

COVID-19 subject 222-TCE

2021-01-10

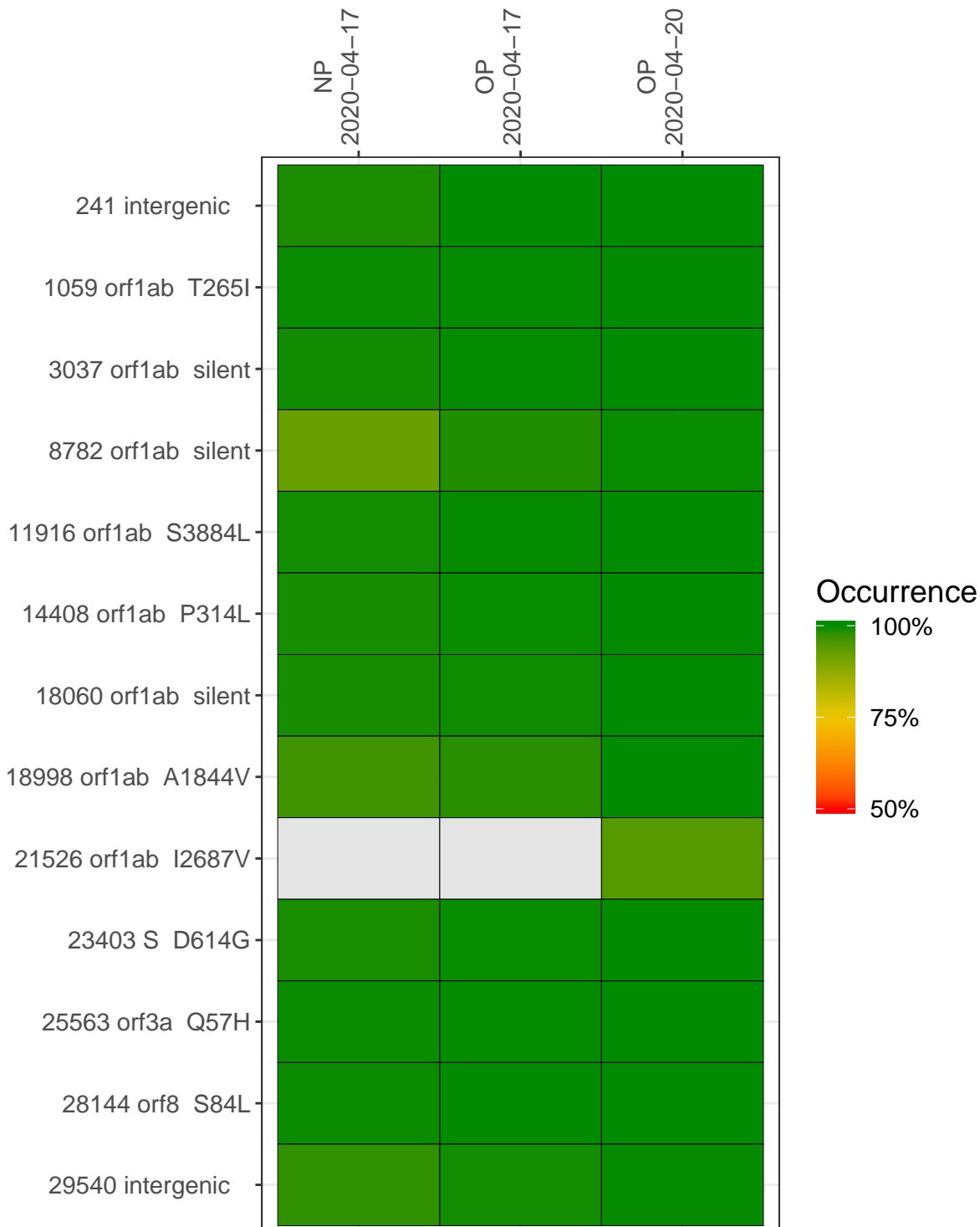
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0181	composite	NA	OP	2020-04-17	29.34	99.9%	99.9%
VSP0182	composite	NA	OP	2020-04-20	29.57	99.9%	99.9%
VSP0180-1m	single experiment	NA	NP	2020-04-17	29.87	99.9%	99.9%
VSP0181-1m	single experiment	NA	OP	2020-04-17	30.04	99.9%	99.9%
VSP0181-2	single experiment	1.510e+11	OP	2020-04-17	29.34	99.9%	99.7%
VSP0182-1m	single experiment	NA	OP	2020-04-20	29.87	99.9%	99.9%
VSP0182-2	single experiment	1.325e+11	OP	2020-04-20	20.33	99.9%	99.4%

