

(19)



**Евразийское
патентное
ведомство**

(11) **047854**

(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2024.09.20

(51) Int. Cl. *C12Q 1/6809* (2018.01)

(21) Номер заявки
202191891

(22) Дата подачи заявки
2020.02.11

(54) СПОСОБЫ НЕИНВАЗИВНОГО ПРЕНАТАЛЬНОГО ТЕСТИРОВАНИЯ АНОМАЛИЙ ПЛОДА

(31) 19156991.2

(32) 2019.02.13

(33) EP

(43) 2021.12.15

(86) PCT/EP2020/053504

(87) WO 2020/165190 2020.08.20

(71)(73) Заявитель и патентовладелец:
МЕДИКАВЕР ПАБЛИК КО ЛТД (СУ)

(72) Изобретатель:
**Кумбарис Джордж, Ахиллеос
Ахиллеас, Тсангарас Кириакос,
Лойзидес Хараламбос, Иоаннидес
Мариос, Патсалис Филиппос (СУ)**

(74) Представитель:
Нилова М.И. (RU)

(56) US-A1-2016340733
WO-A1-2018081130

KUN SUN ET AL.: "Size-tagged preferred ends in maternal plasma DNA shed light on the production mechanism and show utility in noninvasive prenatal testing", PNAS, vol. 115, no. 22, 14 May 2018 (2018-05-14), pages E5106-E5114, XP055612386, US, ISSN: 0027-8424, DOI: 10.1073/pnas.1804134115, page E5107, right-hand column; figures 1, 2, 5, 6

MATTHEW W. SNYDER ET AL.: "Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin", CELL, vol. 164, no. 1-2, 14 January 2016 (2016-01-14), pages 57-68, XP055367434, AMSTERDAM, NL, ISSN: 0092-8674, DOI: 10.1016/j.cell.2015.11.050, the whole document

(57) Настоящее изобретение относится к способу обнаружения генетических и/или геномных аномалий в смешанном образце, включающему этапы биохимического и in silico обогащения подмножества внеклеточных фрагментов ДНК, происходящих из смешанного образца. В изобретении применяют пул длинных ДНК-зондов для обогащения представляющих интерес последовательностей в смешанном образце с последующим массовым параллельным секвенированием и основанным на применении компьютера анализом обогащенной субпопуляции для выявления риска генетических и/или геномных аномалий в указанной субпопуляции смешанного образца. Основанная на применении компьютера часть способа не требует обязательного выравнивания по эталонному геному или калибровочных значений с применением эталонных образцов. Способ также включает набор для осуществления изобретения.

B1

047854

047854 B1

Область техники

Настоящее изобретение относится к области биологии, медицины и химии, в частности к области молекулярной биологии и более конкретно к области молекулярной диагностики.

Уровень техники

Обнаружение внеклеточной фетальной ДНК (cffDNA) в материнской плазме в значительной степени способствовало развитию неинвазивных пренатальных тестов. Однако большинство разработанных тестов для обнаружения у плода анеуплоидии и микроделеций основаны на единичных нормализованных значениях, полученных на основе информации о глубине прочтения. Хотя эти тесты можно рассматривать как значительное улучшение по сравнению с существующими способами, их клиническая чувствительность не превышает 99%. Особенно в том случае, если доля cffDNA в материнском кровотоке ниже 4%, даже с технологией секвенирования следующего поколения (NGS), которая имеет высокую чувствительность, получение достаточной точности для неинвазивного пренатального тестирования (НИПТ, NIPT) является сложной задачей.

Краткое описание изобретения

Согласно настоящему изобретению, предложен способ двойного обогащения фрагментов внеклеточной ДНК (вкДНК, cfDNA), происходящих из плаценты, в смешанном биологическом образце, содержащем внеклеточную фетальную и материнскую ДНК, с применением биохимических и *in silico* подходов, которые увеличивают отношение сигнал/шум за счет применения длинных зондов захвата, анализа размера фрагментов, применения горячих точек неслучайной фрагментации (HSNRF) и новой концепции в биоинформатике. Указанный способ обеспечивает высокочувствительное неинвазивное обнаружение аномалий у плода за счет применения информации, полученной в результате обогащения фрагментов внеклеточной фетальной ДНК в пренатальном образце с применением длинных ДНК-зондов, которые захватывают фрагменты внеклеточной ДНК и/или обогащают в отношении HSNRF.

В настоящем изобретении также предложен новый основанный на применении компьютера способ, который упорядочивает наблюдаемые данные (прочтения секвенирования) в значимые структуры, которые улучшают отношение сигнал/шум при анализе образцов внеклеточной ДНК, содержащих смесь фрагментов вкДНК. В способе согласно настоящему изобретению множество последовательностей ДНК группируют таким образом, что степень гомологии между указанными последовательностями ДНК является максимальной, если они действительно происходят из одной и той же структуры, и минимальной в противном случае. Конкретные паттерны последовательностей идентифицируют в указанных данных и применяют для их распределения в заданные представляющие интерес области. Указанный способ можно применять для обнаружения конкретных структур в данных о последовательности ДНК без необходимости получения каких-либо предварительных знаний, таких как выравнивание на эталонном геноме человека и/или калибровочные значения с применением эталонных образцов, и указанный способ применяют в настоящей заявке для неинвазивного пренатального обнаружения хромосомных аномалий у плода, таких как анеуплоидии, микроделеции, микродупликации и точечные мутации.

Таким образом, в первом аспекте настоящего изобретения относится к способу обнаружения генетических/геномных аномалий у плода в смешанном образце, содержащем материнскую и фетальную вкДНК, причем указанный способ включает следующие этапы:

- (i) получение смеси вкДНК от индивидуума;
- (ii) приготовление библиотеки секвенирования из вкДНК;
- (iii) обогащение библиотеки вкДНК с применением длинных ДНК-зондов;
- (iv) секвенирование обогащенной библиотеки вкДНК;
- (v) осуществление статистического анализа для определения риска хромосомных и/или других генетических аномалий в фетальной ДНК;

при этом риск хромосомной и/или другой генетической аномалии в фетальной ДНК классифицируют на основе способа двойного обогащения, включающего геномную гибридизацию и основанный на применении компьютера способ, который увеличивает отношение сигнал/шум в указанном анализе.

