

Research Article

Parisa Sharafi, F. Seyma Gokdemir, Mesut Akyol, Yasemin Ardicoglu-Akisin, J. Sedef Gocmen and Aysegul Taylan-Ozkan*



Comparison of efficacy and reliability of six commercial COVID-19 diagnostic PCR kits

<https://doi.org/10.1515/tjb-2023-0075>

Received April 7, 2023; accepted January 4, 2024;

published online February 29, 2024

Abstract

Objectives: All countries have been deeply affected by the coronavirus disease 2019 pandemic, both economically and in situations that strain health systems, such as workforce and workload. Therefore, various measures should be taken to control the disease and prevent its spread. Since the disease onset, real-time PCR tests have been used as the gold standard for disease diagnosis. Owing to the rapid progress of the pandemic and the spread of the disease, validation, consistency, and optimization tests of some commercial kits have been conducted directly in the field. Therefore, it is important to compare the results of these kits and improve the existing ones.

Methods: We compared five kits (Bioexen, Polgen, Coronex, Diagen, and Anatolia) donated to the TOBB Economics and Technology University Hospital PCR laboratory with the KrosGen kit to detect severe acute respiratory syndrome coronavirus 2. A total of 244 samples were selected and analyzed using five different severe acute respiratory syndrome coronavirus 2 PCR detection kits.

Results: Positive and negative results from the six kits were compared using the working protocols of the kits, primers, and cycle threshold (Ct) values. Five of the six kits have reliable compatibility for Ct<30 but decreases for Ct≥30. Therefore, it is important to evaluate the performance of these kits for reduced viral loads.

Conclusions: Using a suitable kit with high compatibility for Ct≥30 is important for detecting patients with a low viral load and helping prevent disease spread.

Keywords: pandemics; COVID-19; SARS-CoV-2 detection kits; RT-PCR; Ct values

***Corresponding author: Aysegul Taylan-Ozkan**, TOBB University of Economics and Technology, Faculty of Medicine, Department of Medical Microbiology, Ankara, Türkiye, Phone: +905053177676, E-mail: aysegultaylanozkan@etu.edu.tr. <https://orcid.org/0000-0001-8421-3625>

Parisa Sharafi, TOBB University of Economics and Technology, Faculty of Medicine, Department of Medical Biology and Genetics, Ankara, Türkiye, E-mail: psharafi@etu.edu.tr. <https://orcid.org/0000-0002-7400-4851>

F. Seyma Gokdemir, Başkent University, Institute of Food, Agriculture and Livestock Development, Ankara, Türkiye; and Başkent University, Kahramankazan Vocational School, Department of Food Processing, Ankara, Türkiye, E-mail: fsgokdemir@gmail.com. <https://orcid.org/0000-0003-2951-848X>

Mesut Akyol, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Ankara, Türkiye, E-mail: makyol@aybu.edu.tr. <https://orcid.org/0000-0003-1449-7874>

Yasemin Ardicoglu-Akisin, TOBB University of Economics and Technology, Faculty of Medicine, Department of Medical Biochemistry, Ankara, Türkiye, E-mail: yardicoglu@tobbetuhastanesi.com.tr. <https://orcid.org/0000-0002-4109-0220>

J. Sedef Gocmen, TOBB University of Economics and Technology, Faculty of Medicine, Department of Medical Microbiology, Ankara, Türkiye, E-mail: jgocmen@etu.edu.tr. <https://orcid.org/0000-0001-8207-8749>

Introduction

The coronavirus disease 2019 (COVID-19) pandemic, which emerged in Wuhan, China, in 2019, was the largest pandemic of the last century regarding its spread and the number of people infected. All countries have been severely affected by the pandemic, which has strained their health systems and impacted the economy and workforce. Therefore, various measurement methods have been implemented to control this disease and prevent its spread. One of these measures is early diagnosis of COVID-19. Probe-based real-time polymerase chain reaction is the gold standard for disease diagnosis [1]. Reverse transcriptase-polymerase chain reaction (RT-PCR) is preferred because of its advantages in early-stage disease detection, high sensitivity, high specificity, and ease of use. Since the pandemic outbreak, more than 250 known COVID-19 PCR kits have been developed. All these kits use cycle threshold (Ct) values to determine the patient's viral load, which provides information for detecting severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The

Ct value is a quantitative PCR measure used to calculate the initial concentration or copy number of a target nucleic acid sequence in a sample. The Ct value represents the relative amount compared with the standard reference sample. By comparing the Ct values of different samples, users can evaluate the copy numbers of the samples. The Ct value was calculated starting with a known amount of the target sequence and then monitoring the increase in fluorescence (called the relative fluorescence unit (RFU)) as the increased number of amplifications. The RFU is an important measurement used to assess the intensity of fluorescence produced when a sample is amplified and provides a valid relative reference point to determine the concentration and abundance of target sequences. The Ct value is the point at which the fluorescence increases above the background level. In general, the lower the Ct value, the higher the abundance of the target sequence in the sample and, therefore, the faster the amplification of the target sequence. Some studies have claimed that SARS-CoV-2 has no actual genome and that RT-PCR testing can be misleading [2]. Although these claims are unfounded, most newly produced SARS-CoV-2 diagnostic kits show inconsistent performance, contradicting lung tomography findings. Patients with severe infection without pulmonary tomography findings may have negative PCR results, increasing the possibility that COVID-19-positive cases are causing viral spread. The inconsistent performance of COVID-19 RT-PCR kits may be due to the pandemic's rapid evolution and the disease's spread. Therefore, some kits' validation, consistency, and optimization tests were performed directly in the field. These events require a comprehensive comparison of the results between kits to select and improve the best one [3].

