



Identification of Novel Mutations in the N Gene of SARS-CoV-2 That Adversely Affect the Detection of the Virus by Reverse Transcription-Quantitative PCR

Rashedul Hasan,^a  Mohammad Enayet Hossain,^a Mojnu Miah,^a Md Mahmudul Hasan,^a  Mustafizur Rahman,^a Mohammed Ziaur Rahman^a

^aInternational Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh

KEYWORDS SARS-CoV-2, COVID-19, RT-qPCR, N gene, mutation

COVID-19 originated in late 2019 in Wuhan, China, engendering a pandemic. As of 4 June 2021, there had been more than 173 million confirmed cases and more than 3.7 million deaths reported globally. Laboratory diagnosis of SARS-CoV-2 is the mainstay for alleviating viral spread, since it facilitates the identification and isolation of infected individuals. Since reverse transcription-quantitative PCR (RT-qPCR) oligonucleotides, primers, and probes bind to a diminutive, ~20- to 25-bp region, mutations in this region can potentiate impaired primer and/or probe binding and subsequent inefficient PCR amplification, thereby spawning false negatives. Given the high transmissibility of SARS-CoV-2 (1), such false negatives could have affected the containment of COVID-19 outbreaks. Genetic evolution is a widespread and natural phenomenon for RNA viruses. SARS-CoV-2 is an RNA virus. The nature of its evolution contributes to the emergence of new variants and its demonstrated diagnostic nonuniformity, especially when conducting RT-qPCR using different primers and probes (2).

Here, we report the detection of novel mutations in the N gene of SARS-CoV-2 corresponding to the China CDC N RT-qPCR assay and the determination of the frequency in sequences submitted from Bangladesh.

The International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), in association with the government of Bangladesh, has been actively testing for SARS-CoV-2 since March 2020 as part of the countrywide COVID-19 laboratory network. We have been using the World Health Organization (WHO)-recommended RT-qPCR protocol for the detection of the ORF1ab and N genes of SARS-CoV-2 (3). Occasionally, we have observed two issues during the RT-qPCR: (i) positive results with cycle threshold (C_T) values of 21.66 and 17.48 for the ORF1ab region, but no curve for the N genes, and (ii) a higher C_T value for the N gene compared to the ORF1ab region (31.55 versus 15.10, respectively). The target samples were retested using the US CDC novel coronavirus (nCoV) N1 RT-qPCR assay (3), which yielded positive results for all the targets. Because of the failure of the current assay to amplify the N gene, mutations in the primer- and/or probe-binding sites on the SARS-CoV-2 N gene were suspected. To identify the possible mutations responsible for the PCR failure, we sequenced the entire N gene of two controls and three selected samples using Sanger sequencing. cDNA was prepared using the high-capacity cDNA reverse transcription kit (Thermo Fisher Scientific, CA, USA) as per the kit instructions. The N gene was amplified using the ARTIC primer sets (4) with GoTaq G2 Hot Start *Taq* polymerase (Promega, WI, USA). Nucleotide sequencing was performed in an automated ABI 3500xL genetic analyzer (Applied Biosystems, Foster City, CA) using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems). The sequenced data were inspected by BLAST search, contigs were generated using the SeqMan tool (DNASTAR, Inc.), and multiple sequence alignment was performed using the BioEdit sequence alignment editor. Triple nucleotide substitutions (GGG28881AAC)

Citation Hasan R, Hossain ME, Miah M, Hasan MM, Rahman M, Rahman MZ. 2021. Identification of novel mutations in the N gene of SARS-CoV-2 that adversely affect the detection of the virus by reverse transcription-quantitative PCR. *Microbiol Spectr* 9:e00545-21. <https://doi.org/10.1128/Spectrum.00545-21>.

Editor S. Wesley Long, Houston Methodist Hospital

Copyright © 2021 Hasan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Mohammad Enayet Hossain, enayet.hossain@icddr.org.

