# COVID-19 subject 228

*2021-01-08*

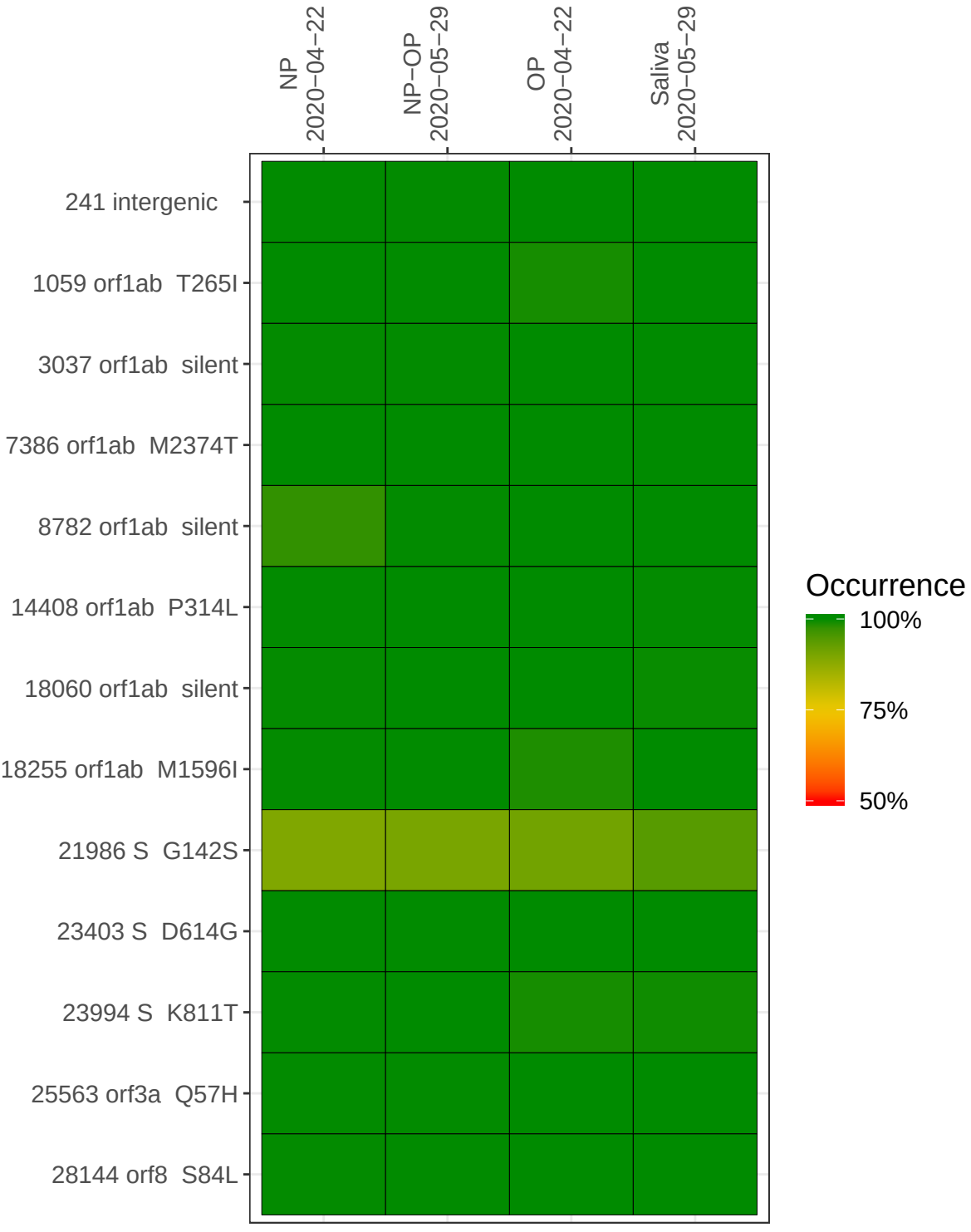
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

| Experiment | Type | Input genomes | Sample type | Sample date | Largest contig (KD) | Reference read coverage | Reference read coverage (>= 5 reads) |
|---|---|---|---|---|---|---|---|
| VSP0021 | composite | NA | NP | 2020-04-22 | 29.82 | 100.0% | 99.8% |
| VSP0022 | composite | NA | OP | 2020-04-22 | 29.68 | 100.0% | 98.8% |
| VSP0187 | composite | NA | ETA | 2020-05-29 | 0.30 | 100.0% | 0.0% |
| VSP0188 | composite | NA | NP-OP | 2020-05-29 | 29.90 | 100.0% | 99.8% |
| VSP0021-1m | single experiment | NA | NP | 2020-04-22 | 29.82 | 100.0% | 99.7% |
| VSP0021-2 | single experiment | 1.10e+04 | NP | 2020-04-22 | 29.82 | 100.0% | 99.8% |
| VSP0022-1a | single experiment | 2.68e+05 | OP | 2020-04-22 | 1.80 | 100.0% | 44.7% |
| VSP0022-2 | single experiment | 1.34e+06 | OP | 2020-04-22 | 0.24 | 100.0% | 0.0% |
| VSP0022-3 | single experiment | 1.34e+06 | OP | 2020-04-22 | 29.68 | 100.0% | 98.8% |
| VSP0187-1 | single experiment | 6.02e+01 | ETA | 2020-05-29 | 0.30 | 100.0% | 0.0% |
| VSP0187-2 | single experiment | 3.01e+02 | ETA | 2020-05-29 | NA | 100.0% | 0.0% |
| VSP0188-1 | single experiment | 2.04e+03 | NP-OP | 2020-05-29 | 9.60 | 100.0% | 93.5% |
| VSP0188-2 | single experiment | 1.02e+04 | NP-OP | 2020-05-29 | 22.61 | 100.0% | 99.1% |
| VSP0188-3 | single experiment | 1.02e+04 | NP-OP | 2020-05-29 | 29.87 | 100.0% | 99.8% |
| VSP0189-1 | single experiment | 8.51e+04 | Saliva | 2020-05-29 | 29.82 | 100.0% | 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.

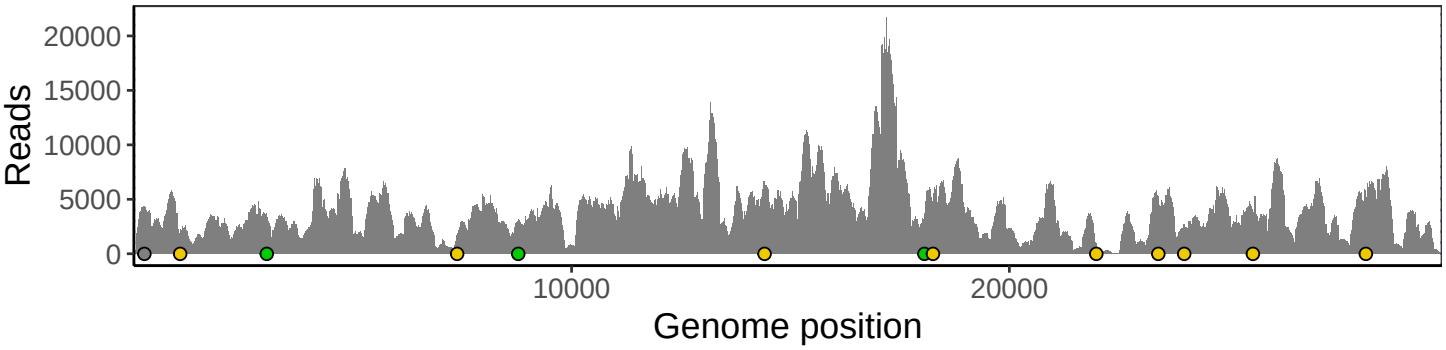



NP
2020−04−22
NP−OP
2020−05−29
OP
2020−04−22
Saliva
)20−05−2
241 intergenic
1059 orf1ab T265I
3037 orf1ab silent
7386 orf1ab M2374T
8782 orf1ab silent
14408 orf1ab P314L
18060 orf1ab silent
18255 orf1ab M1596I
21986 S G142S
23403 S D614G
23994 S K811T
25563 orf3a Q57H
28144 orf8 S84L
715 3376 952 2393 2054 10 429 525
526 1367 1084 967 1267 2 159 300
761 2478 808 1092 2805 526 718
399 1272 1422 171 561 4 259 320
358 2690 927 727 3339 244 338
1126 5107 1254 3384 3087 4 165 1362
687 2681 1451 1013 2892 200 493
597 3806 1484 1056 3890 203 579
285 583 225 374 521 56 260
683 4556 5997 8078 11471 43 516 1428
732 1981 1462 1130 843 11 1 305 270
470 3550 1782 2110 3106 5 324 1016
714 4499 1899 2377 1800 23 355 572
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

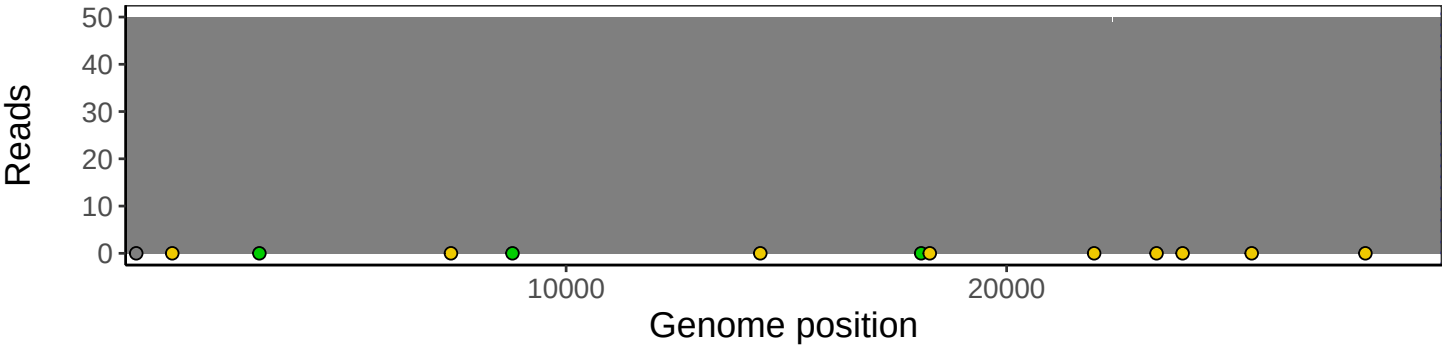
## Analyses of individual experiments and composite results.

