

COVID-19 subject 242

2021-01-10

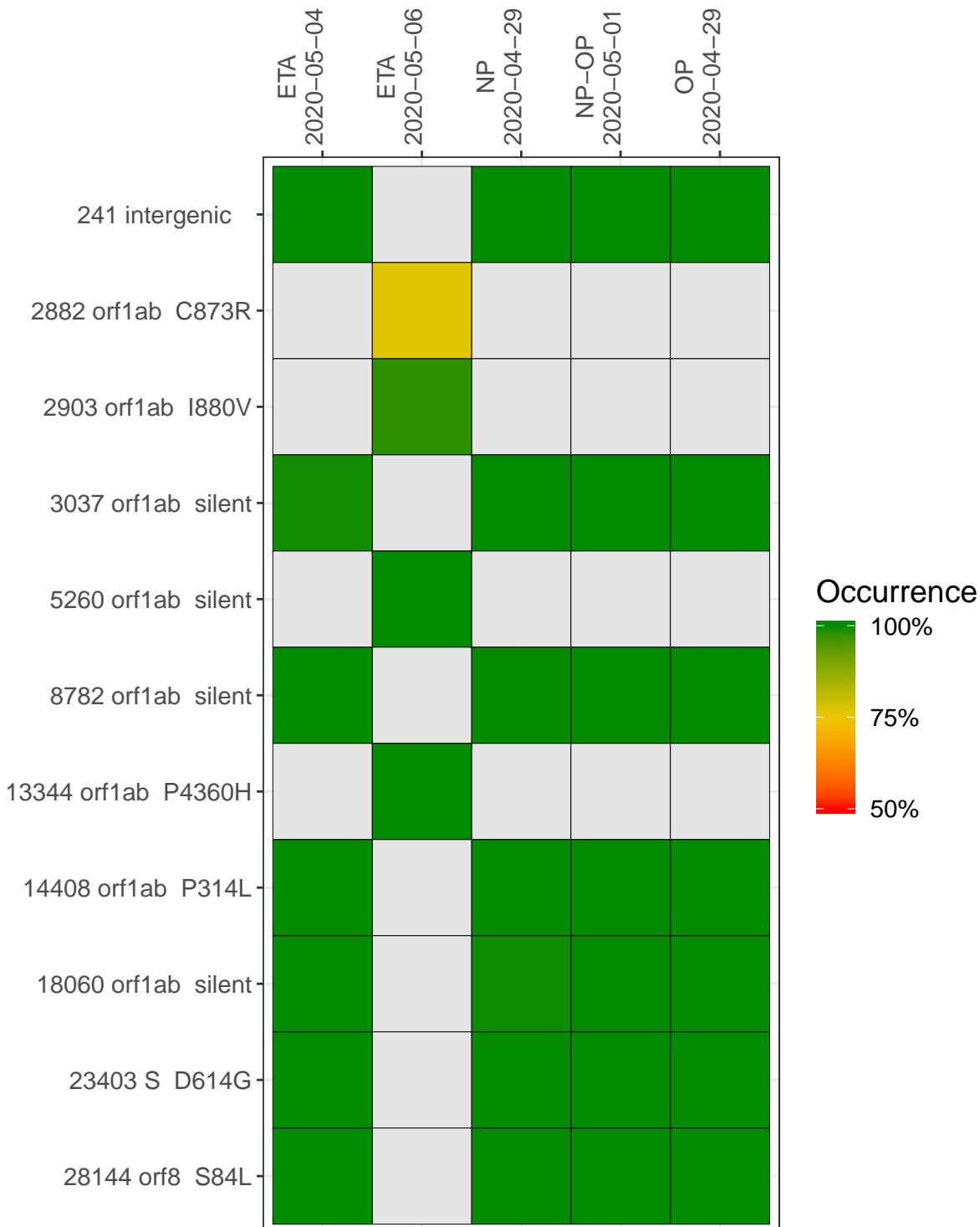
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0044	composite	NA	NP	2020-04-29	29.63	99.9%	99.8%
VSP0085	composite	NA	ETA	2020-05-06	NA	21.1%	2.9%
VSP0044-1m	single experiment	NA	NP	2020-04-29	29.17	99.8%	99.7%
VSP0044-2	single experiment	1.975e+05	NP	2020-04-29	16.80	99.9%	99.7%
VSP0045-1	single experiment	6.670e+02	OP	2020-04-29	29.72	99.6%	99.2%
VSP0047-1	single experiment	2.220e+05	NP-OP	2020-05-01	30.01	99.9%	99.8%
VSP0052-1	single experiment	1.420e+06	ETA	2020-05-04	29.95	99.9%	99.8%
VSP0085-1	single experiment	8.540e+00	ETA	2020-05-06	NA	16.2%	2.3%
VSP0085-2	single experiment	4.270e+01	ETA	2020-05-06	0.27	0.2%	0.0%
VSP0085-3	single experiment	4.270e+01	ETA	2020-05-06	0.20	2.4%	0.0%
VSP0085-4	single experiment	4.270e+01	ETA	2020-05-06	0.27	4.6%	0.0%

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.



	ETA 2020-05-04	ETA 2020-05-06	NP 2020-04-29	NP-OP 2020-05-01	OP 2020-04-29	
241 intergenic	2266		2732	1124	1628	1337
2882 orf1ab C873R	1698	52	648	1583	907	1751
2903 orf1ab I880V	361	150	113	1824	713	1311
3037 orf1ab silent	458		97	2007	824	1396
5260 orf1ab silent	52	5	42	585	46	600
8782 orf1ab silent	293		150	629	57	1006
13344 orf1ab P4360H	2680	27	1274	1171	4610	3488
14408 orf1ab P314L	497		326	467	1802	4963
18060 orf1ab silent	179		71	637	310	1277
23403 S D614G	25216	1	17547	659	18956	2638
28144 orf8 S84L	3169		1740	911	3272	3017
	VSP0052-1	VSP0085-1	VSP0044-1m	VSP0044-2	VSP0047-1	VSP0045-1

