

COVID-19 subject 219

2021-01-10

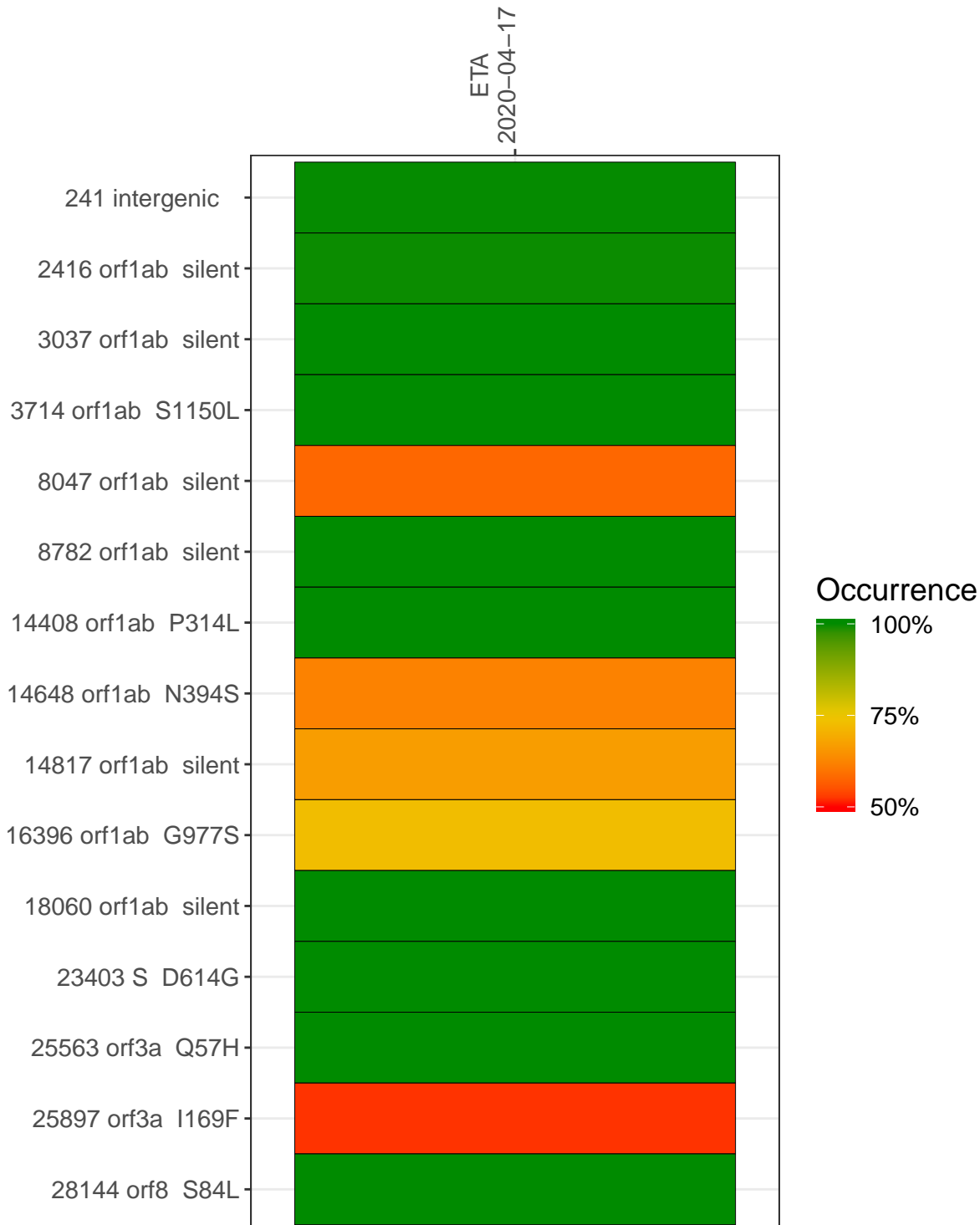
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0017	composite	NA	ETA	2020-04-17	7.66	93.3%	92.0%
VSP0017-1m	single experiment	NA	ETA	2020-04-17	0.93	39.4%	31.9%
VSP0017-2m	single experiment	NA	ETA	2020-04-17	1.23	47.3%	42.8%
VSP0017-3	single experiment	9350	ETA	2020-04-17	1.53	48.8%	40.3%
VSP0017-4	single experiment	1870	ETA	2020-04-17	1.06	49.2%	46.5%
VSP0017-5	single experiment	1870	ETA	2020-04-17	1.69	44.0%	38.6%
VSP0017-6	single experiment	1870	ETA	2020-04-17	3.40	64.2%	61.8%

