



OPEN ACCESS

EDITED BY

Trevor Lucas,
Medical University of Vienna, Austria

REVIEWED BY

Guoxing Zhong,
Huizhou First Maternal and Child Healthcare
Hospital, China
Xiaofeng Wei,
Southern Medical University, China

*CORRESPONDENCE

Fengxiang Wei,
✉ haowei727499@163.com

RECEIVED 28 October 2024

ACCEPTED 20 February 2025

PUBLISHED 10 March 2025

CITATION

Xu J, Hu L, Wen L, Cao X, Xu H, Luo Q, Long Y,
Ji T, Sun L and Wei F (2025) Detection of a novel
large fragment deletion in the alpha-globin
gene cluster using the CNVplex technology.
Front. Genet. 16:1518392.
doi: 10.3389/fgene.2025.1518392

COPYRIGHT

© 2025 Xu, Hu, Wen, Cao, Xu, Luo, Long, Ji, Sun
and Wei. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). 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.

Detection of a novel large fragment deletion in the alpha-globin gene cluster using the CNVplex technology

Jianfei Xu^{1,2}, Liang Hu², Lijuan Wen², Xianzhen Cao²,
Hongyan Xu¹, Qi Luo¹, Yuhong Long¹, Tingyu Ji¹, Lifang Sun³ and
Fengxiang Wei^{1,2*}

¹Jiamusi University, Jiamusi, China, ²Longgang District Maternity & Child Healthcare Hospital of Shenzhen City (Affiliated Shenzhen Women and Children's Hospital (Longgang) of Shantou University Medical College), Shenzhen, China, ³Laboratory of Shenzhen Children's Hospital, Shenzhen, China

Objective: To describe the characterization of a novel deletion causing α -thalassemia.

Methods: The proband was a 4-year-old boy who presented with abnormal hematological parameters identified during routine blood investigation conducted for a cold. Three common α -globin gene deletions, three mutations, and 17 mutations in the β -globin gene were detected using PCR-flow fluorescence hybridization. Next-generation sequencing (NGS) and CNVplex technologies were employed to identify potential rare pathogenic mutation types. The CNVplex technology leverages variations in the lengths of linkage sequences of differential sequences at the same locus to produce linkage products of varying lengths, thereby enabling the detection of multiple loci within the same system. The newly identified deletions were further validated using customized third-generation sequencing (TGS) and Sanger sequencing.

Conclusion: In this study, hematological analysis indicated a potential diagnosis of thalassemia in the proband, characterized by typical microcytic hypochromic features. A novel 134-kb deletion in the α -globin gene cluster was identified in this proband using the CNVplex technology. This deletion encompasses the genes *HBZ*, *HBM*, *HBA2*, *HBA1*, and *HBQ1*. Furthermore, we confirmed the gene deletion through customized TGS testing and Sanger sequencing, allowing us to determine the size of the deletion. The results suggest that this represents a new deletion of 146 kb that has not been previously reported, and we hypothesize that this deletion is likely the primary cause of the α -thalassemia trait observed in the proband.

