COVID-19 subject 242

2020-11-30

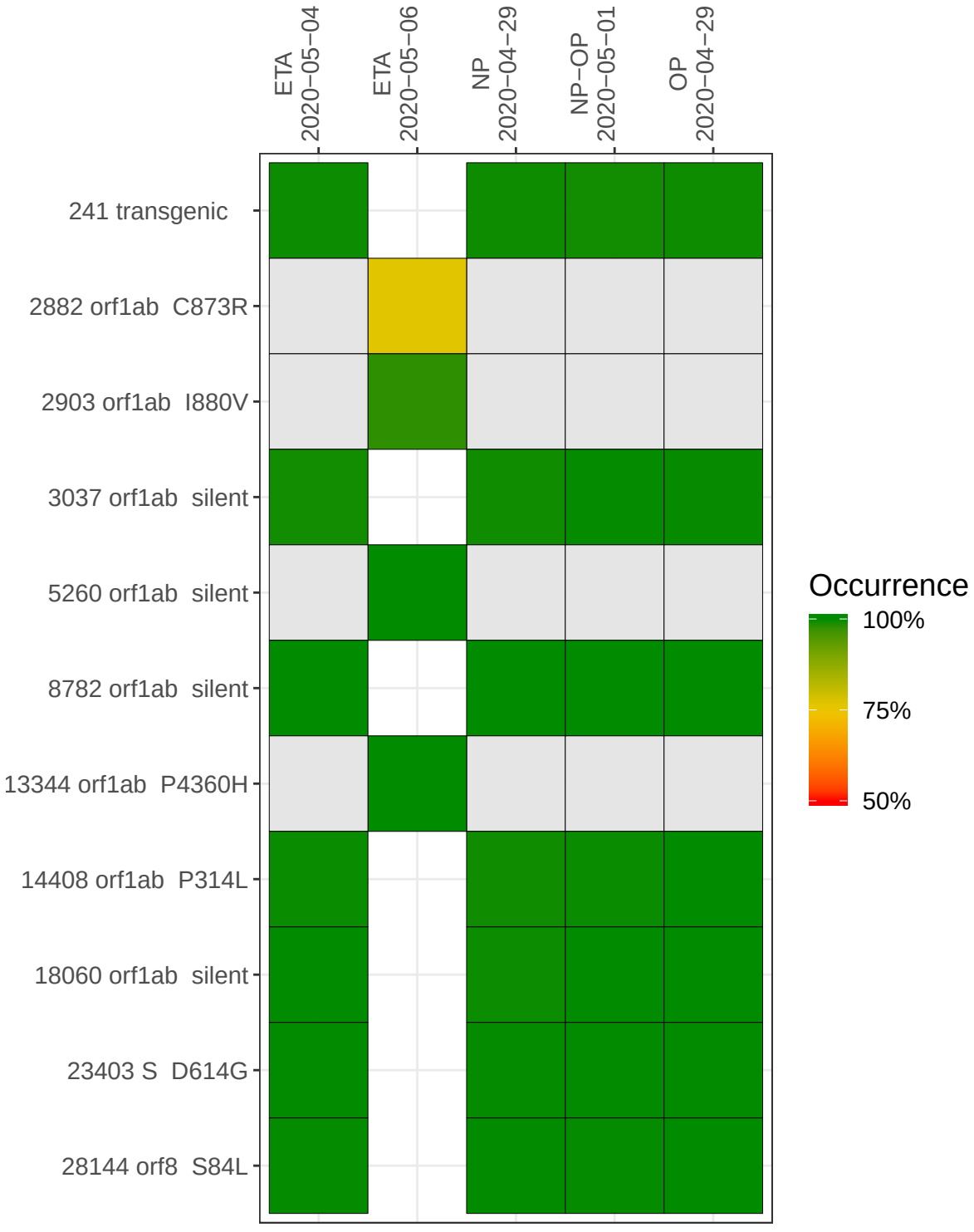
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

