

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА

БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

Раева Анастасия Сергеевна

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МИКРОБНЫХ СООБЩЕСТВ,
АССОЦИИРОВАННЫХ С МОРСКИМИ КОПЕПОДАМИ, И ПОИСК
ГЕНОВ ПОТЕНЦИАЛЬНЫХ ПУТЕЙ БИОСИНТЕЗА ЦЕЛЕНТЕРАЗИНА

Выпускная квалификационная работа магистра

Научный руководитель:

Кандидат биологических наук

Д. А. Сутормин

Москва – 2023

ОГЛАВЛЕНИЕ

1. ВВЕДЕНИЕ	4
2. ОБЗОР ЛИТЕРАТУРЫ	8
2.1. Микробиомы морских беспозвоночных как источник вторичных метаболитов	8
2.1.1. Прикладной потенциал вторичных метаболитов бактерий	8
2.1.2 Морские бактериальные сообщества как источник новых вторичных метаболитов с прикладной значимостью	11
2.1.3 Метагеномные исследования и поиск вторичных метаболитов	13
2.2. Целентеразин-зависимая биолюминесценция	15
2.2.1. Общие представления о целентеразин-зависимой биолюминесценции	15
2.2.2 Представления о биосинтезе целентеразина в живых организмах	20
2.2.3. Применение целентеразина и целентеразин-зависимых люцифераз в биотехнологии и биомедицине	22
2.3. Микробиомы морских копепод	24
2.3.1. Общая характеристика и экологическая роль класса копеподы в морских экосистемах	24
2.3.2. Бактериальные сообщества морских копепод	24
2.4. Описание объектов исследования	27
3. МАТЕРИАЛЫ И МЕТОДЫ	30
3.1. Сбор материала, идентификация и вскрытие объектов	30
3.2. Экстракция тотальной ДНК из образцов морских копепод и морской воды ...	31
3.3. Секвенирование ампликонных библиотек переменного участка V3-V4 гена 16S рРНК и шот-ган метагеномное секвенирование	34
3.4. Обработка данных секвенирования переменного участка V3-V4 гена 16S рРНК	35
3.5. Обработка данных шот-ган метагеномного секвенирования	36
4. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ	38
4.1. Выделение ДНК из образцов морских копепод и морской воды для дальнейшего секвенирования, выбор оптимального метода	38

4.2. Сравнительный анализ микробных сообществ на основе данных секвенирования ампликонов переменного участка V3-V4 гена <i>16S рРНК</i>	40
4.2.1. Анализ первичных данных и удаление контаминации	40
4.2.2. Относительные частоты таксонов микробиомов <i>Metridia longa</i> , <i>Calanus glacialis</i> , кишечника <i>Metridia longa</i> и морской воды	46
4.2.4. Анализ бета-разнообразия микробиомов	54
4.2.5. Дифференциальный анализ микробных сообществ.....	56
4.3. Анализ микробиома на основе данных шот-ган метагеномного секвенирования.....	59
4.3.2. Кластеры генов биосинтеза вторичных метаболитов, обнаруженные в образце <i>Metridia longa</i> и в образце морской воды	60
4.3.3. Кандидатные гены биосинтеза целентеразина на основе данных предсказания открытых рамок считывания	62
5. ЗАКЛЮЧЕНИЕ	67
6. ВЫВОДЫ	68
7. СПИСОК ЛИТЕРАТУРЫ	70

1. ВВЕДЕНИЕ

Морские микробные сообщества сравнительно мало изучены и являются богатым источником новых биологически-активных веществ с потенциалом использования в медицине и биотехнологии. Одной из групп таких веществ являются вторичные метаболиты бактерий, среди которых можно обнаружить, например, новые антибиотики и соединения с противоопухолевыми активностями. Исключительный интерес для изучения представляют микробные сообщества, ассоциированные с морскими организмами, что связано с высокой плотностью таких сообществ.

Морские копеподы (или веслоногие ракообразные) - консументы I и II порядков - являются наиболее массовыми представителями зоопланктона и составляют до 80% его биомассы. Несмотря на этот факт, микробные сообщества копепод относительно мало изучены. Кроме того, для нескольких родов веслоногих ракообразных зарегистрирована способность к биолюминесценции. Обнаруженные у них биолюминесцентные системы используют в качестве субстрата-люциферина целентеразин, также широко встречающихся у других таксономически далёких групп морских беспозвоночных. Существующие данные свидетельствуют в пользу того, что целентеразин синтезируется в представителях рода *Metridia de novo*, однако механизм биосинтеза этого соединения остаётся всё ещё не раскрытым. Неизвестно также, синтезируется ли целентеразин непосредственно ракообразным или его бактериями-симбионтами. На текущий момент сформулированы две альтернативные гипотезы биосинтеза целентеразина. Первая заключается в том, что люциферин является производным короткого рибосомально синтезируемого трипептида FYY, который расположен на С-конце более длинного пептида-прекурсора и получается путем отщепления, циклизации и модификации. Другая гипотеза предполагает, что целентеразин синтезируется нерибосомально путем соединения свободных аминокислот в трипептид FYY по механизму нерибосомного синтеза пептидов, одного из основных механизмов биосинтеза вторичных метаболитов у бактерий. Расшифровка механизма биосинтеза целентеразина в будущем может позволить расширить применимость данного люциферина в биотехнологии и биомедицине, в частности, позволит разрабатывать новые методы для биоимиджинговых исследований.

Целями данной работы являются описание микробных сообществ массовых копепод Белого моря и поиск потенциальных путей биосинтеза целентеразина в микробиоме *Metridia longa* при помощи методов метагеномики.

Для достижения целей данной работы был проведен сбор живых представителей симпатрических веслоногих ракообразных *Metridia longa* и *Calanus glacialis* на ББС МГУ. Для метагеномного анализа была выделена ДНК из цельных ракообразных, кишечника *M. longa* и фильтратов морской воды с места сбора объектов. Из экстрактов ДНК были подготовлены и секвенированы ампликонные библиотеки переменного участка V3-V4 гена *16S rRNA* и шот-ган метагеномные библиотеки тотальной ДНК. По результатам работы были получены 40 экстрактов тотальной ДНК (18 образцов *M. longa*, 3 образца *C. glacialis*, 10 морской воды, а также 9 образцов негативных контролей вскрытия и выделения ДНК), которые в дальнейшем были использованы для профилирования микробного сообщества копепод и морской воды по данным секвенирования ампликонов переменного участка V3-V4 гена *16S rRNA*. Два образца, соответствующие целым особям *M. longa* и фильтрату морской воды, были дополнительно секвенированы по технологии шот-ган для последующего поиска кластеров генов биосинтеза.

Профилирование микробных сообществ методами *16S rRNA* метагеномики показало, что сообщества, ассоциированные с копеподами, значительно отличаются по составу и разнообразию от сообществ морской воды. Представители Cyanobacteria доминировали в морской воде, тогда как Proteobacteria вносили наибольший вклад в сообщества копепод. Альфа-разнообразие микробиомов морской воды было значительно выше разнообразия сообществ, ассоциированных с веслоногими ракообразными. Интересно, что микробиомы *M. longa* и *C. glacialis* практически не различались, что свидетельствует о решающей роли эко-географических факторов в формировании данных микробиомов, в отличие от различий в диете данных видов. Были отмечены достоверные различия между эпи- и эндомикрофлорой *M. longa*. *Cyanobiaceae* и *Vibrionaceae* были обогащены в образцах цельных ракообразных, что, вероятно, отражает состав эпимикробиоты *M. longa*. В образцах кишечника были обогащены *Shewanellaceae* и SAR86, вероятно, являющиеся симбиотическими представителями кишечного микробиома.

Шот-ган секвенирование и сборка метагеномов *M. longa* и морской воды позволило провести аннотацию и поиск кластеров генов биосинтеза, который не выявил потенциальных путей биосинтеза целентеразина. Поиск открытых рамок считывания в метагеноме *M. longa*, содержащих трипептид FYY на С-конце, выявил 32 кандидата. Поиск гомологов данных рамок считывания, однако, не подтверждает существование

консервативного С-концевого мотива и свидетельствует о внесении преждевременных стоп-кодонов за счет мутаций. Таким образом, гипотеза о синтезе целентеразина микробиотой, ассоциированной с *M. longa*, на основании полученных данных не подтверждается.

Личный вклад автора заключается в поиске и анализе литературы по тематике работы, участии в сборе объектов исследования, выделении экстрактов ДНК, анализировании данных секвенирования ампликонов варибельного участка V3-V4 гена *16S pPHK* и данных шот-ган метагеномного секвенирования, а также написании текста выпускной квалификационной работы. Все представленные данные получены лично и совместно с научным руководителем.

Цель работы

Целью данной работы является сравнение микробиомов морских копеод *Metridia longa*, *Calanus glacialis* и морской воды на основе данных секвенирования ампликонов варибельного участка V3-V4 гена *16S pPHK* и данных шот-ган метагеномного секвенирования, предположение кандидатных генов биосинтеза целентеразина на основе данных шот-ган метагеномного секвенирования.

Задачи работы

1. Сбор материала в Кандалакшском заливе Белого моря;
2. Выделение ДНК из образцов копеод *Metridia longa*, *Calanus glacialis* и морской воды, секвенирование ампликонных библиотек варибельного участка V3-V4 гена *16S pPHK*, шот-ган метагеномное секвенирование тотальной ДНК образца морской воды и образца *Metridia longa*;
3. Сравнение состава микробиомов копеод *Metridia longa*, *Calanus glacialis* и морской воды на основании данных секвенирования ампликонных библиотек варибельного участка V3-V4 гена *16S pPHK*;
4. Поиск кластеров генов биосинтеза, потенциально участвующих в биосинтезе целентеразина, на основании данных шот-ган метагеномного секвенирования тотальной ДНК образца морской воды и образца *Metridia longa*;
5. Поиск открытых рамок считывания, содержащих мотив FYY на С-конце и являющихся потенциальными прекурсорами целентеразина, на основании

данных шот-ган метагеномного секвенирования тотальной ДНК образца морской воды и образца *Metridia longa*.

2. ОБЗОР ЛИТЕРАТУРЫ

2.1. Микробиомы морских беспозвоночных как источник вторичных метаболитов

2.1.1. Прикладной потенциал вторичных метаболитов бактерий

Бактерии производят множество метаболитов в результате катаболических и анаболических процессов (Doelle, 2014). Метаболиты можно разделить на две крупные группы, в зависимости от их функциональных свойств и механизма биосинтеза — первичные и вторичные метаболиты. Первичные метаболиты весьма консервативны среди микроорганизмов различных таксономических групп и являются основным источником энергии для базовых биохимических и физиологических функций живых клеток, таких как рост, деление и развитие (Gokulan et al., 2014). В качестве примеров первичных метаболитов можно привести такие соединения, как нуклеотиды, витамины, аминокислоты, интермедиаты основных метаболитических путей. Эта группа метаболитов может быть обнаружена в качестве промежуточных и конечных продуктов метаболизма, а также в качестве конечных продуктов в биосинтезе макромолекул и коферментов. Первичные метаболиты синтезируются, главным образом, в лаг-фазе роста клеточной популяции. Вторичные метаболиты, напротив, не являются жизненно необходимыми соединениями, имеют более сложное химическое строение и выполняют специализированные функции. Часто они необходимы микроорганизмам для развития стратегий выживания при неблагоприятных условиях. К примеру, вторичные метаболиты выступают в роли антибактериальных агентов в условиях конкуренции за пищевые и пространственные ресурсы, обеспечивают межклеточную коммуникацию и защиту клетки-хозяина (Strachan, Davies, 2016; van der Meij et al., 2017). Также они вовлечены в синтез других метаболитических соединений. Вторичные метаболиты образуются, главным образом, в конце экспоненциальной и в стационарной фазе роста бактериальной популяции, и их продукция ассоциирована со стрессовыми условиями (Gokulan et al., 2014).

Вторичные метаболиты бактерий являются обширным источником для поиска важных с точки зрения медицины и биотехнологии веществ (Fouillaud et al., 2022). Среди них обнаружены антибиотики, противогрибковые, противопаразитарные, противоопухолевые вещества, иммунодепрессанты, пигменты, токсины, гормоны, пестициды, ингибиторы ферментов, сидерофоры и т. д. (Arun et al., 2021). Пожалуй, самым известным антибиотиком, являющимся бактериальным вторичным метаболитом, можно назвать стрептомицин. Этот аминогликозидный антибиотик исторически

является вторым обнаруженным антибиотиком после пенициллина и первым обнаруженным антибиотиком бактериального происхождения. Стрептомицин впервые был выделен в лаборатории Зельмана Ваксмана из актиномицета *Streptomyces griseus* в 1943 году и до сих пор успешно применяется в схемах лечения туберкулёза (Arun et al., 2021). Зельман Ваксман также закрепил термин “антибиотик”, указывая на установленную роль вторичных метаболитов как элементов стратегии конкурирования микроорганизмов. Однако многочисленные исследования на протяжении десятилетий после его открытий продемонстрировали, что потенциал этих веществ гораздо шире (Clardy et al., 2009). Актиномицеты, в целом, считаются основными продуцентами антибиотиков — около 60% известных антибиотиков были выделены из бактерий рода *Streptomyces* и около 15% из других родов актиномицетов (Basavaraj et al., 2010). Из *Streptomyces nodusus* выделен и описан Амфотерицин В — полиеновый макроциклический антибиотик с противогрибковой активностью. Применяется при лечении кандидоза внутренних органов и криптококкового менингита, а также является препаратом второго ряда при висцеральном лейшманиозе (Kafetzis et al., 2005). Авермектины — группа макролидных антибиотиков с 16-членным макроциклическим лактонным кольцом, являются вторичными метаболитами *Streptomyces avermitilis*. Эти вещества имеют нейротоксичные свойства по отношению к беспозвоночным, что делает их эффективными инсектицидами, акарицидами и нематицидами (Pitterna et al., 2009). Среди вторичных метаболитов актиномицетов известны и противоопухолевые вещества, успешно применяемые на практике, например, блеомицин, продуцируемый *Streptomyces verticillus*. Это группа гликопептидных антибиотиков, проявляющая цитостатическое действие. Препараты, имеющие блеомицин в качестве действующего вещества, применяются в лечении лимфом, рака яичек, рака яичников и рака шейки матки (Gothelf et al., 2003). Вещества с иммунодепрессантной активностью, применяемые в трансплантологии, сиролимус (рапамицин) и такролимус также выделены из *Streptomyces* (Perez-Simon et al., 2013).

Синтез вторичных метаболитов является сложным процессом, который включает в себя множество этапов и требует участия специфических ферментов или мультиферментных комплексов (Sánchez-Guijo et al., 2014). Данные компоненты являются промежуточными или конечными продуктами внутриклеточного метаболизма. Наиболее распространенными классами вторичных метаболитов являются нерибосомные пептиды (НРП) и поликетиды (ПК), многие важные с медицинской точки

зрения соединения, полученные из микроорганизмов, относятся к этим двум классам (Andryukov et al., 2019). У бактерий, в связи с общей компактизацией генома гены, ответственные за биосинтез вторичных метаболитов, обычно сгруппированы вместе в виде генных кластеров с общей регуляцией. Ниже мы остановимся на кратком описании только механизма нерибосомного синтеза пептидов, поскольку в раннем исследовании для целентеразина был предположен именно такой путь биосинтеза (Oba et al., 2009).

НРП синтезируются крупными мультимодульными ферментами — нерибосомными пептид синтетазами (НРПС) (Strieker et al., 2010). В нерибосомном пептидном синтезе, в отличие от классической трансляции с мРНК с участием рибосомы, используется не РНК матрица, а шаблон в виде структуры модулей НРПС, каждый из которых ответственен за присоединение очередной аминокислоты (Рисунок 1). Каждый модуль разделён на несколько доменов, несущих определённую функцию. Домены НРПС можно разделить на три основных типа: 1) домены аденилирования, 2) домены тиолирования и 3) домены конденсации. Синтез продукта начинается с активирования аминокислот доменом аденилирования (домен А) путем присоединения аденозинмонофосфатной (АМФ) группы к карбоксильной части аминокислоты. Данные домены, в том числе, имеют специфичность к определённому типу аминокислоты и, соответственно, определяют какой структурный блок будет присоединяться к полипептиду. Затем активированная аминокислота переносится в домен тиолирования (также можно встретить название пептидил переносящий белок, домен РСР), отвечающий за перенос аминокислот из доменов аденилирования в домены конденсации (домен С). Далее домен конденсации катализирует образование пептидных связей между растущим пептидом и присоединяемой аминокислотой. Далее пептид переносится в следующий модуль НРПС и цикл повторяется. В дополнение к основным доменам можно обнаружить дополнительные, отвечающие за модификацию аминокислот или конечной пептидной цепи, например, домены эпимеризации, домены метилирования и домены окисления. Терминальный модуль НРПС содержит домен тиюэстеразы (домен ТЕ), катализирующий высвобождение пептида путем гидролиза или макроциклизации. Примечательно, что помимо классических 20 протеиногенных аминокислот, НРПС для синтеза конечного продукта используют непротеиногенные аминокислоты, в том числе D-аминокислоты, а также мономеры жирных кислот и α -гидроксикислоты. Конечные НРП зачастую имеют сложную циклическую и разветвлённую структуру.

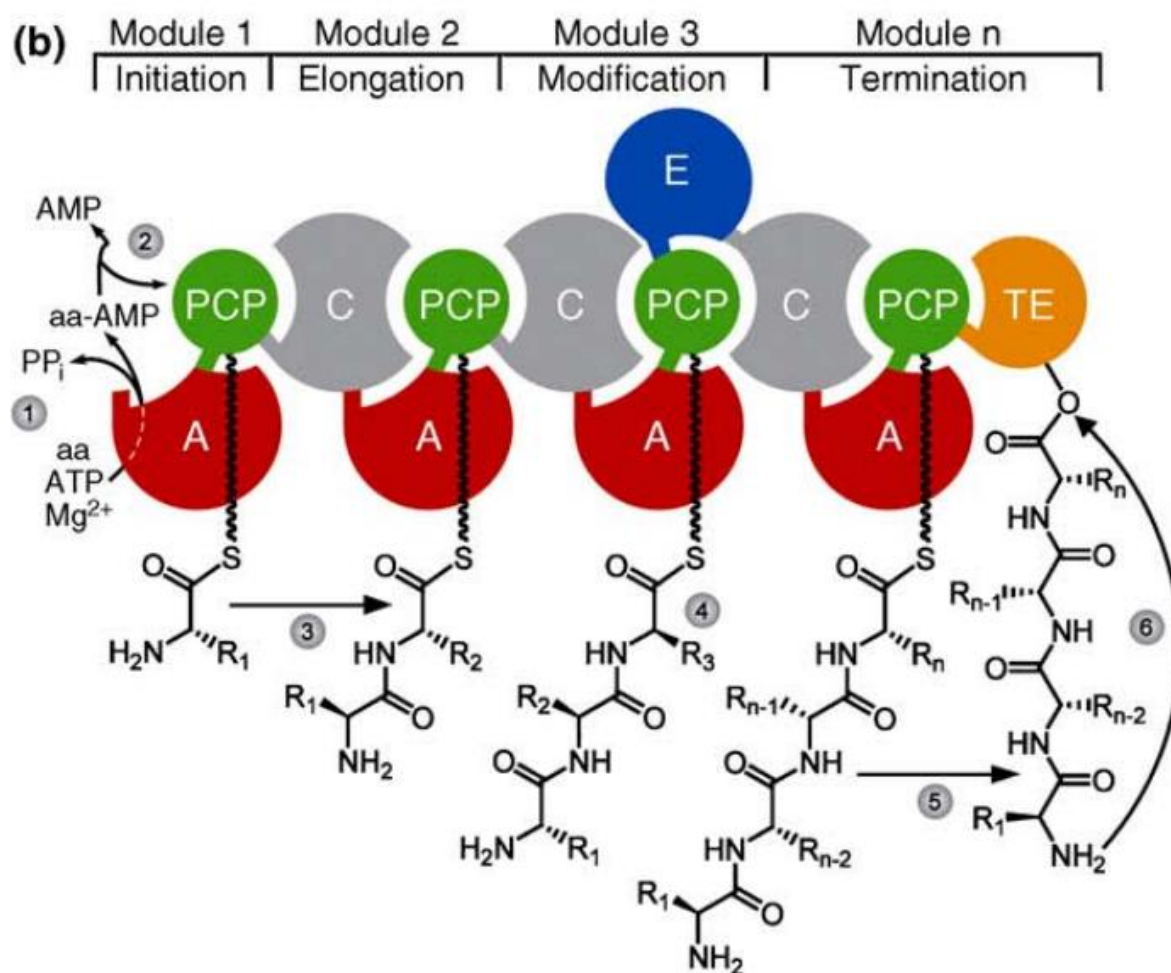


Рисунок 1. Схема синтеза нерибосомных пептидов (НРП) и организация модулей нерибосомной пептид синтетазы (НРПС). (1) Аминокислота активируется как аминоктил-АМФ доменом аденилирования (А). (2) Перенос аминокислоты на домен тиолирования (РСР). (3) Конденсация аминокислот, связанных с РСР. (4) Модификация аминокислот, например, доменами эпимеризации (Е). (5) Переэтерификация пептидной цепи с концевой РСР на домен тиюэстеразы (ТЕ). (6) Катализируемое тиюэстеразой высвобождение продукта путем гидролиза или макроциклизации (Strieker et al., 2010).

2.1.2 Морские бактериальные сообщества как источник новых вторичных метаболитов с прикладной значимостью

Большинство известных на сегодняшний день вторичных метаболитов с антибактериальной активностью обнаружены в почвенных микроорганизмах, однако в последнее время обнаружение новых биологически активных веществ (в частности, антибиотиков) из этой экологической группы микроорганизмов сильно замедлилось, что, вероятно, говорит об исчерпании данного ресурса. На сегодняшний день морские

микробные сообщества, как гораздо менее изученные, представляют перспективную альтернативу почвенным (Srinivasan et al., 2021). Океаны занимают около 70% поверхности Земли, и данные местообитания характеризуются большим разнообразием экологических ниш из-за таких параметров, как pH воды, солёность, температурные градиенты, режимы питания водоёмов, глубина и прочее (Liu et al., 2019). Кроме того, морские микроорганизмы демонстрируют большее филогенетическое разнообразие. Морские микроорганизмы синтезируют порядка 23000 различных вторичных метаболитов (Blunt et al., 2011). Шесть типов морских бактерий (Actinobacteria, Bacteroidetes, Cyanobacteria, Firmicutes, Planctomycetes и Proteobacteria), по-видимому, являются доминантными продуцентами антибактериальных соединений в морских экосистемах (Srinivasan et al., 2021). Особый интерес в поиске биологически активных веществ представляют ассоциации, формируемые между морскими микроорганизмами и морскими животными, поскольку численность микроорганизмов в таких сообществах обычно гораздо выше в сравнении со свободно живущими (Nealson, Hastings, 1979). Этот факт увеличивает вероятность обнаружения новых потенциально полезных штаммов. Охарактеризованы различные поликетидные вторичные метаболиты с прикладным потенциалом, выделенные из морских беспозвоночных, таких как губки, оболочники и мшанки (Blunt et al., 2011). Данные организмы ведут прикреплённый образ жизни и питаются, главным образом, за счет фильтрации, что делает их богатым источником микроорганизмов и их вторичных метаболитов (Haygood et al, 1999). Микроорганизмы могут в изобилии населять их пищеварительные тракты случайным образом из-за характера питания, а также формировать симбиотические взаимоотношения. Примечательным примером вторичного метаболита с медицинской значимостью являются бриостатин (Рисунок 2).

