

COVID-19 subject 235

2020-09-14

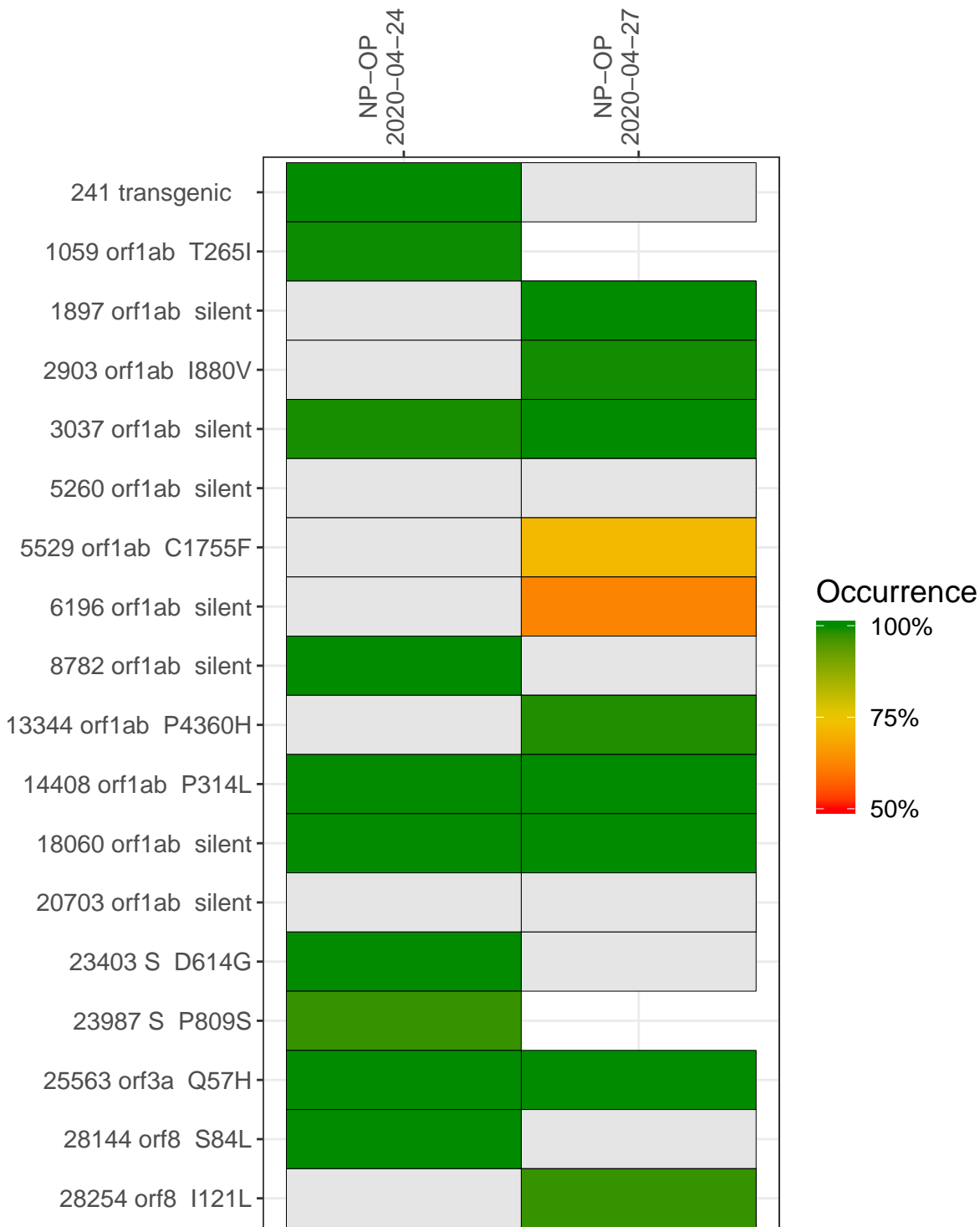
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0030	composite	NA	NP-OP	2020-04-27	2.47	76.0%	54.2%
VSP0029-1m	single experiment	NA	NP-OP	2020-04-27	0.50	16.6%	12.1%
VSP0030-1m	single experiment	NA	NP-OP	2020-04-27	1.05	50.3%	20.0%
VSP0030-2	single experiment	NA	NP-OP	2020-04-27	1.96	55.0%	52.1%
VSP0080-1	single experiment	176000	NP-OP	2020-04-24	29.87	99.9%	99.8%

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.



