COVID-19 subject 256

2021-01-10

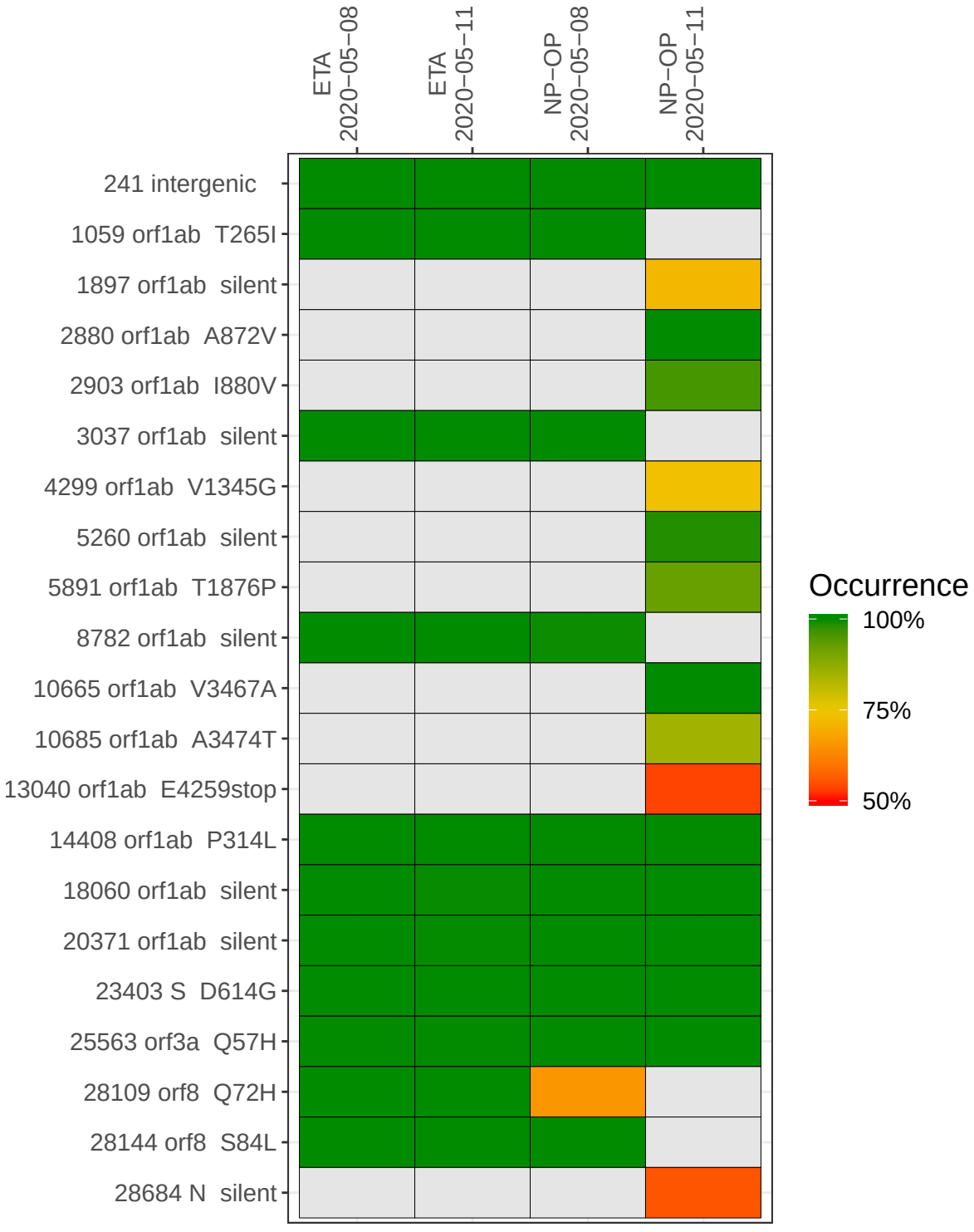
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

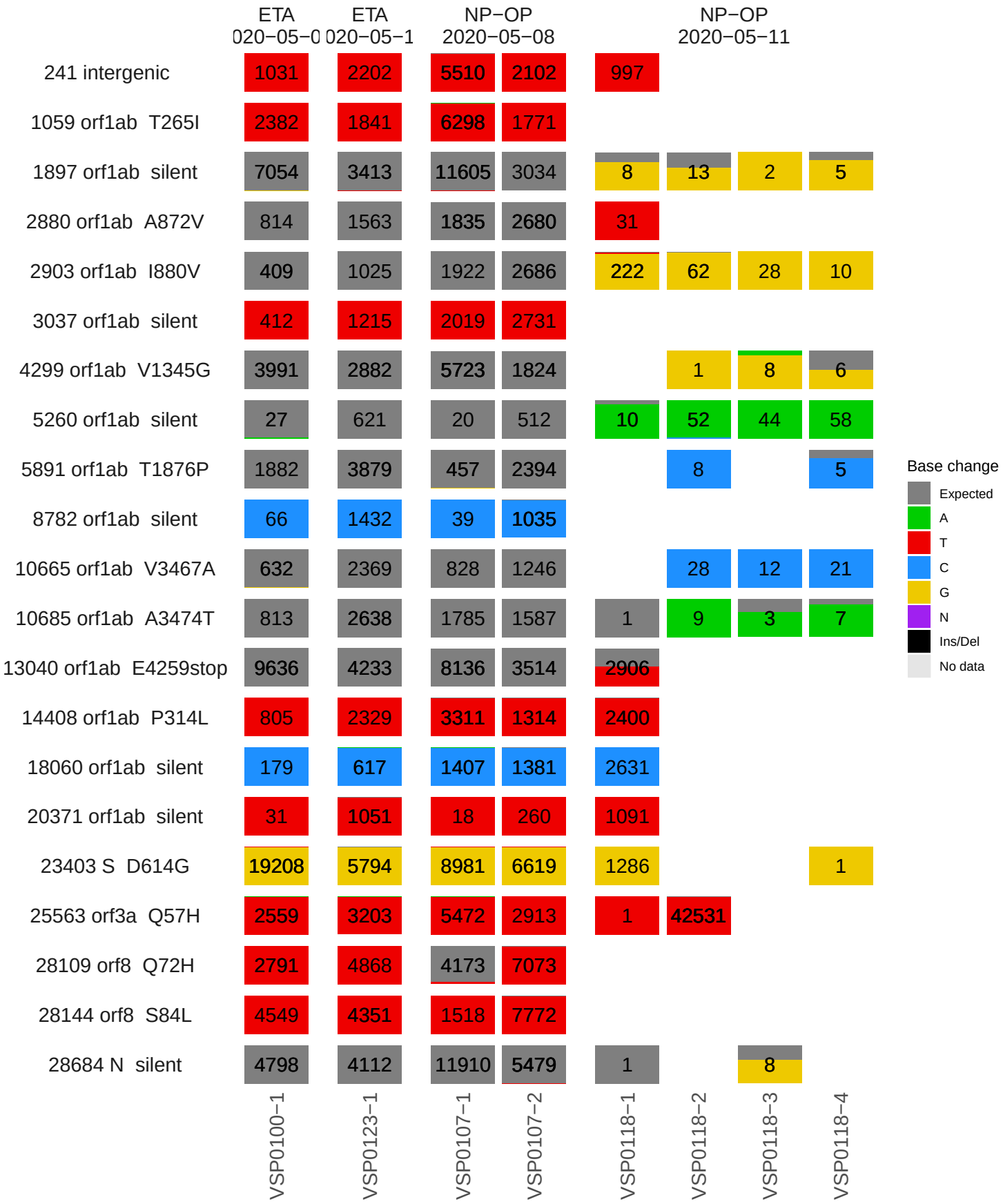
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0107	composite	NA	NP-OP	2020-05-08	29.68	99.9%	99.9%
VSP0118	composite	NA	NP-OP	2020-05-11	2.63	62.5%	58.2%
VSP0100-1	single experiment	2760000	ETA	2020-05-08	29.82	99.7%	99.7%
VSP0107-1	single experiment	595000	NP-OP	2020-05-08	29.89	99.9%	99.9%
VSP0107-2	single experiment	595000	NP-OP	2020-05-08	29.68	99.9%	99.8%
VSP0118-1	single experiment	269	NP-OP	2020-05-11	2.67	56.6%	53.5%
VSP0118-2	single experiment	1345	NP-OP	2020-05-11	0.60	9.3%	5.7%
VSP0118-3	single experiment	1345	NP-OP	2020-05-11	0.56	11.1%	7.8%
VSP0118-4	single experiment	1345	NP-OP	2020-05-11	0.60	10.9%	6.8%
VSP0123-1	single experiment	123000	ETA	2020-05-11	29.26	99.9%	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.

