

Comparison of gene mutation spectrum of thalassemia in different regions of China and Southeast Asia

Zhuo Yang^{1*} | Quexuan Cui^{2*} | Wenzhe Zhou^{2*}  | Ling Qiu¹ | Bing Han² 

¹Department of Clinical Laboratory, Peking Union Medical College, Chinese Academy of Medical Science, Beijing, China

²Department of Hematology, Peking Union Medical College, Chinese Academy of Medical Science, Beijing, China

Correspondence

Bing Han, Department of Hematology, Peking Union Medical College, Chinese Academy of Medical Science, Beijing, China.

Email: hanbing_li@sina.com

Funding information

Chinese Academy of Medical Sciences innovation fund for medical sciences, Grant/Award Number: 2016-I2M-3-004; MANDARIN study from Novartis, Grant/Award Number: C1CL670AAU05; National Natural Science Foundation of China, Grant/Award Number: 81401725

Abstract

Background: Thalassemia is a common genetic disorder. High prevalence of thalassemia is found in South China, Southeast Asia, India, the Middle East, and the Mediterranean regions. Thalassemia was thought to exist only in southern China, but an increasing number of cases from northern China have been recently reported.

Methods: During 2012 to 2017, suspected thalassemia people were detected for common α - and β -thalassemia mutations by gap-Polymerase Chain Reaction (PCR) and reverse dot blot (RDB) analysis in Peking Union Medical College Hospital. One thousand and fifty-nine people with thalassemia mutations were analyzed retrospectively. We picked mutated individuals who originally came from northern areas, and conducted telephone follow-up survey in order to collect their ancestral information. Besides, we used “thalassemia”, “mutation”, and “Southeast Asian countries” as keywords to search the relevant studies in PubMed and Embase databases.

Results: All carriers included in our study were resided in northern China. Among them, 17.3% were native northerners and 82.7% were immigrants from southern China. Although substantial difference was found in α - and β -thalassemia ratio and detailed spectrum of α - and β -globin mutation spectrum between our data and data obtained from a previous meta-analysis literature focused on southern China, the most common gene mutations were the same. Similar β -thalassemia mutation spectrum was found among Thai, Malaysian Chinese, and Guangdong people, however, no other similarities in gene profile were found between Chinese and other ethnic groups in Southeast Asia.

Conclusion: Chinese people in different areas had similar gene mutation, whereas they had significantly different mutation spectrums from other ethnic groups in Southeast Asia.

KEYWORDS

China, comparison, gene mutation, Southeast Asia, thalassemia

*These authors contributed equally to this paper.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals, Inc.

1 | INTRODUCTION

Thalassemia is a common monogenic disease including two major types, α - and β -thalassemia, according to mutations or deletions in α - and β -globin genes (HBA1, HBA2 and HBB; OMIM: 141800, 141850 and 141900), respectively. The mutated or deleted genes produce impaired globin protein subunits and affect oxygen transportation. As estimated by a previous research, at least 20% of the world population carry α^+ -thalassemia, and 5.2% of the population carry a significant variant of β -thalassemia and α^0 -thalassemia, who typically have microcytosis and may have mild anemia as well (Modell & Darlison, 2008). However, the distribution of thalassemia is not homogeneous. High prevalence of thalassemia was mainly reported in southern China, Southeast Asia, India, the Middle East, Africa and Mediterranean region (Li, 2017). As reported by Kuesap, Chaijaroenkul, Rungsahirunrat, Pongjantharasatien, and Na-Bangchang (2015), thalassemia has protective effects on the development of hyperparasitemia and severe anemia in malaria patients, which may justify why the prevalence of thalassemia is high in malarious areas.

Thalassemia is also common in China. Nearly all patients with thalassemia come from the southern China, however, an increasing number of sporadic cases who belong to northern China have been reported (Gan, Yang, Chen, Zhang, & Cui, 2014). We have found a few cases of thalassemia patients whose ancestral home was north China as well. Although immigration from south to north always happens, however, *de novo* mutations can also happen in people who originated in the northern area. The actual rate of prevalence of thalassemia in northern China is not currently clear, however, we believe that adequate knowledge about thalassemia symptoms and gene distribution can help clinicians screen and diagnose thalassemia candidates (Lai, Huang, Li, & He, 2017). Apart from geographical factors, ethnic difference may be another important factor for the gene diversity.

In this study, we compared mutation spectrum of α - and β -thalassemia in indigenous northern people and immigrants from southern China, and also compared the mutation spectrum of immigrants from southern China with the largest meta-analysis in southern China, and the data belong to Southeast Asian countries. The results may help understand the similarities and differences among people who came from different areas and ethnic groups.

2 | METHODS

2.1 | Genetic testing

Venous blood samples from 2,136 cases who were suspected of thalassemia were collected and tested in Peking Union Medical College Hospital (Beijing, China), the largest place for

thalassemia gene analysis in northern China, from 2012 to 2017. Blood tests included blood cell analysis, biochemical parameters, hemoglobin electrophoresis, and thalassemia mutation detection using gap-PCR and reverse dot blot (RDB) analysis.

