

# COVID-19 subject 234

*2020-11-30*

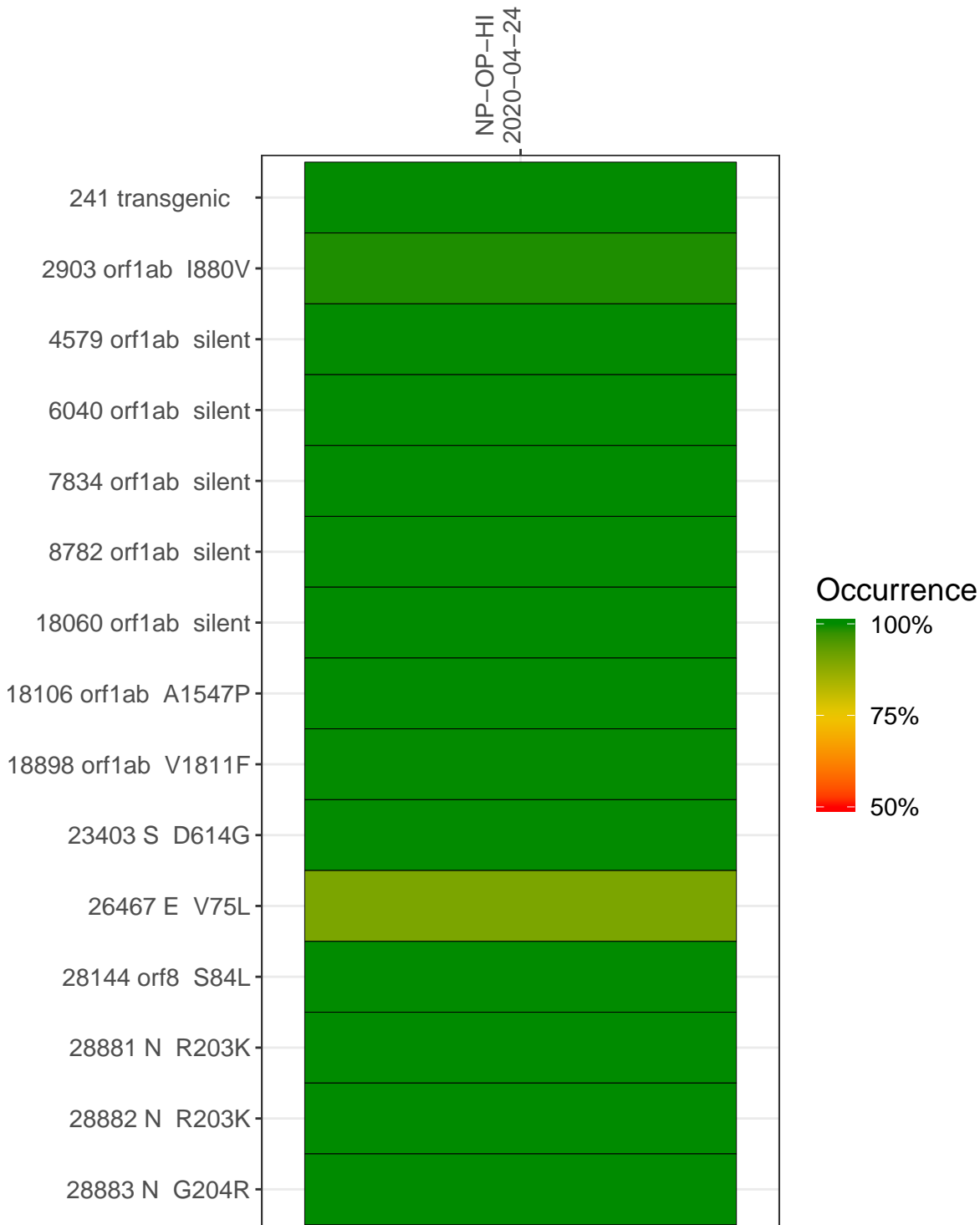
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0079	composite	NA	NP-OP-HI	2020-04-24	10.28	94.1%	92.9%
VSP0079-1	single experiment	147.5	NP-OP-HI	2020-04-24	2.11	57.9%	46.9%
VSP0079-2	single experiment	147.5	NP-OP-HI	2020-04-24	1.96	58.0%	55.1%
VSP0079-3	single experiment	147.5	NP-OP-HI	2020-04-24	2.93	73.6%	71.9%
VSP0079-4	single experiment	147.5	NP-OP-HI	2020-04-24	2.92	55.6%	53.5%

