COVID-19 subject 251

2021-01-08

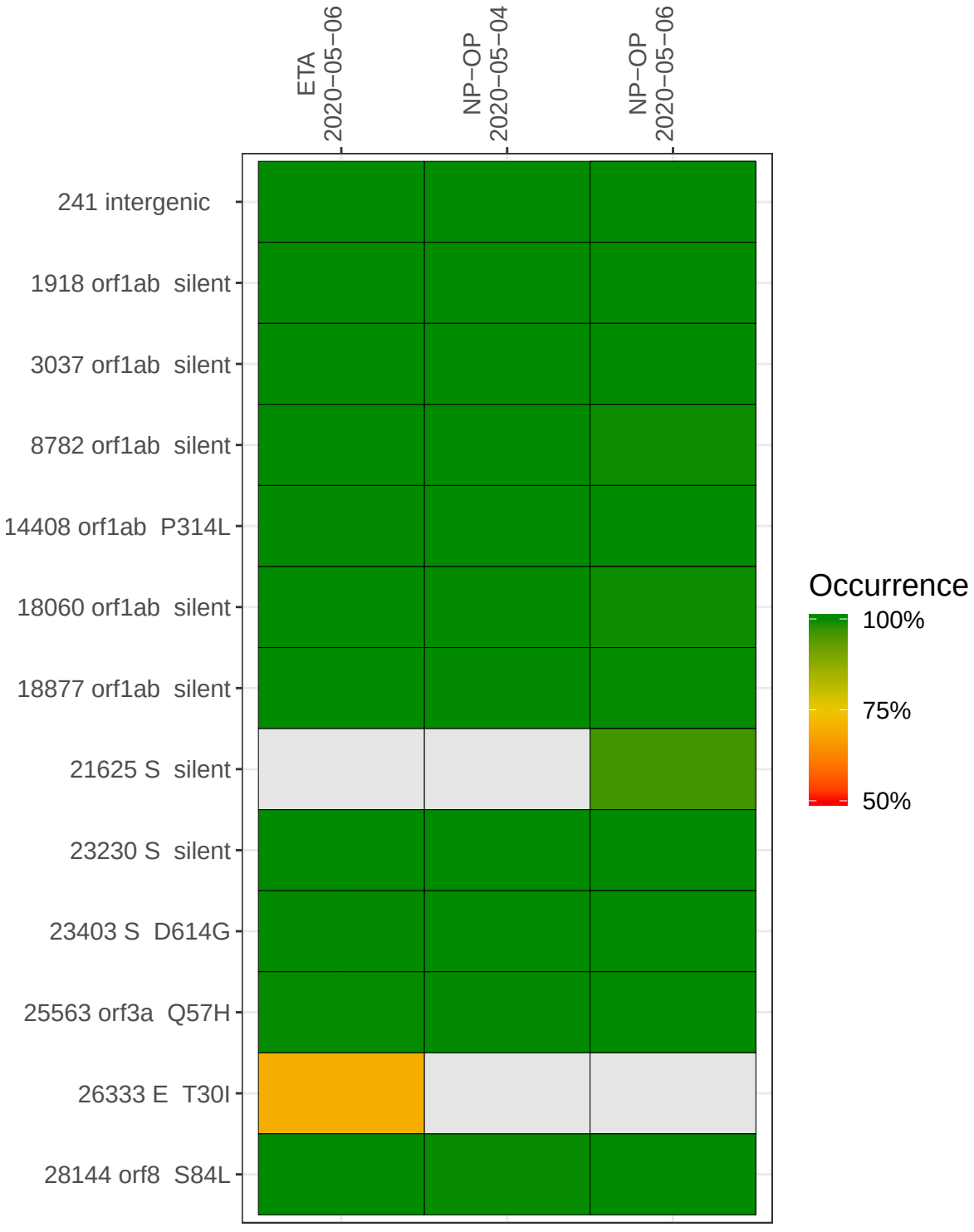
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

