

МЕТОДЫ ОБОГАЩЕНИЯ ДНК-ФРАГМЕНТОВ ДЛЯ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ

© 2023 г. А. Н. Синяков^а, *, Е. В. Костина^а

^аИнститут химической биологии и фундаментальной медицины, Сибирское отделение Российской академии наук, Новосибирск, 630090 Россия

*e-mail: sinyakov@niboch.nsc.ru

Поступила в редакцию 08.11.2022 г.

После доработки 16.12.2022 г.

Принята к публикации 16.12.2022 г.

В обзоре описано использование олигонуклеотидов, в основном полученных с помощью синтезаторов ДНК нового поколения (микрочиповых синтезаторов ДНК), для обогащения целевых геномных фрагментов. Рассмотрены способы обогащения целевых мишеней с использованием методов молекулярной гибридизации, ПЦР и системы CRISPR-Cas9. Приведены примеры практического применения разработанных методов для исследовательских и диагностических целей.

Ключевые слова: высокопроизводительное секвенирование, олигонуклеотиды, микрочипы, амплификация, гибридизация, диагностика, экзом, CRISPR-Cas9

DOI: 10.31857/S0026898423030126, **EDN:** CHPOFS

ВВЕДЕНИЕ

Несмотря на стремительное развитие технологии ДНК-секвенирования нового поколения (NGS), позволяющей определять полную нуклеотидную последовательность геномов, ее использование для прочтения определенных областей генома затруднительно. Это касается исследований, связанных с определением однонуклеотидных полиморфизмов (SNP), анализом перестройки ДНК, а также при изучении подмножеств участков генома, а не всего генома, например кодирующих белок экзонов. Для выполнения таких задач необходимо проводить “целевое обогащение” анализируемой ДНК и создание специализированных библиотек целевых последовательностей для последующего их секвенирования на выбранной платформе NGS.

Удаление той части геномной ДНК, которая не представляет интереса для проводимого исследования, позволяет обеспечить большую глубину покрытия при секвенировании целевых районов генома, увеличить число мишеней, а также значительно сэкономить время и стоимость эксперимента, в том числе за счет упрощения биоинформатического анализа результатов секвенирования.

Методы обогащения мишеней можно разделить на несколько категорий в зависимости от способа захвата целевого участка ДНК. В одних способах обогащения мишеней используют метод гибридизации нуклеиновых кислот. В этом

случае денатурированную геномную ДНК “отжигают” с одноцепочечными (оц) ДНК- или РНК-зондами к целевым участкам. Связавшись с зондами геномную ДНК выделяют и секвенируют. Другие способы обогащения мишеней основаны на методе ПЦР. Используют мультиплексную ПЦР с десятками тысяч праймеров. Эмульсионная ПЦР позволяет одновременно проводить миллионы цепных реакций и получать миллионы мишеней для секвенирования. В последние годы появились методы обогащения целевых фрагментов ДНК, основанные на использовании РНК-программируемого комплекса CRISPR-Cas9.

СПОСОБЫ ОБОГАЩЕНИЯ ЦЕЛЕВЫХ ФРАГМЕНТОВ ДНК С ПОМОЩЬЮ МЕТОДА МОЛЕКУЛЯРНОЙ ГИБРИДИЗАЦИИ

На первом этапе проводят фрагментацию (физическую или ферментативную) геномной ДНК с последующей ферментативной репарацией концов полученных ДНК-дуплексов и пришивкой к ним специфичных для секвенирующей платформы адаптеров. Адаптеры содержат в своем составе “штрих-код” – уникальную последовательность для каждого образца. После секвенирования биоинформатический анализ позволяет с помощью этого штрих-кода группировать данные, относящиеся к одному образцу. Производители наборов для целевого обогащения ДНК поставляют от 96

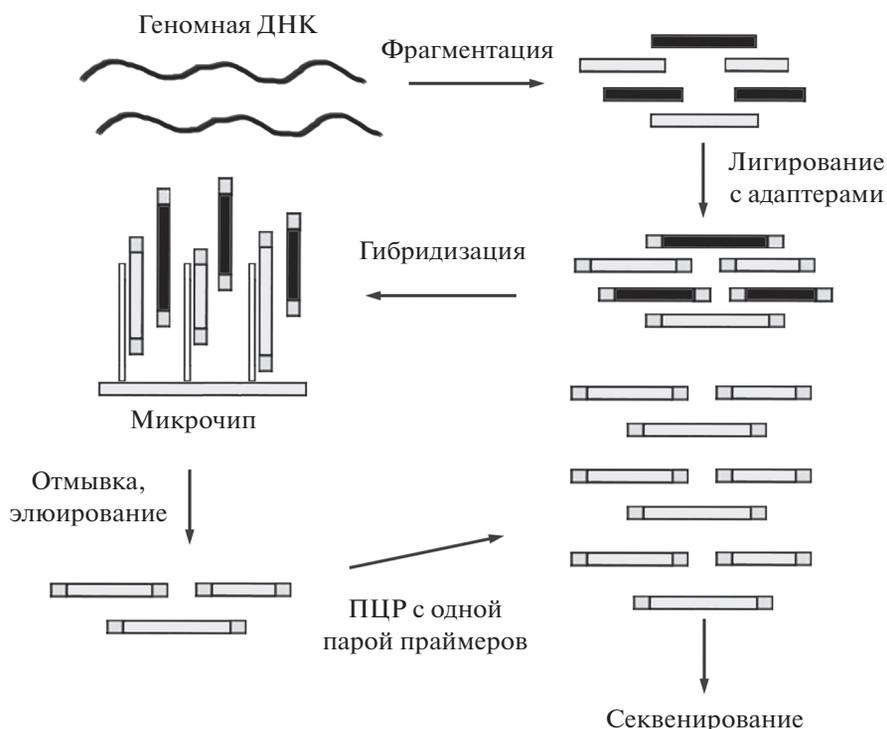


Рис. 1. Схема обогащения геномных фрагментов ДНК с помощью слайдов, содержащих олигонуклеотидные зонды (метод MGS).

до 384 различных адаптеров, содержащих различные штрих-коды. Использование штрих-кодов обеспечивает одновременное секвенирование нескольких образцов и тем самым значительное сокращение временных и финансовых затрат.

Кроме штрих-кода в адаптерах содержатся последовательности праймеров для амплификации. С помощью ПЦР нарабатывается достаточное количество ДНК для последующей стадии гибридизации с одноцепочечными зондами, комплементарными целевым областям генома. Как правило, это ДНК-зонды. Фирма “Agilent” (США) использует в своих наборах SureSelect гибридационные РНК-зонды. Основное преимущество РНК-зондов — большая стабильность РНК:ДНК-гибридов по сравнению с ДНК:ДНК-дуплексами [1].

Гибридизация на твердой поверхности

М. Zwick с сотрудниками [2] первыми использовали слайды “Affymetrix Chip Design Group” (США) с ковалентно связанными олигонуклеотидами для селекции геномных фрагментов (метод MGS, microarray-based genomic selection).

Протокол MGS включает пять основных этапов (рис. 1):

1) расщепление геномной ДНК на случайные фрагменты длиной 300 п.н;

2) репарация концевых фрагментов ДНК с добавлением 3'-выступающих остатков аденина и присоединение уникальных адаптеров с 5'-выступающим тимидином;

3) гибридизация фрагментов ДНК с комплементарными олигонуклеотидами на микрочипе высокой плотности;

4) элюирование фрагментов ДНК, связанных зондами;

5) амплификация выделенных фрагментов с помощью праймеров, интегрированных в адаптеры.

В методе использовали перекрывающиеся гибридационные зонды к обеим цепям дуплекса, при этом каждая цепь перекрывалась зондами от 1.5 до 4 раз. На одном слайде содержалось 385 тыс. гибридационных зондов длиной от 50 до 93 нуклеотидов, позволяющих проводить изотермическую гибридизацию. Зонды покрывали фрагмент генома размером от 4 до 5 млн.п.н. Позднее их число увеличили до 2.1 млн. Эти зонды потенциально могут связаться с 34 млн.п.н. геномной последовательности [3, 4].

Аналогичным методом [5] выделяли более 200 тыс. экзонов для секвенирования. В качестве гибридационных зондов были выбраны олигонуклеотиды к кодирующей последовательности экзона длиной >60 н., перекрывающие последовательность экзона со сдвигом 20 н. Всего использовали в работе 7 слайдов (“Roche NimbleGen”, США), каж-

дый из которых содержал 385 тыс. зондов. В зависимости от применяемого протокола от 55 до 85% выделенных фрагментов ДНК содержали целевые последовательности. Метод позволяет секвенировать до 98% всех целевых экзонов.

Т. Albert и соавт. [6] ввели в схему дополнительную процедуру – амплификацию фрагментов ДНК перед гибридизацией на микрочипе. Секвенирование всего экзона с использованием гибридизации на микрочиповых слайдах доказало свою полезность для выявления редких вариантов генов и мутаций, вызывающих заболевания [7, 8]. Для обогащения 1000 генов из транскриптома хлопка А. Salmon и др. [9] использовали 135 тыс. зондов, содержащихся на слайдах “Roche NimbleGen”.

Nelson с сотрудниками [10] удалось повысить эффективность обогащения целевых мишеней при блокировании побочной гибридизации адаптерных частей зондов с геномной ДНК и дублировании стадии гибридизационного обогащения целевых мишеней.

Для выделения последовательностей генов *BRCA1* и *TP53* в микрофлюидных кассетах микрочипового ДНК-синтезатора Geniom (“Geniom”, Германия) синтезировали 50-членные олигонуклеотиды [11, 12]. На основе этого синтезатора разработали автоматизированный метод HybSelect для подготовки образцов для высокопроизводительного секвенирования. Метод позволяет параллельно проводить обогащение ≤8 образцов размером от 125 т.п.н. до 1 млн.п.н. Метод был апробирован на анализе 115 генов, ассоциированных с развитием злокачественных опухолей [13].

Несмотря на удовлетворительную производительность, разработанный метод HybSelect не лишен существенных недостатков. Во-первых, для проведения селективного выделения целевых дуплексов требуется много (от 10 до 15 мкг) исходного материала ДНК. Во-вторых, необходимость дорогостоящего специализированного оборудования (станции для проведения гибридизации). В-третьих, метод хорошо работает при обогащении геномных фрагментов размером ~500 п.н., но значительно менее эффективен в случае коротких экзонов (~120 п.н.) [5]. В-четвертых, HybSelect сложно масштабировать и автоматизировать.

Гибридизация в растворе

Из-за недостатков обогащения целевых фрагментов ДНК на твердой поверхности “Roche NimbleGen” и другие компании перешли на обогащение мишеней в растворе, где гибридизация идет с избытком гибридизационных зондов. Фирма “Roche NimbleGen” прекратила выпуск слайдов для гибридизации геномных фрагментов непосредственно на микрочипе.

Компания “Agilent” была первой, разработавшей коммерческий продукт, использующий гибридизацию олигонуклеотидов с геномными фрагментами ДНК в растворе [14]. В этом методе применяют синтезированные на микрочиповом ДНК-синтезаторе 200-членные олигонуклеотиды, содержащие 170-членную мишеньспецифическую последовательность, фланкированную двумя праймерами для амплификации. Набор синтезированных олигонуклеотидов отщепляют от слайда и подвергают двураундной ПЦР. В первом раунде используют интегрированные в последовательность олигонуклеотидов праймеры, во втором – в мишеньспецифические дуплексы вводят последовательность промотора фага T7. Затем, используя транскрипцию *in vitro* в присутствии биотин-УТР, получают биотинилированные оцРНК-зонды для выделения целевых геномных фрагментов. Геномную ДНК случайным образом расщепляют на фрагменты длиной ~250 п.н., к которым присоединяют адаптеры. После 12 циклов ПЦР полученные ампликоны гибридизуют с РНК-зондами в растворе. Биотинилированные РНК, содержащие и не содержащие комплементарную геномную ДНК, выделяют из реакционной смеси с помощью магнитных шариков с иммобилизованным стрептавидином. ДНК:РНК-гибриды разрушают и полученную ДНК амплифицируют с помощью ранее введенных в состав фрагментов геномной ДНК адаптеров (рис. 2).