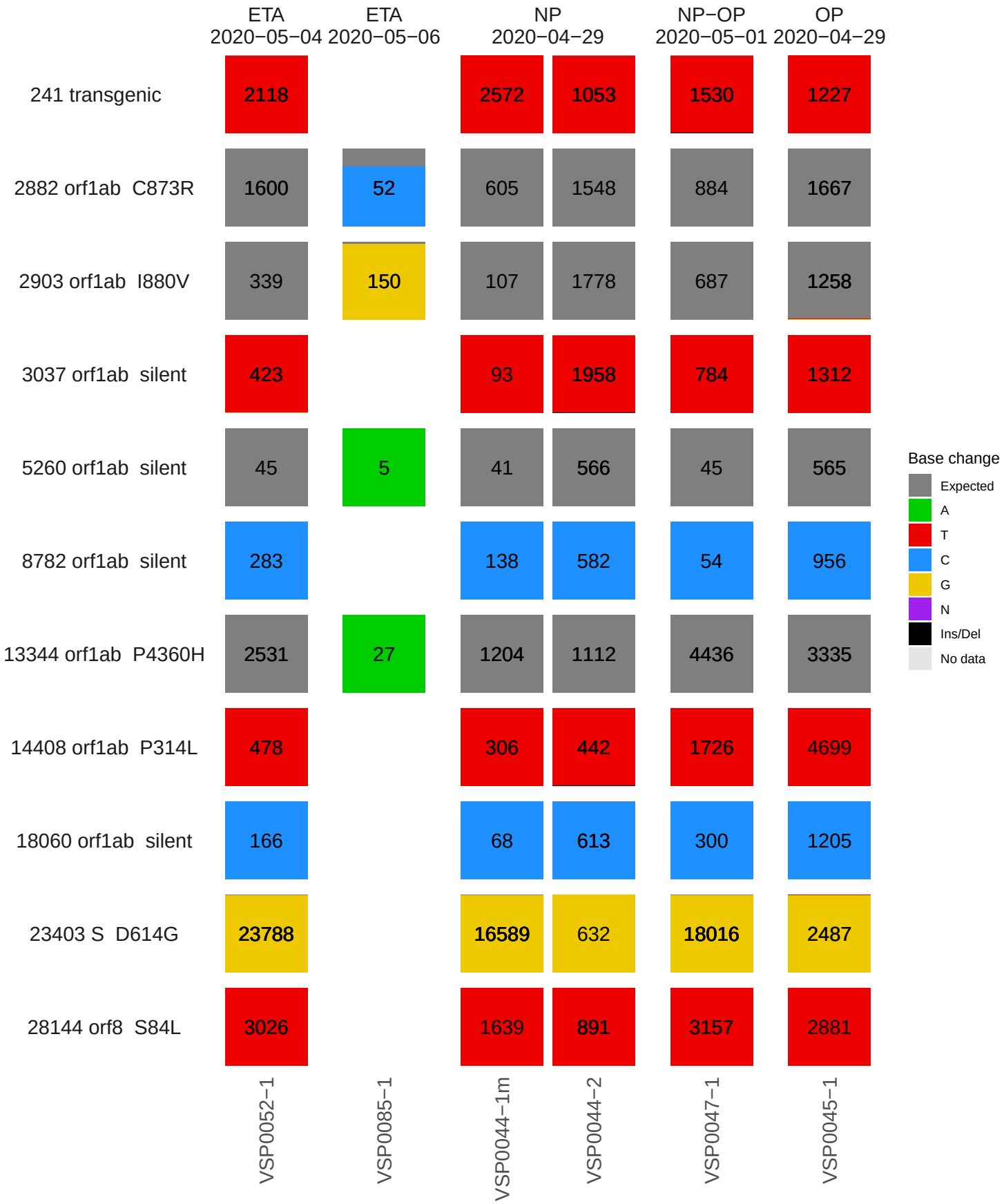
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0044	composite	NA	NP	2020-04-29	29.88	99.9%	99.8%
VSP0085	composite	NA	ETA	2020-05-06	NA	6.4%	1.8%
VSP0044-1m	single experiment	NA	NP	2020-04-29	29.84	99.8%	99.7%
VSP0044-2	single experiment	1.975e+05	NP	2020-04-29	16.60	99.9%	99.7%
VSP0045-1	single experiment	6.670e+02	OP	2020-04-29	29.72	99.6%	99.2%
VSP0047-1	single experiment	2.220e+05	NP-OP	2020-05-01	29.89	99.9%	99.8%
VSP0052-1	single experiment	1.420e+06	ETA	2020-05-04	29.31	99.9%	99.7%
VSP0085-1	single experiment	8.540e+00	ETA	2020-05-06	NA	6.4%	1.8%
VSP0085-2	single experiment	4.270e+01	ETA	2020-05-06	NA	NA	NA
VSP0085-3	single experiment	4.270e+01	ETA	2020-05-06	NA	NA	NA
VSP0085-4	single experiment	4.270e+01	ETA	2020-05-06	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.

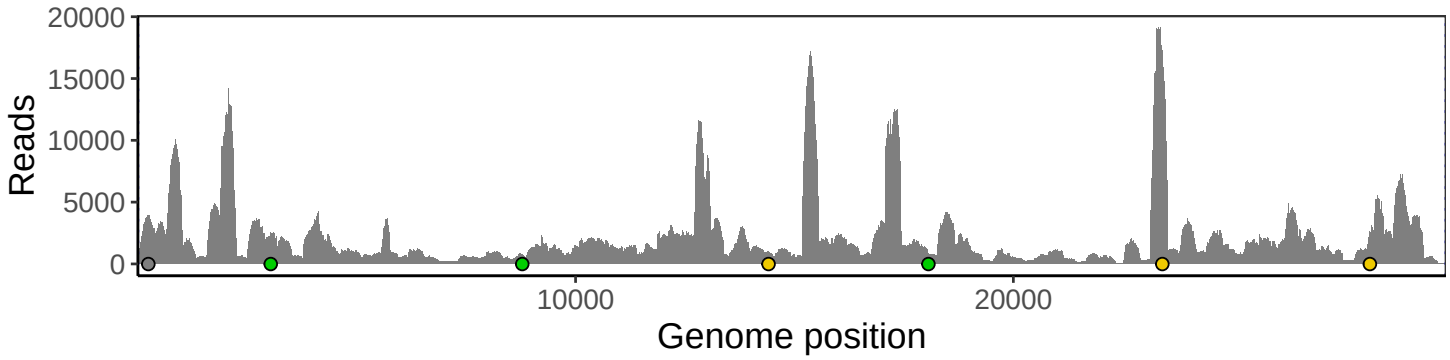



ETA 2020−05−04
ETA 2020−05−06
NP 2020−04−29
NP−OP 2020−05−01
OP 2020−04−29
241 transgenic
2882 orf1ab C873R
2903 orf1ab I880V
3037 orf1ab silent
5260 orf1ab silent
8782 orf1ab silent
13344 orf1ab P4360H
14408 orf1ab P314L
18060 orf1ab silent
23403 S D614G
28144 orf8 S84L
2118
2572
1053
1530
1227
1600
52
605
1548
884
1667
339
150
107
1778
687
1258
423
93
1958
784
1312
45
5
41
566
45
565
283
138
582
54
956
2531
27
1204
1112
4436
3335
478
306
442
1726
4699
166
68
613
300
1205
23788
16589
632
18016
2487
3026
1639
891
3157
2881
VSP0052−1
VSP0085−1
VSP0044−1m
VSP0044−2
VSP0047−1
VSP0045−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

