Enhancing routine noninvasive prenatal testing with cell-free DNA end characteristics

Mengqi Yang[1,2,#], Linfeng Yang[3,#], Zhe Lin[3,#], Shuo Wang[4,#], Yunyun An[2], Yuqi Pan[2,5,6], Xiaoyi Liu[2,5,6], Zhenyu Zhang[2], Jiguang Wang[1], Xin Jin[3,7,*], Kun Sun[2,*]

[1]Division of Life Science, Department of Chemical and Biological Engineering, and State Key Laboratory of Nervous System Disorders, Hong Kong University of Science and Technology, Hong Kong SAR 999077, China

[2]Institute of Cancer Research, Shenzhen Bay Laboratory, Shenzhen 518132, China

[3]BGI Genomics, Shenzhen 518083, China

[4]Tianjin Women and Children's Health Center, Tianjin 300070, China

[5]Shenzhen Medical Academy of Research and Translation, Shenzhen 518107, China

[6]School of Life Sciences, Westlake University, Hangzhou 310030, China

[7]School of Medicine, South China University of Technology, Guangzhou 510006, Guangdong, China

[#]These authors contribute equally to this work

[*]Corresponding authors: jinxin@genomics.cn (X.J.), sunkun@szbl.ac.cn (K.S.).

Keywords: NIPT; cfDNA fragmentomics; Down Syndrome; nucleosome footprint

Abstract

Objective: Noninvasive prenatal testing (NIPT) based on cell-free DNA (cfDNA) analysis is widely used for detecting fetal aneuploidies globally, such as Trisomy 21 (T21) testing. Despite its high sensitivity, current NIPT methods have a non-negligible rate of false-negative results, primarily due to low fractions of fetal-derived cfDNA in maternal plasma, posing challenges for affected families and public health.

Methods and Analysis: We propose a computational approach to enrich fetal-derived cfDNA by leveraging end characteristics that does not require any modifications to existing experimental protocols. We have evaluated this method using three independent datasets, including over 2,200 samples from diverse ethnic backgrounds and experimental platforms.

Results: Here we show that through end selection, we significantly increase Z-scores in all T21 samples from 3 independent datasets, which shows potential in reducing false-negative results while not introducing any false positives in the euploid samples. Our method is compatible with current routine NIPT workflows that generate low-depth, short, and single-end whole genome sequencing data, therefore allowing for seamless integration with minimal additional cost.

Conclusion: Our method offers translational potential for enhancing routine NIPT by reducing the false negatives, addressing a critical limitation in current clinical practice.

Introduction

In human physiology, dying cells degrade their chromatin and release fragmented DNA into the bloodstream, resulting in cell-free DNA (cfDNA) in plasma. The majority of cfDNA originate from the hematopoietic system [1]; however, other tissues can also significantly contribute to the cfDNA pool under specific conditions. For instance, in pregnant women and cancer patients, the placentae and tumors, respectively, release fetal- and tumor-derived cfDNA into circulation [2, 3]. These phenomena have enabled the development of noninvasive prenatal testing (NIPT) and liquid biopsy approaches for cancer diagnostics. NIPT, particularly for detecting Trisomy 21 (T21, also known as Down Syndrome), has become widely adopted, with over 10 million tests performed annually worldwide [4]. The fraction of fetal cfDNA in maternal plasma is a critical parameter influencing the performance of NIPT, especially the sensitivity [5, 6]. Several factors, including maternal obesity, medical conditions, and therapeutic interventions during pregnancy, are associated with reduced fetal DNA fractions, potentially leading to unreportable (or "no calls") or false-negative results (ranged between 0.02-0.26%) [6, 7]. In real practice, false negative results often lead to missed diagnoses in the absence of following confirmatory invasive testing, and therefore remain a critical issue in NIPT and to the society. To address this issue, various experimental and computational strategies aiming to enrich apparent fetal-derived cfDNA fractions have been proposed over the past decade, most of which leverage the relatively shorter fragment length of fetal-derived cfDNA molecules [8-12]. For example, Liang and colleagues utilized experimental size-selection during library preparation to enrich fetal-derived DNA, and validation studies demonstrated the potential of this approach to reduce "no calls" [13, 14]. Similarly, Budis et al., Kwon et al., and Hu et al. developed and validated that *in silico* size-selection is also helpful [10, 15, 16]. In a previous study, we had developed an alternative algorithm that compares the fraction of reads mapped to chr21 in short and long cfDNA fragments (named "COFFEE") [11], which showed elevated Z-scores compared to the conventional approach [10]. However, these existing methods require additional either experimental procedures, dedicated equipment, or paired-end sequencing, limiting their applicability in routine commercialized NIPT workflows that commonly employ shallow-depth, short, single-end whole-genome sequencing to minimize complexity, turn-around time, and cost [17-21]. Thus, there remains a need for approaches that could enrich fetal cfDNA in routine NIPT protocols without incurring substantial additional expenses.

