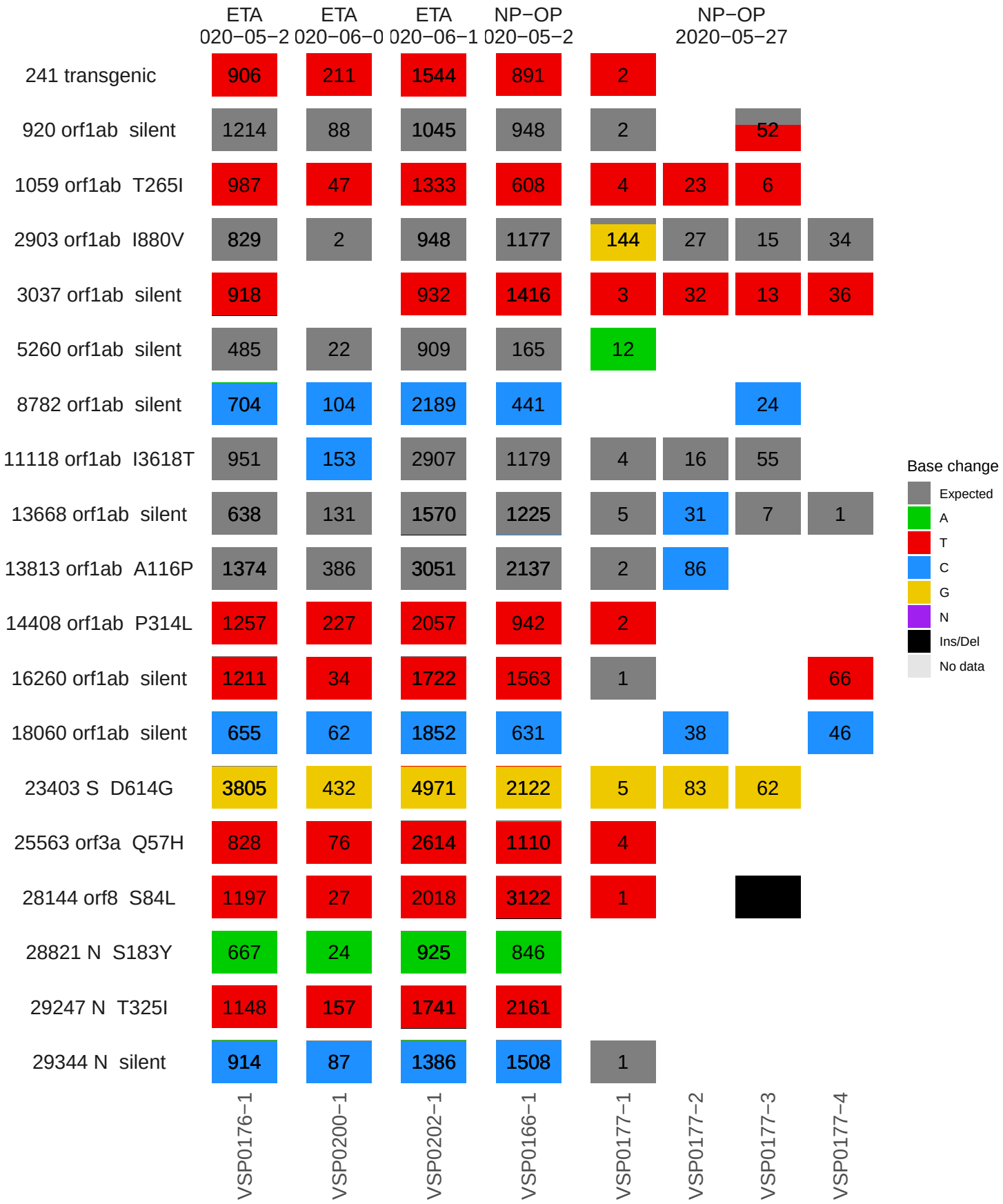
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0177	composite	NA	NP-OP	2020-05-27	3.70	89.8%	72.9%
VSP0166-1	single experiment	NA	NP-OP	2020-05-22	29.84	99.8%	99.5%
VSP0176-1	single experiment	166000.0	ETA	2020-05-27	29.82	99.8%	99.8%
VSP0177-1	single experiment	59.7	NP-OP	2020-05-27	1.03	72.3%	17.5%
VSP0177-2	single experiment	NA	NP-OP	2020-05-27	1.00	31.5%	26.0%
VSP0177-3	single experiment	NA	NP-OP	2020-05-27	1.02	40.8%	35.0%
VSP0177-4	single experiment	NA	NP-OP	2020-05-27	0.93	38.4%	35.6%
VSP0200-1	single experiment	NA	ETA	2020-06-08	5.92	84.5%	82.0%
VSP0202-1	single experiment	16300.0	ETA	2020-06-12	29.83	99.8%	99.8%

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.

