МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА

На правах рукописи

Евсеев Петр Владимирович

Биоинформатические подходы к таксономической классификации бактериофагов

1.5.8 – математическая биология, биоинформатика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата биологических наук

Москва — 2023

Работа выполнена в лаборатории молекулярной биоинженерии Федерального государственного бюджетного учреждения науки Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова Российской академии наук.

Научный руководитель:
Мирошников Константин Анатольевич
доктор химических наук, член-корреспондент РАН

Официальные оппоненты:
Никитин Николай Александрович
доктор биологических наук, профессор кафедры вирусологии биологического факультета Московского государственного университета имени М.В. Ломоносова
Шайтан Константин Вольдемарович
доктор физико-математических наук, профессор, профессор кафедры биоинженерии биологического факультета Московского государственного университета имени М.В. Ломоносова
Белалов Илья Шамильевич
кандидат биологических наук, научный сотрудник лаборатории вирусов микроорганизмов ФИЦ «Фундаментальные основы биотехнологии» РАН

Защита диссертации состоится 22 июня 2023 года в 17:00 на заседании диссертационного совета МГУ.015.10 Московского государственного университета имени М.В. Ломоносова по адресу: 119234, Москва, Ленинские горы, д. 1, стр. 73, Факультет биоинженерии и биоинформатики, ауд. 221.
E-mail: dissovet@belozersky.msu.ru
С диссертацией можно ознакомиться в отделе диссертаций Научной библиотеки МГУ имени М.В. Ломоносова (Москва, Ломоносовский просп., д. 27) и на сайте портале: https://dissovet.msu.ru/dissertation/015.10/2523.

Автореферат разослан «___» мая 2023 года.

Ученый секретарь диссертационного совета,
кандидат химический наук И.В. Шаповалова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность и степень разработанности исследования

Бактериофаги (сокращённо – «фаги») – это вирусы, инфицирующие бактерии. Бактериофаги вездесущи – они обитают в воде, почве, во множестве живых организмов (Wommack and Colwell, 2000). По некоторым оценкам, в течение всего лишь одной секунды фаги вызывают порядка 10^{18} успешных заражений бактериальных клеток (Rohwer et al., 2014). Общее количество бактериофагов можно оценить в 10^{31} вирусных частиц, что больше количества клеток в 10-100 раз (Simmonds et al., 2017). Общая масса этих частиц составляет порядка триллиона тонн (Hendrix et al., 1999). Фаги также являются участниками микробиомов растений и животных, в том числе, и человека – например, его желудочно-кишечный тракт человека содержит более 10^{12} фаговых вирионов (Shkoporov and Hill, 2019)

Способность бактериофагов уничтожать клетки патогенных бактерий привлекла внимание учёных ещё в начале XX века. В последние десятилетия интерес к лечению бактериофагами стал расти, в первую очередь, из-за распространения устойчивости к антибиотикам. Фаговая терапия обладает важными преимуществами (Loc-Carrillo and Abedon, 2011), в том числе устойчивым бактерицидным действием, «автодозированием», заключающимся в том, что количество фагов положительно коррелирует с количеством бактерий-хозяев, фагам присуща низкая собственная токсичность, фаговую терапию отличает минимальное нарушение нормальной флоры и отсутствие перекрестной устойчивости с антибиотиками.

Рост интереса к изучению бактериофагов и их применению требует совершенствования методов характеризации фагов, включая таксономическую классификацию. Исторически фаги классифицировались в соответствии с их морфологией, но во времена первых классификационных схем, описывающих бактериофаги, ещё не существовало ПЦР, секвенирования и многих молекулярных методов, известных нам сегодня (Ackermann, 2004; Turner et al., 2021). Первая распространённая схема таксономической классификации наиболее многочисленных изученных представителей фагов, хвостатых фагов, относящихся к классу *Caudoviricetes*, была предложена Дэвидом Брэдли в 1967 году, расширена А.С. Тихоненко в 1968 году, и усовершенствована Г.-В. Аккерманном и А. Айзенстарком в 1974 году (Ackermann and Eisenstark, 1974; Bradley, 1967). На основе данных электронной микроскопии

бактериофаги классифицировались по трём морфотипам: морфотип А (фаги с сокращающимся хвостом), морфотип В (фаги с длинным, несокращающимся хвостом), морфотип С (фаги с коротким несокращающимся хвостом). Важную роль в развитии систем классификации вирусов, в том числе бактериальных, сыграла Балтиморская классификация – система, представленная Д. Балтимором в 1971 году (Baltimore, 1971). В Балтиморской системе все вирусы сгруппированы в шесть разрозненных групп без каких-либо дальнейших подразделений, которые обычно называются балтиморскими классами (БК). Балтиморские классы были установлены на основе типа нуклеиновой кислоты, включенной в вирионы, и, следовательно, на типе репродукции вируса. Классификация бактериофагов долгое время строилась в соответствии с Балтиморской системой и морфологией фагов.

Совершенствование методов молекулярной биологии, секвенирование фаговых геномов и бурное развитие биоинформатики в начале 2000-х годов выявило гораздо более высокое геномное разнообразие, чем считалось ранее, особенно у хвостатых бактериофагов. По мере увеличения количества секвенированных геномов бактериофагов стало понятным, что три семейства хвостатых фагов, *Myoviridae,* (морфотип А), *Podoviridae* (морфотип С) и *Siphoviridae* (морфотип В) не являются монофилетическими. Парафилия этих семейств была показана биоинформатическими методами, благодаря которым было выделено несколько новых семейств. Биоинформатические подходы позволили уточнить классификацию семейства *Microviridae,* бактериофагов с одноцепочечной ДНК, значительно увеличив количество подсемейств. С начала 2000-2010 годов всё большую роль в описании новых организмов стала играть метагеномика. Метагеномный анализ позволил получить последовательности геномов некультивируемых организмов, которых, как представляется, существенно больше, чем культивируемых. На основе метагеномных данных были выделены фаги группы crAss (Dutilh et al., 2014) и некоторые гигантские бактериофаги (Al-Shayeb et al., 2020; Devoto et al., 2019).

Последние годы ознаменовались взрывным ростом количества новых таксонов бактериофагов. В 2018-2020 годах ICTV (International Committee on Taxonomy of Viruses, Международный комитет по таксономии вирусов) официально ратифицировал три новых семейства миовирусов (*Ackermannviridae, Chaseviridae, Herelleviridae*), два новых семейств сифовирусов (*Demerecviridae* и *Drexlerviridae*) а также два семейства

подовирусов (*Autographiviridae* и *Schitoviridae*). В 2019 году решением рабочей группы ICTV количество рангов таксономической классификации вирусов было увеличено до пятнадцати, включая восемь основных (первичных) и семь производных (или вторичных) (рис. 1). Восемь основных рангов включают четыре, которые уже использовались раньше (отряд, семейство, род и вид) и четыре новых ранга (реалм, царство, тип и класс), которые выше ранга отряда. Базовый ранг в вирусной таксономии предложено называть реалмом, а не доменом (как в таксономических системах клеточных организмов), чтобы показать сложную взаимосвязь между таксономией вирусов и клеточных организмов.

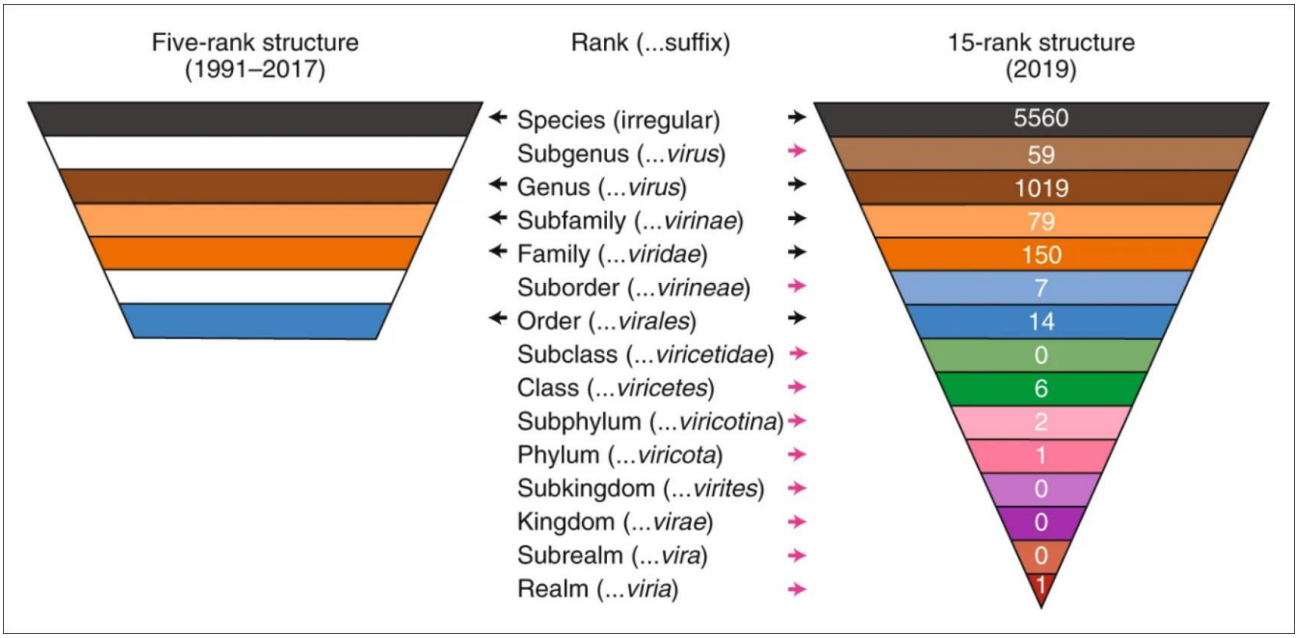


Рисунок 1. 15-ранговая схема таксономической классификации, принятая ICTV в 2020 г. и предыдущая пятиранговая схема. Количество таксонов, присвоенных каждому рангу согласно Основному списку видов ICTV (выпуск 2018b, MSL34) показано белым шрифтом. Когда ранги описываются как иерархия, ранг вида часто называют низшим рангом, а ранг области - высшим рангом. Чёрные стрелки указывают на ранги, общие для пяти- и 15-ранговой схем; розовые стрелки указывают на ранги, введенные в 15-ранговую схему (International Committee on Taxonomy of Viruses Executive Committee, 2020).

В 2021 году была принята новая схема таксономической классификации хвостатых бактериофагов, ратифицированная в марте 2022 года. Новая классификация ликвидировала семейства *Myoviridae, Podoviridae* и *Siphoviridae*, основанные на основании морфологических характеристик (Walker et al., 2021). Все хвостатые фаги были отнесены к классу *Caudoviricetes*, в свою очередь, принадлежащего типу *Uroviricota* царства *Heunggongvirae* реалма *Duplodnaviria*. Отряд *Caudovirales*, включавший ранее в

себя все хвостатые фаги, был ликвидирован, взамен были образованы новые таксоны, в том числе высокоранговые. В новообразованный отряд *Crassvirales* были включены фаги группы crAss, ранее обнаруженные в метагеноме человеческого пищеварительного тракта и инфицирующие бактерии типа *Bacteroidetes*. Также было образовано три отряда, включающие вирусы архей; остальные хвостатые бактериофаги были отнесены к 35 семействам, 37 подсемействам и 493 родам, которые приписаны к классу *Caudoviricetes* напрямую, без отнесения к таксонам промежуточного ранга. Тем не менее, большинство из более чем двадцати тысяч известных науке бактериофагов, геномы которых находятся в базе данных Genome NCBI, на настоящий момент не классифицировано, что отчасти объясняется трудностями обоснованного таксономического описания фагов.