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

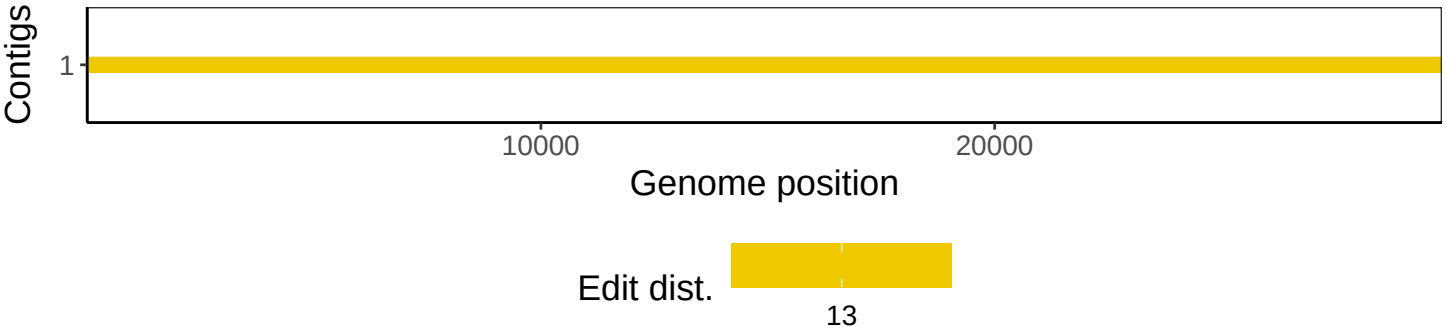
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

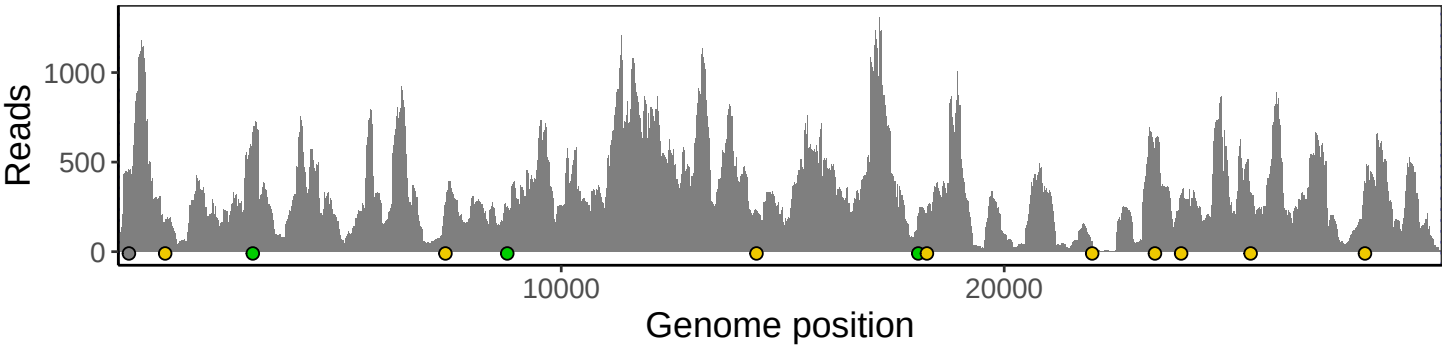



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

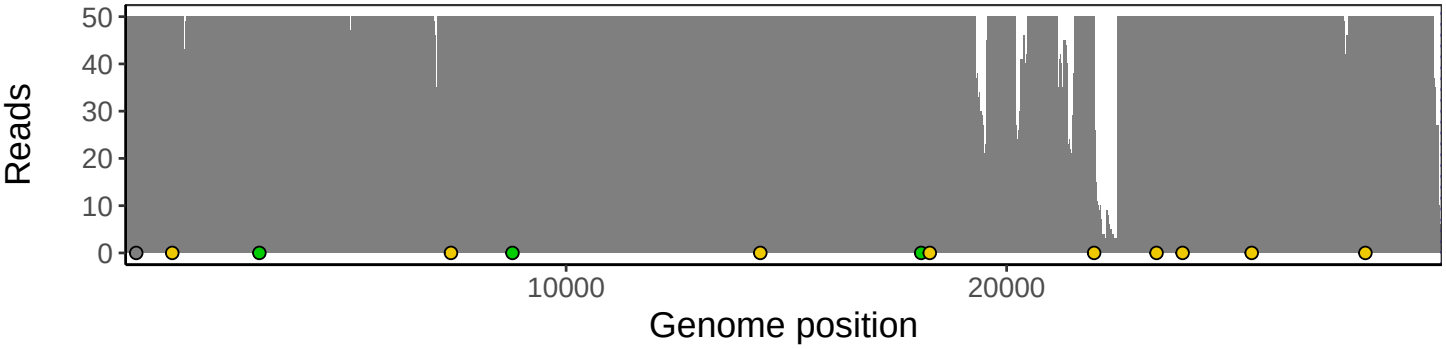


**VSP0022 | 2020-04-22 | OP | 228o | composite result**

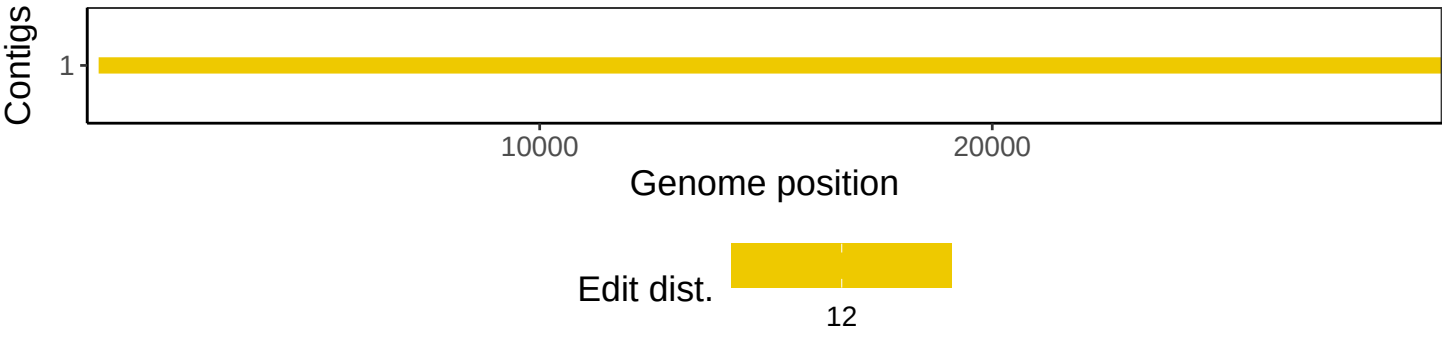
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

