

# COVID-19 subject 251

*2020-11-30*

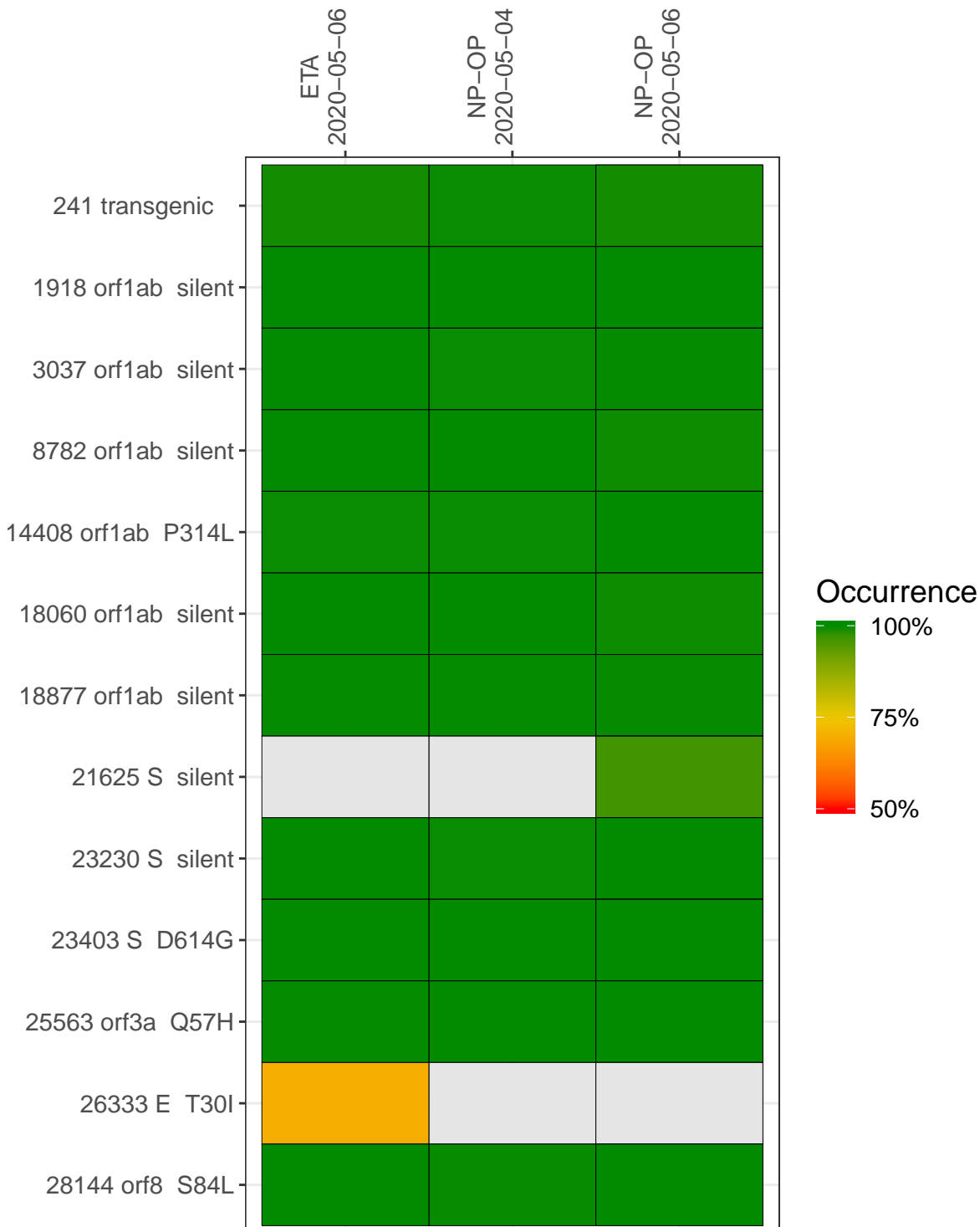
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

| Experiment | Type              | Input genomes | Sample type | Sample date | Largest contig (KD) | Reference read coverage | Reference read coverage (>= 5 reads) |
|------------|-------------------|---------------|-------------|-------------|---------------------|-------------------------|--------------------------------------|
| VSP0065    | composite         | NA            | NP-OP       | 2020-05-04  | 29.94               | 99.9%                   | 99.8%                                |
| VSP0089    | composite         | NA            | NP-OP       | 2020-05-06  | 30.00               | 99.9%                   | 99.7%                                |
| VSP0065-1  | single experiment | 7550000       | NP-OP       | 2020-05-04  | 29.88               | 99.9%                   | 99.8%                                |
| VSP0065-2  | single experiment | 7550000       | NP-OP       | 2020-05-04  | 29.85               | 99.8%                   | 99.3%                                |
| VSP0088-1  | single experiment | 255500        | ETA         | 2020-05-06  | 29.82               | 99.8%                   | 99.8%                                |
| VSP0089-1  | single experiment | 570000        | NP-OP       | 2020-05-06  | 29.91               | 99.9%                   | 99.7%                                |
| VSP0089-2  | single experiment | 570000        | NP-OP       | 2020-05-06  | 29.86               | 99.3%                   | 98.6%                                |
| VSP0318-1  | single experiment | 890000        | NP-OP       | 2020-05-06  | 24.79               | 99.8%                   | 99.1%                                |

