Original Article

Association of N and E Genes with ORF1b Gene for Early Detection of COVID-19 in Pakistani Patients

Yar Muhammad Waryah[1*], Shaista Ehsan[2], Roohi Nigar[3]

ABSTRACT

OBJECTIVE: To investigate the role of the first line and additional confirmatory assay to detect coronavirus based on Ct RT-PCR values.

METHODOLOGY: A collaborative cross-sectional study was conducted using a systematic sampling type at Hyderabad's diagnostic and research laboratory and the SIOVS, Hyderabad, after getting ethical approval from the LUMHS ERC committee from July 2021 to June 2022. A total of 4822 patients have enrolled; 1200 patients were positive for COVID-19, and a nasopharyngeal swab was collected from patients for RNA extraction; the patients positive with COVID-19 RT-PCR test were included, and patients negative with COVID-19 RT-PCR were excluded out from this study. The RT-PCR was performed, and the Ct values of each potential target were recorded.

RESULTS: The RT-PCR finding of potential targets revealed that the N and E genes are correlated and simultaneously positive with OFRb1 and RdRP genes and are crucial for detecting coronavirus. In 77.5% (310/400) of patients, the E gene was positive, and the N gene was positive in 85.8% (1030/1200) patients with ORF1b and Rdrp genes. 71% (850/1200) of males were infected in our cohort, and young patients carried the highest viral load than old age patients with asymptomatic features. Each target's obtained cycler threshold values by different kits were recorded and analyzed in SPSS-17.

CONCLUSION: The data showed N and E genes are simultaneously associated with the ORFb1 gene for detecting COVID-19 in our patients.

KEY WORDS: SARS CoV-2, COVID-19, E gene, N gene, ORF1b gene, RT-PCR

INTRODUCTION

The novel coronavirus emerged from animals to humans around December 2019 in Wuhan city of Hubei province of China and spread worldwide[1,2]. Due to the rapid worldwide transmission of coronavirus from one person to another and increased infection rate[2], the WHO declared novel coronavirus (COVID-19) to the pandemic (WHO, 2020)[3,4]. Almost 6.78 billion people have been infected with this virus, and more than 6.7 million deaths have been recorded worldwide; the ratio of recovered patients is 6.5 billion. In the past two decades, a family of coronaviruses caused three epidemic diseases, including COVID-19, SARS, and MERS. These are all infections caused by members of a group of viruses called coronaviruses and manifesting similar common symptoms of cold fever, fatigue and loss of test in humans[5,6].

Coronaviruses are RNA-positive sense single-stranded RNA viruses; they are divided into four genera (α.β.γ.δ). Human coronavirus (HCoVs) belong to α (HCoV-229E, NL63) and β (MERS-CoV and SARS-CoV) genera and these viruses carry similarities to each other among all known viruses; coronaviruses possess long genome[7]. China sequenced the RNA genome and the genome result of n.covid-19-positive China patients revealed a previously unknown CoV strain in patients[8]. The 88% of sequenced 2019-nCoV showed the genome pattern of SARS, and 55% of the genome consists of an identical way of MERS-like coronaviruses. The genome of SARS-CoV-2 encode four different conserved functional protein, including spike (S), envelop (E), nucleocapsid (N) and membrane (M) proteins. A molecular method was developed to detect 2019-nCoV through RT-PCR[9]. The specific targets of COVID-19 were selected by using different bioinformatics tools. Other available commercialized kits are used worldwide for detecting COVID-19. WHO recommends the ORF1b, RdRp, E and N genes for screening of n.coivd-19 by using COVID-19 RT-PCR kits, and these genes are reported as potential targets for the detection of coronavirus[10].

This study describes the contribution of the E and N genes to the early detection of coronavirus with ORF1b and RdRp genes through RT-PCR findings of Pakistani COVID-19 patients. This study may help detect COVID-19 based on additional and second lines of the genes.

