

А. М. Курильщикова, Н. В. Тикунова, М. Р. Кабилов

Институт химической биологии и фундаментальной медицины СО РАН
пр. Акад. Лаврентьева, 8, Новосибирск, 630090, Россия

E-mail: kabilov@niboch.nsc.ru

МЕТОДЫ И ОБЪЕКТЫ МЕТАГЕНОМНЫХ ИССЛЕДОВАНИЙ *

Метагеномика – один из самых развивающихся разделов геномики, посвященный изучению генетического материала (метагенома) сообществ микроорганизмов в совокупности [1]. Объектами изучения метагеномики могут являться любые популяции микроорганизмов, обитающих в воде, почве, организме животного, человека или любой другой среде. Данное направление стало логическим продолжением геномики индивидуальных микроорганизмов, связанным с исследованием каждого генома в отдельности. Главной целью метагеномики является получение и анализ всех геномов для установления видового состава и метаболических взаимосвязей в сообществе [2]. Однако в настоящий момент эта цель практически недостижима по ряду причин.

Сборка даже одного бактериального генома является нетривиальной задачей, так как современные методы секвенирования позволяют получать нуклеотидную последовательность не целого генома, а его относительно коротких участков, из которых его еще необходимо собирать. Если же количество анализируемых микроорганизмов в сообществе превышает несколько тысяч, то задача сборки из трудной превращается в практически неразрешимую [3]. Однако в сборке полных геномов, как правило, нет необходимости [4], так как зачастую информацию о видовом составе популяции

можно получить из анализа отдельных генов, который также позволяет выстроить и метаболические сети [5; 6].

Важной особенностью метагеномных исследований можно считать отсутствие необходимости в изоляции и культивировании микроорганизмов, что является принципиальным моментом, поскольку не все из них растут на микробиологических средах. К тому же это позволяет включить в анализ присутствующие в популяции вирусы и бактериофаги, что, несомненно, расширяет представление о метагеноме.

Методы секвенирования

Первичной информацией для метагеномных исследований являются нуклеотидные последовательности, получаемые при секвенировании нуклеиновых кислот (НК). Золотым стандартом секвенирования до сих пор является метод Сэнгера, в основе которого лежит статистическое терминирование флуоресцентно меченными дидезоксирибонуклеозид трифосфатами растущей цепи при элонгации праймера с помощью термостабильной ДНК-полимеразы. Полученные меченные ДНК-фрагменты разделяются по длине капиллярным гель-электрофорезом, а с помощью лазера детектируется флуоресцентная метка на их 3'-конце, соответствующая одному из четырех нуклеотидов.

* Работа выполнена при финансовой поддержке Министерства образования и науки РФ (ГК 16.512.11.2113).

Автоматические капиллярные секвенаторы разработаны несколькими фирмами («GE-MegaBACE», «Beckman Coulter – SEQ»), однако наибольшее распространение получили приборы «Applied Biosystems». Их последняя модель секвенатора, 3 500xl с 24 капиллярами позволяет за 2 ч прочитать около 1 000 нуклеотидов с каждого образца, что суммарно составляет почти 0,3 Мб/день.

Однако поскольку размер метагенома зачастую может превосходить даже размер генома человека, использование метода Сэнгера неоправданно с точки зрения временных и финансовых затрат. В последние годы на первый план вышло высокопроизводительное параллельное секвенирование (ВПС), позволяющее получать миллиарды нуклеотидов в день. Появление платформ ВПС нуклеиновых кислот стало новым «толчком» к развитию метагеномных исследований. В настоящий момент используются три основных технологии ВПС, обладающие максимальной производительностью: 454 («Roche»), SOLiD («Applied Biosystems») и HiSeq («Illumina»). Неоспоримыми плюсами данных приборов является стоимость секвенирования одного нуклеотида, объем получаемых данных и скорость их получения. Все три технологии ВПС можно свести к нескольким этапам, а именно: получение библиотеки ДНК-фрагментов, ее амплификация и определение нуклеотидных последовательностей.

На первом этапе проводят статистическое фрагментирование анализируемой ДНК с помощью ультразвука или других методов с последующим присоединением к полученным ДНК-фрагментам олигонуклеотидных адаптеров с известной последовательностью. Для разных платформ и разных типов библиотек (*fragment*, *pair-end* и *mate-paired*) адаптеры, необходимые для дальнейшей амплификации фрагментов метагенома, добавляются различными способами.

Второй этап предполагает проведение амплификации каждого из этих фрагментов с помощью полимеразной цепной реакции (ПЦР), но не просто в смеси, а физически изолированно друг от друга. Разобщенность фрагментов ДНК обеспечивается с помощью твердофазной иммобилизации одного из двух праймеров, участвующих в ПЦР. В случае секвенаторов 454 и SOLiD праймер присоединен к микрошарику. Для физической изоляции микрошариков друг от друга,

необходимой для того, чтобы на поверхности каждого из них был амплифицирован только один фрагмент, используется эмульсионная полимеразная цепная реакция (ePCR). Создается микроэмульсия (*water in oil*), т. е. водные микрореакторы, которые располагаются в масле; в идеальном случае каждый из них будет содержать один микрошарик и один ампликон. В результате получается набор микрошариков, каждый из которых должен содержать на своей поверхности только один амплифицированный фрагмент ДНК. В случае секвенатора 454 микрошарики помещаются в фиксированное количество лунок с определенной геометрией на поверхности проточного чипа. Стекланный чип используется и в SOLiD, однако в этом случае микрошарики случайным образом ковалентно иммобилизуются уже на ровной поверхности, их максимальное количество ограничивается только способностью детектирующей системы различить два соседних микрошарика [7].

