



## RESEARCH ARTICLE

# Molecular epidemiological and hematological profile of thalassemia in the Dongguan Region of Guangdong Province, Southern China

Qi Peng<sup>1,2,3</sup> | Zhenhong Zhang<sup>4</sup> | Siping Li<sup>1,2,3</sup> | Chuyun Cheng<sup>5</sup> | Wenrui Li<sup>1,2,3</sup> | Chunbao Rao<sup>1,2,3</sup>  | Baimao Zhong<sup>1,3,6</sup> | Xiaomei Lu<sup>1,2,3</sup> 

<sup>1</sup>Department of Medical and Molecular Genetics, Dongguan Institute of Pediatrics, Dongguan, China

<sup>2</sup>Medical Laboratory, Dongguan Children's Hospital, Dongguan, China

<sup>3</sup>Key Laboratory for Children's Genetics and Infectious Diseases of Dongguan City, Dongguan, Guangdong, China

<sup>4</sup>Department of Pediatrics Hematology, Dongguan Children's Hospital, Dongguan, China

<sup>5</sup>Department of Gynecology and Obstetrics, Dongguan Eighth People's Hospital, Dongguan, China

<sup>6</sup>Department of Neonates, Dongguan Children's Hospital, Dongguan, China

## Correspondence

Xiaomei Lu, Department of Medical and Molecular Genetics, Dongguan Institute of Pediatrics, Dongguan, Guangdong, China.  
Email: lxmdgeys@163.com

Baimao Zhong, Department of Neonates, Dongguan Children's Hospital, Dongguan, Guangdong, China.  
Email: zbomao@163.com

## Funding information

The study was supported by the Key Projects of Science and Technology Plan of Dongguan (Project Number: 201750715028452 and 20185071501001628), Guangdong Natural Science Foundation Project (Project Number: 2017A030310588), Guangdong Basic and Applied Basic Research Fund Project (2019A1515010396) and Medical Scientific Research Foundation of Guangdong Province of China (Project Number: A2020261). The funders had no role in the design of the study or in the collection, analysis, and interpretation of data or in writing the manuscript.

## Abstract

**Background:** Thalassemia is a common inherited hematological disease in tropical and subtropical regions. This study aimed to investigate the mutation spectrum of thalassemia in the Dongguan region of southern China and comprehensively analyze hematologic features of thalassemia carriers with various types of globin mutations.

**Methods:** A hematological screening including hematological indices such as mean corpuscular volume (MCV), mean corpuscular hemoglobin content (MCH), and mean corpuscular hemoglobin concentration (MCHC) was conducted in 19 442 people from Dongguan region, Guangdong province of China. Then, 4891 suspected thalassemia carriers were further investigated by genetic analysis of combined NGS and gap-PCR.

**Results:** Totally, 2319 (11.9%) cases were diagnosed as carriers of thalassemia, of which 1483 cases (7.6%) were  $\alpha$ -thalassemia, 741 cases (3.8%) were  $\beta$ -thalassemia, and 95 cases (0.5%) were co-inheritance of  $\alpha$ - and  $\beta$ -thalassemia. In  $\alpha$ -thalassemia carriers, the phenotypic severity increases with the number of nonfunctional  $\alpha$ -globin genes. The patients with  $^{-SEA}/\alpha^{WS}$  genotype have less severe clinical phenotypes than those with other Hb H diseases. As for  $\beta$ -thalassemia, the MCV and MCH in both  $\beta^0$  and  $\beta^+$  carriers are markedly reduced.

**Conclusions:** This is the first comprehensive molecular epidemiological survey and hematological profiling of thalassemia in Dongguan area. This study will be benefit

Qi Peng and Zhenhong Zhang contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Journal of Clinical Laboratory Analysis* Published by Wiley Periodicals LLC

for genetic counseling in the clinic and may help pediatricians to make a correct diagnosis of different types of thalassemia.

#### KEYWORDS

hematological, molecular epidemiological, mutation, thalassemia

## 1 | INTRODUCTION

Thalassemia is one of the most frequently inherited recessive disorders caused by mutations in  $\alpha$ - or  $\beta$ -globin gene clusters.<sup>1,2</sup> It is classified into  $\alpha$ -thalassemia and  $\beta$ -thalassemia based on the defects in the globin genes.<sup>3</sup> The  $\alpha$ -thalassemia is caused by absent or decreased production of the  $\alpha$ -chain of hemoglobin, and the severity varies from an asymptomatic condition to severe anemias depending on the number of alpha genes deleted.<sup>4</sup> Fetus suffers from  $\alpha$ -thalassemia major will die in utero or soon after birth.<sup>5</sup>  $\beta$ -thalassemia is caused by depletion or absence of  $\beta$ -globin synthesis, and patients with  $\beta$ -thalassemia major have severe anemia and serious complications, including liver injury, heart disease, and endocrine dysfunction.<sup>6</sup> To date, there is no effective treatment for thalassemia patients. Severe thalassemia patients depend heavily on blood transfusions to maintain their lives, which brings huge burdens to affected families and society.<sup>7</sup>

The thalassemia is widely distributed in Southeast Asia, the Middle East, Africa, and the Mediterranean region of the world.<sup>3,8</sup> In China, the thalassemia mainly occurs in the southern part of the Yangtze River, especially in Guangdong, Guangxi, and Hainan provinces.<sup>9,10</sup> Previous studies have confirmed that screening and genetic counseling for thalassemia carrier are the most effective ways to reduce the incidence of severe thalassemia in children.<sup>11-13</sup> As populations in different geographical regions have different globin gene mutation spectrums, a comprehensive understanding of the molecular epidemiological characteristics of the incidence and distribution of thalassemia is important for developing suitable preventive strategies in areas with high incidence of thalassemia.