## 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-OP-HI  
2020-04-24

241 transgenic	657		9904	
2903 orf1ab I880V	608	6	2	4
4579 orf1ab silent	1		1977	
6040 orf1ab silent		1507		
7834 orf1ab silent			885	
8782 orf1ab silent	2	2324	6754	2077
18060 orf1ab silent				387
18106 orf1ab A1547P				487
18898 orf1ab V1811F		2532	4073	1111
23403 S D614G	1	6518	2326	13960
26467 E V75L	1		9	
28144 orf8 S84L	1	321	11226	7032
28881 N R203K	206	203	378	
28882 N R203K	205	203	378	
28883 N G204R	205	204	379	
	VSP0079-1	VSP0079-2	VSP0079-3	VSP0079-4

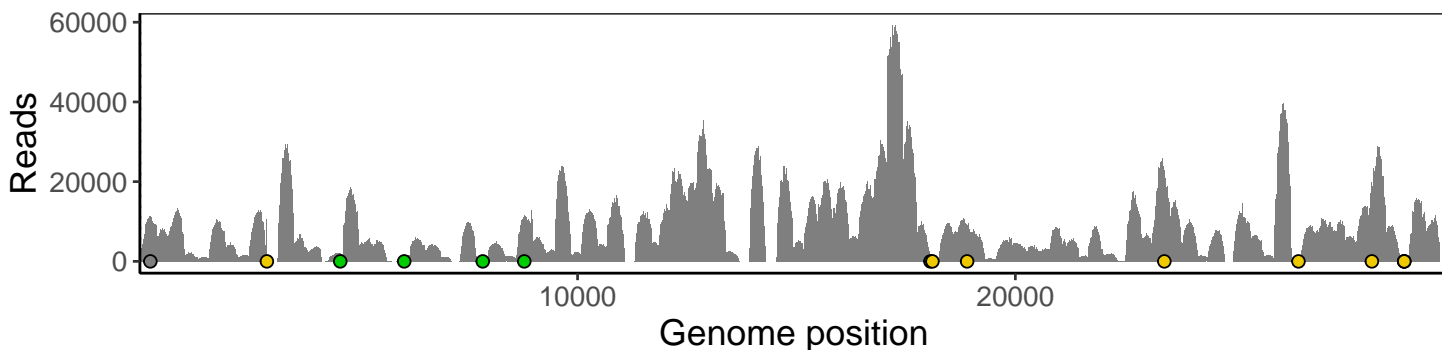
Base change

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

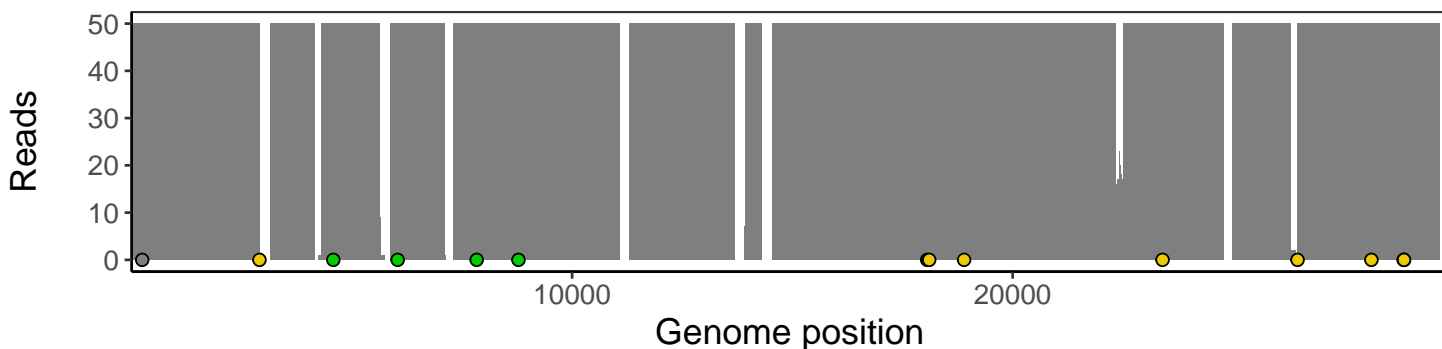
## Analyses of individual experiments and composite results.

VSP0079 | 2020-04-24 | NP-OP-HI | 234noh-t | 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.



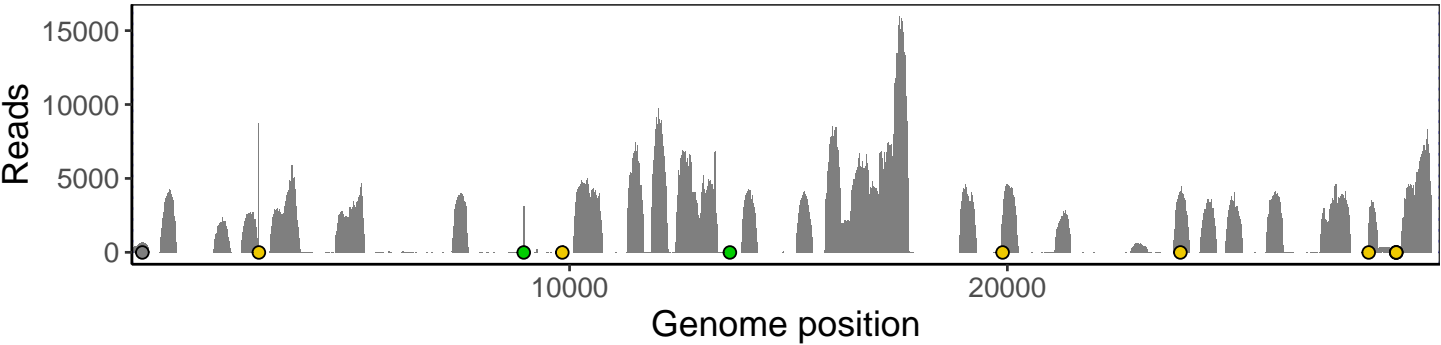
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



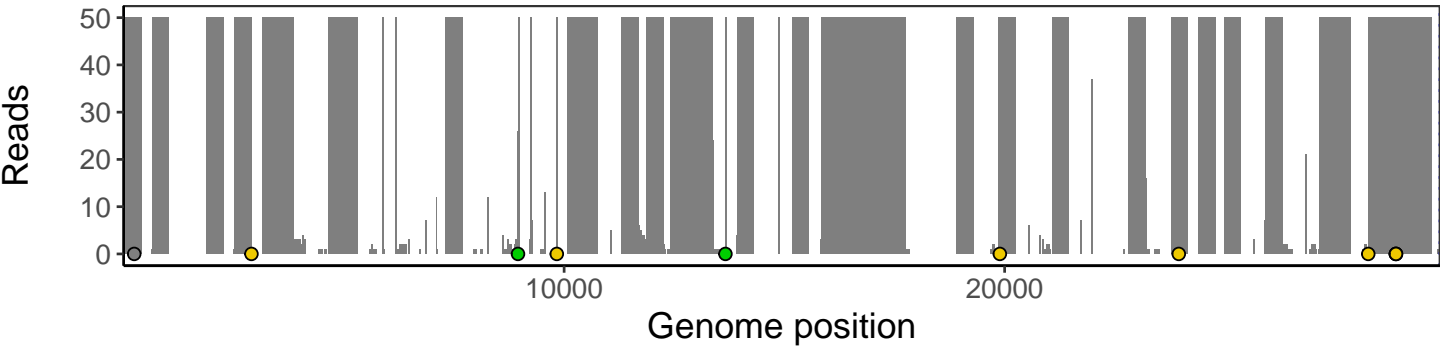
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