*[1]*Department of Molecular Biology & Genetics, Sindh Institute of Ophthalmology & Visual Sciences, Hyderabad*
[2]Department of Pediatrics, Zaiuddin University Clifton Karachi, Sindh-Pakistan.
[3]Senior Registrar, Bilawal Medical Collage, Kotri, Jamshoro, Sindh-Pakistan.
***Correspondence:** yarmwaryah@hotmail.com*
doi: 10.22442/jlumhs.2023.01021
Received: 20-02-2023
Accepted: 27-02-2023 Published Online: 28-02-2023

METHODOLOGY

A collaborative cross-sectional study was conducted using systematic sampling at Hyderabad's diagnostic and research laboratory and the Department of Molecular Biology & Genetics, SIOVS, Hyderabad. After getting the approval of the ethical review committee of LUMHS from July 2021 to June 2022 and confirmation of COVID-positive patients, informed consent was taken from patients, and data were recorded based on RT-PCR results. The obtained cycler threshold values of each target by different kits were recorded and analyzed in SPSS-17, and frequencies/percentages and mean ± SD of categorical were recorded. A total of 4822 patients were enrolled in this study from different parts of Pakistan; 1200 patients tested positive for COVID-19; most patients belonged to the province of interior Sindh. Nasopharyngeal swabs of patients were collected and transported into UTM/VTM tubes for RNA extraction.

Real-time PCR assay for COVID-19

Three commercialized kits were used for RNA extraction and virus detection using auto- RNA/DNA extractor and RT-PCR. The RT-PCR is the most advanced technology for the amplification and detection of specific targets by using primers and probes. RT-PCR works in three steps: in the denaturation step, the cycler temperature becomes higher than the melting point temperature, the RNA template converts into complementary DNA (cDNA), and cDNA is used as a template for the amplification of specific targets, the second step targeted primers anneal to their target at lower temperature and starts the synthesis of complementary DNA and third step

The Abbott RT auto extractor system and Abbott real-time PCR system were used to extract and amplify the desired RNA virus. The Abbott SARS-CoV2 kit was used for RNA extraction and amplification of SARS-CoV-2 targets. Three targets were detected through two different fluorescent probes. RdRp and N-gene targets were detected on FAM-fluorophore dye, and internal control was seen on VIC fluorophore dye. The Ct value <40 was recorded as positive, and >40 was considered negative.

The Sansure novel coronavirus (2019-nCoV) nucleic acid diagnostic kit was used to extract and amplify three targets; each target was detected through three different fluorescent probes. FAM probe for ORF1 gene, ROX probe for N- gene and internal control VIC probe were used.

The maccura SARS-CoV-2 Fluorescent PCR kit detected and amplified ORF1b, E and N genes. Zybio nucleic acid extraction kit was used for RNA extraction through an automated RNA/DNA extractor. Reaction system and amplification conditions were performed according to the manufacturer's specifications (Shanghai BioGerm Medical Technology Co. LTD, China).

RESULTS

In the present study, four different conserved targets of the novel SARS-COV2 virus were screened on the recommendation of the World Health Organization to detect 2019-nCoV. E gene is used for the first line of screening, whereas ORF1b and RdRP genes were used as a confirmatory assay, and the N gene was used for an additional confirmatory assay. Most of the patients were asymptomatic, and only a few were symptomatic. The travelling history of patients was recorded; only 1% (10/1200) of patients had a travelling history during this outbreak, and they don't have any symptoms associated with n.coivd-19. The retrospective data of performed cases showed that 24% (1200/4822) cases were positive with 2019-nCoV, and Ct values of all potential targets were analyzed according to age and gender. Viral load distribution within the patients was also noted.

In this study, we analyzed the data according to the available panel of genes in commercial kits. The probe of the E gene was known only by the **Maccura kit**. The E gene was detected in 77.5% (310/400) of patients and undetected in 22.5% (90/400) of patients; 8.7% of patients carried the highest viral load between a range of Ct 15-20, while 40.2% of patients lies between the range of moderate viral load Ct 21-30 and 28.5% having lowest viral load between Ct of 31-40 (**Table I)**. The young patient between the ages of (21-40) was positive within a moderate range of Ct values, and the viral load of the E gene decreased in old patients. Primarily, young patients carried the highest viral load compared to the old (**Table II)**.