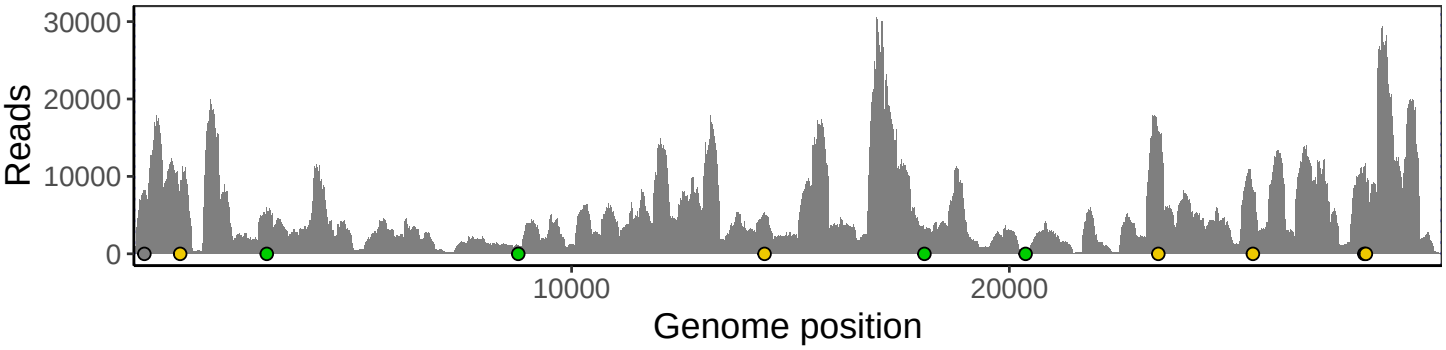



ETA 020−05−0
ETA 020−05−1
NP−OP 2020−05−08
NP−OP 2020−05−11
241 intergenic
1059 orf1ab T265I
1897 orf1ab silent
2880 orf1ab A872V
2903 orf1ab I880V
3037 orf1ab silent
4299 orf1ab V1345G
5260 orf1ab silent
5891 orf1ab T1876P
8782 orf1ab silent
10665 orf1ab V3467A
10685 orf1ab A3474T
13040 orf1ab E4259stop
14408 orf1ab P314L
18060 orf1ab silent
20371 orf1ab silent
23403 S D614G
25563 orf3a Q57H
28109 orf8 Q72H
28144 orf8 S84L
28684 N silent
VSP0100−1
VSP0123−1
VSP0107−1
VSP0107−2
VSP0118−1
VSP0118−2
VSP0118−3
VSP0118−4
Base change
Expected
A
T
C
G
N
Ins/Del
No data

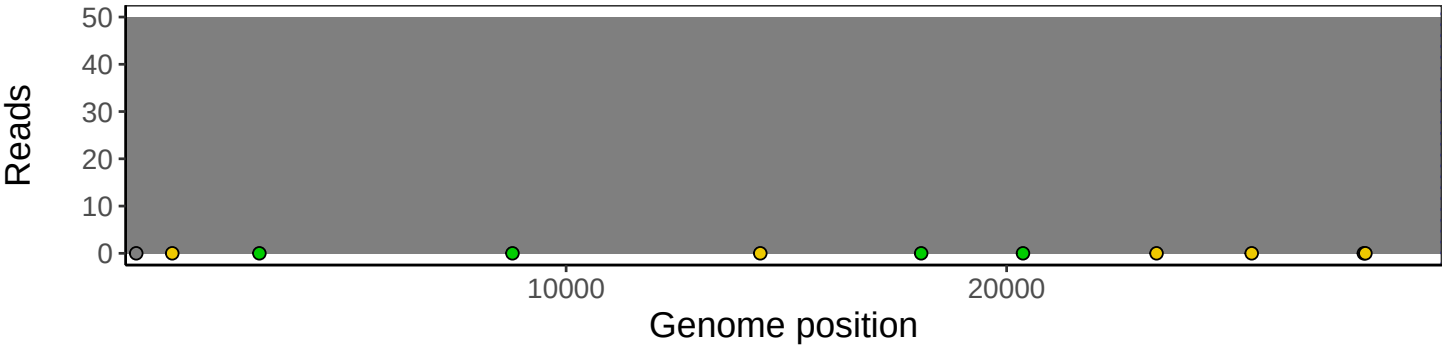
Analyses of individual experiments and composite results.

VSP0107 | 2020-05-08 | NP-OP | 256no-t | composite result

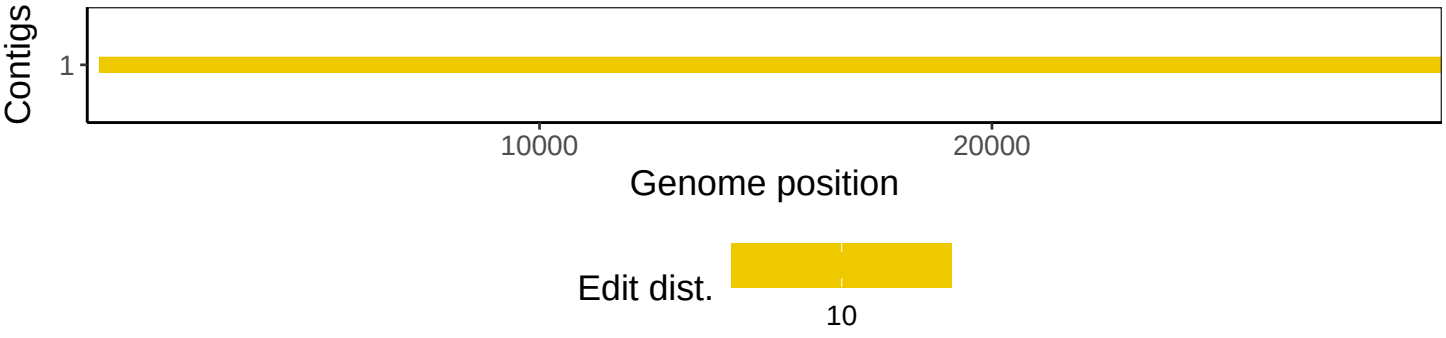
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

