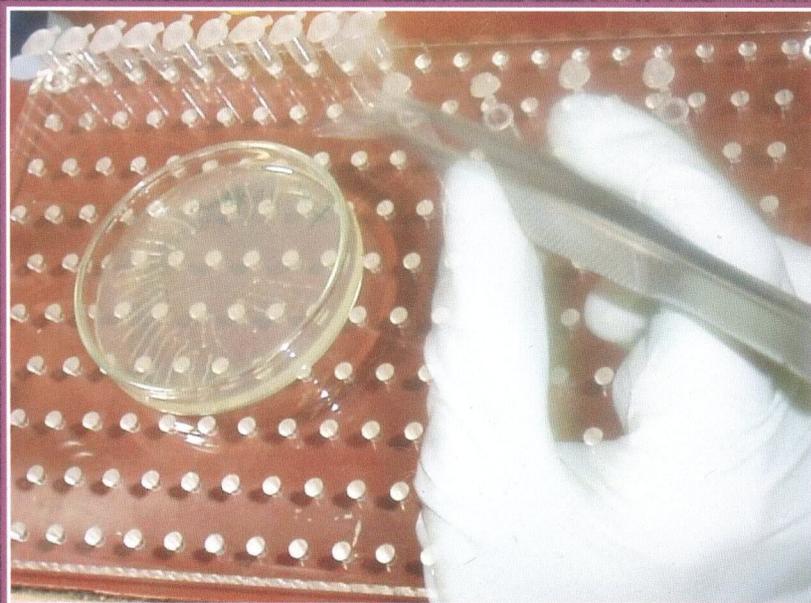


**РАЗНООБРАЗИЕ  
МИКРОБНЫХ СООБЩЕСТВ  
ВНУТРЕННИХ ВОДОЕМОВ  
РОССИИ**

*УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ*



**2009**

**РОССИЙСКАЯ АКАДЕМИЯ НАУК**  
**ИНСТИТУТ БИОЛОГИИ ВНУТРЕННИХ ВОД им. И.Д. ПАПАНИНА РАН**  
**УЧЕБНО-НАУЧНЫЙ ЦЕНТР**



**РАЗНООБРАЗИЕ МИКРОБНЫХ СООБЩЕСТВ**  
**ВНУТРЕННИХ ВОДОЕМОВ РОССИИ**  
*УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ*

**Борок 2009**

УДК 579.26 (07) (063)

ББК 28.4

P-17

Ответственный редактор: д.б.н. А.М. Андреева

Рецензент: д.б.н. В.В. Алешин

Разнообразие микробных сообществ внутренних водоемов России: Учебно-методическое пособие. - Ярославль: Изд-во ООО «Принтхаус», 2009. - 115 с.

ISBN

В сборнике опубликованы материалы четырех Всероссийских школ молодых ученых по проблемам молекулярной экологии микробных сообществ, посвященные практическим вопросам определения разнообразия микробных сообществ внутренних водоемов России: Рыбинского водохранилища, озера Байкал, озер Доронинское и Арахлей (Забайкалье), термальных источников Восточных Саян (Бурятия), подземных вод Пермского края. Исследования выполнены на базе Центра коллективного пользования «Молекулярные технологии» ИБВВ РАН. Основное внимание в материалах уделено адаптации методов пробоотбора к различным объектам – водным пробам, биоматам, накопительным культурам, донным пробам и ассоциатам с макрофитами. Пособие рассчитано на студентов, аспирантов и молодых ученых.

Библиограф. 151 назв. Ил. 23. Табл. 6.

УДК 579.26 (07) (063)

ББК 28.4

ISBN

© Издательство «Принтхаус», 2009

© Учреждение Российской Академии наук  
Институт биологии внутренних вод им. И.Д. Папанина РАН, 2009

## СОДЕРЖАНИЕ

	Стр.
<b>Предисловие</b>	4
<b>Молекулярные методы учета биоразнообразия</b> <i>В.В. Алёшин</i>	5
<b>Молекулярная экология микроорганизмов: факты, проблемы и перспективы</b> <i>Н.Л. Белькова, Н.Н. Деникина</i>	39
<b>Молекулярно-генетические методы анализа микробных сообществ</b> <i>Н.Л. Белькова</i>	53
<b>Определение генетического разнообразия в накопительных культурах микроорганизмов</b> <i>Э.В. Данилова, Н.Л. Белькова</i>	64
<b>Определение условий пробоотбора для получения корректных молекулярно-генетических данных</b> <i>О.П. Дагурова, Е.В. Суханова, И.В. Рыбакова, Е.В. Дзюба, Н.Л. Белькова</i>	75
<b>Селективная детекция живых организмов в природных ассоциациях молекулярно-генетическим методом</b> <i>И.В. Рыбакова, Р.А. Федоров, И.П. Рябцева, В.В. Большаков, Н.Л. Белькова</i>	90
<b>Влияние условий первичной обработки водных проб на разнообразие получаемых генотипов на примере микробных сообществ пресного озера Арахлей (Забайкалье)</b> <i>Е.Б. Матюгина, Н.Л. Белькова</i>	101
<b>Генетическое разнообразие микробных сообществ подземных вод с разным типом минерализации</b> <i>Н.М. Кашеварова, Н.Л. Белькова</i>	108

# Молекулярные методы учета биоразнообразия

В.В. Алешин

*Научно-исследовательский институт физико-химической биологии*

*им. А.Н. Белозерского Московского государственного университета им. М.В. Ломоносова*

## 1. Зачем изучать ДНК планктона и бентоса

С точки зрения зоолога или ботаника дело это бессмысленное – в таком суммарном препарате ДНК перемешаны фрагменты геномов сотен и тысяч разных видов, и неизвестно каких. Метагеном Тихого океана объединяет все виды Пацифики, от бактериофагов до китов. Даже если мы найдем в метагеноме какой-то ген и поймем его функциональное назначение, мы в настоящее время не узнаем, какому виду этот ген принадлежит (разве только нам особенно повезет наткнуться на известный вид). Для традиционно мыслящих систематиков, находящихся в плену рутины, это равносильно бесполезности такой работы. Тем не менее, метагеномика уверенно входит в арсенал современной биологии. Это предопределено уровнем технического развития и требованиями практики. Даже в локальных местообитаниях, как ротовая полость человека, новые методы оценки дают четырехзначные числа видового разнообразия только бактерий, постоянно или временно там проживающих (Keijser et al., 2008), без учета грибов и протистов. Для практики это означает потенциальную возможность доступа к неограниченному разнообразию ферментов и биологически активных препаратов с различными свойствами, для медицины – идентификацию новых, неизвестных прежде патогенных микроорганизмов. Но этим дело не ограничивается. Новые данные заставляют пересматривать устоявшиеся положения общего характера. Например, появляются указания о патогенности определенных комплексов видов, каждый из которых в отдельности безвреден, то есть речь идет о необходимости ревизии правила Коха, более 100 лет лежавшего в основе медицинской микробиологии.

Одновременно мы расширяем коллекцию антимикробных препаратов – естественных средств управления микробным сообществом в природе. Из метагенома термальных вод методами биоинформатики (или прямым экспериментальным скринингом библиотек) извлекают гены термостабильных ферментов для биотехнологии, и а из метагенома глубоководных частей океана (там круглый год минус один градус ниже нуля) – такие же ферменты, активные при низких температурах. И это не только стиральный порошок для мытья посуды или стирки в холодной воде, это другой уровень биотехнологии, получающей доступ к неизвестному ранее разнообразию биопрепаратов. Биотехнологические фирмы понимают эти перспективы и не скупятся на расходы по метагеномике. Для академической науки еще важнее другие аспекты. Обнаруживая в составе метагеномов те или иные гены, кодирующие ключевые ферменты биохимических путей, можно представить функционирование сообщества, устройство биологических циклов элементов и потоков энергии в нем, а в будущем – их регуляцию, как бы физиологию «вырезанного» участка живого покрова Земли – Геомериды (Старынкевич, 1931). На новом уровне наука вновь приходит к эвристичности такого общего понимания организации жизни на Земле, когда весь живой покров следует рассматривать, в известном смысле, как живой организм с его жидкой соединительной тканью морей и океанов (Беклемишев, 1928) и его «метагеномом».

Другая, более привычная сторона, находящаяся полностью в рамках традиционной парадигмы зоологии и ботаники, может быть охарактеризована как доступный метод учета видового разнообразия. Мы сосредоточимся на этой стороне вопроса, понимая ее ограниченность, равно как и неизбежность такой постановки в виде перехода к будущему более общему пониманию биоразнообразия как характеристики живого покрова Земли.