Published 25 August 2021

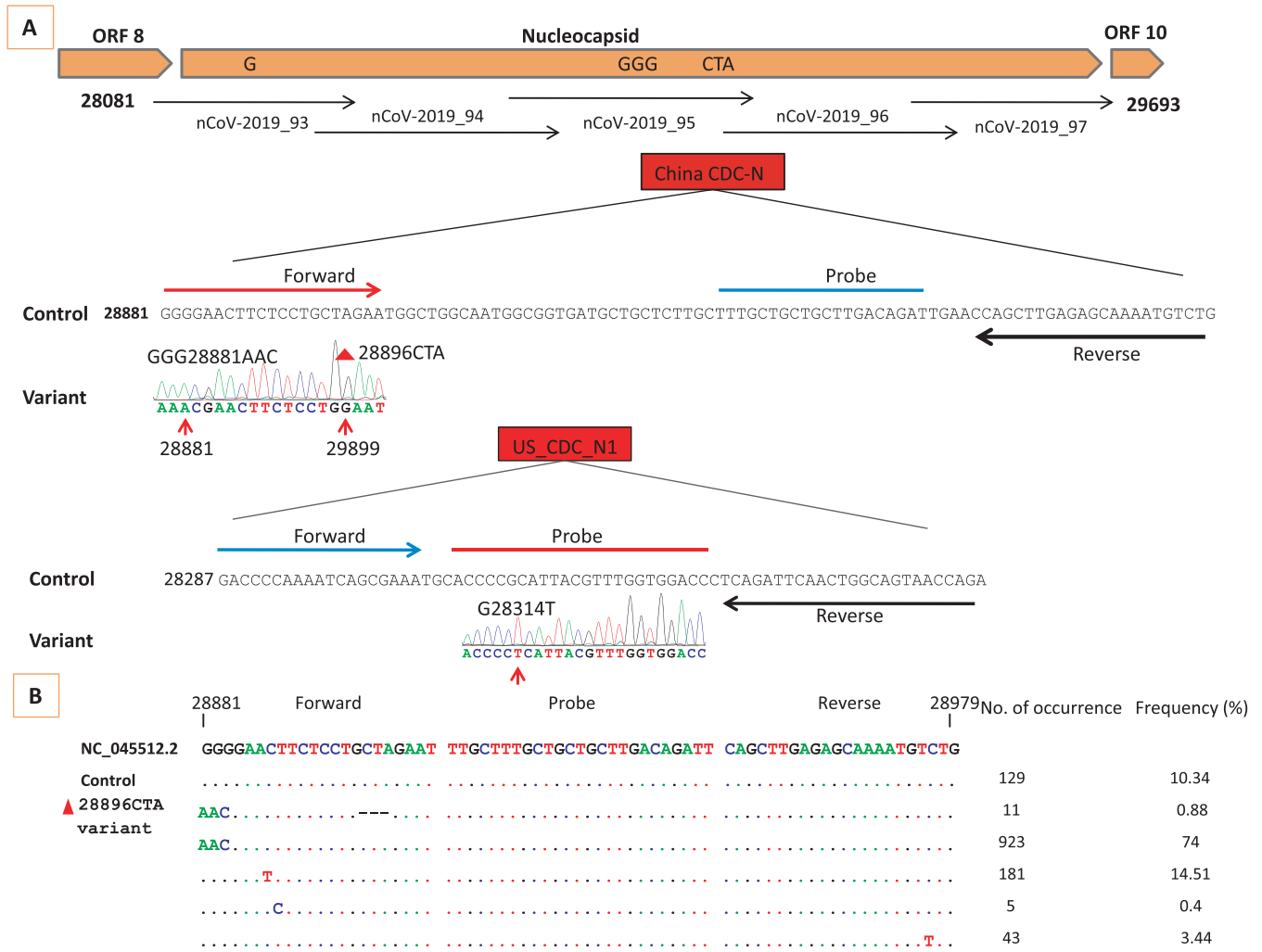


FIG 1 (A) Graphic representation of mutations in the N gene of a SARS-CoV-2 virus and strategy to unveil the suspected mutations linked to detection failure in the China CDC N RT-qPCR assay. The arrows and numbers correspond to the fragments of the ARTIC primer sets for N gene amplification. (B) Frequency of the N gene mutations in the primer-binding regions detected in the study in 1,245 viral genome sequences from Bangladesh. The dots indicate nucleotides identical to those of the primers and probes.

at the 5' end and a triple nucleotide deletion (28896CTA_{del}) at the 3' end of the forward primer-binding site (3) were detected in the suspected samples (Fig. 1A). Interestingly, one nucleotide substitution (G28314T) was also observed in the probe-binding region for the U.S. CDC nCoV N1 RT-qPCR assay, although the mismatch did not impair detection in the confirmatory test (Fig. 1A).