SARS-CoV-2 is a positive-stranded RNA (+ssRNA) virus with a single open reading frame called *ORF1ab* that encodes non-structural proteins involved in viral RNA transcription, replication, and immune evasion. One of these non-structural proteins is RNA-dependent RNA polymerase. Replication and transcription also occur within the *ORF1ab* framework. The structural proteins of SARS-CoV-2, including the membrane (M), envelope (E), nucleocapsid (N), and spike (S) proteins, are expressed through the production of sub-genomic messenger RNAs. *ORF1ab*, *E*, *N*, and *S* genes are the most commonly used primer targets for detecting SARS-CoV-2 [4, 5]. The *E* gene target region is used to detect SARS-CoV-2 and genetically similar viruses (pan-Sarbecovirus) for general screening, whereas the *ORF1ab* target region and *N* gene specifically identify SARS-CoV-2. Some kits use more than one primer to detect multiple SARS-CoV-2 genes simultaneously. Housekeeping genes, such as *RNase P* (*ribonuclease P*) and *GAPDH* (*glyceraldehyde 3-phosphate dehydrogenase*), which are highly expressed in human epithelial tissues, were used as

internal controls. In addition, unlike classical PCR methods, current SARS-CoV-2 detection kits use probe-based primers. The number of nucleic acids copied during the reaction was measured in real-time using fluorescence irradiation in each cycle. Probe-based RT-PCR can simultaneously detect multiple targets in a sample. However, this requires the design of target-specific probes in addition to primers. Several probe designs are available; however, hydrolysis probes are the most commonly used. The hydrolysis probe comprised two primers and one probe. The probes were designed to complement the target sequences. Each probe contained a fluorophore (donor) and a quencher (acceptor) at each end. During the annealing step of PCR, the probe binds to a specific target sequence. However, no fluorescent signals are detected because the fluorophores and quencher are close. In the next step, as an extension, the probe is hydrolyzed by the 5–3' exonuclease activity of the polymerase. Probe hydrolysis separates the fluorophore from the quencher, increasing amplification-dependent fluorescence. Thus, the fluorescent signal from RT-PCR was proportional to the amount of the probe-target sequence present in the sample. Probe-based RT-PCR is more specific than dye-based RT-PCR and is widely used for diagnosis [6]. Currently, kits used to detect SARS-CoV-2 in many countries are probe-based. Depending on the developer's and financial infrastructure's initiatives, more than one probe may be used to detect viral genes. Specific primers and labeled probes were used to detect the virus's conserved *ORF1ab* and *N* genes, and *RNase I* was used as the internal control.

In this study, we compared six commercial kits using 244 patients. Ct values and patient viral loads were compared. We also compared the compatibility of the kits based on the accuracy of the strongly positive (Ct<30), weakly positive (Ct<30), and negative results.

Materials and methods

Patient population and sample collection

Nasopharyngeal aspirate/lavage/swab specimens from 244 individuals (62 negative and 182 positive) who underwent PCR-COVID-19 testing at the TOBB University of Economy and Technology Hospital, Ankara, between June and December 2022, were used. All patients with positive PCR results were confirmed to have COVID-19 symptoms, such as fever, cough, fatigue, loss of taste or smell, and sore throat. All patients were selected in a double-blind manner according to their PCR results (the PCR results were confirmed by a specialist who examined the patients for symptoms). Patients with positive results were grouped according to the Ct values. Patients with a Ct<30 or strong positivity carry a high SARS-CoV-2 load. Patients with Ct≥30 or low positivity carry a low SARS-CoV-2 virus load. Patients with negative PCR results were defined

as those without symptoms and no contact with PCR-positive patients. Information on age, sex, religion, etc., is unavailable to the researchers. Ethical approval [Approval number. KAEK-118/143 (dated 11.11.2022)] was obtained from the Ethics Committee of the Faculty of Medicine.

Power and sample size

The sample size of the study was calculated using PASS 15 Power Analysis and Sample Size Software (2017, NCSS, LLC. Kaysville, Utah, USA, ncss.com/software/pass). The required sample size for 95 % power, $\alpha=0.05$ type I error, $\beta=0.05$ type II error, 20 % expected prevalence of COVID-19, 50 % expected sensitivity, and 50 % expected specificity was calculated as a minimum of 145. Our research team only received the Anatolia kit during the final stages of the study, which limited our ability to conduct a comprehensive study with a sufficiently robust sample size.

Molecular and real-time PCR analysis of SARS-CoV-2

Samples for the reactions were collected using a viral nucleic acid transfer buffer (Extracted Sample Transport Medium, eSTM-02, MAYGEN laboratory services, Turkey), which allowed viral nucleic acid extraction from nasopharyngeal aspirate/lavage/swabs without the use of an RNA extraction kit. Clinical viral nucleic acid transfer samples were used as RNA samples for these reactions.

The KrosQuanT SARS-CoV-2 (2019-nCoV) V2 RT-PCR kit (KRM-136-002; KrosGen Biotechnology, Turkey) was routinely used to detect COVID-19 in the nasopharyngeal aspirate/lavage/swab from patients in our laboratory. We used KrosGen as the reference kit in this study and accepted the results obtained with this kit as accurate and reliable. The same samples were tested using five different SARS-CoV PCR detection kits named: Bio-Speedy Direct RT-PCR SARS-CoV-2 (BS-SY-SC2-500; Bioexen research and development (R&D) Technologies, Turkey), Senteligo SARS-CoV-2 (COVID-19) Multiplex RT-PCR Kit, (Coronex (Ver:2. 0), SentebioLab, Turkey), SARS-CoV-2 (COVID-19) Multiplex Detection Kit (COVID-PPB, Polgen Biotechnology, Turkey), SARS-CoV-2 OneStep RT-PCR Kit (Dia-CoV19-2, Diagen Biyoteknolojik Sistemler, Turkey), and Bosphore Novel Coronavirus (2019-nCoV) Detection Kit v4 (Anatolia Geneworks, MB416v2, Global Technology, Turkey). All kits were used and evaluated according to the manufacturers' recommendations. All information regarding the kits, including the user manual, reaction protocol, thermal program, primers, and probes, is provided in Table 1. The RT-PCR and Ct values of the positive results obtained using the different kits are shown in Table 2.