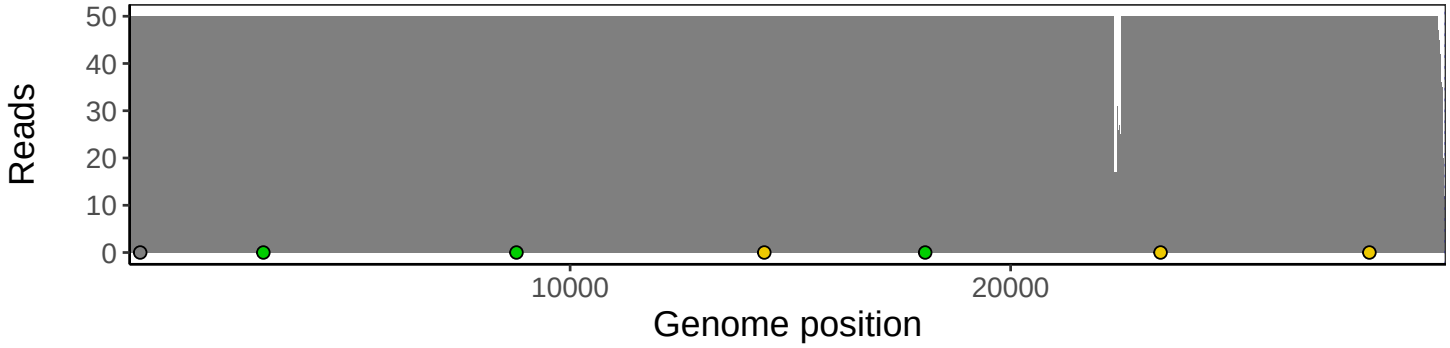
Analyses of individual experiments and composite results.

VSP0044 | 2020-04-29 | NP | 242n-tri | composite result

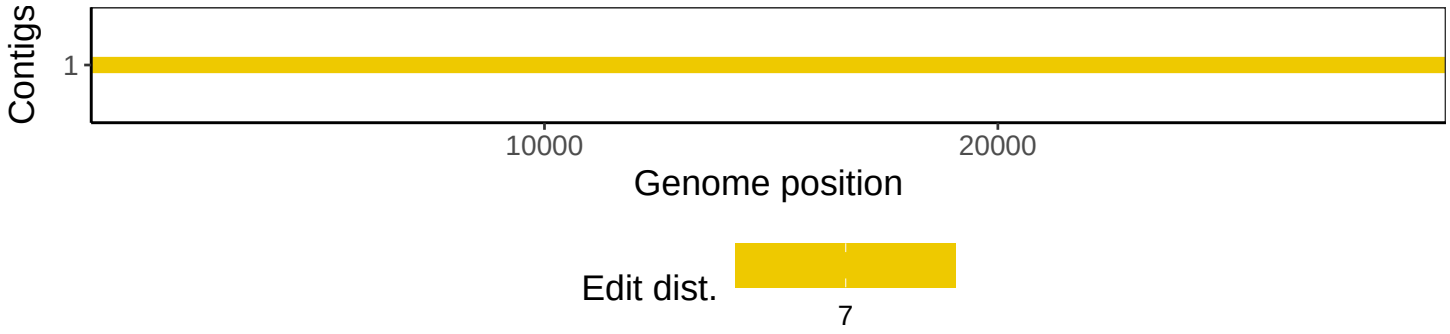
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

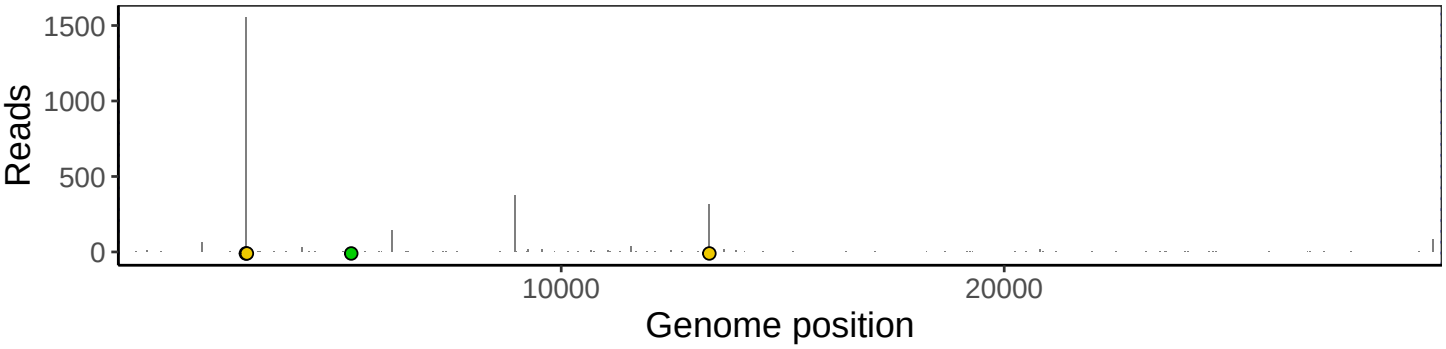


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

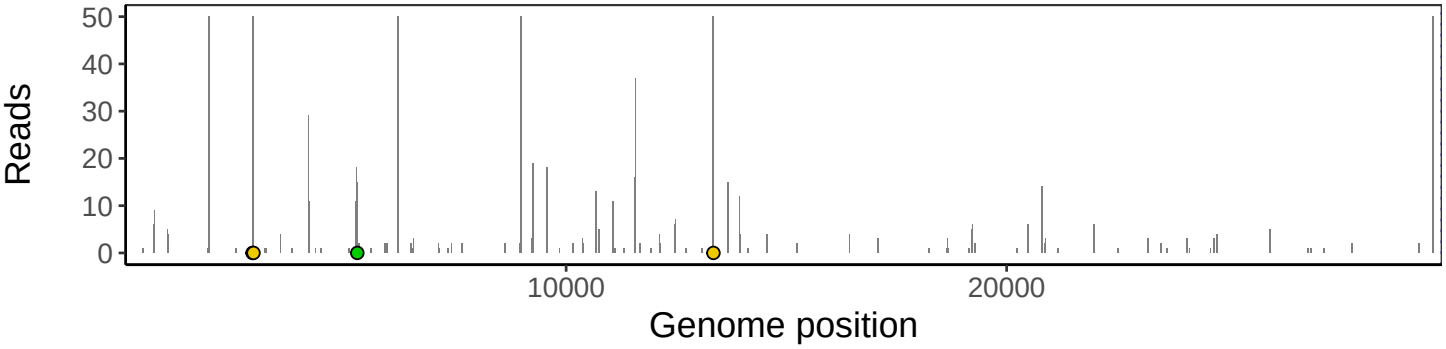


VSP0085 | 2020-05-06 | ETA | 242e-q | composite result

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.



