ORIGINAL RESEARCH ARTICLE UHD JOURNAL OF SCIENCE AND TECHNOLOGY

Enhancing COVID-19 Detection Accuracy: Optimal Gene Combinations, Kit Performance, and Reliable Detection Intervals

Dara A. Tahir[1], Sehand Kamaludeen Arif[2]

[1,2]*Department of Biology, University of Sulaimani, Al sulaimaniyah, Iraq*

ABSTRACT

A significant challenge and threat to public health have been created by the COVID-19 pandemic for the entire global population. The study aimed to compare the SARS-CoV-2 RNA detection capabilities of available primers and probes to identify the most reliable, efficient, and affordable method. From 200 previously detected samples of SARS CoV-2, 94 samples were selected randomly and used for the optimization of our primers and probes. We compared our results with two kits that have been approved by the health authority. In addition, we evaluated the detectability of each gene. The study compared the diagnostic performance of different gene combinations for COVID-19 detection using kits A and B and a novel approach combining *RdRp, N*, and *E* genes. Results showed that the combined approach exhibited superior discriminatory power, particularly with the inclusion of the *E* gene, boasting area under the curve (AUC) values of 83.3%, 79.1%, and 93.7% for the respective genes. Kit B, with Orf1ab and *N* genes, outperformed Kit A (*RdRp* and *S* genes), with AUC values of 81.2% and 90.6% versus 80.2% and 75%, respectively. The chart representation highlighted gene detection frequencies across various cycle threshold (Ct) ranges, demonstrating robust identification within the 20.1–30 Ct range across all kits and genes, emphasizing the reliability of detection within specific intervals. Combining *RdRp, N*, and *E* genes showed the highest accuracy for COVID-19 diagnosis, particularly with the *E* gene. Detection was most reliable within the 20.1–30 Ct range across all gene combinations and kits.

Index Terms: SARS-CoV-2, *RdRp* Gene, *E* Gene, *N* Genes, Kit Comparison

1. INTRODUCTION

Coronaviruses cause various illnesses from common colds to severe diseases such as syndrome coronavirus 2 (SARS-CoV) and COVID-19. These viruses carry a single-stranded RNA genome ranging from 26 to 32 kb. They are classified into four genera – Alpha, Beta, Gamma, and Delta, and they infect both humans and animals, with around 30 species found across human, mammal, and bird samples. Human coronaviruses fall into alpha and beta species [1]. Following 229E, NL63, OC43, HKU1, Middle East Respiratory Syndrome (MERS)-CoV, and the initial SARS-CoV outbreaks, SARS-CoV-2 marks the eighth known coronavirus to impact human populations [2].

The SARS-CoV outbreak in 2002–2003 stemmed from class B beta-coronaviruses originating from bats, whereas the MERS resulted from a class C beta-coronavirus linked to camels in 2012 and beyond [3]. The pneumonia case identified in Wuhan, China, in late December 2019, initially attributed to a coronavirus due to respiratory symptoms, was designated as severe acute respiratory SARS-CoV-2 by the World Health Organization [4], [5].

Access this article online

DOI: 10.21928/uhdjst.v8n1y2024.pp84-92

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Corresponding author's e-mail: Dr. Sehand Kamaludeen Arif, Department of Biology, University of Sulaimani, Al sulaimaniyah, Iraq. E-mail: sehand.arif@univsul.edu.iq

Received: 07/12/2023 Accepted: 20/12/2023 Published: 11/03/2024

The novel virus' genomic sequence is 96.2% identical to that of the bat. In Wuhan, China, researchers discovered the SARS-related coronavirus (SARSrCoV; RaTG13), which contains a genome that is not entirely identical to either the SARS-CoV (about 79%) or the MERS-CoV (about 50%) genomes [6], [7].

Developing specific primers and probes in regions with minimal similarity to other viruses is crucial to avoid misidentification of SARS-CoV-2 [8]. As the most sensitive and specific assay, real-time reverse transcription-quantitative polymerase chain reaction (RT-qPCR) is the "gold standard" diagnostic technique for the detection of SARS-CoV-2 infection [9].

SARS-CoV-2 displays the typical genome organization of the β-coronaviruses. There are 14 functional open reading frames in the genome, including two noncoding regions at each end and other areas that code for structural, accessory, and nonstructural proteins [10].

This study aims to detect SARS-CoV-2 using various probes and primers that target different genes (such as *RdRp, E*, and *N* genes) through reverse transcription real-time and conventional PCR assays. The goal is to compare the efficacy of these primers/probes in detecting SARS-CoV-2 RNA and identify the most accurate, speedy, and cost-effective detection method.

2. MATERIALS AND METHODS

2.1. Sample Collection

Between September 2020 and September 2021, ninety-four samples were selected randomly from a total of 200 samples (Nasopharyngeal swabs) of SARS-CoV-2 previously gathered by Shahid Tahir Ali Wali Bag laboratory in Sulaimani City/Iraqi Kurdistan Region. The (94) selected samples were used for optimization of our primers and probes and for comparing the obtained results with two available kits (AddMedi/South Korea and AeHealth/UK) proved by health authority and evaluation of detectability of each gene.

2.2. RNA Extraction

RNA extractions were carried out using an Addprep viral nucleic acid extraction kit according to the manufacturer's protocol (AddBio Company/South Korea).

2.3. Storage of the RNA

To avoid degradation, eluted RNA was stored at −80°C after the RNA extraction process was completed.

