Microbes and Infectious Diseases 2025; 6(3): 1015-1023

Microbes and Infectious Diseases

Journal homepage: https://mid.journals.ekb.eg/

Original article

Molecular sequence comparison for genome and S protein of COVID-19 virus from different countries

Hiba M. Nasir [1], Ali I. Ibraheem[2], Waleed K. Ibraheem[3]

1- Department of Medical Techniques, Faculty of Health and Medical Techniques, Middle Technical University, Baghdad, Iraq

2- Otolaryngologist, ENT department, Al-Furat General Hospital, Iraq

3- community medicine, Public health department , Al-Furat General Hospital, Iraq

ARTICLE INFO

Article history:

Received 11 January 2025

Received in revised form 26 January 2025

Accepted 4 February 2025

Keywords:

SARS-CoV-2 Genome

Spike Protein

Phylogenetic Analysis

Mutations

Genomic Diversity

ABSTRACT

Background: COVID-19 is caused by the novel virus SARS-CoV-2, which led to an unexpected worldwide pandemic. Numerous researchers worldwide have strived to achieve genome sequencing of the virus and share highly conservative data about Covid-19 genome with very few mutations and no recombination. **Methods:** Tthis study was carried out to align eleven genomes sequence (29000) of SARS-CoV-2 from different countries (Iraq, Turkey, Iran, Jordan, Egypt, Morocco, India and china) and (Basra, Erbil) with a reference genome called Covid virus in NCBI by using many software (MAFFT, Genome detective, Next clade, Neighbor joining for all existing alignments and estimate a relationship between them. **Results:** Results of MAFFT alignment for the sequence of SARS-CoV-2 genome found that the 11 sample had a nucleotide identity comparable to 99 with reference genome. While the results of mutation rate indicated that many variations between these sequences and indicated high rate of mutation for Egypt, china, Iraq, Erbil, Iran, India, Morocco, Basra, Turkey and Jordan, with most frequent mutation 23403A>G. While the phylogenic tree found the all "SARS-CoV-2" sequences accumulated in a different clade, Sequences from a variety of nations were found in some big clades. Although there were some mutation in low frequencies, with common mutation D614G found in most sequence (Iraq, Erbil, Turkey, Jordan, Egypt, Morocco, India and china). This mutation makes a virus transmit more efficiently and 10-fold higher aggressive than the original Chinese strain. Followed by the second most frequent mutation L452R indicates that the binding to the selected monoclonal antibody (mAb) is reduced and may affect their neutralization potential. **Conclusion:** The S protein of the delta and kappa variants demonstrated improved viral replication kinetics when compared to the wild-type virus.All these data should be considered for developing vaccines and antiviral treatment strategies and tracking the diversity of viruses around the world. The initial candidates for the COVID-19 vaccine were developed to specifically target the D614 sequence position of the SARS-CoV-2 virus, the efficacy of the existing antibodies and vaccines may be carefully tested with the G614 variant, immunogenicity and protective efficacy of a DNA vaccine encoding spike protein with D614G mutation were tested. A segment of nucleic acid-based, vectored, protein subunit vaccines was developed utilizing the S protein of SARS-CoV-2, incorporating these specific mutations.

Introduction

The recent pandemic of coronavirus disease 2019 (COVID-19) was caused by the novel enveloped single-stranded RNA virus associated with various respiratory viruses. It was quickly spreading to more than 200 countries. SARS-CoV-2 is a member of the

DOI: 10.21608/MID.2025.352093.2454

* *Corresponding author:* Hiba Nasir

E-mail address: *mansour.hiba@yahoo.com*

-coronavirus genus possessing a RNA strand genome 26-32 kb in length [1]**. Coronaviruses are classified as alpha, beta, gamma, and delta**. SARS-CoV2 -19 and SARS-CoV1-2003; both of them are coronaviruses[2]. All Wuhan coronavirus gene were by analyzed based on grouping of (ORF1ab, S, ORF3, E, M, ORF6, 7a, 8, and ORF10[3]. The viral protein such as spike (S) protein permits the virus to attach to cell membrane, while the (N) Nucleocapsid protein contains viral genome, (E)envelope,(M) membrane [4]. In order to promote the virus to enter the host cell, the S protein was responsible for SARS-attachment to the cell surface receptor and subsequent fusion [5] . The virus specifically affects organs that contain angiotensin converting enzyme 2 (ACE2)[6]. The mutation and evolution rates in RNA viruses are considerably higher than those of their hosts, up to a million times , and the high level of it is attributed to the connection with virulence variation, both of which are considered beneficial for viral persistence [7]. Mutational analysis of SARS-CoV-2 genomes sourced from various geographical locations revealed several distinctive characteristics, including significant country-specific modifications in the virus's proteins, such as the replication polyprotein, spike glycoprotein, envelope protein, and nucleocapsid protein. Coronavirus nucleocapsid protein is required for RNA replication, transcription, and genome packing [8].These viral nucleocapsid proteins ((NCs) encapsulate the RNA genomes of viruses, creating NC–RNA complexes that serve as templates and are crucial for viral replication and transcription. Multiple mutations have been discovered in the receptor-binding motifs (RBMs) of the S protein, which plays a crucial role in enabling viral entry by interacting with the human angiotensin-converting enzyme 2 receptor on host cells [9]. The D614G mutation was found to be the most common mutation across different cell types, such as those from the colon, liver, and lungs [10]. In contrast, the L452R mutation reduced the effectiveness of antibody neutralization and enhanced the virus's infectivity [11]. So, this study was carried out to align eleven genomes sequence (29000) of SARS-CoV-2 from different location (Iraq, Turkey, Iran, Jordan, Egypt, Morocco, India and china)and (Basra, Erbil) with a reference genome called Covid virus in NCBI by using many software (MAFFT, Genome detective, Next clade, Neighbor joining for all existing alignments and estimate a relationship between them.