## 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<br>2020-05-06 | NP-OP<br>2020-05-04 |           | NP-OP<br>2020-05-06 |           |           |
|---------------------|-------------------|---------------------|-----------|---------------------|-----------|-----------|
| 241 transgenic      | 1422              | 1303                | 4964      | 2277                | 575       | 584       |
| 1918 orf1ab silent  | 1605              | 1449                | 4334      | 1728                | 1347      | 515       |
| 3037 orf1ab silent  | 1227              | 2589                | 1542      | 1850                | 278       | 887       |
| 8782 orf1ab silent  | 1129              | 2305                | 230       | 1900                | 107       | 316       |
| 14408 orf1ab P314L  | 2583              | 2062                | 64        | 3791                | 186       | 206       |
| 18060 orf1ab silent | 1249              | 2529                | 478       | 1816                | 184       | 302       |
| 18877 orf1ab silent | 3374              | 4242                | 175       | 4399                | 753       | 745       |
| 21625 S silent      | 264               | 1529                | 16        | 529                 | 38        | 81        |
| 23230 S silent      | 2725              | 2075                | 3342      | 3509                | 1050      | 488       |
| 23403 S D614G       | 3810              | 2573                | 4155      | 4784                | 1276      | 579       |
| 25563 orf3a Q57H    | 1114              | 1375                | 4101      | 1487                | 487       | 365       |
| 26333 E T30I        | 1331              | 1531                | 566       | 1975                | 281       | 271       |
| 28144 orf8 S84L     | 2363              | 1034                | 115       | 3792                | 508       | 465       |
|                     | VSP0088-1         | VSP0065-1           | VSP0065-2 | VSP0089-1           | VSP0089-2 | VSP0318-1 |

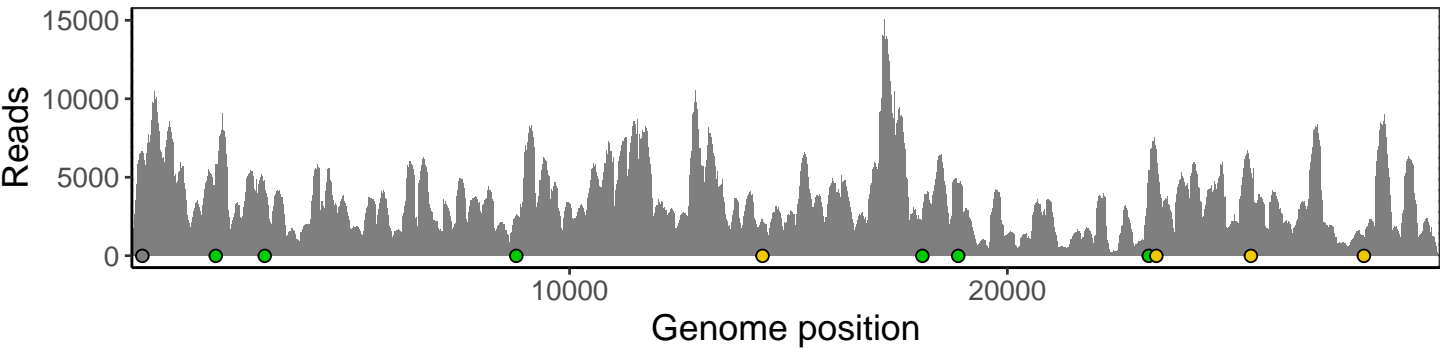
Base change

- Expected
- A
- T
- C
- G
- N
- Ins/Del
- No data

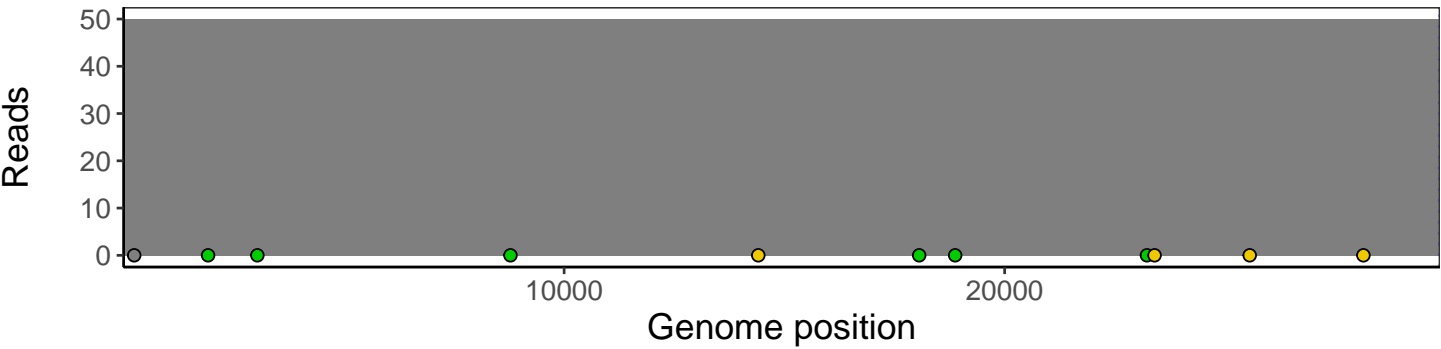
Analyses of individual experiments and composite results.

