

RESEARCH

Open Access



A two-step process for *in silico* screening to assess the performance of qRT-PCR kits against variant strains of SARS-CoV-2

Swati Gupta^{1†}, Amit Kumar^{2†}, Nivedita Gupta^{1*}, Deepak R. Bharti³, Neeraj Aggarwal¹ and Vasanthapuram Ravi⁴

Abstract

Background: Since inception of the COVID-19 pandemic, early detection and isolation of positive cases is one of the key strategies to restrict disease transmission. Real time reverse transcription polymerase chain reaction (qRT-PCR) has been the mainstay of diagnosis. Most of the qRT-PCR kits were designed against the target genes of original strain of SARS-CoV-2. However, with the emergence of variant strains of SARS-CoV-2, sensitivity of the qRT-PCR assays has reportedly reduced. In view of this, it is critical to continuously monitor the performance of the qRT-PCR kits in the backdrop of variant strains of SARS-CoV-2. Real world monitoring of assay performance is challenging. Therefore, we developed a two-step *in-silico* screening process for evaluating the performance of various qRT-PCR kits used in India.

Results: We analysed 73 qRT-PCR kits marketed in India, against the two SARS-CoV-2 VoCs. Sequences of both Delta (B.1.617.2) and Omicron (B.1.1.529) VoCs submitted to GISAID within a specific timeframe were downloaded, clustered to identify unique sequences and aligned with primer and probe sequences. Results were analysed following a two-step screening process. Out of 73 kits analysed, seven were unsatisfactory for detection of both Delta and Omicron VoCs, 10 were unsatisfactory for Delta VoC whereas 2 were unsatisfactory for only Omicron VoC.

Conclusion: Overall, we have developed a useful screening process for evaluating the performance of qRT-PCR assays against Delta and Omicron VoCs of SARS-CoV-2 which can be used for detecting SARS-CoV-2 VoCs that may emerge in future and can also be redeployed for other evolving pathogens of public health importance.

Keywords: Delta, Omicron, SARS-CoV-2 variant, *In silico* analysis, qRT-PCR kits

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a betacoronavirus with RNA strand size of ~29.9kb [1]. After release of its genome sequence by China CDC on January 12, 2020 [2], several nucleic acid amplification tests (NAAT) have been developed for virus

detection. Among the available NAATs, real-time reverse transcription polymerase chain reaction (qRT-PCR) is considered to be the gold standard test for detection of SARS-CoV-2 RNA in clinical samples [3]. Common genes of SARS-CoV-2 targeted for the development of qRT-PCR kits include Envelop (E), Nucleocapsid (N), Spike (S), Membrane (M) and Open Reading Frames (ORFs) i.e. ORF1 ab which also contains the gene encoding for RNA dependent RNA polymerase (RdRP/nsp12) [4]. As the COVID-19 pandemic evolved, the SARS-CoV-2 virus accumulated multiple mutations, most abundant being in the S gene, thus giving rise to variants of concern (VoCs) and variants of interest (VoIs) [5].

[†]Swati Gupta and Amit Kumar are equal authors.

*Correspondence: drguptanivedita@gmail.com

¹ Division of Epidemiology and Communicable Disease, V. Ramalingaswami Bhawan, Indian Council of Medical Research, Ansari Nagar, New Delhi 110029, India

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

S gene encoded spike protein is located on the surface of the virus, abundantly expressed and facilitates virus entry into the host cell. Earlier in the pandemic, S gene target was increasingly selected for inclusion in the molecular diagnostic tests, due to high sensitivity and specificity of detection of this gene. Later, due to the high mutation rates observed in the S gene, other genes having conserved sequences and relatively lesser rates of mutation were preferred [6]. Diagnostic kits with two or more target genes are now increasingly being used to improve the sensitivity of qRT-PCR assays.

At times, mismatch of even a single nucleotide can drastically affect the performance of a test kit. Mutations in regions specific for the hybridization of primers and probes affects the primer-probe binding efficiency, thus giving false negative or inconclusive results. SARS-CoV-2 Alpha /B.1.1.7 variant, Delta /B.1.617.2 variant and Omicron /B.1.1.529 variant had several mutations in the S gene, due to which S gene target failure (SGTF) or S gene drop out in qRT-PCR tests was observed with Alpha and Omicron VoCs [7–9]. However, despite the mutations in S gene, B.1.617.2 (Delta variant) had no SGTF [10]. Similarly, multiple mutations in N gene affected its detection and thus generated false negative RT-PCR results [11, 12]. Most of the available qRT-PCR tests were designed to have acceptable sensitivity and specificity against the early reported sequences of SARS-CoV-2. However, with the introduction of VoCs and VoIs, the sensitivity of these assays has reportedly reduced [13].

It is therefore imperative to monitor the performance of available NAAT tests in the backdrop of emerging VoCs and VoIs. Here we describe the development of a two-step screening process to assess the performance of existing qRT-PCR test kits for their ability to detect the circulating Delta and Omicron variants across the world, using an *in silico* approach. The screening process developed by us can be used for assessing the utility of various NAAT tests against the SARS-CoV-2 variants in future.

Methodology

Primer and probe sequence retrieval

Since inception of the COVID-19 pandemic, the Indian Council of Medical Research (ICMR) was designated as the nodal agency for validation of the qRT-PCR diagnostic test kits for SARS-CoV-2 in India. Earlier in the pandemic, manufacturers were not mandated to submit their primer probe sequences to ICMR, when applying for validation. Later on, due to the repeated emergence of SARS-CoV-2 VoCs and VoIs, there was a critical need to assess the performance of various test kits already marketed in India. Therefore, ICMR contacted the manufacturers of already approved and marketed qRT-PCR kits and requested them to share the primer and probe sequences.

Subsequently it was made mandatory for all qRT-PCR kit manufacturers to submit the primer-probe sequences at the time of making a validation request to ICMR.

Kit selection and database creation

Out of the 200 companies contacted retrospectively, 76 kit manufacturers reverted back with the sequence of their primers and probes. Out of these, only 73 kits were selected for analysis as kits based on LAMP assay and having spacer in their primer and probe sequences were excluded. A database of all the selected kits was prepared with details of name of the kit, name of the manufacturer, target genes and length of the primer/probes along with their sequences. All selected kits were anonymized and primer/probe sequences were converted into FASTA format for further analysis.