Способ, в котором обогащенные фрагменты подвергают процедуре, включает этапы:

- (i) группирование секвенированных прочтений, которые могут быть спарены на основе нуклеотидных паттернов, то есть но не ограничиваясь указанным, степени перекрытия с точки зрения нуклеотидного состава;
- (ii) группирование и/или аннотация по меньшей мере подмножества множества секвенированных прочтений на основе нуклеотидных паттернов, то есть но не ограничиваясь указанным, сходства последовательности по меньшей мере 20 наиболее удаленных от середины последовательности нуклеотидов;
- (iii) сопоставление заданных нуклеотидных последовательностей с прочтениями согласно (i) и (ii);
- (iv) удаление повторяющихся секвенированных прочтений;
- (v) применение информации, полученной на этапах (i)-(iv), для осуществления статистического анализа по определению риска хромосомной и/или другой генетической аномалии в фетальной ДНК.

В другом аспекте согласно настоящему изобретению предложен набор для осуществления указанного изобретения. В одном варианте реализации набор содержит:

- а) зонды, которые гибридизуются по меньшей мере с одним участком во фрагментах нуклеиновой

кислоты, при этом по меньшей мере один указанный участок частично или полностью охватывает фрагмент нуклеиновой кислоты, где находится горячая точка неслучайной фрагментации (HSNRF), и, необязательно,

б) реагенты и/или программное обеспечение для осуществления изобретения и обнаружения генетических и/или геномных аномалий в образце.

Краткое описание фигур

На фиг. 1 показан эксперимент проверки концепции, в котором 92 образца с диплоидным набором хромосом и четыре образца с трисомией по 21-й хромосоме были классифицированы с применением указанного способа.

На фиг. 2 показан эксперимент проверки концепции, в котором 91 нормальный образец и два случая синдрома делеции хромосомы 22q11.2 были классифицированы в отношении микроделеции (синдром делеции хромосомы 22q11.2) с применением указанного способа.

На фиг. 3 показана ядерная оценка плотности распределения размеров фрагментов из репрезентативной выборки с применением (А) всех фрагментов и (В) только фрагментов, выбранных указанным способом.

Подробное описание изобретения

Согласно первому аспекту настоящего изобретения в настоящей заявке раскрыт способ обнаружения хромосомной аномалии в смешанном образце, включающий этапы:

(a) получение биологического образца, который содержит смесь фрагментов внеклеточной ДНК (вкДНК),

(b) приготовление библиотеки секвенирования из фрагментов вкДНК,

(c) гибридизация одного или более зондов по меньшей мере с одним или более фрагментами вкДНК,

(d) выделение из библиотеки фрагментов вкДНК, которые связываются с зондами,

(e) секвенирование фрагментов вкДНК из библиотеки, которые связываются с зондами,

(f) применение информации о размере, начале и/или конце каждого или подмножества обогащенных фрагментов вкДНК согласно этапам (c-e) для выбора фракции фрагментов вкДНК, гибридованных с указанными одним или более зондами.

при этом этап (f) связан с вычислением нескольких статистических тестов.

Таким образом, настоящее изобретение относится к способу обнаружения генетических/геномных аномалий у плода в смешанном образце, содержащем материнскую и фетальную вкДНК, причем способ включает следующие этапы:

(i) получение смеси вкДНК от индивидуума; (ii) приготовление библиотеки секвенирования из вкДНК; (iii) обогащение библиотеки вкДНК с применением длинных ДНК-зондов; (iv) секвенирование обогащенной библиотеки вкДНК;

(v) осуществление статистического анализа для определения риска хромосомной и/или другой генетической аномалии в фетальной ДНК.

В настоящей заявке HSNRF определяют как геномную область, содержащую на расстоянии менее 300 пар оснований (п.о.) (предпочтительно менее 200 п.о., более предпочтительно менее 100 п.о.) предпочтительные сайты, различающие два типа тканей, присутствующих в смеси вкДНК, и где указанные предпочтительные сайты присутствуют в областях HSNRF с большей частотой, чем в других областях, не относящихся к HSNRF. В настоящей заявке предпочтительные сайты называют геномными основаниями, в которых частота нахождения конечной точки прочтения достоверно различается (р-значение по меньшей мере менее 0.05) между двумя типами тканей, присутствующими в смеси вкДНК.

В настоящей заявке смесь фрагментов нуклеиновой кислоты выделяют предпочтительно из образца, взятого из эукариотического организма, предпочтительно примата, более предпочтительно человека.

В контексте настоящего изобретения выражения "фрагменты нуклеиновой кислоты" и "фрагментированные нуклеиновые кислоты" можно применять как взаимозаменяемые.

В одном варианте реализации зонды представляют собой длинные молекулы ДНК и:

(i) длина каждого зонда составляет от 100 до 500 пар оснований,

(ii) каждый денатурированный зонд имеет 5'-конец и 3'-конец,

(iii) предпочтительно каждый зонд связывается с HSNRF на расстоянии не менее 10 пар оснований как на 5'-конце, так и на 3'-конце, от областей, несущих вариации числа копий (CNV), сегментные дупликации или повторяющиеся последовательности ДНК, и

(iv) GC-состав (гуанин-цитозиновый) в каждом зонде составляет от 19 до 80%.

Применяемый в настоящей заявке термин "длинные ДНК-зонды" относится к зондам с размером в диапазоне от 100 до 500 п.о.

В одном варианте реализации зонды имеют размер в диапазоне от 150 до 250 п.о. В другом варианте реализации зонды имеют размер в диапазоне от 160 до 180 п.о.

В одном варианте реализации способа согласно настоящему изобретению, длинные ДНК-зонды охватывают HSNRF.

В другом варианте реализации способа согласно настоящему изобретению, обогащенная библиоте-

ка вкДНК включает HSNRF.

В предпочтительном варианте реализации способа согласно изобретению, фрагменты нуклеиновой кислоты представляют собой циркулирующую вкДНК или РНК.

В одном варианте реализации образец представляет собой образец материнской плазмы, содержащий внеклеточную материнскую ДНК и внеклеточную фетальную ДНК (cffDNA).

