

# COVID-19 subject 209

*2020-08-26*

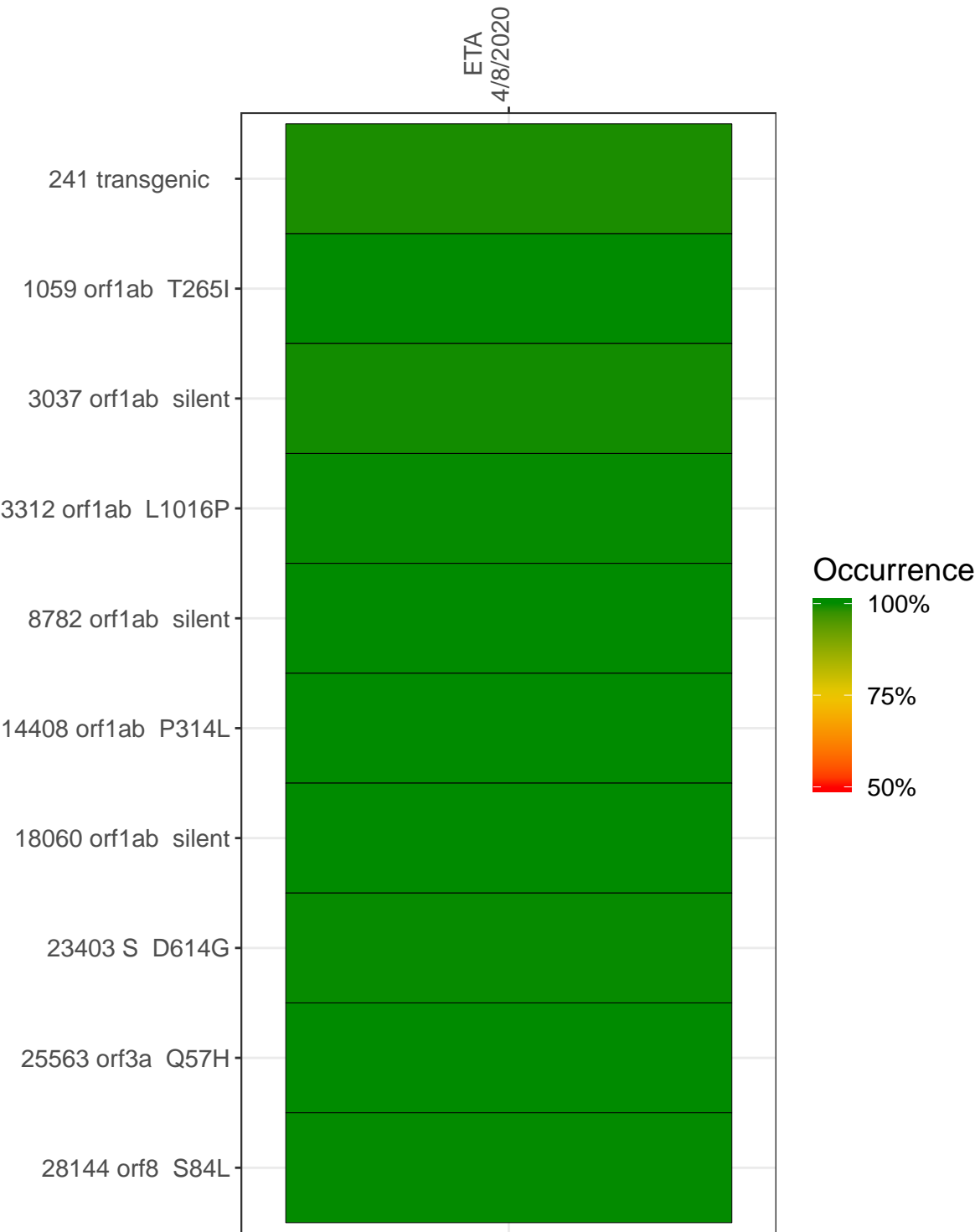
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0011	composite	NA	ETA	4/8/2020	29.85	99.8%	99.4%
VSP0011-1a	single experiment	430000	ETA	4/8/2020	NA	NA	NA
VSP0011-1b	single experiment	430000	ETA	4/8/2020	0.72	51.6%	21.4%
VSP0011-2	single experiment	2150000	ETA	4/8/2020	29.85	99.8%	99.1%