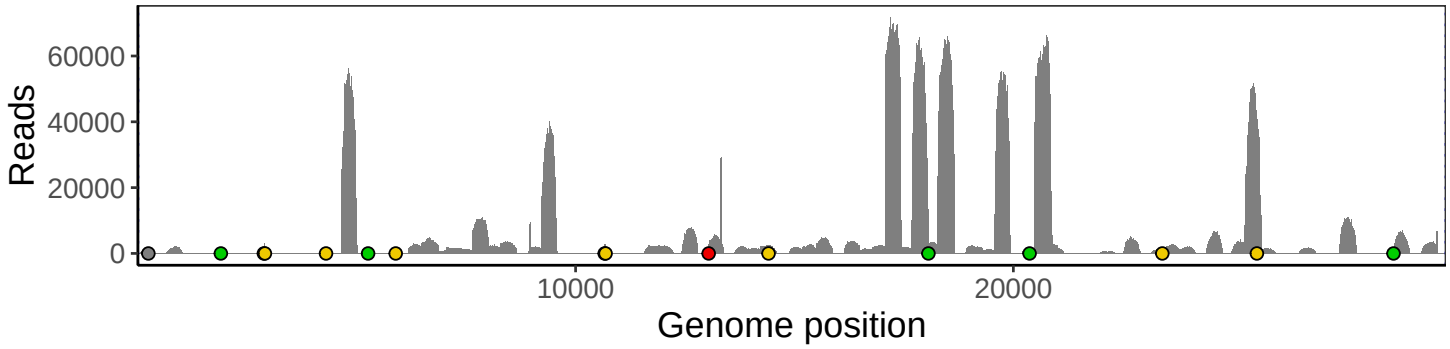


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

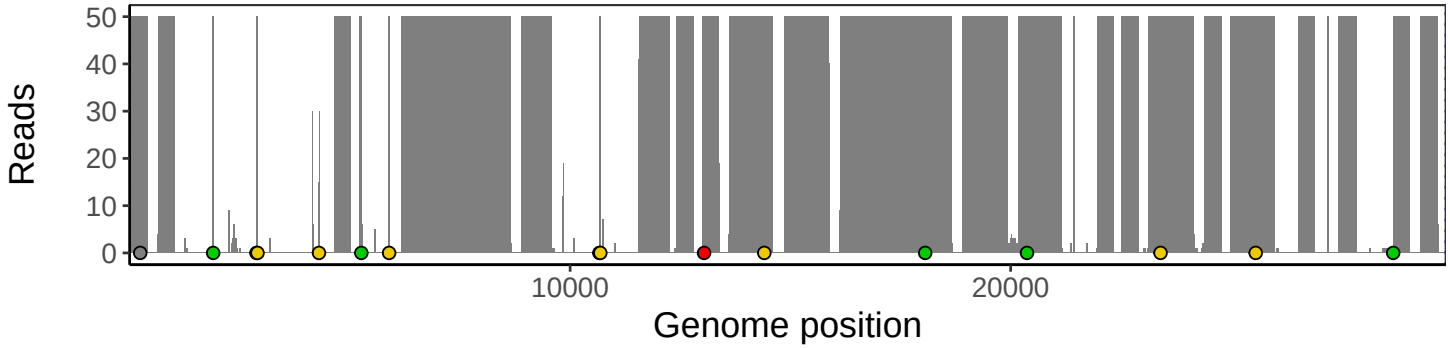


VSP0118 | 2020-05-11 | NP-OP | 256no-t2 | composite result

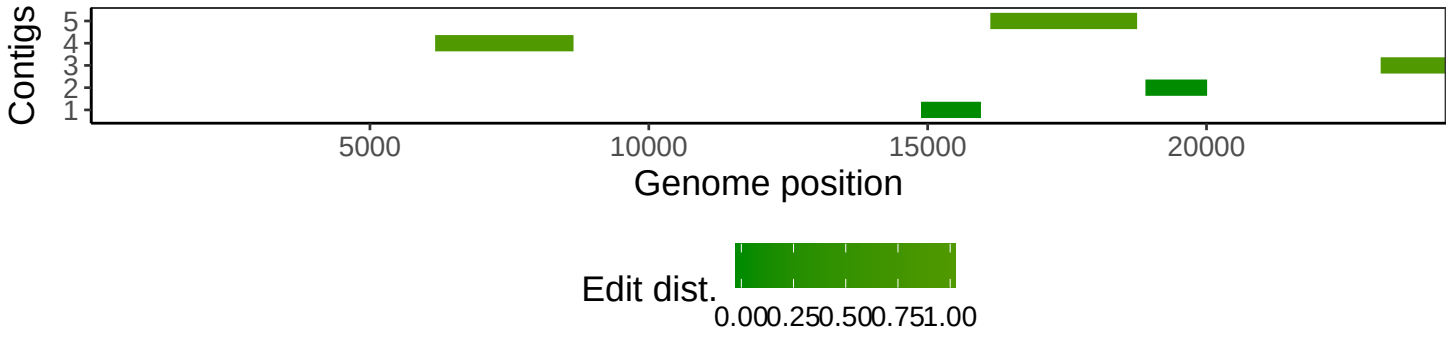
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

