Sequences of S-surface of human COVID-19

Dr. Fadyia Mahdi Muslim Alameedy
Faculty of Science, Kufa University, Najaf, Iraq
Corresponding Author:

.

ABSTRACT

Samples were collected from patients infected with the Coronavirus, according to official approvals, for the purpose of diagnosing the virus through qRT- PCR. Primers were designed based on the (NCBI) and primers of research Myungsun et al., 2020, where the two-step method was adopted, and then a s-spike of singer sequences was conducted for different samples in terms of age, gender and virus concentration where the selection of samples took place in different months of the year, where samples were taken from the April ,September and November months of 2020 year as well as the January, February, April ,June and July months of 2021 year . The number of samples that was carried out for the singer sequences about 10 samples of both selected for study . In addition, viruses that share with the including (Coronavirus, Influenza virus, Parainfluenzavirus, Rhinovirus and Metapneumovirus) were diagnosed, which were isolated from 300 human infected cases and mix cases in same sample of infected pateint.

SAMPLE ABSTRACT:
Aims: Diagnosing the Coronavirus and viruses that share infection for different age groups through real time PCR, as well as using the Sanger sequencing method to homology sequence identity between local human Coronavirus isolate and NCBI-BLAST submitted Coronavirus isolate.
Study design: Cross section.
Place and Duration of Study: From the central health laboratory and quarantine centers, work was carried out in different places for the purpose of diagnosis in Al-Hakim and Al-Sadr Teaching Hospital and the Central Health Center in addition to the private laboratories, between April 2020 and July 2021.
Methodology: Diagnosing the virus through qRT- PCR, PCR and sangersequencing. Sample: We included 300 patients (174men, 128 women; age range 18- 83years) with from people who were infected for the first time with the Coronavirus and repeated infections with the virus with other viruses that affect the respiratory system.
Results: The current study showed that the number of cases of Coronavirus infection (110) case, Influenza virus (90) case, Parainfluenzavirus (65) case, Metapneumovirus (108)case and Rhinovirus (95)case for the period from 4-4-2020 up to 26-26 7-2021 all viruses were diagnosed through qReal time PCR technique by designing primers according to the location NCBIRegarding the sequences test, the results showed the percentage of similarity with the studied strains at a rate ranging between (99.92 - 78 %).
Conclusion: Most of the infections with the Coronavirus are common with respiratory viruses in different proportions, in addition to the fact that the age group (51-61) year is more infected, and (110) case of Influenza shares with the Coronavirus more infection and the results of Sanger sequences between(99.92 - 78 %).

Keywords: Middle East Respiratory Syndrome (MERS) , Severe Acute Respiratory Syndrome (SARS) , COVID-19

1. INTRODUCTION

In December of 2019, a novel strain of coronaviruses, SARS-CoV-2, was revealed prominent to an occurrence of transmittable illness affecting a global pandemic. Coronaviruses were a great family of viruses well-known for contaminate together persons and different organisms. The persons, SARS-CoV-2 coronavirus contagions can source several infections beginning the corporate cold to other acute infections such as (middle east and severe acute respiratory Syndrome) [1].

A healthy cell is infected by the attachment of coronaviruses to a healthy target cell by binding to special receptors located on the membrane of an uninfected cell. The glycoprotein Spike (S) is the single membrane that helps bind to the target cells, which can fuse with the viral cells[2].

Through studies about appearance of new variants of the virus and the presence of the SARS-CoV-2 Spike (S) protein gene that contains one of the main mutations that have an effect on the function of the protein and its ability to infect uninfected cells [3].

In calculation to the variation that causes the mutation in the Spike (S) protein gene, added mutations have currently been discovered . Where all parts of the genome represent the genes that contain mutations, which are important in the preparation of the vaccine as well as the therapeutic research of the virus[4].

