(19) Евразийское патентное ведомство

(21) 202290342 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.09.16

(22) Дата подачи заявки
2020.07.30

(51) Int. Cl. ***C12N 15/09*** (2006.01)
C12N 15/11 (2006.01)
C12Q 1/68 (2018.01)
C12Q 1/6806 (2018.01)
C12Q 1/6844 (2018.01)
C12Q 1/686 (2018.01)

(54) АНАЛИЗ ОТДЕЛЬНЫХ КЛЕТОК

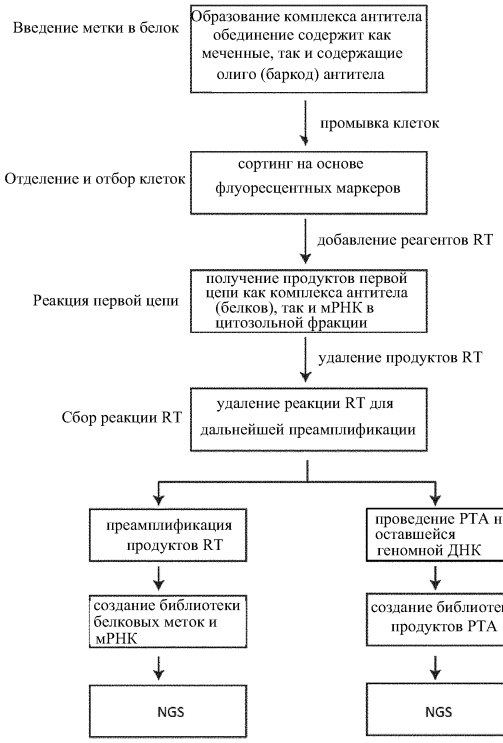
(31) **62/881,183**
(32) **2019.07.31**
(33) **US**
(86) **PCT/US2020/044338**
(87) **WO 2021/022085 2021.02.04**
(88) **2021.03.11**

(71) Заявитель:
БАЙОСКРИБ ДЖЕНОМИКС, ИНК. (US)

(72) Изобретатель:
Гавад Чарльз, Уэст Джэй А.А. (US)

(74) Представитель:
Строкова О.В., Гизатуллин Ш.Ф., Гизатуллина Е.М., Прищепный С.В., Парамонова К.В., Джермакян Р.В., Христофоров А.А., Угрюмов В.М., Костюшенкова М.Ю. (RU)

(57) Изобретение относится к композициям и способам для точных и масштабируемых способов амплификации нуклеиновых кислот посредством амплификации, управляемой первичной матрицей, (PTA) и секвенирования, и их применению для мутационного анализа в исследованиях, диагностике и лечении. Кроме того, настоящее изобретение относится к мультиомическим способам для параллельного анализа ДНК, РНК и/или белков из отдельных клеток.



202290342 A1

202290342 A1

АНАЛИЗ ОТДЕЛЬНЫХ КЛЕТОК

Описание

Ссылка на родственную заявку

[0001] Согласно настоящей заявке испрашивается приоритет в соответствии с предварительной заявкой на выдачу патента США № 62/881183, поданной 31 июля 2019 года, содержание которой включено в настоящий документ посредством ссылки.

Предшествующий уровень техники настоящего изобретения

[0002] Способы исследования, в которых используется амплификация нуклеиновых кислот, например, секвенирование нового поколения, предоставляют большой объем информации о сложных образцах, геномах и других источниках нуклеиновых кислот. В некоторых случаях эти образцы получают в небольших количествах из отдельных клеток. Существует потребность в высокоточных, масштабируемых и эффективных способах амплификации и секвенирования нуклеиновых кислот для исследований, диагностики и лечения с использованием небольших образцов, особенно в способах одновременного анализа РНК, ДНК и белков.

Сущность настоящего изобретения

[0003] Настоящее изобретение относится к способам мультиомического анализа отдельных клеток, предусматривающим: (a) выделение отдельной клетки из популяции клеток, (b) секвенирование библиотеки кДНК, содержащей полинуклеотиды, амплифицированные из транскриптов мРНК из отдельной клетки, и (c) секвенирование генома отдельной клетки, где секвенирование генома предусматривает: (i) контакт генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по меньшей мере один терминаторный нуклеотид, который останавливает репликацию нуклеиновых кислот полимеразой, и (ii) амплификацию по меньшей мере части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи, (iii) лигирование молекул, полученных на стадии (ii), с адапторами, с получением таким образом библиотеки геномной ДНК, и (iv) секвенирование библиотеки геномной ДНК. Кроме того, настоящее изобретение относится к способам, в которых транскрипты мРНК содержат полиаденилировнные транскрипты мРНК. Кроме того, настоящее изобретение относится к способам, в которых транскрипты

мРНК не содержат полиаденилировнные транскрипты мРНК. Кроме того, настоящее изобретение относится к способам, в которых секвенирование библиотеки кДНК предусматривает амплификацию транскриптов мРНК с переключающими матрицу праймерами. Кроме того, настоящее изобретение относится к способам, в которых по меньшей мере некоторые из полинуклеотидов библиотеки кДНК содержат баркод. Кроме того, настоящее изобретение относится к способам, в которых баркод содержит баркод клетки или баркод образца. Кроме того, настоящее изобретение относится к способам, в которых библиотеку кДНК и библиотеку геномной ДНК объединяют перед секвенированием. Кроме того, настоящее изобретение относится к способам, в которых отдельная клетка представляет собой первичную клетку. Кроме того, настоящее изобретение относится к способам, в которых отдельные клетки происходят из печени, кожи, почки, крови или легкого. Кроме того, настоящее изобретение относится к способам, в которых отдельную клетку выделяют посредством проточной цитометрии. Кроме того, настоящее изобретение относится к способам, в которых способ дополнительно предусматривает удаление по меньшей мере одного терминаторного нуклеотида из терминированных продуктов амплификации. Кроме того, настоящее изобретение относится к способам, в которых множество терминированных продуктов амплификации содержат в среднем 1000-2000 оснований в длину. Кроме того, настоящее изобретение относится к способам, в которых множество терминированных продуктов амплификации составляют 250-1500 оснований в длину. Кроме того, настоящее изобретение относится к способам, в которых множество терминированных продуктов амплификации содержат по меньшей мере 97% генома отдельной клетки. Кроме того, настоящее изобретение относится к способам, в которых по меньшей мере некоторые из продуктов амплификации содержат баркод клетки или баркод образца. Кроме того, настоящее изобретение относится к способам, в которых секвенирование библиотеки кДНК предусматривает цитозольный лизис отдельной клетки и обратную транскрипцию. Кроме того, настоящее изобретение относится к способам, в которых транскрипты мРНК амплифицируют посредством обратной транскрипции с переключением матрицы. Кроме того, настоящее изобретение относится к способам, в которых библиотека кДНК содержит по меньшей мере 10000 генов. Кроме того, настоящее изобретение относится к способам, в которых секвенирование генома отдельной клетки дополнительно предусматривает лизис ядра отдельной клетки. Кроме того, настоящее изобретение относится к способам, в которых способ дополнительно предусматривает дополнительную стадию амплификации с применением ПЦР. Кроме того, настоящее изобретение относится к способам, в которых по меньшей мере одна мутация идентифицирована в геноме клетки, где мутация отлична от

соответствующего положения в эталонной последовательности. Кроме того, настоящее изобретение относится к способам, где по меньшей мере одна мутация происходит в менее 1% популяции клеток. Кроме того, настоящее изобретение относится к способам, где по меньшей мере одна мутация происходит в не более 0,1% популяции клеток. Кроме того, настоящее изобретение относится к способам, где по меньшей мере одна мутация происходит в не более 0,001% популяции клеток. Кроме того, настоящее изобретение относится к способам, где по меньшей мере одна мутация происходит в не более 1% последовательностей продукта амплификации. Кроме того, настоящее изобретение относится к способам, где по меньшей мере одна мутация происходит в не более 0,1% последовательностей продукта амплификации. Кроме того, настоящее изобретение относится к способам, где по меньшей мере одна мутация происходит в не более 0,001% последовательностей продукта амплификации.

[0004] Настоящее изобретение относится к способам мультиомического анализа отдельных клеток, предусматривающим: (a) выделение отдельной клетки из популяции клеток, (b) идентификацию по меньшей мере одного белка на поверхности отдельной клетки и (c) секвенирование генома отдельной клетки, где секвенирование генома предусматривает: (i) контакт генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по меньшей мере один терминаторный нуклеотид, который останавливает репликацию нуклеиновых кислот полимеразой, (ii) амплификацию по меньшей мере части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи, (iii) лигирование молекул, полученных на стадии (ii), с адапторами, с получением таким образом библиотеки геномной ДНК, и (iv) секвенирование библиотеки геномной ДНК. Кроме того, настоящее изобретение относится к способам, в которых идентификация по меньшей мере одного белка на поверхности клетки предусматривает контакт клетки с меченным антителом, которое связывается с по меньшей мере одним белком. Кроме того, настоящее изобретение относится к способам, в которых меченное антитело содержит по меньшей мере одну флуоресцентную метку или массовую метку. Кроме того, настоящее изобретение относится к способам, в которых меченное антитело содержит по меньшей мере один баркод нуклеиновой кислоты.

[0005] Настоящее изобретение относится к способам мультиомического анализа отдельных клеток, предусматривающему: (a) выделение отдельной клетки из популяции клеток, (b) секвенирование генома отдельной клетки, где секвенирование генома клетки предусматривает: (i) расщепление генома чувствительным к метилированию ферментом

рестрикции с получением фрагментов генома, (ii) контакт по меньшей мере некоторых фрагментов генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по меньшей мере один терминаторный нуклеотид, который останавливает репликацию нуклеиновых кислот полимеразой, (iii) амплификацию по меньшей мере некоторой части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи, (iv) амплификацию по меньшей мере некоторых фрагментов генома посредством специфической к метилированию ПЦР, (v) лигирование молекул, полученных на стадиях (iii и iv), с адапторами, с получением таким образом библиотеки геномной ДНК и библиотеки ДНК-метиломы, и (vi) секвенирование библиотеки геномной ДНК и библиотеки метиломы.

Включение посредством ссылки

[0006] Все публикации, патенты и патентные заявки, упомянутые в настоящем описании, включены в настоящий документ посредством ссылки в той же степени, как если бы каждая отдельная публикация, патент или патентная заявка были конкретно и индивидуально указаны как включенные посредством ссылки.

Краткое описание чертежей

[0007] Новые признаки настоящего изобретения подробно изложены в прилагаемой формуле изобретения. Лучшее понимание признаков и преимуществ настоящего изобретения может быть получено при обращении к следующему подробному описанию, в котором раскрыты иллюстративные варианты осуществления, в которых используются принципы настоящего изобретения, и приложенным чертежам, где:

[0008] На **фиг. 1A** проиллюстрирована общая схема, раскрывающая в общем анализ выделения белков, ДНК и РНК из отдельных клеток.

[0009] На **фиг. 1B** проиллюстрирована схема анализа выделения белков, ДНК и РНК из отдельных клеток с использованием разделения образцов для сведения к минимуму взаимного загрязнения.

[0010] На **фиг. 1C** проиллюстрирована схема анализа выделения белков, ДНК и РНК из отдельных клеток с использованием предварительной амплификации в одной пробирке.

[0011] На **фиг. 1D** проиллюстрирована схема анализа выделения белков, ДНК и РНК из отдельных клеток с использованием предварительной амплификации в одной пробирке с терминаторами для уменьшения размера ампликона.

[0012] На **фиг. 1E** проиллюстрирована схема анализа выделения белков, ДНК и РНК из отдельных клеток с использованием совместной амплификации.

[0013] На **фиг. 1F** проиллюстрирована схема обработки данных, объединяющую данные эксперимента с белком/ДНК/РНК отдельной клетки, описанного в настоящем документе.

[0014] На **фиг. 1G** проиллюстрировано сравнение способа MDA и способа PTA-необратимой терминации, поскольку они относятся к распространению мутаций. Способ PTA приводит к увеличению числа прямых копий исходной ДНК-матрицы.

[0015] На **фиг. 2A** проиллюстрированы стадии способа, выполняемые после амплификации, которые включают удаление терминатора, репарацию концов и полиаденилирование перед лигированием адаптера. Библиотека объединенных клеток может быть затем подвергнута опосредованному гибридизацией обогащению для всех экзонов или других конкретных представляющих интерес областей перед секвенированием. Клетка происхождения каждого прочтения идентифицируется баркодом клетки (показан в виде зеленой и синей последовательностей).

[0016] На **фиг. 2B** (GC) показано сравнение содержания GC секвенированных оснований для экспериментов посредством MDA и PTA.

[0017] На **фиг. 2C** показаны оценки качества карты (e) (mapQ) при картировании на геноме человека (p_mapped) после того, как отдельные клетки подверглись PTA или MDA.

[0018] На **фиг. 2D** приведен процент прочтений при картировании на геноме человека (p_mapped) после того, как отдельные клетки подверглись PTA или MDA.

[0019] На **фиг. 2E** (ПЦР) показано сравнение процента прочтений, которые являются дубликатами ПЦР, для 20 миллионов прочтений подвыборки после того, как отдельные клетки подверглись MDA и PTA.

[0020] На **фиг. 2F** показана схема амплификации RT отдельных клеток для использования с PTA.

[0021] На **фиг. 2G** показано создание библиотеки из кДНК, полученных с помощью RT.

[0022] На **фиг. 3A** показаны оценки качества карты (c) (mapQ) при картировании на геноме человека (p_mapped) после того, как отдельные клетки подверглись PTA с обратимыми или необратимыми терминаторами.

[0023] На **фиг. 3B** приведен процент прочтений при картировании на геноме человека (p_mapped) после того, как отдельные клетки подверглись PTA с обратимыми или необратимыми терминаторами.

[0024] На **фиг. 3C** показана серия блочных диаграмм, на которых представлены выровненные прочтения для среднего процента прочтений, перекрывающихся с элементами Alu, с использованием различных способов. PTA характеризуется наибольшим количеством прочтений, выровненных с геномом.

[0025] На **фиг. 3D** показана серия блочных диаграмм, на которых представлены дупликации ПЦР для среднего процента прочтений, перекрывающихся с элементами Alu, с использованием различных способов.

[0026] На **фиг. 3E** показана серия блочных диаграмм, на которых представлено содержание GC прочтений для среднего процента прочтений, перекрывающихся с элементами Alu, с использованием различных способов.

[0027] На **фиг. 3F** показана серия блочных диаграмм, на которых представлено качество картирования среднего процента прочтений, перекрывающихся с элементами Alu, с использованием различных способов. PTA характеризуется самым высоким качеством картирования исследованных способов.

[0028] На **фиг. 3G** показано сравнение широты охвата митохондриального генома SC с помощью различных способов WGA при фиксированной глубине секвенирования, составляющей 7,5X.

[0029] На **фиг. 4A** показана средняя глубина охвата 10 т.п.н. окон по хромосоме 1 после выбора клетки высококачественной MDA (представляющей ~ 50% клеток) по сравнению с клеткой, амплифицированной PTA со случайным праймером после субдискретизации каждой клетки до 40 миллионов парных прочтений. На чертеже показано, что MDA имеет меньшую однородность с гораздо большим количеством окон, которые имеют больше (прямоугольник A) или меньше (прямоугольник C), чем удвоенная средняя глубина охвата. Отсутствует охват как MDA, так и PTA на центромере из-за высокого содержания GC и низкого качества картирования повторяющихся областей (прямоугольник B).

[0030] На **фиг. 4B** показаны графики зависимости охвата секвенированием от положения в геноме для способов MDA и PTA (вверху). Нижние блочные диаграммы показывают частоты аллелей для способов MDA и PTA по сравнению с совокупным образцом.

[0031] На **фиг. 5A** показан график доли охваченного генома в зависимости от числа прочтений генома для оценки охвата при увеличении глубины секвенирования для различных способов. Способ PTA приближается к двум совокупным образцам на каждой глубине, что является улучшением по сравнению с другими исследованными способами.

[0032] На **фиг. 5B** показан график зависимости коэффициента вариации охвата генома от количества прочтений для оценки однородности охвата. Было обнаружено, что способ PTA характеризуется наивысшей однородностью среди исследованных способов.

[0033] На **фиг. 5C** показан график Лоренца кумулятивной доли общего числа прочтений по сравнению с кумулятивной долей генома. Было обнаружено, что способ PTA характеризуется наивысшей однородностью среди исследованных способов.

[0034] На **фиг. 5D** показана серия блочных диаграмм рассчитанных коэффициентов Джини для каждого из исследованных способов, чтобы оценить отличие каждой реакции амплификации от идеальной однородности. Было обнаружено, что способ PTA воспроизводимо более однороден, чем другие исследованные способы.

[0035] На **фиг. 5E** показан график доли обнаруженных совокупных вариантов в зависимости от количества прочтений. Частоту обнаружения вариантов для каждого из способов сравнивали с соответствующим совокупным образцом при увеличении глубины секвенирования. Для оценки чувствительности рассчитывали процент вариантов, обнаруженных в соответствующих совокупных образцах, которые были подвергнуты субдискретизации до 650 миллионов прочтений, обнаруженных в каждой клетке на каждой глубине секвенирования (фиг. 3A). Улучшенный охват и однородность PTA привели к обнаружению на 30% больше вариантов по сравнению со способом Q-MDA, который являлся следующим по чувствительности способом.

[0036] На **фиг. 5F** показана серия блочных диаграмм среднего процента прочтений, перекрывающихся с элементами Alu. Способ PTA значительно уменьшил аллельный перекос в этих гетерозиготных сайтах. Способ PTA более равномерно амплифицирует два аллеля в одной и той же клетке по сравнению с другими исследованными способами.

[0037] На **фиг. 5G** показан график зависимости специфичности обнаружений вариантов от числа прочтений для оценки специфичности обнаружений мутаций. Варианты, обнаруженные с помощью различных способов, которые не были обнаружены в совокупных образцах, считались ложноположительными. Способ PTA привел к наименьшему количеству ложноположительных обнаружений (наивысшая специфичность) исследованных способов.

[0038] На **фиг. 5H** показана доля ложноположительных замен оснований для каждого типа замены оснований с помощью различных способов. Без ограничения какой-либо теорией такие паттерны могут зависеть от полимеразы.

[0039] На **фиг. 5I** показана серия блочных диаграмм среднего процента прочтений, перекрывающихся с элементами Alu для ложноположительных обнаружений вариантов.

Способ PTA привел к самым низким частотам аллелей для ложноположительных обнаружений вариантов.

[0040] На **фиг. 6** (часть A) показаны шарики с олигонуклеотидами, прикрепленными с помощью расщепляемого линкера, уникальным клеточным баркодом и случайным праймером. В части B показана отдельная клетка и шарик, инкапсулированные в одной и той же капле, с последующим лизисом клетки и расщеплением праймера. Затем капля может быть слита с другой каплей, содержащей смесь для амплификации PTA. В части C показано, что после амплификации капли разбивают, а ампликоны из всех клеток объединяют. Затем протокол в соответствии с настоящим раскрытием используют для удаления терминатора, репарации концов и полиаденилирования перед лигированием адаптера. Затем библиотека объединенных клеток подвергается опосредованному гибридизацией обогащению представляющими интерес экзонами перед секвенированием. Затем клетку происхождения каждого прочтения идентифицируют с помощью баркода клетки.

[0041] На **фиг. 7A** проиллюстрирована схема мультиомического (или полиомного) анализа отдельных клеток с использованием PTA. Стадия A: клетки контактируют с антителами, содержащими флуоресцентные метки и олигонуклеотидные метки баркода. Стадия B: клетки сортируют на основе флуоресцентных маркеров. Стадия C: пробирки покрывают антителами, которые связываются с ядром, клетки лизируют, цитозольная мРНК подвергается обратной транскрипции, в то время как интактное ядро связывается со стенкой пробирки.

[0042] На **фиг. 7B** показана схема мультиомического анализа одиночных клеток с использованием PTA, продолжение стадии C на **фиг. 7A**. Стадия D: после обратной транскрипции фракцию RT удаляют для анализа последовательности. Стадия E: ядро лизируют, и способ PTA применяют к геномной ДНК. Стадия F: PTA приводит к пулу коротких фрагментов кДНК с приблизительно 1000-кратной амплификацией.

[0043] На **фиг. 8A** показаны праймеры, используемые для обратной транскрипции и преамплификации в схеме мультиомического анализа ДНК/РНК отдельных клеток.

[0044] На **фиг. 8B** проиллюстрирована схема обратной транскрипции и предварительной амплификации для схемы мультиомического анализа ДНК/РНК отдельных клеток. Применяют праймеры согласно фиг. 8A.

[0045] На **фиг. 9A** показан график скоростей роста линий родительских клеток, обработанных 2 нМ квизартиниба в течение трех недель, с получением линий клеток AML, которые интенсивно растут в присутствии ингибитора FLT3. Затем отдельные устойчивые

и родительские клетки (обогащенные FACS) анализировали посредством секвенирования РНК и низкочастотного секвенирования ДНК.

[0046] На **фиг. 9B** проиллюстрирована экспрессия РНК как из родительских, так и из резистентных культур, демонстрирующая способность создавать пулы кДНК (C) с использованием химии секвенирования РНК в одном сосуде, а гены, экспрессируемые в этих клетках, создавали различные паттерны, которые позволяют визуализировать клеточные популяции по экспрессии генов с в среднем ~10K генов, обнаруживаемых на клетку. В отдельной схеме геномы отдельных клеток амплифицировали с использованием метода PTA.

[0047] На **фиг. 9C** проиллюстрированы нормализованные профили экспрессии генов относительно контрольного эксперимента с последовательностью РНК только.

[0048] На **фиг. 9D** проиллюстрирован график количества амплифицированной ДНК с помощью PTA по сравнению с различными протоколами. Транскрипты, получаемые на стадии RT (R), неэффективно амплифицируются реакцией PTA по сравнению с ДНК, а ДНК в отдельных клетках эффективно амплифицируется с использованием комбинированного протокола (SC1-SC8) по сравнению со стандартными геномами, амплифицированными посредством PTA из отдельных клеток (D, RD). NTC = без контроля матрицей, R = стадия RT, D = стадия ДНК PTA, RD = двойная RT/PTA.

[0049] На **фиг. 10A** показано количество митохондриальных хромосом (%) для двух разных протоколов (двойное секвенирование РНК/PTA, стандартное секвенирование РНК) с использованием протокола низкочастотного секвенирования (~ 5 миллионов прочтений/клетка).

[0050] Оцениваемый размер генома превышал 3 миллиарда оснований.

[0051] На **фиг. 10B** показан процент дубликатов для двух разных протоколов (двойное секвенирование РНК/PTA, стандартное секвенирование РНК) с использованием протокола низкочастотного секвенирования (~ 5 миллионов прочтений/клетка).

[0052] На **фиг. 10C** проиллюстрирован оцененный размер генома для двух разных протоколов (двойное секвенирование РНК/PTA, стандартное секвенирование РНК) с использованием протокола низкочастотного секвенирования (~ 5 миллионов прочтений/клетка).

[0053] На **фиг. 10D** показано назначение признаков 3 наборов данных scRNAseq из клеток molm13 с использованием протокола двойного секвенирования РНК/PTA.

[0054] На **фиг. 10E** показан график нормализованных профилей экспрессии для клеточной линии Sum159, полученных с использованием стандартного протокола секвенирования РНК. P = родительские клетки. R = резистентные клетки.

[0055] На **фиг. 10F** показан график нормализованных профилей экспрессии для клеточной линии Sum159, полученных с использованием протокола двойного секвенирования РНК/РТА. P = родительские клетки. R = резистентные клетки.

[0056] На **фиг. 11A** показаны результаты глубокого секвенирования 7 родительских и 5 резистентных клеток molm13, выполненного на глубине приблизительно 25× (K). Прочтения выравнивали с Hg38 с использованием bwa mem. Контроль качества и вызов SNV осуществляли с использованием лучших практических выполнений GATK4. SNV рассматривали только в том случае, если они были ограничены по меньшей мере 2 резистентными клетками, ни в одной родительской клетке не вызывали альтернативные аллели, и по меньшей мере 6 родительских клеток были генотипированы. Все клетки имели по меньшей мере 96% генома, покрытого при 1-кратном покрытии, и по меньшей мере 76% при 10-кратном покрытии. На вставке показано, что известная индель Flt3 в клетках molm13 обнаруживается во всех клетках (4 показаны для ясности).

[0057] На **фиг. 11B** показана тепловая карта профилей экспрессии генов, включая сверхэкспрессированный ген GAS6, который является известным механизмом устойчивости к квизартинибу. Gas6 является лигандом для AXL, который представляет собой клинически значимый механизм резистентности у пациентов с рецидивом, у которых лечение квизартинибом оказалось неэффективным.

[0058] На **фиг. 12A** проиллюстрирован график доли покрытых экзонов в общей массе по сравнению с образцами отдельных клеток.

[0059] На **фиг. 12B** проиллюстрирован график доли экзонов без покрытия в массе по сравнению с образцами отдельных клеток.

[0060] На **фиг. 12C** проиллюстрирован график процента выбранных оснований в массе по сравнению с образцами отдельных клеток.

[0061] На **фиг. 12D** проиллюстрирован график доли оснований, покрытых при 20X массе по сравнению с образцами отдельных клеток.

[0062] На **фиг. 13A** проиллюстрирован график расположения картированных оснований прочтений в геноме, стратифицированных по обработке и заштрихованных по типу образца.

[0063] На **фиг. 13B** проиллюстрирован график зависимости интенсивности образца от размера захваченной вставки.

[0064] На **фиг. 14A** проиллюстрирован график процента дубликатов по сравнению с процентом выбранных оснований для эксперимента 12-plex.

[0065] На **фиг. 14B** проиллюстрирован график зависимости числа целевых оснований от уровня охвата.

Подробное описание настоящего изобретения

[0066] Существует необходимость в разработке новых масштабируемых, точных и эффективных способов амплификации нуклеиновых кислот (включая амплификацию генома отдельных клеток и множества клеток) и секвенирования, которые позволили бы преодолеть ограничения в существующих способах посредством увеличения представления последовательности, единообразия и точности в воспроизводимом способе. Настоящее изобретение относится к композициям и способам для обеспечения точной и масштабируемой амплификации, управляемой первичной матрицей (PTA), и секвенирования. Кроме того, настоящее изобретение относится к способа мультиомического анализа, включая анализ белков, ДНК и РНК из отдельных клеток и соответствующих посттранскрипционных или посттрансляционных модификаций в сочетании с PTA. Такие способы и композиции способствуют высокоточной амплификации целевых (или «матричных») нуклеиновых кислот, что повышает точность и чувствительность последующих применений, таких как секвенирование следующего поколения.

[0067] Определения

[0068] Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют то же значение, которое обычно понимается специалистом в настоящей области техники, к которой принадлежат настоящие изобретения.

[0069] В этом раскрытии числовые признаки представлены в формате диапазона. Следует понимать, что описание в формате диапазона дано просто для удобства и краткости и не должно толковаться как жесткое ограничение объема любых вариантов осуществления. Соответственно, следует считать, что описание диапазона конкретно раскрывает все возможные поддиапазоны, а также отдельные числовые значения в пределах этого диапазона с точностью до десятых долей единицы нижнего предела, если контекст явно не диктует иное. Например, описание диапазона, такого как от 1 до 6, следует рассматривать как конкретно раскрывающие поддиапазоны, такие как от 1 до 3, от 1 до 4, от 1 до 5, от 2 до 4, от 2 до 6, от 3 до 6 и т.д., а также отдельные значения в этом диапазоне, например, 1,1, 2, 2,3, 5 и 5,9. Это применимо независимо от широты диапазона. Верхний и нижний пределы этих промежуточных диапазонов могут независимо включаться в меньшие диапазоны, а также включены в настоящее изобретение, с учетом любого специально исключенного предела в указанном диапазоне. Если указанный диапазон включает в себя один или оба предела, диапазоны, исключающие один или оба из этих

включенных пределов, также включены в настоящее изобретение, если контекст явно не диктует иное.

[0070] Терминология, используемая в настоящем документе, предназначена только для описания конкретных вариантов осуществления и не предназначена для ограничения какого-либо варианта осуществления. Подразумевается, что используемые в настоящем документе формы единственного числа включают в себя также формы множественного числа, если контекст явно не указывает иное. Кроме того, следует понимать, что термины «содержит» и/или «содержащий» при использовании в настоящем описании определяют наличие заявленных признаков, целых чисел, стадий, операций, элементов и/или компонентов, но не исключают наличие или добавление одного или несколько других признаков, целых чисел, стадий, операций, элементов, компонентов и/или их групп. Используемый в настоящем документе термин «и/или» включает в себя любые и все комбинации одного или несколько связанных перечисленных элементов.

[0071] Если специально не указано или не очевидно из контекста, используемый в настоящем документе термин «приблизительно» в отношении числа или диапазона чисел понимают как означающий указанное число и числа +/- 10% от них, или на 10% ниже нижнего указанного предела и на 10% выше указанного верхнего предела для значений, перечисленных для диапазона.

[0072] Используемые в настоящем документе термины «субъект», «пациент» или «индивидуум» относятся к животным, включая млекопитающих, таких как, например, люди, ветеринарные животные (например, кошки, собаки, коровы, лошади, овцы, свиньи и т.д.) и используемые в качестве моделей заболеваний экспериментальные животные (например, мыши, крысы). Согласно настоящему изобретению можно использовать общепринятые техники молекулярной биологии, микробиологии и рекомбинантной ДНК в пределах квалификации в настоящей области техники. Такие техники полностью описаны в литературе. См., например, Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual,* Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (в настоящем документе «Sambrook *et al.,* 1989»), *DNA Cloning: A practical Approach,* Volumes I and II (D.N. Glover ed. 1985), *Oligonucleotide Synthesis* (MJ. Gait ed. 1984), *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. (1985», *Transcription and Translation* (B.D. Hames & S.J. Higgins, eds. (1984», *Animal Cell Culture* (R.I. Freshney, ed. (1986», *Immobilized Cells and Enzymes* (IRL Press, (1986», B. Perbal, *A practical Guide To Molecular Cloning* (1984), F.M. Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology,* John Wiley & Sons, Inc. (1994), среди прочих.

[0073] Термин «нуклеиновая кислота» охватывает как многоцепочечные, так и одноцепочечные молекулы. В двухцепочечных или трехцепочечных нуклеиновых кислотах цепи нуклеиновой кислоты не обязательно должны быть одинаковой протяженности (т.е. двухцепочечная нуклеиновая кислота не обязательно должна быть двухцепочечной по всей длине обеих цепей). Описанные в настоящем документе матрицы нуклеиновых кислот могут характеризоваться любым размером в зависимости от образца (от небольших фрагментов свободно-циркулирующей ДНК до полных геномов), включая без ограничения 50-300 оснований, 100-2000 оснований, 100-750 оснований, 170-500 оснований, 100-5000 оснований, 50-10000 оснований или 50-2000 оснований в длину. В некоторых случаях матрицы имеют длину по меньшей мере 50, 100, 200, 500, 1000, 2000, 5000, 10000, 20000 50000, 100000, 200000, 500000, 1000000 или больше чем 1000000 оснований. Описанные в настоящем документе способы обеспечивают амплификацию нуклеиновых кислот, таких как матрицы нуклеиновых кислот. Описанные в настоящем документе способы дополнительно обеспечивают получение выделенных и по меньшей мере частично очищенных нуклеиновых кислот и библиотек нуклеиновых кислот. В некоторых случаях описанные в настоящем документе способы обеспечивают экстрагированные нуклеиновые кислоты (например, экстрагированные из тканей, клеток или среды). Нуклеиновые кислоты включают без ограничения содержащие ДНК, РНК, кольцевую РНК, мтДНК (митохондриальную ДНК), сцДНК (свободно-циркулирующую ДНК), сцРНК (свободно-циркулирующую РНК), киРНК (короткую интерферирующую РНК), сцфДНК (свободно-циркулирующую фетальную ДНК), мРНК, тРНК, рРНК, миРНК (микроРНК), синтетические полинуклеотиды, аналоги полинуклеотидов, любую другую нуклеиновую кислоту, соответствующую описанию, или любые их комбинации. Длина полинуклеотидов, если она указана, описывается как количество оснований и сокращается, например, н. (нуклеотиды), п.н. (пары нуклеотидных оснований), т.п.н. (тысячи пар нуклеотидов) или млрд.п.н. (миллиарды пар нуклеотидов).

[0074] Используемый в настоящем документе термин «капля» относится к объему жидкости на капельном приводе. В некоторых случаях, например, капли могут являться водными или неводными или могут являться смесями или эмульсиями, включающими водные и неводные компоненты. В отношении неограничивающих примеров капельных жидкостей, которые могут быть подвергнуты операциям с каплями, см., например, международную патентную публикацию № W02007/120241. Любую подходящую систему для формирования капель и управления ими можно использовать в вариантах осуществления, представленных в настоящем документе. Например, в некоторых случаях используют капельный привод. В отношении неограничивающих примеров капельных

приводов, которые можно использовать см., например, патент США № 6911132, 6977033, 6773566, 6565727, 7163612, 7052244, 7328979, 7547380, 7641779, публикации заявок на выдачу патента США №№ US20060194331, US20030205632, US20060164490, US20070023292, US20060039823, US20080124252, US20090283407, US20090192044, US20050179746, US20090321262, US20100096266, US20110048951, международную патентную публикацию № W02007/120241. В некоторых случаях шарики предусмотрены в капле, в зазоре для операций с каплями или на поверхности для операций с каплями. В некоторых случаях шарики предусмотрены в резервуаре, который является внешним по отношению к зазору для операций с каплями или расположен отдельно от поверхности для операций с каплями, и резервуар может быть связан с проточным каналом, который позволяет капле, включая в себя шарики, попадать в зазор для операций с каплями или в контакт с поверхностью для операций с каплями. Неограничивающие примеры техник капельных приводов для иммобилизации магниточувствительных шариков и/или немагнитно-чувствительных шариков и/или проведения протоколов операций с каплями с использованием шариков описаны в публикации на выдачу патента США № US20080053205, международных патентных публикациях W02008/098236, WO2008/134153, W02008/116221, W02007/120241. Характеристики шариков можно использовать в вариантах осуществления мультиплексирования способов, описанных в настоящем документе. Примеры шариков, обладающих характеристиками, подходящими для мультиплексирования, а также способы обнаружения и анализа сигналов, излучаемых такими шариками, можно найти в публикации на выдачу патента США № US20080305481, US20080151240, US20070207513, US20070064990, US20060159962, US20050277197, US20050118574.

[0075] Праймеры и/или переключающие матрицу олигонуклеотиды также могут быть прикреплены к твердой подложке для облегчения обратной транскрипции и переключения матрицы полинуклеотидов мРНК. При таком расположении часть реакции RT или переключения матрицы происходит в объеме раствора устройства, а вторая стадия реакции происходит вблизи поверхности. В других схемах праймер переключающего матрицу олигонуклеотида высвобождается из твердой подложки, чтобы вся реакция могла происходить над поверхностью в растворе. При полиомном подходе праймеры для многостадийной реакции в некоторых случаях прикрепляют к твердой подложке или комбинируют с шариками для создания комбинаций многостадийных праймеров.

[0076] Определенные микрожидкостные устройства также поддерживают полиомные подходы. Например, устройства, изготовленные на основе PDMS, часто имеют смежные камеры для каждой стадии реакции. Такие многокамерные устройства часто

разделяют с помощью конструкции микроклапана, которым можно управлять с помощью давления воздуха или текучей среды, такой как вода или инертный углеводород (т.е. фторинертный). При мультиомическом подходе каждую стадию реакции можно изолировать и проводить отдельно. По завершении определенной стадии клапан между соседней камерой может быть высвобожден на подложки для последующей реакции, которая может быть добавлена серийным образом. Результатом является возможность модулировать последовательный набор реакций, такой как мультиомический набор (белок/РНК/ДНК/эпигеномный) реакций с использованием отдельной клетки в качестве исходного матричного материала. Различные микрофлюидные платформы могут использоваться для анализа отдельных клеток. С клетками в некоторых случаях работают с помощью гидродинамики (капельная микрогидродинамика, инерционная микрогидродинамика, вортексирование, микроклапаны, микроструктуры (например, микролунки, микроловушки)), электрических способов (диэлектрофорез (DEP), электроосмос), оптических способов (оптический пинцет, оптически индуцированный диэлектрофорез (ODEP), оптико-термокапиллярные), акустических способов или магнитных способов. В некоторых случаях микрофлюидная платформа содержит микролунки. В некоторых случаях микрофлюидная платформа включает устройство на основе PDMS (полидиметилсилоксан). Неограничивающими примерами платформ для анализа отдельных клеток, совместимых с описанными в настоящем документе способами, являются: ddSEQ Single-Cell Isolator, (Bio-Rad, Hercules, CA, USA, and Illumina, San Diego, CA, USA)), Chromium (10x Genomics, Pleasanton, CA, USA)), Rhapsody Single-Cell Analysis System (BD, Franklin Lakes, NJ, USA), Tapestri Platform (MissionBio, San Francisco, CA, USA)), Nadia Innovate (Dolomite Bio, Royston, UK), C1 and Polaris (Fluidigm, South San Francisco, CA, USA), ICELL8 Single-Cell System (Takara), MSND (Wafergen), Puncher platform (Vycap), CellRaft AIR System (CellMicrosystems), DEPArray NxT and DEPArray System (Menarini Silicon Biosystems), AVISO CellCelector (ALS) и InDrop System (1CellBio), и TrapTx (Celldom).

[0077] Используемый в настоящем документе термин «уникальный молекулярный идентификатор (UMI)» относится к уникальной последовательности нуклеиновой кислоты, которая прикреплена к каждой из множества молекул нуклеиновой кислоты. При включении в молекулу нуклеиновой кислоты UMI в некоторых случаях используется для корректировки смещения последующей амплификации путем прямого подсчета UMI, которые секвенируются после амплификации. Дизайн, встраивание и применение UMI описаны, например, в международной патентной публикации WO 2012/142213, Islam et al. Nat. Methods (2014) 11:163-166, Kivioja, T. et al. Nat. Methods (2012) 9: 72-74, Brenner et al.

(2000) PNAS 97(4), 1665, и Hollas and Schuler, (2003) Conference: 3rd International Workshop on Algorithms in Bioinformatics, Volume: 2812.

[0078] Используемый в настоящем документе термин «баркод» относится к метке нуклеиновой кислоты, которую можно использовать для идентификации образца или источника материала нуклеиновой кислоты. Таким образом, если образцы нуклеиновой кислоты происходят из нескольких источников, нуклеиновые кислоты в каждом образце нуклеиновой кислоты в некоторых случаях метят разными метками нуклеиновой кислоты, чтобы можно было идентифицировать источник образца. Баркоды, также обычно называемые индексами, метками и т.п., хорошо известны специалистам в настоящей области техники. Можно использовать любой подходящий баркод или набор баркодов. См., например, неограничивающие примеры, приведенные в патенте США № 8053192 и международной патентной публикации № W02005/068656. Кодирование посредством баркода отдельных клеток можно проводить, как описано, например, в публикации заявки на выдачу патента САШ № 2013/0274117.

[0079] Термины «твердая поверхность», «твердая подложка» и другие грамматические эквиваленты в настоящем документе относятся к любому материалу, который подходит или может быть модифицирован, чтобы быть подходящим для прикрепления праймеров, баркодов и последовательностей, описанных в настоящем документе. Иллюстративные субстраты включают без ограничения стекло и модифицированное или функционализированное стекло, пластмассы (включая в себя акрилы, полистирол и сополимеры стирола и других материалов, полипропилен, полиэтилен, полибутилен, полиуретаны, Teflon™ и т.д.), полисахариды, нейлон, нитроцеллюлозу, керамику, смолы, диоксид кремния, материалы на основе диоксида кремния (например, кремний или модифицированный кремний), углерод, металлы, неорганические стекла, пластмассы, пучки оптических волокон и множество других полимеров. Согласно некоторым вариантам осуществления твердая подложка содержит узорчатую поверхность, подходящую для иммобилизации праймеров, баркодов и последовательностей в упорядоченной структуре.

[0080] Используемый в настоящем документе термин «биологический образец» включает без ограничения ткани, клетки, биологические жидкости и их изоляты. Клетки или другие образцы, используемые в описанных в настоящем документе способах, в некоторых случаях выделяют из организма пациентов-людей, животных, растений, почвы или других образцов, содержащих микроорганизмы, такие как бактерии, грибы, простейшие и т.д. В некоторых случаях биологический образец получен от человека. В некоторых случаях биологический образец получен не от человека. В некоторых случаях

клетки подвергают способам PTA, описанным в настоящем документе, и секвенированию. Варианты, обнаруженные по всему геному или в определенных местах, можно сравнить со всеми другими клетками, выделенными от этого субъекта, чтобы проследить историю клеточной родословной в исследовательских или диагностических целях. В некоторых случаях варианты подтверждают посредством дополнительных способов анализа, таких как прямое секвенирование с ПЦР.

[0081] Анализ отдельных клеток

[0082] В настоящем документе описаны способы и композиции для анализа отдельных клеток. Анализ клеток в массе дает общую информацию о клеточной популяции, но часто не может обнаружить низкочастотные мутанты при наличие фона. Такие мутанты могут обладать важными свойствами, такими как резистентность к лекарственному средству или мутации, связанные с раком. В некоторых случаях ДНК, РНК и/или белки из одной и той же клетки анализируются параллельно. Анализ может предусматривать идентификацию эпигенетических посттрансляционных (например, гликозилирование, фосфорилирование, ацетилирование, убиквитинирование, гистонная модификация) и/или посттранскрипционных (например, метилирование, гидроксиметилирование) модификаций. Такие способы могут предусматривать «амплификацию, управляемую первичной матрицей» (PTA) для получения библиотек нуклеиновых кислот для секвенирования. В некоторых случаях PTA сочетается с дополнительными стадиями или способами, такими как ПЦР с обратной транскрипцией или способы количественного определения протеома/белка (например, масс-спектрометрия, окрашивание антителами и т.д.). В некоторых случаях различные компоненты клетки физически или пространственно отделены друг от друга на отдельных стадиях анализа. Например, схема в некоторых случаях включает общие стадии, показанные на **фиг. 1А**. Белки сначала метят антителами. В некоторых случаях по меньшей мере некоторые из антител содержат метку или маркер (например, нуклеиновую кислоту/олигометку, массовую метку или флуоресцентную метку). В некоторых случаях часть антител содержит олигометку. В некоторых случаях часть антител содержит флуоресцентный маркер. В некоторых случаях антитела помечены двумя или более метками или маркерами. В некоторых случаях часть антител сортируют по флуоресцентным маркерам. После ПЦР с обратной транскрипцией получают продукты первой цепи мРНК, которые затем удаляют для анализа. Затем из продуктов ПЦР с обратной транскрипцией и баркодов, присутствующих на белок-специфических антителах, получают библиотеки, которые затем секвенируют. Параллельно геномную ДНК из той же клетки подвергают PTA, получая библиотеку, и секвенируют. Результаты секвенирования

генома, протеома и транскриптома в некоторых случаях объединяют с использованием способов биоинформатики. Описанные в настоящем документе способы в некоторых случаях включают любую комбинацию мечения, сортировки клеток, аффинного разделения/очистки, лизиса конкретных клеточных компонентов (например, внешней мембраны, ядра и т.д.), амплификации РНК, амплификации ДНК (например, PTA) или других стадий, связанных с выделением или анализом белка, РНК или ДНК. В некоторых случаях способы, описанные в настоящем документе, включают одну или несколько стадий обогащения, таких как обогащение экзомом.

[0083] В настоящем документе описан первый способ анализа отдельных клеток, предусматривающий анализ РНК и ДНК из отдельной клетки (**фиг. 1B**). Согласно некоторым вариантам осуществления способ предусматривает выделение отдельных клеток, лизис отдельных клеток и обратную транскрипцию (RT). Согласно некоторым вариантам осуществления обратную транскрипцию выполняют с переключающими матрицу олигонуклеотидами (TSO). Согласно некоторым вариантам осуществления TSO содержат молекулярный TAG, такой как биотин, что позволяет последующее выделение продуктов RT кДНК, и амплификацию с ПЦР продуктов RT с получением библиотеки кДНК. Альтернативно или в комбинации, применяют центрифугирование для отделения РНК в супернатанте от кДНК в пеллете клетки. Оставшуюся кДНК в некоторых случаях фрагментируют и удаляют с помощью UDG (урацил-ДНК-гликозилазы), и применяют щелочной лизис для расщепления РНК RNA и денатурации генома. После нейтрализации, добавления праймеров и PTA, продукты амплификации согласно некоторым вариантам осуществлениям очищали на SPRI (твердофазная обратимая иммобилизация) шариках и лигировали с адапторами с получением библиотеки гДНК.

[0084] В настоящем документе описан второй способ анализа отдельных клеток, предусматривающий анализ РНК и ДНК из отдельной клетки (**фиг. 1C**). Согласно некоторым вариантам осуществления способ предусматривает выделение отдельных клеток, лизис отдельных клеток и обратную транскрипцию (RT). Согласно некоторым вариантам осуществления обратную транскрипцию выполняют с переключающими матрицу олигонуклеотидами (TSO). Согласно некоторым вариантам осуществления TSO содержат молекулярный TAG, такой как биотин, что позволяет последующее выделение продуктов RT кДНК, и амплификацию с ПЦР продуктов RT с получением библиотеки кДНК. Согласно некоторым вариантам осуществления щелочной лизис затем применяют для расщепления РНК и денатурации генома. После нейтрализации, добавления случайных праймеров и PTA, продукты амплификации согласно некоторым вариантам осуществлениям очищали на SPRI (твердофазная обратимая иммобилизация) шариках и

лигировали с адапторами с получением библиотеки гДНК. Продукты RT в некоторых случаях выделяют путем выделения, такого как выделение со стрептавидиновыми шариками.

[0085] В настоящем документе описан третий способ анализа отдельных клеток, предусматривающий анализ РНК и ДНК из отдельной клетки (**фиг. 1D**). Согласно некоторым вариантам осуществления способ предусматривает выделение отдельных клеток, лизис отдельных клеток и обратную транскрипцию (RT). Согласно некоторым вариантам осуществления обратную транскрипцию выполняют с переключающими матрицу олигонуклеотидами (TSO) в присутствии терминаторных нуклеотидов. Согласно некоторым вариантам осуществления TSO содержат молекулярный TAG, такой как биотин, что позволяет последующее выделение продуктов RT кДНК, и амплификацию с ПЦР продуктов RT с получением библиотеки кДНК. Согласно некоторым вариантам осуществления щелочной лизис затем применяют для расщепления РНК и денатурации генома. После нейтрализации, добавления случайных праймеров и PTA, продукты амплификации согласно некоторым вариантам осуществлениям очищали на SPRI (твердофазная обратимая иммобилизация) шариках и лигировали с адаптерами с получением библиотеки ДНК. Продукты RT в некоторых случаях выделяют путем выделения, такого как выделение со стрептавидиновыми шариками.

[0086] В настоящем документе описан четвертый способ анализа отдельных клеток, предусматривающий анализ РНК и ДНК из отдельной клетки (**фиг. 1E**). Согласно некоторым вариантам осуществления способ предусматривает выделение отдельных клеток, лизис отдельных клеток и обратную транскрипцию (RT). Согласно некоторым вариантам осуществления обратную транскрипцию выполняют с переключающими матрицу олигонуклеотидами (TSO). Согласно некоторым вариантам осуществления TSO содержат молекулярный TAG, такой как биотин, что позволяет последующее выделение продуктов RT кДНК, и амплификацию с ПЦР продуктов RT с получением библиотеки кДНК. Согласно некоторым вариантам осуществления щелочной лизис затем применяют для расщепления РНК и денатурации генома. После нейтрализации, добавления случайных праймеров и PTA, продукты амплификации в некоторых случаях подвергают амплификации РНКазой и кДНК с использованием блокированных и меченых праймеров. гДНК очищают на SPRI (твердофазная обратимая иммобилизация) шариках и лигируют с адаптерами с получением библиотеки гДНК. Продукты RT в некоторых случаях выделяют путем выделения, такого как выделение со стрептавидиновыми шариками.

[0087] В настоящем документе описан пятый способ анализа отдельных клеток, предусматривающий анализ РНК и ДНК из отдельной клетки (**фиг. 7A** и **7B**). Популяцию

клеток подвергают контакту с библиотекой антител, где антитела являются мечеными. В некоторых случаях антитела метят флуоресцентными метками, баркодами нуклеиновых кислот или тем и другим. Меченые антитела связываются по меньшей мере с одной клеткой в популяции, и такие клетки сортируют, помещая по одной клетке на контейнер (например, пробирку, флакон, микролунку и т.п.). В некоторых случаях контейнер содержит растворитель. В некоторых случаях область поверхности контейнера покрыта фрагментом захвата. В некоторых случаях фрагмент захвата представляет собой небольшую молекулу, антитело, белок или другой агент, способный связываться с одной или несколькими клетками, органеллами или другим компонентом клетки. В некоторых случаях по меньшей мере одна клетка или отдельная клетка или ее компонент связываются с участком поверхности контейнера. В некоторых случаях ядро связывается с областью контейнера. В некоторых случаях внешняя мембрана клетки лизируется, высвобождая мРНК в раствор в контейнере. В некоторых случаях ядро клетки, содержащее геномную ДНК, связывается с участком поверхности контейнера. Затем часто проводят RT с использованием мРНК в растворе в качестве матрицы для создания кДНК. В некоторых случаях переключающие матрицу праймеры содержат 5' - 3' область TSS (сайт начала транскрипции), якорную область, область BC РНК и поли-dT-хвост. В некоторых случаях поли-dT-хвост связывается с поли-A-хвостом одной или нескольких мРНК. В некоторых случаях переключающие матрицу праймеры содержат 3' - 5' область TSS, якорную область и поли-G-область. В некоторых случаях поли-G-область содержит рибоG. В некоторых случаях поли-G-область связывается с поли-C-областью на транскрипте мРНК. В некоторых случаях рибоG добавляют к транскриптам мРНК с помощью терминальной трансферазы. После удаления продуктов ПЦР с обратной транскрипцией для последующего секвенирования любую оставшуюся в клетке РНК удаляют с помощью UNG. Затем ядро лизируют, а высвобожденную геномную ДНК подвергают способу PTA с использованием случайных праймеров с изотермической полимеразой. В некоторых случаях праймеры имеют длину 6-9 оснований. В некоторых случаях PTA генерирует геномные ампликоны длиной 100–5000, 200–5000, 500–2000, 500–2500, 1000–3000 или 300–3000 оснований. В некоторых случаях PTA генерирует геномные ампликоны со средней длиной 100-5000, 200-5000, 500-2000, 500-2500, 1000-3000 или 300-3000 оснований. В некоторых случаях PTA генерирует геномные ампликоны длиной 250-1500 оснований. В некоторых случаях описанные в настоящем документе способы создают пул коротких фрагментов кДНК с приблизительно 500-кратной, приблизительно 750-кратной, приблизительно 1000-кратной, приблизительно 5000-кратной или приблизительно 10000-кратной амплификацией. В некоторых случаях описанные в настоящем документе способы генерируют пул коротких фрагментов кДНК с

500-5000-кратной, 750-1500-кратной или 250-10000-кратной амплификацией. Продукты PTA необязательно подвергают дополнительной амплификации и секвенированию.

[0088] *Получение образца и выделение отдельных клеток*

[0089] Описанные в настоящем документе способы могут потребовать выделения отдельных клеток для анализа. Совместно с PTA можно использовать любой способ выделения отдельных клеток, например, пипетирование ртом, микропипетирование, проточную цитометрию/FACS, микрогидродинамику, способы сортировки ядер (тетраплоидные или другие) или ручное разведение. Таким способам помогают дополнительные реагенты и стадии, например, обогащение на основе антител (например, циркулирующих опухолевых клеток), другие способы обогащения на основе малых молекул или белков или флуоресцентное мечение. В некоторых случаях описанный в настоящем документе способ мультиомического анализа предусматривает механическую или ферментативную диссоциацию клеток из более крупных тканей.