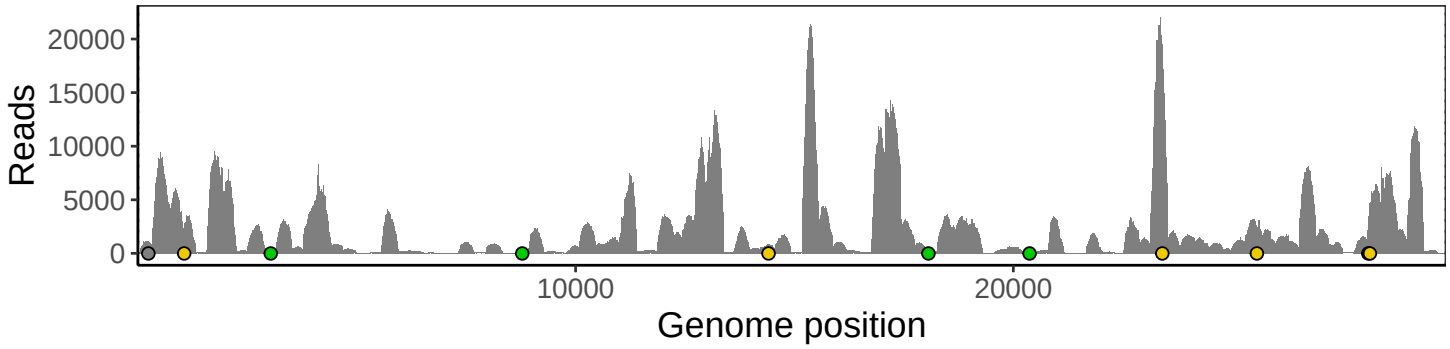


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

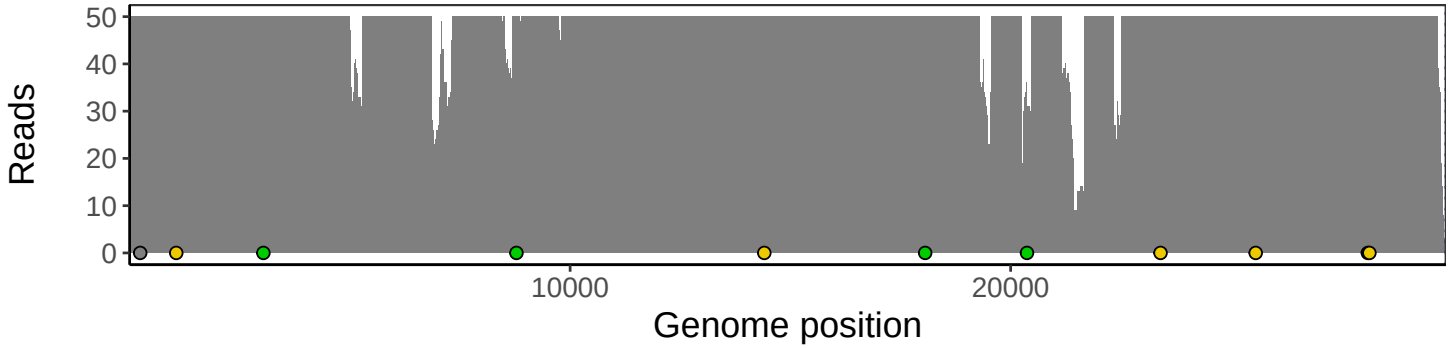


VSP0100-1 | 2020-05-08 | ETA | 256e-q | 2760000 genomes | single experiment

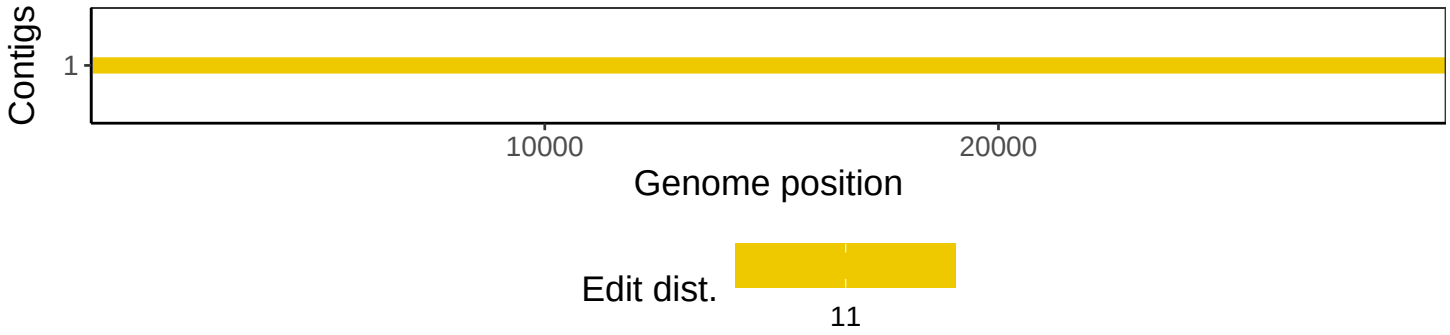
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

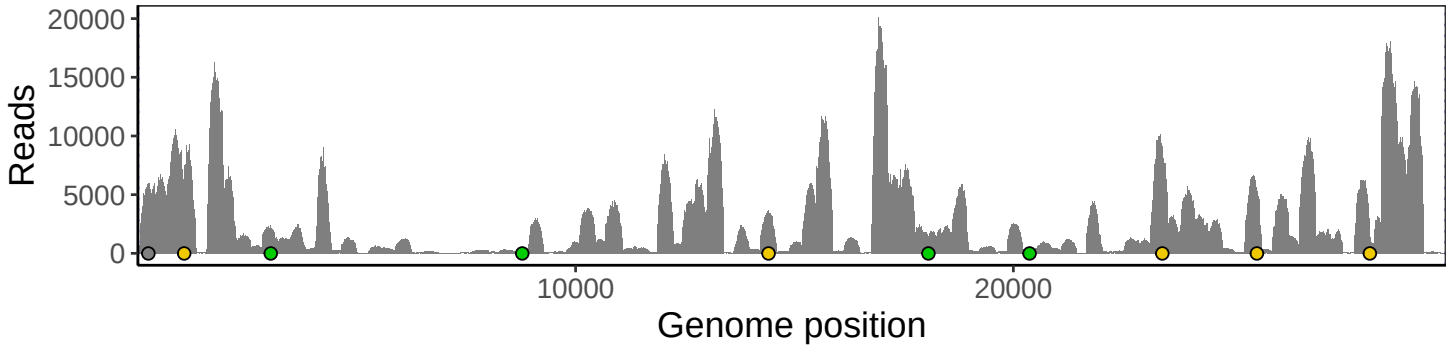


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

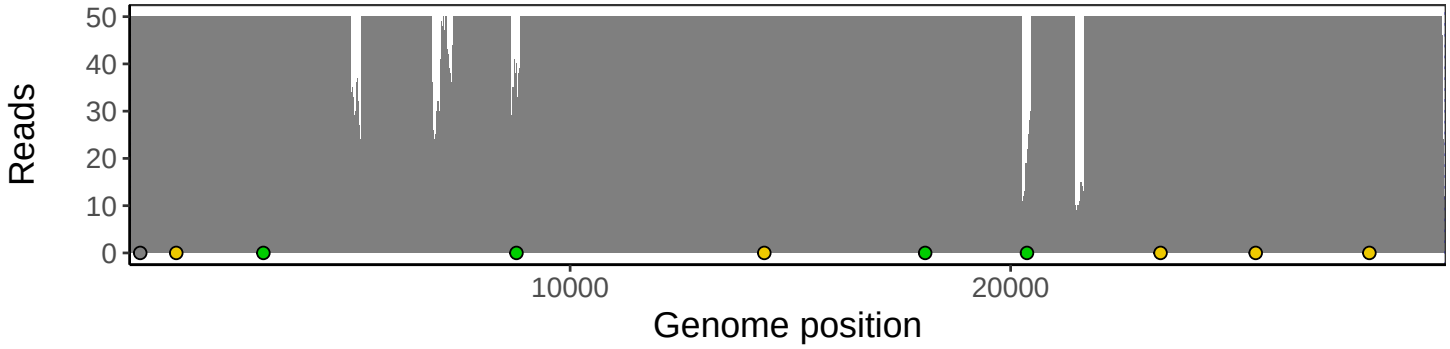


VSP0107-1 | 2020-05-08 | NP-OP | 256no-t | 595000 genomes | single experiment

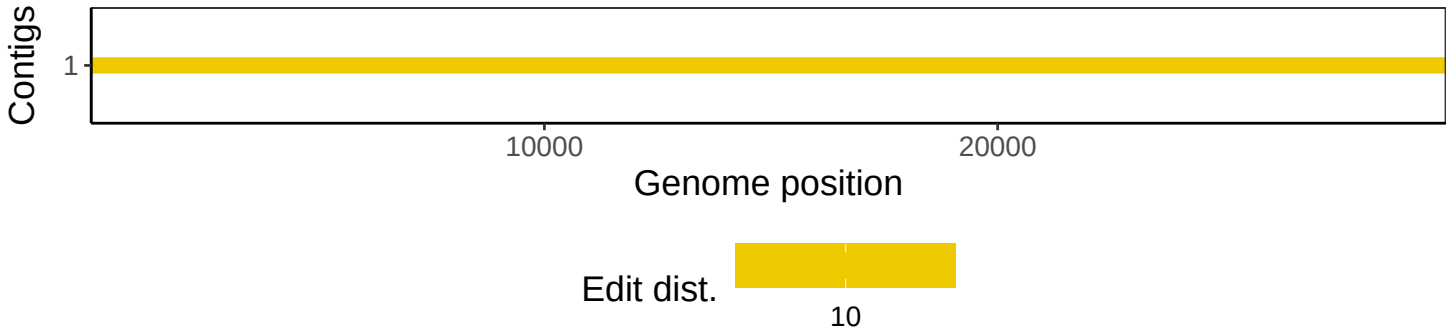
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

