



The Brazilian Journal of INFECTIOUS DISEASES

www.elsevier.com/locate/bjid



Original Article

Sequencing and mutations analysis of the first recorded SARS-CoV-2 Omicron variant during the fourth wave of pandemic in Iraq

Jivan Qasim Ahmed ^{a,*}, Sazan Qader Maulud ^b, Rawand Al-Qadi ^c,
Teroj Abdulrahman Mohamed ^d, Gahin Abdulraheem Tayib ^a,
Akheenk Mustafa Hassan ^d, Luqman Saleh Taha ^c, Khairi Mohammed Qasim ^c,
Mohammed Abid Tawfeeq ^c

^a University of Duhok, Department of Pathology and Microbiology, Kurdistan Region, Iraq

^b Salahaddin University-Erbil, College of Education, Department of Biology, Kurdistan Region, Iraq

^c Directorate General of Health, Kurdistan Region, Iraq

^d University of Duhok, Dental Basic Science Department, Kurdistan Region, Iraq

ARTICLE INFO

Article history:

Received 5 May 2022

Accepted 25 July 2022

Available online 8 August 2022

Keywords:

COVID-19

SARS-CoV-2

Variant of concern

Omicron variant

ABSTRACT

Despite vaccine development and vaccination programs underway around the globe, the coronavirus disease 2019 (COVID-19) pandemic has not been controlled as the SARS-CoV-2 virus is evolving and new variants are emerging. This study was conducted to sequence and molecularly characterize the representing samples from the early fourth SARS-CoV-2 wave in Iraq. Here, we have performed next-generation sequencing of whole-genome sequencing of two representing samples from the country's early beginning of the fourth pandemic wave. The samples were sequenced using Illumina Miseq system, and the reference sequences were retrieved from GISAID database. Phylogenetic analysis was performed through Mega software. This study provides an initial sequence analysis and molecular characterization of the first Omicron variant cases recorded in the country. Our analysis revealed many mutations on the spike glycoprotein, especially on the receptor binding domain, with potential impact on immune escape and infectivity. The study findings suggest considering the highly mutated immunogenic epitope of the Omicron variant as a reference for developing a new vaccine for combating the ongoing pandemic.

© 2022 Sociedade Brasileira de Infectologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Introduction

Approximately two years have passed since Wuhan's first contagious agent was reported. The pandemic coronavirus disease 2019 (COVID-19), caused by the positive sense single-

* Corresponding author.

E-mail address: jivan.ahmed@uod.ac (J.Q. Ahmed).

<https://doi.org/10.1016/j.bjid.2022.102677>

1413-8670/© 2022 Sociedade Brasileira de Infectologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

stranded RNA virus denoted as SARS-CoV-2.¹ Despite the continuous emergence of new variants of the virus and the short-lived nature of vaccine-induced immunity, the peak of the daily new cases was declining, and the results raised hope that the coronavirus pandemic could be contained. Consequently, the current pandemic could get one step closer to an end after vaccine development.² Though globally, COVID-19 led to significant burden of illness and unpredictable death in different parts of the world.¹

Presently, the world faces another recent variant that led researchers to attempt to collect primary data on that new variant initiated pathogenesis and its mechanisms. It has been revealed that Omicron variants with new mutations could affect the transmissibility of the virus and the host's immunity.³

The SARS-CoV-2 Omicron variant was first recorded in South Africa on November 24th, 2021, and was assigned as a variant of concern (VOC) within two days by the World Health Organization (WHO).⁴ The Omicron variant has the largest number of mutations among all other SARS-CoV-2 variants, harbouring more than 50 mutations on the virus genome. S gene was found to be the most mutated gene, and 15 of those mutations are on the receptor-binding domain (RBD) of the virus.⁵ The S gene codes for Spike glycoprotein that is used by the virus to interact with the cellular angiotensin converting-2 enzyme (ACE2).⁶ It was suggested that such mutations in the Spike glycoprotein might lead to escape from the developed immune response due to infection or vaccination and may cause many breakthrough infections or reinfection with mutated viral strains.^{7,8}

Like other places, as of January 25th, 2022, the Kurdistan region had accumulated laboratory-confirmed cases of SARS-CoV-2 omicron variant; presently, it is considered the fourth wave of infection. The first wave started in early March 2020 and was followed by a wave that topped in July 2020 and ended in September.⁹ The second wave peaked in January 2021 and ended in March 2020. The third wave, which peaked in July and ended in September 2021, was overwhelmed by the Delta variant. In late December 2021, the Omicron variant was detected in Duhok Province and associated with rapidly increasing cases; the Kurdistan region is in the red zone of infection.