Previous studies have demonstrated that fetal-derived cfDNA molecules in maternal plasma exhibit distinct cleavage patterns compared to maternal-derived ones [8, 22-25]. Specifically, cfDNA of maternal origin tend to be cleaved at linker regions between nucleosomes, whereas fetal-derived cfDNA more frequently exhibit fragmentation within nucleosome cores [8, 22, 24]. Based on this principle, Straver et al. developed an algorithm to predict fetal DNA fraction in maternal plasma using the ratio of cfDNA ends located within nucleosomes to those in linkers [22]. Interestingly, similar cfDNA end patterns have also been observed in tumor-derived cfDNA from cancer patients. Enrichment strategies based on cfDNA end positioning, such as selectively retaining cfDNA molecules with ends located within hematopoietic nucleosomes, have been shown to improve the detection of tumor-derived cfDNA and facilitate cancer diagnostics [8, 26, 27]. These works suggest the feasibility of incorporating cfDNA end signatures towards improving the performance of routine NIPT assays. In this study, we develop and validate a computational approach that selects cfDNA with ends located within hematopoietic nucleosomes. Our results show that this approach could improve the Z-scores for trisomy testing as well as reduce the false negative calls in conventional Z-scores, therefore holding high translational potential in improving routine NIPT for better serving the society.

Materials and Methods

Ethics approval and sample processing

This work has been approved by the Ethics Committee of Shenzhen Bay Laboratory, the Ethics Committee of BGI-research, and the Ethics Committee of Tianjin Women and Children's Health Center. A total of 2,196 patients with elevated risk for T21 (assessed using ultrasound or known risk factors [28]) were recruited as testing cases, and an additional panel of 20 pregnancies with known euploid fetuses were recruited as the control panel. Both subjects were recruited in Tianjin Women and Children's Health Center, Tianjin, China, during 2024 to 2025 with written-informed consents. For each patient, 6 mL peripheral blood was collected using EDTA-containing tubes and processed within 4 hours. Briefly, blood samples were centrifuged at 1,600 g, 4 °C for 15 min, and then the plasma portion was harvested and re-recentrifuged at 16,000 g, 4 °C for 15 min to remove blood cells. 300 μL plasma was used to extract cfDNA with MGIEasy Circulating DNA Isolation Kit (MGI, #1000017017) and DNA library was constructed using MGI Cell-free DNA Library Prep kit (MGI, #94000018500) following the manufacturer's instructions. All cfDNA libraries were sequenced on an DNBSEQ-T7 sequencer (MGI) in single-end 36 bp mode in multiple runs. After routine NIPT, the euploidy statuses of the fetuses were confirmed using either invasive approaches or after birth.

Sequencing data analysis

CfDNA whole genome sequencing data was processed as reported previously [8, 29]. Briefly, raw cfDNA sequencing data was firstly preprocessed using Ktrim software [30] to remove adapters and low-quality bases. The preprocessed reads were subsequently aligned to the NCBI GRCh38 human reference genome using Bowtie2 software [31]. PCR duplicates, i.e., sequencing reads with identical start and end positions, were identified and filtered out using in-house scripts. An implementation of the whole procedure, including read preprocessing, alignment, and filtering, is freely available from our previous study [29]. Reads with mapping quality scores lower than 30 were discarded from downstream analyses. Moreover, to get rid of dosage biases related to fetal sex, chromosomes X, Y, and mitochondria DNA were excluded from the analysis. Of note, we calculate read counts for chromosomes suffered from common aneuploidies (i.e., 13, 18, and 21) separately, which allows one to utilize these numbers for T13, T18, and T21 testing, as well provides the option to include or exclude chromosomes 13 and 18 into T21 testing (Suppl. Fig. S2). For T21 testing, Z-scores with or without taken chromosomes 13 and 18 into analysis were both reported, and the Z-scores without these two chromosomes were used as the primary results. For T18 testing, Z-scores without taken chromosomes 13 and 21 into the analysis were reported.

Hematopoietic nucleosome track and end selection

Fig. 1 illustrated the principle of cfDNA ending patterns in relative to nucleosome positioning. The maternal-derived cfDNA molecules are primarily released from the hematopoietic system, and therefore the ends of these cfDNA molecules are expected to be concordant with the nucleosome positions of the hematopoietic system, i.e., the majority of the cfDNA ends should locate in linker DNA. In contrast, the fetal-derived cfDNA molecules are shorter and primarily originated from placental tissues, which possess distinct nucleosome positioning patterns with the hematopoietic system, and thus their ends should exhibit lower consistencies to the nucleosome positioning of hematopoietic system. Hence, selecting the cfDNA molecules whose ends are located within hematopoietic nucleosomes would enrich fetal-derived cfDNA.

In this study, the nucleosome track for GM12878 cell line (lymphoblastoid lineage; RRID: CVCL_7526) was utilized to define the hematopoietic nucleosomes, which had been proven applicable for cfDNA fragmentomic analyses in previous studies [8, 32, 33]. The nucleosome track for GM12878 cell line was obtained from NucMap database [34], with accession number hsNuc0390101. This track was constructed from MNase-seq experiments and nucleosome center positions were called using DANPOS algorithm [35]. We defined nucleosome regions as ±73 bp around the annotated center loci of the nucleosome centers as in previous studies [8, 22]. To perform end selection, for a given sample, we compared all the cfDNA 5'-end with the nucleosome track and only reported those reads with 5'-end falling within nucleosome regions. Note that this end selection is performed for both testing samples and the control panel.