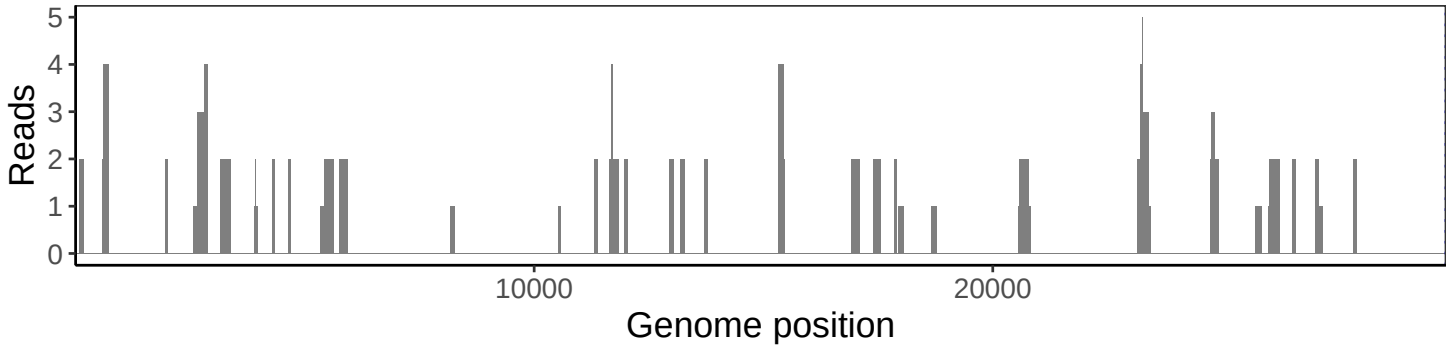



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

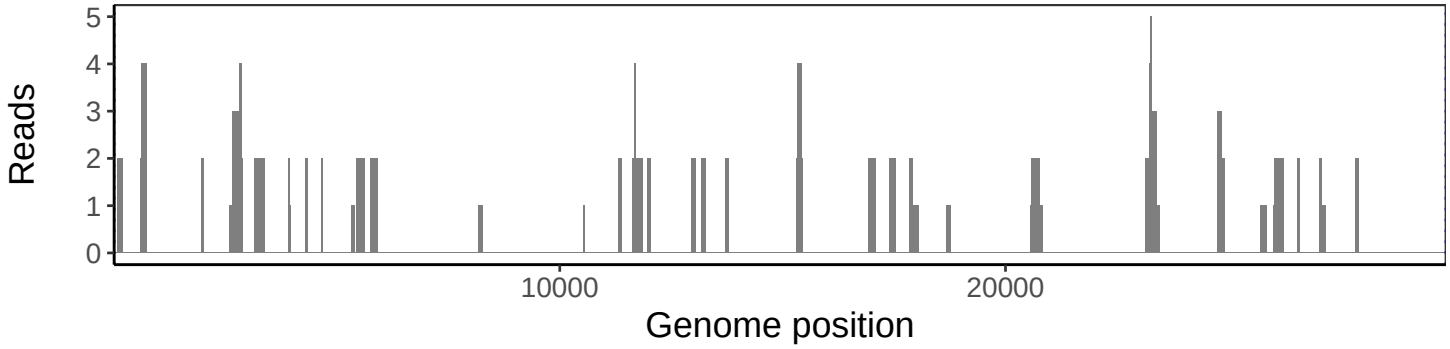


**VSP0187 | 2020-05-29 | ETA | 228e-q | composite result**

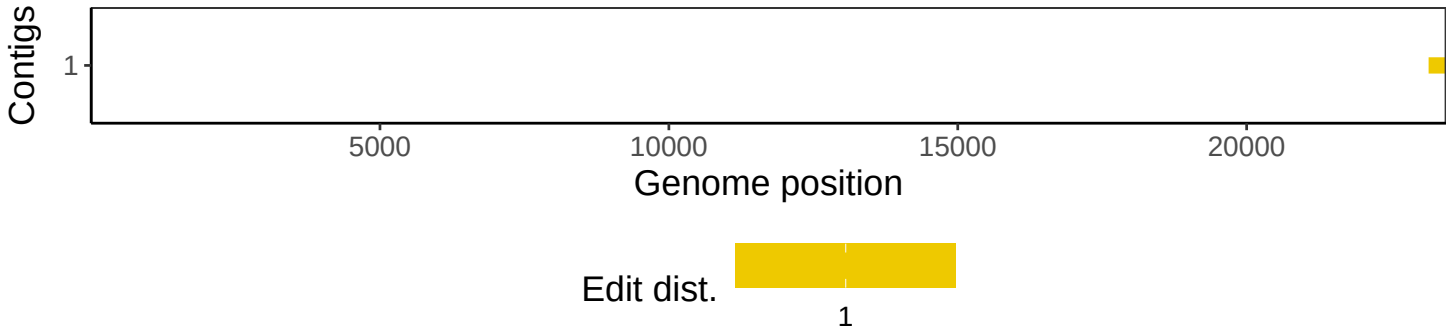
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

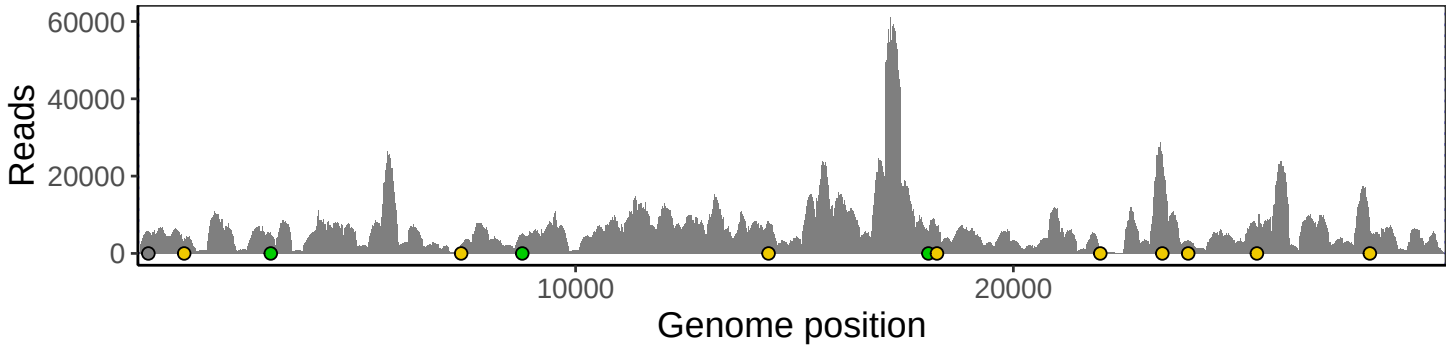



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

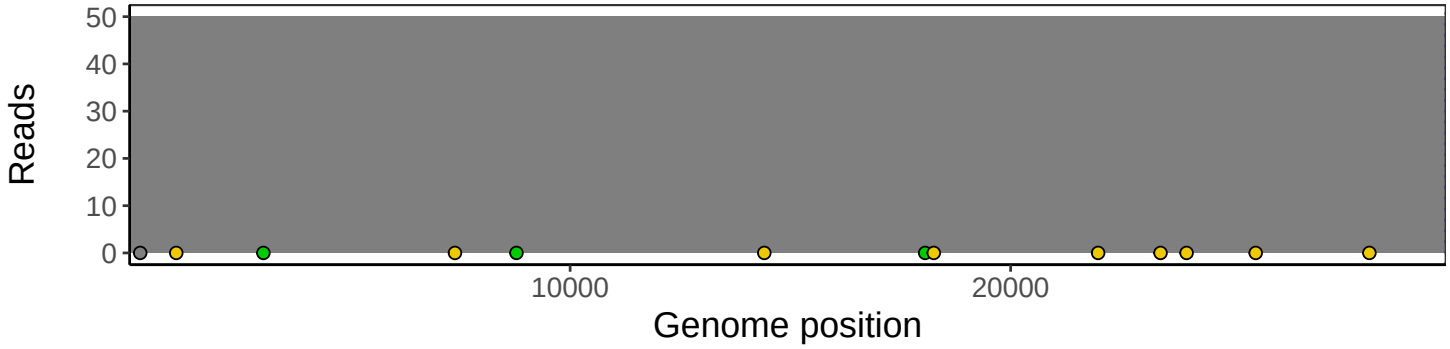


**VSP0188 | 2020-05-29 | NP-OP | 228no-q | composite result**

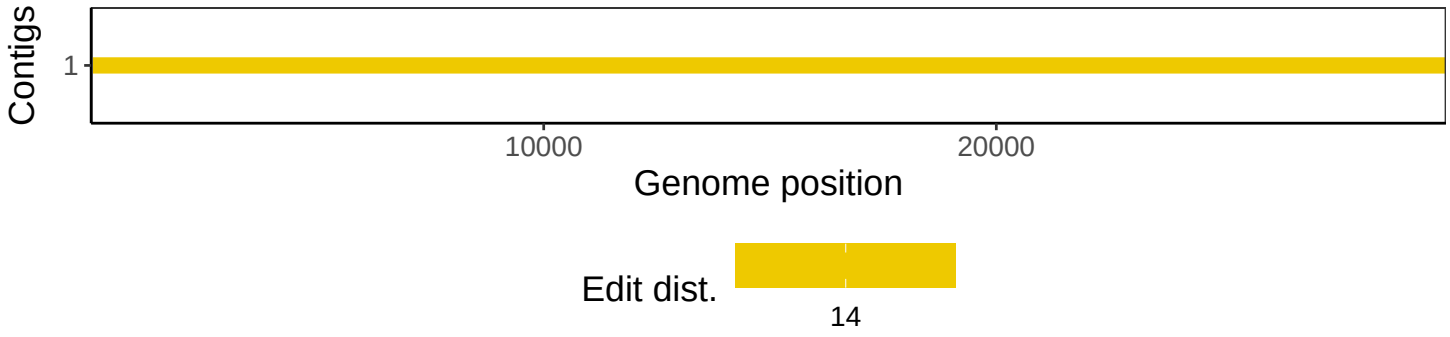
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