SARS-CoV-2 whole genome sequence (WGS) dataset retrieval

Complete genome sequences with less than 1% ambiguous bases (Ns) of Delta and Omicron VoC of SARS-CoV-2 were downloaded from global initiative on sharing all influenza data (GISAID) database [14–16]. For downloading Delta sequences available in GISAID, keywords Delta, time-period till July 15, 2021, complete sequences (less than 1% ambiguous bases (Ns)) and high coverage were used. For downloading Omicron sequences, keywords, Omicron, time-period till January 19, 2022, complete sequences (less than 1% ambiguous bases (Ns)) and high coverage were used. Incomplete sequences were excluded. A total number of 186,355 and 392,855 complete sequences of Delta and Omicron were included in this study respectively.

Removal of repeat sequences and identification of unique sequences

To spot identical sequences of Delta and Omicron VoCs, all selected genome sequences were clustered using CD-HIT software [17] with sequence identity cut-off equal to 1.0 (other parameters were left at default settings). Unique sequences thus obtained were selected for further analysis/ BLAST search [18]. Similarly, all primer and probe sequences were also subjected to CD-HIT clustering for selection of unique sequences and further BLAST search. Since BLAST search is a computation expensive program, this method was used to reduce computational time.

SARS-CoV-2 genome alignment with primers and probes

Selected unique genome sequences were used to create a database for local BLAST search using “makeblastdb” command. Primer and probe sequences were searched in the database using “blastn” program of BLAST. All the above analysis were done on Intel Xeon Gold server with 64 processor and 256 GB RAM.

Figure 1 represents the process of database creation. Results were parsed and a comparative table was generated for further analysis. Pairwise alignment of each forward & reverse primer and probe sequence with the full genome sequences of the two VoCs were retrieved under the following categories:

- Percentage of alignment of each primer/probe with the selected VoCs of SARS-CoV-2.
- Length of the primer/probe sequence match with the full genome of SARS-CoV-2, against the total length of the primer/probe.
- Number of mismatches and gaps between the primer/probe and the VoC sequence on their alignment.
- 5' and 3' start and end location of primer/probe alignment with the VoC sequences.

Mismatch identification

To obtain the positions of mismatches in primers and probes, EMBOSS Water (EMBOSS: 6.6.0.0) package [19] was used. EMBOSS Water is a modified version of the Smith-Waterman algorithm [20] for local alignment.

Results were parsed to obtain the mismatched positions for each database entry and later summarised to obtain position-wise mismatches across the length of the primers and probe.

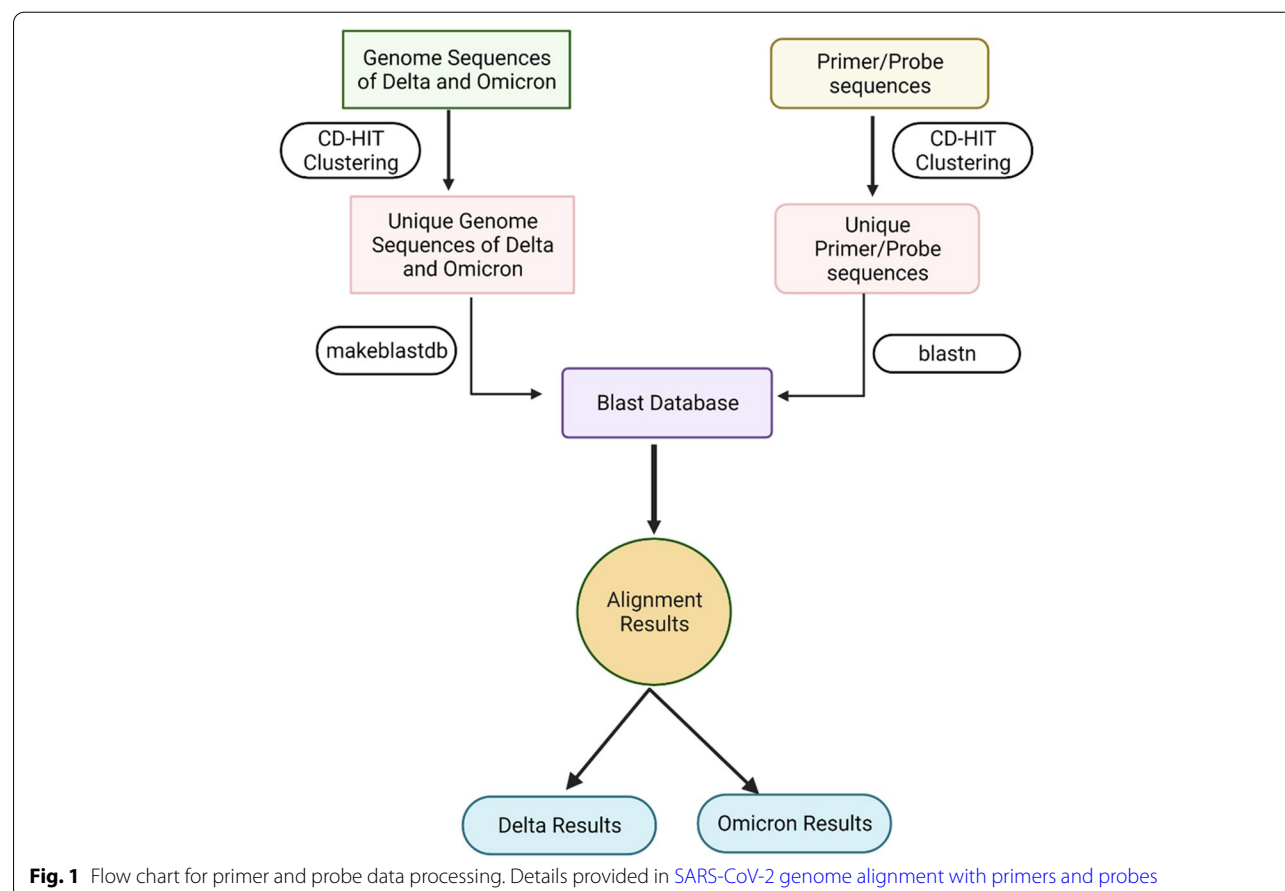
Two-step screening process for assessing the performance of qRT-PCR kits by *in silico* approach

Based on the available evidence and published literature [21–27], a two-step screening process was developed. This two-step screening process is the first of its kind where all the parameters for screening of primers and probes have been included. In the first level screening, complete global sequences of VoC submitted to GISAID were considered and subsequently, in the second level screening, complete sequences submitted to GISAID from India were analysed.

