

МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

На правах рукописи

УДК **xxx.xxx**

Болдырева Дарья Игоревна

НАЗВАНИЕ ДИССЕРТАЦИОННОЙ РАБОТЫ

Специальность **XX.XX.XX** —
«**Название специальности**»

Диссертация на соискание учёной степени
бакалавра

Научный руководитель:

уч. степень, уч. звание

Фамилия И.О.

Долгопрудный — 2015

Оглавление

1	Материалы и методы	12
1.1	Материалы	12
1.2	Методы	12
1.2.1	Выделение ДНК	12
1.2.2	Секвенирование переменных участков V3-V4, V5-V6 гена 16S рРНК с помощью секвенатора Illumina MiSeq	13
1.2.3	Секвенирование переменных участков V2-4-8/ V3-V6,V7-9 гена 16S рРНК с помощью секвенатора IonTorrent PGM	15
1.2.4	Shotgun-секвенирование	16
1.2.5	Анализ полученных данных	18
2	Результаты	19
2.1	Секвенирование переменных участков V3-V4, V5-V6 гена 16S рРНК	19
2.2	Секвенирование переменных участков V2-4-8/ V3-V6,V7-9 гена 16S рРНК	20
3	Выводы	22
3.1	Таблица обыкновенная	22
	Заключение	23
	Список литературы	24
	Список рисунков	25
	Список таблиц	26
A	Название первого приложения	27

Микроорганизмы отвечают за большинство биогеохимических циклов, которые оказывают влияние на атмосферу Земли и ее океанов. Пока еще данные организмы плохо исследованы, так как изучение и понимание метаболических путей микроорганизмов ограничивается отсутствием возможности получения чистых культур. В последних исследованиях ученые приступают к изучению бактерий из окружающей среды путем культуро-независимых методов: извлечение ДНК из образцов и трансформация ее в большие вставки клонов (large insert clones). В настоящее время уже существует и продолжают появляться множество новых и различных методов по изучению микроорганизмов из окружающей среды. Самыми известными и популярными являются метагеномные подходы, которые дают наиболее информативные данные по видовому разнообразию сообщества микроорганизмов, а также его функциональным возможностям. Образцы для исследования по своей природе весьма разнообразны. На данный момент уже исследованы метагеномы кишечника человека, почвы, водоемов...И на этом наука не останавливается и продолжает двигаться дальше.

Объектом нашего исследования был меромиктический водоем - водоем, в котором практически отсутствует циркуляция воды между слоями различной минерализации, разделенные хемоклином, вследствие чего вода нижнего слоя более минерализована и плотная, чем верхнего. Меромиктические озера возникают по двум причинам. Во-первых, в тех случаях, когда породы, слагающие дно озера, являются растворимыми, например, гипсы. В этом случае они частично растворяются в озерной воде, повышая минерализацию придонных слоев. Часто бывает, что подземные воды, протекающие по растворимым породам, также становятся минерализованными, и поступают в глубинные слои в виде подводных ключей с повышенной минерализацией воды. Во-вторых, меромиктия может возникать в результате опреснения поверхностных слоев соленых водоемов, например, глубоко врезающихся в сушу морских заливов, связанных с морем через относительно мелководные проливы. Пресные воды рек, впадающих в такие заливы, частично замещают соленую морскую воду в поверхностных слоях, тогда как в глубинных слоях соленость остается близкой к морской. К таким водоемам относится Черное море. Это самый большой в мире меромиктический водоем.

В меромиктических водоемах нижние слои воды могут обладать значительно более высокой соленостью, чем поверхностные. Поверхностный менее соленый слой называется миксолимнионом. Нижний, более соленый – монимолимнионом. Интервал глубин, в котором происходит переход от миксолимниона к монимолимниону, т.е резко возрастает концентрация солей, называется хемоклином.

Озерная вода, кроме солей, содержит в большем или меньшем количестве растворенные газы: кислород, азот, углекислоту, метан, сероводород и водород. Водород, метан и сероводород поступают в воду из иловых отложений в результате анаэробного (т.е., без присутствия кислорода) распада органики, осуществляемого специфическими микроорганизмами.

Рассмотрим круговорот наиболее важных из этих элементов в стратифицированном водоеме. Кислород поступает из атмосферы, растворяясь в поверхностных слоях воды. Вторым источником кислорода – фотосинтез водорослей и высших растений, но об этом поговорим позже, в следу-

ющих главах. Содержание кислорода в воде в значительной степени определяет очень важный показатель среды – окислительно-восстановительный потенциал (или редокс-потенциал). Коротко можно сказать, что этот показатель говорит, какая обстановка наблюдается в данной среде – окислительная (положительный редокс-потенциал) или восстановительная (отрицательный редокс- потенциал). В придонных слоях достаточно глубоких водоемов фотосинтез невозможен, т.к водная толща поглощает почти весь свет. Следовательно, кислород может проникнуть в глубинные слои воды только за счет диффузии из верхних слоев, либо в процессе перемешивания верхних слоев воды с нижними слоями.

Сероводород образуется сульфатредуцирующими бактериями в донных отложениях любых озер, т.к. даже в самом насыщенном кислородом озере с низкой продукцией органики (так называемом олиготрофном озере) в иловых отложениях на глубине в несколько сантиметров кислорода уже нет. В таких озерах сероводород, произведенный в донных отложениях, окисляется кислородом уже в верхних слоях донных отложений, поэтому его практически нет даже в придонных слоях водной толщи. В меромиктических озерах глубинные слои не содержат кислорода, поэтому в таких озерах весь микробиоценоз насыщается сероводородом.

Углекислый газ всегда содержится в природных водоемах как в виде CO_2 так и в виде молекул угольной кислоты разной степени диссоциации: H_2CO_3 , HCO_3^- , CO_3^{2-} . Она потребляется фотосинтезирующими (так называемыми автотрофными организмами) в качестве источника углерода для синтеза биомассы. Она же вновь образуется в результате разложения этой биомассы живыми организмами (так называемыми гетеротрофными организмами). Таким образом осуществляется биотический круговорот углерода в водной экосистеме, как и в биосфере в целом.

