BMC Medical Genomics

RESEARCH

Open Access

Application of third-generation sequencing technology for identifying rare α- and β-globin gene variants in a Southeast Chinese region

Jianlong Zhuang[1*†], Junyu Wang[1*†], Nan Huang[2], Yu Zheng[3] and Liangpu Xu[4]

Abstract

Background Third-generation sequencing (TGS) based on long-read technology has been gradually used in identifying thalassemia and hemoglobin (Hb) variants. The aim of the present study was to explore genotype varieties of thalassemia and Hb variants in Quanzhou region of Southeast China by TGS.

Methods Included in this study were 6,174 subjects with thalassemia traits from Quanzhou region of Southeast China. All of them underwent common thalassemia gene testing using the DNA reverse dot-blot hybridization technology. Subjects who were suspected as rare thalassemia carriers were further subjected to TGS to identify rare or novel α- and β-globin gene variants, and the results were verified by Sanger sequencing and/or gap PCR.

Results Of the 6,174 included subjects, 2,390 (38.71%) were identified as α- and β-globin gene mutation carriers, including 40 carrying rare or novel α- and β-thalassemia mutations. The $\alpha^{CD30(-GAG)}\alpha$ and Hb Lepore-Boston-Washington were first reported in Fujian province Southeast China. Moreover, the $\beta^{CD15(TGG>TAG)}$, $\beta^{IVS-II-761}$, β^{0}-Filipino(~45 kb deletion), and Hb Lepore-Quanzhou were first identified in the Chinese population. In addition, 35 cases of Hb variants were detected, the rare Hb variants of Hb Jilin and Hb Beijing were first reported in Fujian province of China. Among them, one case with compound $\alpha\alpha\alpha^{anti3.7}$ and Hb G-Honolulu variants was identified in this study.

Conclusion Our findings may provide valuable data for enriching the spectrum of thalassemia and highlight the clinical application value of TGS-based α- and β-globin genetic testing.

Keywords Thalassemia, Hb variants, Third-generation sequencing, Sanger sequencing

†Jianlong Zhuang, Junyu Wang these authors contribute equally to this work.

*Correspondence:
Jianlong Zhuang
415913261@qq.com
Junyu Wang
86813685@qq.com
[1]Prenatal Diagnosis Center, Quanzhou Women's and Children's Hospital, Quanzhou 362000, Fujian Province, China
[2]The teaching and research office of clinical laboratory medicine, Quanzhou Medical College, Quanzhou 362000, China
[3]Research and Development Department, Yaneng BIOscience (Shenzhen) Co. Ltd, Shenzhen 518000, Guangdong Province, China
[4]Medical Genetic Diagnosis and Therapy Center, Fujian Key Laboratory for Prenatal Diagnosis and Birth Defect, Fujian Maternity and Child Health Hospital College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou 350001, Fujian Province, China

Introduction

Thalassaemia is a worldwide hereditary blood disorder affecting globin chain synthesis, which results in the unbalanced globin chain production, leading to ineffective erythropoiesis, increased hemolysis, and deranged iron homoeostasis [1, 2]. The α- and β- thalassemia are the most common thalassemia, showing high prevalences in the regions such as Mediterranean, Middle East, Indian subcontinent, Southeast Asia and South China [3, 4]. The incidence of thalassemia in Fujian province of Southeast China is about 6.8% [5, 6]. Quanzhou, which has the largest population in Fujian, shows a notable diversity and complexity of thalassemia and hemoglobin (Hb) variants. Rare and novel α- and β-globin gene variants have been increasingly identified in Quanzhou region [7–10] .

No effective medical treatment for thalassemia intermedia or major is available at present. Conventional thalassemia gene detection before marriage or pregnancy or prenatal diagnosis is the most effective and low-cost intervention mean to prevent and control of the birth defect of thalassemia major or intermedia. However, conventional thalassemia genetic detection based on DNA reverse dot blot hybridization and gap-PCR often leads to misdiagnosis of rare and novel thalassemia [11]. In recent years, the third-generation sequencing (TGS) based on long reads has been used in identifying α- and β-globin gene variants, manifesting remarkable advantages in identifying single-nucleotide variants, indels and structural variants [12, 13]. A prospective study conducted by Xu et al. [14] has demonstrated the application value of TGS technology in molecular diagnosis of thalassemia. Subsequently, a series of studies has reported more rare or novel thalassemia by using TGS [15–17] .

The aim of the present study was intended to identify rare α- and β-globin gene variants in Quanzhou of Southeast China by using the TGS technology, hoping that our experience and practice could help establish a more optimized method for the prevention and treatment of thalassemia.

Materials and methods

Samples

This prospective study enrolled 6,174 subjects with positive hematological screening results or one of the couples was identified as a thalassemia carrier from April 2021 to October 2023 at Quanzhou Women's and Children's Hospital, excluding those reported in our previous case reports from January 2013 to March 2021 [8]. They included 2,595 males and 3,579 females ranging in age from newborns to 54 years. All cases deny recent blood transfusions within 14 days, and no genetic relationship was exists in all samples. No included subject received transfusion within 14 days before testing, and no genetic relation existed between the samples. Informed consent was obtained from all subjects before conventional thalassemia gene detection. Subjects who were suspected as a rare thalassemia or Hb variants carriers (individuals with positive hematological screening results but elicited negative conventional thalassemia gene testing results) were subjected to α- and β-globin gene sequencing based on TGS technology. This study was approved by the ethics committee of The Women's and Children's Hospital of Quanzhou (No.2021No.61).

Hematological screening analysis

Routine blood analysis and Hb electrophoresis were carried out for hematological screening analysis. Approximately 4 ml peripheral blood was collected from each subject and anticoagulated with EDTA-K_2. Routine blood detection was conducted using an automated cell counter (Sysmex XS-1000i; Sysmex Co., Ltd., Kobe, Japan). The Hb components were analysed by Hb electrophoresis (Sebia, Evry Cedex, France). Positive hematological screening results that met one of the following indexes were diagnosed (which was only suitable for individuals older than 2 years) [8]: a mean corpuscular volume (MCV) ≤ 82 fl., and/or a mean corpuscular hemoglobin (MCH) ≤ 27 pg, and/or Hb A2 levels ≥ 3.4% or ≤ 2.6%, or an Hb F value ≥ 2.0%. All patients with positive hematological analysis results underwent thalassemia gene testing.

