

COVID-19 subject 237

2020-10-23

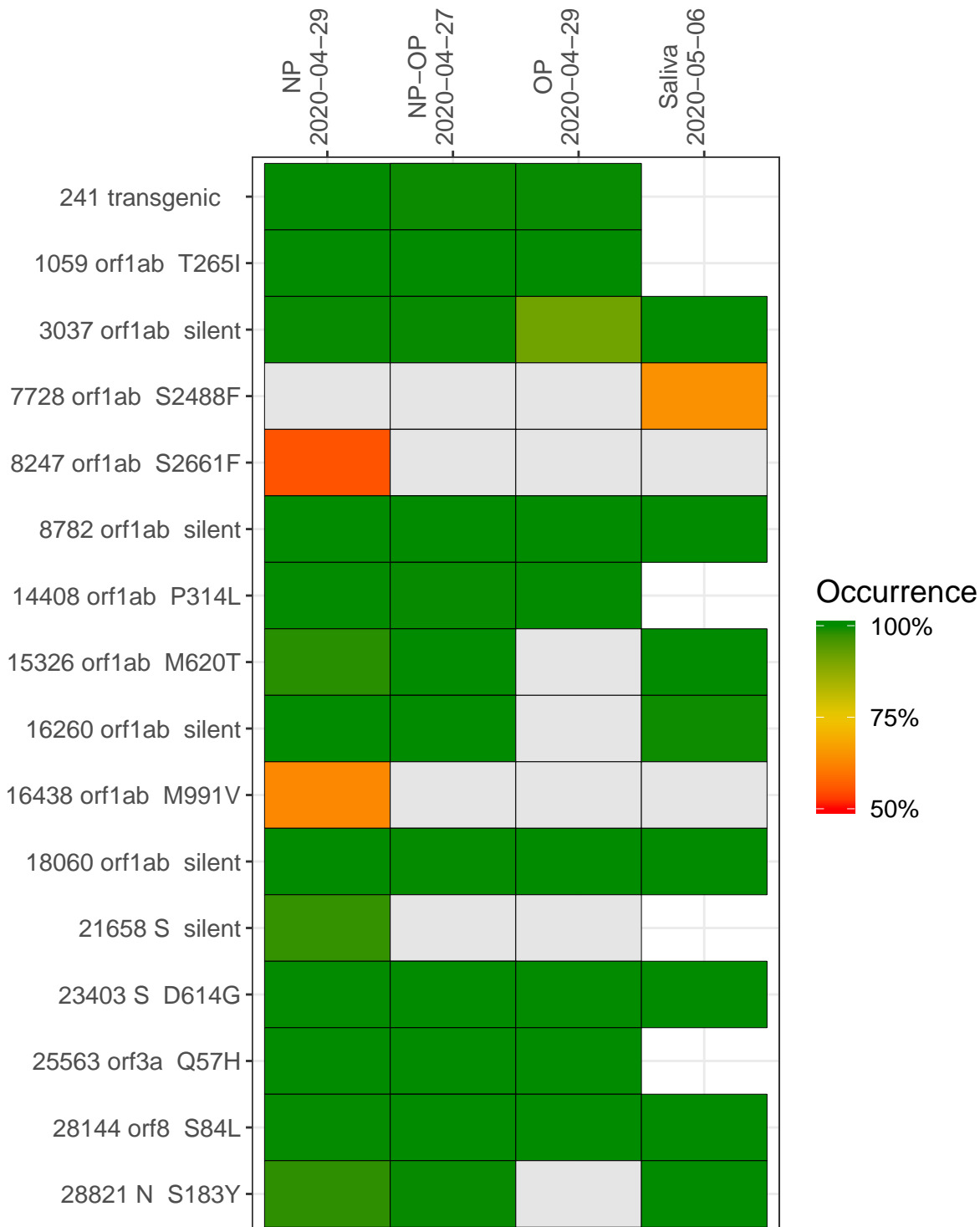
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0033	composite	NA	NP-OP	2020-04-27	29.97	99.9%	99.8%
VSP0034	composite	NA	NP-OP	2020-04-27	29.84	99.8%	99.8%
VSP0037	composite	NA	NP	2020-04-29	24.92	99.3%	98.9%
VSP0038	composite	NA	OP	2020-04-29	29.92	99.9%	99.5%
VSP0033-1m	single experiment	NA	NP-OP	2020-04-27	28.58	99.9%	99.8%
VSP0033-2	single experiment	2985	NP-OP	2020-04-27	29.89	99.9%	99.5%
VSP0034-1m	single experiment	NA	NP-OP	2020-04-27	29.81	99.8%	99.8%
VSP0034-2	single experiment	1895000	NP-OP	2020-04-27	22.45	98.8%	98.5%
VSP0037-1m	single experiment	NA	NP	2020-04-29	6.66	89.3%	86.2%
VSP0037-2	single experiment	770	NP	2020-04-29	11.42	95.2%	94.4%
VSP0038-1m	single experiment	NA	OP	2020-04-29	22.46	99.9%	99.3%
VSP0038-2	single experiment	2060	OP	2020-04-29	22.32	99.7%	98.9%
VSP0084-1	single experiment	NA	Saliva	2020-05-06	4.70	78.5%	77.0%

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 2020-04-29		NP-OP 2020-04-27				OP 2020-04-29		Saliva 2020-05-01
241 transgenic	699	5697	1323	1605	707	47	1444	296	
1059 orf1ab T265I	2049	1689	2434	1449	8306	73	320	86	
3037 orf1ab silent	752	598	1505	2071	4390	169	458	131	188
7728 orf1ab S2488F		2104	2158	2148	7073	108	384	153	124
8247 orf1ab S2661F	24	1000	765	1583	541	71	713	331	191
8782 orf1ab silent	18	1204	740	1697	316	102	841	165	151
14408 orf1ab P314L	2664	2008	2338	1697	8853	107	1259	228	
15326 orf1ab M620T	30	1938	1839	2687	1186	135	1381	635	293
16260 orf1ab silent		266	1236	1777	4160	140	345	131	280
16438 orf1ab M991V	6	292	1130	1918	3058	139	494	167	256
18060 orf1ab silent	2504	1397	1410	1189	4627	75	677	125	111
21658 S silent	9	360	167	238	94	27	188	44	
23403 S D614G	324	4742	1624	4370	1080	216	768	429	201
25563 orf3a Q57H	1687	2698	3166	2321	11480	114	587	121	
28144 orf8 S84L	1087	3472	2493	6260	3018	313	1092	212	428
28821 N S183Y	289	161	1130	1808	584	69	196	10	36
	VSP0037-1m	VSP0037-2	VSP0033-1m	VSP0033-2	VSP0034-1m	VSP0034-2	VSP0038-1m	VSP0038-2	VSP0084-1

