COVID-19 subject 239

2020-08-26

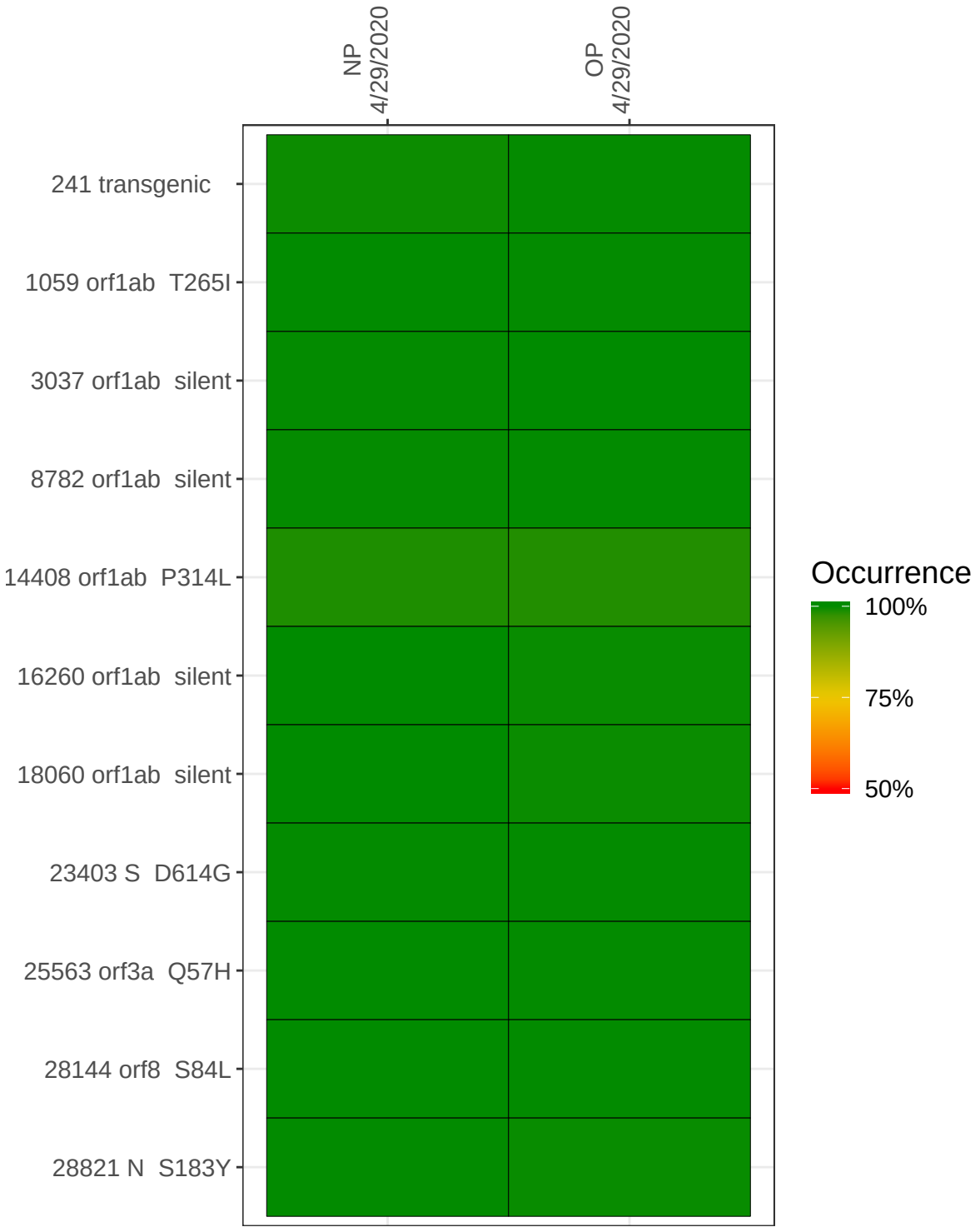
The table below provides a summary of subject samples for which sequencing data is available.
The experiments column shows the number of sequencing experiments performed for each specimen.
Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0041	composite	NA	NP	4/29/2020	30.06	99.9%	99.8%
VSP0042	composite	NA	OP	4/29/2020	27.48	99.8%	99.8%
VSP0041-1a	single experiment	12500	NP	4/29/2020	29.83	99.8%	99.8%
VSP0041-1b	single experiment	12500	NP	4/29/2020	NA	NA	NA
VSP0041-2	single experiment	NA	NP	4/29/2020	30.13	99.8%	99.5%
VSP0042-1a	single experiment	6490	OP	4/29/2020	1.12	70.2%	63.2%
VSP0042-1b	single experiment	6490	OP	4/29/2020	NA	NA	NA
VSP0042-2	single experiment	6490	OP	4/29/2020	1.09	81.0%	64.7%