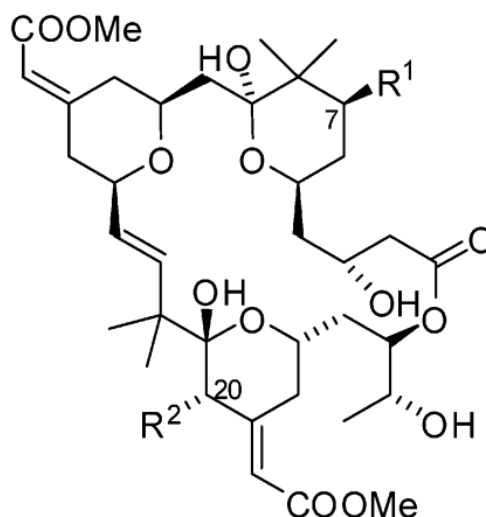


Рисунок 2. Структурная формула бриостатинов (Sudek et al., 2007).

Данное вещество относится к макролидам и было выделено из морской мшанки *Bugula neritina*, однако синтезируется некультивируемым симбионтом мшанок *Candidatus Endobugula sertula* (Sudek et al., 2007). Экологическая роль бриостатина заключается, предположительно, в защите личинок-хозяинов от хищников (Sudek et al., 2007). Продемонстрировано, что бриостатин является модулятором протеинкиназы C (Kollár et al., 2014). Установленная активность сподвигла провести клинические испытания в качестве противоракового препарата, однако они не увенчались успехом ввиду токсичности. Бриостатин также показал потенциал в лечении болезни Альцгеймера и изучался в качестве анти-ВИЧ агента (Kollár et al., 2014). В качестве примера антибиотика, обнаруженного у морской бактерии, можно привести солонамиды — группу циклодепсипептидов, обнаруженную у штамма *Photobacterium galathea* S2753, изолированного из зелёной мидии в Соломоновом море (Zhang et al., 2022). Солонамиды являются примером НРП, продемонстрирована их способность снижать вирулентность золотистого стафилококка (Mansson et al., 2011).

2.1.3 Метагеномные исследования и поиск вторичных метаболитов

Исследование природных микробных сообществ на предмет вторичных метаболитов является основной стратегией в поиске новых биологически активных веществ. Большинство коммерчески используемых антибиотиков являются естественными вторичными метаболитами микроорганизмов, полусинтетическими

производными из них, либо химически синтезированы на основе структуры естественных вторичных метаболитов (Katz et al, 2016). Классические методы поиска новых значимых с прикладной точки зрения вторичных метаболитов долгое время были основаны на культуральных подходах и скрининге антибактериальной активности. Однако такие подходы существенно ограничивают поиск в природных образцах, поскольку подавляющее большинство микроорганизмов является не культивируемыми в виду сложности или даже невозможности обеспечить оптимальные условия для роста колоний в лаборатории. Развитие методов секвенирования нового поколения совершило революцию в предсказании и обнаружении новых вторичных метаболитов, поскольку обеспечило возможность исследовать скрытое биоразнообразие микроорганизмов и его потенциал в продукции новых антибиотиков и прочих биологически активных веществ. В частности, метагеномное секвенирование ДНК окружающей среды или ДНК-экстрактов из животных позволяет получить совокупность геномов микроорганизмов, населяющих интересующий образец, и предсказывать кластеры генов биосинтеза в полученных метагеномах (Zoccarato et al., 2022). Впоследствии, предсказанные кластеры генов биосинтеза возможно клонировать в удобно культивируемые лабораторные штаммы, что позволяет проводить наработку целевого вторичного метаболита (Arun et al., 2021). Подобным образом, например, был идентифицирован кластер биосинтеза сидерофора вибриоферрина (Рисунок 3) из приливно-отливного осадка в заливе Ариаке. Впоследствии обнаруженный кластер был успешно клонирован и экспрессирован в лабораторной культуре *Escherichia coli* (Fujita et al., 2010)

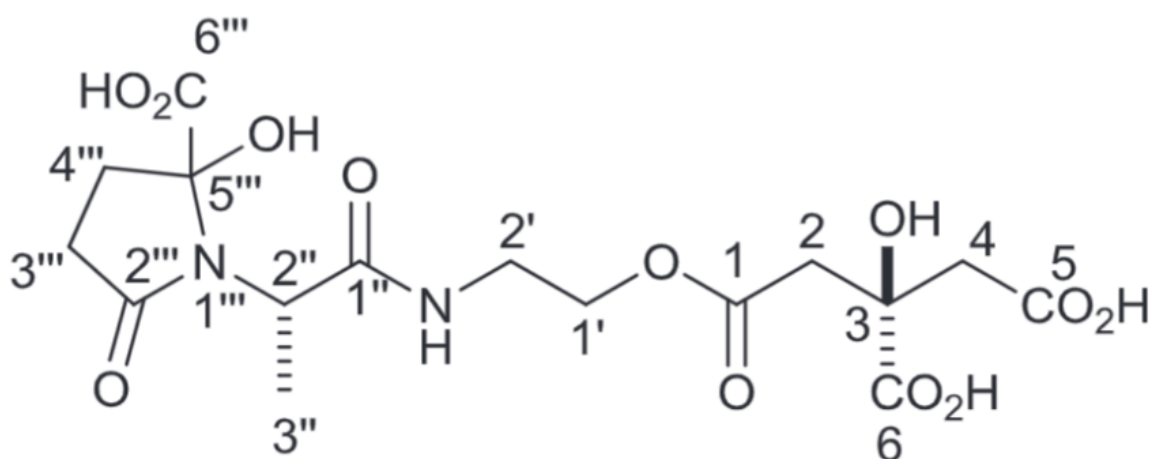


Рисунок 3. Структурная формула вибриоферрина 1 (Fujita et al., 2010).

2.2. Целентеразин-зависимая биолюминесценция

2.2.1. Общие представления о целентеразин-зависимой биолюминесценции

Биолюминесценция — это способность биологических объектов генерировать свечение в результате опосредованных энзимами химических реакций, является частным случаем хемилюминесценции (Shimomura, Yampolsky, 2019). В основе феномена лежит окисление специфического субстрата, называемого люциферин, в присутствии катализирующего фермента — люциферазы. В результате реакции образуется окисленный продукт, оксилуциферин, в возбуждённом электронном состоянии, переход из которого в основное сопровождается эмиссией фотона в видимом спектральном диапазоне (Shimomura, Yampolsky, 2019). На сегодняшний день известно около 30 биолюминесцентных систем и только для 7 из них выделены и охарактеризованы химические структуры люциферин (Tsarkova et al., 2016). В общем случае, химическую реакцию биолюминесценции можно разделить на две стадии:

1) образование возбуждённого продукта C^* из исходных реагентов (А и Б):



2) и переход возбуждённой молекулы C^* в основное состояние с испусканием кванта света:



Биолюминесцентное свечение может производить как сам наблюдаемый биологический объект, так и его бактериальные симбионты (Haddock, 2010). Распределение биолюминесцентных видов по таксономическим группам демонстрирует широкое распространение этого феномена в природе. Почти в каждом крупном таксоне можно обнаружить виды, способные к свечению (Shimomura, Yampolsky, 2019). Таковые известны среди бактерий, одноклеточных эукариот, макромицетов и животных, исключение составляют только растения и наземные позвоночные. Однако наблюдается корреляция от места обитания. Так биолюминесценция встречается гораздо чаще у морских планктонных организмов и наиболее широко распространена среди представителей типа Стрекающие (Cnidaria) и типа Гребневики (Stenophora).

Целентеразин — самый распространённый и изученный люциферин, обнаруженный в морских биолюминесцентных организмах (Hastings, Johnson, 2003). Среди них представители радиолярий, щетинкочелюстных, иглокожих, стрекающих, гребневиков, головоногих моллюсков, рыб и ракообразных, в частности, копепод (Campbell, Herring, 1990; Haddock, Case, 1999; Kaskova et al., 2016). Впервые химическая

структура целентеразина была описана в 1974 году (Shimomura, Yampolsky, 2019). Изоляты вещества были получены из гидромедузы *Aequorea victoria* (Shimomura, Yampolsky, 2019), из которой также был выделен зелёный флуоресцентный белок (GFP), за что Осаму Шимомура был удостоен Нобелевской премии по химии в 2008 году. Позже, в 1975 году, целентеразин был впервые химически синтезирован (Inoue et al., 1975). В основе структуры целентеразина лежит имидазопиразиноновое ядро — комбинация 5- и 6-членных азотсодержащих колец (Рисунок 4). Целентеразин-зависимая биолюминесценция инициируется связыванием молекулы кислорода в положении С-2 имидазопиразинонового ядра с быстрым образованием пероксида. Затем пероксид образует четырехчленное диоксетановое кольцо. При распаде диоксетанона образуется амидный аниона целентерамида в возбужденном состоянии и углекислый газ. Эмиссия кванта света является следствием релаксации амидного аниона целентерамида на высоком энергетическом уровне в основное состояние (Jiang et al., 2016).

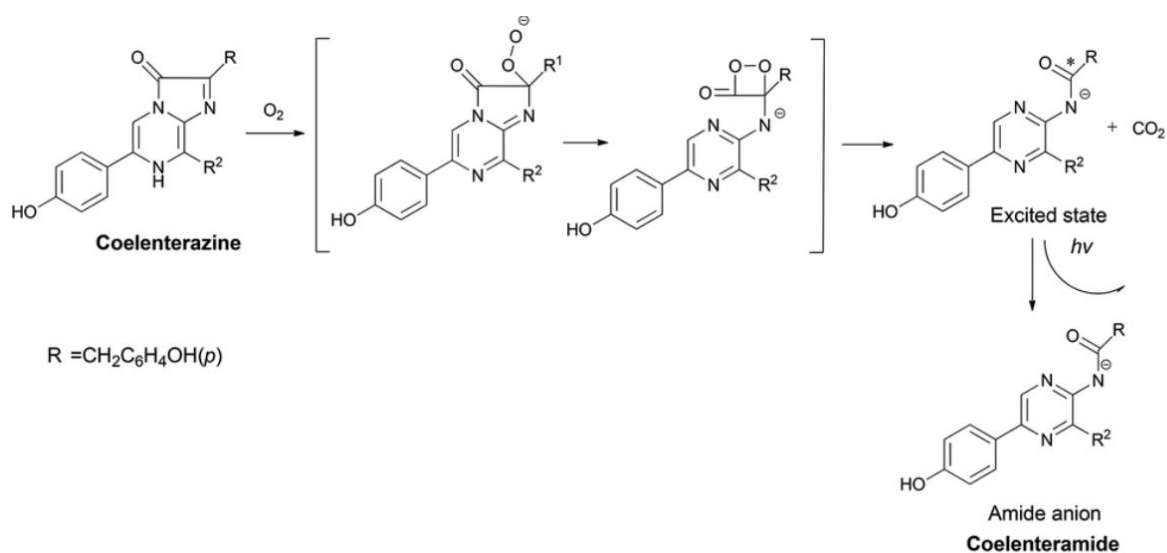


Рисунок 4. Схема механизма химической реакции биолюминесценции целентеразина (Jiang et al., 2016).

Описанные на данный момент биолюминесцентные системы, содержащие в качестве субстрата целентеразин, могут быть разделены на два различных с биохимической точки зрения класса — 1) люциферин-люциферазные системы (Рисунок 5а и 5с) Ca^{2+} -регулируемые фотопротеины (Рисунок 5b) (Krasitskaya et al., 2020).

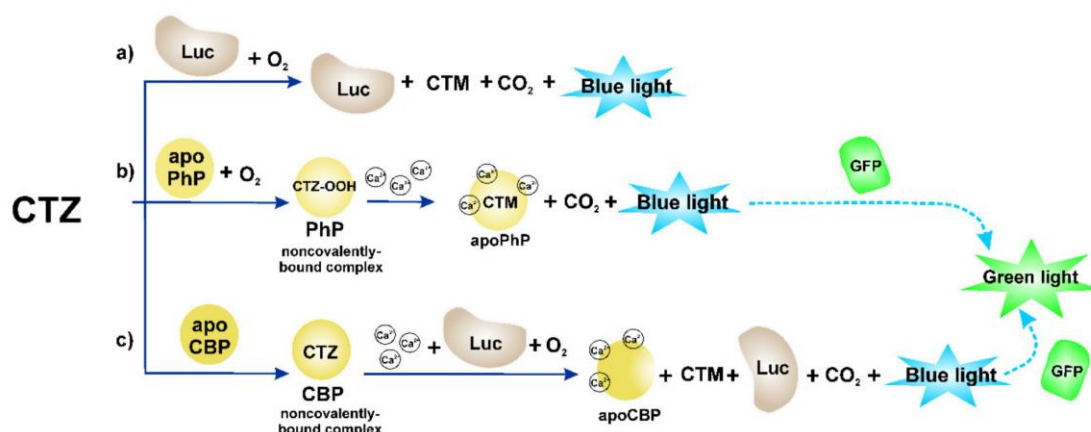


Рисунок 5. Разнообразие целентеразин-зависимых биOLUMИнесцентных систем. Вариант “a” и “c” в тексте описан как целентеразин-зависимая система первого класса, вариант “b” и как система второго класса. Luc – люцифераза; apoPhP – апофотопротейн; apoCBP - Ca^{2+} -зависимый целентеразин-связывающий белок, CTZ – целентеразин и CTM - целентерамид (Krasitskaya et al., 2020).

В системах первого класса окисление целентеразина кислородом катализируется кофактор-независимой люциферазой, относящейся к группе монооксигеназ. Конечными продуктами работы системы являются целентерамид (декарбоксилированная форма целентеразина), углекислый газ и эмиссия кванта света с пиком около 480 нм в голубом спектре видимого излучения (Рисунок 4). Такие варианты описаны, например, у мягких кораллов *Renilla* (Matthews, Hori, Cormier, 1977; Titushin et al., 2008) и копепод *Gaussia* (Inouye, Sahara, 2008) и *Metridia* (Markova et al., 2004; Markova et al., 2012). Люциферазы *Gaussia* (GLuc) и *Metridia* (MLuc) довольно хорошо изучены, соответствующие локусы кДНК успешно клонированы, получены рекомбинантные аналоги, а также получены модифицированные варианты с улучшенными свойствами (Tannous, 2009; Verhaegen, Christopoulos, 2002; Markova et al., 2004; Markova et al., 2012). БиOLUMИнесценция копепод происходит экскреторно. Секрет воспроизводится в специализированных эпидермальных железах и экстрекретируется наружу через кутикулярные поры. Железы, как правило, состоят из 1-3 вытянутых клеток, наполненных секреторными гранулами, и распределены по поверхности тела ракообразного (Рисунки 6 и 7).



Рисунок 6. Паттерн расположения билюминесцентных желёз у *Metridia okhotensis*. Фотография снята с помощью стимуляции флуоресценции желёз источником УФ-излучения (Takenaka et al., 2017).

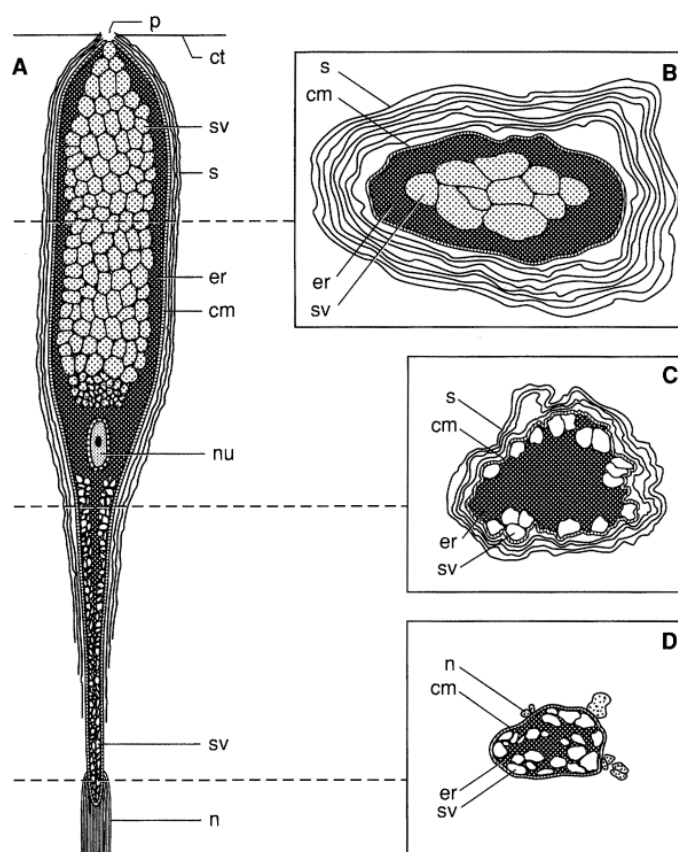


Рисунок 7. Схематическое изображение клетки билюминесцентной железы копеподы *Gaussia princeps*. (A) Продольный разрез клетки. (B) Поперечный срез секреторной гранулы пузырька, заключённого в клетку. (C) и (D) поперечные срез участка клетки, связанного с нервом. cm - клеточная мембрана, ct - кутикула; er - эндоплазматический ретикулум; n - нерв; nu - ядро; p - пора; s - оболочка клетки; sv - секреторные пузырьки (Browlby, Case, 1991)

Свечение в голубом спектре происходит при экскреции содержимого желёз во внешнюю среду (Евстигнеев, 1990). На сегодняшний момент неизвестно, разделены ли люцифераза и целентеразин пространственно в разных секреторных гранулах желёз

копепод, и в какой момент запускается окисление целентеразина. Было продемонстрировано, что биолюминесцентные железы иннервируются, электрическая стимуляция брюшной нервной цепочки вызывала экскрецию секрета у препаратов абдомена копепод (Bowlby, Case, 1991). В отличие от копепод *Gaussia* и *Metridia*, биолюминесценция мягкого коралла *Renilla* происходит внутриклеточно и биолюминесцентная система устроена несколько сложнее, однако также контролируется нервной системой (Hastings, Morin, 1969). Целентеразин у данного объекта не ковалентно упакован в полость Ca^{2+} -зависимого целентеразин-связывающего белка (apoCBP). Окисление целентеразина в таком комплексе катализируется *Renilla*-люциферазой (RLuc) при связывании ионов Ca^{2+} с CBP (Anderson et al., 1974). Эмиссия света в результате биолюминесценции затем стимулирует флуоресценцию зеленого флуоресцентного белка (GFP). Реакция запускается вследствие повышения внутриклеточной концентрации ионов кальция в результате механической стимуляции.

Второй класс целентеразиновых биолюминесцентных систем включает устойчивые фермент-субстратные комплексы, которым не требуется наличие молекул кислорода для реакции биолюминесценции, а эмиссия света происходит при взаимодействии белка с ионами металлов (Shimomura, 2005). Примером подобной системы служит Ca^{2+} -регулируемые фотопротеины — экворин, выделенный из медузы *Aequorea victoria* (Shimomura et al., 1962), и обелин, обнаруженный в гидроидном полипе *Obelia* (Campbell, 1974). Фотопротеины состоят из апофотопротеина (apoPHP) и простетической группы — 2-гидропероксицелентеразина (Shimomura, Johnson, 1978). Эмиссия света осуществляется за счёт конформационных изменений субстрат-связывающей полости белка вследствие связывания с ионами Ca^{2+} . Это нарушает систему водородных связей, стабилизирующих 2-гидропероксицелентеразина и инициирует окислительное декарбоксилирование 2-гидропероксицелентеразина с выделением голубого кванта света. Эмиссия света в результате биолюминесценции с максимумом эмиссии в 465 нм затем стимулирует флуоресценцию зеленого флуоресцентного белка (GFP).

Как уже было сказано ранее, целентеразин-зависимые биолюминесцентные системы обнаруживаются у разнообразных групп организмов, таксономически далёких друг от друга. По-видимому, целентеразин стал синтезироваться ещё древними морскими билатериями. Более того, люминесцентные системы демонстрируют разнообразие функциональных элементов и механизмов, а клонированные на текущий

момент целентеразин-зависимые люциферазы различных организмов демонстрируют низкую степень гомологии (Markova, Vysotski, 2015). Вероятно, способность использовать целентеразин для биолюминесценции было приобретено независимо в нескольких группах живых организмов.

2.2.2 Представления о биосинтезе целентеразина в живых организмах

Несмотря на то, что целентеразин является субстратом биолюминесцентных систем у широкого разнообразия морских организмов, только для четырёх из них был продемонстрирован его *de novo* синтез, в то время как большинство организмов, видимо, получают целентеразин по пищевой цепочке (Haddock et al., 2001). В частности, передача целентеразина по пищевой цепочке была отмечена у гидромедузы *Aequorea victoria*, из которой целентеразин исторически впервые был выделен.

Случай *de novo* синтеза впервые был зарегистрирован у креветки *Systellaspis debilis* в экспериментах по последовательному количественному измерению целентеразина и тотальной РНК в яйцах этого ракообразного. Было показано, что концентрация целентеразина растёт в ходе эмбрионального развития, что можно объяснить только *de novo* биосинтезом, поскольку на ранних стадиях личинка не питается (Thomson et al., 1995). Присутствие биолюминесценции у ранних эмбриональных стадий гребневика *Mnemiopsis leidyi* также свидетельствует о самостоятельном синтезе соединения (Freeman, Reynolds, 1973). Эксперименты по кормлению копепод *Metridia pacifica* и *Metridia lucens* мечеными аминокислотами показали, что эти ракообразные синтезируют целентеразин *de novo* из 1 молекулы L-фенилаланина (F) и 2 молекул L-тирозина (Y) (Рисунок 8) (Oba et al., 2009; Tessler et al., 2018). Эксперименты по кормлению гребневиков *Bolinopsis infundibulum* и *Mnemiopsis leidyi* не биолюминесцентными объектами также продемонстрировали способность этих животных к *de novo* биосинтезу целентеразина (Bessho-Uehara et al., 2020). На основании этих данных было выдвинуто предположение о возможной циклизации целентеразина из линейного трипептида-предшественника вида $H_2N-F-Y-Y-COOH$ (Oba et al., 2009). Это предположение сподвигло начать поиски мотива FYY на концах закодированных в геноме пептидов путём анализа транскриптомов биолюминесцентных и небиолюминесцентных видов гребневиков (Francis et al., 2015). Обнаруженные кандидатные пептиды с мотивом FYY на С-конце были гомологичны с

оксидоредуктазами: изопенициллин-N-синтазами и оксигеназами. Интересно, что кандидатные пептиды не были обнаружены у неблюминесцентных видов гребневиков. В этом же исследовании (Francis et al., 2015) авторы дополняют гипотезу о возможном механизме. Они предполагают, что наиболее очевидный вариант — циклизация и дальнейшая модификация аминокислотных остатков FYY, как части более длинного пептида-прекурсора, который синтезируется рибосомально и затем расщепляется и циклизуется. Стоит отметить, что аналогичный поиск мотива FYY в транскриптах *Metridia lucens* не увенчался успехом (Tessler et al., 2018). Был выдвинут и альтернативный вариант: трипептид синтезируется в результате нерибосомного синтеза путем соединения свободных аминокислот. Подобный синтез теоретически мог бы осуществляться либо рядом отдельных ферментов, которые формируют ди- и трипептидные интермедиаты, которые затем циклизуются их в финальную структуру, либо единой нерибосомной пептид синтазой (НРПС), связывающей аминокислотные остатки и циклизующей их. Таким образом, различные исследования биосинтеза целентеразина в биологических объектах проводятся уже десятки лет, однако остаётся всё ещё не расшифрованным. Неизвестно также синтезируется ли целентеразин животными или их бактериальными симбионтами. Обнаружение генов биосинтеза целентеразина позволило бы проводить биотехнологический синтез данного вещества, а также расширило бы применимость целентеразина для разработки биосенсоров на основе люминесценции.