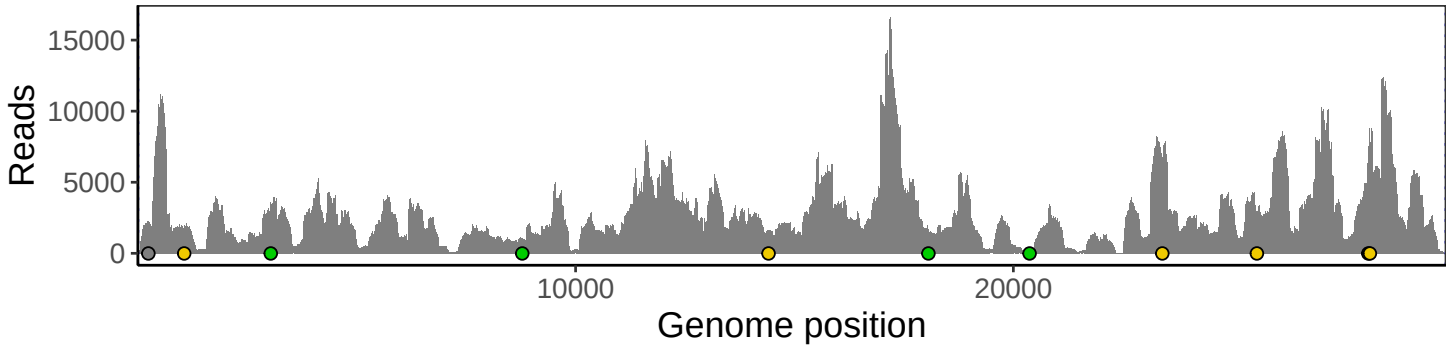


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

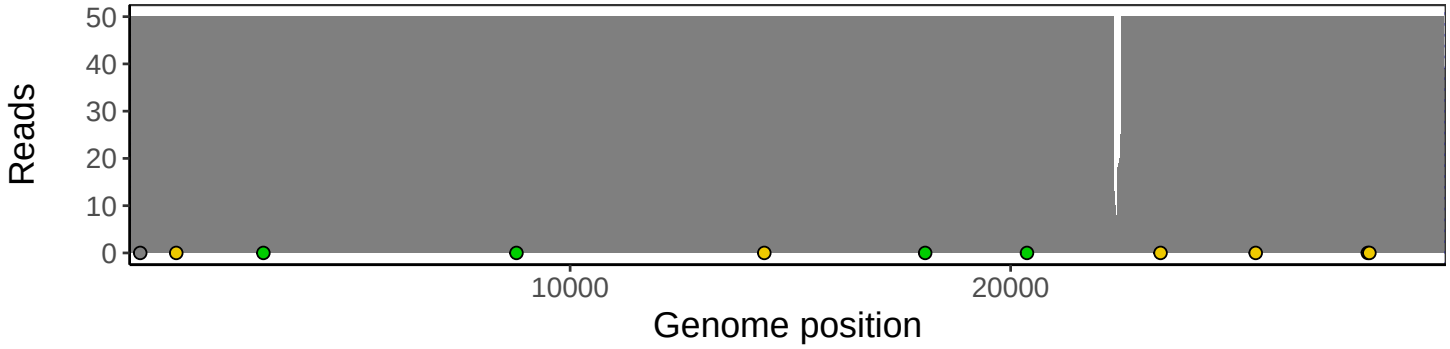


VSP0107-2 | 2020-05-08 | NP-OP | 256no-t | 595000 genomes | single experiment

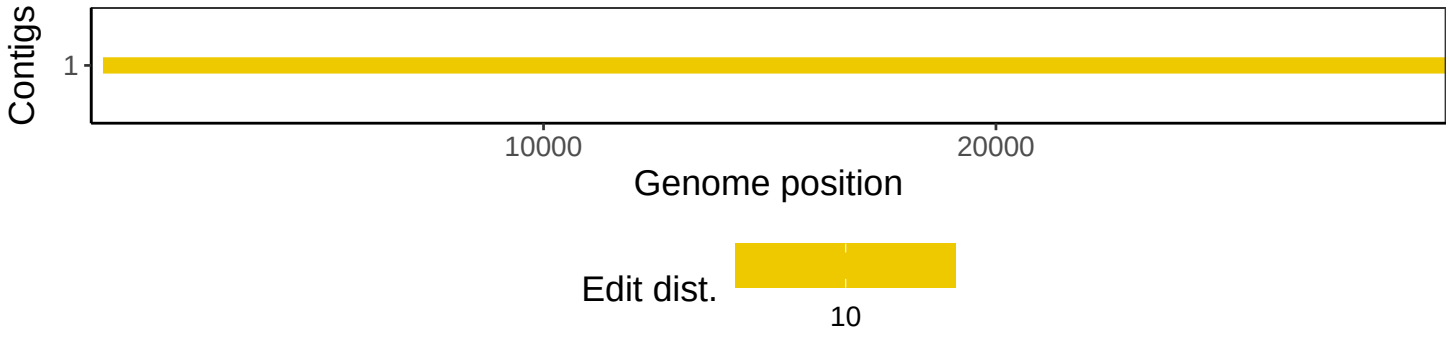
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