Conventional thalassemia gene testing

About 2 ml peripheral blood was collected from each of the enrolled subjects for conventional thalassemia gene testing. Genomic DNA was extracted using the automatic nucleic acid extractor (Ruibao Biological Co., Ltd., Taiwan). Twenty-three common genotypes of α- and β-thalassemia mutations were identified in the Chinese population using the PCR reverse dot-blot hybridization technique (PCR-RDB) [18], including the deletional ($-\alpha^{3.7}$, $-\alpha^{4.2}$, $--^{SEA}$) and non-deletional α-thalassaemia mutations ($\alpha^{CS}\alpha$, $\alpha^{QS}\alpha$ and $\alpha^{Westmead}\alpha$), and 17 common β-thalassemia mutations: CD41-42(-TCTT), IVS-II-654(C > T), −28(A > G), CD71/72(+A), CD17(AAG > TAG), CD26(GAG > AAG), CD43(GAG > TAG), −29(A > G), CD31(-C), −32(C > A), IVS-I-1(G > T), CD27/28(+C), −30(T > C), CD14-15(+G), Cap + 40–43(−AAAC), initiation codon(ATG > AGG) and IVS-I-5(G > C) [7].

TGS and data analysis

The peripheral blood samples from 127 subjects who were suspected as rare globin gene variant carriers were sent to Berry Genomics laboratory for TGS, which was performed based on the PacBio Sequel II platform for α- and β-globin gene detection as described in our previous

report [16]. In brief, the purified DNA samples were quantified using the Qubit dsDNA BR assay kit (ThermoFisher Scientific). Then, optimized primers were used to generate specific amplicons that encapsulate known structural variation regions, and single nucleotide variation in the α- and β-globin gene. Double barcode adapters were ligated to the 5' and 3' ends after purification and end repair, and then Sequel Binding and Internal Ctrl Kit 3.0 (PacBio) was used to prepare SMRT bell libraries. Finally, third-generation sequencing was performed on the PacBio Sequel II System after loading primed DNA-polymerase complexes onto the SMRT cells [16].

By referring to alignment of the subreads, the consensus circular sequence was mapped to the GRCh38 reference and variants called (FreeBayes software, version 1.2.0). WhatsHap (version 0.18) software was used for Linkage analysis (in cis or trans) in the long read based phasing. Alignments of variant and wild-type molecules was manifested by Integrative Genomics Viewer [16].

Sanger sequencing

Sanger sequencing was performed to verify the globin gene variants detected by TGS. Specific primers according to the known DNA sequences were designed to perform PCR. All the designed primers were synthesized by Invitrogen (Shanghai) Trade Co., Ltd. The operation process of Sanger sequencing was performed following our previous study [16].

Results

The detection rate of α- and β-thalassaemia

In our study, 2,390 subjects carrying α- and β-thalassemia mutations were identified, including 1,615 cases with α-thalassemia variants, 719 cases harboring β-thalassaemia variants, and 56 cases with both α- and β-thalassaemia, reaching a total detection rate of 38.71% (2,390/6,174), and the α-thalassemia, β-thalassaemia and complex α- and β-thalassaemia detection rate were 26.16%, 11.65% and 0.91%, respectively. Among these, 2,349 subjects with common α- and β-thalassemia mutations were identified by conventional thalassemia gene testing, reaching a positive thalassemia detection rate of 38.05% (2,349/6,174). In addition, 40 cases with rare or novel thalassemia mutations were identified by TGS, and a case with $--^{SEA}/\alpha\alpha$ that missed the diagnosis by conventional PCR-RDB technology but later was identified by TGS.

Common and rare α-thalassemia variants

As delineated in Table 1, in the 1,615 subjects diagnosed with α-thalassemia, the $--^{SEA}/\alpha\alpha$ (70.84%) genotype was the most prevalent deletional mutation, followed by $-\alpha^{3.7}/$

Table 1 Distribution of α-thalassemia genotypes in Quanzhou region of Southeast China

Genotypes	Cases	Frequency	Class
$--^{SEA}/\alpha\alpha$	1144	70.84%	Common
$-\alpha^{3.7}/\alpha\alpha$	301	18.64%	Common
$-\alpha^{4.2}/\alpha\alpha$	59	3.65%	Common
$\alpha^{QS}\alpha/\alpha\alpha$	25	1.55%	Common
$\alpha^{CS}\alpha/\alpha\alpha$	21	1.30%	Common
$-\alpha^{3.7}/--^{SEA}$	19	1.18%	Common
$-\alpha^{3.7}/-\alpha^{3.7}$	10	0.62%	Common
$\alpha^{Westmead}\alpha/\alpha\alpha$	7	0.43%	Common
$-\alpha^{4.2}/--^{SEA}$	5	0.31%	Common
$-\alpha^{3.7}/-\alpha^{4.2}$	3	0.19%	Common
$\alpha^{CS}\alpha/--^{SEA}$	2	0.12%	Common
$\alpha^{QS}\alpha/--^{SEA}$	1	0.06%	Common
$\alpha^{CS}\alpha/-\alpha^{3.7}$	1	0.06%	Common
$\alpha^{CS}\alpha/\alpha^{CS}\alpha$	1	0.06%	Common
$--^{THAI}/\alpha\alpha$	5	0.31%	Rare
$\alpha^{IVS-II-55(T>G)}\alpha/\alpha\alpha$	2	0.12%	Rare
$--^{FIL}/\alpha\alpha$	2	0.12%	Rare
$--^{THAI}/-\alpha^{3.7}$	1	0.06%	Rare
$\alpha^{CD30}\alpha/\alpha\alpha$	1	0.06%	Rare
$\alpha^{CD15}\alpha/\alpha\alpha$	1	0.06%	Rare
$\alpha\alpha^{CD117/118}/\alpha\alpha$	1	0.06%	Rare
$\alpha^{IVS-II-119(-G)(+CTCGGCCC),\ IVS-II-55(T>G)}\alpha\ /--^{SEA}$	1	0.06%	Rare
$\alpha^{IVS-II-119(-G)(+CTCGGCCC),\ IVS-II-55(T>G)}\alpha\ /\alpha\alpha$	1	0.06%	Rare
$\alpha\alpha\alpha^{anti3.7}/\alpha\alpha^{*}$	1	0.06%	Rare
Total	1615	100.00%	/

*:In this case, Hb G-Honolulu was also compounded

αα (18.64%) and -$\alpha^{4.2}$/αα (3.65%). The most common non-deletional mutation was α^{QS}α/αα (1.55%), followed by α^{CS}α/αα, and $\alpha^{Westmead}$α/αα, with frequencies 1.30% and 0.43%, respectively. Besides, 29 cases that may cause Hb H disease were detected, among which, -$\alpha^{3.7}$/--^{SEA} was the most common (1.18%). A patient with homozygous α^{CS} was also classified as having Hb H disease, who was a female and exhibited a remarkably low level of Hb (86 g/L), slightly low level of MCV (81fL) and MCH (25.2pg), and a slightly increased level of serum iron (28.36 μmol/L).