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

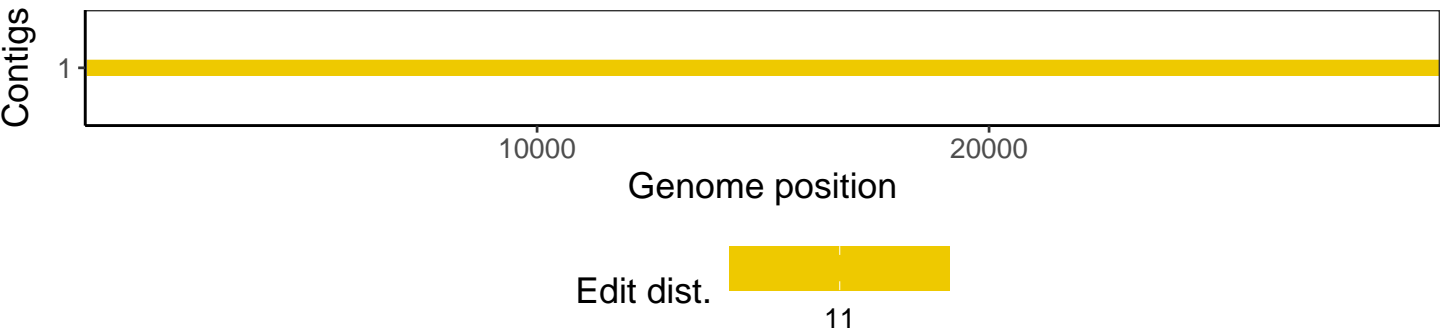
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



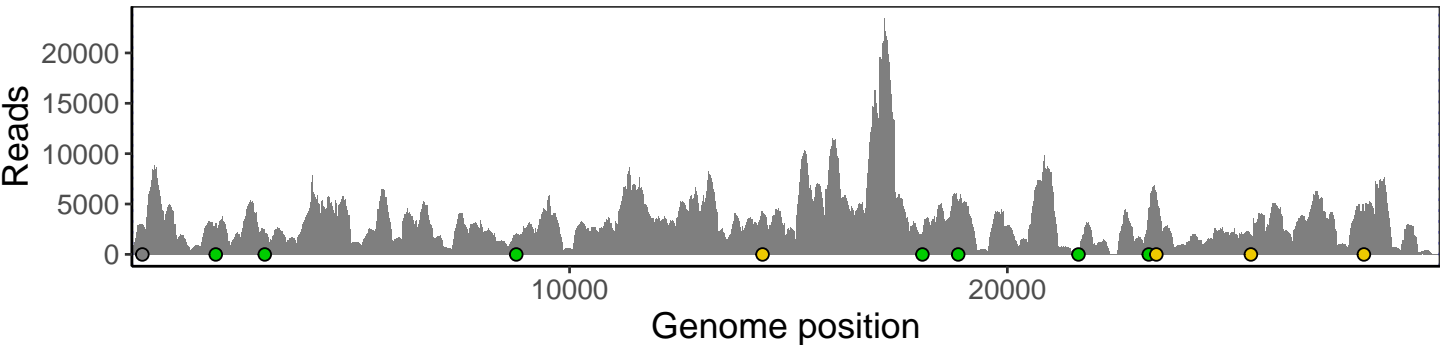
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



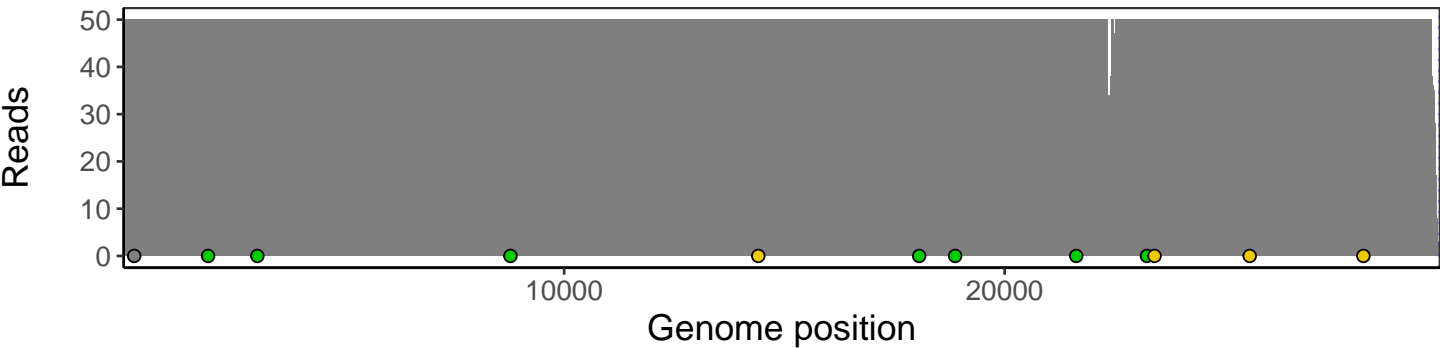
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



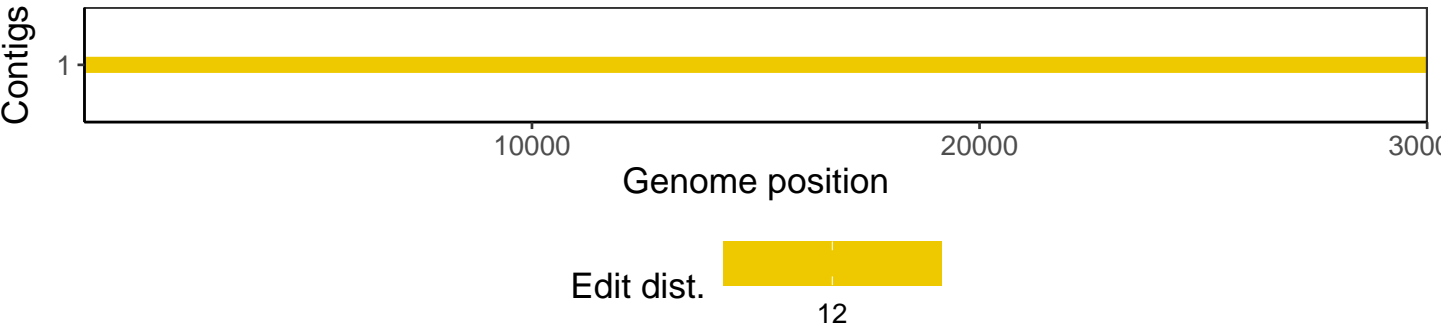
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