Живое население озер в основном представлено планктонными организмами. Планктон – это все живые организмы, которые плавают в водной толще, но не способны противостоять течению. Состав планктонных организмов чрезвычайно сложен, их большое разнообразие позволяет лишь условно выделить основные группы.

Фитопланктон – к фитопланктону относятся фотосинтезирующие (фототрофные микроорганизмы). Это в основном одноклеточные зеленые, диатомовые водоросли, а также цианобактерии (сине-зеленые водоросли). Фитопланктон является автотрофным звеном водной экосистемы, осуществляет первичный синтез органики, а все остальные группы планктона потребляют эту органику, являясь участниками трофической цепи.

Бактериопланктон – к этому типу планктона относятся все основные группы бактерий, взвешенные в водной толще и осуществляющие разные функции в водной экосистеме. Это питающиеся органическими веществами бактерии (гетеротрофные бактерии). Это также хемоавтотрофные бактерии, строящие свою биомассу из углекислоты, так же как водоросли и растения, но в отличие от них, получающие энергию не от солнечного света, а от реакций окисления некоторых неорганических соединений, например H_2S , NH_4^+ , NO_2^- .

Протозоопланктон – инфузории, жгутиковые и прочие крупные одноклеточные микроорганизмы, поедающие фито- и бактериопланктон.

Зоопланктон – многоклеточные существа, в основном рачки и коловратки, которые потребляют фито-, бактерио-, и протозоопланктон. Представители зоопланктона сильно различаются по размерам, от почти невидимых глазом коловраток и циклопов, до рачков-бокоплавов, которые достигают нескольких сантиметров в длину.

Рыбы – питаются фито- и зоопланктоном, а хищные рыбы являются верхним трофическим звеном, питаясь планктоядными рыбами. Молодь рыб питается также протозоопланктоном.

Кроме того, в экосистемах водоемов большую роль играют бентосные организмы, населяющие донные отложения. Это моллюски, черви, личинки насекомых. Разумеется, высшие водные растения, развивающиеся в прибрежной полосе, также вносят значительный вклад в первичную продукцию [1].

Таким образом, мы видим, что несмотря на неоднородность минерализации в стратифицированных водоемах существует множество живых организмов. При изучении микробных сообществ возникает вопрос, как мы можем их исследовать и описать? Обычно описание и классифицирование микроорганизмов происходит по их фенотипу. Фенотип — широкое понятие, которое охватывает наблюдаемые характеристики организма, такие как морфология, физиологическая активность, структура компонент клетки и иногда экологическая ниша, которую занимает микроорганизм. Однако такие характеристики дают мало информации об эволюционном родстве микроорганизмов, что должно быть основой классификационных систем. Более того большинство анализов фенотипических характеристик требуют клеточных функций или роста, следовательно, для определения фенотипа необходимы чистые культуры, создание которых для многих видов организмов затруднено или невозможно. С другой стороны, сравнение организмов для определения их эволюционного сродства может быть достигнуто путем сравнения нуклеотидной последовательности геномов. Но такую информацию достаточно тяжело получить и интерпретировать. Также взаимосвязь между организмами может быть получена из сравнительного анализа последовательностей отдельных генов в геноме [2].

Достижения в такой области анализа (секвенирование ДНК) заключается в возможности изучения генетического разнообразия некультивируемых микроорганизмов. Например, широко используемым методом для изучения разнообразия микробиоты является секвенирование ампликонов. В таком случае из образца выделяют ДНК, затем выделенную ДНК амплифицируют методом ПЦР, далее производят секвенирование, после чего определяют, какие микроорганизмы представлены в образце и в каком количественном соотношении. В случае бактерий и архей вышеприведенным методом изучают малую рибосомальную субъединицу 16S, которая несет в себе как таксономическую, так и филогенетическую информацию. Секвенирование ампликонов 16S рРНК выявило огромное количество микробиологического разнообразия и используется для определения характеристик биоразнообразия микроорганизмов в большом диапазоне окружающей среды, включающей кишечник человека, термальные источники, антарктические минеральные почвы вулканов. Сравнения последовательностей 16S разных образцов показывает, как разнообразие микроорганизмов связано с условиями окружающей среды. Но у данного метода есть свои ограничения и недостатки. Во-первых, можно неверно определить значительную долю

разнообразие в сообществе из-за различных ошибок, связанных с ПЦР. Во-вторых, секвенирование ампликонов может совершить большой разброс оценок биоразнообразия. В дополнении, ошибки секвенирования и неправильно собранные риды могут выдать неверные последовательности, которые будет тяжело определить. В-третьих, секвенирование ампликонов обычно дает только представление о таксономическом составе микробиологического сообщества. Невозможно, точно определить биологические функции сообщества, используя данный подход. Точность, с которой данный метод оценивает действительные функциональные различия сообществ, связана с тем, насколько хорошо сообщество представлено в существующих базах данных последовательностей геномов. Наконец, секвенирование ампликонов ограничивается анализом таксонов, для которых таксономически информативные генетические маркеры известны и могут быть амплифицированы.

Однако существует альтернативный подход, с помощью которого можно изучить некультивируемые микроорганизмы и который помогает избежать предыдущие ограничения. Данный метод называется метод шотган секвенирования. Аналогично секвенированию специфических участков генома, мы выделяем из клеток ДНК образца. Отличие состоит в том, что впоследствии ДНК режется на небольшие участки определенной длины, которые затем секвенируются. Это дает последовательности ДНК, которые выравниваются на различные геномные участки для множества геномов, представленных в образце. Некоторые из полученных ридов будут отобраны из таксономически информативных участков генома, другие — из кодирующих последовательностей, дающие представление о биологических функциях, закодированных в геноме. В результате, данные метагеномики дают возможность одновременно изучать два аспекта микробиологических сообществ: кто это? и какова его функция?