Statistical analysis

The results obtained using the kits were transferred to an Excel spreadsheet. The Ct values were recorded as <30 and ≥ 30 . Error checks and corrections were then performed. Crosstabs were created to compare the reference kit (KrosGen) with the kits used in the present study. Notations n and % are used in the tables. The characteristics of the kits, such as sensitivity, specificity, positive predicted value, and negative predicted value, were calculated based on the results for the reference kit. The Youden index was used to compare the accuracy and compatibility of these detection kits in discriminating between true-positive and true-negative COVID-19 PCR tests (Table 3). The values of min-max, mean $\pm$ SD, and median (IQR: InterQuartile Range) were used to present Ct values (Table 4). IBM SPSS Statistics 22.0 (IBM Corp.

Released 2013 IBM SPSS Statistics for Windows, version 22.0. Armonk, NY: IBM Corp.) was used for statistical analyses. Data were analyzed at a 95 % confidence level, and significance was set at $p<0.05$.

Results

In this study, nasopharyngeal aspirate/lavage/swab specimens from 244 individuals collected at the PCR-COVID-19 laboratory of TOBB University of Economy and Technology Hospital, Ankara, were analyzed using six different SARS-CoV-2 detection kits. The KrosGen kit is routinely used to detect SARS-CoV-2 infection in patients. All patients with positive PCR results were followed up and examined by a specialist for associated symptoms. The stage of SARS-CoV-2 (whether the patients were in the early, acute, or late phases of the disease) was also monitored. Details of the primers, reaction protocols, and thermal programs used for all six commercial kits are listed in Table 1.

Table 2 shows three of the five kits (Bioexen, Coronex, and Anatolia) showed consistent results for $Ct<30$ compared with KrosGen. Conversely, for $Ct\geq 30$, the consistency of the kits fluctuated with positive results. For $Ct\geq 30$, all patients were asked if they had symptoms, and positive results were verified. In addition, all positive samples for $Ct\geq 30$ were diluted by 1:2 to eliminate false positives, and the reaction was repeated. The result was considered positive if the second reaction was also positive; otherwise, a new sample was requested from the patient the following day. Table 2 shows four of the five kits (Bioexen, Polgen, Coronex, and Diagen) had higher false negative results than the reference kit (KrosGen). As stated in the manufacturer's instructions, the evaluation of the positive results of the Polgen kit with the detection of one of the targets showed that the false positive percentage decreased. Only the Coronex kit showed a high false positive rate (87.1 %). The Anatolia kit showed the highest kit concordance for detecting COVID-19 compared with the KrosGen kit.

The number and percentage of positive and negative samples, positive predicted value, negative predictive value, positive likelihood ratio (LR+), negative likelihood ratio (LR-), accuracy rate, and Youden index of the KrosGen kit are shown in Table 3. When comparing the target genes used in the six commercial kits, we found that all the kits used *RNase P* as an internal control. However, each kit uses a different target gene to detect SARS-CoV-2. The KrosGen and Polgen kits use the *N1* and *N2* genes for detection. KrosGen used both in a single channel, and the presence of either one gave a positive result. However, in the Polgen kit, *N1* amplification was measured in 6-carboxyfluorescein (FAM), while *N2* was measured in hexachlorocarboxyfluorescein

Table 1: Summary of manufacturer's instructions for six different COVID-19 detection kits.

Abbreviated name of kits	KrosGen	Bioexen	Polgen	Coronex	Diagen	Anatolia
Targets and fluorescence channels	Gene 1 N1 (FAM) Gene 2 N2 (FAM) IC RNP (HEX)	ORF1ab (FAM) RNP (HEX)	N1 (FAM) N2 (HEX) RNP (Cy5)	ORF1ab (FAM) N (FAM) RNP (HEX)	ORF1ab (HEX) S (FAM) RNP (Texas red)	ORF1ab (FAM) N (FAM) RNP (HEX)
Ct value for positivity	<35	<38	<35 (both should be positive)	<35	<35 (one is enough to be positive)	<35
Reaction protocol	Master mix: 15 µL RNA sample: 5 µL	2× prime script mix: 10 µL Oligo mix (ORF1ab/RNP): 5 µL RNA sample: 5 µL	Master mix: 25 µL RNA sample: 5 µL	Coronex Covid-19-DS mix E: 12.5 µL Coronex Covid-19-DS PPI: 2.5 µL RNA sample: 5 µL	RT PCR 2× master mix: 10 µL S/ORF1ab/HIC mix: 5 µL RNA sample: 5 µL	PCR 2× master mix: 19.75 µL RT mix: 0.25 µL RNA sample: 5 µL
Thermal program	45 °C for 9 min 95 °C for 2 min 40 cycle for: 95 °C for 3 s 55 °C for 10 s	52 °C for 5 min 95 °C for 10 s 40 cycle for: 95 °C for 1 s 55 °C for 30 s	55 °C for 15 min 95 °C for 30 s 35 cycle for: 95 °C for 10 s 60 °C for 30 s	48 °C for 20 min 95 °C for 2 min 35 cycle for: 95 °C for 5 s 60 °C for 10 s	50 °C for 20 min 95 °C for 5 min 38 cycle for: 95 °C for 10 s 57 °C for 30 s	50 °C for 17 min 95 °C for 6 min 38 cycle for: 97 °C for 30 s 62 °C for 30 s

Table 2: Comparison of Ct values of 5 different SARS-COV-19 detection kits with those of reference kits (KrosGen).