	NP-OP 2020-04-24		NP-OP 2020-04-27	
241 transgenic	3189		1	
1059 orf1ab T265I	992			
1897 orf1ab silent	1994	15	6	1
2903 orf1ab I880V	1831	1274	83	203
3037 orf1ab silent	1831			242
5260 orf1ab silent	1147	2	6	2
5529 orf1ab C1755F	3081			405
6196 orf1ab silent	2418	8	20	709
8782 orf1ab silent	2466		2	
13344 orf1ab P4360H	4536	87	172	348
14408 orf1ab P314L	3451		253	580
18060 orf1ab silent	2195		338	386
20703 orf1ab silent	3645	180	249	595
23403 S D614G	7352		1	
23987 S P809S	819			
25563 orf3a Q57H	1959			479
28144 orf8 S84L	3552		1	
28254 orf8 I121L	4157	112	1	
	VSP00080-1	VSP00029-1m	VSP00030-1m	VSP00030-2

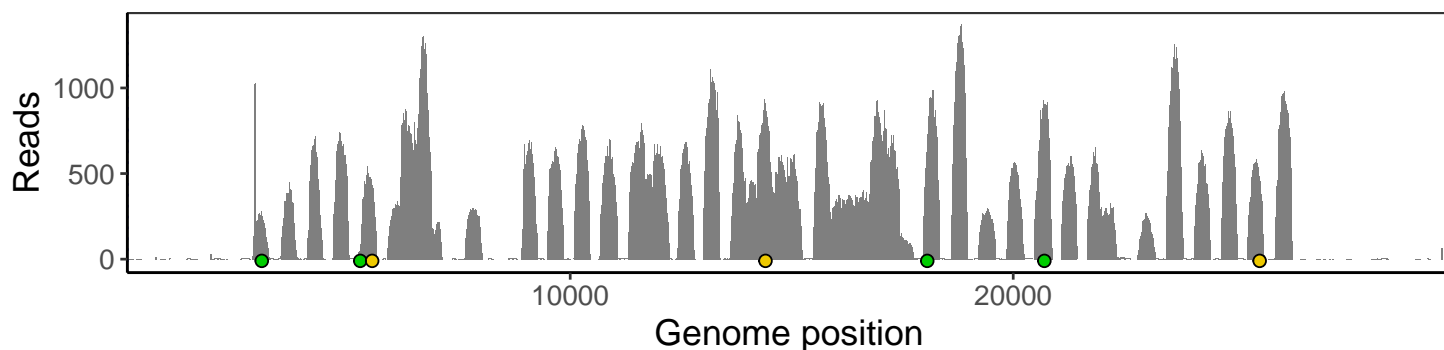
Base change



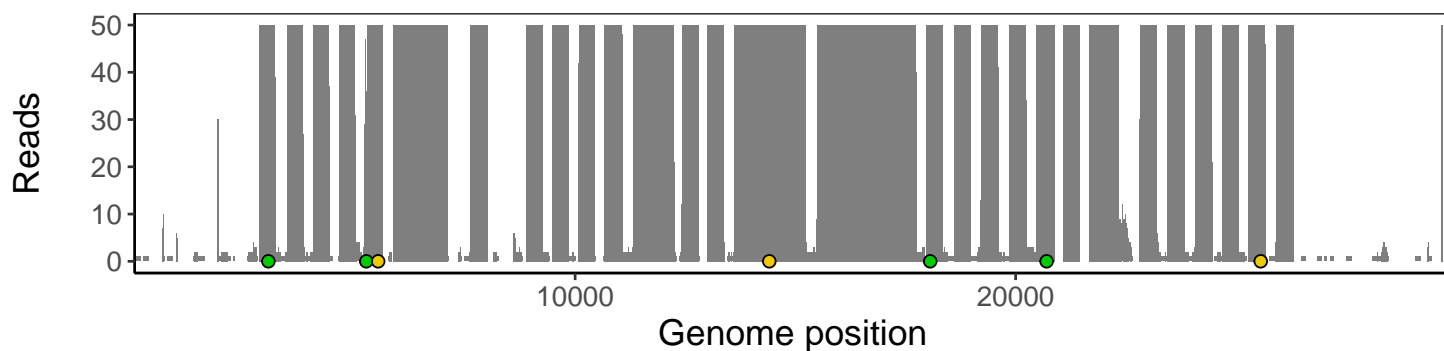
Analyses of individual experiments and composite results.