KEYWORDS

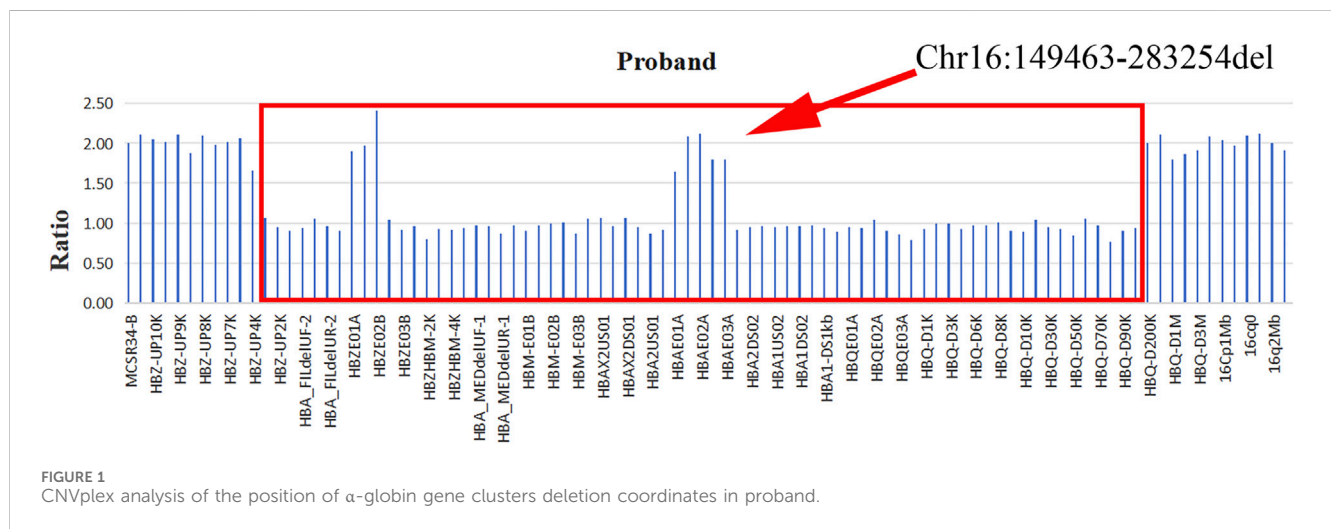
thalassemia, α -globin gene, CNVplex, single-molecule real-time sequencing, deletion

Introduction

Among the types of thalassemia, α -thalassemia is primarily characterized by deletions and duplications of large genomic segments (Colosimo et al., 2011). The phenotype of α -thalassemia is directly related to the number of α -globin genes affected and can be classified into several categories: α^+ -thalassemia ($-\alpha/\alpha\alpha$ or $\alpha^T\alpha/\alpha\alpha$), α^0 -thalassemia ($--/\alpha\alpha$ or $-\alpha/-\alpha$), HbH disease ($--/-\alpha$ or $--/\alpha^T\alpha$), and Hb Bart's hydrops fetalis syndrome ($---/---$). Hb Bart's fetal

TABLE 1 Hematological data, serum ferritin, and iron levels of the family members.

Sample	Sex-age	RBC ($\times 10^{12}/L$)	HGB (g/L)	MCV (fL)	MCH (pg)	HBA (%)	HBA ₂ (%)	HbF (%)	Serum ferritin ($\mu\text{g/L}$)	Serum iron ($\mu\text{mol/L}$)
Proband(II ₁)	M-4 years	6.17 \uparrow	124	60.8 \downarrow	20.1 \downarrow	97.4	2.6	0	40.67	16.6
Father(I ₁)	M-37 years	4.94	154	89.3	31.2	97.5	2.5	0	ND	ND
Mother (I ₂)	F-38 years	4.90	151 \uparrow	90.6	30.8	97.20	2.6	0.2	ND	ND
Sister(II ₂)	F-11 years	4.35	123	83.4	28.3	97.4	2.6	0	ND	ND



edema syndrome occurs in fetuses presenting with severe anemia, often resulting in fetal demise late in pregnancy or at birth (Piel and Weatherall, 2014). The routine test kits currently employed in the majority of hospitals are designed to target high-frequency variants within the local population, thereby failing to detect low-frequency variants. This limitation can pose challenges in clinical counseling and prenatal diagnosis (Li, 2017). With continuous advancement in genetic testing technology, the detection of novel mutations is expected to become increasingly easier and more efficient (Jiang et al., 2023). The CNVplex high-throughput linkage-dependent probe amplification (HLPA) technology is used to accurately detect deletion or duplication mutations in each exon of a target gene (or any deletion or duplication mutation in any non-duplicated target segment greater than 60 bp). The CNVplex technology is based on the use of a ligase-specific ligation reaction to make the target fragment through hybridization and ligation so as to introduce different lengths of non-specific sequences at the end of the ligation probe and to obtain different lengths of ligation products corresponding to different sites, and then through the use of universal primers with different markers of fluorescence through the PCR will be ligated to amplify the product, and then the PCR product on the sequencer to run capillary electrophoresis to separate out the different lengths of fragments. Subsequently, the PCR products were amplified using various labeled fluorescent primers. These products were then subjected to capillary electrophoresis on a sequencer to separate fragments of different lengths. The resulting capillary electrophoresis profiles were

analyzed to identify the peaks corresponding to each site, allowing for the calculation of the copy number of the region of interest in determining the ratio of the target region peaks to the reference peaks (Liang et al., 2017; Wei et al., 2024). This test has been employed in the molecular diagnosis of other genetic disorders, such as congenital deafness and congenital heart disease (Huang et al., 2018; Hu et al., 2022). In this study, we identified a novel deletion of the α -globin gene using the CNVplex technology, determining the size of the deletion to be 146 kb through customized TGS for further validation and precise localization.

