

 pismavavilov.ru

DOI 10.18699/LettersVJ2021-7-05

Методы

## Создание библиотек баркодированных плазмид с помощью метода клонирования по Гибсону

А.В. Смирнов<sup>1</sup>, А.М. Юнусова<sup>1</sup>, А.А. Муравьева<sup>2</sup>, Э.С. Валеев<sup>1</sup>, В.С. Фишман<sup>1,2</sup>, Н.Р. Баттулин<sup>1,2</sup> ✉

**Аннотация:** Развитие методов секвенирования ДНК нового поколения стало одним из основных двигателей прогресса в биологических исследованиях последних лет. Эти методы позволили активно применять ДНК-баркодирование для решения широкого спектра биологических задач. Ключевым этапом работ с баркодированными молекулами ДНК является создание библиотеки баркодированных плазмид. Данная статья посвящена описанию методики с использованием клонирования по Гибсону. Клонирование методом Гибсона позволяет создавать сложные генетические конструкции, состоящие из большого числа фрагментов. Кроме того, этот метод допускает использование олигонуклеотидов с вырожденными позициями, что удобно для разработки баркодированных библиотек. Применение данной методики позволяет получать библиотеки с высоким разнообразием (более 105) баркодированных молекул. В статье описаны основные этапы создания баркодированной библиотеки, в частности рассмотрены принципы дизайна последовательностей баркодированных олигонуклеотидов и особенности сборки библиотек методом Гибсона. Приведены условия очистки реакционной смеси на магнитных частицах и оптимальные параметры электропорации созданных конструкций в клетки бактерий. Описаны протоколы наращивания культур бактериальных клонов, выделения плазмидной ДНК и очистки ее от молекул вектора, не несущих вставки, с помощью дополнительного гидролиза эндонуклеазами рестрикции и обработки препарата плазмидной ДНК экзонуклеазой Plasmid-Safe. Важным этапом работы являются характеристика полученного препарата и оценка разнообразия баркодированных молекул в библиотеке. В статье также представлено описание подготовки библиотеки для последующего секвенирования на платформе Illumina без использования амплификации. Кроме того, приведен алгоритм биоинформатического анализа данных секвенирования для оценки разнообразия баркодированных молекул. Полученная таким способом библиотека может быть использована в исследованиях для анализа индивидуальных молекул ДНК. Представленная в настоящем исследовании библиотека будет применена для изучения рекомбинации в ранних эмбрионах мыши.

**Ключевые слова:** ДНК-баркодирование; плазмидная библиотека; клонирование по Гибсону.

**Благодарности:** Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-34-70087.

**Для цитирования:** Смирнов А.В., Юнусова А.М., Муравьева А.А., Валеев Э.С., Фишман В.С., Баттулин Н.Р. Создание библиотек баркодированных плазмид с помощью метода клонирования по Гибсону. *Письма в Вавиловский журнал генетики и селекции*. 2021;7(1):34-45. DOI 10.18699/LettersVJ2021-7-05

Methods

## Creation of barcoded plasmids libraries using the Gibson cloning method

A.V. Smirnov<sup>1</sup>, A.M. Yunusova<sup>1</sup>, A.A. Muravyova<sup>2</sup>, E.S. Valeev<sup>1</sup>, V.S. Fishman<sup>1,2</sup>, N.R. Battulin<sup>1,2</sup> ✉

**Abstract:** The development of next-gen DNA sequencing methods has become one of the main drivers of progress in biological research in recent years. These methods made it possible to actively use DNA barcoding to solve a wide range of biological problems. The key step for working with barcoded DNA molecules is the creation of a library of barcoded plasmids. This protocol is devoted to the description of such a technique using the Gibson cloning approach. Gibson cloning allows the creation of complex genetic constructs from a large number of fragments. The method allows the use of oligonucleotides with degenerate positions, which is convenient for creating barcoded libraries. The application of the described technique allows one to obtain libraries with high diversity (more than 105) barcoded molecules. The article describes the main stages of the process of creating a barcoded library. In particular, the principles of the design of barcoded oligonucleotides are discussed. We also feature technical details of assembling libraries using the Gibson method. A protocol for growing bacterial clones, isolation of plasmid DNA and its purification from empty vectors using treatment of


<sup>1</sup> Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

<sup>2</sup> Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

<sup>1</sup> Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

<sup>2</sup> Novosibirsk State University, Novosibirsk, Russia

 battulin@gmail.com

 © Смирнов А.В., Юнусова А.М., Муравьева А.А., Валеев Э.С., Фишман В.С., Баттулин Н.Р., 2021

the plasmid DNA preparation with plasmid-safe nuclease is described. An important stage in the work is the characterization of the obtained library and the assessment of the diversity of barcoded molecules in the library. We describe the preparation of the library for sequencing on the Illumina platform, and we present an algorithm for computational analysis of sequencing data to assess the real diversity of barcoded molecules. Library obtained by such method could be used in studying individual DNA molecules. Library from the current research will be used for studying recombination in early mouse embryos.

**Key words:** DNA barcoding; plasmid library; Gibson cloning.

**For citation:** Smirnov A.V., Yunusova A.M., Muravyova A.A., Valeev E.S., Fishman V.S., Battulin N.R. Creation of barcoded plasmids libraries using the Gibson cloning method. *Pisma v Vavilovskii Zhurnal Genetiki i Seleksii = Letters to Vavilov Journal of Genetics and Breeding*. 2021;7(1):34-45. DOI 10.18699/LettersVJ2021-7-05 (in Russian)

## Введение

Развитие методов секвенирования ДНК нового поколения привело к значительному прогрессу в биологических исследованиях последних лет. Рост доступности методов секвенирования не только позволил определить последовательности нуклеотидов в геномах человека, модельных и немодельных организмов, но и приступить к масштабным проектам чтения геномов, таким как The Darwin Tree of Life, нацеленным на прочтение геномов всех 60 тыс. эукариотических организмов, населяющих Великобританию и Ирландию. Однако помимо классического применения – для исследования существующих биологических молекул (геномы и транскриптомы организмов) – технологии секвенирования нового поколения можно использовать для узкоспециализированных задач. Например, с помощью антител, конъюгированных с ДНК-олигонуклеотидами, возможно количественно характеризовать антигенный профиль отдельных клеток (Arnold et al., 2018). В данном случае секвенирование олигонуклеотидов позволяет определять, какие из множества антител связались с клеткой. Такие олигонуклеотиды используют и в других экспериментах, когда необходимо детектировать отдельные молекулы. Искусственно синтезированная последовательность нуклеотидов ДНК служит уникальным идентификатором, который можно опознать после секвенирования. Эта последовательность и получила название ДНК-баркода (англ. barcode – штрихкод, маркировка).

Баркодированные молекулы позволяют решать широкий спектр биологических задач (Юнусова, Баттулин, 2016). В частности, с помощью мечения клеток ДНК-баркодированными лентивирусами показано, что клетки могут наследовать способность к репрограммированию в индуцированные плюрипотентные клетки (Yunusova et al., 2017). Благодаря баркодированным молекулам ДНК возможен многочисленный скрининг функциональных вариантов, влияющих на активность генов в массовых параллельных репортерных анализах (massively parallel reporter assays, MPRA) (Klein et al., 2020). Кроме того, недавно с помощью инъекций баркодированных молекул ДНК в пронуклеус зиготы мыши мы впервые показали чрезвычайно высокую активность гомологичной рекомбинации в ранних эмбрионах (Smirnov et al., 2020). Ключевым этапом работ с баркодированными молекулами ДНК является создание библиотеки баркодированных плазмид. Данная работа посвящена описанию такой методики с использованием клонирования по Гибсону (Gibson et al., 2009).