Noninvasive prenatal testing for fetal aneuploidies

Z-score-based T21 tests were conducted as reported previously [36]. For each cfDNA sample from maternal plasma, the proportion of reads mapped to chromosome 21 (referred to as "%chr21") was first calculated, either using the raw alignment results to calculate standard Z-score or by focusing on the reads whose ends were located within the hematopoietic nucleosomes to calculate Z-scores after end selection. Then, for the control samples, the mean and standard deviation (s.d.) of their %chr21 values were calculated. Subsequently, the Z-score for each test sample was measured by subtracting the mean %chr21 of the control panel from that of the test sample and then dividing the result by the S.D. of the control group's %chr21 values, as illustrated in the following formula:

$$Z\text{-}score = \frac{\%chr21 \text{ in testing sample - } mean \text{ of } \%chr21 \text{ in controls}}{s.d. \text{ of } \%chr21 \text{ in controls}}$$

Similarly, for T18 testing, we counted the reads mapped to chr18 in both the testing cases and controls samples; we then calculated the mean and s.d. of the proportion of reads mapped to chr18 as %chr18 values, and used the following formula to calculate the Z-scores for testing cases:

$$Z\text{-}score = \frac{\%chr18 \text{ in testing sample} - mean \text{ of } \%chr18 \text{ in contols}}{s.d. \text{ of } \%chr18 \text{ in controls}}$$

A cutoff of 3 was used: samples with Z-scores > 3 were considered as T21 or T18, samples with Z-scores < -3 were considered as positive for monosomy 21 or 18, and the rest were considered as euploidies. For all cohorts investigated in this study, cfDNA end selection procedure was also applied to the control panels, and we calculated the mean and s.d. of fraction of %chr21 using on the reads passing end selection. A similar approach was applied for T18 testing, where chr18 was selected to analyze its read counts in the sample-of-interest compared to the control panel to calculate Z-scores.

Statistics and reproducibility

No statistical method was used to predetermine the sample size. All samples were treated as biological replicates. Paired t-tests were used to compare conventional Z-scores and those with end selection. Mann-Whitney U test was used to compare the changes in Z-scores with end selection between T21 cases and euploidies. For Karlsson dataset, one sample with accession SRR1705799 was omitted due to extra-low read number.

Results

Schematic workflow of the end selection algorithm for NIPT

In routine T21 screening, the Z-score-based approach is typically used [10, 11]. This involves calculating the proportion of sequencing reads mapped to chromosome 21 in a test sample and comparing to a reference panel of known euploid pregnancies. A Z-score exceeding 3 (i.e., three standard deviations above the reference mean) is indicative of T21 case [10, 11]. Fig. 1 illustrates the schematic workflow of our method, which incorporates cfDNA end selection. The key modification involves evaluating the 5'-end position of each sequencing read in relation to hematopoietic nucleosome coordinates: reads with 5'-ends located in linker regions are discarded, while those with 5'-ends within nucleosomes are retained. Since maternal cfDNA molecules are primarily originated from the hematopoietic system and longer in size, they tend to be cleaved in linker regions; in contrast, fetal cfDNA derived from the placental tissues do not follow hematopoietic nucleosome positioning and is more frequently cut within nucleosomes [8, 24]. Hence, this end selection process would preferentially enrich fetal-derived cfDNA fragments and result in elevated apparent fetal DNA fractions in the filtered data, potentially improving NIPT sensitivity.

