Li and Ye *Molecular Cytogenetics* (2024) 17:32
https://doi.org/10.1186/s13039-024-00701-4

Molecular Cytogenetics

REVIEW Open Access

Application of third-generation sequencing technology in the genetic testing of thalassemia

Weihao Li[1] and Yanchou Ye[2*]

Abstract
Thalassemia is an autosomal recessive genetic disorder and a common form of Hemoglobinopathy. It is classified into α-thalassemia and β-thalassemia. This disease is mainly prevalent in tropical and subtropical regions, including southern China. Severe α-thalassemia and intermediate α-thalassemia are among the most common birth defects in southern China. Intermediate α-thalassemia, also known as Hb H disease, is characterized by moderate anemia. Severe α-thalassemia, also known as Hb Bart's Hydrops fetalis syndrome, is a fatal condition. Infants with severe β-thalassemia do not show symptoms at birth but develop severe anemia later, requiring expensive treatment. Most untreated patients with severe β-thalassemia die in early childhood. Screening for thalassemia carriers and genetic diagnosis in high-prevalence areas significantly reduce the incidence of severe thalassemia. This review aims to summarize the genetic diagnostic approaches for thalassemia. Conventional genetic testing methods can identify 95–98% of thalassemia carriers but may miss rare thalassemia genotypes. Third-Generation Sequencing offers significant advantages in complementing other genetic diagnostic approaches, providing a basis for genetic counseling and prenatal diagnosis.

Keywords A-thalassemia, B-thalassemia, Genetic diagnosis, Third-generation sequencing, Rare thalassemia genotypes

Background

Thalassemia is an inherited and autosomal recessive disease, and is the most prevalent monogenic disorder worldwide, affecting a larger population. Thalassemia can be classified into α-thalassemia and β-thalassemia, which occur due to an imbalance in the ratio of α to β-globin chains [1]. It is caused by mutation or deletion of α or β- globin chains, however, it can also result due to molecular defects in the α- or β- globin genes or in the cluster of these globin genes such locus control region or main regulatory element [2]. α-thalassemia is primarily caused by abnormal synthesis of α-globin chains in *HBA1* gene (OMIM*141800) and *HBA2* gene (OMIM*141850), with four normal α-globin genes present in individuals. β-thalassemia is primarily caused by abnormalities in the β- globin chains in the *HBB* gene (OMIM*141900), with two normal β-globin genes present in individuals.

α-thalassemia is caused by α-globin gene deletion/mutation which results in reduced or absent production of α-globin chains [3]. α-globin gene has 4 alleles and disease severity ranges from mild to severe depending on the number of deletions of the alleles. Four allele deletion

*Correspondence:
Yanchou Ye
yeyanchou@sysush.com
[1]Reproductive Medicine Center, The Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen, China
[2]Prenatal Diagnostic Center, Medical Genetics Center, Department of Obstetrics and Gynaecology, The Seventh Affiliated Hospital of Sun Yat-sen University, 628 Zhenyuan Rd., Guangming Dist., Shenzhen, PR China.

is the most severe form in which no α-globin genes are produced and the excess γ-globin chains (present during the fetal period) form tetramers. It is incompatible with life and results in hydrops fetalis. Based on the type of α-globin gene mutation, α-globin genes can be classified into deletions and non-deletions. Deletions are further categorized as α^+ thalassemia (deletion of one α-globin gene, -α/) and α^0 thalassemia (deletion of two α-globin genes, --/). Non-deletions involve point mutations in the α-globin gene or the loss of several nucleotides ($\alpha^T\alpha$ or $\alpha\alpha^T$). β-thalassemia results from point mutations in the β-globin gene [4]. It is divided into three categories based on the zygosity of the β-gene mutation. A heterozygous mutation (β^+ thalassemia) results in β-thalassemia minor in which β chains are underproduced. It is mild and usually asymptomatic. β-thalassemia major is caused by a homozygous mutation (β^0 thalassemia) of the β-globin gene, resulting in the total absence of β chains. It manifests clinically as jaundice, growth retardation, hepatosplenomegaly, endocrine abnormalities, and severe anemia requiring life-long blood transfusions [5]. The condition in between these two types is called β-thalassemia intermedia with mild to moderate clinical symptoms. It is now common practice to classify patients as having transfusion-dependent β-thalassemia (TDT) or non-transfusion-dependent β-thalassemia (NTDT) [6]. Patients with TDT include β-thalassemia major, a few β-thalassemia intermediate, and severe forms of HbE/β-thalassemia. The NTDT Patients include β-thalassemia intermediate and mild moderate forms of HbE/β-thalassemia [7].

