

# COVID-19 subject 256

2020-09-09

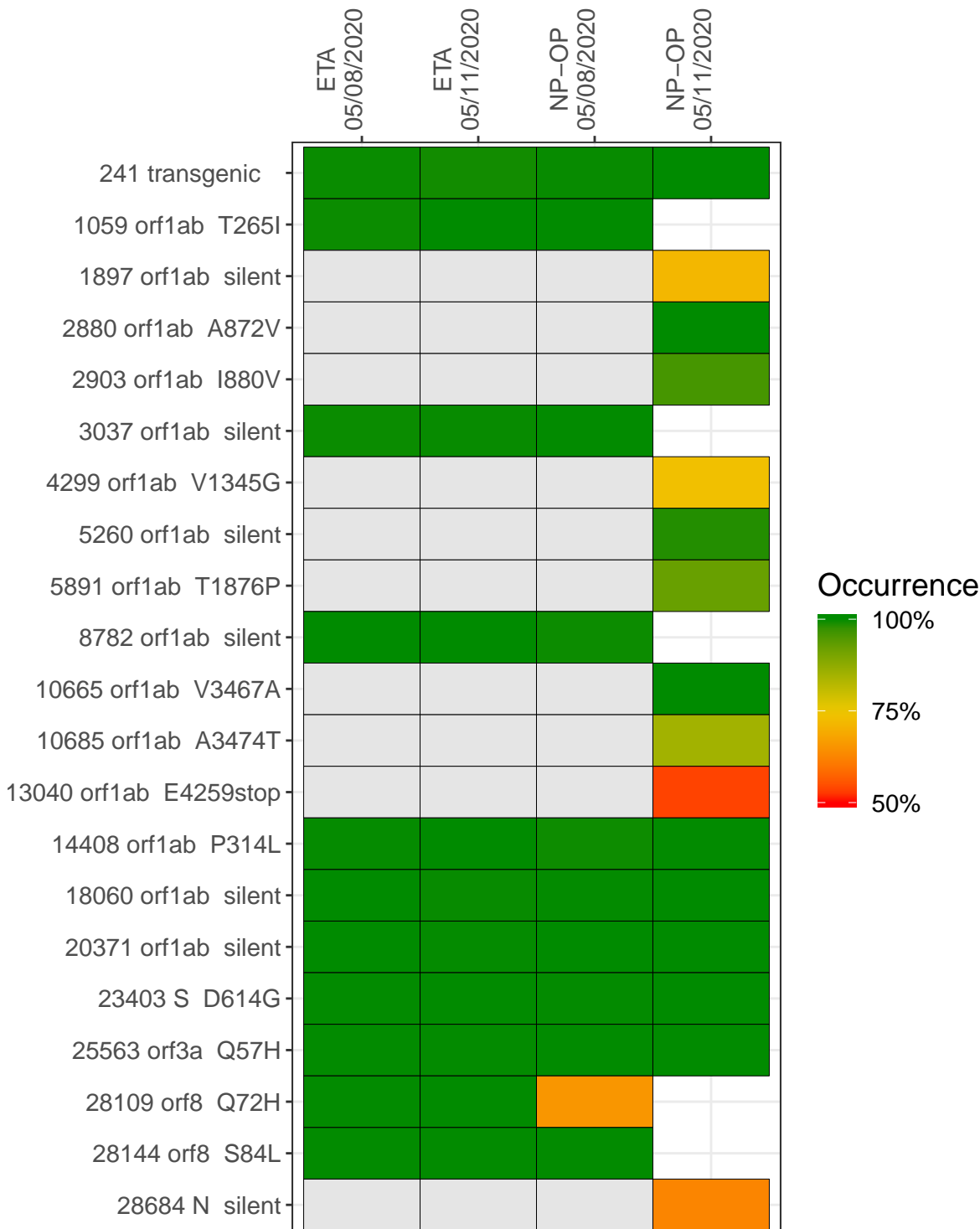
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0107	composite	NA	NP-OP	05/08/2020	29.99	99.9%	99.9%
VSP0118	composite	NA	NP-OP	05/11/2020	2.63	62.0%	58.2%
VSP0100-1	single experiment	2760000	ETA	05/08/2020	29.82	99.7%	99.7%
VSP0107-1	single experiment	595000	NP-OP	05/08/2020	29.88	99.9%	99.8%
VSP0107-2	single experiment	595000	NP-OP	05/08/2020	29.99	99.9%	99.8%
VSP0118-1	single experiment	269	NP-OP	05/11/2020	2.67	56.6%	53.5%
VSP0118-2	single experiment	1345	NP-OP	05/11/2020	0.58	9.2%	5.7%
VSP0118-3	single experiment	1345	NP-OP	05/11/2020	0.53	10.9%	7.7%
VSP0118-4	single experiment	1345	NP-OP	05/11/2020	0.60	10.6%	6.8%
VSP0123-1	single experiment	123000	ETA	05/11/2020	29.89	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.



	ETA 05/08/2020		ETA 05/11/2020		NP-OP 05/08/2020		NP-OP 05/11/2020			
241 transgenic	955		2086		5416 1966		921			
1059 orf1ab T265I	2253		1752		6183 1697					
1897 orf1ab silent	6774		3246		11384 2922		8	13	2	5
2880 orf1ab A872V	775		1517		1801 2618		31			
2903 orf1ab I880V	395		984		1886 2606		222	62	28	10
3037 orf1ab silent	400		1167		1979 2651					
4299 orf1ab V1345G	3852		2788		5611 1759			1	8	6
5260 orf1ab silent	25		600		20 490		9	51	44	56
5891 orf1ab T1876P	1796		3683		443 2357			8		5
8782 orf1ab silent	63		1367		39 998					
10665 orf1ab V3467A	619		2291		812 1133			28	12	21
10685 orf1ab A3474T	789		2560		1747 1466		1	9	3	7
13040 orf1ab E4259stop	9304		4098		7995 3424		2800			
14408 orf1ab P314L	779		2264		3269 1258		2256			
18060 orf1ab silent	174		596		1395 1319		2494			
20371 orf1ab silent	29		1003		16 241		1035			
23403 S D614G	18281		5494		8818 6370		1208		1	
25563 orf3a Q57H	2424		3050		5380 2764		1	40865		
28109 orf8 Q72H	2707		4710		4079 6933					
28144 orf8 S84L	4380		4192		1487 7643					
28684 N silent	4593		3961		11661 5348		1		7	
	VSP0100-1	VSP0123-1	VSP0107-1	VSP0107-2	VSP0118-1	VSP0118-2	VSP0118-3	VSP0118-4		