## 2. Проблема видовой идентификации

2.1. Общая постановка проблемы. Для определения видовой принадлежности бактерии требуется вывести ее в культуру (что далеко не всегда удается), изучать биохимические и культуральные свойства; в этом случае идентификация традиционными методами может растянуться на недели и месяцы, когда нужна экспресс-диагностика (например, возбудителя болезни). В систематике одноклеточных эвкариотов – протистов, а также «трудных» группах животных или растений по силам разобраться только специалисту, затратившему многие годы на ее изучение. Если группа не принадлежит к числу особо важных в хозяйственном отношении, то специалистов по ней может оказаться один-два в мире, а то и вовсе не оказаться в настоящий момент. В результате коллекционные сборы остаются десятилетиями не определенными. Более того, косвенные оценки биоразнообразия показывают, что во многих группах, особенно микроорганизмов, экологических группах пикопланктона и мейобентоса и др. большинство видов остаются не описанными, а, с учетом трудозатрат на описание и наличия трудовых ресурсов в виде квалифицированных специалистов, многие виды так и останутся навсегда неописанными. Рационализм науки требует изъять идентификацию из сферы деятельности систематиков, вывести ее из области искусства (в случае «трудных» групп) и сделать общедоступной, передать ее техническому персоналу. В числе прочего, это позволит обрабатывать массовые пробы, сократить стоимость определения за счет экономии оплаты труда специалистов, первоначально высвободить рабочее время специалистов для творческой работы по ревизии, а в перспективе сократить число систематиков, поскольку основная часть их работы (видовое определение) перейдет в разряд анализов, проводимых техническим персоналом. Молекулярные диагностические признаки оказались идеальными для осуществления этой программы. Они многочисленны, их легко формализовать и интерпретировать, нетрудно

добывать, применяя стандартные процедуры согласно опубликованным методическим указаниям. Цена лабораторного анализа (без учета затрат на сбор образцов в природе) при условии массовых анализов может быть сокращена до 10 – 20 долларов США (если проводить его в РФ) или до 1 доллара США (если проводить его за границей).

Наибольшее практическое значение имеет идентификация микроорганизмов, но по понятным причинам наиболее широкую известность получили методы видовой идентификации, примененные к животным и растениям. Широкую популярность в газетах и журналах приобрела концепция «ДНК-штрихкода». Наиболее популярной меткой для видовой идентификации и различения близких видов животных становится нуклеотидная последовательность фрагмента гена субъединицы I цитохром с оксидазы (*coI* или *coxI*), кодируемого митохондриальной ДНК. Ряд свойств этого гена (для краткости мы здесь опускаем причины того, почему эти свойства именно таковы) обеспечивает его низкую внутривидовую и значительную межвидовую изменчивость. И здесь необходимо важное уточнение: вид здесь понимается не в типологическом, а в генетическом смысле – как генетически связанное посредством скрещиваний сообщество, относительно обособленное в природе от остальных. Есть у *coxI* и другие преимущества, например, высокая консервативность кодируемого белка и то обстоятельство, что огромное большинство закрепляющихся в эволюции мутаций представляют собой замены, а не инсерции и делеции. Замены преобладают и в эволюции и других белков, но в тех случаях, когда белки функционируют в составе мультимолекулярного комплекса, как COI, размер которого строго ограничен его положением внутри мембраны митохондрии, инсерции и делеции особенно редки по понятным причинам. Эти особенности облегчают выравнивание последовательностей и делают его более достоверным. В то же время многочисленные синонимичные мутации фиксируются с большой свободой. Это делает ген *coxI* прекрасным кандидатом на роль «штрихкода», пригодного для автоматической

идентификации видов животных (Hebert et al., 2003; Шнеер 2009; Лухтанов, Кузнецова, 2009). Для высших растений, по-видимому, для повышения достоверности идентификации потребуется дополнительная метка, в качестве которой предлагается последовательность интрона хлоропластного гена матуразы, *matK*, а для грибов данных по *coxI* мало, и наиболее перспективными в настоящее время считаются последовательности внутренних транскрибируемых спейсеров генов рибосомных РНК, ITS1 и ITS2 (см. ниже). Международный консорциум (“Consortium for the Barcode of Life”, CBOL), сначала работавший на энтузиазме, а теперь получивший серьезное финансирование через гранты, обеспечивает поддержание общедоступной базы данных (<http://barcoding.si.edu/>) и ее расширение. Вскоре, по мере пополнения базы данных, видовая идентификация превратится в стандартную процедуру, включающую четыре этапа: 1) взятие образца ткани; 2) выделение ДНК; 3) амплификация нужного участка генома и определение его нуклеотидной последовательности; 4) сравнение полученной последовательности с базой данных.

Практические наработки CBOL не годятся для бактерий: у них нет митохондрий и интронов хлоропластного генова *matK*. Но благодаря этим работам происходит общий пересмотр отношения к систематике, замена линнеевской систематики филокодом (индексом, показывающим родство с другими видами), составляемым по простым, понятным, объективным критериям. Этот процесс теоретического переосмысления в равной степени важен и для ботаники, и для микробиологии.

2.2. Гены рибосомной РНК и методы их изучения. Задолго до создания CBOL в качестве удобных и универсальных молекул для филогенетических сравнений использовали рибосомные РНК. Рибосомы есть у всех живых существ, от человека до бактерий, и содержат гомологичные РНК. Если рассматривать ситуацию не с уровня сегодняшнего дня, а учитывать исторические обстоятельства, то важное преимущество рРНК состояло в том, что позволяло решить серьезные методические трудности, которые много лет

назад стояли на пути исследования генов. Тогда быстрого, удобного и дешевого метода полимеразной цепной реакции (ПЦР) не существовало, и первой и главной трудностью было выделить целевые последовательности из геномной ДНК. Для рРНК ее можно было обойти, выделив РНК вместо ДНК (либо тотальную РНК – в ней около 85% по массе приходится на рРНК, либо сначала рибосомы, а из них чистую рРНК). Вторая методическая трудность состояла в определении длинных нуклеотидных последовательностей, и она решалась тем, что первоначально для анализа использовали 5S и 5,8S рРНК, размер которых 120-150 нуклеотидов. Нуклеотидную последовательность 5S и 5,8S рРНК определяли по методу Максама и Гилберта. В дальнейшем методики были адаптированы для секвенирования фрагментов 18S и 28S рРНК. Используя затравки к консервативным участкам рРНК, проводили с помощью ревертазы секвенирование по методу Сэнгера. Полная длина рРНК малой субчастицы рибосом эукариотов (18S рРНК, SSU rRNA, 16S-like rRNA) варьирует в диапазоне 1700 – 1800 нуклеотидов. РНК такой длины как раз соответствует в седиментационных опытах 18-ти единицам Сведберга (18S). Изредка размеры рРНК малой субчастицы выходят за указанные рамки (они меньше у микроспоридий (1350 н.), некоторых видов грегариин (1500 н.), приближаясь по размеру к прокариотическим, и больше у некоторых видов амёб, фораминифер, членистоногих (до 3500 н.)). Эти более чем двукратные различия обусловлены изменением длины отдельных наиболее вариабельных участков в составе рРНК и не меняют решительным образом ее архитектуры, почти не мешают выравниванию консервативных участков. При сравнении рРНК удаленных видов из разных типов или царств обнаруживается до 30% различий в последовательностях, это сотни специфических признаков-нуклеотидов; эти данные позволяют не только провести достоверную идентификацию, но и делать содержательные и обоснованные филогенетические выводы, точность которых впервые стала превосходить традиционные гипотезы. Но у первоначальных методик были и серьезные

ограничения: они требовали большого количества свежего материала и были неприменимы для исследования музейных образцов, мелких и редких организмов, кроме тех, что выведены в культуру.

Решительный методический прорыв произошел с применением полимеразной цепной реакции (ПЦР) для амплификации целевых фрагментов ДНК (более подробно методика описана ниже). И вновь здесь сказалось преимущество генов рРНК перед белоккодирующими генами. К последним трудно подобрать универсальные затравки (праймеры) для ПЦР (см. ниже): эволюционно консервативные и даже идентичные участки белка кодируются у разных видов неодинаковыми нуклеотидными последовательностями из-за возможностей, предоставляемых вырожденностью генетического кода. Однако рибосомные РНК являются не посредником для трансляции, а макромолекулами, напрямую занятыми выполнением клеточных функций, поэтому их функционально важные участки консервируются на уровне нуклеотидных последовательностей. Еще одной удачей стало обстоятельство, что к числу консервативных участков относятся и концы 18S рРНК. К ним в лаборатории М. Согина были сконструированы более 20 лет назад (Medlin et al., 1988) универсальные праймеры А и В (рис. 1), пригодные для амплификации генов рРНК малой субчастицы самых разных видов эукариотов.

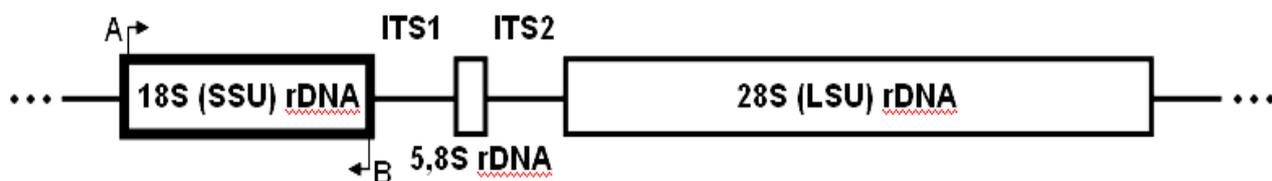


Рис. 1. Структура повторяющейся единицы (рибосомного оперона) в ядерном геноме эукариотов. ITS1, ITS2 – внутренние транскрибируемые спейсеры; А, В – положение праймеров для амплификации 18S рРНК.

Сконструированные праймеры не подходят к прокариотическому гену 16S рРНК, что позволяет не опасаться бактериальных загрязнений и избирательно амплифицировать эукариотические гены из сложных смесей. Точно так же были придуманы специфические праймеры к прокариотическому гену 16S рРНК, а также праймеры к другим участкам, не столь консервативным, которые позволяют специфически амплифицировать фрагменты гена рРНК животных, инфузорий, некоторых порядков бактерий и т. д.