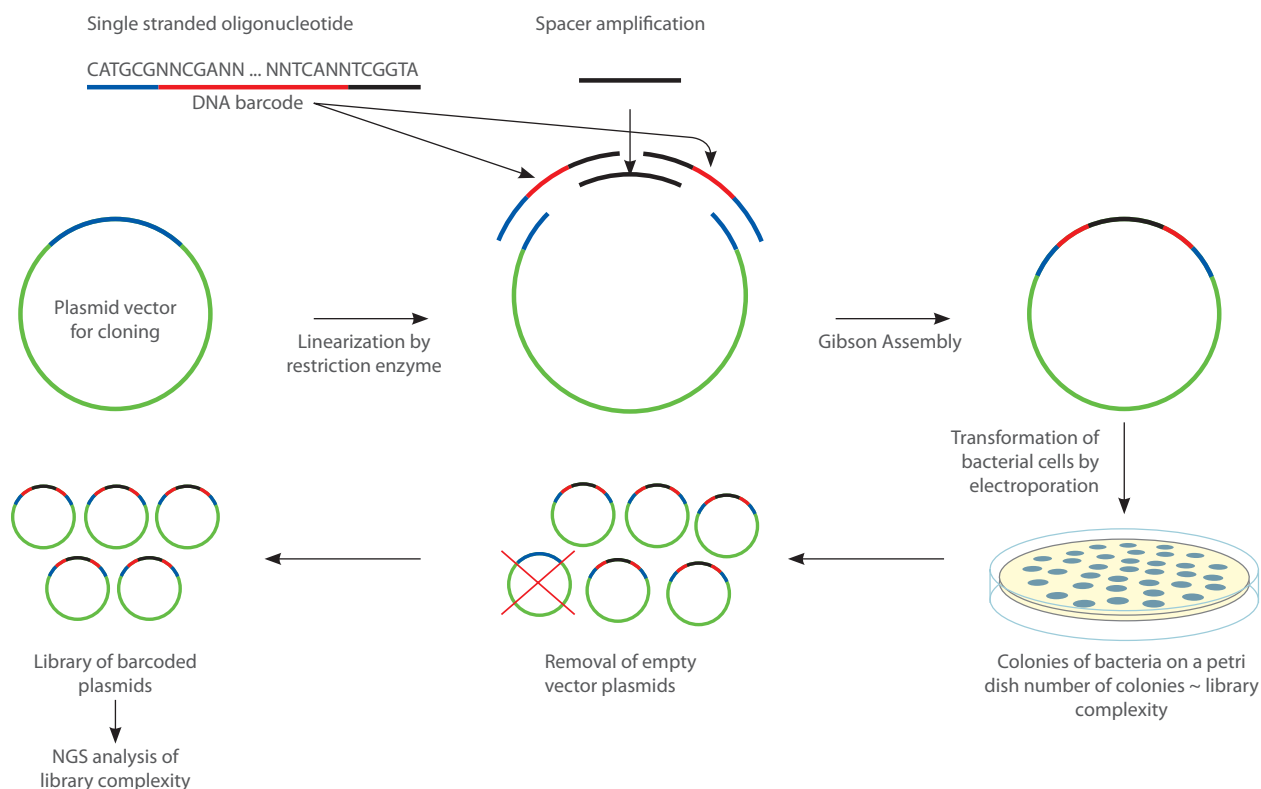
Клонирование по Гибсону – метод, который позволяет объединять несколько молекул ДНК, имеющих частично перекрывающиеся последовательности на концах (так называемые плечи). В настоящей работе мы применили этот способ для получения кольцевых молекул с парами баркодов, что позволило сохранить высокое разнообразие (более 10<sup>5</sup>) и равномерное распределение уникальных меток. Полученную библиотеку мы планируем использовать для инъекций в пронуклеус зигот, чтобы проанализировать процесс рекомбинации концов молекул при трансгенезе.

## Материалы и методы

Целью исследования является получение конструкций, в которых в плазмидном векторе содержатся два баркода, разделенных спейсерным участком. Для сборки библиотеки мы разработали оптимальные для двух методов секвенирования (Nanopore и Illumina) последовательности баркодов, которые включили в одноцепочечные олигонуклеотиды с выступающими концами для перекрывания с плазмидным вектором и спейсерным участком (фрагменты, обозначенные синим и черным цветами, на рис. 1). Спейсерный участок (обозначен черным цветом) амплифицирован из человеческого генома (локус AAVS1). Для клонирования методом Гибсона мы смешали четыре фрагмента: линейризованный плазмидный вектор CAG-Cherry, два одноцепочечных олигонуклеотида с выступающими концами для перекрывания и спейсерный участок. После реакции Гибсона (NEBuilder HiFi DNA Assembly) мы провели электропорацию клеток Top10 *E.coli* разработанными конструкциями и получили около 160000 колоний (итоговое разнообразие библиотеки). После выделения плазмидной ДНК из клеток мы очистили ее от плазмид без встройки (исходного плазмидного вектора) и провели NGS-анализ библиотеки. Этапы исследования схематично представлены на рис. 1.

### Этап 1: дизайн баркодированных олигонуклеотидов

Выбранные принципы конструирования баркодов зависят от целей эксперимента и планируемого разнообразия библиотеки. Плазмидная библиотека, протокол для клонирования которой представлен ниже, применяется для изучения рекомбинации молекул ДНК в эмбрионах с последующим анализом на платформах секвенирования третьего поколения (Oxford Nanopore, PacBio). К сожалению, у них есть значительный недостаток – высокая вероятность ошибочного прочтения нуклеотида. По нашему опыту, вероятность ошибки в конкретной букве составляет около 15%,



**Рис. 1.** Схема последовательности экспериментов по созданию библиотек баркодированных плазмид с помощью метода клонирования по Гибсону

**Fig. 1.** Experimental pipeline for creating barcoded plasmid libraries by using the Gibson cloning method

это означает, что ошибки будут возникать при чтении почти каждого баркода независимо от его длины. Разработанный нами дизайн баркода вида NNCGANNGCANNTGCNNAGCNNCTGNNTCANN отличается большей длиной (32 нуклеотида и из них 14 вырожденных позиций) по сравнению с обычно используемыми баркодами (Yunusova et al., 2017; Lebedev et al., 2019). При этом вероятность возникновения ошибки при чтении баркода составляет  $14 * 0.15 = 2.1\%$ . Однако реальное разнообразие готовой библиотеки составило около 160000 молекул – при таких параметрах вероятность из-за ошибок в прочтении перепутать его с другим баркодом из библиотеки составляет менее 0.1%. В баркодах постоянные триплеты нуклеотидов чередуются с вырожденными позициями, чтобы избежать неравновесия по GC-составу между баркодами. Несмотря на то что последовательности баркодов разработаны для секвенирования методами третьего поколения для изучения рекомбинации в эмбрионах, такой дизайн подходит и для секвенирования методом Illumina для анализа баркодированной библиотеки после сборки, который мы и описываем в настоящей работе.