Разработанный метод не требует специального оборудования и совместим с различными секвенирующими платформами. Он был апробирован на секвенировании 15 тыс. экзонов (2.5 млн.п.н.) человеческого генома.

К основным достоинствам метода относятся:

- 1) высокая концентрация оцРНК-зондов, что позволяет повысить эффективность процесса гибридизации с целевыми ДНК-фрагментами;
- 2) относительно небольшое количество геномной ДНК (0.5–3.0 мкг) для секвенирования;
- 3) проведение гибридизации в растворе, что позволяет масштабировать и автоматизировать процесс в отличие от твердофазной гибридизации на слайдах [2];
- 4) нивелирование различий в гибридизации разных аллелей благодаря использованию протяженных зондов;
- 5) возможность готовить и характеризовать большие партии РНК-зондов, что позволяет иметь стандартизованный материал для использования в большом числе экспериментов;
- 6) эффективность при таргетном секвенировании как множества несмежных геномных фрагментов, так и протяженных геномных районов;
- 7) высокая специфичность – успешно определяют 85–90% целевых последовательностей.

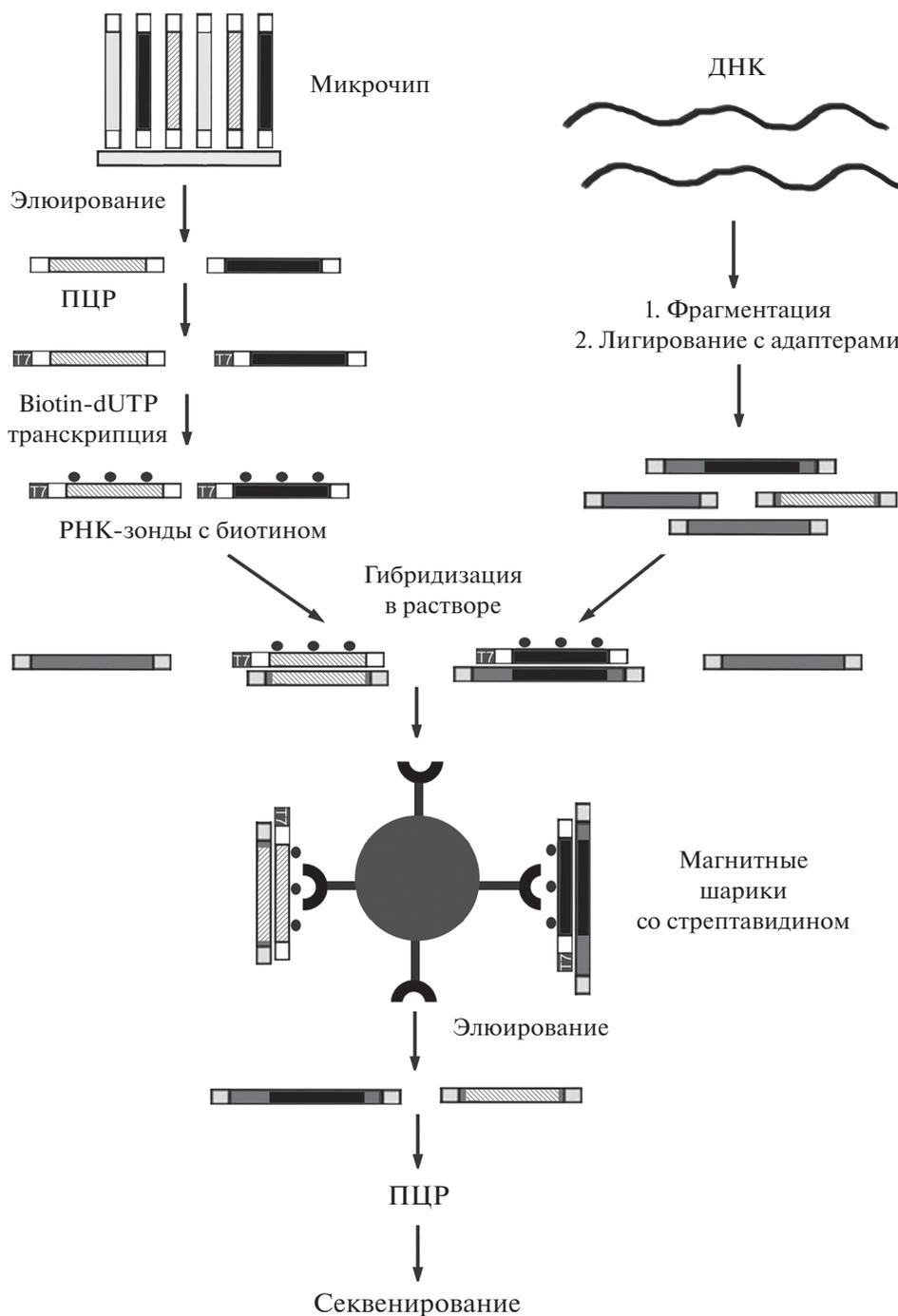


Рис. 2. Схема обогащения геномных фрагментов ДНК с помощью олигонуклеотидных зондов в растворе, применяемая в наборах Agilent.

Созданы коммерческие наборы, основанные на различных вариантах обогащения последовательностей ДНК в растворе: например, Agilent SureSelect (“Agilent”) [15], Illumina TruSeq (“Illumina”, США) [16], Roche NimbleGen SeqCap EZ (“Roche NimbleGen”) [17]. Принципиальная разница между “Agilent” и другими коммерческими наборами заключается в природе гибридизационных

зондов: “Agilent” использует 120–170-членные РНК-пробы, в то время как “Roche NimbleGen” 60–90-членные, а “Illumina” 95-членные ДНК-зонды [3, 16]. Для процедуры проведения обогащения геномной ДНК в растворе применяются стандартные 96-луночные планшеты и термоциклер, то есть специализированного оборудования в этом случае не требуется.

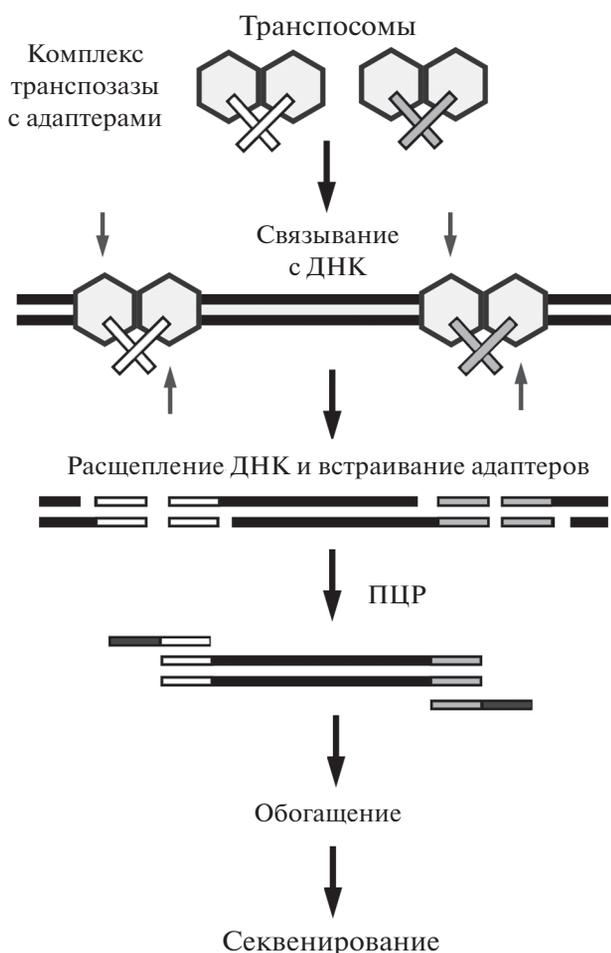


Рис. 3. Схема использования транспозоноопосредованной фрагментации анализируемой ДНК. Комплекс транспозазы с двухцепочечными олигонуклеотидами, содержащими адаптерные участки, связывается с геномной ДНК. Транспозаза расщепляет двухцепочечную геномную ДНК и встраивает олигонуклеотиды на участке разрезания. Полученные молекулы ДНК используют в ПЦР с праймерами, необходимыми для секвенирования.

Модификация метода обогащения целевых фрагментов путем гибридизации олигонуклеотидов с геномными фрагментами ДНК в растворе — транспозоноопосредованная фрагментация анализируемой ДНК (рис. 3).

Данная модификация значительно упрощает процесс получения фрагментов анализируемой ДНК, поскольку фрагментация и присоединение адаптеров происходит за одну стадию с применением фермента транспозазы [18, 19]. Транспозомы (комплекс фермента транспозазы с двухцепочечными олигонуклеотидами, содержащими участок связывания с транспозазой и часть последовательности праймера для последующей амплификации) обладают способностью случайным образом связываться с целевыми последовательностями

ми двухцепочечной геномной ДНК. Транспозазы в транспозоме расщепляют ДНК и одновременно встраивают на участке разрезания олигонуклеотиды из комплекса.

Разработанные методы обогащения экзома оказались эффективными для обнаружения мутаций, служащих причиной редких менделевских заболеваний [20–23], сложных расстройств [24–27], митохондриальных нарушений [28–30]. Также их применяли при скрининге потенциальных генетических мутаций, связанных с раковыми заболеваниями [31–36].

Метод гибридизационного обогащения успешно используют при анализе древней ДНК человека (ancient DNA; аДНК), сохранившейся в человеческих останках. Первым генетическим маркером, проанализированным в палеогенетических исследованиях человека, была митохондриальная ДНК (мтДНК), так как в клетках ее копияемость гораздо выше, чем ядерной ДНК. Для гибридизации использовали биотинилированные ДНК или РНК-зонды, направленные на два гипервариабельных сегмента контрольной области мтДНК [37–43].

Селективное обогащение фрагментов Y-хромосомы аДНК проводили как на твердой поверхности [44], так и в растворе [45].

Carpenter и соавт. [46] разработали метод полногеномного обогащения (whole-genome in-solution capture, WISC) аДНК, используя современную ДНК человека в качестве зондов.

В настоящее время разработаны коммерческие наборы, ориентированные на мтДНК, индивидуальные локусы или целые геномы, такие как myBaitsR³ (“Daicel Arbor Biosciences”, США), которые используют для секвенирования аДНК [47, 48].

Применение смеси олигонуклеотидов, полученных с помощью микрочипового синтезатора, содержащей 962438 зондов, позволило секвенировать экзоны бурундуков [49]. Из целевых 11975 экзонов обогатить удалось более 99%. Зонды были рассчитаны на основе известной последовательности генома бурундука *Tamias alpinus*. Однако эти же зонды были успешно использованы для селективного выделения экзонов родственных видов: *Tamias amoenus*, *Tamias ruficaudus* и *Tamias striatus*.

