# COVID-19 subject 436

2021-04-17

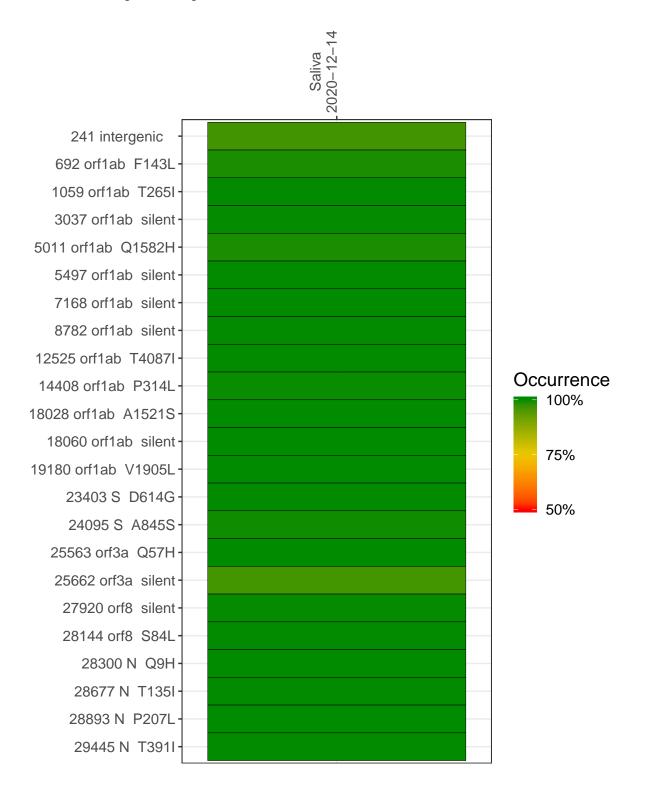
The table below provides a summary of subject samples for which sequencing data is available. The experiments column shows the number of sequencing experiments performed for each specimen. Experiment specific analyses are shown at the end of this report. Lineages are called with the Pangolin software tool (Rambaut et al 2020) for genomes with > 90% sequence coverage.

Table 1. Sample summary.

Experiment	Туре	Genomes	Sample type	Sample date	Largest contig (KD)	Lineage	Reference read coverage	Reference read coverage (>= 5 reads)
VSP0554	composite	NA	Saliva	2020-12-14	29.84	B.1.311	99.8%	99.8%
VSP0554-1	single experiment	NA	Saliva	2020-12-14	29.84	B.1.311	99.8%	99.8%
VSP0554-2	single experiment	NA	Saliva	2020-12-14	19.58	B.1.311	99.8%	99.7%

#### Variants shared across samples

The heat map below shows how variants (reference genome /home/everett/projects/SARS-CoV-2-Philadelphia/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.



