Severe acute respiratory syndrome coronavirus 2 spike ARTIC amplicon 76 dropout in relation to primers, viral load, and variants

Maha G. Haggag[a,b], Massimiliano S. Tagliamonte[b,c], Melanie N. Cash[b,c], Carla N. Mavian[b,c]

[a]Microbiology & Immunology Unit, Research Institute of Ophthalmology, Giza, Egypt, [b]Pathology, Immunology and Laboratory Medicine, Emerging Pathogens Institute, Gainesville, Florida, USA, [c]Department of Pathology, Immunology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, FL, USA

Correspondence to Maha G. Haggag, MD, Allergy Laboratory, Microbiology and Immunology Unit, Department of Microbiology and Parasitology, Research Institute of Ophthalmology (RIO), Giza, Cairo 12211, Egypt. Tel: +20 112 309 2244; Fax: 35735688; e-mail: mghaggag@gmail.com

Received: 30 May 2022
Revised: 23 June 2022
Accepted: 30 June 2022
Published: 24 December 2022

Journal of The Arab Society for Medical Research 2022, 17:180–187

Background/aim

In January 2020, the ARTIC Network designed 98 pairs of primers divided into two pools for targeting amplification of the severe acute respiratory syndrome coronavirus 2 genome by multiplex PCR. However, in using this protocol, several users, including our team at Emerging Pathogens Institute at University of Florida, noticed a systematic dropout of reads covering amplicon 18, which overlaps the gene for nonstructural protein 3 (nsp3) in ORF1a, and amplicon 76, targeting the spike (S) gene. The aim of this work was to verify the design of primers for the dropout region of amplicon 76 in the V3 ARTIC protocol covering the severe acute respiratory syndrome coronavirus 2 spike gene, to evaluate correlation of dropout to viral load and determine if there is a correlation between dropouts and viral load and viral lineages and as well as to further explore potential mutations in the primer regions.

Materials and methods

A total of 544 samples presenting dropout at amplicon 76 from the V3 ARTIC primer set were collected over time from December 2020 to December 2021 in Alachua County, Florida. RNA was extracted, followed by cDNA synthesis and PCR amplification, which targeted the dropout region within the spike. PCR amplification was performed with forward and reverse primers designed in-house. Sequencing was performed to detect potential mutations at primer sites. Viral load was performed for 96 samples showing a dropout of the spike genome at amplicon 76. The results were compared and correlated to lineage/variant of concern.

Results

A total of 544 dropout samples were amplified with the primers designed in-house. Overall, 381 (70%) of these samples showed bands visible by gel electrophoresis. This result, along with Sanger sequencing, was enough to verify the efficacy of our designed primers in amplifying the dropout region. Sequencing revealed that primer regions were conserved; no mutations were observed at the site corresponding to the Illumina primers for amplicon 76. The mean Ct value of dropouts was 32.53, and the mean Ct value of nondropouts was 26.11. Welch test on Ct values of dropouts versus nondropout showed significant difference ($P<10^{-6}$). The 96 dropout samples tested for viral load included 17 different lineages, with delta lineage (B.1.617.2-like) being the most common (21.8%), followed by B.1.2 (10.4%), alpha (B.1.1.7-like) (9.4%), B.1.234 (9.4%), iota (B.1.526-like) (5.2%), and gamma (P.1-like) (4.2%). Unexpectedly, when determining the effect the lineage on the chance of a sample being a dropout, the delta lineage had an odds ratio of less than 1, implying that a sequence belonging to delta lineage was less likely to produce a dropout.

Conclusions

Our study showed that dropouts are not due to mutation at the primer sites. Viral loads likely affect the odds of a sample presenting a dropout of amplicon 76 but was not the only cause of dropout. Several dropout samples could not be classified because of undefined genomic segments, which may have skewed the resulting odds of lineage on sequencing. It is possible, as previously suggested, that dropouts may be due to interaction and dimer formation between primers of the multiplexed PCR reaction.

Keywords:

amplicon 76, ARTIC, delta lineage, severe acute respiratory syndrome coronavirus 2, spike Illumina dropout, viral load

J Arab Soc Med Res 17:180–187

1687-4293

Introduction

Coronavirus disease 2019, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), emerged in the province of Wuhan in China in November 2019 [1]. Soon after its rapid spread around the world, coronavirus disease 2019 was declared a global pandemic by the WHO [2]. SARS-CoV-2 is a positive-strand RNA virus of the Betacoronavirus genus from *Coronaviridae* family. The spike (S) protein, encoded by the S gene, has a key role in receptor recognition and cell membrane fusion process, and the emergence of new variants often relies on the appearance of mutations within this region [3].