В случае секвенаторов «Illumina» на поверхности чипа кластерно иммобилизуется праймер. Одиночная молекула ДНК-фрагмента с присоединенными адаптерами при денатурации попадает на один из кластеров и связывается с праймером. В результате ПЦР каждый из кластеров содержит иммобилизованные копии фрагмента. Таким образом, на всех платформах в конце второго этапа имеется проточный чип с физически изолированными на его поверхности клонами ДНК-фрагментов, которые далее и будут секвенироваться [7].

Принципы секвенирования, которые используются на представленных платформах, значительно отличаются. Общим можно считать только наличие пошагового считывания каждого типа буквы с помощью чувствительных CCD-камер. На платформе 454 используется пиросеквенирование, основанное на последовательном пошаговом пропускании через чип каждого из четырех дезоксирибонуклеотид трифосфатов (dNTP) с ансамблем ферментов и детекцией соответствующего количественного хемиллюминесцентного сигнала. Сигнал формируется в результате каскадного превращения $dNTP \rightarrow PPi \rightarrow ATP \rightarrow \text{свет}$ с помощью ДНК-полимеразы, апиразы и люцефиразы соответственно, где PPi – пиродифосфат, ATP – аденозинтрифосфат. Таким образом, при пропускании каждого типа dNTP прибор

детектирует люминесценцию в лунках и соотносит ее с положением на чипе, причем количество света пропорционально количеству встроившихся букв. Длина чтения может достигать 1 000 нуклеотидов, что вполне сравнимо с методом Сэнгера. К недостаткам платформы 454 можно отнести проблемы с чтением гомоповторов, что связано с нелинейностью люминесцентного сигнала при одновременном встраивании в цепь большого числа дезоксинуклеотид трифосфатов одного типа [7].

На платформе «Illumina» используется так называемое секвенирование при синтезе (Sequencing by Synthesis). ДНК-полимераза встраивает один из четырех типов dNTP, каждый из которых отмечен соответствующим флуорофором, наличие которого, с одной стороны, не позволяет встраиваться следующей букве, а с другой – дает возможность идентифицировать нуклеотид. Встроившаяся буква по флуоресценции соотносится с конкретным кластером на поверхности чипа. После детекции сигналов слайд обрабатывается реагентами, которые удаляют флуоресцентные метки, что позволяет начать новый цикл чтения. Эффективность встраивания нуклеотидов по мере элонгации цепи снижается, что ограничивает максимальную длину прочтения. В настоящий момент для HiSeq 2000 максимальная длина чтения составляет 100 букв.

В отличие от всех остальных подходов к секвенированию нуклеиновых кислот, в платформе SOLiD 5500x1 (Sequencing by Oligonucleotide Ligation and Detection) используется не ДНК-полимераза, а ДНК-лигаза – фермент, образующий ковалентную связь между 5'-фосфатом и 3'-гидроксильной группой в одноцепочечном разрыве ДНК-дуплекса. Все иммобилизованные на поверхности микрошариков одноцепочечные фрагменты ДНК первоначально формируют комплементарный комплекс с универсальным адаптером. Для чтения нуклеотидной последовательности используется набор олигонуклеотидов следующего вида: 3'-XYNNNZZZ-F-5', где XY – один из 16 возможных динуклеотидов; N – любой нуклеотид (вырожденная буква); Z – универсальное основание; F – один из четырех флуорофоров. Один флуорофор кодирует четыре различных динуклеотида (XY, X'Y',

YX, Y'X', где $X < > Y$ и отдельная группа XX, X'X'), например AG, TC, GA и CT. Из добавленного набора зондов с ДНК-матрицей гибридизуется олигонуклеотид, содержащий комплементарные димер (XY) и тример (NNN). ДНК-лигаза формирует фосфодиэфирную связь между универсальным праймером и комплементарным зондом. Далее происходит считывание флуоресцентных сигналов и соотнесение их с иммобилизованными на чипе микрошариками. Для начала нового шага лигирования удаляются три нуклеотида и флуорофор с 5'-конца (ZZZ-F). После нескольких шагов лигирования проводится денатурация и удаление комплементарной цепи с микрошариков, что позволяет начать новый цикл секвенирования, но делается это уже с использованием адаптера, смещенного на один нуклеотид. Смещение адаптера ($n, n - 1, n - 2, n - 3, n - 4$) в каждом новом цикле позволяет выстроить последовательность из флуоресцентных сигналов (*color space*) и восстановить нуклеотидную последовательность, соответствующую каждой микросфере. Тот факт, что каждая буква дважды читается перекрывающимися динуклеотидами, значительно повышает точность секвенирования [7].

Обобщенная информация о каждой из описанных секвенирующих платформ представлена в таблице. Наиболее производительными приборами на текущий момент являются HiSeq 2000 и SOLiD, позволяющие за один запуск получить до 600 и 180 гигабайт данных соответственно. Общим для обеих платформ является относительно короткая длина прочтения, пока что не превышающая 100 нт. Платформа 454 FLX+ генерирует значительно меньше данных, однако у нее есть большое преимущество – средняя длина прочтения, составляющая 700 нуклеотидов, что вполне сравнимо со средней длиной прочтения методом Сэнгера.

В настоящий момент на рынке присутствуют и другие коммерчески доступные секвенаторы, которые также относятся к ВПС: Ion Torrent («Applied Biosystems»), MiSeq («Illumina») и Junior («Roche»). Однако производительность этих приборов значительно ниже, и позиционируются они, скорее, для изучения отдельных геномов, чем для метагеномных исследований.

