# COVID-19 subject 237

*2020-08-26*

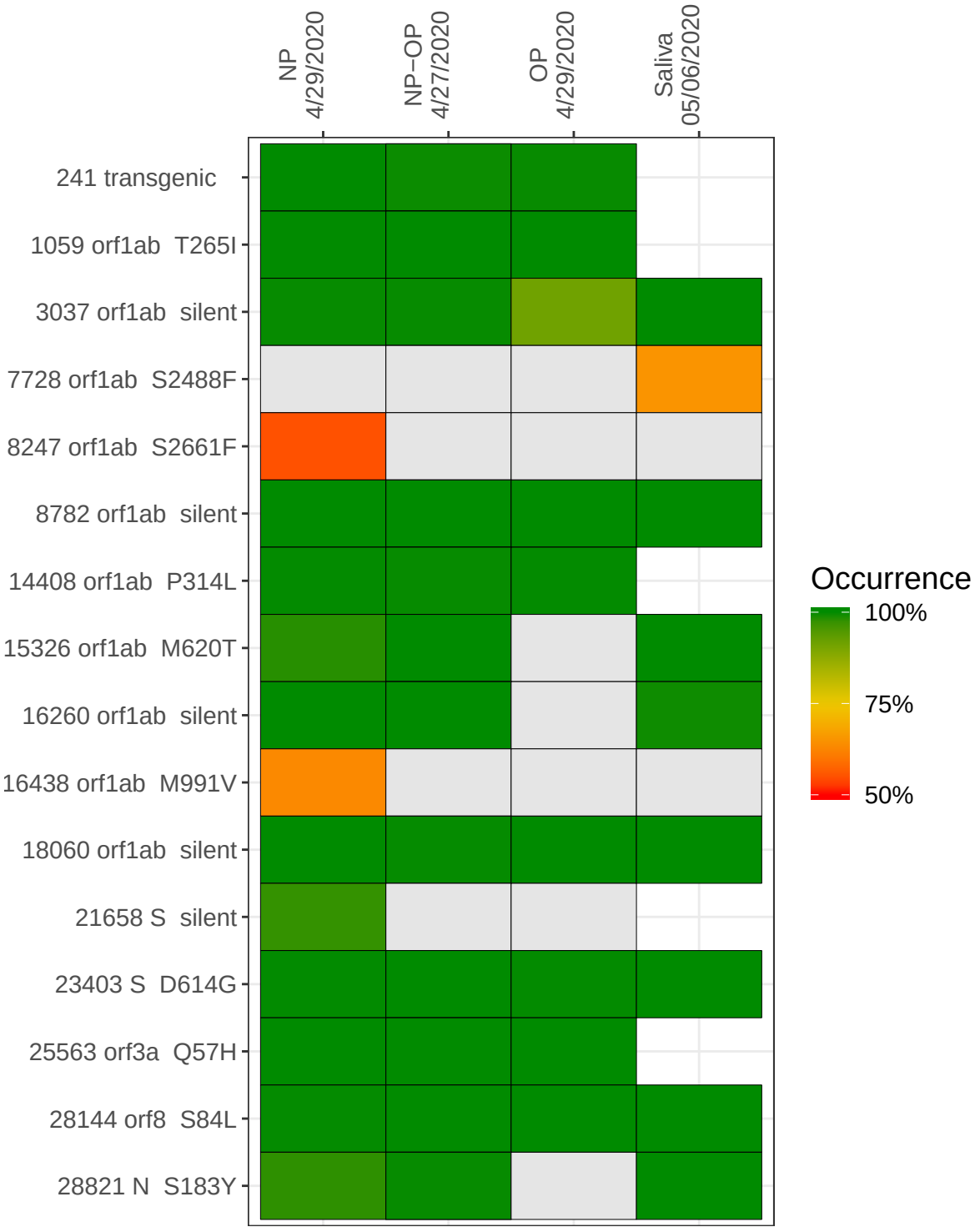
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

| Experiment | Type | Input genomes | Sample type | Sample date | Largest contig (KD) | Reference read coverage | Reference read coverage (>= 5 reads) |
|---|---|---|---|---|---|---|---|
| VSP0033 | composite | NA | NP-OP | 4/27/2020 | 29.97 | 99.9% | 99.8% |
| VSP0034 | composite | NA | NP-OP | 4/27/2020 | 29.84 | 99.8% | 99.8% |
| VSP0037 | composite | NA | NP | 4/29/2020 | 24.92 | 99.3% | 98.9% |
| VSP0038 | composite | NA | OP | 4/29/2020 | 29.92 | 99.9% | 99.5% |
| VSP0033-1a | single experiment | 597 | NP-OP | 4/27/2020 | 28.58 | 99.9% | 99.8% |
| VSP0033-1b | single experiment | 597 | NP-OP | 4/27/2020 | 0.20 | 12.0% | 0.0% |
| VSP0033-2 | single experiment | 2985 | NP-OP | 4/27/2020 | 29.89 | 99.9% | 99.5% |
| VSP0034-1a | single experiment | 379000 | NP-OP | 4/27/2020 | 29.81 | 99.8% | 99.8% |
| VSP0034-1b | single experiment | 379000 | NP-OP | 4/27/2020 | NA | NA | NA |
| VSP0034-2 | single experiment | 1895000 | NP-OP | 4/27/2020 | 22.45 | 98.8% | 98.5% |
| VSP0037-1a | single experiment | 154 | NP | 4/29/2020 | 6.66 | 89.3% | 86.2% |
| VSP0037-1b | single experiment | 154 | NP | 4/29/2020 | NA | NA | NA |
| VSP0037-2 | single experiment | 770 | NP | 4/29/2020 | 11.47 | 95.2% | 94.4% |
| VSP0038-1a | single experiment | 412 | OP | 4/29/2020 | 0.80 | 64.3% | 61.0% |
| VSP0038-1b | single experiment | 412 | OP | 4/29/2020 | 0.56 | 64.2% | 62.1% |
| VSP0038-2 | single experiment | 2060 | OP | 4/29/2020 | 22.32 | 99.7% | 98.9% |
| VSP0084-1 | single experiment | NA | Saliva | 05/06/2020 | 4.70 | 78.4% | 77.0% |

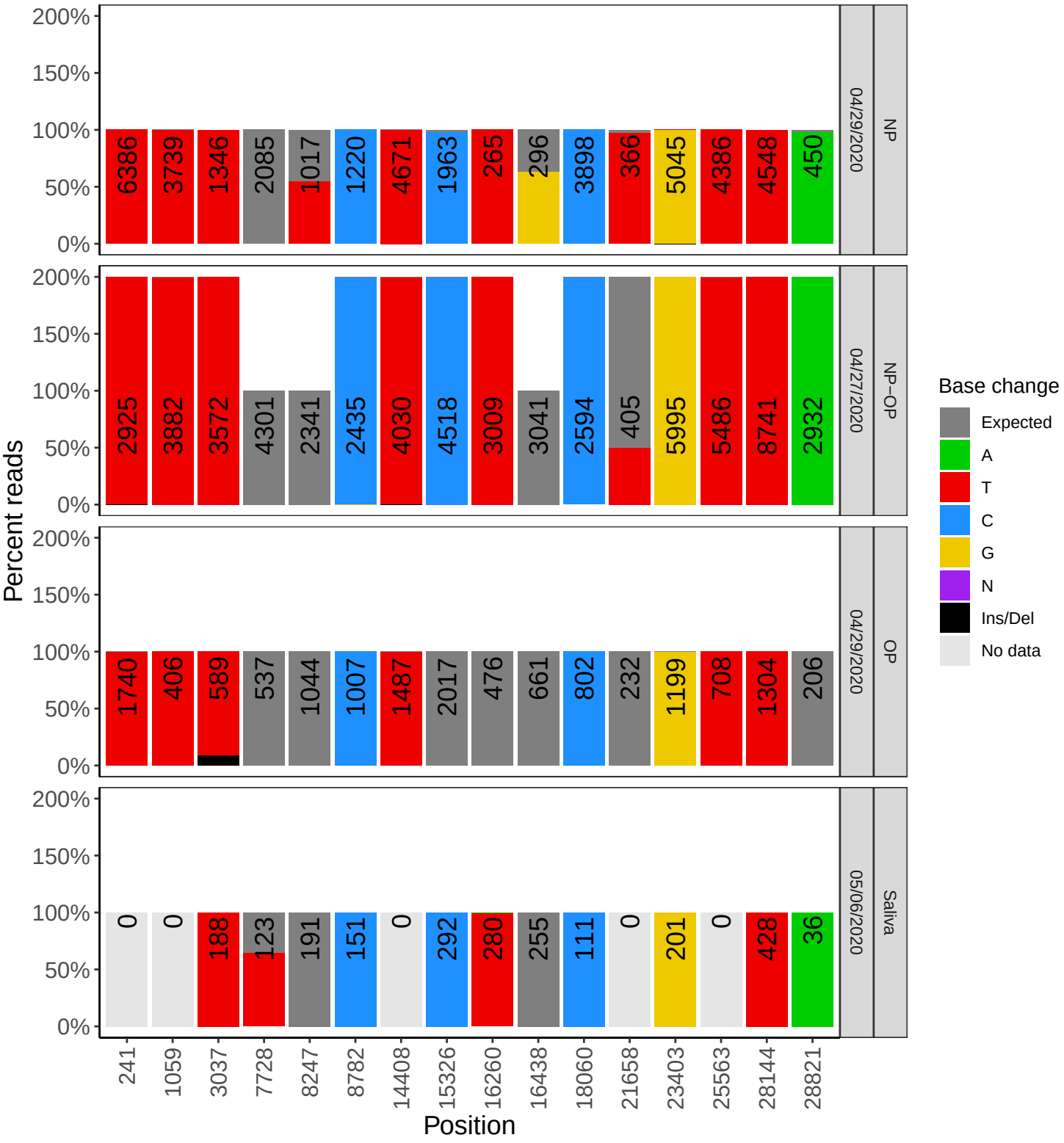
**Variants shared across samples**

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 base composition of tiles are shown in the following plot.