However, the number of shared sequences diagnosed from Iraq in the international database is still scarce (GISAID, <https://www.gisaid.org/>). Mutations of SARS-CoV-2 may happen over time to escape the host immune response, especially with the robust immune pressure in humans.¹⁰ This possibly leads to the emergence of new virus variants with higher infectivity, transmissibility, and virulence potentials.¹¹

Herein, we aimed to perform whole-genome sequencing for the samples of SARS-CoV-2 from the early beginning of the fourth wave of the pandemic in Iraq, in order to characterize the mutation patterns of the Omicron variant in the country.

Materials and methods

Samples

After the emergence of the omicron variant in South Africa and its rapid spread to other countries worldwide, the daily

cases in the Kurdistan region started to surge again; this has brought about the need to sequence representative new cases and identify the circulating variant and its mutation patterns. Two positive samples that were collected on December 21st 2021 for routine SARS-CoV-2 testing at Duhok Central public health laboratory have been selected for whole-genome sequencing.

Viral RNA extraction and Real-Time PCR

The selected samples were subjected to extraction and quantitative detection using RNA extraction and real-time PCR-based virus detection system QIAprep & amp Viral RNA UM Kit (Qiagen). Ct value of selected samples was (< 20) as demanded by the sequencing laboratory. The RNA was then extracted from the viral transport medium by QIAamp Viral RNA Mini Kit (Qiagen), and transported to the Intergen commercial company (Ankara, Turkey) on dry ice for whole-genome sequencing.

SARS-CoV-2 Next Generation Sequencing (NGS) Bioinformatics Analysis

The samples were retested for target confirmation and RNA integrity following arrival at Intergen. Ipsogen RT Kit and nanomere primers from (Qiagen, Germany) were used for reverse transcription and cDNA synthesis. Indexing and tagmentation of the individual samples was performed using illumina sample preparation kit. Next-generation sequencing was performed using illumina Miseq instrument and equipment (Illumina, San Diego, CA, USA) following manufacturers' instructions. IGV 2.8.9 (Broad Institute) software was used to evaluate sequencing data. The short reads sequencing was assembled and aligned with the reference genome (NC_045512.2) through BWA-MEM alignment algorithm.¹² The annotation of the assembled consensus sequences was completed by Annotvar.¹³ Together, Lofreq (version 2) was used for the identification of mutation and variant calling.¹⁴ The accepted sequencing read underwent quality control. The sequences had coverage of (> 99%) and gap length of (< 30) bps that were subjected to further analysis. The obtained whole-genome sequence was submitted to the GISAID database, and the assigned accession numbers were received (EPI_ISL_9084009, EPI_ISL_9765388).

Variants, lineage and phylogenetic analysis

The whole-genome sequences of this study (n = 2), samples from the surrounding countries, and blast search result sequences around the world have been retrieved from GISAID and the National Centre for Biotechnology Information data banks. The criteria considered for sequence selection were good quality sequences, without running Ns, collected during December 2021, and the characterized as Omicron variants. The developed dataset has been subjected to quality control, and low-quality sequences have been excluded. The phylogenetic tree was constructed to reveal the relationship between the sequences of this data set and track down the source of the emerging Omicron variant in Iraq. The phylogenetic tree was built using the neighbour-joining method implemented

in the Molecular Evolutionary Genetics Analysis program (MEGA 11).¹⁵ The variants and lineage identification were made using the Nextclade sequence analysis and Pangolin system version (v3.1.14), respectively.^{16,15}

Results

Two isolates of this study were found to be Omicron VOC using Nextclade platform,¹⁶ and the Pangolin tool has identified these sequences as BA.1.1 lineage of Omicron.¹⁵ The mutation analysis has revealed that Iraqi isolates of this study have a total of 62 different mutations; among these, a total of 45 mutations were non-synonymous, 10 synonymous, one upstream, and deletions at six different positions that resulted in the deletion of 39 single nucleotides (Table 1). Interestingly, 34 (54.8%) of these mutations were found on Spike protein, 33 non-synonymous mutations, followed by 15 (24.2%) mutations in the ORF1ab region that constituting of total mutations. Other genomic regions N, ORF3a, E, M, ORF6, and ORF7b were less frequently mutated. Non-synonymous mutations on the SARS-CoV-2 genome revealed that the mutations hot spot was Spike protein (Fig. 1). The synonymous mutations at nucleotide position 3373 C > T and 16744 G > A were found to be a signature of the two isolates of our study that were not found in the tested data set. The phylogenetic analysis for the two isolates of this study and retrieved sequences from both GISAID and NCBI databases revealed several sub-cladding sequences of the tested data set (Fig. 2). The sequences of this study were found to group with isolates from USA making one to speculate that active international travel could have been the potential source of the Omicron found in Iraq.