VSP0030 | 2020-04-27 | NP-OP | 235-qia | composite result

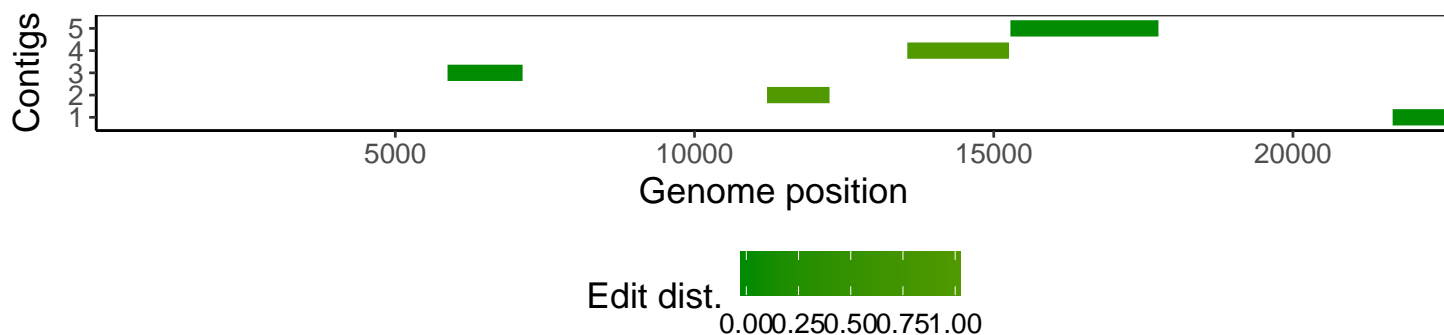
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



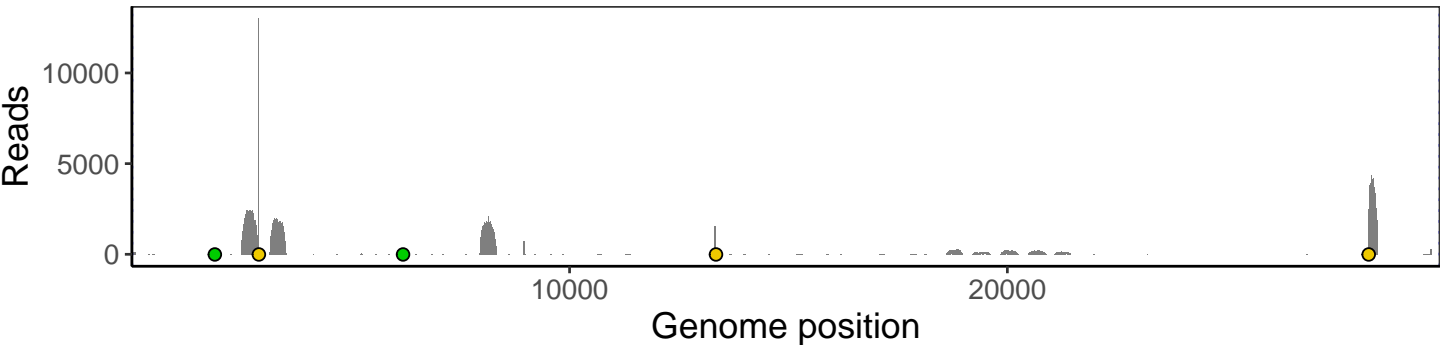
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