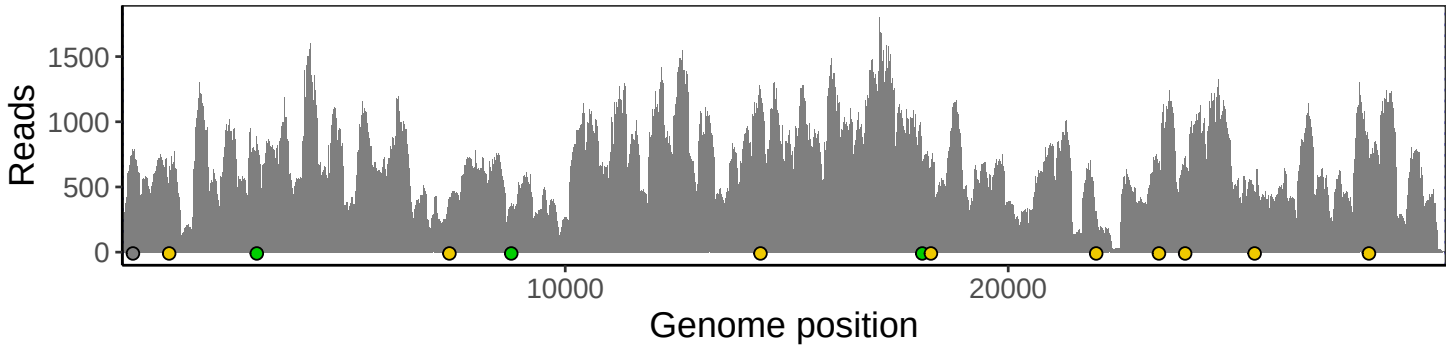



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

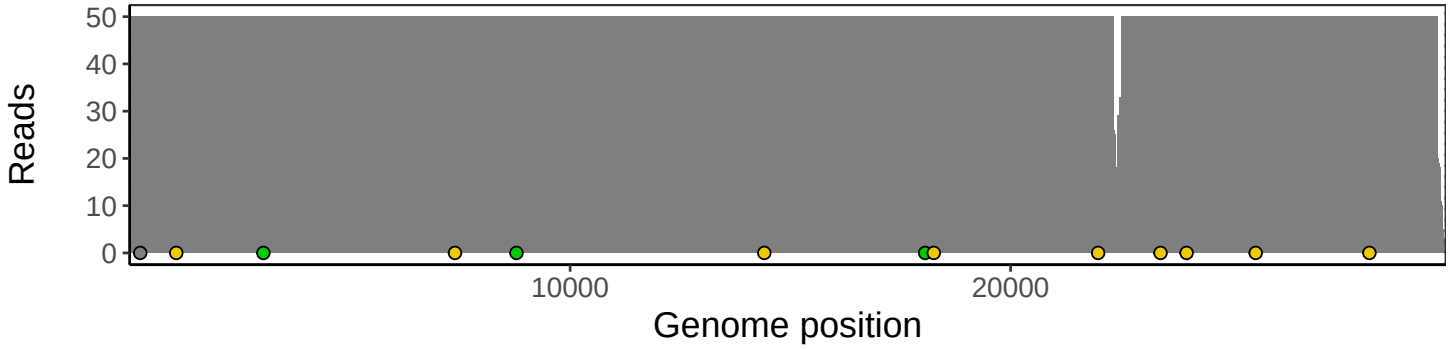


**VSP0021-1m | 2020-04-22 | NP | 228n | genomes | single experiment**

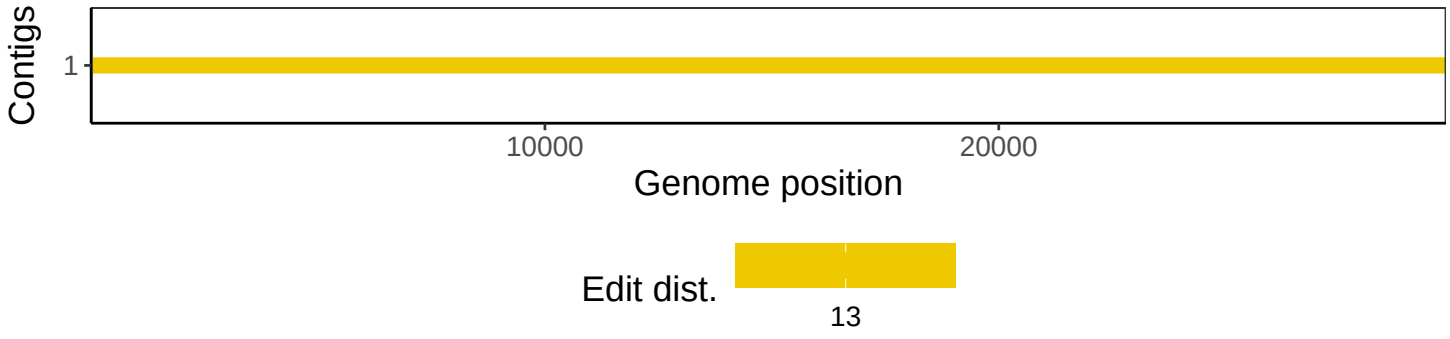
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

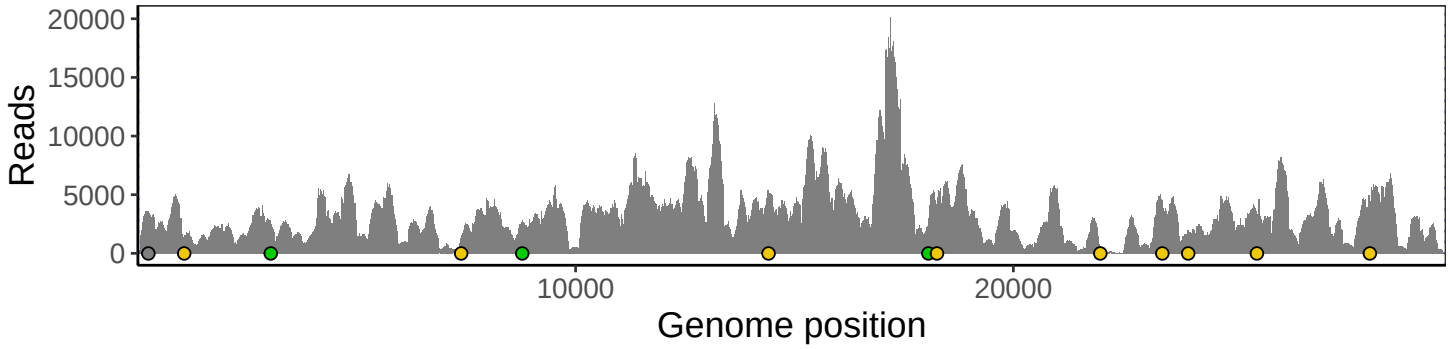



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

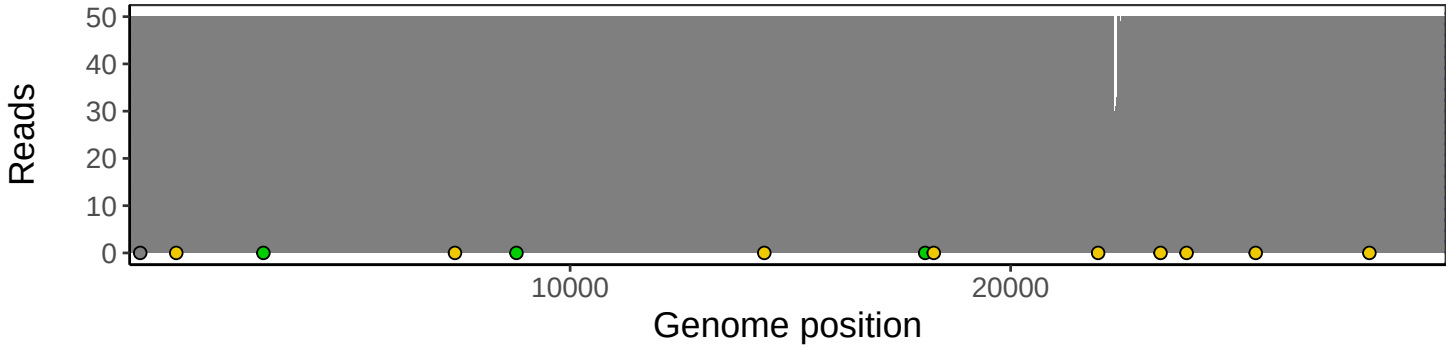


**VSP0021-2 | 2020-04-22 | NP | 228n | 11000 genomes | single experiment**

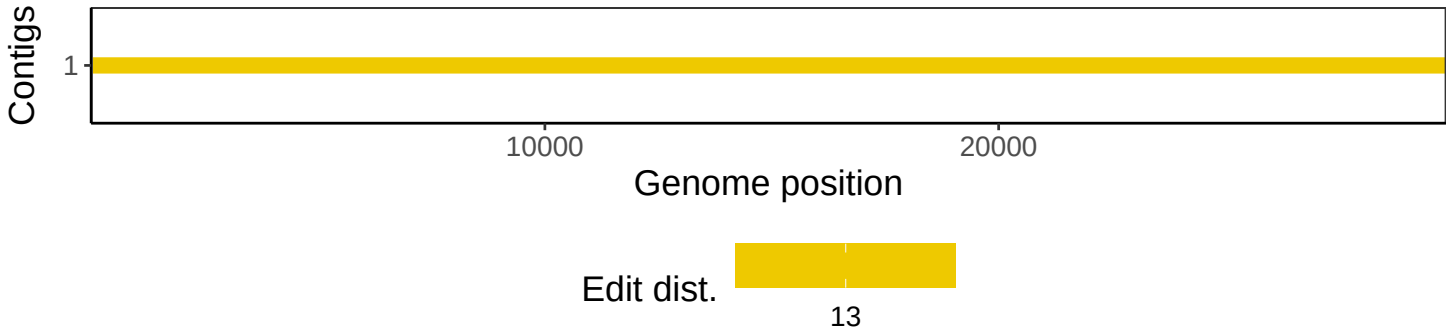
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