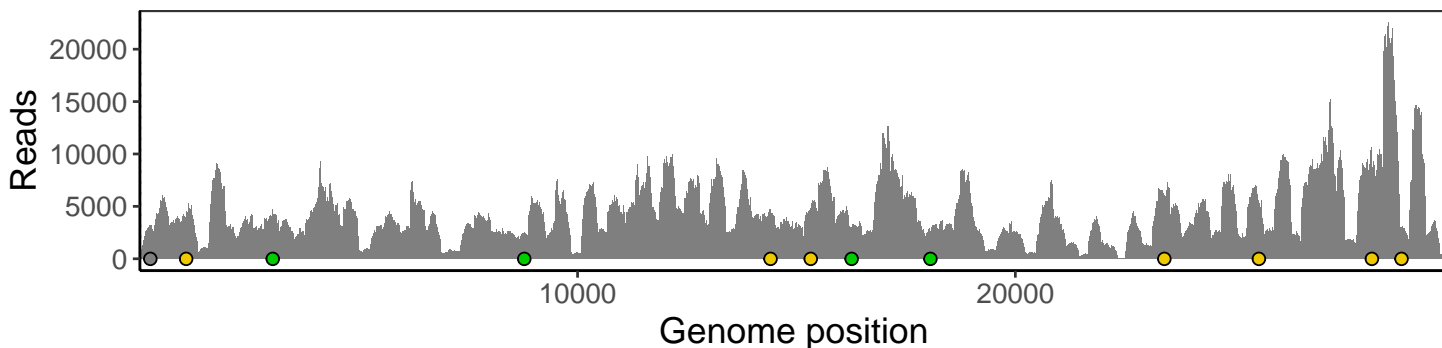
Base change

Expected	
A	
T	
C	
G	
N	
Ins/Del	
No data	

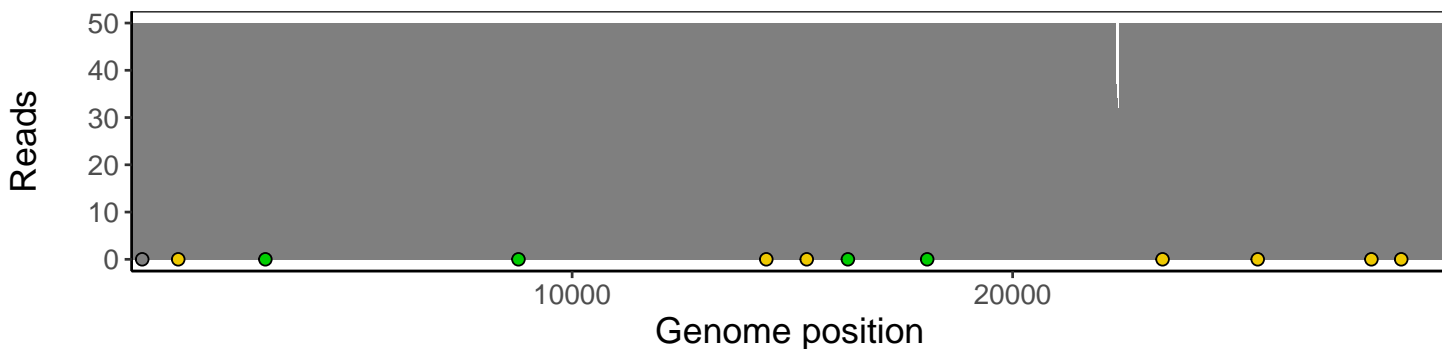
Analyses of individual experiments and composite results.

VSP0033 | 2020-04-27 | NP-OP | 237-tri | composite result

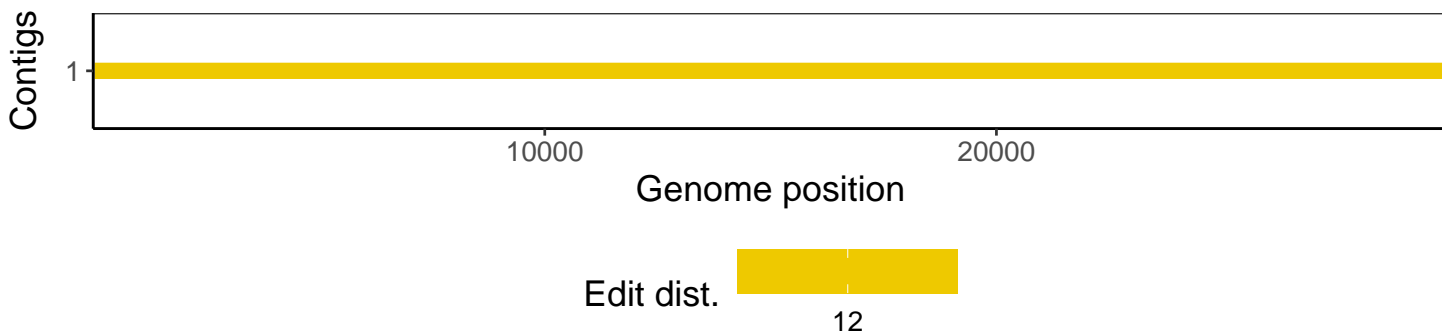
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



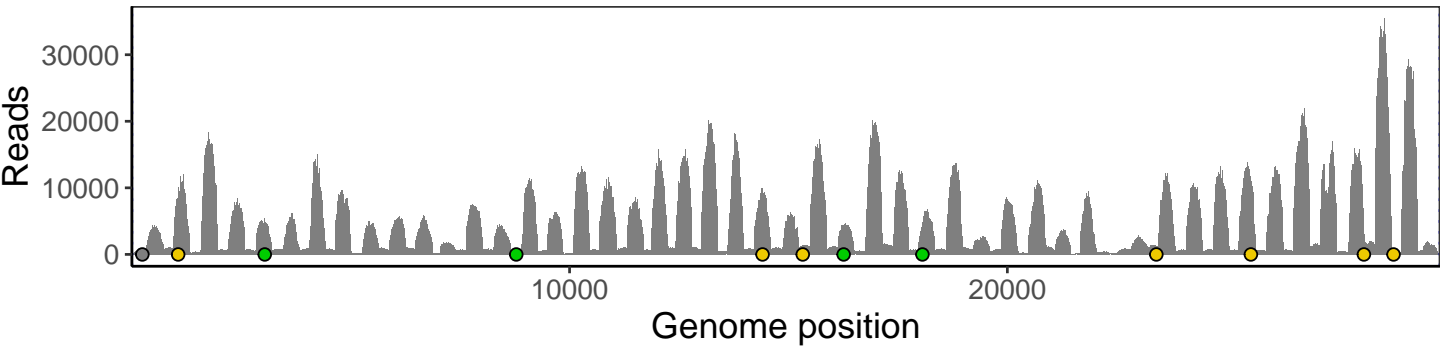
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



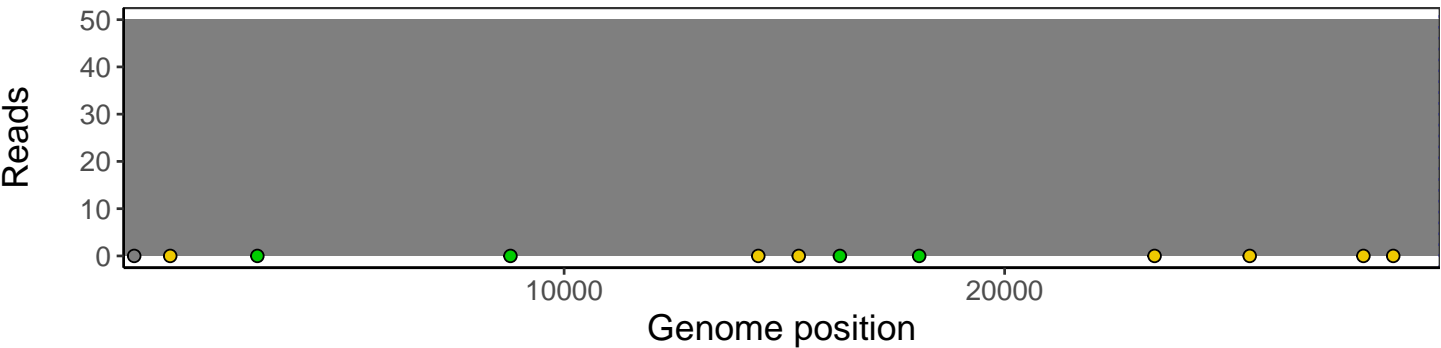
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



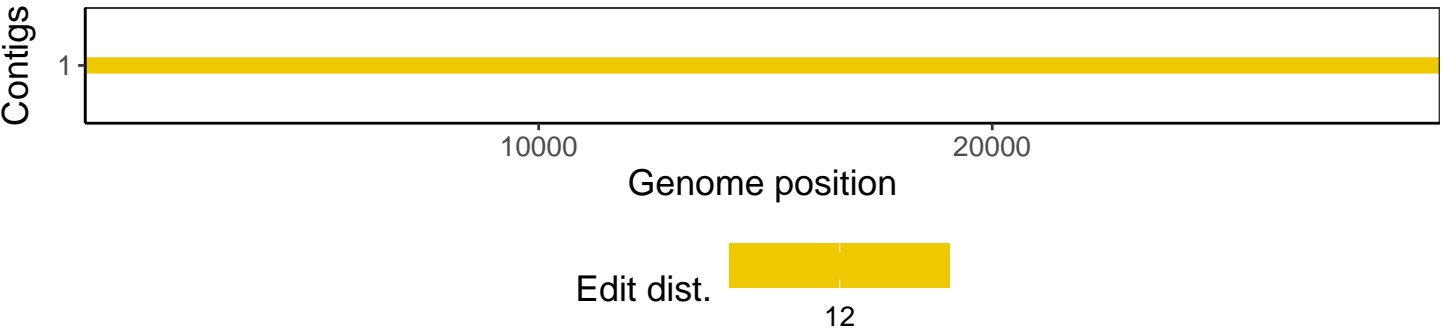
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