The estimated population in Dongguan area was more than 8 million, and the annual live births are about 140 thousand (<http://www.dg.gov.cn/>). Previous studies reporting the prevalence of thalassemia in this area revealed that the carrier frequency of  $\alpha$ -thalassemia was about 6.18%-7.19%, whereas the frequency of  $\beta$ -thalassemia was about 2.77%-3.51%.<sup>14-16</sup> Given the limitation of sample size and technology availability, these studies may not reflect the current and true prevalence of thalassemia. On the other hand, limited information is available on the hematological characterization of thalassemia carrier with different mutation. This study is the first large-scale epidemiological investigation and full-spectrum analysis of mutations and hematological profile in thalassemia in Dongguan, southern China. The results of this study will provide reference for genetic counseling and preventive measures to reduce cases of the severe thalassemia. In addition, the genotype-phenotype correlation in this study might

help the clinicians to have rapid and accurate referrals in genetic counseling.

## 2 | MATERIALS AND METHODS

### 2.1 | Population samples

A total of 19 442 unrelated subjects for routine health examinations between January 2016 and December 2018 were included in this study. Participants with following cases were excluded from the study: (a) lack of informed consent, (b) consanguinity, and (c) incomplete information. The subjects ranged in age from 2 to 70 years, and the male-to-female ratio was 1.02:1. All of them are Chinese Han people.

This study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Institutional Medical and Ethics Committee of Dongguan Children's Hospital. Written informed consents were obtained from all the studied subjects or their parents.

### 2.2 | Hematological analyses

Peripheral blood samples of 2.5 ml volume were taken from all subjects and collected into EDTA-containing vacutainers. Hematological data were determined according to standard laboratory procedures using an automated hematology analyzer (Beckman Coulter, Krefeld, Germany). If mean corpuscular volume (MCV) <80 fl and/or mean corpuscular hemoglobin (MCH) <27 pg, subjects were considered as suspected carriers of thalassemia. Suspected thalassemia subjects were subjected to further genetic analysis.

### 2.3 | Genetic analysis

The samples were preceded directly to genetic analysis in our study. As hemoglobin electrophoresis has a higher diagnosis failure for thalassemia detection, it is not sensitive enough for the diagnosis of  $\alpha$ -thalassemia carriers. For example, patients with certain genotypes, such as  $-\alpha^{3.7}/\alpha\alpha$ ,  $-\alpha^{4.2}/\alpha\alpha$ ,  $\alpha^{CS}\alpha/\alpha\alpha$ , and  $\alpha^{WS}\alpha/\alpha\alpha$ , present normal values by hemoglobin electrophoresis. The false-negative rate of  $\alpha$ -thalassemia detection was 79.8% by hemoglobin electrophoresis, and the false-negative rate of  $\beta$ -thalassemia detection by hemoglobin electrophoresis was 10.5%.<sup>17</sup> Therefore, we chose target-based

next-generation sequencing (Shenzhen BGI Tech, Guangdong, China), which have features of high throughput and low cost.

Genomic DNA was extracted from blood samples using Blood DNA Kit (Tiangen Biotech, Beijing, China) following the manufacturer's protocol. The DNA concentration and purity were determined using NanoDrop™ 8000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA).

DNA analysis was performed by routine gap-polymerase chain reaction (gap-PCR) combined with target-based next-generation sequencing (Shenzhen BGI Tech, Guangdong, China) described by He *et al.*<sup>17</sup>

In brief, multiplex PCR method was used to amplify three genes (HBA1, HBA2, and HBB) related to  $\alpha$ -thalassemia and  $\beta$ -thalassemia. PCR reactions were carried out in 96-well plates and each sample corresponding to one library. After the preparation of library, the DNA sequences of samples in every library were be supplemented with a linker sequence for sequencing and library identification. Libraries were pooled at equal molar concentration. Sequencing was performed using the paired-end tag (PE100) protocol with Illumina HiSeq2000. A bioinformatics pipeline that integrates multiple in-house programs and open-source softwares for quality controls of raw reads data reads alignment to human reference genome (UCSC build hg19) and duplicates marking, SNVs, and Indels calling, and annotation was used in this study.

Moreover, multiplex gap-PCR was used to detect deletion-type  $\alpha$  thalassemia:  $^{-SEA}$ ,  $^{-\alpha^{3.7}}$ ,  $^{-\alpha^{4.2}}$ ,  $^{-FIL}$ , and  $^{-THAI}$  and  $\beta$  thalassemia deletions: Chinese Ggamma (Agammadeltabeta) 0, SEA-HPFH, and Taiwanese.

All of the  $\alpha$ -thalassemia subjects in this study were sub-classified into three groups: three  $\alpha$ -globin genes active ( $-\alpha/\alpha\alpha$ ), two  $\alpha$ -globin genes active ( $-\alpha/\alpha$ ), and one  $\alpha$ -globin gene active ( $-\alpha$ ). Based on the degree of quantitative reduction in the synthesis of normal  $\beta$ -globin,  $\beta$ -thalassemia mutations are classified into two groups: (a)  $\beta^0$ -thalassemia mutation ( $\beta^0/$ ), which results in the absence of  $\beta$ -globin; (b)  $\beta^+$ -thalassemia mutation ( $\beta^+/$ ), which reduces the output of  $\beta$ -globin.<sup>18</sup>

## 2.4 | Statistical analysis

Data were analyzed using the SPSS version 22 (IBM, Armonk, NY, USA). Results were expressed as mean  $\pm$  SD values, and differences among groups were compared for each hematological index using Student's *t* test or analysis of variance (ANOVA). A *P*-value of .05 or less was considered statistically significant.