Если нуклеотидные последовательности структурных генов 18S, 5,8S и 28S рРНК, входящих в состав молекулярной машины – рибосомы, отличаются высокой эволюционной консервативностью, то промежутки между ними – внутренние транскрибируемые спейсеры 1 и 2 (internal transcribed spacers, ITS1 и ITS2, см. рис. 1), далеко не так консервативны. Это определяется их физиологической ролью. Как известно, рибосомный оперон транскрибируется целиком, в виде единого предшественника, включающего рРНК и фрагменты между ними. Затем происходит созревание предшественника – вырезание из него тех фрагментов РНК, которые не войдут в состав рибосом, то есть спейсеров. Их главная роль, как считается, – правильно вырезаться, а с такой функцией оказывается совместима хотя и не любая последовательность, но практически бесчисленное их множество. Большинство мутаций в структурных генах рРНК оказываются вредными, поэтому эволюция рРНК происходит медленно, а в ITS1 и ITS2 большинство мутаций безразличны для функции. В результате значение рРНК и спейсеров для практических нужд таксономии разное. Например, 18S рРНК человека и лошади, двух относительно близких видов, отличаются всего 10-ью нуклеотидами из 1800 (примерно 7 замен и 3 делеции/вставки). Это сопоставимо с внутривидовой изменчивостью (и внутригеномной – ведь почти у всех эукариотов в геноме несколько сотен почти одинаковых копий генов рРНК). Этот пример показывает малую пригодность 18S рРНК для различения близких видов, по крайней мере в некоторых таксонах. Напротив,

последовательности ITS1 и ITS2 даже у близких видов всегда заметно отличаются.

В микробиологии понятие вида несколько отличается от действующего в зоологии и ботанике. Так и должно быть, потому что генетический обмен (именно он поддерживает генетическое единство вида животных или растений) у бактерий больше похож на манипуляции, применяемые в генетической инженерии, чем на половой процесс эвкариотов. Для генетической инженерии межвидовой перенос изолированного гена представляет в основном техническую проблему. Штаммы бактерий с отличиями в 16S рРНК до 3% принято относить к одному виду; это соглашение не вызывает возражений у микробиологов, хотя в зоологии и ботанике такой масштаб показался бы неудобным.

Если стоит задача филогенетического анализа, а не видовой идентификации в пучке, «букете» близких видов, – лучше рРНК ничего нет. То же, еще в большей мере, справедливо, если требуется отнести образец не к какому-то виду, а таксону более высокого ранга (роду, семейству, отряду, типу). Под этим лежат как исторические, так и более содержательные причины. Исторические обусловлены тем, что ни один из других генов не сравнится с рРНК по широте изученности. По сути 20 лет, начиная с прямого секвенирования, рРНК активно изучали в сравнительном плане. Каждая новая работа делала все более и более привлекательным для последующих авторов изучать именно рРНК, чтобы сравнить новые данные с непрерывно растущей базой данных, обеспечивалась своего рода положительная кооперативность в росте научного знания. Никакие другие гены в этом не могли конкурировать с рРНК. Среди слабо генетически изученных таксонов очень немного таких, где известные последовательности не включали бы генов рРНК, зато много таксонов, где гены рРНК – единственные, которые известны. Правда, в отношении бактерий, с прогрессом полногеномного секвенирования, это положение почти потеряло значение и быстро теряет его для эвкариотов. Но для прокариотов сохраняется особая роль генов рРНК как

потенциально наименее подверженных «горизонтальному» переносу. Гены рРНК все равно остаются особенно привлекательными ввиду наличия баз данных с выравниваниями, скорректированными на основании учета моделей вторичной структуры, интегрированных со специализированными средствами филогенетического анализа, поддерживающими различные сервисные функции в режиме удаленного доступа анонимных пользователей (Van de Peer et al., 2002; Cannone et al., 2002; Pruesse et al., 2007; Cole et al., 2009). Базы данных позволяют легко делать нужную таксономическую выборку. Например, база данных Ribosomal database project – II, версия 10, (<http://rdp.cme.msu.edu/>) на 5 октября 2009 г. содержит более 1 млн. 100 тыс. последовательностей 16S рРНК.

Часть ресурсов в настоящее время не поддерживается, но доступны в Интернете, как The European ribosomal RNA database (Van de Peer et al., 2002) (<http://bioinformatics.psb.ugent.be/webtools/rRNA/>). Другие, как Greengenes (<http://greengenes.lbl.gov/>), активно развиваются, в том числе разрабатывается специализированное программное обеспечение (DeSantis et al., 2003, 2006a, 2006b, Dalevi et al., 2007). Конечно, такое предпочтение в отношении рРНК не осуществилось, если бы сравнение рРНК имело только технические преимущества и было малосодержательным (Лебединский и др., 2007). Относительно большая протяженность 16S/18S рРНК давала простор для сравнения, наличие в составе макромолекулы участков с разной степенью консервативности позволяет проводить сравнения на разных таксономических уровнях. Это, естественно, не означает, что рРНК являются универсальными филогенетическими маркерами всех узлов филогенетического дерева. В некоторых группах, где рРНК изменялась в эволюции слишком быстро или слишком медленно, более выгодными оказываются другие маркеры. И все же из разных маркеров рРНК – наиболее универсальный. Такова сегодняшняя практика.

### 2.3. ПЦР: прорыв в новые области исследования и мелкие трудности.

Методика полимеразной цепной реакции (ПЦР) позволяет умножить

количество целевого продукта в миллионы раз. Это делает доступными для исследований микроскопически малые количества биологического материала, вплоть до единичных клеток, образцы плохой сохранности с частично деградированной ДНК (если сохранился хотя бы один протяженный фрагмент ДНК с полным геном рРНК на пробирку, то есть шанс «выловить» его с помощью ПЦР). По мере совершенствования методик рядовым делом стали анализы ДНК сухих гербарных образцов столетней давности, высохших капель крови, слюны, экскрементов (благодаря слущенным клеткам ротового или кишечного эпителия), волосяных луковиц и перьев, субфоссильных остатков и виде костей или мягких тканей, сохранявшихся в вечной мерзлоте. Примерно через 10 лет после адаптации техники ПЦР для амплификации рДНК она была применена для оценки разнообразия эукариотических микроорганизмов в природных сообществах путем анализа сложных смесей ДНК, выделенных из разнородной клеточной популяции.

Схематически такой анализ (рис. 2) включает ряд последовательных стадий и начинается с отбора проб и, в случае необходимости, их фракционирования (например, концентрации планктонных организмов, фракционирования их по размеру путем фильтрации сначала через грубый префильтр, затем сборе мельчайших частиц на мембранном фильтре, и т. п.).

На следующем этапе проводится выделение ДНК (или РНК) по одной из стандартных методик: субстрат, содержащий клетки, суспендируют в буфере, и добавлением детергента проводят лизис клеток; затем нуклеиновые кислоты очищают одним из способов. Необходимо учитывать, что некоторые субстраты могут содержать примеси, неблагоприятно сказывающиеся на последующих процедурах, а некоторые минеральные примеси, например, силикаты, обычные в почве, способны при определенных условиях сорбировать ДНК. В таких случаях может потребоваться оптимизация выделения и очистки ДНК.