The current SARS-CoV-2 pandemic has significantly affected populations and health care systems worldwide. One of the challenges created by SARS-CoV-2 like other RNA viruses is its ability of genomic divergence, which may lead to the emergence of variants with improved transmissibility, increased virulence, and/or immune-escape ability. Variants for which obvious evidence is provided about such modified virus properties, which are likely to have an epidemiological impact, have been categorized as variants of concern (VOCs) by the WHO [4].

Next-generation sequencing (NGS) genomic surveillance has greatly contributed to the investigation of the epidemiology and dynamics of the disease escalating genomic investigations and accelerating the response to outbreaks. NGS genomic surveillance has been a crucial tool for detection and tracking the evolution of SARS-CoV-2, allowing molecular epidemiology surveillance, and identifying emerging new lineages and VOC [1].

Global SARS-CoV-2 full genome sequencing was achieved thanks to the 196 primers (98 pairs) designed by the ARTIC Network [5], using the SARS-CoV-2 genome, with accession number MN908947.3 as reference [6]. The method is based on targeted amplification of the SARS-CoV-2 genome by multiplex PCR: each of the 98 pairs of proposed PCR primers in the protocol produces amplicons of ~400 bp [3]. This protocol was rapidly adopted by laboratories worldwide and resulted in a large number of nearly complete genomes, 98% of which are uploaded into the Global Initiative on Sharing Avian Influenza Data (GISAID) database [7].

However, 1 year into the pandemic, several users worldwide noticed a systematic dropout, meant as failing of producing reads with Illumina sequencing, of amplicons 18 and 76 [8].

Itokawa *et al.* [9] investigated the issue and concluded that the acute dropout of amplicons 18 and 76 was due to the formation of primer dimers between the forward primer for amplicon 18 and the reverse primer for amplicon 76, exactly owing to an interaction between 18_LEFT and 76_RIGHT, which were complementary to one another by 10 nucleotides at their 3′ ends. They also demonstrated how replacing 76_RIGHT, in the Pool 2 reaction with a newly designed primer 76_RIGHTv2 (5′-TCTCTGCCAA ATTGTTGGAAAGGCA-3′), which is located 48-nt downstream from 76_RIGHT, would solve the issue. Quick and Loman [8] on behalf of the ARTIC Consortium have incorporated the suggestion by Itokawa *et al.* [9] (2019-76_RIGHTv2) in their V2 protocol and substituted nCoV-2019_18_LEFT for nCoV-2019_18_LEFT_alt2.

However, this change caused other amplicons to dropout instead [8]. Therefore, a V3 protocol was further developed, which included the addition of alternative primers (alts), to improve coverage in problematic amplicons.

Baker *et al.* [10] presented a new platform technique, Coronavirus high-throughput, which is a more flexible, cost-effective, and simpler method for sequencing of SARS-CoV-2 genomes and has the flexibility to react with the pandemic at the local and national level. They demonstrated that all methods are unpredictable at producing high-quality consensus genomes from positive clinical samples with diagnostic RT-qPCR Cts above 32 (~100 viral genome copies), and they reported that Coronavirus high-throughput performed better in those samples, possibly owing to the additional rounds of PCR during barcoding.

With the evolution of new dominant variants propagating in June 2021 (such as B.1.1.7, B.1.351, B.1.429, B.1.525, B.1.617.1, B.1.617.2, and P.1.), updating the ARTIC V3 primer design was needed. The ARTIC V4 design produces 99 overlapping amplicons and has been optimized to amplify across the entire viral genome with the same efficiency through variable concentrations of the primer pairs [11].

The aim of this work was to verify primers' design for dropout region of amplicon 76 in the SARS-CoV-2 spike gene collected between December 2020 and December 2021, to explore if different lineages presented mutations in the primer regions, and finally, to evaluate whether there was a correlation between viral load and dropout.

Materials and methods

Samples

We investigated 544 samples that presented dropout in amplicon 76 from cases in Alachua County, Florida, USA, collected throughout the timeframe of December 2020 to December 2021. The samples used comprised 363 saliva and 181 nasal swabs in viral transport medium.

