

# COVID-19 subject 228

2020-08-28

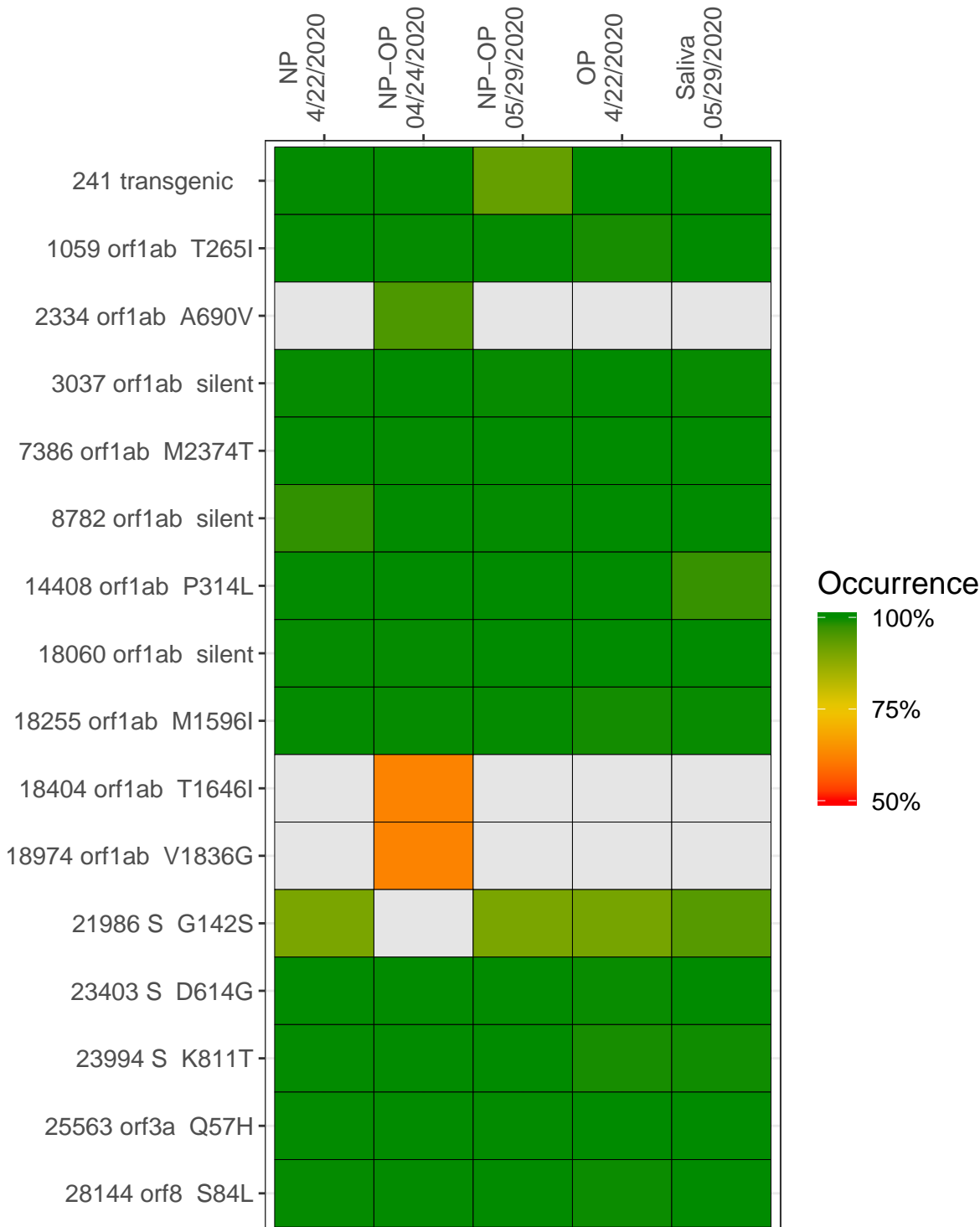
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0021	composite	NA	NP	4/22/2020	29.82	99.8%	99.8%
VSP0022	composite	NA	OP	4/22/2020	29.68	99.8%	98.8%
VSP0069	composite	NA	NP-OP	04/24/2020	18.75	98.5%	98.5%
VSP0187	composite	NA	ETA	05/29/2020	NA	NA	NA
VSP0188	composite	NA	NP-OP	05/29/2020	29.90	99.8%	99.8%
VSP0021-1a	single experiment	2.20e+03	NP	4/22/2020	0.52	64.8%	61.9%
VSP0021-1b	single experiment	2.20e+03	NP	4/22/2020	0.57	63.0%	62.1%
VSP0021-2	single experiment	1.10e+04	NP	4/22/2020	29.82	99.8%	99.8%
VSP0022-1a	single experiment	2.68e+05	OP	4/22/2020	1.75	78.3%	37.7%
VSP0022-2	single experiment	1.34e+06	OP	4/22/2020	0.24	6.1%	0.0%
VSP0022-3	single experiment	1.34e+06	OP	4/22/2020	29.68	99.8%	98.8%
VSP0069-1	single experiment	9.75e+01	NP-OP	04/24/2020	6.54	92.2%	88.2%
VSP0069-2	single experiment	9.75e+01	NP-OP	04/24/2020	9.29	90.5%	90.0%
VSP0069-3	single experiment	9.75e+01	NP-OP	04/24/2020	6.37	87.1%	85.5%
VSP0187-1	single experiment	6.02e+01	ETA	05/29/2020	NA	NA	NA
VSP0187-2	single experiment	3.01e+02	ETA	05/29/2020	NA	NA	NA
VSP0188-1	single experiment	2.04e+03	NP-OP	05/29/2020	9.60	93.7%	93.5%
VSP0188-2	single experiment	1.02e+04	NP-OP	05/29/2020	22.61	99.1%	99.1%
VSP0188-3	single experiment	1.02e+04	NP-OP	05/29/2020	29.87	99.8%	99.8%
VSP0189-1	single experiment	8.51e+04	Saliva	05/29/2020	29.82	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.



	NP 4/22/2020			NP-OP 04/24/2020			NP-OP 05/29/2020			OP 4/22/2020			Saliva 05/29/2020		
241 transgenic	655		3341	1877	1400		878	2357	2035	6		396	497		
1059 orf1ab T265I		485	1348	4	537	3971	998	958	1255	2		153	293		
2334 orf1ab A690V	1	752	1340	235		4862		562	281	1		149	296		
3037 orf1ab silent		694	2447	89	753		763	1076	2758			517	686		
7386 orf1ab M2374T	317	47	1260	2351	1298	1625	1370	170	547	4		257	311		
8782 orf1ab silent	322		2672	4828		2747	867	721	3292			239	318		
14408 orf1ab P314L		1022	5052	1383	1934	6921	1197	3331	3058	4		161	1315		
18060 orf1ab silent		622	2638	940	928	3863	1364	1001	2850			197	471		
18255 orf1ab M1596I		555	3751	922	881	4339	1433	1042	3833			200	568		
18404 orf1ab T1646I	403		5285	2048	130	4223	984	952	1427	8		256	303		
18974 orf1ab V1836G	387	2	2768	2	27	8	518	1596	2796	3		753	244		
21986 S G142S	64	197	577	1826	719	826	209	370	509			53	249		
23403 S D614G	622		4502	4973	7106	3	5603	7958	11305	33		501	1354		
23994 S K811T	667		1963		1206	2730	1406	1112	830	9	1	292	263		
25563 orf3a Q57H		421	3511	1	4082	8363	1679	2091	3061	4		311	981		
28144 orf8 S84L	509	146	4436	3320	1707	2570	1815	2352	1768	16		349	550		
	VSP0021-1a	VSP0021-1b	VSP0021-2	VSP0069-1	VSP0069-2	VSP0069-3	VSP0188-1	VSP0188-2	VSP0188-3	VSP0022-1a	VSP0022-2	VSP0022-3	VSP0189-1		

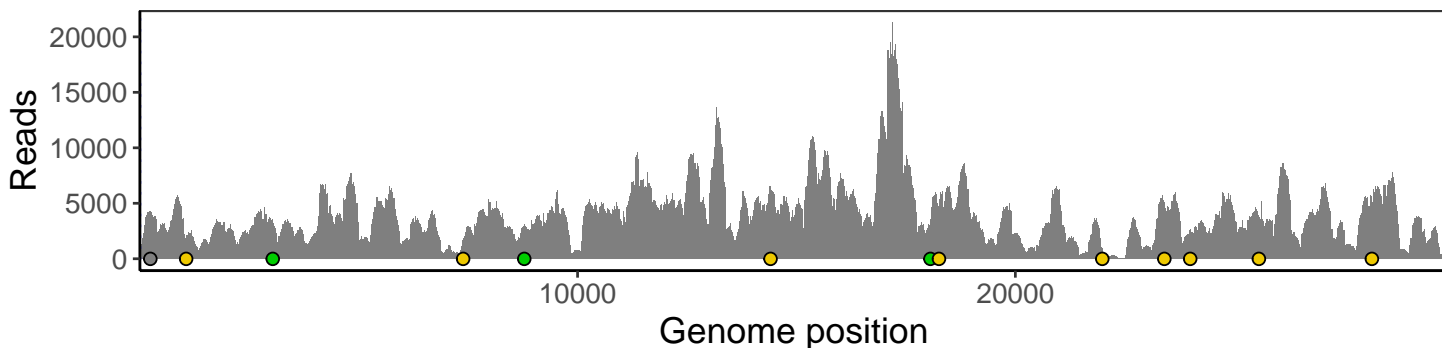
Base change



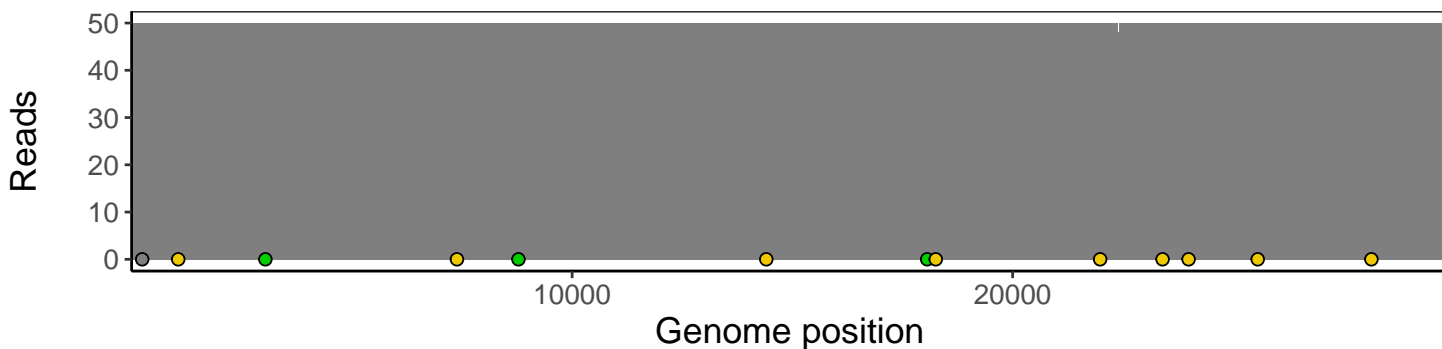
## Analyses of individual experiments and composite results.

