

COVID-19 subject 203

2020-09-14

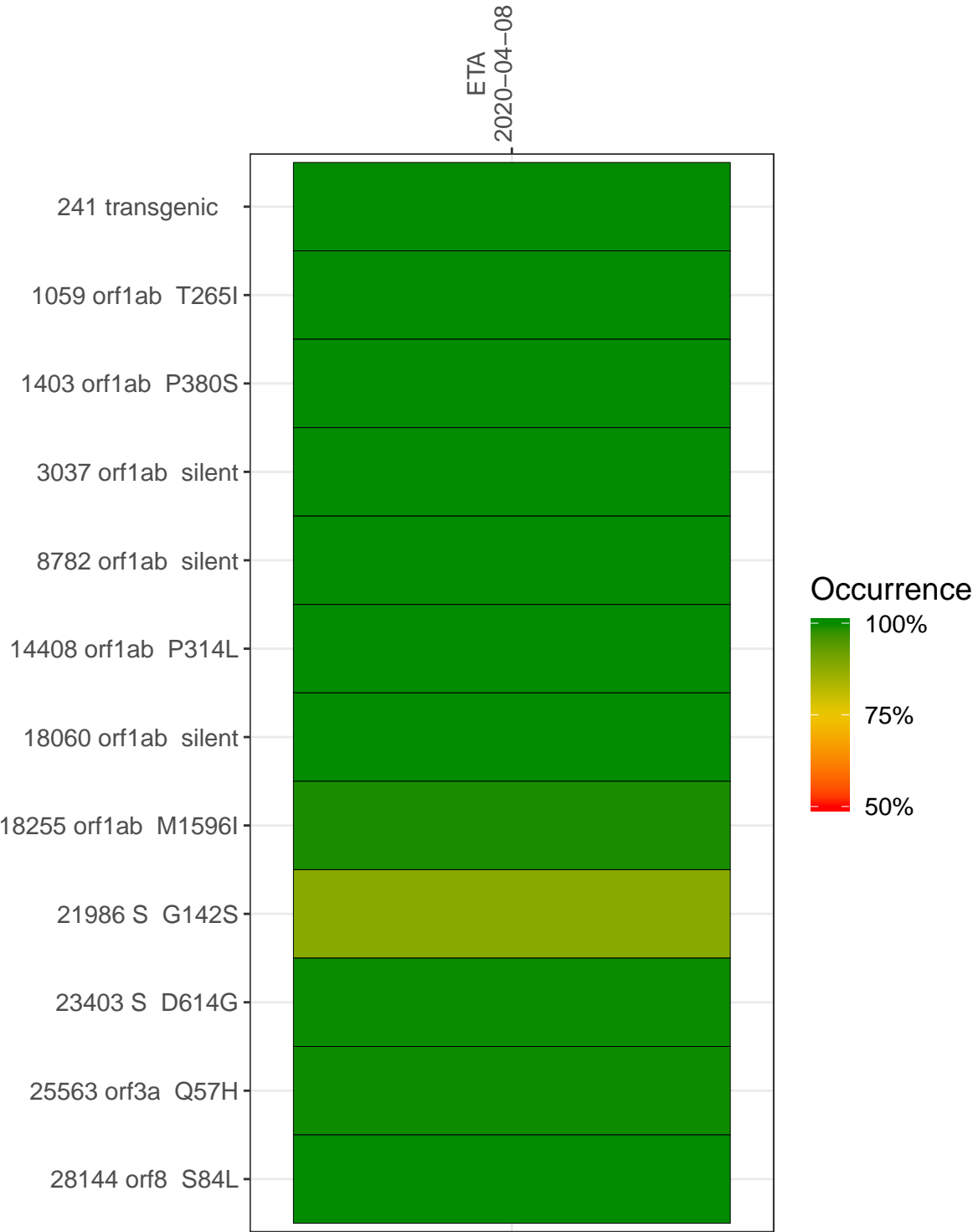
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0010	composite	NA	ETA	2020-04-08	29.84	99.8%	99.4%
VSP0010-1m	single experiment	NA	ETA	2020-04-08	4.43	93.1%	79.2%
VSP0010-2	single experiment	14750	ETA	2020-04-08	0.60	39.1%	4.8%
VSP0010-3	single experiment	2950	ETA	2020-04-08	9.76	96.9%	91.8%
VSP0010-4	single experiment	2950	ETA	2020-04-08	10.36	97.4%	93.6%
VSP0010-5	single experiment	2950	ETA	2020-04-08	17.66	98.5%	91.3%