**Relative read abundances of variants**

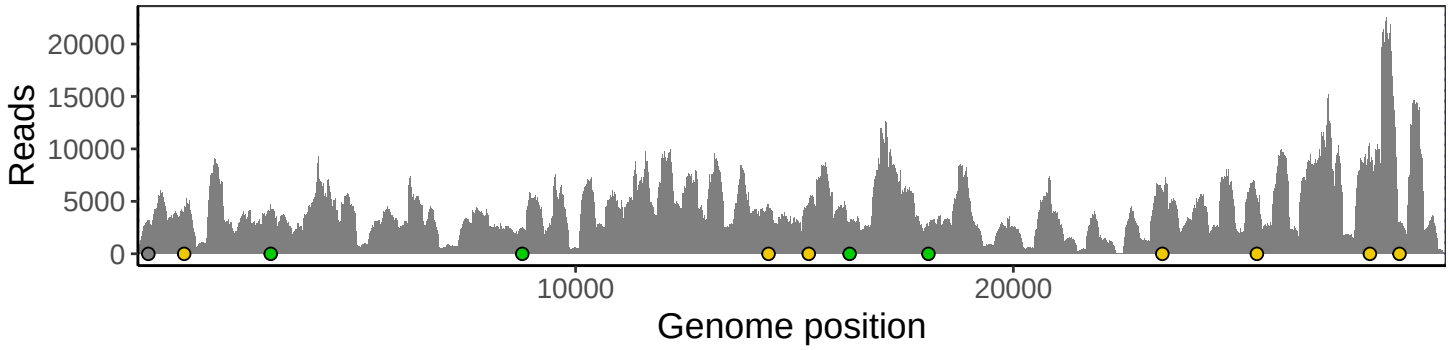
The plot below shows the relative abundances of bases read for each position in the previous variant heatmap where the total number of read pairs covering each position is printed on the stacked bar plots.



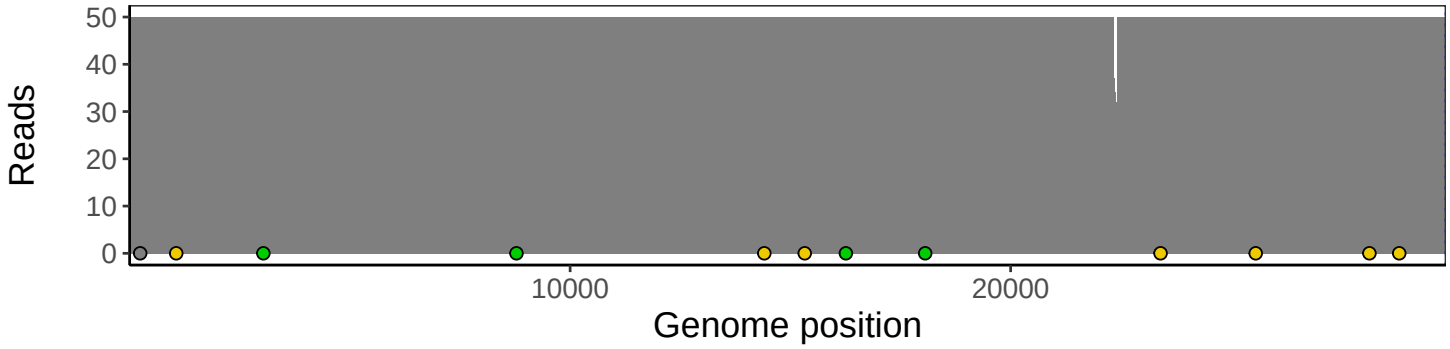
## Analyses of individual experiments and composite results.

### VSP0033 | 4/27/2020 | NP-OP | 237-tri | composite result

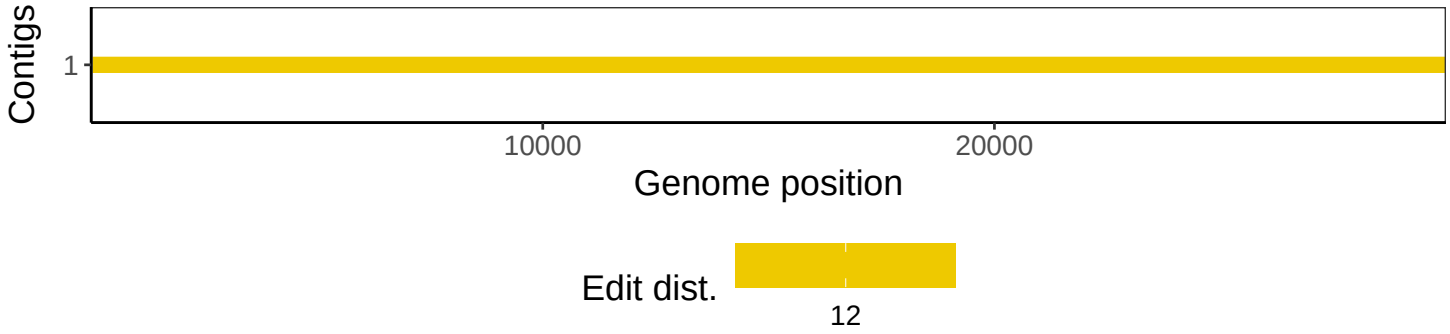
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

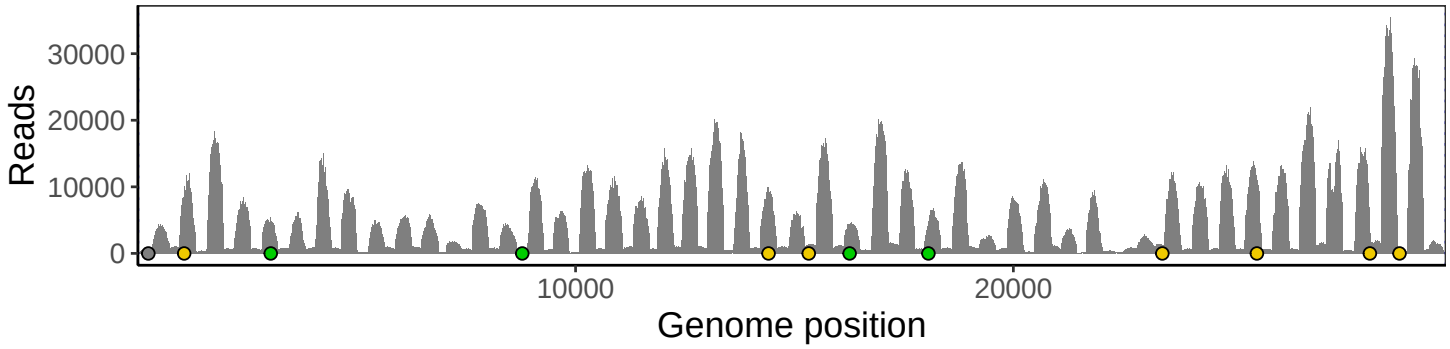



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

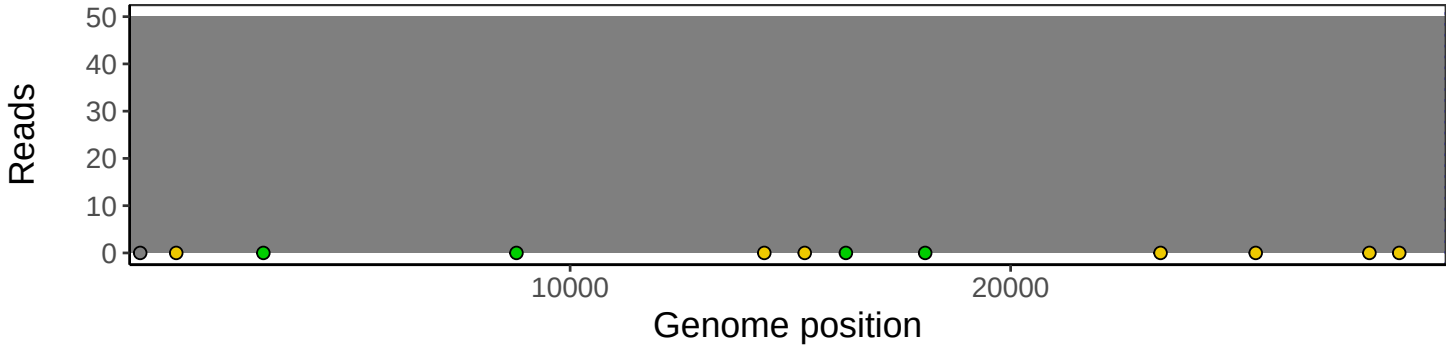


**VSP0034 | 4/27/2020 | NP-OP | 237-qia | composite result**

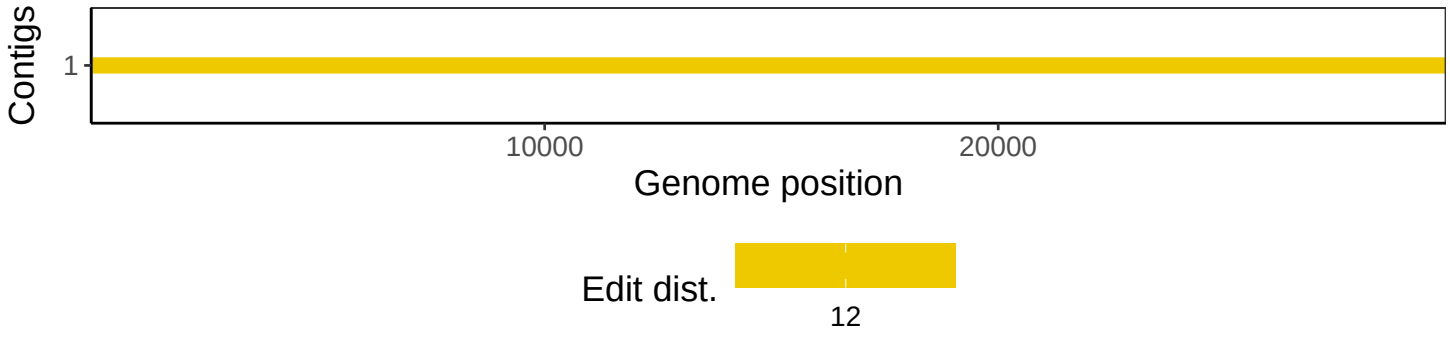
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

