MEDICAL GENETICS
Original article

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2026;30(4):669-675

doi 10.18699/vjgb-26-67

Selection and evaluation of DNase I hypersensitive sites for prenatal screening of trisomy 21 in the fetus

A.M. Mazur [1], A.S. Starshin [1], N.V. Bogush [2], E.B. Prokhortchouk [1, 3]

[1] Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, Russia
[2] N.I. Pirogov Russian National Research Medical University, Moscow, Russia
[3] Sirius University of Science and Technology, Sirius Federal Territory, Krasnodar region, Russia
mazur.am@gmail.com

Abstract. This study introduces a novel approach for noninvasive prenatal testing (NIPT) of chromosomal abnormalities, based on analysis of epigenetic features in circulating cell-free DNA (cfDNA). The core innovation of our method leverages fundamental differences in chromatin organization between maternal and fetal cells. Specifically, we focused on genomic regions that exhibit open chromatin configuration in maternal blood cells but remain tightly packed in fetal tissues (DNase I hypersensitive sites or DHSs). These epigenetic differences create distinct cfDNA fragmentation signatures that allow selective identification of fetal DNA within the maternal cfDNA pool. The study workflow comprised several key steps: performing genome-wide screening to identify differentially accessible chromatin regions, selecting the most informative markers using a machine learning algorithm, and targeted sequencing of the selected epigenetic markers using molecular barcodes. Subsequently, a LASSO regression model was constructed and validated. As a proof of concept, the study demonstrates the method's efficacy in identifying trisomy 21 (Down syndrome), though the underlying principles can be readily adapted to other abnormalities. Complementing its robust performance, the technique offers practical advantages in terms of platform compatibility – the same epigenetic markers can be assessed using either next-generation sequencing or simpler, more cost-efficient methods like digital PCR. With further refinement, the approach could be extended to screen for additional aneuploidies (trisomies 13 and 18) and microdeletion syndromes. Therefore, this approach offers new opportunities for developing cost-effective testing systems suitable for widespread routine clinical implementation, combining high diagnostic accuracy with reduced analysis costs.

Key words: chromatin; cell free DNA; aneuploidy; trisomy; NIPT

For citation: Mazur A.M., Starshin A.S., Bogush N.V., Prokhortchouk E.B. Selection and evaluation of DNase I hypersensitive sites for prenatal screening of trisomy 21 in the fetus. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2026;30(4):669-675. doi 10.18699/vjgb-26-67

Funding. This work was supported by a grant from the state program of the "Sirius" Federal Territory "Scientific and Technological Development of the "Sirius" Federal Territory" (Agreement No. 21-03 dated September 27, 2024) (for the section "Whole-genome sequencing data analysis algorithm") and the Ministry of Science and Higher Education of the Russian Federation.

Отбор и анализ DNase I-гиперчувствительных сайтов для пренатального определения трисомии 21 у плода

А.М. Мазур [1], А.С. Старшин [1], Н.В. Богуш [2], Е.Б. Прохорчук [1, 3]

[1] Институт биоинженерии, Федеральный исследовательский центр «Фундаментальные основы биотехнологии» Российской академии наук, Москва, Россия
[2] Российский национальный исследовательский медицинский университет им. Н.И. Пирогова Министерства здравоохранения Российской Федерации (Пироговский Университет), Москва, Россия
[3] Научно-технологический университет «Сириус», федеральная территория «Сириус», Краснодарский край, Россия
mazur.am@gmail.com

Аннотация. В данной работе представлен новый подход к неинвазивной пренатальной диагностике хромосомных аномалий, основанный на анализе эпигенетических особенностей циркулирующей внеклеточной ДНК (вкДНК). Ключевая идея метода заключается в использовании фундаментальных различий в организации хроматина между материнскими и фетальными клетками. В частности, мы сосредоточились на участках генома, которые в материнских клетках крови находятся в открытом состоянии (DHS), но остаются закрытыми в тканях плода. Эти эпигенетические различия создают уникальные паттерны вкДНК, позволяющие выделить фетальную ДНК из общего пула материнской вкДНК. Работа включает в себя несколько основных шагов: проведение полногеномного скрининга для выявления