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
2020-04-08

	VSP0010-1m	VSP0010-2	VSP0010-3	VSP0010-4	VSP0010-5
241 transgenic	48		45	50	21
1059 orf1ab T265I	84		17	31	10
1403 orf1ab P380S	5		19	17	6
3037 orf1ab silent	68		33	28	26
8782 orf1ab silent	3		55	23	20
14408 orf1ab P314L	118		31	47	22
18060 orf1ab silent	54		26	15	23
18255 orf1ab M1596I	47		28	25	34
21986 S G142S	48		8	4	1
23403 S D614G	51	7	120	132	128
25563 orf3a Q57H	130	1	47	64	64
28144 orf8 S84L	58		37	54	32

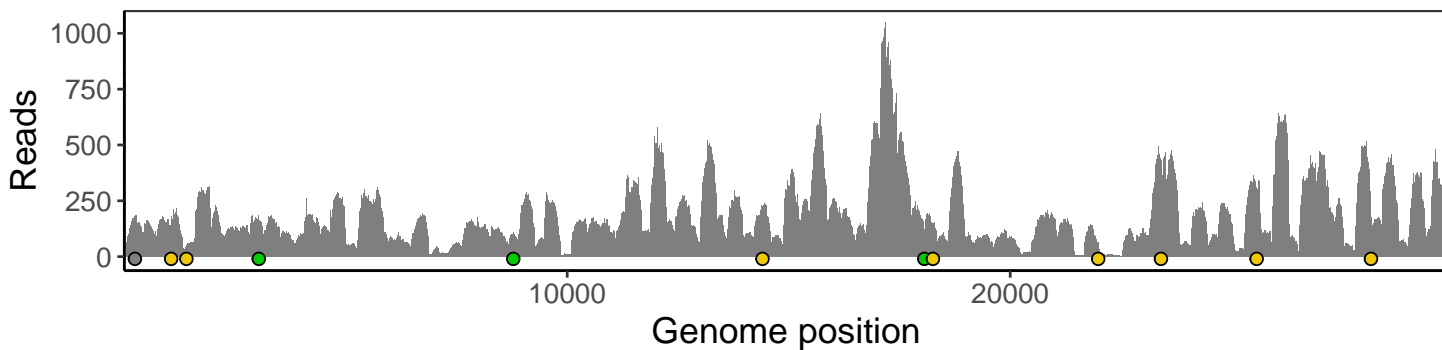
Base change



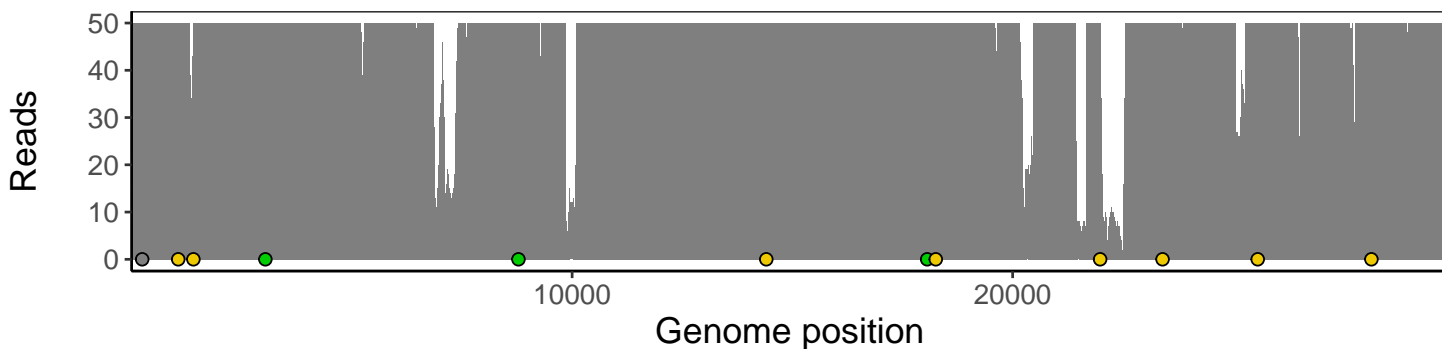
Analyses of individual experiments and composite results.