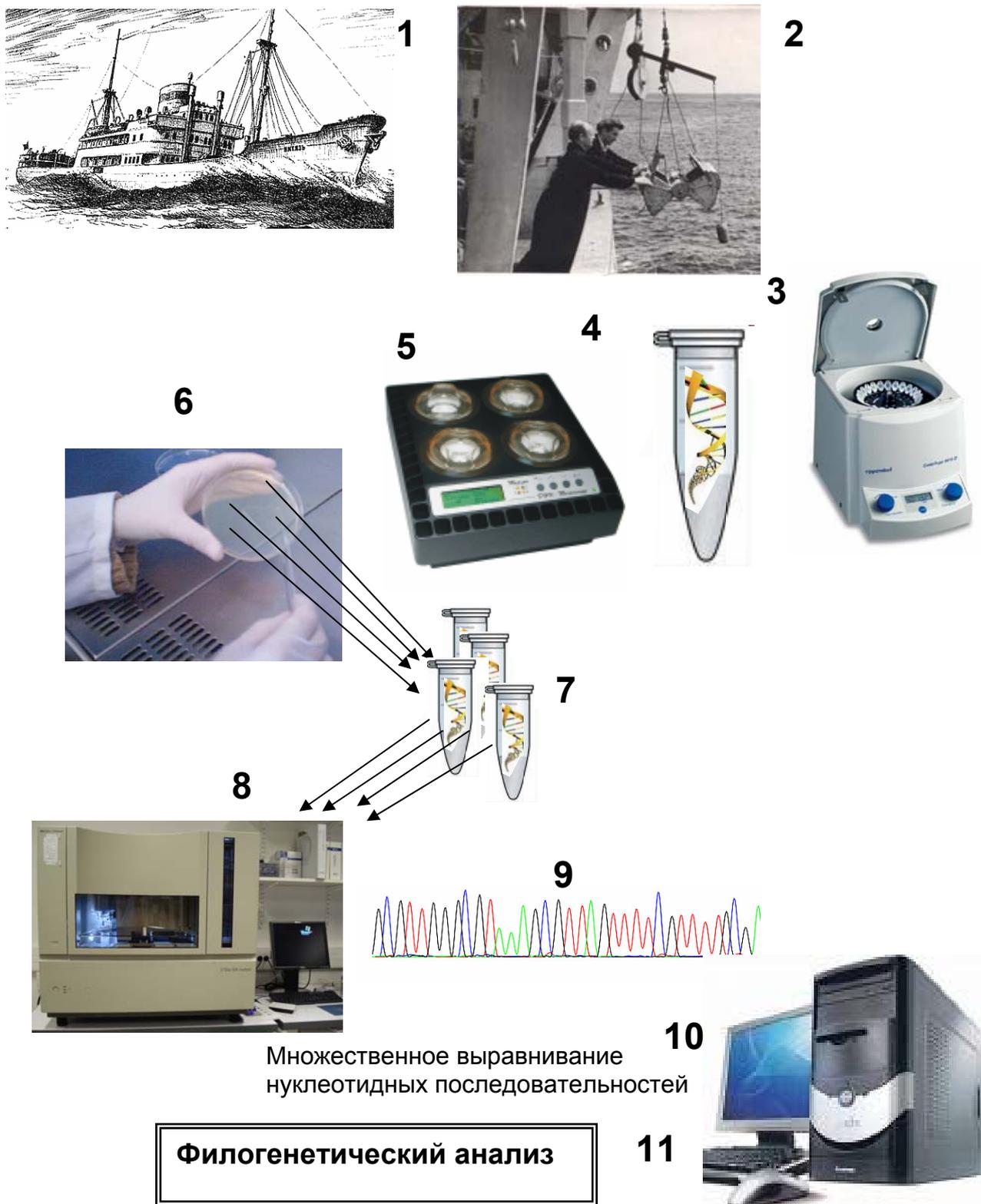


Рис. 2. Общая схема исследований ДНК из природных субстратов:  
 1, 2 – отбор биологических проб, фракционирование; 3, 4 – выделение ДНК;  
 5 – амплификация целевых последовательностей с помощью ПЦР; 6 –  
 клонирование фрагментов ДНК в *E.coli*; 7 – отбор клонов и выделение  
 плазмидной ДНК; 8, 9 – определение нуклеотидных последовательностей  
 плазмидных ДНК отдельных клонов; 10, 11 – анализ последовательностей.

Полученный препарат ДНК используют для ПЦР с парой олигонуклеотидных затравок (праймеров), комплементарных эволюционно консервативным участкам гена рРНК. Такие праймеры способны отжигаться с ДНК разных видов, присутствующих в смеси. Результатом ПЦР является смесь фрагментов близкого размера, но разной (поскольку они амплифицированы с ДНК разных видов), нуклеотидной последовательности. Для того, чтобы определить индивидуальную нуклеотидную последовательность, ее надо выделить из смеси. Традиционный способ – клонировать фрагменты ДНК в клетках кишечной палочки (рис. 2). Для этого полученные фрагменты ДНК, амплифицированные в ходе ПЦР, очищают от других продуктов реакции, лигируют с векторной плазмидной ДНК, трансформируют клетки кишечной палочки лигазной смесью и высевают на чашки Петри с селективной средой. Из выросших колоний отбирают рекомбинантные, из бактерий выделяют плазмидную ДНК и определяют нуклеотидную последовательность клонированного фрагмента. Теперь ее можно сравнить с имеющимися в базе данных последовательностями рРНК для идентификации, а также с последовательностями других рекомбинантных клонов, полученных в этом или других экспериментах. Учитывая, что до 99% бактерий не удается вывести в культуру, да и культивирование эукариотических микроорганизмов немногим проще, исследователи получили мощный инструмент исследования биологического разнообразия. И, между прочим, с удивлением обнаружили, что те штаммы бактерий, которые они выделяют из природы, как правило, не обнаруживаются при исследовании ДНК, выделенной непосредственно из субстрата. Это значит, что культивируемые штаммы бактерий, как правило, играют в природных сообществах незначительную роль из-за их низкой численности и являются высоко специализированными формами, способными к активному размножению в редких условиях, которые моделируются при лабораторном культивировании.

Несмотря на простоту изложенной схемы, имеется несколько источников неприятностей, возможных при проведении опытов. Они, в общем, выяснились вскоре после начала применения ПЦР для целей исследования микробных сообществ (von Wintzingerode et al., 1997). Одно из затруднений связано с неоднородностью микробного населения, преобладанием в нем немногих видов. На эти, наиболее массовые виды, может приходиться значительная доля клонов в клонотеке. Если поставлена задача оценить разнообразие, а не относительную численность видов, то, прежде чем проводить секвенирование сотен плазмид, сначала надо проверить три-четыре десятка. Если в пробной партии преобладает несколько типов последовательностей, то разумнее отказаться от дальнейшего изучения полученной клонотеки ради экономии средств. Если преобладающим будет три-четыре типа, то их можно выбраковывать перед секвенированием. Для этого перед выделением плазмидной ДНК полезно скринировать колонии: провести ПЦР с колоний, полученный продукт (фрагмент ДНК) гидролизовать мелкощепящей эндонуклеазой рестрикции (узнающей четырехнуклеотидную мишень в ДНК), полученный набор фрагментов ДНК анализировать электрофорезом в 1,5 – 2% агарозном геле, и колонии, порождающие мажорный тип распределения полос, исключить из дальнейшего анализа. Однако такая предварительная проверка сильно увеличивает трудозатраты и бывает недостаточно эффективна, если надо выбраковывать более чем 3-4 типа фрагментов. Другая методическая опасность, подстерегающая при ПЦР – опасность получения химер. Представим, что в ходе синтеза ДНК на какой-то конкретной матрице *Taq*-полимераза не успела скопировать участок между праймерами полностью (или произошел обрыв матрицы в результате фрагментации ДНК в ходе выделения или многократного нагревания при ПЦР). Такой неполный продукт сам сможет на следующем цикле выступить как длинный «праймер», а, учитывая высокое сходство рРНК разных видов, он может отжечься не только к «своему», но и к чужеродному гену рРНК (если обрыв

произошел не в вариабельной, а в консервативной области гена). В таком случае к последовательности вида X будут добавлены нуклеотиды, комплементарные матрице вида Y. Такие химеры, в зависимости от качества ДНК и таксономического разнообразия пробы, могут составлять от доли процента до нескольких процентов от числа клонов. Их химерность выявляется при анализе последовательностей. Разработаны компьютерные программы, облегчающие выявление химер. Однако простейшие автоматические процедуры не гарантируют выявление химерного происхождения последовательностей, возникшие на основе более чем двух матриц или же относящихся к таксонам, не представленным в базе данных.

Традиционный метод позволяет провести не только состав микробного населения, но и дать количественную оценку соотношения микроорганизмов в пробе – то, о чем раньше микробиолог не мог и мечтать, учитывая то обстоятельство, что более 90% микроорганизмов не удается вывести в культуру и идентифицировать. Теперь можно более адекватно оценить, какие виды более разнообразны в природе, а не те, которые лучше культивируются. Однако при учете населения по рРНК имеются источники ошибок и искажений. Самое главное – хотя праймеры и считаются «универсальными», нет такого закона природы, запрещающего эволюционные изменения даже самых важных макромолекул и их самых консервативных участков. При особых стечениях обстоятельств мутации в таких областях генов оказываются витальными, и могут фиксироваться. К такой ДНК «универсальный» праймер будет хуже подходить, этот ген будет менее эффективно амплифицироваться в ПЦР и окажется недопредставлен в смеси продуктов. К тому же результату приведет удлинение гена за счет протяженных инсерций в вариабельные участки (как в рРНК некоторых амёб, кинетопластид, эвгленовых и др.) или наличие в таких участках GC-богатых последовательностей, способных образовывать тугоплавкие «шпильки», затрудняющие «продвижение» *Taq*-полимеразы. Менее чем 15% разница в эффективности ПЦР приведет за 30 циклов к 100-кратной разнице

в количестве продукта (Sogin et al., 2006). Особо длинные, GC-богатые и тем более последовательности с измененными мишенями для отжига хотя бы одного праймера могут вообще потеряться при традиционной постановке ПЦР. Таким образом, количественные оценки по соотношению клонов в клонотеках следует рассматривать как предварительные, да и качественные оценки следует считать заниженными. Если у какого-то вида рибосомный оперон имеет много или мало копий в геноме (по сравнению с «средним» для эвкариотов 200-400 копиями), то это тоже скажется на выходе продукта и соотношении клонов. По этой же причине затруднено сравнение численности одноклеточных и многоклеточных. У бактерий также бывает по несколько копий гена 16S рРНК, причем, в этих случаях, обычны внутригеномные различия, которые вполне могут быть приняты за межвидовые. Необходимо также учитывать, что описание сообществ с высоким разнообразием микроорганизмов могут потребовать весьма обширных исследований для выявления более редких видов. По приводимым оценкам (Dunbar et al., 2002), для обнаружения половины видов из общего состава почвенной микрофлоры одной из станций в штате Аризона достаточно секвенировать 10 – 40 тыс. клонов. А для получения статистически значимых и воспроизводимых количественных данных (пренебрегая перечисленными выше источниками артефактов) может потребоваться на порядок больше определений, что является чересчур затратным. Новые данные, полученные с помощью методов высокопроизводительного секвенирования, заставляют, однако, считать такие оценки чересчур оптимистичными (см. ниже).