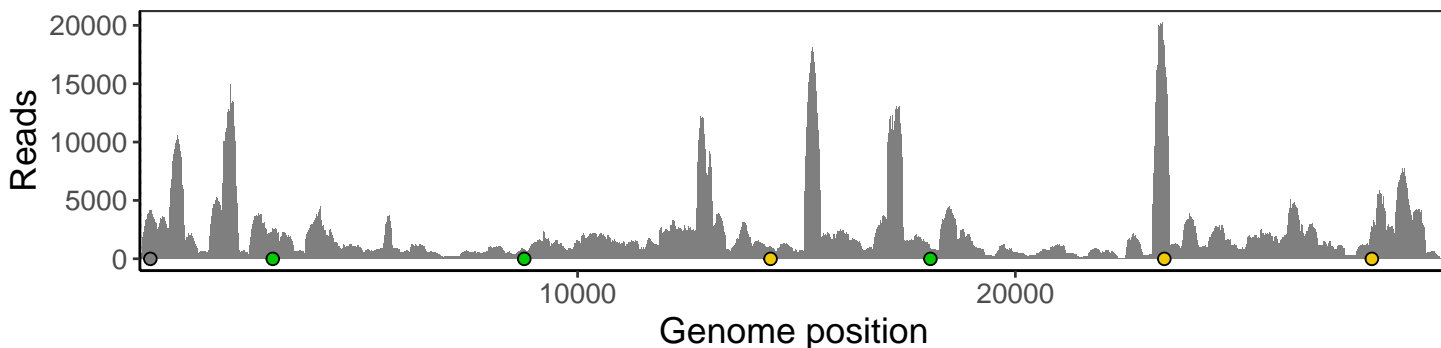
Base change

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

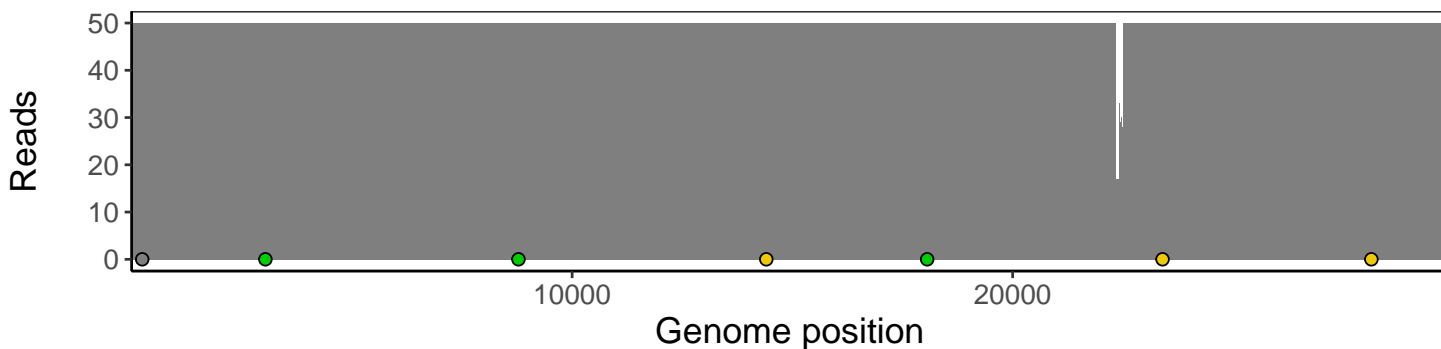
Analyses of individual experiments and composite results.

VSP0044 | 2020-04-29 | NP | 242n-tri | composite result

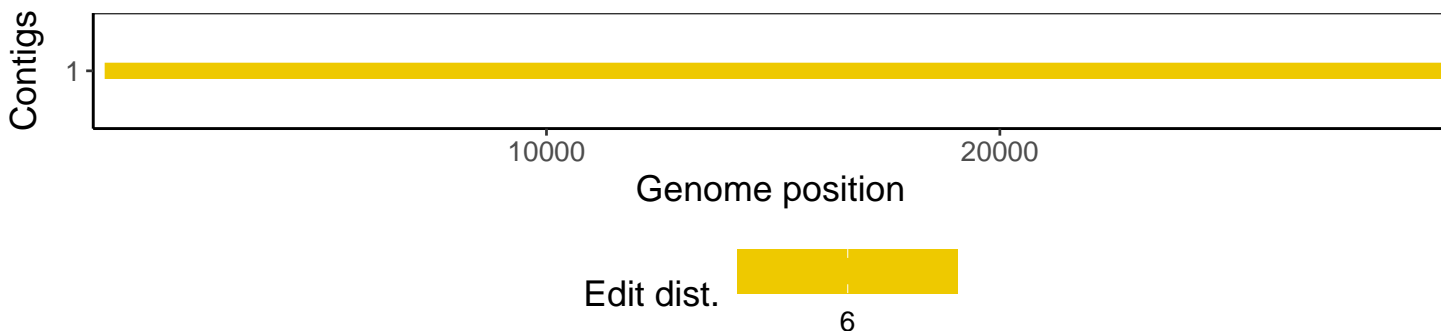
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

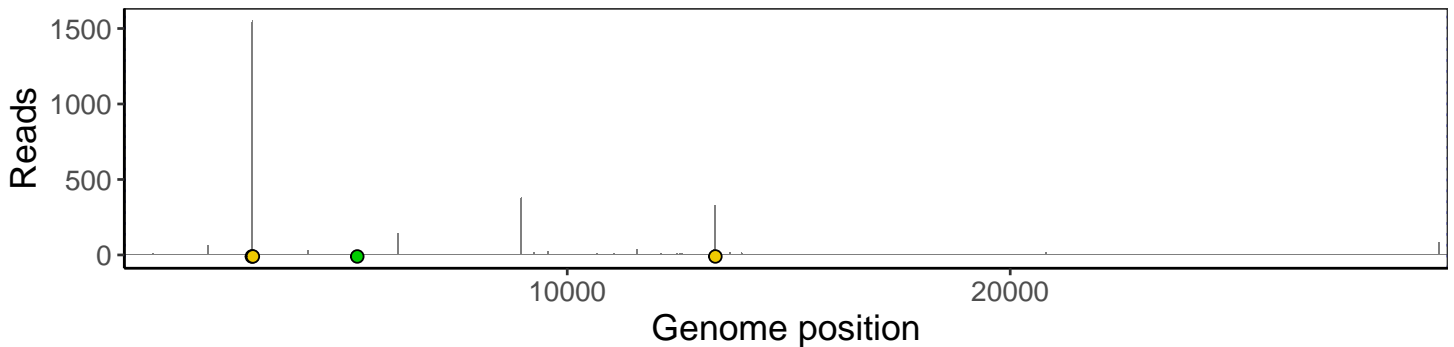


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

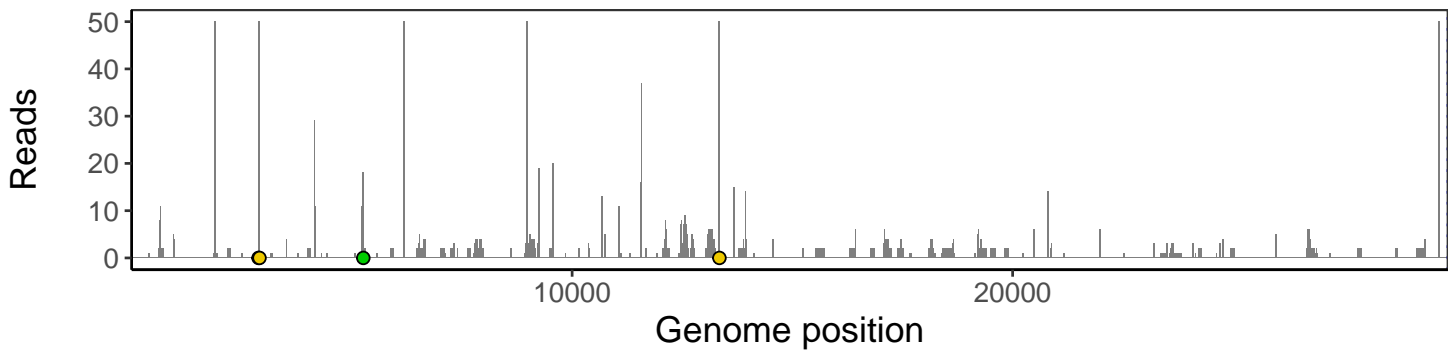


VSP0085 | 2020-05-06 | ETA | 242e-q | composite result

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.