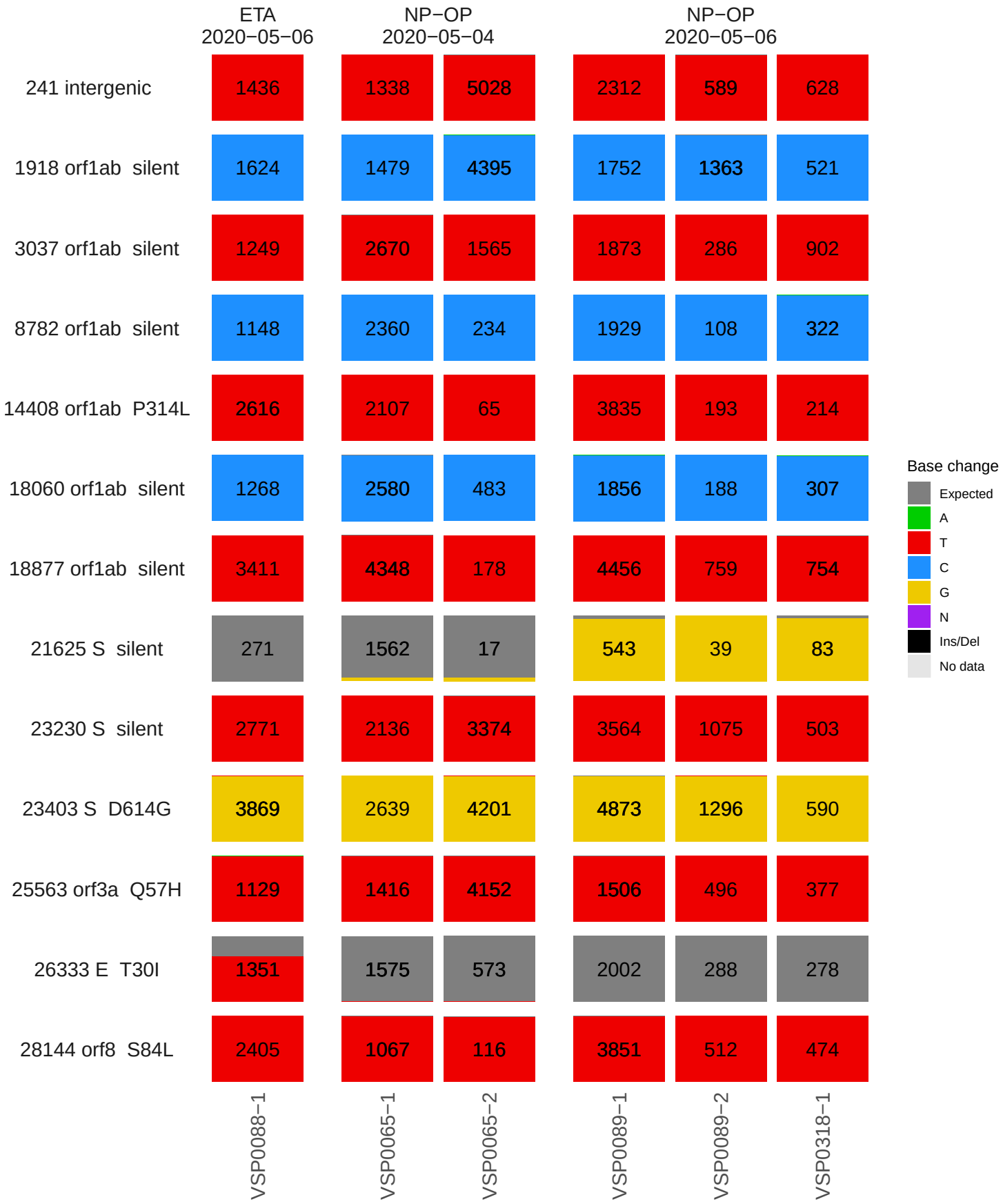
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0065	composite	NA	NP-OP	2020-05-04	16.53	100.0%	99.8%
VSP0089	composite	NA	NP-OP	2020-05-06	30.00	100.0%	99.7%
VSP0065-1	single experiment	7550000	NP-OP	2020-05-04	16.47	100.0%	99.8%
VSP0065-2	single experiment	7550000	NP-OP	2020-05-04	29.85	100.0%	99.3%
VSP0088-1	single experiment	255500	ETA	2020-05-06	29.82	100.0%	99.8%
VSP0089-1	single experiment	570000	NP-OP	2020-05-06	29.91	100.0%	99.7%
VSP0089-2	single experiment	570000	NP-OP	2020-05-06	29.45	100.0%	98.6%
VSP0318-1	single experiment	890000	NP-OP	2020-05-06	24.51	100.0%	99.2%

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.

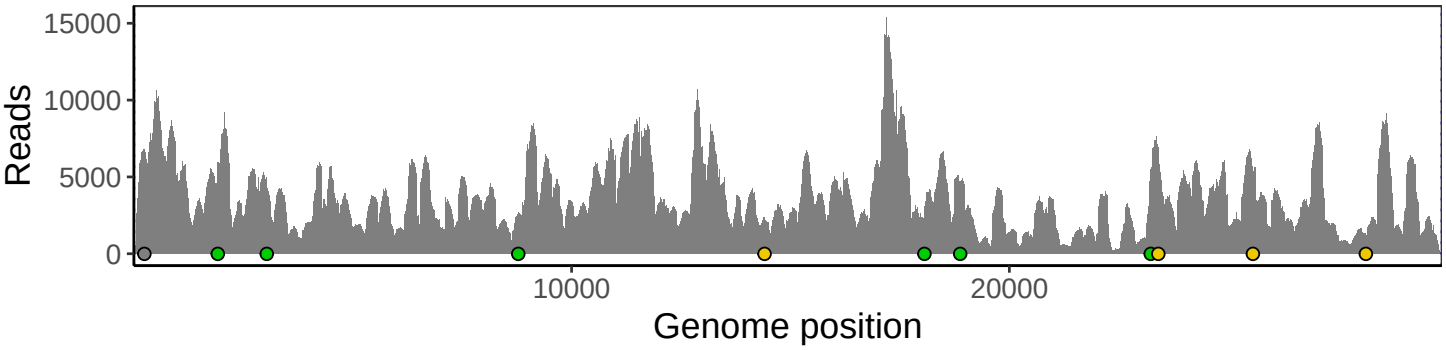



ETA
2020−05−06
NP−OP
2020−05−04
NP−OP
2020−05−06
241 intergenic
1436
1338
5028
2312
589
628
1918 orf1ab silent
1624
1479
4395
1752
1363
521
3037 orf1ab silent
1249
2670
1565
1873
286
902
8782 orf1ab silent
1148
2360
234
1929
108
322
14408 orf1ab P314L
2616
2107
65
3835
193
214
18060 orf1ab silent
1268
2580
483
1856
188
307
18877 orf1ab silent
3411
4348
178
4456
759
754
21625 S silent
271
1562
17
543
39
83
23230 S silent
2771
2136
3374
3564
1075
503
23403 S D614G
3869
2639
4201
4873
1296
590
25563 orf3a Q57H
1129
1416
4152
1506
496
377
26333 E T30I
1351
1575
573
2002
288
278
28144 orf8 S84L
2405
1067
116
3851
512
474
VSP0088−1
VSP0065−1
VSP0065−2
VSP0089−1
VSP0089−2
VSP0318−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

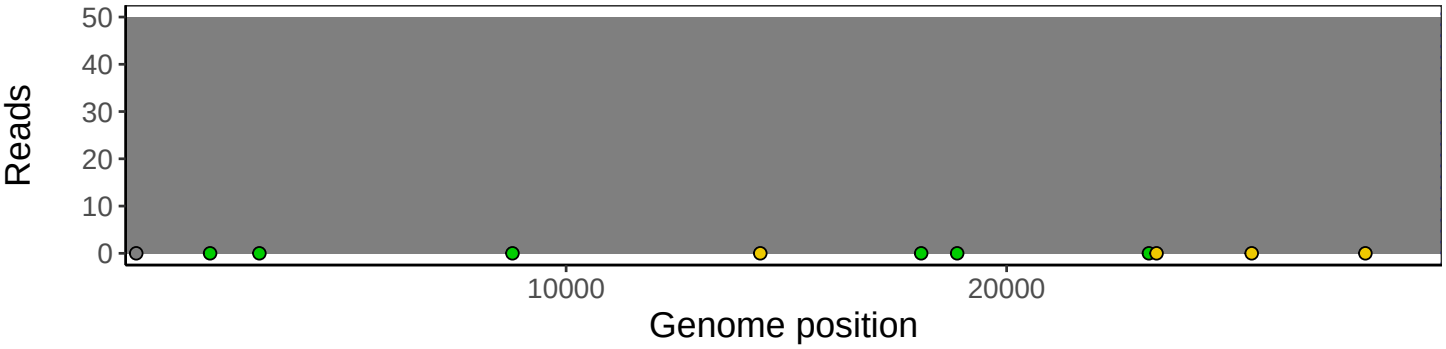
Analyses of individual experiments and composite results.