2.4. Методы высокопроизводительного секвенирования. Часть описанных выше проблем ПЦР (различная эффективность амплификации матриц, создание химер) до некоторой степени преодолеваются новыми методами высокопроизводительного секвенирования. Эти методы заменяют клонирование в живых клетках «молекулярным клонированием» с применением нанотехнологий – эмульсионной или твердофазной амплификацией фрагментов ДНК, или секвенированием отдельных

неамплифицированных фрагментов ДНК в пиколитровых реакторах. Первая из получивших распространение – технология эмульсионной ПЦР и пиросеквенирования, разработанная до уровня практического применения компанией 454 Life Sciences (см. обзор: Натальин, 2008). Технология 454 Life Sciences включает несколько этапов. Вначале ДНК фрагментируют и к концам фрагментов присоединяют адаптеры. Эти адаптеры позволяют ДНК присоединиться к специально приготовленным мельчайшим стеклянным шарикам. Далее раствор ДНК в буфере сильно разбавляют и смешивают с шариками, так чтобы на один шарик приходилось не более одного фрагмента ДНК. Такую суспензию шариков смешивают с маслом и получают тонкую эмульсию: взвесь мельчайших капелек буфера в масле. А в каждой капельке буфера – по одному стеклянному шарик с присоединенной молекулой ДНК. Затем проводят ПЦР. Затравками (праймерами) выступают иммобилизованные на стекло дезоксиолигонуклеотды, комплементарные адаптерам. В каждой капельке один шарик и одна молекула матрицы, в ходе ПЦР она амплифицируется в миллионы раз, и ее копии окажутся присоединены к шарик. Здесь исключается конкуренция матриц за *Taq*-полимеразу и создание химер. Остается отмыть шарики от масла, отделить те из них, которые по случайности остались без ДНК и переходить к следующему этапу – секвенированию. Для этого шарики помещают на пластинку с лунками, по одному шарик на лунку. На сегодня применяют пластики с плотностью 480 лунок на квадратный миллиметр и общим числом 1 млн. 600 тыс. лунок на пластинку. Именно столько фрагментов ДНК одновременно анализируется в опыте. Далее на пластинку циклически подают реагенты для проведения секвенирующей реакции. О встраивании подаваемого нуклеотида судят по освобождаемому пирофосфату, который детектируют с помощью дополнительных ферментативных реакций, сопровождающихся люминесценцией. Дно пластинки соединено с оптоволоконным световодом для регистрации люминесценции (при подаче какого нуклеотида она происходит в каждой лунке).

Технология 454 позволяет одновременно анализировать 1 млн. 600 тыс. фрагментов ДНК и за один опыт полностью определить небольшой бактериальный геном. Если в смеси присутствует ДНК нескольких видов, то, по результатам нескольких опытов, по перекрывающимся фрагментам можно составить полный (или почти полный) геном преобладающего вида и более или менее протяженные контиги (наборы перекрывающихся фрагментов) более редких в пробе видов. Такие данные будут метагеномом. Природный метагеном морского планктона или кишечника человека включает геномы бактерий, вирусов и бактериофагов, эукариотических микроорганизмов и многоклеточных, хотя, обычно, на стадии подготовки проб ограничиваются микроскопически малыми объектами из размерного класса бактерий (рыб и червей из пробы выбрасывают). Определить метагеном – задача в тысячи раз сложнее, чем, скажем, геном человека. Видовая принадлежность большинства найденных генов остаются неизвестной, но тип кодируемого фермента (эстераза, гликозилгидролаза и т. д.) может быть установлена методами биоинформатики, а найденные гены – экспрессированы в бактериальных клетках и испытаны для нужд биотехнологии. Небольшая часть метагенома (примерно 0,01%) приходится на гены рРНК, и они, с большей или меньшей точностью, могут быть отнесены к известным (или неизвестным) таксонам. Ведь только для единственного гена, рРНК малой субъединицы, собрана достаточно представительная выборка, позволяющая определить систематическое положение исследуемого организма с приемлемой точностью. Вначале технология 454 была применена для изучения разнообразия населения «необычных» местообитаний: глубоководных морских станций (Sogin et al., 2006; Huber et al., 2007), железорудной шахты в Миннесоте (Edwards et al., 2006), бактериальных матов горячих источников Йеллоустонского парка (Miller et al., 2009). Затем эта же технология применена к учету бактериальной флоры ротовой полости, гортани и фекалий человека (Andersson et al., 2008; Keijsers et al., 2008) и других сообществ.

Если технологию 454 применить не к геномной ДНК, а к амплифицированным с помощью «обычной» ПЦР фрагментам рРНК, то почти каждый из 1 млн. 600 тыс. образцов можно будет отнести к таксону более или менее высокого ранга, и население изучаемой пробы будет определено с большой точностью. Ввиду огромного числа анализируемых фрагментов, будут обнаружены даже очень редкие организмы, представленные небольшим числом клеток (если их гены рРНК амплифицируются в ПЦР). Насколько полно такая техника позволяет учесть микробное разнообразие? Рассмотрим в качестве наглядного примера исследование с применением технологии 454 двух близко расположенных, но сильно отличающихся по химизму и микробному населению глубоководных гидротерм (Huber et al., 2007). В этой работе мишенью для ПЦР служил небольшой, около 100 н. п. длиной, гипервариабельный фрагмент гена 16S рРНК. Предполагается, что ввиду малого размера матрицы разница в эффективности амплификации разных последовательностей пренебрежимо мала. Клетки были собраны фильтрацией одного или двух литров морской воды через мембрану с диаметром пор 0,2 мкм. Общий титр в обеих станциях составлял около  $10^5$  клеток на 1 мл. Всего из двух библиотек было секвенировано около 900 тыс. последовательностей эвбактерий и архей, всего обнаружено более 36 тыс. различных вариантов. Если исключить варианты, отличающиеся единичными нуклеотидами и установить порог в 3% различий (такой порог обычно микробиологи принимают для разных «видов»), то число обнаруженных вариантов оказывается чуть больше 20 тыс. Более 98% вариантов были очень редкими – встречались менее чем 100 раз на 900 тыс. образцов каждый и, в основном, не могли быть обнаружены методом капиллярного секвенирования клонотек. Экстраполяция кривой насыщения в зависимости от размера выборки дает оценку суммарного бактериально-архейного разнообразия в двух станциях (с порогом не менее 3% различий в сравниваемой вариабельной области 16S рРНК) в 40 тыс.

видов – при условии, если б было секвенировано не 900 тыс., а «бесконечное» число образцов с этих двух станций.

Таким образом, для исчерпывающего анализа микробного разнообразия даже производительность технологии 454 недостаточна, хотя для анализа простой смеси генов или выявления наиболее массовых компонентов запуск прибора даже является избыточным. Для этих случаев разработаны методы мультиплексного анализа, например, индивидуальная, специфическая для пробы метка в составе адаптера, их могут быть десятки, а в перспективе – тысячи (Hamady et al., 2008), или можно, приготовив несколько независимых библиотек (несколько геномов, несколько проб природных субстратов) нанести их на разные поля рабочей ячейки. Ячейка может быть разделена на 2, 4 или 8 ролей.

Пока, однако, и технология 454, и другие технологии высокопроизводительного секвенирования, не вытеснили традиционного анализа клонотек, как атомное оружие не вытеснило стрелковое и даже холодное оружие, для применения которых остается своя ниша. Это определяется дороговизной приборов (до полумиллиона долларов США) и расходных материалов. Например, затраты на расходные материалы по технологии 454 составляют 10 – 20 тыс. долларов США на один опыт. Хотя цена, в пересчете на одну сиквенсную реакцию, меньше примерно в 1000 раз по сравнению с капиллярным секвенированием, общая высокая цена опыта заставляет каждый раз задумываться о целесообразности его проведения.

### **3. Виды-двойники и их молекулярная идентификация**

Необходимость дополнительных методов учета видового разнообразия, кроме световой микроскопии, была поставлена на повестку открытием так называемых видов-двойников. В протистологии эта проблема известна с 50-х гг., начиная с работ Соннеборна по «типам спаривания» инфузорий *Paramecium* (Sonneborn, 1975). Для форм, внешне неотличимых, но

изолированных генетически, сравнение генетического материала, ДНК, является прямым и наиболее адекватным способом выявления различий. Заметим, с развитием программы «ДНК-штрихкодирования» (Hebert et al., 2003; Шнеер 2009; Лухтанов, Кузнецова, 2009) этот способ постепенно становится и наиболее дешевым, в перспективе выводящим видовую диагностику из искусства отдельных высококвалифицированных или уникальных специалистов по группе в разряд дешевого лабораторного анализа, выполняемого лаборантом. Программа автоматической видовой идентификации по молекулярно-генетическим маркерам обещает перевести обработку биологического материала, в первую очередь массовых проб в экологии, на совершенно новый уровень.