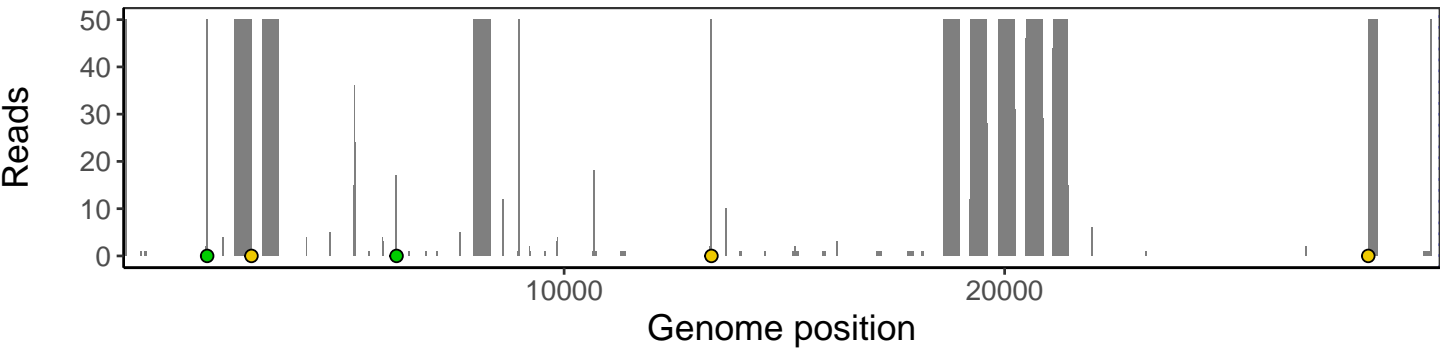
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



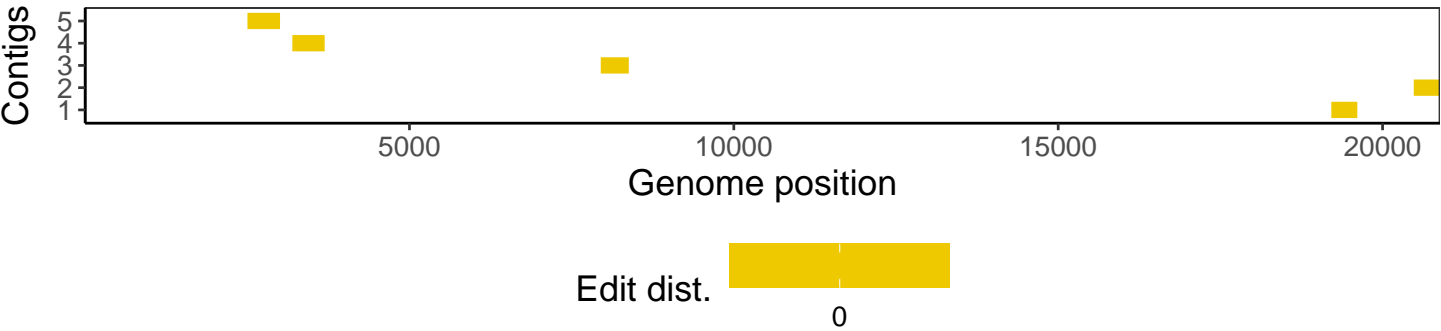
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



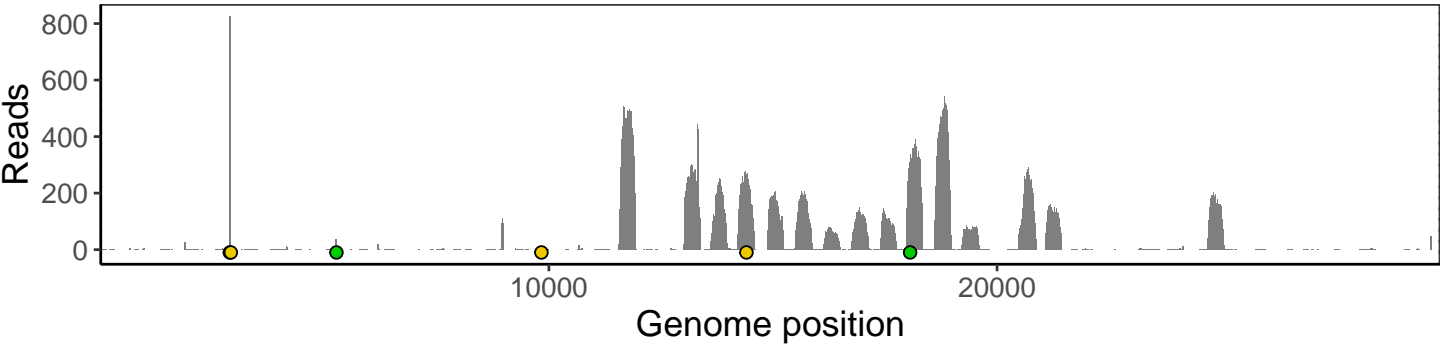
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



