

COVID-19 subject 248

2020-09-14

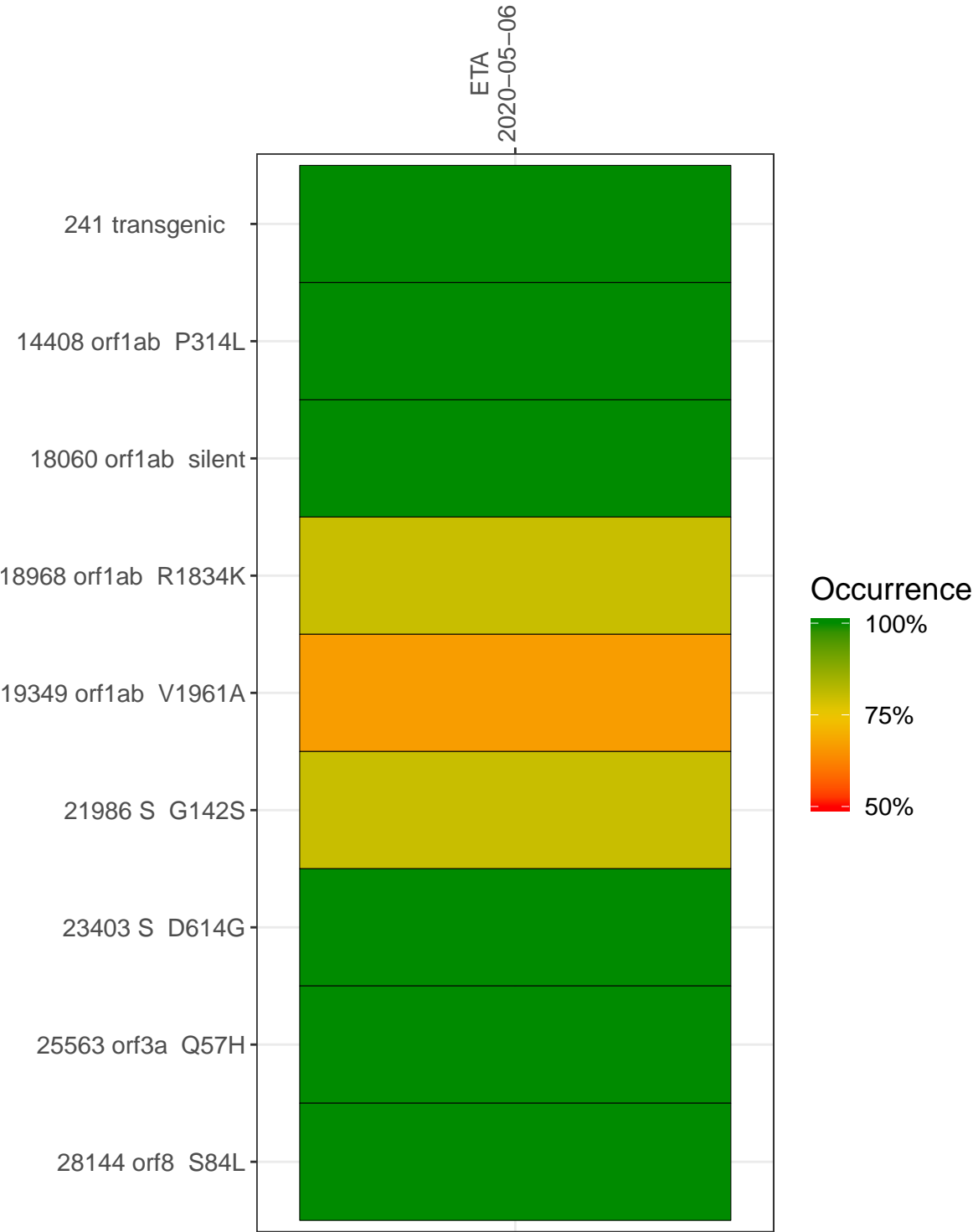
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found ([here](#)).

Table 1. Sample summary.