VSP0021 | 4/22/2020 | NP | 228n | composite result

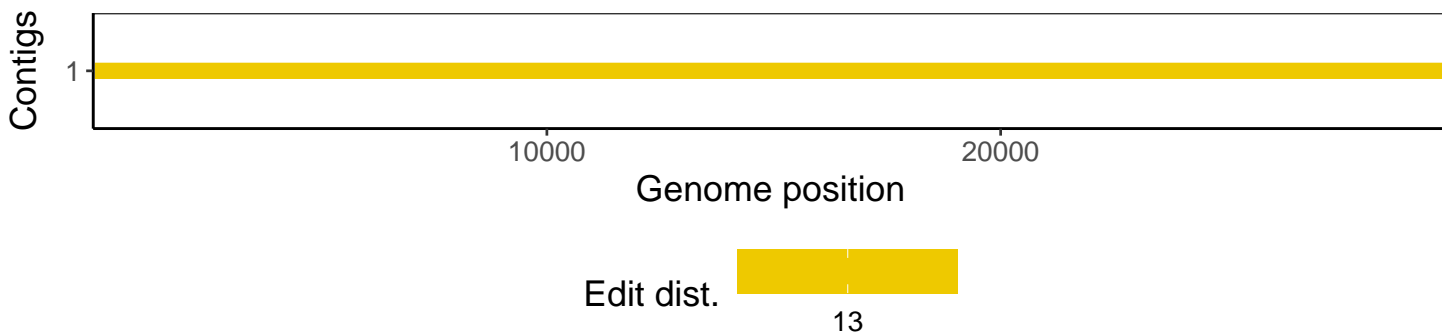
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



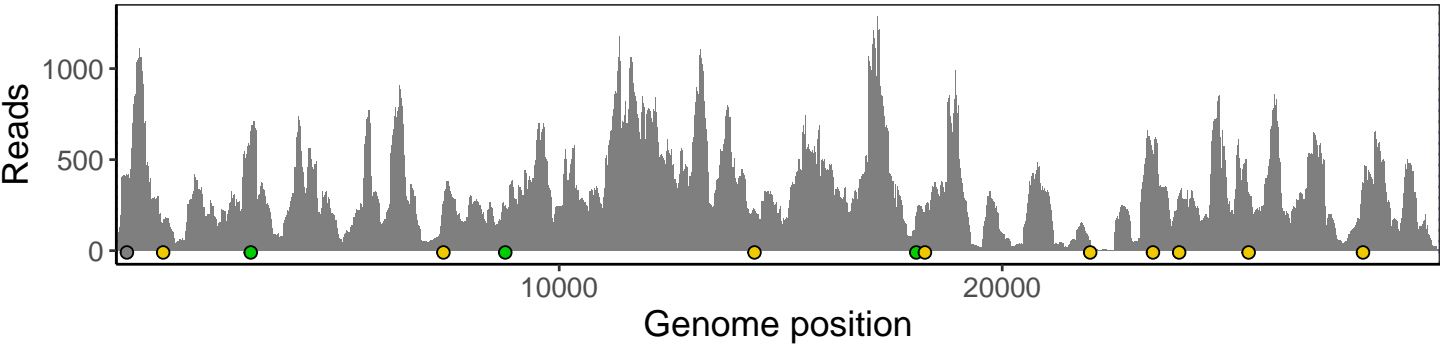
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



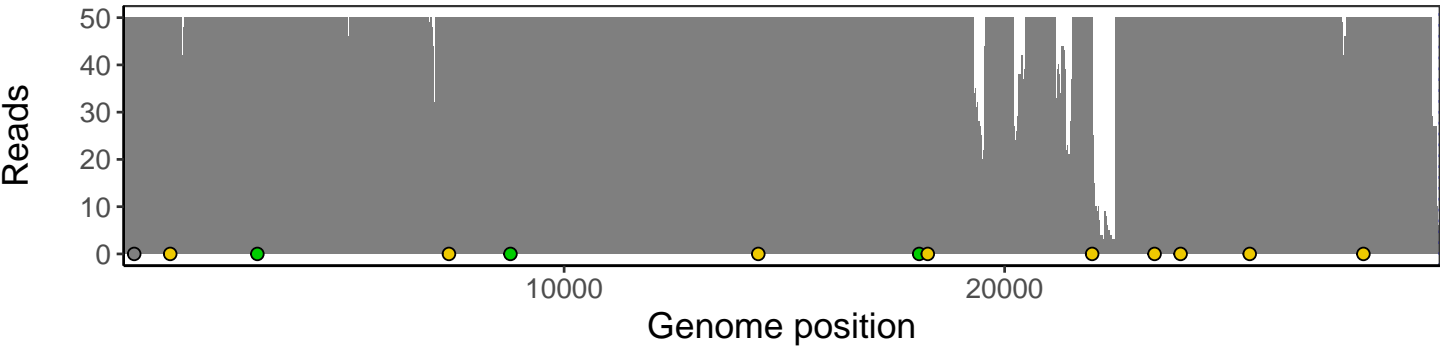
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



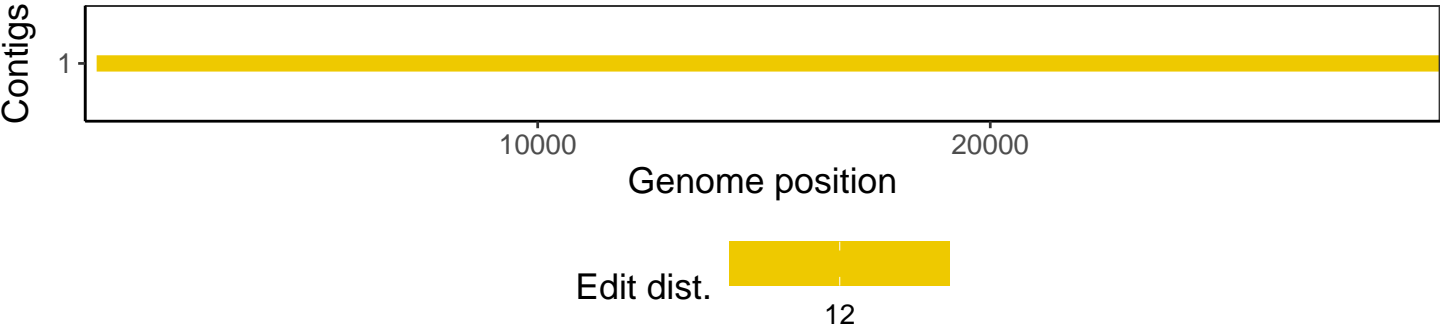
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



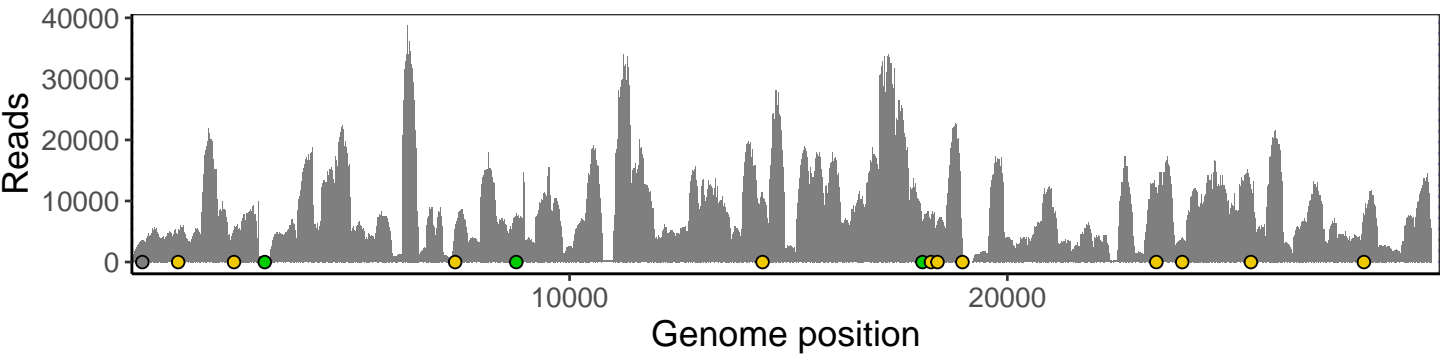
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



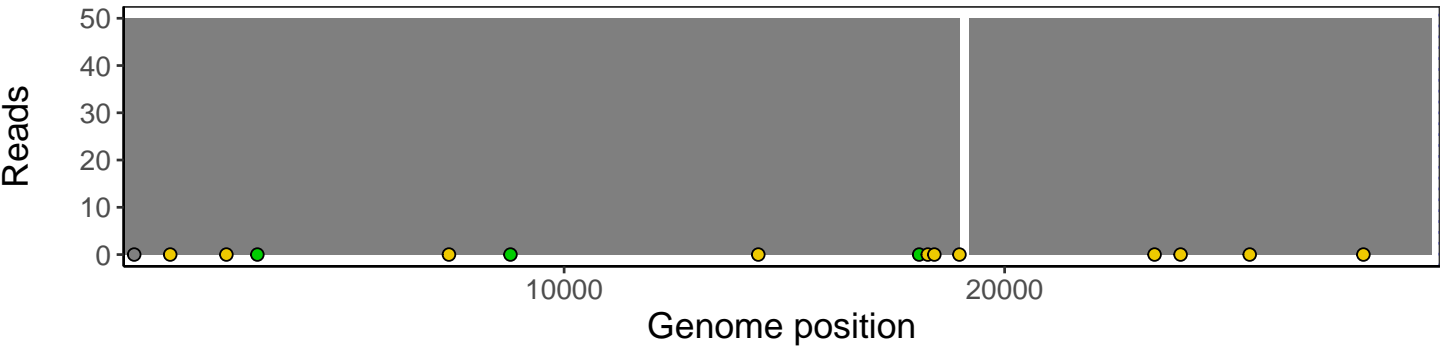
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



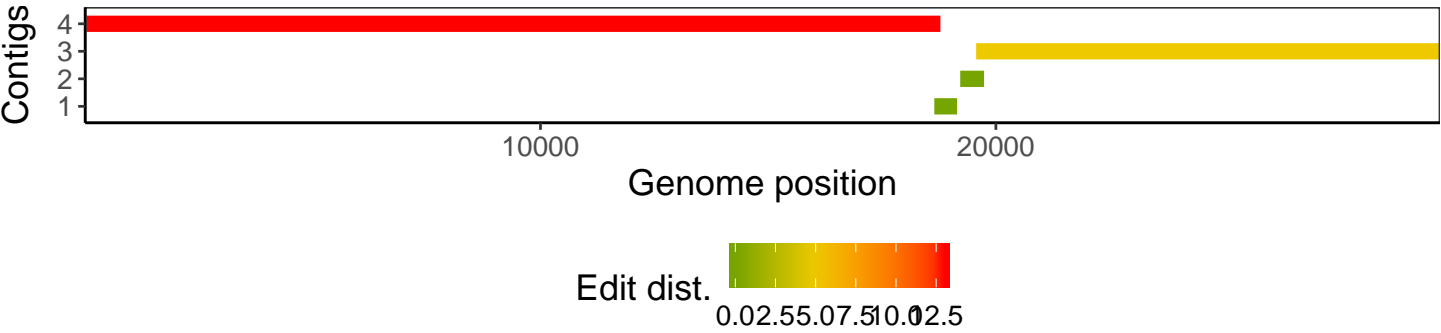
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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.