In order to test the mutation or deletion of thalassemia, we collected 5 ml venous blood from each patient. We used QIAamp DSP DNA Blood Mini Kit to extract DNA from blood. Thirty-five cycles of PCR were performed with index primers. Three common HBA1 deletions were tested by gap-PCR compared to positive controls by electrophoresis through 1.2% agarose gels. Three common HBA2 mutations and 17 mutations of HBB were tested through RDB analysis. The amplification products were denatured, spotted onto nylon membranes and then hybridized with specific oligonucleotide probes. All the deletions or mutations we tested are shown in Table 1.

2.1.1 | Ancestral home

Using hospital's administration data, 523 individuals who came from northern China, were selected from 1,372 α - and β -thalassemia carriers. Telephone follow-up survey was conducted on the basis of their ancestral information, including a household registration and prevalence of their ancestors in three generations of family, to confirm whether these northerners have southern lineage. Carriers with no follow-up results or searchable identity cards were excluded, and 1,059 cases with ancestral information were involved in the final analysis of results.

2.2 | Literature review

We used “thalassemia”, “mutation”, and “Southeast Asian countries” as keywords to find out relevant studies in PubMed and EMBASE updated on 27 December 2017. We considered Southeast Asia which consisted of these countries: Philippines, Malaysia, East Timor, Indonesia, Brunei, Singapore, Cambodia, Laos, Myanmar, Thailand, and Vietnam. The included studies should present in English and cover the majority of the nation. Furthermore, studies on special populations (e.g. minor ethnic groups) were excluded.

2.3 | Statistical analysis

Statistical analysis was conducted using SPSS 22.0 software (IBM, NY). The ratio of α - and β -thalassemia alleles was calculated. In addition, data from included studies and our study were analyzed by Fisher's exact test and Pearson's chi-squared test. $p < 0.05$ was statistically considered significant.

2.4 | Ethical statement

The Ethics Committee of Peking Union Medical College Hospital approved the study (Reference No. S-K438). All participants provided written informed consent as well.

TABLE 1 Target mutations in this study

	Gene	Mutation	HGVS name
α-globin deletions	-α ^{3.7}	-α ^{3.7}	NG_000006.1:g.34164_37967del13804
	-α ^{4.2}	-α ^{4.2}	HGVS not attributable
	-- _{SEA}	--(SEA)	NG_000006.1:g.26264_45564del119301
α-globin mutations	α ^{CS}	Hb Constant Spring (Hb CS)	HBA2:c.427T>C
	α ^{QS}	Hb Quong Sze	HBA2:c.377T>C
	α ^{WS}	Hb Westmead	HBA2:c.369C>G
β-globin mutations	CD41/42	Codons 41/42 (-TTCT)	HBB:c.124_127delTTCT
	CD43	Codon 43 (G->T)	HBB:c.130G>T
	IVS-II-654	IVS-II-654 (C->T)	HBB:c.316-197C>T
	-28	-28 (A->C)	HBB:c.-78A>C
	-29	-29 (A->G)	HBB:c.-79A>G
	-30	-30 (T->C)	HBB:c.-80T>C
	-32	-32 (C->A)	HBB:c.-82C A
	CD71/72	Codons 71/72 (+A)	HBB:c.216_217insA
	βE	Hb E	HBB:c.79G>A
	17	Codon 17 (A->T)	HBB:c.52A>T
	CD14/15	Codons 14/15 (+G)	HBB:c.45_46insG ^a
	CD31	Codon 31 (-C)	HBB:c.94delC
	CD27/28	Codons 27/28 (+C)	HBB:c.84_85insC
	IVS-I-1	1. IVS-I-1 (G->A) 2. IVS-I-1 (G->T)	1. HBB:c.92+1G>A 2. HBB:c.92+1G>T
	IVS-I-5	IVS-I-5 (G->C)	HBB:c.92+5G>C
	Cap+1	1. CAP+1 (A->C) 2. 5'UTR; +43 to +40 (-AAAC) beta+	HBB:c.-50A>C HBB:c.-11_-8delAAAC
	Initiation Condon	Initiation codon ATG->AGG	HBB:c.2T>G

3 | RESULTS

3.1 | Comparison of our thalassemia mutation spectrum in northern and southern China

Data of 1,059 individuals who had thalassemia gene abnormalities were collected during November 2012 to July 2017. Among them, 33.9% (359/1,059) were α-thalassemia carriers or patients, with 330 carrying one deletion or mutation, and 29 carrying two deletions or mutations; 64.5% (683/1,059) were β-thalassemia carriers or patients, with 682 carrying one mutation, and 1 carrying two mutations; 1.6% (17/1,059) were both α- and β-thalassemia carriers or patients. All these people had settled in northern China for more than 3 years.