N gene is used as an additional confirmation assay; the probe of this assay was available in three commercial kits that have been used (Sansure, Maccura and Abbot). This study indicates that the N-gene was positive in 85.8% (1030/1200) patients and negative in 14.1% (170/1200) patients. 20.25% (243/1200) of patients carried the highest viral load between Ct values 15-20, 4.25% (543/1200) patients had a moderate viral load between ranges of Ct values 21-30, and the lowest viral load was recorded in 20% (240/1200) of patients. Only 9.8% (118/1200) of patients observed that the N gene was negative and the ORF1b gene was positive.

ORF1b gene is used as a specific target for (2019-nCoV); detection of this assay was available in (Sansure and maccura) kits. In 93% (746/800) of patients, the ORF1b gene was positive, and only 7% (54/800) were negative. The highest viral load of the ORF1b gene was recorded in 13.5% (108/800) of patients in the range of Ct values 15-20, and 46% (369/800) carried the moderate viral load in the range of Ct values 21-30, and 34% (269/800) patients carried the lowest load in the range of Ct values 31-40 (**Table I)**.

The panel of RdRp gene was available in the Abbott SARS-CoV2 kit; the Rdrp gene was detected in 100% (400/400), the viral load was highest in 40.2%

(161/400) of patients with Ct values 02-20, moderate range of viral load seems in 59.3% (237/400) and lowest viral load recorded in 0.5% (2/400) patients. The distribution of viral load according to Ct values and percentages is described in (**Table I)**, the age-wise distribution of Ct value is defined in (**Table II)**, and the Demographic distribution of Ct values in patients is illustrated in the bar graph (**Figure I)**.

prime (orf1/ab)-structural proteins, Spike (S), Envelope (E), Membrane (M) and Nucleocapsid (N)] −3′ prime. The different RNA viruses affect human health in epidemic transmission, and identifying causative disease agents could be necessary to prevent disease[12]. The present study showed that all these four potential targets are crucial for detecting 2019-nCoV.