VSP0010 | 2020-04-08 | ETA | 203-tri | composite result

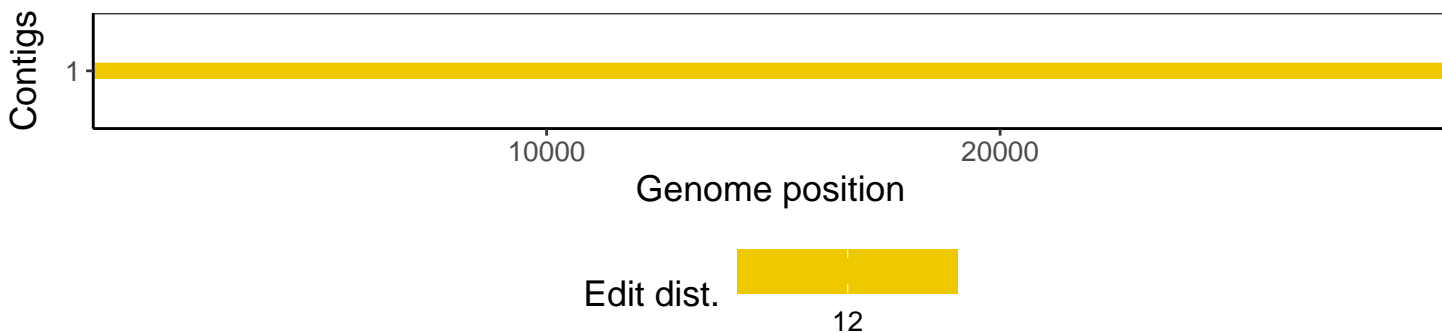
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



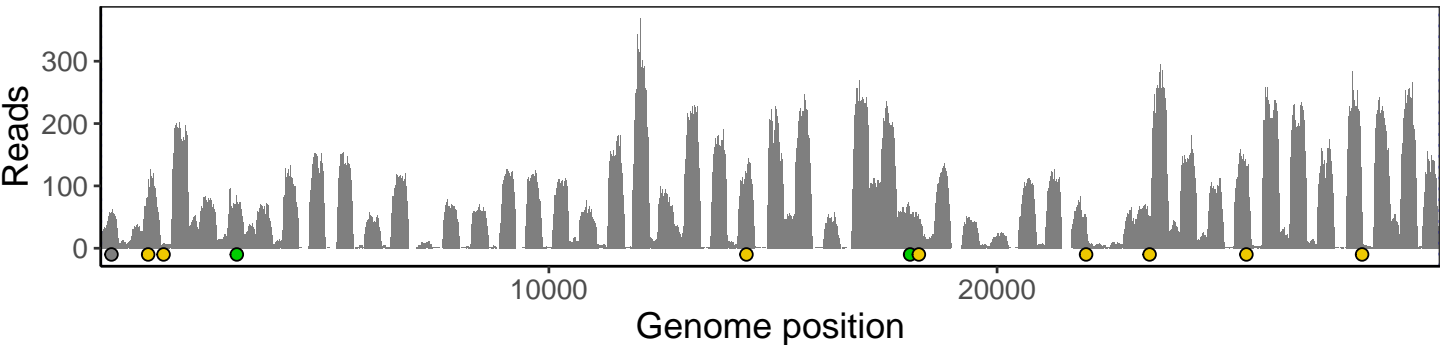
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



