

doi 10.18699/vjgb-26-67


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

А.М. Мазур , А.С. Старшин , Н.В. Богущ², 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

Для цитирования: Мазур А.М., Старшин А.С., Богущ Н.В., Прохорчук Е.Б. Отбор и анализ DNase I-гиперчувствительных сайтов для пренатального определения трисомии 21 у плода. *Вавиловский журнал генетики и селекции*. 2026;30(4): 669-675. doi 10.18699/vjgb-26-67

Финансирование. Работа выполнена при поддержке гранта государственной программы Федерального округа «Сириус» «Научно-технологическое развитие Федерального округа «Сириус» (соглашение № 21-03 от 27.09.2024) (раздел Алгоритм анализа данных полногеномного секвенирования) и Министерства науки и высшего образования Российской Федерации.

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

А.М. Mazur , А.С. Starshin , N.V. Bogush², 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 Selekcii = Vavilov J Genet Breed.* 2026;30(4): 669-675. doi 10.18699/vjgb-26-67

Введение

Неинвазивный пренатальный тест (НИПТ), использующий технологию секвенирования нового поколения (NGS), считается революционным методом диагностики хромосомных нарушений у плода, основанным исключительно на анализе материнской крови (Fan et al., 2008). До возникновения NGS проведение подобного анализа было невозможно ввиду отсутствия эффективных способов выявления крошечных фрагментов фетальной ДНК среди огромного количества собственной ДНК матери (Dhallan et al., 2007). Благодаря высокой производительности и чувствительности NGS стало возможным выделение и точное определение небольших количеств внеклеточной ДНК (вкДНК) плода прямо из кровотока матери (Albergy et al., 2021). Это сделало НИПТ быстрым, безопасным и удобным способом оценки рисков анеуплоидий, таких как синдром Дауна, Эдвардса и Патау (Gil et al., 2017). Таким образом, именно геномная революция, связанная с появлением методов NGS, открыла путь к широкому применению НИПТ в современной медицинской практике (Satam et al., 2023).

Однако, несмотря на впечатляющие возможности NGS, эффективность НИПТ по-прежнему зависит от выбора оптимальных биологических маркеров, позволяющих различать материнскую и фетальную ДНК (Hui, Bianchi, 2020). В этом контексте особый интерес представляют эпигенетические различия, в частности профили доступности хроматина.

DNase I-гиперчувствительные сайты (DHS) представляют собой геномные участки с открытой структурой хроматина, обладающие повышенной доступностью для фермента DNase I. Такие участки часто ассоциируются с регуляторными элементами генов и демонстрируют клеточно-специфичные профили экспрессии (Thurman et al., 2012).

На предыдущем этапе данного исследования использование кластеров DHS существенно ограничивало количество доступных кандидатных регионов, снижая чувствительность и специфичность тест-системы. Для преодоления этого ограничения мы перешли от анализа кластеров к изучению индивидуальных DHS. Такой подход значительно расширил набор кандидатных регионов, позволив проводить более точную фильтрацию и в итоге получить достаточное количество высококачественных локусов для надежной диагностики.

Наше исследование сосредоточено на разработке усовершенствованного подхода к НИПТ, основанного на анализе отдельных DHS – участков генома с характерной

для плода и матери структурой хроматина. Предлагаемый подход должен обеспечить высокую точность детекции фетальной ДНК за счет учета эпигенетических особенностей DHS и снижение стоимости анализа благодаря их целенаправленному анализу.

Материал и методы

Когорта исследования. 1012 образцов крови беременных женщин были собраны в период с апреля 2014 г. по апрель 2015 г. Пациентки были направлены из 47 медицинских учреждений Российской Федерации или обратились самостоятельно. Средний возраст участниц составил 35 лет; самой младшей было 20 лет, самой старшей – 51 год. Срок беременности на момент забора крови составлял от 10 до 24 недель, в среднем 14 недель. Средняя концентрация внеклеточной ДНК плода (вкДНК плода) составила 11 % (Pantiuhk et al., 2016). От всех женщин было получено информированное согласие на использование их образцов в научных целях. Исследование одобрено комиссией ученого совета медико-биологического факультета ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России (протокол № 4 от 26 декабря 2022 г.).