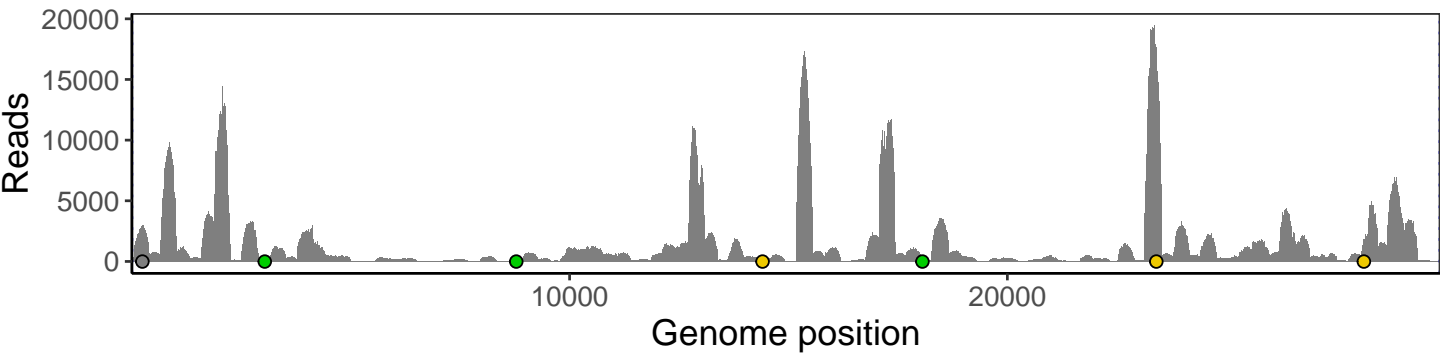


Excerpt from plot above focusing on reads coverage from 0 to 50 NT.

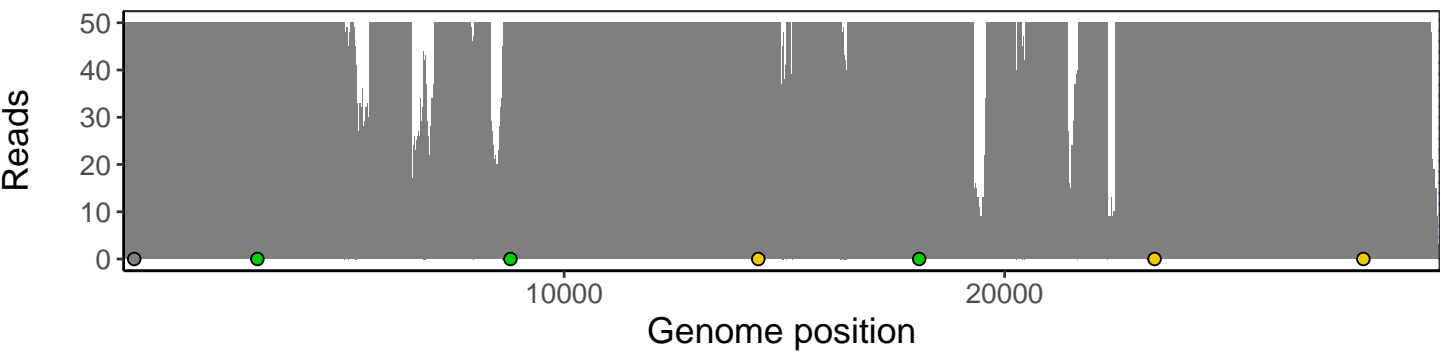


No contig data available.

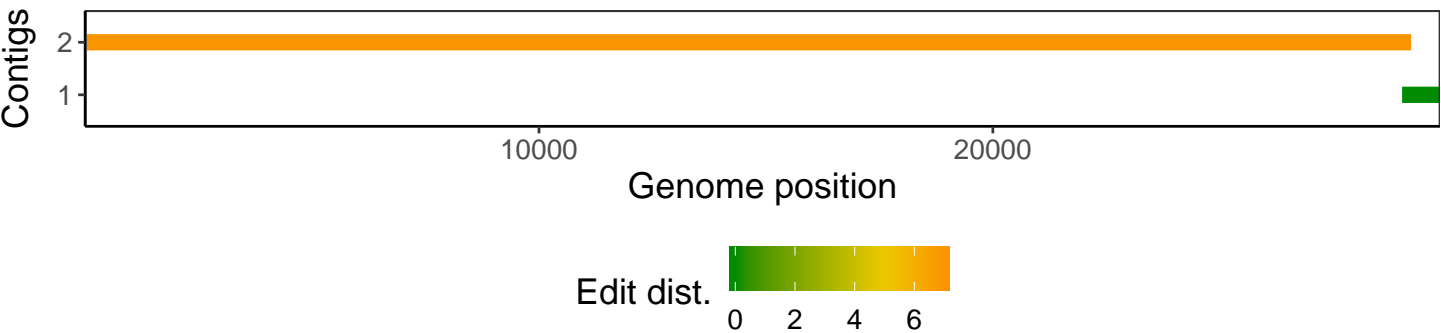
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



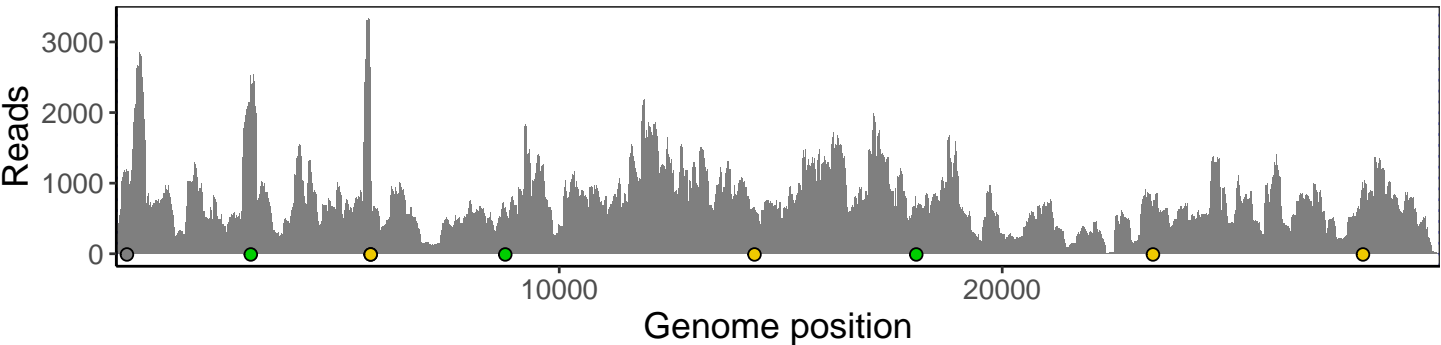
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



