

Research Article

Large-Scale Screening of Thalassemia in Ji'an, P.R. China

Yu Qiu^{1#}, Liangwei Mao^{2,3#}, Shiping Chen^{4#}, Hao Li², Hairong Wang², Liping Guan^{5,6}, Jin Huang², Xuan Wu², Yu Liu², Jie Xiao^{7*}, Yuan Fang^{8*}

¹Obstetrics Clinic, Department of Obstetrics and Gynecology, Ji'an Maternal and Child Health Hospital, Ji'an, China

²BGI-Wuhan Clinical Laboratory, BGI-Shenzhen, Wuhan, China

³The State Key Laboratory of Biocatalysis and Enzyme Engineering, College of Life Sciences, Hubei University, Wuhan, China

⁴Clinical laboratory of BGI Health, BGI-Shenzhen, Shenzhen, China

⁵BGI Genomics, BGI-Shenzhen, Shenzhen, China

⁶Department of Biology, University of Copenhagen, Copenhagen, Denmark

⁷Department of Genetics, Ji'an Maternal and Child Health Hospital, Ji'an, China

⁸Obstetrics Clinic, Department of Obstetrics and Gynecology, Jiujiang Maternal and Child Health Hospital, Jiujiang, China

#Equally contributed authors

***Corresponding author:** Jie Xiao, Department of Genetics, Ji'an Maternal and Child Health Hospital, Ji'an 343000, China

Yuan Fang, Obstetrics Clinic, Department of Obstetrics and Gynecology, Jiujiang Maternal and Child Health Hospital, Jiujiang 33200, China

Received: 24 February 2022; **Accepted:** 04 March 2022; **Published:** 04 April 2022

Citation: Yu Qiu, Liangwei Mao, Shiping Chen, Hao Li, Hairong Wang, Liping Guan, Jin Huang, Xuan Wu, Yu Liu, Jie Xiao, Yuan Fang. Large-Scale Screening of Thalassemia in Ji'an, P.R. China. *Journal of Women's Health and Development* 5 (2022): 109-124.

Abstract

Background: To evaluate the prevalence of alpha- and beta-thalassemia in Ji'an City, Jiangxi Province, 28,941 residents in the region were genetically screened to identify various thalassemia genotypes.

Methods: High-throughput amplicon sequencing was used to screen 301 thalassemia alleles in 28,941 people in the region. Pregnant women were the focus of this screening, and if a pregnant woman had mutations in a thalassemia-causing gene, her spouse was also genetically screened.

Results: Of the participants, 2,380 people were carriers of thalassemia, with at least one thalassemia allele, including 1,694 alpha-thalassemia carriers, 638 beta-thalassemia carriers and 48 composite alpha and beta-thalassemia carriers. In Ji'an, the total carrying rate of thalassemia was 8.22%, while alpha- and beta-thalassemia were 5.85% and 2.20%, respectively. In addition, the first measured carrier rate of composite alpha- and beta-thalassemia in Ji'an was 0.17%. According to the geographical distribution of the 1,694 alpha -thalassemia carriers, the city with the highest carrier rate was Suichuan, followed by Wan'an and Taihe. According to the geographical distribution of the 638 beta-thalassemia carriers, the top three cities with high carrier rates were Suichuan, Wan'an and Xiajiang, sequentially.

Conclusions: This research demonstrates the critical nature of large-scale population screening and that comprehensive molecular epidemiology data are necessary for the proper prevention and treatment of thalassemia. The updated epidemiological data from this study may help the local government to focus on the severity of this disease and find some methods for

resource allocation that is successful under limited resource conditions.

Keywords: Thalassemia; Next-Generation Sequencing; Epidemiology; Prevalence

Abbreviations: PCR: Polymerase Chain Reaction; NGS: Next-generation Sequencing; MCV: Mean Corpuscular Volume; MCH: Mean Corpuscular Hemoglobin; MCHC: Mean Corpuscular Hemoglobin Concentration; Hb: Hemoglobin; MLPA: Multiplex Ligation-dependent Probe Amplification; CGH: Comparative Genome Hybridization

1. Background

Hemoglobinopathy is a term that refers to any hereditary blood disease caused by an abnormal hemoglobin molecular structure or an abnormal rate of globin peptide synthesis (thalassemia) [1]. According to the affected globin gene, thalassemia is classified as alpha-thalassemia and beta-thalassemia. Clinically, there are many symptoms of thalassemia, from asymptomatic to lethal. Patients with such symptoms might be classified as having thalassemia minor, intermedia and major according to the clinical severity. Thalassemia is a condition that affects patients with the later two categories. The severity of this condition is primarily determined by the degree to which the α :non- α chain is imbalanced [2]. This monogenic disease is one of the most prevalent and harmful in the world, affecting the greatest number of people [3, 4]. It is estimated that 1-5% of the global population are carriers of thalassemia mutations [5]. Alpha- and beta-thalassemia are a growing health burden in a number of Asian nations [6]. Plans aimed at preventing and effectively managing these diseases have the potential to considerably improve health

indexes in many developing countries [7]. Accurate population frequency data are necessary for development of these programs.

In China, the high-incidence areas of thalassemia include Guangdong, Guangxi, Hainan, Yunnan, Guizhou, Sichuan, Chongqing, Hunan, Jiangxi and other southern regions [8, 9]. At present, apart from hematopoietic stem cell transplantation, there is no effective treatment for thalassemia major. Numerous individuals require lifelong blood transfusions and chelation therapy, and the only effective preventive methods for thalassemia major and thalassemia intermedia are carrier screening and prenatal diagnosis [10]. For the correct prevention and treatment of thalassemia illness, comprehensive molecular epidemiology data are required. Thus far, significant volumes of genomic data have been generated by next-generation sequencing (NGS) in order to characterize people's genetic makeup and assess potential health concerns. The scope of variation in thalassemia appears to be more than previously described [11], and NGS is an excellent approach for screening for thalassemia-related variation.

Although a large-scale investigation of thalassemia has been conducted in recent years in various regions of China [4, 8, 12], the epidemiologic characteristics of thalassemia in people from Jiangxi Province, one of the high-prevalence areas, and particularly in those from Ji'an City, remain unknown [13]. Ji'an is located in Jiangxi Province's central region, covering an area of 25,300 square kilometers and housing a population of 4,956,600 people. It covers two districts, one city, and ten counties, including Jizhou District, Qingyuan District, Jinggangshan City, Ji'an County, Taihe County, and Wan'an County, Suichuan County, Yongxin County, Yongfeng County, and Jishui

County, Xiajiang County, Anfu County, and Xingan County. The goal of this study was to determine the prevalence of thalassemia mutations in Ji'an City, Jiangxi Province, China, and to learn more about thalassemia's epidemiological characteristics.