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
2020-04-17

241 intergenic	20	315	13	185		240
2416 orf1ab silent	32	100				241
3037 orf1ab silent		1	27	38		
3714 orf1ab S1150L					18	
8047 orf1ab silent					268	368
8782 orf1ab silent					359	680
14408 orf1ab P314L	26	120				
14648 orf1ab N394S				200		327
14817 orf1ab silent				180		361
16396 orf1ab G977S	15	43				154
18060 orf1ab silent						405
23403 S D614G	12	408				365
25563 orf3a Q57H	36	86		160	255	243
28144 orf8 S84L	8	27	20	23	232	473
	VSP0017-1m	VSP0017-2m	VSP0017-3	VSP0017-4	VSP0017-5	VSP0017-6

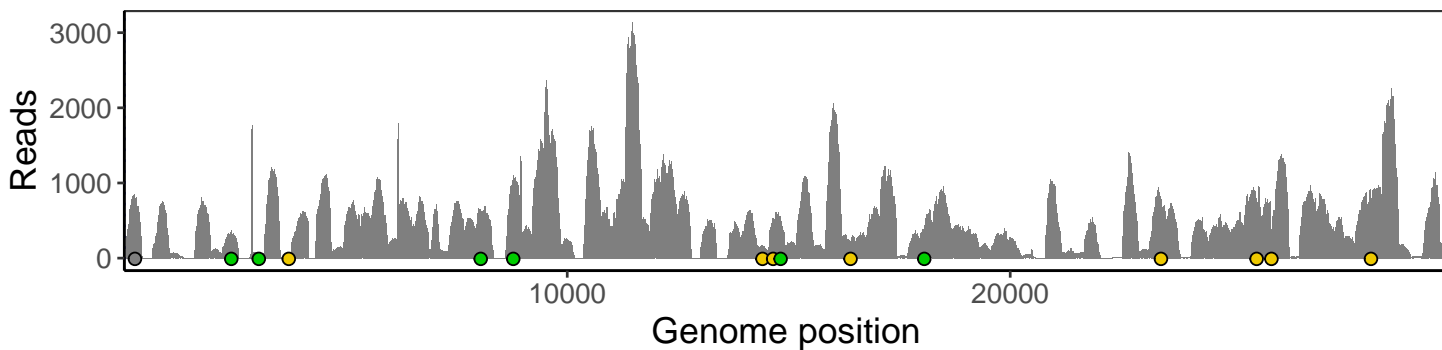
Base change

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

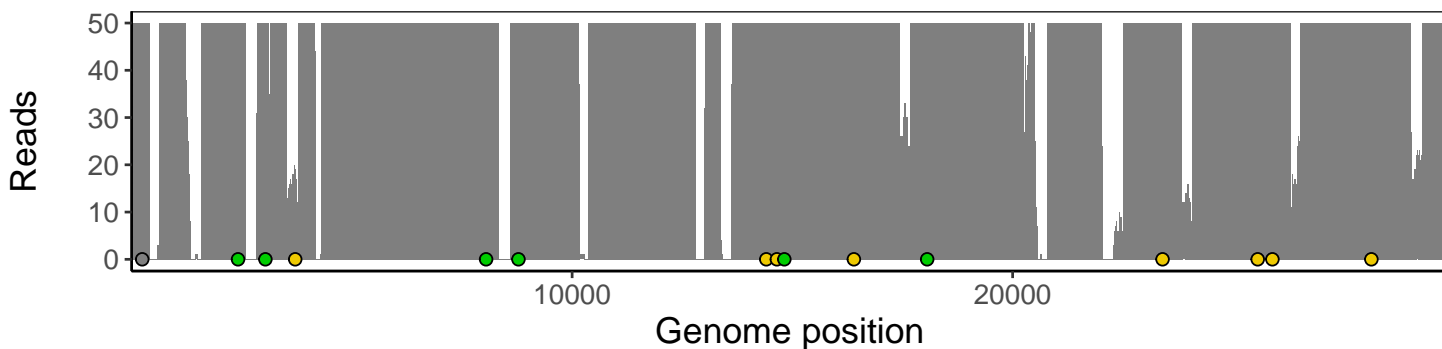
Analyses of individual experiments and composite results.