VSP0065 | 2020-05-04 | NP-OP | 251-q | composite result

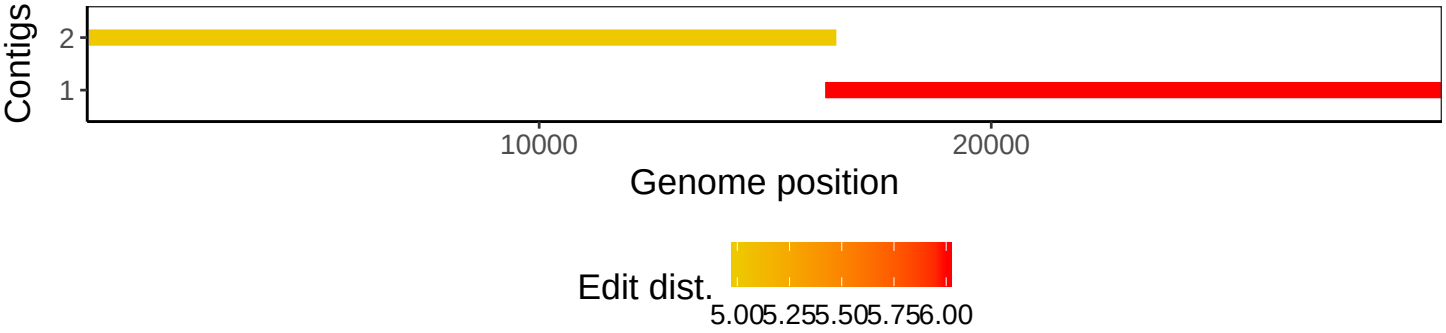
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

