

## ОСОБЕННОСТИ ПОДГОТОВКИ БИБЛИОТЕК ДЛЯ МЕТАГЕНОМНОГО СЕКВЕНИРОВАНИЯ ОБРАЗЦОВ НА ПЛАТФОРМЕ ILLUMINA

А. Ю. Красненко<sup>1</sup>✉, А. Ю. Елисеев<sup>2</sup>, Д. И. Борисевич<sup>1</sup>, К. Ю. Цуканов<sup>1</sup>, А. И. Давыдова<sup>1</sup>, В. В. Ильинский<sup>1</sup>

<sup>1</sup> ООО «Генотек», Москва

<sup>2</sup> Московская гимназия на Юго-западе № 1543, Москва

Метагеномное секвенирование, позволяющее определять таксономическую принадлежность и долю прокариотических организмов в сообществах в разных средах, широко используется не только в научной деятельности, но и в медицинской практике. В настоящее время для изучения микробиома человека применяется так называемое метагеномное секвенирование маркерных генов, при котором секвенируется не весь геном, а лишь те его регионы, по которым можно установить родовую и иногда видовую принадлежность микроорганизмов. Чаще всего для амплификации выбирают участки гена 16S рибосомальной РНК (16S рРНК), последовательность которого, с одной стороны, высококонсервативна, а с другой — содержит вариабельные участки, которые отличаются однонуклеотидными заменами в случае с разными микроорганизмами. При этом для амплификации таких участков генома возможен подбор универсальных праймеров, что значительно снижает стоимость и время исследования. Точность секвенирования в этом случае обеспечивается точностью подбора универсальных праймеров и оптимальностью условий проведения ПЦР на этапе подготовки библиотеки для секвенирования. В статье мы описываем подход к подбору универсальных праймеров и условий ПЦР для амплификации регион-специфической части V3 и V4 участков гена 16S рРНК для дальнейшего секвенирования образцов на платформах Illumina.

**Ключевые слова:** микробиом, метагеномное секвенирование, рибосомальная РНК, 16S рРНК, полимеразная цепная реакция, универсальные праймеры, библиотеки для секвенирования, двойное баркодирование

**Благодарности:** авторы благодарят Дарью Плахину и Ивана Стеценка из ООО «Генотек» за помощь в работе, а также Сергея Глаголева из Московской гимназии на Юго-Западе № 1543 — за ценные советы и замечания.

✉ **Для корреспонденции:** Красненко Анна Юрьевна  
Наставнический пер., 17, стр. 1, подъезд 14, г. Москва, 105120; annakrasnenko@gmail.com

**Статья получена:** 12.04.2017 **Статья принята к печати:** 24.04.2017

## METAGENOMIC LIBRARY PREPARATION FOR ILLUMINA PLATFORM

Krasnenko AY<sup>1</sup>✉, Eliseev AY<sup>2</sup>, Borisevich DI<sup>1</sup>, Tsukanov KY<sup>1</sup>, Davydova AI<sup>1</sup>, Ilinsky VV<sup>1</sup>

<sup>1</sup> Genotek Inc., Moscow, Russia

<sup>2</sup> Moscow South-West High School No. 1543, Moscow, Russia

Metagenomic sequencing is widely used in both scientific research and clinical practice for characterization of taxonomic profiles including estimation of relative abundance of prokaryotes in microbial communities in various media. Metagenomic sequencing of single marker genes is an excellent tool for studying the human microbiome. Unlike whole-genome sequencing, it targets those genome regions that can be instrumental in identification of microorganism species and genus. The 16S ribosomal RNA (16S rRNA) gene sequence is highly conserved but at the same time there are regions containing species-specific sequences that can discriminate between different bacteria and archaea. These regions can be amplified using universal primers, which makes the whole procedure more cost-effective and less time-consuming. Good primers and protocol design for PCR at the step of library preparation is crucial for achieving high data accuracy. Below we describe how to choose the optimal PCR protocol and universal primers to amplify V3 and V4 regions of the 16S rRNA gene for further sample sequencing using Illumina platform.

**Keywords:** microbiome, metagenomic sequencing, ribosomal RNA, 16S rRNA, polymerase chain reaction, universal primers, library preparation, double barcode

**Acknowledgements:** the authors thank Daria Plakhina and Ivan Stetsenok of Genotek for their help and Sergey Glagolev of Moscow South-West High School No. 1543 for his valuable advice and comments.