2.4. Designing of primers and Probes for RT-qPCR

Based on the sequencing analysis, four sets of primers and probes were designed as follows: One set of forward primers, reverse primer, and dual modified fluorescence probe for each of RdRp, N, E, and GAPDH gene for internal control detection. Amplicons were selected for *RdRp*, *N,* and *E* genes by Primer-BLAST, which were (110, 110, and 80 bp), respectively, and *in silico* analyzed using NCBI Primer BLAST to exclude miss annealing with other respiratory viruses. *In silico* predictions also demonstrated that those primers do not bind to non-specific targets for human, bacterial, fungal, and apicomplexan.

The lyophilized form of primers and probes was produced by a humanized genomic Macrogen company from South Korea (Tables 1 and 2).

2.5. RT-qPCR Procedure

Samples containing RNA were subjected to a reverse transcription step to convert RNA into cDNA using reverse transcriptase enzymes. The reaction mixes, containing the designed primers, probes, nucleotides, and specialized polymerase enzymes, were prepared and inserted into the (applied biosystems RT-qPCR) instrument. The instrument cycled through denaturation, annealing, and extension stages, amplifying the target sequences exponentially. Fluorescence emitted during each cycle was monitored, with the software recording and analyzing the signal, determining the initial amount of viral genetic material present in the samples based

TABLE 1: Primers and probes for real-time polymerase chain reaction (*RdRp, N, E* gene)

Genes	Primers and probes
RdRp F	TTGATTGTTACGATGGTGGCT
RdRp R	CATAATAAAGTCTAGCCTTACCCCA
RdRp-Probe	FAM- GGTTGTTGACGATGACTTGGTTAGCA-BHQ
N F	ATTCGTGGTGGTGACGGTAA
N R	ATGCCGTCTTTGTTAGCACC
N Probe	HEX-GGGAAGTCCAGCTTCTGGCCC-BHQ
E F	ACTTCTTTTTCTTGCTTTCGTGGT
E R	GCAGCAGTACGCACACAATC
E Probe	ROX- AGCGCAGTAAGGATGGCTAGTGT -BHQ

TABLE 2: Primers and probes for real-time polymerase chain reaction (internal control GAPDH)

Internal control GAPDH	
IC F	GGTGAAGGTCGGAGTCAACG
IC R	TGAAGGGGTCATTGATGGCAA
IC Probe	Cy5- CTGGTGACCAGGCGCCCAAT-BHQ

on the Ct values. Results were interpreted by comparing Ct values against known standards or controls to ascertain the presence and approximate quantity of SARS-CoV-2 genetic material in the samples.

2.6. Statistical Analysis

Statistical analysis was conducted using two software programs: IBM SPSS (Version 26) and RStudio (Version 1.1.463). These programs were employed to perform various statistical tests (area under the curve [AUC], z – test, and *P*-value) and analyses, ensuring a comprehensive and rigorous approach to data analysis.

2.7. AUC

AUC refers to the graphical representation of a receiver operating characteristic (ROC) curve. In this context, it reflects the assay's ability to distinguish between positive and negative samples based on Ct values. A higher AUC value indicates better discrimination between these samples, signifying the test's accuracy in correctly identifying the presence or absence of SARS-CoV-2 genetic material. This statistical measure helps assess the overall performance and reliability of the PCR assay by quantifying its ability to differentiate between infected and non-infected samples based on Ct values, contributing crucial insights into the assay's diagnostic potential.

Z-test: Z-values represent the number of standard deviations; a data point is from the mean. In this case, the more negative the Z-value, the lower the Ct value, indicating higher viral RNA levels or better detection sensitivity.

P-value: The low *P*-values indicate high statistical significance. It suggests strong evidence against the null hypothesis, reinforcing the reliability of the observed differences.

3. RESULTS

3.1. Detectability of Primers and Probes in Clinical Samples

Out of 200 previously identified SARS-CoV-2 samples, 94 were randomly chosen for the optimization of our primers and probes. The obtained findings were compared with two health authority-endorsed kits to evaluate the detectability of each gene: Kit A with CE and IVD certificate from South Korea with detection channel HEX (RdRp gene) and FAM (S gene), while Kit B with CE and IVD certificate from United Kingdom with detection channel FAM (ORF1ab gene) and VIC (N gene) and the current designed primers as well as probes FAM (RdRp), HEX (N), ROX (E), and CY5 (IPC).

Kit A, which consists of RdRp and S genes, demonstrated an AUC of 80.2% and 75%, respectively. This indicated that Kit A's performance, as measured by the AUC, is 80.2% for one gene and 75% for the other. The AUC value served as a metric for the ability of this kit to distinguish between positive and negative COVID-19 samples. Higher AUC values generally indicated better discriminatory power.

In contrast, Kit B, comprising the Orf1ab and N genes, exhibited AUC values of 81.2% and 90.6%. These AUC values suggested that Kit B is quite effective, with the first gene achieving an AUC of 81.2% and the second gene reaching an even higher AUC of 90.6%. The higher AUC for the N gene in Kit B may indicate its higher accuracy in diagnosing COVID-19 cases. The approach of the current study, which combines the RdRp, N, and E genes, achieved the AUC values of 83.3%, 79.1%, and 93.7%, respectively. These AUC values demonstrated the effectiveness of this approach when using different gene combinations. The AUC of 93.7% for the E gene suggested that it may be a particularly valuable component for COVID-19 diagnosis within this approach.

To visualize the data and facilitate a more comprehensive understanding, the information is presented in a bar chart (Fig. 1).