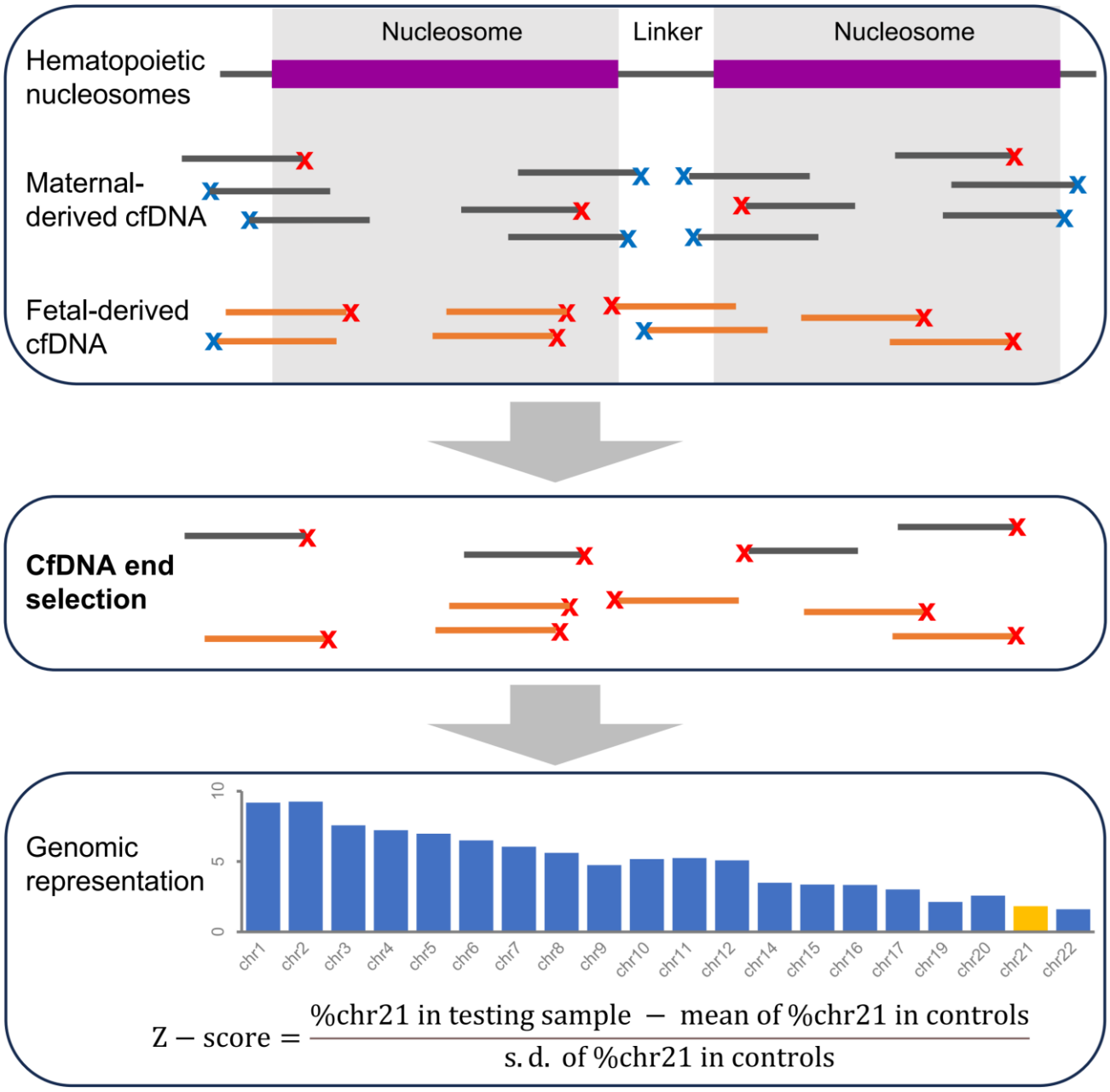

Fig. 1. Cell-free DNA (cfDNA) end selection for enhanced noninvasive prenatal testing. Grey and orange lines represent cfDNA molecules originating from the mother (mostly from the hematopoietic system) and fetus, respectively; the "X" marks represent cfDNA ends located in linker DNA (blue) or within nucleosomes (red). "%chr21" means the fraction of reads mapped to chr21.

Performance evaluation on a cohort with 2,196 samples

To evaluate the effectiveness of our approach, we applied it to a cohort with 2,196 pregnancies recruited in Tianjin, China, of which 207 cases were T21 and 1,989 cases were euploid. For all samples, we collected the maternal blood, then extracted and sequenced the cfDNA to an average of ~9 million (range: 5.0-17.6 million) reads in single-end 36 bp mode using the DNBSEQ platform, following the routine NIPT setting by BGI group, one of the largest NIPT service providers in China [17-21]. We calculated the Z-scores using a control panel with known euploidies, and then applied end selection on all the samples. Of note, the median fraction of reads passing the end selection procedure was 61.16% (range: 58.47-63.95%), and therefore enough number of reads (~5 million) were kept for reliable chromosome counting and Z-score calculation. As a result, the Z-scores after end selection were significantly elevated in all T21 samples, with a median increase of 1.59 ($P<10^{-10}$, paired t-test); the Z-scores were also elevated in euploidies ($P<10^{-10}$, paired t-test), while the increases in Z-scores were very subtle (median: 0.19, $P<10^{-10}$ compared to T21 cases, Mann-Whitney U test; Fig. 2a and Suppl. Fig. S1). Moreover, in the conventional approach, four T21 cases (1.93%) were misclassified as false negatives with Z-scores lower than 3; in contrast, our method reduced the number to two (0.97%; Fig. 2b). Meanwhile, in euploidies, our approach corrected one sample with Z-score > 3 (i.e., potential false-positive) and one sample with Z-score < -3 (i.e., potential false-monosomy) called by the conventional Z-score approach while did not introduce any false positives (i.e., specificity increased from 99.90% to 100.00%; Suppl. Fig. S1), suggesting an improvement of overall accuracy in T21 testing accuracy. Additionally, for the conventional Z-score approach, we tried lowering the diagnostic threshold (i.e., change to 2.64 instead of 3 in this cohort) to rescue the two false-negative cases; however, this parameter would cause two additional false positives (Suppl. Fig. S1), suggesting that this strategy is not appropriate for routine usage and therefore justify the value of improving Z-scores in T21 testing.

