COVID-19 subject 228

2021-01-19

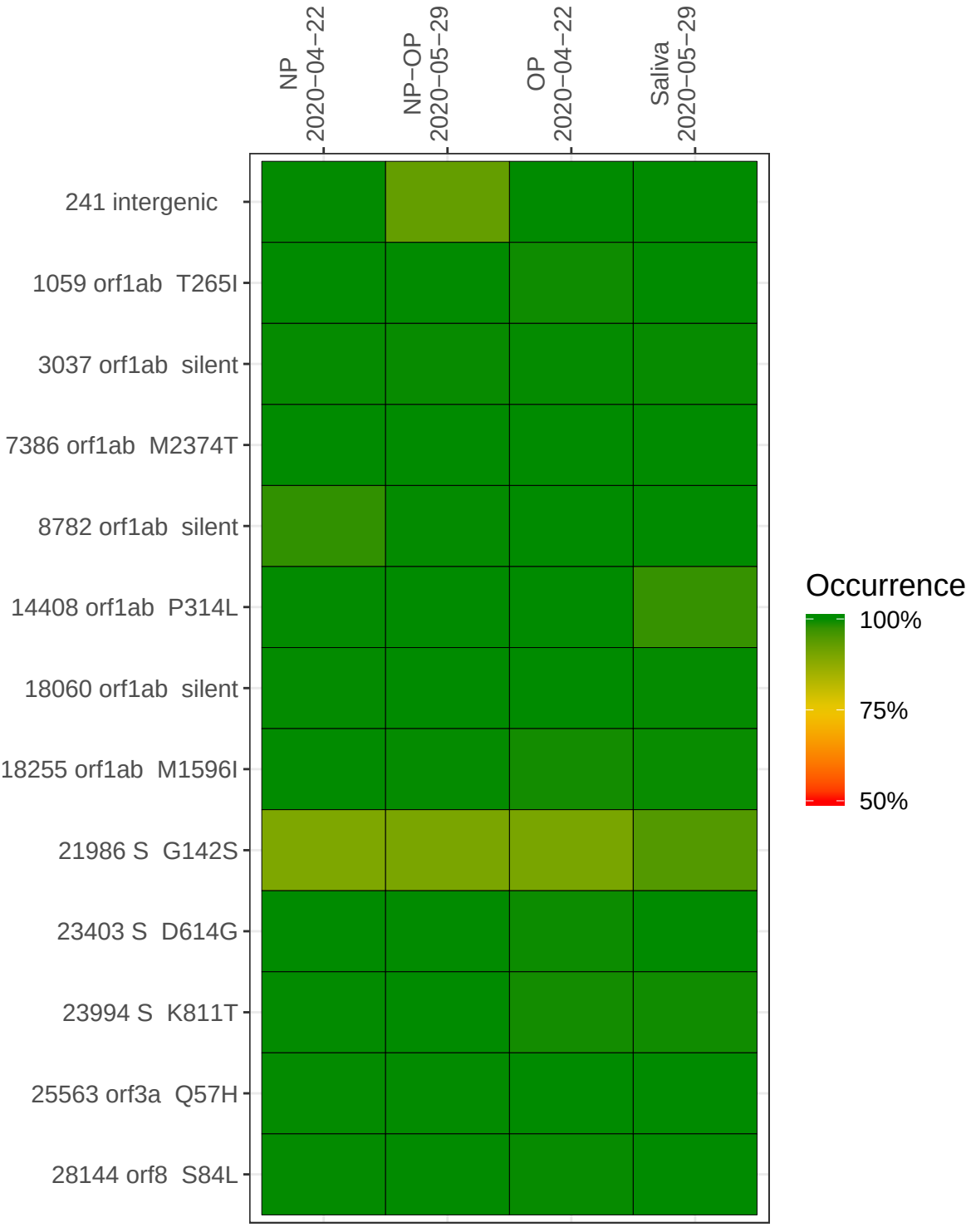
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

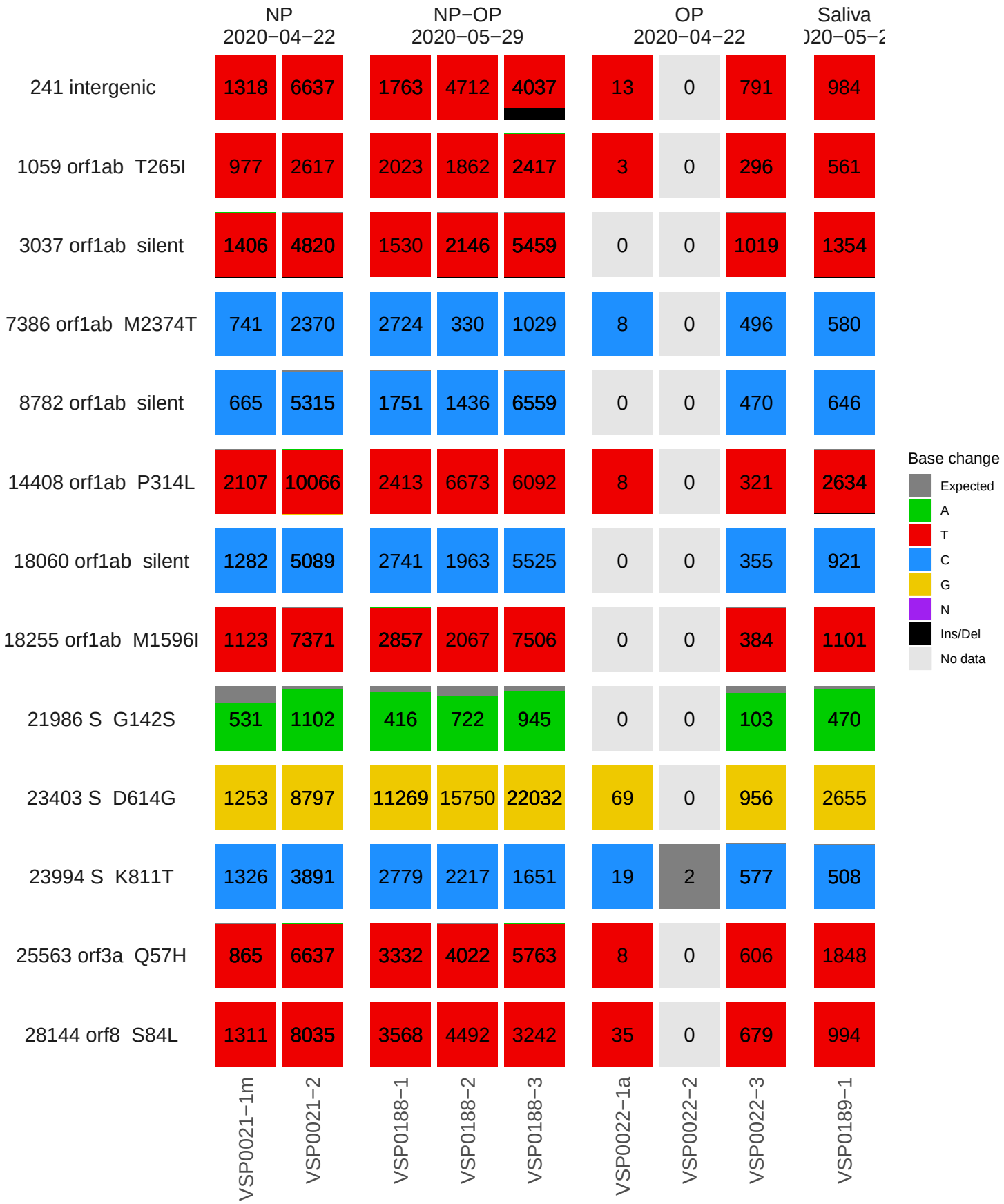
Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0021	composite	NA	NP	2020-04-22	29.82	99.8%	99.8%
VSP0022	composite	NA	OP	2020-04-22	29.68	99.8%	99.7%
VSP0187	composite	NA	ETA	2020-05-29	0.30	14.0%	0.0%
VSP0188	composite	NA	NP-OP	2020-05-29	29.90	99.8%	99.8%
VSP0021-1m	single experiment	NA	NP	2020-04-22	29.82	99.8%	99.7%
VSP0021-2	single experiment	1.10e+04	NP	2020-04-22	29.82	99.8%	99.8%
VSP0022-1a	single experiment	2.68e+05	OP	2020-04-22	1.80	81.4%	58.7%
VSP0022-2	single experiment	1.34e+06	OP	2020-04-22	0.24	6.2%	0.0%
VSP0022-3	single experiment	1.34e+06	OP	2020-04-22	29.68	99.8%	99.7%
VSP0187-1	single experiment	6.02e+01	ETA	2020-05-29	0.30	12.0%	0.0%
VSP0187-2	single experiment	3.01e+02	ETA	2020-05-29	NA	2.3%	0.0%
VSP0188-1	single experiment	2.04e+03	NP-OP	2020-05-29	9.60	93.7%	93.6%
VSP0188-2	single experiment	1.02e+04	NP-OP	2020-05-29	22.61	99.1%	99.1%
VSP0188-3	single experiment	1.02e+04	NP-OP	2020-05-29	29.87	99.8%	99.8%
VSP0189-1	single experiment	8.51e+04	Saliva	2020-05-29	29.82	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 or 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.

