Open access • Posted Content • DOI:10.1101/2021.03.10.21253321

Specific allelic discrimination of N501Y and other SARS-CoV-2 mutations by ddPCR detects B.1.1.7 lineage in Washington State — Source link

Garrett A. Perchetti, Haiying Zhu, Margaret G. Mills, Lasata Shrestha ...+15 more authors

Institutions: University of Washington, Fred Hutchinson Cancer Research Center

Published on: 12 Mar 2021 - medRxiv (Cold Spring Harbor Laboratory Press)

Topics: Lineage (genetic), DNA sequencing, Mutation and Genotyping

Specific allelic discrimination of N501Y and other SARS-CoV-2 mutations by ddPCR detects B.1.1.7 lineage in Washington State

Garrett A. Perchetti[1,a], Haiying Zhu[1,a], Margaret G. Mills[1], Lasata Shrestha[1], Cassia Wagner[2,3], Shah Mohamed Bakhash[1], Michelle Lin[1], Hong Xie[1], Meei-Li Huang[1], Patrick Mathias[1,4], Trevor Bedford[2,3], Keith R. Jerome[1,3], Alexander L. Greninger[1,3,*,#], Pavitra Roychoudhury[1,3,*,#]

[1]Department of Laboratory Medicine and Pathology, Virology Division, University of Washington, Seattle, WA, United States

[2]Department of Genome Sciences, University of Washington, Seattle, WA, United States

[3]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, United States

[4]Department of Biomedical Informatics and Medical Education, University of Washington School of Medicine, Seattle, WA, United States

[a]These authors contributed equally

*Co-senior authors

[#]Corresponding authors

Pavitra Roychoudhury, proychou@uw.edu

1616 Eastlake Ave E, Suite 320, Seattle, WA 98102

Box 358080

Phone: 206 667 7801

Alexander L. Greninger, agrening@uw.edu

1616 Eastlake Ave E, Suite 320, Seattle, WA 98102

Phone: 415 439 3448

Fax: 206 616 4340

Running Title: ddPCR assay for SARS-CoV-2 N501Y mutation

Keywords: ddPCR, N501Y, B.1.1.7, SARS-CoV-2, Spike protein, COVID-19

Abstract Word Count: 234

Manuscript Word Count: 2,370

ABSTRACT

Real-time epidemiological tracking of variants of interest can help limit the spread of more contagious forms of SARS-CoV-2, such as those containing the N501Y mutation. Typically, genetic sequencing is required to be able to track variants of interest in real-time. However, sequencing can take time and may not be accessible in all laboratories. Genotyping by RT-ddPCR offers an alternative to sequencing to rapidly detect variants of concern through discrimination of specific mutations such as N501Y that is associated with increased transmissibility. Here we describe the first cases of the B.1.1.7 lineage of SARS-CoV-2 detected in Washington State by using a combination of

RT-PCR, RT-ddPCR, and next-generation sequencing. We screened 1,035 samples positive for SARS-CoV-2 by our CDC-based laboratory developed assay using ThermoFisher's multiplex RT-PCR COVID-19 assay over four weeks from late December 2020 to early January 2021. S gene dropout candidates were subsequently assayed by RT-ddPCR to confirm four mutations within the S gene associated with the B.1.1.7 lineage: a deletion at amino acid (AA) 69-70 (ACATGT), deletion at AA 145, (TTA), N501Y mutation (TAT), and S982A mutation (GCA). All four targets were detected in two specimens, and follow-up sequencing revealed a total of 10 mutations in the S gene and phylogenetic clustering within the B.1.1.7 lineage. As variants of concern become increasingly prevalent, molecular diagnostic tools like RT-ddPCR can be utilized to quickly, accurately, and sensitively distinguish more contagious lineages of SARS-CoV-2.

BACKGROUND

The first known case of the SARS-CoV-2 B.1.1.7 variant in the United States was reported in Colorado on December 29th 2020 and the next day, it was confirmed in California (1, 2). University of Washington (UW) Virology had begun surveillance a few days prior using PCR to screen SARS-CoV-2 positive samples for the absence, or "dropout" of the S gene, which encodes the spike (S) protein on the surface of the viral particle. The B.1.1.7 variant is characterized by 17 mutations, eight of which occur within the S gene domain (3, 4). This region of the SARS-CoV-2 genome is of interest due to the B.1.1.7 lineage being associated with increased transmissibility, but also

because the FDA emergency use authorization for COVID-19 vaccines in the United States target the S protein (5, 6).

Over the course of four weeks, we screened more than a thousand SARS-CoV-2 positive samples for the S gene target failure (SGTF). We selected random clinical specimens that were positive using a CDC-based laboratory developed test (LDT) for SARS-CoV-2, and typically had a cycle threshold (C_T) under 35 (7–11). These SARS-CoV-2 positive samples were then amplified with the TaqPath COVID-19 assay (ThermoFisher Scientific, Waltham, MA, USA), a multiplex RT-PCR targeting the S gene, as well as regions within the N gene and ORF1ab (12). Candidates for the B.1.1.7 variant have a positive detection for the N gene and ORF1ab, with a negative for the S gene. SGTFs are candidates for the B.1.1.7 lineage, but the TaqPath assay is not necessarily specific for that variant exclusively due to other signature mutations, so genetic sequencing is used to confirm the B.1.1.7. lineage (13). However, sequencing can be a time-consuming and resource-intensive process that not all laboratories have integrated into their clinical workflow. Analysis of publicly available sequencing data by the Broad Institute revealed that it takes a median 85 days to get from sample to publicly available sequence in the United States (14, 15).

Here we describe a novel droplet reverse-transcription digital-PCR (RT-ddPCR) assay that specifically detects four mutations associated with the B.1.1.7 variant, particularly the N501Y mutation. This mutation in the S gene is shared by the U.K. and the South African variant (B.1.351) (16). Preliminary data has indicated B.1.1.7 to be more transmissible, while B.1.351 is considered to be less well neutralized by antibodies induced by certain vaccines, such as the AstraZeneca and Novavax

vaccines (3, 16–18). This RT-ddPCR assay can distinguish SARS-CoV-2 positive samples that carry this important N501Y mutation, as well specific allelic discrimination of the B.1.1.7 lineage, without genetic sequencing.

