



OPEN ACCESS

EDITED BY

Yasir Hameed,
Islamia University of Bahawalpur, Pakistan

REVIEWED BY

Michael Diamantidis,
General Hospital of Larissa, Greece
Muhammad Rizwan,
COMSATS University Islamabad, Pakistan

*CORRESPONDENCE

Liyun Song
✉ chinesu@126.com

[†]These authors have contributed equally to this work

RECEIVED 29 July 2025

ACCEPTED 07 October 2025

PUBLISHED 30 October 2025

CITATION

Shen S, Huang J, Qi H, Liu Z,
Zhang W, Yuan X, Zhang Z, Chen H,
Xie X, Xiao L, Chen J and Song L (2025)
Molecular identification and phenotypic study
of a novel *HBB*: c.-23A>G mutation in the 5'
untranslated region.
Front. Med. 12:1675600.
doi: 10.3389/fmed.2025.1675600

COPYRIGHT

© 2025 Shen, Huang, Qi, Liu, Zhang, Yuan,
Zhang, Chen, Xie, Xiao, Chen and Song. This
is an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Molecular identification and phenotypic study of a novel *HBB*: c.-23A>G mutation in the 5' untranslated region

Shichun Shen^{1†}, Jungao Huang^{1†}, Haimei Qi², Zezhang Liu^{3,4},
Wenqian Zhang^{3,4}, Xianping Yuan⁵, Zhuling Zhang⁵,
Haijun Chen¹, Xinxing Xie¹, Lin Xiao¹, Junkun Chen¹ and
Liyun Song^{6*}

¹Department of Medical Genetics, Ganzhou Maternal and Child Health Hospital, Ganzhou, China,

²Department of Clinical Laboratory, Ganzhou Maternal and Child Health Hospital, Ganzhou, China,

³BGI Genomics, Shenzhen, China, ⁴Clin Lab, BGI Genomics, Wuhan, China, ⁵Obstetrical Department,
Ganzhou Maternal and Child Health Hospital, Ganzhou, China, ⁶Department of Clinical Laboratory,
Ganzhou People's Hospital, Ganzhou, China

Background: β -thalassemia is a prevalent genetic disorder in the Gannan region, Southern China. Mutations in the 5' untranslated region of the β -globin gene are associated with diverse clinical phenotypes, posing challenges for effective prevention strategies in this region.

Methods: In this study, carriers of the *HBB*: c.-23A>G mutation were identified from a cohort of 192,720 individuals who underwent thalassemia gene testing in the Gannan region. Hematological data from these carriers were collected, and pedigree information was gathered for further analysis.

Results: Among the 192,720 individuals tested, 75 carriers of the *HBB*: c.-23A>G mutation were identified, yielding a carrier frequency of 3.89 per 10,000. Statistical analysis showed no significant differences in hematological parameters between *HBB*: c.-23A>G heterozygotes and normal individuals. Furthermore, the minimum free energy of mRNA with the *HBB*: c.-23A>G mutation showed no significant difference compared to that of the wild-type mRNA.

Conclusion: The carrier frequency of *HBB*: c.-23A>G in the Gannan region is non-negligible. Hematological data analyses suggested that this mutation may be a likely benign variant. Overall, this study elucidates the molecular and phenotypic characteristics of the *HBB*: c.-23A>G mutation, providing crucial evidence for genetic counseling in clinical practice.

KEYWORDS

β -globin gene, 5' untranslated region, novel mutation, hematological parameters, minimum free energy

1 Introduction

Thalassemia is a group of autosomal recessive genetic disorders caused by reduced or absent synthesis of the globin chains that constitute the hemoglobin (Hb) tetramer (1). Based on the type of globin chain affected, thalassemia is primarily classified into two categories: α -thalassemia and β -thalassemia. β -thalassemia is prevalent in southern China, with a carrier frequency ranging from 2.27 to 6.43% (2–4). Gannan, a region in southern China with a population of approximately 8.99 million, has previously reported a β -thalassemia carrier rate

of 4.06% (5). Patients with severe β -thalassemia (β -thalassemia major, β -TM) require regular blood transfusions or hematopoietic stem cell transplantation to survive, posing significant medical and public health challenges worldwide, particularly in Southeast Asia and southern China (6).

The transcriptional initiation of human β -globin mRNA occurs at the canonical Cap site (+1) located within the gene's promoter region (7). Between this Cap site and the translation initiation codon (ATG) lies a 50-nucleotide 5' untranslated region (UTR) that serves critical regulatory functions in mRNA stability, ribosomal scanning efficiency, and translational control (8, 9). For example, well-characterized mutations in the Kozak sequence within this region have been established as a significant molecular etiology of β -thalassemia (10). Although this region was historically considered to have low mutational susceptibility, the widespread implementation of next-generation sequencing (NGS) in thalassemia carrier screening has revealed an expanding spectrum of pathogenic variants within this regulatory domain, such as *HBB*: c.-40C>G, *HBB*: c.-10A>T, *HBB*: c.-8C>G, and *HBB*: c.-29G>T (11, 12). These newly identified mutations exhibit remarkable phenotypic variability, ranging from silent carrier states to severe transfusion-dependent anemia, thereby complicating genetic counseling and posing challenges to existing public health strategies for thalassemia prevention in endemic regions (13, 14).

In this study, we investigated the clinical phenotype of the *HBB*: c.-23A>G mutation (CAP+28 [A>G]) located in the 5' UTR of the β -globin gene. Notably, this mutation has not been reported in the specialized globin gene variant databases HbVar¹ and IthaGenes². Through NGS, we identified 75 carriers of this mutation among 192,720 individuals. We studied the hematological data of these carriers to determine the clinical phenotype of this variant. Our study provides a theoretical basis for the screening, prevention, and treatment of thalassemia in this region and other areas.

2 Materials and methods

2.1 Participant recruitment

This study included a total of 192,720 participants, covering thalassemia gene testing data from two groups. The first group consisted of 136,312 individuals of childbearing age who were either registered residents of Gannan or lived in Gannan on a regular basis, and the data were collected from April 2019 to April 2021. The second group included 56,408 pregnant women who were either registered residents of Gannan or lived in Gannan on a regular basis, with data collected from April 2023 to December 2024. In addition, the study integrated hematological data from 75 individuals carrying the *HBB*: c.-23A>G mutation, as well as relevant information from 22 pedigrees. All participants signed written informed consent forms before the study began. The study was approved by the Ethics Committee of Ganzhou Maternal and Child Health Hospital (Approval Number:

2024108) and was conducted in accordance with the ethical guidelines for research involving human subjects.

2.2 Sample collection and genomic DNA extraction

First, 2 mL of peripheral blood samples were collected using ethylene diamine tetraacetic acid K2 (EDTA-K2) anticoagulated tubes. Genomic DNA was extracted from 200 μ L whole blood samples using the QIAamp DNA blood Mini kit (Qiagen, Hilden, Germany). DNA extracts were then arrayed in 96-well plates, and concentration was quantified using a qubit 3.0 fluorometer (Thermo fisher scientific, Waltham, MA, USA). All samples were required to have a DNA concentration >10 ng/ μ L and an A260/A280 ratio between 1.8 and 2.0 for subsequent analysis.

