

COVID-19 subject 239

2021-01-10

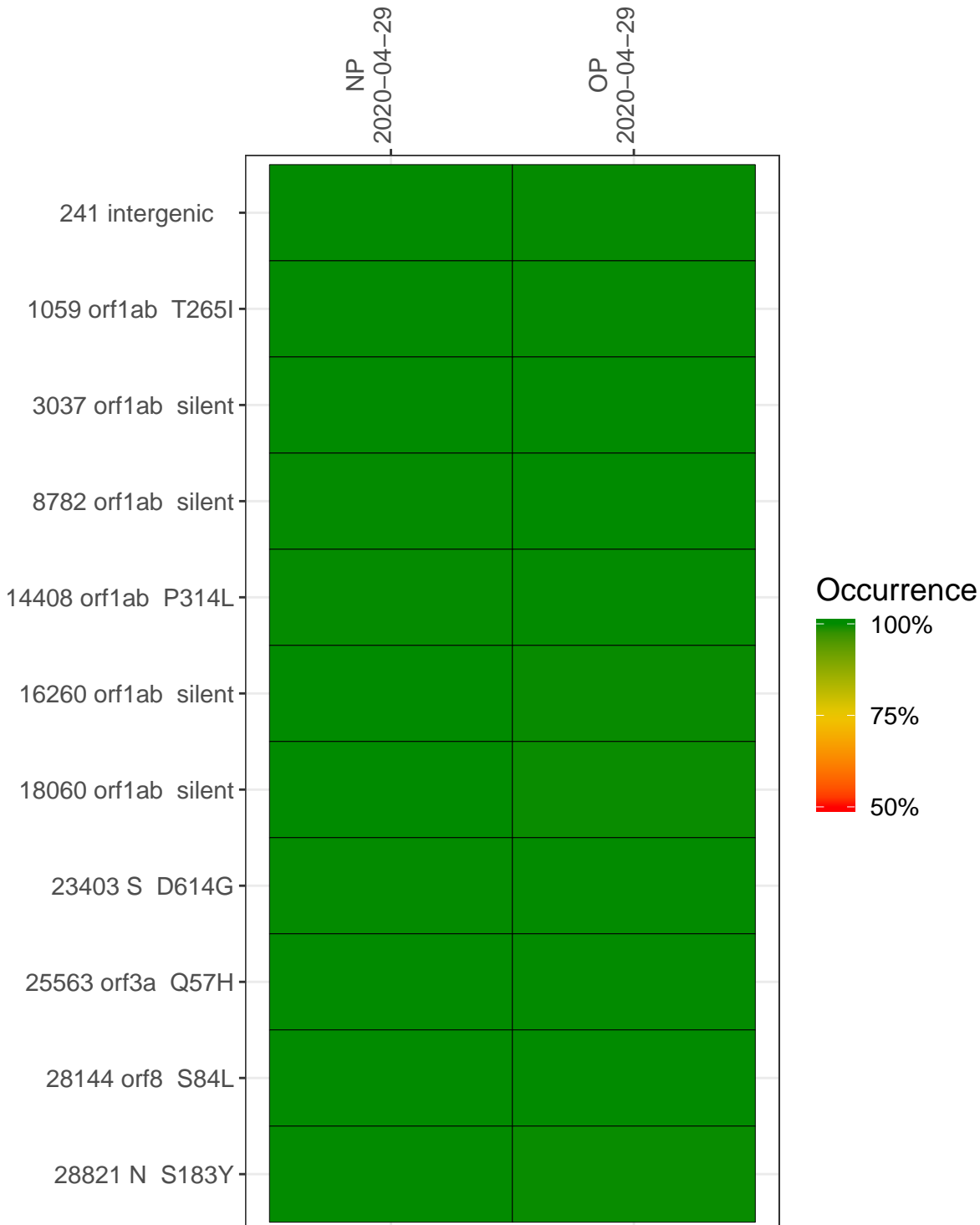
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

