COVID-19 subject 211

2020-08-28

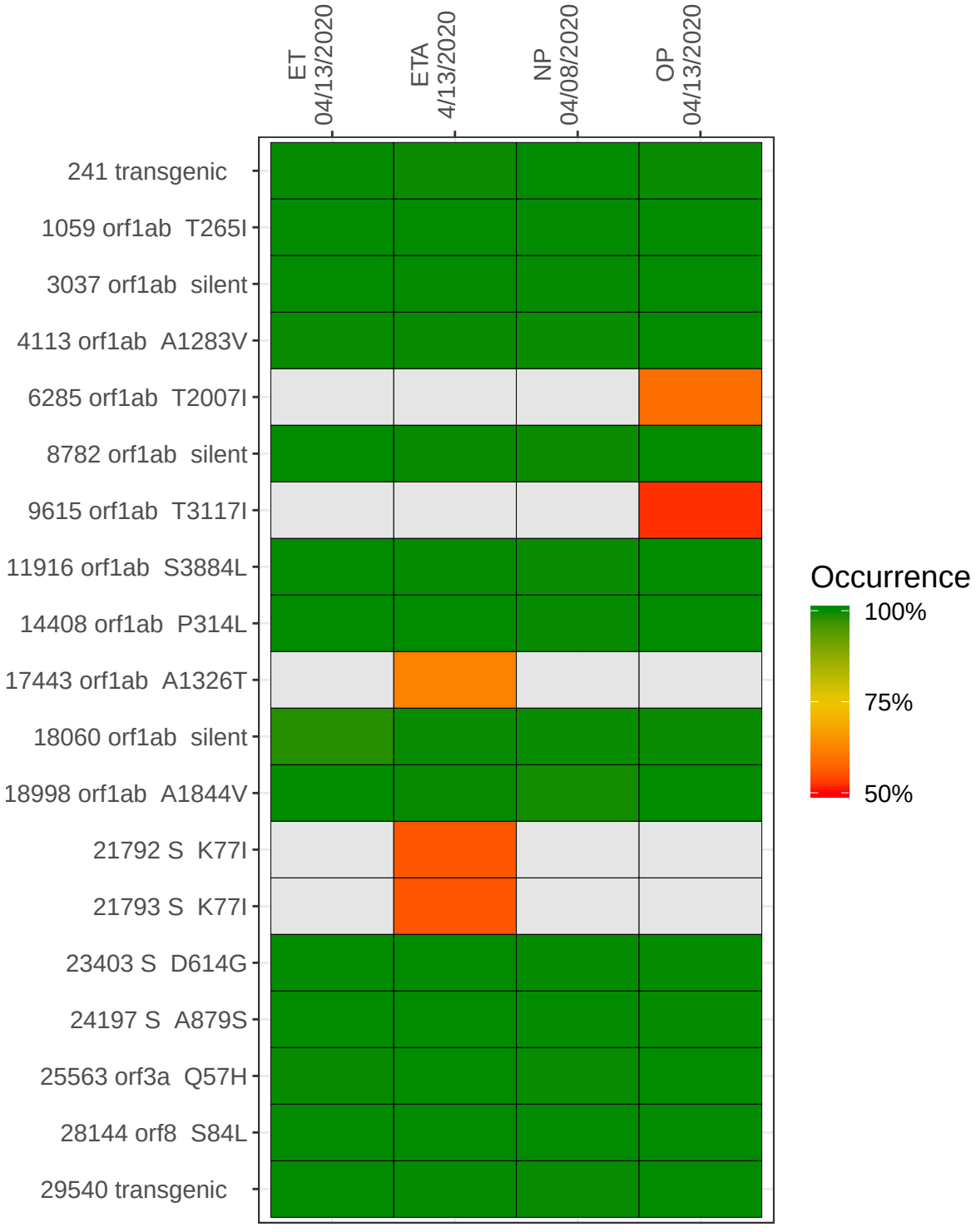
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

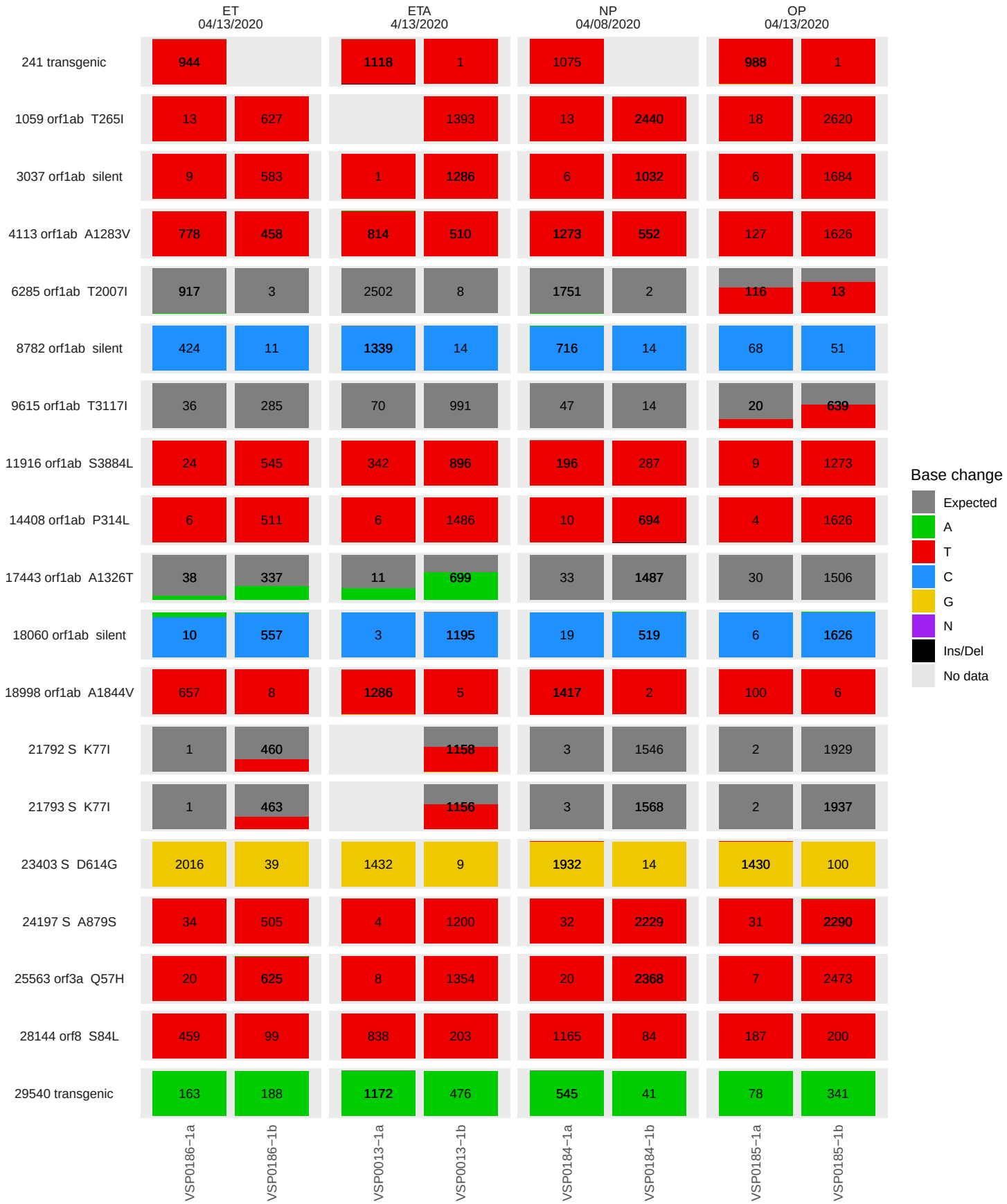
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0013	composite	NA	ETA	4/13/2020	29.90	99.8%	99.8%
VSP0184	composite	NA	NP	04/08/2020	29.86	99.9%	99.8%
VSP0185	composite	NA	OP	04/13/2020	29.88	99.9%	99.8%
VSP0186	composite	NA	ET	04/13/2020	29.80	99.9%	99.8%
VSP0013-1a	single experiment	99600000	ETA	4/13/2020	2.82	92.9%	76.1%
VSP0013-1b	single experiment	99600000	ETA	4/13/2020	3.75	97.5%	83.0%
VSP0184-1a	single experiment	NA	NP	04/08/2020	6.17	99.0%	91.4%
VSP0184-1b	single experiment	NA	NP	04/08/2020	4.78	93.0%	79.9%
VSP0185-1a	single experiment	NA	OP	04/13/2020	5.32	98.0%	86.8%
VSP0185-1b	single experiment	NA	OP	04/13/2020	18.45	98.9%	94.5%
VSP0186-1a	single experiment	NA	ET	04/13/2020	4.83	98.6%	87.7%
VSP0186-1b	single experiment	NA	ET	04/13/2020	7.96	97.3%	89.4%

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.