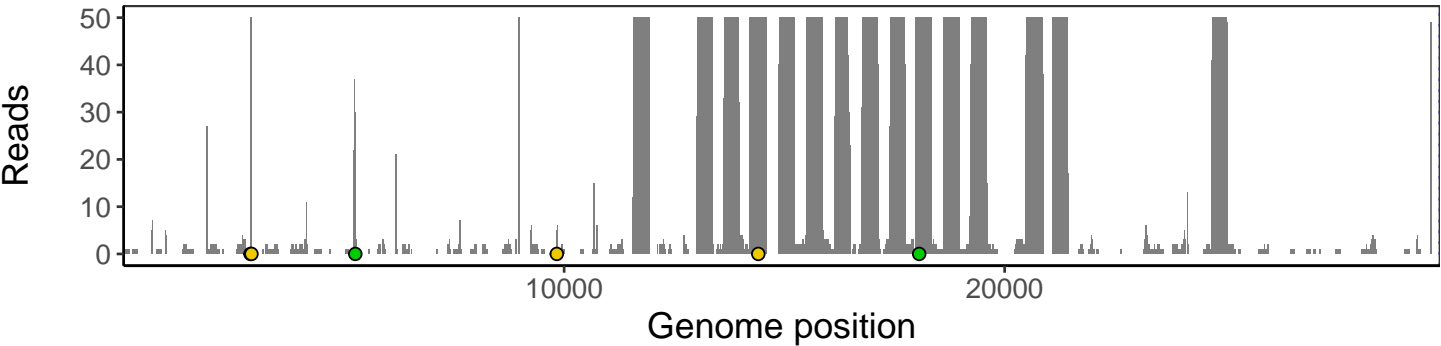
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



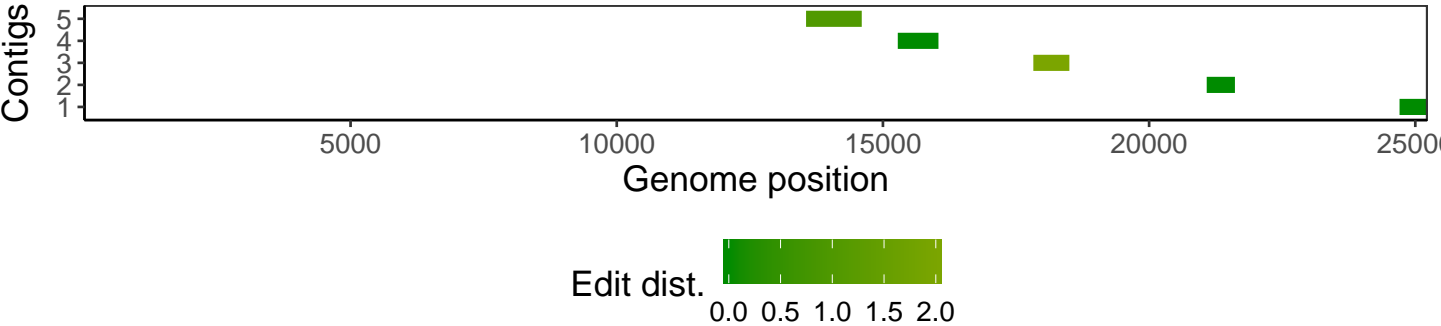
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