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	2020-04-29	29.84	99.9%	99.8%
VSP0042	composite	NA	OP	2020-04-29	27.48	99.8%	99.8%
VSP0041-1m	single experiment	NA	NP	2020-04-29	29.83	99.8%	99.8%
VSP0041-2	single experiment	NA	NP	2020-04-29	29.84	99.8%	99.5%
VSP0042-1m	single experiment	NA	OP	2020-04-29	1.12	71.6%	63.3%
VSP0042-2	single experiment	6490	OP	2020-04-29	1.18	81.3%	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 or 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 2020-04-29		OP 2020-04-29	
241 intergenic	3042	6800		3423
1059 orf1ab T265I	967	6826	1556	1
3037 orf1ab silent	788	520	1113	
8782 orf1ab silent	2065	300		2289
14408 orf1ab P314L	1655	539	1775	1
16260 orf1ab silent	663	507	1122	1
18060 orf1ab silent	963	305	1354	
23403 S D614G	4773	11657		4596
25563 orf3a Q57H	1669	13532	2486	8
28144 orf8 S84L	4570	1525	562	4944
28821 N S183Y	1453	2621	1	2168
	VSP0041-1m	VSP0041-2	VSP0042-1m	VSP0042-2

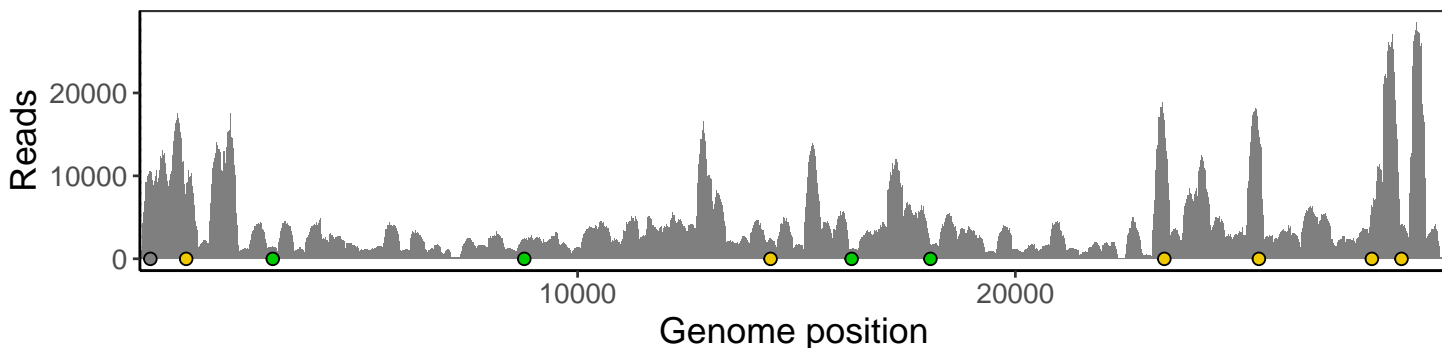
Base change



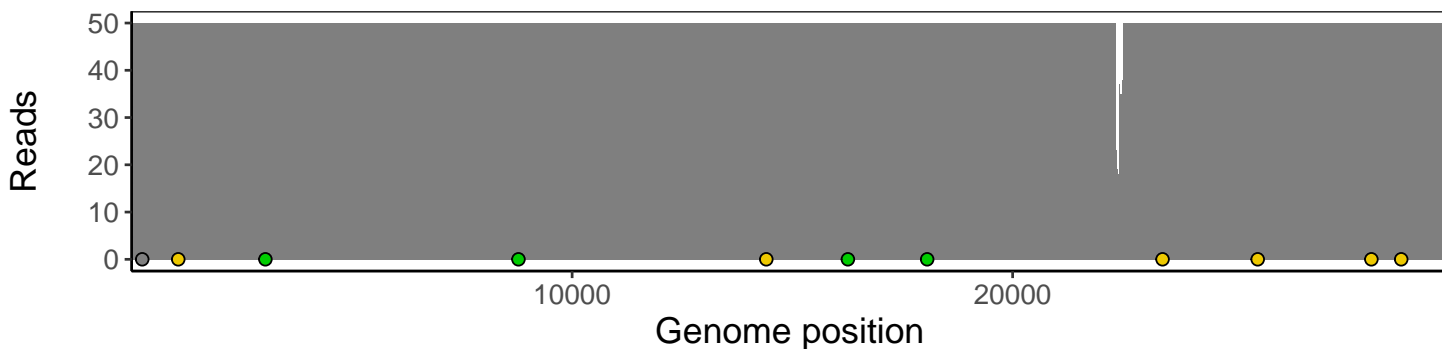
Analyses of individual experiments and composite results.