Настоящее изобретение также можно применять с множеством биологических образцов. По существу любой биологический образец, содержащий генетический материал, например РНК или ДНК и, в частности, вкДНК, можно применять в качестве образца в настоящем изобретении. В одном варианте реализации источником образца ДНК является образец плазмы, содержащий вкДНК. В частности, для пренатального тестирования образец ДНК содержит фетальную ДНК (например, cffDNA). В одном варианте реализации для НИПТ образец представляет собой смешанный образец, который содержит как материнскую ДНК, так и фетальную ДНК (например, cffDNA), как, например, образец материнской плазмы, полученный из материнской периферической крови. Обычно для смешанных образцов материнской/фетальной ДНК образец представляет собой образец плазмы матери, хотя можно применять другие источники ткани, содержащие как материнскую ДНК, так и фетальную ДНК. Применяемый в настоящей заявке термин "смешанный образец" относится к смеси, по меньшей мере, двух биологических образцов, происходящих из разных источников, например, образцов материнской/фетальной ДНК.

В зависимости от условий биологический образец включает эмбриональную ДНК и материнскую ДНК, происходящую из опухоли ДНК и не происходящую из опухоли ДНК, ДНК патогена или ДНК хозяина и ДНК, происходящую из трансплантированного органа, и ДНК, происходящую от хозяина.

Таким образом, в контексте неинвазивной диагностики образец представляет собой смешанный образец, где указанный смешанный образец выбран из группы, содержащей (i) эмбриональную ДНК и материнскую ДНК, (ii) происходящую из опухоли ДНК и происходящую из опухоли ДНК, (iii) ДНК патогена и ДНК хозяина и (iv) ДНК, происходящую из трансплантированного органа, и ДНК, происходящую от хозяина.

Материнскую плазму можно получить из образца периферической цельной крови беременного субъекта, и плазму можно получить стандартными способами. Всего 1-4 мл плазмы достаточно для обеспечения подходящего материала ДНК для анализа в соответствии со способом, описанным в настоящем изобретении. Затем тотальную вкДНК можно экстрагировать из образца с применением стандартных методик, неограничивающие примеры которых включают протокол QIASymphony (QIAGEN), подходящий для выделения cffDNA, или любой другой ручной или автоматический способ экстракции, подходящий для выделения внеклеточной ДНК.

В контексте настоящего изобретения термин "субъект" относится к животным, предпочтительно млекопитающим и, более предпочтительно, человеку. Упомянутый в настоящей заявке "субъект" является беременным субъектом и, следовательно, предпочтительно субъектом женского пола. Беременный субъект может находиться на любом сроке беременности. "Субъект" может забеременеть естественным путем или с помощью искусственных способов. Применяемый в настоящей заявке термин "субъект" также относится к субъекту, страдающему опухолью или у которого есть подозрение на наличие опухоли. Вышеуказанный субъект может быть подвергнут трансплантации органа или быть инфицированным патогеном после трансплантации или независимо от трансплантата.

Для подготовки биологических образцов, как правило, ДНК экстрагируют стандартными способами, известными в данной области техники, неограничивающим примером которых является протокол QIASymphony (QIAGEN).

После выделения вкДНК образца применяют для конструирования библиотеки секвенирования, чтобы сделать образец совместимым с технологией последующего секвенирования, такой как секвенирование следующего поколения (NGS). Обычно это включает лигирование адаптеров на концах фрагментов вкДНК с последующей амплификацией. Наборы для приготовления библиотек для секвенирования коммерчески доступны или могут быть разработаны.

Предпочтительно в способе согласно настоящему изобретению первый и второй фрагменты нуклеиновой кислоты выбраны из групп, содержащих:

- i) эмбриональную ДНК и материнскую ДНК,
- ii) происходящую из опухоли ДНК и не происходящую из опухоли ДНК,
- iii) ДНК патогена и ДНК хозяина,
- iv) ДНК, происходящую из трансплантированного органа, и ДНК, происходящую от хозяина.

В контексте настоящего изобретения термины "плод" и "эмбрион" применяют взаимозаменяемо.

В одном варианте реализации способа выбор фракции фрагментов вкДНК, гибридизирующихся с одним или более зондами, охватывающими области HSNRF, включает следующие этапы:

- (i) отнесение фрагментов вкДНК к первому и второму кластерному распределению,
- (ii) обнаружение горячих точек неслучайной фрагментации (HSNRF) с применением фрагментов вкДНК первого кластерного распределения,
- (iii) отнесение фрагментов вкДНК из второго кластерного распределения к третьему кластеру,
- (iv) объединение фрагментов вкДНК первого кластерного распределения с фрагментами вкДНК

третьего кластерного распределения.

Применяемый в данной заявке термин "кластерное распределение" относится к подмножеству (группе) фрагментов вкДНК, обладающих определенными свойствами, такими как, но не ограничиваясь указанным, длина фрагмента или последовательность фрагмента. В контексте настоящего изобретения "кластерное распределение" основано на размере и последовательности фрагментов, приписанной каждой группе.

В одном варианте реализации первое кластерное распределение содержит фрагменты вкДНК, имеющие длину, меньшую или равную 120 п.о., или 125 п.о., или 130 п.о., или 135 п.о., или 140 п.о., или 145 п.о., или 150 п.о., или 155 п.о., и второе кластерное распределение содержит фрагменты вкДНК, имеющие длину более 120 п.о., или 125 п.о., или 130 п.о., или 135 п.о., или 140 п.о., или 145 п.о., или 150 п.о., или 155 п.о., соответственно.

В другом варианте реализации третий кластер включает выбор фрагментов вкДНК из второго кластера, концы которых перекрываются с HSNRF, обнаруженной на этапе (ii), описанном выше.

В одном варианте реализации предпочтительно зонды охватывают горячую точку сайта неслучайной фрагментации (HSNRF), так что только 5'-конец фрагментированной нуклеиновой кислоты захватывается зондом.

В другом варианте реализации зонды охватывают сайт HSNRF, так что только 3'-конец внеклеточных нуклеиновых кислот, происходящих из HSNRF, может связываться с зондом.

В другом предпочтительном варианте реализации зонды охватывают оба сайта HSNRF, связанные с фрагментированной нуклеиновой кислотой, так что зонд захватывает как 5'-, так и 3'-конец внеклеточной нуклеиновой кислоты, связанной с данным сайтом HSNRF.

В другом варианте реализации применяют комбинации вышеперечисленного.

В идеальном случае, в способе согласно настоящему изобретению, зонды, которые применяют для обогащения вкДНК, представляют собой двухцепочечные фрагменты ДНК, и (i) длина каждого зонда составляет 100-500 пар оснований, (ii) каждый зонд имеет 5'-конец и 3'-конец, (iii) предпочтительно каждый зонд связывается с HSNRF на расстоянии не менее 10 пар оснований, как на 5'-конце, так и на 3'-конце, от областей, несущих вариации числа копий (CNV), сегментные дубликации или повторяющиеся последовательности ДНК, и (iv) GC-состав в каждом зонде составляет от 19 до 80%.