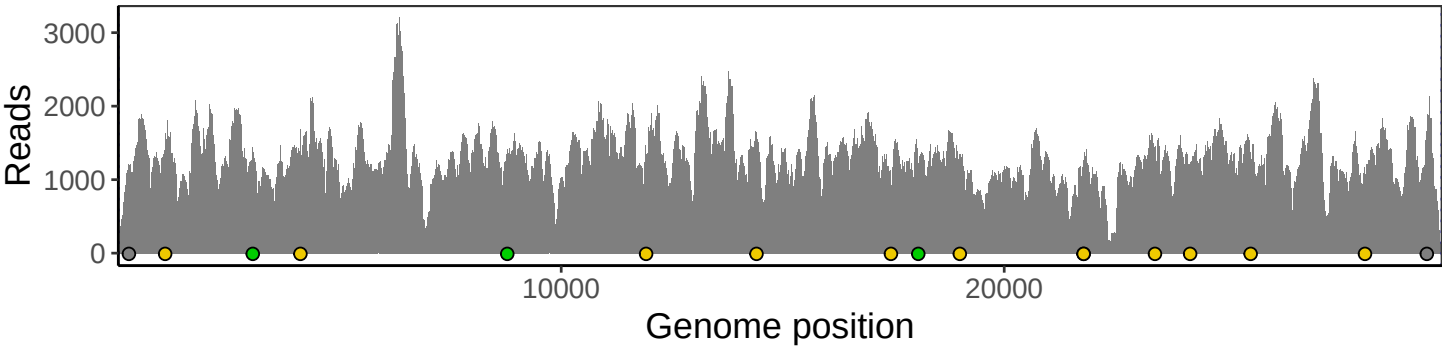



ET
04/13/2020
ETA
4/13/2020
NP
04/08/2020
OP
04/13/2020
241 transgenic
1059 orf1ab T265I
3037 orf1ab silent
4113 orf1ab A1283V
6285 orf1ab T2007I
8782 orf1ab silent
9615 orf1ab T3117I
11916 orf1ab S3884L
14408 orf1ab P314L
17443 orf1ab A1326T
18060 orf1ab silent
18998 orf1ab A1844V
21792 S K77I
21793 S K77I
23403 S D614G
24197 S A879S
25563 orf3a Q57H
28144 orf8 S84L
29540 transgenic
944
1118
1
1075
988
1
13
627
1393
13
2440
18
2620
9
583
1
1286
6
1032
6
1684
778
458
814
510
1273
552
127
1626
917
3
2502
8
1751
2
116
13
424
11
1339
14
716
14
68
51
36
285
70
991
47
14
20
639
24
545
342
896
196
287
9
1273
6
511
6
1486
10
694
4
1626
38
337
11
699
33
1487
30
1506
10
557
3
1195
19
519
6
1626
657
8
1286
5
1417
2
100
6
1
460
1158
3
1546
2
1929
1
463
1156
3
1568
2
1937
2016
39
1432
9
1932
14
1430
100
34
505
4
1200
32
2229
31
2290
20
625
8
1354
20
2368
7
2473
459
99
838
203
1165
84
187
200
163
188
1172
476
545
41
78
341
VSP0186−1a
VSP0186−1b
VSP0013−1a
VSP0013−1b
VSP0184−1a
VSP0184−1b
VSP0185−1a
VSP0185−1b
Base change
Expected
A
T
C
G
N
Ins/Del
No data

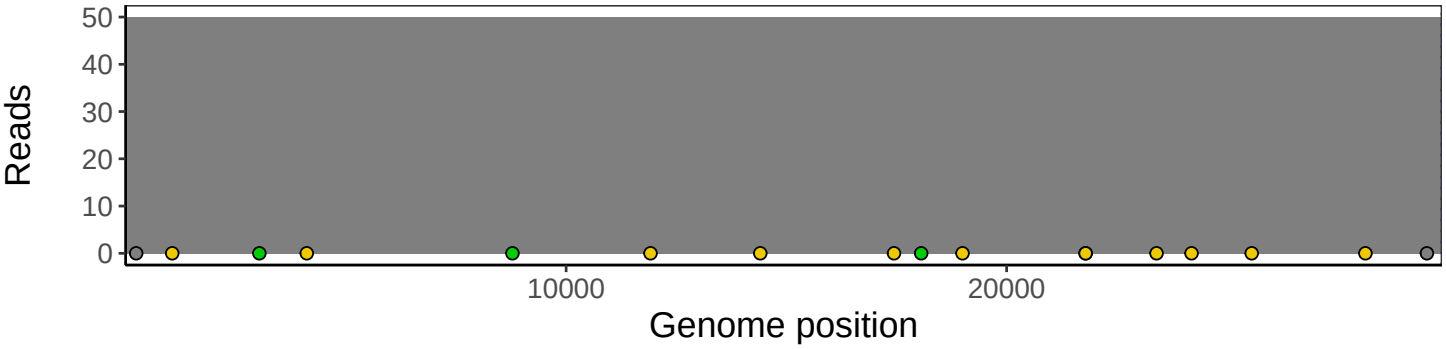
Analyses of individual experiments and composite results.

VSP0013 | 4/13/2020 | ETA | 211-tri | composite result

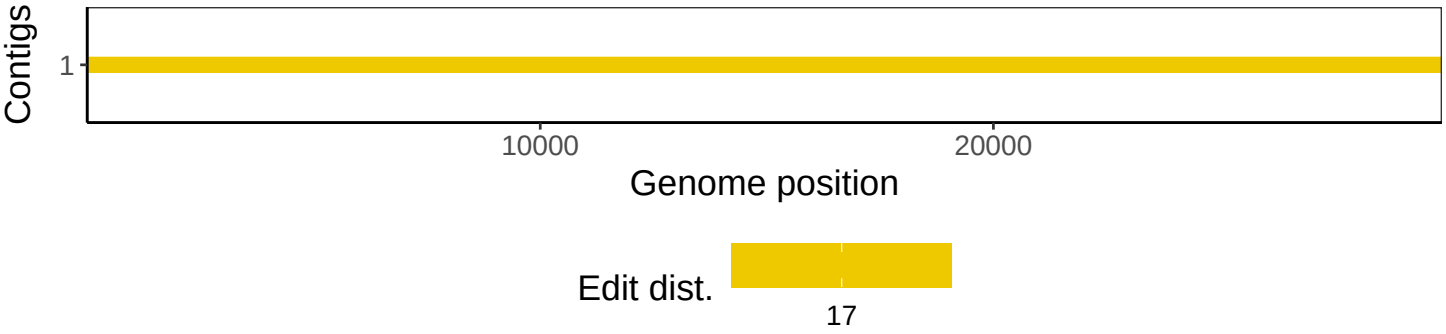
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