Современная таксономическая классификация бактериофагов требует привлечения целого комплекса биоинформатических методов, включая анализ состава и организации генома, кластеризацию на основе средненуклеотидного геномного сходства, филогенетический анализ с использованием последовательностей коровых генов и консервативных белков, анализ протеома. Определение возможной таксономической классификации может затрудняться из-за частых генетических обменов, которые могут включать как отдельные гены, так и их группы. Другая проблема биоинформатических классификационных подходов – это быстрая эволюция вирусных белков, затрудняющая создание иерархической кластеризации, особенно для высокоранговых таксонов.

Задача определения таксономического положения фагов имеет как практическое значение, позволяя предсказывать биологические свойства бактериофагов (в том числе, для целей фаговой терапии), так и теоретическое фундаментальное значение для эволюционной биологии и вирусологии. Эта задача особенно актуальна для малоисследованных и новых бактериофагов, инфицирующих микроорганизмы, которые являются патогенами человека, либо представляют другой интерес для экономики и народного хозяйства, в том числе в качестве перспективных средств биоконтроля патогенов сельскохозяйственных культур.

Цели работы:

– проанализировать применимость биоинформатических методов для таксономического описания бактериофагов на примере новых бактериофагов, инфицирующих патогены растений *Curtobacterium* и *Pectobacterium*, а также на примере малоизученного фага *Pseudomonas* MD8.

Задачи исследования:

1. Предложить обоснованную таксономическую классификацию на основе геномных данных для новых бактериофагов, инфицирующих бактерии родов *Curtobacterium* и *Pectobacterium*.
2. Оценить таксономическое разнообразие бактериофагов, инфицирующих бактерии семейства *Pectobacteriaceae*, которые являются патогенами сельскохозяйственных культур.
3. Проанализировать профаговые области фитопатогенных бактерий рода *Curtobacterium* с целью классификации потенциальных умеренных фагов и поиска генов белков, способных к разрушению клеточных оболочек этих бактерий.
4. Проанализировать применимость биоинформатических подходов для таксономической классификации умеренного бактериофага *Pseudomonas* MD8.
5. Проанализировать применимость биоинформатических подходов, основанных на сравнении структурного сходства белков, моделированных с помощью современных методов глубокого обучения, для выявления эволюционных взаимосвязей бактериофагов в целях построения высокоранговой таксономической иерархии.

Научная новизна работы

Впервые проведён таксономический и геномный анализ новых бактериофагов, инфицирующих грамположительные бактерии рода *Curtobacterium* (фаг Ayka, геномные области профагового происхождения) и грамотрицательные бактерии рода *Pectobacterium* (Horatius, Possum, PP47, PP81, Q19). Проанализировано таксономическое разнообразие бактериофагов, инфицирующих бактерии семейства *Pectobacteriaceae*, вызывающую мягкую гниль сельскохозяйственных культур. На примере фага *Pseudomonas* MD8 впервые детально показан процесс мозаичного формирования геномов умеренных фагов псевдомонад, обсуждены трудности, возникающие при таксономической классификации, и предложено их возможное решение. С использованием предсказанных структур фаговых белков методами глубокого обучения проанализированы возможности улучшения описания функций фаговых белков и эволюционных взаимосвязей между бактериофагами.

Теоретическая и практическая значимость полученных результатов

Результаты диссертационной работы расширяют представления о таксономическом

разнообразии бактериофагов, в том числе, фагов, инфицирующих бактерии, являющиеся опасными для человека и экономически важных сельскохозяйственных культур. Полученные в ходе работы аннотированные геномные последовательности депонированы в международной базе данных NCBI GenBank.

Предложенные подходы к биоинформатическому анализу бактериофагов и методы решения проблем, вызванных генетическим мозаицизмом и быстрой эволюцией вирусных белков, могут быть использованы для построения более точной таксономической иерархии.

Положения, выносимые на защиту:

1. Биоинформатические методы позволяют уверенно классифицировать бактериофаги *Pectobacterium* PP47, PP81, Q19 и предложить классификационную схему на уровне рода, подсемейства, семейства.
2. Использование биоинформатических методов для таксономической классификации нового фага *Curtobacterium* Ayka позволяет предложить классифицировать этот фаг как представителя нового вирусного семейства или подсемейства.
3. Использование биоинформатических методов для геномного анализа исследованных умеренных профагов может быть затруднено в связи с ярко выраженным генетическим мозаицизмом этих фагов.
4. Новые алгоритмы структурного моделирования белков могут быть использованы в целях таксономической классификации.

Степень достоверности данных

Данные, представленные в работе, получены с использованием современных программ и программных пакетов. Обзор литературы и обсуждение результатов подготовлены с использованием актуальных литературных источников. Достоверность полученных результатов определяется достаточным объемом проведенных исследований, использованием в работе современных экспериментальных и биоинформатических методов. Достоверность результатов также подтверждается публикациями в рецензируемых отечественных и международных журналах, депонированием генетических последовательностей в международную базу данных NCBI GenBank.

Апробация результатов

Результаты диссертационной работы представлены и обсуждены на следующих

международных и российских научных конференциях: «The Future Applications of Bacteriophages» (Зевейл, Египет, 2021), «Bioinformatics: from algorithms to applications (BIATA)» (С.-Петербург, Россия, 2021), «IEEE Ural-Siberian Conference on Computational Technologies in Cognitive Science, Genomics and Biomedicine (CSGB)» (Новосибирск, Россия, 2021), «24th Evergreen Phage Meeting» (Олимпия, США, 2021), III Всероссийской конференции «Высокопроизводительное секвенирование в геномике» (Новосибирск, Россия, 2022).

Личный вклад автора

Личный вклад автора заключается в постановке цели исследования и задач, анализе литературных данных, получении данных и обработке полученных результатов, подготовке публикаций и научных докладов. Соискателем был проведен всесторонний биоинформатический анализ геномных данных, включая детальную аннотацию фаговых геномов. Все этапы биоинформатического анализа от сборки геномов до обработки результатов вычислений и выбора задействованных биоинформатических алгоритмов осуществлялись соискателем лично. Все биоинформатические иллюстрации в диссертации были сделаны соискателем. Соискатель также принимал участие в постановке задач и планировании экспериментов (в частности, в экспериментах по индукции профагов и характеризации фагов). Соискателем написаны все главы диссертации, сформулированы выводы и практические рекомендации. Электронная микроскопия выполнена Е.Е. Куликовым (Московский физико-технический институт, Институт микробиологии им. С.Н. Виноградского РАН) и В.А. Кадыковым (Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова).

Публикации по теме диссертации

По результатам исследования опубликовано 9 научных работ: 8 статей в рецензируемых журналах, включённых в системы цитирования Scopus, Web of Science и RSCI, рекомендованных для защиты в диссертационном совете МГУ имени М.В. Ломоносова, 1 расширенный тезис конференций.

Структура и объем диссертации

Диссертация состоит из введения, 3 глав, заключения, выводов и списка литературы. Работа изложена на 246 страницах, содержит 6 таблиц, 103 рисунка, 6 приложений. Список литературы включает 278 источников, из которых 3 на русском

языке, 275 на иностранных языках и 6 интернет-ресурсов.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

В обзоре литературы описаны история становления и развития подходов к таксономической классификации бактериофагов, современные биоинформатические методы, использующиеся в таксономической классификации фагов.

ГЛАВА 2. ОБЪЕКТЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Объектами исследования являлись бактериофаги, в том числе, выделенные в Лаборатории молекулярной биоинженерии Института биоорганической химии РАН и в Лаборатории водной микробиологии Лимнологического института СО РАН.

Геномные последовательности новых бактериофагов были получены сборкой *de novo* данных секвенирования. Геномные последовательности выделенных ранее бактериофагов загружали из базы данных NCBI GenBank и при наличии исходных данных секвенирования проверяли пересборкой. Экспериментально определённые структуры белков загружали из базы данных RCSB Protein Data Bank.

Протокол таксономического анализа включал следующие этапы характеризации фагового генома с использованием методов, рекомендованных ICTV:

- анализ организации генома, с учётом состава геномных модулей и их расположения;
- поиск родственных фагов, включая фаги, близкие по средненуклеотидному геномному составу и содержащие гомологичные гены;
- получение кластерной тепловой карты межгеномного сходства исследуемого и родственных фагов;
- филогенетический анализ с использованием последовательностей консервативных белков;
- анализ протеома;
- анализ характерных генов и белков.

2.3.2. Аннотирование фаговых геномов

Геномы собирали с помощью программ SPAdes и CLC Genomic Workbench. Аннотирование геномов проводили с помощью программы Prokka с использованием

пользовательских баз данных, базы данных RVDB и встроенных баз данных Prokka. Открытые рамки считывания (open reading frames, ORF) были предсказаны с помощью Prodigal 2.6.1, Glimmer 3.02b и Geneious Prime и проверены вручную.

Функции кодируемых в геноме белков предсказывали поиском гомологичных последовательностей и сравнением мотивов HMM. Поиск гомологичных последовательностей выполняли с помощью BLAST с использованием баз данных NCBI и пользовательских баз данных BLAST с использованием фаговых последовательностей GenBank. Поиск с использованием HMM-мотивов проводили с помощью Phyre2 и HHpred с использованием баз данных ECOD_F70, PDB_mmCIF70, Pfam_A, SCOPe70, и UniProt-SwissProt-viral70, а функции присваивали сравнением с аналогичными белками, используя порог 95% достоверности Phyre2 или вероятности HHpred. Наличие генов тРНК проверяли с помощью tRNAscan-SE и ARAGORN. Генетические карты визуализировали в Geneious Prime.

Вычисления среднегеномного нуклеотидного сходства (average nucleotide identity, ANI) фаговых геномов проводили с помощью orthoANIu с использованием алгоритма USEARCH для поиска сходства со всеми полными и частичными последовательностями геномов бактериофагов, размещёнными в базе данных NCBI Genome на момент проведения анализа (примерно 24 тысячи последовательностей в 2019 г. и 31 тыс. последовательностей в 2022 году). Кластерную тепловую карту на основе значений ANI получали с помощью BIONJ. Кластерную тепловую карту межгеномного сходства с учётом доли выровненных участков геномов, рассчитанного программой VIRIDIC, получали с помощью встроенных инструментов VIRIDIC с использованием настроек по умолчанию.

Визуальное сравнение геномов бактериофагов выполняли с помощью программ EasyFig, использующего для анализа сходства отдельных генов TBLASTX.

Выравнивания аминокислотных и нуклеотидных последовательностей для построения филогенетических деревьев делали с помощью Clustal Omega 1.2.3 с настройками «ten refinement iterations, evaluating full distance matrix for initial and guide trees», MAFFT 7.48 с настройками по умолчанию и с использованием алгоритма L-INS-i и MUSCLE 3.8 с настройками по умолчанию.