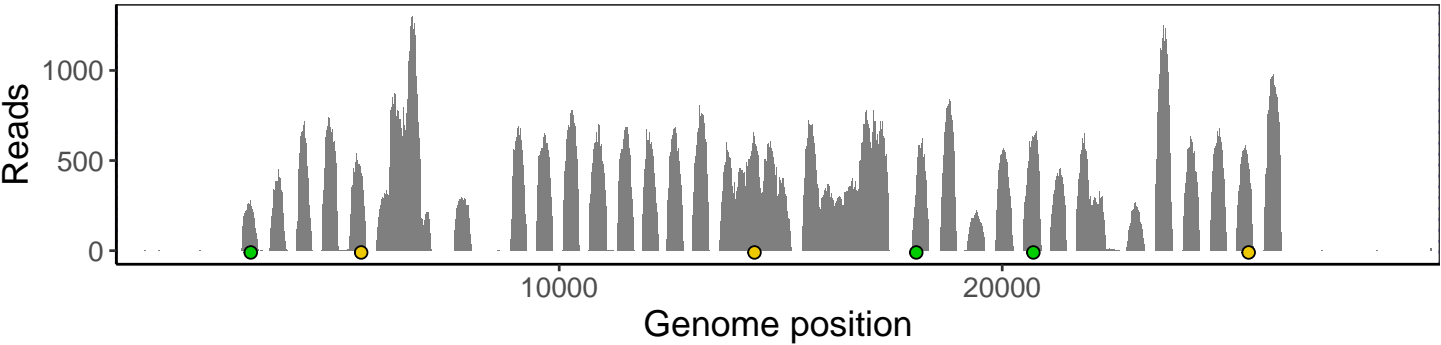
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



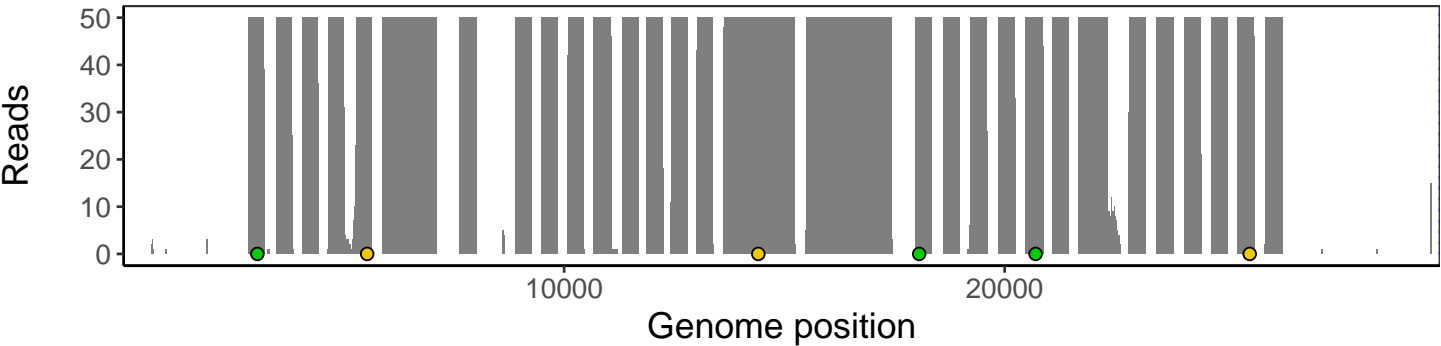
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



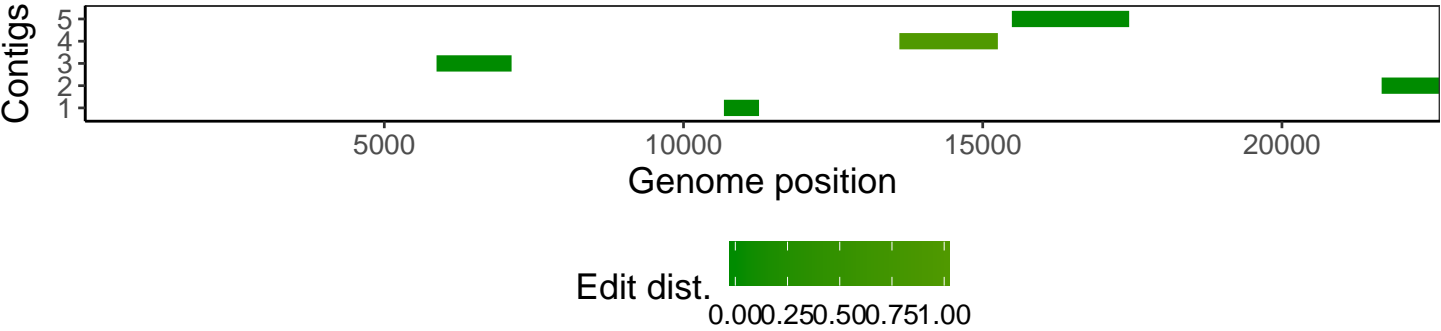
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