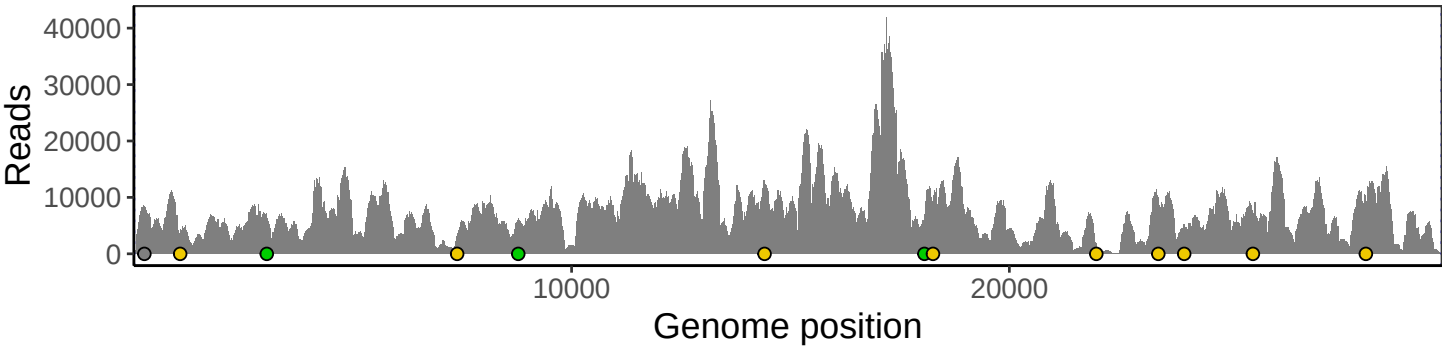



NP
2020−04−22
NP−OP
2020−05−29
OP
2020−04−22
Saliva
)20−05−2
241 intergenic
1059 orf1ab T265I
3037 orf1ab silent
7386 orf1ab M2374T
8782 orf1ab silent
14408 orf1ab P314L
18060 orf1ab silent
18255 orf1ab M1596I
21986 S G142S
23403 S D614G
23994 S K811T
25563 orf3a Q57H
28144 orf8 S84L
1318 6637 1763 4712 4037 13 0 791 984
977 2617 2023 1862 2417 3 0 296 561
1406 4820 1530 2146 5459 0 0 1019 1354
741 2370 2724 330 1029 8 0 496 580
665 5315 1751 1436 6559 0 0 470 646
2107 10066 2413 6673 6092 8 0 321 2634
1282 5089 2741 1963 5525 0 0 355 921
1123 7371 2857 2067 7506 0 0 384 1101
531 1102 416 722 945 0 0 103 470
1253 8797 11269 15750 22032 69 0 956 2655
1326 3891 2779 2217 1651 19 2 577 508
865 6637 3332 4022 5763 8 0 606 1848
1311 8035 3568 4492 3242 35 0 679 994
VSP0021−1m
VSP0021−2
VSP0188−1
VSP0188−2
VSP0188−3
VSP0022−1a
VSP0022−2
VSP0022−3
VSP0189−1
Base change
Expected
A
T
C
G
N
Ins/Del
No data

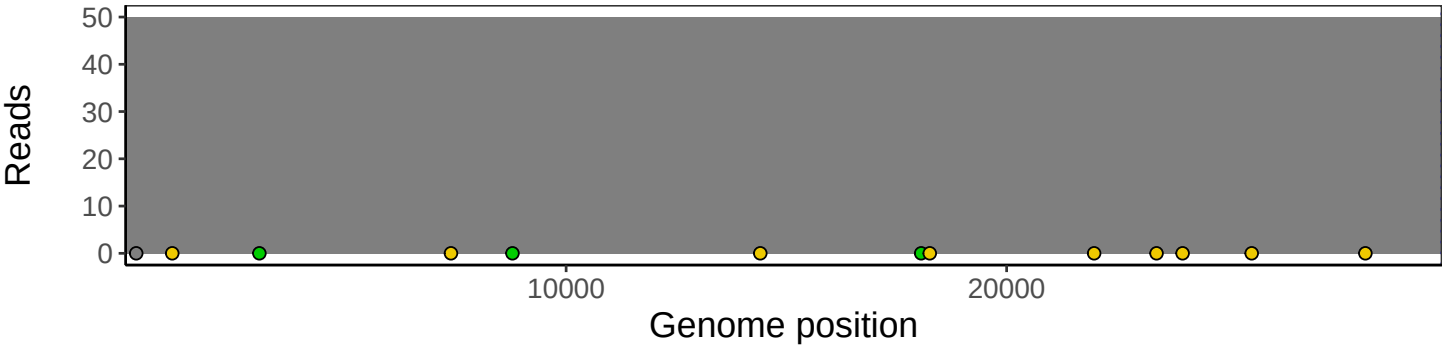
Analyses of individual experiments and composite results.

VSP0021 | 2020-04-22 | NP | 228n | composite result

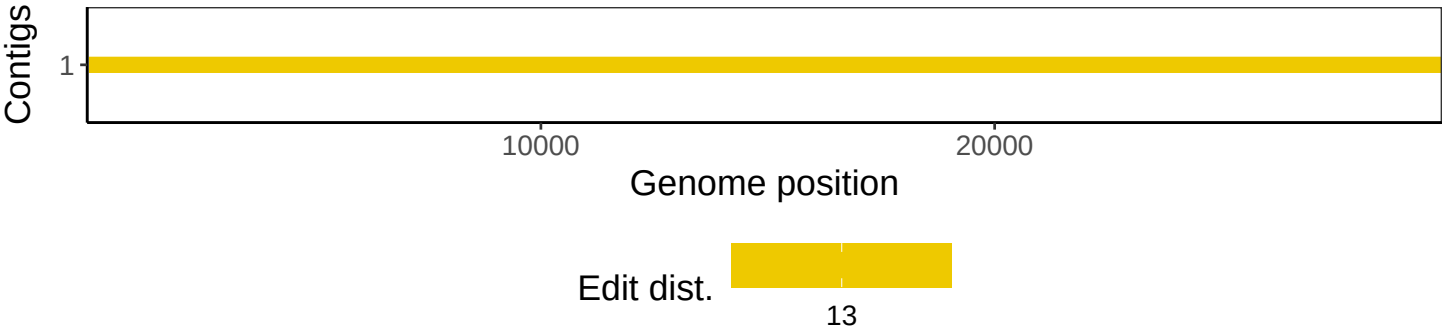
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

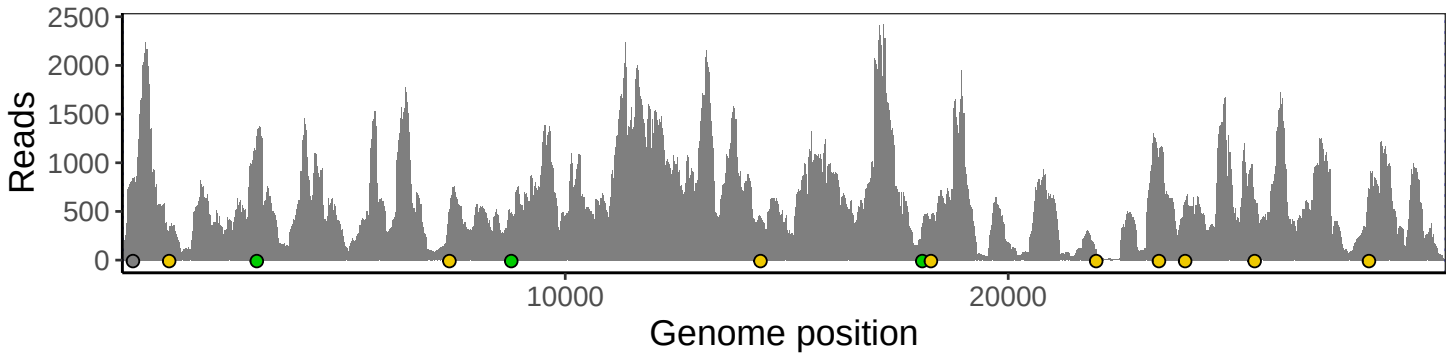


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

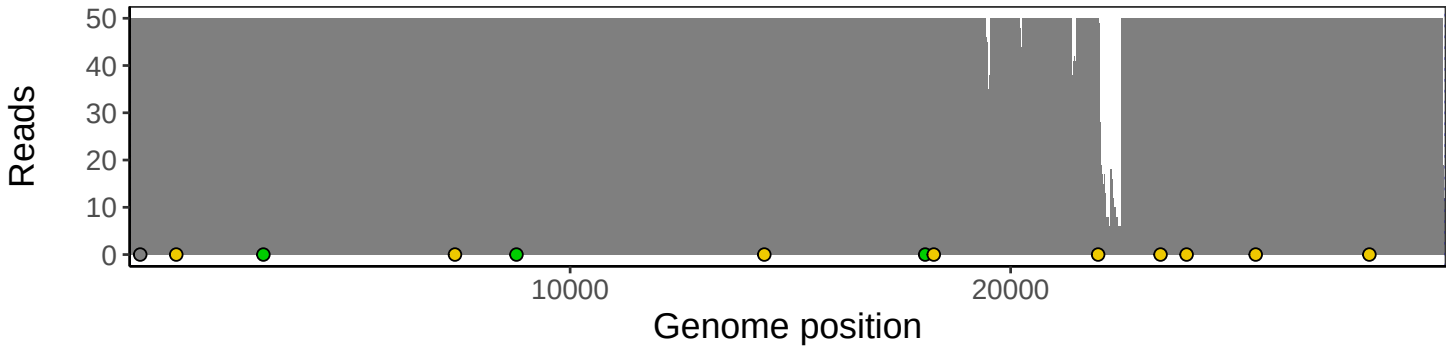


VSP0022 | 2020-04-22 | OP | 228o | composite result

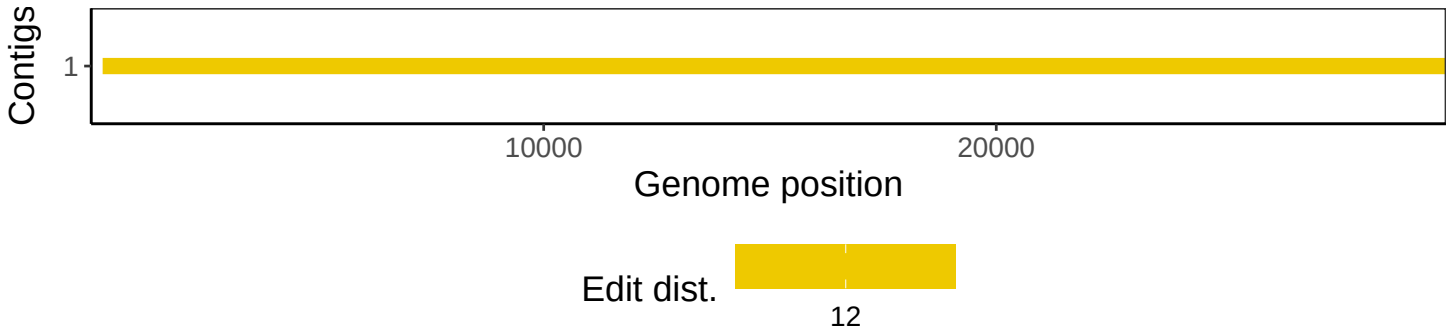
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

