



Comparison of Sequencing and Phylogenetic Analysis of SARS-CoV-2 Spike Proteins Extracted from Patients and Travelers in Duhok-Iraq

Omar Mohammed Younus ¹, Arif Parmaksiz ² and Amer Abdalla Goreal ^{3,*}

¹Preventive Health Affairs, Directorate General of Health, Duhok, Iraq

²Department of Biology, Faculty of Science-Literature, Harran University, Sanliurfa, Turkey

³Department of Medical Microbiology, College of Medicine, University of Duhok, Duhok, Iraq

*Corresponding author: Department of Medical Microbiology, College of Medicine, University of Duhok, Duhok, Iraq. Email: amer.goreal@uod.ac

Received 2023 June 03; Revised 2023 July 12; Accepted 2023 July 15.

Abstract

Background: SARS-CoV-2 is a single-stranded RNA virus and a member of a large family of *Coronaviruses* that are important human pathogens. This virus caused severe acute respiratory syndrome and was initially identified to be transmitted between humans on November 17, 2019.

Objectives: To investigate the lineage, mutational patterns, variants, and serotypes of SARS-CoV-2 viruses circulating in the Duhok governorate population and to compare them with those identified in travelers crossing the border from Turkey in order to trace the epidemiological patterns.

Methods: Nasopharyngeal swabs were collected from 700 individuals living in Duhok and 700 travelers crossing the border to Duhok-Iraq from Turkey. The subjects were recruited by random sampling and questioned about demographic features and symptoms of upper or lower respiratory tract infections. Exclusion criteria included vaccination with COVID-19 vaccines of any approved previous infection. Samples were subjected to RT-PCR, and 30 positive samples with the highest viral load (lowest Ct values) were chosen for sequencing of the complete S gene by next-generation sequencing (NGS). Three platforms of Nextstrain, GISAID, and PANGO were used to identify variants, clades, and lineages and analyze sequences.

Results: Out of 1400 participants, 353 (25.21%) positive samples were identified by RT-PCR, of which 30 representative positive samples (15 from each group: Patients and travelers) were sent for complete sequencing of the S spike gene using NGS. Nineteen samples were successfully sequenced and retrieved, including nine samples from Duhok residents and ten samples from travelers. Nextclade results revealed that 12 samples belonged to the delta strain (Pango lineages: B.1.617.2.78, B.1.617.2, B.1.617.126, and B.1.617.121) distributed among the two groups while 5 omicron (BA.1.1) and 2 alpha (B.1.1.7) strains were found among travelers. A total of 76 mutations, including 52 non-synonymous, 16 synonymous, and 8 deletions, were detected without identifying a unique mutation. Sequencing results were submitted to GISAID, and accession numbers were obtained. A phylogenetic tree was constructed using the sequences obtained from Iraqi and non-Iraqi variants from GISAID.

Conclusions: The present research presents a description and observation of the genetic and epigenetic status of SARS-CoV-2 in Iraq based on sequencing results. The study revealed the impact of travels in introducing new variants to the country, including those with mutations in the S1 domain of the spike protein that can enhance viral attachment to receptors.

Keywords: COVID-19, Clades, GISAID, Phylogenetic Relationship, Mutations, Variants

1. Background

Coronaviruses are enveloped single-stranded positive-sense RNA viruses with pleomorphic or spherical forms and projections resembling spikes on the surface, displaying the appearance of a crown (1). The SARS-CoV that was initially found in China in 2002 caused the SARS outbreak in 2002 - 2003. Middle East respiratory syndrome (MERS) was first identified as a serious human disease in Arabian Peninsula in 2012 (2). On December

31, 2019, Wuhan Municipal Health Commission in China discovered a sizable number of pneumonia patients in Wuhan, Hubei (3, 4). This was when the first SARS-CoV-2 outbreak began. The new coronavirus (or SARS-CoV-2) epidemic was classified as a Public Health Emergency of International Concern (PHEIC) by the director general of the World Health Organization (WHO) on January 30, 2020 (5). The WHO reported on January 30 that there were 7818 confirmed SARS-CoV-2 cases worldwide, with 7736 of those

cases occurring in China and only 82 cases being related to 18 other countries.

The spike (S) protein of the SARS-CoV-2 is the main protein involved in the interaction of the virus with human angiotensin-converting enzyme2 receptor (ACE2) and transmembrane serine (TMPRSS2) receptors on target cells. Many efforts have been undertaken to create COVID-19 vaccines targeting this viral component (6). Natural selection causes any virus to accumulate random mutations. It has been discovered that the S1 subunit of the "S" protein harbors the majority of alterations responsible for transmission, virulence, and host immune evasion, which besides giving rise to several variations, this phenomenon has numerous therapeutic implications (7, 8).

The most contagious mutant of SARS-CoV-2, delta type (B.1.617.2) and delta plus (AY1) strains emerged in the second COVID-19 outbreak wave and were first identified in late 2020 in India, facilitating its rapid spread throughout the globe. The delta plus variant, which was discovered in India on October 5, 2020, was classified later as the variant of concern (VOC) because it was more contagious, could bind lung cell receptors more tightly, and produced a weaker antibody response than delta (9). According to previously published data, a common K-417N mutation between delta plus (AY1) and beta (B.1.351) variants was found to be the cause of neutralizing resistance. In Vietnam, another delta variant sub-lineage was detected as delta-V, which caused a rise in virus propagation at the time. As a result of these changes in the "S" protein of the alpha variant, this variant was identified as a hybrid virus.