Несмотря на преимущества данного метода перед другими, в метагеномном секвенировании также существуют некоторые проблемы. Во-первых, метагеномные данные достаточно сложны и велики с точки зрения биоинформатического анализа. Например, определение генома, из которого были получены риды, может оказаться трудной задачей. Кроме того, большинство сообществ настолько разнообразны, что многие геномы не могут быть полностью представлены ридами. В результате два рида одного и того же гена могут не перекрываться, следовательно их невозможно напрямую сравнить, просто выравнивая последовательности. Если риды перекрываются, то это еще не значит, что они относятся к одному геному или к разным, которые могут выстраивать последовательность. Так же метагеномный анализ, как правило, требует большой объем данных для определения значимых результатов из-за большого количества выбранной геномной информации, что может вызывать вычислительные трудности. К счастью, современные вычислительные технологии развиваются достаточно быстро, чтобы обойти данные проблемы. Во-вторых, метагеном может содержать нежелательную ДНК хозяина, особенно при изучении микробиоты. В некоторых ситуациях ДНК хозяина может подавлять ДНК сообщества молекулярными методами, используемыми для выборочного обогащения ДНК микроорганизмов до процесса секвенирования. В настоящее время разрабатываются молекулярные и биоинформатические методы для фильтрации ДНК хозяина до процесса секвенирования или после него. В-третьих, пока

загрязнение является общей проблемой для исследования окружающей среды, определение и удаление метагеномных последовательностей загрязняющих веществ особенно проблематично. Загрязнение метагенома может привести к неправильному анализу генетического разнообразия сообщества, если количество загрязнений преобладает в исследуемом сообществе, особенно когда их очень много или когда они имеют большой геном. Однако уже существуют программы, которые могут справиться с данной проблемой. Наконец, создание метагенома дороже по сравнению с секвенированием ампликонов, особенно в сложных сообществах или когда ДНК хозяина значительно превосходит микробиологическую ДНК. К счастью, достижения в области ДНК секвенирующих технологий повышают доступность метагеномного секвенирования.

Вышеприведенные проблемы, конечно, ограничивают применение метагеномных исследований, но благодаря упомянутым научным достижениям такой аналитический подход становится все более популярным для большинства лабораторий. За последние годы секвенирование метагенома было использовано для определения новых вирусов, установления характеристик и функций некультивируемых бактерий, выявления новых и экологически важных белков и идентификации таксономического и метаболического путей микробиоты кишечника для здоровых и больных людей. Также метагеномный анализ может быть использован для изучения микробиоты растений [3].

Помимо вышеприведенных методов исследования некультивируемых сообществ существуют другие не менее эффективные подходы для их изучения.

Так например, микроскопические методы исследования, которые используются непосредственно для визуализации нуклеиновых кислот. РНК-FISH (fluorescent in situ hybridization) флуоресцентно меченные олигонуклеотидные пробы для маркировки экспрессированных молекул мРНК в фиксированных клетках. Пробы могут быть разработаны для любого гена с известной последовательностью, интересующие клетки могут быть изучены вместе с окружающей их тканью, что немаловажно при исследовании тканей и опухолей. Например, методом РНК-FISH были посчитаны единичные молекулы мРНК в единичных клетках и при помощи стохастического моделирования cell-cell variability (? перевод) были получены биохимические параметры, задействованные в транскрипции генов. Преимущество микроскопического метода изучения в том, что он сохраняет геометрию исследуемой ткани или клетки, что дает возможность понять, как различные клетки или клеточные компоненты взаимодействуют между собой и как они связаны с экспрессией генов. Ограничениями данного метода являются пропускная способность и возможность автоматизировать и распараллелить геномные измерения. Однако их можно избежать, разделив образцы по микрокамерам и используя автоматическую микроскопию.

Метод FACS (fluorescence-activated cell sorting) был разработан для определения характеристик единичных клеток. Суспензия клеток протекает через узкий поток жидкости. Клетки проходят лазерный луч, и рассеянный на клетках свет попадает на детекторы, которые обрабатывают его и переводят в клеточные характеристики, такие как размер и зернистость. После поток пропускается через сопло и разделяется на маленькие шарики, таким образом, что в каждом шарике находится одна клетка. Шарики приобретают электрический заряд, и затем электрически

разделяют по различным емкостям, соответствующим определенным клеточным характеристикам. FACS-приборы могут обрабатывать десятки тысяч клеток за час и измерять приблизительно до 18 поверхностных маркеров одновременно. После анализов клетки остаются жизнеспособными и могут быть использованы в биологических экспериментах, например, трансплантация стволовых клеток. FACS становится основной технологией в биомедицине, особенно в установлении характеристик различных типов клеток крови и иммунной системы. Использование красителей, связывающихся с ДНК, позволяет изучать содержание ДНК в единичных клетках. Что, с одной стороны, может быть использовано для последующего каритипирования, а с другой стороны, для определения различий между стадиями клеточных циклов. Основное ограничение данного метода заключается в количестве различных флуоресцентных маркеров, которые могут быть измерены одновременно [4].

Одним из наиболее высокопроизводительных методов анализа является метод, использующий микрочипы. ДНК микрочипы включают сотни тысяч ДНК фрагментов, расположенных на стеклянной подложке. Изначально микрочипы были разработаны для профилирования экспрессии генов. Впоследствии их стали использовать в изучении различных аспектах микробной экологии, включающей круговорот метана, общее микробное разнообразие, диапазон биогеохимических функций. Первые олигонуклеотидные микрочипы были применены для изучения микробной экологии, путем использования 9 проб 15-20-мерных олигонуклеотидов для определения азотопотребляющих бактерий. В дальнейшем это позволило разработать высокочувствительные 16S рРНК ген-специфичные (Gene targeting ?) микрочипы, которые могли анализировать около 30 000 20-мерных и 300 000 25-мерных олигонуклеотидных проб для исследования бактериального и архейного разнообразия. кДНК чипы, изначально сделанные для установления таксономических взаимоотношений среди вида *Pseudomonas* и использующие случайно созданные фрагменты генов длиной 1kb, позволили получать результаты с более высокой чувствительностью и лучшим разрешением, по сравнению с олигонуклеотидными микрочипами. Кроме того, CGAs (community genome arrays) и GPMs (genome probing microarrays), которые используют микробные геномы, были разработаны с использованием RSGP (reverse sample genome probing). Данный метод может избежать ошибок ПЦР, а так же повысить их специфичность и эффективность. Методы CGA и GPM используются для определения характеристик сложных микробных составов почвы, рек, морских отложений, а так же ферментированных растительных кормов, контролируемые в течение процесса ферментации. Метагеномными микрочипами (MGA - metagenome microarray) были созданы для быстрой характеристики метагеномных библиотек, содержащие полный геном микробиотического сообщества. От предыдущих данный вид микрочипов отличается понятием зонда и мишени. Зонд, как правило, расположен на стеклянной подложке, где MGA обрабатывает содержащиеся на подложке мишени и использует специфически меченные гены как зонд. Использование обратного подхода совместно с метагеномными микрочипами позволило быстро анализировать метагеномные библиотеки, которые сравнивают cosmid (? перевод) клоны, взятые из различных подземных вод, и fosmid (?) клоны из морских отложений. Наконец, применение изотопных микрочипов, основанных на их возможности использовать