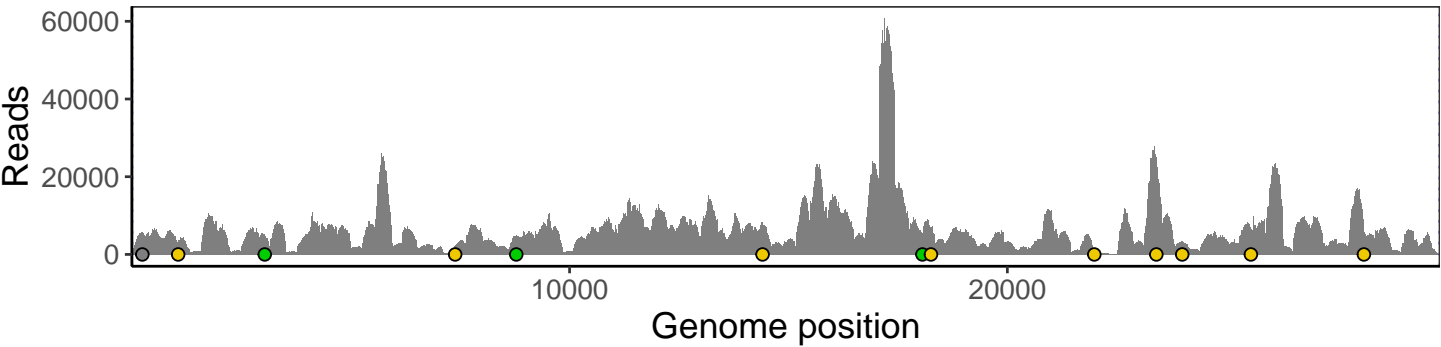


VSP0187 | 05/29/2020 | ETA | 228e-q | composite result

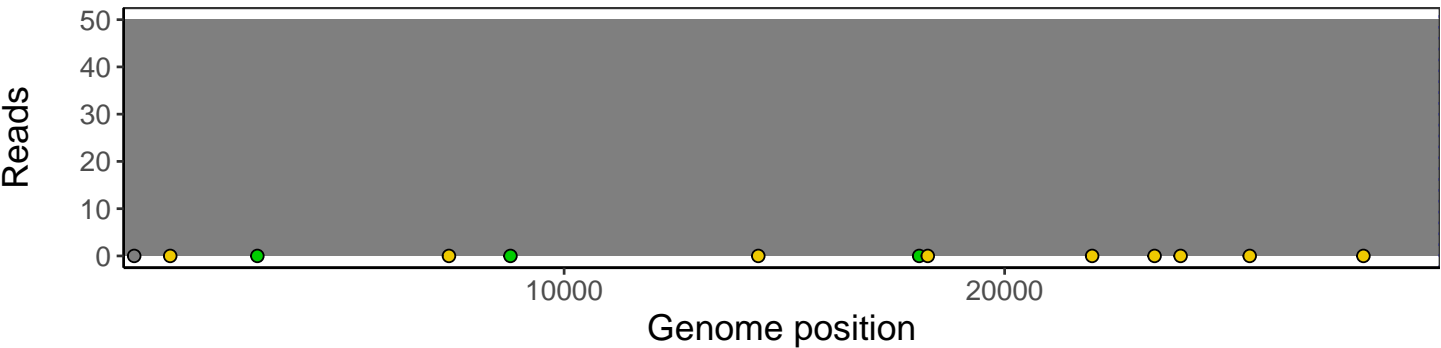
No pileup data available.

No contig data available.

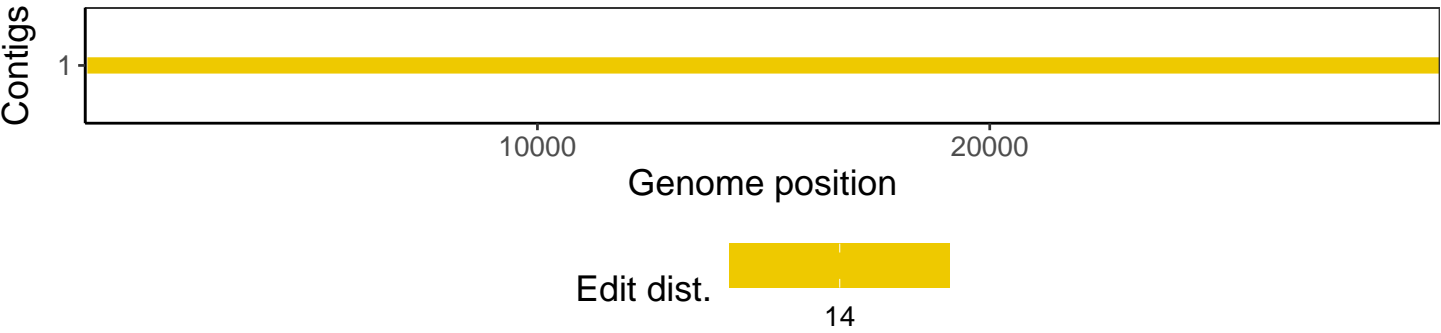
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

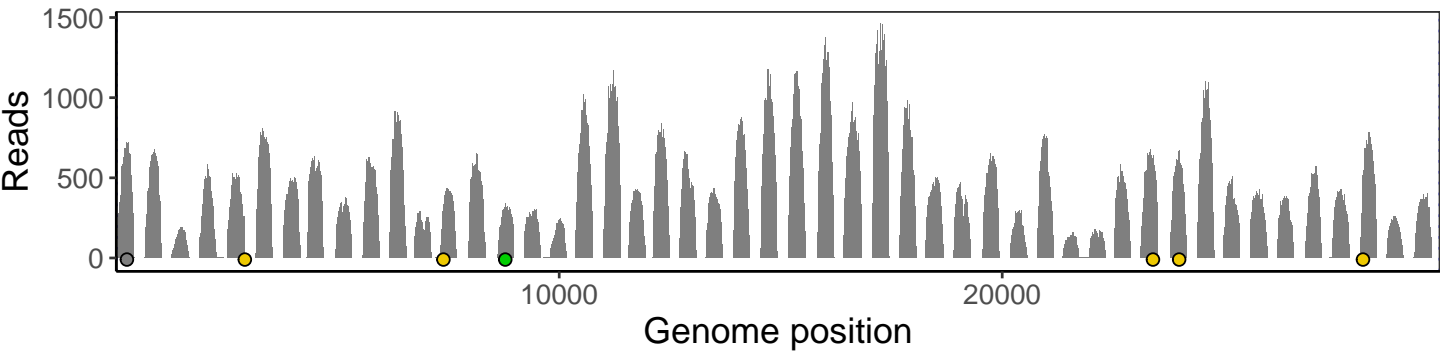


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

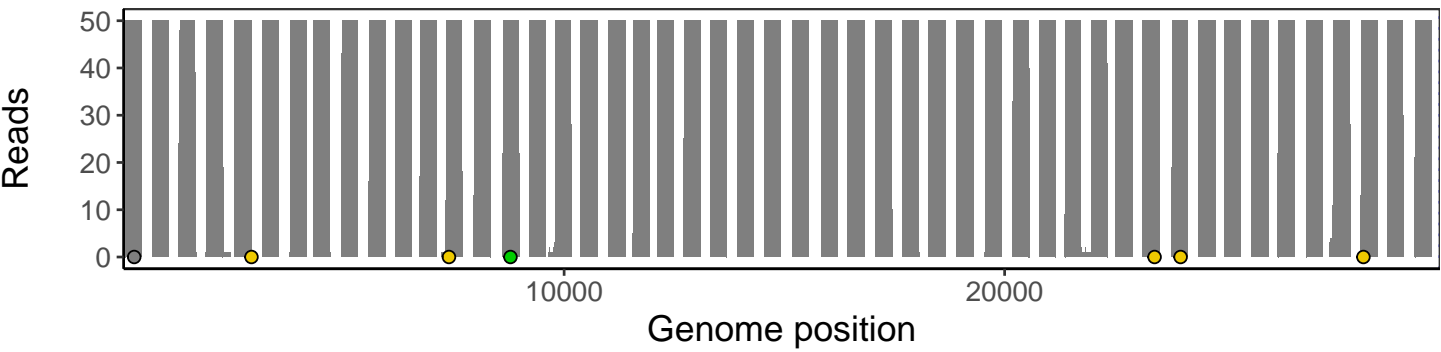




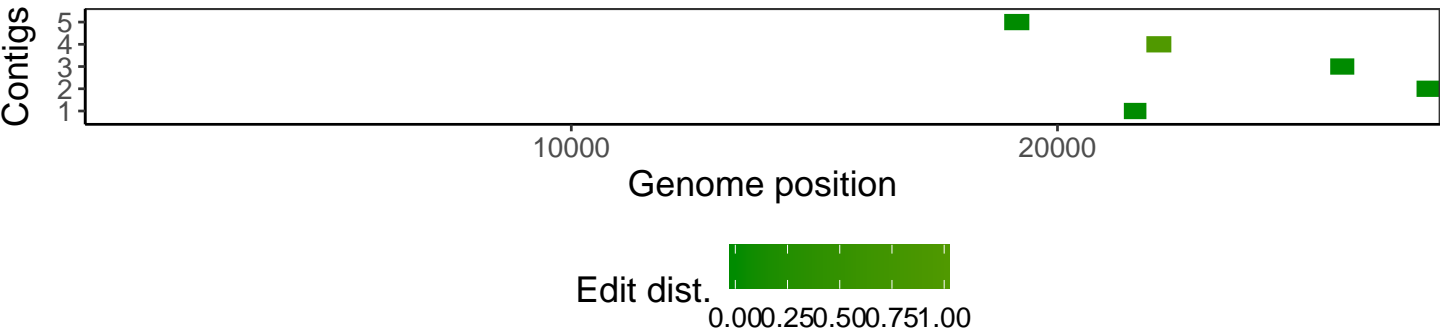
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



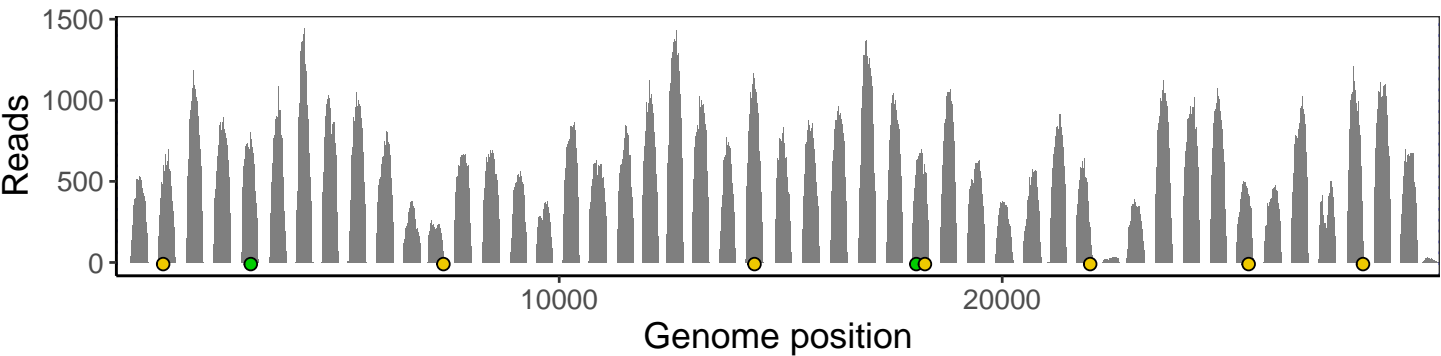
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