Material and Method:

The complete viral genome of 11 viral sequence from different countries recorded in NCBI were studied. Only viruses that infect humans were chosen, with low-quality sequences were removed and only full-length sequences used (>29,000 nt). SARS genome sequence, with the reference sequence (ID: NC. _045512), compare with Sequence of Iraq: Basra, Erbil, Turkey, Iran, Jordan, Egypt, Morocco, India and china with accession number lcl|MZ366458.1, lcl|MW546610.1, lcl|MW633517.1, lcl|MT605818.1, lcl|MT646036.1, lcl|MZ266636.1, lcl|MW737845.1, lcl|MW582699.1, lcl|MZ336033.1 and lcl|MW691153.1. were filtered and analyzed. The sequence taken from NCBI converting into Fasta format in order to analyzed. Multiple sequence alignment was done by MAFFT (bioinformatics software download from internet) to compare Sequences of SARS-CoV-2 from 11 distinct nations. To distinguish between different countries' identities and mutations, a genomic detective was used in addition to the phylogenic tree and different geographical viruses were done by using the neighbor-joining (NJ) and evolutionary distances were estimated. Finally amino acid sequencing alignment of S protein and detection of the mutation for 11 fasta format of S protein to previous accession number sample were analyzed by using Next clade (bioinformatics software download from internet).

Results:

We partitioned the SARS-CoV-2 genome suc]cession information as indicated by their geographic beginnings from nine country and (Erbil ,Basrah) for example, Iraq, Basra, Erbil, Turkey, Iran, Jordan, Egypt, Morocco, India and china. These SARS-CoV-2 successions had a place with the infected patients from 8 nations and 3 from Iraq of the total genome grouping. As a form of perspective genome, the SARS-CoV-2 Wuhan-Hu-1 strain was used. The result of filter sequence alignment as shown in figure (1) revealed that the 11 sequence with length comparable to (29000) were similar in many sequence with variable differences between viral in different country.

The result of the nucleotide mutation as in figure (2) which revealed different types of mutation such as substitution, insertion, deletion, present on sequence tested is compared to the reference of Wuhan, which affects the viral effectiveness this

result. Figue 2(a) revealed the frequency of all mutation in all sample and the (**3037C>T), (23403A>G)** most common mutation in the all sample as in figure 2(b). It was in agreement with [12] who found the many nucleotide change in sequence from different country and that the main cause made SARS-CoV-2 was accountable alterations for death and clinical symptoms. In addition, variations in UTRs have been found to alter the activity, replication, and packaging of genomes, as well as immunological regulation and gene expression in some viral genomes, according to the research [13,14] .

While the phylogenic tree found the 11 SARS-CoV-2 sequences of different area accumulated in a different clade with reference SARS-CoV, many country's sequences accumulate in one clade. Genetic variability within SARS-CoV-2 strains from various countries could thus be connected to their geographic dispersion. COVID-19 was linked to meteorological factors, with dry and cold conditions appearing to hasten virus propagation. Result of figure (3) phylogenic tree revealed the variation in genomes from various geographical areas used in the study, with a phylogenetic analysis. Neighbor joining method was used for the construction of consensus tree.

Result show distinctive phylogenetic distances on clades of covid-19 presented evolutionary relationships among coronaviruses from different country. The tree can be categorized into two primary branches. The lower branch is directly linked to the Iraqi sequence, while the upper branch corresponds to sequences found in the earliest recorded clusters from Wuhan, China, which are closely related to samples from other countries. The results indicate that Egypt and India belong to the same clade.

The phylogenic tree of S protein (figure 5) by using neighbor phylogenic tree revealed that sequence cluster into two main clades. The trees, arranged in diverse sequences, were designed to visually illustrate the transmission across various regions. The first clade includes sequences from Erbil and Jordan, while the second clade encompasses the remaining sequences from other areas.