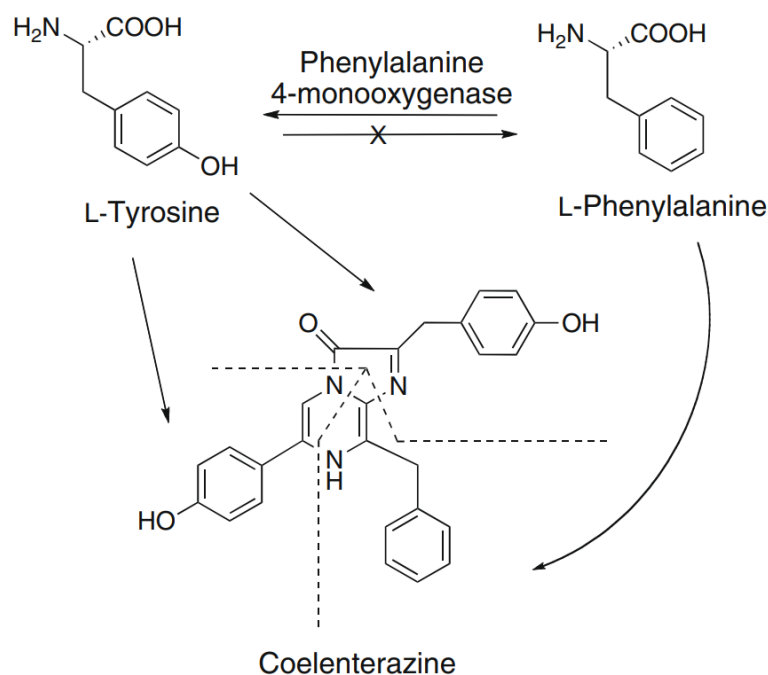


Рисунок 8. Предполагаемый механизм биосинтеза целентеразина из двух аминокислотных остатков тирозина и одного остатка фенилаланина (Oba, 2009).

2.2.3. Применение целентеразина и целентеразин-зависимых люцифераз в биотехнологии и биомедицине

Высокий квантовый выход реакций целентеразин-зависимой люминесценции, низкий уровень фоновой люминесценции и обширные участки линейной зависимости между концентрацией фотопротейна или люциферазы, окисляющих целентеразин, и сигналом люминесценции являются причинами широкого и успешного применения целентеразиновых систем в исследованиях, биомедицине и биотехнологии. На данный момент получены сотни химических модификаций целентеразина, обладающих большей стабильностью, улучшенными свойствами и измененным спектром люминесценции (Contant, et al., 2020).

Для практического применения используются как фотопротейиновые, так и люцифериновые системы, клонированные из различных морских беспозвоночных. Так, например, для детекции активных форм кислорода (АФК) *in vitro* и *in vivo*, применяют неканонический фотопротейин из моллюсков, который опосредует реакцию окисления целентеразина и запускает люминесценцию в присутствии АФК (Kuse, 2014).

Аналогично активно используются Ca^{2+} -зависимые фотопротеины экворин (из *Aequorea victoria*) и обелин (из *Obelia longissima*) из кишечнорастворимых для детекции уровня внутриклеточного кальция (Bonora et al., 2013). Широко применяются фотопротеин-целентеразиновые репортеры для измерения концентраций веществ (гормонов, маркеров опухолей, цитокинов, антител) в медицинской диагностике методами иммуноферментного анализа (ИФА) и измерения в растворе (Krasitskaya, 2020). Данные системы обладают высочайшей чувствительностью, что обусловлено высоким квантовым выходом фотопротеин-целентеразиновых репортеров и чувствительностью спектрофотометров, способных детектировать аттомолярные концентрации фотопротеина (Frank, 2010).

Люциферин-целентеразиновые репортеры активно применяются в *in vivo* исследованиях, в том числе для биолуминесцентной микроскопии. Небольшой размер люцифераз (36 кДа Rluc из *Renilla*, 19 кДа OLuc из *Oplophorus*, 16.5 кДа из *Metridia*) позволяет экспрессировать их в клетках и живых организмах, а также сливать последовательность люциферазы с белком-мишенью. На основе OLuc ведется активная разработка более активных, стабильных и компактных люцифераз (серия NanoLuc). Классическим примером использования таких систем является детекция межмолекулярного белок-белкового взаимодействия *in vitro* и *in vivo*. В данном подходе один белок слит с люциферазой, а второй белок, взаимодействие которого с первым проверяется, слит с флуоресцентным белком (GFP или любым другим, с которым возможен резонансный перенос энергии люминесценции - BRET). В случае взаимодействия между белками происходит BRET и детектируются фотоны определенной длины волны (Kobayashi et al., 2019). Наилучшая чувствительность данного метода достигнута с использованием люциферазы NanoLuc (технология NanoBRET) (Machleidt et al., 2015). Схожие цели по исследованию белок-белковых и белок-лигандных взаимодействий достигаются системой NanoBiT. В данном подходе используется разрезанная на две части люцифераза NanoLuc, каждая из которых не обладает способностью вызывать люминесценцию. В случае взаимодействия белков или белка и лиганда, каждый из которых слит с фрагментом NanoLuc, структура люциферазы восстанавливается и активируется способность к люминесценции (Cooley et al., 2020).

Помимо перечисленных выше примеров множество биосенсоров на основе целентеразин-зависимых люцифераз используется для измерения уровня экспрессии

генов, сплайсинга и поиска транскрипционных факторов. В этих подходах уровень итоговой люминесценции отражает уровень экспрессии репортерного гена люциферазы, закодированного в геноме.

2.3. Микробиомы морских копепод

2.3.1. Общая характеристика и экологическая роль класса копеподы в морских экосистемах

Веслоногие ракообразные или копеподы (Copepoda) — класс ракообразных, повсеместно распространённых в морских акваториях. Среди данного класса ракообразных встречаются как свободноживущие пелагические и бентосные представители, так и паразитические. Большинство свободноживущих копепод имеют каплевидное тело длиной около 0,5–15 мм и длинные антеннулы (Bron et al., 2011). Развитие от яйца до взрослой стадии включает шесть науплиарных и пять копеподитных стадий. Класс копеподы включает 10 отрядов, среди которых Calanoida, Cyclopoida и Harpacticoida являются доминирующими по численности. На текущий момент насчитывается порядка 10800 видов морских копепод (Brun et al., 2017). В морских пелагических сообществах веслоногие ракообразные являются ключевой группой зоопланктона и могут составлять до 80% его биомассы (Verity, Smetacek 1996). С экологической точки зрения, копеподы — консументы второго порядка, основные потребители первичной продукции фитопланктона, а также кормовая база для консументов высших порядков, таких как креветки, молодь рыб и киты (Bron et al., 2011). Хитиновые кутикулы и фекальные гранулы данных ракообразных массово обнаруживаются в толще воды и являются источником питательных веществ для бактерий (Tang et al. 2009). В течение суток копеподы совершают вертикальные миграции, внося вклад в углеродный баланс на разных глубинах (Frangoulis et al., 2005).

2.3.2. Бактериальные сообщества морских копепод

Веслоногие ракообразные заселены различными эписимбиотическими и эндосимбиотическими бактериальными сообществами, различающимися по таксономическому составу между собой и микробиотой окружающей воды (Shoemaker, Moisaner, 2015). Эписимбиотические бактериальные сообщества колонизируют

преимущественно область рта, анального отверстия, межсегментарные участки тела и различные придатки тела, такие как антеннулы, мандибулы и яйцевые мешки взрослых самок (Carman, Dobbs 1997). Копеподы, будучи доминирующим компонентом зоопланктона, и продукты их жизнедеятельности (хитиновые экзувии, фекальные гранулы) формируют центр продукции бактериальной биомассы в пелагиали (De Corte et al., 2014; Tang, 2005; Tang et al. 2009). Между толщей воды и веслоногими ракообразными постоянно происходит циркуляция микробиоты, поскольку скорость пищеварения копепод весьма быстрая и может занимать от нескольких минут до часов. Вследствие ежедневных вертикальных миграций копепод, бактерии довольно эффективно распространяются в толще морской воды, используя зоопланктон как “автостоп” (Grossart et al., 2010; Steinberg et al., 2000), что обеспечивает взаимодействие между пространственно разделёнными сообществами.

Ассоциации между копеподами и бактериями могут быть постоянными или временными, зависят от стадии жизненного цикла хозяина и физиологической активности, а также от условий окружающей среды, таких как диета и глубина воды. Природа ассоциаций, формируемых между копеподами и бактериями, доподлинно неизвестна. В природе ассоциаций, формируемых между копеподами и бактериями, имеет место симбиоз, паразитизм и комменсализм. Колонизация непатогенными бактериями может предотвращать вторичную колонизацию патогенными водорослями и бактериями. Также эписимбиотические сообщества могут разлагать скапливающийся на поверхности кутикулы детрит, препятствующий плавательной активности (Carman, Dobbs, 1997). Первые исследования, посвящённые изучению микробиома копепод, основаны на классических культуральных методах. Подобные работы продемонстрировали преобладание штаммов морских *Vibrio* spp. В том числе были обнаружены потенциально патогенные для человека *V. parahaemolyticus*, *V. alginolyticus*, *V. vulnificus* и *V. cholerae* (Sochard et al., 1979; Kaneko, Colwell, 1975; Montanari et al., 1999). Развитие методов высокопроизводительного секвенирования позволило проводить обширные профилирования бактериомов копепод, учитывающие не только культивируемую микрофлору, но и скрытое разнообразие. Данные мета-баркодирования по вариабельным участкам локуса *16S rRNA* демонстрируют, что у веслоногих ракообразных доминирует ограниченное количество операционных таксономических единиц (OTU). Примерно восемь отрядов бактерий, таких как Actinomycetales, Bacillales, Flavobacteriales, Lactobacillales, Pseudomonales, Rhizobiales и Vibrionales,

идентифицированы как основные представители корового микробиома *Pleuromamma* spp. Анализ 200 представителей *Calanus finmarchicus* установил коровый микробиом, состоящий из 34 OTU, среди которых основными представителями являются бактерии типа Proteobacteria, Actinobacteria и Bacteroidetes. Недавний мета-анализ имеющихся данных о профилях бактериальных сообществ копепод пяти различных родов *Acartia*, *Calanus*, *Centropages*, *Pleuromamma* и *Temora* показал значительные различия в альфа-разнообразии и бета-разнообразии между родами веслоногих (Sadaipappan, 2015).

Кишечник копепод является примечательной средой обитания морских микроорганизмов, поскольку здесь кислотность и насыщение кислородом ниже, чем в окружающей воде (Tang et al., 2011). Как уже обсуждалось, эндобиотическая микрофлора кишечника копепод может варьировать в зависимости от физиологических особенностей исследуемого объекта и условий окружающей среды. Было показано, что бактерии Rhodobacteraceae в изобилии обнаруживаются в особях копепод рода *Acartia* sp. с полным кишечником по сравнению с голодающими особями (Moisander, et al. 2015). Также установлено, что Gammaproteobacteria более многочисленны у голодающих особей *Centropages* sp., *Acartia* sp. (Moisander et al., 2015) и *Pleuromamma* sp. (Cregeen, 2016). Заметное изменение бактериального сообщества между состоянием диапаузы и активно питающейся фазой жизненного цикла отмечено у *Calanus finmarchicus* (Datta et al., 2018). Аналогично количество Flavobacteriaceae скуднее у копепод во время диапаузы и больше у активно питающихся особей (Datta et al., 2018). В этом же исследовании (Datta et al., 2018) отмечено сообщили, что бактерии рода *Marinimicrobium* (Alteromonadaceae) многочисленнее у глубоководных копепод, чем у мелководных. В целом, коровый микробиом копепод зависит не столько от диеты, сколько от места обитания исследуемых веслоногих разнообразных (Moisander et al., 2015). Морская бактериофлора, населяющая кишечника копепод, может играть важную роль в биогеохимических циклах. Эти сообщества могут специализироваться на продукции растворимых соединений железа, анаэробном метагеногенезе (Tang et al., 2011), восстановлении нитритов (Glud et al., 2015), анаэробной азот-фиксации (Scavotto et al, 2015) и денитрификации (De Corte et al., 2018). Более того, данные мета-анализа показывают, что микробиомы разных родов копепод потенциально имеют различное влияние на биогеохимические процессы в морях. Например, ассоциированный с родом *Pleuromamma*, микробиом содержит высокую долю потенциальных генов, ответственных за метаногенез и фиксацию азота, в то время как бактериом *Temora*

содержит высокую долю потенциальных генов, участвующих в восстановлении сульфатов и синтезе цианокобаламина, что является признаками симбиоза (Sadaïappan, 2015).

Несмотря на то, что осуществлялись различные исследования потенциала микробиоты, ассоциированной с копеподами, влиять на биогеохимические процессы в море, изучение вторичных метаболитов данных сообществ не проводилось. Также отсутствуют исследования, изучающие роль микробиоты в биосинтезе целентеразина у копепод способных к люминесценции.

2.4. Описание объектов исследования

В качестве ключевого объекта исследований использовали особей копепод *Metridia longa* (Lübböck, 1854) — массового представителя зоопланктона северных акваторий и, в частности, Белого моря (Рисунок 9). Этот вид широко распространён в северных водах Атлантического океана и повсеместно в водах Северного Ледовитого океана, где зачастую является доминирующим видом зоопланктона. Обнаруживается круглогодично на глубинах от 50 до 1000 м. Взрослые самки *Metridia longa* достигают в длину до 4,5 мм, самцы — до 3,5 мм. Тело у этого вида прозрачное, форма длинная и узкая, abdomen в 1,5 раза короче цефалоторакса. Анализы состава пищи *Metridia longa* показывают, что вид является эврифагом, питается как зоо-, так и фитопланктоном. Наблюдается сезонная динамика в изменении диеты. В летний сезон основу рациона составляет фитопланктон. Осенью, когда продукция фитопланктона сокращается из-за укорачивания светового дня, питание состоит, главным образом, из зоопланктона, в том числе, копепод меньшего размера. Для *Metridia longa* и двух близкородственных видов *Metridia pacifica* и *Metridia lucens* описана целентеразин-зависимая биолюминесценция и наличие специализированных эпидермальных желёз. Представители семейства Metridinidae, в целом, считаются биолюминесцентными.



Рисунок 9. Внешний вид *Metridia longa* (Атлас «Виды - биологические индикаторы состояния морских арктических экосистем», 2020)

Для сравнительного анализа микробиоты был выбран симпатрический не люминесцирующий вид *Calanus glacialis* (Jaschnov, 1955) (Рисунок 10). Данный вид также широко распространён в водах Северного Ледовитого океана на глубинах от 50 до 1000 м и является одним из доминирующих компонентов зоопланктона. Тело прозрачное; придатки тела, антеннулы, мандибулы, последние сегменты абдомена и яйцевые мешки самок, окрашены в оранжевый цвет. Форма тела более овальная, чем у *Metridia longa*, антеннулы, соотношение отделов тела также различается — абдомен в 3 раза короче цефалоторакса. Взрослые самки и самцы в длину достигают до 5,5 мм. Основу рациона питания составляет фитопланктон. Биолюминесценция у семейства Calanidae не была зарегистрирована. Кроме того, предполагается, что данное семейство и ряд других, имеющих миелинизированные нервные волокна и иной характер движения, лишились способности к биолюминесценции в результате эволюционного процесса (Takenaka et al., 2017)



Рисунок 10. Внешний вид *Calanus glacialis* (Атлас «Виды - биологические индикаторы состояния морских арктических экосистем», 2020)

3. МАТЕРИАЛЫ И МЕТОДЫ

3.1. Сбор материала, идентификация и вскрытие объектов

Сбор материала осуществляли 5 сентября 2022 года с 14:00 до 19:30 в Кандалакшском заливе Белого моря на территории Беломорской биологической станции МГУ имени Н.А. Перцова совместно с сотрудниками Центра Морских Исследований МГУ. Для вылова планктона использовали сеть Джеди с фильтрующим конусом из газа с ячейей 180 мкм и диаметром устья 37 см. Для вылова были установлены три локации — за островом Величаиха (координаты: 66°29399' с.ш., 33,35165' в.д.), между островом Кокоиха и островом Величаиха (координаты: 66°29841' с.ш., 33°3202' в.д.), напротив острова Кокоиха (координаты: 66°29935' с.ш., 33°30274' в.д.). Минимальная глубина погружения сети Джеди составила 120 м. После отлова планктонную пробу рассортировали на особей *Metridia longa* и *Calanus glacialis*. Сортировка планктонной пробы производили в течение пяти-шести дней. На протяжении этого промежутка времени ракообразных содержали в стеклянных ёмкостях в холодильнике при температуре 4°C. Все это время копеподы оставались жизнеспособными. Определение таксономической принадлежности планктонных ракообразных, стадии жизненного цикла и половой принадлежности проводили на биостанции совместно с сотрудницей Центра Морских Исследований МГУ, Юриковой Дарьей Андреевной, по внешним морфологическим признакам. Дополнительно видовая принадлежность *Metridia longa* была установлена по флуоресценции биолюминесцентных желёз под действием источника УФ-излучения с длиной волны от 365 до 400 нм. У представителей *Calanus glacialis* флуоресценция отсутствовала. Для дальнейших манипуляций отобрали взрослых самцов и самок и V-копеподитные стадии *Metridia longa* и *Calanus glacialis*. Также были отобраны пробы морской воды с мест вылова зоопланктона. Морскую воду объёмом 3 литра пропускали через стерильные 0,22 мкм колоночные фильтры (Sterivex, Германия) для дальнейшего выделения ДНК из отфильтрованной бактериальной фракции. Особей *Metridia longa* и *Calanus glacialis*, колоночные фильтры с осажённой бактериальной фракцией замораживали в кельвинаторе при температуре -80°C и далее после экспедиции в замороженном виде транспортировали в Москву в Институт Биологии Гена РАН. Для получения препаратов кишечника *Metridia longa* проводили вскрытие стерильными иглами (d = 0,18 мм) с использованием стереоскопического микроскопа, предварительно антисептированным раствором 96% этанола. Последовательность действий при вскрытии была следующей: 1) помещение особей

Metridia longa в центр стерильной пластиковой чашки Петри в каплю воды, очищенной по системе Milli-Q; 2) фиксация ракообразного одной стерильной иглой; 3) иссечение абдоминальных или цефалотораксных сегментов второй стерильной иглой; 4) извлечение кишечника из открывшейся полости и помещение в пробирки объемом 1.5 мл с заранее приготовленным лизис-буфером (состав лизис-буфера зависел от выбора дальнейшей методики выделения ДНК, см. подраздел “Экстракция тотальной ДНК из образцов морских копепод и морской воды”). Помимо капли воды для препарата, в чашку Петри добавляли ещё одну дополнительную каплю в качестве негативного контрольного образца вскрытия для дальнейшего анализа состава микробиоты на предмет лабораторного загрязнения.

3.2. Экстракция тотальной ДНК из образцов морских копепод и морской воды

Экстракцию тотальной ДНК производили несколькими методами, сводная информация представлена в таблице ниже (Таблица 1). Все полученные экстракты относятся к 5 разным типам: 1) морская вода; 2) кишечник *Metridia longa*; 3) целые *Metridia longa* (обозначены в Таблице 1 как *Metridia longa*); 4) целые *Calanus glacialis*; 5) негативные контроли (пояснение к негативным контролям описано в разделе 3.1. выше). Каждому полученному образцу присвоен индивидуальный идентификационный номер (ID).

Для получения ДНК-экстрактов из кишечника *Metridia longa* использовали, в среднем, 40 фрагментов пищеварительной системы на образец. Для получения ДНК-экстрактов из целых ракообразных использовали, в среднем, 20 особей. Аутопсированные фрагменты кишечника *Metridia longa* и непосредственно целые ракообразные *Metridia longa*, *Calanus glacialis* подвергали предварительной гомогенизации пластиковым пестиком в лизис-буфере до исчезновения крупных фрагментов ткани (состав раствора зависел от метода экстракции ДНК). Четыре образца (ID: 5038, 5040, 5042, 10803) получили путём выделения тотальной ДНК с помощью набора реактивов GeneJet Genomic DNA Purification Kit (Thermo Fisher Scientific, США) в соответствии со стандартным протоколом производителя для Грам(+)-бактерий. Четыре образца (ID: 5044, 5046, 5048, 5050) получили путём выделения тотальной ДНК с помощью набора DNA Prep 200 (Diatom, Россия) в соответствии со стандартным протоколом производителя. Двадцать образцов (ID: 10773, 10785, 5652, 8305, 8307,

10775, 10787, 10777, 10789, 10779, 10791, 8309, 10797, 10799, 10801, 8311, 10781, 10793, 10783, 10795) получили с помощью экстракции хлороформом следующим образом. Гомогенизацию осуществляли в 100 мкл рабочего раствора (900 мкл стоковый раствор А (1М Tris Cl (pH 9,0), 0,5 М EDTA, mili-Q вода), 100 мкл 10% SDS, 10 мкл DEPC). Затем добавляли еще 400 мкл рабочего раствора, перемешивали и инкубировали при 70°C 30 мин. После добавляли 116 мкл 5М раствора ацетата калия, перемешивали и инкубировали при 0°C 30 мин. Добавляли 600 мкл хлороформа, перемешивали и полученную смесь центрифугировали 10 мин при 13400 g, 250 мкл водной фазы, содержащей ДНК, переносили в чистые пробирки объемом 1,5 мл, добавляли 3 мкл 1% раствора гликогена и 350 мкл изопропанола, перемешивали, инкубировали в течение ночи при температуре -20°C и центрифугировали 15 мин при 13400 g. Осадок промывали 2 раза 70%-ным этанолом (500 мкл), высушивали при комнатной температуре и растворяли в 30 мкл свободной от нуклеаз воды.

Экстракт тотальной ДНК морской воды (ID: 5052 (5650)) выделяли из колоночного фильтра (Sterivex, Германия) с помощью набора реагентов DNeasy Blood and Tissue Kit (Qiagen, Германия) в соответствии с модифицированным протоколом. Последовательность действий была следующей: 1) разбор пластиковой колонки в стерильных условиях на чашке Петри с помощью скальпеля; 2) перемещение пластины 0,22 мкм фильтра из разобранной колонки на чашку Петри и разрезание стерильными ножницами; 3) помещение нарезанных фрагментов в пробирки объемом 2 мл; 4) 60-минутная инкубация фрагментов в качалке при температуре 37°C в 180 мкл раствора STET (8% сахарозы, 5% Triton® X-100, 50мМ EDTA, 50мМ Tris-HCl, pH 8.0) с добавлением 20 мкл лизоцима. Последующие шаги соответствовали стандартному протоколу производителя.

Образцы с методом выделения, помеченным в таблице как “Все”, получали путём объединения всех ДНК-экстрактов соответствующего типа и дальнейшим вакуумным концентрированием в приборе Thermo Scientific™ Savant™ SpeedVac™ DNA130 в течение 30 минут при комнатной температуре.