Thalassemia is more prevalent in the coastal regions of southern China, with Guangxi, Guangdong, Hainan being high-risk areas. The average carrier rate of thalassemia in the general population is approximately 16.45 ~ 24.14% [8–17](Table 1). The spectrum of thalassemia gene mutations exhibits distinct regional and ethnic characteristics. The Human Hemoglobin Variants and Thalassemia (HbVar, http://globin.cse.psu.edu/) contains over 1800 gene mutations, with the majority being related to hemoglobin variants and approximately 500 associated with thalassemia. The treatment of moderate or severe thalassemia is costly, and currently, there are no effective treatment options available, resulting in low cure rates and imposing a heavy burden on families and the healthcare system. While thalassemia is difficult to cure, it can be prevented and controlled. Pre-marital, pre-pregnancy, or prenatal screening is crucial for the prevention of moderate and severe thalassemia. China has established a comprehensive screening and genetic diagnosis process for thalassemia prevention and control [18]. Couples who carry the same kind of thalassemia genes with a normal phenotype have a 25% chance of having a child with minor, intermediate or severe thalassemia whether they carry the same mutation or different mutation. Therefore, pre-marital and pre-pregnancy screening, as well as prenatal genetic diagnosis, are key in preventing the birth of infants with moderate to severe thalassemia.

Table 1 Statistical of Thalassemia carrier rate in southern China

PROVINCE/Region	Thalassemia frequency, %
Guang Xi	24.14 [6]
Guang Dong	16.45 [7]
Hai Nan	21.37 [8]
Hu Nan	7.1 [9]
Gan Nan	14.55 [10]
Fu Jian	6.8 [11]
Yun Nan	12.5 [12]
Gui Zhou	14.66 [13]
Hong Kong	8.4 [14]
Macao	9.8 [15]

Main text

Current status of traditional thalassemia screening

Traditional methods for thalassemia screening and diagnosis include complete blood count (CBC), hemoglobin electrophoresis and genetic testing. Key indicators for thalassemia screening in CBC include mean corpuscular volume (MCV) and mean corpuscular hemoglobin (MCH), while key indicators for hemoglobin electrophoresis screening include Hemoglogbin A2 (HbA_2), Hemoglogbin A (HbA), Hemoglogbin F (HbF) and other abnormal hemoglobin variants [19]. The reference values for MCV and MCH in normal adults are 80–97 fL and 27.5–33.5 pg, respectively, and the relative content of HbA_2 is 2.5-3.5%. Individuals with MCV < 80 fL and/or MCH < 27 pg are considered positive, indicating characteristics of microcytic hypochromic anemia, after ruling out iron deficiency as the cause of anemia. Data shows that thalassemia carriers can be detected even in individuals with completely normal MCV and MCH, with a missed diagnosis rate of approximately 21% [20]. On the other hand, hemoglobin electrophoresis has a missed diagnosis rate of about 57% for a small number of carriers of mild and silent β-thalassemia, as well as some carriers of mild and silent α-thalassemia genes [20]. The α-globin gene is located on 16p13.3, and 104 mutations have been identified in the Chinese population. The three common deletion types [$-\alpha^{3.7}$ (rightward), $-\alpha^{4.2}$ (left-ward), $--^{SEA}$ (Southeast Asian)] and three non-deletion types [Constant Spring (Hb CS or HBA2: c.427T > C), Hb Quong Sze (Hb QS or HBA2: c.377 > C) and Hb Westmead (Hb WS or HBA2: c.369 C > G)] account for over 99% of cases [8]. The β globin gene is located on 11p15.3, and 129 mutations and 16 deletions have been found in the Chinese population. The 17 common mutations account for over 98% of cases [8, 21]. Currently, routine PCR reagent

kits are used in clinical practice, including gap-PCR, PCR-reverse dot blot hybridization, and liquid bead array technology. These methods can only cover 23 hotspot mutations in thalassemia genes. Common α-globin (--^{SEA}, -$\alpha^{3.7}$, -$\alpha^{4.2}$, α^{CS}, α^{QS}, α^{WS}) and β-globin (HBB: c.316–197 C>T, HBB: c.-11_-8delAAAC, HBB: c.216_217insA, HBB: c.130G>T, HBB: c.124_127delTTCT, HBB: c.94delC, HBB: c.92+5G>C, HBB: c.92+1G>T, HBB: c.84_85insC, HBB: c.79G>A, HBB: c.45_46insG, HBB: c.2T>G, HBB: c.-78 A>G, HBB: c.-79 A>G, HBB: c.-80T>C, HBB: c.-82 C>A), which account for approximately 98% of the overall mutation spectrum in the Chinese population. Rare thalassemia other than the 23 common mutations may account for 2%. Next-Generation Sequencing can increase the detection of mutations and detect abnormal hemoglobin variants additionally, and may detect more deletion types of thalassemia or ααα and αααα variations in α-thalassemia, but highly dependent on coverage, read depth, and CNV analysis tools.