Following were the screening criteria used in analysis:

First level screening criteria

- Each primer or probe sequence (irrespective of the length) under review, must exhibit at least 95%



- alignment with the genome sequences of the VoC, downloaded from GISAID.
- ii. Full length of the primer or probe must match with at least 95% of the sequences with which they aligned in step (i) of screening.
 - iii. Even a single mismatch at 3' or 5' ends was considered unacceptable [21, 23–26].
 - iv. Missing or mismatch of three contiguous nucleotides in primer/ probe alignment was considered unacceptable [27].
 - v. A single nucleotide mismatch in the middle of the probe was considered unacceptable. Therefore, for both odd and even number nucleotides in the probe, central location i.e. -1, 0, +1 nucleotides must align with the genome sequence [21, 25].
 - vi. Notwithstanding any of the above, primers/probes having degenerate nucleotides were considered acceptable if even a single combination of nucleotide meets all the above criteria.

In addition to the above criteria, the following decision-making matrix was developed for assessing the in silico performance of testing kits

- a. E gene was considered as screening whereas S, N, ORF, M and RdRp were considered as confirmatory targets. In addition, for single S gene assays, S gene was considered as confirmatory and in two or more gene assay it was considered as a screening gene. For the two gene assays, wherein E and S genes were targets, the S gene was considered as confirmatory [28, 29].
- b. A strategy was developed to identify the centre of the probe. E.g. for probe with odd number of nucleotides, e.g. 11 nucleotides, nucleotide at number six was considered at central location (0) and -1 and +1 nucleotides were at 5 and 7 location respectively. For probes with even number of nucleotides, e.g. 12 nucleotides, 6.5 was considered as the central location and therefore nucleotides at position 6 and 7 which are -1 and +1 were considered for detecting mismatch location.
- c. Kits with two or more gene targets were assessed on the following basis:
Primers and probes of each individual gene target to essentially meet the criteria i to v and vi, if applicable. For kits using a combination of both screening and confirmatory genes, if the confirmatory gene met the criteria i to v and vi, if applicable, irrespective of the performance of the screening gene target, the kit performance was labelled as satisfactory.

- d. For kits using a combination of both screening and confirmatory genes, if the confirmatory gene did not meet the screening criteria i to v and vi, if applicable, the kit was labelled as unsatisfactory irrespective of the performance of the screening gene target.

Second level screening criteria

Primers or probes found to be unsatisfactory as per criteria number (i) and (ii) in the first level screening analysis, were subjected to second level screening analysis to rule out the possibility due to huge number of global sequences submitted to GISAID. Sequences having mismatch at 3' or 5' ends or probes having mismatch at central location i.e. -1, 0, +1 nucleotides were rejected and were not considered for reanalysis. For reconfirmation, representative complete genome sequences of SARS-CoV-2 submitted to GISAID from India were selected for reanalysis. Criteria (i) to (vi) of second level screening analysis were the same as those used for the first level screening analysis except for the difference of use of only Indian VoC.

The second level screening provided advantage to the kits that did not qualify in first level screening due to stringent criteria of 95% alignment with total number of downloaded sequences and 95% length wise alignment with downloaded sequences. Sequences submitted from India were also part of first level screening but huge number of global sequences resulted in rejection of some kits based on 95% criteria.

Selection of representative SARS-CoV-2 genomes from India

Complete genome sequences with less than 1% ambiguous bases (Ns) submitted to GISAID from India were downloaded for Delta with keywords Delta, time-period till July 15, 2021, complete sequences (less than 1% ambiguous bases (Ns)), high coverage and India. For downloading Omicron sequences, keywords, Omicron, time-period till January 19, 2022, complete sequences (less than 1% ambiguous bases (Ns)), high coverage and India, were used. A total of 17,238 Delta and 7994 Omicron genome sequences were downloaded. These sequences were selected for second level screening. Subsequently, clustering was done to group the identical sequences and identifying the unique sequences using the CD-HIT software. The primers or probes which could not meet the second level screening analysis were finally considered as unsatisfactory.

Results

qRT-PCR kit types based on the number of SARS-CoV-2 target genes

Data obtained from qRT-PCR kit manufacturers was classified into four groups based on the number of genes

used for detection of SARS-CoV-2. Kits having one set of primers and probes for SARS-CoV-2 gene detection were considered as single gene kits. Similarly, two, three and four gene kits had two, three and four sets of primers and probes respectively for SARS-CoV-2 target gene detection. A total of 3, 53, 16 and 1 kits were in the one, two, three and four gene kit category respectively.

Number of unique sequences of SARS-CoV-2 identified for first and second level screening analysis

For the first level screening analysis, a total of 1,86,355 Delta and 3,92,855 Omicron genome sequences were downloaded from GISAID. On clustering, 1,39,352 and 3,15,943 unique sequences were identified for Delta and Omicron respectively. These sequences were finally used to create database for local BLAST search (Fig. 2A and B).

For the second level screening analysis, a total of 17,238 Delta and 7994 Omicron genome sequences were downloaded. On clustering, 8449 and 7685 unique sequences were identified for Delta and Omicron, respectively. Figure 2 describes the process of clustering followed for identification of unique sequences of Delta and Omicron VoCs for the first (Fig. 2A and B) and second level screening analysis (Fig. 2C and D).

Results of the first and second level screening analysis

Primer & probe alignment data for both Delta and Omicron VoCs was obtained and analysed by applying the first level screening criteria. Results of the analysis are summarized in Table 1. The standalone four gene kit was found to perform satisfactorily for detection of both VoCs. 81% of three gene and 74% of two target gene kits were found to be satisfactory for detection of the Delta VoC, whereas 81% of three and 64% of two target gene kits were found to be satisfactory for detection of the Omicron VoC.

Further, after application of the second level screening criteria, the overall satisfactory alignment percentage for two target gene kits increased to 75 and 87% for Delta and Omicron sequences, respectively. Whereas, for three target gene kits for Omicron VoC the satisfactory alignment increased to 94%. However, there was no change for Delta VoCs for three target gene kits.

Overall, irrespective of the number of target genes, by applying the two-step screening process, 77 and 88% of already marketed RT-CPR kits met the *in silico* screening criteria for achieving a satisfactory score for Delta and Omicron VoCs, respectively.