[0090] *Получение и анализ компонентов клетки*

[0091] Способы мультиомического анализа, включающие PTA, описанные в настоящем документе, могут включать один или несколько способов обработки клеточных компонентов, таких как ДНК, РНК и/или белки. В некоторых случаях ядро (содержащее геномную ДНК) физически отделяют от цитозоля (содержащего мРНК), после чего мембранно-селективный лизирующий буфер растворяет мембрану, но сохраняет ядро неповрежденным. Затем цитозоль отделяют от ядра с использованием способов, включающих микропипетирование, центрифугирование или магнитные микрогранулы, конъюгированные с антителами. В другом случае магнитный шарик, покрытый праймером олиго-dT, связывает полиаденилированную мРНК для отделения от ДНК. В другом случае ДНК и РНК предварительно амплифицируют одновременно, а затем разделяют для анализа. В другом случае одна клетка делится на две равные части, при этом обрабатывают мРНК из одной половины и обрабатывают геномную ДНК из другой половины.

[0092] *Мультиомические способы*

[0093] Описанные в настоящем документе способы (например, PTA) можно использовать в качестве замены любого числа других известных в данной области техники способов, которые используются для секвенирования отдельных клеток (мультиомических и т.п.). PTA может заменить способы секвенирования геномной ДНК, такие как MDA, PicoPlex, DOP-PCR, MALBAC или целевые амплификации. В некоторых случаях PTA

заменяет стандартный способ секвенирования геномной ДНК в мультиомическом способе, включая DR-seq (Dey et al., 2015), G&T seq (MacAulay et al., 2015), scMT-seq (Hu et al., 2016), sc-GEM (Cheow et al., 2016), scTrio-seq (Hou et al., 2016), одновременное мультиплексное измерение РНК и белков (Darmanis et al., 2016), scCOOL-seq (Guo et al., 2017), CITE-seq (Stoeckius et al., 2017), REAP-seq (Peterson et al., 2017), scNMT-seq (Clark et al., 2018) или SIDR-seq (Han et al., 2018). Согласно некоторым вариантам осуществления способ, описанный в настоящем документе, предусматривает PTA и способ полиаденилировнных транскриптов мРНК. Согласно некоторым вариантам осуществления способ, описанный в настоящем документе, предусматривает PTA и способ неполиаденилировнных транскриптов мРНК. Согласно некоторым вариантам осуществления способ, описанный в настоящем документе, предусматривает PTA и способ общих (полиаденилировнных и неполиаденилировнных) транскриптов мРНК.

[0094] Согласно некоторым вариантам осуществления PTA комбинируют со стандартным способом секвенирования РНК с получением данных о геноме и транскриптоме. В некоторых случаях мультиомический способ, описанный в настоящем документе, включает PTA и один из следующих способов: Drop-seq (Macosko, et al. 2015), mRNA-seq (Tang et al., 2009), InDrop (Klein et al., 2015), MARS-seq (Jaitin et al., 2014), Smart-seq2 (Hashimshony, et al., 2012, Fish et al., 2016), CEL-seq (Jaitin et al., 2014), STRT-seq (Islam, et al., 2011), Quartz-seq (Sasagawa et al., 2013), CEL-seq2 (Hashimshony, et al. 2016), cytoSeq (Fan et al., 2015), SuPeR-seq (Fan et al., 2011), RamDA-seq (Hayashi, et al. 2018), MATQ-seq (Sheng et al., 2017) или SMARTer (Verboom et al., 2019).

[0095] Для создания библиотек кДНК для анализа транскриптома можно использовать различные реакционные условия и смеси. В некоторых случаях реакционную смесь RT используют для создания библиотеки кДНК. В некоторых случаях реакционная смесь RT содержит реагент краудинга, по меньшей мере один праймер, олигонуклеотид, переключающий матрицу (TSO), обратную транскриптазу и смесь dNTP. В некоторых случаях реакционная смесь RT содержит ингибитор РНКазы. В некоторых случаях реакционная смесь RT содержит одно или несколько поверхностно-активных веществ. В некоторых случаях реакционная смесь RT содержит Tween-20 и/или Triton-X. В некоторых случаях реакционная смесь RT содержит бетаин. В некоторых случаях реакционная смесь RT содержит одну или несколько солей. В некоторых случаях реакционная смесь RT содержит соль магния (например, хлорид магния) и/или хлорид тетраметиламмония. В некоторых случаях реакционная смесь RT содержит желатин. В некоторых случаях реакционная смесь RT содержит PEG (PEG1000, PEG2000, PEG4000, PEG6000, PEG8000 или PEG другой длины).

[0096] Мультиомические способы, описанные в настоящем документе, могут предоставлять как геномную информацию, так и информацию о транскриптах РНК из одной клетки (например, комбинированный или двойной протокол). В некоторых случаях геномную информацию об отдельной клетке получают с помощью способа PTA, а информацию о транскриптах РНК получают с помощью обратной транскрипции с получением библиотеки кДНК. В некоторых случаях для получения библиотеки кДНК использует способ полного транскрипта. В некоторых случаях для получения библиотеки кДНК используют подсчет 3'- или 5'-концов. В некоторых случаях библиотеки кДНК не получают с использованием UMI. В некоторых случаях мультиомический способ предоставляет информацию о транскриптах РНК из отдельной клетки для по меньшей мере 500, 1000, 2000, 5000, 8000, 10000, 12000 или по меньшей мере 15000 генов. Согласно некоторым вариантам осуществления мультиомический способ представляет информацию о транкриптах РНК из отдельной клетки для приблизительно 500, 1000, 2000, 5000, 8000, 10000, 12000 или приблизительно 15000 генов. Согласно некоторым вариантам осуществления мультиомический способ представляет информацию о транкриптах РНК из отдельной клетки для 100-12000 1000-10000, 2000-15000, 5000-15000, 10000-20000, 8000-15000 или 10000-15000 генов. Согласно некоторым вариантам осуществления мультиомический способ представляет информацию о геномной последовательности для по меньшей мере 80%, 90%, 92%, 95%, 97%, 98% или по меньшей мере 99% генома отдельной клетки. Согласно некоторым вариантам осуществления мультиомический способ представляет информацию о геномной последовательности для приблизительно 80%, 90%, 92%, 95%, 97%, 98% или приблизительно 99% генома отдельной клетки.

[0097] Мультиомические способы могут предусматривать анализ отдельных клеток из популяции клеток. Согласно некоторым вариантам осуществления по меньшей мере 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000 или по меньшей мере 8000 клеток анализируют. Согласно некоторым вариантам осуществления приблизительно 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000 или приблизительно 8000 клеток анализируют. Согласно некоторым вариантам осуществления 5-100, 10-100, 50-500, 100-500, 100-1000, 50-5000, 100-5000, 500-1000, 500-10000, 1000-10000 или 5000-20000 клеток анализируют.

[0098] Мультиомические способы могут генерировать выход геномной ДНК из реакции PTA в зависимости от типа отдельной клетки. В некоторых случаях количество ДНК, получаемой из отдельной клетки, составляет приблизительно 0,1, 1, 1,5, 2, 3, 5 или приблизительно 10 микрограмм. В некоторых случаях количество ДНК, получаемой из отдельной клетки, составляет приблизительно 0,1, 1, 1,5, 2, 3, 5 или приблизительно 10 фемтограмм. В некоторых случаях количество ДНК, получаемой из отдельной клетки,

составляет по меньшей мере 0,1, 1, 1,5, 2, 3, 5 или по меньшей мере 10 микрограмм. В некоторых случаях количество ДНК, получаемой из отдельной клетки, составляет по меньшей мере 0,1, 1, 1,5, 2, 3, 5 или по меньшей мере 10 фемтограмм. В некоторых случаях количество ДНК, получаемой из отдельной клетки, составляет приблизительно 0,1-10, 1-10, 1,5-10, 2-20, 2-50, 1-3 или 0,5-3,5 микрограмм. В некоторых случаях количество ДНК, получаемой из отдельной клетки, составляет приблизительно 0,1-10, 1-10, 1,5-10, 2-20, 2-4, 1-3 или 0,5-4 фемтограмм.

[0099] *Анализ метиломы*

[00100] В настоящем документе описаны способы, включающие PTA, в которых сайты метилированной ДНК в отдельных клетках определяют с использованием способа PTA. В некоторых случаях эти способы дополнительно предусматривают параллельный анализ транскриптома и/или протеома одной и той же клетки. Способы обнаружения метилированных геномных оснований предусматривают селективную рестрикцию чувствительными к метилированию эндонуклеазами с последующей обработкой способом PTA. Сайты, разрезаемые такими ферментами, определяют посредством секвенирования и идентифицируют метилированные основания. В другом случае бисульфитная обработка библиотек геномной ДНК превращает неметилированные цитозины в урацил. Затем библиотеки в некоторых случаях амплифицируют с помощью специфических к метилированию праймеров, которые селективно гибридизуются с метилированными последовательностями. В качестве альтернативы проводится специфическая к отсутствию метилирования ПЦР с последующим применением одного или нескольких способов для различения оснований, прореагировавших с бисульфитом, включая прямое пиросеквенирование, MS-SnuPE, HRM, COBRA, MS-SSCA или специфическое для основания расщепление/MALDI-TOF. В некоторых случаях образцы геномной ДНК разделяют для параллельного анализа генома (или его обогащенной части) и анализа метилома. В некоторых случаях анализ генома и метилома включает обогащение геномных фрагментов (например, экзома или других мишеней) или секвенирование всего генома.

[00101] *Биоинформатика*

[00102] Данные, полученные из способов анализа отдельных клеток с использованием описанных в настоящем документе PTA, могут быть объединены в базу данных. Описаны способы и системы интеграции биоинформационных данных. Данные для протеома, генома, транскриптома, метилома или другие данные в некоторых случаях

объединяют/интегрируют в базу данных и анализируют. Способы и системы интеграции биоинформационных данных в некоторых случаях включают одно или несколько из обнаружения белков (FACS и/или NGS), обнаружения мРНК и/или обнаружения вариаций генома. В некоторых случаях эти данные коррелируют со статусом заболевания или состоянием. В некоторых случаях данные из множества отдельных клеток компилируют для описания свойств более крупной клеточной популяции, такой как клетки из определенного образца, области, организма или ткани. В некоторых случаях данные о белках получают с помощью флуоресцентно меченных антител, которые избирательно связываются с белками клетки. В некоторых случаях способ обнаружения белка предусматривает группировку клеток на основе флуоресцентных маркеров и сообщение местоположения образца после сортировки. В некоторых случаях способ обнаружения белка предусматривает обнаружение баркодов образцов, обнаружение баркодов белков, сравнение с сконструированными последовательностями и группирование клеток на основе баркода и количества копий. В некоторых случаях данные о белках получают из антител с баркодом, которые избирательно связываются с белками на клетке. В некоторых случаях данные для транскриптома получают из баркодов образцов и РНК. В некоторых случаях способ обнаружения мРНК предусматривает обнаружение баркодов, специфичных для образца и РНК, выравнивание по геному, выравнивание по RefSeq/кодированию, представление последовательностей экзона/интрона/межгенных последовательностей, анализ экзон-экзонных соединений, группирование клеток на основе баркода и вариации экспрессии и кластерный анализ вариации и высоковариабельных генов. В некоторых случаях геномные данные получают из баркодов, специфических для образцов и ДНК. В некоторых случаях способ обнаружения вариабельности генома предусматривает обнаружение специфических для образца и ДНК баркодов, выравнивание с геномом, определение скорости извлечения генома и картирования SNV, фильтрацию прочтений экзон-экзонных соединений, создание файла вызова вариантов (VCF) и кластерный анализ вариации и высоковариабельных мутаций.

[00103] *Мутации*

[00104] В некоторых случаях способы (например, мультиомический PTA), описанные в настоящем документе, обеспечивают более высокую чувствительность обнаружения и/или более низкую частоту ложноположительных результатов при обнаружении мутаций. В некоторых случаях мутация представляет собой разницу между анализируемой последовательностью (например, с использованием описанных в настоящем документе способов) и эталонной последовательностью. Эталонные

последовательности в некоторых случаях получают из других организмов, других особей того же или подобного вида, популяций организмов или других областей того же генома. В некоторых случаях мутации идентифицируют на плазмиде или хромосоме. В некоторых случаях мутация представляет собой SNV (изменение одного нуклеотида), SNP (полиморфизм одного нуклеотида) или CNV (изменение числа копий или CNA/аберрация числа копий). В некоторых случаях мутация представляет собой замену, вставку или делецию основания. Согласно некоторым вариантам осуществления мутация представляет собой транзицию, трансверсию, несмысловую мутацию, молчащую мутацию, синонимическую или несинонимическую мутацию, непатогенную мутацию, миссенс-мутацию или мутацию сдвига рамки считывания (делецию или вставку). В некоторых случаях PTA приводит к более высокой чувствительности обнаружения и/или более низкой частоте ложноположительных результатов при обнаружении мутаций по сравнению с такими методами, как прогноз in-silico, ChIP-seq, GUIDE-seq, circle-seq, HTGTS (полногеномное транслокационное секвенирование с высокой пропускной способностью), IDLV (интеграционно-дефицитный лентивирус), Digenome-seq, FISH (флуоресцентная гибридизация in situ) или DISCOVER-seq.

[00105] Амплификация, управляемая первичной матрицей

[00106] В настоящем документе описаны способы амплификации нуклеиновых кислот, такие как «амплификация, управляемая первичной матрицей, (PTA)». Например, описанные в настоящем документе способы PTA схематически представлены на фиг. 1G. С помощью способа PTA ампликоны предпочтительно создаются из первичной матрицы («прямые копии») с использованием полимеразы (например, полимеразы, замещающей цепь). Следовательно, ошибки передаются от дочерних ампликонов во время последующих амплификаций с меньшей скоростью по сравнению с MDA. Результатом является легко выполняемый способ, который, в отличие от существующих протоколов WGA, может точно и воспроизводимо усиливать низкий входной сигнал ДНК, включая геномы отдельных клеток, с высокой широтой охвата и однородностью. Более того, терминированные продукты амплификации можно подвергать направленному лигированию после удаления терминаторов, что позволяет прикрепить баркод клетки к праймерам для амплификации, так что продукты из всех клеток можно объединить после прохождения параллельных реакций амплификации. В некоторых случаях матричные нуклеиновые кислоты не связаны с твердой подложкой. В некоторых случаях прямые копии матричных нуклеиновых кислот не связаны с твердой подложкой. В некоторых случаях один или несколько праймеров не связаны с твердой подложкой. В некоторых случаях

никакие праймеры не связаны с твердой подложкой. В некоторых случаях праймер прикрепляют к первой твердой подложке, а матричную нуклеиновую кислоту прикрепляют ко второй твердой подложке, при этом первая и вторая твердые подложки не являются одной и той же. В некоторых случаях РТА используют для анализа отдельных клеток из большей популяции клеток. В некоторых случаях РТА используют для анализа более чем одной клетки из большей популяции клеток или всей популяции клеток.

[00107] В настоящем документе описаны способы, в которых для амплификации используют полимеразы нуклеиновых кислот с активностью замещения цепи. В некоторых случаях такие полимеразы обладают активностью замещения цепи и низкой частотой появления ошибок. В некоторых случаях такие полимеразы обладают активностью замещения цепи и корректирующей экзонуклеазной активностью, такой как 3'->5' корректирующая активность. В некоторых случаях полимеразы нуклеиновых кислот используют в сочетании с другими компонентами, такими как обратимые или необратимые терминаторы или дополнительные факторы замещения цепи. В некоторых случаях полимераза обладает активностью замещения цепи, но не обладает корректирующей экзонуклеазной активностью. Например, в некоторых случаях такие полимеразы включают полимеразу бактериофага phi29 (Ф29), которая также характеризуется очень низкой частотой появления ошибок в результате 3'->5' корректирующей экзонуклеазной активности (см., например, патенты США №№ 5198543 и 5001050). В некоторых случаях неограничивающие примеры замещающих цепи полимераз нуклеиновых кислот включают, например, генетически модифицированную ДНК-полимеразу phi29 (Ф29), фрагмент Кленова ДНК-полимеразы I (Jacobsen et al., Eur. J. Biochem. 45:623-627 (1974)), ДНК-полимеразу фага M2 (Matsumoto et al., Gene 84:247 (1989)), ДНК-полимеразу фага phiPRDl (Jung et al., Proc. Natl. Acad. Sei. USA 84:8287 (1987), Zhu and Ito, Biochim. Biophys. Acta. 1219:267-276 (1994)), ДНК-полимеразу Bst (например, ДНК-полимеразу большого фрагмента Bst (Exo(-) Bst, Aliotta et al., Genet. Anal. (Netherlands) 12:185-195 (1996)), ДНК-полимеразу экзо(-) Bca (Walker and Linn, Clinical Chemistry 42:1604-1608 (1996)), ДНК-полимеразу Bsu, ДНК-полимеразу VentR, включая ДНК-полимеразу VentR(3K3O-) (Kong et al., J. Biol. Chem. 268:1965-1975 (1993)), ДНК-полимеразу Deep Vent, включая ДНК-полимеразу Deep Vent (экзо-), ДНК-полимеразу IsoPol, ДНК-полимеразу I, ДНК-полимеразу Therminator, ДНК-полимеразу T5 (Chatterjee et al., Gene 97:13-19 (1991)), секвеназу (U.S. Biochemicals), ДНК-полимеразу T7, T7-секвеназу, ДНК- полимеразу T7 gp5, PRDI ДНК-полимеразу, ДНК-полимеразу T4 (Kaboord and Benkovic, Curr. Biol. 5:149-157 (1995)). Дополнительные замещающие цепи полимеразы нуклеиновых кислот также совместимы со способами, описанными в настоящем документе. Способность данной

полимеразы осуществлять репликацию с замещением цепи можно определить, например, с использованием полимеразы в анализе репликации с замещением цепи (например, как описано в патенте США № 6977148). Такие анализы в некоторых случаях проводят при температуре, подходящей для оптимальной активности используемого фермента, например, 32°C для ДНК-полимеразы phi29, от 46°C до 64°C для ДНК- полимеразы экзо(-) Bst или от приблизительно 60°C до 70°C для фермента из гипертермофильного организма. Другой применимый анализ для выбора полимеразы представляет собой анализ с блокированием праймера, описанный в Kong et al., J. Biol. Chem. 268:1965-1975 (1993). Анализ состоит из анализа удлинения праймера с использованием матрицы оцДНК M13 в присутствии или при отсутствии олигонуклеотида, который гибридизуется в направлении 5' от удлиняющего праймера, чтобы блокировать его продвижение. Другие ферменты, способные замещать блокирующий праймер в этом анализе, в некоторых случаях можно использовать для раскрытого способа. В некоторых случаях полимеразы включают dNTP и терминаторы приблизительно с одинаковой скоростью. В некоторых случаях отношение скоростей включения для dNTP и терминаторов для полимеразы, описанной в настоящем документе, составляет приблизительно 1:1, приблизительно 1,5:1, приблизительно 2:1, приблизительно 3:1 приблизительно 4:1 приблизительно 5:1, приблизительно 10:1, приблизительно 20:1 приблизительно 50:1, приблизительно 100:1, приблизительно 200:1, приблизительно 500:1 или приблизительно 1000:1. В некоторых случаях отношение скоростей включения для dNTP и терминаторов для полимеразы, описанной в настоящем документе, составляет от 1:1 до 1000:1, от 2:1 до 500:1, от 5:1 до 100:1, от 10:1 до 1000:1, от 100:1 до 1000:1, от 500:1 до 2000:1, от 50:1 до 1500:1 или от 25:1 до 1000:1.

[00108] В настоящем документе описаны способы амплификации, при которых замещение цепей можно облегчить с помощью фактора замещения цепи, такого как, например, хеликаза. Такие факторы в некоторых случаях используют в сочетании с дополнительными компонентами амплификации, такими как полимеразы, терминаторы или другой компонент. В некоторых случаях фактор замещения цепи используют с полимеразой, не обладающей активностью замещения цепи. В некоторых случаях фактор замещения цепи используют с полимеразой, обладающей активностью замещения цепи. Без ограничения какой-либо теорией, факторы замещения цепи могут увеличивать скорость повторного праймирования более мелких двухцепочечных ампликонов. В некоторых случаях любая ДНК-полимераза, которая может выполнять репликацию с замещением цепи в присутствии фактора замещения цепи, подходит для использования в способе PTA, даже если ДНК-полимераза не выполняет репликацию с замещением цепи при отсутствии такого фактора. Факторы замещения цепи, применимые в репликации с замещением цепи, в

некоторых случаях включают без ограничения дополнительную субъединицу полимеразы BMRF1 (Tsurumi et al., J. Virology 67(12):7648-7653 (1993)), белок, связывающий ДНК аденовируса (Zijderveld and van der Vliet, J. Virology 68(2): 1158-1164 (1994)), вирусный белок простого герпеса ICP8 (Boehmer and Lehman, J. Virology 67(2):711- 715 (1993), Skaliter and Lehman, Proc. Natl. Acad. Sei. USA 91(22):10665-10669 (1994)), белки, связывающие одноцепочечную ДНК (SSB, Rigler and Romano, J. Biol. Chem. 270:8910-8919 (1995)), белок гена 32 фага T4 (Villemain and Giedroc, Biochemistry 35:14395- 14404 (1996), хеликаза-праймазу T7, белок SSB gp2.5 T7, Tte-UvrD (от Thermoanaerobacter tengcongensis), хеликазу тимуса теленка (Siegel et al., J. Biol. Chem. 267:13629-13635 (1992)), бактериальный SSB (например, SSB *E. coli),* репликационный белок A (RPA) у эукариот, митохондриальный SSB человека (mtSSB) и рекомбиназы (например, белки семейства рекомбиназа A (RecA), T4 UvsX, Sak4 фага HK620, Rad51, Dmcl или Radb). Комбинации факторов, которые облегчают замещение цепей и праймирование, также являются сопоставимыми со способами, описанными в настоящем документе. Например, хеликазу используют вместе с полимеразой. В некоторых случаях способ PTA включает применение белка, связывающего одноцепочечную ДНК (SSB, gp32 T4 или другого связывающего одноцепочечную ДНК белка), хеликазы и полимеразы (например, SauДНК-полимераза, Bsu-полимераза, Bst2.0, GspM, GspM2.0, GspSSD или другая подходящая полимераза). В некоторых случаях обратные транскриптазы используют в сочетании с факторами замещения цепи, описанными в настоящем документе. В некоторых случаях амплификацию проводят с применением полимеразы и фермента никования (например, “NEAR”), такого как описанные, например, в патенте США № 9617586. В некоторых случаях ферментом никования является Nt.BspQI, Nb.BbvCi, Nb.BsmI, Nb.BsrDI, Nb.BtsI, Nt.AlwI, Nt.BbvCI, Nt.BstNBI, Nt.CviPII, Nb.Bpu10I или Nt.Bpu10I.

[00109] В настоящем документе описаны способы амплификации, предусматривающие применение терминаторных нуклеотидов, полимераз и дополнительных факторов или условий. Например, такие факторы используют в некоторых случаях для фрагментации матрицы (матриц) нуклеиновой кислоты или ампликонов во время амплификации. В некоторых случаях такие факторы включают эндонуклеазы. В некоторых случаях факторы включают транспозазы. В некоторых случаях для фрагментации нуклеиновых кислот во время амплификации используют механический сдвиг. В некоторых случаях во время амплификации добавляют нуклеотиды, которые можно фрагментировать путем добавления дополнительных белков или условий. Например, в ампликоны включают урацил, обработка урацил-D-гликозилазой фрагментирует нуклеиновые кислоты в урацил-содержащих положениях. В некоторых

случаях также используют дополнительные системы для селективной фрагментации нуклеиновых кислот, например, сконструированную ДНК-гликозилазу, которая расщепляет модифицированные пары оснований цитозин-пирен. (Kwon, et al. Chem Biol. 2003, 10(4), 351).

[00110] В настоящем документе описаны способы амплификации, включающие применение терминаторных нуклеотидов, которые останавливают репликацию нуклеиновой кислоты, таким образом, уменьшая размер продуктов амплификации. Такие терминаторы в некоторых случаях используют в сочетании с полимеразами, факторами замещения цепи или другими компонентами амплификации, описанными в настоящем документе. В некоторых случаях терминаторные нуклеотиды снижают или уменьшают эффективность репликации нуклеиновой кислоты. Такие терминаторы в некоторых случаях снижают скорости удлинения по меньшей мере на 99,9%, 99%, 98%, 95%, 90%, 85%, 80%, 75%, 70% или по меньшей мере 65%. Такие терминаторы в некоторых случаях снижают скорости удлинения на 50%-90%, 60%-80%, 65%-90%, 70%-85%, 60%-90%, 70%-99%, 80%-99% или 50%-80%. В некоторых случаях терминаторы снижают среднюю длину продукта ампликона по меньшей мере на 99,9%, 99%, 98%, 95%, 90%, 85%, 80%, 75%, 70% или по меньшей мере 65%. Терминаторы в некоторых случаях снижают среднюю длину ампликона на 50%-90%, 60%-80%, 65%-90%, 70%-85%, 60%-90%, 70%-99%, 80%-99% или 50%-80%. В некоторых случаях ампликоны, содержащие терминаторные нуклеотиды, образуют петли или шпильки, которые снижают способность полимеразы использовать такие ампликоны в качестве матриц. Использование терминаторов в некоторых случаях снижает скорость амплификации в начальных сайтах амплификации посредством включения терминаторных нуклеотидов (например, дидезоксинуклеотидов, которые были модифицированы, чтобы сделать их устойчивыми к экзонуклеазам для прекращения удлинения ДНК), что приводит к меньшим по размеру продуктам амплификации. Путем получения продуктов амплификации меньшего размера, чем в используемых в настоящее время способах (например, средняя длина, составляющая 50-2000 нуклеотидов для способов PTA по сравнению со средней длиной продукта, составляющей > 10 000 нуклеотидов для способов MDA), продукты амплификации PTA в некоторых случаях подвергаются прямому лигированию с адаптерами без необходимости фрагментации, что позволяет эффективно встраивать баркоды клеток и уникальные молекулярные идентификаторы (UMI) (см. фиг. 2A).

[00111] Терминаторные нуклеотиды присутствуют при различных концентрациях в зависимости от таких факторов, как полимераза, матрица или другие факторы. Например, количество терминаторных нуклеотидов в некоторых случаях

выражают как отношение терминаторных нуклеотидов к нетерминаторным нуклеотидам в способе, описанном в настоящем документе. Такие концентрации в некоторых случаях позволяют контролировать длины ампликонов. В некоторых случаях отношение терминаторных нуклеотидов к нетерминаторным нуклеотидам модифицируют для количества присутствующих матриц или размера матрицы. В некоторых случаях отношение терминаторных нуклеотидов к нетерминаторным нуклеотидам уменьшают для более маленьких размеров образцов (например, диапазон от фемтограмма до пикограмма). В некоторых случаях отношение терминаторных нуклеотидов к нетерминаторным нуклеотидам составляет приблизительно 2:1, 5:1, 7:1, 10:1, 20:1, 50:1, 100:1, 200:1, 500:1, 1000:1, 2000:1 или 5000:1. В некоторых случаях отношение терминаторных нуклеотидов к нетерминаторным нуклеотидам составляет 2:1-10:1, 5:1-20:1, 10:1-100:1, 20:1-200:1, 50:1-1000:1, 50:1-500:1, 75:1-150:1 или 100:1-500:1. В некоторых случаях по меньшей мере один из нуклеотидов, присутствующий в ходе амплификации с использованием способа, описанного в настоящем документе, представляет собой терминаторный нуклеотид. Необязательно, чтобы каждый терминатор присутствовал в приблизительно одинаковой концентрации, в некоторых случаях отношения каждого терминатора, присутствующего в способе, описанном в настоящем документе, оптимизируют для конкретного набора условий реакции, типа образца или полимеразы. Без ограничения какой-либо теорией каждый терминатор может обладать разной эффективностью для включения в растущую полинуклеотидную цепь ампликона в ответ на спаривание с соответствующим нуклеотидом на матричной цепи. Например, в некоторых случаях терминатор, спаривающийся с цитозином, присутствует при концентрации на приблизительно 3%, 5%, 10%, 15%, 20%, 25% или 50% более высокой концентрации, чем средняя концентрация терминатора. В некоторых случаях терминатор, спаривающийся с тимином, присутствует при концентрации на приблизительно 3%, 5%, 10%, 15%, 20%, 25% или 50% более высокой концентрации, чем средняя концентрация терминатора. В некоторых случаях терминатор, спаривающийся с гуанином, присутствует при концентрации на приблизительно 3%, 5%, 10%, 15%, 20%, 25% или 50% более высокой концентрации, чем средняя концентрация терминатора В некоторых случаях терминатор, спаривающийся с аденином, присутствует при концентрации на приблизительно 3%, 5%, 10%, 15%, 20%, 25% или 50% более высокой концентрации, чем средняя концентрация терминатора. В некоторых случаях терминатор, спаривающийся с урацилом, присутствует при концентрации на приблизительно 3%, 5%, 10%, 15%, 20%, 25% или 50% более высокой концентрации, чем средняя концентрация терминатора. Любой нуклеотид, способный прекращать удлинение нуклеиновой кислоты полимеразой нуклеиновой кислоты, в некоторых случаях используют в качестве

терминаторного нуклеотида в способах, описанных в настоящем документе. В некоторых случаях для прекращения репликации нуклеиновой кислоты используют обратимый терминатор. В некоторых случаях необратимый терминатор используют для прекращения репликации нуклеиновой кислоты. В некоторых случаях неограничивающие примеры терминаторов включают обратимые и необратимые нуклеиновые кислоты и аналоги нуклеиновых кислот, такие как, например, 3' блокированный обратимый терминатор, содержащий нуклеотиды, 3' неблокированный обратимый терминатор, содержащий нуклеотиды, терминаторы, содержащие 2' модификации дезоксинуклеотидов, терминаторы, содержащие модификации азотистого основания дезоксинуклеотидов, или любую их комбинацию. Согласно одному варианту осуществления терминаторные нуклеотиды представляют собой дидезоксинуклеотиды. Другие нуклеотидные модификации, которые останавливают репликацию нуклеиновой кислоты и могут являться подходящими для осуществления на практике настоящего изобретения, включают без ограничения любые модификации группы г 3' углерода дезоксирибозы, такие как инвертированные дидезоксинуклеотиды, 3' биотинилированные нуклеотиды, 3' аминонуклеотиды, 3'-фосфорилированные нуклеотиды, З'-О-метилнуклеотиды, 3' углеродные спейсерные нуклеотиды, включая 3' С3 спейсерные нуклеотиды, 3' С18 нуклеотиды, 3' гександиоловые спейсерные нуклеотиды, ациклонуклеотиды и их комбинации. В некоторых случаях терминаторы представляют собой полинуклеотиды, содержащие 1, 2, 3, 4 или более оснований в длину. В некоторых случаях терминаторы представляют собой полинуклеотиды, содержащие 1, 2, 3, 4 или более оснований в длину. В некоторых случаях терминаторы не содержат обнаруживаемый фрагмент или метку (например, массовую метку, флуоресцентную метку, краситель, радиоактивный атом или другой обнаруживаемый фрагмент). В некоторых случаях терминаторы не содержат химический фрагмент, обеспечивающий возможность прикрепления обнаруживаемого фрагмента или метки (например, «клик» азид/алкин, партнер для присоединения конъюгата или другой химический элемент для прикрепления метки). В некоторых случаях все терминаторные нуклеотиды содержат одну и ту же модификацию, которая снижает амплификацию до области (например, сахарный фрагмент, основный фрагмент или фосфатный фрагмент) нуклеотида. В некоторых случаях по меньшей мере один терминатор характеризуется другой модификацией, которая снижает амплификацию. В некоторых случаях все терминаторы характеризуются по существу одинаковыми длинами волн флуоресцентного возбуждения или излучения. В некоторых случаях терминаторы без модификации фосфатной группы используют с полимеразами, которые не обладают корректирующей экзонуклеазной активностью. Терминаторы, при использовании с

полимеразами, которые обладают 3’->5’ корректирующей экзонуклеазной активностью (такая как, например, phi29), которая может удалить терминаторный нуклеотид, в некоторых случаях дополнительно модифицируют, чтобы сделать их устойчивыми к экзонуклеазам. Например, дидезоксинуклеотиды модифицируют альфа-тиогруппой, которая создает фосфоротиоатную связь, которая делает эти нуклеотиды устойчивыми к 3’->5’ корректирующей экзонуклеазной активности полимераз нуклеиновых кислот. Такие модификации в некоторых случаях снижают корректирующую экзонуклеазную активность по меньшей мере на 99,5%, 99%, 98%, 95%, 90% или по меньшей мере на 85%. Неограничивающие примеры других модификаций терминаторных нуклеотидов, обеспечивающих устойчивость к 3’->5’ корректирующей экзонуклеазной активности, включают в некоторых случаях нуклеотиды с модификацией альфа-группы, такие как альфа-тио-дидезоксинуклеотиды, создающие фосфоротиоатную связь, С3 спейсерные нуклеотиды, запертые нуклеиновые кислоты (LNA), инвертированные нуклеиновые кислоты, 2’ фторооснования, 3’ фосфорилирование, 2’-О-метиловые модификации (или другие 2’-О-алкильные модификации), модифицированные пропином основания (например, дезоксицитозин, дезоксиуридин), нуклеотиды L-ДНК, нуклеотиды L- =РНК, нуклеотиды с инвертированными связями (например, 5’-5’ или 3’-3’), 5’ инвертированные основания (например, 5’ инвертированные 2’,3’-дидезокси dT), метилфосфонатные скелеты и транснуклеиновые кислоты. В некоторых случаях нуклеотиды с модификацией включают нуклеиновые кислоты с модифицированными основаниями, содержащие свободные 3’ ОН-группы (например, 2-нитробензилалкилированные трифосфаты HOMedU, основания, содержащие модификацию с большими химическими группами, такими как твердые подложки или другой большой фрагмент). В некоторых случаях полимеразу с активностью замещения цепи, но без 3’->5’ корректирующей экзонуклеазной активности используют с терминаторными нуклеотидами с модификациями или без них, чтобы сделать их устойчивыми к экзонуклеазам. Такие полимеразы нуклеиновых кислот включают без ограничения ДНК-полимеразу Bst, ДНК-полимеразу Bsu, ДНК-полимеразу Deep Vent (экзо-), ДНК-полимеразу фрагмента Кленова (экзо-), ДНК-полимеразу Therminator и $Vent_R$ (экзо-).

[00112] Библиотеки праймеров и ампликонов

[00113] В настоящем документе описаны библиотеки ампликонов, полученные в результате амплификации по меньшей мере одной целевой молекулы нуклеиновой кислоты. Такие библиотеки в некоторых случаях получают с использованием способов, описанных в настоящем документе, таких как способы с использованием

терминаторов. Такие способы предусматривают применение полимераз или факторов замещения цепи, терминаторных нуклеотидов (обратимых или необратимых) или других признаков и вариантов осуществления, описанных в настоящем документе. В некоторых случаях библиотеки ампликонов, полученные путем применения терминаторов, описанных в настоящем документе, дополнительно амплифицируют в последующей реакции амплификации (например, ПЦР). В некоторых случаях последующие реакции амплификации не содержат терминаторы. В некоторых случаях библиотеки ампликонов содержат полинуклеотиды, где по меньшей мере 50%, 60%, 70%, 80%, 90%, 95% или по меньшей мере 98% полинуклеотидов содержат по меньшей мере один терминаторный нуклеотид. В некоторых случаях библиотека ампликонов содержит целевую молекулу нуклеиновой кислоты, из которой была получена библиотека ампликонов. Библиотека ампликонов содержит множество полинуклеотидов, где по меньшей мере некоторые из полинуклеотидов представляют собой прямые копии (например, реплицируются непосредственно из целевой молекулы нуклеиновой кислоты, такой как геномная ДНК, РНК или другая целевая нуклеиновая кислота). Например, по меньшей мере 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% или больше чем 95% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях по меньшей мере 5% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях по меньшей мере 10% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях по меньшей мере 15% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях по меньшей мере 20% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях по меньшей мере 50% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях 3%-5%, 3-10%, 5%-10%, 10%-20%, 20%-30%, 30%-40%, 5%-30%, 10%-50% или 15%-75% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях по меньшей мере некоторые из полинуклеотидов представляют собой прямые копии целевой молекулы нуклеиновой кислоты или дочернего (первая копия целевой нуклеиновой кислоты) потомства. Например, по меньшей мере 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% или больше чем 95% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой

кислоты или дочернего потомства. В некоторых случаях по меньшей мере 5% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты или дочернего потомства. В некоторых случаях по меньшей мере 10% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты или дочернего потомства. В некоторых случаях по меньшей мере 20% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты или дочернего потомства. В некоторых случаях по меньшей мере 30% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты или дочернего потомства. В некоторых случаях 3%-5%, 3%-10%, 5%-10%, 10%-20%, 20%-30%, 30%-40%, 5%-30%, 10%-50% или 15%-75% полинуклеотидов ампликона представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты или дочернего потомства. В некоторых случаях прямые копии целевой нуклеиновой кислоты имеют длину 50-2500, 75-2000, 50-2000, 25-1000, 50-1000, 500-2000 или 50-2000 оснований. В некоторых случаях дочернее потомство имеет длину 1000-5000, 2000-5000, 1000-10000, 2000-5000, 1500-5000, 3000-7000 или 2000-7000 оснований. В некоторых случаях средняя длина продуктов амплификации РТА составляет 25-3000 нуклеотидов в длину, 50-2500, 75-2000, 50-2000, 25-1000, 50-1000, 500-2000 или 50-2000 оснований в длину. В некоторых случаях ампликоны, созданные с помощью РТА, составляют не более 5000, 4000, 3000, 2000, 1700, 1500, 1200, 1000, 700, 500 или не более 300 оснований в длину. В некоторых случаях ампликоны, созданные с помощью РТА, составляют 1000-5000, 1000-3000, 200-2000, 200-4000, 500-2000, 750-2500 или 1000-2000 оснований в длину. Библиотеки ампликонов, созданные с использованием способов, описанных в настоящем документе, в некоторых случаях содержат по меньшей мере 1000, 2000, 5000, 10000, 100000, 200000, 500000 или больше чем 500000 ампликонов, содержащих уникальные последовательности. В некоторых случаях библиотека содержит по меньшей мере 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 2000, 2500, 3000 или по меньшей мере 3500 ампликонов. В некоторых случаях по меньшей мере 5%, 10%, 15%, 20%, 25%, 30% или больше чем 30% полинуклеотидов ампликона с длиной, составляющей меньше чем 1000 оснований, представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях по меньшей мере 5%, 10%, 15%, 20%, 25%, 30% или больше чем 30% полинуклеотидов ампликона с длиной, составляющей не более 2000 оснований, представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях по меньшей мере 5%, 10%, 15%, 20%, 25%, 30% или больше чем 30% полинуклеотидов

ампликона с длиной, составляющей 3000-5000 оснований, представляют собой прямые копии по меньшей мере одной целевой молекулы нуклеиновой кислоты. В некоторых случаях отношение ампликонов - прямых копий к целевым молекулам нуклеиновой кислоты составляет по меньшей мере 10:1, 100:1, 1000:1, 10000:1, 100000:1, 1000000:1, 10000000:1 или больше чем 10000000:1. В некоторых случаях отношение ампликонов - прямых копий к целевым молекулам нуклеиновой кислоты составляет по меньшей мере 10:1, 100:1, 1000:1, 10000:1, 100000:1, 1000000:1, 10000000:1 или больше чем 10000000:1, где ампликоны - прямые копии составляют не более 700-1200 оснований в длину. В некоторых случаях отношение ампликонов - прямых копий и дочерних ампликонов к целевым молекулам нуклеиновой кислоты составляет по меньшей мере 10:1, 100:1, 1000:1, 10000:1, 100000:1, 1000000:1, 10000000:1 или больше чем 10000000:1. В некоторых случаях отношение ампликонов - прямых копий и дочерних ампликонов к целевым молекулам нуклеиновой кислоты составляет по меньшей мере 10:1, 100:1, 1000:1, 10000:1, 100000:1, 1000000:1, 10000000:1 или больше чем 10000000:1, где ампликоны - прямые копии составляют 700-1200 оснований в длину, и дочерние ампликоны составляют 2500-6000 оснований в длину. В некоторых случаях библиотека содержит приблизительно 50-10000, приблизительно 50- 5000, приблизительно 50-2500, приблизительно 50-1000, приблизительно 150-2000, приблизительно 250-3000, приблизительно 50-2000, приблизительно 500-2000 или приблизительно 500-1500 ампликонов, которые представляют собой прямые копии целевой молекулы нуклеиновой кислоты. В некоторых случаях библиотека содержит приблизительно 50-10000, приблизительно 50-5000, приблизительно 50-2500, приблизительно 50-1000, приблизительно 150-2000, приблизительно 250-3000, приблизительно 50-2000, приблизительно 500-2000 или приблизительно 500-1500 ампликонов, которые представляют собой прямые копии целевой молекулы нуклеиновой кислоты или дочерних ампликонов. Число прямых копий в некоторых случаях можно контролировать количеством циклов ПЦР-амплификации. В некоторых случаях для получения копий целевой молекулы нуклеиновой кислоты используют не более 30, 25, 20, 15, 13, 11, 10, 9, 8, 7, 6, 5, 4 или 3 циклов ПЦР. В некоторых случаях для получения копий целевой молекулы нуклеиновой кислоты используют приблизительно 30, 25, 20, 15, 13, 11, 10, 9, 8, 7, 6, 5, 4 или приблизительно 3 цикла ПЦР. В некоторых случаях для получения копий целевой молекулы нуклеиновой кислоты 3, 4, 5, 6, 7 или 8 циклов ПЦР. В некоторых случаях 2-4, 2-5, 2-7, 2-8, 2-10, 2-15, 3-5, 3-10, 3-15, 4-10, 4-15, 5-10 или 5-15 циклов ПЦР используют для получения копий целевой молекулы нуклеиновой кислоты. Библиотеки ампликонов, созданные с использованием способов, описанных в настоящем документе, в некоторых случаях подвергают дополнительным

стадиям, таким как лигирование адаптера и дальнейшая ПЦР-амплификация. В некоторых случаях такие дополнительные стадии предшествуют стадии секвенирования.

[00114] Описанные в настоящем документе способы могут дополнительно включать одну или несколько стадий обогащения или очистки. В некоторых случаях один или несколько полинуклеотидов (таких как кДНК, ампликоны РТА или другие полинуклеотиды) обогащают в ходе описанного в настоящем документе способа. В некоторых случаях полинуклеотидные зонды используют для захвата одного или нескольких полинуклеотидов. В некоторых случаях зонды настроены на захват одного или нескольких геномных экзонов. В некоторых случаях библиотека зондов содержит по меньшей мере 1000, 2000, 5000, 10000, 50000, 100000, 200000, 500000 или более чем 1 миллион различных последовательностей. В некоторых случаях библиотека зондов содержит последовательностей, способных связываться с по меньшей мере 10, 20, 50, 100, 200, 500, 1000, 2000, 5000, 10000 или более чем 10000 генами. В некоторых случаях зонды содержат фрагмент для захвата твердой подложкой, такой как биотин. В некоторых случаях стадия обогащения происходит после стадии РТА. В некоторых случаях стадия обогащения предшествует стадии РТА. В некоторых случаях зонды сконфигурированы для связывания библиотек геномной ДНК. В некоторых случаях зонды сконфигурированы для связывания библиотек кДНК.

[00115] Библиотеки ампликонов полинуклеотидов, созданные с помощью способов РТА и композиций (терминаторы, полимеразы и т.д.), описанных в настоящем документе, в некоторых случаях характеризуются увеличенной однородностью. Однородность в некоторых случаях описывают с использованием кривой Лоренца (например, фиг. 5С) или другого такого способа. Такие увеличения в некоторых случаях приводят к пониженному количеству прочтений секвенирования, необходимых для требуемого охвата целевой молекулы нуклеиновой кислоты (например, геномной ДНК, РНК или другой целевой молекулы нуклеиновой кислоты). Например, не более 50% кумулятивной доли полинуклеотидов содержат последовательности, составляющие по меньшей мере 80% кумулятивной доли последовательностей целевой молекулы нуклеиновой кислоты. В некоторых случаях не более 50% кумулятивной доли полинуклеотидов содержат последовательности, составляющие по меньшей мере 60% кумулятивной доли последовательностей целевой молекулы нуклеиновой кислоты. В некоторых случаях не более 50% кумулятивной доли полинуклеотидов содержат последовательности, составляющие по меньшей мере 70% кумулятивной доли последовательностей целевой молекулы нуклеиновой кислоты. В некоторых случаях не более 50% кумулятивной доли полинуклеотидов содержат последовательности,

составляющие по меньшей мере 90% кумулятивной доли последовательностей целевой молекулы нуклеиновой кислоты. В некоторых случаях однородность описывают с использованием коэффициента Джинни (где коэффициент, составляющий 0, представляет идеальную равномерность библиотеки, а коэффициент, составляющий 1, представляет идеальную неравномерность). В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,55, 0,50, 0,45, 0,40 или 0,30. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,50. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,40. Такие количественные показатели однородности в некоторых случаях зависят от количества полученных прочтений. Например, получают не более 100 млн., 200 млн., 300 млн., 400 млн. или не более 500 млн. прочтений. В некоторых случаях длина прочтения составляет приблизительно 50, 75, 100, 125, 150, 175, 200, 225 или приблизительно 250 оснований в длину. В некоторых случаях количественные показатели однородности зависят от глубины охвата целевой нуклеиновой кислоты. Например, средняя глубина охвата составляет приблизительно 10X, 15X, 20X, 25X или приблизительно 30X. В некоторых случаях средняя глубина охвата составляет 10-30X, 20-50X, 5-40X, 20-60X, 5-20X или 10-20X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,55, где получали приблизительно 300 млн. прочтений. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,50, где получали приблизительно 300 млн. прочтений. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,45, где получали приблизительно 300 млн. прочтений. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,55, где получали не более 300 млн. прочтений. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,50, где получали не более 300 млн. прочтений. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,45, где получали не более 300 млн. прочтений. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,55, где средняя глубина охвата секвенирования составляет

приблизительно 15X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,50, где средняя глубина охвата секвенирования составляет приблизительно 15X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,45, где средняя глубина охвата секвенирования составляет приблизительно 15X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,55, где средняя глубина охвата секвенирования составляет по меньшей мере 15X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,50, где средняя глубина охвата секвенирования составляет по меньшей мере 15X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,45, где средняя глубина охвата секвенирования составляет по меньшей мере 15X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,55, где средняя глубина охвата секвенирования составляет не более 15X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,50, где средняя глубина охвата секвенирования составляет не более 15X. В некоторых случаях библиотеки ампликонов, описанные в настоящем документе, характеризуются коэффициентом Джинни, составляющим не более 0,45, где средняя глубина охвата секвенирования составляет не более 15X. Однородные библиотеки ампликонов, созданные с использованием способов, описанных в настоящем документе, в некоторых случаях подвергают дополнительным стадиям, таким как лигирование адаптера и дальнейшая ПЦР-амплификация. В некоторых случаях такие дополнительный стадии предшествуют стадии секвенирования.

[00116] Праймеры содержат нуклеиновые кислоты, используемые для праймирования описанных в настоящем документе реакций амплификации. Такие праймеры в некоторых случаях включают без ограничения случайные дезоксинуклеотиды любой длины с модификациями или без них, чтобы сделать их устойчивыми к экзонуклеазам, случайные рибонуклеотиды любой длины с модификациями или без них, чтобы сделать их устойчивыми к экзонуклеазам, модифицированные нуклеиновые кислоты, такие как запертые нуклеиновые кислоты, ДНК- или РНК-праймеры, которые нацелены на конкретную область генома, и реакции, которые праймированы ферментами, такими как праймаза. В случае полногеномной РТА предпочтительно использовать набор

праймеров, характеризующихся случайными или частично случайными нуклеотидными последовательностями. В образце нуклеиновой кислоты значительной сложности не обязательно знать конкретные последовательности нуклеиновой кислоты, присутствующие в образце, и нет необходимости конструировать праймеры так, чтобы они являлись комплементарными какой-либо конкретной последовательности. Скорее, сложность образца нуклеиновой кислоты приводит к большому количеству различных целевых последовательностей гибридизации в образце, которые будут комплементарны различным праймерам случайной или частично случайной последовательности. Комплементарная часть праймеров для использования в РТА в некоторых случаях полностью рандомизирована, содержит только часть, которая является рандомизированной, или иным образом выборочно рандомизирована. Количество случайных положений оснований в комплементарной части праймеров в некоторых случаях, например, составляет от 20% до 100% от общего количества нуклеотидов в комплементарной части праймеров. В некоторых случаях количество случайных положений оснований в комплементарной части праймеров составляет 10% - 90%, 15 - 95%, 20% - 100%, 30% - 100%, 50% -100%, 75 - 100% или 90 - 95% от общего количества нуклеотидов в комплементарной части праймеров. В некоторых случаях количество случайных положений оснований в комплементарной части праймеров составляет по меньшей мере 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% или не менее 90% от общего количества нуклеотидов в комплементарной части праймеров. Наборы праймеров, характеризующихся случайными или частично случайными последовательностями, в некоторых случаях синтезируют с использованием стандартных методик, позволяющих рандомизировать добавление любого нуклеотида в каждом положении. В некоторых случаях наборы праймеров состоят из праймеров одинаковой длины и/или гибридизационных характеристик. В некоторых случаях термин «случайный праймер» относится к праймеру, который может проявлять четырехкратную вырожденность в каждом положении. В некоторых случаях термин «случайный праймер» относится к праймеру, который может проявлять трехкратную вырожденность в каждом положении. Случайные праймеры, используемые в описанных в настоящем документе способах, в некоторых случаях содержат случайную последовательность, которая составляет 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 или более оснований в длину. В некоторых случаях праймеры содержат случайные последовательности длиной 3-20, 5-15, 5-20, 6-12 или 4-10 оснований. Праймеры могут также содержать не удлиняемые элементы, которые ограничивают последующую амплификацию созданных ими ампликонов. Например, праймеры с не удлиняемыми элементами в некоторых случаях содержат терминаторы. В некоторых случаях праймеры содержат терминаторные

нуклеотиды, например, 1, 2, 3, 4, 5, 10 или более 10 терминаторных нуклеотидов. Праймеры не должны ограничиваться компонентами, которые добавляются извне в реакцию амплификации. В некоторых случаях праймеры создаются in situ путем добавления нуклеотидов и белков, которые способствуют праймированию. Например, ферменты, подобные праймазе, в комбинации с нуклеотидами в некоторых случаях используют для создания случайных праймеров для описанных в настоящем документе способов. Ферменты, подобные праймазе, в некоторых случаях являются представителями суперсемейства ферментов DnaG или AEP. В некоторых случаях фермент, подобный праймазе, представляет собой TthPrimPol. В некоторых случаях праймазоподобным ферментом является хеликаза-праймаза T7 gp4. Такие праймазы в некоторых случаях используют с полимеразами или факторами замещения цепи, описанными в настоящем документе. В некоторых случаях праймазы инициируют праймирование дезоксирибонуклеотидами. В некоторых случаях праймазы инициируют праймирование рибонуклеотидами.