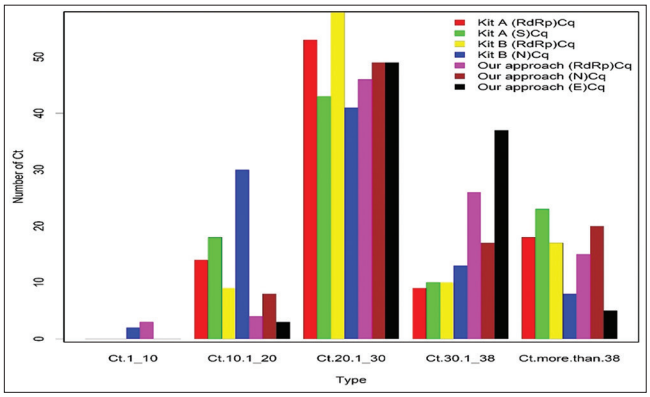

Fig. 1: The X-axis of the chart represents the number of samples tested, while the y-axis indicates the Ct values. Ct values are often used in polymerase chain reaction tests and indicate the number of amplification cycles required to detect a specific gene target in a sample. Gene detection across Ct value ranges in various test kits. The bar chart visually presents gene detection within specific Ct value ranges across different test kits and methodologies, it's important to note that the statistical significance of differences in detection rates was assessed separately through appropriate statistical tests. These tests, including but not limited to t-tests, ANOVA, or non-parametric equivalents, were conducted to evaluate the significance of differences in detection rates between the various gene targets across different kits and methodologies. The results of these statistical analyses provide quantitative insights into the observed differences in detection rates, complementing the visual representation provided by the bar chart. Ct: Cycle Threshold.

The chart allows for a visual comparison of how the AUC values for Kit A, Kit B, and the current approach with various gene combinations vary with the number of samples tested and their corresponding Ct values. This comprehensive comparison of kits and gene combinations could be useful in making informed decisions about which testing method may be the most suitable for COVID-19 diagnosis, taking into account both accuracy and scalability based on the number of samples. It appeared that the current approach with three genes (RdRp, N, and E) showed the highest AUC values, which are 83.3%, 79.1%, and 93.7%, respectively. These AUC values are indicative of the accuracy and discriminatory power of the diagnostic method.

Comparatively, Kit B, which combines Orf1ab and N genes, has AUC values of 81.2% and 90.6%. Kit A, with RdRp and S genes, has AUC values of 80.2% and 75%. Based on these AUC values alone, it appeared that the current approach with three genes was the most effective in diagnosing COVID-19, especially when the E gene is included which appeared to contribute significantly to the accuracy of the test, as reflected in its higher AUC value of 93.7%.

In the comparison between kits with two genes (Kit A and Kit B), Kit B, with AUC values of 81.2% and 90.6%, seemed to be more effective than Kit A, which has AUC values of 80.2% and 75%.

The Ct value ranges are color coded to depict the distribution of gene detection levels. Notably, the range of 20.1–30 exhibits the highest frequency of gene detection across all kits and genes, indicating consistent and reliable identification within this Ct range. The chart highlights the varying detection levels among different genes and kits, emphasizing the robustness of detection within specific Ct value intervals.

The Z-test results indicate significant differences ($P < 0.0001$) among gene targets tested within Kit A, B, and our approach. The *E* gene exhibits the lowest Z-value of −1.49, followed closely by the *N* gene (kit B) with a Z-value of −1.46. These findings suggest that *E* gene demonstrates the highest sensitivity in detecting COVID-19.

4. DISCUSSION

Since the first record of the coronavirus disease 2019 (COVID-19) pandemic, real-time RT-qPCR laboratory testing to identify severe acute respiratory SARS-CoV-2 has been crucial in limiting the spread of the virus [11]. The World Health Organization developed and made a number of RT-qPCR assays accessible to the general public not long after the viral genome sequences became available. The *E* gene, *RdRp* gene, *ORF1ab* gene, and *N* gene are only a few of the target genes within the viral genome that was used to select the primer and probe sequences for these experiments. Based on these primer and probe sequences, numerous commercial and laboratory-developed tests for SARS-CoV-2 detection were created. Several mutational events have resulted from the widespread sustained person-to-person transmission of SARS-CoV-2, some of which may alter the sensitivity and specificity of existing PCR assays [12]. The gold standard for diagnosing COVID-19 is rRT-PCR, but it has been noted that false-positive and false-negative samples may cause issues. Regarding which of the goals is most important for diagnosis, there are differing viewpoints. Testing with three targets is more expensive and time-consuming than testing with two targets. The conclusion drawn from our study is based on a comparative analysis of two different assay methods: one that detects N and RdRp genes (utilizing commercial kits) and another that detects N, E, and RdRp genes (our manual 3-gene method). While it's true that the comparison involves a 3-gene method versus a 2-gene method, the primary focus is on evaluating the sensitivity difference between these two approaches rather than a direct comparison of equal gene sets. Our conclusion is derived from the observation that the assay detecting N, E, and RdRp genes consistently yielded more positive detections across the tested samples compared to the assay detecting only N and RdRp genes. This difference in sensitivity forms the basis of our conclusion that the 3-gene assay is more sensitive than the 2-gene assay.

Barjaktarović *et al.* revealed that the new variations of SARS-CoV-2 are continuously found since it is an RNA virus that mutates more frequently than DNA viruses. Sequences that are less mutable are the focus of diagnostic testing. Yet, it has been discovered that mutations in the *N* gene and the *RdRp* gene increase the possibility of false-negative results, reduce test sensitivity, and result in diagnostic errors [13]. Because of the *E* gene's high stability, it is possible to develop new mRNA vaccines against illnesses caused by newer SARS-CoV-2 strains. Moreover, known coronaviruses have highly conserved *E* and *M* gene sequences. They are less likely to experience mutations than the *S* gene, making the *E* gene a suitable diagnostic target with a high level of specificity. It is vital to use diagnostic tests that target less changeable genes to identify people infected with present and future variations of the SARS-CoV-2 virus since new SARS-CoV-2 variants are continually developing and complicating COVID-19 testing [13].

