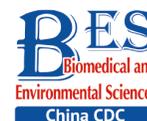


Original Article



Epidemiological Survey of Hemoglobinopathies Based on Next-Generation Sequencing Platform in Hunan Province, China*

XI Hui^{1,&}, LIU Qin^{1,&}, XIE Dong Hua¹, ZHOU Xu¹, TANG Wang Lan¹, TANG De Guo², ZENG Chun Yan², WANG Qiong³, NIE Xing Hui⁴, PENG Jin Ping⁵, GAO Xiao Ya⁶, WU Hong Liang⁷, ZHANG Hao Qing⁸, QIU Li⁹, FENG Zong Hui¹⁰, WANG Shu Yuan¹¹, ZHOU Shu Xiang¹², HE Jun¹³, ZHOU Shi Hao¹³, ZHOU Fa Qun¹⁴, ZHENG Jun Qing¹⁵, WANG Shun Yao¹⁶, CHEN Shi Ping¹⁷, ZHENG Zhi Fen¹⁸, MA Xiao Yuan¹⁹, FANG Jun Qun¹, LIANG Chang Biao¹, and WANG Hua^{1,#}

1. NHC Key Laboratory of Birth Defect for Research and Prevention, Hunan Provincial Maternal and Child Health Care Hospital, Changsha 410028, Hunan, China; 2. Yongzhou Hospital for Maternal and Child Health Care, Yongzhou 425000, Hunan, China; 3. Changde Hospital for Maternal and Child Health Care, Changde 415099, Hunan, China; 4. Hengyang Hospital for Maternal and Child Health Care, Hengyang 421009, Hunan, China; 5. Shaoyang Hospital for Maternal and Child Health Care, Shaoyang 422001, Hunan, China; 6. Yiyang Hospital for Maternal and Child Health Care, Yiyang 413099, Hunan, China; 7. Yueyang Hospital for Maternal and Child Health Care, Yueyang 414021, Hunan, China; 8. Center of Prenatal Diagnosis, Chenzhou No. 1 People's Hospital, Chenzhou 423099, Hunan, China; 9. Loudi Hospital for Maternal and Child Health Care, Loudi 417599, Hunan, China; 10. Huaihua Hospital for Maternal and Child Health Care, Huaihua 418099, Hunan, China; 11. Xiangtan Hospital for Maternal and Child Health Care, Xiangtan 411104, Hunan, China; 12. Zhuzhou Hospital for Maternal and Child Health Care, Zhuzhou 412008, Hunan, China; 13. Changsha Hospital for Maternal and Child Health Care, Changsha 410007, Hunan, China; 14. Xiangxi Tujia and Miao Autonomous Prefecture Hospital for Maternal and Child Health Care, Jishou 416000, Hunan, China; 15. Zhang Jiajie Centre for maternal and child health family planning services, Zhangjiajie 427299, Hunan, China; 16. BGI Genomics, BGI-Shenzhen, Shenzhen 518083, Guangdong, China; 17. Clinical Laboratory of BGI Health, BGI-Shenzhen, Shenzhen 518083, Guangdong, China; 18. BGI-Wuhan Clinical Laboratories, BGI-Shenzhen, Wuhan 430074, Hubei, China; 19. SEBIA Representative Office, Shanghai 200001, China

Abstract

Objective This study was aimed at investigating the carrier rate of, and molecular variation in, α - and β -globin gene mutations in Hunan Province.

Methods We recruited 25,946 individuals attending premarital screening from 42 districts and counties in all 14 cities of Hunan Province. Hematological screening was performed, and molecular parameters were assessed.

Results The overall carrier rate of thalassemia was 7.1%, including 4.83% for α -thalassemia, 2.15% for

*This study was supported by the National Key Research and Development Program of China [2021YFC1005300]; the science and technology innovation Program of Hunan Province—Major Scientific and Technological Projects for Collaborative Prevention and Control of Birth Defects in Hunan Province [2019SK1010 and 2019SK1011]; and Hunan Province Clinical Medical Technology Innovation Guidance Project "Screening, prevention and control of single gene disease carriers and panel research in childbearing age people in Hunan Province" [2021SK50602].

&These authors contributed equally to this work.

#Correspondence should be addressed to WANG Hua, E-mail: wanghua213@aliyun.com

Biographical notes of the first authors: XI Hui, female, born in 1979, Master's Degree, Deputy Chief Physician, majoring in prenatal diagnosis, genetic counseling and other birth defect prevention and control; LIU Qin, female born in 1987, Master's Degree, in Charge Inspector, majoring in birth defects such as thalassemia prevention and control, single gene diseases detection, and so on.

β -thalassemia, and 0.12% for both α - and β -thalassemia. The highest carrier rate of thalassemia was in Yongzhou (14.57%). The most abundant genotype of α -thalassemia and β -thalassemia was $-\alpha^{3.7}/\alpha\alpha$ (50.23%) and $\beta^{IVS-II-654}/\beta^N$ (28.23%), respectively. Four α -globin mutations [CD108 (ACC>AAC), CAP +29 (G>C), Hb Agrinio and Hb Cervantes] and six β -globin mutations [CAP +8 (C>T), IVS-II-848 (C>T), -56 (G>C), beta nt-77 (G>C), codon 20/21 (-TGGA) and Hb Knossos] had not previously been identified in China. Furthermore, this study provides the first report of the carrier rates of abnormal hemoglobin variants and α -globin triplication in Hunan Province, which were 0.49% and 1.99%, respectively.