Сравнение нуклеотидных последовательностей молекулярных маркеров не является исключительно эмпирическим занятием. Оно идет об руку с теорией молекулярной эволюции. В частности, теория указывает, что вслед за генетической изоляцией популяций, неважно по какой причине произошедшей, неизбежно начинается генетическая дивергенция, неуклонно усиливающаяся по ходу независимой эволюции. Сколько-нибудь продолжительная изоляция неизбежно приведет к различиям хотя бы в наиболее эволюционно переменчивых маркерных последовательностях. В дальнейшем будут приобретены различия и в более консервативных последовательностях. Например, прازیнофітова водоросль “*Micromonas pusilla*”, массовый компонент пикопланктона, распространена всесветно и не обнаруживает в строении клеток в разных частях Мирового океана особенностей, позволяющих выделить разные виды. Но анализ нуклеотидных последовательностей маркерных генов культур, хранящихся в международных коллекциях (Šlapeta et al., 2006a), обнаруживает большое число генетически отличающихся в разной степени форм, в том числе симпатрических; с другой стороны, тождественные формы занимают гигантский ареал (от тихоокеанского побережья Канады до южного побережья Австралии). Степень различий позволяет оценить время

расхождения форм «морфовида» “*Micromonas pusilla*”. Для наиболее далеких форм это более 60 млн. лет, то есть немногим меньше времени расхождения семейств цветковых растений и отрядов млекопитающих. А на основании имеющихся в базах данных анонимных последовательностей из планктона различных областей Мирового океана, можно сделать еще один вывод: генетическое разнообразие форм “*Micromonas pusilla*” далеко не исчерпывается штаммами, хранящимися в коллекциях (Šlapeta et al., 2006a).

Не менее впечатляющий пример представляет разнообразие гетеротрофных жгутиконосцев бодонид (Excavata, Kinetoplastida). Среди них найдено две примечательных группы фактов: 1) наличие далеко разошедшихся линий в пределах «морфовидов»; 2) формы с идентичной морфологией, независимо возникшие в различных частях филогенетического дерева (Scheckenbach et al., 2006). Если первая группа расширяет наблюдения над другими таксонами, например, “*Micromonas pusilla*”, то вторая требует отдельного разбора. Для многоклеточных мы знаем о необратимости эволюции (часто именуемый также законом Долло). Одно из следствий закона Долло – запрет на неоднократное возникновение того же самого вида. Очевидно, в случае одноклеточных бедность морфологического диагноза приводит к недостаточной разрешающей способности систематики, к многочисленным ошибкам при определении видов по морфологическим признакам и видимому нарушению закона Долло.

Опираясь на работы, сочетающие выделение культур и их видовую идентификацию в рамках традиционной парадигмы с молекулярно-генетическим анализом, мы получаем основу для правильной оценки результатов анализа природных проб, выполненных только с применением технологий ДНК.

#### 4. Таксоны высокого ранга в анонимных последовательностях

4.1. История “Marine alveolate Group”. Среди вопросов, наиболее естественных по отношению к анонимным нуклеотидным последовательностям генов рРНК, выделяется следующий вопрос: а есть ли среди этих последовательностей такие, что принадлежат неизвестным организмам? То есть новым, неизвестным ранее видам, классам, типам... В одной из ранних работ (Moon-van der Staay et al., 2001) из 35 клонов, полученных после амплификации и клонирования гена 18S рРНК из проб тихоокеанского пикопланктона (клеток размером менее 3 мкм), только два оказались практически идентичными (более 99% сходства) с 18S рРНК известных видов. Остальные в разной степени походили на известные последовательности, но не могли быть отнесены на момент публикации к какому-либо определенному виду. Конечно, одна из причин состояла в том, что далеко не для всех описанных видов (тогда, да и сейчас) известны последовательности 18S рРНК. Но 6 из 35 последовательностей не удалось отнести к какому-либо известному типу (Moon-van der Staay et al., 2001)! Одновременно, в том же номере журнала, в статье других авторов (López-García et al., 2001) были описаны сходные последовательности, найденные в пикопланктоне приполярной зоны Атлантического и Индийского океанов. Станции были взяты на разных глубинах, от поверхности до 3000 м. Стало ясно, что речь идет о всесветно распространенном и богатом виде таксоне. Особенно удивительным казался факт близости новых организмов к панцирным жгутиконосцам – динофлагеллятам, но озадачивал малый размер, а также то, что они обнаруживались в значительном количестве и в афотической зоне, то есть вряд ли могли быть водорослями. Притом неизвестные организмы были гораздо богаче динофлагеллят представлены в клонотеках и отличались между собой в большей степени, чем известные динофлагелляты. Они получили условное название “Marine alveolate Group II” и “Marine alveolate Group I” (López-García et al., 2001). В дальнейшем

“Marine alveolate Group II”, сестринскую группу динокариотических панцирных жгутиконосцев, удалось идентифицировать с «динофлагеллятами» порядка Syndinales (Moon-van der Staay et al., 2001) – паразитами, главным образом, одноклеточных. Их видовое разнообразие и массовое присутствие в планктоне явилось полной неожиданностью и заставляет пересматривать экологическую роль прессы паразитов. “Marine alveolate Group I” – сестринская группа динофлагеллят, включая Syndinales, оставалась загадочной практически до прошлого года, когда были одновременно с применением микроскопической и молекулярной техники были переисследованы *Duboscquella*, паразит инфузорий-тинтинид (Harada et al., 2007), и *Paradinium*, паразит веслоногих ракообразных (Skovgaard, Daugbjerg, 2008). Еще один представитель “Marine alveolate Group I” (Mori et al. 2007; Yuasa et al. 2007) тоже оказался известным видом, считавшимся паразитической динофлагеллятой – *Ichthyodinium chabelardi* (Hollande, Cachon, 1952), и тоже паразитом «одноклеточных» – икры пелагических рыб. Исследования ультраструктуры показали плезиоморфные черты в организации клеток *I. chabelardi* (Gestal et al., 2006). В частности, наличием псевдоконоида они более напоминают *Perkinsus* и споровиков, чем панцирных жгутиконосцев.

Таким образом, история “Marine alveolate Group” I и II оказывается весьма поучительной. Благодаря исследованиям проб ДНК из природных сообществ было обращено внимание на крупную группу – таксон протистов высоко ранга. Эти организмы не оказались новыми для науки, но их статус был, согласно новым данным, пересмотрен и повышен. Это стимулировало дополнительные исследования их морфологии, которые подтвердили высокий таксономический ранг и оказались важны для понимания эволюции Alveolata в целом. Из 500 ссылок в научной литературе на статьи (López-García et al., 2001; Moon-van der Staay et al., 2001) на конец 2009 г. (согласно данным ISI Web of Knowledge), заметная часть приходится на “Marine alveolate Group”. Пересмотрена экологическая роль этих целиком

паразитических одноклеточных, которые оказались неожиданно разнообразными и массовыми в различных сообществах (Moreira, López-García 2003; Lefèvre et al. 2008). Впрочем, количественные оценки, возможно, в будущем потребует коррекции. Различная представленность в библиотеках ДНК и кДНК и некоторые другие наблюдения поддерживают гипотезу о непропорционально высокой копийности рибосомных оперонов в их геномах по сравнению с «микроводорослями» *Pelagomonas*, *Ostreococcus*, *Micromonas* и другими представителями пикопланктона (Massana et al., 2008). Это может объяснить «избыточную» долю “Marine alveolate Group” в клонотеках ДНК. Однако, если оценки численности и могут быть пересмотрены и скорректированы, то оценки высокого разнообразия являются абсолютными и бесспорными.

Подобным образом последовательности рРНК ряда других таксонов высокого ранга впервые были получены из проб природных субстратов. На момент их обнаружения они не были идентифицированы и рассматривались как «независимые эволюционные линии». Первоначально неидентифицированные рРНК из одной только работы (Dawson, Pace, 2002) были позднее отнесены к амебам, грегариам (Verney et al., 2004), кольподеллам (Leander et al., 2003), архигрегариам (Leander et al. 2006), а часть последовательностей остается до сих пор неидентифицированными. Вначале были определены в природных пробах как неидентифицированные, а затем идентифицированы последовательности катаблефарид (Šlapeta et al., 2006a) и некоторых других протистов.

Расширение работ по рРНК природных образцов выдвинуло задачу их обобщения в отношении перспектив обнаружения новых эукариотических групп высокого таксономического ранга, «царств» (Verney et al., 2004). Более тщательный анализ первичных данных за счет поиска химер и использования более совершенных программ построения филогенетических деревьев для правильного размещения на дереве высоко дивергированных последовательностей, а также расширения таксономической выборки в базах

данных, привел к пересмотру начальных оценок в сторону их уменьшения. Так, если первоначально среди 289 различных последовательностей (филотипов) авторы относили 9,7% к «новым» для науки линиям эукариотов, то повторный анализ выявил в них 13,8% химерных последовательностей, и только 3,5% оставались не идентифицированными до типа, представляя собой потенциально «новые» для науки группы (рис. 3). Часть из них сохраняет этот статус до настоящего времени.

4.2. Таксоны, первоначально обнаруженные в анонимных рРНК. В предыдущем разделе мы рассмотрели некоторые примеры, когда нуклеотидные последовательности генов рРНК представителей известных типов или классов были впервые получены из природных субстратов, а не известных, определенных традиционными методами лабораторных культур. В рассмотренных случаях это были, как правило, новые виды, или даже неизвестные науке семейства. Но более интересными кажутся находки таких организмов, которых протистологи никогда не видели в микроскоп – представителей новых для науки типов. Их поиск оказывается занимательной «охотой на микробов», пример которой можно видеть в исследовании глубоководных холодных метановых сипов вблизи Японии (Takishita et al., 2007). Стерильно отобранные пробы из этих обедненных кислородом местообитаний авторы разделяли на две части, одну из которых использовали для выделения ДНК, а другую – для создания накопительных культур на двух типах сред. Накопительные культуры выдерживали 10 суток в анаэробных условиях при 4 °С.