VSP0041 | 2020-04-29 | NP | 239n-tri | composite result

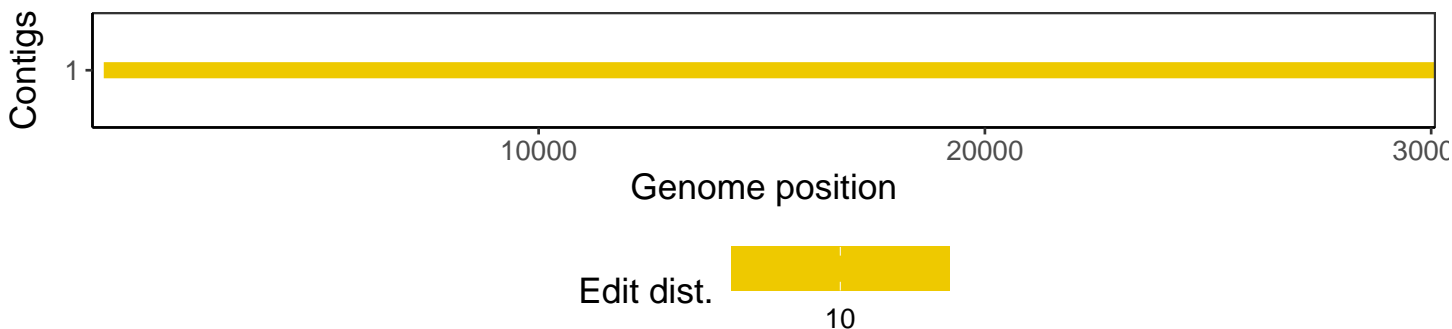
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



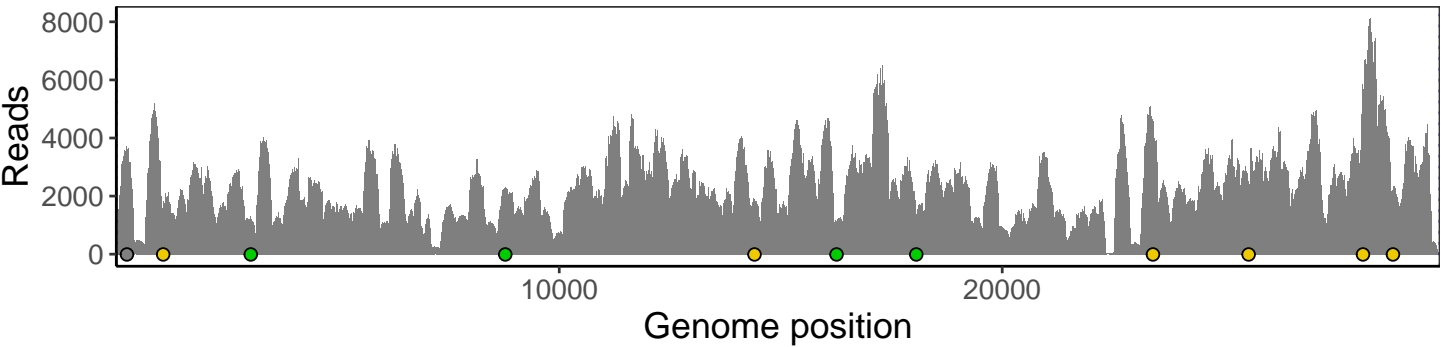
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