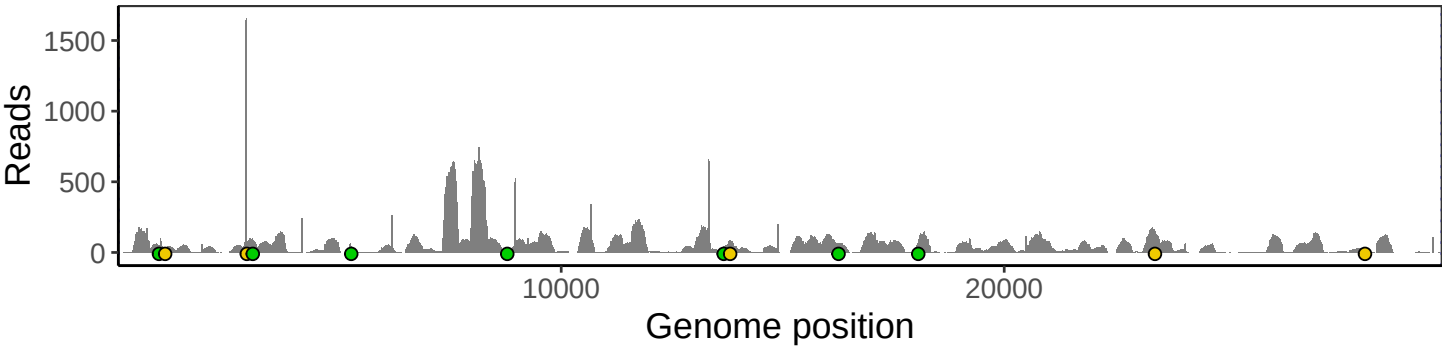



ETA 020−05−2
ETA 020−06−0
ETA 020−06−1
NP−OP 020−05−2
NP−OP 2020−05−27
241 transgenic
920 orf1ab silent
1059 orf1ab T265I
2903 orf1ab I880V
3037 orf1ab silent
5260 orf1ab silent
8782 orf1ab silent
11118 orf1ab I3618T
13668 orf1ab silent
13813 orf1ab A116P
14408 orf1ab P314L
16260 orf1ab silent
18060 orf1ab silent
23403 S D614G
25563 orf3a Q57H
28144 orf8 S84L
28821 N S183Y
29247 N T325I
29344 N silent
906 211 1544 891 2
1214 88 1045 948 2 52
987 47 1333 608 4 23 6
829 2 948 1177 144 27 15 34
918 932 1416 3 32 13 36
485 22 909 165 12
704 104 2189 441 24
951 153 2907 1179 4 16 55
638 131 1570 1225 5 31 7 1
1374 386 3051 2137 2 86
1257 227 2057 942 2
1211 34 1722 1563 1 66
655 62 1852 631 38 46
3805 432 4971 2122 5 83 62
828 76 2614 1110 4
1197 27 2018 3122 1
667 24 925 846
1148 157 1741 2161
914 87 1386 1508 1
VSP0176−1
VSP0200−1
VSP0202−1
VSP0166−1
VSP0177−1
VSP0177−2
VSP0177−3
VSP0177−4
Base change
Expected
A
T
C
G
N
Ins/Del
No data

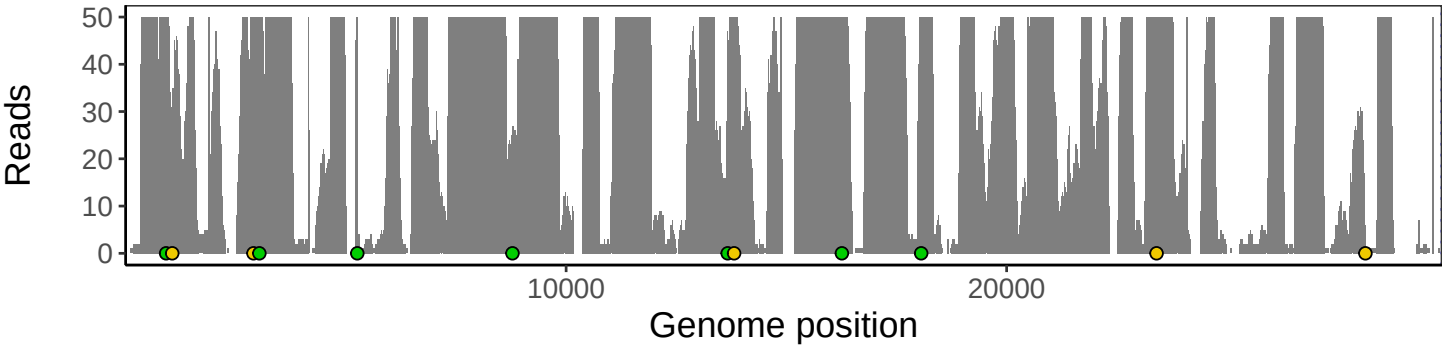
Analyses of individual experiments and composite results.

VSP0177 | 2020-05-27 | NP-OP | 269no-q | composite result

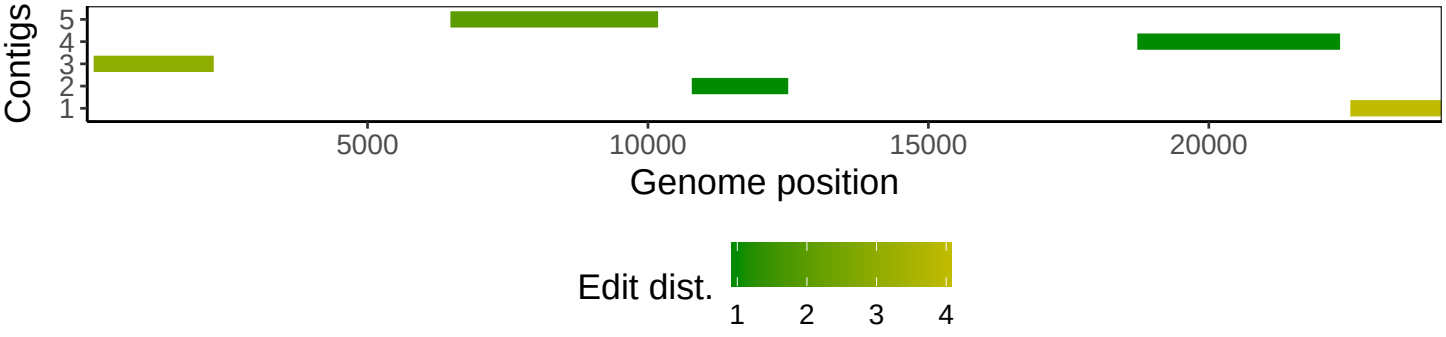
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