Polymerase chain reaction (PCR)-based genetic testing for thalassemia

Gap-polymerase chain reaction (Gap-PCR) is a commonly used method for detecting deletion types of α-thalassemia [22]. The principle involves designing a pair of primers on either side of the deletion sequence. In the normal DNA sample, the distance between the upstream and downstream primers is far apart, resulting in a long amplification fragment that may exceed the effective amplification range and fail to generate PCR product. However, the presence of a deletion brings the ends closer together, allowing for the amplification of a specific fragment of a defined length. By using combinations of multiple primers, wild-type, heterozygous and homozygous can be differentiated based on the amplification products. However, this method can only detect known deletion types, including three α-thalassemia deletion mutations: --^{SEA}, -$\alpha^{3.7}$, and -$\alpha^{4.2}$.

Reverse dot blot (RDB) and Flow-through hybridization are based on the principle of immobilizing a series of allele-specific oligonucleotide (ASO) probes targeting known mutations on a membrane [23, 24]. The PCR amplified target products are then hybridized with the probes on the membrane, and the hybridization signals are visualized through a chemical color reaction. This method allows for the analysis of multiple point mutations that may be present in a specimen through a single hybridization reaction. These methods can currently the most commonly used technique in clinical practice. However, it can only detect known mutations, including α-globin (--^{SEA}, -$\alpha^{3.7}$, -$\alpha^{4.2}$, α^{CS}, α^{QS}, α^{WS}) and β-globin (HBB: c.316–197 C>T, HBB: c.216_217insA, HBB: c.130G>T, HBB: c.124_127delTTCT, HBB: c.94delC, HBB: c.92+5G>C, HBB: c.92+1G>T, HBB: c.84_85insC, HBB: c.79G>A, HBB: c.45_46insG, HBB: c.2T>G, HBB: c.-11_-8delAAAC, HBB: c.-78 A>G, HBB: c.-79 A>G, HBB: c.-80T>C, HBB: c.-82 C>A).

The bead-based suspension array consists mainly of microspheres, with each microsphere fixed with different probe molecules [25]. The principle involves antigen-antibody, enzyme-substrate, and ligand-receptor binding reactions, as well as nucleic acid hybridization reactions, on microspheres with different fluorescent codes. Laser detection is used to measure the fluorescent codes of the microspheres and report the fluorescence signals for qualitative and quantitative analysis. This method can only detect known thalassemia mutation genotypes.

Melting curve analysis (MCA) refers to the change in absorbance (A260) plotted against temperature during the continuous heating denaturation process of DNA molecules, typically resulting in an S-shaped curve [26]. The temperature at which the absorbance reaches 50% of its maximum value is called the melting temperature (Tm), at which 50% of the double-stranded DNA becomes single-stranded. Due to changes in the DNA sequence caused by mutations, the Tm value and melting curve also change. By comparing the Tm values and melting curves of wild-type and mutant DNA and referring to existing data in databases, the specific genotype of the mutant DNA can be determined. This method also has the advantages of being fast and simple, but it can only detect known mutation types.

Multiplex ligation dependent probe amplification(MLPA) involves hybridizing probes with target sequences, followed by ligation and PCR amplification [27, 28]. The PCR products are separated by capillary electrophoresis, and the data is collected and analyzed using software to draw conclusions. This method can simultaneously detect known and unknown copy number variations (deletions or duplications), but it requires a Sanger sequencing instrument and has higher costs for reagents and consumables. The experimental procedure is more complex and requires high experimental standards.

Sanger sequencing is to design primers targeting known pathogenic genes (HBA2, HBA1, HBB) and directly sequence the PCR amplified products. This method can detect known and unknown point mutations and is still considered the gold standard for sequencing technology. However, it cannot detect large deletions and has low throughput, making it less widely applicable.