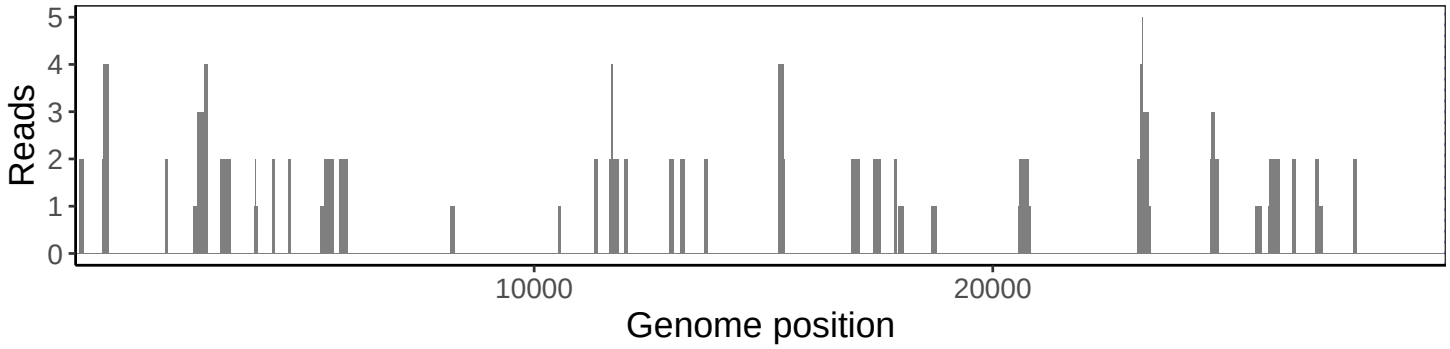


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

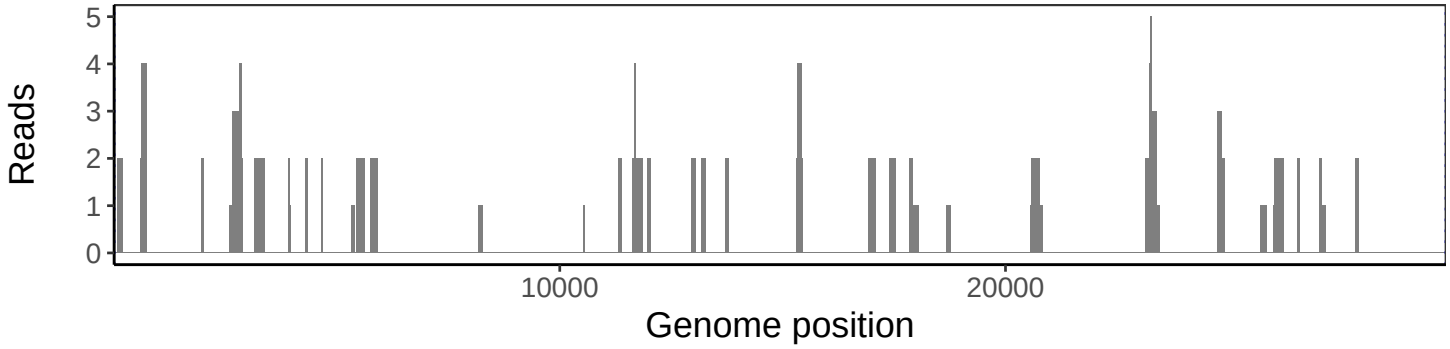


VSP0187 | 2020-05-29 | ETA | 228e-q | composite result

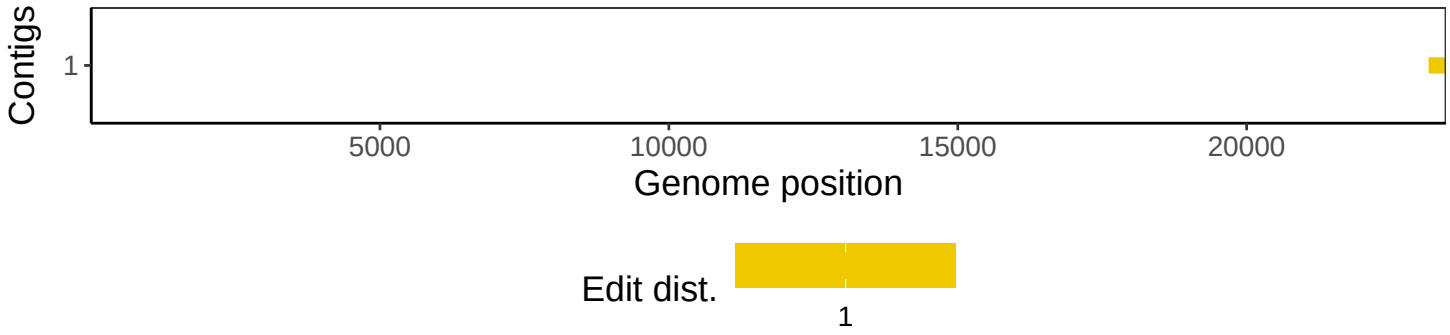
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

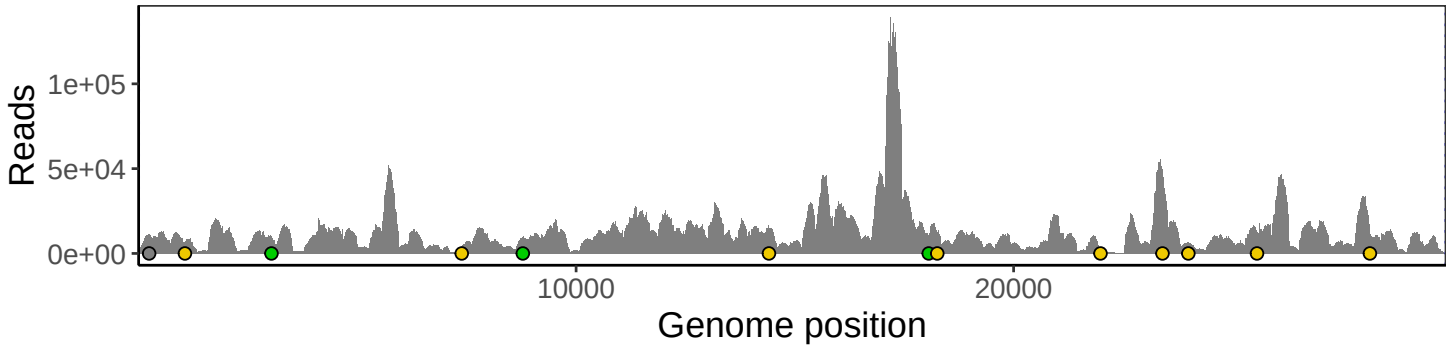


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

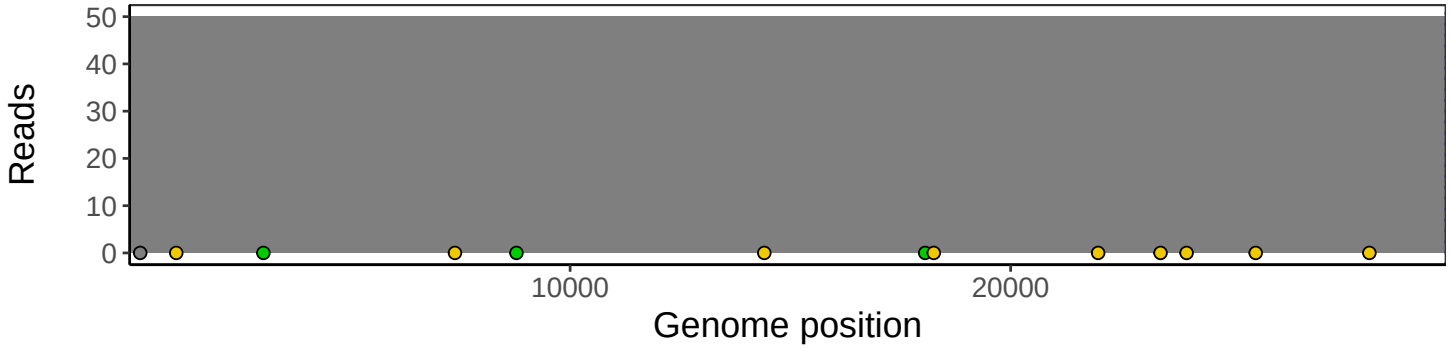


VSP0188 | 2020-05-29 | NP-OP | 228no-q | composite result

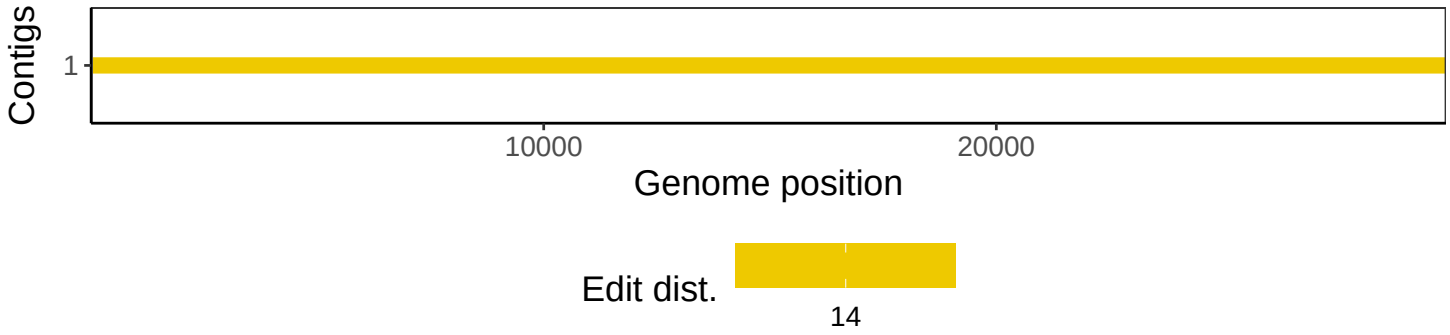
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