Omicron (B.1.1.529), a novel SARS-CoV-2 strain, was first identified in Gauteng, South Africa, in mid-November 2021 (10). As of December 16, 2021, 87 countries had already discovered the omicron version (11). While the 'S' protein of this mutant strain of SARS-CoV-2 has more than 32 mutations, its receptor binding domain (RBD) revealed 15 alterations (12). Omicron's rapid rate of transmission, probably due to its ability to hide from the host's immune system, resulted in a sudden rise in the number of COVID-19 cases worldwide (12). The incubation time of omicron was shorter, lasting just three days (13). The variant caused less damage to the lungs than previously identified strains because it mainly replicated in the upper respiratory tract.

Many scientists believe that omicron's high transmission rate and very low pathogenicity could lead to the development of herd immunity, providing hope for the end of the pandemic (14). This variant's moderate pathogenic traits often include "cold and flu" accompanied by a sore throat, headache, myalgia, throat pain, fever, and lethargy, with no signs of a loss of flavor

or aroma, as observed in previous SARS-CoV-2 strains. Only a few patients had severe symptoms requiring hospitalization, and most of them experienced mild symptoms that did not demand hospitalization (15).

However, some reports asserted that a significant number (21%) of African patients hospitalized due to contracting the omicron variant had disastrous clinical results (16). The patients infected with the omicron variant developed strong immune responses that could neutralize omicron, as well as other variants of the SARS-CoV-2, reducing the chances of re-infections with delta type, dethroning it as the dominant strain (17). Due to its recent emergence, the omicron variant lacked strong evidence regarding the extent of infection-induced immunity; however, Schulze Zur Wiesch predicted that it would be comparable to other variants. Accordingly, people exposed to omicron in the last few weeks would probably be protected for the upcoming few months given that their antibody levels are high enough because omicron spreads more quickly than earlier strains (18).

2. Objectives

The objectives of this study were to determine the clades, variants, and lineage of SARS-CoV-2 strains extracted from residents and travelers in Duhok, Iraq, by sequencing. Furthermore, we explored the mutational patterns of the viruses circulating in the Duhok population and compared them with those identified in travelers crossing the border from Turkey to trace their epidemiological patterns.

3. Methods

3.1. Study Design and Population

This cross-sectional study was carried out from July 2021 to December 2021. The participants were recruited by random sampling and categorized into two groups. The first group consisted of 700 residents of Duhok city, Iraq, who had not had recent travels. The second group consisted of 700 travelers who crossed the Turkish-Iraqi border of Ibrahim Al-Khalil. All participants had upper or lower respiratory infection symptoms. Exclusion criteria for both groups included previous COVID-19 vaccination and already contracting the natural infection. Nasopharyngeal and oropharyngeal swabs were obtained from all participants and transferred into a viral transport medium (VTM). The samples were kept at -70°C in the Central Public Health Laboratory of Duhok city, Iraq, till further processing.

3.2. Real-time PCR and Viral RNA Extraction

The QIAprep and Viral RNA UM Kit (i.e., Qiagen) was used for the detection and confirmation of the SARS-CoV-2 infection using nasopharyngeal samples. This kit combines RNA extraction with a real-time PCR-based reaction to detect the virus. The procedure was run according to the instructions provided by the manufacturer. Among positive samples for SARS-CoV-2 confirmed by real-time PCR from both groups, selected samples with a cycle threshold < 20 (15 samples from each group) were subjected to RNA extraction utilizing QIAamp Viral RNA Mini Kit from Qiagen. The samples were then shipped on dry ice to the INTERGEN Genetics and Rare Diseases Diagnosis Research and Application Center (Turkey) for next-generation sequencing.

3.3. Bioinformatic Analysis and Next Generation Sequencing

After the samples arrived at InterGen/Turkey, they were re-tested for RNA integrity and target confirmation. For reverse transcription and cDNA synthesis, the Ipsogen RT Kit and nanomere primers (Qiagen, Germany) were employed. Using an Illumina sample preparation kit, individual samples were indexed and tagged, and next-generation sequencing (NGS) was conducted using the Illumina Miseq instrument (Illumina, the US) in accordance with the manufacturer's instructions. Software from the Broad Institute's IGV2.8.9 was utilized to analyze sequencing data, and the BWA-MEM alignment technique was used to assemble and align short-read sequencing with the reference genome (NC045512.2) (19). Furthermore, the annotation of the assembled sequences was conducted using Annovar software according to a previous study (20), followed by data analysis in Lofreq (version 2) for mutation detection and variant calling (21). A quality check was conducted on the sequencing reads that were approved. The sequences analyzed exhibited a coverage rate of > 99% and gaps of 30 bps.