Филогенетические деревья, основанные на выравнивании аминокислотных и нуклеотидных последовательностей, строили с помощью нескольких алгоритмов,

включая RAxML-NG, RAxML, MrBayes, FastTree. RAxML-NG использовали вместе с графическим интерфейсом raxmlGUI 2 и настройками «--tree rand{10} --bs-trees 1000», применяя весовую матрицу, выбранную с помощью ModelTest-NG. Надежность деревьев RAxML-NG оценивали с помощью бутстрэп-анализа (количество реплик 1000) и вычисления значений бутстрэпа или ожидаемых результатов бутстрэпа (transfer bootstrap expectation, TBE). Дендрограммы RAxML получали с использованием алгоритма «Rapid Bootstrapping and search for best-scoring tree» и весовых матриц BLOSUM62 и LG. Оценку надёжности деревьев RAxML проводили с помощью бутстрэп-анализа (количество реплик 1000). Дендрограммы MrBayes получали с использованием длины цепи 5500000 и длины прожига 500000. Деревья FastTree строили с применением стандартных настроек и весовой матрицы Whelan and Goldman's 2001.

Дендрограммы, основанные на сходстве протеома и геномной организации фагов, строили с помощью сервера GRAViTy с использованием базы данных DB-B: Baltimore Group Ib - Prokaryotic and archaeal dsDNA viruses (VMRv34). Дендрограммы, основанные на сходстве протеома, строили с помощью VIPtree с использованием встроенной базы данных прокариотических вирусов. Карты генетической сети получали с помощью ConTACT.2.0 с использованием встроенной базы данных ProkaryoticViralRefSeq database v94 и визуализировали в Cytoscape.

Предсказание структур белков делали с помощью алгоритмов глубокого обучения AlphaFold 2 и RoseTTAFold. Для моделирования AlphaFold 2 использовали полные базы данных и параметры командной строки «--monomer_casp14, --monomer и –multimer». Моделирование с помощью RoseTTAFold проводили, используя стандартные настройки сервера Robetta.

Сравнение структур белков проводили с помощью программ структурного выравнивания DALI и mTM-align используя стандартные настройки. Суперимпозицию и визуализацию структур осуществляли с помощью PyMOL 2.

Поиск профаговых участков в геномах проводили с использованием пайплайна PhiSpy и сервера Phaster с настройками по умолчанию. Предсказанные последовательности профагов были аннотированы с использованием Prokka, серверов HHpred и Phyre 2 и поиска BLAST по базам данных NCBI, а также пользовательским базам данных с использованием фаговых геномных последовательностей.

Базы данных последовательностей геномов, генов и белков организовывали в среде Geneious Prime. Для создания поисковых баз данных BLAST использовали инструменты BLAST.

ГЛАВА 3. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

3.1. Таксономический анализ бактериофагов *Pectobacterium*

3.1.1. Фаги семейства *Autographiviridae*

Фаги семейства PP47, PP89 и Q19 инфицируют фитопатогенные бактерии, относящиеся к роду *Pectobacterium*. Они показывают эффективное антибактериальное действие против похожего круга (спектра) хозяев и обладают типичной подовирусной морфологией, схожей с морфологией фага *Escherichia* T7, относящегося к семейству *Autographiviridae* (рис. 2). Геномная организация фагов PP47, PP89 и Q19 также близка к таковой для фагов семейства *Autographiviridae* (рис. 3).

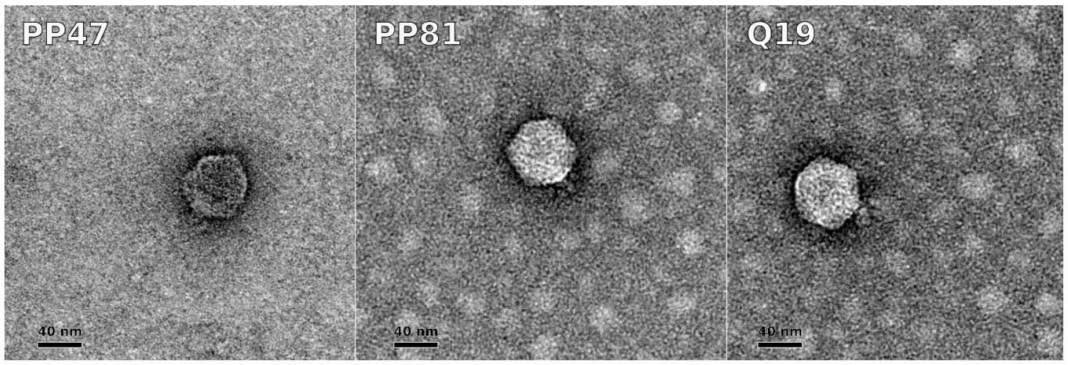


Рисунок 2. Электронная микрофотография фагов *Pectobacterium* PP47, PP81 и Q19. Масштаб шкалы — 40 нм.

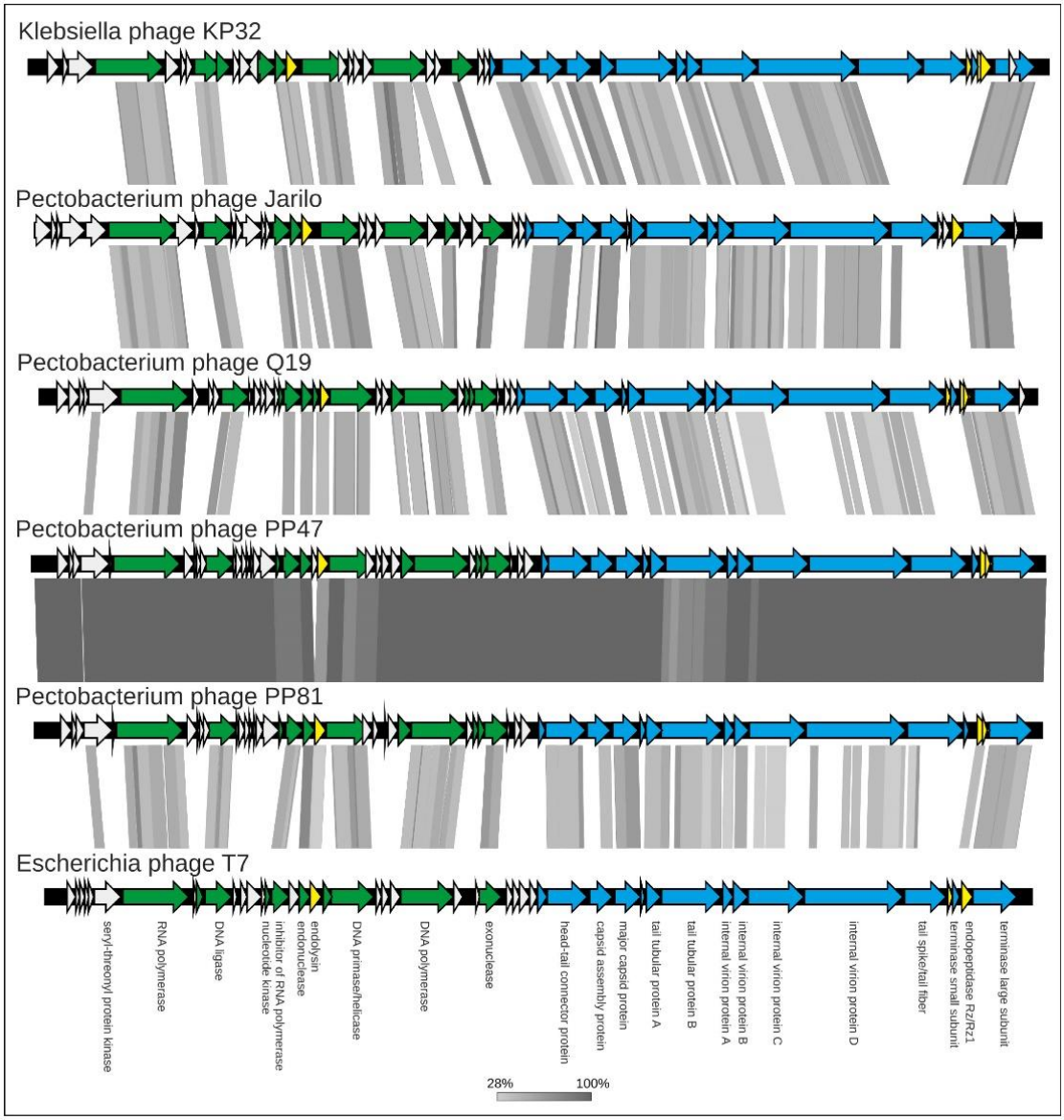


Рисунок 3. Сравнение организации геномов фагов PP47, PP89 и Q19 и других фагов семейства *Autographiviridae*. Гены белков, участвующих в транскрипции и репликации окрашены в зелёный цвет, гены структурных и упаковочных белков окрашены в голубой цвет гены белков лизиса окрашены в жёлтый цвет.

Сравнение нуклеотидного геномного сходства с использованием ортоANIu с использованием всех полных геномных последовательностей фагов *Autographiviridae* Genbank показало значительное сходство между фагами PP47 и PP81 (ANI около 98%), а также и меньшее сходство этих фагов с Q19 и другими фагами *Autographiviridae*. Учитывая 95%-ный видовой порог нуклеотидного сходства, близость геномов PP47 и PP81 позволяет отнести их к одному виду. Кластеризация фаговых геномов с помощью рекомендованной ICTV программы VIRIDIC сгруппировала фаги PP47 и PP81 в один кластер вместе с фагами рода *Pektosvirus* подсемейства *Studiervirinae*. Межгеномное сходство фагов PP47, PP81 и других фагов внутри этого кластера составило более 70%-

ной границы принадлежности к одному роду, что позволяет отнести фаги PP47, PP81, а также PPWS4, MA1A и MA6 к роду *Pektosvirus*.

Согласно результатам вычислений межгеномного сходства VIRIDIC, фаг *Pectobacterium* Q19 ближе всего к фагам родов *Jarilovirus* и *Unyawovirus*, но значения межгеномного сходства Q19 и представителей *Jarilovirus* и *Unyawovirus* составили около 55% и 53% соответственно, что не позволяет отнести Q19 к этим или другим родам. Филогенетический анализ с использованием конкатенированных аминокислотных последовательностей консервативных генов (рис. 4) и анализ протеома подтверждают родство фагов PP47, PP81 и фагов рода *Pektosvirus*, а также указывают на близость фага Q19 и фагов родов *Jarilovirus* и *Unyawovirus*. Согласно критериям ICTV, фаг Q19 может быть классифицирован как представитель нового рода.

