

**Никулин Никита Алексеевич**

**Эволюционная дивергенция T4-родственных бактериофагов,  
связанная с неканоническими азотистыми основаниями ДНК**

1.5.3. – молекулярная биология

**АВТОРЕФЕРАТ**

Диссертации на соискание  
ученой степени  
кандидата биологических наук

Работа выполнена в Федеральном государственном бюджетном учреждении науки «Федеральный исследовательский центр «Пушкинский научный центр биологических исследований Российской академии наук»

**Научный руководитель:**

Зимин Андрей Антонович – кандидат биологических наук

**Официальные оппоненты:**

Рукавцова Елена Борисовна – доктор биологических наук, Федеральное государственное бюджетное учреждение науки Институт биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова РАН, группа биотехнологии растений, старший научный сотрудник

Феоктистова Наталья Александровна – кандидат биологических наук Федеральное государственное бюджетное образовательное учреждение высшего образования «Ульяновский государственный аграрный университет имени П.А. Столыпина», кафедра микробиологии, вирусологии, эпизоотологии и ВСЭ, доцент

**Ведущая организация:**

Федеральное бюджетное учреждение науки “Московский научно-исследовательский институт эпидемиологии и микробиологии им Г. Н. Габричевского” Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека (ФБУН МНИИЭМ им. Г. Н. Габричевского Роспотребнадзора)

Защита диссертации состоится 23 мая 2024 года в 16 ч 00 мин на заседании совета 24.1.232.01 (Д 002.285.01) на базе Федерального государственного бюджетного учреждения науки «Федеральный исследовательский центр «Пушкинский научный центр биологических исследований Российской академии наук» Институт биофизики клетки по адресу: 142290, Московская область, г. Пушкино, ул. Институтская, д. 3

С диссертацией можно ознакомиться в библиотеке «Федеральный исследовательский центр «Пушкинский научный центр биологических исследований Российской академии наук» по адресу: 142290 г. Пушкино, ул. Институтская, д. 3. и на сайте <https://www.pbcras.ru>

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2024 года

Учёный секретарь  
диссертационного совета,  
доктор биологических наук

Дягтерева Ольга Васильевна

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы.** Бактериофаги или вирусы бактерий – это биологические объекты, являющиеся репликаторами, которые участвуют в коэволюционных взаимодействиях «хозяин-паразит» (Koonin, Starokadomskyy, 2016).

Особенностью бактериофагов с двуцепочечной ДНК является наличие большого числа неканонических азотистых оснований, то есть оснований, отличных от аденина, тимина, гуанина, цитозина (Hutinet et al., 2019; Weigele, Raleigh, 2016). Причины этого явления на данный момент не известны, однако предполагается, что это связано с особенностями коэволюционных взаимодействий «бактерия-фаг», в которых такие основания выполняют функцию защиты от действия систем рестрикции-модификации, систем CRISPR-Cas бактерии-хозяина (Hutinet et al., 2019; Vlot et al., 2018; Weigele, Raleigh, 2016). Другой предполагаемой причиной является возможность сохранения бактериофагами остатков ранее существующего разнообразия оснований ДНК, которые были свойственны предкам бактерий, за счет особенной формы существования фагов. Изучение роли неканонических оснований в составе геномной ДНК бактериофагов позволит выявить коэволюционные и эволюционные процессы, происходившие на ранних этапах формирования живых организмов.

Наиболее подходящим объектом для такой цели являются Т4-родственные бактериофаги, архетипом для которых является фаг Т4. Ряд их представителей хорошо изучен (например, фаги Т2, Т4, Т6), а также используется в качестве модельных организмов в молекулярной биологии. Кроме того, у Т4-родственных бактериофагов было определено наличие нескольких неканонических оснований и их модификаций, а также выявлены пути их биосинтеза (Hutinet et al., 2019; Thomas et al., 2018). Также у фага Т4 найдены гены, участвующие в процессах, которые зависят от ДНК, содержащей неканонические основания, или функционирование такой ДНК в клетке зависит от процессов, в которых участвуют эти гены (Petrov et al., 2010). Стоит отметить, что за счет наличия относительно большого числа геномов Т4-родственных бактериофагов в базах данных, а также присутствия этих фагов в различных нишах, возможно исследование их эколого-эволюционных взаимосвязей при помощи биоинформатических инструментов. По вышеприведенным причинам Т4-родственные бактериофаги представляют собой наиболее подходящие модельные объекты для исследования влияния неканонических оснований на эволюционную дивергенцию вирусов. Т4-родственные бактериофаги являются также одними из наиболее перспективных агентов фаговой терапии, которая становится все более актуальна в связи с растущим числом бактерий с множественной устойчивостью к антибиотикам. Их перспективность связана с несколькими аспектами. Существует ряд исследований по конструированию и применению фаговых препаратов на основе представителей Т4-родственных вирусов (Bourdin et al., 2014). Некоторые Т4-родственные бактериофаги способны инфицировать бактерии-патогены. Также, у фага Т4 исследованы гены, участвующие в снижении частоты трансдукции (Young et al., 1982), за счет чего, методами геномики возможен первичный отбор Т4-родственных фагов, предположительно не осуществляющих горизонтальный перенос генов или осуществляющих этот процесс с низкой частотой. Функции этих генов также опосредованы неканоническими основаниями.

Исследования Т4-родственных бактериофагов и их неканонических оснований ДНК являются актуальными как с точки зрения фундаментальных, так и прикладных знаний, в связи с возможностью использования их в изучении эволюции вирусов и коэволюции в системе «фаг-бактерия», а также изучении подбора наиболее перспективных агентов фаговой терапии.

**Цель исследования.** Определение влияния неканонических оснований ДНК на эволюцию Т4-родственных бактериофагов.

**Задачи исследования:**

1. Провести изучение геномов Т4-родственных фагов, не относящихся к виду *Tequatrovirus* Т4, выбранных из охарактеризованной коллекции бактериофагов сточных вод и фекалий зубров, для выявления генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК и определения таксономической принадлежности к родам подсемейства *Tevenvirinae*.
2. Провести анализ геномов и пан-генома вирусов подсемейства *Tevenvirinae* из баз данных, филогенетический анализ генетических маркеров этих вирусов, для определения влияния экологических факторов и влияния генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, на дивергенцию геномов Т4-родственных бактериофагов.
3. Выявить особенности расположения генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, в геномах Т4-родственных вирусов и особенности их наследования.
4. На основе полученных данных предложить концепцию влияния неканонических оснований ДНК на эволюцию Т4-родственных бактериофагов и возможность распространения этой концепции на другие группы вирусов.

**Научная новизна работы.** Впервые определено преобладание в фаговом сообществе, которое инфицирует *E.coli*, фекалий зубров Приокско-Тerrasного биосферного заповедника бактериофагов морфотипов, свойственных “хвостатым” бактериофагам. Найдены и охарактеризованы бактериофаги, принадлежащие к Т4-родственным. Определено наличие у них генетических маркеров, свойственных близкородственным *Tequatrovirus* Т4, определена их способность инфицировать штаммы *E.coli* с системами рестрикции-модификации (RM-системами).

Предложен новый метод скрининга, основанный на спот-тесте на штаммах с RM-системами, ПЦР и электрофорезе фаговых частиц, позволяющий охарактеризовать фаги, выделенные из природных источников и источников антропогенного происхождения, с точки зрения наличия у них неканонических оснований ДНК и принадлежности к Т4-родственным бактериофагам. В результате из сточных вод очистных сооружений города Пущино были выделены Т4-родственные бактериофаги, которые имели неканонические основания в составе ДНК и принадлежали к родам *Tequatrovirus*, *Mosigvirus* подсемейства *Tevenvirinae*.

Впервые отмечено, что степень родства исследуемых вирусов зависит от наличия генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК. Показано, что фаги, принадлежащие к одному роду, обладают одинаковыми неканоническими основаниями, схожими модификациями этих оснований и основными генами, способствующими воспроизведению фагов с неканоническими основаниями ДНК.

Впервые показано, что в процессе эволюции Т4-родственных бактериофагов происходило накопление генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, которые также участвуют в уменьшении

возможностей взаимодействия их ДНК с ДНК бактерии-хозяина, что приводит к уменьшению возможностей горизонтального переноса генов хозяина.

Впервые определено общее расположение большинства генов синтеза 5-гидроксиметилцитозина и его модификаций у T4-родственных вирусов в регионе между двумя коргенами – генами ДНК-полимеразы и хеликазы. Показано наличие внутри этого региона у представителей отдельных групп этих фагов разных хоминг-эндонуклеаз (homing endonucleases). Это, а также ряд других полученных данных может свидетельствовать о высокой частоте рекомбинации в данной области геномов у предков исследуемых бактериофагов.

На основе анализа геномов и информации из литературы и баз данных об источниках выделения представителей T4-родственных вирусов, чьи полногеномные последовательности имелись в NCBI Nucleotide Collection, впервые показана связь эволюции T4-родственных вирусов с экологическими нишами и генами, способствующими воспроизведению фагов с неканоническими основаниями ДНК.

На основе полученных результатов и эволюционной модели Постоянного Разнообразия (Constant-Diversity model) впервые предложена концепция эволюции предков T4-родственных бактериофагов, которая базируется на влиянии неканонических оснований и генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, которые участвуют в уменьшении возможностей горизонтального переноса генов с генетическим пулом экологических ниш, в результате чего в процессе эволюции образовались T4-родственные фаги, наиболее приспособленные к экспансии одной экологической ниши, и фаги, наиболее приспособленные к распространению по различным экологическим нишам.

**Теоретическая и практическая значимость работы.** На основе пангеномного анализа T4-родственных вирусов определено, что близкородственные T4-родственные фаги обладают одинаковыми генами, которые способствуют воспроизведению фагов с неканоническими основаниями ДНК, что позволяет использовать эти гены в качестве генетических маркеров для задач таксономического определения вирусов.

Определено методами сравнительного геномного анализа общее расположение большинства генов синтеза 5-гидроксиметилцитозина и его модификаций у T4-родственных вирусов в регионе между двумя коргенами – генами ДНК-полимеразы и хеликазы, и наличие в данном регионе разных хоминг-эндонуклеаз, что говорит о широких возможностях горизонтального переноса генов синтеза и модификаций 5-гидроксиметилцитозина у предков T4-родственных вирусов, что может способствовать лучшему пониманию особенностей геномной топологии дцДНК вирусов.

Полученные результаты позволяют сформировать новый подход к выбору агентов фаговой терапии, на основе данных о генах, способствующих воспроизведению фагов с неканоническими основаниями ДНК. Кроме того, полученные результаты показывают, что влияние неканонических оснований на эволюцию вирусов может быть куда значительнее, чем коэволюционные контрадаптации между бактериофагами и бактериями.