В целом, этап гибридизации, предпочтительно с зондами, как описано выше, может быть проведен до создания библиотеки секвенирования или после создания библиотеки.

Представляющая(е) интерес область(и) на представляющей(их) интерес хромосоме(ах), где находится HSNRF, обогащает(ют)ся путем гибридизации пула зондов с библиотекой секвенирования с последующим выделением тех последовательностей из библиотеки секвенирования, которые связываются с зондами. В одном варианте реализации зонд охватывает сайт HSNRF, так что только 5'-конец фрагментированных внеклеточных нуклеиновых кислот захватывается зондом. В другом варианте реализации зонд охватывает сайт HSNRF, так что только 3'-конец фрагментированных внеклеточных нуклеиновых кислот, происходящих из HSNRF, может связываться с зондом. В другом предпочтительном варианте реализации зонд охватывает оба сайта HSNRF, связанных с фрагментированной нуклеиновой кислотой, так что зонд захватывает как 5'-, так и 3'-конец внеклеточной нуклеиновой кислоты, связанной с данным сайтом HSNRF.

Чтобы облегчить выделение желаемых обогащенных последовательностей (HSNRF), обычно последовательности зондов модифицируют таким образом, чтобы последовательности, которые гибридизуются с зондами, можно было отделить от последовательностей, которые не гибридизуются с зондами. Обычно это достигается путем фиксации зондов на подложке. Это позволяет физически отделить те последовательности, которые связывают зонды, от тех последовательностей, которые не связывают зонды. В одном варианте реализации каждая последовательность в пуле зондов может быть помечена биотином, и затем этот пул может быть связан с гранулами, покрытыми биотин-связывающим веществом, таким как стрептавидин или авидин. В предпочтительном варианте реализации зонды метят биотином и связывают с магнитными гранулами, покрытыми стрептавидином, что позволяет разделить их за счет применения магнитных свойств гранул. Однако среднему специалисту в данной области будет понятно, что в данной области существуют другие системы аффинного связывания, которые можно применять вместо системы биотин-стрептавидин/авидин. Например, можно применять систему на основе антител, в которой зонды метят антигеном, и затем они связываются с гранулами, покрытыми антителами. Более того, зонды могут включать на одном конце маркер последовательности и могут быть связаны с подложкой через комплементарную последовательность, которая гибридизуется с таким маркером последовательности. Кроме того, в дополнение к магнитным гранулам, могут применяться другие типы подложек, такие как полимерные гранулы и тому подобное.

В некоторых вариантах реализации элементы библиотеки секвенирования, которые связываются с пулом зондов, полностью комплементарны зондам. В других вариантах реализации элементы библиотеки секвенирования, которые связываются с пулом зондов, частично комплементарны зондам. Например, при определенных обстоятельствах желательно применять и анализировать данные, полученные из фрагментов ДНК, которые являются продуктами процесса обогащения, но не обязательно принадлежат

интересующим областям генома (например, такие фрагменты ДНК могут связываться с зондами благодаря частичной гомологии (частичная комплементарность), и при секвенировании будут давать очень низкое покрытие по всему геному в координатах, отличных от зондов).

После обогащения представляющей(их) интерес последовательности (последовательностей) с применением зондов и формирования таким образом библиотеки ДНК, обогащенной сайтами HSNRF, элементы обогащенной библиотеки HSNRF элюируют с твердой подложки, амплифицируют и секвенируют с применением стандартных способов, известных в данной области техники.

Обычно применяют секвенирование следующего поколения (NGS), хотя также можно применять другие технологии секвенирования, которые обеспечивают очень точные вычисления в дополнение к информации о последовательности. Соответственно, другие способы точных вычислений, такие как цифровая полимеразная цепная реакция (ПЦР), секвенирование отдельных молекул, нанопоровое секвенирование и микрочипы также можно применять вместо NGS.

В одном варианте реализации способа обогащенный образец секвенируют с применением способа секвенирования спаренных концов, предпочтительно способа секвенирования 2×75 п.о., чтобы обеспечить сбор данных о начальных/конечных положениях секвенированного фрагмента как с 5'-, так и с 3'-конца, а также для получения достаточной информации о секвенировании фрагмента, чтобы обеспечить *de novo* выравнивание/самовыравнивание.

В другом варианте реализации секвенированные фрагменты группируют по размеру на основе степени перекрытия или отсутствия перекрытия их спаренных концов.

Изобретение относится к способу, в котором фрагмент нуклеиновой кислоты, который должен быть обнаружен, присутствует в смеси в концентрации ниже, чем фрагмент нуклеиновой кислоты из того же генетического локуса, но другого происхождения. Это означает, что, если выбран конкретный локус, например, материнская копия будет присутствовать 100 раз, а копия от плода только один раз, в растворе содержащем выделенную вкДНК. В случае вкДНК, происходящей из плода, компонент смешанного образца, происходящий из плода, может иметь диапазон возможных значений. Например, диапазон фетального материала в смешанном образце может быть от 2 до 30%. Часто фрагменты, происходящие от плода, составляют приблизительно 10% от общей ДНК смешанного образца. Что еще более важно, в некоторых композициях смешанных образцов компонент фетальной ДНК в образце может составлять менее 5%. В частности, в некоторых композициях образцов материал, происходящий из плода, составляет 3% или менее от всего образца.

Настоящий способ особенно подходит для анализа таких низких концентраций целевой вкДНК. В способе, согласно настоящему изобретению, фрагмент нуклеиновой кислоты, который подложит обнаружению или происхождение которого должно быть определено, и фрагмент нуклеиновой кислоты из того же генетического локуса, но другого происхождения, присутствуют в смеси в соотношении, выбранном из группы: 1:2, 1:4, 1:10, 1:20, 1:50, 1:100, 1:200, 1:500, 1:1000, 1:2000 и 1:5000. Соотношения следует понимать как приблизительные отношения, что подразумевает отклонения плюс/минус 30%, 20% или 10%. Специалист в данной области техники знает, что такие отношения не будут иметь точных числовых значений, указанных выше. Эти отношения являются отношением количества молекул, специфичных для локуса, для редкого типа к количеству молекул локуса для распространенного типа.