Iran isolates was more similarity to the original china sequence fallow with Basrah, then china, Iraq and Turkey respectively were . The tree showed detailed evolutionary relationship of SARS-CoV-2, with many sub-branches identified. It is worth noting that some country coronaviruses sampled from the china location, displayed closer genetic distance, which is rational and logical from the perspective of evolutionary progress. The short distance between Erbil and Jordan sample which is located in same clade this may attributed to the distance and transmitted people between them. While the other sample in another clade and Egypt far from them.

The mutation study in all (S) proteins of eleven SARS-CoV-2 genome and in comparison with SARS-CoV reference genome. Result of parent four confirmed many modifications amongst all eleven SARS-CoV-2 Spike proteins are significant surface glycoproteins that are notably recognized for their crucial role in interactions with host cell receptors.. Further, we study the mutations with inside the spike protein of SARS-CoV-2 from specific international locations at diverse positions in the investigated SARS-CoV-2 isolates. The result of SARS-CoV-2 isolated form Turkey shows one mutation, while India shows two amino acid mutation at D614G (23403A>G), D936Y (24368G>T) When compares with reference sequence. While Iran sequence isolate showed amino acid change at 22 (Serine to Threonine). In addition, the sequence belonging to Jordan shows a change at 12 (Serine to phenylalanine); 69-71 (deletion), 176 (Lucien to phenylalanine), 452 (Lucien to arginine), 614 (aspartic acid to glycine) and 899 (Asparagine to serine). Finally the Iraqi isolate shows the mutation at 70 (valine to Lucien), 272 (Proline to Lucien), 501 (asparagine to tyrosine), 570(Alanine to aspartic acid), 681(phenylalanine to histidine), 716(threonine to isoleucine), 982 (serine to alanine); 1118 (aspartic acid to histidine), while India 614 (Asparagine to glycine) 936 (Asparagine to tyrosine) and in addition Egypt and china show only 614 (asparagine to glycine) while Morocco appears mutation at 78 (arginine to methionine); 614 (aspartic acid to glycine). Finally, it is seen that Erbil protein aliment show mutation at 452 (Lucien to arginine); 583 (glutamic acid to aspartic acid); 614 (aspartic acid to glycine). Result of figure 6 declared that with common mutation D614G found in most abundant mutation fallow by the second most frequent mutation L452R mutation.

These mutation rate will change the viral pathogenicity and stability. Result of the characteristics of the S proteins with aspartic acid (SD614) and glycine (SG614) at residue 614

exhibition that SG614 is much more stable than SD614.Fallow with second frequent mutation L452R which reduced affinity to a subgroup of monoclonal antibodies (mAbs) that may have an impact upon their neutralizing capability [15] found that the variant L452R can evade HLA-24-restricted cell - mediated immunity with simultaneously enhancing viral pathogenicity and possibly increasing viremia.

Figure 1. Multiple sequence alignment between different sequence for (11) sequence by using MAFFT alignment to identify the similarity between sample.