By knowing the next generation sequence of the SARS-CoV-2 viral genome, we can identify the variants that occur virus disaster and based on protocol for early ARTIC SARS-CoV-2 sequencing in January 2020 which was approved by several scientific organizations around the world and later published the original protocol in September 2020 which was considered the most common for sequencing SARS-CoV-2[4].

2. MATERIAL AND METHODS / EXPERIMENTAL DETAILS / METHODOLOGY

2.1 Quantitve Real Time PCR Technique: This technique was relied on for the diagnosis of viral infection, and an extraction was used from (CanvaxHigherPurity™ Viral DNA/RNA Extraction Kit AN0605.UK.) .The primers designed according to NCBI by us for the purpose of diagnosing viruses including (Coronavirus, Influenza virus, Parainfluenzavirus, Rhinovirus and Metapneumovirus)are shown in the table (1).

Table (1): Primes design of HHV type 6&8 as well autism spectrum disordersdepending on the NCBI&Myungsun*et al.*,2020.

Type/subtype	Name	Sequences	Bases	PCR product size
S-Covid-19	Primer F	CAAATCGCTCCAGGGCAAAC	20bp	516 bp
	Primer R	CTGTGGATCACGGACAGCAT	20bp	
	Primer F	ACTGTTTTGCCACCTTTGCT	20bp	

S-SARS- CoV-2. IBS_m_S 2 . Myungsun*et al.,*2020	Primer R	AGCTTGTGCATTTTGGTTGA	20bp	300 bp
	Primer F	TTGCTAAAACCCGGAGACAC	20bp	
HA-*Influenza virus*	Primer R	CCTGACGTATTTTGGGCACT	20bp	228 bp
HPIV3gp4 *Parainfluenzavirus* type 3	Primer F	TGCCACCATCTATCAACCAA	20bp	250 bp
	Primer R	CGTGTTCTGGGTTCCATTTT	20bp	
HRV89gp1 *Rhinovirus* type A	Primer F	GCAATGCTAAGTGCTGTCCA	20bp	185 bp
	Primer R	AGGTGGAGGAGATTGGAGGT	20bp	
G - *Metapneumovirus*	Primer F	AGCTCATCACCCATGGAATC	20bp	214 bp
	Primer R	TTGGTGGTGTTGTGTGTGTG	20bp	

2.2 RNA Sequencing Method: Ten samples were selected out of 110 positive for genetic Sequencing. Conventional PCR products of positive S-Covid-19andS-SARS- CoV-2. IBS_m_S2 were sent to Macrogen Company in Korea by for performing the RNA sequencing by (AB RNA sequencing system). The RNA sequencing analysis for S-Covid-19and S-SARS- CoV-2. IBS_mS 2 genotyping. PCR master mix was elaborated by using GoTaq® Green Mater Mix Kit (Promega,LOT.0000401B40,USA),Iqon PCR Ladder Cat. No.: A610641 (dsDNA ladder with bands from 100 bp to 3.000 bp). All PCR products were electrophoresed on agarose gel with ethidium bromide and visualized under UV light.The multiple alignment analysis was based on Clustal alignment analysis, and NCBI-BLAST for the homology sequence identity.

3. RESULTS AND DISCUSSION

The current study showed that the number of cases of Coronavirus infection (110) case, Influenza virus (90) case, Parainfluenzavirus (65) case, Metapneumovirus (108)case and Rhinovirus (95)case for the period from 4-4-2020 up to 26- 7-2021 all viruses were diagnosed through qReal time PCR technique in figure 1 and PCR technique in figure 2 by designing primers according to the location NCBI as shown in table 1. The study included six groups for different age groups (18- 83), where the study showed that the age group from (51-61)year has the highest rate (80)case of infection compared to the rest of the age groups including [(18-28),(29-39),(40-50),(62-72)and (73-83)] year the number of cases of infection was, respectively [(33),(40),(70),(52)and (25)] ,and the number 174 cases of males infected was higher than females 128 cases.Through our study, we noticed a difference in age groups in infection, where the age group (40-50)years was more susceptible to infected. In the study (Russell et al.,2021) In 15 countries, we noticed that the age group (+80)years had the highest infection, while in India the age group (20-29)yearsin

addition to the recovery rates of males were higher than females in India, and the death rate also varied.