The ratio of the most common deletions and mutations of α-thalassemia: --_{SEA}, -α^{3.7}, -α^{4.2}, α^{CS}, α^{WS} and α^{QS}; was 74.8%, 16.1%, 16.1%, 1.7%, 1.2%, and 0.7%, respectively, and the most common five gene mutations of β-thalassemia

were IVS-II-654 (36.8%), CD41-42 (28.3%), CD17 (25.2%), -28 (3.0%), and CD71-72 (2.8%), respectively.

People were classified as indigenous northern population (17.3%) and southern immigrants (82.7%), depending on their ancestral home (Table 2). Since there was no significant difference in individual gene distribution between the two groups either in α-thalassemia ($p = 0.221$) or β-thalassemia ($p = 0.979$), we considered people from northern China and southern China together as a whole in comparison with people of Chinese meta-analysis and Southeast Asia.

3.2 | Comparison of our thalassemia mutation spectrum data and a former meta-analysis in the Chinese population

To date, the largest meta-analysis of thalassemia mutation spectrum in China has been conducted by Lai et al. (Sayani & Kwiatkowski, 2015). Most of the data are obtained from

TABLE 2 Frequency and distribution of α - and β -thalassemia mutations

Mutation	Northern population		Immigrants from southern region		Total	
	N	%	N	%	N	%
α -thalassemia						
Total	49	100.0	356	100.0	405	100.0
-- ^{SEA}	35	71.4	268	75.2	303	74.8
- α ^{3.7}	12	24.5	53	14.9	65	16.1
- α ^{4.2}	2	4.1	20	5.6	22	5.4
α ^{CS}	0	0	7	2.0	7	1.7
α ^{QS}	0	0	5	1.4	5	1.2
α ^{WS}	0	0	3	0.9	3	0.7
χ^2 test	$\chi^2=2.582, p = 0.275$					
β -thalassemia						
Total	131	100.0	541	100.0	672	100.0
IVS-II-654 (C-T)	52	39.7	195	36.0	247	36.8
CD41/42 (-TTCT)	33	25.2	157	29.0	190	28.3
CD17 (A-T)	35	26.7	134	24.8	169	25.3
-28 (A-G)	2	1.5	18	3.3	20	3.0
CD71/72 (+A)	3	2.3	16	3.0	19	2.8
CD27/28 (+C)	1	0.8	6	1.1	7	1.0
CD43 (G-T)	1	0.8	6	1.1	7	1.0
β E (GAG-AAG)	2	1.5	4	0.7	6	0.9
-29 (A-G)	1	0.8	3	0.6	4	0.6
CD14/15 (+G)	0	0	1	0.2	1	0.1
IVS-I-1 (G-T/G-A)	0	0	1	0.2	1	0.1
IVS-I-5 (G-C)	1	0.8	0	0	1	0.1
χ^2 test	$\chi^2 = 2.409, p = 0.661$					

people who settled in the southern China. Our data is the first one about people in northern China.

For α -thalassemia, although the detailed gene compositions showed significant difference ($p < 0.001$) between our data and Lai et al.'s data, the first three most common mutations, which accounted for nearly 90% of patients, were the same: --^{SEA} ranked first, - α ^{3.7} ranked second and - α ^{4.2} ranked third in both groups, respectively (Table 3). The results of patients originated from southern China and the previous meta-analysis were significantly different ($p < 0.001$). However, while comparing indigenous northern population and the previous meta-analysis, we found that they were not significantly different ($p = 0.065$).

For β -thalassemia, IVS-II-654 was the most frequent allele in our study, while CD41/42 gene ranked first in Lai et al.'s study (Table 3). Nevertheless, similar to α -thalassemia, the first three common mutations were IVS-II-654, CD41/42, and CD17 in both studies, which accounted for more than 70% of the cases (the five most frequent alleles of β -thalassemia were listed, and other less frequent alleles were

classified as "other"). Analysis of gene constitutions showed significant difference between the overall population of our study and the Lai et al.'s study ($p < 0.001$). Besides, we compared meta-analysis result with our northern ($p < 0.001$) and southern ($p < 0.001$) subgroups, and the result observed showed that either of them had significant difference.

3.3 | Comparison of our data with data obtained from different Chinese provinces in the literature

We found out all the origin papers from Lai et al.'s meta-analysis (Cai et al., 2002; Chen, Chen, Xia, & Qin, 2004; Li, Lan, & Luo, 2009; Pan, Long, & Li, 2007; Qiu, Chen, & Zhang, 2009; Xiong et al., 2010; Xu et al., 2013; Yao, Yu, et al., 2013; Yao, Zhang, et al., 2013; Yin et al., 2014; Zeng, Chen, Chen, Zhong, & Qiu, 2014; Zhang, Wang, & Gao, 2010). Among them, nine studies involved α -thalassemia and seven involved β -thalassemia gene distribution. We compared allele component between immigrants from southern China in