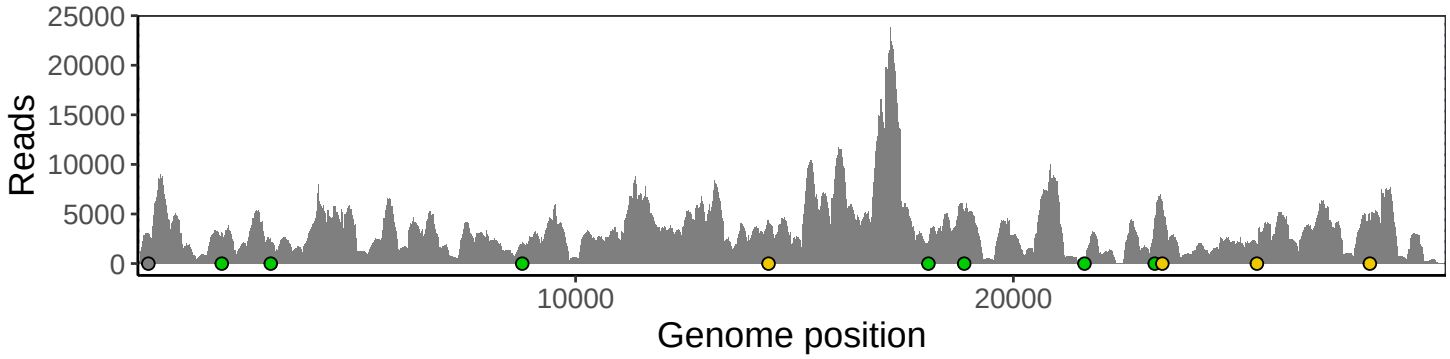


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

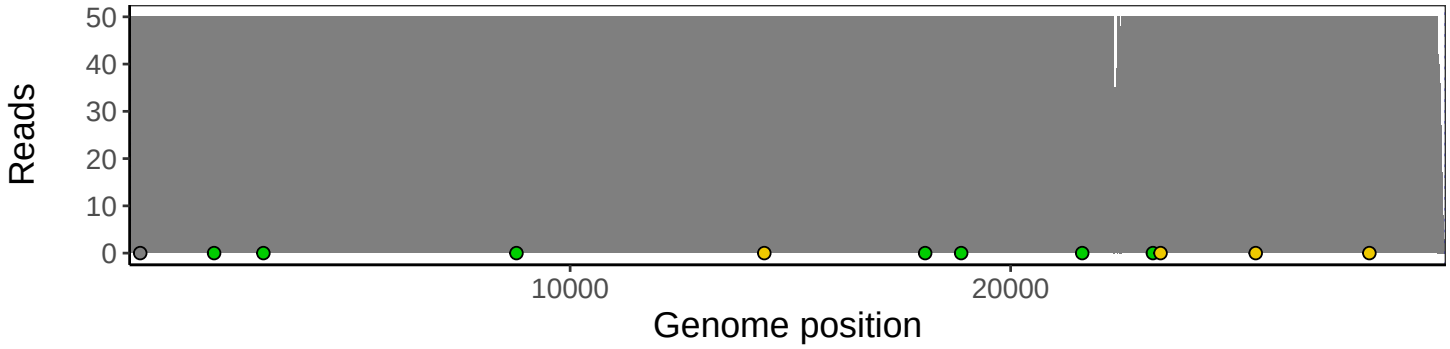


VSP0089 | 2020-05-06 | NP-OP | 251no-q | composite result

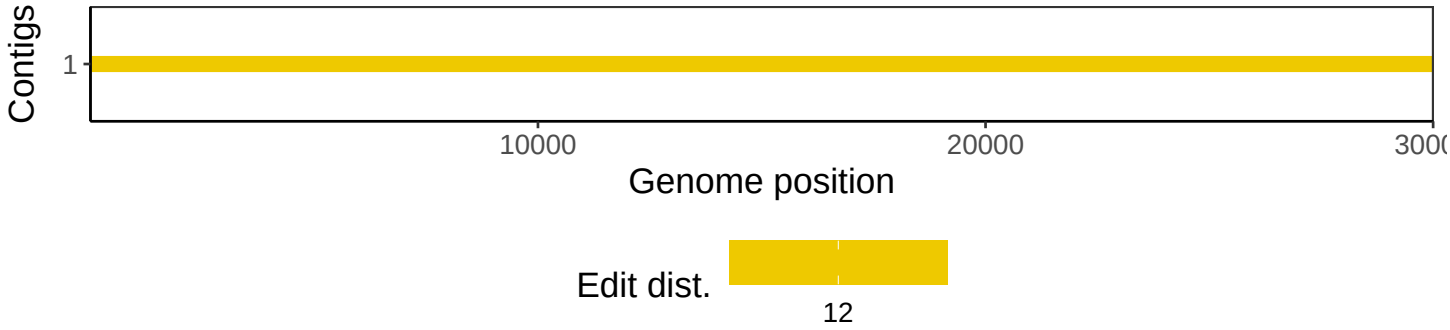
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

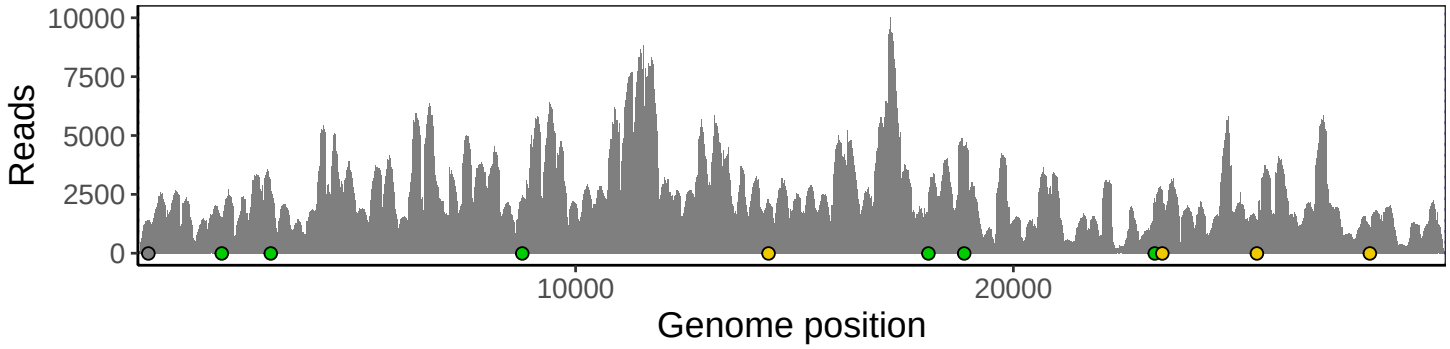


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

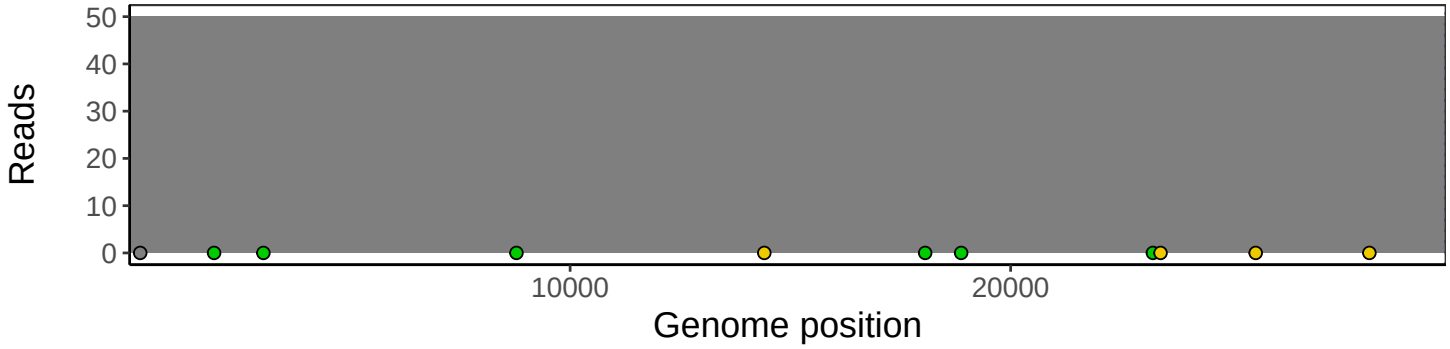


VSP0065-1 | 2020-05-04 | NP-OP | 251-q | 7550000 genomes | single experiment

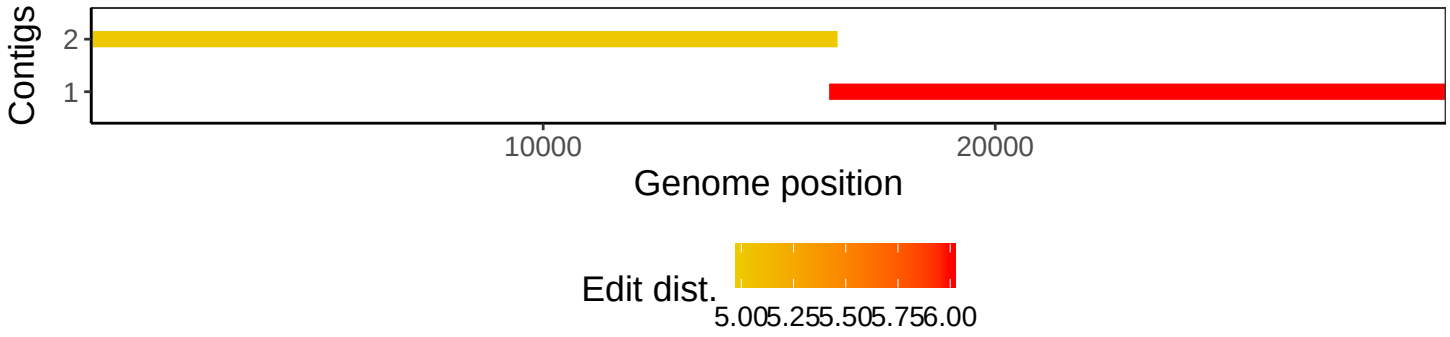
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