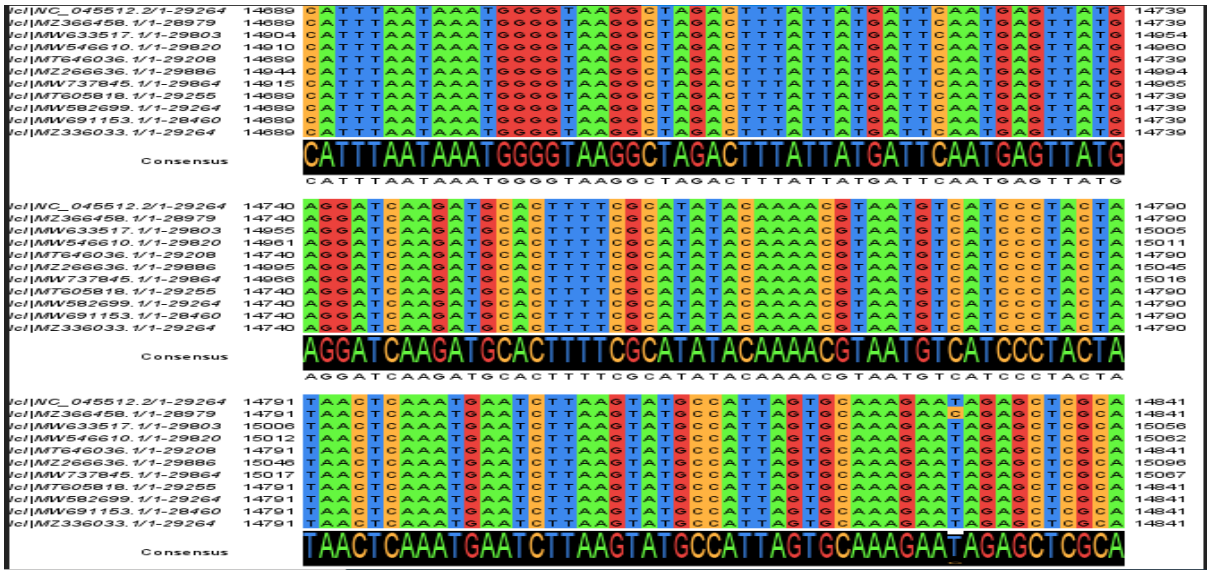
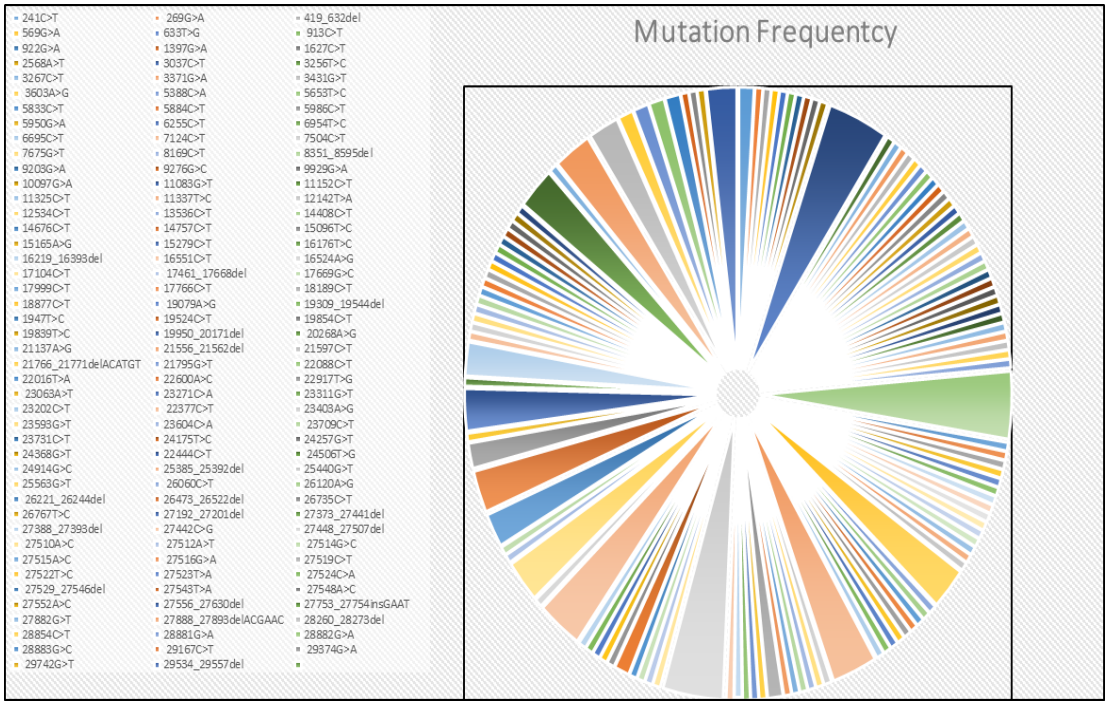
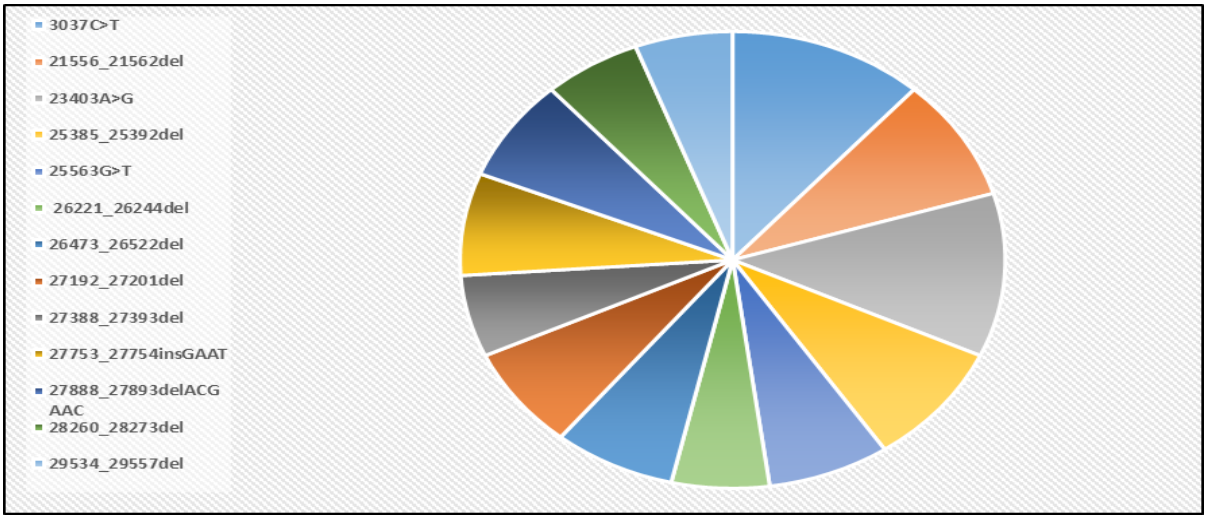


Figure 2. Frequency of mutation in all sample region a) frequency of all mutations (b) frequency of the most common mutation.