The quantitation cycle, which is the primary outcome of qPCR, is the cycle in which fluorescence can be observed. Lower Ct values indicate higher starting copy numbers of the target. This is the fundamental tenet of the quantitative strategy offered by real-time PCR. Previous researchers have recorded comparisons of the effectiveness of various commercial PCR kits for the RT-qPCR diagnosis of SARS-CoV-2 [14], [15], [16], [17].

Lu *et al.* in 2020, compared and evaluated the effectiveness of Sansure and BioGerm, which are both commonly used in Liuzhou people's Hospital in Guangxi, China, and had corresponding effectiveness levels of 80 and 94% [8]. On the other hand, Eberle *et al* in 2021 examined nine RT-qPCR kits that were being utilized in the German city of Bavaria to diagnose viruses. Most of them achieved sensitivity levels of 90–100%, however, two kits claimed efficacy levels of 49% (Fast Track Diagnostics Kit) or 62% (Wells Bio, Inc.) with the greatest proportion of false negatives. Hence, when it comes to the identification of viral variations, kits with more than one target gene are less likely to produce false negative results than tests with a single genetic target [18], [19].

These studies recommended evaluating the effectiveness of commercial RT-qPCR kits used locally to analyze COVID-19. In fact, a poor SARS-COV-2 diagnosis could encourage the future spread of this and other infectious diseases. No standardization, comparison, or efficacy investigations using commercial RT-qPCR kits used in the mass diagnosis of local SARS-CoV-2 or the detection of developing variants have been reported for Sulaimani, the city with a high number of PCR tests performed per thousand residents.

The current study announces the first clinical validations and comparison of two commercially available RT-qPCR Kits for detecting SARS-CoV-2 Kit A and Kit B to primers and probes we designed. For the detection of SARS-CoV-2 genes and diagnosis, the Ct and the relative fluorescence units acquired from their RT-qPCR reactions exhibited significant disparities in the total RNA volume. These differences were significant because they significantly affected the detection of COVID-19-positive cases.

The evaluation of different testing kits and gene combinations in the context of COVID-19 diagnostics, as measured by AUC values, offered crucial insights into their accuracy and discriminatory power [20]. AUC values, representing the ability to distinguish between positive and negative COVID-19 samples, showcased varying performance among the tested kits and gene combinations. Notably, higher AUC values typically indicated better discriminatory ability, serving as a fundamental metric for assessing the efficacy of these diagnostic methods. Comparing Kit A and Kit B revealed intriguing nuances in their performance. While Kit A demonstrated AUC values of 80.2% for RdRp and 75% for the *S* gene, Kit B exhibited a relatively higher accuracy, with AUC values of 81.2% for Orf1ab and an impressive 90.6% for the *N* gene. This comparison suggested that Kit B, particularly the *N* gene within it, displayed a superior capacity in accurately diagnosing COVID-19 cases compared to Kit A.

The findings of this study underscore the pivotal role of gene selection in optimizing the accuracy of COVID-19 diagnosis. Our results revealed compelling evidence supporting the superior performance of the combined *RdRp*, *N*, and *E* genes, showcasing robust discriminatory power with notable area under the curve (AUC) values. Particularly striking was the outstanding 93.7% AUC value for the *E* gene within this multi-gene approach, surpassing individual gene performances within both Kit A and Kit B. The statistically significant differences ($p < 0.0001$) observed in AUC values among these genes signal the potential for the *E* gene as a cornerstone in enhancing diagnostic precision, suggesting its superiority over RdRp and N genes in our tested context. The remarkably high AUC for the E gene highlights its promise as a key component in amplifying the reliability and accuracy of COVID-19 diagnosis, offering a potential avenue for refining

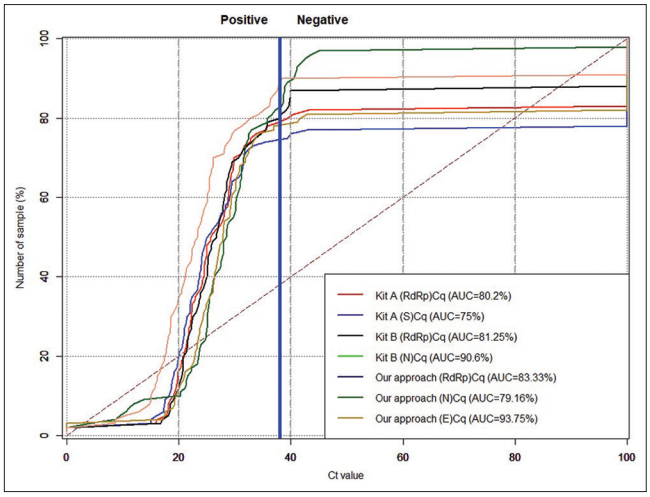

Fig. 2. Line graph shows the performance of different testing kits and our approach using specific genes evaluated in the context of COVID-19 diagnostics. The assessment is based on the calculation of the area under the curve values, which are a common measure of the accuracy of diagnostic tests. The comparison involved three different sets of genes: Kit A (RdRp and S), Kit B (Orf1ab and N), and our approach combining RdRp, N, and E genes.