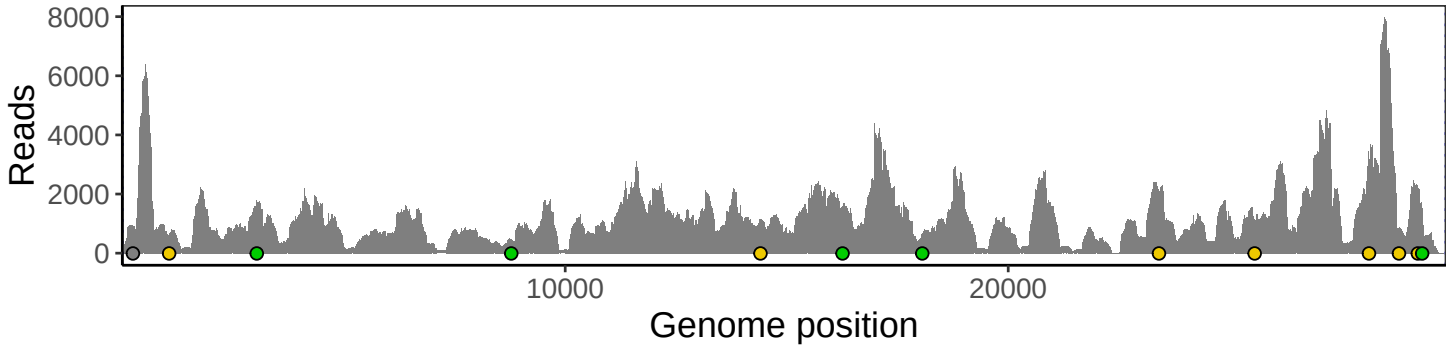


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

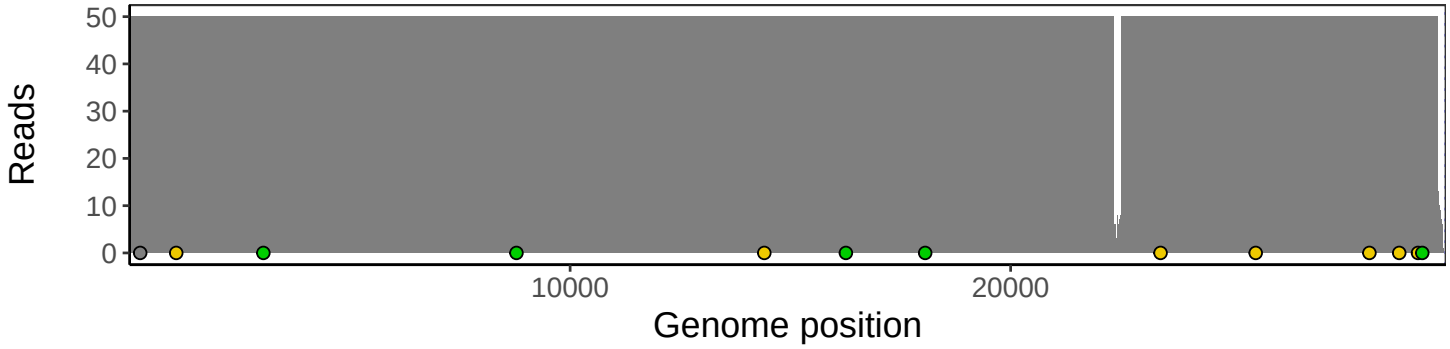


VSP0166-1 | 2020-05-22 | NP-OP | 269no-q | genomes | single experiment

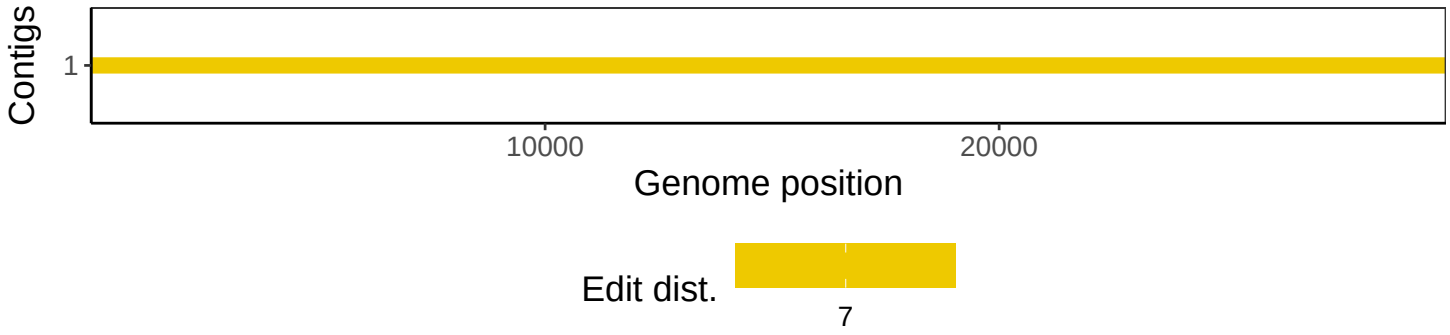
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