**Методология и методы диссертационного исследования.** Методологической основой диссертационного исследования выступают комплексный подход использования аналогий и моделирования, используемый для

научного познания процессов и явлений, происходящих в природе. Диссертационная работа выполнена с использованием современного оборудования, общенаучных и специальных методов исследования, включая методы микробиологии, молекулярной биологии, геномики, биоинформатики, такие как ультрацентрифугирование фаговых частиц, полимеразная цепная реакция с вырожденными и специфическими праймерами, агарозный гель-электрофорез фаговых частиц и ампликонов, спот-тест на ограничение роста бактериофагов на штаммах *E.coli* с системами рестрикции-модификации II типа, секвенирование геномной ДНК фагов, сборка ридов и аннотация собранных геномов вирусов, сравнительный анализ геномов бактериофагов, пан-геномный анализ групп бактериофагов, локальное выравнивание и поиск гомологичных аминокислотных последовательностей из баз данных.

**Положения диссертации, выносимые на защиту:**

1. Существует корреляция между родством T4-родственных вирусов и наличием у них разных генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК.
2. Эволюция *Tevenvirinae* (группы T4-родственных бактериофагов) зависит от экологических ниш и генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК.

**Степень достоверности и апробация результатов.** Результаты работы получены при помощи современных методов и оборудования. Обработка данных и достоверность полученных результатов обеспечивалась использованием современных статистических методов и программ. Результаты являются достоверными и опираются на экспериментальные и литературные данные, имеют статистическую значимость. Результаты, заключения и выводы отвечают целям и задачам диссертационной работы.

Результаты диссертационной работы были доложены и обсуждены на конференциях: I международная (XIV региональная) научная конференция «Техногенные системы и экологический риск» (Обнинск, 2017 г.) (получен диплом II степени); Пущинская школа-конференция «Биохимия, физиология и биосферная роль микроорганизмов» (Пущино, 2019 г., 2021 г., 2022 г.). Также результаты были представлены на конкурсе статей молодых ученых 2021 года МОО «Микробиологическое общество» (получен диплом II степени).

**Публикации.** Материалы диссертационной работы отражены в 12 публикациях, 6 статей из которых опубликованы в журналах, рекомендуемых ВАК Министерства образования и науки РФ для защиты диссертаций, 5 из которых индексированы в библиографической базе Scopus, 2 из которых опубликованы в журналах квартиля Q1.

**Структура диссертации.** Диссертационная работа состоит из введения, обзора литературы, описания материалов и методов, результатов исследования, обсуждения результатов, заключения, выводов, списка цитируемой литературы, приложений. Работа изложена на 172 страницах, включает 46 рисунков, 13 таблиц, 5 приложений. Список литературы включает 186 источников.

**Личный вклад автора.** Личный вклад автора состоял в анализе литературных данных, выполнении и планировании основной части экспериментальной работы и анализе полученных данных, подготовке и публикации работ, представлении результатов на научных конференциях и конкурсах.

Трансмиссионная электронная микроскопия осуществлялась к.б.н. Сузиной Н.Е. (ФИЦ ПНЦБИ РАН (ИБФМ РАН)). Секвенирование геномной ДНК бактериофагов выполнено при участии к.б.н. Воложанцева Н.В., к.б.н. Кисличкиной А.А., к.б.н. Богуна А.Г (ФБУН ГНЦ ПНБ).

**Связь работы с научными программами.** Часть работы выполнена при поддержке гранта РФФИ № 22-25-00669 "Трансдукция плазмид Т4-фагами в условиях, моделирующих природные" (рег. № 122031600276-6).

## **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

### **Материалы и методы**

**Штаммы бактерий и бактериофагов.** Штаммы *Escherichia coli*, использованные в работе: В (Doermann and Hill, 1953; Luria and Anderson, 1942); BL21 (Studier & Moffatt, 1986); 5K (Godessart et al., 1988); 5KRI (Stephanie, 1977); JC5183 (Fouchet et al., 1994); JC5183/pLG13/ (Fouchet et al., 1994; Zakharova et al., 1991). Лабораторные штаммы бактериофагов, использованные в работе: Т4В (Benzer, 1955); Т2L (Hershey, 1946); RB43 (Monod et al., 1997); RB49 (Monod et al., 1997); Т5D7 (Schneider et al., 1985);  $\lambda$  vir (Hayes, 1979)

**Выделение и очистка бактериофагов из фекалий и сточных вод.** Выделение бактериофагов из сточных вод проводилось путем добавления хлороформа к пробе. Для выделения бактериофагов из фекалий к пробе добавляли фаговый раствор (NaCl – 0,2 М, Твин 20 – 1 мг/мл), перемешивали, центрифугированием получали супернатант с фагами и добавляли к нему хлороформ. Очищенные от бактерий пробы высевали методом агаровых слоев. Дополнительно для очистки пересевали отдельную негативную колонию изолята несколько раз.

**ТЭМ фаговых частиц фекалий.** Очищенные от бактерий пробы высевали на чувствительный штамм *E.coli* В, полученные лизаты очищали шприцевыми фильтрами и концентрировали. Проводили трансмиссионную электронную микроскопию. Анализ размеров фаговых частиц производился при помощи ПО ImageJ 1.50i.

**Полимеразная цепная реакция.** Для скрининга изолятов на наличие генетического маркера Т4-родственных бактериофагов использовались вырожденные праймеры, комплиментарные последовательности гена главного белка головки, свойственной данной группе вирусов (Filée et al., 2005). Для скрининга обнаруженных Т4-родственных бактериофагов на выявление у них генетических маркеров фагов близкородственных Т4, использовались праймеры, которые отжигаются на фрагменте гена РНК-лигазы 2 фага Т4 и последовательности 3 домена Ig-подобного белка НОС фага Т4.

**Спот-тест на ограничение роста фагов на штаммах с RM-системами.** Для проведения спот-теста на ограничение роста фагов на штаммах с RM-системами использовались изогенные чувствительные штаммы *E.coli*, отличающиеся по наличию/отсутствию RM-систем. Проверляли наличие или отсутствие роста бактериофагов из очищенных лизатов на парных штаммах.

**Электрофорез фаговых частиц в агарозном геле.** Лизаты наносились на 1% агарозный гель. Проводился горизонтальный электрофорез при напряженности электрического поля 9 В/см. Затем, гель окрашивали раствором кумасси G-250.

**Выделение ДНК.** Выделение ДНК проводили путем стандартного метода фенольной экстракции из очищенных лизатов фагов.

**Секвенирование, сборка и аннотация геномов.** Секвенирование производилось на платформе Illumina MiSeq. Сборку ридов осуществляли Unicycler v.0.4.7 (Wick et al., 2017). Аннотирование контигов геномов осуществлялось при помощи инструмента Genome Annotation сервиса PATRIC (Wattam et al., 2014). Аннотации проверяли при помощи GET\_HOMOLOGUES (Contreras-Moreira and Vinuesa, 2013).

**Сравнительный анализ геномов.** Аннотации геномов из баз данных проверялись на наличие открытых рамок считывания (ORF), при необходимости проводилась аннотация/реаннотация. Гомологи аминокислотных последовательностей ORF геномов были кластеризованы с помощью скрипта GET\_HOMOLOGUES (get\_homologues.pl) с параметрами, подобранными в ходе анализа. Для анализа гомологичных белков использовали пангеномную матрицу наличия/отсутствия кластеризованных последовательностей. Состав пангенома определяли с помощью скрипта parse\_rangenome\_matrix.pl. На основе бинарной матрицы наличия/отсутствия кластеров получали пангеномное дерево с использованием GET\_PHYLOMARKERS (Vinuesa et al., 2018). Филогенетические деревья были получены с помощью W-IQ-TREE (Trifinopoulos et al., 2016) или MEGAX (Kumar et al., 2018). Филогеномные деревья были построены с использованием VICTOR (Meier-Kolthoff and Göker, 2017). В Cytoscape (Shannon et al., 2003) были визуализированы двудольные кластер-геномные сети. Поиск гомологов кластеров, специфичных для определенной клады геномов, полученных на пангеномном дереве, и кластеров, связанных или предположительно связанных с неканоническими основаниями, проводили с помощью PSI-BLAST (Altschul et al., 1997; Schäffer et al., 2001) и HMMER (Zimmermann et al., 2018). Отдельные регионы геномов, содержащие гены, связанные с синтезом неканонических оснований, сравнивали с помощью Easyfig (Sullivan et al., 2011).

### **Результаты и обсуждение**

#### **Бактериофаги из фекалий зубров и сточных вод очистных сооружений.**

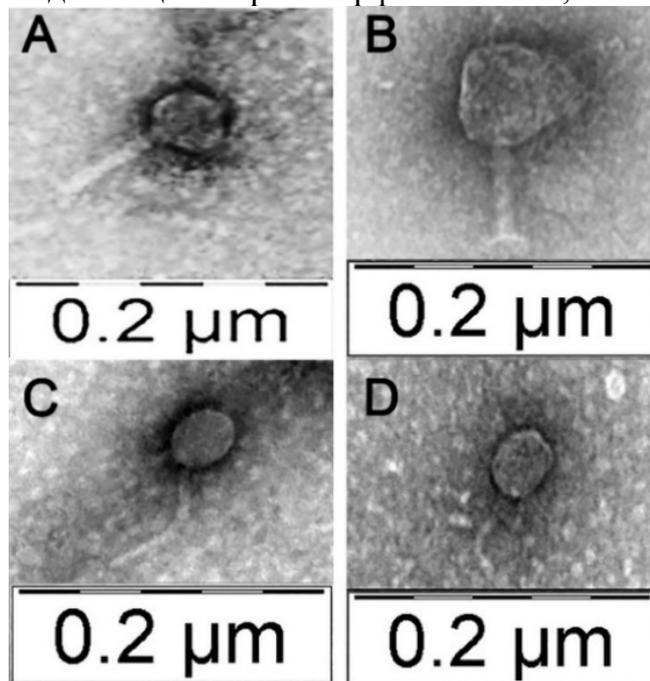
Отбор проб фекалий зубров для выделения бактериофагов производилось в Приокско-Тerrasном государственном биосферном заповеднике имени М. А. Заблочного в питомнике зубров. Отбор проб сточных вод проводился на очистных сооружениях города Пущино Московской области на различных этапах очистки воды: неочищенная вода, первичные отстойники, аэротенки, вторичные отстойники.

При высеве проб из сточных вод различной степени очистки на чувствительный штамм *E.coli* В отмечалось образование в основном негативных колоний от 0,8 до 1,5 мм мутных или прозрачных с нечетко очерченными краями. При высеве проб из сточных вод с очистных сооружений с различных этапов очистки на штамм *E.coli* BL21 в основном обнаруживались прозрачные негативные колонии от 0,5 до 1,7 мм. В целом отмечалось образование схожих колоний на штаммах В и BL21. Учитывая наличие у штамма В системы рестрикции-модификации, что может способствовать увеличению числа фагов с системами антирестрикции, а также то, что данный штамм исходно использовался для выделения первых представителей T4-родственных бактериофагов, для дальнейшего анализа были выделены отдельные негативные колонии, отличающиеся по морфологии, образующиеся на штамме В (50 бляшек).

