Clinical Chemistry 69:3
239–250 (2023)

Evaluating the Clinical Utility of a Long-Read Sequencing-Based Approach in Prenatal Diagnosis of Thalassemia

Qiaowei Liang,[a,†,‡] Jun He,[b,†] Qing Li,[c,†] Yulin Zhou,[d,†] Yanqiu Liu,[e,†] Youqiong Li,[f] Lingfang Tang,[g] Shengwen Huang,[h] Rong Li,[i] Fanqian Zeng,[j] Aiping Mao,[k] Yinyin Liu,[k] Desheng Liang,[a,l,*] and Lingqian Wu[a,l,*]

BACKGROUND: The aim is to evaluate the clinical utility of a long-read sequencing-based approach termed comprehensive analysis of thalassemia alleles (CATSA) in prenatal diagnosis of thalassemia.

METHODS: A total of 278 fetuses from at-risk pregnancies identified in thalassemia carrier screening by PCR-based methods were recruited from 9 hospitals, and PCR-based methods were employed for prenatal diagnosis. CATSA was performed retrospectively and blindly for all 278 fetuses.

RESULTS: Among the 278 fetuses, 263 (94.6%) had concordant results and 15 (5.4%) had discordant results between the 2 methods. Of the 15 fetuses, 4 had discordant thalassemia variants within the PCR detection range and 11 had additional variants identified by CATSA. Independent PCR and Sanger sequencing confirmed the CATSA results. In total, CATSA and PCR-based methods correctly detected 206 and 191 fetuses with variants, respectively. Thus, CATSA yielded a 7.9% (15 of 191) increment as compared with PCR-based methods. CATSA also corrected the predicted phenotype in 8 fetuses. Specifically, a PCR-based method showed one fetus had homozygous *HBB* c.52A > T variants, while CATSA determined the variant was heterozygous, which corrected the predicted phenotype from β-thalassemia major to trait, potentially impacting the pregnancy outcome. CATSA additionally identified α-globin triplicates in 2 fetuses with the heterozygous *HBB* c.316-197C > T variant, which corrected the predicted phenotype from β-thalassemia trait to intermedia and changed the disease prognosis.

CONCLUSIONS: CATSA represents a more comprehensive and accurate approach that potentially enables more informed genetic counseling and improved clinical outcomes compared to PCR-based methods.

Introduction

Thalassemia is the most widely distributed monogenic autosomal recessive disorder characterized by defects in globin synthesis (1, 2). Depending on the impaired globin chain, it is mainly classified into α- and β-thalassemia. In China, thalassemia is most prevalent in the southern provinces such as Guangdong, Guangxi, Guizhou, and Hainan (3–6). The Southeast Asian deletion ($-^{SEA}$), deletion of 4.2 kb ($-\alpha^{4.2}$), deletion of 3.7 kb ($-\alpha^{3.7}$), hemoglobin (Hb) Westmead (*HBA2* c.369C > G), Hb Quong Sze (*HBA2* c.377T > C), and Hb Constant Spring (*HBA2* c.427T > C) are the most common α-thalassemia variants. Nineteen single nucleotide variants (SNVs) and indels of the *HBB* gene are classified as the most common β-thalassemia variants (7).

[a]Department of Medical Genetics, Hunan Jiahui Genetics Hospital, Changsha, Hunan, China; [b]Changsha Hospital for Maternal and Child Health Care, Changsha, Hunan, China; [c]Key Laboratory for Major Obstetric Diseases of Guangdong Province, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong, China; [d]Women and Children's Hospital, School of Medicine and School of Public Health, Xiamen University, Xiamen, China; [e]Jiangxi Maternal and Child Health Hospital, Nanchang, Jiangxi, China; [f]Center for Medical Genetics and Prenatal Diagnosis, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, Guangxi, China; [g]Guilin Women and Children Health Care Hospital, Guilin, Guangxi, China; [h]Guizhou Provincial People's Hospital, Guiyang, Guizhou; [i]Department of Obstetrics and Gynecology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China; [j]Yunnan Maternal and Child Health Care Hospital, Kunming, Yunnan, China; [k]Berry Genomics Corporation, Beijing, China; [l]Center for Medical Genetics and Hunan Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, Changsha, Hunan, China.
*Address correspondence to: D.L. at Center for Medical Genetics and Hunan Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, 110 Xiangya Rd., Changsha 410078, Hunan, China. E-mail liangdesheng@sklmg.edu.cn. L.W. at Center for Medical Genetics and Hunan Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, 110 Xiangya Rd., Changsha 410078, Hunan, China. E-mail wulingqian@sklmg.edu.cn.
†These authors contributed equally to this work.
‡Present address: Department of Human Genetics, Graduate School of Medicine, Yokohama City University, Yokohama, Japan.
Received August 3, 2022; accepted November 2, 2022.
https://doi.org/10.1093/clinchem/hvac200

Normal individuals have 4 functional α-globin genes (αα/αα) and 2 functional β-globin genes (β/β). Phenotypes of thalassemia range from asymptomatic to severe and sometimes lethal anemia, and the severity of this disease is usually correlated with the number of impaired globin genes (8). There are 4 α-thalassemia conditions of increasing severity: silent α-carrier (-α/αα), α-thalassemia trait (α-TT) (−/αα or -α/-α), Hb H disease (−/-α), and Hb Barts hydrops fetalis (−/−) (9). The first 2 types are asymptomatic and often collectively recognized as α-thalassemia trait, while the last 2 are clinically symptomatic. Hb H disease is highly heterogeneous and shows moderate to severe microcytic, hypochromic, and hemolytic anemia. Hb Barts hydrops fetalis is usually fatal as the fetus dies either in utero or shortly after birth. Similarly, β-thalassemia can also be divided into 4 clinical conditions of increasing severity: silent β-carrier (β^{++}/β^{N}), β-thalassemia trait (β-TT) (β^{+}/β^{N} or β^{0}/β^{N}), β-thalassemia intermedia (β-TI) (β^{+}/β^{+} or β^{0}/β^{+}), and β-thalassemia major (β-TM) (β^{+}/β^{+}, β^{0}/β^{+} or β^{0}/β^{0}) (9). Depending on the severity of β^{+}, β^{+}/β^{+} and β^{0}/β^{+} can be classified as either β-TI or β-TM, and it is difficult to distinguish these 2 phenotypes based only on the genotype. β-TI is nontransfusion-dependent thalassemia, and patients with β-TI encompass a wide range of phenotypes that spans from mild anemia to more severe anemia. β-TM is the most severe form of β-thalassemia, and affected individuals usually present with severe anemia and are transfusion-dependent.