радиомеченные подложку, для идентификации групп организмов сделало значительный прорыв в использовании микрочипов в микробной экологии, а так же в физиологическом анализе различных микробных сообществ донных осадков (in activated sludge). Как и все вышеописанные, данный метод имеет свои ограничения. Главными являются его стоимость и пропускная способность, а также количество образцов. Однако сейчас стоимость микрочипового анализа резко падает, что дает ему возможность конкурировать с методами, основанными на секвенировании, для анализа транскриптомов. Считается, что микрочипы представляют более экономичный и более практичный метод для исследований, требующих высокую пропускную способность.

Секвенирование геномов единичных клеток (single-cell genome sequencing) появилось как новый способ, который сделал доступными для анализа большинство некультивируемых организмов. Данный метод был использован для получения генома TM7, кандидат филома с экологической и клинической значимостью. (?) Также для исследований морских микробов, насекомых симбионтов (?), организмов из сложных микробных сообществ, так как рубец скота. Секвенирование единичных клеток используется также для сравнения генетических последовательностей индивидуальных клеток, секвенируемых в популяции. Консорциум клеток, будь то популяция клеток ткани, микробное сообщество микробиоты человека или экологические микроорганизмы, содержат геномные вариации, которые обеспечивают доступ к биологически и экологически интересным процессам (?). Действительно, геномная неоднородность управляет некоторыми аспектами фенотипической неоднородности, наблюдаемыми среди близкородственных клеток. Также геномный анализ единичных клеток используется для установления происхождения опухолей и тканей в многоклеточных организмах. Более того возможно изолировать и амплифицировать единичные хромосомы из единичных клеток. Для этого разработан микрожидкостный прибор способный разделить и амплифицировать гомологичные копии каждой хромосомы из одной метафазной человеческой клетки в независимых емкостях. Это позволило изучить две аллели каждой хромосомы независимо друг от друга. Данный метод может быть использован для получения точной гаплоидной генотипной информации каждой (personal ?) геномной последовательности, для понимания мейотической рекомбинации и непосредственного изучения индивидуальных человеческих лейкоцитарных антигенов. Одним из основных ограничений амплифицированного секвенирования генома единичных клеток являются ошибки процесса амплификации, так как он создает химеры и неспецифические продукты. Надеемся, что эти ограничения будут преодолимы в какой-то степени в скором будущем путем изобретения лучших методов амплификации и пространственной изоляции индивидуальных хромосом [4,5].

Рассмотрим применение некоторых из методов подробнее.

Так, например, используя метод секвенирования гена 16S рРНК и денатурированного градиентного гель-электрофореза, было изучено прокариотическое разнообразие микробного сообщества арктического шельфа. [Prokaryotic diversity. . .] Изучение подобных экосистем предоставляет возможность изучить, как могла существовать жизнь, когда Земля переживала ледниковый период. Микробные экосистемы, существующие в условиях низких температур, также представляют интерес для астробиологов, так как известно, что полярные регионы Марса и спутники Юпитера

и Сатурна содержат запасы замороженной воды способной поддерживать жизнь в прошлом или настоящем. Наконец, организмы, приспособленные к холоду, и их ферменты все чаще используются в еде, химической и текстильной промышленности и для восстановления загрязненной окружающей среды, находящейся в холодных условиях. Подробное исследование экосистем, для которых характерны низкие температуры, привело к описанию микробных сообществ, проживающих в подобных условиях, включающих морские льды, озерные льды, слой вечной мерзлоты, ледниковые системы. Шельфовые ледники являются регионами припайя, которые существуют как многолетние экосистемы в Арктике и Антарктике. В обоих полярных регионах данная окружающая среда обеспечивает уникальное поведение для сложных сообществ микробного мата. Данные сообщества существуют замороженными и остаются нетронутыми внутри шельфовых ледников в течение многих лет и появляются в бассейнах талых вод, образующихся на вершине шельфовых ледников в летние месяцы. Эти сообщества сталкиваются с очень суровыми условиями окружающей среды, включающих постоянное воздействие очень низких температур, изменение концентрации солей и высокое воздействие ультрафиолетового излучения в течение летних месяцев. Целью данного исследования было оценить прокариотическое разнообразие, микробное распределение и метаболическую активность в микробных матах шельфовых ледников Ward Hunt и Markham, используя культуру-зависимые и культуру-независимые методы. Изучение предоставило первое детальное описание прокариотических сообществ в микробных матах шельфовых ледников высоких широтах Арктики, первый молекулярный филогенетический анализ данного сообщества и первое описание метаболических путей данного сообщества при температуре много ниже нуля.