2.3 Next-generation sequencing

The polymerase chain reaction (PCR) technique was employed to amplify the *HBA1*, *HBA2*, and *HBB* genes. The *HBA1* and *HBA2* genes were amplified from 35 bp upstream of the cap site to 150 bp downstream of the termination codon, generating amplicons of approximately 900 bp. Amplification of the *HBB* gene was divided into two segments. The first segment extended from 135 bp upstream of the CAP site at the 5' end to 150 bp into the second intron, generating a DNA fragment of approximately 760 bp. The second segment spanned from 630 bp into the second intron to 150 bp downstream of the termination codon at the 3' end, producing a DNA fragment of approximately 560 bp. After amplification, the DNA fragments were sheared into approximately 200 bp segments by ultrasonication. Following end-repair, sequencing adapters were ligated to construct the sequencing library. Sequencing was carried out on the MGISEQ-200 chip (MGI, Shenzhen, China) in paired-end (PE100) mode. Finally, sequencing data were analyzed using the Halos system to identify mutation sites in the relevant genes.

2.4 Site verification and pedigree verification

For positive samples carrying the *HBB*: c.-23A>G mutation, the variant was verified by Sanger sequencing (ABI 310, Applied Biosystems, USA). To exclude other potential mutations in the *HBB* gene, DNA samples were also subjected to third-generation sequencing (Guangdong HybriBio Biotechnology Co., Ltd., Guangzhou, China). Additionally, 22 family members were tested to determine the cis/trans configuration of two compound heterozygotes and to confirm co-segregation of the *HBB*: c.-23A>G mutation in the pedigrees.

2.5 Hematological data and analysis

Hematological parameters were measured using the Mindray 7,500 blood cell analyzer (Shenzhen, China). Hemoglobin electrophoresis analysis was performed with the Capillarys 3 capillary electrophoresis

¹ <https://globin.bx.psu.edu/hbvar/>

² <https://www.ithanet.eu/db/ithagenes>

instrument (Sebia, France). Hematological data were collected from 80 healthy adults (40 males and 40 females), 60 β^0 heterozygote carriers (30 males and 30 females), 60 β^+ heterozygote carriers (15 males and 15 females), 30 β^N/β^N with $--^{SEA}/\alpha\alpha$ carriers (15 males and 15 females), 20 β^0/β^N with $--^{SEA}/\alpha\alpha$ carriers, 10 β^+/ β^N with $--^{SEA}/\alpha\alpha$ carriers, and 20 $HBB: c.-11_-8delAACA$ and 10 $HBB: c.-29G>A$ (Common mutations of 5' UTR in Gannan region) heterozygotes (10 males and 10 females for the former, 5 males and 5 females for the latter) as the control group (Supplementary Table S1). These data were used to analyze the hematological characteristics of $HBB: c.-23A>G$ carriers and explore the potential impact of this mutation on β -chain expression.

2.6 Prediction of mRNA secondary structure

Using the RNAfold web server (15), we predicted the secondary structures of HBB gene mRNA for the wild type, $HBB: c.-23A>G$ mutant, and $HBB: c.-29G>A$ mutant. A comparative analysis of the minimum free energy and differences in internal stem-loop structures among the three types of mRNA secondary structures was carried out.

2.7 Data analysis and statistics

Statistical analysis was conducted using SPSS 27.0 software. Measured data were expressed as mean $\pm$ standard deviation (SD), and count data were expressed as percentage (%). Differences between groups were examined using the independent samples *t*-test. A *p* value of < 0.05 was considered statistically significant.

3 Results

3.1 Gene testing results of β -thalassemia

After analyzing the screening results of 192,720 individuals, a total of 7,968 β -thalassemia carriers were identified, resulting in a carrier rate of 4.134% in this population. Among these β -thalassemia carriers, 75 individuals were identified as carrying the $HBB: c.-23A>G$ mutation, with a carrier frequency of 3.89 per 10,000 in this region. The genotypic distribution was as follows: 65 were $\beta^{HBB:c.-23A>G}/\beta^N$ with $\alpha\alpha/\alpha\alpha$, 7 were $\beta^{HBB:c.-23A>G}/\beta^N$ with $--^{SEA}/\alpha\alpha$, 1 was $\beta^{HBB:c.-23A>G}/\beta^N$ with $-\alpha^{3.7}/\alpha\alpha$, 1 was $\beta^{HBB:c.-23A>G}/\beta^{HBB:c.126_129delCTTT}$ with $\alpha\alpha/\alpha\alpha$, and 1 was $\beta^{HBB:c.-23A>G}/\beta^{HBB:c.316-197C>T}$ with $\alpha\alpha/\alpha\alpha$. Pedigree follow-up revealed that four carriers had spouses who were also β -thalassemia carriers, and two carriers were sisters (Supplementary Table S2).

3.2 Mutation verification

Sanger sequencing of DNA samples harboring the $HBB: c.-23A>G$ mutation validated all NGS results, confirming the accuracy of the NGS method (Figure 1A). Moreover, third-generation sequencing provided additional verification and confirmed the absence of other pathogenic HBB gene mutations (Figure 1B).

For the carrier with $\beta^{HBB: c.-23A>G}/\beta^{HBB: c.316-197C>T}$ genotype, pedigree verification revealed that $HBB: c.-23A>G$ was inherited from the

father, while $HBB: c.316-197C>T$ was inherited from the mother. Additionally, her daughter inherited the $HBB: c.-23A>G$ mutation (Figure 2A). In another carrier with $\beta^{HBB: c.-23A>G}/\beta^{HBB: c.126_129delCTTT}$ genotype, pedigree verification showed that $HBB: c.-23A>G$ was inherited from the mother, while $HBB: c.126_129delCTTT$ was inherited from the father. In addition, his son inherited the $HBB: c.126_129delCTTT$ mutation (Figure 2B). Notably, among the 18 families with heterozygous carriers, the mutation was transmitted to the next generation in 8 cases (Supplementary Table S3).