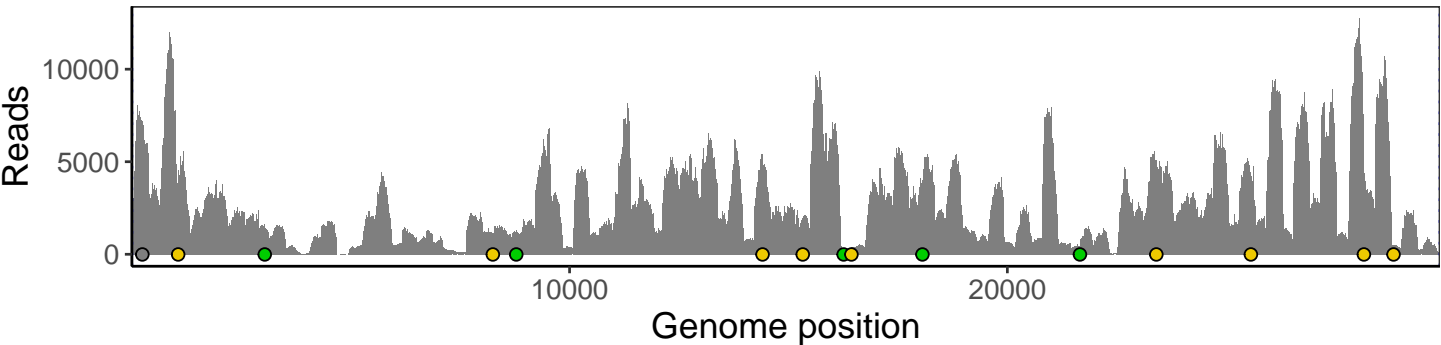
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



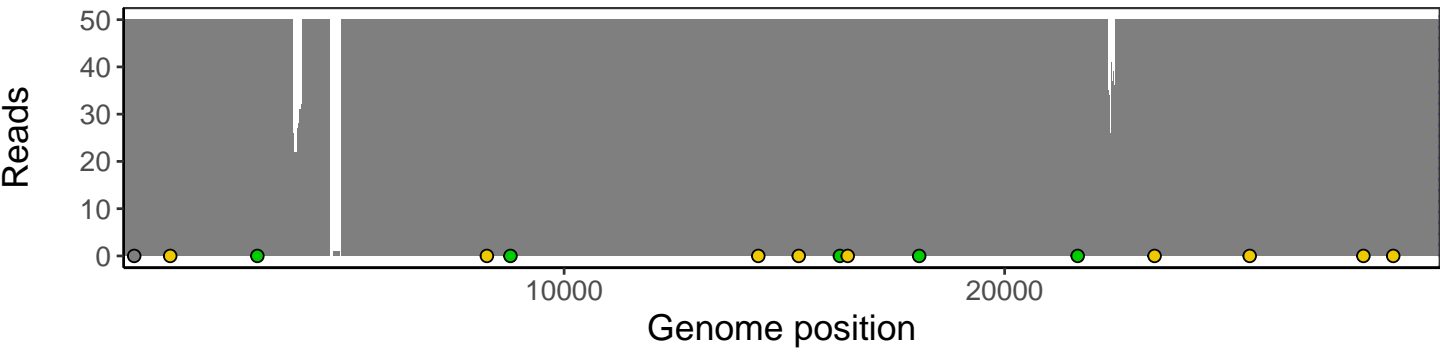
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



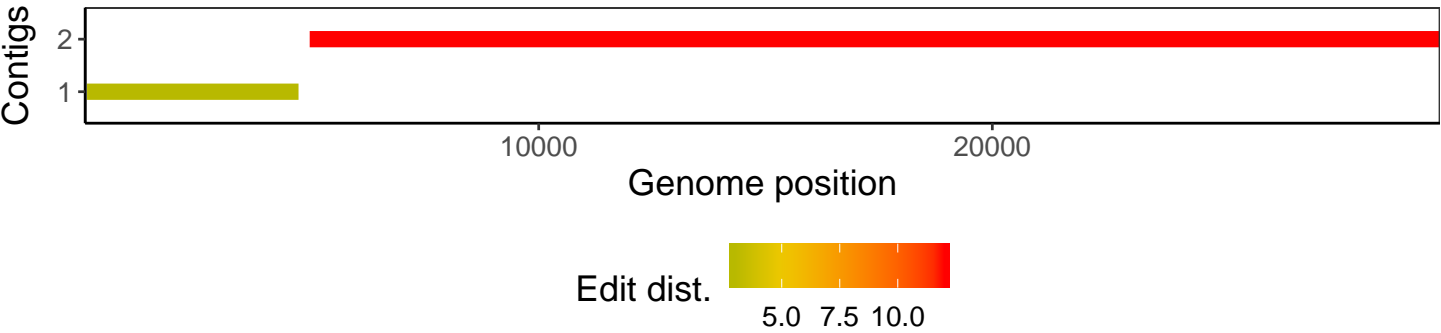
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



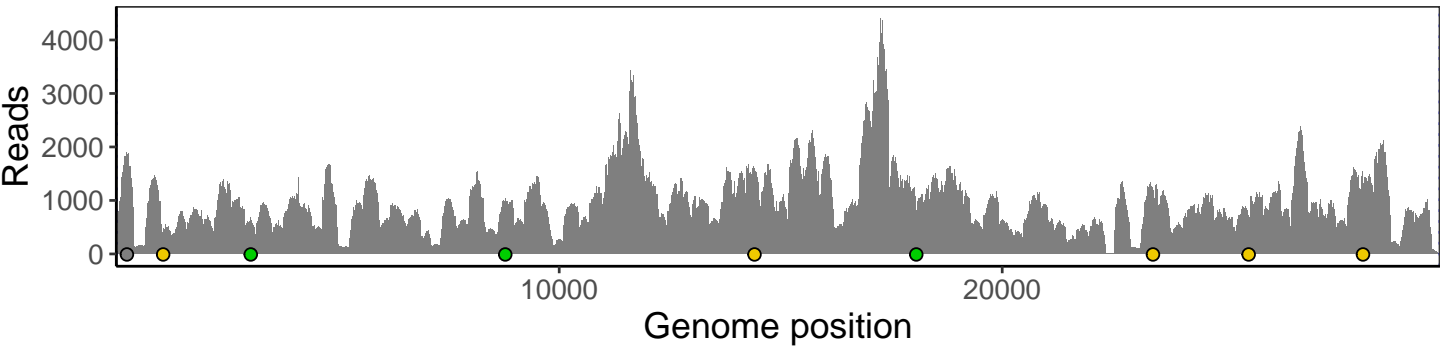
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



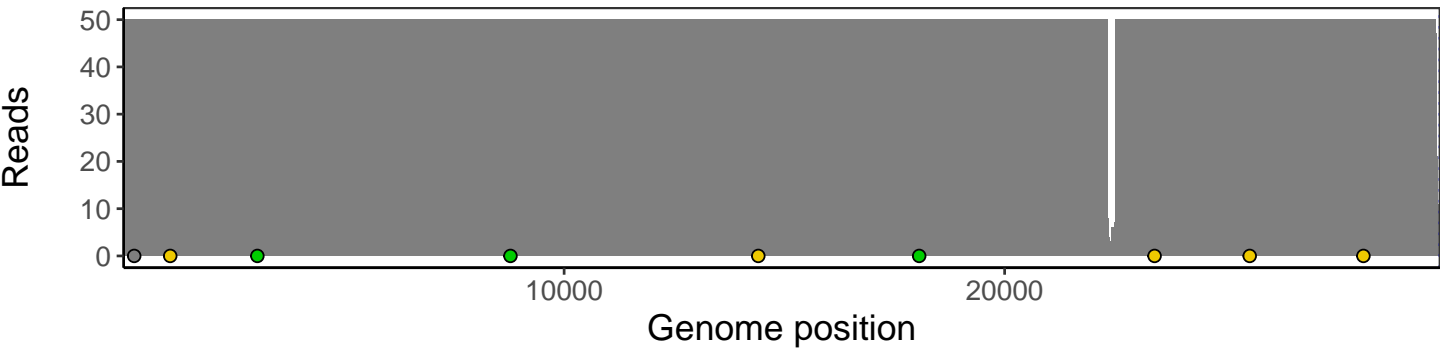
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



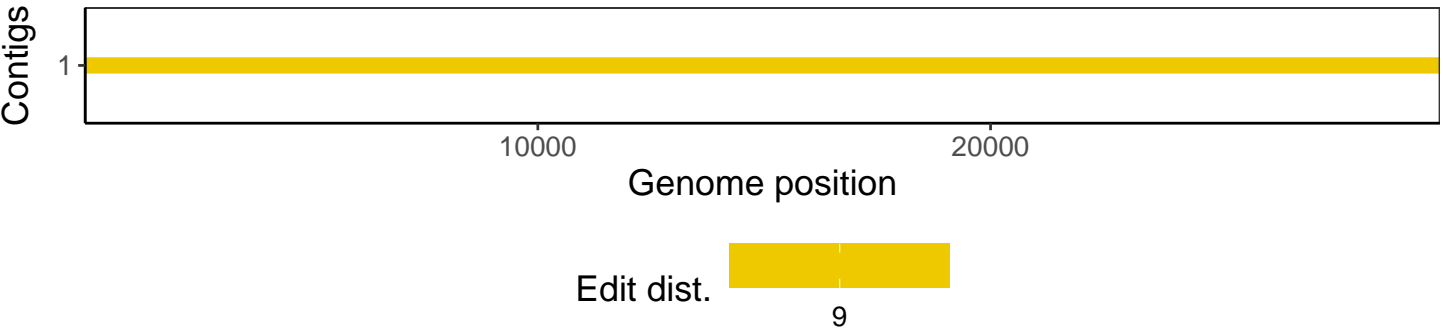
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