Currently, there is no effective way to cure thalassemia, and prenatal diagnosis can be utilized to reduce the risk of having severely affected children (10). European Molecular Genetics Quality Network guidelines suggest prenatal testing for all couples who are at risk for pregnancies with Hb Barts hydrops fetalis and β-TM and occasionally for at-risk pregnancies with Hb H disease and β-TI because of the unpredictable phenotype (11). In China, guidelines suggest prenatal testing for couples at risk for pregnancies with Hb Barts hydrops fetalis, Hb H disease, β-TM, and β-TI under informed consent due to the varied clinical phenotype and substantial medical burden (12, 13).

Worldwide, at least 60 000 severely affected children are born annually (8, 14). In southern China, the prevalence of patients with Hb H disease, β-TM, and β-TI ranges from about 0.2% to 1.5% (15–19). Conventional PCR-based methods, including gap-PCR, PCR-reverse dot blot (RDB), and PCR-flow fluorescence hybridization for the 25 most common thalassemia variants, are routinely employed for carrier screening and prenatal diagnosis of thalassemia in China. However, rare and novel thalassemia variants such as α-globin triplicates and large deletions are missed by PCR-based methods (20, 21). Triplications of α-globin genes can exacerbate the symptoms of β-thalassemia due to the aggravated imbalance of α/β-globin chains, thus detecting α-globin triplicates is of clinical importance (8, 22).

Recently, next-generation sequencing and long-read sequencing (LRS) are emerging as new techniques in genetic testing for thalassemia (21, 23, 24). Though next-generation sequencing has showed improvements of molecular epidemiological characterization of thalassemia (19, 23, 25), the short read length has made it difficult to accurately read long and highly homologous regions such as the α-globin gene locus. LRS can analyze globin genes in single long reads and thus directly detect both common and rare variants including α-globin triplicates and large deletions. Recently, the clinical utility of a LRS-based approach termed *comprehensive analysis of thalassemia alleles* (CATSA), using the PacBio platform in carrier screening of thalassemia, was validated with a large cohort of samples, and CATSA showed great advantages in terms of detection range and accuracy compared to PCR-based methods (21, 26). However, the clinical feasibility of CATSA in prenatal diagnosis of thalassemia has not been explored.

In this retrospective blind clinical study, we evaluated the clinical utility of CATSA in prenatal diagnosis of thalassemia by comparing to PCR-based methods. These 2 methods were both performed for 278 fetuses from at-risk pregnancies and the results were compared for clinical effectiveness in the prenatal diagnosis of thalassemia.

Materials and Methods

STUDY PARTICIPANTS

Between January 2020 and December 2020, a total of 278 fetuses from 278 high-risk pregnancies were recruited (Supplemental Table 1). The study design is detailed in Fig. 1. Amniotic fluid samples were collected from the 278 pregnant women with gestation age between 15 to 26 weeks, and genomic DNA was extracted with DNease Blood & Tissue Kit (Qiagen). Short tandem repeat analysis was performed using genomic DNA extracted from maternal peripheral blood and amniotic fluid samples to exclude maternal cell contamination. In the routine process for prenatal diagnosis, PCR-based methods were performed on genomic DNA extracted from amniotic fluid samples by hospitals for common thalassemia variants. The DNA samples were also retrospectively sent to the genetic testing laboratory of Berry Genomic Corporation for blind CATSA assay. The principal coordinating center was Hunan Jiahui Genetics Hospital. Study ethics approval was granted by the institutional review board of the Hunan Jiahui Genetics Hospital (#2020-01) and the Central South University (#2019-1-30). The other 8

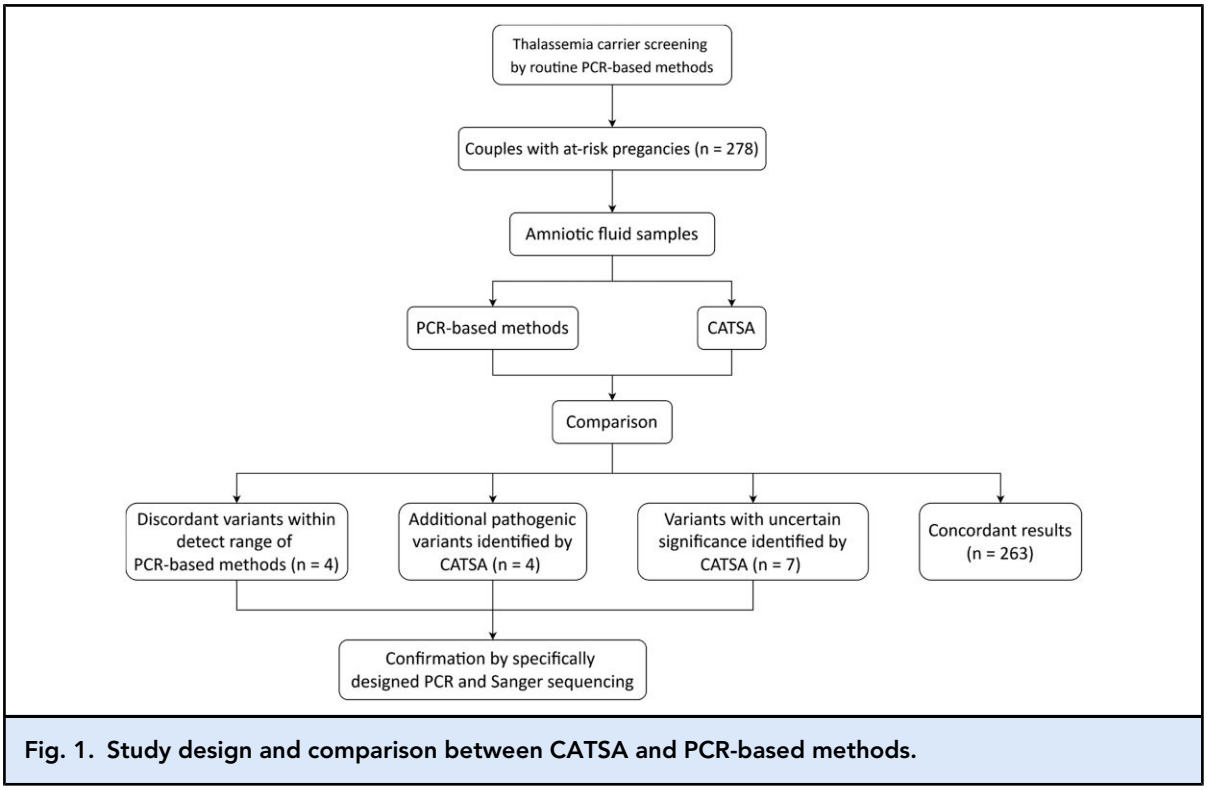

Fig. 1. Study design and comparison between CATSA and PCR-based methods.

study centers also obtained ethics approval from their institutional review boards. All subjects or their legal guardians provided informed written consent.

