

COVID-19 subject 269

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)
VSP0177	composite	NA	NP-OP	2020-05-27	4.03	90.0%	73.8%
VSP0166-1	single experiment	NA	NP-OP	2020-05-22	29.84	99.8%	99.5%
VSP0176-1	single experiment	166000.0	ETA	2020-05-27	29.82	99.8%	99.8%
VSP0177-1	single experiment	59.7	NP-OP	2020-05-27	1.03	72.7%	18.6%
VSP0177-2	single experiment	NA	NP-OP	2020-05-27	1.01	31.7%	26.0%
VSP0177-3	single experiment	NA	NP-OP	2020-05-27	1.02	40.8%	35.5%
VSP0177-4	single experiment	NA	NP-OP	2020-05-27	0.93	38.6%	35.6%
VSP0200-1	single experiment	NA	ETA	2020-06-08	5.92	84.6%	82.1%
VSP0202-1	single experiment	16300.0	ETA	2020-06-12	29.83	99.8%	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 020-05-2	ETA 020-06-0	ETA 020-06-1	NP-OP 020-05-2	NP-OP 2020-05-27			
241 intergenic	916	213	1577	963	2			
920 orf1ab silent	1237	89	1063	976	2		52	
1059 orf1ab T265I	1012	50	1359	631	4	24	7	
2903 orf1ab I880V	846	2	961	1229	144	27	15	34
3037 orf1ab silent	936		948	1489	3	33	13	36
5260 orf1ab silent	494	22	931	170	13			
8782 orf1ab silent	710	104	2234	463			24	
11118 orf1ab I3618T	958	157	2958	1214	4	19	56	
11522 orf1ab del 1	1035	563	2921	2219	8		61	
13668 orf1ab F67T	652	133	1592	1261	5	31	7	1
13813 orf1ab G116P	1398	393	3103	2206	2	87		
14408 orf1ab P314L	1278	230	2101	992	2			
16260 orf1ab silent	1229	36	1748	1643	1			67
18060 orf1ab F1531L	667	65	1883	660		38		46
18060 orf1ab silent	667	65	1883	660		38		46
19904 orf1ab del 2	1288	38	2873	1196	1	26		47
23403 S D614G	3867	440	5079	2188	5	84	62	
23403 S M614G	3867	440	5079	2188	5	84	62	
23403 S silent	3867	440	5079	2188	5	84	62	
24026 S del 3	474		923	801	1			
25563 orf3a Q57H	838	77	2672	1164	6			
25563 orf3a R57H	838	77	2672	1164	6			
28144 orf8 H84L	1218	27	2052	3198	1			
28144 orf8 S84L	1218	27	2052	3198	1			
28821 N L183Y	687	25	952	890				
28821 N S183Y	687	25	952	890				
29247 N H325I	1169	158	1785	2263				
29247 N T325I	1169	158	1785	2263				
29344 N L357I	927	88	1417	1582	1			
29344 N silent	927	88	1417	1582	1			
	VSP0176-1	VSP0200-1	VSP0202-1	VSP0166-1	VSP0177-1	VSP0177-2	VSP0177-3	VSP0177-4