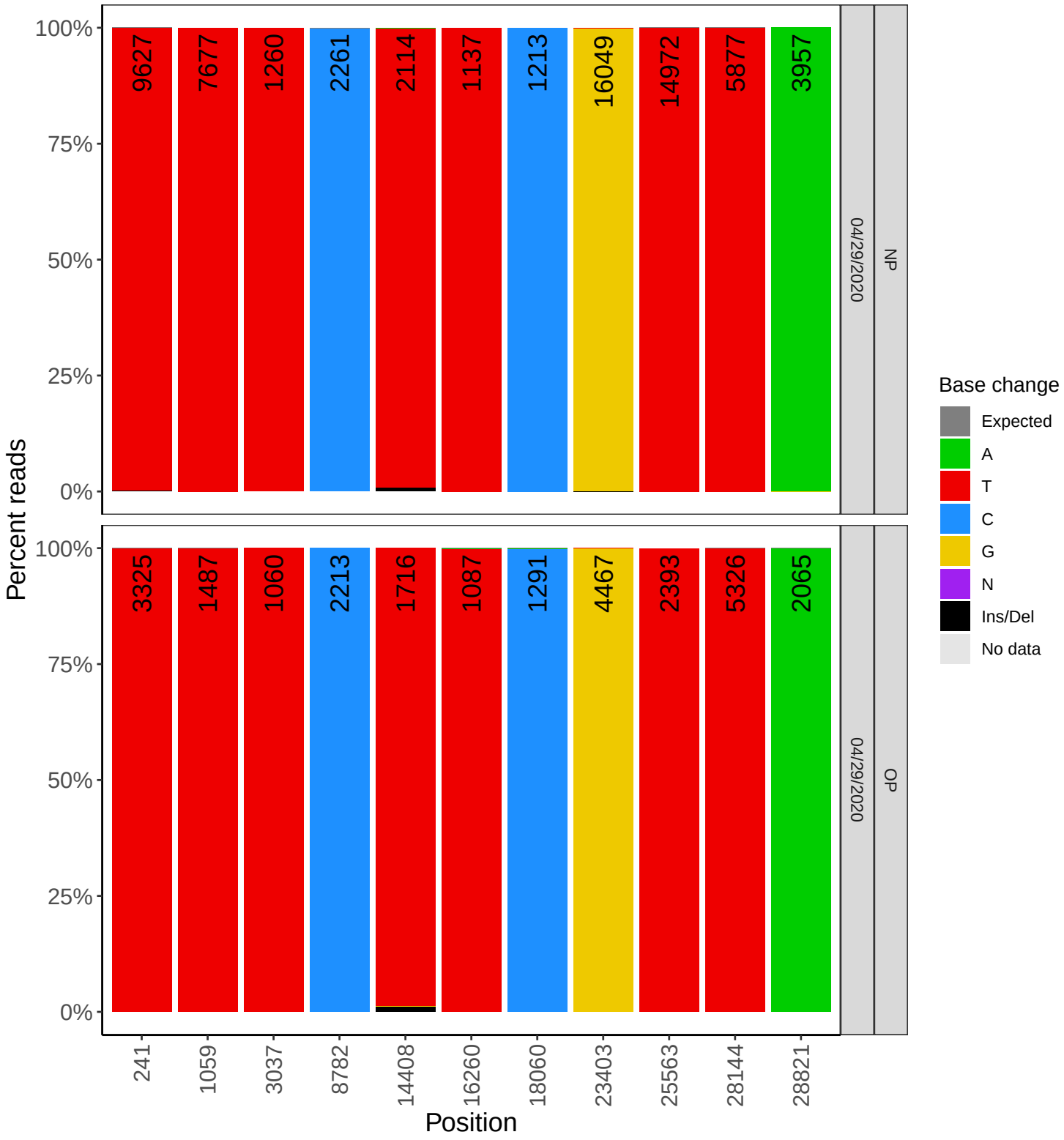
Variants shared across samples

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

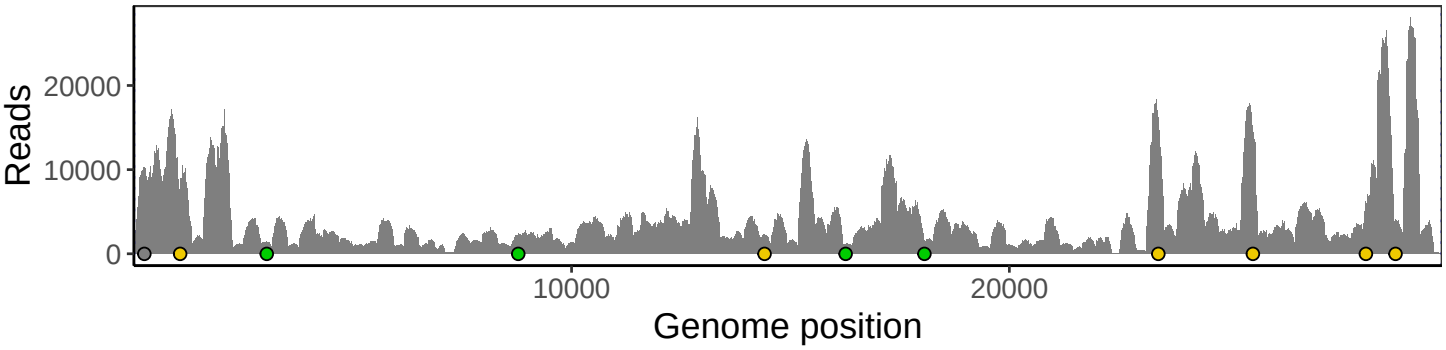
The plot below shows the relative abundances of bases read for each position in the previous variant heatmap where the total number of read pairs