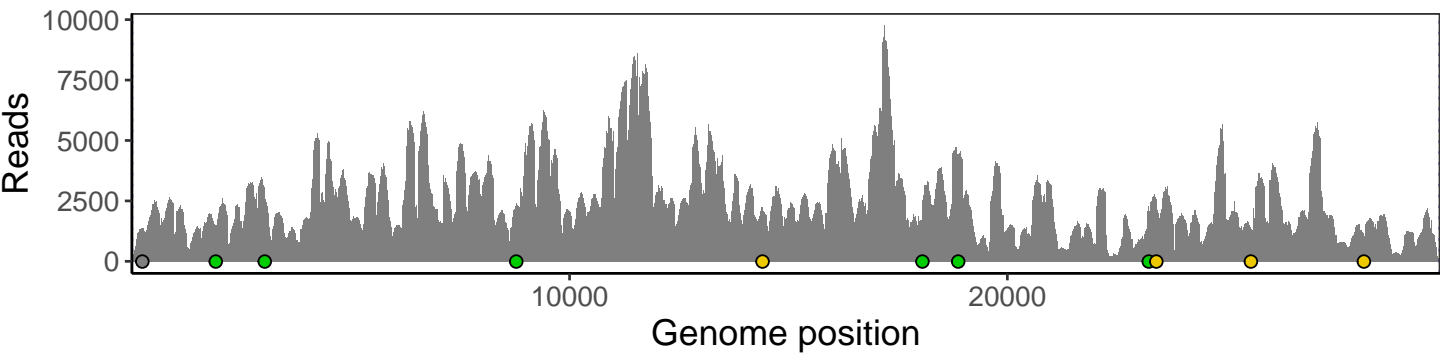
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



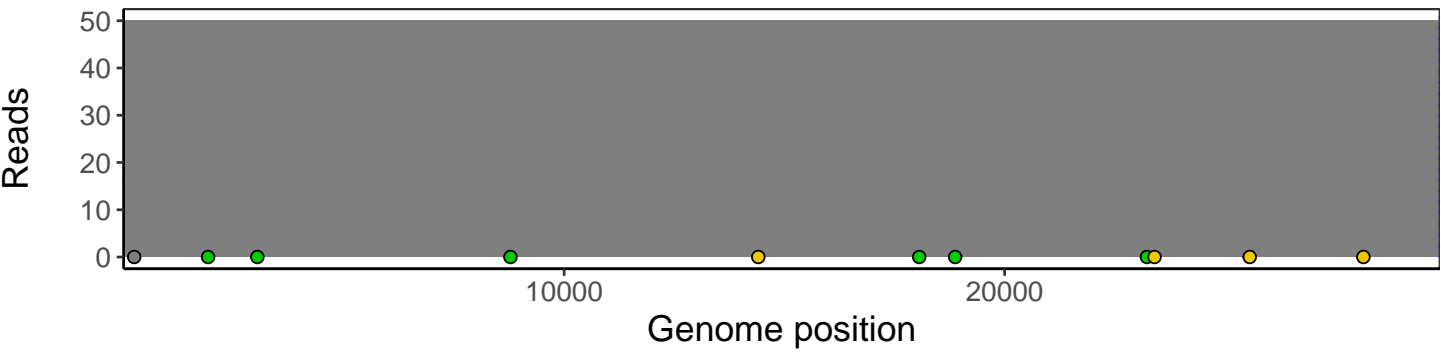
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



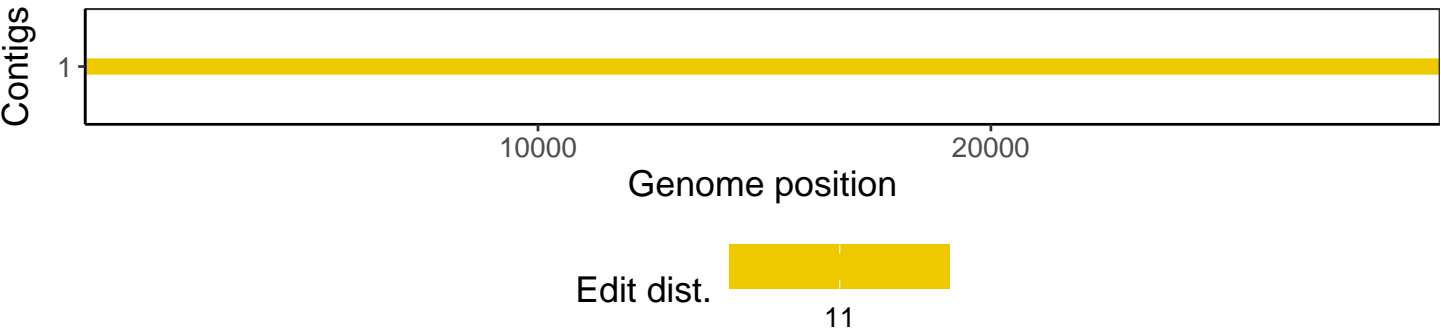
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