Conclusion Our study demonstrates the high complexity and diversity of thalassemia gene mutations in the Hunan population. The results should facilitate genetic counselling and the prevention of severe thalassemia in this region.

Key words: Thalassemia; Carrier rate; Molecular spectrum; Abnormal hemoglobin variants; α -globin triplication

Biomed Environ Sci, 2023; 36(2): 127-134 doi: 10.3967/bes2023.016

ISSN: 0895-3988

www.besjournal.com (full text)

CN: 11-2816/Q

Copyright ©2023 by China CDC

INTRODUCTION

Thalassemia is an autosomal recessive disease involving moderate or severe hemolytic anemia caused by insufficient or absent globin chains. It is widespread in the Mediterranean, Southeast Asia and southern China^[1]. The most common forms are α -thalassemia (OMIM: #604131) and β -thalassemia (OMIM: #613985), which affect the synthesis of the α - and β -globin subunits, respectively^[2]. An estimated 1%–5% of the global population carries a genetic thalassemia mutation^[3]. The frequency of thalassemia gene carriage in southern China is as high as 3%–24%^[4]. People carrying thalassemia variants are concentrated in Guangdong, Guangxi, and Hainan Provinces in China^[5,6]. Notably, as a consequence of recent massive population migrations, thalassemia is no longer restricted to traditional high-incidence regions and is now a relatively common clinical problem in areas including Hunan, Jiangxi, and Fujian. The genetic spectrum of thalassemias in Hunan Province shows high complexity and diversity and remains to be clearly elucidated. Hence, we performed a large-scale next-generation sequencing survey of reproductive age couples covering 42 districts and counties in 14 cities in Hunan Province.

MATERIALS AND METHODS

Population Samples and Statistical Analysis

A cross-sectional study was performed among couples of reproductive age receiving routine health

examinations between June 2020 and April 2021 in 14 cities of Hunan Province. One or both members of each couple were registered in Hunan Province.

Multi-stage stratified random sampling was used. The total required sample size was summed after estimation of the sample size for 14 cities. The number of samples for each city was calculated with a formula^[7]. The thalassemia carrier rate of each city was estimated on the basis of previous studies and the thalassemia carrier rate of neighboring provinces^[8-11]. The number of samples was increased 15% for each city to avoid data deficiencies (Supplementary Table S1, available in www.besjournal.com).

The required samples of 14 cities were distributed to districts and counties according to their proportion of population. The population ratio of districts to counties in the entire province was approximately 1:2. One district and two counties were randomly chosen in each city, and a total of 42 districts and counties were included in the study. Carrier rate ratio of each city was calculated by carrier rate of each city multiplying by their population rate (Carrier rate ratio = Carrier rate \times population rate). Total carrier rate was equal to the sum of the carrier ratio of all cities.

A total of 25,946 individuals of reproductive age receiving routine health examinations between June 2020 and April 2021 were included in this epidemiological survey of thalassemia in Hunan Province. All participants were collected through a random sampling method and came from 14 cities across Hunan Province. The ages of the participants ranged from 20 to 42 years. Informed written consent was provided by all participants after a clear

description of the study objectives was provided. This study was approved by the medical ethics committee of the Hunan Provincial Maternal and Child Health Care Hospital (202030).

Hematological Analyses

Blood samples (3 mL) from all participants were collected in EDTA tubes. Complete blood cell count, hemoglobin electrophoresis and molecular parameters were assessed simultaneously. Hematological parameters were determined with an automated hematology analyzer, and hemoglobin analysis was performed with capillary electrophoresis (Sebia, France and Helena, USA).

Genetic Analysis

DNA Extraction Genomic DNA was extracted from whole blood with a Q Kingfisher Flex system (Thermo Scientific, Rockford, IL) and isolated with a GenMag Nucleic Acid Isolation kit (Magnetic bead method) (GenMagBio, Beijing, China). A NanoDrop-8,000 spectrophotometer (Thermo Scientific, Waltham, MA, USA) was used to quantify the concentrations of DNA samples.

Library Preparation and NGS Sequencing For the variants of the α and β globin genes (*HBA1*, *HBA2*, and *HBB*), five α -thalassemia associated deletions ($-\alpha^{3.7}$, $-\alpha^{4.2}$, $--_{SEA}$, $--_{FIL}$, and $--_{THAI}$) and three β -globin gene deletions [Chinese γ ($\gamma\delta\beta$)⁰-thal, Southeast Asia HPFH(SEA-HPFH), and 1357 bp deletion], a tag sequence was introduced for identification, and the sequences were amplified and enriched through multiplex PCR with primers described in a patent [patent ID CN108796054A] and published article^[12]. The PCR products of samples (≤ 96) were pooled for library preparation. For sequencing and library identification, the DNA sequences were supplemented with a linker sequence in each library. The detailed experimental protocol and bioinformatic analysis methods were as previously described^[13]. We used a next-generation sequencing library protocol for library construction, including purified genomic DNA (Qiagen DNA Purification kit, Hilden, Germany), DNA quantification (NanoDrop 8,000 UV-vis spectrophotometer; Thermo Fisher Scientific, Waltham, MA, USA), DNA fragmentation (excluding CNV amplicons), blunt-ended fragmentation (Enzymatics kits, Qiagen, Hilden, Germany), 3'-dA overhang, paired-end adapter ligation, DNA fragment separation (CNVs not included) and size selection with magnetic beads. Finally, sequencing was performed with the paired-end tag (PE100) on a MGISEQ-2,000 chip. Positive

NGS results were validated with Sanger sequencing.