Современные модели секвенаторов максимальной производительности

Метод	Производитель	Модель секвенатора	Максимальная длина прочтения, нт	Суммарное прочтение за один запуск	Время работы
Секвенирование по Сэнгер	Applied Biosystems	3500xl	1 000	2 400 нт	2 ч
Пиросеквенирование	Roche	GS FLX+	1 000	0,7 Гб	23 ч
Секвенирование лигированием	Applied Biosystems	SOLiD 5500xl	75	180 Гб	6 дней
Секвенирование при синтезе	Illumina	HiSeq 2000	100	600 Гб	11 дней

Метагеномные исследования

Определение видового состава. На начальных этапах развития метагеномики в первую очередь изучался видовой состав сообществ. Для этой цели до сих пор анализируют нуклеотидные последовательности генов 16S и 18S рибосомальной РНК (рРНК), отличающиеся высокой степенью консервативности, что позволяет определять филогенетическую принадлежность прокариот и эукариот соответственно [8]. С помощью праймеров к константным районам гена рРНК проводят ПЦР и получают набор ДНК-фрагментов. Однако секвенирование по Сэнгеру предполагает, что каждый ампликон должен секвенироваться отдельно, что приводит к необходимости их изоляции друг от друга. Классическим способом разделения является клонирование анализируемых фрагментов в плазмидных векторах [9]. В последовательность праймеров с 5'-конца вводится дополнительный участок, необходимый для формирования сайта рестрикции, который позволит провести лигирование ПЦР-фрагмента в вектор. В настоящее время появились и другие подходы, позволяющие разделять фрагменты, однако распространения в метагеномике они пока не получили. В первую очередь, это цифровая ПЦР (digital PCR), основанная на разбавлении анализируемой ДНК до уровня единичных молекул и проведении большого количества параллельных ПЦР [10; 11]. Кроме того, существует метод молекулярных колоний, принцип которого заключается в получении отдельных колоний ДНК-фрагментов при проведении ПЦР в полиакриламидном геле на подложке [12]. Оба ме-

тода потенциально могут позволить получить отдельные молекулы ПЦР-фрагментов для дальнейшего секвенирования.

С появлением платформы 454, длина прочтения на которой вполне сопоставима с методом Сэнгера, стало возможным одновременное секвенирование сотни тысяч ПЦР-фрагментов. Это позволило за несколько дней получать информацию о видовом составе микробиоты [13].

Возможным минусом использования праймеров на гены рРНК для таксономической идентификации может быть тот факт, что их консенсусные последовательности получены исходя из анализа уже известных бактериальных генов. Это потенциально может привести к сложностям выявления тех микроорганизмов, сиквенсы которых отличаются и еще не известны. Также необходимо понимать, что анализ 16S и 18S генов метагенома не позволяет говорить о наличии вирусов и бактериофагов, для которых подобные универсальные консервативные нуклеотидные последовательности отсутствуют.

ВАС-клонирование. Исторически термин «метагеном» впервые использован в работе J. Handelsman и соавт. [14], в которой из почвы выделялась тотальная ДНК, обрабатывалась рестриктазами, после чего полученные фрагменты ДНК клонировались в ВАС-векторах. ВАС-клонирование позволяло работать с индивидуальными протяженными фрагментами размером до нескольких сот тысяч нуклеотидов, однако даже сотни таких фрагментов, как правило, не сравнимы с размерами метагенома.

Shotgun. Одним из первых продуктивных подходов к секвенированию метагенома

был метод «дробовика» (*shotgun sequencing*), основанный на фрагментации ДНК, клонировании полученных коротких фрагментов и их секвенировании методом Сэнгера. Исходно основным назначением этого метода было проведение полногеномного секвенирования индивидуальных организмов, однако и для метагеномных исследований он оказался вполне применим. Shotgun позволяет анализировать метагеном независимо от типа микроорганизмов и вирусов. Одна из первых работ с использованием shotgun-секвенирования выполнена при изучении морской вирусной микробиоты [15]. Среди недостатков подхода необходимо отметить трудоемкость и дороговизну. Также следует указать, что некоторая часть фрагментов ДНК не клонируется вследствие цитотоксичности.

Платформа 454 используется в метагеномных исследованиях не только для изучения видового состава микробиома, но и как более эффективный аналог shotgun-секвенирования [16; 17]. Полученные нуклеотидные последовательности, как в случае shotgun, так и при пиросеквенировании, собираются в более длинные фрагменты (контиги) с помощью специальных алгоритмов. Существует несколько вариантов дальнейшего анализа. Во-первых, это анализ видового разнообразия, например генов 16/18S рРНК; во-вторых, поиск генов, основанный на алгоритмах обнаружения открытых рамок считывания (ORF). Из найденных генов строятся метаболические сети, характеризующие микробиом как единый симбионтный надорганизм.

SOLiD и HiSeq. Как указывалось ранее, принципиальными характеристиками секвенаторов SOLiD и HiSeq являются их высокая производительность и короткая длина читаемых фрагментов. Сотни миллиардов нуклеотидов, выдаваемых этими приборами, соответствуют объемам заложенной информации в метагеноме. Сложность в использовании этих платформ заключается в короткой длине получаемых последовательностей, что затрудняет дальнейший анализ, в особенности анализ последовательностей, не имеющих гомологов в базах данных. Однако к настоящему моменту уже накоплен большой объем информации как по полным геномам, так и по отдельным генам, что делает использование этих платформ все более перспективным [18; 19].

Метагеномика природных экосистем

Природные экосистемы, как естественные, так и искусственные, представляют большой интерес для метагеномных исследований, поскольку микроорганизмы являются обязательными участниками важных биологических процессов. Разнообразие окружающего нас микромира только начинает изучаться в рамках экометагеномики. Три огромных источника биоразнообразия – вода, атмосфера и почва – являются основными объектами ее изучения.

Метагеномные исследования экосистем, существующих в экстремальных условиях (черные курильщики, соленые озера, арктические моря и т. д.), представляют особый интерес, поскольку обитающие там микроорганизмы обладают необычными метаболическими путями, которые нередко находят применение в биотехнологии. Так, именно в термальных источниках открыты термостабильные ДНК-полимеразы, повсеместно используемые в молекулярной биологии и диагностике. Природные источники полезных ископаемых также являются крайне важными объектами для метагеномики, поскольку изучение возможных механизмов образования полезных ископаемых и бактерий, обнаруженных в них, может предоставить важную информацию для разработки перспективных биокатализаторов. Микроорганизмы в окружающей среде нередко несут не только пользу, но и вред. В частности, они являются причиной биогенной коррозии металлов и биодegradации других материалов. Изучение этих сообществ позволит в будущем разработать способы борьбы с ними [20].