covering each position is printed on the stacked bar plots.



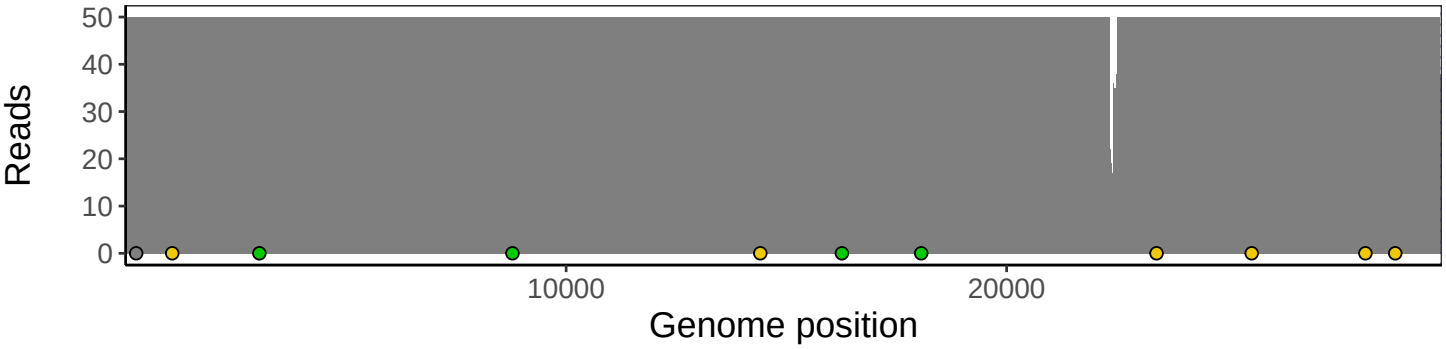
Analyses of individual experiments and composite results.

VSP0041 | 4/29/2020 | NP | 239n-tri | composite result

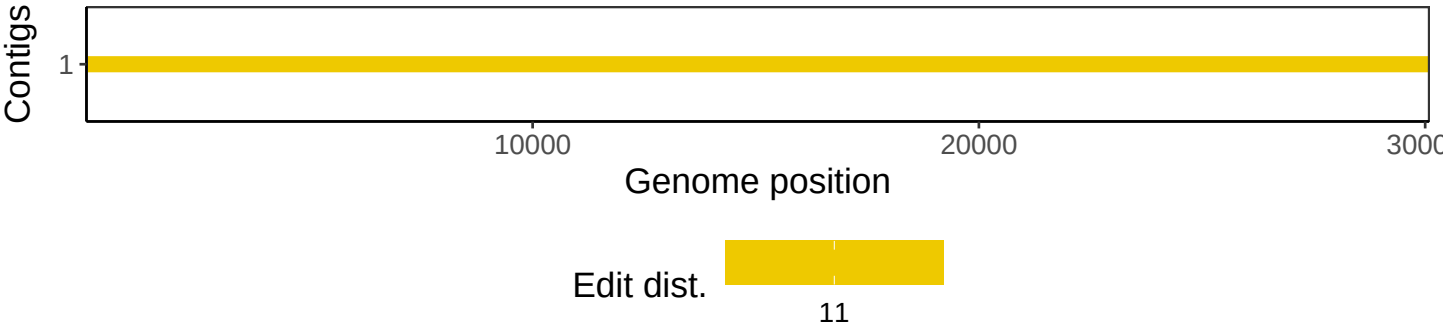
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

