Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Sequencing-free relative haplotype dosage analysis by digital PCR for noninvasive prenatal testing of β-thalassemia: a multicenter blinded validation study

Ying Liu
Xiamen University
Tizhen Yan
Dongguan Maternal and Child Health Care Hospital
Dongzhi Li
Guangzhou Women and Children Medical Center
Siyu Huang
Xiamen University
Haoxian Li
Jiangmen Maternity and Child Health Care Hospital
Jiwu Lou
Dongguan Maternal and Child Health Care Hospital
Yazhe Wu
Xiamen University
Jinmei Xu
Jiangmen Maternity and Child Health Care Hospital
Mingzhu Xu
Xiamen University
Jinjin Zhang
Xiamen University
Xiaoyan Zhang
Dongguan Maternal and Child Health Care Hospital
Ping Chen
First Affiliated Hospital of GuangXi Medical University
Qiuying Huang
Xiamen University
Qingge Li
qgli@xmu.edu.cn

Xiamen University

Research Article

Keywords: Noninvasive prenatal testing, β-thalassemia, Relative haplotype dosage, Digital PCR, Cell-free DNA

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11305925/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Sequencing-free relative haplotype dosage analysis by digital PCR for noninvasive prenatal testing of β-thalassemia: a multicenter blinded validation study

Ying Liu[1,2†], Tizhen Yan[3,4†], Dongzhi Li[5†], Siyu Huang[1], Haoxian Li[6], Jiwu Lou[3,4], Yazhe Wu[1], Jinmei Xu[6], Mingzhu Xu[1], Jinjin Zhang[1], Xiaoyan Zhang[3,4], Ping Chen[7*], Qiuying Huang[1,8*], Qingge Li[1*]

[1] Engineering Research Center of Molecular Diagnostics of the Ministry of Education, State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Xiamen University, Xiamen, Fujian 361102, China

[2] College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, Fujian 361102, China

[3] Prenatal Diagnosis Center, Dongguan Maternal and Child Health Care Hospital, Dongguan, Guangdong 523120, China

[4] Institute of Reproduction and Genetics, Dongguan Maternal and Child Health Care Hospital, Dongguan, Guangdong 523120, China

[5] Prenatal Diagnostic Center, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, Guangdong 510623, China

[6] Medical Genetics Center, Jiangmen Maternity and Child Health Care Hospital, Jiangmen, Guangdong 529000, China

[7] Guangxi Key Laboratory of Thalassemia Research, The First Affiliated Hospital of Guangxi Medical University, Guangxi Medical University, Nanning, Guangxi 530021, China

[8] State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory, Xiamen University, Xiamen, Fujian 361102, China

***Corresponding-authors: Qingge Li, Qiuying Huang, Ping Chen**

E-mail: qgli@xmu.edu.cn; hqying@xmu.edu.cn; cping62@gxmu.edu.cn

Telephone: +8613616064649; +8613950005701; +8613877176778

[†]These authors contributed equally to this work.

Author E-mail

Ying Liu: yingliu@stu.xmu.edu.cn

Tizhen Yan: yantizhen@gdmu.edu.cn

Dongzhi Li: drlidongzhi2014@sina.com

Siyu Huang: syhuang@stu.xmu.edu.cn

Haoxian Li: lihaoxian861212@hotmail.com

Jiwu Lou: loujiwu@163.com

Yazhe Wu: yazhewu@stu.xmu.edu.cn

Jinmei Xu: jimmyoxu@163.com

Mingzhu Xu: mingzhuxu@stu.xmu.edu.cn

Jinjin Zhang: jjzhang1@stu.xmu.edu.cn

Xiaoyan Zhang: 865342@163.com

Ping Chen: cping62@gxmu.edu.cn

Qiuying Huang: hqying@xmu.edu.cn

Qingge Li: qgli@xmu.edu.cn

Abstract

Background: β-thalassemia is among the most common autosomal recessive monogenic disorders in southern China, Southeast Asia and the Mediterranean basin, yet noninvasive prenatal testing (NIPT) for monogenic disorders remains clinically underimplemented because fetal single-nucleotide variants must be quantified against an abundant maternal cell-free DNA (cfDNA) background. Relative mutation dosage, often implemented by digital PCR (dPCR), measures pathogenic variants directly but is vulnerable to maternal background interference, whereas relative haplotype dosage (RHDO) improves robustness by aggregating dosage signals across linked loci but generally requires sequencing-based phasing, limiting accessibility in the resource-limited regions where the disease burden is highest. We hypothesized that patient-specific haplotypes could be resolved experimentally rather than computationally, allowing RHDO to be performed entirely on a benchtop dPCR platform.

Methods: We developed a sequencing-free dPCR-based RHDO assay (dRHDO) for NIPT of β-thalassemia, combining multicolor melting curve analysis to identify informative single-nucleotide polymorphisms (SNPs), a Drop-Phase assay to resolve SNP–mutation linkage experimentally, and multiplex preamplification with dPCR to quantify alleles in maternal plasma cfDNA; fetal fraction was estimated internally from parental genotype-discordant SNPs.

Results: In a blinded multicenter cohort of 96 β-thalassemia families, dRHDO determined fetal genotype with 97.92% (94/96) overall accuracy; all 43 samples meeting the prespecified analytical conditions (cfDNA input ≥5 ng, fetal fraction ≥5%, ≥4 maternally informative loci, ≥1 paternally informative locus) were correctly classified (100%, 43/43), and accuracy was 96.23% (51/53) among the 53 samples not fully meeting these conditions. Both discordant cases were attributable to

insufficient informative-locus redundancy. The workflow is completed in 7.25 h at an estimated cost of approximately $40 per family and requires no sequencing infrastructure.

Conclusions: dRHDO converts relative haplotype dosage from a sequencing-dependent computational strategy into an assay-driven measurement, enabling accurate NIPT for β-thalassemia without sequencing infrastructure.

Keywords: Noninvasive prenatal testing; β-thalassemia; Relative haplotype dosage; Digital PCR; Cell-free DNA

Background

β-thalassemia is one of the most common autosomal recessive monogenic disorders worldwide, with a particularly high carrier frequency in southern China, Southeast Asia and the Mediterranean basin [1-4]. Individuals with β-thalassemia major require lifelong transfusion and iron-chelation therapy [3, 4], and the disease imposes a substantial and largely preventable burden on families and health systems in high-prevalence regions. Because carrier couples face a 25% recurrence risk in each pregnancy [3], prenatal diagnosis is the cornerstone of disease prevention. Current approaches, however, rely on invasive procedures such as chorionic villus sampling and amniocentesis, which carry procedure-related risks [5] and are not always acceptable to pregnant women [6]. An accessible, accurate and rapid noninvasive method for fetal genotyping would therefore have substantial clinical and public-health value in these regions.

Noninvasive prenatal testing (NIPT) using maternal plasma cell-free DNA has transformed prenatal screening for common aneuploidies [7, 8], but monogenic disorders, caused by single-nucleotide variants or small insertions and deletions rather than large copy-number changes, remain a major unmet clinical need [9, 10]. For paternal mutations, fetal inheritance can often be inferred qualitatively by detecting a paternally derived allele absent from the maternal genome, and a range of PCR-based and sequencing-based genotyping approaches have been applied for this purpose [11-17]. Maternally inherited mutations are far more difficult to determine, because fetal alleles must be quantitatively resolved against an abundant maternal background carrying the same allele [18]. This requires dosage-based strategies, principally relative mutation dosage (RMD) and relative haplotype dosage (RHDO) [19].

NGS-based single-molecule counting platforms have improved the analytical performance of RMD analysis [20, 21], but RMD remains fundamentally constrained by its dependence on the pathogenic variant itself rather than the surrounding haplotype context. RHDO addresses this limitation by aggregating dosage information across multiple informative loci and has become the prevailing strategy for maternal mutation inference [19, 22-27]. However, current RHDO approaches generally require sequencing technologies and bioinformatic phasing, including family-based proband analysis, population reference panels [28] and long-read or linked-read sequencing [29, 30]. These requirements increase cost and turnaround time and complicate deployment, particularly in high-prevalence, resource-constrained settings where accessible monogenic NIPT is most needed.

Digital PCR (dPCR) offers absolute molecular quantification with operational simplicity and has emerged as an attractive alternative platform for monogenic NIPT, but existing dPCR-based approaches has focused on RMD analysis or relied on pre-established haplotype information [31-33], and fetal fraction (FF) determination in dPCR-based systems typically requires additional SNP assays and increased cfDNA consumption [33-35]. A central challenge therefore remains: how to achieve reliable multi-locus haplotype inference with integrated FF determination in a simple workflow that requires no sequencing.

We hypothesized that combining experimental haplotype resolution with absolute allele quantification within a dPCR framework would allow RHDO to be implemented entirely at the bench, without sequencing-based phasing or computational haplotype inference. Here, we present dRHDO, a sequencing-free dPCR-based RHDO assay for NIPT of β-thalassemia, in which informative SNPs are identified by multicolor melting curve analysis, SNP–mutation linkage is resolved experimentally by a

Drop-Phase assay, and allele dosage in maternal plasma cfDNA is quantified by multiplex preamplification and dPCR with internally determined fetal fraction. We further define analytical conditions, including cfDNA input, FF and informative-locus availability, that determine inference reliability for each sample. In a blinded multicenter cohort of β-thalassemia carrier couples, we demonstrate accurate fetal genotyping under these conditions, with lower cost and shorter turnaround than sequencing-based workflows.

Methods

Study design and ethical approval

This study developed and validated a dPCR-based RHDO framework for NIPT of β-thalassemia using MMCA genotyping, Drop-Phase phasing, and dPCR quantification. Analytical validation included healthy individuals, carriers, artificial cfDNA standards, and affected families. Clinical validation was performed in a multicenter cohort of 96 β-thalassemia carrier couples recruited from three centers, and the index test was evaluated under a fully blinded design. Plasma samples were analyzed retrospectively using residual plasma obtained during routine clinical testing; reference fetal genotypes were established independently by invasive prenatal diagnosis performed at the participating centers, and these results were withheld from the dRHDO team until test classification was locked. Blinding was implemented with the following steps: (i) samples were re-numbered and coded by a technician who had no access to clinical data or to invasive prenatal diagnosis results; (ii) a second technician performed the dRHDO assay, including genotype calling and haplotype inference, on coded samples, without access to any clinical or reference-standard information; and (iii) dRHDO classifications were

locked before a third analyst compared them with the reference genotypes, so that all statistical analyses were completed before the blind was broken.