Подбор праймеров. Для подбора праймеров использовали программу primer3 (Untergasser et al., 2012) со следующими параметрами: PRIMER_OPT_SIZE = 21, PRIMER_MIN_SIZE = 18, PRIMER_MAX_SIZE = 24, PRIMER_MAX_NS_ACCEPTED = 1, PRIMER_PRODUCT_SIZE_RANGE = 100–140.

Обогащение по отобраным районам и приготовление библиотек. Таргетное секвенирование методом обогащения одноцепочечными праймерами с дуплексными уникальными молекулярными идентификаторами (single-end duplex UMI) проводили в соответствии с протоколом (Peng et al., 2019).

Секвенирование библиотек выполняли по технологии SBS Illumina одноконцевыми чтениями длиной не менее 80 оснований.

Построение и оценка модели. На первом этапе подготовили исходные данные, разделив образцы на две категории: с трисомией и без нее. Образцы с общим числом молекул <15000 отфильтровали.

Затем было проведено 100 независимых испытаний с использованием пакетов R `rsample` и `caret` (Kuhn, 2008). В каждом испытании случайным образом, но с сохранением исходного соотношения классов (стратификацией) данные разделялись на обучающую (60 % образцов), валидационную (20 % образцов) и тестовую (20 % образцов) части.

Для компенсации дисбаланса классов в обучающих наборах применялось взвешивание наблюдений. Веса классов рассчитывали отдельно для каждого обучающего набора по формулам:

$w_t = N_major / N_minor$ – для образцов минорного класса (трисомия),

$w_n = 1$ – для образцов мажорного класса (норма),

где N_major и N_minor – количество образцов мажорного и минорного классов соответственно в обучающей выборке. Такое взвешивание эффективно увеличивает вклад редких случаев трисомии в функцию потерь, предотвращая смещение модели в сторону более распространенного класса.

На валидационной выборке определялся оптимальный порог классификации по критерию Йодена (Youden, 1950), который максимизирует индекс Йодена ($J = \text{чувствительность} + \text{специфичность} - 1$). Данный критерий направлен на нахождение порога, обеспечивающего баланс между способностью метода корректно идентифицировать как положительные случаи (чувствительность), так и отрицательные (специфичность). Такой подход гарантирует, что выбранный порог будет балансировать между двумя типами ошибок: ложноотрицательными результатами (когда анеуплоидия есть, но тест ее не показывает) и ложноположительными (когда тест ошибочно указывает на наличие анеуплоидии). На тестовых данных вычисляли три ключевых показателя: площадь под ROC-кривой (AUC), чувствительность и специфичность модели. Для обучения модели использовали пакет *glmnet* (Friedman et al., 2010). Все результаты визуализировали с помощью *ggplot2* (Wickham, 2016) и *ggpubr*.

Результаты

Определение кандидатных регионов в геноме человека

На первом этапе были отобраны DNase I-гиперчувствительные сайты (DHS) из открытой базы данных (Sheffield et al., 2013). В качестве типа ткани матери были использованы все клеточные линии гематопоэтического и эндотелиального ряда, как типичные источники свободно циркулирующей вкДНК матери (Lui et al., 2002).

Критерием отбора служила дифференциальная доступность этих сайтов: открытое состояние в клеточных линиях, соответствующих материнским тканям (31 линия гематопоэтического и эндотелиального происхождения), и закрытое состояние в фетальных клеточных линиях (Chogion). Такой подход позволил идентифицировать 13637 индивидуальных DHS (рис. S1)¹.

Координаты всех отображенных сайтов были переведены в версию генома человека GRCh38 (табл. S1). Для последующего анализа вокруг каждого DHS длиной 150 п. н. выделяли регион длиной 1000 п. н., располагая сайт в его центре. Хотя сами DHS не перекрывались, фланкирующие регионы могли частично пересекаться при расстоянии между центральными сайтами менее 500 п. н., что было учтено при финальном отборе.