TABLE 3 The thalassemia gene distribution in our study and the literature among Chinese people

	Our data	Meta-analysis conducted by Lai et al. (Sayani & Kwiatkowski, 2015)
α -thalassemia		
<i>N</i>	405	5,806
--SEA	303 (74.8%)	2,958 (51.0%)
$-\alpha^{3.7}$	65 (16.1%)	1598 (27.5%)
$-\alpha^{4.2}$	22 (5.4%)	554 (9.5%)
α^{CS}	7 (1.7%)	249 (4.3%)
α^{QS}	5 (1.2%)	92 (1.6%)
α^{WS}	3 (0.7%)	355 (6.1%)
<i>p</i> value	<0.001	
β -thalassemia		
<i>n</i>	701	2,403
IVS-II-654	258 (36.8%)	448 (18.6%)
CD41/42	204 (29.1%)	931 (38.7%)
CD17	176 (25.1%)	387 (16.1%)
CD71/72	19 (2.7%)	83 (3.5%)
-28	17 (2.4%)	262 (10.9%)
βE	4 (0.6%)	68 (2.8%)
Other	23 (3.3%)	224 (9.3%)
<i>p</i> value	<0.001	

our group and four provinces in southern China that were included in Lai et al.'s meta-analysis (Table 4). For α -thalassemia, our data were similar to that of Fujian province, while that were different from provinces of Guangdong, Guangxi, and Chongqing, which reflected relatively higher ratio of $-\alpha^{3.7}$. For β -thalassemia, mutation spectrum was consistent in our group and the four provinces in southern China as mentioned in the literature (Table 4).

3.4 | Comparison of thalassemia mutation spectrum in China and different Southeast Asian countries

We next compared our data of immigrants from Southern China, collected data from previous Chinese reports, and the data collected from literature focused on Southeast Asia for gene distribution. Studies in Thailand (Boonyawat, Monsereenusorn, & Traivaree, 2014), Indonesia (Rujito, Basalamah, & Mulatsih, 2015), and Malaysia (Dxr et al., 2017; Tan et al., 2006) were finally selected (Figure 1) according to our criteria. The Malaysian report offered gene composition for α -thalassemia and different ethnic data, which included Chinese, Malay, and Indian people for β -thalassemia.

For α -thalassemia, there were few data from Southeast Asia. Different from Chinese people (including our data), in whom --SEA, $-\alpha^{3.7}$, $-\alpha^{4.2}$ were the most common alleles. Malaysians had a high ratio of α^{CS} (44.19%) and sizable

proportions of α^{IniCd} (9.30%) and α^{Cd59} (9.30%), which were rarely seen in Chinese people (Figure 1a).

For β -thalassemia, Thai and Malaysian Chinese people had similar frequent genes (IVS-II-654, CD41/42, and CD17) with Chinese people. Detailed analysis showed that Malaysian Chinese individuals had the same mutation spectrum as Guangdong people, both in our study ($p = 0.135$) and Lai et al.'s research ($p = 0.068$), indicating a possible similar race origin. However, for Indonesian, Malay, and Indian people in Malaysia, the gene constitution was very different from that of Chinese. They had a high percentage of β IVS-5 mutation, which was almost not found in Chinese (Figure 1b).

4 | DISCUSSION

As far as we know, this is the first study that compared α - and β -thalassemia mutation spectrum of people who came from northern and southern China and Southeast Asian countries.

As shown in our study, people came to our hospital to detect thalassemia gene mainly due to microcytic hypochromic anemia, hemolytic anemia of unknown origin, or positive family history of thalassemia. Although there were no strict demographic census data, our study reflected the thalassemia gene distribution in north China to some extent because our hospital is the largest center for thalassemia gene screening in north China.