## 3 | RESULTS

### 3.1 | The overall population prevalence of thalassemia in Dongguan region

Among total 19 442 subjects that were initially screened for hematology abnormality, 4891 were suspected cases that underwent

further genetic analysis. There are 2319 cases diagnosed as thalassemia carriers, including 1483 cases of  $\alpha$ -thalassemia, 741 cases of  $\beta$ -thalassemia, and 95 cases composite  $\alpha$ - and  $\beta$ -thalassemia. The overall prevalence rate of thalassemia in Dongguan was 11.9%, and the prevalence rates of  $\alpha$ - and  $\beta$ -thalassemia were 7.6% and 3.8%, respectively. In addition, carriers of both  $\alpha$ - and  $\beta$ -thalassemia were detected for the first time in Dongguan, and the rate was 0.5%.

### 3.2 | Genotypes and mutation spectrum of $\alpha$ - and $\beta$ -thalassemia

In our cohort, a total of 71 genotypes were identified (Table 1). We identified 12 different variations with 24 distinct genotypes among 1483  $\alpha$ -thalassemia carriers in this study (Table 1). The  $^{-SEA}/\alpha\alpha$  was the most common  $\alpha$ -thalassemia in Dongguan, accounting for 65.8% of all  $\alpha$ -thalassemia genotypes. The next five most common genotypes were  $^{-\alpha^{3.7}}/\alpha\alpha$ ,  $^{-\alpha^{4.2}}/\alpha\alpha$ ,  $\alpha^{CS}/\alpha\alpha$ ,  $\alpha^{WS}/\alpha\alpha$ , and  $\alpha^{QS}/\alpha\alpha$ , with frequencies of 15.4%, 4.9%, 3.7%, 3.4%, and 2.2%, respectively (Table 1). In total, these six genotypes accounted for 95.4% of all  $\alpha$ -thalassemia genotypes.

Some rare genotypes such as  $\alpha^{codon30}/\alpha\alpha$ ,  $\alpha^{codon31}/\alpha\alpha$ , and  $^{-THAI}/\alpha\alpha$  were also identified in this study. Twelve  $\alpha$ -thalassemia mutations were found when calculating the specific mutation frequencies of all  $\alpha$  mutant chromosomes (allele frequency). The  $^{-SEA}$  deletion mutation was the most frequent, accounting for 66.0% of all  $\alpha$  mutant loci. The other high-frequency mutations were  $^{-\alpha^{3.7}}$ ,  $^{-\alpha^{4.2}}$ ,  $\alpha^{CS}$ , and  $\alpha^{WS}$  with allelic frequencies of 17.0%, 6.0%, 4.1%, and 4.1%, respectively (Table 2).

Among the 741 subjects detected with  $\beta$ -thalassemia, we found 18  $\beta$ -thalassemia mutations and 21 genotypes (Table 1). The most frequent genotypes were  $\beta^{Codons 41/42 (-TTCT)}/\beta^N$ , and  $\beta^{IVS-II-654 (C>T)}/\beta^N$ , with a remarkable proportion of 41.7% and 27.5%, respectively. The other common genotypes were  $\beta^{Codon 17 (A>T)}/\beta^N$ ,  $\beta^{-28 (A>G)}/\beta^N$ , and  $\beta^{Codons 71/72 (+A)}/\beta^N$  with corresponding proportions of 12.3%, 9.2%, and 3.1%. As for allele frequency, the most frequent mutations were Codons 41/42 (-TTCT), IVS-II-654 (C>T), Codon 17 (A>T), and -28 (A>G) with corresponding proportions of 41.9%, 27.5%, 12.6%, and 9.1%, which in total accounting for 91.1% of all  $\beta$  mutant chromosomes. The allele frequencies of thalassemia mutations in the populations of different regions in Southern China were shown in Table 2.

Ninety-five subjects carried both  $\alpha$  and  $\beta$ -globin variations. Among them, 78.9% of genotypes consisted of common deletions of  $\alpha$ -globin gene ( $^{-SEA}/\alpha\alpha$ ,  $^{-\alpha^{3.7}}/\alpha\alpha$ ,  $^{-\alpha^{4.2}}/\alpha\alpha$ ) combined with a  $\beta$ -globin gene point mutation, and composite  $^{-SEA}/\alpha\alpha$  and  $\beta^{Codons 41/42 (-TTCT)}/\beta^N$  was the most common genotype (Table 1).

A total of eleven novel mutations were identified from eleven unrelated subjects in this study, which include two mutations in HBA1 gene, two mutations in HBA2 gene, and seven mutations in HBB gene. The hematological characteristics of eleven subjects with novel mutations were shown in Table 3.

