Sequences of S-surface of human COVID-19

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 theApril**,**SeptemberandNovember months of 2020 year as well as theJanuary**,**February, April,June andJulymonths 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 .

Key words: Middle East Respiratory Syndrome (MERS) , Severe Acute Respiratory Syndrome (SARS) ,COVID-19.

Introduction

In December of 2019, a novel strain of coronaviruses, SARS-CoV-2, was revealedprominent to an occurrence of transmittableillnessaffecting a global pandemic. Coronaviruses were a great family of viruses well-knownforcontaminatetogetherpersons and different organisms. Thepersons, SARS-CoV-2 coronavirus contagions can sourceseveralinfectionsbeginning the corporate cold to otheracuteinfections 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 appearanceof 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 theSpike (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 theSARS-CoV-2 viral genome, we can identify the variants that occurvirus disaster and based on protocol for earlyARTIC SARS-CoV-2 sequencing in January 2020 which was approved by several scientific organizations around the world and later published the original protocolin September 2020 which was considered the most common for sequencingSARS-CoV-2[4].

2 .Material and Methods

2.1QuantitveReal Time PCR Technique:This technique was relied on for the diagnosis of viral infection, and an extraction was used from (Canvax HigherPurity™ 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	
S-SARS- CoV-2. IBS_m_S2 . Myungsun*et al.*,2020	**Primer F**	ACTGTTTTGCCACCTTTGCT	20bp	**300 bp**
	Primer R	AGCTTGTGCATTTTGGTTGA	20bp	
HA-*Influenza virus*	**Primer F**	TTGCTAAAACCCGGAGACAC	20bp	**228 bp**
	Primer R	CCTGACGTATTTTGGGCACT	20bp	
HPIV3gp4 *Parainfluenzavirus*type 3	**Primer F**	TGCCACCATCTATCAACCAA	20bp	**250bp**
	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-19andS-SARS- CoV-2.IBS_mS 2genotyping.PCR master mix was elaborated by using GoTaq® Green Mater Mix Kit (Promega,LOT.0000401B40,USA),Iqon PCR LadderCat. 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 Clustalalignment analysis, and NCBI-BLAST for the homology sequence identity.

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 7-2021all viruses were diagnosed through qRealtime PCRtechnique in figure (1) 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) 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)]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 females128 cases.

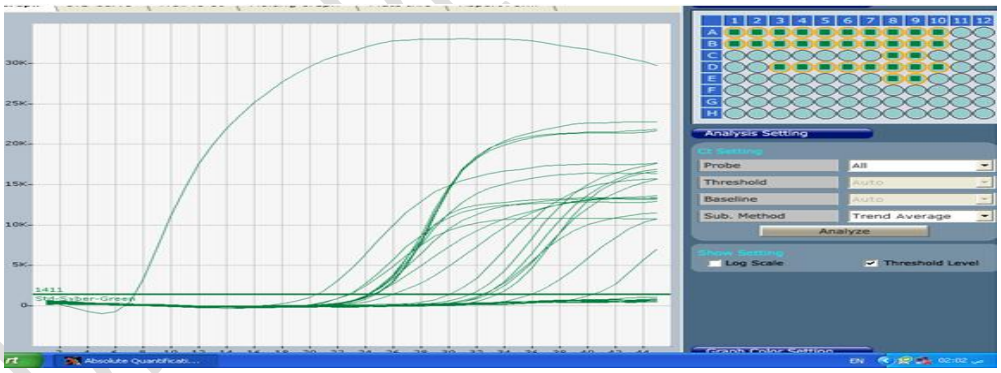

Figure (1): Disgnosis of all virueses (*Coronavirus*, *Influenza virus*, *Parainfluenzavirus*, *Rhinovirus* and *Metapneumovirus*) by qRT-PCR technique

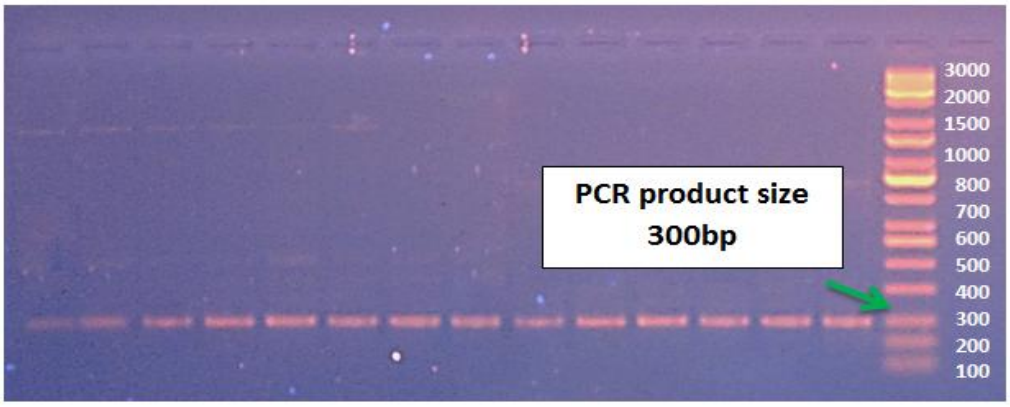

Figure (2): Diagnosis of Human S-SARS- CoV-2by PCR (PCR product size 300bp, 5 µl Iqon PCR Ladder was loaded ona 1.5 % agarose in 1x TBE and stained with ethidium bromide.dsDNA ladder with bands from 100 bp to 3.000 bp.

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.

Sangersequence analysis of the *Coronavirus* variants. All 10 samples that amplified with both the S-Covid-19 in 20-11-2020 and S-SARS- CoV-2IBS_mS 2 primers from 20 -11-2020 up to 10-3-2021 . Table 2 appearthe 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 (2): The NCBI-BLAST homologysequence identity (99.92-78%) between local human Coronavirus isolate and NCBI-BLAST submitted Coronavirus isolate.

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%

Discussion:

According to the official approvals for the purpose of diagnosing the virus through qReal TimePCR , samples were collected from patients infected with the Coronavirus, and then a s-spike singer sequencing was conducted for different samples of age, gender and virus concentration. The fourth, ninth and eleventh of 2020year , 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 [5].

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 [6].

Sanger Sequence analysis of the Human Coronavirus variants. All 10 samples that amplified with both theS-Covid-19 and S-SARS- CoV-2 IBS_mS 2primers clustered with genotype in table (2).In our study, the Sanger Sequence analysis was usedthe NCBI-BLAST homology sequence identity (99.92-78%), while in other studies, whole genomes were used Maria *et al* ., 2020the 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 Nanoporedelivers were placed in the sequence beneathBioProjectagreement 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* ., 2021IlluminaMiSeq technique was used to identify a D614G mutation in spike protein-coding sequence.Studies are continuinguntil now.

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. (2020). Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore.bioRxiv.

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

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

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

5-Wang, M. *et al.* (2020)Wang, M. *et al*. Preprint at medRxiv.Small 16, 2002169 .

6-Bloom, J. D. *et al.* (2021).*Science* 372, 694.

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

8-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.(2021).SARS-CoV-2 complete genome sequencing from the Italian Campania region using a highly automated next generation sequencing system.J Transl Med. 5;19(1):246.

9-Jonathan P., Sara G., Erica L., Navjot S., Dary M., Kirsten St.(2021) Whole-Genome Sequencing of SARS-CoV-2: Assessment of the Ion Torrent AmpliSeq Panel and Comparison with the IlluminaMiSeq ARTIC Protocol. J ClinMicrobiol. 18;59(12):e0064921.

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