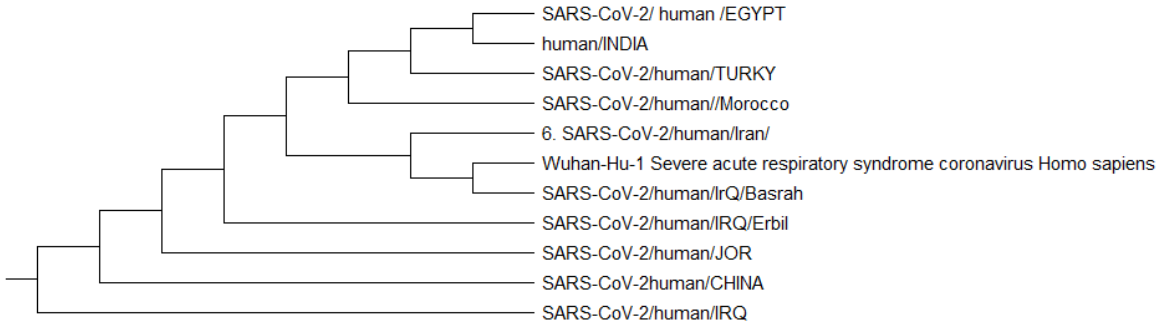


(a)



(b)

Figure 3. Tree to declare the relationship of SARS-CoV-2 complete sequence (nucleotide) from different countries.



While figure (4) of the multiple sequence alignment for nucleotide of S protein sequence revealed some comparability with some variant due to mutation deletion and substitution.

Figure 4. Multiple sequence alignment of amino acid sequence of S protein for Covid 19 from different countries.

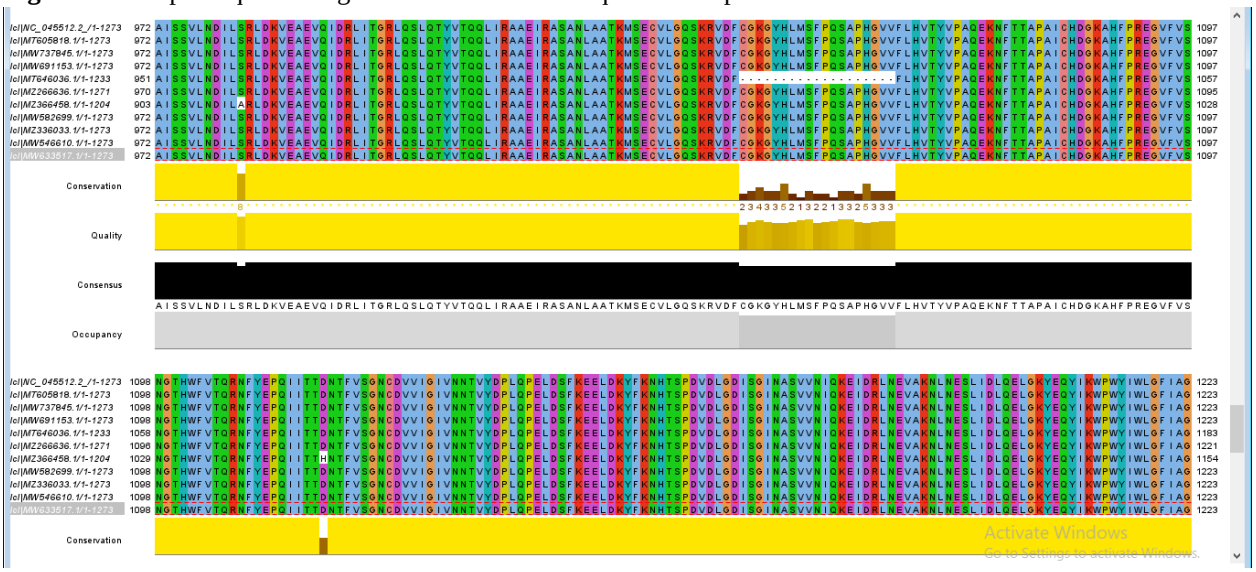


Figure 5. Neighbor joining Phylogenetic tree to declared the SARS-CoV-2 S protein sequence (nucleotide) relationships between countries.