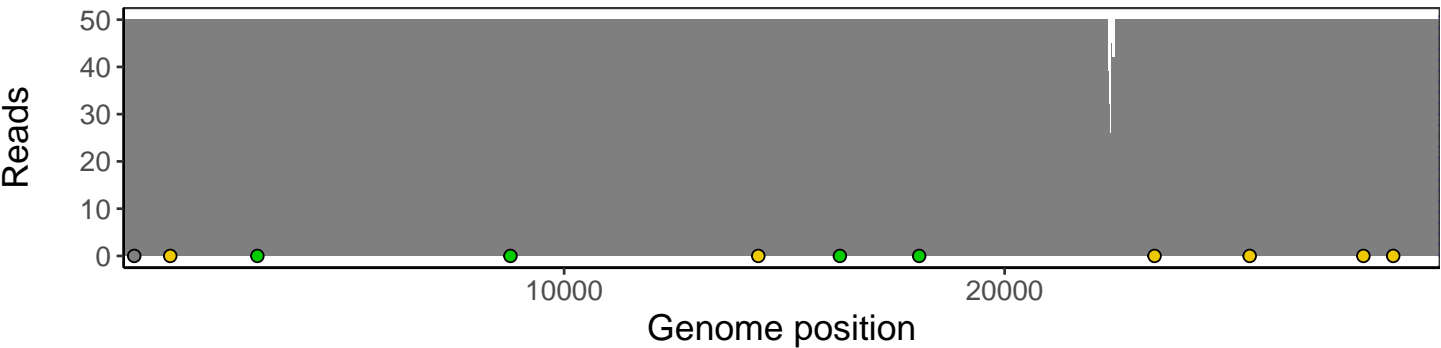
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



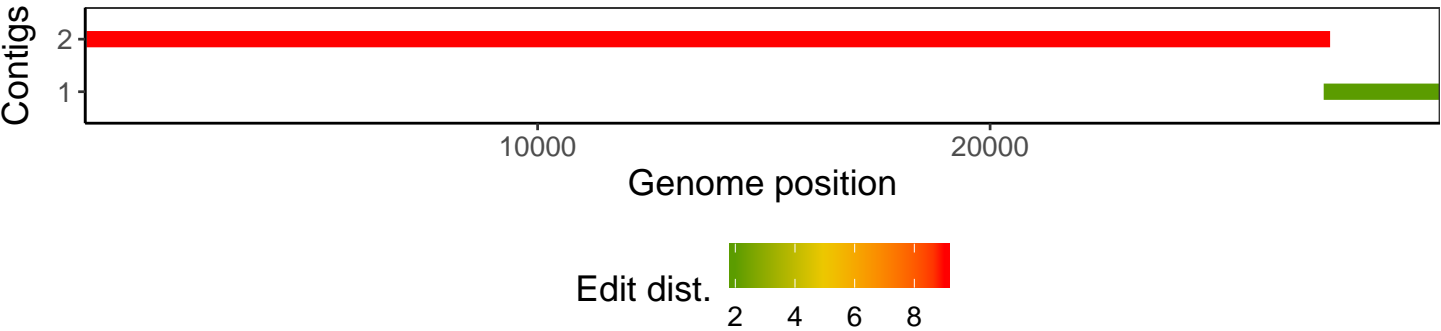
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



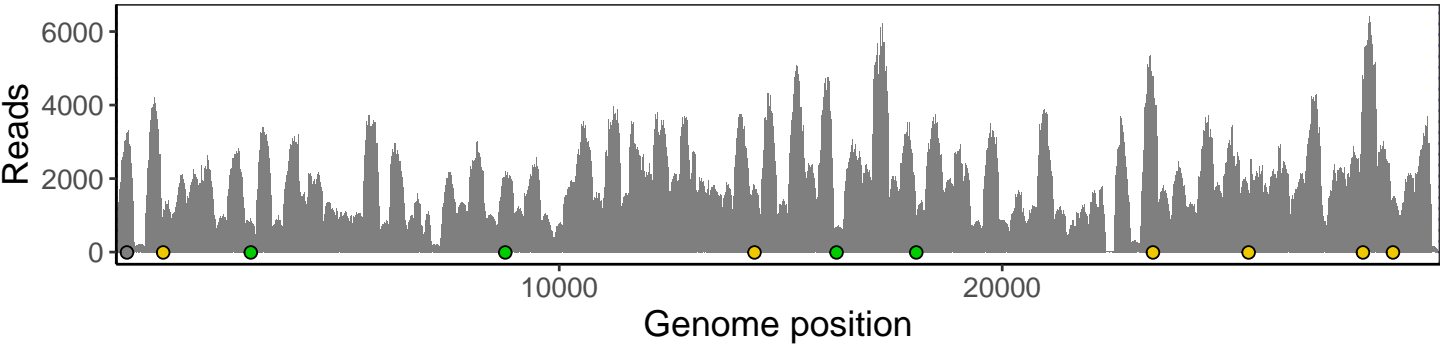
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