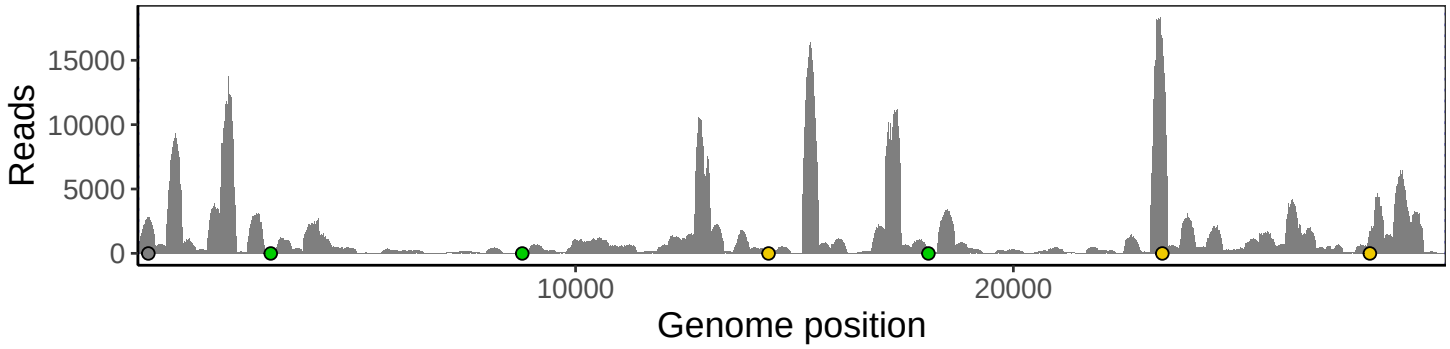
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



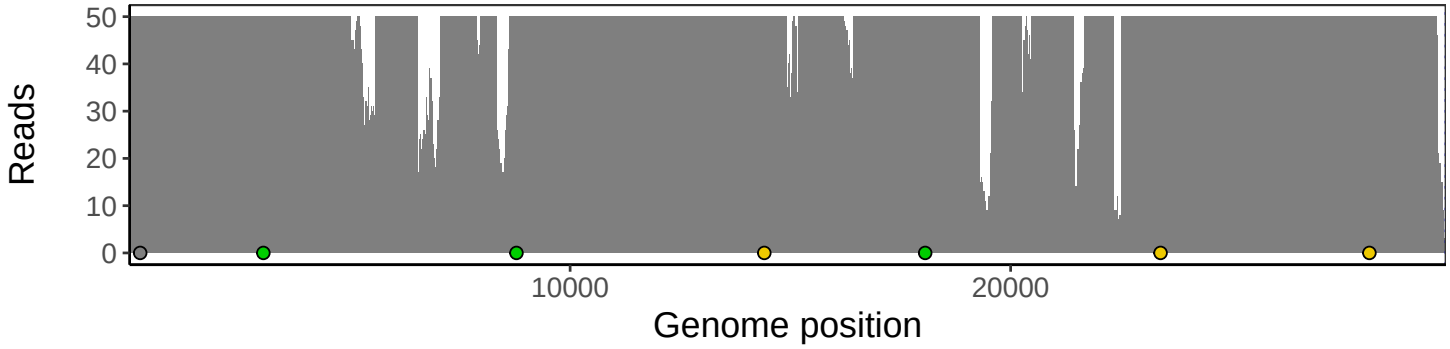
No contig data available.

VSP0044-1m | 2020-04-29 | NP | 242n-tri | genomes | single experiment

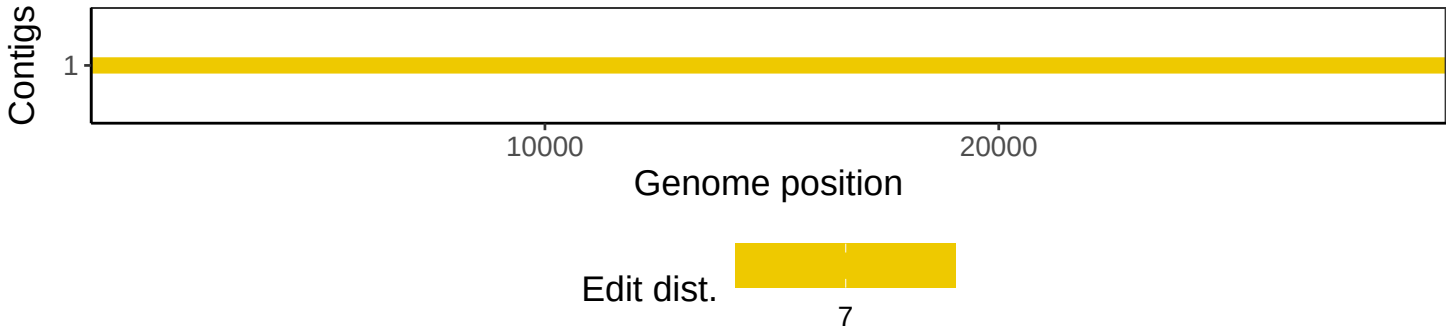
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