VSP0017 | 2020-04-17 | ETA | 219-tri | composite result

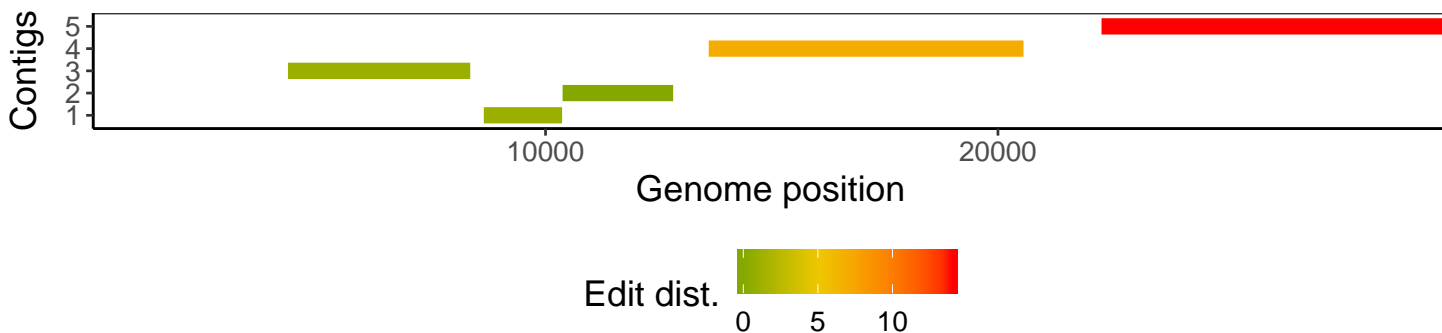
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



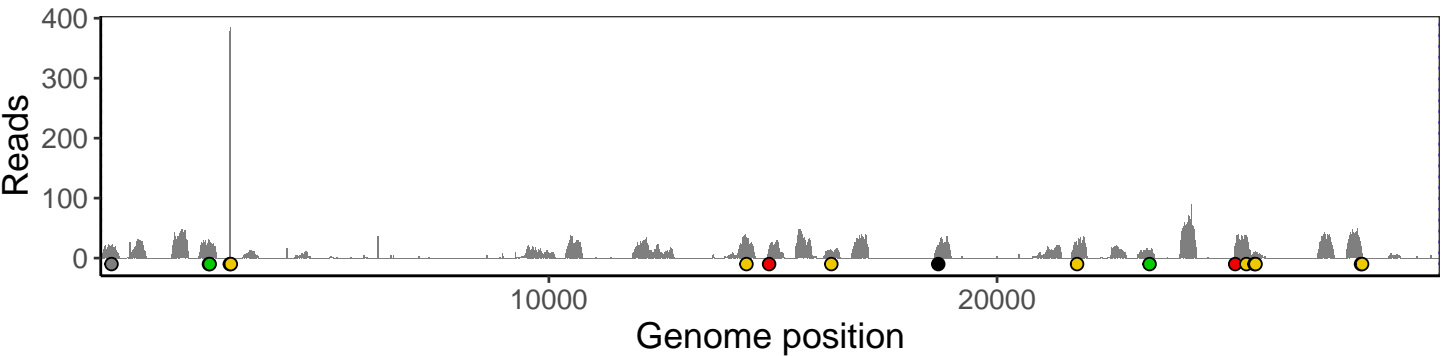
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



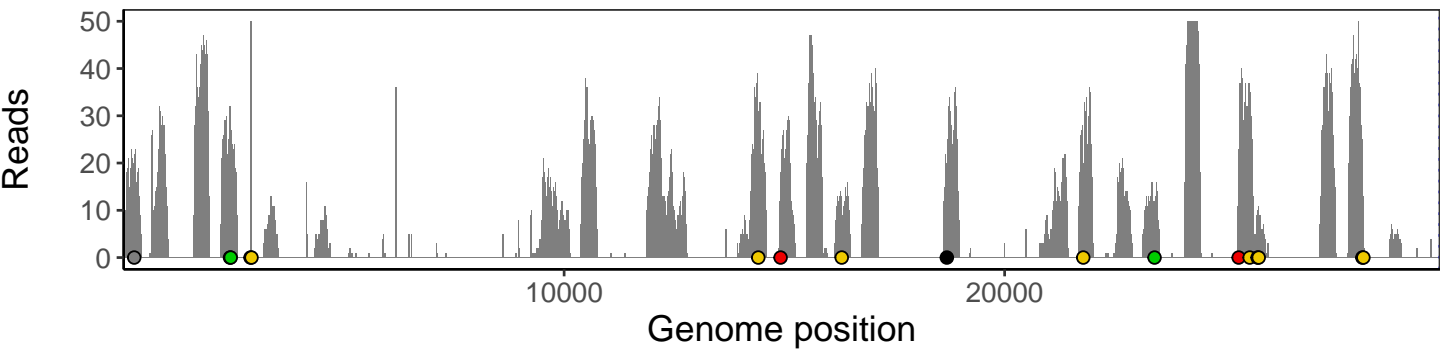
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