Next-generation sequencing (NGS) can generate a large amount of genomic data compared to conventional PCR methods and has significantly improved detection range (mainly SNP/Indel) and accuracy compared to traditional screening methods [20](Supplemental Table S1). However, NGS is limited by read length and cannot

directly obtain complete sequences of the α and β-genes. Therefore, it cannot effectively detect tandem repeat sequences and complex structural rearrangements. NGS relies on PCR amplification, which can result in uneven coverage in regions with extreme GC content. As a result, the sensitivity for detecting structural variations, such as copy number variations, is unstable. It is often necessary to combine NGS with Gap-PCR strategies to improve the detection rate [29]. Furthermore, the application of NGS in the diagnosis of thalassemia faces challenges such as the inability to distinguish homologous region variations in the *HBA2* and *HBA1* genes.

Third-generation sequencing (TGS)

TGS, also known as long-read sequencing (LRS), includes nanopore sequencing technology from Oxford Nanopore Technologies (ONT) and single molecule real time sequencing (SMRT) from Pacific Biosciences (PacBio) [30]. TGS offers several technical advantages, such as ultra-long read lengths and no base preference. It can accurately sequence complex regions with high GC content and highly repetitive regions. TGS overcomes the limitations of NGS in thalassemia gene detection and demonstrates significant advantages in genome assembly, full-length transcript analysis, detection of structural variations, haplotype construction, and epigenetic analysis [31] (Supplemental Table S2).

The ONT sequencing device is lightweight and capable of generating extremely long sequencing reads, with the longest reads reaching up to 300,000 to 400,000 bases. The core component of the sequencing process is the protein nanopore. It forms a channel in a membrane, and the nanopore is embedded in the membrane where a voltage can be applied. When ions pass through the protein pore, moving from one side of the membrane to the other, the flow of ions is hindered when a single-stranded DNA molecule passes through the nanopore. The degree of hindrance varies depending on the different bases. Transducer proteins detect changes in electrical current caused by the five bases (A, C, G, T, U). By analyzing the spectrum of current changes and applying pattern algorithms, the sequence of bases can be determined (Fig. 1). Zhou et al. [32] reported the combination of nanopore sequencing technology with Gap-PCR to target the human globin gene, specifically addressing the challenges posed by high homology and complex secondary structures affecting amplification efficiency. This technology can effectively detect α-deletion types ($--^{SEA}$/, $-\alpha^{3.7}$/, $-\alpha^{4.2}$/, $--^{THAI}$/), β-deletion types in Chinese $G\gamma^{+}(A\gamma\delta\beta)^{0}$ and SEA-HPFH, as well as structural variations of α-globin trios (HKαα, $\alpha\alpha\alpha^{anti3.7}$, $\alpha\alpha\alpha^{anti4.2}$) and various point mutations in α and β-thalassemia genes. Comprehensive analysis of hemoglobin gene variations can be achieved through accurate genotype description and characterization.

The PacBio single-molecule fluorescent sequencing platform has the longest read lengths, reaching up to 20 kb to 30 kb bases, with an average of 10 kb to 15 kb, which is over 100 times longer than the next generation sequencing. The process involves extracting nucleic acids from the sample and performing long fragment multiplex PCR amplification of the target regions. A library is then constructed, and sequencing is carried out using the PacBio sequencing platform. The PacBio sequencing platform has fast sequencing speed, but it has a relatively high error rate, reaching 10–15%, but the errors