Аналогичным образом, используя зонды для экзонов генов домашней коровы (*Bos taurus*), T. Cosart и др. [50] секвенировали 16131 экзон быка (*Bos indicus*) и дикого бизона (*Bison bison*). Предложенный метод селективного обогащения геномной ДНК применили для секвенирования ряда хромосом и митохондриального генома западных шимпанзе (*Pan troglodytes verus*) [51]. R. Tewhey и соавт. [52] для гибридизации использовали перекрывающиеся РНК-зонды длиной 120 нуклеотидов. На мишени размером 3.9 млн.п.н. они

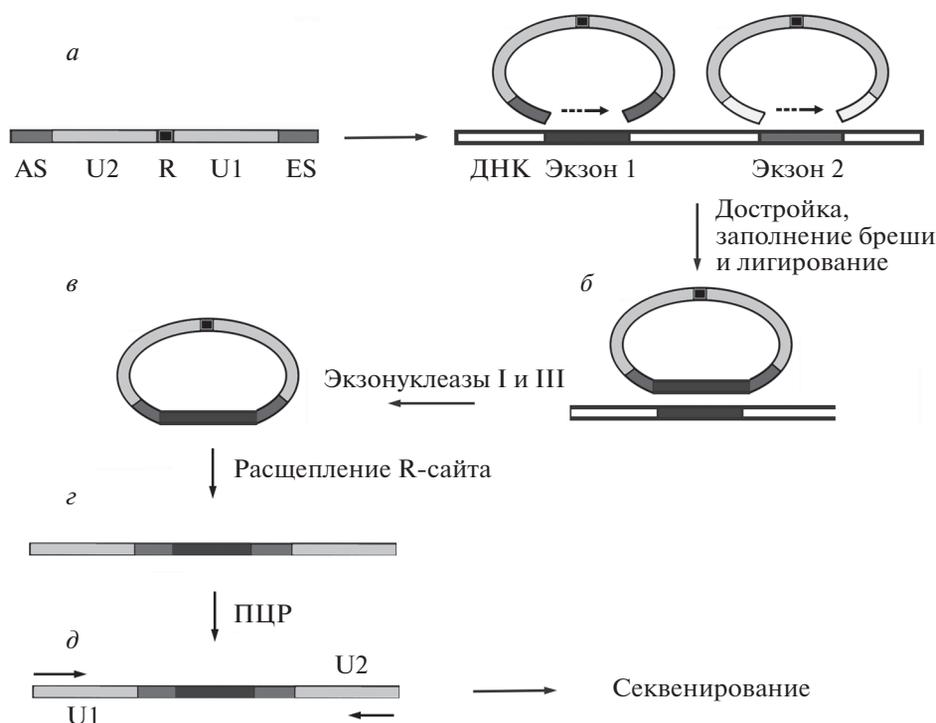


Рис. 4. Схема обогащения целевых фрагментов ДНК с использованием инверсионных зондов CIPer. *а* – Структура CIPer. AS и ES – 20-звенные олигонуклеотиды, гомологичные целевым фрагментам геномной ДНК; R – сайт рестрикции; U1 и U2 – универсальные праймеры для амплификации инвертированного зонда. *б* – Циклизация CIPer путем заполнения бреши ДНК-полимеразой и лигирования. *в* – Обогащение кольцевой ДНК за счет деградации линейной формы ДНК экзонуклеазами I и III. *г* – Линейаризация кольцевой формы ДНК путем расщепления сайта рестрикции. *д* – Амплификация полученной линейной формы ДНК с праймерами U1 и U2.

картировали 93% мутаций с точностью более 99%. Подход Agilent был успешно применен также при проведении секвенирования мутаций в генах, ответственных за потерю слуха [53].

С помощью гибридизации в растворе с ДНК-олигонуклеотидами были получены для секвенирования образцы ретровирусов, интегрированные в геном человека [54].

Добавление коммерческого препарата *Cot1* DNA при проведении гибридизации, как на поверхности, так и в растворе, увеличивает ее специфичность [14, 55]. *Cot1* DNA состоит из коротких фрагментов (50–300 п.н.) плацентарной ДНК человека, обогащенной повторяющимися последовательностями. При гибридизации на слайде или в растворе *Cot1* DNA обычно добавляют соответственно в 5- и 20-кратном избытке по отношению к геномной ДНК.

СПОСОБЫ ОБОГАЩЕНИЯ ЦЕЛЕВЫХ ФРАГМЕНТОВ ДНК, ОСНОВАННЫЕ НА ИСПОЛЬЗОВАНИИ МЕТОДА МОЛЕКУЛЯРНОЙ ИНВЕРСИИ

Разработанные М. Nilsson с соавт. [56] инверсионные зонды в сочетании с мультиплексной ПЦР нашли широкое применение при детектировании

точечных мутаций. Так, М. Akhras и др. [57] использовали их для обогащения мишеней NGS. Схема применения таких зондов (Connector Inversion Probe; CIPer) приведена на рис. 4. Разработанный метод был успешно апробирован авторами при генотипировании вируса папилломы человека (HPV), ассоциированных с онкологическими заболеваниями вирусов, скрининге мутаций антибиотикорезистентности в геноме гонококка (*Neisseria gonorrhoeae*).

Метод коммерциализирован под названием HaloPlex (“Halo Genomics”, Швеция). В настоящее время набор производит компания “Agilent Technologies”. Их использование позволяет выделять 21 тыс. генов общим размером в 37 млн.п.н. Зонды такого типа получили общее название “молекулярные инверсионные пробы” (molecular inversion probes, MIP).

Способ обогащения целевой ДНК с использованием MIP был применен для исследования большой группы пациентов с дистонией на наличие патогенных и редких вариантов гена *ATM* [58], для анализа мутаций гена *GNAS*, вызывающих внутримышечную миксому [59].

Применение MIP в совокупности с таргетным РНК-секвенированием нового поколения позволило авторам создать независимую от гистологии

молекулярную диагностику опухолей головного мозга [60]. Аналогичный подход использовали для анализа профилей экспрессии генов в случае почечно-клеточного рака [61].

Применение МІР для амплификации вариабельных областей V1, V2, V3, V6 и V7 гена 16S рРНК и последующее NGS полученных ампликонов позволило создать метод диагностики бактериальных возбудителей [62]. МІР оказались эффективными при создании автоматизированного процесса анализа мутаций генов *BRCA1* и *BRCA2* с целью диагностики рака молочной железы [63]. Использование МІР при секвенировании гена *SCN1A* позволило создать метод диагностики родительского мозаицизма при синдроме Дауна [64]. Аналогично был проведен анализ мозаичности гена *PIK3CA* при нарушении развития [65].

Y. Wada и соавт. [66], используя МІР, выявляли в генах семейства PPAR/RXR мутации, ассоциированные с развитием шизофрении. NGS с использованием МІР применяли при анализе мутаций генов ионных каналов при диабетической периферической нейропатии [67] и для анализа мутаций в генах *QRDR* [68], определяющих устойчивость к ципрофлоксацину в бактериях, вызывающих сибирскую язву, чуму и туляремию. Разработана методика NGS с применением МІР, предназначенная для исследований в области фармакогеномики [69].

По сравнению с другими способами обогащения мишеней методы на основе МІР имеют ряд преимуществ. Они высокоспецифичны к мишеням, работают с мизерным количеством (≤ 5 пг) геномной ДНК и не требуют этапов предварительной обработки, таких как фрагментация ДНК. Кроме того, при использовании для NGS ограниченного числа МІР можно применять для их синтеза традиционные колоночные и планшетные синтезаторы.

Для создания тысяч МІР, захватывающих большое число целей одновременно, обычно используют микрочиповые синтезаторы. При получении одноцепочечных олигонуклеотидов с помощью микрочиповых синтезаторов необходимо провести ряд последовательных стадий, таких как расщепление эндо- или экзонуклеазами, очистку полученных продуктов в геле и амплификацию. D. Vang с сотрудниками [70] разработали значительно более простой и эффективный протокол применения МІР из синтезированных на микрочиповых синтезаторах олигонуклеотидов — с использованием только одного фермента и двухцепочечных МІР-зондов. Метод был успешно апробирован на 11 510 дуплексных МІР, предназначенных для выделения 3554 экзонов 228 генов в образце геномной ДНК.

СПОСОБЫ ОБОГАЩЕНИЯ ЦЕЛЕВЫХ ФРАГМЕНТОВ ДНК С ПОМОЩЬЮ ПЦР

Моноплексная ПЦР

ПЦР с момента появления в 1980 годах предназначалась для наработки целевого фрагмента ДНК с помощью праймеров для амплификации [71]. Именно поэтому технология ПЦР была применена для NGS. В моноплексной ПЦР в одной пробирке с помощью одной пары праймеров нарабатывают только один ДНК-дуплекс. После амплификации всех целевых фрагментов их объединяют в эквимольных количествах для создания библиотеки для секвенирования. Получение ампликонов с помощью моноплексной ПЦР может быть очень трудоемким процессом, поэтому эта стадия была автоматизирована. Например, для упрощения процедуры получения ампликонов компания “Fluidigm” (Канада) разработала микрофлюидное устройство Access Array system, позволяющее автоматизировать процесс получения целевых ампликонов. Одновременно можно получить 480 различных ампликонов [72].

Эмульсионная технология RainDance (“Bio-Rad”, США) [73], использующая моноплексную ПЦР, значительно более производительна по сравнению с описанной выше. В ней генерируют миллионы капель двух типов: один содержит геномную ДНК (1 молекула геномной ДНК, дезоксирибонуклеозидтрифосфаты, ДНК-полимераза в буфере), а другой только молекулы одной пары праймеров. Затем эти типы капель смешивают в соотношении 1 : 1. Примерно 1.5 млн капель ПЦР собирают в одну пробирку для ПЦР объемом 0.2 мл, амплифицируют в стандартном термоциклере с последующим разрушением эмульсии с высвобождением ампликонов в раствор. После очистки смесь секвенируют.

Технологии как RainDance, так и Fluidigm совместимы с большинством коммерческих секвенаторов и требуют только несколько наногرامмов исходного материала для получения ампликонов для секвенирования. Однако технология RainDance позволяет получать значительно больше ампликонов одновременно: до 20 тыс. на образец по сравнению с 480 для Fluidigm. В продуктах компании “Bio-Rad” есть панели ампликонов, связанных с мутациями в генах, ассоциированных с онкологическими заболеваниями и аутизмом.

Мультиплексная ПЦР

Мультиплексная ПЦР — реакция, в которой несколько пар праймеров одновременно генерируют несколько ампликонов из одного и того же исходного материала. До недавнего времени использование мультиплексной ПЦР было ограничено — из-за высокого уровня неспецифической

амплификации. Однако в последние годы несколько компаний разработали алгоритмы, позволяющие значительно повысить селективность мультиплексной ПЦР.

Компания “Illumina” разработала вариант мультиплексной ПЦР, включающий следующие стадии:

1) гибридизацию олигонуклеотидов к их комплементарным последовательностям в интересующей области;

2) достраивание одного из пары олигонуклеотидов;

3) сшивание достроенного фрагмента со вторым олигонуклеотидом;

4) ПЦР-амплификацию полученного фрагмента ДНК с использованием универсальных праймеров, содержащих индексные последовательности.

Разработанный подход позволяет создавать разнообразные панели, в том числе и нестандартные. Пользовательские панели могут быть созданы для 1536 мишеней в геномах человека, мыши, крысы и коровы. Сконструированы панели TruSeq Amplicon Cancer Panel (“Illumina”), нацеленные на гены, связанные с развитием рака. Одна из них создана для секвенирования 212 ампликонов 48 ассоциированных со злокачественными опухолями генов. TruSight Tumor Panel (“Illumina”) разработана для получения 174 ампликонов, покрывающих сайты мутаций в 26 ассоциированных с развитием рака генах [74].

Компания “Thermo Fisher Scientific” (США) разработали технологию Ion AmpliSeq, позволяющую проводить до 24 тыс. амплификаций в одной пробирке (рис. 5).

В протоколе этого метода предусмотрено использование всего 1 нг ДНК и РНК. Технология AmpliSeq получила широкое распространение во всем мире. Так, ее использовали в исследованиях злокачественных опухолей [75], наследственных заболеваний [76] и бактерий [77].

На основе технологии AmpliSeq разработаны наборы для исследования либо определенных генов (например, Ion AmpliSeq *TP53* Panel, Ion AmpliSeq *BRCA1* и *BRCA2* Panel), либо заболеваний (например, муковисцидоза, деменции, рака толстой кишки и легкого). Создан набор для полноэкзомного секвенирования, содержащий 294 тыс. пар праймеров.

СПОСОБЫ ОБОГАЩЕНИЯ ЦЕЛЕВЫХ ФРАГМЕНТОВ ДНК С ПОМОЩЬЮ СИСТЕМЫ CRISPR-Cas9

В последние годы стали появляться методы обогащения целевых фрагментов ДНК, основанные на использовании РНК-программируемого ком-

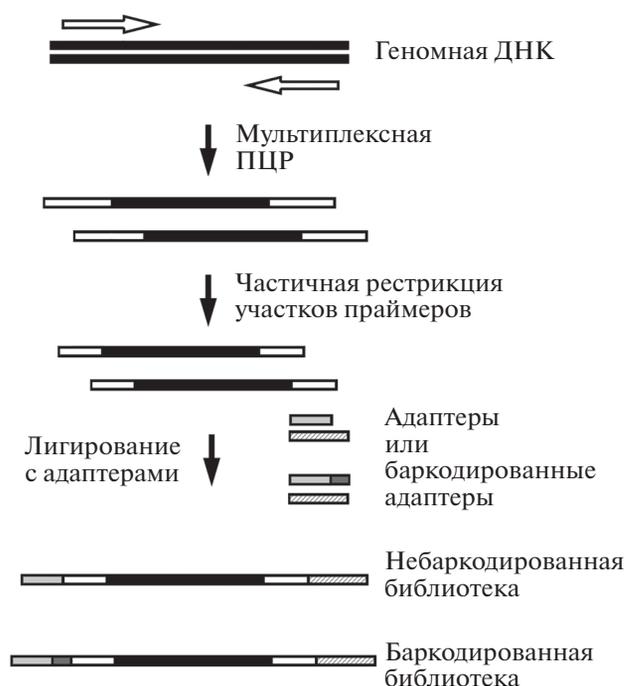


Рис. 5. Схема подготовки библиотек для секвенирования по технологии AmpliSeq.