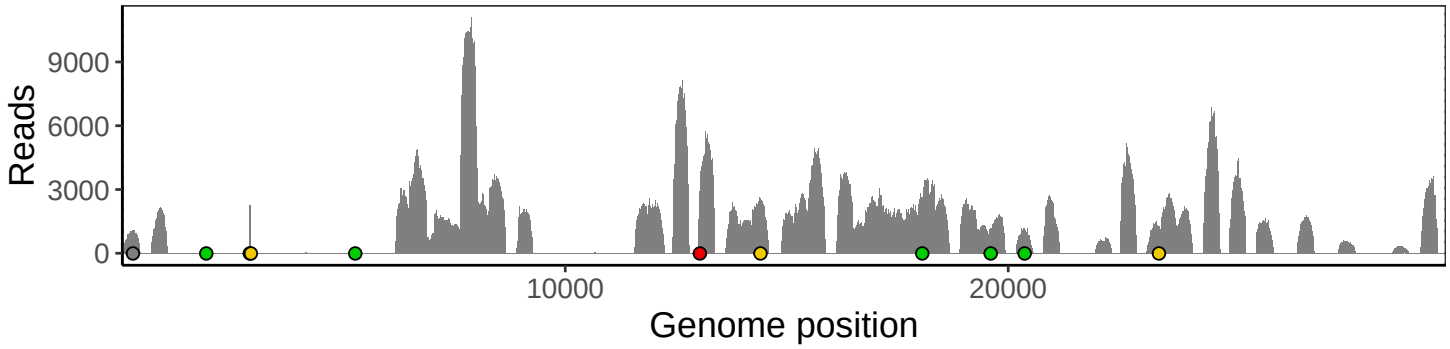


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

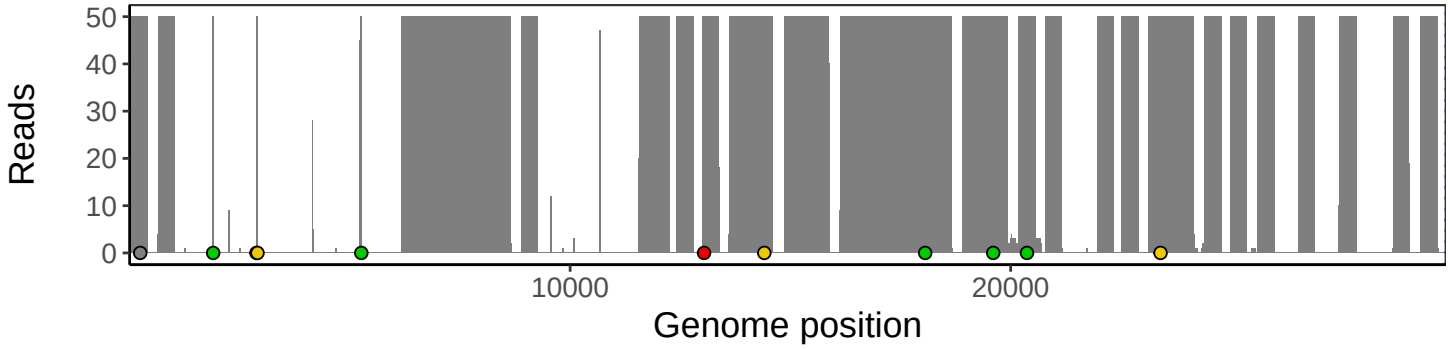


VSP0118-1 | 2020-05-11 | NP-OP | 256no-t2 | 269 genomes | single experiment

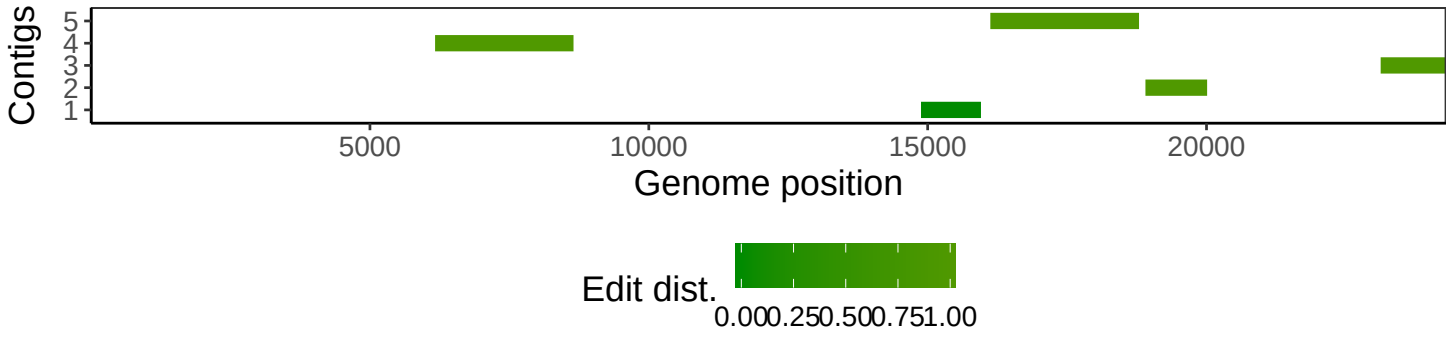
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