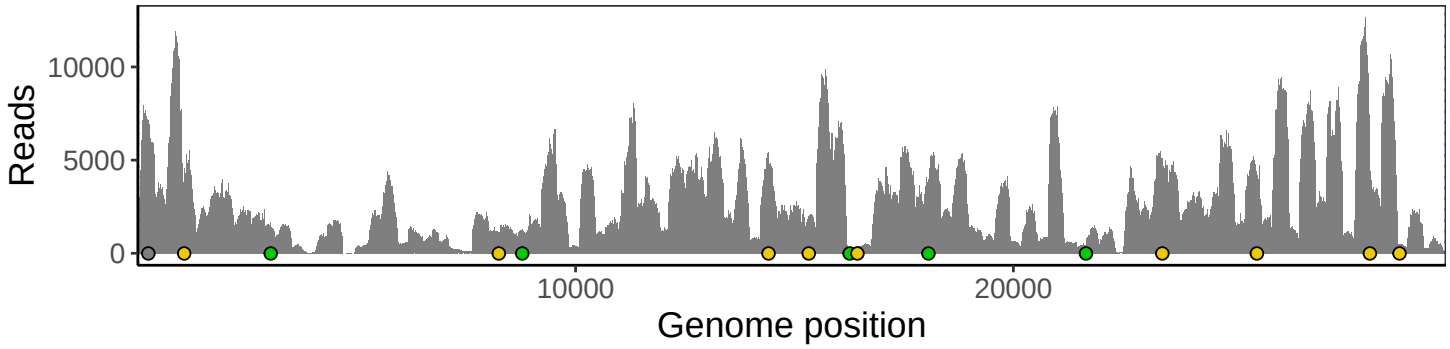



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

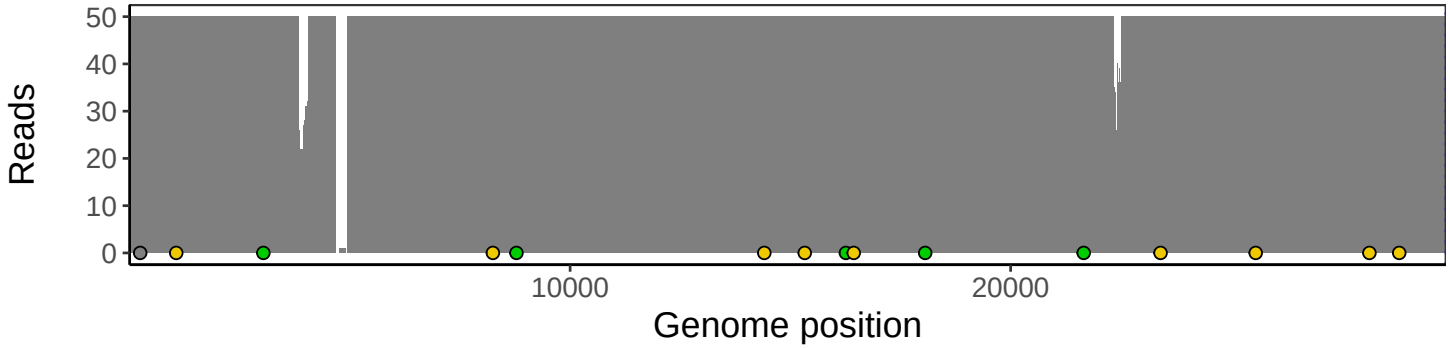


**VSP0037 | 4/29/2020 | NP | 237n-tri | composite result**

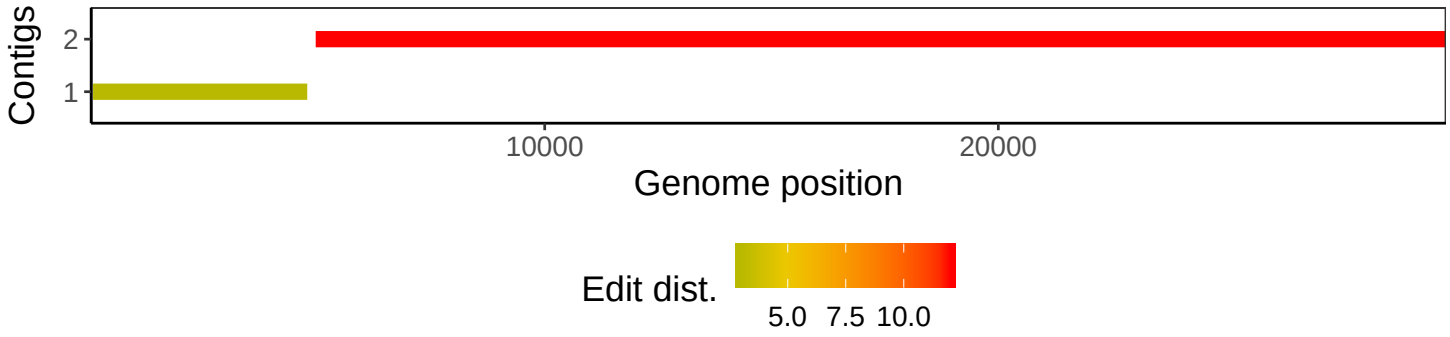
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

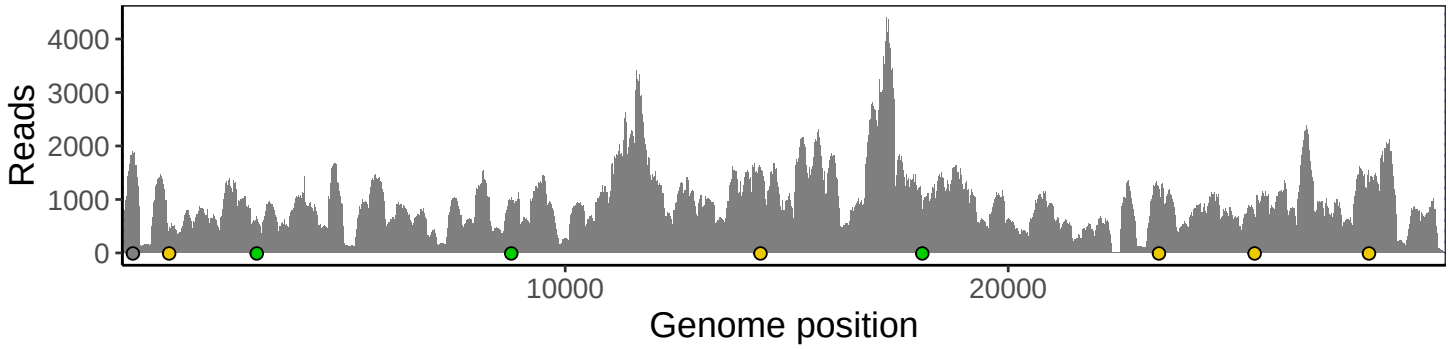



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

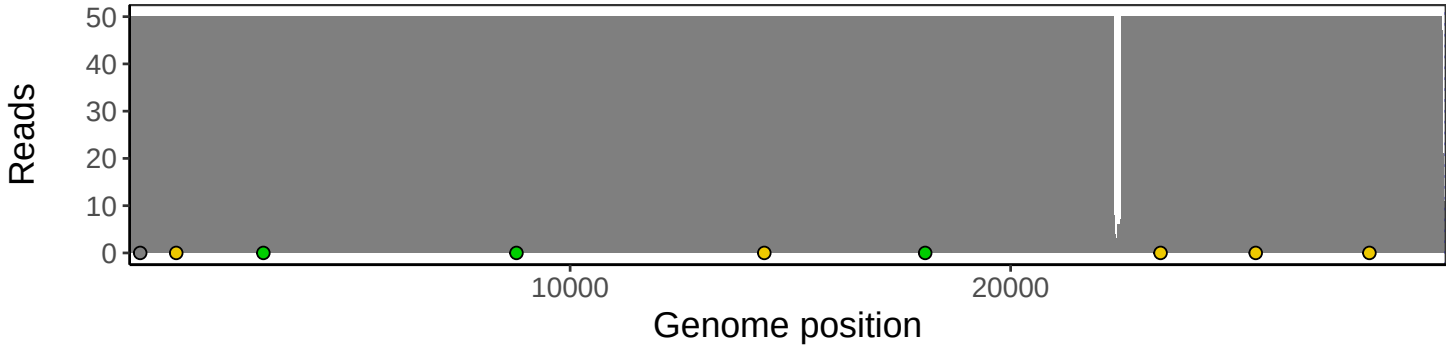


**VSP0038 | 4/29/2020 | OP | 237o-tri | composite result**

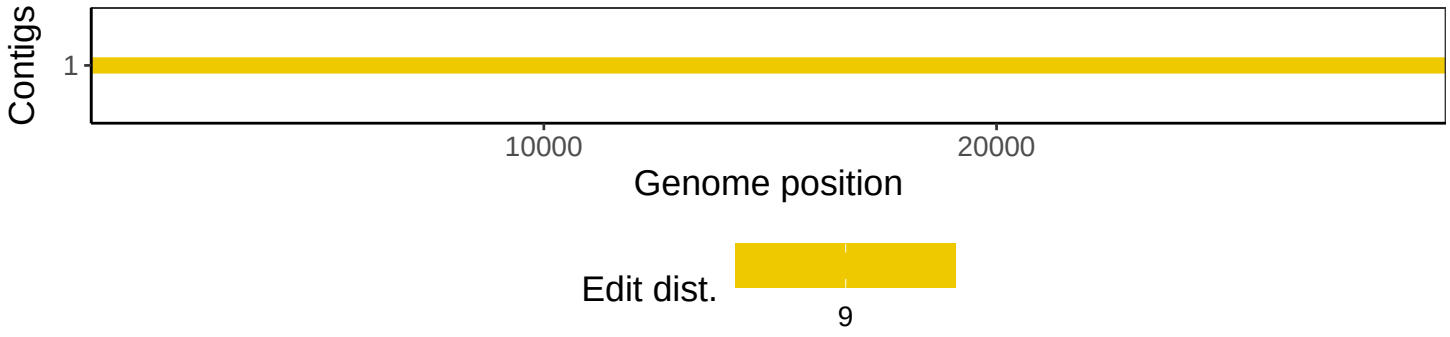
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