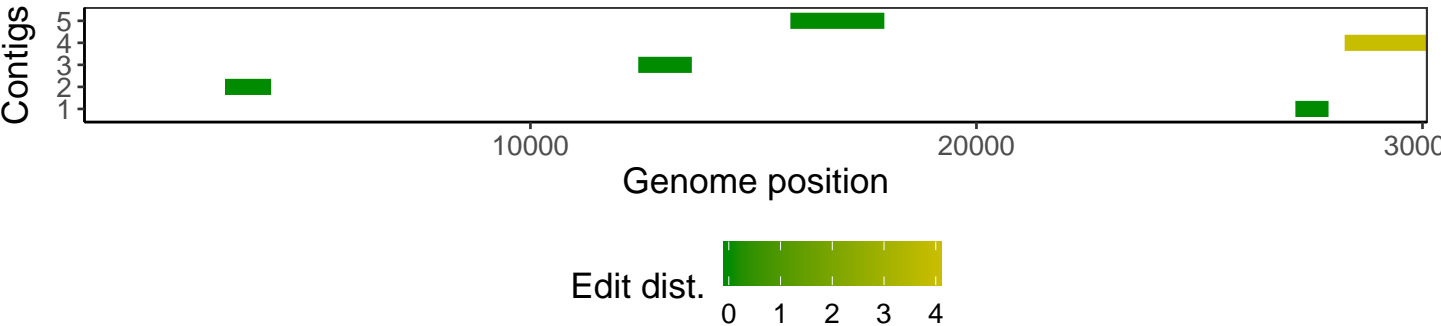
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



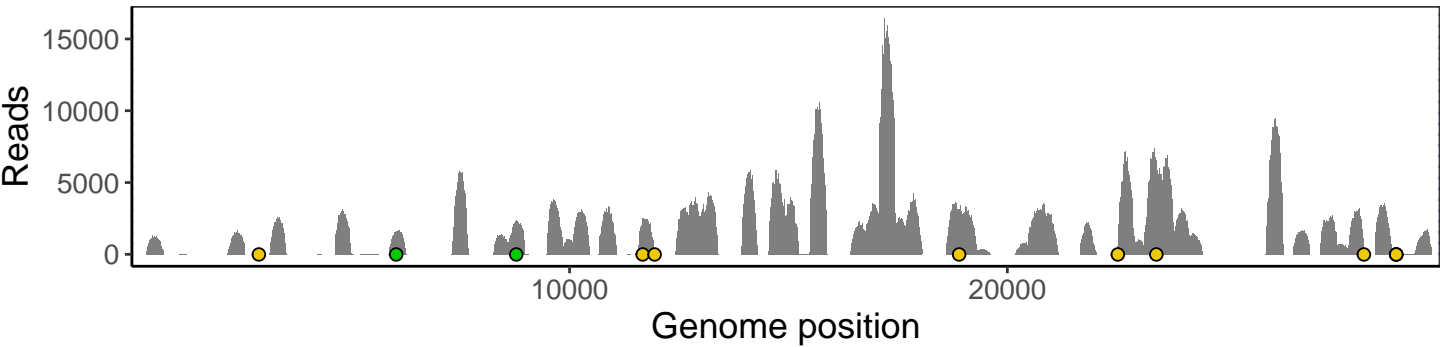
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