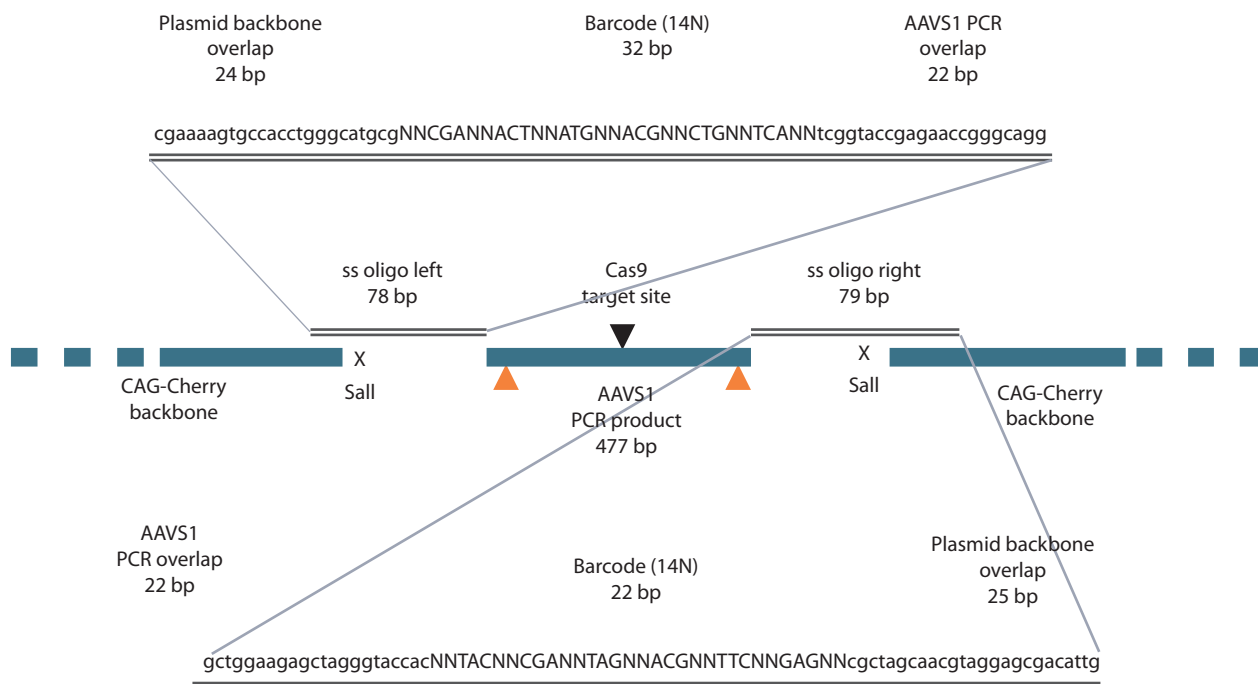
Помимо баркодной части олигонуклеотиды должны содержать выступающие концы (плечи, overlaps) для сборки методом Гибсона. Один из таких концов олигонуклеотида должен перекрываться с участком линейаризованного вектора, другой – с ПЦР-продуктом, служащим спейсером

между баркодами (рис. 2). Температура отжига плеч должна быть выше 50 °C при длине 22–25 п.о. Перед началом клонирования олигонуклеотидов важно оценить разнообразие синтезированных олигонуклеотидов описанным ранее методом (Lebedev et al., 2019). Для этого необходимо включить полученные олигонуклеотиды в ПЦР-фрагмент, после чего провести его секвенирование методом Сэнгера. По полученным пикам в хроматограмме можно оценить представленность каждого нуклеотида в вырожденной позиции: в олигонуклеотидах хорошего качества в каждой из них должно быть четыре пика равной высоты. Финальная последовательность пары олигонуклеотидов представлена на рис. 2.

#### Этап 2: подготовка ДНК-фрагментов для сборки библиотеки

##### Материалы:

- программируемый ПЦР-амплификатор Bio-Rad,
- олигонуклеотиды для амплификации выбранного участка генома,
- реактивы для ПЦР с высокоточной Q5-полимеразой (NEB #M0493),
- набор для выделения ДНК из агарозного геля (Macherey-Nagel NucleoSpin Gel and PCR Clean up, #740609.50),
- геномная ДНК человека с концентрацией ~50 нг/мкл,



**Рис. 2.** Схема клонирования баркодов по методу Гибсона. В центре рисунка показаны ДНК-фрагменты, использованные для получения библиотеки. Одноцепочечные олигонуклеотиды синтезированы в разных ориентациях

Желтые треугольники – сайты KpnI, которые необходимы для инвертированной ПЦР. Cas9 target site – целевой сайт узнавания Cas9 с гРНК в AAVS1-локусе (5'-GGGCCACTAGGGACAGGAT-3' + TGG PAM); Sall – сайт узнавания одноименной эндонуклеазы рестрикции в исходном векторе; plasmid backbone overlap – плечо плазмидного вектора, barcode – баркод; AAVS1 PCR overlap – плечо ПЦР-продукта AAVS1; ss oligo left – левый одноцепочечный олигонуклеотид; ss oligo right – правый одноцепочечный олигонуклеотид; AAVS1 PCR product – ПЦР-продукт AAVS1; CAG-Cherry backbone – вектор CAG-Cherry

**Fig. 2.** Scheme of barcode cloning using Gibson method. DNA fragments used for cloning are depicted in the middle. Single-strand oligonucleotides were synthesized in opposite orientations

Yellow triangles are KpnI sites, which are required for inverse PCR. Cas9 target site is situated in the AAVS1 region (5'-GGGCCACTAGGGACAGGAT-3' + TGG PAM). Sall is the Sall restriction enzyme site in the initial vector

- рестриктаза Sall («СибЭнзим» #E115),
- плазида pCAGGS-Cherry (Addgene #41583) или другой вектор по выбору.

#### ПЦР-амплификация участка генома

Основное назначение описываемой плазмидной библиотеки – изучение рекомбинации на концах баркодированных молекул, поэтому необходимо разделить пары баркодов спейсерным участком (см. рис. 2). Спейсер используют для линеаризации плазмидной библиотеки в районе между баркодами с помощью рестриктаз или нуклеазы Cas9 и направляющих гРНК (single guide RNA). В качестве такого спейсерного участка в нашем случае выбран участок AAVS1 локуса человека (Chr19:55115511-55115987, 477 п.о.), так как он содержит несколько протестированных районов для направляющих гРНК.

Для ПЦР-амплификации участка генома человека применяют высокоточную полимеразу Q5 (NEB). Это необходимо, чтобы избежать мутаций в финальной библиотеке.

Последовательности праймеров для амплификации района AAVS1:

5'-TCGGTACCGAGAACCGGGCAGGTCACGC-3' (F), 5'-GTGGTACCTAGCTCTTCCAGCCCCCTG-3' (R). Протокол проведения

ПЦР представлен в Приложении 1.

Целевой ПЦР-продукт должен иметь размер 477 п.н. и может быть выявлен методом геля-электрофореза в 2% агарозном геле с последующим выделением из геля на колонках (Macherey-Nagel NucleoSpin Gel and PCR Clean up).

#### Линеаризация вектора

В качестве вектора для клонирования баркодированной библиотеки использовали плазмиду pCAGGS-Cherry (Addgene #41583). Мы внесли небольшую модификацию в районе сайта Sall, однако это не повлияло на общую схему клонирования.

Для получения линеаризованного вектора проводится инкубация с избытком фермента Sall. Протокол приведен в Приложении 2.

Реакционная смесь инкубируется при температуре 37 °C в течение 12–16 ч. Как правило, время инкубации можно сократить до нескольких часов, но с учетом того что присутствие в финальной библиотеке плазмид без баркодов («пустые» плазмиды) не желательно, этого лучше не делать. Обратного лигирования липких концов Sall в реакции Гибсона не происходит, о чем говорит низкий фон (менее 0.1%) в контрольной трансформации (см. ниже).

Линеаризованная плазида должна иметь размер 5475 п.о. при детекции методом геля-электрофореза в 1% агарозном геле. После выделения ДНК из геля желательнее сконцентрировать ее до 50–200 нг/мкл для удобного проведения реакции Гибсона.

### Этап 3: сборка библиотеки методом Гибсона

#### Материалы:

- программируемый ПЦР-амплификатор,
- NEBuilder® HiFi DNA Assembly Master Mix (NEB #E2621),
- магнитные частицы AMPure XP (Beckman Coulter #A63880),
- магнитный штатив для 0.2 мкл пробирок,
- этанол 96%.

Клонирование методом Гибсона позволяет создавать сложные генетические конструкции из большого числа фрагментов (Gibson et al., 2009). Реакционная смесь содержит три фермента. Один из них – экзонуклеаза, которая отщепляет нуклеотиды с двуцепочечных концов ДНК в направлении 5'-3', оставляя выступающие 3'-концы длиной до нескольких сотен нуклеотидов. Одноцепочечные концы отжигаются между собой за счет перекрывающихся участков (overlaps), внесенных с праймерами, и высокоточная полимераз застраивает бреши в районах отжига 3'-концов. Наконец, T4-лигаза лигирует одноцепочечные разрывы между молекулами.

