

COVID-19 subject 219

2020-09-14

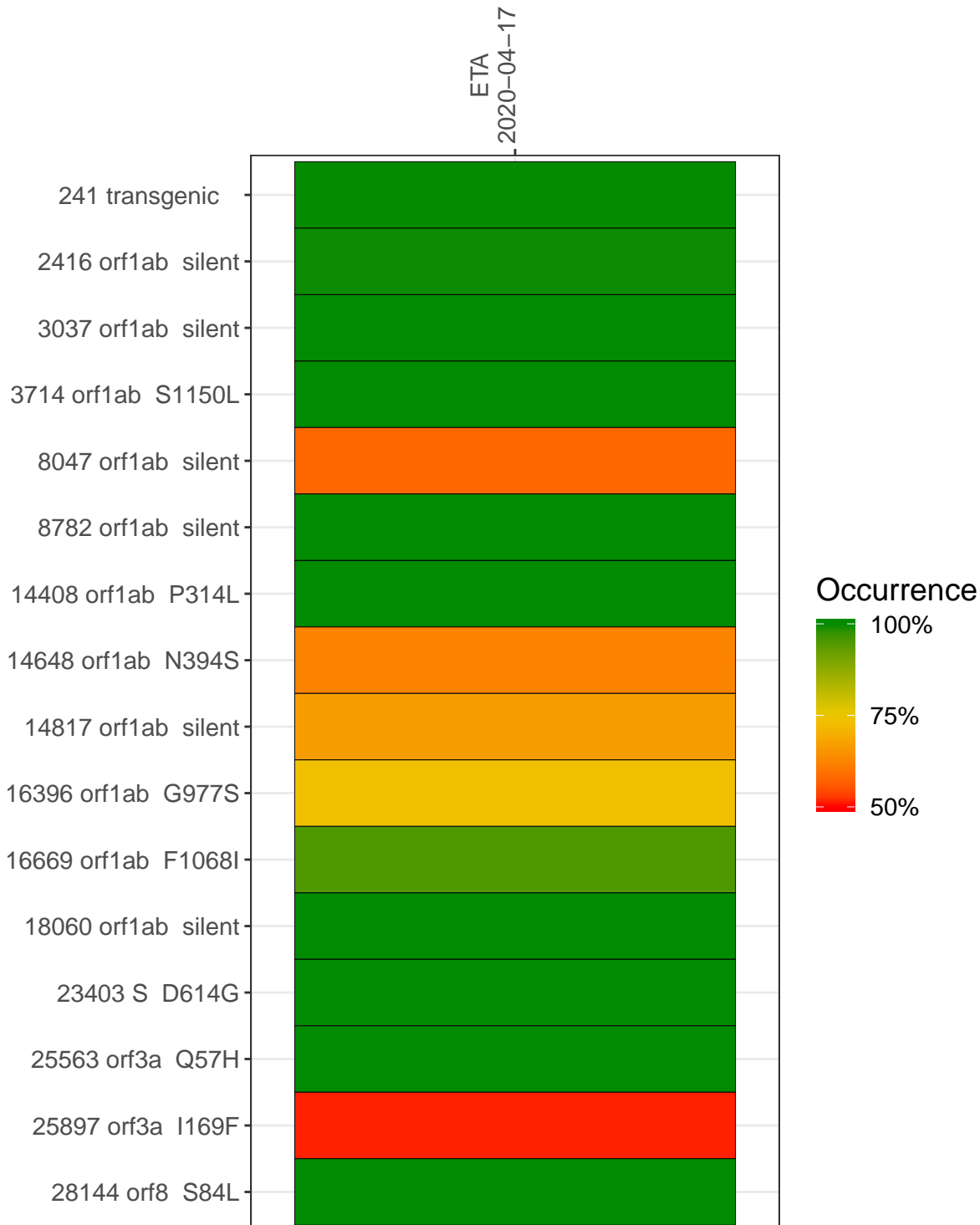
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

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	93.1%	92.0%
VSP0017-1m	single experiment	NA	ETA	2020-04-17	0.93	39.3%	31.2%
VSP0017-2m	single experiment	NA	ETA	2020-04-17	1.23	46.3%	42.8%
VSP0017-3	single experiment	9350	ETA	2020-04-17	1.53	48.6%	40.2%
VSP0017-4	single experiment	1870	ETA	2020-04-17	1.02	49.1%	46.5%
VSP0017-5	single experiment	1870	ETA	2020-04-17	1.69	43.2%	38.4%
VSP0017-6	single experiment	1870	ETA	2020-04-17	3.40	64.1%	61.6%