HEMOGLOBIN TESTING

Hb testing was performed by standard blood assays and Hb electrophoresis. Hematological parameters measured included mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), and Hb. Hb electrophoresis profiling included hemoglobin A2 (Hb A2) and hemoglobin F (Hb F). Normal ranges included MCV $\geq$ 80 fL, MCH $\geq$ 27 pg, Hb A2 levels between 2.5% and 3.5%, and Hb F $\leq$ 5%.

THALASSEMIA GENETIC TESTING BY PCR-BASED METHODS

PCR-based methods, including gap-PCR and PCR-RDB, were carried out at 9 study centers according to the manufacturer's protocol (Yanengbio, Yishengtang, Kaipu Bioscience). A gap-PCR kit was used to detect common deletional α-thalassemia variants including $--^{SEA}$, $-\alpha^{3.7}$, and $-\alpha^{4.2}$. A PCR-RDB kit for α-thalassemia was used to detect variants including *HBA2* c.369C > G, *HBA2* c.377T > C, and *HBA2* c.427T > C. A PCR-RDB kit for β-thalassemia was used to detect variants of *HBB* gene including c.-82C > A, c.-80T > C, c.-79A > G, c.-76A > G, c.-50A > C, c.-11_-8delAAAC, c.2T > G, c.45_46insG, c.52A > T, c.79G > A, c.84_85insC, c.92 + 1G > A, c.92 + 1G > T, c.92 + 5G > C, c.94delC, c.126_129delCTTT, c.130G > T, c. 216_217insA, and c.316-197C > T.

THALASSEMIA GENETIC TESTING BY CATSA

CATSA was performed similarly as described previously (26). Briefly, genomic DNA was amplified by multiplex long-range PCR with primers encapsulating the majority of known structural variations, SNVs, and indels in the *HBA1*, *HBA2,* and *HBB* genes. A one-step end-repair and ligation reaction was conducted to add barcoded adaptors to the PCR products. Then each prelibrary was pooled together by equal mass and converted to single-molecule real-time dumbbell libraries with Sequel Binding and Internal Ctrl Kit 3.0 (Pacific Biosciences) and sequenced on a PacBio Sequel II platform. Raw subreads were converted to high-fidelity circular consensus sequencing reads, demultiplexed by barcodes, and aligned to genome build hg38 in the SMRT Link system (Pacific Biosciences). SNVs and indels were called by FreeBayes1.3.4 (Biomatters, Inc.) and structural variations were identified according to HbVar, Ithanet, and LOVD databases. Pathogenic variants in the detection range of PCR-based methods were classified as common variants, pathogenic variants outside the detection range of PCR-based methods were classified as rare variants, and variants annotated as Hb

variants in HbVar and variants of unknown significance (VUS) in Ithanet and LOVD were classified as VUSs.

CONFIRMATION OF DISCORDANT VARIANTS

Triplications in the α-globin gene identified by CATSA were confirmed by specifically designed PCR to amplify the anti3.7 and anti4.2 segments as previously reported (27). Discordant SNVs and indels between CATSA and PCR-based methods were confirmed by Sanger sequencing.

Results

COMPARISON BETWEEN CATSA AND PCR-BASED METHODS

Among the 278 pregnancies, the number at risk for Hb Barts hydrops fetalis, Hb H disease, β-TM, β-TI/β-TM, and both Hb H disease and β-TI/β-TM was 104 (37.4%), 94 (33.8%), 40 (14.4%), 39 (14.0%), and 1 (0.4%), respectively (Table 1). In prenatal diagnosis by PCR-based methods, 23 fetuses were diagnosed as Hb Barts hydrops fetalis and 6 fetuses were diagnosed as β-TM, and all the 29 pregnancies were terminated following the diagnosis. Hb H disease was diagnosed in 25 pregnancies, of which 10 were terminated and 15 had normal delivery. β-TI/β-TM was diagnosed in 3 pregnancies, and they were all terminated. The remaining 226 fetuses were normal or had mild thalassemia; 223 had normal delivery and 3 were terminated due to 15q11.2 deletion, cleft lip, and heart malformation, respectively (Supplemental Table 1).

CATSA and PCR-based methods had concordant genotypes for 263 (94.6%) fetuses. Of the 15 (5.4%) fetuses with discordant results, 4 had discordant thalassemia variants within the PCR detection range (common variants), 4 had thalassemia variants outside the PCR detection range (rare variants), and 7 had VUSs (Table 2). Independent PCR and Sanger sequencing confirmed the results of CATSA were correct (Fig. 2, Supplemental Fig. 1). In total, CATSA and PCR-based methods correctly detected 206 and 191 fetuses with variants, respectively. CATSA yielded a 7.9% (15 of 191) increment as compared with PCR-based methods. CATSA also corrected the predicted phenotype in 8 fetuses (Table 2). Specifically, CATSA corrected the predicted phenotype from β-TM to β-TT for the fetus in family BDP230 and from β-TT to β-TI for the fetuses in families BDP026 and BDP116.

DISCORDANCE BETWEEN CATSA AND PCR-BASED METHODS FOR COMMON VARIANTS

Of the 278 fetuses, 4 had discordant results of common variants between CATSA and PCR-based methods (Table 2, Fig. 2A–D). Fetus BDP230 was diagnosed as *HBB* c.52A > T (β^0) homozygous (hom) by PCR-based methods, and termination of pregnancy was chosen because the correlated phenotype was β-TM. However, CATSA demonstrated that the genotype of BDP230 was *HBB* c.52A > T heterozygous (het), which was confirmed by Sanger sequencing (Fig. 2A). *HBB* c.52A > T het was corresponded to β-TT, and normally the pregnancy would be continued. For fetus BDP148, the genotype determined by PCR was wild-type while *HBB* c.52A > T het was identified by CATSA (Fig. 2B). PCR-based methods identified the $-^{SEA}$ variant in both the parents and fetus of family BDP092, but β-thalassemia genetic testing was not performed because both parents had normal Hb A2. CATSA and Sanger sequencing showed coincidence of *HBB* c.79G > A (β^+) het variant in fetus BDP092 (Fig. 2C). Both PCR-based methods and CATSA identified the *HBA2* c.377T > C het variant in fetus BDP141. However, CATSA found an additional thalassemia variant of *HBB* c.126_129delCTTT (β^0) het (Fig. 2D). In total, CATSA identified common β-thalassemia variants in 71 fetuses. Of these 71 fetuses, one (1.41%) was not tested for β-globin variants and 3 (4.22%) were misdiagnosed by PCR-based methods.