4. Results

Out of 1400 participants, 353 (25.21%) samples from both groups were demonstrated by RT-PCR to be positive. Among thirty samples with Ct values < 20 sent for the SARS-CoV-2/S gene sequencing, we retrieved the complete sequencing of 19 samples, comprising nine samples from Duhok residents and ten samples from travelers. The S gene sequence was obtained and submitted to the GISAID database and allotted the accession numbers of EPI_ISL_13971728 to EPI_ISL_13971746 (Appendix 1).

A total of 76 mutations, including 52 non-synonymous, 16 synonymous, and eight deletional mutations, were

detected in the sequenced samples, but no unique mutation was found (Appendix 2). Meanwhile, three well-known international platforms (GISAID, PANGO, and Nextstrain) were widely explored to categorize and separate SARS-CoV-2 isolates into different clades, variations, and lineages (22-24), leading to various naming conventions and the requirement for an all-encompassing unified platform. Using the GISAID system for genetic sequencing to classify isolates, the resulting clades included GK (58%), O (5%), GRA (26%), and G (11%) (Figure 1).

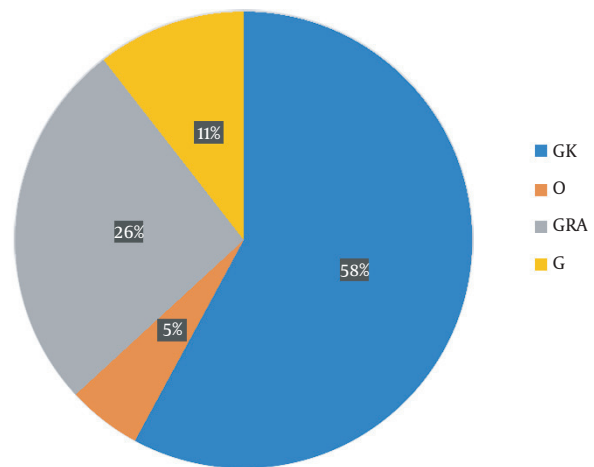


Figure 1. SARS-CoV-2 clades were identified in this study based on sequencing using the GISAID platform.

The variants identified were divided into four different groups as follows: 21J (delta), 21K (omicron), 21A (delta), and 20I (alpha V1), each accounting for 53%, 26%, 10%, and 11%, respectively (Figure 2). For sub-classification to lineages and sub-lineages, the PANGO system was utilized, detecting two B.1.1.7, three AY.78, four B.1.617.2, four AY.126, one AY.121, and five BA.1.1 variants (Figure 3).

4.1. Phylogenetic Analysis

In this study, a total of 48 sequences of the S gene in Iraqi populations and 41 sequences from other countries in the world were retrieved from the GISAID platform. These sequences were used for constructing a phylogenetic tree according to the sequenced isolates that were identified during the second and third waves, as well as at the beginning of the fourth epidemic wave of COVID-19 in the country. The GISAID database was used for phylogenetic evaluations, and a phylogenetic tree was generated (Figures 4 and 5).

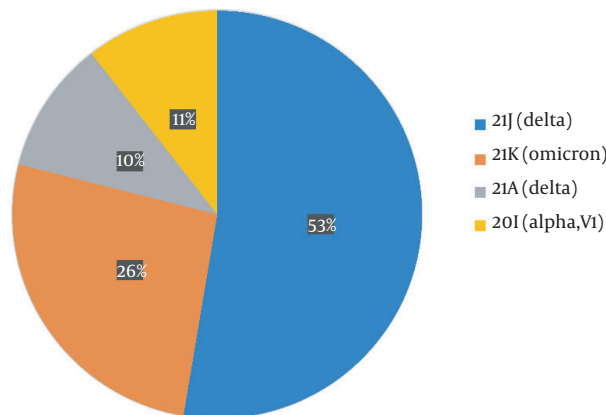


Figure 2. SARS-CoV-2 variants identified in this study based on sequencing using the Nextclade platform.

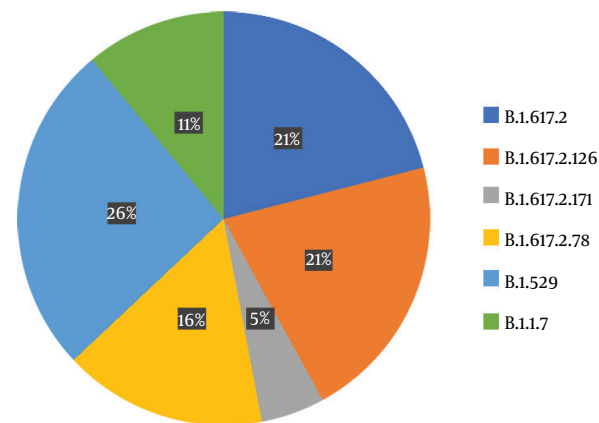


Figure 3. SARS-CoV-2 lineage identification in this study is based on sequencing using pandemic waves.