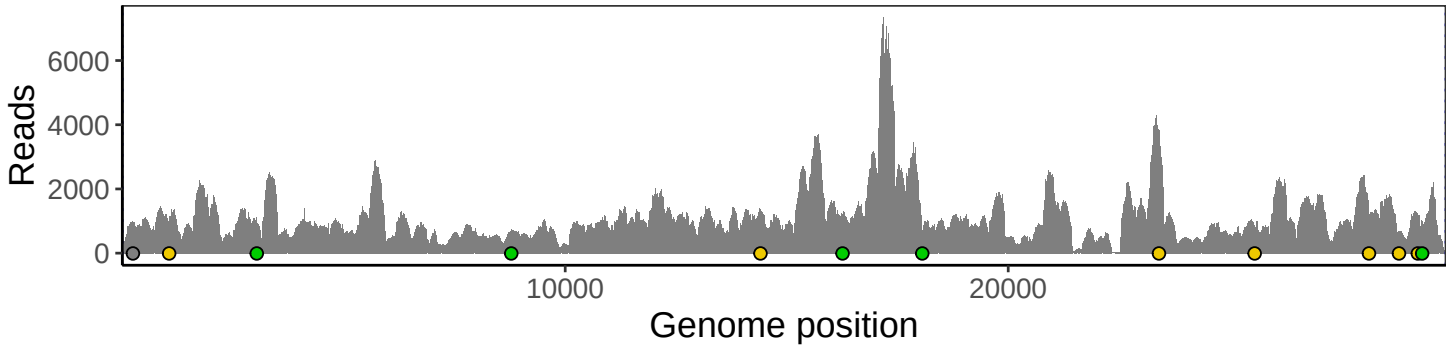


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

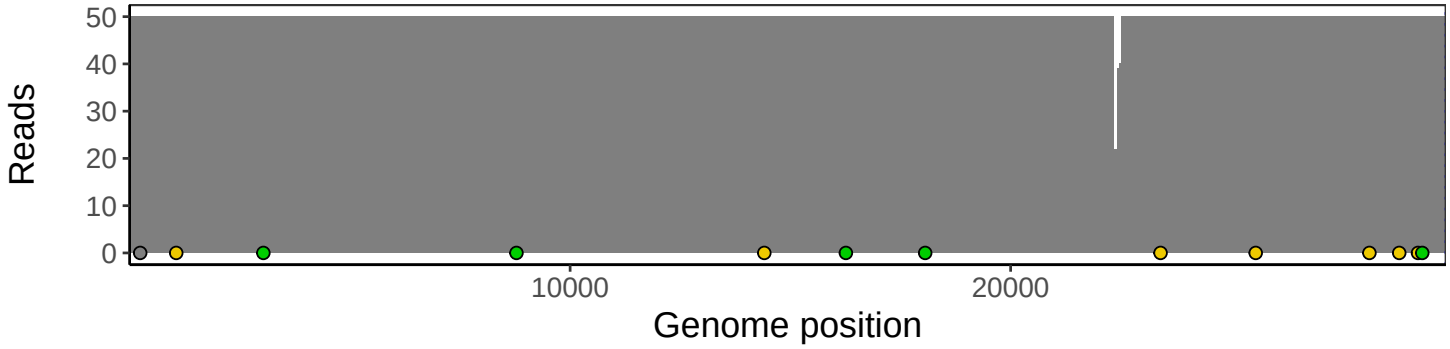


VSP0176-1 | 2020-05-27 | ETA | 269e-q | 166000 genomes | single experiment

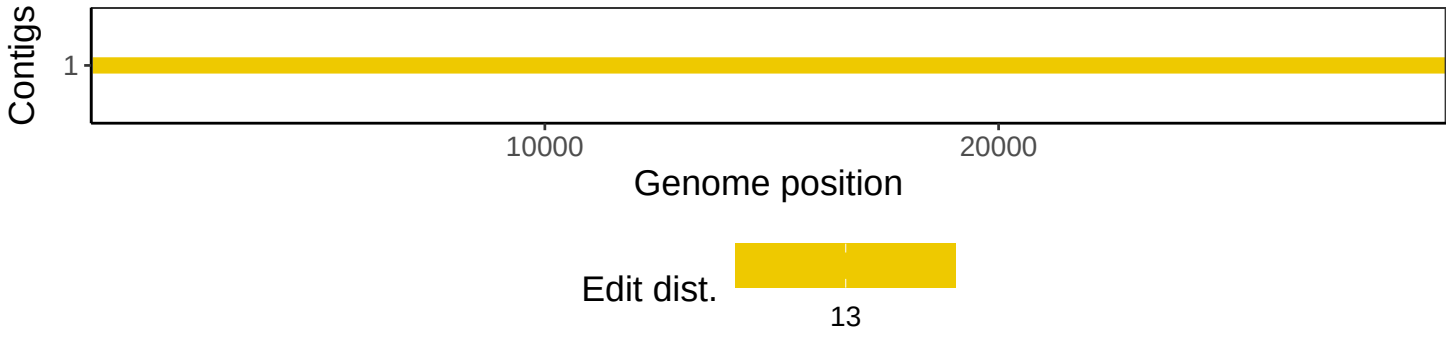
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