дифференциально доступных участков хроматина, отбор наиболее информативных маркеров при помощи алгоритма машинного обучения, таргетное секвенирование отобранных эпигенетических маркеров с применением молекулярных баркодов и построение модели на основе LASSO-регрессии с последующей ее валидацией. Особое внимание уделено диагностике трисомии хромосомы 21 (синдрома Дауна), как наиболее клинически значимого хромосомного нарушения. Разработанный алгоритм демонстрирует высокую точность при анализе отобранных эпигенетических маркеров. Важное преимущество метода – его технологическая гибкость: отобранные маркеры можно анализировать как с помощью NGS, так и более доступными методами, включая цифровую ПЦР. Перспективы дальнейшего развития метода включают расширение диагностической панели для выявления других частых анеуплоидий (трисомии хромосом 13, 18), а также микроделеционных синдромов. Предлагаемый подход открывает новые возможности для создания экономически эффективных тест-систем, которые могут быть внедрены в широкую клиническую практику, обеспечивая высокую точность диагностики при снижении стоимости анализа.
Ключевые слова: хроматин; внеклеточная ДНК; анеуплоидия; трисомия; НИПТ

Introduction

The Non-Invasive Prenatal Test (NIPT), which employs next-generation sequencing (NGS) technology, represents a revolutionary approach to detecting fetal chromosomal abnormalities using only a sample of maternal blood (Fan et al., 2008). Prior to the development of NGS, this type of analysis was not feasible, as there was no reliable method to isolate the small amount of fetal DNA present amidst the vast excess of maternal DNA in the bloodstream (Dhallan et al., 2007). Due to the high throughput and sensitivity of NGS, it is now possible to accurately detect and quantify these trace amounts of fetal cell-free DNA (cfDNA) directly from maternal plasma (Alberry et al., 2021). Consequently, NIPT has emerged as a rapid, safe, and reliable screening tool for assessing the risk of common aneuploidies, including Down, Edwards, and Patau syndromes (Gil et al., 2017). The genomics revolution driven by NGS has therefore paved the way for the widespread clinical implementation of NIPT (Satam et al., 2023).

However, despite the impressive capabilities of NGS, the effectiveness of NIPT still depends on selecting optimal biological markers to distinguish between maternal and fetal DNA (Hui, Bianchi, 2020). In this context, epigenetic differences – particularly chromatin accessibility profiles – are of special interest.

DNase I-hypersensitive sites (DHSs) are genomic regions with an open chromatin structure, exhibiting increased accessibility to the DNase I enzyme. These sites are frequently associated with gene regulatory elements and display cell-specific expression profiles (Thurman et al., 2012).

In the previous stage of this study, the use of DHS clusters significantly restricted the number of available candidate regions, thereby reducing the sensitivity and specificity of the test system. To overcome this limitation, we transitioned from cluster analysis to the investigation of individual DHS. This approach substantially expanded the pool of candidate regions, enabling more precise filtering and ultimately yielding a sufficient number of high-quality loci for reliable diagnostics.

Our study focuses on developing an enhanced NIPT approach based on the analysis of individual DHSs – genomic regions with chromatin structures characteristic of the fetus and mother. The proposed method is expected to ensure high accuracy in fetal DNA detection by leveraging the epigenetic features of DHSs, while also reducing assay costs through their targeted analysis.

Material and methods

Study cohort. A total of 1,012 blood samples from pregnant women were collected between April 2014 and April 2015. The patients were referred from 47 medical institutions across the Russian Federation or were self-referred. The mean age of participants was 35 years, with the youngest being 20 and the oldest 51 years. Gestational age at the time of blood collection ranged from 10 to 24 weeks, with a mean of 14 weeks. The mean fetal cell-free DNA (cfDNA) concentration was 11 % (Pantiuhk et al., 2016). All women provided informed consent for the use of their samples for scientific purposes. The study was approved by the Committee of the Academic Council of the Faculty of Biomedical Sciences, Pirogov Russian National Research Medical University, Ministry of Health of the Russian Federation (protocol No. 4 dated December 26, 2022).

Primer design. Primers were designed using the Primer3 software (Untergasser et al., 2012) with the following parameters: PRIMER_OPT_SIZE = 21, PRIMER_MIN_SIZE = 18, PRIMER_MAX_SIZE = 24, PRIMER_MAX_NS_ACCEPTED = 1, and PRIMER_PRODUCT_SIZE_RANGE = 100–140 bp.