METHODS

Specimen Selection Criteria

From December 25th 2020 to January 20th 2021, 1,035 specimens positive for SARS-CoV-2 by our CDC-based LDT were screened for SGTFs (Fig.1) (19–21). Approximately 50% of these samples came from King County, followed by Pierce County with around 15%, Benton and Franklin Counties at 10%, and the remainder from other counties in Washington State. We selected samples with C_Ts<35 when available to reduce the effect of assay stochasticity, with consideration of initial data showing that the B.1.1.7 variant has been associated with higher viral loads (i.e. lower C_Ts) (22). Extracted nucleic acids were stored at 4°C or -20°C prior to amplification. This work was approved under a waiver of consent by the University of Washington institutional review board (STUDY00000408).

PCR

PCR was performed using TaqPath COVID-19 Combo Kit (ThermoFisher, Waltham, MA, USA) with 10 µL of extracted nucleic acid used as template per 25µL reaction. This multiplex real-time RT-PCR assay targets the S gene, N gene, and ORF1ab of SARS-CoV-2. Reactions utilized a kit-provided positive control ($1x10^4$ copies/µL) diluted with TaqPath COVID-19 Control Dilution Buffer, and dH_2O as a

negative template control. Amplifications were run on Applied Biosystems 7500 Real-Time PCR Systems (ThermoFisher) according to manufacturer's thermocycling parameters.

RT-ddPCR

Four sets of primers and probes were designed based on U.K. variant sequence hCoV-19/England/MILK-9E2FE0/2020 (EPI_ISL_581117, collection date 2020-09-21). To check for variation at primer and probe sites, 216 B.1.1.7 sequences were downloaded from GISAID on December 23rd, 2020, and aligned against the Wuhan-Hu1 reference sequence (NC_045512.2) using MAFFT v7.450 within Geneious Prime (https://www.geneious.com) (Supplementary Table 1). In addition, a G-block of 490bp was designed that includes four target amplicons. All primer and probes were synthesized by ThermoFisher and the G-block was synthesized by IDT (Coralville, IA, USA). Primers were included at 900 nM and probes were used at 250 nM concentrations. All four targets are within the S gene domain: a deletion at amino acid (AA) 69-70 (ACATGT), deletion at AA 145, (TTA), N501Y mutation (TAT), and S982A mutation (GCA). The specific primers and probe sequences and characteristics are outline in Table 1; the G-block sequence:

AACTCAGGACTTGTTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATCTCT
GGGACCAATGGTACTAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGACGC
TACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTTGGGT
GTTTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTAGT
GCGAATAATTGTACACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTT
ACAATCATATGGTTTCCAACCCACTTATGGTGTTGGTTACCAACCATACAGAGTAGT

AGTACTTTCTTTTGAACTTCTACATGCACCAGCAACCCAAACAACTTAGCTCCAATT TTGGTGCAATTTCAAGTGTTTTAAATGATATCCTTGCACGTCTTGACAAAGTTGAGG CTGAAGTGCAAATTGATAGGTTGATCACAGGC.

Two multiplex RT-ddPCR reactions per sample were performed in parallel (as outlined in Table 1) using One-step RT-ddPCR Advanced Kit for Probes (Bio-Rad Laboratories, Hercules, CA, USA) with the Automated Droplet Generator (Bio-Rad) and C1000 Touch thermocycler (Bio-Rad). Template RNA for each clinical sample was diluted to an approximate N1 C_T of 30 before amplification. Thermocycling conditions were as follows: 50°C for 60 min, 40 cycles at 95°C for 30 sec and 60°C for 1 min, then 98°C for 10 min. Droplet detection was performed using the QX200 Droplet Reader (Bio-Rad) and QuantaSoft Pro 1.0.596 version software. For SARS-CoV-2 B.1.1.7 detection, all four targets are amplified beyond the thresholds. For non-B.1.1.7 samples, 0-1 target(s) are amplified beyond the thresholds. Both the synthetic G-block and a known B.1.1.7 clinical sample are used as positive controls, and a SARS-CoV-2 clinical positive control is used as a negative control in the B.1.1.7 assay.

Sequencing

For next-generation sequencing, 11 µL of extracted RNA was used for single-stranded complementary DNA (sscDNA) synthesized using SuperScript IV First-Strand Synthesis System according to the manufacturer's protocol (ThermoFisher). Libraries were prepared using the Swift Normalase Amplicon Panel (SNAP) for SARS-CoV-2 (Swift Biosciences, Ann Arbor, MI, USA) as previously described (23). For standards, a wild type clinical nasopharyngeal specimen positive for SARS-CoV-2 was used as a positive control and water was used as no-template negative control. The resulting

libraries were quantified fluorometrically on Qubit 3.0 using the Quant-iT dsDNA high sensitivity kit (Life Technologies, Carlsbad, CA, USA). The libraries passing quality control, nucleic acid concentrations > [1.3 ng/μL], were normalized manually and sequenced on the Illumina MiSeq platform (Illumina, San Diego, CA, USA) using MiSeq Reagent Kit v2 (2x150 reads).

Raw reads were analyzed using a custom bioinformatic pipeline (TAYLOR, https://github.com/greninger-lab/covid_swift_pipeline) (23). Briefly, raw reads were trimmed to remove adapters and low-quality regions and mapped to the Wuhan-Hu-1 reference sequence (NC_045512.2) using BBMap v38.86 (https:/jgi.doe.gov/data-and-tools/bbtools/). Aligned reads were then soft-clipped of PCR primers using the PrimerClip package from Swift Biosciences (https://github.com/swiftbiosciences/primerclip) and a consensus sequence was called using bcftools v1.9 (24). Consensus sequences were aligned using MAFFT v7.450 within Geneious Prime (https://www.geneious.com) (25). Mutations in the spike protein were manually reviewed in addition to automated variant calling within the pipeline. Clade assignment was performed using Nextclade v0.12.0 (https://github.com/nextstrain/nextclade). Sequences were deposited to Genbank (accessions pending) and GISAID (EPI_ISL_861730 and EPI_ISL_861731); raw reads were deposited to the NIH's Sequence Read Archive (Bioproject PRJNA610428).

