Molecular Biology Reports
https://doi.org/10.1007/s11033-023-08271-1

ORIGINAL ARTICLE

Identification of novel mutations in β-thalassemia patients in Maysan Governorate, Iraq

Abbas Kadhim Jiheel AL-hameedawi[1] · **Ali A. A. Al-Shawi**[1]

Received: 22 November 2022 / Accepted: 10 January 2023

Abstract

Background In homozygous β-thalassemia, the primary genetic modifiers affecting the clinical severity of β-thalassemia are genetic variants and the ability to reduce globin chain imbalance, thus resulting in a milder form of thalassemia. However, there are few reports on the molecular genetics of β-thalassemia in Iraq.

Methods We performed PCR and DNA sequencing on 40 Iraqi patients who were clinically suspected of having β-thalassemia.

Results The first genetic sequencing study was conducted in Maysan Governate, Iraq, using patients from various locations to identify novel mutations. There were five novel mutations: 294.T>C 12% (city center and Almajar district), 205. C>T 25% (city center, Alsalam, and Almashrah districts), 289.G>A 38% (Almaymuna and Gleat Salih districts), 49.T>C 32% (city center), and 624.C>A 32% (city center). These mutations were identified among β-thalassemia patients by two regions of HBB gene 696 bp and 861 bp.

Conclusions The discovery of new genetic variants helps predict the severity of β-thalassemia disease. There are relatively few studies in molecular genetics of β-thalassemia in Iraq, and the new mutations reported here will provide valuable data for the prevention and control of β-thalassemia in Maysan Governate, Iraq. The results can lead to new genetic sequencing investigations for other Iraqi regions.

Keywords β-Thalassemia · DNA sequencing · Hematology diseases · PCR

Introduction

β-Thalassemia is an inherited mutation of the β-globin gene that causes a decrease in hemoglobin's β-globin chain. Most β-thalassemia mutations are found in people of Mediterranean, Middle Eastern, and Asian descent [1]. There are over 200 different thalassemia-causing mutations in the β-globin gene, resulting in the disease's broad genotypic and phenotypic variability [2]. The type of mutation in the β-globin gene determines the severity of the β-thalassemia clinical syndrome. More than 400 different mutations in the β-globin gene have been reported and identified as being responsible for β-thalassemia development [3, 4]. Many of the clinical features of this disorder appear to be caused by the cytotoxic accumulation of free β-globin chains [5]. However, humans have a very limited ability to excrete iron, and thus regular blood transfusions almost always result in iron overload [6]. This is the most serious problem associated with transfusion therapy [7].

The physical health consequences of β-thalassemia include physical malformation, developmental delays, and delayed puberty [8]. A genetic modifier is defined as the use of genetic variants to identify differences in disease phenotype causing variant. These can be classified as primary (which modulates the clinical severity of the disease) and secondary (which affects some disease complications). The

Accession numbers The novel nucleotide mutations sequencing were recorded at the National Science Center for Biotechnology and Information (NCBI) and the Japanese DNA Data Bank (DDBJ). The access numbers are: (LC727508, LC727509, LC727510, LC727511, LC727512, LC727513, LC727514, LC727515, LC727516, LC727517, LC727518, LC727519, LC727520, LC727520, LC727522).

✉ Ali A. A. Al-Shawi
ali.abdulhussein@uobasrah.edu.iq

Abbas Kadhim Jiheel AL-hameedawi
m.abbas.k.j@gmail.com

[1] Department of Chemistry, College of Education for Pure Sciences, University of Basrah, Basrah, Iraq

Published online: 22 January 2023

co-inheritance of other genetic variants mapping outside the globin clusters can also alter the clinical phenotype of homozygous β-thalassemia [9]. Clearly, clinicians, geneticists, clinical researchers, and basic scientists must work together to improve the understanding of the variability and treatment of patients with hemoglobinopathies [10].

According to the World Health Organization, approximately 68,000 people are born with β-thalassemia each year, and the incidence of symptomatic beta thalassemia is estimated to be 1 in 100,000 people in the general population. In Iraq, β-thalassemia is a major public health problem that affects Iraqis due to environmental pollution, particularly those who live near oil refineries that emit petroleum-based gases. There are 20,000 new cases of β-thalassemia per year in Iraq [11, 12]. Several studies have been published in the northern and middle regions of Iraq on the known types of mutations that cause thalassemia based on the demographic distribution of these mutations in different regions [13–17]. These studies, however, have not focused on the thalassemia-causing mutations in Maysan (southern of Iraq).

Moreover, the mutation is essential in the evolution and development of health. The mutation is important as the first step in evolution because it generates a new DNA sequence for a specific gene, creating a new allele. New genetic mutations spread through reproduction, and differential reproduction is the key feature of evolution [18]. A population that adapts most of its modifications are still harmful and will eventually be eliminated by natural selection. Identifying new mutations will assist in understanding the presence of mutations and inheritance patterns that lead to diseases such as thalassemia and other related genes [19, 20].

This study identified novel mutations of β-thalassemia disease in Maysan Governorate, Iraq, using polymer chain reactions and DNA sequencing techniques, and then generated genetic data for future research.

Materials and methods

This study evaluated patients with β-thalassemia major from December 2020 to the end of October 2021. We enrolled 40 β-thalassemia patients (20 males and 20 females) aged 1 to 35 years who were registered at Maysan Maternity and Children Hospital's Genetic Hematology Center. These subjects were all at the hematology center for blood transfusions.