Detection kits		Bioexen FAM (n=244)	Polgen (n=244)				Coronex FAM (n=244)	Diagen (n=244)			Anatolia FAM (n=88)
			Polgen FAM	Polgen HEX	Polgen (both positive)	Polgen (one positive ^a)		Diagen FAM	Diagen HEX	Diagen (one positive)	
Ct<30 (n=122)	Positive (Ct<30)	95 (77.9 %)	22 (18.0 %)	33 (27.1 %)	6 (4.9 %)	70 (57.4 %)	56 (45.9 %)	34 (28.9 %)	34 (27.9 %)	53 (43.4 %)	37 (90.2 %)
	Positive (Ct≥30)	6 (4.9 %)	19 (15.6 %)	19 (15.6 %)	0 (0.0)		15 (12.3 %)	19 (15.6 %)	16 (13.1 %)		4 (9.8 %)
	Negative	21 (17.2 %)	34 (27.9 %)	54 (44.3 %)	116 (95.1 %)	52 (42.6 %)	41 (33.6)	22 (18.0 %)	26 (21.3 %)	23 (18.9 %)	0 (0.0)
	Unavailable to evaluate	0 (0.0)	47 (38.5 %)	16 (13.1 %)	0 (0.0)	0 (0.0)	10 (8.2 %)	47 (38.5 %)	46 (37.7 %)	46 (37.7 %)	0 (0.0)
Ct≥30 (n=60)	Positive (Ct<30)	13 (21.7 %)	0 (0.0)	19 (31.7)	0 (0.0)	28 (46.7 %)	3 (5.0 %)	1 (1.7 %)	1 (1.7 %)	2 (3.3 %)	7 (42.1 %)
	Positive (Ct≥30)	3 (5.0 %)	4 (6.7 %)	9 (7.4)	0 (0.0)		4 (6.7 %)	0 (0.0)	1 (1.7 %)		5 (28.9 %)
	Negative	42 (70.0 %)	18 (30.0 %)	29 (15.0 %)	60 (100.0 %)	32 (53.3 %)	39 (65.0 %)	21 (35.0 %)	19 (31.7 %)	20 (3.3 %)	9 (23.9 %)
	Unavailable to evaluate	2 (3.3 %)	38 (63.3 %)	3 (5.0 %)	0 (0.0)	0 (0.0)	14 (23.3 %)	38 (65.0 %)	39 (65.0 %)	38 (63.3 %)	2 (5.3 %)
Negative (n=62)	Positive (Ct<30)	7 (11.3 %)	0 (0.0)	1 (1.6)	0 (0.0)	1 (1.6 %)	2 (3.2 %)	0 (0.0 %)	0 (0.0)	0 (11.3 %)	1 (11.1 %)
	Positive (Ct≥30)	0 (0.0)	10 (16.1 %)	0 (0.0)	0 (0.0)	10 (16.1)	7 (11.3 %)	4 (6.5 %)	5 (8.1 %)	6 (0.0)	1 (11.1 %)
	Negative	55 (88.7 %)	42 (67.7 %)	59 (95.2 %)	62 (100.0 %)	52 (83.9 %)	52 (83.9 %)	46 (74.2 %)	46 (72.2 %)	45 (70.9 %)	7 (77.8 %)
	Unavailable to evaluate	0 (0.0)	10 (16.1 %)	2 (3.2)	0 (0.0)	0 (0.0)	1 (1.6)	12 (19.4 %)	11 (17.8 %)	11 (17.8 %)	0 (0.0)

^aThis column is not indicated in the user manual of Polgen kit.

Table 3: Comparison of five different COVID-19 detection kits with reference kit (KrosGen).

Detection kits		Standard kit result (Kros-gen [®] , n=244)			Detection kit characteristics							
		Negative (n=62; 25.6 %)	Positive (n=182; 74.4 %)	Total	Sensitivity	Specificity	PPV	NPV	Accuracy rate	Youden index	LR+	LR-
Bioexen [®] – FAM (n=242)	Negative	55 (88.7)	63 (35.0)	118 (48.8)	65.0	88.7	94.4	46.6	71.1	53.7	5.76	0.39
	Positive	7 (11.3)	117 (65.0)	124 (51.2)								
Polgen [®] – FAM (n=223)	Negative	42 (80.8)	51 (53.1)	93 (62.8)	46.9	80.8	81.8	45.2	58.8	27.6	2.44	0.66
	Positive	10 (19.2)	45 (46.9)	55 (37.2)								
Polgen [®] – HEX (n=223)	Negative	60 (100.0)	157 (96.3)	217 (97.3)	3.7	100.0	100.0	27.6	29.6	3.7	N/A	0.96
	Positive	0 (0.0)	6 (3.7)	6 (2.7)								
Polgen [®] – FAM & HEX (n=245)	Negative	63 (100.0)	176 (96.7)	239 (97.6)	3.3	100.0	100.0	26.4	28.2	3.3	N/A	0.97
	Positive	0 (0.0)	6 (3.3)	6 (2.4)								
Polgen [®] – FAM or HEX (n=245)	Negative	11 (17.5)	83 (45.6)	94 (38.4)	54.4	17.5	65.6	11.7	44.9	-28.1	0.66	2.61
	Positive	52 (82.5)	99 (54.4)	151 (61.6)								
Coronex [®] – FAM (n=219)	Negative	8 (13.1)	12 (7.6)	20 (9.1)	92.4	13.1	73.4	40.0	70.3	5.5	1.06	0.58
	Positive	53 (86.9)	146 (92.4)	199 (90.9)								
Diagen [®] – FAM (n=147)	Negative	45 (90.0)	45 (46.4)	90 (61.2)	53.6	90.0	91.2	50.0	66.0	43.6	5.36	0.52
	Positive	5 (10.0)	52 (53.6)	57 (38.8)								
Diagen [®] – HEX (n=148)	Negative	46 (90.2)	44 (45.4)	90 (60.8)	54.6	90.2	91.4	51.1	66.9	44.8	5.57	0.50
	Positive	5 (9.8)	53 (54.6)	58 (39.2)								
Diagen [®] – FAM or HEX (n=149)	Negative	44 (86.3)	43 (43.9)	87 (58.4)	56.1	86.3	88.7	50.6	66.4	42.4	4.09	0.51
	Positive	7 (13.7)	55 (56.1)	62 (41.6)								
Anatolia [®] – FAM (n=86)	Negative	7 (77.8)	9 (11.7)	16 (18.6)	88.3	77.8	97.1	43.8	87.2	66.1	3.97	0.15
	Positive	2 (22.2)	68 (88.3)	70 (81.4)								