**TABLE 1** Distribution of  $\alpha$ - and  $\beta$ -thalassemia genotypes in Dongguan Region

$\alpha$ -Thalassemia			$\beta$ -Thalassemia			$\alpha$ - + $\beta$ -thalassemia		
Genotype	Cases	%	Genotype	Cases	%	Genotype	Cases	%
$_{-SEA}/\alpha\alpha$	975	65.8	$\beta^{\text{Codons 41/42 (-TTCT)}/\beta^N}$	309	41.7	$_{-SEA}/\alpha\alpha; \beta^{\text{Codons 41/42 (-TTCT)}/\beta^N}$	13	13.7
$-\alpha^{3.7}/\alpha\alpha$	229	15.4	$\beta^{\text{IVS-II-654 (C>T)}/\beta^N}$	204	27.5	$-\alpha^{3.7}/\alpha\alpha; \beta^{\text{Codons 41/42 (-TTCT)}/\beta^N}$	9	9.5
$-\alpha^{4.2}/\alpha\alpha$	73	4.9	$\beta^{\text{Codon 17 (A>T)}/\beta^N}$	91	12.3	$-\alpha^{3.7}/\alpha\alpha; \beta^{\text{IVS-II-654 (C>T)}/\beta^N}$	8	8.4
$\alpha^{\text{CS}}\alpha/\alpha\alpha$	55	3.7	$\beta^{28 (A>G)}/\beta^N$	68	9.2	$_{-SEA}/\alpha\alpha; \beta^{28 (A>G)}/\beta^N$	8	8.4
$\alpha^{\text{WS}}\alpha/\alpha\alpha$	51	3.4	$\beta^{\text{Codons 71/72 (+A)}/\beta^N}$	23	3.1	$_{-SEA}/\alpha\alpha; \beta^{\text{Codon 17 (A>T)}/\beta^N}$	7	7.4
$\alpha^{\text{QS}}\alpha/\alpha\alpha$	33	2.2	$\beta^E/\beta^N$	9	1.2	$\alpha^{\text{WS}}\alpha/\alpha\alpha; \beta^{\text{Codons 41/42 (-TTCT)}/\beta^N}$	7	7.4
$_{-SEA}/-\alpha^{3.7}$	19	1.3	$\beta^{\text{Codon 43 (G>T)}/\beta^N}$	7	0.9	$-\alpha^{3.7}/\alpha\alpha; \beta^{\text{Codon 17 (A>T)}/\beta^N}$	6	6.3
$-\alpha^{4.2}/_{-SEA}$	11	0.7	$\beta^{50 (G>A)}/\beta^N$	5	0.7	$-\alpha^{3.7}/\alpha\alpha; \beta^{28 (A>G)}/\beta^N$	6	6.3
$\alpha^{\text{WS}}\alpha/_{-SEA}$	6	0.4	$\beta^{29 (A>G)}/\beta^N$	5	0.7	$_{-SEA}/\alpha\alpha; \beta^{\text{Codons 71/72 (+A)}/\beta^N}$	5	5.3
$-\alpha^{3.7}/-\alpha^{4.2}$	5	0.3	$\beta^{\text{SEA-HPFH}]/\beta^N}$	3	0.4	$-\alpha^{4.2}/\alpha\alpha; \beta^{\text{Codons 41/42 (-TTCT)}/\beta^N}$	5	5.3
$\alpha^{\text{WS}}\alpha/-\alpha^{3.7}$	4	0.3	$\beta^{\text{Codons 27/28 (+C)}/\beta^N}$	3	0.4	$_{-SEA}/\alpha\alpha; \beta^{\text{IVS-II-654 (C>T)}/\beta^N}$	4	4.2
$\alpha^{\text{CS}}\alpha/_{-SEA}$	4	0.3	$\beta^{\text{Codons 14/15 (+G)}/\beta^N}$	2	0.3	$-\alpha^{4.2}/\alpha\alpha; \beta^{\text{IVS-II-654 (C>T)}/\beta^N}$	2	2.1
$\alpha^{\text{codon 30}}\alpha/\alpha\alpha$	3	0.2	$\beta^{\text{IVS-I-1 (G>T)}/\beta^N}$	2	0.3	$\alpha^{\text{QS}}\alpha/\alpha\alpha; \beta^{\text{Codons 41/42 (-TTCT)}/\beta^N}$	2	2.1
$\alpha^{\text{QS}}\alpha/_{-SEA}$	3	0.2	$\beta^{\text{Codon 17 (A>T)}/\beta^{\text{Codons 41/42 (-TTCT)}}$	2	0.3	$_{-SEA}/\alpha\alpha; \beta^{29 (A>G)}/\beta^N$	1	1.1
$-\alpha^{3.7}/-\alpha^{3.7}$	2	0.1	$\beta^{\text{Chinese del}]/\beta^N}$	2	0.3	$_{-SEA}/\alpha\alpha; \beta^{\text{Cap +1 (A>C) (-AAAC)}/\beta^N}$	1	1.1
$\alpha^{\text{CS}}\alpha/-\alpha^{4.2}$	2	0.1	$\beta^{\text{IVS-II-5 (G>C)}/\beta^N}$	1	0.1	$\alpha^{\text{CS}}\alpha/\alpha\alpha; \beta^{\text{Codons 41/42 (-TTCT)}/\beta^N}$	1	1.1
Hb Amsterdam-A1/ $\alpha\alpha$	1	0.1	$\beta^{\text{Codon 37 (TGG>TAG)}/\beta^N}$	1	0.1	$\alpha^{\text{CS}}\alpha/\alpha\alpha; \beta^{50 (G>A)}/\beta^N$	1	1.1
Initiation codon (-T)/ $\alpha\alpha$	1	0.1	$\beta^{90 (C>T)}/\beta^N$	1	0.1	$\alpha^{\text{CS}}\alpha/\alpha\alpha; \beta^{\text{Codon 17 (A>T)}/\beta^N}$	1	1.1
$\alpha^{\text{codon 31}}\alpha/\alpha\alpha$	1	0.1	$\beta^{87 (C>A)}/\beta^N$	1	0.1	$\alpha^{\text{QS}}\alpha/\alpha\alpha; \beta^{\text{IVS-II-654 (C>T)}/\beta^N}$	1	1.1
$\alpha^{\text{CS}}\alpha/-\alpha^{3.7}$	1	0.1	$\beta^{\text{Codons 41/42 (-TTCT)}/\beta^{\text{IVS-II-654 (C>T)}}$	1	0.1	$\alpha^{\text{QS}}\alpha/\alpha\alpha; \beta^{\text{Codon 17 (A>T)}/\beta^N}$	1	1.1
$\alpha^{\text{WS}}\alpha/-\alpha^{4.2}$	1	0.1	$\beta^{\text{Codon 17 (A>T)}/\beta^{87 (C>A)}}$	1	0.1	$-\alpha^{3.7}/\alpha\alpha; \beta^{\text{Codons 71/72 (+A)}/\beta^N}$	1	1.1
$\alpha^{\text{CS}}\alpha/\alpha^{\text{WS}}\alpha$	1	0.1	Total	741	100.0	$\alpha^{\text{WS}}\alpha/\alpha\alpha; \beta^{\text{Codon 17 (A>T)}/\beta^N}$	1	1.1
HK $\alpha\alpha/_{-SEA}$	1	0.1				$\alpha^{\text{WS}}\alpha/\alpha\alpha; \beta^{28 (A>G)}/\beta^N$	1	1.1
$_{-THAI}/\alpha\alpha$	1	0.1				$\alpha^{\text{WS}}\alpha/\alpha\alpha; \beta^{\text{IVS-II-654 (C>T)}/\beta^N}$	1	1.1
Total	1483	100.00				$\alpha^{\text{WS}}\alpha/\alpha\alpha; \beta^{\text{Codons 71/72 (+A)}/\beta^N}$	1	1.1
						$\alpha^{\text{WS}}\alpha/-\alpha^{3.7}; \beta^{\text{Codon 17 (A>T)}/\beta^N}$	1	1.1
						Total	95	100.0