Bioinformatic Analyses A bioinformatic pipeline focused on detecting Hb gene deletions and point mutations was developed. First, we filtered low quality reads and reads with sequencing adaptor contamination, and classified the clean data by using adaptor information (index primer). Then, we aligned the clean data to the human genome reference sequence (hg19) in BWA software^[14] and generated the final bam files with SAMtools. Each consensus sequence's length, coverage and depth were recorded with ReSeqTools^[13]. Finally, the HbVar^[15] and lthaGenes^[16] databases were used to annotate the detected mutation types.

α -globin Triplication The NGS suggested high risk of suspected triplication among participants, as further confirmed by an anti-3.7/4.2 multiplex-PCR method^[17]. The phenotypic data were assessed with factorial ANOVA (factors including RBC, HB, MCV, MCH, Hb_A, Hb_A2, and HbF), and the *P* value indicated no statistical difference ($P > 0.05$).

RESULTS

Carrier Rate of Thalassemia in Hunan Province

The overall carrier rate of thalassemia in Hunan Province was 7.1%. The carrier rates among the 14 cities varied between 4.53% and 14.57% (Figure 1, Table 1). The two regions with the highest carrier rate of thalassemia were Yongzhou (14.57%) and

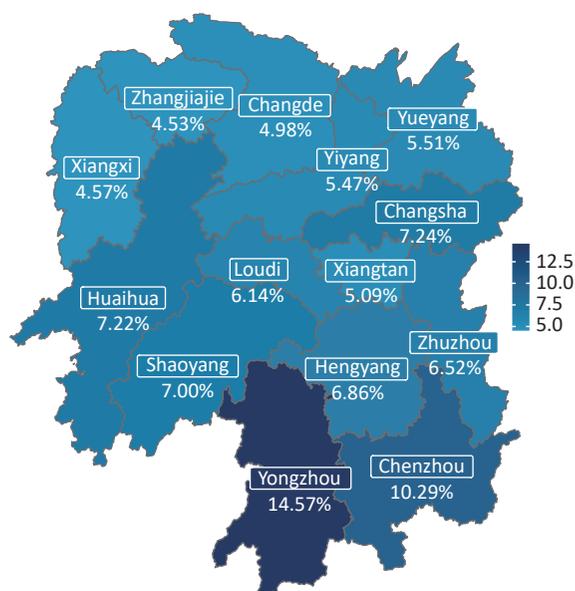


Figure 1. Geographical distribution of thalassemia carriers in Hunan Province.

Chenzhou (10.29%), which are bordered by Guangxi Province and Guangdong Province, respectively. Changsha, the capital city, which is located in the center of Hunan Province, ranked third, with a carrier rate of 7.24%. The lowest carrier rates were detected in Zhangjiajie (4.53%) and Xiangxi Prefecture (4.57%), in the northwestern part of the province.

The carrier rates of α -thalassemia, β -thalassemia, and α combined with β -thalassemia were 4.83%, 2.15% and 0.12%, respectively (Table 1). The top three regions with high α -thalassemia carrier rates were Yongzhou (9.26%), Chenzhou (7.30%) and Changsha (5.62%). β -thalassemia was encountered primarily in Yongzhou (5.06%), Chenzhou (2.88%) and Shaoyang (2.59%).

Mutation Spectrum of α - and β -thalassemia Genes in Hunan

In a total of 25,946 individuals screened for thalassemia, 1,597 were found to be thalassemia

carriers, of whom 1,109 were α -thalassemia carriers, 464 were β -thalassemia carriers, and 24 others were carriers of both α - and β -thalassemia. Among α -thalassemia carriers, 28 different genotypes were detected, and $-\alpha^{3.7}/\alpha$ was the most frequent genotype, accounting for more than half (50.23%).

Other highly prevalent genotypes of α -thalassemia were $-\alpha^{SEA}/\alpha$, $-\alpha^{4.2}/\alpha$, α^{WS}/α , α^{CS}/α , and α^{QS}/α . Overall, these six genotypes accounted for 96.30% of all α -thalassemia genotypes (Table 2). Rare α -globin genotypes were also identified and found to account for 2.43% of all genotypes (27/1,109). The top three uncommon genotypes were α^{CD30}/α (0.45%), α^{CD108}/α (0.45%), and α^{CD61}/α (0.36%). We identified 25 different β -thalassemia genotypes in 464 participants. The three most frequent mutations were seen in genotypes $\beta^{IVS-II-654}/\beta^N$ (28.23%), $\beta^{CD41-42}/\beta^N$ (27.37%), and β^{CD17}/β^N (13.36%). Thirteen rare

Table 1. Carrier rate of α -thalassemia, β -thalassemia, and composite α - and β -thalassemia among 14 cities in Hunan Province