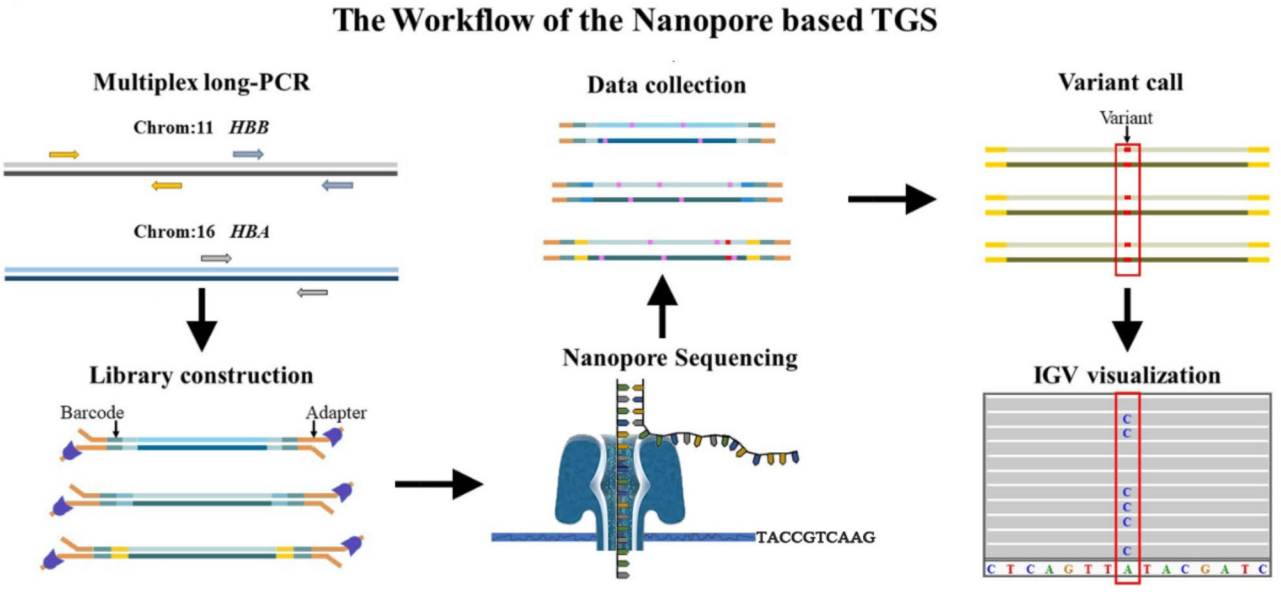

Fig. 1 Workflow and amplicons of ONT TGS method for detection of hemoglobinopathy variants

are random and can be corrected by taking multiple measurements of the same sequence [33]. To address this, the PacBio sequencing platform utilizes circular consensus sequencing (CCS) by sequencing single DNA molecules in a rolling circle manner to achieve self-correction of sequencing errors. When the sequencing depth reaches 10x, the accuracy of CCS reads can reach 99.9% [34]. The PacBio sequencing technology for detecting mutations in α-thalassemia and β-thalassemia genes is called Comprehensive Analysis of Thalassemia Alleles (CATSA) (Fig.2). CATSA meets the requirements of clinical testing and provides ultra-long read lengths (>10 kb) while maintaining sequencing accuracy (QV30, >99.9% accuracy). Leveraging the advantages of long-read sequencing in third-generation sequencing, CATSA can detect the cis/trans relationship of two mutations in the HBA2, HBA1, or *HBB* genes, accurately differentiate highly homologous *HBA1* and *HBA2* genes, achieve single-base resolution, as well as detect α-thalassemia ααα and αααα [35, 36].

Traditional screening and diagnosis of thalassemia require the combination of multiple testing techniques, such as CBC, hemoglobin electrophoresis, Gap-PCR, Sanger sequencing etc(Table 2), but all these methods have limitations. Several clinical studies have confirmed that CATSA enables comprehensive detection in a single test, covering all genes and mutation types of α-thalassemia and β-thalassemia, accurately detecting thalassemia-related variant sites. CATSA addresses the limitations of traditional screening methods, such as known detection range, high false-negative rate, low testing throughput, and difficulties in gene typing, significantly improving the clinical diagnostic rate of thalassemia. Compared to the combined use of traditional PCR, NGS, MLPA, and other methods for thalassemia gene testing, TGS can still improve the diagnostic rate of rare types of α-thalassemia and β-thalassemia, respectively. Xu L et al. [37] published an important retrospective study confirming that CATSA has comprehensive site coverage, high accuracy, high throughput, and the ability to achieve precise gene typing. It can be applied in clinical settings for large-scale screening of thalassemia carriers. The study included 74 clinical samples, including 64 positive samples detected by conventional PCR and 10 negative samples detected by conventional PCR, which were retrospectively tested using CATSA. Inconsistent results were validated using "gold standard" methods such as Sanger sequencing. The results showed 100% accuracy and repeatability of CATSA. One false-negative result was found using conventional PCR, where CATSA detected $-\alpha^{3.7}/\alpha\alpha$ while the traditional PCR result was negative. The CATSA result was validated to be correct. Lingqian Wu et al. [38] published an important