Figure I: Gender-wise distribution of threshold cycle (CT) in patients

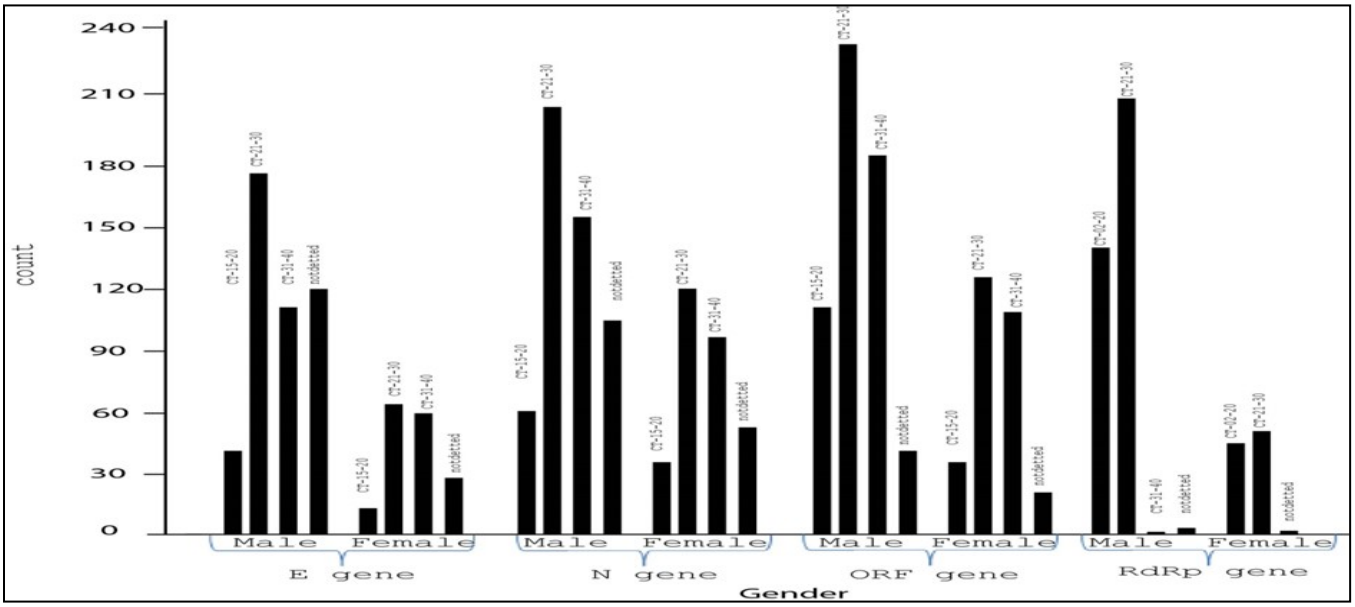

Table I: Percentage-wise distribution of CT values versus covid-19 genes in patients

		Cycler threshold values				
Kits	**Genes**	**0-20 Cycler threshold**	**21-30 Cycler threshold**	**31-40 Cycler threshold**	**Not detected**	**Total**
Group-1 Sansure-kit	N. Gene CT	13%(52/400)	52.7% (211/400)	27.2% (109/400)	7% (28/400)	400
	ORF. Gene CT	10.5%(42/400)	51% (204/400)	38.5% (154/400)	0.00	
Group-2 Maccura-kit	E. Gene CT	8.7%(35/400)	40.2% (161/400)	28.5% (114/400)	22.5% (90/400)	400
	N. Gene CT	7.5%(30/400)	23.7% (95/400)	32.2% (129/400)	36.5% (146/400)	
	ORF. Gene CT	16.5%(66/400)	41.2% (165/400)	28.7% (115/400)	13.5% (54/400)	
Group-3 Abbott-kit	N.Gene/ RdRp. Gene CT	40.2%(161/400)	59.3% (237/400)	0.5% (2/400)	0.00	400
					Grand Total	**1200**

Table II: Percentage-wise distribution of CT (threshold cycle) values versus age of covid-19 patients

Age group	**E. gene- CT %400**			**N. gene- CT %1200**			**ORF1b-gene- CT %1200**		
	0-20	**21-30**	**31-40**	**0-20**	**21-30**	**31-40**	**0-20**	**21-30**	**31-40**
0-20 Years	1.2%	5%	4.7%	7%	9%	6%	4%	6.3%	4.5%
21-40 Years	4.5%	20.5%	14.7%	12.5%	23%	24%	8.8%	25.3%	14.2%
41-60 Years	3%	13.5%	8.5%	0.3%	8%	6.6%	4.3%	15.7%	9.7%
61-above	0	1.2%	0.5%	0.6%	1.6%	1.5%	0.3%	3.6%	2.3%

DISCUSSION

The novel coronavirus 2019 is a single-strand RNA virus, the genome of (2019-nCoV) consists of 29891 nucleotides, and it encodes the 9860 amino acids[11]. The 38% of GC content was similar to other βCoVs, and the genome is arranged in order of 5′-3′ prime, 5′-

We investigated the association of different SARS-CoV2 targets for screening coronavirus in our population. It has been observed that few patients 4.5% (54/1200) were negative on specific targets of ORF1b and RdRp but positive with N and E genes assay with the moderate viral load of Ct value 30-40,

and patients had intense symptoms of 2019-nCoV including fever, fatigue and loss of test. Some of them were admitted to ICU and recovered in time, and some could not be recovered. Only the E gene was positive in 8.2% (33/400) of patients, while N and ORF1b targets were negative in patients. Only 2.6% (21/800) patients were positive with the N gene, and the specific target was undetected. This study supports previously reported that the N gene is essential for screening 2019-nCoV and equally contributes to the confirmation assay and also observed that the E gene plays a vital role in initial virus detection[13]. The majority of screened patients were positive for E, N, ORF1b and Rdrp genes; it was observed that all the genes are the main components for screening. The severity of the disease in the Pakistani population is lower as compared to other people and recorded low death ratio, and 50% (602/1200) of the COVID-19 patients in our cohort are positive with moderate Ct values 21-30, primarily young patients infected more than old age patients.