Base change

Expected

A

T

C

G

N

Ins/Del

No data

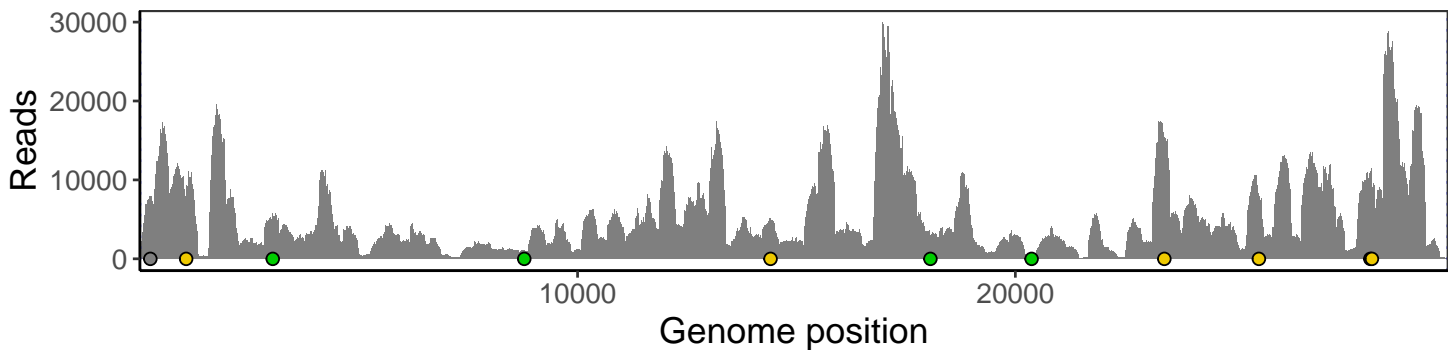
Base change



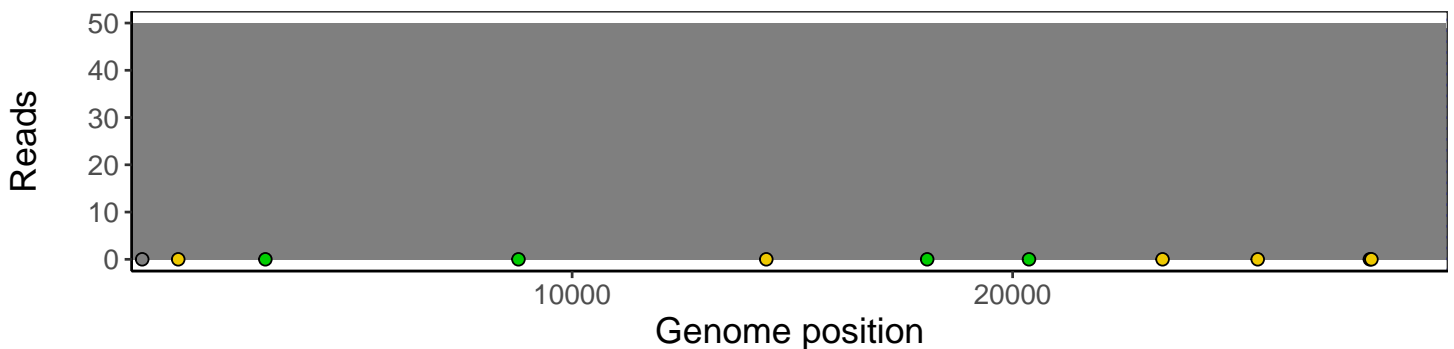
## Analyses of individual experiments and composite results.

VSP0107 | 05/08/2020 | NP-OP | 256no-t | composite result

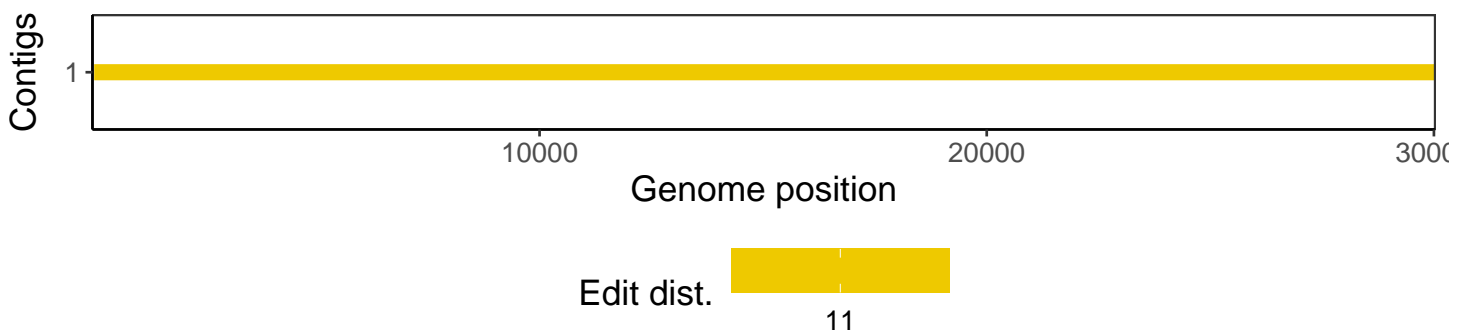
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



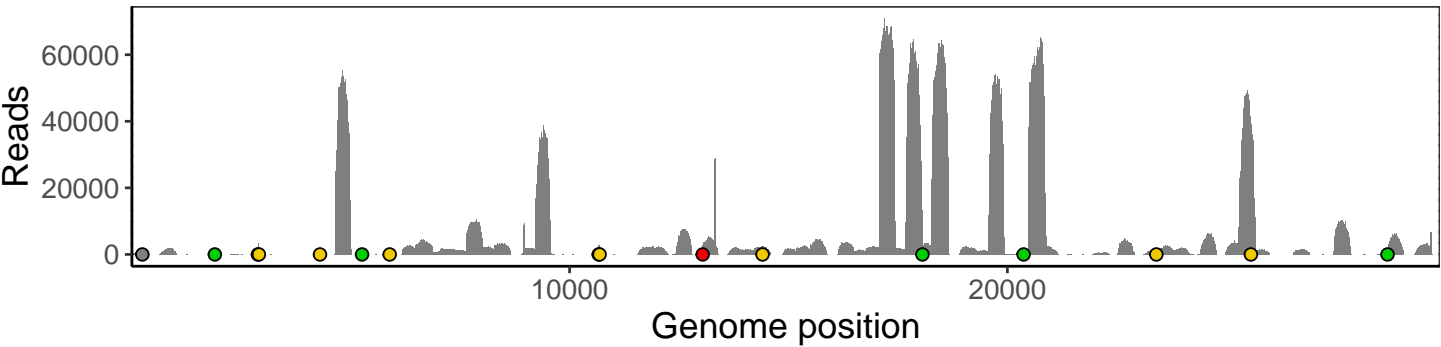
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