Таким же образом были выделены 12 фагов (из-за меньшего разнообразия морфологии негативных колоний), полученных со штамма В, на который

высевались очищенные пробы фекалий зубров. После очистки изолятов путем пересевов, была создана рабочая коллекция.

Также проводился анализ фаговых частиц из пробы фекалий зубров, инфицирующих *E.coli* В. Анализ снимков результатов ТЭМ сконцентрированных и очищенных лизатов со штамма В фаговых частиц из фекалий зубров показал наличие частиц, принадлежащих к трем морфотипам: А1, А2 и В2 (рисунок 1).



**Рисунок 1** – фаговые частицы из фекалий зубров. А – морфотип А1 (длина и ширина головки равны примерно 0,062 мкм, длина хвоста равна примерно 0,075 мкм), В – морфотип А2 (длина головки равна примерно 0,077 мкм, ширина – примерно 0,062 мкм, длина хвоста – примерно 0,075 мкм), С – морфотип В1 (длина головки равна примерно 0,045 мкм, ширина – примерно 0,037 мкм, длина хвоста – примерно 0,089 мкм), D – морфотип В1 (длина головки равна примерно 0,045 мкм, ширина – примерно 0,037 мкм, длина хвоста – примерно 0,061 мкм).

В исследованной пробе преобладали частицы морфотипа А2 (рисунок 1В). Хотя бактериофаги размером, свойственным Т-четным или псевдо-Т-четным (в этих группах были обнаружены фаги, инфицирующие *E.coli*), не были обнаружены, что могло быть связано с отсутствием таких фагов в отобранной пробе, наличие Т4-родственных бактериофагов в фекалиях зубров не исключается, учитывая, что их часто обнаруживают в фекалиях животных.

В целом была произведена первичная характеристика бактериофагов из сточных вод очистных сооружений города Пущино и фекалий зубров. Определено наличие бактериофагов, образующих различные негативные колонии на штамме *E.coli* В из этих проб. Обнаружено преобладание бляшек относительно малого размера, что свойственно Т4-родственным бактериофагам. Получена рабочая коллекция бактериофагов из сточных вод и фекалий зубров, состоящая из 62 изолятов.

**Скрининг бактериофагов рабочей коллекции на принадлежность к Т4-родственным бактериофагам и наличие у них антирестрикционных систем.** При проведении спот-теста на ограничение роста лабораторных штаммов фагов на штаммах *E.coli* с системами рестрикции-модификации было обнаружено три возможных ответа: рост без ограничений по сравнению с изогенным штаммом без RM-систем; отсутствие роста на штаммах с RM-системами; ограничение роста по

сравнению с изогенным штаммом без RM-систем. При проведении спот-теста на ограничение роста лабораторных штаммов фагов на штаммах *E.coli* с системами рестрикции-модификации было обнаружено два возможных ответа: рост без ограничений по сравнению с изогенным штаммом без RM-систем; отсутствие роста на штаммах с RM-системами. По результатам спот-теста можно отметить, что у подавляющего большинства фагов из рабочей коллекции (58 фагов из 62) рост на штаммах с RM-системами не ограничивался.

При проведении ПЦР с вырожденными праймерами, специфическими для генов главного белка капсида T4-родственных фагов, а также ПЦР со специфическими праймерами к последовательности ДНК вида *Tequatrovirus* T4 было обнаружено, что подавляющее большинство фагов рабочей коллекции (58 из 62) имели в составе геномной ДНК последовательность гена главного белка капсида, свойственного T4-родственным бактериофагам, а также 51 из них имели в составе геномной ДНК консервативную последовательность, свойственную близкородственным виду *Tequatrovirus* T4.

Сравнение результатов, полученных при ПЦР и спот-тесте, представлены в таблице 1.

**Таблица 1.** Результаты спот-теста на ограничение роста на штаммах *E.coli* с RM-системами и ПЦР анализов для выявления принадлежности к T4-родственным бактериофагам и к близкородственным виду *Tequatrovirus* T4 лабораторных штаммов вирусов и штаммов рабочей коллекции.

Фаги	Принадлежность к T4-родственным	Принадлежность к близкородственным T4	Рост на штаммах с RM-системами	
			EcoRI	EcoRV
T4B	+	+	не ограничен	не ограничен
RB43	+	-	ограничен	ограничен
RB49	+	-	не ограничен	не ограничен
T2L	+	+	не ограничен	не ограничен
$\lambda$ vir	-	-	отсутствует	отсутствует
1B-2B	-	-	отсутствует	отсутствует
3B-10B	+	+	не ограничен	не ограничен
11B	-	-	отсутствует	отсутствует
12B	+	+	не ограничен	не ограничен
PuMWB1- PuMWB8	+	+	не ограничен	не ограничен
PuMWB9	-	-	неизвестно	отсутствует
PuMWB10	+	+	не ограничен	не ограничен
PuMWB11	+	-	не ограничен	не ограничен
PuMWB12	+	+	не ограничен	не ограничен
PuMWB13	+	-	не ограничен	не ограничен
PuMWB14- PuMWB20	+	+	не ограничен	не ограничен
PuMWB21	+	-	не ограничен	не ограничен
PuMWB22 - PuMWB23	+	+	не ограничен	не ограничен
PuMWB24	+	-	не ограничен	не ограничен
PuMWB25- PuMWB29	+	+	не ограничен	не ограничен
PuMWB30	+	-	не ограничен	не ограничен

Таблица 1. Продолжение

Фаги	Принадлежность к T4-родственным	Принадлежность к близкородственным T4	Рост на штаммах с RM-системами	
			EcoRI	EcoRV
PuMWB25- PuMWB29	+	+	не ограничен	не ограничен
PuMWB30	+	-	не ограничен	не ограничен
PuMWB31 - PuMWB35	+	+	не ограничен	не ограничен
PuMWB36	+	-	не ограничен	не ограничен
PuMWB37- PuMWB39	+	+	не ограничен	не ограничен
PuMWB40	+	-	не ограничен	не ограничен
PuMWB41- PuMWB50	+	+	не ограничен	не ограничен

Первый столбец – через дефис представлены названия фагов, чья нумерация идет по порядку, и они имеют одинаковые результаты спот-теста и ПЦР. Второй столбец – принадлежность фага к T4-родственным по результатам ПЦР; третий столбец – принадлежность фагов к близкородственным виду Tequatrovirus T4; «+» – наличие целевого ампликона, «-» – отсутствие целевого ампликона. «Не ограничен» – фаги одинаково растут на изогенных штаммах с и без RM-системы; «ограничен» – рост на изогенных штаммах с и без RM-системы отличается; «отсутствует» – не растут на штамме с RM-системой, без – растут; «неизвестно» – не растут на изогенных штаммах с и без RM-системы.

В результате было определено, что большинство фагов рабочей коллекции – 58 из 62 фагов – принадлежало к T4-родственным бактериофагам, имеющим антиристрикционные системы. Причем только 4 фага не принадлежали к T4-родственным и не росли на штаммах с RM-системами. Среди T4-родственных бактериофагов рабочей коллекции большинство – 51 из 58 фагов являлись близкородственными виду T4. Лишь 7 фагов: PuMWB11, PuMWB13, PuMWB21, PuMWB24, PuMWB30, PuMWB36, PuMWB40, не имели целевого ампликона, свойственного близкородственным виду T4.

Дополнительно, для определения различий у семи T4-родственных бактериофагов рабочей коллекции, фагов, не являющихся близкородственными виду T4 и содержащих антиристрикционные системы, был проведен электрофорез их частиц (рисунок 2).

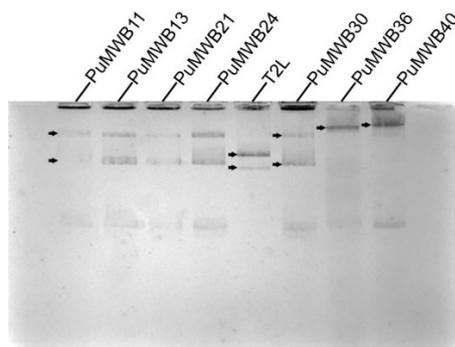
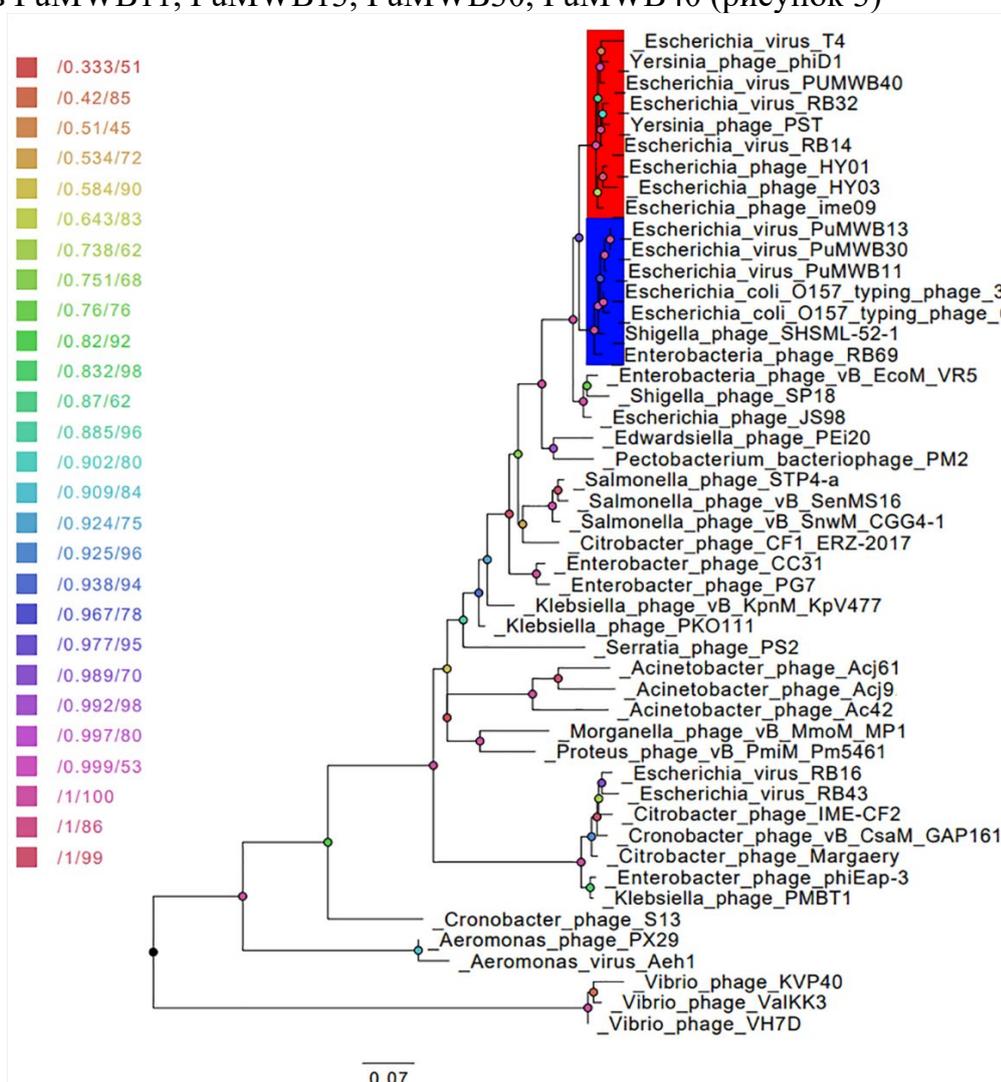


Рисунок 2 – Электрофореграмма фаговых частиц T4-родственных бактериофагов. Стрелками отмечены целевые полосы. T2L – лабораторный штамм фага близкородственного виду T4; PuMWB11, PuMWB13, PuMWB21, PuMWB24, PuMWB30, PuMWB36, PuMWB40 – T4-родственные бактериофаги рабочей коллекции, не являющиеся близкородственными виду T4 и содержащие антиристрикционные системы.