2. Methods

2.1 Participants

In this study, our strategy was to screen pregnant women in the Ji'an region with high-throughput amplicon sequencing for the 301 thalassemia alleles (Table S1). This is a government-funded public health service program, and all pregnant women in the jurisdiction chose to participate in the free testing. If a pregnant woman is found as a thalassemia carrier, her spouse must also consent to a genetic test for thalassemia. The screening procedure in detail is depicted in Figure 1. This research was approved by the Ethics Committee of First People's Hospital of Ji'an, and all the participants signed written informed consent. Parents/legal guardians of participants under the age of 16 years provided written consent.

2.2 NGS library preparation

As previously described, a series of primers based on the features of three genes associated with alpha- and beta-thalassemia, *HBA1*, *HBA2* and *HBB*, were created to facilitate gene amplification [4]. The primers are described in WO/2014/023076, WO/2014/023167, and CN102952877. The nondeletion alpha-thalassemia detection region is situated on chr16:226667-227546 and chr16:222863-223733 in *HBA1* and *HBA2*, respectively; the nondeletion beta-thalassemia variation is located on chr11:5247713-5248438 and chr11:5246655-5247209 in the *HBB* gene. Briefly, multiplex PCR techniques were used to amplify and enrich the *HBA1*, *HBA2*, and *HBB* genes in a sample while also introducing a tag sequence for

sample identification via fragmentation and ligation. The PCR products of a large number of samples (≤ 96 samples) were pooled into a library. Following library preparation, each library's DNA sequences were added with a linker sequence for sequencing and library identification. Libraries were pooled at equal molarity. The pooled libraries were sequenced with 2×100 paired end reads on a MGISEQ-2000.

2.3 Bioinformatics analysis

We employed an in-house pipeline that includes numerous in-house programs and open-source softwares to perform quality controls on raw reads data, reads alignment to the human reference genome (UCSC build hg19), duplication marking, SNVs and Indels calling and annotation. Firstly, we used an in-house tool called GaeaFastqQC to perform quality control on the sequencing data, removing low quality reads (reads with more than 10% unidentified base calls or more than 50% low quality base calls) and reads with sequencing adaptor contamination. We next utilized an open-source BWA (BWA 0.7.10, <http://bio-bwa.sourceforge.net/>) to align the clean

data from the previous phase to the human genome reference sequences (hg19) using BWA-backtrack algorithm with parameters “-L -l 31 -i 10 -a 500 -e 21 -I -t 10”. This was followed by the use of two programs based on Picard tools (<https://broadinstitute.github.io/picard/>) and GATK-Lite-2.3-9 (<ftp://anonymous:anon@ftp.broadinstitute.org/pub/gsa/GenomeAnalysisTK/GenomeAnalysisTKLite-2.3-9-gdcdccb.tar.bz2>) to mark duplication, realign indel and recalibrate the base quality scores. Then program based on samtools was used to sort the bam result and generated the final bam files. SNVs and Indels were exported using GATK-Lite-2.3-9 with parameters “-genotype_likelihoods_model BOTH -stand_call_conf 30.0 -stand_emit_conf 10.0 -dbnp file://\${dbsnp} -noMultiSampleCall”. The two sets of variants were merged, filtered and sorted using an in-house program named mergeVariant. Finally, the HbVar [12] and IthaGenes [14] Database was used to annotate the detected SNPs and Indels. The mutations were named in accordance with the published literature [15]. The complete procedure is depicted in Figure 2.

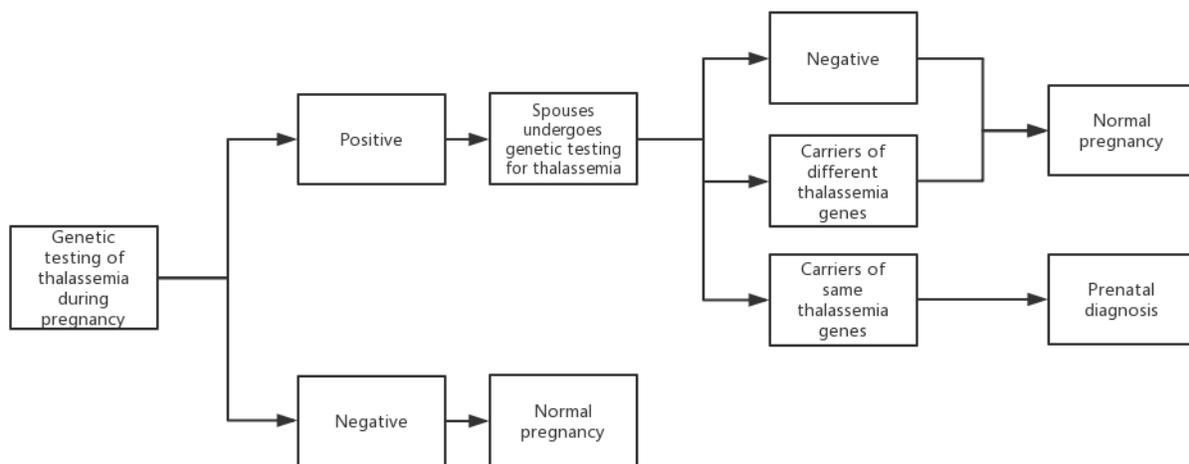


Figure 1: Flow chart for genetic screening of thalassemia.