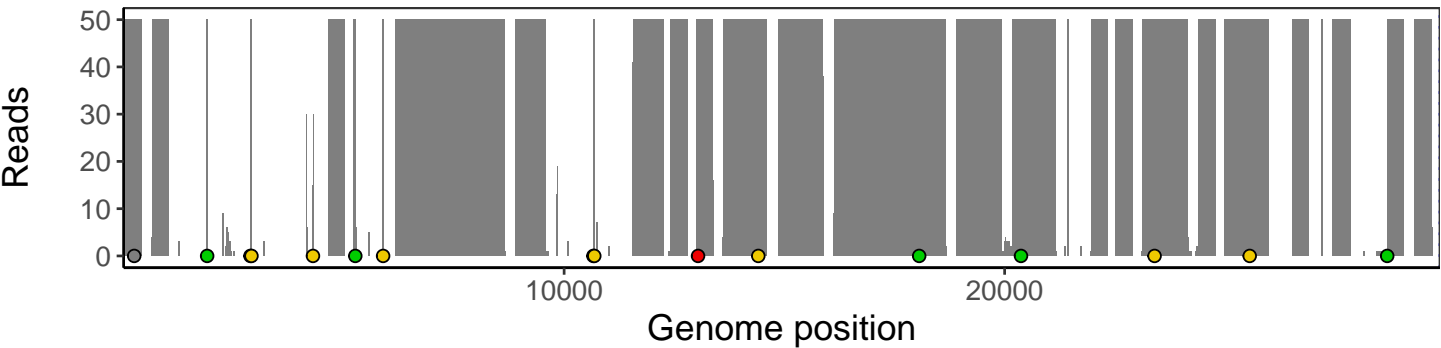
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



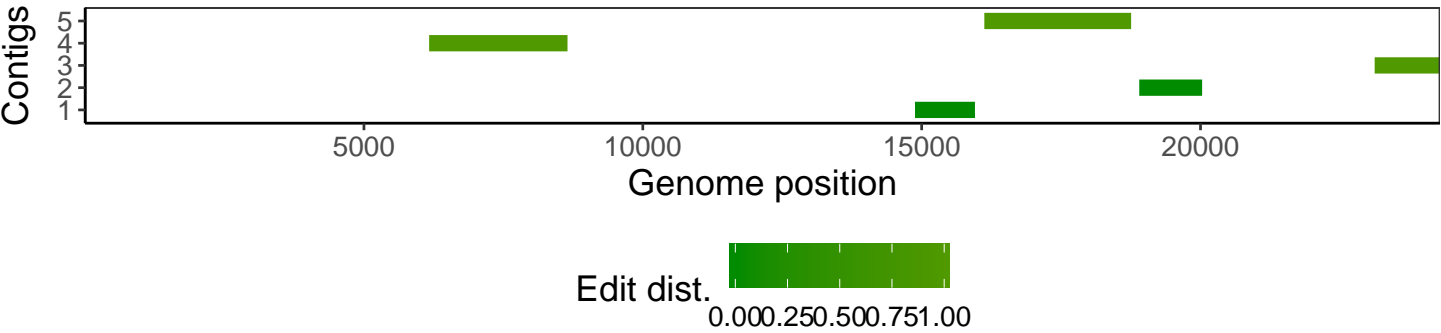
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



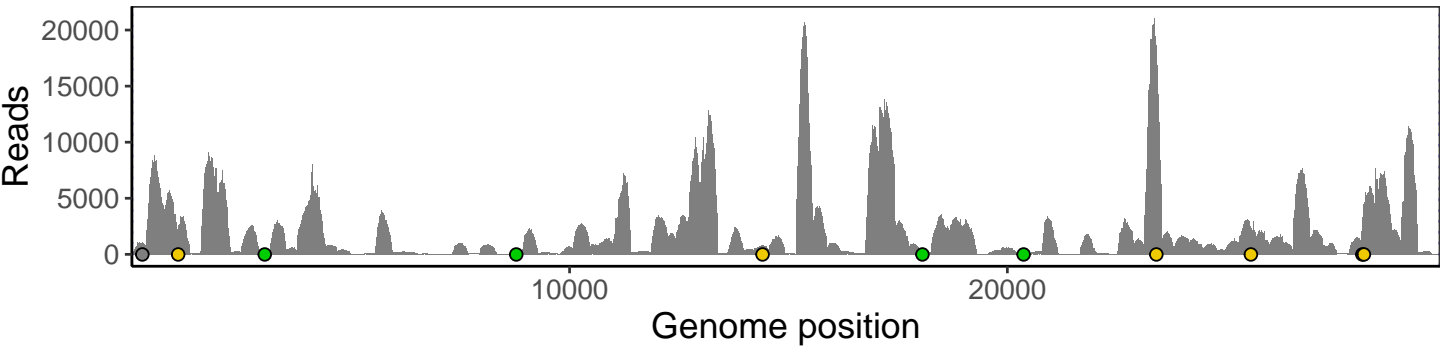
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



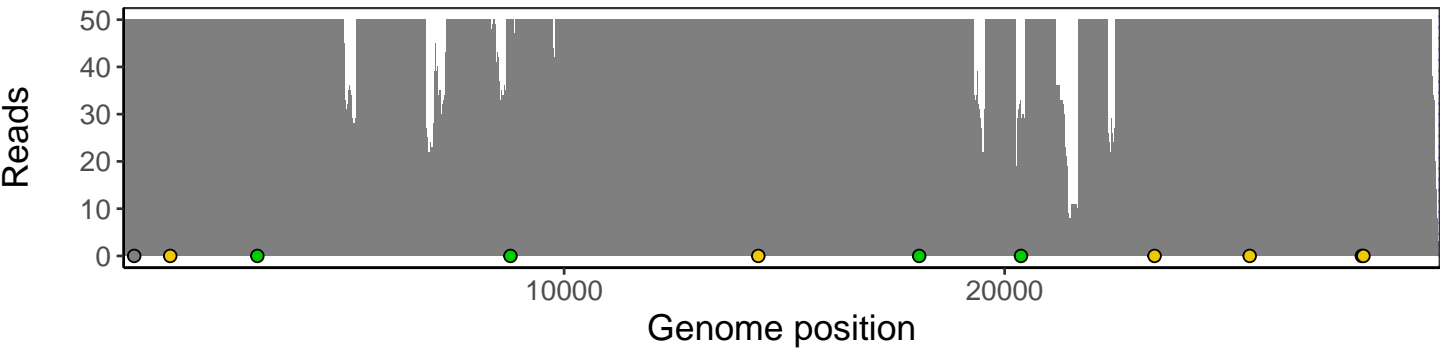
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