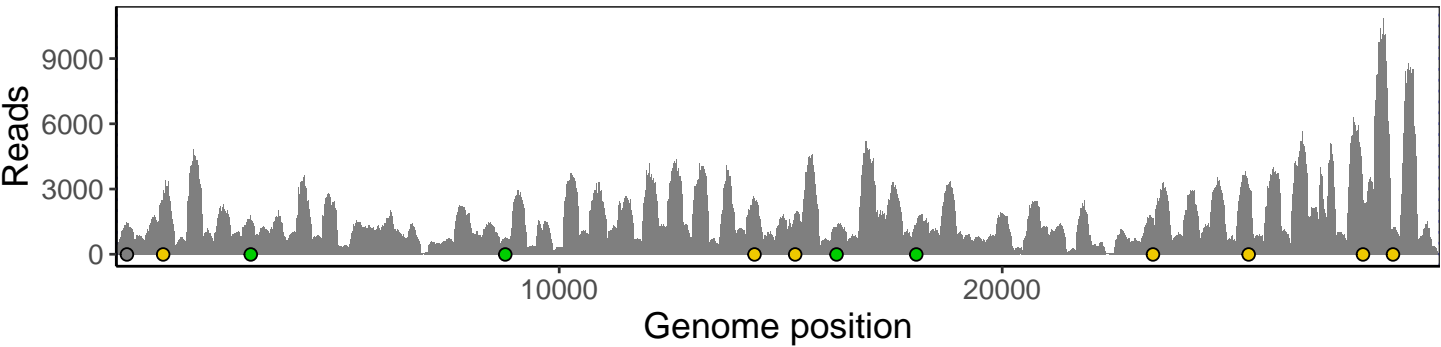
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



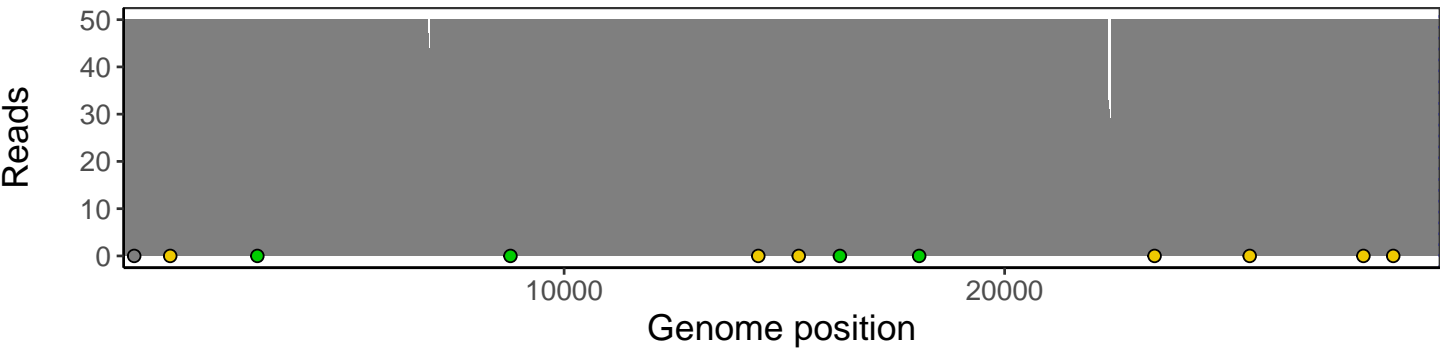
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



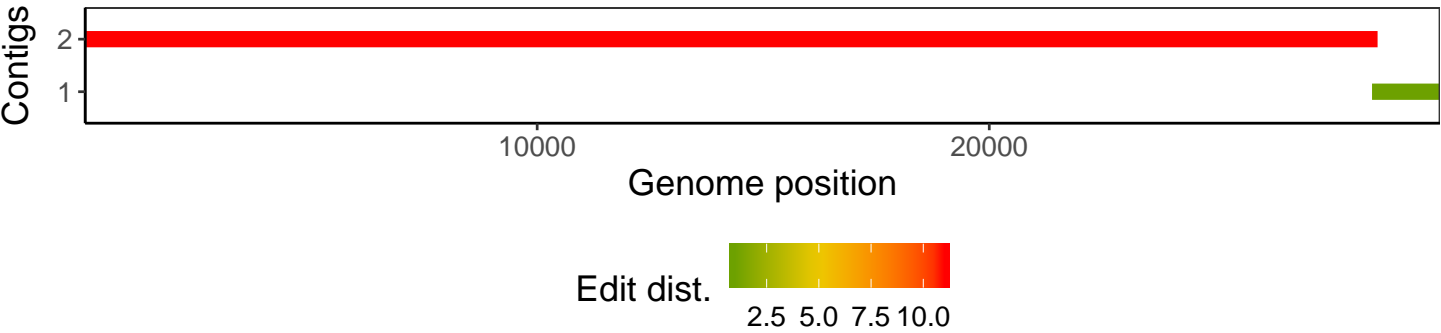
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



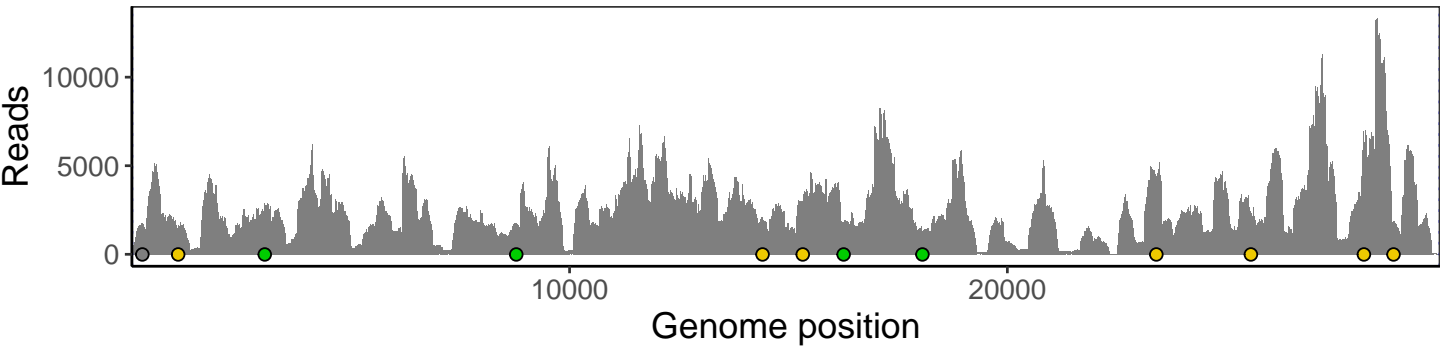
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



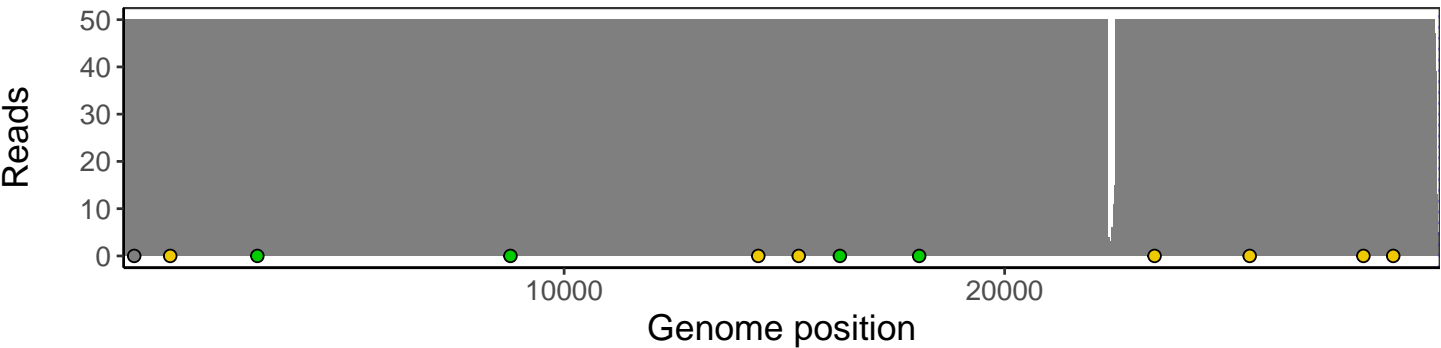
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