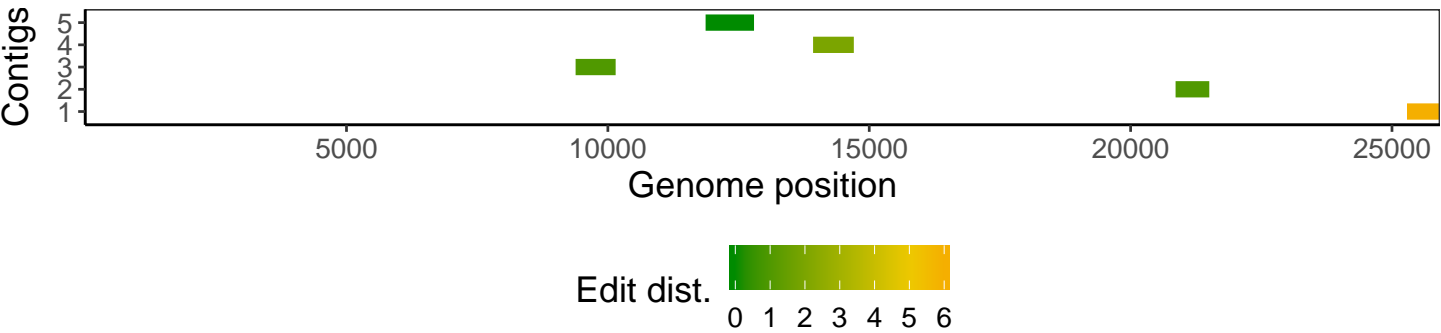
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



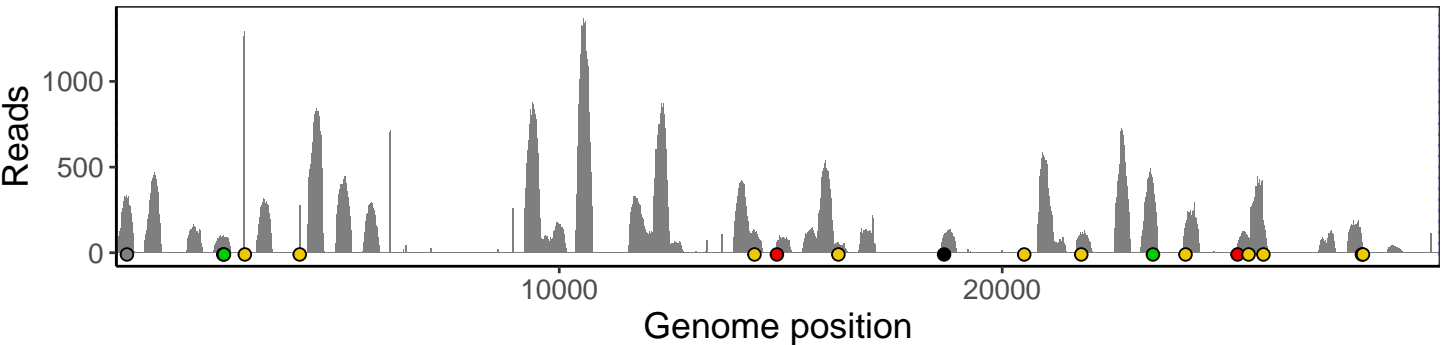
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



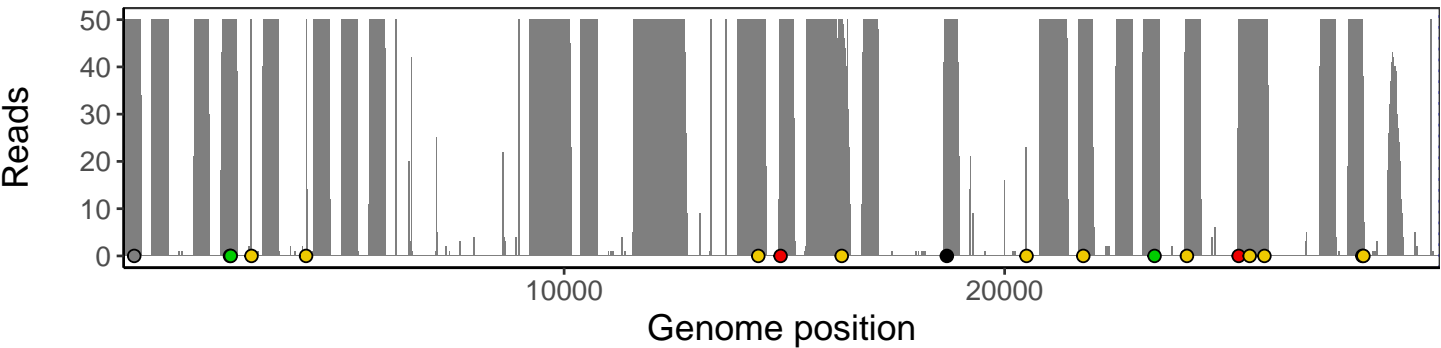
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.