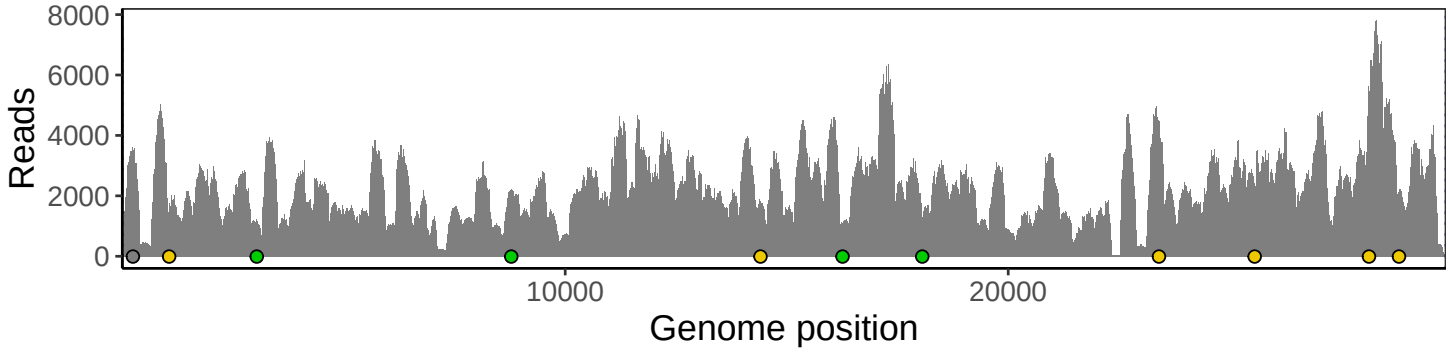


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

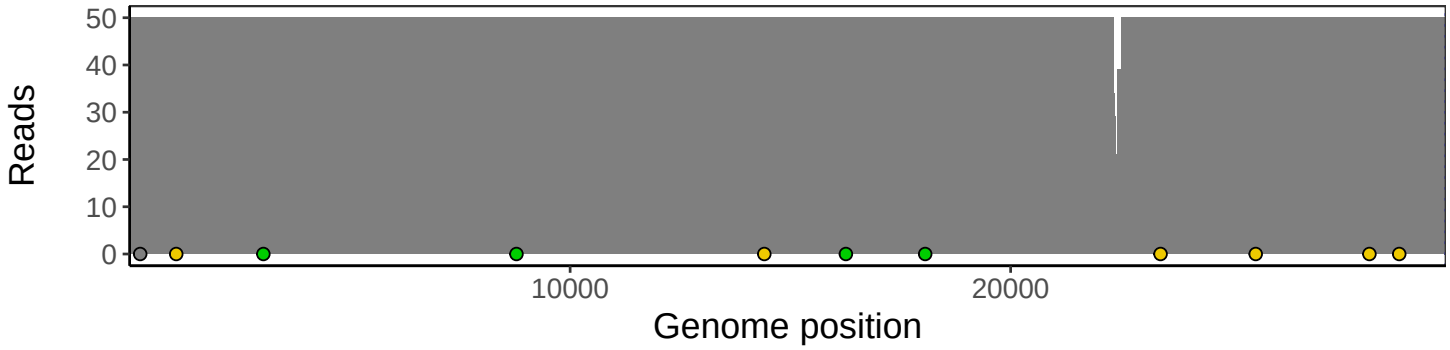


VSP0042 | 4/29/2020 | OP | 239o-tri | composite result

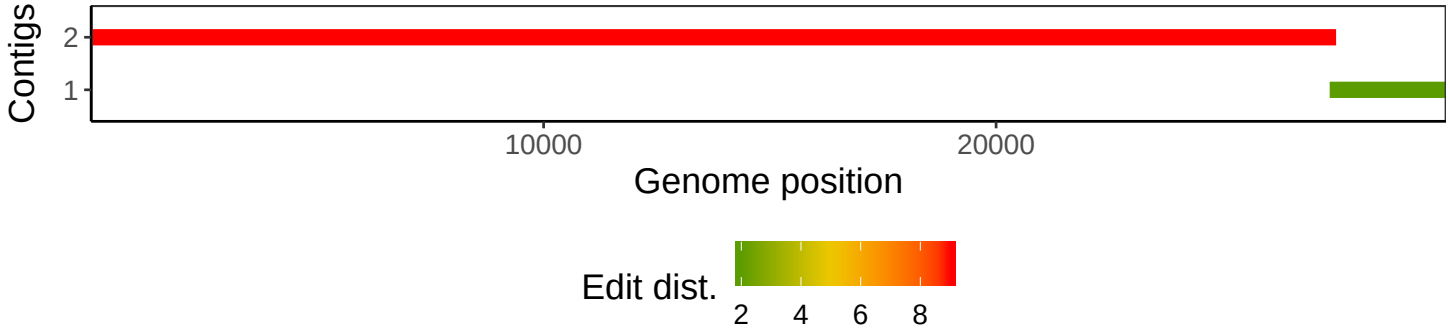
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