Experiment	Type	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0086	composite	NA	ETA	2020-05-06	5.12	95.3%	81.7%
VSP0086-1	single experiment	1730	ETA	2020-05-06	3.39	87.1%	72.5%
VSP0086-2	single experiment	8650	ETA	2020-05-06	NA	NA	NA
VSP0086-3	single experiment	8650	ETA	2020-05-06	1.88	73.8%	29.3%

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
2020-05-06

241 transgenic

2

5

14408 orf1ab P314L

33

2

18968 orf1ab R1834K

19349 orf1ab V1961A

6

21986 S G142S

5

23403 S D614G

29

11

25563 orf3a Q57H

6

28144 orf8 S84L

29

6

VSP0086-1

VSP0086-3

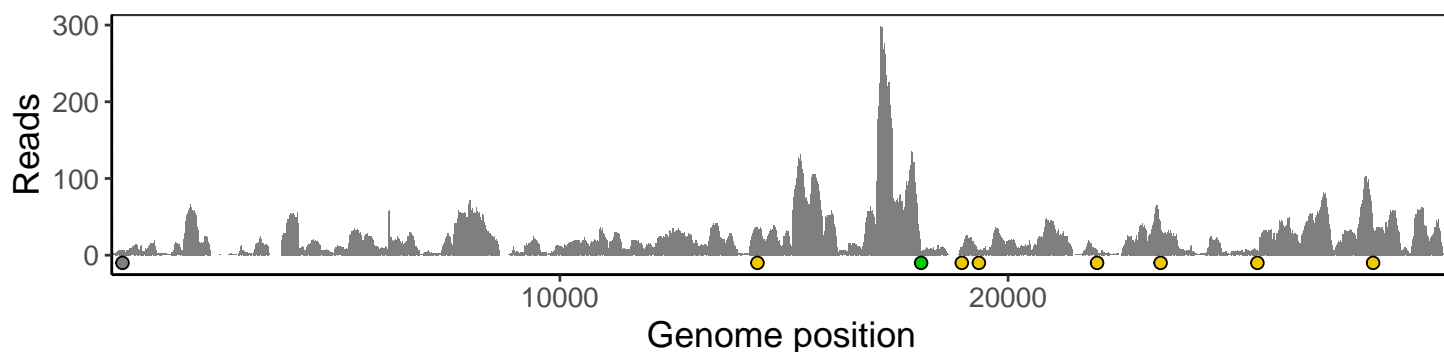
Base change



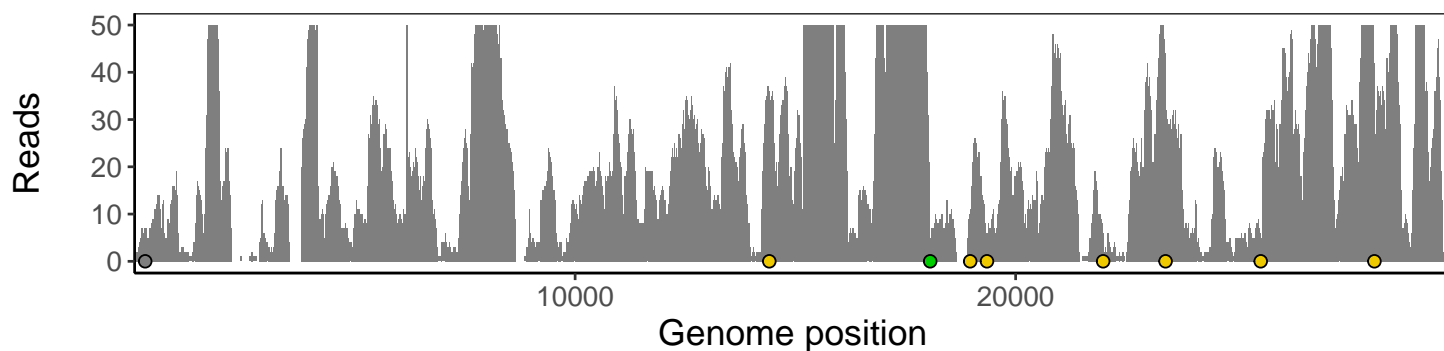
Analyses of individual experiments and composite results.

VSP0086 | 2020-05-06 | ETA | 248e-q | composite result

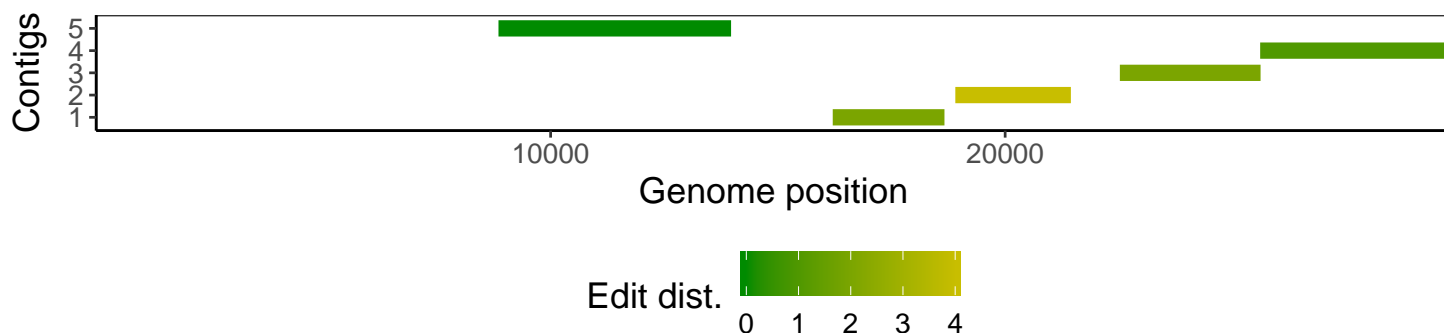
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