плекса CRISPR-Cas9 [78, 79]. Системы CRISPR-Cas обеспечивают бактериям приобретенный иммунитет к вирусам. CRISPR-ассоциированный белок Cas9 — это эндонуклеаза, которая использует направляющую последовательность (РНК-гид), чтобы сформировать комплекс с ДНК-мишенями, что позволяет Cas9 произвести сайтспецифичный двухцепочечный разрыв в ДНК (рис. 6). Эта система направленно действует на любую последовательность геномной ДНК.

Истощение фоновых ДНК

Прежде всего, система CRISPR-Cas была использована для обогащения целевых ДНК путем истощения нежелательных. Метод обогащения получил название DASH (Depletion of Abundant Sequences by Hybridization) [80]. На первом этапе конструируют библиотеку фрагментов ДНК для секвенирования, содержащую в частности праймеры для ее амплификации. Разрушение ненужных для секвенирования последовательностей в полученной библиотеке проводят с помощью РНК-гидов, направленных на неинформативные последовательности. После амплификации дуплексов, содержащих в своем составе оба амплификационных праймера, библиотеку секвенируют. Используя этот метод, J. DeRisi с сотрудниками [80] удалось снизить содержание митохондриальной рРНК в клетках HeLa на два порядка и обогатить содержание целевых последовательностей

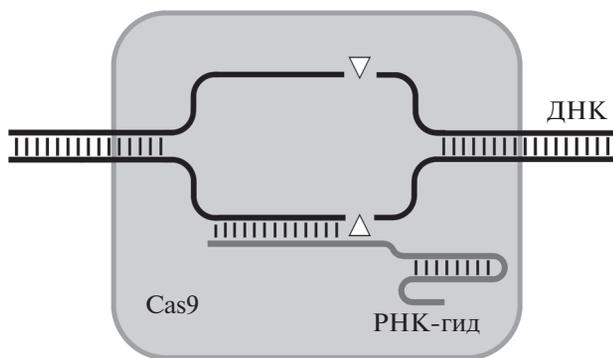


Рис. 6. Схема комплекса РНК-гид:Cas9 с ДНК-мишенью. Треугольниками показаны места сайтспецифичного разрезания ДНК.

патогенов в образцах пациентов. Модифицированный метод DASH был применен S. Вае с сотрудниками [81] для разработки сверхчувствительного способа обнаружения циркулирующих опухолевых ДНК.

R. Stevens и др. [82] разработали метод выделения длинных фрагментов ДНК (10–36 т.п.н.) с помощью система CRISPR-Cas. После расщепления исходной ДНК эндонуклеазой Cas9 оба конца целевого фрагмента ДНК остаются связанными с ферментом. Обработка реакционной смеси экзонуклеазами приводит к разрушению фоновых последовательностей, но не затрагивает защищенные Cas9 целевые последовательности. Разработанный метод позволяет обогатить мишени в 30–600 раз.

Применение гель-электрофореза, аффинной хроматографии и гибридизации

В нижеприведенных методах ДНК разрезают с использованием Cas9 и целевые фрагменты выделяют электрофорезом в агарозном геле. Такие методики позволяют выделять очень большие неповрежденные участки генома. Bennett-Baker & Mueller [83] с помощью пульсирующего электрофореза выделили геномные фрагменты размером 2 млн.п.н. со стократным обогащением. T. Gabrieli с соавт. [84] усовершенствовали метод и апробировали его на секвенировании бактериального фрагмента длиной 200 т.п.н. с помощью нанопорового секвенатора. В рассмотренном методе используют приборы для импульсного гель-электрофореза, а также ДНК с чрезвычайно высокой молекулярной массой, что доступно далеко не во всех лабораториях. Однако есть другие методы, которые обходят эти проблемы.

D. Nachmanson с соавт. [85] разрезали геномную ДНК на фрагменты длиной ~500 п.н. с помощью набора РНК-гидов и выделяли целевые фрагменты гель-электрофорезом либо экстрак-

цией геномной ДНК. Таким способом удалось добиться обогащения в 49 тыс. раз. J. Lee и др. [86] гидролизovali целевую последовательность размером 13 т.п.н. на фрагменты, размеры которых были оптимальны для создания библиотеки для секвенатора Illumina. Целевые фрагменты выделяли с помощью гель-электрофореза либо конъюгированных со стрептавидином шариков. В последнем случае РНК-гиды были биотинилированы.

Идея обогащения целевых ДНК путем выделения комплекса РНК-гид:Cas9:ДНК реализована в нескольких патентах. Например, авторы работ [87, 88] выделяли комплекс, включающий биотинилированный фермент Cas9, с помощью стрептавидинсодержащих шариков. D. Bang и др. [89] использовали для этой цели нуклеазу Cas9, содержащую полигистидиновый тег, и металл-хелатную хроматографию. X. Xu и соавт. [90] выделяли комплекс путем гибридизации удлиненного РНК-гида с комплементарным олигонуклеотидом, иммобилизованным на твердой поверхности. A. Aalipour и др. [91] использовали каталитически неактивный белок Cas9 с полигистидиновым тегом для выделения целевых ДНК и аллельспецифическую количественную ПЦР для выявления редких мутаций ДНК. Разработанный метод позволяет выявить одну такую мутацию на фоне 1000 аллелей дикого типа.

Другой способ применения системы CRISPR-Cas для обогащения целевых последовательностей – FLASH (Finding Low-Abundance Sequences by Hybridization) – связан с подготовкой библиотеки для секвенирования. В этом методе исходную геномную ДНК подвергают дефосфорилированию и разрезают с использованием технологии CRISPR-Cas. Продукты расщепления, содержащие концевые фосфатные группы на целевых фрагментах, способны присоединять адаптеры для секвенирования. После амплификации происходит значительное обогащение библиотеки целевыми последовательностями ДНК. Так, при анализе маркеров устойчивости к антибиотикам удалось достичь 5000-кратного обогащения библиотеки анализируемыми ДНК [92]. Если применять методы секвенирования, позволяющие “прочитывать” протяженные участки ДНК, необходимости в амплификации ДНК нет. Так, N. Hafford-Teag и соавт. [93], используя комплекс CRISPR-Cas и метод одномолекулярного секвенирования в реальном времени (PacBio SMRT), проанализировали мутации, ассоциированные с эндотелиальной дистрофией роговицы (дистрофией Фукса). Варианты метода FLASH получили распространение в нанопоровом секвенировании [94–96].

Систему CRISPR/Cas9 также используют для фрагментации ДНК с целью генерации фрагментов одинаковой длины. В наиболее широко используемом методе – обработке ультразвуком – генериру-

ются фрагменты ДНК произвольного размера, что может приводить к проблемам при секвенировании, связанным с низким уровнем “покрытия”, неравномерным “покрытием” и ложными мутациям. Подход, названный CRISPR-DS, успешно использован для секвенирования экзонных областей *TP53* [85]. Целевые фрагменты ДНК для секвенирования могут быть выделены с помощью простого фракционирования по молекулярному весу, что приводит к обогащению примерно в 49 тыс. раз. Метод позволяет снизить на один–два порядка массу исследуемой ДНК.

На основе системы CRISPR-Cas быстро растет разнообразие методов обогащения целевых последовательностей ДНК, однако до сих пор ни один из них не достиг эффективности таковых с применением гибридизации нуклеиновых кислот.

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ ТАРГЕТНОГО СЕКВЕНИРОВАНИЯ В МЕДИЦИНСКОЙ ДИАГНОСТИКЕ

Методы NGS быстро развиваются и стремительно внедряются в практическую медицину. Многие заболевания человека имеют генетическую основу. Некоторые из них обусловлены отсутствием или дисфункцией определенного белка из-за мутаций в кодирующем гене. Так обстоит дело с заболеваниями менделевской наследственности, такими как болезнь Хантингтона, талассемия и около тысячи других наследственных редких заболеваний [97]. Все больше генетических вариантов и полиморфизмов идентифицируют в качестве факторов риска орфанных заболеваний [98]. Часто развитие злокачественных новообразований тоже опосредовано генетически: мутацией одного или нескольких генов, которые либо повышают риск злокачественной трансформации клеток (например, мутации зародышевой линии), либо способствуют развитию опухоли (онкогены), либо нарушают клеточные механизмы, контролирующие пролиферацию клеток (гены-супрессоры), как это происходит при соматических мутациях [99]. Таргетное или полноэкзомное секвенирование генома пациента позволяют поставить диагноз и выработать адекватную тактику лечения.

Использование в медицинской диагностике NGS не ограничено генетическими заболеваниями. Высокопроизводительное секвенирование применяют и в исследованиях инфекционных заболеваний человека [100]. Эту технологию эффективно используют для выявления малярийных плазмодиев (*Plasmodium* spp.) и возбудителей микозов. Так, для обнаружения генома *Plasmodium falciparum* на фоне геномной ДНК человека применяли метод гибридизационного обогащения в растворе [101, 102]. Предложенный подход был распространен на анализ *Plasmodium vivax* [103] и

Candida albicans [104]. Предварительное обогащение генома бактерий с последующим секвенированием может быть очень полезным в клинической практике. А. Brown с соавт. [105] использовали NGS для определения устойчивости *Mycobacterium tuberculosis* H37Rv к противомикробным препаратам.

Применив для полногеномного обогащения РНК-зонды в растворе, М. Christiansen и др. [106] описали быстрый способ обнаружения и анализа генома *Chlamydia trachomatis*, превышающий на порядок по чувствительности другие известные методы. Аналогичным образом предлагают проводить мониторинг некультивируемого менингококка *Neisseria meningitidis* в клинических образцах [107].

В вирусологических исследованиях технология NGS находит применение при выявлении лекарственной устойчивости патогенов, разработке новых лекарственных средств и вакцин [108].

Первым технологию целевого обогащения вирусных геномов из клинических образцов использовали D. Depledge с соавт. [109]. С помощью 120-членных РНК-зондов авторы провели обогащение геномов вирусов семейства Herpesviridae: варицелла-зостер, Эпштейна–Барр и герпесвируса, ассоциированного с саркомой Капоши. Полноразмерные геномы этих вирусов были реконструированы и использованы в работах по исследованию структуры и разнообразию вирусной популяции. Применяя предложенный подход, многие авторы исследовали геномы вируса Эпштейна–Барр [110], вируса Зика [111], норовируса [112], вируса Ласса [113], вируса гепатита С [114], герпесвируса-7 человека [115], вирусов простого герпеса-1 и -2 [116].

Т. Wylie и др. [117] создали гибридизационную панель ViroCap, предназначенную для обогащения нуклеиновой кислоты ДНК- и РНК-содержащих вирусов 34 семейств, которые инфицируют позвоночных и включают 190 вирусных родов и 337 видов. Панель была создана на основе анализа 1×10^9 геномных последовательностей, а ее применение позволило увеличить число прочтенных вирусных геномов в 296–674 раз. Разработанная панель может быть востребована как в фундаментальных, так и в прикладных исследованиях.

Т. Briese и др. [118] разработали платформу ViroCapSeq-VERT для секвенирования виroma. В систему входит около 2 млн зондов, которые охватывают геномы вирусов, поражающих позвоночных, включая человека. Список всех родов вирусов, которые инфицируют позвоночных, был составлен из списка основных видов Международного комитета по таксономии вирусов (International Committee on Taxonomy of Viruses). Всего для выбора зондов было использовано 342438 белоккодирующих вирусных последовательностей.