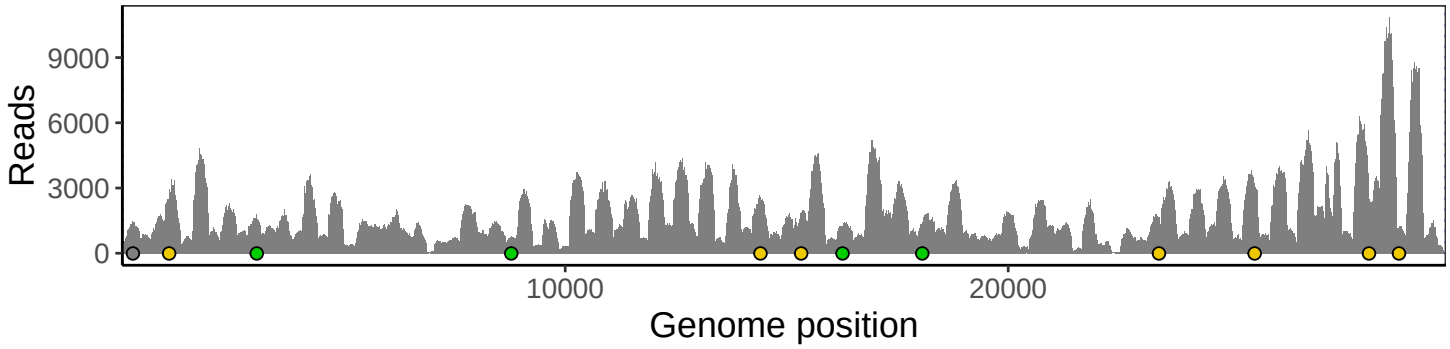



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

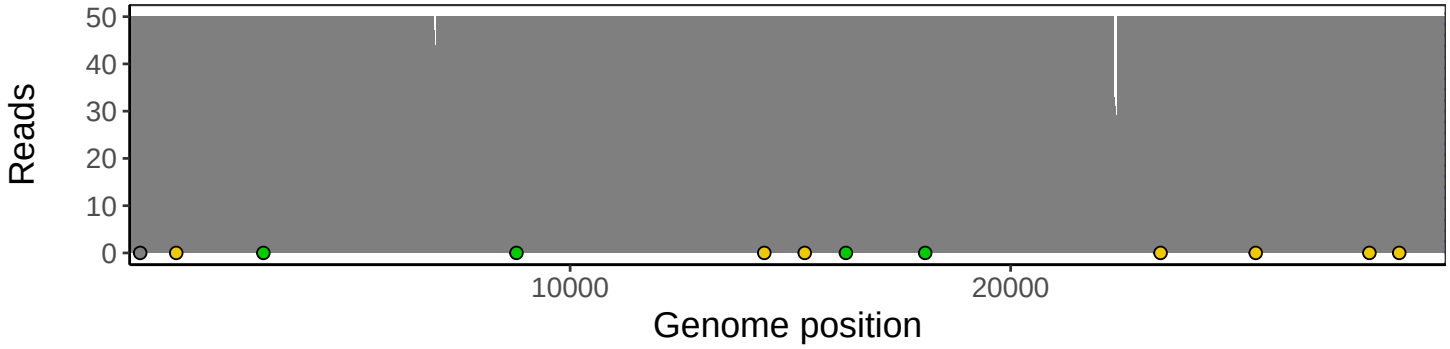


**VSP0033-1a | 4/27/2020 | NP-OP | 237-tri | 597 genomes | single experiment**

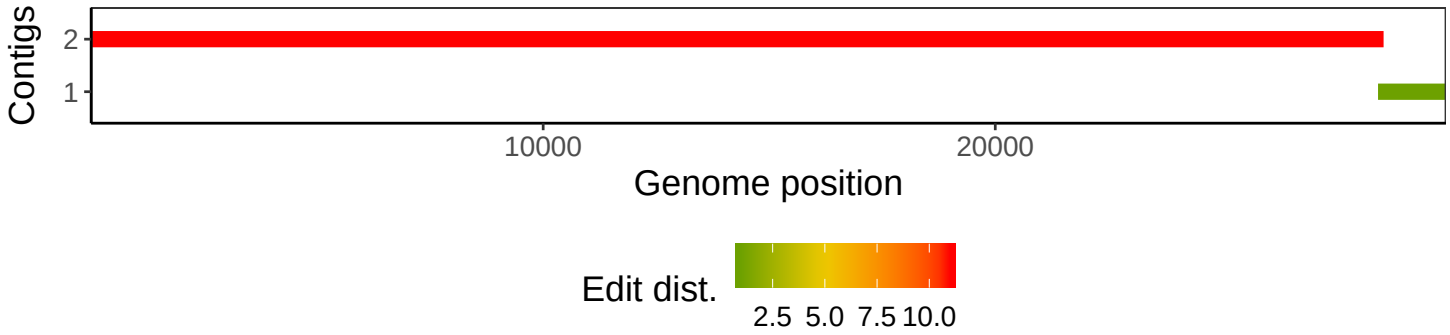
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

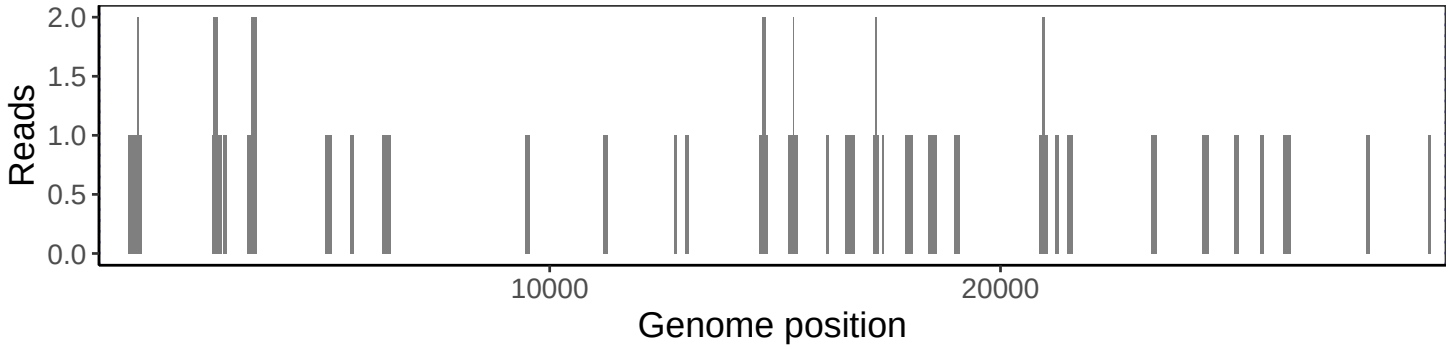



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

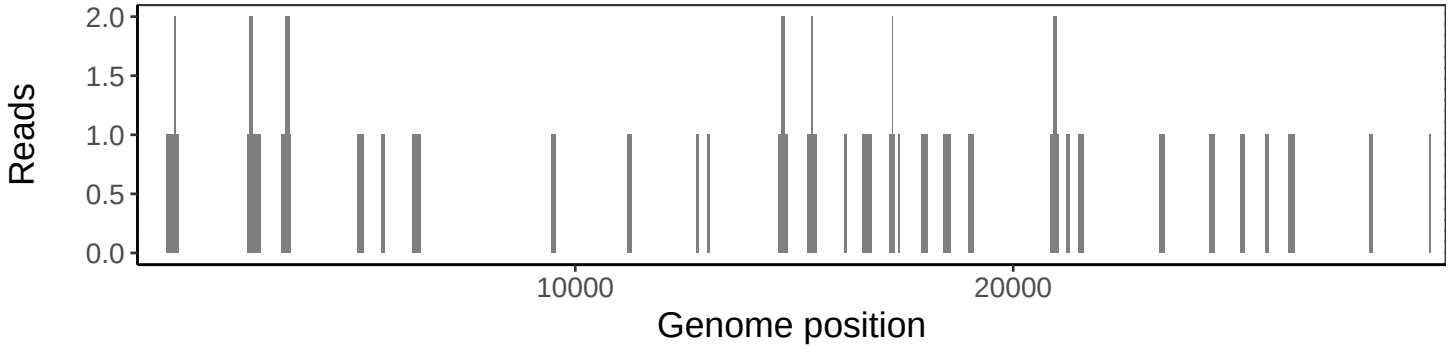


**VSP0033-1b | 4/27/2020 | NP-OP | 237-tri | 597 genomes | single experiment**

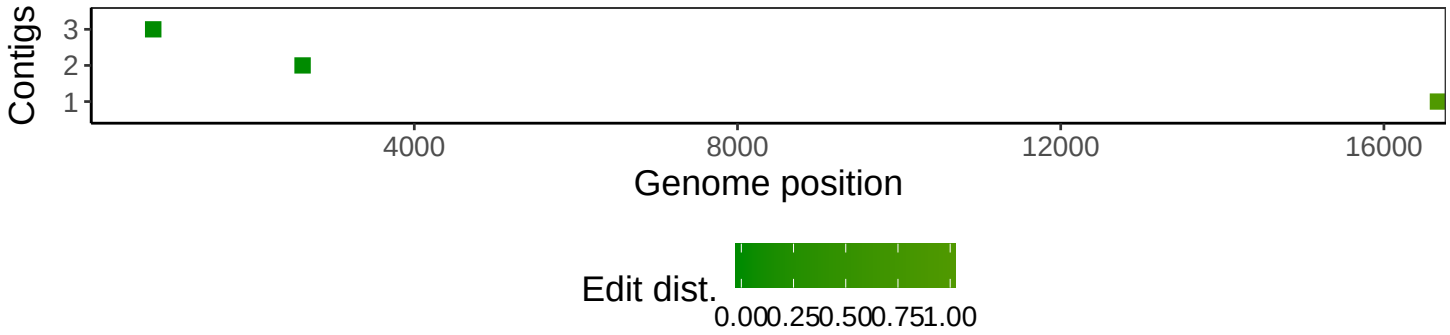
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