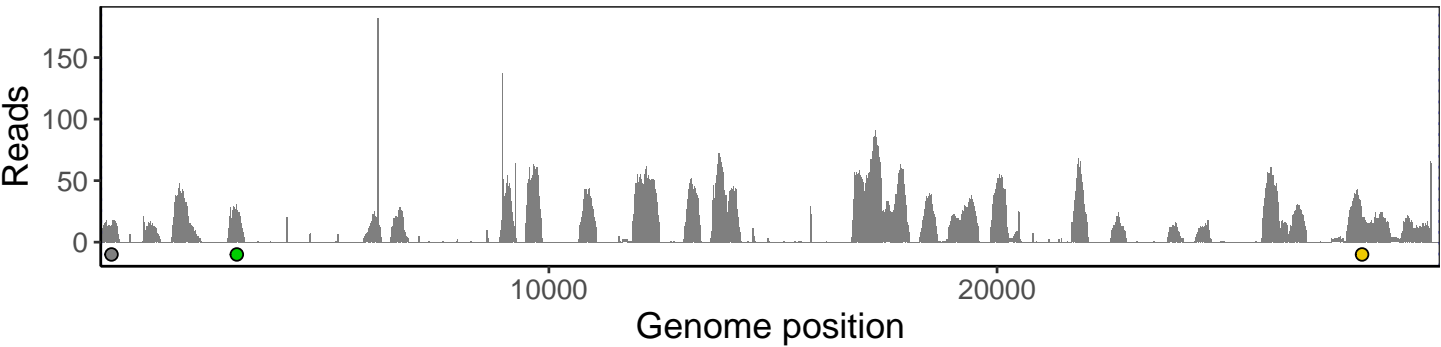
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



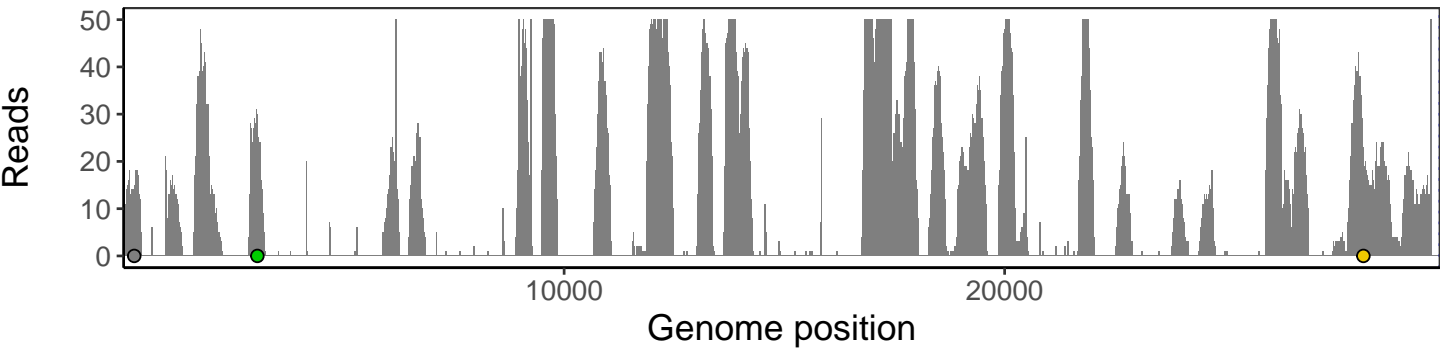
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



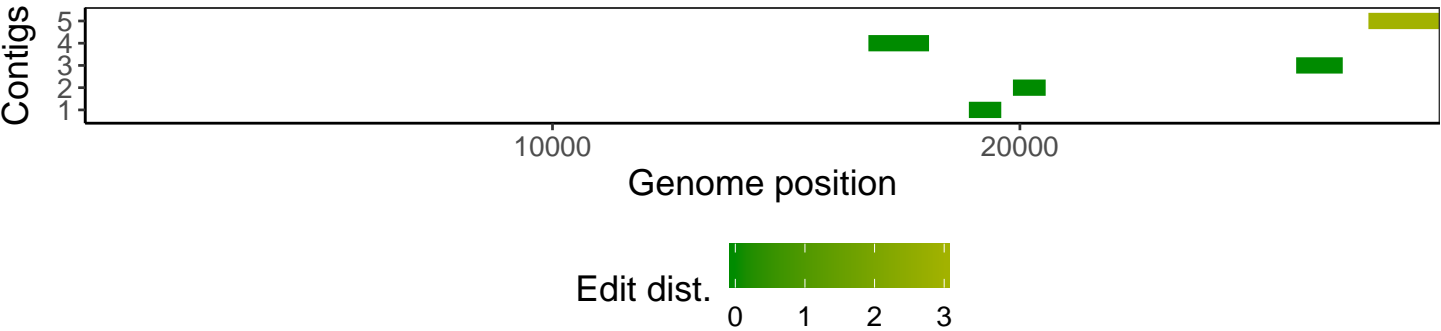
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