Для шот-ган метагеномного секвенирования ДНК-экстракты подвергались дополнительной обработке следующим образом. Образец 5652, соответствующий типу “*Metridia longa*”, очищали от возможной контаминации РНК согласно следующему протоколу: 1) обработка целевого образца РНКазой А; 2) добавление 3М раствора ацетата натрия, равного 0,1 объёма целевого образца; добавление 100% раствора

этилового спирта, охлаждённого при температуре -20°C, равного 2,5-3 объёмам целевого образца; перемешивание смеси с помощью вортекса; 3) преципитация при температуре -20°C в течение ночи; 4) центрифугирование со скоростью 13000 оборотов в минуту при температуре 4°C в течение часа; 5) промывание осадка два раза с помощью добавления 75% раствора этилового спирта, охлаждённого при температуре -20°C, объёмом 0,5 мл и последующим центрифугированием со скоростью 13000 оборотов в минуту при температуре 4°C в течение 10 минут; 6) удаление раствора этилового спирта из пробирки и высушивание осадка при комнатной температуре с открытой крышкой. Далее образец 5652 и образец 5052 (5650) подвергались очистке от коротких фрагментов ДНК с помощью набора реактивов PCR & DNA Cleanup Kit (5 µg) (Monarch, Великобритания) по протоколу, рекомендованному производителем.

Концентрацию ДНК в образцах измеряли с помощью прибора Qubit 3.0 Fluorometer с использованием набора реактивов HS dsDNA HS Assay kit (ThermoFisher Scientific, США). Растворы ДНК хранили в морозильной камере при -20°C.

Таблица 1. Список полученных экстрактов тотальной ДНК из образцов разных типов

Присвоенный ID	Тип образца в зависимости от источника ДНК	Метод выделения
5038	Кишечник <i>Metridia longa</i>	ThermoScientific GeneJet Genomic DNA Purification Kit
5040	Кишечник <i>Metridia longa</i>	ThermoScientific GeneJet Genomic DNA Purification Kit
5042	Кишечник <i>Metridia longa</i>	ThermoScientific GeneJet Genomic DNA Purification Kit
5044	Кишечник <i>Metridia longa</i>	Diatom DNA Prep 200
5046	<i>Metridia longa</i>	Diatom DNA Prep 200
5048	Негативный контроль (капля воды)	Diatom DNA Prep 200
5050	Негативный контроль (капля воды)	Diatom DNA Prep 200
5052 (5650)	Морская вода (колоночный фильтр)	DNeasy Blood&Tissue
10773	Кишечник <i>Metridia longa</i>	Экстракция хлороформом
10785	Негативный контроль (капля воды)	Экстракция хлороформом
5652	<i>Metridia longa</i>	Экстракция хлороформом
8305	<i>Metridia longa</i>	Экстракция хлороформом
8307	<i>Metridia longa</i>	Экстракция хлороформом
10775	Кишечник <i>Metridia longa</i>	Экстракция хлороформом
10787	Негативный контроль (капля воды)	Экстракция хлороформом
10777	Кишечник <i>Metridia longa</i>	Экстракция хлороформом

10789	Негативный контроль (капля воды)	Экстракция хлороформом
10779	Кишечник <i>Metridia longa</i>	Экстракция хлороформом
10791	Негативный контроль (капля воды)	Экстракция хлороформом
8309	<i>Metridia longa</i>	Экстракция хлороформом
10797	<i>Calanus glacialis</i>	Экстракция хлороформом
10799	<i>Calanus glacialis</i>	Экстракция хлороформом
10801	<i>Calanus glacialis</i>	Экстракция хлороформом
8311	<i>Metridia longa</i>	Экстракция хлороформом
10781	Кишечник <i>Metridia longa</i>	Экстракция хлороформом
10793	Негативный контроль (капля воды)	Экстракция хлороформом
10783	Кишечник <i>Metridia longa</i>	Экстракция хлороформом
10795	Негативный контроль (капля воды)	Экстракция хлороформом
8313	Особи <i>Metridia longa</i>	Все
10803	Негативный контроль (капля воды)	ThermoScientific GeneJet Genomic DNA Purification Kit
12758	Кишечник <i>Metridia longa</i>	Все

3.3. Секвенирование ампликонных библиотек переменного участка V3-V4 гена *16S pPHK* и шот-ган метагеномное секвенирование

Все этапы пробоподготовки библиотек для секвенирования и секвенирование осуществлены компанией Евроген (Россия).

Ампликонные библиотеки переменного участка V3-V4 гена *16S pPHK* были подготовлены согласно протоколу 16S Metagenomic Sequencing Library Preparation 15044223 В (Illumina, США) (URL: https://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html). Секвенирование осуществляли на приборе NovaSeq 6000 (длина прочтений - 250 п.о., парноконцевые прочтения) из расчета 100000 кластеров на одну ампликонную библиотеку

Качество образцов метагеномной ДНК проверяли с помощью электрофореза ДНК в агарозном геле. Количественную оценку проводили путем измерения концентрации ДНК на приборе Qubit 2.0. с использованием набора реактивов HS dsDNA HS Assay kit (ThermoFisher Scientific). Подготовку библиотек полных метагеномов проводили с помощью набора Illumina DNA Prep (Illumina). Качество получаемых библиотек оценивали при помощи прибора Fragment Analyzer (Agilent). После получения

библиотеки были очищены и смешаны эквимольно с помощью системы SequalPrep™ Normalization Plate Kit (ThermoFisher, Cat # A10510-01). Контроль качества полученных пулов библиотек проводили с помощью системы Fragment Analyzer, количественный анализ - при помощи количественной ПЦР. Пул библиотек секвенировали на дорожке типа SP для NovaSeq 6000 (длина прочтений - 250 п.о., парноконцевые прочтения) из расчета 100 млн кластеров на одну библиотеку.

Файлы FASTQ для библиотек обоих типов были получены с помощью программного обеспечения bcl2fastq v2.17.1.14 Conversion Software (Illumina, США).

3.4. Обработка данных секвенирования переменного участка V3-V4 гена *16S rPHK*

Помимо перечисленных в подразделе 3.2. образцов для совместного анализа результатов секвенирования были добавлены первичные данные секвенирования переменного участка V3-V4 гена *16S rPHK* из образцов ДНК воды Белого моря, полученные сотрудниками Лаборатории анализа метабеномов Сколковского института науки и технологии. Данный материал был собран в той же экспедиции, что и материал для текущей работы.

Первичные прочтения обрезали и отфильтровали с помощью инструмента Trimmomatic (в. 0.39) (SE -phred 33 HEADCROP 17 ILLUMINACLIP:2:30:10 LEADING:0 TRAILING:0 SLIDINGWINDOW:4:0 HEADCROP:17 MINLEN:150) (Bolger et al., 2014). Необработанные и обработанные с помощью Trimmomatic прочтения анализировали с помощью FastQC (в. 0.11.9) для контроля длины, качества секвенирования и наличия последовательностей адаптеров (URL: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Далее отфильтрованные прямые прочтения обрабатывали с помощью пайплайна DADA2 (в. 1.22.0) (Callahan, 2016) с дополнительной обрезкой, и исправлением ошибок с целью получения вариантов последовательности ампликона (ASV, amplicon sequence variant). Далее полученные ASV кластеризовали с помощью инструмента MMseqs2 (в. 13.45111) (Steinegger, Söding, 2020) (покрытие > 0,95, идентичность > 0,98), репрезентативные последовательности кластеров далее обрабатывались как оперативные таксономические единицы (OTU, operational taxonomic units). После сформированные OTU снова загрузили в DADA2 с целью присвоения таксономической принадлежности для каждой с использованием

данных таксономического обучения SILVA SSU. Перечисленный выше алгоритм был разработан ранее в нашей лаборатории (Rusanova et al., 2021). Полученный датасет, включающий таблицу представленности OTU для каждого образца, таблицу таксономической принадлежности OTU и метаданные для каждого образца, анализировали с помощью пакета phyloseq (в. 1.44.0) (McMurdie, Holmes, 2013) и провели процедуру деконтаминации на основе данных из образцов негативных контролей с помощью пакета decontam (в. 1.20.0) (Davis et al, 2018) по методу “Распространённость” (Prevalence) с пороговым значением = 0.5 в RStudio. Оценку относительных частот таксонов, анализ альфа- и бета-разнообразия, статистические тесты проводили в онлайн сервисе MicrobiomeAnalyst (Dhariwal et al., 2017). Предварительно, было осуществлено удаление низкопредставленных OTU. Были исключены OTU, которые менее чем в 10% образцов имели частоты менее 4.

3.5. Обработка данных шот-ган метагеномного секвенирования

Первичные прочтения обрезали и отфильтровали с помощью Trimmomatic (в. 0.39) (PE-phred 33 LEADING:3 TRAILING:3 ILLUMINACLIP:2:30:10 MINLEN:36). Необработанные и обработанные с помощью Trimmomatic прочтения анализировали с помощью FastQC (в. 0.11.9) для контроля длины, качества секвенирования и наличия последовательностей адаптеров. Отфильтрованные парные прочтения и прямые непарные прочтения собирали с помощью SPAdes (в. 3.15.3) (с включенной опцией metaspades и длиной k-mer 55, 99, 121, 127) для сборки *de novo* (Bankevich et al., 2012). Полученные сборки анализировали с помощью QUAST (Gurevich et al., 2013). Для дальнейшего анализа отбирали контиги длиной более 1 т.п.н. После провели аннотацию полученных контиг инструментом Prokka (в. 1.12) (Seemann, 2014). С целью предсказания кластеров генов биосинтеза вторичных метаболитов бактерий аннотации Prokka и набор контиг длиной более 1 т.п.н. обрабатывали с помощью antiSMASH (в. 6.0.1) (Medema et al., 2011), после чего провели кластеризацию предсказанных кластеров инструментом BiG-SCAPE (в. 1.1.2) (Navarro-Muñoz et al., 2020). Визуализация количества и разнообразия биосинтетических кластеров была проведена скриптами, разработанными ранее в нашей лаборатории (Rusanova et al., 2021). Также провели предсказание открытых рамок считывания для набора контиг с помощью MetaGeneMark (в. 3.25) (Gemayel et al., 2022). Далее пептиды, предположительно кодируемые

предсказанными открытыми рамками считывания, были отобраны по принципу наличия мотива FYY в них с помощью скрипта, написанного на языке программирования Python. Для продуктов, содержащих искомый мотив на С-конце, провели поиск гомологов с помощью алгоритма blastp в базе данных UniProtKB (URL: <https://www.uniprot.org/>).

4. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

4.1. Выделение ДНК из образцов морских копепод и морской воды для дальнейшего секвенирования, выбор оптимального метода

Суммарно были получены 31 образец тотальной ДНК, предназначенных для дальнейшего секвенирования. Из них 9 образцов относятся к типу “Негативный контроль”: 5048, 5050, 10785, 10787, 10789, 10791, 10797, 10795, 10803; 10 образцов к типу “Морская вода”: 4928, 4929, 4930, 4931, 4932, 4933, 4934, 4935, 4936, 5052 (5650) ; 3 образца к типу “*Calanus glacialis*”; 11 образцов к типу “Кишечник *Metridia longa*”: 5038, 5040, 5042, 5044, 10773, 10775, 10777, 10779, 10781, 10783, 12758; 7 образцов к типу “*Metridia longa*”: 5046, 5652, 8305, 8307, 8309, 8311, 8313.

Экстракцию тотальную ДНК осуществляли несколькими разными способами (см. подраздел 3.2 в Материалах и методах), поскольку подготовка библиотек для высокопроизводительного секвенирования требует получения достаточно большого количества ДНК в образце высокого качества и эта задача может оказаться трудновыполнимой при работе природными образцами. В таблице ниже отражена эффективность методов выделения согласно итоговому выходу ДНК по массе в растворе (Таблица 2). Метод экстракции хлороформом был определён как наиболее оптимальный для образцов ракообразных. В итоге большая часть экстрактов ДНК была получена именно таким способом. Однако нельзя не подчеркнуть сложность работы с этим протоколом в случае большого количества образцов. Данный метод гораздо более трудоёмкий и времязатратный, чем использование коммерческих наборов для выделения ДНК. Также использование описанного протокола экстракции хлороформом требует последующей дополнительной очистки от примесей РНК и коротких фрагментов ДНК. Выделение на колонках с помощью коммерческих наборов освобождает от дополнительной постобработки, но даёт риск слабой элюции ДНК.

Все полученные образцы использовали для секвенирования ампликонных библиотек переменного участка V3-V4 гена *16S рРНК*. Образец *Metridia longa* с идентификационным номером 5652 отобрали для дальнейшего шот-ган метагеномного секвенирования, поскольку в нём итоговая масса ДНК составила 3420 нг, что является максимумом среди всех экстрактов данного типа. Также для шот-ган метагеномного секвенирования использовали образец морской воды 5052 (5650) с массой ДНК 1164 нг.

Таблица 2. Результаты выделения ДНК разными способами из всех образцов.

Присвоенный ID	Тип образца	Метод выделения	Концентрация, нг/мл	Объём, мл	Выход ДНК, нг
5038	Кишечник <i>Metridia longa</i>	ThermoScientific GeneJet Genomic DNA Purification Kit	0.3	100	30
5040	Кишечник <i>Metridia longa</i>	ThermoScientific GeneJet Genomic DNA Purification Kit	0.5	100	50
5042	Кишечник <i>Metridia longa</i>	ThermoScientific GeneJet Genomic DNA Purification Kit	0.2	100	20
5044	Кишечник <i>Metridia longa</i>	Diatom DNA Prep 200	-	150	-
5046	<i>Metridia longa</i>	Diatom DNA Prep 200	0.3	150	45
5048	Негативный контроль (капля воды)	Diatom DNA Prep 200	-	150	-
5050	Негативный контроль (капля воды)	Diatom DNA Prep 200	-	150	-
5052 (5650)	Колоночный фильтр	DNeasy Blood&Tissue	19.4	60	1164
10773	Кишечник <i>Metridia longa</i>	Chloroform	9.24	16	147,84
10785	Негативный контроль (капля воды)	Chloroform	-	16	-
5652	<i>Metridia longa</i>	Chloroform	114	30	3420
8305	<i>Metridia longa</i>	Chloroform	4	40	160
8307	<i>Metridia longa</i>	Chloroform	4.7	40	188
10775	Кишечник <i>Metridia longa</i>	Chloroform	1.9	40	76
10787	Негативный контроль (капля воды)	Chloroform	-	40	-
10777	Кишечник <i>Metridia longa</i>	Chloroform	0.2	40	8
10789	Негативный контроль (капля воды)	Chloroform	-	40	-
10779	Кишечник <i>Metridia longa</i>	Chloroform	0.366	40	14,64
10791	Негативный контроль (капля воды)	Chloroform	-	40	-
8309	<i>Metridia longa</i>	Chloroform	66.2	10	662
10797	<i>Calanus glacialis</i>	Chloroform	0.508	20	10,16
10799	<i>Calanus glacialis</i>	Chloroform	0.310	20	6,2
10801	<i>Calanus glacialis</i>	Chloroform	2.420	20	48,4
8311	<i>Metridia longa</i>	Chloroform	1.828	20	36,56
10781	Кишечник <i>Metridia longa</i>	Chloroform	0.368	20	7,36
10793	Негативный контроль (капля воды)	Chloroform	-	20	-
10783	Кишечник <i>Metridia longa</i>	Chloroform	0.5	20	10

10795	Негативный контроль (капля воды)	Chloroform	-	20	-
8313	<i>Metridia longa</i>	Все	30	40	1200
10803	Негативный контроль (капля воды)	ThermoScientific GeneJet Genomic DNA Purification Kit	-	200	-
12758	Кишечник <i>Metridia longa</i>	Все	3.3	30	99

4.2. Сравнительный анализ микробных сообществ на основе данных секвенирования ампликонов варибельного участка V3-V4 гена *16S рPHK*

4.2.1. Анализ первичных данных и удаление контаминации

Всего было проанализировано 40 образцов секвенирования ампликонных библиотек варибельного участка V3-V4 гена *16S рPHK*, относящихся к 5 разным типам: 1) образцы морской воды; 2) образцы кишечника *Metridia longa*, 3) образцы *Metridia longa*, 4) образцы *Calanus glacialis*, 5) негативные контроли вскрытия *Metridia longa*. Из них 31 на этапе выделения ДНК получен лично автором, 9 были включены в набор данных из готовых материалов лаборатории (описано в подразделе 3.4.). Идентификационные номера заимствованных образцов отражены в таблице ниже (Таблица 3).

Таблица 3. Идентификационные номера (ID) дополнительно включенных в набор данных.

ID	Тип образца
4928	Морская вода
4929	Морская вода
4930	Морская вода
4931	Морская вода
4932	Морская вода
4933	Морская вода
4934	Морская вода
4935	Морская вода
4936	Морская вода

Первичные данные суммарно включали 4090414 прочтений и 9045 OTU, среднее количество прочтений на образец составило 102260 со стандартным отклонением 58369 (Рисунок 11). Можно отметить, что количество прочтений в контрольных образцах

значительно ниже, чем в целевых образцах: среднее количество прочтений 40679 и стандартное отклонение 40905 для контрольных образцов; среднее количество прочтений 120139 и стандартное отклонение 50147 для целевых образцов (р-значение t-теста 0,0002).

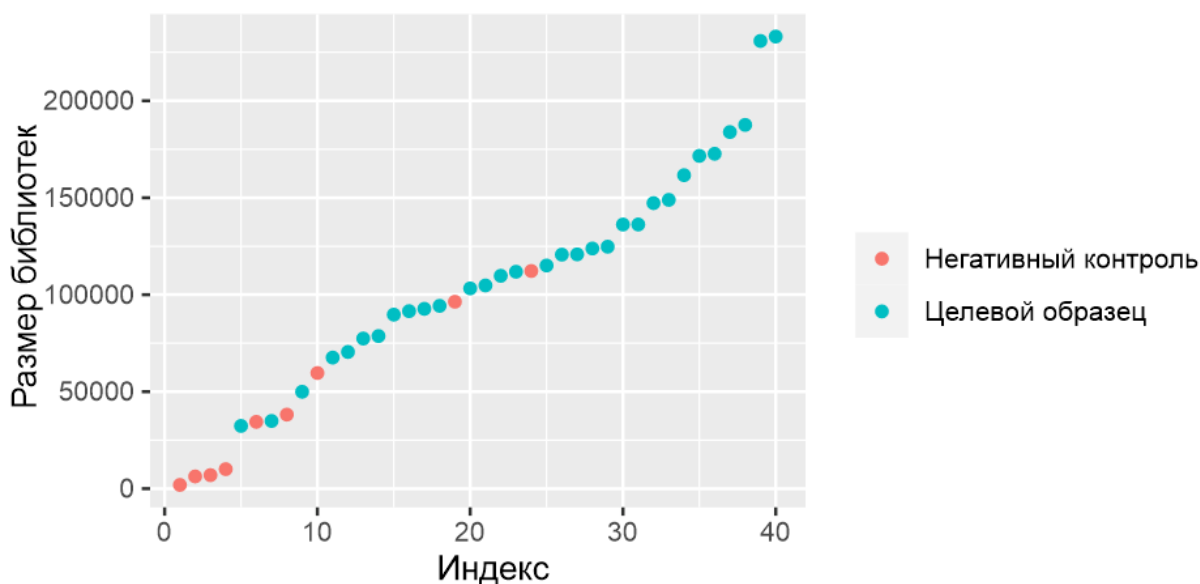


Рисунок 11. Точечная диаграмма размеров библиотек для ампликонов варибельного участка V3-V4 гена *16S рРНК* до процедуры декомнтинатации. Каждая точка соответствует одному образцу. Образцы сгруппированы по типам: негативные контроли и целевые образцы.

График анализа главных координат, рассчитанных по несходствам Брея-Кёртиса, демонстрирует различия в разнообразии микробных сообществ разных типов образцов (Рисунок 12). Данный анализ показал, что образцы морской воды (зелено-коричневые точки) и образцы ракообразных (красные, синие и розовые точки) сгруппированы раздельно друг от друга, что свидетельствует о значительных отличиях между микробиомами. Образцы негативных контролей (зеленые точки) незначительно перекрываются с областью точек образцов кишечника *Metrdia longa*, что означает низкий уровень контаминации и подчёркивает необходимость негативных контролей в исследованиях микробиоты аутопсированных элементов. Выводы о различиях микробиомов разных групп образцов подтверждается перестановочным многомерным дисперсионным анализом (PERMANOVA), проведенным для совокупности образцов (р-значение = 0.001).

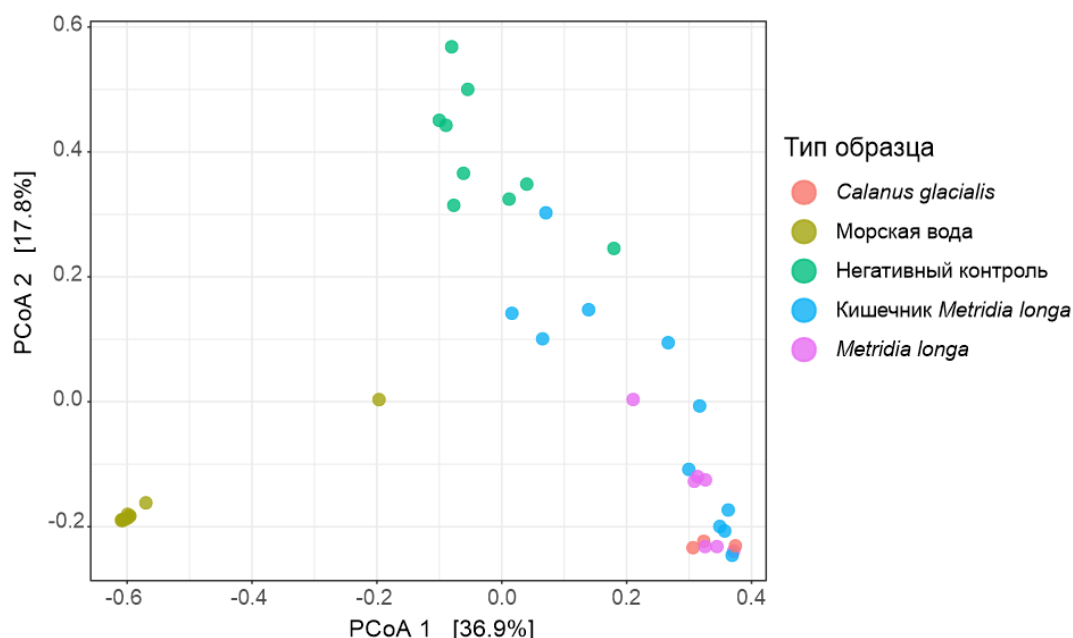


Рисунок 12. График анализа главных координат, рассчитанных по несходствам Брея-Кёртиса, для всех типов исследуемых образцов. Каждая точка соответствует одному образцу.

Удаление контаминации на основе негативных контролей при помощи Decontam показало, что 106 OTU являются контаминантными. После удаления загрязнения суммарное количество прочтений составило 2527584, OTU — 8939. Среднее количество прочтений на образец — 63190 (стандартное отклонение 46575). Размеры библиотек контрольных образцов существенно уменьшились, что видно по соответствующим диаграммам (Рисунки 11 и 13) (p-значение t-теста 0,03). Среднее количество прочтений 7852 и стандартное отклонение 11550 для контрольных образцов; среднее количество прочтений 79255 и стандартное отклонение 39283 для целевых образцов.