Region	Samples	Thalassemia carriers, n (%)				Population, n (%)	Carrier rate ratio*, %			
		Total	α -thalassemia	β -thalassemia	α - and β -thalassemia		Total	α -thalassemia	β -thalassemia	α - and β -thalassemia
Yongzhou	810	118 (14.57)	75 (9.26)	41 (5.06)	2 (0.25)	5,194,275 (7.55)	1.10	0.70	0.38	0.02
Chenzhou	904	93 (10.29)	66 (7.30)	26 (2.88)	1 (0.11)	4,583,531 (6.66)	0.69	0.49	0.19	0.01
Changsha	2,098	152 (7.24)	118 (5.62)	32 (1.53)	2 (0.10)	6,512,131 (9.46)	0.69	0.53	0.14	0.01
Huaihua	1,288	93 (7.22)	72 (5.59)	21 (1.63)	0 (0.00)	5,092,480 (7.40)	0.53	0.41	0.12	0.00
Shaoyang	1,272	89 (7.00)	54 (4.25)	33 (2.59)	2 (0.16)	7,927,558 (11.52)	0.81	0.49	0.30	0.02
Hengyang	918	63 (6.86)	40 (4.36)	21 (2.29)	2 (0.22)	7,902,022 (11.48)	0.79	0.50	0.26	0.03
Zhuzhou	1,212	79 (6.52)	49 (4.04)	27 (2.23)	3 (0.25)	3,898,444 (5.66)	0.37	0.23	0.13	0.01
Loudi	2,083	128 (6.14)	78 (3.74)	46 (2.21)	4 (0.19)	4,323,502 (6.28)	0.39	0.24	0.14	0.01
Yueyang	3,302	182 (5.51)	129 (3.91)	51 (1.54)	2 (0.06)	5,648,840 (8.21)	0.45	0.32	0.13	0.00
Yiyang	2,649	145 (5.47)	103 (3.89)	42 (1.59)	0 (0.00)	4,307,933 (6.26)	0.34	0.24	0.10	0.00
Xiangtan	2,121	108 (5.09)	66 (3.11)	39 (1.84)	3 (0.14)	2,888,294 (4.20)	0.21	0.13	0.08	0.01
Changde	3,572	178 (4.98)	132 (3.70)	44 (1.23)	2 (0.06)	6,225,913 (9.05)	0.45	0.33	0.11	0.01
Xiangxi	1,225	56 (4.57)	41 (3.35)	14 (1.14)	1 (0.08)	2,845,811 (4.13)	0.19	0.14	0.05	0.00
Zhangjiajie	2,492	113 (4.53)	86 (3.45)	27 (1.08)	0 (0.00)	1,478,149 (2.15)	0.10	0.07	0.02	0.00
Total	25,946	1,597 (6.16)	1,109 (4.27)	464 (1.79)	24 (0.09)	68,828,883 (100)	7.10	4.83	2.15	0.12

Note. * Carrier rate ratio = Carrier rate \times population rate. Total carrier rate was equal to the sum of the carrier ratio of all cities.

β -globin genotypes were identified, which accounted for 6.68% (31/464) of the total, of which β^{-50}/β^N accounted for the highest proportion (2.37%). Among the 24 participants who were carriers for both α and β -thalassemia, 83.3% (20/24) of the genotypes were caused by coinheritance of common α -globin gene deletions ($-\alpha^{3.7}/\alpha\alpha$, $--^{SEA}/\alpha\alpha$, or $-\alpha^{4.2}/\alpha\alpha$) with a β -globin gene point mutation (Table 2).

Significant differences in genotype composition were observed among cities in Hunan Province. $-\alpha^{3.7}$ was the main α -thalassemia genotype in most cities. $--^{SEA}$ was the main genotype in Yongzhou and Zhuzhou, accounting for 47.50% and 44%, respectively (Figure 2A). $\beta^{IVS-II-654}$ and $\beta^{CD41-42}$ were the most common β -thalassemia genotypes in the entire

province (Figure 2B).

Results of Structural Hemoglobin Variants

A total of 127 participants were carriers of abnormal hemoglobin variants, with a rate of 0.49%. Hb New York (22.83%), Hb Hekinan II (13.39%), Hb Owari ($\alpha 1$) (7.87%), Hb J-Bangkok (β) (5.51%) and Hb Zengcheng (β) (5.51%) had the highest incidence among the abnormal hemoglobin variants in Hunan Province (Supplementary Table S3, available in www.besjournal.com).

Detection of α Triplication and Analysis of Hematological Phenotype

Triplication of the α -globin gene ($\alpha\alpha\alpha^{anti3.7}$ and $\alpha\alpha\alpha^{anti4.2}$) was detected in 71 participants of 3,569