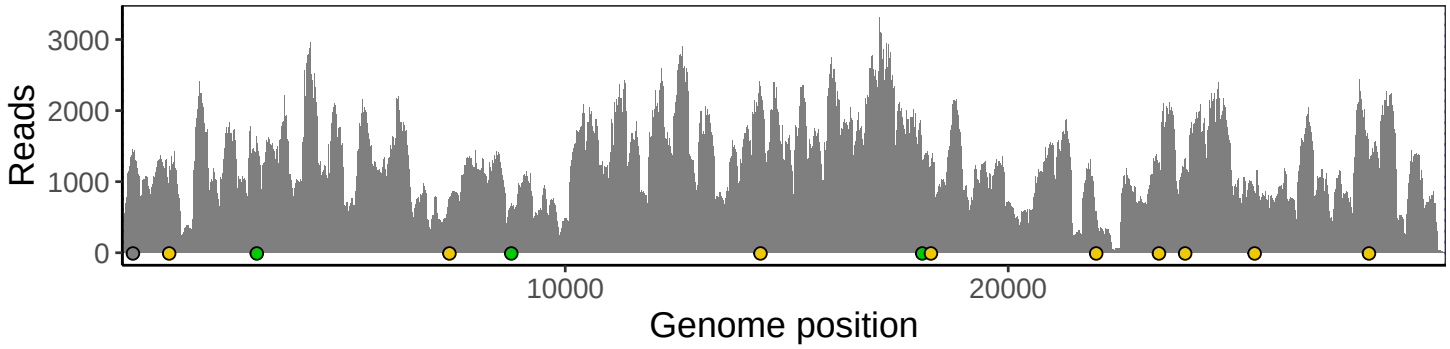


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

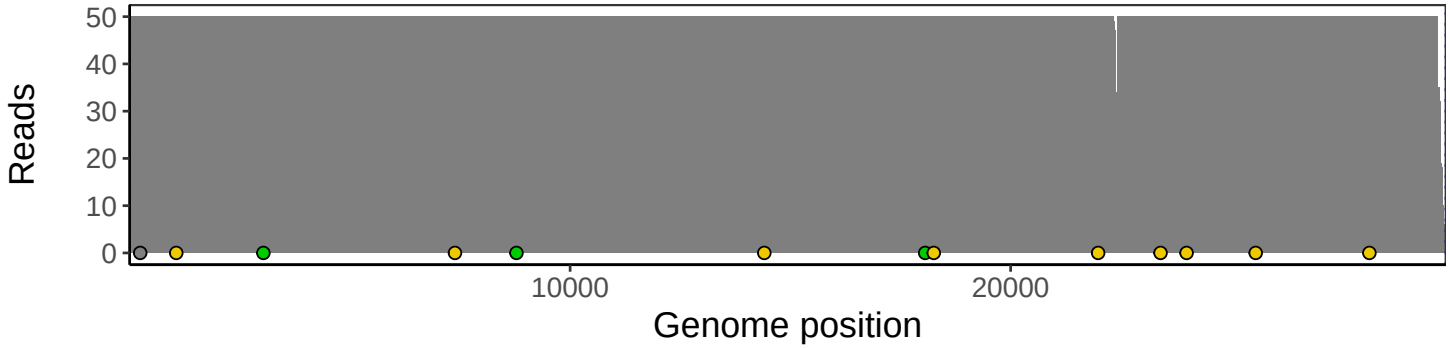


VSP0021-1m | 2020-04-22 | NP | 228n | genomes | single experiment

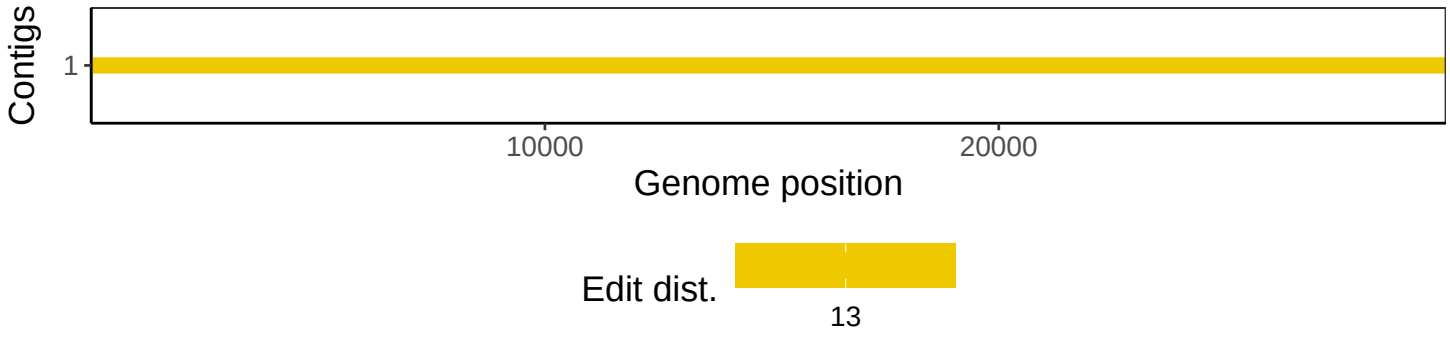
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

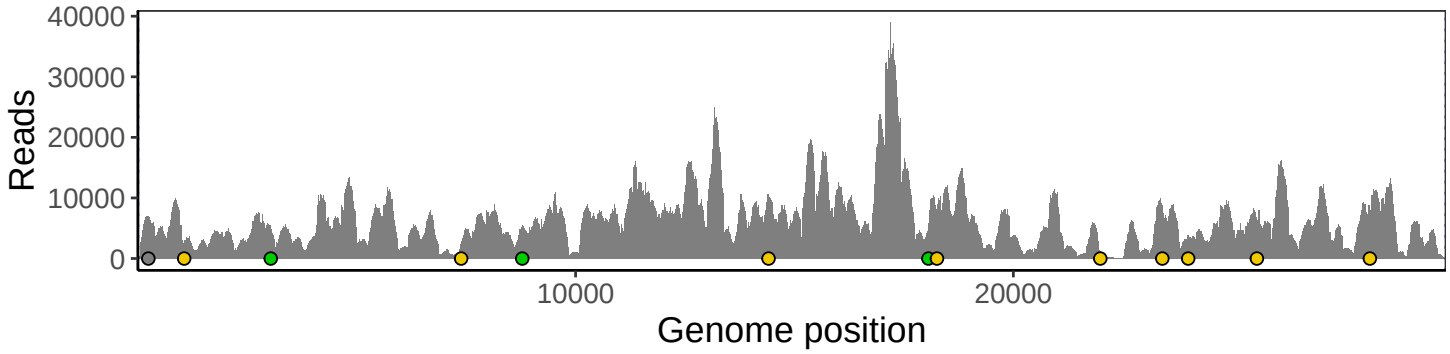


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

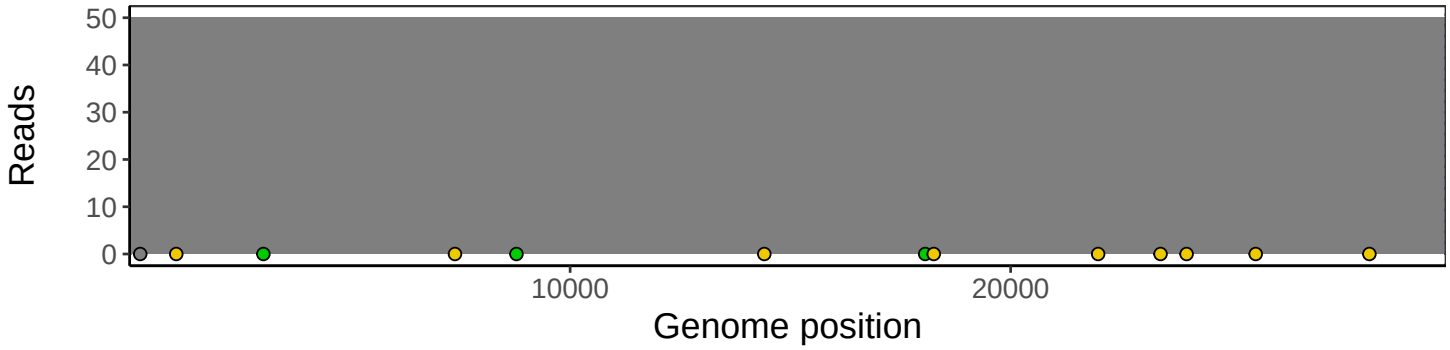


VSP0021-2 | 2020-04-22 | NP | 228n | 11000 genomes | single experiment

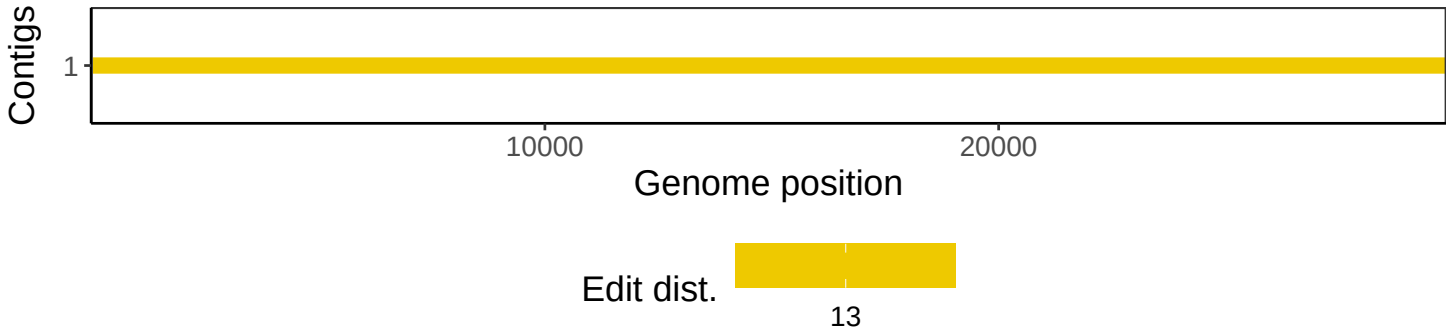
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