Blood collection

Prior to transfusion, three milliliters of venous blood were collected in K3-EDTA tubes. Prior to enrollment in the study, patients and/or parents provided written consent. The Directorate of Maysan Health, College of Science, and University of Maysan all approved the work. An automated method was used to isolate genomic DNA from whole blood via a DNA purification instrument (gSYNC DNA Genomic extraction kits, Geneaid, Taiwan). Following genomic DNA extraction, agarose gel electrophoresis and

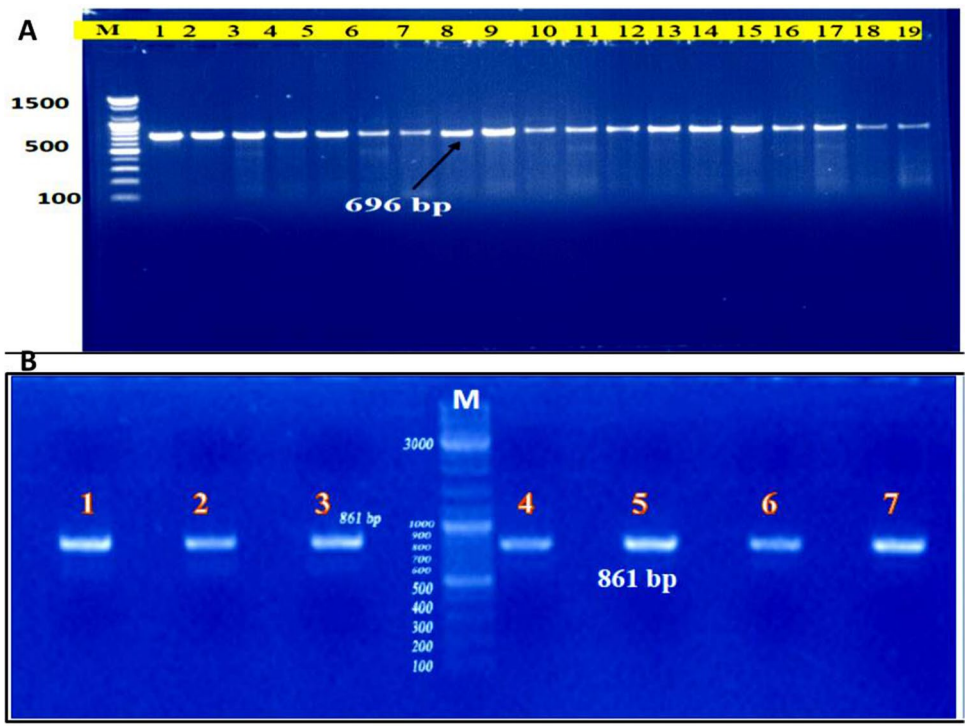

Fig. 1 Showed the gel electrophoresis results of the conventional-PCR products for: **A** The first region of size 696 bp using agarose gel at 0.5% concentration (80 V, for 1 h), M: Marker (Ladder 100 bp). **B** The second region of the β-thalassemia gene 861 bp using 1.5% agarose gel (70 V, 85 mA)

nanodrop instruments were used to confirm the completeness of the extraction.

DNA extraction and amplification

DNA electrophoresis revealed the presence of the isolated genomic material prior to amplification (DNA bundles for chromosomes). The extraction results showed that the amount of DNA ranged from 19 to 89.7 ng/l and the degree of purity (260/280) ranged from 1.7 to 2.05. This is within the normal range between (1.7–2.0) nm detected by Nanodrop devices, and the results of electrophoresis on agarose gel (2%). These results thus confirmed the success of the DNA amplification process for the regions under study. Using primers, the first selected region of the HBB gene yielded a band of 696 base pairs and the second region yielded a band of 861 base, Fig. 1.

Primer selection and design

Primers were chosen based on several studies as well as prior work in Iraq, and elsewhere [21–25]. The primers were designed using the NCBI/Primer-BLAST program and compared to previous research using the same primers to determine their sequence [26]. We included forward and reverse primers for detecting mutations in region I (696 bp) and region II (861 bp). Table 1, shows that primers (P1, P2) were used for the first region 696 bp and primers (P3, P4) for the second region 861 bp.

Conventional-PCR analysis

Conventional PCR technology was used for molecular screening to detect β-thalassemia mutations in two specific regions of the samples under study. We used a specific set of primers for β-globin gene regions I (696 bp) and II (861 bp).

Conventional-PCR programmed

The final PCR reaction volume was 20 µl, which included 5 µl of master mix (Bioneer, Korea), 5 µl of template DNA, and 1 µl of all designed primers. The first zone thermal cycle (696 bp) included 35 cycles: denaturation at 95 °C for 30 s, primer annealing at 58 °C for 45 s, extension at 72 °C for 1 min, and final extension at 72 °C for 7 min. The same program was used for the second region 861 bp, except the annealing temperature was set to 60 °C. The PCR products were then separated using electrophoresis on a 2% agarose gel.

DNA sequencing and data processing

Following confirmation of sample amplification, 20 µl of PCR product was sent to a macrogen humanizing genomics company in South Korea to obtain the target sequences of nitrogenous bases for the gene pieces. By Sanger sequencing using the "dideoxy nucleotide method," single strands of DNA forward and reverses (HBB gene) were sequenced to identify genetic mutations. The process of combining a deoxynucleotide with a large number of nucleotides occurs and either stops without reversal or terminates. The automated sequencer follows the same logic as the Sanger method (deoxynucleotide chain termination). The laser scans the bottom of the gel constantly, detecting bands as they move down the gel, and automated sequencing employs fluorescent tags on ddNTPs (a different dye for each nucleotide). This allows you to run all four reactions (dGTP, dATP, dCTP, and dTTP) in a single pass, allowing you to run a large number of reactions on a single gel. Deoxy-sequencing is a very efficient method that relies on copying one segment of DNA multiple times, but each copy is prematurely terminated. This results in a series of overlapping DNA copies, each one base longer. The DNA sequence is determined by reading the final base of each fragment on the gel after they have been separated on a polyacrylamide gel. After receiving the sequencing results from macrogen humanizing genomics company in South Korea, the researchers used bioinformatics software and NCBI tools to analyze them.

Statistical and bioinformatics analysis

To perform the multiple alignment between the genetic sequences, a statistical analysis of the obtained sequences was performed using the SPSS V.22 statistical analysis program. We also used bioinformatics software such as Bioedit V.7.2.6 [27]. DNA SP V.5.1 program was used to analyze genetic morphology and characterize haplotypes obtained from different samples [28]. The Haplotypes Network is

Table 1 Showed the used primers to detect β-thalassemia mutations in two regions 696 bp and 861 bp

No.	Primers	Primer sequence and orientation 5′–3′
1	P1 (forward)	CTTAGAGGTTCATTGAATCACGGCTGTCATCACTTAGAC
2	P2 (reverse)	TATGACATATTTCGGATCGCCTCCCCTTCCTATGACATGA
3	P3 (forward)	CAATGTATCATGCCTCTTTGCACC
4	P4 (reverse)	GAGTCAAGGCTGAGAGATGCAGGA

plotted using the Network V.10.2 program. We used the Median Joining (MJ) method to plot the haplotypes tuning [29].