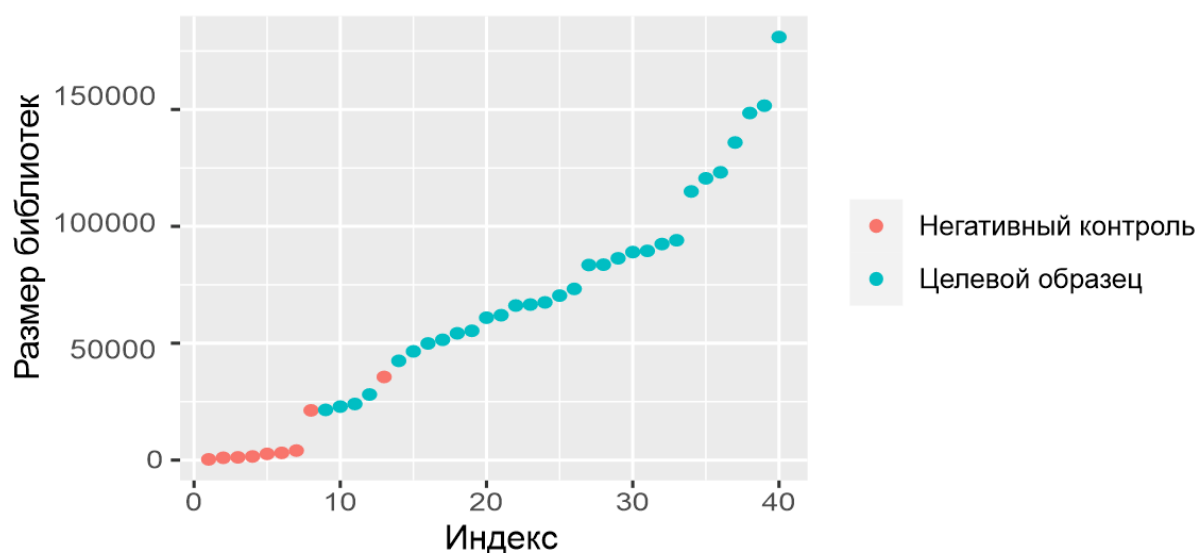


Рисунок 13. Точечная диаграмма размеров библиотек для ампликонов переменного участка V3-V4 гена *16S rRNA* после процедуры декомонтажной. Каждая точка соответствует одному образцу. Образцы сгруппированы по типам: негативные контроли и целевые образцы.

Анализ относительных частот таксонов на уровне типов показывает, что главное отличие в биообразнообразии бактерий в негативных контролях заключается в наличии представителей Actinobacteriota и Firmicutes (Рисунок 14). Также ключевой тип в образцах ракообразных, Proteobacteria, в негативных контролях обнаруживается реже.

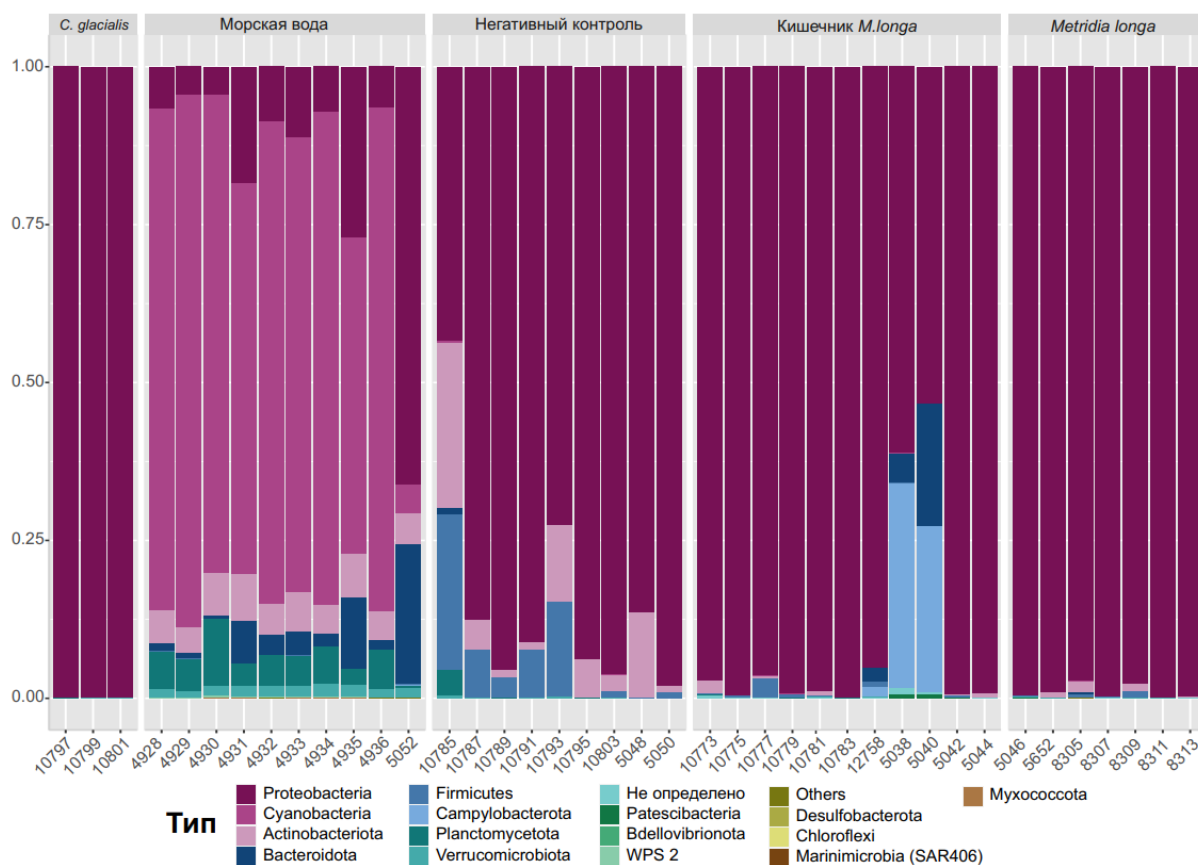


Рисунок 14. Столбчатые диаграммы, отражающие относительные частоты типов бактерий в образцах до удаления контаминации. Диаграммы сгруппированы в зависимости от типа образца.

Анализ разнообразия бактерий в отрицательных контролях продемонстрировал наличие родов бактерий, к которым относятся лабораторные штаммы (*Escherichia/Shigella*, *Bacillus*, *Pseudomonas*), а также таксоны, которые характерны для загрязнения лабораторных реактивов и компонентов китов для выделения ДНК (*Bradyrhizobium*, *Rhizobium*, *Delftia*, *Acinetobacter*, *Acidovorax*, *Stenotrophomonas*, *Phyllobacterium*) (Salter et al., 2014) (Рисунок 15). Для вычитания OTU, относящихся к загрязнению, из природных образцов была использована программа Decontam.

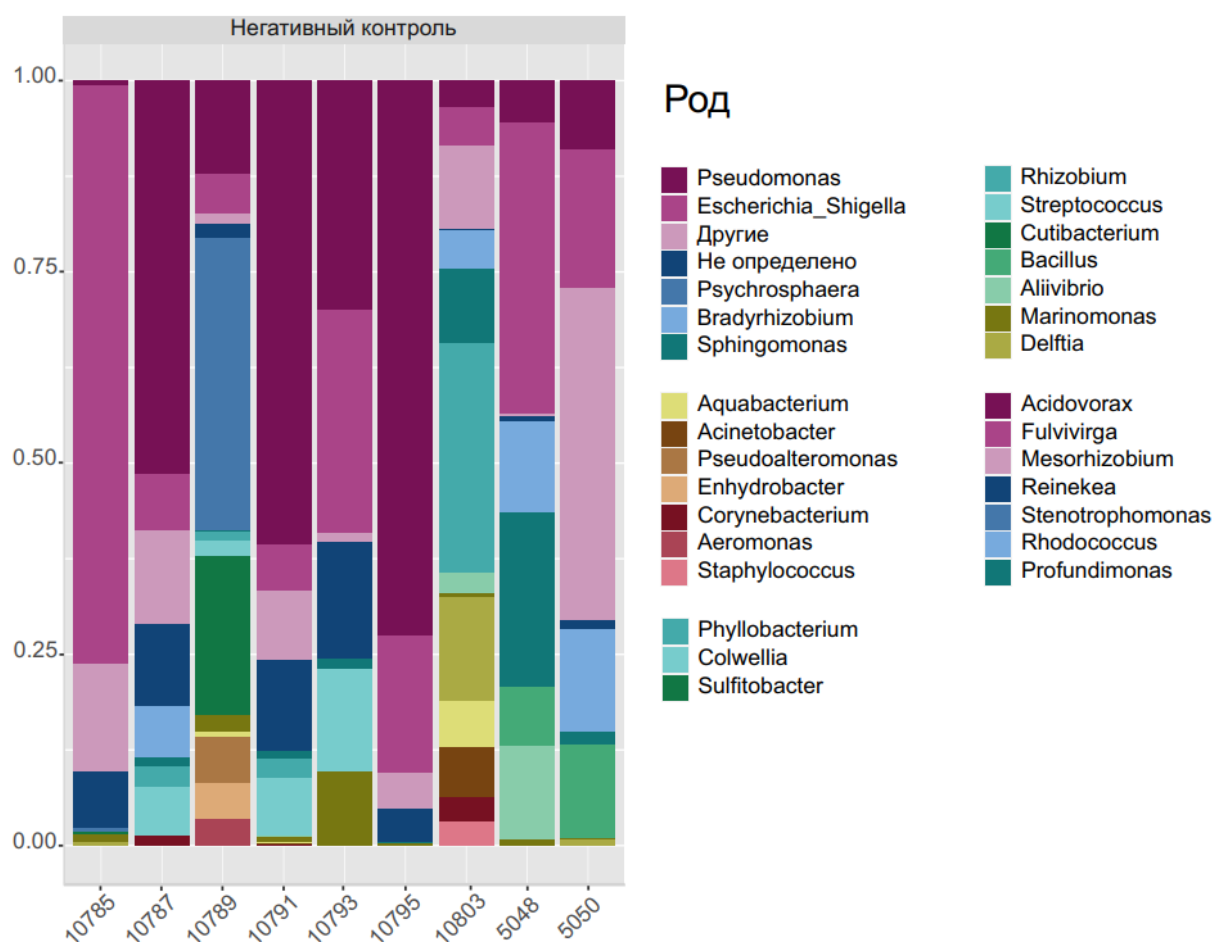


Рисунок 15. Столбчатые диаграммы, отражающие относительные частоты родов бактерий в образцах негативных контролей.

Дальнейший анализ проводили на наборе данных, очищенных от загрязнения и после удаления редких OTU. Согласно анализу кривых разрежения секвенирование было осуществлено на достаточной глубине, о чём свидетельствует выход всех кривых на плато (Рисунок 16).

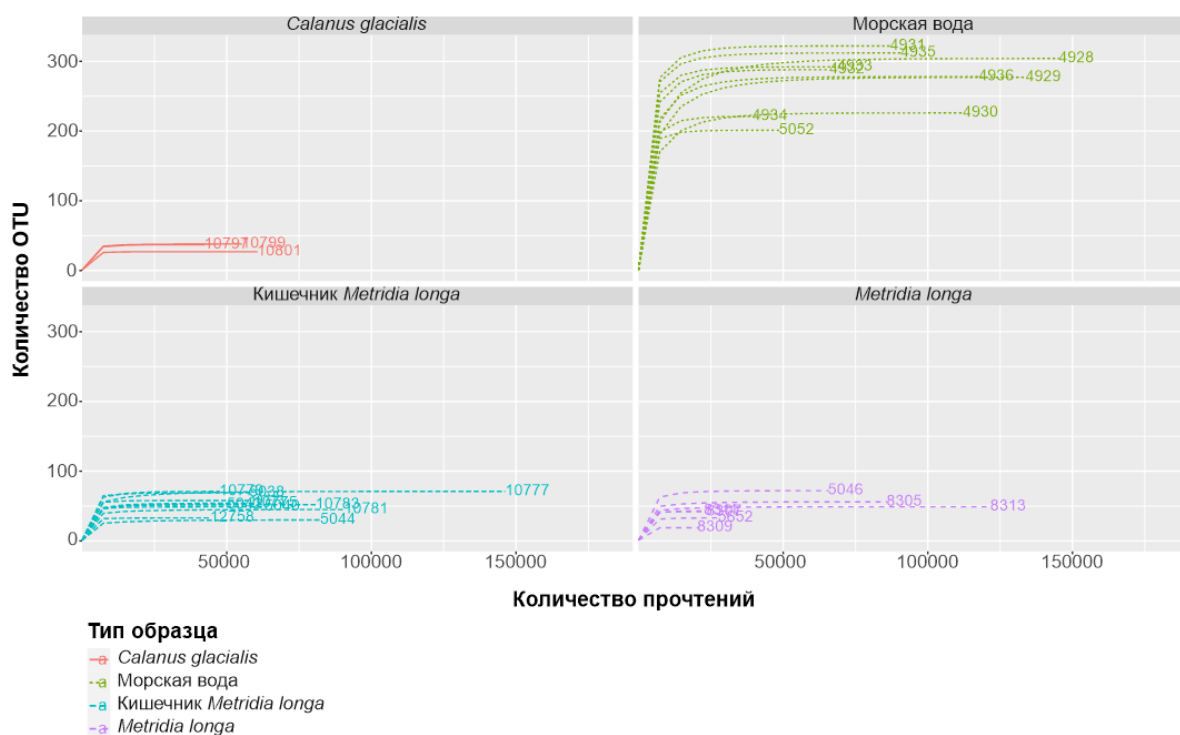


Рисунок 16. Графики кривых разрежения для всех исследуемых образцов разных типов, отражающие зависимость количества идентифицированных OTU от количества прочтений для данного образца. Каждая кривая соответствует одному образцу. Кривые разнесены на 4 графика в зависимости от типа образцов.

4.2.2. Относительные частоты таксонов микробиомов *Metridia longa*, *Calanus glacialis*, кишечника *Metridia longa* и морской воды

Составы микробиомов морской воды и образцов ракообразных существенно различаются (Рисунок 17). В первую очередь, образцы морской воды отличаются большим таксономическим разнообразием. Так доминирующим элементом бактериального сообщества морской воды на уровне типов является Cyanobacteria — один из основных первичных продуцентов в океанах (Vincent, 2012), тогда как сообщество ракообразных представлено, главным образом, Proteobacteria. В образцах ракообразных относительная частота представленности Proteobacteria занимает практически 100%. Согласно исследованию 200 представителей *Calanus finmarchicus* этот тип бактерий относится к коровому микробиому копеподы (Datta et al., 2018). Помимо Cyanobacteria, микробиом морской воды также отличается присутствием Actinobacteriota, Planctomycetota, Verrucomicrobiae. В двух образцах кишечника *Metridia longa* (5038, 5040) отмечено присутствие Campylobacterota.

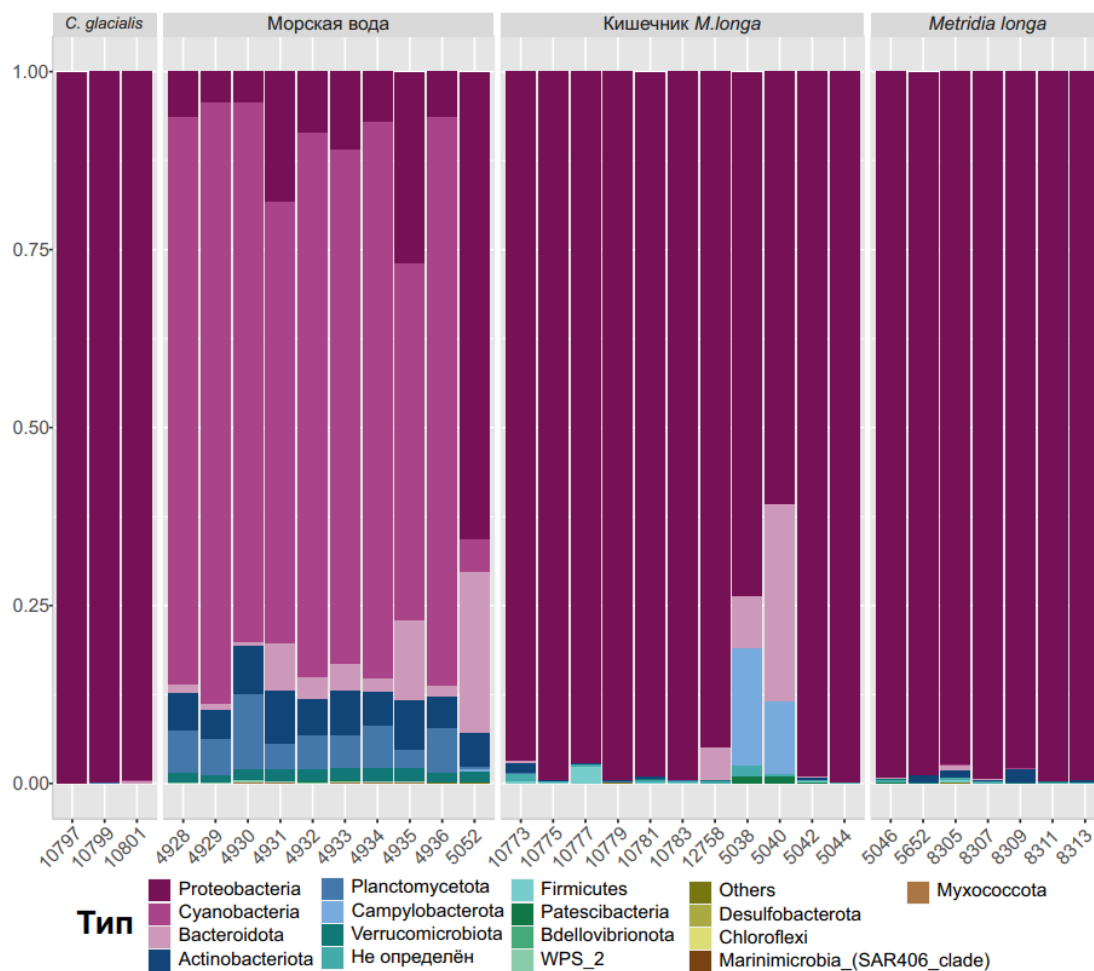


Рисунок 17. Столбчатые диаграммы, отражающие относительные частоты встречаемости типов бактерий в образцах. Диаграммы сгруппированы в зависимости от типа образца.

Анализ разнообразия микробиомов на уровне семейств показал, что в образцах ракообразных наиболее представленными являются семейства *Pseudoalteromonadaceae*, *Moraxellaceae* и *Vibrionaceae* (все принадлежат типу *Proteobacteria*) (Рисунок 18). Среди исследованных образцов микробиом кишечника *Metridia longa* демонстрирует наибольшее разнообразие семейств, в частности, здесь дополнительно обнаруживаются представители *Flavobacteriaceae* (тип *Bacteroidota*), *SAR86* (тип *Proteobacteria*), *Arcobacteraceae* (тип *Campylobacterota*). Доминирующими представителями цианобактерий морской воды определяется *Cyanobiaceae*. Также в образцах морской воды обнаружены семейства “Не определено” и “Митохондрия”. “Не определено” идентифицировано как последовательность гена 16S рРНК хлоропласта. Данные таксономические группы, очевидно, не имеют отношения к разнообразию микробного сообщества.

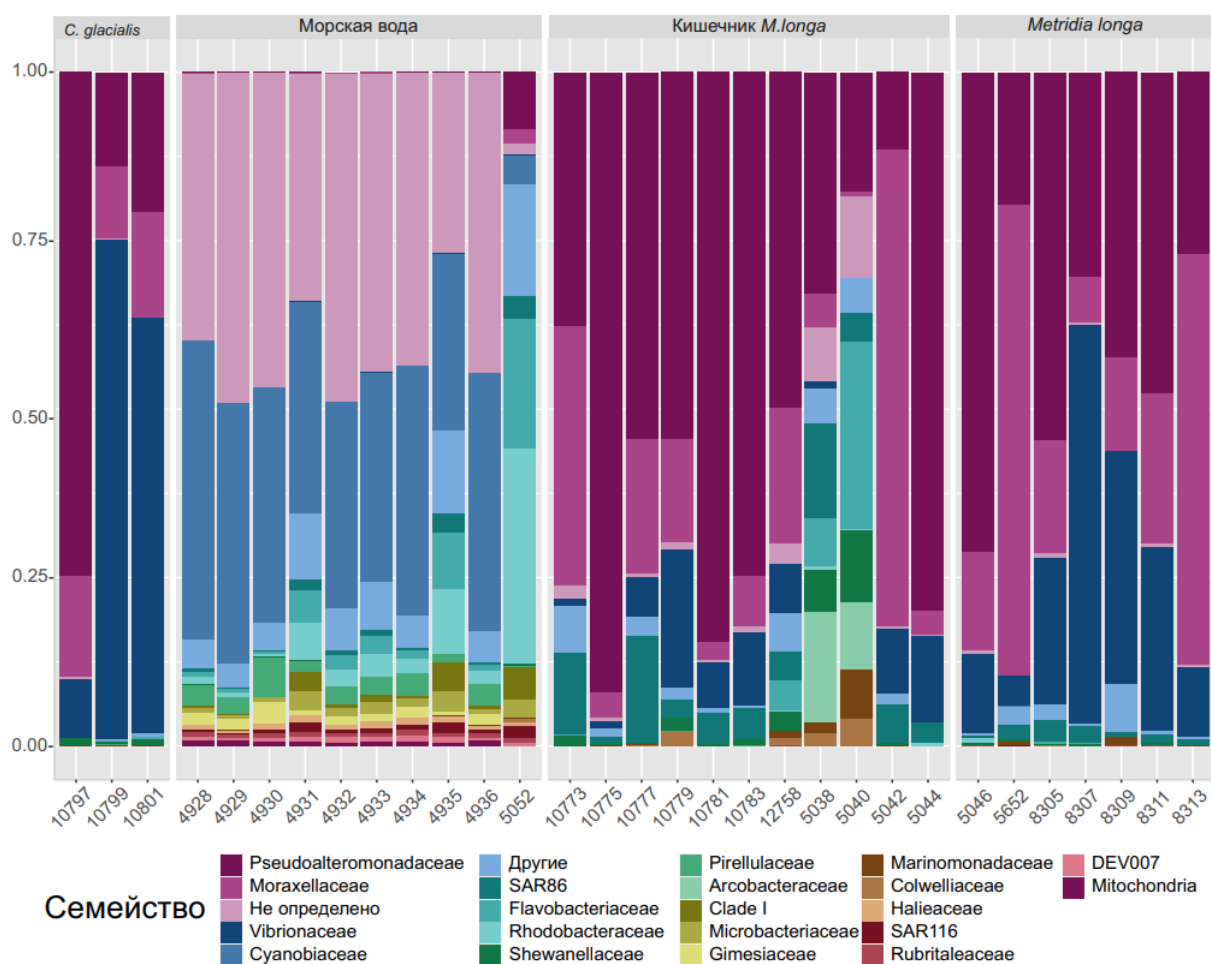


Рисунок 18. Столбчатые диаграммы, отражающие относительные частоты встречаемости семейств бактерий в образцах. Диаграммы сгруппированы в зависимости от типа образца.

4.2.3. Оценка альфа-разнообразия

Для оценки разнообразия исследуемых микробиомов были применены метрики альфа-разнообразия на уровне OTU с использованием индексов Чао1 и Шэннона. Значения индекса Чао1 для всех образцов приведены в таблице (Таблица 4).