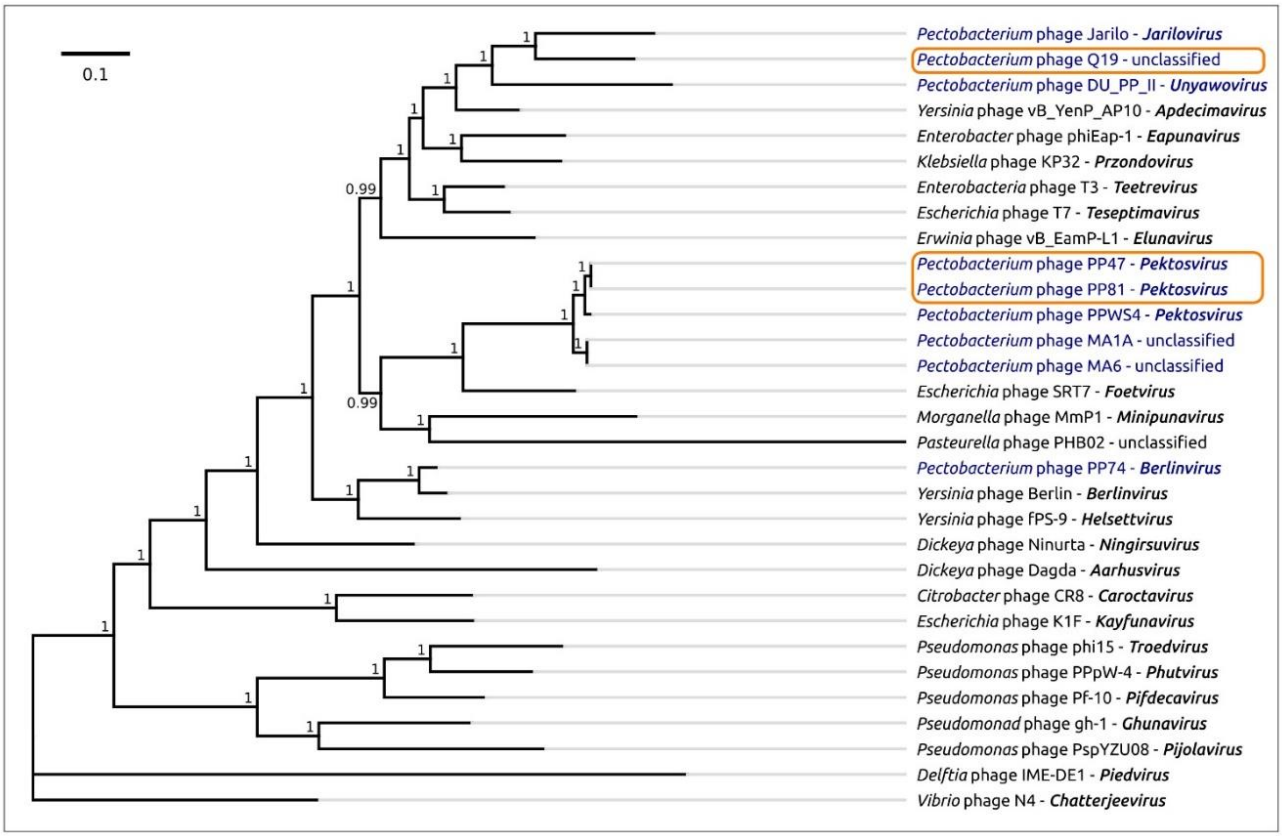


Рисунок 4. Филогенетическое дерево, построенное с использованием конкатенированных аминокислотных последовательностей пяти консервативных белков. Родовая принадлежность указана справа от названия фага.

Филогенетический и структурный анализы также показали интересную особенность белка хвостового шипа, участвующего в адсорбции фага на клетку хозяина, которая может объяснить близкий спектр хозяев фагов PP47, PP89 и Q19. Эти белки характеризуются модульной структурной архитектурой. При этом С-концевой домен

хвостовых шипов, отвечающий за специфичность фагов, и N-концевой домен, отвечающий за крепление шипа к вириону, имеют разную эволюционную историю, включающую горизонтальные переносы с участием бактерий-хозяев. С-концевые домены хвостовых шипов фагов PP47 и PP89 филогенетически ближе к С-концевому домену хвостового шипа фага Q19, чем к С-концевым доменам хвостовых фибрилл фагов *Pektosvirus*, что может объяснить сходный спектр хозяев PP47, PP89 и Q19.

3.1.2. Фаги семейства *Schitoviridae*

Фаги *Pectobacterium* Possum и Horatius (семейство *Schitoviridae*) инфицируют сходный спектр хозяев и морфологически неразличимы (рис. 5). Геномы этих двух фагов почти идентичны, за исключением вставки длиной 15 п.н. в гене *rIIB* Possum и нескольких однонуклеотидных замен, не влияющих на аминокислотные последовательности кодируемых белков. Размер и организация геномов Possum и Horatius близки к таковым у фагов *Pectobacterium* CB1, CB3 и CB4, Nepra, ϕA38 и ϕA41, принадлежащим к группе N4-подобных фагов, недавно выделенных в новое семейство *Schitoviridae* (ратифицировано в марте 2021 года). Геномы фагов Possum и Horatius, как и геномы других фагов семейства *Schitoviridae,* содержат как минимум три гена ДНК-зависимых РНК-полимераз (РНКП). Отличительной чертой этого семейства фагов является ген большой инкапсулированной РНКП (вирионной РНКП, вРНКП). вРНКП вводится в инфицируемую бактериальную клетку вместе с ДНК фага, что способствует быстрому началу транскрипции ранних генов.

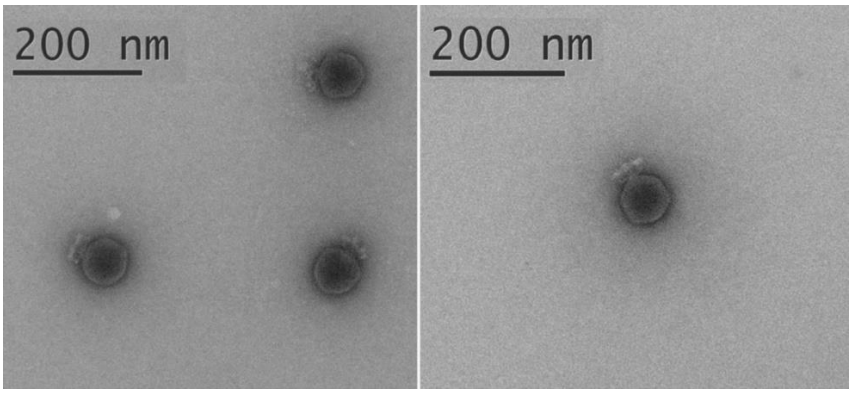


Рисунок 5. Электронная микрофотография фагов *Pectobacterium* Possum (слева) и Horatius (справа). Масштаб шкалы — 200 нм.

Для уточнения таксономической классификации фагов *Pectobacterium* Possum и Horatius были проведены полногеномные сравнения и филогенетические исследования. Расчёты ANI показали, что фаги *Pectobacterium* vB_PatP_CB1, CD3, CB4, ϕA38, ϕA41 и

Nepra, выделенные в род *Cbunavirus*, являются ближайшими родственниками Possum с близкими значениями ANI около 94%. Результаты анализа VIRIDIC продемонстрировали, что все перечисленные выше фаги *Pectobacterium* группируются в один кластер с межгеномным сходством более 70%. Результаты филогенетического анализа с использованием последовательностей главного капсидного белка, портального белка и терминазы, как и анализ протеома (аминокислотных последовательностей фаговых белков), показали (рис. 6), что фаги Possum и Horatius, а также классифицированные фаги рода *Cbunavirus* группируются в одну монофилетическую ветку. Таким образом, все эти фаги можно считать представителями одного рода *Cbunavirus* семейства *Schitoviridae*.

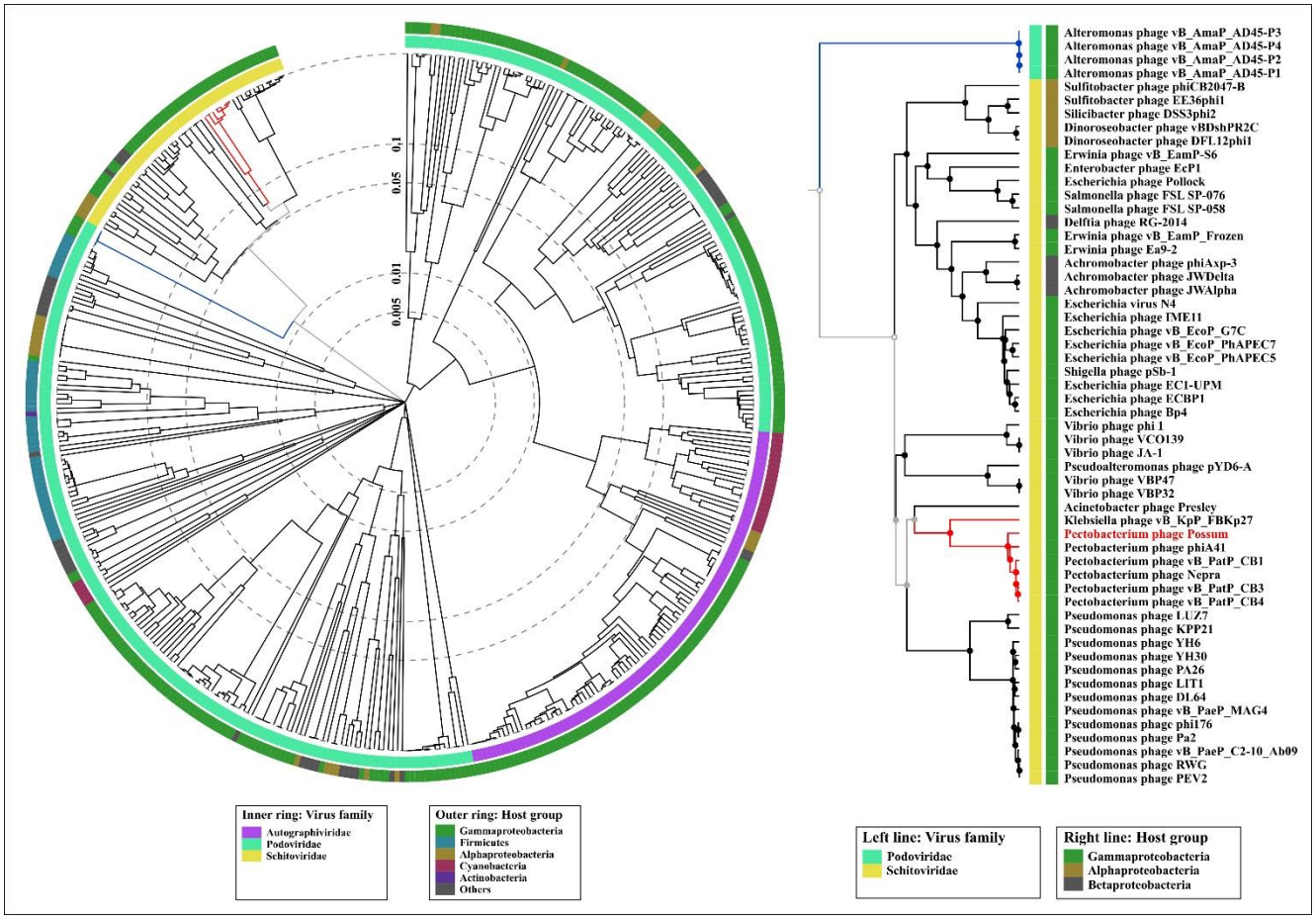


Рисунок 6. Круговая дендрограмма (слева) и её фрагмент (справа), построенная на основании сходства протеомов фагов. В легендах указана таксономическая принадлежность фагов по состоянию на 2020 год.