Для клонирования библиотеки методом Гибсона необходимо смешать ДНК-фрагменты и добавить к ним реакционную смесь ферментов. Приведенные ниже концентрации ДНК-фрагментов рассчитаны исходя из соотношения 1 : 6 : 10 : 10 (вектор : ПЦР-фрагмент : олигонуклеотид 1 : олигонуклеотид 2). Эти условия хорошо зарекомендовали себя для клонирования данной плазмидной библиотеки. В случае замены клонируемых ДНК-фрагментов следует оптимизировать их соотношение, подсчитав количество колоний после контрольных трансформаций.

Смешайте в пробирке на льду в указанном порядке:

1. 2.5 мкл ddH<sub>2</sub>O,
2. 0.5 мкл (11.82 фмоль линеаризованного вектора pCAGGS-Cherry (5475 п.о., 40 нг),
3. 1 мкл (67.84 фмоль) ПЦР-продукта AAVS1 (477 п.о., 20 нг),
4. 0.5 мкл (124.4 фмоль) одноцепочечного олигонуклеотида 1 (78 п.о., 3 нг),
5. 0.5 мкл (122.8 фмоль) одноцепочечного олигонуклеотида 2 (79 п.о., 3 нг),
6. 5 мкл смеси NEBuilder® HiFi DNA Assembly Master Mix.

Общий объем реакции составляет 10 мкл. Реакционную смесь инкубируют при 50 °С в течение часа и затем хранят при температуре –20 °С.

Для контроля также следует поставить реакцию без одноцепочечных олигонуклеотидов, заменив их водой. Это позволит оценить число исходной «пустой» плазмиды в контрольной трансформации.

**Дополнительная опция:** очистка смеси на магнитных частицах. Несмотря на то что в протоколе производителя набора NEBuilder® HiFi DNA Assembly Master Mix не рекомендуется дополнительно очищать реакционную смесь перед электропорацией, это,

как правило, значительно улучшает эффективность трансформации. Для очистки можно использовать преципитацию этанолом или спин-колонки, однако оптимальным является метод связывания ДНК на магнитных частицах. С протоколом такой очистки можно ознакомиться в Приложении 3.

### Этап 4: приготовление плазмидной баркодированной библиотеки

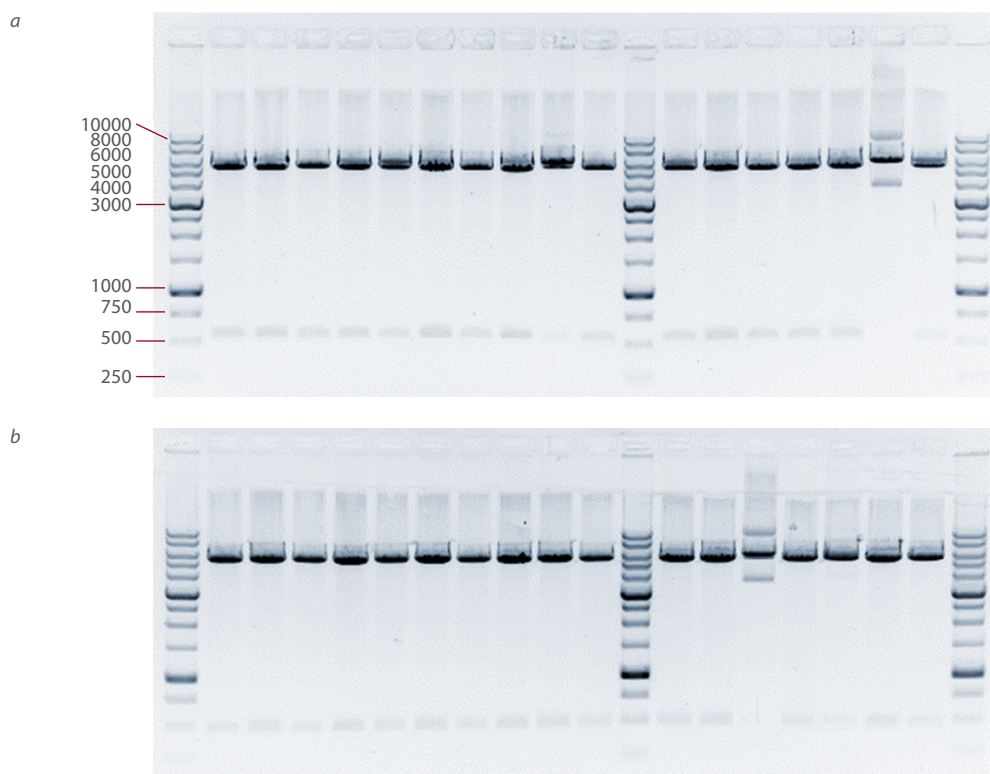
#### Материалы:

- бактериальный термошейкер,
- электропоратор для бактериальных клеток (MicroPulser Electroporator, Bio-Rad),
- кюветы для электропорации 1 мм (Bio-Rad #1652089),
- ультрацентрифуга с адаптерами под 50 мл пробирки,
- глицерин,
- набор для выделения плазмидной ДНК Plasmid Midiprep 2.0 («Евроген» #BC124) или аналог,
- набор diaGene для выделения плазмидной ДНК из бактерий («Диаэм» #3316.0250),
- ДНКаза Plasmid-Safe™ ATP-Dependent DNase (Lucigen #E3101),
- рестриктаза Sall, («СибЭнзим» #E115),
- карбенициллин,
- жидкая среда LB (на 100 мл содержит: бактотриптон 1 г, бактериально-дрожжевой экстракт 0.5 г, NaCl 0.5 г),
- агаризованная среда LB (на 100 мл содержит: бактотриптон 1 г, бактериально-дрожжевой экстракт 0.5 г, NaCl 0.5 г, агар 1.5 г),
- среда SOC (на 100 мл содержит: бактотриптон 2 г, бактериально-дрожжевой экстракт 0.5 г, 10 мМ NaCl, 2.5 мМ KCl, 10 мМ MgCl<sub>2</sub>, 10 мМ MgSO<sub>4</sub>, 20 мМ глюкозы),
- электрокомпетентные клетки Top10 (с компетентностью не ниже 10<sup>9</sup>),
- рестриктазы SphI-HF (NEB #R3182), NheI-HF (NEB #R3131), KpnI-HF (NEB #R3142),
- буфер TE (10 мМ Tris-HCl, 1 мМ EDTA).

#### Контрольная трансформация

Для оценки эффективности клонирования библиотеки следует поставить контрольную электропорацию с небольшим объемом реакции Гибсона. Все манипуляции проводят на льду. Один микролитр реакции (1/3 исходной реакции Гибсона, очищенной магнитными частицами) добавляют к электрокомпетентным клеткам Top10 (50 мкл). Электропорацию проводят в 1 мм кюветах при стандартных настройках (предварительные настройки, Ec1, 1.8 кВ) в приборе MicroPulser (Bio-Rad). После электропорации клетки быстро вымывают из кюветы в 1 мл среды SOC и инкубируют в бактериальном термошейкере (220 об/мин) при температуре 37 °С в течение 30 мин. Затем 1/10 объема клеток равномерно распределяют до выпитывания на чашке Петри с агаризованной средой LB и карбенициллином (100 мкг/мл). Чашку Петри с бактериями помещают в термостат и инкубируют при 37 °С в течение 14–18 ч.

Цель контрольной трансформации – оценить потенциальное число колоний (разнообразие библиотеки баркодов) с учетом всех факторов. Например, в представленном протоколе для контрольной трансформации использовано



**Рис. 3.** Рестрикционный анализ субклонов плазмидной библиотеки. Для оценки выделяли плазмиды из случайных колоний, выросших на чашках после электропорации смеси реакции Гибсона. *a* – рестрикция по сайтам SphI и NheI. Сайты расположены на дальних (относительно ПЦР-фрагмента) 5'-концах баркодированных олигонуклеотидов (см. схему). Ожидаемый размер фрагментов – 5467 + 561 п.о. *b* – рестрикция по сайтам KpnI. Сайты расположены на внутренних (относительно ПЦР-фрагмента) 3'-концах баркодированных олигонуклеотидов. Ожидаемый размер фрагментов – 5545 + 483 п.о.