In Iraq, through our study, no age group is excluded, because the collection of samples did not include all governorates of Iraq and all cities. Samples were taken according to the consent of the patient or the patient for the purpose of conducting the study.

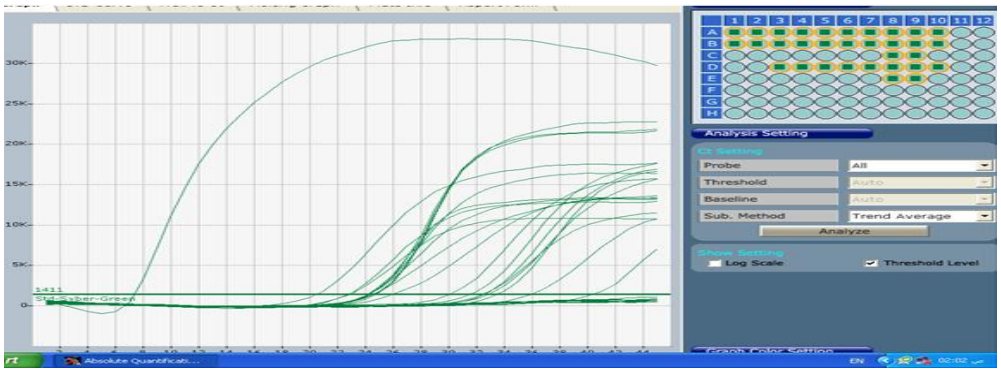

Fig. 1.Disgnosis of all virueses (Coronavirus, Influenza virus, Parainfluenzavirus, Rhinovirus and Metapneumovirus) by qRT-PCR technique

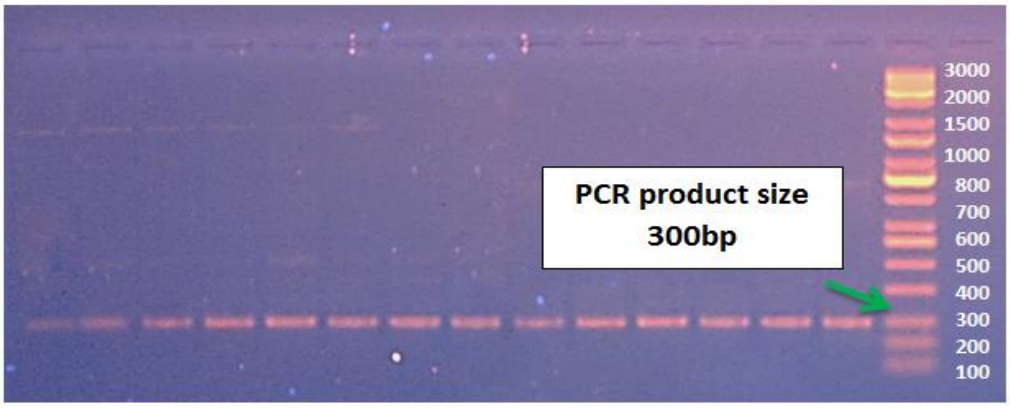

Fig. 2. Diagnosis of Human S-SARS- CoV-2 by PCR (PCR product size 300bp, 5 µl Iqon PCR Ladder was loaded on a 1.5 % agarose in 1x TBE and stained with ethidium bromide. dsDNA ladder with bands from 100 bp to 3.000 bp.

Regarding the sequences test, the results showed the percentage of similarity with the studied strains at a rate ranging between (99.92 - 78 %) as shown in the table (2), ten

samples were selected based on the concentration of the virus cycles , different age groups , gender in addition to the sample source areas at the time of collection.