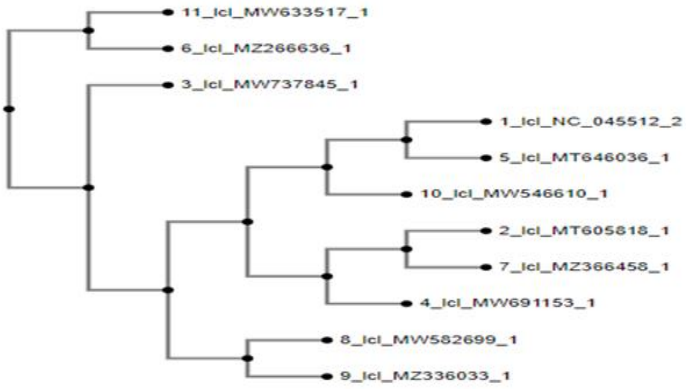
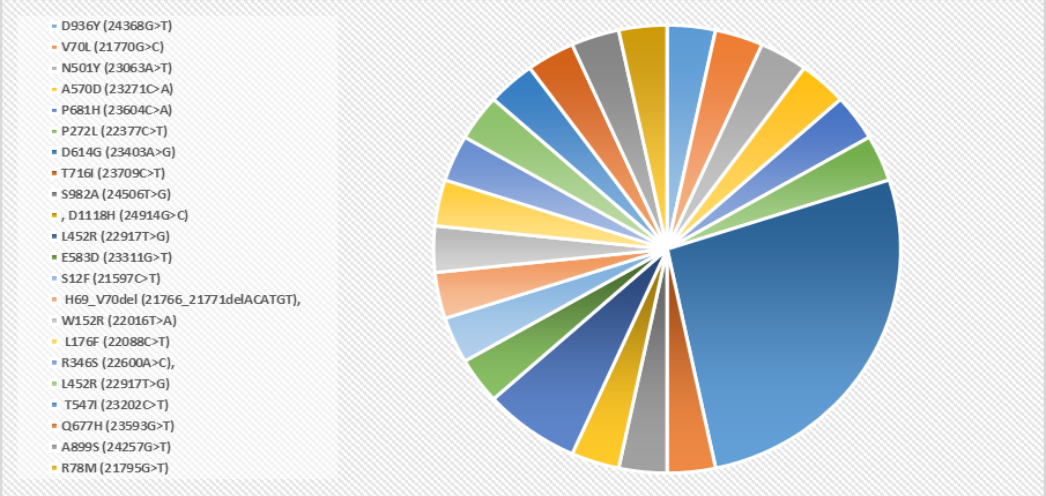


Figure 6. Mutation frequency of amino acid sequence in covid 19 sequence of different countries.



Discussion

Among all known RNA infections, Covid comprise of the biggest genome [16].The enormous genome size gives greater versatility in obliging and altering qualities (26.4 to 31.7 kb). All 11 SARS-CoV-2 sequences for different area accumulated in a different group with reference SARS-CoV-2 sequences according to phylogenetic analyses (Fig 3) with ancestor an]d non-ancestor. Egypt having high similarity with morocco, while turkey sequence with Iraq and china are highly comparable with Jordan while finally Basra is comparable with Iran.Phylogenetic analysis on the coronavirus genomes has founded that SARS-CoV-2 was a replacement member of the betacoronavirus genus, which incorporates SARS-CoV, MERS-CoV, bat severe acute respiratory syndrome connected coronaviruses (SARSr-CoV), further as alternative coronaviruses known i n humans and animal species [17,18].While [19] declared the phylogenomics viral sources, transmission, and potential superinfection in early-stages of COVID-19 patients in Ontario, Canada. In addition, [20] mentions that the COVID-19 prevalent due to a high mutation result a changes in the viral behavior such as its virulent or infectious, and resistance to drugs.

The [21] indicated that the results differ regarding the evolutionary relationships of the Spike protein regions for Lebanon, the West Bank, Iraq, and the Philippines. The D614G amino acid substitution was frequently observed in the spike protein sequences of SARS-CoV-2 genomes from Bangladesh and Iran. Additionally, countries such as India, Nepal, Qatar, Pakistan, Saudi Arabia, Turkey, Bahrain, Bhutan, Jordan, South Korea, China, Georgia, and Hong Kong exhibit closely related evolutionary patterns in their spike protein sequences, characterized by relatively fewer amino acid substitutions. The close associated between different regen may attributed to the work and transmission between different country. And that in

accordance with [22] who revealed that the phylogenetic and evolutionary models successfully linked the human transmission of the SARS-CoV-2 in the South Asian regions through international travel between various neighboring countries as most of these share land borders interactions between ACE2 and S protein.