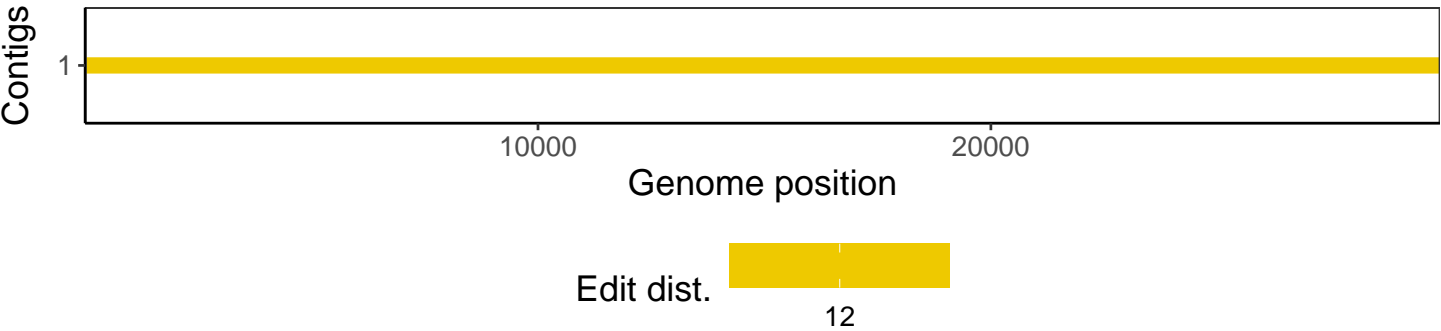
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



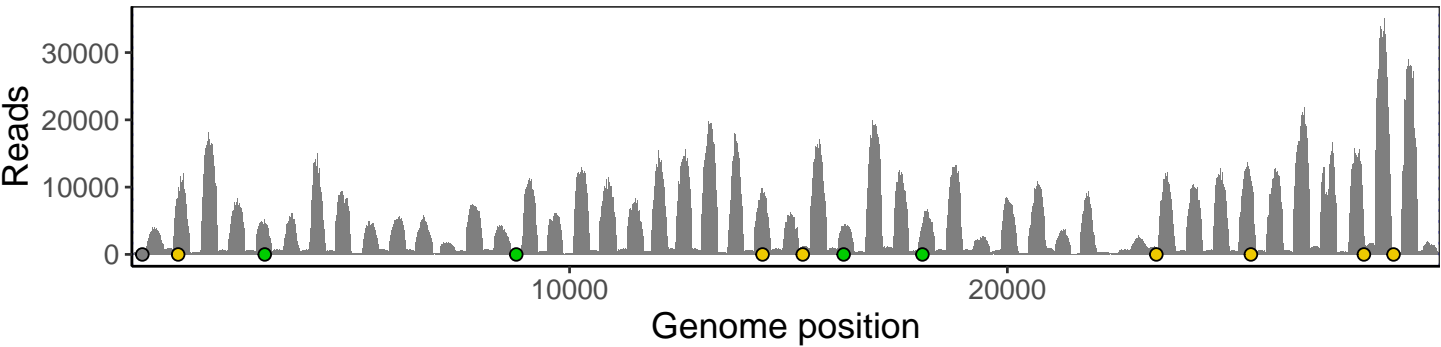
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



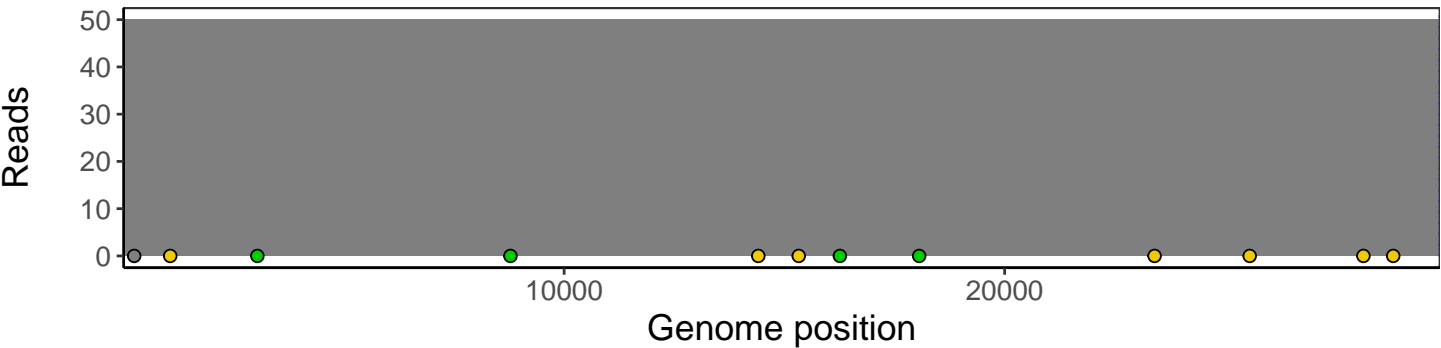
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



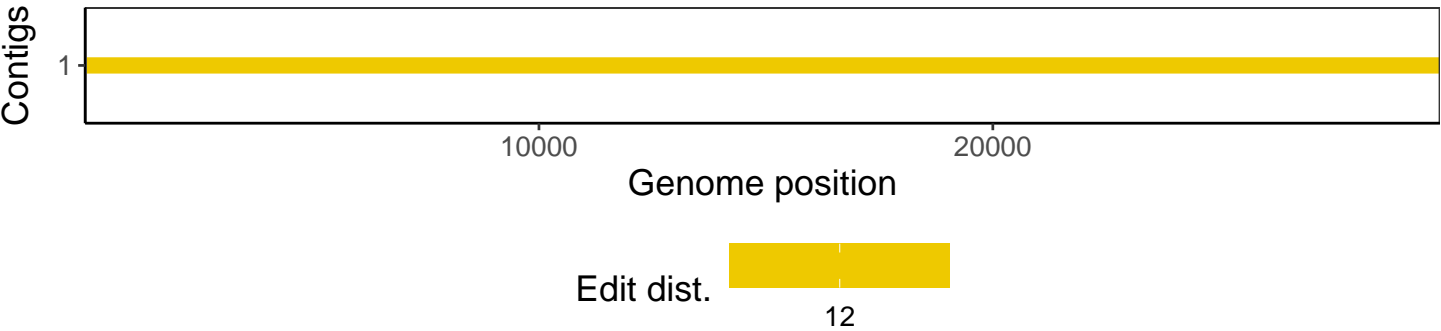
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



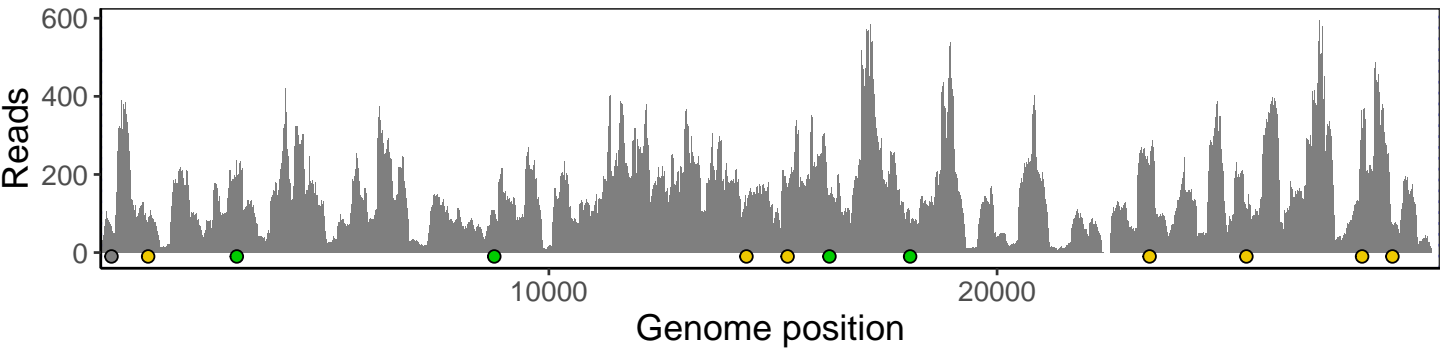
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