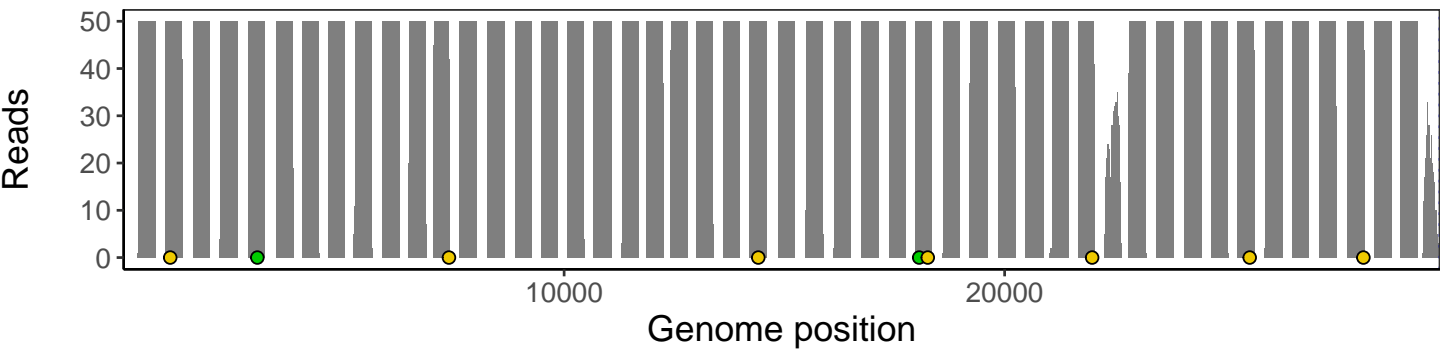
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



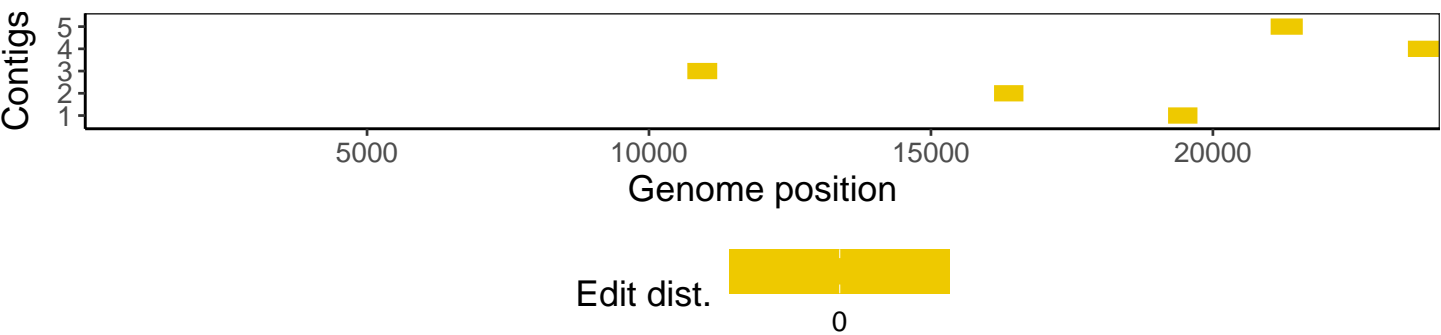
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



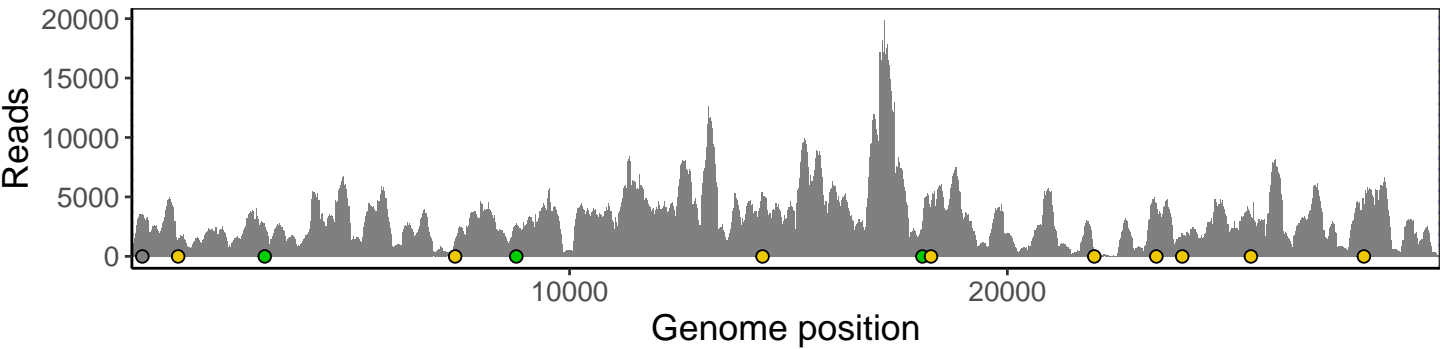
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



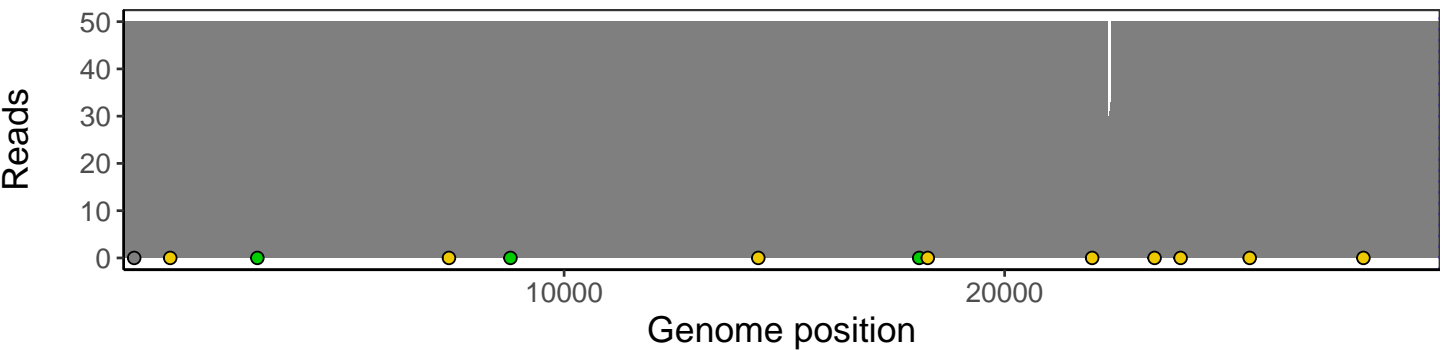
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



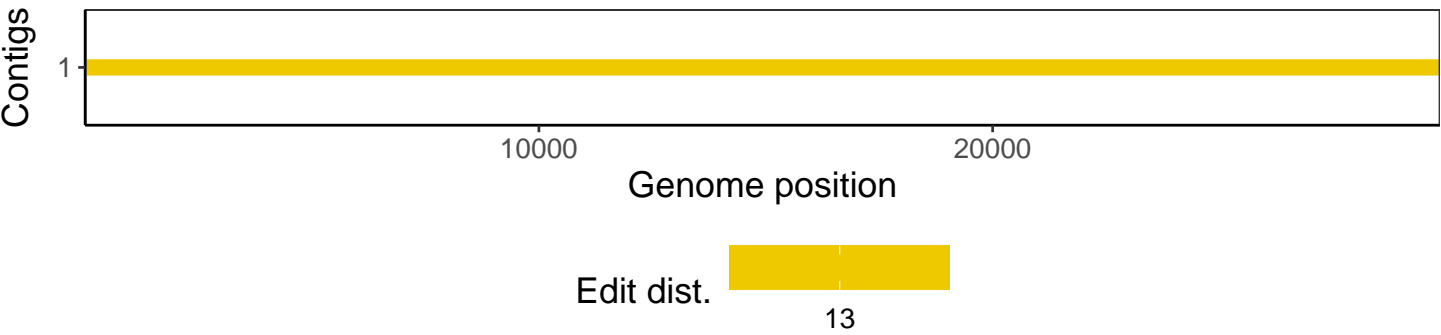
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



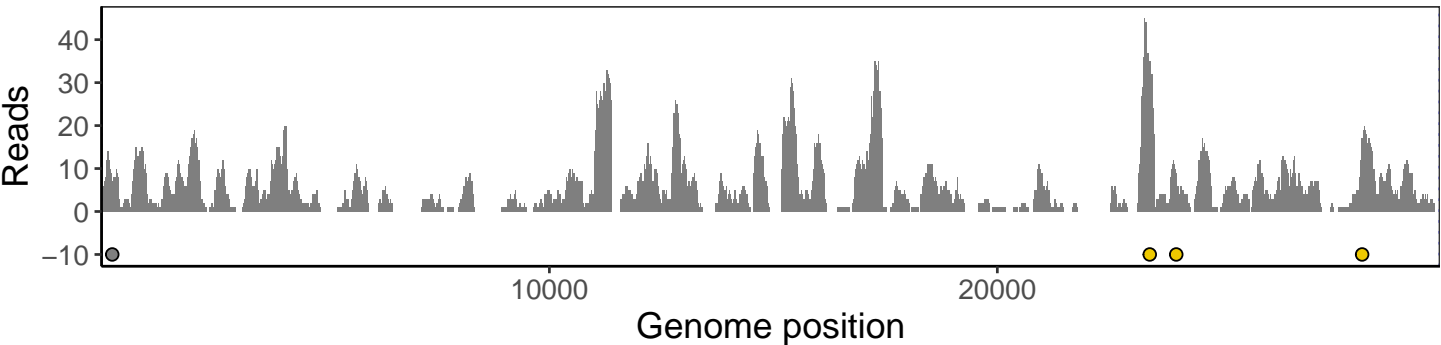
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



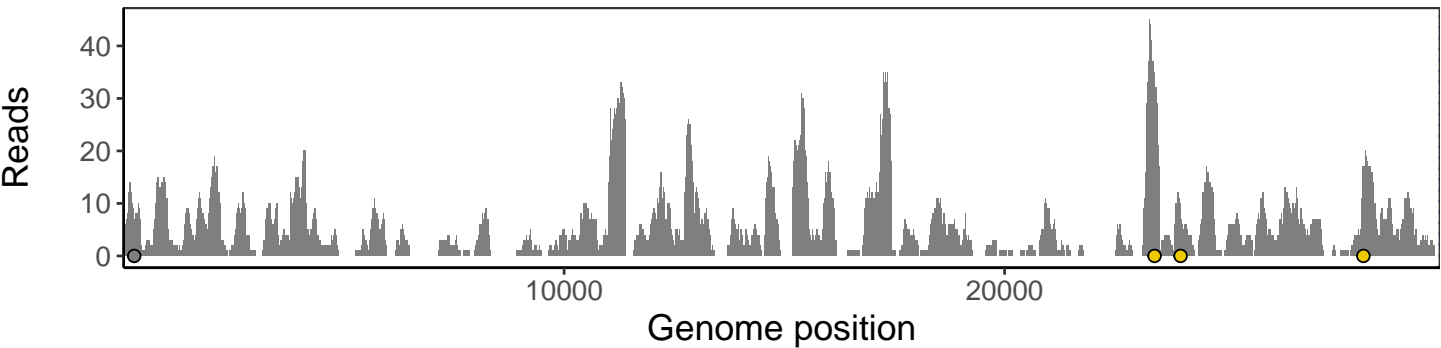
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