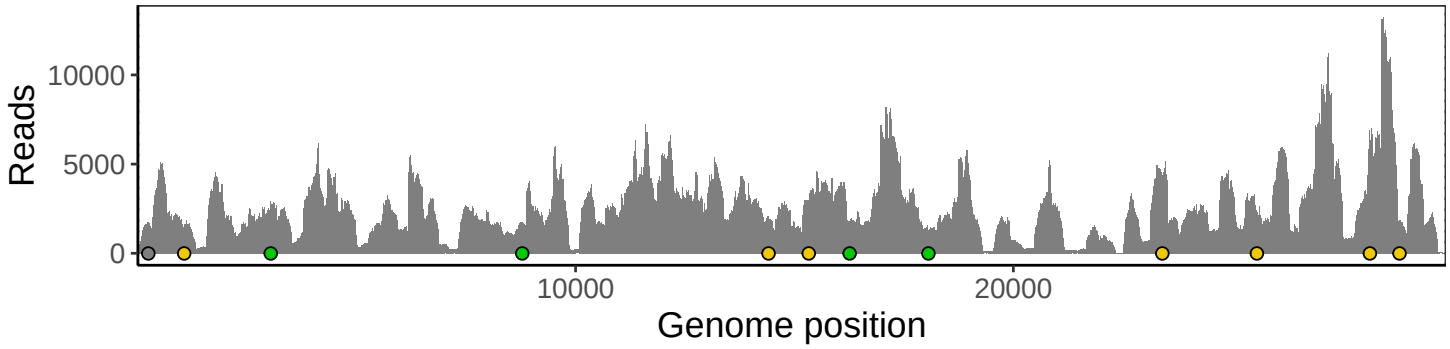



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

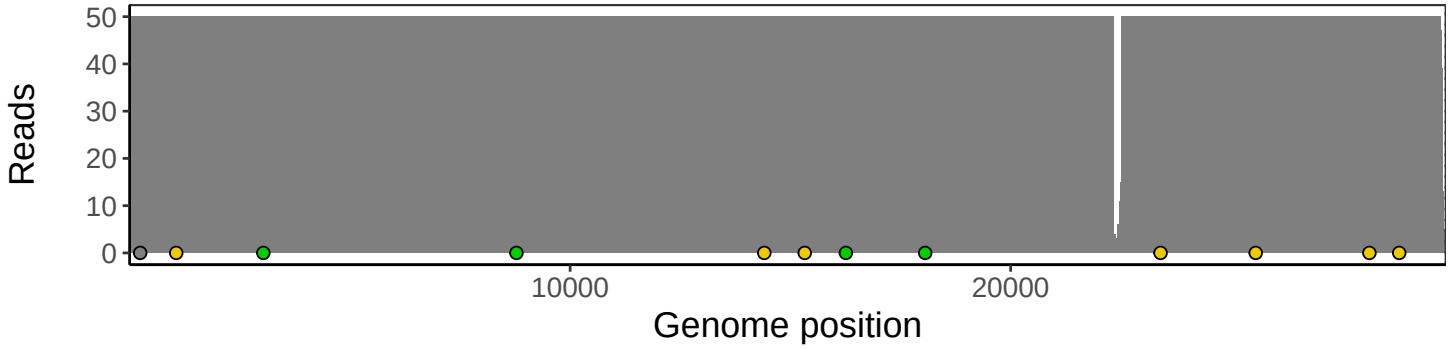


**VSP0033-2 | 4/27/2020 | NP-OP | 237-tri | 2985 genomes | single experiment**

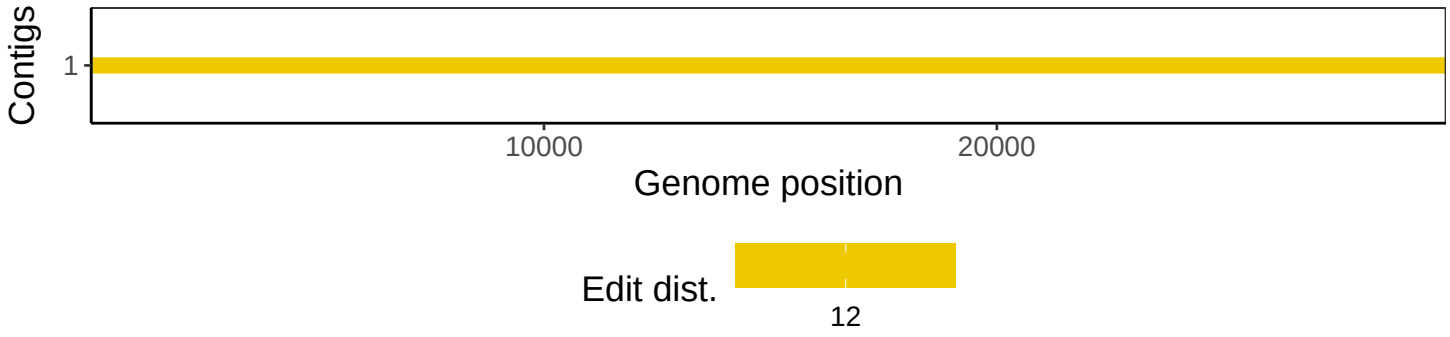
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

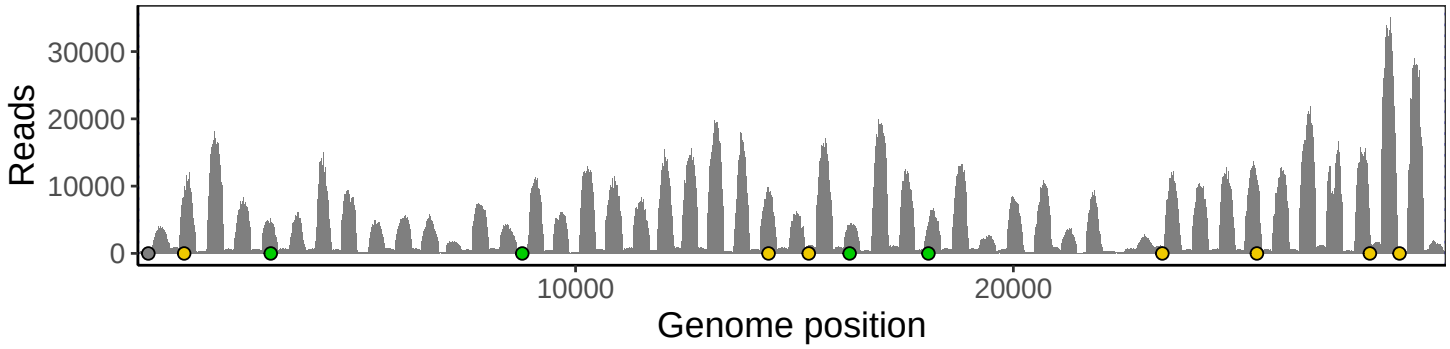



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

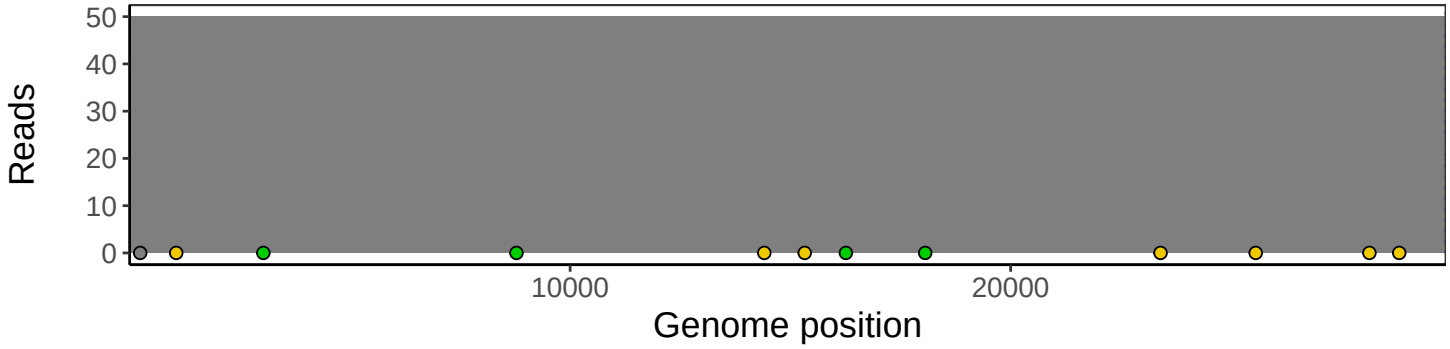


**VSP0034-1a | 4/27/2020 | NP-OP | 237-qia | 379000 genomes | single experiment**

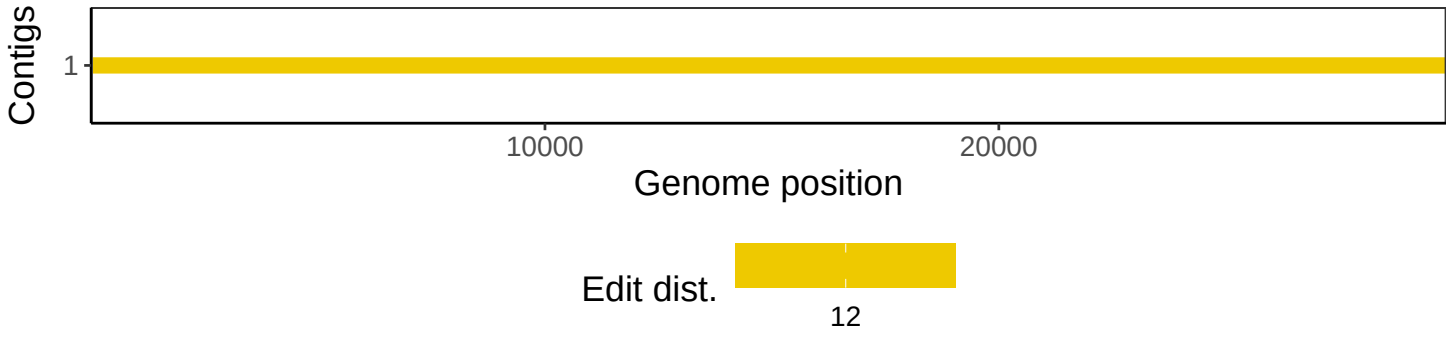
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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.