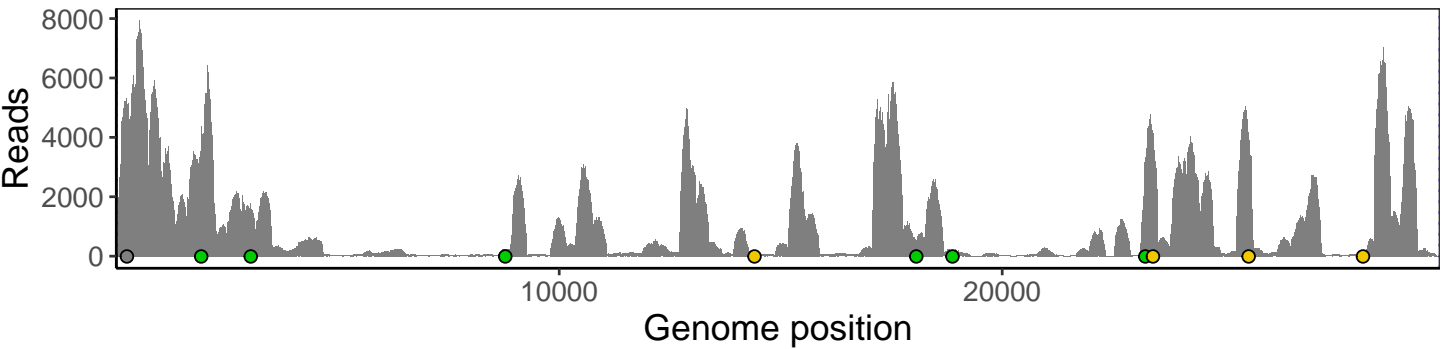
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



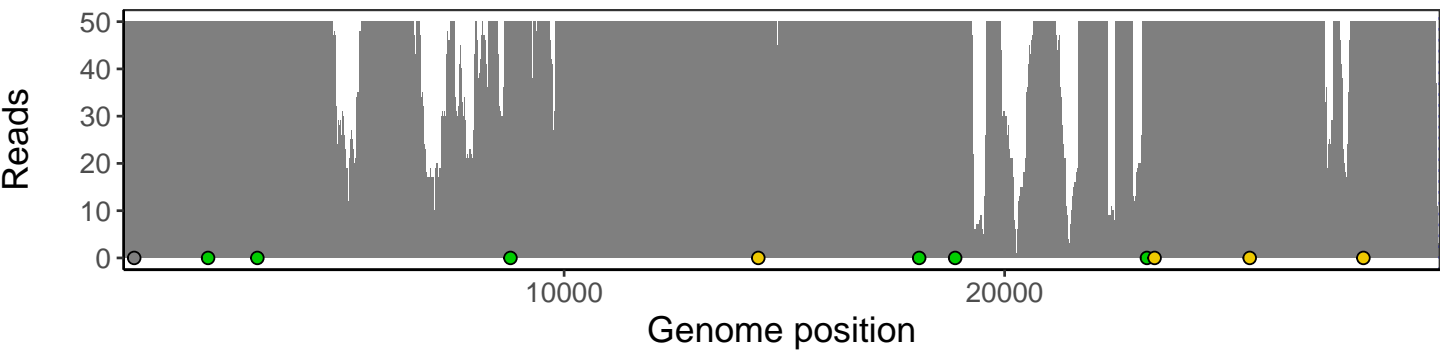
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



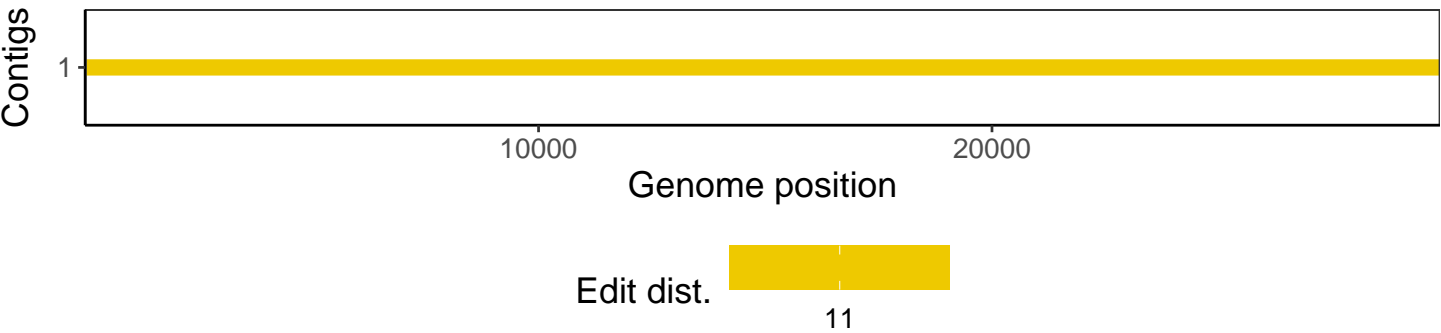
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



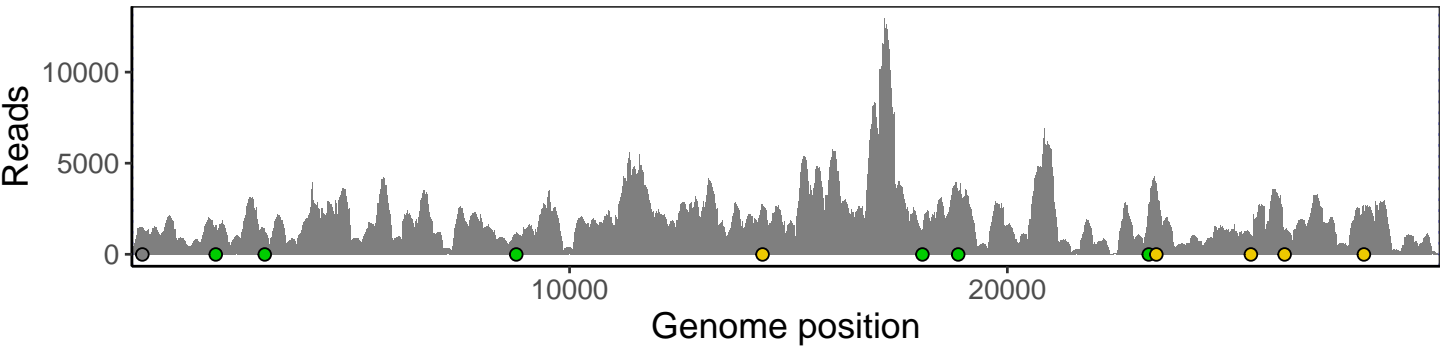
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