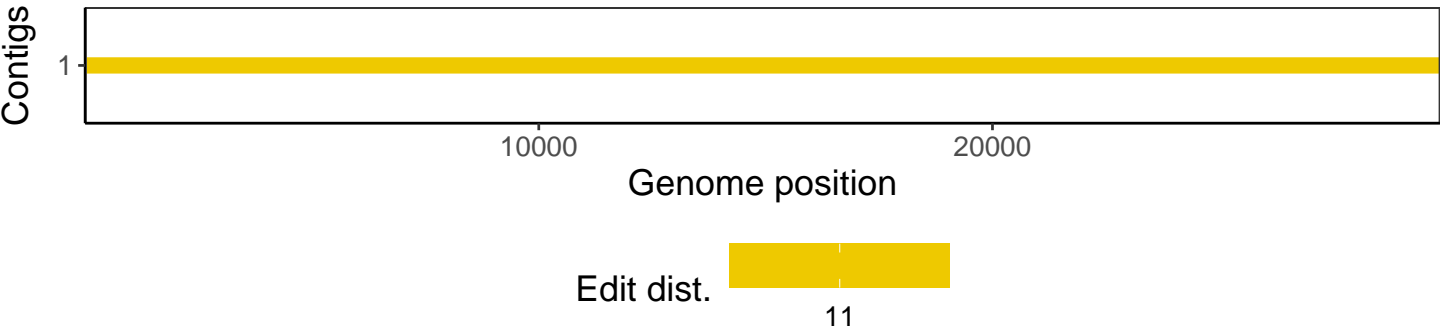
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



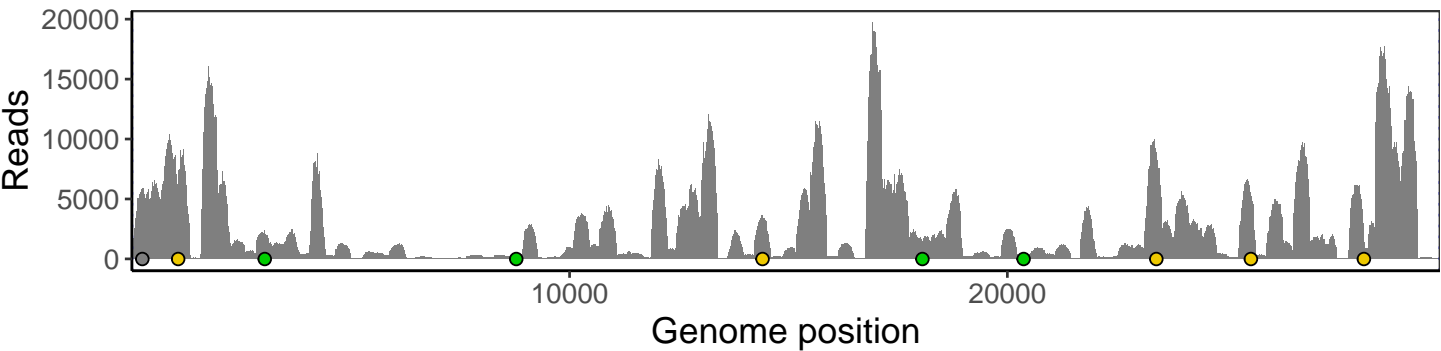
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



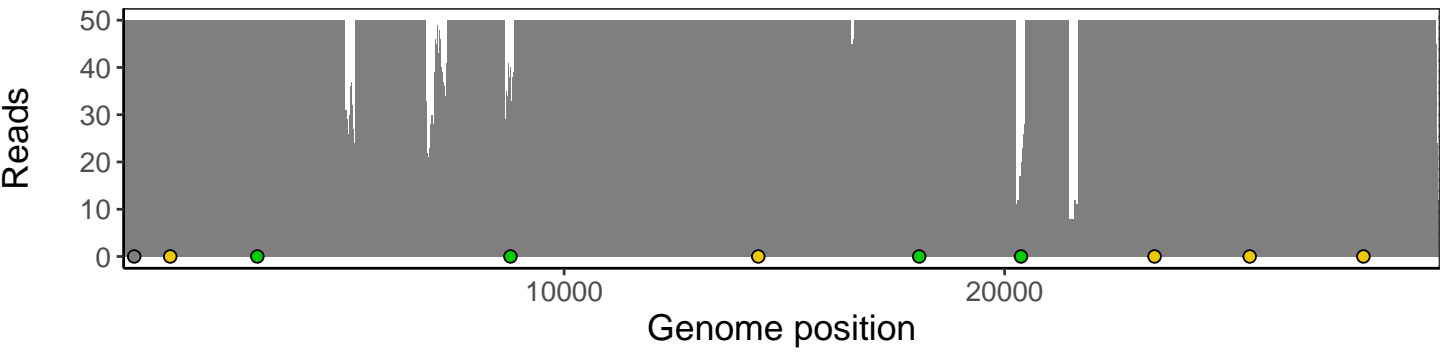
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



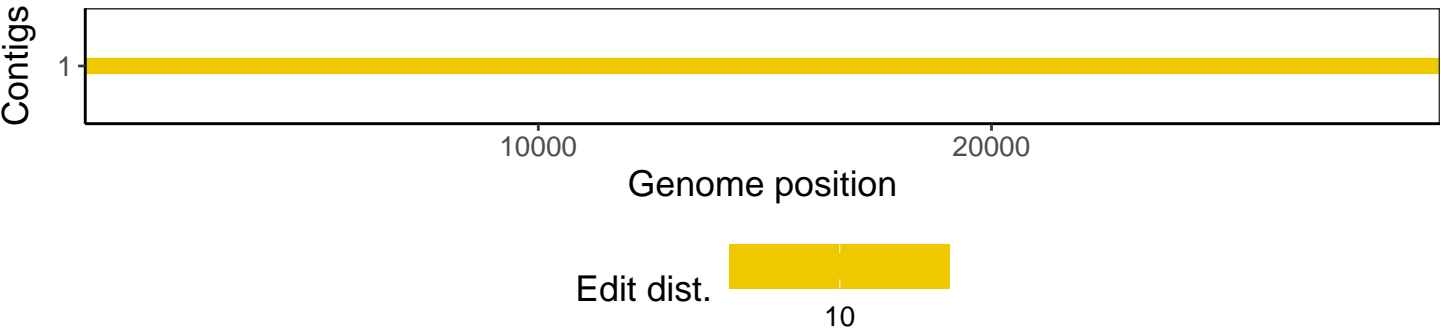
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