Target enrichment and library preparation. Targeted sequencing was performed using single-end enrichment with duplex unique molecular identifiers (UMIs), following the protocol described by Q. Peng et al. (2019).

Library sequencing. Sequencing was carried out using Illumina SBS technology with single-end reads of at least 80 bp.

Model development and evaluation. In the first step, the input data were prepared by dividing samples into two categories: those with trisomy and those without it. Samples with a total molecule count below 15,000 were filtered out.

Subsequently, 100 independent trials were conducted using the R packages rsample and caret (Kuhn, 2008). In each trial, the data were randomly split into training (60 % of samples),

validation (20 %), and test (20 %) sets, while preserving the original class ratio.

To compensate for class imbalance in the training sets, observation weighting was applied. Class weights were calculated separately for each training set according to the formulas:

w_t = N_major / N_minor – for samples of the minority class (trisomy),

w_n = 1 – for samples of the majority class (normal),

where N_major and N_minor are the number of samples in the majority and minority classes, respectively, within the training set. This weighting effectively increases the contribution of rare trisomy cases to the loss function, preventing the model from biasing toward the more common class.

The optimal classification threshold was determined on the validation set using Youden's index (Youden, 1950), which maximizes the J statistic (J = sensitivity + specificity – 1). This criterion aims to find a threshold that balances the method's ability to correctly identify both positive cases (sensitivity) and negative cases (specificity). This approach ensures that the selected threshold balances the two types of errors: false negatives (where aneuploidy is present but the test fails to detect it) and false positives (where the test incorrectly indicates aneuploidy). Three key performance metrics were calculated on the test data: the area under the ROC curve (AUC), sensitivity, and specificity. Model training was performed using the glmnet package (Friedman et al., 2010), and all results were visualized using ggplot2 (Wickham, 2016) and ggpubr.

Results

Identification of candidate regions in the human genome

At the first stage, DNase I-hypersensitive sites (DHSs) were selected from an open-access database (Sheffield et al., 2013). For maternal tissue representation, all hematopoietic and endothelial cell lines were used, as these are typical sources of freely circulating maternal cfDNA (Lui et al., 2002).

The selection criterion was based on differential accessibility of these sites: an open chromatin state in cell lines corresponding to maternal tissues (31 lines of hematopoietic and endothelial origin) and a closed state in fetal cell lines (Chorion). This approach enabled the identification of 13,637 individual DHSs (Fig. S1)[1].

The coordinates of all selected sites were converted to the human genome version GRCh38 (Table S1). For subsequent analysis, a 1,000 bp region was extracted around each 150 bp DHS, with the DHS positioned at its center. Although the DHSs themselves did not overlap, the flanking regions could partially overlap when the distance between central sites was less than 500 bp – a factor that was taken into account during final selection.

[1] Supplementary Figures S1–S3 and Tables S1–S3 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Mazur_Engl_30_4.xlsx

To construct the diagnostic panel, we selected a set of regions containing DHSs on: chromosome 21 – as the direct target for Down syndrome diagnosis; chromosome Y – as a negative control; and chromosomes 1, 2, 3, 4, 5, 6, 8, 10, and 12 – as markers, the relative coverage decrease of which in trisomy cases reflects the redistribution of reads in favor of chromosome 21.

Sample set formation

The study utilized cfDNA sequencing data from a previous work focused on non-invasive prenatal diagnostic methods for aneuploidies (Pantiukh et al., 2016).

To determine fetal cfDNA concentration, two complementary methods were applied: quantitative analysis of the Y chromosome and the SeqFF method, which is based on regression analysis of overrepresented fetal sequences (Kim et al., 2015; van Beek et al., 2017).

The need to group samples arose from the characteristics of the original data, where the mean coverage was only 0.3×. At this coverage level, a standard 150 bp DHS received on average only one read, which made reliable quantitative analysis at the individual level impossible. To address this low coverage issue, samples were sorted by increasing fetal DNA concentration and evenly distributed into four groups (P1–P4) (Fig. S2).