Ethical approval

This study was conducted in accordance with the University of Florida Institutional Review Board (IRB) under project number IRB202000633. UF IRB is a fully accredited human research protection program that protects the rights and welfare of participants in human participant's research studies. UF IRB reviews all research involving human participants to ensure that the welfare and rights of research participants are protected as mandated by federal and state laws, local policies, and ethical principles.

RNA extraction

RNA was extracted from 544 SARS-CoV-2-positive samples using QIAamp 96 Viral RNA Extraction kit (Qiagen, Germantown, Maryland, USA), according to the following protocol: viral RNA was extracted from 180 µl of each sample using a QIAcube HT instrument (Qiagen) with the following settings with a filter plate: the lysed sample was pre-mixed eight times before subjecting to vacuum for 5 min at 25 kPa and vacuum for 3 min at 70 kPa. Following three washes using the same aforementioned vacuum conditions, the samples were eluted in 75 µl of AVE buffer followed by a final vacuum for 6 min at 60 kP. Next, RNA was used for cDNA synthesis and library preparation using the Illumina COVIDSeq Test kit (Illumina, San Diego, California, USA) and Mosquito HV Genomics Liquid Handler (SPT Labtech Inc., Covina, California, USA). The size and purity of each library were determined using the 4200 TapeStation System (Agilent, Santa Carla, California, USA) and the Qubit dsDNA HS Assay Kit (Life Technologies, Santa Carla, California, USA) according to the manufacturer's instructions. Constructed libraries were pooled and sequenced using the NovaSeq. 6000 Sequencing System S1 Reagent Kit and the NovaSeq Xp 2-Lane Kit. Illumina's DRAGEN pipeline was used to derive sample consensus sequences, which were filtered based on a minimum of 70% coverage of the genome.

Synthesis of complementary DNA

Complementary DNA was synthesized using New England BioLabs ProtoScript II First Strand cDNA Synthesis Kit (Ipswich, Massachusetts, USA), according to the following protocol: 2 µl Random Primer Mix was added to 10 µl of reaction mix containing 2 µl of enzyme mix to 6 µl of the extracted RNA with total volume of 20 µl per sample. The first strand synthesis was conducted in a thermocycler using the following conditions: 25°C for 5 min, 42°C for 1 h, 80°C for 5 min, and 4°C (1 h 12 min).

PCR amplification of the dropout region with in-house designed forward and reverse primers

Platinum Green Hot Start PCR Master Mix (Invitrogen, Carlsbad, California, USA) was used according to the manufacturer's protocol. The sequences of our designed forward and reverse primers were:

Dropout-F (5'-3') GGCAAACTGGAAAGATTGC TGA.
Dropout-R (5'-3') GAACACCTGTGCCTGTTA AACC.

The in-house designed primers flanked the dropout region between nucleotides 22 800 and 23 216 for the published reference SARS-CoV-2 genome with accession number MN908947.3 The forward started at bp 22 800 and the reverse started at bp 23 216 (ARTIC primers V3 set No. 76). Stock primers were diluted to 10 µM. PCR was done according to the following protocol using a thermocycler: for each cDNA sample, 25 µl reactions were comprised of 12.5 µl Platinum Green Hot Start PCR Master Mix (2X) (Thermo Fisher, Waltham, Massachusetts, USA), 0.5 µl dropout-F primer, 0.5 µl dropout-R primer, 2 µl cDNA, and 9.5 µl nuclease-free water. PCR conditions were as follows: one cycle at 94°C for 2 min; 35 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 30 s; one cycle at 72°C for 10 min, and 7–4°C (1 : 40 min).

Gel electrophoresis of the PCR products

Five microliters from each 25 µl cDNA sample was run on an agarose gel using the Leader EZ Load 500-bp ladder as a reference marker.

Viral load

A total of 96 samples were selected for viral load evaluation according to the following protocol.

Levels of SARS-CoV-2 RNA were determined using the 2019-nCoV_N1 assay (primer and probe set) with 2019-nCoV_N_positive control (IDT, Coralville, Iowa, USA) per guidelines outlined by the Centers for Disease Control and Prevention (CDC) [12] Viral RNA was extracted as previously described and then subjected to first-strand synthesis using ProtoScript II Reverse Transcriptase according to the manufacturer's instructions (New England Biolabs). Quantitative PCR was performed using TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific) according to the manufacturer's instructions. A standard curve was generated using N1 quantitative standards 10-fold diluted to determine viral copies. The assay was run in triplicate including a nontemplate control.