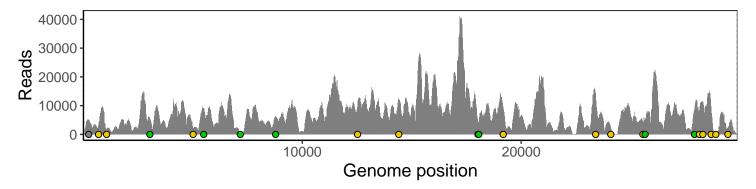
#### Saliva 2020–12–14

241 intergenic 3540 1304 692 orf1ab F143L 2250 816 1059 orf1ab T265I 1241 471 3037 orf1ab silent 3861 1386 5011 orf1ab Q1582H 1316 435 5497 orf1ab silent 6569 2363 7168 orf1ab silent 790 302 8782 orf1ab silent 4618 1456 12525 orf1ab T4087I 7966 2577 14408 orf1ab P314L 8716 2796 Ei-peted A1521S 3472 1030 T		2020		
1059 orf1ab T265I 3037 orf1ab silent 3861 1386 5011 orf1ab Q1582H 1316 435 5497 orf1ab silent 6569 2363 7168 orf1ab silent 790 302 8782 orf1ab silent 4618 1456 12525 orf1ab T4087I 7966 2577 14408 orf1ab P314L 8716 2796 18028 orf1ab A1521S 3472 1030 8 A 118060 orf1ab silent 2870 934 19180 orf1ab V1905L 23403 S D614G 10317 3503 24095 S A845S 1686 535 25563 orf3a Q57H 3307 1123 25662 orf3a silent 28300 N Q9H 7948 2883 28677 N T135I 28893 N P207L 29445 N T391I 1865	241 intergenic	3540	1304	
3037 orf1ab silent 5011 orf1ab Q1582H 5497 orf1ab silent 5497 orf1ab silent 5669 2363 7168 orf1ab silent 790 302 8782 orf1ab silent 4618 1456 12525 orf1ab T4087I 7966 25777 14408 orf1ab P314L 8716 2796 8830 orf1ab silent 180028 orf1ab silent 2870 934 19180 orf1ab V1905L 23403 S D614G 10317 3503 24095 S A845S 1686 535 25563 orf3a Q57H 3307 1123 25662 orf3a silent 2870 3925 1203 27920 orf8 silent 6531 28300 N Q9H 7948 2683 28893 N P207L 613 175 28445 N T391I 1855 534	692 orf1ab F143L	2250	816	
5011 orf1ab Q1582H 5497 orf1ab silent 5669 2363 7168 orf1ab silent 790 302 8782 orf1ab silent 4618 11456 12525 orf1ab T4087I 7986 2577 14408 orf1ab P314L 8716 18028 orf1ab A1521S 3472 1030 18060 orf1ab silent 2870 19180 orf1ab V1905L 7185 24458 1030 24095 S A845S 1686 535 25563 orf3a Q67H 3307 21123 27920 orf8 silent 28300 N Q9H 7948 28893 N P207L 29445 N T391I 1855 534	1059 orf1ab T265I	1241	471	
5497 orf1ab silent 790 302 8782 orf1ab silent 790 302 8782 orf1ab silent 4618 1456 12525 orf1ab T4087I 7966 2577 14408 orf1ab P314L 8716 2796 Base change Expected 18028 orf1ab A1521S 3472 1030 A T 18060 orf1ab silent 2870 934 10 c G G G 19180 orf1ab V1905L 23403 S D614G 10317 3503 24095 S A845S 1686 535 25563 orf3a Q57H 3307 1123 25662 orf3a silent 3925 1203 27920 orf8 silent 28300 N Q9H 7948 2683 28893 N P207L 613 1855 534	3037 orf1ab silent	3861	1386	
7168 orf1ab silent 790 302 8782 orf1ab silent 4618 1456 12525 orf1ab T4087I 7966 2577 14408 orf1ab P314L 8716 2796 Expected 18028 orf1ab A1521S 3472 1030 ☐ A 1	5011 orf1ab Q1582H	1316	435	
8782 orf1ab silent  4618  1456  12525 orf1ab T4087I  7866  2577  14408 orf1ab P314L  8716  2796  18028 orf1ab A1521S  18060 orf1ab silent  2870  934  19180 orf1ab V1905L  23403 S D614G  24095 S A845S  25563 orf3a Q57H  3307  1123  255662 orf3a silent  287920 orf8 silent  6531  27920 orf8 silent  28300 N Q9H  28300 N Q9H  7948  2863  28893 N P207L  613  1876  2577  Base change  Expected  A  1 T  1 C  2796  Base change  II  Expected  A  II  II  II  II  II  II  II  II  II	5497 orf1ab silent	6569	2363	
12525 orf1ab T4087I 7966 2577  14408 orf1ab P314L 8716 2796 Expected  18028 orf1ab A1521S 3472 1030	7168 orf1ab silent	790	302	
14408 orf1ab P314L 8716 2796 Expected  18028 orf1ab A1521S 3472 1030 A  18060 orf1ab silent 2870 934 C  19180 orf1ab V1905L 7185 2458 N  23403 S D614G 10317 3503 No data  24095 S A845S 1686 535  25563 orf3a Q57H 3307 1123  25662 orf3a silent 3925 1203  27920 orf8 silent 6531 2377  28144 orf8 S84L 6330 2093  28300 N Q9H 7948 2683  28677 N T135I 7083 2515  28893 N P207L 613 175  29445 N T391I 1855 534	8782 orf1ab silent	4618	1456	
18028 orf1ab A1521S  18060 orf1ab silent  2870  19180 orf1ab V1905L  23403 S D614G  24095 S A845S  25563 orf3a Q57H  25662 orf3a silent  27920 orf8 silent  28144 orf8 S84L  28300 N Q9H  28677 N T135I  28893 N P207L  613  1030  A  T  C  G  G  N  Ins/Del  No data  Expected  A  T  C  G  G  N  Ins/Del  No data  Expected  A  T  T  C  G  G  A  T  T  A  T  T  A  T  T  A  T  T  A  T  T	12525 orf1ab T4087I	7966	2577	