[00117] За амплификацией PTA может следовать отбор в отношении определенного подмножества ампликонов. Такой отбор в некоторых случаях зависит от размера, аффинности, активности, гибридизации с зондами или другого известного фактора отбора в настоящей области техники. В некоторых случаях отбор предшествует или следует за дополнительными стадиями, описанными в настоящем документе, такими как лигирование адаптера и/или амплификация библиотеки. В некоторых случаях отбор основан на размере (длине) ампликонов. В некоторых случаях выбирают меньшие ампликоны, которые с меньшей вероятностью подверглись экспоненциальной амплификации, которая обогащает продуктами, полученными из первичной матрицы, при дальнейшем преобразовании процесса амплификации из экспоненциальной в квазилинейный процесс амплификации (фиг. 1A). В некоторых случаях выбирают ампликоны, содержащие 50-2000, 25-5000, 40-3000, 50-1000, 200-1000, 300-1000, 400-1000, 400-600, 600-2000 или 800-1000 оснований в длину. Выбор размера в некоторых случаях происходит с использованием протоколов, например, с использованием твердофазной обратимой иммобилизации (SPRI) на карбоксилированных парамагнитных шариках для обогащения фрагментами нуклеиновых кислот определенных размеров или другого протокола, известного специалистам в настоящей области техники. Необязательно или в комбинации, отбор происходит посредством предпочтительного лигирования и амплификации более мелких фрагментов во время ПЦР при подготовке библиотек для секвенирования, а также в результате предпочтительного образования кластеров из более мелких фрагментов библиотеки для секвенирования во время секвенирования (например,

секвенирование путем синтеза, секвенирование с помощью нанопор, или другой способ секвенирования). Другие стратегии отбора более мелких фрагментов также согласуются со способами, описанными в настоящем документе, и включают, помимо прочего, выделение фрагментов нуклеиновых кислот определенных размеров после гель-электрофореза, использование колонок с силикагелем, которые связывают фрагменты нуклеиновых кислот определенных размеров и использование других стратегий ПЦР, которые более сильно обогащают более мелкие фрагменты. Любое число протоколов подготовки библиотеки может быть использовано с методами PTA, описанными в настоящем документе. Ампликоны, полученные с помощью PTA, в некоторых случаях лигируют с адаптерами (необязательно с удалением терминаторных нуклеотидов). В некоторых случаях ампликоны, генерируемые PTA, содержат области гомологии, получаемые в результате фрагментации на основе транспозазы, которые используются в качестве сайтов праймирования. В некоторых случаях библиотеки готовят путем механической или ферментативной фрагментации нуклеиновых кислот. В некоторых случаях библиотеки готовят с использованием тегирования через транспозомы. В некоторых случаях библиотеки готовят путем лигирования адаптеров, таких как Y-адаптеры, универсальные адаптеры или круглые адаптеры.

[00118] Некомплементарная часть праймера, используемого в PTA, может включать последовательности, которые можно использовать для дальнейших манипуляций и/или анализа амплифицированных последовательностей. Примером такой последовательности является «детекторная метка». Детекоторные метки содержат последовательности, комплементарные детекторным зондам, и выявляются с помощью родственных детекторных зондов. На праймере может присутствовать одна, две, три, четыре или более четырех детекторных меток. Не существует фундаментальных ограничений на количество детекторных меток, которые могут присутствовать на праймере, за исключением размера праймера. В некоторых случаях на праймере присутствует одна детекторная метка. В некоторых случаях на праймере присутствуют две детекторные метки. Когда присутствуют несколько детекторных меток, они могут характеризоваться одинаковой последовательностью или могут характеризоваться разными последовательностями, причем каждая отличающаяся последовательность комплементарна другому детекторному зонду. В некоторых случаях несколько детекторных меток характеризуются одинаковой последовательностью. В некоторых случаях несколько детекторных меток характеризуются разной последовательностью.

[00119] Другим примером последовательности, которая может быть включена в некомплементарную часть праймера, является «адресная метка», которая может

кодировать другие признаки ампликонов, такие как расположение в части ткани. В некоторых случаях баркод клетки содержит адресную метку. Адресная метка характеризуется последовательностью, комплементарной адресному зонду. Адресные метки включаются на концах амплифицированных цепей. Если они присутствует, то на праймере может находиться одна или более адресных меток. Нет никаких фундаментальных ограничений на количество адресных меток, которые могут присутствовать в праймере, за исключением размера праймера. Когда присутствуют несколько адресных меток, они могут характеризоваться одинаковой последовательностью или могут характеризоваться разными последовательностями, причем каждая отличающаяся последовательность комплементарна другому адресному зонду. Часть адресной метки может характеризоваться любой длиной, которая поддерживает конкретную и стабильную гибридизацию между адресной меткой и адресным зондом. В некоторых случаях нуклеиновые кислоты из более чем одного источника могут включать в себя вариабельную последовательность метки. Эта последовательность метки может характеризоваться длиной до 100 нуклеотидов, предпочтительно от 1 до 10 нуклеотидов в длину, наиболее предпочтительно от 4, 5 или 6 нуклеотидов в длину и содержит комбинации нуклеотидов. В некоторых случаях последовательность метки составляет 1-20, 2-15, 3-13, 4-12, 5-12 или 1-10 нуклеотидов в длину. Например, если для формирования метки выбирают шесть пар оснований и используют перестановку четырех разных нуклеотидов, то можно создать в общей сложности 4096 якорей нуклеиновых кислот (например, шпильки), каждый с уникальной меткой из 6-ти оснований.

[00120] Описанные в настоящем документе праймеры могут присутствовать в растворе или быть иммобилизованы на твердой подложке. В некоторых случаях праймеры, несущие баркоды образца и/или последовательности UMI, могут быть иммобилизованы на твердой подложке. Твердая подложка может представлять собой, например, один или несколько шариков. В некоторых случаях отдельные клетки контактируют с одним или более шариками, содержащими уникальный набор баркодов образцов и/или последовательностей UMI, чтобы идентифицировать индивидуальную клетку. В некоторых случаях лизаты из отдельных клеток контактируют с одним или более шариками, содержащими уникальный набор баркодов образца и/или последовательностей UMI, чтобы идентифицировать отдельные лизаты клеток. В некоторых случаях экстрагированная нуклеиновая кислота из отдельных клеток контактирует с одним или более шариками, содержащими уникальный набор баркодов образцов и/или последовательностей UMI, чтобы идентифицировать экстрагированную нуклеиновую кислоту из отдельной клетки. С шариками можно проводить манипуляции любым подходящим способом, известным в

настоящей области техники, например, с использованием капельных приводов, как описано в настоящем документе. Шарики могут быть любого подходящего размера, включая в себя, например, микрошарики, микрочастицы, наночастицы и наночастицы. Согласно некоторым вариантам осуществления шарики обладают магнитной чувствительностью, согласно другим вариантам осуществления шарики не обладают значительной магнитной чувствительностью. Неограничивающие примеры подходящих шариков включают микрошарики для проточной цитометрии, микрочастицы и наночастицы полистирола, функционализированные микрочастицы и наночастицы полистирола, микрочастицы и наночастицы полистирола с покрытием, микрошарики кремнезема, флуоресцентные микросферы и наносферы, функционализированные флуоресцентные микросферы и наносферы, флуоресцентные микросферы и наносферы с покрытием, окрашенные микрочастицы и наночастицы, магнитные микрочастицы и наночастицы, суперпарамагнитные микрочастицы и наночастицы (например, DYNABEADS®, поставляемые Invitrogen Group, Carlsbad, CA), флуоресцентные микрочастицы и наночастицы, магнитные микрочастицы и наночастицы с покрытием, ферромагнитные микрочастицы и наночастицы, ферромагнитные микрочастицы и наночастицы с покрытием и описанные в патентных публикациях США №№ US20050260686, US20030132538, US20050118574, 20050277197, 20060159962. Шарики могут быть предварительно связаны с антителом, белком или антигеном, ДНК/РНК-зондом или любой другой молекулой, обладающей аффинностью к требуемой мишени. Согласно некоторым вариантам осуществления праймеры, несущие баркоды образцов и/или последовательности UMI, могут находиться в растворе. Согласно определенным вариантам осуществления может быть представлено множество капель, при этом каждая капля во множестве несет баркод образца, который является уникальным для капли, и UMI, который является уникальным для молекулы, так что UMI повторяются много раз в пределах набора капель. Согласно некоторым вариантам осуществления отдельные клетки контактируют с каплей с уникальным набором баркодов образца и/или последовательностей UMI, чтобы идентифицировать индивидуальную клетку. Согласно некоторым вариантам осуществления лизаты отдельных клеток контактируют с каплей с уникальным набором баркодов образца и/или последовательностей UMI, чтобы идентифицировать лизаты отдельных клеток. Согласно некоторым вариантам осуществления экстрагированная нуклеиновая кислота из отдельных клеток контактирует с каплей с уникальным набором баркодов образца и/или последовательностей UMI, чтобы идентифицировать экстрагированную нуклеиновую кислоту из отдельной клетки.

[00121] Праймеры PTA могут содержать специфический для последовательности или случайный праймер, баркод клетки и/или уникальный молекулярный идентификатор (UMI) (см., например, фиг.10A (линейный праймер) и 10B (шпилечный праймер)). В некоторых случаях праймер включает праймер, специфический для последовательности. В некоторых случаях праймер включает случайный праймер. В некоторых случаях праймер содержит баркод клетки. В некоторых случаях праймер содержит баркод образца. В некоторых случаях праймер содержит уникальный молекулярный идентификатор. В некоторых случаях праймеры содержат два или более баркода клетки. Такие баркоды в некоторых случаях идентифицируют уникальный источник образца или уникальный рабочий процесс. Такие баркоды или UMI в некоторых случаях составляют 5, 6, 7, 8, 9, 10, 11, 12, 15, 20, 25, 30 или более 30 оснований в длину. Праймеры в некоторых случаях содержат по меньшей мере 1000, 10000, 50000, 100000, 250000, 500000, 10^6, 10^7, 10^8, 10^9 или по меньшей мере 10^{10} уникальных баркодов или UMI. В некоторых случаях праймеры содержат по меньшей мере 8, 16, 96 или 384 уникальных баркода или UMI. В некоторых случаях стандартный адаптор затем лигируют с продуктами амплификации перед секвенированием, после секвенирования прочтения сначала присваивают конкретной клетке на основе баркода клетки. Подходящие адаптеры, которые можно использовать вместе со способом PTA, включают, например, адаптеры xGen® Dual Index UMI, доступные от Integrated DNA Technologies (IDT). Затем прочтения из каждой клетки группируются с использованием UMI, и прочтения с одним и тем же UMI могут быть свернуты в «консенсусное прочтение». Использование баркода клетки позволяет объединить все клетки перед подготовкой библиотеки, так как позже они могут быть идентифицированы по баркоду клетки. Использование UMI для формирования консенсусного прочтения в некоторых случаях корректирует смещение ПЦР, улучшая обнаружение вариации числа копий (CNV) (фиг. 11A и 11B). Кроме того, ошибки секвенирования можно скорректировать, потребовав, чтобы фиксированный процент прочтений одной и той же молекулы содержал одну и ту же замену основания, обнаруженное в каждом положении. Этот подход был использован для улучшения обнаружения CNV и исправления ошибок секвенирования в совокупных образцах. В некоторых случаях UMI используют с описанными в настоящем документе способами, например, в патенте США 8835358 раскрыт принцип цифрового подсчета после прикрепления случайного амплифицируемого баркода. Schmitt. et al and Fan et al. раскрывают подобные способы коррекции ошибок секвенирования. Согласно некоторым вариантам осуществления библиотеку получают путем секвенирования с использованием праймеров. Согласно некоторым вариантам осуществления библиотека содержит

фрагменты из 200-700 оснований, 100-1000, 300-800, 300-550, 300-700 или 200-800 оснований в длину. Согласно некоторым вариантам осуществления библиотека содержит фрагменты из по меньшей мере 50, 100, 150, 200, 300, 500, 600, 700, 800 или по меньшей мере 1000 оснований в длину. Согласно некоторым вариантам осуществления библиотека содержит фрагменты из приблизительно 50, 100, 150, 200, 300, 500, 600, 700, 800 или приблизительно 1000 оснований в длину.

[00122] Описанные в настоящем документе способы могут дополнительно предусматривать дополнительные стадии, включая стадии, выполняемые с образцом или матрицей. Такие образцы или матрицы в некоторых случаях подвергают одной или нескольких стадиям перед PTA. В некоторых случаях образцы, содержащие клетки, подвергают стадии предварительной обработки. Например, клетки подвергают лизису и протеолизу для увеличения доступности хроматина с использованием комбинации замораживания- оттаивания, Triton X-100, Tween 20 и протеиназы К. Другие стратегии лизиса также подходят для применения на практике описанных в настоящем документе способов. Такие стратегии включают без ограничения лизис с использованием других комбинаций обработки детергентом, и/или лизоцимом, и/или протеазой, и/или физическое разрушение клеток, такое как обработка ультразвуком и/или щелочной лизис и/или гипотонический лизис. В некоторых случаях первичную матрицу или целевую молекулу (молекулы) подвергают стадии предварительной обработки. В некоторых случаях первичную матрицу (или целевую молекулу) денатурируют с использованием гидроксида натрия с последующей нейтрализацией раствора. Другие стратегии денатурирования также могут быть подходящими для применения на практике описанных в настоящем документе способов. Такие стратегии могут включать в себя без ограничения комбинации щелочного лизиса с другими основными растворами, повышение температуры образца и/или изменение концентрации соли в образце, добавление добавок, таких как растворители или масла, другую модификацию или любую их комбинацию. В некоторых случаях дополнительные стадии включают сортировку, фильтрацию или выделение образцов, матриц или ампликонов по размеру. В некоторых случаях клетки лизируют механическими (например, гомогенизатор высокого давления, помол в шаровой мельнице) или немеханическими (физическими, химическими или биологическими) способами. В некоторых случаях способ физического лизиса предусматривает нагревание, осмотический шок и/или кавитацию. В некоторых случаях химический лизис включает щелочь и/или детергенты. В некоторых случаях биологический лизис включает использование ферментов. Комбинации способов лизиса также совместимы с описанными в настоящем документе способами. Неограничивающие примеры ферментов лизиса включают

рекомбинантный лизоцим, сериновые протеазы и бактериальные лизины. В некоторых случаях лизис ферментами предусматривает использование лизоцима, лизостафина, зимолазы, целлюлозы, протеазы или гликаназы. Например, после амплификации описанными в настоящем документе способами библиотеки ампликонов обогащаются ампликонами желаемой длины. В некоторых случаях библиотеки ампликонов обогащают ампликонами с длиной, составляющей 50-2000, 25-1000, 50-1000, 75-2000, 100-3000, 150-500, 75-250, 170-500, 100- 500 или 75-2000 оснований. В некоторых случаях библиотеки ампликонов обогащают ампликонами с длиной не более 75, 100, 150, 200, 500, 750, 1000, 2000, 5000 или не более 10000 оснований. В некоторых случаях библиотеки ампликонов обогащают ампликонами с длиной, составляющей по меньшей мере 25, 50, 75, 100, 150, 200, 500, 750, 1000 или по меньшей мере 2000 оснований.

[00123] Способы и композиции, описанные в настоящем документе, могут включать буферы или другие составы. Такие буферы в некоторых случаях используют для PTA, RT или других описанных в настоящем документе способов. Такие буферы в некоторых случаях содержат поверхностно-активные вещества/детергенты или денатурирующие агенты (Твин-20, DMSO, DMF, пегилированные полимеры, содержащие гидрофобную группу, или другое поверхностно-активное вещество), соли (фосфат калия или натрия (одноосновный или двухосновный), хлорид натрия, хлорид калия, TrisHCl, хлорид или сульфат магния, соли аммония, такие как фосфат, нитрат или сульфат, EDTA), восстановители (DTT, THP, DTE, бета-меркаптоэтанол, TCEP или другой восстановитель) или другие компоненты (глицерин, гидрофильные полимеры, такие как PEG). В некоторых случаях буферы используют в сочетании с такими компонентами, как полимеразы, факторы замещения цепи, терминаторы или другой компонент реакции, описанными в настоящем документе. Согласно некоторым вариантам осуществления буферы используют в сочетании с такими компонентами, как полимеразы, факторы смещения цепи, терминаторы или другие описанные в настоящем документе компоненты реакции. Буферы могут содержать один или несколько агентов краудинга. В некоторых случаях реагенты краудинга включают полимеры. В некоторых случаях реагенты краудинга включают полимеры, такие как полиолы. В некоторых случаях реагенты краудинга включают полимеры полиэтиленгликоля (PEG). В некоторых случаях реагенты краудинга содержат полисахариды. Без ограничения примеры реагентов краудинга включают фиколл (например, фиколл PM 400, фиколл PM 70 или фиколл с другой молекулярной массой), PEG (например, PEG1000, PEG 2000, PEG4000, PEG6000, PEG8000 или PEG с другой молекулярной массой), декстран (декстран 6, декстран 10, декстран 40, декстран 70, декстран 6000, декстран 138k или декстран с другой молекулярной массой).

[00124] Молекулы нуклеиновой кислоты, амплифицированные описанными в настоящем документе способами, можно секвенировать и проанализировать с использованием способов, известных специалистам в настоящей области техники. Неограничивающие примеры способов секвенирования, которые в некоторых случаях используют, включают, например, секвенирование путем гибридизации (SBH), секвенирование путем лигирования (SBL) (Shendure et al. (2005) Science 309:1728), секвенирование с количественным последовательным добавлением флуоресцентных нуклеотидов (QIFNAS), пошаговое лигирование и расщепление, резонансный перенос энергии флуоресценции (FRET), молекулярные маяки, расщепление репортерного зонда TaqMan, пиросеквенирование, флуоресцентное секвенирование in situ (FISSEQ), шарики FISSEQ (патент США № 7425431), wobble-секвенирование (публикация международной патентной заявки № W02006/073504), мультиплексное секвенирование (публикация заявки на выдачу патента США № US2008/0269068, Porreca et al., 2007, Nat. Methods 4:931), секвенирование полимеризованных колоний (POLONY) (патенты США №№ 6432360, 6485944 и 6511803, и публикация международной патентной заявки № W02005/082098), секвенирование по методу катящегося кольца с использованием наносетки (ROLONY) (патент США № 9624538), аллель-специфические анализы лигирования олигонуклеотидов (например, анализ лигирования олигонуклеотидов (OLA), OLA одиночной матричной молекулы с использованием лигированного линейного зонда и считывание результатов амплификации по методу катящегося кольца (RCA), лигированные замыкающих кольцо зондов и/или OLA одиночной матричной молекулы с использованием лигированного циклического замыкающего кольцо зонда и считывания амплификации по методу катящегося кольца (RCA)), высокопроизводительные способы секвенирования, такие как, например, способы с использованием платформ Roche 454, Illumina Solexa, AB-SOLiD, Hélicos, Polonator и т.п., а также технологии секвенирования на основе использования света (Landegren et al. (1998) Genome Res. 8:769-76, Kwok (2000) Pharmacogenomics 1:95-100, и Shi (2001) Clin. Chem. 0,47:164-172). В некоторых случаях амплифицированные молекулы нуклеиновых кислот секвенируют методом «дробовика». Секвенирование библиотеки секвенирования в некоторых случаях выполняют с использованием любой подходящей технологии секвенирования, включая, помимо прочего, секвенирование одной молекулы в реальном времени (SMRT), секвенирование Polony, секвенирование путем лигирования, секвенирование с обратимым терминатором, секвенирование с обнаружением протонов, секвенирование с ионным полупроводником, нанопоровое секвенирование, электронное секвенирование, пиросеквенирование, секвенирование Максама-Гилберта, секвенирование

обрыва цепи (например, по Сэнгеру), секвенирование +S или секвенирование путем синтеза (на основе матрицы/колонии или на основе наношариков).

[00125] Библиотеки секвенирования, полученные с использованием способов, описанных в настоящем документе (например, PTA или RNAseq), могут быть секвенированы с получением желаемого числа прочтений секвенирования. В некоторых случаях библиотеки получают из одной клетки или образца, содержащего одну клетку (отдельно или в виде части мультиомической схемы). В некоторых случаях библиотеки секвенируют с получением по меньшей мере 0,1, 0,2, 0,4, 0,5, 0,7, 0,8, 0,9, 1, 1,1, 1,2, 1,5, 2, 5 или по меньшей мере 10 миллионов прочтений. В некоторых случаях библиотеки секвенируют с получением не более 0,1, 0,2, 0,4, 0,5, 0,7, 0,8, 0,9, 1, 1,1, 1,2, 1,5, 2, 5 или не более 10 миллионов прочтений. В некоторых случаях библиотеки секвенируют с получением приблизительно 0,1, 0,2, 0,4, 0,5, 0,7, 0,8, 0,9, 1, 1,1, 1,2, 1,5, 2, 5 или приблизительно 10 миллионов прочтений. В некоторых случаях библиотеки секвенируют с получением 0,1-10, 0,1-5, 0,1-1, 0,2-1, 0,3-1,5, 0,5-1, 1-5 или 0,5-5 миллионов прочтений на образец. В некоторых случаях число прочтений зависит от размера генома. В некоторых случаях образцы, содержащие бактериальные геномы, секвенируют с получением 0,5-1 миллионов прочтений. В некоторых случаях библиотеки секвенируют с получением по меньшей мере 2, 4, 10, 20, 50, 100, 200, 300, 500, 700 или по меньшей мере 900 миллионов прочтений. Согласно некоторым вариантам осуществления библиотеки секвенируют с получением не более 2, 4, 10, 20, 50, 100, 200, 300, 500, 700 или не более 900 миллионов прочтений. В некоторых случаях библиотеки секвенируют с получением приблизительно 2, 4, 10, 20, 50, 100, 200, 300, 500, 700 или приблизительно 900 миллионов прочтений. В некоторых случаях библиотеки, содержащие геномы млекопитающих, секвенируют с получением 500-600 миллионов прочтений. В некоторых случаях тип библиотеки секвенирования (библиотеки кДНК или геномные библиотеки) определяют во время секвенирования. В некоторых случаях библиотеки кДНК и геномные библиотеки идентифицируют во время секвенирования с помощью уникальных баркодов.

[00126] Термин «цикл», когда он используется в отношении реакции амплификации, опосредованной полимеразой, используют в настоящем документе для описания стадий диссоциации по меньшей мере части двухцепочечной нуклеиновой кислоты (например, матрица из ампликона или двухцепочечная матрица, денатурация), гибридизации по меньшей мере части праймера с матрицей (отжиг) и удлинения праймера с получением ампликона. В некоторых случаях температура остается постоянной во время цикла амплификации (например, при изотермической реакции). В некоторых случаях количество циклов прямо коррелирует с количеством произведенных ампликонов. В

некоторых случаях количество циклов изотермической реакции регулируется количеством времени, в течение которого реакция может протекать.

[00127] **Способы и применения**

[00128] Описанные в настоящем документе способы идентифицируют мутации в клетках посредством способов мультиомического анализа PTA, таких как отдельные клетки. Применение способа PTA в некоторых случаях приводит к улучшениям по сравнению с известными способами, например, MDA. PTA в некоторых случаях характеризуется более низкой частотой обнаружения ложноположительных и ложноотрицательных вариантов, чем способ MDA. Геномы, такие как платиновые геномы NA12878, в некоторых случаях используют для определения того, приведет ли больший охват генома и однородность PTA к снижению частоты обнаружения ложноположительных вариантов. Без ограничения какой-либо теорией можно определить, что отсутствие распространения ошибок в PTA снижает частоту обнаружения ложноположительных вариантов. Баланс амплификации между аллелями с помощью этих двух способов в некоторых случаях оценивают путем сравнения частот аллелей обнаружений гетерозиготных мутаций в известных положительных локусах. В некоторых случаях библиотеки ампликонов, созданные с использованием PTA, дополнительно амплифицируют с помощью ПЦР. Согласно некоторым вариантам осуществления PTA применяют в схеме с дополнительными способами анализа, такими как секвенирование РНК, анализ метиломы или другой способ, описанный в настоящем документе.

[00129] Клетки, проанализированные с использованием описанных в настоящем документе способов, в некоторых случаях включают опухолевые клетки. Например, циркулирующие опухолевые клетки можно выделить из жидкости, взятой у пациентов, такой как без ограничения кровь, костный мозг, моча, слюна, спинномозговая жидкость, плевральная жидкость, перикардиальная жидкость, асцит или водянистая влага. Затем клетки подвергают описанным в настоящем документе способам (например, PTA) и секвенированию для определения бремени мутаций и комбинации мутаций в каждой клетке. Эти данные в некоторых случаях используют для диагностики конкретного заболевания или в качестве инструментов для прогнозирования ответа на лечение. Аналогичным образом, в некоторых случаях клетки с неизвестным злокачественным потенциалом в некоторых случаях выделяют из жидкости, взятой у пациентов, такой как без ограничения кровь, костный мозг, моча, слюна, спинномозговая жидкость, плевральная жидкость, перикардиальная жидкость, асцит, водянистая влага, жидкость бластоцеля или собранная среда, окружающая клетки в культуре. Согласно некоторым вариантам

осуществления образец получают из собранной среды, окружающей клетки в культуре. После использования описанных в настоящем документе способов и секвенирования такие способы в дальнейшем используют для определения бремени мутаций и комбинации мутаций в каждой клетке. Эти данные в некоторых случаях используют для диагностики конкретного заболевания или в качестве инструментов для прогнозирования прогрессирования предракового состояния в явное злокачественное новообразование. В некоторых случаях клетки можно выделить из образцов первичной опухоли. Затем клетки могут быть подвергнуты PTA и секвенированию для определения бремени мутаций и комбинации мутаций в каждой клетке. Эти данные можно использовать для диагностики конкретного заболевания или в качестве инструментов для прогнозирования вероятности того, что злокачественное новообразование пациента является устойчивым к доступным противораковым лекарственным средствам. При воздействии на образцы различных химиотерапевтических средств обнаружили, что основные и минорные клоны обладают разной чувствительностью к конкретным лекарственным средствам, которая не обязательно коррелирует с наличием известной «мутации-драйвера», что позволяет предположить, что комбинации мутаций в клональной популяции определяют ее чувствительность к конкретным химиотерапевтическим лекарственным средствам. Без ограничения какой-либо теорией эти результаты позволяют предположить, что злокачественное новообразование можно легче устранить, если обнаружены предраковые поражения, которые еще не расширились и превратились в клоны, повышенное количество модификаций генома которых может сделать их более устойчивыми к лечению. См. Ma et al., 2018, «Pan-cancer genome and transcriptome analyses of 1,699 pediatric leukemias and solid tumors». Протокол геномики одиночных клеток в некоторых случаях используют для обнаружения комбинаций соматических генетических вариантов в отдельной раковой клетке или клонотипе в смеси нормальных и злокачественных клеток, выделенных из образцов пациента. Эту технологию в некоторых случаях в дальнейшем используют для идентификации клонотипов, которые подвергаются положительному отбору после воздействия лекарственных средств как in vitro, так и в организме пациентов. Как показано на фиг. 6A, путем сравнения выживших клонов, подвергшихся химиотерапии, с клонами, идентифицированными при постановке диагноза, можно создать каталог клонотипов рака, который документирует их устойчивость к конкретным лекарственным средствам. В некоторых случаях способы PTA обнаруживают чувствительность определенных клонов в образце, состоящем из нескольких клонотипов, к существующим или новым лекарственным средствам, а также к их комбинациям, где способ может определять чувствительность конкретных клонов к лекарственному средству. Этот подход в некоторых

случаях показывает эффективность лекарственного средства для конкретного клона, который может не быть обнаружен с помощью существующих измерений чувствительности к лекарственным средствам, которые учитывают чувствительность всех раковых клонов вместе в одном измерении. Когда PTA, описанный в настоящем документе, применяют к образцам пациентов, собранным во время постановки диагноза, с целью обнаружения раковых клонотипов в раке данного пациента, каталог чувствительности к лекарственным средствам можно использовать для поиска этих клонов и, таким образом, информирования онкологов о том, какое лекарственное средство или комбинация лекарственных средств не будут работать, и какое лекарственное средство или комбинация лекарственных средств, скорее всего, будут эффективны против этого рака пациента. PTA можно применять для анализа образцов, содержащих группы клеток. В некоторых случаях образец содержит нейроны или глиальные клетки. В некоторых случаях образец содержит ядра.

[00130] В настоящем документе описаны способы измерения изменения экспрессии генов в сочетании с мутагенностью фактора окружающей среды. Например, клетки (одиночные или популяция) подвергаются потенциальному воздействию окружающей среды. Например, в некоторых случаях в способе используют клетки, происходящие из органов (печень, поджелудочная железа, легкие, толстая кишка, щитовидная железа или другой орган), тканей (кожа или другие ткани), крови или другого биологического источника. В некоторых случаях условия окружающей среды включают в себя тепло, свет (например, ультрафиолет), облучение, химическое вещество или любую их комбинацию. После определенного воздействия условий окружающей среды, в некоторых случаях минут, часов, дней или дольше, одиночные клетки выделяют и подвергают способу PTA. В некоторых случаях для маркировки образца используют молекулярные баркоды и уникальные молекулярные идентификаторы. Образец секвенируют и затем анализируют для идентификации изменений экспрессии гена и/или результата мутаций, вызванных воздействием условий окружающей среды. В некоторых случаях такие мутации сравнивают с контрольным условием окружающей среды, таким как известное немутагенное вещество, среда/растворитель или отсутствие условия окружающей среды. Такой анализ в некоторых случаях дает не только общее количество мутаций, вызванных условием окружающей среды, но также местонахождение и характер таких мутаций. В некоторых случаях из данных идентифицируют паттерны и их могут использовать для диагностики заболеваний или состояний. В некоторых случаях паттерны используют для прогнозирования будущих заболеваний или состояний. В некоторых случаях описанные в настоящем документе способы измеряют бремя мутаций, местоположения и паттерны в

клетке после воздействия агента окружающей среды, такого как, например, потенциальный мутаген или тератоген. Этот подход в некоторых случаях используют для оценки безопасности данного агента, в том числе его способности вызывать мутации, которые могут способствовать развитию заболевания. Например, способ можно использовать для прогнозирования канцерогенности или тератогенности агента в отношении определенных типов клеток после воздействия определенной концентрации определенного агента.

[00131] В настоящем документе описаны способы идентификации изменения экспрессии гена в комбинации с мутациями в клетках животных, растений или микроорганизмов, которые подверглись редактированию генома (например, с использованием технологий CRISPR). Такие клетки в некоторых случаях можно выделить и подвергнуть PTA и секвенированию для определения бремени мутаций и комбинации мутаций в каждой клетке. Частоту мутаций на клетку и локализацию мутаций, которые возникают в результате протокола редактирования генома, в некоторых случаях используют для оценки безопасности данного способа редактирования генома.

[00132] В настоящем документе описаны способы определения изменения экспрессии гена в комбинации с мутациями в клетках, которые используют для клеточной терапии, такой как без ограничения трансплантация индуцированных плюрипотентных стволовых клеток, трансплантация гемопоэтических или других клеток, которые не подвергались манипуляциям, или трансплантация гемопоэтических или других клеток, подвергшихся редактированию генома. Затем клетки могут подвергать PTA и секвенированию для определения бремени мутаций и комбинации мутаций в каждой клетке. Частоту мутаций на клетку и локализацию мутаций в продукте для клеточной терапии можно использоваться для оценки безопасности и потенциальной эффективности продукта.

[00133] Клетки для применения в способе PTA могут представлять собой фетальные клетки, такие как эмбриональные клетки. Согласно некоторым вариантам осуществления PTA используют в сочетании с неинвазивным преимплантационным генетическим тестированием (NIPGT). Согласно другому варианту осуществления клетки могут быть выделены из бластомеров, созданных в результате оплодотворения in vitro. Затем клетки могут быть подвергнуты PTA и секвенированию для определения нагрузки и комбинации потенциально предрасполагающих к заболеванию генетических вариантов в каждой клетке. Затем изменение экспрессии генов в сочетании с мутационным профилем клетки можно использовать для экстраполяции генетической предрасположенности бластомера к конкретным заболеваниям до имплантации. В некоторых случаях эмбрионы в культуре показывают нуклеиновые кислоты, которые используются для оценки здоровья

эмбриона с помощью низкочастотного секвенирования генома. В некоторых случаях эмбрионы замораживают-оттаивают. В некоторых случаях нуклеиновые кислоты получают из среды, кондиционированной для культуры бластоцитов (BCCM), жидкости бластоцеля (BF) или их комбинации. В некоторых случаях для выявления хромосомных аномалий, таких как анеплоидия плода, используется PTA-анализ фетальных клеток. В некоторых случаях PTA используют для выявления таких заболеваний, как синдромы Дауна или синдром Патау. В некоторых случаях замороженные бластоциты размораживают и культивируют в течение определенного периода времени перед получением нуклеиновых кислот для анализа (например, культуральной среды, BF или биопсии клеток). В некоторых случаях бластоциты культивируют не более 4, 6, 8, 12, 16, 24, 36, 48 или не более 64 часов до получения нуклеиновых кислот для анализа.

[00134] Согласно другому варианту осуществления клетки микроорганизмов (например, бактерии, грибы, простейшие) можно выделить из растений или животных (например, из образцов микробиоты [например, микробиоты ЖКТ, микробиоты кожи и т.д.] или из жидкостей организма, таких как, например, кровь, костный мозг, моча, слюна, спинномозговая жидкость, плевральная жидкость, перикардиальная жидкость, асцит или водянистая влага). Кроме того, клетки микроорганизмов можно выделить из постоянных изделий медицинского назначения, таких как без ограничения внутривенные катетеры, уретральные катетеры, цереброспинальные шунты, искусственные клапаны, искусственные суставы или эндотрахеальные трубки. Затем клетки можно подвергать PTA и секвенированию для определения идентичности конкретного микроорганизма, а также для обнаружения наличия генетических вариантов микроорганизмов, которые предсказывают ответ (или устойчивость) к конкретным антимикробным агентам. Эти данные можно использовать для диагностики конкретного инфекционного заболевания и/или в качестве инструментов для прогнозирования ответа на лечение.

[00135] В настоящем документе описаны способы создания библиотек ампликонов из образцов, содержащих короткую нуклеиновую кислоту, с использованием описанных в настоящем документе способов PTA. В некоторых случаях PTA приводит к повышению точности и однородности амплификации более коротких нуклеиновых кислот. В некоторых случаях длина нуклеиновых кислот не превышает 2000 оснований. В некоторых случаях длина нуклеиновых кислот не превышает 1000 оснований. В некоторых случаях длина нуклеиновых кислот не превышает 500 оснований. В некоторых случаях нуклеиновые кислоты характеризуются длиной, составляющей не более 200, 400, 750, 1000, 2000 или 5000 оснований. В некоторых случаях образцы, содержащие короткие фрагменты нуклеиновых кислот, включают в себя без ограничения древнюю ДНК (возрастом сотни,

тысячи, миллионы или даже миллиарды лет), образцы FFPE (фиксированные формалином, залитые парафином), внеклеточную ДНК или другой образец, содержащий короткие нуклеиновые кислоты.

[00136] **Варианты осуществления**

[00137] В настоящем документе описаны способы амплификации целевой молекулы нуклеиновой кислоты, причем способ предусматривает: a) приведение в контакт образца, содержащего целевую молекулу нуклеиновой кислоты, одного или нескольких праймеров для амплификации, полимеразы нуклеиновых кислот и смеси нуклеотидов, которая содержит один или несколько терминаторных нуклеотидов, которые останавливают репликацию нуклеиновой кислоты полимеразой, и b) инкубацию образца при условиях, которые стимулируют репликацию целевой молекулы нуклеиновой кислоты с получением множества терминированных продуктов амплификации, где репликация происходит путем репликации с замещением цепи. Согласно одному варианту осуществления любого из приведенных выше способов способ дополнительно предусматривает выделение из множества терминированных продуктов амплификации продуктов, которые составляют от приблизительно 50 до приблизительно 2000 нуклеотидов в длину. Согласно одному варианту осуществления любого из приведенных выше способов способ дополнительно предусматривает выделение из множества терминированных продуктов амплификации продуктов, которые составляют от приблизительно 400 до приблизительно 600 нуклеотидов в длину. Согласно одному варианту осуществления любого из приведенных выше способов способ дополнительно предусматривает: c) репарацию концов и полиаденилирование, и d) лигирование молекул, полученных на стадии (c) с адаптерами, получая таким образом библиотеку продуктов амплификации. Согласно некоторым вариантам осуществления способ дополнительно предусматривает удаление терминаторных нуклеотидов из терминированных продуктов амплификации. Согласно одному варианту осуществления любого из приведенных выше способов способ дополнительно предусматривает секвенирование продуктов амплификации. Согласно одному варианту осуществления любого из приведенных выше способов амплификацию проводят при по существу изотермических условиях. Согласно одному варианту осуществления любого из приведенных выше способов полимераза нуклеиновых кислот представляет собой ДНК-полимеразу.

[00138] Согласно одному варианту осуществления любого из приведенных выше способов ДНК-полимераза представляет собой замещающую цепи ДНК-полимеразу. Согласно одному варианту осуществления любого из приведенных выше способов

полимераза нуклеиновых кислот выбрана из полимеразы бактериофага phi29 (Ф29), генетически модифицированной ДНК-полимеразы phi29 (Ф29), фрагмента Кленова ДНК-полимеразы I, ДНК-полимеразы фага M2, ДНК-полимеразы фага phiPRDl, ДНК-полимеразы Bst, ДНК-полимеразы большого фрагмента Bst, полимеразы экзо(-) Bst, ДНК-полимеразы экзо(-) Bca, ДНК-полимеразы Bsu, ДНК-полимеразы VentR, ДНК-полимеразы VentR (экзо-), ДНК-полимеразы Deep Vent, ДНК-полимеразы Deep Vent (экзо-), ДНК-полимеразы IsoPol, ДНК-полимеразы I, ДНК-полимеразы Therminator, ДНК-полимеразы T5, секвеназы, ДНК-полимеразы T7, T7-секвеназы и ДНК-полимеразы T4. Согласно одному варианту осуществления любого из приведенных выше способов полимераза нуклеиновых кислот характеризуется 3'->5' экзонуклеазной активностью, и терминаторные нуклеотиды ингибируют такую 3'->5' экзонуклеазную активность. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из нуклеотидов с модификацией альфа-группы (например, альфа-тио-дидезоксинуклеотиды, создающие фосфоротиоатную связь), C3 спейсерных нуклеотидов, запертых нуклеиновых кислот (LNA), инвертированных нуклеиновых кислот, 2' фторнуклеотидов, 3' фосфорилированных нуклеотидов, 2'-O-метил-модифицированных нуклеотидов и транснуклеиновых кислот. Согласно одному варианту осуществления любого из приведенных выше способов полимераза нуклеиновых кислот не характеризуется 3'->5' экзонуклеазной активностью. Согласно одному конкретному варианту осуществления полимераза выбрана из ДНК-полимеразы Bst, полимеразы экзо(-) Bst, ДНК-полимеразы экзо(-) Bca, ДНК-полимеразы Bsu, ДНК-полимеразы VentR (экзо-), ДНК- полимеразы Deep Vent (экзо-), ДНК-полимеразы фрагмента Кленова (экзо-) и ДНК-полимеразы Therminator. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды содержат модификации группы г 3' углерода дезоксирибозы. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из 3' блокированного обратимого терминатора, содержащего нуклеотиды, 3' неблокированного обратимого терминатора, содержащего нуклеотиды, терминаторов, содержащих 2' модификации дезоксинуклеотидов, терминаторов, содержащих модификации азотистого основания дезоксинуклеотидов, и их комбинации. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из дидезоксинуклеотидов, инвертированных дидезоксинуклеотидов, 3' биотинилированных нуклеотидов, 3' аминонуклеотидов, 3'-фосфорилированных нуклеотидов, 3'-O-метилнуклеотидов, 3' углеродных спейсерных нуклеотидов, включая 3' C3 спейсерные нуклеотиды, 3' C18 нуклеотиды, 3' гександиоловые спейсерные нуклеотиды, ациклонуклеотиды и их комбинаций. Согласно одному варианту осуществления любого из приведенных выше

способов праймеры для амплификации составляют от 4 до 70 нуклеотидов в длину. Согласно одному варианту осуществления любого из приведенных выше способов продукты амплификации составляют от приблизительно 50 до приблизительно 2000 нуклеотидов в длину. Согласно одному варианту осуществления любого из приведенных выше способов целевая нуклеиновая кислота представляет собой ДНК (например, кДНК или геномная ДНК). Согласно одному варианту осуществления любого из приведенных выше способов праймеры для амплификации представляют собой случайные праймеры. Согласно одному варианту осуществления любого из приведенных выше способов праймеры для амплификации содержат баркод. Согласно одному конкретному варианту осуществления баркод содержит баркод клетки. Согласно одному конкретному варианту осуществления баркод содержит баркод образца. Согласно одному варианту осуществления любого из приведенных выше способов праймеры для амплификации содержат уникальный молекулярный идентификатор (UMI). Согласно одному варианту осуществления любого из приведенных выше способов способ предусматривает денатурирование целевой нуклеиновой кислоты или геномной ДНК перед начальным отжигом праймера. Согласно одному конкретному варианту осуществления денатурацию проводят в щелочных условиях с последующей нейтрализацией. Согласно одному варианту осуществления любого из приведенных выше способов образец, праймеры для амплификации, полимераза нуклеиновых кислот и смесь нуклеотидов содержатся в микрожидкостном устройстве. Согласно одному варианту осуществления любого из приведенных выше способов образец, праймеры для амплификации, полимераза нуклеиновых кислот и смесь нуклеотидов содержатся в капле. Согласно одному варианту осуществления любого из приведенных выше способов образец выбран из следующего образцов ткани (тканей), клеток, образцов биологической жидкости (например, кровь, моча, слюна, лимфатическая жидкость, спинномозговая жидкость (CSF), амниотическая жидкость, плевральная жидкость, перикардиальная жидкость, асцит, водянистая влага), образцов костного мозга, образцов спермы, биопсийных образцов, образцов рака, образцов опухоли, образцов клеточного лизата, криминалистических образцов, археологических образцов, палеонтологических образцов, образцов инфекции, производственных образцов, целых растений, частей растения, образцов микробиоты, вирусных препаратов, образцов почвы, образцов морской воды, образцов пресной воды, бытовых или промышленных образцов и их комбинаций и изолятов. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой клетку (например, клетка животного [например, клетка человека], растительная клетка, грибковая клетка, бактериальная клетка и клетка простейшего). Согласно одному конкретному варианту

осуществления клетку лизируют перед репликацией. Согласно одному конкретному варианту осуществления клеточный лизис сопровождается протеолизом. Согласно одному конкретному варианту осуществления клетка выбрана из клетки из предимплантационного эмбриона, стволовой клети, фетальной клетки, опухолевой клетки, предполагаемой раковой клетки, раковой клетки, клетки, подвергнутой процедуре редактирования генов, клетки из патогенного организма, клетки, полученной из криминалистического образца, клетки, полученной из археологического образца, и клетки, полученной из палеонтологического образца. Согласно одному варианту осуществления любого из описанных выше способов образец представляет собой клетку из предимплантационного эмбриона (например, бластомер [например, бластомер, полученный из эмбриона восьмиклеточной стадии, полученного с помощью in vitro оплодотворения]). Согласно одному конкретному варианту осуществления способ дополнительно включает определение наличия предрасполагающих к заболеванию вариантов зародышевой линии или соматических вариантов в эмбриональной клетке. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой клетку из патогенного организма (например, бактерии, гриба, простейшего). Согласно одному конкретному варианту осуществления клетку патогенного организма получают из жидкости, взятой у пациента, образца микробиоты (например, образец микробиоты ЖКТ, образец вагинальной микробиоты, образец микробиоты кожи и т.д.) или постоянного изделия медицинского назначения (например, внутривенный катетер, уретральный катетер, цереброспинальный шунт, искусственный клапан, искусственный сустав, эндотрахеальная трубка и т.д.). Согласно одному конкретному варианту осуществления способ дополнительно включает стадию определения идентичности патогенного организма. Согласно одному конкретному варианту осуществления способ дополнительно предусматривает определение наличия генетических вариантов, ответственных за устойчивость патогенного организма к лечению. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой опухолевую клетку, предполагаемую раковую клетку или раковую клетку. Согласно одному конкретному варианту осуществления способ дополнительно предусматривает определение наличия одной или нескольких диагностических или прогностических мутаций. Согласно одному конкретному варианту осуществления способ дополнительно предусматривает определение наличия вариантов зародышевой линии или соматических вариантов, ответственных за устойчивость к лечению. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой клетку, подвергнутую процедуре редактирования генов. Согласно одному конкретному варианту осуществления способ дополнительно предусматривает

определение наличия незапланированных мутаций, вызванных процессом редактирования генов. Согласно одному варианту осуществления любого из приведенных выше способов способ дополнительно предусматривает определение истории клеточной родословной. Согласно родственному аспекту настоящее изобретение относится к применению любого из приведенных выше способов для идентификации низкочастотных вариантов последовательности (например, вариантов, которые составляют >0,01% от общего числа последовательностей).

[00139] Согласно родственному аспекту настоящее изобретение относится к набору, содержащему полимеразу нуклеиновых кислот, один или несколько праймеров для амплификации, смесь нуклеотидов, содержащую один или несколько терминаторных нуклеотидов, и необязательно инструкции по применению. Согласно одному варианту осуществления наборов согласно настоящему изобретению полимераза нуклеиновых кислот представляет собой замещающую цепи ДНК-полимеразу. Согласно одному варианту осуществления наборов согласно настоящему изобретению полимераза нуклеиновых кислот представляет собой замещающую цепи ДНК-полимеразу. Согласно одному варианту осуществления наборов согласно настоящему изобретению полимераза нуклеиновых кислот выбрана из полимеразы бактериофага phi29 (Ф29), генетически модифицированной ДНК-полимеразы phi29 (Ф29), фрагмента Кленова ДНК-полимеразы I, ДНК-полимеразы фага M2, ДНК-полимеразы фага phiPRDl, ДНК-полимеразы Bst, ДНК-полимеразы большого фрагмента Bst, полимеразы экзо(-) Bst, ДНК-полимеразы экзо(-) Bca, ДНК-полимеразы Bsu, ДНК-полимеразы VentR, ДНК-полимеразы VentR (экзо-), ДНК-полимеразы Deep Vent, ДНК-полимеразы Deep Vent (экзо-), ДНК-полимеразы IsoPol, ДНК-полимеразы I, ДНК-полимеразы Therminator, ДНК-полимеразы T5, секвеназы, ДНК-полимеразы T7, T7-секвеназы и ДНК-полимеразы T4. Согласно одному варианту осуществления наборов согласно настоящему изобретению полимераза нуклеиновых кислот характеризуется 3'->5' экзонуклеазной активностью, и терминаторные нуклеотиды, ингибирующие такую 3'->5' экзонуклеазную активность (например, нуклеотиды с модификацией альфа-группы [например, альфа-тио-дидезоксинуклеотиды], СЗ спейсерные нуклеотиды, запертые нуклеиновые кислоты (LNA), инвертированные нуклеиновые кислоты, 2' фторнуклеотиды, 3' фосфорилированные нуклеотиды, 2'-0-метил-модифицированные нуклеотиды, транснуклеиновые кислоты). Согласно одному варианту осуществления наборов согласно настоящему изобретению полимераза нуклеиновых кислот не характеризуется 3'->5' экзонуклеазной активностью (например, ДНК-полимераза Bst, полимераза экзо(-) Bst, ДНК-полимераза экзо(-) Bca, ДНК-полимераза Bsu, ДНК-полимераза VentR (экзо-), ДНК-полимераза Deep Vent (экзо-), ДНК-полимераза фрагмента

Кленова (экзо-), ДНК-полимераза Therminator). Согласно одному конкретному варианту осуществления терминаторные нуклеотиды содержат модификации группы г 3' углерода дезоксирибозы. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны 3' блокированного обратимого терминатора, содержащего нуклеотиды, 3' неблокированного обратимого терминатора, содержащего нуклеотиды, терминаторов, содержащих 2' модификации дезоксинуклеотидов, терминаторов, содержащих модификации азотистого основания дезоксинуклеотидов, и их комбинаций. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из дидезоксинуклеотидов, инвертированных дидезоксинуклеотидов, 3' биотинилированных нуклеотидов, 3' аминонуклеотидов, 3'-фосфорилированных нуклеотидов, 3'-О-метилнуклеотидов, 3' углеродных спейсерных нуклеотидов, включая 3' С3 спейсерные нуклеотиды, 3' С18 нуклеотиды, 3' гександиоловые спейсерные нуклеотиды, ациклонуклеотиды и их комбинации.

[00140] В настоящем документе описаны способы амплификации генома, причем способ предусматривает a) приведение в контакт образца, содержащего геном, множества праймеров амплификации (например, два или более праймеров), полимеразы нуклеиновых кислот и смеси нуклеотидов, которая содержит один или несколько терминаторных нуклеотидов, которые останавливают репликацию нуклеиновой кислоты полимеразой, и b) инкубацию образца при условиях, которые стимулируют репликацию генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи. Согласно одному варианту осуществления любого из описанных выше способов способ дополнительно предусматривает выделение из множества терминированных продуктов амплификации продуктов, которые составляют от приблизительно 50 до приблизительно 2000 нуклеотидов в длину. Согласно одному варианту осуществления любого из описанных выше способов способ дополнительно предусматривает выделение из множества терминированных продуктов амплификации продуктов, которые составляют от приблизительно 400 до приблизительно 600 нуклеотидов. Согласно одному варианту осуществления любого из описанных выше способов способ дополнительно предусматривает: c) репарацию концов и полиаденилирование и d) лигирование молекул, полученных на стадии (c), с адаптерами, получая таким образом библиотеку продуктов амплификации. Согласно одному варианту осуществления любого из описанных выше способов способ дополнительно предусматривает секвенирование продуктов амплификации. Согласно одному варианту осуществления любого из описанных выше способов амплификацию проводят при по существу изотермических условиях. Согласно

одному варианту осуществления любого из приведенных выше способов полимераза нуклеиновых кислот представляет собой ДНК-полимеразу.