Purification of the PCR products for all 96 samples

Samples that were evaluated for viral load and showed a single band by gel electrophoresis were prepared for Sanger sequencing. Purification was done using QIAquick PCR purification kit (Qiagen) through three steps: first step was binding the DNA by adding 100 μl of binding buffer PB to each 20 μl PCR sample, mixed by pipetting and adding immediately to the QIAquick column, centrifuged for 30 s at 13 000 rpm, discarding flow-through and placing the QIAquick column into the same collection tube. Second step was washing by adding 750 μl of wash buffer PE to the column and centrifugation for 30 s at 13 000 rpm, discarding flow-through and placing the QIAquick column into the same collection tube to be additionally centrifuged for 1 min at 13 000 rpm to ensure no residual ethanol from buffer PE was left, and then the column was placed in a clean 1.5 ml microcentrifuge tube. Third step was elution of DNA by adding 50 μl buffer EB (10 mM tris Cl, pH 8.5) to the center of the QIAquick membrane and the column was centrifuged for 1 min at 13 000 rpm eluting the DNA in the microcentrifuge tube.

Quantification of purified DNA

It was completed for samples that showed intense bands by gel electrophoresis using spectrophotometer/fluorometer (DeNovix Ds-11FX) according to the manufacturer's protocol.

Sequencing

Genewiz Sanger Sequencing was done for the 96 samples that were evaluated for viral load. Two 96-well plates were prepared: one for sequencing with the Dropout-F primer and one for sequencing with the Dropout-R primer. Of 96 samples, 31 showed faint bands by gel electrophoresis. Ten microliters from each sample of those 31 was placed into each of the two plates, and 5 μl from the dropout-F and 5 μl from the dropout-R primers were added to the corresponding plate. A total of 65 samples showed intense bands by gel electrophoresis and were diluted first with nuclease-free water according to quantification results obtained in the previous step. Calculation of dilution was done for every sample so as to get 10 ng DNA in 10 μl. Chromatogram was visualised and analyzed by AliView software program [13].

Results

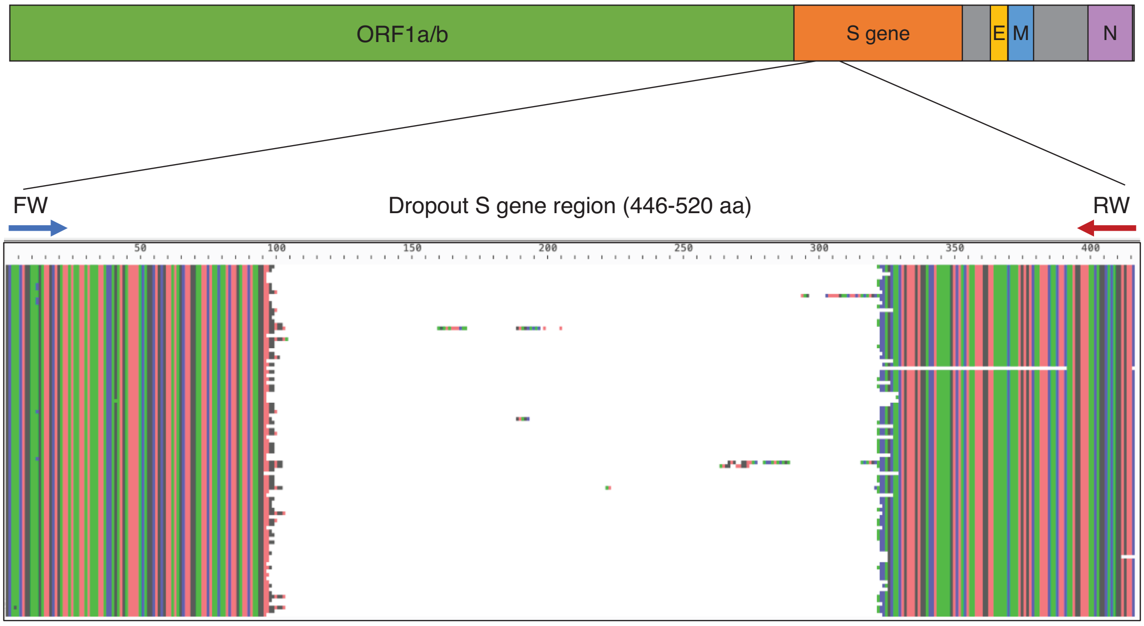
The amplicon dropouts were defined as having gap between nucleotides 22,895–23,128, which correspond to the region amplified by the amplicon 76 from the ARTIC primers V3 set. The dropout region corresponds to the amino acids 446–520 in the spike (Fig. 1).