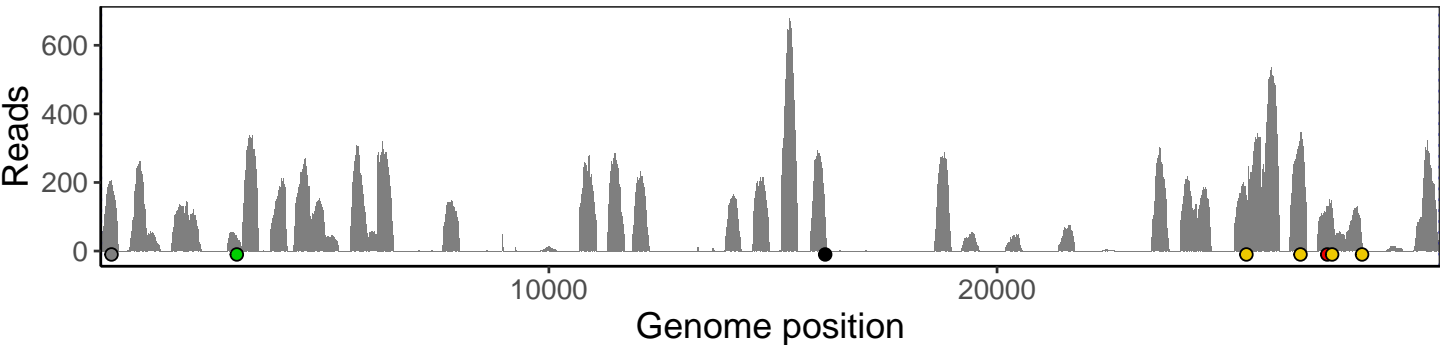
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



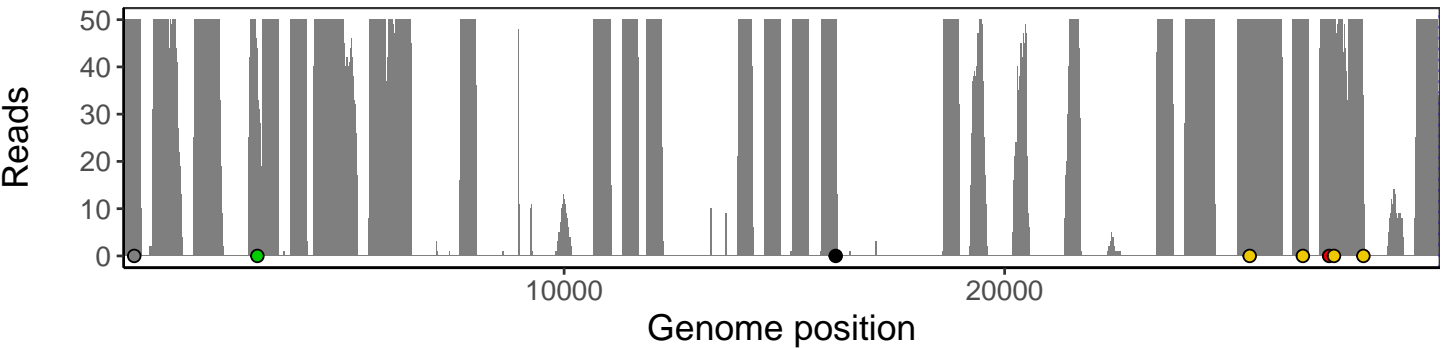
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