Библиотека биотинилированных олигонуклеотидов была синтезирована на основе массивов данных NimbleGen (NimbleGen Systems Inc., США) и использована для захвата в растворе вирусных нуклеиновых кислот. Использование VirCapSeq-VERT привело к 100–10000-кратному увеличению числа прочтений вирусного генома в образцах крови и гомогенатов тканей по сравнению с секвенированием Illumina. Для технологии VirCapSeq-VERT предел обнаружения сравним со специфической ПЦР в реальном времени в сыворотке, крови и экстрактах тканей. Кроме того, этот метод позволяет идентифицировать новые вирусы, геномы которых примерно на 40% отличаются от известных вирусов, использованных для разработки библиотеки зондов.

В 2018 году был разработан аналогичный подход – ViroFind [119]. Библиотека ViroFind включает 165433 вирусных зонда, которые охватывают геномы 535 отобранных ДНК- и РНК-вирусов, инфицирующих человека или способных вызывать зооноз. ViroFind использовали для обнаружения и анализа всех вирусных популяций в головном мозге пяти пациентов с прогрессирующей многоочаговой лейкоэнцефалопатией. По сравнению с прямым глубоким секвенированием, с помощью ViroFind удалось обогатить вирусные последовательности, присутствующие в клинических образцах, до 127 раз. Предложенный подход был распространен на анализ ретровирусных геномов (вирус Т-клеточного лейкоза типа 1 человека и вирус иммунодефицита типа 1 человека), интегрированных в геном человека [54]. При использовании вирусспецифичных зондов обогащение геномных последовательностей вирусов в сотни и тысячи раз превосходило этот показатель при прямом секвенировании.

ЗАКЛЮЧЕНИЕ

Наиболее эффективными методами таргетного обогащения в настоящее время считаются основанные на гибридизации в растворе, мультиплексной ПЦР и молекулярной инверсии. Разработаны эффективные компьютерные программы, позволяющие проводить подбор олигонуклеотидов для таргетного и полногеномного секвенирования. Ряд фирм производит наборы для полноэкзомного обогащения, которые при этом постоянно совершенствуются. Эти методы основаны на гибридизации в растворе. Проведены сравнительные анализы специфичности и производительности таких наборов [120–123].

В целом, все платформы работают хорошо, но между ними выявляют и небольшие различия. Так, РНК-зонды образуют более прочные комплексы с мишенью по сравнению с ДНК-зондами. При сверхглубоком секвенировании РНК-зонды, направленные на обе цепи мишени, работают во

всех случаях лучше, чем РНК-зонды направленные только на одну цепь мишени. ДНК-зонды более эффективны для анализа GC-богатых участков, а РНК-зонды – для AT-областей.

С развитием методов таргетного и полноэкзомного секвенирования появляются новые методы медицинской диагностики. Для развития высокоэффективных современных методов диагностики, основанных на использовании NGS, требуется не только инструментальная база (секвенаторы ДНК), но и современная индустрия для получения массива синтетических олигонуклеотидов (сотен тысяч и миллионов праймеров для амплификации ДНК и зондов для гибридизации). Традиционные колоночные и планшетные синтезаторы ДНК малоприспособлены для получения таких массивов олигонуклеотидов. Для такого рода задач необходимы микрочиповые синтезаторы ДНК. Экспорт высокопроизводительных зарубежных синтезаторов в нашу страну был запрещен с момента их создания. Учитывая исключительную важность таких приборов для современной фундаментальной и прикладной науки, с целью создания макета микрочипового синтезатора олигонуклеотидов Сибирским отделением Российской академии наук был организован консорциум, в который вошли следующие институты: Институт химической биологии и фундаментальной медицины, Институт физики полупроводников, Институт органической химии, Институт автоматизации и электрометрии. В настоящее время эта работа успешно завершена [124, 125]. Разработанный макет позволяет синтезировать до 12 тыс. олигонуклеотидов в одном эксперименте. Полученный прибор открывает возможность проведения исследований с использованием таргетного секвенирования и развивать синтетическую биологию в нашей стране на мировом уровне.

Написание обзора не потребовало специального финансирования.

Настоящая статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Lesnik E.A., Freier S.M. (1995) Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry*. **34**, 10807–10815.
2. Okou D.T., Steinberg K.M., Middle C., Cutler D.J., Albert T.J., Zwick M.E. (2007) Microarray-based genomic selection for high-throughput resequencing. *Nat. Methods*. **4**(11), 907–909.
3. Mamanova L., Coffey A.J., Scott C.E., Kozarewa I., Turner E.H., Kumar A., Howard E., Shendure J., Turner D.J. (2010) Target-enrichment strategies for

- next-generation sequencing. *Nat. Methods*. **7**, 111–118.
4. Kozarewa I., Armisen J., Gardner A.F., Slatko B.E., Hendrickson C.L. (2015) Overview of target enrichment strategies. *Curr. Protoc. Mol. Biol.* **112**, 7.21.1–7.21.23.
 5. Hodges E., Xuan Z., Balija V., Kramer M., Molla M.N., Smith S.W., Middle C.M., Rodesch M.J., Albert T.J., Hannon G.J., McCombie W.R. (2008) Genome-wide *in situ* exon capture for selective resequencing. *Nat. Genet.* **39**(12), 1522–1527.
 6. Albert T.J., Molla M.N., Muzny D.M., Nazareth L., Wheeler D., Song X., Richmond T.A., Middle C.M., Rodesch M.J., J Packard C.J., Weinstock G.M., Gibbs R.A. (2007) Direct selection of human genomic loci by microarray hybridization. *Nat. Methods*. **4**(11), 903–905.
 7. Choi M., Scholl U.I., Ji W., Liu T., Tikhonova I.R., Zumbo P., Nayir A., Lu A.B., Ozen S., Sanjad S., Nelson-Williams C., Farhi A., Mane S., Lifton R.P. (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci. USA*. **106**(45), 19096–19101.
 8. Ng S.B., Turner E.H., Robertson P.D., Flygare S.D., Bigham A.W., Lee C., Shaffer T., Wong M., Bhattacharjee A., Eichler E.E., Bamshad M., Nickerson D.A., Shendure J. (2009) Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*. **461**, 272–276.
 9. Salmon A., Udall J.A., Jeddleloh J.A., Wendel J. (2012) Targeted capture of homoeologous coding and noncoding sequence in polyploid cotton. *G3* (Bethesda). **2**(8), 921–930.
 10. Lee H., O'Connor B.D., Merriman B., Funari V.A., Homer N., Chen Z., Cohn D.H., Nelson S.F. (2009) Improving the efficiency of genomic loci capture using oligonucleotide arrays for high throughput resequencing. *BMC Genomics*. **10**, 646.
 11. Summerer D., Wu H., Haase B., Cheng Y., Schracke N., Stähler C.F., Chee M.S., Stähler P.F., Beier M. (2009) Microarray-based multicycle-enrichment of genomic subsets for targeted next-generation sequencing. *Genome Res.* **19**(9), 1616–1621.
 12. Bau S., Schracke N., Kränzle M., Wu H., Stähler P.F., Hoheisel J.D., Beier M., Summerer D. (2009) Targeted next-generation sequencing by specific capture of multiple genomic loci using low-volume microfluidic DNA arrays. *Anal. Bioanal. Chem.* **393**, 171–175.
 13. Summerer D., Schracke N., Wu H., Cheng Y., Bau S., Stähler C.F., Stähler P.F., Beier M. (2010) Targeted high throughput sequencing of a cancer-related exome subset by specific sequence capture with a fully automated microarray platform. *Genomics*. **95**, 241–246.
 14. Gnirke A., Melnikov A., Maguire J., Rogov P., LeProust E.M., Brockman W., Fennell T., Giannoukos G., Fisher S., Russ C., Gabriel S., Jaffe D.B., Lander E.S., Nusbaum C. (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* **27**, 182–189.
 15. Chen R., Im H., Snyder M. (2015) Whole-exome enrichment with the Agilent SureSelect Human All Exon platform. *Cold Spring Harb. Protoc.* **7**, 626–633.
 16. Chen R., Im H., Snyder M. (2015) Whole-exome enrichment with the Illumina TruSeq exome enrichment platform. *Cold Spring Harb. Protoc.* **7**, 642–648.
 17. Chen R., Im H., Snyder M. (2015) Whole-exome enrichment with the roche NimbleGen SeqCap EZ exome library SR platform. *Cold Spring Harb. Protoc.* **7**, 634–641.
 18. Caruccio N. (2011) Preparation of next-generation sequencing libraries using Nextera™ technology: simultaneous DNA fragmentation and adaptor tagging by *in vitro* transposition. *Methods Mol. Biol.* **733**, 241–255.
 19. Marine R., Polson S.W., Ravel J., Hatfull G., Russell D., Sullivan M., Syed F., Dumas M., Wommack K.E. (2011) Evaluation of a transposase protocol for rapid generation of shotgun high-throughput sequencing libraries from nanogram quantities of DNA. *Appl. Environ. Microbiol.* **77**(22), 8071–8079.
 20. Shearer A.E., DeLuca A.P., Hildebrand M.S., Taylor K.R., Gurrola J., Scherer S., Scheetz T.E., Smith R.J.H. (2010) Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. *Proc. Natl. Acad. Sci. USA*. **107**, 21104–21109.
 21. Martignetti J.A., Tian L., Li D., Ramirez M.C.M., Camacho-Vanegas O., Camacho S.C., Guo Y., Zand D.J., Bernstein A.M., Masur S.K., Kim C.E., Otieno F.G., Hou C., Abdel-Magid N., Tweddle B., Metry D., Fournet J.C., Papp E., McPherson E.V., Zabel C., Vaksmann G., Morisot C., Keating B., Sleiman P.M., Cleveland J.A., Everman D.B., Zackai E., Hakonarson H. (2013) Mutations in *PDGFRB* cause autosomal-dominant infantile myofibromatosis. *Am. J. Hum. Genet.* **92**, 1001–1007.
 22. Nectoux J., de Cid R., Baulande S., Leturcq F., Urtizbera J.A., Penisson-Besnier I., Nadaj-Pakleza A., Roudaut C., Criqui A., Orhant L., Peyroulan D., Yaou R.B., Nelson I., Cobo A.M., Arné-Bes M.C., Uro-Coste E., Nitschke P., Claustres M., Bonne G., Lévy N., Chelly J., Richard I., Cossée M. (2015) Detection of *TRIM32* deletions in LGMD patients analyzed by a combined strategy of CGH array and massively parallel sequencing. *Eur. J. Hum. Genet.* **23**, 929–934.
 23. Rousseau-Nepton I., Okubo M., Grabs R., Mitchell J., Polychronakos C., Rodd C. (2015) A founder *AGL* mutation causing glycogen storage disease type IIIa in Inuit identified through whole-exome sequencing: a case series. *Can. Med. Assoc. J.* **187**, E68–E73.
 24. Poultney C.S., Goldberg A.P., Drapeau E., Kou Y., Harony-Nicolas H., Kajiwaraya Y., De Rubeis S., Durand S., Stevens C., Rehnström K., Palotie A., Daly M.J., Ma'ayan A., Fromer M., Buxbaum J.D. (2013) Identification of small exonic *CNV* from whole-exome sequence data and application to autism spectrum disorder. *Am. J. Hum. Genet.* **93**, 607–619.
 25. Guipponi M., Santoni F.A., Setola V., Gehrig C., Rotharmel M., Cuenca M., Guillin O., Dikeos D., Georgantopoulos G., Papadimitriou G., Curtis L., Méary A., Schürhoff F., Jamain S., Avramopoulos D.,