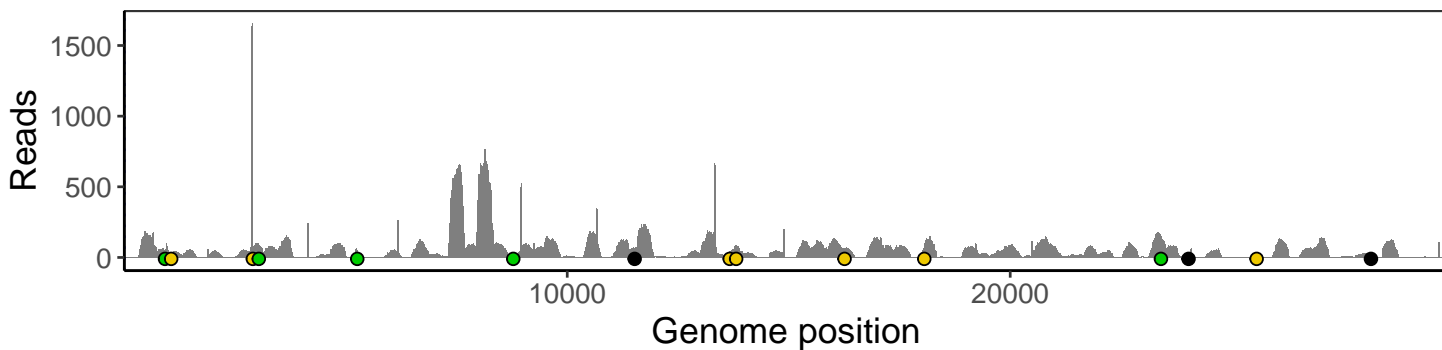
Base change

- Expected
- A
- T
- C
- G
- N
- Ins/Del
- No data

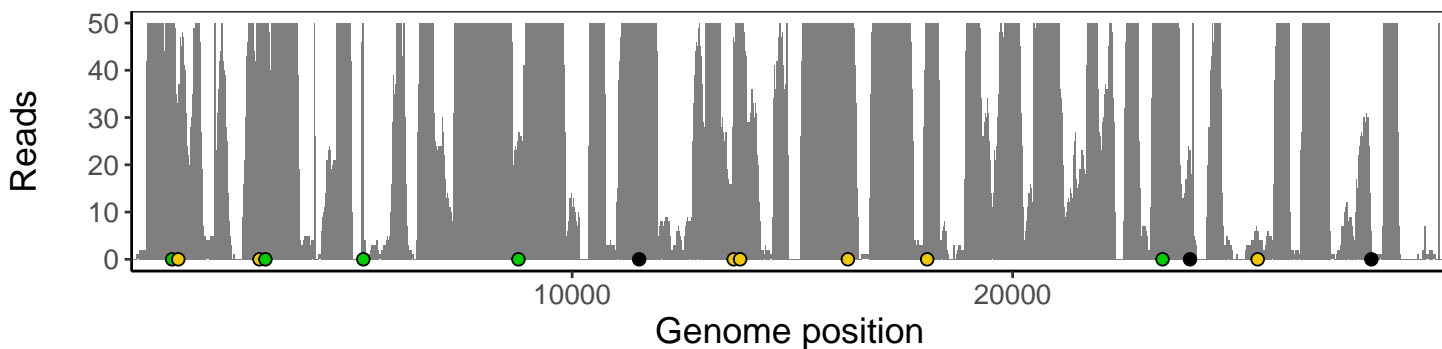
Analyses of individual experiments and composite results.

VSP0177 | 2020-05-27 | NP-OP | 269no-q | composite result

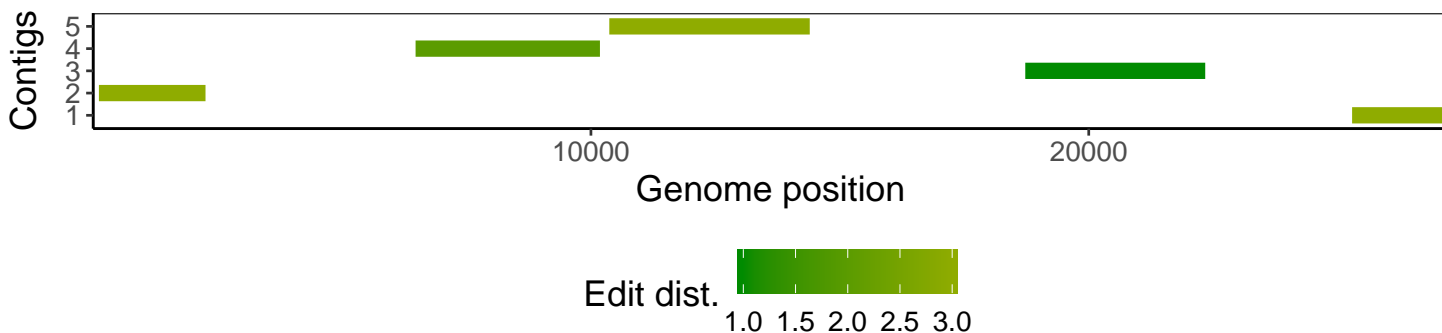
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



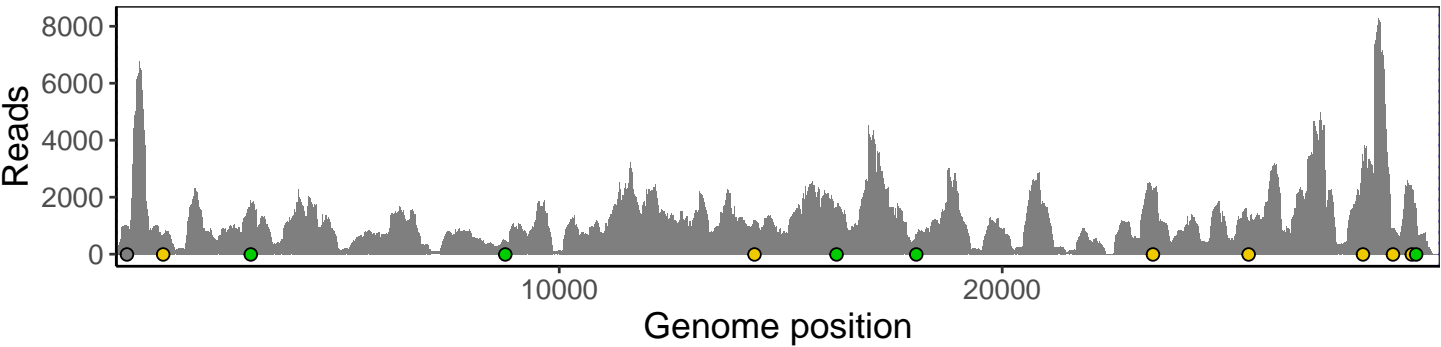
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



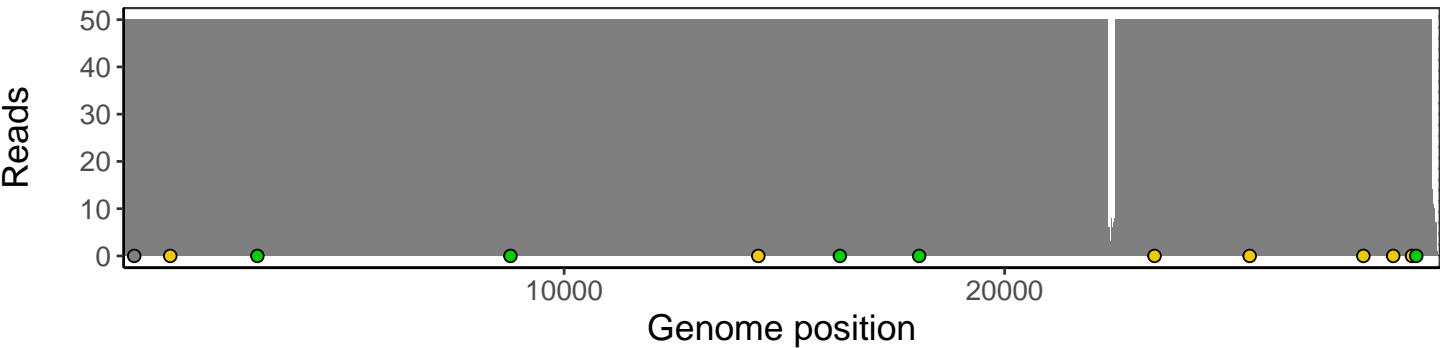
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