To estimate the frequency of both the substitutions and the deletion among Bangladeshi strains, we retrieved all complete SARS-CoV-2 sequences (1,247 sequences) from the GISAID EpiCoV database (28 April 2021). Nucleotide mismatches in the Bangladeshi sequences are described in Fig. 1B. In particular, the GGG28881AAC substitution and the 28896CTA deletion were simultaneously detected in 11 sequences (0.88%). According to a local BLASTN search, 40 isolates carry both the 28881AAC variant and the 28896CTA deletion. Based on the GISAID EpiCoV database, 48.8% (649,115/1,329,560) sequences contain the 28881AAC variant (5). Approximately 0.18% (n = 3,803) of the isolate sequences in GISAID contain the target 28896CTA deletion, and 1.54% of the genomes contain at least one mutation in the last pentanucleotides of the 3' primer region for the China CDC N RT-qPCR assay as reported by the GISAID EpiCoV database (5). Isolates containing both the 28881AAC variant and the 28896CTA deletion, which stopped detection of the N gene using the China CDC N RT-qPCR assay, originated from the United States, Australia, England, and Bangladesh. According to the U.

S. CDC nCoV N1 RT-qPCR assay, the probe region contains about 10,105 single-nucleotide polymorphisms (SNPs). The G28314T substitution is carried by 39 and 25 sequences according to GISAID and BLASTN searches, respectively.

RNA viruses such as SARS-CoV-2 are prone to mutations associated with virulence and evolvability, traits considered beneficial for viruses. Mutations in the primer- and probe-binding regions may lead to false-negative diagnoses for SARS-CoV-2. Previous studies have claimed that a single mutation (C29200T or C29200A) in the probe-binding region of an isolate was associated with failure of the virus detection when using the U.S. CDC nCoV N1 RT-qPCR assay (6, 7).

Our data show that SARS-CoV-2 variants carrying mutations in the primer-/probe-binding regions that impair RT-qPCR detection may have emerged and spread in the community. Our study suggests that mutations in the SARS-CoV-2 N gene RT-qPCR primer-binding sites might cause issues related to sensitivity and potentially underdiagnosis if we rely on an assay with only a single target. Mutations in the primer-/probe-binding regions of both targets could potentially result in false-negative detection. According to the GISAID update on common primer checks, the primer-/probe-binding regions for the targets of all WHO in-house SARS-CoV-2 diagnostic assays carried mutations, suggesting the need for continuous monitoring for mutations, especially when one of the RT-qPCR assays fails to detect SARS-CoV-2 (5). This study also highlights the importance of the use of RT-qPCR assays with at least two targets for SARS-CoV-2 detection. If symptomatic COVID-19 patients test negative using one assay, our findings strongly support retesting with another assay to avoid false-negative results.

Data availability. The sequence data have been submitted to GenBank (accession no. [MZ379157](#) and [MZ379158](#)) and GISAID (accession no. EPI_ISL_1715168).

Ethical statement. This study has been approved by the research review board of icddr,b (protocol no. PR-20102).

ACKNOWLEDGMENTS

This study was conducted with the support of the United States Agency for International Development (USAID) under the terms of USAID's Alliance for Combating TB in Bangladesh activity (cooperative agreement [CA] no. 72038820CA00002). The views expressed herein do not necessarily reflect the views of the U.S. Government or USAID. icddr,b acknowledges with gratitude the commitment of USAID to its research efforts. icddr,b also thanks the governments of Bangladesh, Canada, Sweden, and the United Kingdom for providing core/unrestricted support.

REFERENCES

- Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R. 2020. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis* 26:1470–1477. <https://doi.org/10.3201/eid2607.200282>.
- Pfefferle S, Reucher S, Nörz D, Lütgehetmann M. 2020. Evaluation of a quantitative RT-PCR assay for the detection of the emerging coronavirus SARS-CoV-2 using a high throughput system. *Euro Surveill* 25:2000152. <https://doi.org/10.2807/1560-7917.ES.2020.25.9.2000152>.
- WHO. 2020. WHO in-house assays for 2019-novel coronavirus (2019-nCoV) real-time rRT-PCR. <https://www.who.int/docs/default-source/coronaviruse/whoinhouseassays.pdf>.
- Quick J. 2020. nCoV-2019 sequencing protocol v3 (LoCost). protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye.
- Shu Y, McCauley J. 2017. GISAID: Global initiative on sharing all influenza data — from vision to reality. *Euro Surveill* 22. <https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494>.
- Ziegler K, Steininger P, Ziegler R, Steinmann J, Korn K, Ensser A. 2020. SARS-CoV-2 samples may escape detection because of a single point mutation in the N gene. *Euro Surveill* 25:2001650. <https://doi.org/10.2807/1560-7917.ES.2020.25.9.2001650>.
- Hasan MR, Sundararaju S, Manickam C, Mirza F, Al-Hail H, Lorenz S, Tang P. 2021. A novel point mutation in the N gene of SARSCoV-2 may affect the detection of the virus by reverse transcription-quantitative PCR. *J Clin Microbiol* 59:e03278-20. <https://doi.org/10.1128/JCM.03278-20>.