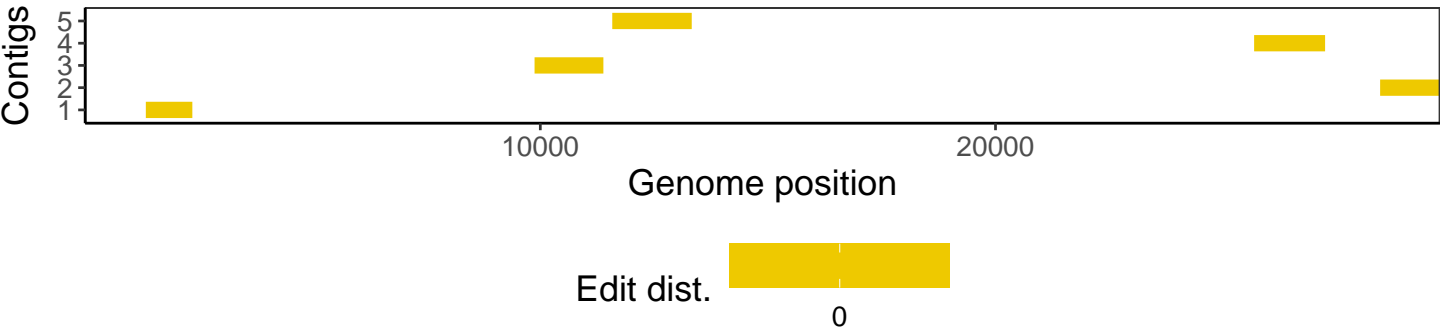
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



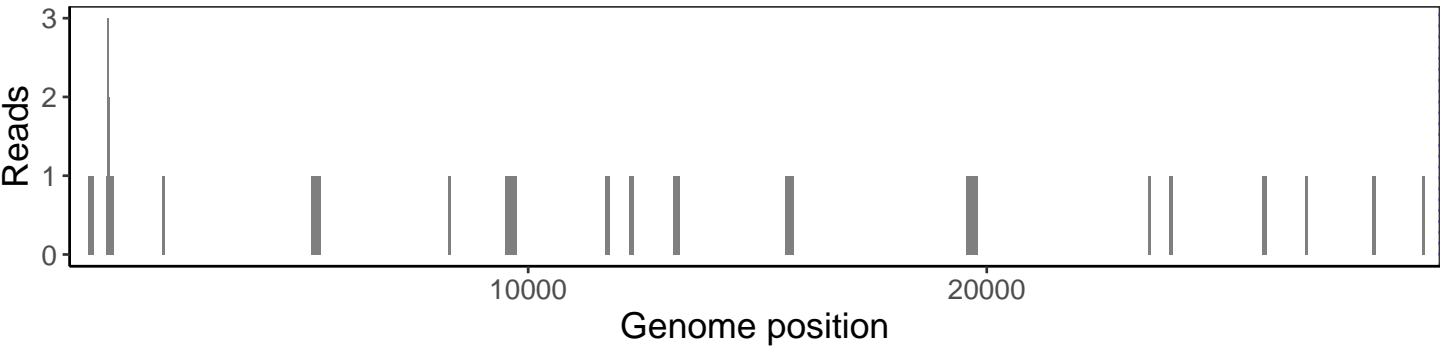
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



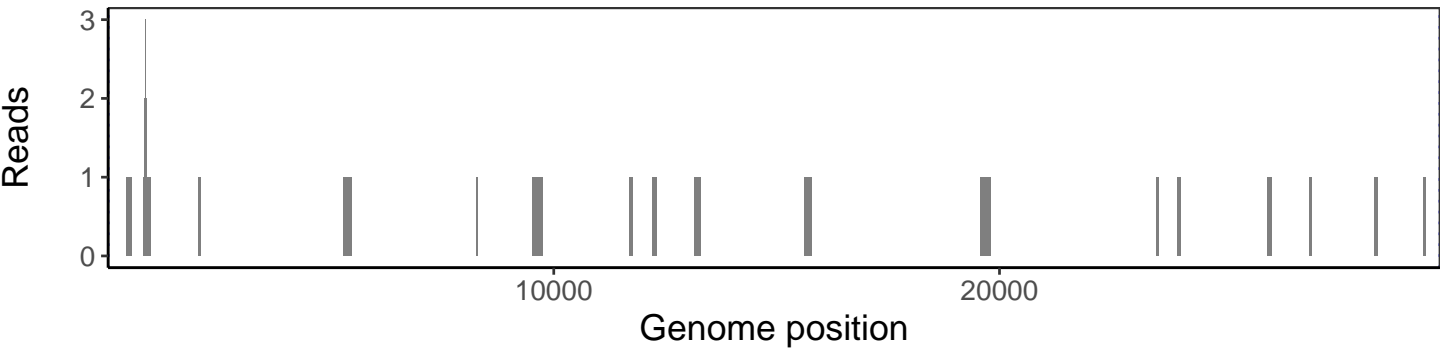
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



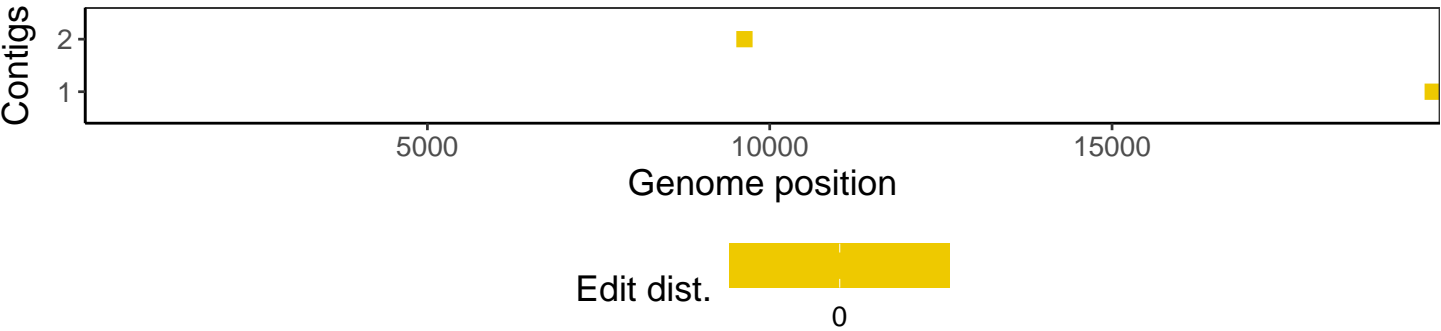
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



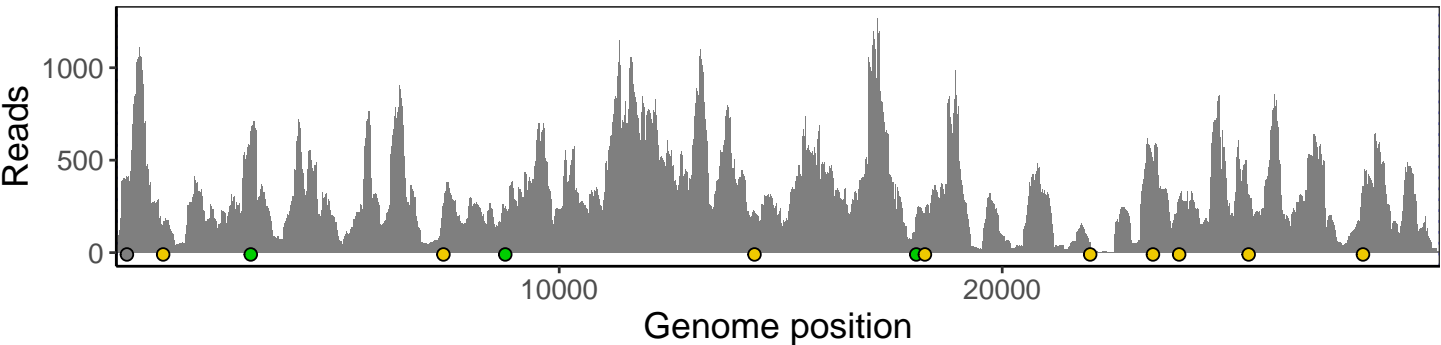
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



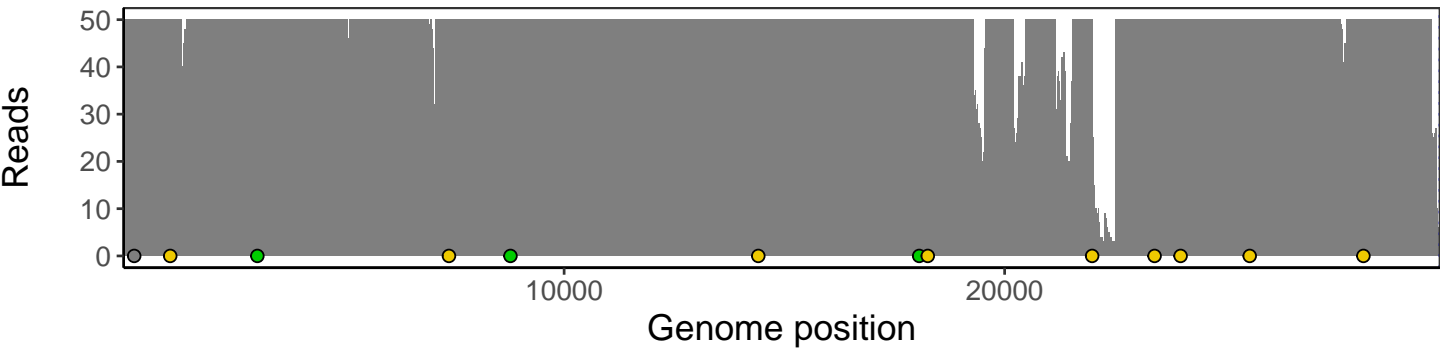
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



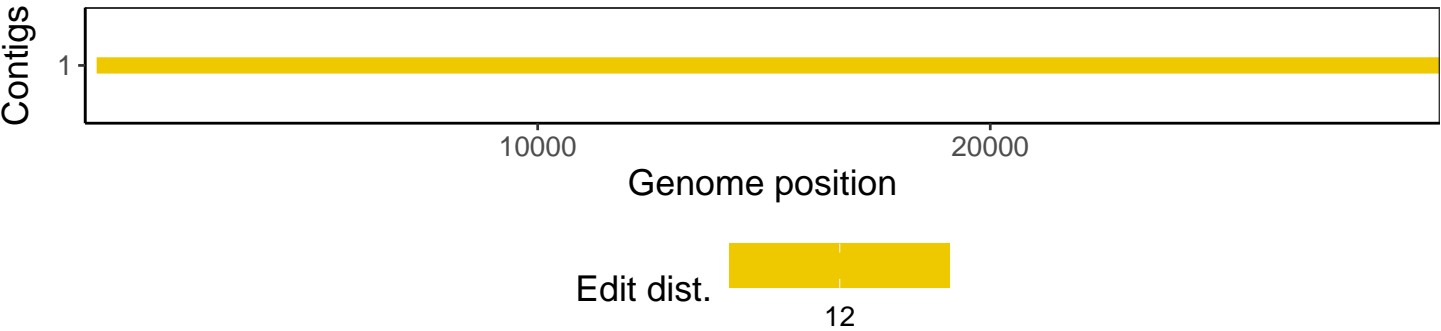
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