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 2020-04-17	OP 2020-04-17		OP 2020-04-20	
241 intergenic	801	731	596	802	738
1059 orf1ab T265I	3584	5390	575	3018	685
3037 orf1ab silent	1327	3597	727	1579	681
8782 orf1ab silent	115	302	493	308	221
11916 orf1ab S3884L	628	2421	265	982	120
14408 orf1ab P314L	968	3338	478	1308	121
18060 orf1ab silent	810	3216	418	1188	266
18998 orf1ab A1844V	90	268	408	343	143
21526 orf1ab I2687V	15	36	32	37	16
23403 S D614G	1566	1999	1055	1905	1084
25563 orf3a Q57H	2815	5020	831	2784	1081
28144 orf8 S84L	362	979	334	669	229
29540 intergenic	192	861	418	360	366
	VSP0180-1m	VSP0181-1m	VSP0181-2	VSP0182-1m	VSP0182-2

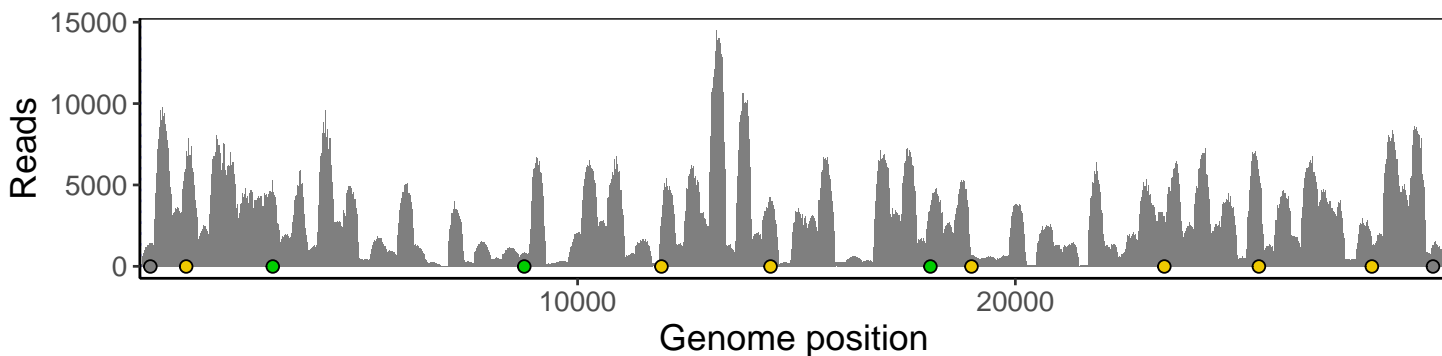
Base change

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

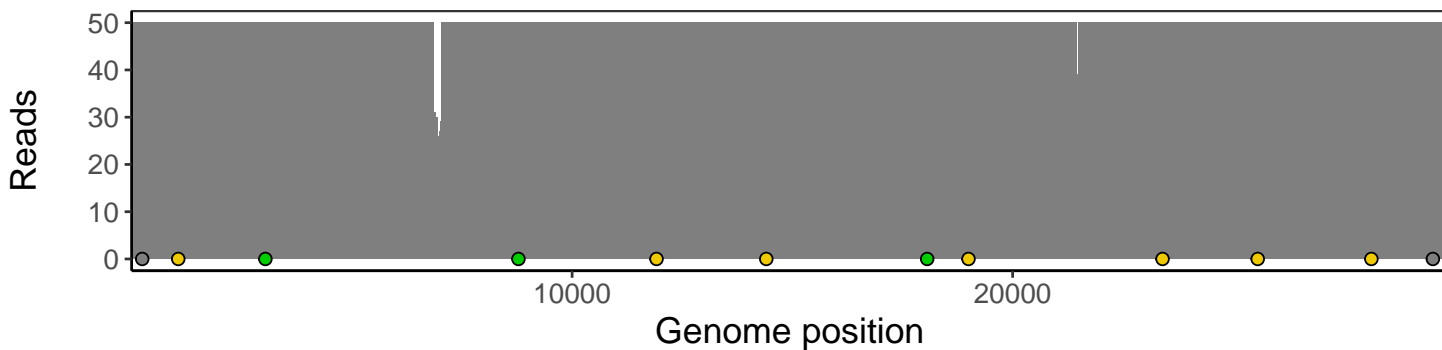
Analyses of individual experiments and composite results.