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 transgenic	20	309	13	183		237
2416 orf1ab silent	32	99				238
3037 orf1ab silent			27	38		
3714 orf1ab S1150L					17	
8047 orf1ab silent					267	364
8782 orf1ab silent					357	674
14408 orf1ab P314L	26	118				
14648 orf1ab N394S				198		325
14817 orf1ab silent				178		355
16396 orf1ab G977S	14	42				153
16669 orf1ab F1068I						
18060 orf1ab silent						402
23403 S D614G	10	399				360
25563 orf3a Q57H	36	85		160	253	239
25897 orf3a I169F	2	180		175		
28144 orf8 S84L	8	25	20	23	227	464
	VSP0017-1m	VSP0017-2m	VSP0017-3	VSP0017-4	VSP0017-5	VSP0017-6

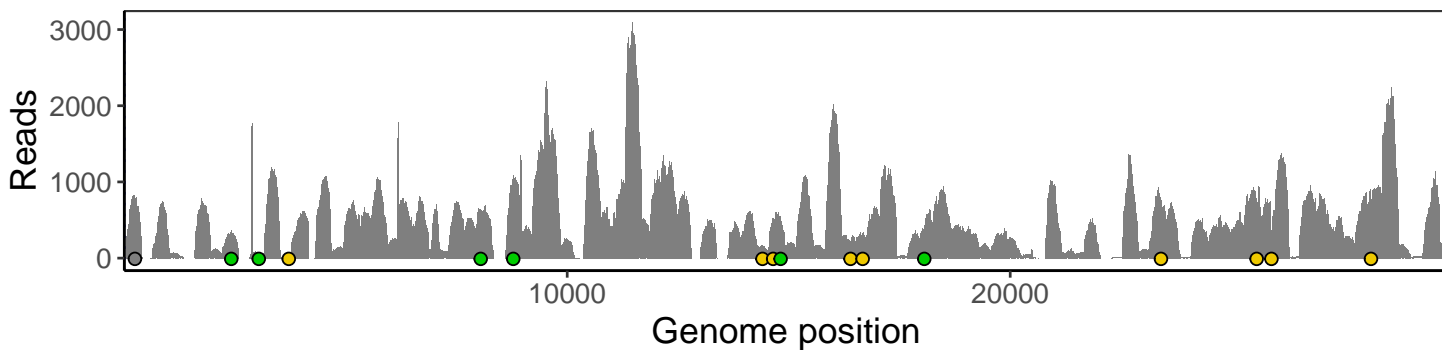
Base change



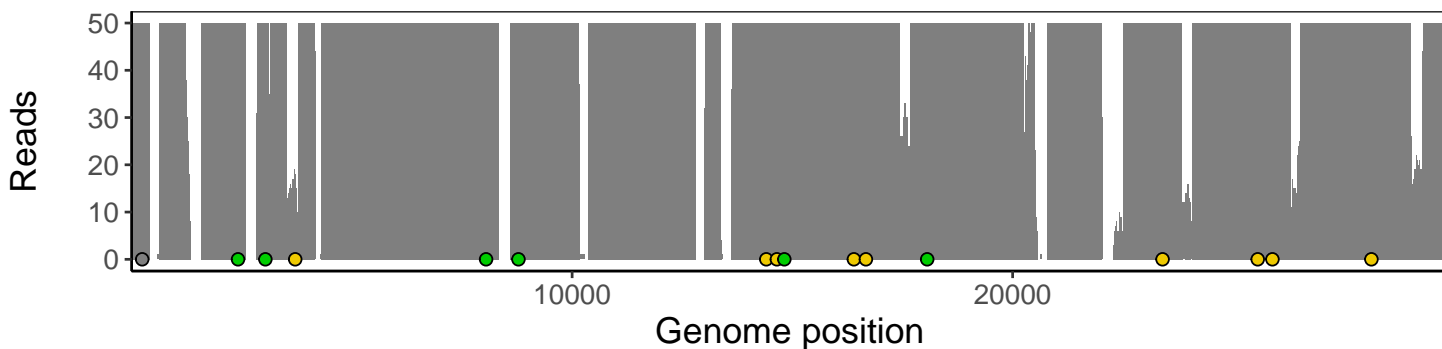
Analyses of individual experiments and composite results.