Sample NO.	Kit A				Kit B				Our approach							
	RdRp		S		Orf1ab		N		RdRp		N		E		IC	
	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq
1	1130	21.06	3546	19.95	898	21.58	837	16.77	92.3	24.95	198	23.01	1072	24.27	145	26.02
2	1921	24.66	5395	24	1006	25.06	955	22.71	115	9.57	49.4	33.99	276	33.72	182	27.18
3	1753	15.92	5351	14.95	975	16.77	1045	10.43	121	28.63	184	29.15	845	29.15	159	28.30
4	1602	20.49	4719	19.67	1059	20.82	960	18.19	2.79	45.00	0.81	0.0	67.3	37.40	196	25.36
5	1755	27.50	3699	26.61	978	28.08	925	25.22	107	30.01	205	27.15	919	26.86	134	29.29
6	0.92	0.0	-2.00	0.0	259	39.91	544	36.43	94.6	23.39	235	21.20	1334	22.44	184	25.14
7	1754	18.12	4612	17.31	981	18.55	862.	16.18	108	27.80	209	26.16	867	28.38	178	27.05
8	1513	28.26	2637	27.32	994	28.58	947	24.05	95.8	23.56	206	26.29	1054	26.67	192	26.05
9	1847	19.32	4032	18.52	873	20.86	792	15.07	102	26.12	208	24.39	1041	25.24	217	28.60
10	1671	26.76	2721	26.06	1031	26.79	954	23.37	2.69	21.14	243	17.62	1451	19.22	224	25.11
11	2053	24.55	3636	24.03	663	26.77	517	25.53	113	24.94	277	23.56	1485	24.71	152	25.01
12	1957	20.54	4240	19.89	1193	20.81	1110	18.48	136	30.18	217	29.45	725	30.51	162	31.14
13	1988	19.83	4786	18.82	1147	19.62	1028	15.30	86.4	23.35	246	20.14	1435	21.87	199	26.59
14	1669	28.87	1736	29.02	1063	28.31	1016	25.51	109	25.17	94.0	29.24	1300	25.07	131	27.25
15	1925	21.79	4122	21.07	1037	22.31	966	19.11	113	28.01	231	26.39	1096	27.86	168.	26.45
16	1750	17.77	4576	16.08	978	18.17	873	12.90	121	29.03	226	27.02	879	29.05	277	27.14
17	1632	24.13	3066	23.53	986	23.86	929	18.67	98	23.66	257	19.73	1394	22.26	142	27.99
18	1811	22.90	3537	22.21	1011	23.38	979	21.05	6.90	31.02	216	28.06	719	30.41	232	28.41
19	1999	25.71	3266	25.06	1233	25.86	1154	20.61	101	21.42	277	20.84	1818	21.07	190	23.09
20	2339	19.58	5366	18.70	1185	20.33	1093	15.86	-2.24	34.36	-2.86	0.0	40.2	42.05	169	27.10
21	1948	21.82	4173	21.09	1094	22.29	979	18.09	127	30.31	191	29.41	785	30.31	219	25.60
22	747	35.60	260	39.95	689	35.83	949	32.01	99.6	25.83	258	24.04	1327	24.33	142	29.22
23	1896	25.07	3396	24.28	1238	25.21	1099	22.91	99.7	20.29	265	17.01	1698	18.75	165	24.08
24	845	33.82	349	35.39	908	33.65	944	29.52	1.79	11.95	9.19	41.07	582	30.93	135	28.51
25	1690	21.09	3806	20.15	1107	21.45	1046	17.80	89.8	23.82	240	22.10	1272	23.55	324	23.72
26	1589	29.75	1876	29.35	1073	29.51	1040	25.19	87.3	31.29	183	30.21	631	31.30	219	27.16
27	1.59	0.0	-1.1	0.0	503	39.75	730	36.98	1.76	40.97	-1.84	18.98	68.4	37.84	581	25.54
28	1.08	0.0	-1.48	0.0	-1.86	0.0	-2.15	0.0	95.4	4.19	-0.22	0.0	165	34.99	161	30.93
29	1966	24.60	3313	24.15	1157	24.87	1114	22.26	113	28.03	230	26.89	1214	27.37	216	24.76
30	1.06	0.0	1.91	0.0	309	39.65	655	36.33	2.78	43.60	6.30	42.82	116	36.07	179	27.79
31	2032	22.29	4164	21.46	1080	22.36	1001	19.72	112	25.07	277	24.46	1475	25.64	162	26.05
32	1458	28.99	1758	28.36	936	29.00	928	26.14	120	30.50	147	30.65	633	31.43	200	25.55
33	1492	21.21	2884	20.80	844	20.82	789	16.63	97.9	24.66	229	22.57	1241	23.56	181	26.29
34	1891	22.19	4145	21.23	1090	21.74	1017	18.50	123	25.15	277	23.37	1475	24.51	198	25.02
35	1.38	0.0	1.41	0.0	0.69	0.0	-1.06	0.0	1.71	40.80	0.69	0.0	60.0	37.89	464	27.07
36	2003	25.63	3467	24.82	1101	25.85	1021	22.45	137	27.53	273	26.67	1097	28.22	345	25.91
37	1781	28.82	2346	28.21	1207	28.89	1145	25.21	126	31.02	208	29.63	757	30.56	356	25.12
38	2106	18.11	5228	17.44	1203	18.52	1072	15.12	87.3	21.59	286	19.52	1739	20.26	156	25.30
39	2059	24.53	3967	23.44	1114	24.58	1002	19.84	133	27.35	296	24.23	1269	26.18	335	26.25
40	317	39.34	104	0.0	799	35.76	868	31.21	-1.76	32.94	2.21	0.0	108	36.13	459	25.31
41	700	35.85	271	39.36	756	34.65	855	30.09	12.4	38.38	4.22		146	34.97	256	29.20
42	2036	23.03	3750	22.37	1038	23.53	995	20.05	119	25.17	275	24.83	1460	25.60	233	26.31
43	1272	32.86	826	33.20	1029	32.98	1033	29.28	54.8	32.46	59.7	33.80	376	33.10	597	25.34
44	2104	16.27	5392	15.01	1128	17.09	1207	8.81	100.4	20.39	300	17.60	1883	18.81	192	24.92
45	67.4	0.0	1.72	0.0	0.23	0.0	-1.08	0.0	-0.42	38.96	2.77	0.0	14.5	0.0	165	30.45
46	2012	19.77	4664	19.09	1103	20.08	1032	17.15	111	25.25	294	22.84	1607	23.86	206	25.78
47	2056	19.11	5007	18.33	1182	19.45	1093	15.78	87.1	23.96	283	20.36	1634	21.52	139	27.22

(Contd...)