In this study, including 10 genotypes of rare α-globin gene variants, were diagnosed by TGS. Notably, --^{THAI}/αα manifested a higher frequency. In addition, $\alpha^{CD30(-GAG)}$α was first reported in Fujian province South China. To the best of our knowledge, the complex genotype of $\alpha^{IVS-II-119(-G)(+CTCGGCCC)}$, $^{IVS-II-55(T>G)}$α/--^{SEA} and a case compound αα$\alpha^{anti3.7}$ with Hb G-Honolulu were first identified in the Chinese population (Figs. 1 and 2).

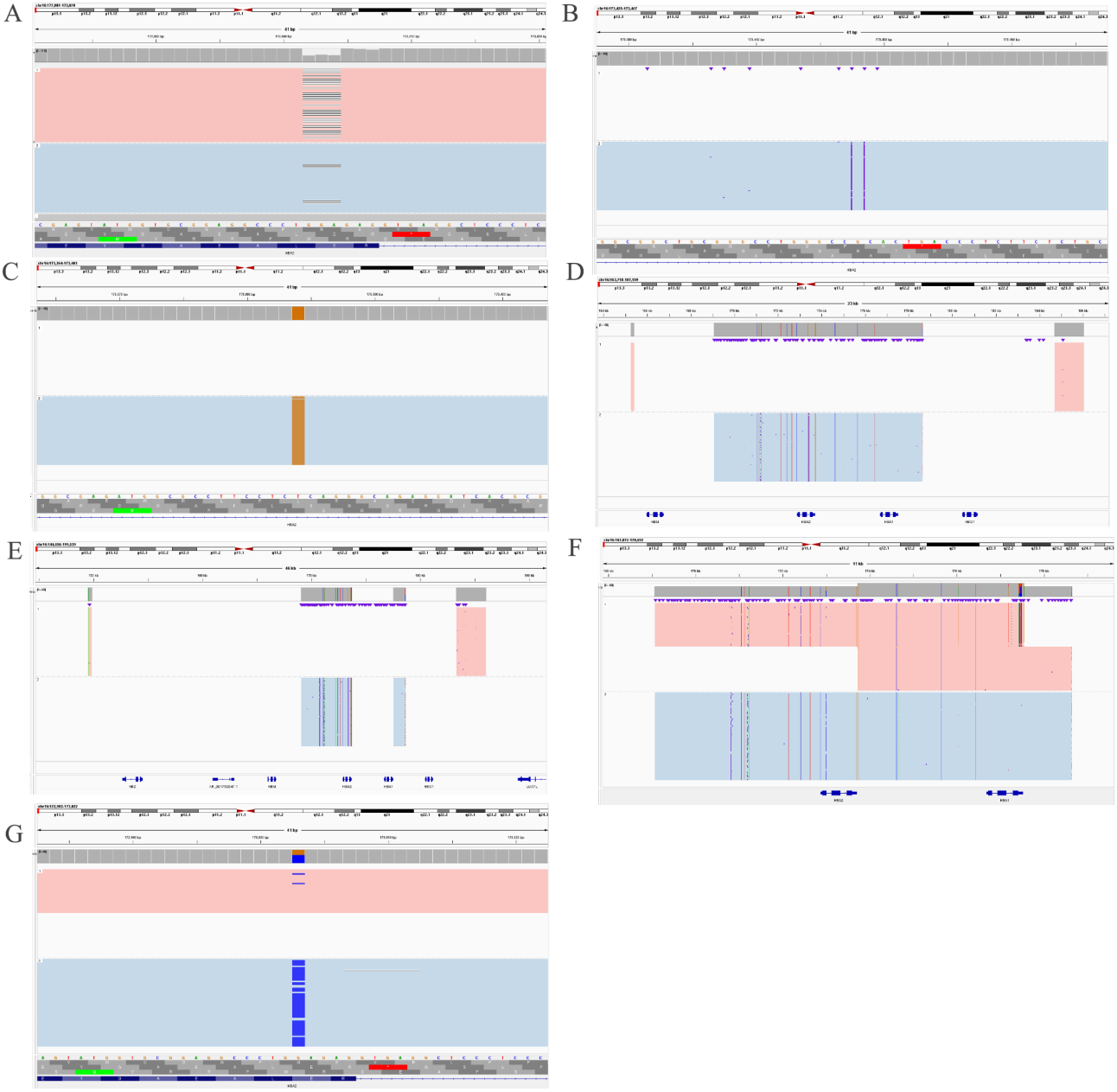

Fig. 1 Identification of rare α-globin gene variants using third-generation sequencing (TGS). **A**: A rare CD30(-GAG) (HBA2:c.91_93delGAG) identified in Fujian province. A case with IVS-II-119(-G;+CTCGGCCC) (HBA2:c.301-24delGinsCTCGGCCC) (**B**) associated with IVS-II-55(T-> G) (HBA2:c.300 + 55T > G) (**C**) in cis and also compounded with --^{SEA} (**D**) in trans was identified by TGS. E: --^{THAI}/-$\alpha^{3.7}$. A case with αα$\alpha^{anti3.7}$ (F) associated with Hb G-Honolulu (HBA2:c.91G > C) (G) in trans was detected by TGS