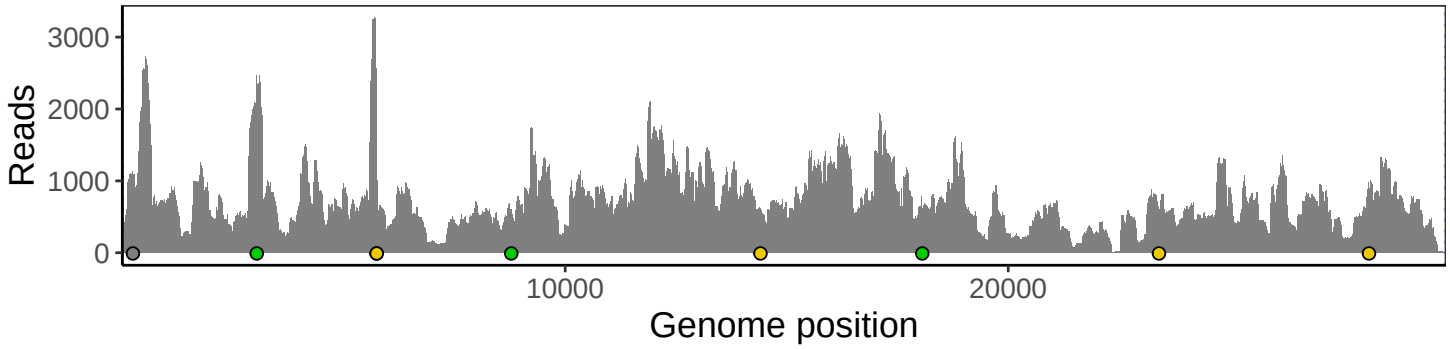


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

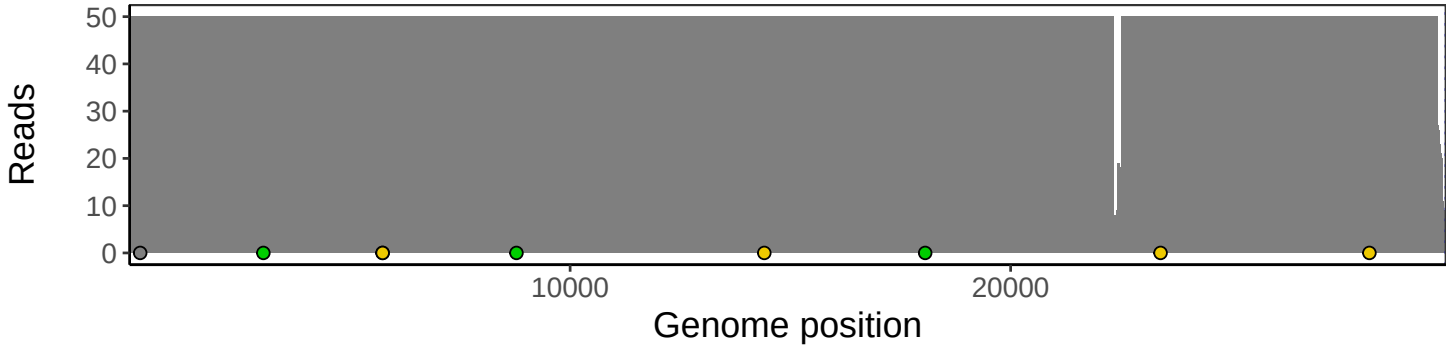


VSP0044-2 | 2020-04-29 | NP | 242n-tri | 197500 genomes | single experiment

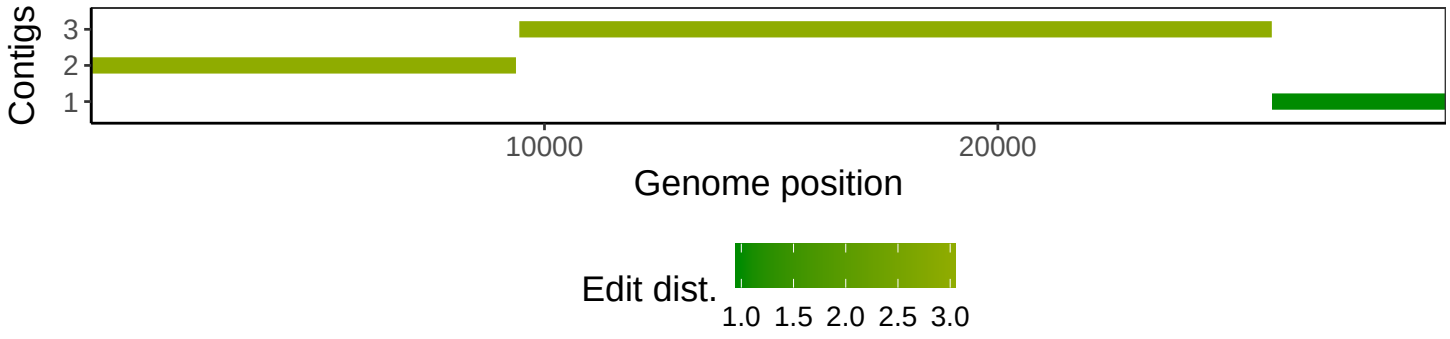
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