Results and discussion

Screening with conventional PCR

Molecular characterization was performed using conventional PCR technology to detect potential mutations within two alleles of the thalassemia gene. Two zone testing is intended to detect a greater variety of mutations. The first regional study included 20 patients with major β-thalassemia and ten healthy controls. Region I (696 bp) consisted of exon (1), the IVSI intron region, exon (2), a portion of the non-coding region (URT5), and region II (861 bp). A total of 20 affected patients were included in the base pair size with thalassemia including the exon (3) and (IVSII intron) regions, the non-coding UTR3 region (2), shown in Fig. 2.

The relationship between phenotype and genotype in the context of identified mutations

The association between the phenotype and the mutations that cause β-thalassemia was investigated and compared to a control group. According to the findings, all identified mutations are β-thalassemia major. The homozygous type had the highest percentage of infection (78%). They were distributed among the genetic mutations in the following proportions: 16%, 17%, 17%, 13%, and 15% for mutations 205.C>T and 294.T>C, 389.G>A, 49.T>C, and 624.C>A, respectively. The percentage of β-thalassemia was heterozygous (22%), and distributed among the mutations: 4%, 3%, 3%, 7%, and 5% for the consecutive mutations 205.C>T and 294.T>C, 389.G>A, 49.T>C, and 624.C>A, respectively (Table 2). Most patients diagnosed with β-thalassemia have homozygous infection, while the minority have heterozygous infection, which may be ineffective due to the effect of the mutation itself, its genetic location, as well as the presence of another healthy gene that performs the function well in heterozygous infections. Even in minor monogenic disorders,

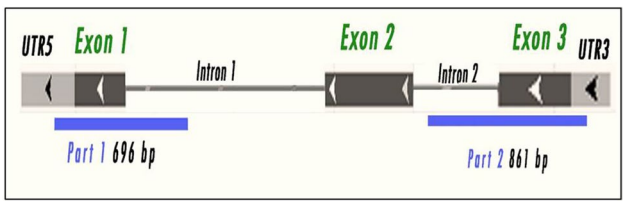

Fig. 2 Showed the studied regions of the HBB gene for thalassemia patients

Table 2 Showed the effect of mutations on genotypes patients with β-thalassemia major

β-Thalassemia mutation	Homozygous		Heterozygous	
	N	%	N	%
HBB: 205. C>T	16	16	4	4
294.T>C HBB:	17	17	3	3
HBB: 289.G>A	17	17	3	3
HBB: 49T>C	13	13	7	7
624.C>A HBB:	15	15	5	5
Total	78	78	22	22

patients with nearly identical genotypes may present with different clinical conditions. The clinical manifestations of β-thalassemia are extremely diverse, ranging from severe anemia to transfusion dependence. The mutation rate at the β-globin locus is the most reliable and predictive factor for disease phenotype. However, the interaction of environmental and genetic factors complicates the causal relationship between phenotype and genotype [30].

Demographic distribution of identified mutations

The demographic distribution of samples is affected by marine anemia (β-thalassemia major) in Maysan Governorate and was investigated here. The study included the city center as well as the districts surrounding it. Our findings revealed that the center ranks first in terms of the highest percentage of infection and the number of detected mutations because it had four mutations, as well as high diversity mutations in terms of genetic forms The center was second in terms of pace and three detected mutations. Al-Musharrah area is in third place, with two detected mutations, and Al-Mejar, Al-Maimiouna, and Qal'at Saleh districts are in fourth place, with one detected mutation in each district. Mutation 294TT had the highest frequency (80%) in the city center, followed by 205CC with a frequency of 65% in the Al-Musharrah sub-district. There were also mutations 624CC and 289GG that had a frequency of 50% in both the center district and the Al-Salam sub-district. Mutations 624AA, 624CA, 289GA, 49TC, and 289AA had a frequency of 25% in the city center, Al-Maimiuna, Al-Salam, and Qala'at Saleh, respectively. For the city center, mutations 205TT, 294TC, and 49CC have a frequency of 15%. These mutations were identified for the first time in Maysan Governorate, Iraq. The genotypes were classified as heterozygous or homozygous. These mutations spread in the city center and some districts and sub-districts but not others, which may be due to the number of mutations. The study's samples and/or the distance to some areas from medical centers and hospitals make it difficult to detect and diagnose thalassemia patients (Fig. 3). When comparing the frequency of mutation

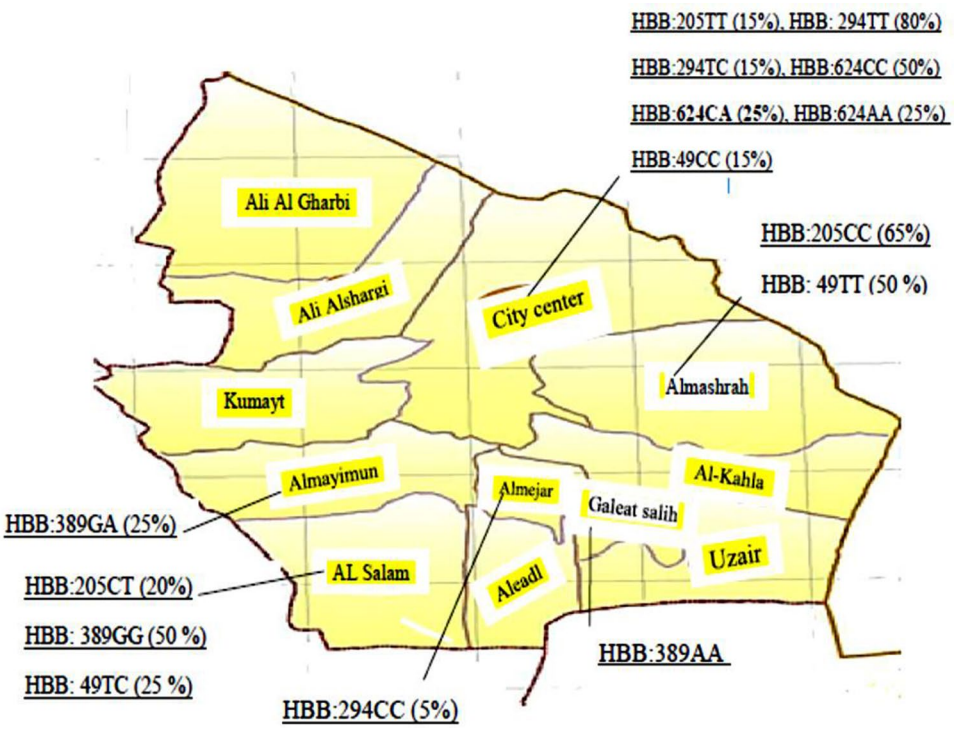

Fig. 3 Showed the map depicts geographical distribution of identified mutations in thalassemia patients in Maysan Governorate, Iraq

genotypes and types in the studied regions, we notice a clear difference between the mutations in terms of type and frequency of the mutation, which is consistent with previous observations. The distribution of alleles is affected by anemia (β-thalassemia), and shows a clear difference between the geographical regions studied [31].