Discussion

In a time of global efforts to develop antivirals and distribute the vaccine, daily cases are rising due to the emergence of new variants in some countries. While the delta variant remains circulating worldwide, the recently discovered Omicron VOC has revealed a tendency to become the upcoming dominant variant. In this study, we are showing a sequence analysis and the mutation pattern of the first Omicron variant sequences from cases recorded in Iraq in December 2021 that subsequently led to the announcement of the fourth SARS-CoV-2 wave in the country by the Ministry of Health (Fig. 3).

The analyzed Omicron VOC of SARS-CoV-2 sequences of this study showed many mutations on S-gene, followed by ORF1ab, N-gene, M-gene. ORF3a, E-gene, ORF6, and ORF7b were the lowest mutated genes. The most mutated S-gene, which encodes a structural protein, functions as a viral binding protein to the host cell receptors and determines the host range.¹⁷ A total of 30 non-synonymous mutations have been found on this gene, including A67V, T95I, G339D, R346K, S371P, S371F, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H,

N969K, L981F, and only D1146D as synonymous mutation along with 68_70del, 142_145del, and 211_212del deletions. This gene was reported to have 4-5 times more mutation potential compared to other genomic positions.¹⁸

It remains uncertain how the Omicron VOC has acquired such vast number of mutations, specifically on the Spike protein. Since this variant was first detected in an immunocompromised patient in South Africa, less selective pressure was on the virus. Still, the prolonged duration of the infection could be the reason for such a huge evolution. In general, an increasing and persistent selective pressure on the virus on both population and individual levels speed up the viral evolution process. The dominant variant D614G in the Spike gene has also been conserved for the Omicron variant; this variant has been found in all other VOCs,¹⁹ which is related to increased transmission and infection rates and viral escape from reactive antibodies.²⁰ The receptor-binding domain (RBD) of the spike protein, which is a primary target for neutralizing antibodies, has been shown to be a mutation hot spot subunit within the Spike glycoprotein having 17 mutations, including G339D, R346K, S371P, S371F, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H. Novel mutations of RBD of Omicron were G339D, R346K, S371P, S371F, S373P, S375F, N440K, G446S, Q493R, G496S, Q498R, Y505H mutations, in addition to other mutations that had been acquired from previous VOCs. Mutations like K417N, N440K, G446S, T478K, E484A, and H69/V70 deletions found previously in other VOCs have been associated with escape from the neutralizing antibodies and immune escapes.⁵ Furthermore, Omicron has many novel mutations all over the spike protein like N-terminal domain (NTD), furin cleavage site, and S2 subunit that could potentially affect binding to ACE2 receptor and reaction to antibodies based on results from previous reports.^{21,22}

ORF1ab, which constitutes 2/3 of the entire genome, was reported to be a mutation hot spot in all other VOCs,^{23,24} while the Omicron sequences of this study in this region were found to be less mutated (24.2%) compared to Spike protein. The ORF1ab gene functions in viral replication by coding for a helicase protein and RNA-dependent RNA polymerase enzyme.²⁵ Among the non-structural proteins of ORF1ab, nsp3 was most mutated with K38R and A1892T variants and deletion at 1265_1266del amino acid positions. Variant T492I was detected on nsp4 and I189V with non-frame shift deletion (104_107del) on nsp6. These mutations and deletions could affect replication and evolution processes. It is believed that nsp3, nsp4, and nsp6 are responsible for double-membrane activation that enables transcription and replication activity²⁶ and allows the evolution process due to selective pressure on the virus.²⁷

N gene encodes the nucleocapsid protein, which is also involved in the replication and transcription process.²⁸ This protein has conserved mutations in three consecutive nucleotides at position 28881-28883 (GGG > AAC) that resulted in a double amino acid change (R203K G204R); these variants were detected at early waves of the pandemic and have been conserved in subsequent emerging variants around the world.^{23,29,30} Nevertheless, ORF6 and ORF7 are accessory

Table 1 – Total mutations, effect on amino acid, and distribution on genomic positions of SARS-CoV-2 (Omicron variant) isolates of this study.