Some sample reactions were unavailable for analysis owing to poor curves; so, they were excluded from the analysis. Therefore, the sum of the values in the Table may be less than that for the total sample. PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; LR-, likelihood ratio.

Table 4: Comparison of Ct values of positive results for the detection kits.

Detection kits	Ct	n	Min-max	Mean $\pm$ SD	Median (IQR)
KrosGen [®] – FAM	<30	122	7.71–29.92	22.75 $\pm$ 5.08	22.33 (8.64)
	$\geq$ 30	60	30.13–34.75	31.82 $\pm$ 1.15	31.74 (1.72)
Bioexen [®] – FAM	<30	115	15.69–29.97	24.04 $\pm$ 3.82	24.42 (6.89)
	$\geq$ 30	9	30.02–36.24	31.31 $\pm$ 2.02	30.63 (1.73)
Polgen [®] – FAM	<30	22	20.18–29.45	26.75 $\pm$ 2.24	26.85 (2.84)
	$\geq$ 30	33	30.09–34.72	32.63 $\pm$ 1.56	33.15 (3.00)
Polgen [®] – HEX	<30	3	28.58–29.28	28.84 $\pm$ 0.39	28.65 (N/A)
	$\geq$ 30	3	30.08–31.08	30.54 $\pm$ 0.50	30.47 (N/A)
Coronex [®] – FAM	<30	61	18.04–29.99	25.40 $\pm$ 3.01	26.46 (4.11)
	$\geq$ 30	25	30.07–34.97	32.29 $\pm$ 1.51	32.04 (2.42)
Diagen [®] – FAM	<30	35	22.61–29.87	27.30 $\pm$ 2.01	27.28 (3.23)
	$\geq$ 30	21	30.13–33.78	31.68 $\pm$ 1.26	31.79 (2.42)
Diagen [®] – HEX	<30	35	23.49–29.75	27.51 $\pm$ 1.77	27.97 (3.13)
	$\geq$ 30	22	30.19–34.51	31.99 $\pm$ 1.35	31.55 (2.28)
Anatolia [®] – FAM	<30	54	15.84–29.85	25.44 $\pm$ 4.20	26.88 (4.88)
	$\geq$ 30	16	30.24–34.19	31.74 $\pm$ 1.29	31.20 (2.22)

(HEX); for a positive result, both *N1* and *N2* genes should be detected according to the kit data sheet. Nevertheless, as shown in Table 3, the simultaneous amplification of *N1* and

N2 decreased the consistency of the Polgen kit to 3.3%. However, assuming that the results of the Polgen kit were evaluated in such a way that the amplification of either *N1* or *N2* was sufficient to achieve a positive result, the compatibility of the kit increased to 54.4%. The Coronex and Anatolia kits were designed to detect *ORF1ab* and *N* (*N1* and *N2*) genes in the same channel (FAM), indicating that detecting one of the target genes was sufficient to produce positive results. Although the Coronex and Anatolia kits used three target genes for viral detection, the accuracy rate of the Coronex kit (70.3) was lower than that of the Anatolia kit (87.2). In contrast, Diagen uses more than one primer (*ORF1ab* in the HEX and *S* genes in the FAM channels) to detect SARS-CoV-2, and detecting only these target genes is sufficient to achieve a positive result. However, 96 of 244 samples (39.3%) could not be evaluated. Bioexen was designed to detect only *ORF1ab*, and the compatibility of the kit for detecting positive results was 65.0%.

The Ct values of positive results for both strong positive (Ct<30) and low positive (Ct $\geq$ 30) were calculated, and the Ct intervals, means, and medians are given in Table 4. The KrosGen kit had the lowest detection Ct compared with the

other kits for $Ct < 30$. In contrast, the Polgen kit (in the hex channel) only detected positive samples, with Ct intervals between 28.58 and 31.08. The mean and range of the Ct values for the other kits were similar.

Discussion

RT-PCR technique has been widely used worldwide to diagnose COVID-19 during the pandemic. This method detects viral RNA during the diagnosis of COVID-19. In basic life sciences, this method detects small amounts of nucleic acids [7, 8]. Currently, it is important to diagnose COVID-19 quickly and accurately. Various COVID-19 diagnostic kits have been used. However, it was observed that different COVID-19 diagnostic kits sometimes gave different results. There are several reasons for these different results; however, they remain controversial [9].

The first reason is the difference in the test methods used. Different manufacturers of COVID-19 diagnostic kits use various testing methods. Some kits are antibody-based and aim to detect the disease by detecting COVID-19 antibodies produced in the body [10]. Other kits use PCR to detect the genetic material of the virus. These methods are based on different principles and can give different results [11, 12].