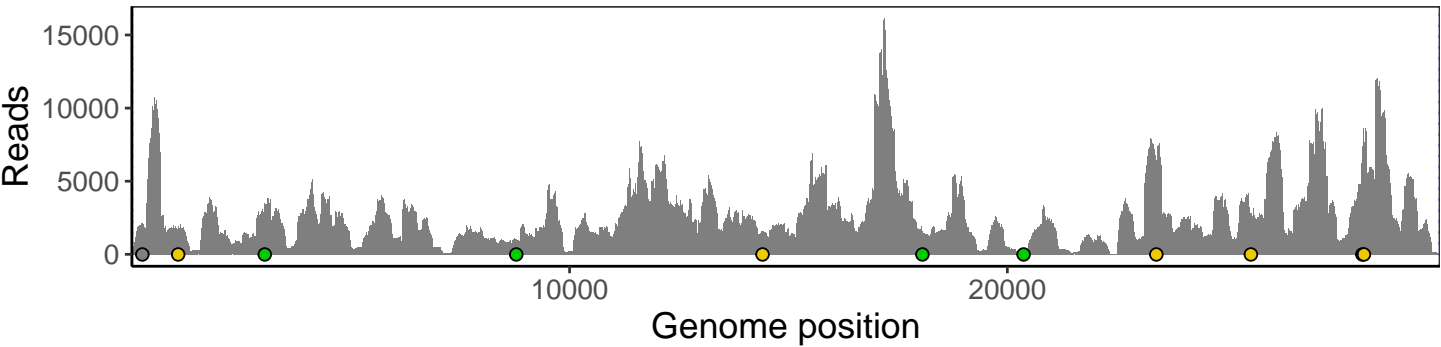
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



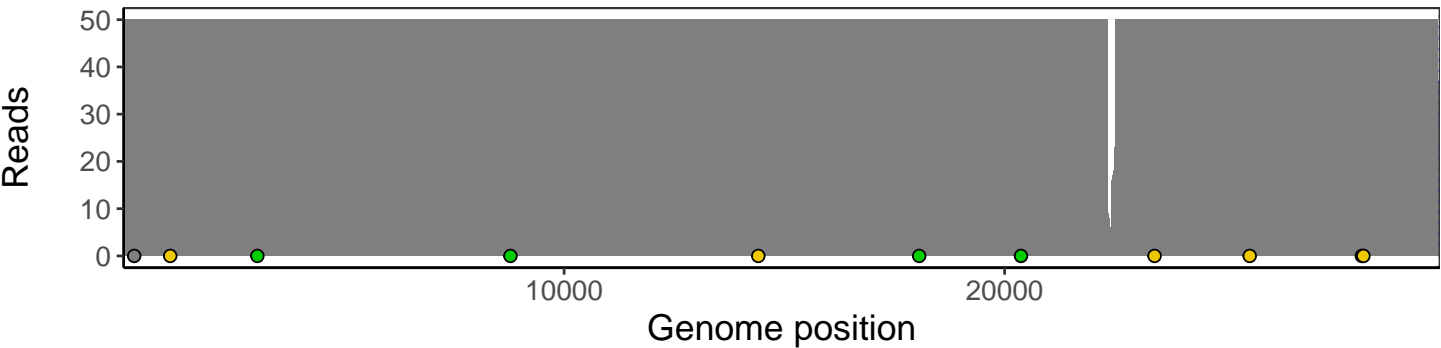
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



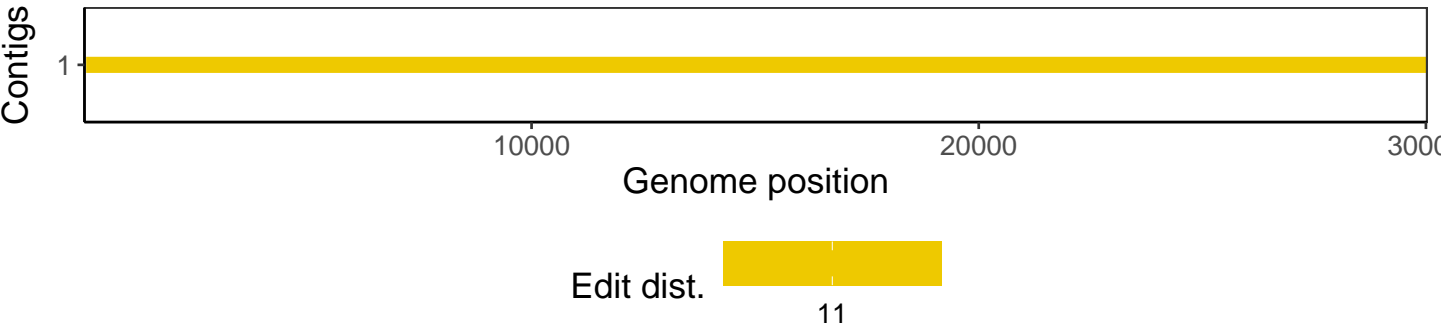
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