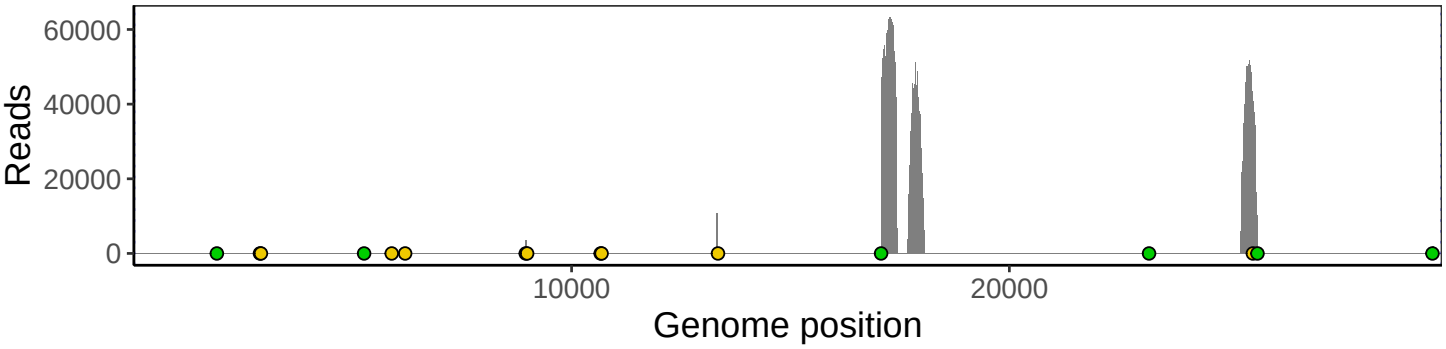


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

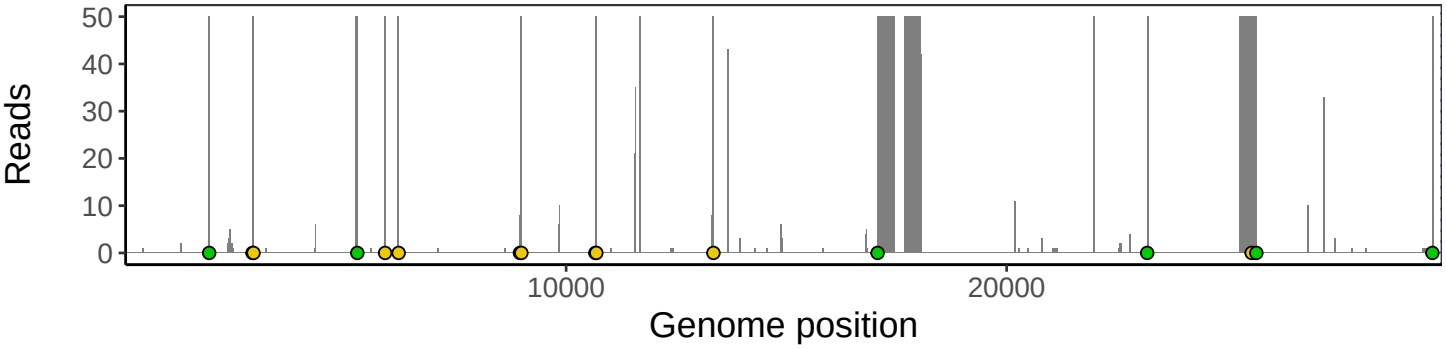


VSP0118-2 | 2020-05-11 | NP-OP | 256no-t2 | 1345 genomes | single experiment

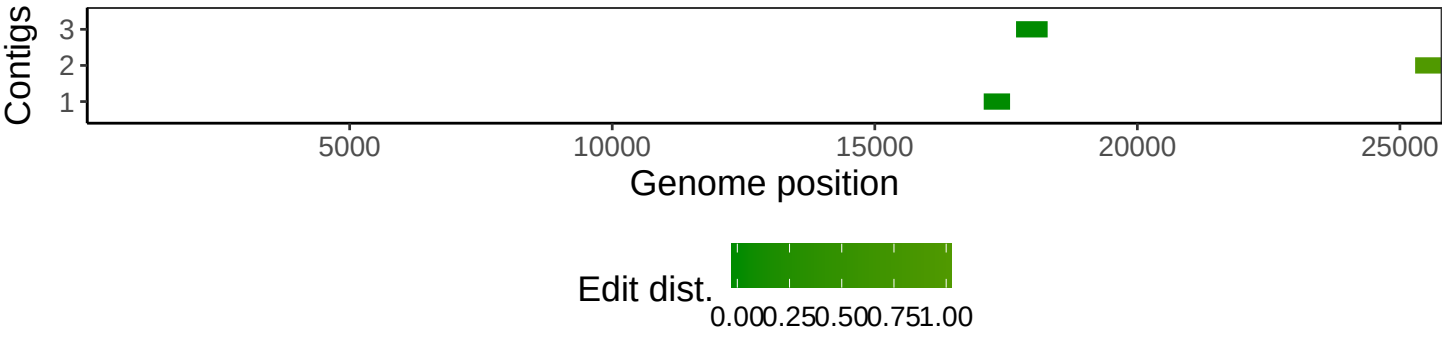
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

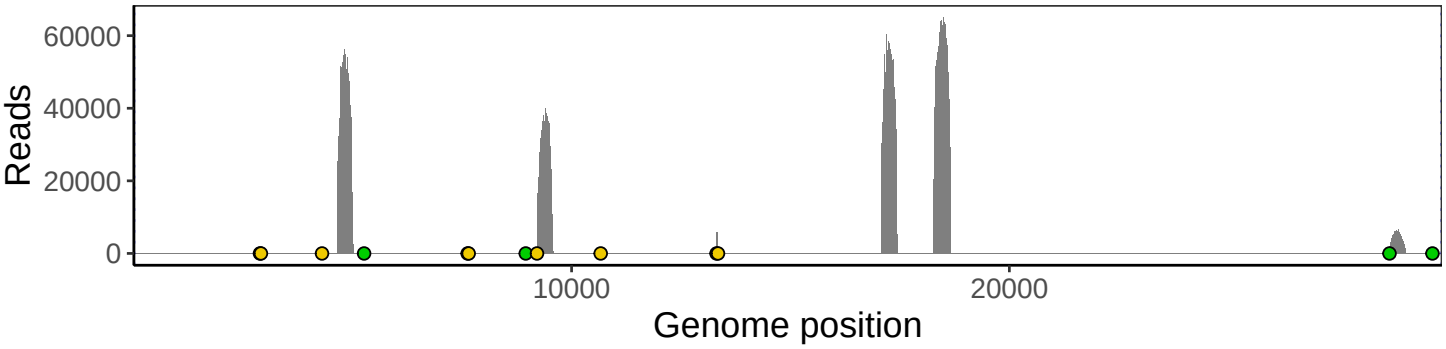


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

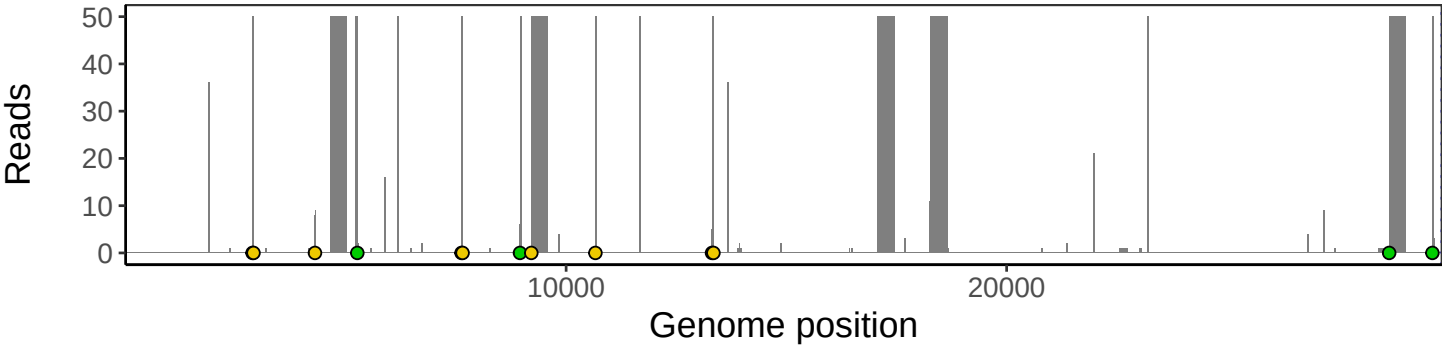


VSP0118-3 | 2020-05-11 | NP-OP | 256no-t2 | 1345 genomes | single experiment

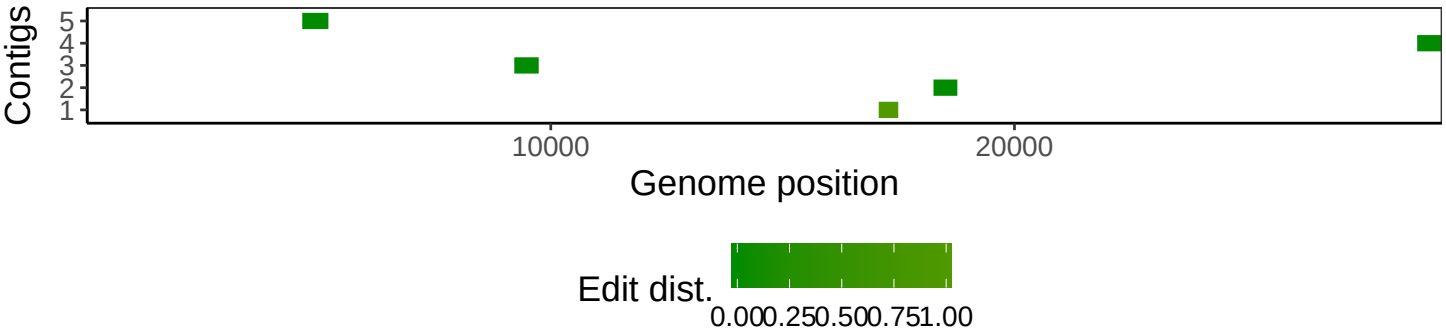
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