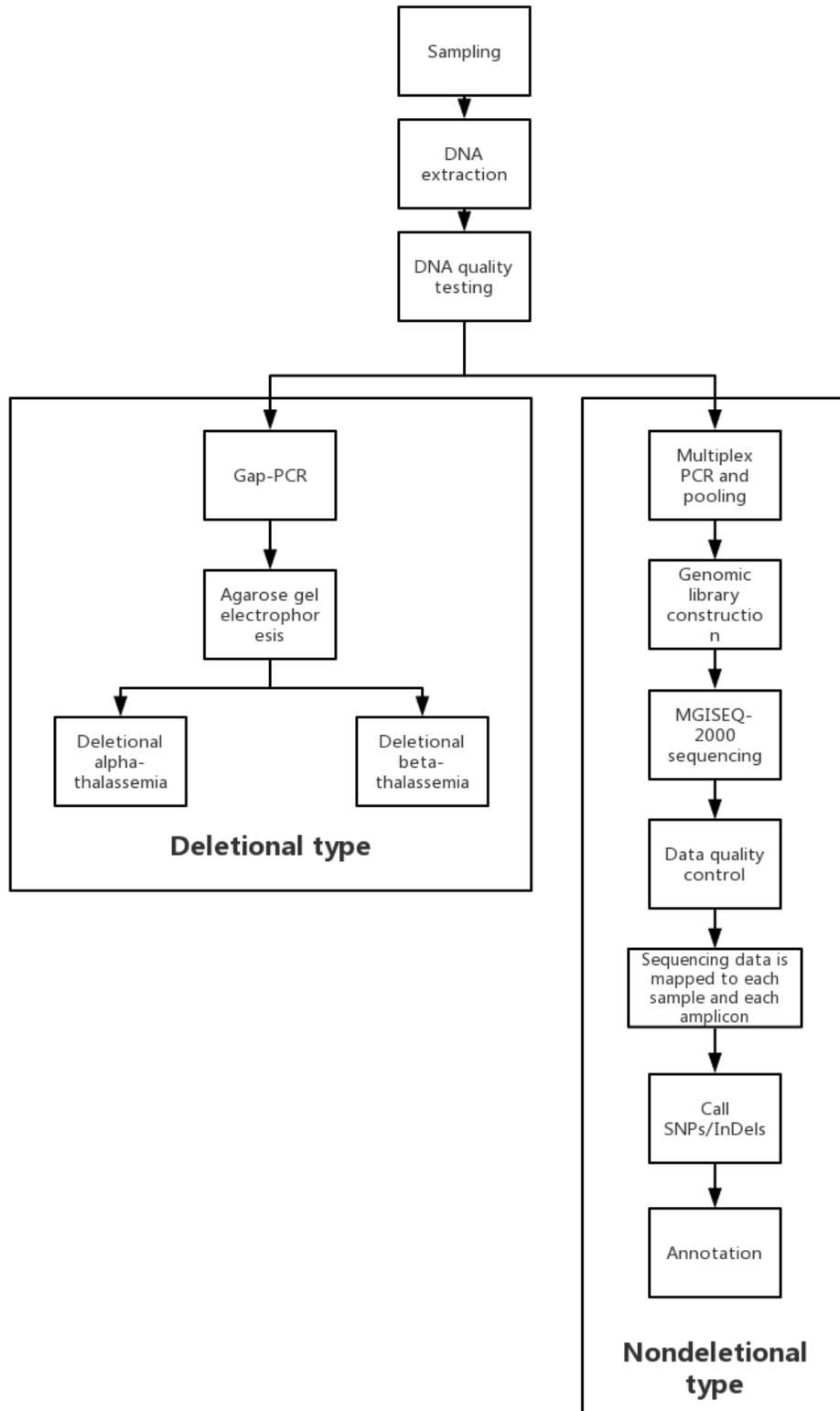


Figure 2: Diagram for the screening of hemoglobin variants and deletional and nondeletional alpha/beta-thalassemia.

2.3.1 Gap-PCR tests: As for suspected α thalassaemia and β thalassaemia carriers, multiplex gap-PCR was utilized to detect deletion-type α Thalassaemia: --SEA, $\alpha^{3.7}$, $-\alpha^{4.2}$, --FIL and --THAI and β thalassaemia deletions : Chinese Ggamma (Agammadeltabeta) 0, South-East Asia type hereditary persistence of fetal hemoglobin (SEA-HPFH) and Taiwanese.

3. Results

3.1 Statistical samples and demographic data

Between May 2018 and March 2019, this study included 28,941 participants (1,370 men, 27,519 women and 52 persons unknown). The age distribution of the participants was as follows: 17–20 years, 438 individuals; 20–30 years, 17,903 individuals; 30–40 years, 9,434 individuals; 40–50 years, 1,159 individuals; and 50 years, 5 individuals.

3.2 Thalassaemia carriers identified by NGS

A total of 28,941 people were involved in thalassaemia gene screening via NGS. Among these participants, 2,380 people were diagnosed as carriers of thalassaemia, including 1,694 carriers of α -thalassaemia, 638 carriers of beta-thalassaemia and 48 carriers of composite alpha- and beta-thalassaemia (Table S2). The total carrier rate of thalassaemia in Ji'an was 8.22%, and the carrier rates of alpha- and beta-thalassaemia were 5.85% and 2.20%, respectively. In addition, the first measured incidence rate of composite alpha- and beta-thalassaemia in Ji'an was 0.17%.

Among 1,694 carriers of alpha-thalassaemia, 15 different variations were identified in this study and accounted for 19 different genotypes (Table 1). $\alpha\alpha/--^{SEA}$ was the most common alpha-thalassaemia genotype, accounting for 41.97% of cases, followed by $\alpha\alpha/-\alpha^{3.7}$,

$\alpha\alpha/-\alpha^{4.2}$ and Hb Westmead/ $\alpha\alpha$, accounting for 38.67%, 10.68% and 4.43%, respectively. Notably, when compared with the conventional kit for alpha-thalassaemia gene testing in China, the screening performed in this study revealed the rare mutations Hb Phnom Penh in 5 cases, and CD 61 AAG>TAG (Lys>Stop), Hb Evanston, initiation codon (A>G), initiation codon (-T), poly A (A>G) and -FIL in 1 case each. A novel alpha-thalassaemia allele, HBA1: C.395insT, which can yield polypeptides completely different from the original alpha-globin peptide, was first identified in this study. In this cohort, we also found 20 beta-thalassaemia mutations and 21 genotypes (Table 2) among 638 participants. IVS-II-654 (C>T)/ β^N , codons 41/42 (-TTCT)/ β^N , codon 17 (A>T)/ β^N , -28 (A>G)/ β^N and codons 27/28 (+C)/ β^N were the most prevalent genotypes in current research, accounting for 31.97%, 23.35%, 13.95%, 12.70% and 4.86%, respectively. Additionally, alleles, such as -50 (G>A) beta+, Chinese Ggamma (Agammadeltabeta) 0, CAP +8 (C>T), SEA-HPFH, -72 (T>A) beta+ and codons 8/9 (+G) beta0, were identified using the screening method utilized in this study but were not detectable using traditional beta-thalassaemia genetic testing kits.