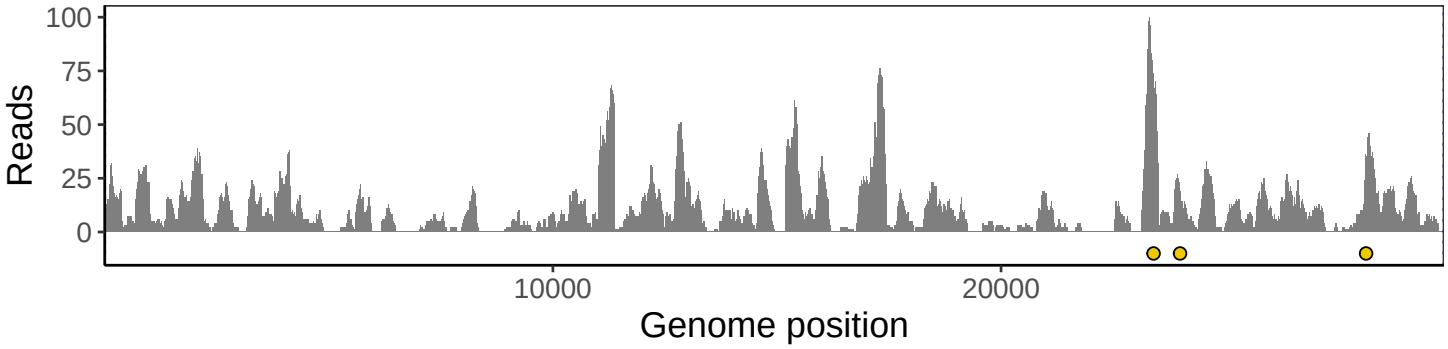


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

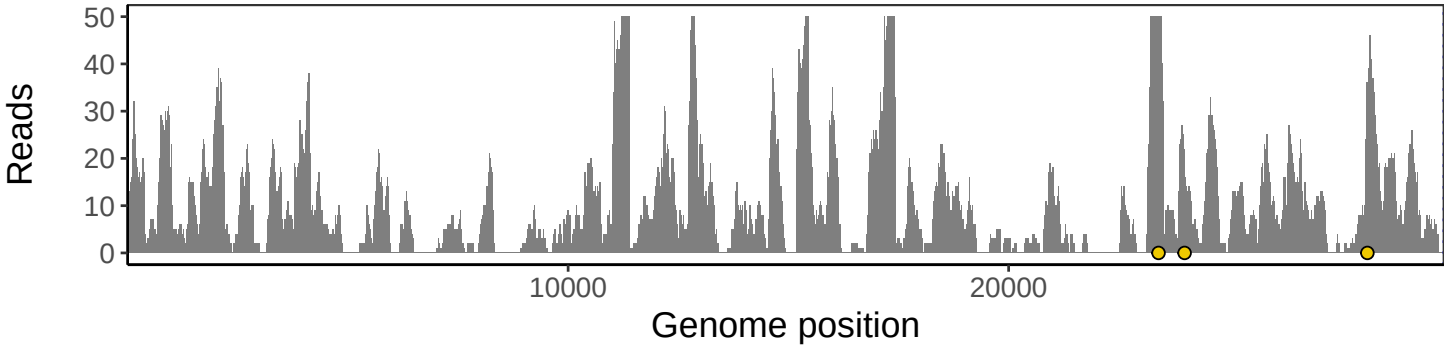


VSP0022-1a | 2020-04-22 | OP | 228o | 268000 genomes | single experiment

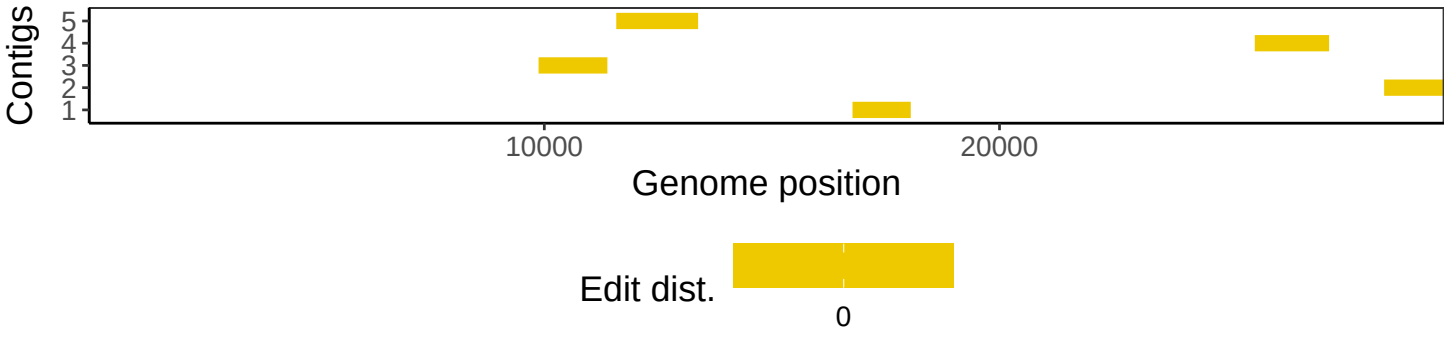
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

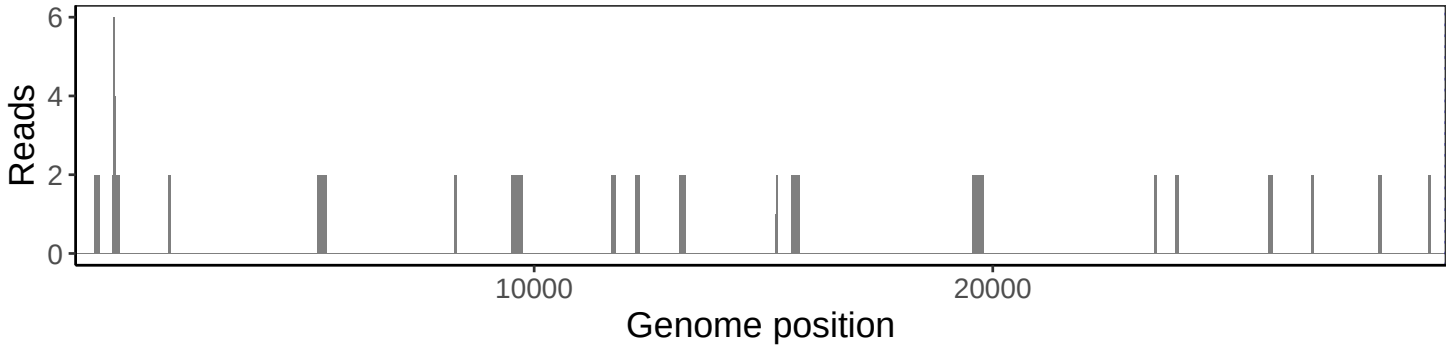


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

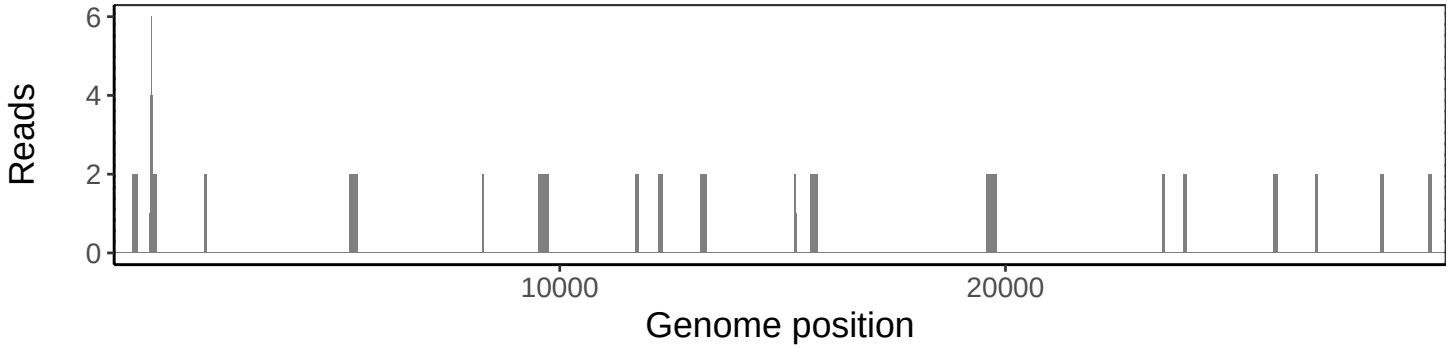


VSP0022-2 | 2020-04-22 | OP | 228o | 1340000 genomes | single experiment

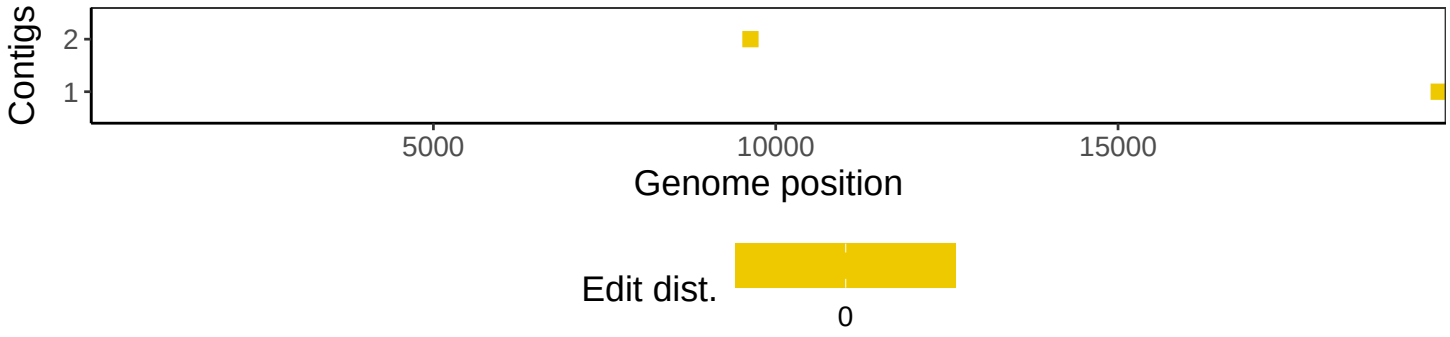
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