Таким образом, дальнейшие исследования, проводимые в рамках экометагеномики, обеспечат получение фундаментальных знаний в области исследования эволюции микробных сообществ, а также послужат основой для прикладных разработок.

Водные экосистемы. Изучение микробных сообществ в водных биоценозах стало первой областью применения метагеномного подхода. По количеству видов микроорганизмов водные экосистемы значительно превышают любые другие. Первая метагеномная работа, посвященная океаническим экосистемам океана, опубликована в 1996 г. [21]. Авторы проводили изучение архей

и для идентификации видового состава популяции использовали тактику клонирования в плазмиде фрагмента ДНК, содержащего ген рибосомной РНК.

Опубликованная в 2004 г. статья, представляющая данные об изучении сообществ микроорганизмов в районе Саргассового моря, увеличила количество известных человеку последовательностей генов почти вдвое – на 1,2 млн [22]. Длина последовательностей, полученных в результате этого исследования, составила 1,5 млрд нуклеотидов. В 2007 г. база данных генов увеличилась еще на 6,25 млрд нуклеотидов как результат значительного по объему метагеномного исследования, охватившего 41 различных регион мирового океана. Исследование метагеномов океана дало информацию для разработки новых лекарств [23], позволило описать многие гены, участвующие в ряде неисследованных метаболических путей [24; 25]. Изучение микробиотической структуры сообществ мирового океана по-прежнему является огромным полем для метагеномных исследований и таит множество информации как фундаментального, так и прикладного характера.

Метагеномика почвы. Почва – это сложная и структурированная экосистема, характеризующаяся очень высоким биоразнообразием [26]. Почвы состоят преимущественно из твердофазных частиц, на поверхности и внутри которых находятся микроорганизмы, образуя кластеры из живой и неживой материи [27]. Количество видов, обитающих на небольшой площади, огромно. Показано, что в 30 г почвы обитает более полумиллиона видов [28]. Однако в культуре способны расти только 0,3 % от всех почвенных микроорганизмов [29].

Столь огромное биоразнообразие является богатым источником новых лекарств – антибиотиков [30], противораковых препаратов [31], иммуносупрессантов [32] и т. д. Кроме того, почва – очень богатый источник микроорганизмов, перспективных для использования в биотехнологической промышленности [33]. К настоящему времени метагеномные исследования почв позволили обнаружить значительное количество новых ферментов – липаз, амилаз, амидаз, оксидоредуктаз, биокатализаторов [34]. Причем практически каждое новое исследование открывает новые белки, тем самым убедительно доказывая тот факт, что биоразнооб-

разие почв на порядки превышает наши знания о нем.

Метагеномика воздушных сред. Биоразнообразие в воздушной среде значительно превышает наши теоретические предположения. Согласно последним оценкам, культивированию поддаются менее чем 1 % микроорганизмов, обитающих в воздухе [35]. Таким образом, использование метода секвенирования некультивированных образцов позволяет расширить представления о микробиологическом составе воздуха более чем в 100 раз.

Обитание в воздушной среде связано для микроорганизмов с рядом адаптивных метаболических путей, нехарактерных для других сообществ. В частности, это касается путей защиты от окисления свободным кислородом воздуха. Следовательно, изучение таких метаболических путей возможно в первую очередь в экосистемах воздушных сред.

Метагеномика человека

До начала проекта «Геном человека» предполагалось, что число генов в человеческом геноме составляет около 100 000. Однако после секвенирования, сборки и аннотации генома человека обнаружено всего лишь около 20 000 генов, кодирующих белки, почти столько же, сколько у плодовой мушки *Drosophila melanogaster*. Вместе с тем, если расширить представление о том, что есть организм человека, и включить в это определение еще и симбиотические организмы, то оценка в 100 000 генов окажется сильно заниженной. Можно сказать, что микробиом – это отдельный орган человеческого организма, столь же обязательный и незаменимый, как и другие [36].

Микробиом человека является одним из основных и наиболее важных объектов изучения метагеномики. Для детального изучения микробиоты человека требуется проведение метагеномных исследований множества выборок образцов, полученных от разных людей в разных регионах в разное время.

В организме человека есть несколько основных зон максимальной концентрации микроорганизмов. Это респираторный и пищеварительный тракты, органы мочеполовой системы, а также кожа. Наиболее изученными в настоящее время являются

метагеномы желудочно-кишечного тракта и верхних дыхательных путей, поскольку именно они вовлечены в наибольшее число процессов, влияющих на состояние организма.

Метагеномика респираторного тракта. Бронхолегочная система и, прежде всего, верхние дыхательные пути содержат высоко разнообразное сообщество микроорганизмов. Значительная их доля является по отношению к человеку комменсалистами, однако широкий спектр бактерий и вирусов, находящихся постоянно или попадающих в респираторный тракт из окружающей среды, являются патогенными и условно-патогенными. Известно, что верхние дыхательные пути являются основным местом инвазии инфекции в организм человека, и, согласно медицинской статистике, наиболее часто регистрируемыми инфекционными заболеваниями являются именно вирусные и бактериальные инфекции респираторного тракта.

Основными бактериальными агентами, находящимися в бронхолегочном тракте, являются такие бактерии, как *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Neisseria meningitidis* и *Staphylococcus aureus*. При этом колонизация слизистых верхних дыхательных путей этими условно-патогенными бактериями, как правило, проходит бессимптомно, хотя их наличие увеличивает риск возникновения пневмоний, сепсиса и даже менингитов. Механизмы такого взаимодействия микроорганизмов в настоящее время неизвестны.