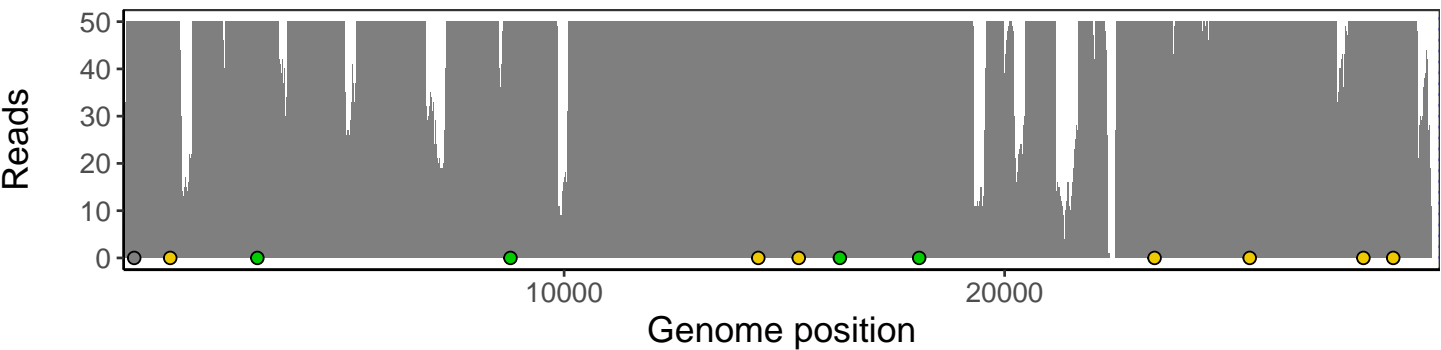
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



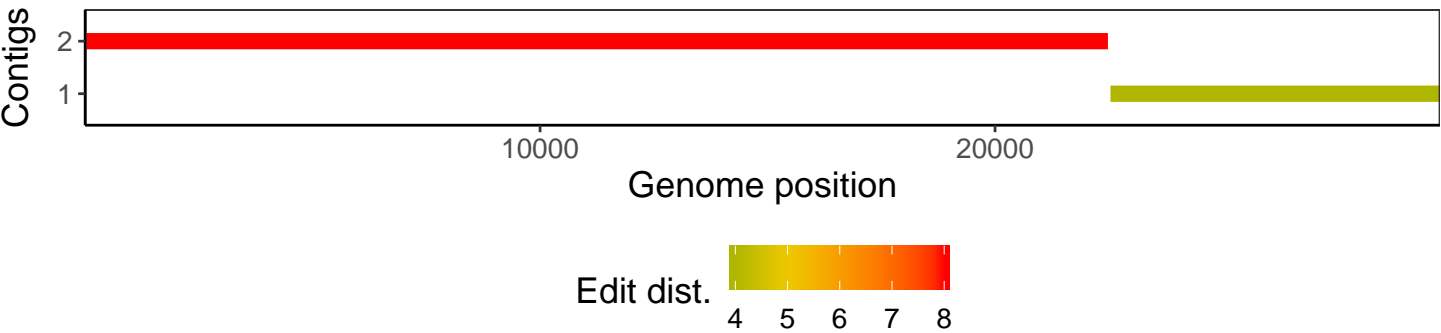
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



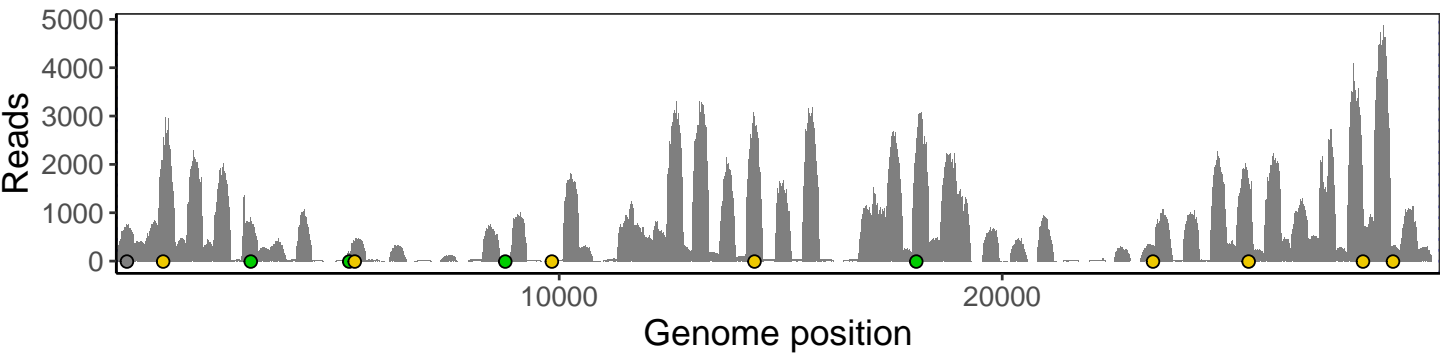
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



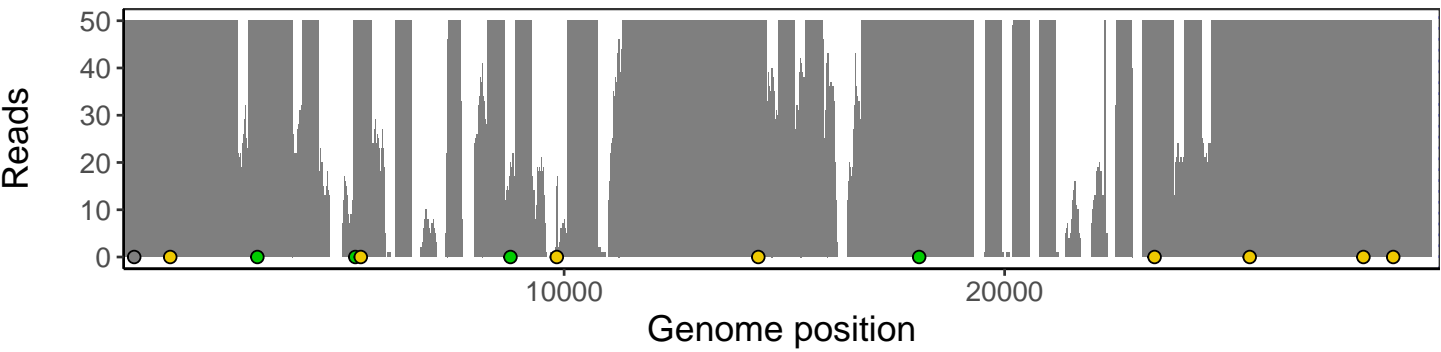
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



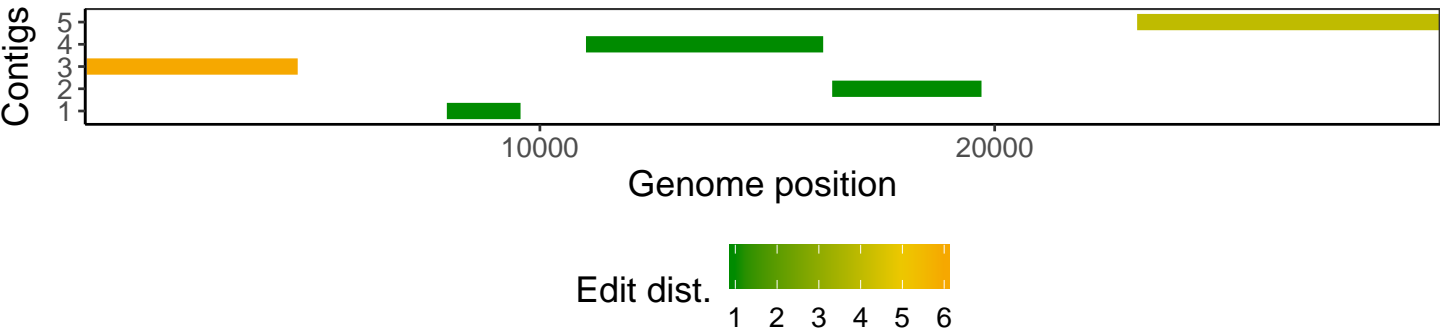
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