Методом shotgun-секвенирования было исследована связь между биоразнообразием и потенциальной функциональной ролью в современных окружающих средах на примере стратифицированных озер в карстовой области Banyoles, северо-восточной Испании. [оч. Крутая статья]. Связь между составом микробного сообщества и природными процессами, такими как круговорот углекислого газа, азота и серы - одна из важнейших целей микробиологических экологов. Данная информация необходима нам для лучшего понимания структуры и функциональных возможностей микробных сообществ, фундаментальных механизмов, контролирующих микробные процессы и взаимодействия *in situ*, также для изучения взаимодействия между биосферой-гидросферой-геосферой. Однако понимание биологического взаимодействия в сложных системах достаточно затруднено. Стратифицированные озера с аноксигенным сернистым дном – упрощенная система изучения для описания общих взаимодействий между биоразнообразием и биогеохимией, так как они имеют высокую активность, большую биомассу и низкое микробное разнообразие. Такие озера также могут потенциально быть смоделированы как современный аналог экосистемы океанов в течение долгого периода истории Земли. В данном примере проводилось изучение между металимнионом и гипolimнионом двух серосодержащих озер путем shotgun-секвенирования метабенома и анализа некоторых метаболических путей *in silico*. После филогенетической и функциональной идентификации была исследована связь между составом микробного сообщества и его функциональностью в круговороте углерода, азота

и серы. Из-за нехватки кислорода, большой микробной биомассы и большого вклада процесса фиксации CO_2 в глубоких темных слоях было сделано предположение о высоком генетическом потенциале хемотрофной фиксации CO_2 и тесной редокс-связи между круговоротами углерода, азота и серы. Функциональный анализ был сфокусирован на трех основных биогеохимических циклах для данного типа озер: круговорот углерода, азота и серы. В целом таксономическое распределение сообщества было оценено при помощи филогенетической аннотации метагеномных ридов. Домен *Bacteria* численно преобладает в генетическом составе микробного сообщества как аэробного-анаэробного взаимодействия, так и в анаэробном хиполимнионе. Более 95 процентов всех таксономически выровненных метагеномных ридов соответствуют бактериям с небольшим количеством представителей архей (более представлены в хиполимнионе, чем в металимнионе), фагов и эукариот. Секвенирование гена 16S рРНК подтвердило данное распределение. Наиболее представленными оказались бактерии вида *Chlorobiales* (серо-зеленые бактерии), *Campylobacteriales* (эпсилон-протеобактерии), *Burkholderiales* (бета-протеобактерии), *Frankiales* (актинобактерии). Данные бактерии были сильно разделены между слоями и озерами и показали таксономическую кластеризацию в соответствии с редокс-потенциалом. По результатам данных показано, что гены, отвечающие за анаэробную фиксацию углерода, фиксацию азота и ассимиляторную сульфатредукцию, составляют значительный процент от всех проаннотированных ридов в хиполимнионе, а гены для аэробного дыхания, ассимиляции азота и минерализации серы - в аэробном-анаэробном слое. Таким образом, метагеномный подход предоставил возможность установить связь между бактериями и биогеохимическими циклами двух стратифицированных озер, которые можно представить как современным аналогом древних океанов.

Глава 1

Материалы и методы

1.1 Материалы

1.2 Методы

1.2.1 Выделение ДНК

Образцы поступили будучи законсервированными в спирт, поэтому первым этапом было центрифугирование образцов при 5000g при температуре 4° в течение 15 минут. Затем супернатант (SN) переносили в новую пробирку для последующего переосаждения остатков ДНК из него*. Далее производилось выделение ДНК из образованного осадка. К осадку добавляли 800 мкл лизирующего буфера (500 mM NaCl, 50 mM Tris HCl pH=8, 50 mM ЭДТАб 4% SDS) и 300 мг кремне-циркониевых бусин диаметром 0,1 мм и 100 мг – диаметром 0,5 мм. Затем гомогенизировали на MiniBeadbeater (BioSpec, США), после чего образцы переносились в термостат для инкубации при 70° в течение 15 минут. После инкубации образцы центрифугировали на максимальных оборотах (14000g) в течение 15 минут, полученный SN переносили в новую 1.5 мл пробирку (Eppendorf). К SN добавили равный объем изопропанола, перемешав переворачиванием. После инкубировали не менее часа при –20°. Далее центрифугировали на максимальных оборотах 30 минут при температуре 4°. Полученный осадок дважды промывали 70% спиртом (этанол) и высушивали при комнатной температуре 15 минут. Осадок растворяли в 100 мкл TE-буфера.

***Процедура переосаждения.**

После первого центрифугирования (5000g, 15 мин, 4°) к SN добавляется два объема спирта и один объем 3М ацетата Na. Далее образцы инкубируются не менее часа при –20°, после чего снова центрифугируются на максимальных оборотах (14000g) в течение 30 минут при 4°. Далее осадок дважды промывается в 70% спирте и высушивается при комнатной температуре 15 минут. Затем осадок растворяется в TE-буфере.

Далее происходило объединение ДНК, выделенной из осадка и из SN.

Электрофоретическое разделение ДНК в агарозном геле.

Выделенный образец ДНК наносили на гель для электрофореза. Предварительно прогретую до 92 – 95°C 1,5% суспензию агарозы в 1×ТАЕ, охлаждали до 40 – 50°C, добавляли бромистый этидий (0,4 мкг/мл) и заливали в специальную форму (Helicon). Образцы ДНК смешивали с 3 мкл буфера для нанесения и наносили на гель для электрофореза. Электрофоретическое разделение проводили в 1×ТАЕ буфере при токе в 20-30 мА и разности потенциалов 5-7 В/см. Фрагменты ДНК визуализировали с помощью прибора УФ-транзлюминатора Imaging System Gel DOC (Bio-Rad, США), используя компьютерную программу Quantity One(version 4.6.8, build 027).

После выделения ДНК производилось приготовление библиотеки к секвенированию.

Определение концентрации ДНК

3 мкл раствора выделенной ДНК анализировали в агарозном геле. Количество продукта оценивали по интенсивности флуоресценции в УФ с помощью Qubit (Broad Range).