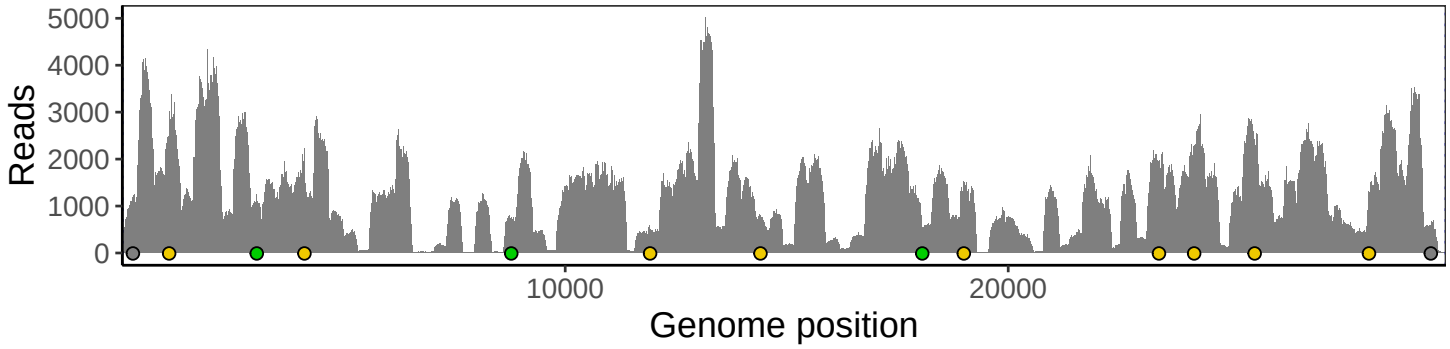


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

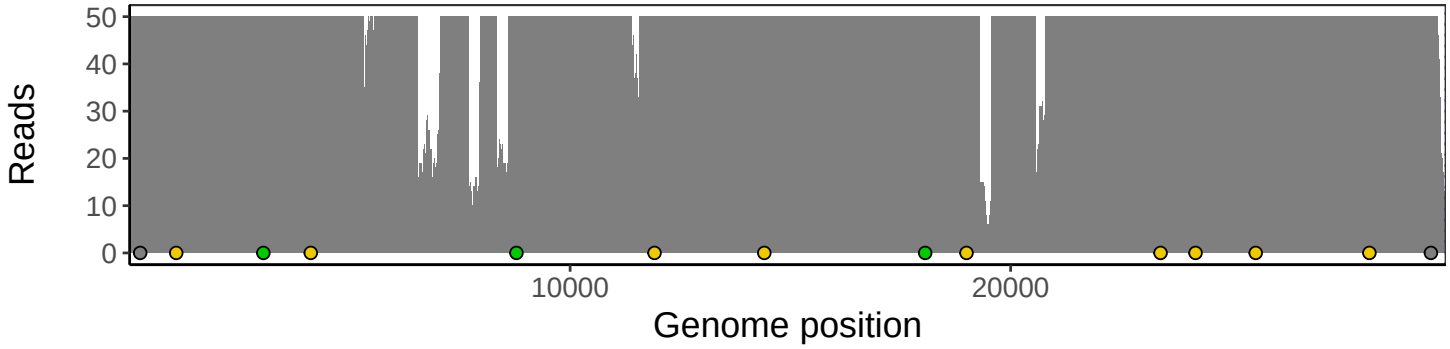


VSP0184 | 04/08/2020 | NP | 5 | composite result

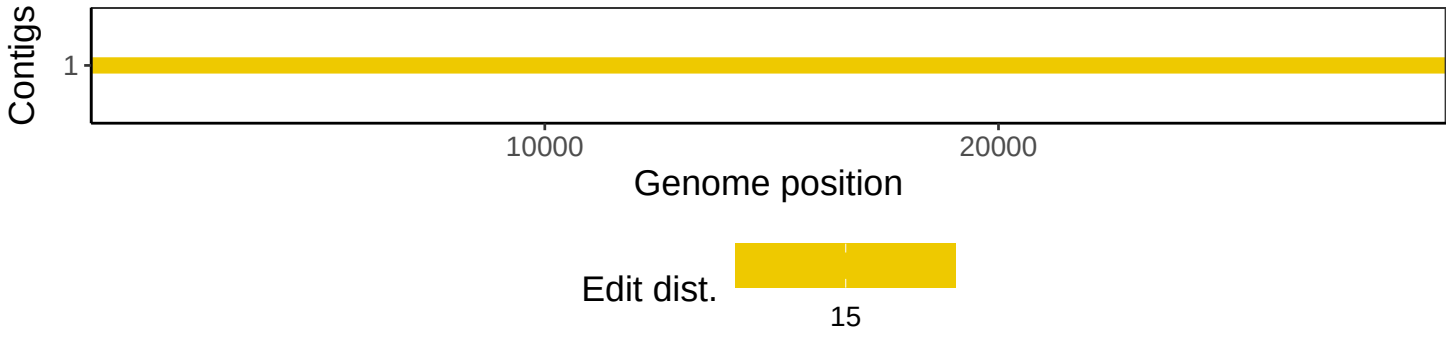
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

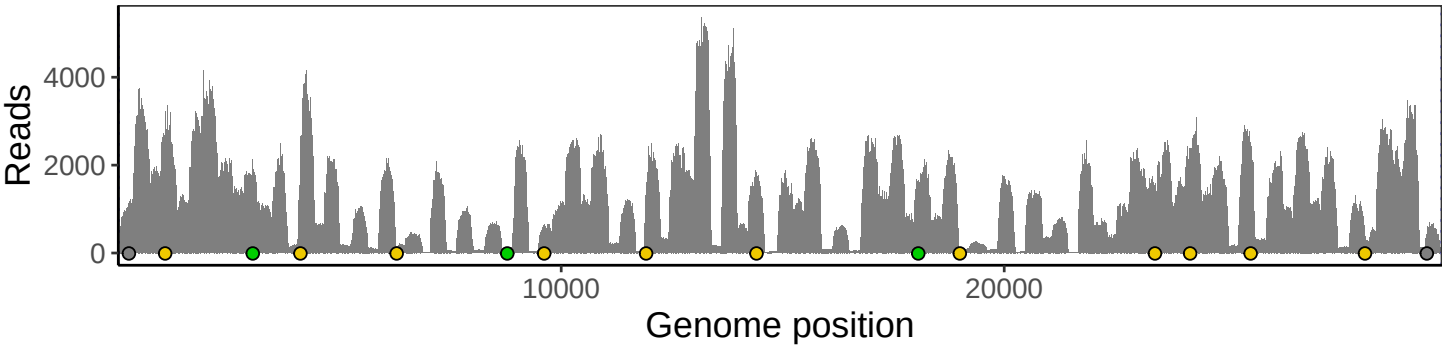


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

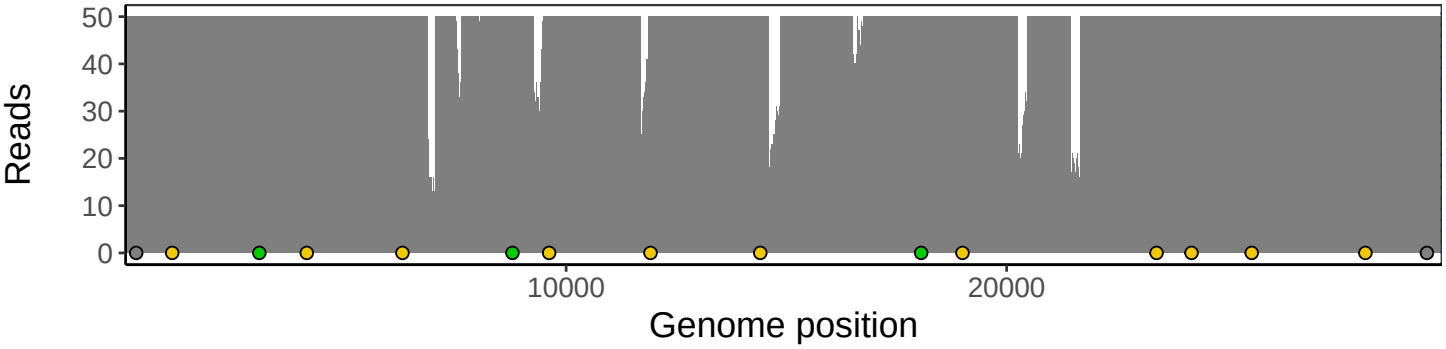


VSP0185 | 04/13/2020 | OP | 6 | composite result

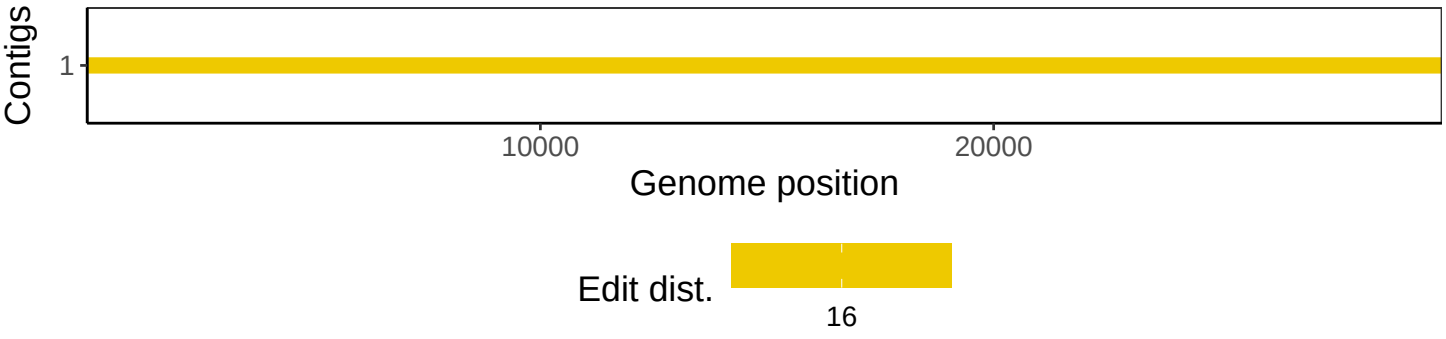
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