	Gene	Start	End	Ref	Alt	Protein Change	Type of mutation
1	5UTR	241	241	C	T		Upstream
2	ORF1a;ORF1ab;nsp3	2832	2832	A	G	K38R	Non-synonymous SNV
3	ORF1a;ORF1ab;nsp3	3037	3037	C	T	F106F	synonymous SNV
4	ORF1a;ORF1ab;nsp3	3373	3373	C	T	D218D	synonymous SNV
5	ORF1a;ORF1ab;nsp3	5386	5386	T	G	A889A	synonymous SNV
6	ORF1a;ORF1ab;nsp3	6513	6515	GTT	-	1265_1266del	Non-frame shift deletion
7	ORF1a;ORF1ab;nsp3	8393	8393	G	A	A1892T	Non-synonymous SNV
8	ORF1a;ORF1ab;nsp4	10029	10029	C	T	T492I	Non-synonymous SNV
9	ORF1a;ORF1ab;nsp5	10449	10449	C	A	P132H	Non-synonymous SNV
10	ORF1a;ORF1ab;nsp6	11285	11293	TTGT5CTGGT	-	104_107del	Non-frame shift deletion
11	ORF1a;ORF1ab;nsp6	11537	11537	A	G	I189V	Non-synonymous SNV
12	ORF1a;ORF1ab;nsp10	13195	13195	T	C	V57V	Synonymous SNV
13	ORF1ab;nsp12	14408	14408	C	T	P323L	Non-synonymous SNV
14	ORF1ab;nsp12	15240	15240	C	T	N600N	Synonymous SNV
15	ORF1ab;nsp13	16744	16744	G	A	G5494S	Non-synonymous SNV
16	ORF1ab;nsp14	18163	18163	A	G	I5967V	Non-synonymous SNV
17	S	21762	21762	C	T	A67V	Non-synonymous SNV
18	S	21765	21770	TACATG	-	68_70del	Non-frame shift deletion
19	S	21846	21846	C	T	T95I	Non-synonymous SNV
20	S	21987	21995	GTGTTTATT	-	142_145del	Non-frame shift deletion
21	S	22194	22196	ATT	-	211_212del	Non-frame shift deletion
22	S	22578	22578	G	A	G339D	Non-synonymous SNV
23	S	22599	22599	G	A	R346K	Non-synonymous SNV
24	S	22673	22673	T	C	S371P	Non-synonymous SNV
25	S	22674	22674	C	T	S371F	Non-synonymous SNV
26	S	22679	22679	T	C	S373P	Non-synonymous SNV
27	S	22686	22686	C	T	S375F	Non-synonymous SNV
28	S	22813	22813	G	T	K417N	Non-synonymous SNV
29	S	22882	22882	T	G	N440K	Non-synonymous SNV
30	S	22898	22898	G	A	G446S	Non-synonymous SNV
31	S	22992	22992	G	A	S477N	Non-synonymous SNV
32	S	22995	22995	C	A	T478K	Non-synonymous SNV
33	S	23013	23013	A	C	E484A	Non-synonymous SNV
34	S	23040	23040	A	G	Q493R	Non-synonymous SNV
35	S	23048	23048	G	A	G496S	Non-synonymous SNV
36	S	23055	23055	A	G	Q498R	Non-synonymous SNV
37	S	23063	23063	A	T	N501Y	Non-synonymous SNV
38	S	23075	23075	T	C	Y505H	Non-synonymous SNV
39	S	23202	23202	C	A	T547K	Non-synonymous SNV
40	S	23403	23403	A	G	D614G	Non-synonymous SNV
41	S	23525	23525	C	T	H655Y	Non-synonymous SNV
42	S	23599	23599	T	G	N679K	Non-synonymous SNV
43	S	23604	23604	C	A	P681H	Non-synonymous SNV
44	S	23854	23854	C	A	N764K	Non-synonymous SNV
45	S	23948	23948	G	T	D796Y	Non-synonymous SNV
46	S	24130	24130	C	A	N856K	Non-synonymous SNV
47	S	24424	24424	A	T	Q954H	Non-synonymous SNV
48	S	24469	24469	T	A	N969K	Non-synonymous SNV
49	S	24503	24503	C	T	L981F	Non-synonymous SNV
50	S	25000	25000	C	T	D1146D	synonymous SNV
51	ORF3a	25584	25584	C	T	T64T	synonymous SNV
52	E	26270	26270	C	T	T9I	Non-synonymous SNV
53	M	26530	26530	A	G	D3G	Non-synonymous SNV
54	M	26577	26577	C	G	Q19E	Non-synonymous SNV
55	M	26709	26709	G	A	A63T	Non-synonymous SNV
56	ORF6	27259	27259	A	C	R20R	Synonymous SNV
57	ORF7b	27807	27807	C	T	L18L	Synonymous SNV
58	N	28311	28311	C	T	P13L	Non-synonymous SNV
59	N	28362	28370	GAGAACGCA	-	30_33del	Non-frame shift deletion
60	N	28881	28881	G	A	R203K	Non-synonymous SNV
61	N	28882	28882	G	A	R203R	Synonymous SNV
62	N	28883	28883	G	C	G204R	Non-synonymous SNV