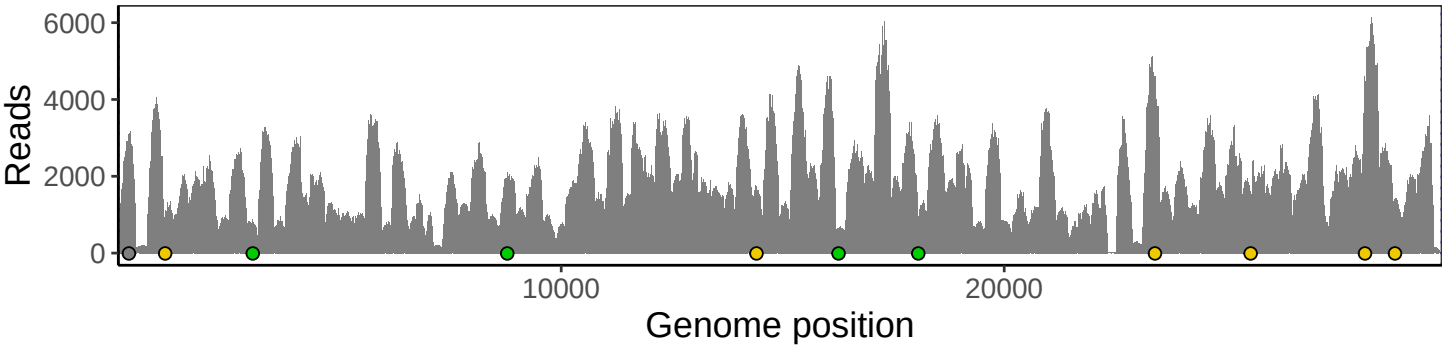


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

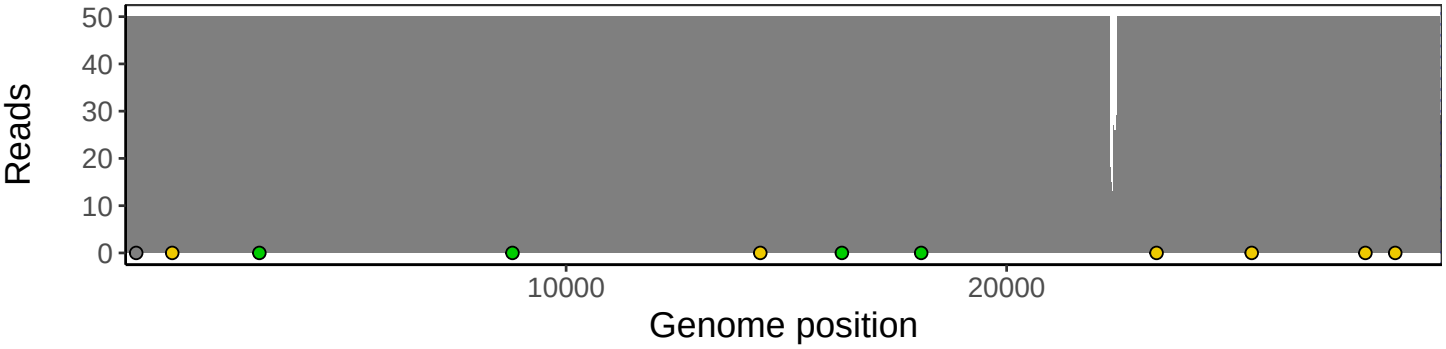


VSP0041-1a | 4/29/2020 | NP | 239n-tri | 12500 genomes | single experiment

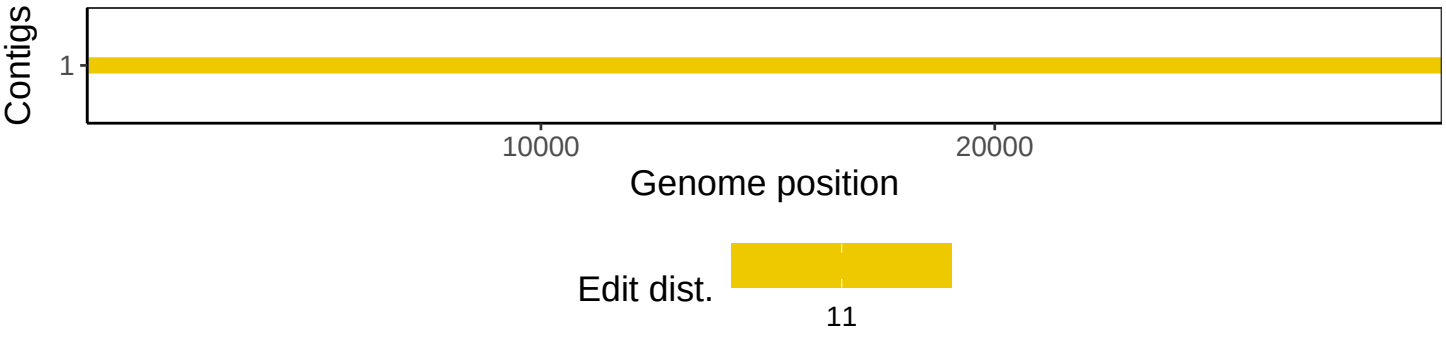
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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.