### 3.3 | Hematological parameters in thalassemia carrier of different genotype

The mean hemoglobin levels, RBC counts, MCV, MCHC, and MCH values of common thalassemia groups are shown in Table 4.

In the cohort with  $\alpha$ -thalassemia, the phenotypic severity increases with the number of nonfunctional  $\alpha$ -globin genes. In our study, the most common clinical condition of  $\alpha$ -thalassemia is  $\alpha$ -thalassemia trait, which loses two  $\alpha$  genes and majority of them is genotype of  $_{-SEA}/\alpha\alpha$ . Hb H disease loses three  $\alpha$  genes ( $-\alpha$  or  $-\alpha^T\alpha$ ) and shows a moderate to severe microcytic hypochromic hemolytic anemia. Among the silent  $\alpha$ -carriers group, the Hb of  $\alpha^{\text{CS}}\alpha/\alpha\alpha$  group was lower than those of  $-\alpha^{3.7}/\alpha\alpha$  and  $\alpha^{\text{WS}}\alpha/\alpha\alpha$  groups ( $P < .05$ ), the RBC of  $\alpha^{\text{CS}}\alpha/\alpha\alpha$  group was lower than those of  $\alpha^{\text{QS}}\alpha/\alpha\alpha$  and  $\alpha^{\text{WS}}\alpha/\alpha\alpha$  groups ( $P < .05$ ), the MCV of  $\alpha^{\text{QS}}\alpha/\alpha\alpha$  group was lower than those of  $-\alpha^{3.7}/\alpha\alpha$  and  $\alpha^{\text{WS}}\alpha/\alpha\alpha$  groups ( $P < .05$ ), and the MCH and MCHC of  $\alpha^{\text{WS}}\alpha/\alpha\alpha$  group were higher than those of

$\alpha^{\text{CS}}\alpha/\alpha\alpha$  and  $\alpha^{\text{QS}}\alpha/\alpha\alpha$  groups ( $P < .05$ ). As shown in Table 4, the patients of  $_{-SEA}/\alpha^{\text{WS}}\alpha$  genotype have a less severe clinical phenotype than patients with other Hb H diseases. The hemoglobin, MCV, and MCH of  $_{-SEA}/\alpha^{\text{WS}}\alpha$  genotype were higher than those average of Hb H diseases ( $P < .05$ ).

The MCV and MCH in both the of  $\beta^0$  and  $\beta^+$  carriers group are markedly reduced. There are no differences among the entire  $\beta^0$  carriers group, but there are significantly differences between  $\beta^{\text{IVS-II-654 (C>T)}/\beta^N}$  and  $\beta^{28 (A>G)}/\beta^N$  group in MCV and MCH.

## 4 | DISCUSSION

Epidemiological data have shown that thalassemia is highly prevalent in Southern China, which seriously affects the lives of children and their families.<sup>9,10</sup> Therefore, early and accurate detection is very important for the diagnosis and management of the disease.

**TABLE 2** Comparison of the allele frequencies of thalassemia mutations in the populations of different regions in Southern China

Mutation	Allele frequencies (%)				
	Dongguan (this study)	Guangdong <sup>14</sup>	Chaoshan <sup>18</sup>	Meizhou <sup>19</sup>	Guangxi <sup>20</sup>
$\alpha$ -thalassemia					
(- <sup>SEA</sup> /) deletion	66.0	48.5	57.6	65.2	44.7
(- $\alpha^{3.7}$ /) deletion	17.0	36.4	23.7	18.0	27.2
(- $\alpha^{4.2}$ /) deletion	6.0	11.1	13.7	8.3	9.2
HBA2:c.427T >C (Hb CS)	4.1	2.1	3.2	4.4	6.9
HBA2:c.369C >G (Hb Westmead)	4.1	-	-	2.5	8.9
HBA2:c.377T >C (Hb QS)	2.3	0.4	1.8	0.7	2.1
Other	0.5	1.5	-	-	1.1
$\beta$ -thalassemia					
Codons 41/42 (-TTCT)	41.9	36.4	30.2	32.4	42.5
IVS-II-654 (C>T)	27.5	28.4	39.6	40.8	8.9
Codon 17 (A>T)	12.6	4.0	7.6	8.1	24.7
-28 (A>G)	9.1	16.7	13.2	10.5	8.6
Codons 71/72 (+A)	3.1	3.5	-	1.0	2.3
Hb E	1.2	5.1	-	1.8	1.2
Codon 43 (G>T)	0.9	3.5	-	0.1	0.1
Other	3.6	2.3	-	5.3	11.7