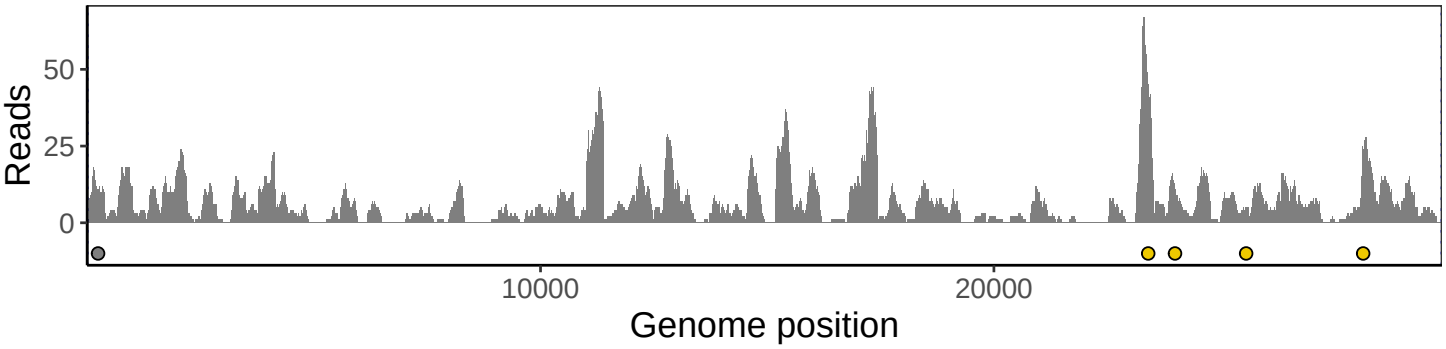



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

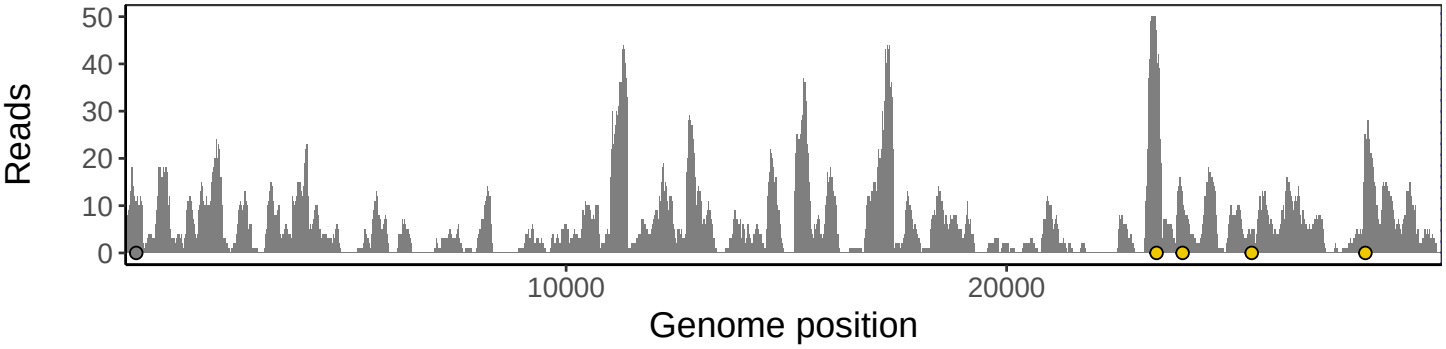


**VSP0022-1a | 2020-04-22 | OP | 228o | 268000 genomes | single experiment**

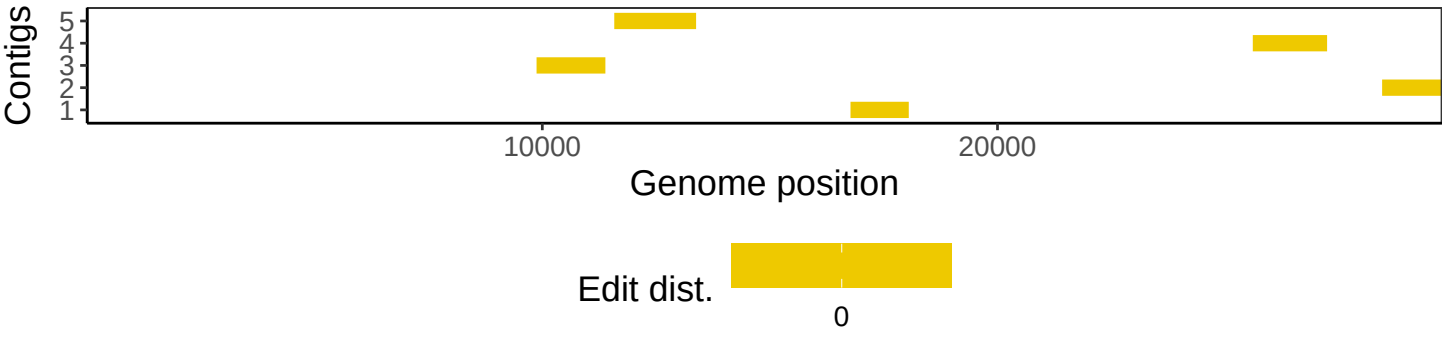
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

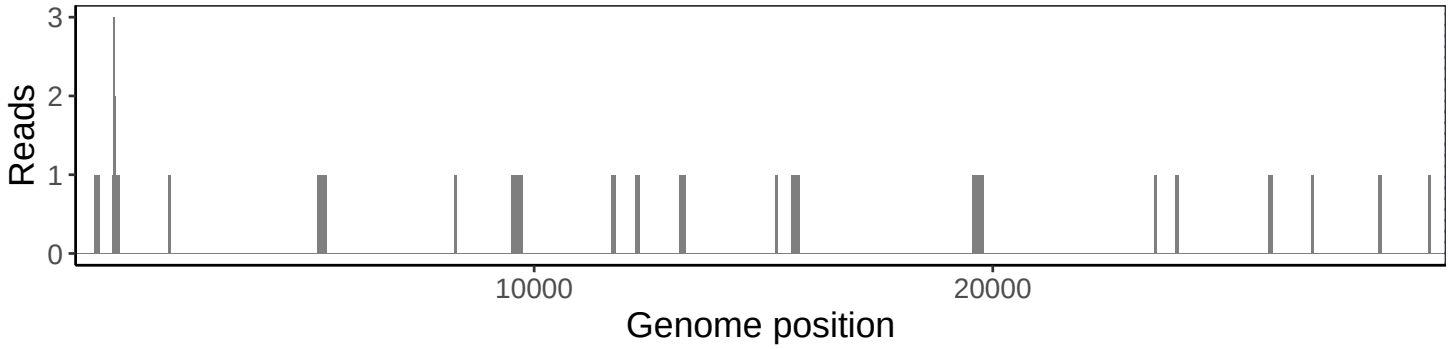



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

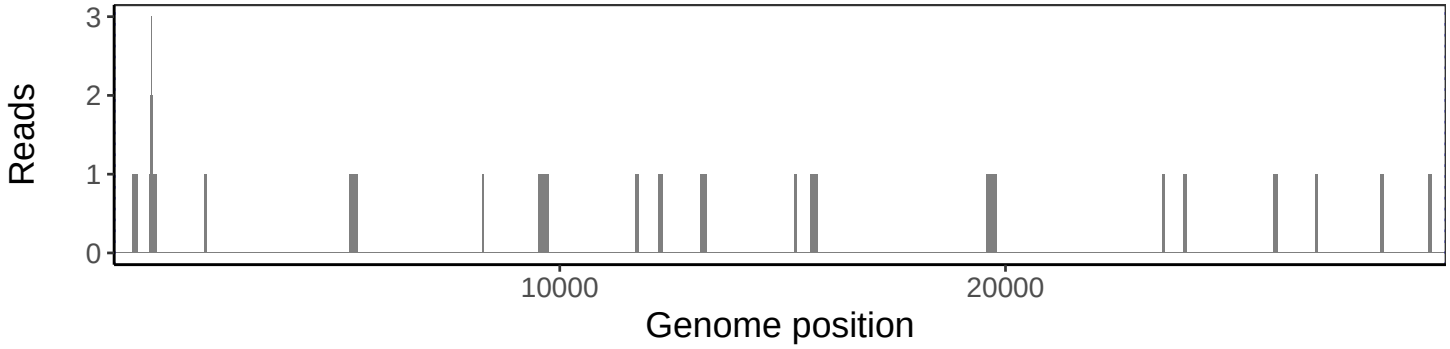


**VSP0022-2 | 2020-04-22 | OP | 228o | 1340000 genomes | single experiment**

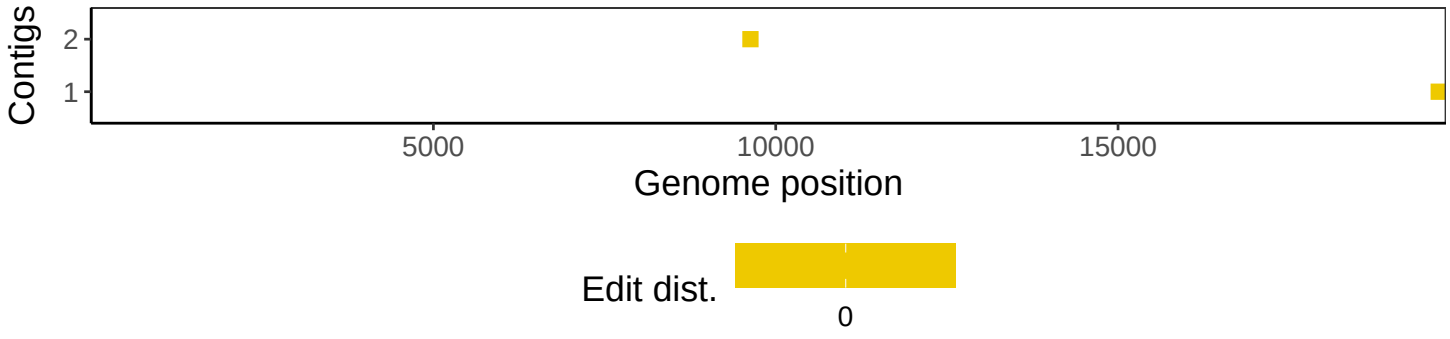
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

