

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР «ПУЩИНСКИЙ
НАУЧНЫЙ ЦЕНТР БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ РОССИЙСКОЙ
АКАДЕМИИ НАУК»

На правах рукописи

Никулин Никита Алексеевич

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

1.5.3. – молекулярная биология

Диссертация на соискание ученой степени

Кандидата биологических наук

Научный руководитель
Кандидат биологических наук
Зимин Андрей Антонович

Пушино - 2024

ОГЛАВЛЕНИЕ

СПИСОК СОКРАЩЕНИЙ.....	5
ВВЕДЕНИЕ.....	7
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ	14
1.1. T4-родственные бактериофаги	14
1.1.1. T-четные фаги.....	14
1.1.2. T4-родственные бактериофаги и их классификация	16
1.1.3. Пангеном T4-родственных бактериофагов.....	23
1.2. Неканонические азотистые основания в составе ДНК бактериофагов	28
1.2.1. Неканонические пурины	28
1.2.2. Неканонические пиримидины	33
1.2.3. Белки, способствующие воспроизведению фагов с неканоническими основаниями ДНК и функции неканонических оснований у T4-родственных бактериофагов.....	41
1.3. Эколого-эволюционные модели для фагов	44
1.3.1. Модель Красной Королевы (RQ)	45
1.3.2. Модель «выезжай на спине победителя» (PtW)	45
1.3.3. Модель «убей победителя» (KtW)	46
1.3.4. Модель Постоянного Разнообразия (C-D)	47
1.4. Заключение по обзору литературы.....	48
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ	50
2.1. Материалы.....	50
2.1.1. Реактивы и реагенты.....	50
2.1.2. Коммерческие наборы, ферменты, маркеры молекулярных масс, ДНК.....	50
2.1.3. Питательные среды	50
2.1.4. Буферы и растворы.....	51

2.1.5. Штаммы бактерий и бактериофагов	51
2.2. Методы	52
2.2.1. Выделение и очистка бактериофагов из фекалий и сточных вод	52
2.2.2. ТЭМ фаговых частиц фекалий.....	53
2.2.3. Полимеразная цепная реакция	53
2.2.4. Спот-тест на ограничение роста фагов на штаммах с RM-системами	57
2.2.5. Электрофорез фаговых частиц в агарозном геле.....	58
2.2.6. Получение фаговых частиц в высокой концентрации, их очистка, выделение ДНК.....	58
2.2.7. Секвенирование, сборка и аннотация геномов	59
2.2.8. Сравнительный анализ геномов	59
ГЛАВА 3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ И ИХ ОБСУЖДЕНИЕ.....	63
3.1. Т4-родственные бактериофаги с неканоническими основаниями из фекалий зубров и сточных вод	63
3.1.1. Бактериофаги из фекалий зубров и сточных вод очистных сооружений	63
3.1.2. Скрининг бактериофагов рабочей коллекции на принадлежность к Т4-родственным бактериофагам и наличие у них антирестрикционных систем	69
3.1.3. Анализ геномов выбранных при скрининге бактериофагов	78
3.2. Анализ геномов Т4-родственных бактериофагов из баз данных.....	84
3.2.1. Кластеризация гомологичных последовательностей из геномов	84
3.2.2. Филогенетические, пангеномные и филогеномные деревья	87
3.2.3. Анализ группоспецифичных кластеров.....	97
3.2.4. Неканонические основания и гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК.....	102
3.2.5. Возможность распространения гипотезы о роли неканонических оснований, определенной у Tevenvirinae, на другие вирусы.....	117

3.2.6. Связь неканонических оснований с классификацией Т4-родственных бактериофагов.....	120
3.3. Перспективы использования Т4-родственных бактериофагов как агентов фаговой терапии	121
ЗАКЛЮЧЕНИЕ	127
ВЫВОДЫ	130
СПИСОК ЛИТЕРАТУРЫ.....	131
СПИСОК РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ.....	147
ПРИЛОЖЕНИЕ 1	149
ПРИЛОЖЕНИЕ 2	153
ПРИЛОЖЕНИЕ 3	155
ПРИЛОЖЕНИЕ 4	162
ПРИЛОЖЕНИЕ 5	167

СПИСОК СОКРАЩЕНИЙ

АТФ – аденозинтрифосфат;

ВОЗ – Всемирная Организация Здравоохранения;

ГПГ – горизонтальный перенос генов;

ДНК – дезоксирибонуклеиновая кислота;

дцДНК – двуцепочечная ДНК;

мРНК- матричная РНК;

п.о. – пар оснований;

ПЦР – полимеразная цепная реакция;

РНК – рибонуклеиновая кислота;

C-D (англ. “Constant Diversity”) – модель “Постоянного Разнообразия”;

т.п.о. – тысяч пар оснований;

CRISPR (англ. “Clustered Regularly Interspaced Short Palindromic Repeats”) – короткие палиндромные повторы, регулярно расположенные группами;

DDBJ (англ. “DNA Data Bank of Japan”) – база данных геномных последовательностей;

dN(C, A, T, G, U)M(D, T)P (дНМ(Д, Т)Ф) – дезоксирибонуклеозид(цитидин, аденозин, тимидин, гуанидин, уридин)монофосфат(дифосфат, трифосфат);

E.coli – *Escherichia coli*;

EMBL (англ. “European Molecular Biology Laboratory”) – база данных биологических последовательностей;

Gly – глицин;

^{hm}5C (^{5hm}C, ^{5hm}dC) – 5-гидроксиметилцитозин;

Нос (англ. “head outer capsid protein”) – внешний белок капсида;

ICTV – (англ. “International Committee on Taxonomy of Viruses”) – международный комитет по таксономии вирусов;

Ig – иммуноглобулин;

IpI (англ. “internal protein I”) – внутренний белок капсида I

KtW (англ. “kill-the-winner”) – модель “Убей Победителя”;

ОРС (англ. “Open Reading Frame”) – открытые рамки считывания;

PDB (англ. “Protein Data Bank”) – база данных последовательностей белков;

PtW (англ. “piggyback-the-winner”) – модель “Выезжай на Спине Победителя”;

RefSeq (англ. “Reference Sequence Database”) – база данных геномных последовательностей;

RM-система – система рестрикции-модификации;

RQ (англ. “Red Queen”) – модель “Красная королева”;

SARS-CoV-2 (англ. “Severe acute respiratory syndrome-related coronavirus 2”) – штамм короновируса, вызвавший пандемию COVID-19 (англ. “Coronavirus disease 2019”).

ВВЕДЕНИЕ

Актуальность темы

Бактериофаги или вирусы бактерий – это биологические объекты, являющиеся репликаторами, которые участвуют в коэволюционных взаимодействиях «хозяин-паразит» [1].

Особенностью бактериофагов с двуцепочечной ДНК является наличие большого числа неканонических азотистых оснований, то есть оснований, отличных от аденина, тимина, гуанина, цитозина [2–4]. 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 [3–5]. Другой предполагаемой причиной является возможность сохранения бактериофагами остатков ранее существовавшего 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 T4-родственные бактериофаги, архетипом для которых является фаг T4 [6]. Ряд их представителей хорошо изучен (например, фаги T2, T4, T6), а также используется в качестве модельных организмов в молекулярной биологии [7]. Кроме того, у T4-родственных бактериофагов было определено наличие нескольких неканонических оснований и их модификаций, а также выявлены пути их биосинтеза [4, 6, 8]. Также у фага T4 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 [6]. Стоит отметить, что за счет наличия относительно большого числа геномов T4-родственных бактериофагов в базах данных, а также присутствия этих фагов в различных нишах, возможно исследование их эколого-эволюционных взаимосвязей при помощи

бионформатических 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 [9]. Некоторые Т4-родственные бактериофаги способны инфицировать бактерии-патогены. Также, у фага Т4 исследованы гены, участвующие в снижении частоты трансдукции [10], 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 T4*, 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 T4-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 Т4-родственных вирусов определено, что близкородственные Т4-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 Т4-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 Т4-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. Результаты являются достоверными и опираются на экспериментальные и литературные данные. 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 индексируются 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).

Благодарности

Автор благодарит своего научного руководителя к.б.н. Зимина А.А. за выбор темы исследования, за наставничество, терпение и труд на протяжении всего исследования и за неоценимый вклад в работу.

Автор благодарит к.б.н. Сузину Н.Е. за помощь в проведении электронной микроскопии.

Автор выражает благодарность к.б.н. Воложанцеву Н.В., к.б.н. Кисличкине А.А., к.б.н. Богуну А.Г за помощь в проведении секвенирования.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Т4-родственные бактериофаги

1.1.1. Т-четные фаги

Для Т4-родственных бактериофагов, как уже упоминалось ранее, архетипом послужил фаг Т4. Это связано с тем, что данный вирус являлся одной из основных используемых биологических моделей при становлении молекулярной генетики [11, 12]. Кроме Т4, родственные ему фаги Т2 и Т6 также использовались учеными во времена возникновения молекулярной генетики [7]. Вместе, Т2, Т4, Т6 составляли группу вирусов, названных Т-четными.

Т-четные фаги являются частью набора Т-фагов (Т1, Т2, Т3, Т4, Т5, Т6, Т7), которые были предложены в качестве экспериментальной системы в исследовании генетических процессов на молекулярном уровне [7]. Предложил эту систему Макс Дельбрюк – Нобелевский лауреат по медицине и физиологии, основатель знаменитой «Фаговой группы», в которую входили многие Нобелевские лауреаты: Сальвадор Лурия, Альфред Херши, Джеймс Уотсон, Френсис Крик, Ренато Дульбекко, Сидней Бреннер.

Первым среди Т-четных фагов в экспериментах Фаговой группы использовался Т2. Исходно в их экспериментах он имел название фаг «γ» [13, 14]. Фаговой группой данный фаг был получен от доктора Жака Бронфенбреннера (Jacques J. Bronfenbrenner) [15]. Бронфенбреннером Т2 назывался «РС» (и другие варианты сочетаний данных литер, например «Р.С.»). Имеется предположение, что исходно данный вирус был выделен в двадцатых годах XX века с большой вероятностью из фекалий [7]. На рисунке 1 представлена одна из первых электронных микрофотографий фага Т2.

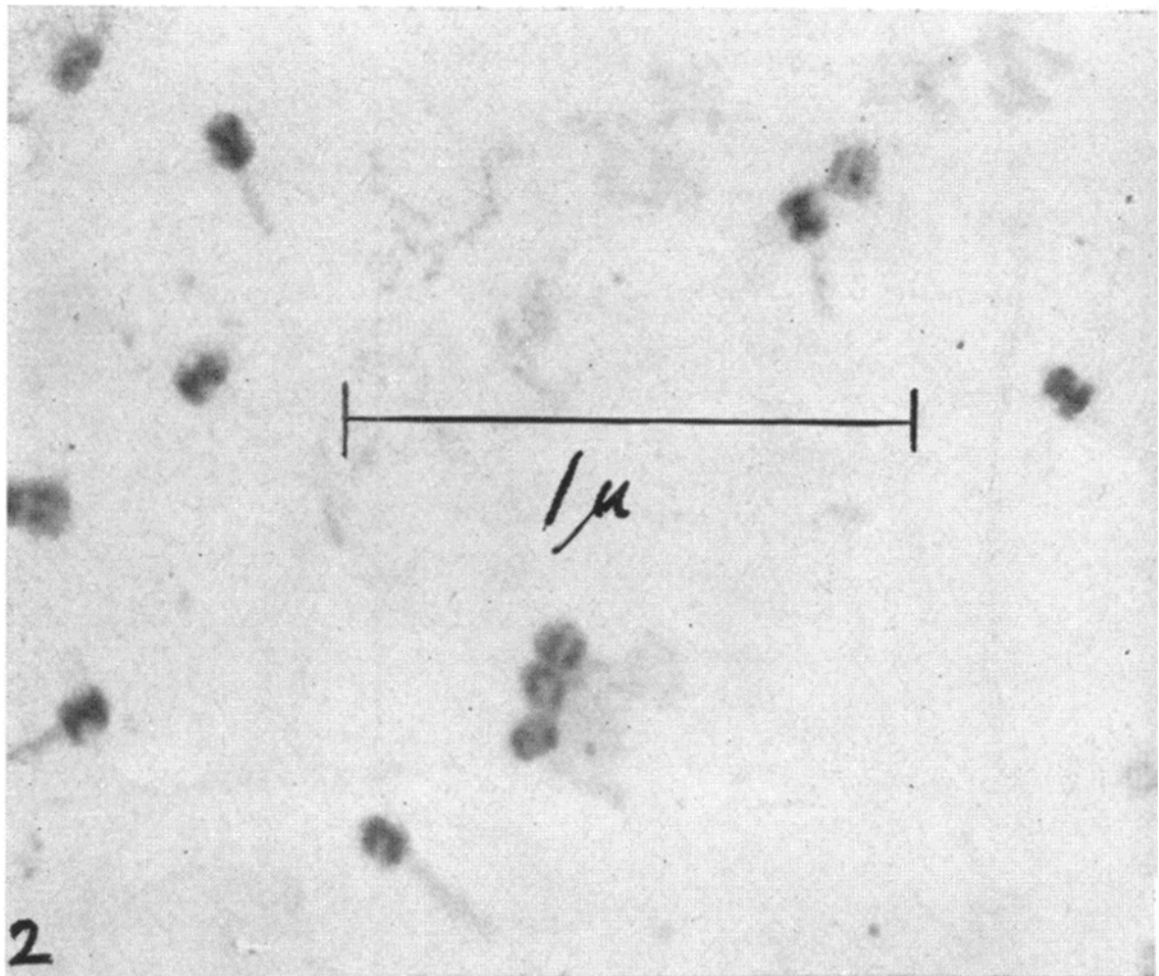


Рисунок 1 – электронно-микроскопическая фотография частиц фага «γ» (T2), опубликованная в 1943 году [15].

В дальнейшем было выделено 3 лабораторные линии фага T2: T2L, T2K, T2H [16]. T2H был получен Херши из единичной бляшки (негативной колонии) оригинального фага «PC» [16], ранее T2H имел название «P₉H» [17]. T2K был выделен Г. Калмансоном (G.M. Kalmanson) также из единичной бляшки «PC» [16], ранее сохранял название «PC» [18]. T2L был выделен из единичной бляшки фага T2K [16], имел название фага «γ» [13].

Фаги T4 и T6 не имели такой длительной истории переименования. Они были выделены из смеси фагов, предоставленных доктором Тони Ракитеном (Tony L. Rakieta) [19]. С большой вероятностью данная смесь была получена из сточных вод [7]. Среди исходных лабораторных штаммов T4 наиболее известными являются T4B (здесь «B» – Benzer) [20] и T4D («D» – Doermann) [21].

Стоит отметить, что основным признаком, по которому были выбраны фаги в Т-серию, являлась способность инфицировать штамм *Escherichia coli* В, название которому дали Лурия и Дельбрюк. Также, как и Т2 данный штамм был получен от Бронфенбреннера [22].

Наиболее часто используемым в экспериментах Т-четным фагом изначально являлся фаг Т2. Так, например, в известном «эксперименте Херши-Чейз» по определению носителя генетической информации использовался именно бактериофаг Т2 [23]. Позднее за счет создания коллекции летальных мутантных штаммов, чувствительных к температуре, и амбер мутантов Ричардом Эпштейном (Richard Epstein) и Робертом Эдгаром (Robert Edgar) [24] и распространении данных штаммов среди ученых, а также ряд других экспериментов, где использовался Т4 и его мутанты, данный вирус стал основным объектом из Т-четных фагов, используемый для изучения молекулярных механизмов бактериофагов. Впоследствии, за счет того, что Т4 фаг стал одним из самых изученных среди родственных вирусов, в честь него была названа группа данных организмов – Т4-родственные [6].

1.1.2. Т4-родственные бактериофаги и их классификация

Первыми фагами, которые можно определить как Т4-родственные, являлись Т-четные. Их родство было выявлено по морфологии вириона, серологических перекрестных реакций и ряда других признаков [25]. Позднее, их родство было подтверждено различными генетическими и геномными исследованиями этих вирусов [11]. Следующим вирусом, для которого было определено родство с Т-четными по морфологии и серологическим тестам, являлся фаг С16 [26], который использовался в качестве типового фага 11 серологической группы одной из первых классификаций бактериофагов, предложенной Фрэнком Макфарлейном Бёрнетом (Frank Macfarlane Burnet) [27]. В дальнейшем их родство также было подтверждено и по другим характеристикам [28].

В дальнейшем, при накоплении данных о фагах, выделенных из различных источников и изолированных на разных бактериях, были найдены и другие

бактериофаги, которые по своей морфологии, участкам ДНК, серологическим характеристикам были схожи с Т-четными вирусами [28–30]. Для ряда фагов, которые имели схожую морфологию с Т-четными бактериофагами, могли инфицировать как энтеробактерий, близкородственных *E.coli*, так и другие бактерии, был использован термин «бактериофаги Т4-типа» [28]. Позже, фаги Т4-типа были разделены на несколько групп. Сначала, по мимо понятия «Т-четные бактериофаги», появилось «Псевдо-Т-четные бактериофаги» (PseudoT-evens) [31]. Псевдо-Т-четные фаги сильнее других представителей Т-четных фагов отличались по морфологии, способности к гидролизу их ДНК различными эндонуклеазами рестрикции, гомологии геномной ДНК, в результате чего были выделены в отдельную группу. К этой группе относились, например, вирусы RB42, RB43, RB49 [31, 32]. Затем была выделена третья группа, названная «Шизо-Т-четные бактериофаги» (SchizoT-evens). Основным отличием вирусов этой группы можно считать то, что они инфицируют *Gamma*proteobacteria, отличных от энтеробактерий (например, *Vibrio*). Кроме того, они имеют больше отличий по гомологии геномной ДНК [33]. К Шизо-Т-четным бактериофагам относятся, например, вирусы Aeh1, KVP40 [33]. Последней выделенной группой по данной системе классификации фагов Т4-типа была группа «Экзо-Т-четных бактериофагов» (ExoT-evens). К данной группе были отнесены цианофаги, имеющие схожий с фагами Т4-типа ген главного белка капсида (один из основных генетических маркеров отдельных таксономических групп дцДНК вирусов) [34]. На рисунке 2 представлена филогенетическая диаграмма фагов Т4-типа.

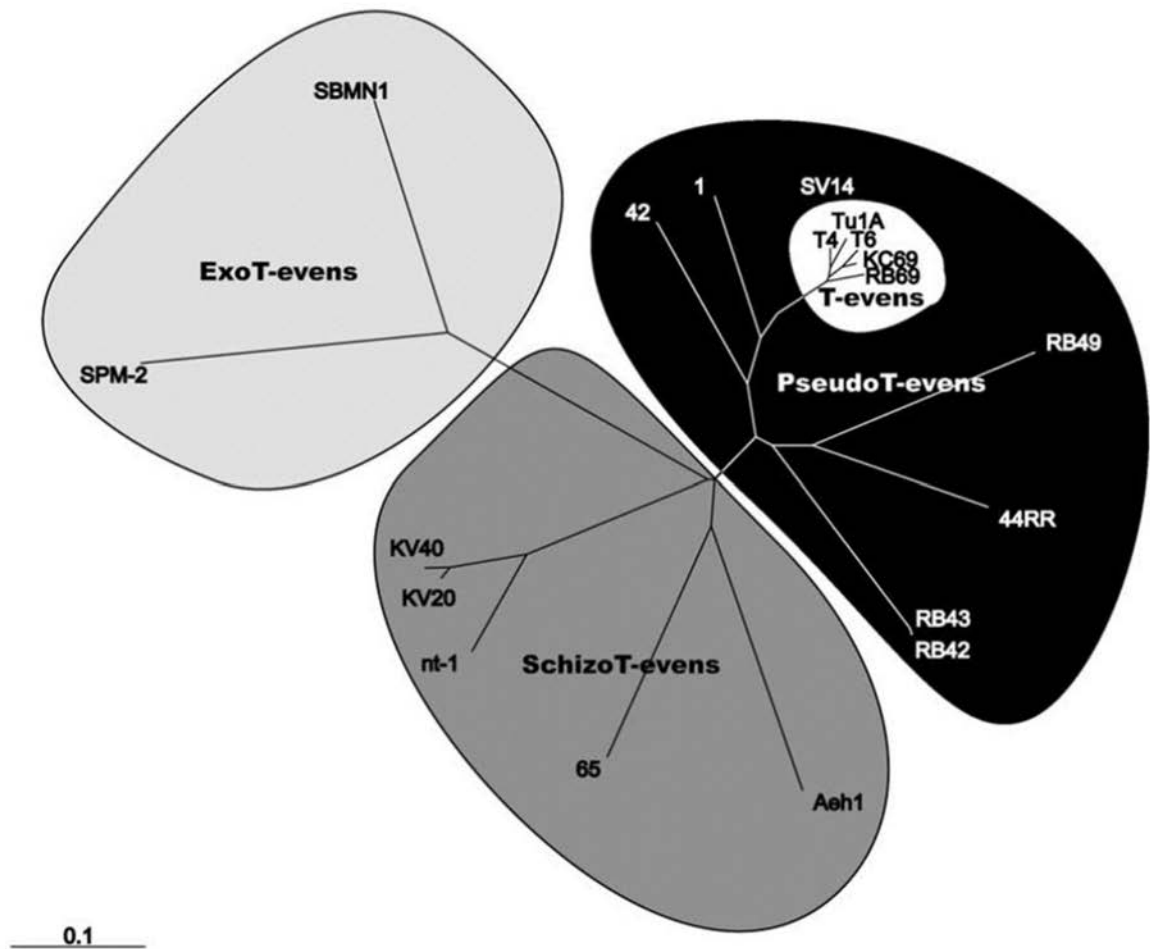


Рисунок 2 – филогенетическая диаграмма бактериофагов Т4-типа. Выделены группы Т-четных (T-evens), Псевдо-Т-четных (PseudoT-evens), Шизо-Т-четных (SchizoT-evens), Экзо-Т-четных (ExoT-evens) бактериофагов. Взято из [34].

Кроме понятия «фаги Т4-типа» в работах использовали понятия «Т4-подобные» (T4-like) и «Т4-родственные» (T4-related). Понятие «Т4-родственные бактериофаги» больше использовалось в работах, связанных со сравнением геномных последовательностей [6, 35, 36]. На основе геномных и метагеномных исследований, связанных с геном главного белка головки, была отдельно выделена группа Циано-Т4 (Cyano-T4) фагов, а также другие Т4-родственные фаги разделены на Приблизительно-Т4-родственные (Near-T4) (включают в себя Т-четные, псевдо-, шизо-Т-четные) и Отдаленно-Т4-родственные (Far-T4) (Экзо-Т-четные и более удаленные по родству вирусы) [37, 38]. На рисунке 3 представлено филогенетическое дерево, в котором была использована данная классификация Т4-родственных вирусов.

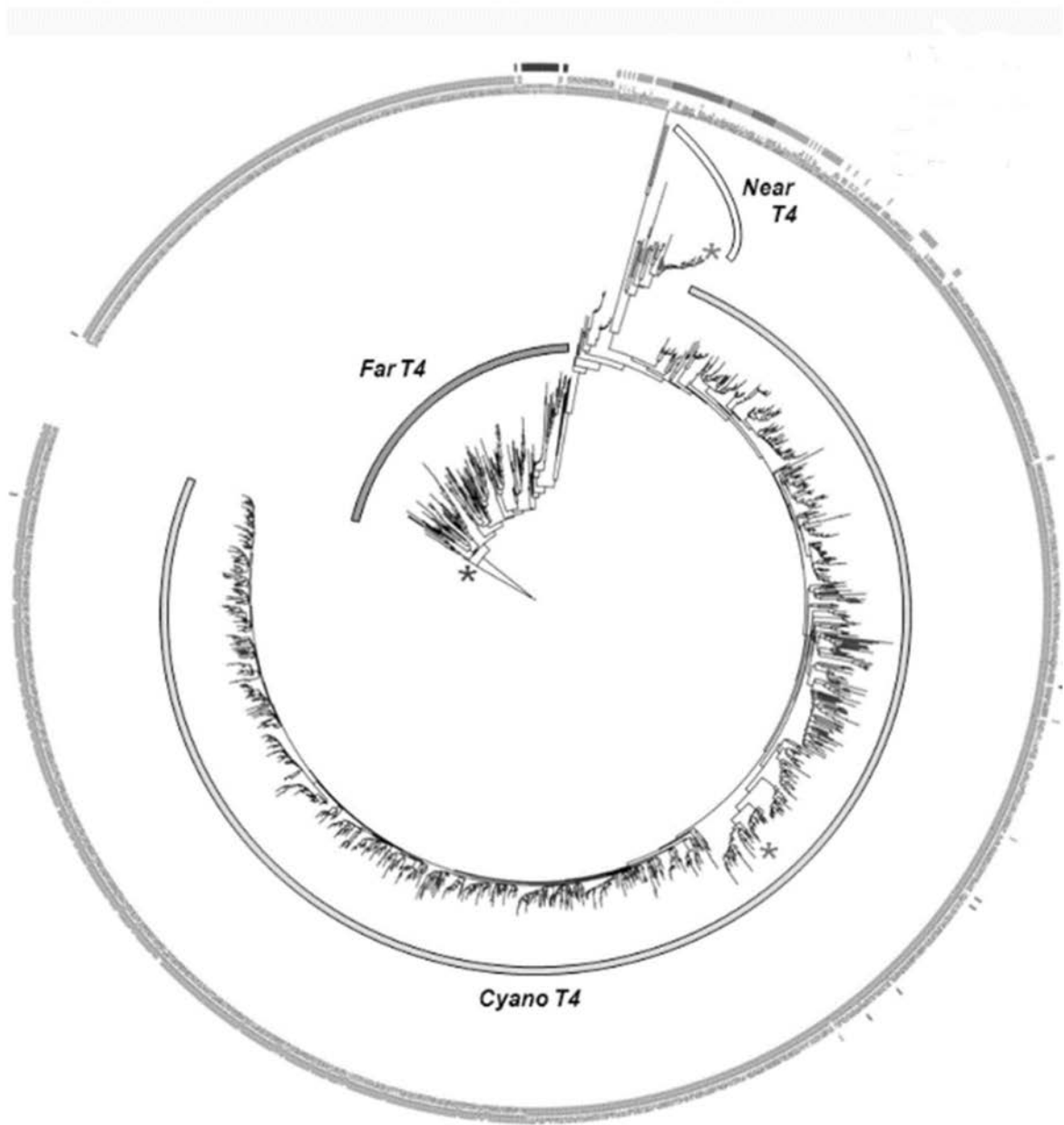


Рисунок 3 – филогенетическое дерево примерно 1400 последовательностей главного белка капсида T4-родственных бактериофагов. Отмечены группы Приблизительно-T4-родственных (Near T4), Отдаленно-T4-родственные (Far T4), Циано-T4 (Cyano T4) фаги. Взято из [37].

Международным Комитетом Таксономии Вирусов (ICTV) [39] таксономическая классификация T-четных и родственных им вирусов ведется достаточно давно – с первого отчета данного комитета, когда ICTV имел название Международного Комитета Вирусной Номенклатуры (INTV) [40]. В таблице 1 представлены изменения таксономии T-четных фагов (T2, T4, T6) с 1971 года по 2021 [41].

Таблица 1.Изменение таксономии Т-четных фагов (по данным из [41])

Год	Предложение об изменении	Таксономия
1971	Первый отчет INTV	Род: <i>T-even phages</i>
1976	Второй отчет ICTV	Род: <i>T-even phage group</i>
1981	Протокол пятого заседания ICTV	Семейство: <i>Myoviridae</i> Род: <i>T-even phage group</i>
1995	Пятый отчет ICTV	Семейство: <i>Myoviridae</i> Род: <i>T4-like phages</i>
1998	Предложение субкомитета Бактериальных вирусов на двадцать седьмом совещании исполнительного комитета ICTV	Отряд: <i>Caudovirales</i> Семейство: <i>Myoviridae</i> Род: <i>T4-like phages</i>
1999	Седьмой отчет ICTV	Отряд: <i>Caudovirales</i> Семейство: <i>Myoviridae</i> Род: <i>T4-like viruses</i>
2011	Предложение по таксономии исполнительному комитету ICTV за 2011 год	Отряд: <i>Caudovirales</i> Семейство: <i>Myoviridae</i> Подсемейство: <i>Tevenvirinae</i> Род: <i>T4-like viruses</i>
2012	Предложение по таксономии исполнительному комитету ICTV за 2012 год	Отряд: <i>Caudovirales</i> Семейство: <i>Myoviridae</i> Подсемейство: <i>Tevenvirinae</i> Род: <i>T4likevirus</i>
2015	Предложение по таксономии исполнительному комитету ICTV за 2012 год	Отряд: <i>Caudovirales</i> Семейство: <i>Myoviridae</i> Подсемейство: <i>Tevenvirinae</i> Род: <i>T4virus</i>
2018	Предложение по таксономии исполнительному комитету ICTV за 2018 год	Отряд: <i>Caudovirales</i> Семейство: <i>Myoviridae</i> Подсемейство: <i>Tevenvirinae</i> Род: <i>Tequatrovirus</i>
2019	Предложение по таксономии исполнительному комитету ICTV за 2019 год	Реалм: <i>Duplodnaviria</i> Царство: <i>Heunggongvirae</i> Тип: <i>Uroviricota</i> Класс: <i>Caudoviricetes</i>

Таблица 1 (продолжение). Изменение таксономии Т-четных фагов

Год	Предложение об изменении	Таксономия
		Отряд: <i>Caudovirales</i> Семейство: <i>Myoviridae</i> Подсемейство: <i>Tevenvirinae</i> Род: <i>Tequatrovirus</i>
2021	Предложение по таксономии исполнительному комитету ICTV за 2021 год	Реалм: <i>Duplodnaviria</i> Царство: <i>Heunggongvirae</i> Тип: <i>Uroviricota</i> Класс: <i>Caudoviricetes</i> Семейство: <i>Straboviridae</i> Подсемейство: <i>Tevenvirinae</i> Род: <i>Tequatrovirus</i>

Используемый подход ICTV для современной таксономии вирусов включает в себя сравнение консервативных мотивов/фолдов белков, коргенов, аминокислотных последовательностей и нуклеотидных последовательностей от таксонов наиболее высокого ранга до таксонов низкого ранга соответственно [42]. На рисунке 4 представлена таксономическая структура, используемая ICTV, и методологии для выделения таксонов.