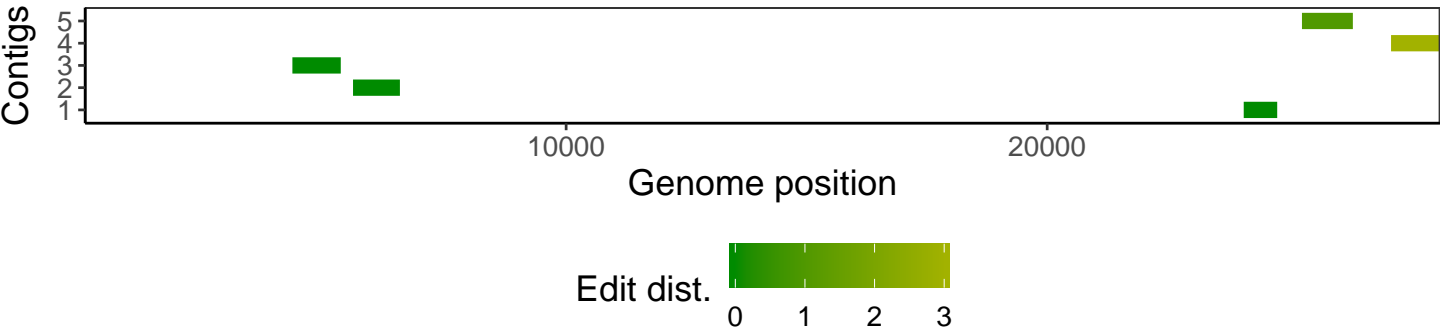
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



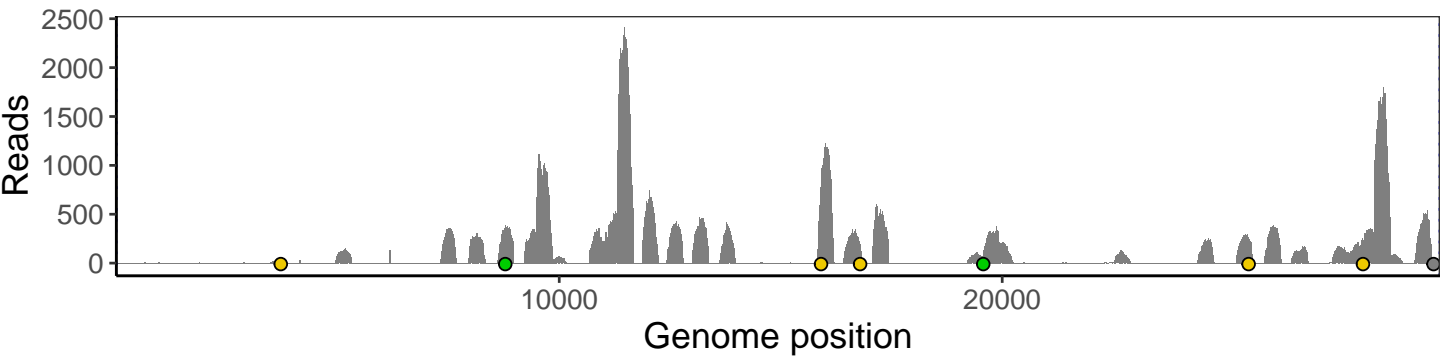
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



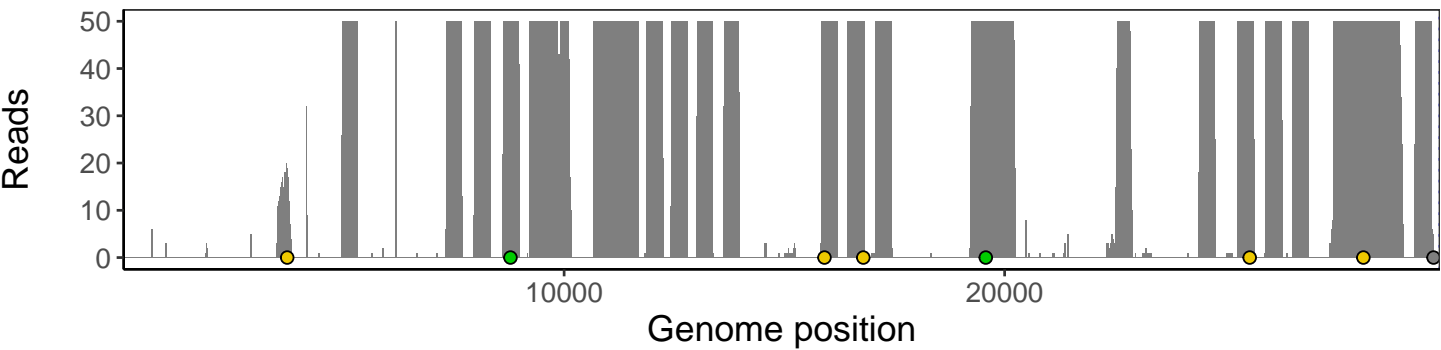
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