The second reason is the quality and accuracy of the kits. There may be differences in the production and quality of the COVID-19 diagnostic kits. Certain kits provide accurate results by identifying sensitive and specific antibodies or genetic targets. However, some kits may give false positive or false negative results [13, 14]. This affects the accuracy of the test. Despite their widespread use in diagnosing COVID-19, there is speculation regarding their reliability. However, this technique is highly sensitive, specific, and preferred for pandemic conditions such as COVID-19. Various diagnostic kits have been developed for this technique in laboratories and R&D centers. The diversity of reagents specially designed for each diagnostic kit, the variety of protocols and analytical methods, and the fact that not every test is performed according to standardized procedures have raised doubts regarding the accuracy and reliability of RT-PCR for the diagnosis of COVID-19. Therefore, it was essential to compare the detection rates of the two kits. Many studies in the literature have compared commercial kits. For example, six different kits were compared for the first time in China [2]. This study found differences between samples with low and high viral loads. In another study, seven different kits were compared, and all showed an efficacy of over 96 %, indicating that these kits could be used for diagnosis [4]. Another study compared two

locally manufactured kits and found them highly compatible, suggesting they were safe for diagnosis [15].

Each kit reported in the literature has its own Ct interval values for detecting COVID-19, which may differ between kits. Although the Centers for Disease Control and Prevention accepts the detection of $Ct < 40$ as positive in some diagnostic kits, many kits used in our study and available on the market evaluate amplification results up to 35 as positive. Reactions with Ct values above 35 should be considered erroneous, and artifact binding and Taq polymerase may be inactivated. Failure to account for these factors significantly affected the results. Patients who tested positive in our study had a Ct value < 35 . However, a comparative study of the different kits revealed that the kits had different Ct values. Problems such as storage conditions, the timing of clinical specimen collection (transfer of specimens to the laboratory or improper specimen collection), and contamination may also affect the positivity of the kits. Therefore, further studies are needed to measure the reliability of different kits.

The third reason is the difference in sampling and processing methods. For COVID-19 diagnostic kits, sampling and processing must be performed correctly to obtain accurate results. For example, incorrect nasal or throat swabs may yield false positive results. In addition, improper storage or handling of the specimens can affect the results. This possibility was eliminated in our study by selecting patient samples, which was also supported by positive clinical findings.

The fourth reason is variations in the virus. Viral mutations and variants emerged during the COVID-19 pandemic. Some COVID-19 diagnostic kits better detect certain variants, whereas others are less sensitive. This could have led to different results. However, this possibility was not assessed because the studies were performed simultaneously on samples from the same patients (Table 1).

In this study, five of the six kits have reliable compatibility at $Ct < 30$, but this decreases at $Ct \geq 30$. When the viral load is high, it is easier to detect the presence of viruses in the nasopharyngeal aspirate/lavage/swab specimens. However, detecting positive results in individuals recently infected with SARS-CoV-2 with a low viral load is critical. Therefore, it is important to evaluate the performance of these kits in reduced viral load.

One of the major differences among the kits is the target gene region used for primer design. The minor differences among the kits may be due to the primers used for the target gene region. These are the regions of the virus used to determine the genes, especially in positive samples. Each commercial kit uses primers that amplify at least one gene region. The most commonly used gene regions for these kits are *ORF1ab*, *N*, and *S*. Our study showed that utilizing one or more gene regions in the same kit did not affect the accuracy.

For example, although a single gene region (*ORF1ab*) was used in the highly reliable Bioexen kit, primers for two gene regions were used in the Diagen kit. However, the results showed that the Bioexen was more consistent than the Diagen. The same conclusion applies to the Polgen and Coronex kits.

Fluorophores are used in the TaqMan probe-based kits. The most common fluorophores used in COVID-19 diagnostic kits are FAM and HEX, which contain fluorescent reporters and quenching dyes. In probe-based assays, the probe must be designed to be approximately 10 °C higher than the primer binding temperature to ensure correct binding. Therefore, long probes have been designed to detect regions with low cytosine-guanine (GC) ratios. However, if the probes are too long, unnecessary background radiation will occur [16]. This phenomenon is a possible reason for the unwanted radiation observed with Diagen and Polgen kits. We conclude that the GC ratios of primers, probes, primer annealing temperatures, and probe selection should be carefully determined when designing kits. The more accurate the combination, the more reliable the kit is. Although using more than one fluorophore is advantageous, optimization can be challenging [17]. Multiple fluorophores can lead to misdiagnosis of the disease, especially in pandemic conditions, and to a loss of confidence in kits, where most R&D optimizations are performed in the field.

Three-dimensional (3D) modeling of the proteins encoded in the viral genome was performed to determine whether the stability or unstable nature of the proteins encoded by the gene regions used in the kits affected the reliability of the preferred target regions. PHYRE2 database (<http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index>) was used to determine whether there is a relationship between the stability of the proteins targeted by the primers used in the kits and the reliability of the kits. A visual comparison of the results, in which 3D modeling was performed using homology-based approaches, showed that the proteins with the most stable structures belonged to the *ORF1ab* (*ORFa* and *ORFb*) regions. Therefore, kits (Polgen, Diagen, Coronex) using primers from the *S* and *N* regions had lower reliability, and the protein models were more unstable (Figure 1). Although KrosGen and Polgen used the same gene regions for positive detection, KrosGen provided more consistent results than Polgen. When comparing these two kits, there was a 10 °C difference in primer annealing temperatures (Table 1). The fact that the Polgen kit uses two different fluorophores and has a 10 °C higher annealing temperature than the KrosGen suggests two possibilities. One is the presence of various fluorophores, which may affect the irradiation. Another possibility could be differences in the GC ratio of primer candidates belonging to the

same gene region. It would be better for kit manufacturers to share their primer sequences. Another difference between the two kits is the ability to evaluate gene amplification in one or two distinct channels. As shown in Tables 2 and 3, the kits assess the presence of COVID-19 in one channel (using one or two genes) exhibited a higher compatibility and accuracy rate than those evaluating more than one gene in two different channels. Kits that used two genes in one channel to detect COVID-19 provided results with the highest accuracy rates and compatibility.