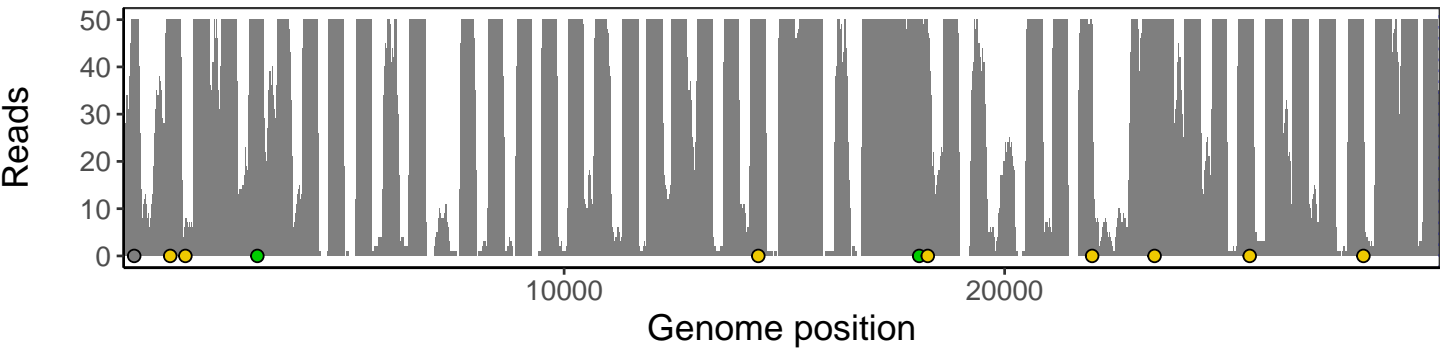
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



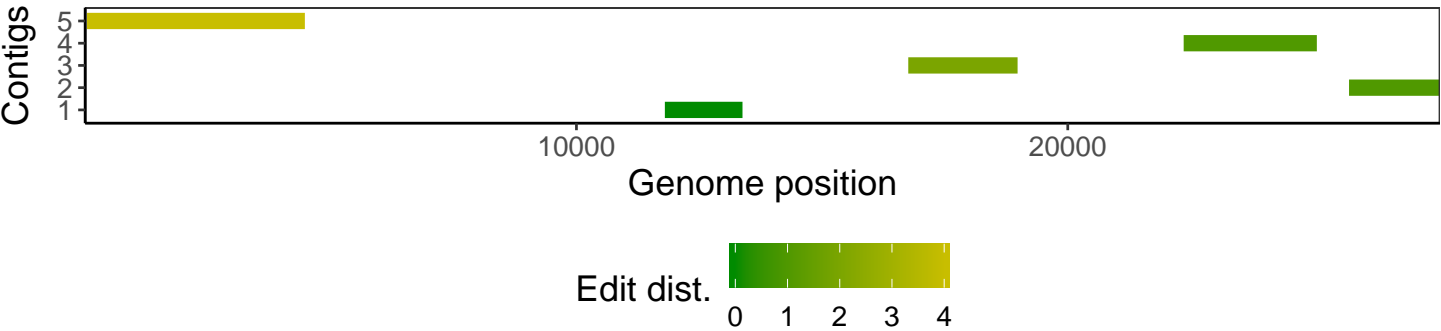
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



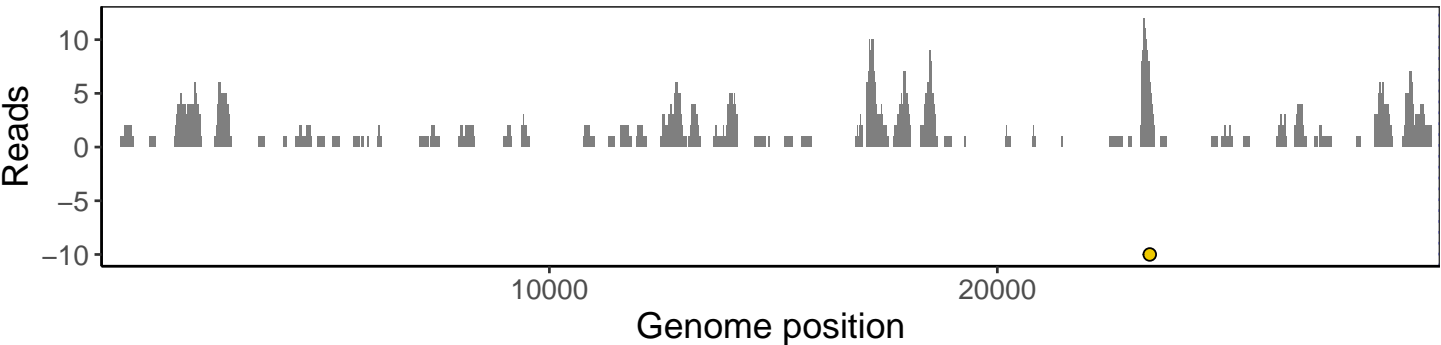
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