Sample NO.	Kit A				Kit B				Our approach							
	RdRp		S		Orf1ab		N		RdRp		N		E		IC	
	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq	RFU	Cq
48	1903	21.77	4077	20.60	1014	21.90	947	17.94	91.7	25.54	253	22.79	1308	24.77	185	25.97
49	1640	21.39	3407	20.69	886	21.59	844	17.77	98.4	25.47	229	23.40	1158	24.56	136	26.60
50	1424	29.15	1180	29.26	1038	29.36	1061	25.80	117	0.0	103	32.49	428	32.74	164	29.36
51	2273	19.87	4874	19.18	1109	20.55	1061	17.53	131	25.22	304	23.74	1622	24.49	203	24.50
52	1578	29.85	1533	29.53	1177	28.55	1194	26.23	78.5	32.47	68	33.07	375	33.23	180	28.76
53	2079	24.58	3584	23.83	1012	25.31	989	22.18	126	28.47	245	27.35	1085	28.07	205	25.68
54	2056	18.53	4769	17.61	999	19.03	970	13.97	114	21.40	293	19.32	1673	21.27	179	26.07
55	1247	32.74	205	43.00	952	32.20	952	28.85	5.59	38.65	106	32.45	352	33.31	162	28.61
56	1223	31.24	954	31.62	864	31.17	898	28.28	89.2	32.17	98.4	31.92	415	32.58	259	27.94
57	1576	22.46	3061	22.06	935	23.09	929	20.75	110	26.41	236	26.09	1234	25.59	177	25.36
58	1.47	0.0	3.88	0.0	-5.46	0.0	532	37.91	125	11.22	1.99	0.0	8.21	0.0	168	29.71
59	113	0.0	2.35	0.0	0.60	0.0	-0.05	0.0	-0.03	41.94	6.38	41.66	102	36.37	525	24.55
60	2057	19.60	4748	18.77	1031	20.13	990	16.16	108	22.27	291	21.04	1647	21.68	448	24.04
61	1234	31.47	1037	31.38	0.02	0.0	-0.20	0.0	8.68	36.69	31.1	33.21	283	33.79	338	23.38
62	1848	29.80	155	0.0	1072	27.24	1069	24.79	124	31.58	269	28.12	918	29.34	140	29.88
63	165	0.0	9.70	0.0	354	37.98	552	35.13	-1.04	35.88	0.26	0.0	79.9	36.84	345	26.30
64	1175	28.87	2106	29.46	903	28.60	897	25.52	112	30.17	172	30.25	649	30.84	193	28.66
65	1222	26.04	3071	25.64	909	26.12	884	21.19	132	28.60	242	26.20	1064	27.16	208	26.26
66	1.27	0.0	-1.91	0.0	391	39.92	410	38.50	70.4	38.75	1.40	0.0	100	36.13	555	24.84
67	1448	29.00	1455	31.79	1171	27.11	1184	24.34	161	29.77	269	29.07	944	29.59	137	30.23
68	1.88	0.0	1.25	0.0	-9.08	0.0	794	36.21	0.85	10.63	-0.07	0.0	13.7	0.0	359	26.83
69	1649	28.15	1648	30.69	1084	26.95	1097	23.82	141	31.28	290	26.77	975	28.87	203	29.38
70	1714	25.08	3728	24.14	1003	25.19	967	21.36	126	26.20	275	25.47	1253	26.61	428	25.03
71	1488	29.42	1023	31.56	1048	28.20	1081	25.47	129	31.23	259	29.13	846	30.22	186	29.43
72	1117	31.73	1328	31.43	979	31.48	959	28.10	76.6	13.80	113	31.88	418	32.59	482	25.62
73	1257	29.34	1580	31.15	865	27.88	903	25.70	96.5	30.24	235	27.98	814	29.70	134	28.76
74	1703	18.01	4162	17.34	889	18.35	867	14.89	-0.16	22.23	302	19.34	1713	20.36	265	26.19
75	1919	24.89	3410	24.50	981	25.29	984	22.45	145	28.07	262	27.77	1230	27.64	223	26.46
76	206	43.13	4.88	0.0	451	39.19	895	34.19	5.77	41.06	0.21	0.0	69.6	37.30	461	25.89
77	1763	22.26	3425	22.08	1004	25.13	977	20.60	0.76	39.42	132	30.85	588	31.75	165	26.69
78	0.74	0.0	-1.41	0.0	335	38.30	399	37.78	-1.05	35.89	-0.24	0.0	74.3	37.18	628	23.74
79	5.16	0.0	0.10	0.0	345	39.81	424	37.60	0.13	8.79	-1.21	0.0	107	35.75	379	26.48
80	1926	22.17	4031	21.43	982	24.35	950	19.45	121	26.78	250	23.84	1212	26.20	254	23.68
81	1509	22.40	2956	22.04	877	22.55	902	18.58	115	28.04	229	25.31	1005	26.43	137	29.04
82	0.34	0.0	-0.21	0.0	0.80	0.0	-1.15	0.0	3.92	42.57	0.63	0.0	8.32	0.0	108	31.93
83	2059	23.44	3822	23.04	1018	23.93	977	21.28	122	25.94	283	24.94	1308	26.49	221	24.78
84	1941	25.07	3380	24.74	1048	25.46	1013	23.39	92.7	31.64	128	30.89	813	30.24	149	27.59
85	311	40.46	3.26	0.0	-1.36	0.0	1.82	0.0	-2.30	30.79	5.54	0.0	84.3	36.83	190	28.56
86	1946	20.33	4277	19.76	1159	20.72	1148	17.45	129	28.70	265	26.38	1157	27.08	162	28.15
87	1656	29.10	1957	28.96	1029	29.23	1033	26.21	100	31.98	152	30.38	690	31.03	190	27.79
88	1592	28.04	2383	27.52	1112	27.77	1168	23.48	103	31.29	206	36.98	821	29.78	165	27.71
89	578	37.92	30.0	0.0	419	35.54	654	32.76	2.69	40.54	17.9	0.0	174	34.53	144	28.73
90	126	0.0	0.57	0.0	0.42	0.0	99.3	8.67	99.9	38.91	0.91	27.90	150	34.75	219	28.84
91	1493	29.88	947.2	32.45	956	27.91	970	24.43	112	31.69	255	27.20	828	29.62	215	29.06
92	2012	23.87	3812	23.39	966	23.65	964	21.75	122	27.62	245	30.20	1136	27.75	158	28.27
93	1777	27.81	2716	27.17	1038	27.45	1028	24.08	94.6	31.83	181	23.28	725	30.59	132	29.46
94	2102	21.47	4567	20.93	997	21.42	990	18.67	110	25.39	280	36.98	1390	24.98	133	27.40
95	2.69	0.0	0.40	0.0	2.23	0.0	0.60	0.0	5.52	38.65	4.61	0.0	95.9	35.96	225	30.1
96	1973	24.29	3786	23.80	750	30.65	561	28.08	6.58	38.17	-0.80	27.90	114.8	35.85	44.3	36.84