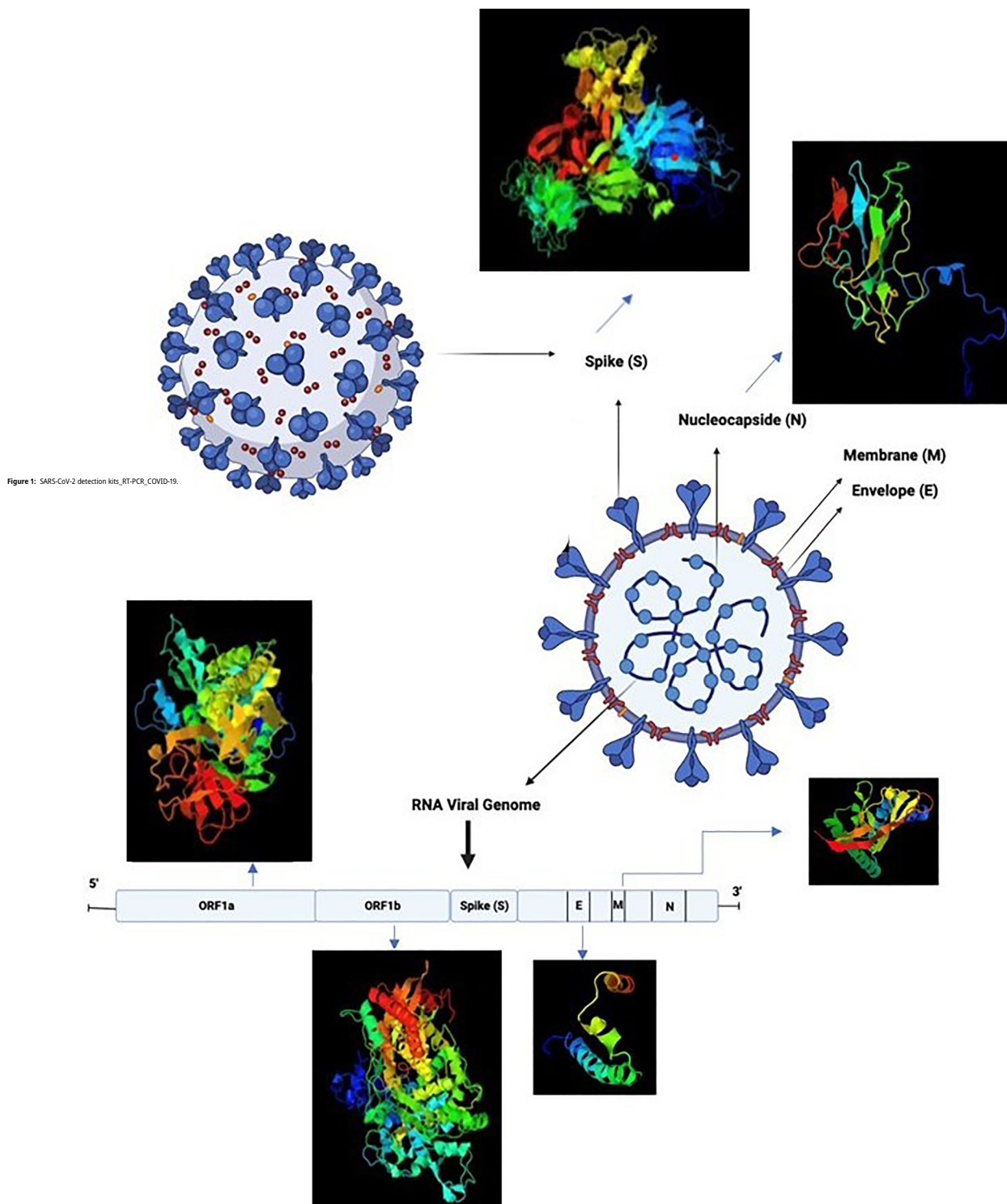
Conclusions

One reason for the uncontrolled global spread of COVID-19 is the false negative test results in patients with a low COVID-19 viral load. Therefore, a sensitive and suitable kit with high compatibility for $Ct \geq 30$ is essential to identify patients with a low viral load. This study compared six commercial COVID-19 detection kits for patients with low viral loads. We found that the KrosGen, Anatolia, and Bioexen kits were the most effective in detecting SARS-CoV-2.

This study has several limitations. One of these limitations is the lack of a gold standard, such as viral culture and/or international standards. Viral culture can only be performed in a Biosafety Level 3 laboratory requiring special conditions. There are fewer than 10 BSL3 laboratories in Turkey. Although a World Health Organization international standard exists, it is difficult to obtain with a limited budget. These kit manufacturers are also expected to use this standard to validate and verify their kits. However, several studies have compared kits that do not use this standard [4, 5].

The second limitation is related to the sample size of the Anatolia kit. The Anatolia kit did not reach the researchers until the final phase of the study, leaving insufficient time and resources for thorough examination with a representative sample size. This unexpected delay in acquiring the Anatolia kit hampered our ability to conduct a comprehensive analysis, as originally intended. Nevertheless, given that our sample size fell below the threshold set determined in our power analysis, we felt it was imperative to include these findings in our study because of their remarkable consistency compared with other kits, as they are more in line with our research findings.