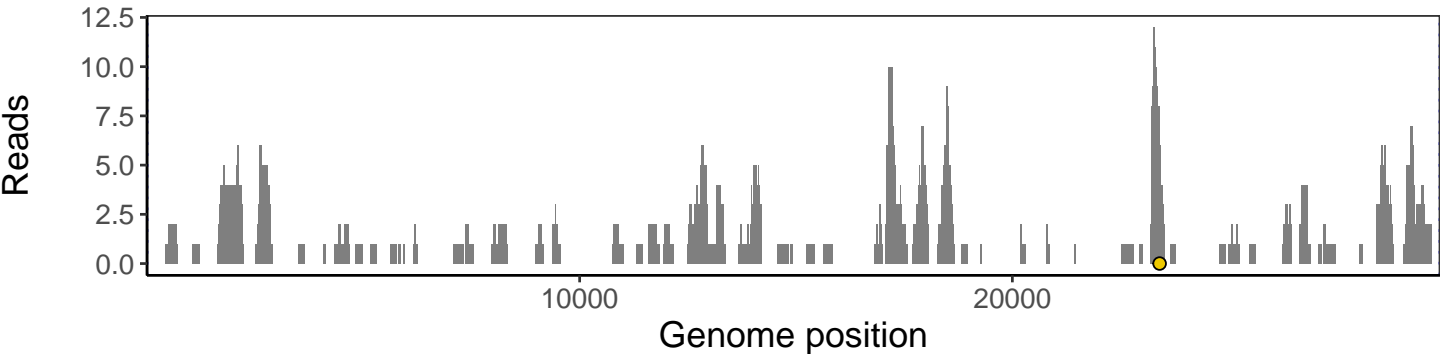
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



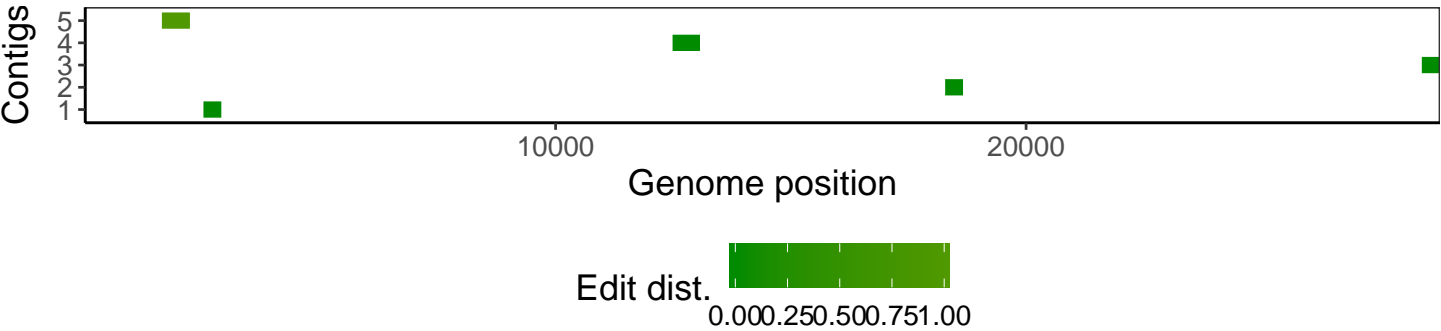
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