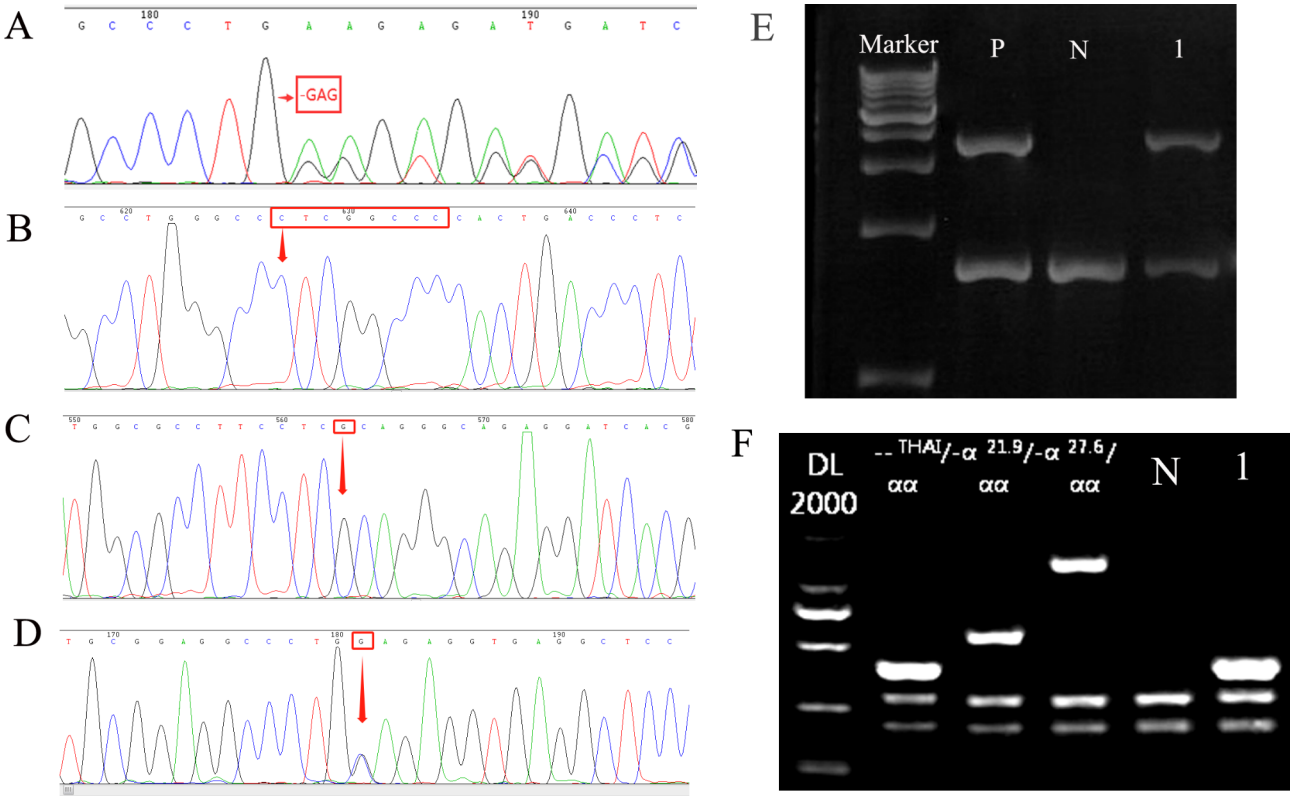

Fig. 2 The rare α-globin gene variant was further confirmed by third-generation Sanger sequencing and gap-PCR. **A**: CD30(-GAG) (HBA2:c.91_93delGAG). **B**:IVS-II-119(-G;+CTCGGCCC) (HBA2:c.301-24delGinsCTCGGCCC). **C**:IVS-II-55(T-> G) (HBA2:c.300 + 55T > G). **D**: Hb G-Honolulu (HBA2:c.91G > C). **E**: Electrophoresis result of $\alpha\alpha^{anti3.7}$. P: Positive control; N: negative control; 1: $\alpha\alpha^{anti3.7}$ carrier. **F**: Electrophoresis result of $--^{THAI}$. N: negative control; 1: $--^{THAI}$ thalassemia carrier

In addition, the rare variants of $--^{FIL}/\alpha\alpha$, $\alpha\alpha^{CD15}/\alpha\alpha$, and $\alpha\alpha^{CD117/118}/\alpha\alpha$ were identified again in Quanzhou region of Southeast China.

Common and rare β-thalassemia variants

A total of 719 cases were diagnosed as β-thalassemia, among which $\beta^{IVS-II-654}/\beta^N$ (36.02%) and $\beta^{CD41-42}/\beta^N$ (26.01%) were the most common β-thalassemia mutations, followed by β^{CD17}/β^N (19.61%), β^{CD26}/β^N (5.70%), and β^{-28}/β^N (3.48%), which accounted for about 90.82% of β-thalassemia variants (Table 2). In this study, we also identified 7 cases of β-thalassemia intermedia or major, with $\beta^{IVS-II-654}/\beta^{IVS-II-654}$ being the most common genotype.

In addition, 24 cases including 12 genotypes of rare β-thalassemia variants were identified, of which SEA-HPFH was the most frequent (Table 2). Among them, the rare $\beta^{CD54-58}/\beta^N$ and β^{CD53}/β^N were diagnosed again in Quanzhou region using TGS, while, the Chinese $^G\gamma^+(^A\gamma\delta\beta)^0$ had not been described in Quanzhou region before (Figs. 3 and 4). The Hb Lepore-Boston-Washington identified in this study was the first reported case in Fujian province of Southeast China. Moreover, the $\beta^{CD15(TGG>TAG)}$, $\beta^{IVS-II-761}$, β^0-Filipino(~45 kb deletion), and Hb Lepore-Quanzhou were first reported in the Chinese population by us.

Compound α- and β-thalassemia variants

Of the 6,174 enrolled subjects, 56 cases (0.91%) were diagnosed as compound α- and β-thalassemia, among which $--^{SEA}/\alpha\alpha/\beta^{IVS-II-654}/\beta^N$, $-\alpha^{3.7}/\alpha\alpha/\beta^{CD41-42}/\beta^N$ and $-\alpha^{3.7}/\alpha\alpha/\beta^{CD17}/\beta^N$ were the three most prevalent genotypes (Table 3).

Hb variant genotypes

TGS was also carried out to confirm the exact hemoglobin variants in subjects with abnormal Hb band(s). Thirty-five cases of Hb variants were detected, including 11 cases of Hb Q-Thailand [CD74(GAC>CAC)], 2 cases of Hb G-Honolulu [CD30(GAG>CAG)], one case of Hb Beijing [CD16(AAG>AAT), and one case of Hb Jilin [CD139(AAA>CAA)], all of which were induced by α-globin gene mutation. In addition, 17 cases of Hb New York [CD113(GTG>GAG)], and three cases of Hb J-Bangkok [CD56(GGC>GAC)] that caused by β-globin gene mutation were identified.

Table 2 Distribution of β-thalassemia genotypes in Quanzhou region of Southeast China

Types	Genotypes	Cases	Frequency	Class
Heterozygote	$\beta^{IVS-II-654}/\beta^{N}$	259	36.02%	Common
	$\beta^{CD41-42}/\beta^{N}$	187	26.01%	Common
	β^{CD17}/β^{N}	141	19.61%	Common
	β^{CD26}/β^{N}	41	5.70%	Common
	β^{-28}/β^{N}	25	3.48%	Common
	$\beta^{CD27/28}/\beta^{N}$	15	2.09%	Common
	$\beta^{CD71-72}/\beta^{N}$	5	0.70%	Common
	β^{CD43}/β^{N}	5	0.70%	Common
	$\beta^{CAP+40-43}\beta^{N}$	3	0.42%	Common
	β^{Int}/β^{N}	3	0.42%	Common
	$\beta^{IVS-I-1}/\beta^{N}$	2	0.28%	Common
	$\beta^{CD14-15}/\beta^{N}$	2	0.28%	Common
	SEA-HPFH	5	0.70%	Rare
	Hb Lepore-Boston-Washington	4	0.56%	Rare
	$\beta^{IVS-II-806}/\beta^{N}$	3	0.42%	Rare
	Chinese $^{G}\gamma^{+}(^{A}\gamma\delta\beta)^{0}$	2	0.28%	Rare
	β^{CD53}/β^{N}	2	0.28%	Rare
	$\beta^{IVS-II-672}/\beta^{N}$	2	0.28%	Rare
	β^{CD15}/β^{N}	1	0.14%	Rare
	$\beta^{CD54-58}/\beta^{N}$	1	0.14%	Rare
	Filipino(~45 kb deletion)	1	0.14%	Rare
	$\beta^{IVS-II-761}/\beta^{N}$	1	0.14%	Rare
	$\beta^{IVS-II-180}/\beta^{N}$	1	0.14%	Rare
	Hb Lepore-Quanzhou	1	0.14%	Novel
Compound heterozygote	$\beta^{IVS-II-654}/\beta^{IVS-II-654}$	2	0.28%	Common
	$\beta^{CD41-42}/\beta^{CD41-42}$	1	0.14%	Common
	β^{CD26}/β^{-28}	1	0.14%	Common
	$\beta^{IVS-II-654M}/\beta^{-28}$	1	0.14%	Common
	$\beta^{CD41-42}/\beta^{CD17}$	1	0.14%	Common
	$\beta^{CD41-42}/\beta^{-28}$	1	0.14%	Common
	Total	719	100.00%	/