Таблица 4. Значение индекса Чао1 для всех исследуемых образцов.

Тип образца	ID	Значение индекса Чао1
<i>Calanus glacialis</i>	10797	37
<i>Calanus glacialis</i>	10799	38
<i>Calanus glacialis</i>	10801	27
Морская вода	4928	304

Морская вода	4929	277
Морская вода	4930	226
Морская вода	4931	322
Морская вода	4932	288
Морская вода	4933	292
Морская вода	4934	221
Морская вода	4935	312
Морская вода	4936	278
Морская вода	5052	201
Кишечник <i>Metridia longa</i>	5038	69
Кишечник <i>Metridia longa</i>	5040	49
Кишечник <i>Metridia longa</i>	5042	52
Кишечник <i>Metridia longa</i>	5044	30
Кишечник <i>Metridia longa</i>	10773	58
Кишечник <i>Metridia longa</i>	10775	54
Кишечник <i>Metridia longa</i>	10777	71
Кишечник <i>Metridia longa</i>	10779	71
Кишечник <i>Metridia longa</i>	10781	45
Кишечник <i>Metridia longa</i>	10783	52
Кишечник <i>Metridia longa</i>	12758	33
<i>Metridia longa</i>	5046	72

<i>Metridia longa</i>	5652	33
<i>Metridia longa</i>	8305	56
<i>Metridia longa</i>	8307	42
<i>Metridia longa</i>	8309	19
<i>Metridia longa</i>	8311	43
<i>Metridia longa</i>	8313	49

Для *Calanus glacialis* среднее значение индекса Чao1 составило 34 со стандартным отклонением 6,1; для *Metridia longa* среднее значение составило 44,9 со стандартным отклонением 16,8; для кишечников *Metridia longa* среднее значение составило 53,1 со стандартным отклонением 13,9; для морской воды среднее значение составило 272,1 со стандартным отклонением 41,6. Высокий индекс разнообразия микробиома морской воды подтверждает предположение о большем таксономическом разнообразии данного сообщества по сравнению с микробиомами копепод (р-значение $1,2 \cdot 10^{-9}$, Таблица 5, Рисунок 19). Среди исследованных микробиомов наименьшее разнообразие отмечается у *Calanus glacialis*. Примечательно, что альфа-разнообразие микробиома кишечников *Metridia longa* незначительно больше, чем микробиом целых *Metridia longa* (р-значение 0,29, Таблица 5), но выше, чем у *Calanus glacialis* (р-значение 0,01, Таблица 5).

Таблица 5. Результаты t-теста и р-значения для попарного сравнения образцов некоторых типов на основе рассчитанного индекса Чao1.

Пары типов образцов	Значение t-теста	р-значение
Морская вода и все образцы ракообразных	-20,2	$1,25 \cdot 10^{-9}$
Кишечник <i>Metridia longa</i> и целые <i>Metridia longa</i>	1,12	0,29
Целые <i>Metridia longa</i> и <i>Calanus glacialis</i>	-1,87	0,01

Боксплоты, построенные на основе рассчитанного индекса Чao1 визуально демонстрируют существенное различие альфа-разнообразия между образцами морской воды и образцами ракообразных, тогда как сами образцы ракообразных между собой достаточно схожи (Рисунок 19).

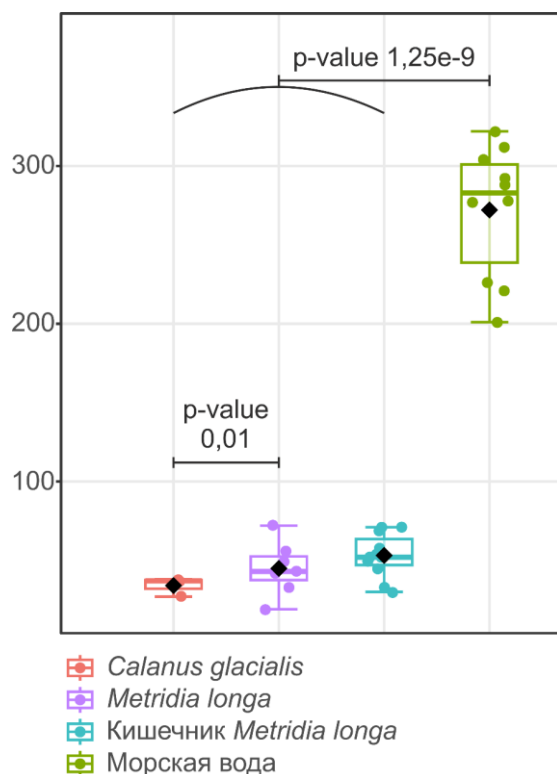


Рисунок 19. Боксплоты, отражающие альфа-разнообразия микробных сообществ разных типов образцов. Оценка альфа-разнообразия осуществлена на основе индекса Чao1. Точки соответствуют исследуемым образцам. Образцы сгруппированы в зависимости от типа.

Значения индекса Шэннона для всех образцов приведены в таблице (Таблица 6).

Таблица 6. Значение индекса Шэннона для всех исследуемых образцов.

Тип образца	ID образца	Значение индекса Шэннона
<i>Calanus glacialis</i>	10797	1,8
<i>Calanus glacialis</i>	10799	1,7
<i>Calanus glacialis</i>	10801	1,9
Морская вода	4928	2,8
Морская вода	4929	2,5

Морская вода	4930	2,7
Морская вода	4931	3,5
Морская вода	4932	3,0
Морская вода	4933	3,1
Морская вода	4934	2,8
Морская вода	4935	3,7
Морская вода	4936	2,7
Морская вода	5052	3,9
Кишечник <i>Metridia longa</i>	5038	2,9
Кишечник <i>Metridia longa</i>	5040	2,7
Кишечник <i>Metridia longa</i>	5042	1,9
Кишечник <i>Metridia longa</i>	5044	1,1
Кишечник <i>Metridia longa</i>	10773	2,5
Кишечник <i>Metridia longa</i>	10775	1,2
Кишечник <i>Metridia longa</i>	10777	2,4
Кишечник <i>Metridia longa</i>	10779	2,3
Кишечник <i>Metridia longa</i>	10781	1,4
Кишечник <i>Metridia longa</i>	10783	1,9
Кишечник <i>Metridia longa</i>	12758	2,3

<i>Metridia longa</i>	5046	1,7
<i>Metridia longa</i>	5652	1,9
<i>Metridia longa</i>	8305	2,6
<i>Metridia longa</i>	8307	1,7
<i>Metridia longa</i>	8309	2,3
<i>Metridia longa</i>	8311	2,3
<i>Metridia longa</i>	8313	2,1

Для *C. glacialis* среднее значение составило 1,8 со стандартным отклонением 0,1; для кишечников *M. longa* среднее значение составило 2,1 со стандартным отклонением 0,6; для *M. longa* среднее значение составило 2,1 со стандартным отклонением 0,3; для морской воды среднее значение составило 3,1 со стандартным отклонением 0,5. Статистический анализ попарного сравнения индекса Шэннона образцов, сгруппированных по типам, выявил закономерности схожие полученным ранее для индекса Чао1. Достоверно более разнообразными сообществами были сообщества морской воды (p -значение $2,21 \cdot 10^{-5}$, Таблица 7, Рисунок 20) Микробиомы целых *M. longa* не отличались по разнообразию от микробиомов кишечников *M. longa* (p -значение 0,90, Таблица 7, Рисунок 20). Микробиомы *M. longa* были более разнообразными чем микробиомы *C. glacialis* (p -значение 0,04, Таблица 7, Рисунок 20).

Таблица 7. Результаты t-теста и p -значения для попарного сравнения образцов некоторых типов на основе рассчитанного индекса Шеннона.

Пары типов образцов	Значение t-теста	p -значение
Морская вода и все образцы ракообразных	-5.87	$2,21 \cdot 10^{-5}$
Кишечник <i>Metridia longa</i> и целые <i>Metridia longa</i>	-0.12	0,90
Целые <i>Metridia</i>	-2.47	0,04

<i>longa</i> и <i>Calanus glacialis</i>		
---	--	--

Боксплоты, построенные на основе рассчитанных индексов Шэннона (Рисунок 20), как и на основе индекса Чao1, показывают большее альфа-разнообразие микробиома морской воды по сравнению с остальными образцами. Также мы наблюдаем самое скудное сообщество у *Calanus glacialis*. Разница между нижним и верхним квартилем боксплотов для кишечников *Metridia longa* и целых *Metridia longa* больше, чем таковые на иллюстрации для альфа-разнообразия, рассчитанного индексом Чao1.

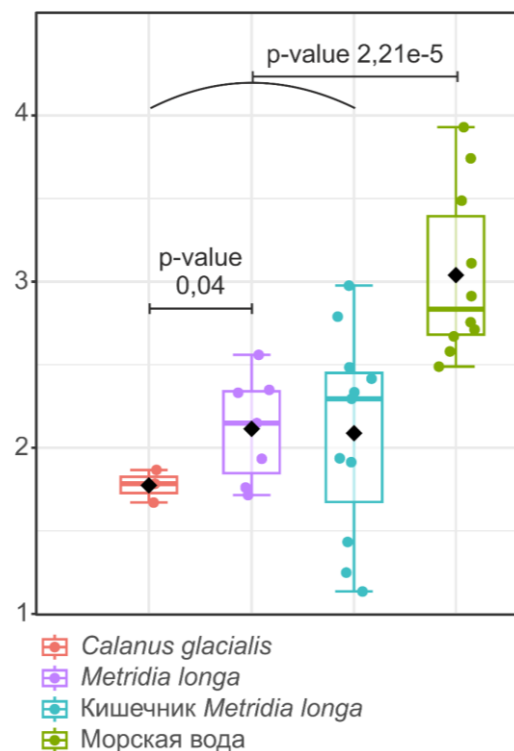


Рисунок 20. Боксплоты, отражающие альфа-разнообразие микробных сообществ разных типов образцов. Оценка альфа-разнообразия осуществлена на основе индекса Шеннона. Точки соответствуют исследуемым образцам. Образцы сгруппированы в зависимости от типа.

4.2.4. Анализ бета-разнообразия микробиомов

Согласно анализу главных координат, рассчитанных по несходствам Брея-Кёртиса по данным разнообразия микробиомов, прошедших фильтрацию от негативных контролей, вычитание загрязнения не повлияло значительно на сходство и различия между образцами (Рисунок 21). Образцы воды сгруппированы отдельно от образцов ракообразных, что подтверждает различия между микробиомами. Микробиомы

ракообразных же группируются вместе, что свидетельствует об их сходстве. Также был проведён PERMANOVA для некоторых пар типов образцов: 1) образцы воды и все образцы ракообразных, p -значение = 0.001, 2) образцы кишечника *Metridia longa* и целые *Metridia longa*, p -значение = 0.038, 3) образцы целых *Metridia longa* и *Calanus glacialis*, p -значение = 0.158. Согласно проведённому статистическому анализу, можно отметить разницу между таксономическим разнообразием эндо- и эпимикробиоты *Metridia longa*. В целом, таксономическое разнообразие микроорганизмов разных видов ракообразных, *Metridia longa* и *Calanus glacialis*, по-видимому, не имеет существенных различий. Данные виды хотя и демонстрируют некоторые различия в характере питания, однако являются симпатрическими. Данное наблюдение соотносится с наблюдениями о том, что коровый микробим копепод зависит, в первую очередь, от их места обитания (Moisander et al., 2015). Опубликованный мета-анализ разнообразия микробиомов, ассоциированных с разными видами копепод, продемонстрировал, что микробиомы *Calanus* (сем. *Calanidae*) и *Pleuromamma* (относятся к сем. *Metridinidae*, как и типовой род *Metridia*) значительно различаются (Sadaiappan, 2021), что противоречит полученным нами данным и данным Moisander и коллег. Вероятная причина различий в том, что представители *Calanus* и *Pleuromamma* были собраны в разных географических точках (*Calanus* были собраны в Атлантическом океане к западу от Британии и в Норвежском море, тогда как *Pleuromamma* были получены из Саргассова моря) и изучены в разных лабораториях.

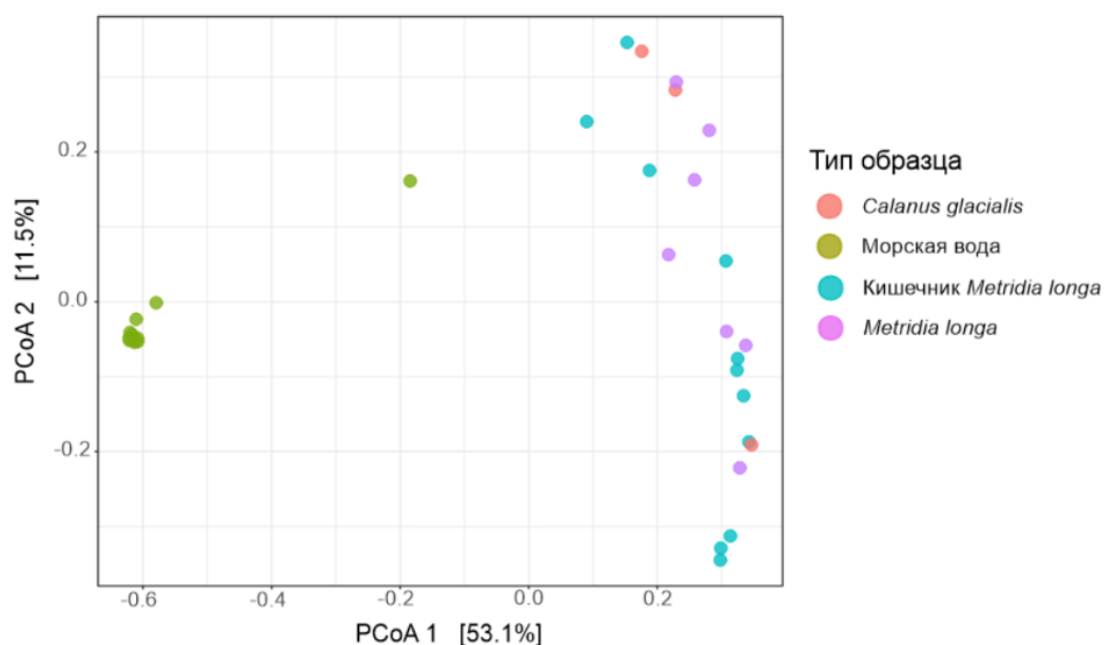


Рисунок 21. График анализа главных координат, рассчитанных по несходствам Брея-Кёртиса, для разных типов исследуемых образцов. Каждая точка соответствует одному образцу.

4.2.5. Дифференциальный анализ микробных сообществ

Для более глубокого исследования различий между микробиомами разных типов образцов, был проведен дифференциальный анализ микробных сообществ с помощью пакета LEFse в онлайн-сервисе MicrobiomeAnalyst.

Дифференциальный анализ микробных сообществ всех образцов ракообразных (*M. longa* и *C. glacialis*) и образцов морской воды вновь продемонстрировал ключевое отличие морской воды, заключающееся в избытии представителей цианобактерий, в частности, таких семейств как *Cyanobiaceae* (LDA = -6.2) и *Rhodobacteraceae* (LDA = -5.46) (Рисунок 22). Аналогичная картина для представителей Planctomycetota (семейство *Pirellulaceae*, LDA = -5.12; *Gimesiaceae*, LDA = -4.86), представителей Bacteroidota (семейство *Flavobacteriaceae*, LDA = -5.04), Actinobacteriota (семейство *Microbacteriaceae*, LDA = -4.86). “Не определено” идентифицировано как последовательность гена *16S rRNA* хлоропласта. Данные таксономические группы, очевидно, не имеют отношения к разнообразию микробного сообщества. В микробиомах веслоногих раков обогащены семейства филума Proteobacteria класса Gammaproteobacteria, такие как *Pseudoalteromonadaceae*, *Moraxellaceae*, *Vibrionaceae* и

Shewanellaceae, к которым, вероятно, относятся представители корового микробиома изучаемых ракообразных.

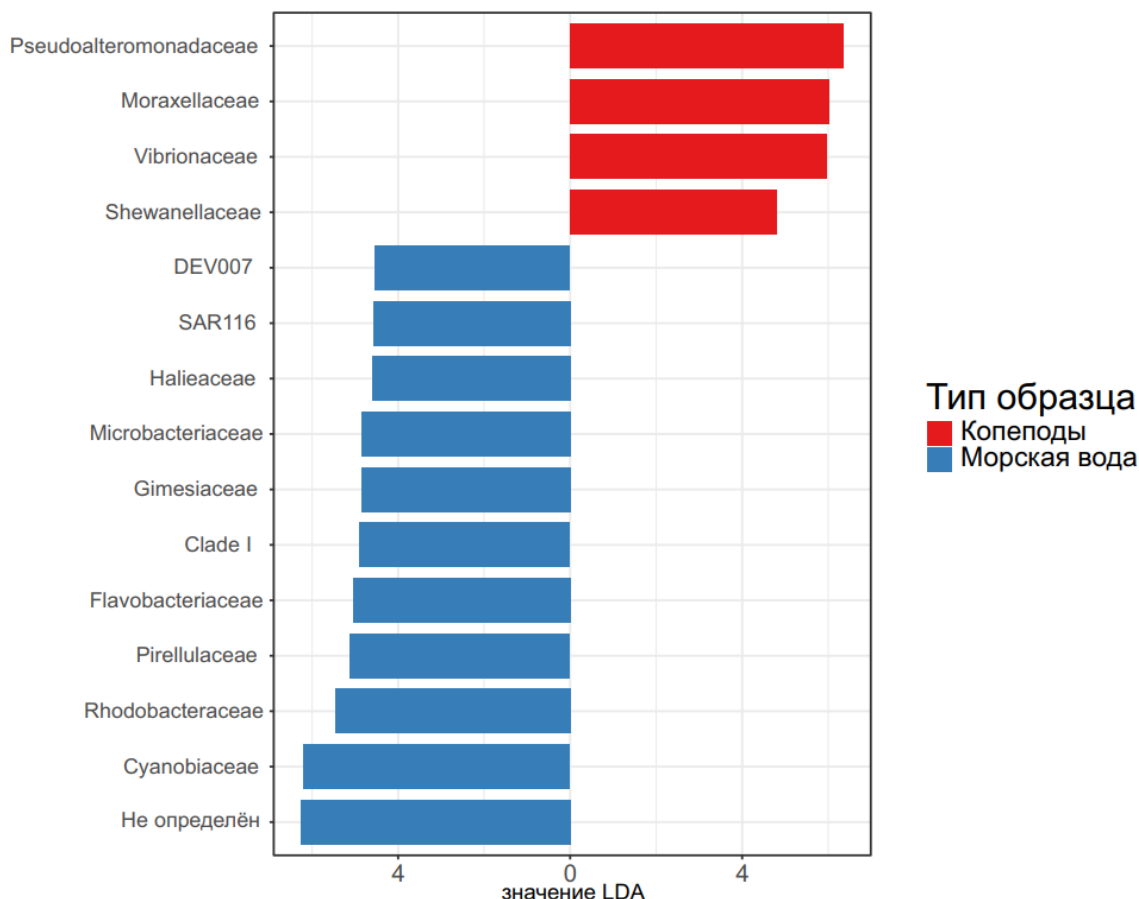


Рисунок 22. Гистограммы LDA значений для наиболее сильно различающихся по частоте семейств бактерий между микробиомами веслоногих ракообразных и микробиомами морской воды. Красные столбики - семейства бактерий обогащенные в микробиоме веслоногих ракообразных, синие столбики - семейства бактерий обогащенные в микробиоме морской воды.

Сравнение состава микробиомов целых *Metridia longa* и образцов кишечника *Metridia longa* с выявил таксоны микроорганизмов, которые наиболее сильно различаются по частоте между сообществами эпи- и эндомикробиоты (Рисунок 23). В образцах кишечника *M. longa* наблюдалось гораздо более высокая частота представителей семейств *Shewanellaceae* (LDA = 5,12) и *SAR86* (LDA = 5,36), а в образцах целых веслоногих рачков — семейств *Cyanobiaceae* (LDA = -3,05) и *Vibrionaceae* (LDA = -5,93). Подобные различия можно объяснить тем, что некоторые представители *Vibrionaceae* проявляют хитиновую активность, следовательно, можно предположить, что хитиновые кутикулы ракообразных могут представлять для них привлекательную экологическую нишу (Suginta et al., 2000). Различные цианобактерии, как уже было

показано, в изобилии обнаруживаются в морской воде. Они могут осаждаться на поверхность *M. longa* и образовывать микроколонии. Бактерии *Shewanellaceae* довольно широко распространены в морских экосистемах и, в частности, уже было показано их присутствие в микробиоме кишечника рыб Белого моря (Burtseva et al., 2021). Представители некультивируемого SAR86 также демонстрируют повсеместное распространение в океанических водах (Hoarfrost et al, 2020). Тем не менее, обогащенные в микробиомах кишечника представители этих семейств могут быть потенциальными симбионтами копепод.

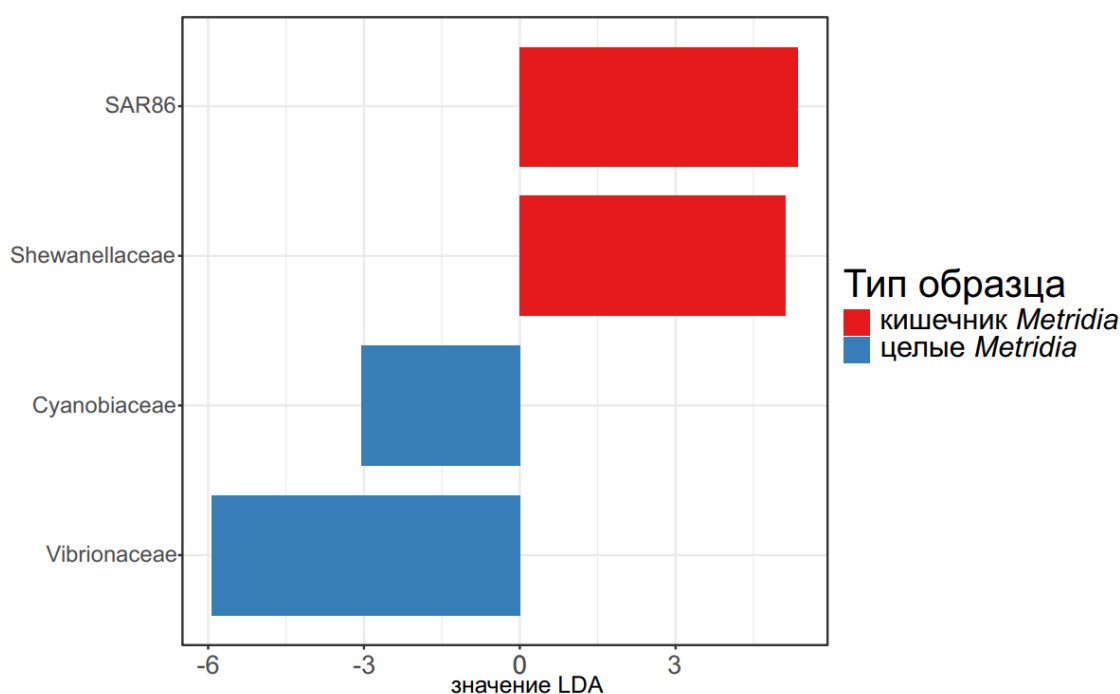


Рисунок 23. Гистограммы LDA значений для наиболее сильно различающихся по частоте семейств бактерий между микробиомами кишечника *M. longa* и микробиомами целых представителей этого вида ракообразных. Красные столбики - семейства бактерий обогащенные в кишечном микробиоме, синие столбики - семейства бактерий обогащенные в микробиоме целых веслоногих ракообразных.