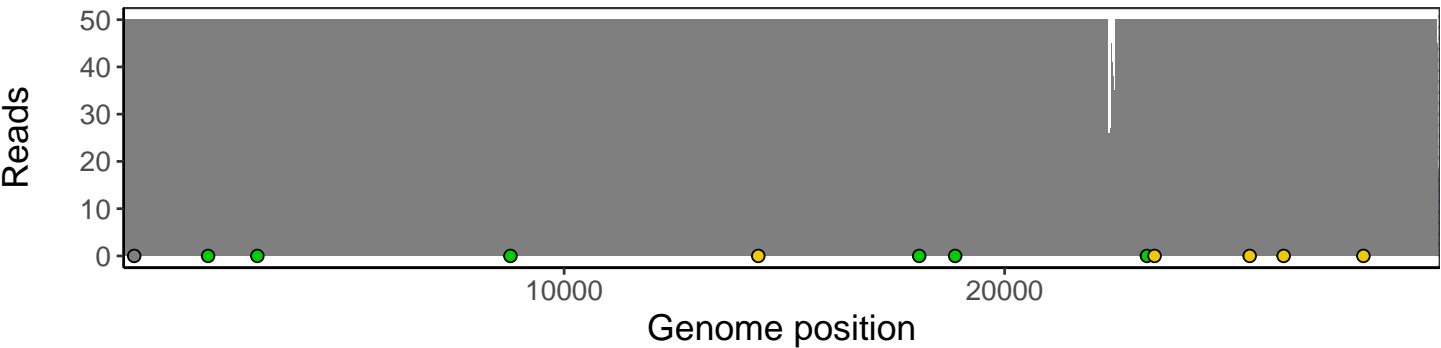
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



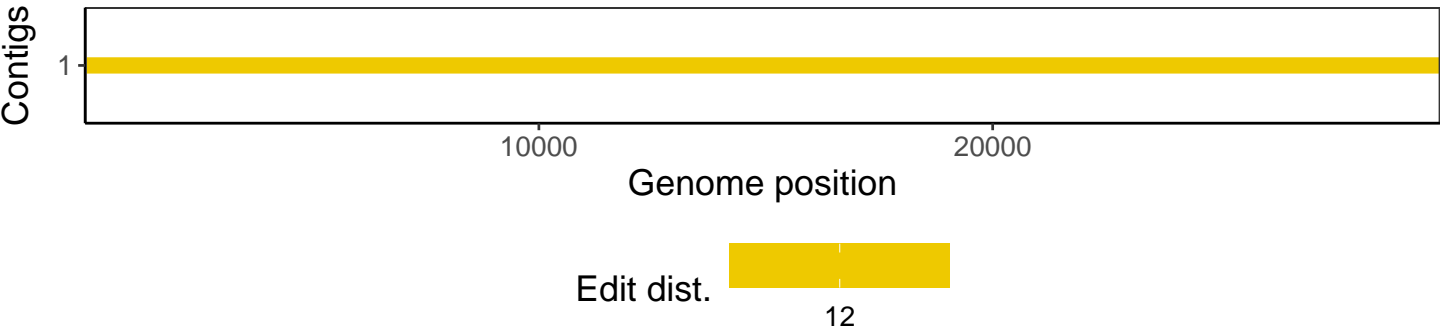
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