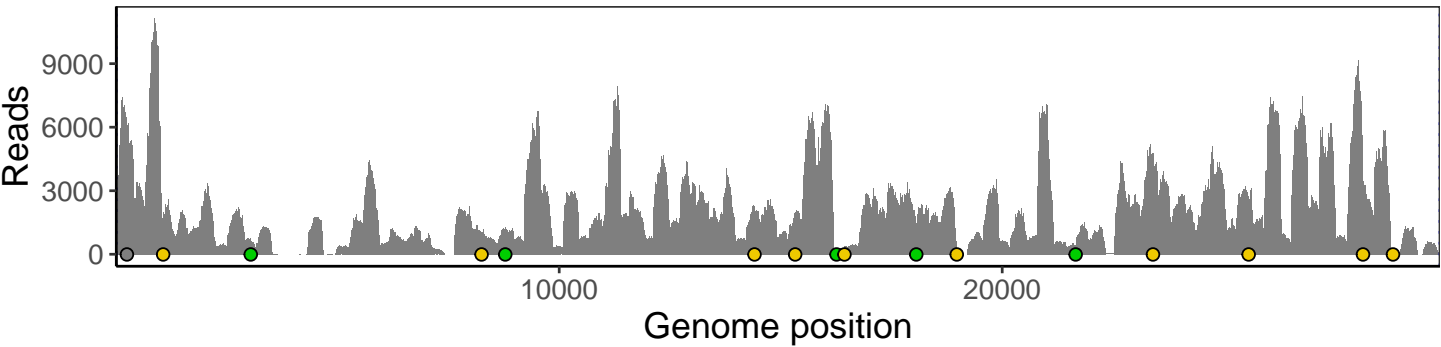
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



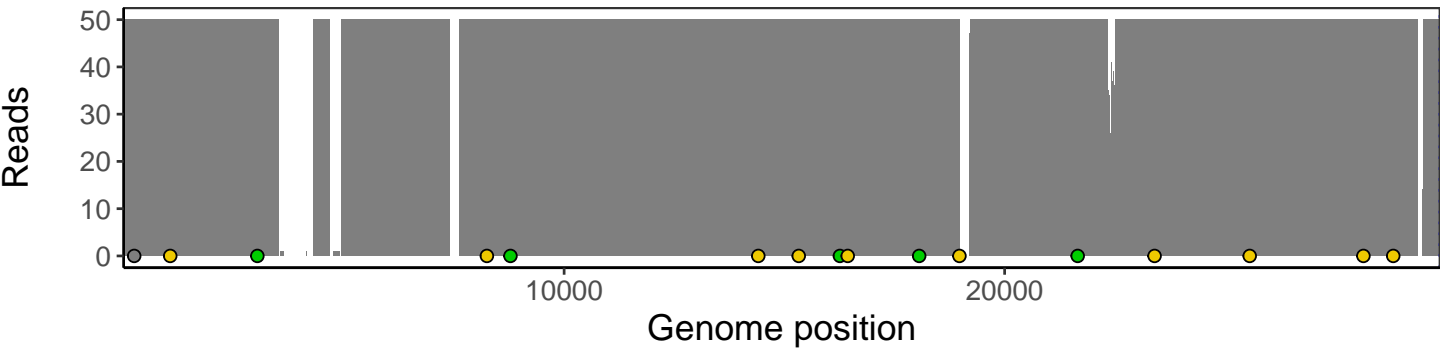
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



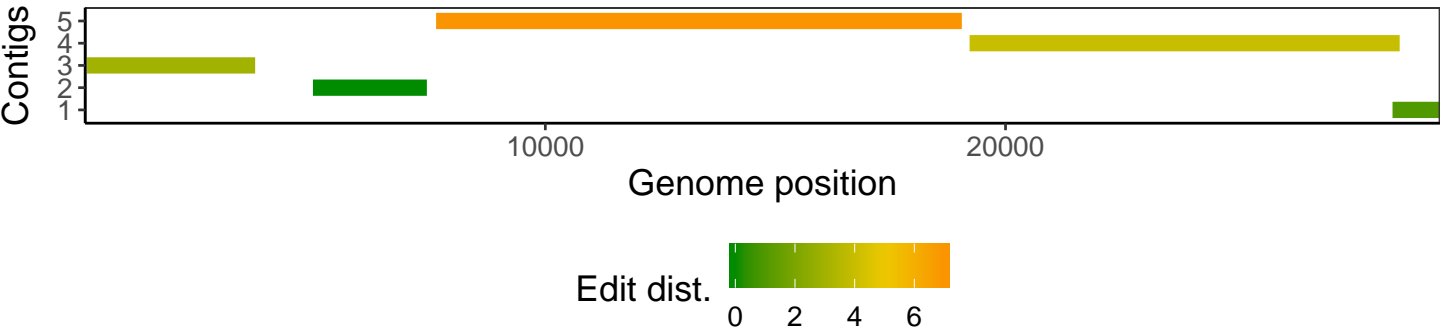
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



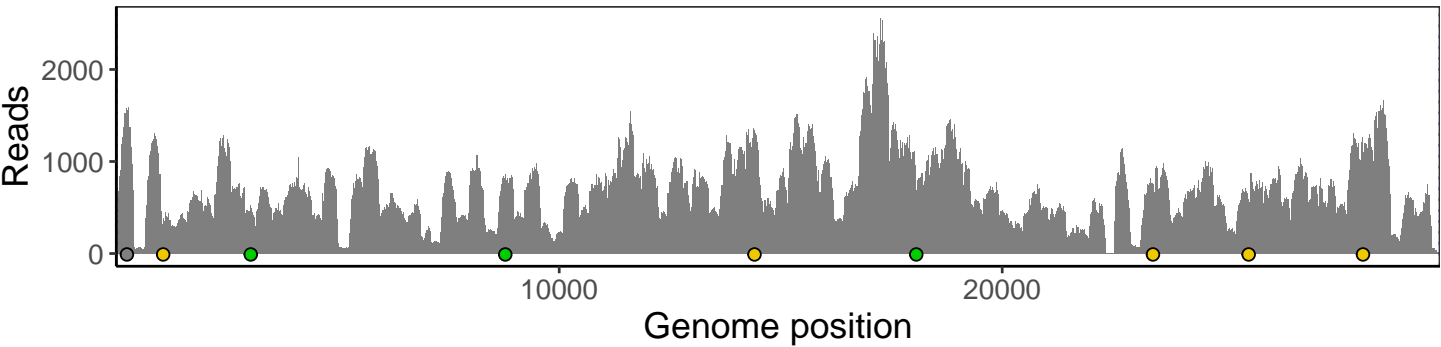
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



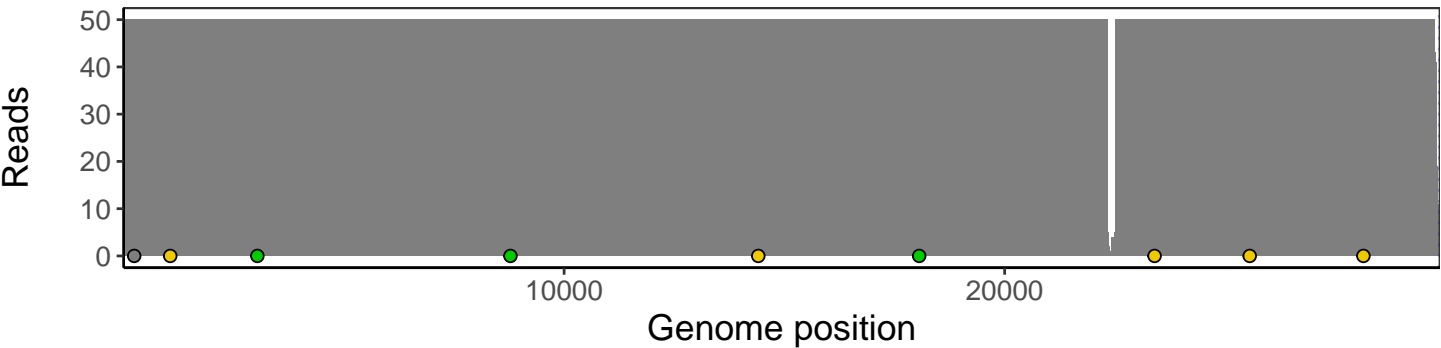
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



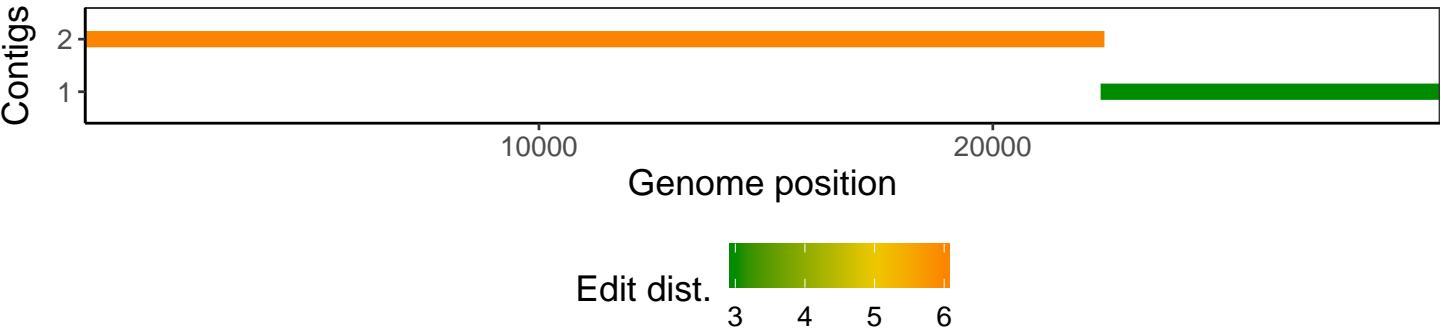
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