Несмотря на слабые различия микробиомов двух видов копепод (согласно PERMANOVA и сравнению индексов альфа-разнообразия), анализ с помощью LEFse выявил четыре семейства, различающиеся по частоте между микробиомами (Рисунок 24). В микробиомах представителей *C. glacialis* были более представлены *Shewanellaceae* и *Alteromonadaceae*, тогда как в *M. longa* более массовыми были представители SAR86 и *Corynebacteriaceae*. Обнаруженные тонкие отличия, вероятно, отражают разницу в составах коровых микробиомов данных видов копепод и/или различия диеты этих организмов.

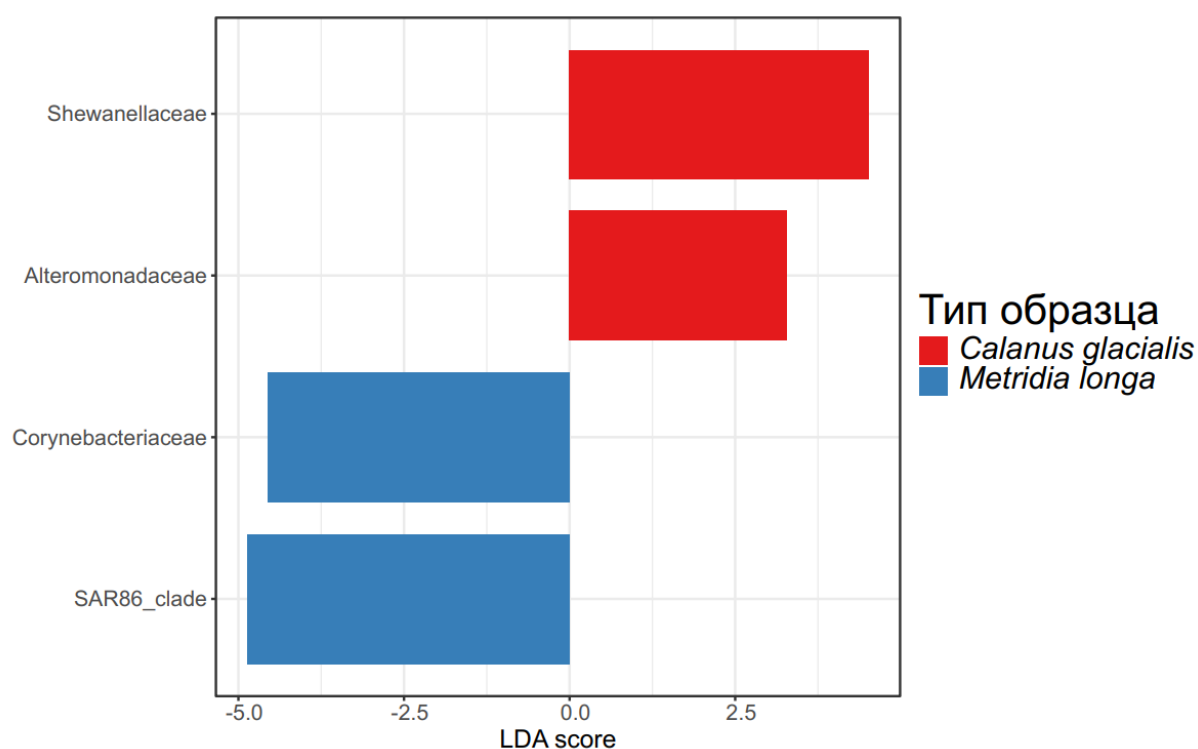


Рисунок 24. Гистограммы LDA значений для наиболее сильно различающихся по частоте семейств бактерий между микробиомами *C. glacialis* и *M. longa*. Красные столбики - семейства бактерий обогащенные в микробиоме *C. glacialis*, синие столбики - семейства бактерий обогащенные в микробиоме *M. longa*.

4.3. Анализ микробиома на основе данных шот-ган метагеномного секвенирования

4.3.1 Сборка и аннотация шот-ган метагеномов

Для поиска кластеров генов биосинтеза были подготовлены и секвенированы шот-ган библиотеки метагеномной ДНК *M. longa* и бактериальной фракции морской воды. После удаления адаптерных последовательностей, фильтрации прочтений по качеству и по длине, в метагеноме *M. longa* осталось 189 млн прочтений и в метагеноме воды - 165 млн прочтений, с которыми проводили сборки *de novo* при помощи metaSPAdes (Таблица 8). После сборок контиги короче 1000 нт были исключены из последующего анализа как слишком короткие и не несущие информации о кластерах генах (кластеры генов биосинтеза могут простираться на десятки тысяч нуклеотидов). В оставшихся контигах были предсказаны открытые рамки считывания (ОРС) (см. подраздел 3.5 в Материалах и Методах): 253 тысячи ОРС было обнаружено в метагеноме

M. longa и 947 тысяч ОРС в метагеноме морской воды (Таблица 8). Несмотря на меньшее количество полученных прочтений, метагеном воды был собран с большей непрерывностью (N50 1249 нт против 739 нт для *M. longa*): для него значительно большее количество контиг было длинее порога в 1000 нт и в нем было обнаружено в 3,7 раза больше ОРС. Причина подобных различий, вероятно, связана с высоким уровнем содержания эукариотической ДНК ракообразного в метагеноме *M. longa*, что усложнило сборку метагенома. В данном случае увеличение непрерывности сборок возможно при помощи скэффолдинга или гибридной сборки с использованием длинных прочтений (полученных по технологии PacBio или Oxford Nanopore Technologies), что планируется сделать в дальнейшем.

Таблица 8. Характеристики сборки шот-ган метагеномов *M. longa* и морской воды.

Образец	Количество ридов исходно	Количество ридов после фильтрации	Количество контиг	Суммарная длина сборки, Мб	N50, нт	Количество контиг > 1000 нт	Суммарная длина сборки, Мб (контиги > 1000 нт)	Количество ORF
<i>M. longa</i>	189354764	147972344	3355732	1026	739	190540	281	252639
Морская вода	165113006	129428125	2500422	1302	1249	301983	759	947543

4.3.2. Кластеры генов биосинтеза вторичных метаболитов, обнаруженные в образце *Metridia longa* и в образце морской воды

В образце *Metridia longa* (5650) при помощи antiSMASH были предсказаны всего 9 кластеров генов биосинтеза вторичных метаболитов (Рисунок 25). Были определены типы потенциально производимых из них продуктов. Трое из них идентифицированы как терпены, два как похожие на рибосомально синтезированные и пост-трансляционно модифицированные пептиды (RiPPs-like), один как гомосерин лактон, один как арилполиен, один как беталактон и один как редокс-кофактор. Кластер арилполиена показал 15% сходство с кластером биосинтеза арилополиена Vf, описанного из штамма *Aliivibrio fischeri ES114* (Cimermanic et al., 2014). Кластер беталактона показал 13% сходство с кластером биосинтеза НРП фенгицина, описанного из штамма *Bacillus velezensis FZB42* (Koumoutsis et al., 2004). Кластер редокс-кофактора показал 13% сходство с кластером биосинтеза ланкацидина С, продукта гибридного нерибосомного и поликетидного синтеза, описанного у *Streptomyces rochei* (Mochizuki et al., 2003).

Исходный дизайн эксперимента предполагал использование предиктора NERPA, предназначенного для поиска кластеров генов, вовлечённых в биосинтез НРП, с использованием структуры целентеразина в формате SMILES и результатов предсказаний antiSMASH. Однако визуальный анализ обнаруженного кластера НРП показал нецелесообразность этого подхода, поскольку в кластере отсутствовала НРП синтаза. Таким образом, нам не удалось обнаружить биосинтетические кластеры НРП, которые могли бы быть потенциально вовлечены в синтез целентеразина. Вероятно, увеличение непрерывности сборки позволит в дальнейшем провести более полный анализ.

Разнообразие и количество кластеров генов биосинтеза в образце воды (5052 (5650)) оказалась существенно больше и составила 722 штуки (Рисунок 25), что является следствием более высокого качества сборки. Наиболее многочисленными из них оказались кластеры, вовлечённые в продукцию терпенов (242 шт.), многие из которых определены как каротиноиды. Такая картина сходится с нашими данными по профилированию таксономического состава образцов морской воды и ракообразных. Микробное сообщество морской воды продемонстрировало существенно большее таксономическое разнообразие, где доминирующими представителями установлены *Cyanobacteria*. Стоит отметить, что цианобактерии используют каротиноиды для функционирования фотосистемы II (Zakar et al., 2016). Среди 49 обнаруженных НРП кластеров, содержащих гены НРП-синтазы, 14 содержали полные последовательности этих генов и 35 кластеров содержали фрагменты генов НРП-синтазы. Ни одна из обнаруженных НРП-синтаз, однако, не обладала предсказанной специфичностью доменов к аминокислотам, вероятно, входящим в состав целентеразина.



Рисунок 25. Столбчатые диаграммы, отражающие разнообразие разных типов кластеров генов биосинтеза вторичных метаболитов в образце морской воды и образце *Metridia longa*.

4.3.3. Кандидатные гены биосинтеза целентеразина на основе данных предсказания открытых рамок считывания

В сборке метагенома *Metridia longa* были обнаружены 1953 предсказанных открытых рамок считывания, чьи продукты содержали мотив FYY в любом положении, среди которых 32 ОРС содержали искомый мотив на С-конце. Для данного набора ОРС был проведен поиск гомологов в базе данных UniProtKB. Интересно, что для двух последовательностей среди гомологов были обнаружены оксигеназы и изопенициллин-N-синтазы, для которых ранее была показана ассоциация с люминесцирующими гребневиками (Francis et al., 2015) (Таблица 9). Интересно, что ранее изопенициллин-N-синтазы, содержащие FYY мотив, не были обнаружены в транскриптоме *Metridia lucens* (Tessler et al., 2018), что может свидетельствовать либо о низком покрытии транскриптома, либо о кодировании данного гена в микробиоме веслоного ракообразного.

Таблица 9. Описание гомологов последовательностей ОРС, обнаруженных в сборке метагенома *M. longa*, содержащих С-концевой мотив FYY, обнаруженных в базе данных UniProtKB. Жирным шрифтом выделены хиты, относящиеся к оксигеназам и изопенициллин-N-синтазам.

Название ОРС	Хиты в UniProtKB
>5032_13144_c5_1 5032_13144_c5 1432 2403 + gene_type=archaea	Copper transporter, Fe2OG dioxygenase domain-containing protein, Isopenicillin N synthase-like Fe(2+) 2OG dioxygenase domain-containing protein
>13042_12311_c18_2 13042_12311_c18 818 952 + gene_type=archaea	Simple sugar transport system permease protein, non-specific serine/threonine protein kinase
>35130_11700_c6_2 35130_11700_c6 1478 1699 - gene_type=archaea partial=01	Zinc finger protein CCCH domain-containing protein, putative, Srp40 C-terminal domain-containing protein
>37741_11664_c1_1 37741_11664_c1 1 1485 + gene_type=archaea partial=10	MULE transposase domain-containing protein, Type 4 fimbrial biogenesis protein PilX N-terminal domain-containing protein, leucine--tRNA ligase
>40451_11630_c2_1 40451_11630_c2 1 159 + gene_type=bacteria partial=10	Golgi to ER traffic protein 4 homolog
>42504_11605_c1_1 42504_11605_c1 1 120 + gene_type=bacteria partial=10	-
>53028_11501_c9_1 53028_11501_c9 619 834 + gene_type=archaea	asparaginase, Rho GTPase-activating protein 21, Leucine-rich repeat protein
>62994_11425_c1_2 62994_11425_c1 536 1126 + gene_type=archaea	Transmembrane protein 185A, Mitochondrial-processing peptidase subunit beta,
>64030_11417_c7_1 64030_11417_c7 3 95 + gene_type=bacteria partial=10	-
>66196_11403_c2_2 66196_11403_c2 919 1401 + gene_type=archaea partial=01	Hypp6972 protein, Reverse transcriptase domain-containing protein, CRB1 protein
>68323_11389_c2_1 68323_11389_c2 1 144 + gene_type=archaea partial=10	Thiol-disulfide isomerase-like thioredoxin, G-protein coupled receptors family 1 profile domain-containing protein, DUF6533 domain-containing protein
>69764_11380_c7_1 69764_11380_c7 3 1238 - gene_type=bacteria partial=10	Reverse transcriptase domain-containing protein, ribonuclease H, Putative RNA-directed DNA polymerase from transposon X-element
>73358_11359_c1_1 73358_11359_c1 1228 1359	Lysophospholipase L1-like esterase, Adhesion G-protein coupled receptor V1, G-protein coupled receptor 98
>76170_11342_c7_2 76170_11342_c7 688 1206 + gene_type=archaea	Putative splicing factor, arginine/serine-rich 7-like Protein, RRM domain-containing protein, Peptidyl-prolyl cis-trans isomerase G isoform X4
>82519_11308_c2_2 82519_11308_c2 1141 1308 - gene_type=archaea partial=01	Kinase-like protein, Addiction module toxin RelE, THH1/TOM1/TOM3 domain-containing protein
>89511_11275_c2_1 89511_11275_c2 1179 1274 - gene_type=bacteria partial=01	-
>90025_11273_c1_1 90025_11273_c1 3 95 + gene_type=bacteria partial=10	-
>99253_11234_c1_1 99253_11234_c1 1136 1234 - gene_type=archaea partial=01	-
>104774_11212_c15_1 104774_11212_c15 3 905 + gene_type=archaea partial=10	Copper transporter, Fe2OG dioxygenase domain-containing protein, Isopenicillin N synthase-like Fe(2+) 2OG dioxygenase domain-containing protein

>114062_11180_c2_1 114062_11180_c2 1 126 + gene_type=archaea partial=10	NAD(P)H-quinone oxidoreductase subunit 2 b chloroplastic,
>134676_11119_c1_2 134676_11119_c1 489 719 + gene_type=archaea	SEC7 domain-containing protein, Regulatory protein SoxS, Fork-head domain-containing protein
>137429_11111_c10_2 137429_11111_c10 204 986 - gene_type=bacteria	Vang-like protein
>140862_11103_c1_3 140862_11103_c1 996 1103 - gene_type=archaea partial=01	-
>141833_11100_c2_2 141833_11100_c2 565 1098 + gene_type=bacteria partial=01	NADH dehydrogenase subunit 5, O7E24 protein, G-protein coupled receptors family 1 profile domain-containing protein
>142617_11098_c2_1 142617_11098_c2 2 121 + gene_type=bacteria partial=10	SCY1 like pseudokinase 1, RING-H2 finger protein ATL70
>144687_11093_c2_1 144687_11093_c2 877 1092 - gene_type=archaea partial=01	Cytochrome c domain-containing protein, UV radiation resistance protein and autophagy-related subunit 14, FecR protein
>145274_11092_c1_1 145274_11092_c1 3 170 + gene_type=bacteria partial=10	AraC family transcriptional regulator, Lysosomal cobalamin transporter
>148327_11084_c7_1 148327_11084_c7 2 127 + gene_type=bacteria partial=10	protein-glutamate O-methyltransferase
>153389_11073_c1_1 153389_11073_c1 310 1071 - gene_type=bacteria partial=01	Muscle-specific protein 300 kDa, Serine/threonine-protein kinase Nek1-like isoform X6, Nesprin-1
>155655_11068_c1_2 155655_11068_c1 788 1066 - gene_type=archaea partial=01	Allantoin permease, Transporter component, Spore germination protein
>184525_11010_c2_1 184525_11010_c2 3 137 - gene_type=archaea partial=10	Cytochrome c domain-containing protein, Permease, Cholesterol oxidase
>188340_11003_c4_1 188340_11003_c4 1 90 - gene_type=bacteria partial=10	Reverse transcriptase, Reverse transcriptase/retrotransposon-derived protein RNase H-like domain-containing protein

В сборке метагенома морской воды были обнаружены 9812 ОРС, чьи продукты содержали мотив FYY в любом положении, среди которых всего 33 ОРС содержали искомый мотив на С-конце. Среди гомологов данных ОРС, в отличие от ОРС из *M. longa*, не были обнаружены оксигеназы и/или изопенициллин-N-синтазы (Таблица 10).

Таблица 10. Описание гомологов последовательностей ОРС, обнаруженных в сборке метагенома морской воды, содержащих С-концевой мотив FYY, обнаруженных в базе данных UniProtKB.

Название ОРС	Хиты в UniProtKB
>62_1173084_c34_21 62_1173084_c34 21003 22277 + gene_type=bacteria	Serine hydroxymethyltransferase
>1546_126642_c6_11 1546_126642_c6 8399 9241 - gene_type=bacteria	Iron permease FTR1 family protein
>2033_122568_c6_14 2033_122568_c6 15044 15673 + gene_type=bacteria	Lysine transporter LysE, Threonine/homoserine/homoserine lactone efflux protein
>3934_115176_c4_4 3934_115176_c4 3004 3933 + gene_type=bacteria	Phosphatidate cytidylyltransferase,

>4457_114033_c35_15 4457_114033_c35 11858 13366 - gene_type=bacteria	Flagellar hook-associated protein 2,
>15162_16168_c5_5 15162_16168_c5 2794 3636 - gene_type=bacteria	Iron permease FTR1 family protein
>16467_15832_c30_7 16467_15832_c30 4920 5831 - gene_type=bacteria partial=01	Iron permease FTR1 family protein
>17554_15596_c9_5 17554_15596_c9 4313 5164 - gene_type=bacteria	CG14292-PA, Lipoprotein, TRAP transporter large permease protein
>36455_13478_c5_1 36455_13478_c5 571 2058 - gene_type=bacteria	Membrane protein (DUF2207), Putative membrane protein, CorA-like Mg ²⁺ transporter protein
>42295_13171_c2_2 42295_13171_c2 158 916 - gene_type=bacteria	Outer membrane protein beta-barrel domain-containing protein, PorT family protein
>45952_13014_c4_4 45952_13014_c4 2709 3014 + gene_type=bacteria partial=01	ATP-dependent helicase, DNA 3'-5' helicase, UvrD-like helicase ATP-binding domain-containing protein
>49059_12897_c47_2 49059_12897_c47 790 1146 - gene_type=bacteria	Snoal-like domain-containing protein, Nuclear transport factor 2 family protein, Ester cyclase
>59301_12588_c2_3 59301_12588_c2 2078 2587 - gene_type=bacteria partial=01	Alpha-E domain-containing protein, DUF403 domain-containing protein
>59325_12587_c3_1 59325_12587_c3 1 1296 - gene_type=bacteria partial=10	Membrane protein, Nucleoside transporter, Purine-cytosine permease
>60824_12550_c4_2 60824_12550_c4 226 768 - gene_type=bacteria	Apoptotic chromatin condensation inducer in the nucleus-like isoform X2, Eukaryotic translation initiation factor 5B, mRNA export factor GLE1
>86958_12059_c4_1 86958_12059_c4 73 621 + gene_type=bacteria	DUF4402 domain-containing protein
>87582_12050_c3_2 87582_12050_c3 688 1479 - gene_type=archaea	Peptidase M14 carboxypeptidase A domain-containing protein, GTPase Der, DNA-directed RNA polymerase
>91235_12001_c2_1 91235_12001_c2 1 477 - gene_type=bacteria partial=10	Peptide ABC transporter permease, ABC transporter permease subunit, Binding-protein-dependent transporter inner membrane component
>91831_11993_c2_1 91831_11993_c2 1 492 - gene_type=archaea partial=10	UDP-3-O-(3-hydroxymyristoyl) glucosamine N-acyltransferase, Probable UDP-N-acetylglucosamine acyltransferase
>142306_11543_c3_2 142306_11543_c3 683 1543 - gene_type=bacteria partial=01	Iron permease FTR1 family protein
>144061_11532_c2_1 144061_11532_c2 1 540 + gene_type=bacteria partial=10	F5/8 type C domain-containing protein, Glucoamylase and related glycosyl hydrolases, Coagulation factor 5/8 type domain-containing protein
>146548_11517_c2_1 146548_11517_c2 2 1516 + gene_type=bacteria partial=11	PKD domain-containing protein, Por secretion system C-terminal sorting domain-containing protein, Glycosyl hydrolase
>152942_11480_c3_1 152942_11480_c3 3 749 - gene_type=bacteria partial=10	Lipoprotein, Calx-beta domain-containing protein, Phosphatidylglycerol lysyltransferase
>162333_11430_c3_1 162333_11430_c3 1 177 + gene_type=archaea partial=10	Transport related membrane protein, Auxin Efflux Carrier, AEC family transporter
>171701_11385_c2_1 171701_11385_c2 1 552 + gene_type=bacteria partial=10	Putative PpiC-type peptidyl-prolyl cis-trans isomerase, Peptidylprolyl isomerase
>206878_11242_c1_2 206878_11242_c1 758 1240 + gene_type=bacteria partial=01	SMP-30/Gluconolactonase/LRE domain protein, Gluconolactonase, Regucalcin
>225433_11181_c3_1 225433_11181_c3 2 493 - gene_type=bacteria partial=10	C1q domain-containing protein, Alpha/beta hydrolase, DUF2793 domain-containing protein
>227258_11176_c2_3 227258_11176_c2 955 1176 + gene_type=bacteria partial=01	NYN domain-containing protein,

>236587_11149_c8_3 236587_11149_c8 656 925 - gene_type=bacteria	DNA primase
>237904_11146_c1_1 237904_11146_c1 1 1014 - gene_type=bacteria partial=10	Sodium-independent anion transporter, Solute carrier 26 family protein, Sulfate transporter
>250424_11112_c4_2 250424_11112_c4 568 1110 - gene_type=bacteria partial=01	CIA30 family NADH ubiquinone oxidoreductase subunit, NADH:ubiquinone oxidoreductase intermediate-associated protein 30 domain-containing protein, Complex I intermediate-associated protein 30 (CIA30)
>274816_11055_c1_2 274816_11055_c1 505 1053 + gene_type=bacteria partial=01	Peptide methionine sulfoxide reductase MsrA,
>285062_11033_c1_1 285062_11033_c1 1 663 - gene_type=bacteria partial=10	Polysaccharide chain length determinant N-terminal domain-containing protein, Lipopolysaccharide biosynthesis protein, Chain length determinant protein

Среди ОРС с С-концевым мотивом FYY можно отметить высокую частоту нахождения гомологичных белков с трансмембранными участками - 7/32 для *M. longa* и 12/33 для морской воды. Вероятно, это связано с гидрофобностью мотива FYY и скоррелированностью данного физического свойства белков с их функцией.

Можно отметить, что доля ОРС с С-концевым мотивом FYY относительно всех ОРС с данным мотивом в метагеноме воды значительно ниже, чем аналогичная доля в метагеноме *M. longa* (0.3% против 1.6% для *M. longa*). При этом, доля FYY-содержащих ОРС относительно всех ОРС примерно одинакова в обоих метагеномах: 0.8% для *M. longa* и 1% для морской воды. Перепредставленность ОРС с С-концевым мотивом в метагеноме *M. longa*, вероятно, отражает высокий уровень эукариотической ДНК, либо связано с биосинтезом целентеразина и биолюминесценцией данного ракообразного.