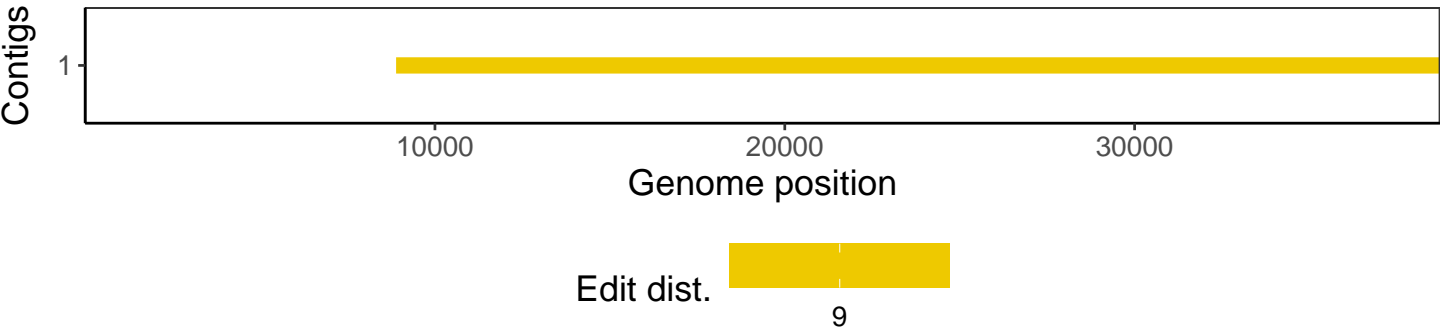
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



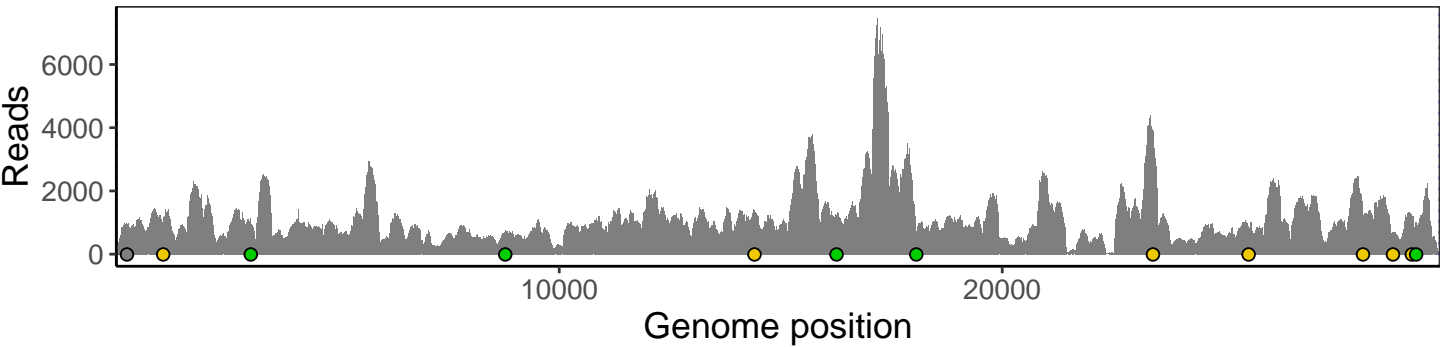
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



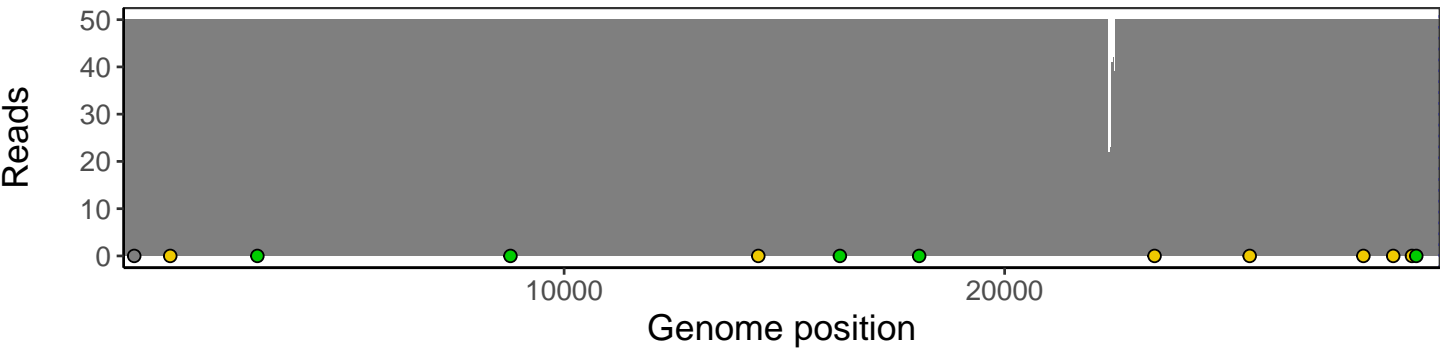
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