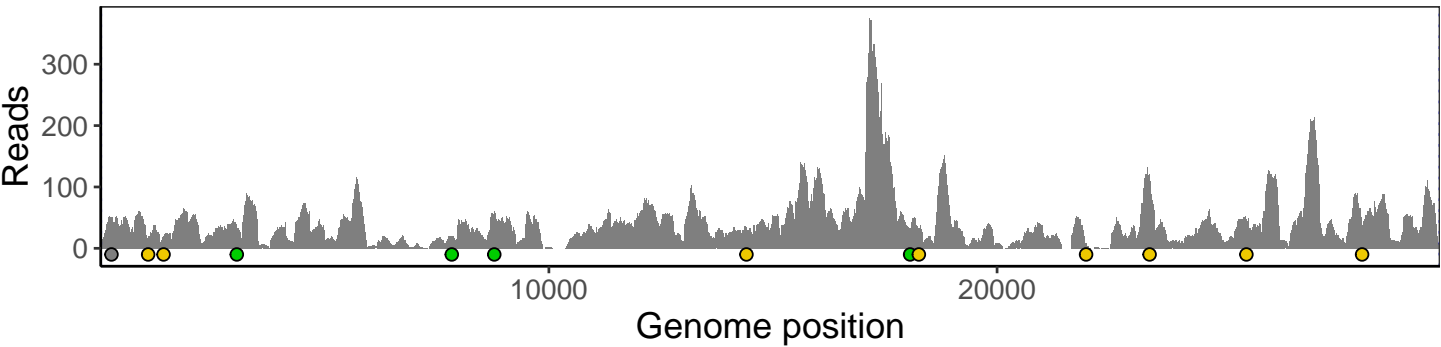
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



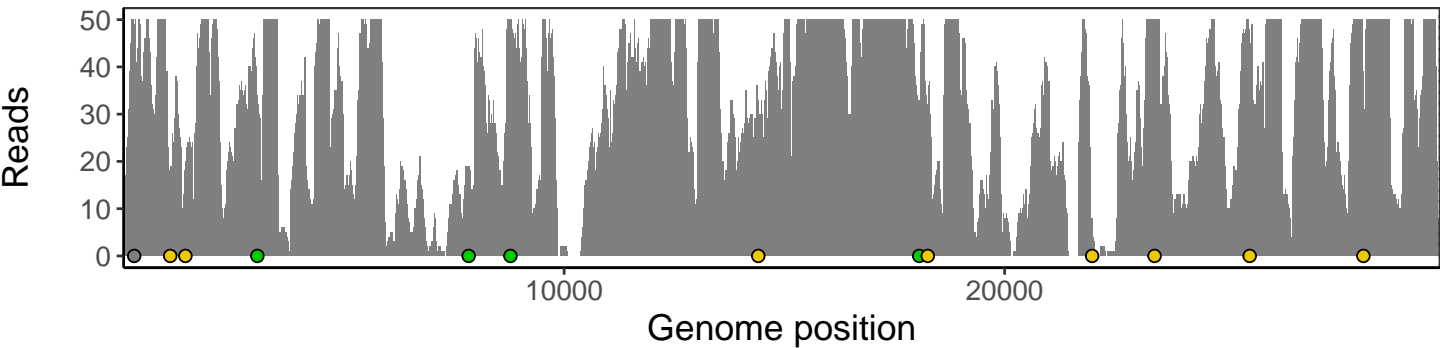
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



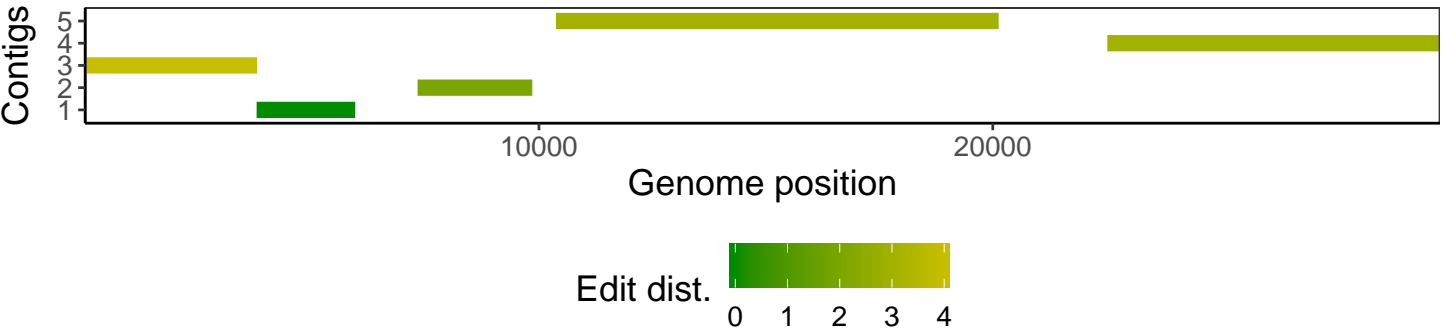
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