3.3 Hematological analysis

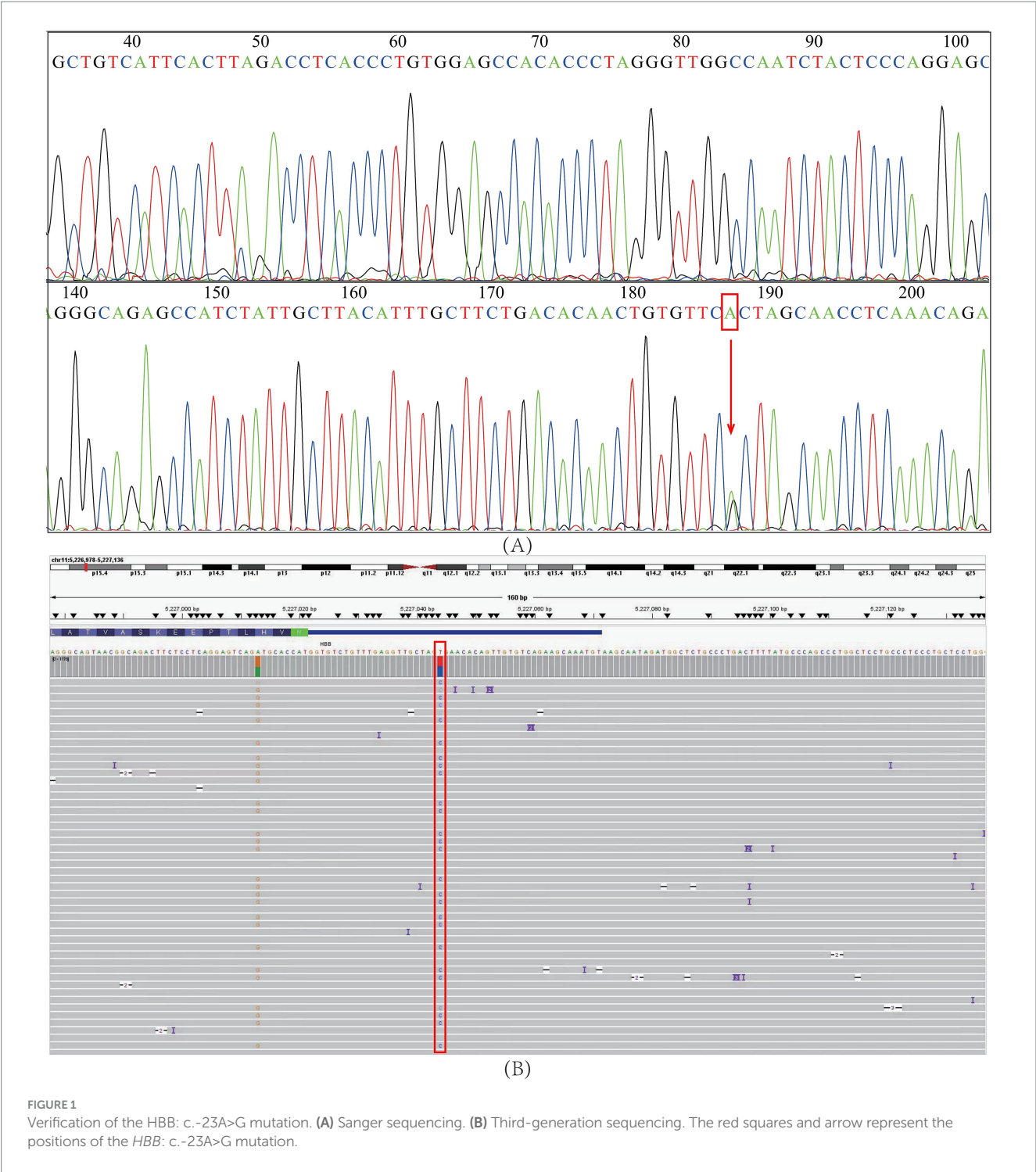
After excluding the carriers with iron-deficiency anemia, we obtained the hematological data of 69 heterozygotes, 6 heterozygotes with $\alpha\alpha/--^{SEA}$, 1 heterozygote with $\alpha\alpha/-\alpha^{3.7}$, and 2 compound heterozygotes (Supplementary Table S4). The mean corpuscular volume (MCV) of the 69 heterozygotes was 92.07 ± 3.94 fL, while the mean corpuscular hemoglobin (MCH) was 30.58 ± 1.49 pg, and the mean hemoglobin A2 (Hb A2) was $2.70 \pm 0.17\%$. In comparison, the MCV, MCH, and Hb A2 of the 6 heterozygotes with $\alpha\alpha/--^{SEA}$, were 70.07 ± 1.99 fL, 21.58 ± 0.67 pg, and $2.28 \pm 0.07\%$, respectively (Table 1).

After comparing with hematological data of control group, the MCV, MCH, and Hb A2 of the 69 heterozygotes showed significant differences ($p < 0.001$) with those of β^0 heterozygote carriers, β^+ heterozygote carriers, and $HBB: c.-29G>A$ heterozygotes. In contrast, there were no significant differences in MCV, MCH, and Hb A2 between the 69 heterozygotes and the normal population or $HBB: c.-11_-8delAACA$ heterozygotes (Figure 3). There were no statistically significant differences in Hb A2 between the 6 compound $\alpha\alpha/--^{SEA}$ heterozygotes and the β^N/β^N with $--^{SEA}/\alpha\alpha$ group. However, significant differences were observed when compared with the normal population, the $\beta^{HBB: c.-23A>G}/\beta^N$ with $\alpha\alpha/\alpha\alpha$ group, the β^0/β^N with $--^{SEA}/\alpha\alpha$ group, and the β^+/ β^N with $--^{SEA}/\alpha\alpha$ group (Figure 3D).

The two compound heterozygotes ($\beta^{HBB: c.-23A>G}/\beta^{HBB: c.316-197C>T}$ and $\beta^{HBB: c.-23A>G}/\beta^{HBB: c.126_129delCTTT}$) exhibited hemoglobin levels of 125 and 130 g/L, respectively, with microcytic hypochromic erythrocytes (MCV: 68.8 fL and 62.8 fL; MCH: 21.6 pg and 20.6 pg) and elevated Hb A2 (4.8 and 5.4%). No elevation of Hb F was found in either of them (Table 1). Both individuals were asymptomatic and did not require transfusion or chelation therapy. Their clinical presentations were consistent with β -thalassemia trait rather than thalassemia intermedia or major.

3.4 Secondary structure prediction of mRNA

Secondary structures of wild-type and mutant ($HBB: c.-23A>G$ and $HBB: c.-29G>A$) HBB gene mRNA were predicted using the RNAfold web server. We analyzed the minimum free energy (MFE) structures, centroid structures, and per-nucleotide binding probabilities (Figures 4A–I). The MFE values for the wild-type, $HBB: c.-23A>G$ mutant, and $HBB: c.-29G>A$ mutant mRNAs were -224.10 kcal/mol (Figure 4A), -222.90 kcal/mol (Figure 4D), and -221.60 kcal/mol (Figure 4G), respectively. Statistical analysis revealed that the differences among these three MFEs were not statistically significant ($p > 0.05$). Compared with the wild type, the HBB gene mRNA of the $HBB: c.-23A>G$ mutant exhibited only a change from



AU pairing to GU pairing in the first internal loop (Figure 4D). In contrast, the HBB gene mRNA of the HBB: c.-29G>A mutant altered the structure of the hairpin loop (Figure 4G).