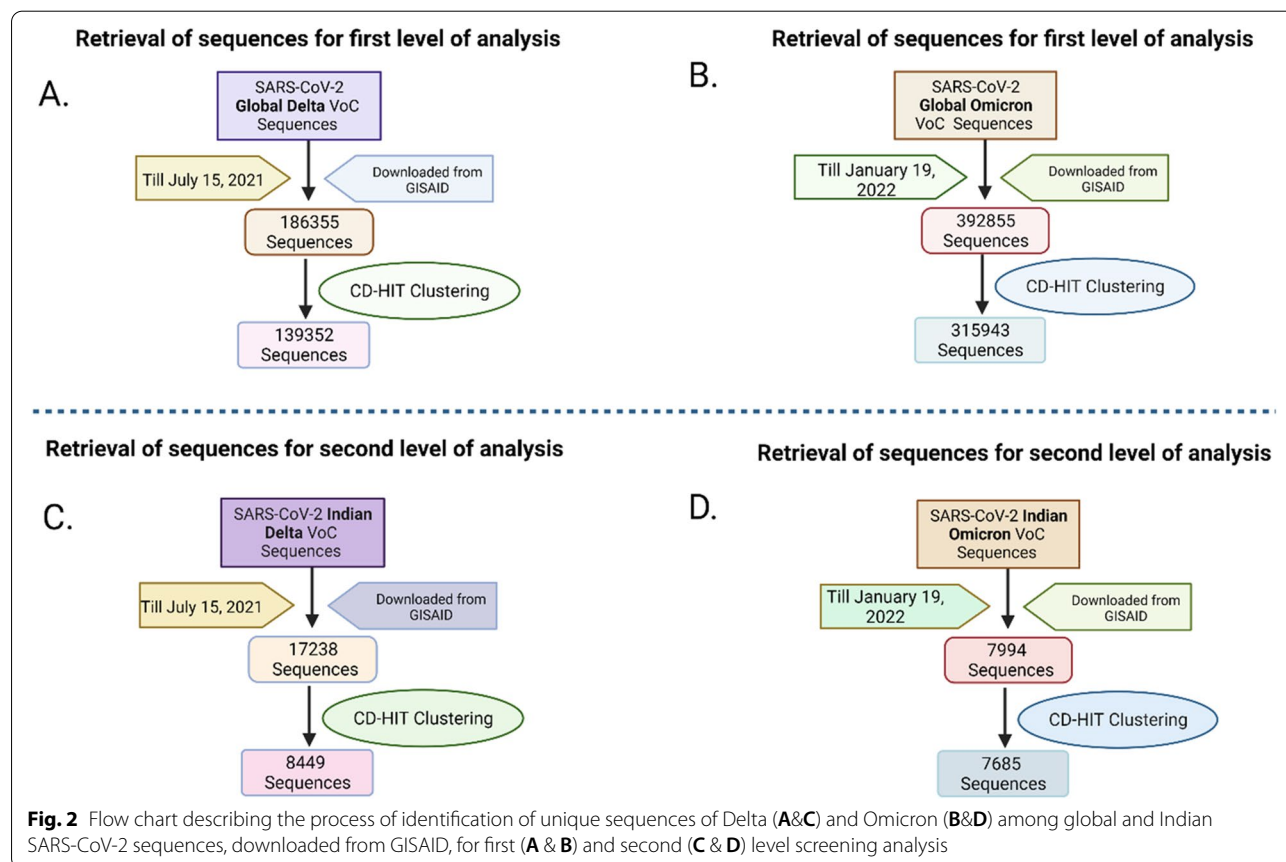


Table 1 A describes the *in silico* qRT-PCR kit validation with SARS-CoV-2 Delta variant and B describes the *in silico* qRT-PCR kit validation with SARS-CoV-2 Omicron variant

A

***In silico* evaluation of qRT-PCR Kits with SARS-CoV-2 Delta Variant**

Type of Kit	Alignment results with global Delta sequences (First level screening analysis)			Alignment results with Indian Delta sequences (Second level screening analysis)			Overall results ^a
	Total number of kits	No. satisfactory	% satisfactory	No. of kits	No. satisfactory	% satisfactory	Overall satisfactory %
One gene	3	2	67	0	NA	NA	67
Two gene	53	39	74	1	1	100	75
Three gene	16	13	81	2	0	0	81
Four gene	1	1	100	0	NA	NA	100
Total	73	55	75	3	1	33	77

B

***In silico* evaluation of qRT-PCR Kits with SARS-CoV-2 Omicron Variant**

Type of Kit	Alignment results with global Omicron sequences (First level screening analysis)			Alignment results with Indian Omicron sequences (Second level screening analysis)			Overall results ^a
	Total number of kits	No. satisfactory	% satisfactory	No. of kits	No. satisfactory	% satisfactory	Overall satisfactory %
One gene	3	1	33	2	1	50	67
Two gene	53	34	64	16	12	75	87
Three gene	16	13	81	2	2	100	94
Four gene	1	1	100	0	NA	NA	100
Total	73	49	67	20	15	75	88

^a The overall results were calculated by adding the number of satisfactory kits in first and second level screening then divided by total number of kits in each category and expressed as percentage

Criteria for classifying qRT-PCR kits as unsatisfactory

With Delta VoC sequences, 18 out of 73 (25%) kits were found to be unsatisfactory after alignment with global Delta sequences, with majority of them found unsatisfactory on the basis of a single common criteria of missing alignment of primers and probes at 3' or 5' end i.e. criteria iii of first level screening analysis (Fig. 3). Three kits were subjected to second level screening analysis with Indian Delta sequences as they were found unsatisfactory as per criteria (i) and (ii) of the first level screening analysis. Two out of three (67%) kits were found to be unsatisfactory after second level screening. In majority of cases, RdRp, N and ORF gene primers and S, ORF and RdRp gene probes did not align with Delta VoC as per the two step screening criteria (Supplementary data file).