ADDITIONAL A- AND B-GLOBIN GENE VARIANTS DETECTED BY CATSA

With sequencing coverage of full α- and β-globin genes, CATSA can detect additional thalassemia variants other than common variants. Of the 278 fetuses, 10 additional variants including 4 rare thalassemia variants and 6 VUSs were identified in 11 fetuses by CATSA (Table 2). These variants included $\alpha\alpha\alpha^{anti4.2}$, $\alpha\alpha\alpha^{anti3.7}$, *HBB* c.-100G > A, *HBB* c.316-90A > G, *HBA1* c.300 + 55G > T, *HBA2* c.46G > A, *HBA2* c.327C > A, *HBA2* c.300 + 34G > A, *HBB* c.315 + 308delA, and *HBB* c.341T > A.

Both CATSA and PCR-based methods identified the *HBB* c.316-197C > T het (β^+/β^N, β-TT) variant in fetus BDP026. However, CATSA additionally identified an $\alpha\alpha\alpha^{anti4.2}$ variant in the fetus (Fig. 3A). Both parents of BDP026 were heterozygous carriers of *HBB* c.316-197C > T, and CATSA showed that the $\alpha\alpha\alpha^{anti4.2}$ allele was inherited from the father (Fig. 3A). Triplication of the α-globin gene can increase the imbalance between α/β-globin in patients with β-thalassemia and cause mild to severe β-TI. This could explain that, while the mother had mild abnormal hematological parameters, the father had hypochromic microcytosis with MCV of 64.9 fl and MCH of 19.2 pg (Table 2). The family also had a 12-year-old son who suffered from hypochromic microcytosis, and CATSA showed that he also was heterozygous for *HBB* c.316-197C > T compounded with $\alpha\alpha\alpha^{anti4.2}$ (Fig. 3A). Fetus BDP026 was born female, and she was found to have hypochromic microcytosis with

Table 1. Summary of prenatal diagnosis for thalassemia by PCR-based methods and CATSA.

			Prenatal diagnosis			Pregnancy outcome[a]	
Type	Disorder at-risk	Number	Predicted phenotype	PCR-based methods	CATSA	TOP[b]	Normal delivery
α-thalassemia	Hb Barts	104	Hb Barts	23	22	23	0
			Hb Barts, β-TT	0	1	0	0
			Hb H	1	1	0	1
			α-TT	45	43	0	45
			α-TT, β-TT	3	5	0	3
			β-TT	2	2	0	2
			Normal	30	30	0	30
	Hb H	94	Hb H	20	20	10	10
			Hb H, β-TT	4	4	0	4
			α-TT	45	44	1[b]	44
			α-TT, β-TT	2	3	0	2
			β-TT	4	4	0	4
			Normal	19	19	0	19
β-thalassemia	β-TM	40	β-TM	6	5	6	0
			β-TT	16	18	0	16
			α-TT, β-TT	3	3	0	3
			α-TT	2	2	0	2
			Normal	13	12	0	13
	β-TI/β-TM	39	β-TI/β-TM	3	3	3	0
			β-TI	0	2	0	0
			β-TT	21	19	1[b]	20
			α-TT, β-TT	2	2	0	2
			α-TT	1	1	0	1
			Normal	12	12	1[b]	11
α- and β-thalassemia	HbH, β-TI/β-TM	1	Normal	1	1	0	1

[a]The pregnancy outcome based on PCR-based methods results.
[b]Termination of pregnancy with reasons other than thalassemia.

MCV of 57.2 fl and MCH of 18.4 pg by the age of 2. CATSA confirmed that the postnatal genotype was the same as prenatal genotype (Fig. 3A).

Similarly, CATSA identified that fetus BDP116 had *HBB* c.316-197C > T het compounded with $\alpha\alpha\alpha^{anti3.7}$, which was not detected by PCR-based methods (Fig. 3B). Family analysis by CATSA suggested that the $\alpha\alpha\alpha^{anti3.7}$ allele was inherited from the mother, who had *HBB* c.126_129delCTTT het compounded with $\alpha\alpha\alpha^{anti3.7}$ and suffered from hypochromic microcytic anemia (Fig. 3B, Table 2). The father was a heterozygous carrier of *HBB* c.316-197C > T and had only hypochromic microcytosis (Table 2). When fetus BDP116 was followed up through 1 year of age after birth, her mother declared that the child had anemia but did not provide detailed hematological parameters.

In addition to heterozygous $-^{SEA}$, CATSA identified the *HBB* c.-100G > A het variant in fetus BDP241 (Fig. 3C), who had normal hematological parameters at 1 day old after birth (Table 2). PCR-based methods identified $-^{SEA}/-^{SEA}$ in fetus BDP242 and termination of pregnancy was performed at gestational age of 18 weeks, when the placenta and embryolemma showed morphological anomalies. CATSA found an additional thalassemia variant of *HBB* c.316-90A > G het in fetus BDP242 (Fig. 3D). Moreover, CATSA identified 6 VUSs in 7 fetuses including BDP262, BDP024, BDP071, BDP138, BDP021, BDP130, and BDP102 (Table 2). These

Table 2. Discordant genetic testing results between PCR-based methods and CATSA.