3.1.3. Фаги, инфицирующие фитопатогенные бактерии *Pectobacterium* и *Dickeya*

Таксономическое разнообразие фагов, инфицирующие фитопатогенные бактерии *Pectobacterium* и *Dickeya* (PD-фаги), было изучено с использованием геномных последовательностей 108 PD-фагов, размещённых в базе данных NCBI Genome. Анализ указал на высокое таксономическое и биологическое разнообразие PD-фагов, включающих представителей 7 семейств и 9 подсемейств (по номенклатуре ICTV на начало 2021 г.) (рис. 7). Бóльшую часть PD-фагов (58 геномов) представляют фаги семейства *Autographiviridae*, интересные с точки зрения применения в фаговой терапии. Филогенетический анализ, проведённый с использованием последовательностей консервативных белков, и анализ характерных белков, позволил уточнить ранее неизученные эволюционные взаимосвязи и характерные особенности ряда PD-фагов.

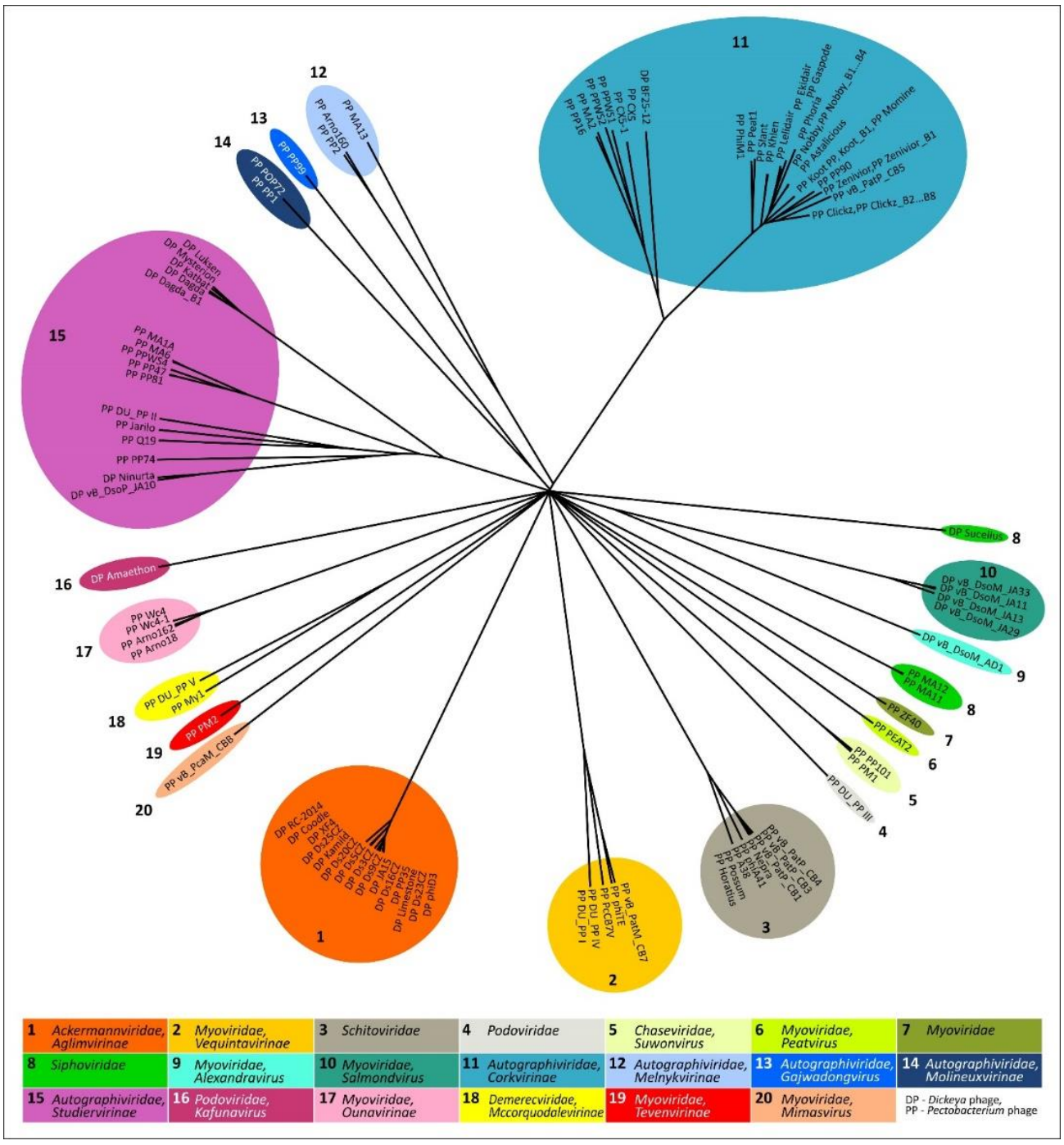


Рисунок 7. Дендрограмма, построенная с использованием значений среднегеномного нуклеотидного сходства ANI. В легендах указана таксономическая принадлежность фагов по состоянию на начало 2021 г.

3.2. Таксономический анализ бактериофагов *Curtobacterium*

3.2.1. Фаг *Curtobacterium* Ayka

Фаг Ayka (рис. 8) является первым описанным представителем бактериофагов, инфицирующих фитопатогенные бактерии, принадлежащие к роду *Curtobacterium*. Сравнительно небольшой геном фага состоит из 18400 н.п. и содержит блок генов процессинга ДНК, включая ген, кодирующий терминальный белок. Наличие этого гена

свидетельствует о механизме репликации, использующем белковый праймер, что характерно для ряда фагов подовирусной морфологии, родственных фагу φ29 и другим фагам семейства *Salasmaviridae*. Сравнение HMM мотивов показало сходство белков морфогенеза с аналогичными белками фага φ29. Тем не менее, сравнение межгеномного сходства показало его уровень по сравнению с родственными фагами около 5% и менее, что значительно ниже 70%-ного порога принадлежности к одному роду.

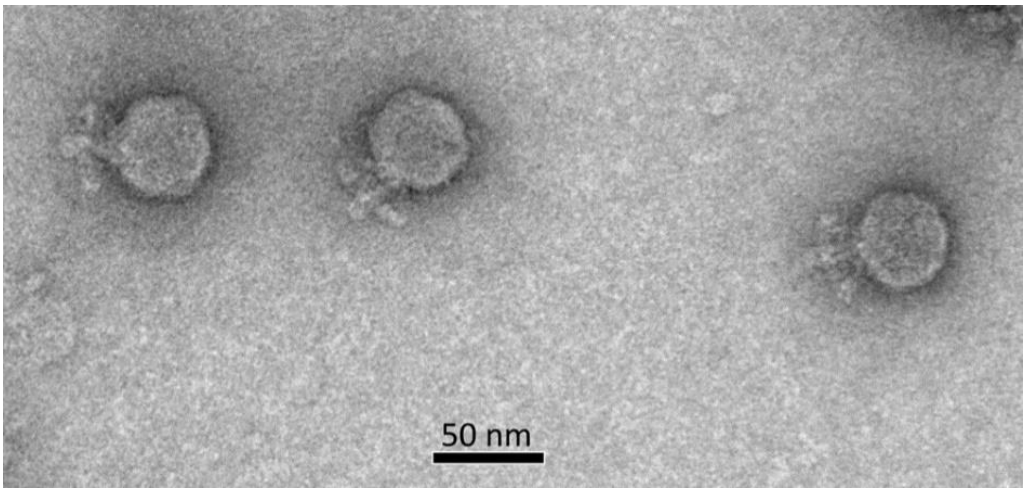
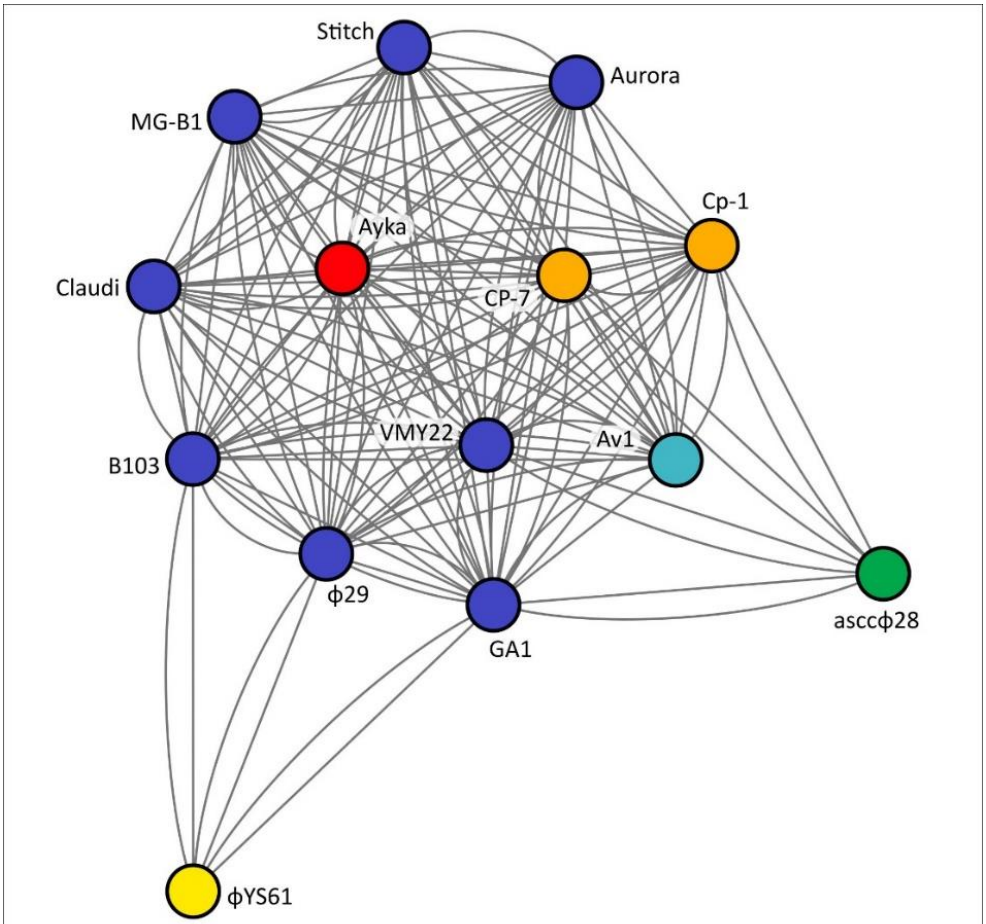


Рисунок 8. Электронная микрофотография фага *Curtobacterium* Ayka. Масштаб шкалы — 50 нм.