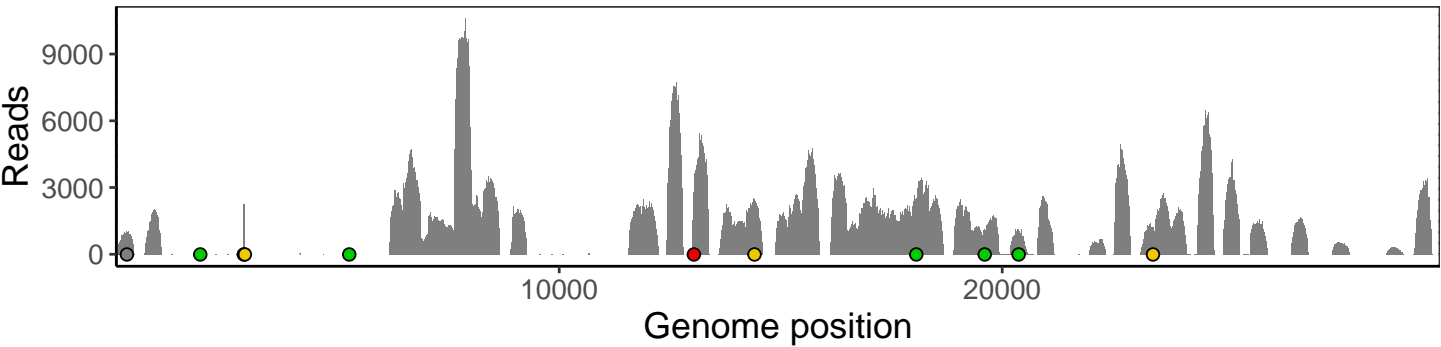


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

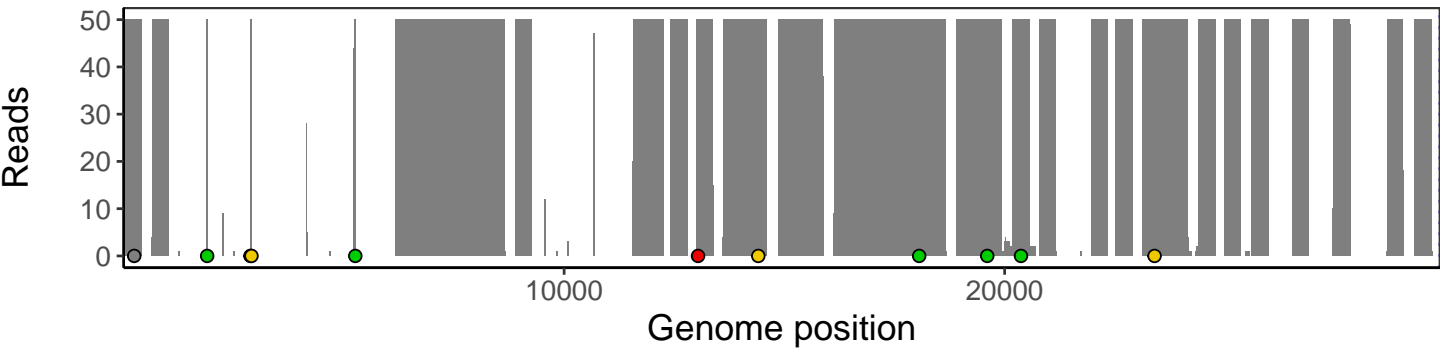




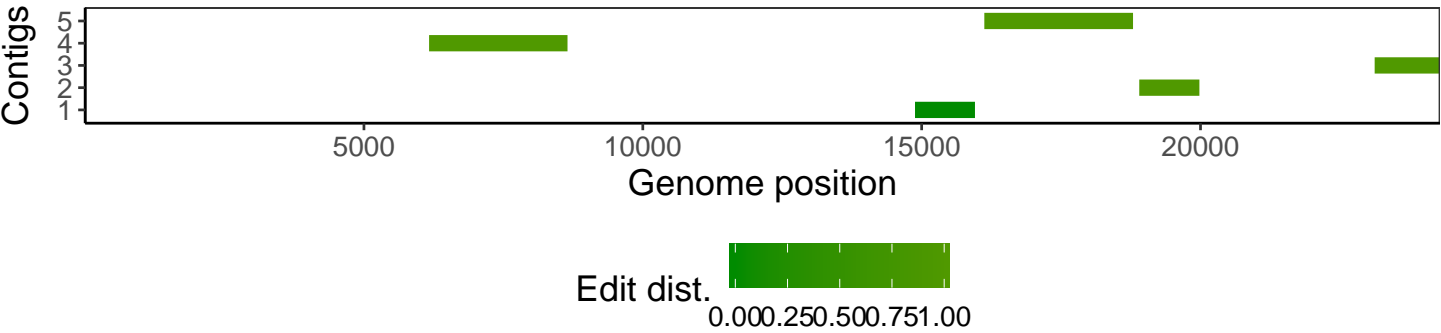
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



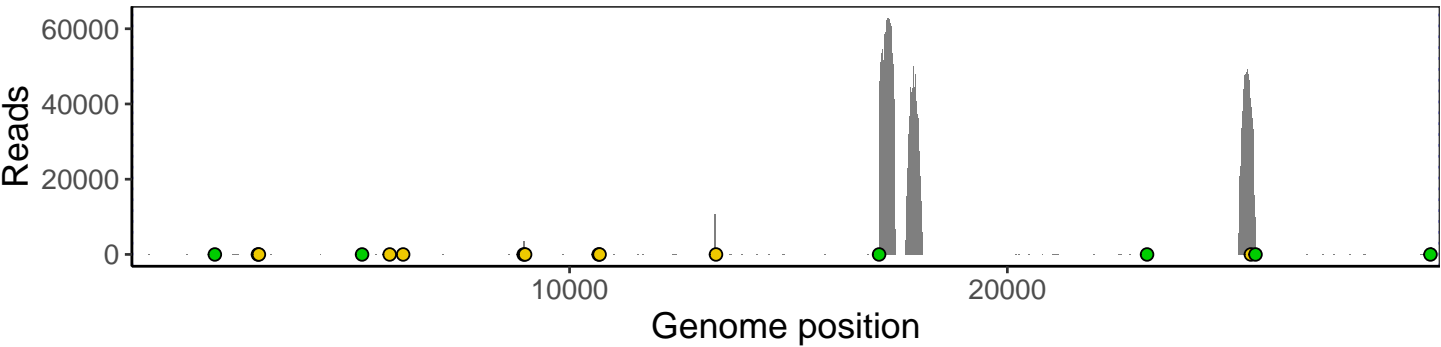
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



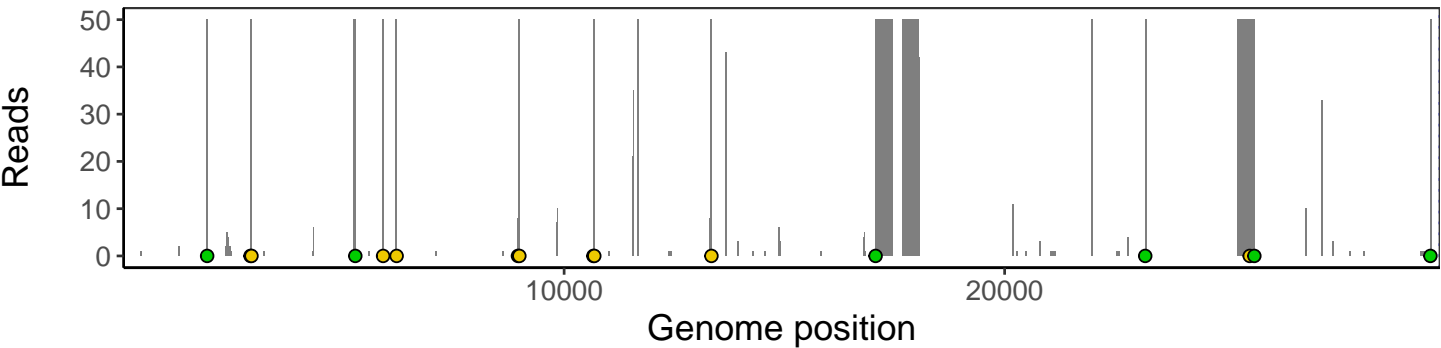
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



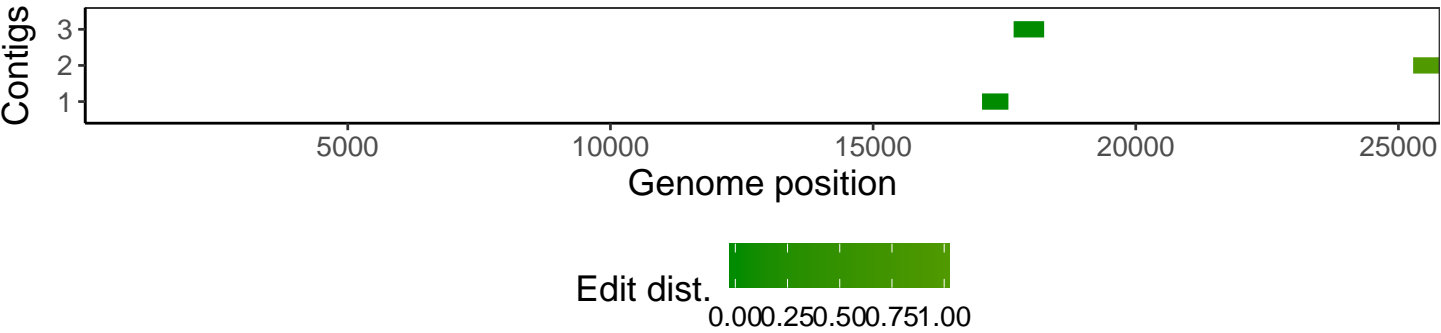
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