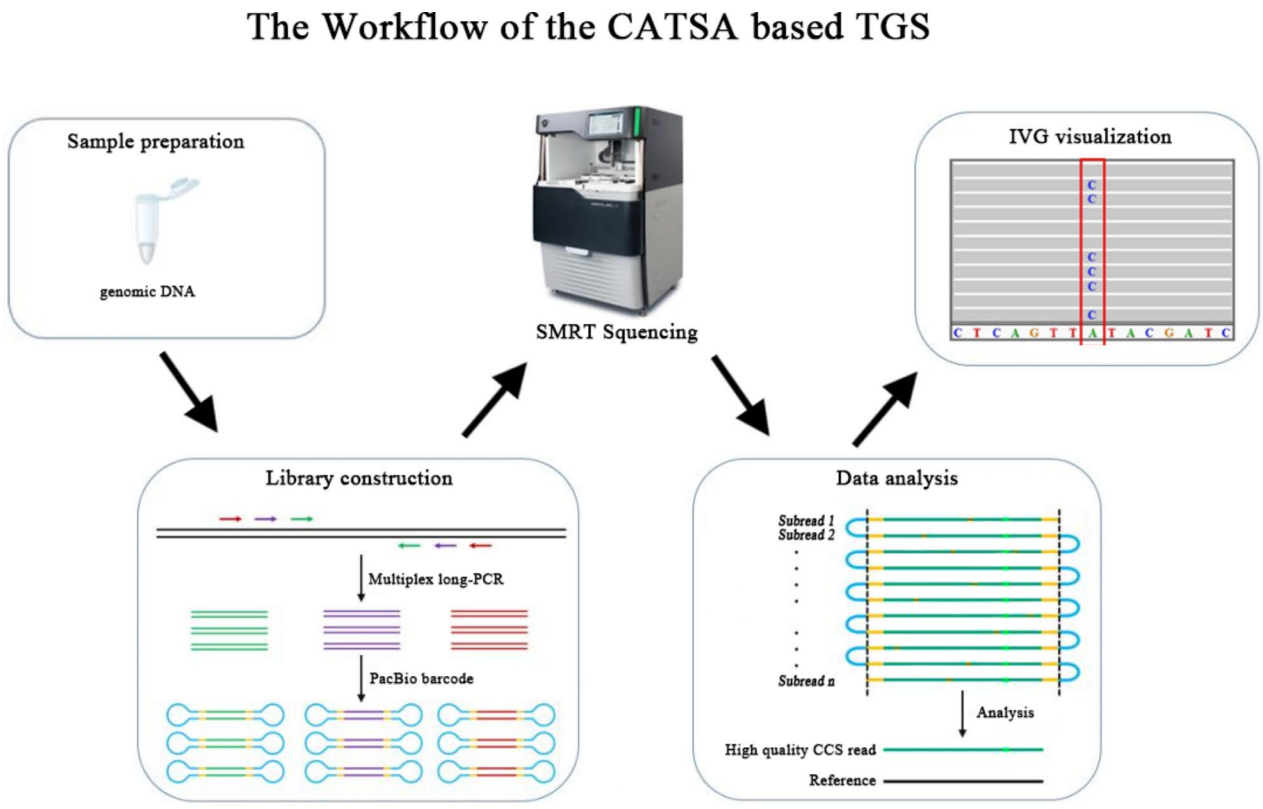

Fig. 2 Workflow and amplicons of CATSA based TGS method detection of hemoglobinopathy variants

Table 2 Genetic testing methods for thalassemia

METHODS	Detection Principle	Detection range
Gap-PCR	Primers are designed on either side of the target deletion fragment, the primers upstream and downstream of the wild-type sequence are too far apart to amplify the product, and if the deletion is present, a fragment of a specific length can be amplified.	known deletion types.
RDB	Based on the principle of immobilizing a series of allele-specific oligonucleotide (ASO) probes targeting known mutations on a membrane. The PCR-amplified target products are then hybridized with the probes on the membrane, and the hybridization signals are visualized through a chemical color reaction.	known mutation types.
MCA	The fluorescence intensity change of the temperature is monitored in real time PCR, and the melting curve is generated to analyze the dissolution temperature (Tm value) of the hybrids formed by the probe and the target sequences, and different target sequences have different Tm values.	Known deletion and mutation types, unknown mutation types within the coverage of probes.
MLPA	The probe is hybridized to the target sequence, followed by ligation, PCR amplification, and the product is passed through capillary electrophoresis to achieve detection and quantitative analysis of the target sequence.	gene copy number variations (CNVs).
Sanger sequencing	Primers were designed at both ends of the sequence to be tested, and after PCR amplification, the sequence was sequenced by double deoxy chain termination method, and then analyzed against the reference sequence.	Known or unknown mutation types within the amplified sequences.
NGS	Mutation types are captured and enriched by target region probes, and screened by library preparation, Next-Generation Sequencing and bioinformatics analysis.	Known or unknown mutation types within the amplified sequences.
TGS	The target region of the α-globin genes clusters is amplified by long-fragment multiplex PCR amplification and Third-Generation Sequencing, deletion variants, mutation variants, and complex structural variants were detected.	Known or unknown mutation, Indel and structural variations.