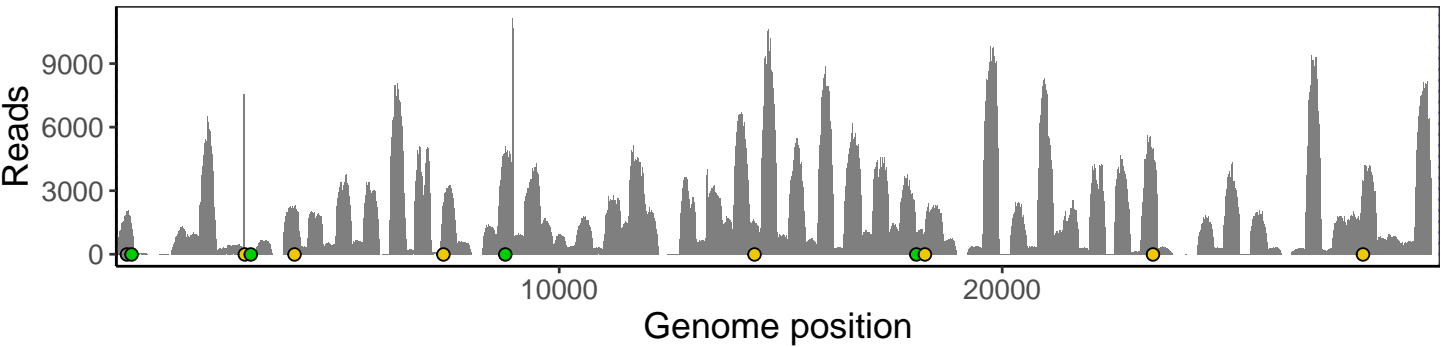
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



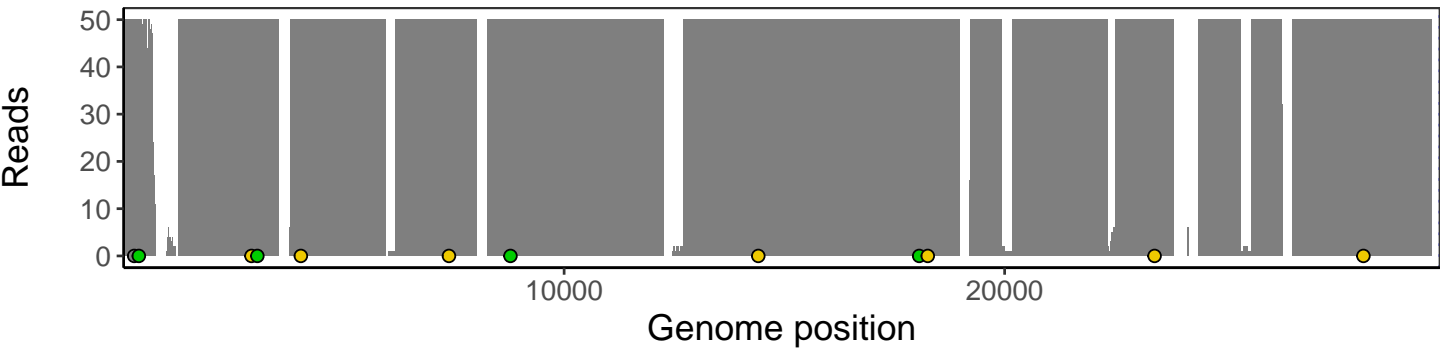
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



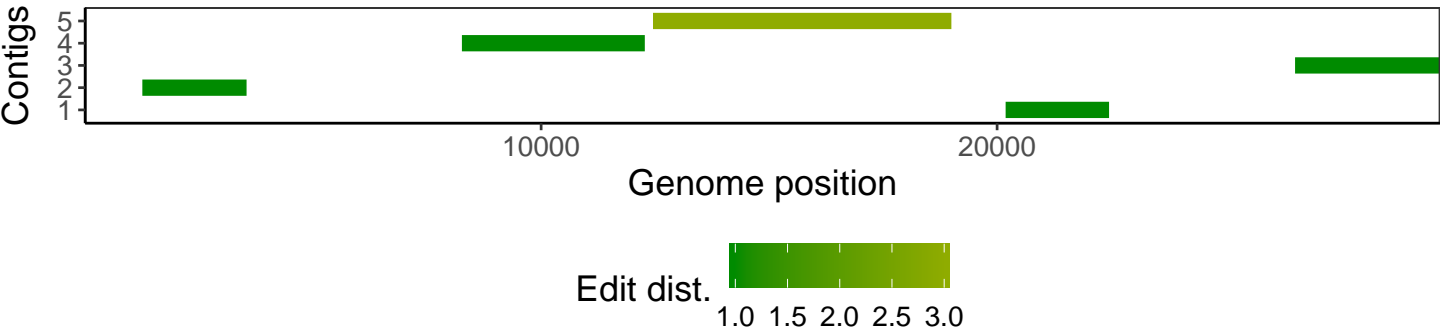
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



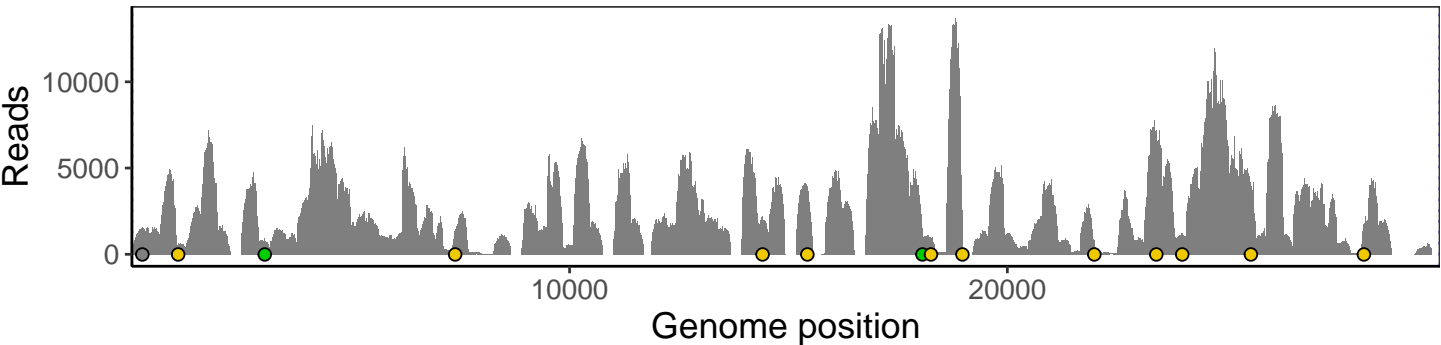
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



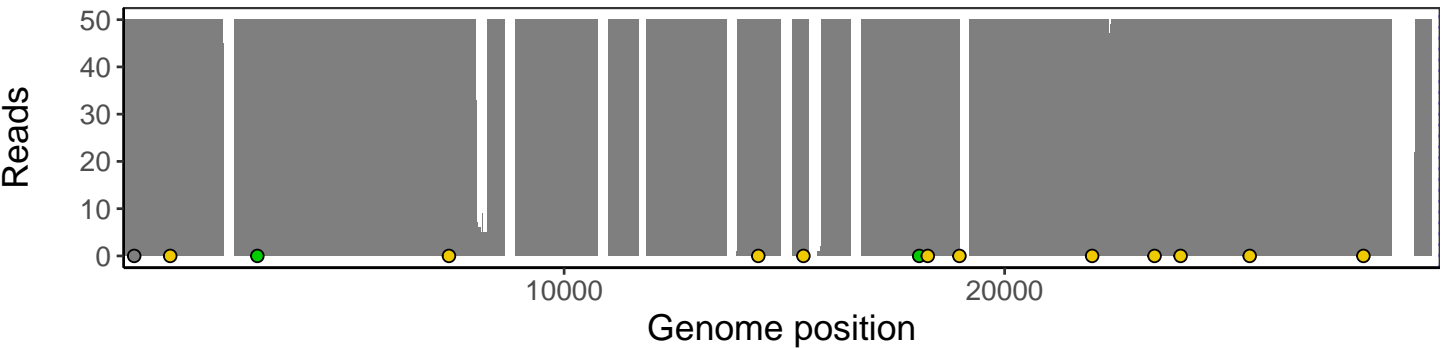
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



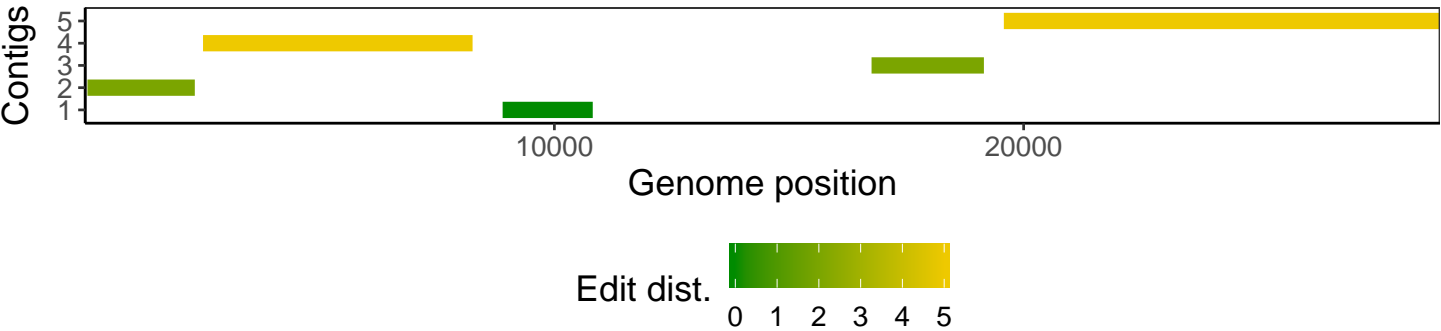
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

