

ОТЗЫВ официального оппонента
на диссертацию на соискание ученой степени
кандидата биологических наук Гарушянц Софьи Константиновны
на тему: «Структура и эволюционная динамика прокариотических сообществ
необычных местообитаний»
по специальности 03.01.09 – «математическая биология, биоинформатика»

Современные биологические, в особенности микробиологические, исследования уже невозможны без использования геномного анализа. Его введение в рутинную лабораторную практику существенно расширило круг доступных для решения задач и сфокусировало внимание на эволюционных аспектах исследуемых явлений. Возможно, самым интригующим в этом направлении является изучение симбиотических бактериальных сообществ, которое, позволяет получить надёжный результат о закономерностях эволюционного процесса за приемлемое время. Рецензируемая работа посвящена филогенетическому анализу природных симбиотических сообществ и анализу геномов мало исследованных таксонов. Она в полной мере своевременна, актуальна и фундаментально значима. Особого упоминания заслуживают результаты поиска бактериальных генов в геномах архей, которые определили диапазон скоростей обмена генетической информацией между этими доменами, а также результаты филогенетического анализа редуцированных рибосом внутриядерных симбионтов парамеций, т.е. молекулярной машины, практически недоступной для изучения восходящей эволюционной траектории из-за отсутствия промежуточных форм.

По структуре, диссертационная работы Софьи Константиновны немного отличается от стандартного образца. Кроме Введения и Обзора литературы, она содержит четыре главы с изложением полученных результатов, каждая из которых включает описание материалов и методов, использованных для решения поставленных задач, а также разделы «Результаты» и «Обсуждение результатов», которые только во второй главе объединены. В конце диссертационной работы имеется общее «Заключение», четыре ключевых вывода и список процитированной литературы, содержащий 326 ссылок. Работа изложена на 137 страницах машинописного текста. Выносимые на защиту утверждения отображены 24 рисунками и 13 таблицами. Ещё 3 рисунка использованы для иллюстрации литературных данных.

В относительно объёмном Введении автор акцентирует внимание на актуальности изучения бактериальных сообществ с использованием метагеномного анализа, результаты которого позволяют приступить к реконструкции метаболических путей даже некультивируемых микроорганизмов; отмечает весомый вклад биоинформатики в понимание масштабности горизонтального переноса генов и обсуждает незаменимость геномного анализа при изучении симбиозов. В отдельных разделах, кроме фактических данных о структуре диссертационной работы и публикациях, обсуждаются степень разработанности диссертационной темы, сформулированы цель, задачи, методология и новизна работы, а также теоретическая и практическая значимость полученных результатов. Т.к. в качестве объектов были выбраны микроорганизмы и сообществ

необычных мест обитания, полученные результаты уникальны, а выносимые на защиту положения имеют большое научное значение.

Обзор литературы написан очень хорошо. Он в точности соответствует сути проделанной работы и информативен. Кроме общих сведений о разнообразии прокариот и их природных популяций, обзор включает 4 раздела, соответствующие результативным главам диссертации. Важно, что материалы, представленные в обзоре использованы для того, чтобы обосновать выбор объектов исследования и решаемых в работе задач.