Peripheral blood samples were residual specimens from routine clinical testing, collected with written informed consent from participants at the First Affiliated Hospital of Guangxi Medical University, Dongguan Maternal and Child Health Care Hospital and Guangzhou Women and Children's Medical Center. All laboratory analyses were performed at Xiamen University. The study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of Xiamen University (approval no. XDYX2021026).

Sample collection and plasma processing

Peripheral blood was collected into EDTA or Cell-Free DNA Storage Tubes. Plasma was separated by double centrifugation and stored at −80 °C until cfDNA extraction. For gDNA extraction, peripheral blood was collected in EDTA tubes.

cfDNA and gDNA extraction

cfDNA was extracted from plasma using the Apostle MiniMax™ High Efficiency cfDNA Extraction Kit (Apostle Inc., San Jose, CA), and gDNA was extracted from buffy coat using the Lab-Aid 824 automated nucleic acid extraction system and matched extraction reagents. Quantification was performed using the Qubit™ 1X dsDNA HS Assay Kit (Thermo Fisher Scientific, Wilmington).

SNP panel architecture and informative SNP definitions

The final analytical panel comprised 96 SNPs, including 48 *HBB*-flanking SNPs for local haplotype inference and 48 genome-wide autosomal SNPs for FF-iSNP discovery. FF-iSNPs were defined as SNPs homozygous for different alleles in the two parents (A/A versus B/B) and were used for FF estimation. Pat-iSNPs were defined as SNPs heterozygous in the father and homozygous in the mother and were used for paternal mutation tracking. Mat-iSNPs were defined as SNPs heterozygous in the mother and homozygous in the father and were used for maternal haplotype dosage inference. Because the pathogenic variant itself can also serve as an informative locus, we use locus-level terms alongside these SNP-level definitions. Paternally informative loci (Pat-iLoci) comprise the available Pat-iSNPs plus, when the paternal mutation differs from the maternal mutation, the paternal mutation site itself, which is then directly detectable in maternal plasma; when both parents carry the same mutation, paternal origin cannot be resolved by mutation detection and Pat-iLoci are therefore restricted to Pat-iSNPs. Maternally informative loci (Mat-iLoci) comprise the available Mat-iSNPs plus the maternal *HBB* mutation site. Informative-locus counts reported in this work refer to these locus-level numbers.

MMCA, Drop-Phase, multiplex preamplification, and allele quantification

MMCA was performed on a SLAN-96S instrument with five fluorescence channels. Drop-Phase assays experimentally phased SNP alleles with *HBB* mutations on a droplet digital PCR system. Multiplex preamplification enriched target loci, followed by dPCR-based allele quantification. The limit of blank (LOB) was determined for each assay.

Artificial cfDNA standards

Artificial cfDNA standards were prepared by mixing two fragmented gDNA samples to simulate maternal plasma cfDNA at FFs of 3%, 5% and 7%. A wild-type fetal standard was composed of maternal *HBB*:c.216_217insA/N and fetal βN/βN genotypes. A mutant fetal standard contained maternal *HBB*:c.216_217insA/N and fetal *HBB*:c.216_217insA/*HBB*:c.52A>T genotypes. The mutation-bearing fetal standards simulated inheritance of both paternal and maternal mutations. Each standard was preamplified using 5 ng total DNA input, with five replicates for wild-type fetal standards and ten replicates for mutant fetal standards.

FF estimation and parental mutation inference

FF was estimated using allele-resolved dPCR copy numbers from FF-iSNPs, and measurement error was evaluated relative to the expected Poisson variation.

Paternal inheritance was determined by direct detection of a paternal-specific mutation when parental mutations differed, or by Pat-iSNP linkage analysis when they were identical; classification used a Z-score framework incorporating the assay-specific LOB. When no Pat-iLocus was available, paternal transmission was inferred from the dosage of the shared parental mutation alone. Maternal inheritance was inferred using a binomial log-likelihood ratio integrated with Wald's sequential probability ratio test (SPRT), applied to Mat-iSNPs and the maternal mutation site. Cumulative likelihood ratios across multiple informative loci determined the classification. Detailed equations, decision thresholds, and handling of inconclusive results are provided in Additional file 1.

Additional methodological details

Detailed experimental protocols, including sample collection, DNA extraction, reaction conditions, thermal cycling parameters, linkage decay, turnaround time and cost estimation, and cumulative recombination probability calculation are described in Additional file 1.

Statistics

All 95% confidence intervals for proportions were calculated using the Wilson method. Inter-center comparisons of continuous variables (fetal fraction and cfDNA input) were performed using the Kruskal-Wallis test, followed by Dunn's post-hoc test with Bonferroni correction when the overall test was significant. Comparisons of continuous variables between the robust-condition and challenging groups were performed using the Mann–Whitney U test, except for the variables used to define the stratification (cfDNA input, fetal fraction, Mat-iLoci number and Pat-iLoci number); because between-group differences in these variables follow from the stratification criteria by definition, they were not formally tested. Fetal genotype composition, which is not part of the stratification criteria, was compared between groups using the chi-square test. No formal sample-size calculation was performed, as this was a method-development and validation study.

Results

A sequencing-free dRHDO framework

To enable sequencing-free RHDO analysis, we developed a dPCR-based framework, termed dRHDO, that integrates multiplex SNP selection, experimental haplotype phasing, FF estimation and dPCR-based allele quantification. Multiplex SNP genotyping identified three classes of informative SNPs:

Pat-iSNPs for paternal mutation detection, Mat-iSNPs for maternal mutation inference and FF-iSNPs for fetal fraction determination (Figure 1A, B). Together with the mutation sites, these SNPs define paternally informative loci (Pat-iLoci), which additionally include the paternal mutation when the parental mutations differ, and maternally informative loci (Mat-iLoci), which additionally include the maternal *HBB* mutation site. They support two core objectives: local haplotype reconstruction around the β-thalassemia locus (via Pat- and Mat-iSNPs) and independent FF estimation from parental genotype discordance (via FF-iSNPs).

Parental gDNA was analyzed using a Drop-Phase assay to experimentally assign informative SNP alleles to mutant-linked or wild-type–linked *HBB* haplotypes (Figure 1A, C) [36]. All candidate SNPs and *HBB* mutation sites were subsequently co-amplified by multiplex preamplification, and then quantified using the allele quantification assay for the previously selected informative SNPs and parental mutation sites. In the allele quantification assay, each locus was amplified with locus-specific primers and distinguished by dual-color allele-specific probes, so that both alleles at each site are enumerated simultaneously by dPCR [34].

Integration of experimentally resolved haplotype phase, allele dosage and FF enabled reconstruction of haplotype-specific cfDNA representation. The resulting dosage imbalance between mutant-linked and wild-type–linked haplotypes was then interpreted using a quantitative RHDO framework to infer fetal inheritance without sequencing.

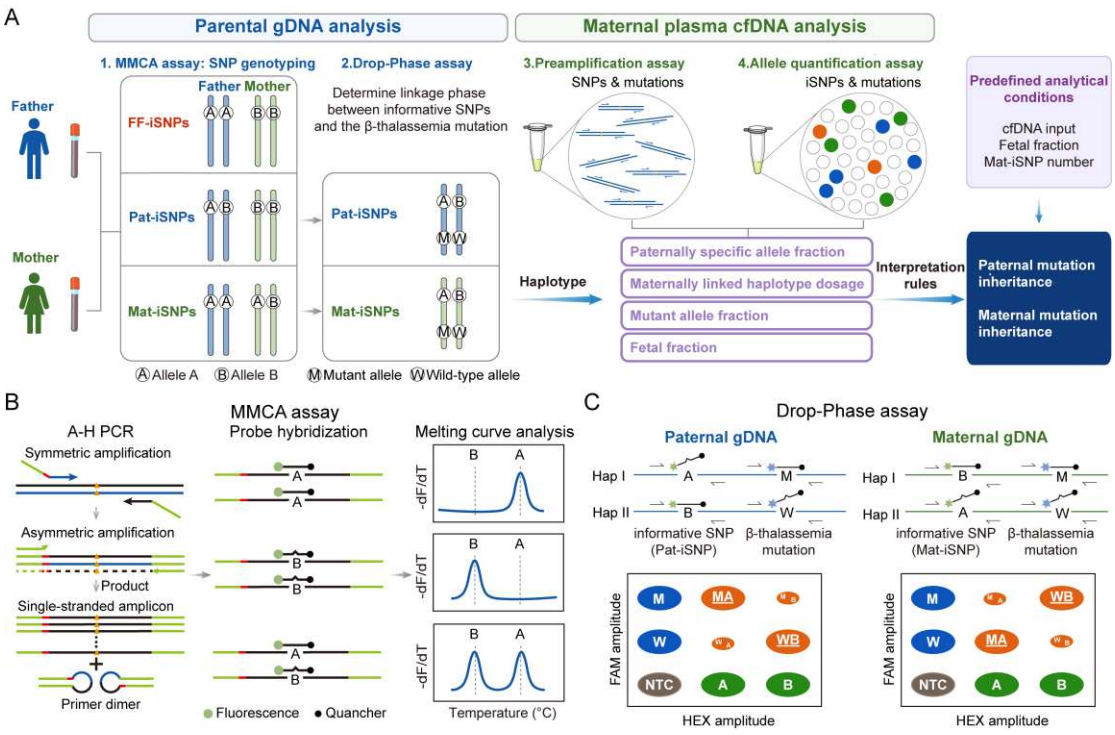


Figure 1. Workflow of the dPCR-based relative haplotype dosage (RHDO) system for noninvasive prenatal testing of β-thalassemia. (A) Overview of the integrated experimental framework for sequencing-independent RHDO analysis. The system combines single nucleotide polymorphism (SNP) genotyping, experimental haplotype phasing, fetal fraction estimation, and allele dosage analysis. Informative SNPs (iSNPs) include fetal fraction-informative SNPs (FF-iSNPs), paternally informative SNPs (Pat-iSNPs), and maternally informative SNPs (Mat-iSNPs), enabling fetal fraction determination and parental haplotype tracking. Maternal plasma cfDNA is subjected to multiplex preamplification followed by dPCR-based allele quantification for Pat-iSNPs, Mat-iSNPs, pathogenic mutations, and FF-iSNPs. Integration of haplotype phase, allele dosage, and fetal fraction enables RHDO-based inference of fetal inheritance. **(B)** Multicolor melting curve analysis (MMCA) for SNP genotyping in parental gDNA. The low-concentration tailed primers and excess universal Tag primer minimize primer-dimer accumulation through the formation of pan-handle structures. Allele-specific self-quenched probes generate distinct melting profiles corresponding to matched,

mismatched, and heterozygous genotypes. A-H PCR, ARMS-HANDS PCR. **(C)** Drop-Phase assay for experimental haplotype phasing. Dual-color dPCR quantifies co-occurrence of SNP and *HBB* mutation signals. Enrichment of double-positive droplets identifies linkage between SNP alleles and mutant- or wild-type–associated haplotypes.