VSP0181 | 2020-04-17 | OP | 2 | composite result

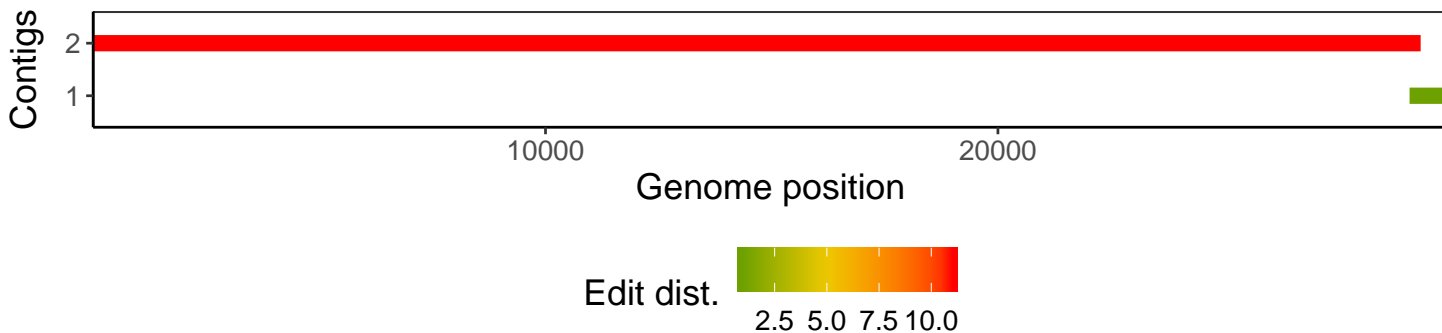
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



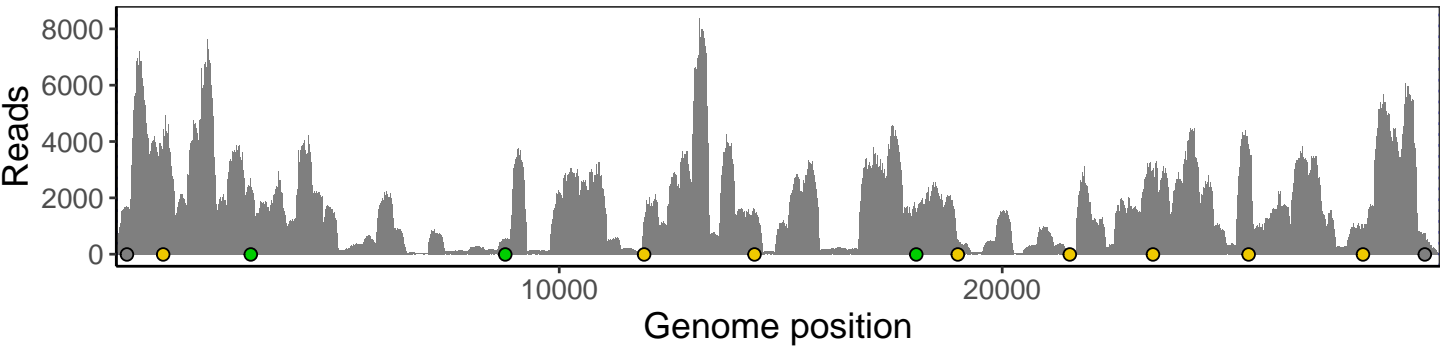
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



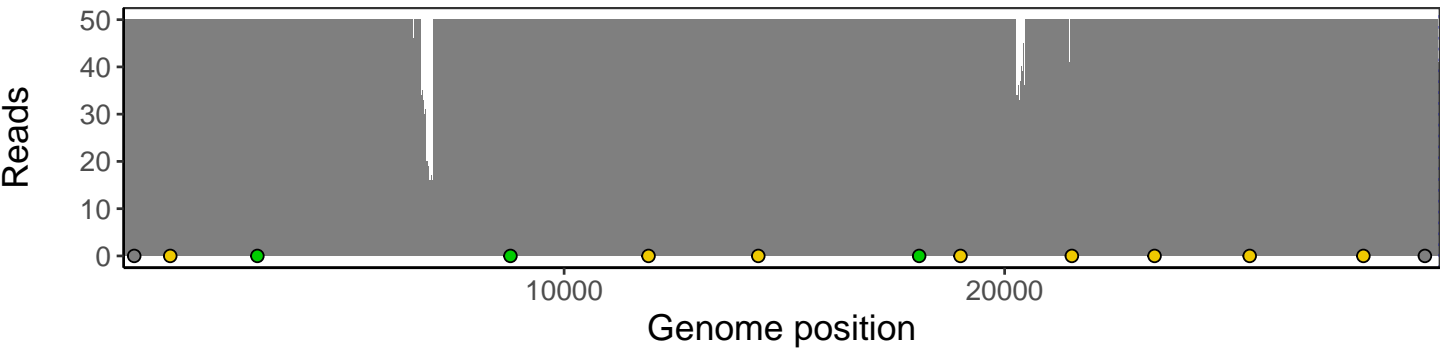
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