[00141] Согласно одному варианту осуществления любого из описанных выше способов ДНК-полимераза представляет собой замещающую цепи ДНК-полимеразу. Согласно одному варианту осуществления любого из приведенных выше способов полимераза нуклеиновых кислот выбрана из полимеразы бактериофага phi29 (Ф29), генетически модифицированной ДНК-полимеразы phi29 (Ф29), фрагмента Кленова ДНК-полимеразы I, ДНК-полимеразы фага M2, ДНК-полимеразы фага phiPRDl, ДНК-полимеразы Bst, ДНК-полимеразы большого фрагмента Bst, полимеразы экзо(-) Bst, ДНК-полимеразы экзо(-) Bca, ДНК-полимеразы Bsu, ДНК-полимеразы VentR, ДНК-полимеразы VentR (экзо-), ДНК-полимеразы Deep Vent, ДНК-полимеразы Deep Vent (экзо-), ДНК-полимеразы IsoPol, ДНК-полимеразы I, ДНК-полимеразы Therminator, ДНК-полимеразы T5, секвеназы, ДНК-полимеразы T7, T7-секвеназы и ДНК-полимеразы T4. Согласно одному варианту осуществления любого из приведенных выше способов полимераза нуклеиновых кислот характеризуется 3'->5' экзонуклеазной активностью, и терминаторные нуклеотиды ингибируют такую 3'->5' экзонуклеазную активность. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из нуклеотидов с модификацией альфа-группы (например, альфа-тио-дидезоксинуклеотиды, создающие фосфоротиоатную связь), C3 спейсерных нуклеотидов, запертых нуклеиновых кислот (LNA), инвертированных нуклеиновых кислоты, 2' фторнуклеотидов, 3' фосфорилированных нуклеотидов, 2'-O-метил-модифицированных нуклеотидов и транснуклеиновых кислот. Согласно одному варианту осуществления любого из приведенных выше способов полимераза нуклеиновых кислот не характеризуется 3'->5' экзонуклеазной активностью. Согласно одному конкретному варианту осуществления полимераза выбрана ДНК-полимеразы Bst, полимеразы экзо(-) Bst, ДНК-полимеразы экзо(-) Bca, ДНК-полимеразы Bsu, ДНК-полимеразы VentR (экзо-), ДНК- полимеразы Deep Vent (экзо-), ДНК-полимеразы фрагмента Кленова (экзо-) и ДНК-полимеразы Therminator. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды содержат модификации группы г 3' углерода дезоксирибозы. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из 3' блокированного обратимого терминатора, содержащего нуклеотиды, 3' неблокированного обратимого терминатора, содержащего нуклеотиды, терминаторов, содержащих 2' модификации дезоксинуклеотидов, терминаторов, содержащих модификации азотистого основания дезоксинуклеотидов, и их комбинаций. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из дидезоксинуклеотидов,

инвертированных дидезоксинуклеотидов, 3’ биотинилированных нуклеотидов, 3’ аминонуклеотидов, 3’-фосфорилированных нуклеотидов, 3’-О-метилнуклеотидов, 3’ углеродных спейсерных нуклеотидов, включая 3’ С3 спейсерные нуклеотиды, 3’ С18 нуклеотиды, 3’ гександиоловые спейсерные нуклеотиды, ациклонуклеотиды и их комбинации. Согласно одному варианту осуществления любого из приведенных выше способов праймеры для амплификации составляют от 4 до 70 нуклеотидов в длину. Согласно одному варианту осуществления любого из приведенных выше способов продукты амплификации составляют от приблизительно 50 до приблизительно 2000 нуклеотидов в длину. Согласно одному варианту осуществления любого из приведенных выше способов целевая нуклеиновая кислота представляет собой ДНК (например, кДНК или геномная ДНК). Согласно одному варианту осуществления любого из приведенных выше способов праймеры для амплификации представляют собой случайные праймеры. Согласно одному варианту осуществления любого из приведенных выше способов праймеры для амплификации содержат баркод. Согласно одному конкретному варианту осуществления баркод содержит баркод клетки. Согласно одному конкретному варианту осуществления баркод содержит баркод образца. Согласно одному варианту осуществления любого из приведенных выше способов праймеры для амплификации содержат уникальный молекулярный идентификатор (UMI). Согласно одному варианту осуществления любого из приведенных выше способов способ предусматривает денатурирование целевой нуклеиновой кислоты или геномной ДНК перед начальным отжигом праймера. Согласно одному конкретному варианту осуществления денатурацию проводят в щелочных условиях с последующей нейтрализацией. Согласно одному варианту осуществления любого из приведенных выше способов образец, праймеры для амплификации, полимераза нуклеиновых кислот и смесь нуклеотидов содержатся в микрожидкостном устройстве. Согласно одному варианту осуществления любого из приведенных выше способов образец, праймеры для амплификации, полимераза нуклеиновых кислот и смесь нуклеотидов содержатся в капле. Согласно одному варианту осуществления любого из приведенных выше способов образец выбран образцов ткани (тканей), клеток, образцов биологической жидкости (например, кровь, моча, слюна, лимфатическая жидкость, спинномозговая жидкость (CSF), амниотическая жидкость, плевральная жидкость, перикардиальная жидкость, асцит, водянистая влага), образцов костного мозга, образцов спермы, биопсийных образцов, образцов рака, образцов опухоли, образцов клеточного лизата, криминалистических образцов, археологических образцов, палеонтологических образцов, образцов инфекции, производственных образцов, целых растений, частей растения, образцов микробиоты, вирусных препаратов, образцов почвы,

образцов морской воды, образцов пресной воды, бытовых или промышленных образцов и их комбинаций и изолятов. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой клетку (например, клетка животного [например, клетка человека], растительная клетка, грибковая клетка, бактериальная клетка и клетка простейшего). Согласно одному конкретному варианту осуществления клетку лизируют перед репликацией. Согласно одному конкретному варианту осуществления клеточный лизис сопровождается протеолизом. Согласно одному конкретному варианту осуществления клетка выбрана из клетки предимплантационного эмбриона, стволовой клетки, фетальной клетки, опухолевой клетки, предполагаемой раковой клетки, раковой клетки, клетки, подвергнутой процедуре редактирования генов, клетки из патогенного организма, клетки, полученной из криминалистического образца, клетки, полученной из археологического образца, и клетки, полученной из палеонтологического образца. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой клетку из предимплантационного эмбриона (например, бластомер [например, бластомер, полученный из эмбриона восьмиклеточной стадии, полученного с помощью in vitro оплодотворения]). Согласно одному конкретному варианту осуществления способ дополнительно предусматривает определение наличия предрасполагающих к заболеванию вариантов зародышевой линии или соматических вариантов в эмбриональной клетке. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой клетку из патогенного организма (например, бактерии, гриба, простейшего). Согласно одному конкретному варианту осуществления клетку патогенного организма получают из жидкости, взятой у пациента, образца микробиоты (например, образец микробиоты ЖКТ, образец вагинальной микробиоты, образец микробиоты кожи и т.д.) или постоянного изделия медицинского назначения (например, внутривенный катетер, уретральный катетер, цереброспинальный шунт, искусственный клапан, искусственный сустав, эндотрахеальная трубка и т.д.). Согласно одному конкретному варианту осуществления способ дополнительно предусматривает стадию определения идентичности патогенного организма. Согласно одному конкретному варианту осуществления способ дополнительно предусматривает определение наличия генетических вариантов, ответственных за устойчивость патогенного организма к лечению. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой опухолевую клетку, предполагаемую раковую клетку или раковую клетку. Согласно одному конкретному варианту осуществления способ дополнительно предусматривает определение наличия одной или нескольких диагностических или прогностических мутаций. Согласно одному

конкретному варианту осуществления способ дополнительно предусматривает определение наличия вариантов зародышевой линии или соматических вариантов, ответственных за устойчивость к лечению. Согласно одному варианту осуществления любого из приведенных выше способов образец представляет собой клетку, подвергнутую процедуре редактирования генов. Согласно одному конкретному варианту осуществления способ дополнительно включает определение наличия незапланированных мутаций, вызванных процессом редактирования генов. Согласно одному варианту осуществления любого из приведенных выше способов способ дополнительно предусматривает определение истории клеточной родословной. Согласно родственному аспекту настоящее изобретение относится к применению любого из приведенных выше способов для идентификации низкочастотных вариантов последовательности (например, вариантов, которые составляют >0,01% от общего числа последовательностей).

[00142] Согласно родственному аспекту настоящее изобретение относится к набору, содержащему обратную транскриптазу, полимеразу нуклеиновой кислоты, один или несколько праймеров амплификации, смесь нуклеотидов, содержащую один или несколько терминаторных нуклеотидов и необязательно инструкцию по применению. Согласно одному варианту осуществления набора согласно настоящему изобретению полимераза нуклеиновой кислоты представляет собой замещающую цепи ДНК-полимеразу. Согласно некоторым вариантам осуществления обратная транскриптаза осуществляет переключение матрицы. Согласно некоторым вариантам осуществления обратная транскриптаза является вариантом MMLV (вирус мышиного лейкоза Молони), HIV-1, AMV (вирус миелобластоза птиц), теломеразы RT, FIV (вирус иммунодефицита кошачьих) или XMRV (ксенотропный вирус, связанный с вирусом лейкемии мышей). Неограничивающие примеры обратных транскриптаз включают SuperScript I (Thermo), SuperScript II (Thermo), SuperScript III (Thermo), SuperScript IV (Thermo), OmniScript (Qiagen), SensiScript (Qiagen), PrimeScript (Takara), Maxima H- (Thermo), AcuuScript Hi-Fi (Agilent), iScript (Bio-Rad), eAMV (Merck KGaA), qScript (Quanta Biosciences), SmartScribe (Clontech) или GoScript (Promega). Согласно одному варианту осуществления наборов согласно настоящему изобретению полимераза нуклеиновых кислот выбрана из полимеразы бактериофага phi29 (Ф29), генетически модифицированной ДНК-полимеразы phi29 (Ф29), фрагмента Кленова ДНК-полимеразы I, ДНК-полимеразы фага M2, ДНК-полимеразы фага phiPRDl, ДНК- полимеразы Bst, ДНК-полимеразы большого фрагмента Bst, полимеразы экзо(-) Bst, ДНК- полимеразы экзо(-) Bca, ДНК-полимеразы Bsu, ДНК-полимеразы VentR, ДНК-полимеразы VentR (экзо-), ДНК-полимеразы Deep Vent, ДНК-полимеразы Deep Vent (экзо-), ДНК- полимеразы IsoPol, ДНК-полимеразы I, ДНК-

полимеразы Therminator, ДНК-полимеразы T5, секвеназы, ДНК-полимеразы T7, T7-секвеназы и ДНК-полимеразы T4. Согласно одному варианту осуществления наборов согласно настоящему изобретению полимераза нуклеиновых кислот характеризуется 3'->5' экзонуклеазной активностью, и терминаторные нуклеотиды ингибируют такую 3'->5' экзонуклеазную активность (например, нуклеотиды с модификацией альфа-группы [например, альфа-тио-дидезоксинуклеотиды], C3 спейсерные нуклеотиды, запертые нуклеиновые кислоты (LNA), инвертированные нуклеиновые кислоты, 2' фторнуклеотиды, 3' фосфорилированные нуклеотиды, 2'-O-метил-модифицированные нуклеотиды, транснуклеиновые кислоты). Согласно одному варианту осуществления наборов согласно настоящему изобретению полимераза нуклеиновых кислот не характеризуется 3'->5' экзонуклеазной активностью (например, ДНК-полимераза Bst, полимераза экзо(-) Bst, ДНК-полимераза экзо(-) Bca, ДНК-полимераза Bsu, ДНК- полимераза VentR (экзо-), ДНК-полимераза Deep Vent (экзо-), ДНК-полимераза фрагмента Кленова (экзо-), ДНК-полимераза Therminator). Согласно одному конкретному варианту осуществления терминаторные нуклеотиды содержат модификации группы г 3' углерода дезоксирибозы. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из 3' блокированного обратимого терминатора, содержащего нуклеотиды, 3' неблокированного обратимого терминатора, содержащего нуклеотиды, терминаторов, содержащих 2' модификации дезоксинуклеотидов, терминаторов, содержащих модификации азотистого основания дезоксинуклеотидов, и их комбинаций. Согласно одному конкретному варианту осуществления терминаторные нуклеотиды выбраны из дидезоксинуклеотидов, инвертированных дидезоксинуклеотидов, 3' биотинилированных нуклеотидов, 3' аминонуклеотидов, 3'-фосфорилированных нуклеотидов, 3'-O-метилнуклеотидов, 3' углеродных спейсерных нуклеотидов, включая 3' C3 спейсерные нуклеотиды, 3' C18 нуклеотиды, 3' гександиоловые спейсерные нуклеотиды, ациклонуклеотиды и их комбинации. Согласно некоторым вариантам осуществления набор содержит по меньшей мере один стабилизатор фермента, нейтрализующий буфер, денатурирующий буфер или их комбинацию. В некоторых случаях набор содержит один или несколько модулей. В некоторых случаях набор содержит модуль генома и модуль транскриптома.

Пронумерованные варианты осуществления

[00143] В настоящем документе описаны следующие пронумерованные варианты осуществления 1-46. 1. В настоящем документе описаны варианты осуществления, относящиеся к способу мультиомического анализа отдельных клеток,

предусматривающему: a. выделение отдельной клетки из популяции клеток, b. секвенирование библиотеки кДНК, содержащей полинуклеотиды, амплифицированные из транскриптов мРНК из клетки, и c. секвенирование генома клетки, где секвенирование генома клетки предусматривает: i. обеспечение генома отдельной клетки, ii. контакт генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по меньшей мере один терминаторный нуклеотид, который останавливает репликацию нуклеиновых кислот полимеразой, и iii. амплификацию по меньшей мере части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи, iv. Лигирование молекул, полученных на стадии (iii), с адапторами, с получением таким образом библиотеки геномной ДНК, и v. секвенирование библиотеки геномной ДНК. 2. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 1, где способ дополнительно предусматривает идентификацию по меньшей мере одного белка на поверхности клетки. 3. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 1, где транскрипты мРНК содержат полиаденилировнные транскрипты мРНК. 4. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 1, где транскрипты мРНК не содержат полиаденилировнные транскрипты мРНК. 5. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-4, где секвенирование библиотеки кДНК содержит амплификация транскриптов мРНК с переключающими матрицу праймерами. 6. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-4, где по меньшей мере некоторые из полинуклеотидов библиотеки кДНК содержат баркод. 7. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-4, где по меньшей мере некоторые из полинуклеотидов библиотеки кДНК содержат по меньшей мере два баркода. 8. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 6 или 7, где баркод содержит баркод клетки. 9. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 6 или 7, где баркод содержит баркод образца. 10. Способ мультиомического анализа отдельных клеток, предусматривающий: a. выделение отдельной клетки из популяции клеток, b. идентификацию по меньшей мере одного белка на поверхности клетки, и c. секвенирование генома клетки, где секвенирование генома клетки предусматривает: i. обеспечение генома отдельной клетки, ii. контакт генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по меньшей мере один терминаторный нуклеотид, который

останавливает репликацию нуклеиновых кислот полимеразой, iii. амплификацию по меньшей мере части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи, iv. лигирование молекул, полученных на стадии (iii), с адапторами, с получением таким образом библиотеки геномной ДНК, и v. секвенирование библиотеки геномной ДНК. 11. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 10, где идентификация по меньшей мере одного белка на поверхности клетки предусматривает контакт клетки с меченным антителом, которое связывается с по меньшей мере одним белком. 12. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 11, где меченное антитело содержит по меньшей мере одну флуоресцентную метку. 13. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 11, где меченное антитело содержит по меньшей мере одну массовую метку. 14. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 11, где меченное антитело содержит по меньшей мере один баркод нуклеиновой кислоты. 15. Способ мультиомического анализа отдельных клеток, предусматривающий: a. выделение отдельной клетки из популяции клеток, b. секвенирование генома клетки, где секвенирование генома клетки предусматривает: i. обеспечение генома отдельной клетки, ii. расщепление генома чувствительным к метилированию ферментом рестрикции с получением фрагментов генома, iii. контакт по меньшей мере некоторых фрагментов генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по меньшей мере один терминаторный нуклеотид, который останавливает репликацию нуклеиновых кислот полимеразой, iv. амплификацию по меньшей мере части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи, v. амплификацию по меньшей мере некоторых фрагментов генома посредством специфической к метилированию ПЦР, vi. лигирование молекул, полученных на стадиях (iv и v) с адапторами, с получением таким образом библиотеки геномной ДНК и библиотеки метиломы ДНК, и vii. секвенирование библиотеки геномной ДНК и библиотеки метиломы. 16. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 15, где идентификация по меньшей мере одного белка на поверхности клетки предусматривает контакт клетки меченным антителом, которое связывается с по меньшей мере одним белком. 17. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 16, где меченное антитело содержит по меньшей мере одну флуоресцентную метку. Кроме того, настоящее изобретение относится к способу

согласно варианту осуществления 16, где меченное антитело содержит по меньшей мере одну массовую метку. 19. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 16, где меченное антитело содержит по меньшей мере один баркод нуклеиновой кислоты. 20. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-19, где отдельная клетка представляет собой клетку млекопитающего. 21. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-19, где отдельная клетка представляет собой клетку человека. 22. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-19, где отдельные клетки происходят из печени, кожи, почки, крови или легкого. 23. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-19, где отдельная клетка представляет собой первичную клетку. 24. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-23, где способ дополнительно предусматривает удаление по меньшей мере одного терминаторного нуклеотида из терминированных продуктов амплификации. 25. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-23, где по меньшей мере некоторые из продуктов амплификации содержит баркод. 26. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-23, где по меньшей мере некоторые из продуктов амплификации содержат по меньшей мере два баркода. 27. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 24 или 26, где баркод содержит баркод клетки. 28. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 24 или 26, где баркод содержит баркод образца. 29. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-28, где по меньшей мере некоторые из праймеров амплификации содержат уникальный молекулярный идентификатор (UMI). 30. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-28, где по меньшей мере некоторые из праймеров амплификации содержат по меньшей мере два уникальных молекулярных идентификатора (UMI). 31. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-30, где способ дополнительно предусматривает дополнительную стадию амплификации с применением ПЦР. 32. Кроме того, настоящее изобретение относится к способу согласно любому из вариантов осуществления 1-30, где по меньшей мере одна мутация идентифицирована в геноме клетки, где мутация отлична от соответствующего положения в эталонной последовательности. 33. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна

мутация происходит в менее 50% популяции клеток. 34. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в менее 25% популяции клеток. 35. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в менее 1% популяции клеток. 36. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 0,1% популяции клеток. 37. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 0,01% популяции клеток. 38. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 0,001% популяции клеток. 39. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 0,0001% популяции клеток. 40. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 50% последовательностей продукта амплификации. 41. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 25% последовательностей продукта амплификации. 42. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 1% последовательностей продукта амплификации. 43. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 0,1% последовательностей продукта амплификации. 44. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 0,01% последовательностей продукта амплификации. 45. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 0,001% последовательностей продукта амплификации. 46. Кроме того, настоящее изобретение относится к способу согласно варианту осуществления 32, где по меньшей мере одна мутация происходит в не более 0,0001% последовательностей продукта амплификации.

Примеры

[00144] Следующие ниже примеры представлены для более ясной иллюстрации принципа и осуществления на практике вариантов осуществления, раскрытых в настоящем документе, специалистам в настоящей области техники, и их не следует

истолковывать как ограничение объема любых заявленных вариантов осуществления. Если не указано иное, все части и проценты являются массовыми.

[00145] Пример 1. Амплификация, управляемая первичной матрицей (PTA)

[00146] Хотя PTA можно использовать для амплификации любой нуклеиновой кислоты, она является особенно применимой для амплификации целого генома, поскольку позволяет охватывать больший процент клеточного генома более однородным и воспроизводимым образом и с более низкими показателями ошибок, чем используемые в настоящее время способы, такие как, например, амплификация с множественным замещением цепи (MDA), позволяя избежать таких недостатков используемых в настоящее время способов, как экспоненциальная амплификация в тех местах, где полимераза сначала удлиняет случайные праймеры, что приводит к случайному избыточному представлению локусов и аллелей и распространению мутации (см. фиг. 1G). PTA также применяют для других методик анализа, таких как анализ транскриптомы.

[00147] *Клеточная культура*

[00148] Клетки NA12878 человека (Coriell Institute) поддерживали в среде RPMI с добавлением 15% FBS и 2 мМ L-глутамина и 100 единиц/мл пенициллина, 100 мкг/мл стрептомицина и 0,25 мкг/мл амфотерицина B (Gibco, Life Technologies). Клетки высевали с плотностью 3,5 x 10^5 клеток/мл. Культуры разделяли каждые 3 дня и поддерживали во влажной камере при 37°C с 5% CO_2.

[00149] *Выделение отдельных клеток и WTA*

[00150] Общий протокол для WTA (полный анализ транскриптома) показан на фиг. 2F. Клетки ресуспендировали при концентрации 150–500 клеток/мкл. Эту клеточную суспензию окрашивали 20 мкл свежеприготовленного буфера для окрашивания (2,5 мкл гомодимера этидия-1 и 0,625 мкл кальцеина AM из набора LIVE/DEAD® Viability/Cytotoxicity Kit от Life Technology), добавляли к 1,25 мл клеточного буфера, содержащего 1X PBS и 0,05% tween-20. Затем клетки сортировали с помощью сортировочного устройства FACS Aria III с внесением клеток в каждую из 96 лунок. Реакционную смесь, содержащую 5x RT буфер, PEG4000, RT праймер (100 мкМ), TS Oligo (20 мкМ), обратную транскриптазу, ингибитор РНКазы, желатин, Tween-20, Triton -X, смесь dNTP, TMAC (1 М), Бетаин (5 М), $MgCl_2$ (50 мМ), ERCC, добавляли в каждую лунку ERCC, затем образец помещали в термоциклер на 90 минут при 42°C, 30 мин при 50°C, а затем

выдерживали при 4°C до тех пор, пока образец не будет обработан для предварительной амплификации. После термоциклирования для RT образец либо обрабатывали для амплификации ДНК, либо для предварительной амплификации первой цепи кДНК, полученной в результате реакции RT. Предварительную амплификацию образца проводили с использованием одного праймера (полусупрессивная ПЦР) со следующим протоколом для амплификации продуктов кДНК. Кратко, 5 мкл реакционной смеси RT добавляли к реакционной смеси объемом 30 мкл, содержащей 2X мастер-микс, 1 микромолярный праймер и 5X буфер Preamp, используя следующие условия термоциклирования: 95°C – 1 мин. 21 цикл 95°C – 15 сек, 60°C – 30 сек, 68°C – 4 мин с последующей выдержкой при 72°C в течение десяти минут. Затем образцы преобразовывали в библиотеку для секвенирования с помощью набора для подготовки библиотеки Nextera XT в соответствии с инструкциями производителя (фиг. 2G). Результаты эксперимента RT показаны для шести образцов в Таблице 1.

[00151] **Таблица 1**

Образец №	**Картирование на геноме**	**Картирование на транскриптоме**	**Обнаружение гена**	**Прочтения/ образец**	**Клетки**
1	99,46%	94,24%	9298	2,530,484	~350
2	99,48%	94,17%	9692	2,239,730	~350
3	99,45%	94,77%	3630	2,466,899	~35
4	98,91%	94,45%	4351	2,235,999	~35
5	98,72%	92,70%	2189	1,788,750	~3
6	96,13%	89,92%	4272	1,565,289	~3

[00152] *Выделение отдельных клеток и WGA*

[00153] После культивирования клеток NA12878 в течение минимум трех дней после посева при плотности $3,5 \times 10^5$ клеток/мл, 3 мл клеточной суспензии осаждали при 300×g в течение 10 минут. Затем среду отбрасывали и клетки промывали три раза 1 мл буфера для промывки клеток (IX PBS, содержащий 2% FBS без Mg^2 или Ca^2), центрифугированных при 300×g, 200×g и, наконец, 100×g в течение 5 минут. Затем клетки ресуспендировали в 500 мкл буфера для промывания клеток. За этим следовало окрашивание с помощью 100 нМ кальцеина AM (Molecular Probes) и 100 нг/мл йодида пропидия (PI, Sigma-Aldrich) для различения популяции живых клеток. Клетки загружали в проточный цитометр BD FACScan (FACSAria II) (BD Biosciences), который тщательно очищали с помощью ELIMINase (Decon Labs) и откалибровывали с использованием

флуоресцентных шариков Accudrop (BD Biosciences) для сортировки клеток. Одиночную клетку из кальцеин AM-положительной, PI-отрицательной фракции отсортировали в каждую лунку 96-луночного планшета, содержащего 3 мкл PBS с 0,2% Tween 20 в клетках, которые должны подвергаться PTA (Sigma-Aldrich). Несколько лунок намеренно оставляли пустыми, чтобы использовать их в качестве контроля без матрицы (NTC). Сразу после сортировки планшеты кратковременно центрифугировали и помещали на лед. Затем клетки замораживали как минимум в течение ночи при -20°C. На следующий день WGA-реакции собирали на установке для предварительной ПЦР, которая обеспечивает постоянное положительное давление воздуха, отфильтрованного HEPA, и которую обеззараживали УФ-светом в течение 30 минут перед каждым экспериментом.

[00154] MDA проводили с модификациями, которые, как было показано ранее, улучшают однородность амплификации. В частности, устойчивые к экзонуклеазам случайные праймеры (ThermoFisher) добавляли в лизисный буфер/смесь до конечной концентрации 125 мкМ. 4 мкл полученной смеси для лизиса/денатурирования добавляли в пробирки, содержащие отдельные клетки, перемешивали вихревым способом, кратковременно центрифугировали и инкубировали на льду в течение 10 минут. Лизаты клеток нейтрализовали добавлением 3 мкл стоп-раствора, перемешивали встряхиванием, кратковременно центрифугировали и помещали при комнатной температуре. За этим следовало добавление 40 мкл смеси для амплификации перед инкубацией при 30°C в течение 8 часов, после чего амплификацию прекращали нагреванием до 65°C в течение 3 минут.

[00155] PTA проводили путем сначала дополнительного лизирования клеток после замораживания-оттаивания путем добавления 2 мкл предварительно охлажденного раствора смеси 1:1 5% Triton X-100 (Sigma-Aldrich) и 20 мг/мл протеиназы K (Promega). Затем клетки перемешивали вихревым способом и кратковременно центрифугировали перед тем, как поместить при 40° в течение 10 минут. Затем к лизированным клеткам добавляли 4 мкл лизисного буфера/смеси и 1 мкл 500 мкМ случайного праймера, устойчивого к экзонуклеазам, для денатурирования ДНК перед перемешивания вихревым способом, центрифугирования и помещения при 65° в течение 15 минут. Затем добавляли 4 мкл стоп-раствора при комнатной температуре и образцы перемешивали вихревым способом и центрифугировали. 56 мкл смеси для амплификации (праймеры, dNTP, полимераза, буфер), которая содержала альфа-тио-ddNTP в равных соотношениях при концентрации 1200 мкМ в конечной реакции амплификации. Затем образцы помещали при 30°C на 8 часов, после чего амплификацию прекращали нагреванием до 65°C в течение 3 минут.

[00156] После стадии амплификации ДНК из обеих реакций, MDA и PTA, очищали с использованием магнитных шариков AMPure XP (Beckman Coulter) при соотношении шариков к образцу 2:1, и выход измеряли с использованием набора Qubit dsDNA HS Assay Kit с помощью флуориметра Qubit 3.0 согласно инструкции производителя (Life Technologies).

[00157] *Получение библиотеки*

[00158] Реакции с MDA привели к продукции 40 мкг амплифицированной ДНК. 1 мкг продукта фрагментировали в течение 30 минут согласно стандартным процедурам. Затем образцы прошли подготовку стандартной библиотеки с помощью 15 мкМ адапторов с двойным индексом (репарация конца посредством T4 полимеразы, T4 полинуклеотидной киназы и Taq полимеразы для полиаденилирования) и 4 цикла ПЦР. Каждая реакция PTA генерировала от 40 до 60 нг материала, который использовали для получения стандартной библиотеки секвенирования ДНК целиком без фрагментации. При лигировании с T4 лигазой использовали 2,5 мкМ адаптеры с UMI и двойными индексами, а в заключительной амплификации использовали 15 циклов ПЦР (полимераза с горячим стартом). Затем библиотеки очищали с использованием двустороннего SPRI с коэффициентами 0,65X и 0,55X для правосторонней и левосторонней селекции соответственно. Конечные библиотеки количественно определяли с использованием набора Qubit dsDNA BR Assay Kit и 2100 Bioanalyzer (Agilent Technologies) перед секвенированием на платформе Illumina NextSeq. Все платформы секвенирования Illumina, включая NovaSeq, также совместимы с этим протоколом.

[00159] *Анализ данных*

[00160] Прочтения секвенирования демультиплексировали на основе баркода клетки с использованием Bcl2fastq. Затем прочтения укорачивали с помощью trimmomatic, после чего проводили выравнивание по hg19 с помощью BWA. Прочтения подвергли параллельной маркировке с помощью Picard с последующим локальным выравниванием и повторной калибровкой базы с использованием GATK 4.0. Все файлы, используемые для расчета показателей качества, подвергали субдискретизации до двадцати миллионов прочтений с помощью Picard DownSampleSam. Количественные показатели качества получали из окончательного файла BAM с использованием qualimap, а также Picard AlignmentSummaryMetrics и CollectWgsMetrics. Общий охват генома также оценивали с помощью Preseq.

[00161] *Обнаружение вариантов*

[00162] Однонуклеотидные варианты и индели обнаруживали с использованием GATK UnifiedGenotyper от GATK 4.0. Стандартные критерии фильтрации с использованием практических рекомендаций GATK использовали для всех стадий процесса (https://software.broadinstitute.org/gatk/best-practices/). Варианты числа копий обнаруживали с использованием Control-FREEC (Boeva et al., Bioinformatics, 2012, 28(3):423-5). Структурные варианты также обнаруживали с помощью CREST (Wang et al., Nat Methods, 2011, 8(8):652-4).

[00163] Результаты

Как показано на фиг. 3A и 3B, показатели картирования и показатели качества картирования амплификации только с дидезоксинуклеотидами («обратимые») составляют 15,0 +/- 2,2 и 0,8 +/- 0,08, соответственно, в то время как включение устойчивых к экзонуклеазам альфа-тио-дидезоксинуклеотидных терминаторов («необратимые») приводит к показателям картирования и показателям качества картирования 97,9 +/- 0,62 и 46,3 +/- 3,18, соответственно. Эксперименты также проводили с использованием обратимого ddNTP и различных концентраций терминаторов. (Фигура 2A, внизу).

[00164] На фиг. 2B-2E показаны сравнительные данные, полученные на отдельных клетках человека NA12878, которые подверглись MDA (в соответствии со способом Dong, X. et al., Nat Methods. 2017, 14(5):491-493) или PTA. В то время как оба протокола давали сравнимо низкие показатели дупликации ПЦР (MDA 1,26% +/- 0,52 по сравнению с PTA 1,84% +/- 0,99) и GC% (MDA 42,0 +/- 1,47 по сравнению с PTA 40,33 +/- 0,45), PTA продуцировала ампликоны меньшего размера. Процент прочтений, которые картировали и у которых показатели качества картирования были также значительно выше для PTA по сравнению с MDA (PTA 97,9 +/- 0,62 по сравнению с MDA 82,13 +/- 0,62 и PTA 46,3 +/- 3,18 по сравнению с MDA 43,2 +/- 4,21, соответственно). В целом, PTA производит более удобные, картируемые данные по сравнению с MDA. На фиг. 4A показано, что по сравнению с MDA, PTA значительно улучшила однородность амплификации с большей широтой охвата и меньшим количеством областей, где охват падает почти до 0. Использование PTA позволяет идентифицировать низкочастотные варианты последовательностей в популяции нуклеиновых кислот, включая варианты, которые составляют >0,01% от общего числа последовательностей. PTA можно успешно использовать для амплификации генома отдельной клетки.

[00165] **Пример 2. Сравнительный анализ PTA**

[00166] *Сопоставительный анализ PTA и поддержания клеток и выделения SCMDA*

[00167] Лимфобластоидные клетки из проекта 1000 Genome Project NA12878 (Coriell Institute, Camden, NJ, USA) поддерживали в среде RPMI с добавлением 15% FBS, 2 мМ L-глутамина, 100 единиц/мл пенициллина, 100 мкг/мл стрептомицина и 0,25 мкг/мл амфотерицина B). Клетки высевали с плотностью $3,5\times10^5$ клеток/мл и разделяли каждые 3 дня. Их поддерживали во влажной камере при 37°C с 5% CO_2. Перед выделением отдельных клеток 3 мл суспензии клеток, которые размножились за предыдущие 3 дня, центрифугировали при 300×g в течение 10 минут. Осажденные клетки промывали три раза 1 мл буфера для промывки клеток (IX PBS, содержащий 2% FBS без Mg^{2+} или Ca^{2+})), где их последовательно центрифугировали при 300×g, 200×g и, наконец, 100×g в течение 5 минут для удаления мертвых клеток. Затем клетки ресуспендировали в 500 мкл буфера для промывания клеток с последующим окрашиванием 100 нМ кальцеина AM и 100 нг/мл йодида пропидия (PI) для различия популяции живых клеток. Клетки загружали в проточный цитометр BD FACScan (FACSAria II), который тщательно очищали с помощью ELIMINase и откалибровывали с использованием флуоресцентных шариков Accudrop. Отдельную клетку из кальцеин-AM-положительной, PI-отрицательной фракции сортировали в каждой лунке 96-луночного планшета, содержащего 3 мкл PBS с 0,2% Tween -20. Несколько лунок намеренно оставляли пустыми, чтобы использовать их в качестве контроля без матрицы. Сразу после сортировки планшеты кратковременно центрифугировали и помещали на лед. Затем клетки замораживали как минимум в течение ночи при -80°C.

[00168] *Эксперименты PTA и SCMDA*

[00169] Реакции WGA собирали на установке для предварительной ПЦР, которая обеспечивает постоянное положительное давление с помощью воздуха, отфильтрованного HEPA, и которую обеззараживают УФ-светом в течение 30 минут перед каждым экспериментом. MDA проводили в соответствии с методологией SCMDA в соответствии с опубликованным протоколом (Dong et al. Nat. Meth. 2017, 14, 491-493). В частности, устойчивые к экзонуклеазам случайные праймеры добавляли в конечной концентрации 12,5 мкМ в буфер для лизиса. 4 мкл полученной смеси для лизиса добавляли в пробирки, содержащие одиночные клетки, пипетировали три раза для перемешивания, быстро центрифугировали и инкубировали на льду в течение 10 минут. Лизаты клеток нейтрализовали добавлением 3 мкл буфера для гашения, перемешивали пипетированием 3

раза, кратковременно центрифугировали и помещали на лед. После этого добавляли 40 мкл смеси для амплификации перед инкубацией при 30°C в течение 8 часов, после чего амплификацию прекращали нагреванием до 65°C в течение 3 минут. PTA проводили путем сначала дополнительного лизиса клеток после замораживания-оттаивания путем добавления 2 мкл предварительно охлажденного раствора смеси 1:1 5% Triton X-100 и 20 мг/мл протеиназы K. Затем клетки перемешивали вихревым способом и кратковременно центрифугировали перед тем, как поместить при 40 градусах в течение 10 минут. Затем к лизированным клеткам добавляли 4 мкл денатурирующего буфера и 1 мкл 500 мкМ устойчивого к экзонуклеазам случайного праймера для денатурирования ДНК перед перемешивания вихревым способом, центрифугированием и помещением при 65°C в течение 15 минут. Затем добавляли 4 мкл гасящего раствора при комнатной температуре, и образцы перемешивали вихревым способом и центрифугировали. 56 мкл смеси для амплификации, которая содержала альфа-тио-ddNTP в равных соотношениях при концентрации 1200 мкМ в конечной реакции амплификации. Затем образцы помещали при 30°C на 8 часов, после чего амплификацию прекращали нагреванием до 65°C в течение 3 минут. После амплификации SCMDA или PTA ДНК очищали с использованием магнитных шариков AMPure XP при соотношении шариков к образцу 2:1, и выход измеряли с помощью набора Qubit dsDNA HS Assay Kit с помощью флуорометра Qubit 3.0 в соответствии с инструкциями производителя.

[00170] *Получение библиотеки*

[00171] 1 мкг продукта SCMDA фрагментировали в течение 30 минут в соответствии с протоколом HyperPlus после добавления кондиционирующего раствора. Затем образцы подвергали подготовке стандартной библиотеки с использованием 15 мкМ уникальных адапторов с двойным индексом и 4 цикла ПЦР. Весь продукт каждой реакции PTA использовали для получения библиотеки секвенирования ДНК с использованием стандартных протоколов амплификации без фрагментации. При лигировании использовали 2,5 мкМ уникального адаптера с двойным индексом, а в окончательной амплификации использовали 15 циклов ПЦР. Затем библиотеки из SCMDA и PTA визуализировали на 1% Agarose E-Gel. Фрагменты размером 400-700 п.н. вырезали из геля и выделяли с помощью набора Gel DNA Recovery. Конечные библиотеки количественно определяли с использованием набора Qubit dsDNA BR Assay Kit и Agilent 2100 Bioanalyzer перед секвенированием на NovaSeq 6000.

[00172] *Анализ данных*

[00173] Данные усекали с использованием trimmomatic, за которым следовало выравнивание по hgl9 с использованием BWA. Прочтения подвергли параллельной маркировке с помощью Picard, после чего производили локальную корректировку и повторную калибровку базы с использованием практических рекомендаций GATK 3.5. Все файлы подвергали субдискретизации до указанного числа прочтений с помощью Picard DownSampleSam. Количественные показатели качества получали из окончательного bam-файла с использованием qualimap, а также Picard AlignmentMetricsAummary и CollectWgsMetrics. Строили кривые Лоренца и коэффициенты Джини рассчитывали с использованием htSeqTools. Обнаружение SNV выполняли с помощью UnifiedGenotyper, которые затем фильтровали с использованием стандартных рекомендуемых критериев (QD < 2,0 || FS > 60,0 || MQ < 40,0 || SOR > 4,0 || MQRankSum < -12,5 || ReadPosRankSum < -8,0). Никакие области не исключали из анализа, и не выполняли никакую другую нормализацию данных или манипуляции. Количественные показатели секвенирования для исследованных способов приведены в таблице 2.

[00174] **Таблица 2. Сравнение показателей секвенирования между протестированными методами**

	PTA	**MDA Kit 2**	**PicoPlex**	**MALBAC**	**LIANTI**	**MDA Kit 1**	**DOP PCR**
% картирования генома	97	88	55	79	92	65	52
% извлечения генома (300M прочтений)	95	75	43	60	82	73	23
CV охвата (300M прочтений)	0,8	1,8	3	2,5	1,1	2	3,5
Чувствительность SNV % (300M прочтений)	76	50	15	34	49	46	5
Чувствительность SNV % (300M прочтений)	93	91	56	47	88	90	35

CV = коэффициент вариации, SNV = однонуклеотидная вариация, значения относятся к охвату 15X.

[00175] *Ширина и однородность охвата генома*

[00176] Проводили всесторонние сравнения PTA со всеми распространенными способами WGA отдельных клеток. Для осуществления этого PTA и улучшенную версию MDA, называемую MDA отдельных клеток ((Dong et al. Nat. Meth. 2017, 14, 491- 493) (SCMDA), проводили на 10 клетках NA12878 каждую. Кроме того, эти результаты для клеток, которые подверглись амплификации с помощью DOP-PCR ((Zhang et al. PNAS 1992, 89, 5847-5851), MDA Kit 1 (Dean et al. PNAS 2002, 99, 5261-5266), MDA Kit 2, MALBAC (Zong et al. Science 2012, 338, 1622-1626), LIANTI (Chen et al., Science 2017, 356, 189-194) или PicoPlex (Langmore, Pharmacogenomics 3, 557-560 (2002)) сравнивали с использованием данных, полученных в рамках исследования LIANTI.

[00177] Для нормализации по образцам необработанные данные из всех образцов выравнивали и подвергали предварительной обработке для обнаружения вариантов с использованием одного и того же конвейера. Затем bam-файлы подвергали субдискретизации до 300 миллионов прочтений каждый перед выполнением сравнений. Важно отметить, что продукты PTA и SCMDA не подвергали скринингу перед выполнением дальнейших анализов, в то время как все другие способы подвергали скринингу на охват генома и однородность перед отбором клеток самого высокого качества, которые использовали в последующих анализах. Следует отметить, что SCMDA и PTA сравнивали с объемными диплоидными образцами NA12878, в то время как все другие способы сравнивали с объемными диплоидными фибробластами BJ1, которые использовали в исследовании LIANTI. Как видно на фигурах 3C-3F, у PTA был самый высокий процент прочтений, выровненных с геномом, а также самое высокое качество картирования. PTA, LIANTI и SCMDA имели сходное содержание GC, все из которых были ниже, чем у других способов. Показатели дупликации ПЦР были сходными для всех способов. Кроме того, способ PTA позволил более мелким матрицам, таким как митохондриальный геном, обеспечить более высокие показатели охвата (аналогичные более крупным каноническим хромосомам) по сравнению с другими исследованными способами (фиг. 3G).

[00178] Затем сравнивали широту и однородность охвата всех способов. Примеры графиков охвата хромосомы 1 показаны для SCMDA и PTA, где показано, что PTA имеет значительно улучшенную однородность охвата и частоту аллелей (фиг. 4B). Затем рассчитывали коэффициенты охвата для всех способов с использованием возрастающего числа прочтений. PTA приближается к двум совокупным образцам на каждой глубине, что является значительным улучшением по сравнению со всеми другими способами (фиг. 5 A). Затем авторы настоящего изобретения использовали две стратегии

для измерения однородности охвата. Первый подход заключался в вычислении коэффициента вариации охвата при увеличении глубины секвенирования, когда было обнаружено, что PTA более однородна, чем все другие способы (фиг. 5B). Вторая стратегия заключалась в вычислении кривых Лоренца для каждого субдискретизированного bam-файла, где снова было обнаружено, что PTA характеризуется наибольшей однородностью (фиг. 5C). Чтобы измерить воспроизводимость однородности амплификации, рассчитывались коэффициенты Джини для оценки отличия каждой реакции амплификации от идеальной однородности (de Bourcy et al., PloS one 9, e105585 (2014)). Опять же было показано, что PTA является воспроизводимо более однородной, чем другие способы (фиг. 5D).

[00179] *Чувствительность SNV*

[00180] Чтобы определить влияние этих различий на производительность способов амплификации на обнаружение SNV, показатели частоты обнаружения вариантов для каждого сравнивали с соответствующим совокупным образцом при увеличении глубины секвенирования. Для оценки чувствительности сравнивали процент вариантов, обнаруженных в соответствующих совокупных образцах, которые подвергли субдискретизации до 650 миллионов прочтений, которые были обнаружены в каждой клетке на каждой глубине секвенирования (фиг. 5E). Улучшенные охват и однородность PTA привели к обнаружению на 45,6% больше вариантов по сравнению с набором для MDA 2, который являлся следующим по чувствительности способом. Исследование сайтов, названных гетерозиготными в совокупном образце, показало, что PTA значительно уменьшила аллельный перекос в этих гетерозиготных сайтах (фиг. 5F). Это открытие подтверждает утверждение, что PTA не только характеризуется более равномерной амплификацией по геному, но также более равномерно амплифицирует два аллеля в одной и той же клетке.

[00181] *Специфичность SNV*

[00182] Чтобы оценить специфичность обнаружений мутаций, варианты, обнаруженные в каждой одиночной клетке, не обнаруженные в соответствующем совокупном образце, считались ложноположительными. Низкотемпературный лизис SCMDA значительно снизил количество ложноположительных обнаружений вариантов (фиг. 5G). Способы с использованием термостабильных полимераз (MALBAC, PicoPlex и DOP-PCR) показали дальнейшее снижение специфичности обнаружения SNV с увеличением глубины секвенирования. Без ограничения какой-либо теорией это, вероятно,

является результатом значительного увеличения частоты ошибок этих полимераз по сравнению с ДНК-полимеразой phi29. Кроме того, паттерны замены оснований, наблюдаемые в ложноположительных обнаружениях, также оказываются зависимыми от полимеразы (фиг. 5H). Как видно на фиг. 5G, модель подавления распространения ошибок в PTA поддерживается более низкой частотой ложноположительных обнаружений SNV в PTA по сравнению со стандартными протоколами MDA. Кроме того, PTA характеризуется самыми низкими частотами аллелей ложноположительных обнаружений вариантов, что снова согласуется с моделью подавления распространения ошибки с PTA (фиг. 5I).

[00183] Пример 3. Массовое параллельное секвенирование ДНК отдельных клеток

[00184] Используя PTA, устанавливают протокол массового параллельного секвенирования ДНК. Сначала к случайному праймеру добавляют случайный праймер. Используют две стратегии для минимизации любого смещения при амплификации, вызванного баркодом клетки: 1) удлинение размера случайного праймера и/или 2) создание праймера, который закольцован на себе, чтобы предотвратить связывание баркода клетки с матрицей (фиг. 10B). После того, как оптимальная стратегия праймера определена, вплоть до 384 отсортированных клеток масштабируют с помощью, например, устройства для манипуляций с жидкостями Mosquito HTS, которое может с высокой точностью дозировать даже вязкие жидкости объемом вплоть до 25 нл. Этот жидкостный манипулятор также снижает затраты на реагенты приблизительно в 50 раз за счет использования 1 мкл реакции PTA вместо стандартного реакционного объема, составляющего 50 мкл.

[00185] Протокол амплификации переводят в капли путем доставки праймера с баркодом клетки в каплю. Необязательно используют твердые подложки, такие как шарики, созданные с использованием стратегии смешения-разделения. Подходящие шарики доступны, например, от ChemGenes. Олигонуклеотид в некоторых случаях содержит случайный праймер, баркод клетки, уникальный молекулярный идентификатор и расщепляемую последовательность или спейсер для высвобождения олигонуклеотида после того, как шарик и клетка инкапсулируются в одной и той же капле. Во время этого процесса оптимизируют концентрации матрицы, праймера, dNTP, альфа-тио-ddNTP и полимеразы для объема в нижнем нанолитровом диапазоне в каплях. Оптимизация в некоторых случаях включает применение более крупных капель для увеличения реакционного объема. Как видно на фиг. 9, этот процесс требует двух последовательных реакций для лизиса клеток, за которыми следует WGA. Первую каплю, которая содержит лизированные клетки и шарик, объединяют со второй каплей с амплификационной смесью.

В качестве альтернативы или в комбинации клетку инкапсулируют в шарик гидрогеля перед лизисом, а затем оба шарика можно добавить к капле масла. См. Lan, F. et al., Nature Biotechnol., 2017, 35:640-646).

[00186] Дополнительные способы включают использование микролунок, которые в некоторых случаях захватывают 140000 отдельных клеток в реакционных камерах емкостью 20 пиколитров на устройстве размером с предметное стекло микроскопа 3 дюйма × 2 дюйма. Подобно способам на основе капель, эти лунки объединяют клетку с шариком, который содержит баркод клетки, что позволяет проводить массовую параллельную обработку. См. Gole et al., Nature Biotechnol., 2013, 31:1126-1132).

[00187] **Пример 4. Параллельный анализ генома и транскриптома в отдельных клетках**

[00188] Отдельные клетки из популяции клеток сортируют, помещая по одной клетке в лунку. Каждая лунка содержала антитело, прикрепленное к участку поверхности, где антитело связывается с ядром клетки. Внешнюю мембрану клетки лизировали, высвобождая мРНК в раствор в лунке, в то время как нуклеаза оставалась интактной и связанной с участком лунки. RT проводят с использованием мРНК в растворе в качестве матрицы для создания кДНК с использованием праймеров на фиг. 8A. Необязательно проводят стадию истощения рРНК (рибосомальной РНК). Первый переключающий матрицу праймер, содержащий 5' - 3' область TSS (сайт начала транскрипции), якорную область, область BC РНК и поли-dT-хвост, и второй переключающий матрицу праймер, содержащий 3' - 5' область TSS, якорную область и поли-G-область, использовали для ПЦР с обратной транскрипцией. После удаления продуктов ПЦР с обратной транскрипцией (библиотеки кДНК) для последующего секвенирования любую оставшуюся в клетке РНК удаляли с помощью UNG. Библиотеку РНК готовили с использованием способа и реагентов секвенирования на основе Nextera/транспозона (фиг. 8B). Библиотека кДНК содержала короткие кДНК с приблизительно 1000-кратной амплификацией. Затем ядро лизировали, а высвобожденную геномную ДНК подвергали способу PTA с использованием случайных праймеров с изотермической полимеразой со случайными праймерами с 6-9 основаниями в длину. Условия амплификации для PTA выбирали таким образом, чтобы генерировать ампликоны длиной 250-1500 оснований. Продукты PTA необязательно подвергали дополнительной амплификации и секвенированию. Данные секвенирования РНК и данные секвенирования ДНК собирали в базу данных для анализа.

[00189] Пример 5. Мультиомический анализ отдельной клетки

[00190] Популяцию клеток вводили в контакт с библиотекой антител, где антитела являются мечеными. Антитела являются мечеными флуоресцентными метками, баркодами нуклеиновых кислот или тем и другим. Меченые антитела связывались по меньшей мере с одной клеткой в популяции, и такие клетки сортировали, помещая по одной клетке в лунку. Некоторые меченые антитела предоставляют специфическую информацию о маркерах белков клеточной поверхности после связывания, которую получают либо с помощью флуоресцентной микроскопии, либо с помощью считывания баркодов, которые вносят метки в антитела. Каждая лунка содержала антитело, прикрепленное к участку поверхности, при этом антитело связывалось с ядром клетки. Внешнюю мембрану клетки лизируют, высвобождая мРНК в раствор в лунке, в то время как нуклеаза оставалась интактной и связанной с участком лунки. Необязательно проводили стадию истощения рРНК (рибосомальной РНК). Затем выполняли RT с использованием мРНК в растворе в качестве матрицы для создания кДНК. Первый переключающий матрицу праймер, содержащий 5' - 3' область TSS (сайт начала транскрипции), якорную область, область BC РНК и поли-dT-хвост, и второй переключающий матрицу праймер, содержащий 3' - 5' область TSS, якорную область и поли-G-область, использовали ПЦР с обратной транскрипцией. После удаления продуктов ПЦР с обратной транскрипцией (библиотеки кДНК) для последующего секвенирования любую оставшуюся в клетке РНК удаляли с помощью UNG. Библиотека кДНК содержали короткие кДНК с приблизительно 1000-кратной амплификацией. Затем ядро лизировали, а высвобожденную геномную ДНК подвергали способу PTA с использованием случайных праймеров с изотермической полимеразой со случайными праймерами 6-9 оснований в длину. Условия амплификации для PTA выбирали таким образом, чтобы генерировать ампликоны длиной 250-1500 оснований. Продукты PTA необязательно подвергали дополнительной амплификации и секвенированию. Данные о белках, данные секвенирования РНК и данные секвенирования ДНК собирали в базу данных для анализа.

[00191] Пример 6. Анализ метилома и транскриптома отдельных клеток

[00192] Отдельные клетки из популяции клеток сортировали, помещая по одной клетке в лунку. Каждая лунка содержала антитело, прикрепленное к участку поверхности, при этом антитело связывалось с ядром клетки. Внешнюю мембрану клетки лизировали, высвобождая мРНК в раствор в лунке, в то время как нуклеаза оставалась интактной и связанной с участком лунки. Транскрипты мРНК вводили в контакт с терминальной трансферазой с добавлением рибогуанина к 5'-концу цепей мРНК. Затем

выполняли RT с использованием мРНК в растворе в качестве матрицы для создания кДНК. Необязательно проводили стадию истощения рРНК (рибосомальной РНК). Первый переключающий матрицу праймер, содержащий 5' - 3' область TSS (сайт начала транскрипции), якорную область, область BC РНК и поли-dT-хвост, и второй переключающий матрицу праймер, содержащий 3' - 5' область TSS, якорную область и поли-G-область, использовали ПЦР с обратной транскрипцией. После удаления продуктов ПЦР с обратной транскрипцией (библиотеки кДНК) для последующего секвенирования любую оставшуюся в клетке РНК удаляли с помощью UNG. Библиотека кДНК содержали короткие кДНК с приблизительно 1000-кратной амплификацией. Затем ядро лизировали, а высвобожденную геномную ДНК подвергали способу PTA с использованием случайных праймеров с изотермической полимеразой со случайными праймерами 6-9 оснований в длину. Условия амплификации для PTA выбирали таким образом, чтобы генерировать ампликоны длиной 250-1500 оснований. Продукты PTA необязательно подвергали дополнительной амплификации и секвенированию. Данные о белках, данные секвенирования РНК и данные секвенирования ДНК собирали в базу данных для анализа, и идентифицировали чувствительные к метилированию участки разреза эндонуклеазы. Эти сайты использовали для картирования мест метилирования исходной геномной ДНК.

[00193] Пример 7. Анализ метилома и генома отдельных клеток

[00194] Отдельные клетки из популяции клеток сортировали, помещая по одной клетке в лунку. Каждая лунка содержала антитело, прикрепленное к участку поверхности, где антитело связывается с ядром клетки. Клетки лизировали ферментами, чувствительными к метилированию, и геномы подвергали способу PTA с использованием случайных праймеров изотермической полимеразой со случайными праймерами 6-9 оснований в длину. Условия амплификации для PTA выбирали таким образом, чтобы генерировать ампликоны длиной 250-1500 оснований. Реакционную смесь разделяли, при этом половину смеси подвергали обогащению экзомом, секвенированию всего генома или другому способу целевого секвенирования. Другую половину реакционной смеси подвергали чувствительной к метилированию ПЦР. Данные метилирования и секвенирования ДНК собирали в базу данных для анализа.