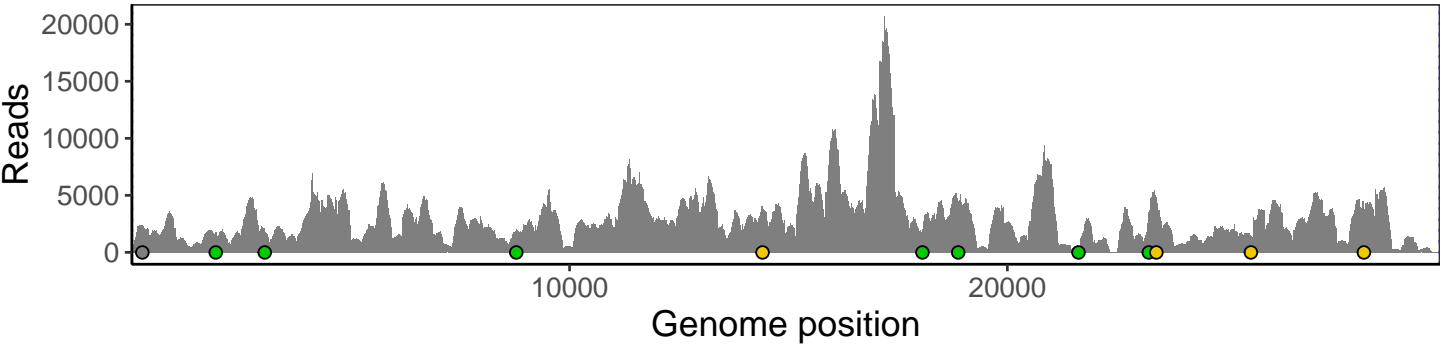


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

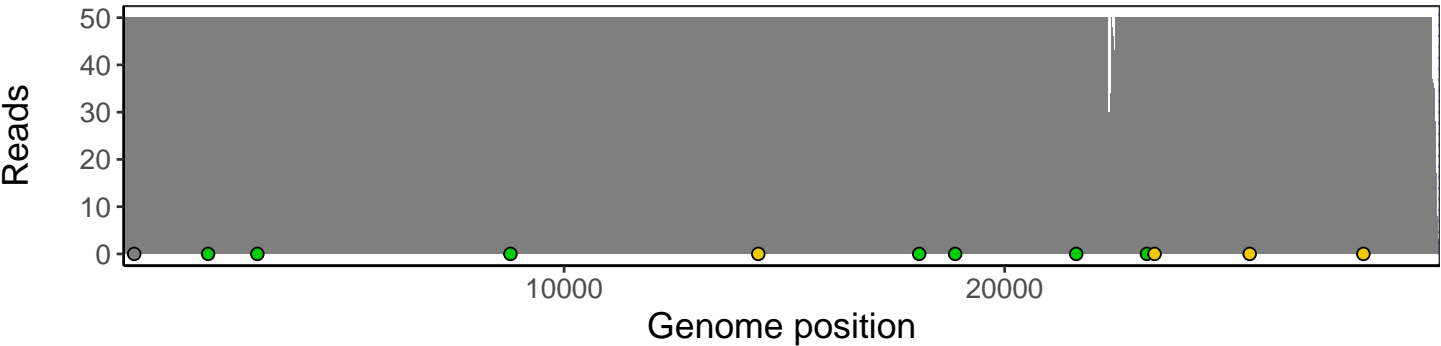




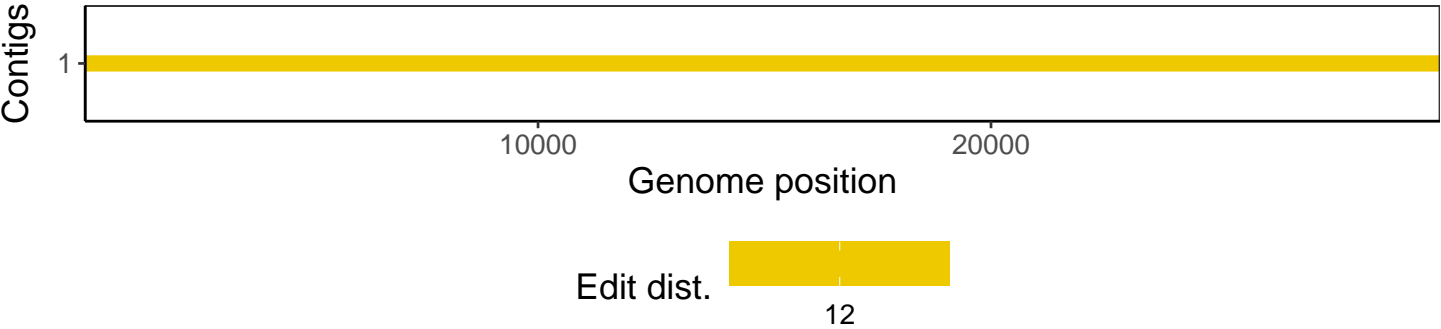
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



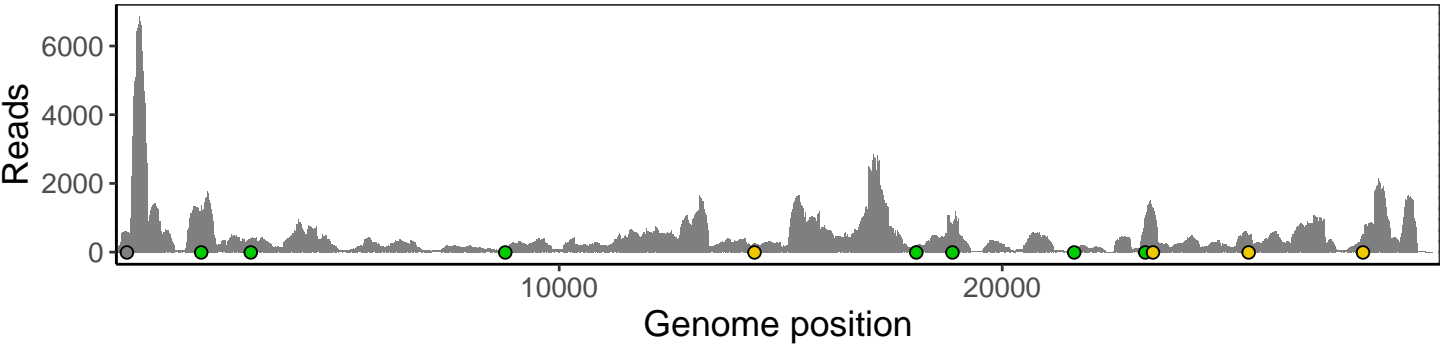
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



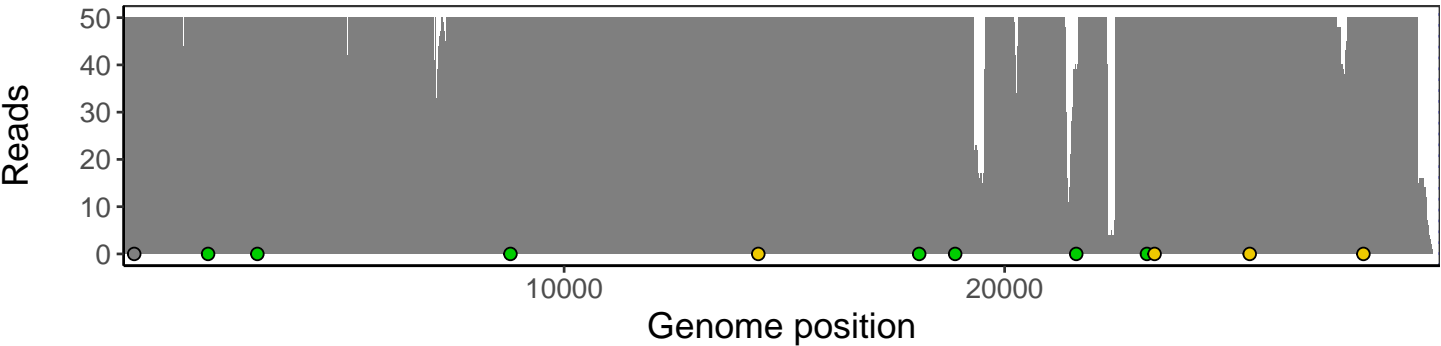
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