The resulting clinical sample set included data from 958 cases with normal fetal karyotype and 54 samples with trisomy 21 (Down syndrome).

Whole-genome sequencing data analysis algorithm

All selected cfDNA samples were aligned to the human reference genome GRCh38 using bowtie2 (Langmead, Salzberg, 2012) in global alignment mode (--end-to-end) with the --sensitive parameter. The mean percentage of successfully mapped reads was 97.05 %, with unmapped reads excluded from further analysis. For each sample, coverage was calculated for every nucleotide within the pre-selected regions.

Aggregated data were divided into groups: non-pregnant women, pregnant women with normal fetal karyotype (further subdivided by fetal DNA concentration), and pregnant women with fetal trisomies. A statistically significant increase in coverage within DHSs was observed between groups of pregnant women with a normal fetal karyotype (Fig. 1*a*) and between pregnant women compared to non-pregnant women – the latter explained by the closed chromatin state of fetal DNA in these genomic regions, rendering it inaccessible to DNase I (Fig. 1*b*). Differences were also noted for pregnant women carrying fetuses with trisomy 21 (Fig. 1*b*). This increase in DHS coverage confirmed that the proposed method enables the detection of fetal aneuploidies.

Interestingly, multiple regions with increased coverage were identified in pregnant women. Verification using the UCSC Genome Browser confirmed the presence of additional DHSs

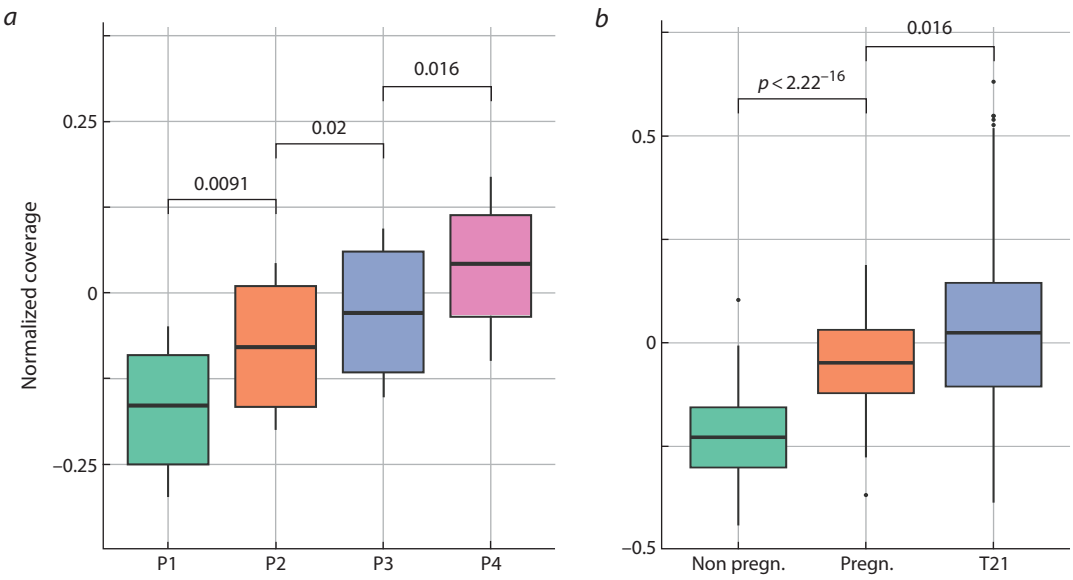

Fig. 1. Normalized coverage (Z-score) in target regions using chromosome 21 as an example: *a* – in subgroups of pregnant women with normal fetal karyotype; *b* – in groups of non-pregnant women, pregnant women with normal fetal karyotype, and pregnant women with fetal trisomy.

Statistical significance of differences between the groups was assessed using the Mann–Whitney U-test (Wilcoxon test for independent samples).

within such regions (Fig. S3), which is consistent with known data on the clustering of regulatory elements in the genome (Thurman et al., 2012).