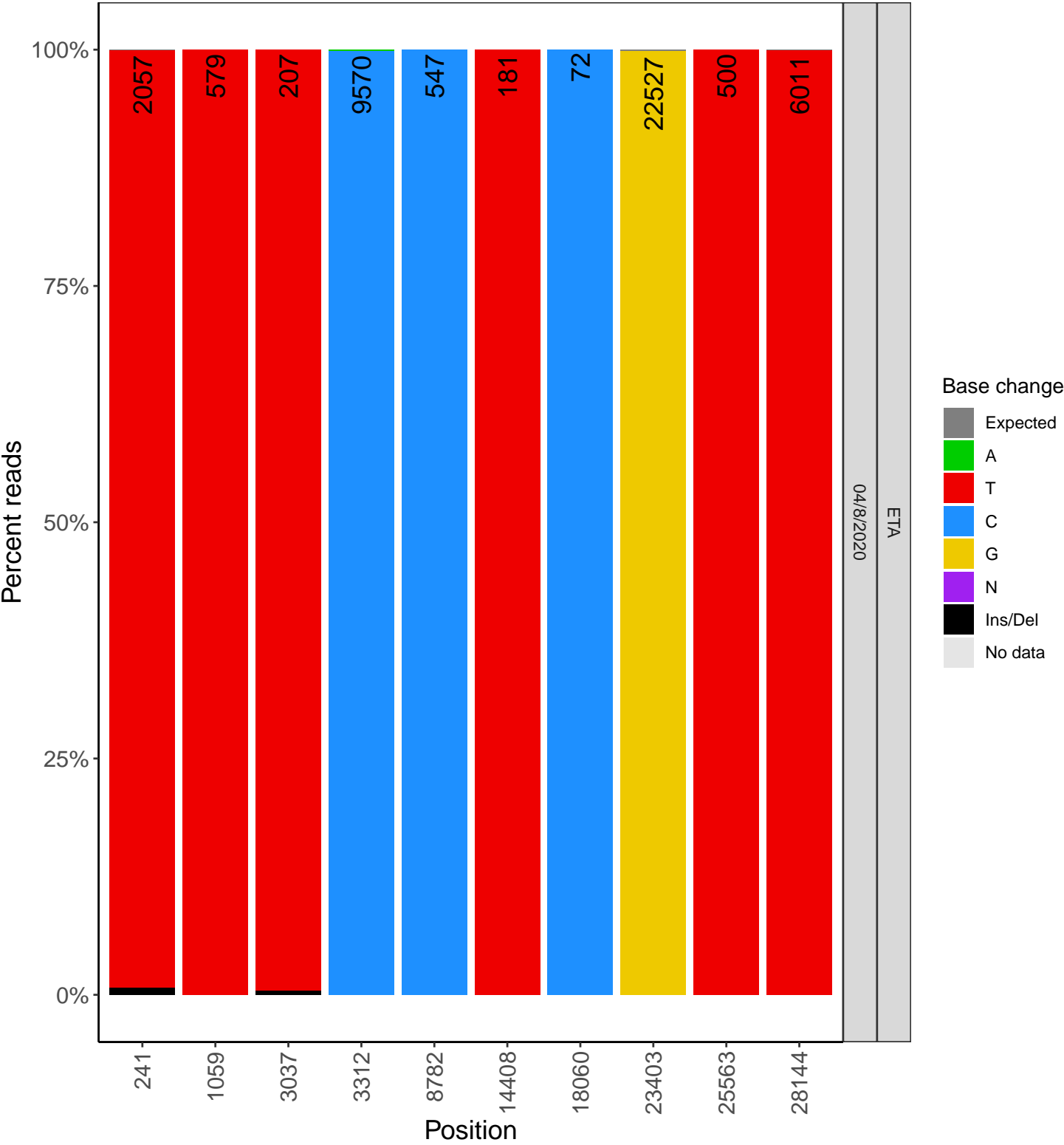
**Variants shared across samples**

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 base composition of tiles are shown in the following plot.