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

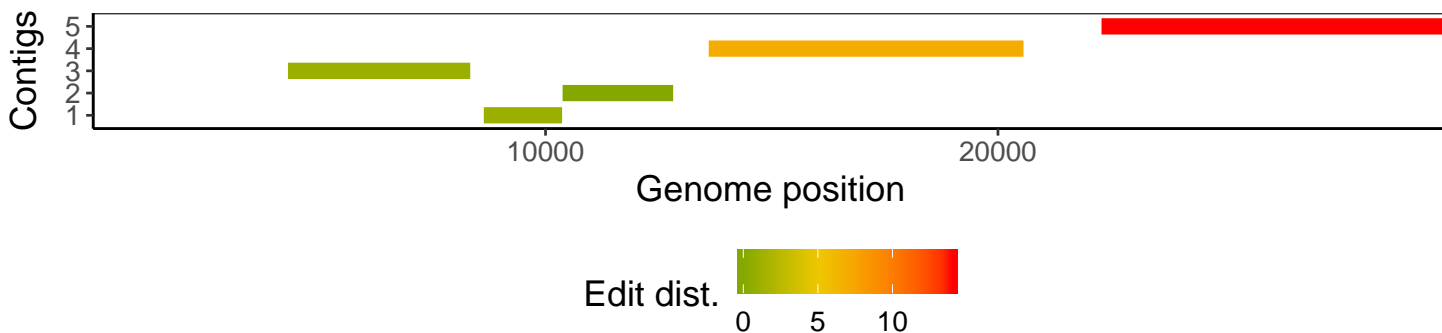
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according to variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



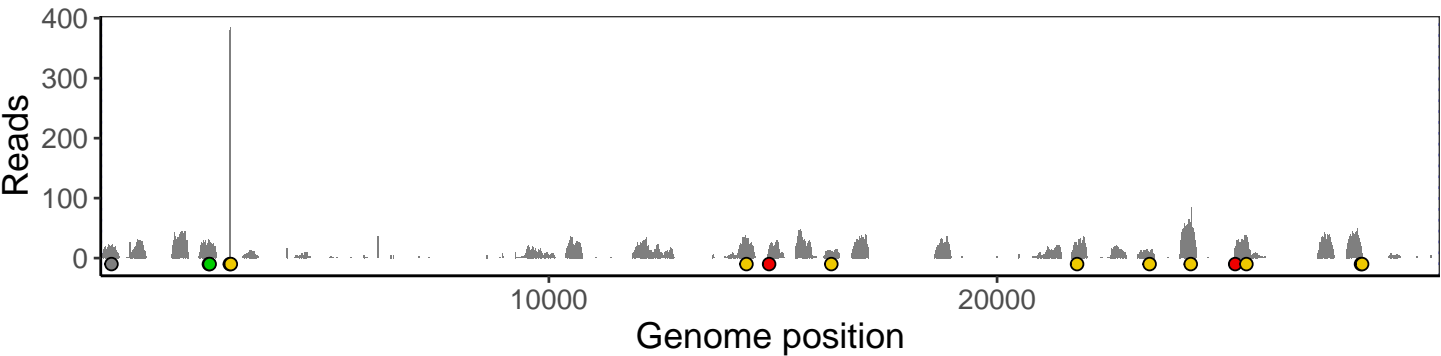
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



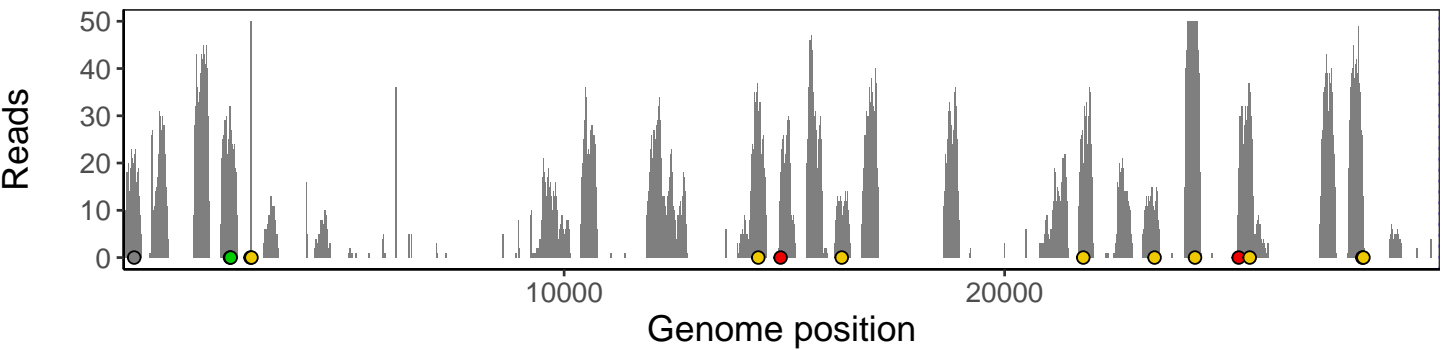
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