✉ **Correspondence should be addressed:** Anna Krasnenko  
Nastavnicheskiy per. 17, str. 1, pod. 14, Moscow, 105120; annakrasnenko@gmail.com

**Received:** 12.04.2017 **Accepted:** 24.04.2017

В человеческом организме существует несколько бактериальных сообществ: в ротовой полости, кишечнике, урогенитальном тракте и т. д. Совокупность микроорганизмов, находящихся в симбиозе с организмом-хозяином, называют микробиомом [1]. Изучение микробиомов человека

позволяет узнать, какие бактерии живут в нашем организме в норме и при развитии различных заболеваний [2].

В настоящее время для исследования микробиома используется секвенирование регионов геномов бактерий, по которым можно установить родовую и иногда

видовую принадлежность микроорганизмов — метагеномное секвенирование маркерных генов, или метагеномное секвенирование [3]. Одними из наиболее удобных регионов для секвенирования являются участки гена 16S рибосомальной РНК (16S рРНК). Это связано с высокой консервативностью указанного гена в геномах всех прокариот, что делает возможным подбор универсальных праймеров для амплификации интересующих участков и уменьшение стоимости и времени исследования [4, 5]. В гене 16S рРНК можно выделить консервативные и варибельные участки. Последние содержат однонуклеотидные замены, по которым можно определить видовую и/или родовую принадлежность микроорганизма путем сравнения обнаруженных в результате секвенирования замен с базой данных, составленной и обновляемой по литературным источникам.

Одним из основных этапов метагеномного секвенирования является подготовка образца для исследования, т. е. превращение исходного материала — нуклеиновой кислоты — в стандартную библиотеку фрагментов ДНК, подходящую для загрузки в секвенатор. При этом необходимо учитывать, что существует большое количество различных платформ секвенирования, и, хотя стратегии пробоподготовки во многом схожи, есть нюансы, обусловленные способом детекции сигнала в процессе секвенирования [6]. Основным шагом при подготовке образца является получение специфических фрагментов, в дальнейшем позволяющих определить видовую и/или родовую принадлежность микроорганизма. Чтобы получить точный результат методом метагеномного секвенирования, важно правильно подобрать праймеры, необходимые для наработки специфических ампликонов с помощью полимеразной цепной реакции (ПЦР) [7, 8]. В 2013 г. Klindworth и соавт. составили 512 пар праймеров, которые были разделены на несколько групп по типу технологии высокопроизводительного секвенирования (next generation sequencing, NGS): первая группа — для технологий Illumina и Ion Torrent (самые короткие продукты амплификации), вторая — для технологии 454 Life Science (средние по размеру продукты амплификации), третья — для технологий PacBio и др. и для создания колониальных библиотек (самые длинные продукты амплификации) [9]. В результате для каждой группы были найдены несколько пар праймеров для архей и бактерий, которые являются универсальными для дифференцировки разных родов и видов.

В статье мы рассматриваем особенности метагеномного секвенирования образцов на платформе Illumina, в первую очередь уделяя внимание подготовке библиотек для секвенирования, имеющей в случае с Illumina ряд особенностей [10]. Дело в том, что все процессы секвенирования на этой платформе проходят в проточной ячейке, покрытой одноцепочечными нуклеотидами, комплементарными последовательности адаптера, прикрепляемого во время приготовления библиотек. Прикрепленные к проточной ячейке фрагменты удлиняются при помощи ПЦР, при этом полимеразы начинают работать только в том случае, когда свободный конец связанного фрагмента гибридизуется с комплементарным олигонуклеотидом на поверхности ячейки. Вследствие этого происходит рост числа копий единичной молекулы в определенной части ячейки. Образуются миллионы кластеров, каждый из которых представляет собой скопление определенного фрагмента. После этого происходит распознавание последовательности: построение комплементарной цепи с использованием меченых нуклеотидов и снятие сигнала после каждого шага

удлинения. Описанная технология определяет особенности подготовки образцов к секвенированию, которые и были учтены в работе.

### Подготовка библиотек для секвенирования на платформе Illumina

Образцы для исследования могут представлять собой биологический материал различных типов: слюна, ушная сера, соскоб со слизистой оболочки полости носа и т. д. В статье мы детально рассматриваем общий подход подготовки библиотек для секвенирования на платформе Illumina независимо от исследуемого образца. Описываемый подход включает несколько этапов: 1) выделение тотальной ДНК из образца; 2) определение участков для секвенирования и подбор праймеров для дальнейшей амплификации с помощью ПЦР (на этом этапе особое внимание уделяется качеству ПЦР, от которого зависит качество последующего секвенирования); 3) двойное баркодирование полученных библиотек для секвенирования; 4) секвенирование; 5) биоинформатический анализ полученных данных.