Phylogenetic tree construction utilized the Nextstrain pipeline to align sequences, reconstruct maximum-likelihood and time-resolved phylogenetic trees and to infer nucleotide and amino acid substitutions across the phylogeny (26). The specific workflow for this analysis is available at: https://github.com/blab/ncov-wa-build. The tree

185 included the B.1.1.7 sequences described here, 1,586 SARS-CoV-2 samples available
186 on GISAID collected in Washington State from November 2020 through February 12th,
187 2021, and an additional 1,824 global contextual sequences from GISAID sampled
188 based on genetic similarity to the Washington sequences (Supplementary Table 2). The
189 tree, continually updated with additional Washington sequences collected between
190 November 2020 and February 2021, can be viewed at:
191 https://nextstrain.org/groups/blab/ncov/wa/nov20-feb21.
192
193
194 **RESULTS**
195 *PCR*
196 Means, medians, and the ranges of C_Ts detected for the S gene, N gene, and
197 ORF1ab targets are characterized in Table 2. Seven samples were candidates for
198 SGTF, out of a total of 1,035 samples screened for the B.1.1.7 variant. Of these, 5/7
199 had C_Ts >33.5 for all targets with a mean and median C_T of 35.7 and 35.6 for N gene,
200 and C_Ts 37.8 and 37.5 for ORF1ab, respectively. Two samples however, had strong
201 fluorescent amplification for the N gene and ORF1ab targets, but no S gene
202 amplification (Fig. 2). Candidate 1 (55538) had C_Ts of 23.4 and 23.7 for N gene and
203 ORF1ab targets respectively, without S gene amplification. Candidate 2 (55545) had
204 C_Ts of 26.7 and 26.4 for N gene and ORF1ab. Consequently, we decided to ddPCR and
205 sequence these specimens.
206 *RT-ddPCR*

Both specimens showed clear fluorescent amplification above analysis thresholds of all four targets for both sets of RT-ddPCR reactions (Fig. 3). Amplification characteristics are outlined in Table 3 with absolute virus copies/µL of RNA included in the ddPCR reaction. With back-calculations considering RNA dilutions and extraction compressions, U.K. Variant #1 (55538) was quantified to have 765,500-860,000 virus copies/mL and U.K. Variant #2 (55545) was quantified to have 60,250-72,250 virus copies/mL depending on target amplicon.

While the primer sites for all four targets are present in wild type SARS-CoV-2 viral RNA as well as B.1.1.7 RNA, the probes for each target are designed to bind specifically to the regions mutated in the B.1.1.7 variant. Depending on the difference in melting temperature between the wild type and variant sequences for a given target, the probe may show some binding to the wild type sequence as well, but at a lower efficiency than it binds to the variant sequence. In RT-PCR, inefficient binding to the wild type strand is indistinguishable from high-efficiency binding to a lower-concentration variant strand. In RT-ddPCR, because individual template strands are amplified in separate droplets, inefficient probe binding can be identified as lower-amplitude fluorescence from each droplet. Thus, even an A to T single nucleotide polymorphism such as that present in the N501Y mutation (S1B) is easily distinguishable by RT-ddPCR by screening for droplets with S1B probe amplitude above a threshold of 5,700.

Sequencing

We obtained good quality consensus genomes for both samples, with more than 20,000X average coverage across the genome (Table 4). Both samples were classified

as 20I/501Y.V1 using Nextclade (ver 0.12.0) and had 100% pairwise nucleotide identity in a whole genome alignment). For each sample, a total of 10 mutations were found in the spike protein: H69-, V70-, Y144-, N501Y, A570D, D614G, P681H, T716I, S982A, D1118H. In a phylogenetic tree, these samples represented a unique cluster within clade 20I/501Y.V1 (Figure 4). They are 6 mutations diverged from the genetically closest sample available in GISAID, England/ALDP-BB47ED/2020, sampled in the United Kingdom in November 2020. Other B.1.1.7 samples collected in Washington did not cluster with them, suggesting that these samples represent a unique introduction into the state.

DISCUSSION

The B.1.1.7 variant of SARS-CoV-2 is more transmissible than the original lineage, which can ultimately lead to an increase in global cases, unfortunately resulting in more death. Detection of this variant has clinical and epidemiological relevance due to its increased transmissibility, as well as the N501Y mutation being shared with the South African (B.1.351) variant (27). B.1.351 is considered to be associated with decreased immunogenicity by provoking a weaker neutralizing antibody response against spike protein mutants (18, 28). Unfortunately, this means recently developed vaccines may have diminished vaccine efficacy. Pfizer reported that via plaque reduction neutralization testing, neutralization of the B.1.351 variant was reduced by two-thirds compared to USA-WA1/2020 (29); in the same day, Moderna reported their vaccine elicited sixfold less antibodies against the B.1.351 variant (30). These mutations

have tremendous real-world impact as South Africa recently halted using the Oxford-AstraZeneca vaccine due to its decreased efficacy against the B.1.351 variant (31).

Our experience with the SGTF screening underscores the enduring utility of multiplex PCR assays and their adaptability. PCR remains a powerful tool in the scalability of screening large volumes of viral samples for variant detection. Other clinical laboratories from San Francisco to Lyon have also recently implemented modified PCR-based screening methods to hone in on samples with potential to be variants of interest (32–34). Using PCR and allelic discrimination ddPCR in conjunction with sequencing, clinical laboratories can efficiently and quickly identify SGTF specimens that are candidate variants of concern. Helix reported that of the positive COVID-19 tests screened, less than 1% had SGTF, however, of those SGTFs, more than one-third were confirmed B.1.1.7 lineages (35).

According to GISAID, the United States has only sequenced 3 out of every 1,000 positive SARS-CoV-2 samples (36). RT-ddPCR has potential to identify samples to prioritize for sequencing, allowing more efficient allocation of strain surveillance resources. However, multiple targets are necessary to accurately detect newer viral mutations, such as those seen in the U.K. and South African variants; these mutations can arise in regions that are targeted by PCR primers (37, 38). Assays that only target the S gene domain of SARS-CoV-2 run the risk of missing newer variants and ultimately lead to an increase in false negatives (39).

Although our surveillance was limited, the variant positivity at time of initial detection was approximately 1 in 500 (2 out of 1,035) randomly selected SARS-CoV-2 samples. However, reports indicate that due to B.1.1.7's increased transmissibility, it is

already increasing in frequency and is soon expected to become the dominant strain in the United States (6, 27, 35, 40). The large number of genomic mutations associated with B.1.1.7, including single nucleotide polymorphisms (SNPs) A570D, D614G, P681H, T716I, S982A, and D1118H, underscores the need to distinguish these subtle mutations in emerging variants of concern (41). RT-ddPCR is more sensitive than RT-PCR at resolving SNPs, and is better suited for allelic discrimination to differentiate the B.1.1.7 lineage or other variants of concern (42–44). The S982A SNP, for instance, is specific for B.1.1.7 lineages, whereas the N501Y mutation is not (40). The S982A SNP is readily detectable using RT-ddPCR, however, would not be distinguishable by RT-PCR.