Первая результативная глава посвящена таксономическому и функциональному анализу сообществ микробных топливных элементов. Исследуемыми биологическими образцами были метагеномы бактериальных популяций, использованных для очистки сточных вод в двух экспериментальных реакторах, и метагеномы биоплёнок, образовавшихся на анодах этих реакторов через 90 (Великобритания) и 70 (Япония) дней после их запуска. Поводом для такого исследования было отсутствие комплексных метагеномных данных для реакторов полупромышленных объёмов. Композиционный состав всех четырёх образцов сначала был определён с использованием классического типирования по 16S рРНК, после чего таксономический ранг был приписан каждой картируемой последовательности в библиотеках методом поиска наименьшего общего предка. В результате, был определён композиционный состав инокулятов и сообществ, колонизовавших аноды. Он отличался для обеих исследованных популяций, в которых, кроме бактерий были впервые идентифицированы археи. Достоинством этого раздела является попытка изучения метаболического потенциала исследованных сообществ, для чего была оценена представленность генов крупных функциональных категорий, а также генов, кодирующих ферменты утилизации ряда органических соединений, и, независимо, генов, способных обеспечить передачу электронов на электрод. Кроме ожидаемого присутствия микроорганизмов, необходимых для разложения органических соединений и микроорганизмов, способных обеспечить поток электронов на анод с участием растворимых переносчиков, в активном иле были найдены представители рода *Geobacter*, способные к прямому переносу электронов, а также метаногенные археи, что может быть использовано для производства биогаза. Кроме этого, в образце, полученном из биоплёнок, было обнаружено повышенное содержание генов, связанных с хемотаксисом, передачей сигнала и мембранным транспортом. Это любопытно как факт, но не понятно, означает ли это, что сформировавшие биоплёнку микроорганизмы имеют большее число генов этих категорий, или их избыточному, по сравнению с контролем, присутствию есть какое-то другое объяснение, ведь популяционный состав изолированной анодной камеры, судя по всему, в ходе эксперимента не менялся.

Замечания к разделу:

1. Осталось не понятным, из какой субстанции были получены референсные метагеномы: из инокулята, загруженного в анаэробную анодную камеру во время пуска биореакторов, или из активного ила, находящегося в этой камере в течении всего эксперимента, но не контактирующего с анодом. В обоих случаях, фраза: «Оба исходных сообщества претерпели значительные структурные изменения после 70-90

дней работы МТЭ» (стр. 38) не корректна, т.к. в экспериментах сравнивали микробиом ила с микробиомом плёнки.

2. Некоторое удивление вызывает очень низкий порог по длине для контигов (500 нп) при минимальном пороговом значении для прочтений - 400 н. Нет ли здесь ошибки?
3. Во второй таблице, также как на Рис. 2.3 показана представленность функции в биоплёнке и её отличие от инокулята. Согласно методическому разделу, она оценивалась как «сумма представленностей каждой ортологической группы, относящейся к данной категории». Т.е. как некая интегральная величина, не предоставляющая возможности для статистической оценки и нивелирующая разницу между ситуациями, когда вклад дают много генов нескольких таксонов и несколько генов многих таксонов. Может быть более информативным было бы показать эти данные с использованием box-plots, отражающих вариации внутри категории. Максимальная обнаруженная разница в 75% для метагеномных данных не очень убедительна.
4. В первом выводе, отражающем результаты этого раздела второе предложение «Изученные сообщества из двух МТЭ способны достаточно эффективно очищать сточные воды» не уместно, т.к. эффективность очистки сточных вод в данной работе не изучалась.

Следующая глава посвящена оценке процента горизонтально перенесённых генов в геномы метаносарцин. По литературным данным, их особенностью является повышенная частота обмена генетическим материалом с бактериями. В качестве моделей использовали две группы архейных геномов (геномы трёх штаммов *Methanosarcina* spp. или геномы шести штаммов семейства *Methanosarcinaceae*, включающие *Methanosarcina* spp). Для каждой из них были построены группы ортологичных белков (ГОб). При этом, из выборки ГОб для *Methanosarcinaceae* были удалены все ГОб, учтённые группами ортологичных белков *Methanosarcina*, т.е. две изученные выборки не перекрывались. Их использовали для поиска гомологов в забранной из NCBI базе данных белковых последовательностей. Для выравнивания отбирали 100 ближайших ортологов каждого члена ГОб, которые после удаления транспозонов были использованы для построения филогенетических деревьев методом ближайших соседей с бутстрэп поддержкой и методом наибольшего правдоподобия. Потенциально перенесёнными генами считали такие, которые формировали единую кладу внутри бактериальной ветви на деревьях, построенных обоими методами. В результате, было найдено 349 генов из 143 ГОб, которые могли быть перенесены из бактерий в общего предка трёх метаносарцин и гены еще 72 ГОб которые могли быть интегрированы в предка *Methanosarcinaceae*. Независимо были проанализированы белки, принадлежащие только одному штамму метаносарцин, перенесённые в геномы после дивергенции от общего предка.