Существует большое количество протоколов и наборов реагентов для эффективного выделения тотальной ДНК [11] в зависимости от типа образца, поэтому мы не будем подробно останавливаться на обсуждении этой задачи. Контроль качества получаемой ДНК мы проводим с помощью электрофореза в агарозном геле, концентрацию определяем на приборе Qubit 3.0 Fluorometer (Thermo Fisher Scientific, США) по протоколу производителя [12].

После выделения тотальной ДНК проводится наработка необходимых участков для секвенирования с помощью ПЦР. Нами были выбраны участки гена 16S рРНК, поскольку, как говорилось выше, последовательность гена позволяет подобрать универсальные праймеры. Качество получаемых фрагментов зависит от степени комплементарности праймеров к участкам гена 16S рРНК [13, 14]. Последовательности праймеров состоят из регион-специфичной части, которая комплементарна границам амплифицируемого региона, и синтетической части, некомплементарной специфичной части и предназначенной для последующей гибридизации адаптеров для секвенирования. Праймеры подбирают таким образом, чтобы как минимум четыре 3'-концевых нуклеотида не были комплементарны самому праймеру или праймеру в паре во избежание образования димеров праймеров. Даже единичное расхождение в комплементарности последних 3–4 нуклеотидов на 3'-конце праймера может значительно снизить эффективность ПЦР, в том числе при условии подбора оптимальной температуры отжига [15, 16].

Существует множество постоянно обновляемых баз данных, содержащих последовательности генов 16S рРНК микроорганизмов различных видов [17, 18] и позволяющих, в том числе с помощью специальных программ [19], самостоятельно подобрать универсальные праймеры. В своей работе мы выбрали последовательности универсальных пар праймеров к регион-специфической части V3 и V4 участков гена 16S рРНК [23]. В качестве синтетической части были выбраны последовательности комплементарные адаптерам двух типов: Nextera и Truseq (синтетическая часть каждого праймера предназначена для последующей гибридизации адаптеров) (табл. 1).

После подбора праймеров необходимо провести оптимизацию условий проведения ПЦР (подобрать концентрацию праймеров, концентрацию ДНК, оптимальную

**Таблица 1.** Универсальные пары праймеров, состоящие из регион специфической части V3 и V4 участков гена 16S рPHK и синтетической части, комплементарной адаптерам Nextera и Truseq

Тип адаптера	Последовательность праймера	
Nextera	прямой	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG
	обратный	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC
Truseq	прямой	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTACGGGNGGCWGCAG
	обратный	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGACTACHVGGGTATCTAATCC

температуру отжига, концентрацию  $Mg^{2+}$ , количество циклов и т. д.) для получения достаточного количества и качества необходимых ампликонов для дальнейшего секвенирования. Проверку качества ПЦР проводят с помощью электрофореза в агарозном геле. При оптимизации условий необходимо помнить об отрицательном и положительном контроле. В качестве отрицательного контроля чаще всего используют ту же ПЦР-смесь, что и для исследуемых образцов, только без добавления ДНК-матрицы. В качестве положительного контроля мы используем два образца ДНК: ДНК одного образца принадлежит бактериям рода *Rhizobium*, а другого — рода *Rhodococcus*.

На количество нарабатываемых ампликонов может влиять образование димеров праймеров во время ПЦР [20]. При этом образование димеров можно увидеть при проведении контроля в агарозном геле (рис. 1). Снизить образование димеров можно путем разбавления праймеров, добавляемых в реакцию, а также изменением температуры их отжига. Оптимальная температура отжига праймеров, кроме того, дает возможность получать наиболее чистый продукт, поскольку обеспечивает точное прикреплени праймера к участку ДНК. Современные приборы позволяют проводить ПЦР с использованием градиента температур, что позволяет подобрать оптимальную температуру отжига за один прогон. Для эффективного проведения ПЦР также важна концентрация ионов  $Mg^{2+}$ , поскольку они связываются с dNTP, праймерами, ДНК-матрицей и хелатирующими агентами (EDTA), которые могут присутствовать в буфере [21]. Известно, что активность полимеразы увеличивается при высокой концентра-

ции  $Mg^{2+}$ , однако при этом снижается точность полимеразы. Как правило, для подбора оптимальной концентрации ионов  $Mg^{2+}$  в реакционной смеси проверяют влияние концентрации в диапазоне от 1 до 4 мМ с шагом в разбавлении 0,5 мМ.