- Leboyer M., Rujescu D., Pulver A., Campion D., Siderovski D.P., Antonarakis S.E. (2014) Exome sequencing in 53 sporadic cases of schizophrenia identifies 18 putative candidate genes. *PLoS One*. **9**, e112745.
26. Griesi-Oliveira K., Acab A., Gupta A.R., Sunaga D.Y., Chailangkarn T., Nicol X., Nunez Y., Walker M.F., Murdoch J.D., Sanders S.J., Fernandez T.V., Ji W., Lifton R.P., Vadasz E., Dietrich A., Pradhan D., Song H., Ming G.L., Gu X., Haddad G., Marchetto M.C.N., Spitzer N., Passos-Bueno M.R., State M.W., Muotri A.R. (2015) Modeling non-syndromic autism and the impact of *TRPC6* disruption in human neurons. *Mol. Psychiatry*. **20**, 1350–1365.
 27. Pérez-Serra A., Toro R., Campuzano O., Sarquella-Brugada G., Berne P., Iglesias A., Mangas A., Brugada J., Brugada R. (2015) A novel mutation in lamin a/c causing familial dilated cardiomyopathy associated with sudden cardiac death. *J. Card. Fail.* **21**, 217–225.
 28. Calvo S.E., Compton A.G., Hershman S.G., Lim S.C., Lieber D.S., Tucker E.J., Laskowski A., Garone C., Liu S., Jaffe D.B., Christodoulou J., Fletcher J.M., Bruno D.L., Goldblatt J., Dimauro S., Thorburn D.R., Mootha V.K. (2012) Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. *Sci. Transl. Med.* **4**, 118ra10.
 29. Gai X., Ghezzi D., Johnson M.A., Biagosch C.A., Shamseldin H.E., Haack T.B., Reyes A., Tsukikawa M., Sheldon C.A., Srinivasan S., Gorza M., Kremer L.S., Wieland T., Strom T.M., Polyak E., Place E., Consugar M., Ostrovsky J., Vidoni S., Robinson A.J., Wong L.-J., Sondheimer N., Salih M.A., Al-Jishi E., Raab C.P., Bean C., Furlan F., Parini R., Lamperti C., Mayr J.A., Konstantopoulou V., Huemer M., Pierce E.A., Meitinger T., Freisinger P., Sperl W., Prokisch H., Alkuraya F.S., Falk M.J., Zeviani M. (2013) Mutations in *FBXL4*, encoding a mitochondrial protein, cause early-onset mitochondrial encephalomyopathy. *Am. J. Hum. Genet.* **93**, 482–495.
 30. Sikkema-Raddatz B., Johansson L.F., de Boer E.N., Almomani R., Boven L.G., van den Berg M.P., van Spaendonck-Zwarts K.Y., van Tintelen J.P., Sijmons R.H., Jongbloed J.D.H., Sinke R.J. (2013) Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum. Mutat.* **34**, 1035–1042.
 31. Drilon A., Wang L., Arcila M.E., Balasubramanian S., Greenbowe J.R., Ross J.S., Stephens P., Lipson D., Miller V.A., Kris M.G., Ladanyi M., Rizvi N.A. (2015) Broad, hybrid capture-based next-generation sequencing identifies actionable genomic alterations in “driver-negative” lung adenocarcinomas. *Clin. Cancer Res.* **21**, 3631–3639.
 32. Xie J., Lu X., Wu X., Lin X., Zhang C., Huang X., Chang Z., Wang X., Wen C., Tang X., Shi M., Zhan Q., Chen H., Deng X., Peng C., Li H., Fang Y., Shao Y., Shen B. (2016) Capture-based next-generation sequencing reveals multiple actionable mutations in cancer patients failed in traditional testing. *Mol. Genet. Genomic Med.* **4**, 262–272.
 33. Rozenblum A.B., Ilouze M., Dudnik E., Dvir A., Soussan-Gutman L., Geva S., Peled N. (2017) Clinical impact of hybrid capture-based next-generation sequencing on changes in treatment decisions in lung cancer. *J. Thorac. Oncol.* **12**, 258–268.
 34. Xu M.-D., Liu S.-L., Feng Y.-Z., Liu, Q., Shen M., Zhi Q., Liu Z., Gu D.M., Yu J., Shou L.-M., Gong F.-R., Zhu Q., Duan W., Chen K., Zhang J., Wu M.-Y., Tao M., Li W. (2017) Genomic characteristics of pancreatic squamous cell carcinoma, an investigation by using high throughput sequencing after in-solution hybrid capture. *Oncotarget*. **8**, 14620–14635.
 35. Clark T.A., Chung J.H., Kennedy M., Hughes J.D., Chennagiri N., Lieber D.S., Fendler B., Young L., Zhao M., Coyne M., Breese V., Young G., Donahue A., Pavlick D., Tsiros A., Brennan T., Zhong S., Mughal T., Bailey M., He J., Roels S., Frampton G.M., Sporerke J.M., Gendreau S., Lackner M., Schleifman E., Peters E., Ross J.S., Ali S.M., Miller V.A., Gregg J.P., Stephens P.J., Welsh A., Otto G.A., Lipson D. (2018) Analytical validation of a hybrid capture-based next-generation sequencing clinical assay for genomic profiling of cell-free circulating tumor DNA. *J. Mol. Diagn.* **20**, 686–702.
 36. Schrock A.B., Pavlick D., Klempner S.J., Chung J.H., Forcier B., Welsh, A., Young L., Leyland-Jones B., Bordoni R., Carvajal R.D., Chao J., Kurzrock R., Sicklick J.K., Ross J.S., Stephens P.J., Devoe C., Braiteh F., Ali S.A., Miller V.A. (2018) Hybrid capture-based genomic profiling of circulating tumor DNA from patients with advanced cancers of the gastrointestinal tract or anus. *Clin. Cancer Res.* **24**, 1881–1890.
 37. Briggs A.W., Good J.M., Green R.E., Krause J., Maricic T., Stenzel U., Lalueza-Fox C., Rudan P., Brajkovic D., Kucan Z., Gusic I., Schmitz R., Doronichev V.B., Golovanova L.V., de la Rasilla M., Fortea J., Rosas A., Pääbo S. (2009) Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science*. **325**, 318–321.
 38. Krause J., Fu Q., Good J.M., Viola B., Shunkov M.V., Derevianko A.P., Pääbo S. (2010) The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature*. **464**, 894–897.
 39. Enk J., Rouillard J.-M., Poinar H. (2013) Quantitative PCR as a predictor of aligned ancient DNA read counts following targeted enrichment. *Biotechniques*. **55**, 300–309.
 40. Kihana M., Mizuno F., Sawafuji R., Wang L., Ueda S. (2013) Emulsion PCR-coupled target enrichment: an effective fishing method for high-throughput sequencing of poorly preserved ancient DNA. *Gene*. **528**, 347–351.
 41. Templeton J.E.L., Brotherton P.M., Llamas B., Soubrier J., Haak W., Cooper A., Austin J.J. (2013) DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investig. Genet.* **4**, 26.
 42. Eduardoff M., Xavier C., Strobl C., Casas-Vargas A., Parson W. (2017) Optimized mtDNA control region primer extension capture analysis for forensically relevant samples and highly compromised mtDNA of different age and origin. *Genes*. **8**, E237.

43. Loreille O., Ratnayake S., Bazinet A.L., Stockwell T.B., Sommer D.D., Rohland N., Mallick S., Johnson P.L.F., Skoglund P., Onorato A.J., Bergman N.H., Reich D., Irwin J.A. (2018) Biological sexing of a 4000-year-old Egyptian mummy head to assess the potential of nuclear DNA recovery from the most damaged and limited forensic specimens. *Genes*. **9**, E135.
44. Fu Q., Meyer M., Gao X., Stenzel U., Burbano H. A., Kelso J., Pääbo S. (2013) DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci. USA*. **110**, 2223–2227.
45. Cruz Dávalos D.I., Nieves-Colón M.A., Sockell A., Poznik D.G., Schroeder H., Stone A.C., Bustamante C.D., Malaspina A.-S., Ávila-Arcos M.C. (2017) In-solution Y-chromosome capture-enrichment on ancient DNA libraries. *BMC Genomics*. **19**, 608.
46. Carpenter M.L., Buenrostro J.D., Valdiosera C., Schroeder H., Allentoft M.E., Sikora M., Rasmussen M., Gravel S., Guillén S., Nekhrizov G., Leshtakov K., Dimitrova D., Theodossiev N., Pettener D., Luiselli D., Sandoval K., Moreno-Estrada A., Li Y., Wang J., Gilbert M.T.P., Willerslev E., Greenleaf W.J., Bustamante C.D. (2013) Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *Am. J. Hum. Genet.* **93**, 852–864.
47. Enk J.M., Devault A.M., Kuch M., Murgha Y.E., Rouillard J.-M., Poinar H.N. (2014) Ancient whole genome enrichment using baits built from modern DNA. *Mol. Biol. Evol.* **31**, 1292–1294.
48. Lindo J., Achilli A., Perego U.A., Archer D., Valdiosera C., Petzelt B., Mitchell J., Worl R., Dixon E.J., Fifield T.E., Rasmussen M., Willerslev E., Cybulski J.S., Kemp B.M., DeGiorgio M., Malhi R.S. (2017) Ancient individuals from the North American Northwest Coast reveal 10000 years of regional genetic continuity. *Proc. Natl. Acad. Sci. USA*. **114**, 4093–4098.
49. Bi K., Vanderpool D., Singhal S., Linderth T., Moritz C., Good J.M. (2012) Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*. **13**, 403.
50. Cosart T., Beja-Pereira A., Chen S., Ng S.B., Shendure J., Luikart G. (2011) Exome-wide DNA capture and next generation sequencing in domestic and wild species. *BMC Genomics*. **12**, 347.
51. Perry G.H., Marioni J.C., Melsted P., Gilad Y. (2010) Genomic-scale capture and sequencing of endogenous DNA from feces. *Mol. Ecol.* **19**(24), 5332–5344.
52. Tewhey R., Nakano M., Wang X., Pabón-Peña C., Novak B., Giuffre A., Lin E., Happe S., Roberts D.N., LeProust E.M., Topol E.J., Harismendy O., Frazer K.A. (2009) Enrichment of sequencing targets from the human genome by solution hybridization. *Genome Biol.* **10**, R116.
53. Brownstein Z., Friedman L.M., Shahin H., Oron-Karni V., Kol N., Rayyan A.A., Parzefall T., Lev D., Shalev S., Frydman M., Davidov B., Shohat M., Rahile M., Lieberman S., Levy-Lahad E., Lee M.K., Shomron N., King M.-C., Walsh T., Kanaan M., Avraham K.B. (2011) Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in Middle Eastern families. *Genome Biol.* **12**, R89.
54. Miyazato P., Katsuya H., Fukuda A., Uchiyama Y., Matsuo M., Tokunaga M., Hino S., Nakao M., Satou Y. (2016) Application of targeted enrichment to next-generation sequencing of retroviruses integrated into the host human genome. *Sci. Rep.* **6**, 28324.
55. Hodges E., Rooks M., Xuan Z., Bhattacharjee A., Gordon D.B., Brizuela L., McCombie W.R., Hanon G.J. (2009) Hybrid selection of discrete genomic intervals on custom designed microarrays for massively parallel sequencing. *Nat. Protocols*. **4**, 960–974.
56. Nilsson M., Malmgren H., Samiotaki M., Kwiatkowski M., Chowdhary B.P., Landegren U. (1994) Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science*. **265**, 2085–2088.
57. Akhras M.S., Unemo M., Thiyagarajan S., Nyren P., Davis R.W., Fire A.Z., Pourmand N. (2007) Connector inversion probe technology: a powerful one-primer multiplex DNA amplification system for numerous scientific applications. *PLoS One*. **2**(9), e915.
58. Pogoda M., Hilke F.-J., Lohmann E., Sturm M., Lenz F., Matthes J., Muyas F., Ossowski S., Hoischen A., Faust U., Sepahi I., Casadei N., Poths S., Riess O., Schroeder C., Grundmann K. (2019) Single molecule molecular inversion probes for high throughput germline screenings in dystonia. *Front. Neurol.* **10**, 1332.
59. Bekers E.M., Eijkelenboom A., Rombout P., van Zwam P., Mol S., Ruijter E., Scheijen B., Flucke U. (2019) Identification of novel GNAS mutations in intramuscular myxoma using next-generation sequencing with single-molecule tagged molecular inversion probes. *Diagn. Pathol.* **14**, 15.
60. Lenting K., van den Heuvel C.N.A.M., van Ewijk A., ElMelik D., de Boer R., Tindall E., Wei G., Kusters B., Te Dorsthorst M., Ter Laan M., Huynen M.A., Leenders W.P. (2019) Mapping actionable pathways and mutations in brain tumours using targeted RNA next generation sequencing. *Acta Neuropathol. Commun.* **7**, 185.
61. van den Heuvel C.N.A.M., van Ewijk A., Zeelen C., de Bitter T., Huynen M., Mulders P., Oosterwijk E., Leenders W.P.J. (2019) Molecular profiling of druggable targets in clear cell renal cell carcinoma through targeted RNA sequencing. *Front. Oncol.* **9**, 117.
62. Stefan C.P., Hall A.T., Minogue T.D. (2018) Detection of 16S rRNA and KPC genes from complex matrix utilizing a molecular inversion probe assay for next-generation sequencing. *Sci. Rep.* **8**(1), 2028.
63. Neveling K., Mensenkamp A.R., Derks R., Kwint M., Ouchene H., Steehouwer M., van Lier B., Bosgoed E., Rikken A., Tychon M., Zafeiropoulou D., Castelein S., Hehir-Kwa J., Thung D.T., Hofste T., Lelieveld S.H., Bertens S.M.M., Adan I.B.J.F., Eijkelenboom A., Tops B.B., Yntema H., Stokowy T., Knappskog P.M., Hoberg-Vetti H., Steen V.M., Boyle E., Martin B., Ligtenberg M.J.L., Shendure J., Nelen M.R., Hoischen A. (2017) BRCA testing by single-molecule molecular inversion probes. *Clin. Chem.* **63**(2), 503–512.