Of 544 samples that were verified as containing the dropout by NGS, 381 (70%) could be amplified with our designed primers showing a product by gel electrophoresis (Fig. 2). This result was enough to verify the efficacy of the designed primers to amplify the dropout region.

Each amplicon produced by the dropout primers was sequenced and confirmed by Illumina. Fasta sequences showed that our designed dropout primer regions are conserved and showed very few mutations (4.3%), as shown in Fig. 3.

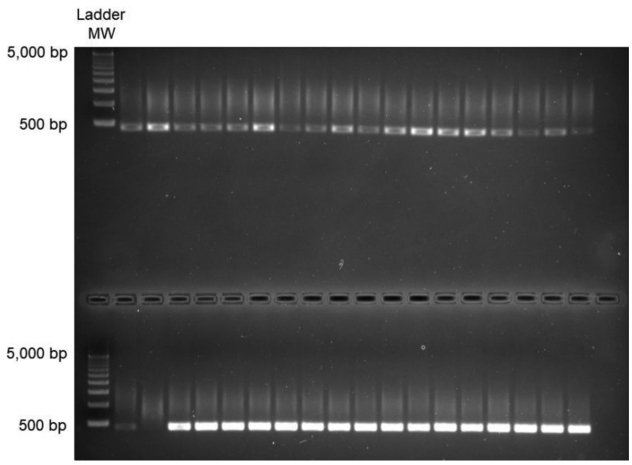
We randomly selected 96 dropout and 96 nondropout samples to investigate possible correlation between Ct value and dropout. After performing qPCR on 96 dropout and 96 nondropout samples, 16 samples of the dropout group resulted below the limit of detection. These samples had Ct values at 38.33, which our detection limit was based on the qPCR standard curve. The distribution of Ct values for each group is shown in Fig. 4. The median Ct value for dropouts was 33.21, whereas the median Ct value for nondropouts was 25. When comparing the means by two-tailed Welch test, where dropout Ct mean was 32.53 and nondropouts Ct mean was 26.11, we obtained a P value of 9.32×10^{-13}. Furthermore, we performed logistic regression to investigate the effect of Ct value and lineage on dropout status. Logistic regression for the dropout being a function of Ct values resulted in odds ratio of 1.19, that is, a 0.19 increase in odds of being a dropout for each unit increase of Ct value. The 96 dropout samples included 17 different lineages classified with

Figure 1



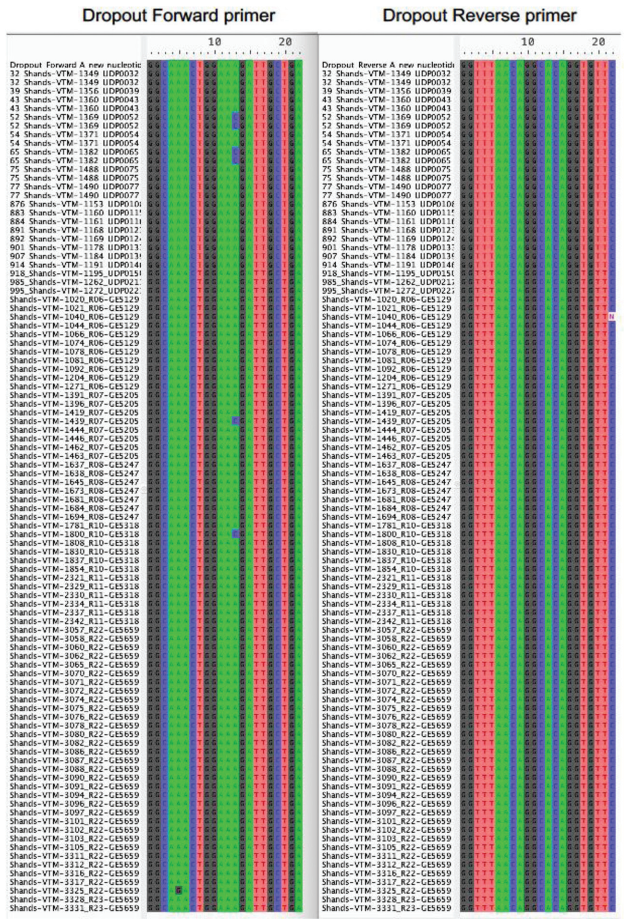
SARS-CoV-2 dropout region of S gene. The dropout region in S amplicon 76 is between amino acids 446 and 520. Forward and reverse primers regions are shown in blue and red, respectively (genbank number QHD43416.1 of SARS-CoV-2 genome accession number MN908947.3). SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

Figure 2



Gel electrophoresis of PCR products of 36 dropout samples. The figure shows the reference marker leader EZ MW range 5000–500 bp. The dropout bands correspond to 500-bp MW.