1.2.2 Секвенирование переменных участков V3-V4, V5-V6 гена 16S рРНК с помощью секвенатора Illumina MiSeq

Приготовление ампликонных библиотек на переменные участки V3-V4, V5-V6 проводилось в соответствии с протоколом, рекомендуемым компанией Illumina , *16S Metagenomic Sequencing Library Preparation*

Была проведена двустадийная полимеразная цепная реакция (ПЦР) (см. рис. 1.1). Для каждого образца составлялась смесь общим объемом 25 мкл, содержащая 1 мкл ДНК, 1мкл каждого из двух праймеров (10 пмоль), 2,5 мкл 10x буфера для ПЦР, 2,5 мкл 10x dNTP, 0,5 мкл Taq-полимеразы и 16,5 мкл воды. Далее проводили реакцию при следующих условиях:

1. 95° , 3 мин – денатурация;
2. 25 циклов:
 - 95°, 30 с – денатурация
 - 55°, 30 с – отжиг праймеров на матрице
 - 72°, 30 с – элонгация
3. 72°, 5 мин – достройка концов ПЦР-фрагментов
4. 4° , ∞

Далее проводилась чистка полученных ампликонов от свободных праймеров и димеров с использованием AMPure XP beads (Beckman Coulter, США). Для наиболее эффективной очистки

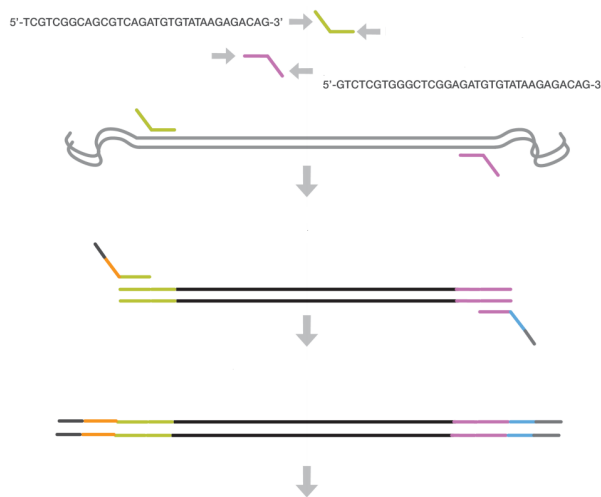


Рисунок 1.1: Двустадийная ПЦР

брали 20 мкл AMPure XP beads. Далее для качественной оценки ПЦР продукта был проведен электрофорез в 2% агарозном геле.

После того, как были выбраны накоплены необходимые вариабельные участки, шла вторая стадия ПЦР, на которой к образцам пришивались индексированные праймеры. Для каждого образца составлялась смесь общим объемом 50 мкл, которая содержала 5 мкл матрицы, 5 мкл буфера для ПЦР, 5 мкл dNTP, по 1 мкл каждого из двух индексированных праймеров, 0,5 мкл полимеразы и 32,5 мкл воды. Далее проводили реакцию при следующих условиях:

1. 95°, 3 мин – денатурация;
2. 8 циклов:
 - 95°, 30 с – денатурация
 - 55°, 30 с – отжиг праймеров на матрице
 - 72°, 30 с – элонгация
3. 72°, 5 мин – ?
4. 4°, ∞

Далее проводилась чистка полученных ампликонов от свободных праймеров и димеров с использованием AMPure XP beads. Для наиболее эффективной очистки брали 56 мкл AMPure XP beads. Далее образцы анализировались с использованием Bioanalyzer DNA 1000 chip для проверки размера полученной библиотеки (630п.о.).

Далее образцы смешивались в эквимольном соотношении: рассчитывается концентрация ДНК в нМ, определенная Agilent Technologies 2100 Bioanalyzer.

$$\frac{\text{концентрация в нг/мкл}}{660 \text{ г/моль} \cdot \text{средний размер библиотеки}} \cdot 10^6 = \text{концентрация в нМ} \quad (1.1)$$

Далее полученная библиотека была отсековирована с помощью прибора Illumina MiSeq (Illumina, США) согласно инструкциям и рекомендациям протокола производителя.

1.2.3 Секвенирование переменных участков V2-4-8/ V3-V6,V7-9 гена 16S рРНК с помощью секвенатора IonTorrent PGM

Приготовление библиотеки происходило в соответствии с протоколом *Ion 16STM Metagenomics Kit (Life Technologies, США)* (см. рис. 1.2). Было использовано два набора праймеров для переменных участков V2-4-8/ V3-V6,V7-9.

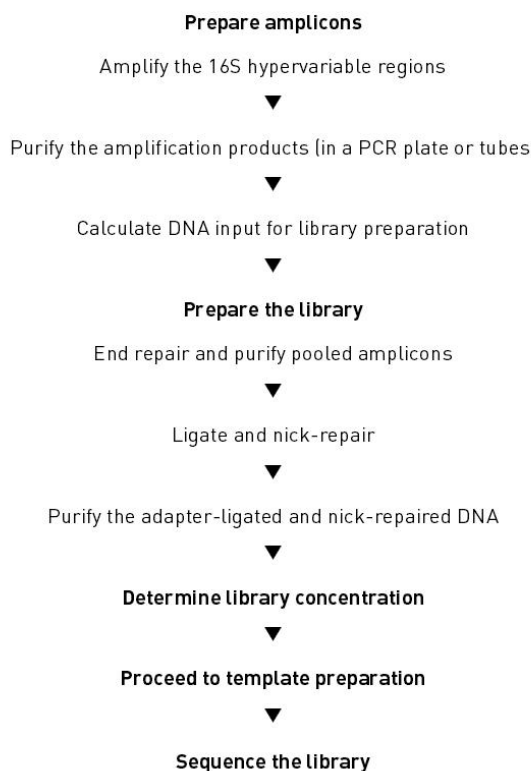


Рисунок 1.2: Алгоритм работы по протоколу Ion 16STM Metagenomics Kit

Для получения необходимых ампликонов была проведена ПЦР. Для каждого образца готовилась смесь, общим объемом 30 мкл, которая содержала 15 мкл 2X Environmental Master Mix, 3 мкл 16S Primer Set (10X)^[1], 2-12 мкл образца, объем воды, необходимый до 30 мкл. Далее проводили реакцию при следующих условиях:

1. 95° , 10 мин – денатурация;
2. 18-25 циклов:
 - 95°, 30 с – денатурация
 - 58°, 30 с – отжиг праймеров на матрице
 - 72°, 20 с – элонгация

3. 72°, 7 мин – ?