As genetic surveillance becomes increasingly relevant in efforts to track and understand new SARS-CoV-2 variants of concern in real-time, RT-ddPCR continues to cement its place in the clinical laboratory armamentarium (15, 17, 26, 41, 45). However, RT-ddPCR technology is still not ubiquitous in clinical laboratory settings. Increased adoption and investment in this technology can allow labs to rapidly estimate prevalence of existing variants and perform sample screening to better allocate limited sequencing resources.

AKNOWLEDGEMENTS

We gratefully acknowledge the authors and laboratories involved in the generation and deposition of sequencing data, which we obtained via the GISAID Initiative (Supplementary Table 2). The authors would also like to thank Victoria Mallett Rachleff for assistance in data deposition.

TABLES

Table 1. Design information for B.1.1.7 identification assay

Target		Primer/Probe Name	Sequence	Length (nt)	T_M (°C)	Amplicon Length (nt)	RT-ddPCR reaction	Analysis Threshold
S1A	Deletion 69-70 (ACATGT)	VUI202012_SFwd1A	tgttcttacctttcttttccaatgttactt	30	60			
		VUI202012_SRev1A	aatggtaggacagggttatcaaacct	26	60.2	91	Set 1	6,000
		VUI202012_SProbe1A	FAM_atgcta**tc**tctgggaccaa_MGB	19	69			
S1B	N501Y (A->T)	VUI202012_SFwd1B	atggtgttgaaggttttaattgttacttt	29	58.8			
		VUI202012_SRev1B	gtgcatgtagaagttcaaaagaaagtacta	30	58	82	Set 1	5,700
		VUI202012_SProbe1B	VIC_atggtttccaacccact**t**at_MGB	20	72			
S2A	Deletion 145 (TTA)	VUI202012_SFwd2A	tgttgttattaaagtctgtgaatttcaatttt	32	59.9			
		VUI202012_SRev2A	tcgcactagaataaactctgaactcact	28	59.1	120	Set 2	7,000
		VUI202012_SProbe2A	FAM_atccatttttgggtg**tt**taccaca_MGB	24	70			
S2B	S982A (T->G)	VUI202012_SFwd2B	aattttggtgcaatttcaagtgttt	25	58.9			
		VUI202012_SRev2B	acctatcaatttgcacttcagcct	24	59.2	113	Set 2	5,000
		VUI202012_SProbe2B	VIC_actttgtcaagacgtg**c**aa_MGB	19	68			

Primer and probe sets used in B.1.1.7 genotyping RT-ddPCR assay for SARS-CoV-2 variant detection. In the "Sequence" column, the bolded and underlined nucleotides reflect the region of deletion or mutation. For instance, for the probe used in S1A to detect the 69-70 deletion, the sequence ACATGT is inserted between the bolded and underlined thymine and cytoside nucleotides. In the probe used for S1B, a thymine nucleotide is included in place of adenosine at position 18 in the probe sequence. Notably, in S2B, although the mutation is a T to G transversion, the probe sequence is a reverse complement and so is reflected by a cytosine nucleotide at position 17 in the S2B probe sequence. Primers were included at 900 nM and probes were used at 250 nM concentrations.

Abbreviations: T_M, primer melting temperature; FAM, 6-carboxyfluorescein; MGB, Minor Groove Binder; VIC, 2'-chloro-7'phenyl-1,4-dichloro-6-carboxy-fluorescein

322 **Table 2.** TaqPath COVID-19 assay cycle thresholds characteristics of SARS-CoV-2

323 specific targets screening for S gene dropouts

Target	C_T mean	C_T median	C_T range	No. Detected
N gene	24.2	22.9	14.4-38.9	1,035
ORF1ab	23.5	22.3	14.3-39.8	1,001
S gene	24.1	22.9	14.5-39.8	994

324

325 Three SARS-CoV-2 specific amplicons, N gene, ORF1ab, and S gene, were targeted

326 with the multiplex RT-PCR ThermoFisher COVID-19 Assay. Over 1,000 samples

327 positive for SARS-CoV-2 were screened for SGTF, with seven candidates having N

328 gene and ORF1ab amplification without S gene detection. Two of these seven

329 candidates were determined to candidates for genotyping by RT-ddPCR and

330 sequencing based on their viral load characteristics (i.e. mean and median C_Ts> 35.0

331 for all targets detected).

332 Abbreviations: C_T, cycle threshold, SGTF, S gene target failure

333

334

335 **Table 3.** RT-ddPCR genotyping results for two SGTF samples

Sample	S1A	S1B	S2A	S2B
U.K. Variant #1_55538	344 (860,000)	306 (765,000)	323 (807,500)	336 (840,000)
U.K. Variant #2_55545	289 (72,250)	241 (60,250)	267 (66,750)	275 (68,750)
Pos Control[1]	705	680	588	630
Pos Control[2]	660	733	654	833
Neg Control[1]	0	0	0	0
Neg Control[2]	0	0	0	0

336

337 Copies per µl RNA for each target amplified in multiplex RT-ddPCR reaction.

338 Parentheticals denote virus copies/mL back-calculated to consider RNA extraction and

RT-ddPCR dilution factors. Pos control = [1]Previously-identified and sequenced B.1.1.7 clinical sample; [2]G-block. Neg control = [1]Clinical (non-B.1.1.7) SARS-CoV-2 positive; [2]water.