Variants	Family	Member[a]	Age (year)/ GA (week + day)	Hematology					PCR-based methods			CATSA			
				MCV (fl)	MCH (pg)	Hb (g/L)	Hb A2 (%)	HbF (%)	α-globin variants	β-globin variants	Predicted phenotype	α-globin variants	β-globin variants	Predicted phenotype	Pregnancy outcome
Common	BDP230	Mother	32	68.4	21.9	93	4.5	1.4	ND	c.52A > T het	β-TT	UK	UK	UK	TOP
		Father	33	63.2	22.6	128	5.9	0	ND	c.52A > T het	β-TT	UK	UK	UK	
		Fetus	18 + 0	NA	NA	NA	NA	NA	ND	**c.52A > T hom**[f]	**β-TM**	ND	**c.52A > T het**	**β-TT**	
	BDP148	Mother	32	70.2	22.8	114	4.8	2.0	ND	c.52A > T het	β-TT	UK	UK	UK	Normal delivery
		Father	27	60.3	22.6	132	4.4	0	ND	c.52A > T het	β-TT	UK	UK	UK	
		Fetus	17 + 2	NA	NA	NA	NA	NA	ND	**ND**	**Normal**	ND	**c.52A > T het**	**β-TT**	
	BDP092	Mother	25	65.5	22.7	118	2.3	1.2	$-^{SEA}/\alpha\alpha$	UK	α-TT	UK	UK	UK	Normal delivery
		Father	27	71.0	21.6	138	2.2	0.5	$-^{SEA}/\alpha\alpha$	UK	α-TT	UK	UK	UK	
		Fetus	18 + 1	NA	NA	NA	NA	NA	$-^{SEA}/\alpha\alpha$	UK	α-TT	$-^{SEA}/\alpha\alpha$	**c.79G > A het**	α-TT, **β-TT**	
	BDP141	Mother	36	72.1	23.2	112	2.1	1.6	*HBA2* c.377T > C het	ND	α-TT	UK	UK	UK	Normal delivery
		Father	39	69.3	25.6	142	4.7	2.8	$-^{SEA}/\alpha\alpha$	c.126_129delCTTT het	α-TT, β-TT	UK	UK	UK	
		Fetus	18 + 0	NA	NA	NA	NA	NA	*HBA2* c.377T > C het	**ND**	α-TT	HBA2c.377T > C het	**c.126_129delCTTT het**	α-TT, **β-TT**	
Rare	BDP026	Mother	35	75.6	24.7	99	4.4	0.5	ND	c.316-197C > T het	β-TT	ND	c.316-197C > T het	β-TT	Normal delivery
		Father	36	64.9	19.2	117	6.5	0.6	ND	c.316-197C > T het	β-TT	$\alpha\alpha\alpha^{anti4.2}/\alpha\alpha$	c.316-197C > T het	**β-TI**	
		Son	12	60.5	19.1	116	3.9	0	ND	c.316-197C > T het	β-TT	$\alpha\alpha\alpha^{anti4.2}/\alpha\alpha$	c.316-197C > T het	**β-TI**	
		Fetus	19 + 1	NA	NA	NA	NA	NA	ND	c.316-197C > T het	β-TT	$\alpha\alpha\alpha^{anti4.2}/\alpha\alpha$	c.316-197C > T het	**β-TI**	
		Child	2	57.2	18.4	110	4.2	0	ND	c.316-197C > T het	β-TT	$\alpha\alpha\alpha^{anti4.2}/\alpha\alpha$	c.316-197C > T het	**β-TI**	
	BDP116	Mother	30	58.6	19.8	89	5.6	1.0	ND	c.126_129delCTTT het	β-TT	$\alpha\alpha\alpha^{anti3.7}/\alpha\alpha$	c.126_129delCTTT het	**β-TI**	Normal delivery
		Father	34	56.4	19.5	129	5.2	0	ND	c.316-197C > T het	β-TT	ND	c.316-197C > T het	UK	
		Fetus	20 + 1	NA	NA	NA	NA	NA	ND	c.316-197C > T het	β-TT	$\alpha\alpha\alpha^{anti3.7}/\alpha\alpha$	c.316-197C > T het	**β-TI**	

Continued

Table 2. (continued)

Variants	Family	Member[a]	Age (year)/ GA (week + day)	Hematology					PCR-based methods			CATSA			
				MCV (fl)	MCH (pg)	Hb (g/L)	Hb A2 (%)	HbF (%)	α-globin variants	β-globin variants	Predicted phenotype	α-globin variants	β-globin variants	Predicted phenotype	Pregnancy outcome
	BDP241	Mother	32	74.1	22.3	110	2.2	2.0	$-^{SEA}/\alpha\alpha$	UK	α-TT	UK	UK	UK	Normal delivery
		Father	33	68.6	21.0	148	2.2	0.5	$-^{SEA}/\alpha\alpha$	UK	α-TT	UK	UK	UK	
		Fetus	16 + 3	NA	NA	NA	NA	NA	$-^{SEA}/\alpha\alpha$	**UK**	α-TT	$-^{SEA}/\alpha\alpha$	**c.-100G > A het**	α-TT, β-**TT**	
		Child	1 day	90.9	28.7	133	UK	UK	$-^{SEA}/\alpha\alpha$	UK	α-TT	UK	UK	UK	
	BDP242	Mother	32	69.1	22.8	114	2.1	3.6	$-^{SEA}/\alpha\alpha$	UK	α-TT	UK	UK	UK	TOP
		Father	34	70.4	22.1	151	2.3	1.2	$-^{SEA}/\alpha\alpha$	UK	α-TT	UK	UK	UK	
		Fetus	17 + 1	NA	NA	NA	NA	NA	$-^{SEA}/-^{SEA}$	UK	Hb Barts	$-^{SEA}/-^{SEA}$	**c.316-90A > G het**	Hb Barts, β-**TT**	
Unknown pathogenicity	BDP262	Mother	24	77.3	25.1	117	6.1	2.0	$-\alpha^{3.7}/\alpha\alpha$	c.-78A > G het	α-TT, β-TT	UK	UK	UK	TOP (cleft lip)
		Father	22	63.7	20	134	5.6	1.4	ND	c.52A > T het	β-TT	UK	UK	UK	
		Fetus	22 + 0	NA	NA	NA	NA	NA	ND	c.-78A > G het	β-TT	***HBA1* c.300 + 55G > T het**	c.-78A > G het	β-TT	
	BD	Mother	25	79.9	25.9	108	4.5	0.3	ND	c.79G > A het	β-TT	UK	UK	UK	Normal delivery
	P02	Father	30	60.9	19.9	115	3.9	0.8	ND	c.52A > T het	β-TT	UK	UK	UK	
	4	Fetus	18 + 6	NA	NA	NA	NA	NA	ND	ND	Normal	***HBA2* c.46G > A het**	UK	Normal	
	BD	Mother	24	68.2	21.5	81.0	4.6	3.2	UK	c.52A > T het	β-TT	UK	UK	UK	Normal delivery
	P071	Father	27	71.2	18.9	132	4.7	0	UK	c.126_129delCTTT het	β-TT	UK	UK	UK	
		Fetus	16 + 1	NA	NA	NA	NA	NA	UK	c.126_129delCTTT het	β-TT	***HBA2* c.327C > A het**	c.126_129delCTTT het	β-TT	
	BDP138	Mother	26	68.0	20.1	113	5.2	1.5	ND	c.52A > T het	β-TT	UK	UK	UK	Normal delivery

Continued

Table 2. (continued)