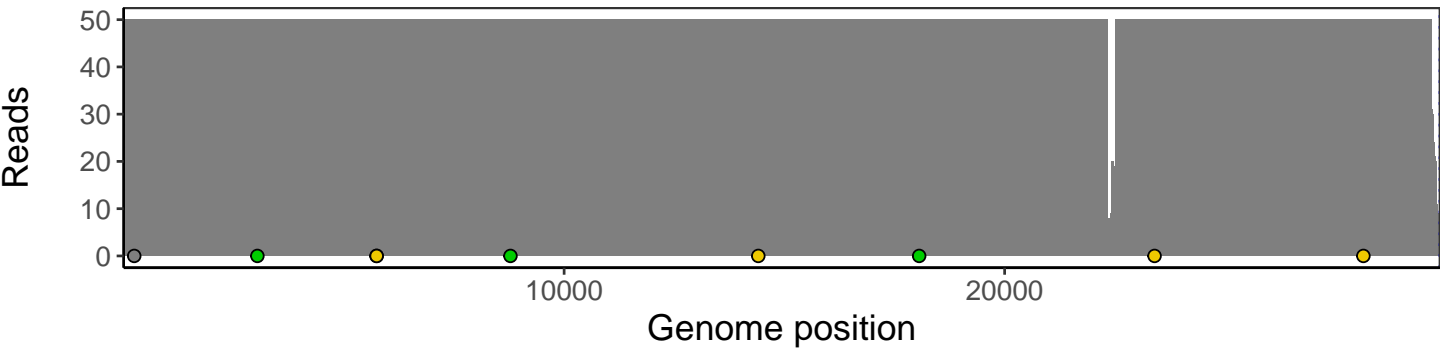
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



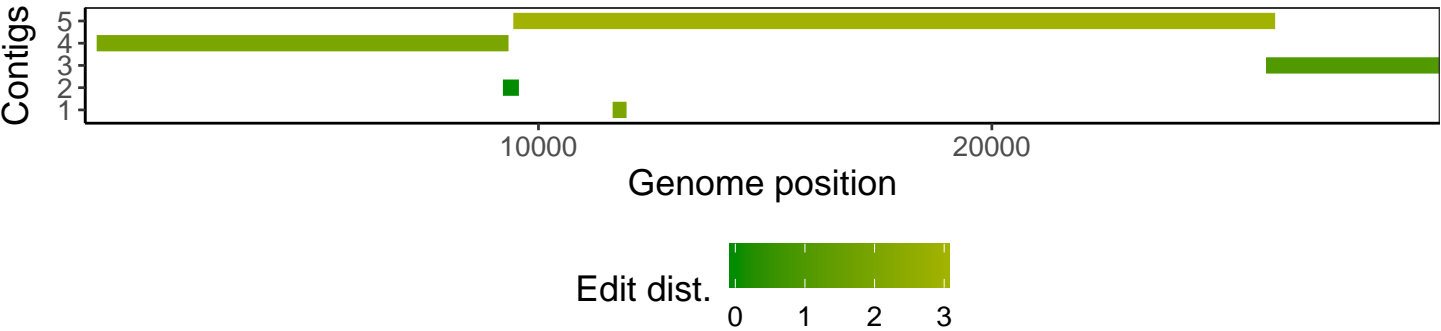
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



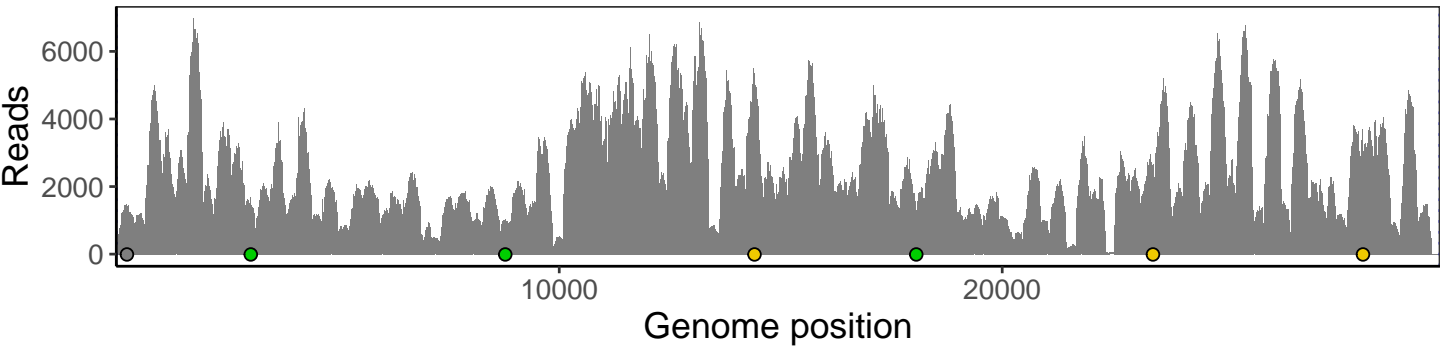
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



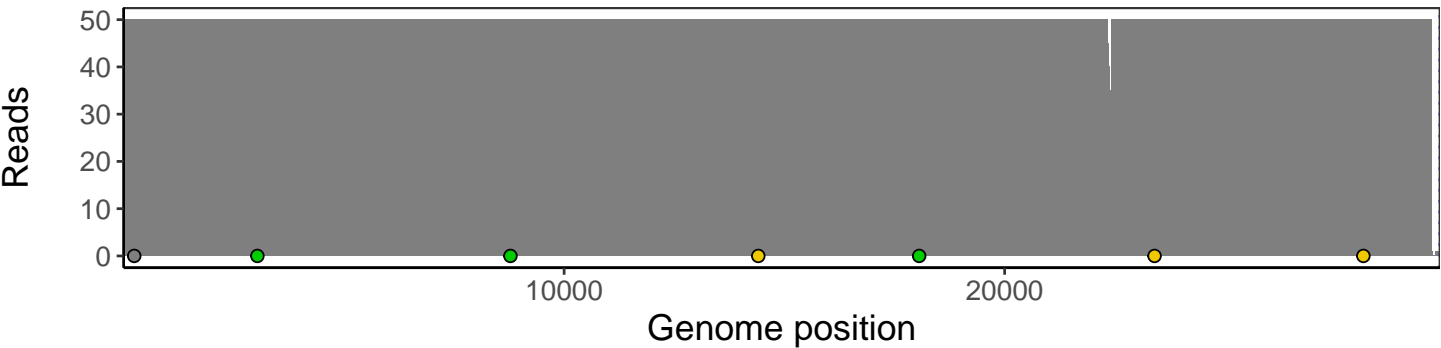
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