4 Discussion

β-thalassemia exhibits a broad global distribution with striking heterogeneity in its mutational spectrum across geographic and ethnic populations. This diversity is evident in the

distinct profiles of common pathogenic alleles, such as the IVS-I-110, IVS-I-1, and IVS-I-6 variants prevalent in Turkey and Lebanon (16, 17), which contrast sharply with the IVS-II-654 (HBB: c.316-197C>T) and CD41-42 (HBB: c.126_129delCTTT) mutations common in China. This principle of geographic stratification extends to the regulatory 5' UTR, where despite a limited number of reported mutations, distinct ethnogeographic patterns are observed: HBB: c.-29G>A is primarily found in Black Sea-bordering countries like Turkey and Bulgaria (18), HBB: c.-50A>C has been identified in South Asian Indian populations

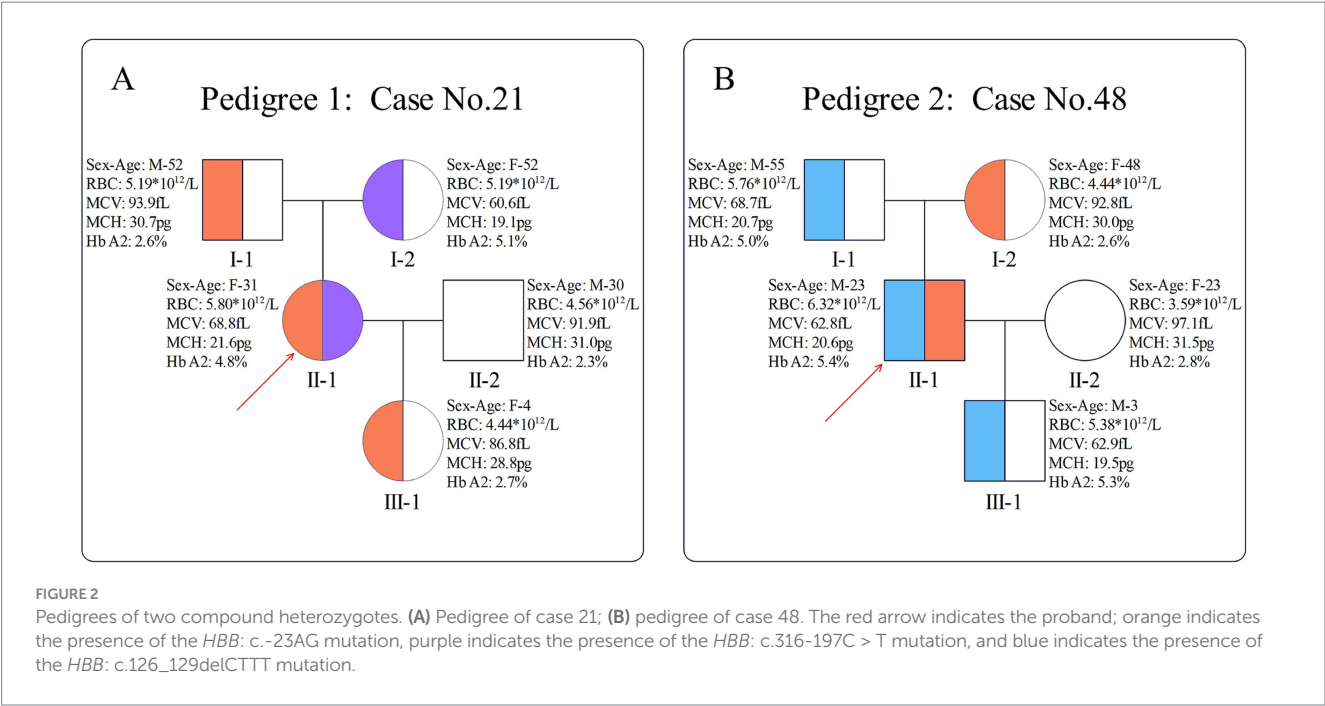


TABLE 1 Hematological data statistics of HBB: c.-23A>G mutation carriers.

Genotype	Number of cases	RBC (10 ¹² /L)	HB (g/L)	MCV (fL)	MCH (pg)	Hb A2 (%)
$\alpha\alpha/\alpha\beta^{HBB: c.-23A>G/\beta^N}$	69	4.45 ± 0.65	135.51 ± 17.84	92.07 ± 3.94	30.58 ± 1.49	2.70 ± 0.17
	25 (Male)	5.07 ± 0.53	153.44 ± 13.97	91.21 ± 3.40	30.30 ± 1.14	2.66 ± 0.17
	44 (Female)	4.09 ± 0.40	125.32 ± 10.10	92.57 ± 4.14	30.73 ± 1.64	2.72 ± 0.17
$\alpha\alpha/-_{SEA}\beta^{HBB: c.-23A>G/\beta^N}$	6	5.60 ± 0.33	121.00 ± 9.42	70.07 ± 1.99	21.58 ± 0.67	2.28 ± 0.07
	2 (Male)	5.36, 6.25	110, 140	6.95, 70.1	20.4, 22.4	2.2, 2.3
	4 (Female)	5.50 ± 0.18	119 ± 2.92	70.20 ± 2.42	21.68 ± 0.40	2.30 ± 0.07
$\alpha\alpha/-\alpha^{3.7}\beta^{HBB: c.-23A>G/\beta^N}$	1 (Male)	5.88	160	82.6	27.2	2.6
$\alpha\alpha/\alpha\beta^{HBB: c.-23A>G/\beta^{HBB: c.316-197C>T}}$	1 (Female)	5.80	125	68.8	21.6	4.8
$\alpha\alpha/\alpha\beta^{HBB: c.-23A>G/\beta^{HBB: c.126_129delCTTT}}$	1 (Male)	6.32	130	62.8	20.6	5.4

(19), and HBB: c.-11_8delAACA demonstrates high frequency in China (20).

Functional studies suggest that within this short 50-nucleotide 5' UTR, different "sub-regions" may be associated with distinct regulatory functions (21). The complex interplay between genotype and phenotype (22), further complicated by population migration spreading thalassemia beyond traditional endemic regions (23), must be considered when evaluating any variant's clinical significance. Phenotypic variability, even for the same mutation across different reports, often presents diagnostic challenges for clinicians (24). Many studies have attempted to address this challenge by using online prediction tools and sequence interpretation databases, such as the Hb Var database (25), Itha Genes (26), and the LOVD database (27). However, the American College of Medical Genetics and Genomics (ACMG) guidelines explicitly caution against overreliance on in silico pathogenicity predictions, stating they "should not be used as standalone

evidence for clinical interpretation" (8). The ACMG framework mandates rigorous evaluation of multiple evidentiary categories, emphasizing that comprehensive evaluation of published clinical data remains imperative when assessing potential variant pathogenicity (28).

Through literature retrieval, we found that the HBB: c.-23A>G variant was first submitted to the ClinVar database by Women's Health and Genetics/Laboratory Corporation of America, LabCorp on March 17, 2018 (Variant ID: rs1010004981) (29). Given the lack of clinical information and functional studies, the variant is classified as a variant of uncertain significance (VUS). In addition, this study identified descriptions of this site in two publications, both of which reported case information of one individual carrying the mutation from China (2, 30). Nonetheless, neither of these two publications conducted in-depth research on this mutation. Based on the information retrieved, it is indeed challenging to determine the pathogenicity of this mutation.