Variants	Family	Member[a]	Age (year)/ GA (week + day)	Hematology					PCR-based methods			CATSA			
				MCV (fl)	MCH (pg)	Hb (g/L)	Hb A2 (%)	HbF (%)	α-globin variants	β-globin variants	Predicted phenotype	α-globin variants	β-globin variants	Predicted phenotype	Pregnancy outcome
		Father	34	69.1	20.5	141	4.3	0	ND	c.52A > T het	β-TT	UK	UK	UK	
		Fetus	19 + 0	NA	NA	NA	NA	NA	ND	c.52A > T het	β-TT	***HBA2* c.300 + 34G > A het**	c.52A > T het	β-TT	
	BD	Mother	20	71.4	24.4	114	4.6	1.0	*HBA2* c.369C > G het	c.79G > A het	α-TT, β-TT	*HBA2* c.369C > G het	c.79G > A het	UK	Normal delivery
	P021	Father	21	60.5	20.8	118	5.6	0	ND	c.126_129delCTTT het	β-TT	ND	c.126_129delCTTT/ **c.315 + 308delA**	UK	
		Fetus	18 + 0	NA	NA	NA	NA	NA	ND	ND	Normal	ND	**c.315 + 308delA het**	Normal	
		Child	1	83.2	32.8	148	UK	UK	UK	UK	UK	UK	**UK**	**UK**	
	BDP130	Mother	22	59.8	23.2	94	5.0	2.5	UK	c.52A > T het	β-TT	UK	UK	UK	Normal delivery
		Father	24	66.2	24.5	129	5.4	0	UK	c.52A > T het	β-TT	UK	UK	UK	
		Fetus	18 + 0	NA	NA	NA	NA	NA	*HBA2* c.377T > C het	c.52A > T het	α-TT, β-TT	***HBA2* c.377T > C het**	c.52A > T/**c.315 + 308delA**	α-TT, β-TT	
	BDP102	Mother	27	80.5	25.3	119	2.4	1.2	$-\alpha^{3.7}/\alpha\alpha$	ND	α-TT	UK	UK	UK	Normal delivery
		Father	29	70.2	20.4	143	2.5	0	$-^{SEA}/\alpha\alpha$	ND	α-TT	UK	UK	UK	
		Fetus	17 + 2	NA	NA	NA	NA	NA	$-^{SEA}/\alpha\alpha$	ND	α-TT	$-^{SEA}/\alpha\alpha$	**c.341T > A het**	α-TT	

[a]Child represents child from the pregnancy in this study.

TOP, termination of pregnancy; ND, not detected (the test was performed but no variants were detected); UK, uncharacterized (the test was not performed); NA, not applicable.

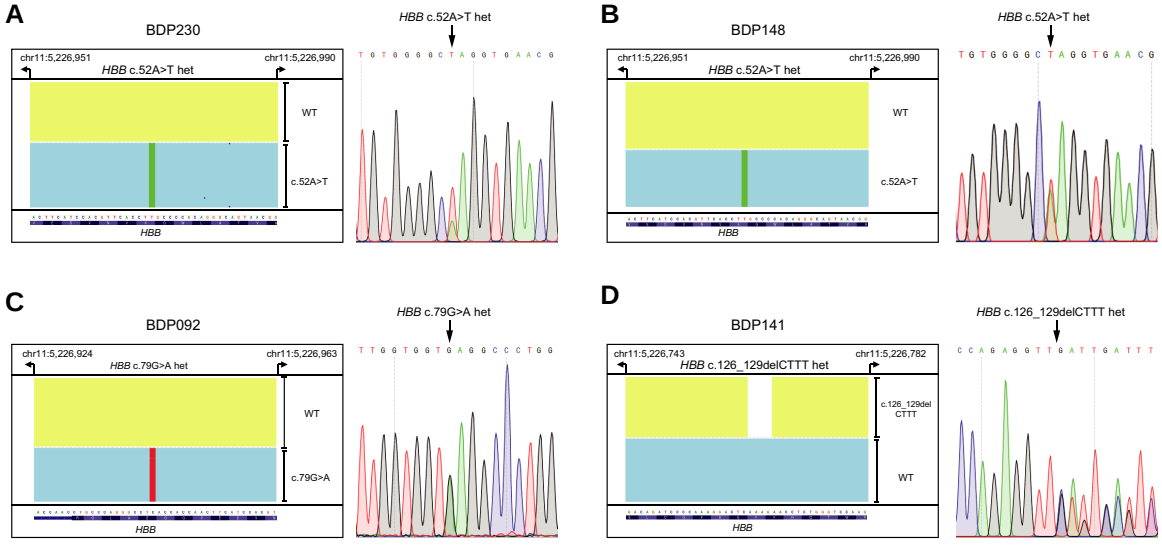

Fig. 2. Discordant common thalassemia variants between CATSA and PCR-based methods. Integrative Genomics Viewer plots of CATSA sequencing reads and Sanger sequencing showing the *HBB* c.52A > T variant in fetus BDP230 (A), *HBB* c.52A > T variant in fetus BDP148 (B), *HBB* c.79G > A variant in fetus BDP092 (C), and *HBB* c.126_129delCTTT variant in fetus BDP141 (D). Each colored area represents one chromosome and each vertical line represents one SNV.

variants were included in the HbVar or Ithanet database, but there was not sufficient evidence for clinical significance. Of note, CATSA identified 2 *HBB* variants including c.52A > T and c.315 + 308delA in fetus BDP130 and directly showed that the 2 variants were in transconfiguration (Fig. 3E). Fetus BDP130 was delivered normally.

Discussion

Using a cohort of 278 amniotic fluid samples, the LRS-based CATSA approach demonstrated advantages in efficient and accurate variant detection and phenotypic prediction of the fetuses as compared to PCR-based methods in this multicenter retrospective study. Compared to PCR-based methods, CATSA correctly identified variants in 15 (5.40%) more samples, including 4 with common variants, 4 with rare variants, and 7 with VUSs. CATSA also corrected the predicted phenotype assigned by PCR-based methods in 8 (2.88%) fetuses, including one from β-TM to β-TT and 2 from β-TT to β-TI. In particular, CATSA could have changed the pregnancy outcome for one fetus.

CATSA showed several advantages compared to PCR-based methods. First, CATSA showed 100% accuracy while PCR-based methods had one false-positive and 2 false-negative results for common β-globin variants. A previous study showed that PCR-based methods had 9 false positive/negative results out of 1759 samples tested for thalassemia variants (26). Second, CATSA can also detect rare clinically relevant variants and VUSs. Given that CATSA additionally identified 6 VUSs in 7 fetuses, it would be of clinical significance to reclassify these variants in further studies, providing better prediction of the phenotype and more precise information for genetic counseling. Third, a single CATSA assay can comprehensively detect both α- and β-globin variants, while at least 2 PCR assays are needed to cover deletion and nondeletion common variants, respectively. Fourth, CATSA can directly determine the cis- and transconfigurations of 2 or more variants.