Forty-eight subjects had both alpha- and beta-globin variants (Table 3). Among these carriers, 79.17% of the genotypes were composed of the common deletion of the alpha-globin gene ($\alpha\alpha/-\alpha^{3.7}$, $\alpha\alpha/--^{SEA}$) and beta-globin gene point mutations, among which composite $\alpha\alpha/--^{SEA}$ and IVS-II-654 (C>T) beta+ heterozygosis was the most common genotype. Forty-six suspected couples were identified in this research. Among them, 22 couples, 16 couples and 8 couples had the risk of giving birth to Hb H disease patients, patients with Hb Bart's Hydrops fetalis syndrome and beta-thalassaemia major patients,

respectively. Among 126 people, 29 Hb variations were found, the majority of which were associated with normal phenotypes, however Hb Groene Hart, Hb Port Phillip, Hb Shenyang, and Hb Zurich-Albisrieden could show as cellule poor pigment anemia (Table 4).

3.3 Geographical distribution of thalassemia carriers in Ji'an

The total carrier rate of thalassemia in Ji'an was 8.22%, and the carrier rate in the southern counties was the highest (Figure 3a). Suichuan has the highest

carrier rate of the 1,694 alpha-thalassemia carriers, followed by Wan'an and Taihe (Figure 3b). Suichuan, Wan'an, and Xiajiang were the top three cities with the highest carrier rates, sequentially, according to the geographical distribution of the 638 beta-thalassemia carriers (Figure 3c). Moreover, we also counted the carriers of $\alpha\alpha/--^{SEA}$, the most prevalent genotype of alpha-thalassemia. As illustrated in Figure 3, a large disparity existed within a relatively limited geographical distance (Figure 3d). IVS-II-654 (C>T)/ β^N , a frequent beta-thalassemia genotype, also demonstrated the similar tendency (Figure 3e).

Genotype	Number	Frequency (%)
$\alpha\alpha/--^{SEA}$	737	42.3077
$\alpha\alpha/-\alpha3.7$	667	38.2893
$\alpha\alpha/-\alpha4.2$	183	10.5052
Hb Westmead/ $\alpha\alpha$	76	4.3628
Hb Constant Spring (Hb CS) / $\alpha\alpha$	36	2.0666
Hb Quong Sze/ $\alpha\alpha$	18	1.0333
Hb Phnom Penh/ $\alpha\alpha$	5	0.287
$\alpha\alpha/--^{THAI}$	4	0.2296
($\alpha\alpha/-\alpha3.7$)(Hb Constant Spring (Hb CS) / $\alpha\alpha$)	3	0.1722
CD 61 AAG>TAG (Lys>Stop) / $\alpha\alpha$	2	0.1148
$-\alpha3.7/-\alpha3.7$	2	0.1148
HK $\alpha\alpha/--^{SEA}$	2	0.1148
Hb Evanston/ $\alpha\alpha$	1	0.0574
Initiation codon (A>G) / $\alpha\alpha$	1	0.0574
Initiation codon (-T) / $\alpha\alpha$	1	0.0574
Poly A (A>G) / $\alpha\alpha$	1	0.0574
$-\alpha4.2/-\alpha4.2$	1	0.0574
$\alpha\alpha/--^{FIL}$	1	0.0574
($\alpha\alpha/--^{SEA}$)(Hb Westmead/ $\alpha\alpha$)	1	0.0574

Table 1: Distribution of α -thalassemia genotypes in Ji'an Region.

Genotype	Number	Frequency (%)
IVS-II-654 (C>T)/β ^N	204	31.974922
Codons 41/42 (-TTCT) /β ^N	149	23.354232
Codon 17 (A>T) /β ^N	89	13.949843
-28 (A>G) /β ^N	81	12.695925
Codons 27/28 (+C) /β ^N	31	4.8589342
5'UTR +43 to +40 (-AAAC) /β ^N	25	3.9184953
-50 (G>A) /β ^N	16	2.507837
Codons 71/72 (+A) /β ^N	14	2.1943574
Hb E/β ^N	7	1.0971787
ChineseGgamma(Agammadeltabeta)0/β ^N	3	0.4702194
-29 (A>G) /β ^N	4	0.6269592
IVS II-761 (A>G) beta (0 or + unclear) /β ^N	2	0.3134796
CAP +8 (C>T) /β ^N	2	0.3134796
Codon 43 (G>T) /β ^N	2	0.3134796
Codons 14/15 (+G) /β ^N	2	0.3134796
SEA-HPFH/β ^N	2	0.3134796
-50 (G>A) /Codon 17 (A>T)	1	0.1567398
-72 (T>A)/β ^N	1	0.1567398
Codons 8/9 (+G) /β ^N	1	0.1567398
Initiation codon ATG>AGG/β ^N	1	0.1567398
IVS-I-5 (G>C) /β ^N	1	0.1567398

Table 2: Distribution of β-thalassemia genotypes in Ji'an Region.

α	β	Number
αα/--SEA	IVS-II-654 (C>T) /β ^N	11
αα/--SEA	Codon 17 (A>T) /β ^N	5
αα/-α3.7	IVS-II-654 (C>T) /β ^N	5
αα/--SEA	-28 (A>G) /β ^N	3
αα/--SEA	ChineseGgamma(Agammadeltabeta)0/β ^N	2
αα/-α3.7	-28 (A>G) /β ^N	2
αα/-α3.7	Codon 17 (A>T) /β ^N	2
αα/-α3.7	Codons 41/42 (-TTCT) /β ^N	2
αα/--SEA	Codons 27/28 (+C) /β ^N	1
αα/--SEA	Codons 41/42 (-TTCT) /β ^N	1