Frequency of genetic polymorphisms

DNA sequencing is the most precise method for determining the exact nature of a mutation or variable position. Because DNA sequencing is regarded as the gold standard for mutation detection, mutations detected by scanning methods must be confirmed by DNA sequencing. Traditional Sanger DNA sequencing has been widely used [10].

Statistical analysis of HBB gene sequences for the selected gene segments in this study, showed 696 base pairs and 861 base pairs and five mutations distributed in different regions of the gene. Three were on the first plot at 205.C>T, 294.T>C, and 389.G>A. There were two, the second plot at 49.T>C and 624.C>A. The computational results showed that the percentage of the mutant allele was higher in the patients group than the control group in all mutations. The allele was mutant T for the first mutation 205.C>T for 25% of cases versus 5%. Allele C was recorded in 12% versus none for the second mutation 294.T>C. Allele C was recorded in 12% versus none for the third mutation 389.G>A. In the fourth mutation, only mutant allele C was seen 32% versus 15%. The fourth mutation 624.C>A only mutant allele A was seen 32% of the patients versus 15% of the controls.

Statistical analysis using the Chi-Square model revealed no significant differences in the distribution of all these mutations between the patient and control groups, except for the third mutation, 389.G>A. Here, the mutant allele A was must better with a P-value of 0.001. The distribution of the mutant allele was 38% in the patient group versus 0% in the control group, with a chi-squared value of 7.5 when the two groups were compared, Table 3 and Supplementary Table 1 demonstrate this. The C, C, A, C, and A alleles of the mutations 205.C>T, 294.T>C, 389.G>A, 49.T>C, and 624.C>A were recorded, with a clear superiority in recurrences in the group of thalassemia patients versus healthy controls for all mutations seen here. The results also revealed a significant (0.01) difference in the frequency of the C allele in the 49.T>C mutation in the patient group versus healthy controls. The statistical analysis revealed no significant differences between the mutation alleles 205.C>T, 294.T>C, and 624.C>A, versus healthy controls, There were only arithmetic differences in the allele frequencies. β-thalassemia allele distribution varies by region and between Maysan regions. This is consistent with previous findings showing that β-thalassemia mutants are extremely heterogeneous, with no distinct distribution patterns that would aid in determining any ethnic background [32], as shown in Table 2.

Table 3 Display the polymorphism frequencies and genotypes of the studied regions

Mutations		(Patients)		(Control)		Chi Square (X^2) *P* value
		Polymorphisms	Polymorphisms frequency %	Polymorphisms	Polymorphisms frequency %	
205.C>T	CC	13	65	9	90	X^2 2.468
	CT	4	20	1	10	P (0.29)
	TT	3	15	0	0	N.S
294.T>C	TT	16	80	10	100	X2 2.308
	TC	3	15	0	0	P (0.31)
	CC	1	5	0	0	N.S
389.G>A	GG	10	50	10	100	X^2 7.5
	GA	5	25	0	0	P (0.02)
	AA	5	25	0	0	*
49.T>C	TT	10	50	10	100	X^2 4.74
	TC	7	25	0	0	P (0.93)
	CC	3	15	0	0	N.S
624.C>A	CC	10	50	7	70	X^2 3.03
	CA	5	25	3	30	P (0.21)
	AA	5	25	0	0	N.S

(* = Significant; ** = High significant)

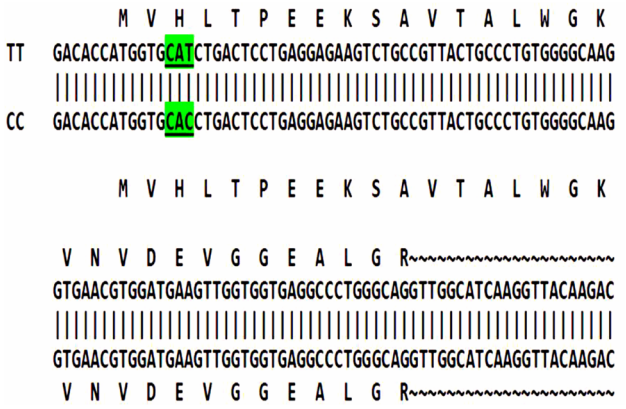

Fig. 4 Showed the impact of the mutation HBB.205 C>T on the amino acid

Identifying of mutations

The mutation HBB.205 C>T

The first mutation, 205.C>T, occurred in exon number one and is found on the entire gene 188. The first exon is the ninth nitrogen base, i.e., the third amino acid of the β-globin protein's peptide chain; this mutation is known as a silent mutation because the base change C histidine has not been changed to the T base in the histidine code because it CAT and CAC both code for the same histidine. This result is similar to the Chauhan results [33], as shown in Fig. 4. In patients, the percentage of repeat genotypes was 13 (65%) for CC, 4 (20%) for genotype CT, and 3 (15%) for the genotype TT, versus 9 (90%) for healthy controls. Genotype CC has one (10%) for the genotype CT, and 0% for the genotype TT, Supplementary Fig. 2. The study found no statistically significant differences between genotypes for patients and healthy subjects (P = 0.29), as shown in Table 3. There was no statistically significant relationship for the alleles C and T for patients and healthy subjects, as shown in Supplementary Table 1. The odds ratio (OR) for genotype TT was 4.20 and for genotype CT was equal to 4.20. For the genotype CC was 0.20, a person carrying the genotype TT is more likely to get the disease, while the structure CT is less likely to get the disease. The odds ratio of infection to allele T was 6.33, which is a high percentage compared to allele C, which amounted to 0.157 (Supplementary Table 2).