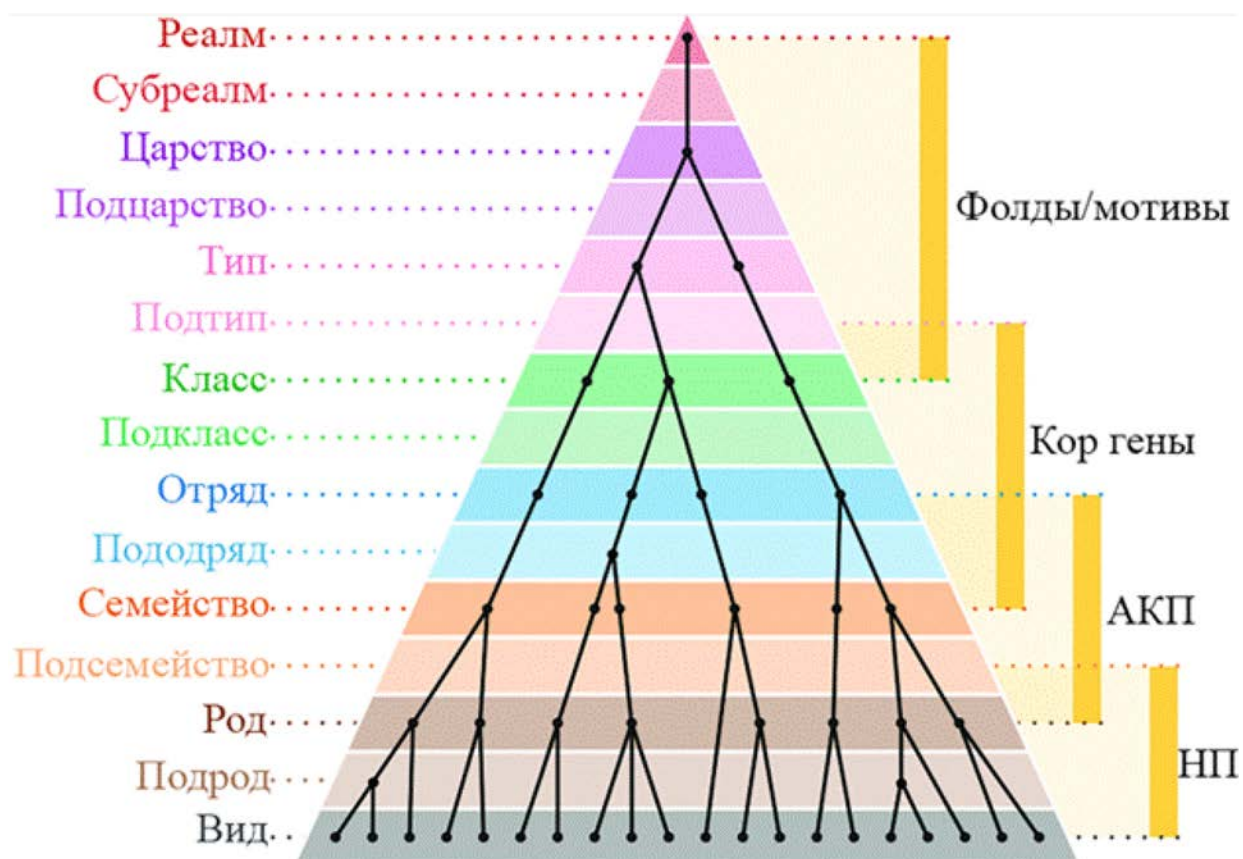


Рисунок 4 – Таксономическая структура, используемая ICTV. Слева представлены используемая ранговая система, справа – методологии, которые используются для определения эволюционных связей вирусов и присвоения им каждого ранга: фолды/мотивы – сравнение консервативных фолдов/мотивов белков; коргены – сравнение генов, свойственных всем представителям ранга; АКП – сравнение аминокислотных последовательностей; НП – сравнение нуклеотидных последовательностей. Адаптировано из статьи [42].

Классификация Т4-родственных бактериофагов, используемая в работах по исследованию эволюции данной группы вирусов, коррелирует с классификацией ICTV 2021 года больше, чем с предыдущими классификациями ICTV. Семейства *Straboviridae* и *Kyanoviridae* представляют собой Приближено-Т4-родственные и Циано-Т4 фаги соответственно. В качестве методов выделения этих фагов в отдельные семейства семейств использовались филогенетический анализ коргенов (определены в [43–45]), сравнение полных геномов [41]. До классификации ICTV 2021 года Приближено-Т4-родственные выделялись в подсемейство *Tevenvirinae*, которое состояло из Т-четных, Псевдо-Т-четных и части Шизо-Т-четных фагов,

остальные представители были разделены на отдельные рода. Циано-T4 бактериофаги не выделялись в единую таксономическую единицу. Таким образом, улучшение подходов таксономической классификации позволило точнее отразить картину взаимосвязей T4-родственных бактериофагов, что подтверждается сопоставимыми результатами исследований по эволюции вирусов этой группы.

Исследования полногеномных последовательностей, коргенома и в целом пангенома определенной группы вирусов кроме таксономии, также позволяет выявлять многие их эволюционные особенности. T4-родственные бактериофаги являются одними из самых изученных вирусов бактерий с точки зрения геномных и пангеномных исследований.

1.1.3. Пангеном T4-родственных бактериофагов

Пангеном – это глобальный репертуар гомологичных генов определённой группы организмов [46, 47]. В основном в качестве гомологичных последовательностей рассматриваются паралоги и ортологи [48], так как определение существующими инструментами более специфических разновидностей гомологов (например, ксенологов) является трудным или невозможным.

Пангеном можно разделить на различные компоненты. В целом можно выделить две системы, использующие схожий принцип деления. В первом случае [49] выделяют:

- Коргеном – пул генов, общий для всех представителей исследуемой группы геномов [50].
- Софткоргеном – пул генов, общий для 95% представителей исследуемой группы геномов [51].
- Оболочка пангенома – пул генов, умеренно сохраняемых у представителей исследуемой группы геномов [52].
- Облако пангенома – пул генов, встречающихся у отдельных представителей исследуемой группы геномов [52].

Данная система в основном используется в биоинформатических исследованиях геномов различных организмов.

Во втором случае [6, 53] выделяют:

- Коргеном – часть пангенома группы организмов, которая состоит из существенных генов, необходимых во всех известных условиях [6].
- Квазикоргеном – часть пангенома группы организмов, которая состоит из существенных генов, которые могут быть заменены или обойдены в определенных генетических фонах организма [6].
- Пластичная периферия пангенома – часть пангенома группы организмов, которая состоит из генов, образующие сегменты, значительно различающиеся по генетическому составу между различными линиями организмов [6, 53].

Данная система использовалась при исследовании геномов Т4-родственных бактериофагов. Кроме того, существуют и другие способы разделения пангенома, однако в основном они отличаются по названиям компонентов и их количеству [54].

Перед рассмотрением пангенома стоит обратить внимание на общие признаки и отличия геномов Т4-родственных бактериофагов. Гены Т4-родственных бактериофагов располагаются в циклически переставленном порядке внутри линейных хромосом двуцепочечной ДНК, что связано с механизмом упаковки ДНК в капсид (headfull package). У Т4-родственных имеется набор из 31-33 генов, которые связаны с возможностью фага осуществлять автономный контроль собственной репродукции. Из основных различий геномов этих вирусов можно выделить следующие: размеры геномов (варьируют от примерно 160 т.п.о. до примерно 250 т.п.о.); тип неканонического основания и его модификации (или их отсутствие) в ДНК геномов; дивергенция последовательностей генов (мутационный дрейф); полный набор генов геномов; топологическое расположение генов и их регуляторные сигналы в разных геномах [6].

Для Т4-родственных бактериофагов в качестве кор- и квазикоргенов выделяют гены, представленные в таблице 2.

Таблица 2. Кор- и квазикоргены пангенома Т4-родственных бактериофагов. Адаптировано из [6].

Гены	Продукты генов (и/или их активность)
Репликация, репарация, рекомбинация ДНК	
<i>43; 45; 44 и 62; 41 и 61; 59; 32; 46 и 47; uvsW; uvsX, uvsY; 30; rnh; 39, 60 и 52; dda; 49</i>	gp43 (ДНК-полимераза); gp45 (скользящий зажим); gp44/gp62 (комплекс загрузчика скользящего зажима); gp41/gp61 (хеликазо-праймазный комплекс); gp59 (загрузчик хеликазы/праймазы и регулятор gp43); gp32 (белок, связывающий одноцепочечную ДНК); gp46-gp47 (субъединицы рекомбинационного нуклеазного комплекса, необходимые для инициации репликации ДНК); белок UvsW (рекомбинационная ДНК-РНК-хеликаза); UvsX (RecA-подобный белок); UvsY (хелперный белок uvsX); gp30 (ДНК-лигаза); Rnh (рибонуклеаза H); gp39, gp60 и gp52 (субъединицы ДНК-топоизомеразы II типа); белок Dda (ДНК-хеликаза ближнего действия); gp49 (эндонуклеаза VII, необходимая для рекомбинации и упаковки ДНК).
Вспомогательный метаболизм	
<i>nrdA и nrdB; nrdC; nrdG; nrdH; 56; cd; frd; td; tk; 1; denA; dexA</i>	NrdA-NrdB (субъединицы аэробного комплекса рибонуклеотидредуктазы); NrdG и NrdH (субъединицы анаэробного комплекса рибонуклеотидредуктазы); NrdC (тиоредоксин); gp56 (дЦТФаза-дУТФаза); Cd (дезаминаза дЦМФ); Frd (дигидрофолатредуктаза); Td (тимидилатсинтетаза), Tk (тимидинкиназа); gp1 (дНМФ-киназа); DenA (эндонуклеаза II); DexA (экзонуклеаза A).
Экспрессия генов	
<i>33; 55; regA</i>	gp33 (белок, который опосредует взаимодействия gp55-gp45 -РНК-полимеразы в поздней транскрипции); gp55 (сигма-фактор поздней транскрипции); RegA (мРНК-связывающий репрессор трансляции)
Морфогенез фаговых частиц	
<i>2; 3; 4; 5; 6; 8; 13; 14; 15; 16 17; 18; 19; 20; 21; 22; 23; 25; 26; 34; 35; 36; 37; 49; 53</i>	gp2 (защищает концы упакованной ДНК от RecBCD); gp3 (терминатор оболочки); gp4 (белок завершения головки); gp5 (компонент хаба лизоцима базальной пластинки); gp6 (клиновой компонент базальной пластинки); gp8 (клиновой компонент базальной пластинки), gp13 (белок завершения головки); gp14 (белок завершения головки); gp15 (белок завершения хвоста); gp16 и gp17 (субъединицы терминазы); gp18 (субъединица оболочки хвоста); gp19 (субъединица трубки хвоста);

Полужирным курсивом отмечены коргены, курсивом отмечены квазикоргены. В качестве названий генов используются названия их гомологов в геноме фага Т4.

Таблица 2. Продолжение.

Гены	Продукты генов (и/или их активность)
Морфогенез фаговых частиц	
	gp20 (белок портала капсида); gp21 (коровый белок и протеаза проголовки); gp22 (коровый белок прокапсида); gp23 (предшественник основного белка капсида); gp25 (клиновой компонент базальной пластинки); gp26 (субъединица хаба базальной пластинки); gp34 (субъединица проксимального отростка хвоста); gp35 (шарнирный белок фибриллы хвоста); gp36 (субъединица малой дистальной фибриллы хвоста); gp37 (субъединица большой дистальной фибриллы хвоста); gp49 (эндонуклеаза VII; требуется для упаковки ДНК в капсид); gp53 (клиновой компонент базальной пластинки)
Другие	
<i>rIIA</i> и <i>rIIB</i>	Точные функции продуктов генов <i>rIIA</i> и <i>rIIB</i> неизвестны

Для Т4-родственных бактериофагов свойственно модульное расположение коргенов в геномах [6, 53]. Например, у представителей этой группы имеются модули репликации, морфогенеза вириона, фибрилл [53]. Модули коргенов перемежаются участками генов пластичной периферии пангенома [6, 53]. На рисунке 5 представлены примеры топологии коргенов в геномах Т4-родственных бактериофагов.

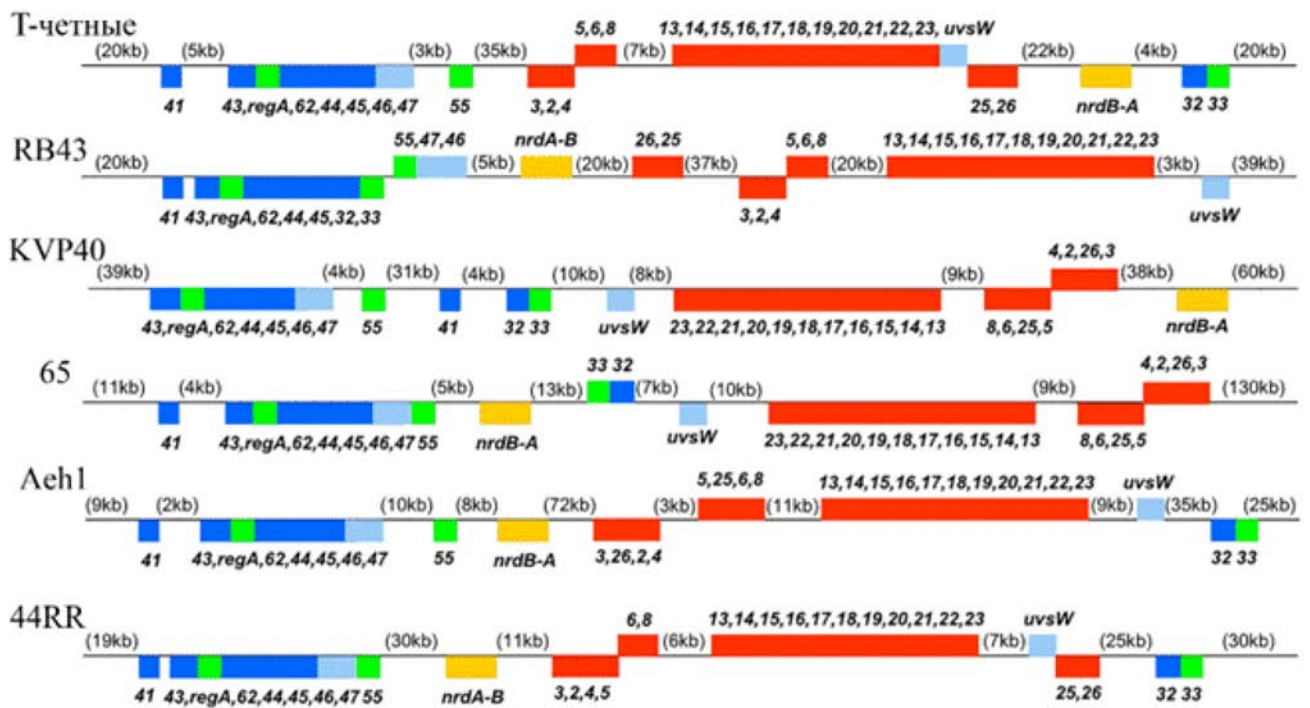


Рисунок 5 – Организация коргенов у представителей Т4-родственных бактериофагов; kb – тысяч пар оснований. Адаптировано из [6].

Для Т4-родственных бактериофагов характерен обмен коргенами между представителями различного родства. Кроме того, для ряда представителей найдена возможность горизонтального обмена квазикоргенами с бактериями-хозяевами [55, 56]. Считается, что высокая частота ГПГ у Т4-родственных бактериофагов связана с наличием большого числа хоминг-эндонуклеаз у различных представителей [6] и механизмов хоминга интрона [57] и безинтронного хоминга [58]. Кроме того это может быть связано с наличием мобильных регуляторных кассет в составе геномов [59, 60].

Хотя влияние ГПГ на эволюцию Т4-родственных вирусов считается высоким, однако это влияние ниже, чем у некоторых других фагов, например “лямбдоидных”. Отражается это в размере коргенов (числе коргенов) этих групп. Кроме того, отмечаются сильные барьеры, которые ограничивают обмен между близкими гомологами генов Т4-родственных фагов и сохранение общей синтении для ряда коргенов [53].

В качестве примера барьеров выделяют исключение из инфекции бактериофагом RB69 фага Т4 [32]. Другим примером барьера может являться возможность инфицирования фагами различных хозяев, имеющих разные системы

защиты от чужеродной ДНК. Оба этих примера связаны с наличием различных неканонических оснований в составе ДНК Т4-родственных вирусов.

1.2. Неканонические азотистые основания в составе ДНК бактериофагов

Среди всех известных биологических объектов бактериофаги имеют самое большое разнообразие неканонических оснований ДНК [3]. На данный момент у вирусов бактерий известно более 20 неканонических оснований и их модификаций и это число постоянно растет [3, 4, 8, 61]. Среди таких оснований присутствуют как пурины, так и пиримидины, являющиеся истинными неканоническими основаниями (неканонические основания встраиваются в ДНК после синтеза соответствующих нуклеотидов-предшественников *de novo*), модифицированными основаниями (канонические основания, подвергнутые изменению пострепликативно), модифицированными неканоническими основаниями или гипермодифицированными основаниями (неканонические основания, подвергнутые изменению пострепликативно)

1.2.1. Неканонические пурины

Н6-метиладенин

Один из вариантов модифицированных оснований встречается у одного из классических объектов молекулярной биологии – бактериофага λ [62], а именно Н6-метиладенин. Модификация происходит при помощи ДНК-метилтрансферазы, катализирующей перенос метильной группы с S-аденозил-L-метионина (рисунок 6) [3].

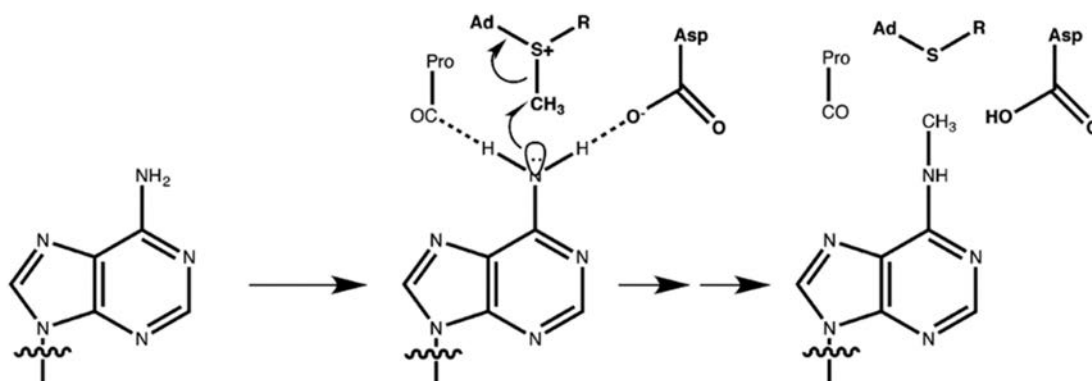


Рисунок 6 – Реакция модификации аденина N6-метилтрансферазой при участии S-аденозил-L-метионина. Адаптировано из [3].

2,6-диаминопурин

Одним из интересных неканонических оснований является 2,6-диаминопурин, который в составе ДНК заменяет аденин и, в отличие от него, образует не две, а три водородные связи с тиминам (рисунок 7).

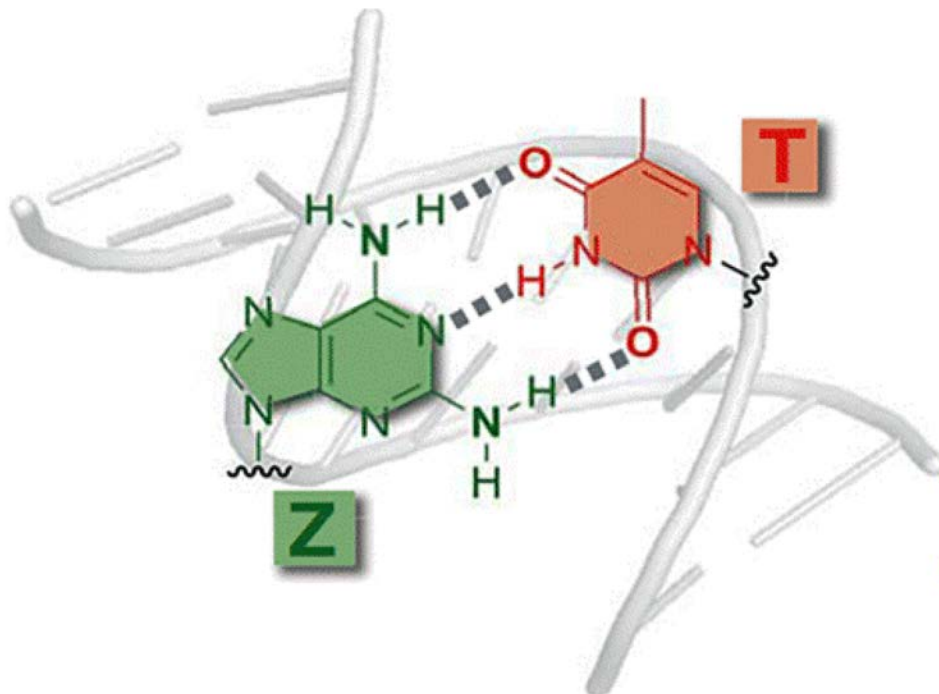


Рисунок 7 – схема взаимодействия 2,6-диаминопурина (Z) с тиминам (T). Адаптировано из [63].

Впервые данное основание было обнаружено у бактериофага S2-L, инфицирующего цианобактерию рода *Synechocystis* [64]. Позднее данное основание также было обнаружено у фага SH-Ab 15497 ацинетобактерий и других вирусов [63, 65]. Начало биосинтеза 2,6-диаминопурина (Z) катализируется кодируемым фагом ферментом PurZ, гомологом бактериальной аденилосукцинатсинтетазы (PurA). PurA катализирует образование сукциниладенозина монофосфата (SMP) из ксантозина монофосфата (XMP), глутамата и АТФ. Второй клеточный фермент, аденилосукцинатлиаза (PurB), отщепляет фумаратную часть от SMP с образованием АМФ. Точно так же PurZ может образовывать сукцинилдезоксигуанозинмонофосфат из dGMP, глутамата и АТФ. Отщепление фумаратной части от этого нуклеотида ферментом PurB, кодируемым хозяином, дает dZMP, который, в свою очередь, фосфорилируется нуклеотидкиназами, кодируемыми фагом и хозяином, с образованием dZTP для

использования кодируемой фагом ДНК-полимеразой [66]. Схема пути биосинтеза 2,6-диаминопурина представлена на рисунке 8.

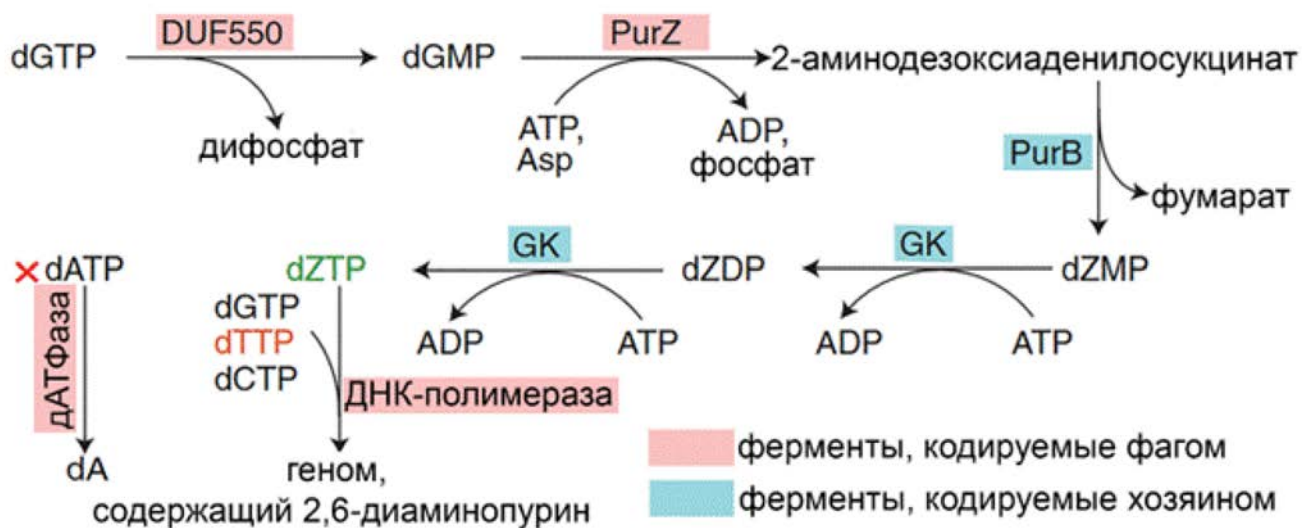


Рисунок 8 – путь биосинтеза dZTP. DUF550 – дАТФ/дГТФ пирогидролаза (содержит домен DUF550); GK – гуанилат киназа. Адаптировано из [63].

N6-карбомоил-метиладенин

Бактериофаг Ми имеет в составе своей ДНК модифицированное основание – N6-карбомоил-метиладенин (рисунок 9).

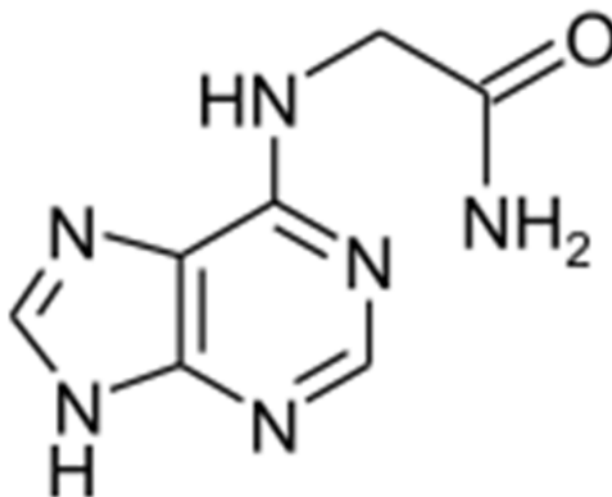


Рисунок 9 – структура N6-карбомоил-метиладенина. По [67].

В составе ДНК данного вируса он заменяет 15% аденина [68]. Фермент, который модифицирует аденин – Mom, является членом семейства GCN5-родственных N-ацетилтрансфераз (GNAT). Что интересно, в качестве косубстрата используется Gly-тРНК^{Gly} (показано в препринте [69]).

7-метилгуанин

У бактериофага DDVI, который инфицирует *Shigella sonnei* и *E.coli*, 1% всех гуанинов заменен на 7-метилгуанин [70] (рисунок 10).

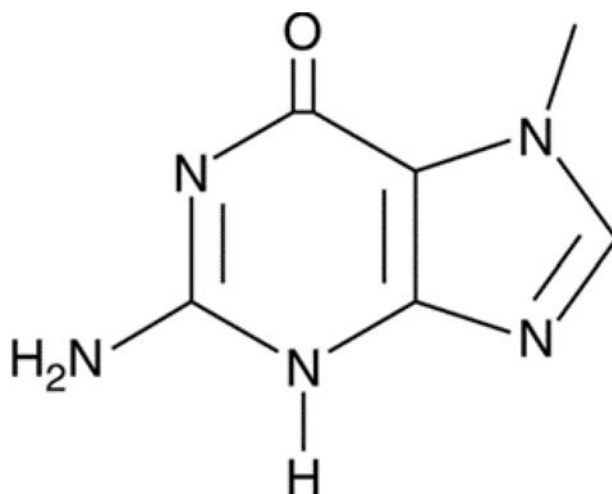


Рисунок 10 – структура 7-метилгуанина. По [71]

Для модификации гуанина в данном случае используется гуанинметилтрансфераза, донором метильной группы является S-аденозилметионин [70].

Производные 7-деазагуанина

Особым случаем среди неканонических оснований является 7-деазагуанин и его модификации, так как данное основание было исходно найдено в составе тРНК бактерий (в виде производного – кьюзина (Q)) [72] и архей (в виде производного – археозина (G^+)) [73]. Впервые у бактериофагов данное основание было найдено у бактериофага 9g [74] в виде производного – 2'-дезоксирхеозина (dG^+) [75] (рисунок 11).

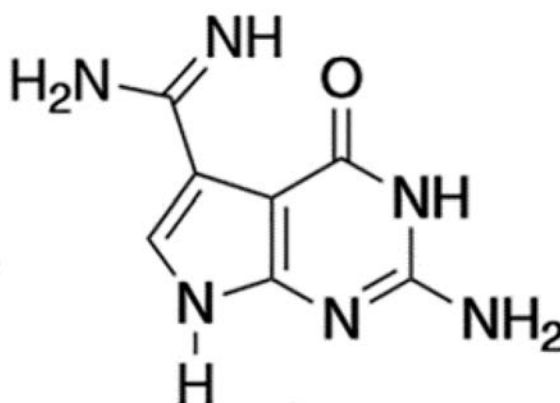


Рисунок 11 – структура dG^+ . Адоптировано из [66]

Для производных 7-деазагуанина используется целый набор ферментов, который кодируется генами, образующими геномные островки. Хотя такие наборы могут отличаться, однако для производных деазагуанина имеются общие этапы – синтез 7-циано-7-деазагуанина (PreQ₀) и его встраивание в ДНК. На рисунке 12 показан путь синтеза 7-циано-7-деазагуанина (dPreQ₀).

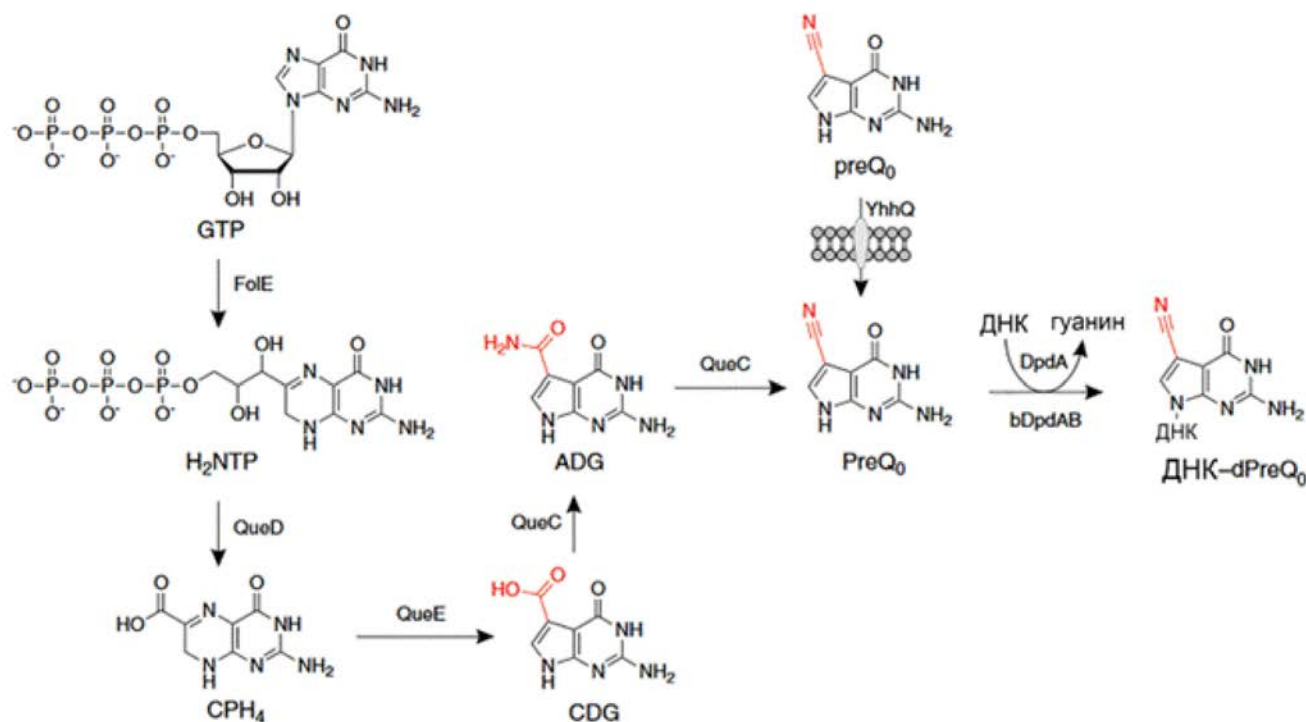


Рисунок 12 – путь синтеза dPreQ₀. Циклогидролаза ГТФ I (FolE) катализирует образование дигдронеоптеринтрифосфата (H₂NTP) из ГТФ, высвобождая муравьиную кислоту. Затем происходит гидролиз H₂NTP до 6-карбокси-5,5,7,8-тетрагидроптерина (CPH₄), катализируемый синтазой CPH₄ (QueD). После чего CPH₄ трансформируется в 7-карбокси-7-деазагуанин (CDG) с помощью CDG-синтазы (QueE). Затем карбоксильная группа CDG заменяется нитрильной группой с образованием PreQ₀. Эта реакция катализируется PreQ₀-синтазой (QueC), потребляющей две молекулы АТФ и проходящей через промежуточное соединение 7-амидо-7-деазагуанин (ADG). Некоторые бактерии и фаги также имеют специализированный транспортер PreQ₀ – YhhQ. В конце транскриптаза DpdA заменяет гуанин в ДНК на dPreQ₀. Адаптировано из [66].