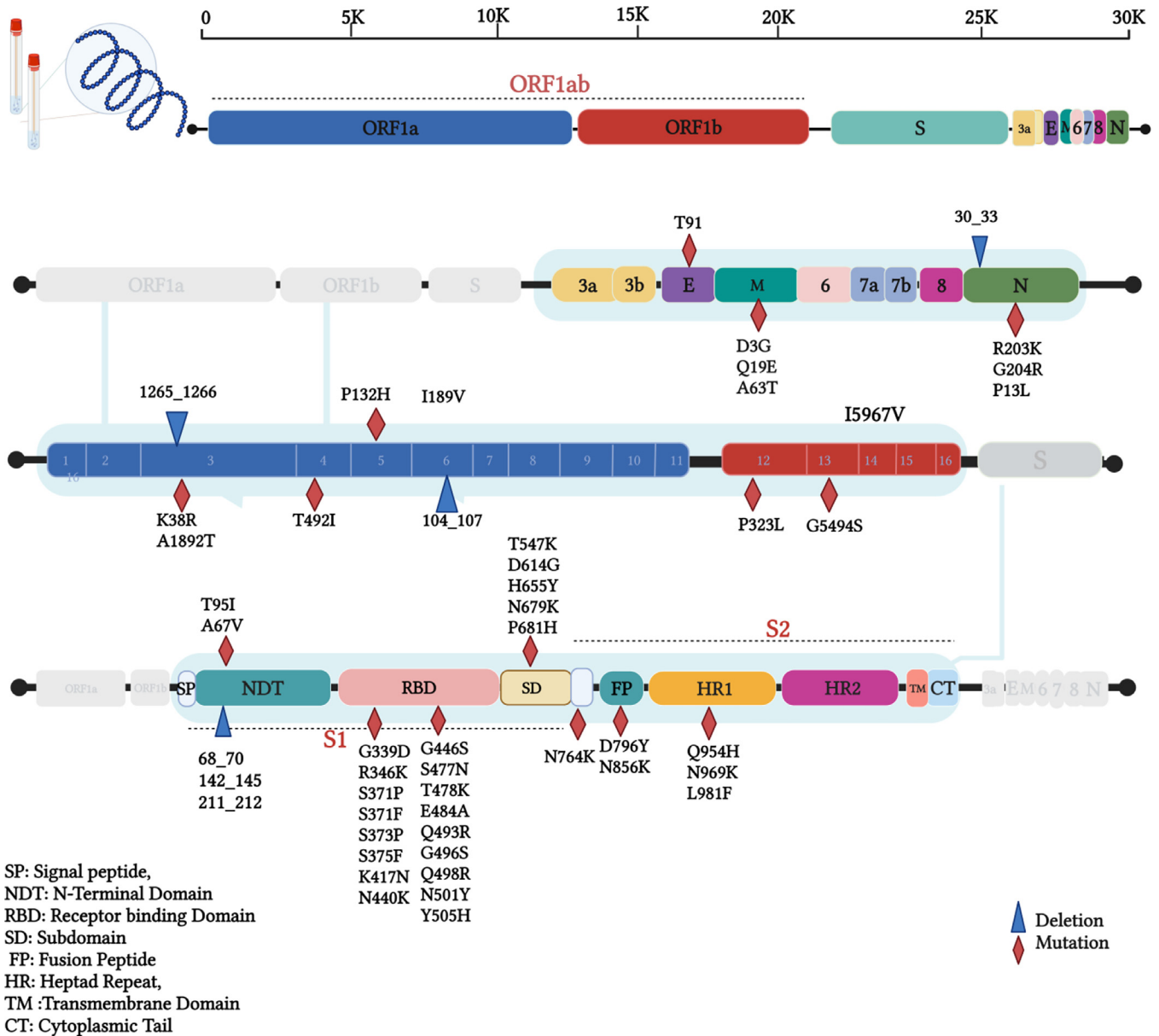


Fig. 1 – Illustration of the mutations and their positions on the Omicron variant isolates of this study.

proteins that have a role in suppressing innate immunity, signalling pathways, and interferon (IFN) expression.³¹ Duhok sequences have shown less frequent mutations on these two proteins, and the two nucleotide changes (27259 A > C and 27807 C > T) were silent; these data are in line with previously analyzed sequences of the other VOCs.³²

Interestingly, a silent mutation 3373 C>T was found to be a signature of our two sequences and had not been identified in any other sequences of our data set. At the same time, non-synonymous mutation G5494S of our isolates was found as a notable change only found in sequences from USA and Germany. Nevertheless, phylogenetic analysis of our data set has shown that our isolates were clustered with Omicron sequences from the USA, suggesting the source of the first

Omicron emergence in the country through active international travelling.