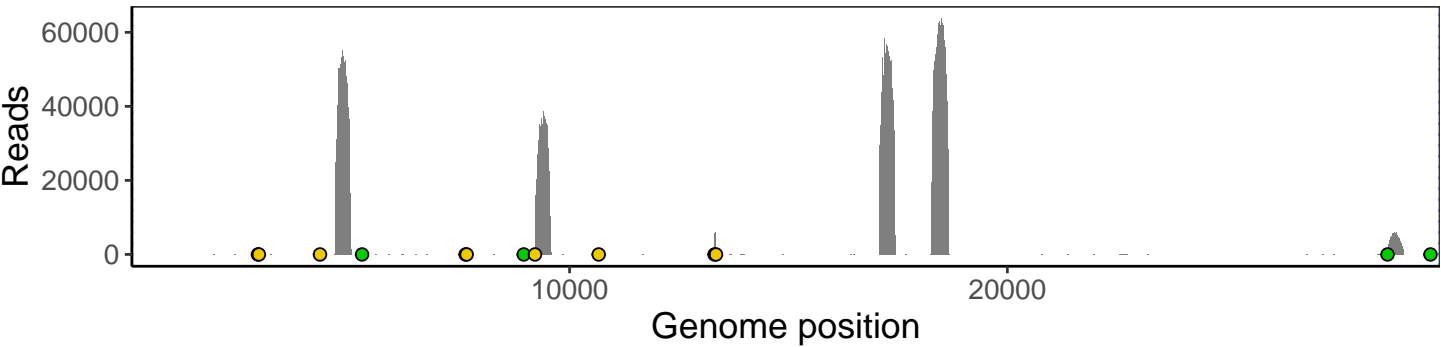
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



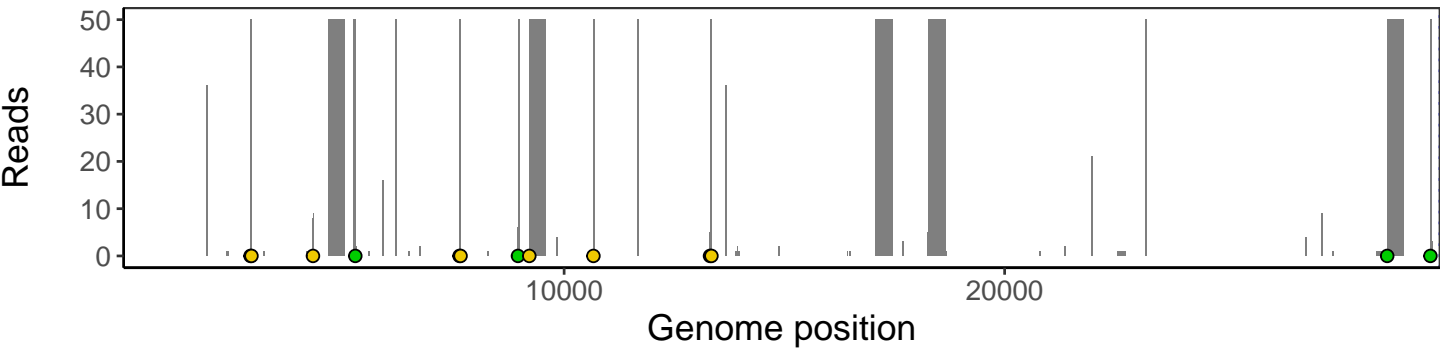
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



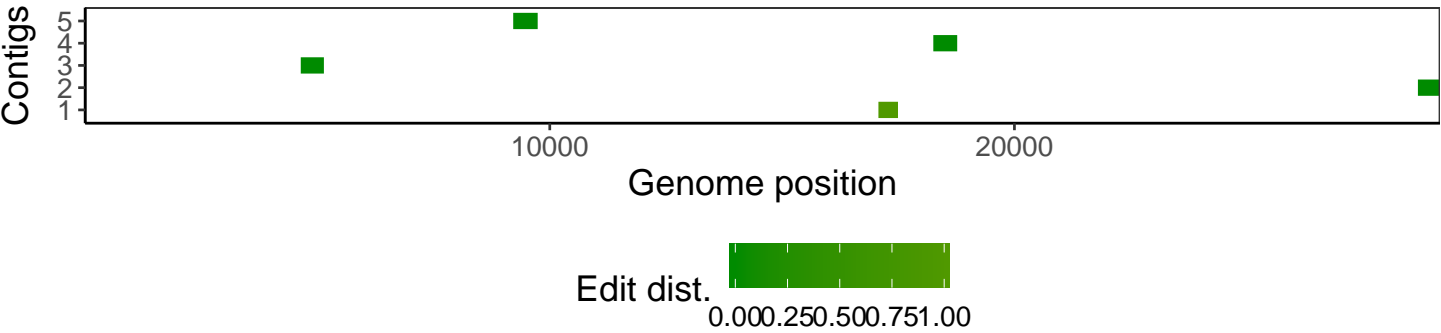
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



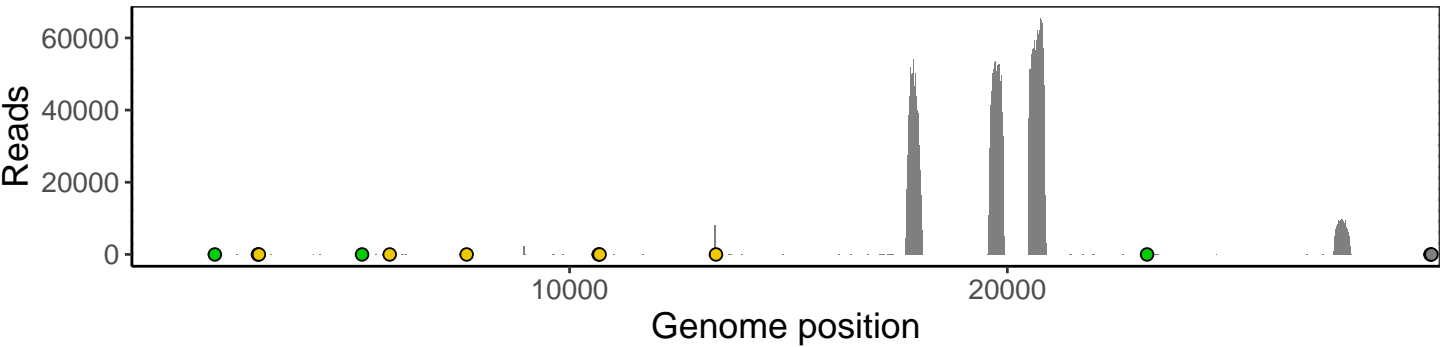
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