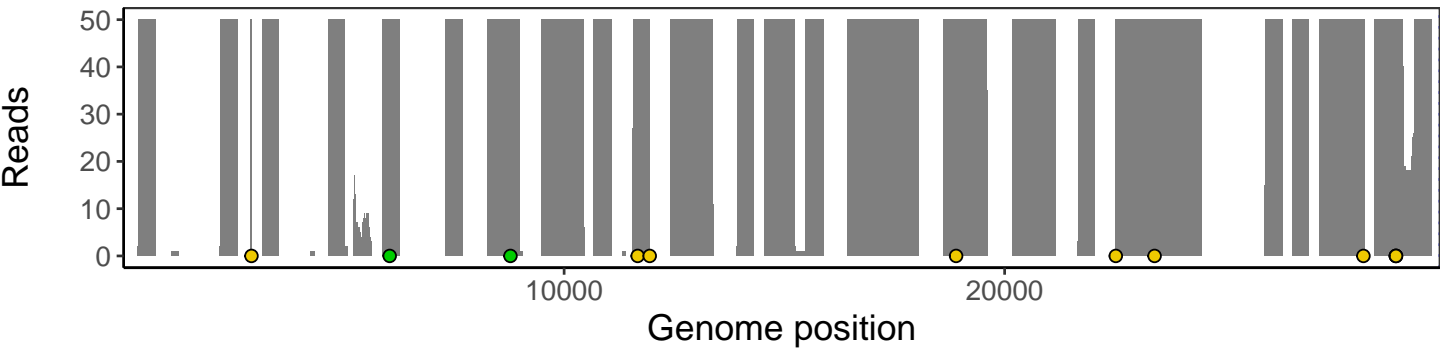
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



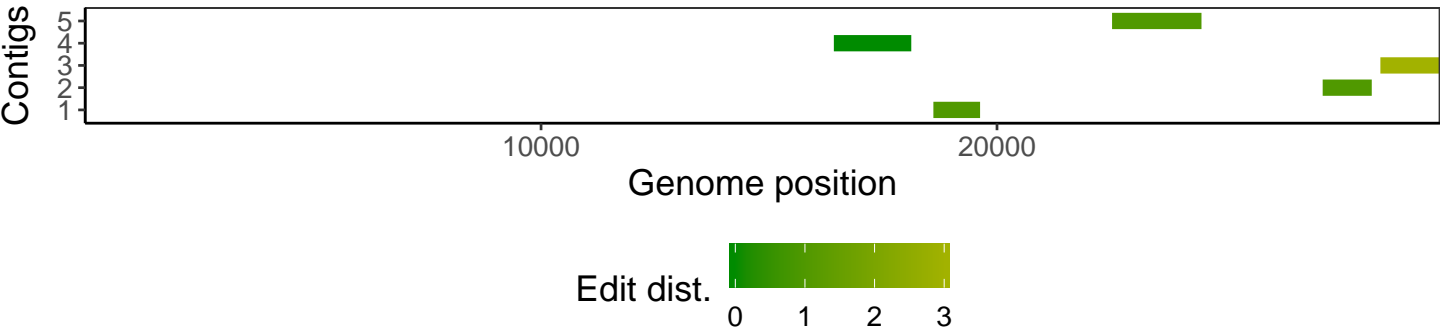
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