This high number of mutations on the SARS-CoV-2 genome, specifically on spike protein, could increase viral transmission and immune escape compared to other VOCs. Furthermore, the accumulation of these bulky mutations on the immunogenic epitopes of Spike protein brought about the necessity of developing new vaccines that include the Omicron variant as a potential reference strain. Meanwhile, further studies are required to identify the infectivity and effectiveness of available vaccines against the Omicron variant.

In conclusion, this study provides an analysis of the first recorded Omicron variant sequences in Iraq that have caused

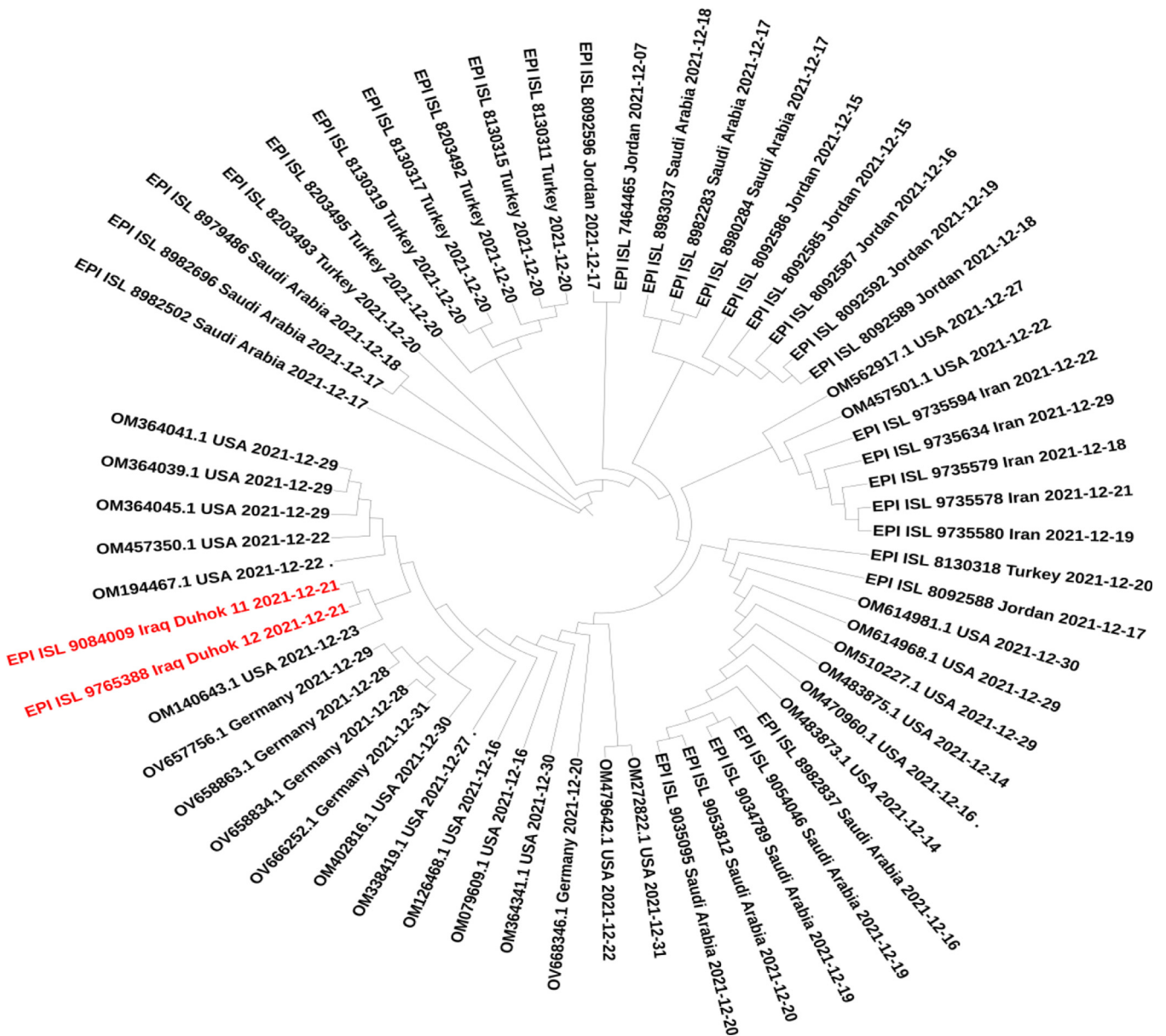


Fig. 2 – Phylogenetic tree constructed using two Omicron isolates of this study and a selected data set that includes sequences from blast search and neighbouring countries.

a sharp increase in daily cases and the fourth wave of the pandemic announcement in the country. This work represents the first publication of Omicron sequence analysis in the country. One mutation was a unique signature of our isolates, in addition to a rare mutation only found in two sequences from the USA and Germany. These findings provide data on the mutation pattern in