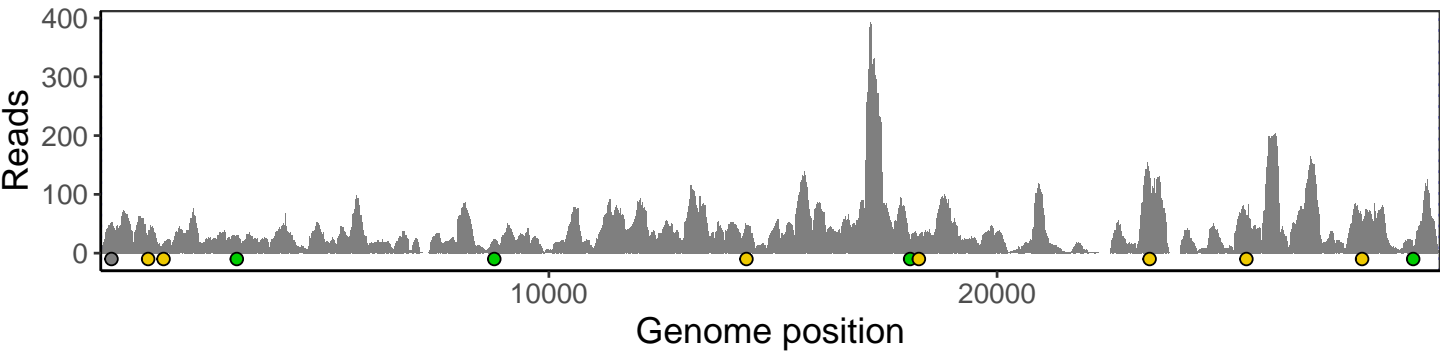
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



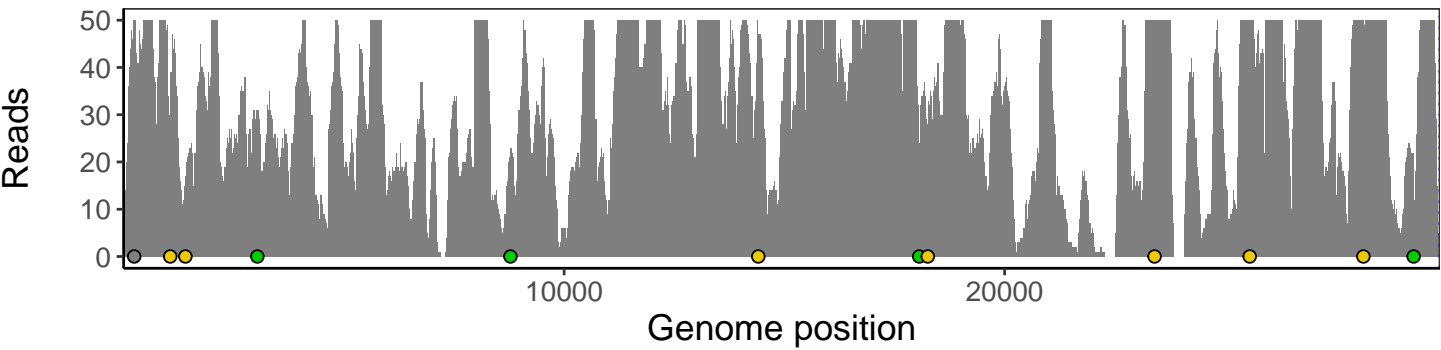
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