В одном варианте реализации зонды обеспечивают в форме, которая позволяет им связываться с подложкой, такой как биотинилированные зонды. В другом варианте реализации зонды обеспечивают вместе с подложкой, такие как биотинилированные зонды, которые обеспечивают совместно с магнитными гранулами, покрытыми стрептавидином.

В отдельном варианте реализации GC-состав этих или других зондов составляет от 10 до 70%, предпочтительно от 15 до 60%, более предпочтительно от 20 до 50%.

В описанном способе результат может быть объединен с дополнительными статистическими тестами из группы, включающей t-критерий Стьюдента, двумерный непараметрический бутстреп-тест, стратифицированный пермутационный тест и тест на пропорции размеров фрагментов.

В одном варианте реализации способ дополнительно включает этап объединения статистических тестов, выбранных из группы, включающей, но не ограничивающейся указанным, t-критерий Стьюдента, двумерный непараметрический бутстреп-тест, стратифицированный пермутационный тест, ANOVA, любые тесты на пропорции и/или регрессионную модель.

В контексте настоящего изобретения термин "частично" относится к области (участку) из 10, 20, 30 или 40 оснований фрагмента нуклеиновой кислоты от 5'-конца или от 3'-конца. Следовательно, применяемый в данной заявке термин "полностью" относится к области (участку), которая охватывает 100% фрагмента нуклеиновой кислоты. Согласно способу настоящего изобретения зонды гибридизуются по меньшей мере с одним участком внутри фрагмента нуклеиновой кислоты. Однако зонды также могут быть нацелены на более чем один участок в одном и том же фрагменте нуклеиновой кислоты.

В одном аспекте данной заявки раскрыты зонды для применения в способе согласно настоящему изобретению.

В другом аспекте данной заявки раскрыт способ для применения при диагностике и/или скрининге генетической аномалии в образце.

В одном варианте реализации генетическая аномалия выбрана из группы, включающей, но не ограничивающейся указанным:

(a) анеуплоидии хромосом 13, 18, 21 и/или X, Y;

(b) структурные аномалии, включая, но не ограничиваясь указанным, изменения числа копий, в том числе микроделеции и микродупликации, вставки, делеции, транслокации и мутации небольшого размера, в том числе точечные мутации.

В одном варианте реализации способ применяют для обнаружения эпигенетических аномалий, включая, но не ограничиваясь указанным, модификации ДНК и гистонов. Применяемый в данной заявке термин "эпигенетические аномалии" относится к изменениям экспрессии гена с изменением или без изменения последовательности ДНК. Соответственно, этот термин включает анеуплоидии, микроделеции, микродупликации и точечные мутации, а также изменения эпигенетических модификаций, таких как метилирование нуклеотидов в ДНК, или модификаций гистонов, таких как ацетилирование или деацетилирование гистонов, метилирование, убиквитилирование, фосфорилирование, сумоилирование и т.д.

В соответствии с другим аспектом настоящего изобретения в данной заявке раскрыт способ двойного обогащения фрагментов, происходящих из плаценты, в смеси внеклеточной ДНК (вкДНК), включающий следующие этапы:

(a) получение биологического образца, который содержит смесь фрагментов внеклеточной ДНК (вкДНК),

(b) приготовление библиотеки секвенирования из фрагментов вкДНК,

(c) гибридизацию библиотеки вкДНК с множеством зондов, причем указанные зонды предпочтительно охватывают по меньшей мере одну HSNRF,

(d) выделение фрагментов вкДНК из библиотеки, которые связываются с зондами,

(e) секвенирование фрагментов вкДНК из библиотеки, которые связываются с зондами,

(f) удаление повторяющихся секвенированных прочтений,

(g) отбор коротких фрагментов вкДНК,

(h) обнаружение HSNRF в коротких фрагментах вкДНК,

(i) отбор длинных фрагментов вкДНК, концы которых перекрываются с HSNRF,

(j) сопоставление выбранных фрагментов вкДНК из (g) и (i) с последовательностями зонда.

В одном варианте реализации короткие фрагменты вкДНК имеют длину меньше или равную 120 п.о., или 125 п.о., или 130 п.о., или 135 п.о., или 140 п.о., или 145 п.о., или 150 п.о., или 155 п.о., и длинные фрагменты вкДНК имеют длину выше, чем 120 п.о., или 125 п.о., или 130 п.о., или 135 п.о., или 140 п.о. или 145 п.о., или 150 п.о., или 155 п.о., соответственно.

В другом варианте реализации фрагменты вкДНК с этапа (j) применяют для обнаружения анеуплоидии у плода в неинвазивном пренатальном диагностическом тесте.

В другом варианте реализации фрагменты вкДНК с этапа (j) применяют для обнаружения микроделеции/микродупликации у плода в неинвазивном пренатальном диагностическом тесте.

В другом варианте реализации фрагменты вкДНК с этапа (j) применяют для обнаружения у плода вставок, делеций, транслокаций и мутаций небольшого размера, включая точечные мутации.

В другом аспекте настоящее изобретение обеспечивает набор для осуществления способов согласно настоящему изобретению, содержащий:

a) зонды, которые гибридизуются по меньшей мере с одним участком во фрагменте нуклеиновой кислоты, причем указанный по меньшей мере один участок частично или полностью охватывает фрагмент нуклеиновой кислоты и, необязательно,

b) реагенты и/или программное обеспечение для осуществления способа определения и/или обнаружения.

В другом аспекте, в настоящей заявке раскрыт набор для осуществления неинвазивного теста для применения в способе согласно настоящему изобретению, при этом набор включает:

a) зонды, которые гибридизуются по меньшей мере с одним участком во фрагментах нуклеиновой кислоты, причем указанный по меньшей мере один участок предпочтительно охватывает по меньшей мере одну HSNRF, и, необязательно,

b) реагенты и/или программное обеспечение для осуществления способа, описанного в аспектах и вариантах реализации настоящего изобретения.

В одном варианте реализации набор содержит контейнер, состоящий из пула зондов, программного обеспечения и инструкций для осуществления способа.

В дополнение к пулу зондов набор может содержать (i) один или более компонентов для выделения вкДНК из биологического образца, (ii) один или более компонентов для приготовления и обогащения библиотеки секвенирования (например, праймеры, адаптеры, буферы, линкеры, модифицирующие ДНК ферменты, ферменты лигирования, зонды, полимеразные ферменты и т.п.), (iii) один или более компонентов для амплификации и/или секвенирования обогащенной библиотеки, и/или (iv) программное обеспечение для осуществления статистического анализа.