В результате такого масштабного анализа в геномах метаносарцин было найдено только ~5% бактериальных генов, что существенно меньше ранее сделанных оценок. Поэтому было реконструировано состояние геномной базы NCBI в отношении архей на все промежуточные годы и поиск чужих генов был повторён в упрощённом варианте. Такая

историческая ретроспектива, с одной стороны, подтвердила ожидаемое уменьшение процента потенциально перенесённых генов с увеличением объёма доступных геномных данных, а с другой, показала перспективность использования процедуры экстраполяции оценочных параметров для таксонов с малой представленностью.

В рамках этого же раздела было выполнено ещё два блока исследований. В одном из них искали бактериальные гены в геномах пироккокков (семейство *Thermococcaceae*), т.е. архей с предположительно низким уровнем генетических интервенций, а в другом, оценивали процент генов, предположительно перенесённых из архей в геномы *Thermotogaceae*. В первом случае, была найдена только одна ГОБ архей, содержащая бактериальные гены. Не исключено, поэтому, что *Methanosarcinaceae* существенно отличаются от других архей по числу геномных интервенций. Во втором случае, было обнаружено 4.6% архейных генов в бактериальных геномах. Это соответствует проценту бактериальных генов в геномах метаносарцин, и может служить разумной оценкой вклада меж-доменного обмена генетической информацией в общее число горизонтально перенесённых генов, который для бактерий оценивается на уровне 20-30%. Тем не менее, строгость использованных критериев могла привести к потере каких-то чужих генов. Так, например, требовалось, чтобы все члены отобранной ГОБ формировали единую кладу внутри бактериальной ветви на деревьях, построенных двумя филогенетическими методами.

Особым достоинством этой части работы является идентификация потенциальных доноров ГПГ, т.к. работы, направленные на оценку процента чужих генов в масштабе всего генома, обычно ориентируются на поиск различного рода текстуальных отличий от среднего по геному, используя филогенетический анализ только как дополнительный или вспомогательный инструмент. По факту, самыми частыми донорами генов, перенесённых в геномы метаносарцин оказались фирмикуты и протеобактерии, и, хотя нормировка на число прочитанных геномов внесла свои корректировки, фирмикуты, населяющие общие с метаносарцинами экологические ниши, всё равно, остались в четвёрке лидирующих доноров. Несмотря на то, что горизонтальный перенос возможен для генов разной функциональной принадлежности, в работе показано, что половина генов, перенесённых из бактерий в *Methanosarcinaceae* ассоциированы с метаболизмом и транспортом различных соединений, причём перенос транспортёров оказался чуть более частым событием, чем перенос ферментов. Классификация перенесённых генов по категориям: гены домашнего хозяйства, ферменты, транспортёры и факторы транскрипции подтвердило преимущественный перенос генов транспортных белков. Оказалось, что чужие гены реже включены в опероны, чем собственные гены архей, но анализ доступных протеомных и экспрессионных данных привёл к заключению, что доля дифференциально экспрессирующихся генов среди них не ниже, чем в среднем по геному. Это значит, что перенесённые гены встроились в регуляторные сети реципиентов и было бы интересно узнать, отличаются ли по этому свойству гены, интегрированные в геномы исследованных метаносарцин после их дивергенции. В любом случае, вывод, сделанный по результатам всего второго блока, полностью и адекватно отражает суть полученных результатов.