18028 orf1ab A1521S  18060 orf1ab silent  2870  934  19180 orf1ab V1905L  23403 S D614G  24095 S A845S  25563 orf3a Q57H  25662 orf3a silent  27920 orf8 silent  28144 orf8 S84L  28300 N Q9H  28677 N T135l  28893 N P207L  613  1030  A T T C G G G N Ins/Del No data  A T C C G G N Ins/Del No data  A T T C C G G N Ins/Del No data  A T T C C G G N Ins/Del No data  A T T C C C G G N Ins/Del No data  A T T C C C G G N Ins/Del No data  A T T C C C G G N Ins/Del No data  A T T C C C G G N Ins/Del No data  A T T C S A T T D S A T T S D Ins/Del No data  A T T D S D Ins/Del No data  A T T D S D Ins/Del No data  A T T D S D Ins/Del No data  A T T D S D Ins/Del No data  A T D D D D D D D D D D D D D D D D D	14408 orf1ab P314L	8716	2796	_
18060 orf1ab silent  19180 orf1ab V1905L  7185  2458  2458  N  23403 S D614G  10317  3503  24095 S A845S  1686  535  25563 orf3a Q57H  3307  1123  25662 orf3a silent  27920 orf8 silent  6531  28144 orf8 S84L  28300 N Q9H  7948  28677 N T135I  28893 N P207L  613  1855  534	18028 orf1ab A1521S	3472	1030	A
19180 orf1ab V1905L 23403 S D614G 10317 3503  24095 S A845S 1686 535  25563 orf3a Q57H 3307 1123  25662 orf3a silent 3925 1203  27920 orf8 silent 6531 28144 orf8 S84L 6330 28300 N Q9H 7948 28677 N T135I 7083 28893 N P207L 613 175 29445 N T391I 1855  2458  N N Ins/Del No data  No data	18060 orf1ab silent	2870	934	С
23403 S D614G       10317       3503       No data         24095 S A845S       1686       535         25563 orf3a Q57H       3307       1123         25662 orf3a silent       3925       1203         27920 orf8 silent       6531       2377         28144 orf8 S84L       6330       2093         28300 N Q9H       7948       2683         28677 N T135I       7083       2515         28893 N P207L       613       175         29445 N T391I       1855       534	19180 orf1ab V1905L	7185	2458	N
25563 orf3a Q57H       3307       1123         25662 orf3a silent       3925       1203         27920 orf8 silent       6531       2377         28144 orf8 S84L       6330       2093         28300 N Q9H       7948       2683         28677 N T135l       7083       2515         28893 N P207L       613       175         29445 N T391l       1855       534	23403 S D614G	10317	3503	
25662 orf3a silent       3925       1203         27920 orf8 silent       6531       2377         28144 orf8 S84L       6330       2093         28300 N Q9H       7948       2683         28677 N T135l       7083       2515         28893 N P207L       613       175         29445 N T391l       1855       534	24095 S A845S	1686	535	
27920 orf8 silent       6531       2377         28144 orf8 S84L       6330       2093         28300 N Q9H       7948       2683         28677 N T135I       7083       2515         28893 N P207L       613       175         29445 N T391I       1855       534	25563 orf3a Q57H	3307	1123	
28144 orf8 S84L       6330       2093         28300 N Q9H       7948       2683         28677 N T135I       7083       2515         28893 N P207L       613       175         29445 N T391I       1855       534	25662 orf3a silent	3925	1203	
28300 N Q9H       7948       2683         28677 N T135I       7083       2515         28893 N P207L       613       175         29445 N T391I       1855       534	27920 orf8 silent	6531	2377	
28677 N T135I       7083       2515         28893 N P207L       613       175         29445 N T391I       1855       534	28144 orf8 S84L	6330	2093	
28893 N P207L 613 175 29445 N T391I 1855 534	28300 N Q9H	7948	2683	
29445 N T391I 1855 534	28677 N T135I	7083	2515	
	28893 N P207L	613	175	
VSP0554-1	29445 N T391I	1855	534	
		VSP0554-1	VSP0554-2	