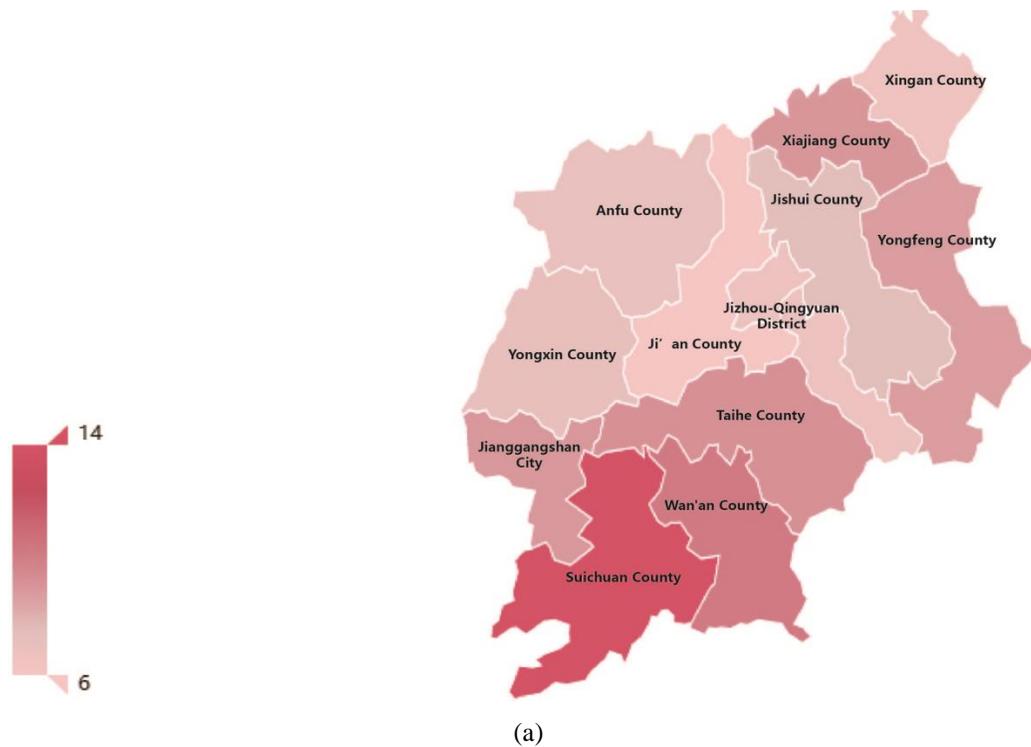
$\alpha\alpha/--SEA$	Codons 71/72 (+A) / β^N	1
$\alpha\alpha/--SEA$	-50 (G>A) / β^N	1
$\alpha\alpha/--SEA$	IVS I-7 (A>T) /Codons 41/42 (-TTCT)	1
- $\alpha3.7/-\alpha3.7$	Hb E/ β^N	1
CD 61 AAG>TAG [Lys>STOP]/ $\alpha\alpha$	IVS-II-654 (C>T) / β^N	1
Hb Constant Spring (Hb CS)/ $\alpha\alpha$	Codons 41/42 (-TTCT) / β^N	1
Hb Quong Sze/ $\alpha\alpha$	Hb E/ β^N	1
Hb Quong Sze/ $\alpha\alpha$	IVS II-761 (A>G) beta (0 or + unclear)/ β^N	1
Hb Westmead/ $\alpha\alpha$	-50 (G>A) / β^N	1
$\alpha\alpha/--THAI$	-50 (G>A) / β^N	1
$\alpha\alpha/--THAI$	IVS-II-654 (C>T) / β^N	1
$\alpha\alpha/-\alpha3.7$	5'UTR +43 to +40 (-AAAC) / β^N	1
$\alpha\alpha/-\alpha4.2$	-28 (A>G) / β^N	1
$\alpha\alpha/-\alpha4.2$	IVS-II-654 (C>T) / β^N	1

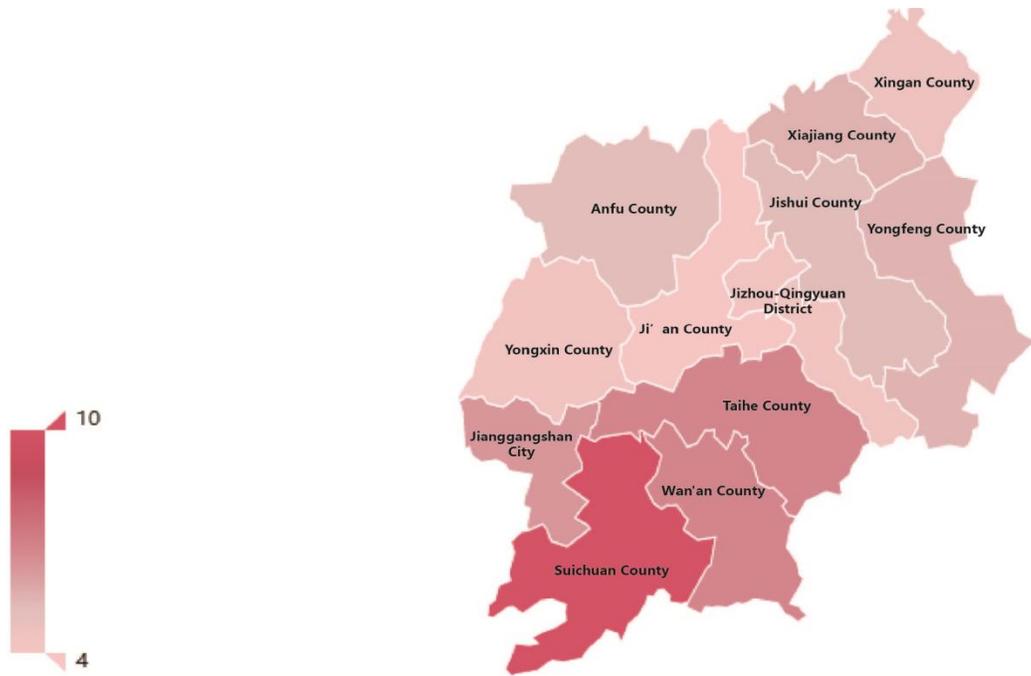
Table 3: Genotypes of composite α and β -thalassemia in Ji'an Region.

Abnormal Hemoglobin	HGVS	Number	Phenotype
Hb Allison Park	HBA1:c.215C>T	1	Normal
Hb Buzen	HBB:c.415G>A	4	Normal
Hb City of Hope	HBB:c.208G>A	1	Normal
Hb Deer Lodge	HBB:c.8A>G	1	Normal
Hb D-Los Angeles	HBB:c.364G>C	1	Normal
Hb G-Coushatta	HBB:c.68A>C	1	Normal
Hb G-Georgia	HBA2:c.287C>T	1	Normal
Hb G-Honolulu	HBA2:c.91G>C	10	Normal
Hb Groene Hart	HBA1:c.358C>T	2	It belongs to thalassaemia and abnormal hemoglobin, which is characterized by small cell hypochromic anemia.
Hb G-Siriraj	HBB:c.22G>A	1	Normal
Hb G-Taipei	HBB:c.68A>G	2	Normal
Hb Hamilton	HBB:c.34G>A	2	Normal
Hb Hekinan II	HBA1:c.84G>T	22	Normal
Hb Iraq-Halabja	HBB:c.32C>T	2	Normal
Hb I	HBA1:c.49A>G	1	Normal
Hb J-Bangkok	HBB:c.170G>A	7	Normal