64. Lange I.M., Koudijs M.J., van 't Slot R., Sonsma A.C.M., Mulder F., Carbo E.C., van Kempen M.J.A., Nijman I.J., Ernst R.F., Savelberg S.M.C., Knoers N.V.A.M., Brilstra E.H., Koelman B.P.C. (2018) Assessment of parental mosaicism in *SCN1A*-related epilepsy by single-molecule molecular inversion probes and next-generation sequencing. *J. Med. Genet.* **56**(2), 75–80.
65. Mirzaa G., Timms A.E., Conti V., Boyle E.A., Girissha K.M., Martin B., Kircher M., Olds C., Juusola J., Collins S., Park K., Carter M., Glass I., Krogeloh-Mann I., Chitayat D., Parikh A.S., Bradshaw R., Torti E., Braddock S., Burke L., Ghedia S., Stephan M., Stewart F., Prasad C., Napier M., Saitta S., Straussberg R., Gabbett M., O'Connor B.C., Keegan C.E., Yin L.J., Lai A.H.M., Martin N., McKinnon M., Addor M.-C., Boccuto L., Schwartz C.E., Lanoel A., Conway R.L., Devriendt K., Tatton-Brown K., Pierpont M.E., Painter M., Worgan L., Reggin J., Hennekam R., Tsuchiya K., Pritchard C.C., Aracena M., Gripp K.W., Cordisco M., Esch H.V., Garavelli L., Curry C., Goriely A., Kayserilli H., Shendure J., Graham J.R., Guerrini R., Dobyns W.B. (2016) *PIK3CA*-associated developmental disorders exhibit distinct classes of mutations with variable expression and tissue distribution. *JCI Insight.* **1**(9), e87623.
66. Wada Y., Maekawa M., Ohnishi T., Balan S., Matsuo S., Iwamoto K., Iwayama Y., Ohba H., Watanabe A., Hisano Y., Nozaki Y., Toyota T., Shimogori T., Itokawa M., Kobayashi T., Yoshikawa T. (2020) Peroxisome proliferator-activated receptor α as a novel therapeutic target for schizophrenia. *EBioMedicine.* **62**, 103130.
67. Sleczkowska M., Almomani R., Marchi M., de Greef B.T.A., Sopacua M., Hoeijmakers J.G.J., Lindsey P., Salvi E., Bönhof G.J., Ziegler D., Malik R.A., Waxman S.G., Lauria G., Faber C.G., Smeets H.J.M., Gerrits M.M. (2022) Peripheral ion channel gene screening in painful- and painless-diabetic neuropathy. *Int. J. Mol. Sci.* **23**, 7190.
68. Almomani R., Marchi M., Sopacua M., Lindsey P., Salvi E., de Koning B., Santoro S., Magri S., Smeets H.J.M., Martinelli B.F., Malik R.R., Ziegler D., Hoeijmakers J.G.J., Bönhof G., Dib-Hajj S., Waxman S.G., Merckies I.S.J., Lauria G., Faber C.G., Gerrits M.M. (2020) Evaluation of molecular inversion probe versus TruSeq® custom methods for targeted next-generation sequencing. *PLoS One.* **15**(9), e0238467.
69. Suzuki O., Dong O.M., Howard R.M., Wiltshire T. (2019) Characterizing the pharmacogenome using molecular inversion probes for targeted next-generation sequencing. *Pharmacogenomics.* **20**(14), 1005–1020.
70. Yoon J.-K., Ahn J., Kim H.S., Han S.M., Jang H., Lee M.G., Lee J.H., Bang D. (2015) microDuMIP: target-enrichment technique for microarray-based duplex molecular inversion probes. *Nucleic Acids Res.* **43**(5), e28.
71. Saiki R.K., Scharf S., Faloona F., Mullis K.B., Horn G.T., Erlich H.A., Arnheim N. (1985) Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science.* **230**, 1350–1354.
72. Halbritter J., Diaz K., Chaki M., Porath J.D., Tarrier B., Fu C., Innis J.L., Allen S.J., Lyons R.H., Stefanidis C.J., Omran H., Soliman N.A., Otto E.A. (2012) High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing. *J. Med. Genet.* **49**(12), 756–767.
73. Tewhey R., Warner J.B., Nakano M., Libby B., Medkova M., David P.H., Kotsopoulos S.K., Samuels M.L., Hutchison J.B., Larson J.W., Topol E.J., Weiner M.P., Harismendy O., Olson J., Link D.R., Frazer K.A. (2009) Microdroplet-based PCR enrichment for large-scale targeted sequencing. *Nat. Biotechnol.* **27**, 1025–1031.
74. Vollbrecht C., Werner R., Walter R.F.H., Christoph D.C., Heukamp L.C., Peifer M., Hirsch B., Burbat L., Mairinger T., Schmid K.W., Wohlschlaeger J., Mairinger F.D. (2015) Mutational analysis of pulmonary tumours with neuroendocrine features using targeted massive parallel sequencing: a comparison of a neglected tumour group. *Br. J. Cancer.* **113**, 1704–1711.
75. Lee A., Lee S.-H., Jung C.K., Park G., Lee K.Y., Choi H.J., Min K.O., Kim T.J., Lee E.J., Lee Y.S. (2018) Use of the Ion AmpliSeq Cancer Hotspot Panel in clinical molecular pathology laboratories for analysis of solid tumours: with emphasis on validation with relevant single molecular pathology tests and the OncoPrint Focus Assay. *Pathol. Res. Pract.* **214**(5), 713–719.
76. Mustafa A.E., Faquih T., Baz B., Kattan R., Al-Issa A., Tahir A.I., Imtiaz F., Ramzan K., Al-Sayed M., Alowain M., Al-Hassan Z., Al-Zaidan H., Abouelhoda M., Al-Mubarak B.R., Al Tassan N.A. (2018) Validation of ion Torrent™ inherited disease panel with the PGM™ sequencing platform for rapid and comprehensive mutation detection. *Genes (Basel).* **9**(5), 267.
77. Cao C., Lou I., Huang C., Lee M.-Y. (2016) Metagenomic sequencing of activated sludge filamentous bacteria community using the Ion Torrent platform. *Desalin. Water Treat.* **57**(5), 2175–2183.
78. Jiang F., Doudna J.A. (2017) CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **46**, 505–529.
79. Schultzhause Z., Wang Z., Stenger D. (2021) CRISPR-based enrichment strategies for targeted sequencing. *Biotechnol. Adv.* **46**, 107672.
80. Gu W., Crawford E.D., O'Donovan B.D., Wilson M.R., Chow E.D., Retallack H., DeRisi J.L. (2016) Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* **17**(1), 41.
81. Lee S.H., Yu J., Hwang G.H., Kim S., Kim H.S., Ye S., Kim K., Park J., Park D.Y., Cho Y.K., Kim J.S., Bae S. (2017) CUT-PCR: CRISPR-mediated, ultrasensitive detection of target DNA using PCR. *Oncogene.* **36**(49), 6823–6829.
82. Stevens R.C., Steele J.L., Glover W.R., Sanchez-Garcia J.F., Simpson S.D., O'Rourke D., Ramsdell J.S., MacManes M.D., Thomas W.K., Shuber A.P. (2019)