**TABLE 3** Hematological characteristics of 11 subjects with novel mutations

ID	Sex-age	HGVS	MCV (fL)	MCH (pg)	Hb (g/L)
1	M-3	HBA1:c.260T>C	83.3	24.7	135
2	F-6	HBA1:c.351G>T	86.6	28.1	130
3	F-50	HBA2:c.184A>T	82.6	26	127
4	M-33	HBA2:c.242delT	81.1	27.5	147
5	F-28	HBB:c.209G>C	68.9	22.1	116
6	M-3	HBB:c.*+129T>A	89.1	25.4	110
7	M-6	HBB:c.-113A>G	83.4	26.6	119
8	F-22	HBB:c.268_281delAGTGAGCTGCACTG	76.5	24.9	106
9	M-5	HBB:c.-23A>G	83.8	27.9	134
10	F-32	HBB:c.-113A>G	81.2	27.6	142
11	M-2	HBB:c.-15C>T	94.2	29.2	101

In this study, for the first time we reported the prevalence and molecular characteristics of thalassemias in Dongguan, a city with a large amount of migrants in Southern China. The results showed that there was a high prevalence of thalassemia in this area, with overall frequency of 11.9%. The overall frequency of thalassemia was similar to the average level of Guangdong province (11.07%), but lower than that of Guangxi Region (24.51%).<sup>10</sup>

In comparison with the previous reports on the same region,<sup>14-16</sup> our data showed higher frequencies of  $\alpha$ -thalassemia carriers and  $\beta$ -thalassemia carriers. We speculated that it was

probably resulted from the combination of rapid increase of population mobility and the coverage of pathogenic loci by the latest NGS technology which is more sensitive and comprehensive than the traditional genotyping analysis. Thus, NGS could be a better large-scale screening method, especially in high prevalence regions. The frequency order of common  $\beta$ -thalassemia mutation types was similar to that reported in the previous studies.<sup>14-16</sup> However, there are differences in that of  $\alpha$ -thalassemia. Our study showed the frequency of  $\alpha^{CS}\alpha/\alpha\alpha$  was higher than  $\alpha^{WS}\alpha/\alpha\alpha$  and  $\alpha^{QS}\alpha/\alpha\alpha$ , which was similar to the results by He.<sup>15</sup> Study

**TABLE 4** Hematological parameters in thalassemia carrier of common genotype

Genotype	Case number	Hb (g/L)	RBC ( $\times 10^{12}/L$ )	MCV (fl)	MCHC (g/L)	MCH (pg)
$\alpha$ -thalassemia						
Silent $\alpha$ -carriers						
$-\alpha^{3.7}/\alpha\alpha$	229	120.4 $\pm$ 19.4	4.8 $\pm$ 0.6	78.9 $\pm$ 6.2	318.5 $\pm$ 17.0	25.1 $\pm$ 2.4
$-\alpha^{4.2}/\alpha\alpha$	73	117.5 $\pm$ 18.9	4.7 $\pm$ 0.5	77.5 $\pm$ 6.9	320.2 $\pm$ 18.9	24.8 $\pm$ 2.6
$\alpha^{CS}\alpha/\alpha\alpha$	55	109.4 $\pm$ 16.8 <sup>a</sup>	4.6 $\pm$ 0.6 <sup>b</sup>	76.5 $\pm$ 7.3	310.7 $\pm$ 18.2	23.8 $\pm$ 2.9
$\alpha^{WS}\alpha/\alpha\alpha$	51	123.6 $\pm$ 17.5	4.9 $\pm$ 0.6	78.7 $\pm$ 8.5	323.6 $\pm$ 15.5 <sup>d</sup>	25.5 $\pm$ 3.3 <sup>d</sup>
$\alpha^{QS}\alpha/\alpha\alpha$	33	114.2 $\pm$ 22.2	5.0 $\pm$ 0.6	75.0 $\pm$ 7.8 <sup>c</sup>	305.2 $\pm$ 26.7	23.0 $\pm$ 3.3
$\alpha$ -thalassemia trait						
$-\text{SEA}/\alpha\alpha$	975	111.1 $\pm$ 15.3	5.4 $\pm$ 0.7	66.0 $\pm$ 5.6	313.1 $\pm$ 20.0	20.7 $\pm$ 2.0
Hb H disease						
$-\text{SEA}/-\alpha^{3.7}$	19	90.4 $\pm$ 13.9	5.4 $\pm$ 0.8	58.0 $\pm$ 7.2	293.7 $\pm$ 23.4	17.0 $\pm$ 2.2
$-\text{SEA}/-\alpha^{4.2}$	11	86.1 $\pm$ 9.3	5.0 $\pm$ 0.7	54.5 $\pm$ 14.4	298 $\pm$ 22.3	17.7 $\pm$ 2.3
$-\text{SEA}/\alpha^{WS}\alpha$	6	103.8 $\pm$ 13.1 <sup>e</sup>	5.3 $\pm$ 0.7	63.0 $\pm$ 6.0	309.5 $\pm$ 11.5	19.5 $\pm$ 1.4 <sup>e</sup>
$-\text{SEA}/\alpha^{CS}\alpha$	4	81.8 $\pm$ 20.0	4.7 $\pm$ 1.1	62.6 $\pm$ 6.0	276.8 $\pm$ 13.7	17.3 $\pm$ 1.1
$\beta$ -thalassemia						
$\beta^0$ carriers						
$\beta^{(e)}/\beta^N$	309	101.9 $\pm$ 14.2	5.3 $\pm$ 0.8	61.0 $\pm$ 5.9	317.3 $\pm$ 16.1	19.4 $\pm$ 2.1
$\beta^{(e)}/\beta^N$	91	102.3 $\pm$ 14.6	5.3 $\pm$ 0.9	61.3 $\pm$ 6.6	320.0 $\pm$ 18.5	19.6 $\pm$ 2.6
$\beta^{(e)}/\beta^N$	23	95.8 $\pm$ 17.2	5.0 $\pm$ 0.7	61.2 $\pm$ 4.2	310.8 $\pm$ 19.6	19.0 $\pm$ 1.5
$\beta^+$ carriers						
$\beta^{IVS-II-654 (C>T)}/\beta^N$	204	98.8 $\pm$ 15.1	5.3 $\pm$ 0.8	62.3 $\pm$ 7.5	313.8 $\pm$ 18.1	19.5 $\pm$ 2.4
$\beta^{-28 (A>G)}/\beta^N$	68	108.3 $\pm$ 16.1 <sup>f</sup>	5.1 $\pm$ 0.8	68.9 $\pm$ 5.9 <sup>f</sup>	312.9 $\pm$ 19.0	21.6 $\pm$ 2.4 <sup>f</sup>