При подборе условий для ПЦР оказалось, что для амплификации регионов V3 и V4 16S рPHK больше подходят праймеры под адаптеры Nextera. Вероятно, это связано с уникальностью последовательности Nextera, которая не комплементарна участкам генома исследуемых бактерий, что исключает образование побочных продуктов. Оптимальные условия для проведения ПЦР на приборе Step One Plus (Applied Biosystems, США) представлены в табл. 2.

После получения необходимого фрагмента для секвенирования проводят двойное баркодирование (индексирование) библиотек также с помощью метода ПЦР. Баркодирование — это «пришивание» последовательности, несущей индекс длиной восемь нуклеотидов и необходимой для последующего различения образцов из одной пробы при секвенировании сразу нескольких образцов [22]. Стоит отметить, что существует множество наборов реагентов для баркодирования при секвенировании на приборах Illumina, например Nextera XT Index kit. В нашем случае олигонуклеотиды были синтезированы в компании «Евроген» (Россия). Подбор условий ПЦР при баркодировании показал, что эффективность реакции не зависит от чистоты ДНК-матрицы и предварительная очистка образцов не требуется. Нами были подобраны оптимальные условия ПЦР для праймеров-индексов Nextera (табл. 3), они представлены в табл. 4.

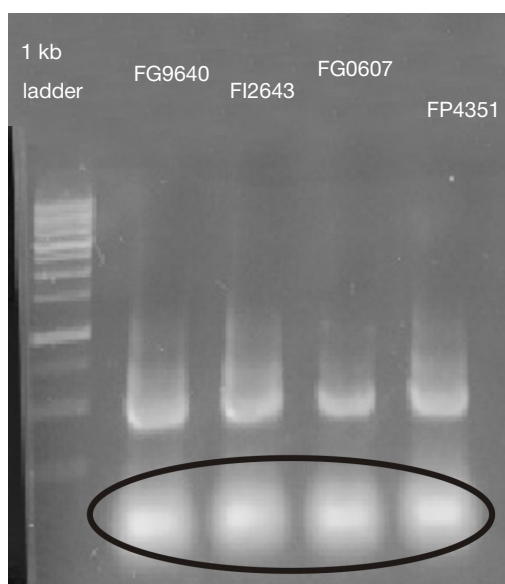
Необходимо отметить важность проведения контроля качества библиотек фрагментов ДНК во избежание получения ошибок на этапе секвенирования. Мы проводим контроль качества с помощью прибора Agilent Bioanalyzer 2100 (Agilent Technologies, США) (рис. 2).

### Секвенирование образцов

В данной статье мы не будем подробно разбирать особенности этапа секвенирования, поскольку оно проводится по стандартным протоколам компании-производителя [24, 25]. Мы проводим секвенирование образцов на приборе MiSeq компании Illumina согласно стандартным протоколам с парноконцевыми прочтениями длиной 250 п. н.

### Биоинформатический анализ

Полученные прочтения обрабатываются и классифицируются в соответствии с базой данных рибосомальной систематики RDP (Мичиганский университет, США) версии 11.5 при помощи пакета RDPTools версии 2016-07-21 [26, 27]. Порог условной вероятности корректности классификации (параметр `-conf`) устанавливается равной 50 %



**Рис. 1.** Агарозный гель после проведения электрофореза с продуктами ПЦР. В первую лунку нанесен маркер длин ДНК, в остальные лунки — образцы. Черным овалом выделены димеры праймеров, образующиеся при постановке ПЦР

Таблица 2. Условия проведения полимеразной цепной реакции для амплификации V3 и V4 регионов с универсальными праймерами для адаптера Nextera

Состав	V, мкл	Протокол		
		T, °C	t, сек	Количество циклов
Прямой праймер (C = 1 мкМ)	5,0	95	180	1
Обратный праймер (C = 1 мкМ)	5,0	95	30	25
Тақ-полимераза	0,5	55	30	
dNTP	1,5	72	30	
ПЦР-буфер	2,5			
ДНК	2,5	72	300	1
mQ	8,0			

Таблица 3. Праймеры-индексы Nextera

Название праймера	Последовательность праймера
i701	CAAGCAGAAGACGGCATAACGAGATTCGCCTTAGTCTCGTGGGCTCGG
i702	CAAGCAGAAGACGGCATAACGAGATCTAGTACGGTCTCGTGGGCTCGG
i703	CAAGCAGAAGACGGCATAACGAGATTTCTGCCTGTCTCGTGGGCTCGG
i704	CAAGCAGAAGACGGCATAACGAGATGCTCAGGAGTCTCGTGGGCTCGG
i501	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGGCAGCGTC
i502	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCTCGTCGGCAGCGTC
i503	AATGATACGGCGACCACCGAGATCTACACTATCCTCTTCGTCGGCAGCGTC
i504	AATGATACGGCGACCACCGAGATCTACACAGAGTAGATCGTCGGCAGCGTC