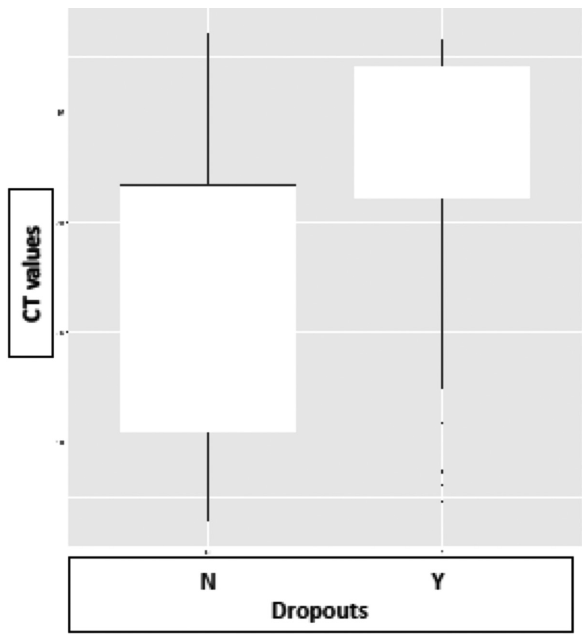
Figure 3



Sequences of forward and reverse primer regions corresponding to the genomes with dropout in amplicon 76. The figure shows the genetic information in FASTA format for the Illumina amplicon 76 primer regions.

Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin), where delta lineage (B.1.617.2-like) was the most common (21.8%), followed by B.1.2 (10.4), alpha (B.1.1.7-like) (9.4%), B.1.234 (9.4), Iota (B.1.526-like) (5.2%), gamma (P.1-like) (4.2%), and others (Table 1). Delta lineage was also the most common (76%) among 13 different lineages in the nondropout samples (Table 2). When factoring in the lineage information, only the delta variant coefficient had significant P value (P=0.00031). The odds ratio of 0.32 though implied that a sequence

Figure 4



Ct values of dropout and nondropout samples. The figure shows the distribution of qPCR Ct values for dropout and non dropout groups. Comparing the means by two-tailed Welch test resulted in a P value of 9.32×10^{-13}.

Table 1 Type and number of lineages of 96 dropout samples

Lineages	Samples	Percentages
Delta (B.1.617.2 like)	21	21.8
B.1.2	10	10.4
Alpha (B.1.1.7-like)	9	9.4
B.1.234	9	9.4
Iota (B.1.526-like)	5	5.2
B.1	5	5.2
B.1.596	5	5.2
Gamma (P.1-like)	4	4.2
B.1.311	3	3.1
B.1.580	2	2.1
B.1.243	2	2.1
B.1.564	1	1.0
B.1.265	1	1.0
B.1.1	1	1.0
B.1.1.294	1	1.0
B.1.604	1	1.0
B.1.1.434	1	1.0
None	15	15.6

belonging to a delta variant is less likely to produce a dropout. Dropout sequences are also more likely to be of an undefined lineage, with Pearson's χ^2 test with Yates' continuity correction P value of 1.21×10^{-5}. In summary, as also visible from the partial overlap in distributions from the boxplots, Ct values do affect the odds of being dropout, but are likely not the only cause.

Table 2 Type and number of lineages of 96 nondropout samples

Lineages	Samples	Percentages
Delta	73	76
B.1.2	7	7.3
B.1.234	5	5.2
B.1.596	2	2.1
Alpha (B.1.1.7-like)	1	1.0
B.1	1	1.0
Gamma (P.1-like)	1	1.0
B.1.580	1	1.0
B.1.243	1	1.0
B.1.577	1	1.0
B.1.1.28	1	1.0
B.1.509	1	1.0
B.1.621	1	1.0

The lineages reported in Tables 1 and 2 are classified with Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin) [14].

Discussion

The ARTIC protocol is the most common amplicon-based protocol for whole-genome amplification of SARS-CoV-2. However, amplicon dropout may occur if a given primer is not hybridizing to its specific complementary sequence. This might happen for multiple reasons, such as dimerization with other primers during the multiplex reaction, or presence of mutations arising at the primer-binding site owing to continuous evolution of SARS-CoV-2. Dropout often requires redesign and revalidation of primers [2,9].