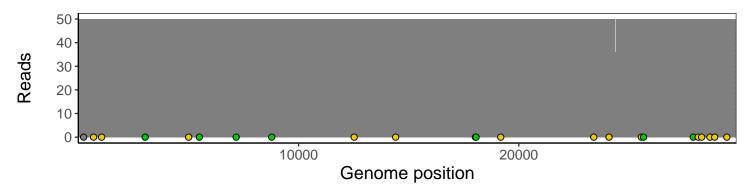
### Analyses of individual experiments and composite results

#### VSP0554 | 2020-12-14 | Saliva | 436s | composite result

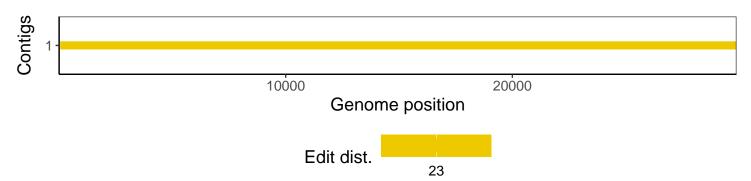
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

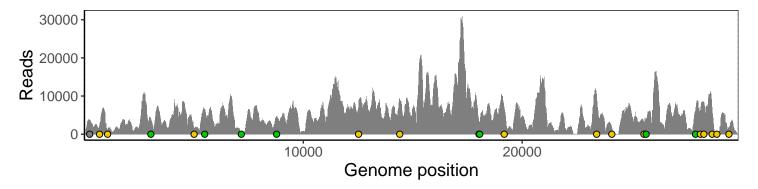


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

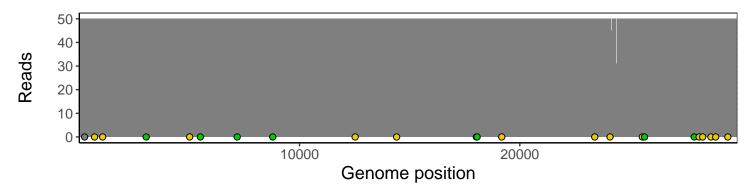


#### VSP0554-1 | 2020-12-14 | Saliva | 436s | genomes | single experiment

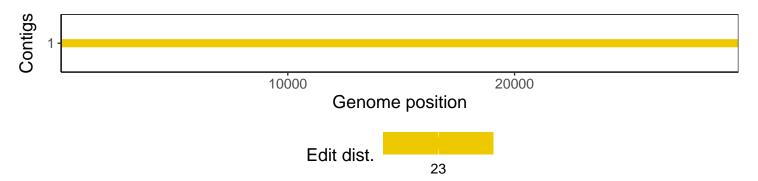
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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

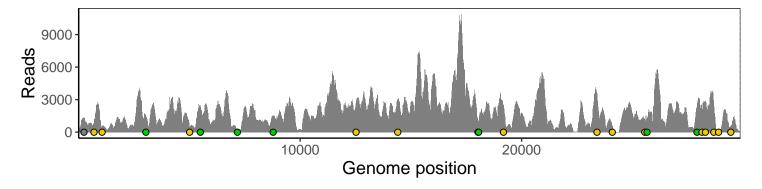


The longest five assembled contigs are shown below colored by their edit distance to the reference genome.

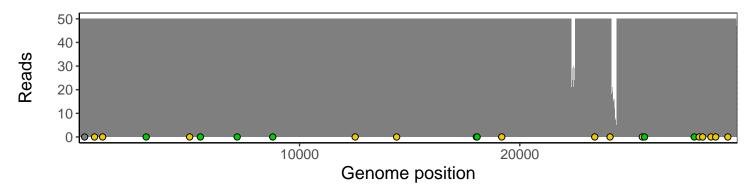


#### VSP0554-2 | 2020-12-14 | Saliva | 436s | genomes | single experiment

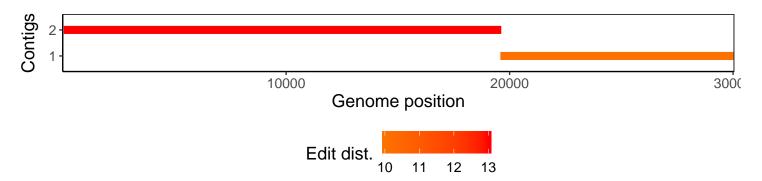
The plot below shows the number of reads covering each nucleotide position in the reference genome. Variants are shown as colored dots along the bottom of the plot and are color coded according by variant types: gray - transgenic, green - silent, gold - missense, red - nonsense, black - indel.



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.



## Software environment

Software/R package	Version
R	3.4.0
bwa	0.7.17-r1198-dirty
samtools	1.10 Using htslib 1.10
bcftools	1.10.2-34-g1a12af0-dirty Using htslib 1.10.2-57-gf58a6f3
pangolin	2.3.8
genbankr	1.4.0
optparse	1.6.0
forcats	0.3.0
stringr	1.4.0
dplyr	0.8.1
purrr	0.2.5
readr	1.1.1
tidyr	0.8.1
tibble	2.1.2
ggplot2	3.0.0
tidyverse	1.2.1
ShortRead	1.34.2
${\it Genomic Alignments}$	1.12.2
SummarizedExperiment	1.6.5
DelayedArray	0.2.7
matrixStats	0.54.0
Biobase	2.36.2
Rsamtools	1.28.0
GenomicRanges	1.28.6
$\operatorname{GenomeInfoDb}$	1.12.3
Biostrings	2.44.2
XVector	0.16.0
IRanges	2.10.5
S4Vectors	0.14.7
BiocParallel	1.10.1
BiocGenerics	0.22.1