Case presentation

The proband was a 4-year-old boy from Shenzhen, Guangdong Province, who presented to our hospital with symptoms of a cold. His hematological parameters were assessed using an automated hematology analyzer (XT-2000i, Japan), which indicated that he was suffering from microcytic hypochromic anemia. Specifically, he exhibited a mean corpuscular volume (MCV) of 60.8 fL (reference range: 76.0~88.0 fL) and a mean hemoglobin concentration (MCH) of 20.1 pg (reference range: 24.0~30.0 pg). Notably, the patient's erythrocyte count was $6.17 \times 10^{12}/L$ (reference range: $4.0\sim 5.50 \times 10^{12}/L$). It is speculated that these findings may represent a compensatory response. Detailed hematological indices are presented in Table 1, and iron deficiency has been excluded. Therefore, we consider the proband to be at risk for thalassemia. Given the

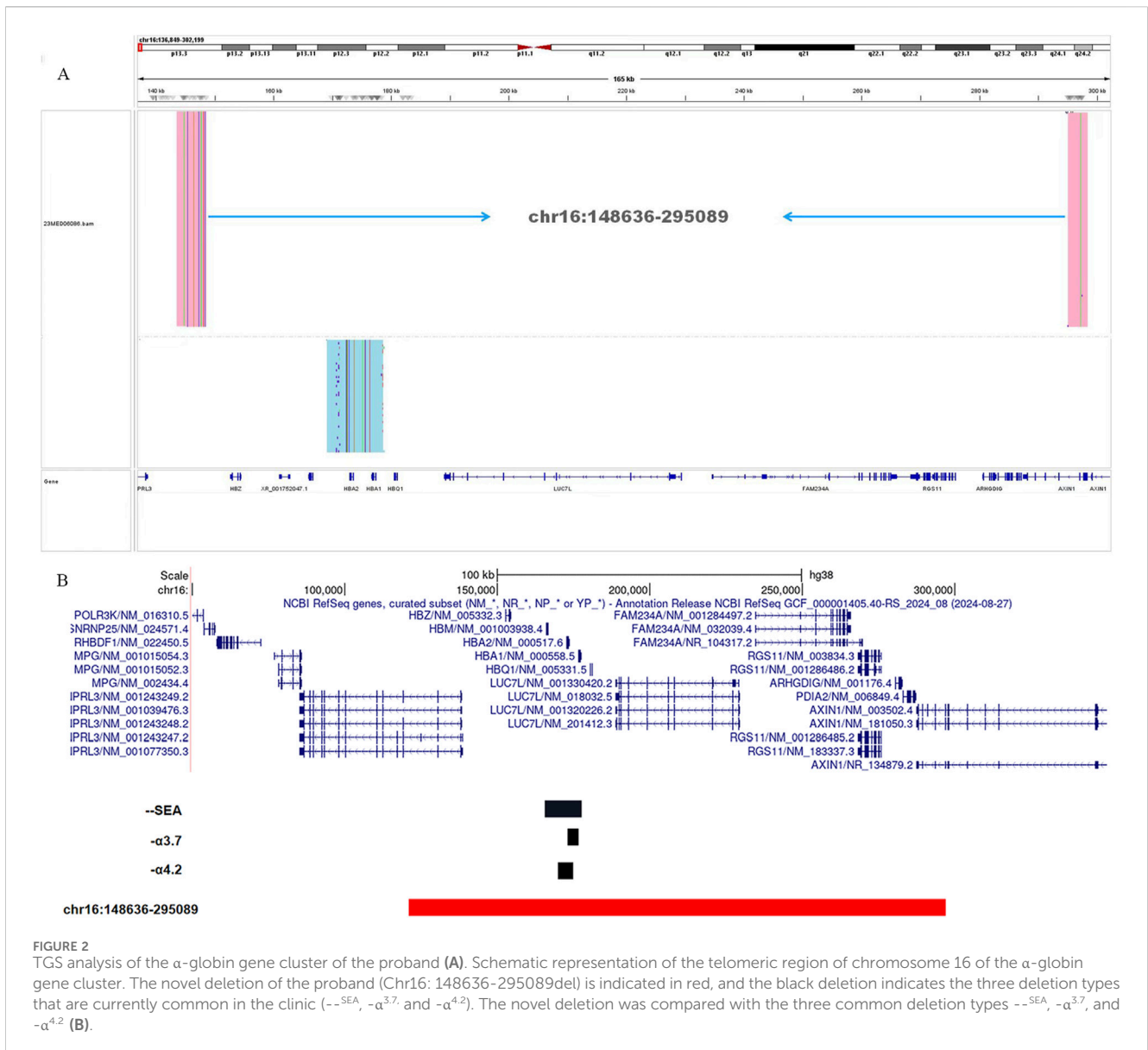


FIGURE 2
TGS analysis of the α -globin gene cluster of the proband (A). Schematic representation of the telomeric region of chromosome 16 of the α -globin gene cluster. The novel deletion of the proband (Chr16: 148636-295089del) is indicated in red, and the black deletion indicates the three deletion types that are currently common in the clinic (--SEA, - $\alpha^{3.7}$ and - $\alpha^{4.2}$). The novel deletion was compared with the three common deletion types --SEA, - $\alpha^{3.7}$ and - $\alpha^{4.2}$ (B).

patient's positive thalassemia phenotype, thalassemia screening was advised. For this screening, genomic DNA was extracted from peripheral blood using the DNeasy extraction kit (QIAGEN Inc., Valencia, CA, United States) and analyzed via PCR-flow fluorescence hybridization (Daan Gene, Guangzhou, China) including three α -thalassemia deletions--SEA (NC_000016.9:g.215,400_234700del), - $\alpha^{3.7}$ (NC_000016.9:g.223300_227103del), and - $\alpha^{4.2}$ (NC_000016.9:g.219817_223755_224,074del); three mutations (CD122(CAC>CAG) HBA2:c.427T>C, CD142 (TAA>CAA) HBA2:377T>C, CD125(CTG>CCG) HBA2:427T>C); and 17 mutations in the β -globin gene (CD41-42 (-CTTT) HBB:c.126_129del, IVS -2-654(C>T) HBB:c.316-197C>T, CD17 (A>T) HBB:c.52A>T, -28 (A>G) HBB:c.-78A>G, CD26 (G>A) HBB:c.79G>A, CD71-72 (+A) HBB:c.217dupA, CD43 (G>T) HBB:c.130G>T, -29 (A>C) HBB:c.-79A>C, Int (ATG>AGG) HBB:c.2T>G, CD14-15 (+G) HBB:c.45dupG, CD27-28 (+C) HBB:c.85dupC, -32(C>A) HBB:c.-82C>A, -30 (T>C) HBB:c.-80T>C, IVS-1-1 (G>T) HBB:c.92 + 1G>T, IVS-1-5 (G>A) HBA1:c.95 + 5G>A, CD31 (-C) HBB:c.94delC, and Cap+40-43 (-AAAC) HBB:c.-

10_-7delAACA). The proband did not exhibit any abnormalities. Given his hematological parameters, there was a suspicion that he might be carrying a rare thalassemia gene. Consequently, after obtaining consent from the proband and his family, blood samples were collected for the targeted sequencing of the HBA1, HBA2, and HBB genes using NGS (MGI Tech Co., Shenzhen, China).

No abnormalities were detected as a result of this testing. Based on the patient's hematological parameters, it was presumed that she had a thalassemia trait (--/aa). Following informed consent from the proband and his family, blood samples were collected from the proband's parents and sister and subjected to CNVplex testing (Suzhou Tianhao Biomedical Laboratory, Suzhou, China). The CNVplex technology was employed to detect deletion and duplication mutations in these samples, and point mutations were analyzed in conjunction with Sanger sequencing. The results revealed an unreported deletion of approximately 134 kb (Chr16:149463-283254) on the HBA gene of this proband (Figure 1). Limitations on the number of probes restricted the precise localization of the deletion, and testing of his family