TABLE 4 α - and β -thalassaemia gene distribution in four southern provinces

	Fujian province			Guangdong province			Guangxi province			Chongqing province		
	Literature (Xu et al., 2013)		Our data	Literature (Chen et al., 2004; Yin et al., 2014; Zhang et al., 2010)		Our data	Literature (Cai et al., 2002; Li et al., 2009; Pan et al., 2007; Qiu et al., 2009; Xiong et al., 2010; Zeng et al., 2014)		Our data	Literature (Yao, Yu, et al., 2013; Yao, Zhang, et al., 2013)		Our data
	n	%		n	%		n	%		n	%	
α-thalassaemia												
Total (N)	40	363	33	3,889	83	1722	4	56				
-SEA	28	70.0	28	84.9	58	69.9	784	45.5	3	75.0	10	17.9
$-\alpha^{3.7}$	9	22.5	3	9.1	13	15.7	437	25.4	0	0.0	30	53.6
$-\alpha^{4.2}$	0	0.0	1	3.0	8	9.6	194	11.3	1	25.0	5	8.9
α^{CS}	0	0.0	5	1.4	3	3.6	136	7.9	0	0.0	0	0.0
α^{QS}	2	5.0	12	3.3	1	1.2	96	5.6	0	0.0	1	1.8
α^{WS}	1	2.5	2	0.6	0	0.0	52	3.0	0	0.0	0	0.0
χ^2	5.98		14.079		18.803		8.267					
<i>p</i>	0.241		0.009		0.001		0.026					
β-thalassaemia												
Total (N)	29	148	35	1,201	66	1,009	15	45				
IVS-II-654	15	51.7	9	25.7	4	6.1	67	6.6	6	40.0	9	20.0
CD41/42	10	34.5	15	42.9	28	42.4	427	42.3	7	46.7	21	46.7
CD17	3	10.3	6	17.1	24	36.4	283	28.1	2	13.3	5	11.1
CD71/72	0	0.0	3	8.6	6	9.1	78	7.7	0	0.0	0	0.0
-28	0	0.0	1	0.7	4	6.1	65	6.4	0	0.0	0	0.0
βE	0	0.0	2	1.4	0	0.0	36	3.6	0	0.0	5	11.1
-29	0	0.0	2	1.4	2	5.7	7	0.7	0	0.0	4	8.9
Other ^a	1	3.5	16	10.8	0	0.0	46	4.6	0	0.0	1	2.2
χ^2	4.673		9.525		6.896		4.409					
<i>p</i>	0.693		0.151		0.391		0.505					

^aAll other rare mutations in β -thalassaemia were categorized as "other."

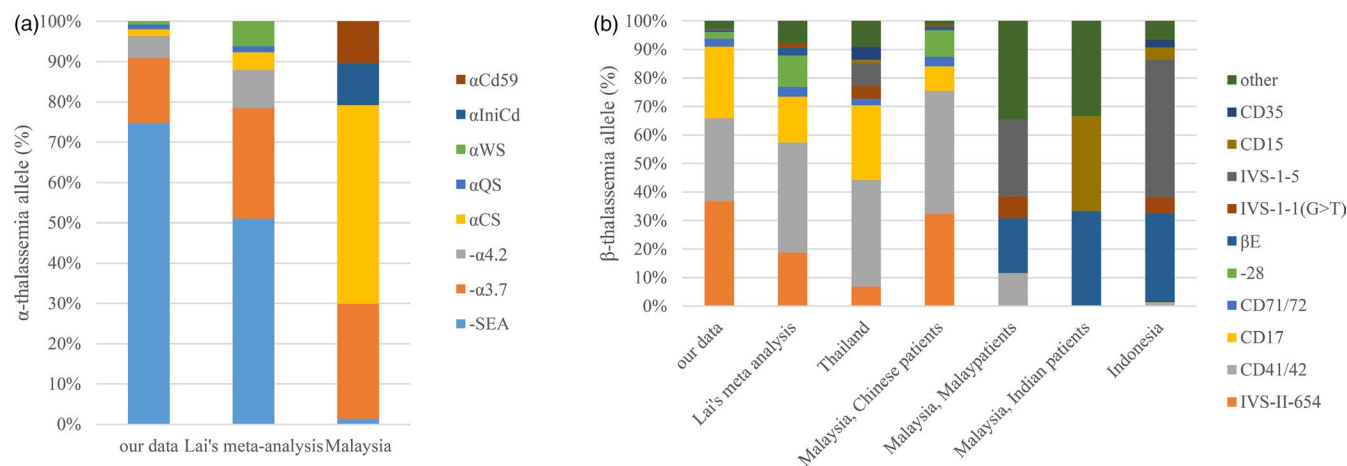


FIGURE 1 Comparison of mutation spectrum between China and Southeast Asia. (a) α -thalassemia and (b) β -thalassemia

There was more β -thalassemia patients than α -thalassemia patients in our study, which is consistent with some of the thalassemia epidemic literature in China (Niu, Huang, An, Wang, & Jiang, 2016; Zhang et al., 2012). However, Lai et al.'s data were in agreement with studies conducted in Taiwan (Peng et al., 2013) and Guilin, Guangxi (Tang et al., 2015). β -thalassemia was distributed in wide regions, such as Southeast Asia, the Middle East, the Mediterranean region, and North and South America. However, cases of α -thalassemia were mostly observed in Southeast Asia, the Middle East, Africa and Mediterranean region (Kaushansky, Beutler, & Kipps, 2010). We presumed that the distribution of α - and β -thalassemia would be different in China as well. Lai et al.'s meta-analysis included four provinces (as shown in Table 4), while our study covered nearly all provinces of mainland China. We had a significant part of patients who originated from provinces of Sichuan, Hunan, and Hubei, which was not included in Lai et al.'s study. In our study, the percentage of β -thalassemia was 72.3%, 71.2% and 73.2% in Sichuan, Hunan, and Hubei, respectively, which indicated that there were more β -thalassemia patients in these provinces. The difference of relative number of α - and β -thalassemia between ours and Lai et al.'s data was probably due to the difference of samples. We hypothesized that α -thalassemia patients mainly concentrated in Southern coastal region of China.