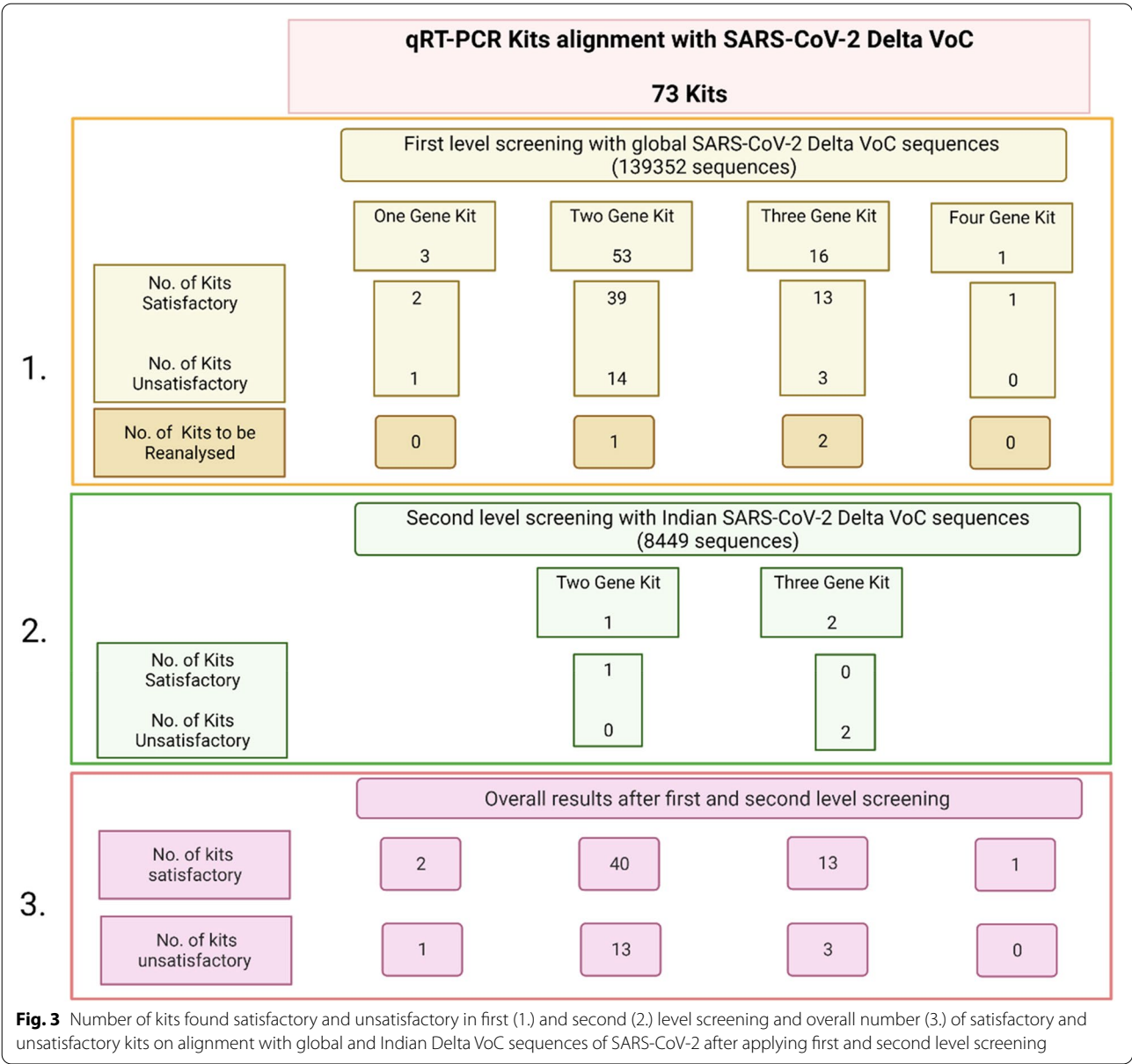
For Omicron VoC, 24 out of 73 (33%) kits were found to be unsatisfactory in the first level screening, of which 4 kits were rejected as there was no alignment of primers and probes at 3' or 5' end. Subsequently, 20 kits underwent second level screening analysis with Indian Omicron sequences as they were found unsatisfactory as per criteria (i) and (ii) of the first level screening analysis. A total 5 out of 20 (25%) kits were found to be unsatisfactory after second level screening. Out of five, three kits were rejected as they failed to meet criteria i and ii and two kits could not meet criteria iii of second

level screening analysis (Fig. 4). In majority of cases, S, N and E gene primer and ORF, S, N and RdRp gene probe did not align with Omicron VoC as per the two step screening criteria (Supplementary data file).

Overall, 17 out of 73 (23%) kits were found to be unsatisfactory on alignment with Delta sequences of SARS-CoV-2 and 9 out of 73 (12%) kits were found unsatisfactory on alignment with Omicron sequences of SARS-CoV-2. Out of 9 kits found unsatisfied to detect Omicron, seven kits were found unsatisfactory to detect Delta variant also.

Discussion

With time, SARS-CoV-2 has evolved into variants of concern which have enabled the virus to escape the host immune response, enhance its transmissibility and evade diagnosis by standard methods. One of the primary reasons for diagnostic evasion is the occurrence of mutations in the target region leading to non-hybridization of primers/probes. The pathognomic feature of the B.1.1.7 or alpha VoC, identified first in UK, was S gene target drop-out in multiplex assays [9, 30]. S gene failure was not observed much with the Delta VoC [31], however with emergence of Omicron VoC, S gene target failure was again consistently seen with BA.1.1.529 [32]. In addition, N gene failure was also reported with the Omicron