Таблица 4. Условия проведения полимеразной цепной реакции для баркодирования для праймеров-индексов Nextera

Состав	V, мкл	Протокол		
		T, °C	t, сек	Количество циклов
Индекс 1 (C = 1 мкМ)	5,0	95	180	1
Индекс 2 (C = 1 мкМ)	5,0	95	30	10
Тақ-полимераза	0,5	55	30	
dNTP	1,5	72	30	
ПЦР-буфер	2,5			
Ампликоны из 1-й ПЦР	5,0	72	300	1
mQ	5,5			

в соответствии с рекомендуемыми параметрами запуска, указанными в [28].

Необходимо отметить важность проведения контроля качества секвенирования. Так, не менее 95 % последовательностей в каждом образце должны быть высокого качества, т. е. не должны содержать адаптерных и иных загрязняющих последовательностей, не картирующихся на геном человека. Кроме того, следует уделять особое внимание количеству прочтений для каждого образца при метагеномном секвенировании. Не существует стандартного значения для данного критерия: число прочтений, прежде всего, зависит от цели исследования. Так, если цель исследования заключается в оценке только основных типов бактерий в образце, то достаточно небольшого числа прочтений (например, в 22 образцах, взятых из кишечника человека, при количестве прочтений 350 можно обнаружить два основных типа бактерий: 75 % *Firmicutes* и 18 % *Bacteroidetes*) [4]. Однако следует помнить, что чем больше прочтений выбрано, тем более малочисленные микроорганизмы удастся обнаружить в образцах и тем незначительнее влияние ошибки сэмпирования. Для описания метагенома с высоким разрешением количество прочтений обычно составляет 10 000 и выше [4].

Значения доли последовательностей, которые не удалось классифицировать, доли последовательностей, классифицированных как «неизвестная бактерия», медианной доли последовательностей, для которых надежно определен род и семейство, должны соответствовать ожидаемым результатам для анализа микробиального метагенома методом 16S рПНК. Так, доля неопределенных ридов не должна превышать ~20 %, семейство должно

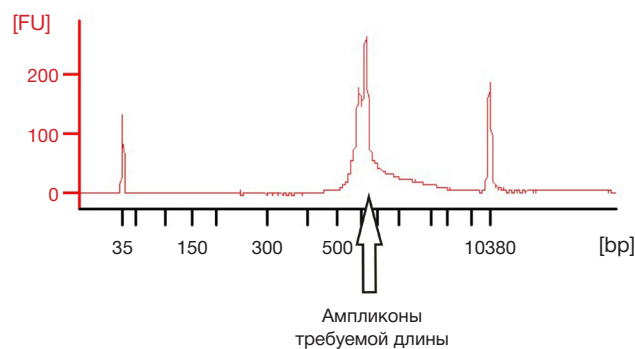


Рис. 2. Результат проверки качества библиотек для секвенирования на приборе Agilent Bioanalyzer 2100



Таблица 5. Вид представления результатов биоинформатического анализа

Rank	Name	rootrank	domain	phylum	class	order	family	genus	Sample 1	Sample 2	Sample ...
1_domain	Archaea	Root	Archaea								
2_phylum	Crenarchaeota	Root	Archaea	Crenarchaeota							
3_class	Thermoprotei	Root	Archaea	Crenarchaeota	Thermoprotei						
4_order	Acidilobales	Root	Archaea	Crenarchaeota	Thermoprotei	Acidilobales					
5_family	Acidilobaceae	Root	Archaea	Crenarchaeota	Thermoprotei	Acidilobales	Acidilobaceae				
6_genus	Acidilobus	Root	Archaea	Crenarchaeota	Thermoprotei	Acidilobales	Acidilobaceae	Acidilobus			

**Примечание.** Первый столбец Rank содержит таксономический уровень, второй столбец Name — название таксона. Эти столбцы предназначены для фильтрации данных. К примеру, управляя фильтром по столбцу Rank, можно получить представление результатов на уровне только видов, только родов и т. д. Следующие 7 столбцов (rootrank, domain, phylum, class, order, family, genus) содержат полный путь классификации таксона. Последние столбцы содержат данные об относительном содержании данного таксона в конкретном образце, в процентах. Название образцов указываются в заголовках столбцов.

определяться для ~70 % последовательностей и более, а род — для ~50 % и более. Однако, приведенные выше значения не являются стандартными и могут варьироваться от исследования к исследованию.