The ARTIC V3 primer design is known as the 'work horse' for SARS-CoV-2 amplification. However, researchers reported that certain amplicons such as 72, 74, and 76 are poorly covered by this primer pool. This reduced performance affects consensus sequence generation and high-resolution coverage of the spike protein gene in recent variants, such as delta, particularly in low viral titer samples [11].

Ct value represents the cycle number at which the signal reaches the threshold for positivity; a lower Ct value is indicative of a higher viral load. Walker *et al.* [15] suggested that marked variation in community SARS-CoV-2 Ct values could be a useful epidemiological early-warning indicator. Waudby-West *et al.* [16] demonstrated an independent relationship between the Ct value (indicative of viral load) of an individual's first positive SARS-CoV-2

PCR test and overall mortality in which a lower Ct value is associated with greater possibility of death.

We are not the only ones that have been investigating dropout regions in the spike. Previously, Borcard *et al.* [3] reported the region surrounding the start of the spike gene of SARS-CoV-2 delta variant genome sequences was systematically under-sequenced. The region they investigated was from positions 21 357–22 246, flanked by primers 71L and 73R of the ARTIC V3 protocol. They provided a series of alternative primers for 72R and 73L, mainly addressing users of the ARTIC V3 protocol and targeting circulating delta variants in Switzerland in summer 2021, and the result was better amplification and sequencing of 75% success rate of the tested delta variant samples.

Despite the evolution of SARS-CoV-2, in the present study, we do not observe any mutation in the region of the primer site, that is, primer sites for amplicon 76 are conserved. The dropout region we have investigated is not to be mistaken for the dropout of Spike gene commonly reported by many researchers in clinical settings that are performing RT-PCR SARS-CoV-2 tests, such as the company Helix, and that is due to mutation at the primer site for that specific PCR amplification [17]. These mutations produce thermodynamic instabilities between primer oligonucleotides and genomic template substance interfering with amplification, eventually causing dropouts in sequencing coverage and hampering the collection of complete, high-quality viral genomes as explained by Kuchinski *et al.* [18]; however, this is not the cause in our study.Lambisia *et al.* [19] reported that the ability to produce almost-complete genomes when using the ARTIC Network's V3 primers is affected by sample quality, viral load quantity, and consistent virus evolution that guarantees mutations on primer-binding sites leading to amplicon drop-offs in up to 12 amplicon primer sites across the delta, alpha, and beta variants. Moreover, the trials to improve genome recovery by using supplemental primers or increasing primer concentrations do not always ensure success and can be a challenge. Hence, the ARTIC Network's V4 primers were released to address mutations in the primer-binding sites that were resulting in amplicon drop-offs in the delta VOC.

Our results showed that low viral load is correlated with amplicon dropout in the spike region, as well as a sequence belonging to a delta variant was less likely to produce a dropout. Isolates of undefined lineage were also present in the dropout group.

Despite the fast evolutionary rate of SARS-CoV-2 that is accumulating mutations leading to the emergence of different lineages and VOC, we found a high degree of conservation – that is absence of mutations – in the primer regions for the amplicon involved in dropout, and a correlation between dropout and low viral load.

Our analyses highlighted how the primer region was conserved in the isolates in our data set. It is therefore likely that a combination of low viral load, together with other factors, such as dimerization, as reported by Itokawa *et al.* [9], are at the root of the issue. Such factors are not necessarily the only ones in play, as the rise of the omicron variant, later to our sample collection, showed.

Conclusions

Our study showed that dropouts are not due to mutations at the primer sites. Viral loads likely affect the odds of amplicon 76 dropping out. The results of delta variant on dropout odds may be biased by the number of undetermined lineages that were also present in the dropout group. It is possible, as previously suggested, that dropouts may be owing to interaction and dimer formation between primers of the multiplexed PCR reaction. Constant surveillance of SARS-CoV-2 evolution is needed as we face continuous emergence of new VOC to promptly address any potential sequencing failures owing to the rise of new mutations.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1 Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, *et al.* A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med 2020; 382:727–733.