¹ Рис. S1–S3 и табл. S1–S3 Приложения см. по адресу: <https://vavilovj-icg.ru/download/pict-2026-30/appx37.zip>

Для формирования диагностической панели мы взяли набор регионов, включающий в себя DHS: на хромосоме 21 – как прямую мишень для диагностики синдрома Дауна, на Y-хромосоме – в качестве отрицательного контроля, и на хромосомах 1, 2, 3, 4, 5, 6, 8, 10, 12 – в качестве маркеров, относительное снижение покрытия которых при трисомии отражает перераспределение ридов в пользу 21-й хромосомы.

Формирование набора образцов

В исследовании были использованы данные секвенирования вкДНК из предыдущей работы (Pantiukh et al., 2016), посвященной методам неинвазивной пренатальной диагностики анеуплоидий.

Для определения концентрации фетальной вкДНК применяли два взаимодополняющих метода: количественный анализ Y-хромосомы и метод SeqFF, основанный на регрессионном анализе перепредставленных фетальных последовательностей (Kim et al., 2015; van Beek et al., 2017).

Необходимость группировки образцов обусловлена особенностями исходных данных, где среднее покрытие составляло всего 0.3×. При таком уровне покрытия на стандартный DHS длиной 150 нуклеотидов приходилось в среднем лишь одно чтение, что делало невозможным надежный количественный анализ на индивидуальном уровне. Для решения проблемы низкого покрытия образцы были отсортированы по возрастанию концентрации фетальной ДНК и равномерно распределены на четыре группы (P1–P4) (рис. S2).

Сформированный набор клинических образцов включал данные 958 случаев с нормальным кариотипом плода и 54 образца с трисомией по хромосоме 21 (синдром Дауна).

Алгоритм анализа данных полногеномного секвенирования

Все выбранные образцы вкДНК были картированы на референсный геном человека GRCh38 с использованием программы *bowtie2* (Langmead, Salzberg, 2012) в режиме глобального выравнивания (--end-to-end) с параметром --sensitive. Доля успешно картированных чтений составила 97.05 %, при этом некартированные чтения были исключены из дальнейшего анализа. Для каждого образца рассчитали покрытие каждого нуклеотида в предварительно отображенных регионах.

Агрегированные данные разделили на группы: небеременные женщины, беременные с нормальным кариотипом плода (с дальнейшим разделением по концентрации фетальной ДНК) и беременные с трисомиями плода. В DHS наблюдалось статистически значимое увеличение покрытия между группами беременных женщин с нормальным кариотипом плода (рис. 1, а) и у беременных женщин по сравнению с небеременными (это объясняется тем, что у фетальной ДНК хроматин на данном участке генома закрыт и не подвергается воздействию DNКазазы I) (см. рис. 1, б). Также различия отмечались для беременных с трисомией 21-й хромосомы у плода (см. рис. 1, б). Увеличение покрытия в DHS позволило установить, что

