



## OPEN ACCESS

## EDITED BY

Giulia Ceglie,  
Bambino Gesù Pediatric Hospital (IRCCS), Italy

## REVIEWED BY

Barbara Eleni Rosato,  
University of Naples Federico II, Italy

## \*CORRESPONDENCE

Baoheng Gui  
✉ BaohengGui@yeah.net

RECEIVED 03 April 2023

ACCEPTED 19 June 2023

PUBLISHED 06 July 2023

## CITATION

Zhan L, Gui C, Wei W, Liu J and Gui B (2023)  
Third generation sequencing transforms the  
way of the screening and diagnosis of  
thalassemia: a mini-review.  
Front. Pediatr. 11:1199609.  
doi: 10.3389/fped.2023.1199609

## COPYRIGHT

© 2023 Zhan, Gui, Wei, Liu and Gui. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in  
other forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Third generation sequencing transforms the way of the screening and diagnosis of thalassemia: a mini-review

Lixia Zhan<sup>1,2</sup>, Chunrong Gui<sup>3,4</sup>, Wei Wei<sup>3,4</sup>, Juliang Liu<sup>3,4</sup>  
and Baoheng Gui<sup>1,3,4\*</sup>

<sup>1</sup>The Second School of Medicine, Guangxi Medical University, Nanning, China, <sup>2</sup>Child Healthcare Department, The Second People's Hospital of Beihai, Beihai, China, <sup>3</sup>Center for Medical Genetics and Genomics, The Second Affiliated Hospital of Guangxi Medical University, Nanning, China, <sup>4</sup>The Guangxi Health Commission Key Laboratory of Medical Genetics and Genomics, The Second Affiliated Hospital of Guangxi Medical University, Nanning, China

Thalassemia is an inherited blood disorder imposing a significant social and economic burden. Comprehensive screening strategies are essential for the prevention and management of this disease. Third-generation sequencing (TGS), a breakthrough technology, has shown great potential for screening and diagnostic applications in various diseases, while its application in thalassemia detection is still in its infancy. This review aims to understand the latest and most widespread uses, advantages of TGS technologies, as well as the challenges and solutions associated with their incorporation into routine screening and diagnosis of thalassemia. Overall, TGS has exhibited higher rates of positive detection and diagnostic accuracy compared to conventional methods and next-generation sequencing technologies, indicating that TGS will be a feasible option for clinical laboratories conducting in-house thalassemia testing. The implementation of TGS technology in thalassemia diagnosis will facilitate the development of effective prevention and management strategies, thereby reducing the burden of this disease on individuals and society.

## KEYWORDS

third-generation sequencing, nanopore sequencing, single-molecule real-time sequencing, thalassemia, screening, diagnosis

## 1. Introduction

Thalassemia is a group of inherited hemolytic anemia diseases resulting from genetic mutations that lead to the absence or deficiency of the synthesis of one or more globin chains in hemoglobin. The clinical symptoms of thalassemia were initially described by Dr. Thomas B. Cooley in 1925 (1). In 1936, Whipple and Bradford created the term “thalassemia” for this anemia disease, deriving it from the Greek words “thalassa”, referring to “sea”, and “haima”, meaning “blood”, as it was primarily prevalent in anemic patients from Mediterranean countries (2). Subsequently, the disease was found to be widespread in the Middle East, India, and Southeast Asia.

Thalassemia is classified into two main types based on the affected globin chains:  $\alpha$ - and  $\beta$ -thalassemia. Clinically, thalassemia can be further classified into the following types: (1) Thalassemia major: It is the most severe form of thalassemia, characterized by severe anemia, growth retardation, hepatosplenomegaly, and skeletal abnormalities. (2) Thalassemia minor: It is a relatively mild form of thalassemia, and patients may only

have mild anemia and minor clinical symptoms. (3) *Thalassemia intermedia*: Patients exhibit a higher degree of anemia compared to *thalassemia minor* but less severe than *thalassemia major*. Patients with *thalassemia major* usually require regular blood transfusions to replenish healthy red blood cells to relieve the symptoms of anemia. However, prolonged transfusions may lead to iron overload and require iron chelation therapy to prevent iron toxicity.

Currently, Hematopoietic Stem Cell Transplantation (HSCT) is a potentially curative treatment for *thalassemia* by transplanting normal hematopoietic stem cells into the patient to replace the damaged hematopoietic system. However, HSCT requires a matched donor and carries certain risks and complications. Recently, there have been significant advancements in gene therapy and gene editing approaches for *thalassemia* treatment. Notably, in August 2022, the U.S. Food and Drug Administration (FDA) approved Zynteglo, a gene therapy developed by bluebird bio incorporation, for the treatment of  $\beta$ -*thalassemia* patients. However, Zynteglo has gained attention due to its high cost, priced at \$2.8 million, making it the most expensive drug on the market. Moreover, on June 8th, 2023, the FDA accepted Biologics License Applications for exagogene autotemcel (exel cel), the first medical product of CRISPR gene-editing, for severe transfusion-dependent  $\beta$ -*thalassemia*.