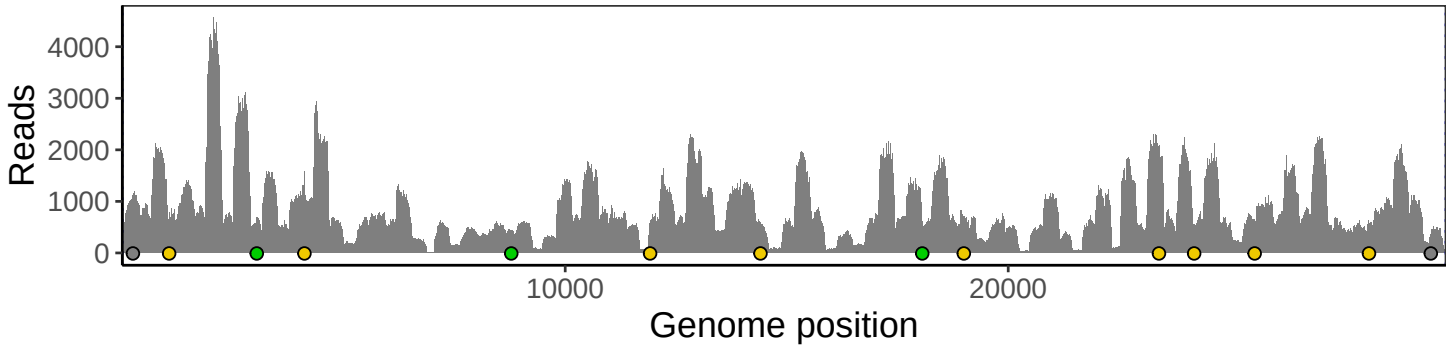


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

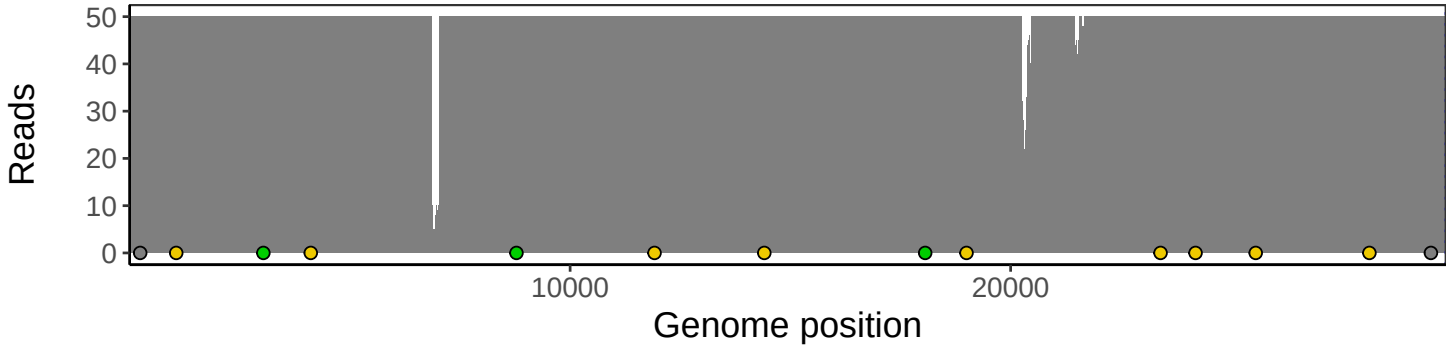


VSP0186 | 04/13/2020 | ET | 7 | composite result

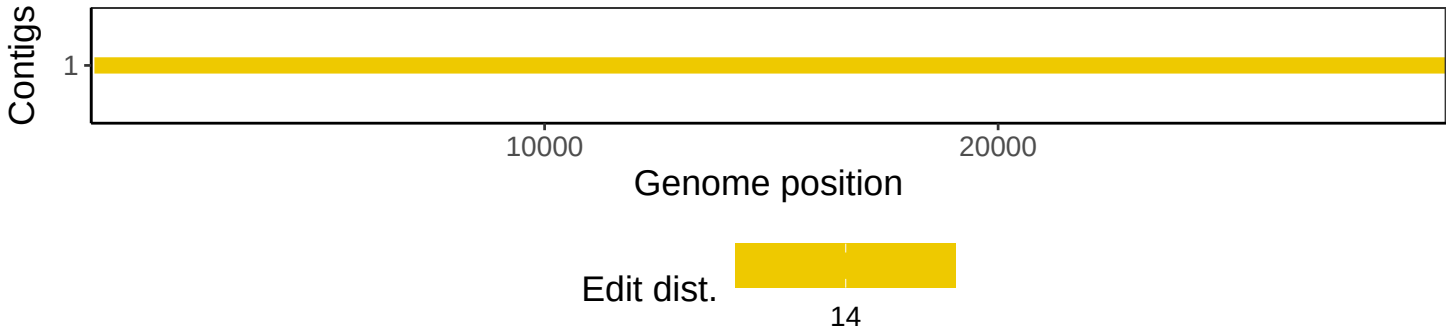
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