По результатам электрофореза частиц T4-родственных бактериофагов рабочей коллекции, не являющихся близкородственными виду T4 и содержащие

антиристрикционные системы, можно судить о наличии 3 вариантов электрофоретической подвижности этих вирусов: вариант 1 (PuMWB11, PuMWB13, PuMWB21, PuMWB24, PuMWB30); вариант 2 (PuMWB36), вариант 3 (PuMWB40). Для дальнейшего анализа была взята три бактериофага с вариантом подвижности 1 и один бактериофаг – с вариантом 3 для того, чтобы дополнительно проверить эффективность метода электрофореза фаговых частиц и провести анализ геномов.

После получения нуклеотидных последовательностей геномов выбранных T4-родственных бактериофагов рабочей коллекции, не являющихся близкородственными виду T4 и содержащих антиристрикционные системы, были определены длины их геномов: PuMWB11 – 166676 п.о., PuMWB13 – 166966 п.о., PuMWB30 – 166964 п.о., PuMWB40 – 167232 п.о. Для определения принадлежности к родам подсемейства *Tevenvirinae* фагов PuMWB11, PuMWB13, PuMWB30, PuMWB40 было построено пангеномное дерево геномов классифицированных представителей T4-родственных бактериофагов, взятых из базы данных RefSeq, и геномов PuMWB11, PuMWB13, PuMWB30, PuMWB40 (рисунок 3)



**Рисунок 3** – Пангеномное дерево геномов классифицированных представителей T4-родственных бактериофагов, взятых из базы данных RefSeq, и геномов PuMWB11, PuMWB13, PuMWB30, PuMWB40. Красным и синим отмечены клады *Tequatrovirus* и *Mosigvirus* соответственно. Цветными кружками отмечена поддержка (значения для aBayes support и ultrafast bootstrap support представлены слева).

Анализ пангеномного дерева показал, что PuMWB11, PuMWB13, PuMWB30 принадлежат к роду *Mosigvirus*, PuMWB40 – *Tequatrovirus*.

На основе полученных данных о родстве PuMWB11, PuMWB13, PuMWB30, PuMWB40 можно судить о том, что проведенный ранее электрофорез фаговых частиц коррелирует с результатами анализа геномов выбранных вирусов, так как он позволил отличить по электрофоретической подвижности бактериофаги одного вида (PuMWB11, PuMWB13, PuMWB30) от фагов другого вида (PuMWB40), относящиеся к разным родам (*Mosigvirus* и *Tequatrovirus* соответственно).

При рассмотрении геномов фагов PuMWB11, PuMWB13, PuMWB30 и PuMWB40, было найдено, что набор генов, участвующих в синтезе и модификации неканонических оснований, отличаются в зависимости от рода. В связи с этим было проведено детальное исследование геномов T4-родственных бактериофагов из базы данных NCBI Nucleotide Collection для выявления особенностей наследования генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК (белки, участвующие в защите от действия RM-систем IV типа; белки, гидролизующие ДНК, не содержащую неканонических замен dC; белки, влияющие на транскрипцию ДНК бактерии-хозяина и бактериофага), и выявления влияния на эволюцию данной группы вирусов этих генов.

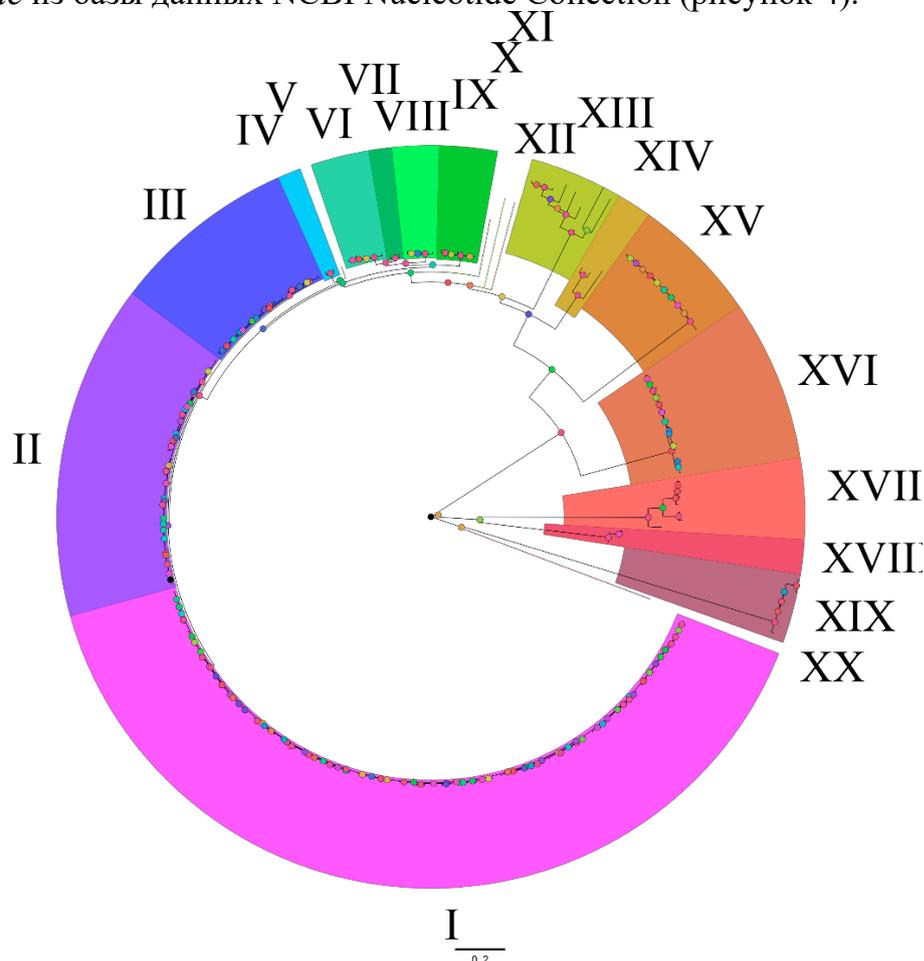
**Анализ геномов T4-родственных бактериофагов из баз данных.** Для исследования вклада ДНК с неканоническими основаниями в эволюцию T4-родственных бактериофагов, использовались геномы из NCBI Nucleotide Collection, являющиеся представителями *Tevenvirinae*, которые принадлежали к данному подсемейству на 2019 год. Проведено несколько кластеризаций геномов, до образования всеми геномами гомологичных кластеров коргенов. Суммарно было использовано 57499 аминокислотных последовательностей, получено 4530 гомологичных кластеров: в кор вошел 61 кластер, в софтвер – 76 (с учетом коргеномных кластеров), в оболочку - 1966, в облако – 2488 кластеров; 79 геномов были реаннотированы, 7 из них удалены из анализа и всего 199 геномов были использованы для дальнейшего анализа.

Для выявления эволюционных взаимосвязей бактериофагов подсемейства *Tevenvirinae* были построены филогенетическое дерево гомологов главного белка капсида, пан-геномное и филогеномное деревья. Филогеномное дерево использовалось для отражения генетического разнообразия наиболее близких по родству вирусов, тогда как пан-геномное дерево использовалось для отражения генетического разнообразия более удаленных по родству вирусов. Филогенетическое дерево гомологов использовалось для лучшего отражения вертикального переноса генов между родственными вирусами, так как данный белок является одним из наиболее консервативных из вирусных белков.

Для выявления экологических взаимосвязей бактериофагов подсемейства *Tevenvirinae* были построены сеть оболочки пан-генома и филогенетическое дерево РНК-лигазы 2. Сетевой анализ полнее, чем деревья, позволяет отразить горизонтальный обмен генами между родственными вирусами, кроме того сетевой анализ оболочки пан-генома отражает экологические взаимосвязи между вирусами за счет того, что гены данного компонента пан-генома в основном имеют функции, связанные с приспособлением к определенным условиям среды. Выбор РНК-лигазы 2 с точки зрения исследования экологических взаимосвязей был выбран по нескольким причинам. Во-первых, гены данного белка принадлежат к оболочке пан-генома. Во-вторых, гомологи генов РНК-лигазы 2 присутствуют в большинстве рассматриваемых геномов (188 геномов из 199) *Tevenvirinae*. За счет этих причин

возможно рассмотрение наследования гена оболочки пан-генома у относительно крупного числа фагов и выявление особенностей наследования таких генов.

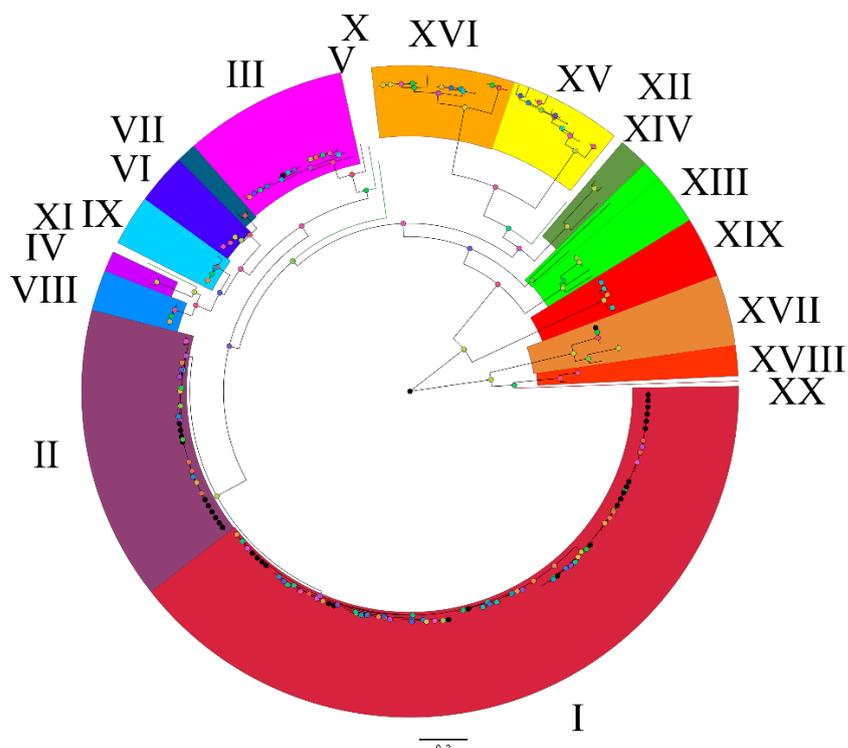
На основе бинарной матрицы наличия/отсутствия кластеров гомологичных аминокислотных последовательностей построено пангеномное дерево геномов *Tevenvirinae* из базы данных NCBI Nucleotide Collection (рисунок 4).