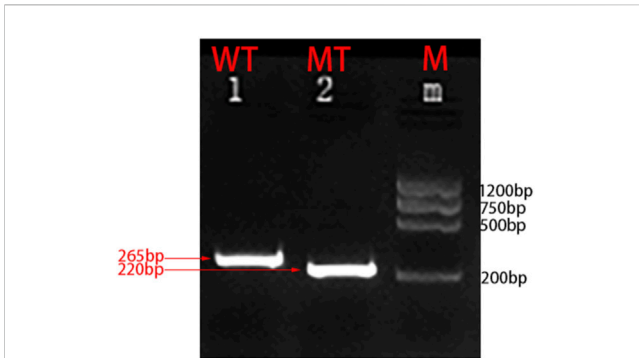


FIGURE 3
Molecular characterization of the large fragment deletion in this proband. Two primers specific for the α -globin gene cluster were designed. Amplification was successfully performed by GAP-PCR: F: GGGTGCTGTCGCTTTCTA and R:GTGAGCCGAGATCCGAGGTC (mutant-type); F:GGTGCTGTCGCTTTCTA and R:AAGTAGA TCCTGTTCCAGGGTAG (wild-type).

members did not yield any abnormal results. The proband was confirmed to have a new large segmental deletion.

To further validate the accuracy of the technique and clarify the exact location of the deletion, we employed long-read single-molecule real-time sequencing technology (SMRT) to analyze the proband. Preliminary results indicated that the *HBA* gene in this sample exhibited a haplotype of 1, which does not rule out the possibility of a deletion occurring outside the detection range. To accurately identify the precise breakpoints of the deletion in this sample, we customized and validated samples using long-read single-molecule real-time sequencing technology. Initially, primers were designed for the breakpoints based on the range of deletions detected by CNVplex. Amplification products were

obtained through GAP-PCR targeting the specified region, and these products were subsequently subjected to TGS sequencing. The customized TGS validation revealed a deletion of approximately 146 kb in size within the α -globin gene of this proband (Figure 2A). Using a relatively accurate range estimation, primers were designed on both sides of the breakpoint, resulting in a unique GAP-PCR product (Figure 3). Sanger sequencing confirmed that the breakpoint was located at chr16:148636-295089 (GRCh38/hg38), indicating a deletion of 146,453 bp (Figure 4). Through database searches, including NCBI and HbVar, we concluded that the deletion represents an unreported large-band deletion of approximately 146 kb located on chromosome 16 (Chr16:148636-295089, GRCh38/hg38) (Figure 4). Subsequently, we obtained the registration number for the novel mutation in the IthaNet database (ID: 4110), which has been designated as a 146-kb deletion.

In comparison to the common types of α -globin gene deletions observed in clinical settings, the deletion fragment in this case is larger and includes genes such as *HBZ*, *HBM*, *HBA1*, *HBA2*, and *HBQ* (Figure 2B). Notably, both key genes associated with thalassemia, *HBA1* and *HBA2*, are completely deleted, leading to the manifestation of α^0 -thalassemia. Additionally, based on the patient's hematological characteristics, we noted a more pronounced decrease in MCV, MCH, and other markers when compared to the typical α -globin gene deletions, including $-\alpha^{3.7}/\alpha\alpha$, $-\alpha^{4.2}/\alpha\alpha$, and $-\alpha^{SEA}/\alpha\alpha$ (Xian et al., 2022).

Discussion and conclusion

In the diagnosis of thalassemia, the presence of thalassemia-associated phenotypes is typically first assessed based on the patient's blood count and hemoglobin electrophoresis.