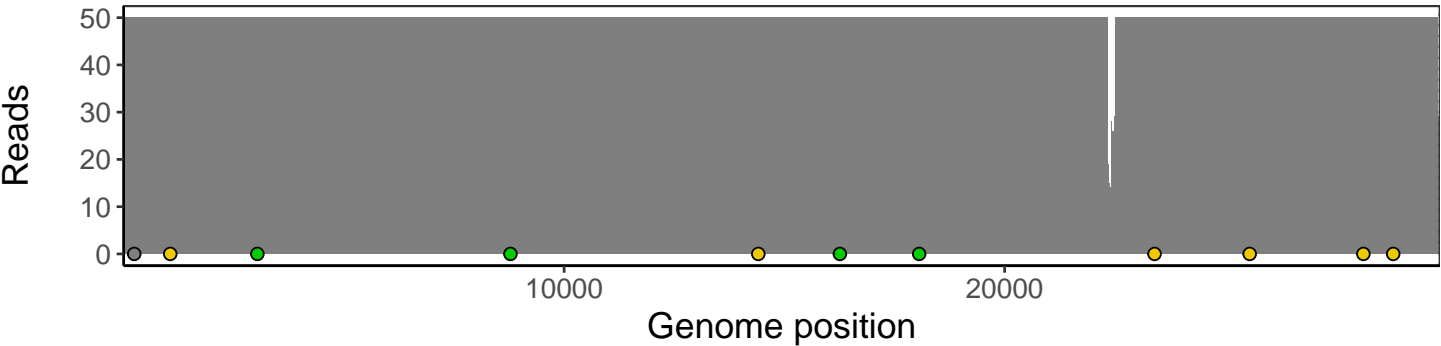
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



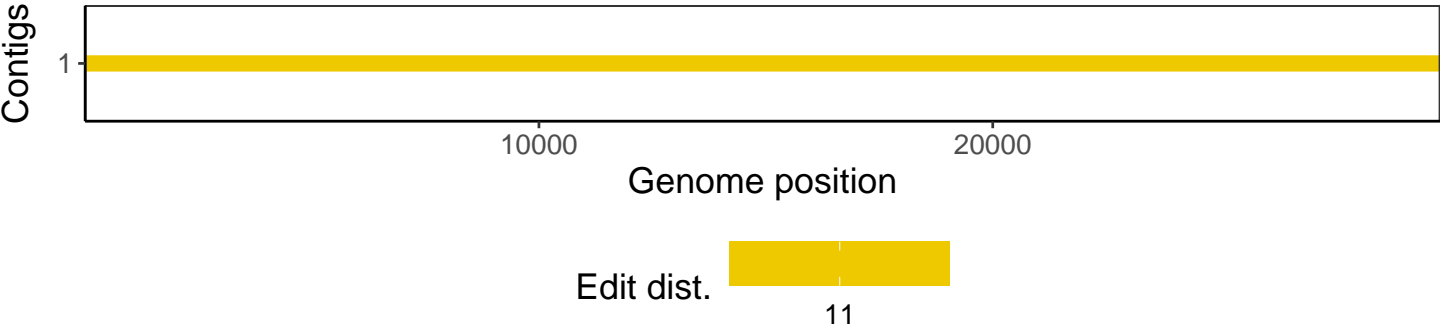
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