Среди патогенных вирусов наиболее известным является вирус гриппа. Благодаря высокой мутационной способности он регулярно вызывает эпидемии. Другой не менее известной вирусной инфекцией, передающейся воздушно-капельным путем, является натуральная оспа. К счастью, в настоящий момент эта инфекция условно побеждена благодаря тотальной вакцинации. Тем не менее человечество не застраховано от возникновения новых вирусов, не менее опасных, чем оспа. Поэтому изучение метагеномов респираторного тракта является одной из важнейших задач современной медицины [37].

Метагеномика желудочно-кишечного тракта. Микробиом ЖКТ является наиболее разнообразным по видовому составу и количеству метаболических путей среди

всех симбиотических систем человеческого тела. Согласно некоторым оценкам, микробиота среднестатистического человека состоит из 10–100 трлн индивидуальных прокариотических клеток, относящихся к 150–800 видам из 1 380–4 000 различных таксономических групп [38–41]. Пищеварительный тракт человека населен преимущественно бактериями (90–95 % генетического материала). Доля вирусных геномов составляет примерно 5 %, геномов архей – до 1 %, других эукариот без учета генома человека – не более 1 % [42]. Большинство бактерий относятся к родам *Bacteroides* (до 30 % всех бактерий в кишечнике), *Clostridium*, *Fusobacterium*, *Eubacterium*, *Ruminococcus*, *Peptococcus*, *Peptostreptococcus* и *Bifidobacterium*. Представительство различных бактерий в кишечной микробиоте сильно варьирует от человека к человеку.

Несмотря на высокую вариабельность микрофлоры кишечного тракта, в ходе последних метагеномных исследований удалось выявить существование всего трех стабильных типов кишечной микробиоты – «базовых» микробиомов («core microbiome»), названных энтеротипами, не зависящих от местообитания человека, его нации и расы, состояния здоровья, пола и биологических показателей [42]. Различия в энтеротипах заключаются в представленности определенных метаболических путей синтеза и спектра преобладающих бактерий: в первом энтеротипе преобладают *Bacteroides*, во втором – *Prevotella*, в третьем – *Ruminococcus*. Очевидно, что этими различиями не ограничивается вариабельность возможных метаболических путей, и можно ожидать дальнейших открытий в этой области.

Кишечная микробиота участвует в метаболических процессах усвоения питательных веществ, является источником некоторых метаболитов, которые человек не может синтезировать сам, влияет на состояние иммунитета и здоровье человека в целом [43]. Одним из наиболее очевидных примеров является усвоение короткоцепочечных жирных кислот, единственным источником получения которых является симбиотическая микробиота кишечника [40; 44].

Микробы вовлечены во множество патологических процессов, происходящих в человеческом организме. В частности, показано, что патогенез ожирения связан с мик-

роорганизмами и сопровождается изменением удельного количества *Bacteroides* и *Firmicutes* [45]. Язвенная болезнь желудка в настоящее время считается следствием жизнедеятельности бактерии *Helicobacter pylori*, пациенты проходят лечение антибиотиками [46].

Показано влияние микробиоты на развитие рака кишечника [47]. Эксперименты на мышцах доказали, что наличие в желудочно-кишечном тракте бактерии *Bacteroides vulgatus* стимулирует развитие рака толстой кишки [48]. Инфицирование мышей бактерией *Helicobacter hepaticus* увеличивает вероятность возникновения аденокарциномы молочных желез [49]. Механизмы связи канцерогенеза и бактерий известны лишь частично, но можно предположить наличие влияния хронического воспалительного процесса в кишечнике, вызванного патогенами, на вероятность возникновения новообразований. Другой возможный механизм индукции канцерогенеза связан с метаболическими путями микробиома, синтезирующими канцерогены [50].

Между развитием рака и микрофлорой существуют примеры не только положительной, но и отрицательной связи. Этот эффект обусловлен преимущественно ферментативной активностью некоторых бактерий, препятствующей активации проканцерогенов [51], в частности лактобактерий *L. casei* и *L. acidophilus* [52; 53].

Однако изучение влияния отдельных бактерий на развитие заболеваний – это лишь первый шаг к пониманию роли микробиоты в патогенезе. Бактерии и вирусы существуют не изолированно друг от друга, они взаимодействуют между собой, обмениваются генетической информацией. Для понимания этих процессов требуются дальнейшие детальные исследования, поэтому можно ожидать значительное увеличение количества метагеномных исследований микрофлоры ЖКТ уже в ближайшем будущем.

Консорциумы

Крупнейшими объединениями по изучению человеческого микробиома являются европейский консорциум MetaHit и американский HMP. MetaHit основан в 2008 г. и финансируется из средств Еврокомиссии.

В консорциум в настоящее время входят 13 промышленных и исследовательских организаций из восьми стран Европы. Консорциум HMP (Human Microbiome Project) основан как инициатива НИИ (National Institute of Health, США). HMP финансирует метагеномные работы, проводимые в США. Оба проекта подразумевают изучение человеческого микробиома в самых различных направлениях: создание баз данных генов и геномов микробиома человека, поиск взаимосвязей между микробиомом и заболеваниями человека, разработка новых методов анализа метагеномных данных и т. д. В ряде стран, в частности в Канаде и России, существуют группы по метагеномным исследованиям. В России такое объединение появилось в 2009 г. под названием «Русский метагеномный проект». В настоящее время в нем участвуют 14 организаций-членов. Главное международное объединение, в которое входят практически все консорциумы, – International Human Microbiome Consortium.

Заключение

Изучение отдельных микроорганизмов является необходимым, но не достаточным условием понимания микробиологических экосистем и не может дать ответы на основополагающие вопросы. Какие организмы образуют данное сообщество? Как они взаимодействуют между собой и окружающей средой? Метагеномика, изучающая совокупность всех генов и геномов сообщества, в перспективе может дать ответы на эти вопросы.