The mutation HBB: 294.T>C

The study findings revealed that the nucleotide changed from T to C. This change results in a silent mutation within site 294, which impacts the transcription of DNA (mRNA) and thus the effect on protein synthesis. There were three mutations 15% for genotype TC and one (5%) for genotype CC for patients compared to healthy subjects, the percentage of repeat genotypes was 10 (100%) for TT and 0 (0%) for genotype TC, Supplementary Fig. 3. The study found no significant statistical differences between the genotypes. Supplementary Table 2 shows no statistically significant relationship for the alleles C and T between patients and healthy subjects. The odds ratio (OR) for genotype TT was 0.17, for genotype TC was 4.20, and for

genotype CC was 1.61. The person carrying genotype TC is more likely to develop the disease, and genotype CC is less likely to develop the disease. Genotype TT is weaker than the other structures, and the percentage of infection for allele C is 6.35, which is high.

The mutation HBB: 389.G>A

Our findings also revealed a change in the nucleotide sequence, with the nitrogenous base switching from G to A. This variation is caused by the presence of a mutation within locus 389. The percentage of genotype recurrence was 12 (60%) for genotype GG, 3 (15%) for genotype GA, and 5 (25%) for genotype GA AA, Supplementary Fig. 4. When patients were compared to healthy controls, the percentage of recurrence of genotypes was 10 (100%) for genotype GG, zero (0%) for genotype GA, and 0% for genotype GA AA. There was a significant difference between patients and healthy people (P = 0.02), Table 3. Table 1, shows a very significant relationship and statistical significance for the alleles G and A between patients and healthy subjects. The results also revealed that the genotype GG has an odds ratio was equal to 0.07. Genotype GA was 4.20 and genotype AA was 7.45. The person carrying genotype AA is more likely to get the disease and genotype GA is less likely to get the disease. GG is less likely to cause disease than the other combinations. Furthermore, the odds of disease for the allele G was 0.049, which is a higher percentage than the allele A. This was 0.415, despite the fact that the odds ratio for the allele A is slightly higher than the Allele ratio G, but both alleles have a low probability of developing the disease, as shown in Supplementary Table 2.

The mutation HBB: 49.T>C

This mutation was caused by a change in the nucleotide sequence, where the nitrogenous base changed from T to C. The percentage of repeat genotypes within locus 49 was 10 (50%) for TT, 7 (15%) for genotype TC, and 3 (35%) for genotype CC for patients compared to healthy subjects, the percentage of repeat genotypes was 9 (90%) for TT, 1 (10%) for genotype TC, and 0% for genotype CC CC, Supplementary Fig. 5. Table 3, shows a significant relationship and statistical significance for the alleles C and T between patients and healthy subjects. TC was 4.84 and for genotype CC was 4.20, person carrying the genotype TC is more likely and prone to contracting the disease, and genotype CC is less likely to contract the disease. The probability of contracting the disease with genotype TT is weaker from the other genotypes. The odds of disease for allele C was 20.12, which is a high percentage compared to the all T (0.049), Supplementary Table 2.

The mutation HBB: 624.C>A

This mutation was caused by a change in the nucleotide sequence, where the nitrogenous base changed from C to A. Within site 624, the percentage of repeat genotypes for the patients was 10(50%) for genotype CC, 5(25%) for genotype CA, and 5(25%) for genotype AA, Supplementary Fig. 6. Table 3 shows that the percentage of recurrence of genotypes was 7(70%) for genotype CC, 3 (30%) for genotype CA, and 0(0%) for genotype AA (P = 0.21). There was no significant relationship or statistical significance for the alleles C and T between patients and healthy subjects (P = 0.065). Here, CA was 0.77, and CC was 0.42, thus indicating that the person carrying genotype AA is more likely and prone to contracting the disease than genotypes CC and TT. Here, the probability of disease is low for both. The probability ratio for allele A is 3.40 and the probability ratio for allele C is 0.294. The OR value suggets that allele A is more likely to be infected or carry the disease, as shown in Supplementary Table 2.

Several Iraqi studies on thalassemia disease in different regions have been reported (1999–2022). Deutsch et al. found an undescribed β-globin variant using members of a family from the Kurdistan region of Halabja. Hb Beta10 Iraq-Halabja: (A7) Ala → Val (GCC → GTC), associated with β-thalassemia IVS-2 nt1 G → A, and either alpha-thal-2–3.7 kb deletion (brother) or anti-3.7 kb type alpha-globin gene triplication (propositus) [34]. Al-Allawi et al. identified eight common mutations: IVS-II-1 (G → A), codon 44 (-C), codon 5 (-CT), IVS-I-1 (G → A), codon 39 (C → T), IVS-I-6 (T → C), codons 8/9 (+G), and IVS-I-5 (G → C) in the Dohuk province (northern Iraq) [11]. Al-Allawi et al. found seven most common in 88.2% of the thalassemia chromosomes from two provinces in northern Iraq (Erbil and Dohuk): IVS-II-1 (G→A), IVS-I-1 (G→A), codon 8 (-AA), codon 39 (G→T), codon 8/9 (+G), codon 44 (-C), and codon 5 (-CT) [35]. Al-Allawi et al. identified the four most common mutations: IVS-I-6 (T>C) [33.3%], IVS-II-I (G>A) [21.1%], codon 82/83 (−G) [10.1%], and codon 8 (-AA) [8.1%] in the northern Iraq [36]. Shamoon et al. identified the four major molecular patterns associated with the β-thalassemia phenotype: (+)/(+) (38.5%), (+)/(0) (21.6%), (0)/(0) (31.3%), and (0)/wild type (8.4%), in the northern Iraq [37]. Eissa et al. identified eleven distinct β-thalassemia mutations in Ninevah province (north Iraq). The most common mutation was IVS-I-110 (G>A), which was found in 34% of cases, followed by IVS-I-6 (T>C) in 9.6%, IVS-I-5 (G>C) in 8.5%, codon 39 (C>T) and codon 44 (-C) in 7.4% each, and IVS-I-1 (G>A) in 6.4%. Other mutations were less common, including codon 8 (-AA),