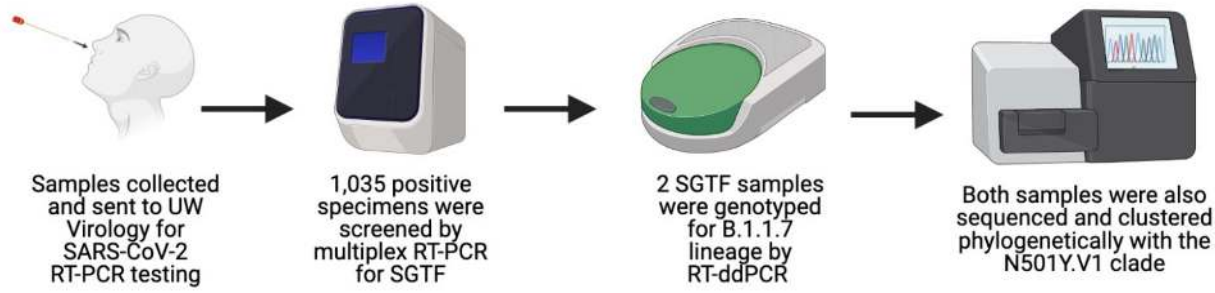
Abbreviations: SGTF, S gene target failure

Table 4. Next-generation sequencing characteristics

Sample	Raw Reads	Mapped Reads	Mean Coverage	Spike Mean Coverage	Min. Spike Coverage	Mean MAPQ	Mean BQ	% N
55538	8,918,592	8,635,583	42,594.10	26,062.40	101	14,262.80	39.63	0
55545	5,054,244	4,657,412	21,376.40	13,034.80	74	14,100.20	39.53	0

FIGURES

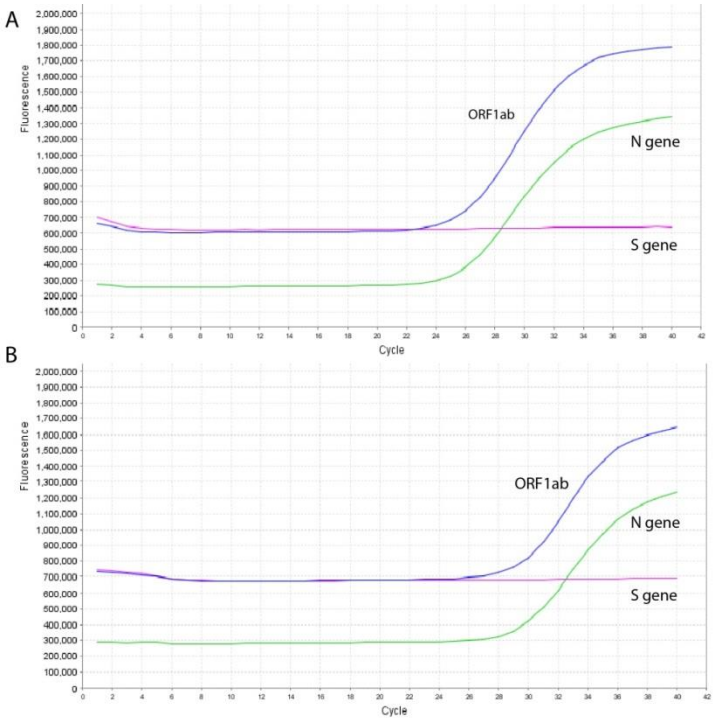
Figure 1. RT-PCR, RT-ddPCR, and sequencing surveillance effort for detection of B.1.1.7 lineage



Samples sent to UW Virology for SARS-CoV-2 molecular detection were screened for S gene transcript failure (SGTF) by multiplex RT-PCR as a proxy for variant detection. SARS-CoV-2 positive samples with SGTF were subsequently genotyped for specific allelic discrimination by RT-ddPCR. Both SGTF samples that genotyped as B.1.1.7

lineage by RT-ddPCR were also sequenced and clustered with N501Y.V1 clades in the United States.

Figure 2. Multiplexed qRT-PCR fluorescence of S gene dropout in SARS-CoV-2 positive samples



Here, multicomponent amplification plots are shown for (A) U.K. Variant #1 (55538) and (B) U.K. Variant #2 (55545). PCR cycle is plotted on the X axis, with quantity of fluorescence detected in real-time on the Y axis. Two out of three fluorophores are detected by qRT-PCR for both samples on the 7500 Real-Time PCR System (Life Technologies, Carlsbad, CA, USA) and analyzed on 7500 software v.2.3 (Life Technologies). VIC and FAM are reporters for N gene and ORF1ab, respectively and are seen here with robust amplification. ABY, the probe for the S gene, is present in the

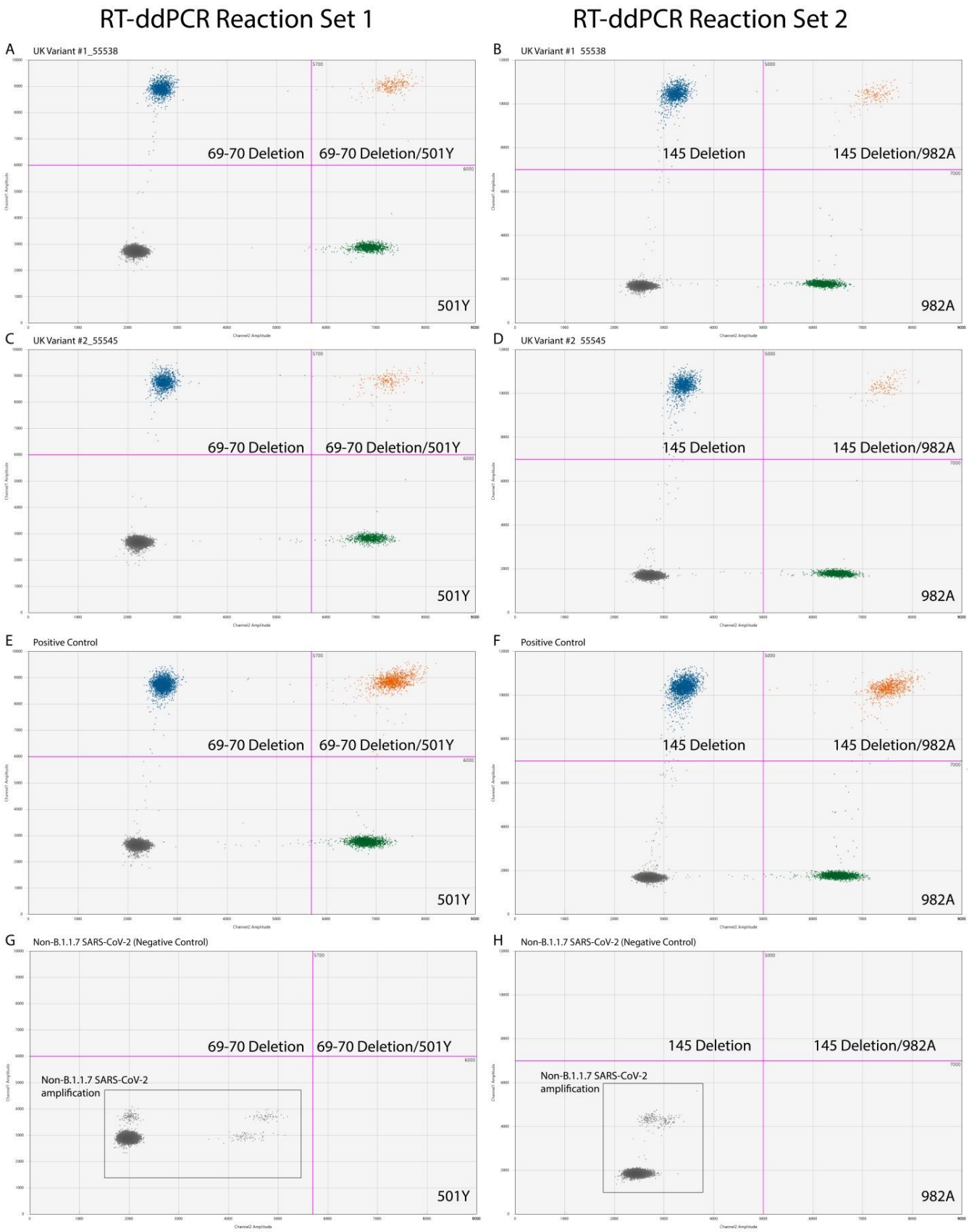
371 reaction mix, but is not detected by fluorescence, indicating a potential candidate for a