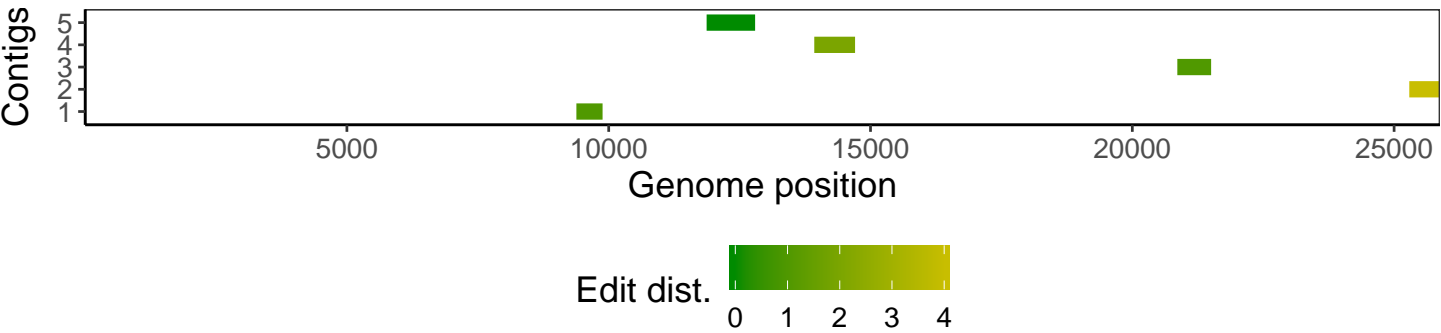
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



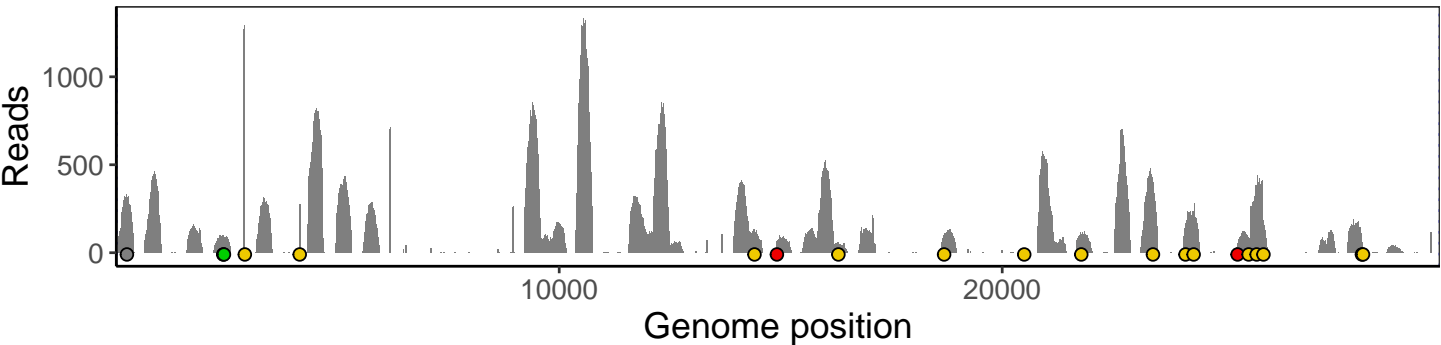
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



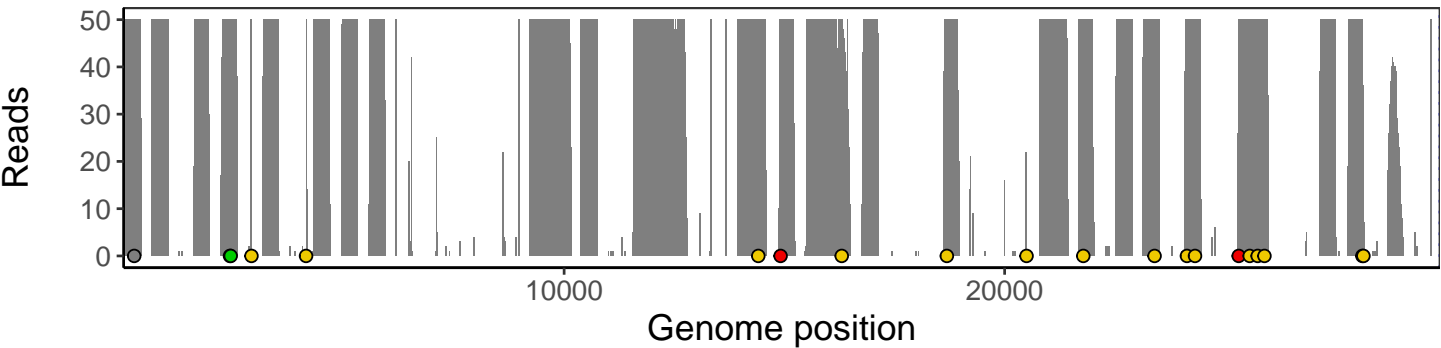
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.