5. Discussion

With regard to COVID-19, the Iraqi Ministry of Health issued a warning stating that "the epidemiological situation has become dangerous," that the number of infections has been "rapidly increasing," and that the latest wave, or what it called the "third wave," was more severe than the previous one. The sharp increase in infections was attributed to peoples' indifference to preventive measures and lack of compliance with safety regulations, particularly the use of face masks and social distancing. In the present study, we investigated the molecular signatures and patterns of the mutations of the circulating strains of SARS-CoV-2 in Duhok Governorate, Iraq, during the end of the second wave and at the beginning of the third wave of the pandemic, when the

number of daily cases reached its highest level since the outbreak began. The worst wave of the pandemic hit the nation at the beginning of the summer of 2021. As a result, a study was necessary to track the evolution of the virus in the country. Compared to other countries, not many studies have been dedicated to tracking and categorizing SARS-CoV-2 variants in Iraq and in the city of Duhok in particular.

The GISAID and Pango lineages used in this study demonstrated that clade G was the most prevalent entity. The dominant genetic amino acid mutation has been found to be D614G, which is found in all sequenced strains. This mutation affects the receptor binding domain of the S spike protein that is responsible for the attachment of the virus to its receptors on host cells. At the same time, this site is where neutralizing antibodies act. In addition, this change seems to enhance the infectivity and stability of SARS-CoV-2 (25).

This study revealed that the GK clade (delta VOC) had the highest prevalence among locals and travelers in Duhok city, followed by the GRA clade (omicron VOC), which was only found in travelers crossing the border from Turkey, confirming the possibility of the transmission of this variant from neighbor countries through traveling. The phylogenetic tree showed high homogeneity between the strains sequenced from Duhok city and the variants registered in GISAID from other countries, including Turkey, Saudi Arabia, Germany, and Switzerland. In contrast to the other variants of SARS-CoV-2, the delta variant first appeared in India in late 2020, which was later found to have greater infectivity and transmissibility (26), as well as a shorter incubation period and the potential to evade neutralizing antibodies (27, 28).

Our analysis of the omicron VOC of SARS-CoV-2 sequences revealed numerous mutations in the S gene, which encodes a structural protein that serves as a viral protein binding to receptors on the host cell's surfaces and, therefore, determines the host range (29). A total of 31 non-synonymous mutations were discovered in this gene. This gene has been reported to be 4-5 times more likely to become mutated in comparison with other genes of the virus (30). Omicron strains share a dominant polymorphism known as D614-G in the Spike gene, which was discovered in all other VOCs (31). This polymorphism is linked with higher rates of infection and transmission, as well as viral escape from reactive antibodies (32). A mutation hot spot within the spike glycoprotein (known as omicron RBD), which is one of the primary targets for neutralizing antibodies, was shown to include 17 mutations, such as R346K, G339D, S371P, S373P, S371F, S375F, N440K, K417N, S477N, G446S, T478K, Q493R, E484A, Q498R, G496S, Y505H, and N501Y (33).