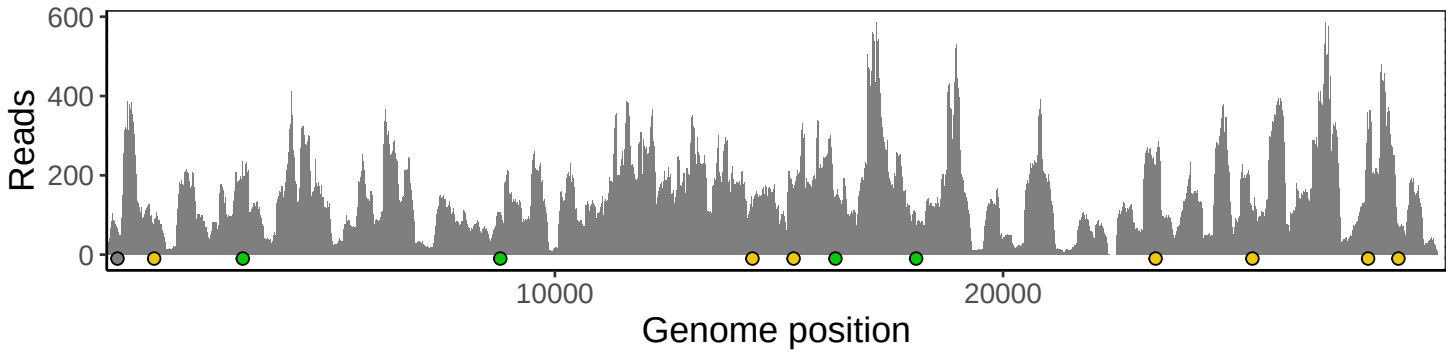


**No pileup data available.**

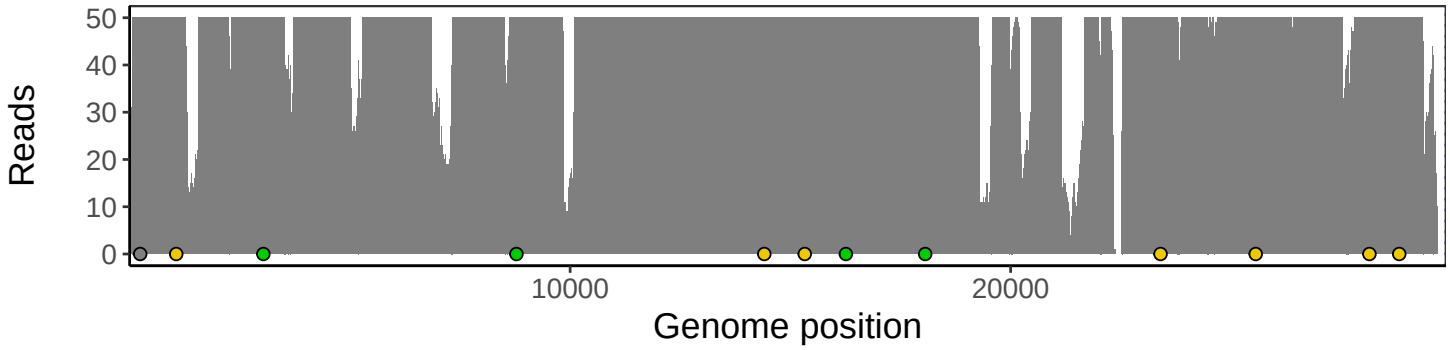
**No contig data available.**

**VSP0034-2 | 4/27/2020 | NP-OP | 237-qia | 1895000 genomes | single experiment**

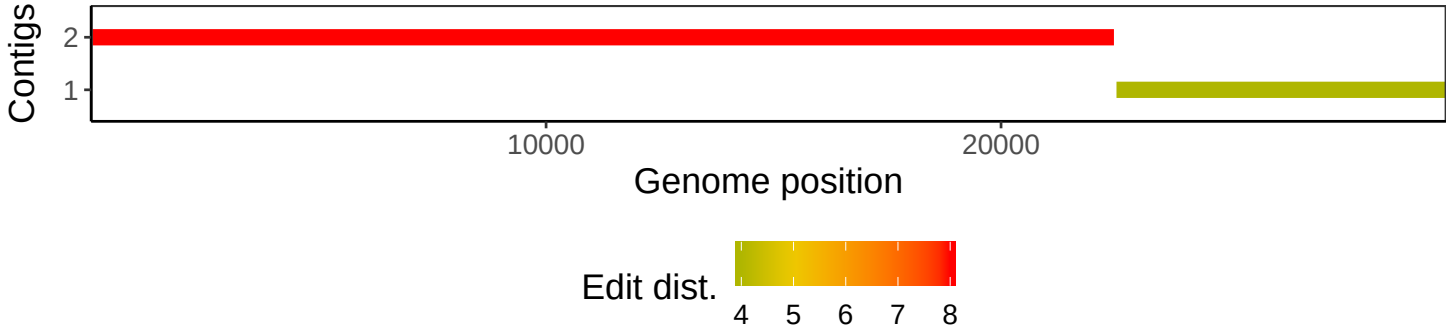
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

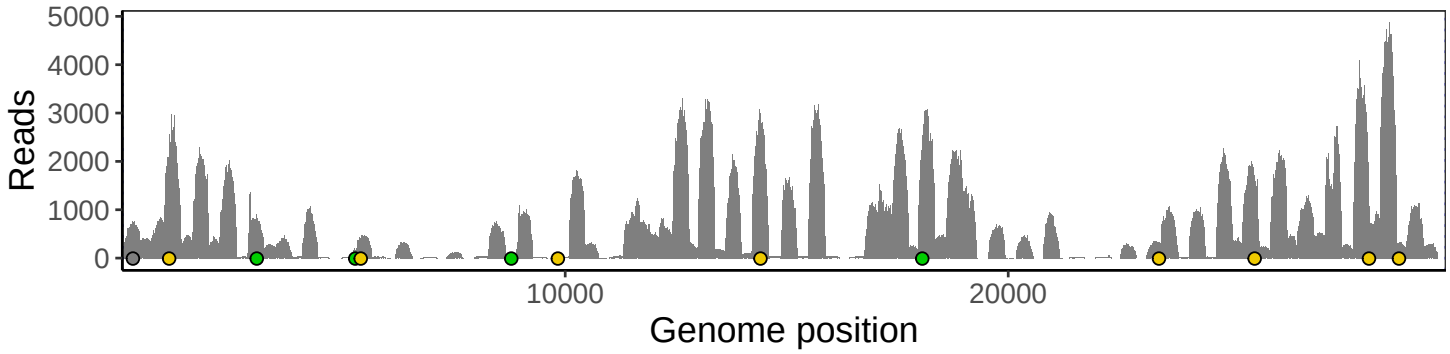



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

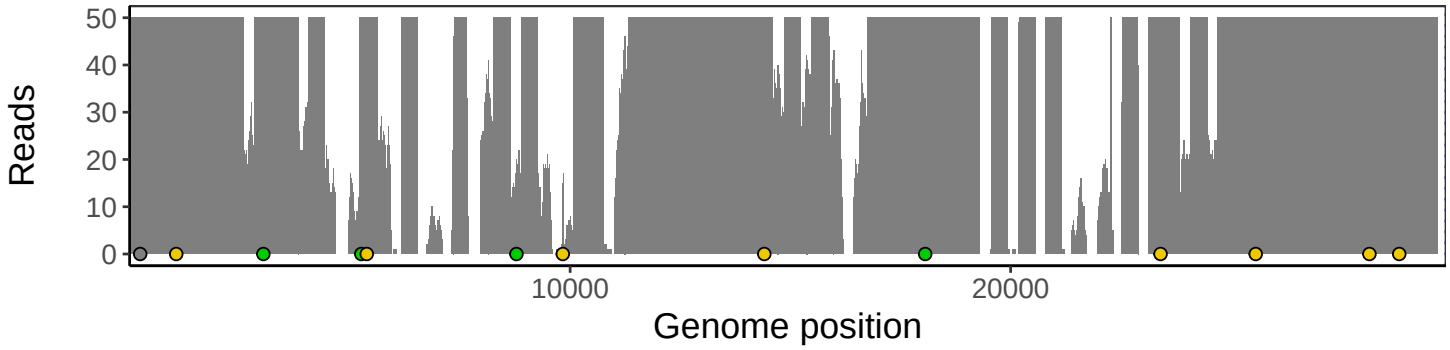


**VSP0037-1a | 4/29/2020 | NP | 237n-tri | 154 genomes | single experiment**

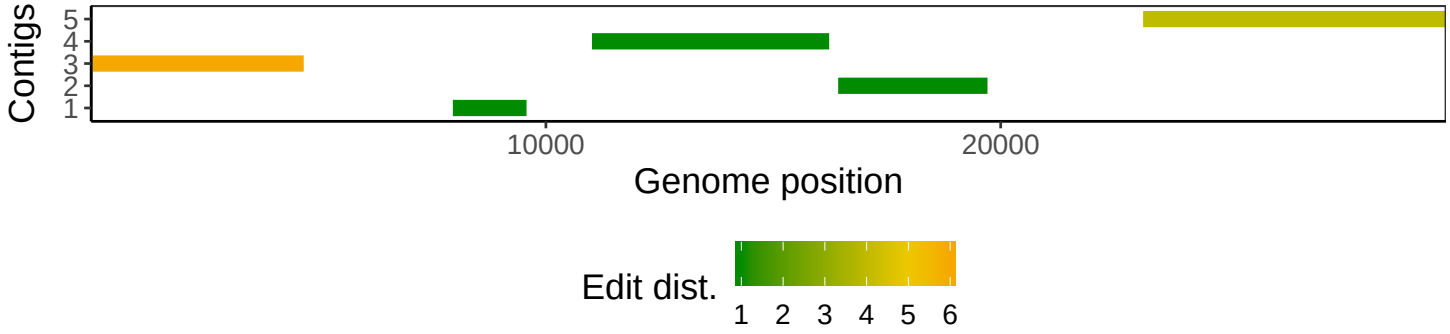
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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.