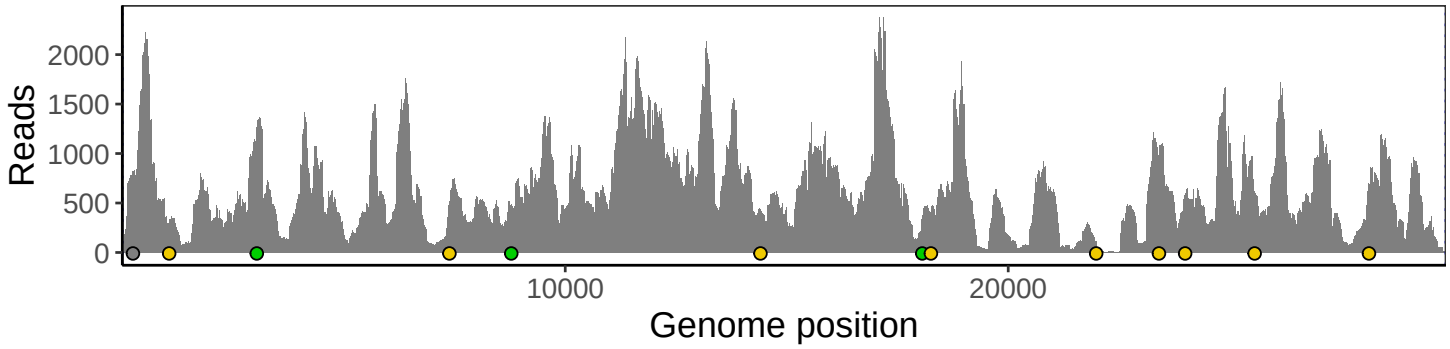


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

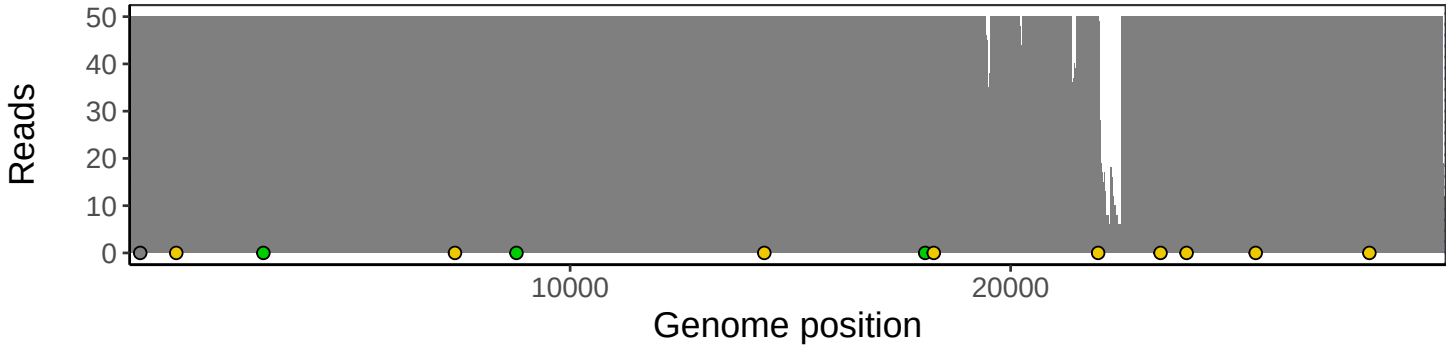


VSP0022-3 | 2020-04-22 | OP | 228o | 1340000 genomes | single experiment

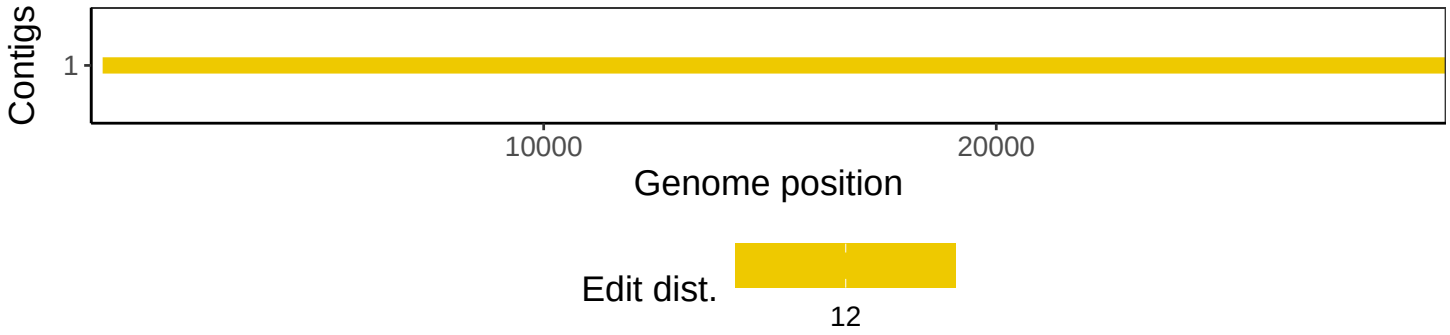
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

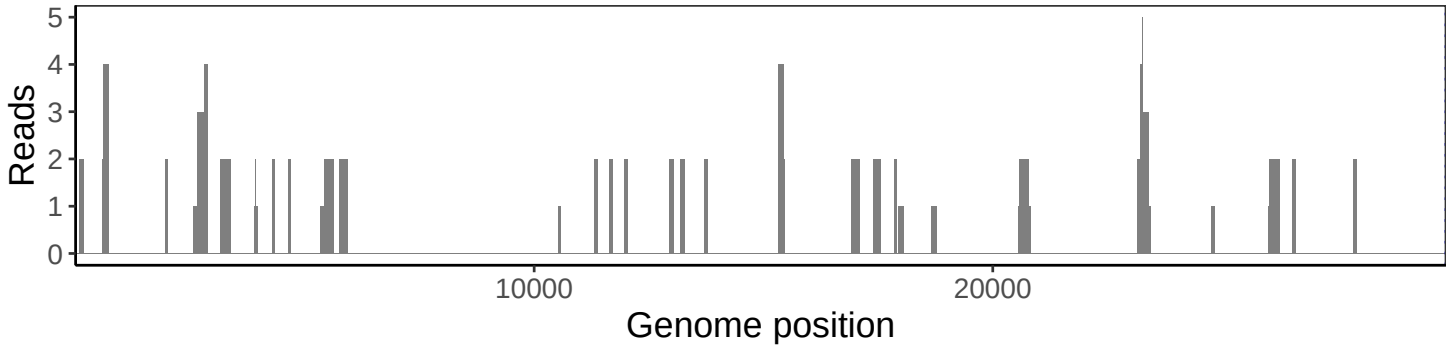


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

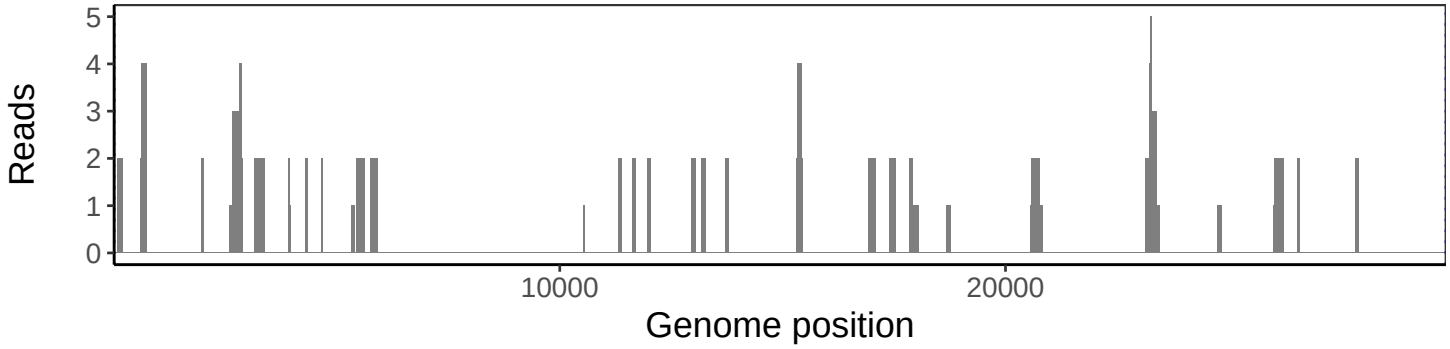


VSP0187-1 | 2020-05-29 | ETA | 228e-q | 60.2 genomes | single experiment

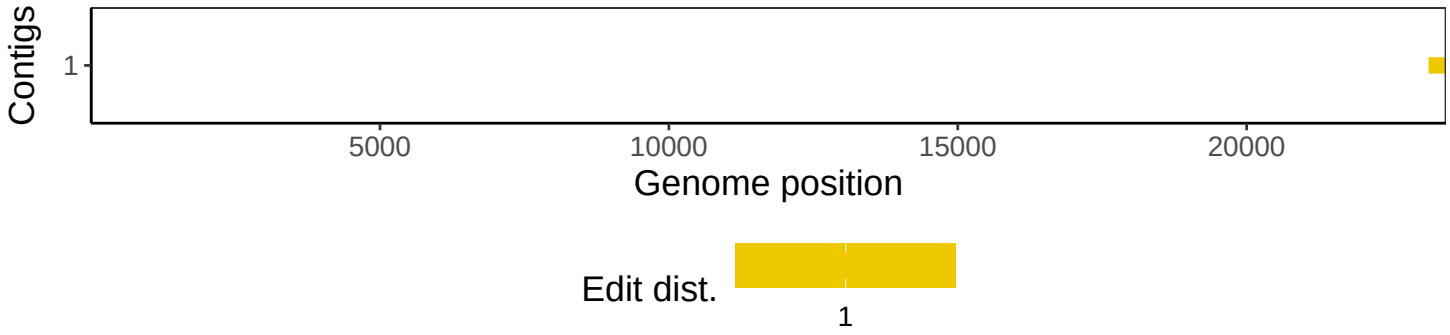
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