Design and MMCA-based genotyping of informative SNP panels

We first defined the genomic window and SNP architecture for haplotype inference around *HBB*. The mutation panel included ten clinically prevalent β-thalassemia variants in the Chinese population, collectively accounting for >96% of β-thalassemia variants [1, 37]. To guide SNP selection for experimental phasing, we characterized linkage decay across increasing inter-locus distances using gDNA and the Drop-Phase assay (Figure 2A). The dPCR-derived proportion of linked molecules remained above 15% at a distance of 30 kb from the reference SNP rs7938317, substantially exceeding the LOB determined using inter-chromosomal controls. These results established the genomic range suitable for selecting *HBB*-linked SNPs for downstream haplotype reconstruction.

Based on this linkage profile, 48 SNPs located within ±25 kb of *HBB* and with minor allele frequencies (MAF) of 0.1–0.5 were selected (Additional file 1: Figure S1) and incorporated into a four-reaction, five-color MMCA assay (Additional file 1: Figure S2). This panel enabled multiplex SNP genotyping with a detection sensitivity of 0.25 ng DNA per reaction. Screening of 192 unrelated individuals, followed by simulation of 36,672 parental pairings (A_{192}^2), yielded an average of more than ten informative SNPs for parental haplotype reconstruction (Figure 2B). However, FF-iSNPs, defined as SNPs homozygous for different alleles in the two parents, were limited when only *HBB*-linked SNPs

were considered, with a mean of 4.6 FF-iSNPs per simulated family (Additional file 1: Figures S3 and S4). Thus, the locus-restricted panel provided sufficient redundancy for local haplotype reconstruction but limited coverage for stable FF estimation.

To overcome this limitation, we incorporated an additional 48 genome-wide autosomal SNPs with high heterozygosity (MAF approximately 0.5) and broad chromosomal distribution, generating an integrated 96-SNP dual panel design. The two 48-SNP subsets were each configured as a separate four-reaction MMCA set, resulting in a total of eight reactions for complete panel genotyping. Population allele-frequency modeling predicted that this expanded panel increased the probability of obtaining at least three FF-iSNPs to ≥98% (Figure 2C). Simulation across the 36,672 parental pairings generated from the 192 genotyped unrelated individuals independently gave a probability of 97% for obtaining at least three FF-iSNPs, concordant with the allele-frequency prediction. Experimental validation in 192 individuals confirmed the predicted improvement in FF-iSNPs availability, with simulated families yielding an average of > ten FF-iSNPs (Additional file 1: Figure S5). This integrated 96-SNP panel, genotyped in a unified MMCA workflow, therefore supplied the markers for both local haplotype reconstruction around *HBB* and stable FF estimation.

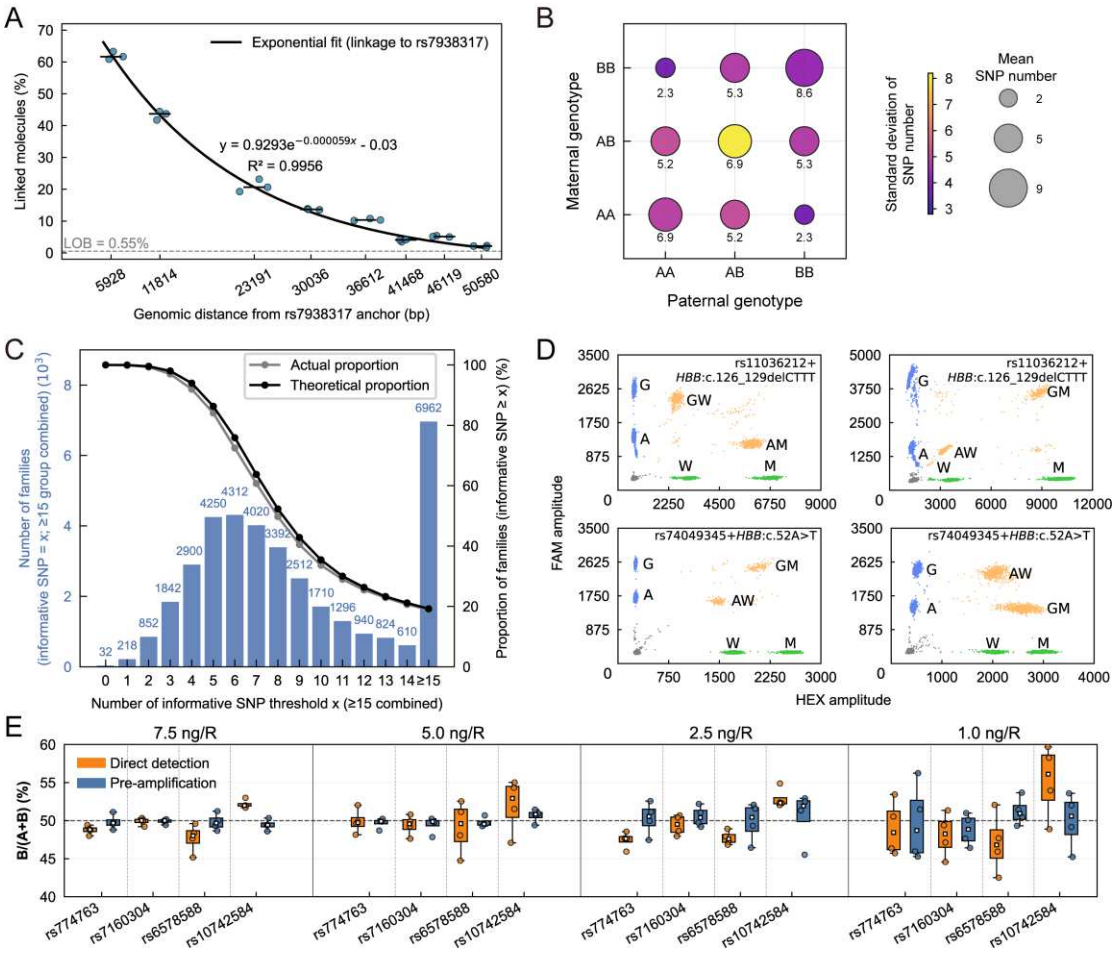


Figure 2. Analytical validation of the dRHDO system. (A) Relationship between genomic distance and molecular linkage proportion using gDNA extracted by the automated magnetic bead method. LOB, limit of blank. **(B)** Bubble plot showing the mean number of informative SNPs across nine parental genotype combinations. The analysis was based on 36,672 ordered parental pairings generated from 192 unrelated individuals (A^2_{192}). **(C)** Distribution of informative SNP numbers in simulated families based on 36,672 ordered parental pairings generated from 192 unrelated individuals. At least three FF-iSNPs were obtained in 97% of these pairings, consistent with the ≥98% predicted by population allele-frequency modeling. **(D)** Representative Drop-Phase assay results for phasing *HBB* mutations with linked distal SNPs. rs11036212 is located 26,168 bp downstream of *HBB*:c.126_129delCTTT, and rs74049345 is located 24,595 bp upstream of *HBB*:c.52A>T. W and M

denote wild-type and mutant alleles, respectively; combined labels indicate the co-occurrence of SNP alleles and mutation or wild-type alleles within droplets, enabling haplotype assignment. **(E)** Comparison of allele ratios measured by direct dPCR and preamplification-based dPCR across different cfDNA input amounts. The dashed horizontal line indicates the expected 50% allelic ratio. Box plots indicate the interquartile range (IQR, 25th–75th percentiles), white squares represent medians, orange or blue dots represent individual data points.

Experimental phasing of SNP–mutation linkage

Having established the composition of informative SNPs within the 96-SNP panel, we next determined whether SNP alleles could be experimentally assigned to mutant-linked or wild-type–linked *HBB* haplotypes. Allele-discriminating probes were designed for each SNP (Additional file 1: Figure S6) and *HBB* mutations (Additional file 1: Figure S7), enabling dual-color Drop-Phase analysis of SNP–mutation pairs (Additional file 1: Figure S8). Among the extraction strategies evaluated, the automated magnetic bead method preserved long-range linkage most efficiently (Additional file 1: Figure S9) and was therefore used for all subsequent phasing analyses.

Phasing performance was assessed using SNP–mutation pairs separated by approximately 25 kb, representing the maximum linkage distance incorporated into the *HBB* haplotype panel. Clear separation of mutant-linked and wild-type–linked haplotypes was achieved for two representative distal SNP–mutation pairs (Figure 2D), supporting the suitability of *HBB*-flanking SNPs for local haplotype reconstruction. Independent validation in β-thalassemia carrier families demonstrated concordant phase assignment with genotyping results (Additional file 1: Figure S10 and Additional file 2: Table S1).

These data establish Drop-Phase as a reliable experimental approach for resolving SNP–mutation linkage and provide the haplotype foundation required for downstream RHDO inference.

Allele quantification and preamplification performance

We next established quantitative allele measurement for low-input cfDNA. Allele quantification assays targeting selected SNPs and *HBB* mutations showed low background signals, with LOBs ranging from 0.0251% to 0.2237% (Additional file 1: Figures S11–S12), which allows sensitive detection of low-abundance alleles.