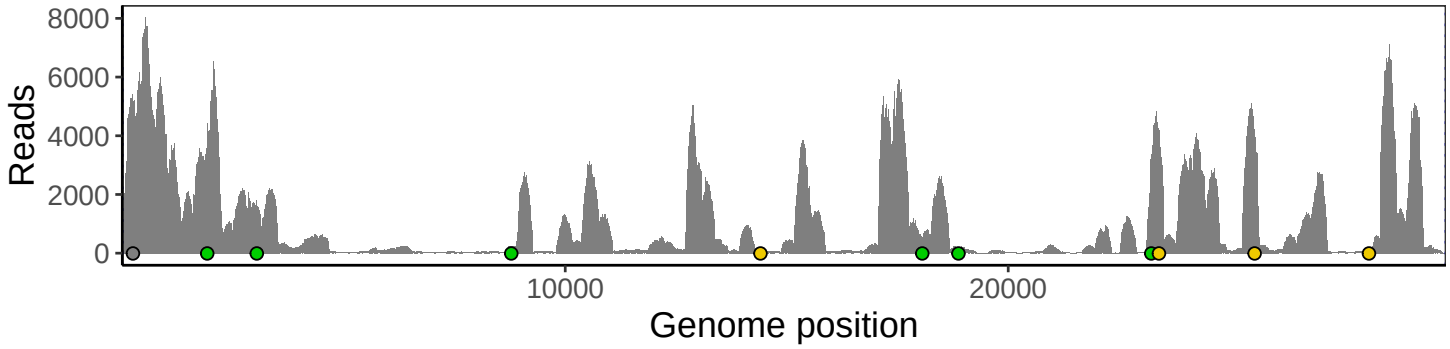


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

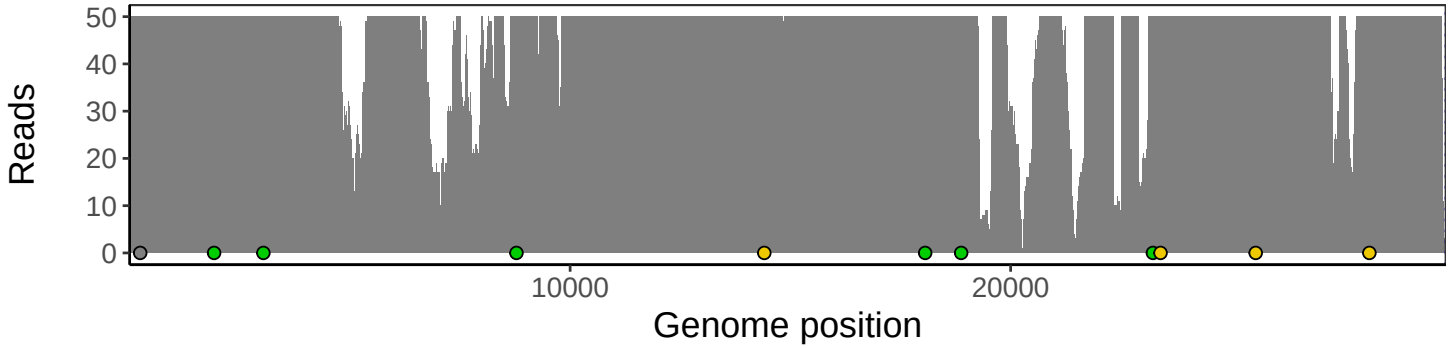


VSP0065-2 | 2020-05-04 | NP-OP | 251-q | 7550000 genomes | single experiment

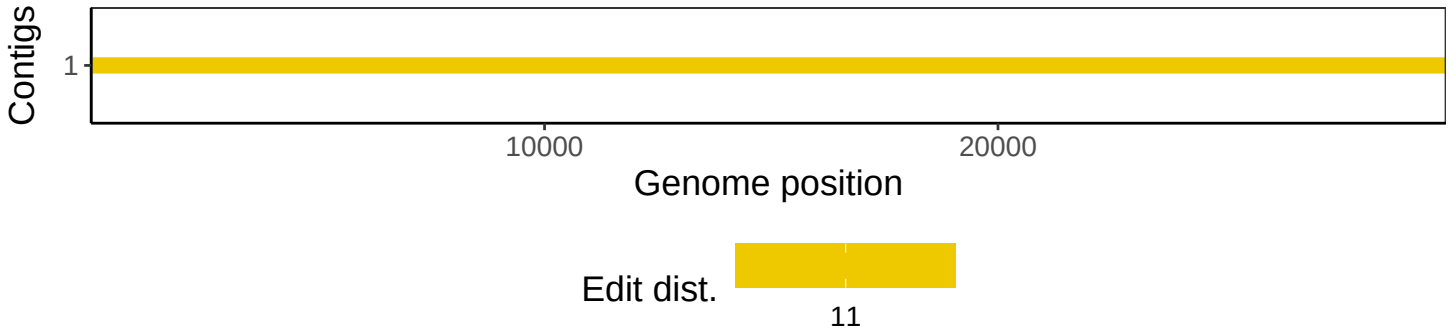
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