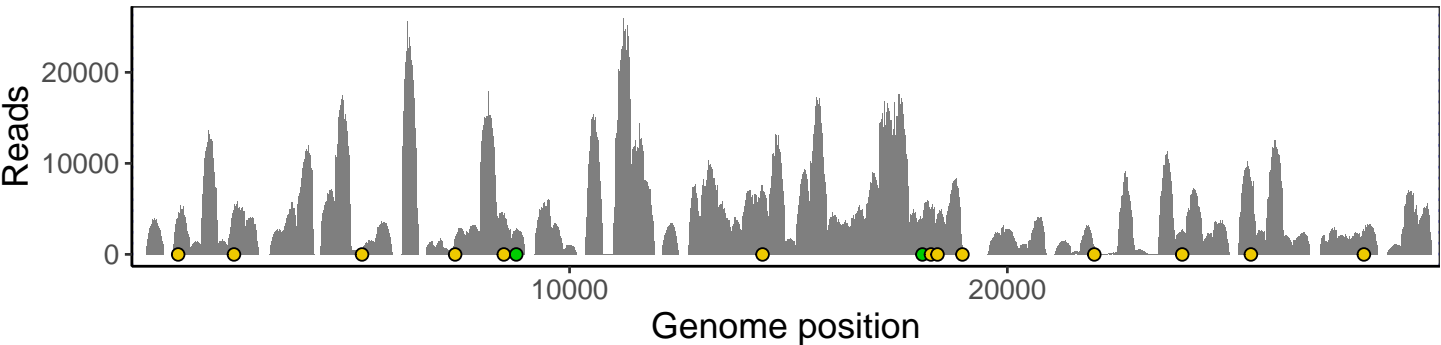


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

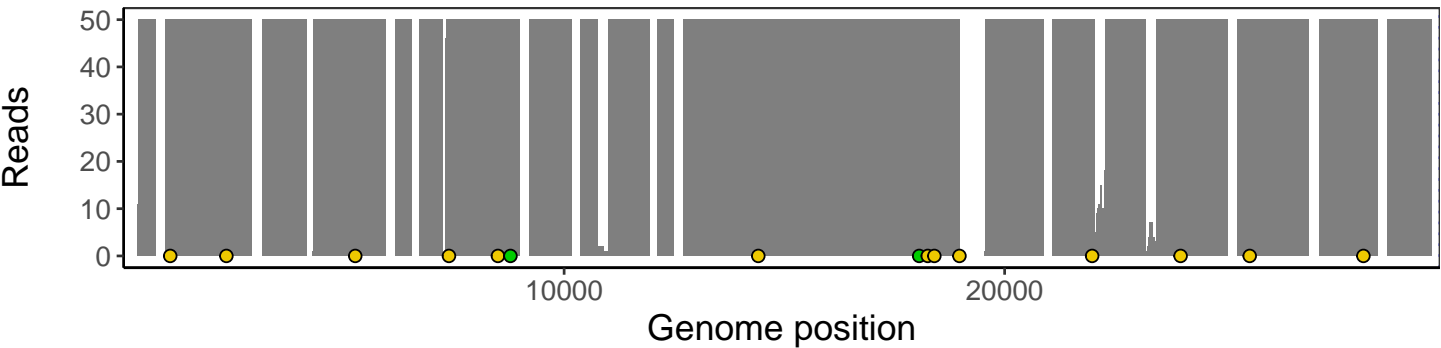




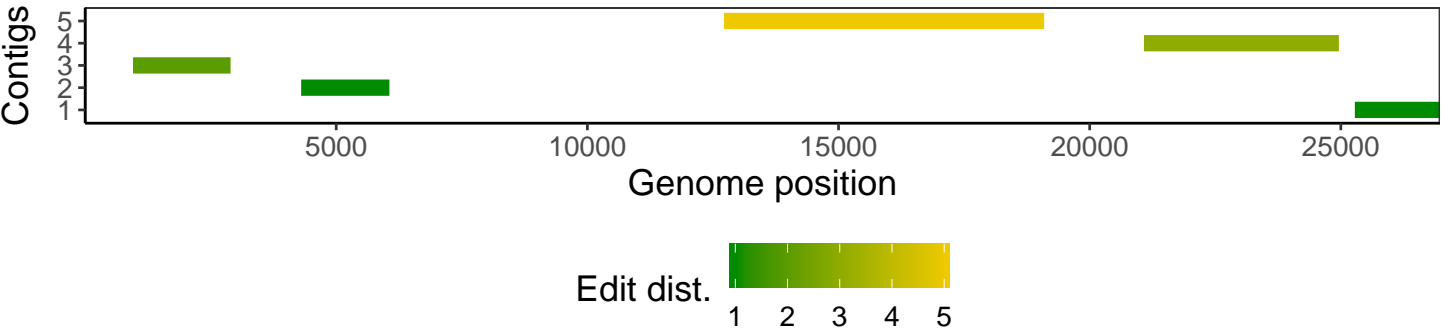
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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.



VSP0187-1 | 05/29/2020 | ETA | 228e-q | 60.2 genomes | single experiment

No pileup data available.

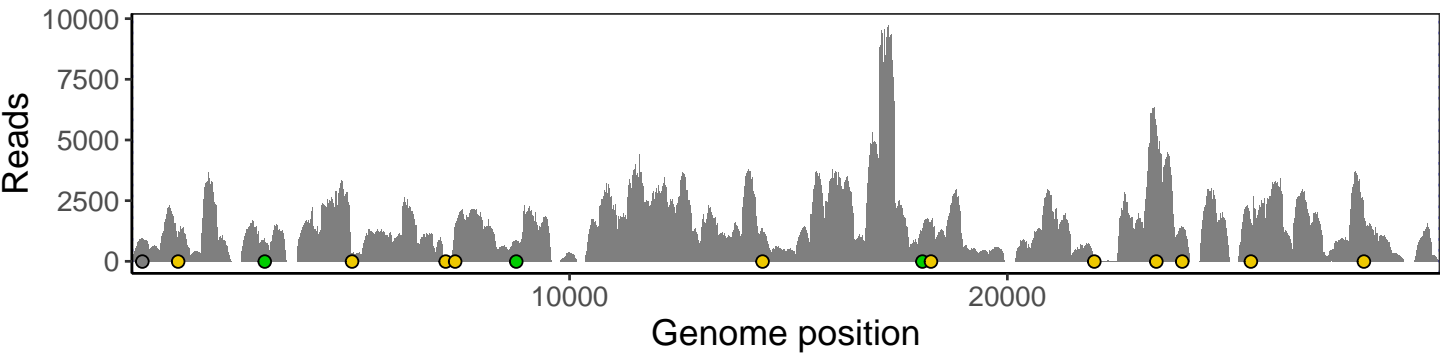
No contig data available.

VSP0187-2 | 05/29/2020 | ETA | 228e-q | 301 genomes | single experiment

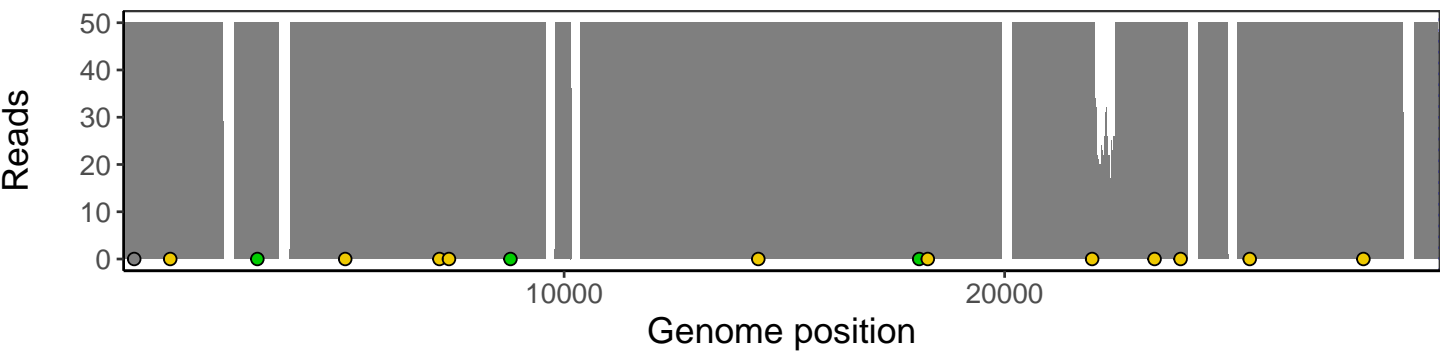
No pileup data available.

No contig data available.

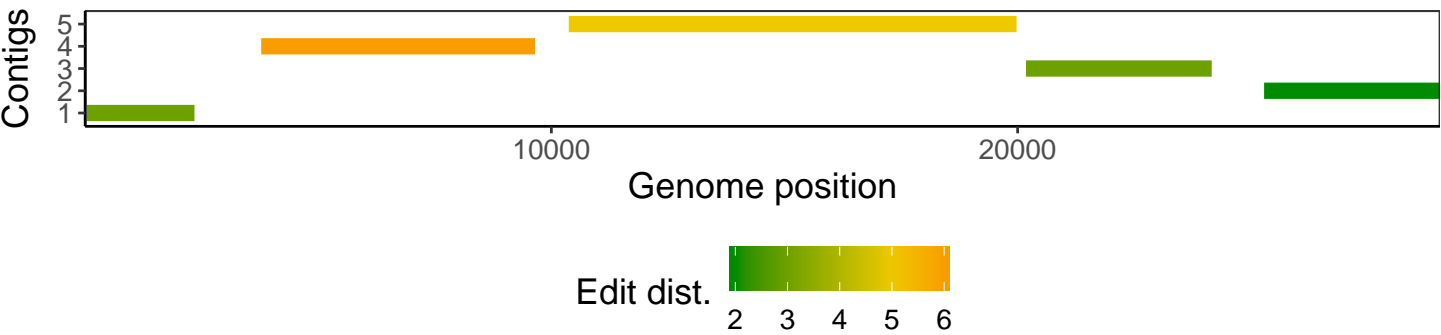
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



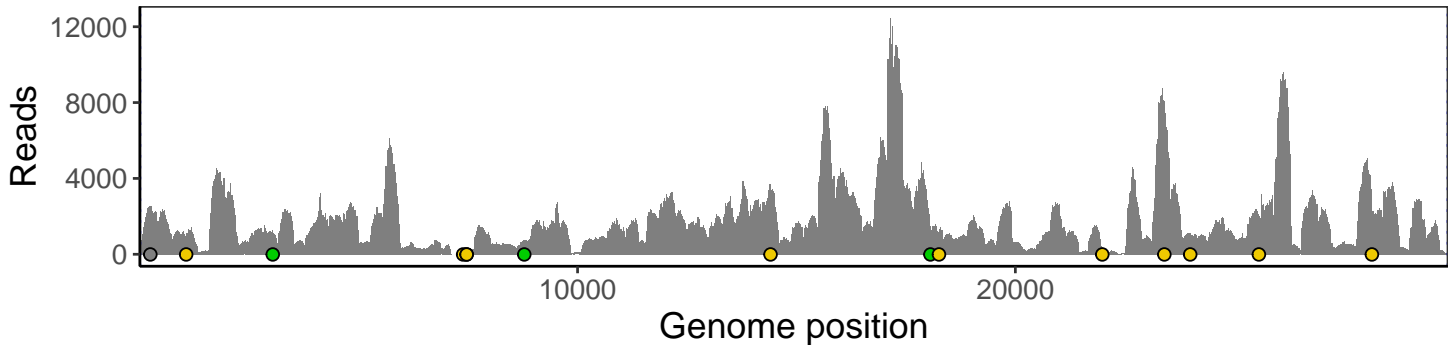
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