circulating variants in the country, which could help public health authorities in issuing and updating the control roles for SARS-CoV-2 emergence. Here we recommend future studies focus on the impact and function of such genomic variants on the virus's infectivity, pathogenesis, and severity. Nevertheless,

developing a new vaccine, especially multivalent vaccines containing multiple VOCs, could be a good step forward in controlling the latest infections.

Ethics approval

Ethical Committee of the Directorate General of Health in Duhok has approved this study. The next-generation sequencing of the SARS-CoV-2 positive samples has been done after the cases underwent regular SARS-CoV-2 diagnostic tests at Duhok Central public health laboratory.

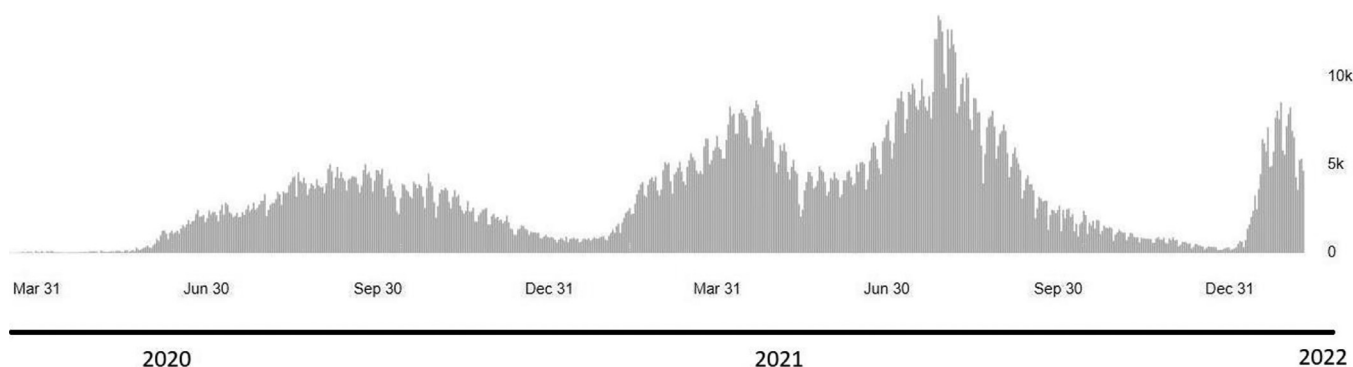


Fig. 3 – SARS-CoV-2 reported daily cases in Iraq, data reported by the WHO Coronavirus dashboard.

Availability of data and materials

The SARS-CoV-2 sequences of this study are deposited in the Global Initiative on Sharing all Individual Data (GISAID).

Funding

No funding has been received.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors acknowledge the Duhok Central public health laboratory for the valuable support of this study.