Оказалось, что в накопительной культуре на среде 5% RYNFH (ATCC) доминирует гетеротрофный жгутиконосец, последовательность рРНК которого практически идентична таковой, полученной из природных проб.

Микроскопическое исследование выявило у клеток признаки внешнего строения, свойственные представителям Excavata (глубокую вентральную бороздку, проксимально расширенный задний жгутик и др.).

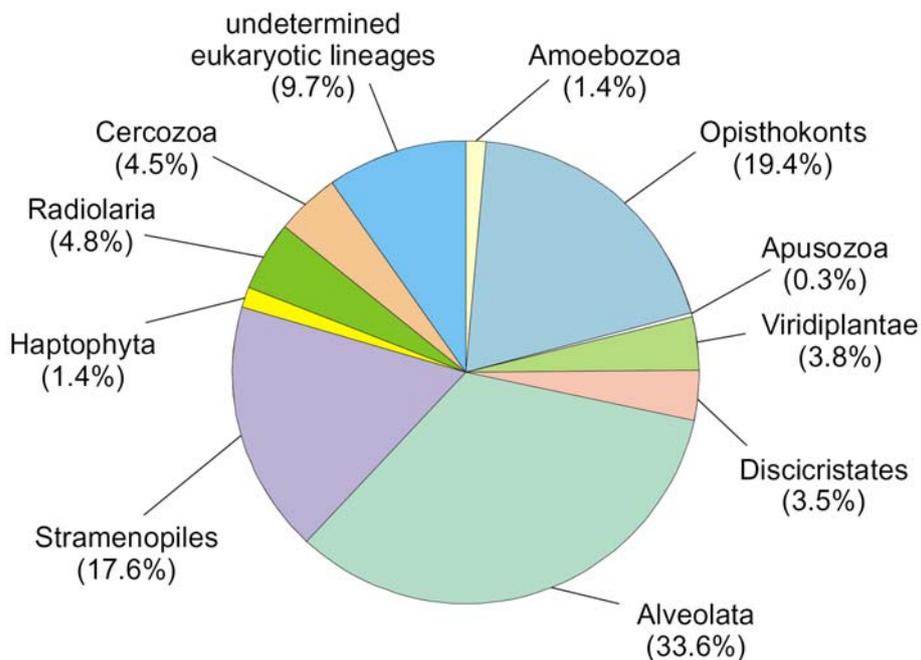
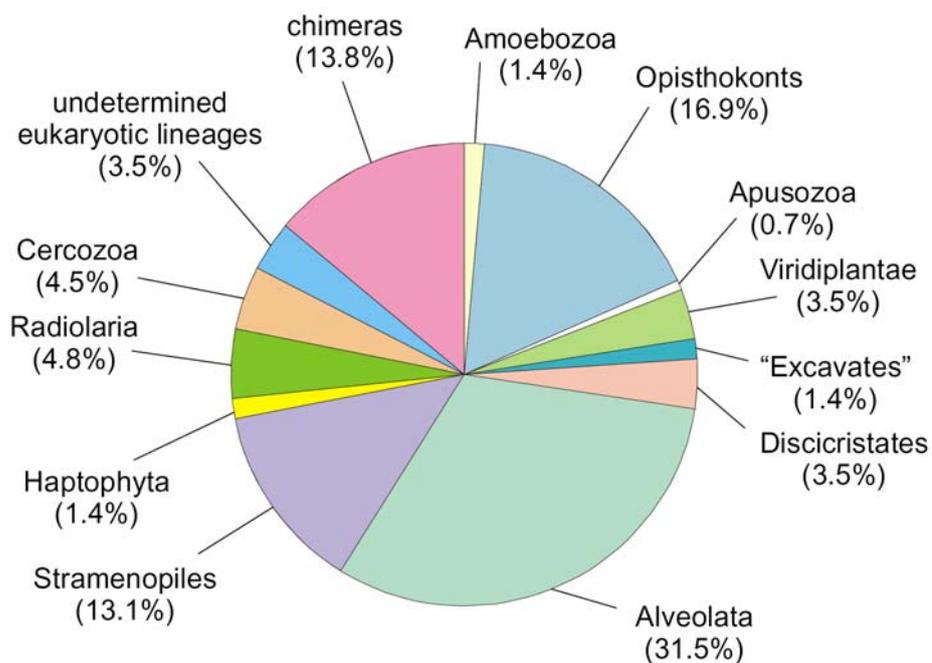
**A****B**

Рис. 3. Соотношение различных групп эукариотов в пробах рДНК из природных субстратов на основании оригинальных оценок (А) и по результатам переисследования оригинальных данных (В). Из: (Berney et al., 2004).

Филогенетический анализ последовательностей 18S рРНК помещает выделенный организм к особую, изолированную кладу Excavata, сестринскую ретортамонадидам и карпедиемонадидам. Ранее 18S рДНК ее представителей была обнаружена в осадке гидротерм в Калифорнийском заливе (Edgcomb et. al., 2002). Такишита с соавт. (Takishita et. al., 2007) не описали выделенного жгутиконосца, обещая сделать это в отдельной работе, однако снабдили статью его фотографиями.

Находки не ограничиваются новыми для науки классами. Давно идет «охота» на «морских stramenoplies», MUST (для обзора см.: Massana et al., 2004), на которых может приходиться до 20% последовательностей клонотек, полученных из кДНК пикопланктона (Not et al., 2009). Недавно представлены доказательства филогенетической обособленности неизвестных пикопланктонных организмов, получивших наименование picobiliphytes (Not et al., 2007) или biliphytes (Cuvelier al., 2008). На филогенетических деревьях 18S рРНК они оказываются сестринской группой криптофитовых, хотя статистическая поддержка такой группировки невысокая (Cuvelier al., 2008). То есть их ранг примерно соответствует рангу криптофитовых и гаптофитовых. Пока что эти организмы идентифицированы по последовательностям 18S рРНК и гибридизацией *in situ* со специфическими зондами (FISH), изучен их пигментный состав, свидетельствующий в пользу их способности к фотосинтезу, нанесены на карту географические координаты мест их нахождения (атлантика, пацифика, полярные и тропические воды), изучена приуроченность к теплым (30 °С) и холодным (5°С) водным массам. Picobiliphytes или biliphytes пока остаются не выделенными в культуру и не описанными организмами.

Другие, более редко встречающиеся в природе чем picobiliphytes представители изолированных групп на филогенетических деревьях 18S рРНК еще ждут приговора, являются ли они членами неизвестных науке клад или дериватами известных. И список таких кандидатов продолжает

пополняться, несмотря на существенное расширение референсной базы данных.

### Список литературы

1. *Беклемишев В.Н.* Методология систематики (рукопись 1928 г.). М.: КМК Ltd., 1994. 250 с.
2. *Лебединский А.В., Черных Н.А., Бонч-Осмоловская Е.А.* 2007. Геносистематика микроорганизмов термальных местообитаний // Биохимия. Т. 72. № 12. С. 1594–1609.
3. *Лухтанов В.А., Кузнецова В.Г.* Молекулярно-генетические и цитогенетические подходы к проблемам видовой диагностики, систематики и филогенетики // Журн. общ. биологии. 2009. Т 70. № 5. С. 415–437.
4. *Натальин П.* 2008. 454-секвенирование (высокопроизводительное пиросеквенирование ДНК) // Биомолекула. (<http://biomolecula.ru/content/214>)
5. *Старынкевич К.Д.* Строение жизни. Прага: Politika, 1931. 36 с.
6. *Шнеер В.С.* ДНК-штрихкодирование видов животных и растений – способ их молекулярной идентификации и изучения биоразнообразия // Журн. общ. биологии. 2009. Т 70. № 4. С. 296–315.
7. *Andersson A.F., Lindberg M., Jakobsson H., Bäckhed F., Nyrén P., Engstrand L.* 2008. Comparative analysis of human gut microbiota by barcoded pyrosequencing // PLoS One. V. 3. № 7. P. e2836.
8. *Berney C., Fahrni J., Pawlowski J.* 2004. How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys // BMC Biol. V. 2. P. 13.
9. *Cannone J.J., Subramanian S., Schnare M.N., Collett J.R., D'Souza L.M., Du Y., Feng B., Lin N., Madabusi L.V., Muller K.M., Pande N., Shang Z., Yu N., Gutell R.R.* 2002. The comparative RNA web (CRW) site: an online database

of comparative sequence and structure information for ribosomal, intron, and other RNAs // BMC Bioinformatics. V. 3. P. 2.