Considering the cost of treatment, *thalassemia* is still a significant social and economic burden, especially in highly endemic areas, with an estimated world prevalence of 5%–7% carriers and an annual birth rate of over 2.4% (3). In China, there are around 30 million *thalassemia* carriers and about 300,000 patients with severe and intermediate forms, with an increasing rate of approximately 10% per year (4). Highly endemic areas in China include Guangxi, Guangdong, Hainan, Yunnan, Sichuan, Hunan, and Jiangxi, with a population carriage rate of 1%–23% (4). Therefore, implementing screening programs for *thalassemia* is crucial to prevent new cases and maintain the desired annual birth rate of *thalassemia major*, decreasing lifetime maintenance costs for patients.

Traditional *thalassemia* screening involves a three-step workflow, starting with a full blood count (FBC) and erythrocyte morphology to observe and calculate mean corpuscular volume (MCV) and mean erythrocyte hemoglobin content (MCH) to identify low hemoglobin levels and abnormal erythrocyte content. Subsequent biochemical analysis is performed using hemoglobin (Hb) electrophoresis, high-performance liquid chromatography (HPLC), or capillary electrophoresis (CE). Confirmatory tests such as spanning break site assay ((gap-polymerase chain reaction, Gap-PCR)), reverse spot hybridization assay ((polymerase chain reaction-reverse dot blot, PCR-RDB)), multiplex linkage probe amplification (MLPA), or conventional genetic tests like direct Sanger sequencing are then used in basis of preliminary results. While these methods are considered the gold standard for *thalassemia* investigations, they are labor-intensive, and over 1,530 hemoglobin-related genomic variants have been identified to date, which can further complicate the interpretation of *thalassemia*, especially in the presence of abnormal hemoglobin genotypes or modifier genes. Moreover, traditional methods are not reliable in accurately diagnosing rare mutations, which necessitates the development of new DNA screening tools.

Next-generation sequencing (NGS), also known as second-generation sequencing (SGS), has found wide application in clinical practice for its ability to identify mutations across the whole human genome, which can reveal various genetic diseases. However, there are several limitations that hinder its application as a stand-alone technology for screening and diagnosing *thalassemia*, requiring the aid of Gap-PCR and Sanger sequencing (Figure 1) (5). Third-generation sequencing (TGS) has emerged as an alternative method that demonstrates great potential for *thalassemia* detection. This article aims to describe the latest and most widespread uses, advantages of TGS technology for *thalassemia* detection, as well as the challenges and solutions associated with its incorporation into standard screening and diagnostic protocols.

## 2. Overview of third-generation sequencing

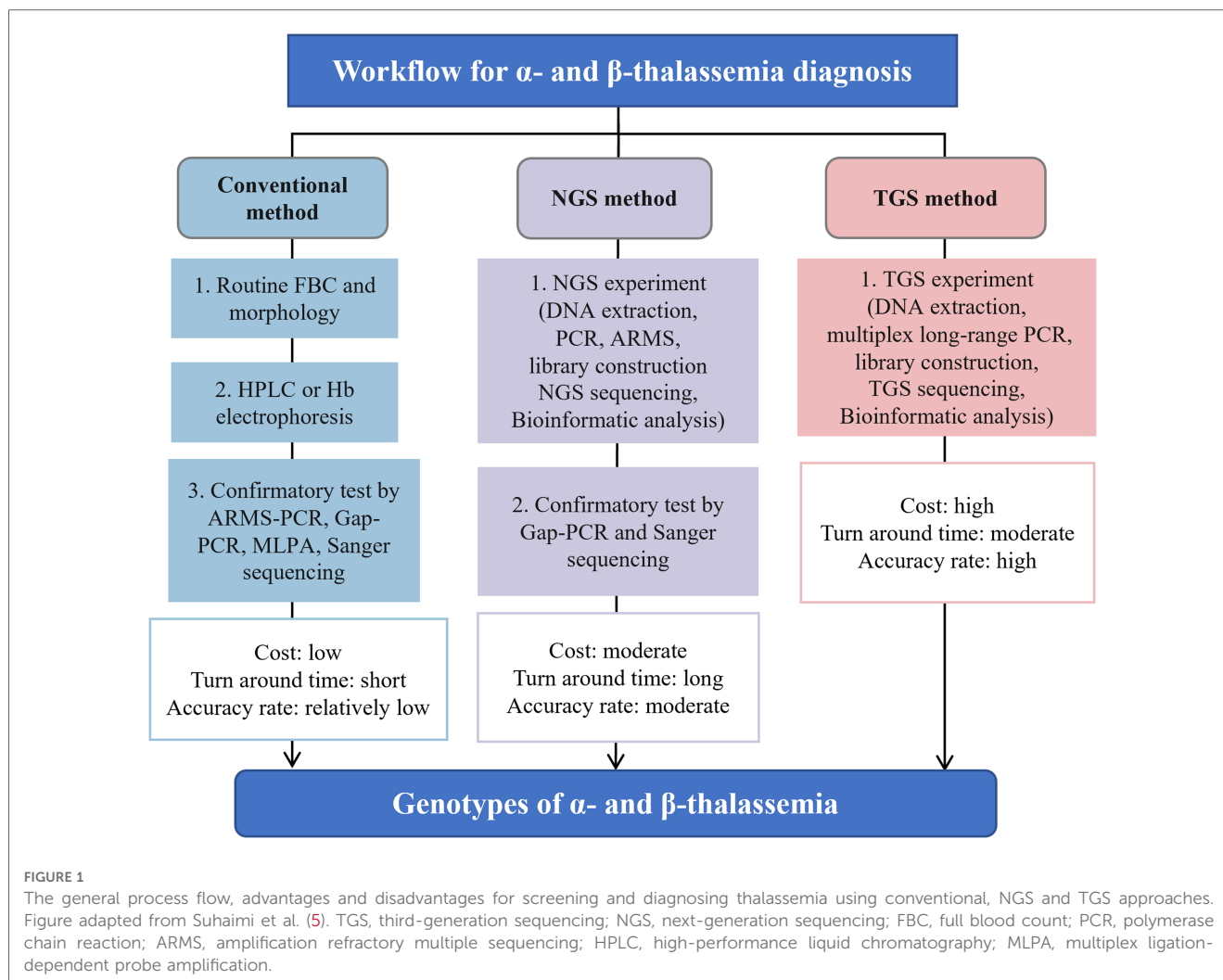
TGS is a newer sequencing technology and has rapidly advanced the field of genomics. TGS technologies differ from traditional second-generation sequencing (SGS) methods by facilitating direct sequencing of long-stranded DNA or RNA molecules without fragmentation or amplification. TGS technologies mainly include single-molecule real-time (SMRT) sequencing, nanopore sequencing, and synthetic long-read (SLR) sequencing.