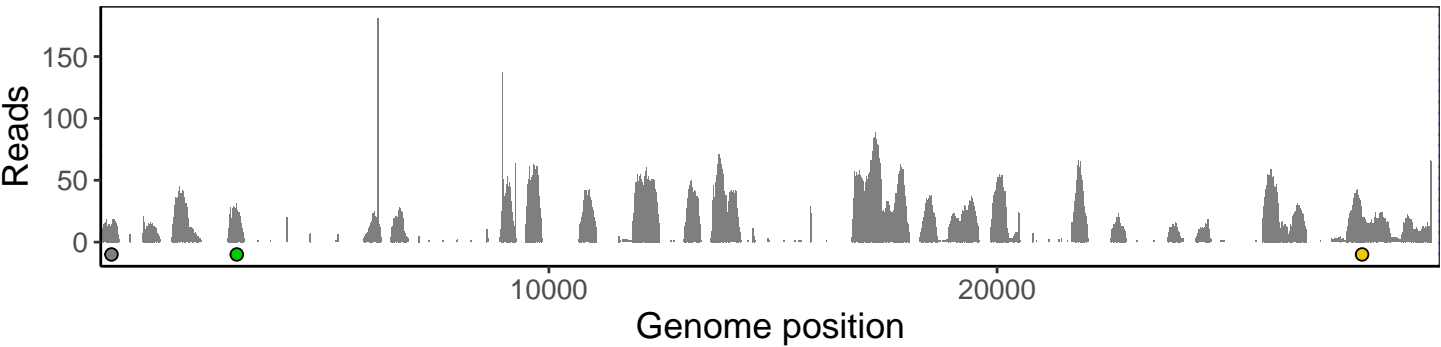
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



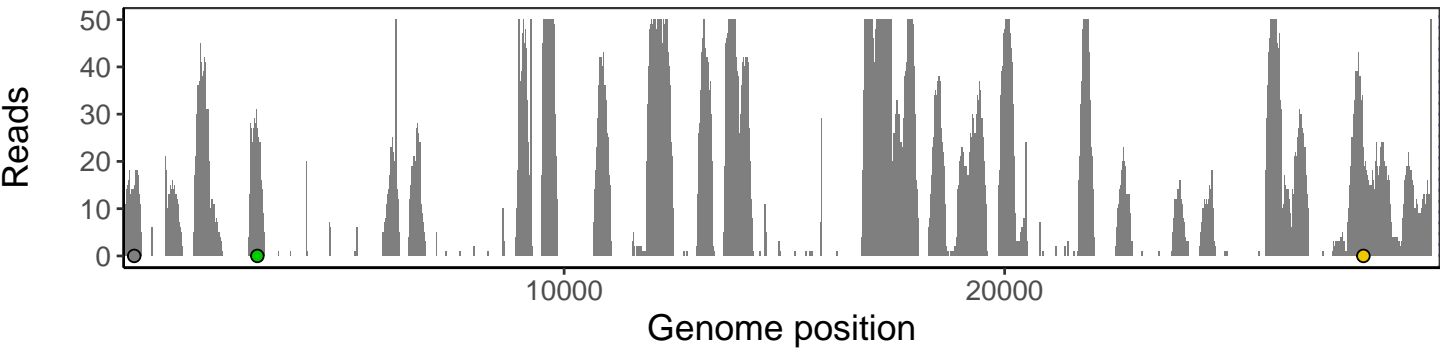
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



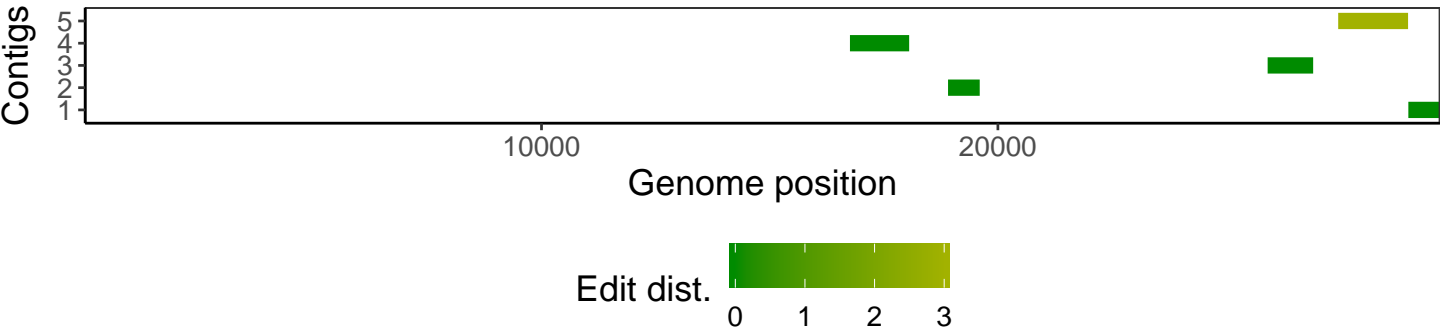
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