Sanger sequence analysis of the Coronavirus variants . All 10 samples that amplified with both the S-Covid-19 in 20-11-2020 and S-SARS- CoV-2 IBS_mS 2 primers from 20 -11-2020 up to 10-3-2021 . Table 2 appear the six results were the percentage of NCBI-BLAST homology sequence identity between (99.92 – 95.32 %), while the S-SARS- CoV-2 IBS_mS from (7-10) samples , the percentage ranged between (89.19 – 78 %), and all samples were sent to the NCBI for the purpose of recording it.

Table 1. The NCBI-BLAST homology sequence identity (99.92-78%) between local human Coronavirus isolate and NCBI-BLAST submitted Coronavirus isolate

112

113 According to the official approvals for the purpose of diagnosing the virus through qReal
114 Time PCR , samples were collected from patients infected with the Coronavirus, and then a
115 s-spike singer sequencing was conducted for different samples of age, gender and virus

Local isolate No.	NCBI-BLAST Homology Sequence identity (%)		
	NCBI-BLAST identical Genotypes	Genbank Accession number	Identity (%)
Human Coronavirus isolate No.1	SARS coronavirus isolate CUHKtc53L spike glycoprotein (S) gene, complete cds	DQ412628.1	99.92%
Human Coronavirus isolate No.2	SARS coronavirus ExoN1, complete genome	FJ882930.1	99%
Human Coronavirus isolate No.3	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/NM-CDC-ASC210522675/2021, complete genome	OL946799.1	99%
Human Coronavirus isolate No.4	Select seq NC_000012.12 Homo sapiens chromosome 12, GRCh38.p13 Primary Assembly	NC_000012.12	95.32%
Human Coronavirus isolate No.5	Select seq NC_000013.11 Homo sapiens chromosome 13, GRCh38.p13 Primary Assembly	NC_000013.11	89.19%
Human Coronavirus isolate No.6	Select seq NC_000023.11 Homo sapiens chromosome X, GRCh38.p13 Primary Assembly.	NC_000023.11	89%
Human Coronavirus isolate No.7	Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/MD-MDH-0567/2020, complete	MW484816.1	85%
Human Coronavirus isolate No.8	Select seq NC_000014.9 Homo sapiens chromosome 14, GRCh38.p13 Primary Assembly.	NC_000014.9	80%
Human Coronavirus isolate No.9	Homo sapiens isolate CHM13 chromosome 15	CP068263.2	79%
Human Coronavirus isolate No.10	Select seq NC_000016.10 Homo sapiens chromosome 16, GRCh38.p13 Primary Assembly	NC_000016.10	78%

concentration. The fourth, ninth and eleventh of 2020 year , as well as the month of the first, second, fourth, sixth and seventh of 2021 year, and the analysis of the results was conducted according to the (NCBI), and the results were compared, as well as the percentages of similarity with the strains that appeared according to the four variables so far of the virus. The study was conducted for the first time in Najaf/Iraq.

Where more than 172 countries have shared the genome sequences of the Corona virus, the viral evolutionary geneticist at the Fred Hutchinson Cancer Research Center in Seattle, Washington explained the importance of these sequences as not being transformative on 22 June [6].

During Bloom's study from May 2020 after searching for genetic data for the early stages of the epidemic, the linkage of sequences through the nuclear sequencing technology for the purpose of revealing the genetic material of different samples of infected people, as this study was published in the journal Small in June 2020 [7].