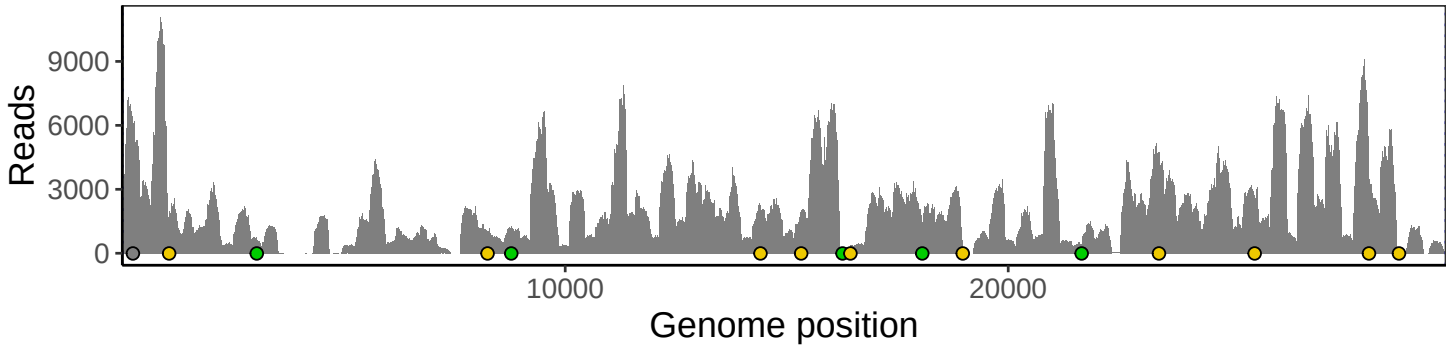


**No pileup data available.**

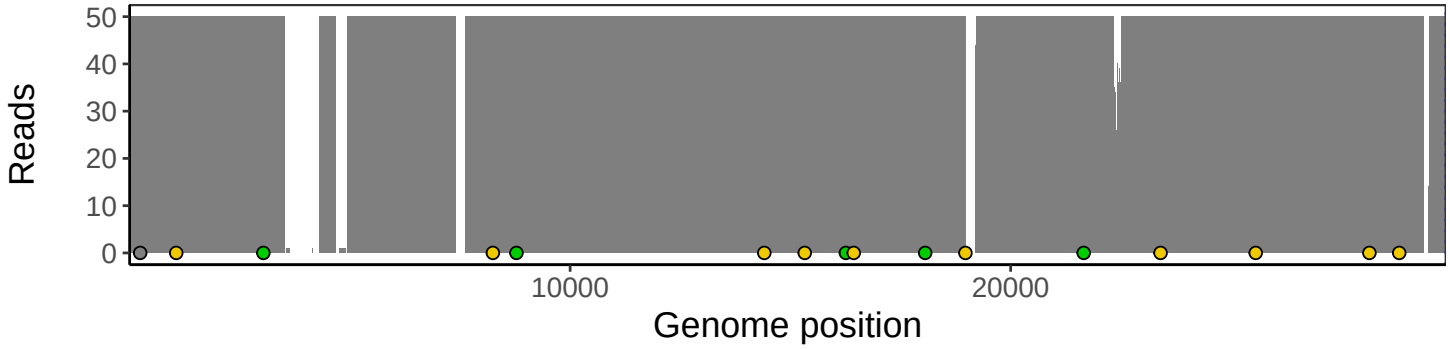
**No contig data available.**

**VSP0037-2 | 4/29/2020 | NP | 237n-tri | 770 genomes | single experiment**

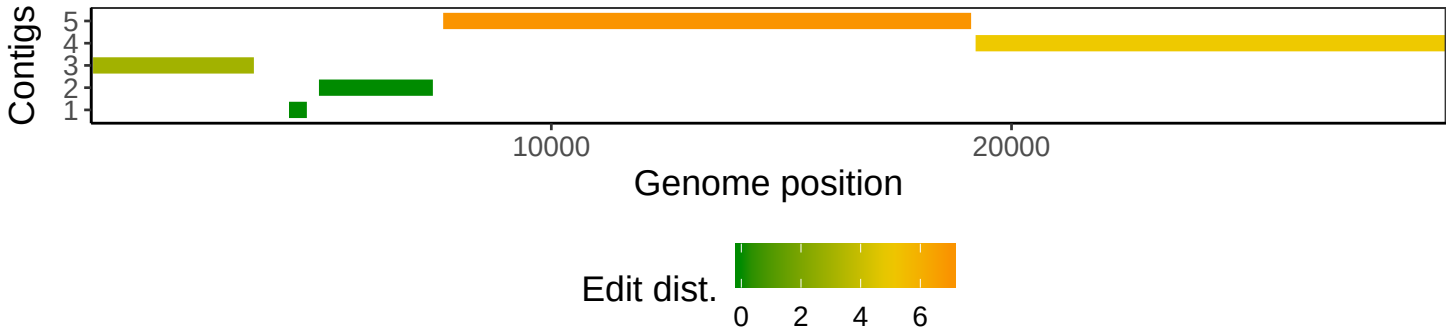
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

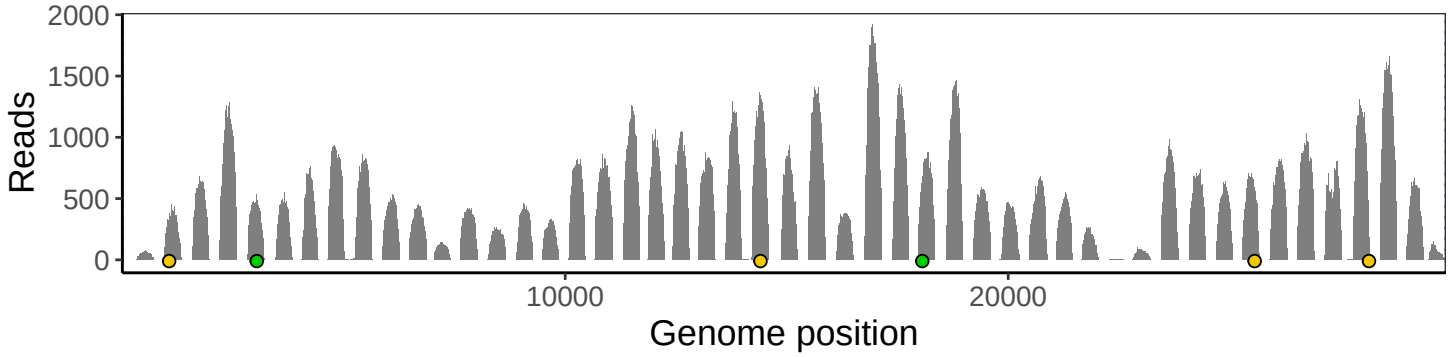



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

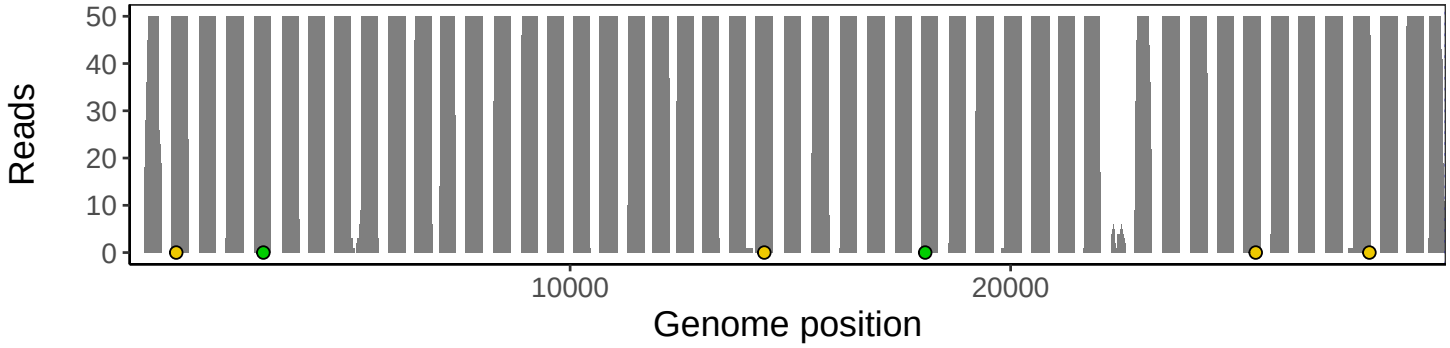


**VSP0038-1a | 4/29/2020 | OP | 237o-tri | 412 genomes | single experiment**

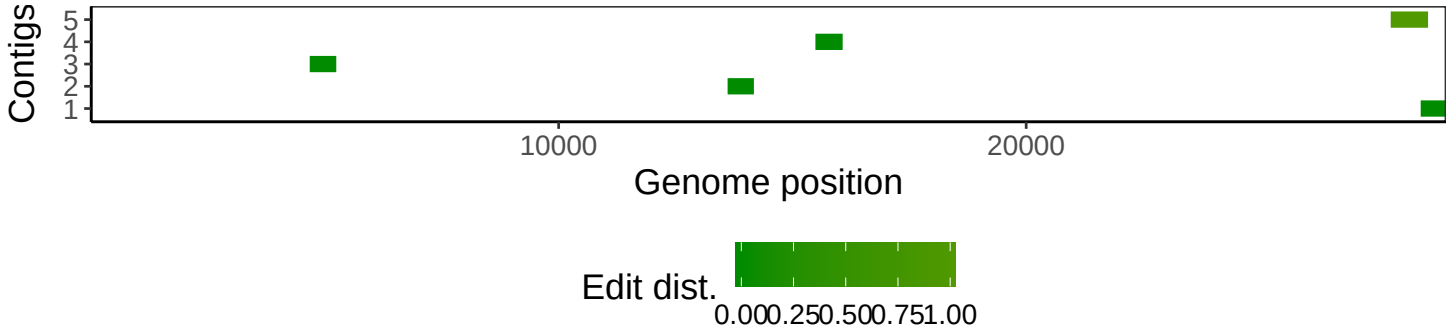
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