prospective clinical study, including 1759 prospective clinical samples with positive results in complete blood count or hemoglobin testing. The study was conducted in a double-blind manner and validated using various "gold standard" methods such as Sanger sequencing and Gap-PCR. The study confirmed the 100% accuracy of CATSA. CATSA detected 33 more positive results compared to traditional PCR technology and identified 1 false-negative and 8 false-positive results in traditional PCR testing. CATSA detected 3 types of single nucleotide variants (SNVs) in the *HBA2* and *HBA1* genes and 8 types of insertions and deletions (Indels) compared to traditional PCR technology, resulting in a 2.8% increase in the detection rate of α-thalassemia. It was found that traditional PCR testing misclassified 3 cases of HKαα/αα as $-\alpha^{3.7}$/αα. CATSA detected 6 cases of SNVs/indels in the *HBB* gene compared to traditional PCR, resulting in a 2.1% increase in the detection rate of β-thalassemia. Li Du et al. [39] used SMRT sequencing, which enables long-read single-molecule sequencing to detect rare and complex variations in Chinese patients with microcytic hypochromic anemia who tested negative for common mutations. Considering the limitations of traditional methods in detecting rare types of thalassemia, which may lead to misdiagnosis or missed diagnosis, SMRT sequencing has been proven to be a feasible method for discovering rare and complex variations in thalassemia.

Conclusions

The large-scale clinical application of TGS technology is still limited. The bioinformatics algorithms based on the TGS platforms are not yet comprehensive enough, and standardized analysis methods are lacking. The interpretation and clinical consultation of the large amount of data obtained pose a major challenge. Currently, the cost of sequencing is reduced by targeting and enriching the regions of interest, thereby reducing data redundancy. It is believed that with platform updates and further development and application of related algorithms and analysis software, the cost of TGS will gradually decrease. With its technical advantages, it will overcome the detection bottleneck of NGS and bring new opportunities for genetic disease diagnosis. In summary, the application of TGS technology in the detection of thalassemia allows for the direct acquisition of complete sequences of α and β-thalassemia. It can detect structural variations in thalassemia genes that cannot be detected by conventional PCR and NGS, thereby increasing the detection rate of thalassemia and reducing the rate of missed diagnosis and misdiagnosis. It represents a new model for thalassemia testing, which can better address the correlation between hematological phenotypes and genotypes, and is helpful for genetic counseling and prenatal diagnosis.

Abbreviations

Gap-PCR	Gap polymerase chain reaction
RDB	Reverse dot blot hybridization
PMCA	Probe melting curve analysis for genotyping

MLPA	Multiplex ligation-dependent probe amplification
NGS	Next-generation sequencing
TGS	Third-generation sequencing
LRS	Long-read sequencing
ONT	Oxford nanopore technologies
SMRT	Single-molecule real-time sequencing
CCS	Circular consensus sequencing
CATSA	Comprehensive analysis of thalassemia alleles
SNVs	Single nucleotide variants
Indels	Insertions and deletions

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13039-024-00701-4.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
Not applicable.

Author contributions
Weihao Li Writing- original draft preparation，Yanchou Ye Writing- review and editing. All authors reviewed the manuscript.

Funding
There is no funding supported for this study.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 25 June 2024 / Accepted: 2 December 2024
Published online: 18 December 2024