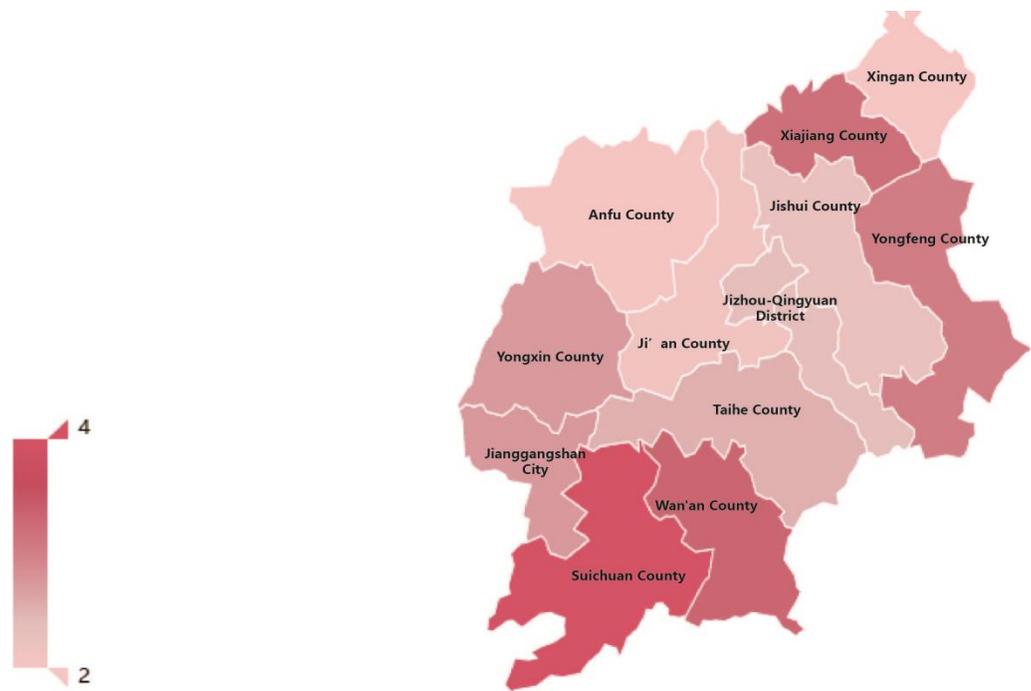
Hb J-Wenchang-Wuming	HBA2:c.34A>C	1	Normal
Hb Las Palmas	HBB:c.149C>T	1	Normal
Hb New York	HBB:c.341T>A	14	Normal
Hb Owari	HBA1:c.364G>A	16	Normal
Hb Port Phillip	HBA2:c.275T>C	6	Small cell hypochromic anemia
Hb Q-Thailand	HBA1:c.223G>C	10	Normal
Hb Russ	HBA2:c.154G>C	2	Normal
Hb Shenyang	HBA2:c.80C>A	1	Small cell hypochromic anemia
Hb South Yorkshire	HBA2:c.151C>T	1	Normal
Hb Ube-2	HBA1:c.205A>G	1	Normal
Hb Zengcheng	HBB:c.343C>A	11	Normal
Hb Zurich-Albisrieden	HBA2:c.178G>C	1	It belongs to thalassaemia and abnormal hemoglobin, which is characterized by small cell hypochromic anemia.
Hb Zurich-Langstrasse	HBB:c.151A>T	2	Normal

Table 4: Hemoglobin variants in this cohort.





(b)



(c)

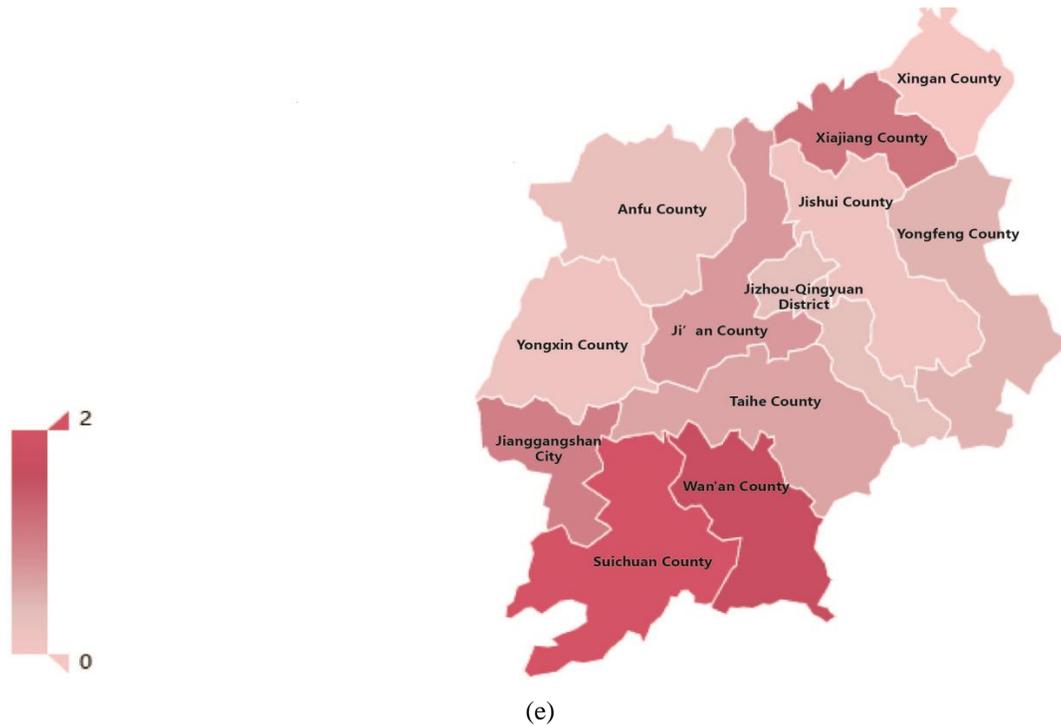
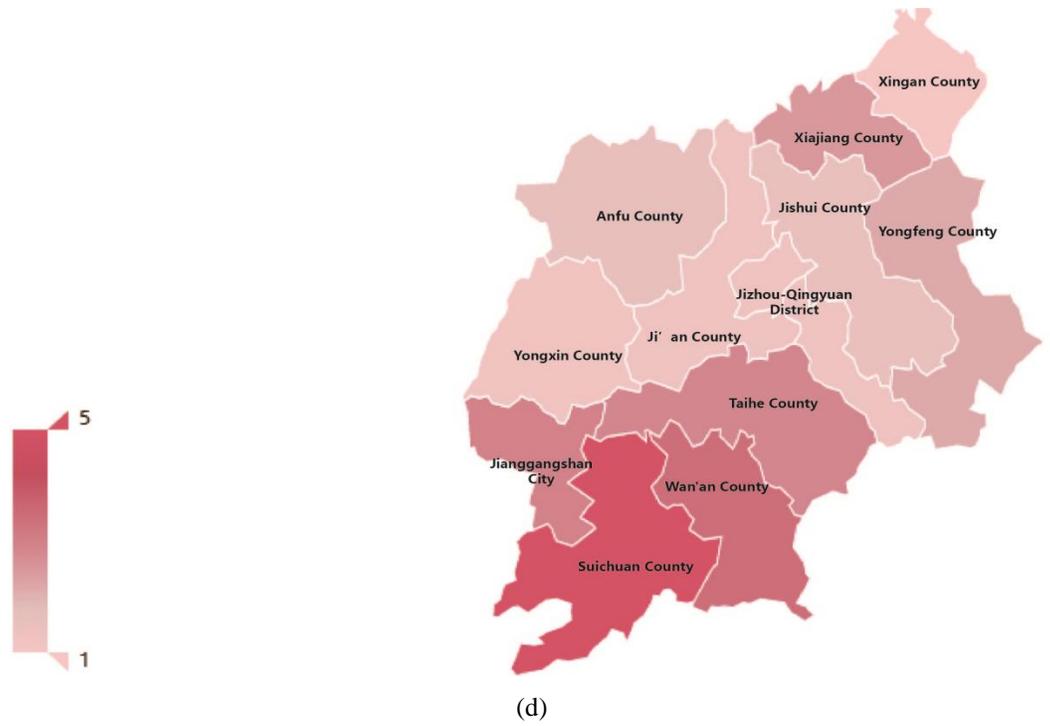