[00195] Пример 8. Анализ отдельной клетки поверхностного протеома и генома

[00196] Клетки из образца, содержащего популяцию клеток, вводили в контакт с библиотекой мишеней, таких как антитела, полинуклеотиды или другие

небольшие молекулы. В некоторых случаях мишени маркировали баркодом (например, антитела с баркодом) для того, чтобы можно было выделить и идентифицировать связывание мишени с белками на клеточных поверхностях. В качестве альтернативы или в комбинации мишени маркировали другими меткими, такими как флуоресцентные метки или массовые метки. Отдельные клетки из популяции клеток сортировали, помещая по одной клетке в лунку. Необязательно, мишени, которые связались с клеточной поверхностью, удаляли для секвенирования или идентификации до подготовки геномной библиотеки. Клетки лизировали, геном высвобождался в раствор и генерировались фрагменты. Фрагменты генома подвергали способу PTA с использованием случайных праймеров с изотермической полимеразой со случайными праймерами 6-9 оснований в длину. В качестве альтернативы геном не фрагментировали перед амплификации с помощью PTA. Условия амплификации для PTA выбирали таким образом, чтобы генерировать ампликоны длиной 250-1500 оснований. Продукты PTA необязательно подвергали дополнительной амплификации и секвенированию. Белки клеточной поверхности и данные секвенирования ДНК собирали в базу данных для анализа.

[00197] Пример 9. Мультиомические способы измерения устойчивости к лекарственному средству

[00198] Монотерапия низкомолекулярными ингибиторами, нацеленными на FLT3, при AML (острый миелоидный лейкоз) показала клиническую пользу, но неизменно возникает резистентность. Ингибитор FLT3 квизартиниб (AC220) является одним из таких ингибиторов, при котором лекарственное средство обеспечивает совокупную полную ремиссию приблизительно в 50% случаев у пациентов с рецидивирующим или рефрактерным AML. Несмотря на этот успех, вторичные мутации FLT3 в петле активации (D835) и в контрольном остатке F691 идентифицировали у пациентов с FLT3-ITD, у которых возникал рецидив на терапию квизартинибом. Клиническая устойчивость к мультикиназному ингибитору PKC412 была определена как результат вторичной мутации в киназном домене FLT3. Дополнительные, FLT3-независимые способы резистентности к нацеленной терапии были идентифицированы при FLT3-ITD AML, включая активацию обходного пути AXL, а также мутации NRAS, TET2 и IDH1/2. Также наблюдали мутации в эпигенетических модифицирующих ферментах и факторах транскрипции, что подчеркивает сложность и разнообразие механизмов устойчивости к ингибированию FLT3.

[00199] Создавали устойчивые к квизартинибу и соответствующие родительские клеточные линии MOLM-13 AML, а также клеточную линию, несущую гетерозиготную мутацию FLT3-ITD. Способ PTA сочетал химию RNAseq и использовался

для геномного и транскрипционного исследования этих устойчивых к лекарственным средствам отдельных клеток для того, чтобы получить представление о механизмах устойчивости после ингибирования FLT3 при AML. Кратко, рабочий процесс включал (1) создание резистентных клеток, (2) изоляцию резистентных клеток, (3) цитозольный лизис для высвобождения мРНК, (4) обратную транскрипцию для получения кДНК из мРНК, (5) лизис ядра для высвобождения геномных ДНК, (6) амплификацию PTA, (7) отдельное обогащение ДНК/РНК, (8) PreAMP кДНК обогащенной мРНК, (9) получение библиотеки, контроль качества и объединение, (10) секвенирование следующего поколения и (11) анализ данных.

[00200] *Клеточная культура*. Клетки острого миелоидного лейкоза MOLM-13, несущие гетерозиготную внутреннюю тандемную дупликацию (ITD)1 FLT3, получали из Немецкой коллекции микроорганизмов и клеточных культур DSMZ (ACC 554). Клетки поддерживали в среде RPMI 1640 (Gibco 11875-093) с добавлением 10% FBS и пенициллина/стрептомицина и субкультивировали каждые 2-3 дня, поддерживая диапазон плотности 2,5 E5 – 1,5 E6 клеток/мл. Для создания линии MOLM-13, устойчивой к квизартинибу, клетки постоянно обрабатывали 2 нМ квизартиниба и пополняли лекарственное средство при каждом субкультивировании до появления устойчивых клонов через 5 недель в культуре (фиг. 9A). Геномную ДНК или общую РНК выделяли из устойчивых к квизартинибу и соответствующих родительских клеток MOLM-13 во время сортировки FACS с созданием контрольных библиотек массового секвенирования для сравнения с наборами данных отдельных клеток.

[00201] *FACS.* Для анализа отдельных клеток ~2,0 E6 MOLM-13-устойчивых к квизартинибу или соответствующих родительских клеток дважды промывали в фосфатно-буферном солевом растворе Дульбекко (Gibco) без кальция и магния с добавлением 2% FBS и хранили на льду до сортировки BD FACSAria III FACS. После окрашивания Calcein AM, йодидом пропидия и DAPI устанавливали гейтирование живых клеток (отрицательные DAPI/PI, максимум 70% положительных Calcein-AM) и отдельные клетки сортировали (установка сопел 130 микрон) в 96-луночные планшеты для ПЦР с низким связыванием (полупокрытые), содержащие клеточный буфер, и немедленно замораживали на сухом льду после кратковременного встряхивания и центрифугирования.

[00202] *Комбинированный геномный/транскриптомный анализ.* Во-первых, конъюгированный с биотином олиго-dT-праймер использовали в реакции обратной транскрипции с переключением матрицы для получения кДНК первой цепи из отдельных родительских клеток MOLM-13 или устойчивых к квизартинибу клеток. Амплификацию, управляемую первичной матрицей, (PTA) проводили последовательно после обратной

транскрипции. Затем кДНК первой цепи подвергали аффинной очистке с использованием шариков стрептавидина M-280 и подвергали двум промывкам с высоким содержанием соли, а затем одной промывке с низким содержанием соли. Было проведено 20 циклов преамплификации для получения кДНК второй цепи, и библиотеки для секвенирования РНК получали с использованием набора для получения библиотек Nextera DNA Flex Library Preparation Kit. Для приготовления библиотек PTA продукт PTA, не связанный с шариками стрептавидина, очищали с использованием шариков и лигировали с адаптерами TruSeq. Продукты амплификации реакций PTA сначала очищали с помощью очистки шариками, измеряли с помощью Qubit и анализировали с помощью электрофореза. Типичные выходы для клеток млекопитающих (~ 6 пг ДНК) составляли 1-3 мкг, тогда как отдельные бактериальные геномы (2-4 фг) генерировали до 50 нг. Размер продукта ампликона для образцов, амплифицированных с помощью PTA, составлял от 0,2 до 4 т.н. (в среднем 1,5 т.н.). Библиотеки PTA получали без фрагментации для способов WGS, что дало выход приблизительно 500 нг с диапазоном размеров 300-550 оснований. Полный геном клеток млекопитающих анализировали с помощью NovaSeq с целью ~550 миллионов прочтений. Затем файлы секвенирования переносили для выравнивания обрезки и создания файла VCF и анализировали с помощью облачного решения биоинформатической платформы Trailblazer™. Время подготовки QC и библиотеки составляло 4-6 часов. Для сравнения проводили параллельные эксперименты с использованием только RNASeq.

[00203] *Результаты*. Экспрессия РНК как из родительских, так и из резистентных культур продемонстрировала способность создавать пулы кДНК (фиг. 9B) с использованием химии секвенирования РНК в одном сосуде, и гены, экспрессируемые в этих клетках, создавали различные паттерны, которые позволяют визуализировать клеточные популяции по экспрессии генов на протяжении в среднем ~ 10 тыс. генов, обнаруженных на клетку. В отдельном рабочем процессе геномы отдельных клеток амплифицировали с использованием способа PTA. Два протокола затем объединяли (результаты на фиг. 9D) с получением комбинированных пулов кДНК транскриптома и генома из каждой клетки. Низкий проход (~ 5 миллионов прочтений/клетка) демонстрирует эффективную амплификацию и получение библиотеки как устойчивых, так и родительских линий, имеющих низкое количество митохондриальных хромосом и высокие оценки полного генома PreSeq (фиг. 10A-10C). Данные продемонстрировали, что транскрипты, полученные на стадии обратной транскрипции, неэффективно амплифицируются посредством реакции PTA по сравнению с ДНК, и что ДНК в отдельных клетках эффективно амплифицируется с использованием комбинированного протокола по сравнению с геномами, амплифицированными стандартной PTA, из отдельных клеток (фиг.

9D). Комбинированный способ RNASeq/PTA давал результаты, сходные (фиг. 10A) со стандартным протоколом PTA, где ChrM и процент дубликатов обычно составляли менее 2%, а оценочный размер генома превышал 3 миллиарда оснований (фиг. 10A-10C). Оценка геномов выявила картирование и охват более 90%, а также специфический вызов более 75% однонуклеотидных вариантов в каждой клетке. В двойном протоколе наблюдали большую вариацию по сравнению со стандартной химией PTA генома. Для транскриптома химия прототипа, по-видимому, обнаруживала около 3000-5000 генов, содержащих экзон-экзонные соединения. Приблизительно 30% генов были обнаружены в двойном протоколе (фиг. 10D) по сравнению с протоколом только для RNAseq (фиг. 9C). Кроме того, двойной/комбинированный протокол RNASeq/PTA использовали со второй резистентной клеточной линией SUM159 (линия клеток тройного негативного рака молочной железы). Данные RNAseq, полученные в обоих протоколах, давали одинаковое распределение PCA, что указывает на то, что комбинированный химический анализ способен обнаруживать дифференциальную экспрессию генов, которая не ограничивается одноклеточным типом родительских и резистентных клеток (фиг. 10E-10F).

[00204] Проводили глубокое секвенирование 7 родительских и 5 резистентных клеток molm13 приблизительно до глубины 25x (фиг. 11). Прочтения выравнивали по Hg38 с использованием bwa mem. Контроль качества и вызов SNV осуществляли с использованием лучших практик GATK4. SNV рассматривали только в том случае, если они были ограничены по меньшей мере 2 резистентными клетками, ни в одной родительской клетке не вызывались альтернативные аллели, и по меньшей мере 6 родительских клеток были генотипированы. Все клетки имели по меньшей мере 96% генома, охваченного при 1-кратном охвате, и не менее 76% при 10-кратном охвате. На вставке показано, что известная индель Flt3 в клетках molm13 обнаруживается во всех клетках (4 показаны для ясности).

[00205] Способы RNAseq и PTA, как правило, сопоставимы, при этом как картирование, так и охват превышали 95%, а дубликаты ChrM и ПЦР обычно составляли менее 2,0%. Кроме того, извлекали более 95% генома в отобранных образцах как родительских, так и резистентных клеточных линий. Для клеточной линии Molm13 идентифицировали сверхэкспрессированный ген GAS6 (L), что является известным механизмом устойчивости к квизартинибу. Gas6 является лигандом для AXL, который представляет собой клинически значимый механизм резистентности у пациентов с рецидивом, у которых лечение квизартинибом оказалось неэффективным (фиг. 11B). Глубокое секвенирование генома как родительских, так и устойчивых клеточных линий MOLM13 из двойного протокола выявило мутации, которые были распределены по всем

хромосомам. В совокупности среди всех отдельных клеток было идентифицировано 5675 SNV, уникальных для резистентной к квизартинибу популяции. Обнаруживали вариацию кодирующей последовательности, однако большинство наблюдаемых вариантов находились в межгенном пространстве. Не ограничиваясь теорией, хотя мутации-пассажиры, несомненно, присутствуют в этой когорте вариантов, это предполагает, что регуляция экспрессии генов на уровне энхансера или промотора способствует устойчивости, а также потенциально регуляции некодирующих РНК. Химия двойная химия транскриптома mRNA seq PTA позволяет обнаруживать более 10 тыс. генов в одной клетке, которые могут быть обогащены с помощью FACS. Способ PTA позволяет извлечь более 97% полного генома отдельной клетки. Возможность извлечения как транскриптома, так и генома не оказывает существенного влияния на способность извлекать большую часть генома. Более 70% экспрессированного гена можно обнаружить во многих клетках при сравнении только транскриптомной или комбинированной химии амплификации транскриптома/генома.

[00206] **Пример 10. Анализ отдельных клеток PTA с захватом экзома**

[00207] Использовали общие способы PTA согласно примеру 3 с модификацией: использовали дополнительную стадию захвата экзома для обогащения ампликонов, сгенерированных PTA. Получали 60 миллионов прочтений как для образцов отдельных клеток (27 образцов), так и для объемного образца (112 образцов). Результаты секвенирования захвата экзома из отдельных клеток сравнивали с результатами для общего образца (фиг. 12A-12D, 13A, 14A и 14B). Результаты секвенирования были одинаковыми для нескольких образцов (фиг. 13A), а средний размер захваченных ампликонов составлял 623 основания (фиг. 13B).

[00208] **Пример 11. Захват экзома + Мультиомические способы**

[00209] Использовали общий способ любого из примеров 5-8 с модификацией: использовали дополнительную стадию захвата для обогащения ампликонов, генерируемых PTA, полученных из геномной ДНК. Стадия захвата включала либо панель экзома, либо другую панель, нацеленную на определенные гены. В некоторых случаях такие панели направлены на очаги рака, вирусные геномы или митохондриальную ДНК.

[00210] В настоящем документе описаны примеры, специалистам в данной области техники очевидно, что такие варианты осуществления приведены только в качестве примера. Многочисленные вариации, изменения и замены теперь будут очевидны специалистам в настоящей области техники, не отклоняясь от настоящего изобретения.

Следует понимать, что различные альтернативы описанным в настоящем документе вариантам осуществления изобретения можно использовать при практическом применении настоящего изобретения. Предполагается, что нижеследующая формула изобретения определяет объем настоящего изобретения, и что способы и структуры в рамках настоящей формулы изобретения и их эквиваленты охвачены ею.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ мультиомического анализа отдельных клеток, предусматривающий:

a. выделение отдельной клетки из популяции клеток,

b. секвенирование библиотеки кДНК, содержащей полинуклеотиды, амплифицированные из транскриптов мРНК из отдельной клетки, и

c. секвенирование генома отдельной клетки, где секвенирование генома предусматривает:

i. приведение в контакт генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по меньшей мере один терминаторный нуклеотид, который останавливает репликацию нуклеиновых кислот полимеразой, и

ii. амплификацию по меньшей мере части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи,

iii. лигирование молекул, полученных на стадии (ii), с адапторами, с получением таким образом библиотеки геномной ДНК, и

iv. секвенирование библиотеки геномной ДНК.

2. Способ по п. 1, где транскрипты мРНК содержат полиаденилировнные транскрипты мРНК.

3. Способ по п. 1, где транскрипты мРНК не содержат полиаденилировнные транскрипты мРНК.

4. Способ по п. 1, где секвенирование библиотеки кДНК предусматривает амплификацию транскриптов мРНК с переключающими матрицу праймерами.

5. Способ по п. 1, где по меньшей мере некоторые из полинуклеотидов библиотеки кДНК содержат баркод.

6. Способ по п. 5, где баркод содержит баркод клетки или баркод образца.

7. Способ по п. 1, где библиотеку кДНК и библиотеку геномной ДНК объединяют перед секвенированием.

8. Способ по п. 1, где отдельная клетка представляет собой первичную клетку.

9. Способ по п. 1, где отдельные клетки происходят из печени, кожи, почки, крови или легкого.

10. Способ по п. 1, где отдельная клетка представляет собой раковую клетку, нейрон, глиальную клетку или фетальную клетку.

11. Способ по п. 1, где отдельную клетку выделяют посредством проточной цитометрии.

12. Способ по п. 1, где способ дополнительно предусматривает удаление по меньшей мере одного терминаторного нуклеотида из терминированных продуктов амплификации.

13. Способ по п. 1, где множество терминированных продуктов амплификации содержат в среднем 1000-2000 оснований в длину.

14. Способ по п. 1, где множество терминированных продуктов амплификации составляют 250-1500 оснований в длину.

15. Способ по п. 1, где множество терминированных продуктов амплификации содержат по меньшей мере 97% генома отдельной клетки.

16. Способ по п. 1, где по меньшей мере некоторые из продуктов амплификации содержат баркод клетки или баркод образца.

17. Способ по п. 1, где секвенирование библиотеки кДНК содержит цитозольный лизис отдельной клетки и обратную транскрипцию.

18. Способ по п. 1, где транскрипты мРНК амплифицируют посредством обратной транскрипции с переключением матрицы.

19. Способ по п. 1, где библиотека кДНК содержит по меньшей мере 10000 генов.

20. Способ по п. 1, где секвенирование генома отдельной клетки дополнительно предусматривает лизис ядра отдельной клетки.

21. Способ по п. 1, где способ дополнительно предусматривает дополнительную стадию амплификации с применением ПЦР.

22. Способ по п. 1, где по меньшей мере одну мутацию идентифицируют в геноме клетки, где мутация отлична от соответствующего положения в эталонной последовательности.

23. Способ по п. 1, где по меньшей мере одна мутация происходит в менее 1% популяции клеток.

24. Способ по п. 1, где по меньшей мере одна мутация происходит в не более 0,1% популяции клеток.

25. Способ по п. 1, где по меньшей мере одна мутация происходит в не более 0,001% популяции клеток.

26. Способ по п. 1, где по меньшей мере одна мутация происходит в не более 1% последовательностей продукта амплификации.

27. Способ по п. 1, где по меньшей мере одна мутация происходит в не более 0,1% последовательностей продукта амплификации.

28. Способ по п. 1, где по меньшей мере одна мутация происходит в не более 0,001% последовательностей продукта амплификации.

29. Способ мультиомического анализа отдельных клеток, предусматривающий:

a. выделение отдельной клетки из популяции клеток,

b. идентификацию по меньшей мере одного белка на поверхности отдельной клетки и

c. секвенирование генома отдельной клетки, где секвенирование генома предусматривает:

i. приведение в контакт генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по меньшей мере один терминаторный нуклеотид, который останавливает репликацию нуклеиновых кислот полимеразой,

ii. амплификацию по меньшей мере части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи,

iii. лигирование молекул, полученных на стадии (ii), с адапторами, с получением таким образом библиотеки геномной ДНК, и

iv. секвенирование библиотеки геномной ДНК.

30. Способ по п. 29, где идентификация по меньшей мере одного белка на поверхности клетки предусматривает контакт клетки с меченным антителом, которое связывается с по меньшей мере одним белком.

31. Способ по п. 30, где меченное антитело содержит по меньшей мере одну флуоресцентную метку или массовую метку.

32. Способ по п. 30, где меченное антитело содержит по меньшей мере один баркод нуклеиновой кислоты.

33. Способ мультиомического анализа отдельных клеток, предусматривающий:

a. выделение отдельной клетки из популяции клеток,

b. секвенирование генома отдельной клетки, где секвенирование генома клетки предусматривает:

i. расщепление генома чувствительным к метилированию ферментом рестрикции с получением фрагментов генома,

ii. контакт по меньшей мере некоторых фрагментов генома с по меньшей мере одним праймером амплификации, по меньшей мере одной полимеразой нуклеиновой кислоты и смесью нуклеотидов, где смесь нуклеотидов содержит по

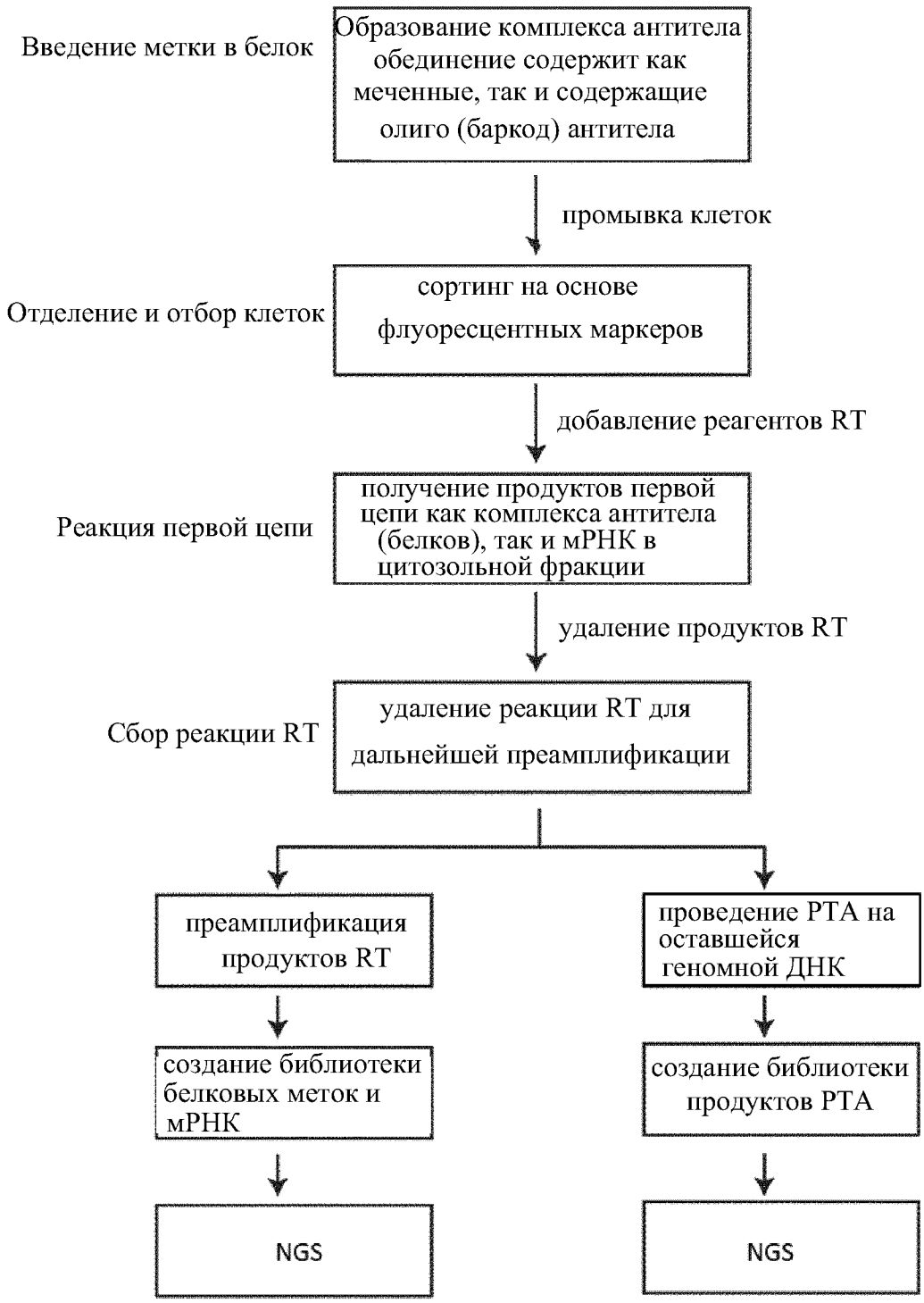
меньшей мере один терминаторный нуклеотид, который останавливает репликацию нуклеиновых кислот полимеразой,

iii. амплификацию по меньшей мере части генома с получением множества терминированных продуктов амплификации, где репликация протекает посредством репликации с замещением цепи,

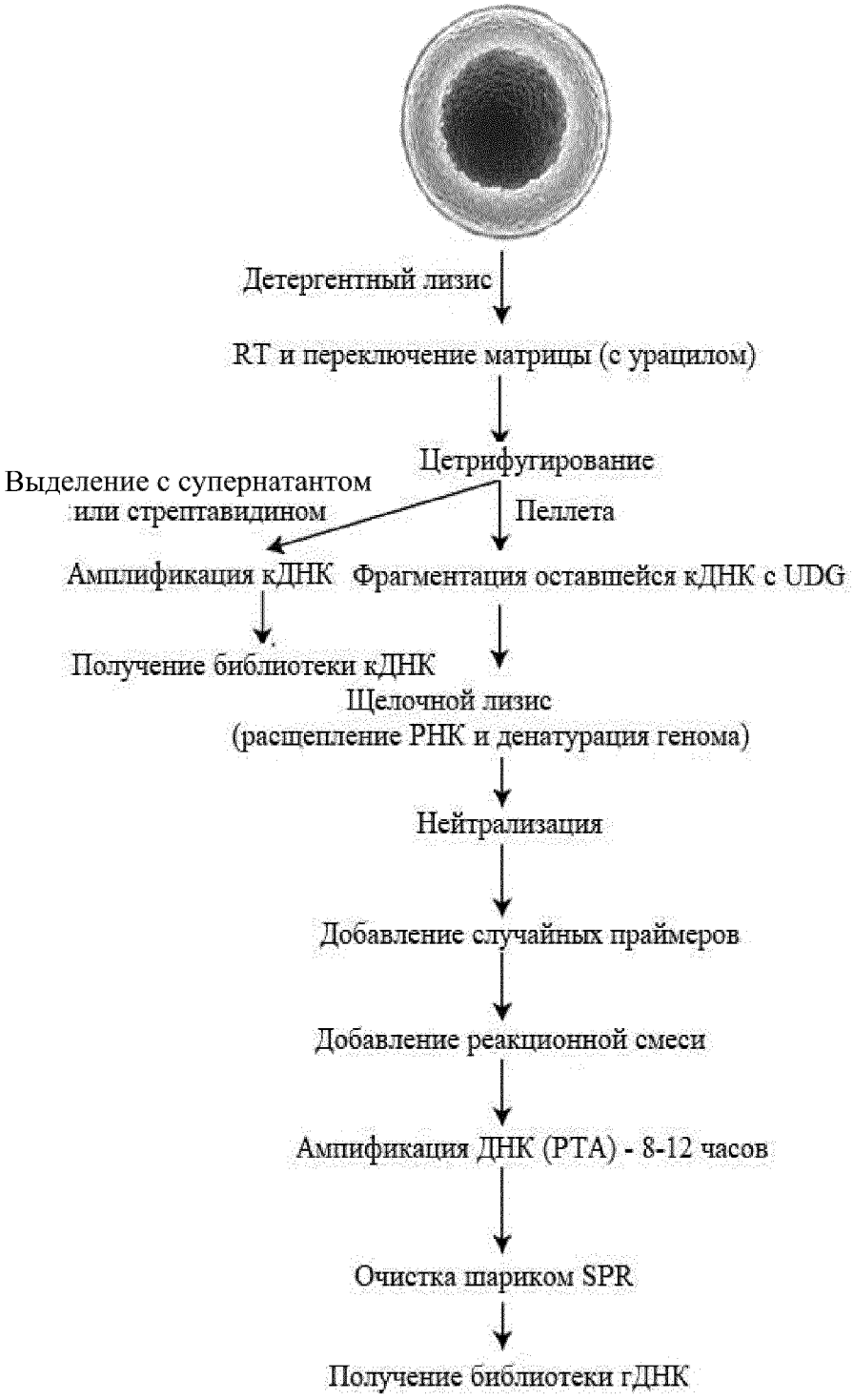
iv. амплификацию по меньшей мере некоторых фрагментов генома посредством специфической к метилированию ПЦР,

v. лигирование молекул, полученных на стадиях (iii и iv), с адапторами, с получением таким образом библиотеки геномной ДНК и библиотеки метиломы ДНК, и

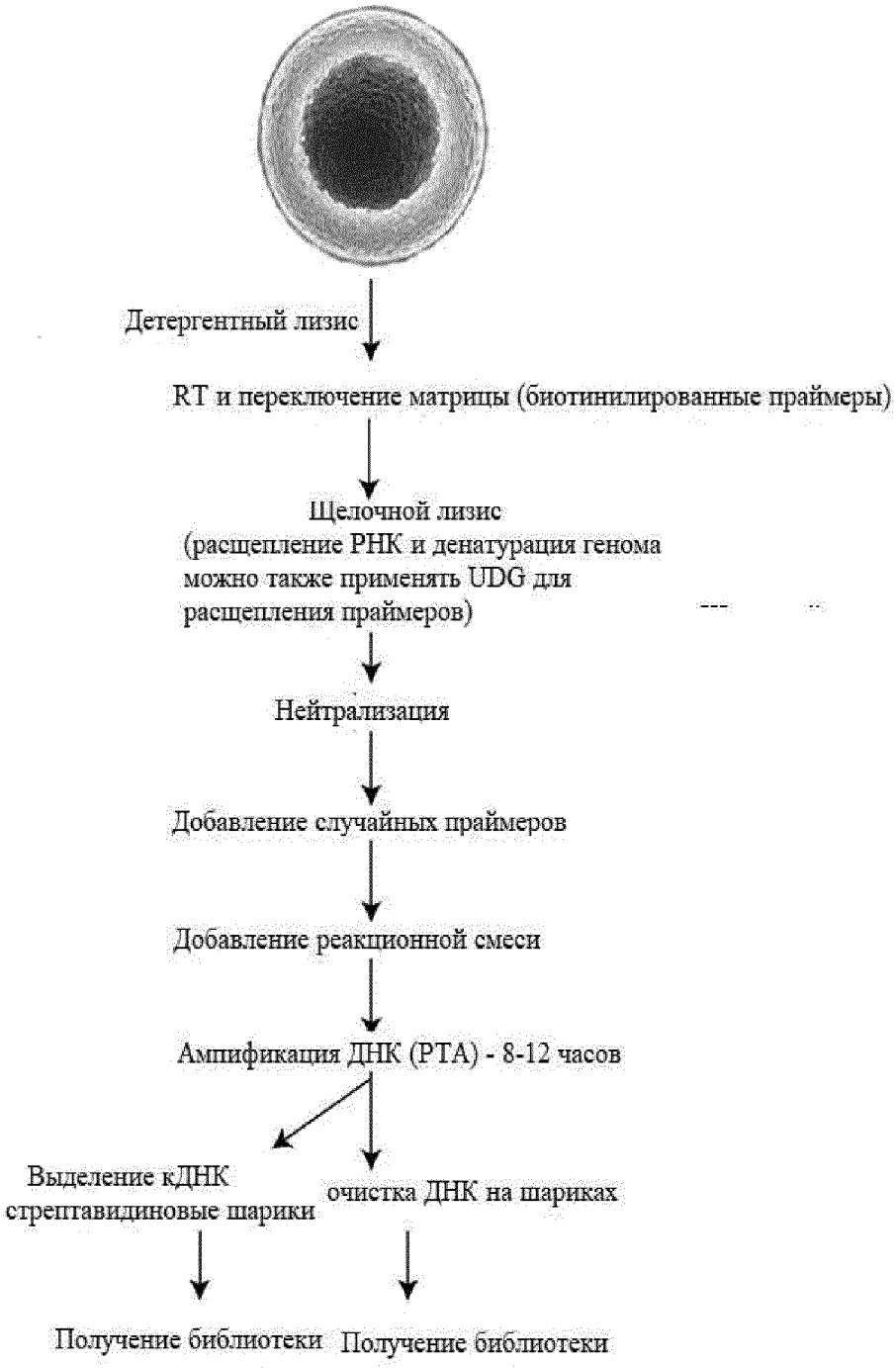
vi. секвенирование библиотеки геномной ДНК и библиотеки метиломы.


Введение метки в белок
Образование комплекса антитела обединение содержит как меченные, так и содержащие олиго (баркод) антитела
промывка клеток
Отделение и отбор клеток
сортинг на основе флуоресцентных маркеров
добавление реагентов RT
Реакция первой цепи
получение продуктов первой цепи как комплекса антитела (белков), так и мРНК в цитозольной фракции
удаление продуктов RT
Сбор реакции RT
удаление реакции RT для дальнейшей преамплификации
преамплификация продуктов RT
проведение PTA на оставшейся геномной ДНК
создание библиотеки белковых меток и мРНК
создание библиотеки продуктов PTA
NGS
NGS

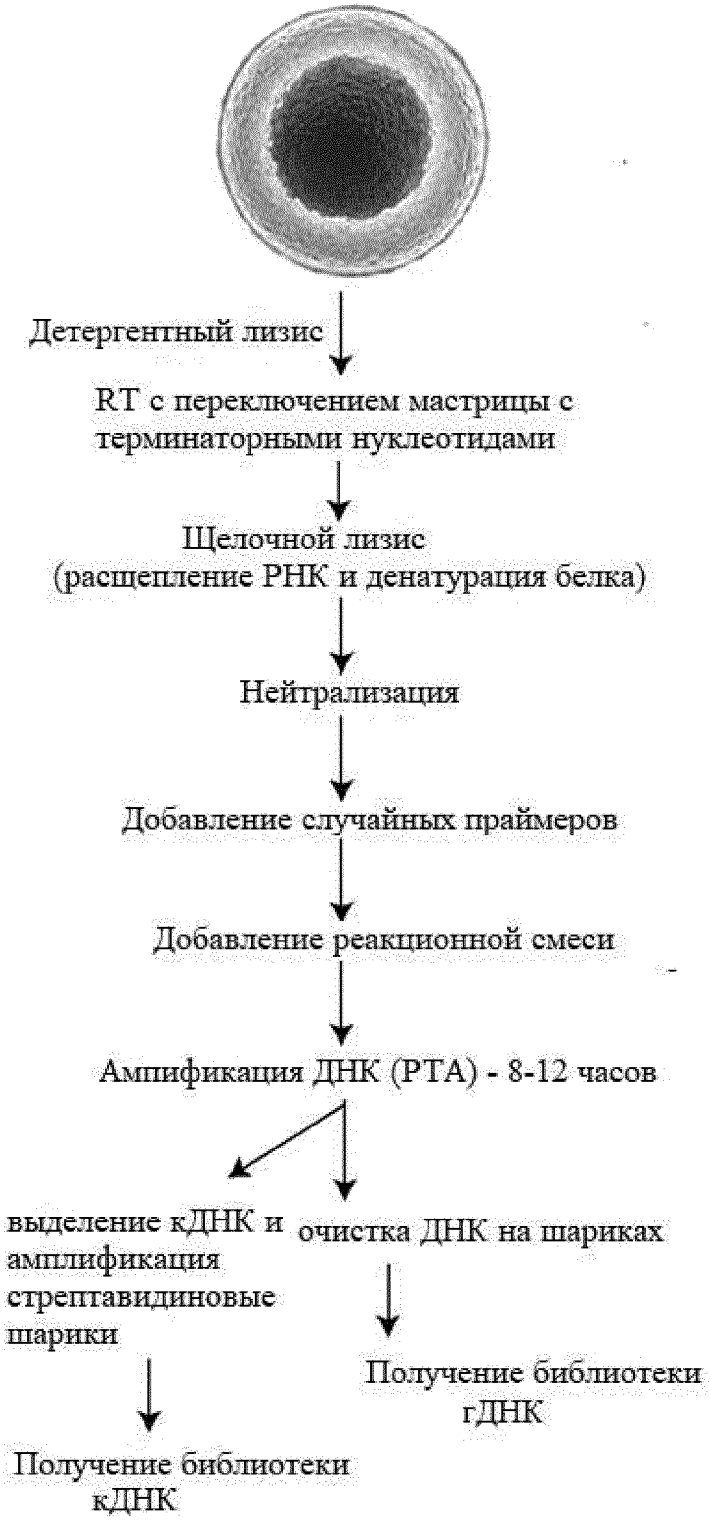
ФИГ. 1А


Детергентный лизис
RT и переключение матрицы (с урацилом)
Цетрифугирование
Выделение с супернатантом
или стрептавидином
Пеллета
Амплификация кДНК
Фрагментация оставшейся кДНК с UDG
Получение библиотеки кДНК
Щелочной лизис
(расщепление РНК и денатурация генома)
Нейтрализация
Добавление случайных праймеров
Добавление реакционной смеси
Ампификация ДНК (PTA) - 8-12 часов
Очистка шариком SPR
Получение библиотеки гДНК

ФИГ. 1В


Детергентный лизис
RT и переключение матрицы (биотинилированные праймеры)
Щелочной лизис
(расщепление РНК и денатурация генома
можно также применять UDG для
расщепления праймеров)
Нейтрализация
Добавление случайных праймеров
Добавление реакционной смеси
Ампификация ДНК (PTA) - 8-12 часов
Выделение кДНК
стрептавидиновые шарики
очистка ДНК на шариках
Получение библиотеки
Получение библиотеки

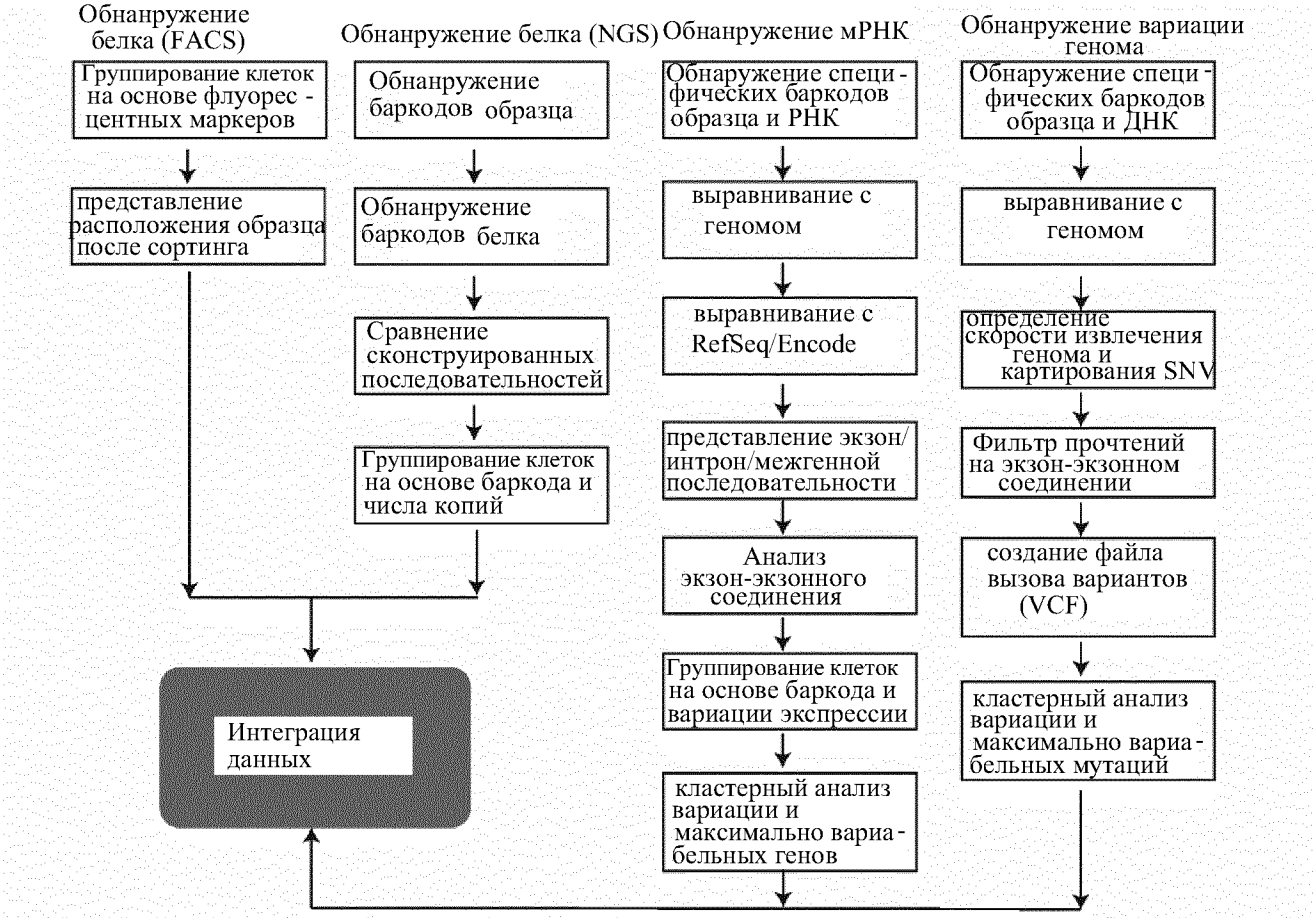
ФИГ. 1С


Детергентный лизис
RT с переключением мастрицы с терминаторными нуклеотидами
Щелочной лизис (расщепление РНК и денатурация белка)
Нейтрализация
Добавление случайных праймеров
Добавление реакционной смеси
Ампификация ДНК (PTA) - 8-12 часов
выделение кДНК и амплификация стрептавидиновые шарики
очистка ДНК на шариках
Получение библиотеки гДНК
Получение библиотеки кДНК

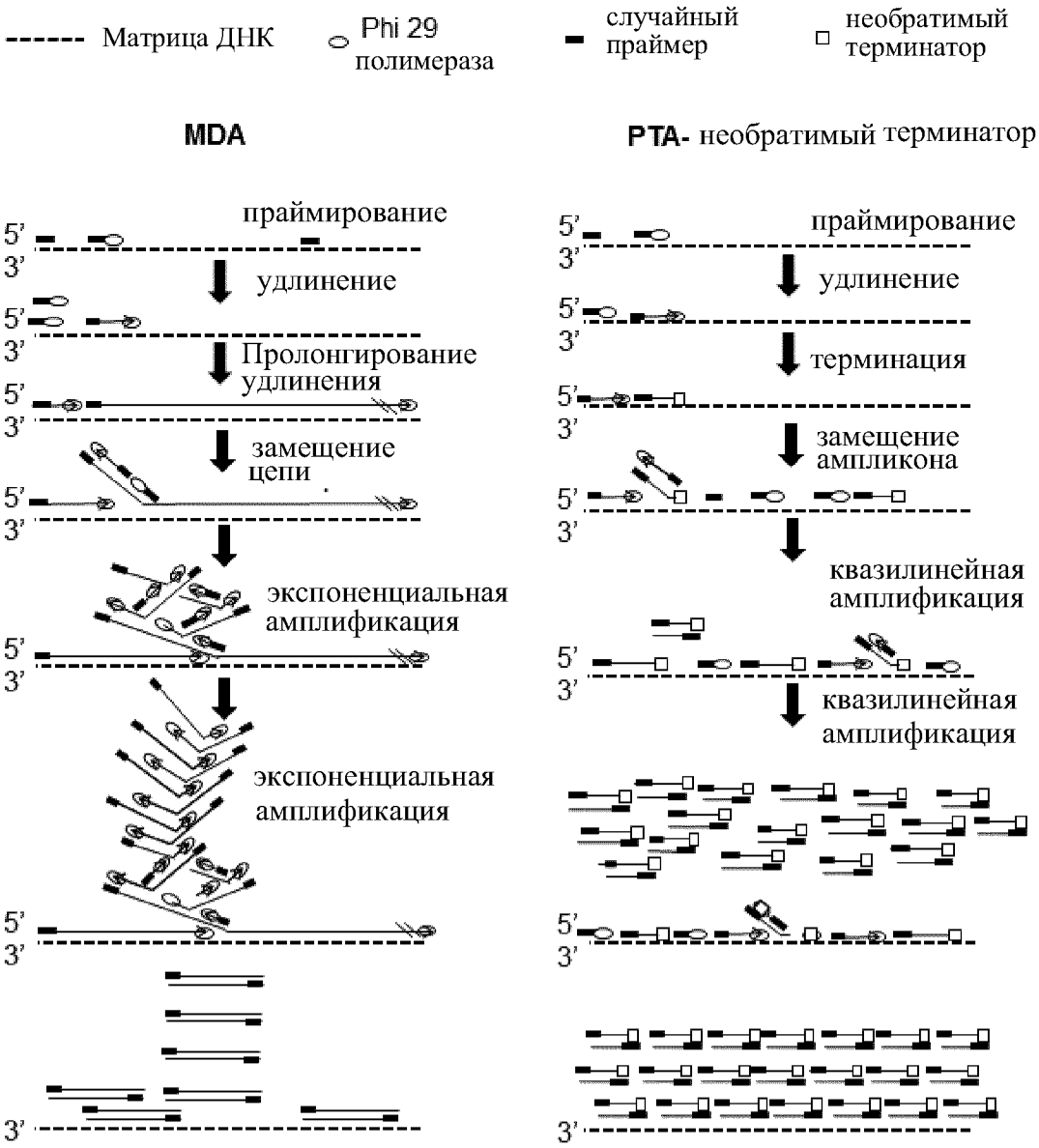
ФИГ. 1D

Детергентный лизис
RT с переключением мастрицы
Щелочной лизис
(расщепление РНК и денатурация белка)
Нейтрализация
Добавление случайных праймеров
(позднее удаление RT праймеров)
Добавление реакционной смеси
Ампификация ДНК (PTA) - 8-12 часов
Амплификация РНКазы кДНК с применением блокированных и биотинилированных праймеров
выделение кДНК и амплификация стрептавидиновые шарики
очистка ДНК на шариках
Получение библиотеки гДНК
Получение библиотеки кДНК

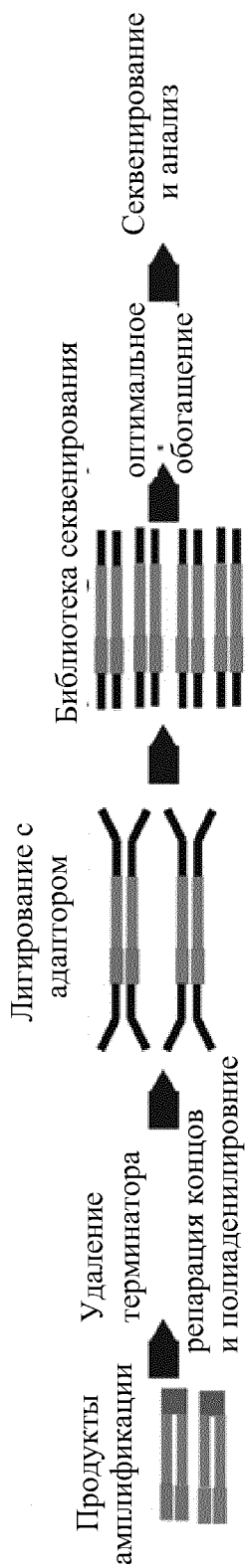
ФИГ. 1E


Обнаружение белка (FACS)
Группирование клеток на основе флуорес - центных маркеров
представление расположения образца после сортинга
Обнаружение белка (NGS)
Обнаружение баркодов образца
Обнаружение баркодов белка
Сравнение сконструированных последовательностей
Группирование клеток на основе баркода и числа копий
Обнаружение мРНК
Обнаружение специ - фических баркодов образца и РНК
выравнивание с геномом
выравнивание с RefSeq/Encode
представление экзон/ интрон/межгенной последовательности
Анализ экзон-экзонного соединения
Группирование клеток на основе баркода и вариации экспрессии
кластерный анализ вариации и максимально вариа - бельных генов
Обнаружение вариации генома
Обнаружение специ - фических баркодов образца и ДНК
выравнивание с геномом
определение скорости извлечения генома и картирования SNV
Фильтр прочтений на экзон-экзонном соединении
создание файла вызова вариантов (VCF)
кластерный анализ вариации и максимально вариа - бельных мутаций
Интеграция данных

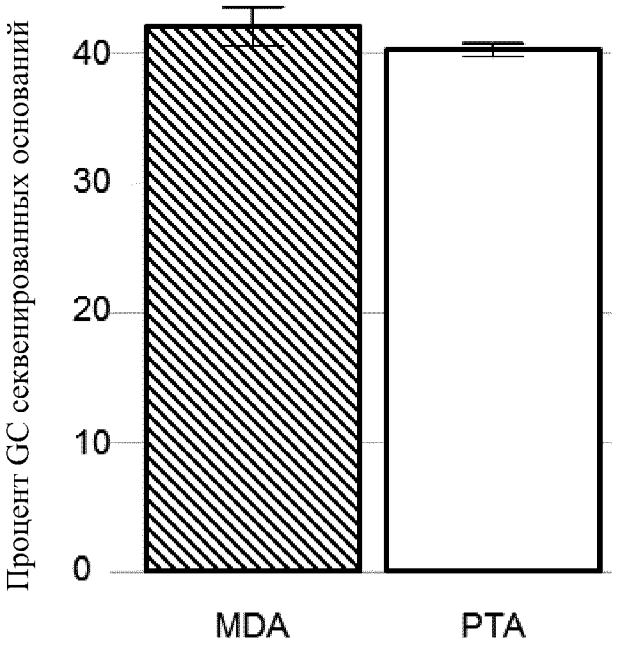
ФИГ. 1F


Матрица ДНК
Phi 29 полимераза
случайный праймер
необратимый терминатор
MDA
PTA- необратимый терминатор
праймирование
удлинение
Пролонгирование удлинения
замещение цепи
экспоненциальная амплификация
экспоненциальная амплификация
праймирование
удлинение
терминация
замещение ампликона
квазилинейная амплификация
квазилинейная амплификация
5'
3'

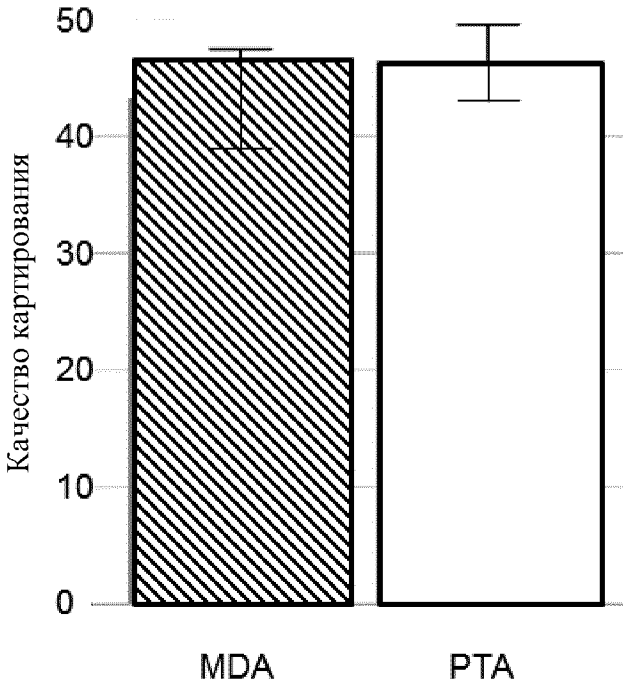
ФИГ. 1G


Продукты амплификации
Удаление терминатора
репарация концов и полиаденилировние
Лигирование с адаптором
Библиотека секвенирования
оптимальное обогащение
Секвенирование и анализ