Abbreviations: Hb, hemoglobin; MCH, mean corpuscular Hb; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RBC, red blood cell count.

<sup>a</sup>Among the silent  $\alpha$ -carriers group, the Hb of  $\alpha^{CS}\alpha/\alpha\alpha$  group was lower than those of  $-\alpha^{3.7}/\alpha\alpha$  and  $\alpha^{WS}\alpha/\alpha\alpha$  groups ( $P < .05$ ).

<sup>b</sup>Among the silent  $\alpha$ -carriers group, the RBC of  $\alpha^{CS}\alpha/\alpha\alpha$  group was lower than those of  $\alpha^{QS}\alpha/\alpha\alpha$  and  $\alpha^{WS}\alpha/\alpha\alpha$  groups ( $P < .05$ ).

<sup>c</sup>Among the silent  $\alpha$ -carriers group, the MCV of  $\alpha^{QS}\alpha/\alpha\alpha$  group was lower than those of  $-\alpha^{3.7}/\alpha\alpha$  and  $\alpha^{WS}\alpha/\alpha\alpha$  groups ( $P < .05$ ).

<sup>d</sup>Among the silent  $\alpha$ -carriers group, the MCH and MCHC of  $\alpha^{WS}\alpha/\alpha\alpha$  group were higher than those of  $\alpha^{CS}\alpha/\alpha\alpha$  and  $\alpha^{QS}\alpha/\alpha\alpha$  groups ( $P < .05$ ).

<sup>e</sup>The Hb and MCH of  $-\text{SEA}/\alpha^{WS}\alpha$  genotype were higher than that of other Hb H diseases ( $P < .05$ ).

<sup>f</sup>Among the  $\beta^+$  carriers groups, there are significantly differences between  $\beta^{IVS-II-654 (C>T)}/\beta^N$  and  $\beta^{-28 (A>G)}/\beta^N$  group.

by Zhong<sup>14</sup> showed  $\alpha^{QS}\alpha/\alpha\alpha$  has an obvious higher frequency than those of  $\alpha^{CS}\alpha/\alpha\alpha$  and  $\alpha^{WS}\alpha/\alpha\alpha$ . This is likely due to the difference in sample size apart from the rapid change in population composition and improved sensitivity in genetic analysis.

Similar to other southern regions of China, the most frequent  $\alpha$ -thalassemia mutations in Dongguan region are  $-\text{SEA}$ ,  $-\alpha^{3.7}$ , and  $-\alpha^{4.2}$ . However, the constituent ratio of mutations in  $\alpha$ -thalassemia is different from the average level of Guangdong province, with the proportion of  $-\text{SEA}$ ,  $-\alpha^{3.7}$ , and  $-\alpha^{4.2}$  in 66.0% vs 48.5%, 17% vs 36.5%, and 6.0% vs 11.0%, respectively.<sup>19</sup> This might be due to the fact that more than 2/3 of Dongguan's populations are immigrants from other areas, such as Guangxi, Jiangxi, and Hunan provinces. The migrants from outside could affect the prevalence in many cities.<sup>20</sup>

As for  $\beta$ -thalassemia genotypes,  $\beta^{\text{Codons 41/42(-TTCT)}/\beta^N}$  and  $\beta^{IVS-II-654 (C>T)}/\beta^N$  were the most two frequent  $\beta$ -thalassemia subtypes, accounting for 69.4% of the genotypes. The frequency ranking order of the two major mutations was consistent with the average level of Guangdong,<sup>19</sup> but different from Chaoshan region in eastern part of Guangdong

province<sup>21</sup> and Meizhou region in northeast of Guangdong province.<sup>22</sup> This may be attributed to the fact that there is a distinct population composition in these areas. People residing in Chaoshan speak a unique dialect and have a distinct lifestyle and are defined as Chaoshanese, while Meizhou area has one of the largest Hakka communities in the world.

Xiong<sup>10</sup> thought  $\alpha^{WS}\alpha/\alpha\alpha$  genotype cause  $\alpha$ -thalassemia silent and  $\alpha^{CS}$  and  $\alpha^{QS}$  genotype lead to  $\alpha$ -thalassemia trait. Traditionally, the patients of  $-\alpha^{CS}\alpha$  genotype, which is also known as Hb H Constant Spring (HCS), have more severe clinical symptoms than patients with other Hb H diseases.<sup>23</sup> This difference was not observed in our cohort, likely explained by the limited sample size. In the future study, we will increase the sample size to explore the genotype-phenotype correlation of thalassemia.