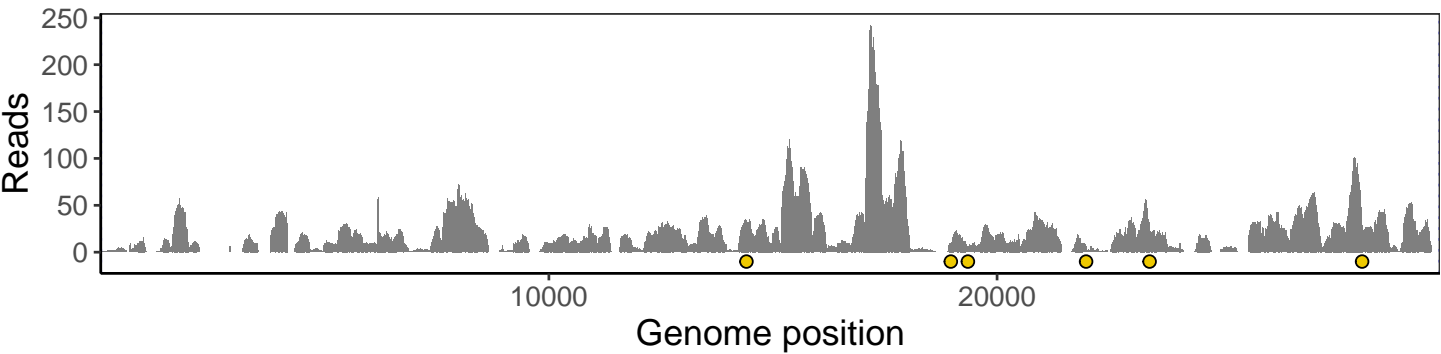
Excerpt from plot above focusing on reads coverage from 0 to 50 NT.



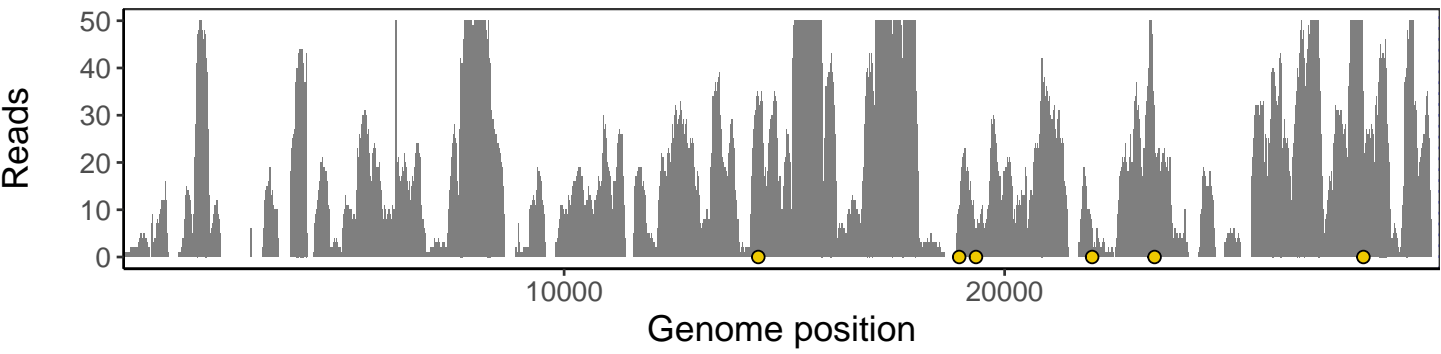
The longest five assembled contigs are shown below colored by their edit distance to the reference genome.