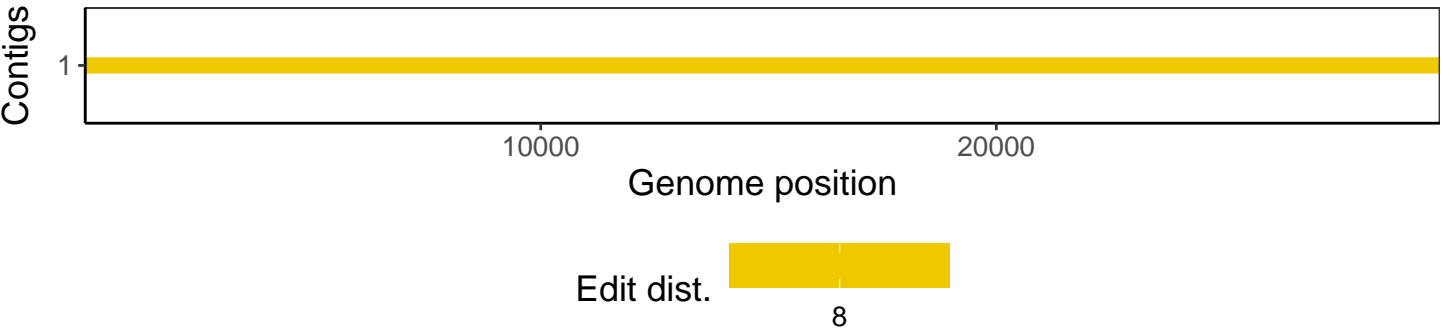
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



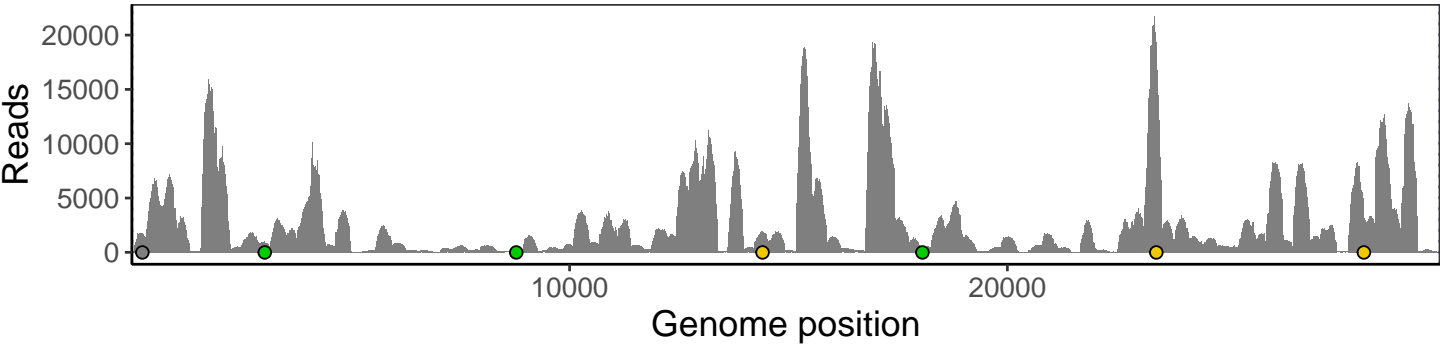
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



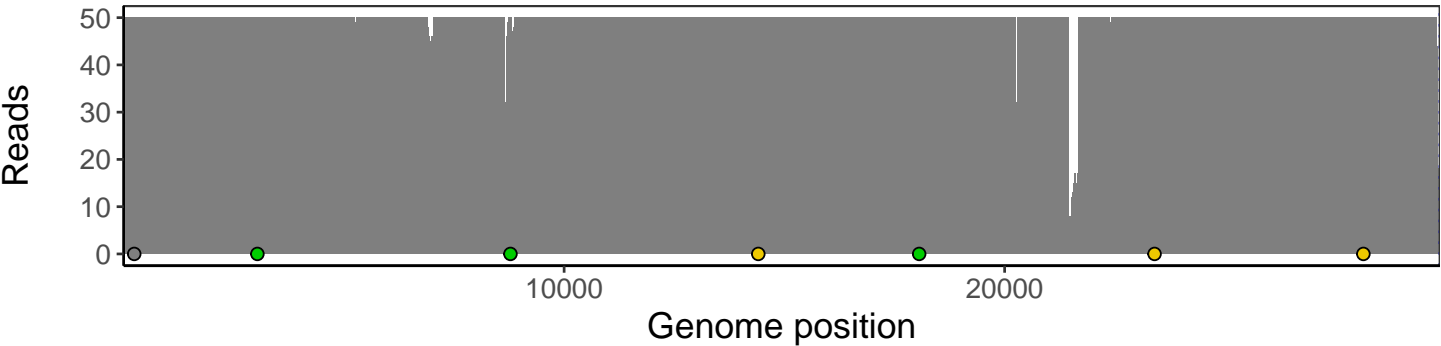
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



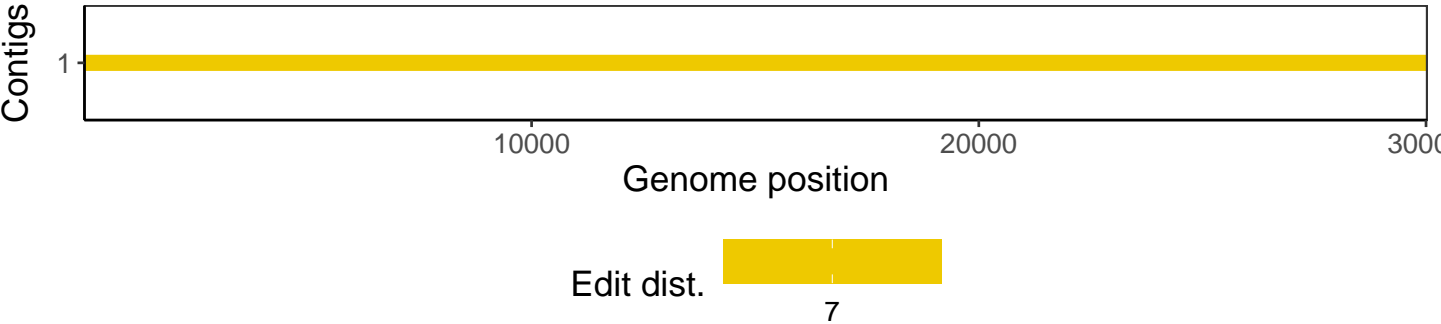
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