Один из важнейших объектов изучения метагеномики – симбиотический микробиом человека. Он представляет собой не просто совокупность микроорганизмов, обитающих в теле человека и на его поверхности, но сложную и многокомпонентную систему с внутренней структурой, динамикой, активно взаимодействующую по ряду аспектов с организмом хозяина. Патогенез множества заболеваний и одновременно способы их лечения прямо или косвенно связаны с ферментативной и биохимической активностью микрофлоры и ее влиянием на организм человека. Все это и объясняет возрастающий интерес к метагеномным исследованиям.

Список литературы

1. Gilbert J. A., Dupont C. L. Microbial Metagenomics: Beyond the Genome // Annual Review of Marine Science. 2011. Vol. 3, № 1. P. 347–371.
2. Simon C., Daniel R. Metagenomic Analyses: Past and Future Trends // Applied and Environmental Microbiology. 2011. Vol. 77, № 4. P. 1153–1161.
3. Haiminen N., Kuhn D. N., Parida L., Rigoutsos I. Evaluation of Methods for de novo Genome Assembly from High-Throughput Sequencing Reads Reveals Dependencies that Affect the Quality of the Results // PLoS ONE. 2011. Vol. 6, № 9. E24182.
4. Raes J., Korb J. O., Lercher M. J., Mering C. von, Bork P. Prediction of Effective Genome Size in Metagenomic Samples // Genome Biology. 2007. Vol. 8, № 1. R. 10.
5. Wooley J. C., Godzik A., Friedberg I. A Primer on Metagenomics // PLoS Computational Biology. 2010. Vol. 6, № 2. E1000667.
6. Liu B., Pop M. MetaPath: Identifying Differentially Abundant Metabolic Pathways in Metagenomic Datasets // BMC Proc. 2011. Vol. 5, Suppl 2. P. 9.
7. Pettersson E., Lundeberg J., Ahmadian A. Generations of Sequencing Technologies // Genomics. 2009. Vol. 93, № 2. P. 105–111.
8. Rajendhran J., Gunasekaran P. Microbial Phylogeny and Diversity: Small Subunit Ribosomal RNA Sequence Analysis and Beyond // Microbiol. Res. 2011. Vol. 166, № 2. P. 99–110.
9. Ward D. M., Weller R., Bateson M. M. 16S rRNA Sequences Reveal Numerous Uncultured Microorganisms in a Natural Community // Nature. 1990. Vol. 345. P. 63–65.
10. Vogelstein B., Kinzler K. W. Digital PCR // PNAS. 1999. Vol. 96, № 16. P. 9236–9241.
11. Tadmor A. D., Ottesen E. A., Leadbetter J. R., Phillips R. Probing Individual Environmental Bacteria for Viruses by Using Microfluidic Digital PCR // Science. 2011. Vol. 333. P. 58–62.
12. Четверина Е. В., Четверин А. Б. Наноклонии: обнаружение, клонирование и анализ индивидуальных молекул // Успехи биол. химии: ежегодник. 2008. Т. 48. С. 3–64.
13. Wu G. D., Lewis J. D., Hoffmann C., Chen Y. Y., Knight R., Bittinger K., Hwang J., Chen J., Berkowsky R., Nessel L., Li H., Bushman F. D. Sampling and Pyrosequencing Methods for Characterizing Bacterial Communities in the Human Gut Using 16S Sequence Tags // BMC Microbiol. 2010. Vol. 10. P. 1–14.
14. Handelsman J., Rondon M. R., Brady S. F., Clardy J., Goodman R. M. Molecular Biological Access to the Chemistry of Unknown Soil Microbes: A New Frontier for Natural Products // Chemistry & Biology. 1998. Vol. 5. P. 245–249.
15. Breitbart M., Salamon P., Andresen B., Mahaffy J. M., Segall A. M., Mead D., Azam F., Rohwer F. Genomic Analysis of Uncultured Marine Viral Communities // Proc. Nat. Acad. USA. 2002. Vol. 99, № 22. P. 14250–14255.
16. Edwards R. A., Rodriguez-Brito B., Wegley L., Haynes M., Breitbart M., Peterson D. M., Saar M. O., Alexander S., Alexander E. C., Rohwer F. Using Pyrosequencing to Shed Light on Deep Mine Microbial Ecology // BMC Genomics. 2006. Vol. 7. P. 1–13.
17. Dong Q., Nelson D. E., Toh E., Diao L., Gao X., Fortenberry J. D., Pol B. van der. The Microbial Communities in Male First Catch Urine Are Highly Similar to Those in Paired Urethral Swab Specimens // PLoS ONE. 2011. Vol. 6, № 5. ID 19709.
18. Lazarevic V., Whiteson K., Huse S., Hernandez D., Farinelli L., Osteras M., Schrenzel J., Francois P. Metagenomic Study of the Oral Microbiota by Illumina High-Throughput Sequencing // J. Microbiol. Methods. 2009. Vol. 79, № 3. P. 266–271.
19. Gilbert J. A., Meyer F., Antonopoulos D., Balaji P., Brown C. T., Brown C. T., Desai N., Eisen J. A., Evers D., Field D., Feng W., Huson D., Jansson J., Knight R., Knight J., Kolkner E., Konstantindis K., Kostka J., Kyrpides N., Mackelprang R., McHardy A., Quince C., Raes J., Sczyrba A., Shade A., Stevens R. Meeting Report: The Terabase Metagenomics Workshop and the Vision of an Earth Microbiome Project // Standards in Genomic Sciences. 2010. Vol. 3. P. 243–248.
20. Struchtemeyer C. G., Davis J. P., Elshahed M. S. Influence of the Drilling Mud Formulation Process on the Bacterial Communities in Thermogenic Natural Gas Wells of the Barnett Shale // Appl. Environ. Microbiol. 2011. Vol. 77, № 14. P. 4744–4753.
21. Stein J. L., Marsh T. L., Wu K. Y., Shizuya H., DeLong E. F. Characterization of Uncultivated Prokaryotes: Isolation and Analysis of a 40-Kilobase-Pair Genome Fragment Front

a Planktonic Marine // *Archaeon. J. Bacteriol.* 1996. Vol. 178. P. 591–599.