Анализ генетической сети, включающей фаг Ayka и геномы прокариотических вирусов базы данных пайплайна ConTACT.2.0 (рис. 9), показали наличие эволюционных связей между фагом Ayka, фагами семейства *Salasmaviridae* и другими небольшими фагами, инфицирующими грамположительные бактерии. Филогенетический анализ с использованием аминокислотных последовательностей пяти консервативных белков, протеомная филогения GRAViTy и филогения, основанная на сравнении предсказанных структур главного капсидного белка и терминазы, поместили фаг Ayka и фаги *Salasmaviridae* в сестринские группы, составляющие вместе с другими родственными фагами общую кладу. Генетические дистанции между фагом Ayka и родственными группами фагов имели примерно тот же порядок, что и генетические дистанции между классифицируемыми подсемействами семейства *Salasmaviridae*, и рядом других классифицированных семейств. Таким образом, фаг Ayka является первым представителем нового семейства или подсемейства вирусов класса *Caudoviricetes*.



	Phage	Family	Subfamily	Genus
●	*Actinomyces* virus Av1	–	–	*Dybvigvirus*
●	*Bacillus* phage Aurora	*Salasmaviridae*	*Northropvirinae*	*Claudivirus*
●	*Bacillus* phage Claudi	*Salasmaviridae*	*Northropvirinae*	*Claudivirus*
●	*Bacillus* phage MG-B1	*Salasmaviridae*	*Northropvirinae*	*Klosterneuburgvirus*
●	*Bacillus* phage Stitch	*Salasmaviridae*	*Northropvirinae*	*Claudivirus*
●	*Bacillus* phage VMY22	*Salasmaviridae*	–	*Mingyongvirus*
●	*Bacillus* virus B103	*Salasmaviridae*	*Picovirinae*	*Beecentumtrevirus*
●	*Bacillus* virus GA1	*Salasmaviridae*	*Tatarstanvirinae*	*Gaunavirus*
●	*Bacillus* virus ϕ29	*Salasmaviridae*	*Picovirinae*	*Salasvirus*
●	*Curtobacterium* phage Ayka	–	–	–
●	*Lactococcus* phage ascc ϕ28	–	–	–
●	*Streptococcus* phage Cp-1	*Salasmaviridae*	–	*Cepunavirus*
●	*Streptococcus* phage CP-7	*Salasmaviridae*	–	*Cepunavirus*
●	*Weissella* phage ϕYS61	–	–	–

Рисунок 9. Карта генетической сети фага Ayka и родственных фагов, полученная с использованием встроенной базы данных ProkaryoticViralRefSeq database v94 пайплайна ConTACT.2.0. В легендах указана таксономическая принадлежность фагов по состоянию на 2022 год.

3.2.2. Профаги *Curtobacterium*

Области профагового происхождения искали с помощью двух популярных инструментов, пайплайна Phaster и сервера PhiSpy, результаты предсказаний которых, тем не менее, приводили к несовпадающим результатам. Ручная проверка предсказанных кодирующих рамок и функциональная аннотация позволила определить границы 64

областей, содержащих гены главного капсидного белка, терминазы и других генов фагового происхождения (рис. 10).

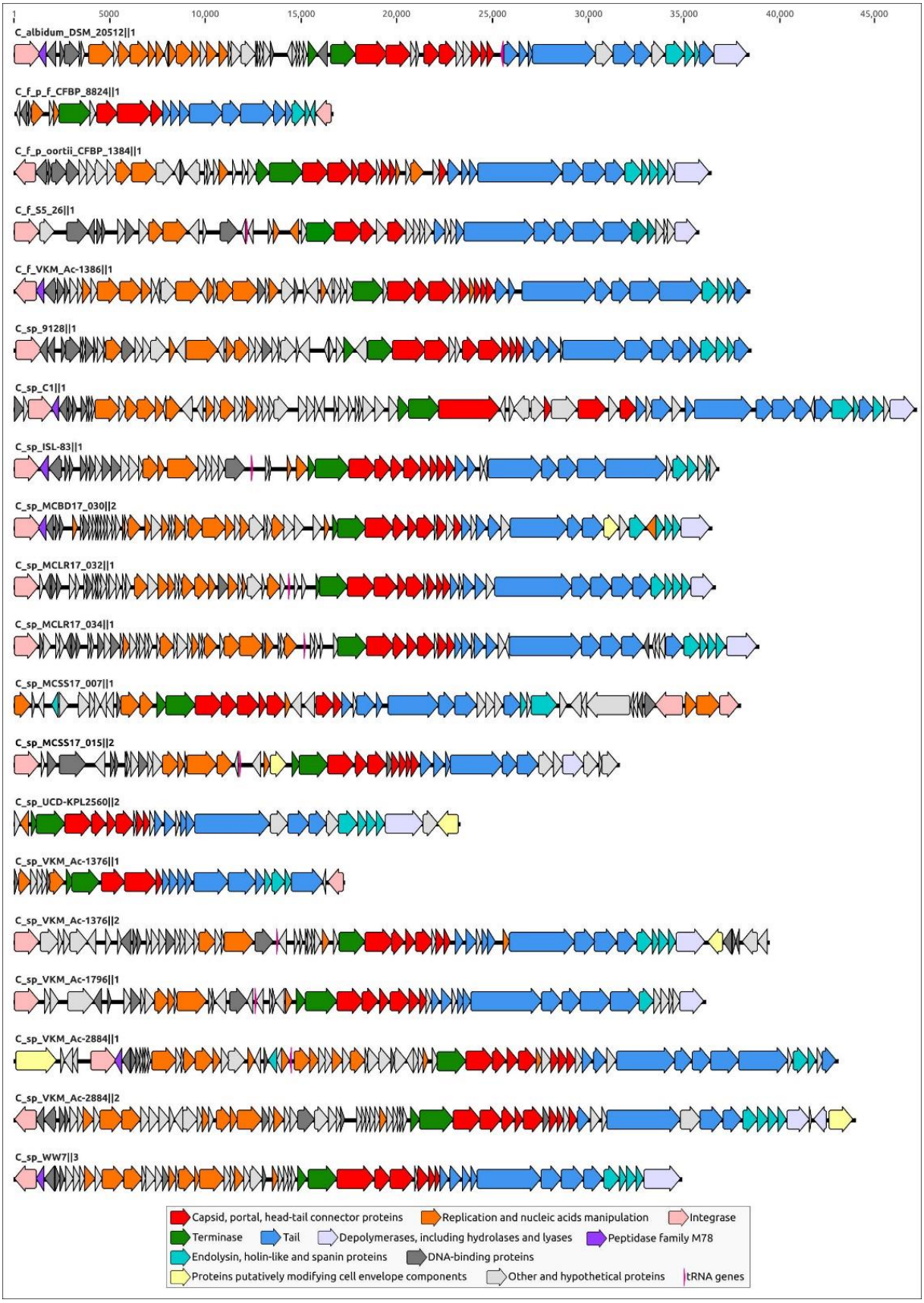


Рисунок 10. Генетические карты профаговых областей в геномах *Curtobacterium*.

Межгеномные сравнения показали сложную картину родственных взаимоотношений предсказанных профагов. Применяя порог сходства 95%, две группы можно рассматривать как один и тот же вид. На основании межгеномного сходства 46 профагов можно сгруппировать в несколько кластеров, причём только два профага можно отнести к одному роду. Это означает, что либо большинство профагов представляют собой либо отдаленные таксономические группы функциональных умеренных фагов, либо сильно мутировавшие дефектные профаги. Некоторые области профагов показали сходство одновременно с профагами, принадлежащими к разным кластерам, что может быть результатом генетического мозаицизма, характерного для эволюции фагов, особенно умеренных. Таким образом, кластеризация предсказанных профагов с использованием межгеномного сходства не позволяет предложить полностью непротиворечивую классификацию.

Филогенетический анализ, проведенный с использованием аминокислотных последовательностей ГКБ (главный капсидный белок) и TerL (terminase large subunit, большая субъединица терминазы), закодированных в выбранных областях профага, и других фаговых белков показал несовпадающую топологию деревьев, хотя состав клад был похожим. Эти различия также могут быть результатом мозаичной эволюции фаговых геномов. Для определения близкородственных таксономических групп фагов проводили поиск близких по ANI фагов с использованием представителей групп, сгруппированных с помощью филогении с использованием последовательностей ГКБ и всех полных геномов фагов *Caudoviricetes*, доступных в базе данных NCBI Genome по состоянию на июль 2022 года. Далее рассчитывали межгеномное сходство для предсказанных профагов и известных фагов с использованием инструмента VIRIDIC. Эти анализы не выявили значимого сходства между профаговыми регионами *Curtobacterium* и известными полными фаговыми геномами. Однако небольшое сходство порядка 10% было выявлено у некоторых фагов, поражающих бактерии *Microbacterium*, эволюционно близкие к роду *Curtobacterium*. В частности, небольшие фаги *Microbacterium* с размером генома менее 20 тыс. н.п., недавно отнесенные к новым таксонам (семейство *Orlajensenviridae*, подсемейство *Pelczarvirinae*, род *Paopuvirus*), показывают некоторое сходство с профагами *Curtobacterium* с геномами близкого размера.

Филогенетический анализ с использованием аминокислотных последовательностей консервативных белков может выявить более отдаленные

эволюционные и таксономические отношения. Список родственных таксономических групп фагов, выявляемых с использованием филогений ГКБ и TerL, в основном включает как неклассифицированные, так и классифицированные умеренные актинофаги сифовирусной морфологии, отнесённые к подсемейству *Nclasvirinae* и нескольким десяткам родов, не выделенных в подсемейство или семейство, в том числе *Bridgettevirus, Britbratvirus, Bronvirus, Coralvirus, Decurrovirus, Fromanvirus, Mapvirus, Timquatrovirus* и др. Родственные фаги инфицируют бактерии, принадлежащие к родам *Arthrobacter, Attisvirus, Bifidobacterium, Corynebacterium, Gordonia, Microbacterium, Mycobacterium, Streptomyces, Propionibacterium, Rathayibacter* и *Rhodococcus*. Поиск BLAST с помощью последовательностей спейсеров CRISPR, найденных в геномах *Curtobacterium*, обнаружил сходные участки, также, в основном, в геномах актинофагов сифовирусной морфологии, в том числе фагов, принадлежащих к подсемействам *Arquatrovirinae, Bclasvirinae, Guernseyvirinae, Mclasvirinae, Nclasvirinae, Nymbaxtervirinae, Weiservirinae*, неклассифицированных фагов и фагов, принадлежащих к разным родам, не отнесенных к подсемействам или семействам.

При анализе профаговых областей было выявлено около 100 генов, кодирующих эндолизины и другие гликополимер-деградирующих белки (ГДБ), которые могут рассматриваться в качестве кандидатов для использования в качестве антибактериальных агентов. Интересно, что кластеризация не эндолизиновых ГДБ, корректно отражающая функции этих белков, не могла быть проведена с использованием филогении, задействующей аминокислотные последовательности, из-за большого различия последовательностей белков и их разного происхождения. Тем не менее, использование структурного сходства моделей, предсказанных AlphaFold, позволило сгруппировать эти ГДБ в соответствии с их предсказанной функцией.

3.3. Таксономический анализ бактериофага *Pseudomonas* MD8

3.3.1. Геномный и филогенетический анализ

Фаг *Pseudomonas* MD8 – умеренный фаг, принадлежащий к классу *Caudoviricetes* с с геномом размера около 43 тыс. н.п. Структура генома фага MD8 напоминает геном фага *Escherichia* λ и другие лямбдоидные бактериофаги. Начиная с 5'-конца, геном включает блок из 30 генов морфогенеза, ориентированных в прямом направлении, блок из 23 генов, включая интегразу, ориентированных в обратном направлении, и блок из 14 генов, начинающийся с репрессора *cro* и содержащий, в том числе, гены репликации и лизиса.