The incorrect genotypic results caused by PCR-based methods suggests that it may be necessary to perform a repeat experiment or confirm the results with a different assay for prenatal diagnosis when PCR-based methods are applied. In this study, CATSA detected 4 rare clinically relevant variants including α-globin triplicates. Of note, the term *rare* was used here only because these variants were not covered by PCR-based methods. The prevalence of $\alpha\alpha\alpha^{anti3.7}$ and $\alpha\alpha\alpha^{anti4.2}$ was actually as high as 1% to 2% among the general population in southern China (23, 28). Since α-globin triplicates and even quadruplets can cause mild to severe β-TI in carriers of heterozygous β^0 or β^+ variants, it would be critical to include these variants in thalassemia carrier screening and prenatal diagnosis.

In China, the routine workflow for thalassemia screening includes Hb testing followed by genetic testing using PCR-based methods. Some hospitals detect both

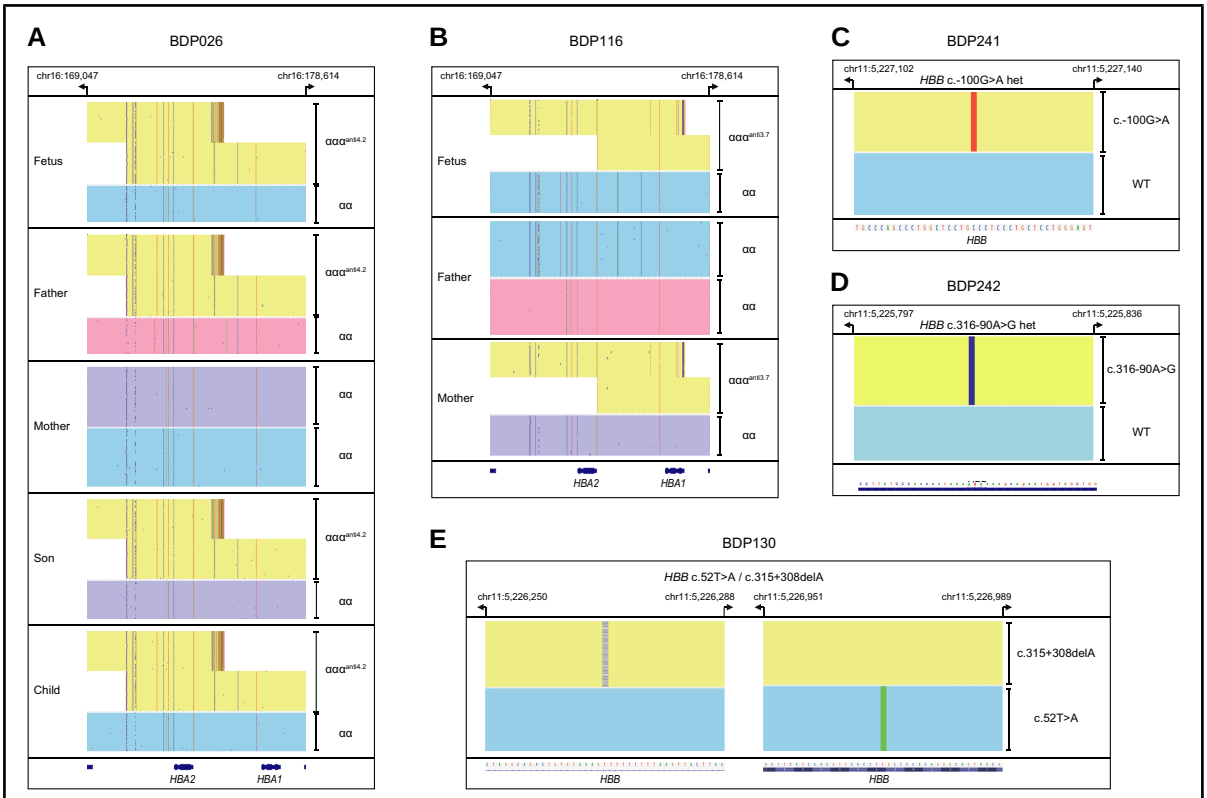

Fig. 3. Additional variants identified by CATSA. Integrative Genomics Viewer plots of CATSA reads showing the $\alpha\alpha\alpha^{anti4.2}$ variant in family BDP026 (A), $\alpha\alpha\alpha^{anti3.7}$ variant in family BDP116 (B), *HBB* c.-100G > A variant in fetus BDP241 (C), *HBB* c.316-90A > G variant in fetus BDP242 (D), and coincidence of *HBB* c.52T > A and *HBB* c.315_308delA in fetus BDP130 (E). Each colored area represents one chromosome and each vertical line represents one SNV.

α- and β-globin variants in patients with abnormal hematological parameters, while others test only α-globin or β-globin variants, depending on Hb A2 levels, for economic reasons. For example, for family BDP092, PCR-based methods were not utilized to test β-globin variants in the parents because their Hb A2 levels were not elevated. Therefore, analysis of β-globin variants was also not performed on the fetus and only the $--^{SEA}$ variant was identified by PCR-based methods, while CATSA identified an additional *HBB* c.79G > A het variant. Though this may not change the pregnancy decision, it could help for better genotype-phenotype correlation and disease management. These results showed it was necessary to test both α-globin and β-globin variants instead of only one for both carrier screening and prenatal diagnosis of thalassemia.

With the total cost of library preparation and sequencing reagent per sample reduced to less than $20, CATSA has been applied at a large scale in southern China for carrier screening. However, the turnaround time of CATSA is about 8 days and involves multiple procedures including DNA extraction, long-range PCR, library preparation, sequencing, and bioinformatics analysis, while the turnaround time of PCR-based methods is less than 3 days. Since prenatal diagnosis is very time-sensitive, it is better to perform the CATSA assay for the fetuses at an earlier gestational age. Although the total cost of library preparation and sequencing reagent per sample is low, the PacBio instrument is very expensive. Developing a less expensive benchtop platform with lower throughput in the future could be a resolution to increase the clinical applications for CATSA in areas with a high prevalence of thalassemia. Though having advantages in hardware cost, the Oxford Nanopore Technologies platform has not been fully explored for its utility and accuracy in thalassemia genetic testing.

Conclusions

This study showed that PCR-based methods, while faster than CASTA, could lead to inaccurate phenotype

prediction due to limited detection scope and accuracy. CATSA represented a more comprehensive and accurate approach, which can make a better phenotype prediction, thus providing benefit for prenatal diagnosis and genetic counseling.

Supplementary Material

Supplementary material is available at *Clinical Chemistry* online.