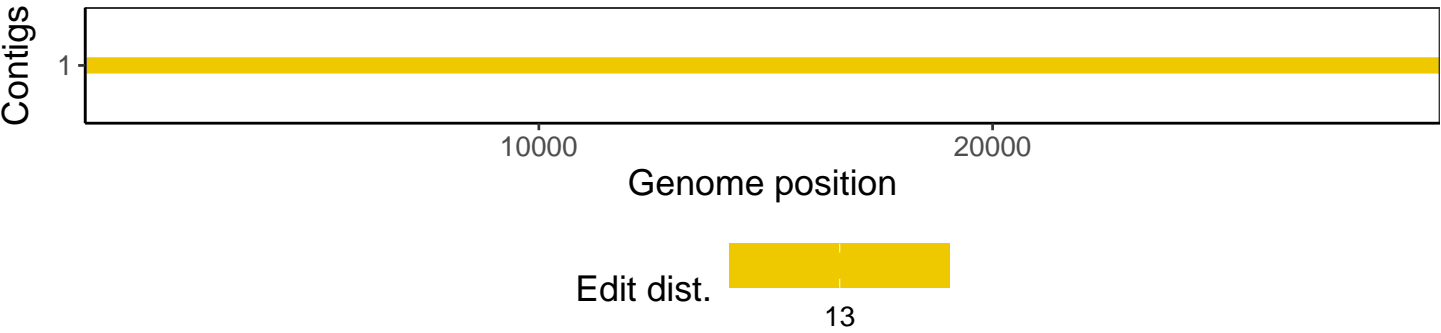
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



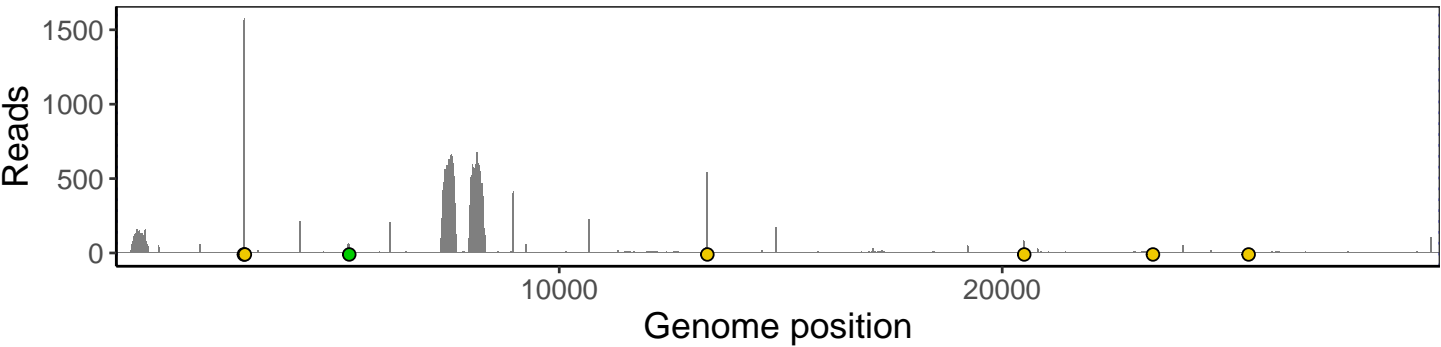
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



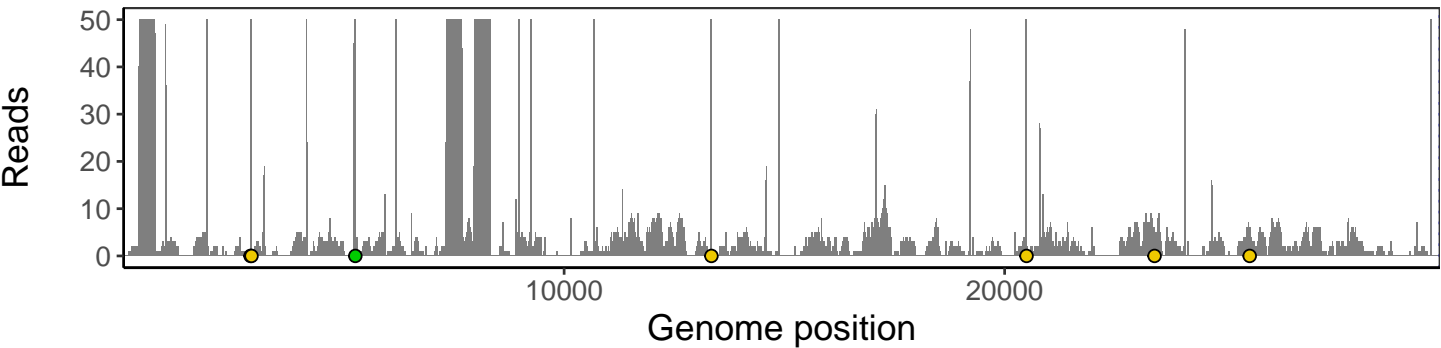
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