Epidemiologic studies proposing that viruses with SG614 transfer more effective [23]. However, the aspartate to glycine mutation (D614G) of the S protein was assumed to be 10-times more infective than the original strain from China (Wuhan-1)[19]. Patients infected by the D614G mutant variant had advanced viral loads [24] .For example, SARS COV-2 with a D614G change in the S protein attacks ACE2 express tissues more efficiently, improving its virulence. As results, an individual exposed with a mutated variant develops higher infection rates and much more acute COVID-19 symptoms. Some primary significance is mutations in the S protein and it is due to the idea that the spike protein defines viral host range and it is often the target of neutralizing antibodies [25,26] S protein facilitates viral invasion into host cells by adhering to a host receptor via the RBD in the S1 subunit and then fusing the viral and host surfaces via stimulating by host tissue proteolytic enzymes of a S2 subunit. Pseudo viruses containing the D614G mutation, either in combination with L452R or W152C, or swithout.The D614G variant, were produced and utilized to infect 293T cells that stably express the ACE2 receptor for cell entry and the TMPRSS2 cofactor necessary for SARS-CoV-2[27]. Additionally, recurrent attenuating mutations, including P323L, L37F, G251V, and Q27stop, have been identified, which are believed to diminish the severity of the disease. The emergence of these attenuating mutations indicates that SARS-CoV-2 may be evolving towards a less pathogenic form in human [28].

1. The first COVID-19 vaccine candidates were developed to specifically target the original D614 strains of SARS-CoV-2. Consequently, there exists a connection between the D614 spike antigen utilized in the vaccines and the G614 variant, which heightens the likelihood of the virus circumventing the immunity produced by the vaccine [29].

Since Severe acute respiratory identifies ACE2 like its host receptor for adhering to virus S proteins, identifying the RBD in Severe acute respiratory S protein in its most possible target for the virus adhesion process, like new inhibitors, neutralizing antibodies, and immunizations, is essential.

Conclusion:

COVID-19 is a global pandemic that has spread worldwide.. Because the virus's adaptability mechanisms are natural, there are so many variations on its genomic that more research is wanted to find a method of controlling the sickness. The MAFFT was used to align eleven SARS-CoV-2 isolate genomes (29000) the relationship between them and the mutations observed were linked to their impact on the primary protein structure. Furthermore, the alignment and mutation presented within S protein show many amino acid substitution and deletion found within S protein sequence and these change result on the change of the stability and activity of the virus. Monitoring the emergence of future variants of SARS-CoV-2 is essential for establishing effective control measures for the COVID-19 pandemic. Mutations that are believed to diminish immune recognition, particularly those occurring in the spike protein, have been shown to decrease sensitivity to both natural immunity and immunity conferred by vaccines.

DECLARATIONS

ACKNOWLEDGMENTS

the author gratefully acknowledges Ministry of higher education and scientific research in Iraq.

CONFLICT OF INTEREST

the authors declare that there is no conflict of interest.

ETHICS STATEMENT

Not Applicable.

AVAILABILITY OF DATA

The data generated during and/or analysed during the current study are included in the manuscript.

References

1- Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a newcoronavirus of probable bat origin. Nature, 2020;579: 270–273.

2- Marra MA, Jones SJ, Astell CR, Holt RA, Brooks WA, Butterfield YS, et al.The genome sequence of the SARS-associated coronavirus.

Science. 2003;300(5624):1399-404.doi: 10.1126/science.1085953.

3- Chang T, Yang DM, Wang ML, Liang KH, Tsai PH, Chiou SH, Lin TH, Wang CT. Genomic analysis and comparative multiple sequences of SARS-CoV2. Journal of the Chinese Medical Association 2020.83 (6): 537-543.

4- Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science.2003;300: 1394-1399.

5- Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. Virology journal.2019; 16: 1-22.

6- Ibraheem AI, Nasir HM, Abdulamir AS, Ali CL, Kabah KK, Hussein IA, Ftak DF, Rasheed AM. EARLY IDENTIFICATION OF HIGH RISK COVID-19 PATIENTS USING HEMATOLOGICAL INDICES. JUMMEC .2021;1:1-6.

7- Pachetti M, Marini B, Benedetti F, Giudici F, Mauro E, Storici P,et al. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. Journal of translational medicine.2020;18: 1-9.

8- Hsin WC, Chang CH, Chang CY, Peng WH, Chung LC, Chang MF, et al. Nucleocapsid protein-dependent assembly of the RNA packaging signal of Middle East respiratory syndrome. J Biomed Sci.2018. 24;25(1):47.doi: 10.1186/s12929-018-0449-x.

9- Mittal A, Manjunath K, Ranjan RK, Kaushik S, Kumar S, Verma V. COVID-19 pandemic: insights into structure, function, and hACE2 receptor recognition by SARS-CoV-2. PLoS Pathog. 2020;16 doi: 10.1371/journal.ppat.1008762. e1008762. [DOI] [PMC free article] [PubMed] [Google Scholar

10- Daniloski Z, Guo X, Sanjana NE. The D614G mutation in SARS-CoV-2 Spike increases transduction of multiple human cell types. BioRxiv Prepr. Serv. Biol. 2020 doi: 10.1101/2020.06.14.151357.