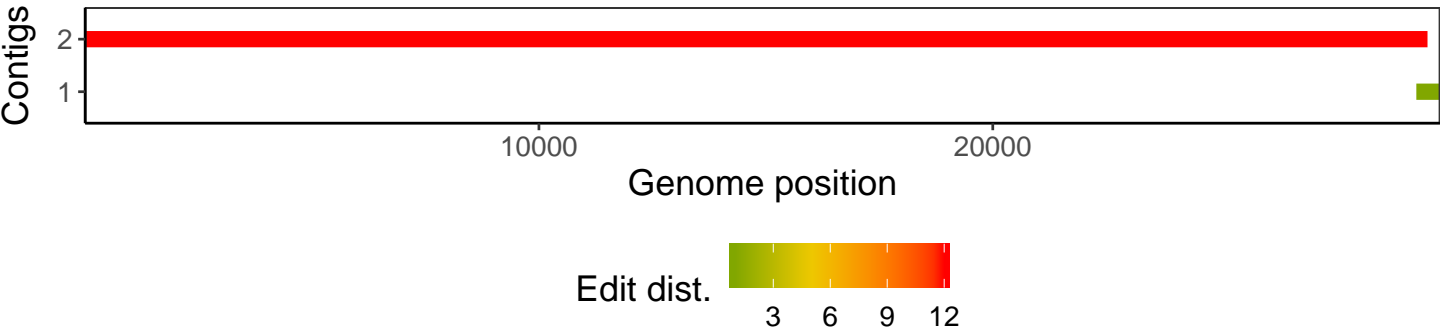
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



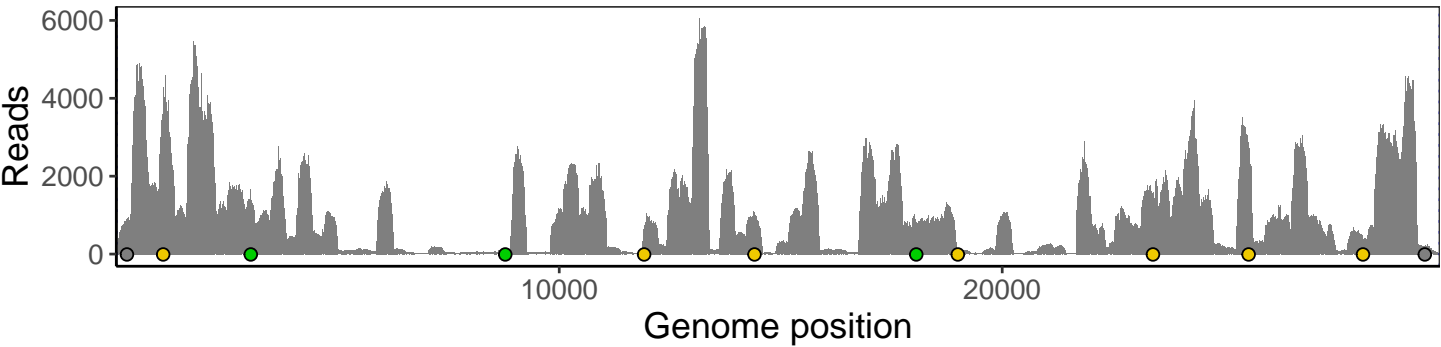
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



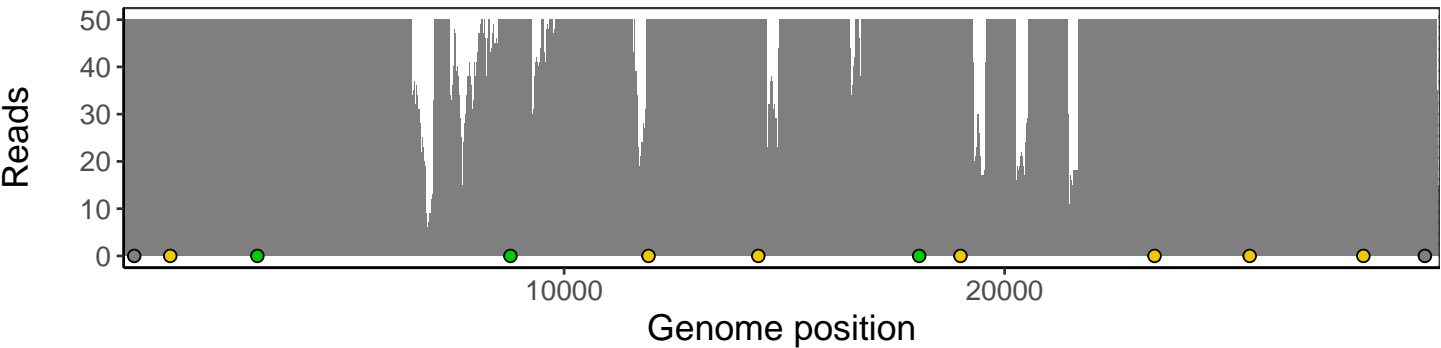
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