Establishing criteria and selecting regions

The primary objective was to identify regions most effective at distinguishing between pregnant and non-pregnant women. To this end, the following criteria were proposed: the length of the maximum window where the difference between mean coverage in non-pregnant women and mean coverage in pregnant women is less than zero; the maximum area under the difference curve within a 100-nucleotide window; the relationship between coverage level and fetal cfDNA concentration; and functional significance. Functional significance was assessed based on criteria proposed in R.E. Thurman et al. (2012). Promoter sites were defined as those the coordinates of which intersected with transcription start sites from the GENCODE human genome reference annotation (Harrow et al., 2012), or as the DHS nearest to a transcription start site in the 5′ direction. Enhancers were defined as those located within 500,000 nucleotides of a promoter site and showing high correlation ($R \geq 0.7$) with that promoter site based on DNase-seq data.

To determine which of the proposed criteria was most suitable for region selection, a binary classifier based on random forest was constructed to distinguish samples from pregnant women versus non-pregnant women, using the mean coverage in the 20 top-ranking regions according to each criterion. We used the Random Forest Classifier algorithm from the scikit-learn library in its default configuration (Pedregosa et al., 2011). Prediction quality was assessed using the F-measure for non-pregnant sample predictions, as the number of such samples was smaller than that of pregnant samples. To identify the most effective region selection rule, we evaluated the predictive performance of each criterion individually and in various combinations. The best F-measure was achieved using regions selected by a combination of three criteria: the area under the difference curve criterion, the presence of a relationship between coverage level and fetal DNA concentration within the region, and the criterion of whether the site was a promoter or enhancer.

Next, for each nucleotide coordinate, the mean normalized coverage (Z-score) was calculated across all classes: for non-pregnant samples; for pregnant samples (by groups with different cfDNA concentrations); and for samples from pregnant women carrying fetuses with trisomy 21. Subsequently, all regions on a given chromosome were sorted in descending order according to the value determined by the selected criterion, and a specified number of regions were selected.

Sequencing of the selected regions

For the analysis of the 40 selected genomic regions (Table S2) across 283 samples, targeted sequencing based on multiplex PCR was employed. Specific primers were designed for each region (Table S3), enabling simultaneous amplification of all target sites in a single reaction. Prior to sequencing, molecular barcoding of the original DNA molecules was performed during the amplification step, ensuring accurate counting of

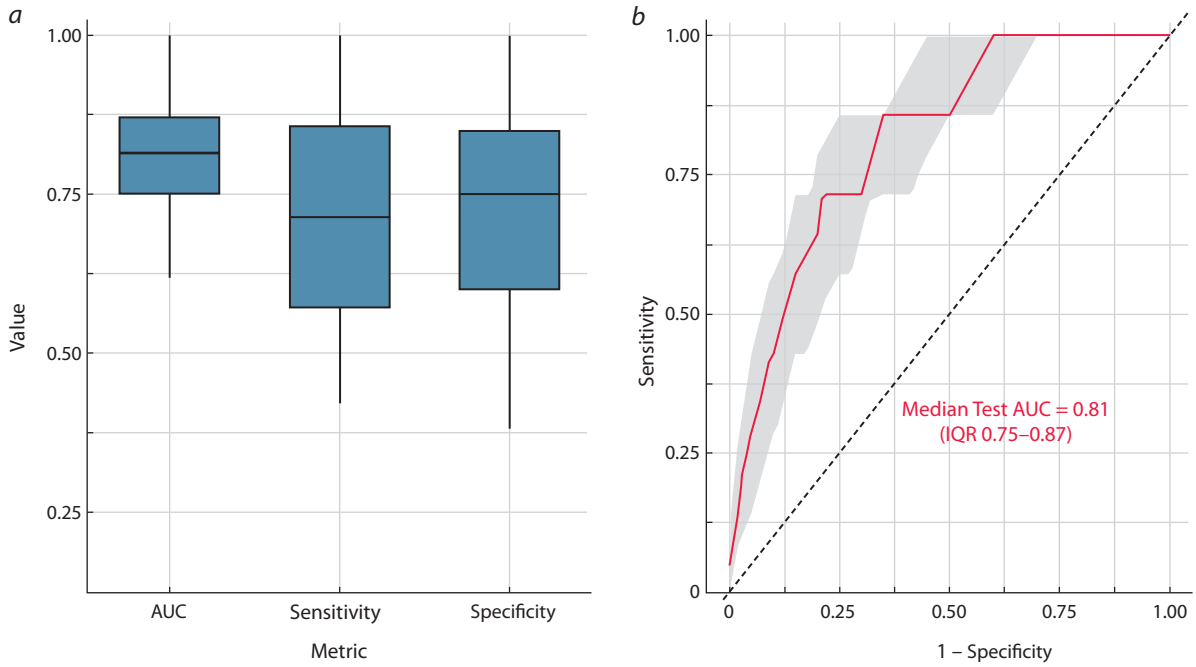

Fig. 2. Model performance: *a* – distribution of metrics (AUC, sensitivity, and specificity); *b* – ROC curve based on 100 independent splits.