В результате анализа обычно формируется таблица, в которой указываются таксоны различного уровня (домен, тип, класс и т. д.) и содержится информация о таксономическом дереве и относительной представленности таксонов в каждом образце. Примером такой таблицы является табл. 5.

## ВЫВОДЫ

При проведении метагеномного секвенирования участков гена 16S рРНК важным этапом является подбор праймеров для амплификации, поскольку от качества полученных

ампликонов зависит точность последующего секвенирования. В данной статье нами описан подход к подбору оптимальных условий ПЦР на стадии подготовки образца, позволяющий самостоятельно подобрать условия для проведения ПЦР на этапе подготовки библиотек, и основные проблемы этапа. Хотя существует достаточное количество публикаций, посвященные созданию библиотек для метагеномного секвенирования [29], нами разработан конкретный протокол, подобраны оптимальные условия для амплификации регионов V3 и V4 16S рРНК с универсальными праймерами для последующего секвенирования на платформе Illumina. Стоит отметить, что качество амплификации зависит от многих факторов, в том числе и от качества используемых реагентов, поэтому условия проведения ПЦР не могут быть универсальными и варьируются в зависимости от реагентов.

## Литература

- Mori H, Maruyama F, Kato H, Toyoda A, Dozono A, Ohtsubo Y, et al. Design and Experimental Application of a Novel Non-Degenerate Universal Primer Set that Amplifies Prokaryotic 16S rRNA Genes with a Low Possibility to Amplify Eukaryotic rRNA Genes. *DNA Res.* 2014; 21 (2): 217–27.
- Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, Garcia-Cobos S, et al. Application of next generation sequencing in clinical microbiology and infection prevention. *J Biotechnol.* 2017 Feb 10; 243: 16–24.
- Oulas A, Pavloudi C, Polymenakou P, Pavlopoulos GA, Papanikolaou N, Kotoulas G., et al. Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinform Biol Insights.* 2015 May 5; 9: 75–88.
- Hamady M, Knight R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res.* 2009 Jul; 19 (7): 1141–52.
- Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol.* 2014 Sep; 12 (9): 635–45.
- Mardis ER. Next-generation sequencing platforms. *Annu Rev Anal Chem (Palo Alto Calif).* 2013; 6: 287–303.
- Fabrice A, Didier R. Exploring microbial diversity using 16S rRNA high-throughput methods. *J Comput Sci Syst Biol.* 2009; 2 (1): 74–92.
- Schloss PD, Gevers D, Westcott SL. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE.* 2011; 6 (12): e27310.
- Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 2013 Jan 7; 41 (1): e1.
- Illumina Inc., интернет-сайт [официальная веб-страница]. c2017 [дата обращения: 10 апреля 2017 г.]. Next-Generation Sequencing (NGS) [дата обращения: 10 апреля 2017 г.]; [примерно 2 экрана]. Доступно по: <https://www.illumina.com/technology/next-generation-sequencing.html>
- Bag S, Saha B, Mehta O, Anbumani D, Kumar N, Dayal M, et al. An Improved Method for High Quality Metagenomics DNA Extraction from Human and Environmental Samples. *Sci Rep.* 2016 May 31; 6: 26775.
- Qubit 3.0 Fluorometer User Guide [файл из интернета]. Версия А.0. CLLA: Thermo Fisher Scientific Inc. c2014 [дата обращения: 10 апреля 2017 г.]; 67 с. Доступно по: [https://tools.thermofisher.com/content/sfs/manuals/qubit\\_3\\_fluorometer\\_man.pdf](https://tools.thermofisher.com/content/sfs/manuals/qubit_3_fluorometer_man.pdf)
- Polz MF, Cavanaugh CM. Bias in template-to-product ratios in multitemplate PCR. *Appl Environ Microbiol.* 1998 Oct; 64 (10): 3724–30.
- Reysenbach AL, Giver LJ, Wickham GS, Pace NR. Differential amplification of rRNA genes by polymerase chain reaction. *Appl Environ Microbiol.* 1992 Oct; 58 (10): 3417–8.
- Bru D, Martin-Laurent F, Philippot L. Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example. *Appl Environ Microbiol.* 2008 Mar; 74 (5): 1660–3.
- Wu JH, Hong PY, Liu WT. Quantitative effects of position and type of single mismatch on single base primer extension. *J Microbiol Methods.* 2009 Jun; 77 (3): 267–75.