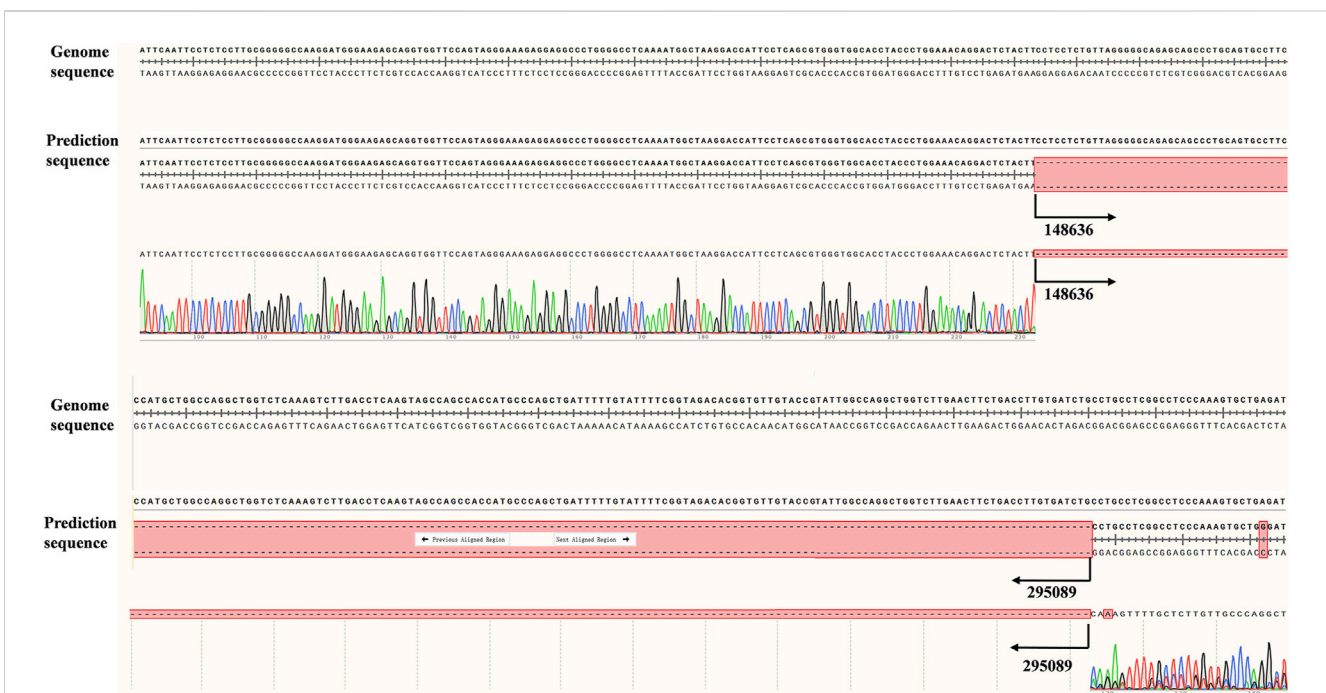


FIGURE 4
Difference between upstream and downstream breakpoints by TGS and Sanger sequencing.

Subsequently, commonly employed thalassemia genetic testing techniques, such as Gap-PCR, PCR-RDB, and flow fluorescence hybridization, are used. However, the conventional assays currently employed in clinical laboratories are limited to detecting high-frequency variants that are specific to the local population. When there is a strong suspicion of a deletion that cannot be identified by conventional methods, we often resort to multiplex ligation-dependent probe amplification (MLPA) for detection. Nonetheless, the density of the MLPA probes restricts its ability to accurately pinpoint the deletion (Li et al., 2023).

After decades of development, next-generation sequencing (NGS) and single-molecule real-time (SMRT) sequencing have emerged as reliable tools for thalassemia, demonstrating both accuracy and stability. These technologies have been further validated in mass screening across various populations in China. NGS enables the detection of not only globin genes but also modifier genes such as *BCL11A*, *MYB*, and *KLF1*; however, it is generally limited by a read length of 300 base pairs (bp) (Wu et al., 2022; Li et al., 2023). While NGS offers high throughput, its sequencing results can be inaccurate in repetitive regions, guanine- and cytosine-rich regions, and highly homologous regions, primarily due to the short read lengths and the PCR amplification that occurs during the sequencing process. Furthermore, NGS struggles to differentiate between homologous genes such as *HBA1* and *HBA2*, as well as *HBB* and *HBD* (Zhou et al., 2023; Huang et al., 2024). In this study, NGS did not successfully detect the genotype of the sample. In contrast, the advantage of long-read single-molecule real-time sequencing technology over NGS lies in its ability to achieve read lengths of up to tens of kilobases. This technology allows for the direct sequencing of long-stranded DNA or RNA molecules without the need for fragmentation or amplification, thereby enabling the accurate identification of large deletions (Long et al., 2022).