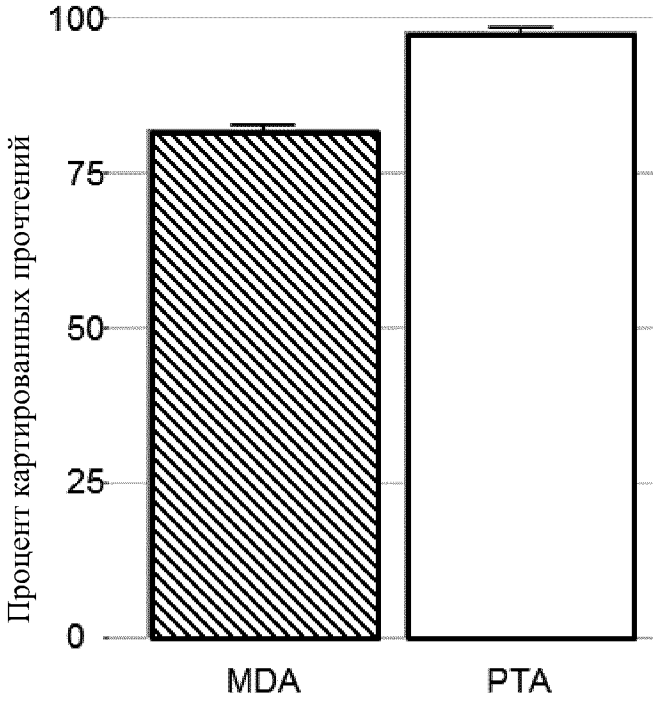
ФИГ. 2А


Процент GC секвенированных оснований
40
30
20
10
0
MDA
PTA

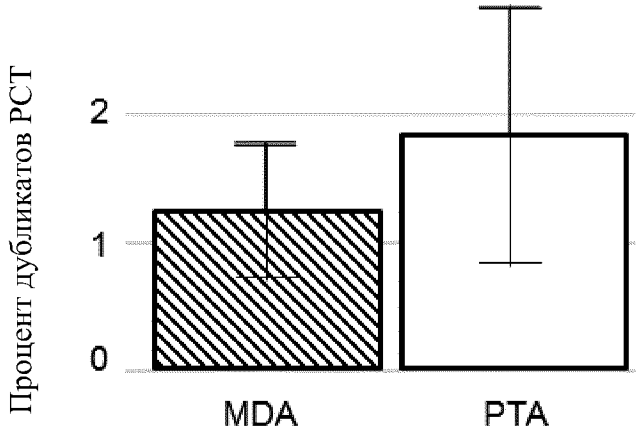
Фиг.2B


Качество картирования
50
40
30
20
10
0
MDA
PTA

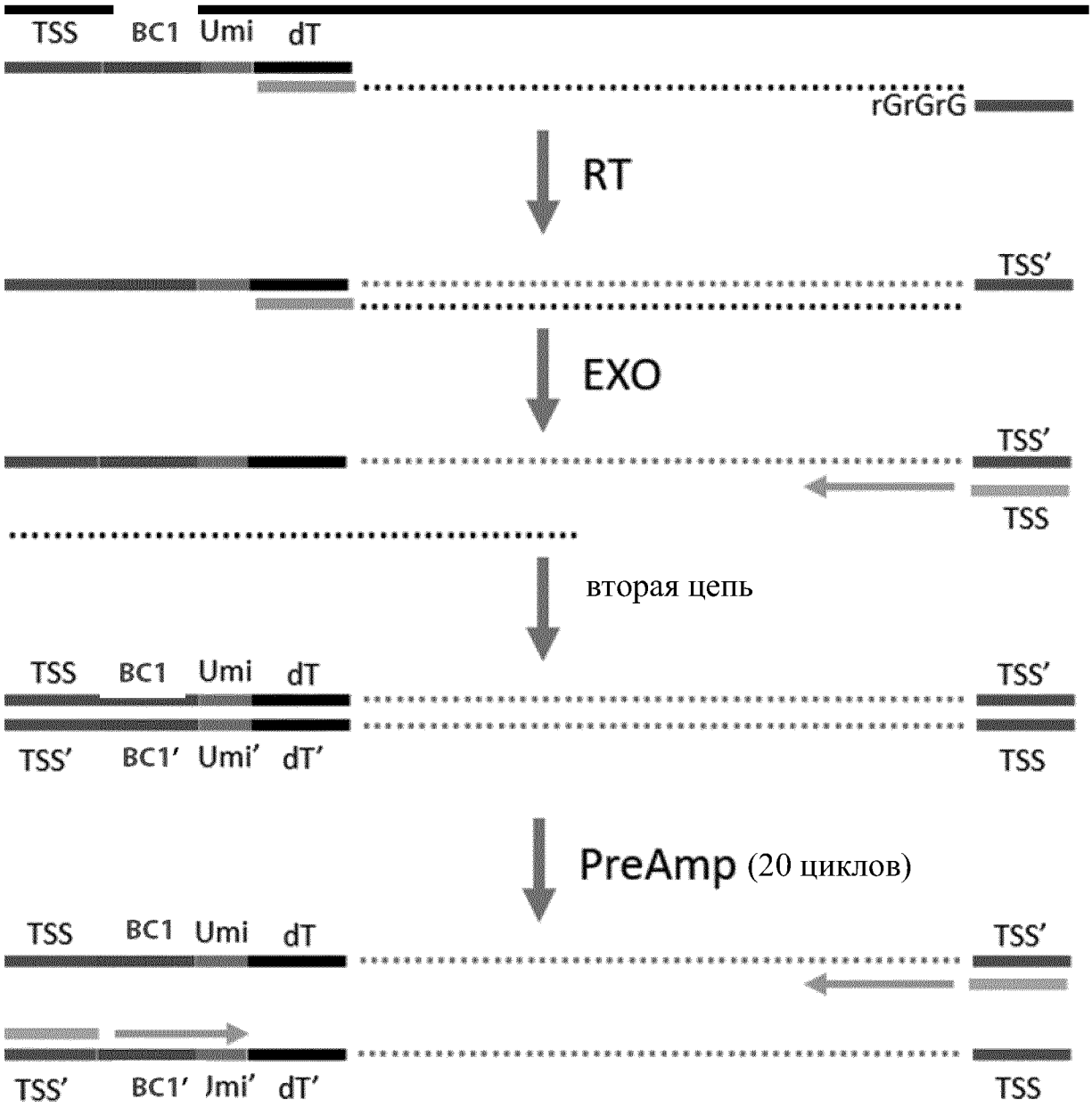
Фиг.2C


Процент картированных прочтений
100
75
50
25
0
MDA
PTA

Фиг.2D

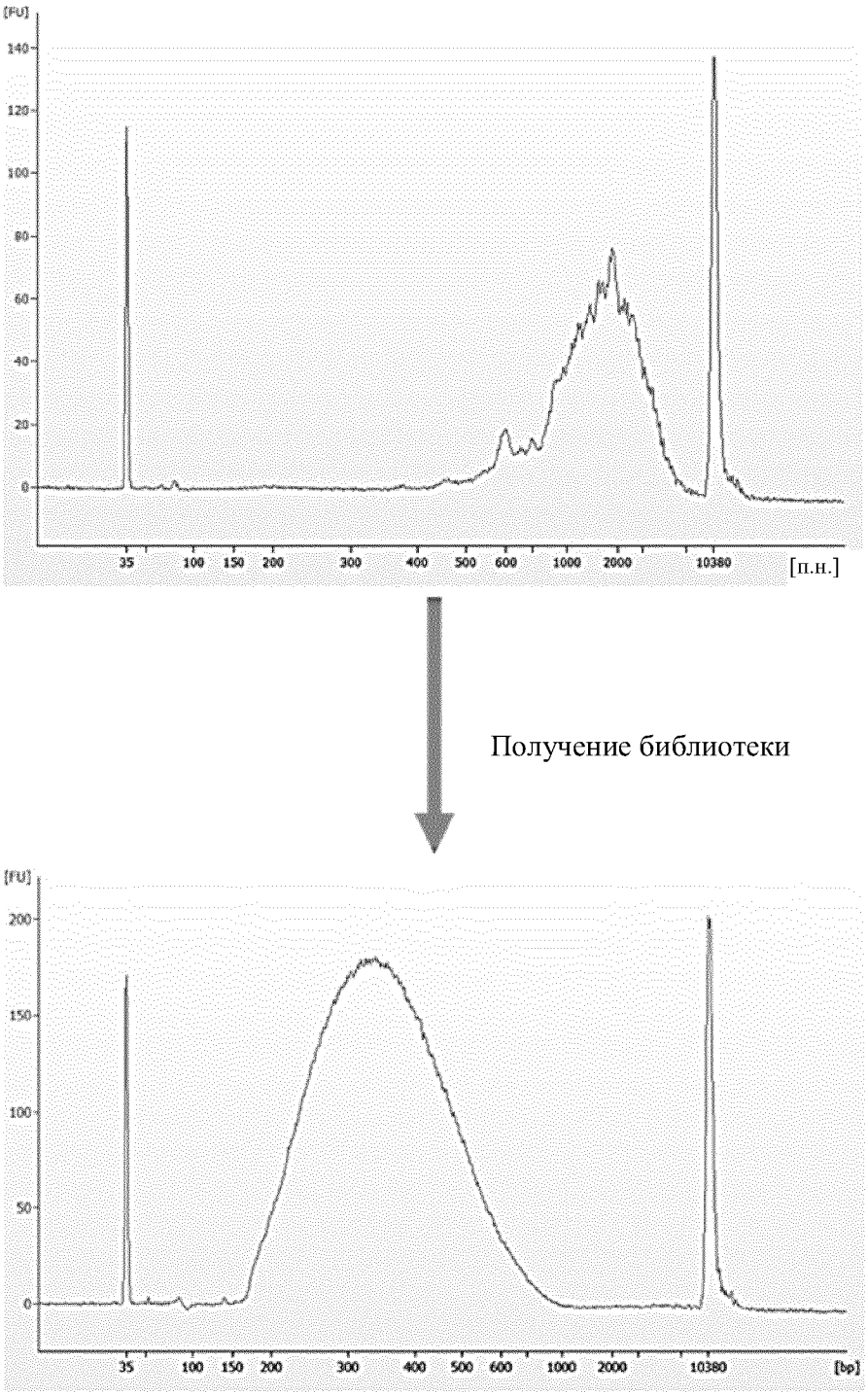

Процент дубликатов PCT
2
1
0
MDA
PTA

Фиг.2E


TSS
BC1
Umi
dT
rGrGrG
RT
TSS'
EXO
TSS'
TSS
вторая цепь
TSS
BC1
Umi
dT
TSS'
TSS'
BC1'
Umi'
dT'
TSS
PreAmp (20 циклов)
TSS
BC1
Umi
dT
TSS'
TSS'
BC1'
Jmi'
dT'
TSS

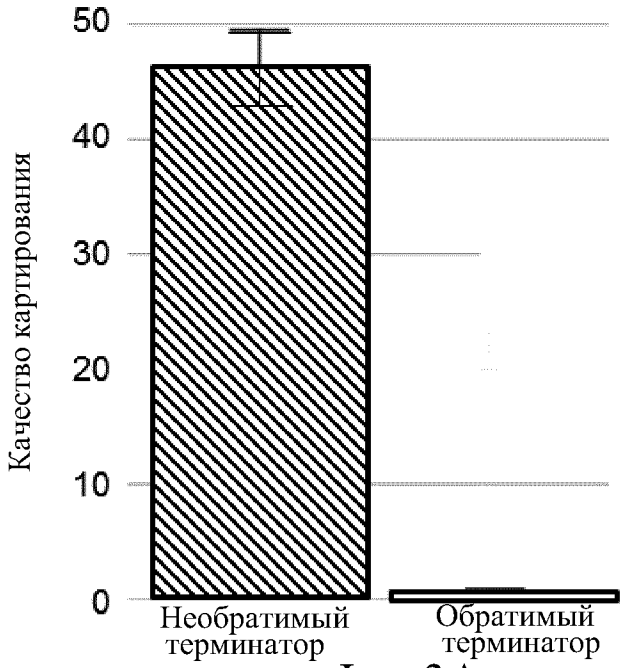
ФИГ. 2F

Объединение кДНК, полученных посредством RT

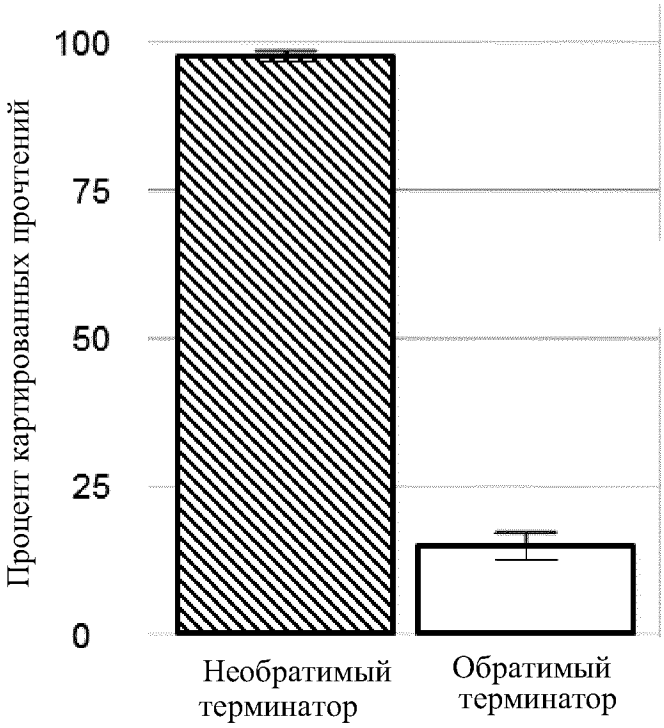


Библиотека, полученная на основе транспозона (nextera XT)

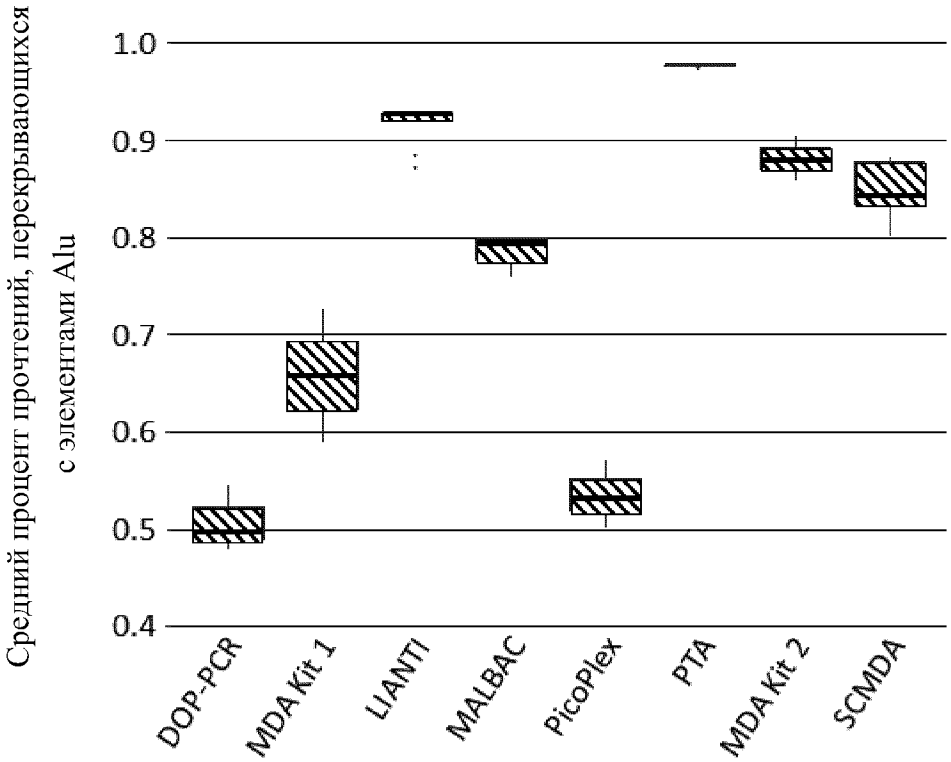
Фиг. 2G


Качество картирования
50
40
30
20
10
0
Необратимый терминатор
Обратимый терминатор

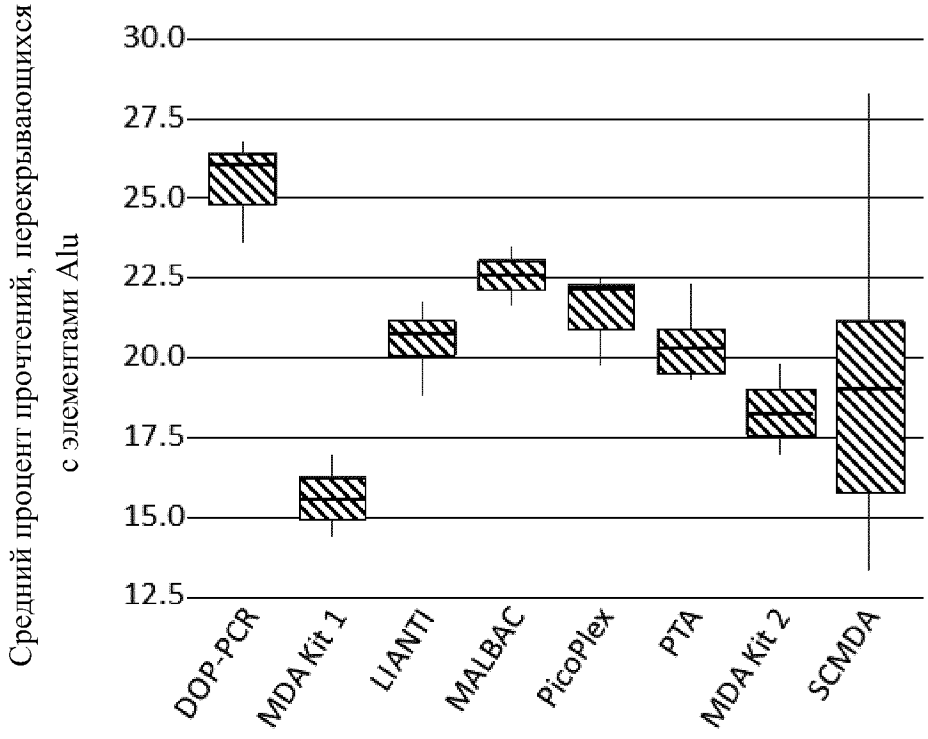
Фиг.3А


Процент картированных прочтений
100
75
50
25
0
Необратимый терминатор
Обратимый терминатор

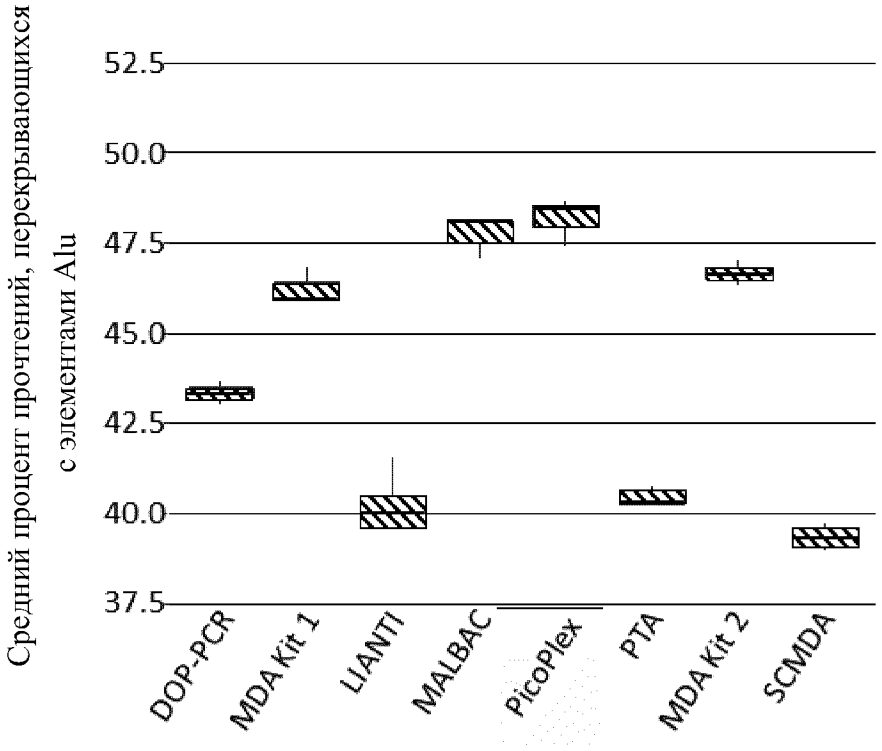
Фиг.3В


Средний процент прочтений, перекрывающихся с элементами Alu
1.0
0.9
0.8
0.7
0.6
0.5
0.4
DOP-PCR
MDA Kit 1
LIANTI
MALBAC
PicoPlex
PTA
MDA Kit 2
SCMDA

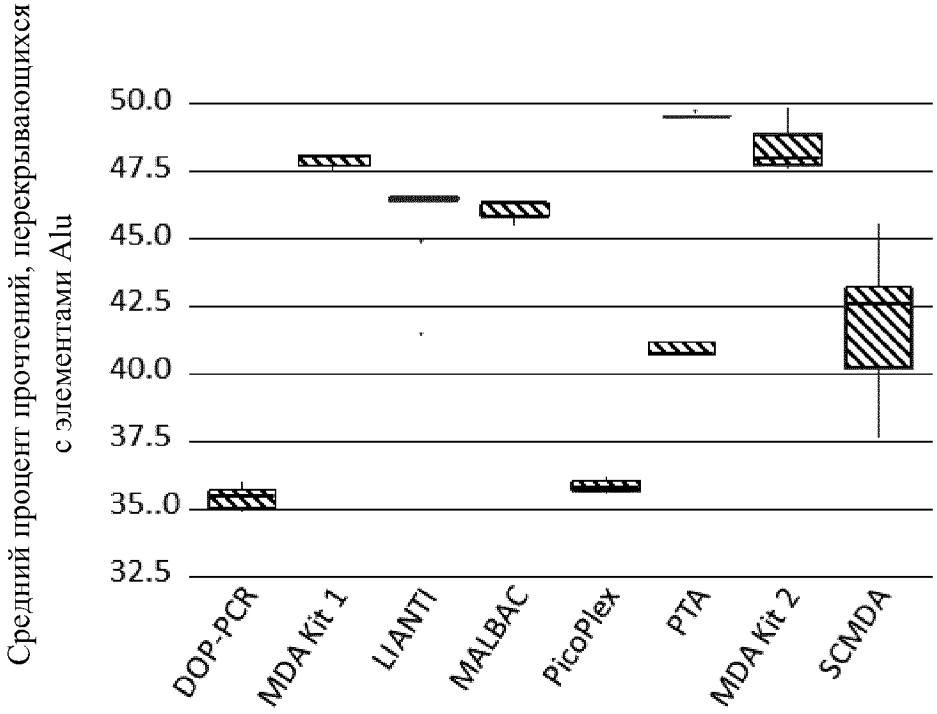
Фиг.3C


Средний процент прочтений, перекрывающихся с элементами Alu
30.0
27.5
25.0
22.5
20.0
17.5
15.0
12.5
DOP-PCR
MDA Kit 1
LIANTI
MALBAC
PicoPlex
PTA
MDA Kit 2
SCMDA

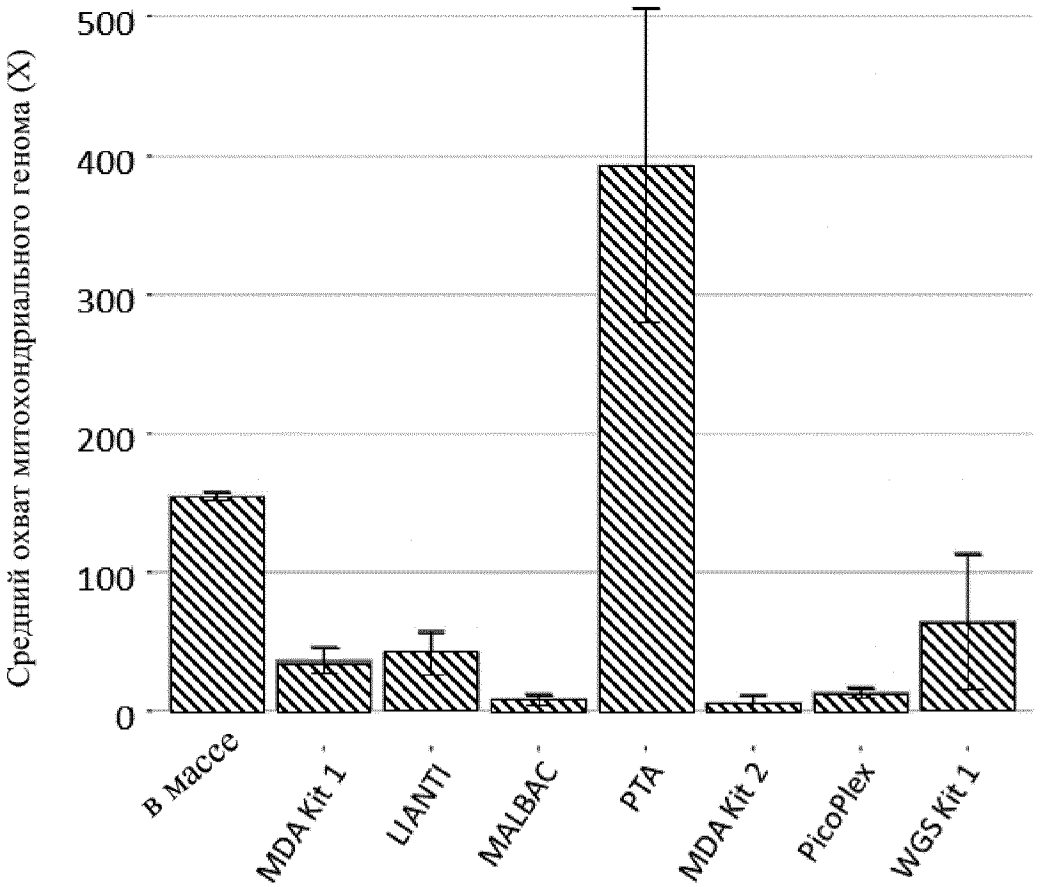
Фиг.3D


Средний процент прочтений, перекрывающихся с элементами Alu
52.5
50.0
47.5
45.0
42.5
40.0
37.5
DOP-PCR
MDA Kit 1
LIANTI
MALBAC
PicoPlex
PTA
MDA Kit 2
SCMDA

Фиг.3E

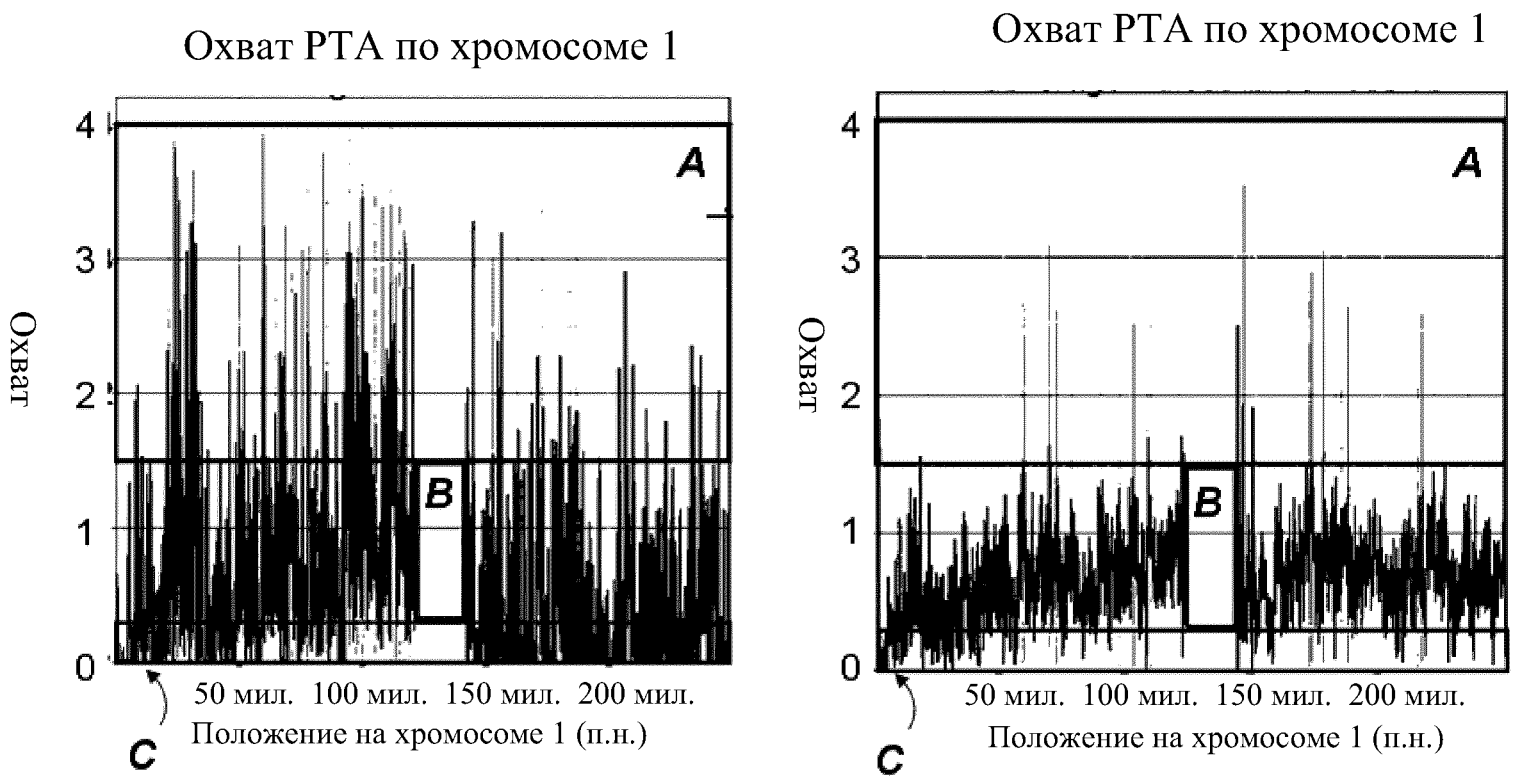

Средний процент прочтений, перекрывающихся с элементами Alu
50.0
47.5
45.0
42.5
40.0
37.5
35..0
32.5
DOP-PCR
MDA Kit 1
LIANTI
MALBAC
PicoPlex
PTA
MDA Kit 2
SCMDA

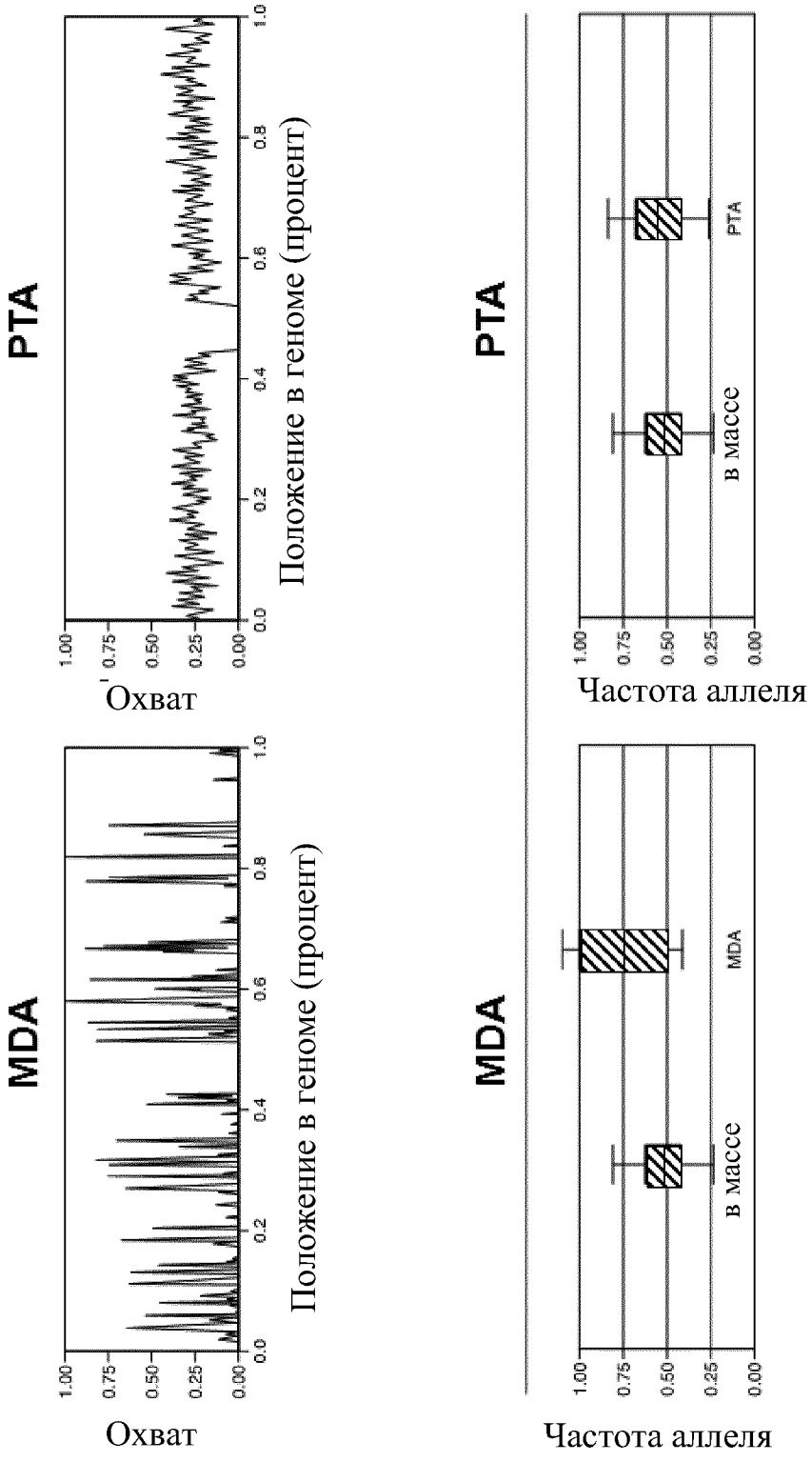
Фиг.3F


Средний охват митохондриального генома (X)
500
400
300
200
100
0
в массе
MDA Kit 1
LIANTI
MALBAC
PTA
MDA Kit 2
PicoPlex
WGS Kit 1

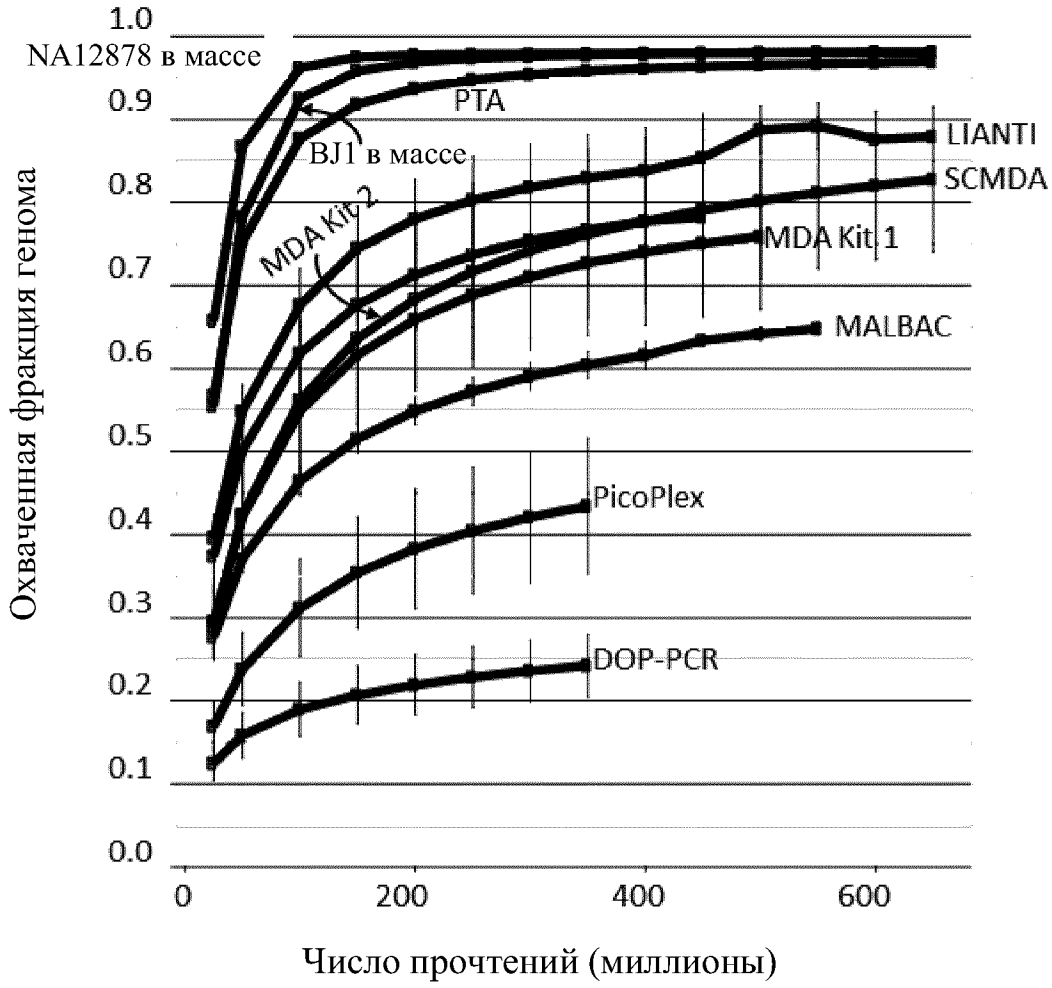
Фиг.3G

Фиг.4А

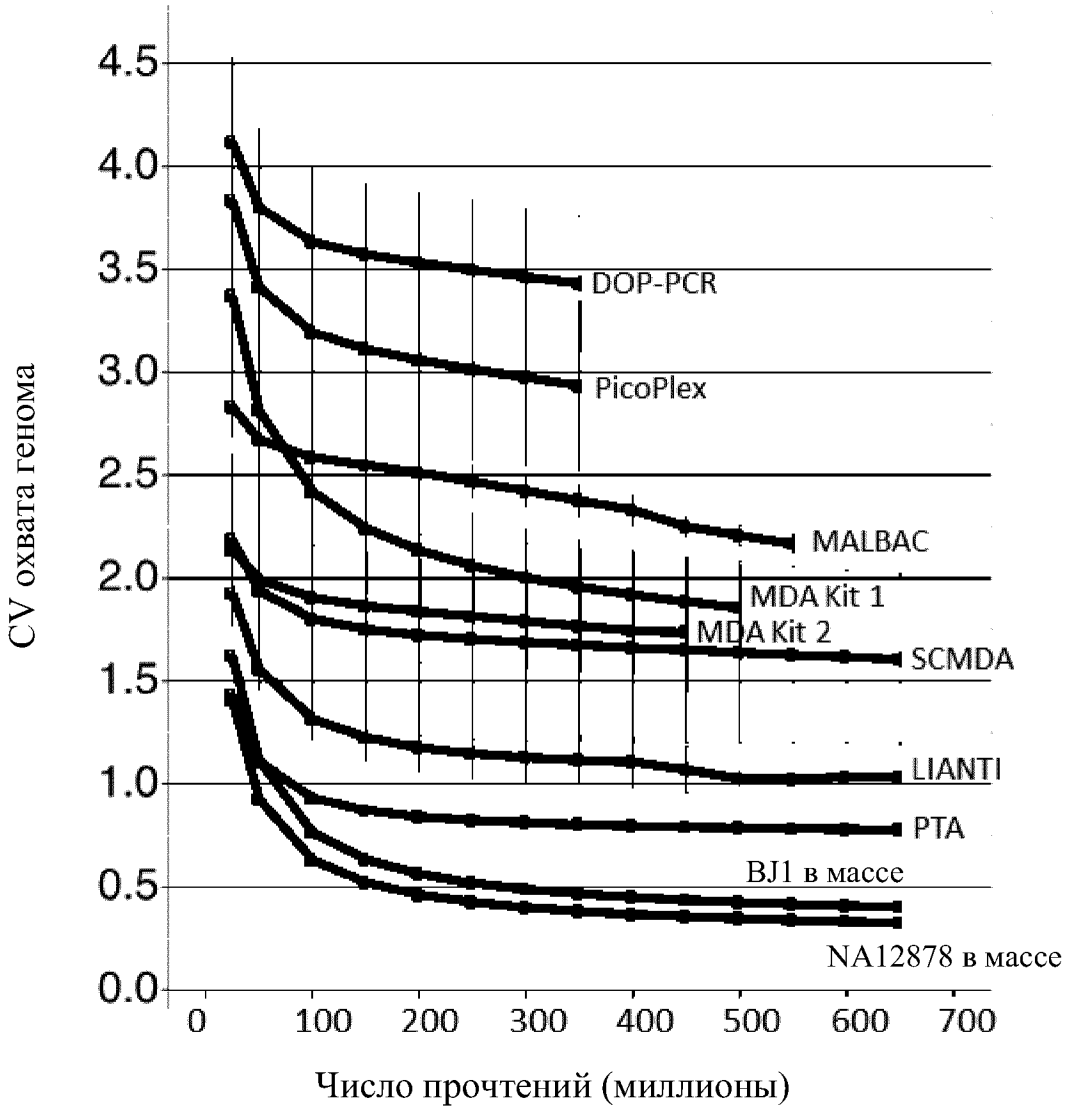

Охват PTA по хромосоме 1
Охват
4
3
2
1
0
A
B
C
50 мил. 100 мил. 150 мил. 200 мил.
Положение на хромосоме 1 (п.н.)
Охват PTA по хромосоме 1
Охват
4
3
2
1
0
A
B
C
50 мил. 100 мил. 150 мил. 200 мил.
Положение на хромосоме 1 (п.н.)


PTA
Охват
Положение в геноме (процент)
MDA
Охват
Положение в геноме (процент)
PTA
Частота аллеля
в массе
PTA
MDA
Частота аллеля
в массе
MDA

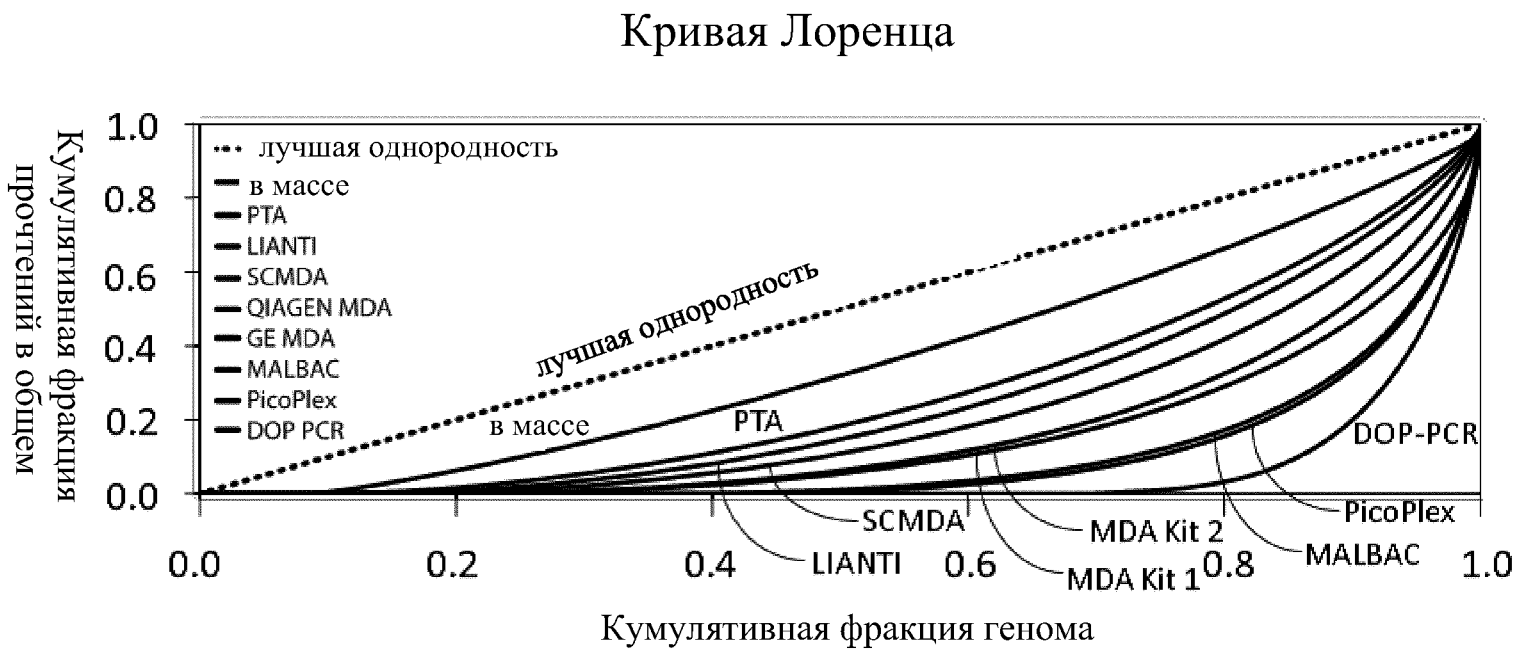
Фиг.4B


1.0
NA12878 в массе
0.9
PTA
BJ1 в массе
0.8
MDA Kit 2
0.7
LIANTI
SCMDA
MDA Kit 1
0.6
MALBAC
0.5
0.4
PicoPlex
0.3
0.2
DOP-PCR
0.1
0.0
0
200
400
600
Охваченная фракция генома
Число прочтений (миллионы)

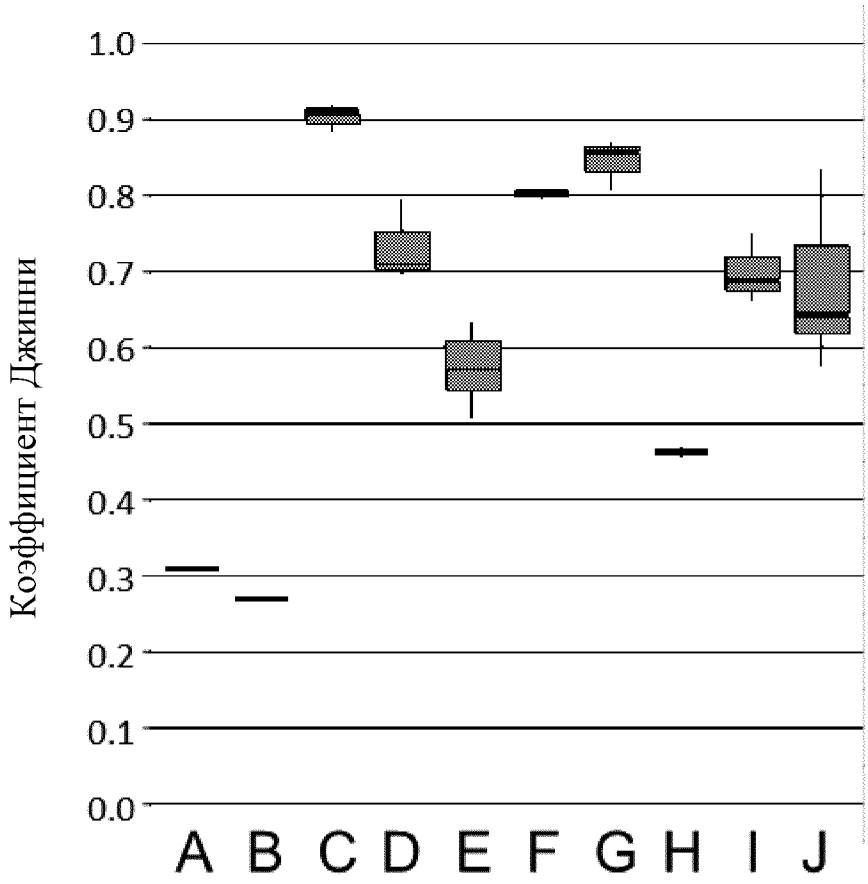
Фиг.5А


CV охвата генома
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
100
200
300
400
500
600
700
DOP-PCR
PicoPlex
MALBAC
MDA Kit 1
MDA Kit 2
SCMDA
LIANTI
PTA
BJ1 в массе
NA12878 в массе
Число прочтений (миллионы)

Фиг.5B

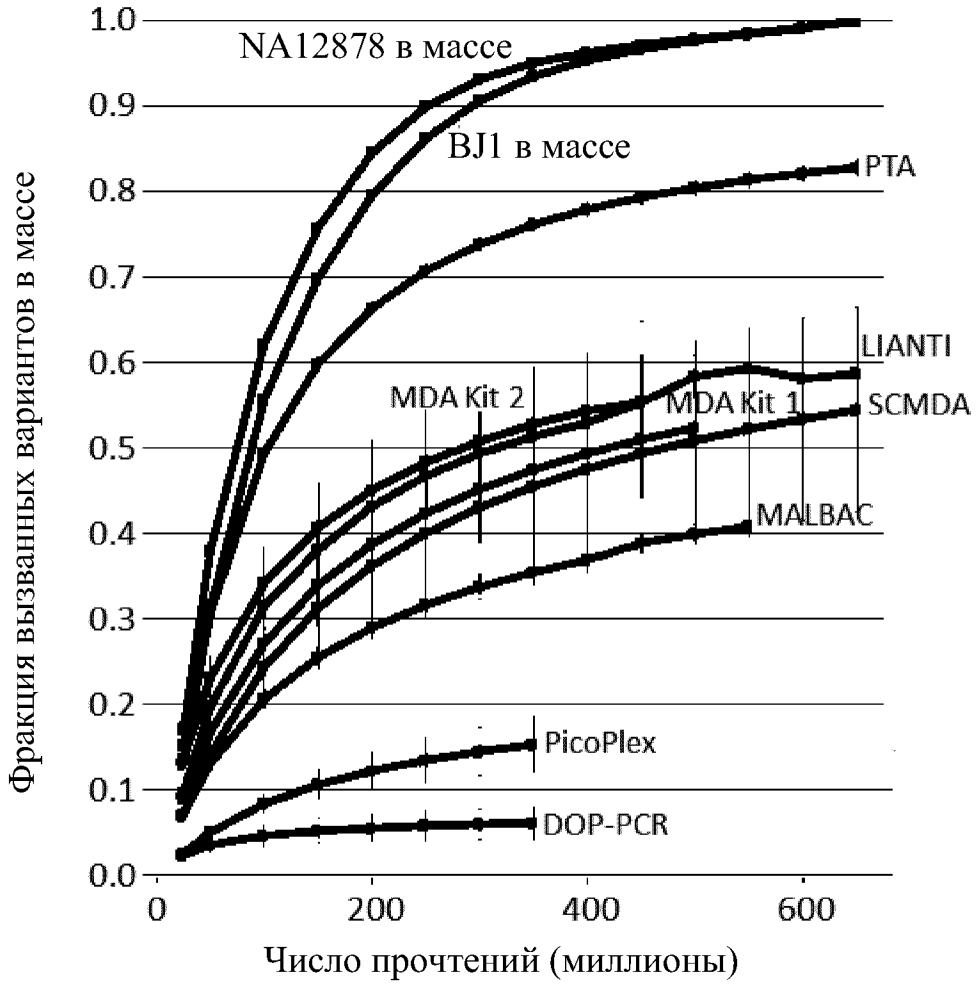

Кривая Лоренца
Кумулятивная фракция прочтений в общем
1.0
0.8
0.6
0.4
0.2
0.0
лучшая однородность
в массе
PTA
LIANTI
SCMDA
QIAGEN MDA
GE MDA
MALBAC
PicoPlex
DOP PCR
лучшая однородность
в массе
PTA
DOP-PCR
SCMDA
MDA Kit 2
PicoPlex
LIANTI
MDA Kit 1
MALBAC
0.0
0.2
0.4
0.6
0.8
1.0
Кумулятивная фракция генома