Sanger Sequence analysis of the Human Coronavirus variants. All 10 samples that amplified with both the S-Covid-19 and S-SARS- CoV-2 IBS_mS 2 primers clustered with genotype in table (2). In our study, the Sanger Sequence analysis was used the NCBI-BLAST homology sequence identity (99.92-78%) , while in other studies, whole genomes were used Maria et al ., 2 020 the B1.1 variant was isolated in Europe and is considered to be more dominant, as the sequence showed the presence of a mutation in the spike protein due to a change in the amino acid sequence of SARS-CoV-2 Siena-1/2020 has been placed in GenBank underneath the succession no. MT531537. The rare Nanopore delivers were placed in the sequence beneath BioProject agreement no. PRJNA658490 with no. SRX8982904 [direct RNA sequencing] and SRX8982905 (amplicon sequencing). ; Anna et al ., 2021 the complete genome (100%) of SARS-CoV-2 was positively gained for 21/27 samples. Jonathan et al ., 2021 appear whole analysis (>98%) of the viral genome ; while study Nihad et al ., 2021 IlluminaMiSeq technique was used to identify a D614G mutation in spike protein-coding sequence. Studies are continuing until nowTables should be explanatory enough to be understandable without any text reference. Double spacing should be maintained throughout the table, including table headings and footnotes. Table headings should be placed above the table. Footnotes should be placed below the table with superscript lowercase letters. Sample table format is given below.

4. CONCLUSION

The number of cases of Coronavirus infection (110) case, Influenza virus (90) case, Parainfluenzavirus (65) case, Metapneumovirus (108)case and Rhinovirus (95)case, and the age group (51-61) year has the highest infection rate, and the incidence of males174 cases is higher than females128 cases.

ACKNOWLEDGEMENTS

I thank all health affiliates for their cooperation in collecting samples and completing the research requirements, in addition to the role of hospitals and health centers among their employees, and the cooperation and approval of patients.

COMPETING INTERESTS DISCLAIMER:

Authors have declared that no competing interests exist. The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors

REFERENCES

1- Tyson, J.., James, P., Stoddart, D., Sparks, N., Wickenhagen, A., Hall, G., & Quick, J. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore.2020. bioRxiv.

2- Bosch, B. J., Van der Zee, R., De Haan, C.., &Rottier, P. J. The Coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. Journal of virology.2003; 77(16): 8801-8811.

3- McNamara, R., Caro-Vegas, C., Landis, J.., Moorad, R., Pluta, L.., Eason, A., & Dittmer, D. High-density amplicon sequencing identifies community spread and ongoing evolution of SARS-CoV-2 in the Southern United States. Cell reports.2020; 33(5): 108352.

4- Huang, Y., Yang, C., Xu, Xf. et al. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. ActaPharmacol Sin .2020;41: 1141–1149.

5-Russell,L., Tin-Suet,J.,Zejia,C., Chelsea ,Z.,&Jianping, X. Global Temporal Patterns of Age Group and Sex Distributions of COVID-19.Infectious Disease Reports.2021;13: 582–596.

6-Wang, M. et al. Wang, M. et al. Preprint at medRxiv .Small.2020; 16: 2002169 .

7-Bloom, J. D. et al.Science .2021;372: 694.

8-Maria G., Claudia G., Gabriele A. , Gianni P. , Francesco S. Whole-Genome Sequence of SARS-CoV-2 Isolate Siena-1/2020. MicrobiolResourAnnounc. 2020 ; 24;9(39):e00944-20.

9- Anna M., Luca De., Cristin R., Michele C., Mariano F., Daniela T., Raffaella P., Francesca B., Ernesta C., Gerardo B., Gabriele V., Giuseppe. , Nicola N.(SARS-CoV-2 complete genome sequencing from the Italian Campania region using a highly automated next generation sequencing system. J Transl Med. 2021;5;19(1):246.

10-Jonathan P., Sara G., Erica L., Navjot S., Dary M., Kirsten St.Whole-Genome Sequencing of SARS-CoV-2: Assessment of the Ion Torrent AmpliSeq Panel and

Comparison with the IlluminaMiSeq ARTIC Protocol. J Clin Microbiol.2021; 18;59(12):e0064921.

11-Nihad A., Danilo L., Sreejith R., Simeone Dal. , Alessandro M. , Murad G ., Basel S., Yasir A. , Laith A . , Ali J., Karar K. , Ali A. , Mohamed B. Genome Sequencing of a Novel Coronavirus SARS-CoV-2 Isolate from Iraq. MicrobiolResour Announc.2021; 28;10(4):e01316-20.