Discussion

Of the 6,174 subjects with suspected thalassemia traits, 1,615 (26.16%) were identified as harboring α-thalassemia variants, 719 (11.16%) as harboring β-thalassemia variants, and 56 (0.91%) as harboring both α- and β-thalassemia, with an overall detection rate of 38.71% (2,390/6,174)]. Notably, α-thalassemia manifests a higher incidence than β-thalassemia, and $--^{SEA}/\alpha\alpha$, $-\alpha^{3.7}/\alpha\alpha$ and $-\alpha^{4.2}/\alpha\alpha$ were the most common α-thalassemia genotypes, which is consistent with the neighboring provinces of China, and also consistent with the other previous studies in Quanzhou region [8, 18–21].

In this study, we identified 1,615 subjects with α-thalassemia mutations, among them, the most common non-deletional mutation was $\alpha^{QS}\alpha/\alpha\alpha$, followed by $\alpha^{CS}\alpha/\alpha\alpha$, and $\alpha^{Westmead}\alpha/\alpha\alpha$. However, α^{CS} was reported to be the most prevalent non-deletional α-thalassemia in Southeast Asia, and $\alpha^{Westmead}$ was displayed a more higher incidence in Guangxi province of China [20]. Patients with compounded heterozygous or homozygous of α^{QS} and α^{CS} are likely to develop moderate anemia similar to Hb H disease [22]. Our present study identified a case of homozygous α^{CS} exhibiting remarkable moderate anemia similar to Hb H disease, which is consistent with a previous study [22]. As such, couples both carrying α^{CS} should be recommended for performing prenatal thalassemia gene testing after pregnant. Of the 16 cases identified by TGS, 10 belong to the rare α-thalassemia genotype. The $--^{THAI}/\alpha\alpha$ manifested a higher frequency in the rare α-thalassemia genotype, which is also consistent with the previous studies [7, 8] suggesting that TGS could be recommended as a routine method for detection of $--^{THAI}$ in Quanzhou region. The $\alpha^{CD30(-GAG)}$ is a rare *HBA2* globin gene variant with few reports available, which could be used as a predictor of α^{+} or α^{0}-thalassemia according to IthaGenes database. Previous studies reported that patients with $\alpha^{CD30(-GAG)}$ combined with deletional α^{0}-thalassemia were likely to develop mild to severe anemia [23, 24]. In the present study, we described the $\alpha^{CD30(-GAG)}$ again in the Chinese population, and for the first time in a cohort in Fujian province. In addition, our finding first displayed a complex genotype of