Relative read abundances of variants

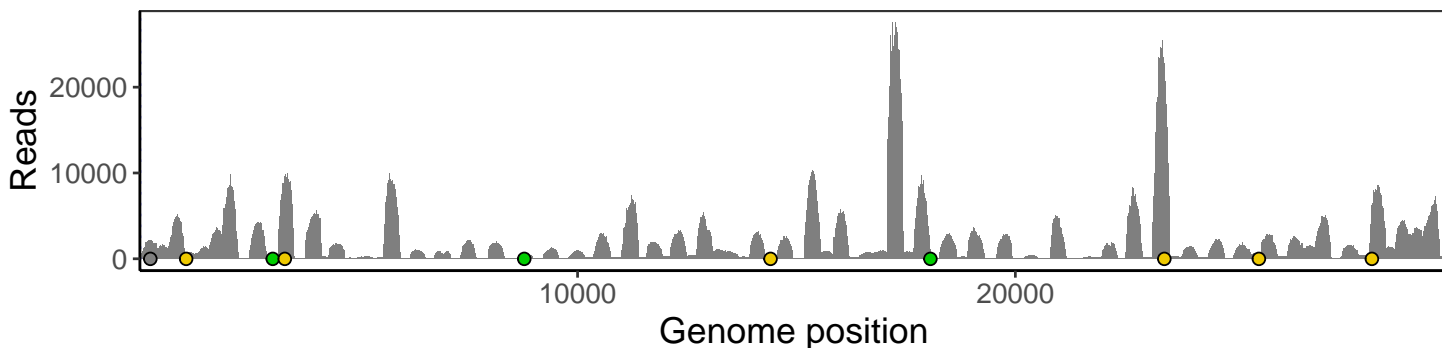
The plot below shows the relative abundances of bases read for each position in the previous variant heatmap where the total number of read pairs covering each position is printed on the stacked bar plots.



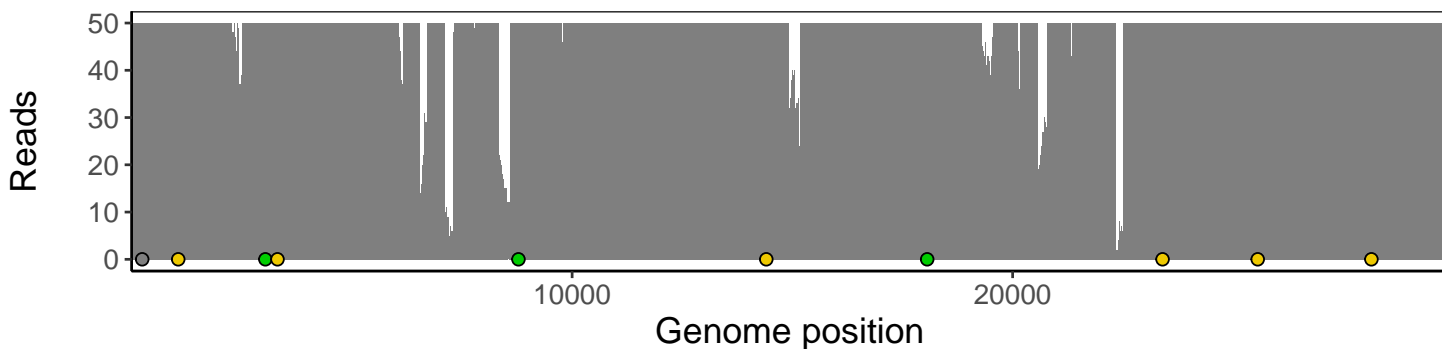
## Analyses of individual experiments and composite results.