Because RHDO inference requires accurate dosage measurements across multiple loci, we introduced multiplex preamplification to overcome the limited template availability of maternal plasma cfDNA. To assess the impact of preamplification, we compared targeted preamplification followed by dPCR-based allele quantification with direct dPCR measurements using cfDNA inputs ranging from 1 to 7.5 ng. Four heterozygous targets (two FF SNPs and two *HBB*-flanking SNPs) were analyzed in this comparison (Figure 2E). Preamplification reduced variability in allele-fraction measurements without introducing detectable bias. Although locus-dependent differences in absolute copy number were observed, consistent with amplicon-specific amplification efficiency, detected copy number remained linearly related to cfDNA input across the tested range (Additional file 1: Figure S13). These results demonstrate that multiplex preamplification preserves quantitative allele information while improving measurement stability. Preamplification was therefore integrated into downstream RHDO analysis. Based on clinical sample availability and stable allele-ratio performance, 5 ng cfDNA was selected as the standard input for subsequent experiments.

Informative-locus requirements and FF effects

We next determined how the number of informative loci influences FF estimation and parental mutation inheritance inference. For FF estimation, we first analyzed a carrier–carrier family with 29 FF-iSNPs. FF estimates progressively stabilized as the number of FF-iSNPs increased, converging at 7–8%, while the standard deviation decreased from 0.020 with a single FF-iSNP to 0.0036 with 15 FF-iSNPs (Figure 3A and Additional file 1: Figure S14). Three FF-iSNPs alone were sufficient to substantially reduce variation, and this number of informative loci is achievable in 97% of simulated parental pairings (Figure 2C). We further evaluated artificial wild-type fetal standards by quantifying nine FF-iSNPs. Measurements based on individual FF-iSNPs showed variability approaching the expected Poisson limit, whereas integration of three FF-iSNPs substantially reduced measurement variation (Figure 3B and Additional file 1: Figure S15). Accordingly, three FF-iSNPs were selected for routine FF estimation when available; when fewer loci were available, all available FF-iSNPs were incorporated.

We next assessed the minimal requirement for paternal mutation inference. Using *HBB*:c.52A>T as a model, artificial wild-type and mutant fetal standards were analyzed by dPCR-based allele quantification. Accurate classification of paternal inheritance was achieved when FF exceeded 3%, and a single Pat-iSNP was sufficient for qualitative inference when direct paternal mutation detection was not applicable (Figure 3C). A single Pat-iLocus was therefore prespecified as the minimum requirement for paternal inference, taking the form of the paternal mutation itself when the parental mutations differ, or of one Pat-iSNP when they are identical.

Maternal mutation inference represents the most challenging component of dRHDO because fetal-

derived alleles must be quantified against a predominant maternal cfDNA background. We therefore systematically evaluated the effects of FF, cfDNA input and Mat-iSNP number using artificial standards representing wild-type and mutant fetal genotypes. Assayed loci were selected such that the maternal genome and the wild-type fetal genome were heterozygous, whereas the mutant fetal genome was homozygous, allowing the homozygous allele in the mutant fetal standard to serve as surrogate mutant-linked allele for modelling Mat-iSNP dosage imbalance.

At 5 ng cfDNA input, four Mat-iSNPs were sufficient for accurate maternal mutation inference at FF of 5% or 7% (Figure 3D). Under the more challenging condition of 3% FF, wild-type and mutant fetal standard were evaluated using 12 Mat-iSNPs. To systematically assess the effect of locus number, n loci were randomly selected from the 12 Mat-iSNPs to generate all possible combinations (C_{12}^{n}), and a locus-combination (robustness) analysis was performed: for each informative-locus number, every combination was scored as correct, incorrect or inconclusive according to whether its cumulative SPRT value crossed the decision boundary corresponding to the reference fetal genotype, the opposite boundary, or neither. Because these combinations are resampled from the same 12 loci rather than drawn from independent pregnancies, the resulting proportions describe how inference stabilizes as informative loci are added and do not constitute an independent estimate of clinical accuracy. Under this low-FF condition, 11 Mat-iSNPs were required for all combinations to be classified correctly (Figure 3E and Additional file 1: Figure S16).

Finally, we validated the informative-locus requirements in eight clinical families in which the mother was a β-thalassemia carrier. In three families, allele quantification was performed for one *HBB* mutation site and 19 Mat-iSNPs. The number of discordant single-locus classifications decreased as FF

and cfDNA input increased (Additional file 1: Figure S17). When FF was ≥5% and cfDNA input was ≥5 ng, integration of the mutation site with three Mat-iSNPs was sufficient for accurate maternal mutation inference. When either analytical condition was not satisfied, additional informative loci were required; in the most challenging tested condition, integration of the mutation site with 11 Mat-iSNPs was required for every tested locus combination to be classified correctly (Figure 3F). In the remaining five families, four Mat-iSNPs were analyzed. Although individual Mat-iSNPs occasionally yielded borderline classifications, multi-locus integration generated results fully concordant with invasive prenatal diagnosis (Figure 3G).

These analyses established the informative-locus requirements for dRHDO inference. These thresholds were analytically prespecified as cfDNA input ≥5 ng, FF ≥5%, at least four Mat-iLoci (the *HBB* mutation site and three Mat-iSNPs) and at least one Pat-iLocus (the paternal mutation when the parental mutations differ, or one Pat-iSNP when they are identical), based on artificial standards and a separate set of clinical samples, and were locked before the clinical validation was performed. When these analytical conditions are met, the predefined criteria are sufficient for accurate maternal mutation inference. When they are not met, analyses of samples were exploratory with all available informative loci incorporated to maximize inference confidence, aimed at defining the assay's performance boundaries. For practical clinical implementation, Mat-iSNPs should ideally be selected from both sides of the *HBB* mutation to minimize potential recombination interference. If discordance persists, invasive prenatal diagnosis is recommended for confirmation.

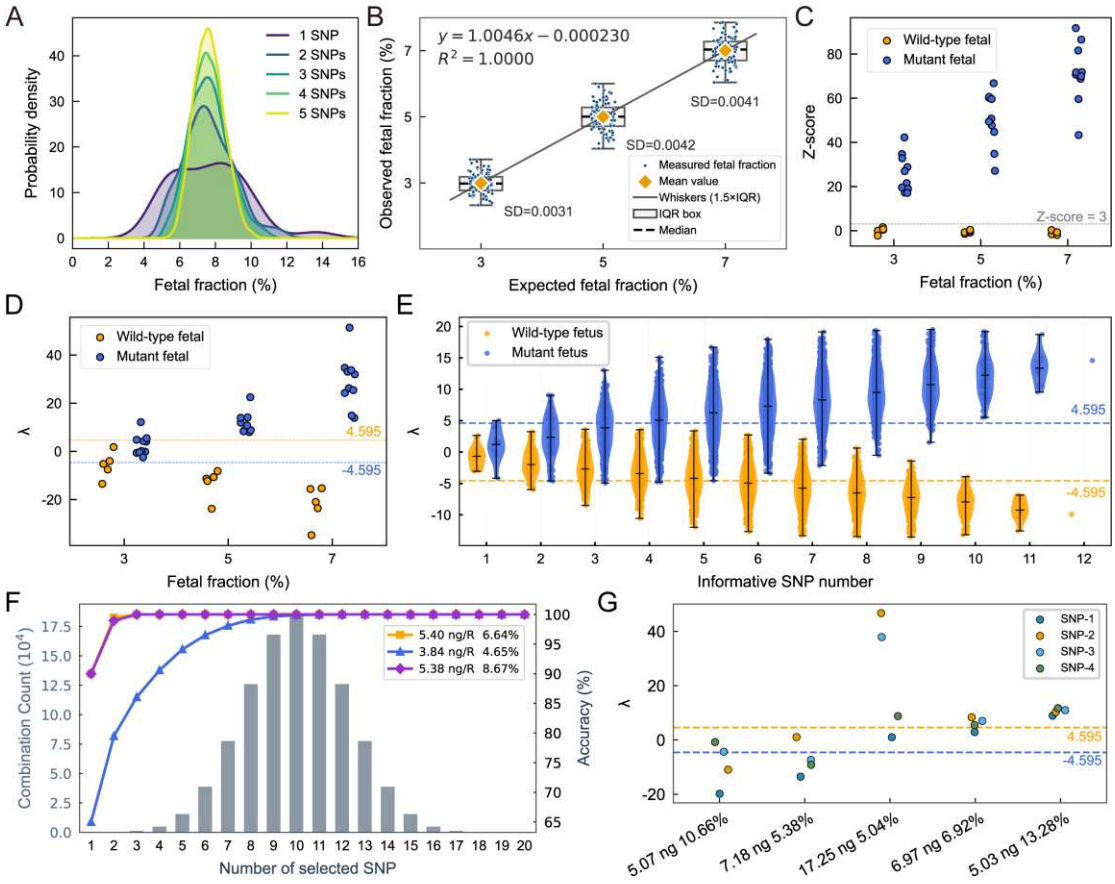


Figure 3. Systematic evaluation of informative SNP requirements for fetal fraction estimation and parental mutation inference. (A) Distribution of fetal fraction estimates derived from different numbers of FF-iSNPs, demonstrating reduced variance with increasing SNP number. **(B)** Correlation between observed and expected fetal fractions, demonstrating linear agreement across simulations. **(C)** Z-score–based classification of paternal inheritance across fetal fraction gradients using 5 ng gDNA input. **(D)** Sequential probability ratio test (SPRT)-based classification of maternal inheritance in simulated standards across different fetal fractions (5 ng cfDNA input). **(E)** SPRT-based locus-combination analysis of maternal inheritance at 3% fetal fraction (5 ng cfDNA input). **(F)** Locus-combination analysis of the relationship between number of selected SNPs and the proportion of correctly classified locus combinations under SPRT analysis. Grey bars indicate the number of all possible combinations from 20 loci; lines represent the proportion of combinations reaching a correct

classification across representative family structures. **(G)** SPRT-based classification results in five clinical families with carrier mothers and wild-type fathers, fully concordant with invasive prenatal diagnosis.

Clinical validation of dRHDO in a multicenter cohort of 96 β-thalassemia families

As the final validation step, we evaluated the clinical performance of the dRHDO framework in a multicenter cohort of 96 β-thalassemia families recruited from three participating centers. The complete analytical workflow is illustrated in a representative family (Additional file 1: Representative family analysis; Additional file 1: Figures S18 and S19).