Table 2. Distribution of α - and β -thalassemia genotypes in Hunan Province

α -thalassemia			β -thalassemia			$\alpha+\beta$ thalassemia		
Genotype	N	%	Genotype	N	%	Genotype	N	%
$-\alpha^{3.7}/\alpha\alpha$	557	50.23	$\beta^{IVS-II-654}/\beta^N$	131	28.23	$-\alpha^{3.7}/\alpha\alpha; \beta^{CD41-42}/\beta^N$	4	16.67
$--^{SEA}/\alpha\alpha$	312	28.13	$\beta^{CD41-42}/\beta^N$	127	27.37	$-\alpha^{3.7}/\alpha\alpha; \beta^{IVS-II-654}/\beta^N$	3	12.50
$-\alpha^{4.2}/\alpha\alpha$	108	9.74	β^{CD17}/β^N	62	13.36	$--^{SEA}/\alpha\alpha; \beta^{CD41-42}/\beta^N$	2	8.33
$\alpha^{WS}/\alpha/\alpha\alpha$	55	4.96	$\beta^{CD71-72}/\beta^N$	22	4.74	$-\alpha^{3.7}/\alpha\alpha; \beta^{-28}/\beta^N$	1	4.17
$\alpha^{CS}/\alpha/\alpha\alpha$	21	1.89	β^{CD26}/β^N	21	4.53	$-\alpha^{3.7}/\alpha\alpha; \beta^{5'UTR+43\ to\ +40}/\beta^N$	1	4.17
$\alpha^{OS}/\alpha/\alpha\alpha$	15	1.35	β^{-28}/β^N	20	4.31	$-\alpha^{3.7}/\alpha\alpha; \beta^{CD17}/\beta^N$	1	4.17
$\alpha^{CD30}/\alpha/\alpha\alpha$	5	0.45	$\beta^{5'UTR+43\ to\ +40}/\beta^N$	17	3.66	$-\alpha^{3.7}/\alpha\alpha; \beta^{CD71-72}/\beta^N$	1	4.17
$\alpha^{CD108}/\alpha/\alpha\alpha$	5	0.45	$\beta^{CD27-28}/\beta^N$	12	2.59	$-\alpha^{4.2}/\alpha\alpha; \beta^{CD17}/\beta^N$	1	4.17
$\alpha^{CD61}/\alpha/\alpha\alpha$	4	0.36	β^{-50}/β^N	11	2.37	$-\alpha^{4.2}/\alpha\alpha; \beta^{IVS-II-654}/\beta^N$	1	4.17
Hb Phnom Penh	3	0.27	β^{CD43}/β^N	10	2.16	$--^{SEA}/\alpha\alpha; \beta^{-28}/\beta^N$	1	4.17
$-\alpha^{3.7}/--^{SEA}$	3	0.27	β^{-29}/β^N	7	1.51	$--^{SEA}/\alpha\alpha; \beta^{-29}/\beta^N$	1	4.17
$\alpha^{WS}/\alpha/-\alpha^{3.7}$	3	0.27	SEA-HPFH	4	0.86	$--^{SEA}/\alpha\alpha; \beta^{5'UTR+43\ to\ +40}/\beta^N$	1	4.17
$\alpha^{IVS-138}/\alpha/\alpha\alpha$	2	0.18	β^{-31}/β^N	3	0.65	$--^{SEA}/\alpha\alpha; \beta^{CD26}/\beta^N$	1	4.17
$-\alpha^{3.7}/-\alpha^{3.7}$	2	0.18	$\beta^{CD14-15}/\beta^N$	3	0.65	$--^{SEA}/\alpha\alpha; \beta^{IVS-II-654}/\beta^N$	1	4.17
$\alpha^{CAP+29}/\alpha/\alpha\alpha$	1	0.09	β^{CAP+8}/β^N	2	0.43	$\alpha^{WS}/\alpha/\alpha\alpha; \beta^{IVS-II-654}/\beta^N$	1	4.17
Hb Agrinio	1	0.09	ChineseGgamma	2	0.43	$\alpha^{CS}/\alpha/\alpha\alpha; \beta^{CD17}/\beta^N$	1	4.17
Hb Cervantes	1	0.09	$\beta^{IVS-II-848}/\beta^N$	2	0.43	$\alpha^{CS}/\alpha/\alpha\alpha; \beta^{IVS-II-654}/\beta^N$	1	4.17
Hb Iberia	1	0.09	$\beta^{5'UTR+43\ to\ +40}/\beta^{IVS-II-654}$	1	0.22	$\alpha^{CD30}/\alpha/\alpha\alpha; \beta^{IVS-II-654}/\beta^N$	1	4.17
$\alpha^{WS}/\alpha/\alpha^{WS}$	1	0.09	β^{-56}/β^N	1	0.22			
$\alpha^{IVS-117}/\alpha/\alpha\alpha$	1	0.09	β^{-72}/β^N	1	0.22			
$\alpha^{CD31}/\alpha/\alpha\alpha$	1	0.09	β^{nt-77}/β^N	1	0.22			
$-\alpha^{3.7}/-\alpha^{4.2}$	1	0.09	$\beta^{CD20-21}/\beta^N$	1	0.22			
$-\alpha^{4.2}/--^{SEA}$	1	0.09	Hb Knossos	1	0.22			
$--^{THAI}/\alpha\alpha$	1	0.09	$\beta^{IVS-II-643}/\beta^N$	1	0.22			
$\alpha^{OS}/\alpha/-\alpha^{3.7}$	1	0.09	β^{polyA}/β^N	1	0.22			
$-\alpha^{4.2}$	1	0.09						
$\alpha^{CS}/\alpha/-\alpha^{4.2}$	1	0.09						
$\alpha^{WS}/\alpha/-\alpha^{4.2}$	1	0.09						
Total	1,109	100	Total	464	100	Total	24	100

Note. * HGVS names of all mutations are shown in Supplementary Table S2 (available in www.besjournal.com).

randomly selected samples, with an estimated rate of 1.99%. They included 34 with $\alpha\alpha/\alpha\alpha^{\text{anti}4.2}$ and 37 with $\alpha\alpha/\alpha\alpha^{\text{anti}3.7}$; in the latter group, one participant had combined $-\alpha^{3.7}$ deletion ($\alpha\alpha^{\text{anti}3.7}/-\alpha^{3.7}$), and two

participants had β -thalassemia. The phenotypic characteristics of all cases are shown in Table 3. (Supplementary Table S4, available in www.besjournal.com).