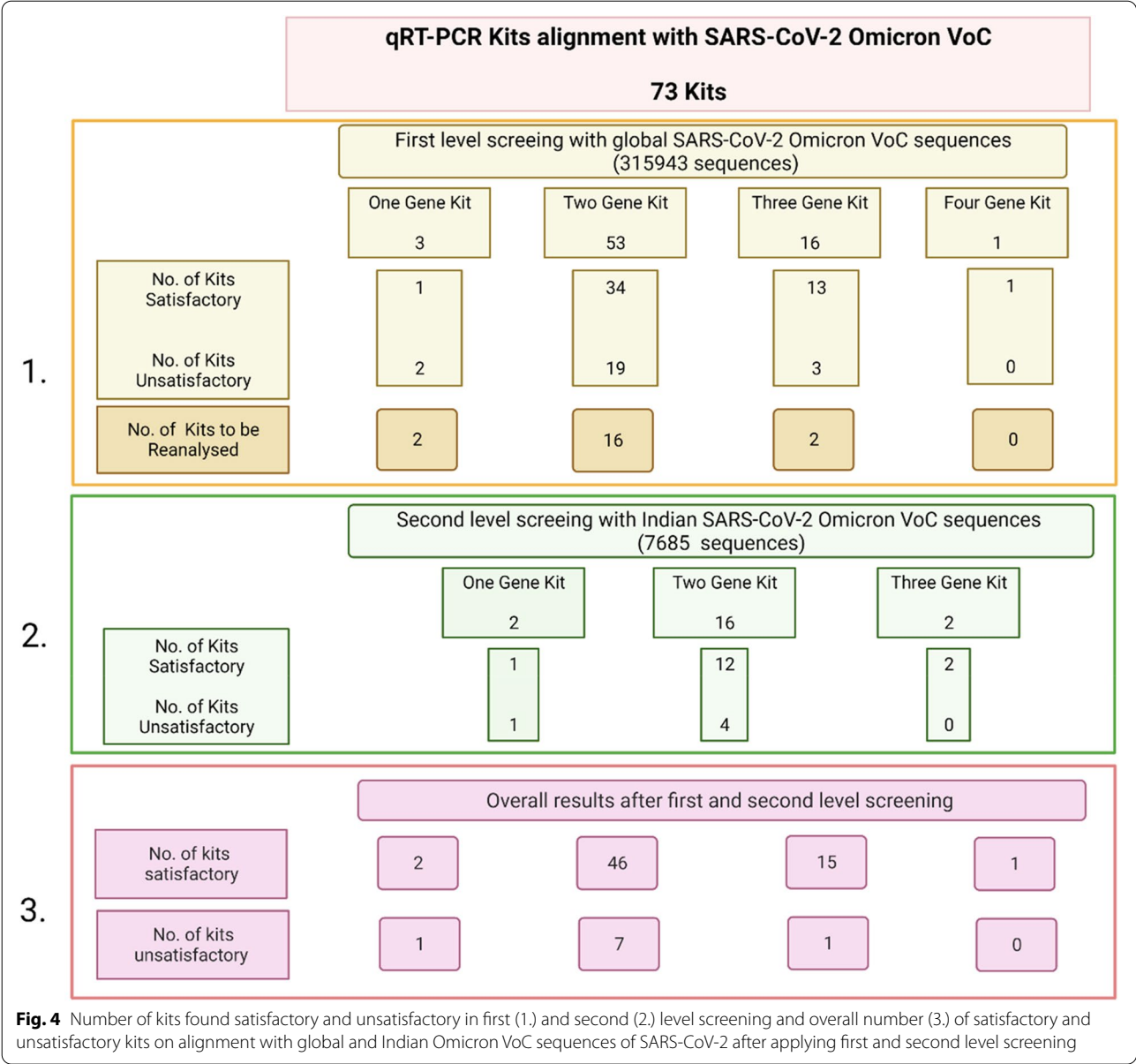


VoC [33]. However, again with the emergence of sublineages of Omicron VoC, the S and N gene target failures became inconsistent and were largely dropped as screening methods for these VoCs [33].

Since inception of the pandemic, early detection, contact tracing and containment of infection were globally identified as key strategies for preventing transmission and controlling spread of the disease. However, over the past 1 year, the focus has shifted towards isolation of infected patients for preventing disease spread to vulnerable individuals and eventually reducing hospitalizations and deaths. In both situations, the role of reliable diagnostics is pivotal. Early detection of infected individuals

and their isolation can only be achieved if the diagnostic tests are accurate and have high sensitivity and specificity for detecting the variant strains. Since majority of tests for detection of SARS-CoV-2 are based on target genes of conventional strains of the virus, their performance has been questioned in the backdrop of emergence of SARS-CoV-2 VoCs.

Early in the pandemic, single-plex kits were recommended by WHO for use in regions with established transmission of SARS-CoV-2 [34]. Singleplex kits are known to have a better sensitivity as compared to multiplex assays [35], with the trade-off of missing diagnosis on emergence of variant strains of SARS-CoV-2. While



single S gene target based kits are likely to miss diagnosis in a significant number of clinical samples, dual-, triple- or multiplex kits with two or more primers/probes directed against the target genes of the virus, reduce chances of RTPCR failure and improve accuracy [36, 37]. Therefore, the use of singleplex assay was not preferred later in the pandemic, as the public health strategy has always necessitated the use of sensitive assays.

With the evolution of SARS-CoV-2 virus it is critical to continuously monitor the performance of existing molecular assays with time. Assessing the performance of each diagnostic assay in field is labour-intensive and challenging. Thus, it is desirable to develop simple *in silico*

process to reliably estimate the performance of different molecular assays in field. *In silico* analysis of the binding efficiency of primers and probes with the VoC sequences of SARS-CoV-2, submitted to available public databases like GISAID is a feasible method to reliably estimate the performance of various assays.

In the present study we have developed a two-step *in silico* screening process for estimating the performance of several qRT-PCR kits used in India. *In silico* evaluation of qRT-PCR kits has been undertaken in other studies also [22, 23]. However, none of the studies had proposed a standard process considering all the parameters for assessing the impact of circulating or new VoC on

performance of molecular assays. In addition, a strategy to address the challenge of locating the centre of probes with odd and even number of nucleotides was also developed which is one of its own kind. Further, the primer or probes having degenerate nucleotides satisfying all the screening criteria with even a single nucleotide combination were given preference. Moreover, the data on mismatch location of a specific primer or probe provide the guidance to the kit manufacturers for redesign at particular location. Also the data of concerning/not aligning primers and probes guide the country to not consider the kits for diagnosis of VoC having sequence similar to the concerning ones.

Overall, our study has several strengths. Besides evaluating the performance of each primer and probe, our two-step screening process has also taken into consideration the sub-optimal performance of kits not satisfying the stringent 95% screening criteria i.e. 95% alignment with the number of downloaded sequences and 95% lengthwise alignment with downloaded sequences, mentioned as criteria (i) and (ii). Alignment of each primer and probe used in India, to the large number of global sequences may also be misleading at times due to difference in SARS-CoV-2 strains circulating globally. Therefore, alignment with lesser number of country specific sequences (with less number of 95% downloaded sequences) gave opportunity to the kit manufacturers by re-evaluating their kits based on the circulating VoC in the given country.

Also, this two-step screening process can help the countries in balancing the risk of epidemiologically missing the cases for a particular variant not pre-dominantly circulating over the pre-dominantly circulating variant of SARS-CoV-2. This improved the stringency of screening.

Our study has a few limitations also. Though the screening process developed by us in the present study has considered various possibilities for evaluation of qRT-PCR kits for detection of SARS-CoV-2, but *in silico* approach needs to be further validated by real word performance. Similarly, multiple softwares are available for clustering and alignment of primers/probes with the downloaded sequences. We could not compare these softwares and provide specific recommendations on the most suitable ones to be used. Also, we could only retrieve primer/probe sequences from 73 kit manufacturers out of 200 kit manufacturers contacted. Thus we were not able to evaluate the overall performance of all qRT-PCR kits available in Indian market.