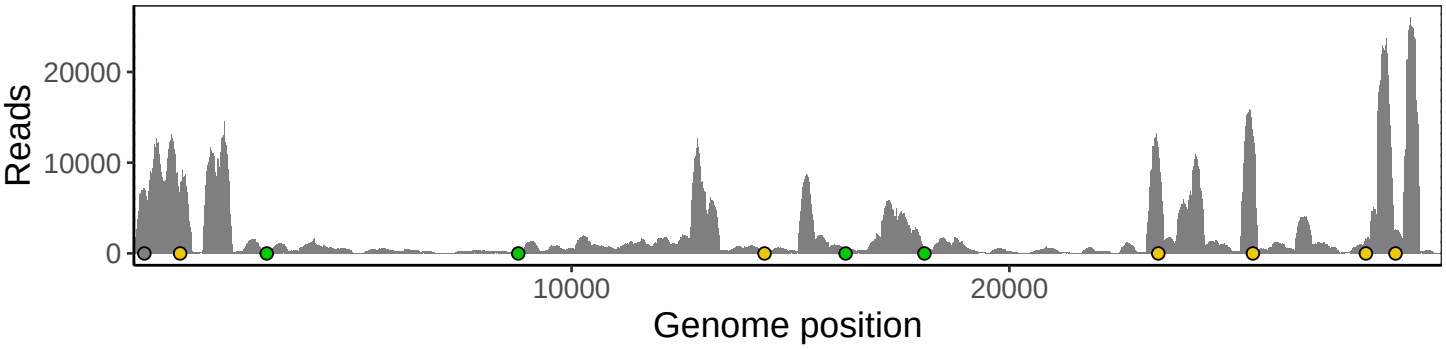


No pileup data available.

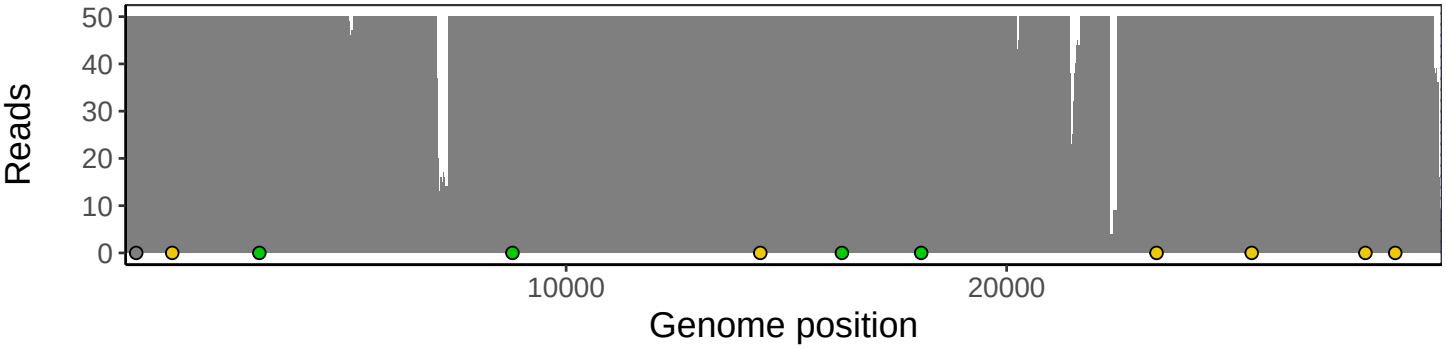
No contig data available.

VSP0041-2 | 4/29/2020 | NP | 239n-tri | genomes | single experiment

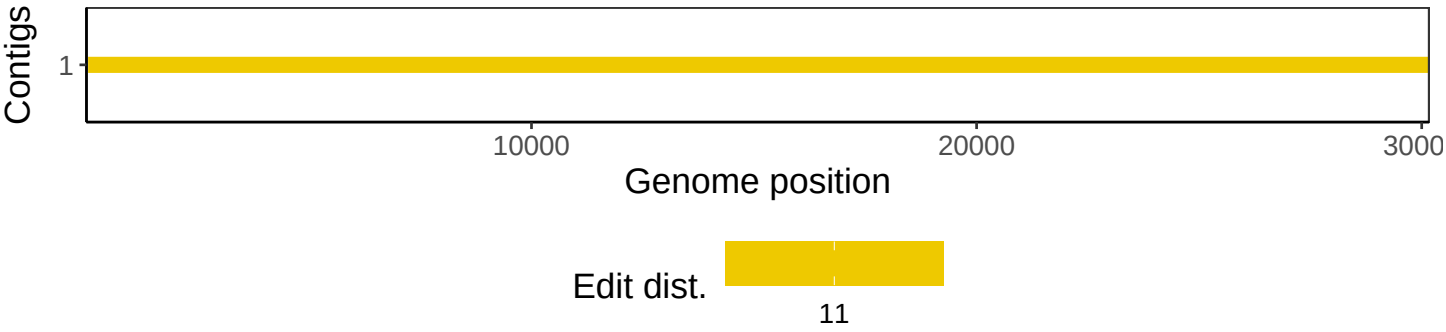
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