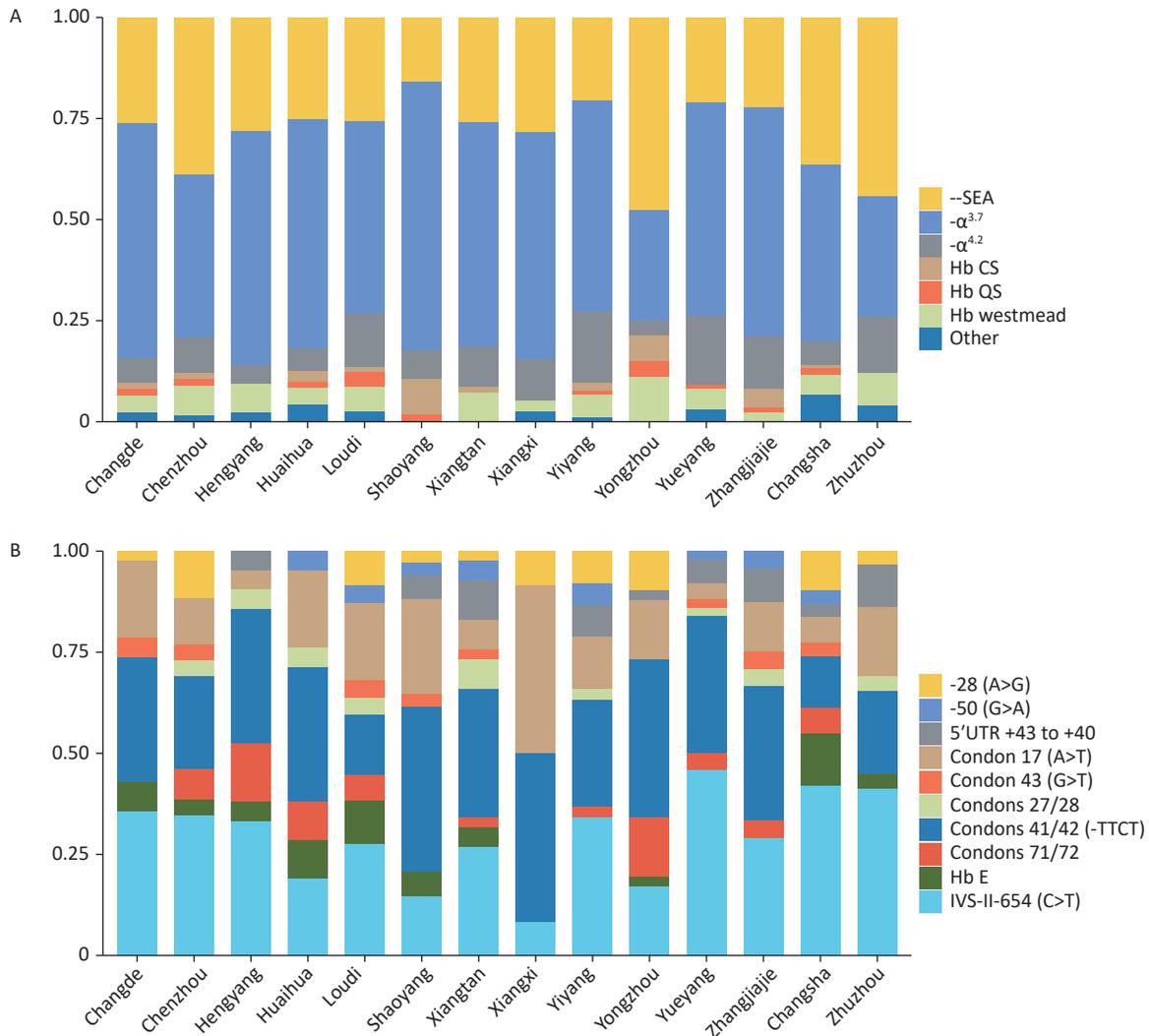


Figure 2. Allele frequencies of thalassemia in different regions of Hunan Province. (A) Allele frequency of α -thalassemia among 14 cities in Hunan Province. (B) Allele frequency of common mutations of β -thalassemia among 14 cities in Hunan Province.

Table 3. Hematological phenotypic analysis of α triplication

Genotype	Number	Hematological phenotype						
		RBC ($\times 10^{12}/L$)	Hb (g/L)	MCV (fl)	MCH (pg)	Hb A (%)	Hb A2 (%)	Hb F (%)
$\alpha\alpha^{\text{anti}3.7}/\alpha\alpha, \beta^N/\beta^N$	34	4.63 \pm 0.53	143.03 \pm 20.13	91.84 \pm 5.49	30.84 \pm 2.26	97.03 \pm 0.64	2.70 \pm 0.25	0.27 \pm 0.60
$\alpha\alpha^{\text{anti}4.2}/\alpha\alpha, \beta^N/\beta^N$	34	4.75 \pm 0.47	146.15 \pm 15.19	91.34 \pm 3.72	30.80 \pm 1.35	97.06 \pm 0.49	2.77 \pm 0.22	0.17 \pm 0.42
$\alpha\alpha^{\text{anti}3.7}/-\alpha^{3.7}, \beta^N/\beta^N$	1	4.33	131	88.2	30.3	97.3	2.4	0.3
$\alpha\alpha^{\text{anti}3.7}/\alpha\alpha, \beta^N/\beta^{\text{CD}14-15}$	1	5.36	109	64	20.3	93.7	5.5	0.8
$\alpha\alpha^{\text{anti}3.7}/\alpha\alpha, \beta^N/\beta^{\text{IVS-II-654}}$	1	5.62	114	67.5	20.2	92.6	4.9	2.5

DISCUSSION

Herein, we conducted a large-scale, large-sample, province-wide study. The overall carrier rate of thalassemia was 7.1%, including 4.83% for α -thalassemia, 2.15% for β -thalassemia, and 0.09% for α - and β -thalassemia (Table 1). These values are lower than the reported carrier rates in the six provinces in South China, Guangdong, Guangxi, Hainan, Yunnan, Jiangxi, and Guizhou^[8,18]. The carrier rates of thalassemia in Hunan and Fujian were similar^[10]. However, the current values were 1.74 times higher than Hunan Province's overall carrier rate first reported in 2007 (4.13%). The carrier rate of thalassemia in Changsha (7.24%) was higher than that in a previous study (4.54%)^[19]. Our study more accurately reflects the carrier rate of thalassemia in various areas, because we performed genetic testing with the NGS method on all recruited individuals, whereas previous studies have used traditional testing methods in individuals with positive hematologic phenotypes.