The earliest TGS technology, SMRT sequencing, was developed by Pacific Biosciences (PacBio) and was introduced in 2010. SMRT sequencing detects nucleotide binding as the DNA polymerase travels along a single DNA molecule, producing long reads of up to tens of kilobases with high accuracy and low error rates. In 2014, the introduction of nanopore sequencing by Oxford Nanopore Technologies (ONT) allowed for the detection of changes in current as DNA or RNA molecules pass through a nanopore, resulting in a new sequencing technology. This technology also produces long reads, with an average length of tens of kilobases, and can be used for real-time sequencing applications. In recent years, several other TGS technologies have been developed, including synthetic long-read sequencing. SLR developed by Illumina combines short-read sequencing and long-range PCR to produce synthetic long reads.

TGS technologies have revolutionized genomics research by allowing the detection of complex genomic changes that were previously challenging or impossible to identify. TGS technologies have also enabled the sequencing of long-read transcriptomes, metagenomes, and epigenomes, broadening the scope of genomic research. TGS technology has a few clinical applications, including the diagnosis and management (e.g., 6–9).

## 3. The application of TGS in thalassemia

As a novel sequencing technology, the use of TGS in *thalassemia* is a recent development, with most reports published after 2020. A search of the authoritative literature database PubMed using keywords such as “*thalassemia* third-generation sequencing”, “*thalassemia* long read sequencing”, “*thalassemia*



single-molecule real-time”, and “thalassemia nanopore” yielded a total of 30 relevant reports (<https://pubmed.ncbi.nlm.nih.gov/>; as of March 11, 2023). These studies demonstrate that TGS is gaining popularity as a technology for thalassemia screening and diagnosis, with 27 out of the 30 reports originating from China (refer to **Table 1**).

### 3.1. TGS in the molecular screening of thalassemia carriers

To reduce the incidence of thalassemia major infants, a mandatory mass screening strategy and appropriate genetic counseling are necessary for at-risk populations (39). Thalassemia screening using TGS has recently emerged (see **Table 1**). A more comprehensive and universal TGS screening framework, called Comprehensive Analysis of Thalassemia Alleles (CATSA), was established for thalassemia in 2021 (6). The CATSA method involves DNA extraction, PCR, library construction, sequencing, and bioinformatics analysis (**Figure 1**). A crucial step in the CATSA method is the utilization of a multiplex long-range PCR approach with primers optimized for specificity. This allows the generation of targeted amplicons that encompass all

known thalassemia regions of insertions, deletions, single nucleotide variants (SNVs), structural variations (SVs) and copy number variations (CNVs). In the CATSA study, a total of 1759 primary screening samples with positive routine blood or hemoglobin tests were subjected to CATSA screening. The results were subsequently validated using double-blind testing and other “gold standard” methods, which demonstrated that the CATSA test achieved a 100% accuracy rate.

Several other studies have been conducted to validate the CATSA method for population screening, with most reporting improved detection rates compared to standard methods (refer to **Table 1**). For instance, Peng et al. (10) performed CATSA and conventional methods on 100 patients with suspected thalassemia and observed that CATSA identified 10 more rare clinical variants than the latter, including the  $-\alpha^{3.7III}$  subtype, which was reported for the first time in China. The findings suggest that rare thalassemia variants are more prevalent than previously thought and are often misdiagnosed by conventional diagnostic methods. Liu et al. (29) conducted CATSA tests on 49 subjects who were negative for specially designed GAP-PCR and Sanger sequencing tests and identified mutations in the thalassemia gene in 8 subjects, indicating a higher positive detection rate of CATSA.

TABLE 1 Reposts using TGS in thalassemia screening and diagnosis from 2021 to 2023.