Вычисления ANI с использованием всех фаговых геномов, содержащихся в базе NCBI Genome и анализ межгеномного сходства VIRIDIC позволили выявить родственные фаги (т.н. «группа MD8», «MD8-подобные фаги»), которые, тем не менее, не отвечают критериям принадлежности к одному роду с фагом MD8 (рис. 11). Геном фага MD8 проявляет сходство с геномами разных групп родственных фагов *Pseudomonas*.

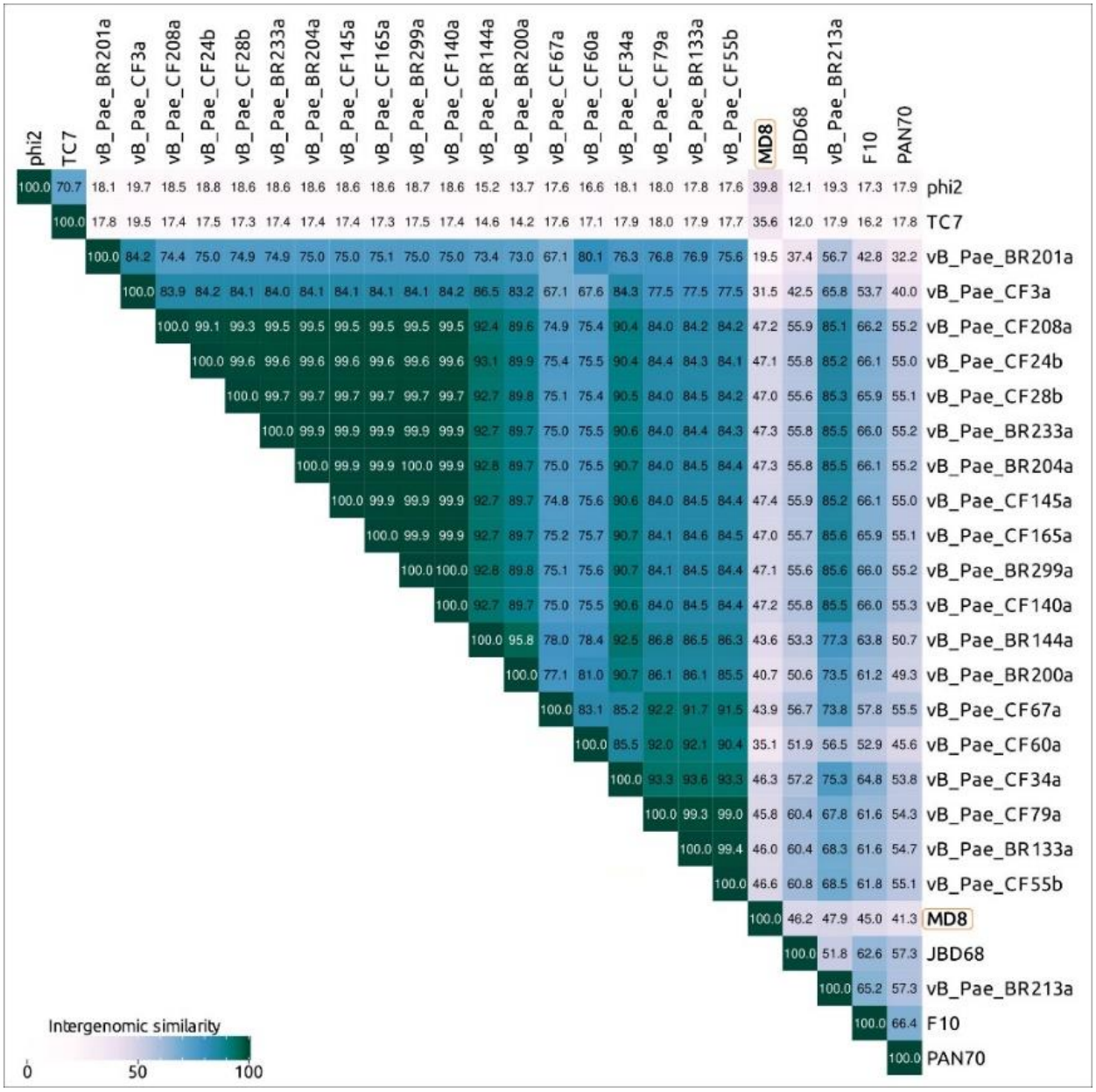


Рисунок 11. Кластерная тепловая карта родственных бактериофагов, инфицирующих представителей *Pseudomonas*, полученная на основании межгеномного сходства, рассчитанного VIRIDIC.

3.3.2. Анализ мозаичности генома

Филогенетический анализ, проведённый на основе аминокислотных последовательностей ГКБ и TerL MD8-подобных и других бактериофагов показал разную

топологию дендрограмм для этих двух белков. Для терминазы результаты анализа указывали на близкое родство и общее происхождение TerL большей части фагов группы MD8 и λ-подобных фагов *Escherichia*, в то время как терминазы фагов φ2 и TC7 эволюционно ближе к терминазе фага *Thermus* φFA1, чем к терминазам фага λ и MD8. Результаты филогенетических исследований главного капсидного белка MD8 указывали на другую эволюционную историю этого белка и его более отдалённое родство с ГКБ фага λ. Главные капсидные белки фагов φ2 и TC7 не показали какого-либо значительного сходства с аналогами остальных фагов группы MD8. Тем не менее, детальный филогенетический анализ, проведённый с последовательностями всех белков, кодируемых в геноме фага MD8, показал сходство большого количества белков, в первую очередь, относящихся к белкам репликации и лизогении, с белками фагов φ2 и TC7. Таким образом, геном фага MD8 имеет мозаичное строение и сложную эволюционную историю.

3.4. Использование предсказаний структуры фаговых белков для эволюционной таксономии

3.4.1. Эволюция чехольных белков

Чехольные белки хвостовой трубки (ЧБ) составляют часть сократительного молекулярного механизма бактериофагов с миовирусной морфологией и системой секреции VI типа (T6SS), обнаруживаемой у многих грамотрицательных бактерий. Экспериментально определённые структуры ЧБ указывали на наличие общих особенностей структуры данных белков, в том числе, консервативного домена, но детали эволюции этих белков и возможности использования их структуры для таксономического анализа были неясны. В связи с ограниченным объёмом экспериментальных данных, использовали более 100 структур ЧБ, представляющих разные группы фагов и T6SS, предсказанных с помощью программы глубокого обучения AlphaFold 2 (рис. 12).

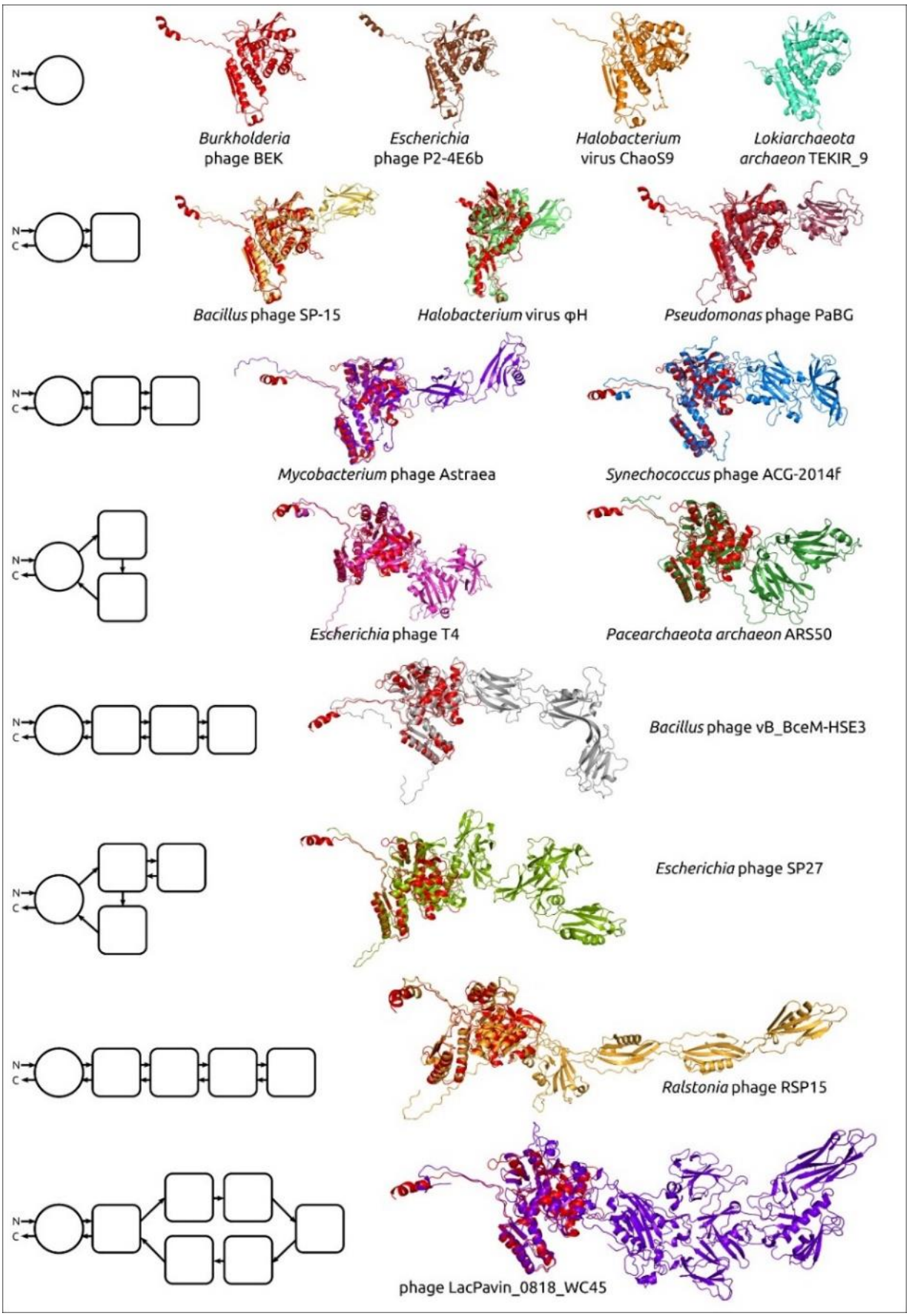


Рисунок 12. Предсказанная с помощью AlphaFold 2 структурная архитектура белков хвостовых чехлов разных фагов.

Анализ структурной архитектуры показал, что все чехольные белки содержат структурно схожий высоко консервативный домен, включающий как N-концевую, так и С-концевую части, тогда как остальные части состоят из одного или нескольких умеренно консервативных доменов, предположительно добавленных в ходе эволюции фагов. Дополнительные домены могут способствовать поддержанию стабильности вириона или адсорбции на клетке-хозяине, а их количество коррелирует с размером генома фага и часто является характерным признаком для таксономических групп.