VSP0011 | 4/8/2020 | ETA | 209-tri | composite result

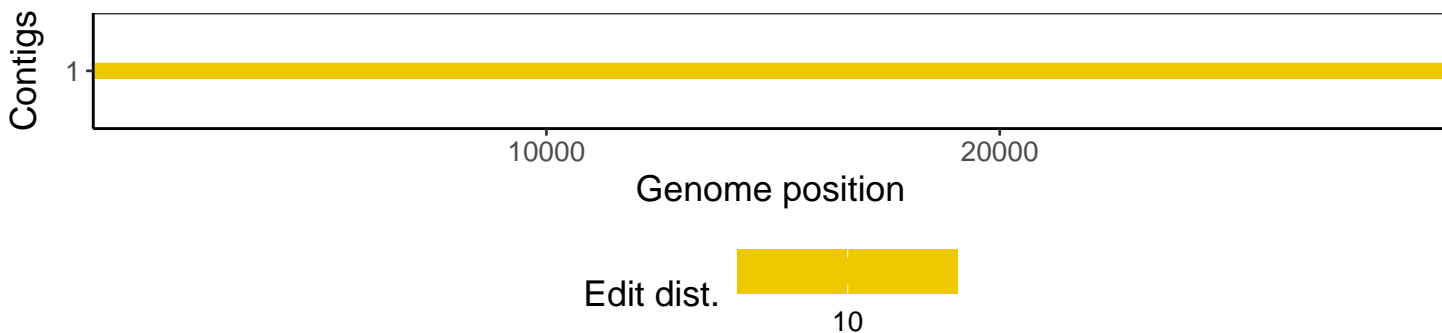
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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.

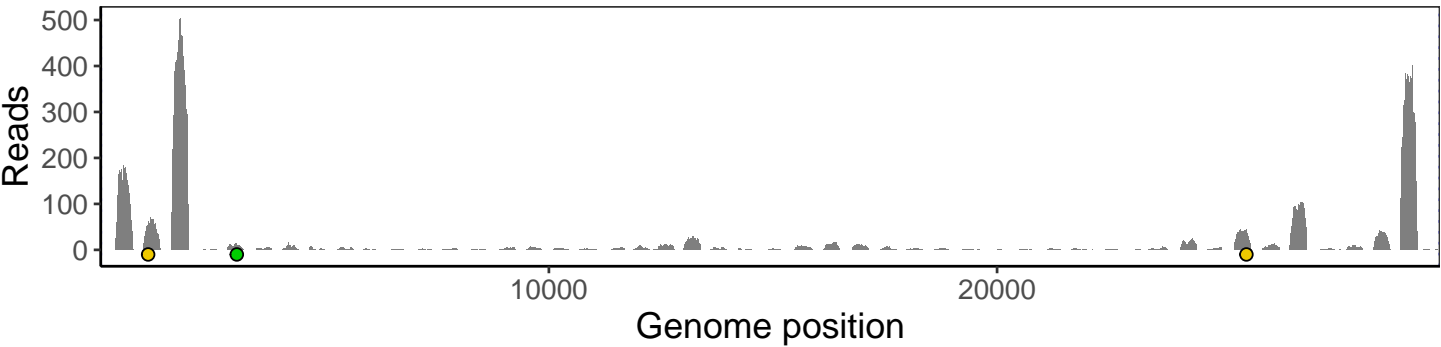


VSP0011-1a | 4/8/2020 | ETA | 209-tri | 430000 genomes | single experiment

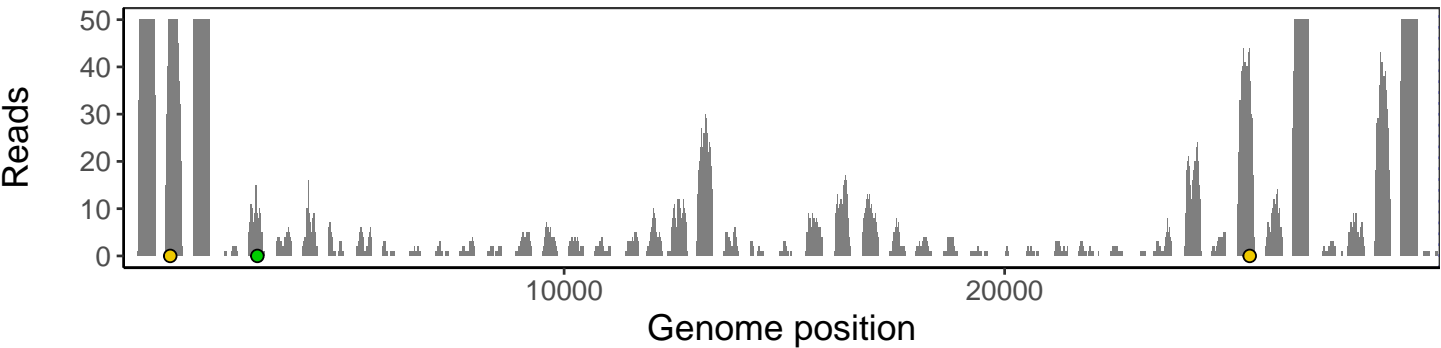
No pileup data available.