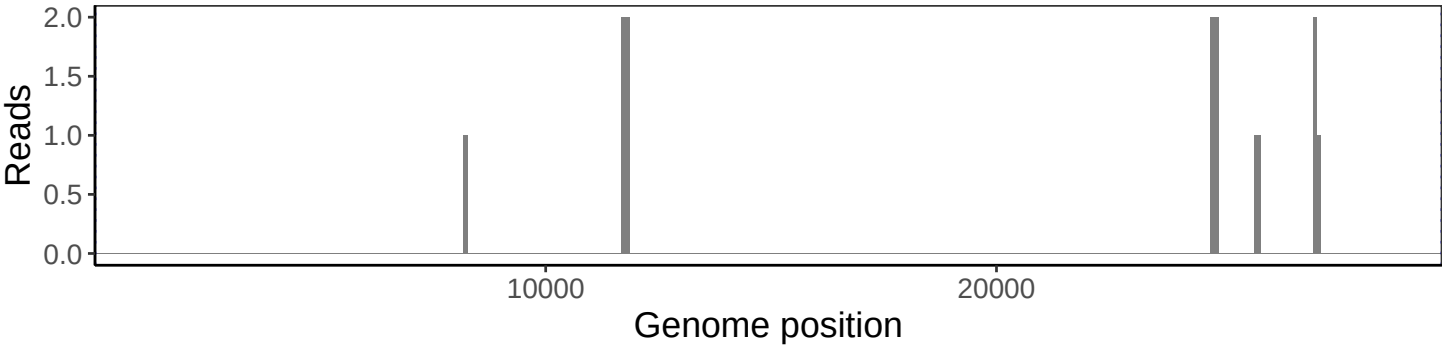


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

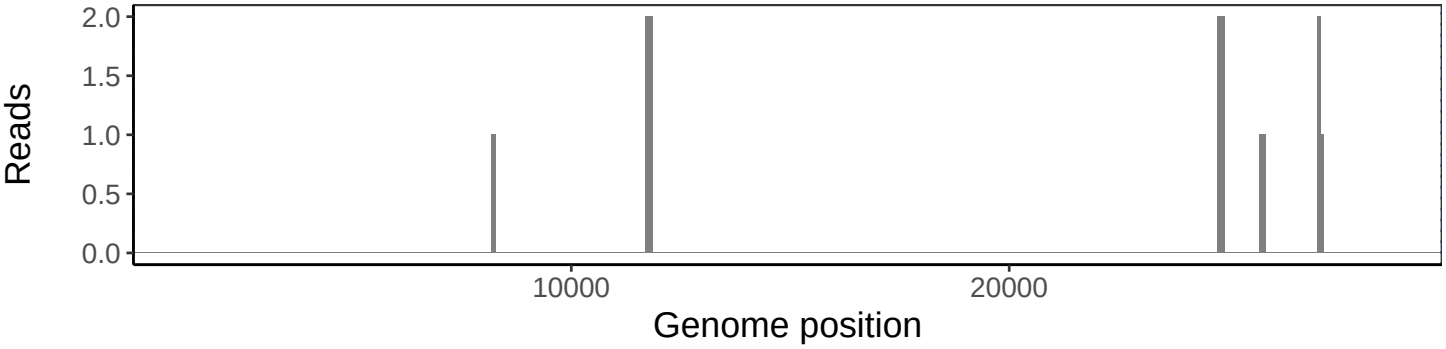


VSP0187-2 | 2020-05-29 | ETA | 228e-q | 301 genomes | single experiment

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.



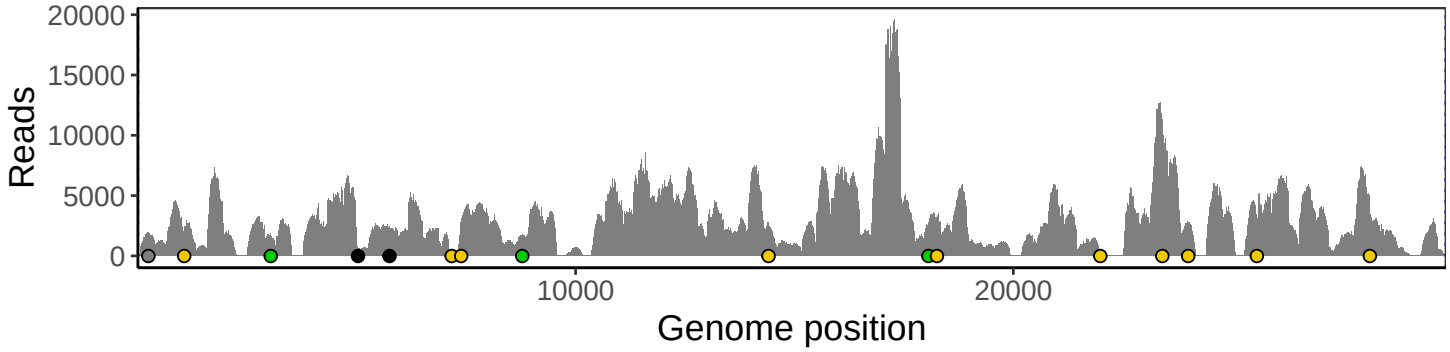
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



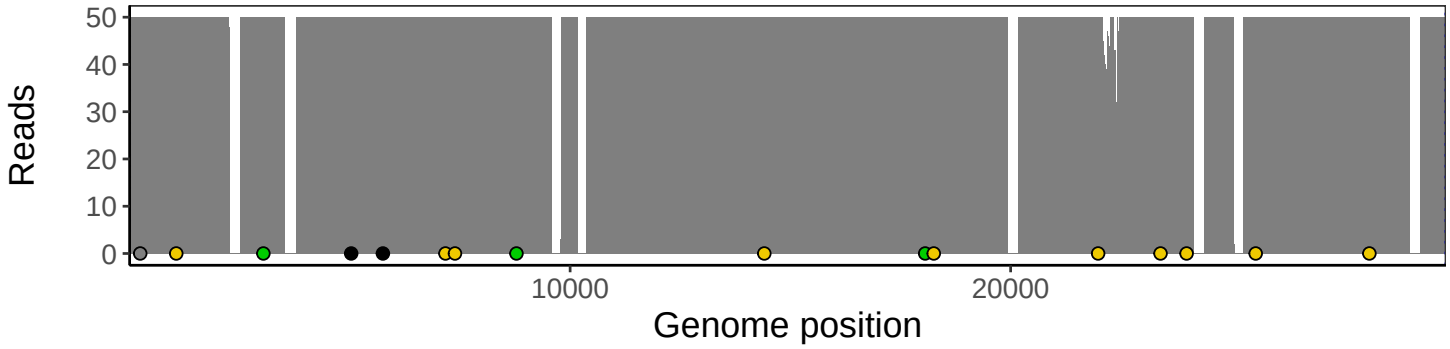
No contig data available.

VSP0188-1 | 2020-05-29 | NP-OP | 228no-q | 2040 genomes | single experiment

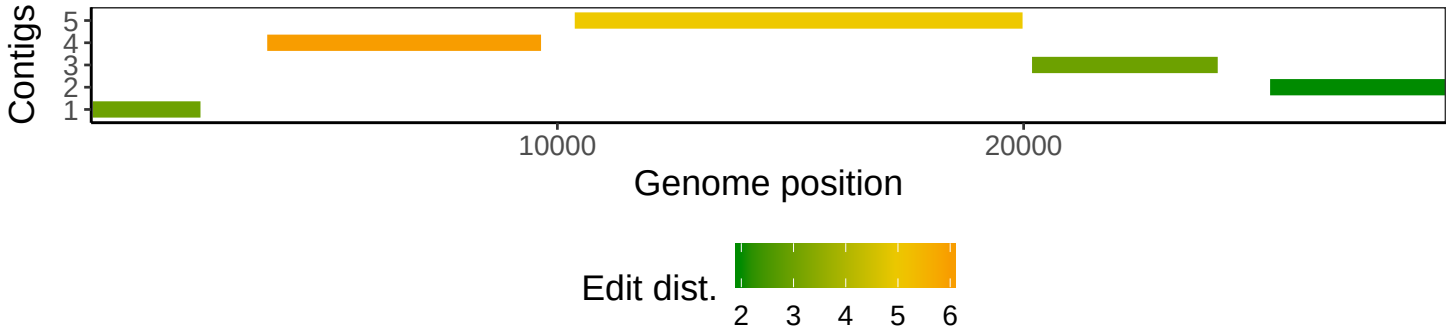
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