Замечания к разделу:

1. Есть противоречие между фразой «После такой нормализации оказалось, что наиболее перепредставленные таксоны-доноры – это *Planctomycetes*, *Synergistetes* и *Firmicutes*» (Стр. 58) и Рис. 3.2, в соответствии с которым доминирующими являются *Planctomycetes*, *Acidobacteria* и *Synergistetes*.
2. В подписи к Таблице 3.5 есть смысловая ошибка: «в столбце “% НТ” показан процент горизонтально перенесённых генов в данном геноме» надо, по-видимому, «в столбце “% ГПГ” показан процент горизонтально перенесённых генов *указанной категории* в данном геноме».

Третья результативная глава диссертации посвящена сравнительному анализу геномов *Holospira* spp, т.е. внутриядерных симбионтов инфузорий, размножающихся в макронуклеусе, но способных выходить из него для заражения новых клеток. В качестве модельных было использовано 4 собранных в контиги/скаффолды генома холоспор, в том числе *Holospira curviuscula*, сборка и тщательная аннотация генома которой были основными задачами этого блока исследований. Попутно, авторами были получены данные, свидетельствующие против необходимости разделения *H. undulata* HU1 и *H. elegans* E1 на два разных вида. В качестве референсной группы были использованы геномы более чем 100 рикетсий, являющихся близкими родственниками холоспор.

В результате очень тщательного анализа, в геноме *H. curviuscula* было обнаружено 1594 гена, и для 683 из них была предсказана функция. Относительно большой размер генома вполне убедительно был объяснён наличием в нём большого числа протяжённых повторов, а последовательностей предположительно фагового происхождения у *H. curviuscula* оказалось меньше, чем в геномах других холоспор. Но самой удивительной находкой стал крайне редуцированный метаболический потенциал всех исследованных холоспор. Ни в одном из четырёх геномов, в частности, не были обнаружены гены, способные обеспечить весь цикл автономного биосинтеза аминокислот. Зато в геноме *H. curviuscula* было найдено 30 генов, кодирующих транспортные белки, в том числе транспортёры олигопептидов и аминокислот. По оценкам авторов, этого не достаточно, для того, чтобы обеспечить полноценную работу трансляционного аппарата без использования импортируемых олигопептидов. Ситуация с биосинтезом азотистых оснований оказалась ненамного лучше, но все известные геномы холоспор имеют рибонуклеотид-редуктазы, которые могут превращать рибонуклеотиды в дезоксирибонуклеотиды, что позволяет им использовать нуклеотиды в качестве источника энергии. Важным этапом аннотации был поиск генов, необходимые для выживания в ядре, для которого был использован сравнительный анализ генного состава всех известных холоспор с генным составом других *Rickettsiales*. В результате, было идентифицировано более сотни ГОБ, специфичных для холоспор, среди которых особое значение могут иметь гены, принимающие участие в образовании репродуктивной формы и развитии инфекции. Результаты, представленные в этой главе адекватно отражены в третьем выводе диссертационной работы.

Последний раздел диссертации посвящён эволюции рибосом у симбиотических бактерий. Он представляет особый интерес в связи с ограниченной возможностью изучения

процесса эволюции трансляционного аппарата. Объектами рецензируемой работы были рибосомные белки 214 симбиотических бактерий с редуцированными геномами. Автором впервые был идентифицирован набор рибосомных белков, часто отсутствующих в трансляционном аппарате, а также изучены укороченные рРНК. О характерной для автора тщательности компьютерного анализа на данном этапе свидетельствует пере-аннотация генов всех рибосомных белков во всех геномах. Было исследовано 57 белков, из которых одиннадцать были найдены во всех штаммах. Скорость эволюции остальных рибосомных белков оценивалась по средней скорости эволюции этого универсального набора. Были идентифицированы 10 белков, чаще других отсутствующих в исследуемых бактериях (9 принадлежат большой субъединице рибосомы) и оказалось, что все они были независимо потеряны несколько раз в разных таксонах. Были также идентифицированы геномы с максимальной потерей генов рибосомных белков, которые принадлежат бактериям с экстремально редуцированными геномами. Оказалось, что часто пропадающие белки формируют меньше контактов с другими белками и РНК. Тем не менее, в их число попали два белка (L34 и L36), глубоко погруженные в структуру рибосомы. В заключении к этой главе автор очень правильно обращает внимание на то, что отсутствие некоторых рибосомных белков может свидетельствовать не только об адаптивном дрейфе редуцированных геномов на минимизацию, но и об их более позднем приобретении в эволюции.