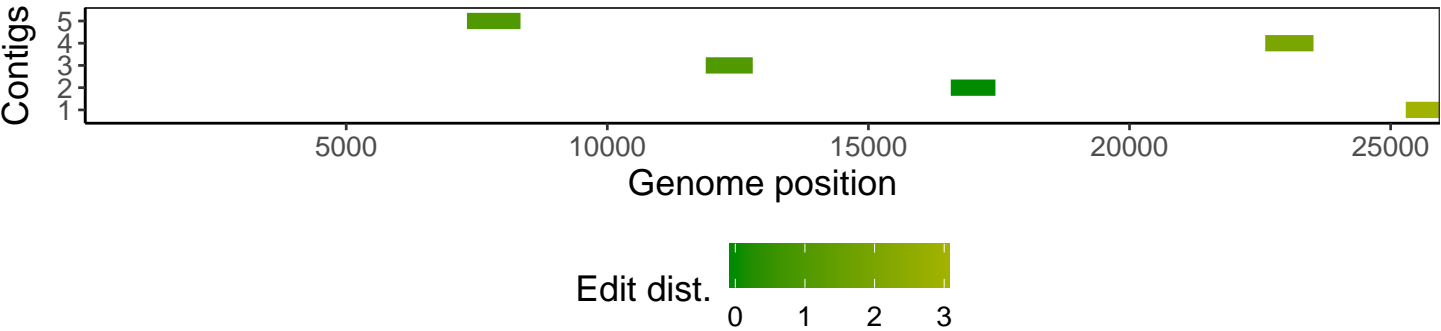
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



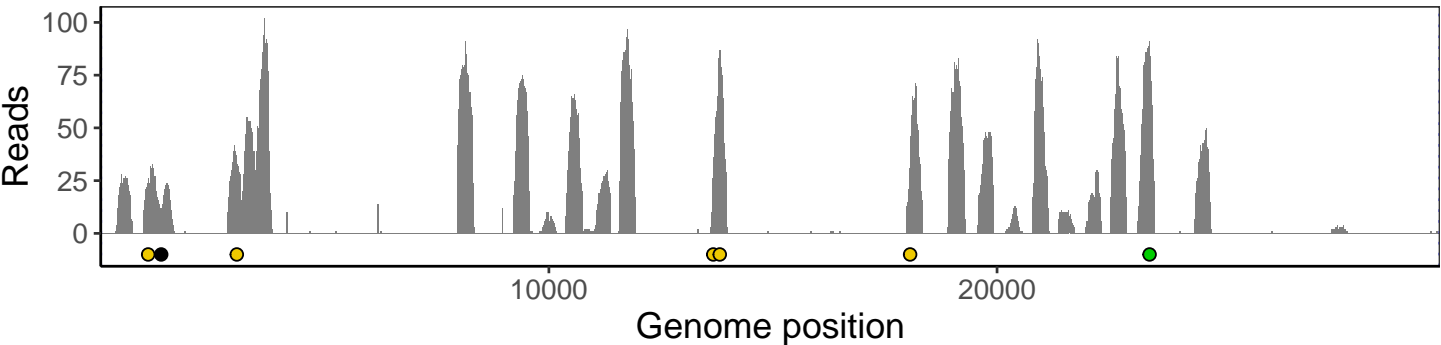
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



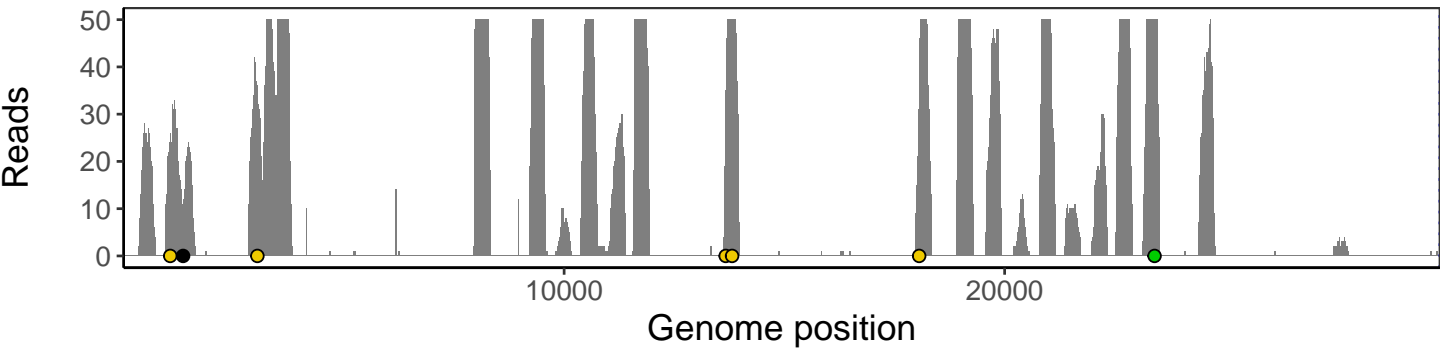
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.