2 Rosenthal SH, Gerasimova A, Ruiz-Vega R, Livingston K, Kagan RM, Liu Y, *et al.* Development and validation of a high throughput SARS-CoV-2 whole genome sequencing workflow in a clinical laboratory. Sci Rep 2022; 12:2054–2069

3 Borcard L, Gempeler S, Terrazos Miani MA, Baumann C, Grädel C, Dijkman R, *et al.* Investigating the extent of primer dropout in SARS-CoV-2 genome sequences during the early circulation of delta variants. Front Virol 2022; 2:840952–840961.

4 Giovacchini N, Coppi M, Aiezza N, Baccani I, Malentacchi F, Pollini S, *et al.* Rapid screening for SARS-CoV-2 VOC-Alpha (202012/01, B.1.1.7) using the Allplex^{TM} SARS-CoV-2/FluA/FluB/RSV Assay. Short communication. Int J Infect Dis 2021; 113:207–209.

5 ARTIC Network. ARTIC Network - real-time molecular epidemiology for outbreak response [Internet]. January 22, 2021. Available at: https://artic.network/. [Accessed date 9 Sep 2020]

6 Wu F, Zhao S, Yu B, Chen Y-M., Wang W, Song Z-G, *et al.* A new coronavirus associated with human respiratory disease in China. Nature 2020; 579:265–269.

7 Elbe S, Buckland-Merrett G. Data, disease and diplomacy: GISAID's innovative contribution to global health. Global Chall 2017; 1:33–46.

8 Quick J, Loman N. hCoV-2019/nCoV-2019 Version 3 Amplicon Set. ARTIC Consortium. March 24, 2020. Available at: https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w. [Accessed date 24 March 2020]

9 Itokawa K, Sekizuka T, Hashino M, Tanaka R, Kuroda M. Disentangling primer interactions improves SARS-CoV-2 genome sequencing by multiplex tiling PCR. PLoS ONE 2020; 15:e0239403.

10 Baker DJ, Aydin A, Le-Viet T, Kay GL, Rudder S, Martins Lde O, *et al.* CoronaHiT: high-throughput sequencing of SARS-CoV-2 genomes. Genome Med 2021; 13:21.

11 Clark CR, Hardison MT, Houdeshell HN, Vest AC, Whitlock DA, Skola DD, *et al.* Evaluation of an optimized protocol and Illumina ARTIC V4 primer pool for sequencing of SARS-CoV-2 using COVIDSeq^{TM} and DRAGEN^{TM} COVID Lineage App workflow bioRxiv preprint. Available at:. https://doi.org/10.1101/2022.01.07.475443. [Accessed date 7 Jan 2022]

12 Centers for Disease Control and Prevention (CDC) National Center for Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases June 6, 2020. Available at: https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html/.

13 Larsson A. A fast and lightweight alignment viewer and editor for large data sets. Bioinformatics 2014; 30:3276–3278.

14 O'Toole A, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, *et al.* Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. Virus Evol 2021; 7:1–9.

15 Walker AS, Pritchard E, House T, Robotham JV, Birrell PJ, Bell I; *et al.* COVID-19 Infection Survey team. Ct threshold values, a proxy for viral load in community SARS-CoV-2 cases, demonstrate wide variation across populations and over time. Elife. 2021;10:e64683. doi: 10.7554/eLife.64683. PMID: 34250907; PMCID: PMC8282332.

16 Waudby-West R, Parcell BJ, Palmer CNA, Bell S, Chalmers JD, Siddiqui MK. The association between SARS-CoV-2 RT-PCR cycle threshold and mortality in a community cohort. Eur Respir J 2021; 58:2100360.

17 Washington NL, White S, Barrett KMS, Cirulli ET, Bolze A, Lu JT. S genedropout patterns in SARS-CoV-2 tests suggest spread of the H69del/V70del mutation in the US. medRxiv preprint doi: https://doi.org/10.1101/2020.12.24.20248814; this version posted December 30, 2020. The copyright holder for this preprint is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

18 Kuchinski KS, Nguyen J, Lee TD, Hickman R, Jassem AN, Hoang LMN, *et al.* Mutations in emerging variant of concern lineages disrupt genomic sequencing of SARS-CoV-2 clinical specimens. medRxiv preprint doi: https://doi.org/10.1101/2021.06.01.21258181; this version posted June 4, 2021. The copyright holder for this preprint is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

19 Lambisia AW, Mohammed KS, Makori TO, Ndwiga L, Mburu MW, Morobe JM, *et al.* Optimization of the SARS-CoV-2 ARTIC Network V4 primers and whole genome sequencing protocol. Front Med (Lausanne) 2022; 9:836728.