5. ЗАКЛЮЧЕНИЕ

По результатам работы было получено 40 образцов тотальной ДНК, относящихся к разным типам: кишечник *Metridia longa*, целые особи *Metridia longa*, целые особи *Calanus glacialis* и морская вода. Первичный материал для экстрактов ДНК был собран в сентябре 2023 года в Кандалакшском заливе Белого моря на территории ББС МГУ. Для всех образцов проведено секвенирование ампликонных библиотек переменного участка V3-V4 гена *16S rRNA*. По результатам профилирования микробных сообществ данным методом были отмечены существенные различия в разнообразии и составе микробиоты воды и ракообразных. Доминирующим типом бактерий в морской воде оказались представители Cyanobacteria, в частности, семейство *Cyanobiaceae*, а в образцах ракообразных - представители типа Proteobacteria, в особенности, семейство *Pseudoalteromonadaceae*. Дифференциальный анализ состава сообществ продемонстрировал различия в составе эпи- и эндомикробиоты *Metridia longa*. Представители семейств *Shewanellaceae* и SAR86 были отмечены как бактерии-эпобионты, а представители *Cyanobiaceae* и *Vibrionaceae* как бактерии-эндобионты. Метагеномные образцы *Metridia longa* и морской воды были дополнительно секвенированы с использованием шот-ган подхода для последующего предсказания кластеров генов биосинтеза и открытых рамок считывания. В сборках метагеномов были обнаружены 9 кластеров генов биосинтеза для образца *M. longa* и 722 кластера для образца морской воды. Согласно этому результату, следует отметить, что в целях поиска вторичных метаболитов с потенциальным прикладным применением фильтраты морской воды представляют больший интерес, чем микробиота, ассоциированная с морскими животными. Среди НРП-кластеров из *M. longa* и морской воды не было обнаружено кластеров-кандидатов потенциально способных к синтезу целентеразина. Среди предсказанных в сборках метагеномов рамок считывания был осуществлён поиск последовательностей кодирующих мотив FYY для последующего поиска потенциальных прекурсоров целентеразина. Всего были обнаружены 32 ОРС с С-концевым мотивом FYY для образца *Metridia longa* и 33 для образца морской воды. Среди ОРС из образца *M. longa* были обнаружены гомологи изопенициллин-N-синтазы, которая ранее была выдвинута в качестве потенциального участника и прекурсора в биосинтезе целентеразина в гребневиках.

6. ВЫВОДЫ

1. Был проведен сбор живых представителей *Metridia longa* и *Calanus glacialis*, а также окружающей морской воды в сентябре 2023 г. на ББС МГУ.
2. Были получены экстракты ДНК из собранных образцов. Оптимальным методом получения ДНК из веслоногих ракообразных оказалась хлороформная экстракция с последующей дополнительной очисткой на спин-колонках. Было проведено секвенирование 40 ампликонных библиотек переменного участка V3-V4 гена *16S rRNA* и шот-ган секвенирование 2 метагеномных образцов - *M. longa* и морской воды.
3. Проанализированы результаты секвенирования 40 ампликонных библиотек переменного участка V3-V4 гена *16S rRNA*, выделенных из беломорских копепод *Metridia longa* и *Calanus glacialis*, кишечников *Metridia longa*, воды и негативных контролей лабораторного загрязнения. Были продемонстрированы различия между микробиотой ракообразных и морской воды, а также выявлены бактерии-эпibiонты *M. longa*.
4. Количество кластеров генов биосинтеза (722 шт.) в метагеноме морской воде существенно больше, чем в таковом у *Metridia longa* (9 шт.). Не были обнаружены НРП кластеры биосинтеза, потенциально вовлечённые в биосинтез целентеразина.
5. Среди гомологов открытых рамок считывания, предсказанных в сборке метагенома *Metridia longa*, были обнаружены оксигеназы/ изопенициллин-N-синтазы, потенциально участвующие в биосинтезе целентеразина. В метагеноме морской воды подобного не было обнаружено.

7. БЛАГОДАРНОСТИ

Автор выражает исключительные благодарности научному руководителю, Сутормину Дмитрию Александровичу, за чуткое и терпеливое курирование выпускной квалификационной работы и неоценимую моральную поддержку. Всему коллективу лаборатории молекулярной генетики микроорганизмов Института Биологии Гена РАН, в частности, Павлову Константину Денисовичу, за помощь в проведении экспериментов и моральную поддержку. Сотруднику Института Биоинформатики, Райко Михаилу Петровичу, за освоение онлайн-сервиса MicrobiomeAnalyst. Сотрудникам Центра Морских Исследований МГУ, Коноваловой Ольге Петровне и Юриковой Дарье Андреевне, за помощь в организации экспедиции и сборе материала. А также Константину Викторовичу Северинову за предоставленную возможность.

7. СПИСОК ЛИТЕРАТУРЫ

1. Виды — биологические индикаторы состояния морских арктических экосистем / ПАО «НК «Роснефть», ООО «Арктический Научный Центр», Фонд «НИР». — 2020.
2. Евстигнеев П. В., Битюков Э. П. Билюминесценция морских копепод. — 1990.
3. Anderson J., Charbonneau H., Cormier M. Mechanism of calcium induction of Renilla bioluminescence. Involvement of a calcium-triggered luciferin binding protein //Biochemistry. — 1974. — Т. 13. — №. 6. — С. 1195-1200.
4. Andryukov B., Mikhailov V., Besednova N. The biotechnological potential of secondary metabolites from marine bacteria //Journal of Marine Science and Engineering. — 2019. — Т. 7. — №. 6. — С. 176.
5. Arun K. B. et al. Bacterial bioactive metabolites as therapeutic agents: From production to action //Sustainable Chemistry and Pharmacy. — 2022. — Т. 27. — С. 100650.
6. Bankevich A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing //Journal of computational biology. — 2012. — Т. 19. — №. 5. — С. 455-477.
7. Basavaraj K. N. et al. Isolation and morphological characterization of antibiotic producing actinomycetes //Tropical Journal of Pharmaceutical Research. — 2010. — Т. 9. — №. 3.
8. Blunt J. W. et al. Marine natural products //Natural product reports. — 2011. — Т. 28. — №. 2. — С. 196-268.
9. Bolger A. M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data //Bioinformatics. — 2014. — Т. 30. — №. 15. — С. 2114-2120.
10. Bonora M. et al. Subcellular calcium measurements in mammalian cells using jellyfish photoprotein aequorin-based probes //Nature protocols. — 2013. — Т. 8. — №. 11. — С. 2105-2118.
11. Bowlby M. R., Case J. F. Flash kinetics and spatial patterns of bioluminescence in the copepod *Gaussia princeps* //Marine Biology. — 1991. — Т. 110. — С. 329-336.
12. Bron J. E. et al. Observing copepods through a genomic lens //Frontiers in zoology. — 2011. — Т. 8. — №. 1. — С. 1-15.
13. Brun P., Payne M. R., Kiørboe T. A trait database for marine copepods //Earth System Science Data. — 2017. — Т. 9. — №. 1. — С. 99-113.

14. Burtseva O. et al. Gut microbiome of the White Sea fish revealed by 16S rRNA metabarcoding //Aquaculture. – 2021. – T. 533. – C. 736175.
15. Callahan B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data //Nature methods. – 2016. – T. 13. – №. 7. – C. 581-583.
16. Campbell A. K. Extraction, partial purification and properties of obelin, the calcium-activated luminescent protein from the hydroid *Obelia geniculata* //Biochemical Journal. – 1974. – T. 143. – №. 2. – C. 411-418.
17. Campbell A. K., Herring P. J. Imidazolopyrazine bioluminescence in copepods and other marine organisms //Marine Biology. – 1990. – T. 104. – C. 219-225.
18. Carman K. R., Dobbs F. C. Epibiotic microorganisms on copepods and other marine crustaceans //Microscopy research and technique. – 1997. – T. 37. – №. 2. – C. 116-135.
19. Carman K. R., Dobbs F. C. Epibiotic microorganisms on copepods and other marine crustaceans //Microscopy research and technique. – 1997. – T. 37. – №. 2. – C. 116-135.
20. Cimermancic P. et al. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters //Cell. – 2014. – T. 158. – №. 2. – C. 412-421.
21. Clardy J., Fischbach M. A., Currie C. R. The natural history of antibiotics //Current biology. – 2009. – T. 19. – №. 11. – C. R437-R441.
22. Cooley R. et al. Development of a cell-free split-luciferase biochemical assay as a tool for screening for inhibitors of challenging protein-protein interaction targets //Wellcome Open Research. – 2020. – T. 5.
23. Coutant E. P. et al. Bioluminescence profiling of NanoKAZ/NanoLuc luciferase using a chemical library of coelenterazine analogues //Chemistry–A European Journal. – 2020. – T. 26. – №. 4. – C. 948-958.
24. Coutant E. P. et al. Bioluminescence profiling of NanoKAZ/NanoLuc luciferase using a chemical library of coelenterazine analogues //Chemistry–A European Journal. – 2020. – T. 26. – №. 4. – C. 948-958.
25. Cregeen S. J. J. Microbiota of dominant Atlantic copepods: *Pleuromamma* sp. as a host to a betaproteobacterial symbiont : дис. – University of Southampton, 2016.
26. Datta M. S. et al. Inter-individual variability in copepod microbiomes reveals bacterial networks linked to host physiology //The ISME journal. – 2018. – T. 12. – №. 9. – C. 2103-2113.

27. Datta M. S. et al. Inter-individual variability in copepod microbiomes reveals bacterial networks linked to host physiology //The ISME journal. – 2018. – T. 12. – №. 9. – C. 2103-2113.
28. Davis N. M. et al. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data //Microbiome. – 2018. – T. 6. – C. 1-14.
29. Dhariwal A. et al. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data //Nucleic acids research. – 2017. – T. 45. – №. W1. – C. W180-W188.
30. Doelle H. W. Bacterial metabolism. – Academic Press, 2014.
31. Fouillaud M., Dufossé L. Microbial secondary metabolism and biotechnology //Microorganisms. – 2022. – T. 10. – №. 1. – C. 123.
32. Francis W. R. et al. Occurrence of isopenicillin-N-synthase homologs in bioluminescent ctenophores and implications for coelenterazine biosynthesis //PLoS One. – 2015. – T. 10. – №. 6. – C. e0128742.
33. Frangoulis C. et al. Importance of copepod carcasses versus faecal pellets in the upper water column of an oligotrophic area //Estuarine, Coastal and Shelf Science. – 2011. – T. 92. – №. 3. – C. 456-463.
34. Frank L. A. Ca²⁺-regulated photoproteins: effective immunoassay reporters //Sensors. – 2010. – T. 10. – №. 12. – C. 11287-11300.
35. Freeman G., Reynolds G. T. The development of bioluminescence in the ctenophore *Mnemiopsis leidyi* //Developmental biology. – 1973. – T. 31. – №. 1. – C. 61-100.
36. Gemayel K., Lomsadze A., Borodovsky M. MetaGeneMark-2: Improved Gene Prediction in Metagenomes //bioRxiv. – 2022. – C. 2022.07. 25.500264.
37. Glud R. N. et al. Copepod carcasses as microbial hot spots for pelagic denitrification //Limnology and Oceanography. – 2015. – T. 60. – №. 6. – C. 2026-2036.
38. Gokulan K., Khare S., Cerniglia C. METABOLIC PATHWAYS | Production of Secondary Metabolites of Bacteria//Encyclopedia of Food Microbiology. – Elsevier, 2014. – P. 561-569.
39. Gothelf A., Mir L. M., Gehl J. Electrochemotherapy: results of cancer treatment using enhanced delivery of bleomycin by electroporation //Cancer treatment reviews. – 2003. – T. 29. – №. 5. – C. 371-387.

40. Gurevich A. et al. QUASt: quality assessment tool for genome assemblies //Bioinformatics. – 2013. – T. 29. – №. 8. – C. 1072-1075.
41. Haddock S. H. D., Case J. F. Bioluminescence spectra of shallow and deep-sea gelatinous zooplankton: ctenophores, medusae and siphonophores //Marine Biology. – 1999. – T. 133. – C. 571-582.
42. Haddock S. H. D., Moline M. A., Case J. F. Bioluminescence in the sea //Annual review of marine science. – 2010. – T. 2. – C. 443-493.
43. Haddock S. H. D., Rivers T. J., Robison B. H. Can coelenterates make coelenterazine? Dietary requirement for luciferin in cnidarian bioluminescence //Proceedings of the National Academy of Sciences. – 2001. – T. 98. – №. 20. – C. 11148-11151.
44. Hastings J. W., Johnson C. H. [3] Bioluminescence and chemiluminescence //Methods in enzymology. – Academic press, 2003. – T. 360. – C. 75-104.
45. Hastings J. W., Morin J. G. Calcium-triggered light emission in Renilla. A unitary biochemical scheme for coelenterate bioluminescence //Biochemical and Biophysical Research Communications. – 1969. – T. 37. – №. 3. – C. 493-498.
46. Haygood M. G. et al. Microbial symbionts of marine invertebrates: opportunities for microbial biotechnology //Journal of molecular microbiology and biotechnology. – 1999. – T. 1. – №. 1. – C. 33-43.
47. Hoarfrost A. et al. Global ecotypes in the ubiquitous marine clade SAR86 //The ISME journal. – 2020. – T. 14. – №. 1. – C. 178-188.
48. Inouye S., Sahara Y. Identification of two catalytic domains in a luciferase secreted by the copepod *Gaussia princeps* //Biochemical and biophysical research communications. – 2008. – T. 365. – №. 1. – C. 96-101.
49. Jiang J., Xiong Y. L. Natural antioxidants as food and feed additives to promote health benefits and quality of meat products: A review //Meat science. – 2016. – T. 120. – C. 107-117.
50. Kafetzis D. A. et al. Treatment of paediatric visceral leishmaniasis: amphotericin B or pentavalent antimony compounds? //International journal of antimicrobial agents. – 2005. – T. 25. – №. 1. – C. 26-30.
51. Kaskova Z. M., Tsarkova A. S., Yampolsky I. V. 1001 lights: luciferins, luciferases, their mechanisms of action and applications in chemical analysis, biology and medicine //Chemical Society Reviews. – 2016. – T. 45. – №. 21. – C. 6048-6077.

52. Kobayashi H. et al. Bioluminescence resonance energy transfer–based imaging of protein–protein interactions in living cells //Nature protocols. – 2019. – T. 14. – №. 4. – C. 1084-1107.
53. Kollár P. et al. Marine natural products: bryostatins in preclinical and clinical studies //Pharmaceutical biology. – 2014. – T. 52. – №. 2. – C. 237-242.
54. Koumoutsis A. et al. Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. – 2004.
55. Koumoutsis A. et al. Structural and functional characterization of gene clusters directing nonribosomal synthesis of bioactive cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. – 2004.
56. Krasitskaya V. V., Bashmakova E. E., Frank L. A. Coelenterazine-dependent luciferases as a powerful analytical tool for research and biomedical applications //International Journal of Molecular Sciences. – 2020. – T. 21. – №. 20. – C. 7465.
57. Kuse M. Chromophores in photoproteins of a glowing squid and mollusk //Bioscience, Biotechnology, and Biochemistry. – 2014. – T. 78. – №. 5. – C. 731-736.
58. Machleidt T. et al. NanoBRET A Novel BRET Platform for the Analysis of Protein–Protein Interactions //ACS chemical biology. – 2015. – T. 10. – №. 8. – C. 1797-1804.
59. Mansson M. et al. Inhibition of virulence gene expression in *Staphylococcus aureus* by novel depsipeptides from a marine photobacterium //Marine drugs. – 2011. – T. 9. – №. 12. – C. 2537-2552.
60. Markova S. V. et al. Cloning and expression of cDNA for a luciferase from the marine copepod *Metridia longa*: a novel secreted bioluminescent reporter enzyme //Journal of Biological Chemistry. – 2004. – T. 279. – №. 5. – C. 3212-3217.
61. Markova S. V., Vysotski E. S. Coelenterazine-dependent luciferases //Biochemistry (Moscow). – 2015. – T. 80. – C. 714-732.
62. Markova S. V., Vysotski E. S. Coelenterazine-dependent luciferases //Biochemistry (Moscow). – 2015. – T. 80. – C. 714-732.
63. Matthews J. C., Hori K., Cormier M. J. Purification and properties of *Renilla reniformis* luciferase //Biochemistry. – 1977. – T. 16. – №. 1. – C. 85-91.
64. McMurdie P. J., Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data //PloS one. – 2013. – T. 8. – №. 4. – C. e61217.

65. Medema M. H. et al. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences //Nucleic acids research. – 2011. – T. 39. – №. suppl_2. – C. W339-W346.
66. Mochizuki S. et al. The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism //Molecular microbiology. – 2003. – T. 48. – №. 6. – C. 1501-1510.
67. Moisander P. H., Sexton A. D., Daley M. C. Stable associations masked by temporal variability in the marine copepod microbiome //PLoS one. – 2015. – T. 10. – №. 9. – C. e0138967.
68. Navarro-Muñoz J. C. et al. A computational framework to explore large-scale biosynthetic diversity //Nature chemical biology. – 2020. – T. 16. – №. 1. – C. 60-68.
69. Nealson K. H., Hastings J. W. Bacterial bioluminescence: its control and ecological significance //Microbiological reviews. – 1979. – V.43. – №. 4. – P.496.
70. Oba Y. et al. Biosynthesis of coelenterazine in the deep-sea copepod, *Metridia pacifica* //Biochemical and biophysical research communications. – 2009. – T. 390. – №. 3. – C. 684-688.
71. Pitterna T. et al. New ventures in the chemistry of avermectins //Bioorganic & medicinal chemistry. – 2009. – T. 17. – №. 12. – C. 4085-4095.
72. Rusanova A. et al. An interplay between viruses and bacteria associated with the White Sea sponges revealed by metagenomics //Life. – 2021. – T. 12. – №. 1. – C. 25.
73. Sadaïappan B. et al. Meta-analysis cum machine learning approaches address the structure and biogeochemical potential of marine copepod associated bacteriobiomes //Scientific Reports. – 2021. – T. 11. – №. 1. – C. 1-17.
74. Sánchez-Guijo F. et al. Sequential third-party mesenchymal stromal cell therapy for refractory acute graft-versus-host disease //Biology of Blood and Marrow Transplantation. – 2014. – T. 20. – №. 10. – C. 1580-1585.
75. Scavotto R. E. et al. Nitrogen-fixing bacteria associated with copepods in coastal waters of the North Atlantic Ocean //Environmental microbiology. – 2015. – T. 17. – №. 10. – C. 3754-3765.
76. Seemann T. Prokka: rapid prokaryotic genome annotation //Bioinformatics. – 2014. – T. 30. – №. 14. – C. 2068-2069.
77. Shimomura O. The discovery of aequorin and green fluorescent protein //Journal of microscopy. – 2005. – T. 217. – №. 1. – C. 3-15.

78. Shimomura O., Johnson F. H., Saiga Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea* //Journal of cellular and comparative physiology. – 1962. – T. 59. – №. 3. – C. 223-239.
79. Shimomura O., Yampolsky I., Bioluminescence: chemical principles and methods. – World Scientific, 2019.
80. Shoemaker K. M., Moisander P. H. Microbial diversity associated with copepods in the North Atlantic subtropical gyre //FEMS Microbiology Ecology. – 2015. – T. 91. – №. 7.
81. Srinivasan R. et al. Marine bacterial secondary metabolites: a treasure house for structurally unique and effective antimicrobial compounds //Marine Drugs. – 2021. – T. 19. – №. 10. – C. 530.
82. Strachan C. R., Davies J. Antibiotics and evolution: food for thought //Journal of Industrial Microbiology and Biotechnology. – 2016. – T. 43. – №. 2-3. – C. 149-153.
83. Strieker M., Tanović A., Marahiel M. A. Nonribosomal peptide synthetases: structures and dynamics //Current opinion in structural biology. – 2010. – T. 20. – №. 2. – C. 234-240.
84. Sudek S. et al. Identification of the putative bryostatin polyketide synthase gene cluster from “*Candidatus Endobugula sertula*”, the uncultivated microbial symbiont of the marine bryozoan *Bugula neritina* //Journal of natural products. – 2007. – T. 70. – №. 1. – C. 67-74.
85. Suginta W. et al. Chitinases from *Vibrio*: activity screening and purification of chiA from *Vibrio carchariae* //Journal of applied microbiology. – 2000. – T. 89. – №. 1. – C. 76-84.
86. Takenaka Y., Yamaguchi A., Shigeri Y. A light in the dark: ecology, evolution and molecular basis of copepod bioluminescence //Journal of plankton research. – 2017. – T. 39. – №. 3. – C. 369-378.
87. Takenaka Y., Yamaguchi A., Shigeri Y. A light in the dark: ecology, evolution and molecular basis of copepod bioluminescence //Journal of plankton research. – 2017. – T. 39. – №. 3. – C. 369-378.
88. Tang K. W. et al. Copepod guts as biogeochemical hotspots in the sea: evidence from microelectrode profiling of *Calanus* spp //Limnology and Oceanography. – 2011. – T. 56. – №. 2. – C. 666-672.

89. Tannous B. A. Gaussia luciferase reporter assay for monitoring biological processes in culture and in vivo //Nature protocols. – 2009. – T. 4. – №. 4. – C. 582-591.
90. Tessler M. et al. Luciferin production and luciferase transcription in the bioluminescent copepod *Metridia lucens* //PeerJ. – 2018. – T. 6. – C. e5506.
91. Thomson C. M., Herring P. J., Campbell A. K. Evidence for de novo biosynthesis of coelenterazine in the bioluminescent midwater shrimp, *Systellaspis debilis* C //Journal of the Marine Biological Association of the United Kingdom. – 1995. – T. 75. – №. 1. – C. 165-171.
92. Titushin M. S. et al. Coelenterazine-binding protein of *Renilla muelleri*: cDNA cloning, overexpression, and characterization as a substrate of luciferase //Photochemical & Photobiological Sciences. – 2008. – T. 7. – C. 189-196.
93. Tsarkova A. S., Kaskova Z. M., Yampolsky I. V. A tale of two luciferins: fungal and earthworm new bioluminescent systems //Accounts of chemical research. – 2016. – T. 49. – №. 11. – C. 2372-2380.
94. Van der Meij A. et al. Chemical ecology of antibiotic production by actinomycetes //FEMS microbiology reviews. – 2017. – T. 41. – №. 3. – C. 392-416.
95. Verhaegen M., Christopoulos T. K. Recombinant Gaussia luciferase. Overexpression, purification, and analytical application of a bioluminescent reporter for DNA hybridization //Analytical chemistry. – 2002. – T. 74. – №. 17. – C. 4378-4385.
96. Verity P. G., Smetacek V. Organism life cycles, predation, and the structure of marine pelagic ecosystems //Marine Ecology Progress Series. – 1996. – T. 130. – C. 277-293.
97. Vincent W. F., Quesada A. Cyanobacteria in high latitude lakes, rivers and seas //Ecology of cyanobacteria II: their diversity in space and time. – 2012. – C. 371-385.
98. Zakar T. et al. Carotenoids assist in cyanobacterial photosystem II assembly and function //Frontiers in plant science. – 2016. – T. 7. – C. 295.
99. Zhang S. D. et al. Solonamides, a Group of Cyclodepsipeptides, Influence Motility in the Native Producer *Photobacterium galathea* S2753 //Applied and Environmental Microbiology. – 2022. – T. 88. – №. 17. – C. e01105-22.