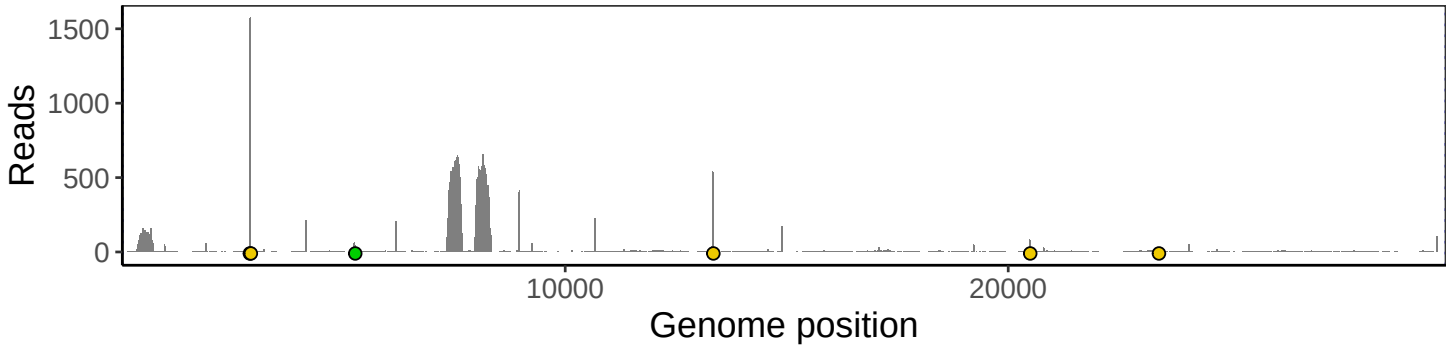


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

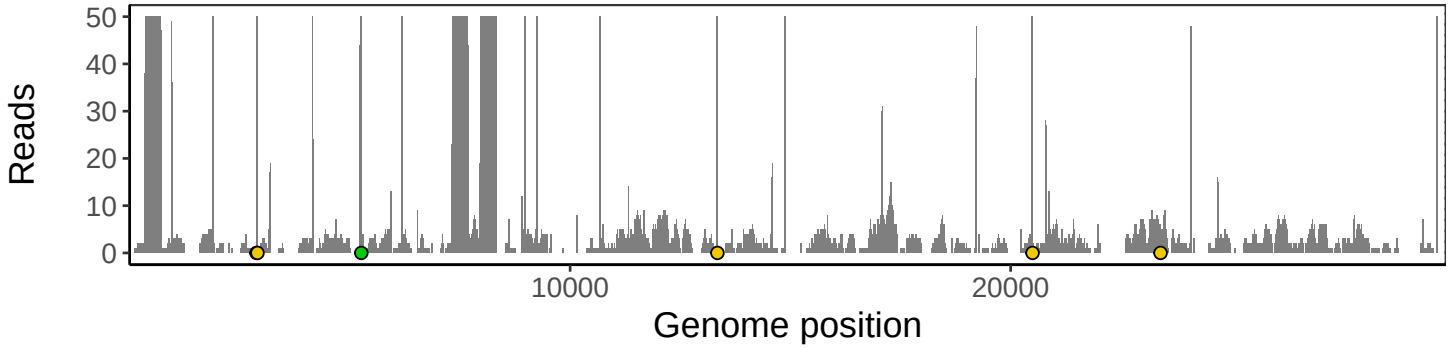


VSP0177-1 | 2020-05-27 | NP-OP | 269no-q | 59.7 genomes | single experiment

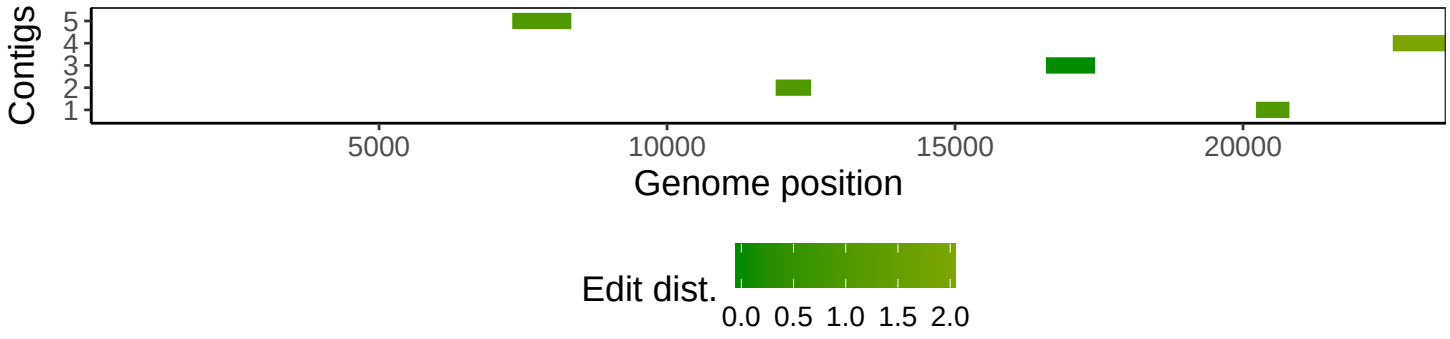
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

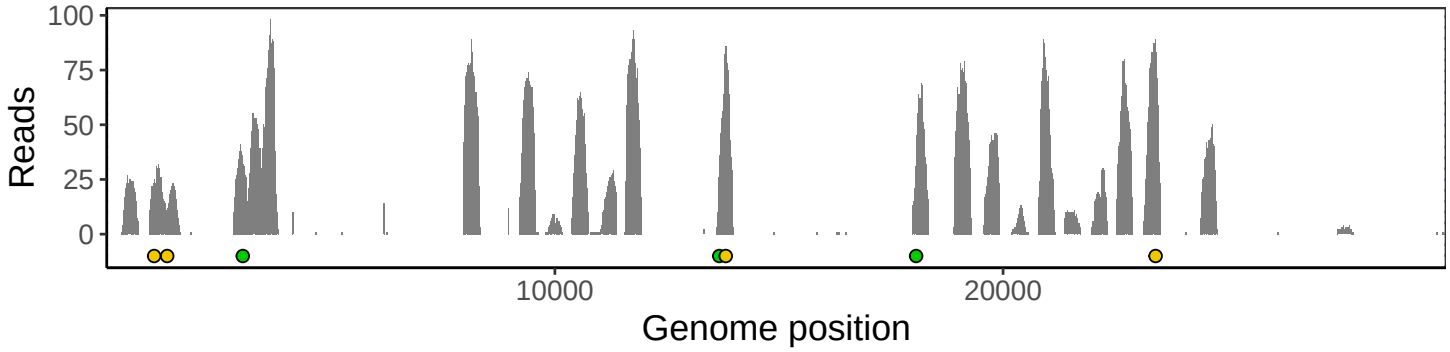


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

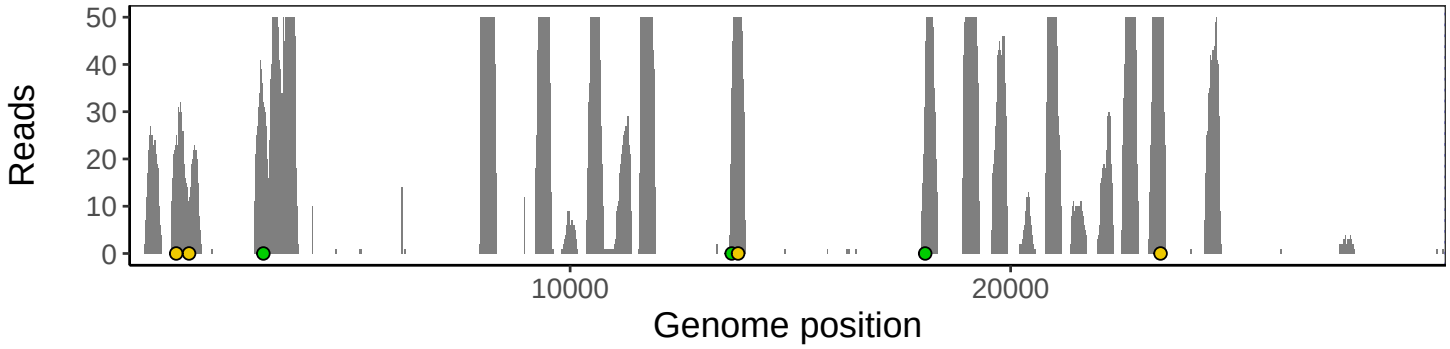


VSP0177-2 | 2020-05-27 | NP-OP | 269no-q | genomes | single experiment

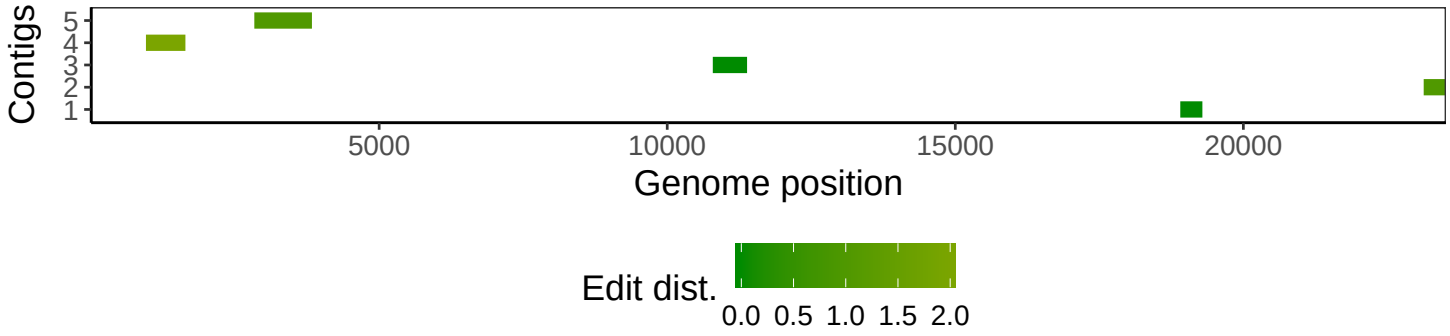
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