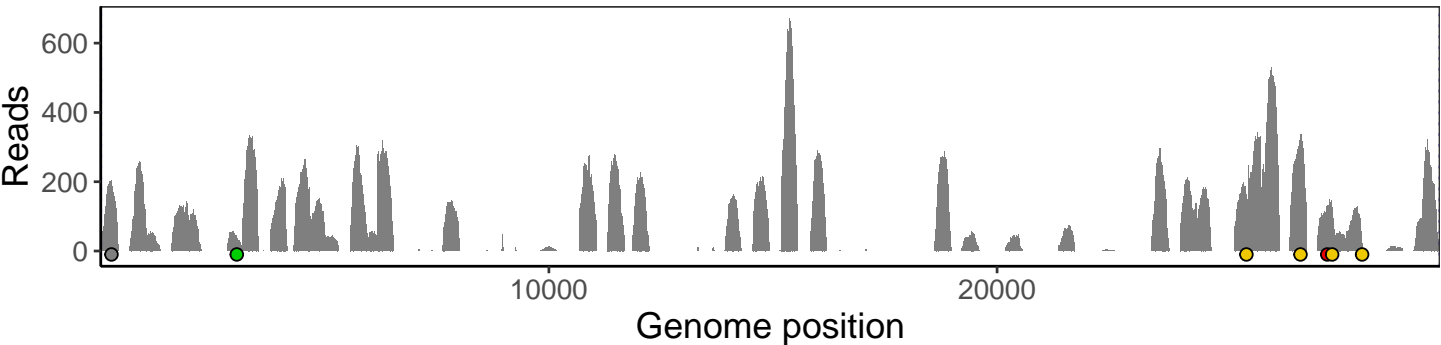
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



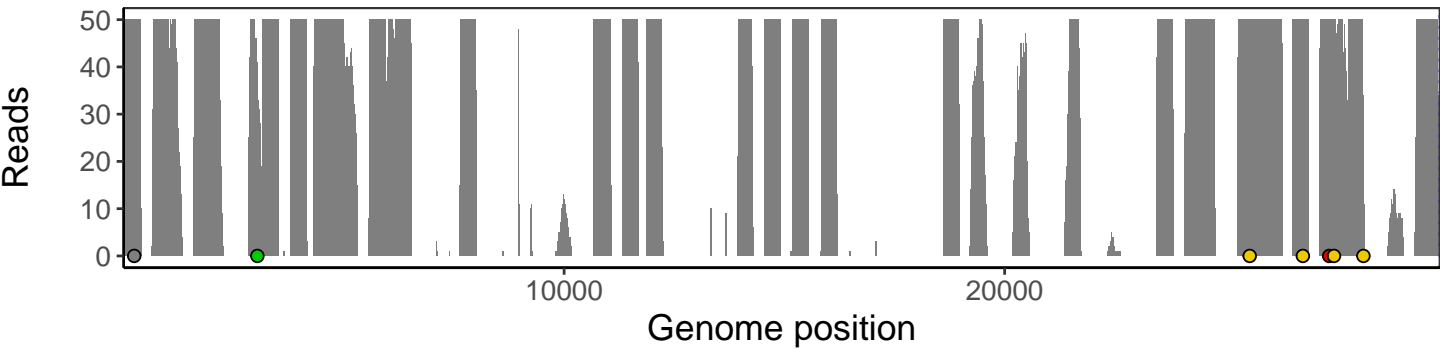
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