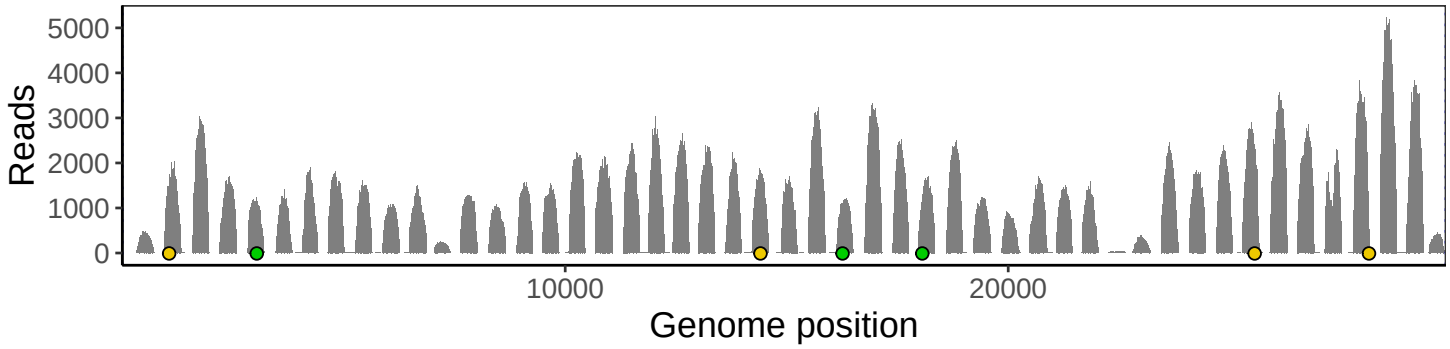


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

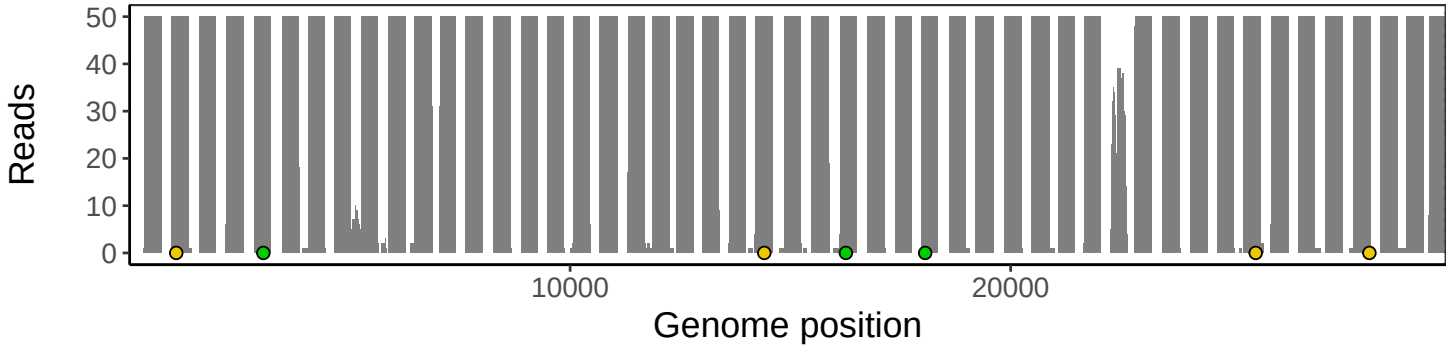


VSP0042-1a | 4/29/2020 | OP | 239o-tri | 6490 genomes | single experiment

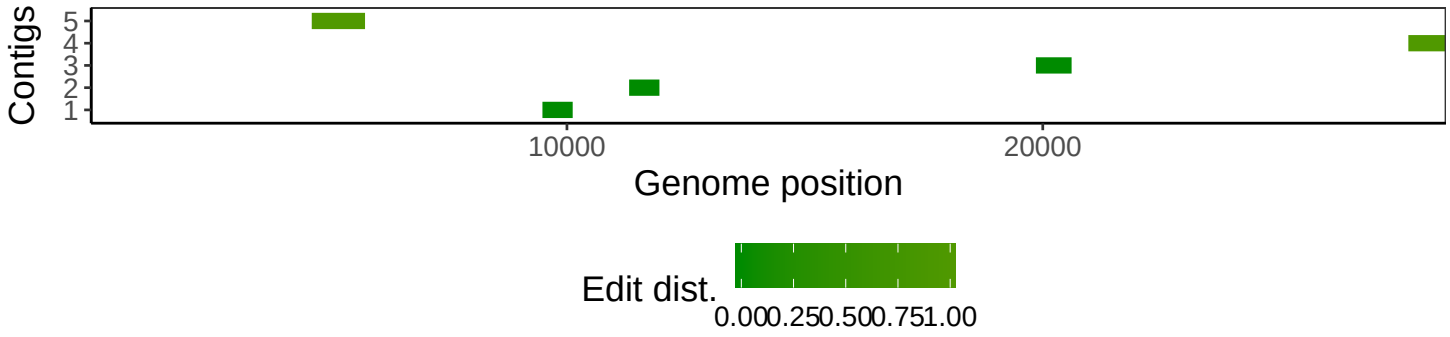