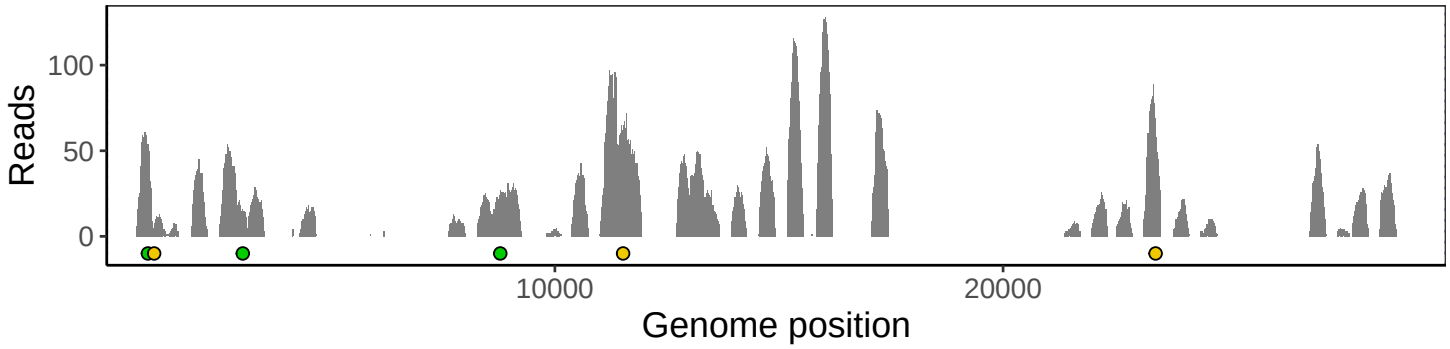


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

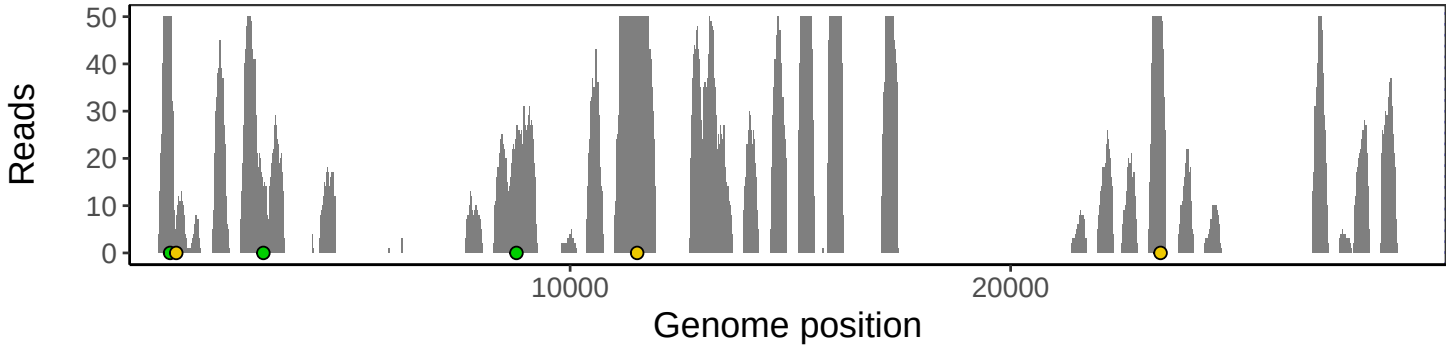


VSP0177-3 | 2020-05-27 | NP-OP | 269no-q | genomes | single experiment

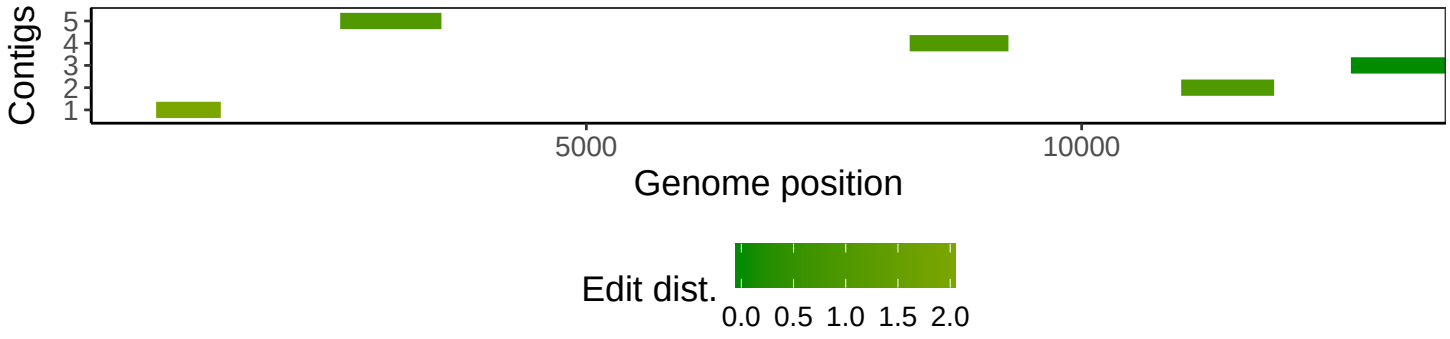
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