Our results also indicated a higher carrier rate of thalassemia in southern than northern Hunan Province, in agreement with the general geographical distribution of thalassemia in China. The higher carrier rate in Yongzhou (14.57%) and Chenzhou (10.29%), located in the southern region, was partly due to their geographical proximity to Guangdong and Guangxi; people in this area migrate to and from Guangdong and Guangxi (Table 1). In our study, the carrier rate of thalassemia in the Chenzhou region (10.29%) was close to that in a 2018 study (10.78%)^[15].

Previous research has revealed that, in contrast to $--^{SEA}/\alpha\alpha$, which is most common in Guangxi, Guangdong and Yunnan^[20,21], $-\alpha^{3.7}/\alpha\alpha$ is the most common α -thalassemia genotype in Hunan^[18,22]. In addition to the six common genotypes ($-\alpha^{3.7}/\alpha\alpha$, $--^{SEA}/\alpha\alpha$, $-\alpha^{4.2}/\alpha\alpha$, $\alpha^{WS}/\alpha\alpha$, $\alpha^{CS}/\alpha\alpha$, and $\alpha^{QS}/\alpha\alpha$), we identified 13 rare genotypes of α -globin genes, accounting for 2.43% of all genotypes (27/1,109), including CD108 (ACC>AAC), CAP +29(G>C), Hb Agrinio and Hb Cervantes, which had not previously been reported in China. Among 489 β -thalassemia carriers, 25 genotypes were identified, and the most common mutations were $\beta^{IVS-II-654}/\beta^N$ (28.23%), $\beta^{CD41-42}/\beta^N$ (27.37%), and β^{CD17}/β^N (13.36%). The top three β -thalassemia mutations in Hunan were similar to those in the neighboring Guizhou Province but differed from those in Guangdong, Guangxi, Hainan, and Yunnan^[2].

In contrast to PCR-RDB method, 13 additional

rare genotypes of β -globin genes has been identified by NGS method accounting for 6.68% of all genotypes. β^{-50}/β^N (2.37%) was the most prevalent rare genotype, and CAP+8 (C>T), IVS-II-848 (C>T), -56 (G>C), beta nt-77 (G>C), codon 20/21 (-TGGGA), and Hb Knossos had not previously been reported in China. These results indicated that the type and frequency of thalassemia gene mutations in Hunan Province have clear regional specificity, and that screening and detection of rare gene mutations should be emphasized.

We found 127 cases of abnormal hemoglobin mutation in 25,946 individuals, with a carrier rate of 0.49%, a value higher than previously reported in Guangdong (0.358%) and Fujian (0.26%)^[10,16,19,22-24]. We detected 40 different abnormal hemoglobin variants; this is the first report of simultaneous detection of a wide range of hemoglobin variants in China. Most detected variants were apparently silent and did not affect hemoglobin function and stability, or cause clinical effects. However, these findings enrich the known genetic landscape of hemoglobinopathies in Hunan Province.

The carrier rate of α -globin triplication in Hunan Province was 1.99%, a rate higher than reported previously in five southern provinces^[13,18,20]. We observed no significant abnormalities in the hematological phenotype of the simple triad in our study, in agreement with previous research^[21].

According to many reports, coinheritance of α -globin gene triplication with β -thalassemia mutations may lead to moderate to severe anemia, and some patients may require intermittent blood transfusion therapy^[25]. However, in this study, two cases of heterozygous β -thalassemia combined with the α -triad showed only mild anemia. A related study in Hong Kong, China has also reported two cases with β^0 mutation combined with α triplication showing a phenotype of only mild thalassemia^[26]. An Iranian study of 67 cases with similar genotypes has reported that only four cases had hemoglobin lower than 90 g/L, and no cases required blood transfusion^[27]. These results suggest that the clinical phenotypes of people with α -globin triplication combined with heterozygous β -thalassemia are quite diverse, and therefore, must be fully elucidated in clinical work.

In conclusion, this large-scale, large-sample and province-wide study is the first comprehensive investigation of the spectrum of thalassemia mutations in Hunan Province. Our findings showed the great complexity and diversity of thalassemia gene mutations in the Hunan population. The findings should be important for genetic counseling and the

prevention of severe thalassemia in this region.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

XI Hui, WANG Hua, LIANG Chang Biao, and FANG Jun Qun performed the research. XI Hui, LIU Qin, and XIE Dong Hua designed the research study. WANG Shun Yao, CHEN Shi Ping, and MA Xiao Yuan contributed essential reagents or tools. TANG De Guo, ZENG Chun Yan, WANG Qiong, NIE Xing Hui, PENG Jin Ping, GAO Xiao Ya, WU Hong Liang, ZHANG Hao Qing, QIU Li, FENG Zong Hui, WANG Shu Yuan, ZHOU Shu Xiang, HE Jun, ZHOU Shi Hao, ZHOU Fa Qun, and ZHENG Jun Qing enrolled the participants and collected the blood samples. LIU Qin and ZHENG Zhi Fen analyzed the data. XI Hui and LIU Qin wrote the paper. TANG Wang Lan, ZHOU Xu, and XI Hui managed the budget plan.