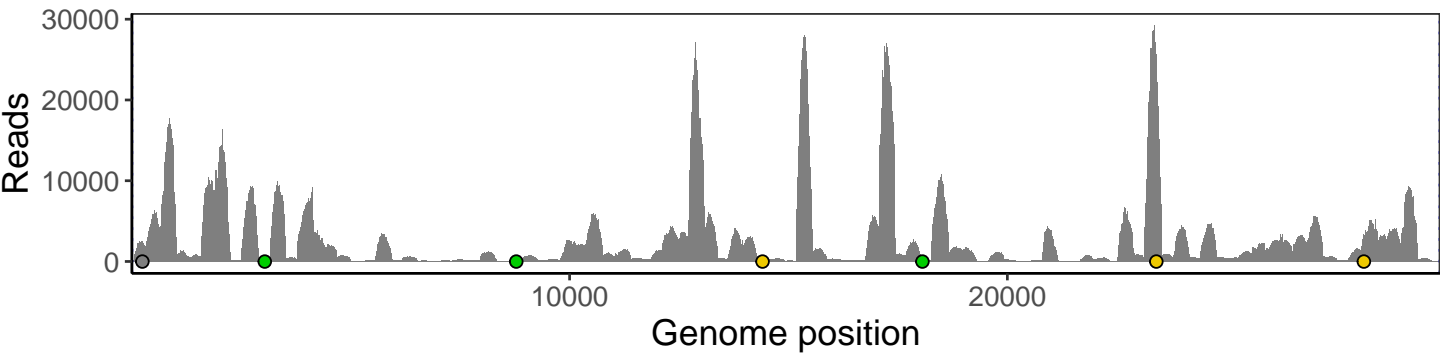
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



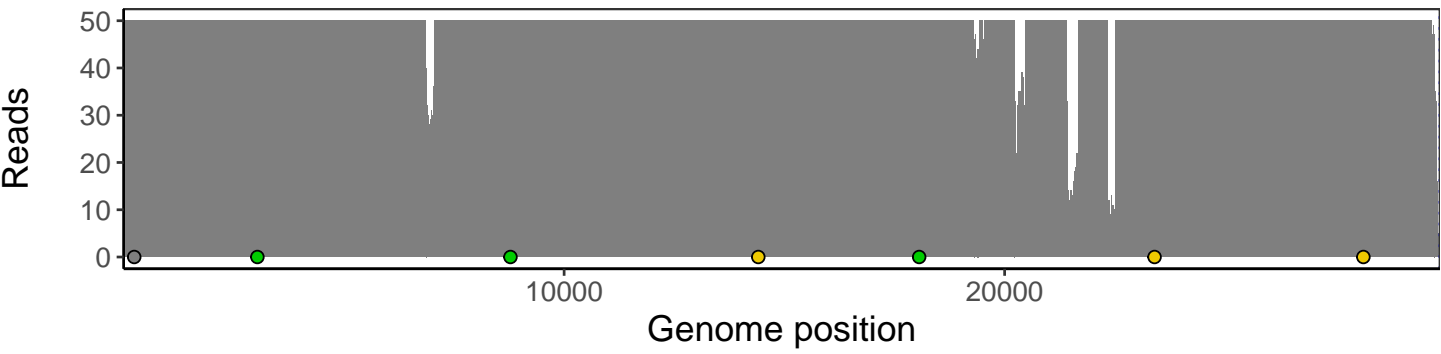
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



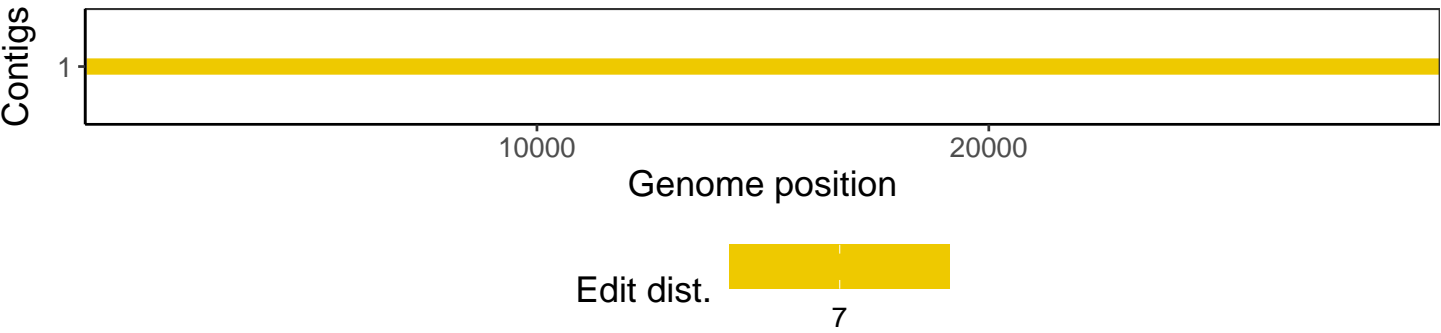
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

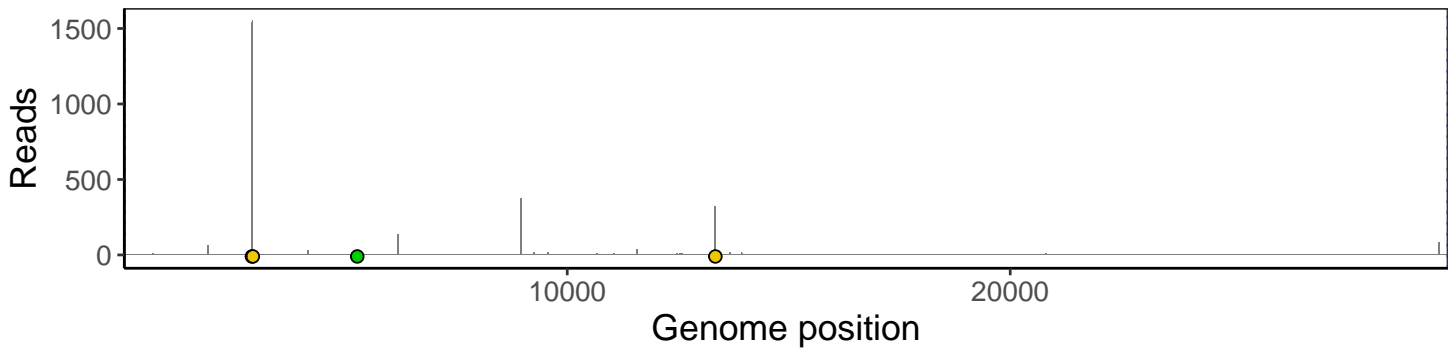


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

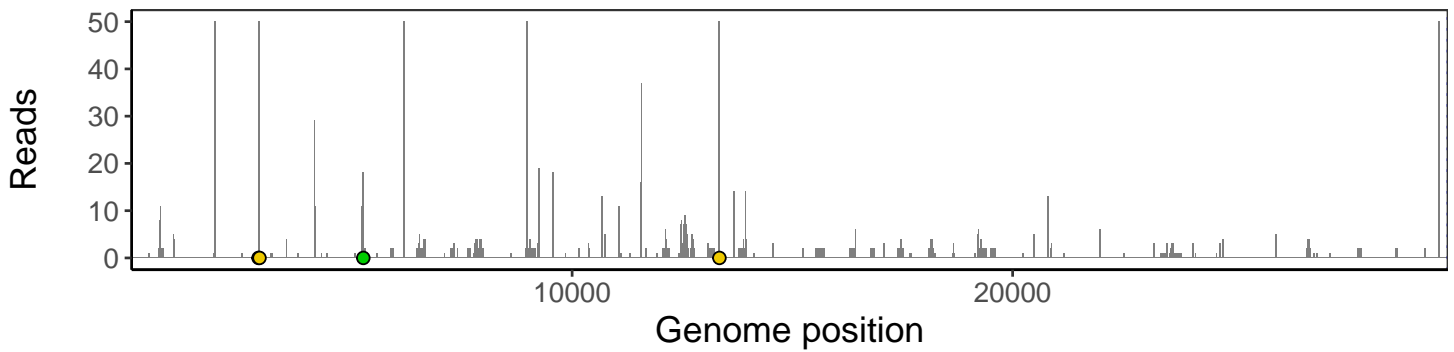


VSP0085-1 | 2020-05-06 | ETA | 242e-q | 8.54 genomes | single experiment

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.