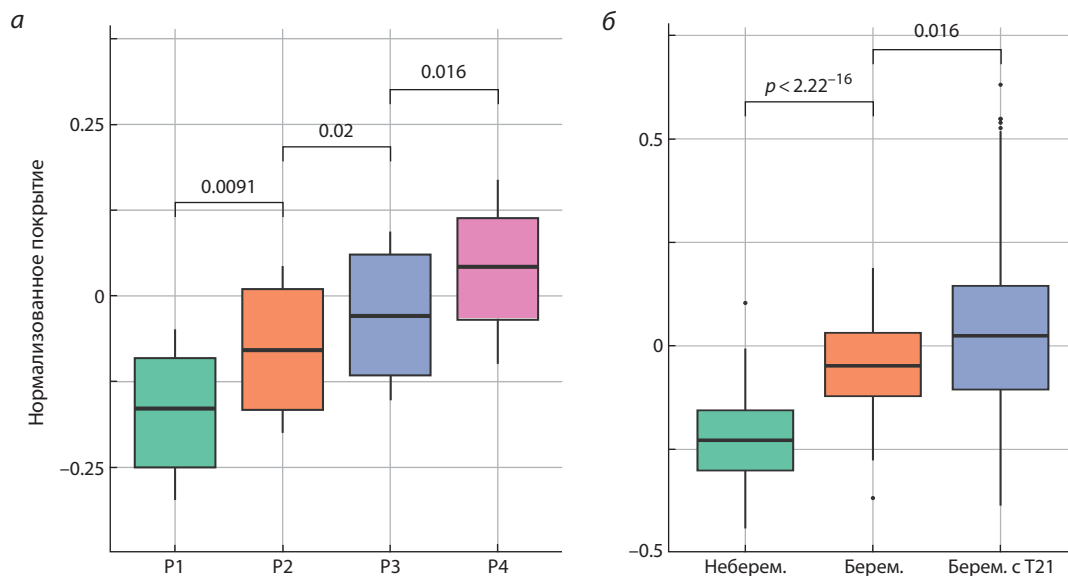


Рис. 1. Нормализованное покрытие (Z-score) в целевых регионах на примере хромосомы 21: а – в подгруппах беременных с нормальным кариотипом плода; б – в группах небеременных, беременных с нормальным кариотипом плода и беременных с трисомией у плода.

Статистическая значимость различий между группами оценивалась с помощью U-критерия Манна–Уитни (критерий Уилкоксона для независимых выборок).

с помощью предложенного метода можно детектировать фетальные анеуплоидии.

Интересно, что в некоторых регионах выявлялись множественные участки с повышенным покрытием у беременных. Проверка в геномном браузер UCSC подтвердила наличие дополнительных DHS в таких регионах (рис. S3), что согласуется с известными данными о кластеризации регуляторных элементов в геноме (Thurman et al., 2012).

Формирование критериев и выбор регионов

Основная задача состояла в поиске регионов, наиболее эффективно различающих беременных и небеременных. Для этого были предложены следующие критерии: длина максимального окна, в котором разность между средним покрытием у небеременных и средним покрытием у беременных меньше нуля; максимальная площадь под кривой разности в окне из 100 нуклеотидов; зависимость величины покрытия от концентрации фетальной вкДНК; а также функциональная значимость. Функциональная значимость оценивалась на основе критериев, предложенных в работе (Thurman et al., 2012). Промоторными сайтами считались те, у которых координаты пересекаются с сайтом начала транскрипции из референсной аннотации генома человека GENCODE (Harrow et al., 2012), или ближайший DHS к сайту начала транскрипции в 5'-направлении. Эхансерами считались те, что находятся на расстоянии не более 500000 нуклеотидов от промоторного сайта и имеют с данным промоторным сайтом высокую корреляцию ($R \geq 0.7$) по данным DNase-seq.

Чтобы определить, какой из предложенных критериев лучше всего подходит для выбора регионов, был построен бинарный классификатор на основе метода случайного

леса, который отличает образцы беременных женщин от образцов небеременных на основании среднего покрытия в 20 наилучших по данным критериям участках. Мы использовали алгоритм случайного леса Random Forest Classifier из библиотеки scikit-learn в стандартной конфигурации (Pedregosa et al., 2011). Качество предсказаний определяли F-мерой для предсказаний образцов небеременных, поскольку количество таких образцов меньше, чем беременных. Для определения наиболее эффективного правила отбора регионов мы оценивали предсказательную способность каждого критерия в отдельности и их возможных комбинаций. Наилучшая F-мера была получена на регионах, отобранных по сочетанию трех критериев: критерия подсчета площади под кривой разности; критерия наличия в регионе зависимости величины покрытия от концентрации плодной ДНК; критерия того, является ли сайт промоторным или эхансерным.

Затем для координаты каждого нуклеотида было найдено среднее значение нормализованного покрытия (Z-score) по всем классам: для образцов небеременных; для образцов беременных (по группам с разной концентрацией вкДНК); для образцов беременных с трисомией по 21-й хромосоме. После этого все регионы конкретной хромосомы были отсортированы по убыванию величины, определенной в выбранном критерии, и отбиралось заданное количество регионов.