Затем, синтез неканонического основания может закончиться или может произойти модификация основания. Для фага 9g показано наличие слитого гена, имеющего гомологи доменов глутаминамидотрансферазы класса II (GAT) и QueC, которые необходимы для синтеза dG^+ из $PreQ_0$ [4]. Для T4-родственного *Vibrio phage nt-1* также показано наличие гомологов генов, необходимых для синтеза dG^+ , однако у него имеется не гомолог GAT-QueC, а гомолог амидотрансферазы ArcS. Было обнаружено, что в связи с этим, кроме dG^+ в составе генома этого вируса имеются $dPreQ_0$ и ADG [4]. Для ряда других T4-родственных бактериофагов показано наличие различных наборов генов, необходимый для синтеза производных 7-деазагуанина [4]. Всего по анализу геномов бактериофагов на данный момент определено 4 производных 7-деазагуанина: $preQ_0$, ADG, G^+ и 7-аминометил-7-деазагуанин ($preQ_1$) [4, 76].

1.2.2. Неканонические пиримидины

Дезоксиурацил

У бактериофага PBS1 тимин в составе ДНК полностью заменен на дезоксиурацил.

Особенностью неканонических пиримидинов является то, что в большинстве случаев (не считая дезоксиурацила, 5-гидроксицитозина) в их синтезе у бактериофагов участвуют гомологи тимидилатсинтазы (TS). Синтез дУТФ происходит путем ингибирования хозяйских дУТФазы, тимидилатсинтазы, урацил-ДНК-гликозилазы, а также индуцирования дТТФазы, цитозиндезаминазы, дУМФ-киназы белками, кодируемыми фагом. Встраивание дУТФ в ДНК осуществляется за счет ДНК-полимеразы бактериофага [3, 77]. Путь синтеза дезоксиурацила у фага PBS1 представлен на рисунке 13.

5-гидроксицитозин

Nc1c(O)cnc(=O)n1

Рисунок 14 – структура 5-гидроксиметилцитозина. Взято из [79].

5-гидроксипентаурацил

Одним из сильно модифицированных неканонических оснований является 5-гидроксипентаурацил (рисунок 15).

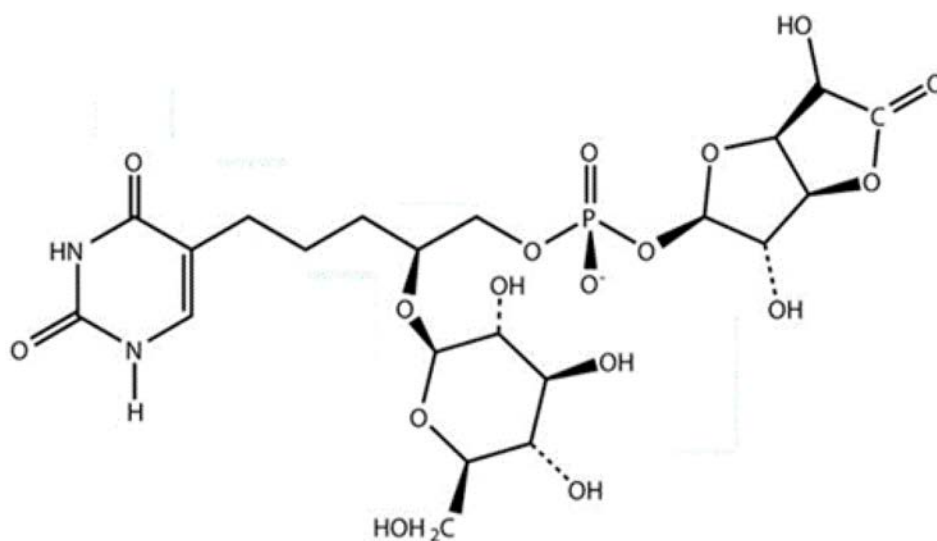


Рисунок 15 – структура 5-гидроксипентаурацила. Адаптировано из [80].

Данное основание было найдено у *Bacillus* phage SP-15 [2, 81]. Путь биосинтеза данного основания не известен, однако был найден кластер генов, который может участвовать в процессе модификации данного основания [80].

Неканонические пиримидины, синтезируемые тимидилатсинтаза-подобными ферментами

Синтез большинства неканонических пиримидинов фагов происходит при участии гомологов тимидилатсинтазы, которые присоединяют метильную или гидроксиметильную группы к пиримидиновому кольцу дНМФ. На рисунке 16 представлена общая схема синтеза неканонических пиримидинов.

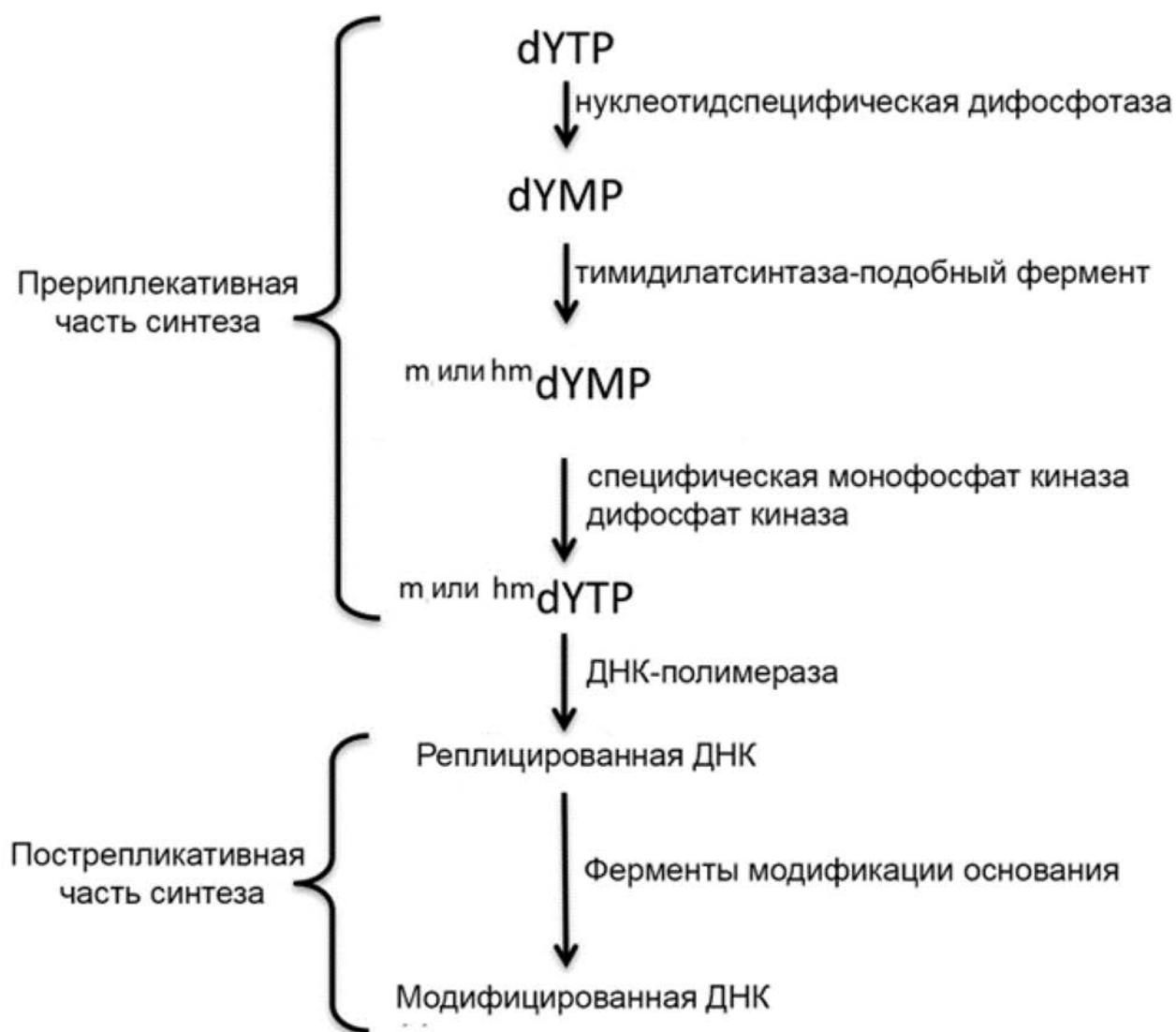


Рисунок 16 – общая схема биосинтеза неканонических пиримидинов: dYTP – пиримидин, «^m» – метилирование, «^{hm}» – гидроксиметилирование. Адаптировано из [3].

Кроме гомологов тимидилатсинтазы в геномах фагов кодируются и другие гены ферментов, необходимых для синтеза неканонических пиримидинов: специфические дифосфатазы оснований, которые заменяют неканонические; специфические монофосфат киназы, имеющие активность в отношении к неканоническому дНМФ [3]. Примерами неканонических пиримидинов, в синтезе которых участвуют тимидилатсинтаза-подобные ферменты являются 5-гидроксиметилдезоксисурацил (^{hm5}dU), 5-метилцитозин (^{m5}C), 5-гидроксиметилцитозин (^{hm5}C) [3] (рисунок 17).

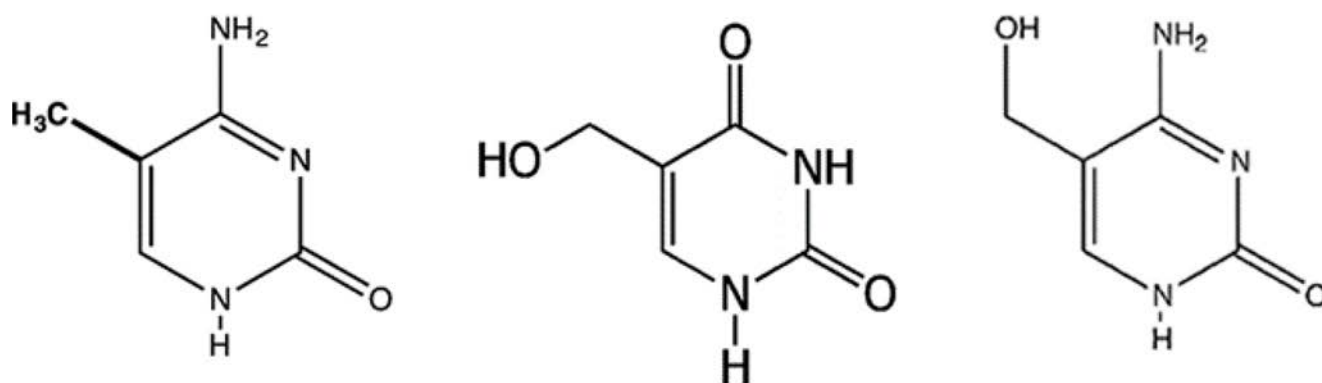


Рисунок 17 – неканонические пириимидины, в чьем синтезе участвуют гомологи тимидилатсинтазы. Слева – m^5C , по центру – hm^5dU , справа – hm^5C . Адаптировано из [3] и [82].

Особенностью hm^5dU и hm^5C является наличие ряда их модификаций в составе ДНК бактериофагов.

Модификации гидроксиметилдезоксимурацила

Для hm^5dU в геномах бактериофагов известно 5 модификаций (рисунок 18)

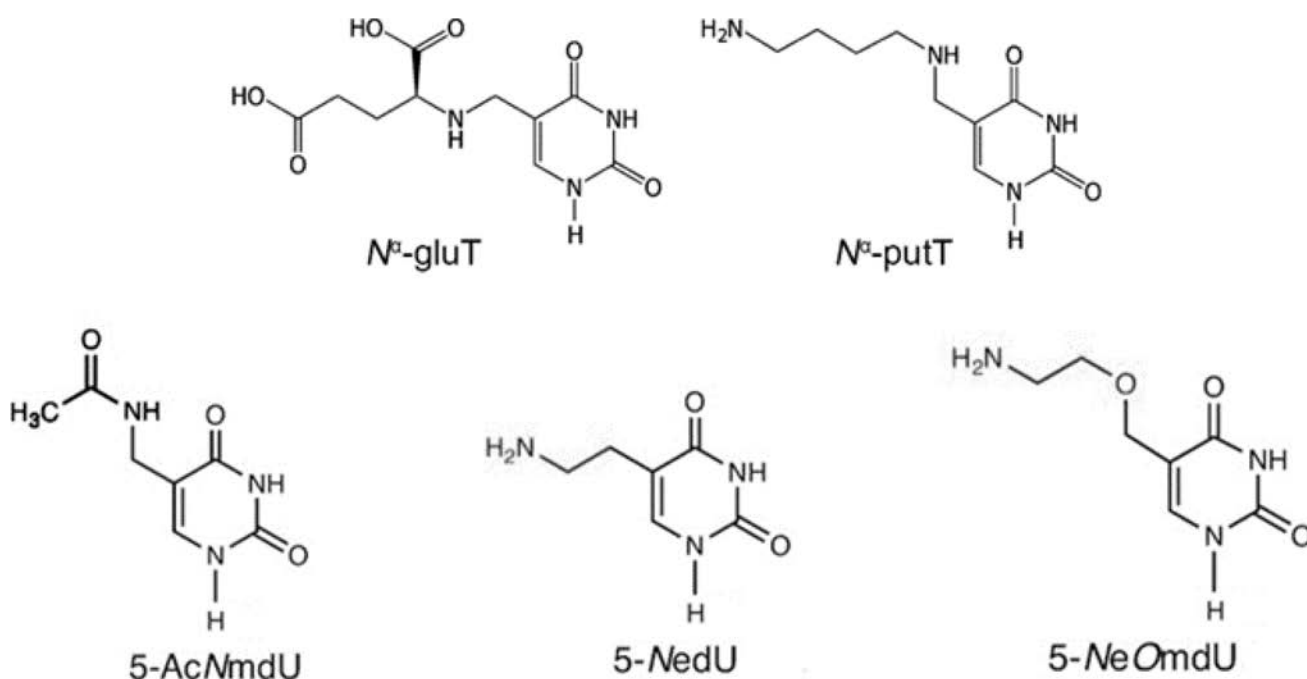


Рисунок 18 – модификации 5- гидроксиметилдезоксимурацила. N^{α} -gluT – α -глутамилтимин, N^{α} -putT – α -путресцинилтимин, 5-NedU – 5-аминоэтил-2'-дезоксимуридин, 5-NeOmdU – 5-аминоэтоксидезоксиметилуридин, 5-AcNmdU – 5-ацетиламинометил-2'-дезоксимурацил. Адаптировано из [3, 83].

Общей стадией синтеза для данных модификаций является образование 5-фосфометил-2'-дезоксимуридина (5-PmdU) при помощи ДНК-киназы 5-

гидроксиметилдезоксимурацила. Кроме того, общим признаком является образование модулей синтеза неканонических оснований и модификаций в геномах фагов [83]. В синтезе α -глутамилтимина у фага SP10 [84], у которого модификация была найдена, участвует только данный фермент, он дополнительно содержит домен глутамилтимидинпирофосфорилазы [61]. У фага ФW-14, у которого была найдена модификация α -путресцинилтимин [85], в отличие от SP10, содержит отдельный ген тимидинпирофосфорилазы, чьим продуктом является путресцинилтимидинпирофосфорилаза [83]. Другим общим признаком двух этих модификаций является то, что они оба начинаются от 5-метильной группы тимина от связи N-C.

Фагом М6, содержащим 5-NedU, кроме ДНК-киназы, также кодируются ДНК-трансфераза, пиридоксальфосфатзависимая (PLP-зависимая) декарбоксилаза, глицинилтиминизомераза. Таким же набором ферментов обладает фаг ViI, у которого в составе генома имеется 5-NeOmdU [61], однако ДНК трансферазы М6 и ViI используют разные субстраты (аминокислоты глицин и серин соответственно), что в последствии приводит и к другим отличиям в синтезе. Фаг РаМх11 имеет в составе ДНК 5-AcNmdU [83]. Начальные этапы синтеза данного основания схожи с таковыми у фага М6, одним из промежуточных продуктов является 5-N ^{α} -глицинилтимидин. Затем, под действием ферментов флавинзависимой лиазы и ацетилтрансферазы образуется 5-AcNmdU [83].

Модификации 5-гидроксиметилцитозина

У Т4-родственных бактериофагов было идентифицировано четыре модификации 5-гидроксиметилцитозина (рисунок 19).

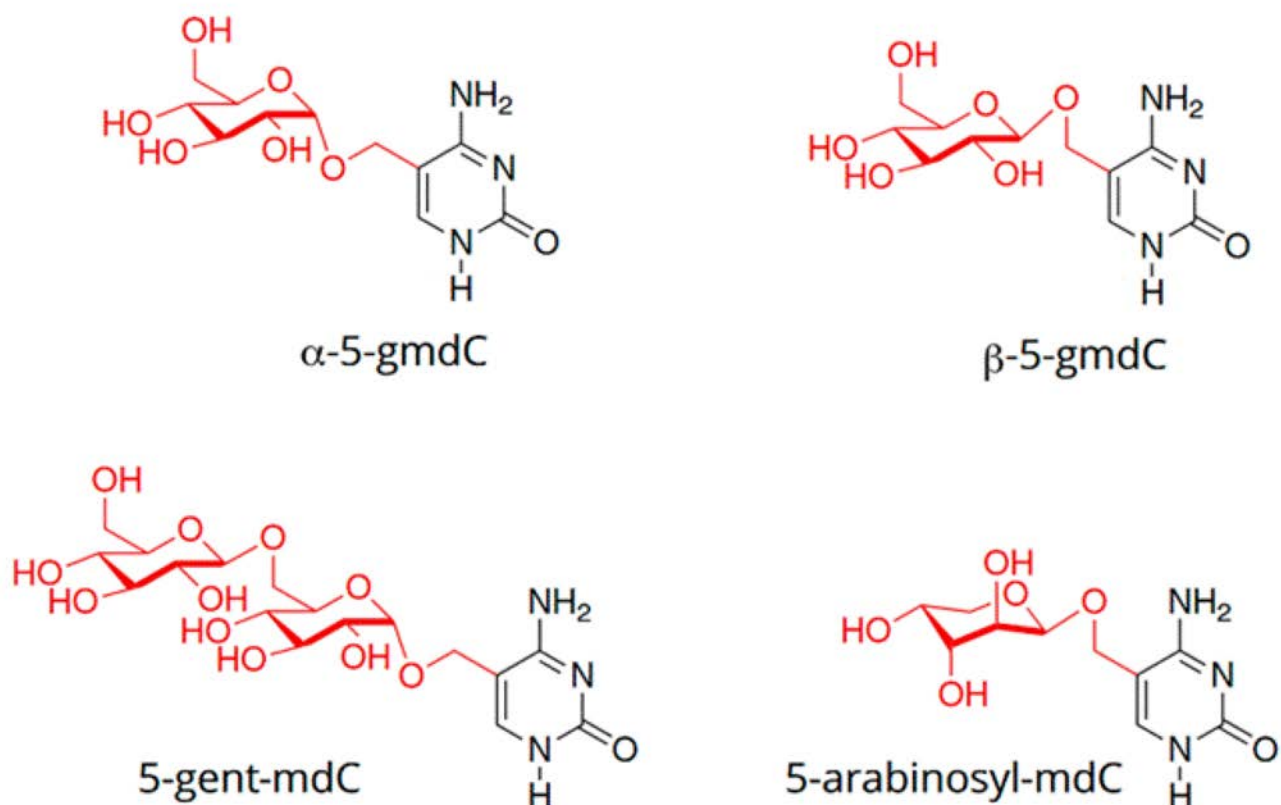


Рисунок 19 – модификации ^{hm}5C: α -5-gmdC – 5-(α -D-глюкозил)оксиметилцитозин, β -5-gmdC – 5-(β -D-глюкозил)оксиметилцитозин, 5-gent-mdC – 5-(α -генциобиозил)оксиметилцитозин, 5-arabinosyl-mdC – 5-(арабинозил)оксиметилцитозин. Адаптировано из [66].

Три из четырех этих оснований связано с добавлением к основанию глюкозы (α -5-gmdC, β -5-gmdC, 5-gent-mdC), одно – арабинозы. В случае арабинозы путь биосинтеза не известен, однако определены модули генов, участвующих в данном процессе и основной фермент – арабинозилтрансфераза. Для модификаций с участием глюкозы такой путь известен (рисунок 20).

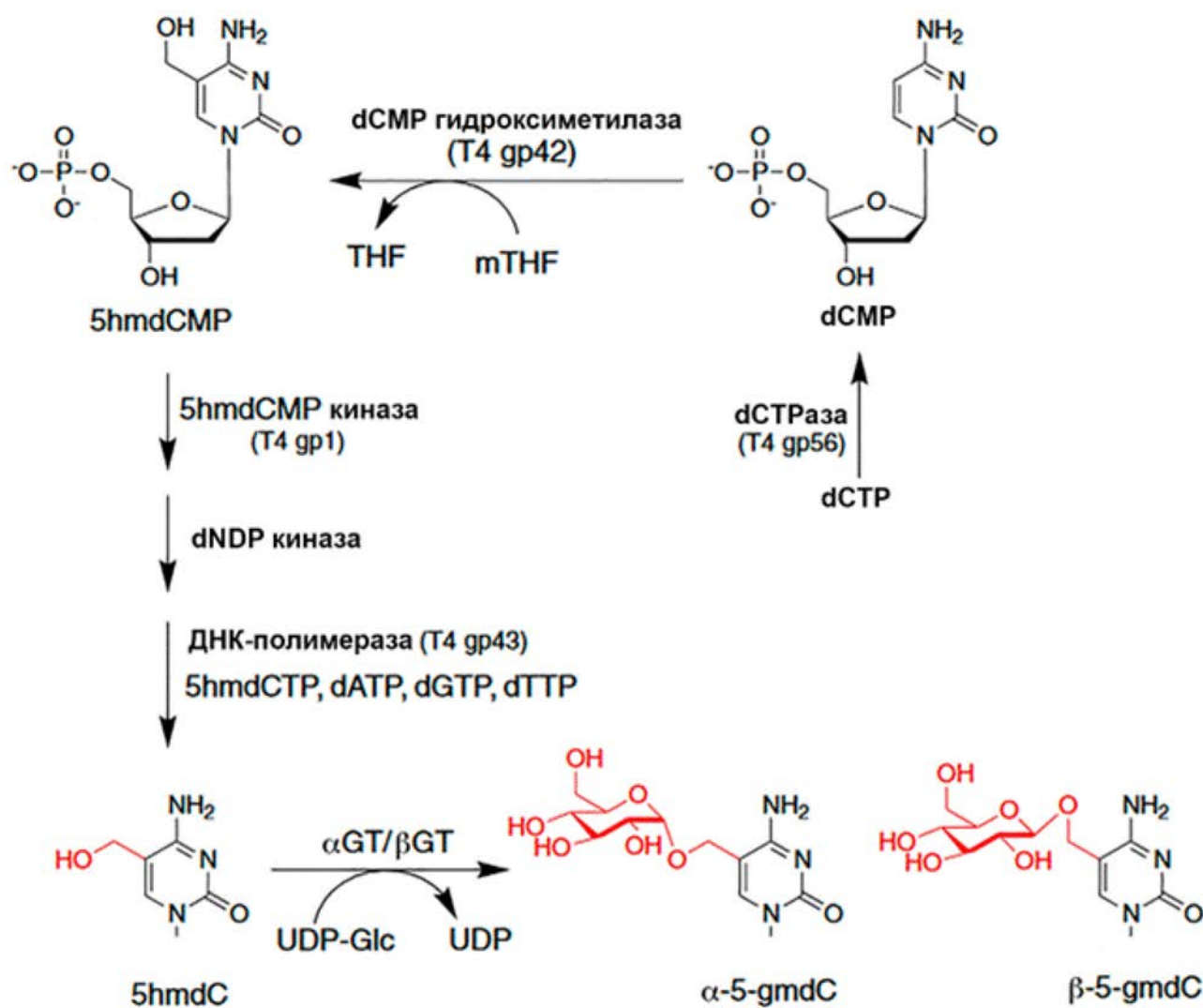


Рисунок 20 – путь синтеза гликозилированных ^{hm5}dC . THF – тетрагидрофолат, mTHF – 5,10-метилентетрагидрофолат, α GT/ β GT – α -гликозилтрансфераза/ β -гликозилтрансфераза, UDP-Glc – УДФ-глюкоза. Адаптировано из [66].

Синтез данных гликозилированных оснований начинается с изменения пула дНТФ, путем конвертирования дЦТФ в дЦМФ дЦТФазой, а затем гидроксиметилированием дЦМФ гомологом тимидилатсинтазы и mTHF [66]. После действия дНМФ/дНДФ киназ и встраивании в ДНК остатка $^{hm5}dCTP$ происходит гликозилирование ^{hm5}dC под действием α - или β -гликозилтрансфераз (в качестве донора они используют УДФ-глюкозу) [86]. В дальнейшем, может происходить дополнительное β -гликозилирование α -5-gmdC под действием β -1,6-гликозил- α -гликозотрансферазы [83].

1.2.3. Белки, способствующие воспроизведению фагов с неканоническими основаниями ДНК и функции неканонических оснований у T4-родственных бактериофагов

Широкое разнообразие неканонических оснований и их модификаций может быть связано с несколькими причинами. Во-первых, это «гонка вооружений» в процессе коэволюции бактериофагов и их бактерий-хозяев. В результате этой гонки бактерии приобретали новые системы защиты от чужеродной ДНК (CRISPR-Cas системы, системы рестрикции-модификации), бактериофаги же приобретали способы защиты или ухода от этих систем, одним из которых являются неканонические основания и их модификации [3]. Во-вторых, учитывая особую форму существования вирусов, они могут выступать в качестве носителей реликтов более ранних этапов эволюции живого. На это может указывать обнаружение 2,6-диаминопурина на углеродном метеорите [87]. Кроме того, неканонические основания могут выполнять различные функции в процессах репликации и реализации генетической информации вирусов.

У T4-родственных бактериофагов известно несколько белков, способствующих воспроизведению фагов с неканоническими основаниями ДНК: белки синтеза и модификаций неканонических оснований; белки, участвующие в защите от действия RM-систем IV типа; белки, гидролизующие ДНК, содержащую только канонический dC; белки, влияющие на транскрипцию ДНК бактерии-хозяина и бактериофага.

Белки, участвующие в защите от действия RM-систем IV типа

Особым типом RM-систем является IV тип, так как он выделяется на основе способности его представителей узнавать и гидролизовать сайты ДНК, содержащие неканонические основания [88]. Для ДНК, содержащей ^{hm5}dC известны несколько ферментов, способных ее гидролизовать: EcoKMcra, EcoKMcrc, SauUSI, семейство MspJI [89]. Данные ферменты кроме ^{hm5}dC способны гидролизовать ДНК, содержащие и другие неканонические основания (например, ^{m5}dC) [3]. Однако сайты с гликозилированными ^{hm5}dC (^{ghm5}dC) данные белки не узнают [89]. Таким образом ДНК ряда T4-родственных бактериофагов уже

защищены от этих эндонуклеаз. Существует и другие семейства рестриктаз, которые узнают и гидролизуют сайты, содержащие ghm^5dC . Речь идет о семействе PvuRts1I и GmrSD [89]. Семейство PvuRts1I способно узнавать как сайты с ghm^5dC , так и с hm^5dC (в меньшей степени с m^5dC) [90]. GmrSD же специфичен к сайтам с ghm^5dC [91].

Дополнительно к глюкозилированию у ряда T4-родственных бактериофагов имеется белок Arn, который ингибирует действие EcoKMcBVC путем мимикрии ДНК. Кроме того, данный белок ингибирует H-NS белок *E.coli* [92]. Также у ряда T4-родственных бактериофагов был найден IPI (internal protein I), который в процессе морфогенеза вириона попадает внутрь капсида и затем попадает внутрь другой клетки в процессе выхода ДНК из капсида в бактерию. Он способен ингибировать действие GmrSD по неизвестному механизму [93].

Белки, гидролизующие ДНК, не содержащую неканонических замен dC

В составе геномов представителей T4-родственных вирусов имеются гомологи генов, чьи продукты гидролизуют сайты ДНК, содержащие канонический dC – *denA* и *denB*. Ген *denA* кодирует эндонуклеазу II фага T4, которая представляет собой GIY-YIG эндонуклеазу. Она узнает сильно вариабельный ассиметричный сайт в дцДНК, в котором строго консервативными являются лишь пара цитозин, гуанин (CG), находящиеся в трех позициях от разрезаемой связи, и вносит ник на нижней (относительно сайта) цепи [94, 95]. Ген *denB* кодирует эндонуклеазу IV. В отличие от эндонуклеазы II она гидролизует оцДНК по механизму расщепления “цитозин-зависимого тракта” (dC-зависимого тракта). В случае DenB гидролиз происходит исключительно в 5'-проксимальном положении (dC1) внутри dC-зависимого тракта (представляет собой последовательность, где как минимум в первом и последнем положении находится цитозин), имеющего оптимальный размер 6 оснований [96]. Действие обеих эндонуклеаз ингибируется неканоническими заменами цитозина, в результате их основной функцией является участие в процессе гидролиза ДНК бактерии-хозяина (ее продолжают экзонуклеазы и другие ферменты фага) [97]. Последним ферментом данного процесса можно назвать gp56 фага T4 (или dCTPазу), которая использует

полученные dСТР для образования dСМР. За счет данного фермента происходит изменение пула dС и преобразование его в пул ^{hm5}dС.

Белки, влияющие на транскрипцию ДНК бактерии-хозяина и бактериофага

У Т4-родственных бактериофагов известно 2 белка, которые участвуют в модулировании транскрипции бактерии-хозяина или бактериофага. Белок Alc ингибирует транскрипцию ДНК, которая содержит цитозин. Данный белок является сайт-зависимым, хотя эти сайты являются сильно вырожденными и встречаются, в случае если хозяином является *E.coli*, с интервалами 80 – 200 пар нуклеотидов. Точный механизм действия Alc на данный момент не известен [98]. Другим белком, способствующим воспроизведению фагов с неканоническими основаниями ДНК, является белок MotA. Он является активатором транскрипции средних промоторов фага. Он связывается с ремоделированной σ^{70} *E.coli* другим фаговым белком – AsiA, а также со своим сайтом на ДНК (находится в средних промоторах фага), образуя комплекс, обеспечивающий узнавание средних промоторов фага, вместо промоторов бактерии-хозяина. Кроме того, ^{hm5}С и ^{ghm5}С усиливают аффинность данного комплекса [99].

Влияние на частоту трансдукции и горизонтальный перенос генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК

Как уже было отмечено ранее, для фага Т4 показано влияние белков, способствующих воспроизведению фагов с неканоническими основаниями ДНК, на частоту трансдукции. Исследовались следующие белки: эндонуклеаза IV, Alc, dСТРаза [10, 100]. В другой работе, также косвенно показано влияние эндонуклеазы II на трансдукцию ДНК плазмид [101]. Наличие большого числа белков, связанных с уменьшением частоты трансдукции, которые также влияют на возможности взаимодействия с ДНК клетки, может свидетельствовать о приспособлении фага Т4 к преобладанию вертикального пути наследования генетической информации. Выявление генетических и геномных особенностей у Т4-родственных вирусов, связанных с наличием/отсутствием гомологов этих

белков у них, откроет возможность определения влияния горизонтальной и вертикальной передачи генов на эволюцию этой группы вирусов.

Неканонические основания в составе ДНК имеют широкое влияние на Т4-родственные бактериофаги, учитывая наличие целой «генетической сети» [6], гены которой напрямую или опосредованно взаимодействуют с фаговой ДНК с неканоническими основаниями. Такое влияние может отмечаться в экологической эволюции данных вирусов. Для фагов, существует несколько эколого-эволюционных моделей, в которых может отразиться влияние неканонических оснований.

1.3. Эколого-эволюционные модели для фагов

Для экологических взаимоотношений в процессе коэволюции бактерий и фагов выделяют три основные модели: Красная Королева (Red Queen), убей победителя (kill-the-winners), выезжай на спине (используй) победителя (piggyback-the-winner) (рисунок 21).