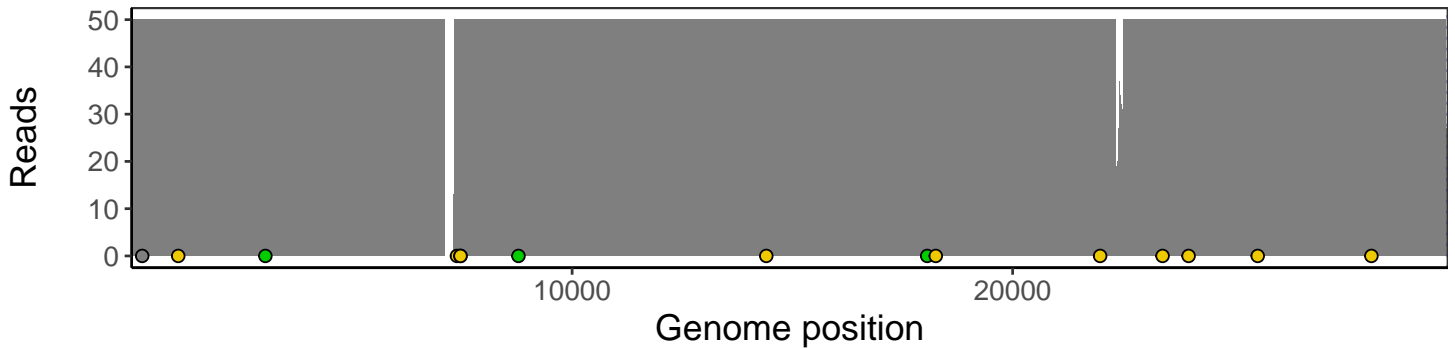
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



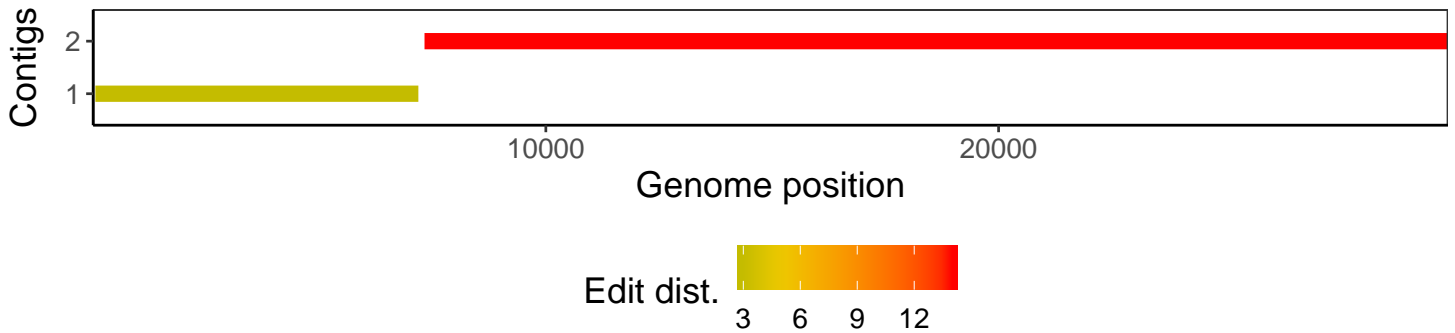
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



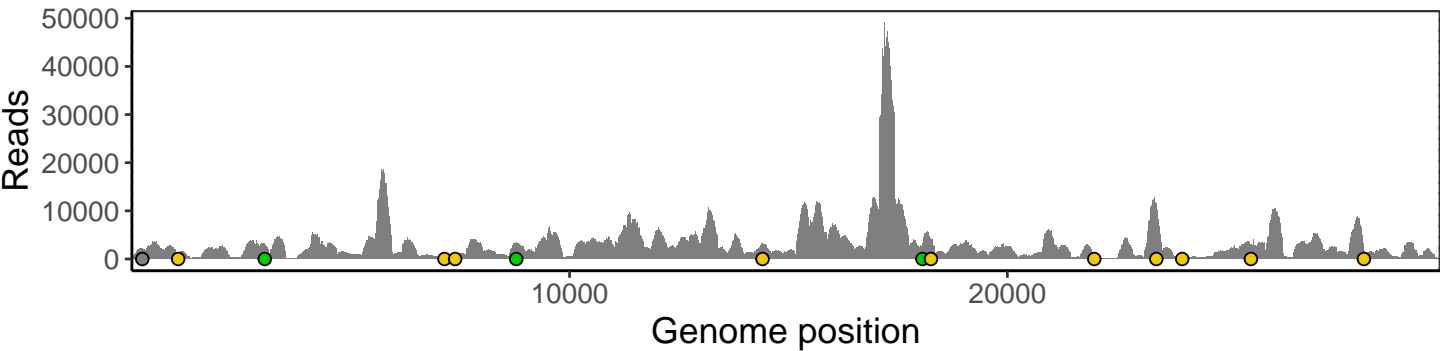
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



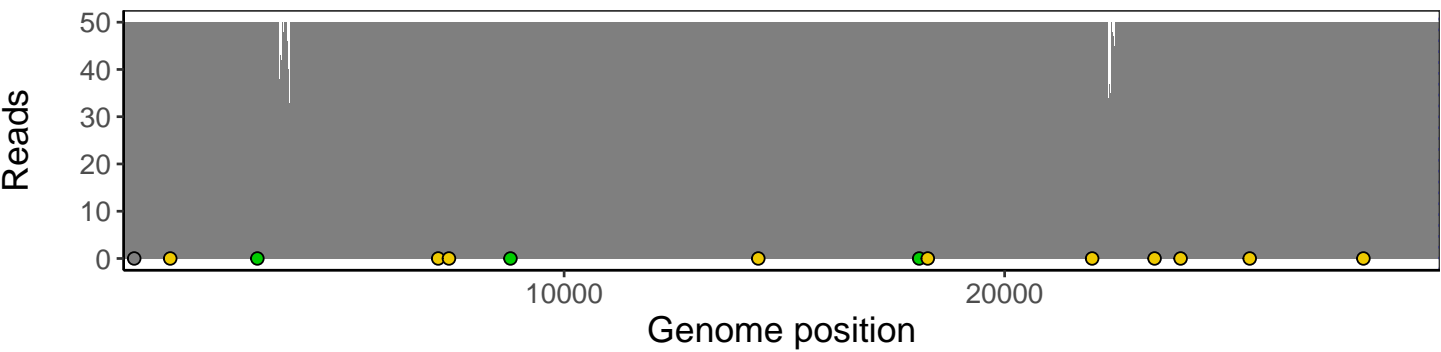
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



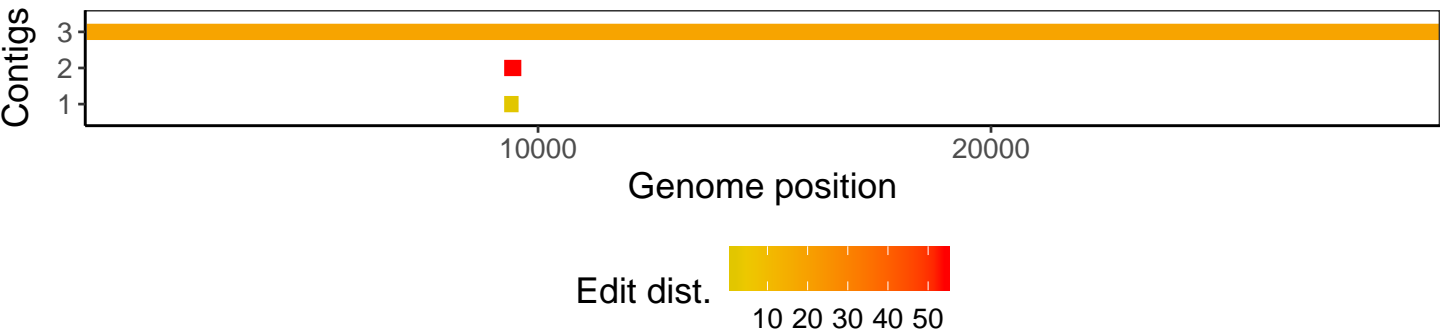
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



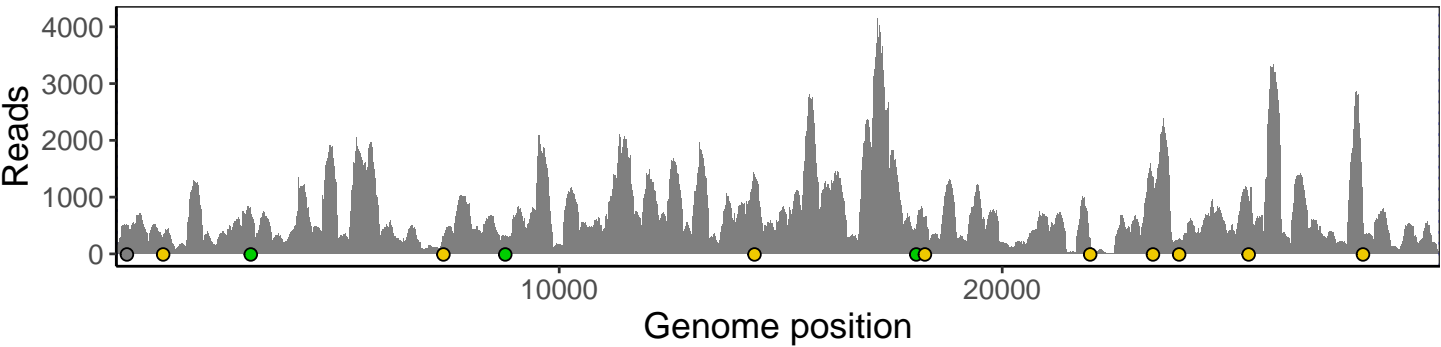
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



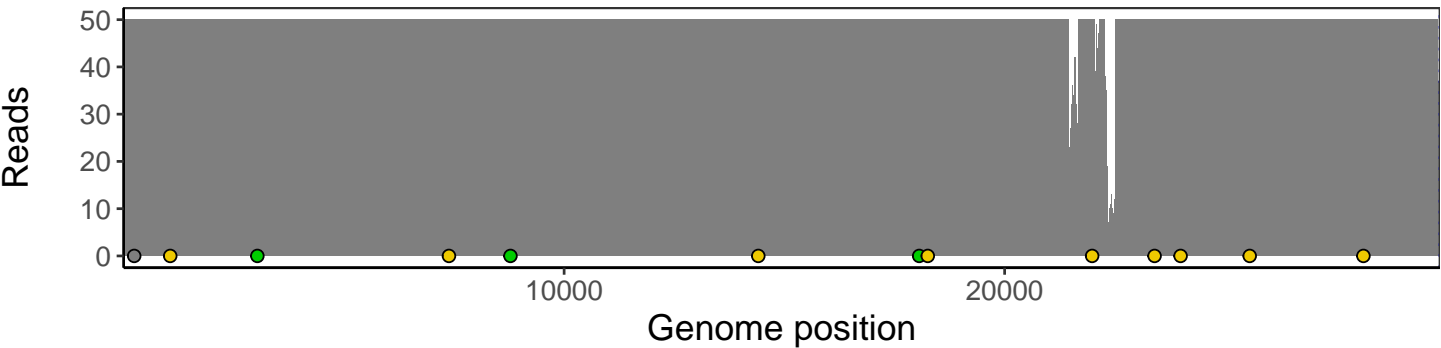
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