Figure 3: Geographical distribution of thalassemia carriers in this study. (a) thalassemia carrier rate; (b) α -thalassemia carrier rate ; (c) $\alpha\alpha$ --SEA genotype carrier rate; (d) β -thalassemia carrier rate; (e) IVS-II-654 (C>T)/ β N carrier rate. The map depicted in figure is under copyright and its use was granted by <https://www.tubiaoixiu.com>. The authors have the appropriate permissions from the copyright holders.

4. Discussion

This was the first study to conduct a molecular epidemiology investigation of thalassemia among the population in Ji'an. In our research, 28,941 people were screened for alpha/beta-thalassemia. Women aged 20-30 years made up the majority of participants, accounting for 61.86%. Our findings indicate that 5.85% and 2.20% of people in this area were carriers of alpha- and beta-thalassemia, respectively, and that the incidence rate of composite alpha- and beta-thalassemia was 0.17%. Ji'an's southern counties had a high proportion of thalassemia carriers and varied geographically.

The carrier rate of alpha-thalassemia has been found to be substantially higher than in Nanchang (1.49%) and Xinyu (2.2%), but slightly lower than in Ganzhou (7.19%) in Jiangxi [13]. The carrier rate of beta-thalassemia is higher than in Nanchang (1.14%), Xinyu (1.7%) and also slightly lower than Ganzhou (2.3%) [13]. Furthermore, the alpha-thalassemia and beta-thalassemia composite carrier rates in Jiangxi are similar to those in Ganzhou (0.18%) [13]. The most prevalent alpha-thalassemia mutation was $\alpha\alpha/--^{SEA}$, which was consistent with earlier observations [16, 17]. Codons 41/42 (-TTCT) and IVS-II-654 (C>T) were the two most prevalent beta-thalassemia alleles. The ranking of the two main alleles differed from those found in prior Jiangxi Province studies [13]. We think this may be caused by the new genetic screening method. Our research indicated that the carrier rates of alpha- and beta-thalassemia in this area were significantly lower than those in Guangdong Province (11.07%) [16] and Guangxi Zhuang Autonomous Region (24.51%) [18], two high-incidence areas of alpha/beta-thalassemia. The frequency spectrum of alpha- and beta-thalassemia mutations in this study is similar to that previously

described in South China, e.g., Chenzhou [12]. For the composite genotypes, $\alpha\alpha/--^{SEA}$ and IVS-II-654 (C>T) / β^N were more than other genotypes. The thalassemia carrier rate in South China was higher than that in North China, similar to the general geographical distribution of people with thalassemia in China. We also found a significant difference between the common mutation of alpha-thalassemia, $\alpha\alpha/--^{SEA}$, and the common mutation of beta-thalassemia, IVS-II-654 (C>T) beta+, within a narrow geographical radius. Suichuan had a greater rate of alpha/beta-thalassemia carriers than other cities in the province. Suichuan, in southwest Ji'an, is close to Ganzhou, which borders Guangdong Province's Meizhou. The alpha/beta-thalassemia carrier rate in Meizhou is greater than the average carrier rate in Guangdong Province, and one of the key causes for the higher alpha/beta-thalassemia carrier rate in Suichuan could be population movement.

High-throughput amplicon sequencing were adopted to screen the 301 thalassemia alleles in this research, and the molecular epidemiology data of Ji'an was first comprehensively reported. In recent years, with the advancement and development of sequencing techniques, high-throughput sequencing can now test almost all mutations in thalassemia gene sequences; additionally, it has a high throughput and low cost, making it ideal for screening thalassemia genes and identifying rare alleles of thalassemia. At present, the three-step thalassemia screening program of routine blood samples (indexes: MCV < 80 fL, MCH < 27 and MCHC < 320), Hb electrophoresis (indexes: HbA2 < 2.5%, HbA2 > 3.5% and HbF increases with abnormal Hb strips) and routine genetic testing is primarily used in clinical practice. Although the

three-step screening program plays a significant role in the prevention and control of thalassemia, as screening technology has progressed, the method's flaws have become more evident, which are mostly reflected in the following features. (1) A routine blood test can only hint at the possibility of thalassemia; it cannot be used to make a conclusive diagnosis, and it may fail to detect thalassemia carriers who have normal MCV and MCH. In 2017, Zhu Baosheng of Yunnan Province's First People's Hospital led a team to investigate the thalassemia-carrying condition of 951 Dai Ethnic Minority Group members in Yunnan using routine blood and high-throughput sequencing strategies, and they discovered that the failed detection rate in routine blood screening was 17.1-23.4% [4]. (2) Hb electrophoresis testing is suitable for the definitive diagnosis of most carriers of β -thalassemia minor and thalassaemia patients (α and β), but it is insufficiently sensitive for the diagnosis of α -thalassemia carriers. According to research by Zhu et al., due to HbA₂ >2.5, more than 90% of $-\alpha^{3.7}/\alpha\alpha$ -type carriers failed to be detected using hemoglobin electrophoresis [4]. When it comes to detecting thalassemia, hemoglobin electrophoresis has a significant failure rate. (3) Routine genetic testing is mainly directed to the testing of 23 common gene types (3 deletion types and 3 mutation types of α -thalassemia and 17 mutation types of β -thalassemia); therefore, it cannot be utilized to screen for unusual kinds, resulting in a 2% failure rate. However, the NGS technique still has some limitations. For instance, Zebisch et al. identified the new mutation of ϵ - γ - δ - β thalassemia with MLPA and CGH, while NGS cannot test this mutation [19]. In light of the different variations in thalassemia, based on cost, different methods should be used in combination to comprehensively test for all variations of thalassemia.