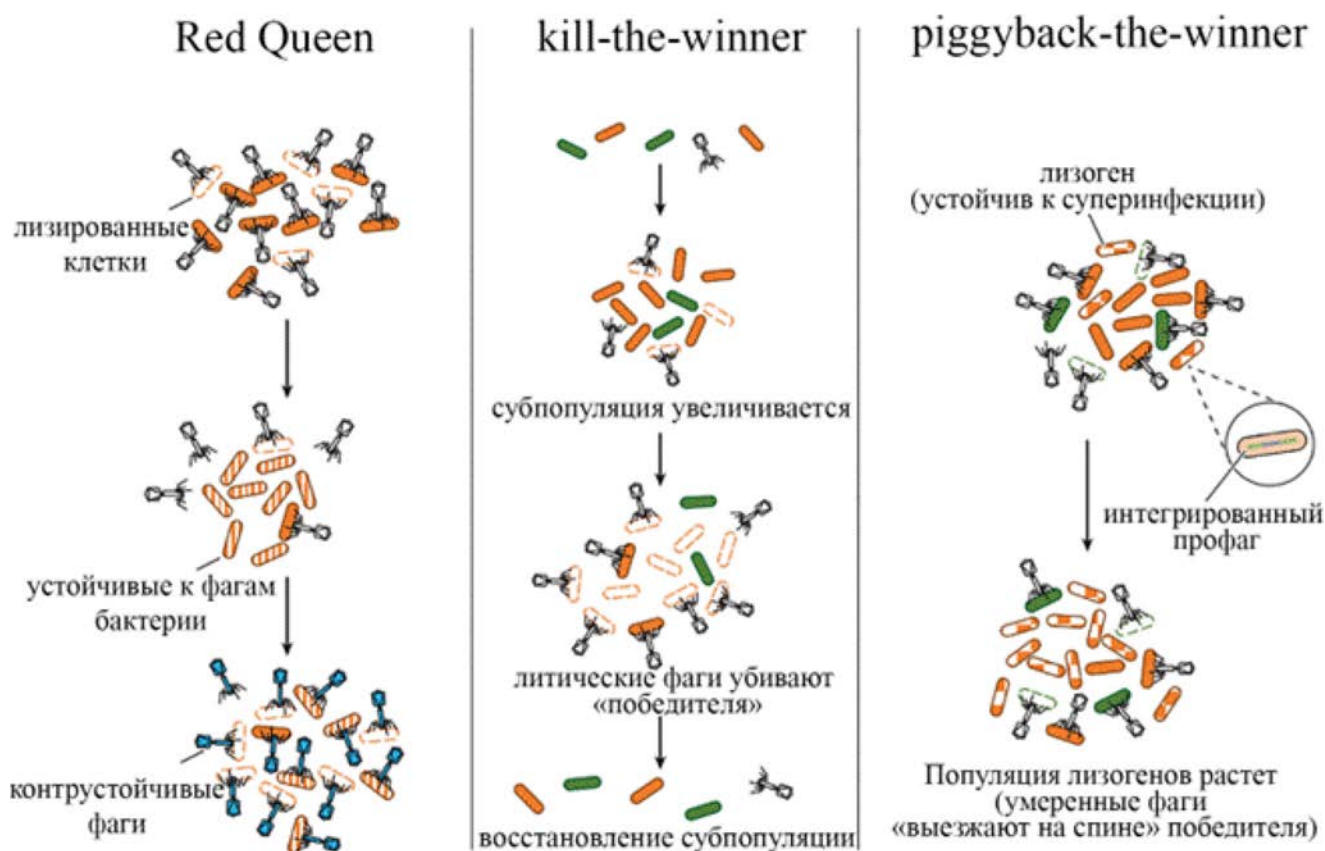


Рисунок 21 – взаимодействия фаг-бактерия в эволюционных моделях Red Queen (RQ), kill-the-winner (KtW), piggyback-the-winner (PtW). Адаптировано из [102].

1.3.1. Модель Красной Королевы (RQ)

Гипотеза (модель) красной королевы (или гонки вооружений) утверждает, что конкурентные взаимодействия с окружающей средой, такие как взаимодействия между хозяевами и паразитами, приведут к постоянной изменчивости и отбору в сторону адаптации хозяина и контрадаптации со стороны паразита. Первоначально она предполагала, что в тесных коэволюционных взаимодействиях, таких как отношения жертвы и хищника, изменения одной стороны могут привести к почти исчезновению другой стороны. Единственный способ, которым вторая сторона может поддерживать свое выживание— это контрадаптация. Это приведет к неустойчивому балансу между добычей и хищником, когда виды должны будут постоянно развиваться, чтобы оставаться на одном и том же уровне приспособленности. Было обнаружено, что эта модель «эволюционной гонки вооружений» применима ко многим биологическим процессам, но больше всего она подходит к отношениям паразит-хозяин. Для взаимоотношений бактериофагов и бактерий данная эволюционная тенденция, как считается, проявляется наиболее ярко, что связано с быстрой эволюцией вирусов и «оборотом» фаговых частиц [103].

На данный момент эта модель, например, используется для рассмотрения механизмов защиты, участвующих в «гонке вооружений», таких как системы рестрикции-модификации, системы CRISPR/Cas [103].

1.3.2. Модель «выезжай на спине победителя» (PtW)

Модель «выезжай на спине (используй) победителя» описывает экологические параметры, которые можно применять к умеренным фагам в микробных сообществах с высокой плотностью. Многие экологические ниши, в которых обитают бактериальные сообщества с высокой плотностью (кишечник, коралловые рифы, почва), демонстрируют более низкое, чем ожидалось, соотношение фаг/бактерия [104]. Сниженное соотношение фагов к бактериям может быть объяснено лизогенией умеренных фагов. Например, большинство бактерий в кишечнике лизогенизируется по крайней мере одним умеренным фагом

[105, 106]. Умеренные фаги защищают своего хозяина от повторного заражения фагом посредством исключения суперинфекции [107]. Лизогенная конверсия (изменение фенотипа бактерии-хозяина, связанная с внедрением фага) также может дать бактериям-хозяевам множество преимуществ. Модель PtW постулирует, что высокая плотность микробов способствует лизогении, а не лизису [102]. В результате, лизогены имеют большую вероятность выжить, относительно других бактерий, которые могут инфицировать вирулентные фаги. Можно отметить, что PtW отражает в большей мере взаимоотношения бактерий и умеренных фагов, для вирулентных фагов (к которым относятся T4-родственные бактериофаги) существует другая модель.

1.3.3. Модель «убей победителя» (KtW)

Модель «убей победителя» в контексте литических вирулентных фагов зависит от плотности популяции бактерий. Когда плотность бактериальной популяции хозяина низкая, взаимодействие фага с хозяином происходит реже, что приводит к уменьшению частоты инфекции. Если популяция бактерий-хозяев увеличивается, шансы того, что фаг столкнется с восприимчивым хозяином, увеличиваются. Более высокие темпы заражения замедляют рост бактериальной популяции. Таким образом, фаги могут стабилизировать состав бактериального сообщества за счет инфекции по модели «убей победителя», которая предотвращает захват сообщества быстрорастущими видами бактерий. Однако такая инфекция применительно к вирулентным фагам, заражающим различные бактериальные популяции, может привести к резкому коллапсу популяции. В бактериальных популяциях, состоящих из нескольких видов, динамика роста популяции, в которой происходит такой резкий коллапс, циклична для каждого вида [108].

Существует большое число интерпретаций и модификаций модели KtW. Одной из таких является модель Постоянного Разнообразия (Constant-Diversity).

1.3.4. Модель Постоянного Разнообразия (C-D)

Модель Постоянного Разнообразия утверждает, что несколько клональных линий сосуществуют, обогащая генетическое богатство популяции, и они удерживаются под контролем благодаря стабилизирующему влиянию аналогичного разнообразия вирусов, которые охотятся на популяцию [109]. Отличие от KtW состоит в том, что модель C-D рассматривает взаимоотношения фагов и бактерий на уровне популяции и предполагает постоянное разнообразие клональных линий популяции бактерии за счет стабилизирующего отбора фагами. В результате чего на уровне популяции не обнаруживается доминирующей линии. Это позволяет популяции более эффективно овладевать нишей за счет того, что отдельные линии приобретают разные, но дополняющие друг друга возможности использования ниши. Ожидаемое более эффективное использование ресурсов сообществом может приводить к лучшему функционированию экосистемы, за счет отсутствия критических коллапсов популяций, участвующих в биогеохимических процессах [110]. На рисунке 22 представлена C-D модель на примере нескольких симпатрических клонов с различными способностями использовать ресурсы окружающей среды, которые сосуществуют в популяции в водной среде обитания.

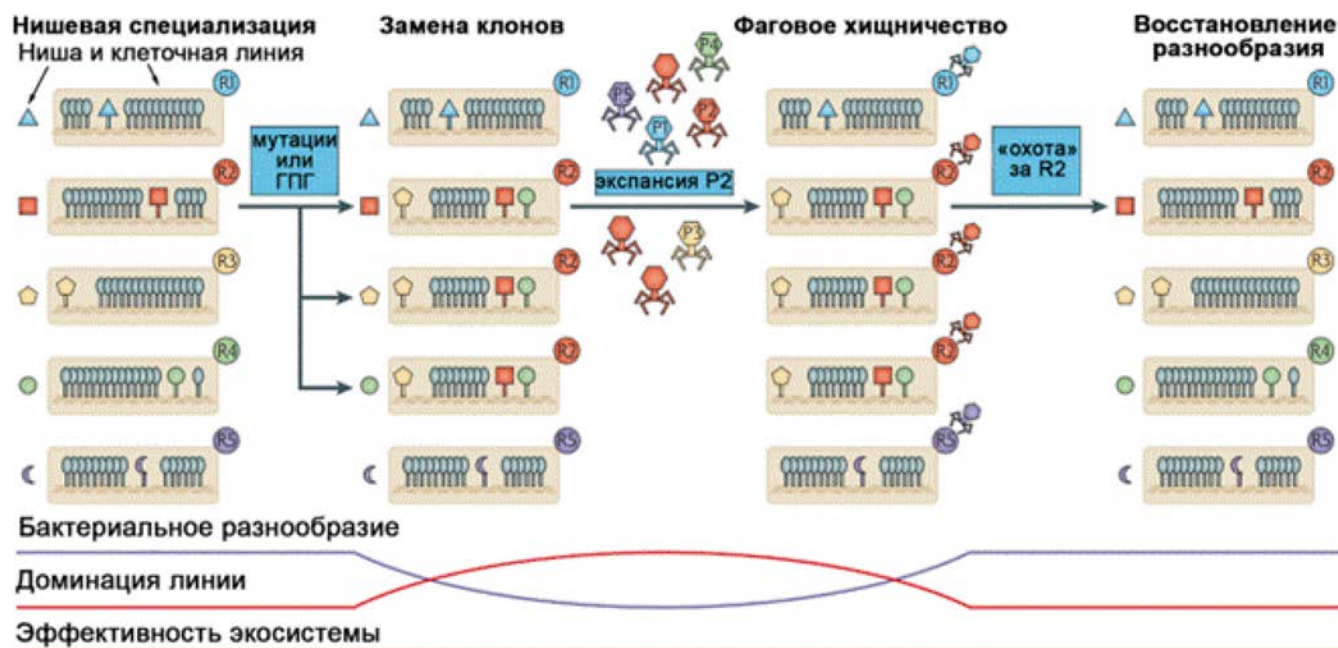


Рисунок 22 – Динамика популяции в условиях модели C-D. Различные ресурсы обозначены геометрическими фигурами, каждый из которых представляет свою

нишу. Адаптации, необходимые для использования ниши, представлены одинаковыми геометрическими фигурами в геноме линии. Эти линии различаются также набором рецепторов (R1–R5) к фагам (P1–P5), которые они несут. Каждый штамм содержит только один основной фаговый рецептор. Одна линия может повысить приспособленность за счет мутаций в геноме, введения новой ДНК путем горизонтального переноса генов (ГПГ) или изменения ниши, благоприятной для одного конкретного генотипа, и заменить другие линии. Внизу представлены изменения бактериального разнообразия, доминанции отдельной линии, эффективности экосистемы. Адаптировано из [110].

Одним из выводов данной модели является то, что многие версии молекулярных модулей, участвующих во взаимодействии фаг-бактерия, присутствуют одновременно в популяции вирусов, занимающей одну экологическую нишу. Примером может служить наличие у популяции фагов, занимающих одну нишу, вариаций доменов белков узнавания рецепторов бактерий, что позволяет прикрепляться к бактериям популяции, с таким же разнообразием рецепторов. Учитывая, что белки, способствующие воспроизведению фагов с неканоническими основаниями ДНК, могут также представлять модули взаимодействия фаг-бактерия (например, за счет участия в противодействии RM-системам бактерии-хозяина), модель Постоянного Разнообразия может быть использована при отражении процессов эволюционной дивергенции предков T4-родственных, связанных с неканоническими основаниями.

1.4. Заключение по обзору литературы

В обзоре литературы были рассмотрены современные представления о группе вирусов – T4-родственных бактериофагах, а также о неканонических азотистых основаниях ДНК фагов. Была рассмотрена история изменений в классификации данной группы вирусов и соответствие классификации T4-родственных бактериофагов с современной международной классификацией ICTV. Показано, что более поздние версии классификации ICTV лучше отражают родство

данной группы вирусов, чем предыдущие версии, что выражается в выделении отдельных семейств *Straboviridae* и *Kyanoviridae*, которые соответствуют Near-T4-родственным и Циано-T4-родственным.

Продемонстрированы проведенные ранее исследования пангенома T4-родственных бактериофагов, в которых были выделены кор- и квазикоргены T4-родственных бактериофагов, а также изучена топология коргенов в отдельных геномах (образование ими отдельных модулей, внутри которых расположены гены, чьи продукты выполняют схожие функции или участвуют в одном процессе).

В обзоре представлены известные неканонические основания ДНК бактериофагов, а также их изученные пути синтеза. Отдельно отмечено высокое разнообразие неканонических оснований и их модификаций как известных, так и гипотетических у T4-родственных бактериофагов. Рассмотрены белки, способствующие воспроизведению фагов с неканоническими основаниями ДНК.

Рассмотрены основные эколого-эволюционные модели, использующиеся при рассмотрении коэволюции бактерий-хозяев и их бактериофагов. Отдельно отмечена модель Постоянного Разнообразия, основанная на модели «Убий Победителя», которая может быть использована при рассмотрении влияния неканонических оснований ДНК на эволюцию предков T4-родственных бактериофагов.

Таким образом обзор литературы демонстрирует, что T4-родственные бактериофаги возможно использовать в качестве объектов для изучения того, как неканонические основания влияли и влияют на геномное разнообразие и эволюцию вирусов.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Материалы

2.1.1. Реактивы и реагенты

Триптон (Helicon, Москва), дрожжевой экстракт (Helicon, Москва), агар (Helicon, Москва), Трис (Helicon, Россия), агароза Standard Low -m_r (Bio-Rad Laboratories Inc, США), физиологический раствор хлористого натрия, 0,9 % (ОАО НПК “Эском”, Россия), борная кислота (ООО АО “Реахим”, Россия), этидий бромид (SERVA Electrophoresis GmbH, Германия), хлороформ (АО “ЭКОС-1», Россия), фенол (Panreac Química SLU, Испания), кумасси G-250 (ООО “ДиаМ”, Россия), бромфеноловый синий (ООО АО “Реахим”, Россия), уксусная кислота ледяная (ООО “Альдоса”, Россия), Твин 20 (Ferak Berlin GmbH, Германия), EDTA (SERVA Electrophoresis GmbH, Германия). Дополнительно использованные реактивы (марок “хч” и “чда”) получены от отечественных производителей.

2.1.2. Коммерческие наборы, ферменты, маркеры молекулярных масс, ДНК

ДНК λ (СибЭнзим, Россия), ДНК-маркер M1 (СибЭнзим, Россия), набор dNTP (СибЭнзим, Россия), Taq ДНК полимеразы с SE-буфером (СибЭнзим, Россия), набор реагентов для подготовки библиотек Nextera DNA Sample Preparation Kit (Illumina Inc., США), набор реагентов для секвенирования MiSeq® Reagent Kit v3 (600 cycle) (Illumina Inc., США).

2.1.3. Питательные среды

Среда LB, бульон, на 1 л: триптон – 10 г, NaCl – 10 г, дрожжевой экстракт – 4 г, дистиллированная вода – до 1 л общего объема среды.

Среда LB, полужидкая, на 1 л: триптон – 10 г, NaCl – 10 г, дрожжевой экстракт – 4 г, агар – 7 г, дистиллированная вода – до 1 л общего объема среды.

Среда LB, плотная, на 1 л: триптон – 10 г, NaCl – 10 г, дрожжевой экстракт – 4 г, агар – 15 г, дистиллированная вода – до 1 л общего объема среды.

Питательные среды стерилизовали при 121 °C и 1 атмосфере в течение 30 минут. Водородный показатель не учитывался.

2.1.4. Буферы и растворы

Фаговый раствор для выделения фагов из фекалий: NaCl – 0,2 М, Твин 20 – 1 мг/мл.

Фаговый буфер: Tris-HCl – 50 мМ, NaCl – 0,2 М, MgCl₂ – 10 мМ; pH = 7,4.

TBE буфер для электрофореза в агарозном геле: Tris – 0,1 М, борная кислота – 0,1 М, EDTA – 2 мМ; pH = 8,0.

Раствор этидий бромид (10 мг/мл).

Растворы CsCl плотностью 1,3 г/мл, 1,4 г/мл, 1,5 г/мл, 1,7 г/мл.

Раствор кумасси G-250 для окрашивания фаговых частиц в агарозном геле: кумасси G-250 – 0,05%, уксусная кислота – 10%.

2.1.5. Штаммы бактерий и бактериофагов

Штаммы *Escherichia coli*, использованные в работе:

- В [21, 111];
- BL21 ($F^- ompT gal dcm lon hsdSB(r_B^- m_B^-)$) [112];
- 5K ($F^- (r_K^- m_K^-) rpsL thr thi leu lacZ$) [113];
- 5KRI ($F^- (r_K^- m_K^-) (r_{RI}^+ m_{RI}^+) rpsL thr thi leu lacZ$) [114];
- JC5183 ($hr leuC argH his rpsL lacI phoS tol^+$) [115];
- JC5183/pLG13/ ($hr leuC argH his rpsL lacI phoS tol^+ pLG13 (r_{RV}^+ m_{RV}^+)$) [115, 116].

Лабораторные штаммы фагов, использованные в работе:

- T4B [20];
- T2L [16];
- RB43 [31];
- RB49 [31];
- T5D7 [117];
- λ vir (точечные мутации в oL2, oR1, и oR3, rep λ) [118].

2.2. Методы

2.2.1. Выделение и очистка бактериофагов из фекалий и сточных вод

Для выделения бактериофагов из сточных вод к пробе добавляли хлороформ в соотношении 10 к 1 (проба/хлороформ). Затем пробирку с пробой и хлороформом аккуратно перемешивали переворачиванием и оставляли при комнатной температуре на 10 мин.

Для выделения бактериофагов из фекалий к 2 г пробы добавляли фаговый раствор до общего объема 10 мл. Пробирку с полученной суспензией перемешивали на лабораторном ротаторе Rotamix RM-1 в течение часа при комнатной температуре. Отбирали супернатант и получали аликвоты. Аликвоты центрифугировали при 12000 об/мин на центрифуге Eppendorf MiniSpin plus. Отбирали супернатант и добавляли к нему хлороформ в соотношении 10 к 1 (супернатант/хлороформ).

Высевали очищенные пробы методом агаровых слоев на чашки Петри с питательной средой LB. Для этого необходимый объем пробы (100 мкл или 1 мл) и 100 мкл культуры чувствительного штамма бактерии добавляли к полужидкой питательной среде LB (предварительно растопленной и разогретой до 42 °С на водяном термостате) необходимого объема (4,5 – 5 мл среды). Аккуратно перемешивали и разливали смесь в чашки Петри с плотной питательной средой LB. После застывания верхнего слоя с полужидкой средой чашки инкубировали при 37 °С в течение ночи. На следующий день производили учет результатов. Негативные колонии различной морфологии пересеивали методом истощающегося штриха при помощи стерильных полосок бумаги с острым концом на чашки Петри со слоями плотной и полужидкой питательными средами (в верхний слой с полужидкой средой добавляли культуру с чувствительным штаммом *E.coli*). Пересев отдельной негативной колонии методом истощающегося штрихом производили 3-4 раза. Отдельные очищенные негативные колонии разных изолятов бактериофагов вырезались при помощи бактериальной петли и помещались в пробирку с 1 мл

жидкой питательной среды LB и 100 мкл хлороформа. Хранение штаммов бактериофагов осуществляли при температуре не выше 8 °С.

2.2.2. ТЭМ фаговых частиц фекалий

Пробы фекалий, очищенных по методике, представленной в «Выделение и очистка бактериофагов из фекалий и сточных вод», высевали методом агаровых слоев на штамм *E.coli* В. Инкубировали при 37 °С в течение ночи и добавляли 5 мл жидкой среды LB. Инкубировали при 4 °С в течение ночи. Отбирали жидкость сверху полужидкой среды, фильтровали ее шприцевыми фильтрами с диаметром пор 0,22 мкм. Фильтрат объемом 200 мкл помещали в центрифужные ультраконцентраторы (300000 MWCO) и центрифугировали при 10000 об/мин в течение 2 мин на MiniSpin plus. Отбирали концентрат. На металлическую сеточку, покрытую формваровой пленкой наносили концентрат и высушивали. Проводили контрастирование 1% раствором уранил ацетата. Трансмиссионная электронная микроскопия была произведена в ФИЦ ПНЦБИ РАН (ИБФМ РАН) к.б.н. Сузиной Натальей Егоровной на микроскопе JEM-150.

Анализ размеров фаговых частиц производился при помощи ПО ImageJ 1.50i [119].

2.2.3. Полимеразная цепная реакция

Для определения принадлежности бактериофагов к Т4-родственным бактериофагам и виду *Tequatrovirus T4* использовались вырожденные и специфические праймеры соответственно.

Для скрининга рабочей коллекции бактериофагов на наличие у них генетического маркера Т4-родственных бактериофагов использовались вырожденные праймеры, комплиментарные последовательности гена главного белка головки, свойственной данной группе вирусов. Последовательности праймеров (таблица 3) были взяты из литературных данных [120].

Таблица 3. Праймеры, отжигающиеся на последовательностях генов главного белка головки Т4-родственных бактериофагов

Название праймера	Последовательность праймера, комплементарная искомой	Tm, °C
MZIA1bis	5'-GATATTTGIGGIGTTCAGCCATGA-3'	56
MZIA6	5'-CGCGGTTGATTTCCAGCATGATTTC-3'	65,9

Для скрининга Т4-родственных бактериофагов рабочей коллекции на выявление у них генетических маркеров фага Т4, использовались праймеры, представленные в таблице 4.

Таблица 4. Праймеры, специфически отжигающиеся на консервативных участках генов РНК-лигазы 2 и Нос фага *Escherichia virus* Т4

Название праймера	Последовательность праймера, комплементарная искомой	Tm, °C
hocfR	5'-TTCAGCAAGATGCTTCGG-3'	54,08
inhrC	5'-ATCGGTCCAGTGCGTTTAGC-3'	59,43

Праймер inhrC отжигается на фрагменте гена РНК-лигазы 2 фага Т4, hocfR – последовательности 3 домена иммуноглобулин-подобного (Ig-подобного) белка Нос (head outer capsid protein) фага Т4 (один из трех Ig-подобных доменов, единственный, схожий с доменом I-set иммуноглобулинов) [121].

Выбор данного региона генома фага Т4 обусловлен тем, что белки Нос и РНК-лигаза 2 представляли интерес с точки зрения исследования особенностей вириона Т4-родственных бактериофагов, в случае Нос, и с точки зрения исследования компонентов пан-генома, в случае РНК-лигазы 2.

Белок Нос, как отмечалось выше, является иммуноглобулин-подобным и вносит основной вклад в иммуногенность вирионов Т4-родственных фагов в организме животных [122]. Считается, что данный белок позволяет вирионам прикрепляться к поверхностным гликопротеинам эукариотических и прокариотических клеток, а также, в зависимости от окружающих условий, позволяет агрегироваться или диспергироваться вирионам за счет Нос-Нос

взаимодействий. В результате, Нос способствует переживанию негативных условий среды и способствует более эффективному нахождению бактерий-хозяев [121]. Таким образом, Нос представляет интерес с точки зрения эволюции и экологии Т4-родственных бактериофагов. За счет того, что Т4-родственные бактериофаги занимают различные экологические ниши, отдельные домены гомологов Нос у представителей этой группы вирусов сильно варьируют, из-за необходимости взаимодействия с различными поверхностными гликопротеинами. Кроме того, показано, что Нос вносит существенный вклад в электрофоретическую подвижность фаговых частиц Т4-родственных фагов, за счет значительного заряда этого белка, влияющего на общий заряд частиц вирусов. Учитывая высокую вариабельность данного белка, одинаковых размеров частиц близкородственных вирусов, возможна дифференциация частиц близкородственных вирусов при помощи гель-электрофореза, что было продемонстрировано на ряде представителей Т4-родственных бактериофагов [123]. Метод гель-электрофореза фаговых частиц был также использован в данной работе.

РНК-лигаза 2 является ферментом, катализирующим реакцию репарации одонитевого разрыва между 5'-фосфорильной и 3'-гидроксильной группами соседних нуклеотидов в РНК дуплексах [124]. Интересно, что данный фермент широко используется как инструмент молекулярной биологии, например, в качестве компонента для секвенирования малых РНК. Хотя ферментативная активность данного фермента известна и используется в исследовательских целях, его биологическая роль у Т4-родственных бактериофагов до сих пор не выявлена. Исследование принадлежности гомологов гена данного белка к определенному компоненту пан-генома и распространения этого среди определенных групп Т4-родственных бактериофагов может способствовать определению этой роли, а также определению влияния данного фермента на эволюцию Т4-родственных вирусов. Такое исследование было также проведено в данной работе.

Специфичность праймеров, возможное образование вторичных шпилечных структур, содержание GC-оснований и температура плавления проверялась при помощи Primer-BLAST [125]. Специфичность проверялась по базе данных pr

(включает нуклеотидные последовательности баз GenBank, EMBL, DDBJ, PDB, RefSeq). Обнаружено, что полная специфичность имеется только у нуклеотидных последовательностей вида *Tequatrovirus T4*, размер которых также специфичен – 1276 п.о., а также отсутствие комплементарности в 1-2 основания имеется у наиболее близкородственных вирусов (например, T2). Таким образом, можно использовать данные праймеры для отсева T4 и наиболее близких по родству вирусов.

ПЦР проводили в объеме 15 мкл. Для обеих пар праймеров концентрации компонентов реакционной смеси для ПЦР были одинаковы: буфер для полимеразы – однократный, праймеры – 10 пмоль, dNTP – 200 мкМ (каждого), ДНК-полимераза – 2 единицы активности на реакцию.

ПЦР проводили на амплификаторе MJ Mini Bio-Rad, программы представлены в таблице 5.

Таблица 5. Программы амплификации, использованные для ПЦР с праймерами, отжигающиеся на нуклеотидных последовательностях главного белка капсида T4-родственных бактериофагов (пара MZIA1bis/MZIA6) и специфически отжигающиеся на консервативной нуклеотидной последовательности фага T4 (пара hocfR/inhrC).

Стадии (циклы)	Режим для MZIA1bis/MZIA6	Режим для hocfR/inhrC
1 стадия	95°C, 45 с	95°C, 45 с
2 стадия (3 цикла) Денатурация	95°C, 15 с	95°C, 15 с
Отжиг	50°C, 45 с	46°C, 45 с
Элонгация	72°C, 60 с	72°C, 90 с

Таблица 5. Продолжение.

Стадии (циклы)	Режим для MZIA1bis/MZIA6	Режим для hocfR/inhrC
3 стадия (30 циклов)		
Денатурация	95°C, 15 с	95°C, 15 с
Отжиг	55°C, 30 с	51°C, 30 с
Элонгация	72°C, 30 с	72°C, 45 с
4 стадия	72°C, 2 мин	72°C, 2 мин
5 стадия	4 °C	4 °C

Анализ результатов проводили при помощи горизонтального агарозного гель электрофореза (1,5% агарозы) при напряженности электрического поля 9 В/см.

2.2.4. 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. Получали методом агаровых слоев чашки с плотной и полужидкой питательной средой LB (в верхнем слое полужидкой среды находился чувствительный штамм бактерии). На поверхность среды наносились 5 мкл исходного лизата штамма бактериофага рабочей коллекции и его 100-кратное и 1000-кратное разведение. Чашки подсушивали при комнатной температуре и затем инкубировали при 37 °C в течение ночи. На следующий день производили учет результатов, по наличию или отсутствию ограничения роста бактериофагов.

Дополнительно, после отбора, для T4-родственных фагов, не являющихся видом *Tequatrovirus* T4, сравнивалась концентрация частиц на изогенных штаммах методом агаровых слоев.

2.2.5. Электрофорез фаговых частиц в агарозном геле

Методика электрофореза фаговых частиц взята из литературных данных [123]. Однако вместо комбинированного геля использовался агарозный. Как было показано позднее, такой гель также может быть использован [126]. Лизаты вместе с бромфеноловым синим наносились на 1% агарозный гель. Проводился горизонтальный электрофорез при напряженности электрического поля 9 В/см. Затем, гель окрашивали раствором кумасси G-250 при 37 °С в течении 15 часов. Обесцвечивание геля осуществлялось при помощи 10% уксусной кислоты. Производили анализ результатов.

2.2.6. Получение фаговых частиц в высокой концентрации, их очистка, выделение ДНК

Для получения фаговых частиц штаммов рабочей коллекции в высокой концентрации к 250 мл жидкой питательной среде LB добавляли 5 мл ночной культуры чувствительного штамма *E.coli*. Выращивали культуру до плотности $2,5 \cdot 10^8$ КОЕ/мл ($OD_{590} = 0,5$) при 37 °С при качании. Инокулировали бактериофаг с множественностью (MOI) 0,1 и выращивали в течение 5 часов при качании при 37 °С. После добавляли NaCl до 200 мМ и 5 мл хлороформа, перемешивали еще 15 минут. Добавляли РНКазу и ДНКазу до 20 мкг/мл. Лизат переносили в центрифужные бутылки и центрифугировали на роторе угловом Beckman JA-10 центрифуге Beckman J2-21 10 минут при 5000 об/мин при 10 °С. Отбирали супернатант и добавляли к нему ПЭГ6000 до 10%, перемешивали и оставляли при 8 °С на ночь. Затем производили центрифугирование на роторе угловом Beckman JA-10 центрифуге Beckman J2-21 30 минут при 10000 об/мин при 10 °С. Отбирали супернатант. Осадок растворяли в 15 мл фагового буфера. Полученный концентрат очищали.

Концентрат очищали центрифугированием в преформированном градиенте CsCl. Для этого наслаивали друг на друга растворы CsCl плотностью 1,7 г/мл – 2 мл, 1,5 г/мл – 3 мл, 1,4 г/мл – 3 мл –, 1,3 г/мл – 1 мл. Сверху наслаивали 3 мл концентрата. Ультрацентрифугирование производили на бакетном роторе Beckman

SW-41Ti на центрифуге Beckman L7-65 в течение 6 часов при 17 °C при 25000 об/мин. Отбирали фаговые частицы, расположенные примерно между плотностями 1,4 и 1,5 г/мл. Проводили диализ очищенных частиц от CsCl против 1000 объемов фагового буфера. Всего было произведено 2-3 смены. Фаг хранили при 4 °C.

Для выделения ДНК использовали 400 мкл очищенного фага и 200 мкл фенола, насыщенного Tris-HCl (pH=8,0). Центрифугировали на MiniSpin plus при 13000 об/мин в течение 90с. Отбирали водную фазу. Очистку ДНК производили 2-3 раза, до отсутствия видимой интерфазы. Добавляли 3 объема 96% этилового спирта и 200 mM NaCl, выдерживали 40 мин при -70 °C. Затем производили центрифугирование на MiniSpin plus при 13000 об/мин в течение 3 минут. Отбирали супернатант. Осадок промывали 100 мкл 70% этилового спирта. Осадок растворяли в деионизованной воде.

2.2.7. Секвенирование, сборка и аннотация геномов

Секвенирование и обработка ридов производилась к.б.н. Воложанцевым Н.В., Кисличкиной А.А., к.б.н. Богуном А.Г (ФБУН «ГНЦ ПМБ»).