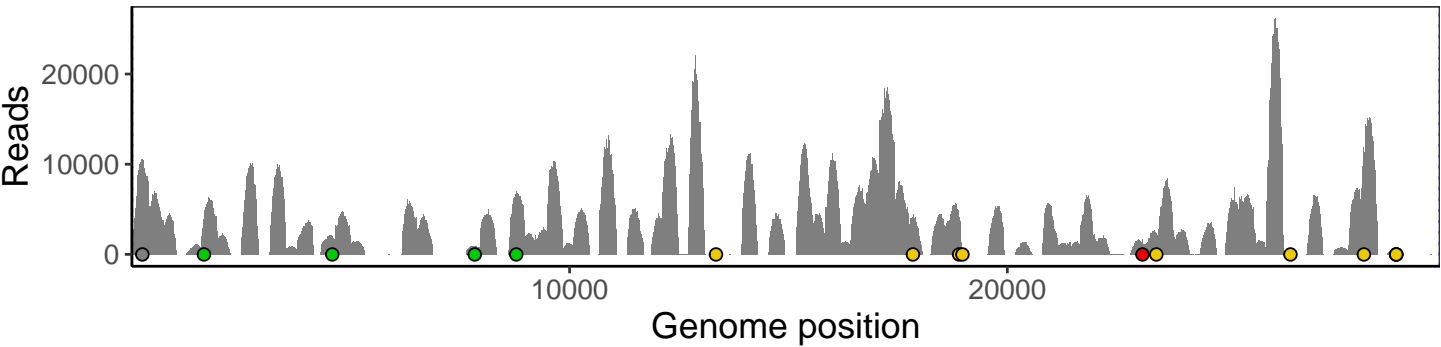
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



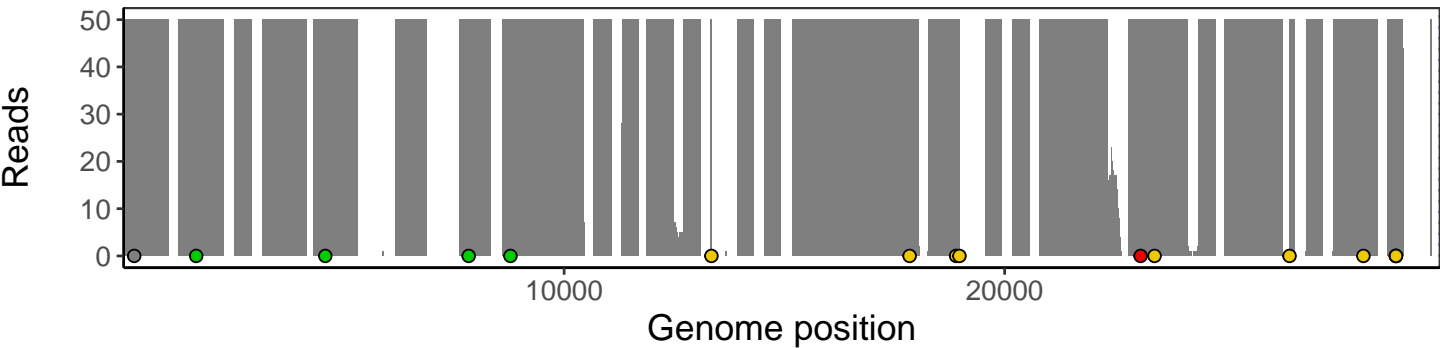
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



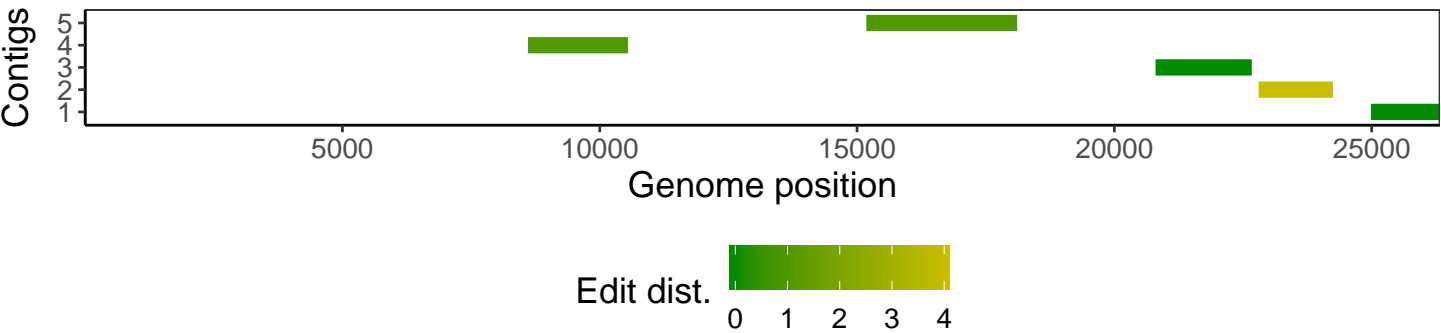
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