10. *Cole J.R., Wang Q., Cardenas E., Fish J., Chai B., Farris R.J., Kulam-Syed-Mohideen A.S., McGarrell D.M., Marsh T., Garrity G.M., Tiedje J.M.* 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis // Nucleic Acids Res. V. 37. Database issue. D141–D145.
11. *Cuvelier M.L., Ortiz A., Kim E., Moehlig H., Richardson D.E., Heidelberg J.F., Archibald J.M., Worden A.Z.* 2008. Widespread distribution of a unique marine protistan lineage // Environ. Microbiol. V. 10. № 6. P. 1621–1634.
12. *Dawson S.C., Pace N.R.* 2002. Novel kingdom-level eukaryotic diversity in anoxic environments // Proc. Natl. Acad. Sci. USA. V. 99. № 12. P. 8324–8329.
13. *Dalevi D., DeSantis T.Z., Fredslund J., Andersen G.L., Markowitz V.M., Hugenholtz P.* 2007. Automated group assignment in large phylogenetic trees using GRUNT: GRouping, Ungrouping, Naming Tool // BMC Bioinformatics. V. 8. № 1. P. 402.
14. *DeSantis T.Z., Dubosarskiy I., Murray S.R., Andersen G.L.* 2003. Comprehensive aligned sequence construction for automated design of effective probes (CASCADE-P) using 16S rDNA // Bioinformatics. V. 19. № 12. P. 1461-1468.
15. *DeSantis T.Z., Hugenholtz P., Keller K., Brodie E.L., Larsen N., Piceno Y.M., Phan R., Andersen G.L.* 2006a. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes // Nucleic Acids Res. V. 34. Web Server issue. W394-W399.
16. *DeSantis T.Z., Hugenholtz P., Larsen N., Rojas M., Brodie E.L., Keller K., Huber T., Dalevi D., Hu P., Andersen G.L.* 2006b. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB // Appl. Environ. Microbiol. V. 72. № 7. P. 5069–5072.

17. *Dunbar J., Barns S.M., Ticknor L.O., Kuske C.R.* 2002. Empirical and theoretical bacterial diversity in four Arizona soils // *Appl. Environ. Microbiol.* V. 68. № 6. P. 3035–3045.
18. *Edgcomb V.P., Kysela D.T., Teske A., de Vera Gomez A., Sogin M.L.* 2002. Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment // *Proc. Natl. Acad. Sci. USA.* V. 99. № 11. P. 7658–7662.
19. *Edwards R.A., Rodriguez-Brito B., Wegley L., Haynes M., Breitbart M., Peterson D.M., Saar M.O., Alexander S., Alexander E.C. Jr, Rohwer F.* 2006. Using pyrosequencing to shed light on deep mine microbial ecology // *BMC Genomics.* V. 7. P. 57.
20. *Gestal C., Novoa B., Posada D., Figueras A., Azevedo C.* 2006. *Perkinsoide chabelardi* n. gen., a protozoan parasite with an intermediate evolutionary position: possible cause of the decrease of sardine fisheries? // *Environ. Microbiol.* V. 8. № 6. P. 1105–1114.
21. *Hamady M., Walker J.J., Harris J.K., Gold N.J., Knight R.* 2008. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex // *Nat. Methods.* V. 5. № 3. P. 235–237.
22. *Harada A., Ohtsuka S., Horiguchi T.* 2007. Species of the parasitic genus *Duboscquella* are members of the enigmatic Marine Alveolate Group I // *Protist.* V. 158. P. 337–347.
23. *Hebert P.D., Cywinska A., Ball S.L., de Waard J.R.* 2003. Biological identifications through DNA barcodes // *Proc. Biol. Sci.* V. 270. P. 313–321.
24. *Hollande A., Cachon J.* 1952. Un parasite des oeufs de sardine: l'*Ichthyodinium chabelardi*, nov. gen. nov. sp. (Péridinien parasite) // *C. R. Acad. Sci. Paris.* V. 235. S. 976–977.
25. *Huber J.A., Mark Welch D.B., Morrison H.G., Huse S.M., Neal P.R., Butterfield D.A., Sogin M.L.* 2007. Microbial population structures in the deep marine biosphere // *Science* 318:97-100.
26. *Keijser B.J., Zaura E., Huse S.M., van der Vossen J.M., Schuren F.H., Montijn R.C., ten Cate J.M., Crielaard W.* 2008. Pyrosequencing analysis of

- the oral microflora of healthy adults // *J. Dent. Res.* V. 87. № 11. P. 1016–1020.
27. *Leander B.S., Kuvardina O.N., Aleshin V.V., Myl'nikov A.P., Keeling P.J.* 2003. The free-living ancestry of apicomplexans: phylogeny of colpodellids (Alveolata) as inferred from small subunit rDNA // *J. Eukaryot. Microbiol.* V. 50. № 5. P. 334–340.
  28. *Leander B.S., Lloyd S.A.J., Marshall W., Landers S.C.* 2006. Phylogeny of Marine Gregarines (Apicomplexa) – *Pterospora*, *Lithocystis* and *Lankesteria* – and the origin(s) of coelomic parasitism // *Protist.* V. 157. № 1. P. 45–60.
  29. *Lefèvre E., Roussel B., Amblard C., Sime-Ngando T.* 2008. The molecular diversity of freshwater picoeukaryotes reveals high occurrence of putative parasitoids in the plankton // *PLoS ONE.* V. 3. P. E2324.
  30. *López-García P., Rodríguez-Valera F., Pedrós-Alió C., Moreira D.* 2001. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton // *Nature.* V. 409. № 6820. P. 603–607.
  31. *Massana R., Castresana J., Balagué V., Guillou L., Romari K., Groisillier A., Valentin K., Pedrós-Alió C.* 2004. Phylogenetic and ecological analysis of novel marine stramenopiles // *Appl. Environ. Microbiol.* V. 70. № 6. P. 3528–3534.
  32. *Massana R., Karniol B., Pommier T., Bodaker I., Bèjà O.* 2008. Metagenomic retrieval of a ribosomal DNA repeat array from an uncultured marine alveolate // *Environ. Microbiol.* V. 10. № 5. P. 1335–1343.
  33. *Medlin L., Elwood H.J., Stickel S., Sogin M.L.* The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions // *Gene.* 1988. V. 71. № 2. P. 491–499.
  34. *Miller S.R., Strong A.L., Jones K.L., Ungerer M.C.* 2009. Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park // *Appl. Environ. Microbiol.* V. 75. № 13. P. 4565–4572.

35. Moon-van der Staay S.Y., De Wachter R., Vaolot D. 2001. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity // Nature. V. 409. № 6820. P. 607–610.
36. Moreira D., López-García P. 2003. Are hydrothermal vents oases for parasitic protists // Trends Parasitol. V. 19. P. 556–558.
37. Mori K., Yamamoto K., Teruya K., Shiozawa S., Yoseda K., Sugaya T., Shirakashi S., Itoh N., Ogawa K. 2007. Endoparasitic dinoflagellate of the genus *Ichthyodinium* infecting fertilized eggs and hatched larvae observed in the seed production of leopard coral grouper *Plectropomus leopardus* // Fish Pathology. V. 42. P. 49–57.
38. Not F., Valentin K., Romari K., Lovejoy C., Massana R., Töbe K., Vaolot D., Medlin L.K. 2007. Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes // Science. V. 315. № 5809. P. 253–255.
39. Not F., del Campo J., Balagué V., de Vargas C., Massana R. 2009. New Insights into the Diversity of Marine Picoeukaryotes // PLoS ONE. V. 4. № 9. e7143.
40. Pruesse, E., Quast C., Knittel K., Fuchs B., Ludwig W., Peplies J., Glöckner F. O. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB // Nucleic Acids Res. V. 35. № 21. P. 7188–7196.
41. Scheckenbach F., Wylezich C., Mylnikov A.P., Weitere M., Arndt H. 2006. Molecular comparisons of freshwater and marine isolates of the same morphospecies of heterotrophic flagellates // Appl. Environ. Microbiol. V. 72. № 10. P. 6638–6643.
42. Skovgaard A., Daugbjerg N. 2008. Identity and Systematic Position of *Paradinium poucheti* and Other *Paradinium*-Like Parasites of Marine Copepods Based on Morphology and Nuclear-Encoded SSU rDNA // Protist. V. 159. P. 401–413.

43. Šlapeta J., López-García P., Moreira D. 2006a. Global dispersal and ancient cryptic species in the smallest marine eukaryotes // *Mol. Biol. Evol.* V. 23. № 1. P. 23–29.
44. Šlapeta J., López-García P., Moreira D. 2006b. Present Status of the Molecular Ecology of Kathablepharids // *Protist.* V. 157. № 1. P. 7–11.
45. Sogin M.L., Morrison H.G., Huber J.A., Mark Welch D., Huse S.M., Neal P.R., Arrieta J.M., Herndl G.J. 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere” // *Proc. Natl. Acad. Sci. USA.* V. 103. № 32. P. 12115–12120.
46. Sonneborn T.M. 1975. The *Paramecium aurelia* complex of fourteen sibling species // *Trans. Am. Micro. Soc.* 94. № 2. P. 155–178.
47. Takishita K., Yubuki N., Kakizoe N., Inagaki Y., Maruyama T. 2007. Diversity of microbial eukaryotes in sediment at a deep-sea methane cold seep: surveys of ribosomal DNA libraries from raw sediment samples and two enrichment cultures // *Extremophiles.* V. 11. № 4. P. 563–576.
48. Van de Peer Y., De Rijk P., Wuyts J., Winkelmanns T., De Wachter R. 2000. The European small subunit ribosomal RNA database // *Nucleic Acids Res.* V. 28. № 2. P. 175–176.
49. von Wintzingerode F., Göbel U.B., Stackebrandt E. 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis // *FEMS Microbiol. Rev.* V. 21. № 3. P. 213–229.
50. Yuasa K., Kamaishi T., Mori K., Hutapea J.H., Permana G.N., Nakazawa A. 2007. Infection by a protozoan endoparasite of the genus *Ichthyodinium* in embryos and yolk-sac larvae of yellowfin tuna *Thunnus albacares* // *Fish Pathology.* V. 42. P. 59–66.