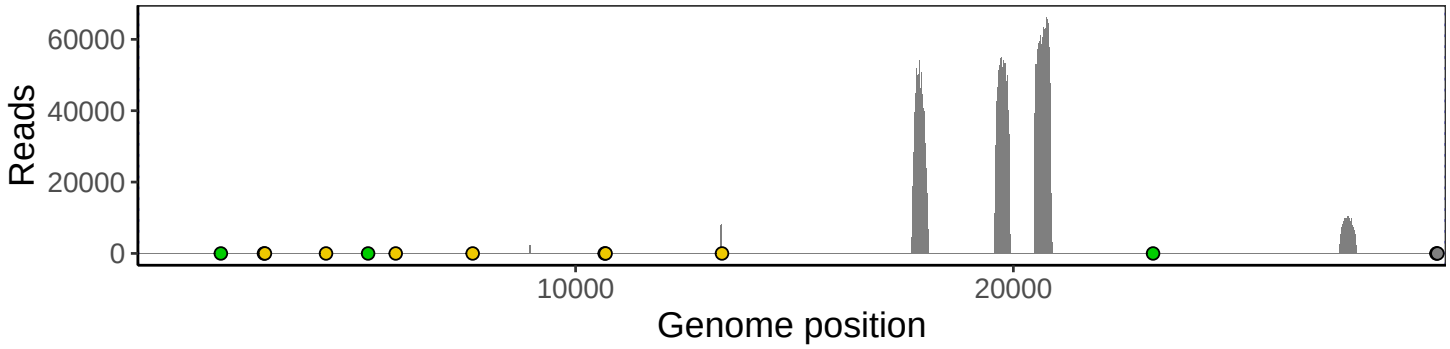


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

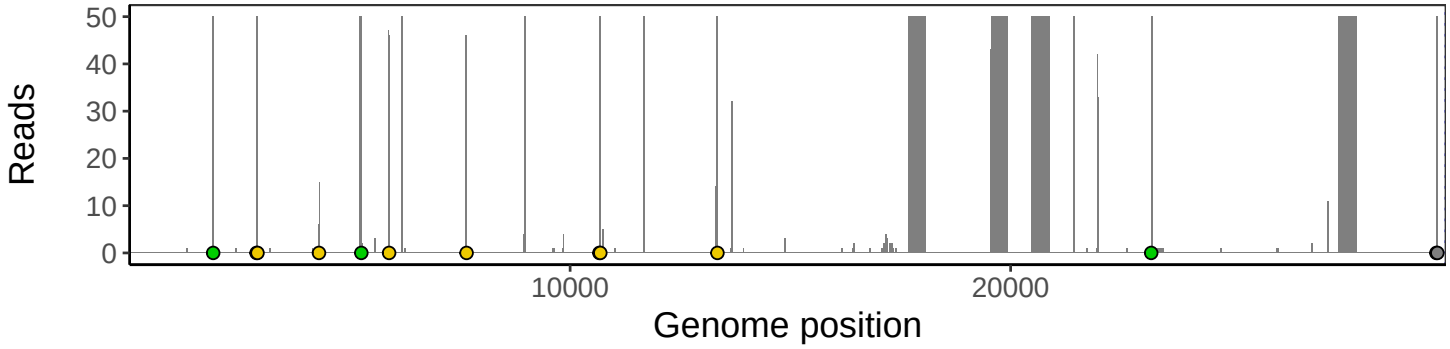


VSP0118-4 | 2020-05-11 | NP-OP | 256no-t2 | 1345 genomes | single experiment

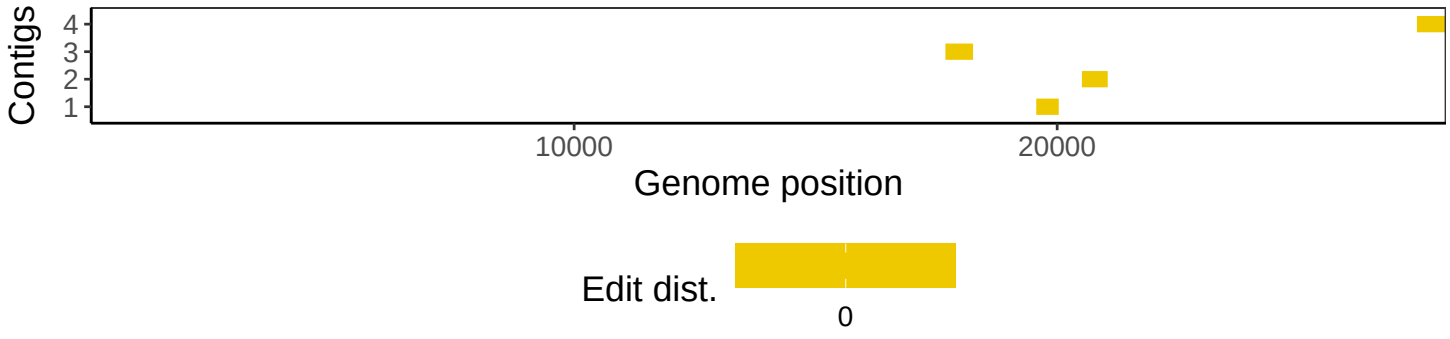
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

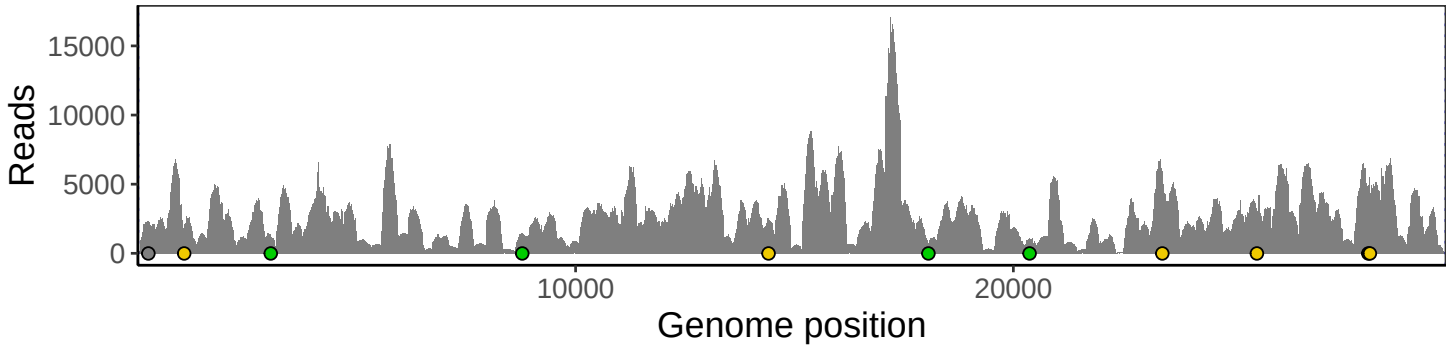


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

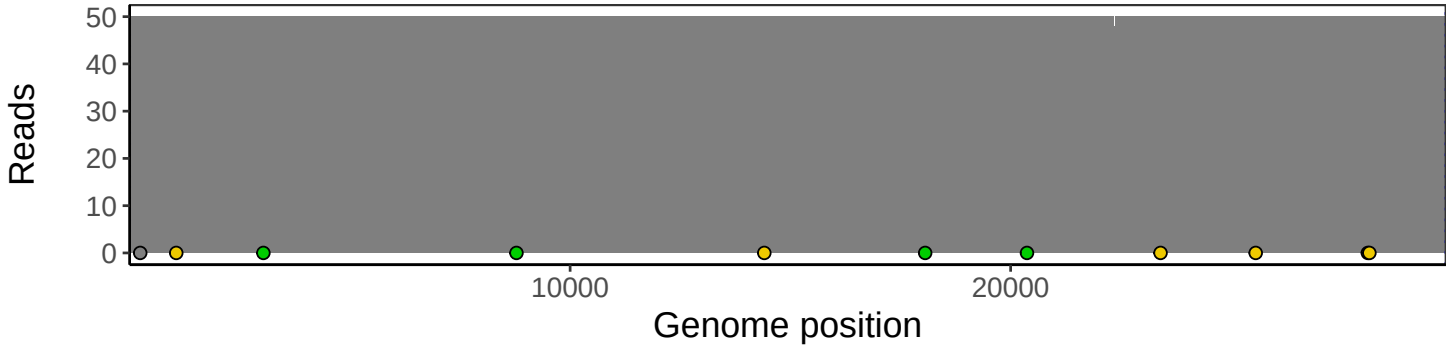


VSP0123-1 | 2020-05-11 | ETA | 256e-q | 123000 genomes | single experiment

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.