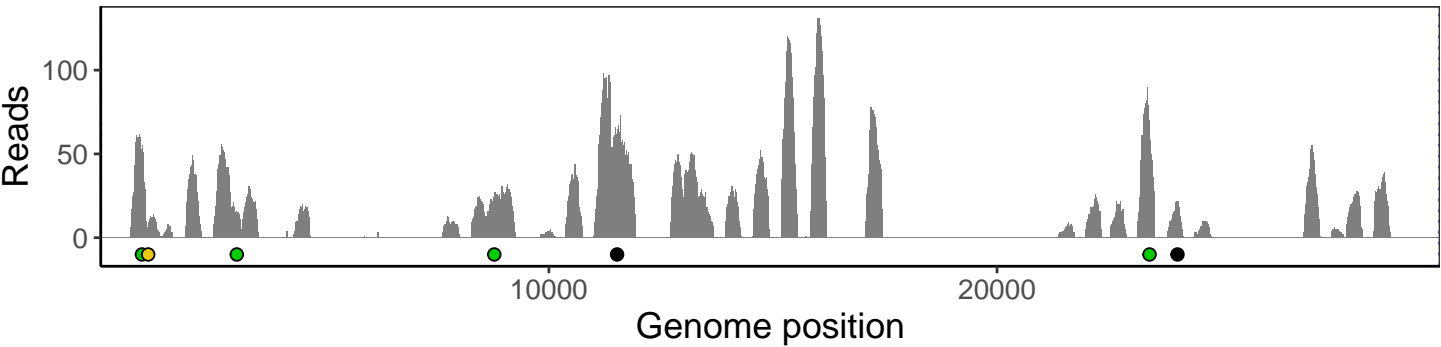
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



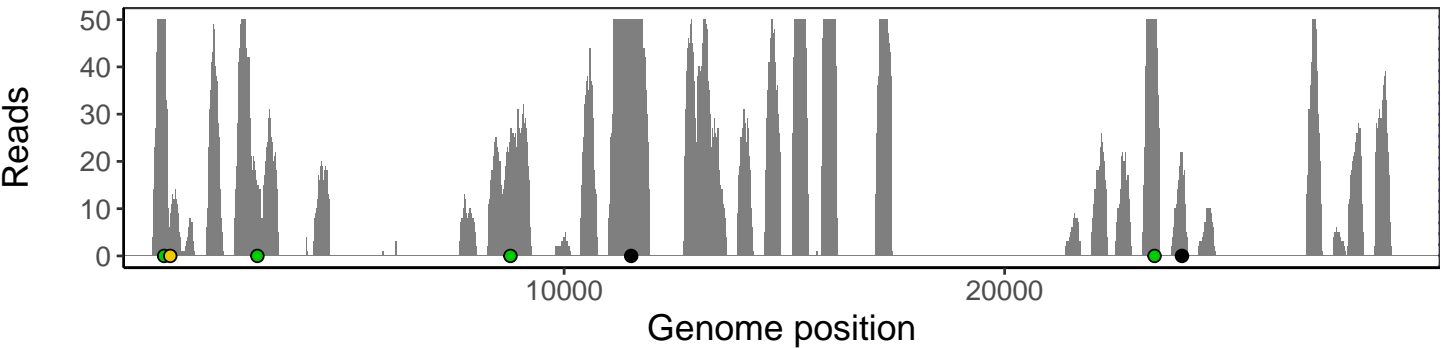
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.



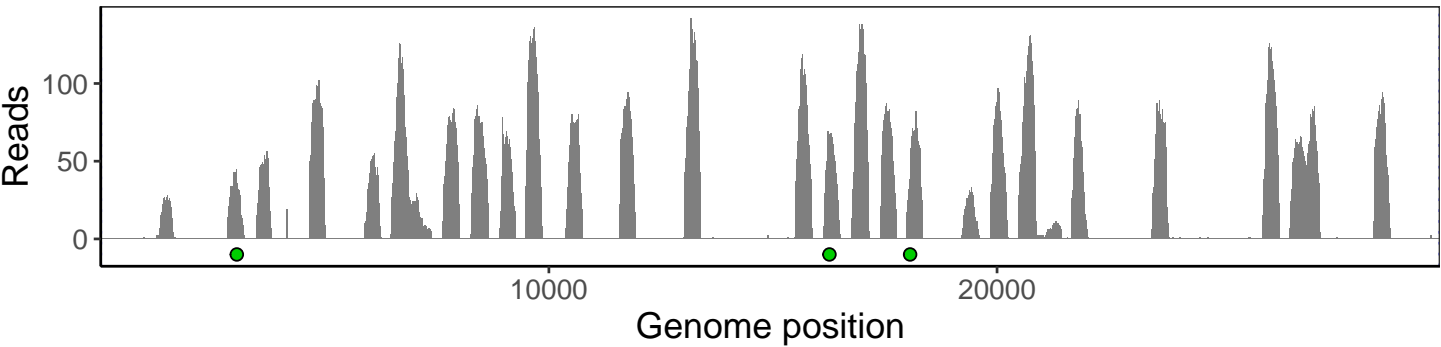
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