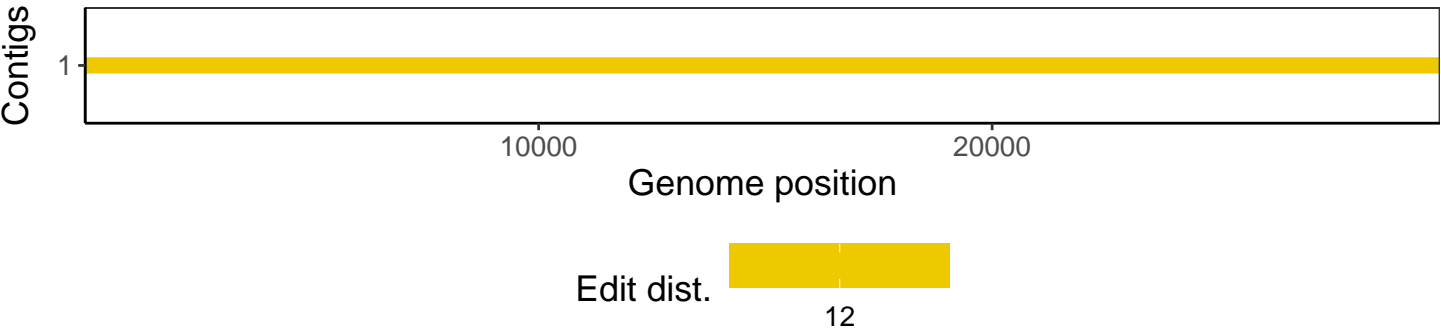
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



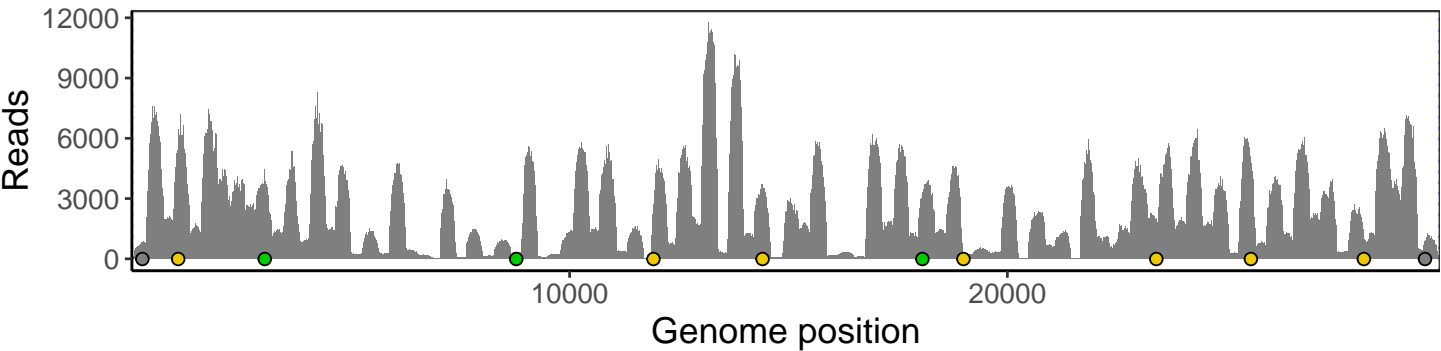
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



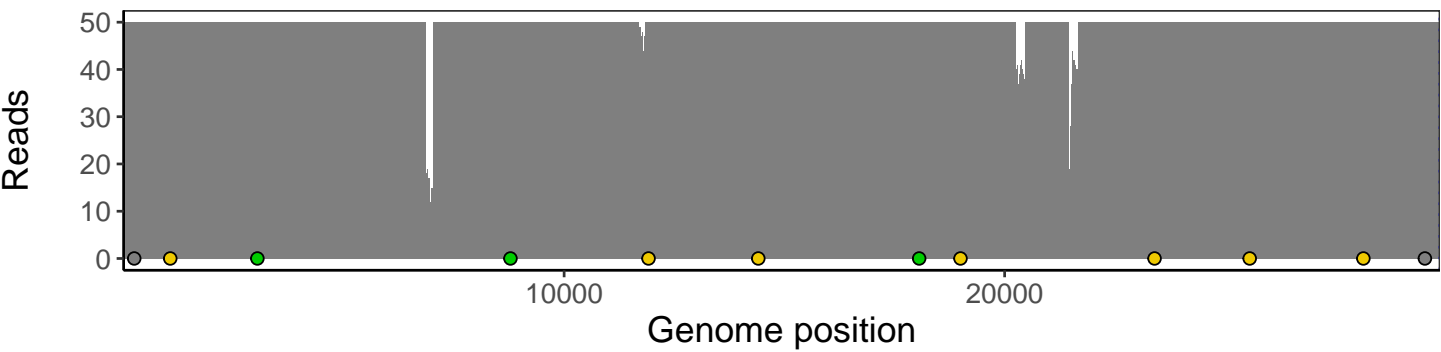
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