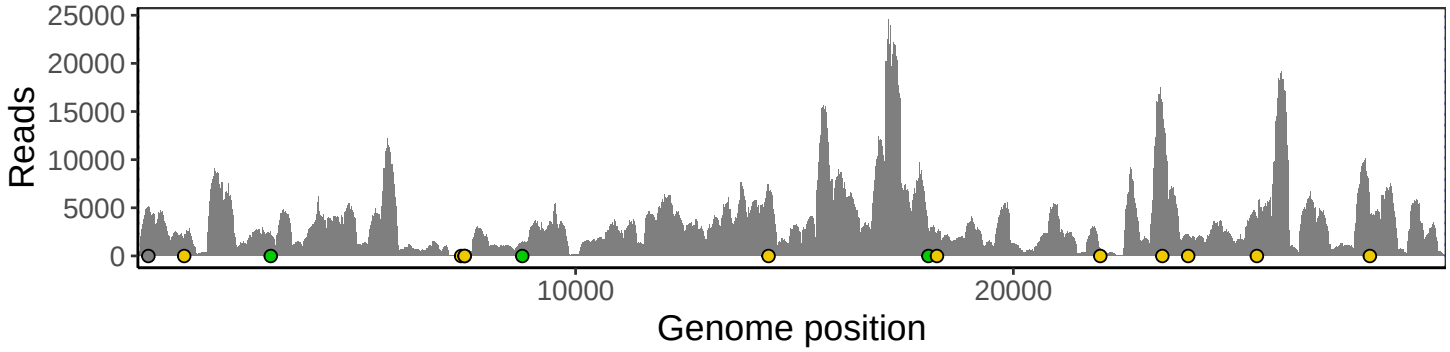


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

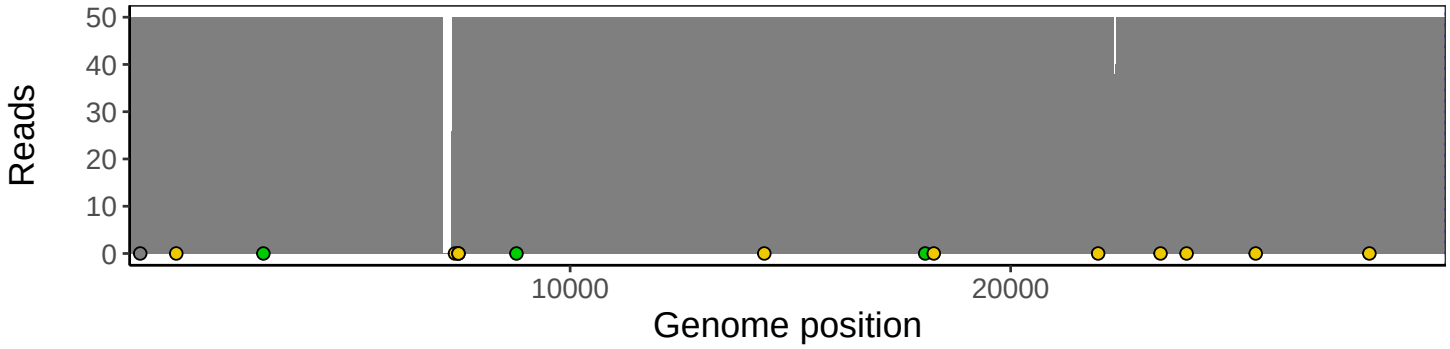


VSP0188-2 | 2020-05-29 | NP-OP | 228no-q | 10200 genomes | single experiment

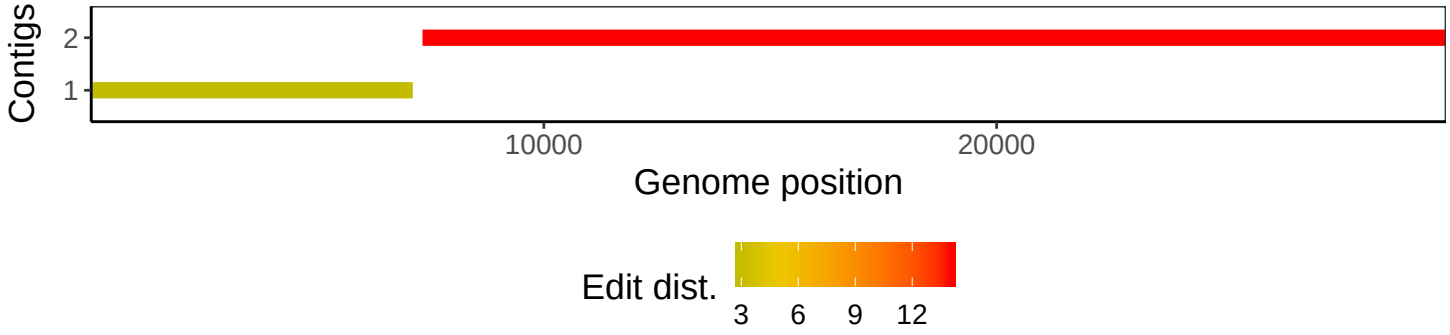
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

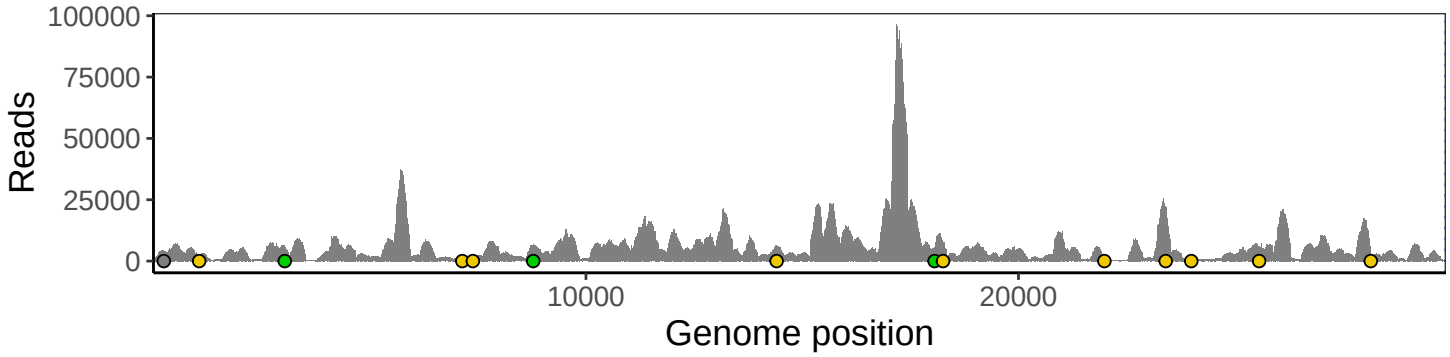


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

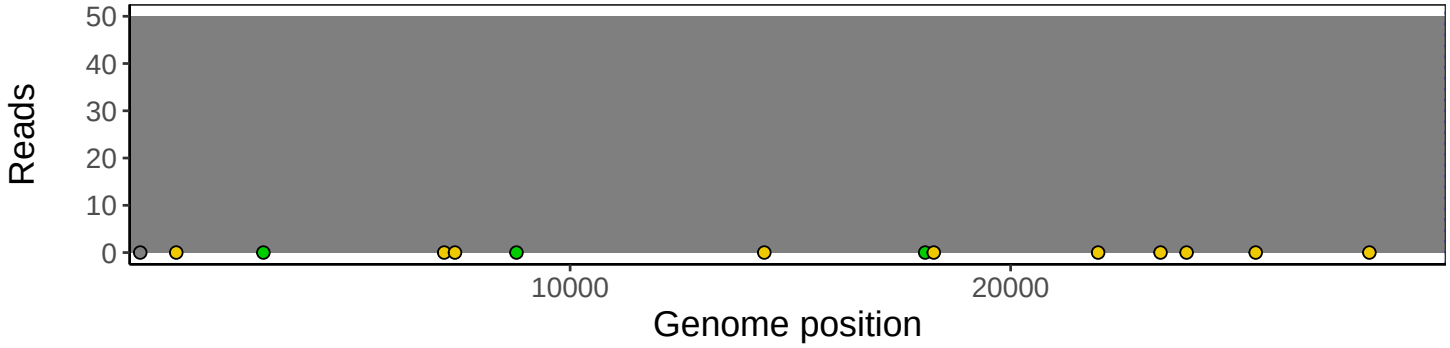


VSP0188-3 | 2020-05-29 | NP-OP | 228no-q | 10200 genomes | single experiment

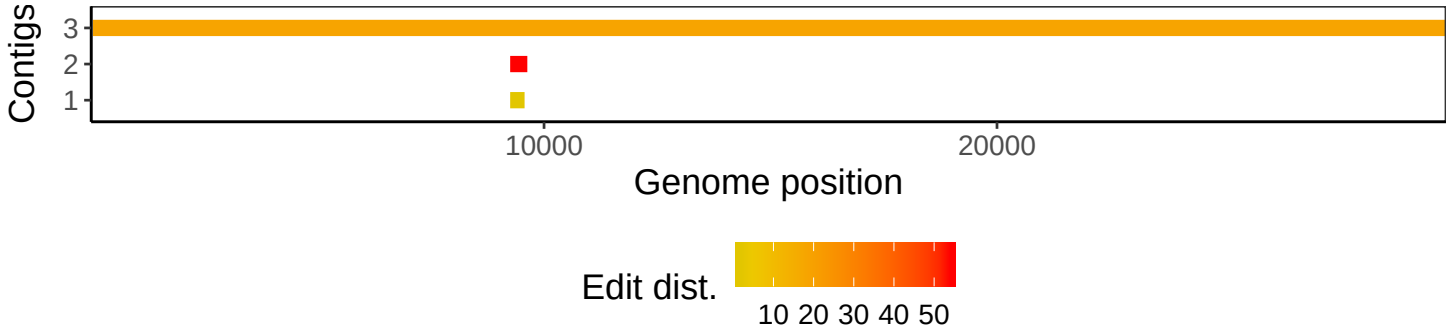
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

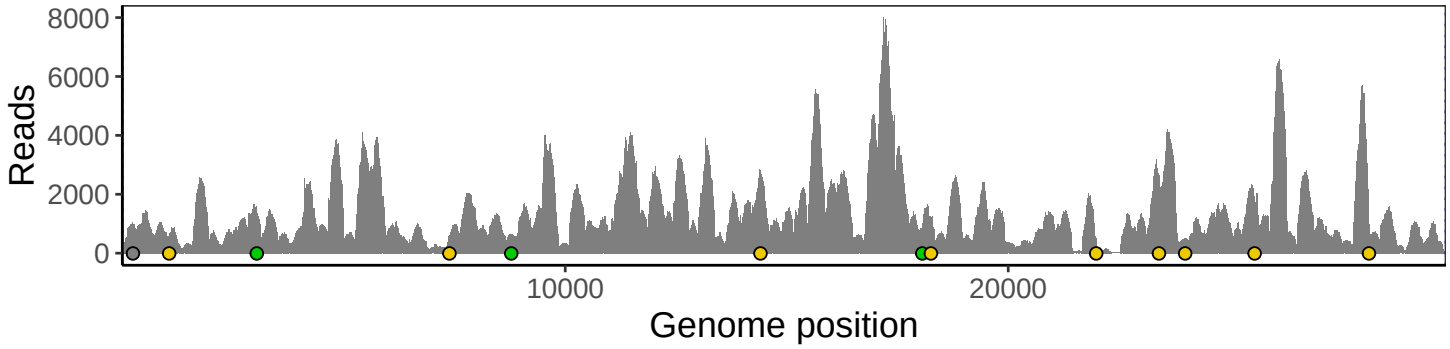


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

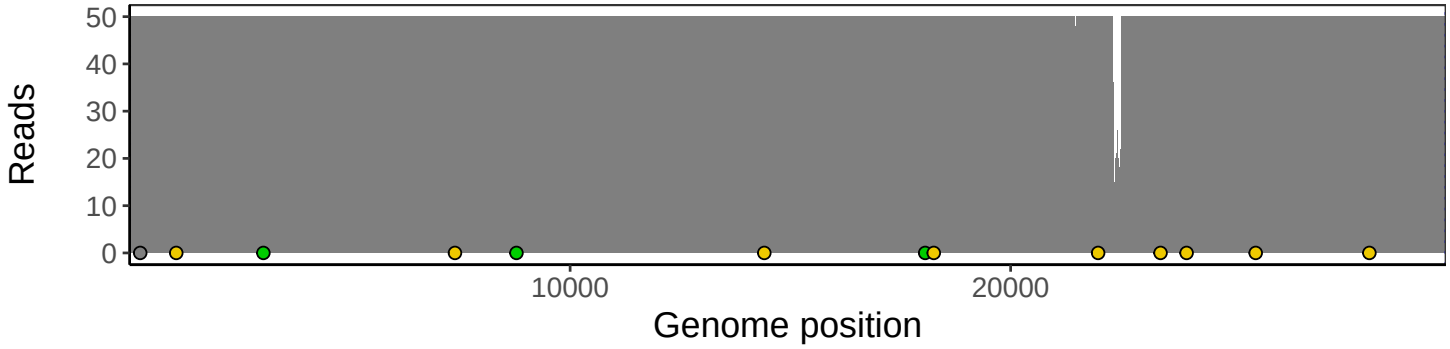


VSP0189-1 | 2020-05-29 | Saliva | 228s-q | 85100 genomes | single experiment

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.