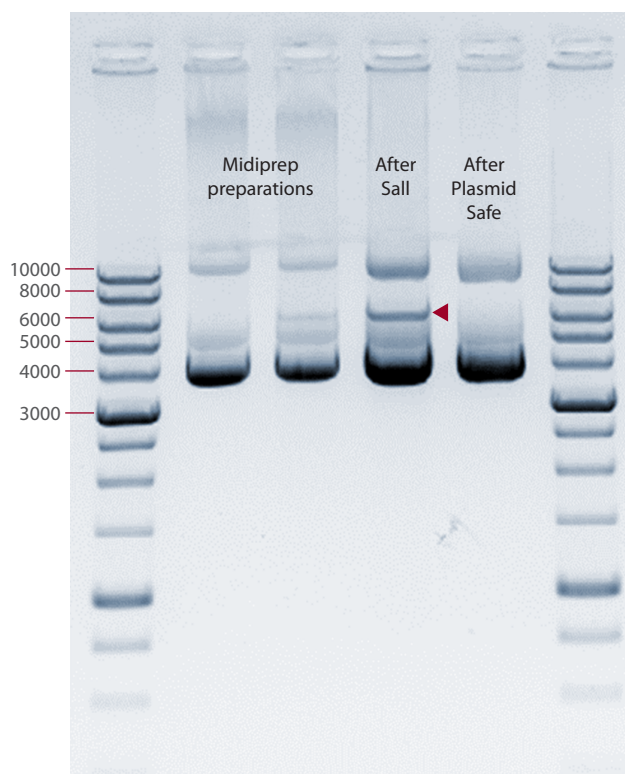
**Fig. 3.** Restriction analysis of the subclones from the barcoded plasmid library. Plasmids were purified from overnight incubations of random bacterial colonies, which were obtained after electroporation with Gibson reaction mix. *(a)* Restriction digestion at SphI and NheI sites. Sites are located at the distant (relative to the PCR fragment) 5'-ends of barcoded oligonucleotides (scheme). Predicted size of fragments 5467 + 561 b.p. *(b)* Restriction digestion at KpnI sites. Sites are located at the inner (relative to the PCR fragment) 3'-ends of barcoded oligonucleotides (scheme). Predicted size of fragments 5545 + 483 b.p.

1/3 стандартной реакции Гибсона и 1/8 от 50 мкл (разбавление 10% глицерином) с компетентностью  $2 \times 10^9$  (Morrison, 2001). На чашку высажено 1/10 клеток из 1 мл SOC. Всего на чашке выросло 2 тыс. колоний. Это означает, что ожидаемое разнообразие библиотеки составляет  $3 \times 8 \times 10 \times 2000 = 480000$  на 40 нг вектора (стандартная реакция Гибсона) при электропорации в 50 мкл неразбавленных клеток. Такая эффективность клонирования библиотеки является хорошим результатом, сравнимым с клонированием одиночных фрагментов методом Гибсона или рестрикции и лигирования (Lebedev et al., 2019).

Для предварительной характеристики библиотеки следует также провести рестрикционный анализ плазмидных субклонов. Случайным образом выбрать 15–20 колоний, нарастить ночные культуры в небольшом объеме LB и выделить плазмидную ДНК на колонках (miniprep) (набор diaGene для выделения плазмидной ДНК из бактерий). Для характеристики библиотеки следует выполнять рестрикционный анализ клонов с парой рестриктаз NheI-HF и SphI-HF, а также KpnI-HF, сайты для которых присутствуют в баркоди-

рованных олигонуклеотидах. С протоколом рестрикционного анализа можно ознакомиться в Приложении 4. Затем проводится детекция методом геля-электрофореза в 1% агарозном геле.

Предварительный рестрикционный анализ библиотеки показывает, правильно ли происходит встройка баркодов методом Гибсона (рис. 3). Следует обратить внимание на ожидаемый размер встройки, наличие «пустых» плазмид, а также на встречающиеся иногда некорректные варианты плазмид (см. рис. 3, два случая), которые связаны с мутациями в сайтах рестрикции ферментов KpnI, SphI или NheI. Детектируемые в районе олигонуклеотидов мутации (5–10% олигонуклеотидов) представляют собой нуклеотидные замены и небольшие делеции и связаны как с ошибками синтеза олигонуклеотидов, так и рекомбинацией при встройке олигонуклеотидов методом Гибсона. Если число мутантных олигонуклеотидов больше, необходимо заказать их у нового поставщика. Присутствие небольшого фона мутаций неизбежно и не влияет на общее качество библиотеки, так как при дизайне эксперимента сайты для рестриктаз или прай-



**Рис. 4.** Удаление «пустых» векторов из баркодированной библиотеки. На фореже представлены образцы баркодированной библиотеки в равных количествах. Выделенная ДНК-библиотека последовательно обработана рестриктазой Sall и экзонуклеазой Plasmid-Safe. После обработки линейные молекулы размером 6028 п.о. исчезают (соответствующий фрагмент обозначен черным треугольником). Число линейных молекул искусственно завышено, так как Sall дополнительно гидролизует молекулы по сайтам в небольшом проценте баркодов (3%), однако это создает заметный фрагмент на электрофореграмме, наглядно иллюстрирующий активность ДНКазы

Midiprep preparations – выделение midiprep; after Sall – после гидролиза ферментом Sall; after Plasmid-Safe – после обработки Plasmid-Safe

**Fig. 4.** Elimination of empty vectors from the barcoded plasmid library. Equal amounts of barcoded plasmid library samples were successively treated with Sall and Plasmid-Safe exonuclease. As can be seen, linear molecules of 6028 b.p. length disappear (fragment is indicated with black triangle). The amount of linear copies was increased artificially because Sall sites can be found in 3% of barcodes

меров можно выбрать на расстоянии от олигонуклеотидов. В некотором смысле мутации даже создают дополнительное разнообразие в библиотеке, но это нужно учитывать при биоинформатическом анализе данных NGS (см. ниже).

#### Финальная электропорация

После оценки примерного разнообразия библиотеки нужный объем реакции Гибсона добавляют к электрокомпетентным клеткам, чтобы получить планируемое число колоний. В нашем случае взяты 1 мкл очищенной реакции Гибсона (1/3 исходной реакции) и одна фасовка 50 мкл клеток с компетентностью  $2 \times 10^9$ . Электропорацию проводят по описанному ранее протоколу. При работе с большим числом коло-

ний ( $>100000$ ) оптимально использовать 6–8 крупных чашек Петри размером  $23 \times 23$  см. Для получения качественной библиотеки клетки должны быть распределены по чашкам равномерно. Необходимо также аккуратно заливать чашки с агаром, так как неравномерная глубина агара приведет к различию в размере колоний. Затем чашки с бактериями помещают в термостат и инкубируют при  $37^\circ\text{C}$  в течение 14–18 ч. Подсчет колоний можно провести разделив чашку на равные сектора  $4 \times 4$  см и оценив среднее число колоний в 5–6 секторах. При значительном разнообразии библиотеки такой метод позволяет достаточно точно оценить количество баркодов. В нашем случае число колоний составило около 150 тыс.