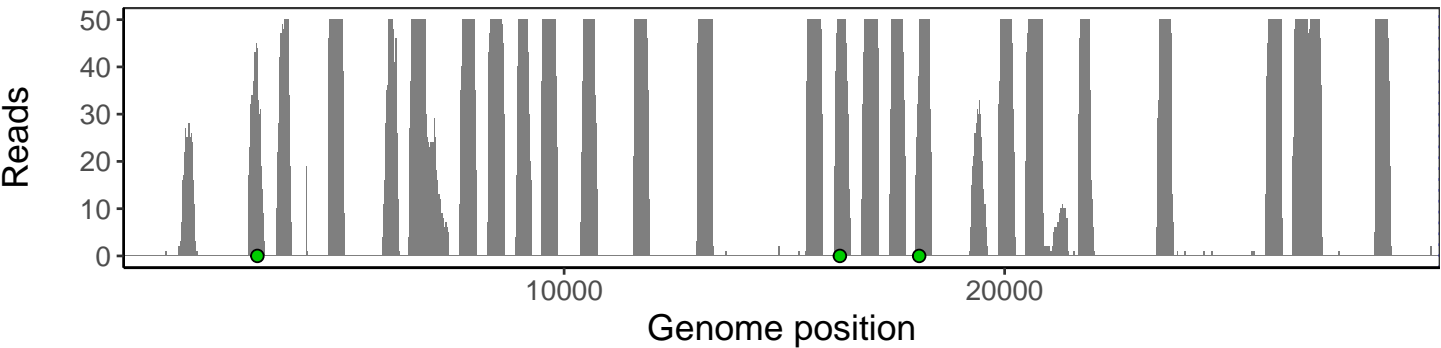
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



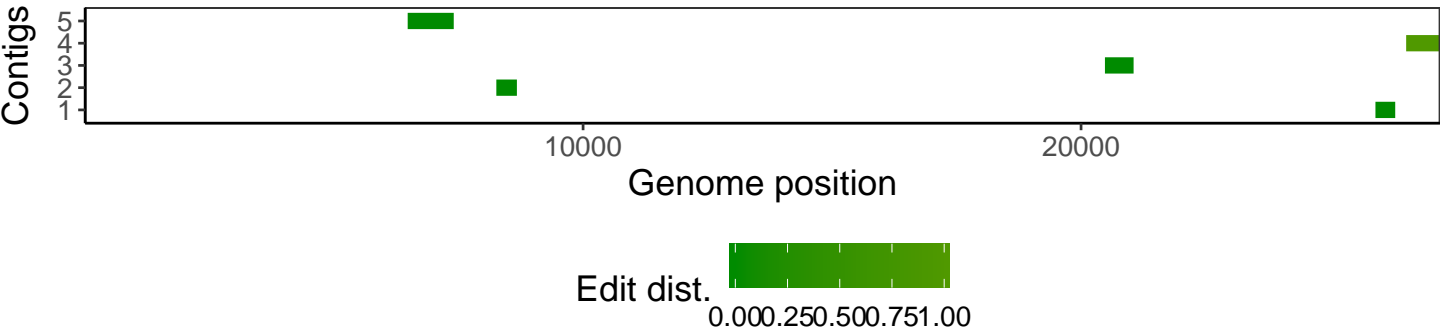
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



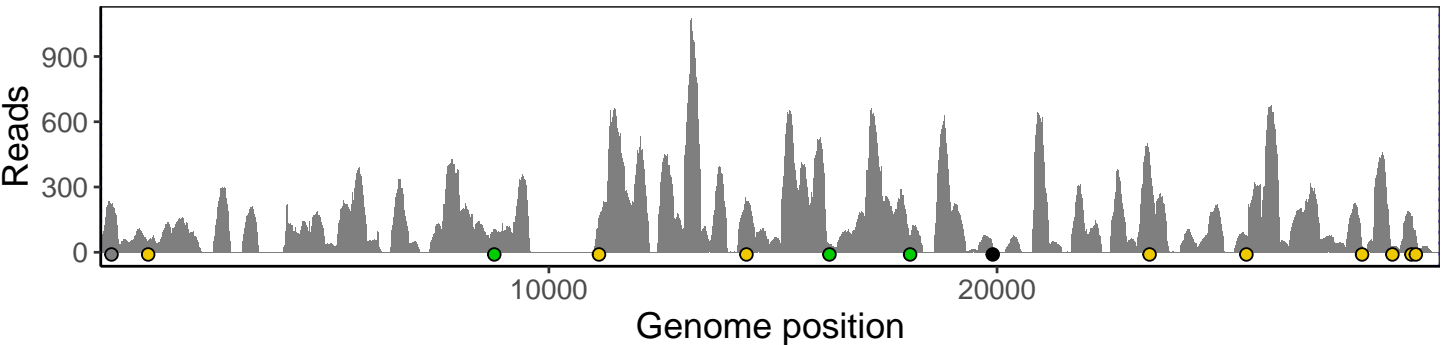
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



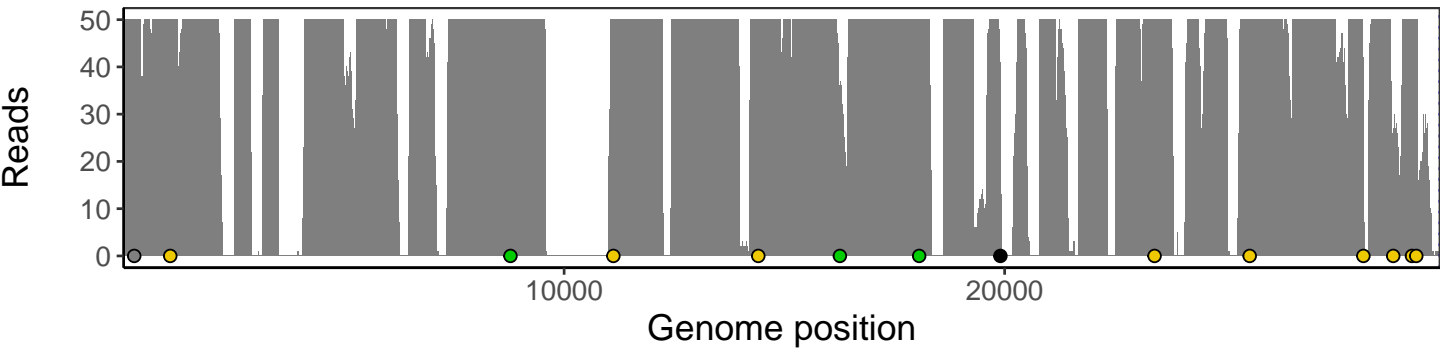
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