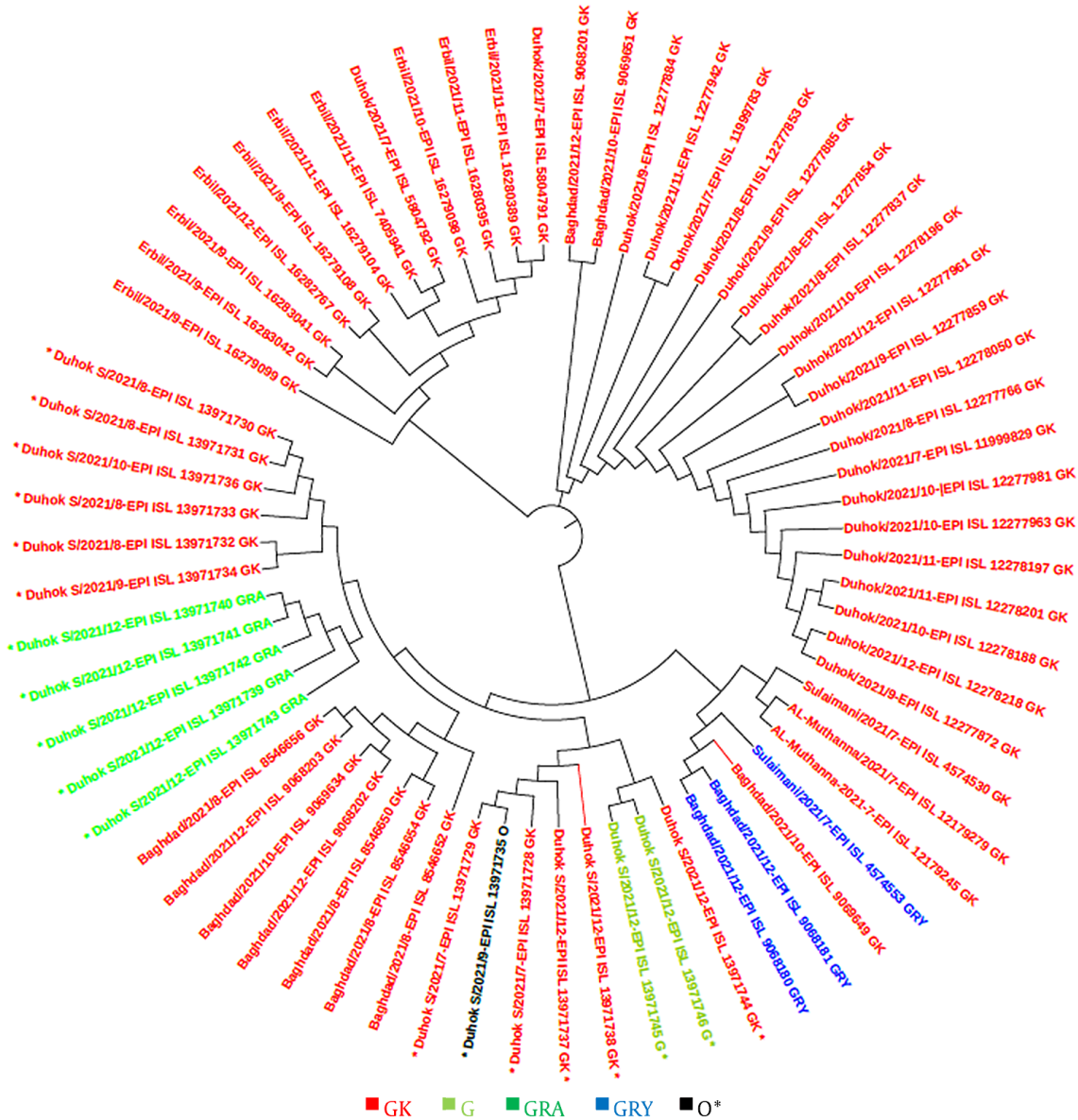


Figure 4. The phylogenetic tree representing the relationship between the SARS-CoV-2 S gene sequences detected in this study and various strains found in Iraqis (SARS-CoV-2 strain grouping was based on GISAID clades). *: Study samples.

Based on findings from earlier studies, omicron also possesses numerous novel mutations throughout the gene encoding the spike protein, including in the furin cleavage site, N-terminal domain (NTD), and S2 sub-unit. These mutations may have an impact on the spike protein's capability to bind the ACE-2 receptor and respond to antibodies (34, 35). Interestingly, the Ins215ped mutation already occurred 127 times (0.00% of all samples with the spike sequence) in 16 countries. The first strain

with this amino acid change, identified in December 2021, was hCoV-19/France/ARA-HCLO22005965401/2021, and most recently, it was detected in the hCoV-19/England/ALDP-37D7BF2/2022 strain in February 2022.

When compared to other VOCs, the SARS-CoV-2 genome represents a high rate of mutations, particularly those affecting the spike protein, which may boost viral transmission and immunological escape. Additionally,