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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.

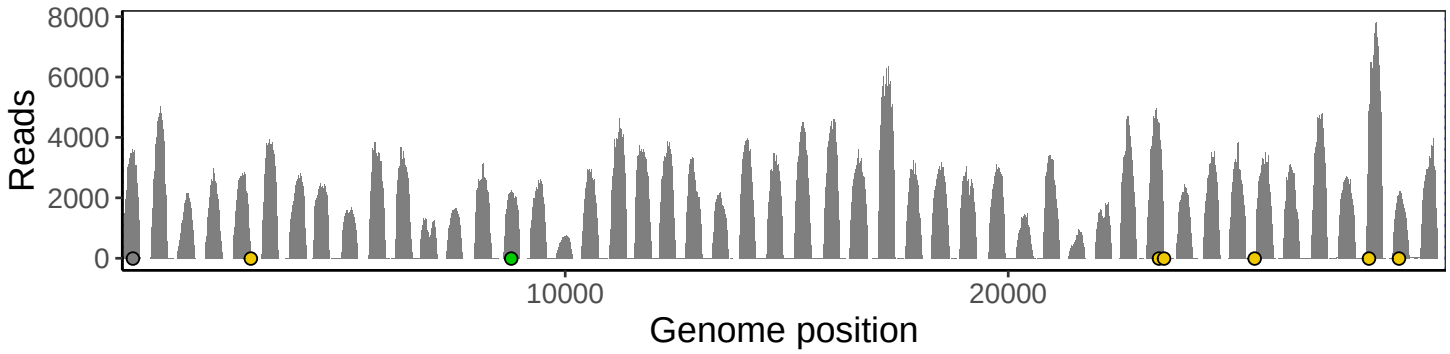


No pileup data available.

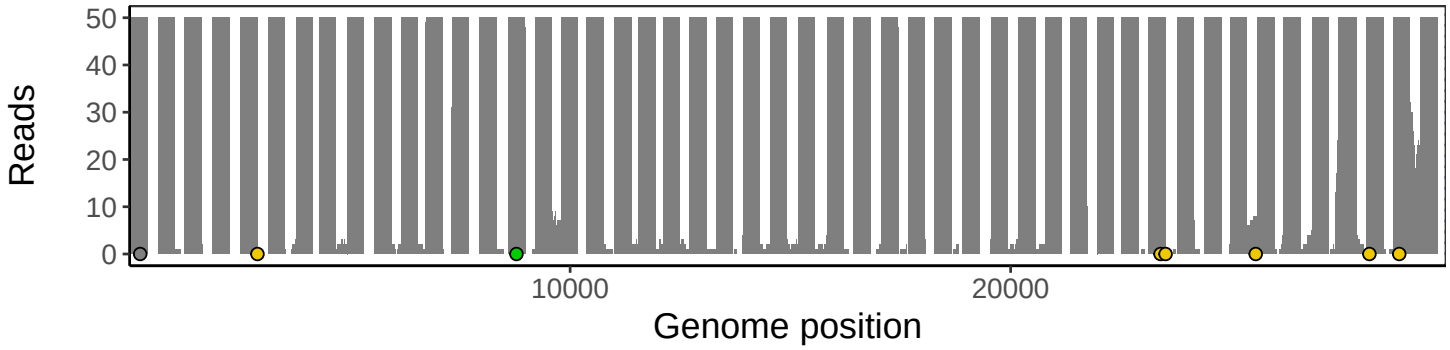
No contig data available.

VSP0042-2 | 4/29/2020 | OP | 239o-tri | 6490 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.