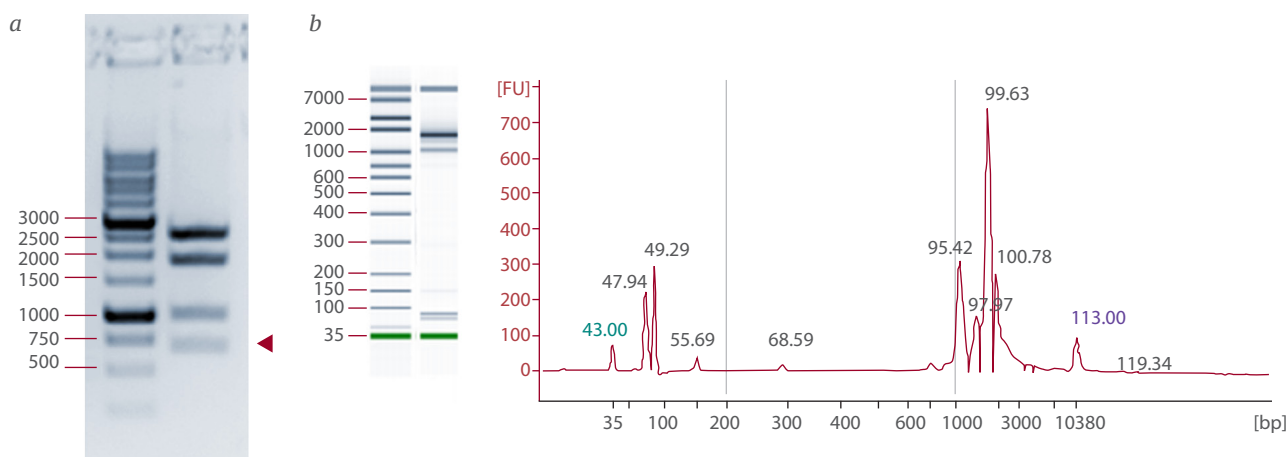
Для выделения плазмидной библиотеки из колоний их смывают с культуральных чашек средой LB (200–400 мл) и осаждают центрифугированием ( $5000g$  10 мин). Затем осадок клеток обрабатывают в зависимости от выбранного набора для выделения плазмидной ДНК. Смывого количества клеток хватает для получения ДНК-библиотеки около 400 мкг. Для увеличения выхода следует разбавить смывые клетки дополнительным объемом LB до низких значений оптической плотности (OD) (около 0.5) и инкубировать библиотеку несколько часов при  $37^\circ\text{C}$  для достижения оптимального OD ( $>2$ ). Стоит отметить, что продолжительные инкубации  $>1$  ч нежелательны, так как несинхронные деления клеток приводят к сдвигу представленности баркодов.

**Дополнительная опция:** заморозка клеток с глицерином для длительного хранения.

Необходимо ресуспендировать часть клеток с чашек в среде LB, смешанной с глицерином (финальная концентрация глицерина – 33%). Затем ресуспендированные клетки можно разделить на аликвоты по 1 мл, заморозить в азоте и поместить для длительного хранения при  $-80^\circ\text{C}$ . После характеристики библиотеки глицериновые стоки можно использовать для наработки новой библиотеки с известными баркодами.

#### Удаление «пустых» векторов ДНКазой Plasmid-Safe

Полученную плазмидную библиотеку можно сразу использовать для экспериментов. Однако для некоторых задач, например для изучения рекомбинации между баркодированными молекулами, присутствие даже небольшого (менее 0.1%) числа исходных «пустых» векторов, без встройки, является нежелательным. Сначала исходные «пустые» векторы конвертируют в линейные молекулы за счет обработки плазмидной библиотеки рестриктазой Sall (сайты узнавания Sall должны разрушаться при встройке баркодированных фрагментов). ДНКазы Plasmid-Safe представляет собой экзонуклеазу, которая деградирует двухцепочечные линейные молекулы ДНК. В дополнение к удалению «пустых» векторов ДНКазная обработка Plasmid-Safe помогает удалить фрагменты бактериальной геномной ДНК и случайно линейризованные молекулы баркодированных плазмид, которые всегда присутствуют в плазмидных выделениях (см. например одну из двух дорожек геля плазмидной библиотеки midiprep на рис. 4). Полный протокол очистки экзонуклеазой Plasmid-Safe представлен в Приложении 5.



**Рис. 5.** Подготовка фрагмента с парой баркодов из плазмидной библиотеки для NGS. *a* – рестрикция плазмидной библиотеки по сайтам *SpeI* и *BspHI*. ДНК-фрагмент, отмеченный стрелкой (681 п.о.), содержит пары баркодов, разделенные сайтом AAVS1. *b* – анализ ДНК-библиотеки после лигирования адаптеров на приборе Bioanalyzer 2100

Исходя из этих данных, основная часть ДНК-фрагментов имеет размер в диапазоне 1000–2000 п.о., однако это является артефактом детекции из-за наличия на концах молекул одноцепочечных адаптеров, которые меняют подвижность молекул в геле. Как показывает дальнейший NGS-анализ, реальный размер библиотеки соответствует ожидаемому

**Fig. 5.** Preparation of fragments with barcode pairs for NGS. (*a*) Restriction digestion of initial library at *SpeI* and *BspHI* sites. Indicated DNA fragment (681 b.p.) contains barcode pairs separated by the AAVS1 site. (*b*) DNA library analysis on the Bioanalyzer 2100

These data imply that the main fraction of the fragments is in a 1000–2000 b.p. range. This is a mere artefact of detection, because attaching single-stranded adapters to the ends of the DNA molecules change mobility of the molecules in gel. The following NGS analysis revealed that the real size of the library matches well the predicted outcomes

### Этап 5: подготовка библиотеки для NGS-секвенирования на приборе Illumina HiSeq X Ten в режиме парных чтений 2 × 150 п.о.

#### Материалы:

- КАРА HyperPrep Kit (Roche #07962312001) с адаптерами КАРА Unique Dual-Indexed Adapters (Roche #08278555702),
- магнитные частицы AMPure XP (Beckman Coulter #A63880),
- магнитный штатив для 0.2 мл пробирок,
- прибор Qubit 4.0 (Thermo Fisher Scientific) с расходниками,
- прибор Bioanalyzer 2100 (Agilent) с расходниками,
- рестриктазы *SpeI*-HF (NEB #R3133) и *BspHI* (NEB #R0517),
- набор для выделения ДНК из агарозного геля (Macherey-Nagel NucleoSpin Gel and PCR Clean up, #740609.50).

#### Рестрикция плазмидной библиотеки

Чтобы избежать рекомбинации или мутаций в парах баркодов, подготовку библиотеки для NGS рекомендуется проводить без стадии ПЦР-амплификации. Поэтому фрагмент с парой баркодов, разделенных сайтом AAVS1, вырезают из плазмиды при помощи рестриктаз. Важно отметить, что сайты рестрикции не должны располагаться дальше, чем 100 п.о. от баркодов (предел прочтения большинства приборов для массового параллельного секвенирования составляет 150 п.о.). Для плазмидной библиотеки на основе pCAGGS-Cherry, описанной в данном протоколе, следует использовать рестриктазы *SpeI*-HF и *BspHI*, сайты рестрикции которых расположены на расстоянии 7 и 62 п.о. от баркоди-

рованных олигонуклеотидов соответственно. С учетом того что целевой рестрикционный фрагмент составляет около 10% размера плазмиды, в реакцию рестрикции необходимо взять большое количество библиотеки. Полный протокол рестрикции представлен в Приложении 6.

Фрагмент ожидаемой длины (681 п.о.) (рис. 5, *a*) выделяют из геля с помощью соответствующего набора и используют для последующего лигирования адаптеров для NGS.

#### Лигирование фрагмента библиотеки с адаптерами для NGS

Для лигирования адаптеров на приборе Illumina HiSeq X Ten в режиме парных чтений 2 × 150 п.о. используют набор КАРА HyperPrep. Подготовка образцов проводят без стадии амплификации (Library Amplification). Подробный протокол представлен в Приложении 7 или на сайте производителя: <https://sequencing.roche.com/en/products-solutions/products/sample-preparation/dna-reagents/library-preparation/kapa-hyperprep.html>

#### Анализ ДНК-библиотеки и расчеты необходимого количества ДНК

Итоговый анализ концентрации ДНК-библиотеки проводят на приборах Qubit 4.0 и Bioanalyzer 2100. Определяемый размер итоговой библиотеки отличается от ожидаемого: 1000–2000 вместо ~800 п.о. (см. рис. 5, *b*). Это артефакт, который характерен для библиотек, приготовленных без стадии амплификации, и связан с наличием на концах целевых молекул одноцепочечных участков Y-адаптеров, которые замедляют движение молекул в геле.