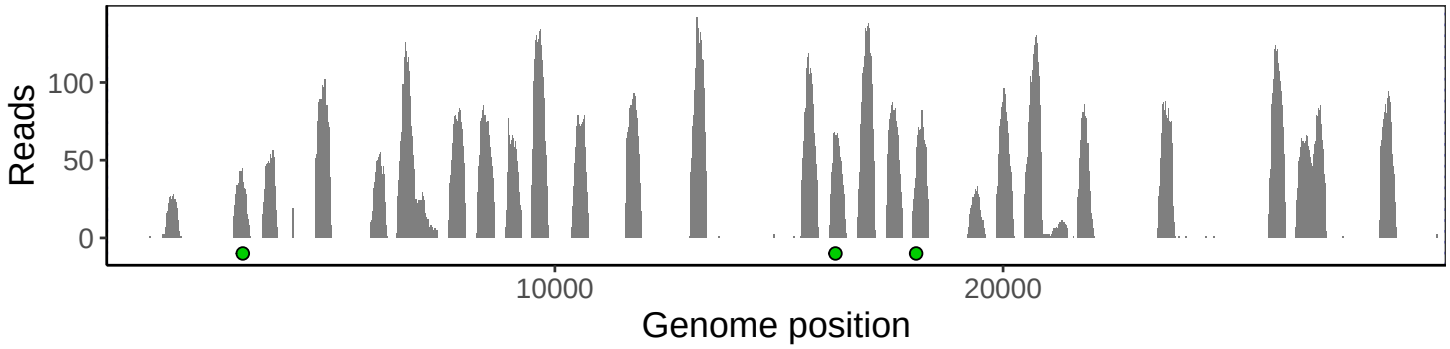


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

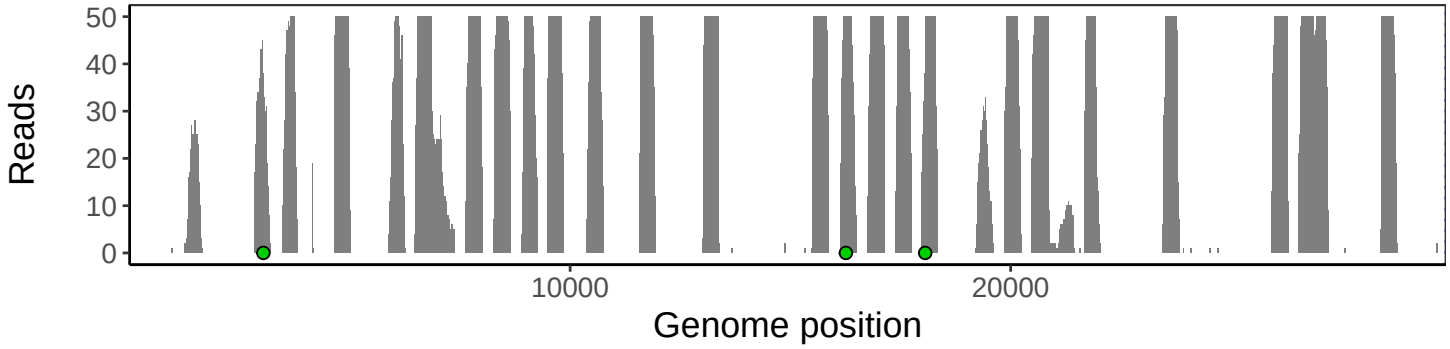


VSP0177-4 | 2020-05-27 | NP-OP | 269no-q | genomes | single experiment

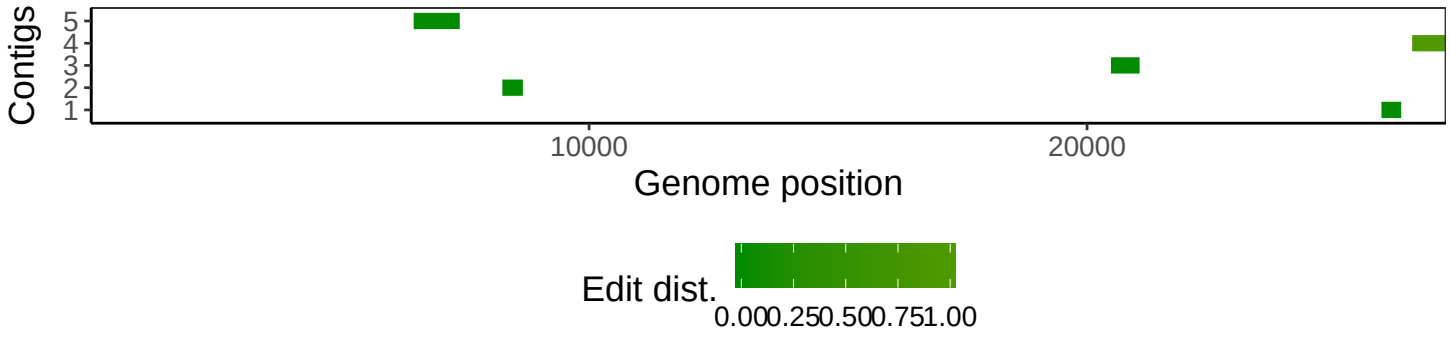
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