22. *Venter J. C., Remington K., Heidelberg J. F., Halpern A. L., Rusch D., Eisen J. A., Wu D. Y., Paulsen I., Nelson K. E., Nelson W., Fouts D. E., Levy S., Knap A. H., Lomas M. W., Nealson K., White O., Peterson J., Hoffman J., Parsons R., Baden-Tillson H., Pfannkoch C., Rogers Y. H., Smith H. O.* Environmental Genome Shotgun Sequencing of the Sargasso Sea // *Science*. 2004. Vol. 304. P. 66–74.

23. *Moore B. S., Kalaitzis J. A., Xiang L. K.* Exploiting Marine Actinomycete Biosynthetic Pathways for Drug Discovery // *Antonie van Leeuwenhoek*. 2005. Vol. 87. P. 49–57.

24. *Badger M. R., Price G. D., Long B. M., Woodger F. J.* The Environmental Plasticity and Ecological Genomics of the Cyanobacterial CO₂ Concentrating Mechanism // *J. Exp. Botany*. 2006. Vol. 57. P. 249–265.

25. *Kagan J., Sharon I., Beja O., Kuhn J. C.* The Tryptophan Pathway Genes of the Sargasso Sea Metagenome: New Operon Structures and the Prevalence of Non-Operon Organization // *Genome Biology*. 2008. Vol. 9 (1). P. 20.

26. *Robe P., Nalin R., Capellano C., Vogel T. M.* Extraction of DNA from Soil of DNA from Soil // *Eur. J. Soil. Boil.* 2003. Vol. 39. P. 183–190.

27. *Daniel R.* The Soil Metagenome—a Rich Resource for the Discovery of Novel Natural Products // *Current Opinion in Biotechnology*. 2004. Vol. 15, № 3. P. 199–204.

28. *Doolittle W.* Phylogenetic Classification and the Universal Tree // *Science*. 1999. Vol. 284. P. 2124–2128.

29. *Amann R. I., Ludwig W., Schleifer K. H.* Phylogenetic Identification and in situ Detection of Individual Microbial Cells without Cultivation // *Microbiol.* 1995. Vol. 59. P. 143–169.

30. *Raaijmakers J. M., Weller D. M., Thomashow L. S.* Frequency of Antibiotic-Producing *Pseudomonas* spp. in Natural Environments // *Appl. Environ. Microbiol.* 1997. Vol. 63. P. 881–887.

31. *Shen B., Du L., Sanchez C., Edwards D. J., Chen M., Murrell J. M.* The Biosynthetic Gene Cluster for the Anticancer Drug Bleomycin from *Streptomyces verticillus* ATCC15003 as a Model for Hybrid Peptide-Polyketide Natural Product Biosynthesis // *J. Ind. Microbiol. Biotechnol.* 2001. Vol. 27. P. 378–385.

32. *Skoko N., Vujovic J., Savic M., Papic N., Vasiljevic B., Ljubijankic G.* Construction of

Saccharomyces Cerevisiae Strain FAV20 Useful in Detection of Immunosuppressants Produced by Soil Actinomycetes // *J. Microbiol. Methods*. 2005. Vol. 61. P. 137–140.

33. *Ullrich R., Nuske J., Scheibner K., Spantzel J., Hofrichter M.* Novel Haloperoxidase from The Agaric Basidiomycete *Agrocybe Aegerita* Oxidizes Aryl Alcohols and Aldehydes // *Appl. Environ. Microbiol.* 2004. Vol. 70. P. 4575–4581.

34. *Ghazanfar S., Azim A., Ghazanfar M. A., Iqbal M., Anjum I. B.* Metagenomics and Its Application in Soil Microbial Community Studies: Biotechnological Prospects // *J. Animal. Plant. Sci.* 2010. Vol. 6, № 2. P. 611–622.

35. *Tringe S. G., Zhang T., Liu X., Yu Y., Lee W. H., Yap J., Yao F., Suan S. T., Ing S. K., Haynes M., Rowher F., Wei C. L., Tan P., Bristow J., Rubin E. M., Ruan Y.* The Airborne Metagenome in an Indoor Urban Environment // *PloS ONE*. 2008. Vol. 3, № 4. ID 1862.

36. *Turnbaugh P. J., Ley R. E., Hamady M., Fraser-Liggett C. M., Knight R., Gordon J. I.* The Human Microbiome Project // *Nature*. 2007. Vol. 449. P. 804–810.

37. *Bogaert D., Keijser B., Huse S., Rossen J., Veenhoven R., Gils E. van, Bruin J., Montijn R., Bonten M., Sanders E.* Variability and Diversity of Nasopharyngeal Microbiota in Children: A Metagenomic Analysis // *PloS ONE*. 2011. Vol. 6, № 2. E17035.

38. *Ley R. E., Peterson D. A., Gordon J. I.* Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine // *Cell*. 2006. Vol. 124. P. 837–848.

39. *Turnbaugh P. J., Quince C., Faith J. J., McHardy A. C., Yatsunenkov T., Niazi F., Affourtit J., Egholm M., Henrissat B., Knight R., Gordon J. I.* Organismal, Genetic and Transcriptional Variation in the Deeply Sequenced Gut Microbiomes of Identical Twins // *Proc. Natl. Acad. Sci. USA*. 2010. Vol. 107. P. 7503–7508.