testing strategies and elevating diagnostic efficacy (Fig. 2). The bar chart provided a clear overview of gene detection across various Ct value ranges for different genes tested using multiple kits (Kit A, Kit B) and a distinct approach. It vividly showed the frequency of gene detection within each Ct range, presenting a color-coded comparison across kits and genes. One striking observation was the consistency in gene detection within the 20.1–30 Ct range, which emerged as the most reliable across all kits and genes. This range consistently exhibited the highest frequency of gene detection, implying a robust and dependable identification within this specific Ct interval. Moreover, the chart highlighted the disparities in detection levels among different genes and kits. It underscored the importance of considering both the gene being targeted and the testing methodology, as they influence the detection efficacy within specific Ct value intervals (Fig. 1).

The Z-values reflect the deviation from the mean and indicate the sensitivity of each target in detection. Lower Z-values suggest a higher abundance or greater ease of detection. Notably, the E gene Ct stands out with the lowest Z-value of −1.49, followed closely by Kit B N Ct with a Z-value of −1.46. These values imply a potentially heightened sensitivity for detecting COVID-19 viral RNA in these specific gene regions. Moreover, all recorded *P*-values are extremely low (0.00001), signaling strong statistical significance. This suggests substantial evidence against the null hypothesis and underscores the reliability of the observed differences in sensitivity among gene targets.

The practical implications of these findings extend to the clinical realm of COVID-19 diagnostics. The superiority of the current approach suggested a promising avenue for enhancing accuracy in detecting COVID-19 cases, particularly emphasizing the significance of incorporating the *E* gene in diagnostic methodologies. The insights gained from this study advocate for the critical consideration of gene combinations in optimizing diagnostic accuracy, potentially influencing decision-making in clinical settings.

Despite the significance of these findings, there are inherent limitations that warrant consideration. Factors such as sample size, variations in testing methodologies, or specific population characteristics might have influenced the observed results. Therefore, future research endeavors should focus on expanding sample sizes, validating findings in diverse populations, as well as exploring additional gene combinations to further solidify these findings and translate them into clinically applicable solutions.

In conclusion, the assessment of AUC values among different testing kits and gene combinations elucidated a hierarchy of effectiveness in COVID-19 diagnostics. This approach, particularly with the inclusion of the *E* gene, emerges as a highly promising diagnostic method with enhanced accuracy, laying a foundation for improved COVID-19 testing methodologies and potentially contributing to more effective disease management strategies. The most dependable detection occurred within the 20.1–30 Ct range across various gene combinations and testing kits.

5. ACKNOWLEDGMENTS

We would like to thank the Kurdistan Regional Government, the Ministry of Higher Education and Scientific Research as well as the Presidency of Sulaimani University.

6. ETHICS APPROVAL

The study was approved by the Ethics Committee of the Biology Department, College of Science, University of Sulaimai (protocol code UoS-Sci-Bio 0010/ September/19/2023

7. CONFLICTS OF INTERESTS

The authors declared that they have no competing interests.

8. AUTHOR CONTRIBUTIONS

Conceptualization, S.A; methodology, S.A. and D.A; software, Z.H.; validation, S.A., D.A., and Z.H.; formal analysis, D.A. and Z.H; investigation, D.A.; resources, D.A and Z.H.; data curation, S.A.; writing—original draft preparation, D.A.; writing—review and editing, S.A and Z.H.; visualization, D.A.; supervision, S.A. All authors have read and agreed to the published version of the manuscript.

9. FUNDING

This research received no external funding.