Hospital location	Application field	Number of subject test	Detection method	TGS platform	TGS efficacy finding	Reference and year
Hunan, China	PS	1,759	TGS vs. TM	PacBio	100% detection rate with higher genotype accuracy for individual carriers	Liang et al. 2021 (6)
–	PS	100	TGS vs. TM	PacBio	Detection of 10 more rare ground-poor mutations	Peng et al. 2022 (10)
Guangdong, China	PS	9	TGS vs. TM	PacBio	Detection more mutations and one-step determination of cis or trans conformation	Jiang et al. 2022 (11)
Fujian, China	PS	70	TGS vs. TM	PacBio	High detection rate of 7.14% in rare thalassemia and the first record in China with three variants found	Zhuang et al. 2022 (12)
Fujian, China	PS	33	TGS vs. TM	PacBio	The same detection rate of HK $\alpha\alpha$ carriers, and detection of two more $\beta$ -variants	Zhang et al. 2023 (13)
Guangxi, China	PS	434	TGS vs. TM	PacBio	Higher detection rate (83.18% vs. 73.27%) and more detection of rare genotypes	Luo et al. 2022 (14)
Guizhou, China	PS and prenatal screening	176 and 53 couples	TGS vs. TM	PacBio	Higher detection rate, and finding that five of couples had offspring at high risk of severe thalassemia	Wu et al. 2022 (15)
Hainan, China	PS	1,122	TGS vs. NGS	PacBio	Improved detection rates for common, rare and unknown genotypes	Huang et al. 2023 (16)
Guangdong, China	PS	23	TGS vs. NGS	PacBio	The same detection rate and correction of NGS misdiagnosis on 2 heterozygotes	Zhou et al. 2022 (17)
Guangdong, China	PGD	32	TGS vs. NGS	PacBio	Consistent with the NGS results on haplotype linkage analysis	Wu et al. 2022 (18)
Yunnan, China	PGD	5	TGS vs. NGS vs. TM	ONT	Consistent detection rates, but sequencing equipment is more affordable than NGS	Liu et al. 2021 (19)
Guangdong, China	NIPD	26	TGS vs. NGS	ONT	Haplotype typing was achieved in all families based on a 20 kb amplicon library	Jiang et al. 2021 (20)
Hunan, China	DRTV	278	TGS vs. TM	PacBio	Improved the detection rate by 19%, detected more rare genotypes, and corrected the diagnosis of 1 heterozygote by traditional methods	Liang et al. 2023 (21)
Guangxi, China	DRTV	32	TGS vs. TM	PacBio	Genotype detection rate of 100% and one-step differentiation of different genotypes	Li et al. 2022a (22)
Hubei, China	DRTV	2	TGS vs. NGS	PacBio	Diagnosis of a novel 10.3 kb deletion causing $\alpha^0$ -thalassemia	Xu et al. 2023 (23)
Guangxi, China	DRTV	206	TGS vs. TM	PacBio	Same detection rate, and correction of 3 cases of misidentification based on traditional methods	Li et al. 2022c (24)
Guangxi, China	DRTV	4	TGS vs. TM	PacBio	Same detection rate and ability to directly identify the structure of genotypes	Long et al. 2022 (25)
Guangxi, China	DRTV	4	TGS vs. TM	PacBio	Diagnosis of four more rare mutations of thalassemia	Luo et al. 2022 (26)
Guangxi, China	DRTV	1	TGS vs. TM	PacBio	Diagnosis of a novel 4.9 kb deletion at beta-globin gene	Chen et al. 2022 (27)
Guangxi, China	DRTV	1	TGS vs. TM	PacBio	Diagnosis of a novel 107 kb deletion in the alpha-globin gene	Li et al. 2022c (28)
Hunan, China	DRTV	49	TGS vs. TM	PacBio	Diagnosis of eight more rare mutations of thalassemia	Liu et al. 2023 (29)
Guangdong, China	DRTV	1	TGS vs. TM	PacBio	Diagnosis of a novel 7.2 kb deletion at $\beta$ -globin gene	Zhong et al. 2022a (30)
Guangdong, China	DRTV	1	TGS vs. TM	PacBio	Diagnosis of a novel 15.8 kb deletion at $\alpha$ -globin gene	Zhong et al. 2022b (31)
Guangdong, China	DRTV	1	TGS vs. TM	PacBio	Successful diagnosis of $\alpha^+$ thalassemia subtype $-\alpha^{3.7III}$	Bao et al. 2022a (32)
Guangdong, China	DRTV	1	TGS vs. TM	PacBio	Diagnosis of a novel 5 kb deletion at $\beta$ -globin gene	Bao et al. 2022b (33)
Minnesota, USA	DRTV	4	TGS vs. TM	PacBio	Improving the identification of structural variants in complex $\beta$ -globin gene cluster	Rangan et al. 2021 (34)
Guangxi, China	DRTV	1	TGS vs. TM	PacBio	Diagnosis of novel rearrangements of the $\alpha$ -globin gene cluster	Ning et al. 2022 (35)
Guangdong, China	DRTV	2	TGS vs. TM	PacBio	Conformation of Hb Q-Thailand heterozygosity not typically associated with the ( $-\alpha^{4.2}$ ) $\alpha^+$ -thalassemia deletion allele	Qin et al. 2022 (36)
/	Editorial	/	TGS vs. TM	/	Improved sensitivity and specificity, with important advantages in the prenatal diagnosis	Toledo and Lafferty 2023 (37)
/	Review	/	TGS vs. NGS	/	Advantages in homologous gene variant discovery and copy number variation calling	Hasson et al. 2023 (38)

PS, Population screening; PGD, preimplantation genetic diagnosis; NIPD, non-invasive prenatal diagnosis; DRTV, diagnosis of rare thalassemia variants; TGS, third-generation sequencing; NGS, next-generation sequencing; TM, traditional method; PacBio, Pacific Biosciences; ONT, Oxford Nanopore Technologies; –, unavailable.