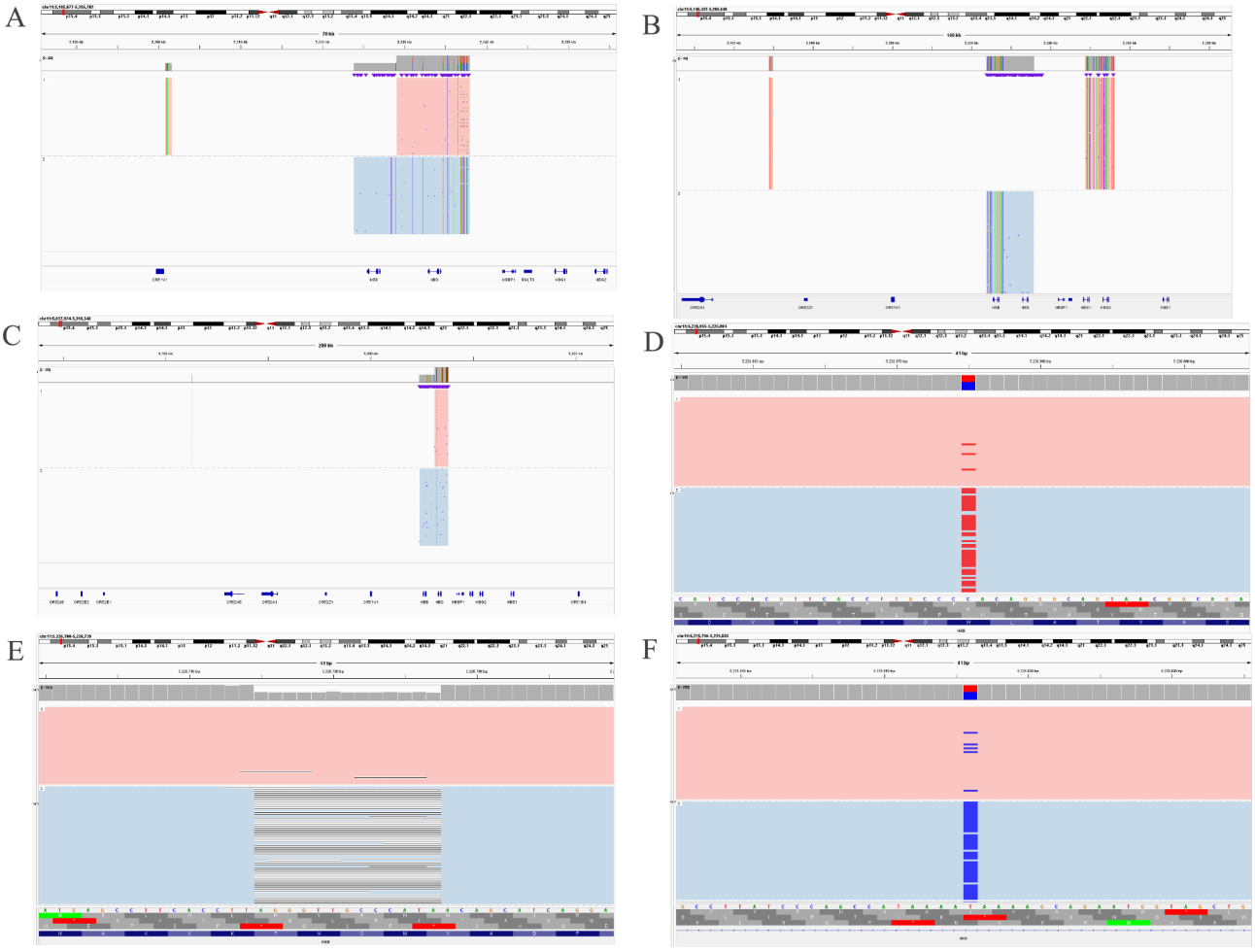

Fig. 3 Identification of rare β-globin gene variants using third-generation sequencing (TGS). (**A**, **B**): Identification of the rare δβ-thalassemia of SEA-HPFH and Chinese $^{G}\gamma^{+}(^{A}\gamma\delta\beta)^{0}$ by TGS. (**C**, **D**): Identification of the rare Filipino (~45 kb deletion) and rare CD15(TGG > TAG) (HBB: c.47G > A) by TGS. **E**: Detection of the rare CD54-58 (-TATGGGCAACCCT) mutation in *HBB* gene by TGS. **F**: Identification of the rare variant of IVS-II-761 (HBB: c.316–90 A > G) by TGS.

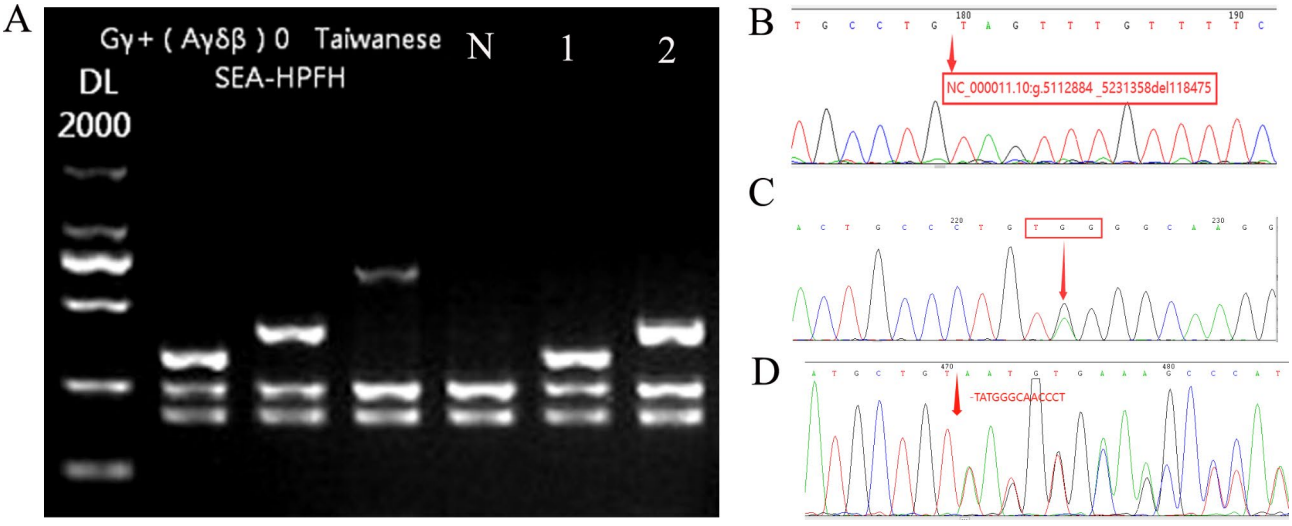

Fig. 4 The rare β-globin gene variant was further confirmed by third-generation Sanger sequencing technology (TGS) and gap-PCR. **A**: Electrophoresis result of SEA-HPFH and Chinese $^{G}\gamma^{+}(^{A}\gamma\delta\beta)^{0}$. N: negative control; 1: SEA-HPFH; 2: Chinese $^{G}\gamma^{+}(^{A}\gamma\delta\beta)^{0}$. **B**: The Filipino (~45 kb deletion) was further confirmed by Sanger sequencing (NC_000011.10:g.5112884 _5231358del118475). **C**: CD15(TGG > TAG) (HBB: c.47G > A). **D**:CD54-58(-TATGGGCAACCCT)

Table 3 Distribution of compound α and β-thalassemia in Quanzhou region of Southeast China

Genotypes	$--^{SEA}/\alpha\alpha$	$-\alpha^{3.7}/\alpha\alpha$	$-\alpha^{4.2}/\alpha\alpha$	$\alpha\alpha^{WS}/\alpha\alpha$	$\alpha\alpha^{CS}/\alpha\alpha$	$\alpha\alpha^{QS}/\alpha\alpha$	$-\alpha^{3.7}/--^{SEA}$
$\beta^{IVS-II-654}/\beta^{N}$	12	6	0	0	0	1	0
$\beta^{CD41-42}/\beta^{N}$	5	8	1	5	1	0	1
β^{CD17}/β^{N}	1	7	0	0	1	0	0
β^{CD26}/β^{N}	1	3	0	0	0	0	0
$\beta^{IVS-I-1}/\beta^{N}$	0	0	0	0	1	0	0
$\beta^{CD71-72}/\beta^{N}$	1	0	0	0	0	0	0
$\beta^{CAP+40-43}\beta^{N}$	1	0	0	0	0	0	0

$\alpha^{IVS-II-119(-G)(+CTCGGCCC),\ IVS-II-55(T>G)}\alpha/--^{SEA}$ in a Chinese pediatric patient with normal hemoglobin value (114 g/L) and a slightly low level of mean corpuscular hemoglobin (MCH, 26.8pg). In addition, we identified a patient with IVS-II-119(-G)(+CTCGGCCC) and IVS-II-55(T>G) in *HBA2* located in cis by TGS, who exhibited normal hematological screening results, suggesting that both variants in α2 may be benign variants. Moreover, a rare case of compound $\alpha\alpha\alpha^{anti3.7}$ with Hb G-Honolulu was first identified in Chinese population with normal hematological screening results.