Фиг.5С

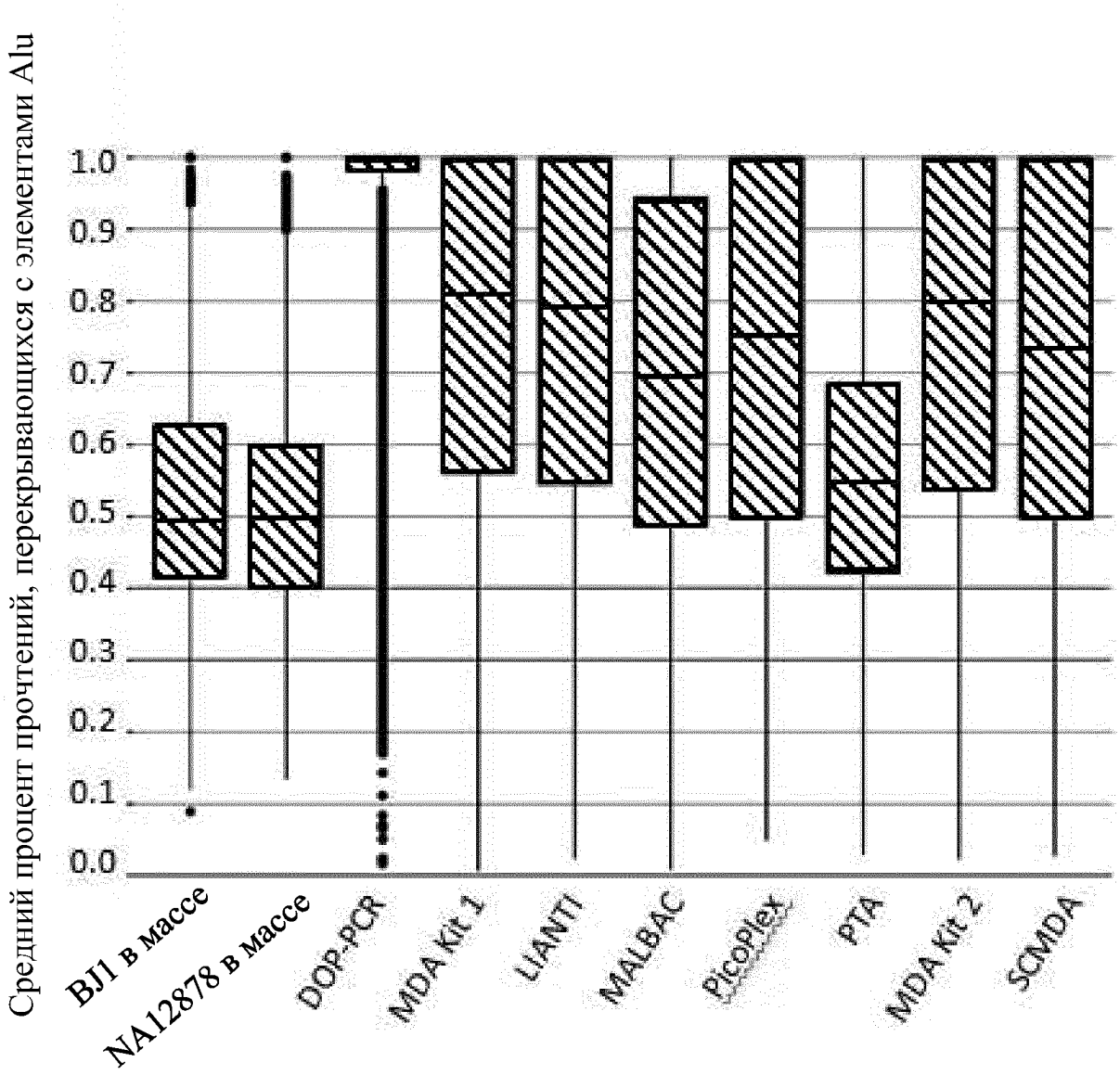


A. BJ1 в массе
B. NA12878 в массе
C. DOP-PCR
D. MDA Kit 1
E. LIANTI
F. MALBEC
G. PicoPLex
H. PTA
I. MDA Kit 2
J. SCMDA

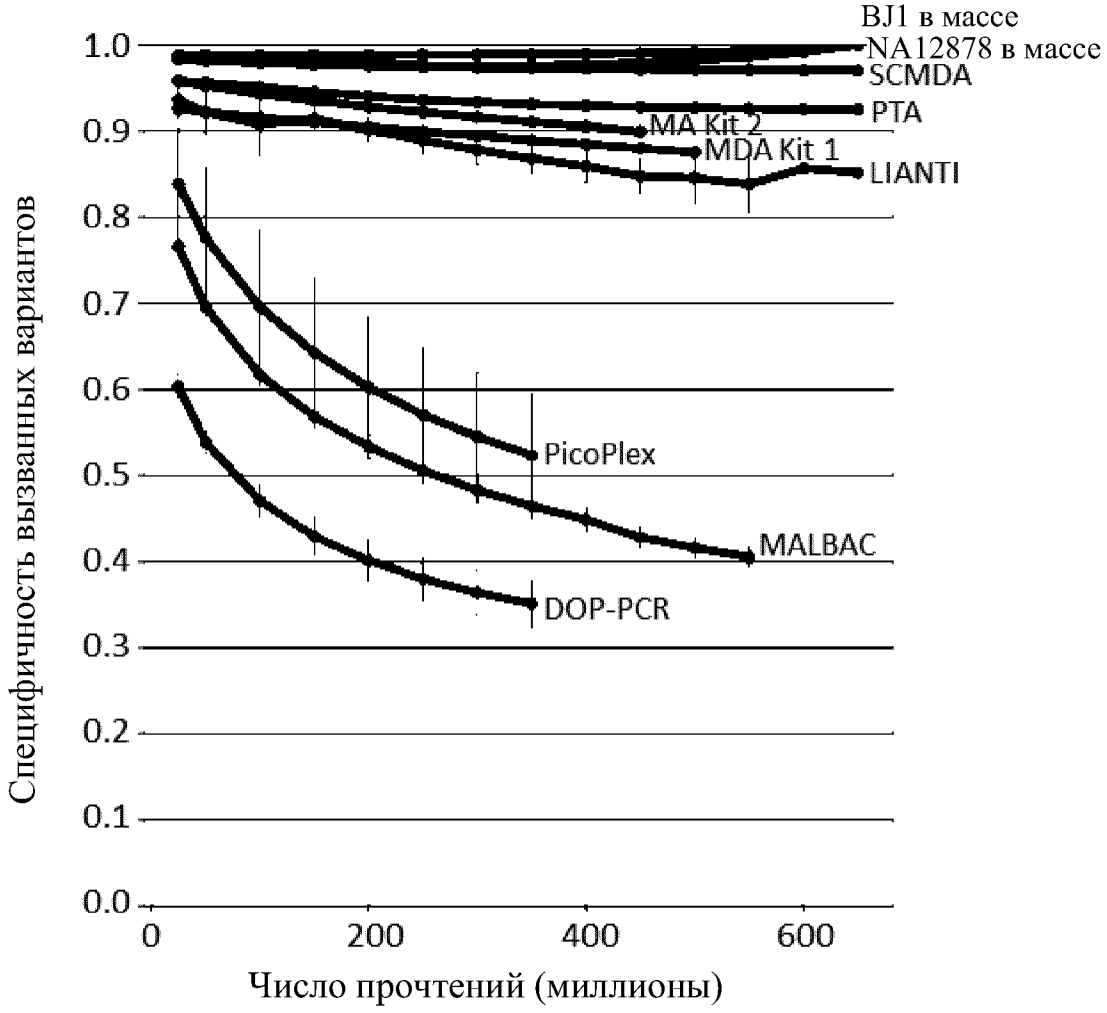
Фиг.5D


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
200
400
600
Фракция вызванных вариантов в массе
Число прочтений (миллионы)
NA12878 в массе
BJ1 в массе
PTA
LIANTI
MDA Kit 2
MDA Kit 1
SCMDA
MALBAC
PicoPlex
DOP-PCR

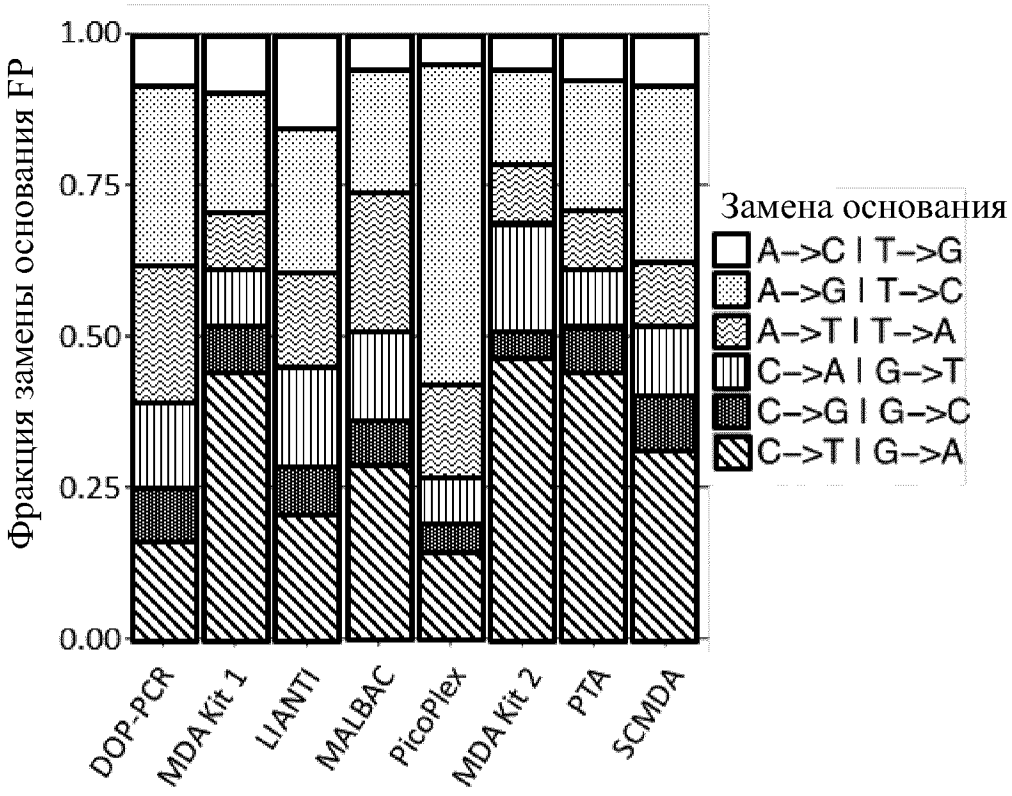
Фиг.5E


Средний процент прочтений, перекрывающихся с элементами Alu
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
BJ1 в массе
NA12878 в массе
DOP-PCR
MDA Kit 1
LIANTI
MALBAC
PicoPlex
PTA
MDA Kit 2
SCMDA

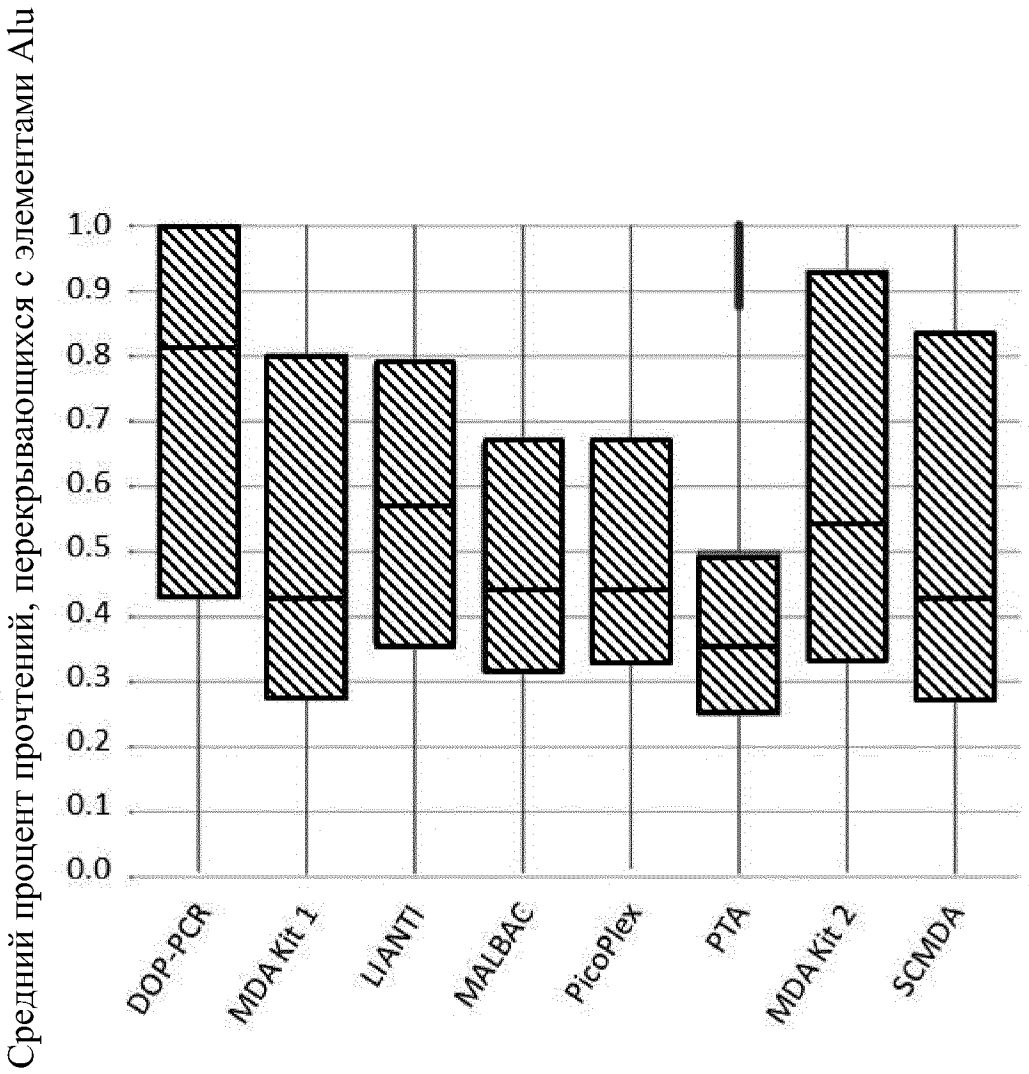
Фиг.5F


BJ1 в массе
NA12878 в массе
SCMDA
PTA
MA Kit 2
MDA Kit 1
LIANTI
PicoPlex
MALBAC
DOP-PCR
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
200
400
600
Специфичность вызванных вариантов
Число прочтений (миллионы)

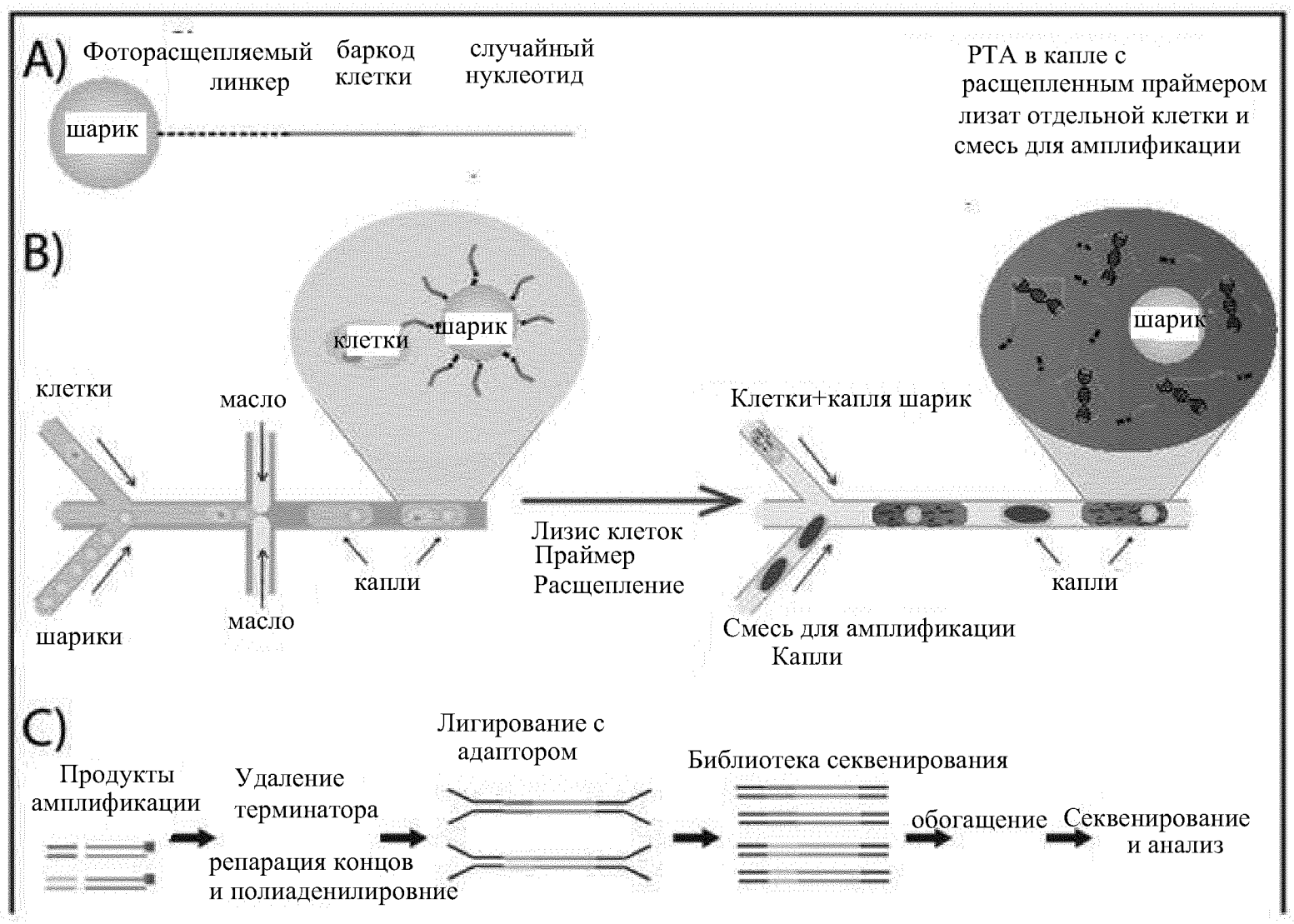
Фиг.5G


Фракция замены основания FP
1.00
0.75
0.50
0.25
0.00
DOP-PCR
MDA Kit 1
LIANTI
MALBAC
PicoPlex
MDA Kit 2
PTA
SCMDA
Замена основания
A->C | T->G
A->G | T->C
A->T | T->A
C->A | G->T
C->G | G->C
C->T | G->A

Фиг.5H

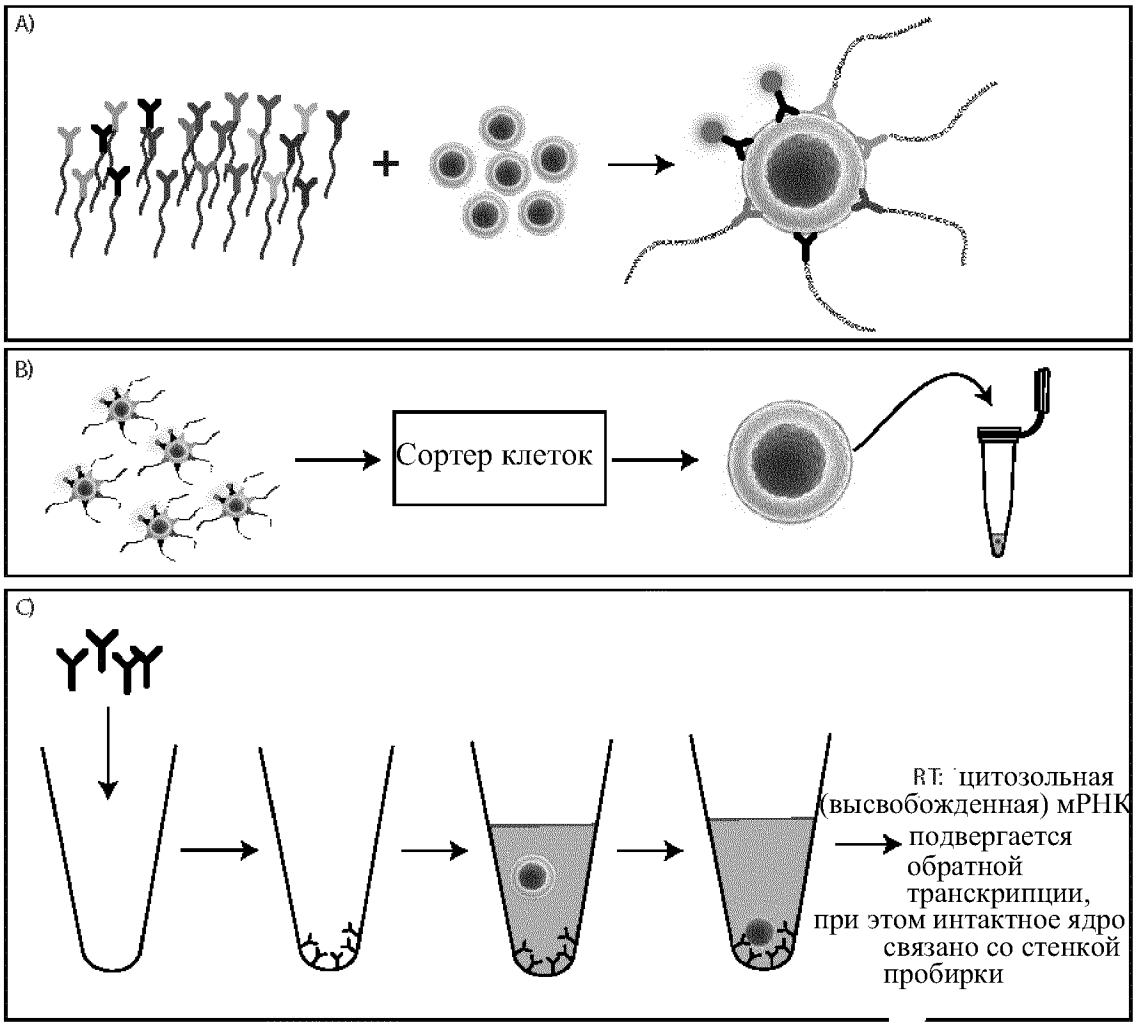

Средний процент прочтений, перекрывающихся с элементами Alu
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
DOP-PCR
MDA Kit 1
LIANTI
MALBAC
PicoPlex
PTA
MDA Kit 2
SCMDA

Фиг.5I

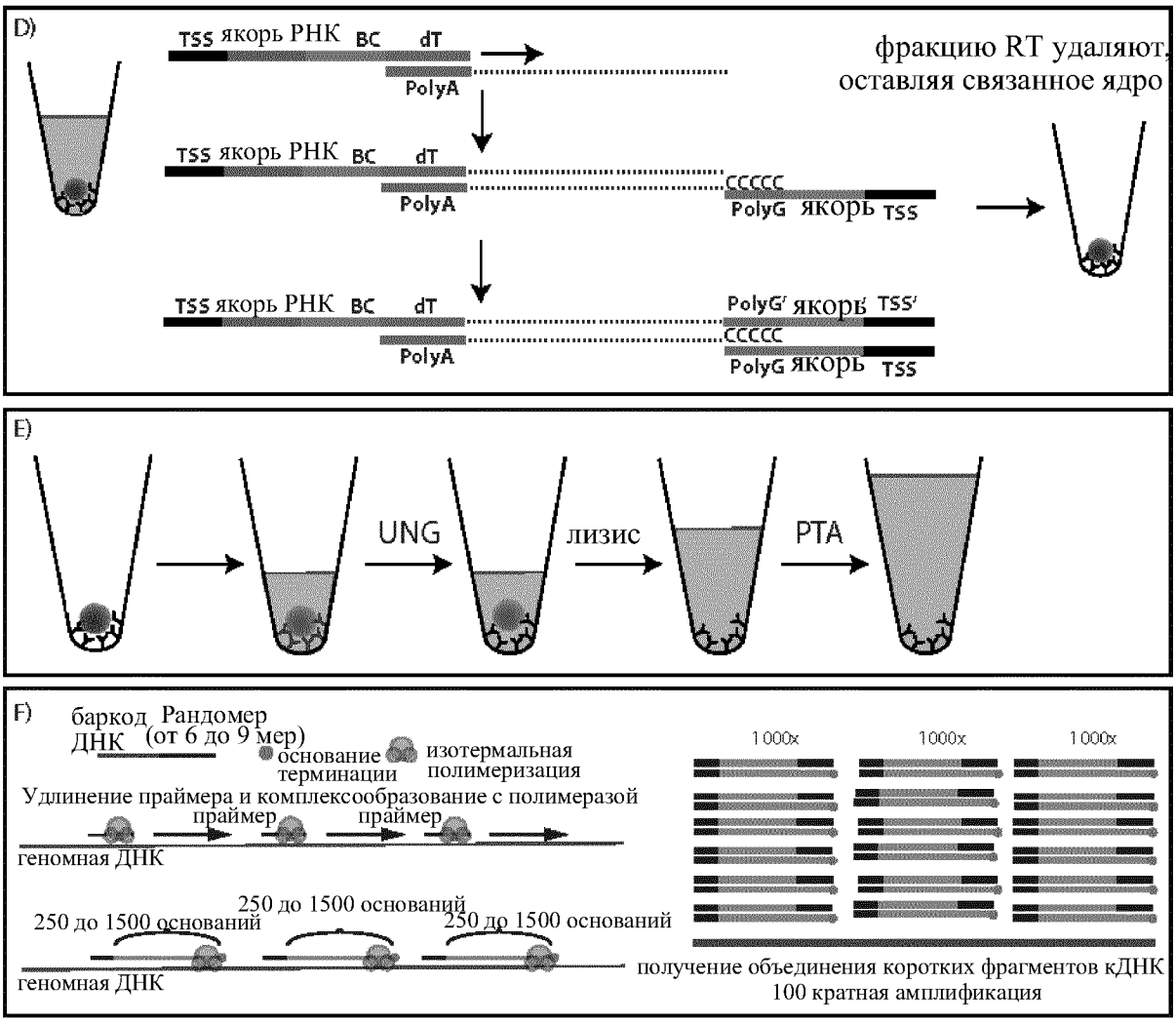

A)
Фоторасщепляемый линкер
баркод клетки
случайный нуклеотид
шарик
РТА в капле с расщепленным праймером
лизат отдельной клетки и смесь для амплификации
B)
клетки
шарик
клетки
масло
капли
шарики
масло
Лизис клеток
Праймер
Расщепление
Клетки+капля шарик
шарик
капли
Смесь для амплификации
Капли
C)
Продукты амплификации
Удаление терминатора
репарация концов и полиаденилировние
Лигирование с адаптором
Библиотека секвенирования
обогащение
Секвенирование и анализ

ФИГ. 6

Способ мультиомического анализа отдельных клеток



ФИГ. 7А


D)
TSS якорь РНК BC dT
PolyA
фракцию RT удаляют,
оставляя связанное ядро
CCCCC
PolyG якорь TSS
PolyG' якорь TSS'
E)
UNG
лизис
PTA
F)
баркод ДНК
Рандомер (от 6 до 9 мер)
основание терминации
изотермальная полимеризация
Удлинение праймера и комплексообразование с полимеразой
праймер
геномная ДНК
250 до 1500 оснований
1000x
получение объединения коротких фрагментов кДНК
100 кратная амплификация

ФИГ. 7В

RT и преамплификация

RT баркод олигодT праймер

TGTCGTTGTTGTCGGCCTTCTCTGGAAGCAGTGGTATCAACGCAGAGTACGGACGTTCAGAGNNNNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTVN

SS праймер

AAGCAGTGGTATCAACGCAGAGT

переключающий матрицу олигонуклеотид

/5BiosG/AAGCAGTGGTATCAACGCAGAGTACATrGrGrG

N701 Custom олиго

CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGG GTCTCGTGGGCTCGG

адаптор P7 праймер прочтения 2 адаптор Nextera

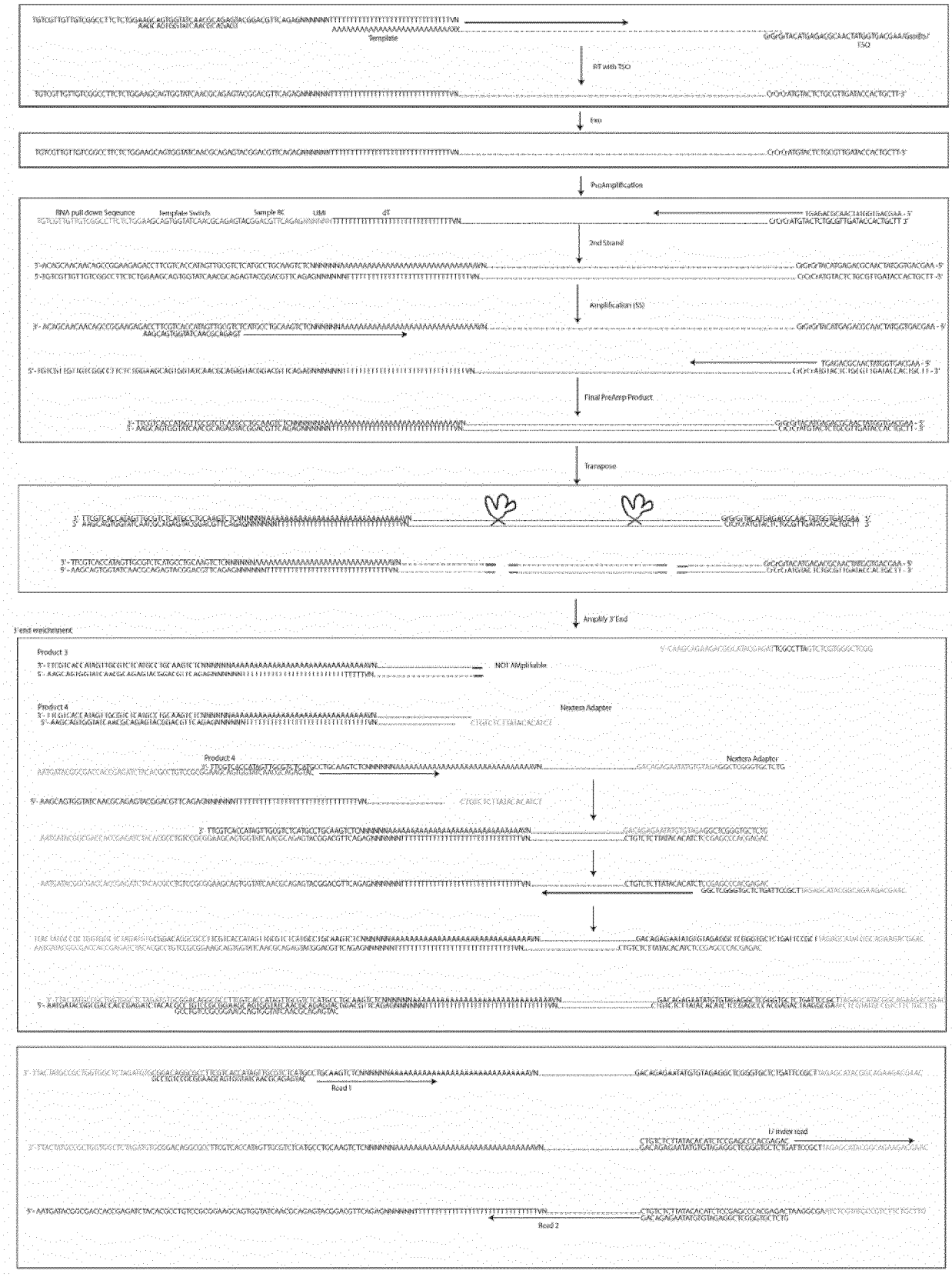
AATGATACGGCGACCACCGAGATCTACACGCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGTAC

адаптор P5 праймер прочтения 1 переключение матрицы

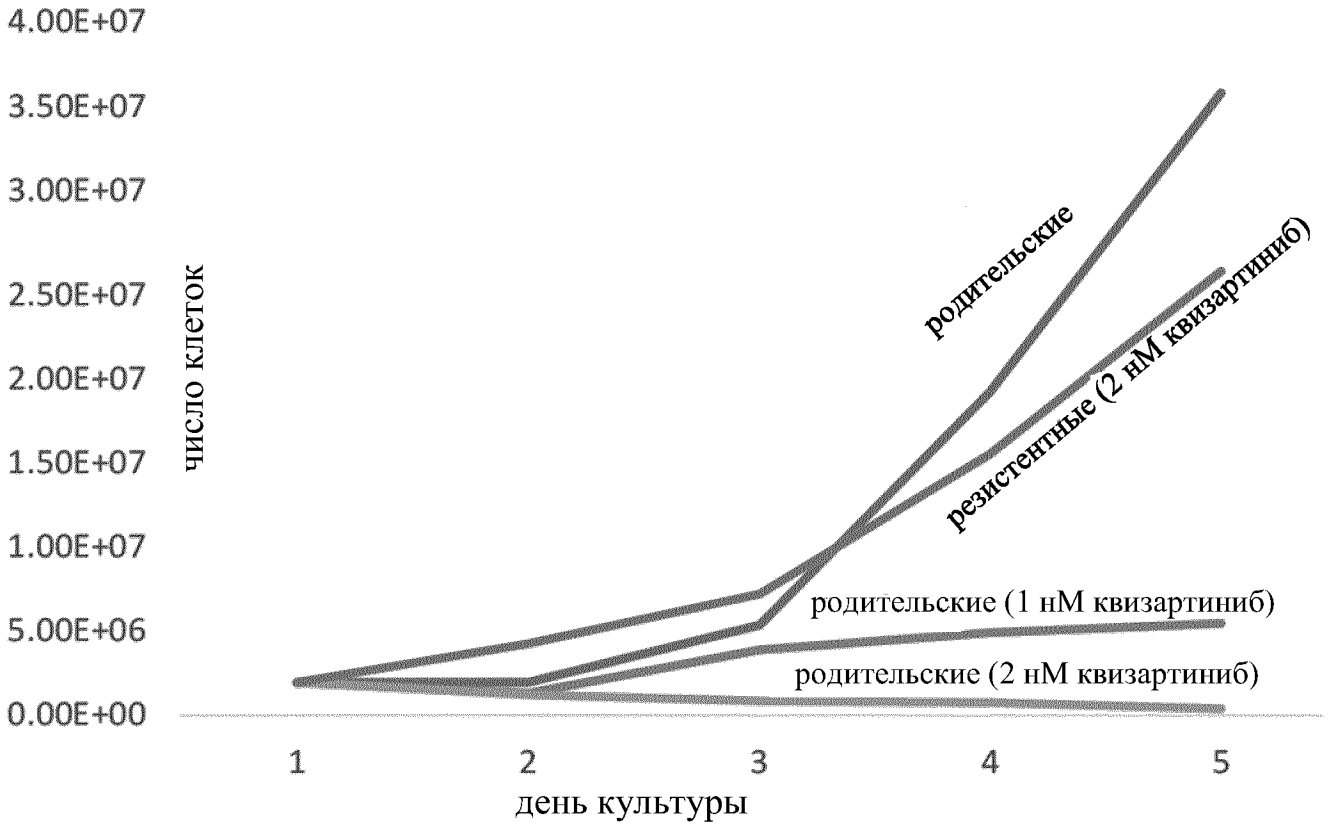
AATGATACGGCGACCACCGAGATCTACACGCCTGTCCGCGGAAGCAGTGGTATCAACGCAGAGTACGG

адаптор P5 праймер прочтения 1 переключение матрицы

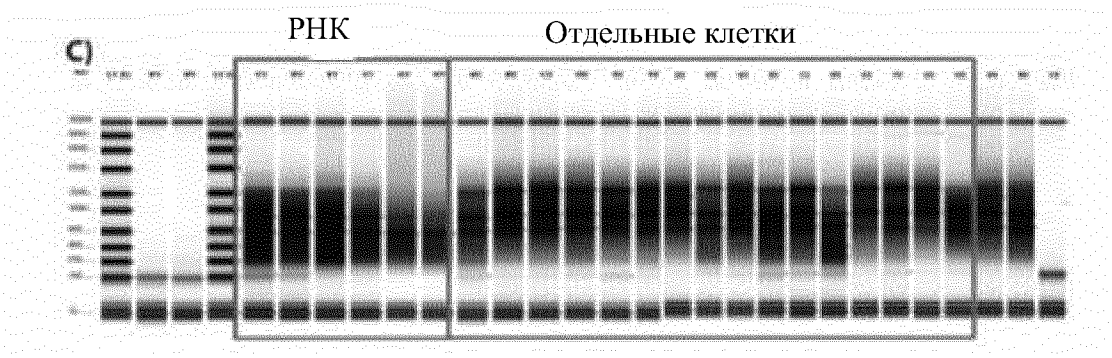
ФИГ. 8А


RT with TSO
Template
TSO
Exo
PreAmplification
RNA pull down Sequence
Template Switch
Sample BC
UMI
dT
2nd Strand
Amplification (ss)
Final PreAmp Product
Transpose
Amplify 3' End
3' end enrichment
Product 3
NOT Amplifiable
Product 4
Nextera Adapter
Read 1
i7 index read
Read 2

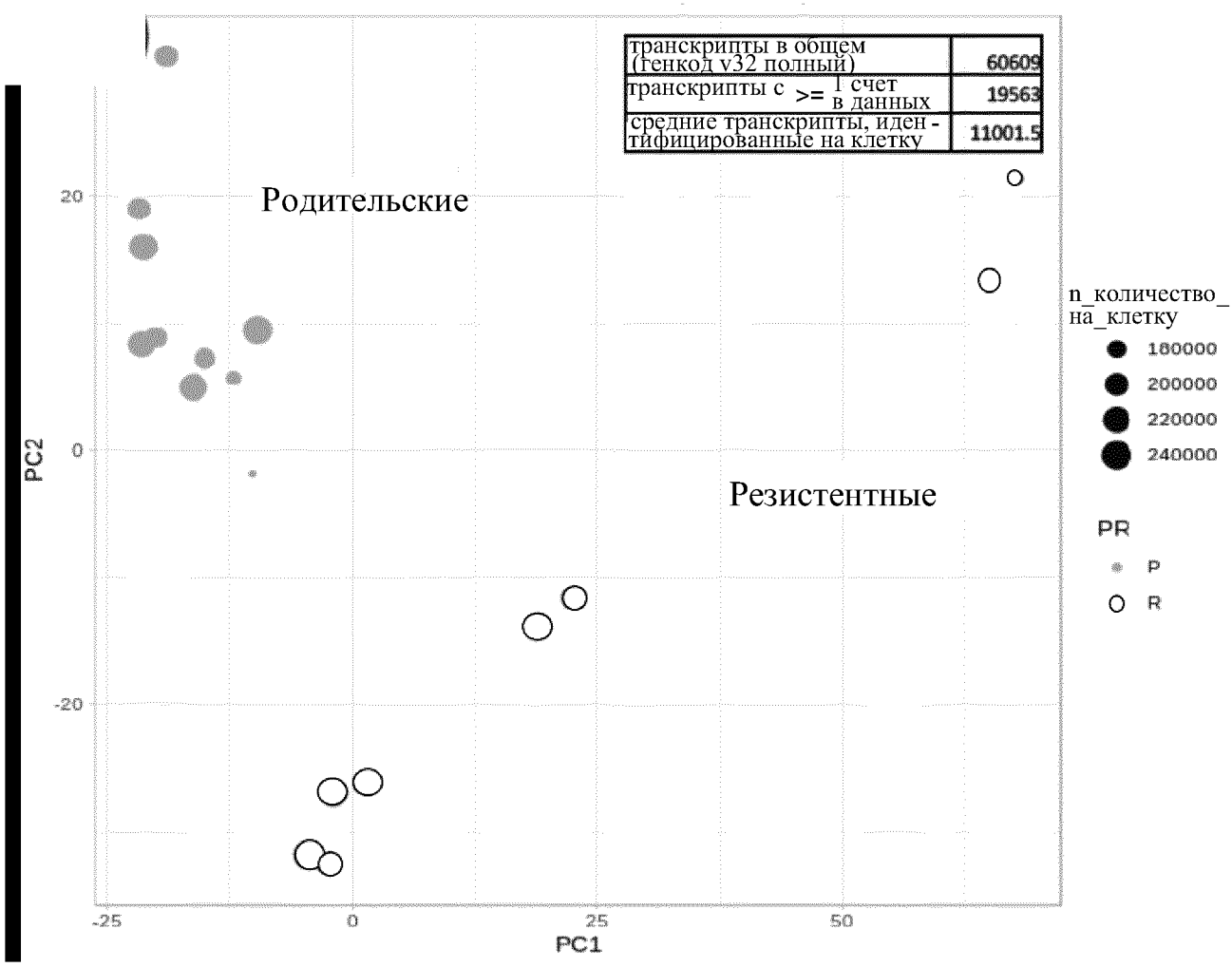
Фиг. 8B


4.00E+07
3.50E+07
3.00E+07
2.50E+07
2.00E+07
1.50E+07
1.00E+07
5.00E+06
0.00E+00
число клеток
родительские
резистентные (2 нМ квизартиниб)
родительские (1 нМ квизартиниб)
родительские (2 нМ квизартиниб)
1
2
3
4
5
день культуры

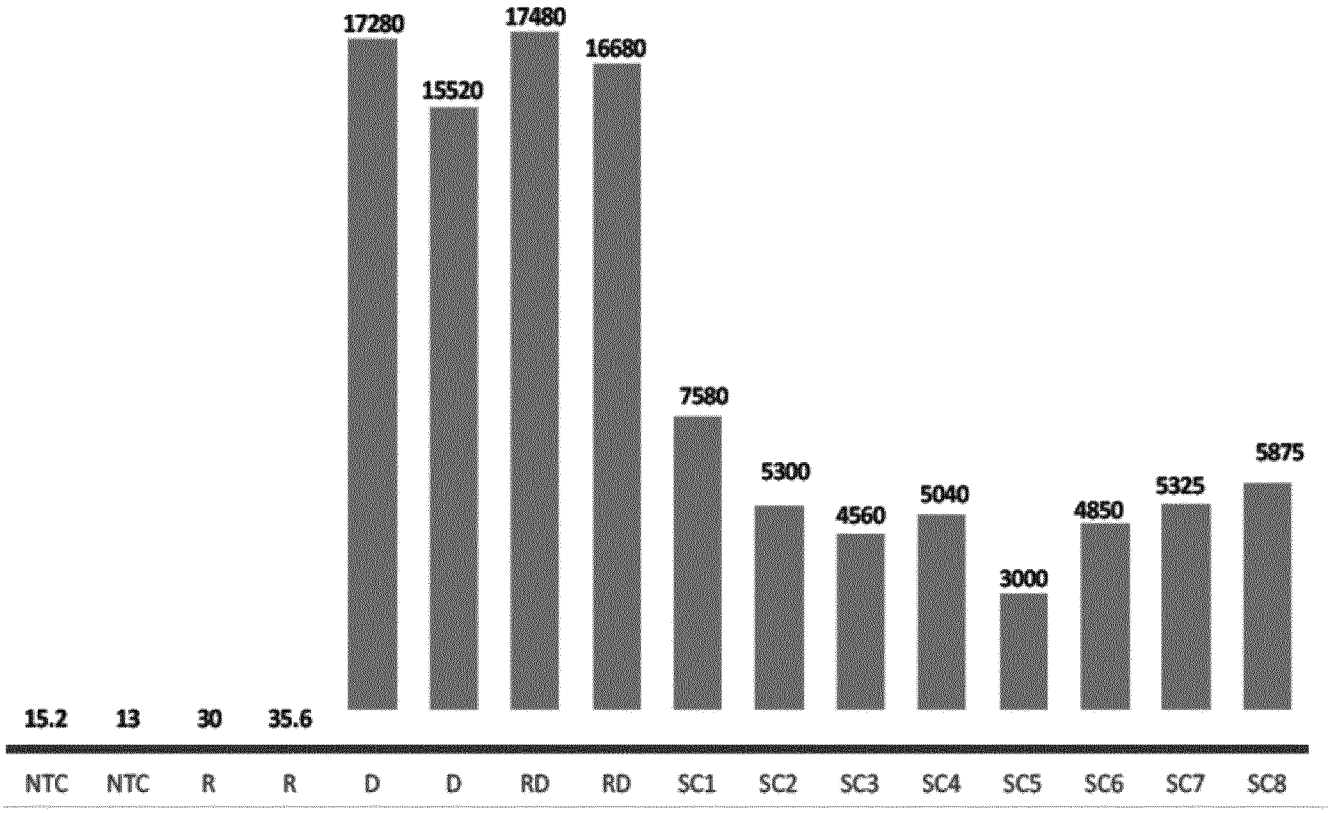
ФИГ. 9А


РНК
Отдельные клетки

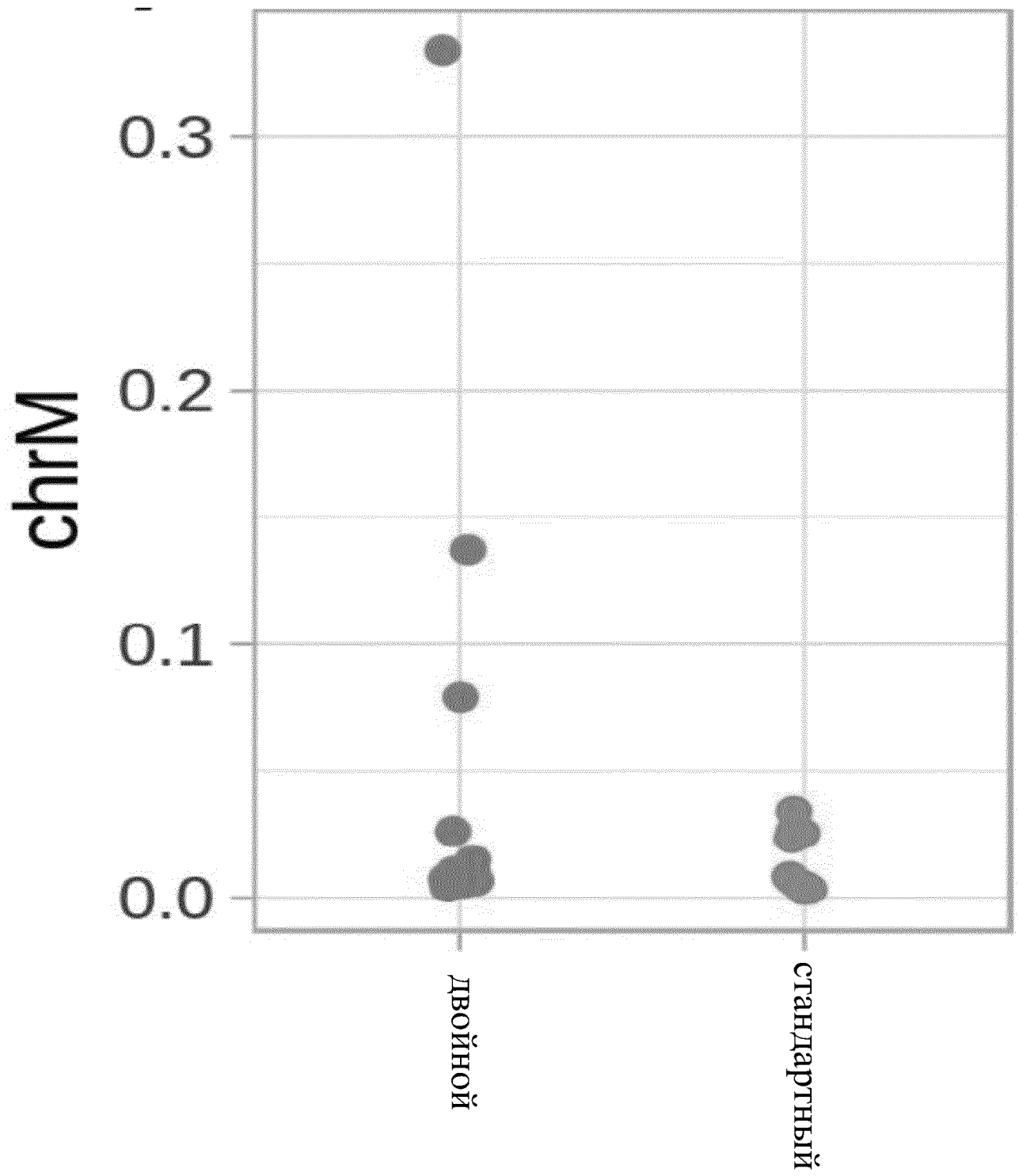
ФИГ. 9В


транскрипты в общем (генкод v32 полный) 60609
транскрипты с >= 1 счет в данных 19563
средние транскрипты, идентифицированные на клетку 11001.5
Родительские
Резистентные
PC2
PC1
20
0
-20
-25
0
25
50
n_количество_на_клетку
160000
200000
220000
240000
PR
P
R

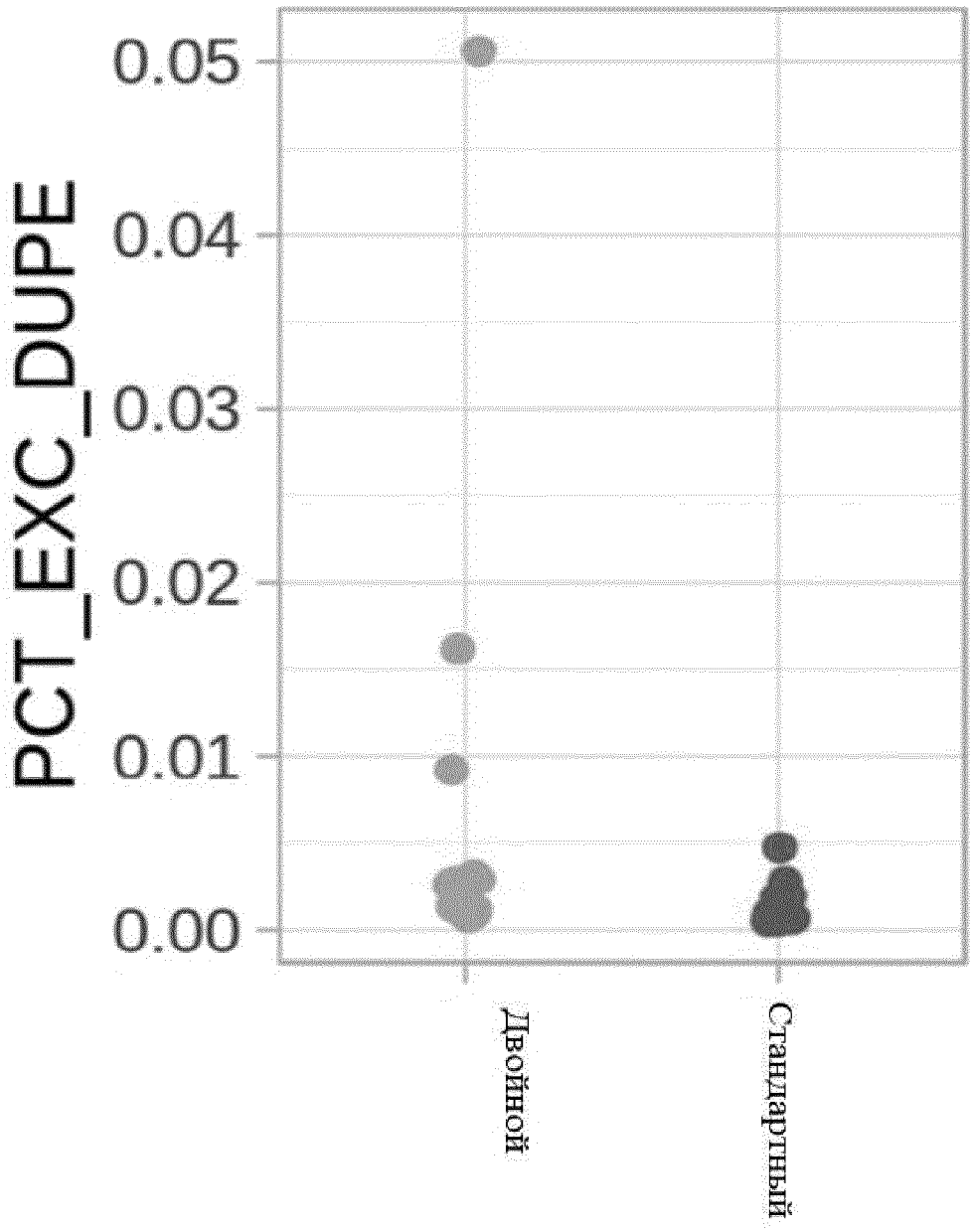
ФИГ. 9С


15.2
13
30
35.6
17280
15520
17480
16680
7580
5300
4560
5040
3000
4850
5325
5875
NTC
NTC
R
R
D
D
RD
RD
SC1
SC2
SC3
SC4
SC5
SC6
SC7
SC8

