COVID-19 subject 272

2021-01-19

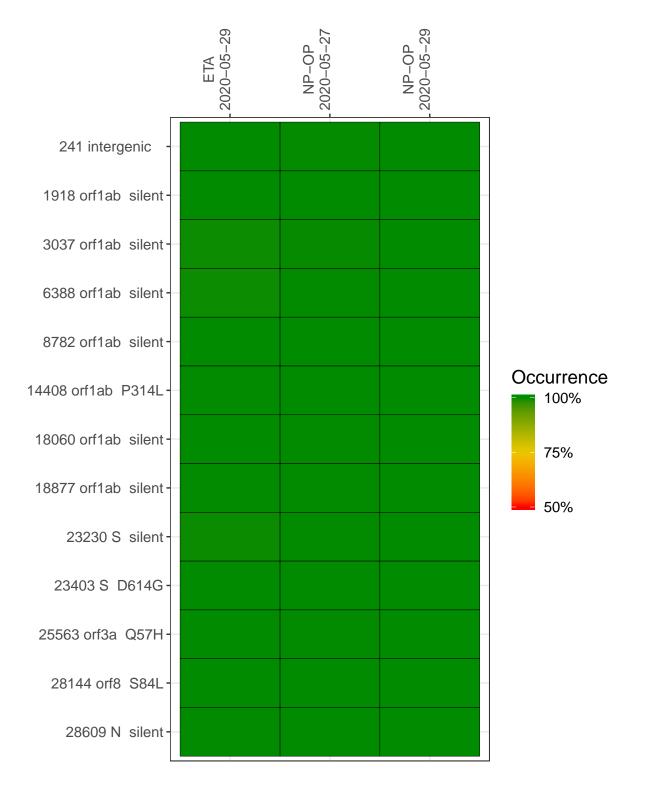
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. The code base for this analysis can be found (here).

Table 1. Sample summary.

Experiment	Туре	Input genomes	Sample type	Sample date	Largest contig (KD)	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0179-1	single experiment	14500	NP-OP	2020-05-27	29.41	99.8%	99.8%
VSP0195-1	single experiment	1530000	ETA	2020-05-29	25.63	99.8%	99.2%
VSP0196-1	single experiment	14300	NP-OP	2020-05-29	22.29	99.5%	99.1%
VSP0260-1	single experiment	NA	Stool	2020-06-17	NA	3.5%	0.0%

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 or 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.

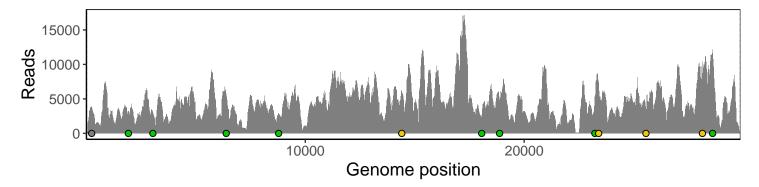


	ETA 2020-05-29	NP-OP 2020-05-27	NP-OP 2020-05-29	
241 intergenic	487	3572	60	
1918 orf1ab silent	5043	3229	53	
3037 orf1ab silent	2148	2920	78 230	
6388 orf1ab silent	1071	5330		
8782 orf1ab silent	1661	2841	57	
14408 orf1ab P314L	1877	5734	90	Base change Expected A
18060 orf1ab silent	876	2494	41	T C G
18877 orf1ab silent	3600	4784	87	N Ins/Del No data
23230 S silent	1554	5315	227	
23403 S D614G	2122	7246	267 41 198	
25563 orf3a Q57H	4872	3767		
28144 orf8 S84L	2051	8746		
28609 N silent	3531	12130	401	
	VSP0195-1	VSP0179-1	VSP0196-1	

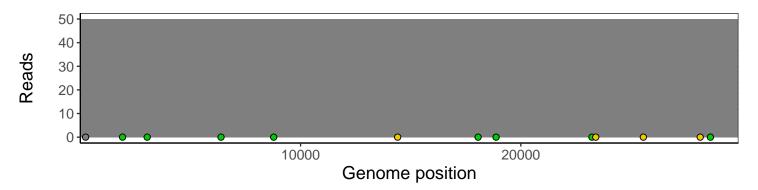
Analyses of individual experiments and composite results.