40. *Qin J., Li R., Raes J., Arumugam M., Burgdorf K. S., Manichanh C., Nielsen T., Pons N., Levenez F., Yamada T., Mende D. R., Li J., Xu J., Li S., Li D., Cao J., Wang B., Liang H., Zheng H., Xie Y., Tap J., Lepage P., Bertalan M., Batto J., Hansen T., Le Parslier D., Linneberg A., Nielsen H. B., Pelletier E., Renault P., Sicheritz-Ponten T., Turner K., Zhu H., Yu C., Li S., Jian M., Zhou Y., Li Y., Zhang X., Li S., Qin N., Yang H., Wang J., Brunak S., Dore J., Guarner F., Kristiansen K., Pedersen O., Parkhill J., Weissembach J., Bork P., Ehrlich S. D., Wang J.* A Hu-

man Gut Microbial Gene Catalogue Established by Metagenomic Sequencing // *Nature*. 2010. Vol. 464. P. 59–65.

41. Tap J., Mondot S., Levenez F., Pelletier E., Caron C., Furet JP., Ugarte E., Muñoz-Tamayo R., Paslier D.L., Nalin R., Dore J., Leclerc M. Towards the Human Intestinal Microbiota Phylogenetic Core // *Environ. Microbiol.* 2009. Vol. 11. P. 2574–2584.

42. Arumugam M., Raes J., Pelletier E., Le Paslier D., Yamada T., Mende D. R., Fernandes G. R., Tap J., Bruls T., Batto J. M., Bertalan M., Borruel N., Casellas F., Fernandez L., Gautier L., Hansen T., Hattori M., Hayashi T., Kleerebezem M., Kurokawa K., Leclerc M., Levenez F., Manichanh C., Nielsen H. B., Nielsen T., Pons N., Poulain J., Qin J., Sicheritz-Ponten T., Tims S., Torrents D., Ugarte E., Zoetendal E. G., Wang J., Guarner F., Pedersen O., Vos de W. M., Brunak S., Dore J.; MetaHIT Consortium, Antolin M., Artiguenave F., Blottiere H. M., Almeida M., Brechot C., Cara C., Chervaux C., Cultrone A., Delorme C., Denariac G., Dervyn R., Foerster K. U., Friss C., Guchte M. van de, Guedon E., Haimet F., Huber W., Hylckama-Vlieg J. van, Jamet A., Juste C., Kaci G., Knol J., Lakhdari O., Layec S., Le Roux K., Maguin E., Merieux A., Melo Minardi R., M'riani C., Muller J., Oozeer R., Parkhill J., Renault P., Rescigno M., Sanchez N., Sunagawa S., Torrejon A., Turner K., Vandemeulebrouck G., Varela E., Winogradsky Y., Zeller G., Weissenbach J., Ehrlich S.D., Bork P. Enterotypes of the Human Gut Microbiome // *Nature*. 2011. Vol. 473. P. 174–180.

43. Sellon R. K., Tonkonogy S., Schultz M., Dieleman L., Grenther W., Balish E., Rennie D. M., Sartor R. B. Resident Enteric Bacteria Are Necessary for Development of Spontaneous Colitis and Immune System Activation in Interleukin-10-Deficient Mice // *Inf. Immun.* 1998. Vol. 66, № 11. P. 5224–5231.

44. Lupton J. R. Microbial Degradation Products Influence Colon Cancer Risk: The Butyrate Controversy // *J. Nutr.* 2004. Vol. 134. P. 479–482.

45. Ley R. E., Turnbaugh P. J., Klein S., Gordon J. I. Microbial Ecology: Human Gut Microbes Associated with Obesity // *Nature*. 2006. Vol. 444. P. 1022–1023.

46. Pietroiusti A., Luzzi I., Gomez M. J., Magrini A., Bergamaschi A., Forlini A., Galante A. *Helicobacter pylori* Duodenal Colonization Is a Strong Risk Factor for the Development of Duodenal Ulcer // *Aliment. Pharmacol.* 2005. Vol. 21, № 7. P. 909–915.

47. Wu S., Rhee K. J., Albesiano E., Rabizadeh S., Wu X., Yen H. R., Huso D. L., Brancati F. L., Wick E., McAllister F., Housseau F., Pardoll D. M., Sears C. L. A Human Colonic Commensal Promotes Colon Tumorigenesis via Activation of T-helper Type 17 T-cell Responses // *Nat. Med.* 2009. Vol. 15. P. 1016–1022.

48. Uronis J. M., Muhlbauer M., Herfarth H. H. et al. Modulation of the Intestinal Microbiota Alters Colitis-Associated Colorectal Cancer Susceptibility // *PLoS ONE*. 2009. Vol. 4. P. e6026.

49. Rao V. P., Poutahidis T., Ge Z., Nambiar P. R., Boussahmain C., Wang Y. Y., Horwitz B. H., Fox J. G., Erdman S. E. Innate Immune Inflammatory Response against Enteric Bacteria *Helicobacter hepaticus* Induces Mammary Adenocarcinoma in Mice // *Cancer Res.* 2006. Vol. 66. P. 7395–7400.

50. Laqueur G. L., McDaniel E. G., Matsumoto H. Tumor Induction in Germfree Rats with Methylazoxymethanol (MAM) and Synthetic MAM Acetate // *J. Natl. Cancer Inst.* 1967. Vol. 39. P. 355–371.

51. Geier M. S., Butler R. N., Howarth G. S. Probiotics, Prebiotics and Synbiotics: A Role in Chemoprevention for Colorectal Cancer? // *Cancer Biol. Ther.* 2006. Vol. 5. P. 1265–1269.

52. Goldin B. R., Gorbach S. L. Alterations of the Intestinal Microflora by Diet, Oral Antibiotics, and *Lactobacillus*: Decreased Production of Free Amines from Aromatic Nitro Compounds, Azo Dyes, and Glucuronides // *J. Natl. Cancer Inst.* 1984. Vol. 73. P. 689–695.

53. Goldin B. R., Swenson L., Dwyer J., Sexton M., Gorbach S. L. Effect of Diet and *Lactobacillus acidophilus* Supplements on Human Fecal Bacterial Enzymes // *J. Natl. Cancer Inst.* 1980. Vol. 64. P. 255–261.