Across the 96 clinical families, informative-locus availability varied substantially among individuals (Additional file 2: Table S2). The median number of FF-iSNPs was 7 (interquartile range [IQR], 5.00–9.25), Pat-iLoci was 8 (IQR, 3.00–14.25), and Mat-iLoci was 10 (IQR, 3.75–17.00) (Figure 4A, Additional file 1: Figure S20 and Table 1). Consistent with the dual-flanking SNP design, 76 of the 96 families contained at least one Mat-iSNP on each side of the *HBB* mutation, which supports local haplotype reconstruction. cfDNA input ranged from 2.49 to 21.30 ng (median, 8.34 ng), and FF ranged from 2.31% to 13.14% (median, 6.44%). Samples were subsequently classified according to the predefined analytical conditions for robust inference established above, incorporating cfDNA input, FF and informative-locus availability (Figure 4B). Forty-three samples satisfied these conditions, whereas the remaining 53 samples did not fully meet one or more analytical requirements (Table 1). All 43 robust-condition samples had at least one Pat-iLocus (35 through the paternal mutation itself and 8 through Pat-iSNPs), and all six families with no available Pat-iLocus were assigned to the challenging

group. As expected from the prespecified analytical conditions, FF, cfDNA input and Mat-iLoci counts were higher in the robust-condition group than in the challenging group (Table 1). Because these variables define the stratification, the differences follow from the classification criteria rather than representing independent findings, and were not formally tested; they nevertheless confirm that the stratification separated samples with high analytical robustness from those with limited informativeness. The number of FF-iSNPs, which is not part of the criteria, did not differ between the groups (Mann–Whitney U test, $P = 0.446$). Fetal genotype composition also did not differ between the groups (chi-square test, $P = 0.372$). Table 2 shows the per-center distribution of enrolled families, FF, cfDNA input, and diagnostic accuracy. No significant differences were observed across the three centers (Kruskal–Wallis test, FF: $P = 0.24$; cfDNA input: $P = 0.19$). The cohort covered nine of the ten panel mutations among the 192 mutant alleles; the four most prevalent mutations (*HBB*:c.126_129delCTTT, 38.0%; *HBB*:c.316-197C>T, 30.7%; *HBB*:c.52A>T, 14.6%; and *HBB*:c.-78A>G, 10.4%) accounted for 93.7% of mutant alleles, consistent with the reported mutation spectrum in southern China and with the panel design, which covers >96% of β-thalassemia variants.

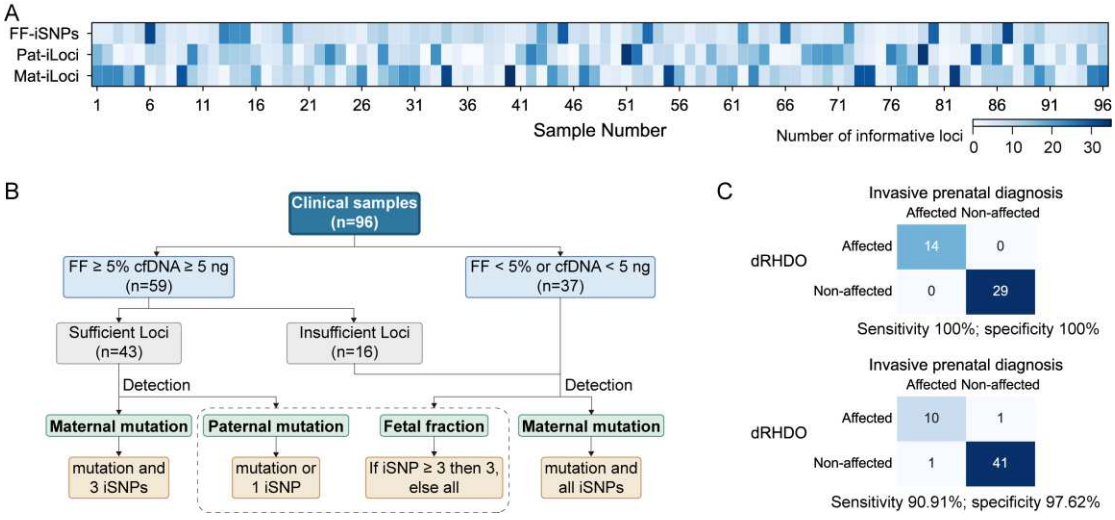


Figure 4. Clinical characteristics and diagnostic performance of the dRHDO system in 96 β-thalassemia pregnancies. (A) Distribution of informative loci across 96 clinical samples. **(B)** Sample classification and detection rule for clinical interpretation. **(C)** Confusion matrices of fetal genotype classification by dRHDO versus invasive prenatal diagnosis in the robust-condition group (upper panel) and the challenging group (lower panel). FF-iSNP, fetal fraction-informative SNP; Pat-iLoci, paternally informative loci; Mat-iLoci, maternally informative loci; iSNP, informative SNP.

Table 1. Characteristics of the 96 β-thalassemia families stratified by analytical robustness.

Characteristic	Overall cohort n=96	Robust-condition samples n=43	Challenging samples n=53	*P*
cfDNA input, median (IQR)	8.34 (5.46-9.86)	8.79 (7.25-10.11)	7.55 (4.86-9.66)	NA
Fetal fraction, median (IQR)	6.44 (4.94-8.17)	7.41 (6.30-8.40)	5.29 (4.25-7.31)	NA
FF-iSNPs, median (IQR)	7.00 (5.00-9.25)	7.00 (5.00-9.50)	6.00 (4.00-9.00)	0.446
Pat-iLoci, median (IQR)	8.00 (3.00-14.25)	9.00 (3.00-17.50)	8.00 (3.00-13.00)	NA
Mat-iLoci, median (IQR)	10.00 (3.75-17.00)	12.00 (7.50-22.00)	5.00 (2.00-14.00)	NA
Same parental mutation, n (%)	31 (32.29)	8 (18.60)	23 (43.40)	NA
Different parental mutations, n (%)	65 (67.71)	35 (81.40)	30 (56.60)	NA
Affected fetus, n (%)	25 (26.04)	14 (32.56)	11 (20.75)	0.372
Carrier fetus, n (%)	49 (51.04)	21 (48.84)	28 (52.83)	NA
Unaffected fetus, n (%)	22 (22.92)	8 (18.60)	14 (26.42)	NA
Correct dRHDO classification, n (%)	94 (97.92)	43 (100)	51 (96.23)	NA

Data are presented as median (IQR) or n (%). The analytical conditions for robust inference were prespecified as cfDNA input ≥5 ng, fetal fraction ≥5%, at least four Mat-iLoci (the *HBB* mutation site and three Mat-iSNPs) and at least one Pat-iLocus (the paternal mutation when the parental mutations differ, or one Pat-iSNP when they are identical). Samples meeting all four criteria were classified as robust-condition (n=43); those failing one or more criteria were classified as challenging (n=53). Because cfDNA input, fetal fraction, Mat-iLoci number and Pat-

iLoci number define the stratification itself, between-group differences in these variables follow from the prespecified criteria rather than representing independent findings; they were therefore not formally compared and are reported as NA. Only the number of FF-iSNPs, which is not part of the stratification criteria, was formally compared between the two groups, using the Mann–Whitney U test. Fetal genotype composition, which is also not part of the stratification criteria, was compared between the two groups using the chi-square test ($\chi^2 = 1.976$; df = 2; $P = 0.372$). Parental mutation concordance was not formally tested, because it is structurally coupled to the Pat-iLocus requirement; all remaining categorical variables are presented for descriptive purposes only. IQR, Interquartile range; NA, not applicable. Pat-iLoci and Mat-iLoci counts include the mutation site.

All 96 samples yielded a conclusive genotype call, and no sample remained in the SPRT inconclusive zone. For samples meeting all four predefined analytical conditions (cfDNA input ⩾5 ng, FF ⩾5%, ⩾4 Mat-iLoci and ⩾1 Pat-iLocus), the minimal prespecified locus set was used for analysis: the *HBB* mutation site plus three Mat-iSNPs for maternal inference (four Mat-iLoci in total) and one Pat-iLocus for paternal inference. All 43 such samples were analyzed with this minimal set and crossed the decision thresholds on first-pass testing, without additional loci or repeat measurements. For the 53 samples not fully meeting the conditions, all available informative loci were incorporated as prespecified and conclusive calls were likewise obtained in every case (see Table 2 for performance in this subgroup). Of the 96 families, 65 had different parental mutations and were resolved by direct paternal allele detection, the paternal mutation itself providing the Pat-iLocus; the remaining 31 carried identical mutations and required haplotype-based inference (Additional file 2: Table S2). Among these 31, 25 had informative Pat-iSNPs and were analyzed using the RHDO approach, and the six families lacking Pat-iSNPs were analyzed using an LLR-based approach.

Table 2. Per-center distribution of enrolled families, sample characteristics, and diagnostic performance.

Center	Families (n)	Median FF (%) (IQR)	Median cfDNA input (ng) (IQR)	Correct calls/total
Guangxi	21	6.61 (4.92-7.79)	9.00 (7.99-9.66)	20/21
Dongguan	41	6.82 (5.29-8.54)	6.86 (5.00-10.30)	41/41
Guangzhou	34	5.91 (4.14-7.52)	8.47 (7.05-10.44)	33/34
Total	96	6.44 (4.94-8.17)	8.34 (5.46-9.86)	94/96

FF, fetal fraction; IQR, Interquartile range. P values refer to comparisons across the three participating centers and were calculated using the Kruskal–Wallis test (fetal fraction, H = 2.885, df = 2, P = 0.236; cfDNA input, H = 3.296, df = 2, P = 0.192). Diagnostic accuracy was not formally compared across centers because only two discordant cases occurred in the entire cohort. Guangxi, the First Affiliated Hospital of Guangxi Medical University; Dongguan, Dongguan Maternal and Child Health Care Hospital; Guangzhou, Guangzhou Women and Children's Medical Center.