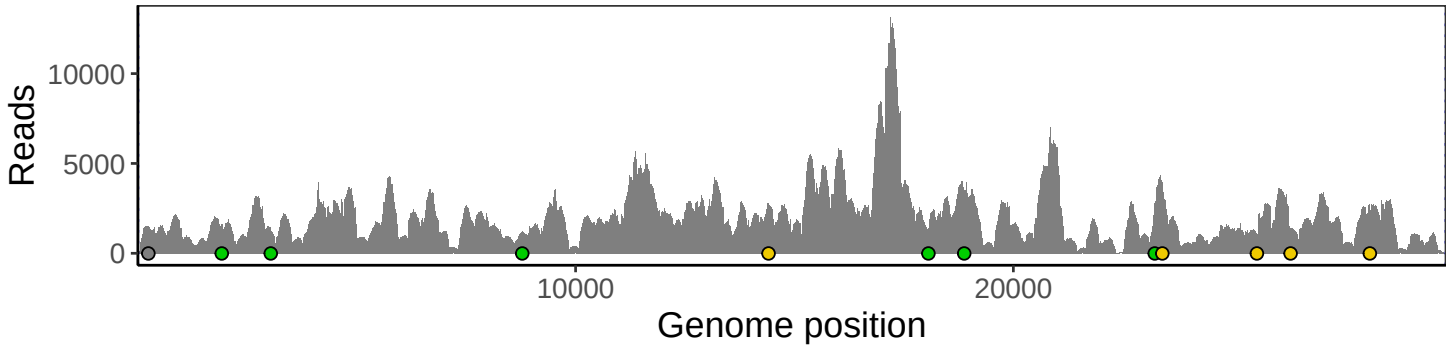


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

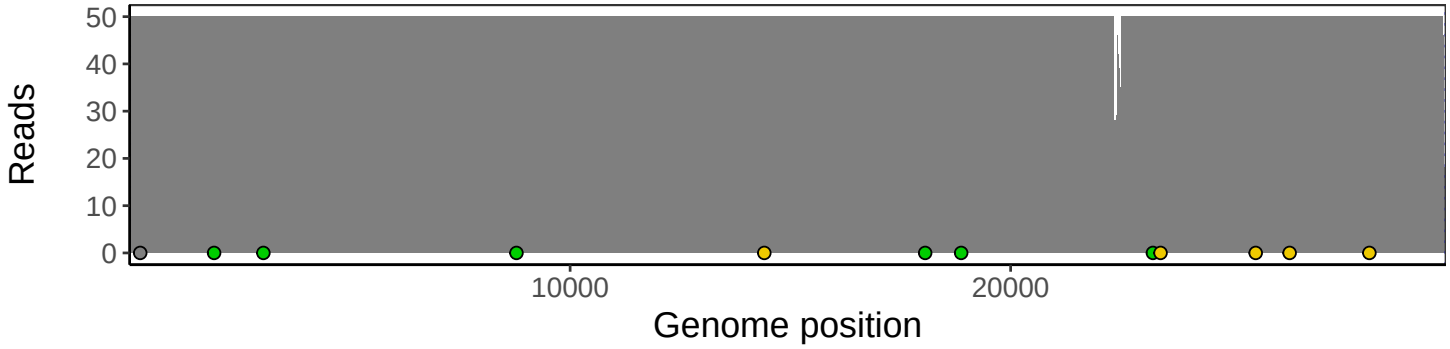


VSP0088-1 | 2020-05-06 | ETA | 251e-q | 255500 genomes | single experiment

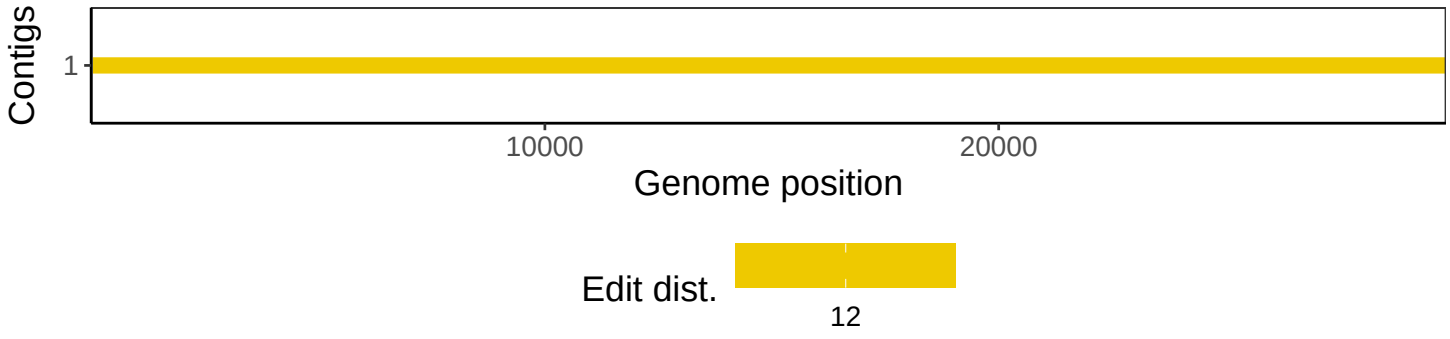
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