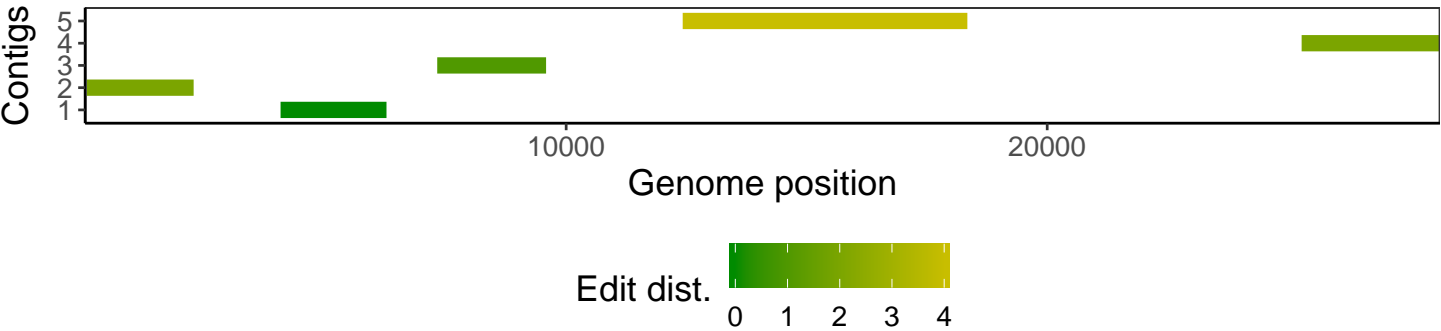
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



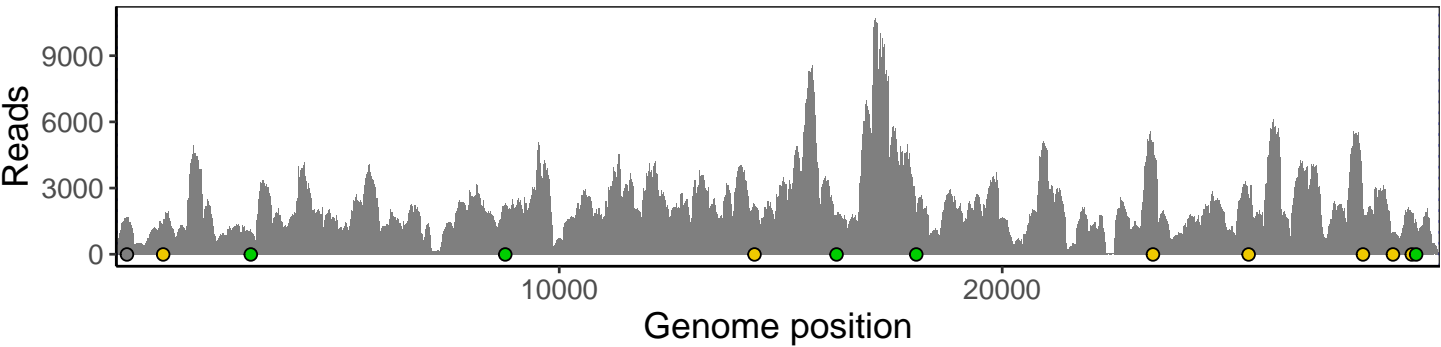
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



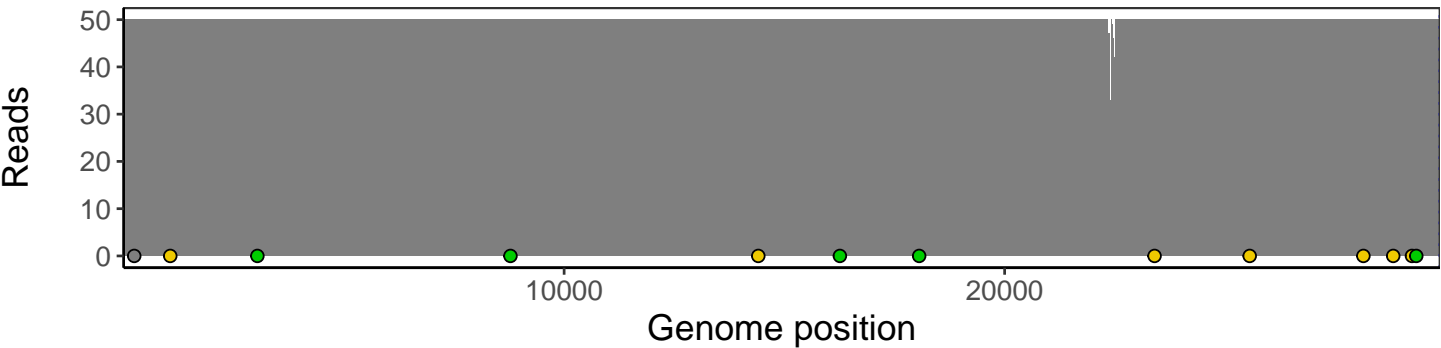
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