Overall fetal genotype accuracy was 97.92% (94/96; 95% CI, 92.72–99.43%). For the 43 samples meeting the predefined analytical conditions (cfDNA input ≥5 ng, FF ≥5%, ≥4 Mat-iLoci and ≥1 Pat-iLocus), dRHDO achieved 100% accuracy (43/43; 95% CI, 91.80–100%). Among the 53 samples that did not fully meet these conditions, accuracy was 96.23% (51/53; 95% CI, 87.25–98.96%), with sensitivity for affected-fetus detection of 90.91% (10/11; 95% CI, 62.26–98.38%) and specificity of 97.62% (41/42; 95% CI, 87.68–99.58%) (Figure 4C). The two discordant cases were attributable to

insufficient maternal-informative locus redundancy. One false-negative case had adequate FF (6.62%) and cfDNA input (8.61 ng) but only one available Mat-iSNP, limiting multi-locus SPRT aggregation. One false-positive case had FF of 7.07% and cfDNA input of 8.30 ng but lacked available Mat-iSNPs, requiring inference primarily from the mutation site alone. Accordingly, samples not meeting the prespecified analytical conditions would not be issued a clinical report; instead, they would be flagged as challenging, prompting repeat sampling, expanded locus testing or confirmatory invasive diagnosis. Consistent with autosomal recessive inheritance in carrier couples, the observed distribution of fetal genotypes (25 affected, 49 heterozygous carriers and 22 unaffected) was close to the expected 1:2:1 Mendelian ratio.

We further assessed the operational characteristics of the dRHDO workflow. The complete assay, including parental MMCA genotyping, Drop-Phase phasing, multiplex preamplification and dPCR-based cfDNA allele quantification, could be completed in 7.25 h (within an 8-h working day) from extracted DNA to final interpretation, with an estimated reagent and consumable cost of approximately $40 per family. Compared with sequencing-based RHDO workflows, which typically require library preparation, sequencing and bioinformatic haplotype analysis, dRHDO provides a simplified workflow with reduced operational complexity and shorter turnaround time. A detailed breakdown of assay time and cost is provided in Additional file 2: Table S3.

Together, these multicenter clinical data demonstrate that dRHDO enables accurate noninvasive fetal genotyping of β-thalassemia across diverse clinical sample conditions while defining the analytical requirements for reliable inference. More broadly, these findings establish that MMCA genotyping, experimentally resolved haplotypes, FF determination and dPCR-based allele dosage can be integrated

to transform RHDO from a sequencing-dependent computational strategy into a scalable assay-driven measurement framework.

Discussion

In this study, we established a sequencing-free dRHDO framework for NIPT of β-thalassemia by integrating multiplex SNP selection, experimental haplotype phasing, FF determination and dPCR-based allele quantification. By combining MMCA-based genotyping of a 96-SNP dual-panel design with Drop-Phase–based physical phasing and allele-resolved dPCR measurement, dRHDO enables haplotype reconstruction and quantitative inference of fetal inheritance without sequencing readouts or population-based haplotype modelling. In dRHDO, haplotype structure is resolved at the bench and dosage imbalance is quantified directly, rather than inferred computationally from sequencing data. We also define analytical conditions, including cfDNA input, FF and informative-locus availability, that determine inference reliability and give the assay explicit operational requirements. This should simplify clinical evaluation and implementation of monogenic NIPT.

Accurate FF estimation is essential for cfDNA-based diagnosis. Existing approaches include Y-chromosome–based methods, methylation-based assays [38, 39], and digital PCR- or SNP-based strategies [34, 35], each with limitations in applicability, cost, or robustness. SNP-based approaches are broadly applicable but typically require additional parental genotyping or cfDNA consumption [34, 35]. In contrast, dRHDO integrates FF estimation directly into the SNP-selection workflow through MMCA-based identification of FF-iSNPs derived from parental homozygous discordance, avoiding additional cfDNA allocation or epigenetic assumptions. By incorporating FF estimation into the same

analytical pipeline used for RHDO inference, dRHDO provides an internally consistent measurement strategy. We further demonstrate that three FF-iSNPs are sufficient for stable FF estimation, with variability approaching the expected Poisson limit and a coefficient of variation below 10%. FF determination therefore adds neither sample consumption nor additional assays.

Paternal inheritance is the easier case in monogenic NIPT. Previous approaches include real-time quantitative PCR [11], PCR-restriction endonuclease digestion [40], dPCR and NGS-based assays [21]. When parents carry different pathogenic variants, paternal inheritance can be inferred directly by detecting a paternally derived allele that is absent from the maternal background. However, when both parents carry the same mutation, parental origin cannot be resolved by mutation detection alone and requires haplotype-based inference, as demonstrated for phenylketonuria [41]. To address this, we established a unified strategy that combines direct mutation detection when parental variants differ with SNP-mediated haplotype inference when mutations are identical. In the latter setting, Pat-iSNPs identified by MMCA and phased using Drop-Phase enabled indirect paternal inheritance inference through dPCR-based haplotype reconstruction. An LLR-based contingency method kept the strategy informative even when no Pat-iLocus was available, a situation that would otherwise defeat SNP-mediated paternal haplotype inference. Together with the scalability of the dPCR-based workflow, this flexibility supports dRHDO as a practical alternative to sequencing-based RHDO for prospective clinical implementation.

Maternal inheritance is harder to resolve, because fetal-derived alleles must be quantified against a predominant maternal cfDNA background. Although dPCR provides high analytical sensitivity, conventional dPCR-based approaches have primarily focused on RMD analysis and lack the ability to

incorporate multi-locus haplotype information [42]. A previous study demonstrated that linked haplotype markers could improve allelic dosage analysis using dPCR; however, this approach relied on pre-established haplotype information and was demonstrated only in limited cases [33]. Here, we extend dPCR-based analysis by incorporating experimentally resolved SNP haplotypes into an RHDO framework, enabling haplotype dosage inference without sequencing-based phasing. By integrating mutant-linked Mat-iSNP dosage with phased SNP information, dRHDO infers maternal inheritance from multi-locus dosage imbalance rather than single-site measurements. Systematic analytical modelling and experimental validation established the informative-locus requirements across different FF and cfDNA inputs. Under predefined analytical conditions (cfDNA input ≥5 ng, FF ≥5%, at least four Mat-iLoci and at least one Pat-iLocus), four Mat-iLoci (the *HBB* mutation site and three Mat-iSNPs) were sufficient for accurate inference. The two discordant cases observed in the full cohort were attributable to insufficient Mat-iSNP redundancy; multi-locus aggregation, rather than assay sensitivity alone, determined RHDO performance.

Compared with existing strategies for monogenic NIPT, dRHDO combines elements that each of them lacks. Sequencing-based haplotype phasing approaches, spanning family- or proband-informed inference, targeted capture sequencing, linked-read and long-read phasing, population reference panels and machine-learning-assisted genotype prediction [22-30], provide genome-wide haplotype information and are broadly applicable across diseases, but they require sequencing infrastructure, specialized bioinformatic phasing and longer turnaround times. As noted earlier, RMD-based approaches remain centered on the pathogenic variant itself. [20, 21]. Conventional dPCR-based RMD assays [31, 32] offer absolute quantification with operational simplicity, yet are restricted to a small

number of loci and lack integrated haplotype information. dPCR-based haplotype approaches relied on pre-established population-specific haplotypes [33]. A recently reported optimized NIPT for thalassemia based on changes in haplotype doses, which combines cell-free fetal DNA enrichment with hybrid-capture sequencing and requires a reference offspring for family-based haplotyping, achieved 100% concordance with invasive testing in 93 clinical families [43]; that workflow nonetheless depends on sequencing-based haplotyping, which dRHDO replaces with experimental droplet phasing. dRHDO addresses these gaps by experimentally resolving patient-specific haplotypes and aggregating dosage across multiple informative loci within a single dPCR workflow. In the present cohort, this translated into a turnaround time of 7.25 h and an estimated cost of approximately $40 per family. Sequencing-based haplotype phasing approaches, by contrast, require sequencing instrumentation, specialized bioinformatics personnel and multi-day workflows; because most of these methods have so far been implemented as research protocols rather than established clinical services, directly comparable cost figures are scarce. The present comparison is therefore based on operational characteristics rather than a formal head-to-head cost evaluation, which remains to be performed.

Recombination between pathogenic variants and linked SNPs represents a potential limitation of RHDO-based approaches. Previous studies using broad genomic regions have identified occasional recombination-associated discordant signals during haplotype inference [44, 45]. In contrast, dRHDO minimizes this risk through a restricted SNP design strategy, using *HBB*-flanking SNPs within a ±25 kb region and selecting informative loci on both sides of the mutation. This dual-flanking design reduces the likelihood that a single recombination event would disrupt all informative signals simultaneously. The cumulative recombination probability across the panel of 48 *HBB*-flanking SNPs (a panel-level

estimate, not a per-family value) was <0.0047 from the deCODE population-based recombination map (Additional file 1) [46], supporting a low estimated recombination risk within the interrogated region. Consistent with this design principle, no recombination-related inconsistencies were observed among the 96 clinical families. Nevertheless, when discordant inheritance signals occur across multiple loci, increasing informative-locus coverage and evaluating locus-specific recombination probability will improve confidence; if the discordance persists, invasive prenatal diagnosis is recommended for confirmation [45]. For paternal inference, the current implementation relies on a single Pat-iSNP when parental mutations are identical; future expansion to multiple flanking Pat-iSNPs may further reduce dependence on single-locus measurements.

Several limitations should be considered. First, some samples exhibited low cfDNA input or low FF, reducing inference robustness. Across the cohort, cfDNA input median 8.34 ng and FF median 6.44%, reflecting the use of residual plasma from routine clinical testing (1–2 mL per sample), including standard EDTA tubes and pre-analytical delays, rather than dedicated cfDNA preservation tubes. Maternal gDNA contamination, a recognized consequence of such pre-analytical conditions, was directly visualized by capillary electrophoresis in the sample with the lowest FF in the cohort (FF = 2.31%; Additional file 1: Figure S21), although its overall contribution to the reduced FF across the cohort was not systematically quantified. The median FF in this cohort (6.44%) was lower than the ~11.6% typically reported in routine NIPT [47]. Both limitations are addressable prospectively by using cfDNA preservation tubes, collecting 5–10 mL of plasma, and ensuring timely processing. For cases in which FF remains suboptimal despite proper collection, enrichment of fetal-derived cfDNA fragments [27, 43] or real-time fragment-length assessment could further improve reliability. Second,

occasional insufficient informative SNP coverage occurred in some families, likely due to population-specific linkage disequilibrium patterns. Expanding the SNP panel or integrating population haplotype reference databases [28] may improve performance across diverse populations and additional monogenic disorders. Third, this study was conducted in southern China and focused on β-thalassemia, where mutation spectra and linkage disequilibrium are population-specific. SNP informativeness may differ in other ethnic groups, and the β-thalassemia mutation spectrum varies considerably outside East and Southeast Asia. Thus, the SNP panel and analytical conditions should be prospectively evaluated before clinical use in other populations. Nevertheless, the experimental framework (multiplex SNP genotyping, physical haplotype phasing, and dPCR-based allele quantification) is readily transferable to other diseases and populations with appropriate SNP adaptation.