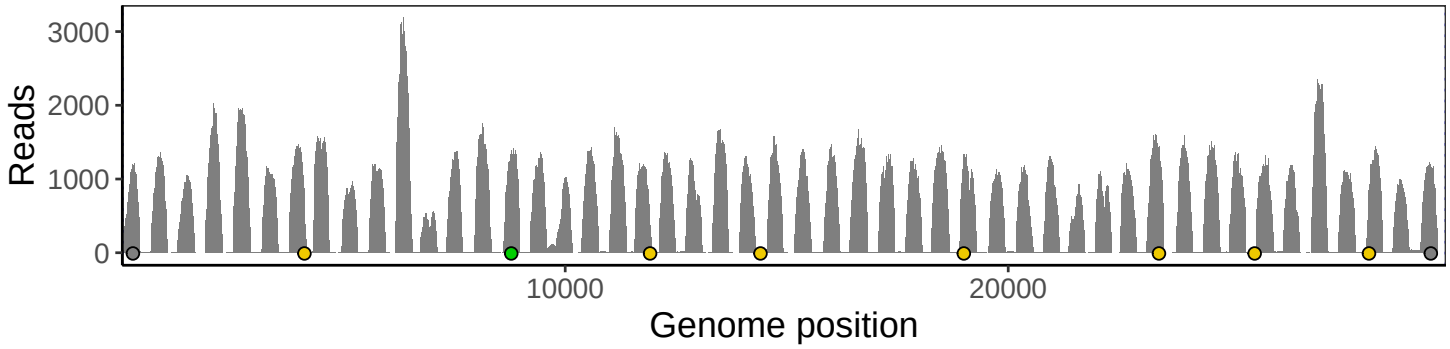


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

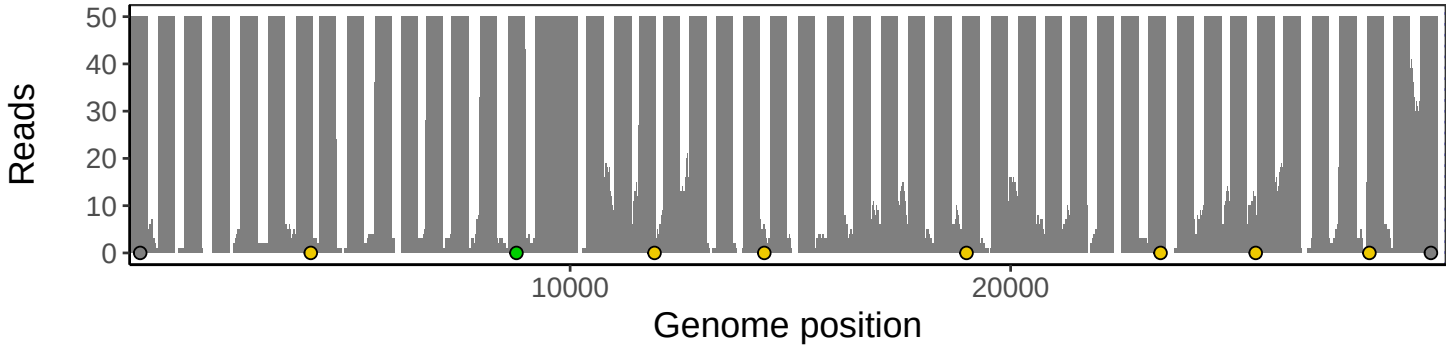


VSP0013-1a | 4/13/2020 | ETA | 211-tri | 99600000 genomes | single experiment

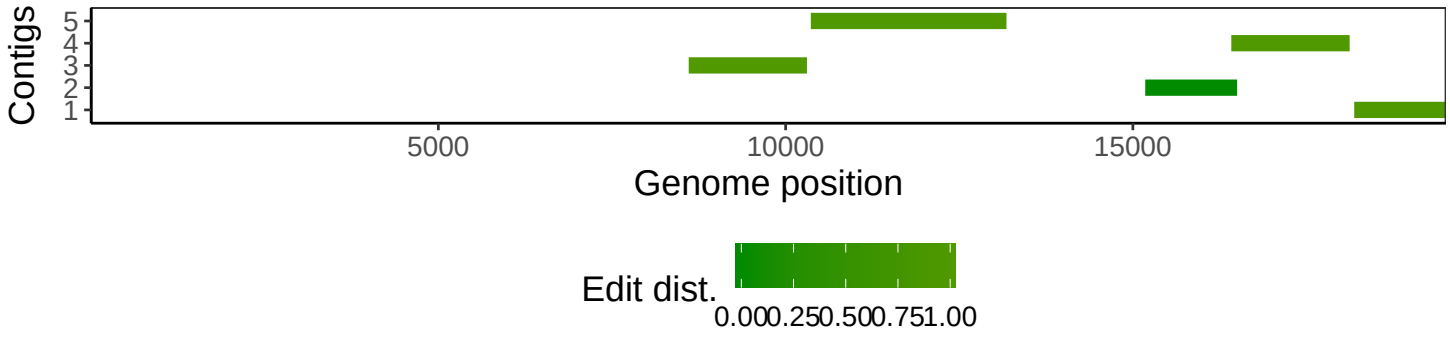
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

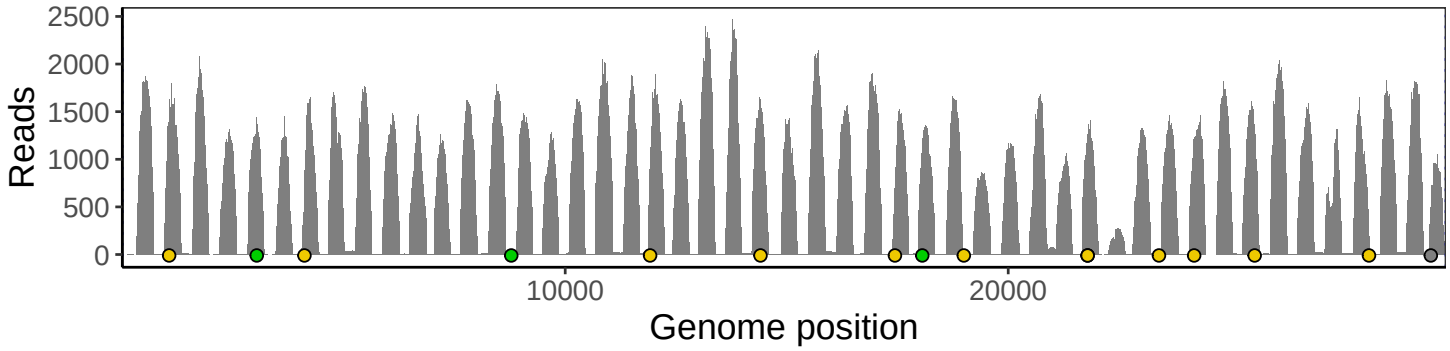


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

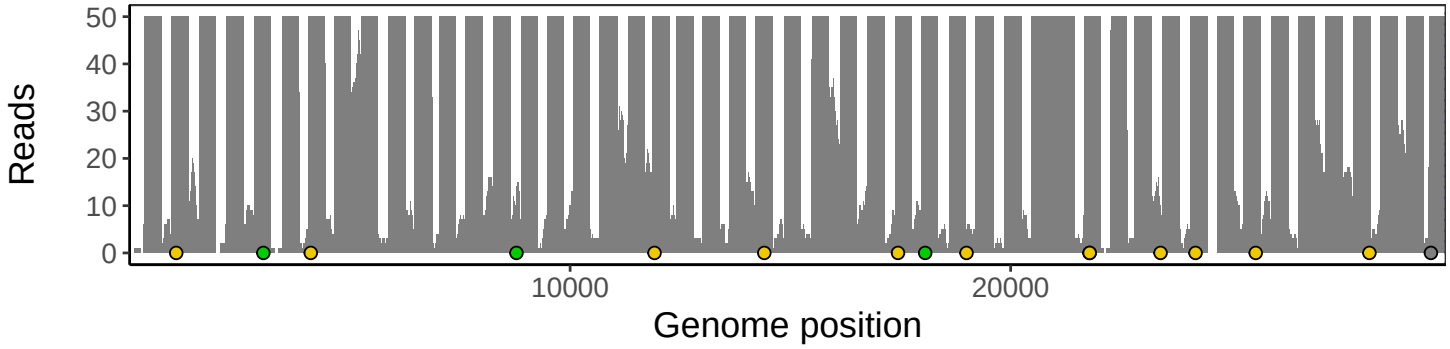


VSP0013-1b | 4/13/2020 | ETA | 211-tri | 99600000 genomes | single experiment

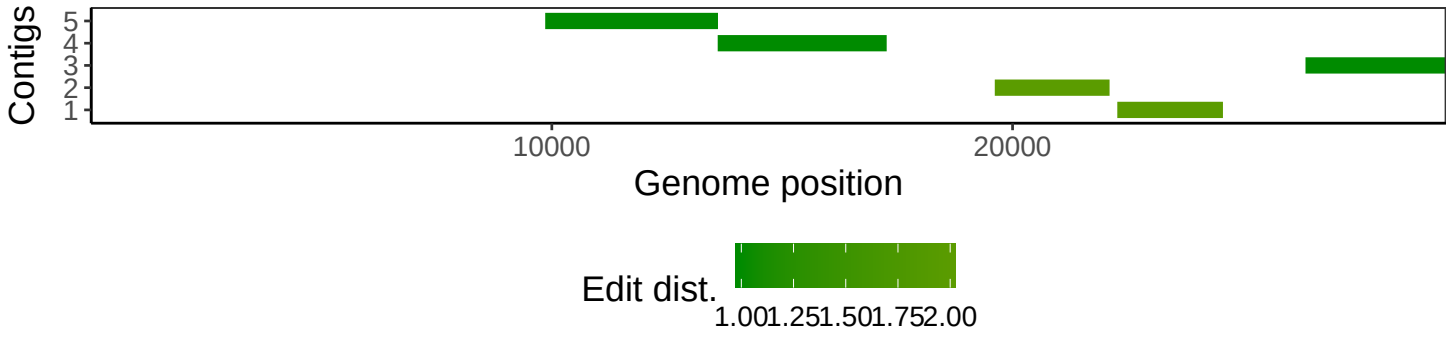
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