Difference of α -thalassemia mutation spectrum between northern China (our data) and southern China (Lai et al.'s data) was statistically significant ($p < 0.001$). We noticed that patients who lived in the north had higher percentage of $--_{SEA}$ than those who originated from the same province, but lived in southern China. This phenomenon was observed in the provinces of Fujian (Xu et al., 2013), Guangdong (Chen et al., 2004; Yin et al., 2014; Zhang et al., 2010), Guangxi (Cai et al., 2002; Li et al., 2009; Pan et al., 2007; Xiong et al., 2010; Zeng et al., 2014), and Chongqing (Yao, Yu, et al., 2013) (Table 4). The high gene frequency of $--_{SEA}$ in northern China could be partly explained by the difference of sampling

methods. In Lai et al.'s meta-analysis, all included studies used cluster sampling or stratified sampling, while in our study, only suspected thalassemia patients were involved in genetic testing. It is known that people with only one $-\alpha^{3.7}$ or $-\alpha^{4.2}$ deletions were silent carriers, and may not have clinical manifestations or anemia. Thus, probably less carriers with $-\alpha^{3.7}$ or $-\alpha^{4.2}$ were included in our study, and the percentage may decrease as well. Furthermore, our study covered nearly all the provinces in mainland China, and the proportion of people from some provinces was different from Lai et al.'s meta-analysis. Even though, the three most common mutations, which accounted for nearly 90% of patients, were the same in ours and Lai et al.'s report, indicating the consistency of gene pattern of Chinese people.

Similarly, the three most common mutations of β -thalassemia in our group, IVS-II-654, CD41/42 and CD17, were similar to those of Lai et al.'s data although the ratios of them were different. Nevertheless, we did not observe the difference in the four provinces that we chose to analyze in both groups. β -Thalassemia mutation spectrum was also different in different provinces. For example, IVS-II-654 was the most common mutation in Fujian, while CD41/42 was common in the other three provinces. Consequently, the difference between our data and Lai et al.'s data was probably due to difference in samples collected from different provinces. Many provinces were not included in the detailed comparison because data of these provinces were absent in Lai et al.'s study.

The mutation spectrum of α - and β -thalassemia is diverse in different areas, however, the race difference may also play a great role. Although Chinese and Southeast Asian people were both yellow races, they had significant difference in thalassemia mutation spectrum. For α -thalassemia, the most common allele in Malaysia was α^{CS} (44.19%) instead of $--_{SEA}$ in China. We also paid attention to the sizable proportions of α^{IniCd} (9.30%) and α^{Cd59} (9.30%) in Malaysians, which were rare in Chinese patients. As for β -thalassemia, IVS-II-654, a common mutation in China, was not very common

in Southeast Asia countries, except for Chinese Malaysians. IVS-I-5 and β E were common in Malaysia and Indonesia (Rujito et al., 2015; Tan et al., 2006), but not in China.

Interbreeding of indigenous people and immigrants may increase the diversity of local mutation spectrum. For instance, HbS mutation in Central West Africa flew to North Africa due to the Ottoman rule in the 17th century, and then flew to Europe through slavery roads. As a result, thalassemia patients in the Mediterranean region have a high ratio of Codon 39 (C \rightarrow T) and IVS-I-110 (G \rightarrow A) mutation in HBB, which are the most frequent gene in Tunisia and Algeria (Anwar, Khyatti, & Hemminki, 2014). Migration from China to Southeast Asia altered the mutation spectrum of thalassemia. As shown in Figure 1b, Chinese people and Malaysian Chinese had similar gene distribution diagram. Detailed analysis even showed that for β -thalassemia, Malaysian Chinese had the same mutation spectrum as Guangdong people, indicating a possibly similar race origin. Since Ming Dynasty, a large number of Chinese people had immigrated to Southeast Asia, and most of them originated from coastal areas, including Guangdong province. Similar explanation may also be true for Thai people, although the study that we chose did not clarify the ration of patients, thus we could only see the overall gene distribution instead of the separated ones (Boonyawat et al., 2014).

In addition to Chinese people, Indians were also a large immigrant group in Southeast Asia (Wu, 2015), which also made an impact on thalassemia mutation spectrum in Southeast Asia. The most frequent allele of α -thalassemia in Indians was $-\alpha^{3.7}$ (Purohit, Dehury, Patel, & Patel, 2014), which was observed in Malaysians as well. As for β -thalassemia, IVS-I-5 (G \rightarrow C), Codon 15 (G \rightarrow A), and Codon 30 (G \rightarrow C), which were common in India (Edison et al., 2008), were also seen in Thailand, Malaysia, and Indonesia. However, IVS-I-5 mutation was also observed in Malay patients, thus it is not clear whether this mutation was originated from native people or Indian immigrants.