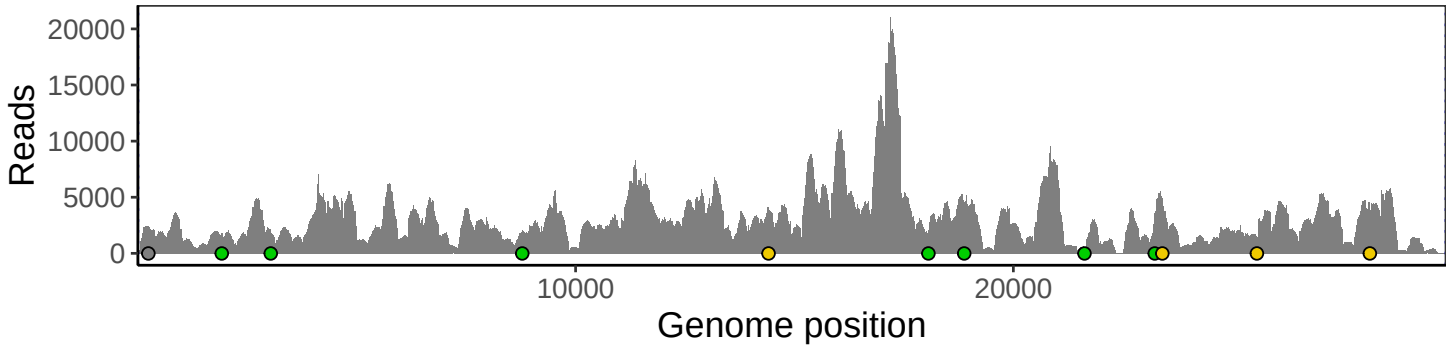


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

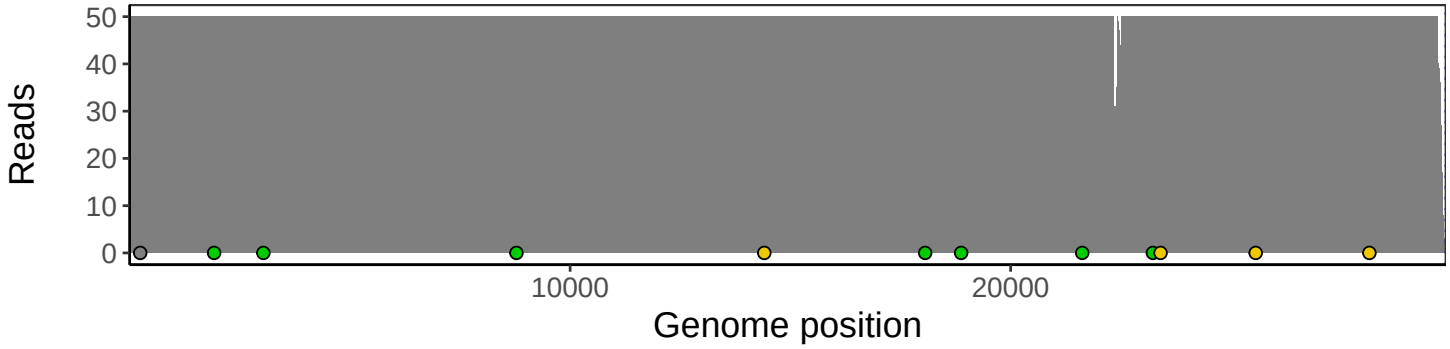


VSP0089-1 | 2020-05-06 | NP-OP | 251no-q | 570000 genomes | single experiment

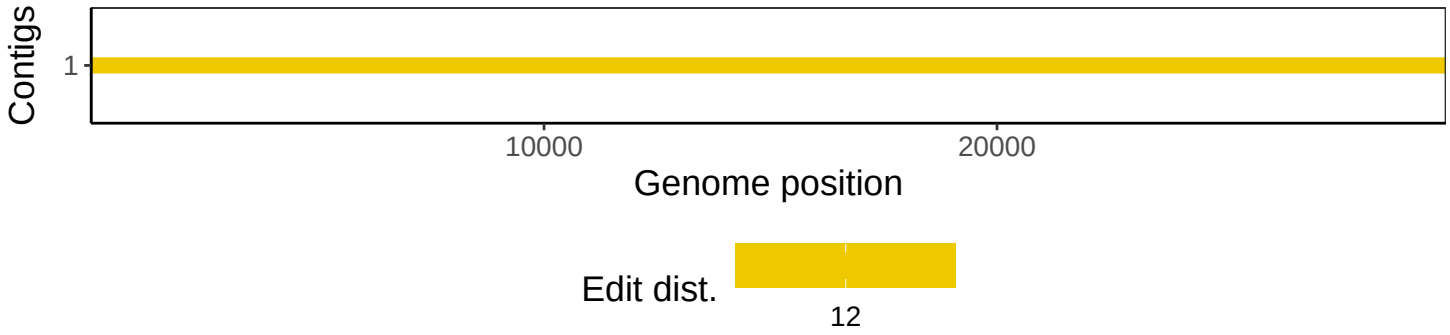
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