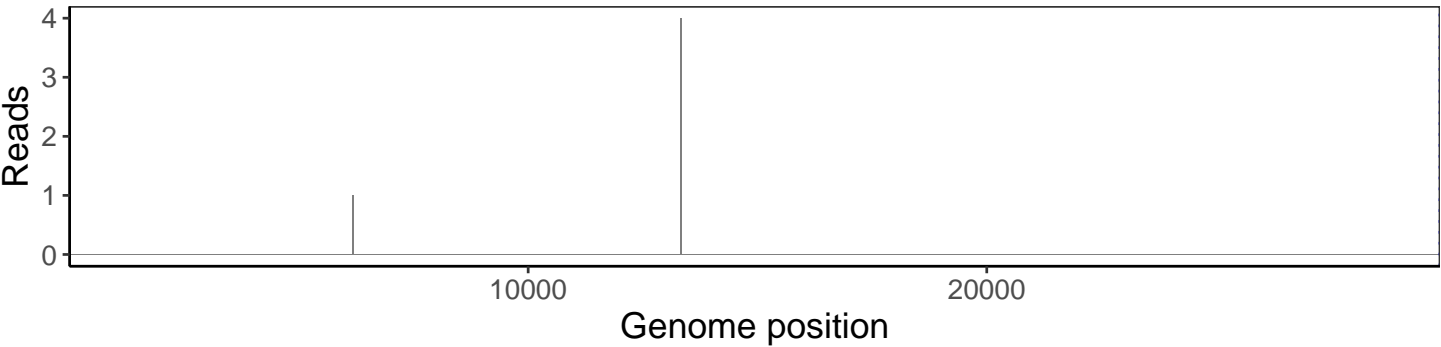


Excerpt from plot above focusing on reads coverage from 0 to 50 NT.

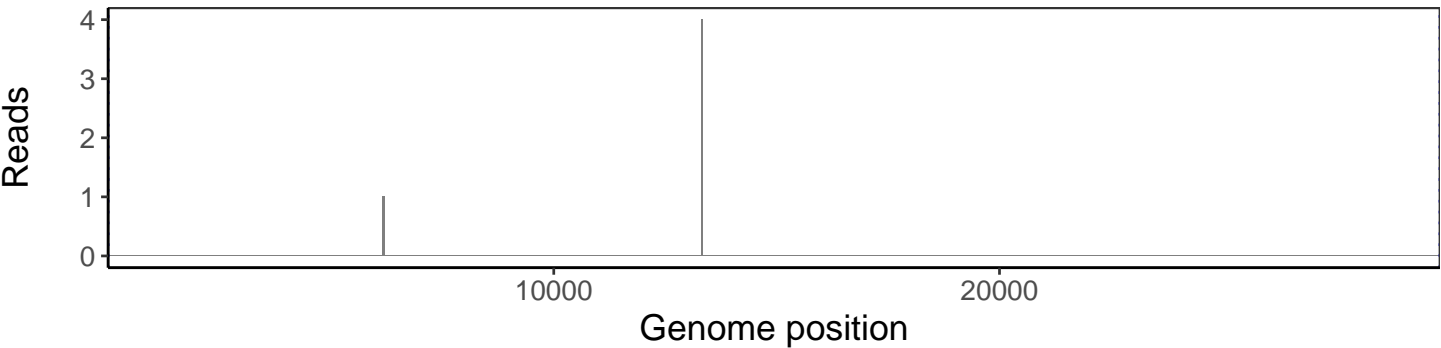


No contig data available.

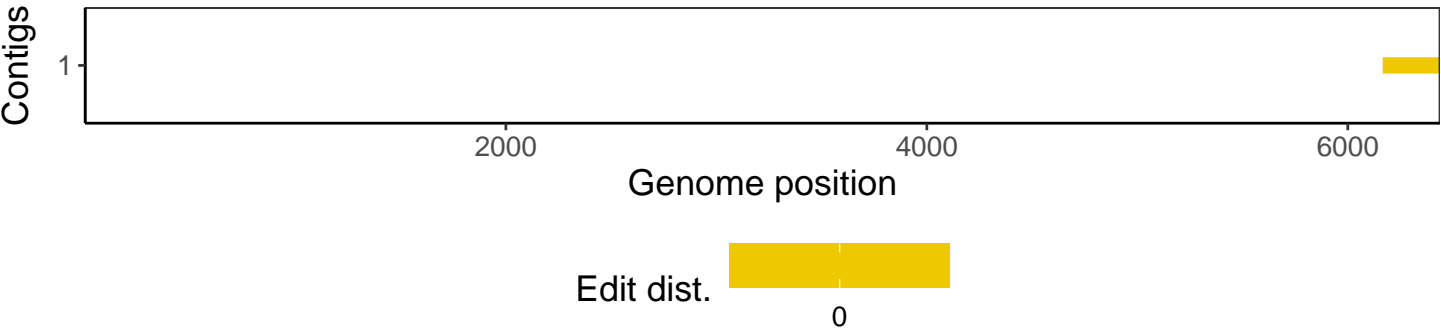
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



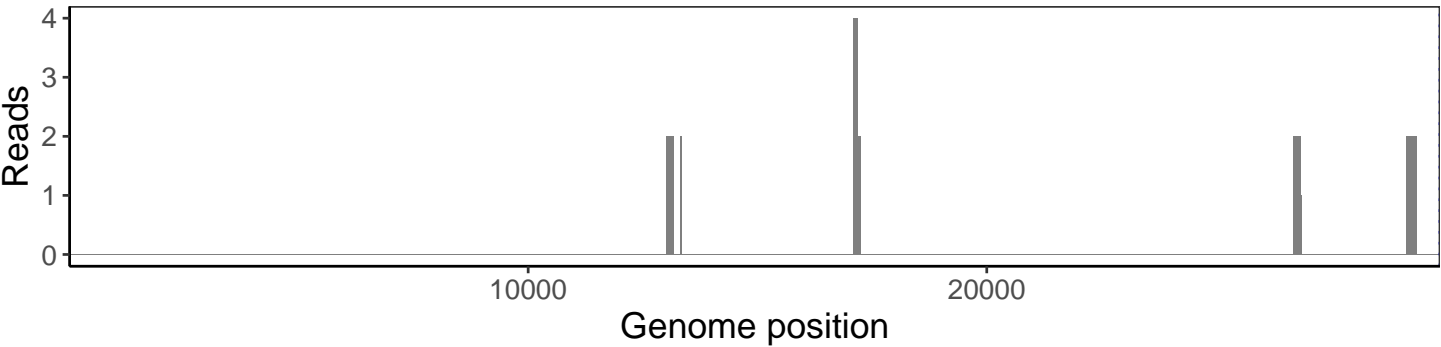
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