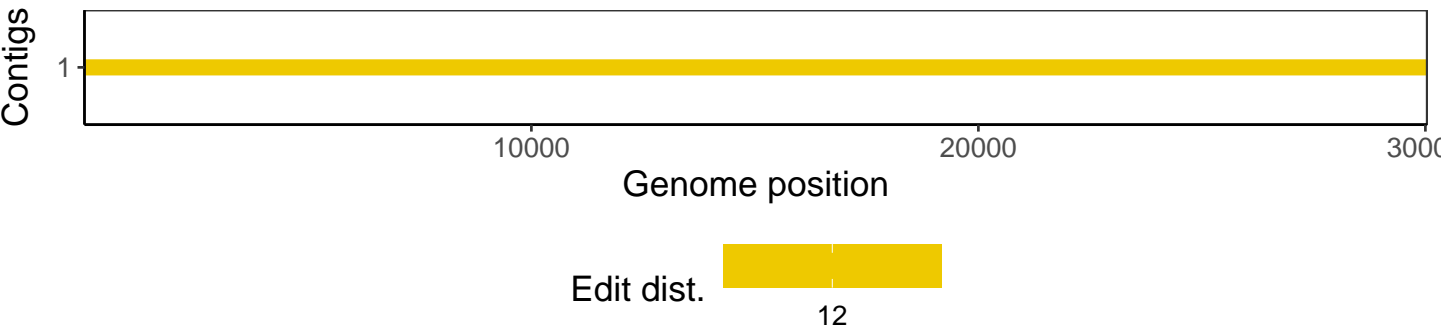
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



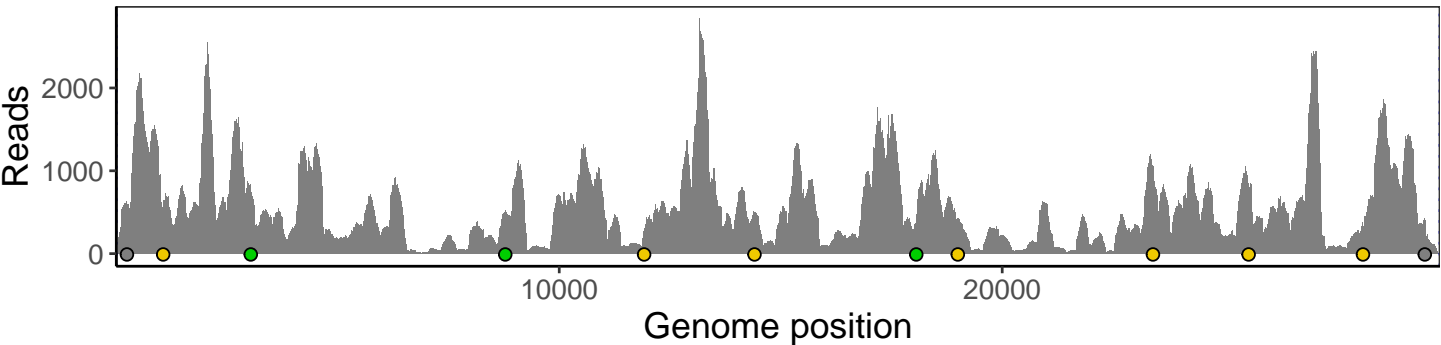
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



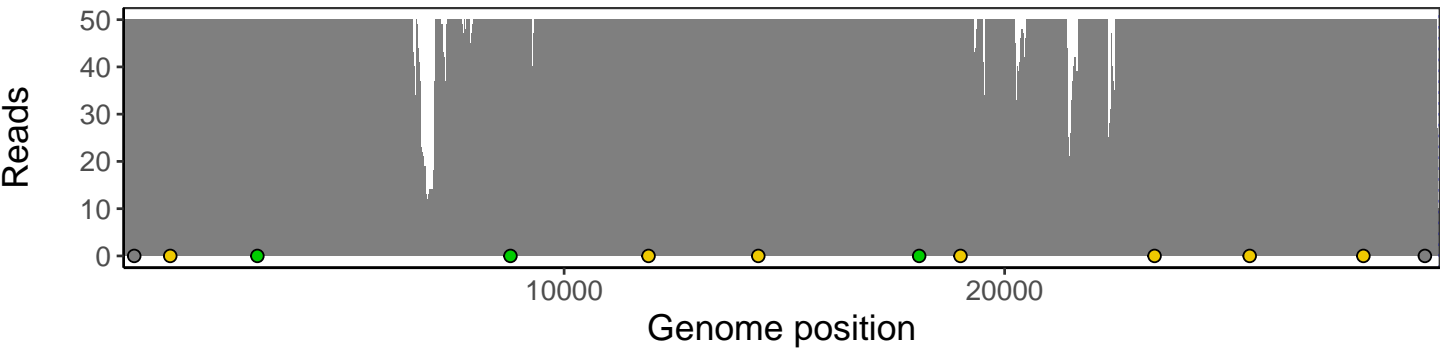
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