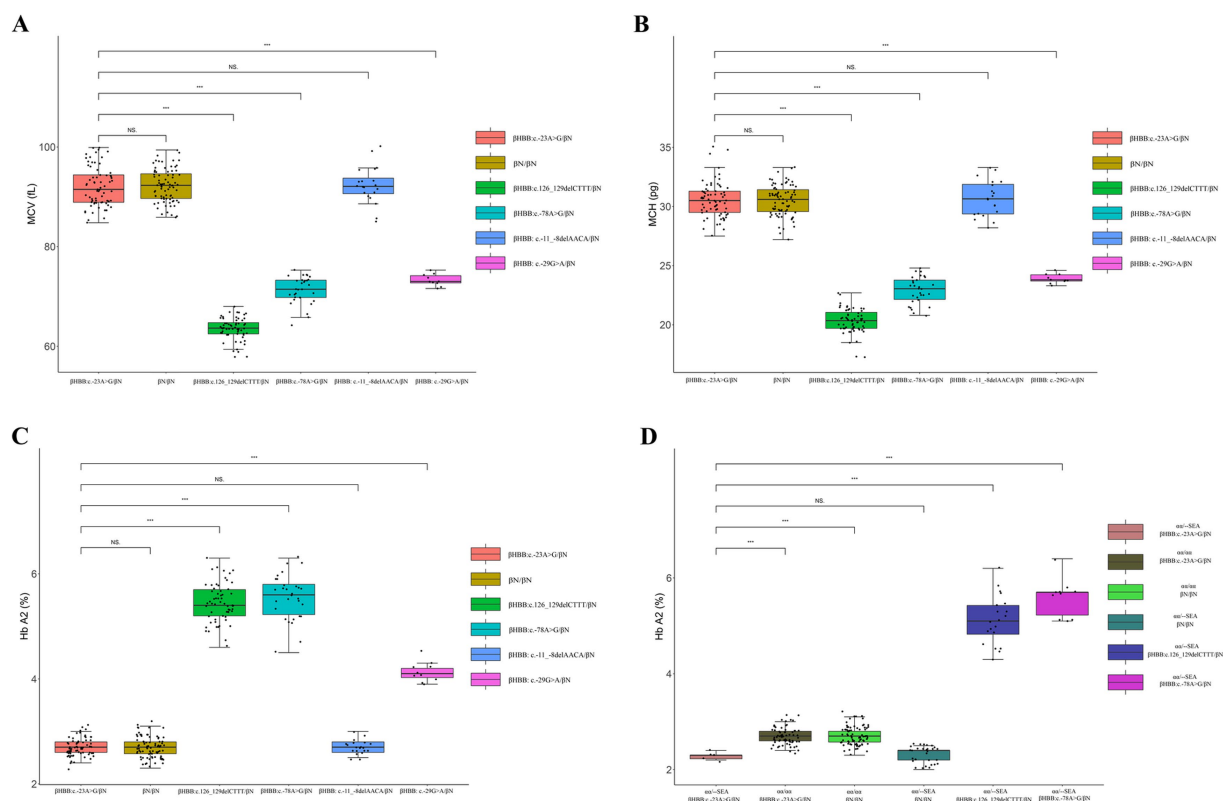


FIGURE 3

Comparison of hematological data between *HBB* c.-23A>G mutation carriers and the control group. (A) Comparison of MCV between *HBB* c.-23A>G heterozygotes and the control group; (B) comparison of MCH between *HBB* c.-23A>G heterozygotes and the control group; (C) comparison of Hb A2 between *HBB* c.-23A>G heterozygotes and the control group; (D) comparison of Hb A2 between the $\beta^{\text{HBB: c.-23A>G}}/\beta^{\text{N}}$ with $\alpha\alpha/\alpha\alpha$ group and the control group. "NS" indicates no statistically significant difference between the two groups, while "****" signifies an extremely significant statistical difference between the two groups ($p < 0.001$).

In previous studies, the carrier rate of β -thalassemia in this region was relatively high (5), and the carrier rate in this study is consistent with the previously reported data (4.134% vs. 4.056%). The prevalence of the *HBB* c.-23A>G variant in this region also represents a non-negligible frequency. Through follow-up, we found that among 75 carriers, four of them had spouses who were also carriers of β -thalassemia. Given that when both husband and wife are carriers of β -thalassemia, there is a risk of having a child with β -TM (31), we performed prenatal diagnosis for one couple when we first identified this variant, due to the unclear phenotype and limitations in genetic counseling capabilities. Fortunately, the diagnosis showed that the fetus is $\beta^{\text{HBB: c.316-197C>T}}/\beta^{\text{N}}$, highlighting the particular necessity to study the clinical phenotype of *HBB* c.-23A>G. The literature indicates that globin chain synthesis studies are the established standard for determining the pathogenicity of a variant, as they help to determine whether the variant has led to a reduction in β -globin chains (8). In this study, we performed hemoglobin electrophoresis analysis and hemoglobin component analysis on *HBB* c.-23A>G carriers in our region to assess the hematological phenotype and indirectly infer whether *HBB* c.-23A>G truly causes a reduction in β -globin chains.

This study revealed that the MCV, MCH, and Hb A2 levels of heterozygotes with this mutation showed no significant statistical differences compared with those of the normal population, aligning with a previous report (2). Additionally, we observed that the MCV,

MCH, and Hb A2 levels of compound heterozygotes carrying this variant and $\beta^{0/+}$ showed no statistical differences compared with those of $\beta^{0/+}$ heterozygotes alone. The MCV, MCH, and Hb A2 levels of a previously reported case of $\beta^{\text{HBB: c.316-197C>T}}/\beta^{\text{HBB: c.-23A>G}}$ compound heterozygote also were consistent with the results of this study (30). Crucially, none of these cases exhibited the compensatory high Hb F expression observed in β -TM. The level of Hb A2 is considered an important indicator for assessing the expression of α - and β -globin chains (32). When α -thalassemia occurs, the reduced synthesis of α -globin chains leads to a decrease in Hb A2 levels. In contrast, when β -thalassemia occurs, the reduced synthesis of β -globin chains results in an increase in Hb A2 levels. However, in patients who carry both α - and β -thalassemia, despite the reduced synthesis of both α - and β -globin chains, the Hb A2 levels are elevated. We observed that in the heterozygotes for this variant, their Hb A2 levels did not increase. In the cases of $\beta^{\text{HBB: c.-23A>G}}/\beta^{\text{N}}$ with $--^{\text{SEA}}/\alpha\alpha$, the Hb A2 levels did not increase but decreased instead. Based on the above data, this study concluded that the *HBB* c.-23A>G mutation does not lead to a reduction in the production of β -globin chains. This mutation is likely to be mainly distributed in the Chinese population, and its phenotype is consistent with silent β^{++} phenotype. According to the ACMG guidelines, genetic interpretation experts, in conjunction with the data from this study, classified the *HBB* c.-23A>G mutation as likely benign (BS4 and BP4).