Screening of prospective parents for thalassemia carriers has received equally attention in many countries, especially in Israel, Sardinia and Cyprus, where mandatory premarital screening has been implemented in recent decades and through which the burden of thalassemia has been successfully reduced (40). Recently, TGS has also been applied to prenatal screening for thalassemia. In Wu et al. (15), 53 couples were examined for homozygous thalassemia mutations using TGS. The result showed that seven of them were identified as positive carriers, with five of them at risk of severe thalassemia during pregnancies, highlighting the potential of TGS for prenatal screening in regions with a high incidence of thalassemia.

### 3.2. TGS in the molecular diagnosis of thalassemia

Accurately determining the severity of a disease is a critical step in selecting the appropriate treatment, particularly in cases where patients exhibit atypical clinical manifestations or where conventional genotyping methods yield uncertain diagnoses. TGS technology, similar to NGS, can directly detect thalassemia genotypes and provide efficient interpretation of thalassemia variants. TGS has been successfully applied in various clinical scenarios, including preimplantation genetic diagnosis (PGD), non-invasive prenatal diagnosis (NIPD), and rare thalassemia diagnosis (as summarized in Table 1).

In the context of PGD, Liu et al. (19) evaluated the effectiveness of TGS using the Oxford Nanopore sequencing platform and observed consistent diagnostic outcomes with those obtained through conventional methods and NGS. Additionally, the cost of sequencing equipment was lower than that of NGS. Wu et al. (18) employed TGS to conduct haplotype linkage analysis on 32 embryos from three couples and succeeded in establishing haplotype linkage in 68.75% of them. These results align with those obtained via NGS, suggesting that TGS has promising potential for PGD applications. In a study examining NIPD applications, Jiang et al. (20) analyzed  $\beta$ -thalassemia in 13 families via TGS combined with relative haplotype dosage (RHDO) analysis using the Oxford Nanopore sequencing platform. The results indicated successful haplotyping in all 13 families when using a 20 kb amplicon library, demonstrating the feasibility of the TGS protocol. The advantages of TGS in detecting and interpreting thalassemia genotypes highlight its potential for improving the accuracy of diagnosis and facilitating personalized treatment strategies for patients with thalassemia.

### 4. Advantages of TGS in screening and diagnosis of thalassemia

Compared to traditional methods, TGS has shown significant advantages in thalassemia detection. TGS can cover over hundreds of  $\alpha$ - and  $\beta$ -deletion variants in a single test, which surpasses the coverage of all previous methods for thalassemia detection. Moreover, TGS improves the detection rate of positive thalassemia

genes, especially rare ones, by up to 19%. This can help avoid underdiagnoses caused by traditional methods. Additionally, TGS enhances the accuracy of diagnosis and can avoid misdiagnosis by traditional methods, especially for complex heterozygotes. TGS allows for the direct detection of new thalassemia genetic variants and precise identification of cis-trans variants in a “one-step” manner, without the need for family verification. This greatly facilitates the interpretation of thalassemia variants. Furthermore, TGS has a simple operation process and eliminates the need for multiple methods of cross-validation and inference, thus avoiding possible contamination from tedious steps.

Compared to NGS, recent studies with larger sample sizes indicate that the detection rate of TGS is higher, especially for deletion variants and heterozygotes (16, 17, 23). This is due to the significant advantage of TGS in read length, which is much longer than that of NGS, making it easier to detect deletion variants in thalassemia genes (23). Medical practitioners have started to recognize TGS as the preferred method over NGS for detecting long deletion variants in the thalassemia gene (Table 1) (23, 27, 28, 30, 31, 33). On the other hand, NGS faces difficulties in accurately sequencing regions with abnormal GC content and highly repetitive sequences, which can lead to missed mutations and false-negative results. In contrast, TGS can address this limitation due to its longer read length and less preference for GC content, leading to higher detection rates and diagnostic accuracy for thalassemia genes than NGS. Additionally, another advantage of TGS over NGS is its shorter turnaround time. While NGS strategies typically require additional confirmation through Gap-PCR and Sanger sequencing, which can prolong the overall testing process, TGS provides more rapid results (Figure 1).

### 5. Challenges and solutions of TGS in routine detection of thalassemia

The accuracy of TGS has been a longstanding topic of debate and improvement for many years, which has also raised doubts about its application in disease diagnosis. However, recent advancements have demonstrated that the accuracy of TGS sequencing can reach levels comparable to other sequencing technologies, such as Illumina sequencing. For instance, the earlier versions of PacBio sequencing platform produced high native error rates (ca. 13%) (41), but the current sequencing accuracy is high, especially high-fidelity (HiFi) reads generated by the Circular Consensus Sequencing (CCS) mode with an accuracy of up to 99.999% (42). Similarly, earlier ONT's equipment also had a high native sequencing error rate (5%–25%) (43), while ONT claims that the new version 10 flow cells can deliver up to 99.99% base-calling accuracy. Furthermore, using an amplicon sequencing approach combining unique molecular identifiers (UMIs) with ONT or PacBio CCS, the accuracy rate can be over 99.99% (44). Consequently, the accuracy of TGS will no longer pose a challenge for the proper detection of thalassemia.