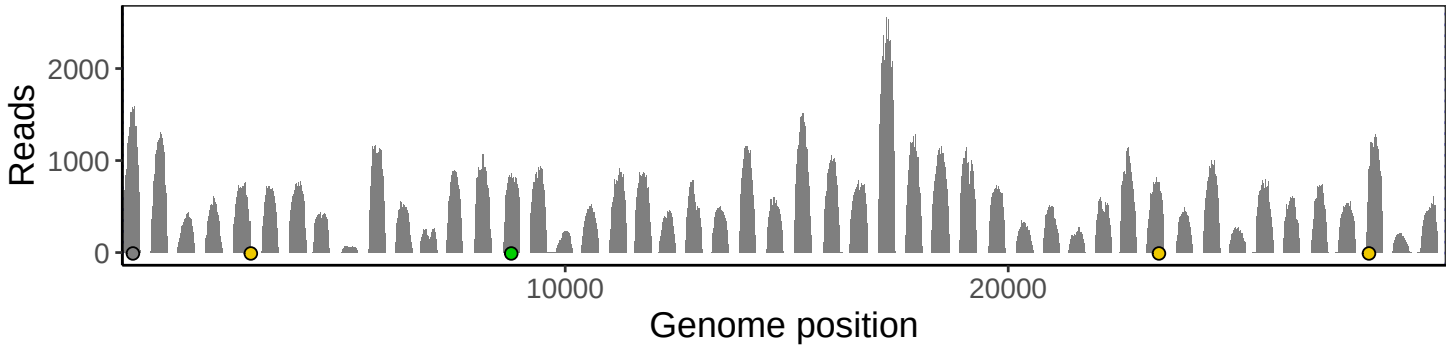



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

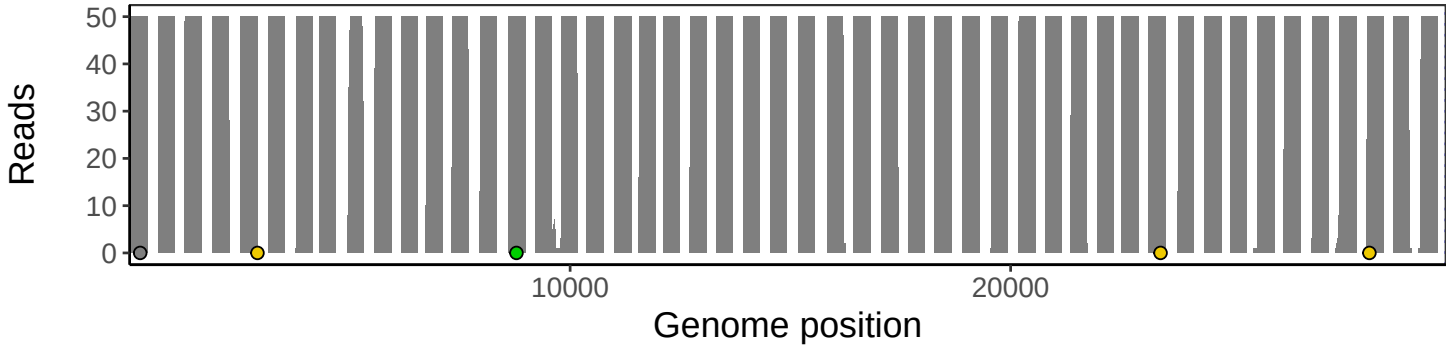


**VSP0038-1b | 4/29/2020 | OP | 237o-tri | 412 genomes | single experiment**

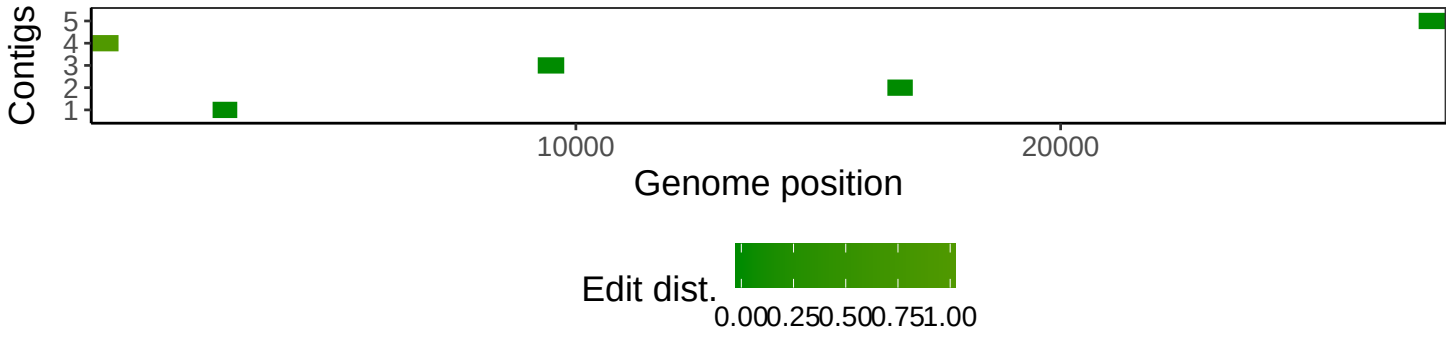
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

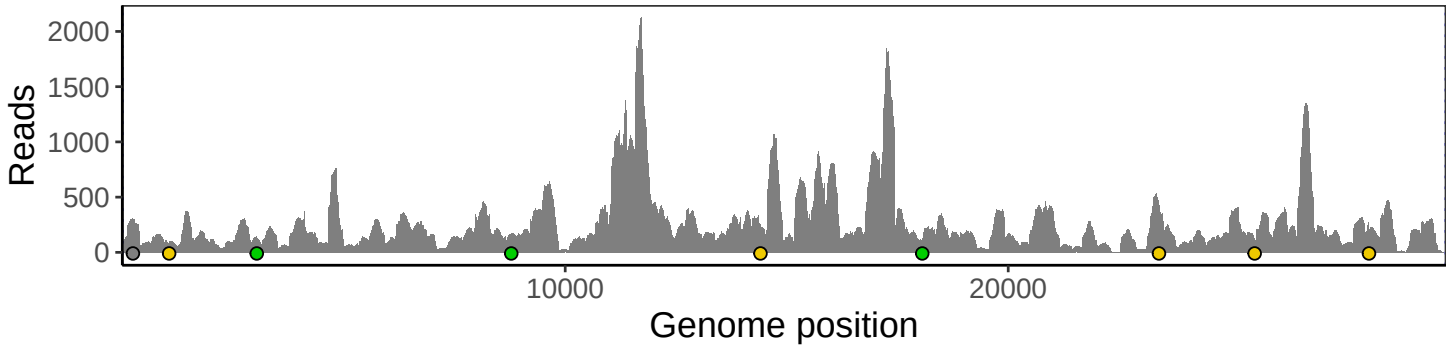



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

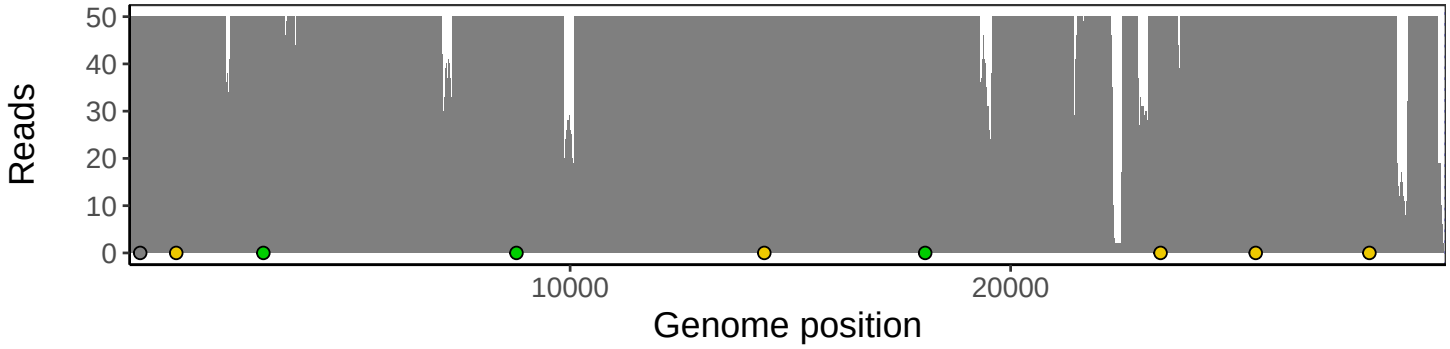


**VSP0038-2 | 4/29/2020 | OP | 237o-tri | 2060 genomes | single experiment**

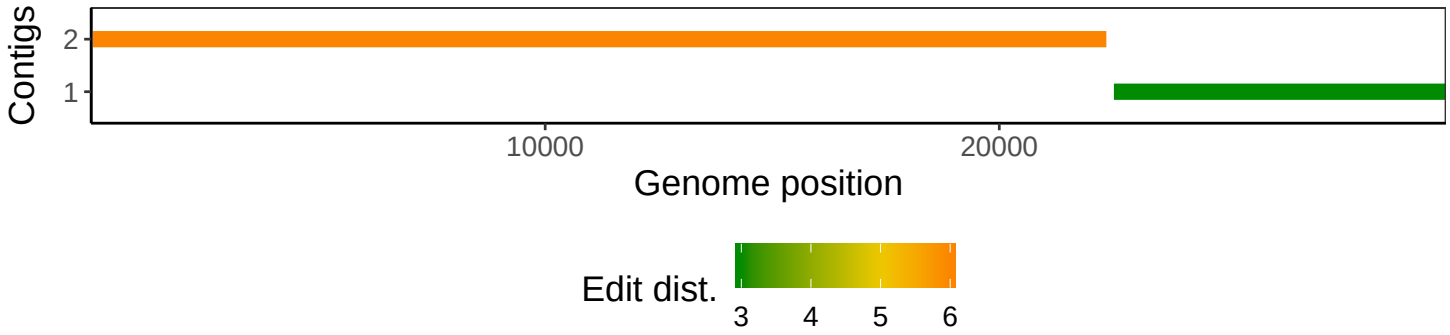
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.




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

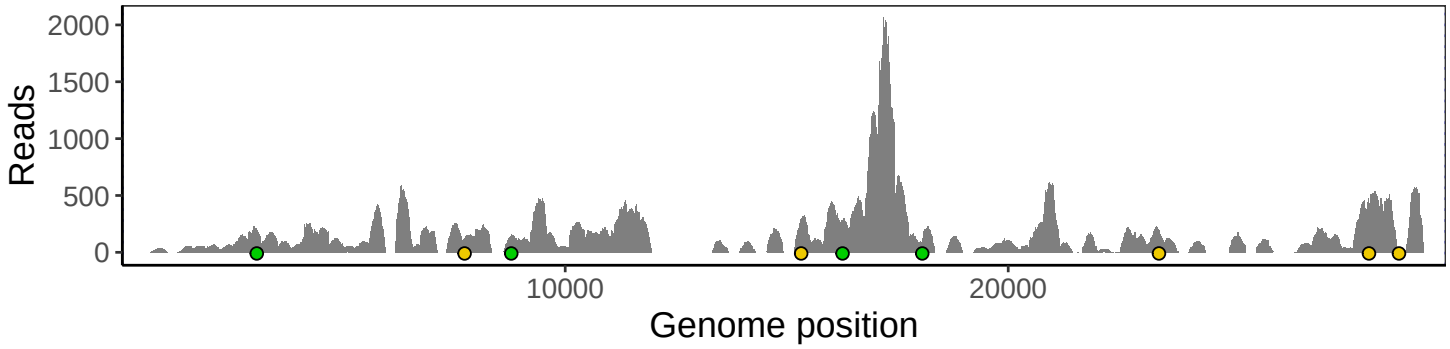



The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

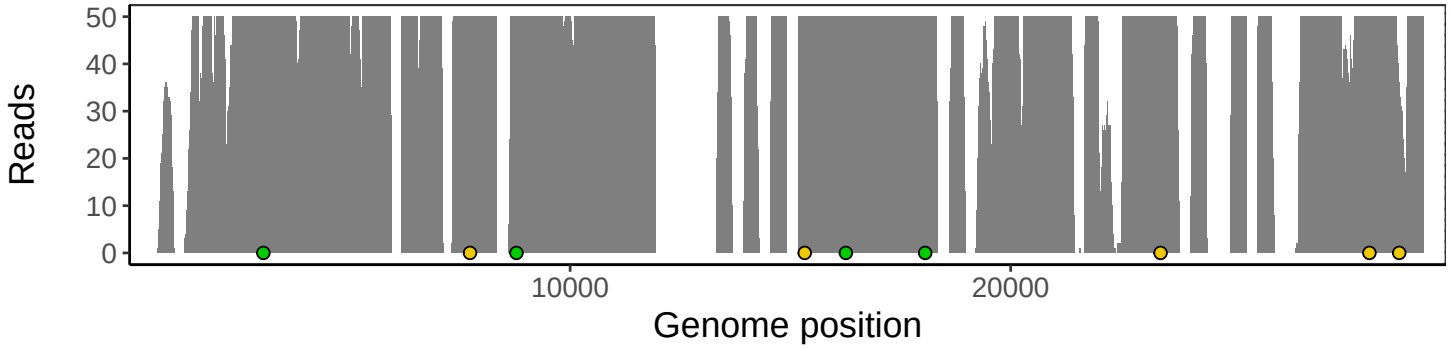


**VSP0084-1 | 05/06/2020 | Saliva | 237s-q | genomes | single experiment**

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.