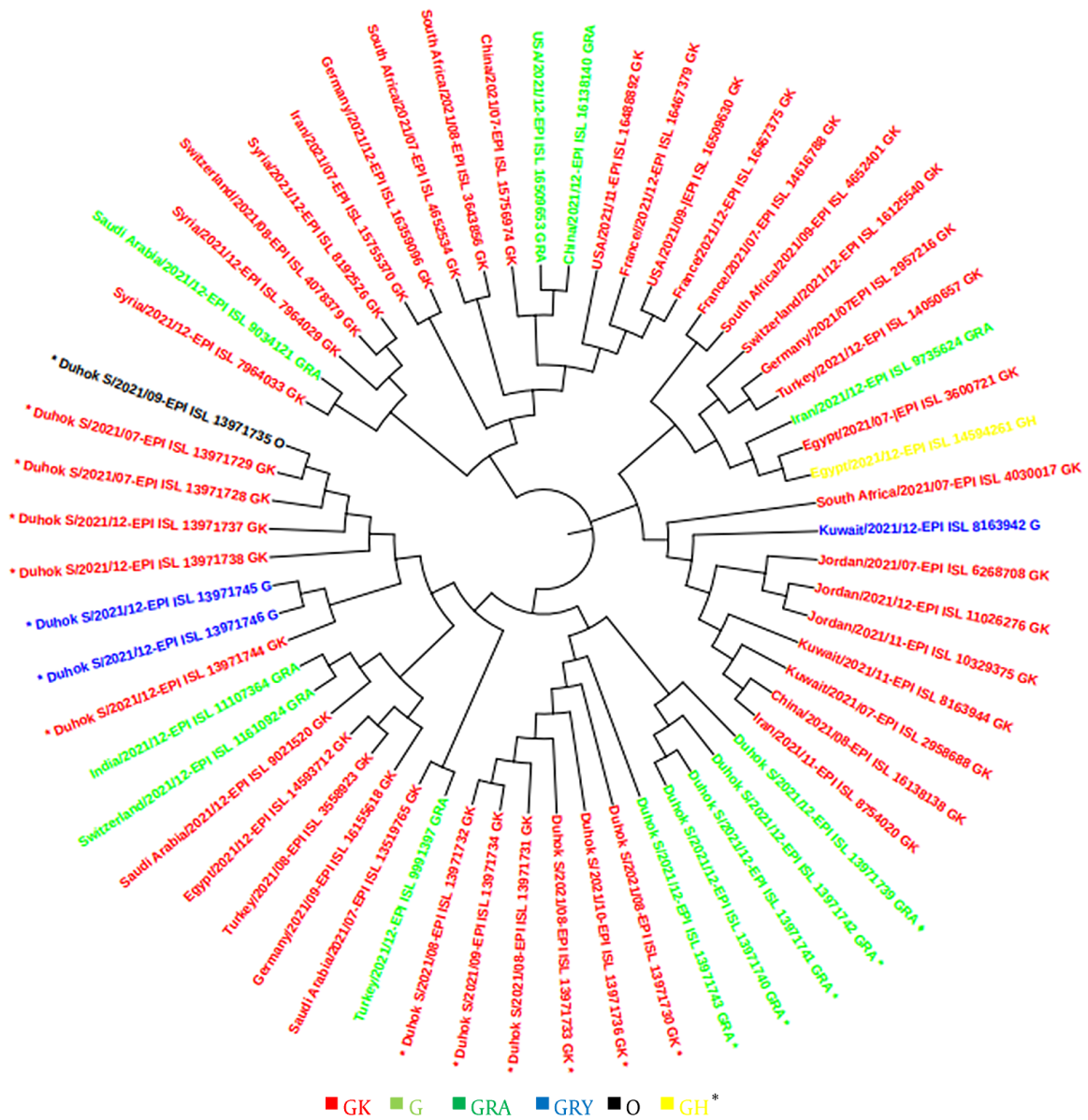


Figure 5. The phylogenetic tree represents the relationship between the SARS-CoV-2 S gene sequences identified in this study with the sequences reported from other countries. SARS-CoV-2 strains grouping was based on GISAID clades. *: Study samples.

the accumulation of numerous mutations on the immunogenic epitopes of the spike protein necessitates the production of novel vaccines using the omicron as a viable reference strain. More research is necessary to determine the efficacy of current vaccines against the omicron variant. There are a few deletions and more than 30 changes respective to the original sequence in the SARS-CoV-2 omicron variant (36), many of which seem to facilitate the spread of the virus and, actually, are the same

mutations found in previous SARS-CoV-2 variants, affecting binding affinity and transmissibility of the virus (37). The consequences of most other omicron mutations are unknown (38). The symptoms caused by the omicron and delta variants are different, and public health and medical professionals need to understand these differences. This not only foretells possible signs to look out for but also helps understand prognosis and outcomes.

5.1. Conclusions

The SARS-CoV-2 outbreak in late 2019 forced all countries' health systems to pay attention to important public health issues to fight the virus and stop its propagation. The incidence, distribution, and infectivity of COVs have increased in recent years because of the introduction of unique mutations in their genomes. Most novel COVs' variants have changes in the S1 domain of the spike protein, which is responsible for interaction with ACE2, rendering neutralizing antibodies (monoclonal antibodies, convalescent plasma, and sera from vaccinated people) less effective. Particularly, omicron is an emerging variant with around 30 changes in the amino acid sequence of the spike protein, enhancing its interaction with ACE II and removing the epitopes detected by neutralizing antibodies. Epidemiological monitoring is a crucial tactic for detecting new SARS-CoV-2 variants and characterizing them in terms of the fatality rate. The development of pan-coronavirus vaccines, as alternatives to present vaccines, and monoclonal antibody therapies based on new viral variations can reduce the mortality and morbidity inflicted by the SARS-CoV-2 infection.

Supplementary Material

Supplementary material(s) is available [here](#) [To read supplementary materials, please refer to the journal website and open PDF/HTML].

Footnotes

Authors' Contribution: Conceptualization and supervision: A. G., and A. P.; methodology: A. G., and O. Y.; validation, and writing, reviewing and editing: A. G., A. P., and O. Y.; software, investigation, data curation, and visualization: O. Y.; formal analysis, writing the original draft, and project administration: A. G.; resources: O. Y., and A. P.

Conflict of Interests: The authors declare no conflicts of interest.

Ethical Approval: The present study has been approved by the General Directorate of Health in Dohuk, Ministry of Health of the Republic of Iraq (reference number: 10-8-24102021).

Funding/Support: This research received no external funding.

Informed Consent: Informed consent was obtained from all participants enrolled in the study.