372 SARS-CoV-2 variant

373

374

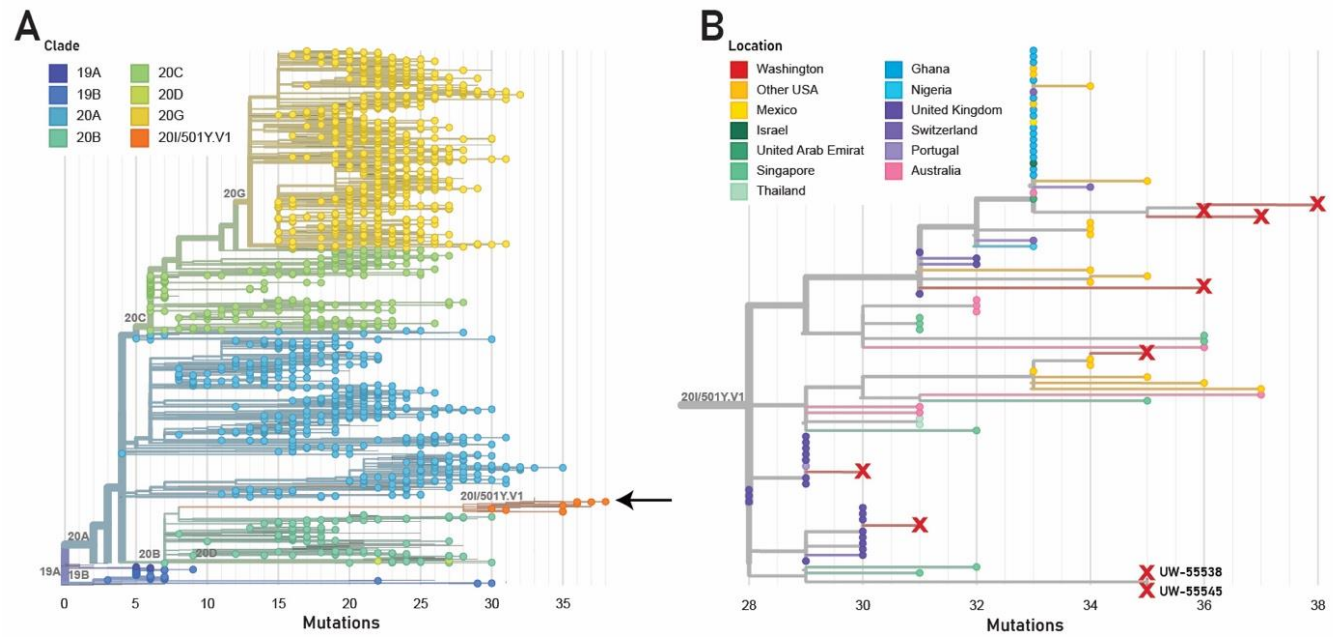
375 **Figure 3.** RT-ddPCR amplification results for SARS-CoV-2 B.1.1.7 lineage



376

(A) U.K. Variant Sample #1 (55538) and (C) U.K. Variant Sample #2 (55545) have amplification for AA69-70del and N501Y mutation. (E) Positive control demonstrates amplification for AA69-70del and N501Y mutation and (G) B.1.1.7 Negative control (Non-B.1.1.7 SARS-CoV-2) shows no amplification for AA69-70del and N501Y mutation for RT-ddPCR amplicon set 1. (B) U.K. Variant Sample #1 (55538) and (D) U.K. Variant Sample #2 (55545) have amplification for AA145 deletion and 982A mutation. (F) Positive control demonstrates amplification for AA145 deletion and 982A mutation and (G) B.1.1.7 Negative control (Non-B.1.1.7 SARS-CoV-2) shows no amplification for AA145 deletion and 982A mutation for RT-ddPCR amplicon set 2.

Figure 4. Phylogenetic tree focused on Washington SARS-CoV-2 samples collected from November 2020 through February 2021



In (A) the phylogenetic tree is filtered to only show Washington samples; 501Y.V1 (B.1.1.7) samples are shown in orange (arrow). (B) 501Y.V2 clade showing Washington samples (red X) in context of global SARS-CoV-2, selected by genetic proximity to the Washington samples. UW-55538 & UW-55545 separately cluster at the bottom of the clade.

REFERENCES

1. 2020. 1st reported US case of COVID-19 variant found in Colorado. AP NEWS.

2. California Identifies A Case Of Coronavirus Variant First Seen In U.K. NPR.org.

3. CDC. 2020. Coronavirus Disease 2019 (COVID-19). Cent Dis Control Prev.

4. WHO | SARS-CoV-2 Variant – United Kingdom of Great Britain and Northern Ireland. WHO. World Health Organization.

5. Dai L, Gao GF. 2021. Viral targets for vaccines against COVID-19. Nat Rev Immunol 21:73–82.

6. Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday J, Pearson CAB, Russell TW, Tully DC, Washburne AD, Wenseleers T, Gimma A, Waites W, Wong KL, van Zandvoort K, Silverman JD, CMMID COVID-19 Working Group, Diaz-Ordaz K, Keogh R, Eggo RM, Funk S, Jit M, Atkins KE, Edmunds WJ. 2020. Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in England. preprint, Epidemiology.