Соответственно, в другом варианте реализации настоящего изобретения предпочтительно, чтобы фрагмент, который необходимо обнаружить или определить, имел меньший размер, чем второй фраг-

мент. В зависимости от ткани происхождения ожидают разные размеры фрагментов в участках расположения HSNRF. Поскольку клетки каждой ткани дифференцированы для выполнения определенных функций, то ожидают, что для каждого типа ткани будут наблюдаться разные паттерны генетической активности. Поскольку активация гена зависит от третичной структуры ДНК, разные ткани будут иметь несколько разные конформации, что приводит к разным сайтам расщепления и, следовательно, разным размерам фрагментов. Например, было продемонстрировано, что внеклеточная ДНК плацентарного происхождения, вероятно, будет иметь меньший размер по сравнению с внеклеточной ДНК материнского происхождения.

Примеры

Следующие далее примеры иллюстрируют настоящее изобретение, их не следует рассматривать как дополнительное ограничение.

Пример 1. Сбор образцов и приготовление библиотеки.

Объяснена общая методология двойного обогащения фрагментов ДНК, происходящих из плаценты, в смешанном биологическом образце, содержащем внеклеточную фетальную и материнскую ДНК, для целей неинвазивной пренатальной диагностики. В этом примере описаны способы сбора и обработки образца материнской плазмы (содержащей материнскую и фетальную ДНК). Тот же подход можно применять в других клинически применимых случаях, таких как, но не ограничиваясь указанным, онкология, генетическая мутация, трансплантация и оценка патогенной нагрузки. В другом аспекте тот же подход можно применять для обнаружения эпигенетических модификаций.

Сбор образцов.

Образцы плазмы были получены анонимно у беременных женщин после 10-й недели беременности. Протоколы, применяемые для сбора образцов, были одобрены соответствующим комитетом по биоэтике, и от всех участников было получено информированное согласие.

Экстракция образца.

Внеклеточную ДНК экстрагировали из плазмы каждого индивидуума с применением ручного или автоматического способа экстракции, подходящего для выделения внеклеточной ДНК, такого как, но не ограничиваясь указанным, протокол QIAasympphony, подходящего для выделения внеклеточной ДНК (QIAGEN).

Приготовление библиотеки секвенирования.

Выделенную внеклеточную ДНК из образцов материнской плазмы применяли для построения библиотеки секвенирования. Первоначально 5'- и 3'-"липкие" концы были восстановлены, в то время как 5'-концы были фосфорилированы. Продукты реакции очищали с применением гранул AMPure XP (Beckman Coulter). Затем к обоим концам ДНК лигировали адаптеры для секвенирования с последующей очисткой с применением гранул AMPure XP (Beckman Coulter). Одноцепочечные разрывы удаляли в реакционной смеси с полимеразой и затем очищали с применением гранул AMPure XP (Beckman Coulter). Амплификацию библиотеки проводили с применением другого фермента полимеразы (Koumbaris et al. (2016) *Clinical Chemistry* 62 (6): 848-855). Конечные продукты библиотеки очищали с применением гранул AMPure XP (Beckman Coulter) и измеряли спектрофотометрически.

Пример 2. Конструирование и приготовление зонда.

В этом примере описано приготовление специальных зондов для обнаружения HSNRF. Геномные целевые локусы, применяемые для конструирования зондов, были выбраны на основе их GC-состава и расстояния от них до повторяющихся элементов (минимум 50 п.о.). Размер зондов может варьировать. В одном варианте реализации способы зонды имеют размер от 100 до 500 п.о. и их генерируют с помощью основанного на ПЦР подхода. Зонды получали с помощью симплексной ГЩР с применением стандартной полимеразы Taq, праймеров, предназначенных для амплификации локусов-мишеней, и нормальной ДНК, применяемой в качестве матрицы. В предпочтительном варианте реализации зонд охватывает сайт HSNRF, таким образом, что только 5'-конец фрагментированной нуклеиновой кислоты захватывается зондом. В другом варианте реализации зонд охватывает сайт HSNRF, так что только 3'-конец внеклеточных нуклеиновых кислот, происходящих из HSNRF, может связываться с зондом. В другом предпочтительном варианте реализации зонд охватывает оба сайта HSNRF, связанные с фрагментированной нуклеиновой кислотой, так что зонд захватывает как 5'-, так и 3'-конец внеклеточной нуклеиновой кислоты, связанной с данным сайтом HSNRF.

Пример 3. Гибридизация и амплификация зонда.

В этом примере описан способ целевого захвата нуклеиновых кислот путем гибридизации с применением зондов, причем данные зонды предпочтительно охватывают HSNRF, с последующим количественным определением захваченных последовательностей с помощью секвенирования следующего поколения (NGS).

Биотинилирование зонда.

Зонды были подготовлены для гибридизации, начиная с тупого конца, с последующей очисткой. Затем их лигировали с помощью адаптера биотина и очищали. Зонды денатурировали перед иммобилизацией на магнитных гранулах, покрытых стрептавидином.

Гибридизация зондов.

Аmplифицированные библиотеки смешивали с блокирующими олигонуклеотидами, ДНК Cot-1, ДНК спермы лосося, гибридизационным буфером, блокирующим агентом и денатурировали. За денатурацией следовала 30-минутная инкубация при 37°C. Затем полученную смесь добавляли к биотинилированным зондам и инкубировали в течение 12-48 ч при 60-70°C. После инкубации обогащенные библиотеки промывали, как описано ранее, и элюировали ДНК путем нагревания. Элюированные продукты амплифицировали с применением адаптерных праймеров с внешней связью. Обогащенные амплифицированные продукты объединяли эквимолярно и секвенировали на подходящей платформе.

Пример 4. Биоинформационный анализ.

Авторы настоящего изобретения разработали основанный на применении компьютера способ, который организывает наблюдаемые данные (прочтения секвенирования) в структуры, которые позволяют выбирать подмножества секвенированных прочтений из обогащенного образца, чтобы увеличить отношение сигнал/шум в образце, содержащем смесь фрагментов внеклеточной ДНК (вкДНК). Способ согласно настоящему изобретению способ группировать и/или аннотирует множество последовательностей ДНК таким образом, чтобы степень гомологии между этими последовательностями ДНК была максимальной, если они действительно происходят из одной и той же группы, и минимальной в противном случае. Затем в указанных данных идентифицируют конкретные паттерны последовательностей и их применяют для их распределения в заданные представляющие интерес области. В настоящем изобретении способ применяют для неинвазивного обнаружения хромосомных аномалий у плода, таких как анеуплоидии, микроделеции и микродупликации в образце, содержащем смесь внеклеточной фетальной ДНК (cffDNA) и материнской ДНК.