REFERENCES

[1] Y. Chen, Q. Liu and D. Guo. "Emerging coronaviruses: Genome structure, replication, and pathogenesis". *Journal of Medical Virology*, vol. 92, no. 4, pp. 418-423, 2020.

[2] R. A. Khailany, M. Safdar and M. Ozaslan. "Genomic characterization of a novel SARS-CoV-2". *Gene Reports*, vol. 19, p. 100682, 2020.

[3] M. J. Loeffelholz and Y. W. Tang. "Laboratory diagnosis of emerging human coronavirus infections-the state of the art". *Emerging Microbes and Infections*, vol. 9, no. 1, pp. 747-756, 2020.

[4] A. E. Gorbalenya and S. C. Baker. "The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2". *Nature Microbiology*, vol. 5, no. 4, pp. 536-544, 2020.

[5] C. Sohrabi, Z. Alsafi, N. O'Neill, M. Khan, A. Kerwan, A. Al-Jabir, C. Iosifidis and R. Agha. "World health organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)". *International Journal of Surgery*, vol. 76, no. 1, pp. 71-76, 2020.

[6] S. U. Rehman, L. Shafique, A. Ihsan and Q. Liu. "Evolutionary trajectory for the emergence of novel coronavirus SARS-CoV-2". *Pathogens*, vol. 9, no. 3, p. 240, 2020.

[7] L. T. Phan, T. V. Nguyen, Q. C. Luong, T. V. Nguyen, H. T. Nguyen, H. Q. Le, T. T. Nguyen, T. M. Cao and Q. D. Pham. "Importation and human-to-human transmission of a novel coronavirus in Vietnam". *New England Journal of Medicine*, vol. 382, no. 9, pp. 872-874, 2020.

[8] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang., & N. Zhu. "Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding". *The Lancet*, vol. 395, no. 10224, pp. 565-574, 2020.

[9] A. Banko, G. Petrovic, D. Miljanovic, A. Loncar, M. Vukcevic, D. Despot and A. Cirkovic. "Comparison and sensitivity evaluation of three different commercial real-time quantitative PCR Kits for SARS-CoV-2 detection". *Viruses*, vol. 13, no. 7, p. 1321, 2021.

[10] Y. Lu, L. Li, S. Ren, X. Liu, L. Zhang, W. Li and H. Yu. "Comparison of the diagnostic efficacy between two PCR test kits for SARS-CoV-2 nucleic acid detection". *Journal of Clinical Laboratory Analysis*, vol. 34, no. 10, p. e23554, 2020.

[11] V. M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. Chu and T. Bleicker. "Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR". *Eurosurveillance*, vol. 25, no. 3, p. 2000045, 2020.

[12] M. Y. Wang, R. Zhao, L. J. Gao, X. F. Gao, D. P. Wang and J. M. Cao. "SARS-CoV-2: Structure, biology, and structure-based therapeutics development". *Frontiers in Cellular and Infection Microbiology*, vol. 10, pp. 587269, 2020.

[13] I. Barjaktarović, M. Y. Wang, R. Zhao, L. J. Gao, X. F. Gao, D. P. Wang and J. M. Cao. "Diagnosing COVID-19: Diagnostic importance of detecting E gene of the SARS-CoV-2 genome". *Future Virology*, vol. 18, no. 1, pp. 31-38, 2023.

[14] P. B. Van Kasteren, B. Van der Veer, S. Van Den Brink and L. Wijsman. "Comparison of commercial RT-PCR diagnostic kits for COVID-19". *Journal of Clinical Virology*, vol. 128. p. 104412.

[15] Z. Iglói, M. Leven, Z. A. K. Abou-Nouar and B. Weller B. "Comparison of commercial realtime reverse transcription PCR assays for the detection of SARS-CoV-2". *Journal of Clinical Virology*, vol. 129, p. 104510, 2020.

[16] A. Tastanova, C. I. Stoffel, A. Dzung, P. F. Cheng, E. Bellini, P. Johansen, A. Duda, S. Nobbe, R. Lienhard, P. P. Bosshard and M. P. Levesque. "A comparative study of real-time RT-PCR-based SARS-CoV-2 detection methods and its application to human-derived and surface swabbed material". *The Journal of Molecular Diagnostics*, vol. 23, no. 7, pp. 796-804, 2021.

[17] J. Zhao, Y. Zhang and B. Cao. "Comparison of the performance of Six SARS-CoV-2 nucleic acid detection kit in positive samples using RT-PCR". *Journal of Clinical Chemistry and Laboratory Medicine*, vol. 4, no. 3, p. 170, 2021.

[18] U. Eberle, C. Wimmer, I. Huber, A. Neubauer-Juric, G. Valenza and N. Ackermann, A. Sing and Bavarian SARS-CoV-2-Public Health Laboratory Team. "Comparison of nine different commercially available molecular assays for detection of SARS-CoV-2 RNA". *European Journal of Clinical Microbiology and Infectious Diseases*, vol. 40, pp.1303-1308, 2021.

[19] D. Kalita and S. Deka. "Effectiveness of different gene-target strategies for SARS-CoV-2 screening by RT-PCR and other modalities". *Journal of Medical Diagnostic Methods*, vol. 9, no. 4, p. 298, 2020.

[20] C. Hernández, C. Florez, S. Castañeda, N. Ballesteros, D. Martínez, A. Castillo, M. Muñoz, S. Gomez, A. Rico, L. Pardo, A. Paniz-Mondolfi and J. D. Ramírez. "Evaluation of the diagnostic performance of nine commercial RT-PCR kits for the detection of SARS-CoV-2 in Colombia". *Journal of Medical Virology*, vol. 93, no. 9, pp. 5618-5622, 2021.