**Рисунок 4** – пангеномное дерево геномов фагов подсемейства *Tevenvirinae* из баз данных NCBI Nucleotide Collection. Лучшая модель для построения: GTR2+FO+I+G4. Log-likelihood консенсусного дерева:  $-41953,89217$ . Цветом отмечены группы, цветными кружками – поддержка ветвей. Римскими цифрами отмечены группы бактериофагов.

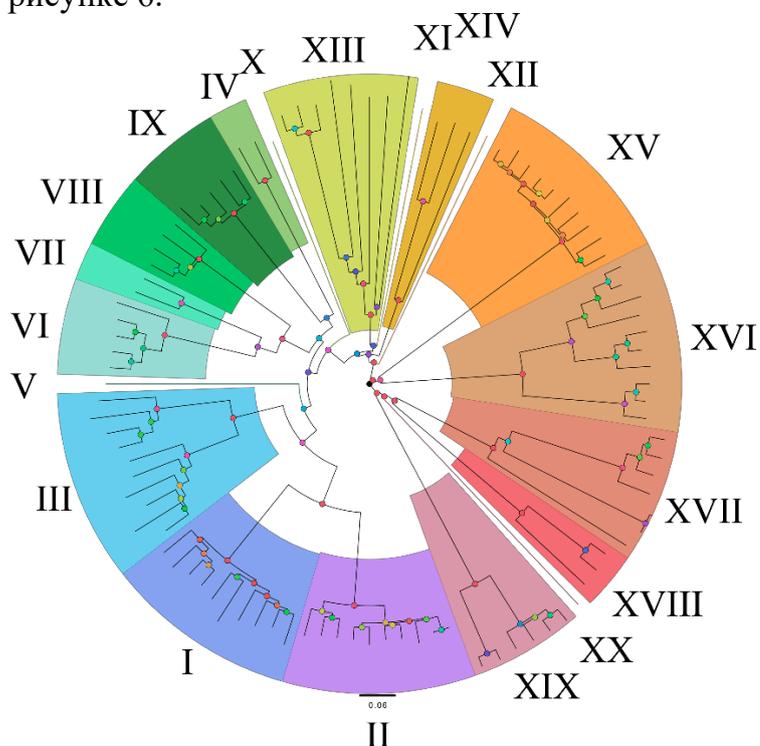
Анализируя пангеномное дерево, было выделено 20 групп *Tevenvirinae*. Для разделения геномов на отдельные группы использовались следующие критерии: каждая группа образовывала отдельную кладу на пангеномном дереве, каждая группа имеет ГСК (группоспецифические кластеры – гомологичные кластеры с последовательностями нуклеотидов, присутствующими во всех геномах группы и отсутствующими в геномах других групп), принималась во внимание таксономическая классификация фагов. Как показал анализ пан-геномного дерева и данные об источниках выделения фагов и чувствительных бактериях (представлены в приложении 3 кандидатской диссертации), фаги из групп, представленных в данном анализе более чем одним геномом, инфицировали родственных (на уровне порядка, семейства или рода) бактерий-хозяев: *Enterobacteriales* ( I–III, VIII, XV, XVI), *Morganellaceae* (XIV), *Klebsiella* (IV, IX), *Salmonella* (VI), *Citrobacter* (VII), *Acinetobacter* (XIII), *Aeromonas* (XVII, XVIII), *Vibrio* (XIX), а также выделены из схожих экологических ниш.

На рисунке 5 показано филогенетическое дерево, построенное на основе гомологичных последовательностей основного белка капсида.



**Рисунок 5** – филогенетическое дерево гомологов главного белка капсида фагов подсемейства *Tevenvirinae* из базы данных NCBI Nucleotide Collection. Лучшая модель для построения: LG+F+I+G4. Найден лучший результат:  $-15512,735$ . Цветом отмечены группы, цветными кружками – поддержка ветвей.

Было построено филогеномное дерево 199 геномов *Tevenvirinae*. Оно представлено на рисунке 6.

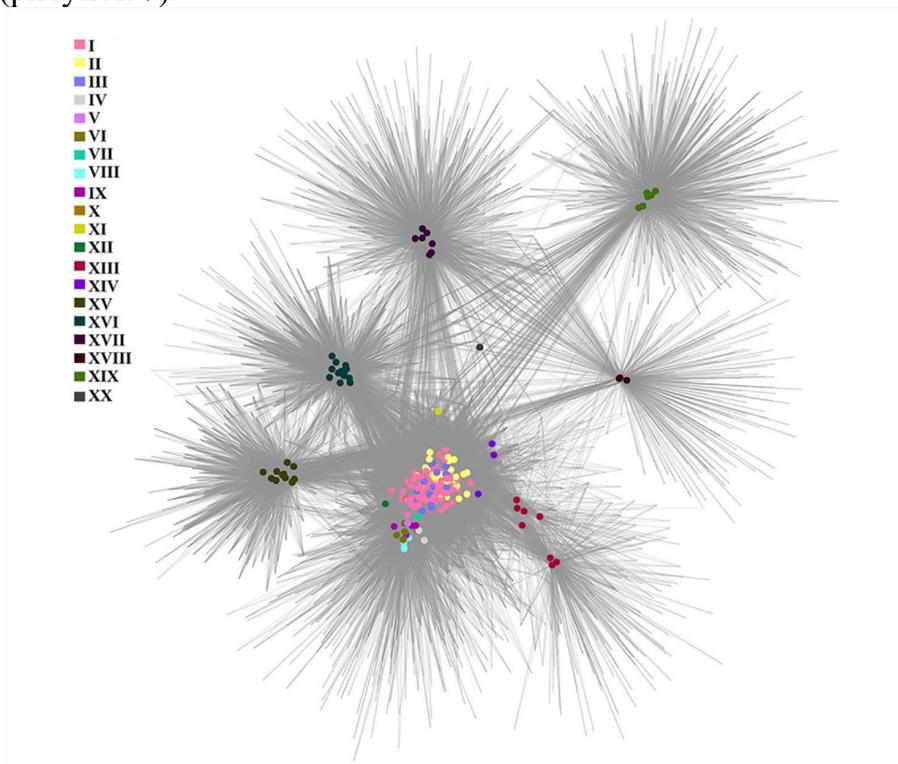


**Рисунок 6** – филогеномное дерево геномов подсемейства *Tevenvirinae* из базы данных NCBI Nucleotide Collection. Цветом и римскими цифрами отмечены группы, цветными кружками – поддержка ветвей.

Основные различия между деревьями заключались в иерархии клад и иерархии ветвей отдельных представителей внутри клад. Эти различия могут быть связаны с тем, что пангеномное и филогеномное деревья отражают как

горизонтальный, так и вертикальный обмен генетической информацией между фагами *Tevenvirinae*, тогда как филогенетическое дерево главного белка капсида отражает вертикальный обмен генетической информации между ними.

Была построена двудольная сеть геном/гомологичный кластер оболочки пангенома (рисунок 7).



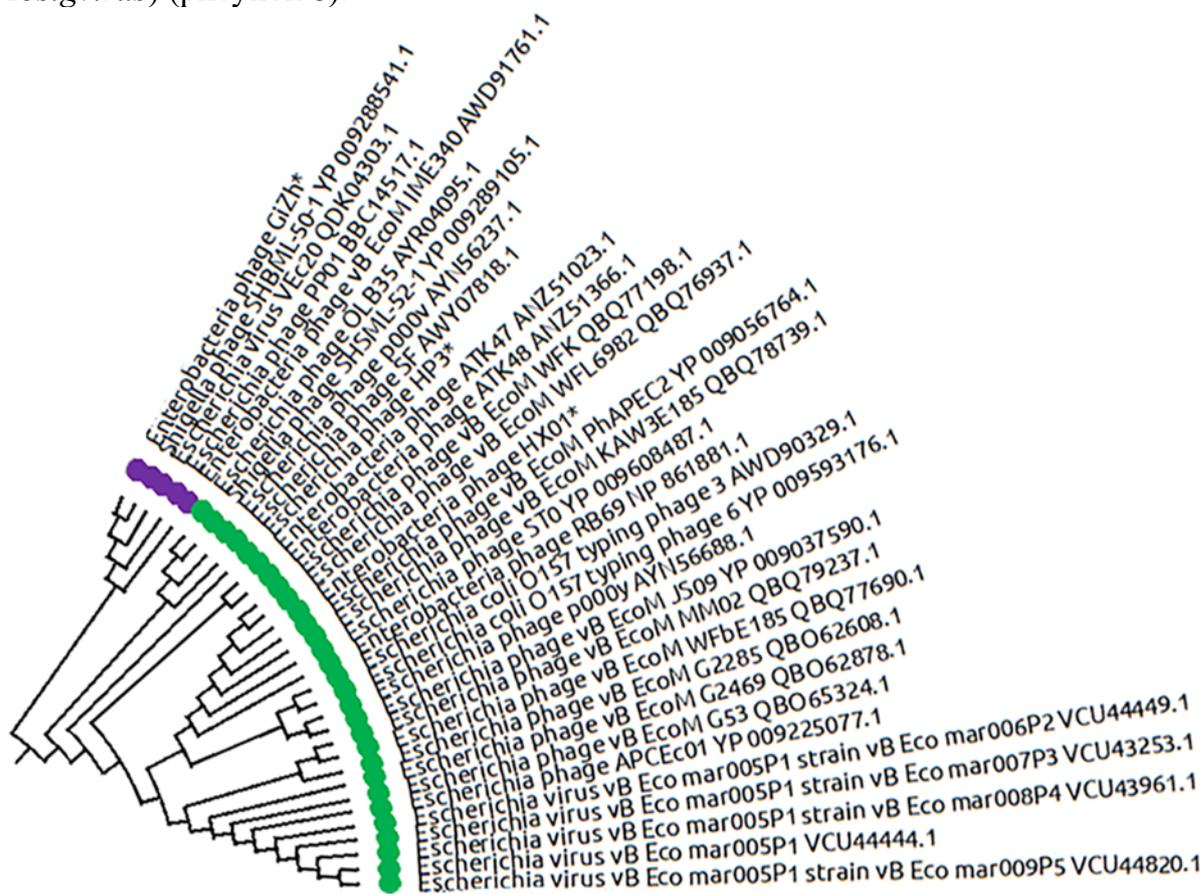
**Рисунок 7** – двудольная сеть оболочки пангенома. Цветными кружками отмечены узлы геномов *Tevenvirinae* из базы данных NCBI Nucleotide Collection, цвет кружка отражает отношение к группе, ребра отражают связи между отдельными геномами и кластерами, узлы кластеров не отмечены.