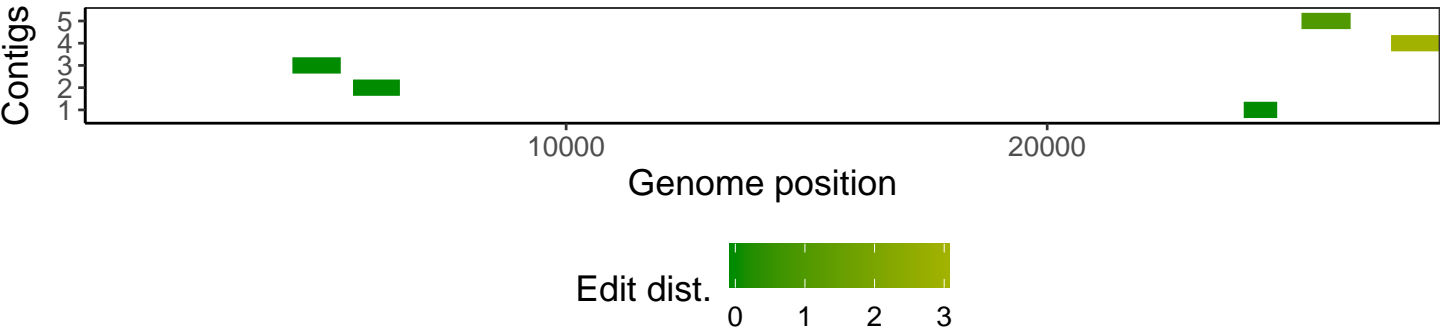
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



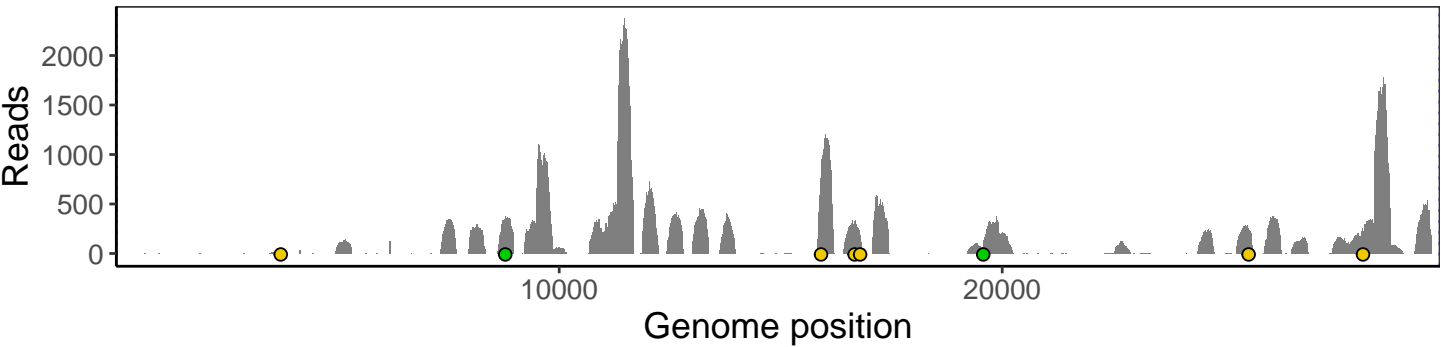
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



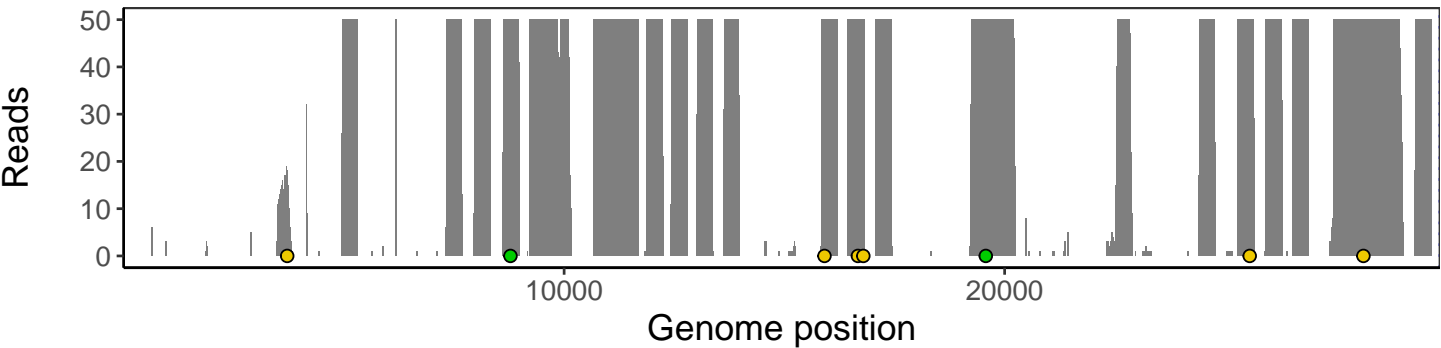
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