17. Seshadri R, Kravitz SA, Smarr L, Gilna P, Frazier M. CAMERA: a community resource for metagenomics. *PLoS Biol.* 2007 Mar; 5 (3): e75.
18. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013 Jan; 41 (Database issue): D590–6.
19. Mao D-P, Zhou Q, Chen C-Y, Quan Z-X. Coverage evaluation of universal bacterial primers using the metagenomic datasets. *BMC Microbiol.* 2012 May 3; 12: 66.
20. Brownie J, Shawcross S, Theaker J, Whitcombe D, Ferrie R, Newton C, et al. The elimination of primer-dimer accumulation in PCR. *Nucleic Acids Res.* 1997 Aug 15; 25 (16): 3235–41.
21. Lorenz TC. Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies. *J Vis Exp.* 2012 May 22; (63): e3998.
22. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* 2012 Jan; 40 (1): e3.
23. 16S Metagenomic Sequencing Library Preparation. Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System [файл из интернета]. [дата обращения: 10 апреля 2017 г.]: 28 с. Доступно по: [https://www.illumina.com/content/dam/illumina-support/documents/documentation/chemistry\\_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf](https://www.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf)
24. MiSeq System Denature and Dilute Libraries Guide [файл из интернета]. Сан-Диего: Illumina Inc.; c2016 [дата обращения: 10 апреля 2017 г.]: 14 с. Доступно по: [http://support.illumina.com/content/dam/illumina-support/documents/documentation/system\\_documentation/miseq/miseq-denature-dilute-libraries-guide-15039740-01.pdf](http://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/miseq/miseq-denature-dilute-libraries-guide-15039740-01.pdf)
25. Руководство по системе MiSeq [файл из интернета]. Сан-Диего: Illumina Inc.; c2015 [дата обращения: 10 апреля 2017 г.]: 106 с. Доступно по: [http://support.illumina.com/content/dam/illumina-support/documents/documentation/system\\_documentation/miseq/translations/miseq-system-guide-15027617-01-rus.pdf](http://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/miseq/translations/miseq-system-guide-15027617-01-rus.pdf)
26. База данных рибосомальной систематики (RDP) [база данных в интернете]. Релиз 11. Ист-Лансинг, США: Мичиганский университет. c1992–2014 [дата обращения: 10 апреля 2017 г.]. Доступно по: <https://rdp.cme.msu.edu/>.
27. Collection of commonly used RDP Tools for easy building. На GitHub [хостинг проектов]. c2017 [дата обращения: 10 апреля 2017 г.]. Доступно по: <https://github.com/rdpstaff/RDPTools>
28. RDP extensible sequence classifier for fungal lsu, bacterial and archaeal 16s. На GitHub [хостинг проектов]. c2017 [дата обращения: 10 апреля 2017 г.]. Доступно по: <https://github.com/rdpstaff/classifier/tree/701e229dde7cbe53d4261301e23459d91615999d>
29. Jones MB, Highlander SK, Anderson EL, Li W, Dayrit M, Klitgord N, et al. Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proc Natl Acad Sci U S A.* 2015 Nov 10; 112 (45): 14024–9.