REFERENCES

- Zhou P, Yang X-L, Wang X-G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579:270–3.
- Gazit S, Shlezinger R, Perez G, et al. Comparing SARS-CoV-2 natural immunity to vaccine-induced immunity: reinfections versus breakthrough infections. *medRxiv*. 2021. <https://doi.org/10.1101/2021.08.24.21262415>.
- Harvey WT, Carabelli AM, Jackson B, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol*. 2021;19:409–24.
- WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data. <https://covid19.who.int/>.
- Wang L, Cheng G. Sequence analysis of the emerging SARS-CoV-2 variant Omicron in South Africa. *J Med Virol*. 2022;94:1728–33.
- Lauber C, Goeman JJ, Parquet M del C, et al. The footprint of genome architecture in the largest genome expansion in RNA viruses. *Stern A, editor PLoS Pathog*. 2013;9:e1003500.
- Callaway E. Heavily mutated Omicron variant puts scientists on alert. *Nature*. 2021;600:21.
- Li Q, Nie J, Wu J, et al. SARS-CoV-2 501Y.V2 variants lack higher infectivity but do have immune escape. *Cell*. 2021;184:2362–2371.e9.
- Qadir Maulud S, Omar Majed S, Abdulqadir Ali B, Jamal Jalal P, Hasan Azeez S, Abdulmuhsin Mohammad K. Epidemiological approach of SARS-CoV2 in the first month of appearance in the Kurdistan Region of Iraq. *Eur J Mol Clin Med*. 2021;7:2853–65.
- Molina-Mora JA, Cordero-Laurent E, Godínez A, et al. SARS-CoV-2 genomic surveillance in Costa Rica: evidence of a divergent population and an increased detection of a spike T1117I mutation. *Infect Genet Evol*. 2021;92:104872.
- Nguyen TT, Pathirana PN, Nguyen T, et al. Genomic mutations and changes in protein secondary structure and solvent accessibility of SARS-CoV-2 (COVID-19 virus). *Sci Rep*. 2021;11(1):3487.
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Oxford Univ Press; 2013 <https://www.scienceopen.com/document?vid=e623e045-f570-42c5-80c8-ef0aea06629c>.
- Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164–e164.
- Wilm A, Aw PPK, Bertrand D, 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;40:11189–201.
- O'Toole Á, Scher E, Underwood A, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evol*. 2021;7(2).
- Aksamentov I, Roemer C, Hodcroft E, Neher R. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *J Open Source Softw*. 2021;6:3773.
- 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.
- Amicone M, Borges V, João Alves M, et al. Mutation rate of SARS-CoV-2 and emergence of mutators during experimental evolution. *bioRxiv*. 2021. <https://doi.org/10.1101/2021.05.19.444774>.
- Ahmad L. Implication of SARS-CoV-2 Immune Escape Spike Variants on Secondary and Vaccine Breakthrough Infections. *Front Immunol*. 2021;12:1–7.
- Li Q, Wu J, Nie J, et al. The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell*. 2020;182:1284–1294.e9.
- Li Y, Ma M, Lei Q, 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. <https://doi.org/10.1016/j.celrep.2021.108915>.
22. Harvey WT, Carabelli AM, Jackson B, et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol.* 2021;19:409–24.
 23. Ahmad J, Tayib G, Mohamed T. SARS-CoV-2 mutation hotspots incidence in different geographic regions. *Microb Biosyst.* 2020;5:1–8.
 24. Vidanović D, Tešović B, Volkening JD, et al. First whole-genome analysis of the novel coronavirus (SARS-CoV-2) obtained from COVID-19 patients from five districts in Western Serbia. *Epidemiol Infect.* 2021;149:e246.
 25. Graham RL, Sparks JS, Eckerle LD, Sims AC, Denison MR. SARS coronavirus replicase proteins in pathogenesis. *Virus Res.* 2008;133:88–100.
 26. Sakai Y, Kawachi K, Terada Y, Omori H, Matsuura Y, Kamitani W. Two-amino acids change in the nsp4 of SARS coronavirus abolishes viral replication. *Virology.* 2017;510:165–74.
 27. Wu C, Liu Y, Yang Y, et al. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm Sin B.* 2020;10:766–88.
 28. McBride R, van Zyl M, Fielding B. The Coronavirus nucleocapsid is a multifunctional protein. *Viruses.* 2014;6:2991–3018. 1.
 29. Sahin E, Bozdayi G, Yigit S, et al. Genomic characterisation of SARS-CoV-2 isolates from patients in Turkey reveals the presence of novel mutations in spike and nsp12 proteins. *J Med Virol.* 2021;93:6016–26.
 30. Rahman MM, Kader SB, Rizvi SMS. Molecular characterisation of SARS-CoV-2 from Bangladesh: implications in genetic diversity, possible origin of the virus, and functional significance of the mutations. *Heliyon.* 2021;7:e07866.
 31. Frieman M, Yount B, Heise M, Kopecky-Bromberg SA, Palese P, Baric RS. Severe acute respiratory syndrome coronavirus ORF6 antagonises STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/golgi membrane. *J Virol.* 2007;81:9812–24.
 32. Miljanovic D, Milicevic O, Loncar A, Abazovic D, Despot D, Banko A. The first molecular characterization of Serbian SARS-CoV-2 isolates from a Unique Early Second Wave in Europe. *Front Microbiol.* 2021;12:1–11.