Замечания к разделу:

1. Т.к. для изучения трансляционного аппарата были выбраны бактерии с редуцированными геномами, в Таблице 5.1. не плохо было бы указать их размер.
2. На мой взгляд, очень важным результатом работы является обнаружение делеций в рибосомных РНК. Очень жаль, что они не указаны.
3. Результаты, представленные на Рис. 5.7а визуально свидетельствуют против наличия корреляции между длиной генома и утратой последовательности анти-Шайна – Дальгарно в 16S рРНК, а в тексте (Стр. 106) есть утверждение, что её исчезновение не связано с количеством потерянных р-белков. Однако приведённые значения коэффициентов корреляции (0,24 – 0,28) при $p = 9 \times 10^{-5} - 8 \times 10^{-4}$, формально, свидетельствуют о слабой корреляции, что требует менее жёсткой формулировки последнего предложения в четвёртом выводе.

В кратком итоговом Заключении автор суммирует основные результаты, полученные по всем четырём направлениям и даже предлагает перспективный путь практического совершенствования микробных топливных элементов с использованием протеобактерий рода *Geobacter*, способных к прямому переносу электронов. Вся работа является цельным, хорошо спланированным и законченным в рамках поставленных задач исследованием. Она очень хорошо оформлена и содержит мало опечаток (стр. 17, 25, 41, 42, 68, 79, 99), несогласованных предложений (стр. 5, 37, 56, 64, 77, 87, 103) и других технических погрешностей (на Рис. 5.1, например, отсутствующие белки S22 и Thx помечены красным, а не синим цветом). Это несколько не снижает прекрасного впечатления от самой работы, которая вызывает большой интерес и желание её в полной мере оценить. Достоверность

результатов никаких сомнений не вызывает, а выводы отражают суть полученных данных. Автореферат соответствует содержанию диссертации, а в 8 публикациях автора отражены практически все результаты. Актуальность, научная новизна и практическая значимость проведенных в работе исследований полностью соответствует п. 9 «Положения о присуждении ученой степени», утвержденного постановлением Правительства Российской Федерации от 29.09.2013 № 842, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор, Гарушянц Софья Константиновна, безусловно, заслуживает присуждения ученой степени кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика.

Доктор биологических наук, профессор, заведующий лабораторией функциональной геномики и клеточного стресса Института биофизики клетки - обособленного структурного подразделения Федерального государственного бюджетного учреждения науки Федерального исследовательского центра «Пущинский научный центр биологических исследований» Российской академии наук (ФИЦ ПНЦ БИ РАН)

Подпись сотрудника ФИЦ ПНЦ БИ РАН
Ольги Николаевны Озолинь удостоверяю:

Нач. отд. Награды
19 марта 2019 г.



Т. Н. Степанова

Контактные данные:

тел.: +7(4967)739-140, e-mail: ozoline@rambler.ru

Специальность, по которой официальным оппонентом защищена диссертация: 03.01.02 «Биофизика»

Адрес места работы:

142290, Московская область г. Пущино, ул. Институтская, д. 3,
ФИЦ ПНЦ БИ РАН, Институт биофизики клетки РАН,
Лаборатория функциональной геномики и клеточного стресса
Тел.: +7(4967)739-140; e-mail: ozoline@icb.psn.ru