Nonstandard Abbreviations: $-^{SEA}$, Southeast Asian deletion; $-\alpha^{4.2}$, deletion of 4.2 kb; $-\alpha^{3.7}$, deletion of 3.7 kb; SNV, single-nucleotide variant; Hb, hemoglobin; TT, thalassemia trait; TI, thalassemia intermedia; TM, thalassemia major; RDB, reverse dot blot; LRS, long-read sequencing; CATSA, comprehensive analysis of thalassemia alleles; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; VUS, variant of uncertain significance; hom, homozygous; het, heterozygous.
Human Genes: *HBA1*, hemoglobin subunit alpha 1; *HBA2*, hemoglobin subunit alpha 2; *HBB*, hemoglobin subunit beta.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:*

Employment or Leadership: A. Mao and Y. (Yinyin) Liu are employees of Berry Genomics Corporation.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: This work was funded by the National Key Research and Development Program of China (2021YFC1005300) and the Science and Technology Innovation Program of Hunan Province (2019SK1014).
Expert Testimony: None declared.
Patents: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, preparation of manuscript, or final approval of manuscript.

Acknowledgments: We thank all subjects who participated in this study.

References

1. Taher AT, Weatherall DJ, Cappellini MD. Thalassaemia. Lancet 2018;391:155–67.
2. Williams TN, Weatherall DJ. World distribution, population genetics, and health burden of the hemoglobinopathies. Cold Spring Harb Perspect Med 2012;2:a011692.
3. Xiong F, Sun M, Zhang X, Cai R, Zhou Y, Lou J, et al. Molecular epidemiological survey of haemoglobinopathies in the Guangxi Zhuang Autonomous Region of southern China. Clin Genet 2010;78:139–48.
4. Yao H, Chen X, Lin L, Wu C, Fu X, Wang H, et al. The spectrum of alpha- and beta-thalassemia mutations of the Li people in Hainan Province of China. Blood Cells Mol Dis 2014;53:16–20.
5. Yin A, Li B, Luo M, Xu L, Wu L, Zhang L, 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.
6. Jiang F, Chen GL, Li J, Xie XM, Zhou JY, Liao C, Li DZ. Pregestational thalassemia screening in Mainland China: the first two years of a preventive program. Hemoglobin 2017;41:248–53.
7. Xu XM, Zhou YQ, Luo GX, Liao C, Zhou M, Chen PY, 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;57:517–22.
8. Modell B, Darlison M. Global epidemiology of haemoglobin disorders and derived service indicators. Bull World Health Organ 2008;86:480–7.
9. Shang X, Xu X. Update in the genetics of thalassemia: what clinicians need to know. Best Pract Res Clin Obstet Gynaecol 2017;39:3–15.
10. Vrettou C, Kakourou G, Mamas T, Traeger-Synodinos J. Prenatal and preimplantation diagnosis of hemoglobinopathies. Int J Lab Hematol 2018;40:74–82.
11. Traeger-Synodinos J, Harteveld CL, Old JM, Petrou M, Galanello R, Giordano P, et al. EMQN Best practice guidelines for molecular and haematology methods for carrier identification and prenatal diagnosis of the haemoglobinopathies. Eur J Hum Genet 2015;23:560.
12. Writing Group for Practice Guidelines for Diagnosis and Treatment of Genetic Diseases, Medical Genetics Branch of Chinese Medical Association, Shang X, Zhang X, Yang F, Xu X. [Clinical practice guidelines for alpha-thalassemia] [in Chinese]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 2020;37:235–42.
13. Writing Group for Practice Guidelines for Diagnosis and Treatment of Genetic Diseases, Medical Genetics Branch of Chinese Medical Association, Shang X, Wu X, Zhang X, Feng X, Xu X. [Clinical practice guidelines for beta-thalassemia] [in Chinese]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 2020;37:243–51.
14. Higgs DR, Engel JD, Stamatoyannopoulos G. Thalassaemia. Lancet 2012;379:373–83.
15. Xu G, Wang C, Wang J, Lin M, Chang Z, Liang J, et al. Prevalence and molecular characterization of common thalassemia among people of reproductive age in the border area of Guangxi-Yunnan-Guizhou province in southwestern China. Hematology 2022;27:672–83.
16. Peng Q, Zhang Z, Li S, Cheng C, Li W, Rao C, 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.
17. Tan M, Bai Y, Zhang X, Sun J, Huang C, Tian R, et al. Early genetic screening uncovered a high prevalence of thalassemia among 18 309 neonates in Guizhou, China. Clin Genet 2021;99:704–12.
18. He S, Li D, Yi S, Huang X, Zhou C, Chen B, et al. Molecular characterization of alpha- and beta-thalassaemia among children from 1 to 10 years of age in Guangxi, a multi-ethnic region in southern China. Front Pediatr 2021;9:724196.
19. He J, Song W, Yang J, Lu S, Yuan Y, Guo 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;19:1022–31.
20. Jiang F, Mao AP, Liu YY, Liu FZ, Li YL, Li J, et al. Detection of rare thalassemia mutations using long-read single-molecule real-time sequencing. Gene 2022;825:146438.
21. Luo S, Chen X, Zeng D, Tang N, Yuan D, Zhong Q, et al. The value of single-molecule real-time technology in the diagnosis of rare thalassemia variants and analysis of phenotype-genotype correlation. J Hum Genet 2022;67:183–95.
22. Wang M, Zhang X, Zhao Y, Lu Z, Xiao M. Prevalence and genetic analysis of

thalassemia in childbearing age population of Hainan, the free trade island in southern China. J Clin Lab Anal 2022;36: e24260.
23. Shang X, Peng Z, Ye Y, Asan, Zhang X, Chen Y, et al. Rapid targeted next-generation sequencing platform for molecular screening and clinical genotyping in subjects with hemoglobinopathies. EBioMedicine 2017;23:150–9.
24. Xu L, Mao A, Liu H, Gui B, Choy KW, Huang H, et al. Long-molecule sequencing: a new approach for identification of clinically significant DNA variants in alpha-thalassemia and beta-thalassemia carriers. J Mol Diagn 2020;22:1087–95.
25. Zhang H, Li C, Li J, Hou S, Chen D, Yan H, et al. Next-generation sequencing improves molecular epidemiological characterization of thalassemia in Chenzhou Region, P.R. China. J Clin Lab Anal 2019; 33:e22845.
26. Liang Q, Gu W, Chen P, Li Y, Liu Y, Tian M, et al. A more universal approach to comprehensive analysis of thalassemia alleles (CATSA). J Mol Diagn 2021;23:1195–204.
27. Wang W, Ma ESK, Chan AYY, Prior J, Erber WN, Chan LC, 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.
28. Long J, Liu E. The carriage rates of alphaalphaalpha(anti3.7), alphaalphaalpha(anti4.2), and HKalphaalpha in the population of Guangxi, China measured using a rapid detection qPCR system to determine CNV in the alpha-globin gene cluster. Gene 2021;768: 145296.