It is the first time that we describe the similarities and differences between Chinese people in different areas and Chinese people with other races in Southeast Asia. Further studies with larger sample size and more included countries with human species composition are helpful for the complete understanding of gene transition in these countries and races as well.

ACKNOWLEDGMENT

We are grateful to every patient who consented to disclose their medical records and answered our review calls. Their cooperation and helpful comments throw light upon our data collection and statistical analysis. This study was supported by grants from the Chinese Academy of Medical Sciences innovation fund for medical sciences (2016-I2M-3-004)

and MANDARIN study from Novartis (Protocol Number: CICL670AAU05).

CONFLICT OF INTEREST

The authors declare that there are no conflict of interests regarding the publication of this article.

ORCID

Wenzhe Zhou  <https://orcid.org/0000-0002-3191-4003>

Bing Han  <https://orcid.org/0000-0001-9970-0733>

REFERENCE

- Anwar, W. A., Khyatti, M., & Hemminki, K. (2014). Consanguinity and genetic diseases in North Africa and immigrants to Europe. *European Journal of Public Health*, 24(suppl_1), 57–63. <https://doi.org/10.1093/eurpub/cku104>
- Boonyawat, B., Monsereenusorn, C., & Traivaree, C. (2014). Molecular analysis of beta-globin gene mutations among Thai beta-thalassemia children: Results from a single center study. *Application of Clinical Genetics*, 7, 253–258. <https://doi.org/10.2147/tacg.s73058>
- Cai, R., Li, L., Liang, X., Liu, Z., Su, L., Li, W., ... Xu, X. (2002). Prevalence survey and molecular characterization of alpha and beta thalassemia in Liuzhou city of Guangxi. *Chinese Journal of Epidemiology*, 23(4), 281–285.
- Chen, P. Y., Chen, Y. J., Xia, Y. Y., & Qin, C. R. (2004). Molecular epidemiological study of α -thalassemia in Shaoguan city of Guangdong. *Chinese Journal of Primary Medicine Pharmacy*, 11(5), 518–520.
- Edison, E. S., Shaji, R. V., Devi, S. G., Moses, A., Viswabandhya, A., Mathews, V., ... Chandy, M. (2008). Analysis of β globin mutations in the Indian population: Presence of rare and novel mutations and region-wise heterogeneity. *Clinical Genetics*, 73(4), 331.
- Gan, Y., Yang, Z., Chen, Q., Zhang, Y., & Cui, W. (2014). Analysis of gene mutation spectrum of α - and β -thalassemia patients in Beijing. *Chinese Journal of Health Laboratory Technology*, 24(12), 1730–1753.
- Kaushansky, K., Beutler, E., & Kipps, T. J. (2010). *Williams Hematology*, 8th ed. New York, NY: McGraw-Hill Education / Medical.
- Koh, D. X. R., Sabudin, R. Z. A. R., Yusoff, M. M., Hussin, N. H., Ahmad, R., Othman, A., & Ismail, E. (2017). Molecular characterisation of α - and β -thalassaemia among indigenous Senoi Orang Asli communities in Peninsular Malaysia. *Annals of Human Genetics*, 81(5), 205–212.
- Kuesap, J., Chaijaroenkul, W., Rungsihirunrat, K., Pongjantharasatien, K., & Na-Bangchang, K. (2015). Coexistence of malaria and thalassemia in malaria endemic areas of Thailand. *The Korean Journal of Parasitology*, 53(3), 265. <https://doi.org/10.3347/kjp.2015.53.3.265>
- Lai, K. T., Huang, G. F., Li, S., & He, Y. Y. (2017). The prevalence of thalassemia in mainland China: Evidence from epidemiological surveys. *Scientific Reports*, 7(1), 920. <https://doi.org/10.1038/s41598-017-00967-2>
- Li, C. K. (2017). New trend in the epidemiology of thalassaemia. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 39, 16–26. <https://doi.org/10.1016/j.bpobgyn.2016.10.013>