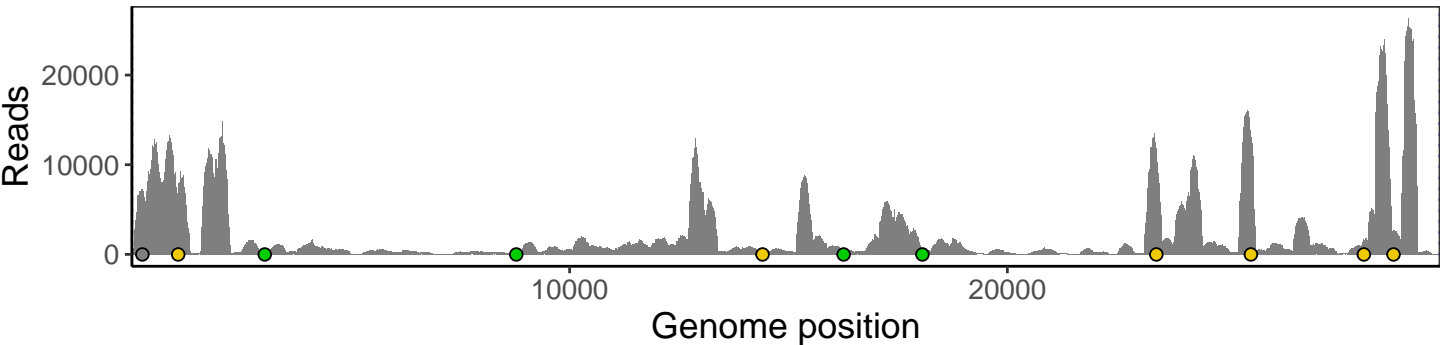
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



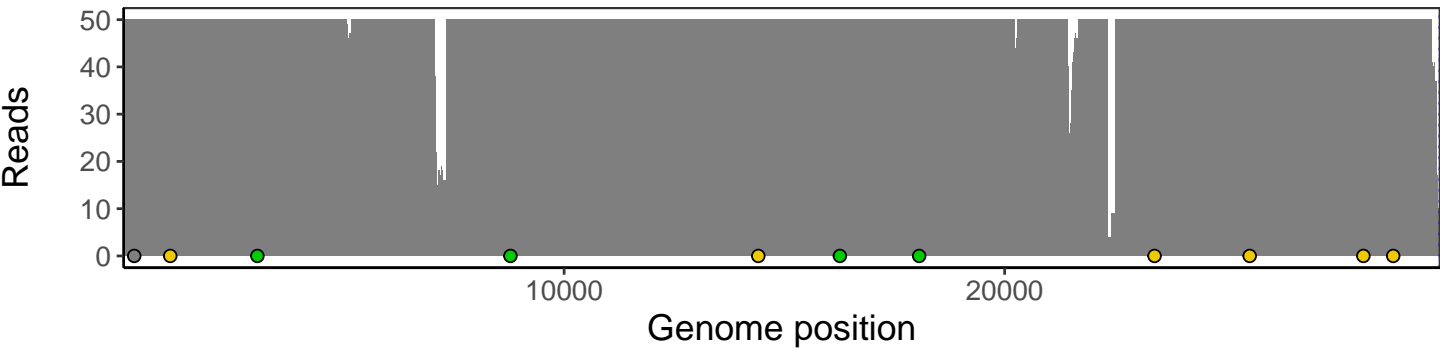
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



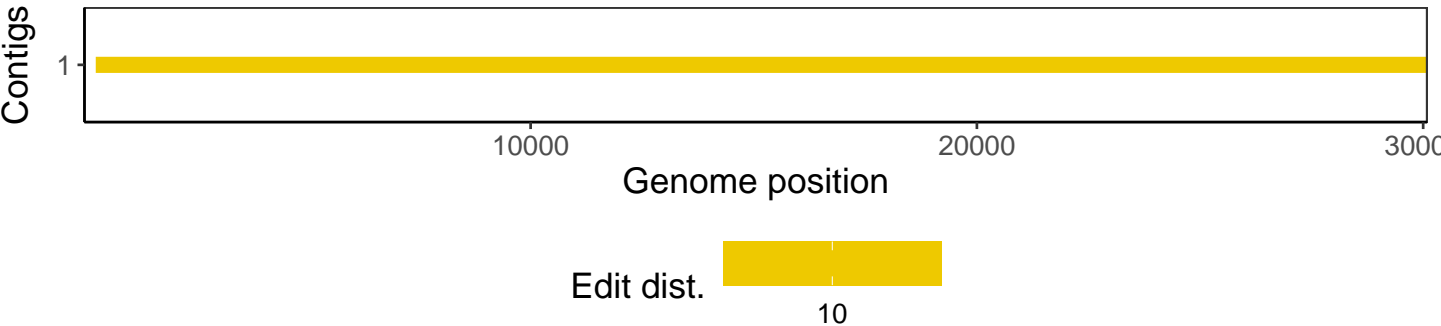
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