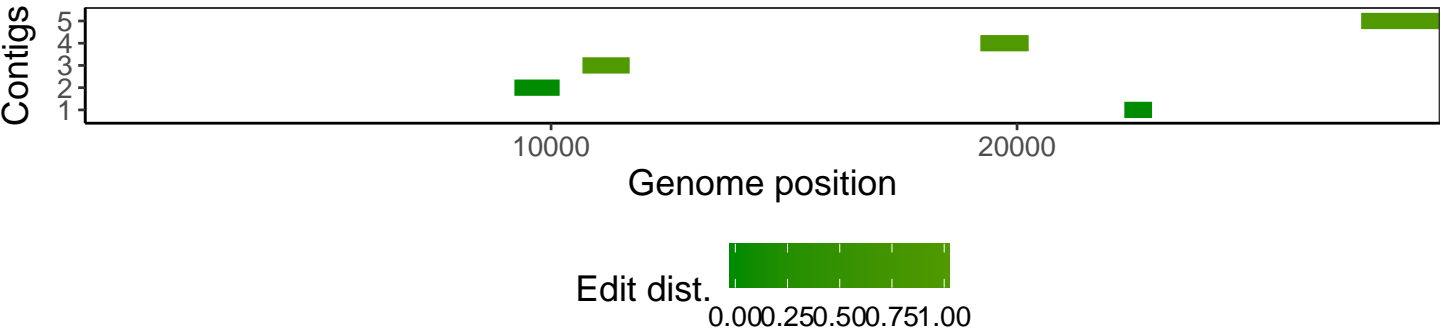
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



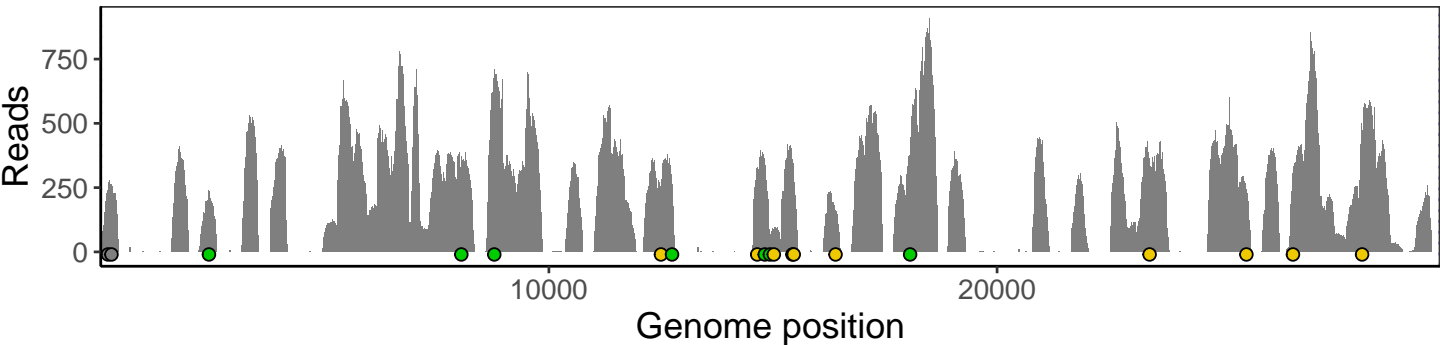
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



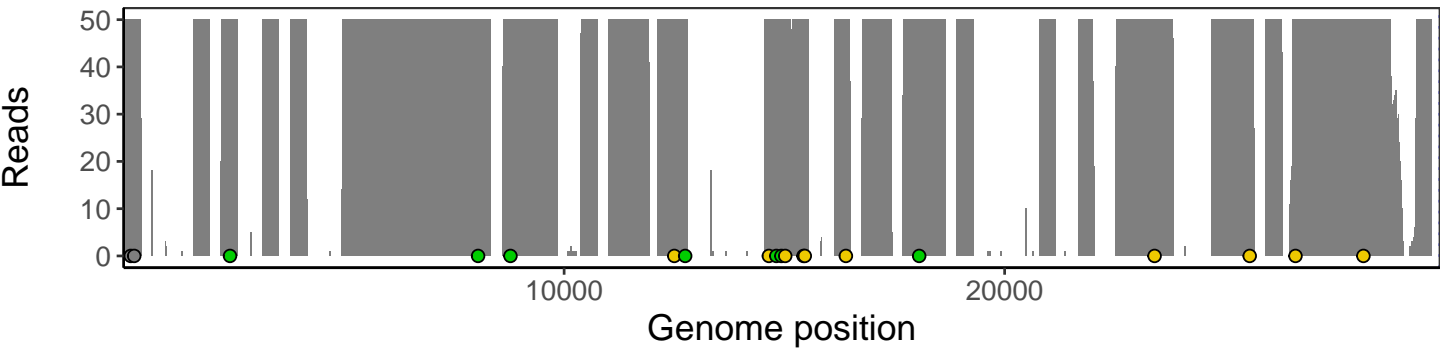
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