Currently, the main platform for thalassemia detection using TGS is the PacBio platform (see Table 1). Despite its significant

technical advantages, the high equipment costs and long turnaround times of PacBio limit its integration into routine screening and diagnostic procedures. While the total cost of preparing TGS libraries and sequencing reagents for each sample has dropped to below \$20 (21), mainstream PacBio sequencing equipment is still expensive (e.g., the PacBio Sequel II device costs approximately \$350,000), limiting its adoption by many medical institutions. If TGS library preparation samples are sent to biotech companies for sequencing, the overall testing cycle takes about 8 days (including sample transportation time) in China, whereas the testing cycle for traditional methods is no more than 3 days (21). It is expected that PacBio will release sequencing equipment that covers medium and low throughput with greatly reduced costs, and this equipment can fully meet the low throughput testing requirements for thalassemia currently. At that time, if medical institutions can purchase this equipment for in-house sequencing, they can eliminate transportation time to the company and arrange sequencing work promptly, thereby shortening the testing cycle.

An alternative solution to address the issues of high equipment costs and long turnaround times is to use the ONT platform for thalassemia detection. The ONT devices available on their official website range from the cheapest MinION at \$1,000 to the most expensive GridION at \$49,955 (as of March 27, 2023), which are significantly more affordable than PacBio devices, and thus more accessible to most medical institutions. On the other hand, ONT's turnaround time is greatly reduced. For example, a study reported a total turnaround time of 6 h for detecting viruses such as Ebola, Chikungunya and Hepatitis C virus (45). Although there is still limited research on the use of ONT in thalassemia detection, with only two studies available (see Table 1). Both the studies have demonstrated promising clinical potential for ONT in NIPD and PGD (19, 20). It is believed that in the near future, it will be a good alternative to traditional methods and facilitate the transition to third-generation sequencing technology for screening of thalassemia.

## 6. Conclusion

Third-generation sequencing (TGS) is a novel technology that has not yet been widely implemented for the screening and diagnosis of thalassemia in clinical practice worldwide. However, recent validation trials and preliminary studies have demonstrated that TGS using the PacBio platform provides significant advantages over traditional methods and NGS in terms of positive detection rate and diagnostic accuracy for thalassemia (see Table 1). Nevertheless, the high equipment cost and lengthy turnaround time currently limit the integration of PacBio-based TGS into routine thalassemia testing (summarized

in Figure 1). Although ONT is a promising TGS platform with lower equipment cost and shorter turnaround time, its efficacy for thalassemia genetic testing has not been fully explored. Large-scale comparative trials are still required to validate its utility and accuracy. With the ongoing improvement in the reduced cost and faster turnaround time for PacBio, as well as the increased efficacy of ONT, it is anticipated that TGS will become a viable and cost-effective alternative for clinical laboratories to perform in-house thalassemia testing in the near future.

## Author contributions

LZ, CG, WW, JL, BG contributed to writing, editing, and finalizing the manuscript. BG organized and reviewed the final version. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported in part by the National Natural Science Foundation of China (82001531 and 81860272 to BG), Guangxi Major Research Programme (AB22035013 to BG), Guangxi Natural Science Foundation (2023GXNSFBA026124 to CG, 2018GXNSFAA281067 to BG), Initial Scientific Research Fund for Advanced Talents from The Second Affiliated Hospital of Guangxi Medical University (2019112 to BG), Special Scientific Research Fund of Guangxi Ten-Hundred-Thousand Talents Project (2021186 to BG), and the Science Foundation for Young Scholars of Guangxi Medical University (GXMUYSF202115 to CG), Beihai Science and Technology Plan Project (2019D06 to LZ).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Cooley TB. Anemia in children: with splenomegaly and peculiar changes in the bones report of cases. *Am J Dis Child.* (1927) 34:347. doi: 10.1001/archpedi.1927.04130210022002
2. Whipple GH, Bradford WL. Mediterranean disease-thalassemia (Erythroblastic anemia of cooley). *J Pediatr-US.* (1936) 9:279–311. doi: 10.1016/S0022-3476(36)80021-3