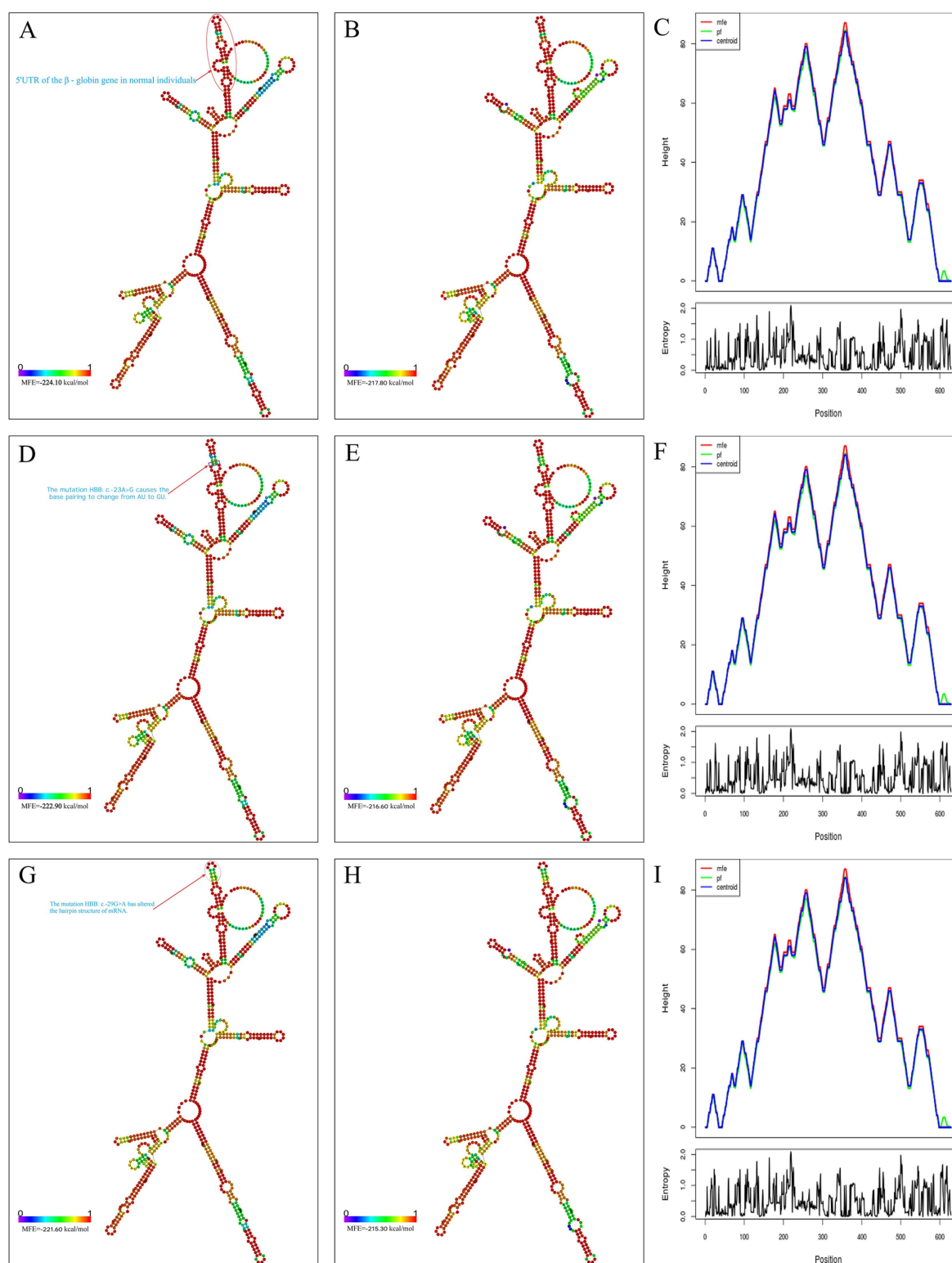


FIGURE 4

Prediction of mRNA secondary structure. **(A)** MFE structure drawing encoding base-pair probabilities in normal individuals; **(B)** centroid structure drawing encoding base-pair probabilities in normal individuals; **(C)** the binding strength at each site on mRNA in normal individuals; **(D)** MFE structure drawing encoding base-pair probabilities in *HBB*: c.-23A>G mutants; **(E)** centroid structure drawing encoding base-pair probabilities in *HBB*: c.-23A>G mutants; **(F)** the binding strength at each site on mRNA in *HBB*: c.-23A>G mutants; **(G)** MFE structure drawing encoding base-pair probabilities in *HBB*: c.-29G>A mutants; **(H)** centroid structure drawing encoding base-pair probabilities in *HBB*: c.-29G>A mutants; **(I)** the binding strength at each site on mRNA in *HBB*: c.-29G>A mutants.

Compared to other known 5' UTR mutations, *HBB*: c.-23A>G demonstrates a distinct phenotypic profile. This region exhibits a spectrum of phenotypic outcomes across different populations: the *HBB*: c.-29G>A mutation, prevalent in Turkey and Bulgaria, is associated with a typical β^+ thalassemia phenotype (18); the *HBB*: c.-50A>C mutation, identified in South Asian Indian populations, presents a β^{++} phenotype (19); and the *HBB*: c.-11_-8delAACA mutation, common in China, manifests as a silent β^{++} phenotype (20). Our study establishes *HBB*: c.-23A>G as a silent β^{++} variant, based on the absence of significant hematological alterations in heterozygotes. This marked variability underscores the functional diversity within the 5' UTR and cautions against overgeneralizing the effects of a given mutation across different populations.

The 5' UTR region downstream of the cap site can be transcribed into mRNA. Mutations in this region primarily reduce β -chain synthesis by affecting mRNA stability and translational efficiency (33). As early as 2003, Sgourou found that the stability of mRNA containing four mutation sites in the 5' UTR of the β -globin gene (*HBB*: c.-43C>T, *HBB*: c.-41delT, *HBB*: c.-29G>A, *HBB*: c.-11_-8delAACA) was reduced to varying degrees (34). However, research in recent years and data from our region indicated that the phenotype of *HBB*: c.-11_-8delAACA may be a benign variant (20), suggesting that the stability of mRNA at this site may not change significantly.

It is widely acknowledged that mRNA sequences are typically composed of permutations of four nucleotides (A, C, G, U). In base pairing interactions, AU and GC pairs form the most stable configurations through Watson-Crick base pairing, while GU pairs constitute less stable wobble base pairings (33). These nucleotide pairings serve as fundamental components of RNA secondary structure (35). The thermodynamic stability of Watson-Crick base pairs (GC > AU) significantly influences RNA structural formation (36), with GU wobble pairs providing structural flexibility while maintaining partial pairing stability (37). This structural organization plays crucial roles in post-transcriptional regulation, particularly in mRNA processing and translation initiation mechanisms (38). Using the RNAfold web server to predict the secondary structure of mRNA (15), we found that the secondary structure of mRNA with the *HBB*: c.-23A>G mutation did not significantly differ in stability compared to the normal sequence, and it did not alter the hairpin loop or stem-loop structures. The only change was that the alteration of this nucleotide caused the first internal loop of the mRNA to change from a stable Watson-Crick pair (AU) to a less stable wobble base pair (GU). This may be the reason why this variant has almost no impact on β -chain synthesis. In contrast, another mutation in this region (*HBB*: c.-29G>A), manifests as typical β^+ thalassemia (39), likely because this mutation affects the hairpin loop of the mRNA.