Двудольная сеть в большинстве своем аналогична пангеномному и филогеномному деревьям, но имеются некоторые отличия. Самым большим отличием было расположение отдельных представителей групп I–XIV, которые образовывали плотную подсеть отдельно от других групп. Скорее всего, это можно объяснить спецификой их приспособления к среде, зависящей от генов оболочки пангенома и их продуктов. Однако группы XV и XVI были далеки от групп, представители которых занимали схожие экологические ниши. Группы XVII и XVIII, фаги которых заражают *Aeromonas* и были выделены из сходных источников, также были далеко друг от друга. Учитывая особенности оболочки пангенома, можно предположить, что общие предки вирусов из групп XV и XVI, XVII и XVIII имели барьер, который блокировал активный обмен генов между фагами одной экологической ниши.

Было построено филогенетическое дерево гомологов РНК-лигазы 2.

На основе результатов анализа оболочки пан-генома можно отметить, что РНК-лигаза 2 отсутствует лишь у одной выделенной группы *Tevenvirinae* – группы XV, которую представляют вирусы рода *Krischvirus*. Роль данного фермента у рассматриваемого подсемейства вирусов неизвестна, однако по результатам является важной для существования в различных экологических нишах. Лишь особые адаптации вирусов группы XV, не свойственные другим группам, позволили убрать гомологи гена РНК-лигазы 2 из геномов, дальнейшие исследования отличий

фагов *Krischvirus* от других представителей *Tevenvirinae* может помочь в выяснении биологической роли РНК-лигазы 2. В отличие от филогенетического дерева гомологов главного белка капсида, на филогенетическом дереве гомологов РНК-лигазы 2 часть группы I (*Tequatrovirus*) образовывали общую кладу с группой II (*Mosigvirus*) (рисунок 8).



**Рисунок 8** – часть дерева гомологов РНК-лигазы 2 у *Tevenvirinae* из базы данных NCBI Nucleotide Collection. Цветными кругами отмечены рода *Tevenvirinae*: фиолетовые – *Tequatrovirus* (группа I), зеленые – *Mosigvirus* (группа II).

Такое явление также отражается на сети оболочки пан-генома, где вирусы данной группы образуют общую плотную подсеть. Как уже отмечалось, это связано с тем, что представители данных групп занимают схожие экологические ниши и инфицируют схожие бактерии (представителей *Enterobacteriales*) (приложение 3). Таким образом, можно отметить, что ген РНК-лигазы 2 подвергся горизонтальному переносу между фагами *Tevenvirinae*, занимающих схожие экологические ниши.

Чтобы получить больше информации об идентифицированных группах фагов, были проанализированы кластеры гомологов, специфичных для каждой группы (ГСК – группоспецифичные кластеры) и проверено наличие их у других организмов. Рассмотрение организмов и вирусов, белки которых имели последовательности, гомологичные с ГСК, показало, что большинство этих организмов и вирусов имеют сходные экологические ниши с фагами той или иной группы *Tevenvirinae*.

Сравнение результатов анализа пангеномного дерева, филогеномного дерева, ГСК и двудольной сети показывает, что эволюционная дивергенция геномов *Tevenvirinae* в большей степени определяется характеристиками их экологических ниш. Анализируя ГСК группы XIV, мы обнаружили гомологи, участвующие в синтезе и модификации 7-деазагуанина. Принимая во внимание это, а также наличие

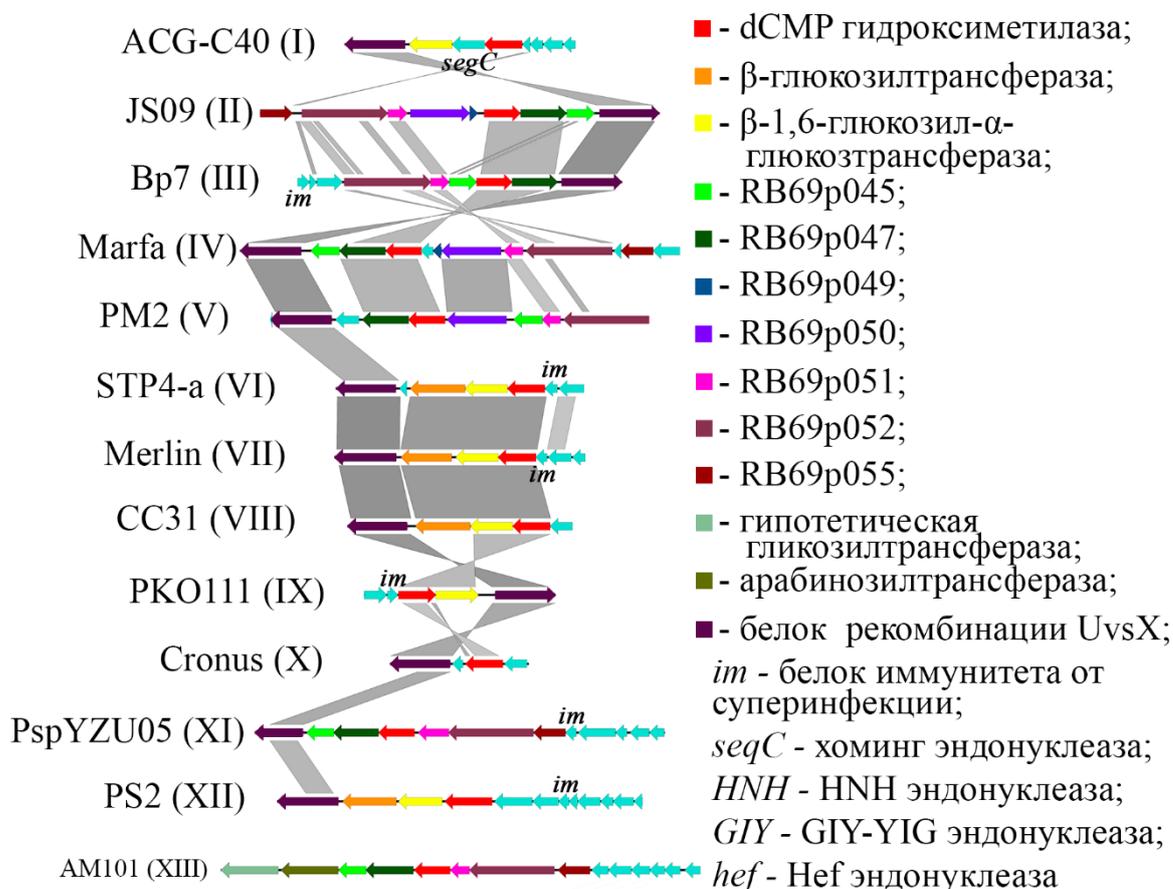
других неканонических оснований у фагов *Tevenvirinae*, найденные особенности в геномах выделенных фагов, было предположено, что появление разных оснований и их модификаций у определенных групп предков *Tevenvirinae* было одним из факторов, обуславливающих дивергенцию геномов у фагов этого подсемейства. Чтобы подтвердить это, мы рассмотрели гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК. В таблице 2 представлены эти гены и их продукты.

**Таблица 2.** Гены и их белки, способствующие воспроизведению фагов *Tevenvirinae* с неканоническими основаниями ДНК.

Ген/локус	Белок/возможный белок (домен)	Функция/возможная функция
42	НМС-трансфераза	синтез <sup>5hm</sup> dCMP
<i>α-gt</i>	α-глюкозилтрансфераза	α-глюкозилирование <sup>5hm</sup> dCMP в ДНК
<i>β-gt</i>	β-глюкозилтрансфераза	β-глюкозилирование <sup>5hm</sup> dCMP в ДНК
<i>βα-gt</i>	β-глюкозил-1,6-α-глюкозилтрансфераза	β-глюкозилирование α-глюкозилированного <sup>5hm</sup> dCMP в ДНК
<i>RB69p003c</i>	арабинозилтрансфераза	арабинозилирование <sup>5hm</sup> dCMP в ДНК
<i>denB</i>	эндонуклеаза IV	расщепление немодифицированной (dCMP-содержащей) ДНК
<i>Acj61p076</i>	предположительно гликозилтрансфераза	модификация неканонического основания
<i>RB69p052</i>	гомологи среди аминокликозид-3'-фосфотрансфераз	вероятно, участвует в синтезе УДФ-арабинозы
<i>RB69p055</i>	арабинозил-5-фосфатизомераза	
56	дЦТФаза-дУТФаза	увеличивает пул dCMP, уменьшает пул dCTP;
<i>RB69p047</i>	n/a	предположительно участвуют в синтезе арабинозил- <sup>5hm</sup> dCMP.
<i>RB69p049</i>	n/a	
<i>RB69p050</i>	гомологи среди пептидаз	
<i>RB69p051</i>	присутствует фосфатазный домен полинуклеотидкиназы	
<i>PX29p085</i>	N-ацетилтрансфераза семейства GNAT	
<i>Aeh1ORF087c</i>	глюкозилтрансфераза	Предположительно модифицирует ДНК
<i>folE</i>	белок FolE	необходимы для синтеза 7-циано-7-дезагуанина (PreQ0)
<i>KVP40.0120</i>	белок QueD	
<i>KVP40.0124</i>	белок QueC	
<i>KVP40.0284</i>	белок QueE	
<i>KVP40.0122</i>	белок DpdA2	необходим для вставки PreQ0 и/или PreQ1 в ДНК
<i>KVP40.0123</i>	белок QueF или QueF-L	синтез дезоксиархеозина или 2'-дезоксидезокси-7-аминометил-7-дезагуанина

Обнаружено, что близкородственные бактериофаги содержат одинаковый набор генов, способствующих воспроизведению фагов *Tevenvirinae* с неканоническими основаниями ДНК.

Для фага RB69 показано, что большинство генов, продукты которых участвуют или предположительно участвуют в арабинозилровании  $^{5\text{hm}}\text{dC}$  в ДНК, расположены в области между генами, кодирующими ДНК-полимеразу и UvsX. Анализ фагов, имеющих гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, показал, что расположение генов, кодирующих эти белки, весьма сходно. Данную область фланкировали гены шаперона сборки верхушек капсида (в некоторых случаях отсутствовал UvsX) и ДНК-полимеразы. Были проанализированы фрагменты геномов между ними. На рисунке 9 показаны вариации этой области, выявленные у близкородственных фагов *Tevenvirinae*.