References

1. Kattamis A, Kwiatkowski JL, Aydinok Y. Thalassaemia Lancet. 2022;399(10343):2310–24.
2. Ji X, et al. Both locus control region and proximal regulatory elements direct the developmental regulation of beta-globin gene cluster. J Cell Biochem. 2000;76(3):376–85.
3. Harewood J, Azevedo AM. Alpha Thalassemia. 2024.
4. Needs T, Gonzalez-Mosquera LF, Lynch DT. Beta Thalassemia. 2024.
5. Farmakis D, et al. Thalassaemia International Federation guidelines for the management of transfusion-dependent thalassemia. Hemasphere. 2021;2022(68):pe732.
6. Taher AT, Musallam KM, Cappellini MD. beta-thalassemias. N Engl J Med. 2021;384(8):727–43.
7. Musallam KM, et al. 2021 update on clinical trials in beta-thalassemia. Am J Hematol. 2021;96(11):1518–31.
8. Xiong F, et al. Molecular epidemiological survey of haemoglobinopathies in the Guangxi Zhuang Autonomous Region of southern China. Clin Genet. 2010;78(2):139–48.
9. Li B, et al. High prevalence of thalassemia in migrant populations in Guangdong Province, China. BMC Public Health. 2014;14:905.
10. Lai Y, et al. Molecular spectrum of thalassemia in tropical Hainan Island of southern China: high allele frequency with low health burden. J Genet Genomics. 2022;49(12):1162–4.
11. Xi H, et al. Epidemiological survey of Hemoglobinopathies based on next-generation sequencing platform in Hunan Province, China. Biomed Environ Sci. 2023;36(2):127–34.
12. Yang T, et al. Next-generation sequencing analysis of the molecular spectrum of Thalassemia in Southern Jiangxi, China. Hum Genomics. 2023;17(1):77.
13. Huang H, et al. Molecular characterization of Thalassemia and hemoglobinopathy in Southeastern China. Sci Rep. 2019;9(1):3493.
14. Zhang J, et al. The spectrum of alpha- and beta-thalassemia mutations in Yunnan Province of Southwestern China. Hemoglobin. 2012;36(5):464–73.
15. Wu LS, et al. Prevalence of thalassemia-carrier couples and fertility risk assessment. Int J Hematol. 2024;119(4):374–82.
16. Tamagnini GP, Kuam B, Fai WK. Congenital anemias in Macau. Hemoglobin. 1988;12(5–6):637–43.
17. Lau YL, et al. Prevalence and genotypes of alpha- and beta-thalassemia carriers in Hong Kong -- implications for population screening. N Engl J Med. 1997;336(18):1298–301.
18. Wang WD, et al. Thalassaemia in China. Blood Rev. 2023;60:101074.
19. Li LY, et al. [The value of MCV, MCH and HbA(2) in laboratory screening of Thalassemia]. Zhonghua Fu Chan Ke Za Zhi. 2012;47(2):96–100.
20. He 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(9):1022–31.
21. Xu XM, 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(5):517–22.
22. Zhou YQ, et al. Evaluation of clinical application of gap-PCR as a routine method for alpha-thalassemia carrier detection. Di Yi Jun Yi Da Xue Xue Bao. 2002;22(5):434–6.
23. Liu CL, et al. [Clinical value of PCR-flow fluorescence hybridization in prenatal genetic diagnosis of Thalassemia]. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2021;29(1):221–7.
24. Liang HF, et al. Clinical validation of a single-tube PCR and reverse dot blot assay for detection of common alpha-thalassaemia and beta-thalassaemia in Chinese. J Int Med Res. 2022;50(2):3000605221078785.
25. Yin A, et al. Development of bead-based suspension array technology for the diagnosis of Thalassemia. Am J Hematol. 2014;89(12):1158–9.
26. Qin J, Nested Asymmetric A, et al. Melting curve assay for one-step genotyping of nondeletional alpha-thalassemia mutations. J Mol Diagn. 2020;22(6):794–800.
27. Colosimo A, et al. Application of MLPA assay to characterize unsolved alpha-globin gene rearrangements. Blood Cells Mol Dis. 2011;46(2):139–44.
28. Jomoui W, et al. Revisiting and updating molecular epidemiology of alpha-thalassemia mutations in Thailand using MLPA and new multiplex gap-PCR for nine alpha-thalassemia deletion. Sci Rep. 2023;13(1):9850.
29. Tan M, et al. Early genetic screening uncovered a high prevalence of thalassemia among 18 309 neonates in Guizhou, China. Clin Genet. 2021;99(5):704–12.
30. Zeng P, et al. Comparison of ONT and CCS sequencing technologies on the polyploid genome of a medicinal plant showed that high error rate of ONT reads are not suitable for self-correction. Chin Med. 2022;17(1):94.
31. Long J, et al. Third-generation sequencing: a novel tool detects complex variants in the alpha-thalassemia gene. Gene. 2022;822:146332.
32. Huang W, et al. Nanopore Third-Generation sequencing for Comprehensive Analysis of Hemoglobinopathy variants. Clin Chem. 2023;69(9):1062–71.
33. Karst SM, et al. High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. Nat Methods. 2021;18(2):165–9.
34. Wenger AM, et al. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nat Biotechnol. 2019;37(10):1155–62.
35. Liu Q, et al. Identification of rare thalassemia variants using third-generation sequencing. Front Genet. 2022;13:1076035.
36. Feng J, et al. The comprehensive analysis of Thalassemia alleles (CATSA) based on single-molecule real-time technology (SMRT) is a more powerful strategy in the diagnosis of thalassemia caused by rare variants. Clin Chim Acta. 2023;551:117619.
37. Xu L, 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(8):1087–95.

38. Liang Q, et al. A more Universal Approach to Comprehensive analysis of thalassemia alleles (CATSA). J Mol Diagn. 2021;23(9):1195–204.
39. Bao X, et al. Identification of four novel large deletions and complex variants in the alpha-globin locus in Chinese population. Hum Genomics. 2023;17(1):38.

Publisher's note