3. Williams TN, Weatherall DJ. World distribution, population genetics, and health burden of the hemoglobinopathies. *Cold Spring Harb Perspect Med.* (2012) 2: a011692–a011692. doi: 10.1101/cshperspect.a011692
4. Beijing AngelMom Charity Foundation, China Philanthropy Research Institute. *Blue book on thalassemia in China—survey report on the prevention and treatment of thalassemia in China (2020)*. 1st ed. Beijing, China: China Society Press (2021). 1–54. (in Chinese).
5. Suhaimi SA, Zulkupli IN, Ghani H, Abdul-Hamid MRW. Applications of next generation sequencing in the screening and diagnosis of thalassemia: a mini-review. *Front Pediatr.* (2022) 10:1015769. doi: 10.3389/fped.2022.1015769
6. 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. doi: 10.1016/j.jmoldx.2021.06.008
7. Hoenen T, Groseth A, Rosenke K, Fischer RJ, Hoenen A, Judson SD, et al. Nanopore sequencing as a rapidly deployable ebola outbreak tool. *Emerg Infect Dis.* (2016) 22:331–4. doi: 10.3201/eid2202.151796
8. Mimori T, Yasuda J, Kuroki Y, Shibata TF, Katsuoka F, Saito S, et al. Construction of full-length Japanese reference panel of class I HLA genes with single-molecule, real-time sequencing. *Pharmacogenomics J.* (2019) 19:136–46. doi: 10.1038/s41397-017-0010-4
9. Hu T, Li J, Long M, Wu J, Zhang Z, Xie F, et al. Detection of structural variations and fusion genes in breast cancer samples using third-generation sequencing. *Front Cell Dev Biol.* (2022) 10:854640. doi: 10.3389/fcell.2022.854640
10. Peng C, Zhang H, Ren J, Chen H, Du Z, Zhao T, et al. Analysis of rare thalassemia genetic variants based on third-generation sequencing. *Sci Rep.* (2022) 12:9907. doi: 10.1038/s41598-022-14038-8
11. Jiang F, Mao A-P, Liu Y-Y, Liu F-Z, Li Y-L, Li J, et al. Detection of rare thalassemia mutations using long-read single-molecule real-time sequencing. *Gene.* (2022) 825:146438. doi: 10.1016/j.gene.2022.146438
12. Zhuang J, Chen C, Fu W, Wang Y, Zhuang Q, Lu Y, et al. Third-generation sequencing as a new comprehensive technology for identifying rare  $\alpha$ - and  $\beta$ -globin gene variants in thalassemia alleles in the Chinese population. *Arch Pathol Lab Med.* (2023) 147:208–214. doi: 10.5858/arpa.2021-0510-OA
13. Zhang M, Lin Z, Chen M, Pan Y, Zhang Y, Chen L, et al. Application of the single-molecule real-time technology (SMRT) for identification of HK $\alpha\alpha$  thalassemia allele. *Lab Med.* (2023) 54:65–71. doi: 10.1093/labmed/lmac065
14. 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. doi: 10.1038/s10038-021-00983-1
15. Wu J, Xie D, Wang L, Kuang Y, Luo S, Ren L, et al. Application of third-generation sequencing for genetic testing of thalassemia in guizhou province, southwest China. *Hematology.* (2022) 27:1305–11. doi: 10.1080/16078454.2022.2156720
16. Huang R, Liu Y, Xu J, Lin D, Mao A, Yang L, et al. Back-to-back comparison of third-generation sequencing and next-generation sequencing in carrier screening of thalassemia. *Arch Pathol Lab Med.* (2023) doi: 10.5858/arpa.2022-0168-OA [Epub ahead of print]
17. Zhou Q-M, Jiang F, Xu J, Lin D, Zhou J-Y, Qu Y-X, et al. High accuracy of single-molecule real-time sequencing in detecting a rare  $\alpha$ -globin fusion gene in carrier screening population. *Ann Hum Genet.* (2022) 1:9–17. doi: 10.1111/ahg.12486
18. Wu H, Chen D, Zhao Q, Shen X, Liao Y, Li P, et al. Long-read sequencing on the SMRT platform enables efficient haplotype linkage analysis in preimplantation genetic testing for  $\beta$ -thalassemia. *J Assist Reprod Genet.* (2022) 39:739–46. doi: 10.1007/s10815-022-02415-1
19. Liu S, Wang H, Leigh D, Cram DS, Wang L, Yao Y. Third-generation sequencing: any future opportunities for PGT? *J Assist Reprod Genet.* (2021) 38:357–64. doi: 10.1007/s10815-020-02009-9
20. Jiang F, Liu W, Zhang L, Guo Y, Chen M, Zeng X, et al. Noninvasive prenatal testing for  $\beta$ -thalassemia by targeted nanopore sequencing combined with relative haplotype dosage (RHDO): a feasibility study. *Sci Rep.* (2021) 11:5714. doi: 10.1038/s41598-021-85128-2
21. Liang Q, He J, Li Q, Zhou Y, Liu Y, Li Y, et al. Evaluating the clinical utility of a long-read sequencing-based approach in prenatal diagnosis of thalassemia. *Clin Chem.* (2023) 69:239–50. doi: 10.1093/clinchem/hvac200
22. Li J, Ye G, Zeng D, Tian B, Wang W, Feng Q, et al. Accurate genotype diagnosis of Hong Kong  $\alpha\alpha$  thalassemia based on third-generation sequencing. *Ann Transl Med.* (2022) 10:1113. doi: 10.21037/atm-22-4309
23. Xu R, Li H, Yi S, Du J, Jin J, Qin Y, et al. Identification of a novel 10.3 kb deletion causing  $\alpha^0$ -thalassemia by third-generation sequencing: pedigree analysis and genetic diagnosis. *Clin Biochem.* (2023) 113:64–9. doi: 10.1016/j.clinbiochem.2022.12.018
24. Li Y, Liang L, Qin T, Tian M. Detection of hemoglobin H disease by long molecule sequencing. *J Clin Lab Anal.* (2022) 36:e24687. doi: 10.1002/jcla.24687