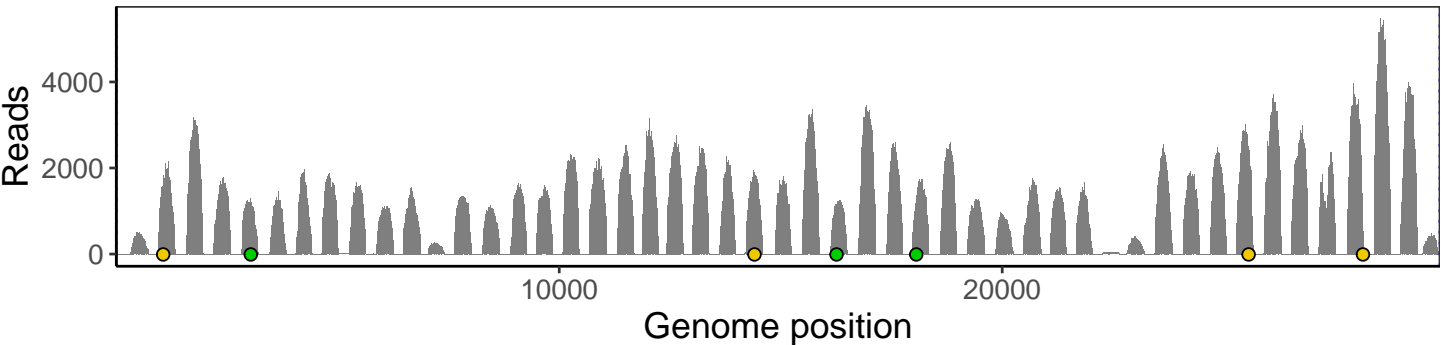
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



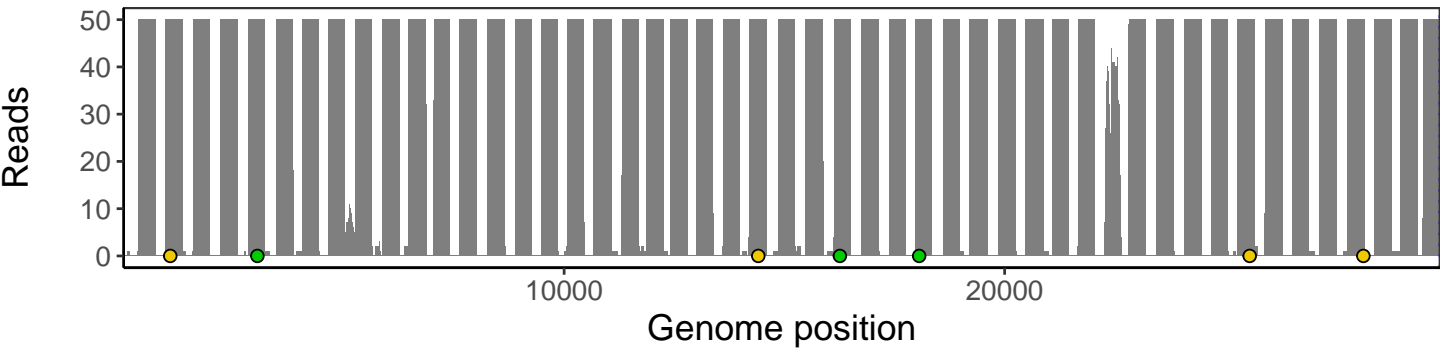
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