412 7. La Scola B, Le Bideau M, Andreani J, Hoang VT, Grimaldier C, Colson P, Gautret
413 P, Raoult D. 2020. Viral RNA load as determined by cell culture as a management
414 tool for discharge of SARS-CoV-2 patients from infectious disease wards. Eur J
415 Clin Microbiol Infect Dis 39:1059–1061.

416 8. Kim M-C, Cui C, Shin K-R, Bae J-Y, Kweon O-J, Lee M-K, Choi S-H, Jung S-Y,
417 Park M-S, Chung J-W. 2021. Duration of Culturable SARS-CoV-2 in Hospitalized
418 Patients with Covid-19. N Engl J Med NEJMc2027040.

419 9. Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, Ladhani S,
420 Zambon M, Gopal R. 2020. Duration of infectiousness and correlation with RT-PCR
421 cycle threshold values in cases of COVID-19, England, January to May 2020.
422 Eurosurveillance 25.

423 10. Kampf G, Lemmen S, Suchomel M. 2020. Ct values and infectivity of SARS-CoV-2
424 on surfaces. Lancet Infect Dis 0.

425 11. van Kampen JJA, van de Vijver DAMC, Fraaij PLA, Haagmans BL, Lamers MM,
426 Okba N, van den Akker JPC, Endeman H, Gommers DAMPJ, Cornelissen JJ,
427 Hoek RAS, van der Eerden MM, Hesselink DA, Metselaar HJ, Verbon A, de
428 Steenwinkel JEM, Aron GI, van Gorp ECM, van Boheemen S, Voermans JC,
429 Boucher CAB, Molenkamp R, Koopmans MPG, Geurtsvankessel C, van der Eijk
430 AA. 2021. Duration and key determinants of infectious virus shedding in
431 hospitalized patients with coronavirus disease-2019 (COVID-19). Nat Commun
432 12:267.

433 12. 2020. Solutions for surveillance of the S gene mutation in the B.1.1.7 (501Y.V1)
434 SARS-CoV-2 strain lineage. Bench.

435 13. 2020. Identification of a novel SARS-CoV-2 Spike 69-70 deletion lineage
436 circulating in the United States - SARS-CoV-2 coronavirus / SARS-CoV-2
437 Molecular Evolution. Virological.

438 14. Correspondent EC CNN Senior Medical. Much of US data to catch newest
439 coronavirus variants is several months old. CNN.

440 15. Shu Y, McCauley J. 2017. GISAID: Global initiative on sharing all influenza data –
441 from vision to reality. Eurosurveillance 22.

442 16. Xie X, Zou J, Fontes-Garfias CR, Xia H, Swanson KA, Cutler M, Cooper D,
443 Menachery VD, Weaver S, Dormitzer PR, Shi P-Y. 2021. Neutralization of N501Y
444 mutant SARS-CoV-2 by BNT162b2 vaccine-elicited sera. preprint, Microbiology.

445 17. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J,
446 Doolabh D, Pillay S, San EJ, Msomi N, Mlisana K, von Gottberg A, Walaza S,
447 Allam M, Ismail A, Mohale T, Glass AJ, Engelbrecht S, Van Zyl G, Preiser W,
448 Petruccione F, Sigal A, Hardie D, Marais G, Hsiao M, Korsman S, Davies M-A,
449 Tyers L, Mudau I, York D, Maslo C, Goedhals D, Abrahams S, Laguda-Akingba O,
450 Alisoltani-Dehkordi A, Godzik A, Wibmer CK, Sewell BT, Lourenço J, Alcantara
451 LCJ, Pond SLK, Weaver S, Martin D, Lessells RJ, Bhiman JN, Williamson C, de
452 Oliveira T. 2020. Emergence and rapid spread of a new severe acute respiratory

453 syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike
454 mutations in South Africa. preprint, Epidemiology.

455 18. Mahase E. 2021. Covid-19: Novavax vaccine efficacy is 86% against UK variant
456 and 60% against South African variant. BMJ n296.

457 19. Nalla AK, Casto AM, Huang M-LW, Perchetti GA, Sampoleo R, Shrestha L, Wei Y,
458 Zhu H, Jerome KR, Greninger AL. 2020. Comparative Performance of SARS-CoV-
459 2 Detection Assays Using Seven Different Primer-Probe Sets and One Assay Kit. J
460 Clin Microbiol 58:e00557-20, /jcm/58/6/JCM.00557-20.atom.

461 20. Perchetti GA, Nalla AK, Huang M-L, Jerome KR, Greninger AL. 2020. Multiplexing
462 primer/probe sets for detection of SARS-CoV-2 by qRT-PCR. J Clin Virol
463 129:104499.

464 21. Perchetti GA, Nalla AK, Huang M-L, Zhu H, Wei Y, Stensland L, Loprieno MA,
465 Jerome KR, Greninger AL. 2020. Validation of SARS-CoV-2 detection across
466 multiple specimen types. J Clin Virol 128:104438.

467 22. Kidd M, Richter A, Best A, Mirza J, Percival B, Mayhew M, Megram O, Ashford F,
468 White T, Moles-Garcia E, Crawford L, Bosworth A, Plant T, McNally A. 2020. S-
469 variant SARS-CoV-2 is associated with significantly higher viral loads in samples
470 tested by ThermoFisher TaqPath RT-QPCR. preprint, Infectious Diseases (except
471 HIV/AIDS).

472 23. Addetia A, Lin MJ, Peddu V, Roychoudhury P, Jerome KR, Greninger AL. 2020.
473 Sensitive Recovery of Complete SARS-CoV-2 Genomes from Clinical Samples by

474 Use of Swift Biosciences' SARS-CoV-2 Multiplex Amplicon Sequencing Panel. J
475 Clin Microbiol 59:JCM.02226-20, e02226-20.

476 24. Li H. 2011. A statistical framework for SNP calling, mutation discovery, association
477 mapping and population genetical parameter estimation from sequencing data.
478 Bioinformatics 27:2987–2993.

479 25. Katoh K. 2002. MAFFT: a novel method for rapid multiple sequence alignment
480 based on fast Fourier transform. Nucleic Acids Res 30:3059–3066.