Conclusion

Overall, we have developed a useful process for screening the performance of qRT-PCR assays for detection of SARS-CoV-2. This process has been standardized against

Delta and Omicron VoCs, however it has the potential of use for other variants of SARS-CoV-2 that may emerge in future. The two-step screening process developed by us also has the potential to be used for other evolving pathogens of public health importance.

Abbreviations

SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; NAAT: Nucleic acid amplification test; E: Envelop; ICMR: Indian Council of Medical Research; N: Nucleocapsid; S : Spike; M: Membrane; ORFs: Open Reading Frames; RdRP : RNA dependent RNA polymerase; VOCs: Variant of Concerns; VOIs: Variants of Interest; qRT-PCR: Real-Time Reverse Transcription Polymerase Chain Reaction; SGTF: S gene target failure; WGS: Whole Genome Sequence.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-022-08999-3>.

Additional file 1.

Acknowledgements

Authors would like to thank Dr. Harpreet Singh, Head and Scientist F, Biomedical Informatics Division, Indian Council of Medical Research for his support for the execution of this study.

Authors' contributions

SG: Data retrieval, Data analysis, Writing - Original Draft, Validation. AK: Software development, Data analysis, Writing - Original Draft. DB: Data analysis. NG: Conceptualization, Visualization, Writing - Review & Editing, Data Validation. NA: Conceptualization, Writing - Review & Editing, Supervision. VR: Conceptualization, Writing - Review & Editing, Supervision. All authors have read and approved the final draft of the manuscript.

Funding

The study does not involve any funding support.

Availability of data and materials

The unidentified raw data supporting the conclusions of this article will be made available by the corresponding author without undue reservation.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors have no competing interests to declare.

Author details

¹Division of Epidemiology and Communicable Disease, V. Ramalingaswami Bhawan, Indian Council of Medical Research, Ansari Nagar, New Delhi 110029, India. ²Biomedical Informatics Division, Indian Council of Medical Research, Ansari Nagar, New Delhi, India. ³Department of Forensic Medicine & Toxicology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi, India. ⁴Department of Neurosciences, National Institute of Mental Health and Neuro-Sciences, Bengaluru 560029, India.