initial molecule numbers and minimizing the impact of PCR artifacts. Read mapping was conducted in the same manner as described in the "Whole-genome sequencing data analysis algorithm" section.

Model development and validation

For 201 samples with a normal fetal karyotype and 82 samples with trisomy, counts of cfDNA molecules corresponding to the 40 selected regions were obtained. To reliably assess the suitability of these regions for use in a test system, we trained and tested a LASSO regression model using stratified repeated random subsampling with class weighting (see Material and methods).

The high sensitivity and specificity values (Fig. 2*a*), along with the AUC (Fig. 2*b*) achieved by the model, allow us to conclude that it is possible to predict the presence of fetal trisomy based on sequencing data from DHS regions of circulating cfDNA.

Discussion

Before the analysis of plasma cfDNA became the dominant approach in non-invasive prenatal testing, the scientific community was actively searching for alternative methods to obtain fetal genetic material without invasive procedures (Simpson, Elias, 1993).

Two main cell types served as primary targets in these studies: trophoblasts (Bruch et al., 1991) and nucleated red blood cells (fetal erythroblasts) (Simpson, Elias, 1994). Trophoblasts, being placental derivatives, seemed a logical target as they contain the complete fetal genome and could theoretically serve as ideal material for analysis (Covone et al., 1984). Concurrently, research was conducted on isolating fetal erythroblasts (Bianchi et al., 1990). The advantage of this approach lay in the fact that immature nucleated red blood cells are normally virtually absent from the peripheral blood of healthy adults, theoretically allowing them to be distinguished against the background of maternal cells.

However, practical application of the cellular approach encountered serious limitations (Holzgreve et al., 1992). The main problem was the extremely low concentration of fetal cells in maternal blood, making their isolation technically challenging and unreliable.

Against this backdrop, the discovery made by Dennis Lo and colleagues in 1997 – demonstrating the presence of significant amounts of fetal cfDNA in maternal plasma – marked the beginning of a fundamentally different path (Lo et al., 1997). Despite the fact that cfDNA represented fragmented placental DNA rather than the complete fetal genome, this approach had decisive practical advantages: a relatively high concentration of genetic material and technological accessibility for analysis using PCR (Lo et al., 1998).

Current commercial NIPTs utilizing high-throughput sequencing methods demonstrate high accuracy, but their widespread adoption is still constrained by relatively high costs (Gil et al., 2017).

Our method, based on targeted analysis of 40 epigenetic markers (DHSs), offers a cost-effective solution. Such approaches require sequencing less than 1 % of the genome, substantially reducing reagent and data processing costs (Schmitt et al., 2015). It is worth noting that not only NGS platforms but

also the more accessible digital PCR (dPCR) technology can be used for this analysis, further lowering costs and simplifying implementation into routine clinical practice (Li et al., 2022). This is particularly relevant for smaller laboratories without access to expensive sequencing systems. At the same time, all key advantages of NIPT are preserved: non-invasiveness, the possibility of early diagnosis, and high accuracy in detecting chromosomal abnormalities (Zhang et al., 2026).

An important aspect of this study was the development of rigorous criteria for selecting informative regions. Our method incorporates a comprehensive set of parameters, including not only statistical measures of differential accessibility but also the functional significance of sites and the dependence of the signal on fetal DNA concentration. Such an integrated approach enables the creation of a robust diagnostic panel resilient to technical noise. Furthermore, the application of molecular barcoding addresses a crucial challenge in quantitative cfDNA analysis – accurate counting of original molecules in the presence of PCR amplification.