Секвенирование отобранных районов

Для анализа 40 отобранных геномных регионов (табл. S2) у 283 образцов использовали метод таргетного секвенирования на основе мультиплексной ПЦР. Специфичные праймеры были разработаны для каждого региона (табл. S3),

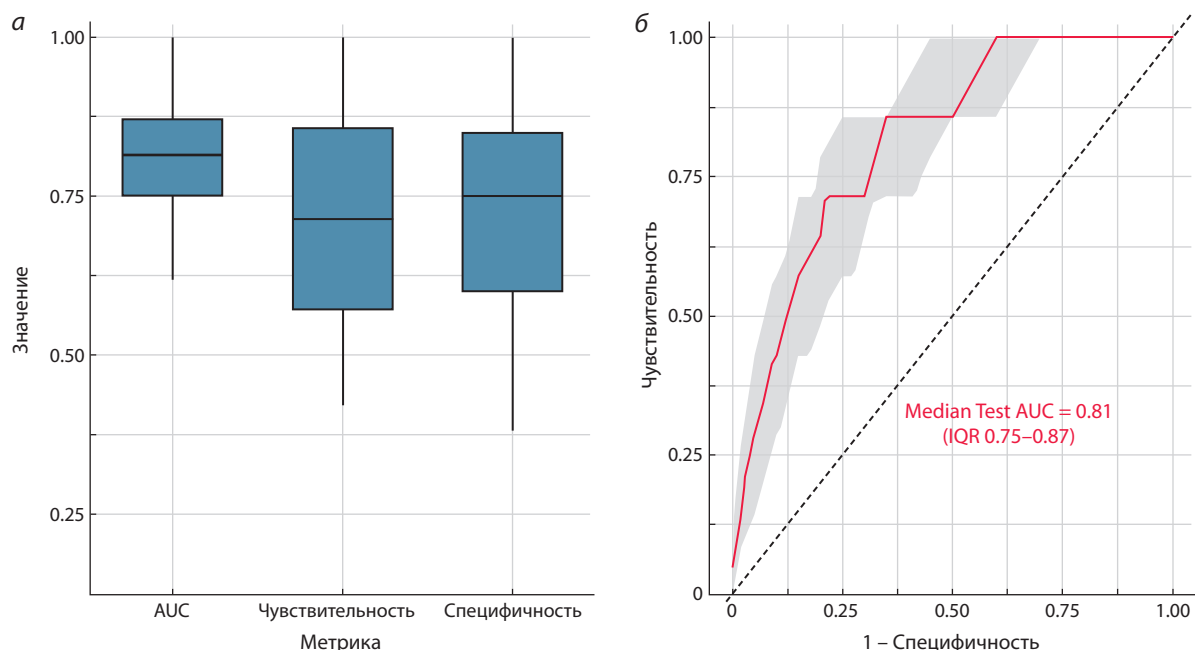


Рис. 2. Характеристики модели: *а* – распределение метрик (AUC (площадь под кривой), чувствительность и специфичность); *б* – ROC, построенная на 100 независимых разбиениях.

что позволило одновременно амплифицировать все целевые участки в одной реакции. Перед секвенированием провели молекулярное баркодирование исходных молекул ДНК на этапе амплификации, что обеспечило точный подсчет количества исходных молекул и минимизировало влияние ПЦР-артефактов. Картирование чтений было проведено таким же образом, как описано в разделе «Алгоритм анализа данных полногеномного секвенирования».

Построение и валидация модели

Для 201 образца с нормальным кариотипом плода и 82 с трисомией был произведен подсчет количества молекул вкДНК, соответствующих 40 выбранным регионам. Чтобы достоверно оценить пригодность этих регионов для использования в тест-системе, мы обучили и протестировали модель на основе LASSO-регрессии, применяя стратифицированный метод повторных случайных подвыборок и взвешивание классов (см. Материалы и методы).

Высокие показатели чувствительности и специфичности (рис. 2, *а*), а также AUC (см. рис. 2, *б*) модели позволяют сделать вывод о возможности предсказания наличия трисомии у плода по данным секвенирования DHS свободно циркулирующей вкДНК.

Обсуждение

До того момента, как анализ вкДНК плазмы крови стал доминирующим подходом в неинвазивном пренатальном тестировании, научное сообщество находилось в активном поиске альтернативных методов, позволяющих получить генетический материал плода без инвазивного вмешательства (Simpson, Elias, 1993).