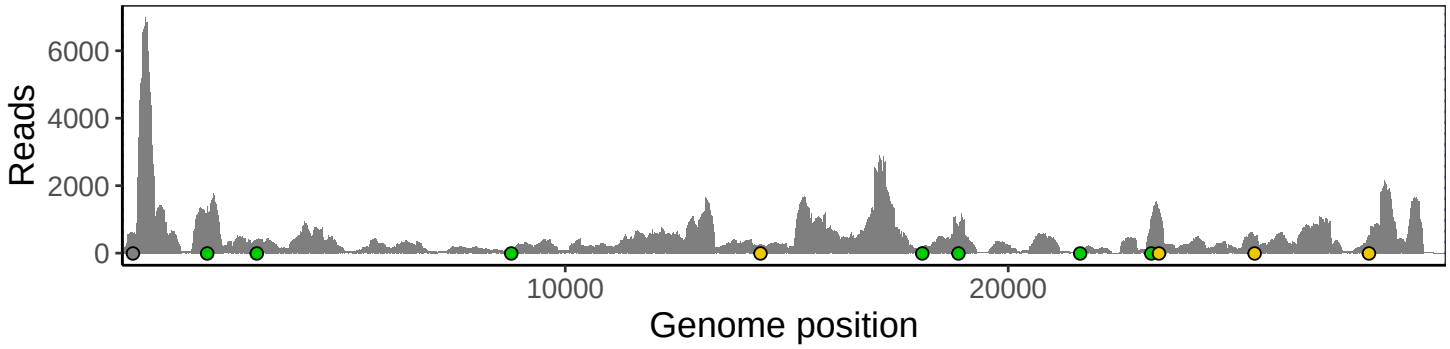


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

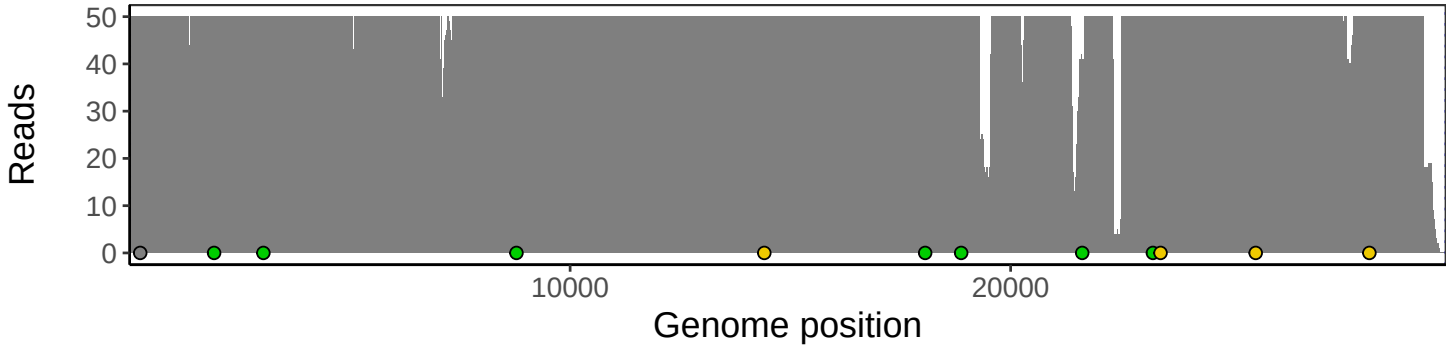


VSP0089-2 | 2020-05-06 | NP-OP | 251no-q | 570000 genomes | single experiment

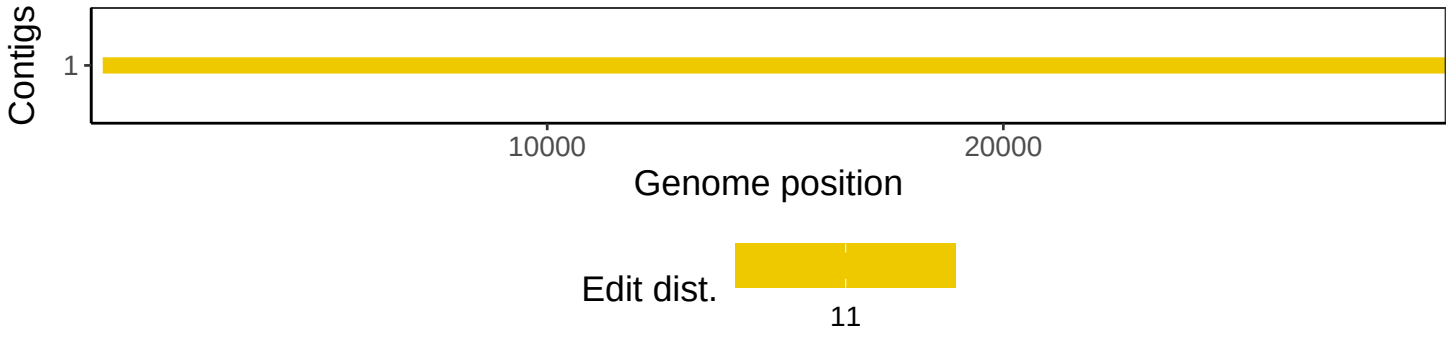
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

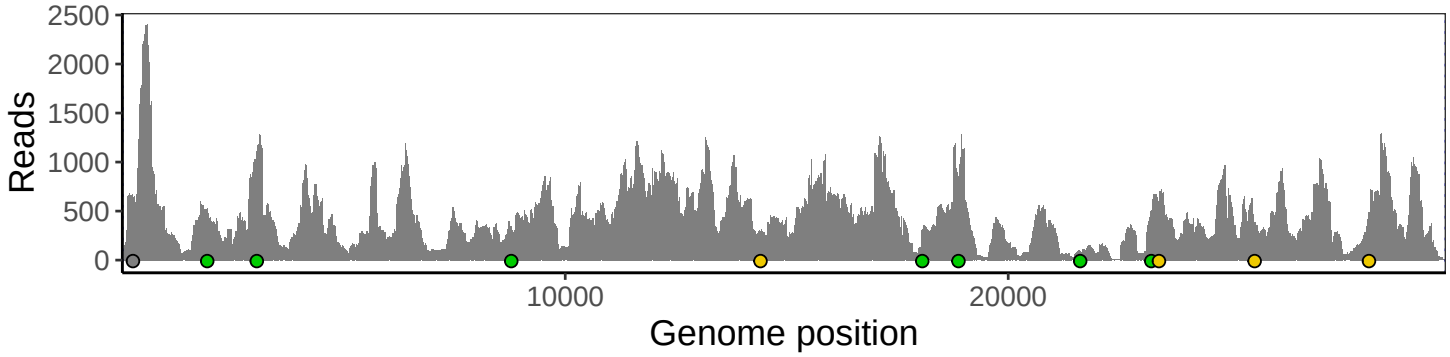


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

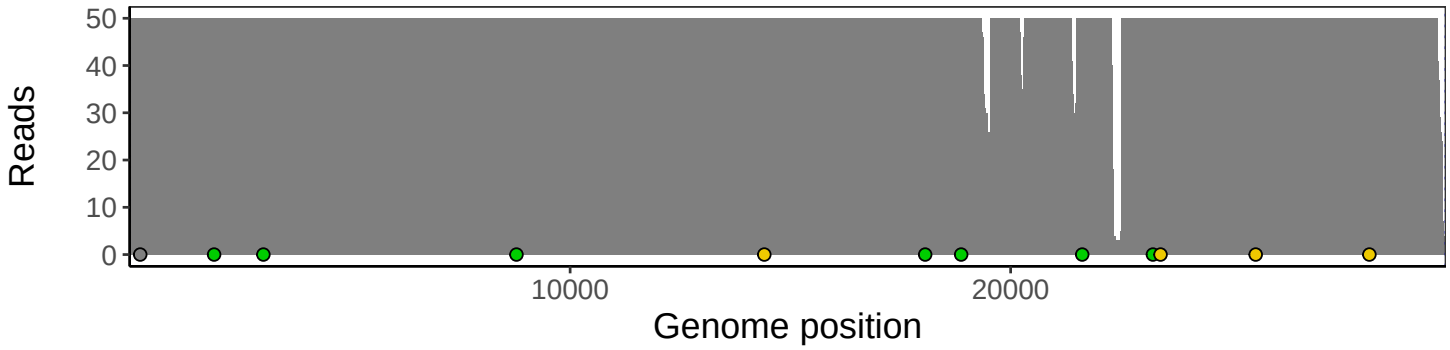


VSP0318-1 | 2020-05-06 | NP-OP | 251no-q | 890000 genomes | single experiment

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.