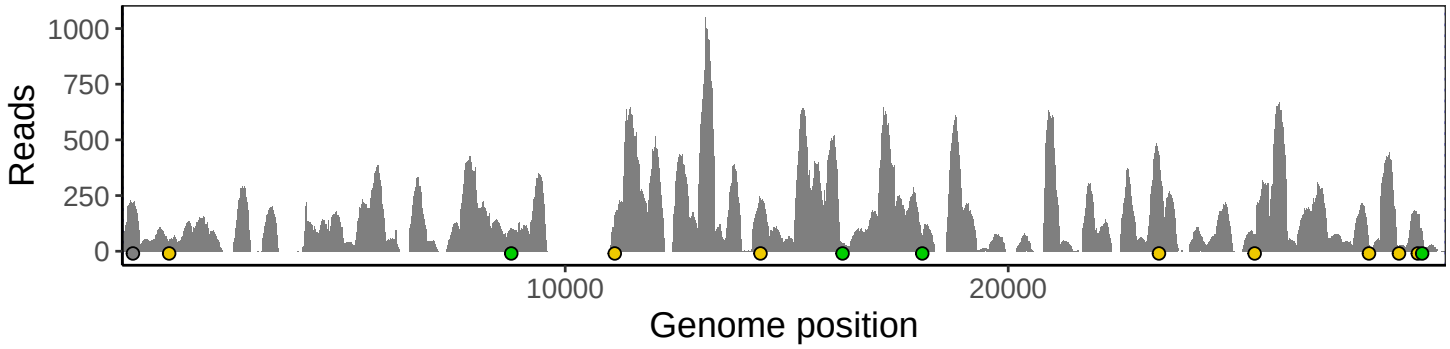


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

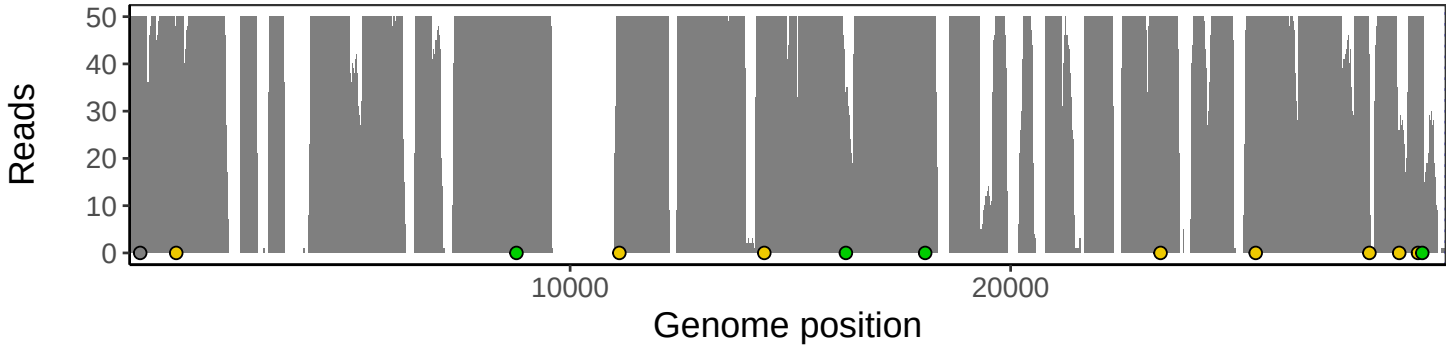


VSP0200-1 | 2020-06-08 | ETA | 269e-q | genomes | single experiment

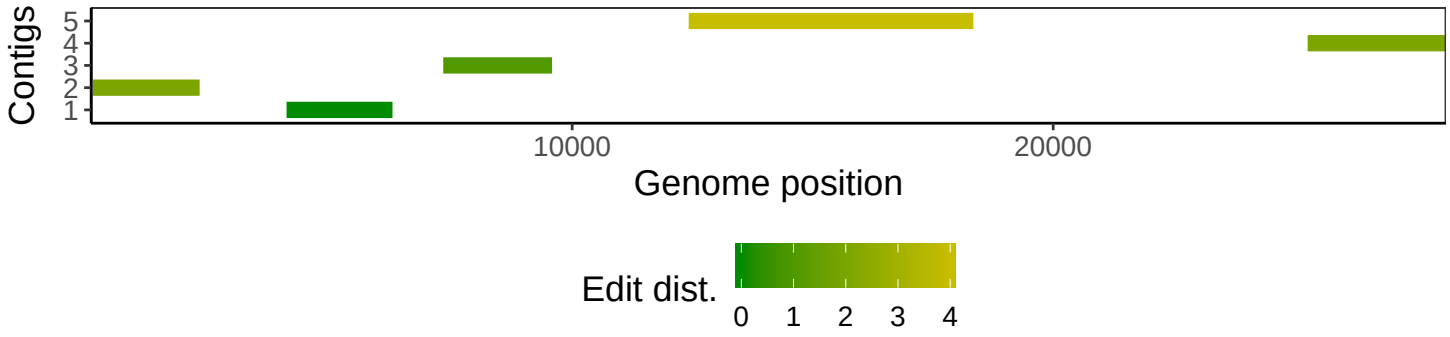
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

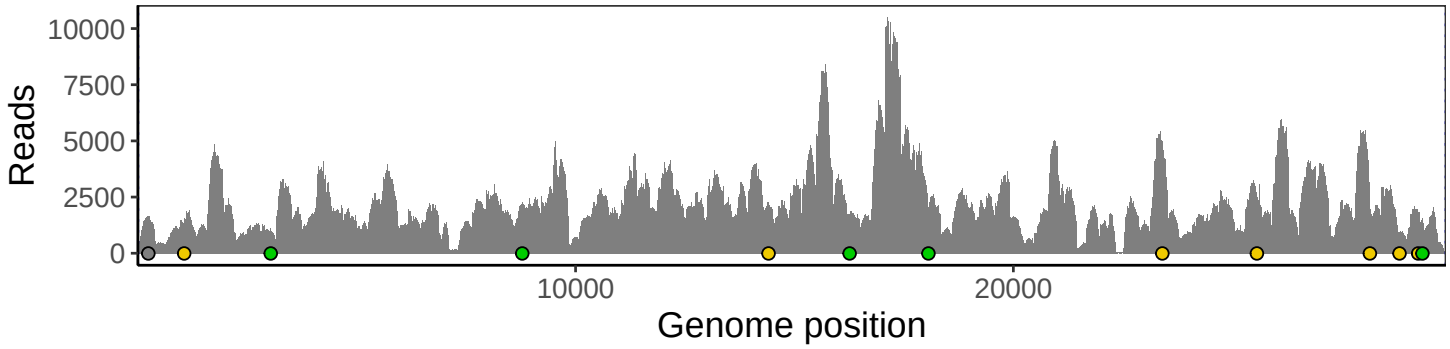


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

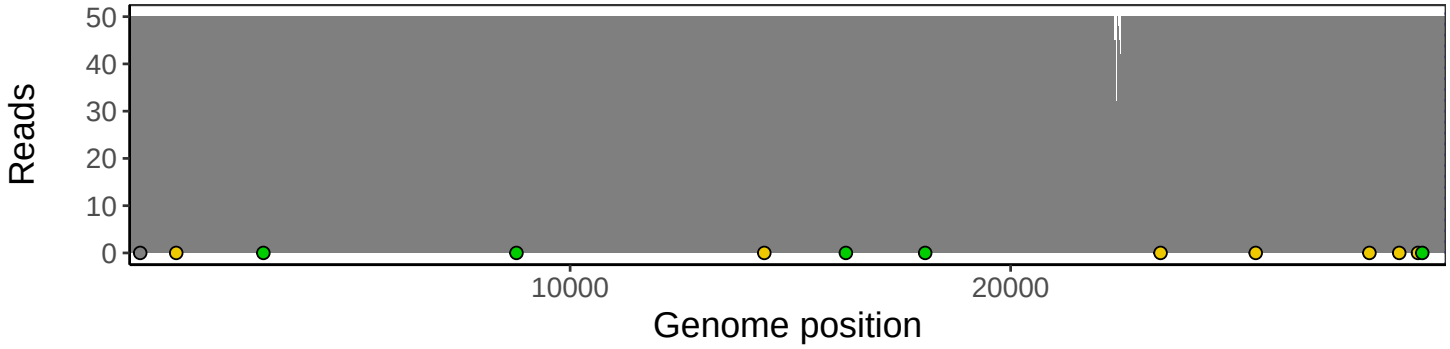


VSP0202-1 | 2020-06-12 | ETA | 269e-q | 16300 genomes | single experiment

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.