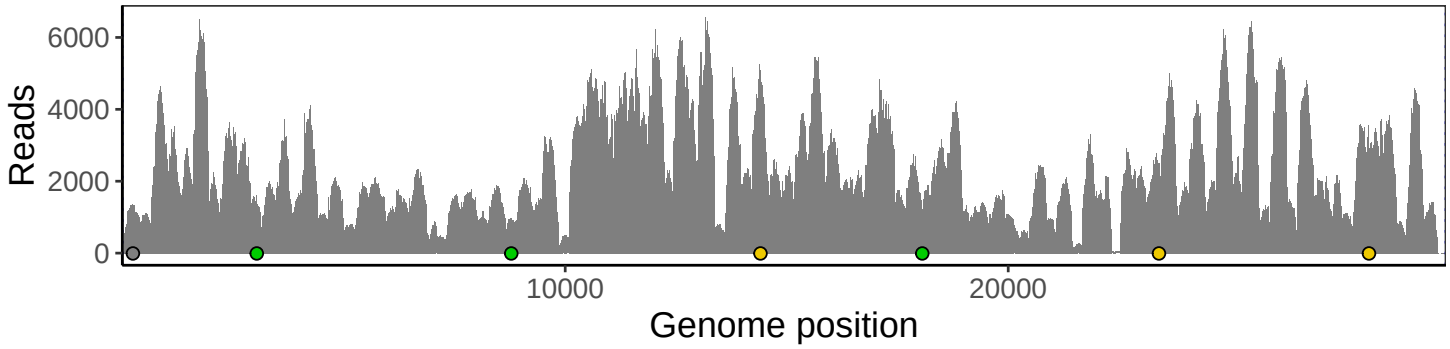


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

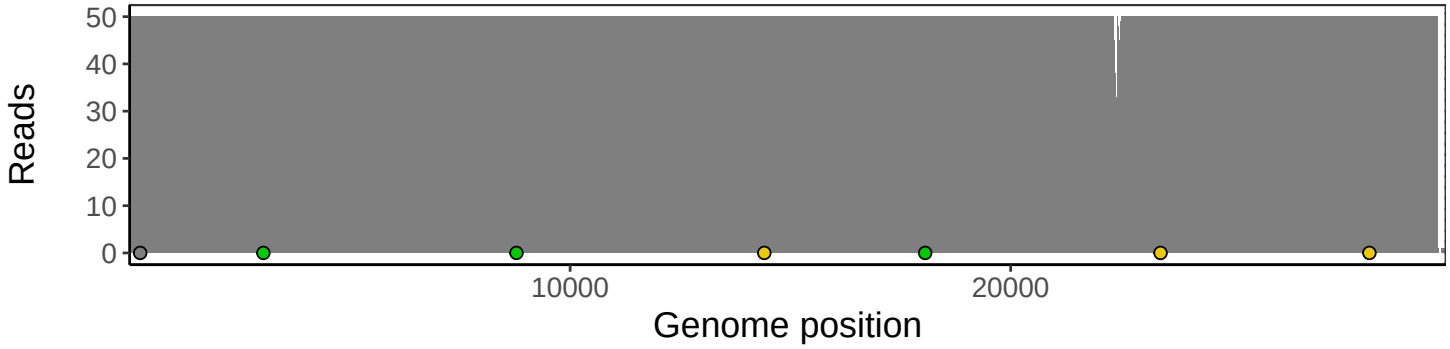


VSP0045-1 | 2020-04-29 | OP | 242o-tri | 667 genomes | single experiment

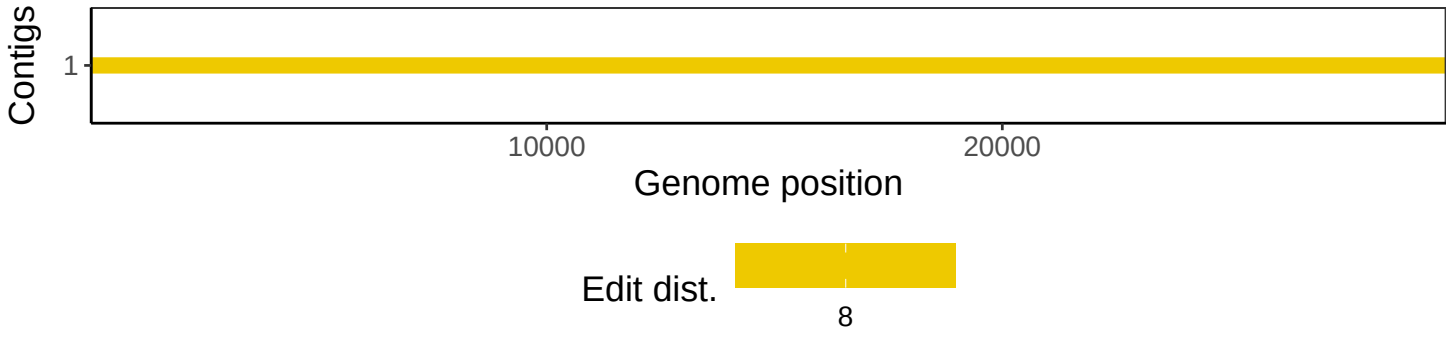
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