Основными мишенями в этих исследованиях выступали два типа клеток: трофобласты (Bruch et al., 1991) и ядро-

содержащие эритроциты плода (фетальные эритроциты) (Simpson, Elias, 1994). Трофобласты, являющиеся производными плаценты, представлялись логичной целью, так как они содержат полный геном плода и теоретически могли бы служить идеальным материалом для анализа (Covone et al., 1984). Параллельно велись исследования в области выделения фетальных эритроцитов (Bianchi et al., 1990). Преимущество этого подхода заключалось в том, что незрелые ядросодержащие эритроциты в норме практически отсутствуют в периферической крови здорового взрослого человека, что теоретически позволяло отличить их на фоне материнских клеток.

Однако практическое применение клеточного подхода столкнулось с серьезными ограничениями (Holzgreve et al., 1992). Основной проблемой была чрезвычайно низкая концентрация фетальных клеток в материнской крови, что делает их выделение технически сложным и ненадежным.

На этом фоне начало принципиально иному пути положило открытие, сделанное Деннисом Ло и его коллегами в 1997 г., которое продемонстрировало наличие значительного количества вкДНК плода в плазме крови матери (Lo et al., 1997). Несмотря на то что вкДНК представляла собой фрагментированную плацентарную ДНК, а не полный геном плода, этот подход обладал решающими практическими преимуществами: относительно высокой концентрацией генетического материала и технологической доступностью для анализа с помощью ПЦР (Lo et al., 1998).

Современные коммерческие НИПТ, использующие методы высокопроизводительного секвенирования, показывают высокую точность, но их широкое применение все еще сдерживается относительно высокой стоимостью (Gil et al., 2017).

Наш метод, основанный на таргетном анализе 40 эпигенетических маркеров (DHS), предлагает экономически выгодное решение. Подобные подходы требуют секвенирования менее 1 % генома, что существенно сокращает расходы на реагенты и обработку данных (Schmitt et al., 2015). Следует отметить, что для такого анализа может быть использована не только платформа NGS, но и более доступная технология цифровой ПЦР (dPCR), что дополнительно снижает стоимость и упрощает внедрение метода в рутинную клиническую практику (Li et al., 2022). Это особенно актуально для небольших лабораторий, не имеющих доступа к дорогостоящим системам секвенирования. При этом сохраняются все ключевые преимущества НИПТ: неинвазивность, возможность ранней диагностики и высокая точность определения хромосомных аномалий (Zhang et al., 2026).

Важным аспектом исследования стала разработка строгих критериев отбора информативных регионов. Наш метод учитывает комплекс параметров, включая не только статистические показатели дифференциальной доступности, но и функциональную значимость сайтов, а также зависимость сигнала от концентрации фетальной ДНК. Такой комплексный подход позволяет создать надежную диагностическую панель, устойчивую к техническому шуму. А применение в работе молекулярных баркодов решает такую важную проблему количественного анализа вкДНК, как точный подсчет исходных молекул при наличии ПЦР-амплификации.

Заключение

Перспективы практического применения разработанного подхода выглядят многообещающе. Во-первых, метод принципиально позволяет расширить спектр детектируемых хромосомных аномалий без существенного увеличения стоимости анализа. Во-вторых, его адаптация для различных технологических платформ открывает возможности для создания гибких диагностических систем, которые могут быть оптимизированы под конкретные клинические задачи и бюджетные ограничения.

Дальнейшие исследования могут быть направлены на оптимизацию панели маркеров для одновременного выявления широкого спектра хромосомных аномалий и стандартизацию протоколов для использования в рутинной клинической практике. Предложенный метод имеет все предпосылки стать основой для нового поколения скрининговых тестов, сочетающих высокую надежность и экономическую эффективность.

Список литературы / References

- Albery 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., Metzeau 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., Raymond 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/bfpg/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., Waghoos 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., Huijssens-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-cnrcr2820030106%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

Конфликт интересов. Мазур А.М. является директором по науке в ООО «Геноаналитика» – компании, занимающейся разработкой НИПТ. Однако данное исследование не финансировалось компанией «Геноаналитика» и не было связано с коммерческими интересами компании. Поступила в редакцию 18.12.2025. После доработки 06.02.2026. Принята к публикации 09.02.2026.