**Рисунок 9** – Примеры вариации области между генами ДНК-полимеразы и шаперона сборки верхушек капсида у близкородственных *Tevenvirinae* из базы данных NCBI Nucleotide Collection, имеющих гомологи белков синтеза  $^{5\text{hm}}\text{dC}$ .

Следует отметить, что у *Acinetobacter* phage Aсj9 гены, участвующие в синтезе и модификации неканонических оснований, расположены в более широкой области генома: между генами хеликазы и ДНК-полимеразы (оба являются коргенами). У остальных фагов эти гены располагались в рассмотренном выше регионе. Интересно, что у фагов *Mosigvirus*, ген арабинозилтрансферазы находится вне рассматриваемой области. Однако у *Acinetobacter* phage AM101 и фага *Acinetobacter* phage Aсjб1 есть два гомолога гена этого фермента, и их гены расположены в пределах рассмотренной области. У большинства близкородственных фагов вариации в области генома между генами хеликазы и ДНК-полимеразы связаны с

наличием или отсутствием генов, которые предположительно кодируют хоминг-эндонуклеазы. Можно предположить, что гены, участвующие в синтезе и модификации неканонических оснований, подвергались или подвергаются частому горизонтальному переносу.

Сравнительный анализ наборов белков, ассоциированных с неканоническими основаниями, и взаимоотношений между фагами *Tevenvirinae* показал, что близкородственные фаги имеют ферменты, участвующие в синтезе одних и тех же или химически схожих неканонических оснований. Как известно, бактериофаг T4 не способен к общей трансдукции (Wilson et al., 1979); кроме того, он исключается из инфекции бактериофагом RB69, имеющим другую модификацию <sup>hm5</sup>dC (Russell, 1967). Также в (Wilson et al., 1979) на мутантах фага T4 было показано, что частота трансдукции меняется в зависимости от наличия или отсутствия белков, способствующих воспроизведению фагов с неканоническими основаниями ДНК. Учитывая, что некоторые из этих белков влияют на ДНК хозяина, можно предположить, что эти белки также будут изменять частоту других событий горизонтального переноса генов. Эти факты и полученные результаты позволяют предположить, что неканонические основания и гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, образовали барьер для обмена генетической информацией между родственными фагами, что привело к их дивергенции в процессе эволюции.

Принимая во внимание, что неканонические основания и гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, являются одним из механизмов регуляции частоты горизонтального переноса генов у *Tevenvirinae*, меняют специфичность узнавания их ДНК системами рестрикции-модификации бактерий-хозяев, и на основании результатов, полученных в нашем исследовании, было предположено, что они играют одну из ключевых ролей в разнообразии геномов этих вирусов.

Суммируя результаты проведенного анализа и данные о барьерных функциях неканонических оснований и генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, можно представить возможную цепочку событий, приведшую к дивергенции геномов предков *Tevenvirinae*. Первым этапом было заражение различных бактерий-хозяев: *Vibrio*, *Aeromonas*, *Enterobacteriales* и др. близкородственными фагами, и, как следствие, подобные фаги широко распространились в различных типах местообитаний: морской среде, почве, растениях и животных. Адаптация к специфическим условиям местообитаний предполагала изменения в геномах фагов и бактерий-хозяев, что позволило бы выработать специфические метаболические пути.

Вторым этапом стало появление различных неканонических оснований в ДНК предков *Tevenvirinae*. Это событие сыграло ключевую роль в дивергенции фаговых геномов в пределах одной экологической ниши. Фаги с различным набором неканонических оснований и их модификаций все больше специализировались на заражении определенных близкородственных бактерий (отличающихся по системам защиты от чужеродной ДНК). Далее предки *Tevenvirinae*, населявшие одни и те же экологические ниши, начали расходиться в способах обмена генетической информацией. Некоторые из них сохранили способность обмениваться генетической информацией с организмами, ДНК которых имеет канонический состав оснований. Другие начали приобретать гены, которые препятствовали ГПГ. В результате геномы таких фагов стали более “консервативными”. Кроме того,

появление генов, продукты которых защищали фагов с неканоническими основаниями от действия RM-систем типа IV, расширило возможности этих фагов по инфицированию бактерий. В результате, появлялись популяции, специализирующиеся на экспансии одной экологической ниши, популяции эффективно занимающие разные экологические ниши.

Для рассмотрения такой стратегии эволюции *Tevenvirinae* с точки зрения эколого-эволюционных моделей была использована модификация модели «убить победителя» – модель Постоянного Разнообразия. Модель Постоянного Разнообразия предполагает сосуществование конкурирующих популяций фагов в одной и той же биологической нише, где каждая популяция состоит из клональных линий, различающихся наличием определенных генов (метавирусных островков). Белки, кодируемые такими генами, могут, например, приспособлять специфичность фага к клональным линиям хозяина. Сосуществование различных клональных линий бактерий обогащает генетический фонд бактериальной популяции, а сопоставимое разнообразие фагов, инфицирующие эти бактерии, оказывает на него стабилизирующее действие. Применительно к стратегии эволюции *Tevenvirinae*, связанной с неканоническими основаниями, это приводит к выводу, что предковые фаги *Tevenvirinae* могли использовать разные неканонические основания и гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, для своего развития в клетках-хозяевах, которые могли отличаться как, например, генами систем рестрикции-модификации (контрадаптация), так и генами метаболизма (адаптация).

**Возможность распространения гипотезы о роли неканонических оснований, определенной у *Tevenvirinae*, на другие вирусы.** Было проведено исследование геномов бактериофагов Actinobacteria из базы данных NCBI Nucleotide Collection, которые имеют аминокислотные последовательности, гомологичные белкам Streptomyces phage Gilgamesh – бактериофага, принадлежащего к группе фагов-плазмид с высоким содержанием ГЦ-пар в геномной ДНК, относительно крупными геномами (более 120 т.п.о.), имеющие циклические перестановки геномов и гомологию больших субъединиц терминазы с белками терминаз фагов, осуществляющих общую трансдукцию с высокой частотой. Среди наиболее родственных фагу Gilgamesh по филогеномному анализу был найден фаг Saccharomonospora phage PIS 136. Особенностью PIS 136 является то, что он относится к фагам-одиночкам (singleton phage) – бактериофагам, которые не имеют близкородственных вирусов. Другой его особенностью является наличие среди белков цитозин-специфичной метилтрансферазы, гомологичной цитозин-специфичной метилтрансферазы группы XVI (группа, выделенная при анализе геномов *Tevenvirinae* из базы данных NCBI Nucleotide Collection). Как оказалось, ген цитозин-специфичной метилтрансферазы Saccharomonospora phage PIS 136 находится в регионе генома между ДНК-полимеразой и хеликазой, что, как было показано выше при анализе генов синтеза и модификации 5-гидроксиметилцитозина фагов *Tevenvirinae*, свойственно расположению генов синтеза и модификации 5-гидроксиметилцитозина. Учитывая предположения о возможности частого горизонтального переноса генов в данной области геномов у предков *Tevenvirinae*, исходно, гомолог гена цитозин-специфичной метилтрансферазы PIS 136 у предков группы XVI также мог располагаться в данной области. Такое расположение генов синтеза и модификаций пиримидинов могло иметь эволюционное преимущество для предков этих вирусов. Дальнейшие исследования особенностей генов, участвующих

в синтезе и модификации пиримидиновых оснований у вирусов, могут объяснить закономерность такого расположения этих генов. Кроме того, изучение организации генов синтеза и модификаций пиримидинов, может способствовать выяснению распространения стратегии эволюции *Tevenvirinae*, связанной с неканоническими основаниями, на другие вирусы.

При исследовании одного из белков, связанных с репарацией ДНК у T4-родственных бактериофагов – DenV, было предположено, что его действие связано с неканоническими основаниями ДНК. При рассмотрении его гомологов у других вирусов, в результате, отмечалось наличие гомологов не только между вирусами бактерий, но и *Tevenvirinae* и вирусами хлорелл рода *Chlorovirus*, причем не только между кластером DenV, но и другими. Что самое интересное, между ними образовался общий гомологичный кластер тимидилатсинтаз и НМС-трансферазы (является TS-подобным ферментом). Неизвестны данные о наличии у данных эукариотических вирусов неканонических оснований, но наличие гомологии среди нескольких ферментов свидетельствует о возможности ГПГ между ними и возможных схожих механизмов метаболизма оснований, влияющий на другие процессы активности вирусов. Учитывая это, а также данных о наличии у ряда вирусов бактерий схожее расположение генов, участвующих в синтезе неканонических оснований за счет тимидилатсинтаз-подобных ферментов [61], определенные в работе закономерности эволюции T4-родственных вирусов могут распространяться на огромное число других дцДНК вирусов.

**Связь неканонических оснований с классификацией T4-родственных бактериофагов.** Используемые геномы бактериофагов из базы данных NCBI Nucleotide Collection, являющихся на 2019 год представителями *Tevenvirinae*, для сравнительного геномного анализа, на данный момент представляют часть более широкой группы – семейство *Straboviridae*. Однако, выделенные рода сохранились и добавились новые. При сравнении этих родов, с полученными нами данными по набору генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, определено, что каждый род, представляющий отдельную экологическую нишу, имеет свой набор.

Таким образом, гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, также могут быть использованы в качестве генетических маркеров родов T4-родственных бактериофагов.

**Перспективы использования T4-родственных бактериофагов как агентов фаговой терапии.** Для T4-родственных бактериофагов известен широкий спектр исследований по использованию их в качестве агентов фаговой терапии (Chibeu et al., 2012; Merabishvili et al., 2012; Ronner, Cliver, 1990; Carter et al., 2012; Coffey et al., 2014; Cowley et al., 2015; Howard-Varona et al., 2018; Kropinski et al., 2012; Lee et al., 2016; Liao et al., 2019; Morita et al., 2002; Pham-Khanh et al., 2019; Ronner, Cliver, 1990), созданию препаратов на основе этих вирусов (Bourdin et al., 2014; Sarker et al., 2012; Loh et al., 2021; Moghtader et al., 2017; Batalha et al., 2021; Śliwka et al., 2019; Stanford et al., 2010) и испытаний этих препаратов на животных (Chibani-Chennoufi et al., 2004; Weiss et al., 2009; Dąbrowska et al., 2014), клинических испытаний препаратов (Bruttin, Brussow, 2005; Sarker et al., 2012) и клиническому использованию (Terwilliger et al., 2021).