Conclusion

The prospects for practical application of the developed approach appear promising. First, the method fundamentally allows for expanding the range of detectable chromosomal abnormalities without substantially increasing analysis costs. Second, its adaptability to various technological platforms opens up possibilities for creating flexible diagnostic systems that can be optimized for specific clinical tasks and budget constraints.

Future research may focus on optimizing the marker panel for simultaneous detection of a wide range of chromosomal abnormalities and standardizing protocols for use in routine clinical practice. The proposed method has every reason to become the foundation for a new generation of screening tests that combine high reliability with cost-effectiveness.

References

Alberry M.S., Aziz E., Ahmed S.R., Abdel-Fattah S. Non invasive prenatal testing (NIPT) for common aneuploidies and beyond. *Eur J Obstet Gynecol Reprod Biol.* 2021;258:424-429. doi 10.1016/j.ejogrb.2021.01.008

Bianchi D.W., Flint A.F., Pizzimenti M.F., Knoll J.H., Latt S.A. Isolation of fetal DNA from nucleated erythrocytes in maternal blood. *Proc Natl Acad Sci USA.* 1990;87(9):3279-3283. doi 10.1073/pnas.87.9.3279

Bruch J.F., Metezeau P., Garcia-Fonknechten N., Richard Y., Tricottet V., Hsi B.L., Kitzis A., Julien C., Papiernik E. Trophoblast-like cells sorted from peripheral maternal blood using flow cytometry: a multiparametric study involving transmission electron microscopy and fetal DNA amplification. *Prenat Diagn.* 1991;11(10):787-798. doi 10.1002/pd.1970111007

Covone A.E., Mutton D., Johnson P.M., Adinolfi M. Trophoblast cells in peripheral blood from pregnant women. *Lancet.* 1984;2(8407):841-843. doi 10.1016/s0140-6736(84)90875-4

Dhallan R., Guo X., Emche S., Damewood M., Bayliss P., Cronin M., Barry J., Betz J., Franz K., Gold K., Vallecillo B., Varney J. A non-invasive test for prenatal diagnosis based on fetal DNA present in maternal blood: a preliminary study. *Lancet.* 2007;369(9560):474-481. doi 10.1016/S0140-6736(07)60115-9

Fan H.C., Blumenfeld Y.J., Chitkara U., Hudgins L., Quake S.R. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc Natl Acad Sci USA.* 2008;105(42):16266-16271. doi 10.1073/pnas.0808319105

Friedman J., Hastie T., Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Statist Software*. 2010;33(1):1-22. doi 10.18637/jss.v033.i01

Gil M.M., Accurti V., Santacruz B., Plana M.N., Nicolaides K.H. Analysis of cell-free DNA in maternal blood in screening for aneuploidies: updated meta-analysis. *Ultrasound Obstet Gynecol.* 2017;50(3):302-314. doi 10.1002/uog.17484

Harrow J., Frankish A., Gonzalez J.M., Tapanari E., Diekhans M., Kokocinski F., Aken B.L., … Valencia A., Reymond A., Gerstein M., Guigó R., Hubbard T.J. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* 2012;22(9):1760-1774. doi 10.1101/gr.135350.111

Holzgreve W., Garritsen H.S., Ganshirt-Ahlert D. Fetal cells in the maternal circulation. *J Reprod Med.* 1992;37(5):410-418

Hui L., Bianchi D.W. Fetal fraction and noninvasive prenatal testing: what clinicians need to know. *Prenat Diagn.* 2020;40(2):155-163. doi 10.1002/pd.5620

Kim S.K., Hannum G., Geis J., Tynan J., Hogg G., Zhao C., Jensen T.J., Mazloom A.R., Oeth P., Ehrich M., van den Boom D., Deciu C. Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts. *Prenat Diagn.* 2015;35(8):810-815. doi 10.1002/pd.4615

Kuhn M. Building predictive models in R using the caret package. *J Stat Softw.* 2008;28(5):1-26. doi 10.18637/jss.v028.i05

Langmead B., Salzberg S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357-359. doi 10.1038/nmeth.1923

Li Y.Q., Tan G.J., Zhou Y.Q. Digital PCR and its applications in noninvasive prenatal testing. *Brief Funct Genomics.* 2022;21(5):376-386. doi 10.1093/bfgp/elac024