CONCLUSION

This study describes the contribution of the E and N genes to the early detection of coronavirus with ORF1b and RdRp genes through RT-PCR findings of Pakistani COVID-19 patients. This study is based on Pakistani ethnicity and genetic pattern. The result showed that additional and second line of targets (E and N) are simultaneously positive with ORF1b and Rdrp genes; both E and N genes are the main components of early detection of COVID-19. This study may help to improve laboratory quality and generate reliable conclusive reports in those cases where only additional and second lines of the targets were detected in RT-PCR.

ACKNOWLEDGEMENTS

The authors thank all technical and non-technical staff of the diagnostic and research laboratory, Hyderabad, for their support throughout the study.

Ethical permission: Liaquat University of Medical & Health Sciences Jamshoro ERC Letter No. LUMHS/ REC/Ch/031.

Conflict of Interest: No conflicts of interest, as stated by authors.

Financial Disclosure / Grant Approval: No funding agency was involved in this research.

Data Sharing Statement: The corresponding author can provide the data proving the findings of this study on request. Privacy or ethical restrictions bound us from sharing the data publically.

AUTHOR CONTRIBUTIONS

Waryah YM: Study design, collection of data and drafting
Ehsan S: Data analysis and drafting a critical review
Nigar R: Help in a study design and critical review

REFERENCES

1. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020; 579(7798): 270-3. Epub 2020/ 02/06.
2. Jiatong S, Wenjun L. Epidemiological characteristics and prevention and control measures of Corona Virus Disease 2019 in children. J Trop Med. 2020; 20(2): 153-6.
3. Organization WH. WHO characterizes COVID-19 as a pandemic [EB/OL]. 2020.
4. Rodriguez-Morales AJ, Bonilla-Aldana DK, Balbin-Ramon GJ, Rabaan AA, Sah R, Paniz-Mondolfi A, et al. History is repeating itself: Probable zoonotic spillover as the cause of the 2019 novel Coronavirus Epidemic. Le infezioni in medicina. 2020; 28(1): 3-5. Epub 2020/02/06.
5. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet. 2020; 395(10225): 689-97. Epub 2020/02/06.
6. Dhama K, Khan S, Tiwari R, Sircar S, Bhat S, Malik YS et al. Coronavirus Disease 2019-COVID-19. Clin Microbiol Rev. 2020; 33(4). Epub 2020/06/26.
7. Woo PC, Huang Y, Lau SK, Yuen KY. Coronavirus genomics and bioinformatics analysis. Viruses. 2010; 2(8): 1804-20. Epub 2010/08/01.
8. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H et al. Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet (Londs, Engl). 2020; 395(10224): 565-74. Epub 2020/02/03.
9. 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(6). Epub 2020/05/30.
10. Ward S, Lindsley A, Courter J, Assa'ad A. Clinical testing for COVID-19. J Allergy Clin Immun. 2020; 146(1): 23-34. Epub 2020/05/24.
11. Li X, Geng M, Peng Y, Meng L, Lu S. Molecular immune pathogenesis and diagnosis of COVID-19. J Pharmaceut Analysis. 2020; 10(2): 102-8. Epub 2020/04/14.
12. Kakhki RK, Kakhki MK, Neshani A. COVID-19 target: A specific target for novel coronavirus detection. Gene Reports. 2020; 20: 100740. Epub 2020/06/09.
13. Chu DKW, Pan Y, Cheng SMS, Hui KPY, Krishnan P, Liu Y et al. Molecular Diagnosis of a Novel Coronavirus (2019-nCoV) Causing an Outbreak of Pneumonia. Clin Chem. 2020; 66(4): 549-55. Epub 2020/02/08.