In summary, we investigated the carrier rates of variations related to thalassemia in Ji'an using NGS and demonstrated the diversity of relevant variations. In this study, a novel mutation was discovered, demonstrating that applying NGS to routine thalassemia gene screening can effectively reduce the failed detection ratio of unique or rare genotypes and contribute to the improvement of thalassemia prevention and control. Domestic and foreign practices have illustrated that the birth rate of babies with thalassemia major can be effectively reduced through premarital and pregnancy physical examination, prenatal screening, prenatal diagnosis and other prevention and control measures. During the implementation process, the per capita cost of this program was kept within RMB 200. The screening strategy of first examining pregnant women and then deciding whether or not to test their spouses depending on the results was adopted, and this method may be a better method for thalassemia gene screening.

5. Conclusions

Our findings have significant implications for thalassemia prevention and treatment in this and other high-prevalence areas. The epidemiological data updated in this study may allow the local government to concentrate on the severity of the disease and devise methods for effective resource allocation under resource constraints.

Declarations

Ethics approval and consent to participate

The experimental protocol was established, according to the ethical guidelines of the Helsinki Declaration and was approved by the Human Ethics Committee of First People's Hospital of Ji'an. This research was approved by the Ethics Committee of First People's

Hospital of Ji'an, and all the participants signed written informed consent. Written consent was obtained from parents/legal guardians of any participants under 16 years of age.

Consent to publish

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

No funding was obtained for this study.

Authors' Contributions

LWM performed the data generation, processing, and analysis and prepared the paper and figures; YQ,SPC,HL and HRW contributed to data analysis and interpretation; LPG, JH, XW, YL contributed to result interpretation; JX, YF designed the study and carried out sample collection, storage and transport; All authors have read and approved the manuscript.

Acknowledgements

We are grateful to patients and their families who kindly consented to join the study.

References

1. Shang X, Peng Z, Ye Y, et al. Rapid Targeted Next-Generation Sequencing Platform for Molecular Screening and Clinical Genotyping in Subjects with Hemoglobinopathies. *EBioMedicine* 23 (2017): 150-159.
2. Shang X, Xu X. Update in the genetics of thalassemia: What clinicians need to know. *Best Pract Res Clin Obstet Gynaecol* 39 (2017): 3-15.
3. Kassebaum NJ, Jasrasaria R, Naghavi M, et al. A systematic analysis of global anemia burden from 1990 to 2010. *Blood* 123 (2014): 615-624.
4. He J, Song W, Yang J, et al. Next-generation sequencing improves thalassemia carrier screening among premarital adults in a high prevalence population: the Dai nationality, China. *Genet Med* 19 (2017): 1022-1031.
5. Weatherall DJ. The inherited diseases of hemoglobin are an emerging global health burden. *Blood* 115 (2010): 4331-4336.
6. Colah R, Gorakshakar A, Nadkarni A. Global burden, distribution and prevention of beta-thalassemsias and hemoglobin E disorders. *Expert Rev Hematol* 3 (2010): 103-117.
7. Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ* 86 (2008): 480-487.
8. Lai K, Huang G, Su L, et al. The prevalence of thalassemia in mainland China: evidence from epidemiological surveys. *Sci Rep* 7 (2017): 920.
9. Lin YH, Lin W, Wang XX. Genotyping of Patients with alpha and beta Thalassemia in Fujian Province Area in China. *Zhongguo Shi Yan Xue Ye Xue Za Zhi* 27 (2019): 899-903.
10. Schrier SL, Angelucci E. New strategies in the treatment of the thalassemsias. *Annu Rev Med* 56 (2005): 157-171.
11. Pappasava T, van Ijcken WF, Kockx CE, et al. Next generation sequencing of SNPs for non-invasive prenatal diagnosis: challenges and feasibility as illustrated by an application to beta-thalassaemia. *Eur J Hum Genet* 21 (2013): 1403-1410.

12. Zhang H, Li C, Li J, et al. Next-generation sequencing improves molecular epidemiological characterization of thalassemia in Chenzhou Region, P.R. China. *J Clin Lab Anal* 33 (2019): e22845.
13. Lin M, Zhong TY, Chen YG, et al. Molecular epidemiological characterization and health burden of thalassemia in Jiangxi Province, P. R. China. *PLoS One* 9 (2014): e101505.
14. Kountouris P, Lederer CW, Fanis P, et al. IthaGenes: an interactive database for haemoglobin variations and epidemiology. *PLoS One* 9 (2014): e103020.
15. Patrinos GP, Giardine B, Riemer C, et al. Improvements in the HbVar database of human hemoglobin variants and thalassemia mutations for population and sequence variation studies. *Nucleic Acids Res* 32 (2004): D537-D541.
16. 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* 57 (2004): 517-522.
17. Lin M, Wen YF, Wu JR, et al. Hemoglobinopathy: molecular epidemiological characteristics and health effects on Hakka people in the Meizhou region, southern China. *PLoS One* 8 (2013): e55024.
18. Xu XM, Li ZQ, Liu ZY, et al. Molecular characterization and PCR detection of a deletional HPFH: application to rapid prenatal diagnosis for compound heterozygotes of this defect with beta-thalassemia in a Chinese family. *Am J Hematol* 65 (2000): 183-188.
19. Zebisch A, Schulz E, Grosso M, et al. Identification of a novel variant of epsilon-gamma-delta-beta thalassemia highlights limitations of next generation sequencing. *Am J Hematol* 90 (2015): E52-E54

Supplementary

Here, is the list of [supplementary files](#). Click to download.



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC-BY\) license 4.0](#)