Секвенирование производилось на платформе Illumina MiSeq (длина ридов составляла 300 п.о.). Для подготовки библиотек использовался набор Nextera DNA Library Preparation Kit, для секвенирования – MiSeq Reagent Kits v3.

Сборку ридов осуществляли Unicycler v.0.4.7 [127].

Аннотирование контигов геномов осуществлялось при помощи инструмента Genome Annotation (использует RASTtk [128]) сервиса PATRIC [129]. Аннотации проверяли при помощи GET_HOMOLOGUES [49], путем сравнения с аннотациями геномов родственных организмов. Геномные карты получали в CGview Server [130].

2.2.8. Сравнительный анализ геномов

Кластеризация гомологичных последовательностей геномов и аннотация

Аннотации геномов из баз данных проверялись на наличие открытых рамок считывания (ОРС) и, в случае их отсутствия, геномы автоматически аннотировались с помощью сервиса Genome Annotation PATRIC с параметрами

(Domain: Viruses) и (Annotation recipe: Phage). Аминокислотные последовательности, отобранные из аннотаций геномов, были распределены по гомологичным кластерам с помощью скрипта GET_HOMOLOGUES (get_homologues.pl) с параметрами, подобранными в ходе анализа. Для этого, проводилась кластеризация с параметрами минимальной идентичности и минимального покрытия равными 50% (данный параметр использовался в оригинальной работе по пангеному [46]). Затем проверялось наличие в геномах коргенов, определенных в [6]. Проводилось изменение параметров минимальной идентичности и минимального покрытия (с шагом 5 %) для получения наилучшего выделения кластеров коргенов.

Проверялись аннотации геномов на наличие всех коргенов. Если ОРС корового генома отсутствовали в анализируемом геноме, его последовательность была автоматически повторно аннотирована с помощью PATRIC, а затем снова проверена. Если ОРС корового гена была смещена, нарушена или отсутствовала, геном исключался из дальнейшего анализа.

В качестве названий коргенов были использованы соответствующие названия генов фага T4 [12].

Пангеномный анализ

Для анализа гомологичных белков, полученных скриптом get_homologues.pl, использовали пангеномную матрицу наличия/отсутствия кластеризованных последовательностей, сгенерированную с помощью скрипта compare_clusters.pl. Состав пангенома: коргеном, софткоргеном, оболочка, облако – определяли с помощью скрипта parse_pangenome_matrix.pl.

Филогеномные, филогенетические, пангеномные деревья и двудольные сети

На основе бинарной матрицы наличия/отсутствия кластеров получали пангеномное дерево с использованием скрипта estimate_pangenome_phylogenies.sh GET_PHYLOMARKERS [131] со 100 независимыми повторениями (-r). Филогенетические деревья были построены из множественных выровненных в MUSCLE [132] последовательностей с помощью W-IQ-TREE [133]. Также для

филогенетических деревьев использовалось ПО MEGAX [134], для исходного дерева применялся алгоритм максимальной экономии (Maximum Parsimony), далее – алгоритм максимального правдоподобия (Maximum likelihood), с использованием матрицы замен JTT [135] и bootstrap-анализа (число итераций – 1000). Филогеномные деревья были построены с использованием VICTOR [136] из 100 нуклеотидных последовательностей геномов (ограничение инструмента), которые сформировали отдельные клады вдоль пангеномного дерева. В ПО Cytoscape [137] были визуализированы двудольные кластер-геномные сети, в основе которых лежат белки, гены которых включены в необходимые для анализа компоненты пангенома.

Анализ гомологичных кластеров

Поиск гомологов кластеров, специфичных для определенной клады геномов, полученной на пангеномном дереве, и кластеров, способствующих воспроизведению фагов с неканоническими основаниями ДНК, проводили с помощью PSI-BLAST [138, 139] и HMMER [140].

Для первоначального поиска гомологов в PSI-BLAST использовали аминокислотную последовательность из кластера, принадлежащую одному из вирусных геномов. Затем был получен список гомологичных белков. Были отобраны последовательности вирусов, принадлежащих к *Tevenvirinae*. Поиск был повторен (1 итерация). Если кластер был образован одним белком и после первоначального поиска не было найдено других гомологичных последовательностей *Tevenvirinae*, то поиск не повторялся. Параметры поиска были следующими: (Max target sequences: 1000), (Word size: 2), (Matrix: BLOSUM45 (в случае превышения CPU – BLOSUM62)). Для уточнения аннотаций кластеров, в которых гомологи обнаружены только среди гипотетических белков, дополнительно использовали HMMER. Поиск проводили с использованием стандартных параметров, базы данных nr30 и матрицы на основе белковых последовательностей, выровненных в MUSCLE.

Отдельные регионы геномов, содержащие гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, сравнивали с помощью Easyfig [141].

ГЛАВА 3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ И ИХ ОБСУЖДЕНИЕ

3.1. Т4-родственные бактериофаги с неканоническими основаниями из фекалий зубров и сточных вод

3.1.1. Бактериофаги из фекалий зубров и сточных вод очистных сооружений

Источники для выделения фагов

Выделение бактериофагов проводилось из двух источников, которые обычно используют для этого – фекалии и сточные воды. Это обусловлено высоким разнообразием и легкостью выделения из сточных вод, а также подбором фагов для терапии кишечных заболеваний из фекалий.

Отбор проб фекалий зубров для выделения бактериофагов производилось в Приокско-Тerrasном государственном биосферном заповеднике имени Михаила Заблoцкого в питомнике зубров у взрослых особей (рисунок 23) и молодняка вида *Bison bonasus*.



Рисунок 23 – взрослый самец *Bison bonasus*.

Отбор проб сточных вод проводился на очистных сооружениях города Пущино Московской области на различных этапах очистки воды: первичные отстойники, аэротенки, вторичные отстойники. Также отбиралась неочищенная вода.

Первичный микробиологический анализ негативных колоний сточных вод

Были определены концентрации фаговых частиц на различных этапах очистки сточных вод очистных сооружений (таблица 6).

Таблица 6. Концентрация фаговых частиц из проб сточных вод очистных сооружений, полученных 07.07.2016 на разных штаммах *E.coli*.

Штамм <i>E.coli</i>	Место отбора воды	Концентрация, БОЕ/мл
В	Вход	$6,3 \cdot 10^2$
	Первичный отстойник	$1,08 \cdot 10^3$
	Аэротенк	$7 \cdot 10^1$
	Вторичный отстойник	$7 \cdot 10^1$
BL21	Вход	$1,25 \cdot 10^3$
	Первичный отстойник	$1,21 \cdot 10^3$
	Аэротенк	$2 \cdot 10^2$
	Вторичный отстойник	0

Вход – неочищенная вода.

Можно отметить, что концентрация фаговых частиц на этапе первичного отстойника выше или близка к концентрации в неочищенной воде. Такой результат вероятнее всего связан с отбором воды в первичном отстойнике с верхних слоев до достаточного осаждения частиц, также возможно попадание частиц с поверхности воды.

Негативные колонии бактериофагов из проб сточных вод, образующиеся на штамме *E.coli* В показаны на рисунке 24.

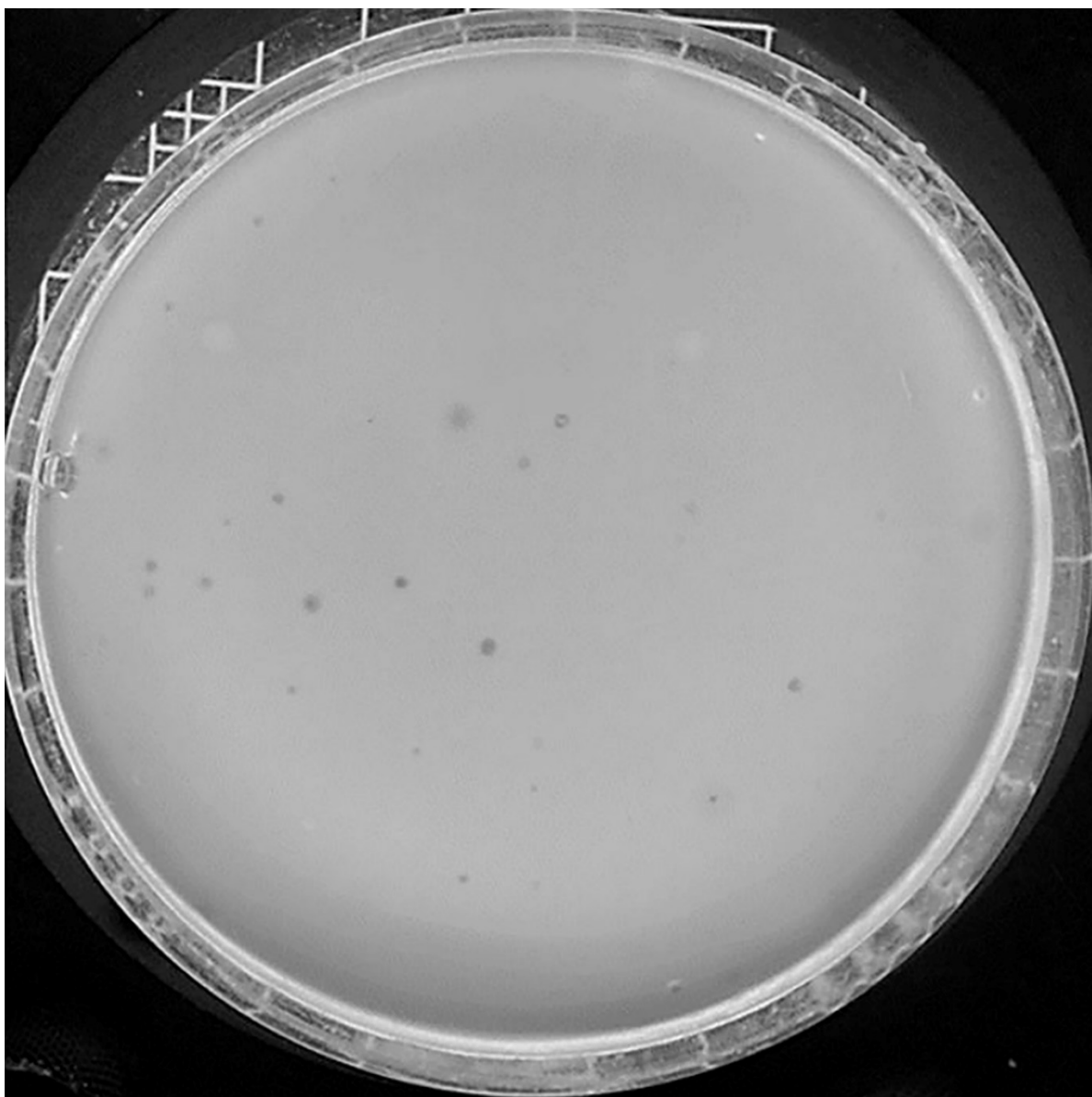


Рисунок 24 – негативные колонии бактериофагов из сточных вод, образующиеся на штамме В. По [142].

При высеве сточных вод различной степени очистки на чувствительный штамм В отмечалось образование в основном негативных колоний от 0,8 до 1,5 мм мутных или прозрачных с нечетко очерченными краями. Также обнаружены бляшки размером до 2,8 мм мутные с расплывчатым ореолом. Отмечались негативные колонии общим диаметром 3,7 мм, с прозрачным центром диаметром 0,6 - 0,7 мм и мутным ореолом.

Негативные колонии бактериофагов из проб сточных вод, образующиеся на штамме *E.coli* BL21 показаны на рисунке 25.

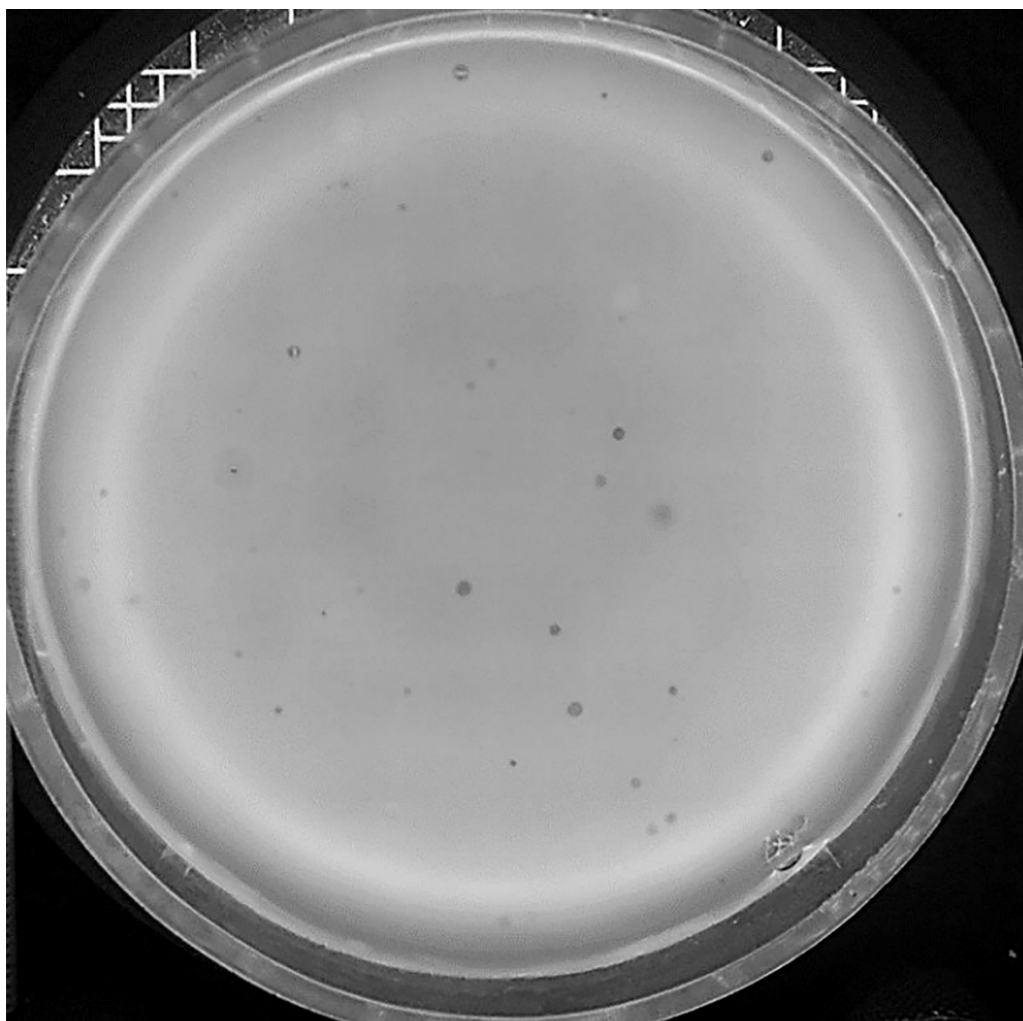


Рисунок 25 – негативные колонии бактериофагов из сточных вод, образующиеся на штамме BL21.

При высеве сточных вод с очистных сооружений с различных этапов очистки на штамм *E.coli* BL21 в основном обнаруживались прозрачные негативные колонии 0,5 до 1,7 мм. Были отмечено образование бляшек с расплывчатым краем диаметром 3 мм. Также найдены негативные колонии общим диаметром 4 мм, с прозрачным центром диаметром 0,4 - 0,5 мм и мутным ореолом.

В целом отмечалось образование схожих колоний на штаммах В и BL21. Учитывая наличие у штамма В системы рестрикции-модификации, что может способствовать увеличению представленности фагов с системами антирестрикции, а также то, что данный штамм исходно использовался для выделения первых представителей Т4-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 на штамме В, на который высевались очищенные пробы фекалий зубров.

После очистки изолятов путем пересевов, была создана рабочая коллекция со следующим названием штаммов:

PuMWBn – штаммы бактериофагов, выделенные из сточных вод очистных сооружений на штамме *E.coli* В, n – номер.

nВ – штаммы бактериофагов, выделенные из фекалий зубров на штамме *E.coli* В, n – номер.

ТЭМ фаговых частиц из фекалий зубров из лизатов штамма В

Так как фекалии зубров являлись относительно новым источником для выделения бактериофагов, произведена трансмиссионная электронная микроскопия фагов из фекалий. Анализ снимков результатов ТЭМ фаговых частиц из фекалий зубров из сконцентрированных и очищенных лизатов *E.coli* В показал наличие частиц, принадлежащих к трем морфотипам: А1, А2 и В2 (морфотипы представлены на рисунке 26, фаговые частицы – на рисунке 27).

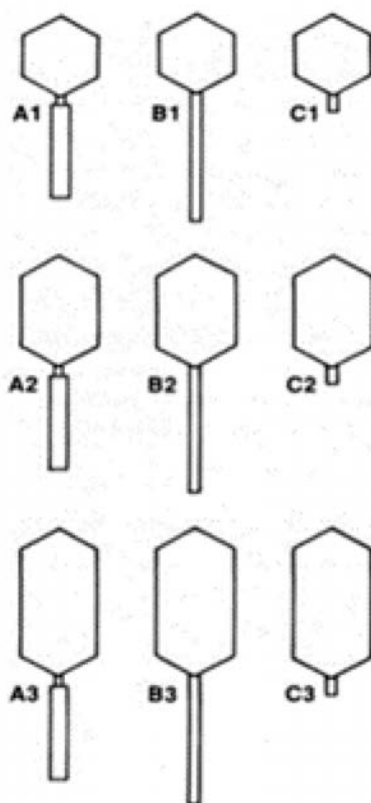


Рисунок 26 – морфотипы хвостатых фагов.

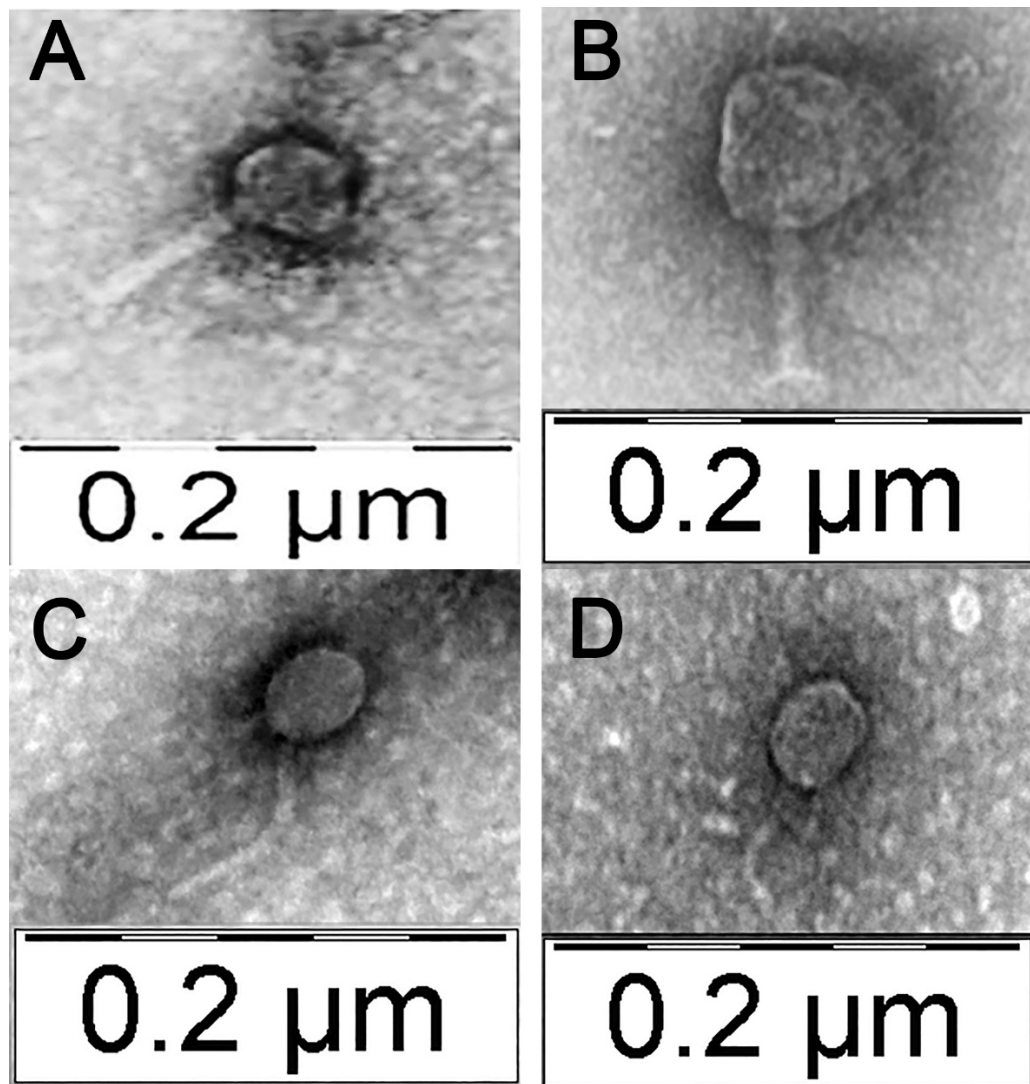


Рисунок 27 – фаговые частицы из фекалий зубров. А – морфотип А1, В – морфотип А2, С и D – морфотип В1.

В таблице 7 представлены размеры обнаруженных частиц.

Таблица 7. Размер частиц фагов из фекалий зубров.

Пример частицы на рисунке 27	Морфотип	Число использованных частиц для анализа размеров	Размеры частиц		
			Длина головки, мкм (SD)	Ширина головки, мкм (SD)	Длина хвоста, мкм (SD)
A	A1	19	0,062 (0,001)	0,062 (0,001)	0,075 (0,002)
B	A2	10	0,077 (0,002)	0,058 (0,001)	0,078 (0,002)
C	B1	11	0,045 (0,002)	0,037 (0,002)	0,089 (0,002)
D	B1	6	0,045 (0,004)	0,037 (0,002)	0,061 (0,005)

Отмечено соответствие частиц на рисунке 27 и в таблице, морфотип частиц, число использованных частиц для анализа размеров, размеры частиц (в скобках указано стандартное отклонение).

Для раздела “Бактериофаги из фекалий зубров и сточных вод очистных сооружений” можно отметить следующее. Определено наличие бактериофагов, образующих различные негативные колонии на штамме *E.coli* B из проб сточных вод. Обнаружено преобладание бляшек относительно малого размера, что свойственно T4-родственным бактериофагам, обладающим белками R. За счет этих белков у данных вирусов имеется особый механизм – синхронное ингибирование лизиса [143], в результате действия которого высокая концентрация вирусных частиц ингибирует выход новообразованных вирусов, что проявляется в размерах бляшек. Для бактериофагов из фекалий зубров была проведена ТЭМ, по результатам которой было обнаружено, что в преобладающем большинстве бактериофаги, способные инфицировать штамм *E.coli* B, представлены хвостатыми фагами. Хотя бактериофаги размером, свойственным T-четным или псевдо-T-четным (в этих группах были обнаружены фаги, инфицирующие *E.coli*) [144], не были обнаружены, что могло быть связано с отсутствием таких фагов в отобранной пробе, наличие T4-родственных фагов в фекалиях зубров не исключается, учитывая, что их часто обнаруживают в фекалиях животных.

3.1.2. Скрининг бактериофагов рабочей коллекции на принадлежность к T4-родственным бактериофагам и наличие у них антирестрикционных систем

Спот тест на ограничение роста фагов на штаммах с системами рестрикции-модификации

При проведении спот-теста на ограничение роста фагов на штаммах с системами рестрикции-модификации было обнаружено три возможных ответа: рост без ограничений по сравнению с изогенным штаммом без RM-систем (рисунок 28); отсутствие роста на штаммах с RM-системами (рисунок 29); ограничение роста по сравнению с изогенным штаммом без RM-систем (рисунок 30).

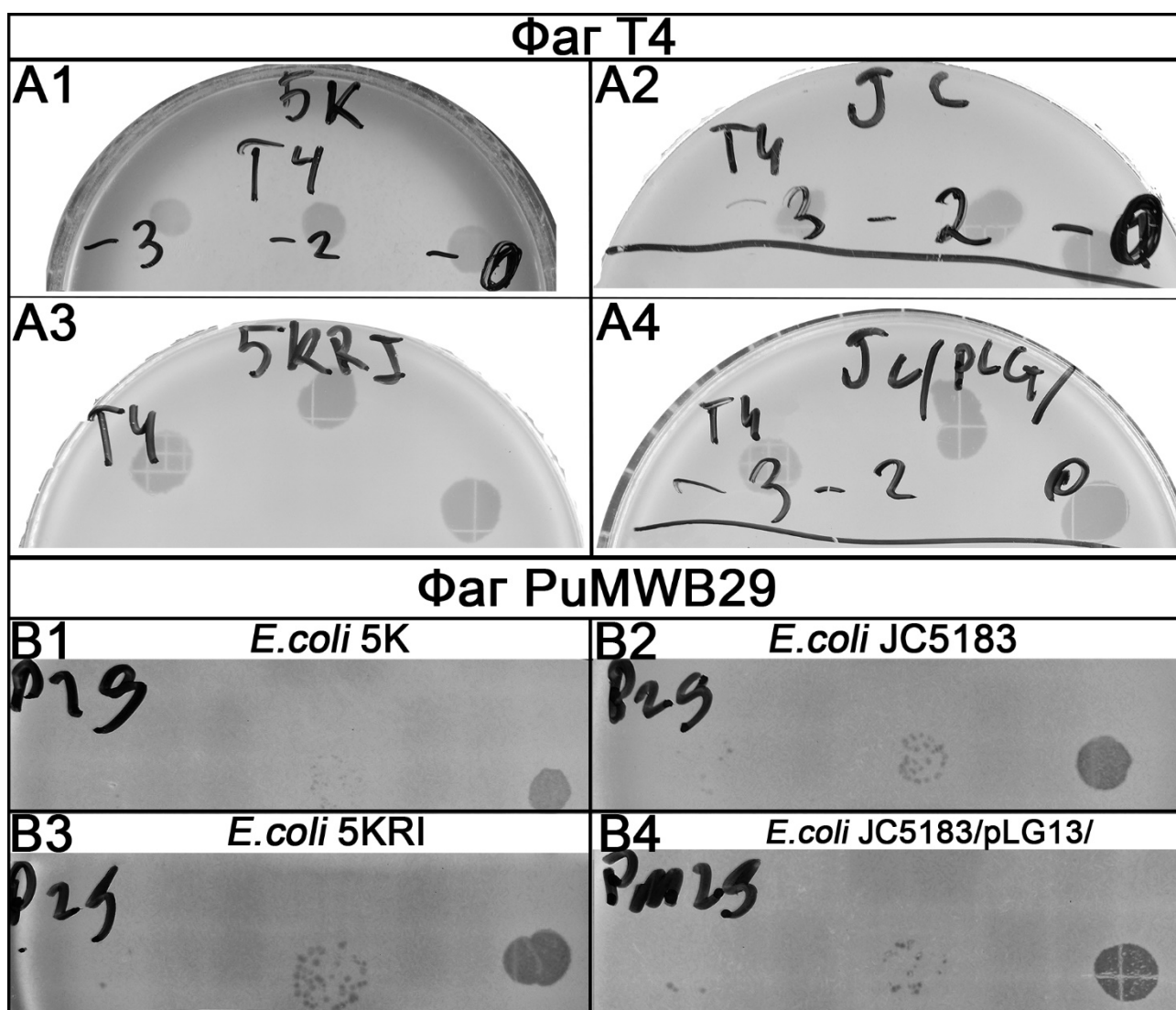


Рисунок 28 – Спот-тест фагов на ограничение роста на штаммах с RM-системами (ограничение роста отсутствует). Фаг T4 (A1-A4)– бактериофаг T4B, использовался в качестве контроля, фаг PuMWB29 (B1-B4)– пример результата с фагом из рабочей коллекции. A1-A4, B1-B4 – продемонстрированы чашки с газонм чувствительного штамма *E. coli*, на чашках с газонм слева наносилось тысячекратное разведение лизата с фагом, по центру – стократное разведение, справа – исходный лизат с фагом. A1, B1 – использовался штамм *E. coli* 5K; A2, B2 – *E. coli* JC5183; A3, B3 – *E. coli* 5KRI (обладает эндонуклеазой рестрикции EcoRI); A4, B4 – JC5183/pLG13/ (обладает эндонуклеазой рестрикции EcoRV).

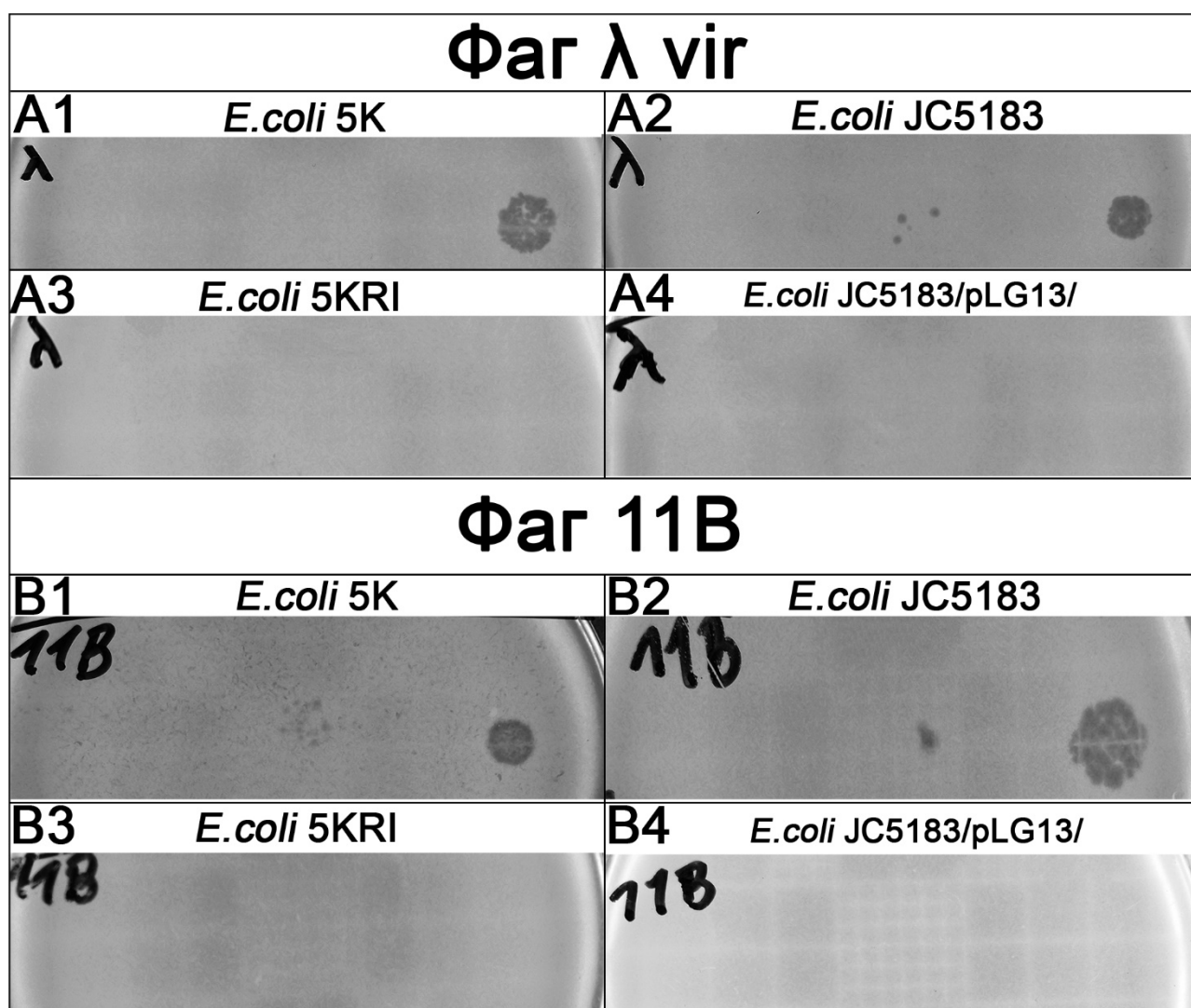


Рисунок 29 – Спот-тест фага на ограничение роста на штаммах с RM-системами (отсутствие роста на штаммах с RM-системами). Фаг λ – пример результата с фагом, который использовался в качестве контроля, фаг 11B – пример результата с фагом из рабочей коллекции. A1-A4, B1-B4 – продемонстрированы чашки с газоном чувствительного штамма *E.coli*, на чашках с газоном слева наносилось тысячекратное разведение лизата с фагом, по центру – стократное разведение, справа – исходный лизат с фагом. A1, B1 – использовался штамм *E.coli* 5K; A2, B2 – *E.coli* JC5183; A3, B3 – *E.coli* 5KRI (обладает эндонуклеазой рестрикции *EcoRI*); A4, B4 – JC5183/pLG13/ (обладает эндонуклеазой рестрикции *EcoRV*).