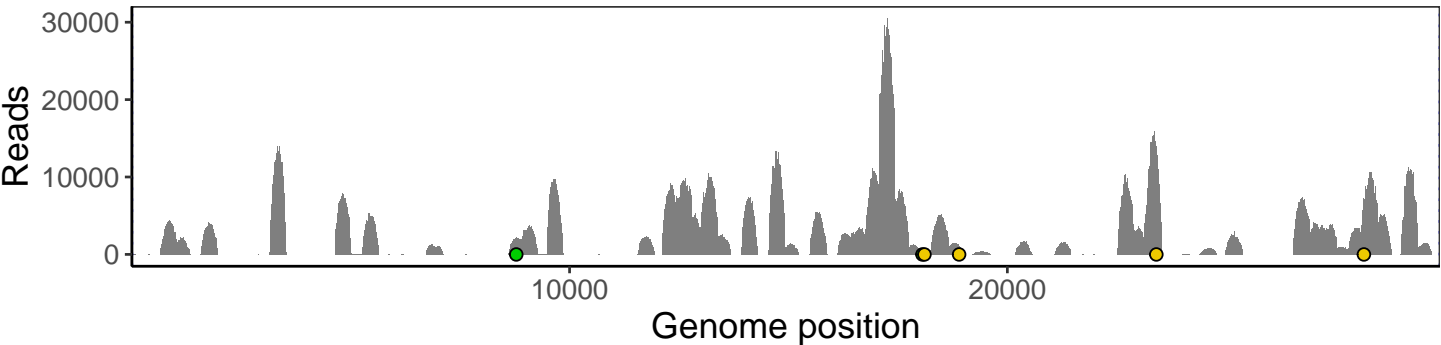
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



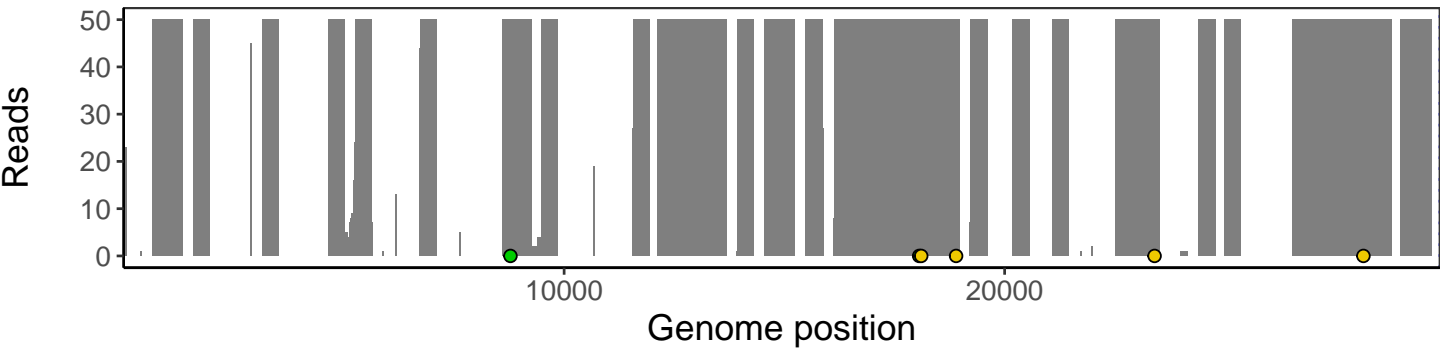
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



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.