Lo Y.M., Corbetta N., Chamberlain P.F., Rai V., Sargent I.L., Redman C.W., Wainscoat J.S. Presence of fetal DNA in maternal plasma and serum. *Lancet.* 1997;350(9076):485-487. doi 10.1016/S0140-6736(97)02174-0

Lo Y.M., Tein M.S., Lau T.K., Haines C.J., Leung T.N., Poon P.M., Wainscoat J.S., Johnson P.J., Chang A.M., Hjelm N.M. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *Am J Hum Genet.* 1998;62(4):768-775. doi 10.1086/301800

Lui Y.Y., Chik K.W., Chiu R.W., Ho C.Y., Lam C.W., Lo Y.M. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem.* 2002;48(3):421-427. doi 10.1093/clinchem/48.3.421

Pantiukh K.S., Chekanov N.N., Zaigrin I.V., Zotov A.M., Mazur A.M., Prokhortchouk E.B. Report on noninvasive prenatal testing: classical and alternative approaches. *F1000Res.* 2016;5:722. doi 10.12688/f1000research.8243.1

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., … Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12(85):2825-2830

Peng Q., Xu C., Kim D., Lewis M., DiCarlo J., Wang Y. Targeted single primer enrichment sequencing with single end duplex-UMI. *Sci Rep*. 2019;9(1):4810. doi 10.1038/s41598-019-41215-z

Satam H., Joshi K., Mangrolia U., Waghoo S., Zaidi G., Rawool S., Thakare R.P., Banday S., Mishra A.K., Das G., Malonia S.K. Next-

generation sequencing technology: current trends and advancements. *Biology.* 2023;12(7):997. doi 10.3390/biology12070997

Schmitt M.W., Fox E.J., Prindle M.J., Reid-Bayliss K.S., True L.D., Radich J.P., Loeb L.A. Sequencing small genomic targets with high efficiency and extreme accuracy. *Nat Methods.* 2015;12(5):423-425. doi 10.1038/nmeth.3351

Sheffield N.C., Thurman R.E., Song L., Safi A., Stamatoyannopoulos J.A., Lenhard B., Crawford G.E., Furey T.S. Patterns of regulatory activity across diverse human cell types predict tissue identity, transcription factor binding, and long-range interactions. *Genome Res.* 2013;23(5):777-788. doi 10.1101/gr.152140.112

Simpson J.L., Elias S. Isolating fetal cells from maternal blood. *JAMA.* 1993;270(19):2357-2361

Simpson J.L., Elias S. Isolating fetal cells in maternal circulation for prenatal diagnosis. *Prenat Diagn.* 1994;14(13):1229-1242. doi 10.1002/pd.1970141308

Thurman R.E., Rynes E., Humbert R., Vierstra J., Maurano M.T., Haugen E., Sheffield N.C., … Kaul R., Furey T.S., Dekker J., Crawford G.E., Stamatoyannopoulos J.A. The accessible chromatin landscape of the human genome. *Nature.* 2012;489(7414):75-82. doi 10.1038/nature11232

Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B.C., Remm M., Rozen S.G. Primer3 – new capabilities and interfaces. *Nucleic Acids Res*. 2012;40(15):e115. doi 10.1093/nar/gks596

van Beek D.M., Straver R., Weiss M.M., Boon E.M.J., Huijsdens-van Amsterdam K., Oudejans C.B.M., Reinders M.J.T., Sistermans E.A. Comparing methods for fetal fraction determination and quality control of NIPT samples. *Prenat Diagn.* 2017;37(8):769-773. doi 10.1002/pd.5079

Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag, 2016

Youden W.J. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32-35. doi 10.1002/1097-0142(1950)3:1%3C32::aid-cncr2820030106%3E3.0.co;2-3

Zhang L., Chen L., Yu Y., Wu Z., Zhao M., Wang P., Liu Z., Dai J., Li Y., Wang M. Digital PCR in prenatal diagnosis. *Clin Chim Acta.* 2026;578:120535. doi 10.1016/j.cca.2025.120535

Conflict of interest. Mazur A.M. is the science director at Genoanalytica LLC, a company involved in NIPT development. However, this study was not funded by Genoanalytica LLC and is not related to the company's commercial interests.

Received December 18, 2025. Revised February 6, 2026. Accepted February 9, 2026.