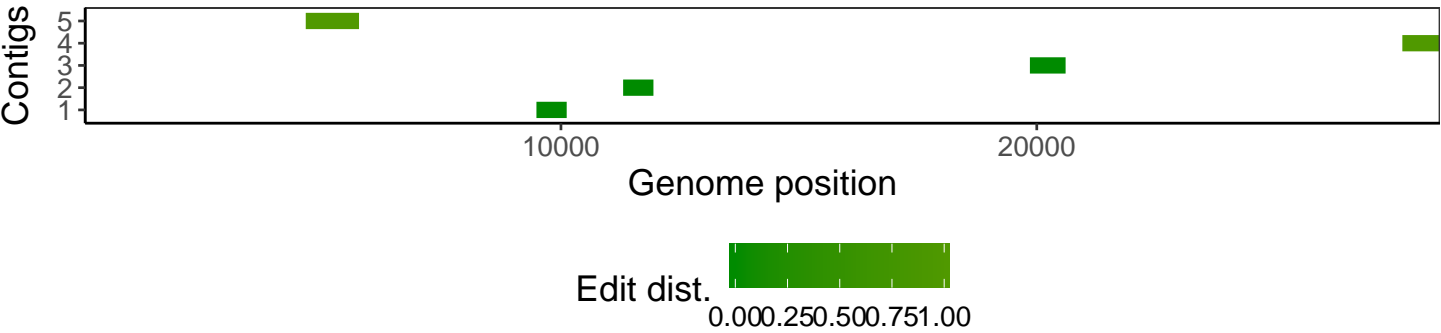
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



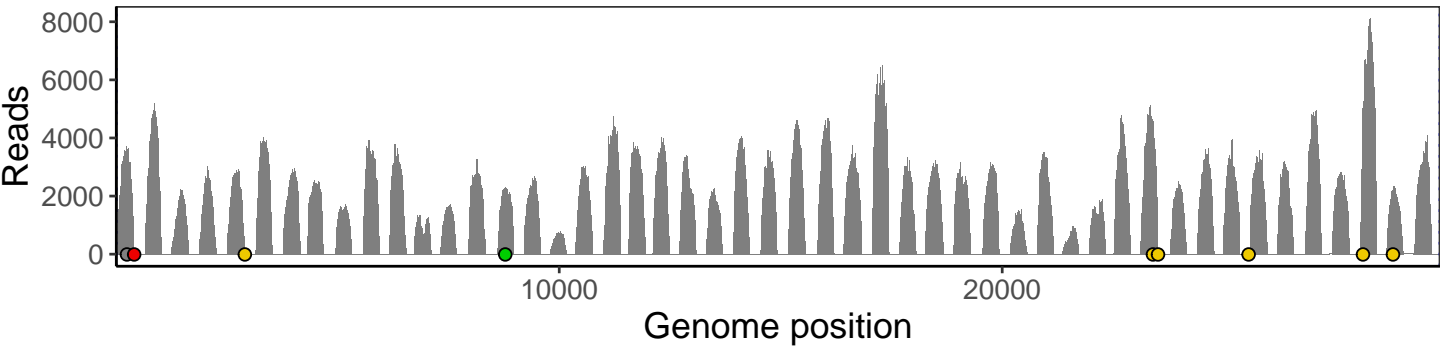
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



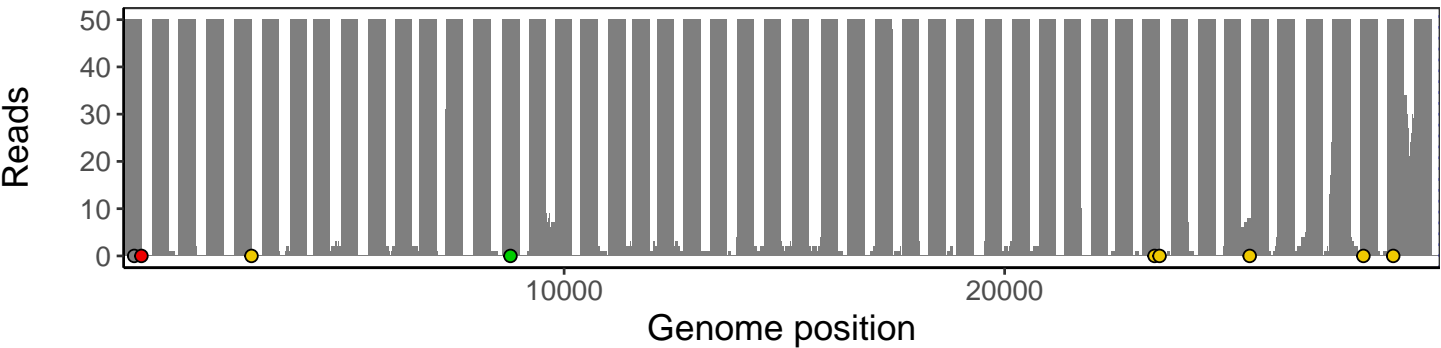
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