Перед отправкой образцов на NGS необходимо оценить оптимальное число прочтений, с которым нужно секвенировать библиотеку. Для описанной в данном протоколе библиотеки с ожидаемым разнообразием ~150000 вариантов (см. оценку числа колоний ранее) оптимальное число прочтений можно оценить как «разнообразие» × 1000 ридов, т. е. ~150 млн ридов. Это соответствует ~100 пг ДНК-библиотеки в финальном образце для NGS.

#### Этап 6: анализ баркодов в данных NGS

Анализ проведен с использованием операционной системы Linux. При обработке данных использованы следующие вспомогательные программы:

1. Cutadapt v2.8 – распознавание и удаление технических последовательностей;
2. Python v3.7:
  - biopython: работа с данными секвенирования,
  - networkx: построение графов,
  - pandas/pandarallel: работа с таблицами,
  - scipy.stats: пакет для статистической обработки данных,
  - Levenstein: вычисление расстояния Левенштейна,
  - matplotlib: построение графиков.

#### Определение последовательности баркода в составе прочтения

Анализ данных секвенирования необходимо начинать с определения последовательности ДНК-баркода в составе рида. Для этого можно использовать два альтернативных подхода: определение баркода с помощью утилиты cutadapt (Martin, 2011) либо выравнивание рида на ожидаемую последовательность баркода с помощью алгоритма bwa (Li, Durbin, 2009). По нашему опыту, решить эту задачу можно с помощью обоих подходов, однако cutadapt позволяет сделать это быстрее и проще. Cutadapt запускали со следующими параметрами:

```
"cutadapt -j $threads -e $error_rate -g $seq_1R -g $seq_1F -G $seq_1R -G $seq_1F -O $search_length -o test_R1.fq -p test_R2.fq input_R1.fq input_R2.fq"
```

где:

- error\_rate = 0.1 – допустимая доля ошибок в искомой последовательности;
- seq\_1F, seq\_1R – 15-буквенные фрагменты плазмиды, примыкающие к баркодам;
- search\_length = 15 – минимальная длина перекрытия, равная длине искомого фрагмента (т. е. поиск фрагмента только внутри рида).

#### Определение правого и левого баркодов

Правый и левый баркоды в ридах определяли по уникальным последовательностям в окружении баркода. Если позиция обрезки совпала с предполагаемой ( $\pm 2$  буквы) – рид считали подходящим: ему присваивали соответствующую букву (F или R) и с краю вырезали 34 буквы (длина баркода + 2).

#### Объединение баркодов в пары

Для определения пар баркодов использовали метод построения графа с помощью библиотеки NetworkX (Hagberg et al., 2008). Вершинами графа были последовательности баркодов, весами ребер – количество попаданий баркодов в пару. После отсека пар баркодов, подтвержденных не-

достаточным количеством ридов (cutoff level), составляли список баркодов и их пар, отсортированных по числу связей. Анализ этого списка позволяет получить оценку разнообразия библиотеки баркодированных молекул.

#### Заключение

Наша недавняя работа показала, что ДНК-баркоды не блокируют рекомбинацию на концах молекул трансгенов и даже помогают детектировать специфические паттерны рекомбинации (Smirnov et al., 2020). ДНК-баркоды нечасто применяют для изучения процессов рекомбинации. Как правило, в работах ведут анализ полиморфизмов в ДНК-последовательностях (SNP, сайты рестрикции), перенос которых между молекулами указывает на копирование генетического материала (genetic conversion) (Kan et al., 2014; 2017). Баркоды создают пул из тысяч индивидуально меченых молекул, каждая из которых может давать информацию о рекомбинации. Стоит отметить, что использование метода Гибсона позволяет осуществлять сборку большого числа фрагментов. К примеру, в описанном нами протоколе два фрагмента ДНК объединяются с двумя олигонуклеотидами. После подбора условий можно провести клонирование трех фрагментов ДНК и трех олигонуклеотидов, получив библиотеку плазмид с тремя сайтами баркодирования. Объединив две такие независимые библиотеки, можно получить мультибаркодированные плазмиды, содержащие пять – шесть баркодов по всей длине. Такой дизайн поможет в исследовании рекомбинации, которая происходит вдали от двухцепочечных концов молекул, что является приоритетной задачей для объяснения патологических путей рекомбинации (Verma, Greenberg, 2016; Piazza, Heyer, 2019). Прочтение мультибаркодированных библиотек можно осуществлять методами секвенирования третьего поколения (Nanopore, PacBio) для упрощения анализа баркодов.

#### Список литературы / References

- Юнусова А.М., Баттулин Н.Р. Методы маркирования клеток для изучения судьбы клеточных поколений. *Вавиловский журнал генетики и селекции*. 2016;20(6):909-917. DOI 10.18699/VJ16.211.
- [Yunusova A.M., Battulin N.R. Cell-marking techniques for cell lineage tracing. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2016;20(6):909-917. DOI 10.18699/VJ16.211. (in Russian)]
- Arnold A.E., Malek-Adamian E., Le P.U., Meng A., Martínez-Montero S., Petrecca K., Damha M.J., Shoichet M.S. 2018. Antibody-Antisense Oligonucleotide Conjugate Downregulates a Key Gene in Glioblastoma Stem Cells. *Mol Ther Nucleic Acids*. 2018;11:518-527. DOI 10.1016/j.omtn.2018.04.004.
- Gibson D.G., Young L., Chuang R.Y., Venter J.C., Hutchison C.A., Smith H.O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*. 2009;6(5):343-345. DOI 10.1038/nmeth.1318.
- Hagberg A., Swart P., Chult D. Exploring Network Structure, Dynamics, and Function Using NetworkX. *Proceedings of the 7th Python in Science Conference*. 2008;11-16.
- Kan Y., Ruis B., Lin S., Hendrickson E.A. The mechanism of gene targeting in human somatic cells. *PLoS Genet*. 2014;10(4):e1004251. DOI 10.1371/journal.pgen.1004251.
- Kan Y., Ruis B., Takasugi T., Hendrickson E.A. Mechanisms of precise genome editing using oligonucleotide donors. *Genome Res*. 2017;27(7):1099-1111. DOI 10.1101/gr.214775.116.
- Klein J.C., Agarwal V., Inoue F., Keith A., Martin B., Kircher M., Ahituv N., Shendure J. A systematic evaluation of the design and context dependencies of massively parallel reporter assays. *Nat. Methods*.

- 2020;17:1083-1091. DOI 10.1038/s41592-020-0965-y.
- Lebedev M.O., Yarinich L.A., Ivankin A.V., Pindyurin A.V. Generation of barcoded plasmid libraries for massively parallel analysis of chromatin position effects. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2019;23(2):203-211. DOI 10.18699/VJ19.483.
- Li H., Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754-1760. DOI 10.1093/bioinformatics/btp324.
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 2011;17:10. DOI 10.14806/ej.17.1.200.
- Morrison S.L. Transformation of *E. coli* by Electroporation. In: Current Protocols in Immunology. USA, Hoboken, NJ: John Wiley & Sons, Inc., 2001.
- Piazza A., Heyer W.-D. Homologous recombination and the formation of complex genomic rearrangements. *Trends Cell Biol*. 2019;29(2):135-149. DOI 10.1016/j.tcb.2018.10.006.
- Smirnov A., Fishman V., Yunusova A., Korablev A., Serova I., Skryabin B.V., Rozhdestvensky T.S., Battulin N. DNA barcoding reveals that injected transgenes are predominantly processed by homologous recombination in mouse zygote. *Nucleic Acids Res*. 2020;48(2):719-735. DOI 10.1093/nar/gkz1085.
- Verma P., Greenberg R.A. Noncanonical views of homology-directed DNA repair. *Genes Dev*. 2016;30(10):1138-54. DOI 10.1101/gad.280545.116.
- Yunusova A.M., Fishman V.S., Vasiliev G.V., Battulin N.R. Deterministic versus stochastic model of reprogramming: new evidence from cellular barcoding technique. *Open Biol*. 2017;7:160311. DOI 10.1098/rsob.160311.