В предпочтительном варианте реализации файлы FASTQ каждого образца не выравнивают против версии эталонного генома человека. Информацию о размере и/или классе фрагмента получают с применением длины отдельных прочтений (если она больше 150 п.о.) или информации о перекрывающихся последовательностях прочтений на спаренных концах.

Идентичные прочтения удаляют, и осуществляют либо сборку de novo, либо сопоставление с заранее определенными мишенями коротких последовательностей.

Средний специалист в данной области техники оценит наличие многих легкодоступных и хорошо зарекомендовавших себя инструментов для осуществления сопоставления последовательностей.

Алгоритм представляет собой набор процедур обработки данных, математических и статистических моделей, организованных в виде серии шагов. Шаг алгоритма направлен на определение относительного числа копий представляющей интерес хромосомы и/или представляющей интерес субхромосомной области в обогащенном образце и его применяют для обнаружения полных или частичных хромосомных аномалий, таких как анеуплоидии хромосом 13, 18, 21, X, Y или любой другой хромосомы, а также синдромов микроделеции/микродупликации и других мутаций небольшого размера.

Ключевой характеристикой способа является двойное обогащение образца по фетальной (плацентарной) фракции путем выбора подмножества обогащенных фрагментов, что приводит к увеличению отношения сигнал/шум (фетальная фракция к материнской фракции, например, пропорция фрагментов плацентарного происхождения). С этой целью разработанный способ идентифицирует и применяет фрагменты, с высокой вероятностью происходящие из плаценты. На фиг. 3 показано предполагаемое распределение по размерам фрагментов, выбранных заявленным способом, по сравнению с распределением по размерам всех фрагментов показательного образца, что демонстрирует обогащение более коротких фрагментов, которые в основном происходят из фетальных (плацентарных) тканей.

В одном варианте реализации способа обогащенный образец секвенируют с применением способа секвенирования 2x75 п.о. В указанном варианте реализации секвенированные фрагменты сгруппированы по размеру на основе степени перекрытия или отсутствия перекрытия прочтений их спаренных концов. Каждый фрагмент классифицируют, соответственно, на две группы: короткие (Группа 1) и длинные (Группа 2). Затем HSNRF идентифицируют, применяя сходство последовательностей, по крайней мере, 20 наиболее удаленных от середины последовательности нуклеотидов фрагментов в Группе 1. Фрагменты из Группы 2, концы которых перекрываются с идентифицированными HSNRF, затем применяют для создания Группы 3. Затем все фрагменты из Группы 1, и все фрагменты из Группы 3 оставляют для последующего анализа. Указанные фрагменты вкДНК Группы 1 меньше по размеру и обогащены внеклеточной фетальной ДНК. Указанные фрагменты вкДНК Группы 3 больше по размеру, но также обогащены внеклеточной фетальной ДНК, поскольку их концы перекрываются с HSNRF. Фрагменты вкДНК Группы 2, концы которых не перекрываются с HSNRF, имеют более высокую вероятность происхождения от матери и исключаются из всех последующих анализов. На фиг. 3 показано распределение по размеру выбранных заявленным способом фрагментов (Группа 1 и Группа 3) по сравнению с распределением по размеру всех фрагментов в типичном образце, что показывает обогащение фетальными (плацентарными) фрагментами вкДНК.

После сопоставления всех фрагментов из Группы 1 и Группы 3 с последовательностями зонда для каждого тестируемого образца вычисляют классификационную оценку путем сравнения количества оставшихся фрагментов на целевой хромосоме, например, хромосоме 21 (фиг. 1) и/или целевой субхромо-

сомной области, например, 22q11.2 (фиг. 2) с количеством фрагментов на эталонных хромосомах и/или субхромосомных областях. В другом варианте реализации способа указанные шаги могут быть осуществлены после выравнивания секвенированных прочтений с эталонным геномом.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ обнаружения хромосомной аномалии в смешанном образце, включающий следующие этапы:

(a) получение биологического образца, указанный образец содержит смесь фрагментов внеклеточной ДНК (вкДНК),

(b) приготовление библиотеки секвенирования из указанных фрагментов вкДНК,

(c) гибридизация одного или более зондов по меньшей мере с одним или более фрагментами вкДНК, причем указанные один или более зондов охватывают по меньшей мере одну геномную область, содержащую на расстоянии менее чем 300 п.о. геномные основания, различающие два типа тканей, присутствующих в смеси вкДНК (HSNRF),

(d) выделение фрагментов вкДНК из указанной библиотеки, которые связываются с указанными зондами,

(e) секвенирование указанных фрагментов вкДНК из библиотеки, которые связываются с зондами,

(f) применение информации о размере, начале и/или конце каждого или подмножества обогащенных фрагментов вкДНК с этапов (с-е) для выбора фракции фрагментов вкДНК, гибридованных с указанным одним или более зондами,

причем этап (f) дополнительно содержит следующие этапы:

(i) отнесение фрагментов вкДНК к первому и второму подмножеству, причем указанное первое подмножество содержит фрагменты вкДНК, имеющие длину меньше или равную 145 п.о., и указанное второе подмножество содержит фрагменты вкДНК, имеющие длину больше 145 п.о.,

(ii) обнаружение по меньшей мере одной геномной области, содержащей на расстоянии менее чем 300 п.о. геномные основания, различающие два типа тканей, присутствующих в смеси вкДНК (HSNRF) с применением фрагментов вкДНК первого подмножества,

(iii) отнесение фрагментов вкДНК из второго подмножества к третьему подмножеству, причем указанное третье подмножество включает выбор фрагментов вкДНК из указанного второго подмножества, концы которых перекрываются с HSNRF, обнаруженными на этапе (ii),

(iv) объединение фрагментов вкДНК первого подмножества с фрагментами вкДНК третьего подмножества,

(v) сопоставление всех фрагментов вкДНК из первого и третьего подмножеств с последовательностями зондов, и

(vi) вычисление классификационной оценки на основании фрагментов вкДНК согласно (v) путем сравнения количества фрагментов вкДНК, оставшихся на целевой хромосоме, на которой тестируют аномалию, с количеством фрагментов вкДНК на эталонной хромосоме, причем хромосомную аномалию обнаруживают, когда классификационная оценка выше порогового значения, составляющего 1.