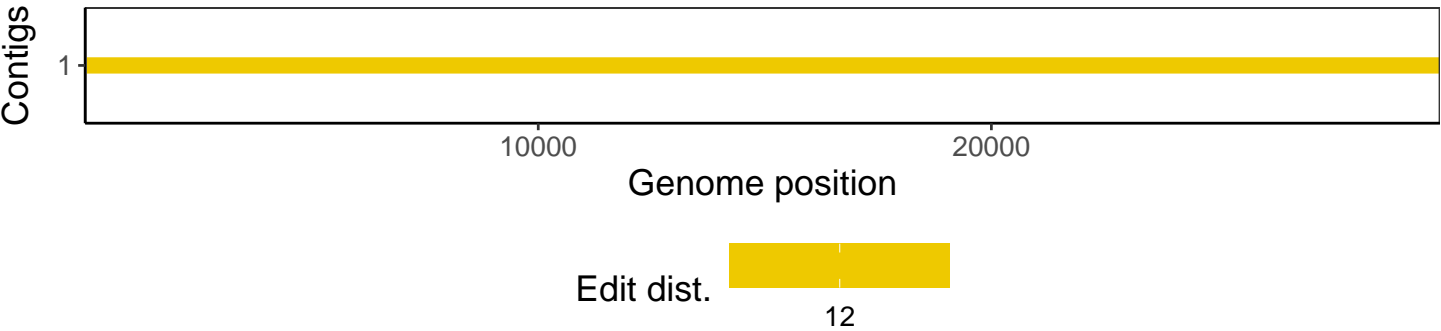
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



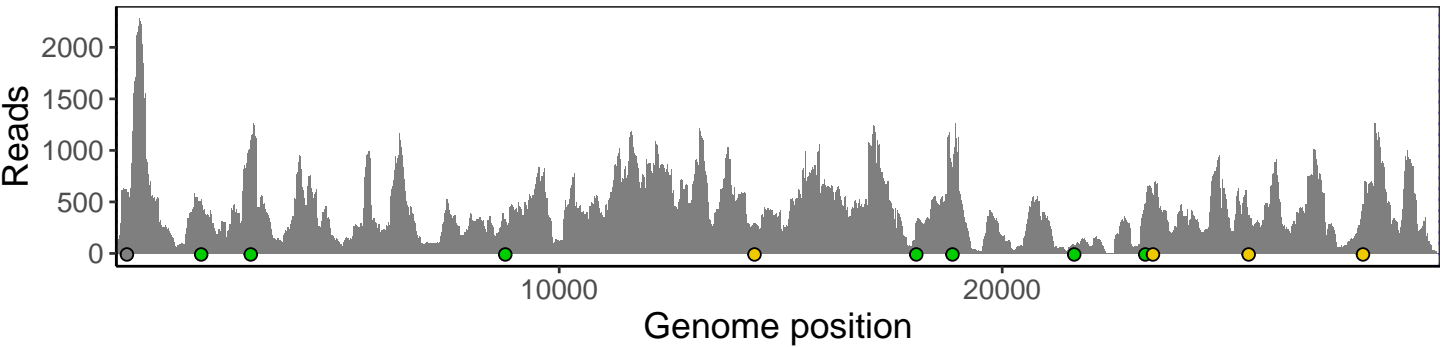
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



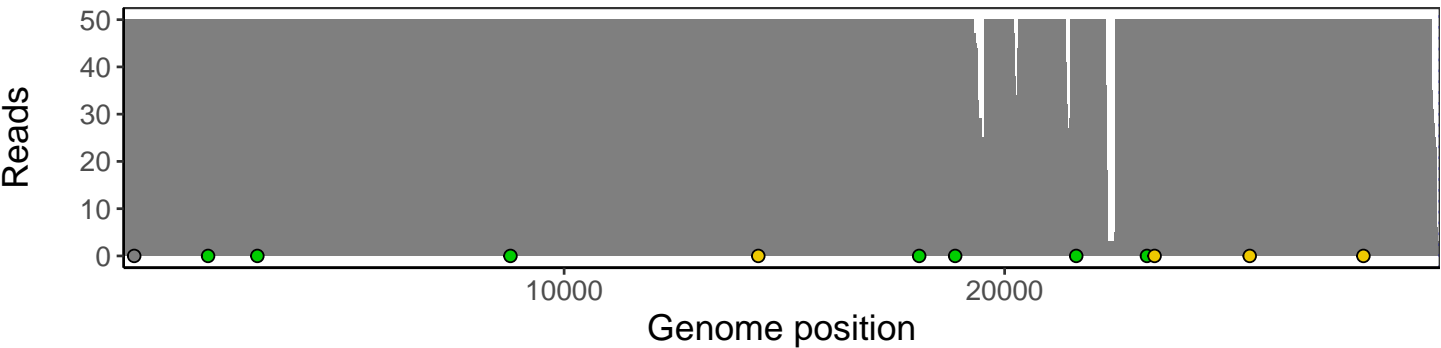
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