In the present study, we identified 719 cases carrying β-thalassemia variants, and the most frequent mutation was IVS-II-654(C>T) and CD41-42(-TCTT), which is similar with the finding in other regions of South China [5, 19, 21]. However, in Yunnan province, CD26(G>A) was reported to be the most frequency β-thalassemia mutation [25]. In the present study, we also identified 7 cases with β-thalassemia intermedia or major were also identified, with a detection rate of 0.11%(7/6,174), indicating that a more optimized and comprehensive thalassemia prevention and control method should be carried out. Moreover, we also identified 12 genotypes of rare β-thalassemia variants, including $\beta^{CD15(TGG>TAG)}$ which is known as rare variant condition that may lead to β^0-thalassemia and has been frequently reported in Iran, Pakistan and India [26–28]. Interestingly, it is for the first time that we described the rare $\beta^{CD15(TGG>TAG)}$ variant in the Chinese population, and it is also for the first time that we displayed an extremely rare variant of IVS-II–761 (HBB: c.316–90 A>G) in the Chinese population, knowing that $\beta^{CD15(TGG>TAG)}$ may lead to β^+ or β^0-thalassemia, which was first discovered in the United States of America [29]. Interestingly, the patient with IVS-II-761 (HBB: c.316–90 A>G) exhibited low levels of MCV (68fL), MCH (19.7pg), and Hb A2 (1.5%), which may result in β-thalassemia. Furthermore, this study also firstly present a rare deletional Filipino (~45 kb deletion) in the Chinese population, which was mentioned in our previous study [30]. This finding may further confirm the mobility and complexity of the population in Quanzhou region. The Filipino (~45 kb deletion) was first reported in a Filipino family [31]. A previous study conducted by Tan et al. [32] reported that the Filipino (~45 kb deletion) caused 73.3% of β-thalassemia in the Orang Asli in Sabah and Sarawak, which are indigenous/aboriginal peoples in East Malaysia. It was found in our study that the variants of IVS-II–806 (HBB: c.316–45G>C) and IVS-II–672 (HBB: c.316–179 A>C) exhibited variable hematological profiles, but they would lead to β-thalassemia or not needs more investigation.

Notably, we identified 11 cases of δβ-thalassemia in this study, among which SEA-HPFH was the most frequently detected mutation, which may lead to δβ-thalassemia. This finding is consistent with our previous study [8]. It was also for the first time that we identified another rare δβ-thalassemia of Chinese $^{G}\gamma^{+}(^{A}\gamma\delta\beta)^{0}$ in Quanzhou region, although it has been reported in Chinese populations of other regions [33]. Also, we identified a novel 7.414 kb deletion NG_000007.3:g.63511_70924del that partially covers *HBB* and *HBD* globin genes and may cause δβ-thalassemia in Quanzhou region, and we named it as Hb Lepore-Quanzhou [34]. In addition, we first described four cases of Hb Lepore variant (NG_000007.3:g.63633_71046 del) with one bp different with Hb Lepore-Boston-Washington, which has been reported by our previous study [35]. More attention should be paid to the screening and diagnosis of the rare δβ-thalassemia, knowing that homozygous δβ-thalassemia or co-occurring heterozygosity with β-thalassemia may lead to β-thalassemia intermedia or major.

In this study, we identified 35 cases of Hb variants, including Hb New York and Hb Q-Thailand as the most common two hemoglobin variants. Notably, both the rare Hb Jilin [α139(HC1)Lys>Gln (AAA>CAA); HBA2:c.418 A>C] and Hb Beijing [α16(A14)Lys>Asn (AAG>AAT); HBA2:c.51G>T] detected in the present study were first reported in Fujian province, though they were mentioned in our previous study [36]. The Hb Jilin is an extremely rare α-chain variant, which was only occasionally reported in the literature [37]. In addition, Hb Beijing is also a rare condition, with few reports available in the literature [38, 39].

By using the TGS technique, we identified a diverse range of α- and β-globin gene variants in the present study, among which $--^{SEA}/\alpha\alpha$ was diagnosed as a common genotype by TGS, which was missed in the conventional

thalassemia gene testing. This study further has strengthened the application value of TGS, suggesting that TGS may prove to be a powerful tool for the diagnosis of common and rare α- and β-globin gene variants.

In this study, we succeeded in identifying rare α- and β-globin gene variants by TGS in Quanzhou region of Southeast China, including a variety of thalassemia and Hb variants such as $\beta^{CD15(TGG>TAG)}$, $\beta^{IVS-II-761}$, β^{0}-Filipino (~45 kb deletion), Hb Lepore-Quanzhou, Hb Jilin and Hb Beijing. These findings expanded the prevalence and spectrum of thalassemia in Fujian province of Southeast China, suggesting that the detection of thalassemia genes based on TGS could be used as a first-line tool for globin gene variant detection.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12920-024-02014-2.

Supplementary Material 1

Acknowledgements
We wish to express our appreciation to Quanzhou Science and Technology, Fujian Provincial Key Laboratory of Prenatal Diagnosis and Birth Defects and Huaqiao University Joint of Hospital and University projects for funding this work. We also express our appreciation to the patients who participated in this study.

Author contributions
JZ and JW wrote the article; NH performed participants recruitment and analyzed the data; YZ performed specific gap-PCR amplification and DNA sequencing; JW, LX and JZ revised and polished the paper. All authors have approved the final article.

Funding
This research was supported by Quanzhou City Science and Technology Project (no.2020N049s); Huaqiao University Joint of Hospital and University Innovation Project (no.2021YX005) and Fujian Provincial Key Laboratory of Prenatal Diagnosis and Birth Defects open project (no.cqzdsys-2022-01).

Data availability
The datasets used and/or analyzed during the current study are available from the corresponding author by reasonable request.

Declarations

Ethics approval and consent for publication
This study was approved by the ethics committee of the Quanzhou Women's and Children's Hospital (2021No.61). We ensured that all patients participating in this study signed written informed consent for publication of their genetic data and other relevant information.

Consent to participate
All patients signed written informed consent to participate in this study and agreed for their relevant data and information to be used for scientific research.

Conflicting interests
The authors declare that they have no conflict of interests.

Received: 8 February 2024 / Accepted: 17 September 2024
Published online: 01 October 2024