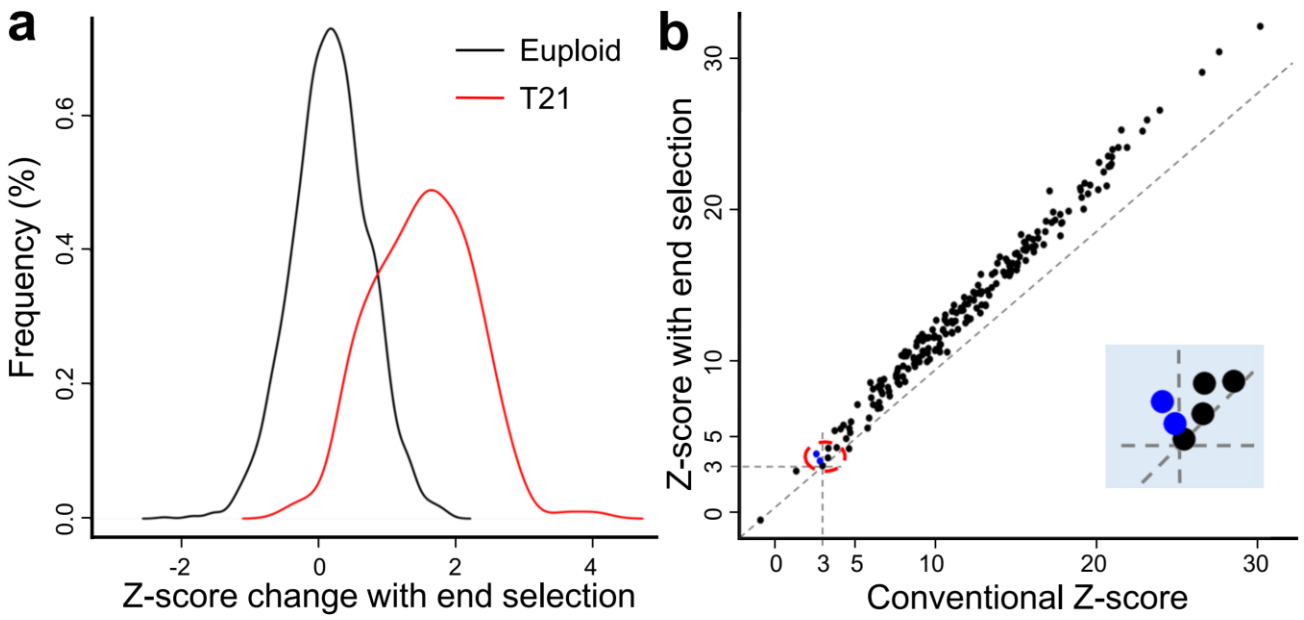

Fig. 2. Validation of our approach for improving routine noninvasive Trisomy 21 (T21) testing. (a) Distribution of changes in Z-scores with end selection between T21 and euploid cases. (b) Comparison of Z-scores in T21 cases. Each dot represents one sample. In (b), red circles highlight the samples with false negative results by conventional Z-scores while are corrected by our approach, and these samples were zoomed-in with light blue background at the bottom-right corner.

Validation on external cohorts

To validate the findings, two publicly available maternal plasma cfDNA datasets were analyzed: the Karlsson dataset [37] comprises 8 T21 (1 sample discarded due to extra-low depth, and 7 kept in analysis), 1 Trisomy 13 (T13), 4 Trisomy 18 (T18), 2 sex chromosome aneuploidies, 12 euploid samples, and a pre-defined control panel with 4 samples; the Chandrananda dataset [38] includes 9 T21 and 20 euploid samples (of which 10 samples were randomly selected to form the control panel). Notably, both datasets were generated using Illumina platforms in single-end mode with different experimental protocols. In these cohorts, the median fraction of reads passing the end selection procedure was 71.24% (range: 69.02-72.38%; Suppl. Fig. S2). The results, summarized in Fig. 3 and Suppl. Fig. S3, demonstrated that Z-scores for T21 samples were significantly elevated following end selection ($P < 0.01$ in both datasets, paired t-tests), while Z-scores for euploid samples remained unchanged ($P > 0.1$ in both datasets, paired t-tests). Importantly, in the Karlsson dataset, conventional Z-scores for 3 out of 7 (42.86%) T21 samples were below the diagnostic threshold (i.e., potential false negative results), while all of them received Z-scores higher than 3 with end selection, and therefore would be correctly reported as T21 by our approach; meanwhile, in Chandrananda dataset, our approach successfully corrected 2 out of 4 (50.00%) potential false negative results by conventional Z-scores.

Of note, there are four T18 cases in Karlsson dataset, and we therefore applied end selection to evaluate the performance of T18 testing on this dataset. As a result, Z-scores for T18 testing were also significantly elevated for the T18 cases with end selection ($P=0.0022$, paired t-test; Suppl. Fig. S4); interestingly, end selection also corrected 1 out of 2 false positive calls in T18 testing (Suppl. Fig. S4) while did not introduce any new false positives in these 2 datasets for both T21 and T18 testing (all specificity=100.00%). Collectively, these results indicate that our cfDNA end selection strategy improves the sensitivity of T21 test in NIPT, particularly by elevating Z-scores and thereby reducing the risk of false-negative outcomes.