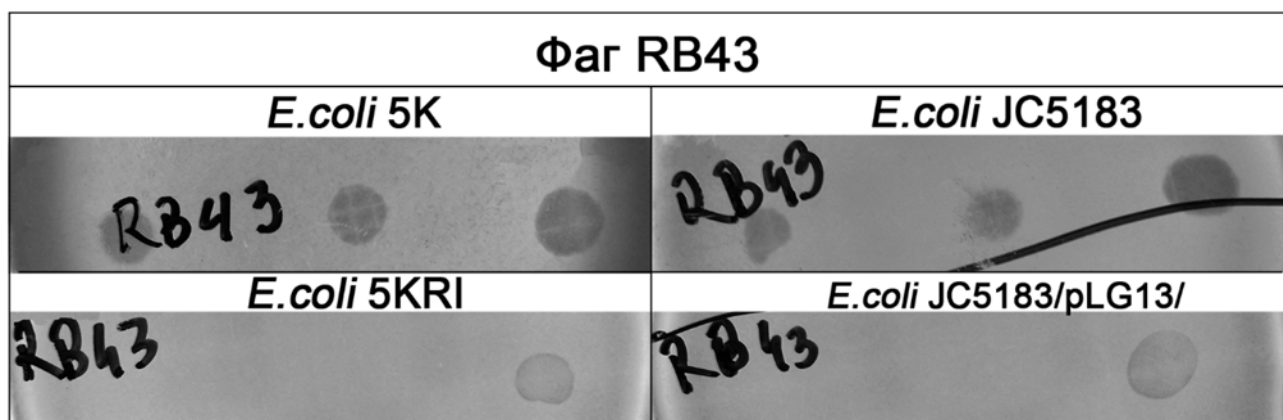


Рисунок 30 – Спот-тест фага RB43, который использовался в качестве контроля, на ограничение роста на штаммах с RM-системами. Обнаружено частичное ограничение роста на штаммах с RM-системами по сравнению с изогенными штаммами без них (ограничение на 2 порядка, а также образование мутного литического пятна). Продемонстрированы чашки с газоном чувствительных штаммов *E.coli*, на рисунке отмечены их названия. На чашках с газоном слева наносилось тысячекратное разведение лизата с фагом, по центру – стократное разведение, справа – исходный лизат с фагом. Схожий ответ у фагов из рабочей коллекции обнаружен не был.

При рассмотрении роста лабораторных штаммов бактериофагов (T2L, T4B, λ , RB43, RB49), обнаружено, что рост T2L, T4B, которые обладают известными неканоническими основаниями, не ограничен на штаммах с рассматриваемыми RM-системами, тогда как фаг λ , обладающий модифицированным основанием не образовывал негативные колонии. В случае RB43 обнаруживалось частичное ограничение роста, однако неизвестно имеются ли неканонических основания у данного вируса, также, как и у RB49, у которого рост не ограничивался. Таким образом у данных вирусов могут присутствовать неизвестные на данный момент механизмы защиты от этих RM-систем, включая неканонические или модифицированные основания.

Спот-тест на контрольных бактериофагах показал пригодность его использования в целях скрининга для поиска вирусов с антиристрикционными механизмами, способных защитить от действия определенной RM-системы.

В таблице 8 представлены результаты спот-теста на ограничение роста фагов рабочей коллекции на штаммах *E.coli* с RM-системами.

Таблица 8. Результаты спот-теста на ограничение роста фагов рабочей коллекции и лабораторных фагов на штаммах с RM-системами.

Фаги	Рост на штаммах с RM-системами	
	EcoRI	EcoRV
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-PuMWB50	не ограничен	не ограничен

«Не ограничен» – фаги одинаково растут на изогенных штаммах с и без RM-системы; «отсутствует» – фаги не растут на штамме с RM-системой, без – растут; «неизвестно» – не растут на изогенных штаммах с и без RM-системы. PuMWBn – бактериофаги, выделенные из сточных вод, nB – выделенные из фекалий зубров. Через дефис представлены названия фагов, чья нумерация идет по порядку, и они имеют одинаковый результат спот-теста.

По результатам спот-теста можно отметить, что у подавляющего большинства фагов (58 фагов из 62) рост на штаммах с RM-системами не ограничивался. Лишь 4 фага не росли на этих штаммах. Кроме того, PuMWB9 не рос как на штамме *E.coli* 5KRI, так и 5K.

ПЦР для поиска T4-родственных фагов, отличных от вида T4

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 получены результаты, представленные в таблице 9.

Таблица 9. Результаты ПЦР для определения принадлежности фагов рабочей коллекции к Т4-родственным бактериофагам и близкородственным виду *Tequatrovirus* T4.

Фаги	Принадлежность к Т4-родственным	Принадлежность к близкородственным виду Т4
1В-2В	-	-
3В-10В	+	+
11В	-	-
12В	+	+
PuMWB1-PuMWB8	+	+
PuMWB9	-	-
PuMWB10	+	+
PuMWB11	+	-
PuMWB12	+	+
PuMWB13	+	-
PuMWB14-PuMWB20	+	+
PuMWB21	+	-
PuMWB22 - PuMWB23	+	+
PuMWB24	+	-
PuMWB25-PuMWB29	+	+
PuMWB30	+	-
PuMWB31 - PuMWB35	+	+
PuMWB36	+	-
PuMWB37-PuMWB39	+	+
PuMWB40	+	-
PuMWB41-PuMWB50	+	+

1 столбец – через дефис представлены названия фагов, чья нумерация идет по порядку, и они имеют одинаковый результат ПЦР; 2 столбец – наличие или отсутствие целевого ампликона, полученного при ПЦР с праймерами, отжигающихся на генах главного белка капсида Т4-родственных фагов; 3 столбец – наличие целевого ампликона, полученного при ПЦР со специфическими праймерами к последовательности вида *Tequatrovirus* T4. «+» – наличие целевого ампликона по результатам ПЦР, «-» – отсутствие целевого ампликона. PuMWBn – бактериофаги, выделенные из сточных вод, nВ – выделенные из фекалий зубров.

Электрофореграммы результатов ПЦР представлены в приложении 1.

По результатам ПЦР можно отметить, что подавляющее большинство фагов рабочей коллекции (58 из 62) имели в составе геномной ДНК последовательность гена главного белка капсида, свойственного Т4-родственным бактериофагам, а также 51 из них имели в составе геномной ДНК консервативную последовательность, свойственную близкородственным виду *Tequatrovirus* T4.

Результаты скрининга рабочей коллекции методами ПЦР и спот-теста представлены в сводной таблице 10.

Таблица 10. Результаты спот-теста на ограничение роста на штаммах *E.coli* с RM-системами и ПЦР анализов для выявления принадлежности к Т4-родственным бактериофагам и к близкородственным виду *Tequatrovirus* T4 лабораторных штаммов вирусов и штаммов рабочей коллекции.

Фаги	Принадлежность к Т4-родственным	Принадлежность к близкородственным Т4	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
λ 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

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 Т4-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.

Таблица 10. Продолжение

Фаги	Принадлежность к Т4-родственным	Принадлежность к близкородственным Т4	Рост на штаммах с RM-системами	
			EcoRI	EcoRV
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
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

Результаты скрининга для каждого фага рабочей коллекции по отдельности представлен в приложении 2.

В результате было определено, что большинство фагов рабочей коллекции – 58 из 62 фагов – принадлежало к Т4-родственным бактериофагам. Причем только 4 фага, не принадлежащих к Т4-родственным, не росли на штаммах с RM-системами. Среди Т4-родственных бактериофагов рабочей коллекции большинство – 51 из 58 фагов являлись близкородственными виду Т4. Лишь 7 фагов: PuMWB11, PuMWB13, PuMWB21, PuMWB24, PuMWB30, PuMWB36, PuMWB40, не имели целевого ампликона, свойственного близкородственным виду Т4.

Подавляющее преобладание фагов, близкородственных виду Т4, может объясняться во-первых, выбором в качестве чувствительного штамма бактерий *E.coli* В, во-вторых, при рассмотрении геномов Т4-родственных бактериофагов в

базах данных, можно отметить преобладание близкородственных Т4 вирусов среди фагов данной группы, что также свидетельствует о их большей встречаемости в природе, по сравнению с другими представителями данной группы вирусов.

Электрофорез фаговых частиц

Для ряда Т4-родственных бактериофагов показано, что основным белком, влияющим на общий заряд частицы является поверхностный белок Нос. За счет высокой вариабельности данного белка, разница в общем заряде частиц имеется даже у близкородственных вирусов, причем эта разница отражается в электрофоретической подвижности частиц [123]. За счет этого возможно проведение дифференциации Т4-родственных бактериофагов при помощи гель-электрофореза.

Фаговые частицы PuMWB11, PuMWB13, PuMWB21, PuMWB24, PuMWB30, PuMWB36, PuMWB40 использовались для проведения нативного электрофореза частиц. Результат представлен на рисунке 31.

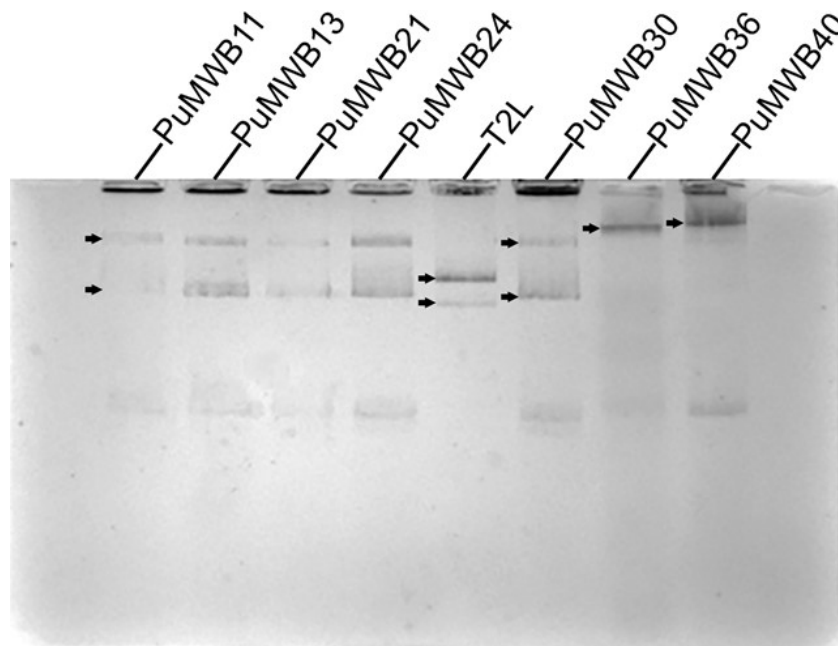


Рисунок 31 – Электрофореграмма фаговых частиц Т4-родственных бактериофагов. Стрелками отмечены целевые полосы. T2L – лабораторный штамм фага близкородственного виду Т4; PuMWB11, PuMWB13, PuMWB21, PuMWB24, PuMWB30, PuMWB36, PuMWB40 – Т4-родственные бактериофаги рабочей коллекции, не являющиеся близкородственными виду Т4 и содержащие антиристрикционные системы.

На рисунке 31 для ряда фагов отмечены две целевые полосы, которые представляют собой целые частицы и головки фагов без хвостов. Как показано для фага T4 такое явление часто встречается при электрофорезе частиц [145].

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 было взято 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.

3.1.3. Анализ геномов выбранных при скрининге бактериофагов

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. Карты аннотированных геномов представлены на рисунках 32-35. В качестве названия генов использованы названия гомологичных генов фага T4.

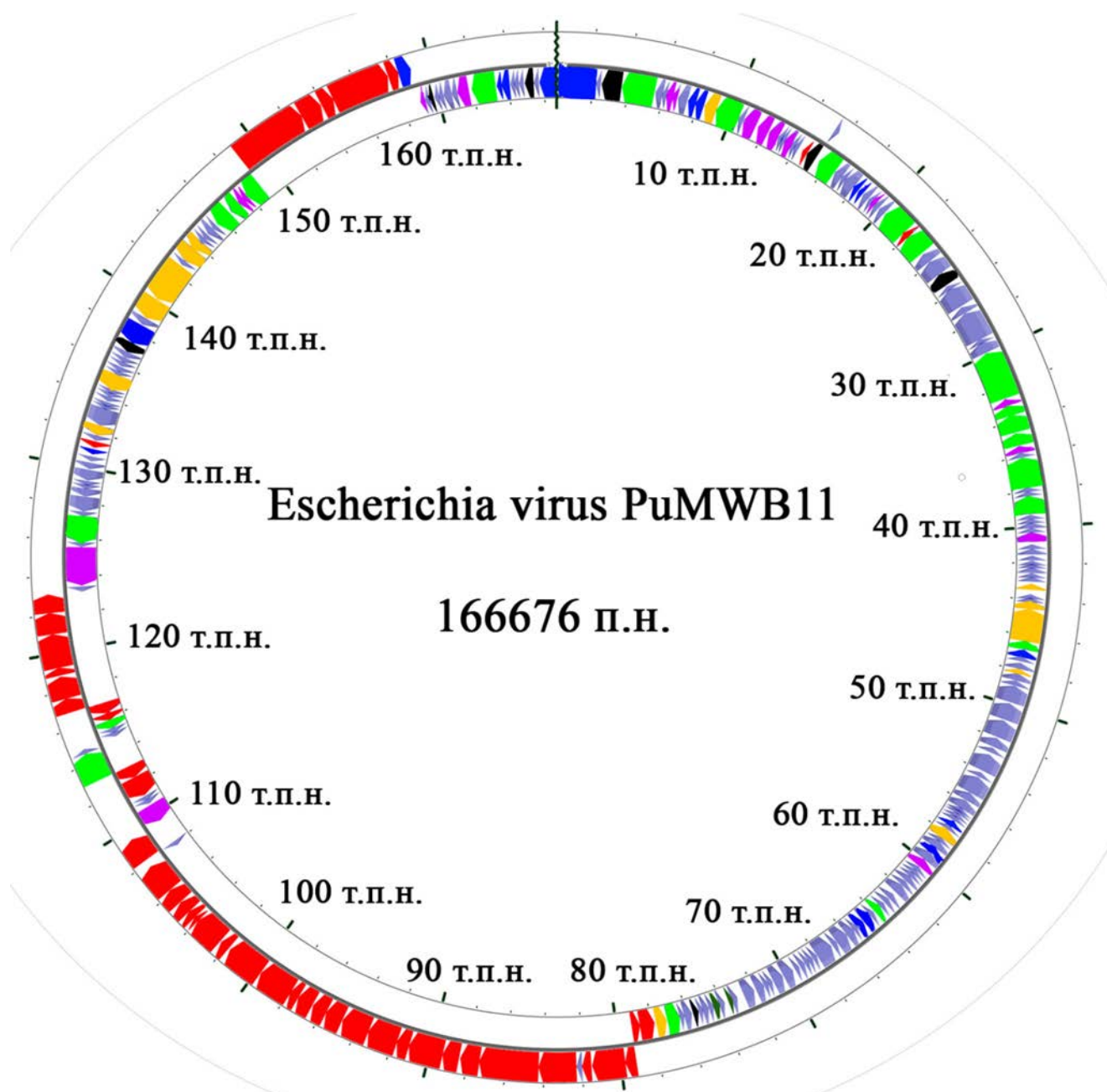


Рисунок 32 – геномная карта фага PuMWB11. Красным отмечены гены, чьи продукты участвуют в морфогенезе частиц; желтым – в метаболизме нуклеотидов; зеленым – в репликации, репарации рекомбинации ДНК; фиолетовым – в экспрессии генов; черным – способствуют воспроизведению фагов с неканоническими основаниями ДНК; синим – выполняют другие функции; голубым – гипотетические белки. «П.н. – пар нуклеотидов».

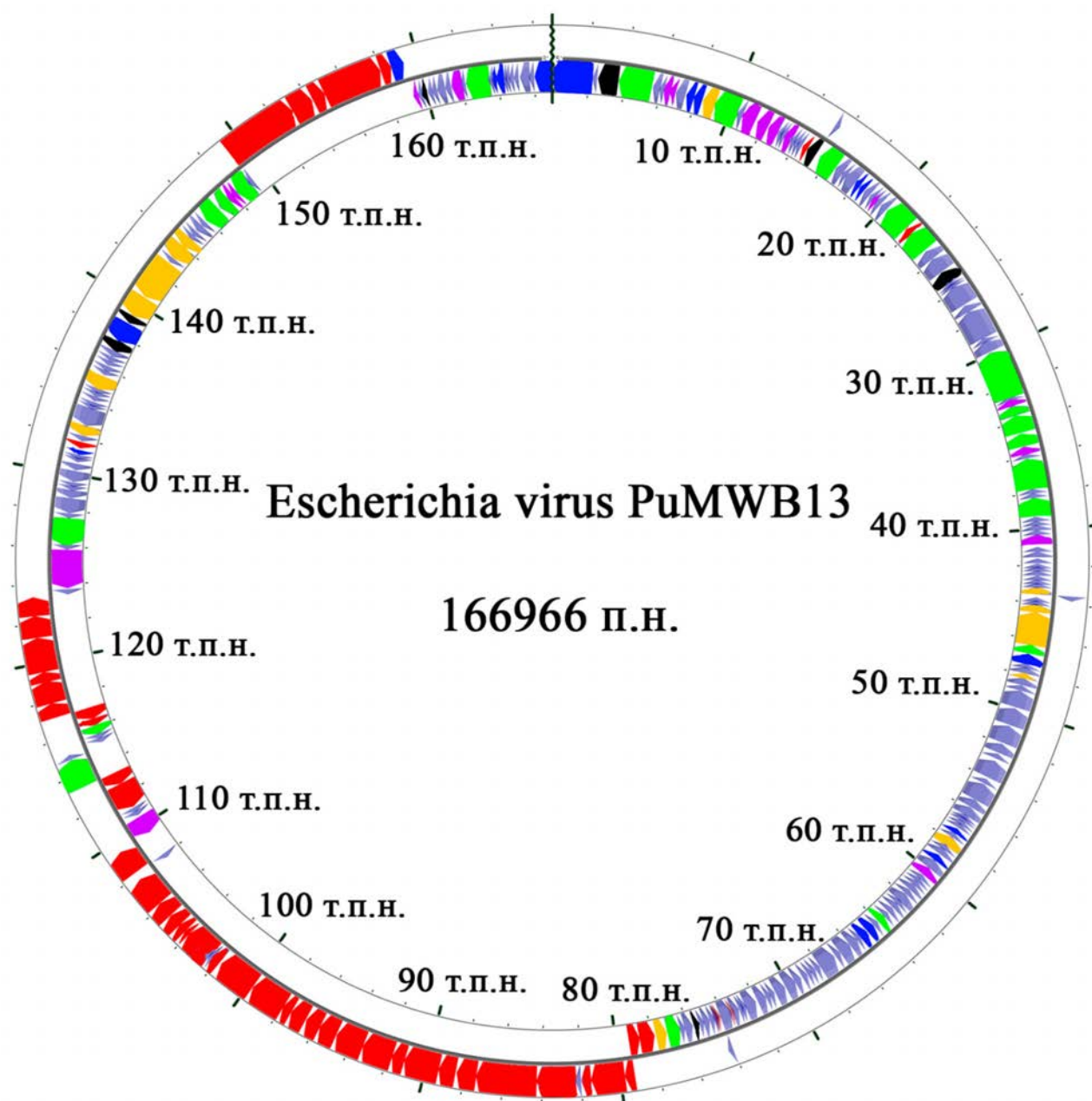


Рисунок 33 – геномная карта фага PuMWB13. Красным отмечены гены, чьи продукты участвуют в морфогенезе частиц; желтым – в метаболизме нуклеотидов; зеленым – в репликации, репарации рекомбинации ДНК; фиолетовым – в экспрессии генов; черным – способствуют воспроизведению фагов с неканоническими основаниями ДНК; синим – выполняют другие функции; голубым – гипотетические белки. «П.н. – пар нуклеотидов.

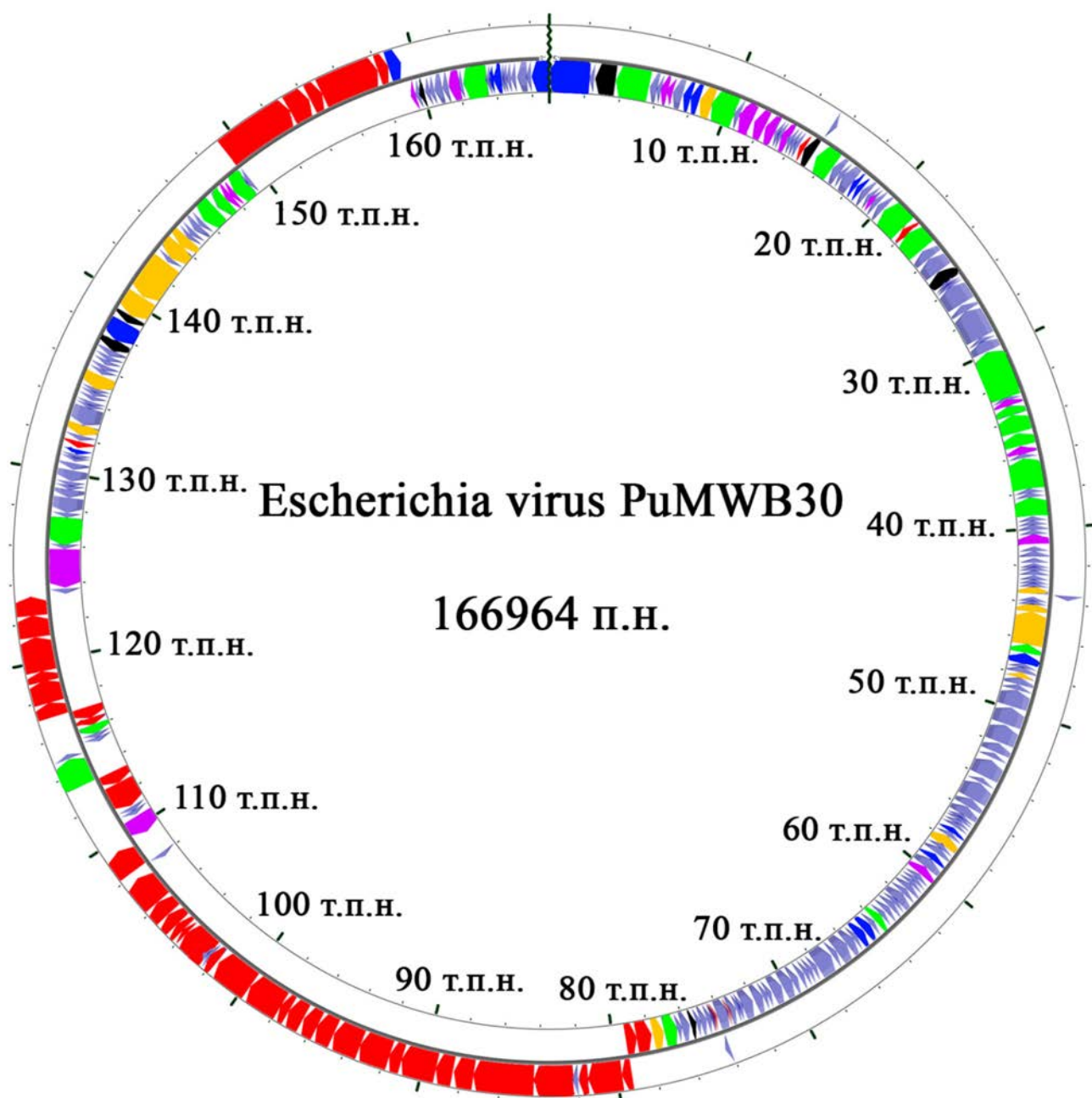


Рисунок 34 – геномная карта фага PuMWB30. Красным отмечены гены, чьи продукты участвуют в морфогенезе частиц; желтым – в метаболизме нуклеотидов; зеленым – в репликации, репарации рекомбинации ДНК; фиолетовым – в экспрессии генов; черным – способствуют воспроизведению фагов с неканоническими основаниями ДНК; синим – выполняют другие функции; голубым – гипотетические белки. «П.н. – пар нуклеотидов.

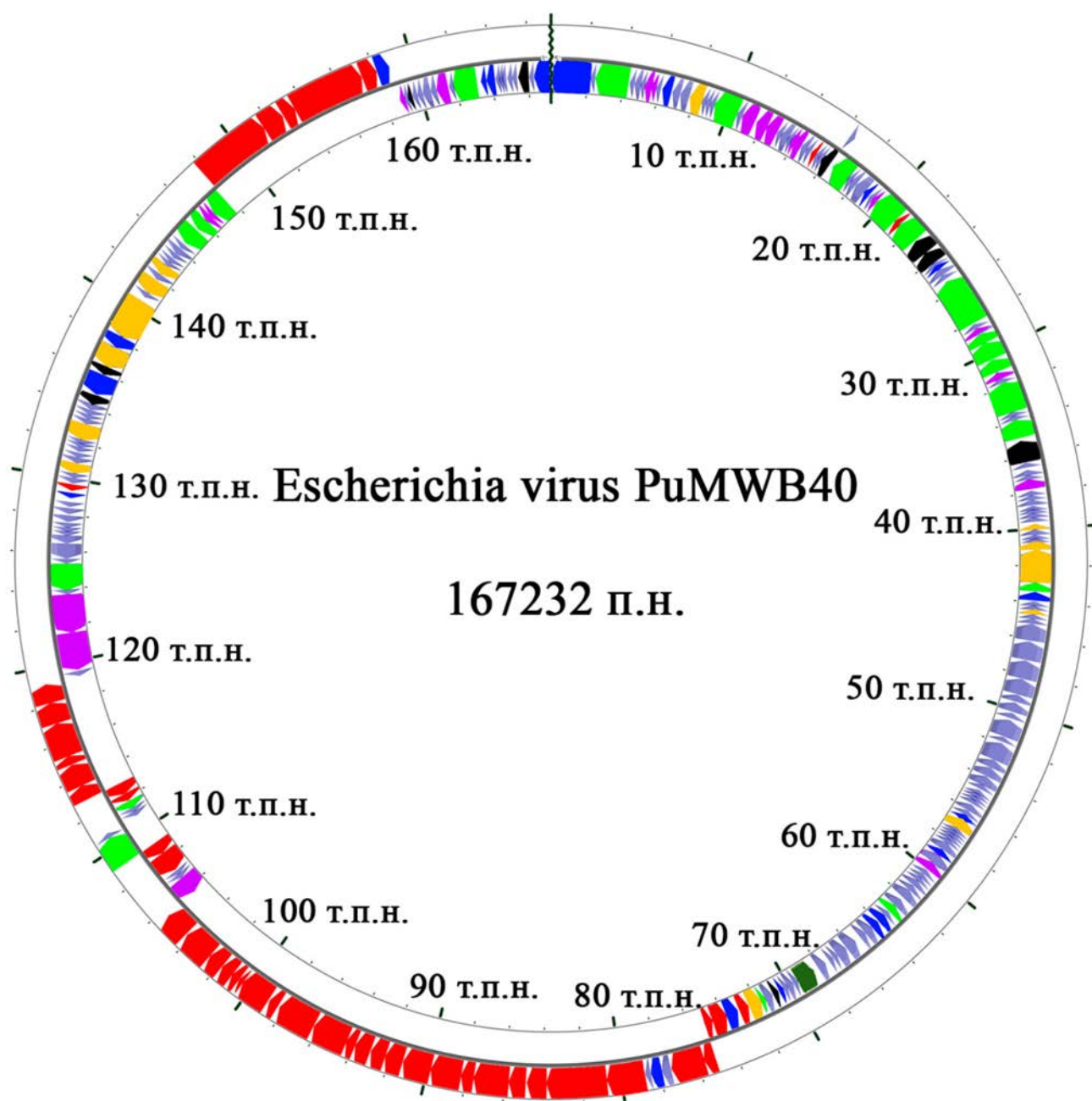


Рисунок 35 – геномная карта фага PuMWB40. Красным отмечены гены, чьи продукты участвуют в морфогенезе частиц; желтым – в метаболизме нуклеотидов; зеленым – в репликации, репарации рекомбинации ДНК; фиолетовым – в экспрессии генов; черным – способствуют воспроизведению фагов с неканоническими основаниями ДНК; синим – выполняют другие функции; голубым – гипотетические белки. «П.н. – пар нуклеотидов.

У PuMWB11, PuMWB13, PuMWB30 найдены гены синтеза 5-(α -генциобиозил)оксиметилцитозина, у PuMWB40 – 5-(арабинозил)оксиметилцитозина.

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 36).

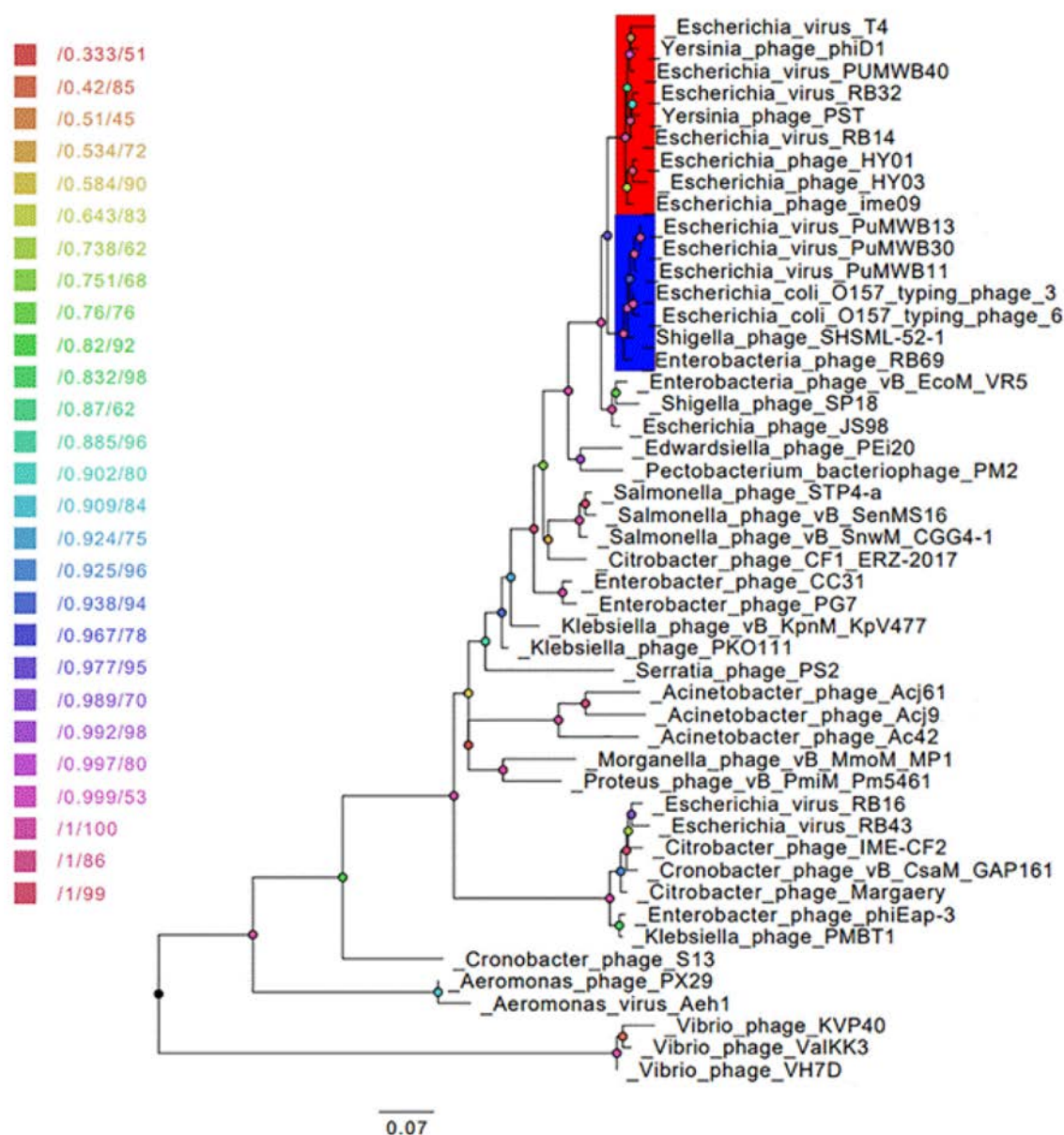


Рисунок 36 – Пангеномное дерево геномов классифицированных представителей T4-родственных бактериофагов, взятых из базы данных RefSeq, и геномов PuMWB11, PuMWB13, PuMWB30, PuMWB40. Красным и синим отмечены клады *Tequatrovirus* и *Mosigvirus* соответственно. Цветными кружками отмечена поддержка (значения для aBayes support и ultrafast bootstrap support представлены слева).