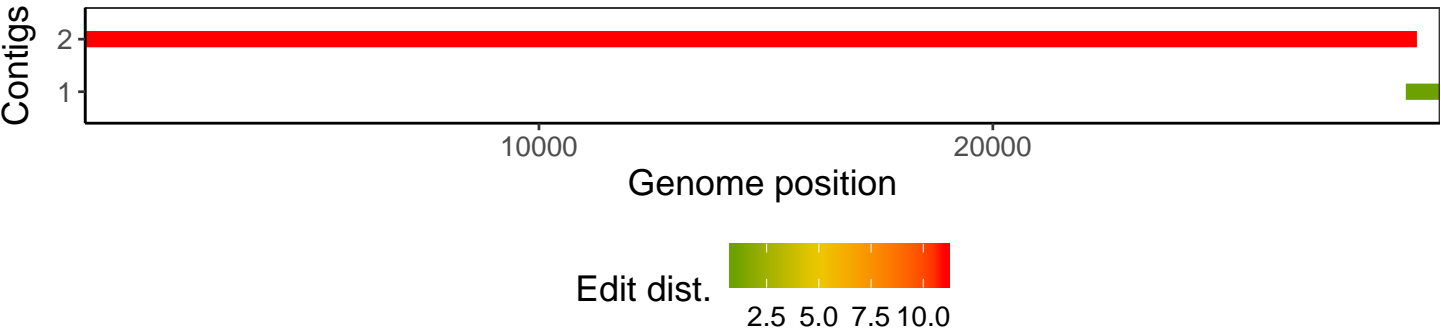
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



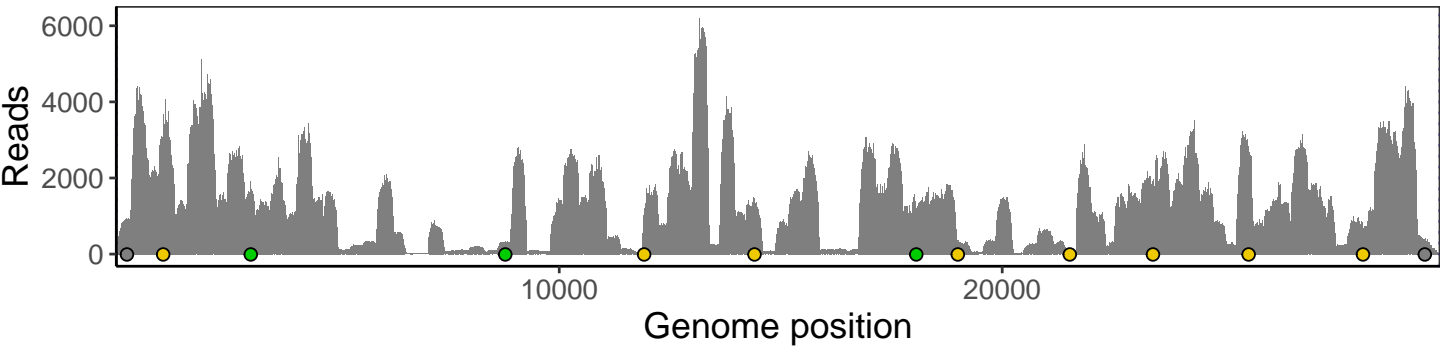
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



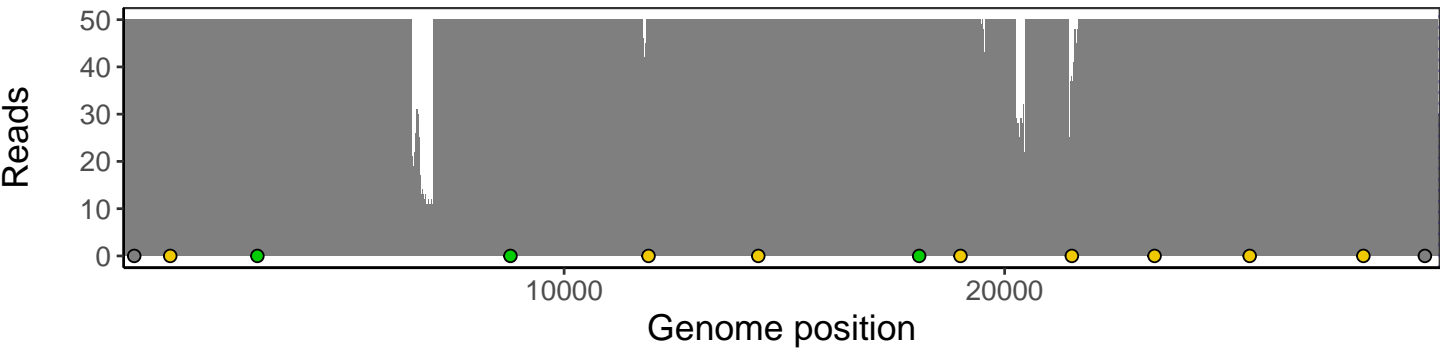
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