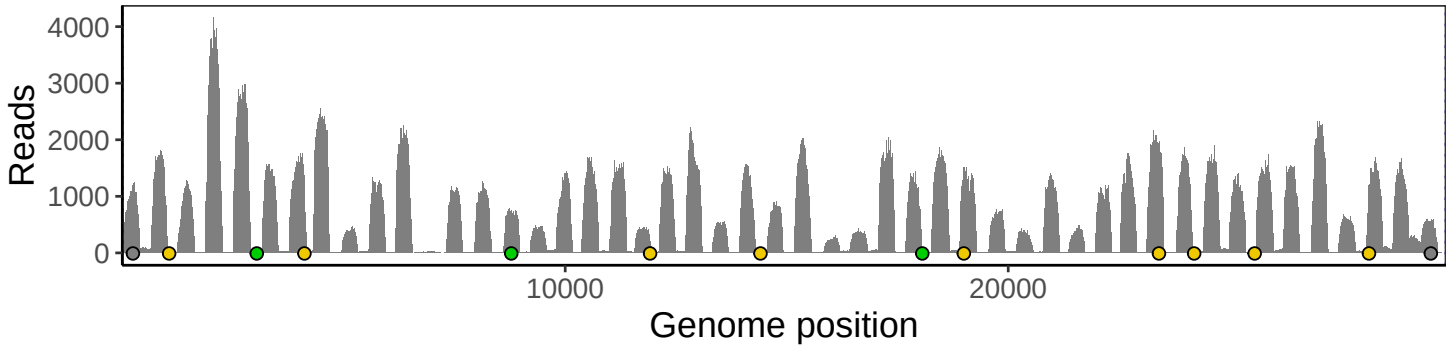


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

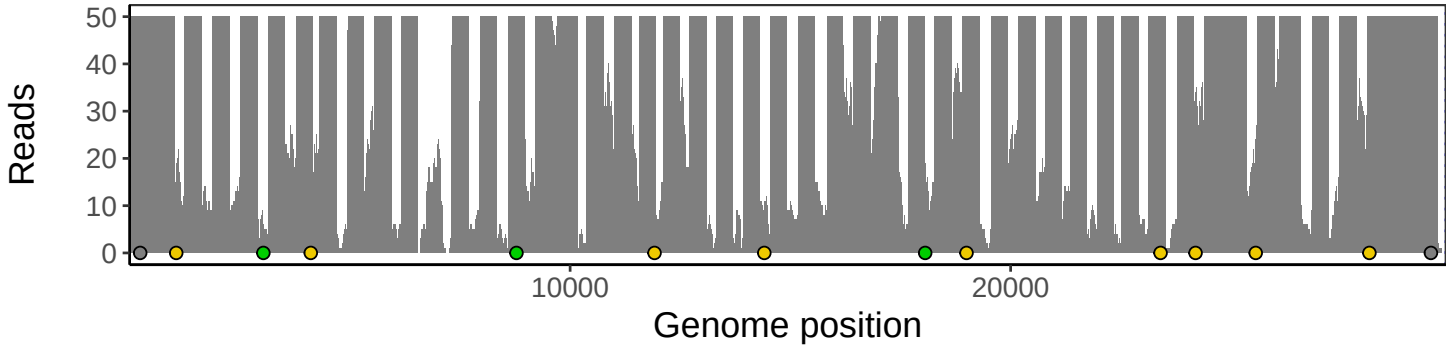


VSP0184-1a | 04/08/2020 | NP | 5 | NA genomes | single experiment

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.



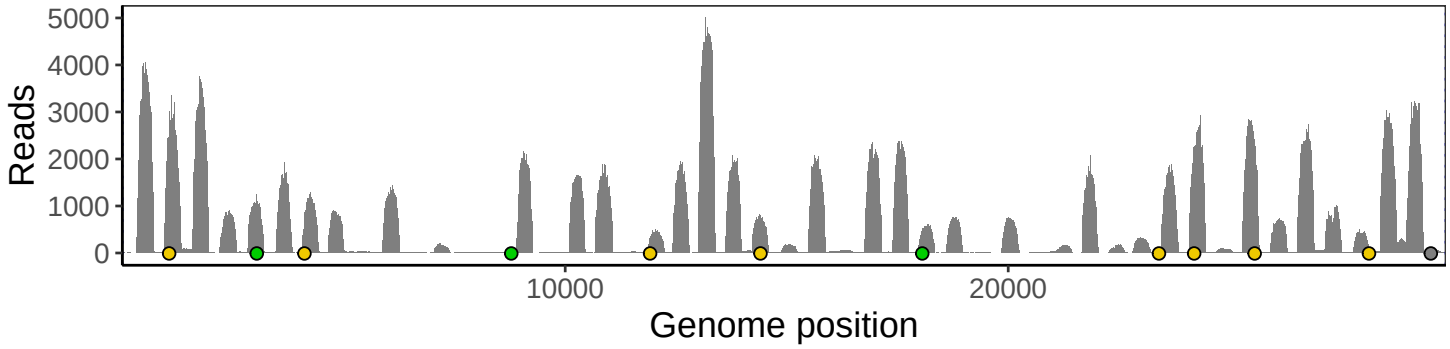
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



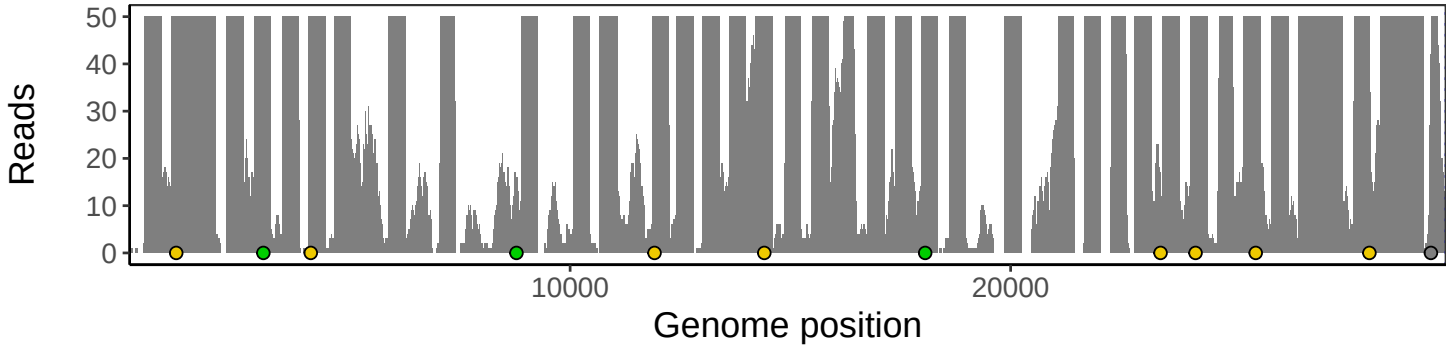
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

VSP0184-1b | 04/08/2020 | NP | 5 | NA genomes | single experiment

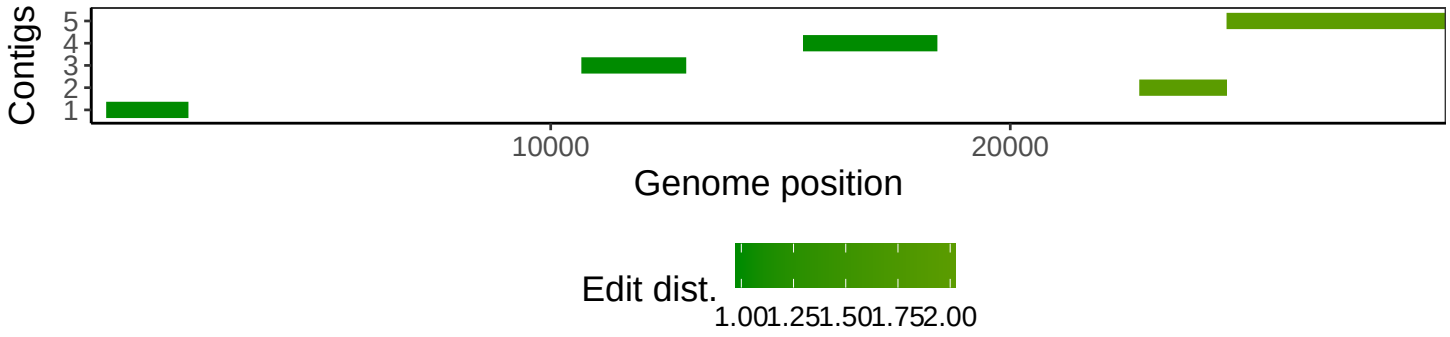
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