$VSP0179\text{-}1 \mid 2020\text{-}05\text{-}27 \mid NP\text{-}OP \mid 272\text{no-}q \mid 14500 \; genomes \mid single \; experiment$

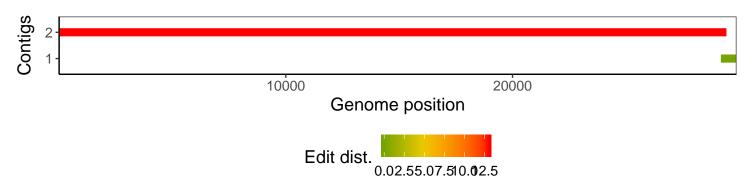
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

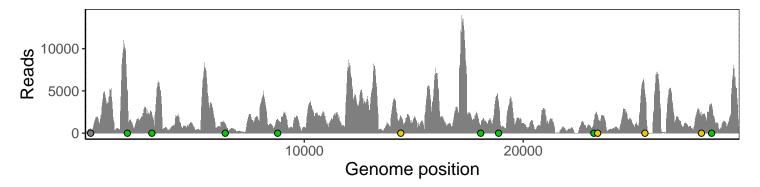


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

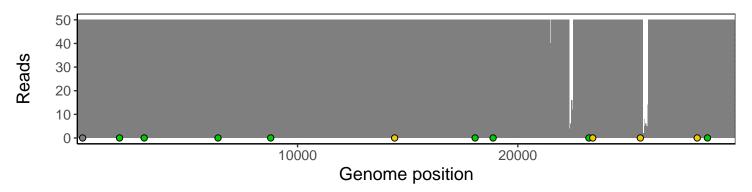


VSP0195-1 | 2020-05-29 | ETA | 272e-q | 1530000 genomes | single experiment

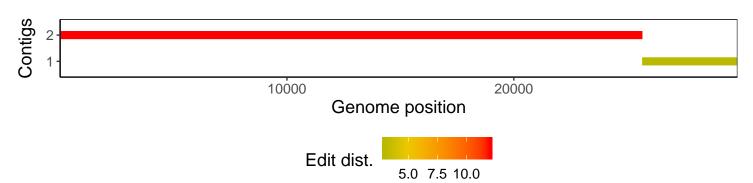
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

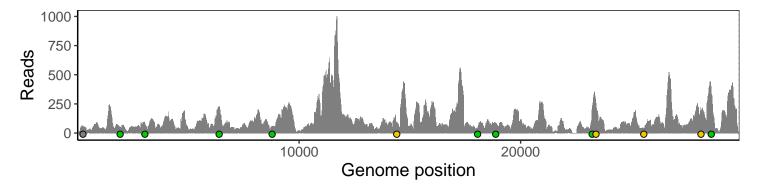


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

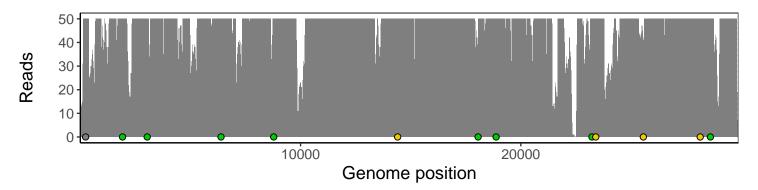


VSP0196-1 | 2020-05-29 | NP-OP | 272
no-q | 14300 genomes | single experiment

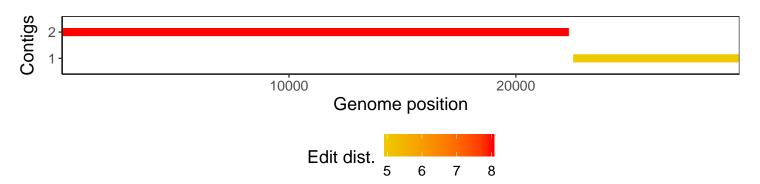
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

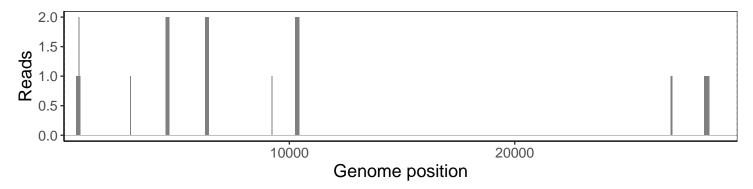


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

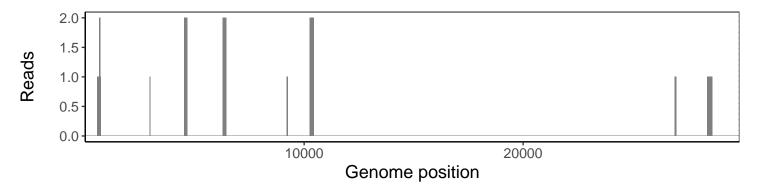


VSP0260-1 | 2020-06-17 | Stool | 272
p | NA genomes | single experiment

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.



No contig data available.