Although TGS currently possesses significant technical advantages in diagnosing thalassemia, the high cost of equipment and the lengthy turnaround time for tests limit its widespread application in screening and diagnosing the condition (Liang et al., 2023). In this study, a deletion of approximately 134 kb on the α -globin gene was identified using the CNVplex technique; this deletion encompassed the *HBZ*, *HBM*, *HBA2*, and *HBQ1* genes. We hypothesize that the complete deletion of the *HBA1/HBA2* gene is the primary cause of the abnormal phenotype observed in this proband. Furthermore, we detected a deletion size of 146 kb using long-read SMRT sequencing technology. Notably, the 12-kb fragment that differed from the CNVplex results contained the *AXIN1* gene, which is a component of the *AXIN1-HIPK2-TP53* complex that regulates cell growth, apoptosis, and development (Li et al., 2007).

The CNVplex technology can be used to detect deletion or repeat mutations in each exon of a target gene. However, this technology is not sufficiently accurate to estimate the extent of large deletions in the α -globin gene when compared to SMRT, primarily due to the limitations associated with the number of probes available. In this study, the SMRT sequencing technology was employed based on the approximate deletion extent provided by CNVplex for further detection. Without this approximate range of deletions for the subject, the SMRT sequencing technology identified only one *HBA* haplotype in the sample. We speculate that the types of mutations detectable by the SMRT sequencing system are constrained by the system's limitations in analyzing amplified LD-PCR products, which typically range from 10 to 20 kb.

Consequently, if the triplex system fails to amplify this disease-causing haplotype, larger fragment deletions exceeding 80 kb may remain undetected.

Based on this study, we conclude that the CNVplex technique offers significant advantages, including its relative operational simplicity and its enhanced capability to detect various large-band deletions present in the α -globin gene associated with thalassemia. Thus, we assert that CNVplex serves as an effective primary screening method for identifying rare and novel α -thalassemia deletions. However, to further elucidate the true coverage of the α -globin gene cluster, technologies such as long-read SMRT sequencing are necessary. Our findings from the hematological tests, in conjunction with the SMRT sequencing results, suggest that this deletion represents a *de novo* deletion in the α -globin gene, approximately 146 kb in length, and that its clinical presentation may be more severe compared to the $\alpha\text{-}^{\text{SEA}}/\alpha\alpha$ genotype.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ithanet.eu/db/ithagenes?ithaID=4110>, 4110.

Ethics statement

The studies involving humans were approved by the Ethics Review Committee of Shenzhen Longgang District Maternal and Child Health Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

JX: writing—original draft, writing—review and editing, conceptualization, data curation, investigation, methodology, project administration, software, supervision, and validation. LH: investigation, project administration, software, and writing—review and editing. LW: investigation, methodology, project administration, and writing—review and editing. XC: data curation, investigation, and writing—review and editing. HX: data curation, investigation, methodology, and writing—review and editing. QL: data curation, investigation, and writing—review and editing. YL: writing—review and editing. TJ: writing—review and editing. LS: writing—review and editing. FW: conceptualization, funding acquisition, investigation, methodology, resources, supervision, and writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This

work was supported by grants from Shenzhen Science and Technology Innovation Committee (KJYY20180703173402020) and Shenzhen City and Longgang District Science and Technology Innovation Projects (LGKCYLW2022008) and Shenzhen Science and Technology Research Projects (JCYJ20220530162412029).

Acknowledgments

The authors would like to thank Weiqiang Liu and Yuanyuan Pei for their support of this study. The authors thank Xin Yang for helping with this study.

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.