11- Liu Z, VanBlargan LA, Bloyet M, Rothlauf PW, Chen RE, Stumpf Set al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. Cell Host Microbe. 2021;29(3):477-488.e4. doi: 10.1016/j.chom.2021.01.014

12- Angeletti S, Benvenuto D, Bianchi M, Giovanetti M, Pascarella S, Ciccozzi M, COVID-2019: The role of the nsp2 and nsp3 in its pathogenesis. J Med Virol. 2020;92(6):584-588. doi: 10.1002/jmv.25719.

13- De Lorenzo G. Drikic M. Pap G. Eichwald C. Burrone O R, and Arnoldi F. An inhibitory motif on the 5’UTR of several rotavirus genome segments affects protein expression and reverse genetics strategies. PLoS ONE.2016; 11. Article e0166719, 10.1371/journal.pone.0166719.

14- Silveira CJ, Equestre M, Pagnotti P, Gradi A, Sonenberg N, Perez BR. 5'UTR Of Hepatitis A Virus RNA: Mutations In The 5'-Most Pyrimidine-Richtract Reduce Its Ability To Direct Internal Initiation Of Translation. J Gen Virol 2009;76:1189-1196.

15- Motozono C, Toyoda M, Zahradnik J, Ikeda T, Saito A, Tan TS, et al. An emerging SARS-CoV-2 mutant evading cellular immunity and increasing viral infectivity. BioRxiv.2021;https://doi.org/10.1101/2021.04.02.438288.

16- Woo PC, Huang Y, Lau SK, Yuen KY. Coronavirus genomics and bioinformatics

analysis. Viruses. 2010 7;2(8):1804-1820. doi: 10.3390/v2081803.

17- Li Q, Wu J, Nie J, Zhang L, Hao H, Liu S,et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. Cell. 2020 Sep 3;182(5):1284-1294.e9.doi: 10.1016/j.cell.2020.07.012.

18- Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. The lancet. 2020;395(10224):565-574.doi: 10.1016/S0140-6736(20)30251-8.

19- Calvin PS, Nazneen R, Gerald AE, David H, Santiago P, Melissa LH et al. Phylogenomics reveals viral sources, transmission, and potential superinfection in early-stage COVID-19 patients in Ontario, Canada. Scientific Reports.2021; 11:3697.

20- Wabalo EK, Dubiwak AD, Senbetu MW, Gizaw TS, Effect of Genomic and Amino Acid Sequence Mutation on Virulence and Therapeutic Target of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS COV-2). Infect Drug Resist. 2021 Jun 14:14:2187-2192. doi: 10.2147/IDR.S307374

21- Kamaleldin BS, Ahmed A, Fawaz A, Khalid FA, Meshari A, Tulika B, Mohammad ZN, Rajeev S, Mohd A K. Molecular evolutionary model based on phylogenetic and mutation analysis of SARS-CoV-2 spike protein sequences from Asian countries: A phylogenomic approach. . Informatics in Medicine Unlocked.2023; 38:101221. https://doi.org/10.1016/j.imu.2023.101221.

22- Wang Q, Zhang Y, Lili W, Niu S, Song C, Zhang Z, Guangwen L, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell. 2020;181(4):894–904.e9. doi: 10.1016/j.cell.2020.03.045.

23- Zhang L, Jackson CB, Mou H, Ojha A, Rangarajan FS, Izard T, Farzan M, Choe H,1 The D614G mutation in the SARS-CoV - 2020. shedding and increases infectivityNature Communications doi: 10.1038/s41467-020-19808-4

24- Korber B, Fischer,W, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2. BioRxiv;2020. doi: https://doi.org/10.1101/2020.04.29.069054.

25- Chen WH, Hotez PJ, and Bottazzi ME. Potential for developing a SARS-CoV receptor-binding domain (RBD) recombinant protein as a heterologous human vaccine against coronavirus infectious disease (COVID)-19. Hum Vacc immunother 2020; 16:1239-1242.

26- Yuan M, Wu NC, Zhu X, Lee CC, So RT, Lv H, et al. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. Science.2020;368(6491):630-633. doi: 10.1126/science.abb7269.

27- Hoffmann M, Kleine-Weber H,· Schroeder S.SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 2020;181(2):271-280.e8.doi: 10.1016/j.cell.2020.02.052.

28- Jalen S, Pranav P, Andrew G. Mc, Arinjay B, Karen M. Evolutionary trajectory of SARS-CoV-2 and emerging variants. VIROLOGY .2021; 18: 166 .

29- Jing Z, Xuping X, Camila R. F, Kena AS, Isis K, Kristin T, et al.The effect of SARS-CoV-2 D614G mutation on BNT162b2 vaccine-elicited neutralization. Vaccines .2021; 6.