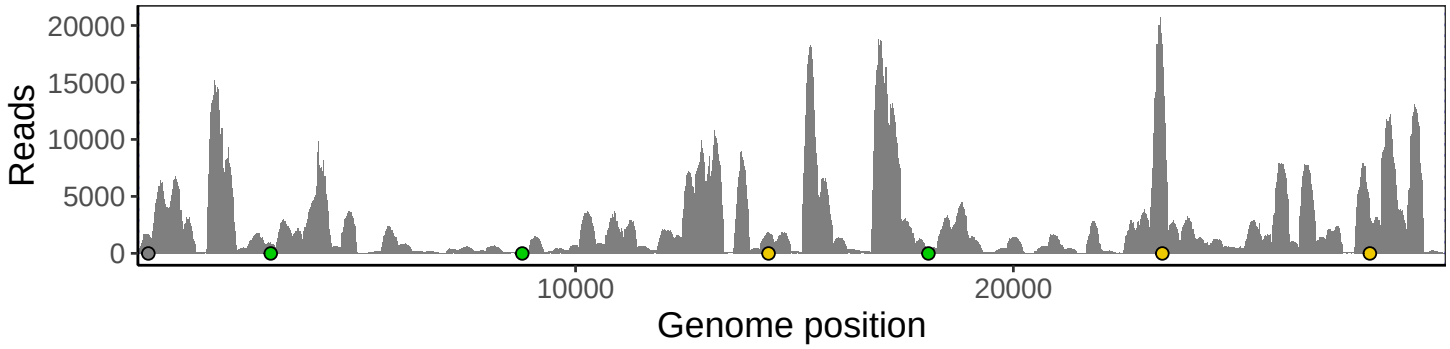


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

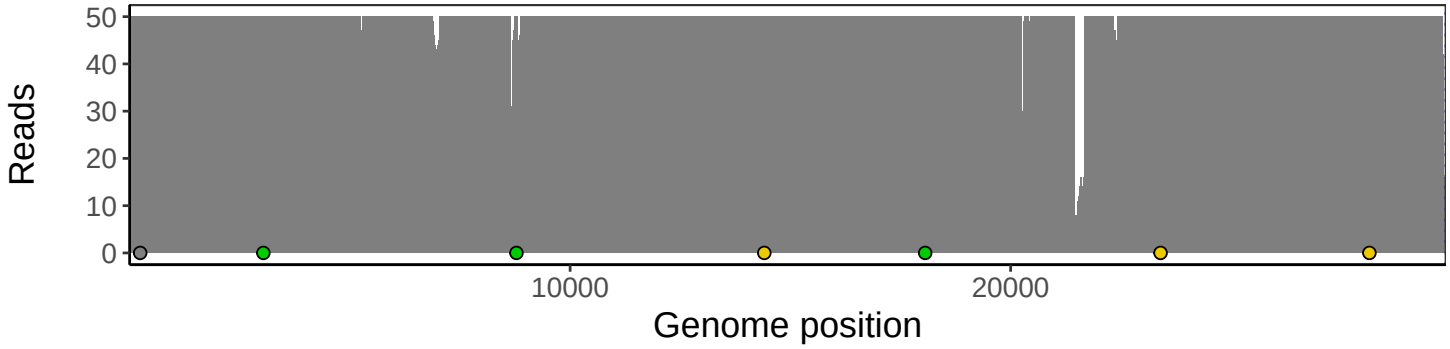


VSP0047-1 | 2020-05-01 | NP-OP | 242-qia | 222000 genomes | single experiment

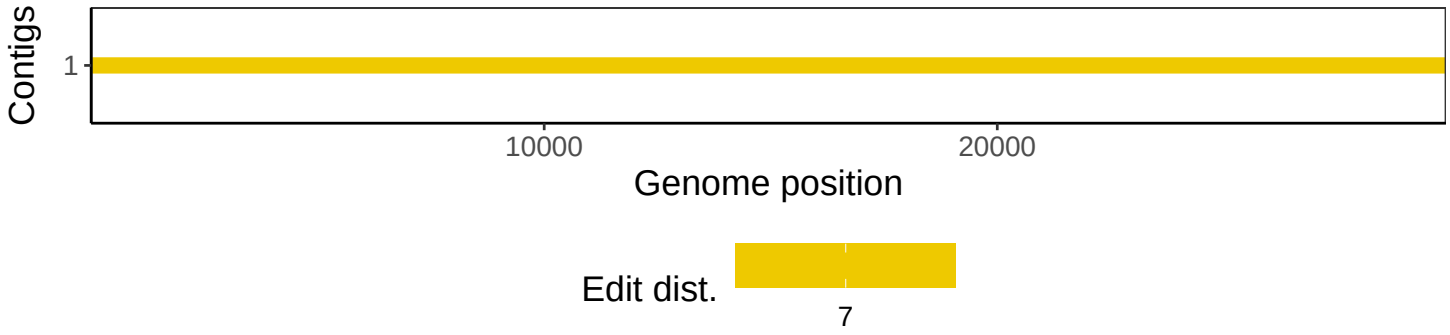
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