References

- Kumar D, Malviya R, Kumar Sharma P. Corona virus: A review of COVID-19. *Eurasian J Med Oncol*. 2020;**4**(1):8-25. <https://doi.org/10.14744/ejmo.2020.51418>.
- da Silva PG, Mesquita JR, de Sao Jose Nascimento M, Ferreira VAM. Viral, host and environmental factors that favor anthroponotic spillover of coronaviruses: An opinionated review, focusing on SARS-CoV, MERS-CoV and SARS-CoV-2. *Sci Total Environ*. 2021;**750**:141483. [PubMed ID: 32829257]. [PubMed Central ID: PMC7405882]. <https://doi.org/10.1016/j.scitotenv.2020.141483>.
- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with Pneumonia in China, 2019. *N Engl J Med*. 2020;**382**(8):727-33. [PubMed ID: 31978945]. [PubMed Central ID: PMC7092803]. <https://doi.org/10.1056/NEJMoa2001017>.
- Samanta S, Banerjee J, Rahaman SN, Ali KM, Ahmed R, Giri B, et al. Alteration of dietary habits and lifestyle pattern during COVID-19 pandemic associated lockdown: An online survey study. *Clin Nutr ESPEN*. 2022;**48**:234-46. [PubMed ID: 35331497]. [PubMed Central ID: PMC8848725]. <https://doi.org/10.1016/j.clnesp.2022.02.007>.
- Jee Y. WHO International Health Regulations Emergency Committee for the COVID-19 outbreak. *Epidemiol Health*. 2020;**42**:e2020013. [PubMed ID: 32192278]. [PubMed Central ID: PMC7285442]. <https://doi.org/10.4178/epih.e2020013>.
- Dhama K, Sharun K, Tiwari R, Dadar M, Malik YS, Singh KP, et al. COVID-19, an emerging coronavirus infection: Advances and prospects in designing and developing vaccines, immunotherapeutics, and therapeutics. *Hum Vaccin Immunother*. 2020;**16**(6):1232-8. [PubMed ID: 32186952]. [PubMed Central ID: PMC7103671]. <https://doi.org/10.1080/21645515.2020.1735227>.
- Khateeb J, Li Y, Zhang H. Emerging SARS-CoV-2 variants of concern and potential intervention approaches. *Crit Care*. 2021;**25**(1):244. [PubMed ID: 34253247]. [PubMed Central ID: PMC8274962]. <https://doi.org/10.1186/s13054-021-03662-x>.
- Das S, Kar SS, Samanta S, Banerjee J, Giri B, Dash SK. Immunogenic and reactogenic efficacy of Covaxin and Covishield: A comparative review. *Immunol Res*. 2022;**70**(3):289-315. [PubMed ID: 35192185]. [PubMed Central ID: PMC8861611]. <https://doi.org/10.1007/s12026-022-09265-0>.
- Malabadi RB, Kolkar KP, Meti NT, Chalannavar RK. Outbreak of coronavirus (SARS-CoV-2) Delta variant (B.1.617.2) and Delta Plus (AY. 1) with fungal infections, Mucormycosis: Herbal medicine treatment. *Int J Res Sci Innov*. 2021;**8**(6):59-70. <https://doi.org/10.51244/ijrsi.2021.8603>.
- Thakur V, Ratho RK.OMICRON (B.1.1.529): A new SARS-CoV-2 variant of concern mounting worldwide fear. *J Med Virol*. 2022;**94**(5):1821-4. [PubMed ID: 34936120]. <https://doi.org/10.1002/jmv.27541>.
- Patil BB. Omicron: The tsunami?? *Indian J Health Sci Biomed Res*. 2022;**15**(1):1-3. https://doi.org/10.4103/kleuhsj.kleuhsj_76_22.
- Flemming A. Omicron, the great escape artist. *Nat Rev Immunol*. 2022;**22**(2):75. [PubMed ID: 35017722]. [PubMed Central ID: PMC8749340]. <https://doi.org/10.1038/s41577-022-00676-6>.
- Brandal LT, MacDonald E, Veneti L, Ravlo T, Lange H, Naseer U, et al. Outbreak caused by the SARS-CoV-2 Omicron variant in Norway, November to December 2021. *Eurosurveillance*. 2021;**26**(50):pii=2101147. <https://doi.org/10.2807/1560-7917.es.2021.26.50.2101147>.
- Kozlov M. Omicron's feeble attack on the lungs could make it less dangerous. *Nature*. 2022;**601**(7892):177. [PubMed ID: 34987210]. <https://doi.org/10.1038/d41586-022-00007-8>.
- Wolter N, Jassat W, Walaza S, Welch R, Moultrie H, Groome M, et al. Early assessment of the clinical severity of the SARS-CoV-2 Omicron variant in South Africa. Preprint. *medRxiv*. Posted online December 21, 2021. <https://doi.org/10.1101/2021.12.21.21268116>.
- Nealon J, Cowling BJ. Omicron severity: Milder but not mild. *Lancet*. 2022;**399**(10323):412-3. [PubMed ID: 35065007]. [PubMed Central ID: PMC8769661]. [https://doi.org/10.1016/S0140-6736\(22\)00056-3](https://doi.org/10.1016/S0140-6736(22)00056-3).