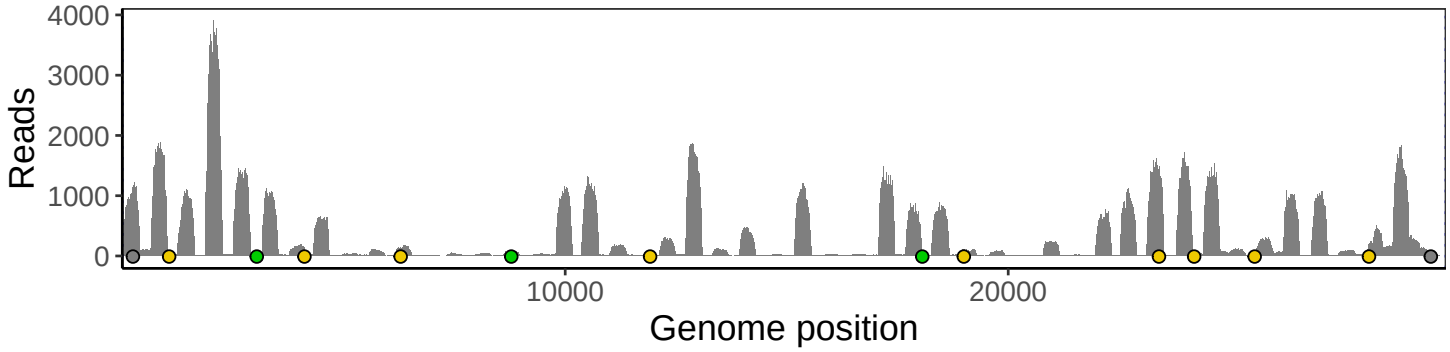


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

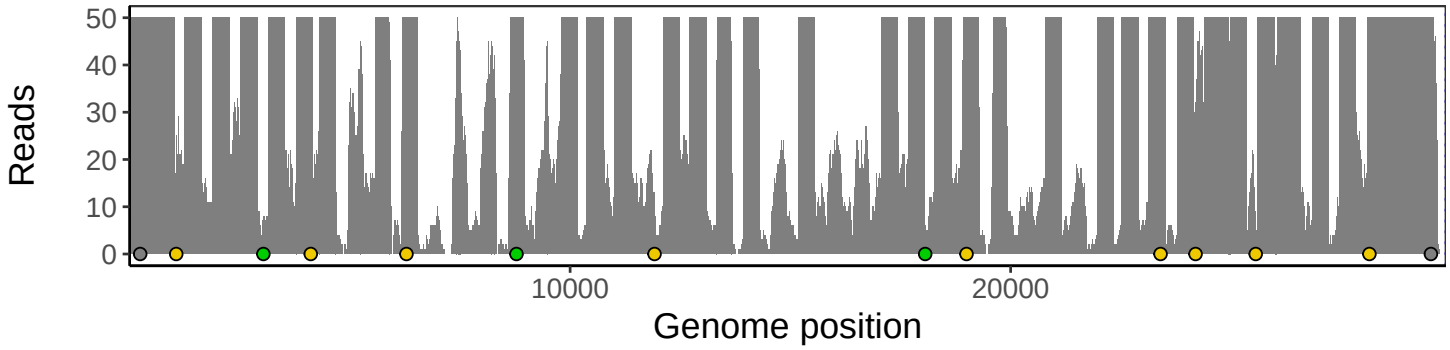


VSP0185-1a | 04/13/2020 | OP | 6 | NA genomes | single experiment

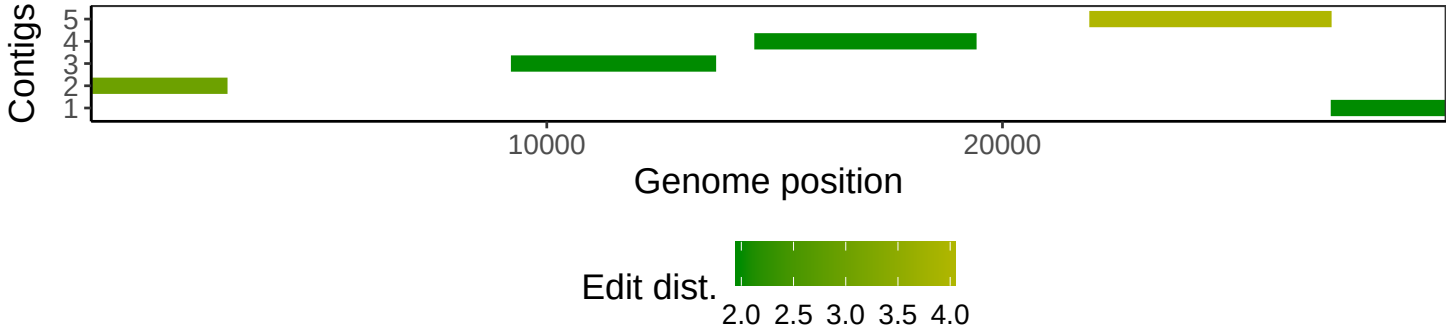
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

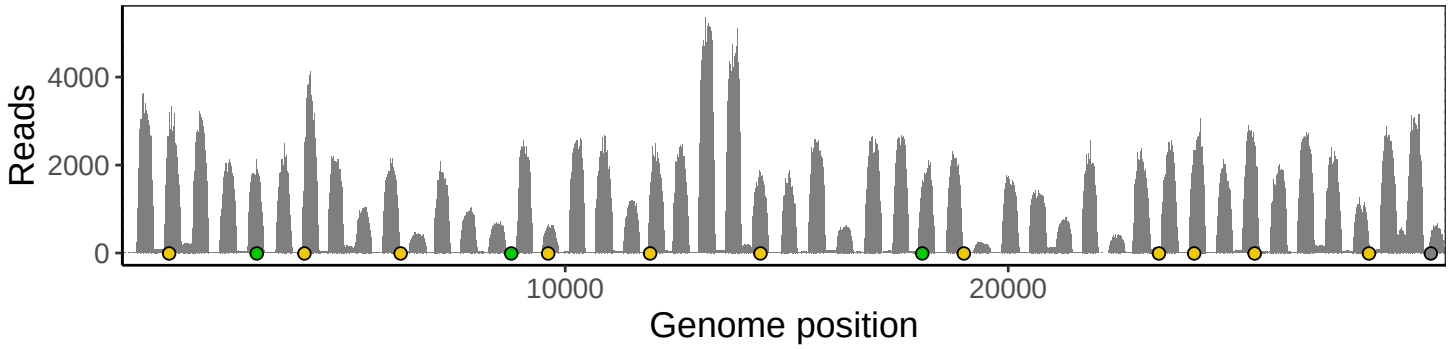


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

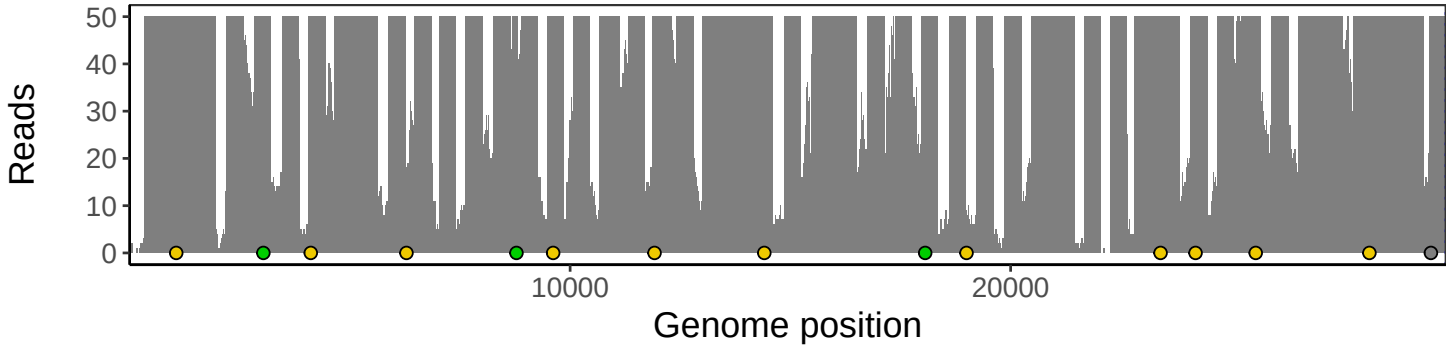


VSP0185-1b | 04/13/2020 | OP | 6 | NA genomes | single experiment

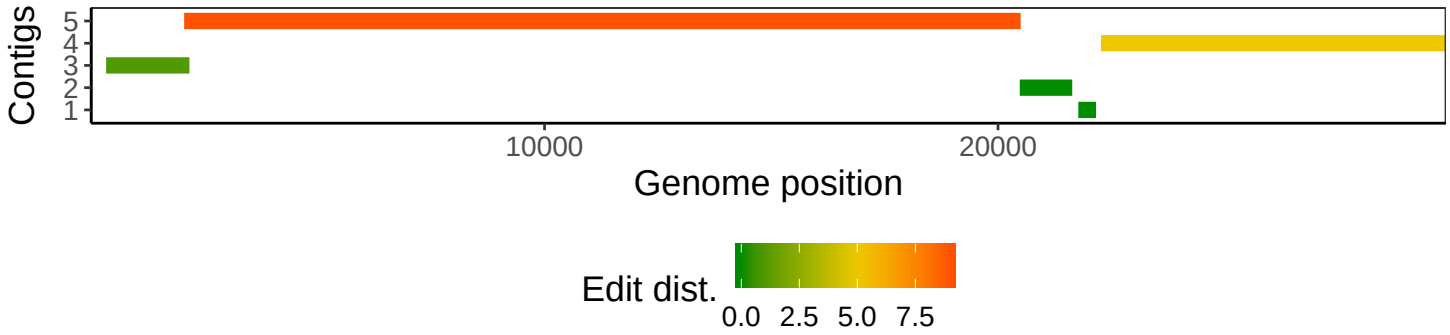
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