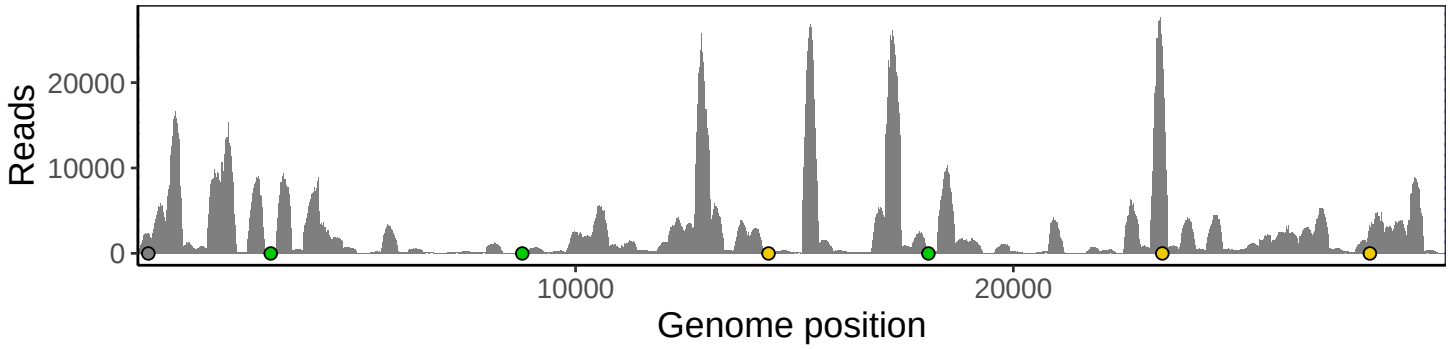


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

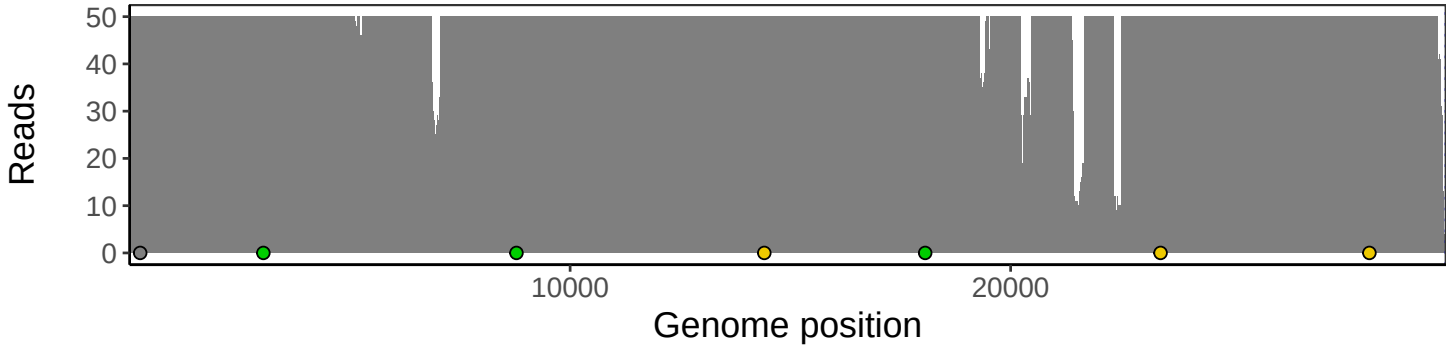


VSP0052-1 | 2020-05-04 | ETA | 242-t | 1420000 genomes | single experiment

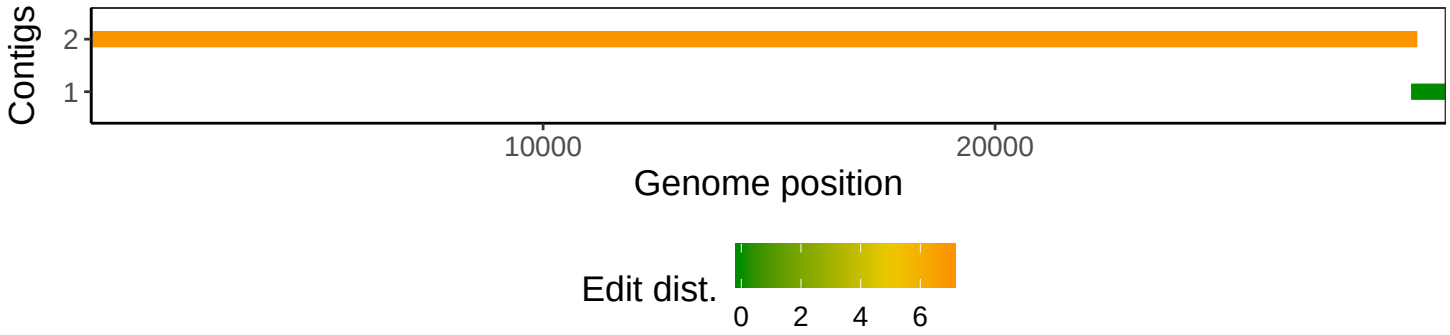
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

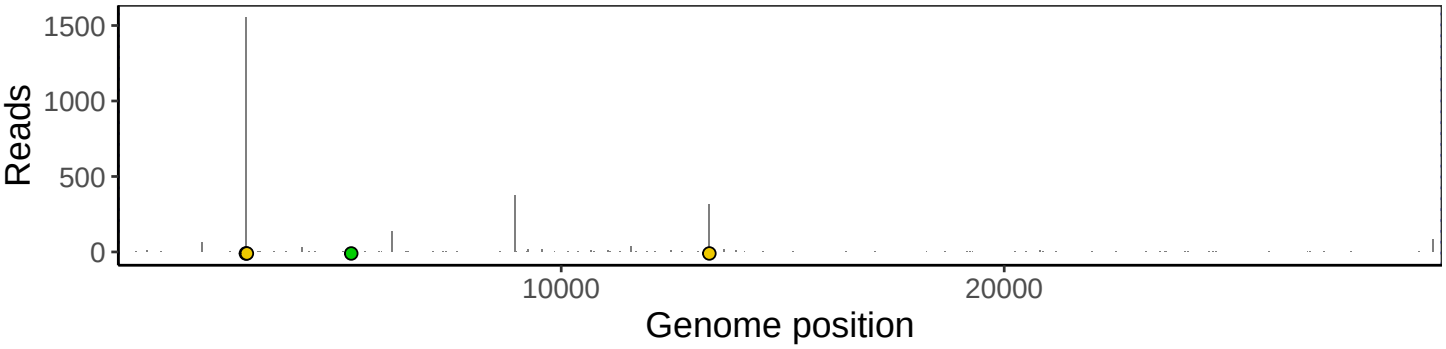


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

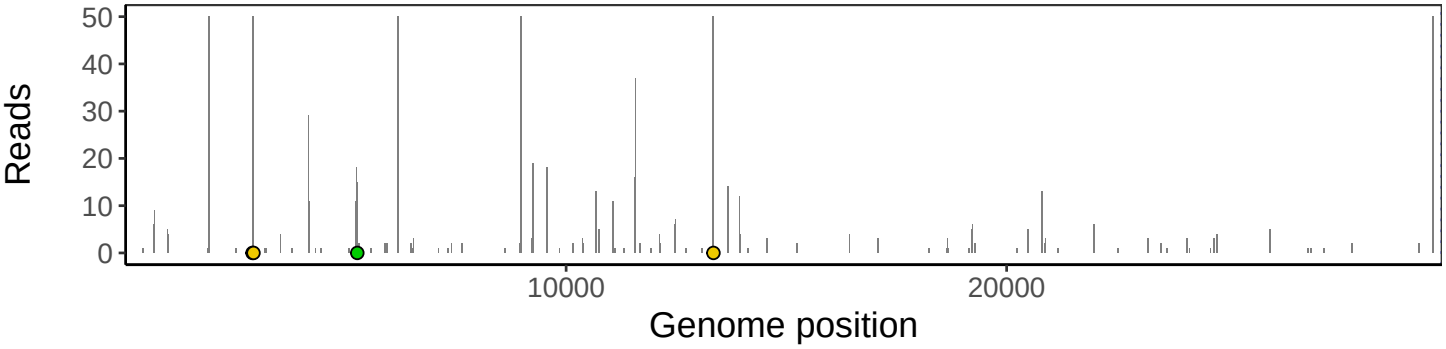


VSP0085-1 | 2020-05-06 | ETA | 242e-q | 8.54 genomes | single experiment

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.



No contig data available.

No pileup data available.

No contig data available.

No pileup data available.

No contig data available.

No pileup data available.

No contig data available.