IVS-I-130 (G>C), codon 5 (-CT), and IVS-II-745 (C>G) [38]. Jiang et al. compared twenty-two FSC8 homozygote patients to Iraqi twin men who were homozygous for HBB codon 8. They found that the twins inherited homozygosity for HMIP 3-bp deletion at rs66650371 and heterozygosity for Hph β-thalassaemia mutation, but not the other twenty-two patients [39]. Amin et al. used patients in Sulaymaniyah province (northern Iraq). They identified nineteen different β-globin gene mutations distributed across 37 different genotypes. IVS-II-I (G>A) (47.2%) was the most common, followed by IVS-I-6 (T>C) (23.3%) and IVS-I-110 (G>A) (5%) [40]. Amin et al. found a total of 22 β-globin mutations in 53 different genotypes in northern of Iraq (Sulaymaniyah province). The most common was IVS-II-1 (G>A) (35.7%), IVS-I-6 (T>C) (18.0%), and codon 8/9 (+G) (8.5%) [41]. AlMosawi et al. used molecular analysis to detect nine types of mutations in the β-thalassemia patients in the middle region of Iraq (Al-Muthanna province): IVS II-1 G>A, IVS 1–5 G>C, IVS-II-666 C>T, CD2 CAT>CAC, IVS-II-850 G>A, IVS-II-16 G<C, Hb King's Mill, Hb Saveh, and IVS-II-81 C>T [42]. Atroshi et al. used patients from the Iraqi northern region (Dohuk province) to identify twenty different thalassemia mutations, the most frequent being IVS-II-1 (G>A) (HBB: c.315+1G>A), IVS-I-6 (T>C) (HBB: c.92+6T>C), codon 5 (-CT) (HBB: c.17 18delCT), IVS-I-110 (G>A) (HBB: c.93-21G>A), codon 44 (-C) (HBB: c.135delC), codon 8 (-AA) (HBB: c.25 26delAA) and IVS-I-1 (G>A) (HBB: c.92+1G>A) [43]. Adnan et al. identified seven different types of β-thalassemia mutations: codon 15 (G-A), IVS1nt -5 (GC), codon 8/9 (+G), codon 30 (G-C), -88 (C-T), codon 41/42 (-TCTT), and codon 8/9 (+G) (AA). The most common β-thalassemia mutations are codon-15(G-A) and IVS 1nt5(G-C) in 31 (37.3%) and 18(21.7%), in the south of Iraq (Basrah province) [16]. In the present study, we selected patients diagnosed with β-thalassemia from different regions in Maysan province (southern Iraq). PCR and DNA sequencing technologies aided in the identification of five new mutations: 294.T>C 12%, 205. C>T 25%, 289. G>A 38%, 49.T>C 32%, 624. C>A 32% via two regions of the HBB gene: 696 bp and 861 bp. The new mutations were recorded at the National Science Center for Biotechnology and Information (NCBI) and the Japanese DNA Data Bank (DDBJ). These new mutations will aid researchers in better understanding the importance of premarital screening, prenatal diagnostic services, and developing appropriate prevention programs to control thalassemia disease in Iraq, particularly in the south.

Haplotype network

According to the findings of the first region study 696 bp, the total number of HBB haplotypes is seven, with two haplotypes belonging to the Iraqi patients under study. One haplotype belongs to healthy Iraqis, four haplotypes belong to American patients with thalassemia, two haplotypes belong to British patients, one haplotype is from Russian patients, and one haplotype is from French patients (Fig. 5). The findings revealed that the haplotype H1 was common among American patients, Iraqi patients, and healthy Iraqi patients under study. Haplotype H7 was common in Russian and French patients, and the haplotype H4 was common in Iraqi and American patients. The study results for the second region 861 bp showed ten total haplotypes of the β-globulin gene HBB, four haplotypes belong to the Iraqi patients under study, one haplotype was for healthy Iraqis, and five haplotypes were seen in Iraqis with thalassemia among other studies, and two haplotypes were seen for French patients (Supplementary Fig. 1). Haplotype H7 was shared by the patients under study and the Iraqi health control system. Haplotype H1 was shared by the French patients and the Iraqi patients.

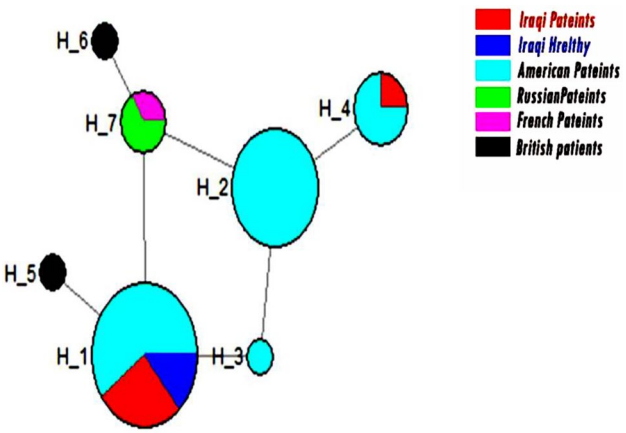

Fig. 5 Showed the distribution of haplotypes in the HBB gene's first region 696 bp genomic sequence and some similar sequences from the NCBI GenBank database

Conclusion

The current study demonstrates the diversity of molecular defects causing β-thalassemia in Maysan Governate, Iraq. Most β-thalassemia mutations result from point mutations, small deletions or insertions within coding regions, and exon–intron junctions. The heterogeneity of the mutations makes identifying the mutation in some β-thalassemia patients difficult. More than 400 β-thalassemia mutations have been reported in the genes database to date. Therefore, discovering new β-thalassemia mutations will increase the possibility of controlling and regulating this disease. We used direct PCR and DNA sequencing to characterize all of the β-thalassemia alleles, and 15 novel mutations accounted for the β-thalassemia genes. The frequency represents β-globin gene mutations among Iraqi patients, who

primarily lived in various locations throughout Maysan Governate. Further research can help facilities studies in Iraqi Governates near Maysan Governate. The results can also identify new mutations to connect the relationships, causes, and risk factors of β-thalassemia development.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11033-023-08271-1.

Acknowledgements This research project is funded by the Ministry of Higher Education and Scientific Researches, University of Basrah, College of Education for Pure Sciences, Department of Chemistry, Iraq, and is part of the Ph.D program's graduation requirements. The authors would like to express their gratitude to the Maysan Health Department and the Genetic Blood Disease Center for their assistance and cooperation. Dr. Asaad Yahya, University of Basrah. Department of Animal Production is grateful for his assistance with statistical analyses. Dr. Azhar. A. Jaffar Marsh Research Center, Thi-Qar University deserves a heartfelt thank you. Dr. Maitham-Abdul-Kadhim Al-Shamri, Head of Department of Life Sciences, College of Science, Maysan University is thanked and grateful for his support of the research project.