- A novel CRISPR/Cas9 associated technology for sequence-specific nucleic acid enrichment. *PLoS One*. **14**(4), e0215441.
83. Bennett-Baker P.E., Mueller J.L. (2017) CRISPR-mediated isolation of specific megabase segments of genomic DNA. *Nucleic Acids Res.* **45**(19), e165.
 84. Gabrieli T., Sharim H., Fridman D., Arbib N., Michaëli Y., Ebenstein Y. (2018) Selective nanopore sequencing of human *BRCA1* by Cas9-assisted targeting of chromosome segments (CATCH). *Nucleic Acids Res.* **46**(14), e87.
 85. Nachmanson D., Lian S., Schmidt E.K., Hipp M.J., Baker K.T., Zhang Y., Tretiakova M., Loubet-Seneac K., Kohrn B.F., Salk J.J., Kennedy S.R., Risques R.A. (2018) Targeted genome fragmentation with CRISPR/Cas9 enables fast and efficient enrichment of small genomic regions and ultra-accurate sequencing with low DNA input (CRISPR-DS). *Genome Res.* **28**(10), 1589–1599.
 86. Lee J., Lim H., Jang H., Hwang B., Lee J.H., Cho J., Lee J.H., Bang D. (2019) CRISPR-Cap: multiplexed double-stranded DNA enrichment based on the CRISPR system. *Nucleic Acids Res.* **47**(1), e1.
 87. Peter B.J., Ach R.A. (2014) US Patent No. US 20140356867 A1. patents.google.com/patent/US20140356867A1/en?q=US+20140356867+A1.
 88. Carpenter M.L., Bustamante C.D. (2018) US Patent No. US 20180298421 A1. patents.google.com/patent/US20180298421A1/en?q=US+20180298421
 89. Bang D., Ji Won Lee J.W., Lim H.S. (2016) US Patent No. US 20160244829 A1. patents.google.com/patent/US20160244829A1/en?q=US20160244829A1
 90. Xu X., Luo T., Gao J., Lin N., Li W., Xia X., Wang J. (2020) CRISPR-assisted DNA detection: a novel dCas9-based DNA detection technique. *CRISPR J.* **3**(6), 487–502.
 91. Aalipour A., Dudley J.C., Park S.M., Murty S., Chabon J.J., Boyle E.A., Diehn M., Gambhir S.S. (2018) Deactivated CRISPR associated protein 9 for minor-allele enrichment in cell-free DNA. *Clin. Chem.* **64**(2), 307–316.
 92. Quan J., Langelier C., Kuchta A., Batson J., Teyssier N., Lyden A., Caldera S., McGeever A., Dimitrov B., King R., Wilhelm J., Murphy M., Ares L.P., Trivisano K.A., Sit R., Amato R., Mumbengegwi D.R., Smith J.L., Bennett A., Gosling R., Mourani P.M., Calfee C.S., Neff N.F., Chow E.D., Kim P.S., Greenhouse B., DeRisi J.L., Crawford E.D. (2019) FLASH: a next-generation CRISPR diagnostic for multiplexed detection of antimicrobial resistance sequences. *Nucleic Acids Res.* **47**(14), e83.
 93. Hafford-Tear N.J., Tsai Y.-C., Sadan A.N., Sanchez-Pintado B., Zarouchlioti C., Maher G.J., Liskova P., Tuft S.J., Hardcastle A.J., Clark T.A., Davidson A.E. (2019) CRISPR/Cas9-targeted enrichment and long-read sequencing of the Fuchs endothelial corneal dystrophy-associated *TCF4* triplet repeat. *Genet. Med.* **21**(9), 2092–2102.
 94. Stangl C., de Blank S., Renkens I., Westera L., Verbeek T., Valle-Inclan J.E., González R.C., Henssen A.G., van Roosmalen M.J., Stam R.W., Voest E.E., Kloosterman W.P., van Haaften G., Monroe G.R. (2020) Partner independent fusion gene detection by multiplexed CRISPR-Cas9 enrichment and long read nanopore sequencing. *Nat. Commun.* **11**(1), 2861.
 95. Watson C.M., Crinnion L.A., Hewitt S., Bates J., Robinson R., Carr I.M., Sheridan E., Adlard J., Bonthron D.T. (2019) Cas9-based enrichment and single-molecule sequencing for precise characterization of genomic duplications. *Lab. Inv.* **100**(1), 135–146.
 96. Gilpatrick T., Lee I., Graham J.E., Raimondeau E., Bowen R., Heron A., Sedlazeck F.J., Timp W. (2020) Targeted Nanopore Sequencing with Cas9 for studies of methylation, structural variants, and mutations. *Nat. Biotechnol.* **38**(4), 433–438.
 97. Kuhlensäuer G., Hullmann J., Appenzeller S. (2011) Novel genomic techniques open new avenues in the analysis of monogenic disorders. *Hum. Mutat.* **32**(2), 144–151.
 98. Kiezun A., Garimella K., Do R., Stitzel N.O., Neale B.M., McLaren P.J., Gupta N., Sklar P., Sullivan P.F., Morn J.L., Hultman C.M., Lichtenstein P., Magnusson P., Lehner T., Shugart Y.Y., Price A.L., de Bakker P.I., Purcell S.M., Sunyaev S.R. (2012) Exome sequencing and the genetic basis of complex traits. *Nat. Genet.* **44**(6), 623–630.
 99. Vogelstein B., Papadopoulos N., Velculescu V.E., Zhou S., Diaz L.A.J., Kinzler K.W. (2013) Cancer genome landscapes. *Science*. **339**(6127), 1546–1558.
 100. Gaudin M., Desnues C. (2018) Hybrid capture-based next generation sequencing and its application to human infectious diseases. *Front. Microbiol.* **9**, 2924.
 101. Melnikov A., Galinsky K., Rogov P., Fennell T., Van Tyne D., Russ C., Daniels R., Barnes K.G., Bochicchio J., Ndiaye D., Sene P.D., Wirth D.F., Nusbaum C., Volkman S.K., Birren B.W., Gnirke A., Neafsey D.E. (2011) Hybrid selection for sequencing pathogen genomes from clinical samples. *Genome Biol.* **12**, R73.
 102. Smith M., Campino S., Gu Y., Clark T.G., Otto T.D., Maslen G., Manske M., Imwong M., Dondorp A.M., Kwiatkowski D.P., Quail M.A., Swerdlow H. (2012). An in-solution hybridisation method for the isolation of pathogen DNA from human DNA-rich clinical samples for analysis by NGS. *Open Genom. J.* **5**, 18–29.
 103. Bright A.T., Tewhey R., Abeles S., Chuquiyauri R., Llanos-Cuentas A., Ferreira M.U., Schork N.J., Vinetz J.M., Winzeler E.A. (2012) Whole genome sequencing analysis of *Plasmodium vivax* using whole genome capture. *BMC Genomics.* **13**, 262.
 104. Amorim-Vaz S., Tran V.D.T., Pradervand S., Pagni M., Coste A.T., Sanglard D. (2015) RNA enrichment method for quantitative transcriptional analysis of pathogens *in vivo* applied to the fungus *Candida albicans*. *mBio.* **6**(5), e00942-15.
 105. Brown A.C., Bryant J.M., Einer-Jensen K., Holdstock J., Houniet D.T., Chan J.Z.M., Depledge D.P., Nikolayevskyy V., Broda A., Stone M.J., Christiansen M.T., Williams R., McAndrew M.B., Tutill H., Brown J., Melzer M., Rosmarin C., McHugh T.D., Shorten R.J., Drobniowski F., Speight G., Breuer J. (2015) Rapid whole-genome sequencing of *Mycobac-*

- terium tuberculosis* isolates directly from clinical samples. *J. Clin. Microbiol.* **53**(7), 2230–2237.
106. Christiansen M.T., Brown A.C., Kundu S., Tutill H.J., Williams R., Brown J.R., Holdstock J., Holland M.J., Stevenson S., Dave J., Tong C.Y., Einer-Jensen K., Depledge D.P., Breuer J. (2014) Whole-genome enrichment and sequencing of *Chlamydia trachomatis* directly from clinical samples. *BMC Infect. Dis.* **14**, 591.
 107. Clark S.A., Doyle R., Lucidarme J., Borrow R., Breuer J. (2018) Targeted DNA enrichment and whole genome sequencing of *Neisseria meningitidis* directly from clinical specimens. *Int. J. Med. Microbiol.* **308**, 256–262.
 108. Houldcroft C.J., Beale M.A., Breuer J. (2017) Clinical and biological insights from viral genome sequencing. *Nat. Rev. Microbiol.* **15**, 183–192.
 109. Depledge D.P., Palser A.L., Watson S.J., Lai I.Y.-C., Gray E.R., Grant P., Kanda R.K., Leproust E., Kellam P., Breuer J. (2011) Specific capture and whole-genome sequencing of viruses from clinical samples. *PLoS One.* **6**, e27805.
 110. Kwok H., Wu C.W., Palser A.L., Kellam P., Sham P.C., Kwong D.L.W., Chiang A.K.S. (2014) Genomic diversity of Epstein–Barr virus genomes isolated from primary nasopharyngeal carcinoma biopsy samples. *J. Virol.* **88**, 10662–10672.
 111. Metsky H.C., Matranga C.B., Wohl S., Schaffner S.F., Freije C.A., Winnicki S.M., West K., Qu J., Baniecki M.L., Gladden-Young A., Lin A.E., Tomkins-Tinch C.H., Ye S.H., Park D.J., Luo C.Y., Barnes K.G., Shah R.R., Chak B., Barbosa-Lima G., Delatorre E., Vieira Y.R., Paul L.M., Tan A.L., Barcellona C.M., Porcelli M.C., Vasquez C., Cannons A.C., Cone M.R., Hogan K.N., Kopp E.W., Anzinger J.J., Garcia K.F., Parham L.A., Ramírez R.M.G., Montoya M.C.M., Rojas D.P., Brown C.M., Hennigan S., Sabina B., Scotland S., Gangavarapu K., Grubaugh N.D., Oliveira G., Robles-Sikisaka R., Rambaut A., Gehrke L., Smole S., Halloran M.E., Villar L., Mattar S., Lorenzana I., Cerbino-Neto J., Valim C., Degraeve W., Bozza P.T., Gnirke A., Andersen K.G., Isern S., Michael S.F., Bozza F.A., Souza T.M.L., Bosch I., Yozwiak N.L., MacInnis B.L., Sabeti P.C. (2017) Zika virus evolution and spread in the Americas. *Nature.* **546**, 411–415.
 112. Brown J.R., Roy S., Ruis C., Yara Romero E., Shah D., Williams R., Breuer J. (2016) Norovirus whole-genome sequencing by sureselect target enrichment: a robust and sensitive method. *J. Clin. Microbiol.* **54**, 2530–2537.
 113. Matranga C.B., Andersen K.G., Winnicki S., Busby M., Gladden A.D., Tewhey, R., Stremlau M., Berlin A., Gire S.K., England E., Moses L.M., Mikkelsen T.S., Odiya I., Ehiane P.E., Folarin O., Goba A., Kahn S.H., Grant D.S., Honko A., Hensley L., Happi C., Garry R.F., Malboeuf C.M., Birren B.W., Gnirke A., Levin J.Z., Sabeti P.C. (2014) Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. *Genome Biol.* **15**, 519.
 114. Thomson E., Ip C.L.C., Badhan A., Christiansen M.T., Adamson W., Ansari M.A., Bibby D., Breuer J., Brown A., Bowden R., Bryant J., Bonsall D., Da Silva Filipe A., Hinds C., Hudson E., Klenerman P., Lythgow K., Mbisa J.L., McLauchlan J., Myers R., Piazza P., Roy S., Trebes A., Sreenu V.B., Witteveldt J., STOP-HCV Consortium, Barnes E., Simmonds P. (2016) Comparison of next generation sequencing technologies for the comprehensive assessment of full-length hepatitis C viral genomes. *J. Clin. Microbiol.* **54**, 2470–2484.
 115. Donaldson C.D., Clark D.A., Kidd I.M., Breuer J., Depledge D.D. (2013) Genome sequence of human herpesvirus 7 strain UCL-1. *Genome Announc.* **1**, e00830-13.
 116. Greninger A.L., Roychoudhury P., Xie H., Casto A., Cent A., Pepper G., Koelle D.M., Huang M.L., Wald A., Johnston C., Jerome K.R. (2018) Ultrasensitive capture of human herpes simplex virus genomes directly from clinical samples reveals extraordinarily limited evolution in cell culture. *mSphere.* **3**, e00283-18.
 117. Wylie T.N., Wylie K.M., Herter B.N., Storch G.A. (2015) Enhanced virome sequencing using targeted sequence capture. *Genome Res.* **25**, 1910–1920.
 118. Briese T., Kapoor A., Mishra N., Jain K., Kumar A., Jabado O., Lipkin W.I. (2015) Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *mBio.* **6**, e01491-15.
 119. Chalkias S., Gorham J.M., Mazaika E., Parfenov M., Dang X., DePalma S., McKean D., Seidman C.E., Seidman J.G., Koralknik I.J. (2018) ViroFind: a novel target-enrichment deep-sequencing platform reveals a complex JC virus population in the brain of PML patients. *PLoS One.* **13**, e0186945.
 120. Chilamakuri C.S.R., Lorenz S., Madoui M.-A., Vodák D., Sun J., Hovig E., Myklebost O., Meza-Zepeda L.A. (2014) Performance comparison of four exome capture systems for deep sequencing. *BMC Genomics.* **15**(1), 449.
 121. Clark M.J., Chen R., Lam H.Y.K., Karczewski K.J., Chen R., Euskirchen G., Butte A.J., Snyder M. (2011) Performance comparison of exome DNA sequencing technologies. *Nat. Biotechnol.* **29**(10), 908–914.
 122. Shigemizu D., Momozawa Y., Abe T., Morizono T., Borojevich K.A., Takata S., Ashikawa K., Kubo M., Tsunoda T. (2015) Performance comparison of four commercial human whole-exome capture platforms. *Sci. Rep.* **5**(1), 12742.
 123. Zhou J., Zhang M., Li X., Wang Z., Pan D., Shi Y. (2021) Performance comparison of four types of target enrichment baits for exome DNA sequencing. *Hereditas.* **158**, 10.
 124. Синяков А.Н., Бессмельцев В.П., Камаев Г.Н. (2018) Микрочиповый синтезатор ДНК. II Всероссийская научно-практическая конференция “Научное приборостроение – современное состояние и перспективы развития”. Казань 4–7 июня. Сб. материалов конференции. с. 280–282.
 125. Sinyakov A.N., Ryabinin V.A., Kostina E.V., Zaytsev D.E., Chukanov N.V., Kamayev G.N. (2021) Linkers for oligonucleotide microarray synthesis. *J. Saudi Chem. Soc.* **25**, 101382.

DNA Fragment Enrichment for High-Throughput Sequencing

A. N. Sinyakov¹, * and E. V. Kostina¹

¹*Institute of Chemical Biology and Fundamental Medicine, Siberian Branch, Russian Academy of Sciences,
Novosibirsk, 630090 Russia*

**e-mail: sinyakov@niboch.nsc.ru*

Application of oligonucleotides, mainly obtained using new generation DNA synthesizers (microarray DNA synthesizers), for the enrichment of targeted genomic fragments are described. Methods of molecular hybridization, polymerase chain reaction and CRISPR-based methods for targets enrichment are considered. Examples of the practical use of the developed methods for research and diagnostic purposes are given.

Keywords: high-throughput sequencing, oligonucleotides, microarray, amplification, diagnostics, exome, CRISPR-Cas