При сравнительном анализе геномов фагов подсемейства *Tevenvirinae* из базы данных NCBI Nucleotide Collection было отмечено, что близкородственные фаги имеют одинаковый набор генов, способствующих воспроизведению фагов с

неканоническими основаниями ДНК. Более того, схожие наборы таких генов имеются у фагов, принадлежащих к одному роду. Данные литературы показывают, что в зависимости от того, какие гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, присутствуют в геноме фага, частота трансдукции будет изменяться до полного отсутствия трансдукции у фагов, которые имеют гены синтеза и модификации неканонических оснований, ген дУТФазы-дЦТФазы, *denA*, *denB*, *alc*. Не все T4-родственные бактериофаги могут быть использованы для фаговой терапии. Например, представители родов *Krischvirus* и *Pseudotevenvirus*, обладают относительно высокой частотой трансдукции (Tanyashin et al., 2003).

Для фаговой терапии наиболее перспективным родом подсемейства *Tevenvirinae* является род *Tequatrovirus*, так как представители этого рода встречаются часто при выделении фагов из экологических ниш, свойственным данным вирусам; имеют одно из самых больших для *Tevenvirinae* число генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, в результате чего у них не выявляется процесс трансдукции; способны инфицировать патогенные бактерии рода *Escherichia*, *Shigella*, *Yersinia*.

Также, гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, можно использовать в качестве генетических маркеров для выявления групп вирусов, которые способны или не способны к трансдукции, с целью отбора подходящих агентов для фаговой терапии.

### ЗАКЛЮЧЕНИЕ

Эволюция вирусов бактерий представляет огромный интерес с точки зрения фундаментальных и прикладных исследований. Учитывая особенности их формы существования, многие способы хранения, передачи, реализации, репликации генетической информации, которые были полностью или по большей мере утрачены у клеточных организмов в процессе эволюции, сохранились у них. Примером этого могут служить неканонические основания ДНК, чье наибольшее разнообразие было найдено у вирусов бактерий. Исследования особенностей этих оснований и влияние их на процессы существования фагов помогут в большей мере понять процессы, которые могли происходить на ранних этапах эволюции живого. С другой стороны, неканонические основания и основанные на них механизмы, снижающие частоту трансдукции до полного или почти полного ее отсутствия, могут представлять интерес с точки зрения конструирования фаговых препаратов. Возможность отбирать бактериофаги, которые исходно имеют механизмы, уменьшающие взаимодействие вирусов с ДНК бактерий, является важным пунктом для развития фаговой терапии. Для изучения различных аспектов неканонических оснований бактериофагов необходима модельная группа вирусов. Такой являются T4-родственные бактериофаги.

Наличие большого числа геномов в базах данных, присутствие в различных экологических нишах, широкая изученность ряда их представителей, наличие нескольких неканонических оснований, их модификаций и белков, способствующие воспроизведению фагов с неканоническими основаниями ДНК, позволяют в полной мере использовать их для изучения экологии и эволюции вирусов, как модельные организмы. Кроме того, на ряде T4-родственных бактериофагах исследовались способы получения и производства фаговых препаратов для терапии.

В ходе проведенного исследования были рассмотрены способы скрининга бактериофагов для отбора T4-родственных с неканоническими основаниями. В

результате при помощи спот-теста на штаммах с RM-системами, ПЦР, гель-электрофореза частиц удалось выделить T4-родственные бактериофаги, принадлежащие к различным родам и обладающие разными неканоническими основаниями. На основе их анализа найдено, что близкородственные бактериофаги имеют одинаковый набор генов, способствующие воспроизведению фагов с неканоническими основаниями ДНК, тогда как более удаленные по родству – разные.

Для исследования этой закономерности был проведен сравнительный геномный анализ представителей T4-родственных бактериофагов – *Tevenvirinae*. Было найдено, что фаги разных родов внутри рода имеют схожий набор генов, способствующие воспроизведению фагов с неканоническими основаниями ДНК, одинаковые неканонические основания и схожие модификации таких оснований, тогда как между родами они отличаются. Были определены особенности наследования таких генов, возможное влияние на эволюцию T4-родственных бактериофагов. Предложена стратегия эволюции предков T4-родственных бактериофагов, основанная на меж- и внутринишевых отличиях в способностях горизонтального обмена информации между T4-родственными бактериофагами и окружающей средой. Такая стратегия также подтверждается эколого-эволюционной моделью Постоянного Разнообразия, которая постулирует наличие разнообразных молекулярных модулей в популяции вирусов, необходимых для взаимодействия с бактериями-хозяевами. Такими модулями также могут являться гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, так как они позволяли по-разному взаимодействовать с бактериями и другими вирусами предкам T4-родственных фагов.

Также было определено возможное наличие схожих механизмов у других вирусов как бактерий, так и у эукариот. Учитывая современную классификацию, полученные данные также могут быть использованы для выбора генетических маркеров, позволяющих отбирать T4-родственных бактериофагов определенного рода. Более того, эти данные позволяют отбирать бактериофаги с необходимой частотой трансдукции.

Полученные результаты показывают, что неканонические основания сыграли одну из важнейших ролей в эволюции T4-родственных бактериофагов и их дивергенции. Кроме того, основываясь на них, возможно использование данной группы бактериофагов в качестве модели для исследования различных аспектов неканонических оснований.

## ВЫВОДЫ

1. Установлена корреляция между наличием генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, и родством у выбранных T4-родственных вирусов с неканоническими основаниями ДНК из охарактеризованной коллекции бактериофагов сточных вод и фекалий зубров.
2. Проведен сравнительный геномный анализ бактериофагов подсемейства *Tevenvirinae*. Выявлена зависимость эволюции *Tevenvirinae* от экологических ниш и генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК. Впервые обнаружено, что специфический набор генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, может определять принадлежность того или иного бактериофага к определенному роду подсемейства *Tevenvirinae*.

3. Впервые определено, что большинство генов синтеза и модификаций <sup>5</sup>hmдС у Т4-родственных бактериофагов располагаются в геномной области между двумя коргенами – генами ДНК-полимеразы и хеликазы, а также наличие в этой области у представителей Т4-родственных бактериофагов различных генов хоминг-эндонуклеаз. Полученные результаты в совокупности с литературными данными могут указывать на активный горизонтальный перенос генов синтеза неканонических оснований данной группы вирусов.
4. Предложена концепция эволюции Т4-родственных бактериофагов, связанная с распространением между различными экологическими нишами и затем экспансии ниши путем накопления генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, которые уменьшали возможность горизонтального переноса генов между фагами и пулом генов ниши. Определена возможность распространения данной концепции на вирусы с двуцепочечной ДНК других таксономических групп.

### **СПИСОК РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ**

#### **Статьи из изданий, входящих в список ВАК или приравненных им:**

1. Конструирование терапевтических фаговых коктейлей на основе бактериофагов Т4-типа: преимущества и недостатки / Н. А. Никулин, С. И. Кононенко, А. Г. Кощаев, А. А. Зимин // Политематический Сетевой Электронный Научный Журнал Кубанского Государственного Аграрного Университета. – 2017. – № 133. – С. 823-849
2. Zimin A. A. Homologs of the Bacteriophage T4 RNA Ligase 2 in Metagenomes of Ocean Microbiota / A. A. Zimin, N. A. Nikulin, N. N. Nazipova // *Mathematical Biology and Bioinformatics*. – 2020. – Vol. 15. – № S. – P. 88-108.
3. Nikulin N. A. Influence of Non-canonical DNA Bases on the Genomic Diversity of Tevenvirinae / N. A. Nikulin, A. A. Zimin // *Frontiers in Microbiology*. – 2021. – Vol. 12. – P. 632686.
4. Karmanova A. N. Structural organization, evolution, and distribution of viral pyrimidine dimer-DNA glycosylases / A. N. Karmanova, N. A. Nikulin, A. A. Zimin // *Biophysical Reviews*. – 2022. – Vol. 14. – № 4. – P. 923-932.
5. Phages for treatment of Escherichia coli infections / N. Nikulin, A. Nikulina, A. Zimin, R. Aminov // *Progress in Molecular Biology and Translational Science*. – 2023. – Vol. 200. – P. 171-206.
6. Comparative analysis of Actinobacteria phage-plasmids and their transduction potential / N. A. Nikulin, S. S. Kiselev, V. V. Panyukov [et al.] // *Mathematical Biology and Bioinformatics*. – 2023. – Vol. 18. – № 2. – P. 323-346.

#### **Статья из издания, входящего в базу данных РИНЦ:**

Никулин Н. А. Анализ Колифагов Зубров На Наличие Антирестрикционных Систем / Н. А. Никулин, М. Г. Шляпников, А. А. Зимин // Сборник Научных Трудов Краснодарского Научного Центра По Зоотехнии И Ветеринарии. – 2019. – Т. 8. – № 1. – С. 123-128

#### **Публикации в материалах научных конференций:**

1. Никулин Н. А. Бактериофаговая флора очистных сооружений города Пушкино Московской области / Н. А. Никулин, А. А. Зимин // в книге: Биохимия, физиология и биосферная роль микроорганизмов. Сборник тезисов III Пушкинской школы-конференции. Под редакцией Т.А. Решетиловой. 2016. С. 143-146.
2. Зимин А. А. Изучение бактериофаговой флоры европейских зубров *Bison bonasus* (L, 1758) и американских бизонов *Bison bison* (L, 1758) (Bovinae, Bovidae,

- Artiodactyla) Приокского-террасного государственного заповедника / А. А. Зимин, Н. А. Никулин, В. И. Землянко // в книге: Биохимия, физиология и биосферная роль микроорганизмов. Сборник тезисов III Пущинской школы-конференции. Под редакцией Т.А. Решетиловой. 2016. С. 141-143.
3. Никулин Н. А. Биоиндикация загрязненности воды при помощи бляшкообразующих колифагов, инфицирующих штаммы E.coli C600, DN1, B, VL21 / Н. А. Никулин, А. А. Зимин // В книге: Техногенные системы и экологический риск. Тезисы докладов I международной (XIV региональной) научной конференции. Под общей редакцией А.А. Удаловой. 2017. С. 147.
4. Отбор и идентификация бактериофагов подсемейства Tevenvirinae с неканоническими основаниями ДНК из природных источников / Н. А. Никулин, Н. В. Воложанцев, А. А. Кисличкина [и др.] // В книге: VI Пущинская школа-конференция «Биохимия, физиология и биосферная роль микроорганизмов». Сборник тезисов конференции. Под редакцией Т.А. Решетиловой. Москва, 2019. С. 137-139.
5. Никулин Н. А. Особенности эволюции вирусов бактерий подсемейства Tevenvirinae / Н. А. Никулин, А. А. Зимин // В книге: VII Пущинская конференция «Биохимия, физиология и биосферная роль микроорганизмов». Школа-конференция для молодых ученых, аспирантов и студентов «Генетические технологии в микробиологии и микробное разнообразие». Сборник тезисов конференции. Под редакцией Т.А. Решетиловой. Москва, 2021. С. 68-70.