Declarations

Competing interests The authors declare that they have no competing interests.

Ethical approval All participants provided informed consent, while the parent or legal guardian of participants under the legal age of majority in Iraq provided with written informed consent. The present study was commenced after obtaining approval from the ethical committee at the Maysan Health Department and the Genetic Blood Disease Center, and approved by the University of Basrah, College of Education for Pure Sciences, Department of Chemistry. All methods were performed in accordance with the Helsinki Declaration.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Jaing TH, Chang TY, Chen SH, Lin CW, Wen YC, Chiu CC (2021) Molecular genetics of β-thalassemia: a narrative review. Medicine (Baltimore) 100(45):e27522. https://doi.org/10.1097/MD.0000000000027522
2. Macharia AW, Mochamah G, Uyoga S, Ndila CM, Nyutu G, Tendwa M, Nyatichi E, Makale J, Ware RE, Williams TN (2020) β-Thalassemia pathogenic variants in a cohort of children from the East African coast. Mol Genet Genomic Med 8(7):e1294. https://doi.org/10.1002/mgg3.1294
3. Xu L, Mao A, Liu H, Gui B, Choy KW, Huang H, Yu Q, Zhang X, Chen M, Lin N, Chen L, Han J, Wang Y, Zhang M, Li X, He D, Lin Y, Zhang J, Cram DS, Cao H (2020) Long-Molecule sequencing: a new approach for identification of clinically significant DNA variants in α-thalassemia and β-thalassemia carriers. J Mol Diagn 22(8):1087–1095. https://doi.org/10.1016/j.jmoldx.2020.05.004
4. Needs T, Gonzalez-Mosquera LF, Lynch DT (2022) Beta Thalassemia. StatPearls Publishing, Treasure Island
5. Belmokhtar I, Lhousni S, Elidrissi Errahhali M, Ghanam A, Elidrissi Errahhali M, Sidqi Z, Ouarzane M, Charif M, Bellaoui M, Boulouiz R, Benajiba N (2022) Molecular heterogeneity of β-thalassemia variants in the Eastern region of Morocco. Mol Genet Genomic Med 10(8):e1970. https://doi.org/10.1002/mgg3.1970
6. Arab-Zozani M, Kheyrandish S, Rastgar A, Miri-Moghaddam E (2021) A systematic review and meta-analysis of stature growth complications in β-thalassemia major patients. Ann Glob Health 87(1):48. https://doi.org/10.5334/aogh.3184
7. Abdul-Zahra HA, Hassan MK, Ahmed BA (2016) Health-related quality of life in children and adolescents with β-thalassemia major on different iron chelators in Basra, Iraq. J Pediatr Hematol Oncol 38(7):503–511. https://doi.org/10.1097/MPH.0000000000000663
8. Thein SL (2018) Molecular basis of β thalassemia and potential therapeutic targets. Blood Cells Mol Dis 70:54–65. https://doi.org/10.1016/j.bcmd.2017.06.001
9. Taher AT, Cappellini MD (2018) How I manage medical complications of β-thalassemia in adults. Blood 132(17):1781–1791. https://doi.org/10.1182/blood-2018-06-818187
10. Chen J, Lin S, Gan J, Xin X, Huang J (2020) A novel β-thalassemia variant at HBB:c.14delC (Codon 4, -C) identified via next-generation sequencing. Hematology 25(1):400–404. https://doi.org/10.1080/16078454.2020.1841920
11. Al-Allawi NA, Jubrael JM, Hughson M (2006) Molecular characterization of beta-thalassemia in the Dohuk region of Iraq. Hemoglobin 30(4):479–486. https://doi.org/10.1080/03630260600868097
12. Jalal SD, Al-Allawi NA, Bayat N, Imanian H, Najmabadi H, Faraj A (2010) β-Thalassemia mutations in the Kurdish population of northeastern Iraq. Hemoglobin 34(5):469–476. https://doi.org/10.3109/01676830.2010.513591
13. Omer WA (2010) Molecular Characterization of beta-thalassemia mutations in Baghdad. Iraqi J Comm Med 2:90–95
14. Katta HK, Salman ED, Saud AM (2013) Molecular spectrum of beta- globin mutations in transfusion- dependent patients with Thalassemia in Wasit province p Iraq. Curr Res Microbiol and Biotech 1(16):174–277
15. Al-Fartosi KG, Aziz HA (2014) Molecular detection of some mutation which causes β-thalassemia in Al-Muthanna province – Iraq. Int J Adv Res 2(10):321–329
16. Adnan IA, Meaad KH, Assad FW (2016) β-Thalassemia mutations among thalassemia major patients in Basrah Province – Iraq. Int J Curr Microbiol Appl Sci 5(5):448–457. https://doi.org/10.20546/ijcmas.2016.505.047
17. Abdulwahid DA, Hassan MK (2013) β- and α-Thalassemia intermedia in Basra. Southern Iraq Hemoglobin 37(6):553–563. https://doi.org/10.3109/03630269.2013.825841
18. Schell R, Hale JJ, Mullis MN, Matsui T, Foree R, Ehrenreich IM (2022) Genetic basis of a spontaneous mutation's expressivity. Genetics 220(3):iyac013. https://doi.org/10.1093/genetics/iyac013
19. Jiang F, Lyu GZ, Zhang VW, Li DZ (2021) Identification of thalassemia gene cluster deletion by long-read whole-genome sequencing (LR-WGS). Int J Lab Hematol 43(4):859–865. https://doi.org/10.1111/ijlh.13452
20. Arpaci A, Gul BU, Ozcan O, Ilhan G, El C, Dirican E, Elmacioglu S, Kaya H (2021) Presentation of two new mutations in the 3'untranslated region of the β-globin gene and evaluating the molecular spectrum of thalassemia mutations in the Mediterranean region of Turkey. Ann Hematol 100(6):1429–1438. https://doi.org/10.1007/s00277-021-04509-9
21. Al-Allawi NA, Puehringer H, Raheem RA, Oberkanins C (2015) Genetic modifiers in β-thalassemia intermedia: a study on 102 Iraqi Arab patients. Genet Test Mol Biomarkers 19(5):242–247. https://doi.org/10.1089/gtmb.2014.0310