From a translational perspective, several requirements should be met before routine clinical adoption. First, the assay needs parental blood samples for MMCA genotyping and Drop-Phase phasing in addition to maternal plasma, so implementation is best suited to prenatal diagnostic centers that already manage carrier couples. Second, pre-analytical conditions matter: the low median FF observed here reflected residual plasma and pre-analytical delays with standard EDTA tubes, and prospective use of cfDNA preservation tubes with 5–10 mL plasma collection is expected to improve both FF and informative-locus availability. Third, the current SNP panel is optimized for the Chinese population, and panel adaptation guided by population-specific allele frequencies will be required in other regions. Operationally, dRHDO is intended not to replace invasive prenatal diagnosis in all pregnancies, but to provide a rapid noninvasive option for carrier couples when the predefined analytical conditions are met; samples falling outside these conditions are reported as challenging, prompting repeat sampling,

expanded locus testing or confirmatory invasive diagnosis. This decision structure is particularly relevant for high-prevalence regions where large numbers of carrier couples require accessible prenatal genetic testing.

Conclusions

In summary, we established a sequencing-free RHDO strategy on a dPCR platform for noninvasive prenatal testing of β-thalassemia. Under predefined analytical conditions, the assay provided accurate results in the multicenter cohort, supporting its analytical and clinical validity in eligible cases. Compared with sequencing-based RHDO approaches [27, 48, 49], dRHDO does not require sequencing, avoids complex bioinformatic phasing, and offers substantially shorter turnaround times, enabling same-day analysis from extracted DNA to result. Compared with conventional dPCR-based RMD strategies, dRHDO introduces experimentally resolved haplotype information, converting single-locus dosage measurement into multi-locus haplotype dosage inference. Although challenges remain for samples with limited FF or insufficient informative loci, the analytical conditions and classification framework established here provide a practical foundation for clinical translation. More broadly, dRHDO offers a scalable strategy for extending RHDO principles to a wider range of autosomal recessive inherited disorders.

List of abbreviations

cfDNA: cell-free DNA; CI: confidence interval; dPCR: digital PCR; dRHDO: dPCR-based relative haplotype dosage; FF: fetal fraction; FF-iSNP: fetal fraction-informative SNP; gDNA: genomic DNA;

IQR: interquartile range; LLR: log-likelihood ratio; LOB: limit of blank; MAF: minor allele frequency; Mat-iLoci: maternally informative loci; Mat-iSNP: maternally informative SNP; MMCA: multicolor melting curve analysis; NGS: next-generation sequencing; NIPT: noninvasive prenatal testing; Pat-iLoci: paternally informative loci; Pat-iSNP: paternally informative SNP; RHDO: relative haplotype dosage; RMD: relative mutation dosage; SNP: single nucleotide polymorphism; SPRT: sequential probability ratio test

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Xiamen University (approval no. XDYX2021026). Written informed consent was obtained from all participants.

Consent for publication

Not applicable. All data are presented in de-identified form and no individually identifiable information is included.

Acknowledgements

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files.

Funding

This work was supported by the grant from National Natural Science Foundation of China [82271729]; the Guangxi Science and Technology Major Project — Construction of China-ASEAN Science and Technology City [GuikeAA24206002]; the Scientific Research Foundation of State Key Laboratory of Vaccines for Infectious Diseases, Xiang An Biomedicine Laboratory [2023XAKJ0103068]; the Guangxi Science and Technology Program [GuikeAD23026025]; the Open Project of Guangxi Key Laboratory of Thalassemia Research; and the Guangdong Basic and Applied Basic Research Foundation, Dongguan Joint Fund (Regional Cultivation Project) [2025A1515140138].

Authors' contributions

Y.L. and Q.H. conceived and designed the study. Y.L., S.H., J.L., J.X., Y.W., M.X., J.Z., and X.Z. acquired the data and performed the analyses. Y.L., T.Y., D.L., S.H., and H.L. performed the statistical analysis. Y.L. drafted the manuscript. T.Y., D.L., P.C., Q.H., and Q.L. supervised the study. Y.L. and Q.H. directly accessed and verified the underlying data reported in the manuscript. All authors read and

approved the final manuscript.

Additional files

Additional file 1: Supplementary Methods and Figures.

File format: PDF.

Title of data: Supplementary Methods and Figures.

Description of data: This file contains detailed supplementary methods, including sample collection and plasma processing, cfDNA and gDNA extraction, MMCA assay for SNP genotyping, Drop-Phase assay for experimental haplotype phasing, multiplex preamplification of cfDNA, allele quantification by dPCR, linkage decay analysis, FF estimation, paternal and maternal mutation inference, turnaround-time and cost analysis, cumulative recombination probability calculation, validation of linkage inference, and representative family analysis. It also includes Additional file 1: Figures S1–S21.

Additional file 2: Supplementary Tables S1–S3.

File format: XLSX.

Title of data: Supplementary Tables S1–S3.

Description of data: This file contains three supplementary tables. Table S1: Results of the dPCR phasing assay for the validation family. Table S2: Characteristics of the 96 β-thalassemia families, including maternal and paternal genotypes, cfDNA concentration and input, fetal DNA fraction, number of informative loci, robust condition status, dRHDO results, fetal genotype by invasive prenatal diagnosis, concordance, and notes. Table S3: Detailed assay time and cost of dRHDO.

References

1. Lai K, Huang G, Su L, He Y: **The prevalence of thalassemia in mainland China: evidence from epidemiological surveys**. *Sci Rep* 2017, **7**(1):920.
2. Weatherall D: **The inherited disorders of haemoglobin: an increasingly neglected global health burden**. *Indian J Med Res* 2011, **134**(4):493–497.
3. Taher AT, Weatherall DJ, Cappellini MD: **Thalassaemia**. *Lancet* 2018, **391**(10116):155–167.
4. Kattamis A, Forni GL, Aydinok Y, Viprakasit V: **Changing patterns in the epidemiology of β-thalassemia**. *Eur J Haematol* 2020, **105**(6):692–703.
5. Salomon LJ, Sotiriadis A, Wulff CB, Odibo A, Akolekar R: **Risk of miscarriage following amniocentesis or chorionic villus sampling: systematic review of literature and updated meta-analysis**. *Ultrasound Obstet Gynecol* 2019, **54**(4):442–451.
6. Canh Chuong N, Minh Duc D, Anh ND, Thi Vui L, Pham Sy Cuong L, Thi Thuy Duong D, Thi Thu Ha B: **Amniocentesis test uptake for congenital defects: Decision of pregnant women in Vietnam**. *Health Care Women Int* 2018, **39**(4):493–504.
7. Chiu RW, Chan KC, Gao Y, Lau VY, Zheng W, Leung TY, Foo CH, Xie B, Tsui NB, Lun FM *et al*: **Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma**. *Proc Natl Acad Sci U S A* 2008, **105**(51):20458–20463.
8. Lo YM, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, Wainscoat JS: **Presence of fetal DNA in maternal plasma and serum**. *Lancet* 1997, **350**(9076):485–487.
9. Hanson B, Scotchman E, Chitty LS, Chandler NJ: **Non-invasive prenatal diagnosis (NIPD): how analysis of cell-free DNA in maternal plasma has changed prenatal diagnosis for monogenic disorders**. *Clin Sci (Lond)* 2022, **136**(22):1615–1629.
10. Wang C, Hou D, Wang G, Wu X, Zhou Z, Yang L, Liu Y, Wang X: **Progress, clinical application and challenges of non-invasive prenatal testing for monogenic diseases**. *Front Pediatr* 2026, **14**:1734842.
11. Chiu RW, Lau TK, Leung TN, Chow KC, Chui DH, Lo YM: **Prenatal exclusion of beta thalassaemia major by examination of maternal plasma**. *Lancet* 2002, **360**(9338):998–1000.
12. van den Oever JM, van Minderhout IJ, Harteveld CL, den Hollander NS, Bakker E, van der Stoep N, Boon EM: **A Novel Targeted Approach for Noninvasive Detection of Paternally Inherited Mutations in Maternal Plasma**. *J Mol Diagn* 2015, **17**(5):590–596.
13. Perlado S, Bustamante-Aragonés A, Donas M, Lorda-Sánchez I, Plaza J, Rodríguez de Alba M: **Fetal Genotyping in Maternal Blood by Digital PCR: Towards NIPD of Monogenic Disorders Independently of Parental Origin**. *PLoS One* 2016, **11**(4):e0153258.
14. Pacault M, Verebi C, Lopez M, Vaucouleur N, Orhant L, Deburgrave N, Leturcq F, Vidaud D, Girodon E, Bienvenu T *et al*: **Non-invasive prenatal diagnosis of single gene disorders by paternal mutation exclusion: 3 years of clinical experience**. *Bjog* 2022, **129**(11):1879–1886.
15. Tan Y, Jian H, Zhang R, Wang J, Zhou C, Xiao Y, Liang W, Wang L: **Applying amplification refractory mutation system technique to detecting cell-free fetal DNA for single-gene disorders purpose**. *Front Genet* 2023, **14**:1071406.