## References

1. Mori H, Maruyama F, Kato H, Toyoda A, Dozono A, Ohtsubo Y, et al. Design and Experimental Application of a Novel Non-Degenerate Universal Primer Set that Amplifies Prokaryotic 16S rRNA Genes with a Low Possibility to Amplify Eukaryotic rRNA Genes. *DNA Res.* 2014; 21 (2): 217–27.
2. Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, Garcia-Cobos S, et al. Application of next generation sequencing in clinical microbiology and infection prevention. *J Biotechnol.* 2017 Feb 10; 243: 16–24.
3. Oulas A, Pavlouidi C, Polymenakou P, Pavlopoulos GA, Papanikolaou N, Kotoulas G., et al. Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinform Biol Insights.* 2015 May 5; 9: 75–88.
4. Hamady M, Knight R. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res.* 2009 Jul; 19 (7): 1141–52.
5. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol.* 2014 Sep; 12 (9): 635–45.
6. Mardis ER. Next-generation sequencing platforms. *Annu Rev Anal Chem (Palo Alto Calif).* 2013; 6: 287–303.
7. Fabrice A, Didier R. Exploring microbial diversity using 16S rRNA high-throughput methods. *J Comput Sci Syst Biol.* 2009; 2 (1): 74–92.
8. Schloss PD, Gevers D, Westcott SL. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE.* 2011; 6 (12): e27310.
9. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 2013 Jan 7; 41 (1): e1.
10. Illumina Inc., Web Site [homepage on the Internet]. c2017 [cited 2017 Apr 10]. Next-Generation Sequencing (NGS) [cited 2017 Apr 10]; [about 2 screens]. Available from: <https://www.illumina.com/technology/next-generation-sequencing.html>
11. Bag S, Saha B, Mehta O, Anbumani D, Kumar N, Dayal M, et al. An Improved Method for High Quality Metagenomics DNA Extraction from Human and Environmental Samples. *Sci Rep.* 2016 May 31; 6: 26775.
12. Qubit 3.0 Fluorometer User Guide [file on the Internet]. Revision A.0. US: Thermo Fisher Scientific Inc. c2014 [cited 2017 Apr 10]; 67 p. Available from: [https://tools.thermofisher.com/content/sfs/manuals/qubit\\_3\\_fluorometer\\_man.pdf](https://tools.thermofisher.com/content/sfs/manuals/qubit_3_fluorometer_man.pdf)
13. Polz MF, Cavanaugh CM. Bias in template-to-product ratios in multitemplate PCR. *Appl Environ Microbiol.* 1998 Oct; 64 (10): 3724–30.
14. Reysenbach AL, Giver LJ, Wickham GS, Pace NR. Differential amplification of rRNA genes by polymerase chain reaction. *Appl Environ Microbiol.* 1992 Oct; 58 (10): 3417–8.
15. Bru D, Martin-Laurent F, Philippot L. Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example. *Appl Environ Microbiol.* 2008 Mar; 74 (5): 1660–3.
16. Wu JH, Hong PY, Liu WT. Quantitative effects of position and type of single mismatch on single base primer extension. *J Microbiol Methods.* 2009 Jun; 77 (3): 267–75.
17. Seshadri R, Kravitz SA, Smarr L, Gilna P, Frazier M. CAMERA: a community resource for metagenomics. *PLoS Biol.* 2007 Mar; 5 (3): e75.
18. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013 Jan; 41 (Database issue): D590–6.
19. Mao D-P, Zhou Q, Chen C-Y, Quan Z-X. Coverage evaluation of universal bacterial primers using the metagenomic datasets. *BMC Microbiol.* 2012 May 3; 12: 66.
20. Brownie J, Shawcross S, Theaker J, Whitcombe D, Ferrie R, Newton C, et al. The elimination of primer-dimer accumulation in PCR. *Nucleic Acids Res.* 1997 Aug 15; 25 (16): 3235–41.
21. Lorenz TC. Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies. *J Vis Exp.* 2012 May 22; (63): e3998.
22. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* 2012 Jan; 40 (1): e3.
23. 16S Metagenomic Sequencing Library Preparation. Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System [file on the Internet]. [cited 2017 Apr 10]; 28 p. Available from: [https://www.illumina.com/content/dam/illumina-support/documents/documentation/chemistry\\_documentation/16s/16s-](https://www.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-)

- metagenomic-library-prep-guide-15044223-b.pdf
24. MiSeq System Denature and Dilute Libraries Guide [file on the Internet]. San Diego: Illumina Inc.; c2016 [cited 2017 Apr 10]: 14 p. Available from: [http://support.illumina.com/content/dam/illumina-support/documents/documentation/system\\_documentation/miseq/miseq-denature-dilute-libraries-guide-15039740-01.pdf](http://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/miseq/miseq-denature-dilute-libraries-guide-15039740-01.pdf)
  25. Rukovodstvo po sisteme MiSeq [file on the Internet]. San Diego: Illumina Inc.; c2015 [cited 2017 Apr 10]: 106 p. Available from: [http://support.illumina.com/content/dam/illumina-support/documents/documentation/system\\_documentation/miseq/translations/miseq-system-guide-15027617-01-rus.pdf](http://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/miseq/translations/miseq-system-guide-15027617-01-rus.pdf). Russian.
  26. The Ribosomal Database Project (RDP) [database on the Internet]. Release 11. East Lansing, MI: Michigan State University. c1992–2014 [cited 2017 Apr 10]. Available from: <https://rdp.cme.msu.edu/>
  27. Collection of commonly used RDP Tools for easy building. In: GitHub [Git repository hosting service]. c2017 [cited 2017 Apr 10]. Available from: <https://github.com/rdpstaff/RDPTools>
  28. RDP extensible sequence classifier for fungal lsu, bacterial and archaeal 16s. In: GitHub [Git repository hosting service]. c2017 [cited 2017 Apr 10]. Available from: <https://github.com/rdpstaff/classifier/tree/701e229dde7cbe53d4261301e23459d91615999d>
  29. Jones MB, Highlander SK, Anderson EL, Li W, Dayrit M, Klitgord N, et al. Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proc Natl Acad Sci U S A*. 2015 Nov 10; 112 (45): 14024–9.