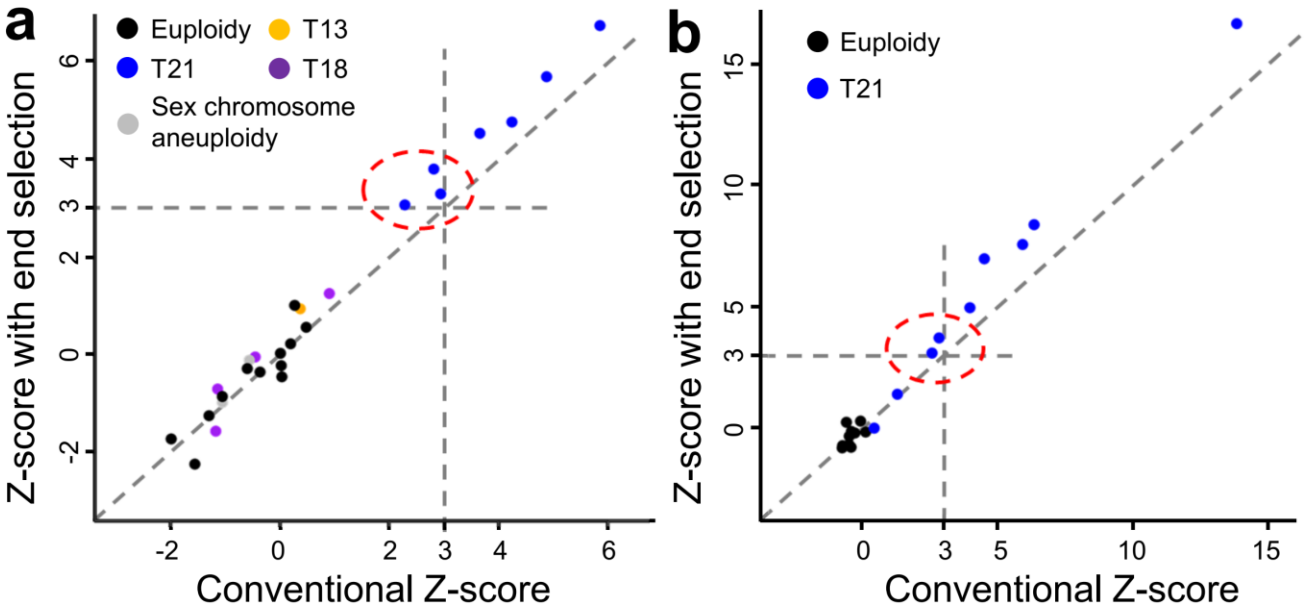

Fig. 3. Validation results in external datasets. Comparison of Z-scores in (a) Karlsson, and (b) Chandrananda datasets. Each dot represents one sample; red circles highlight the samples with false negative results in conventional analysis while are corrected by our approach. Blue and black dots represent T21 and euploidies, respectively. In (a), orange, purple, and grey dots represent Trisomy 13 (N=1), Trisomy 18 (N=4), and sex chromosome aneuploid (N=2) cases, respectively.

Discussion

In clinical practice, false negatives in T21 screening are especially of concern, because these cases often do not undergo further confirmatory examinations and result in severe emotional and societal consequences. Here, we propose a computational method that leverages cfDNA end characteristics to improve T21 detection without altering experimental protocols at all. As validated on 3 datasets, our end selection procedure generated elevated Z-scores compared to the conventional approach in T21 cases. As Z-scores are positively correlated with fetal DNA fractions in T21 cases [39, 40], the data suggested that our end selection improved (apparent) fetal DNA fractions. Crucially, our approach is compatible with conventional shallow-depth, short, single-end sequencing workflows and causes negligible additional cost, making it highly suitable for seamless integration into current whole genomewide sequencing-based NIPT workflows. As a contrast, current size-based computational approaches require paired-end or long-read sequencing data, and therefore could not work on such routine NIPT data generated in this study; hence, head-to-head comparisons are impossible. However, as paired-end sequencing might be available in some regions or service providers, it would be interesting to benchmark our approach versus the current algorithms using such data in future studies.

As shown in 3 cohorts, the evaluated Z-scores generated by our approach could rescue false negative results, thus possessing translational merit in NIPT. Interestingly, our approach also shows the ability to reduce false positive results in trisomy tests (Suppl. Fig. S1 and S4). Furthermore, given that our approach enriches fetal-derived cfDNA in a genome-wide manner, it might be applicable to other chromosomal aneuploidy tests beyond T21; due to data limitations, this feasibility was only roughly explored on four T18 samples in the Karlsson dataset (Suppl. Fig. S4), and therefore it represents an important avenue for future evaluations. Moreover, broader validations with large-scale retrospective or prospective studies are still needed before clinical deployment on specific NIPT platforms and/or settings.

In addition, it would be interesting to explore the feasibility of integrating end selection into other Z-score calculation methods besides the conventional approach. Of note, it is known that different chromosomes have their own unique compositional properties [41], which might introduce biases to the end selection procedure. As evaluated in 3 cohorts, such bias might not significantly affect end selection when comparing read counts from the same chromosome between testing cases and controls in conventional Z-score definitions; however, approaches utilizing within-sample comparisons that use genomic regions from different chromosomes as controls (e.g., the WISECONDOR algorithm [42]) might be affected and therefore requires future investigations. In conclusion, by exploiting cfDNA end characteristics, we present and validate a novel and costless strategy for enhancing routine NIPT assays. Our work expands the utility and highlights translational merit of cfDNA fragmentomics in noninvasive prenatal diagnostics.

Data availability

Raw cfDNA sequencing data used in this study was downloaded from Sequence Read Archive (SRA) with accession numbers SRP028828 and SRP050360. Data generated in this study are available from the corresponding authors upon reasonable requests.

Code availability

Pre-compiled programs and scripts for common x86_64 GNU/Linux operating system to reproduce this work has been deposited to Zenodo (https://zenodo.org/records/17107133; the package would be

made publicly accessible upon acceptance of this manuscript). Source codes under a non-commercial license are available from the corresponding authors upon reasonable requests.

Acknowledgements

This work was supported by National Natural Science Foundation of China (32401206), Guangdong Basic and Applied Basic Research Foundation (2023B1515120073), Shenzhen Clinical Research Center for Oral Diseases (20210617170745001), and National Key R&D Program of China (2022YFA0912700). J.W. was supported by an ITC grant (ITCPD/17-9) and the Padma Harilela Professorship. We'd like to thank Ms. Qi Wang for technical assistance, and Shenzhen Bay Laboratory HPC facility for computational supports.

Author contributions

Conception and design: K.S.; Supervision: X.J., K.S.; Methodology: M.Y., K.S.; Patient recruitment and analysis of clinical data: S.W.; Sequencing data analysis: M.Y., L.Y., Z.L.; Result interpretation: all authors; Writing the manuscript: M.Y., K.S.; Review and edit of the manuscript: M.Y., J.W., X.J., K.S.