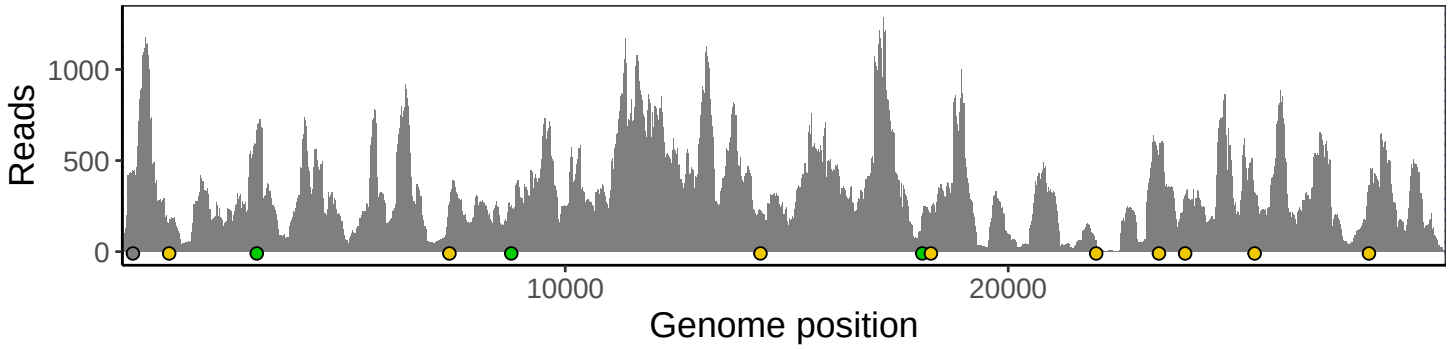



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

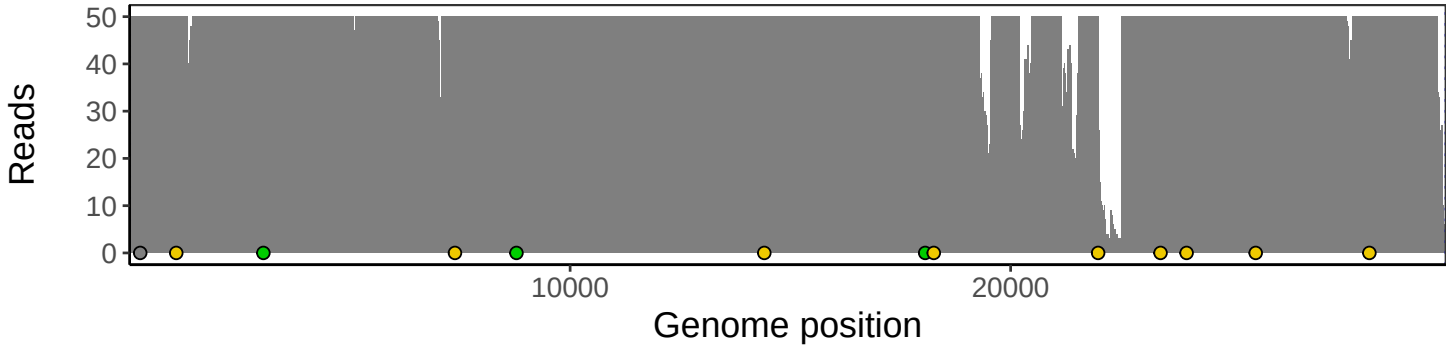


**VSP0022-3 | 2020-04-22 | OP | 228o | 1340000 genomes | single experiment**

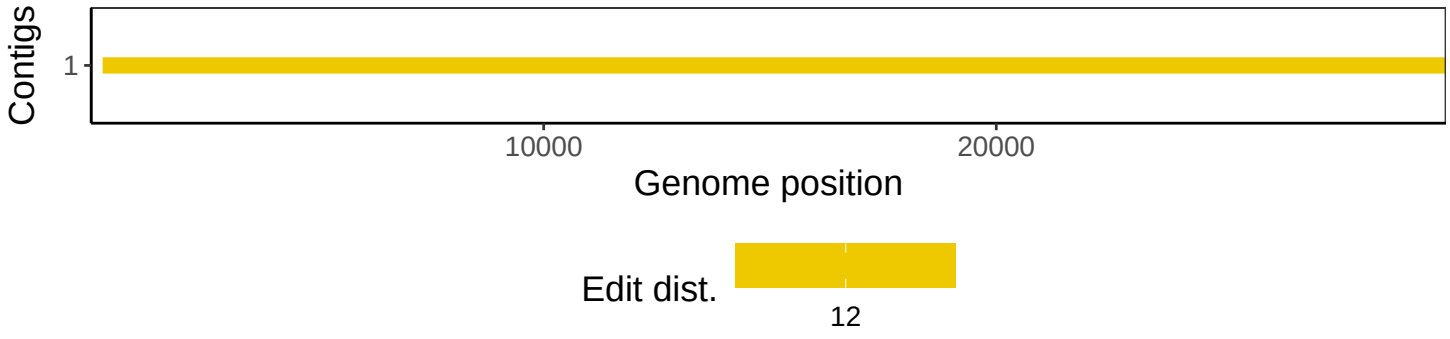
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

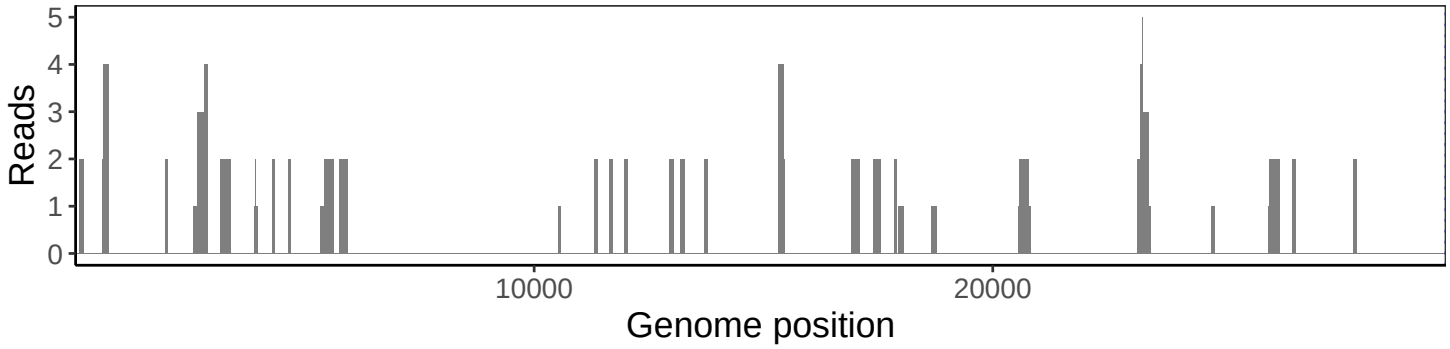



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

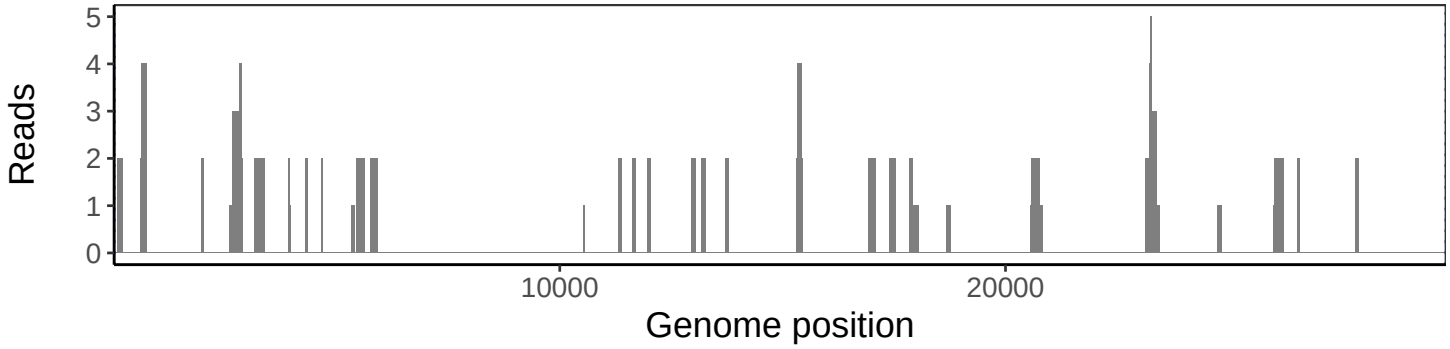


**VSP0187-1 | 2020-05-29 | ETA | 228e-q | 60.2 genomes | single experiment**

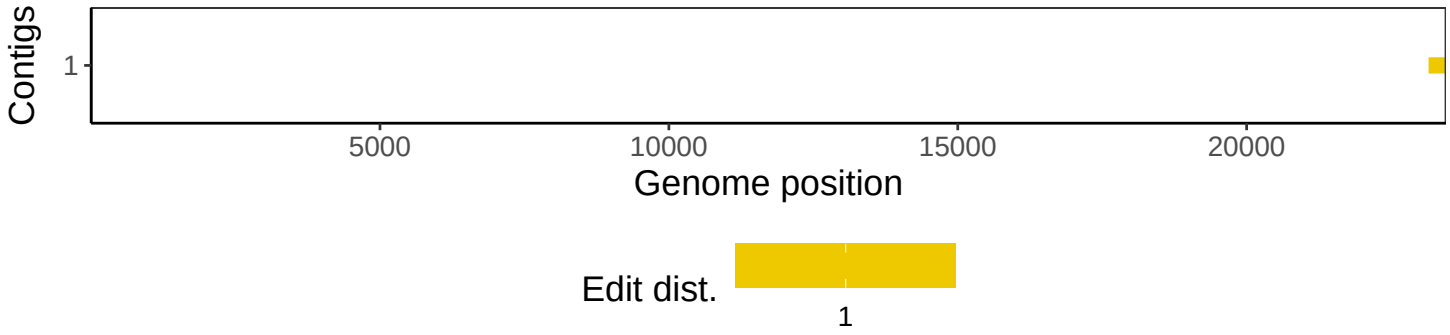
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

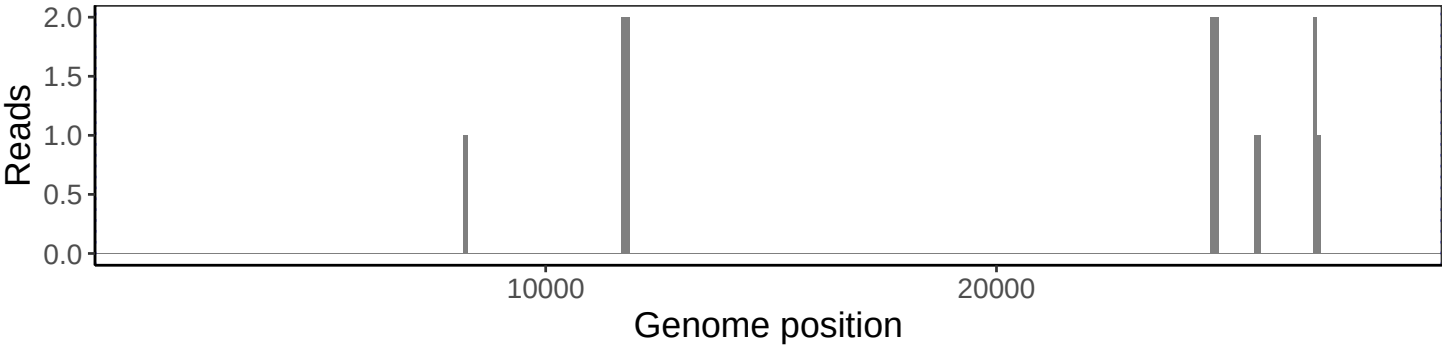



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

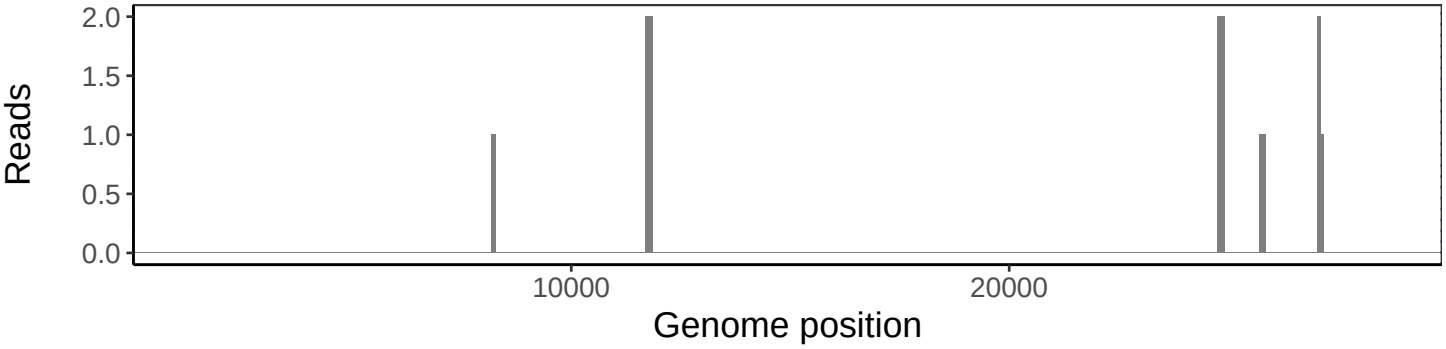


**VSP0187-2 | 2020-05-29 | ETA | 228e-q | 301 genomes | single experiment**

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.