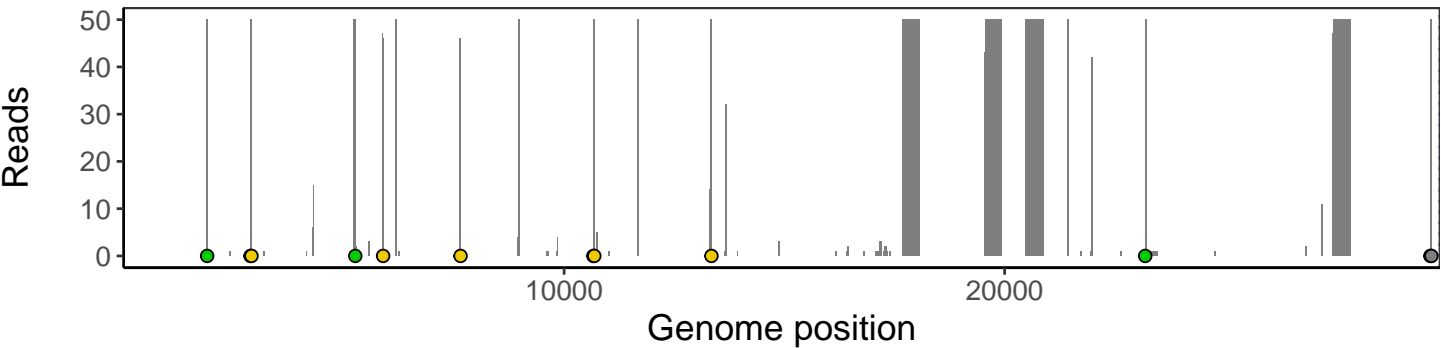
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



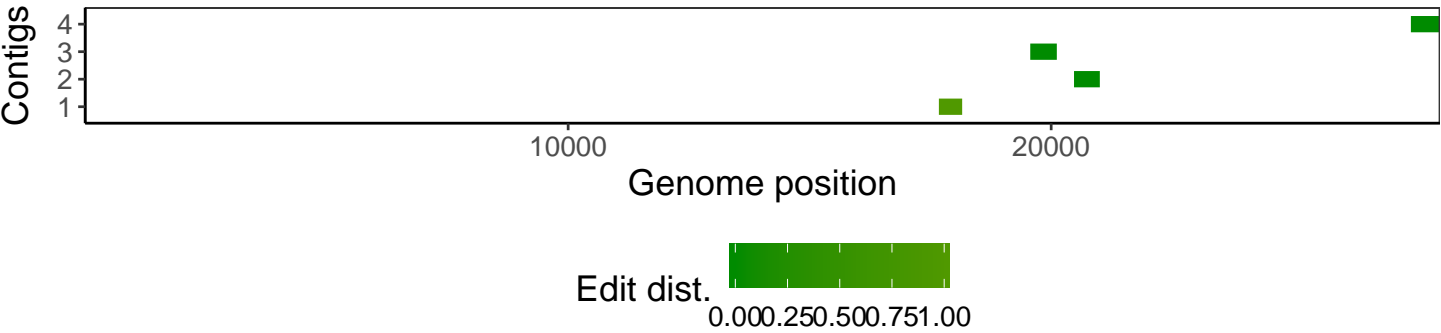
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



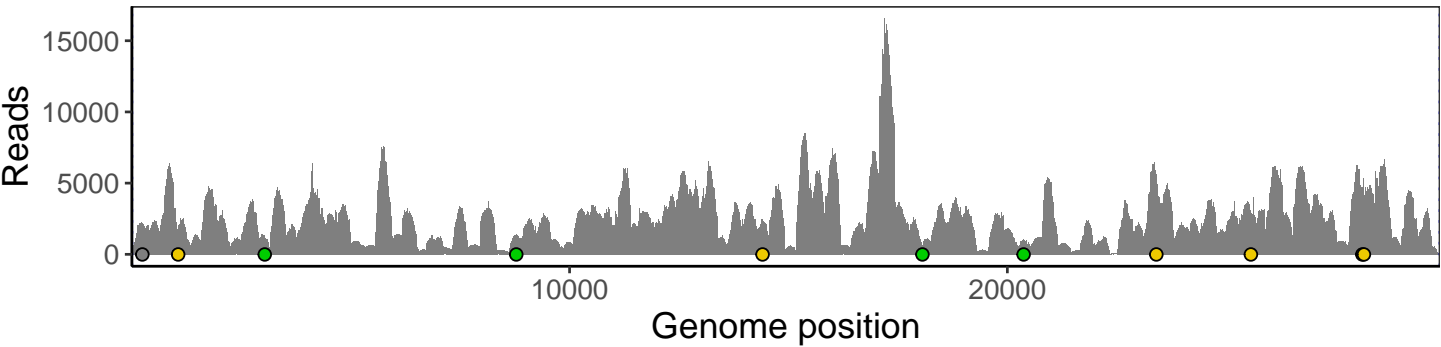
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



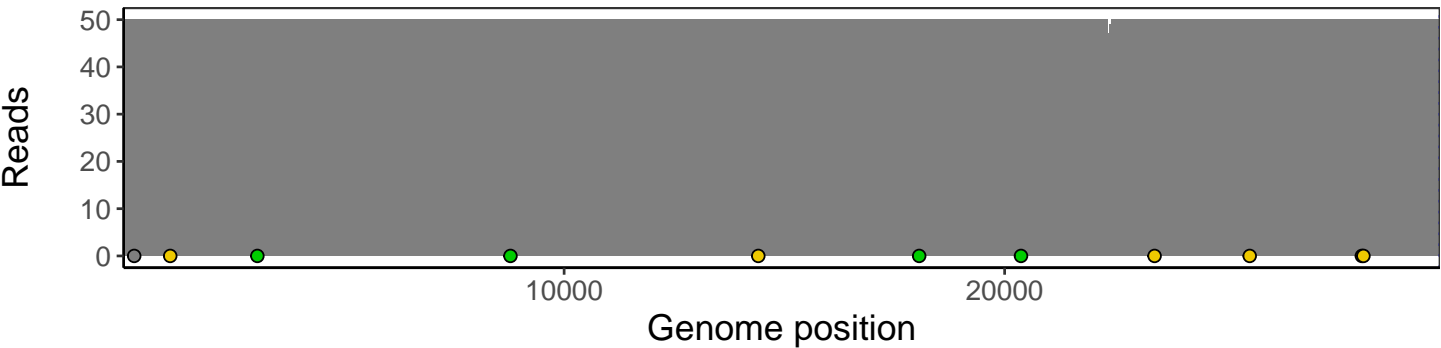
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