References

- Colosimo, A., Gatta, V., Guida, V., Leodori, E., Foglietta, E., Rinaldi, S., et al. (2011). Application of MLPA assay to characterize unsolved α -globin gene rearrangements. *Blood Cells Mol. Dis.* 46 (2), 139–144. doi:10.1016/j.bcmd.2010.11.006
- Hu, P., Tan, J., Yu, F., Shao, B., Zhang, F., Zhang, J., et al. (2022). A capillary electrophoresis-based multiplex PCR assay for expanded carrier screening in the eastern Han Chinese population. *NPJ Genom. Med.* 7 (1), 6. doi:10.1038/s41525-021-00280-y
- Huang, R., Liu, Y., Xu, J., Lin, D., Mao, A., Yang, L., et al. (2024). Back-to-Back comparison of third-generation sequencing and next-generation sequencing in carrier screening of thalassemia. *Arch. Pathol. Lab. Med.* 148 (7), 797–804. doi:10.5858/arpa.2022-0168-OA
- Huang, T., Cheng, S., Feng, Y., Sheng, Z., and Gong, Y. (2018). A copy number variation generated by complicated organization of PCDHA gene cluster is associated with egg performance traits in Xinhua E-strain. *Poult. Sci.* 97 (10), 3435–3445. doi:10.3382/ps/pey236
- Jiang, F., Zhou, J., Zuo, L., Tang, X., Li, J., Li, F., et al. (2023). Utilization of multiple genetic methods for prenatal diagnosis of rare thalassemia variants. *Front. Genet.* 14, 1208102. doi:10.3389/fgene.2023.1208102
- Li, C. K. (2017). New trend in the epidemiology of thalassaemia. *Best. Pract. Res. Clin. Obstet. Gynaecol.* 39, 16–26. doi:10.1016/j.bpobgyn.2016.10.013
- Li, Q., Wang, X., Wu, X., Rui, Y., Liu, W., Wang, J., et al. (2007). Daxx cooperates with the Axin/HIPK2/p53 complex to induce cell death. *Cancer Res.* 67 (1), 66–74. doi:10.1158/0008-5472.Can-06-1671
- Li, Y., Liang, L., Guo, W., Wu, X., Qin, T., and Tian, M. (2023). Identification of a novel 107 kb deletion in the alpha-globin gene cluster using third-generation sequencing. *Clin. Biochem.* 113, 36–39. doi:10.1016/j.clinbiochem.2022.12.010
- Liang, Q., He, J., Li, Q., Zhou, Y., Liu, Y., Li, Y., et al. (2023). Evaluating the clinical utility of a long-read sequencing-based approach in prenatal diagnosis of thalassemia. *Clin. Chem.* 69 (3), 239–250. doi:10.1093/clinchem/hvac200
- Liang, Q., Qin, H., Ding, Q., Xie, X., Wu, R., Wang, H., et al. (2017). Molecular and clinical profile of VVD in a large cohort of Chinese population: application of next generation sequencing and CNVplex[®] technique. *Thromb. Haemost.* 117 (8), 1534–1548. doi:10.1160/th16-10-0794
- Long, J., Sun, L., Gong, F., Zhang, C., Mao, A., Lu, Y., et al. (2022). Third-generation sequencing: a novel tool detects complex variants in the α -thalassemia gene. *Gene* 822, 146332. doi:10.1016/j.gene.2022.146332
- Piel, F. B., and Weatherall, D. J. (2014). The α -thalassemias. *N. Engl. J. Med.* 371 (20), 1908–1916. doi:10.1056/NEJMra1404415
- Wei, X., Wang, X., Xiong, F., Zhang, X., Liu, D., Zhou, W., et al. (2024). SNPscan combined with CNVplex as a high-performance diagnostic method for thalassemia. *Prenat. Diagn.* 44 (11), 1310–1317. doi:10.1002/pd.6661
- Wu, H., Chen, D., Zhao, Q., Shen, X., Liao, Y., Li, P., et al. (2022). Long-read sequencing on the SMRT platform enables efficient haplotype linkage analysis in preimplantation genetic testing for β -thalassemia. *J. Assist. Reprod. Genet.* 39 (3), 739–746. doi:10.1007/s10815-022-02415-1
- Xian, J., Wang, Y., He, J., Li, S., He, W., Ma, X., et al. (2022). Molecular epidemiology and hematologic characterization of thalassemia in Guangdong Province, southern China. *Clin. Appl. Thromb. Hemost.* 28, 10760296221119807. doi:10.1177/10760296221119807
- Zhou, Q. M., Jiang, F., Xu, J., Lin, D., Huang, R. L., Zhou, J. Y., et al. (2023). High accuracy of single-molecule real-time sequencing in detecting a rare α -globin fusion gene in carrier screening population. *Ann. Hum. Genet.* 87 (1-2), 9–17. doi:10.1111/ahg.12486

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2025.1518392/full#supplementary-material>