22. Munkongdee T, Chen P, Winichagoon P, Fucharoen S, Paiboonsukwong K (2020) Update in laboratory diagnosis of thalassemia. Front Mol Biosci 7:74. https://doi.org/10.3389/fmolb.2020.00074
23. Al-Abboodi MJ (2011) Molecular study for some beta-globin mutations in sample of Iraqi thalassemic patients. Dissertation, University of Baghdad, Iraq
24. Ali S, Mumtaz S, Shakir HA, Khan M, Tahir HM, Mumtaz S, Mughal TA, Hassan A, Kazmi SAR, Sadia, Irfan M, Khan MA (2021) Current status of beta-thalassemia and its treatment strategies. Mol Genet Genomic Med 9(12):e1788. https://doi.org/10.1002/mgg3.1788
25. Kalkan BM, Kala EY, Yuce M, Karadag Alpaslan M, Kocabas F (2020) Development of gene editing strategies for human β-globin (HBB) gene mutations. Gene 734:144398. https://doi.org/10.1016/j.gene.2020.144398
26. Topfer SK, Feng R, Huang P, Ly LC, Martyn GE, Blobel GA, Weiss MJ, Quinlan KGR, Crossley M (2022) Disrupting the adult globin promoter alleviates promoter competition and reactivates fetal globin gene expression. Blood 139(14):2107–2118. https://doi.org/10.1182/blood.2021014205
27. Hall TA (1999) BioEdit:a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp 41:95–98
28. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11):1451–1452. https://doi.org/10.1093/bioinformatics/btp187
29. Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16(1):37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036
30. Jaing TH, Chang TY, Chen SH, Lin CW, Wen YC, Chiu CC (2021) Molecular genetics of β-thalassemia: a narrative review. Medicine 100(45):27522. https://doi.org/10.1097/MD.0000000000027522
31. Saha J, Panja A, Nayek K (2021) The prevalence of *HBB* mutations among the transfusion-dependent and non transfusion-dependent Hb E/β-thalassemia children in a Tertiary Center of West Bengal, India. Hemoglobin 45(3):157–162. https://doi.org/10.1080/03630269.2021.1933023
32. Lin W, Zhang Q, Shen Z, Qu X, Wang Q, Wei L, Qiu Y, Yang J, Xu X, Lao J (2021) Molecular and phenotype characterization of an elongated β-globin variant produced by HBB:C.313delA. Int J Lab Hematol 43(6):1620–1627. https://doi.org/10.1111/ijlh.13639
33. Chauhan W, Afzal M, Zaka-Ur-Rab Z, Noorani MS (2021) A Novel Frameshift mutation, deletion of HBB: c.199_202delAAAG [Codon 66/67 (-AAAG)] in β-thalassemia major patients from the Western Region of Uttar Pradesh, India. Appl Clin Genet 14:77–85. https://doi.org/10.2147/TACG.S294891
34. Deutsch S, Darbellay R, Offord R, Frutiger A, Kister J, Wajcman H, Beris P (1999) Hb Iraq-Halabja beta10 (A7) Ala–>Val (GCC–>GTC): a new beta-chain silent variant in a family with multiple Hb disorders. Am J Hematol 61(3):187–193. https://doi.org/10.1002/(sici)1096-8652(199907)61:3%3c187::aid-ajh5%3e3.0.co;2-7
35. Al-Allawi NA, Hassan KM, Sheikha AK, Nerweiy FF, Dawood RS, Jubrael J (2010) β-Thalassemia mutations among transfusion-dependent thalassemia major patients in Northern Iraq. Mol Biol Int 2010:479282. https://doi.org/10.4061/2010/479282
36. Al-Allawi NA, Jalal SD, Mohammad AM, Omer SQ, Markous RS (2014) β -thalassemia intermedia in Northern Iraq: a single center experience. Biomed Res Int 2014:262853. https://doi.org/10.1155/2014/262853
37. Shamoon RP, Al-Allawi NA, Cappellini MD, Di Pierro E, Brancaleoni V, Granata F (2015) Molecular basis of β-thalassemia intermedia in Erbil Province of Iraqi Kurdistan. Hemoglobin 39(3):178–183. https://doi.org/10.3109/03630269.2015.1032415
38. Eissa AA, Kashmoola MA, Atroshi SD, Al-Allawi NA (2015) Molecular characterization of β-thalassemia in Nineveh Province illustrates the relative heterogeneity of mutation distributions in Northern Iraq. Indian J Hematol Blood Transfus 31(2):213–217. https://doi.org/10.1007/s12288-014-0369-1
39. Jiang Z, Luo HY, Huang S, Farrell JJ, Davis L, Théberge R, Benson KA, Riolueang S, Viprakasit V, Al-Allawi NA, Ünal S, Gümrük F, Akar N, Başak AN, Osorio L, Badens C, Pissard S, Joly P, Campbell AD, Gallagher PG, Steinberg MH, Forget BG, Chui DH (2016) The genetic basis of asymptomatic codon 8 frameshift (HBB:c25_26delAA) β(0)-thalassaemia homozygotes. Br J Haematol 172(6):958–965. https://doi.org/10.1111/bjh.13909
40. Amin SS, Jalal SD, Ali KM, Mohammed AI, Rasool LK, Osman TJ (2020) Beta-thalassemia intermedia: a single thalassemia center experience from Northeastern Iraq. Biomed Res Int 2020:2807120. https://doi.org/10.1155/2020/2807120
41. Amin S, Jalal S, Ali K, Rasool L, Osman T, Ali O, M-Saeed A (2020) Molecular characterization and disease-related morbidities of β-thalassemia patients from the Northeastern part of Iraq. Int J Gen Med 13:1453–1467. https://doi.org/10.2147/IJGM.S277947
42. AlMosawi RHN, Al-Rashedi NAM, Ayoub NI (2020) Clinical laboratory manifestation and molecular diagnosis of β-thalassemia patients in Iraq. J Pediatr Hematol Oncol 42(1):27–31. https://doi.org/10.1097/MPH.0000000000001656
43. Atroshi SD, Al-Allawi NAS, Eissa AA (2021) Updated molecular spectrum of β-thalassemia mutations in Duhok Province, Northern Iraq: ethnic variation and the impact of immigration. Hemoglobin 45(4):239–244. https://doi.org/10.1080/03630269.2021.1984250