No contig data available.

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.



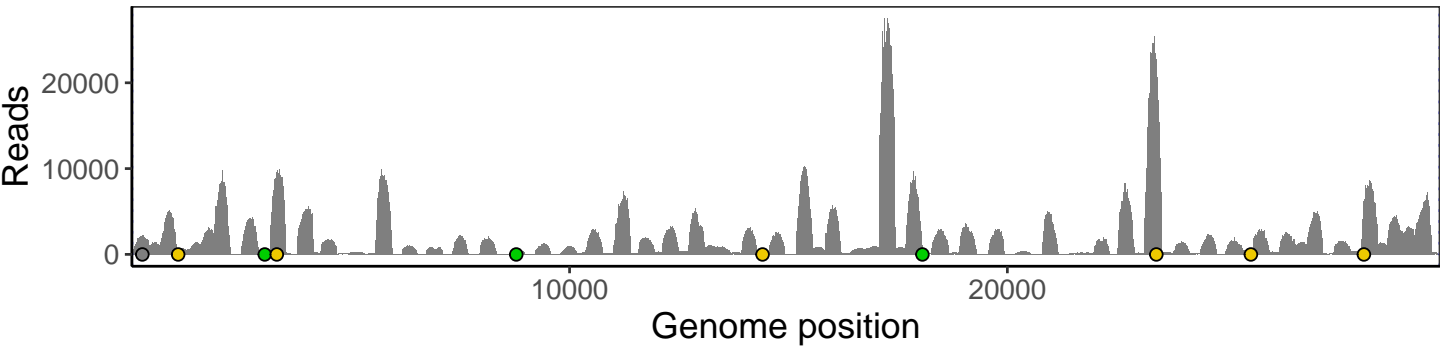
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



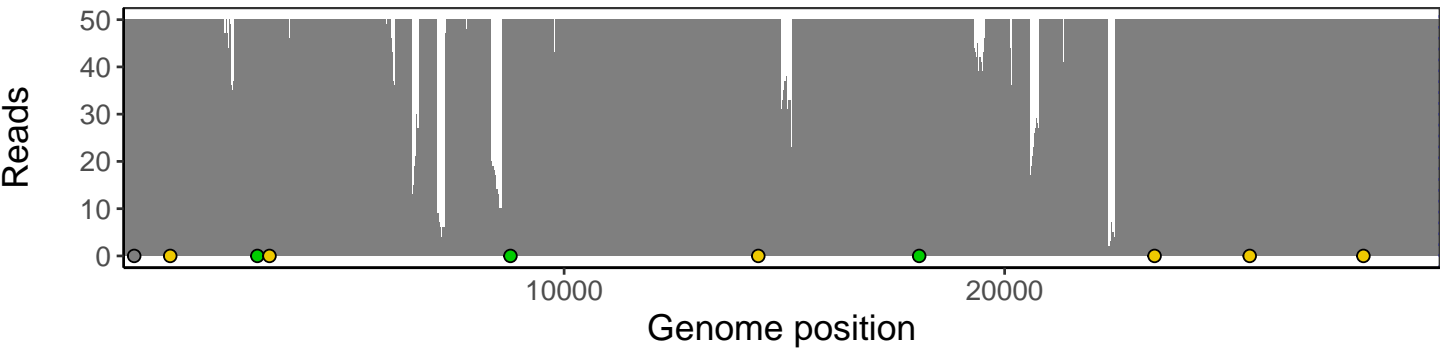
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