Meanwhile, ninety-five subjects were identified with both  $\alpha$ - and  $\beta$ -globin variations in this study. In this case, the synthesis of both chains has been impaired, and the individuals manifested significantly improved phenotypes. However, the diagnosis of double heterozygote state is important. Unlike typical  $\beta$ -thalassemia

carriers, these individuals will be at risk of having offspring with severe  $\alpha$ -thalassemia if their partners are also  $\alpha$ -thalassemia heterozygotes.

## 5 | CONCLUSIONS

We analyzed the genetic profiles and hematological phenotypes of thalassemia in Dongguan, southern China. Based on this study, preventive measures and strategies could be formulated according to the thalassemia genotyping in order to reduce the incidence of severe thalassemia in children. Our findings will provide valuable reference for genetic counseling and prevention of severe thalassemia in this area. The limitation of our study is the possible missed-out carriers during primary screening of routine blood test, which will be addressed in the future study.

## ACKNOWLEDGMENTS

We thank all the participants for their cooperation during this work.

## CONFLICTS OF INTEREST

We declare that we have no conflicts of interest.

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study was approved by the Ethics Committee of Dongguan Children's Hospital in agreement with the Declaration of Helsinki. Written informed consent was obtained from the guardians of the study subject.

## DATA AVAILABILITY STATEMENT

All relevant data are included in the manuscript. The datasets used and/or analyzed during the current study are available from the corresponding author upon request.

## ORCID

Chunbao Rao  <https://orcid.org/0000-0002-1065-3158>

Xiaomei Lu  <https://orcid.org/0000-0001-5454-1548>

## REFERENCES

- Weatherall DJ. Thalassemia as a global health problem: recent progress toward its control in the developing countries. *Ann New York Acad Sci.* 2010;1202(1):17-23. <https://doi.org/10.1111/j.1749-6632.2010.05546.x>
- Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ.* 2008;86:480-487.
- Freedman WL. Alpha and beta thalassemia and pregnancy. *Clin Obstet Gynecol.* 1969;12:115-122.
- Taher AT, Weatherall DJ, Cappellini MD. Thalassaemia. *Lancet.* 2018;391:155-167.
- Lai K, Huang G, Su L, He Y. The prevalence of thalassemia in mainland China: evidence from epidemiological surveys. *Sci Rep.* 2017;7: <https://doi.org/10.1038/s41598-017-00967-2>
- Origa R.  $\beta$ -Thalassemia. *Genet Med.* 2017;19:609-619.
- Suchdev PS, Ruth LJ, Earley M, Macharia A, Williams TN. The burden and consequences of inherited blood disorders among young children in western Kenya. *Matern Child Nutr.* 2014;10:135-144.
- Weatherall DJ, Clegg JB. Inherited haemoglobin disorders: an increasing global health problem. *Bull World Health Organ.* 2001;79:704-712.
- Yin A, Li B, Luo M, et al. The prevalence and molecular spectrum of  $\alpha$ - and  $\beta$ -globin gene mutations in 14,332 families of Guangdong Province, China. *PLoS One.* 2014;9:e89855.
- Xiong F, Sun M, Zhang X, et al. Molecular epidemiological survey of haemoglobinopathies in the Guangxi Zhuang autonomous region of southern China. *Clin Genet.* 2010;78:139-148.
- Cao A, Rosatelli MC, Monni G, Galanello R. Screening for thalassemia: a model of success. *Obstet Gynecol Clin North Am.* 2002;29:305-328.
- Cao A, Cristina Rosatelli M, Galanello R. Control of  $\beta$ -thalassaemia by carrier screening, genetic counselling and prenatal diagnosis: the Sardinian experience. *Ciba Found Symp.* 2007;197:137-155.
- Cao A, Kan YW. The prevention of thalassemia. *Cold Spring Harb Perspect Med.* 2013;3: <https://doi.org/10.1101/cshperspect.a011775>
- Zhong M, Jiang B, Liu H, et al. Analysis on screening of thalassemia among 26 906 peoples reproductive age in Dongguan area. *Chinese J Birth Health Hered.* 2015;23(5):25-26.
- He Y, Lou JW, Tang L, et al. Prenatal screening and prenatal diagnosis of thalassemia in Dongguan. *Chin J Clin.* 2013;7(16):7597-7598.
- Jiang B, Yao QY, Wei SS, et al. Analysis on thalassaemia screening results of couples during pre-pregnancy health. *Maternal Child Health Care China.* 2016;31(7):1475-1477.
- 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.* 2017;19:1022-1031.
- Shang X, Xu X. Update in the genetics of thalassemia: what clinicians need to know. *Best Pract Res: Clin Obstetr Gynaecol.* 2017;39:3-15.
- 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;57:517-522.
- Li B, Zhang X-Z, Yin A-H, et al. High prevalence of thalassemia in migrant populations in Guangdong Province, China. *BMC Public Health.* 2014;14: <https://doi.org/10.1186/1471-2458-14-905>
- Zheng X, Lin M, Yang H, et al. Molecular epidemiological characterization and health burden of thalassemias in the Chaoshan Region, People's Republic of China. *Hemoglobin.* 2016;40:138-142.
- Zhao P, Weng R, Wu H. Molecular spectrum of  $\alpha$ - and  $\beta$ -thalassaemia mutations in a large ethnic Hakka population in Southern China. *Hemoglobin.* 2018;42:117-121.
- Lal A, Goldrich ML, Haines DA, Azimi M, Singer ST, Vichinsky EP. Heterogeneity of hemoglobin H disease in childhood. *N Engl J Med.* 2011;364:710-718.

**How to cite this article:** Peng Q, Zhang Z, Li S, et al. Molecular epidemiological and hematological profile of thalassemia in the Dongguan Region of Guangdong Province, Southern China. *J. Clin. Lab. Anal.* 2021;35:e23596. <https://doi.org/10.1002/jcla.23596>