A limitation of this study is the absence of functional assays, such as in vitro globin chain synthesis studies or luciferase reporter assays, to directly quantify the impact of the *HBB*: c.-23A>G mutation on translational efficiency or mRNA stability. Furthermore, the potential clinical consequence of co-inheritance with α -globin triplication remains unexplored. Previous reports have indicated that the coexistence of β -thalassemia and α -globin triplication can exacerbate the α/β -globin chain imbalance, potentially leading to a more severe anemic phenotype (40). In a prior investigation from our laboratory, 1,443 carriers of α -globin triplication were identified among 73,967 individuals in this region, yielding a carrier rate of 1.95% (41).

Regrettably, due to the current sample size limitations, none of the 75 carriers of the *HBB*: c.-23A>G mutation in this study were found to co-inherit α -globin triplication. Consequently, we were unable to investigate the clinical manifestations of the combination of *HBB*: c.-23A>G and $\alpha\alpha\alpha$. It is anticipated that future studies with larger sample sizes may identify such compound cases, allowing for a comprehensive assessment of their phenotypic impact.

5 Conclusion

In conclusion, we systematically characterized the clinical phenotype associated with the *HBB*: c.-23A>G mutation and elucidated the molecular rationale for its likely benign clinical manifestations. Additionally, comparative analysis of mRNA secondary structures demonstrated minimal thermodynamic destabilization compared to the wild-type sequence, with preserved hairpin loop and stem-loop architectures. Collectively, our findings provide valuable insights for improving genetic counseling for carriers of this mutation in clinical practice.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material. Raw sequencing data are available from the European Nucleotide Archive (ENA) under accession number PRJEB88162 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB88162>).

Ethics statement

The studies involving humans were approved by The Ethics Committee of Ganzhou Maternal and Child Health Hospital (No. 2024108). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

SS: Conceptualization, Data curation, Investigation, Writing – original draft. JH: Funding acquisition, Investigation, Project administration, Writing – review & editing. HQ: Formal analysis, Methodology, Validation, Writing – original draft. ZL: Formal analysis, Investigation, Software, Writing – original draft. WZ: Formal analysis, Resources, Visualization, Writing – review & editing. XY: Methodology, Validation, Writing – review & editing. ZZ: Data curation, Validation, Writing – review & editing. HC: Formal analysis, Software, Writing – original draft. XX: Investigation, Supervision, Validation, Writing – review & editing. LX: Funding acquisition, Supervision, Writing – review & editing. JC: Funding acquisition,

Project administration, Supervision, Writing – review & editing. LS: Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by the Health Commission Program of Ganzhou (no. GZWJW202502388) and Jiangxi Provincial Natural Science Foundation (no. 20232BAB216098).

Acknowledgments

We thank the patients for their willingness to participate in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Tesio N, Bauer DE. Molecular basis and genetic modifiers of thalassemia. *Hematol Oncol Clin North Am.* (2023) 37:273–99. doi: 10.1016/j.hoc.2022.12.001
2. Peng Q, Zhang Z, Li S, Cheng C, Li W, Rao C, et al. Molecular epidemiological and hematological profile of thalassemia in the Dongguan region of Guangdong Province, southern China. *J Clin Lab Anal.* (2021) 35:e23596. doi: 10.1002/jcla.23596
3. Xi H, Liu Q, Xie DH, Zhou X, Tang WL, Tang G, et al. Epidemiological survey of hemoglobinopathies based on next-generation sequencing platform in Hunan Province, China. *Biomed Environ Sci.* (2023) 36:127–34. doi: 10.3967/bes2023.016
4. Xiong F, Sun M, Zhang X, Cai R, Zhou Y, Lou J, et al. Molecular epidemiological survey of haemoglobinopathies in the Guangxi Zhuang autonomous region of southern China. *Clin Genet.* (2010) 78:139–48. doi: 10.1111/j.1399-0004.2010.01430.x
5. Yang T, Luo X, Liu Y, Lin M, Zhao Q, Zhang W, et al. Next-generation sequencing analysis of the molecular spectrum of thalassemia in southern Jiangxi, China. *Hum Genomics.* (2023) 17:77. doi: 10.1186/s40246-023-00520-5
6. Angastiniotis M, Petrou M, Loukopoulos D, Modell B, Farmakis D, Englezos P, et al. The prevention of thalassemia revisited: a historical and ethical perspective by the thalassemia international federation. *Hemoglobin.* (2021) 45:5–12. doi: 10.1080/03630269.2021.1872612
7. Babendure JR, Babendure JL, Ding JH, Tsien RY. Control of mammalian translation by mRNA structure near caps. *RNA.* (2006) 12:851–61. doi: 10.1261/rna.2309906
8. Colaco S, Nadkarni A. Borderline HbA(2) levels: dilemma in diagnosis of beta-thalassemia carriers. *Mutat Res Rev Mutat Res.* (2021) 788:108387. doi: 10.1016/j.mrrev.2021.108387
9. Leppek K, Das R, Barna M. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol.* (2018) 19:158–74. doi: 10.1038/nrm.2017.103
10. Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell.* (1986) 44:283–92. doi: 10.1016/0092-8674(86)90762-2
11. Hamid M, Zargan Nezhad E, Keikhaei B, Galehdari H, Saberi A, Sedaghat A, et al. Two novel and five rare mutations in the non coding regions of the β -globin gene in the Iranian population. *Hemoglobin.* (2020) 44:225–30. doi: 10.1080/03630269.2020.1790384
12. Vinciguerra M, Passarello C, Cassarà F, Leto F, Cannata M, Calvaruso G, et al. Coheredity of a new silent mutation: c.-29G>T, with a severe β -thal mutation in a patient with β -thalassaemia intermediate. *Int J Lab Hematol.* (2018) 40:e17–20. doi: 10.1111/ijlh.12777
13. Russo R, Marra R, Rosato BE, Iolascon A, Andolfo I. Genetics and genomics approaches for diagnosis and research into hereditary Anemias. *Front Physiol.* (2020) 11:613559. doi: 10.3389/fphys.2020.613559
14. Suhaimi SA, Zulkipli IN, Ghani H, Abdul-Hamid MRW. Applications of next generation sequencing in the screening and diagnosis of thalassemia: a mini-review. *Front Pediatr.* (2022) 10:1015769. doi: 10.3389/fped.2022.1015769
15. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA websuite. *Nucleic Acids Res.* (2008) 36:W70–4. doi: 10.1093/nar/gkn188
16. Aydinok Y, Oymak Y, Atabay B, Aydoğan G, Yeşilipek A, Ünal S, et al. A National Registry of thalassemia in Turkey: demographic and disease characteristics of patients, achievements, and challenges in prevention. *Turk J Haematol.* (2018) 35:12–8. doi: 10.4274/tjh.2017.0039
17. Makhoul NJ, Wells RS, Kaspar H, Shbaklo H, Taher A, Chakar N, et al. Genetic heterogeneity of Beta thalassemia in Lebanon reflects historic and recent population migration. *Ann Hum Genet.* (2005) 69:55–66. doi: 10.1046/j.1529-8817.2004.00138.x
18. Oner R, Oner C, Birben E, Sözen M, Gümrük F, Gürgey A, et al. beta-thalassaemia intermedia in a Turkish girl: homozygosity for G-→A substitution at +22 relative to the beta-globin cap site. *Br J Haematol.* (2001) 115:90–4. doi: 10.1046/j.1365-2141.2001.03071.x
19. Karim MU, Moinuddin M, Babar SU. Cap +1 mutation; an unsuspected cause of beta thalassaemia transmission in Pakistan. *Turk J Haematol.* (2009) 26:167–70. Available at: <https://jag.journalagent.com/z4/vi.asp?pdire=jh&plng=eng&un=TJH-25338&look4=>
20. Li DZ, Liao C, Li J, Xie XM, Zhong HZ. The 4-bp deletion (-AAAC) in the 5' untranslated region of the beta-globin gene: a simple polymorphism? *Ann Hematol.* (2009) 88:709–10. doi: 10.1007/s00277-008-0651-0
21. Ho PJ, Sloane-Stanley J, Athanassiadou A, Wood WG, Thein SL. An in vitro system for expression analysis of mutations of the beta-globin gene: validation and application to two mutations in the 5' UTR. *Br J Haematol.* (1999) 106:938–47. doi: 10.1046/j.1365-2141.1999.01659.x
22. Diamantidis MD, Karanikola RA, Polyzoudi C, Delicou S, Manafas A, Savera H, et al. Clinical significance of mutational variants in beta and alpha genes in patients with hemoglobinopathies from two large Greek centers: a complex interplay between genotype and phenotype. *J Mol Med (Berl).* (2023) 101:1073–82. doi: 10.1007/s00109-023-02342-3
23. Shchemeleva E, Salomashkina VV, Selivanova D, Tsvetaeva N, Melikyan A, Doronina L, et al. Active spread of β -thalassaemia beyond the thalassaemia belt: a study on a Russian population. *Clin Genet.* (2025) 107:23–33. doi: 10.1111/cge.14606
24. Stephanou C, Petrou M, Kountouris P, Makariou C, Christou S, Hadjigavriel M, et al. Unravelling the complexity of the +33 C>G [HBB:c.-18C>G] variant in Beta thalassaemia. *Biomedicine.* (2024) 12:296. doi: 10.3390/biomedicine12020296