Анализ пангеномного дерева на рисунке 36 показал, что PuMWB11, PuMWB13, PuMWB30 принадлежат к роду *Mosigvirus*, PuMWB40 – *Tequatrovirus*.

Учитывая принадлежность PuMWB11, PuMWB13, PuMWB30 к одному роду, было проведено сравнение нуклеотидных последовательностей их геномов в BLASTN [146]. Оказалось, что PuMWB11 и PuMWB13 имели идентичность 99,99%, PuMWB11 и PuMWB13 имели с PuMWB30 идентичность 99,95%. При этом покрытие составляло 100% во всех случаях. По признаку идентичности, данные вирусы можно отнести к одному виду, [147]. При этом PuMWB11, PuMWB13 вероятнее всего являются разными изолятами одного фага.

На основе полученных данных о родстве PuMWB11, PuMWB13, PuMWB30, PuMWB40 можно судить о том, что проведенный ранее электрофорез фаговых коррелирует с результатами анализа выбранных геномов, 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 отмечено, 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, и выявления влияния на эволюцию данной группы вирусов этих генов.

3.2. Анализ геномов T4-родственных бактериофагов из баз данных

3.2.1. Кластеризация гомологичных последовательностей из геномов

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. В приложении 3 отмечены названия геномов

этих вирусов; их номера доступа; виды/штаммы бактерий, которые способны инфицировать эти вирусы; источники выделения бактериофагов. Первоначальная кластеризация аминокислотных последовательностей показала, что в 79 аннотациях геномов отсутствовали некоторые гены коргенома, что было связано с отсутствием этих генов или со сдвигом рамки считывания. После повторной кластеризации с реаннотированными геномами было обнаружено, что корбелок – коактиватор поздней транскрипции (gp33 у фага T4), фагов *Vibrio*, принадлежащих к роду *Schizotequatrovirus*, не образует общего гомологичного кластера с коактиваторами других *Tevenvirinae*. Поскольку эти кластеры белков имели низкую гомологию друг с другом, гомологи коактиватора поздней транскрипции вибриофагов как коргены не учитывались при отборе геномов. Аналогичная ситуация была с генами белков фибрилл хвоста (gp35, gp36 и gp37). Эти гены обладают высоким мозаицизмом и часто кластеризуются с негомологичными генами фибрилл других фагов [6]. При второй кластеризации также не использовались геномы *Aeromonas* phage phiAS5, *Shigella* phage SHFML-26, *Shigella* phage Sf25, *Vibrio* phage nt-1 и *Vibrio* phage VH7D из-за сдвига рамок считывания в ряде коргенов. Кроме того, при исследовании софткора пангенома мы обнаружили, что в геноме *Aeromonas* phage Ah1 наблюдается сдвиг рамки считывания в гене белка сборки головки (gp40 у фага T4); поэтому этот фаг также не использовался для анализа. В результате использовалось 199 отобранных геномов для дальнейшего анализа.

Суммарно было использовано 57499 аминокислотных последовательностей, получено 4530 гомологичных кластеров: в кор вошел 61 кластер, в софткор – 76 (с учетом коргеномных кластеров), в оболочку – 1966, в облако – 2488 кластеров; 79 геномов были реаннотированы, 7 из них удалены из дальнейшего анализа; всего 199 геномов были использованы для дальнейшего анализа. В приложении 4 представлено число компонентов пангенома в каждом геноме.

В таблице 11 отражены гены, чьи гомологи образуют кластеры софткоргенома, а также их белки или функции.

Таблица 11. Компоненты софтверного генома *Tevenvirinae*. Названия генов взяты из аннотации генома T4. Адаптировано из [6, 148]

Название гена	Белок или функция
<i>rIIA</i>	защита от раннего лизиса, индуцированного профагами
39	большая субъединица ДНК топоизомеразы II
<i>dexA</i>	ДНК экзонуклеаза
<i>dda</i>	ДНК-зависимая АТФаза, ДНК-хеликаза
61	ДНК праймаза
44	большая субъединица загрузчика зажима
41	репликативная ДНК-хеликаза
40	ассоциированный с мембраной инициатор сборки вершин головки
<i>UvsX</i>	RecA-подобный белок рекомбинации
43	ДНК-полимераза реплисомы
<i>regA</i>	репрессор трансляции RegA
62	малая субъединица загрузчика зажима
45	скользящий зажим
45.2	гипотетический белок
46	субъединица рекомбинантного комплекса, необходимая для инициации репликации ДНК
47	субъединица рекомбинантного комплекса, необходимая для инициации репликации ДНК
<i>a-gl.4</i>	гипотетический белок
55	сигма фактор поздней транскрипции
<i>nrdH</i>	субъединица комплекса анаэробной рибонуклеотидредуктазы
<i>nrdC.11</i>	гипотетический белок
<i>tk</i>	тимидин киназа
<i>tk.4</i>	гипотетический белок
<i>vs.1</i>	гипотетический белок
57B	гипотетический белок
1	дезоксинуклеозидмонофосфаткиназа
3	терминатор оболочки
2	защищает концы упакованной ДНК от нуклеазы RecBCD
4	белок завершения головки
53	клиновой компонент базальной пластинки
5	компонент хаба лизоцима базальной пластинки
6	клиновой компонент базальной пластинки
7	клиновой компонент базальной пластинки
8	клиновой компонент базальной пластинки
9	коннектор хвостовых фибрилл и триггер для сокращения хвостовой оболочки
10	субъединица шипа базальной пластинки
11	субъединица шипа базальной пластинки
12	коротки фибриллы хвоста
<i>wac</i>	фибрин
13	белок шеи
14	белок шеи
15	стабилизация и завершение оболочки хвоста
16	малая субъединица терминазы
17	большая субъединица терминазы
18	белок внешней сократительной оболочки хвоста
19	белок внутренней трубки хвоста
20	белок портала
68	кор белок каркаса проголовки
21	кор белок каркаса проголовки и протеаза
22	белок каркаса головки

Таблица 11. Продолжение

Название гена	Белок или функция
23	главный белок капсида
24	белок вершины капсида
<i>inh</i>	ингибитор протеазы проголовки
<i>UvsW</i>	рекомбинационная ДНК-РНК-хеликаза
<i>UvsY</i>	хелперный белок <i>UvsX</i>
25	клиновой компонент базальной пластинки
26	субъединица хаба базальной пластинки
<i>pseT</i>	полинуклеотид киназа
48	вспомогательный белок хвостовой трубки
54	белок соединения хвостовой трубки базальной пластинки
30	АТФ-зависимая ДНК-лигаза
30.3	гипотетический белок
31	кошаперон GroES
<i>cd</i>	дезоксцитидилат дезаминаза
<i>rnlA</i>	РНК-лигаза А
<i>nrdB</i>	субъединица аэробного комплекса рибонуклеотидредуктазы
<i>nrdA</i>	субъединица аэробного комплекса рибонуклеотидредуктазы
<i>td</i>	хоминг-эндонуклеаза I-TevI
<i>frd</i>	дигидрофолат редуктаза
32	белок SSB
59	загрузчик хеликазы
33	вспомогательный белок поздней транскрипции
<i>dsbA</i>	двухцепочечный ДНК-связывающий белок
<i>rnh</i>	РНКазы Н
34	длинные фибриллы хвоста, проксимальная субъединица
52	средняя субъединица топоизомеразы II
<i>rIIB</i>	защита от раннего лизиса, индуцированного профагами

Было определено большее число генов софторгенома, чем в работе по исследованию кор- и квазикоргенома Т4-родственных фагов [6] (76 софторгенов в данной работе и 63 в [6]), что вероятно связано с тем, что была использована подгруппа Т4-родственных бактериофагов, а именно *Tevenvirinae*. Кроме того, при реаннотации ряда использованных геномов было найдено, что часть ОРС коргенов не были исходно аннотированы, что также могло влиять на число исходно определенных коргенов Т4-родственных бактериофагов.

3.2.2. Филогенетические, пангеномные и филогеномные деревья

Для выявления эволюционных взаимосвязей бактериофагов подсемейства *Tevenvirinae* были построены филогенетическое дерево гомологов главного белка капсида, пан-геномное и филогеномное деревья. Филогеномное дерево позволяет лучше отразить генетическое разнообразие наиболее близких по родству вирусов,

за счет использования полной нуклеотидной последовательности геномов, тогда как пан-геномное дерево позволяет лучше отразить генетическое разнообразие более удаленных по родству вирусов, что связано с использованием всех аминокислотных последовательностей вирусов. Выбор аминокислотных последовательностей главного белка головки для построения филогенетического дерева обусловлен тем, что данный белок является одним из основных филогенетических маркеров дцДНК вирусов, так как его аминокислотная последовательность является одной из самых консервативных для дцДНК вирусов, а также ген этого белка менее других генов вирусов подвержен горизонтальному переносу генов. Таким образом, филогенетическое дерево на основе последовательностей главного белка головки позволяет наиболее точно отразить вертикальный перенос генов между родственными вирусами.

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 считается [149], что 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 37).

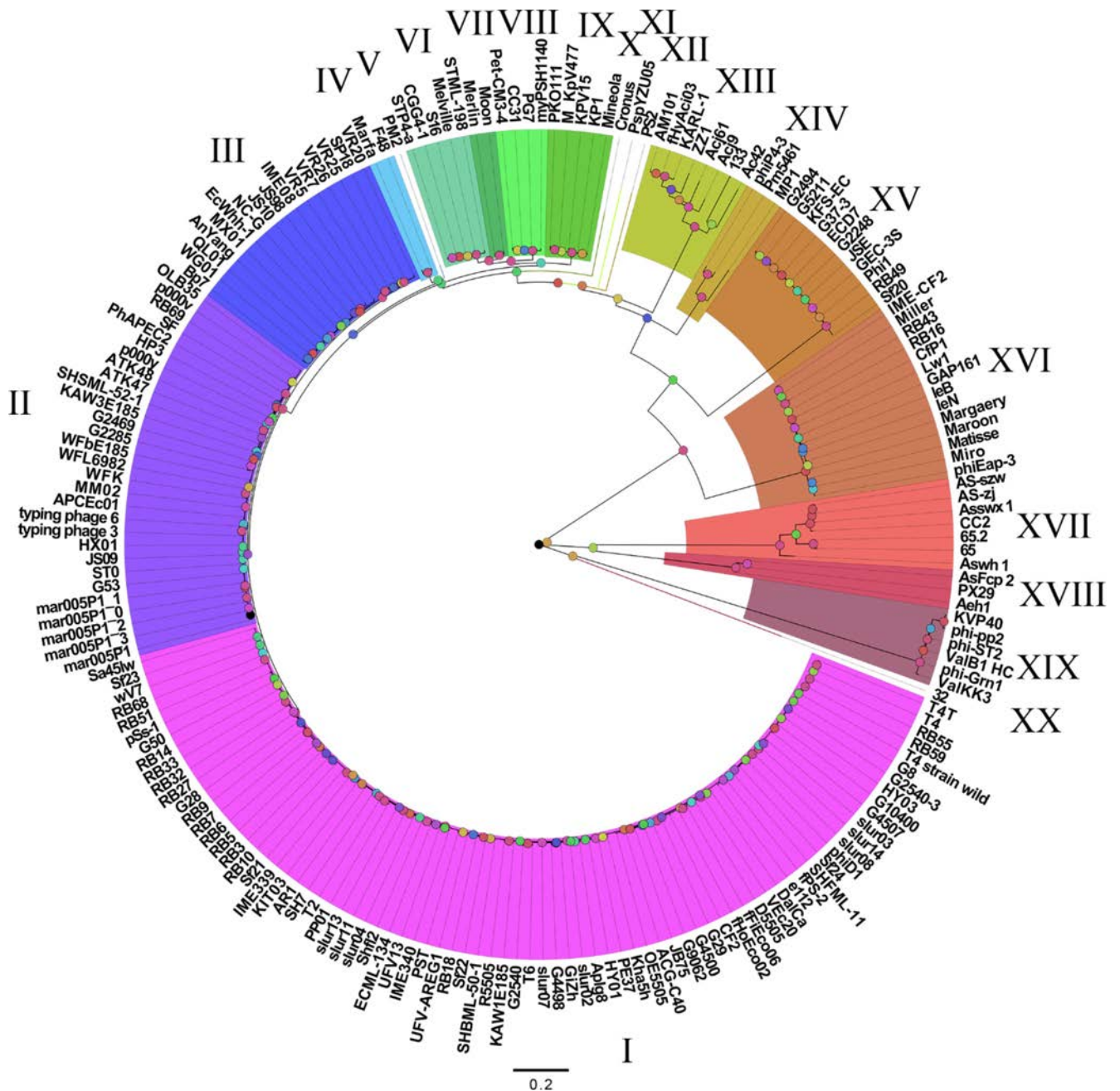


Рисунок 37 – пангеномное дерево геномов фагов подсемейства *Tevenvirinae* из баз данных NCBI Nucleotide Collection. Лучшая модель для построения: GTR2+FO+I+G4. Log-likelihood консенсусного дерева: $-41953,89217$. Цветом и римскими цифрами отмечены группы, цветными кружками – поддержка ветвей (значения представлены в приложении 5). Полные названия фагов представлены в приложении 4. Адаптировано из [148].

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 (группоспецифичные кластеры, которые представляют собой кластеры аминокислотных гомологичных последовательностей, присутствующие только у одного набора фагов).
- Принималась во внимание таксономическая классификация фагов.

В одну группу были отнесены фаги, принадлежащие к одному роду и образующие отдельную кладу с четко идентифицируемыми ГСК. Единственным исключением является группа 3, в которую входят представители двух родов фагов, *Dhakavirus* и *Gaprivervirus*. Это связано с тем, что клада, образованная представителями рода *Dhakavirus*, не имеет идентифицируемых ГСК.

Следует отметить, что, как показал анализ пан-геномного дерева и данные об источниках выделения фагов и чувствительных бактериях (представлены в приложении 3), 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 38 показано филогенетическое дерево, построенное на основе гомологичных последовательностей основного белка капсида *Tevenvirinae* – одного из коргенов и основного маркера для отдельных родственных групп дцДНК вирусов.

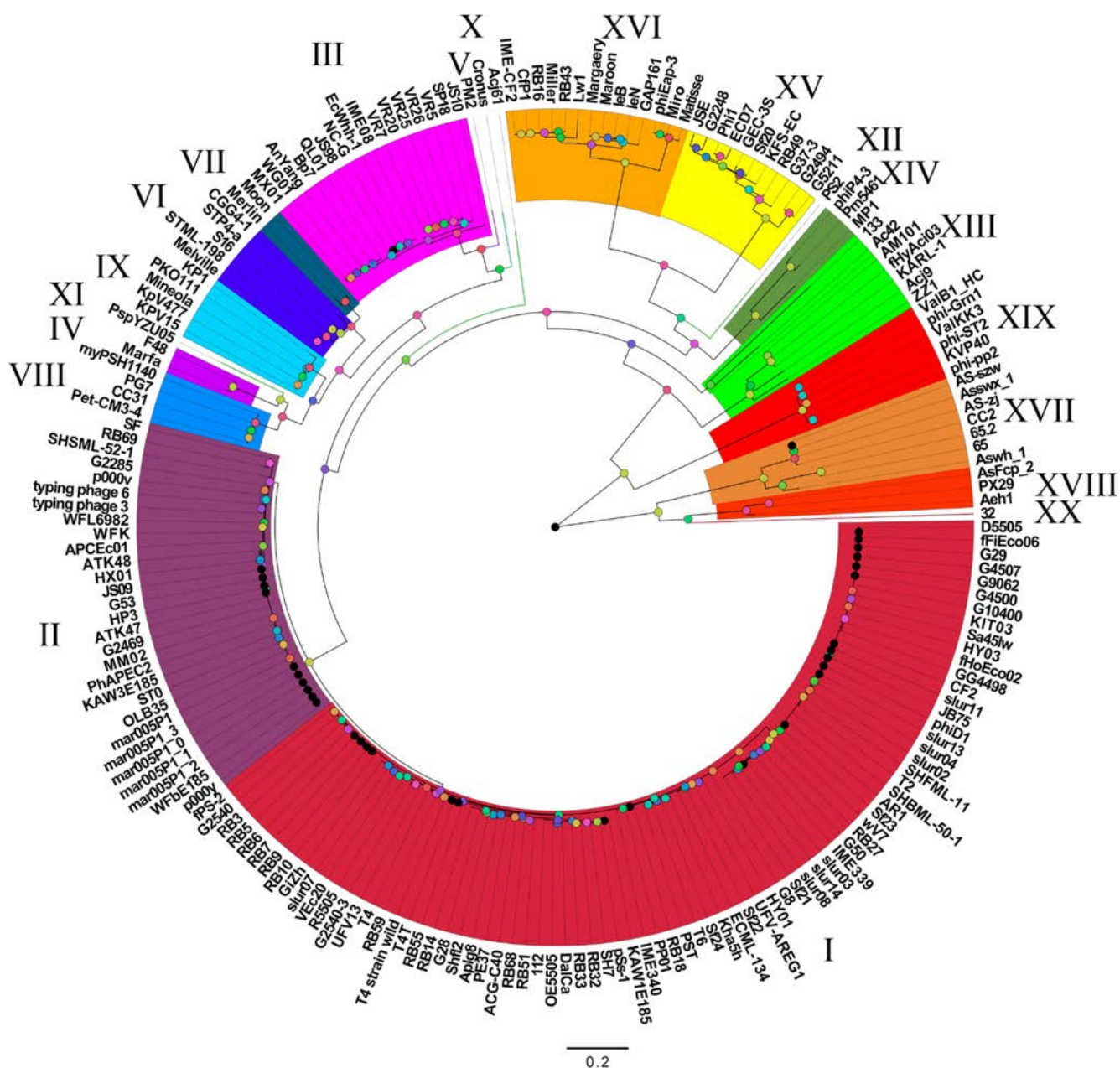


Рисунок 38 – филогенетическое дерево гомологов главного белка капсида фагов подсемейства *Tevenvirinae* из базы данных NCBI Nucleotide Collection. Лучшая модель для построения: LG+F+I+G4. Найден лучший результат: $-15512,735$. Цветом и римскими цифрами отмечены группы, цветными кружками – поддержка ветвей (значения представлены в приложении 5) Полные названия фагов представлены в приложении 4. Адаптировано из [148].

Было построено филогеномное дерево 199 геномов *Tevenvirinae*. Оно представлено на рисунке 39

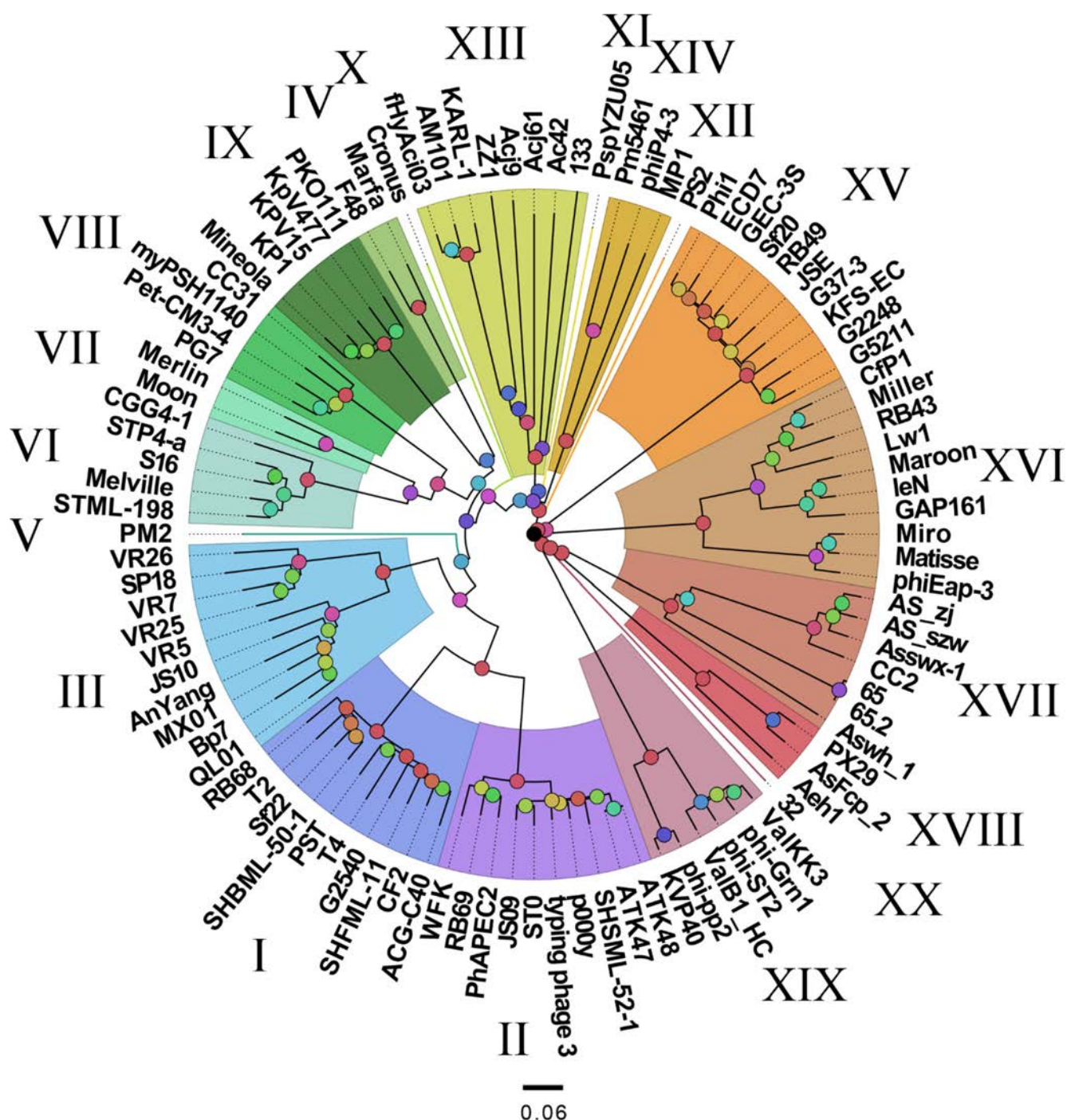


Рисунок 39 – филогеномное дерево геномов подсемейства *Tevenvirinae* из базы данных NCBI Nucleotide Collection. Цветом и римскими цифрами отмечены группы, цветными кружками – поддержка ветвей (значения представлены в приложении 5). Полные названия фагов представлены в приложении 4. Адаптировано из [148].

Сравнительный анализ пангеномного, филогеномного и филогенетического дерева главного белка капсида показал, что во всех случаях представители выделенных групп находились в одних и тех же кладах – за исключением группы

XIII, представитель которой (*Acinetobacter* phage Acj61) сформировал отдельную ветвь на филогенетическом дереве гомологов главного белка капсида. 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. Она представлена на рисунке 40.

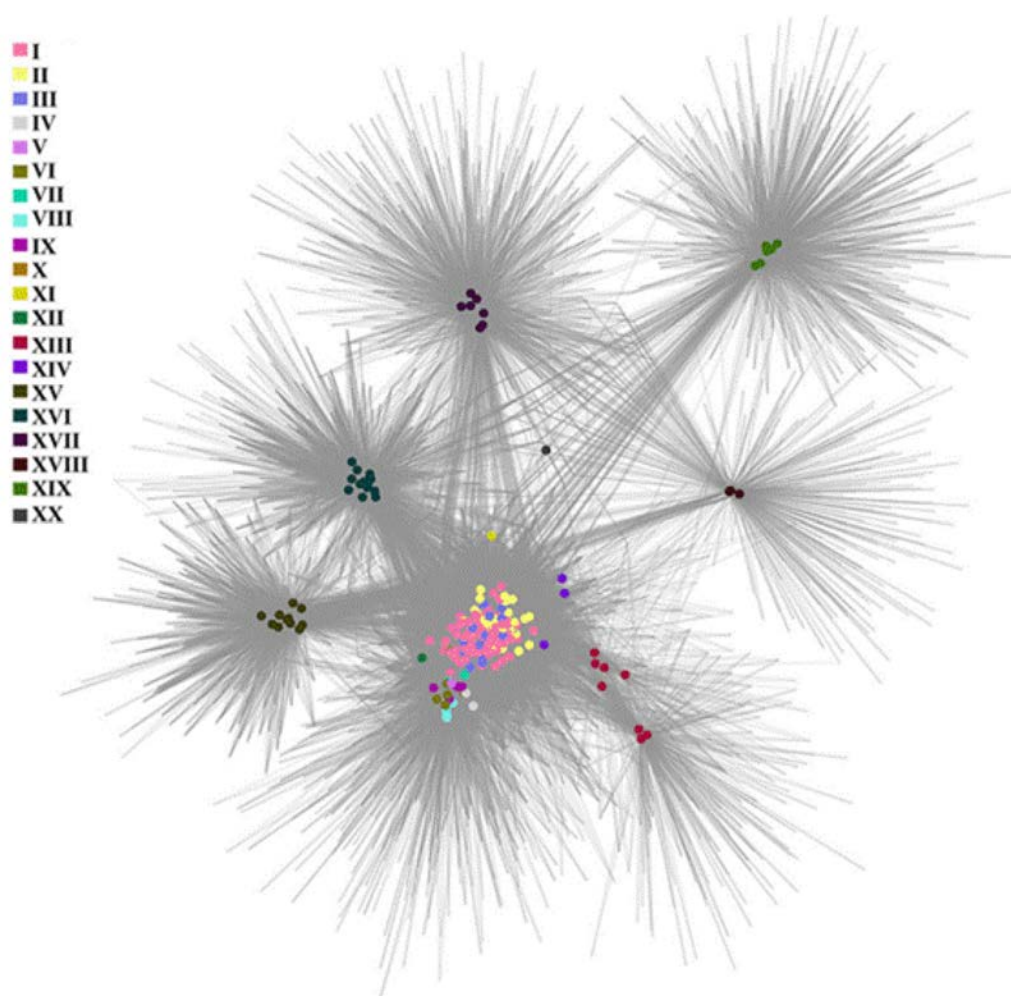


Рисунок 40 – двудольная сеть оболочки пангенома. Цветными кружками отмечены узлы геномов *Tevenvirinae* из базы данных NCBI Nucleotide Collection, цвет кружка отражает отношение к группе, ребра отражают связи между отдельными геномами и кластерами, узлы кластеров не отмечены. Адаптировано из [148].

Как отмечено выше, оболочка пангенома представляет наибольший интерес для изучения взаимодействия фагов с экологической нишей. Для анализа оболочки была построена двудольная сеть на основе кластеров этого компонента пангенома. В целом сеть совпадала с пангеномным деревом, но имелся ряд отличий. Например, группы I–III (что отмечено выше) на сети образуют общую плотную подсеть, что неудивительно, учитывая специфику экологических условий их местообитаний. В этом отношении стоит отметить геном *Vibrio phage vB_VmeM-32*, единственного использованного в данном анализе представителя группы XX. Он имеет наименьшее количество общих кластеров с другими геномами. Однако большинство этих кластеров представлены гомологами генов, которые обнаруживаются у большинства вирусов, представленных в приложении 3, на что указывает расположение фага в сети и его связи. Это может быть связано с особенностями его местообитания (приложение 3). Данный фаг был выделен из щелочного озера, из-за чего данный фаг мог находиться долгое время в изоляции от других представителей *Tevenvirinae*, которым свойственны другие места обитания (приложение 3), из-за чего не происходил генетический обмен между ними. В результате у данного фага сохранились лишь наиболее необходимые гены оболочки пангенома, тогда как остальные гены приспособления к условиям обитания он получил из пула генов в своей экологической нише.

Также было построено филогенетическое дерево гомологов РНК-лигазы 2 фага T4. Данный фермент включен в оболочку пангенома, в результате присутствует только у части геномов (188 из 199 геномов), однако у подавляющей части.

На рисунке 41 представлено дерево гомологов данного фермента.



Рисунок 41 – дерево гомологов РНК-лигазы 2 у *Tevenvirinae* из базы данных NCBI Nucleotide Collection. Цветными кругами отмечены рода *Tevenvirinae*. Номера доступов белков и названия родов (на 2019 год) представлены в приложении 5. Взято из [150].

Функцией данного фермента является репарация одностороннего разрыва между 5'-фосфорильной и 3'-гидроксильной группами соседних нуклеотидов в РНК дуплексах. Однако биологическая роль данного фермента у *Tevenvirinae* неизвестна. Известна она у гомологов данного фермента у трипаносом – участие в

редактировании мРНК кинетопластид [124]. 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*. Учитывая это, можно предположить, что роль данного фермента у рассматриваемого подсемейства вирусов 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 вирусов рода *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*). 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 подвергался горизонтальному переносу между фагами *Tevenvirinae*, занимающих схожие экологические ниши.

Чтобы получить больше информации об особенностях идентифицированных группах фагов *Tevenvirinae*, были проанализировали кластеры гомологов, специфических для каждой группы (ГСК) и проверили наличие их у других организмов. В результате поиска были выявлены ГСК, не имеющие гомологов среди белков, депонированных в базу данных nr (NCBI); ГСК, имеющие гомологи среди гипотетических белков других организмов; и ГСК, которые имели гомологи среди белков других организмов с предсказанной функцией. Число кластеров, выделенных по этим критериям, приведены в таблице 12.

Таблица 12. Количество группоспецифических кластеров (ГСК), выделенных групп *Tevenvirinae*.

Группа	Общее число ГСК	Количество ГСК с гомологами среди гипотетических белков	Количество ГСК с гомологами белков, функции которых известны
I	2	1	1
II	4	2	0
III	1	0	0
IV	27	7	1
V	43	14	2
VI	12	1	0
VII	22	5	0
VIII	28	23	1
IX	23	23	0
X	134	7	3
XI	75	8	6
XII	128	15	4
XIII	7	2	1
XIV	20	14	4
XV	99	25	5
XVI	69	13	4
XVII	104	27	4
XVIII	146	10	8
XIX	219	39	32
XX	137	31	19

Во втором столбце отмечено общее найденное число ГСК для каждой группы; в третьем – число ГСК, имеющих гомологи среди гипотетических белков, депонированных в базу данных nr (NCBI), других групп вирусов и организмов, в четвертом – число ГСК, имеющих гомологи среди белков других групп вирусов и организмов, депонированных в базу данных nr (NCBI) и имеющих известные функции.

3.2.3. Анализ группоспецифических кластеров

Группа XX включала только один фаг в данном анализе – *Vibrio phage vB_VmeM-32*. В основном ГСК этого фага имели гомологов у бактериофагов, инфицирующих бактерии *Vibrio* и *Vibrionaceae* и фагах *Caudovirales*, выделенных из морских источников. Гомологические последовательности этих ГСК в основном кодировали белки, предположительно связанные с фаговой активностью (экспрессия фаговых генов, морфогенез вириона и т.п.).