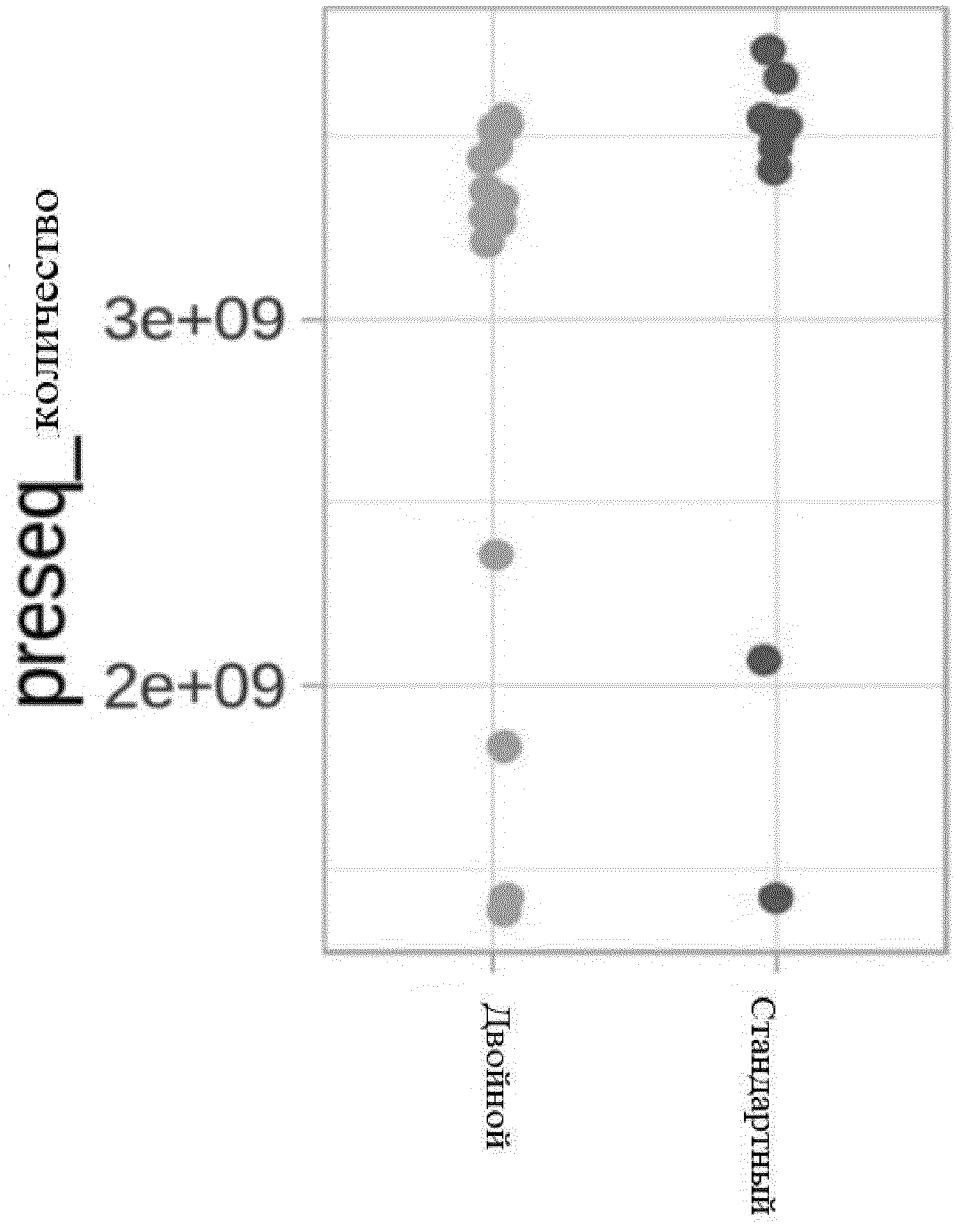
ФИГ. 9D



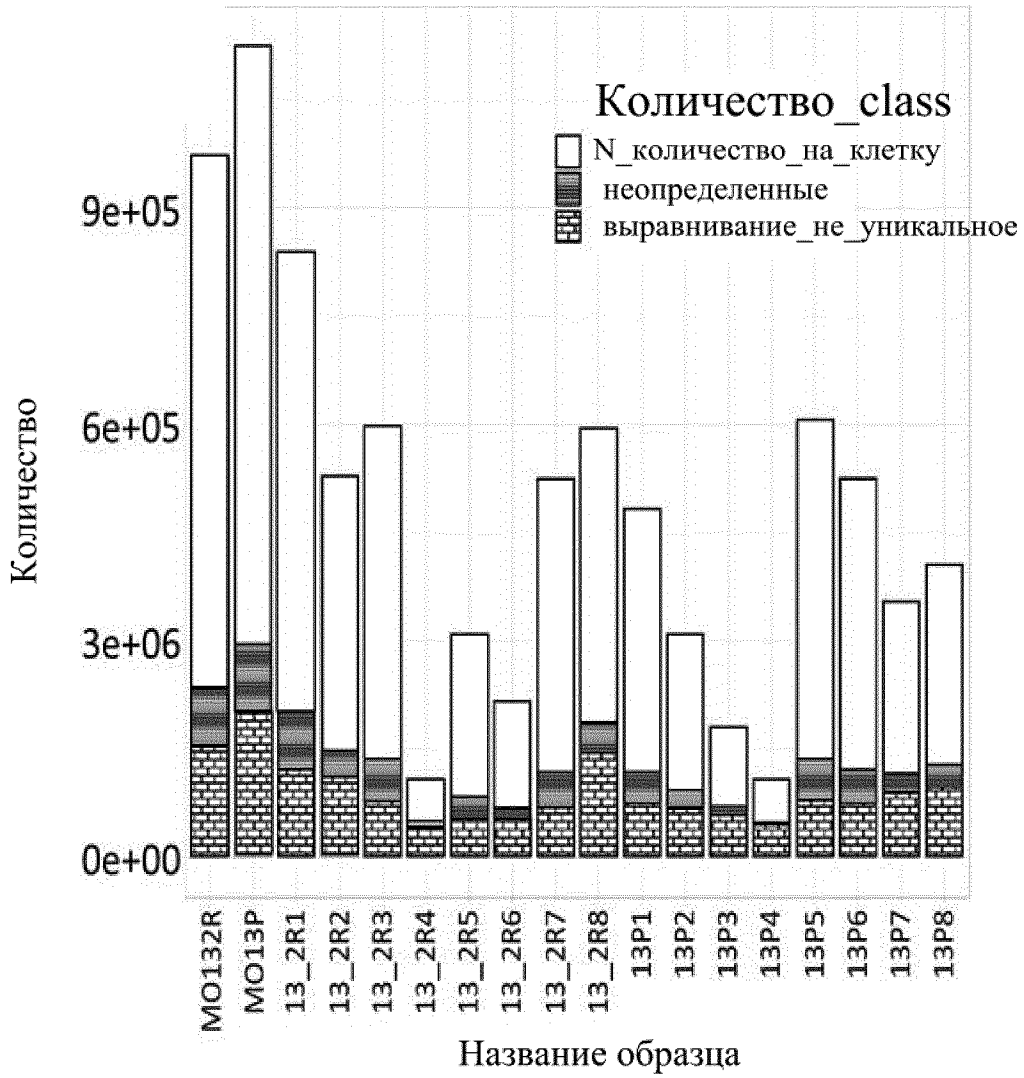
ФИГ. 10А



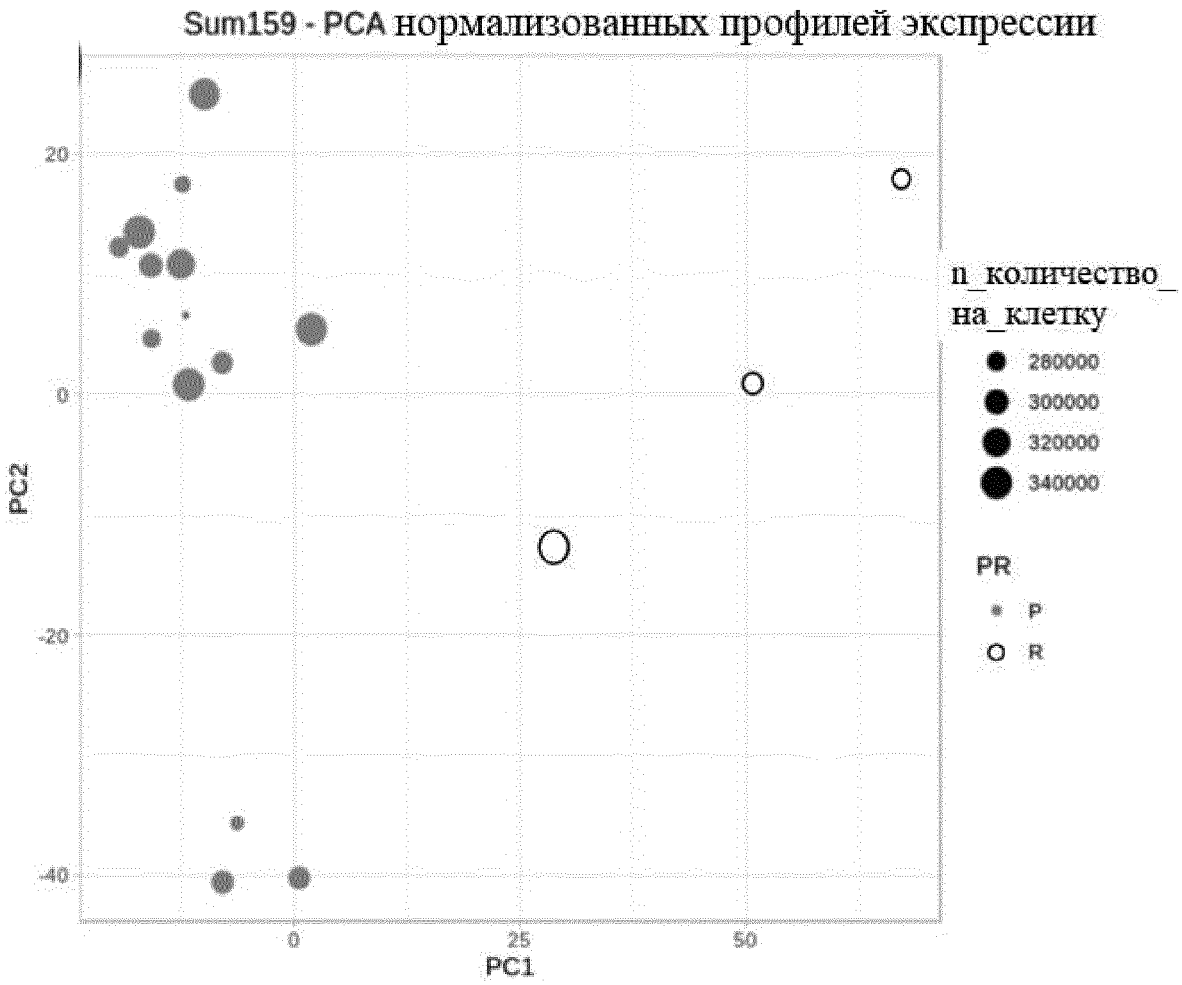
ФИГ. 10В



ФИГ. 10C

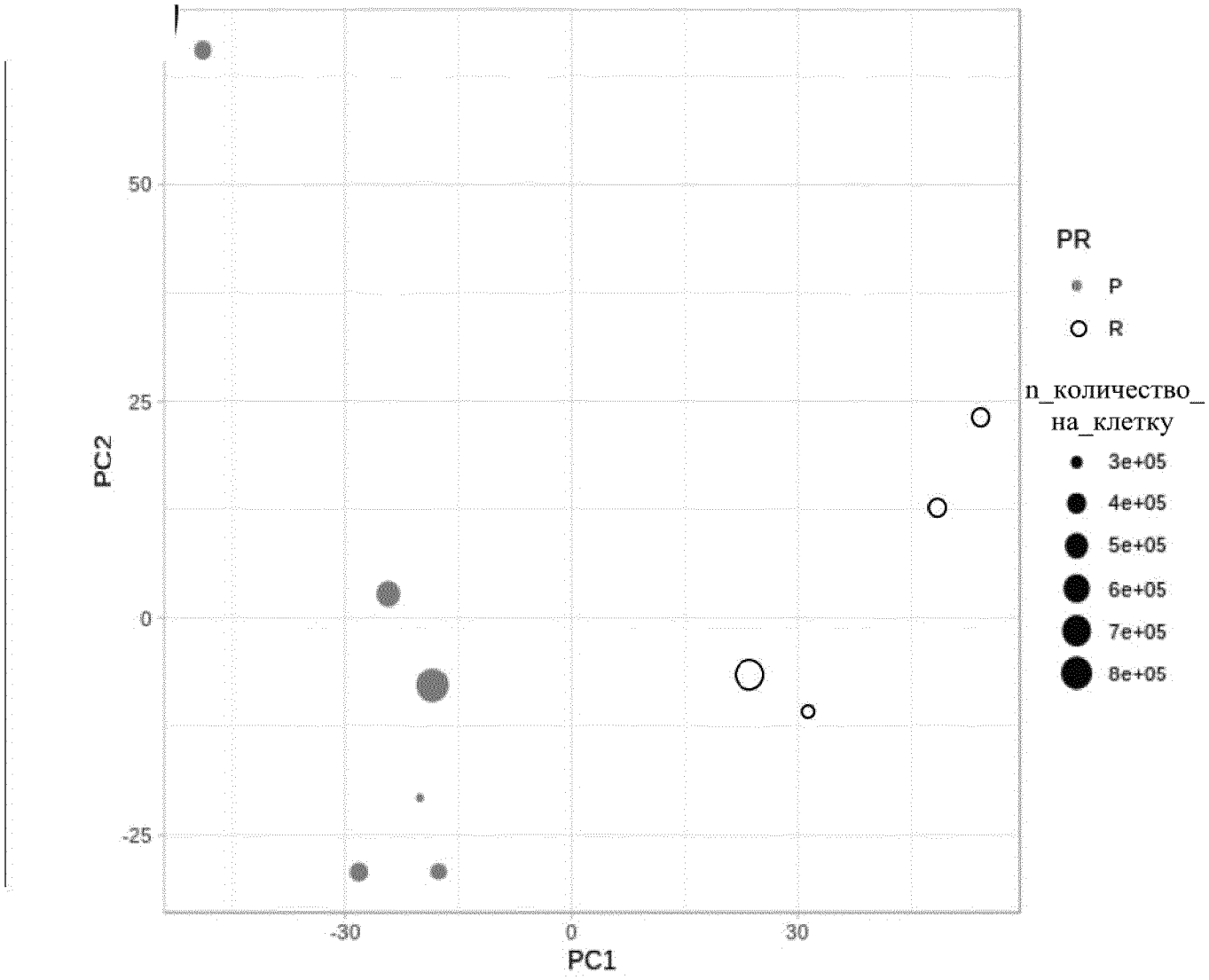

Количество_class
N_количество_на_клетку
неопределенные
выравнивание_не_уникальное
Количество
9e+05
6e+05
3e+06
0e+00
MO132R
MO13P
13_2R1
13_2R2
13_2R3
13_2R4
13_2R5
13_2R6
13_2R7
13_2R8
13P1
13P2
13P3
13P4
13P5
13P6
13P7
13P8
Название образца

ФИГ. 10D

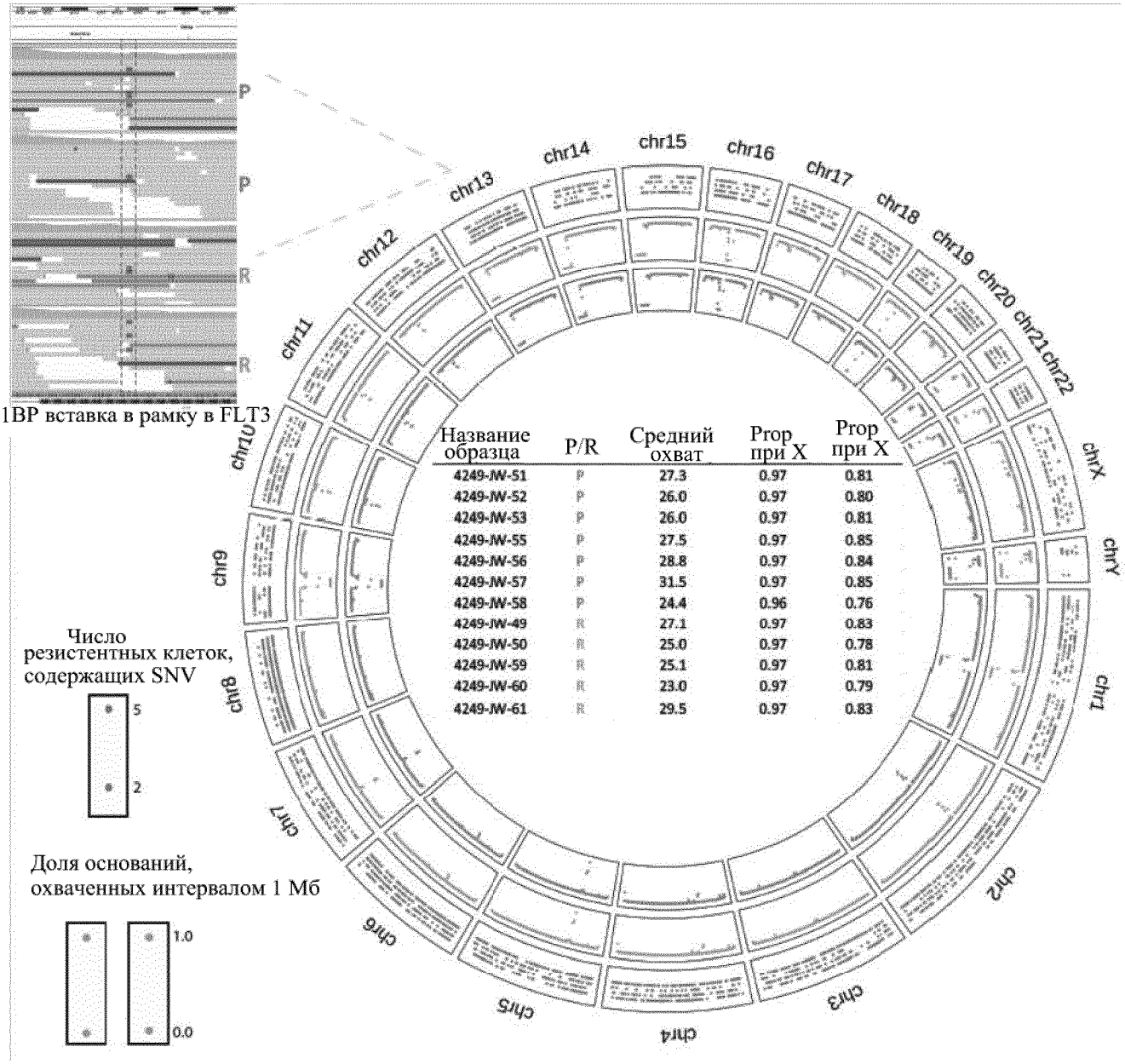

Sum159 - PCA нормализованных профилей экспрессии
20
0
-20
-40
PC2
0
25
50
PC1
n_количество_
на_клетку
280000
300000
320000
340000
PR
P
R

ФИГ. 10E

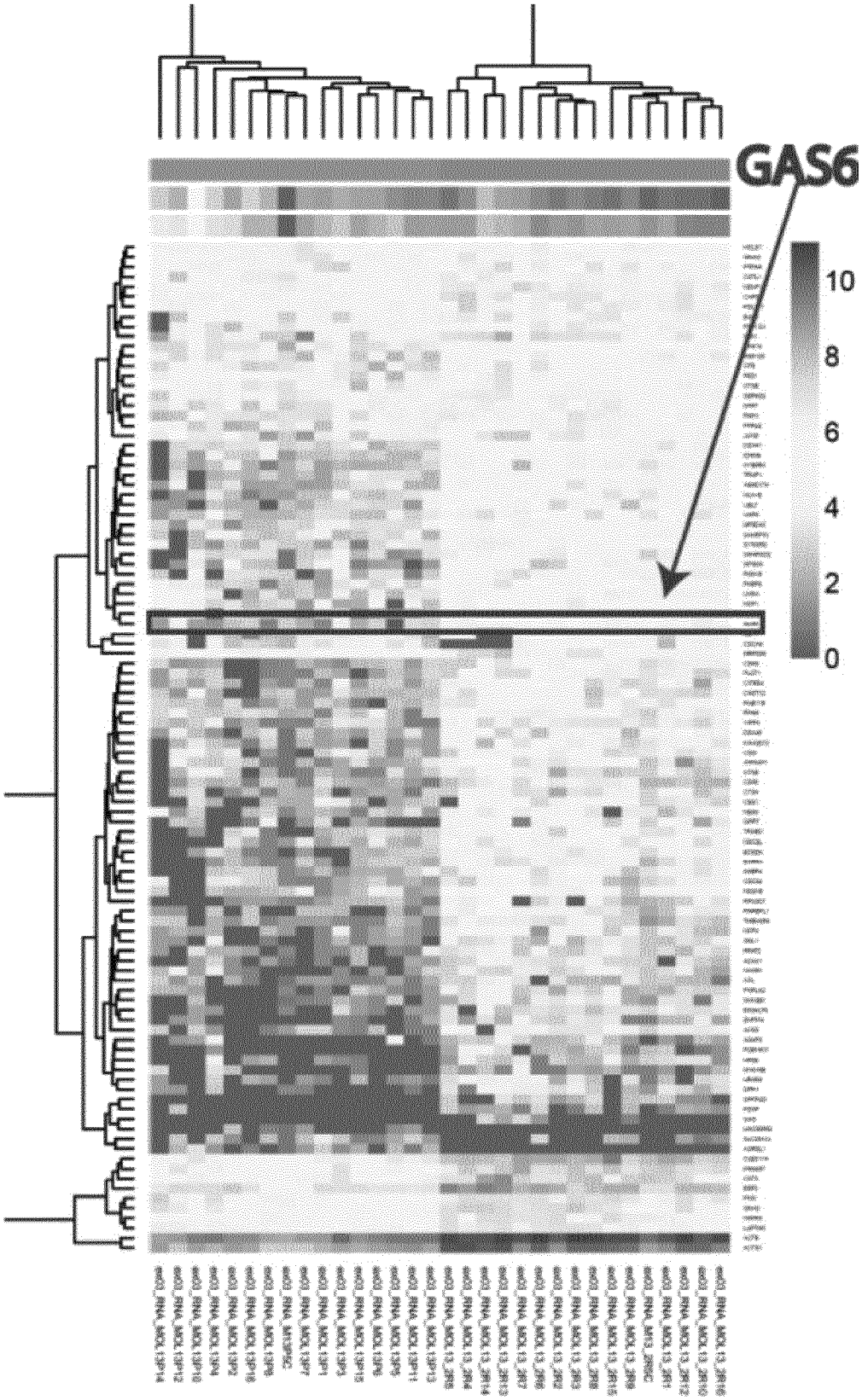
Двойной протокол Sum159 - PCA нормализованных профилей экспрессии



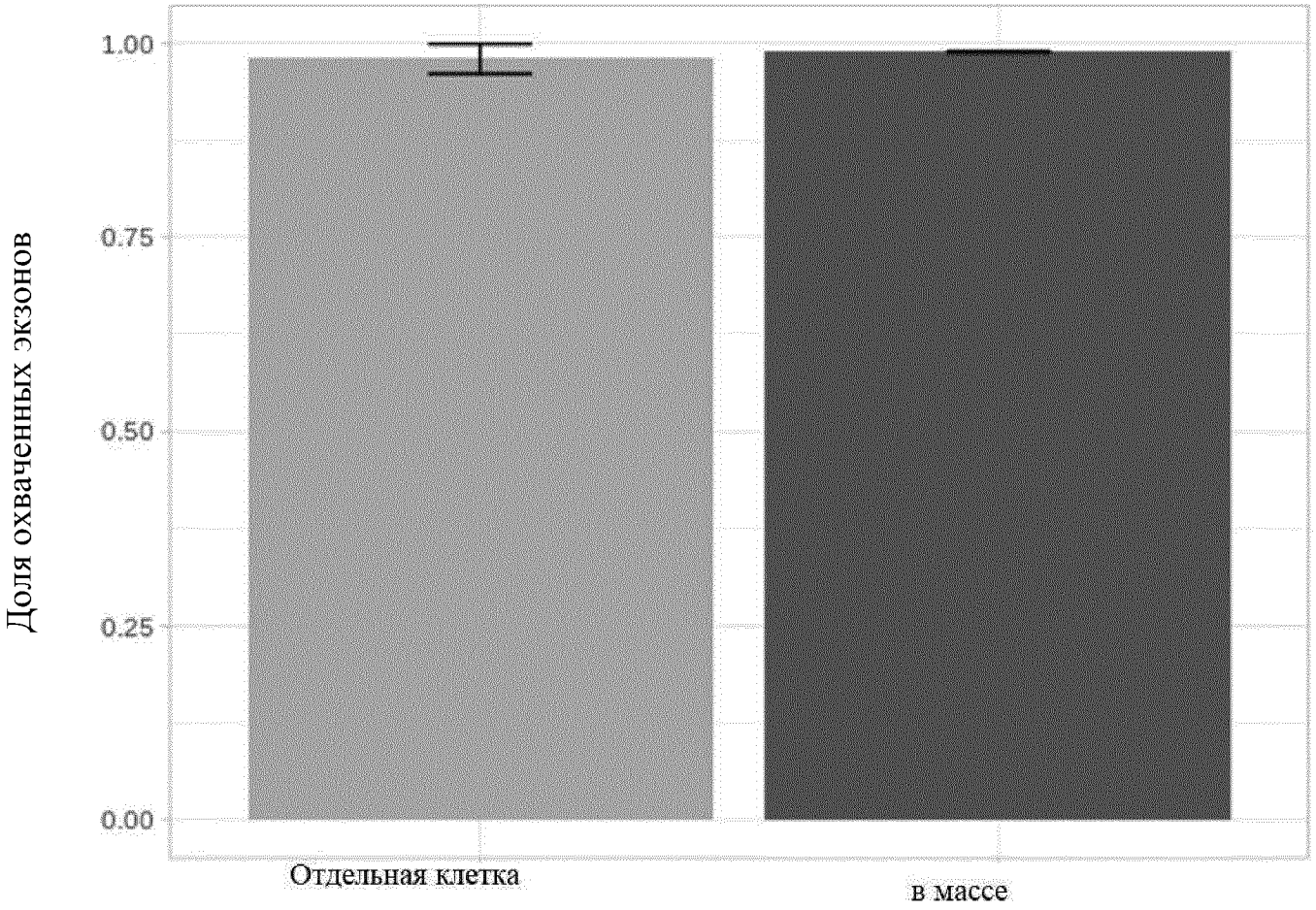
ФИГ. 10F


P
P
R
R
21BP вставка в рамку в FLT3
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
chrX
chrY
Название образца
P/R
Средний охват
Prop при X
Prop при X
4249-JW-51 P 27.3 0.97 0.81
4249-JW-52 P 26.0 0.97 0.80
4249-JW-53 P 26.0 0.97 0.81
4249-JW-55 P 27.5 0.97 0.85
4249-JW-56 P 28.8 0.97 0.84
4249-JW-57 P 31.5 0.97 0.85
4249-JW-58 P 24.4 0.96 0.76
4249-JW-49 R 27.1 0.97 0.83
4249-JW-50 R 25.0 0.97 0.78
4249-JW-59 R 25.1 0.97 0.81
4249-JW-60 R 23.0 0.97 0.79
4249-JW-61 R 29.5 0.97 0.83
Число резистентных клеток, содержащих SNV
5
2
Доля оснований, охваченных интервалом 1 Мб
1.0
0.0

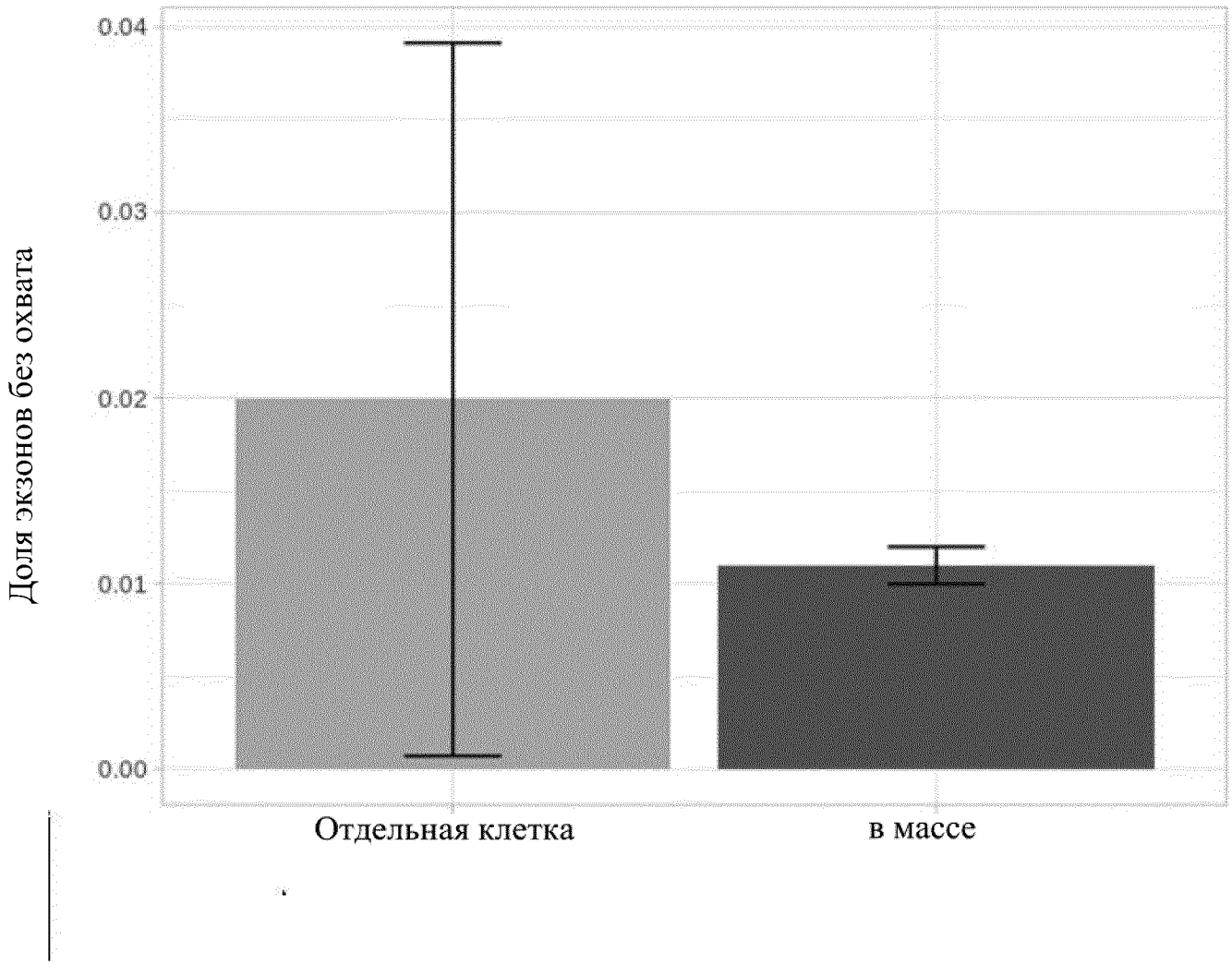
ФИГ. 11A



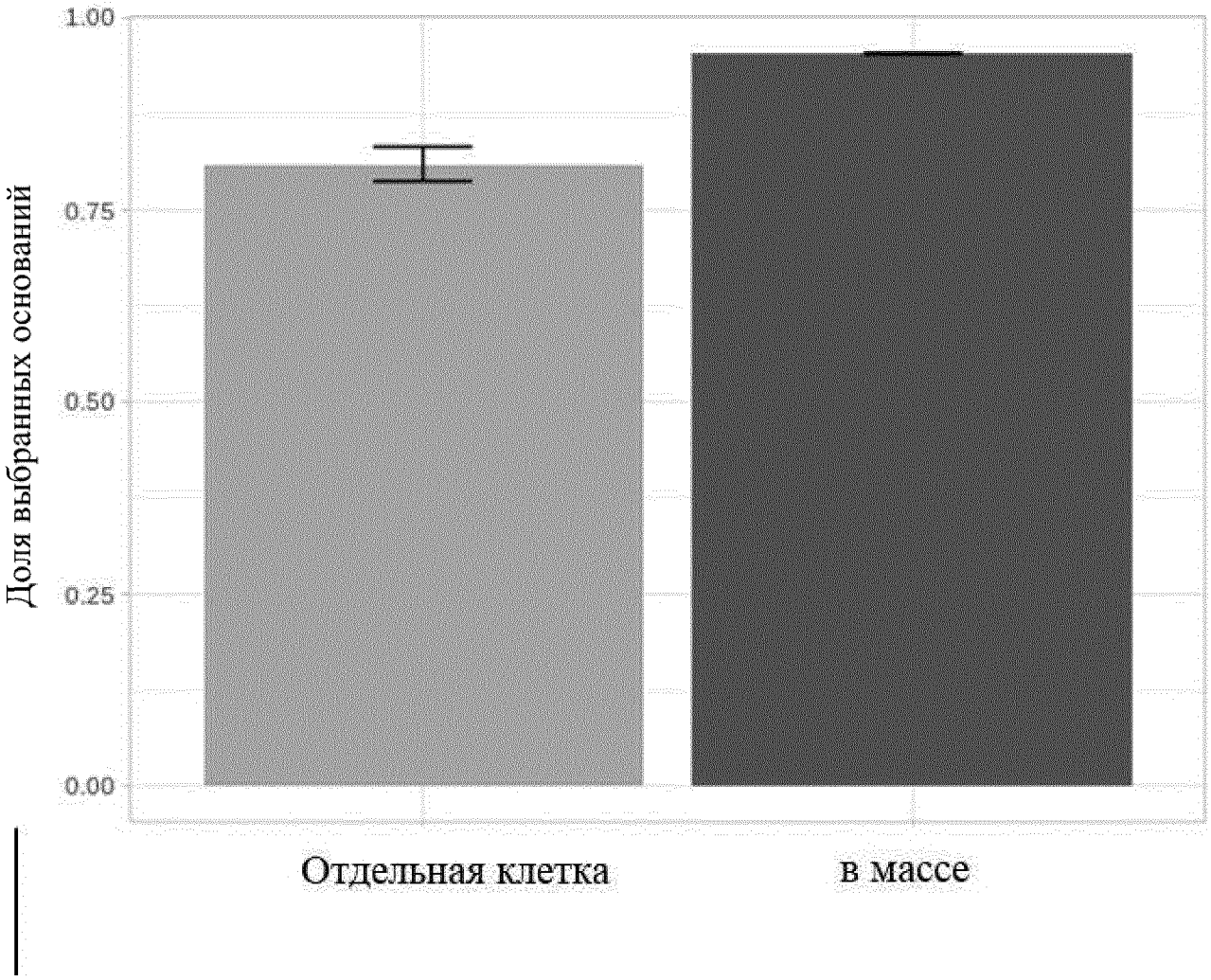
ФИГ. 11B


Доля охваченных экзонов
1.00
0.75
0.50
0.25
0.00
Отдельная клетка
в массе

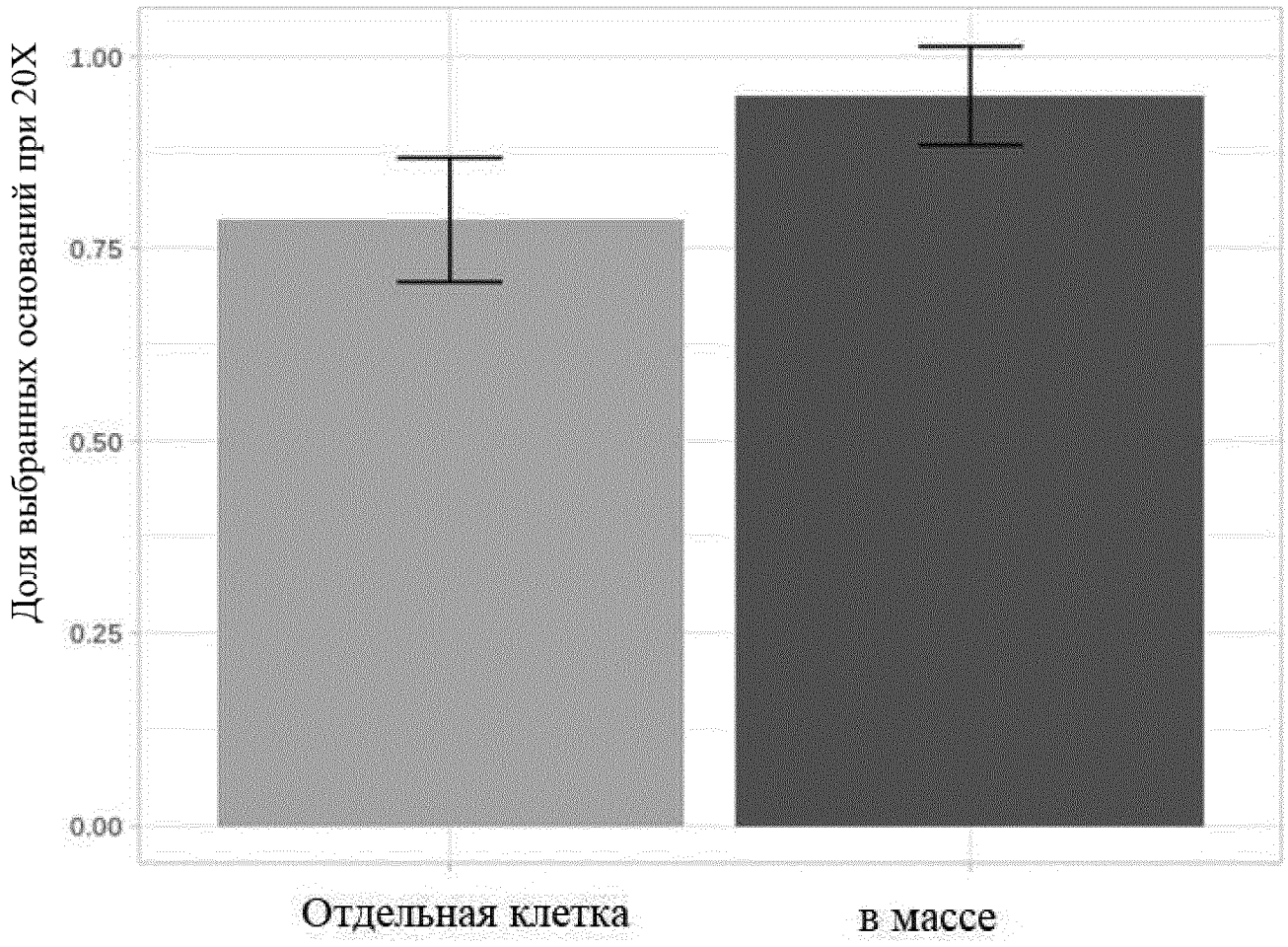
ФИГ. 12А


Доля экзонов без охвата
0.04
0.03
0.02
0.01
0.00
Отдельная клетка
в массе

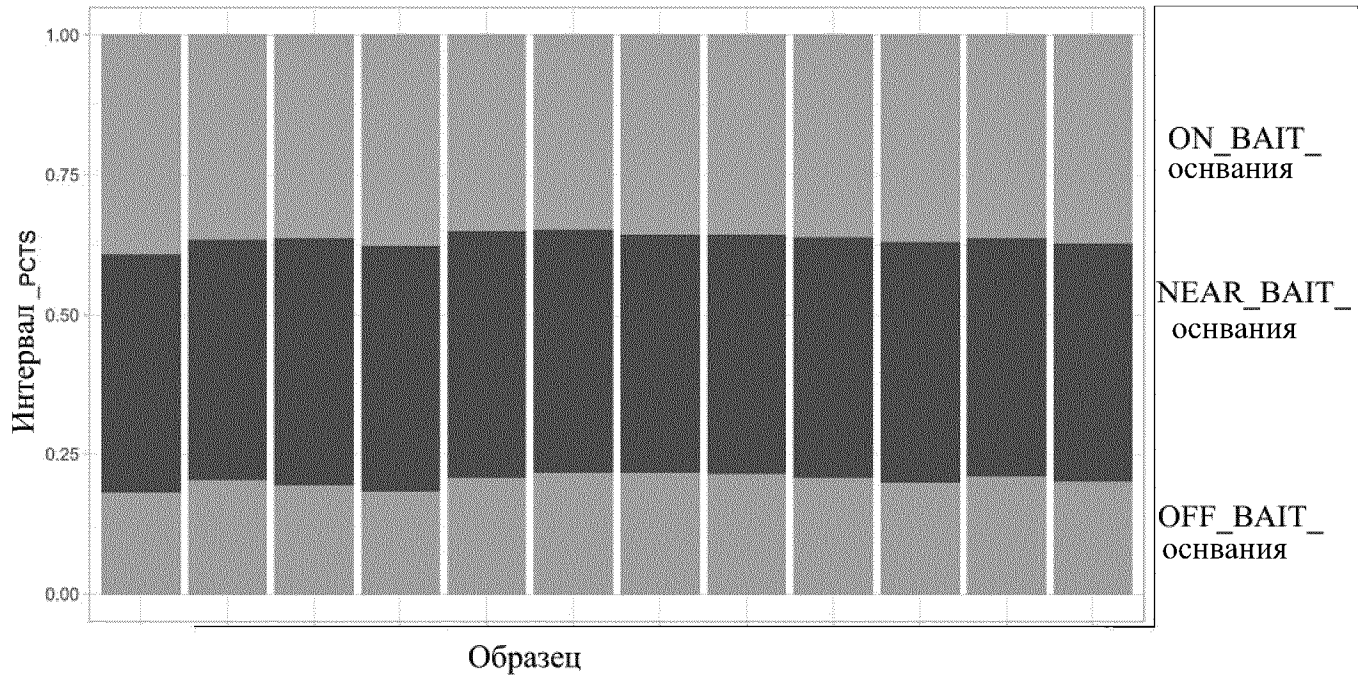
ФИГ. 12B


1.00
0.75
0.50
0.25
0.00
Доля выбранных оснований
Отдельная клетка
в массе

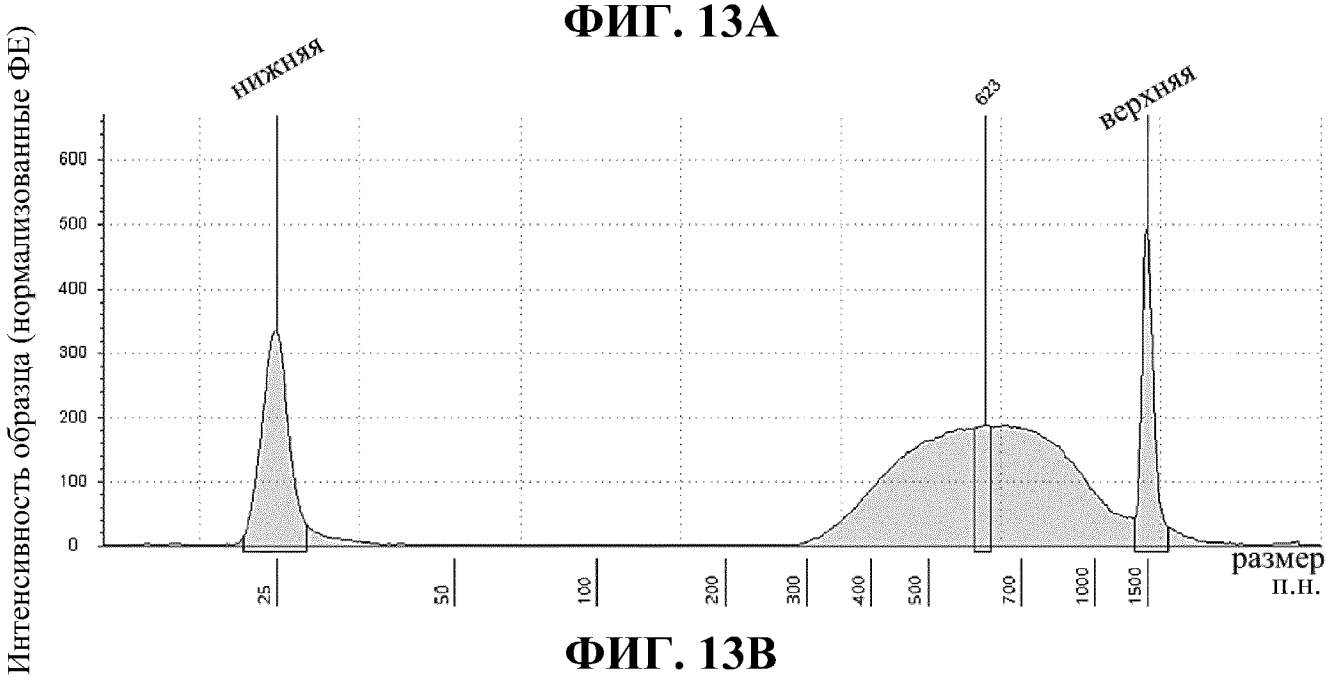
ФИГ. 12C


Доля выбранных оснований при 20X
1.00
0.75
0.50
0.25
0.00
Отдельная клетка
в массе

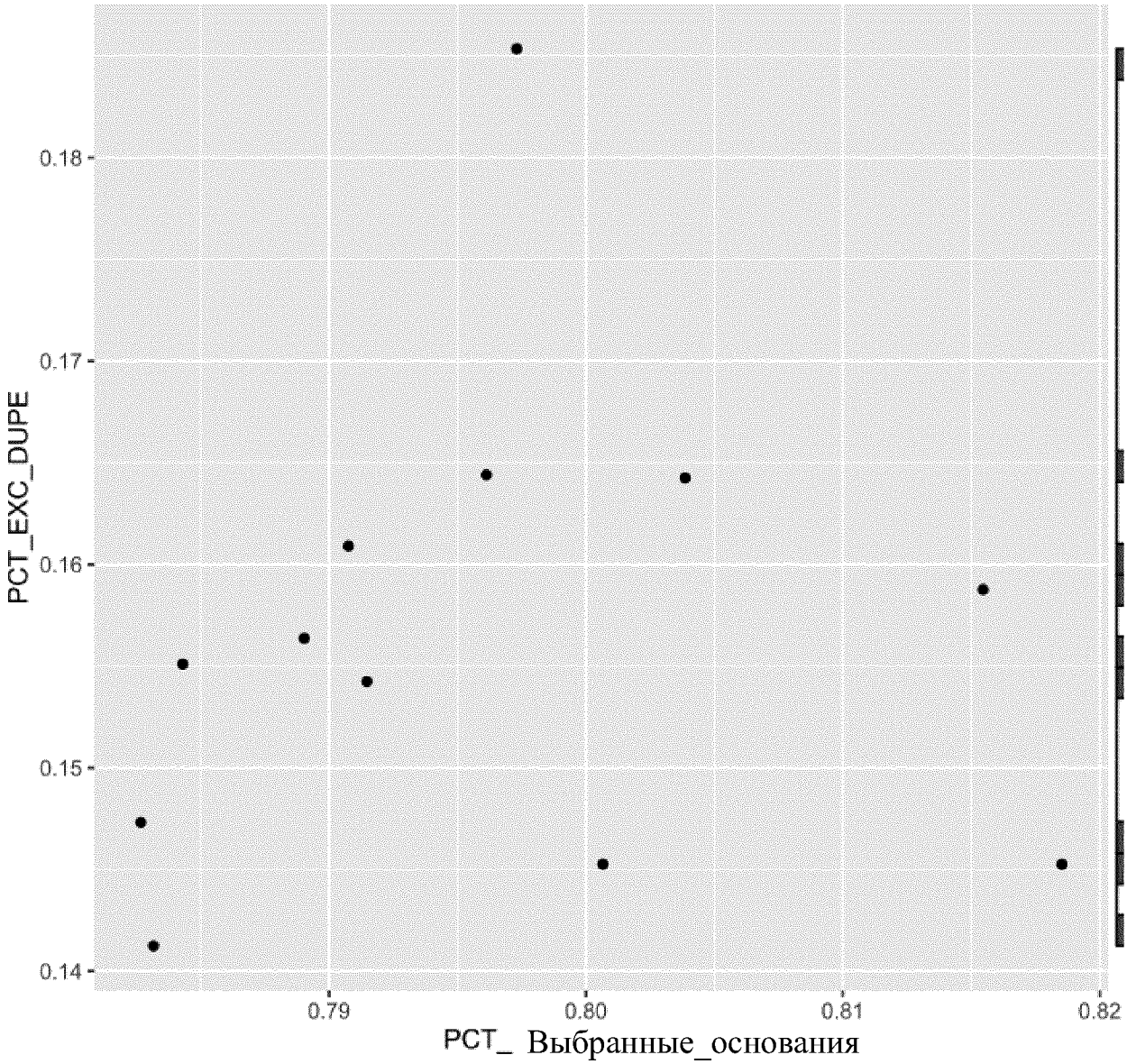
ФИГ. 12D



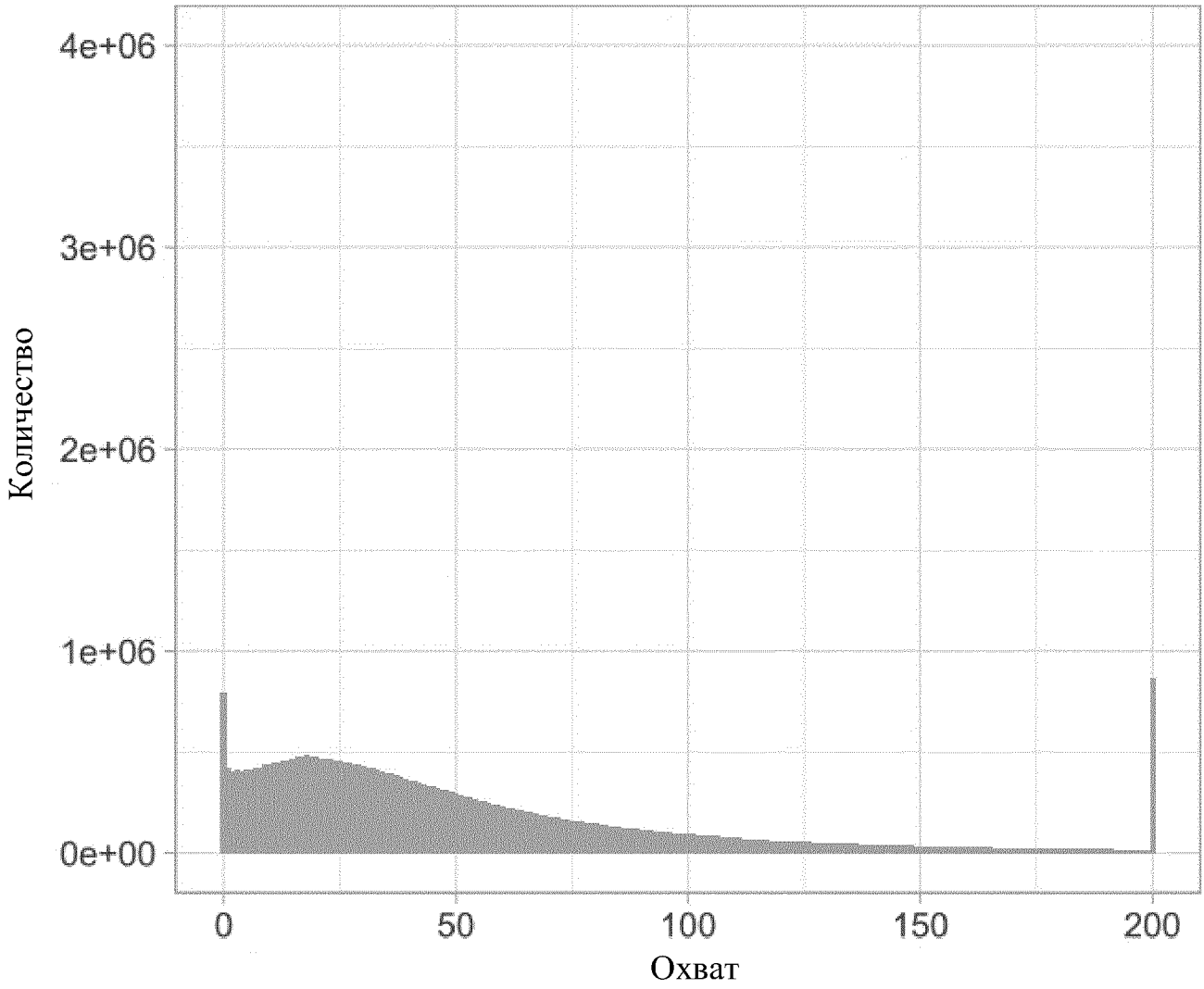
ФИГ. 13A



ФИГ. 13B


0.18
0.17
0.16
0.15
0.14
PCT_EXC_DUPE
0.79
0.80
0.81
0.82
PCT_ Выбранные_основания

ФИГ. 14А



ФИГ. 14B