4. 4°, ∞

Далее проводилась чистка полученных ампликонов от свободных праймеров и димеров с использованием AMPure XP beads. Для наиболее эффективной очистки брали 1,8 V от образца AMPure XP beads. Далее из очищенных ампликонов готовили библиотеку для секвенирования. На первом этапе происходила достройка концов ампликонов, при помощи инкубации их при комнатной температуре в течение 20 минут с использованием 5X End Repair Buffer и End Repair Enzyme. Далее проводилась чистка полученных ампликонов от свободных праймеров и димеров с использованием AMPure XP beads. Затем были лигированы адаптеры и проведена нуклеотидная транскрипция. Для этого была проведена ПЦР. Для каждого образца замешивалась смесь, общим объемом 100 мкл, которая содержала 25 мкл матрицы, 10 мкл 10X Ligase Buffer, 2 мкл Ion P1 Adapter, 2 мкл Ion Xpress™ Barcode X[1], 2 мкл dNTP Mix, 49 мкл Nuclease-free Water, 2 мкл DNA Ligase, 8 мкл Nick Repair Polymerase. Далее проводили реакцию при следующих условиях:

1. 25°, 15 мин;

2. 72°, 5 мин;

3. 4°, ∞.

После следовала чистка ампликонов от свободных праймеров и димеров с использованием AMPure XP beads. Затем измеряли концентрацию, используя Qubit, проверяли качество библиотеки на Bioanalyzer DNA HS chip. Разводили библиотеки до концентрации 100 пмоль, смешивали в эквимольном соотношении (аналогично протоколу Illumina). Далее следовала эмульсионная ПЦР, чтобы посадить фрагментную библиотеку на специальные шарики, которые впоследствии будут нанесены на чип для секвенирования. Так как данный процесс имеет погрешность, то есть не все шарики оказываются связанными с библиотечкой, то была проведена процедура обогащения, в процессе которой мы избавились от пустых шариков. Затем следовало нанесение уже обогащенной библиотеки на чип, и далее было проведено секвенирование.

1.2.4 Shotgun-секвенирование

Приготовление библиотеки происходило в соответствии с протоколом *IonXpress™ Plus gDNA Fragment Library Preparation*.

Выделенную ДНК необходимо было поделить на фрагменты длиной 200 пар оснований. Данный этап происходил с использованием Covaris (). После необходимо выравнивать концы фрагментов, для этого была проведена инкубация при комнатной температуре с использованием 5X End Repair Buffer и End Repair Enzyme. Избавление от димеров и фрагментов малой длины (50 п.о.) происходило при чистке с использованием *Agencourt AMPure XP Kit*. Следующим этапом было лигирование адаптеров и нуклеотидная транскрипция. Для этого ставилась ПЦР. Для каждого образца

замешивалась смесь общим объемом 100 мкл, которую впоследствии делили на две пробирки одинакового объема для лучшего протекания реакции. Каждая пробирка содержала 25 мкл ДНК, 10 мкл 10X Ligase Buffer, 2 мкл Ion P1 Adapter, 2 мкл Ion Xpress Barcode X, 2 мкл dNTP Mix, 49 мкл Nuclease-free Water, 2 мкл DNA Ligase, 8 мкл Nick Repair Polymerase.

Далее проводили реакцию при следующих условиях для одного цикла:

1. 25° , 15 мин;
2. 72° , 5 мин;
3. 4° , ∞.

Далее была чистка от димеров с использованием *Agencourt AMPure XP Kit*.

Так как существует большая погрешность при фрагментации и не все фрагменты получаются одинаковой длины, то очищение от фрагментов большой длины (~100 п.о.) происходит с использованием агарозного E-Gel. Данный этап работы называется size-select. Затем снова проводилась амплификация, чтобы увеличить концентрацию фрагментов с лигированными адаптерами. Для каждого образца замешивалась смесь общим объемом 130 мкл, которую впоследствии делили на две пробирки одинакового объема для лучшего протекания реакции. Каждая пробирка содержала 100 мкл Platinum PCR SuoerMix High Fidelity, 5 мкл Library Amplification Primer Mix, 25 мкл неамплифицированной библиотеки.

Далее проводили реакцию при следующих условиях для одного цикла:

1. 95° , 5 мин
2. n циклов:
 - 95°, 15 с – денатурация
 - 58°, 15 с – отжиг праймеров на матрице
 - 70°, 1 мин – элонгация
3. 4° , ∞

Далее снова чистили библиотеки, измеряли концентрацию, используя Qubit, проверяли качество библиотеки на Bioanalyzer DNA HS chip. Разводили библиотеки до концентрации 100 пмоль, смешивали в эквимольном соотношении (аналогично протоколу Illumina). Далее следовала эмульсионная ПЦР, чтобы посадить фрагментную библиотеку на специальные шарики, которые впоследствии будут нанесены на чип для секвенирования. Так как данный процесс имеет погрешность, то есть не все шарики оказываются связанными с библиотечкой, то была проведена процедура обогащения, в процессе которой мы избавились от пустых шариков. Затем следовало нанесение уже обогащенной библиотеки на чип, и далее было проведено секвенирование.

1.2.5 Анализ полученных данных

Данные, полученные с Illumina MiSeq, были обработаны программным обеспечением, встроенным в прибор. Для классификации переменных фрагментов гена 16S рРНК была использована база данных GreenGenes (<http://greengenes.lbl.gov/cgi-bin/nph-index.cgi>) [1].

Для реконструкции гена 16S рРНК из ридов был использован геномный сборщик IDBA-UD (ссылки). Контиги были классифицированы с использованием онлайн-сервера RDP-Classifer (<https://rdp.cme.msu.edu/classifier/classifier.jsp>).

Для визуализации данных была использована программа MEGAN5 (<https://rdp.cme.msu.edu/classifier/classifier.jsp>).

Риды, полученные в результате shotgun-секвенирования, были собраны SPAdes Genome Assembler [2]. Полученная сборка была проаннотирована с помощью программы Prokka [3]. Для визуализации метаболических путей был использован онлайн ресурс KEGG [4].

Глава 2

Результаты

2.1 Секвенирование переменных участков V3-V4, V5-V6 гена 16S рРНК

Для каждого образца были получены следующие характеристики (см. таблицу 2.1)

По анализу переменных участков V3-V4, V5-V6 гена 16S рРНК было установлено видовое разнообразие меромектических озер бассейна Белого моря. Наиболее представленными оказались бактерии, относящиеся к следующим типам *Chlorobi*, *Proteobacteria*, *Actinobacteria*. Также было отмечено, что переменный участок V3-V4 хуже идентифицирует *Cyanobacteria*, чем переменный участок V5-V6 (см. рис. 2.1).