3.4.2. Использование структурных предсказаний для классификации таксонов высокого ранга

В работе проведён комплексный анализ главного капсидного белка и большой субъединицы терминаз 50 представителей всех классифицированных семейств вирусов, отнесённых к реалму *Duplodnaviria*, включая бактериофаги, хвостатые вирусы архей и вирусы герпеса, а также несколько гигантских фагов, фагов λ, HK-97, Ayka и MD8. Анализ включал предсказание структур с помощью AlphaFold и RoseTTAFold, а также филогенетический анализ с использованием выравниваний, полученных разными алгоритмами. Сравнение результатов моделирования ГКБ с помощью AlphaFold и RoseTTAFold показало, в целом, более высокую точность AlphaFold. Сравнение результатов предсказаний структур ГКБ и TerL с помощью AlphaFold показало более высокую точность предсказаний для терминазы.

Анализ результатов сравнения предсказанных структур (рис. 13) в ряде случаев позволил получить биологически более осмысленную кластеризацию, чем филогенетический анализ, основанный на выравниваниях аминокислотных последовательностей без учёта структурного сходства. В то же время, и результаты сравнения предсказанных структур, и результаты филогенетического анализа указывают на существование противоречий с результатами анализа GRAViTy, основанного на сравнении состава протеома и геномной организации. Обсуждены причины этих противоречий и сделаны предложения по совершенствованию существующей таксономической классификации.

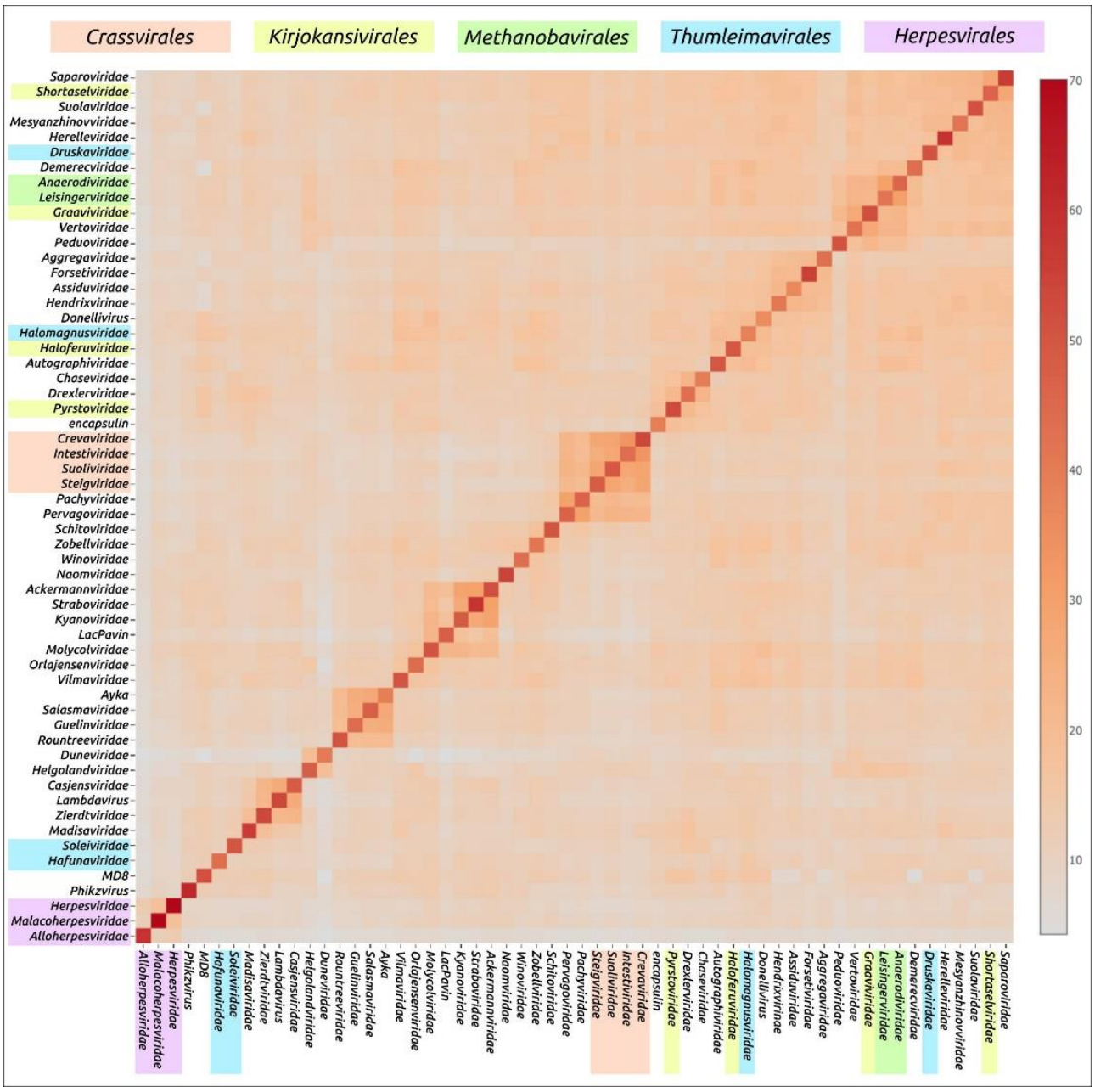


Рисунок 13. Кластерная тепловая карта, основанная на попарных сравнениях предсказанных структур 57 главных капсидных белков с использованием DALI Z-score.

ВЫВОДЫ

1. Таксономический анализ свидетельствует о принадлежности фагов *Pectobacterium* PP47, PP81, Q19 к семейству *Autographiviridae* и подсемейству *Studiervirinae*. Несмотря на схожую морфологию, спектр хозяев и организацию геномов, фаги PP47 и PP81, с одной стороны, и фаг Q19, с другой, относятся к разным близким таксономическим группам. Фаг Q19 должен быть отнесён к новому роду.
2. Фаги *Pectobacterium* Possum и Horatius являются представителями семейства *Schitoviridae* и рода *Cbunaviruses*.

3. Биоинформатический анализ указывает на высокое таксономическое разнообразие бактериофагов, инфицирующих бактерии рода *Pectobacterium* и другие фитопатогены, вызывающие болезнь мягкой гнили растений.
4. Фаг *Curtobacterium* Ayka представляет новое семейство или подсемейство.
5. Профаговые области геномов бактерий рода *Curtobacterium* представляют собой либо интактные профаги, родственные умеренным актинофагам, требующие классификации в новые таксоны ранга рода и выше, либо являются необратимо интегрированными нефункциональными профагами родственными умеренным актинофагам.
6. Геномный и филогенетический анализ указывают на высокую интенсивность горизонтальных переносов между различными умеренными фагами, инфицирующими *Pseudomonas aeruginosa*, вызвавшую ярко выраженный генетический мозаицизм фага MD8, что затрудняет его таксономическую классификацию. Высокий уровень генетического мозаицизма требует выработки новых критериев для классификации подверженных ему бактериофагов.
7. Предсказание структуры белков с помощью новых алгоритмов глубокого обучения может способствовать построению эволюционно-биологически осмысленной иерархической классификации. Кластеризация и использование структурного сходства таких моделей могут быть использованы для выявления глубоких эволюционных связей и построения классификационной системы таксонов высокого ранга.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в рецензируемых журналах Scopus, Web of Science и RSCI (в скобках приведен объем публикации в печатных листах и вклад автора в печатных листах):

1. **Evseev P.V.**, Lukianova A.A., Shneider M.M., Korzhenkov A.A., Bugaeva E.N., Kabanova A.P., Miroshnikov K.K., Kulikov E.E., Toshchakov S.V., Ignatov A.N., Miroshnikov K.A. Origin and Evolution of *Studiervirinae* Bacteriophages Infecting *Pectobacterium*: Horizontal Transfer Assists Adaptation to New Niches // Microorganisms, 2020, Vol. 8, No. 11, P. 1707. IF 4,782, (1,64/1,15).

2. Miroshnikov K.A., **Evseev P.V.**, Lukianova A.A., Ignatov A.N. Tailed Lytic Bacteriophages of Soft Rot *Pectobacteriaceae*. Microorganisms // Microorganisms, 2021, Vol. 9, No. 9, P. 1819. IF 4,782, (2,26/0,90).
3. **Evseev P.**, Lukianova A., Sykilinda N., Gorshkova A., Bondar A., Shneider M., Kabilov M., Drucker V., Miroshnikov K. *Pseudomonas* Phage MD8: Genetic Mosaicism and Challenges of Taxonomic Classification of Lambdoid Bacteriophages // International Journal of Molecular Sciences, 2021, Vol. 22, No. 19, P. 10350. IF 6,009, (2,03/1,42).
4. **Evseev P.**, Shneider M., Miroshnikov K. Evolution of Phage Tail Sheath Protein // Viruses, 2022, Vol. 14, No. 6, P. 1148. IF 5,712, (2,29/1,61).
5. Lukianova A.A., **Evseev P.V.**, Shneider M.M., Dvoryakova E.A., Tokmakova A.D., Shpirt A.M., Kabilov M.R., Obraztsova E.A., Shashkov A.S., Ignatov A.N., Knirel Y.A., Dzhalilov F.S.-U., Miroshnikov K.A. *Pectobacterium* versatile Bacteriophage Possum: A Complex Polysaccharide-Deacetylating Tail Fiber as a Tool for Host Recognition in Pectobacterial *Schitoviridae* // International Journal of Molecular Sciences, 2022, Vol. 23, No. 19, P. 11043. IF 6,009, (1,12/0,34).
6. Tarakanov R.I., Lukianova A.A., **Evseev P.V.**, Pilik R.I., Tokmakova A.D., Kulikov E.E., Toshchakov S.V., Ignatov A.N., Dzhalilov F.S.-U., Miroshnikov K.A. Ayka, a Novel Curtobacterium Bacteriophage, Provides Protection against Soybean Bacterial Wilt and Tan Spot // International Journal of Molecular Sciences, 2022, Vol. 23, No. 18, P. 10913. IF 6,009, (1,11/0,33).
7. **Evseev, P.**; Lukianova, A.; Tarakanov, R.; Tokmakova, A.; Popova, A.; Kulikov, E.; Shneider, M.; Ignatov, A.; Miroshnikov, K. Prophage-Derived Regions in Curtobacterium Genomes: Good Things, Small Packages // International Journal of Molecular Sciences 2023, Vol. 24, No. 2, P. 1586. IF 6,009, (1,78/1,25).
8. **Evseev, P.**; Gutnik, D.; Shneider, M.; Miroshnikov, K. Use of an Integrated Approach Involving AlphaFold Predictions for the Evolutionary Taxonomy of Duplodnaviria Viruses // Biomolecules 2023, Vol. 13, No. 1, P. 110. IF 5,880, (1,97/1,38).

В материалах конференций:

1. **Евсеев П.В.**, Лукьянова А.В., Токмакова А.Д., Шнейдер М.М., Игнатов А.Н., Попова А.В., Мирошников К.А. Профаговые области в геномах *Curtobacterium* spp. и *Curtobacterium flaccumfaciens* pv. *flaccumfaciens*: геномика и белки, разрушающие клеточную стенку//Сборник тезисов III Всероссийской конференции «Высокопроизводительное секвенирование в геномике (HSG-2022)» Новосибирск, 2022. С. 48. (0,10/0,07).