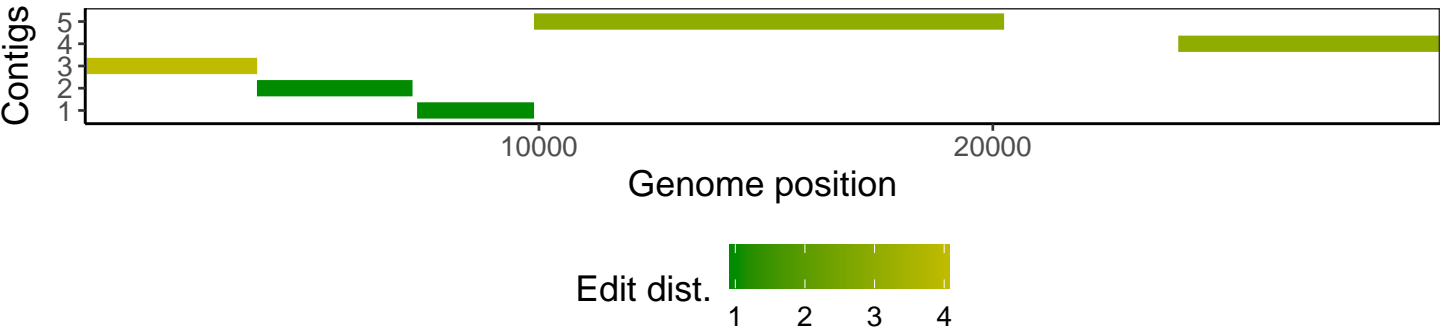
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



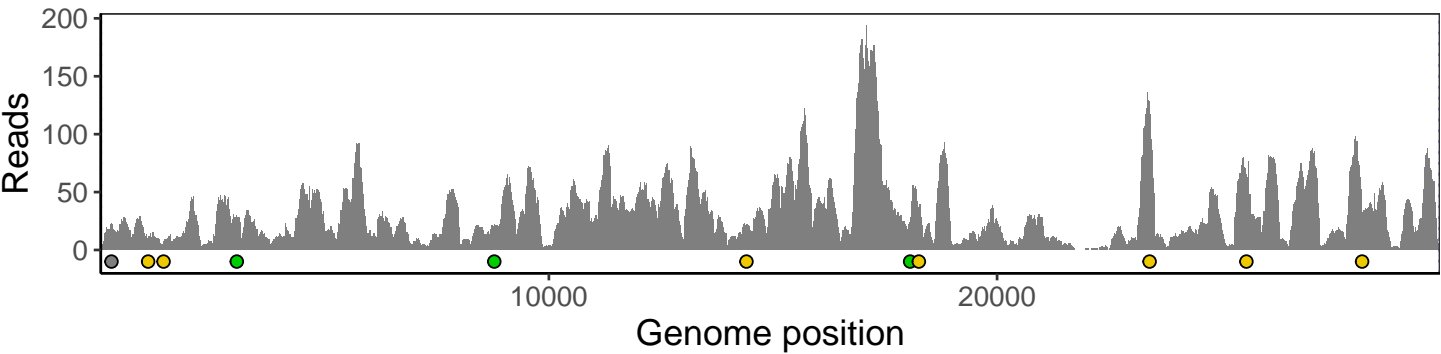
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



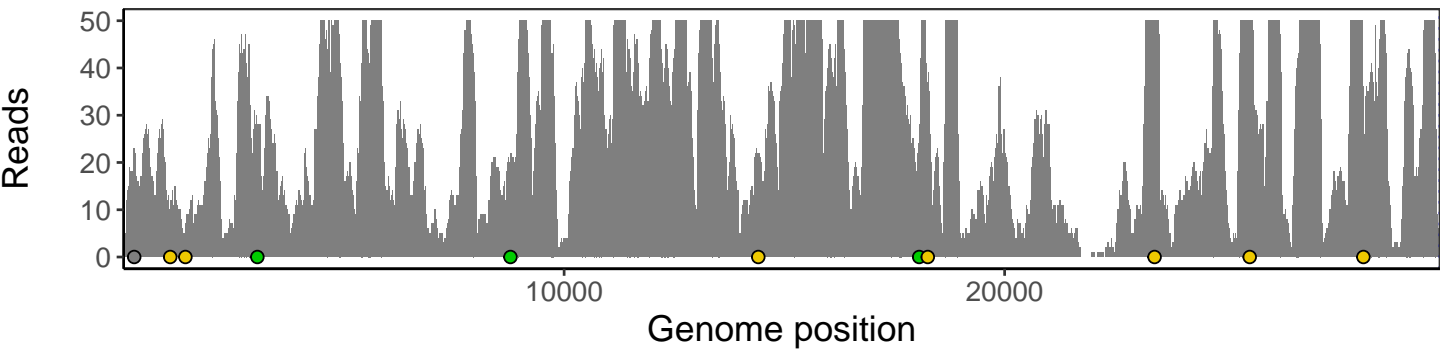
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