ACKNOWLEDGEMENTS

We thank the patients for their willingness to participate in this study. We are grateful to Prof. SUN Zhen Qiu and Prof. TAN Li Hong from Xiangya School of Public Health, Central South University for their guidance.

Received: June 2, 2022;

Accepted: August 19, 2022

REFERENCES

- Yang Z, Cui QX, Zhou WZ, et al. Comparison of gene mutation spectrum of thalassemia in different regions of China and Southeast Asia. *Mol Genet Genomic Med*, 2019; 7, e680.
- Shang X, Xu XM. Update in the genetics of thalassemia: what clinicians need to know. *Best Pract Res Clin Obstet Gynaecol*, 2017; 39, 3–15.
- Taher AT, Weatherall DJ, Cappellini MD. Thalassaemia. *Lancet*, 2018; 391, 155–67.
- Yang Z, Zhou WZ, Cui QX, et al. Gene spectrum analysis of thalassemia for people residing in northern China. *BMC Med Genet*, 2019; 20, 86.
- Li B, Zhang XZ, Yin AH, et al. High prevalence of thalassemia in migrant populations in Guangdong Province, China. *BMC Public Health*, 2014; 14, 905.
- Lai KT, Huang GF, Su L, et al. The prevalence of thalassemia in mainland China: evidence from epidemiological surveys. *Sci Rep*, 2017; 7, 920.
- Bhattacharya D. Asymptotic inference from multi-stage samples. *J Econometr*, 2005; 126, 145–71.
- Lin M, Zhong TY, Chen YG, et al. Molecular epidemiological characterization and health burden of thalassemia in Jiangxi Province, P. R. China. *PLoS One*, 2014; 9, e101505.
- Niu Q, Huang XB, An YF, et al. Genotype of thalassemia in Han Chinese and Tibetans in Sichuan province, China. *J Sichuan Univ (Med Sci Ed)*, 2016; 47, 941–4. (In Chinese)
- Huang HL, Xu LP, Chen MH, et al. Molecular characterization of thalassemia and hemoglobinopathy in Southeastern China. *Sci Rep*, 2019; 9, 3493.
- Zhu YW, Shen N, Wang X, et al. Alpha and beta-Thalassemia mutations in Hubei area of China. *BMC Med Genet*, 2020; 21, 6.
- Tan ASC, Quah TC, Low PS, et al. A rapid and reliable 7-deletion multiplex polymerase chain reaction assay for α -thalassemia. *Blood*, 2001; 98, 250–1.
- He J, Song WH, Yang JL, et al. Next-generation sequencing improves thalassemia carrier screening among premarital adults in a high prevalence population: the Dai nationality, China. *Genet Med*, 2017; 19, 1022–31.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 2009; 25, 1754–60.
- 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*, 2019; 33, e22845.
- Kountouris P, Lederer CW, Fanis P, et al. IthaGenes: an interactive database for haemoglobin variations and epidemiology. *PLoS One*, 2014; 9, e103020.
- Wang W, Ma ESK, Chan AYY, et al. Single-tube multiplex-PCR screen for anti-3.7 and anti-4.2 alpha-globin gene triplications. *Clin Chem*, 2003; 49, 1679–82.
- Shang X, Peng ZY, Ye YH, et al. Rapid targeted next-generation sequencing platform for molecular screening and clinical genotyping in subjects with hemoglobinopathies. *eBioMedicine*, 2017; 23, 150–9.
- He J, Zeng HL, Zhu L, et al. Prevalence and spectrum of thalassaemia in Changsha, Hunan province, China: discussion of an innovative screening strategy. *J Genet*, 2017; 96, 327–32.
- Xie XM, Wu MY, Li DZ. Evidence of selection for the α -globin gene deletions and triplications in a southern Chinese Population. *Hemoglobin*, 2015; 39, 442–4.
- Harteveld CL, Refaldi C, Cassinerio E, et al. Segmental duplications involving the α -globin gene cluster are causing β -thalassemia intermedia phenotypes in β -thalassemia heterozygous patients. *Blood Cells, Mol, Dis*, 2008; 40, 312–6.
- Zhao JH, Li J, Lai QH, et al. Combined use of gap-PCR and next-generation sequencing improves thalassaemia carrier screening among premarital adults in China. *J Clin Pathol*, 2020; 73, 488–92.
- Zhou BY, Wang YX, Shan XS, et al. Molecular spectrum of α - and β -thalassemia among young individuals of marriageable age in Guangdong Province, China. *Biomed Environ Sci*, 2021; 34, 824–9.
- Heng G, Wang YX, Du MX, et al. Effectiveness of using mean corpuscular volume and mean corpuscular hemoglobin for beta-thalassemia carrier screening in the Guangdong population of China. *Biomed Environ Sci*, 2021; 34, 667–71.
- Moosavi SF, Amirian A, Zarbakhsh B, et al. The carrier frequency of α -globin gene triplication in an Iranian population with normal or borderline hematological parameters. *Hemoglobin*, 2011; 35, 323–30.
- Ma SK, Au WY, Chan AY, et al. Clinical phenotype of triplicated α -globin genes and heterozygosity for β 0-thalassemia in Chinese subjects. *Int J Mol Med*, 2001; 8, 171–5.
- Hamid M, Keikhaei B, Galehdari H, et al. Alpha-globin gene triplication and its effect in beta-thalassemia carrier, sickle cell trait, and healthy individual. *eJHaem*, 2021; 2, 366–74.