25. Long J, Sun L, Gong F, Zhang C, Mao A, Lu Y, et al. Third-generation sequencing: a novel tool detects complex variants in the  $\alpha$ -thalassemia gene. *Gene.* (2022) 822:146332. doi: 10.1016/j.gene.2022.146332
26. Luo S, Chen X, Zeng D, Tang N, Yuan D, Liu B, et al. Detection of four rare thalassemia variants using single-molecule realtime sequencing. *Front Genet.* (2022) 13:974999. doi: 10.3389/fgene.2022.974999
27. Chen X, Luo M, Pan L, Huang Y, Yan Z, Shen K, et al. A novel 4.9 kb deletion at beta-globin gene is identified by the third-generation sequencing: case report from baoan, China. *Clin Chim Acta.* (2022) 529:10–6. doi: 10.1016/j.cca.2022.01.024
28. Li Y, Liang L, Guo W, Wu X, Qin T, Tian M. Identification of a novel 107 kb deletion in the alpha-globin gene cluster using third-generation sequencing. *Clin Biochem.* (2023) 113:36–9. doi: 10.1016/j.clinbiochem.2022.12.010
29. Liu Q, Chen Q, Zhang Z, Peng S, Liu J, Pang J, et al. Identification of rare thalassemia variants using third-generation sequencing. *Front Genet.* (2023) 13:1076035. doi: 10.3389/fgene.2022.1076035
30. Zhong G, Zhong Z, Guan Z, Chen D, Wu Z, Yang K, et al. Case report: the third-generation sequencing confirmed a novel 7.2 kb deletion at  $\beta$ -globin gene in a patient with rare  $\beta$ -thalassemia. *Front Genet.* (2022) 13:984996. doi: 10.3389/fgene.2022.984996
31. Zhong Z, Zhong G, Guan Z, Chen D, Wu Z, Yang K, et al. A novel 15.8 kb deletion  $\alpha$ -thalassemia confirmed by long-read single-molecule real-time sequencing: hematological phenotypes and molecular characterization. *Clin Biochem.* (2022) 108:46–9. doi: 10.1016/j.clinbiochem.2022.06.015
32. Bao X, Wang J, Qin D, Zhang R, Yao C, Liang J, et al. The  $-\alpha^{3.7III}$  subtype of  $\alpha^+$ -thalassemia was identified in China. *Hematology.* (2022) 27:826–30. doi: 10.1080/16078454.2022.2101913
33. Bao X-Q, Wang J-C, Qin D-Q, Yao C-Z, Liang J, Du L. A novel 5 kb deletion in the  $\beta$ -globin gene cluster identified in a Chinese patient. *Hemoglobin.* (2022) 46:245–8. doi: 10.1080/03630269.2022.2118604
34. Rangan A, Hein MS, Jenkinson WG, Koganti T, Aleff RA, Hilker CA, et al. Improved characterization of Complex  $\beta$ -globin gene cluster structural variants using long-read sequencing. *J Mol Diagn.* (2021) 23:1732–40. doi: 10.1016/j.jmoldx.2021.08.013
35. Ning S, Luo Y, Liang Y, Xie Y, Lu Y, Meng B, et al. A novel rearrangement of the  $\alpha$ -globin gene cluster containing both the  $-\alpha^{3.7}$  and  $\alpha\alpha\alpha^{anti4.2}$  crossover junctions in a Chinese family. *Clin Chim Acta.* (2022) 535:7–12. doi: 10.1016/j.cca.2022.07.020
36. Qin D, Wang J, Yao C, Bao X, Liang J, Du L. Hb Q-Thailand heterozygosity unlinked with the  $(-\alpha^{4.2})$   $\alpha^+$ -thalassemia deletion allele identified by long-read SMRT sequencing: hematological and molecular analyses. *Hematology.* (2023) 28:2184118. doi: 10.1080/16078454.2023.2184118
37. Toledo DM, Lafferty KA. Clinical perspective on use of long-read sequencing in prenatal diagnosis of thalassemia. *Clin Chem.* (2023) 69:211–2. doi: 10.1093/clinchem/hvac223
38. Hassan S, Bahar R, Johan MF, Mohamed Hashim EK, Abdullah WZ, Esa E, et al. Next-generation sequencing (NGS) and third-generation sequencing (TGS) for the diagnosis of thalassemia. *Diagnostics.* (2023) 13:373. doi: 10.3390/diagnostics13030373
39. Goonasekera HW, Paththinige CS, Dissanayake VHW. Population screening for hemoglobinopathies. *Annu Rev Genom Hum Genet.* (2018) 19:355–80. doi: 10.1146/annurev-genom-091416-035451
40. Shafie AA, Wong JHY, Ibrahim HM, Mohammed NS, Chhabra IK. Economic burden in the management of transfusion-dependent thalassaemia patients in Malaysia from a societal perspective. *Orphanet J Rare Dis.* (2021) 16:157. doi: 10.1186/s13023-021-01791-8
41. Wick RR, Judd LM, Holt KE. Deepbinner: demultiplexing barcoded Oxford nanopore reads with deep convolutional neural networks. *PLoS Comput Biol.* (2018) 14:e1006583. doi: 10.1371/journal.pcbi.1006583
42. Wenger AM, Peluso P, Rowell WJ, Chang P-C, Hall RJ, Concepcion GT, et al. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nat Biotechnol.* (2019) 37:1155–62. doi: 10.1038/s41587-019-0217-9
43. Ardui S, Ameer A, Vermeesch JR, Hestand MS. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res.* (2018) 46:2159–68. doi: 10.1093/nar/gky066
44. Karst SM, Ziels RM, Kirkegaard RH, Sørensen EA, McDonald D, Zhu Q, et al. High-accuracy long-read amplicon sequences using unique molecular identifiers with nanopore or PacBio sequencing. *Nat Methods.* (2021) 18:165–9. doi: 10.1038/s41592-020-01041-y
45. Greninger AL, Naccache SN, Federman S, Yu G, Mbala P, Bres V, et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med.* (2015) 7:99. doi: 10.1186/s13073-015-0220-9