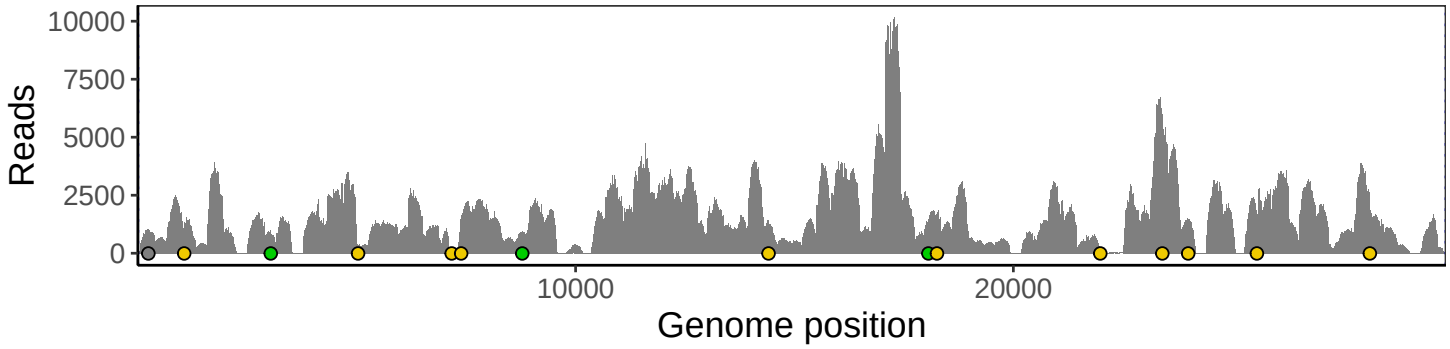

Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



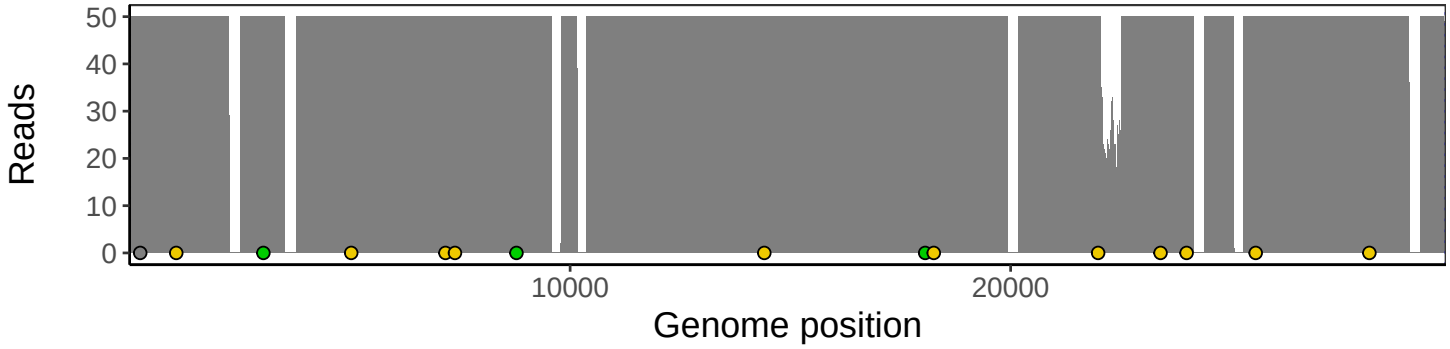

## No contig data available.

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

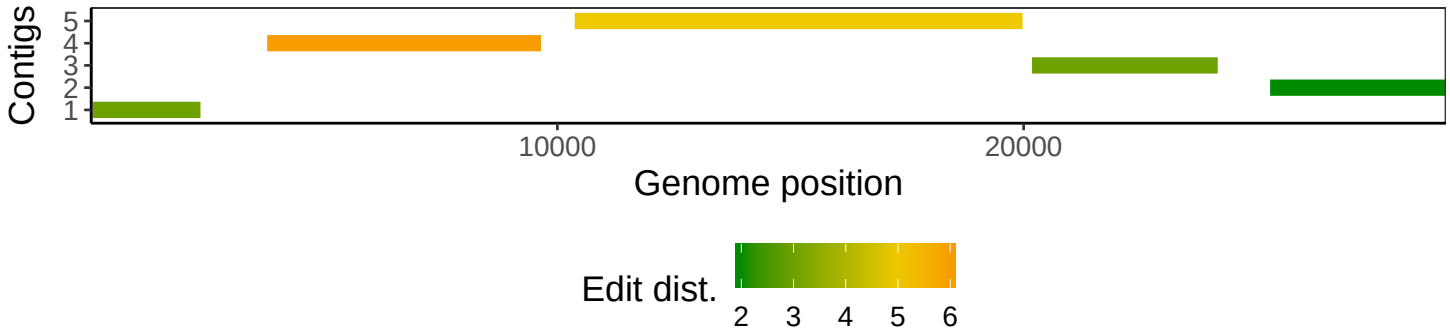
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

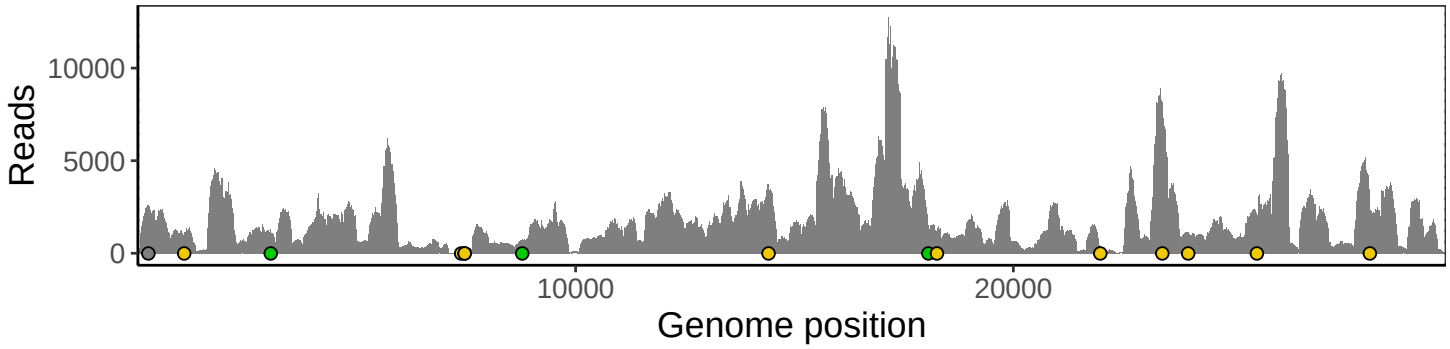



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

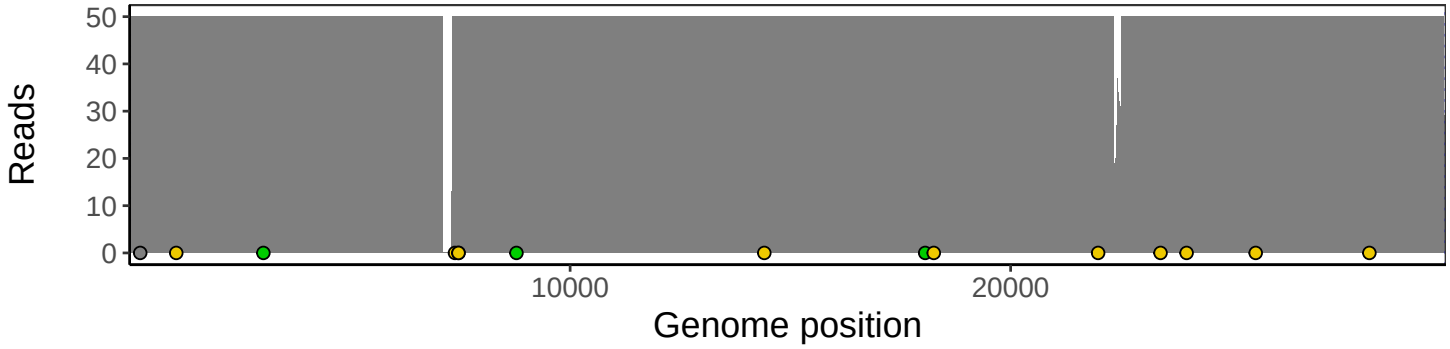


**VSP0188-2 | 2020-05-29 | NP-OP | 228no-q | 10200 genomes | single experiment**

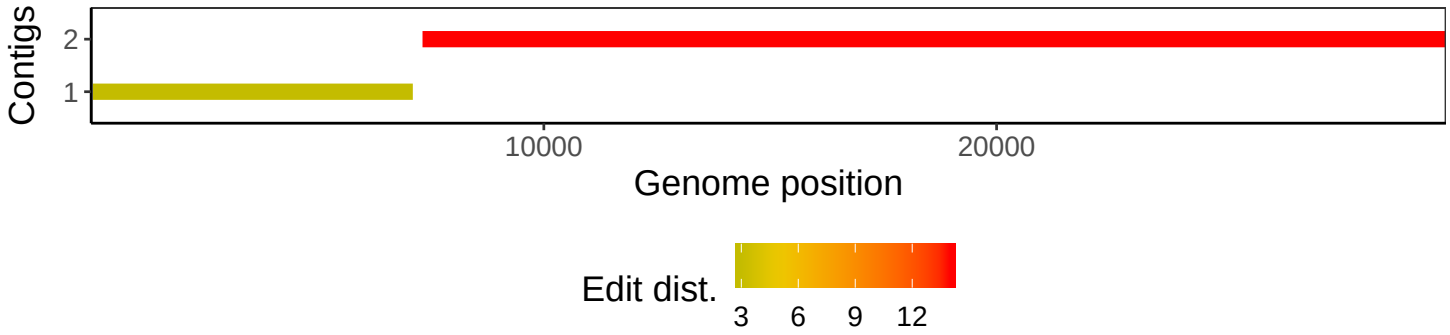
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