481 26. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P,
482 Bedford T, Neher RA. 2018. Nextstrain: real-time tracking of pathogen evolution.
483 Bioinformatics 34:4121–4123.

484 27. SARS-CoV-2 Variants of Concern | CDC.

485 28. Wu K, Werner AP, Moliva JI, Koch M, Choi A, Stewart-Jones GBE, Bennett H,
486 Boyoglu-Barnum S, Shi W, Graham BS, Carfi A, Corbett KS, Seder RA, Edwards
487 DK. 2021. mRNA-1273 vaccine induces neutralizing antibodies against spike
488 mutants from global SARS-CoV-2 variants. preprint, Immunology.

489 29. Liu Y, Liu J, Xia H, Zhang X, Fontes-Garfias CR, Swanson KA, Cai H, Sarkar R,
490 Chen W, Cutler M, Cooper D, Weaver SC, Muik A, Sahin U, Jansen KU, Xie X,
491 Dormitzer PR, Shi P-Y. 2021. Neutralizing Activity of BNT162b2-Elicited Serum —
492 Preliminary Report. N Engl J Med NEJMc2102017.

493 30. Wu K, Werner AP, Koch M, Choi A, Narayanan E, Stewart-Jones GBE, Colpitts T,
494 Bennett H, Boyoglu-Barnum S, Shi W, Moliva JI, Sullivan NJ, Graham BS, Carfi A,
495 Corbett KS, Seder RA, Edwards DK. 2021. Serum Neutralizing Activity Elicited by
496 mRNA-1273 Vaccine — Preliminary Report. N Engl J Med NEJMc2102179.

497 31. Mahase E. 2021. Covid-19: South Africa pauses use of Oxford vaccine after study
498 casts doubt on efficacy against variant. BMJ n372.

499 32. Stanford Medicine launches large-scale surveillance of coronavirus variants in Bay
500 Area | News Center | Stanford Medicine.

501 33. Bal A, Destras G, Gaymard A, Stefic K, Marlet J, Eymieux S, Regue H, Semanas
502 Q, d'Aubarede C, Billaud G, Laurent F, Gonzalez C, Mekki Y, Valette M,
503 Bouscambert M, Gaudy-Graffin C, Lina B, Morfin F, Josset L, the COVID-Diagnosis
504 HCL Study Group. 2021. Two-step strategy for the identification of SARS-CoV-2
505 variant of concern 202012/01 and other variants with spike deletion H69–V70,
506 France, August to December 2020. Eurosurveillance 26.

507 34. Washington NL, Gangavarapu K, Zeller M, Bolze A, Cirulli ET, Barrett KMS, Larsen
508 BB, Anderson C, White S, Cassens T, Jacobs S, Levan G, Nguyen J, Ramirez JM,
509 Rivera-Garcia C, Sandoval E, Wang X, Wong D, Spencer E, Robles-Sikisaka R,
510 Kurzban E, Hughes LD, Deng X, Wang C, Servellita V, Valentine H, De Hoff P,
511 Seaver P, Sathe S, Gietzen K, Sickler B, Antico J, Hoon K, Liu J, Harding A,
512 Bakhtar O, Basler T, Austin B, Isaksson M, Febbo PG, Becker D, Laurent M,
513 McDonald E, Yeo GW, Knight R, Laurent LC, de Feo E, Worobey M, Chiu C,
514 Suchard MA, Lu JT, Lee W, Andersen KG. 2021. Genomic epidemiology identifies

515 emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States.
516 preprint, Infectious Diseases (except HIV/AIDS).

517 35. 2021. Update on the Helix, Illumina surveillance program: B.1.1.7 variant of SARS-
518 CoV-2, first identified in the UK, spreads further into the US. Helix.

519 36. The US is way behind on tracking COVID-19 variants. Pop Sci.

520 37. Wang R, Hozumi Y, Yin C, Wei G-W. 2020. Mutations on COVID-19 diagnostic
521 targets. Genomics 112:5204–5213.

522 38. Khan KA, Cheung P. 2020. Presence of mismatches between diagnostic PCR
523 assays and coronavirus SARS-CoV-2 genome. R Soc Open Sci 7:200636.

524 39. Rana DR, Pokhrel N. 2020. Sequence mismatch in PCR probes may mask the
525 COVID-19 detection in Nepal. Mol Cell Probes 53:101599.

526 40. Galloway SE, Paul P, MacCannell DR, Johansson MA, Brooks JT, MacNeil A,
527 Slayton RB, Tong S, Silk BJ, Armstrong GL, Biggerstaff M, Dugan VG. 2021.
528 Emergence of SARS-CoV-2 B.1.1.7 Lineage — United States, December 29,
529 2020–January 12, 2021. MMWR Morb Mortal Wkly Rep 70:95–99.

530 41. Long SW, Olsen RJ, Christensen PA, Subedi S, Olson R, Davis JJ, Saavedra MO,
531 Yerramilli P, Pruitt L, Reppond K, Shyer MN, Cambric JE, Finkelstein IJ, Gollihar J,
532 Musser J. 2021. Sequence Analysis of 20,453 SARS-CoV-2 Genomes from the
533 Houston Metropolitan Area Identifies the Emergence and Widespread Distribution
534 of Multiple Isolates of All Major Variants of Concern. preprint, Pathology.

535 42. Pinheiro-de-Oliveira TF, Fonseca-Júnior AA, Camargos MF, Laguardia-Nascimento
536 M, Giannattasio-Ferraz S, Cottorello ACP, de Oliveira AM, Góes-Neto A, Barbosa-
537 Stancioli EF. 2019. Reverse transcriptase droplet digital PCR to identify the
538 emerging vesicular virus Senecavirus A in biological samples. Transbound Emerg
539 Dis 66:1360–1369.

540 43. Li H, Bai R, Zhao Z, Tao L, Ma M, Ji Z, Jian M, Ding Z, Dai X, Bao F, Liu A. 2018.
541 Application of droplet digital PCR to detect the pathogens of infectious diseases.
542 Biosci Rep 38:BSR20181170.

543 44. Tong Y, Shen S, Jiang H, Chen Z. 2017. Application of Digital PCR in Detecting
544 Human Diseases Associated Gene Mutation. Cell Physiol Biochem 43:1718–1730.

545 45. White RA, Blainey PC, Fan HC, Quake SR. 2009. Digital PCR provides sensitive
546 and absolute calibration for high throughput sequencing. BMC Genomics 10:116.

547