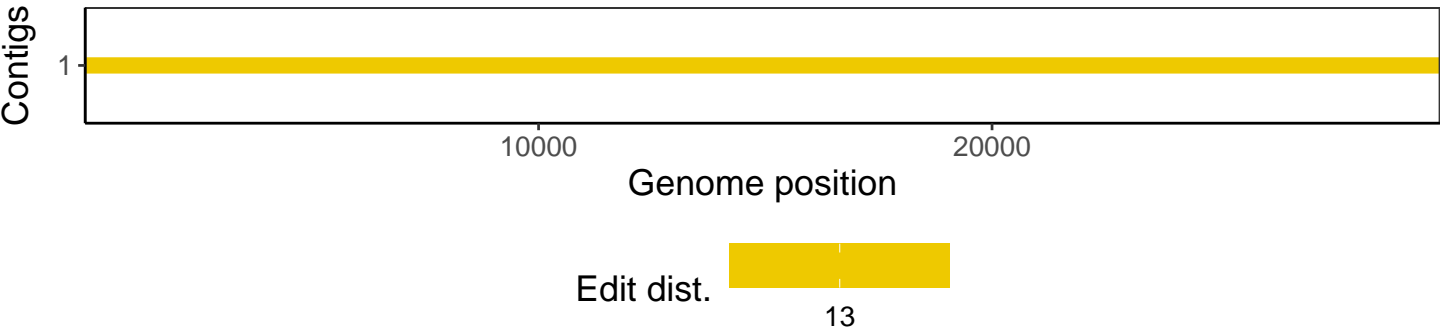
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



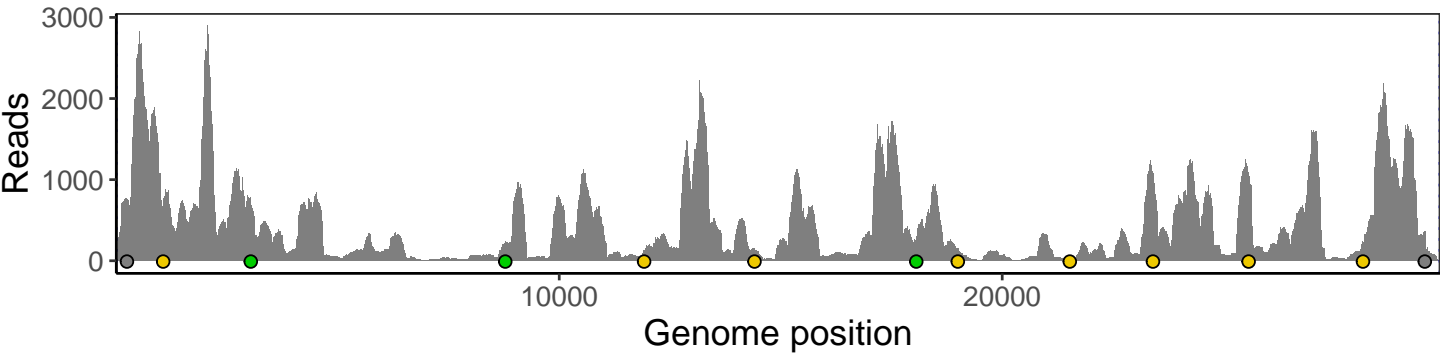
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



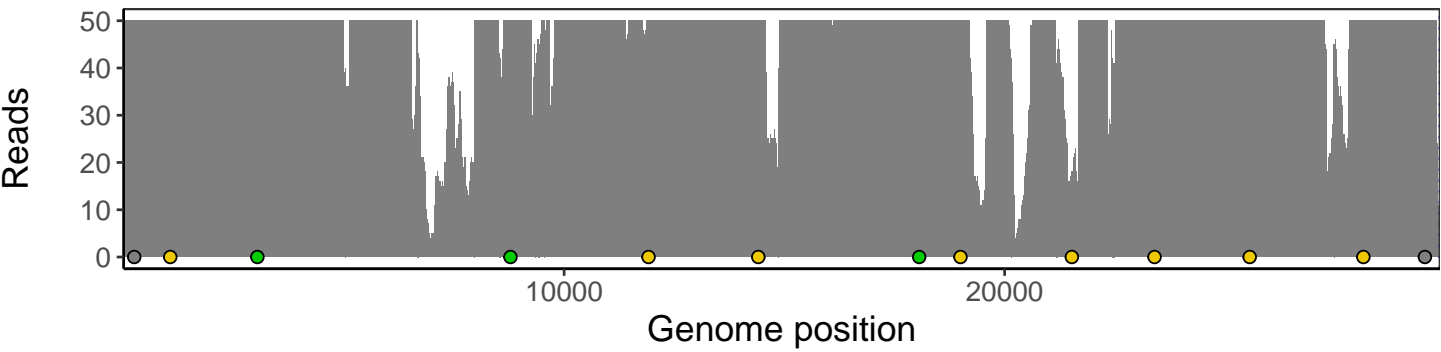
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