16. Yang L, Wu Y, Hu Z, Zhang H, Pu D, Yan H, Zhang S, Jiang H, Liu Q, Yuan Y *et al*: **Simultaneous detection of fetal aneuploidy, de novo FGFR3 mutations and paternally derived β-thalassemia by a novel method of noninvasive prenatal testing**. *Prenat Diagn* 2021, **41**(4):440–448.
17. Brand H, Whelan CW, Duyzend M, Lemanski J, Salani M, Hao SP, Wong I, Valkanas E, Cusick C, Genetti C *et al*: **High-Resolution and Noninvasive Fetal Exome Screening**. *N Engl J Med* 2023, **389**(21):2014–2016.
18. Lun FM, Tsui NB, Chan KC, Leung TY, Lau TK, Charoenkwan P, Chow KC, Lo WY, Wanapirak C, Sanguansermsri T *et al*: **Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma**. *Proc Natl Acad Sci U S A* 2008, **105**(50):19920–19925.
19. Lo YM, Chan KC, Sun H, Chen EZ, Jiang P, Lun FM, Zheng YW, Leung TY, Lau TK, Cantor CR *et al*: **Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus**. *Sci Transl Med* 2010, **2**(61):61ra91.
20. Lv W, Wei X, Guo R, Liu Q, Zheng Y, Chang J, Bai T, Li H, Zhang J, Song Z *et al*: **Noninvasive prenatal testing for Wilson disease by use of circulating single-molecule amplification and resequencing technology (cSMART)**. *Clin Chem* 2015, **61**(1):172–181.
21. Yang X, Zhou Q, Zhou W, Zhong M, Guo X, Wang X, Fan X, Yan S, Li L, Lai Y *et al*: **A Cell-free DNA Barcode-Enabled Single-Molecule Test for Noninvasive Prenatal Diagnosis of Monogenic Disorders: Application to β-Thalassemia**. *Adv Sci (Weinh)* 2019, **6**(11):1802332.
22. Li J, Lu J, Su F, Yang J, Ju J, Lin Y, Xu J, Qi Y, Hou Y, Wu J *et al*: **Non-Invasive Prenatal Diagnosis of Monogenic Disorders Through Bayesian- and Haplotype-Based Prediction of Fetal Genotype**. *Front Genet* 2022, **13**:911369.
23. Liautard-Haag C, Durif G, VanGoethem C, Baux D, Louis A, Cayrefourcq L, Lamairia M, Willems M, Zordan C, Dorian V *et al*: **Noninvasive prenatal diagnosis of genetic diseases induced by triplet repeat expansion by linked read haplotyping and Bayesian approach**. *Sci Rep* 2022, **12**(1):11423.
24. Wang J, Gao P, Cao Q, Chen F, Song J, Wang C, Dou J, Wu Y, Niu Q, Li J *et al*: **Haplotype-based non-invasive prenatal diagnosis of recessive dystrophic epidermolysis bullosa via targeted capture sequencing of maternal plasma**. *J Dermatol* 2023, **50**(7):946–950.
25. Hanson B, Shaw J, Povarnitsyn N, Bowns B, Young E, Gerrish A, Allen S, Scotchman E, Chitty LS, Chandler NJ: **Expanding Access to Noninvasive Prenatal Diagnosis for Monogenic Conditions to Consanguineous Families**. *Clin Chem* 2024, **70**(5):727–736.
26. Liscovitch-Brauer N, Mesika R, Rabinowitz T, Volkov H, Grad M, Matar RT, Basel-Salmon L, Tadmor O, Beker A, Shomron N: **Machine learning-enhanced noninvasive prenatal testing of monogenic disorders**. *Prenat Diagn* 2024, **44**(9):1024–1032.
27. Zhang L, Hua R, Wu Y, Han X, Wu Y, Fei H, Zhao X, Chang C, Gao L, Chen Y *et al*: **Universal noninvasive prenatal diagnosis for monogenic disorders using cell-free plasma DNA**. *Genome Med* 2025, **18**(1):4.
28. Chen C, Li R, Sun J, Zhu Y, Jiang L, Li J, Fu F, Wan J, Guo F, An X *et al*: **Noninvasive prenatal testing of α-thalassemia and β-thalassemia through population-based parental haplotyping**. *Genome Med* 2021, **13**(1):18.

29. Chen C, Chen M, Zhu Y, Jiang L, Li J, Wang Y, Lu Z, Guo F, Wang H, Peng Z *et al*: **Noninvasive prenatal diagnosis of monogenic disorders based on direct haplotype phasing through targeted linked-read sequencing**. *BMC Med Genomics* 2021, **14**(1):244.
30. Lee JS, Lee KB, Song H, Sun C, Kim MJ, Cho SI, Lee YK, Park SS, Seong MW: **Noninvasive prenatal test of single-gene disorders by linked-read direct haplotyping: application in various diseases**. *Eur J Hum Genet* 2021, **29**(3):463–470.
31. Constantinou CG, Karitzi E, Byrou S, Stephanou C, Michailidou K, Makariou C, Hadjilambi G, Christofides A, Kleanthous M, Papasavva T: **Optimized Droplet Digital PCR Assay on Cell-Free DNA Samples for Non-Invasive Prenatal Diagnosis: Application to Beta-Thalassemia**. *Clin Chem* 2022, **68**(8):1053–1063.
32. Tangwerapornpong C, Traisrisilp K, Charoenkwan P, Tongprasert F, Phusua A, Fanhchaksai K: **Non-invasive prenatal diagnosis of beta-thalassemia disease using digital PCR**. *Front Med (Lausanne)* 2026, **13**:1771405.
33. Gu W, Koh W, Blumenfeld YJ, El-Sayed YY, Hudgins L, Hintz SR, Quake SR: **Noninvasive prenatal diagnosis in a fetus at risk for methylmalonic acidemia**. *Genet Med* 2014, **16**(7):564–567.
34. Camunas-Soler J, Lee H, Hudgins L, Hintz SR, Blumenfeld YJ, El-Sayed YY, Quake SR: **Noninvasive Prenatal Diagnosis of Single-Gene Disorders by Use of Droplet Digital PCR**. *Clinical Chemistry* 2018, **64**(2):336–345.
35. Tan C, Dai P, Zhao Z, Chen X, Wang D, Ren X, Zhu L, Yan T, Lou J, Kong X *et al*: **A Digital PCR Assay for Fetal Fraction Quantification Using Multiplex SNP and Y Chromosome Detection**. *Clin Chem* 2026.
36. Regan JF, Kamitaki N, Legler T, Cooper S, Klitgord N, Karlin-Neumann G, Wong C, Hodges S, Koehler R, Tzonev S *et al*: **A rapid molecular approach for chromosomal phasing**. *PLoS One* 2015, **10**(3):e0118270.
37. Yang Z, Zhou W, Cui Q, Qiu L, Han B: **Gene spectrum analysis of thalassemia for people residing in northern China**. *BMC Med Genet* 2019, **20**(1):86.
38. Chen S, Luan X, Cao X, Liu C, Zhang L, Shang W, Zhang Y, Yang Z, Xu C: **A Fetal Fraction Optimized 106-Plex Digital PCR Assay for Non-Invasive Prenatal Testing of Fetal Trisomy**. *Diagnostics (Basel)* 2026, **16**(11).
39. Chan KC, Ding C, Gerovassili A, Yeung SW, Chiu RW, Leung TN, Lau TK, Chim SS, Chung GT, Nicolaides KH *et al*: **Hypermethylated RASSF1A in maternal plasma: A universal fetal DNA marker that improves the reliability of noninvasive prenatal diagnosis**. *Clin Chem* 2006, **52**(12):2211–2218.
40. Chitty LS, Griffin DR, Meaney C, Barrett A, Khalil A, Pajkrt E, Cole TJ: **New aids for the non-invasive prenatal diagnosis of achondroplasia: dysmorphic features, charts of fetal size and molecular confirmation using cell-free fetal DNA in maternal plasma**. *Ultrasound in Obstetrics & Gynecology* 2011, **37**(3):283–289.
41. Zhu J, Zhao Z, Li S, Zhou Y, Kong L, Fu X, Li H, Feng J, Tang W, Wu D *et al*: **High-Resolution Haplotyping of the PAH Gene Enables Early Gestation Noninvasive Prenatal Diagnosis of Phenylketonuria and Evolution Analysis of Recurrent Pathogenic Variations**. *Prenat Diagn* 2024, **44**(10):1198–1209.
42. Guo Y, Charoenkwan P, Traisrisilp K, Piyamongkol W, Tongprasert F: **Application of Digital**

Polymerase Chain Reaction (dPCR) in Non-Invasive Prenatal Testing (NIPT). *Biomolecules* 2025, **15**(3).

43. Sun F, Zhou Y, Zhao X, Jie Q, Ma L, Lin D, Li Y, Mai Y, Gao G, Zhang Y *et al*: **Application of an optimized non-invasive prenatal testing for thalassemia based on change of haplotype doses**. *J Genet Genomics* 2026, **53**(3):488–496.
44. Hui WW, Jiang P, Tong YK, Lee WS, Cheng YK, New MI, Kadir RA, Chan KC, Leung TY, Lo YM *et al*: **Universal Haplotype-Based Noninvasive Prenatal Testing for Single Gene Diseases**. *Clin Chem* 2017, **63**(2):513–524.
45. Li H, Du B, Jiang F, Guo Y, Wang Y, Zhang C, Zeng X, Xie Y, Ouyang S, Xian Y *et al*: **Noninvasive prenatal diagnosis of β-thalassemia by relative haplotype dosage without analyzing proband**. *Mol Genet Genomic Med* 2019, **7**(11):e963.
46. Kong A, Thorleifsson G, Gudbjartsson DF, Masson G, Sigurdsson A, Jonasdottir A, Walters GB, Jonasdottir A, Gylfason A, Kristinsson KT *et al*: **Fine-scale recombination rate differences between sexes, populations and individuals**. *Nature* 2010, **467**(7319):1099–1103.
47. Deng C, Liu J, Liu S, Liu H, Bai T, Jing X, Xia T, Liu Y, Cheng J, Wei X *et al*: **Maternal and fetal factors influencing fetal fraction: A retrospective analysis of 153,306 pregnant women undergoing noninvasive prenatal screening**. *Front Pediatr* 2023, **11**:1066178.
48. Chen C, Zhu Y, Jiang L, Tang C, Peng Z, Huang J, Yang Y: **A novel and accessible proband-free approach for noninvasive prenatal diagnosis of single-gene disorders**. *Eur J Obstet Gynecol Reprod Biol* 2026, **319**:114988.
49. Jang SS, Lim BC, Yoo SK, Shin JY, Kim KJ, Seo JS, Kim JI, Chae JH: **Targeted linked-read sequencing for direct haplotype phasing of maternal DMD alleles: a practical and reliable method for noninvasive prenatal diagnosis**. *Sci Rep* 2018, **8**(1):8678.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1submitted.pdf
- Additionalfile2submitted.xlsx