- Li, X. Y., Lan, J., Luo, R. G., Zhang, T. L., Nong, Q. T., Yang, W. W., ... Zhang, X. H. X. (2009). Epidemiological investigation and genotype of thalassemia on middle school students in Chongzuo, Guangxi. *Chinese Journal of Epidemiology*, *30*, 567–570.
- Modell, B., & Darlison, M. (2008). Global epidemiology of haemoglobin disorders and derived service indicators. *Bulletin of the World Health Organization*, *86*, 480–487. <https://doi.org/10.2471/BLT.06.036673>
- Niu, Q., Huang, X. B., An, Y. F., Wang, J., & Jiang, H. (2016). Genotype of Thalassemia in Han Chinese and Tibetans in Sichuan Province, China. *Journal of Sichuan University. Medical Science Edition*, *47*(6):941.
- Pan, H. F., Long, G. F., Li, Q., et al. (2007). Current status of thalassemia in minority populations in Guangxi, China. *Clinical Genetics*, *71*(5), 419–426.
- Peng, C.-T., Liu, S.-C., Peng, Y.-C., Lin, T.-H., Wang, S.-J., Le, C.-Y., ... Lin, C.-Y. (2013). Distribution of thalassemias and associated hemoglobinopathies identified by prenatal diagnosis in Taiwan. *Blood Cells Molecules & Diseases*, *51*(3), 138–141.
- Purohit, P., Dehury, S., Patel, S., & Patel, D. K. (2014). Prevalence of deletion alpha thalassemia and sickle gene in a tribal dominated malaria endemic area of Eastern India. *ISRN Hematology*, *2014*, 745245.
- Qiu, X. Y., Chen, P., Zhang, N., Zeng, X. Y., Huang, C. Q., Chen, W. Q., ... Li, S. Q. (2009). Study on the incidence of β -Thalassemia and genotypes among children under 7 year-olds in Nanning, Liuzhou and Baise areas, Guangxi province. *Chinese Journal of Epidemiology*, *30*, 1021–1024.
- Rujito, L., Basalamah, M., Mulatsih, S., & Sofro, A. S. (2015). Molecular scanning of β -thalassemia in the southern region of central java, indonesia; a step towards a local prevention program. *Hemoglobin*, *39*(5), 330–333.
- Sayani, F. A., & Kwiatkowski, J. L. (2015). Increasing prevalence of thalassemia in America: Implications for primary care. *Annals of Medicine*, *47*(7), 592.
- Tan, J. A., Chin, P. S., Wong, Y. C., Tan, K. L., Chan, L. L., & George, E. (2006). Characterisation and confirmation of rare beta-thalassaemia mutations in the Malay, Chinese and Indian ethnic groups in Malaysia. *Pathology*, *38*(5), 437–441.
- Tang, W., Zhang, C., Lu, F., Tang, J., Lu, Y., Cui, X., ... Li, S. (2015). Spectrum of α -thalassemia and β -thalassemia mutations in the Guilin Region of southern China. *Clinical Biochemistry*, *48*(16–17), 1068.
- Wu, Y. (2015). *Comparative Study on the Chinese and Indians in Contemporary Southeast Asia*. Kunming: Yunnan University.
- Xiong, F., Sun, M., Zhang, X., Cai, R., Zhou, Y., Lou, J., ... Xu, X. (2010). Molecular epidemiological survey of haemoglobinopathies in the Guangxi Zhuang Autonomous Region of southern China. *Clinical Genetics*, *78*, 139–148. <https://doi.org/10.1111/j.1399-0004.2010.01430.x>
- Xu, L. P., Huang, H. L., Wang, Y., Zheng, L., Wang, L. S., Xu, J. B., ... Lin, Y. (2013). Molecular epidemiological analysis of α - and β -thalassemia in Fujian province. *Chinese Journal of Medical Genetics*, *30*(4), 403.
- Yao, X. Y., Yu, J., Chen, S. P., Xiao, J. W., Zheng, Q. C., Liu, H. Y., ... Zou, L. (2013). Prevalence and genetic analysis of α -thalassemia and β -thalassemia in Chongqing area of China. *Gene*, *532*(1), 127–131.
- Yao, X. Y., Zhang, Y. M., Qin, Z. Z., Fan, R., Zou, L., Chen, S. P., ... Yu, J. (2013). Prevalence and molecular analysis of β -thalassemia in children of Han ethnicity in Chongqing city. *Chinese Journal of Pediatrics*, *51*(7), 518.
- Yin, A., Li, B., Luo, M., Xu, L., Wu, L., Zhang, L., ... Zhang, X. (2014). The prevalence and molecular spectrum of α - and β -globin gene mutations in 14,332 families of Guangdong Province, China. *PLoS One*, *9*(2), e89855.
- Zeng, X. Y., Chen, C. C., Chen, P., Zhong, Q. A., & Qiu, X. Q. (2014). The research of α -thalassemia genotypes on children under 7 years old in Nanning, Liuzhou and Baise areas of Guangxi province. *Maternal and Child Health Care of China*, *29*, 413–416.
- Zhang, C. M., Wang, Y., Gao, J. H., Liao, J. L., Feng, H. J., Wan, B., ... Huang, G. Z. (2010). Molecular epidemiology survey of α -thalassemia in Zhongshan city, Chinese. *Journal of Child Health Care*, *18*(5), 361–364.
- Zhang, J., Zhu, B. S., He, J., Zeng, X. H., Su, J., Xu, X. H., ... Zhang, Y. H. (2012). The Spectrum of α - and β -thalassemia mutations in Yunnan Province of Southwestern China. *Hemoglobin*, *36*(5), 464–473.

How to cite this article: Yang Z, Cui Q, Zhou W, Qiu L, Han B. Comparison of gene mutation spectrum of thalassemia in different regions of China and Southeast Asia. *Mol Genet Genomic Med*. 2019;7:e680. <https://doi.org/10.1002/mgg3.680>