Declaration of interests

K.S. had filed patent applications to China National Intellectual Property Administration (202311138167.X) based on the method developed in this work; the remaining authors declare no conflict of interests.

References

1. Sun K, Jiang P, Chan KCA, Wong J, Cheng YK, Liang RH, Chan WK, Ma ES, Chan SL, Cheng SH *et al*: **Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments**. *Proc Natl Acad Sci U S A* 2015, **112**(40):E5503-5512.
2. Lo YMD, 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.
3. Stroun M, Anker P, Maurice P, Lyautey J, Lederrey C, Beljanski M: **Neoplastic characteristics of the DNA found in the plasma of cancer patients**. *Oncology* 1989, **46**(5):318-322.
4. Norton ME, Jacobsson B, Swamy GK, Laurent LC, Ranzini AC, Brar H, Tomlinson MW, Pereira L, Spitz JL, Hollemon D *et al*: **Cell-free DNA analysis for noninvasive examination of trisomy**. *N Engl J Med* 2015, **372**(17):1589-1597.
5. Lopes JL, Lopes GS, Enninga EAL, Kearney HM, Hoppman NL, Rowsey RA: **Most noninvasive prenatal screens failing due to inadequate fetal cell free DNA are negative for trisomy when repeated**. *Prenat Diagn* 2020, **40**(7):831-837.
6. Hui L, Bianchi DW: **Fetal fraction and noninvasive prenatal testing: What clinicians need to know**. *Prenat Diagn* 2020, **40**(2):155-163.
7. Van Opstal D, Srebniak MI, Polak J, de Vries F, Govaerts LC, Joosten M, Go AT, Knapen MF, van den Berg C, Diderich KE *et al*: **False negative NIPT results: risk figures for chromosomes 13, 18 and 21 based on chorionic villi results in 5967 cases and literature review**. *PLoS One* 2016, **11**(1):e0146794.

8. An Y, Zhao X, Zhang Z, Xia Z, Yang M, Ma L, Zhao Y, Xu G, Du S, Wu X *et al*: **DNA methylation analysis explores the molecular basis of plasma cell-free DNA fragmentation**. *Nat Commun* 2023, **14**(1):287.
9. Qiao L, Yu B, Liang Y, Zhang C, Wu X, Xue Y, Shen C, He Q, Lu J, Xiang J *et al*: **Sequencing shorter cfDNA fragments improves the fetal DNA fraction in noninvasive prenatal testing**. *Am J Obstet Gynecol* 2019, **221**(4):345 e341-345 e311.
10. Budis J, Gazdarica J, Radvanszky J, Szucs G, Kucharik M, Strieskova L, Gazdaricova I, Harsanyova M, Duris F, Minarik G *et al*: **Combining count- and length-based z-scores leads to improved predictions in non-invasive prenatal testing**. *Bioinformatics* 2019, **35**(8):1284-1291.
11. Sun K, Chan KCA, Hudecova I, Chiu RWK, Lo YMD, Jiang P: **COFFEE: control-free noninvasive fetal chromosomal examination using maternal plasma DNA**. *Prenat Diagn* 2017, **37**(4):336-340.
12. Lo YMD, Chan KCA, 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.
13. Qiao L, Zhang B, Wu X, Zhang C, Xue Y, Tang H, Tang H, Shi J, Liang Y, Yu B *et al*: **A fetal fraction enrichment method reduces false negatives and increases test success rate of fetal chromosome aneuploidy detection in early pregnancy loss**. *J Transl Med* 2022, **20**(1):345.
14. Liang B, Li H, He Q, Li H, Kong L, Xuan L, Xia Y, Shen J, Mao Y, Li Y *et al*: **Enrichment of the fetal fraction in non-invasive prenatal screening reduces maternal background interference**. *Sci Rep* 2018, **8**(1):17675.
15. Kwon HJ, Yun S, Joo J, Park D, Do WJ, Lee S, Lee MS: **Improving the accuracy of noninvasive prenatal testing through size-selection between fetal and maternal cfDNA**. *Prenat Diagn* 2023, **43**(13):1581-1592.
16. Hu P, Liang D, Chen Y, Lin Y, Qiao F, Li H, Wang T, Peng C, Luo D, Liu H *et al*: **An enrichment method to increase cell-free fetal DNA fraction and significantly reduce false negatives and test failures for non-invasive prenatal screening: a feasibility study**. *J Transl Med* 2019, **17**(1):124.
17. Liu S, Liu Y, Gu Y, Lin X, Zhu H, Liu H, Xu Z, Cheng S, Lan X, Li L *et al*: **Utilizing non-invasive prenatal test sequencing data for human genetic investigation**. *Cell Genom* 2024, **4**(10):100669.
18. Liu S, Huang S, Chen F, Zhao L, Yuan Y, Francis SS, Fang L, Li Z, Lin L, Liu R *et al*: **Genomic Analyses from Non-invasive Prenatal Testing Reveal Genetic Associations, Patterns of Viral Infections, and Chinese Population History**. *Cell* 2018, **175**(2):347-359 e314.
19. Zheng L, Yin N, Wang M, Huang H, Zhang S, Su L, Xu L: **Comparative performance and health economic analysis of prenatal screening for down syndrome in Fujian province, China**. *Sci Rep* 2025, **15**(1):23940.
20. Zhen J, Zhang L, Wang H, Chen X, Wang W, Li L, Zhang Q: **Clinical experience of genome-wide non-invasive prenatal testing as a first-tier screening test in a cohort of 59,771 pregnancies**. *PLoS One* 2025, **20**(8):e0329463.
21. Huang Q, Xu Q, Chen M, Fan W, Huang H: **Application of non-invasive prenatal testing for fetal chromosomal disorders in low-risk pregnancies: a follow-up study in central China**. *Front Genet* 2025, **16**:1574775.