References

1. Weatherall DJ. Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. Nat Rev Genet. 2001;2(4):245–55.
2. Kattamis A, Kwiatkowski JL, Aydinok Y, Thalassaemia. Lancet. 2022;399(10343):2310–24.
3. Cao A, Kan YW. The prevention of Thalassemia. Cold Spring Harb Perspect Med. 2013;3(2):a011775.
4. Li CK. New trend in the epidemiology of Thalassaemia. Best Pract Res Clin Obstet Gynaecol. 2017;39:16–26.
5. Xu LP, Huang HL, Wang Y, Zheng L, Wang LS, Xu JB, Huang XX, Lin Y. [Molecular epidemiological analysis of α- and β-thalassemia in Fujian Province]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi. 2013;30(4):403–6.
6. Huang H, Xu L, Chen M, et al. Molecular characterization of Thalassemia and hemoglobinopathy in Southeastern China. Sci Rep. 2019;9(1):3493.
7. Zhuang J, Jiang Y, Wang Y, et al. Molecular analysis of α-thalassemia and β-thalassemia in Quanzhou Region Southeast China. J Clin Pathol. 2020;73(5):278–82.
8. Zhuang J, Zhang N, Wang Y, et al. Molecular characterization analysis of Thalassemia and Hemoglobinopathy in Quanzhou, Southeast China: a large-scale retrospective study. Front Genet. 2021;12:727233.
9. Zhuang J, Tian J, Wei J, Zheng Y, Zhuang Q, Wang Y, Xie Q, Zeng S, Wang G, Pan Y, Jiang Y. Molecular analysis of a large novel deletion causing α+-thalassemia. BMC Med Genet. 2019;20(1):74.
10. Chen YB, Jiang YC, Chen ZX, Zhang ZS. [Epidemiological Survey of Thalassemia in women of childbearing age in Quanzhou Area of Fujian Province in China]. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2018;26(5):1453–8.
11. Luo H, Huang T, Lu Q, et al. Molecular prevalence of HBB-associated hemoglobinopathy among reproductive-age adults and the prenatal diagnosis in Jiangxi Province, southern central China. Front Genet. 2022;13:992073.
12. Liu Q, Chen Q, Zhang Z, et al. Identification of rare thalassemia variants using third-generation sequencing. Front Genet. 2023;13:1076035.
13. Cretu Stancu M, Van Roosmalen MJ, Renkens I, et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing[J]. Nat Commun. 2017;8(1):1326.
14. Xu L, Mao A, Liu H, et al. Long-molecule sequencing: a New Approach for Identification of clinically significant DNA variants in α-Thalassemia and β-Thalassemia carriers. J Mol Diagn. 2020;22(8):1087–95.
15. Long J, Sun L, Gong F, et al. Third-generation sequencing: a novel tool detects complex variants in the α-thalassemia gene. Gene. 2022;822:146332.
16. Zhuang J, Chen C, Fu W, et al. Third-generation sequencing as a New Comprehensive Technology for identifying rare α- and β-Globin gene variants in thalassemia alleles in the Chinese Population. Arch Pathol Lab Med. 2023;147(2):208–14.
17. Peng C, Zhang H, Ren J, et al. Analysis of rare thalassemia genetic variants based on third-generation sequencing. Sci Rep. 2022;12(1):9907.
18. Liu Q, Jia ZJ, Xi H, et al. [Analysis on the genotype of 5018 cases of Thalassemia in Hunan Area]. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2019;27(6):1938–42.
19. Xu XM, Zhou YQ, Luo GX, et al. The prevalence and spectrum of alpha and beta thalassaemia in Guangdong Province: implications for the future health burden and population screening. J Clin Pathol. 2004;57(5):517–22.
20. Zheng CG, Liu M, Du J, Chen K, Yang Y, Yang Z. Molecular spectrum of α- and β-globin gene mutations detected in the population of Guangxi Zhuang Autonomous Region, people's Republic of China. Hemoglobin. 2011;35(1):28–39.
21. Yang T, Luo X, Liu Y, et al. Next-generation sequencing analysis of the molecular spectrum of Thalassemia in Southern Jiangxi, China. Hum Genomics. 2023;17(1):77.
22. Writing Group For Practice Guidelines For Diagnosis And Treatment Of Genetic Diseases Medical Genetics Branch Of Chinese Medical Association, Shang X, Zhang X, Yang F, Xu X. Zhonghua Yi Xue Yi Chuan Xue Za Zhi. 2020;37(3):235–42.
23. Chan V, Chan VW, Tang M, Lau K, Todd D, Chan TK. Molecular defects in hb H hydrops fetalis. Br J Haematol. 1997;96(2):224–8.
24. Ren ZM, Li WJ, Xing ZH, et al. Detecting rare thalassemia in children with anemia using third-generation sequencing. Hematology. 2023;28(1):2241226.
25. Zhang J, Zhu BS, He J, et al. The spectrum of α- and β-thalassemia mutations in Yunnan Province of Southwestern China. Hemoglobin. 2012;36(5):464–73.
26. Abbasali FH, Mahmoud KS, Hengameh N, et al. Rare and new mutations of B-Globin in Azari Population of Iran, a considerable diversity. Balkan J Med Genet. 2023;25(2):51–62.

27. Mahmood Baig S, Sabih D, Rahim MK, et al. β-Thalassemia in Pakistan: a pilot program on prenatal diagnosis in Multan. J Pediatr Hematol Oncol. 2012;34(2):90–2.
28. Colah R, Gorakshakar A, Nadkarni A, et al. Regional heterogeneity of beta-thalassemia mutations in the multi ethnic Indian population. Blood Cells Mol Dis. 2009;42(3):241–6.
29. Chan OT, Westover KD, Dietz L, Zehnder JL, Schrijver I. Comprehensive and efficient HBB mutation analysis for detection of beta-hemoglobinopathies in a pan-ethnic population. Am J Clin Pathol. 2010;133(5):700–7.
30. Chen M, Lv A, Zhang S et al. First Report of Filipino β0-Thalassemia/β-Thalassemia in a Chinese Family. *Hemoglobin*. Published online January 9, 2024.
31. Motum PI, Kearney A, Hamilton TJ, Trent RJ. Filipino beta zero thalassaemia: a high hb A2 beta zero thalassaemia resulting from a large deletion of the 5′ beta globin gene region. J Med Genet. 1993;30(3):240–4.
32. Tan JA, George E, Tan KL, et al. Molecular defects in the beta-globin gene identified in different ethnic groups/populations during prenatal diagnosis for beta-thalassemia: a Malaysian experience. Clin Exp Med. 2004;4(3):142–7.
33. Ju AP, Li N, Lin K, Huang HH, Liu SX, Jiang F. [Molecular epidemiological characteristics and Differential diagnosis of common δβ-Thalassemia/HPFH]. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2022;30(4):1182–7.
34. Jianlong Zhuang Y, Zheng Y, Jiang J, Wang S, Zeng N, Liu. Long-Read Sequencing Identified a Large Novel δ/β-Globin Gene Deletion in a Chinese Family, Human Mutation, vol. 2023, Article ID 2766625, 8 pages, 2023. https://doi.org/10.1155/2023/2766625
35. Zhuang J, Zhang N, Zheng Y, et al. Molecular characterization of similar Hb Lepore Boston-Washington in four Chinese families using third generation sequencing. Sci Rep. 2024;14(1):9966.
36. Zhuang J, Jiang Y, Chen Y, Mao A, Chen J, Chen C. Third-generation sequencing identified two rare α-chain variants leading to hemoglobin variants in Chinese population. Mol Genet Genomic Med. 2024;12(1):e2365.
37. Xu A, Li J, Chen W, et al. Identification of a novel hemoglobin variant Hb Jilin [α139(HC1)lys > gln; HBA2:C.418 A > C] in a Chinese family. Int J Lab Hematol. 2019;41(3):e73–5.
38. Liang CC, Chen S, Yang K, et al. Hemoglobin Beijing [alpha 16 (A14) Lys replaced by Asn]: a new fast-moving hemoglobin variant. Hemoglobin. 1982;6(6):629–33.
39. Lin M, Han ZJ, Wang Q, et al. Molecular epidemiological survey of hemoglobinopathies in the Wuxi region of Jiangsu Province, eastern China. Hemoglobin. 2013;37(5):454–66.

Publisher's note