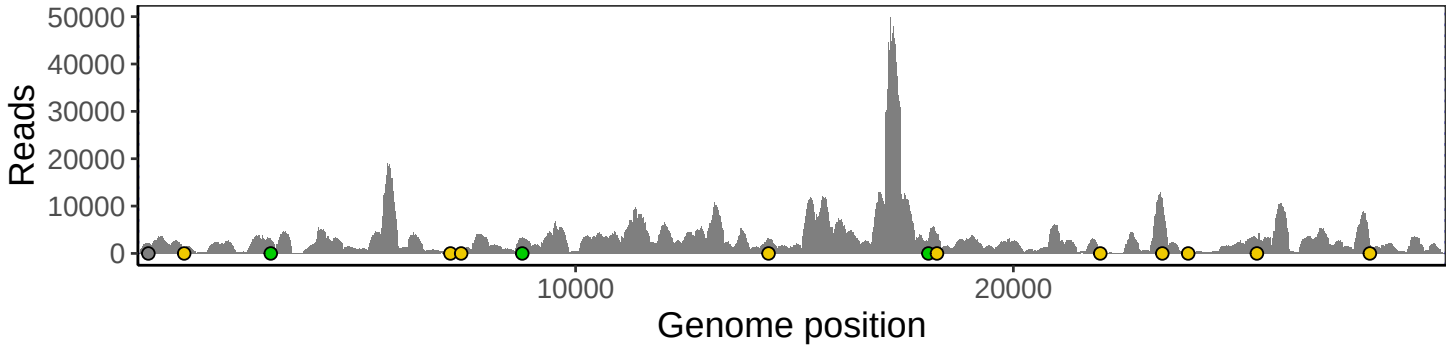



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

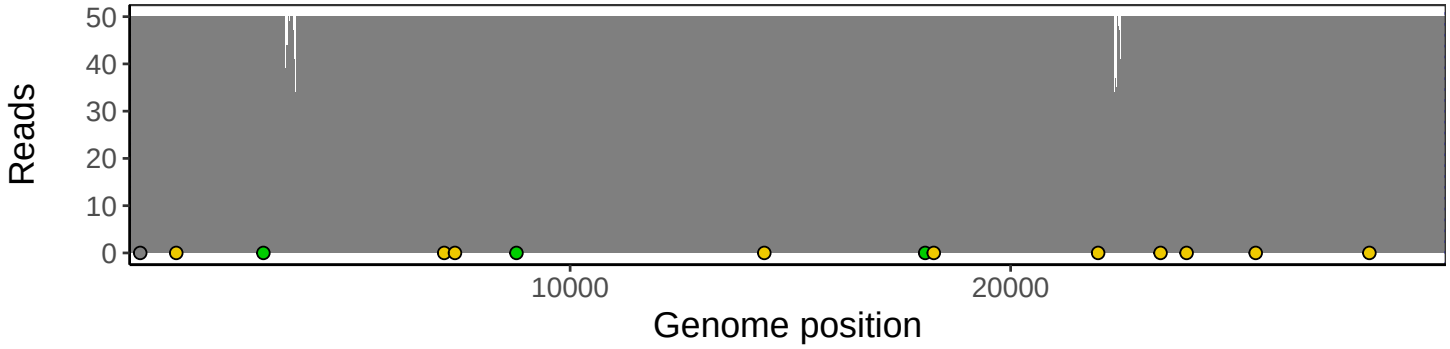


**VSP0188-3 | 2020-05-29 | NP-OP | 228no-q | 10200 genomes | single experiment**

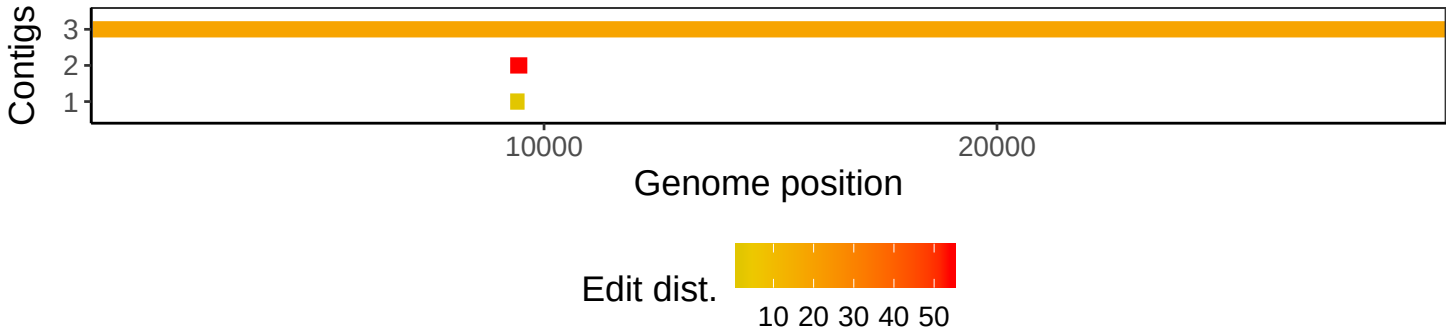
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

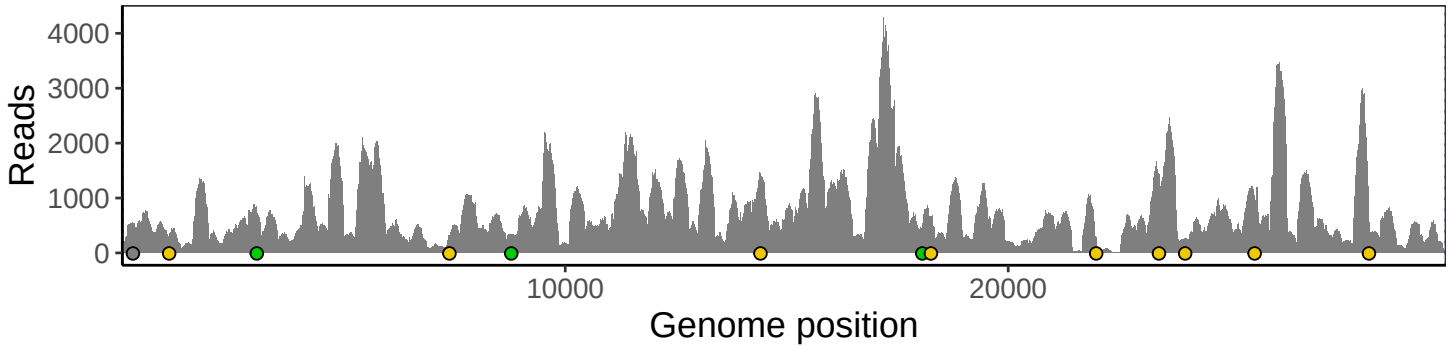



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

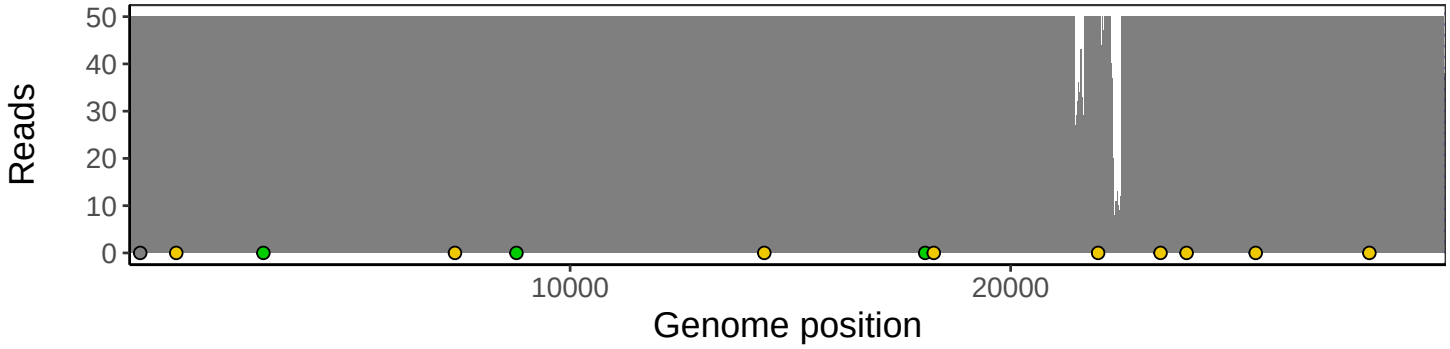


**VSP0189-1 | 2020-05-29 | Saliva | 228s-q | 85100 genomes | single experiment**

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.