22. Straver R, Oudejans CB, Sistermans EA, Reinders MJ: **Calculating the fetal fraction for noninvasive prenatal testing based on genome-wide nucleosome profiles**. *Prenat Diagn* 2016, **36**(7):614-621.
23. Chan KCA, Jiang P, Sun K, Cheng YK, Tong YK, Cheng SH, Wong AI, Hudecova I, Leung TY, Chiu RWK *et al*: **Second generation noninvasive fetal genome analysis reveals de novo mutations, single-base parental inheritance, and preferred DNA ends**. *Proc Natl Acad Sci U S A* 2016, **113**(50):E8159-8168.
24. Sun K, Jiang P, Wong AIC, Cheng YKY, Cheng SH, Zhang H, Chan KCA, Leung TY, Chiu RWK, Lo YMD: **Size-tagged preferred ends in maternal plasma DNA shed light on the production mechanism and show utility in noninvasive prenatal testing**. *Proc Natl Acad Sci U S A* 2018, **115**(22):E5106-E5114.
25. Sun K, Jiang P, Cheng SH, Cheng THT, Wong J, Wong VWS, Ng SSM, Ma BBY, Leung TY, Chan SL *et al*: **Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin**. *Genome Res* 2019, **29**(3):418-427.
26. Ju J, Zhao X, An Y, Yang M, Zhang Z, Liu X, Hu D, Wang W, Pan Y, Xia Z *et al*: **Cell-free DNA end characteristics enable accurate and sensitive cancer diagnosis**. *Cell Rep Methods* 2024, **4**(10):100877.
27. Liu X, Gong F, Lin H, An Y, Sun K: **Protocol for the analysis of cell-free DNA end characteristics for accurate cancer diagnosis**. *STAR Protoc* 2025, **6**(2):103757.
28. Coppede F: **Risk factors for Down syndrome**. *Arch Toxicol* 2016, **90**(12):2917-2929.
29. Liu X, Yang M, Hu D, An Y, Wang W, Lin H, Pan Y, Ju J, Sun K: **Systematic biases in reference-based plasma cell-free DNA fragmentomic profiling**. *Cell Rep Methods* 2024, **4**(6):100793.
30. Sun K: **Ktrim: an extra-fast and accurate adapter- and quality-trimmer for sequencing data**. *Bioinformatics* 2020, **36**(11):3561-3562.
31. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2**. *Nat Meth* 2012, **9**(4):357-359.
32. Gaffney DJ, McVicker G, Pai AA, Fondufe-Mittendorf YN, Lewellen N, Michelini K, Widom J, Gilad Y, Pritchard JK: **Controls of nucleosome positioning in the human genome**. *PLoS Genet* 2012, **8**(11):e1003036.
33. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J: **Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin**. *Cell* 2016, **164**(1-2):57-68.
34. Zhao Y, Wang J, Liang F, Liu Y, Wang Q, Zhang H, Jiang M, Zhang Z, Zhao W, Bao Y *et al*: **NucMap: a database of genome-wide nucleosome positioning map across species**. *Nucleic Acids Res* 2019, **47**(D1):D163-D169.
35. Chen K, Xi Y, Pan X, Li Z, Kaestner K, Tyler J, Dent S, He X, Li W: **DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing**. *Genome Res* 2013, **23**(2):341-351.
36. Chiu RWK, Chan KCA, 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.
37. Karlsson K, Sahlin E, Iwarsson E, Westgren M, Nordenskjold M, Linnarsson S: **Amplification-free sequencing of cell-free DNA for prenatal non-invasive diagnosis of chromosomal aberrations**. *Genomics* 2015, **105**(3):150-158.

38. Chandrananda D, Thorne NP, Ganesamoorthy D, Bruno DL, Benjamini Y, Speed TP, Slater HR, Bahlo M: **Investigating and correcting plasma DNA sequencing coverage bias to enhance aneuploidy discovery**. *PLoS One* 2014, **9**(1):e86993.
39. Hudecova I, Sahota D, Heung MM, Jin Y, Lee WS, Leung TY, Lo YM, Chiu RW: **Maternal plasma fetal DNA fractions in pregnancies with low and high risks for fetal chromosomal aneuploidies**. *PLoS One* 2014, **9**(2):e88484.
40. Balslev-Harder M, Richter SR, Kjaergaard S, Johansen P: **Correlation between Z score, fetal fraction, and sequencing reads in non-invasive prenatal testing**. *Prenat Diagn* 2017, **37**(9):943-945.
41. Costantini M, Clay O, Auletta F, Bernardi G: **An isochore map of human chromosomes**. *Genome Res* 2006, **16**(4):536-541.
42. Straver R, Sistermans EA, Holstege H, Visser A, Oudejans CB, Reinders MJ: **WISECONDOR: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme**. *Nucleic Acids Res* 2014, **42**(5):e31.