---

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 02.12.2020. После доработки 08.02.2021. Принята к публикации 24.02.2021.

### Приложение 1

Условия реакции с Q5-полимеразой:

1. 32.5 мкл ddH<sub>2</sub>O,
2. 10 мкл буфера Q5 5x,
3. 1 мкл дНТФ 10 мМ (финальная концентрация 200 мкМ),
4. 2.5 мкл праймера F 10 мкМ (финальная концентрация 0.5 мкМ),
5. 2.5 мкл праймера R 10 мкМ (финальная концентрация 0.5 мкМ),
6. 0.5 мкл Q5-полимеразы (финальная концентрация 0.02 ед/мкл),
7. 1 мкл (50 нг) геномной ДНК.

Итоговый объем 60 мкл.

Условия ПЦР: начальная денатурация ДНК – 98 °С в течение 30 с с последующей амплификацией (35 циклов), включая денатурацию при 98 °С – 15 с, отжиг праймеров при 60 °С – 30 с и элонгацию при 72 °С – 1 мин. Финальная элонгация: 72 °С – 3 мин.

### Приложение 2. Протокол проведения реакции рестрикции с ферментом Sall

Смешать в пробирке на льду в указанном порядке:

1. 50-Х мкл ddH<sub>2</sub>O, где Х – объем остальных компонентов;
2. 5 мкл буфера O («СибЭнзим») 10x;
3. 2 мкл (20U) Sall;
4. 0.5–5 мкл (3 мкг) плазмиды pCAGGS-Cherry.

### Приложение 3

Для очистки 10 мкл реакции Гибсона на магнитных частицах AMPure XP (Beckman Coulter) необходимо:

1. Предварительно нагреть раствор магнитных частиц до комнатной температуры (30 мин).
2. Перемешать раствор на вортекс-миксере. Добавить раствор частиц к реакционной смеси Гибсона в соотношении 1 : 1 (в нашем случае 10 мкл). Ресуспендировать.
3. Инкубировать 10–15 мин при комнатной температуре, чтобы ДНК связалась с частицами.
4. Поставить смесь в магнитный штатив и дождаться осаждения частиц на стенке пробирки. Это занимает 5–10 мин.
5. Не вынимая пробирки из штатива дважды промыть частицы большим объемом свежеприготовленного 80% этанола. Убрать этанол.
6. Дождаться испарения излишков этанола при комнатной температуре в течение 3–5 мин. Важно не допустить пересыхания осадка, поэтому нужно контролировать процесс высыхания визуально. Осадок магнитных частиц должен иметь матовый блеск без трещин. Пересушивание частиц приводит к снижению выхода ДНК после элюции.
7. Ресуспендировать подсушенные частицы в подходящем объеме ddH<sub>2</sub>O (3–10 мкл) для элюции ДНК. Инкубировать раствор 5–10 мин при комнатной температуре. Вернуть пробирки в магнитный штатив и дождаться осаждения частиц на стенке пробирки. Отобрать чистый раствор ДНК – его можно сразу использовать для электропорации клеток или хранить при –20 °С.

### Приложение 4. Протокол рестрикционного анализа библиотеки

Смешайте в пробирке на льду в указанном порядке:

1. 30-Х мкл ddH<sub>2</sub>O, где Х – объем остальных компонентов;
2. 3 мкл буфера CutSmart 10x;
3. 0.1 мкл SphI-HF (2U) + 0.1 мкл NheI-HF (2U) или 0.1 мкл KpnI-HF (2U);
4. 0,5–5 мкл плазмидного клона (1 мкг).

Реакция инкубируется при 37 °С в течение 2 ч.

### Приложение 5. Протокол очистки библиотеки от линейных молекул экзонуклеазой Plasmid-Safe

Для удаления линейных плазмидных молекул из раствора плазмидной библиотеки сначала смешайте плазмидную библиотеку с рестриктазой Sall в пробирке на льду в указанном порядке:

1. 200-Х мкл ddH<sub>2</sub>O, где Х – объем остальных компонентов;
2. 20 мкл буфера O («СибЭнзим») 10x;
3. 5 мкл (50U) Sall;
4. 10–100 мкл (50 мкг) плазмидной библиотеки.

Инкубировать при 37 °С в течение 2 ч и очистить ДНК переосаждением этанолом на спин-колонках или магнитных частицах, элюировав в 50 мкл буфера TE. Затем смешать следующую реакцию на льду:

1. 250-Х мкл ddH<sub>2</sub>O, где Х – объем остальных компонентов;

2. 25 мкл буфера Plasmid-Safe 10x;
3. 10 мкл АТФ 25 мМ (финальная концентрация 1 мМ);
4. 15 мкл (150U) ДНКазы Plasmid-Safe;
5. 10–100 мкл (~50 мкг) Sall-обработанной плазмидной библиотеки.

Общий объем реакции составляет 250 мкл. Реакция инкубируется при 37 °С в течение 2 ч, затем проводят инактивацию фермента при 70 °С в течение 30 мин. Далее реакцию необходимо переосадить спиртом, очистить на спин-колонке или магнитных частицах. Финальный раствор плазмидной библиотеки в буфере TE можно использовать для дальнейших экспериментальных задач, включая NGS-секвенирование библиотеки.

#### Приложение 6. Протокол линеаризации библиотеки перед секвенированием

Смешайте в пробирке на льду в указанном порядке:

1. 50-X мкл ddH<sub>2</sub>O, где X – объем остальных компонентов;
2. 5 мкл буфера CutSmart 10x;
3. 1 мкл (20U) SpeI-HF;
4. 2 мкл (20U) BspHI;
5. 4–40 мкл (20 мкг) плазмидной библиотеки.

#### Приложение 7. Протокол подготовки библиотеки к секвенированию

##### Шаг 1. Достройка концов и аденилирование (End Repair and A-Tailing)

Смешайте в пробирке на льду в указанном порядке:

1. 50 мкл ДНК (300 нг), разбавленной в ddH<sub>2</sub>O;
2. 7 мкл буфера End Repair & A-Tailing;
3. 3 мкл End Repair & A-Tailing Enzyme Mix.

Общий объем реакции составляет 60 мкл. Смесь инкубируется при 20 °С в течение 30 мин, затем при 65 °С в течение 30 мин.

##### Шаг 2. Лигирование адаптеров (Adapter Ligation)

Смешайте в пробирке на льду в указанном порядке:

1. 60 мкл предыдущей реакции (шаг 1),
2. 5 мкл выбранного адаптера (15 мкМ сток) (Kapa Adapters),
3. 5 мкл ddH<sub>2</sub>O,
4. 30 мкл лигазного буфера,
5. 10 мкл ДНК-лигазы.

Общий объем реакции составляет 110 мкл. Смесь инкубируется при 20 °С в течение 30 мин.

**Шаг 3.** Очистка на магнитных частицах, как описано ранее, с соотношением магнитных частиц к объему реакции 0,8 : 1 (0,8×)

**Шаг 4** (опционально). Дополнительная очистка на магнитных частицах для удаления димеров адаптеров (короткие фрагменты длиной менее 100 п.о., см. рис. 5, b) в соотношении 1 : 1 (1×)

Элюировать ДНК в 30 мкл ddH<sub>2</sub>O. Далее раствор можно использовать для NGS, предварительно оценив концентрацию ДНК.