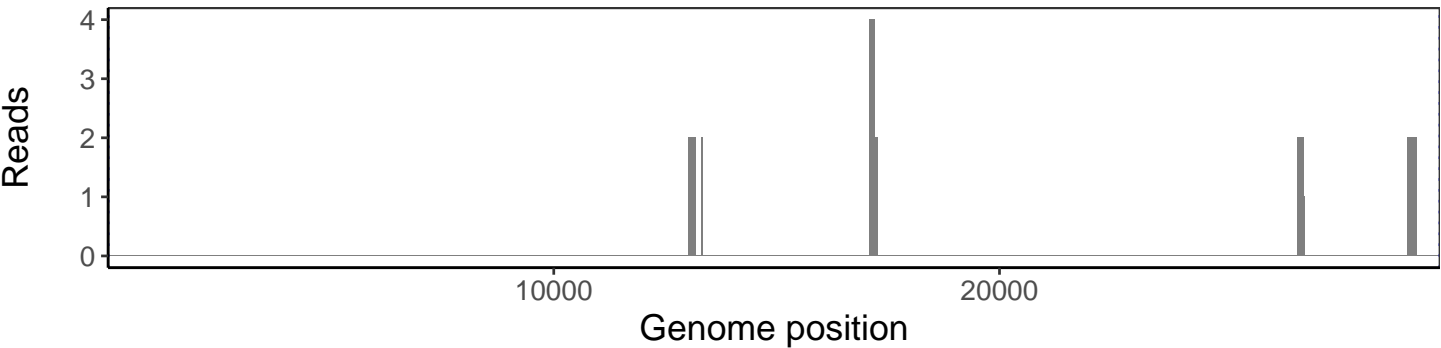
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



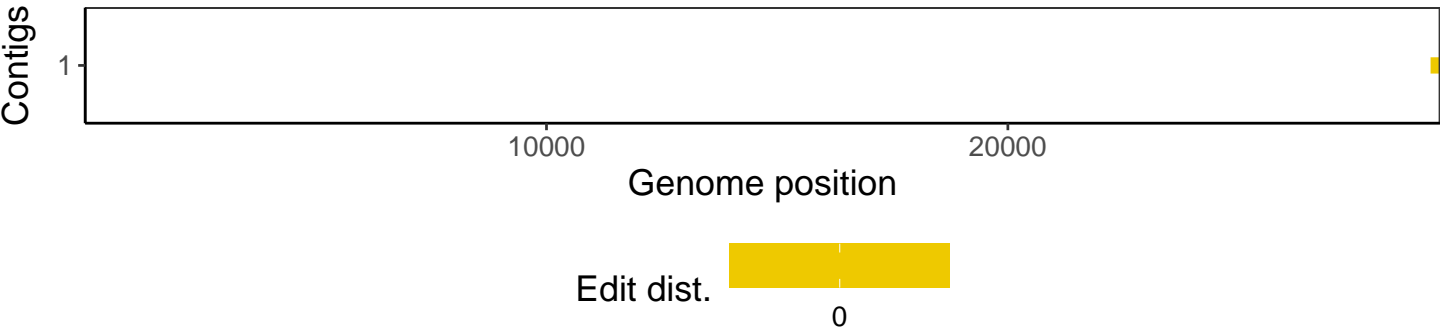
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



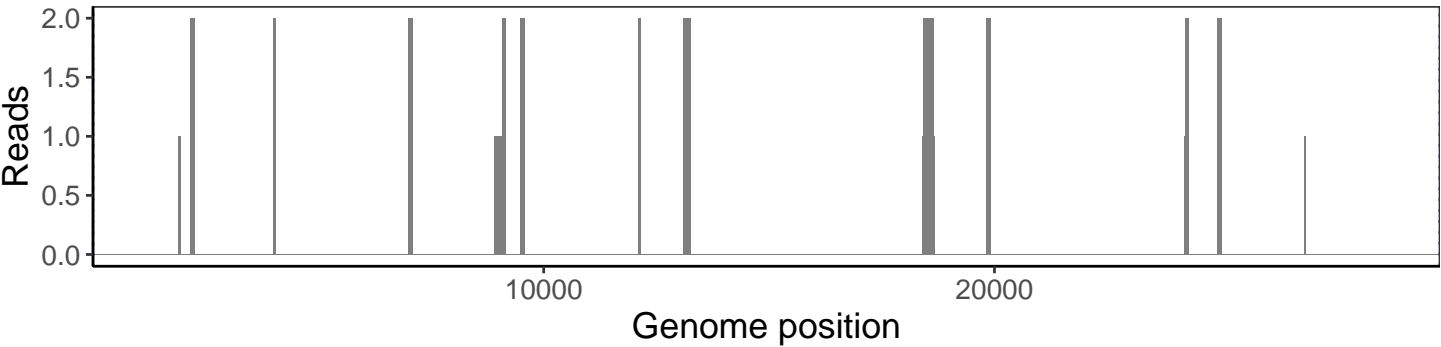
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



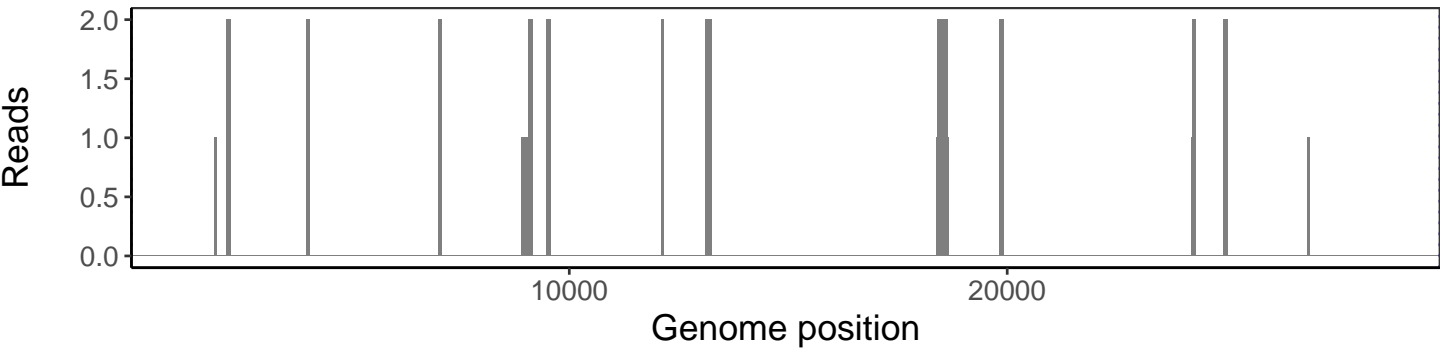
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