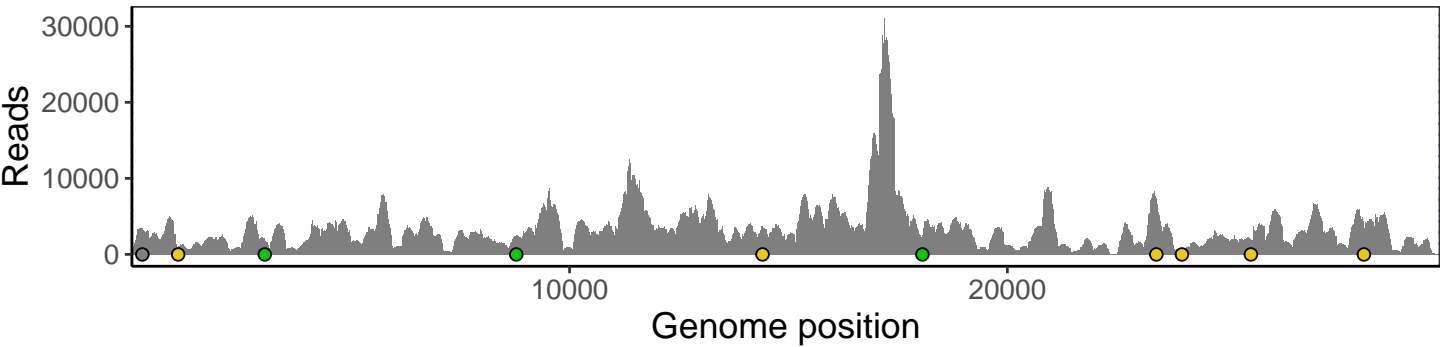
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



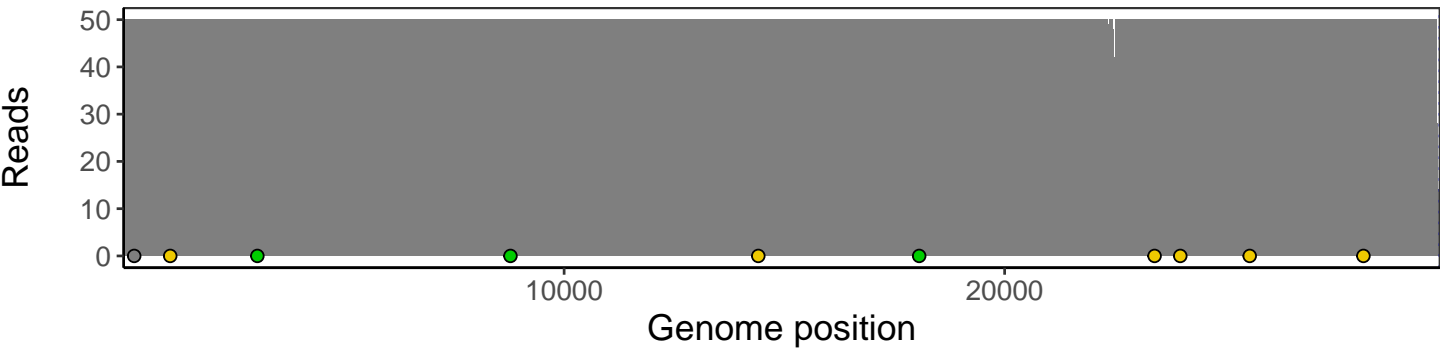
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