Рисунок 2.1: Определение видового разнообразия по переменным участкам а) V3-V4, б) V5-V6

Для озера Трехцветное был построен бактериальный профиль (видовое разнообразие в зависимости от глубины забора образца). Данный профиль показал, что с увеличением глубины видовое разнообразие достаточно резко сокращается и увеличивается количество бактерий, относящихся к типу *Chlorobi*. Таким образом, мы можем сделать предположение о том, что на глубине около двух метров начинается зона хемоклина. Наше предположение подтверждается тем, что на данной глубине происходит резкое уменьшение концентрации кислорода и увеличение концентрации сероводорода (см. рис. 2.2).



Рисунок 2.2: а) Зависимость концентрации кислорода и сероводорода от глубины, б) Бактериальный профиль оз. Трехцветное

2.2 Секвенирование переменных участков V2-4-8/ V3-V6,V7-9 гена 16S рНК

Повторное секвенирование гена 16S рНК показало, что доминантным представителем является *Chlorobium*. Последующий сбор ридов в контиги позволил лучше определить таксономическое разнообразие. Для более точного установления из данных по shotgun-секвенированию были выбраны риды, которые соответствуют переменным участкам V1, V5.

По собранным контигам из shotgun-секвенирования была установлена метаболическая карта для *Chlorobium*, что дало возможность анализировать и сравнивать данный тип бактерий с другими представителями рода (см. рис. 2.3).

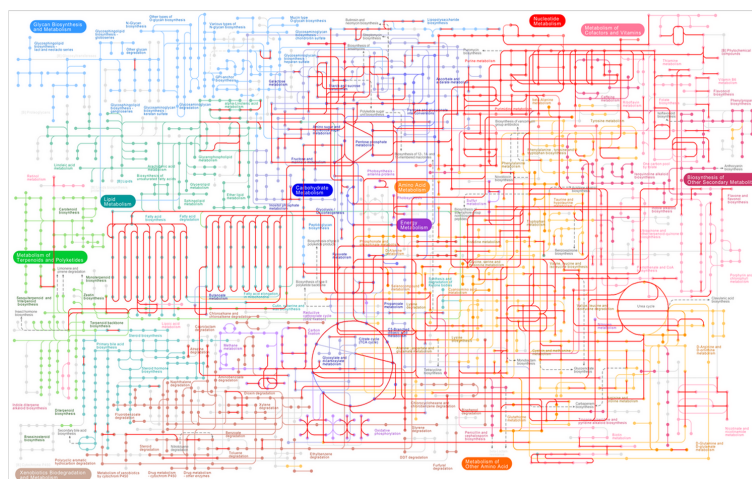


Рисунок 2.3: Карта метаболических путей

Также для них были получены карты метаболических путей серы и азота. По метаболическому пути азота удалось установить, что данные бактерии являются азотофиксирующими. Метаболический путь серы указал на способность бактерий к анаэробному фотосинтезу с использованием серосодержащих веществ как доноров серы.

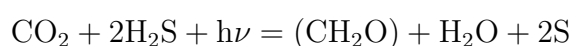


Таблица 2.1: Характеристики образцов, секвенированных на Illumina MiSeq.

Образец	Количество ридов участка V3-V4	Количество ридов участка V5-V6	Проверка качества ри- дов участка V3-V4	Проверка качества ри- дов участка V5-V6
10 КС 0,4	133 870	842 956	63,6 %	88,6 %
30 ТЦ 1,9-4	189 853	914 442	72,1 %	91,0 %
40 ЕЛ 3,0	159 276	883 811	72,7 %	90,9 %
50 ТЦ 1,75	134 000	696 095	77,5 %	89,8 %
60 КС 0,3	218 728	1 073 838	77,5 %	90,8 %
80 ТЦ 1,9	220 225	1 304 206	73,9 %	90,3 %
110 ЕР 2,5	225 734	1 306 371	76,5 %	91,0 %
120 ЗМ 4,5	166 293	1 014 921	67,2 %	90,0 %
130 ЕЛ 2,9	49 838	157 883	75,0 %	87,8 %

Глава 3

Выводы

3.1 Таблица обыкновенная

Заключение

Основные результаты работы заключаются в следующем.

1. На основе анализа ...
2. Численные исследования показали, что ...
3. Математическое моделирование показало ...
4. Для выполнения поставленных задач был создан ...

И какая-нибудь заключающая фраза.

Список литературы

1. Свидерская И.В., Рогозин Д.Ю. Биофизика. 10 класс. Модуль 5. Биофизика наземных и водных экосистем: учебно-методическая часть. — Красноярск: РИО КрасГУ, 2006. — 32 с.
2. Hugenholtz Philip, Pace Norman R. Identifying microbial diversity in the natural environment: a molecular phylogenetic approach // *Trends Biotechnol.* — 1996. — 6. — Vol. 14, no. 6. — Pp. 190–197.
3. Sharpton Thomas J. An introduction to the analysis of shotgun metagenomic data // *frontiers in PLANT SCIENCE.* — 2014. — 6. — Vol. 5, no. 6. — P. 14.
4. Kalisky Tomer, Blainey Paul, Quake Stephen R. Genomic Analysis at the Single-Cell Level // *PubMed Central.* — 2011. — 2. — Vol. ?, no. 45. — P. 22.
5. Comparing microarrays and next-generation sequencing technologies for microbial ecology research / Seong Woon Roh, Guy C.J. Abell, Kyoung-Ho Kim et al. // *Trends in Biotechnology.* — 2010. — 3. — Vol. 28, no. 6. — P. 9.

Список рисунков

1.1	Двустадийная ПЦР	14
1.2	Алгоритм работы по протоколу Ion 16S TM Metagenomics Kit	15
2.1	Определение видового разнообразия по вариабельным участкам а) V3-V4, б) V5-V6	19
2.2	а) Зависимость концентрация кислорода и сероводорода от глубины , б) Бактери- альный профиль оз. Трехцветное	20
2.3	Карта метаболических путей	20

Список таблиц

2.1	Характеристики образцов, секвенированных на Illumina MiSeq.	21
-----	---	----

Приложение А

Название первого приложения