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2025.1675600/full#supplementary-material>

25. Giardine B, Borg J, Viennas E, Pavlidis C, Moradkhani K, Joly P, et al. Updates of the HbVar database of human hemoglobin variants and thalassemia mutations. *Nucleic Acids Res.* (2014) 42:D1063–9. doi: 10.1093/nar/gkt911
26. Kountouris P, Lederer CW, Fanis P, Feleki X, Old J, Kleanthous M. Ithagenes: an interactive database for haemoglobin variations and epidemiology. *PLoS One.* (2014) 9:e103020. doi: 10.1371/journal.pone.0103020
27. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat.* (2011) 32:557–63. doi: 10.1002/humu.21438
28. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* (2015) 17:405–24. doi: 10.1038/gim.2015.30
29. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* (2018) 46:D1062–d1067. doi: 10.1093/nar/gkx1153
30. Yang Y, Lu YH, Zhang Y, Zhao AL, Zhan XY, Wu B, et al. The genotypes and clinical characteristics of thalassemia on children in Wuhan region. *Zhongguo Shi Yan Xue Ye Xue Za Zhi.* (2021) 29:1875–80. doi: 10.19746/j.cnki.issn.1009-2137.2021.06.031
31. Zhang R, Zhang S, Ming J, Xie J, Liu B, Jiang W, et al. A cross-sectional study: caregiver burden and related determinants of adult patients with β -thalassemia major in mainland China. *BMC Nurs.* (2024) 23:151. doi: 10.1186/s12912-024-01826-y
32. Giambona A, Passarello C, Renda D, Maggio A. The significance of the hemoglobin a(2) value in screening for hemoglobinopathies. *Clin Biochem.* (2009) 42:1786–96. doi: 10.1016/j.clinbiochem.2009.06.026
33. Dever TE, Dinman JD, Green R. Translation elongation and recoding in eukaryotes. *Cold Spring Harb Perspect Biol.* (2018) 10:2649. doi: 10.1101/cshperspect.a032649
34. Sgourou A, Routledge S, Antoniou M, Papachatzopoulou A, Psiouri L, Athanassiadou A. Thalassemia mutations within the 5'UTR of the human beta-globin gene disrupt transcription. *Br J Haematol.* (2004) 124:828–35. doi: 10.1111/j.1365-2141.2004.04835.x
35. Higgs PG. RNA secondary structure: physical and computational aspects. *Q Rev Biophys.* (2000) 33:199–253. doi: 10.1017/s0033583500003620
36. Takahashi S, Sugimoto N. Watson-crick versus Hoogsteen Base pairs: chemical strategy to encode and express genetic information in life. *Acc Chem Res.* (2021) 54:2110–20. doi: 10.1021/acs.accounts.0c00734
37. Xu D, Landon T, Greenbaum NL, Fenley MO. The electrostatic characteristics of G.U wobble base pairs. *Nucleic Acids Res.* (2007) 35:3836–47. doi: 10.1093/nar/gkm274
38. Kozak M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene.* (2005) 361:13–37. doi: 10.1016/j.gene.2005.06.037
39. Akar E, Ozdemir S, Hakki Timur I, Akar N. First observation of homozygous hemoglobin Hamadan (B 56 (D7) GLY-ARG) and beta thalassemia (–29 G>a)-hemoglobin Hamadan combination in a Turkish family. *Am J Hematol.* (2003) 74:280–2. doi: 10.1002/ajh.10404
40. Ropero P, González Fernández FA, Nieto JM, Torres-Jiménez WM, Benavente C. β -Thalassemia intermedia: interaction of α -globin gene triplication with β -thalassemia heterozygous in Spain. *Front Med (Lausanne).* (2022) 9:866396. doi: 10.3389/fmed.2022.866396
41. Xie X, Gan J, Liu Z, Zhou Y, Yuan K, Chen Z, et al. Prevalence and genetic analysis of triplicated α -globin gene in Ganzhou region using high-throughput sequencing. *Front Genet.* (2023) 14:1267892. doi: 10.3389/fgene.2023.1267892