Received: 24 June 2022 Accepted: 10 November 2022

Published online: 17 November 2022

References

- Scrima M, Cossu AM, D'Andrea EL, Bocchetti M, Abruzzese Y, Iannarone C, et al. Genomic characterization of the emerging SARS-CoV-2 lineage in two districts of Campania (Italy) using next-generation sequencing. *Front Virol*. 2022;0:4.
- Tan W, Zhao X, Ma X, Wang W, Niu P, Xu W, et al. A novel coronavirus genome identified in a cluster of pneumonia cases — Wuhan, China 2019–2020. *China CDC Weekly*. 2020;2(4):61–2 Available from: <https://weekly.chinacdc.cn/en/article/doi/10.46234/ccdcw2020.017> [cited 16 Mar 2022].
- Kevadiya BD, Machhi J, Herskovitz J, Oleynikov MD, Blomberg WR, Bajwa N, et al. Diagnostics for SARS-CoV-2 infections. *Nat Mater*. 2021;20(5):593–605 Available from: <https://www.nature.com/articles/s41563-020-00906-z> [cited 22 Mar 2022].
- Barreto HG, de Pádua Milagres FA, de Araújo GC, Daúde MM, Benedito VA. Diagnosing the novel SARS-CoV-2 by quantitative RT-PCR: variations and opportunities. *J Mol Med (Berl)*. 2020;98(12):1727 Available from: <https://pmc/articles/PMC7567654/> [cited 18 May 2022].
- WHO. Tracking SARS-CoV-2 variants. Available from: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>. [cited 23 Mar 2022].
- Mardian Y, Kosasih H, Karyana M, Neal A, Lau CY. Review of current COVID-19 diagnostics and opportunities for further development. *Front Med*. 2021;8:562.
- Li A, Maier A, Carter M, Guan TH. Omicron and S-gene target failure cases in the highest COVID-19 case rate region in Canada—December 2021. *J Med Virol*. 2022;94(5):1784–6 Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.27562> [cited 23 Mar 2022].
- Rosato AE, Msiha E, Weng B, Mesisca M, Gnass R, Gnass S, et al. Rapid detection of the widely circulating B.1.617.2 (Delta) SARS-CoV-2 variant. *Pathology*. 2022;54:351–6 Available from: <https://pubmed.ncbi.nlm.nih.gov/35221043/> [cited 23 Mar 2022].
- Walker AS, Vihta K-D, Gethings O, Pritchard E, Jones J, House T, et al. Tracking the Emergence of SARS-CoV-2 Alpha Variant in the United Kingdom. *N Engl J Med*. 2021;385(27):2582–5 Available from: <https://pubmed.ncbi.nlm.nih.gov/34879193/> [cited 25 Mar 2022].
- Wolter N, Jassat W, Walaza S, Welch R, Moultrie H, Groome M, et al. Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. *Lancet*. 2022;399(10323):437–46.
- Laine P, Nihtilä H, Mustanoja E, Lyyski A, Ylinen A, Hurme J, et al. SARS-CoV-2 variant with mutations in N gene affecting detection by widely used PCR primers. *J Med Virol*. 2022;94(3):1227–31. <https://doi.org/10.1002/jmv.27418>. Epub 2021 Nov 2.
- Lesbon JCC, Poletti MD, de Mattos Oliveira EC, Patané JSL, Clemente LG, Viala VL, et al. Nucleocapsid (N) gene mutations of sars-cov-2 can affect real-time rt-pcr diagnostic and impact false-negative results. *Viruses*. 2021;13(12):2474 Available from: <https://pmc/articles/PMC8707239/> [cited 23 Mar 2022].
- Khateeb J, Li Y, Zhang H. Emerging SARS-CoV-2 variants of concern and potential intervention approaches. *Crit Care*. 2021;25(1):1–8 Available from: <https://ccforum.biomedcentral.com/articles/10.1186/s13054-021-03662-x> [cited 13 May 2022].
- Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance*. 2017;22(13):2–4.
- Elbe S, Buckland-Merrett G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Challenges (Hoboken, NJ)*. 2017;1(1):33–46 Available from: <https://pubmed.ncbi.nlm.nih.gov/31565258/> [cited 20 Apr 2022].
- Khare S, Gurry C, Freitas L, Schultz MB, Bach G, Diallo A, et al. GISAID's Role in Pandemic Response. *China CDC Wkly*. 2021;3(49):1049–51 Available from: <https://pubmed.ncbi.nlm.nih.gov/34934514/> [cited 21 Apr 2022].
- Li W, Jaroszewski L, Godzik A. Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics*. 2001;17(3):282–3 Available from: <https://pubmed.ncbi.nlm.nih.gov/11294794/> [cited 21 Apr 2022].
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10 Available from: <https://pubmed.ncbi.nlm.nih.gov/2231712/> [cited 21 Apr 2022].
- Rice P, Longden L, Bleasby A. EMBL: the European molecular biology open software suite. *Trends Genet*. 2000;16(6):276–7 Available from: <https://pubmed.ncbi.nlm.nih.gov/10827456/> [cited 25 Oct 2022].
- Smith TF, Waterman MS. Identification of common molecular subsequences. *J Mol Biol*. 1981;147(1):195–7 Available from: <https://pubmed.ncbi.nlm.nih.gov/7265238/> [cited 25 Oct 2022].
- Khan KA, Cheung P. Presence of mismatches between diagnostic PCR assays and coronavirus SARS-CoV-2 genome: sequence mismatches in SARS-CoV-2 PCR. *R Soc Open Sci*. 2020;7(6):200636.
- Gand M, Vanneste K, Thomas I, Van Gucht S, Capron A, Herman P, et al. Deepening of in silico evaluation of sars-cov-2 detection rt-qpcr assays in the context of new variants. *Genes (Basel)*. 2021;12(4):565 Available from: <https://pmc/articles/PMC8069896/> [cited 16 Jan 2022].
- Sharma D, Martinez-sobrido L, Gromiha MM, Henry BM. In silico evaluation of the impact of the Omicron variant on the sensitivity of RT-qPCR assays for SARS-CoV-2 detection using whole genome sequencing. *J Med Virol*. 2022:e28241. <https://doi.org/10.1002/jmv.28241>. Online ahead of print.
- Chow CK, Qin K, Lau LT, Yu ACH. Significance of a single-nucleotide primer mismatch in hepatitis B virus real-time PCR diagnostic assays. *J Clin Microbiol*. 2011;49(12):4418–9 Available from: <https://journals.asm.org/doi/full/10.1128/JCM.05224-11> [cited 19 May 2022].
- da Santos RS, Bret RSC, de OM MAC, Campos AR, Silva ARAE, Lima DM, et al. In silico analysis of mismatches in RT-qPCR assays of 177 SARS-CoV-2 sequences from Brazil. *Rev Soc Bras Med Trop*. 2020;53:1–5 Available from: <http://www.scielo.br/j/rsbmt/a/x6XRX3HLDrGrFFcplCh7Lm/?lang=en> [cited 16 Jan 2022].
- Mushtaq MZ, Shakoor S, Kanji A, Shaheen N, Nasir A, Ansar Z, et al. Discrepancy between PCR based SARS-CoV-2 tests suggests the need to re-evaluate diagnostic assays. *BMC Res Notes*. 2021;14(1):1–5.
- Miranda P, Weber G. Thermodynamic evaluation of the impact of DNA mismatches in PCR-type SARS-CoV-2 primers and probes. *Mol Cell Probes*. 2021;56:101707 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/33609730>.
- Ghoshal U, Vasanth S, Tejan N. A guide to laboratory diagnosis of Corona Virus Disease-19 for the gastroenterologists. *Indian J Gastroenterol*. 2020;39(3):1 Available from: <https://pmc/articles/PMC7462729/> [cited 14 May 2022].
- Basawarajappa SG, Rangaiah A, Padukone S, Yadav PD. Performance evaluation of Truenat™ Beta CoV & Truenat™ SARS-CoV-2 point-of-care assays for coronavirus disease 2019; 2020.
- WHO. Guidance on tools for detection and surveillance of SARS CoV-2 Variants Interim Guidance. 2021. Available from: <https://www.afro.who.int/sites/default/files/Covid-19/Technicaldocuments/GuidanceontoolsfordetectionandsurveillanceofSARS-CoV-2Variants-EN.pdf> [cited 29 May 2022].
- Alhudiri IM, Ramadan AM, Ibrahim KM, Abdalla A, Eljilani M, Salem MA, et al. Whole-genome sequencing of SARS-CoV-2 showed wide spread of B.1.525 in February 2021 in Libya. *Libyan J Med*. 2021;16(1):2001210 Available from: <https://pmc/articles/PMC8583741/> [cited 29 May 2022].
- WHO. Classification of Omicron (B.1.1.529): SARS-CoV-2 Variant of Concern. 2021. Available from: [https://www.who.int/news/item/26-11-2021-classification-of-omicron-\(b.1.1.529\)-sars-cov-2-variant-of-concern](https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern) [cited 29 May 2022].
- Harankhedkar S, Chatterjee G, Rajpal S, Khan A, Srivastava T, Mirgh S, et al. N Gene Target Failure (NGTF) for detection of Omicron: a way out for the 'stealth' too? *medRxiv*. 2022:2022.01.28.22269801 Available from: <https://www.medrxiv.org/content/10.1101/2022.01.28.22269801v2> [cited 29 May 2022].
- World Health Organization. Maintaining surveillance of influenza and monitoring SARS-CoV-2: GISRS. WHO - Interim Guid; 2020. <https://apps.who.int/iris/handle/10665/336689>.
- Hu X, Zhu L, Luo Y, Zhao Q, Tan C, Chen X, et al. Evaluation of the clinical performance of single-, dual-, and triple-target SARS-CoV-2 RT-qPCR methods. *Clin Chim Acta*. 2020;511:143 Available from: <https://pmc/articles/PMC7550050/> [cited 19 May 2022].
- Brown KA, Gubbay J, Hopkins J, Patel S, Buchan SA, Daneman N, et al. S-Gene Target Failure as a Marker of Variant B.1.1.7 Among SARS-CoV-2 Isolates in the Greater Toronto Area, December 2020 to March 2021. *JAMA*. 2021;325(20):2115 Available from: <https://pmc/articles/PMC8033504/> [cited 14 May 2022].
- Banko A, Petrovic G, Miljanovic D, Loncar A, Vukcevic M, Despot D, et al. Comparison and sensitivity evaluation of three different commercial real-time quantitative PCR kits for SARS-CoV-2 detection. *Viruses*. 2021;13(7):1321. <https://doi.org/10.3390/v13071321>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.