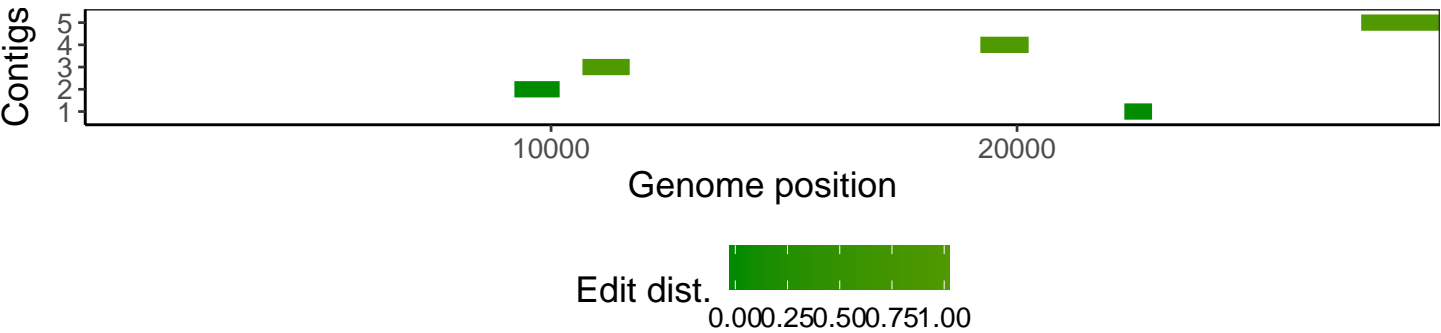
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



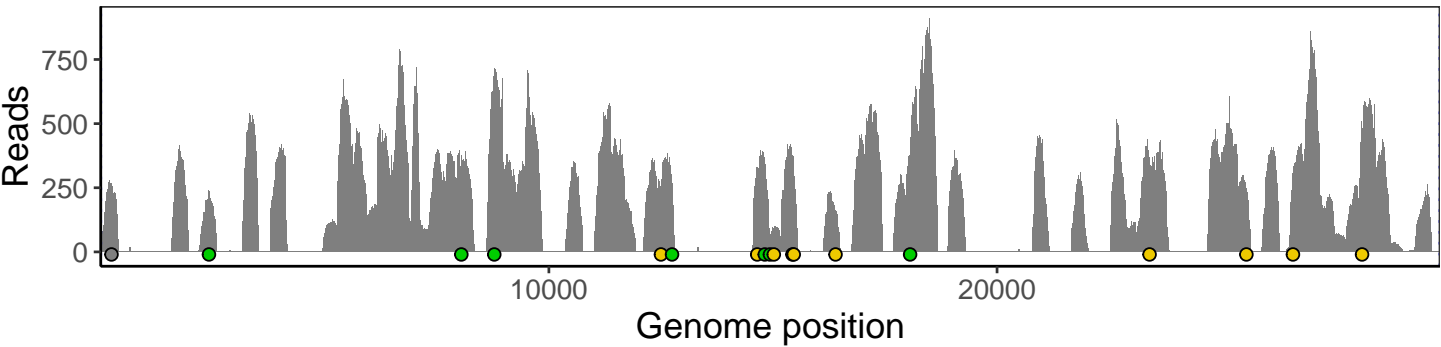
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



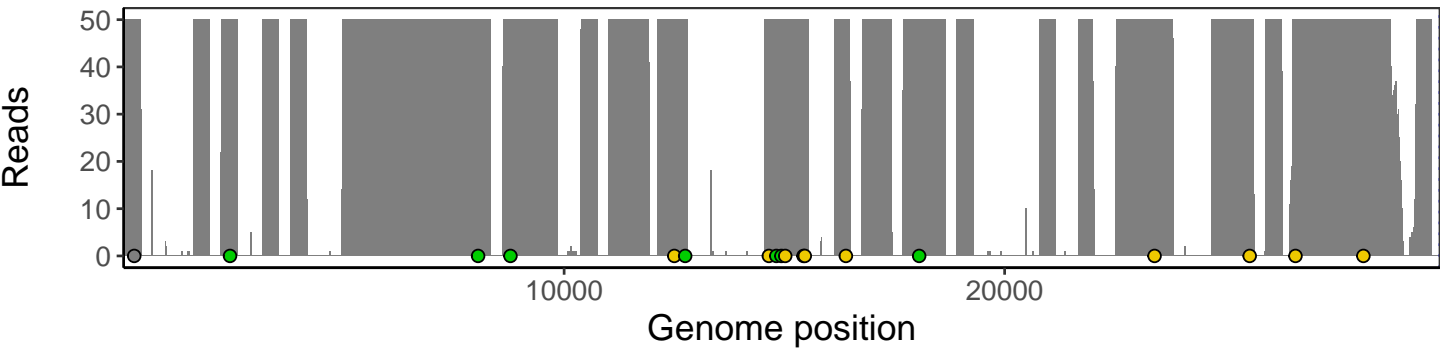
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