In conclusion, to improve the accuracy and compatibility of kits, we suggest using primers to detect two genes with a single probe evaluated in a single channel. In addition, for those samples with $Ct \geq 30$, it is better to repeat the reactions with 1:1 and, if needed, with 1:9 dilutions. A consistently positive result was considered reliable. However, in cases where



the results are inconsistent, it is advisable to repeat the reaction using a new sample from the patient. Nevertheless, it is important to keep in mind that regardless of the brand of the kit, a validation study should be performed

before routine use of the kit for COVID-19 detection. Finally, it is strongly recommended that interlaboratory comparisons and external quality control measures validate the results of kit reactions.

Acknowledgments: We thank Aysin Acar, Sabri Onder and Ercan Gundemir, the staff of PCR-SARS-CoV-19 laboratory who helped us during the experiments.

Research ethics: The study protocol was approved by the Institutional Ethics Committee of TOBB University of Technology and Economy, ANKARA (permit no. KAEK-118/143 (dated 11.11.2022)). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Informed consent: Informed consent was obtained from all individuals included in this study.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission. The main idea of this project and conceptualization developed by Aysegul Taylan Ozkan. The experiments were designed by Aysegul Taylan Ozkan, Parisa Sharafi and F. Seyma Gokdemir. The results reaction for Covid-PCR reaction and questionnaire for symptoms were carried out by J. Sedef Gocmen and Yasemin Yardicoglu Akisin and Parisa Sharafi. The experiments performed by F. Seyma Gokdemir and Parisa Sharafi. F. Seyma Gokdemir and Parisa Sharafi wrote the main manuscript text and results. F. Seyma Gokdemir and Parisa Sharafi prepared the figures and tables respectively. The statistics and power analysis was calculated by Mesut Akyol.

Competing interests: The authors state no conflict of interest.

Research funding: The study was financially supported by TOBB Economy and Technology University Hospital.

Data availability: The raw data can be obtained on request from the corresponding author.

References

- Rodrigues CO, Olivares IRB, Reischak D, Catozzi MB, Cardoso FG, Pinhel MFM. Development and application of proficiency testing (PT) to evaluate the diagnostic capacity of SARS-CoV-2 by RT-qPCR – a practical and metrological approach. *Diagn Microbiol Infect Dis* 2023; 107:116021.
- Guo Y, Wang K, Zhang Y, Zhang W, Wang L, Liao P. Comparison and analysis of the detection performance of 6 domestic novel coronavirus nucleic acid detection reagents. *Chongqing Med J* 2020; 49:2435–9.
- Yüce M, Filiztekin E, Özkaya KG. COVID-19 diagnosis – a review of current methods. *Biosens Bioelectron* 2021;172:112752.
- van Kasteren PB, van der Veer B, van den Brink S, Wijsman L, de Jonge J, van den Brandt A, et al. Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. *J Clin Virol* 2020;128:104412.
- Lu Y, Li L, Ren S, Liu X, Zhang L, Li W, et al. Comparison of the diagnostic efficacy between two PCR test kits for ratio of the target gene molecules -CoV-2 nucleic acid detection. *J Clin Lab Anal* 2020;34:e23554.
- Kudo E, Israelow B, Vogels CBF, Lu P, Wyllie AL, Tokuyama M, et al. Detection of SARS-CoV-2 RNA by multiplex RT-qPCR. *PLoS Biol* 2020;18: e3000867.
- Lau LT, Fung YW, Wong FP, Lin SS, Wang CR, Li HL, et al. A real-time PCR for SARS-coronavirus incorporating target gene pre-amplification. *Biochem Biophys Res Commun* 2003;312:1290–6.
- Sule WF, Oluwayelu DO. Real-time RT-PCR for COVID-19 diagnosis: challenges and prospects. *Pan Afr Med J* 2020;35(2 Suppl):121.
- Younes N, Al-Sadeq DW, Al-Jighefee H, Younes S, Al-Jamal O, Daas HI, et al. Challenges in laboratory diagnosis of the novel coronavirus SARS-CoV-2. *Viruses* 2020;12:582.
- Xu J, Kerr L, Jiang Y, Suo W, Zhang L, Lao T, et al. Rapid antigen diagnostics as frontline testing in the COVID-19 pandemic. *Small Sci* 2022;2:2200009.
- Falzone L, Gattuso G, Tsatsakis A, Spandidos DA, Libra M. Current and innovative methods for the diagnosis of COVID-19 infection (Review). *Int J Mol Med* 2021;47:100.
- Artik Y, Coşgun AB, Cesur NP, Hızal N, Uyar Y, Sur H, et al. Comparison of COVID-19 laboratory diagnosis by commercial kits: effectivity of RT-PCR to the RT-LAMP. *J Med Virol* 2022;94:1998–2007.
- Gong F, Wei HX, Li Q, Liu L, Li B. Evaluation and comparison of serological methods for COVID-19 diagnosis. *Front Mol Biosci* 2021;23: 682405.
- Rivera-Olivero IA, Henríquez-Trujillo AR, Kyriakidis NC, Ortiz-Prado E, Laglaguano JC, Vallejo-Janeta AP, et al. Diagnostic performance of seven commercial COVID-19 serology tests available in South America. *Front Cell Infect Microbiol* 2022;12:65.
- Şenol Akar Ş, Akçali S. Comparison of two different commercial SARS-CoV-2 PCR diagnostic kits. *Flora* 2022;27:177–82.
- Arya M, Shergill IS, Williamson M, Gommersall L, Arya N, Patel HR. Basic principles of real-time quantitative PCR. *Expert Rev Mol Diagn* 2005;5: 209–19.
- Blacket MJ, Robin C, Good RT, Lee SF, Miller AD. Universal primers for fluorescent labelling of PCR fragments – an efficient and cost-effective approach to genotyping by fluorescence. *Mol Ecol Resour* 2012;12: 456–63.