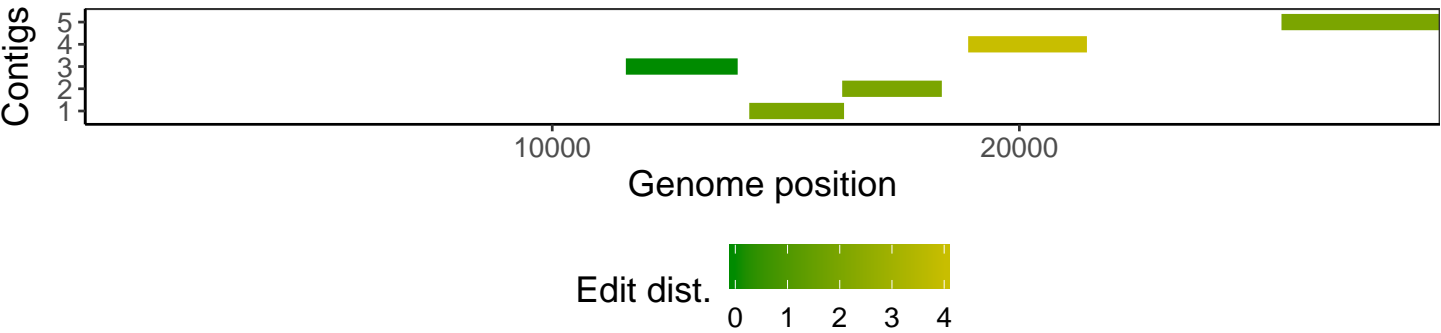
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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.

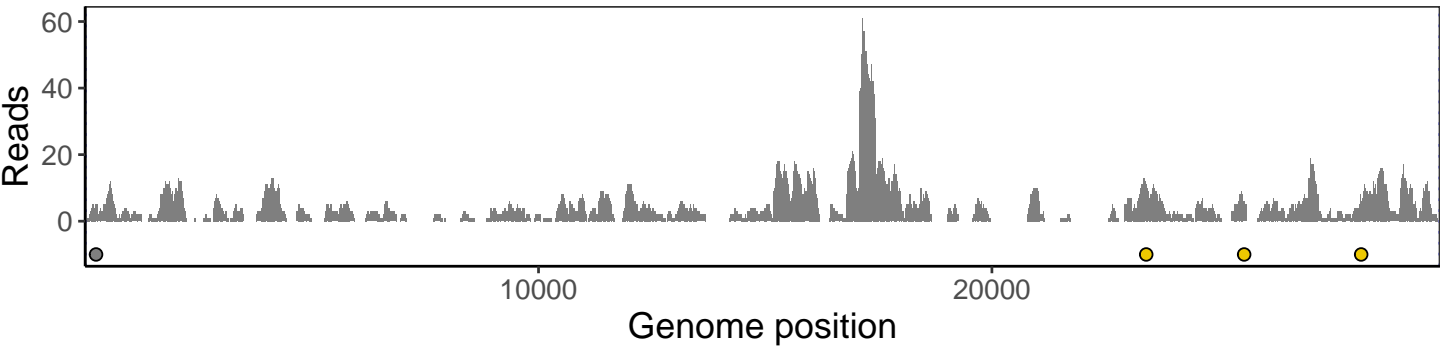


VSP0086-2 | 2020-05-06 | ETA | 248e-q | 8650 genomes | single experiment

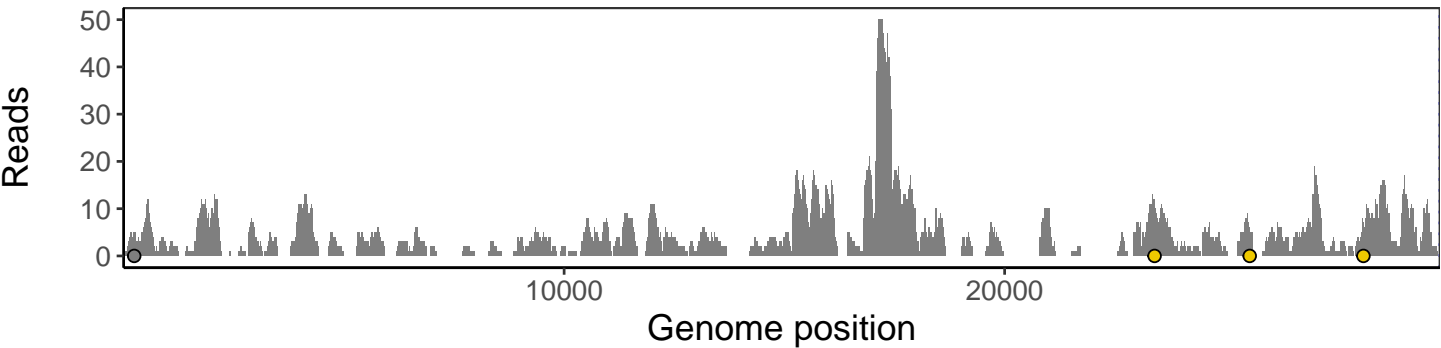
No pileup data available.

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.