17. Economic Times. *Immune response induced by Omicron effectively neutralise Delta variant: ICMR study*. 2022, [cited 12 March 2022]. Available from: <https://economictimes.indiatimes.com/industry/healthcare/biotech/healthcare/Immune-response-induced-by-omicroneffectively-neutralised-delta-variant-icmr-study/articleshow/89136215.cms>.
18. Roth C. *'Natural immunity': How effective is it against Omicron?* 2022, [cited 12 March 2022]. Available from: <https://www.dw.com/en/omicron-is-natural-immunity-better-than-a-vaccine/a-60425426>.
19. Li H. *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM*. 2013. Available from: <https://arxiv.org/abs/1303.3997>.
20. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;**38**(16):e164. [PubMed ID: 20601685]. [PubMed Central ID: PMC2938201]. <https://doi.org/10.1093/nar/gkq603>.
21. Wilm A, Aw PP, Bertrand D, Yeo GH, Ong SH, Wong CH, et al. LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res*. 2012;**40**(22):1189-201. [PubMed ID: 23066108]. [PubMed Central ID: PMC3526318]. <https://doi.org/10.1093/nar/gks918>.
22. Rambaut A, Holmes EC, O'Toole A, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol*. 2020;**5**(11):1403-7. [PubMed ID: 32669681]. [PubMed Central ID: PMC7610519]. <https://doi.org/10.1038/s41564-020-0770-5>.
23. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data - from vision to reality. *Eurosurveillance*. 2017;**22**(13):pii=30494. <https://doi.org/10.2807/1560-7917.es.2017.22.13.30494>.
24. O'Toole A, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol*. 2021;**7**(2):veab064. [PubMed ID: 34527285]. [PubMed Central ID: PMC8344591]. <https://doi.org/10.1093/ve/veab064>.
25. Mansbach RA, Chakraborty S, Nguyen K, Montefiori DC, Korber B, Gnanakaran G. The SARS-CoV-2 spike variant D614G favors an open conformational state. *Biophys J*. 2021;**120**(3):298A. <https://doi.org/10.1016/j.bpj.2020.11.1904>.
26. Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, et al. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Eurosurveillance*. 2021;**26**(24):pii=2100509. <https://doi.org/10.2807/1560-7917.es.2021.26.24.2100509>.
27. Grant R, Charmet T, Schaeffer L, Galmiche S, Madec Y, Von Platen C, et al. Impact of SARS-CoV-2 Delta variant on incubation, transmission settings and vaccine effectiveness: Results from a nationwide case-control study in France. *Lancet Reg Health Eur*. 2022;**13**:100278. [PubMed ID: 34849500]. [PubMed Central ID: PMC8616730]. <https://doi.org/10.1016/j.lanepe.2021.100278>.
28. Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, et al. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature*. 2021;**596**(7871):276-80. [PubMed ID: 34237773]. <https://doi.org/10.1038/s41586-021-03777-9>.
29. Amoutzias GD, Nikolaidis M, Tryfonopoulou E, Chlichlia K, Markoulatos P, Oliver SG. The remarkable evolutionary plasticity of coronaviruses by mutation and recombination: Insights for the COVID-19 pandemic and the future evolutionary paths of SARS-CoV-2. *Viruses*. 2022;**14**(1):78. [PubMed ID: 35062282]. [PubMed Central ID: PMC8778387]. <https://doi.org/10.3390/v14010078>.
30. Amicone M, Borges V, Alves MJ, Isidro J, Zé-Zé L, Duarte S, et al. Mutation rate of SARS-CoV-2 and emergence of mutators during experimental evolution. Preprint. *bioRxiv*. Posted online November 24, 2021. <https://doi.org/10.1101/2021.05.19.444774>.
31. Ahmad L. Implication of SARS-CoV-2 immune escape spike variants on secondary and vaccine breakthrough infections. *Front Immunol*. 2021;**12**:742167. [PubMed ID: 34804022]. [PubMed Central ID: PMC8596465]. <https://doi.org/10.3389/fimmu.2021.742167>.
32. 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;**182**(5):1284-94-e9. [PubMed ID: 32730807]. [PubMed Central ID: PMC7366990]. <https://doi.org/10.1016/j.cell.2020.07.012>.
33. Wang L, Cheng G. Sequence analysis of the emerging SARS-CoV-2 variant Omicron in South Africa. *J Med Virol*. 2022;**94**(4):1728-33. [PubMed ID: 34897752]. <https://doi.org/10.1002/jmv.27516>.
34. Li Y, Ma ML, Lei Q, Wang F, Hong W, Lai DY, et al. Linear epitope landscape of the SARS-CoV-2 Spike protein constructed from 1,051 COVID-19 patients. *Cell Rep*. 2021;**34**(13):108915. [PubMed ID: 33761319]. [PubMed Central ID: PMC7953450]. <https://doi.org/10.1016/j.celrep.2021.108915>.
35. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol*. 2021;**19**(7):409-24. [PubMed ID: 34075212]. [PubMed Central ID: PMC8167834]. <https://doi.org/10.1038/s41579-021-00573-0>.
36. Global Initiative on Sharing All Influenza Data. *Tracking of hCoV-19 Variants*. 2022, [cited 27 January 2022]. Available from: <https://gisaid.org/hcov19-variants/>.
37. Greaney AJ, Starr TN, Gilchuk P, Zost SJ, Binshtein E, Loes AN, et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe*. 2021;**29**(1):44-57-e9. [PubMed ID: 33259788]. [PubMed Central ID: PMC7676316]. <https://doi.org/10.1016/j.chom.2020.11.007>.
38. Karim SSA, Karim QA. Omicron SARS-CoV-2 variant: A new chapter in the COVID-19 pandemic. *Lancet*. 2021;**398**(10317):2126-8. [PubMed ID: 34871545]. [PubMed Central ID: PMC8640673]. [https://doi.org/10.1016/S0140-6736\(21\)02758-6](https://doi.org/10.1016/S0140-6736(21)02758-6).