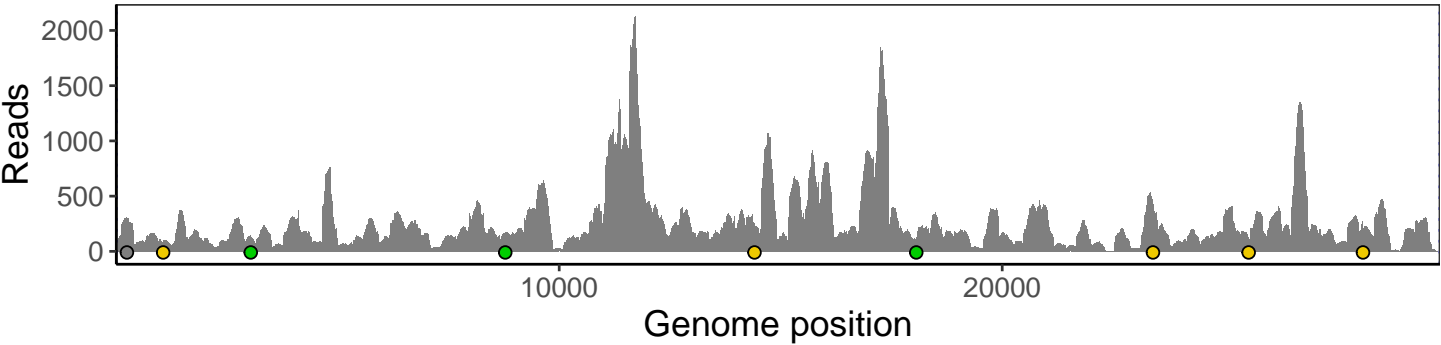
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



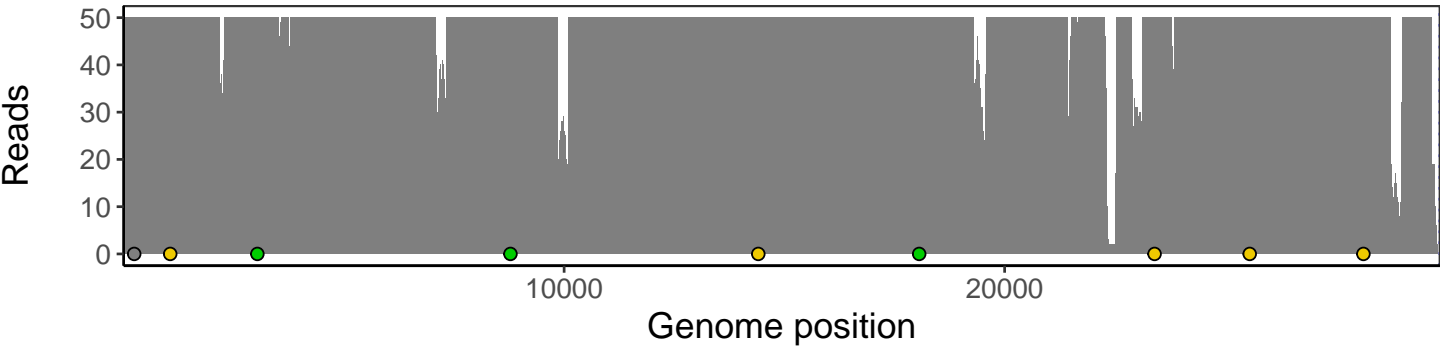
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



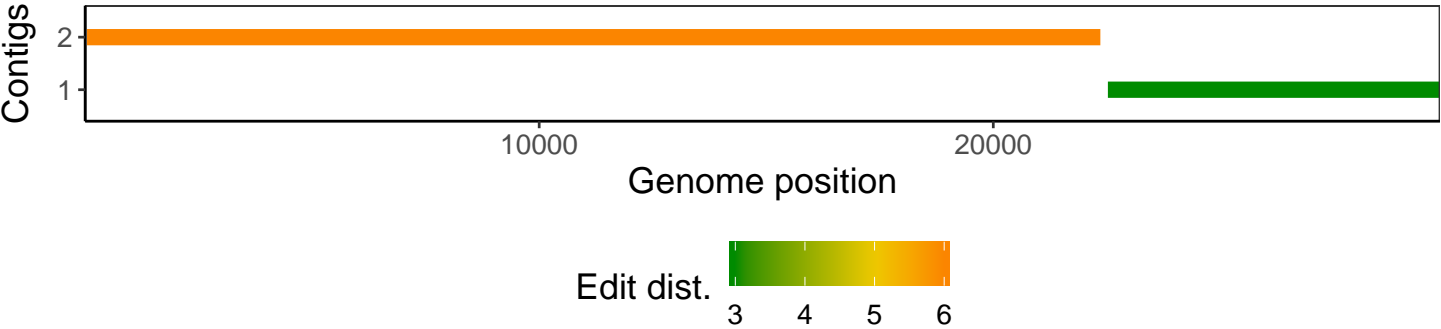
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



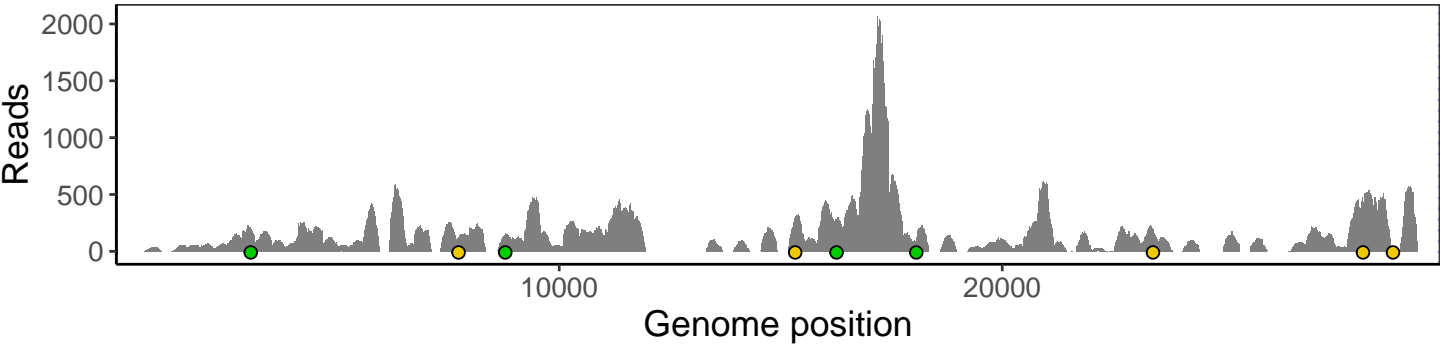
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



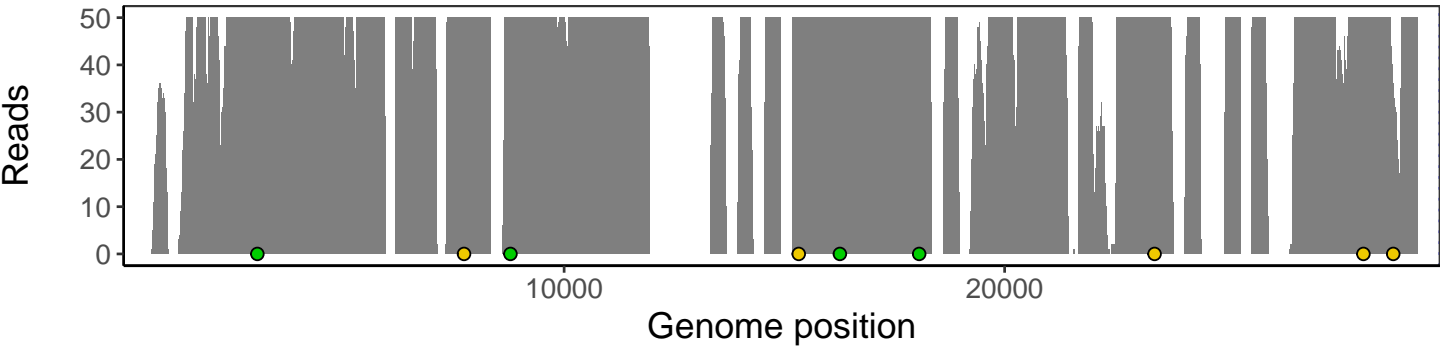
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