2. Способ по п.1, отличающийся тем, что указанные зонды представляют собой длинные молекулы ДНК и:

(i) длина каждого зонда составляет от 100 до 500 пар оснований,

(ii) каждый денатурированный зонд имеет 5'-конец и 3'-конец,

(iii) предпочтительно, каждый зонд связывается с HSNRF на расстоянии не менее 10 пар оснований как на 5'-конце, так и на 3'-конце, от областей, несущих вариации числа копий (CNV), сегментные дубликации или повторяющиеся элементы ДНК, и

(iv) GC-состав каждого зонда составляет от 19 до 80%.

3. Способ по любому из пп.1, 2, отличающийся тем, что указанный образец представляет собой смешанный образец, выбранный из группы, содержащей: (i) эмбриональную ДНК и материнскую ДНК, (ii) происходящую из опухоли ДНК и не происходящую из опухоли ДНК, (iii) ДНК патогена и ДНК хозяина и (iv) ДНК, происходящую из трансплантированного органа, и ДНК, происходящую от хозяина.

4. Способ согласно любому из пп.1-3, отличающийся тем, что указанный способ дополнительно включает этап объединения статистических тестов, выбранных из группы, содержащей, но не ограничивающейся указанным, t-критерий, двумерный непараметрический бутстреп-тест, стратифицированный пермутационный тест, непараметрический тест, ANOVA и тест на пропорции размеров фрагментов.

5. Способ по любому из пп.1-4 для применения при диагностике и/или скрининге генетической аномалии в образце.

6. Способ по п.5, отличающийся тем, что указанная генетическая аномалия выбрана из группы, содержащей, но не ограничивающейся указанным:

(a) анеуплоидии хромосом 13, 18, 21 и/или X, Y,

(b) структурные аномалии, включая, но не ограничиваясь указанным, изменения числа копий, в том числе микроделеции и микродубликации, вставки, делеции, транслокации и мутации небольшого разме-

ра, в том числе точечные мутации.

7. Способ двойного обогащения фрагментов, происходящих из плаценты, в смеси внеклеточной ДНК (вкДНК), включающий следующие этапы:

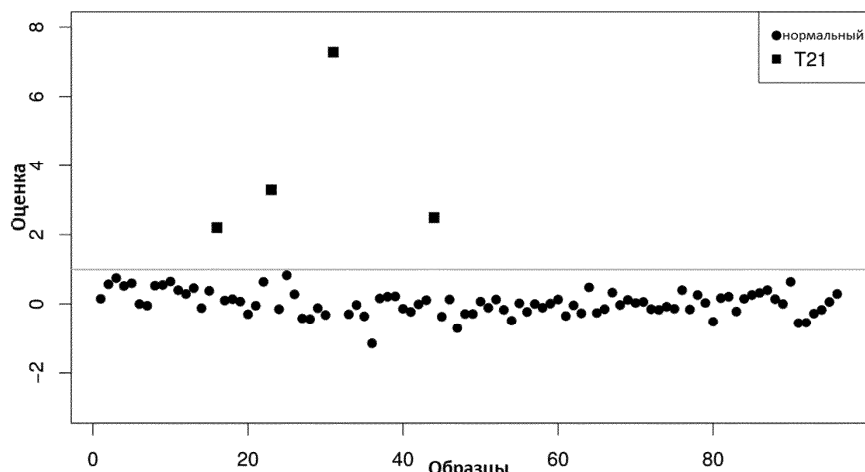
- (a) получение биологического образца, указанный образец содержит смесь фрагментов внеклеточной ДНК (вкДНК),
- (b) приготовление библиотеки секвенирования из указанных фрагментов вкДНК,
- (c) гибридизация указанной библиотеки вкДНК с множеством зондов, причем указанные зонды охватывают по меньшей мере одну геномную область, содержащую на расстоянии менее чем 300 п.о. геномные основания, различающие два типа тканей, присутствующих в смеси вкДНК (HSNRF),
- (d) выделение из указанной библиотеки фрагментов вкДНК, которые связываются с указанными зондами,
- (e) секвенирование указанных фрагментов вкДНК из библиотеки, которые связываются с зондами,
- (f) удаление повторяющихся секвенированных прочтений,
- (g) отбор коротких фрагментов вкДНК,
- (h) обнаружение указанной по меньшей мере одной HSNRF в коротких фрагментах вкДНК,
- (i) отбор длинных фрагментов вкДНК, концы которых перекрываются с указанной по меньшей мере одной HSNRF,
- (j) сопоставление выбранных фрагментов вкДНК из (g) и (i) с последовательностями зондов, и
- (k) вычисление классификационной оценки на основании фрагментов вкДНК согласно (j) путем сравнения количества фрагментов вкДНК, оставшихся на целевой хромосоме, с количеством фрагментов вкДНК на эталонной хромосоме, причем хромосомную аномалию обнаруживают, когда классификационная оценка выше порогового значения, составляющего 1.

8. Способ по п.7, отличающийся тем, что указанные короткие фрагменты вкДНК имеют длину меньше или равную 120 п.о., или 125 п.о., или 130 п.о., или 135 п.о., или 140 п.о., или 145 п.о., или 150 п.о., или 155 п.о., и указанные длинные фрагменты вкДНК имеют длину более 120 п.о., или 125 п.о., или 130 п.о., или 135 п.о., или 140 п.о., или 145 п.о., или 150 п.о., или 155 п.о., соответственно.

9. Способ по п.7, отличающийся тем, что указанные фрагменты вкДНК этапа (j) применяют для обнаружения анеуплоидии у плода в неинвазивном пренатальном диагностическом тесте.

10. Способ по п.7, отличающийся тем, что указанные фрагменты вкДНК этапа (j) применяют для обнаружения микроделеции/микродупликации у плода в неинвазивном пренатальном диагностическом тесте.

11. Способ по п.7, отличающийся тем, что указанные фрагменты вкДНК этапа (j) применяют для обнаружения у плода вставок, делеций, транслокаций и мутаций небольшого размера, включая точечные мутации.



Фиг. 1