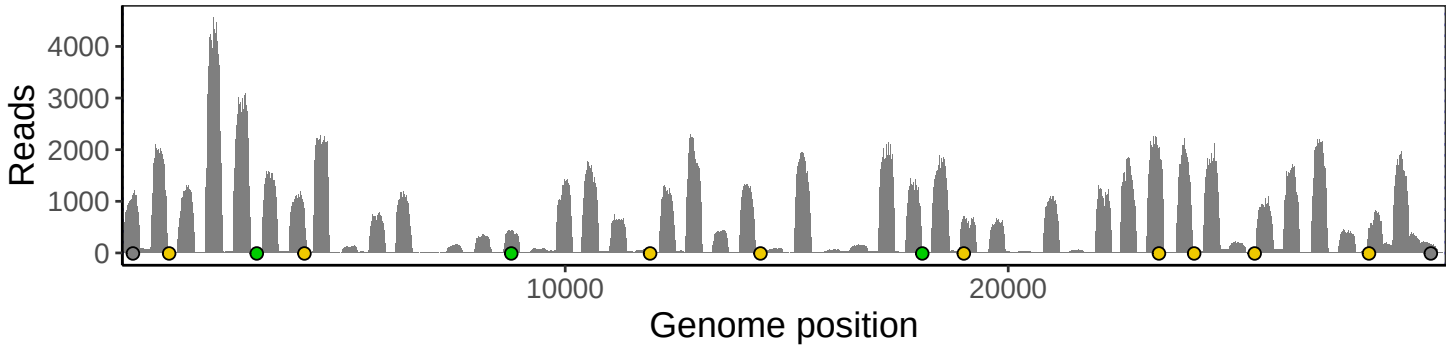


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

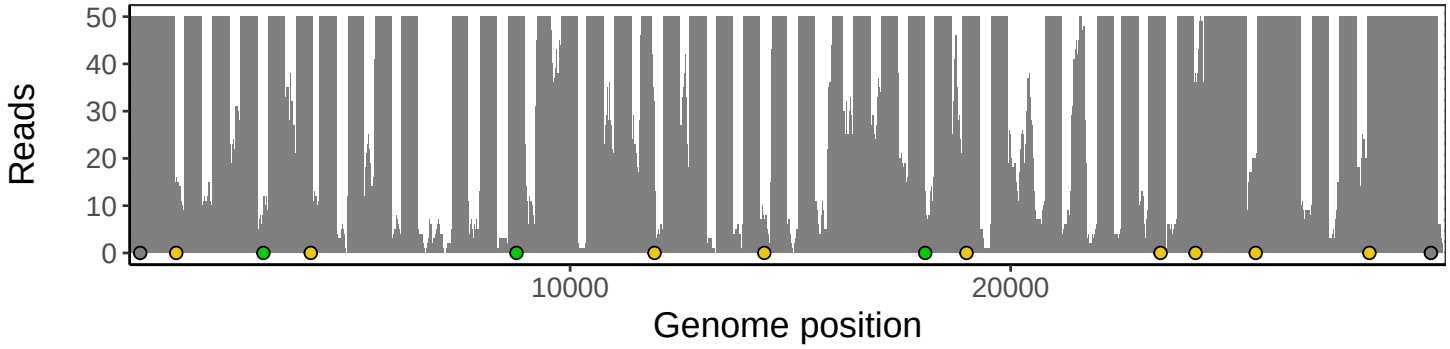


VSP0186-1a | 04/13/2020 | ET | 7 | NA genomes | single experiment

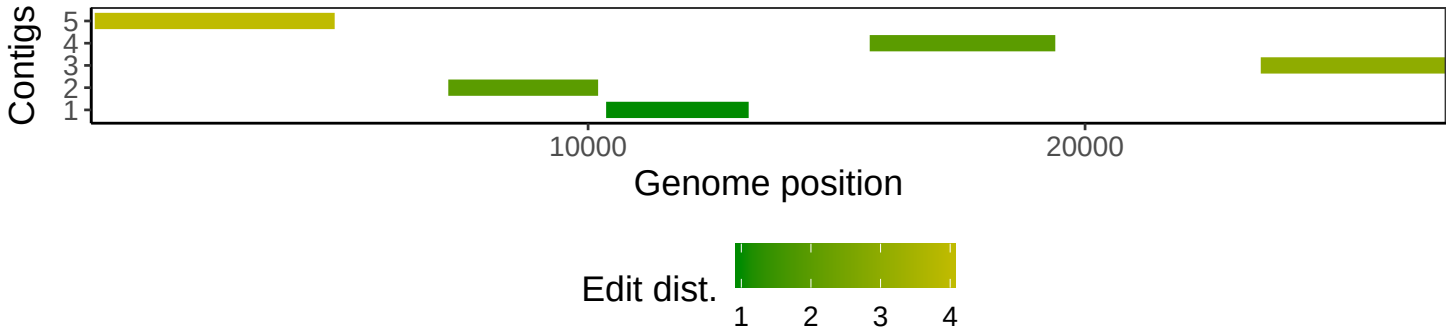
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

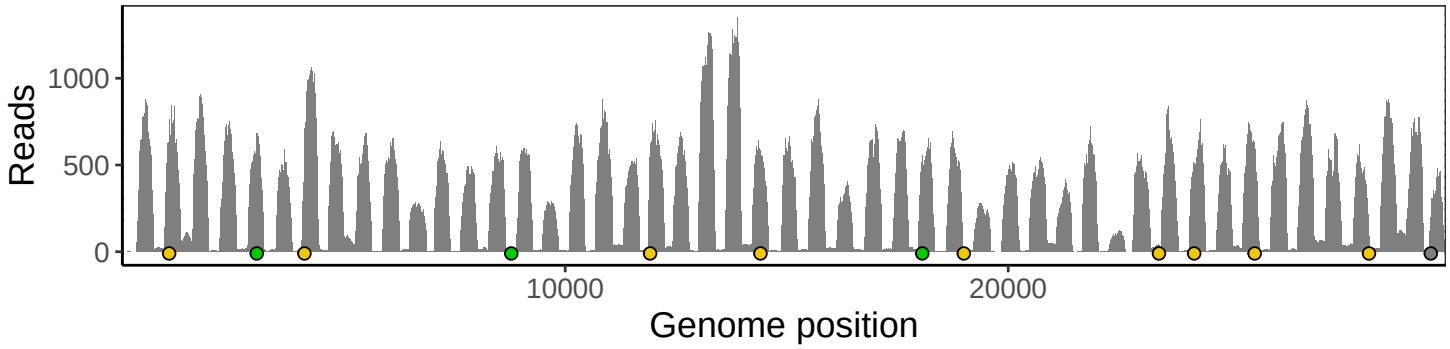


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

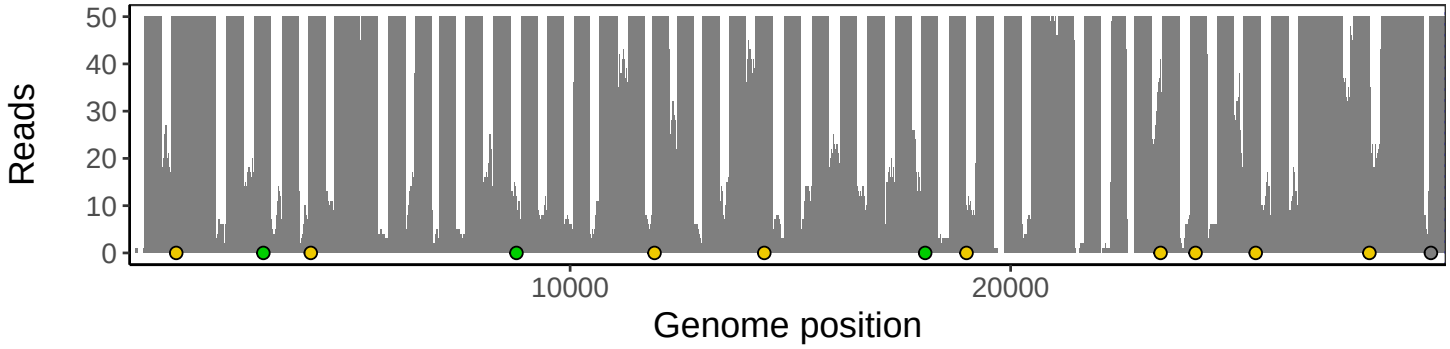


VSP0186-1b | 04/13/2020 | ET | 7 | NA genomes | single experiment

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.