Группа XIX содержала представителей рода *Schizotequatrovirus*. Наиболее схожие гомологи ГСК в основном были обнаружены у фагов, инфицирующих

бактерии, выделенные из морских источников. Следует отметить, что ГСК этой группы содержали набор белков, предположительно участвующих в биосинтезе модификаций 7-дезазагуанина [4]. Имелись также гомологи белков, участвующих в биосинтезе нуклеотидов [151], в морфогенезе фаговых частиц и в других процессах.

В группу XVIII входили фаги, инфицирующие *Aeromonas*. Наиболее схожие гомологи ГСК были обнаружены в основном у фагов *Caudovirales*, инфицирующих гаммапротеобактерии, и бактерий из морских метагеномов и микрофлоры человека. ГСК этой группы были гомологичны таким белкам, как АДФ-рибозилтрансфераза, АДФ-рибозилгликогидролаза, белок CreA, белок утилизации гема, N(6)-аденин-специфичная ДНК-метилтрансфераза, УДФ-3-О-(3-гидроксимиристоил)глюкозамин-N-ацилтрансфераза и другие.

Группа XVII также была представлена фагами, инфицирующими *Aeromonas*. Гомологи ГСК этой группы были обнаружены в основном в геномах различных фагов, инфицирующих *Aeromonas*, и бактерий из морских источников и почв. ГСК имели гомологи среди таких предполагаемых белков, как шарнирный коннектор дистальных длинных хвостовых фибрилл, ингибитор протеаз хозяина, аденилатциклаза и белок суперсемейства РНКаз Н.

В группу XVI входили фаги, инфицирующие *Enterobacteriales*; одним из представителей этой группы является *Enterobacteria* phage RB43 [152]. Гомологи этих ГСК были обнаружены у фагов рода *Caudovirales*, инфицирующих *Aeromonas*, *Bacillus*, *Enterobacteriales*, *Clostridioides difficile* и бактерий, выделенных из фекалий, сточных вод и почвы, а также у *Bacillales* и *Enterobacteriales* из неизвестных источников. Кластеры содержали гомологи таких белков, как ДНК-полимераза, SAM, цитозин-специфичная метилтрансфераза и хаб базальной пластинки.

Группа XV содержала род *Krischvirus*. Наиболее схожими гомологами этих ГСК оказались белки фагов *Myoviridae* (преимущественно *Tunavirinae* и *Vequintavirinae*), инфицирующих *Enterobacteriales* (преимущественно *Salmonella enterica*, выделенных из различных источников).

Группа XIV включала фаги, инфицирующие *Proteus* и *Morganella*. Большинство гомологов было обнаружено у вирусов *Caudovirales*, инфицирующих *Enterobacteriales*, *Aeromonas*, *Stenothrophomonas*, *Vibrio* и *Cellulomonas oligotrophica*. В ГСК этой группы были обнаружены гомологи нуклеозидтрифосфатпирофосфогидролазы, РПВ, АДФ-рибозо-связывающего белка и HNH-эндонуклеазы.

Группа XIII включала фаги, инфицирующие *Acinetobacter*. Ближайшие гомологи ГСК этой группы были обнаружены среди белков фагов, выделенных из *Acinetobacter* и морских источников. Один из гомологов был идентифицирован как АДФ-рибозопирофосфатаза.

Группа XII в данном исследовании содержала один геном – *Serratia phage PS2*. Наиболее схожие гомологи были обнаружены в основном среди белков фагов *Myoviridae*, инфицирующих энтеробактерии, и белков самих энтеробактерий: эндонуклеаза HNH, дистальная субъединица длинных фибрилл, катализатор сборки дистальных длинных фибрилл, t-холин.

Группа XI состояла в данном исследовании из одного генома – *Pseudomonas phage PspYzu05*. Большинство ГСК этой группы имели схожие гомологи среди метагеномов, выделенных из морских источников. Имелись также гомологи белков фагов *Caudovirales*, инфицирующих *Vibrio*, *Aeromonas* и *Enterobacteriales*, и бактерий микрофлоры человека. Гомологи включали ряд предсказанных белков: нуклеозидтрифосфатпирофосфогидролазу, LtrC-подобный белок, белок фибрилл и ABC-транспортёрную АТФазу.

Группа X в данном анализе состояла из одного генома – *Erwinia phage Cronus*. Наиболее схожие гомологи были обнаружены в основном у фагов *Caudovirales*, выделенных из энтеробактерий (в случае некоторых ГСК, из *Erwinia*) и у самих энтеробактерий. Гомологи включали EPS-деполимеразу, SSB и ингибитор протеазы Lon.

Группа IX содержала фаги *Jiaodavirus*. Большинство кластеров имели гомологи среди белков *Klebsiella*; несколько кластеров среди белков *Winklervirus* и *Kolesnikvirus*, поражающих энтеробактерии.

Группа VIII была представлена фагами *Karamvirus*. Наиболее схожие гомологи ГСК были обнаружены в основном у фагов *Winklervirus*, инфицирующих энтеробактерии. Среди гомологов были белки Нос.

Группа VII содержала фаги *Moonvirus*. Наиболее схожие гомологи обнаружены среди фагов *Vequintavirinae* и *Tunavirinae*, выделенных из энтеробактерий, и среди бактерий *Dickeya*.

Группа VI состояла из фагов *Gelderlandvirus*. Только у одного кластера из этой группы обнаружены гомологи среди предполагаемых белков вирусов *Tegunavirus*, *Winklervirus* и *Sugarlandvirus*, поражающих энтеробактерии.

Группа V содержала в данном исследовании один геном – *Pectobacterium phage PM2*. Наиболее схожие гомологи обнаружены среди белков фагов *Winklervirus*, выделенных из энтеробактерий. ГСК имели гомологи среди белков биосинтеза молибдоптерин-гуанидинуклеотида MobD и белков завершения головки.

Группа IV состояла из двух фагов, инфицирующих *Klebsiella*. Наиболее схожие гомологи ГСК обнаружены у фагов *Winklervirus*, инфицирующих энтеробактерии, *Yersinia* и *Clostridioides difficile*. Гомологи включали I-спанин.

Группа III включала два рода *Tevenvirinae* (*Gaprivervirus* и *Dhakavirus*). Был обнаружен только один ГСК.

Группа II была представлена *Mosigvirus*. ГСК имели гомологи гипотетических белков *Salmonella enterica* subsp. *enterica*, выделенных из человека.

Группа I была представлена *Tequatrovirus*. Среди представителей бактерий *Clostridioides difficile* и *Salmonella enterica* были обнаружены гомологи двух ГСК. Один из кластеров состоял из гомологов тиоредоксина.

На основе наличия/отсутствия общих белковых гомологов было выделено двадцать групп фагов *Tevenvirinae*. Анализ имеющихся данных о представителях этих групп показал, что они были выделены из бактерий, занимающих сходные экологические ниши. Также эти бактерии принадлежали к гаммапротеобактериям. Кроме того, большинство представителей одной группы были выделены из схожих экологических ниш.

Чтобы более подробно рассмотреть обмен генетической информацией внутри *Tevenvirinae*, мы проанализировали связи между кластерами, формирующими оболочку пангенома, а также проанализировали ГСК. Отмечено, что 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.

Для характеристики геномов идентифицированных групп *Tevenvirinae* был проведен поиск гомологов их ГСК среди организмов, аминокислотные последовательности которых депонированы в базе данных nr. Как оказалось, большинство ГСК не имеют гомологов среди других организмов, за исключением групп I, II, VIII и XIV. Не считая групп, представленных в данном исследовании одним фагом (V, X, XI, XII, XX), наибольшее количество ГСК обнаружено у фагов групп XV–XIX. Большое количество ГСК в группах, представленных в данном анализе одним вирусным геномом, объясняется тем, что они включали по мимо ГСК белки, уникальные для этого единственного генома. 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*. Это может указывать на то, что предки этих организмов имели общий пул генов.

Таким образом, сравнение результатов анализа пангеномного дерева, филогеномного дерева, филогенетических деревьев, 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 и их продукты, способствующие воспроизведению фагов с неканоническими основаниями ДНК.

3.2.4. Неканонические основания и гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК

Как отмечалось выше, у бактериофагов группы XIX имеются ГСК, участвующие в синтезе и модификации 7-деазагуанина. Экспериментально подтверждено, что некоторые фаги *Tevenvirinae* могут иметь и другие неканонические основания: глюкозилированный ⁵hm dC (Т-четные бактериофаги из группы I); арабинозилированный ⁵hm dC (RB69 из группы II). Соответственно, был проведен поиск гомологов *Tevenvirinae*, которые по экспериментальным данным или *in silico* [4, 6, 8, 153], участвуют в синтезе таких оснований (или связаны с ними каким-либо иным образом). Их список представлен в таблице 13.

Таблица 13. Гены и их белки, способствующие воспроизведению фагов *Tevenvirinae* с неканоническими основаниями ДНК. Адаптировано из [148]

Ген/локус (фаг)	Белок/возможный белок (домен)	Функция/возможная функция	Основание
42 (Enterobacteria phage T4)	НМС-трансфераза	синтез ^{5hm} dCMP	модификации ^{5hm} C
α -gt (Enterobacteria phage T4)	α -глюкозилтрансфераза	α -глюкозилирование ^{5hm} dCMP в ДНК	
β -gt (Enterobacteria phage T4)	β -глюкозилтрансфераза	β -глюкозилирование ^{5hm} dCMP в ДНК	
$\beta\alpha$ -gt (Enterobacteria phage T2)	β -глюкозил-1,6- α -глюкозилтрансфераза	β -глюкозилирование α -глюкозилированного ^{5hm} dCMP в ДНК	
RB69p003c (Enterobacteria phage RB69)	арабинозилтрансфераза	арабинозилирование ^{5hm} dCMP в ДНК	
Acj61p076 (Acinetobacter phage Acj61)	предположительно гликозилтрансфераза	модификация неканонического основания	
RB69p052 (Enterobacteria phage RB69)	гомологи среди аминокликозид-3'-фосфотрансфераз	вероятно, участвует в синтезе УДФ-арабинозы	
RB69p055 (Enterobacteria phage RB69)	арабинозил-5-фосфатизомераза	вероятно, участвует в синтезе УДФ-арабинозы	
56 (Enterobacteria phage T4)	дЦТФаза-дУТФаза	увеличивает пул dCMP, уменьшает пул dCTP; обеспечивает dUMP для синтеза dTMP	
denA (Enterobacteria phage T4)	эндонуклеаза II	расщепление немодифицированной (dCMP-содержащей) ДНК	
denB (Enterobacteria phage T4)	эндонуклеаза IV	расщепление немодифицированной (dCMP-содержащей) ДНК	
alc (Enterobacteria phage T4)	белок Alc	блокирует транскрипцию немодифицированной (dCMP-содержащей) ДНК	
arn (Enterobacteria phage T4)	белок Arn	ингибирует действие RM-системы McrBC	
ipI (Enterobacteria phage T4)	внутренний белок (internal protein) I	ингибирует активность эндонуклеазы IV типа gmrSD	
RB69p045 (Enterobacteria phage RB69)	гомологи среди тимидилаткиназ	предположительно участвуют в синтезе арабинозил- ^{5hm} dCMP.	

В первом столбце в скобках отмечены фаги, названия генов или локусов которых представлены в столбце один.

Таблица 13. Продолжение

Ген/локус (фаг)	Белок/возможный белок (домен)	Функция/возможная функция	Основание
<i>RB69p047 (Enterobacteria phage RB69)</i>	n/a	предположительно участвуют в синтезе арабинозил- ⁵ hmCMP.	модификации ⁵ hmC
<i>RB69p049 (Enterobacteria phage RB69)</i>	n/a		
<i>RB69p050 (Enterobacteria phage RB69)</i>	гомологи среди пептидаз		
<i>RB69p051 (Enterobacteria phage RB69)</i>	присутствует фосфатазный домен полинуклеотидкиназы		
<i>PX29p085 (Aeromonas phage PX29)</i>	N-ацетилтрансфераза семейства GNAT	предположительно, модифицирует ДНК	неизвестное основание
<i>Aeh1OPC087c (Aeromonas phage Aeh1)</i>	глюкозилтрансфераза	предположительно, модифицирует ДНК	неизвестное основание
<i>folE (Vibrio phage KVP40)</i>	белок FolE	необходимы для синтеза 7-циано-7-дезагуанина (PreQ0)	модификации 7-дезагуанина
<i>KVP40.0120 (Vibrio phage KVP40)</i>	белок QueD		
<i>KVP40.0124 (Vibrio phage KVP40)</i>	белок QueC		
<i>KVP40.0284 (Vibrio phage KVP40)</i>	белок QueE		
<i>KVP40.0122 (Vibrio phage KVP40)</i>	белок DpdA2	необходим для вставки PreQ0 и/или PreQ1 в ДНК	
<i>KVP40.0123 (Vibrio phage KVP40)</i>	белок QueF или QueF-L	синтез дезоксиархеозина или 2'-дезоксигуанина или 2'-дезоксигуанина-7-дезагуанина	

Было рассмотрено наличие гомологов генов, представленных в таблице 13, у исследуемых фагов (таблица всех гомологов представлена в дополнительных материалах в [148], здесь не приводится из-за слишком больших размеров таблицы). Обнаружено, что близкородственные бактериофаги содержат одинаковый набор генов, способствующих воспроизведению фагов *Tevenvirinae* с неканоническими основаниями ДНК.

В [8] показали, что большинство генов, продукты которых участвуют или предположительно участвуют в арабинозилировании $^5\text{hm}d\text{C}$ 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 42 показаны вариации этой области, выявленные у близкородственных фагов *Tevenvirinae*.

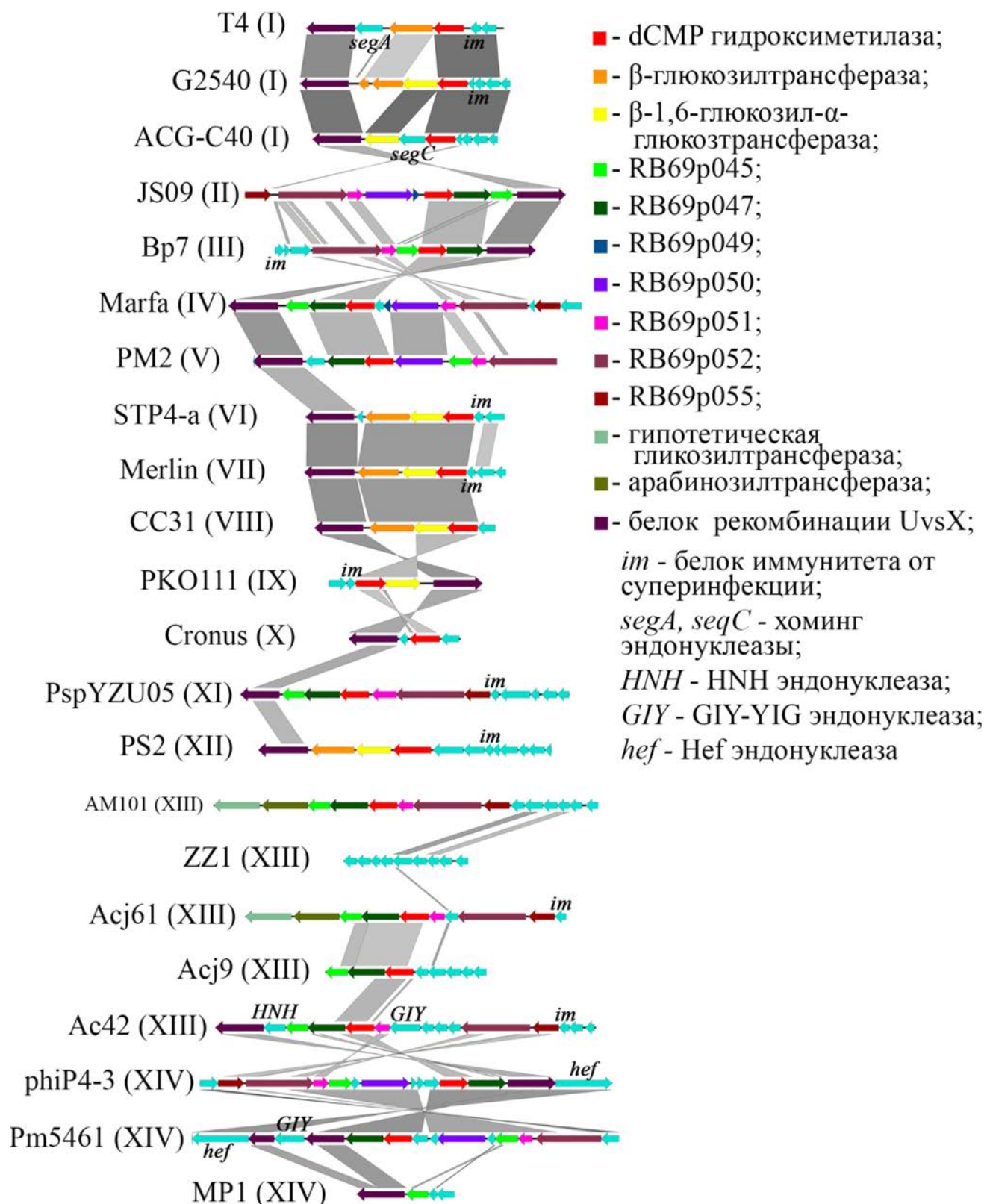


Рисунок 42 – Вариации области между генами ДНК-полимеразы и шаперона сборки верхушек капсида у близкородственных *Tevenvirinae* из базы данных NCBI Nucleotide Collection, имеющих гомологи белков синтеза $^5\text{hm dC}$.

Следует отметить, что у *Acinetobacter* phage Acj9 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 Acj61 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. Ранее предполагалось участие хоминг-эндонуклеаз в горизонтальном переносе генов *Tevenvirinae* [6]. Поэтому можно предположить, что гены, участвующие в синтезе и модификации неканонических оснований, подвергались или подвергаются частому горизонтальному переносу.

Поиск гомологов белков, способствующих воспроизведению фагов с неканоническими основаниями ДНК, показал, что они присутствовали у всех представителей *Tevenvirinae*. Однако у фагов групп XV и XVI, а также у некоторых представителей групп III (*Gaprivervirus*), XIII (*Acinetobacter* phage ZZ1, *Acinetobacter* phage 133) и XIV (*Morganella* phage vB_MmoM_MP1) отсутствовали какие-либо известные ферменты, участвующие в синтезе таких оснований.

Фаги групп I–XIV (за исключением упомянутых выше представителей и *Morganella* phage vB_MmoM_MP1) содержали гомологи гена, кодирующего тимидилатсинтазу, которая участвует в синтезе $^{hm5}dCMP$. Фаги групп I, VI, VII, VIII, IX, XII дополнительно содержали гены, продукты которых участвуют в гликозилировании $^{hm5}dCMP$, хотя состав этих генов варьировал. Большинство фагов группы I имели α -глюкозилтрансферазу и β -глюкозил-1,6- α -глюкозотрансферазу (у бактериофага T4 и пяти близкородственных фагов вместо β -глюкозил-1,6- α -глюкозотрансферазы имелась β -глюкозилтрансфераза, а фаг *Escherichia* vB_EcoM_G2540 имел гены всех трех белков). Фаги группы IX содержали гены, кодирующие α -глюкозилтрансферазу и β -глюкозил-1,6- α -

глюкозотрансферазу – как и большинство фагов группы I. Фаги групп VI, VII, VIII, XII имели гомологи генов β -глюкозилтрансферазы и β -глюкозил-1,6- α -глюкозотрансферазы. Функция второго белка в этих вирусах до сих пор неясна, так как фермент не может дополнительно модифицировать основания по какому-либо известному механизму в отсутствии α -глюкозил-^{hm5}dC в ДНК. Следовательно, либо этот белок по-другому модифицирует неканоническое основание, либо его ген остается в геномах в виде атавизма.

Фаги группы II и XIII (за исключением *Acinetobacter* phage ZZ1, *Acinetobacter* phage 133 и *Acinetobacter* phage Ac42) имели гомологи белков, участвующие с арабинозилированием ^{5hm}dC в ДНК. Фаги групп IV, V, XI и частично группы XIV (*Proteus* phage phiP4-3, *Proteus* phage vB_PmiM_Pm5461) содержали белки, предположительно участвующие в этом процессе, но у них отсутствовали какие-либо известные арабинозилтрансферазы. Среди упомянутых белков два предположительно участвуют в синтезе УДФ-арабинозы, а один является предполагаемой тимидилаткиназой.

Бактериофаги групп XVII (за исключением *Aeromonas* phage AS-szw, *Aeromonas* phage AS-zj и *Aeromonas* phage Asswx 1), XVIII и XX имели гомологи двух генов, которые участвуют в модификации оснований. Гомолог одного из генов был общим для фагов группы XX и некоторых фагов группы XVIII (*Aeromonas* phage AsFcp 2 и *Aeromonas* phage PX29); а гомолог второго гена обнаруживался у представителей группы XVIII и характерен представителям группы XVII (*Aeromonas* phage 65.2, *Aeromonas* phage 65 и *Aeromonas* phage Aeh1).

Как было отмечено выше, фаги группы XIX имеют гомологи белков, участвующих в синтезе и модификации 7-диазагуанина. По экспериментальным данным из [4], у *Vibrio* phage nt-1 имеется небольшое количество этих неканонических оснований, замещающих гуанин (у фага *Vibrio* phage nt-1 гуанин лишь частично и редко замещается этой модификацией). Для остальных представителей группы XIX таких данных нет. Поскольку они отличаются от *Vibrio* phage nt-1 функцией гомологичных генов, которые участвуют в синтезе и

модификации 7-диазагуанина (имеют разные механизмы модификации 7-деазагуанина), процент замен также может различаться.

У некоторых фагов *Tevenvirinae* отсутствовали белки, участвующие в синтезе неканонических оснований, однако у них были гомологи генов, продукты которых зависят от этих оснований. А именно, это гомологи DenA, DenB, Alc и dCTPазы-dUTPазы. Их функции приведены в таблице 13. Как показано экспериментально, представители групп XV и XVI (а именно фаги RB43 и RB49) способны осуществлять общую трансдукцию в бактериях, содержащих каноническую ДНК [154]. В то же время эти фаги имеют гены, кодирующие гомологи DenA и dCTPазы-dUTPазы, чьи функции отрицательно влияют на возможность переноса и взаимодействия с канонической ДНК бактерий, но, видимо, не ограничивают эту возможность полностью. Активность белка Alc может быть ингибирована, как показано в [98], путем метилирования последовательностей CpG вблизи его локуса. Что касается DenB, то мы не нашли примеров, демонстрирующих ингибирование его эндонуклеазной активности. Кроме того, показано, что DenB играет существенную роль в гидролизе ДНК с каноническим составом оснований [97]. Таким образом, наличие генов гомологов DenB и Alc может свидетельствовать о наличии неканонических оснований в вирусной ДНК, тогда как наличие гомологов DenA и dCTPазы-dUTPазы указывает на это лишь частично – особенно у наиболее удаленных от фагов вирусов группы I. Стоит, однако, отметить наличие у представителей групп XV и XVI ГСК гомологов цитозинметилтрансфераз, что может противодействовать действию DenA.

У представителей III группы, образующих род *Gaprivervirus*, обнаружены гомологи всех четырех описанных выше белков. Не исключено, что у этих фагов есть неизвестные гены, участвующие в синтезе неканонических оснований. Такое же предположение можно сделать в отношении *Acinetobacter* phage ZZ1 и фага *Acinetobacter* phage 133 (имеющих гомологи DenB, Alc и dCTPазы-dUTPазы) и фаги *Aeromonas* AS-szw, AS-zj и Asswx 1 (эти фаги имеют гомологи DenA и Alc). Фаги групп XV и XVI, как было показано ранее, имеют только гомологи DenA и dCTP – dUTPase. Фаги этих двух групп не имеют каких-либо известных белков,

участвующие в синтезе неканонических оснований, однако анализ их ГСК выявил гомологи двух цитозин(C5)-ДНК-метилтрансфераз. Продукты этих ферментов потенциально могут ингибировать активность ферментов DenA и dСТРазы-dUTРазы. Дальнейшее изучение этих фагов может прояснить, как они защищают себя от действия ферментов, способствующих воспроизведению фагов с неканоническими основаниями ДНК

Сравнительный анализ наборов белков, ассоциированных с неканоническими основаниями, и взаимоотношений между фагами *Tevenvirinae* показал, что близкородственные фаги имеют ферменты, участвующих в синтезе одних и тех же или химически схожих неканонических оснований. Как известно, бактериофаг Т4 не способен к общей трансдукции [100]; кроме того, он исключается из инфекции бактериофагом RB69, имеющим другую модификацию ^{hm5}dC [32]. Также в [100] на мутантах фага Т4 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.

Также установлено, что гены синтеза и модификации ^{hm5}dC, а также гены синтеза и модификации 7-дезагуанина расположены близко друг к другу (большинство из них располагаются в одной области генома). Было предположено, что гены синтеза и модификации ^{5hm}dC у предков *Tevenvirinae* представляли собой геномные островки, способные переноситься горизонтально. На это также могут указывать частая встречаемость хоминг-эндонуклеаз в этом регионе и наличие генетических элементов, свойственных мобильным кассетам в этом регионе, что продемонстрировано в [59, 60].

Принимая во внимание, что неканонические основания и гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, являются одним из механизмов регуляции частоты ГПГ у *Tevenvirinae*, и на основании результатов, полученных в нашем исследовании, было предположено, что они играют одну из ключевых ролей в разнообразии геномов этих вирусов.

Еще одной важной функцией неканонических оснований является защита ДНК фагов от систем бактерий-хозяев, ограничивающих развитие фагов. Бактерии обладают RM-системами, которые в одних случаях расщепляют каноническую, а в других — неканоническую ДНК. Кроме того, у них есть Cas-системы, которые могут ингибироваться неканоническими основаниями, причем ингибирование зависит от состава оснований [3–5]. Известно, что у *Tevenvirinae* ^{hm5}dC и его модификации обладают разной степенью защиты от рестрикционной активности RM-систем, а также Cas-систем [3, 4, 91]. Также было показано, что фаги, содержащие 7-деазагуанин, обладают защитой от RM-систем [4]. Кроме того, представители групп I–III обладают белками IpI [155] и Arn [156], которые защищают их от действия RM-систем типа IV. В зависимости от наличия у предков *Tevenvirinae* тех или иных оснований (канонических или неканонических) в составе ДНК и наборов белков, защищающих от антивирусных систем хозяев, эти вирусы могли воспроизводиться в разном спектре бактерий-хозяев, в зависимости от наличия у бактерий определенных антивирусных систем. В результате создавался барьер для обмена генетической информации у близкородственных вирусов с разными основаниями в составе ДНК и наборов белков, защищающих от антивирусных систем хозяев.

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.

Использование такой стратегии обмена генетической информацией и является одной из причин распространения *Tevenvirinae* в самых разных экологических нишах. С одной стороны, это позволяет им приобретать гены, продукты которых необходимы для выживания в определенных условиях. С другой стороны, это позволяет им сохранять эти гены в своих геномах. Под давлением естественного отбора популяции, в которых ГПГ менее ограничен, имели неотъемлемое преимущество для размножения своего потомства в начале процесса отбора. Однако по мере того, как они становились все более приспособленными к конкретным условиям среды, популяции, в которых ГПГ ограничен, будут иметь преимущество в сохранении генов, наиболее критичных для выживания в той или иной экологической нише. В конце концов, популяции, которым удастся сохранить «лучший» набор генов, связанных с выживанием, будут распространяться внутри ниши, но в то же время станут менее способными к распространению в других

нишах. Другими словами, история эволюционного развития *Tevenvirinae* отражает их постепенную специализацию к специфическим средам, причем неканонические основания и ассоциированные с ними белки играют свою роль в процессе специализации.

Концептуальная схема стратегии представлена на рисунке 43

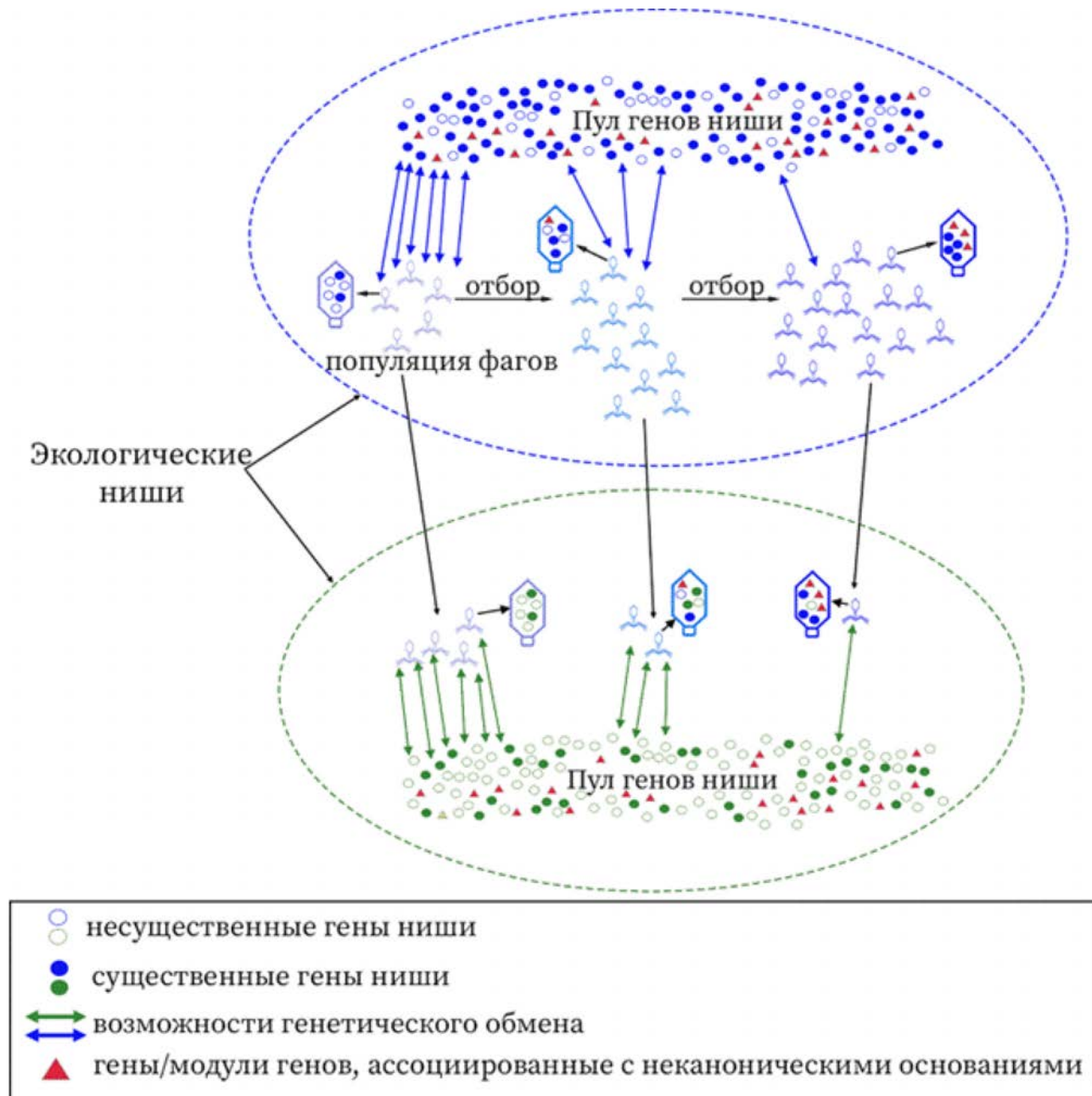


Рисунок 43 – концептуальная схема стратегии эволюции предков бактериофагов подсемейства *Tevenvirinae*. Популяция предков, имеющая наибольшие возможности к горизонтальному переносу генетической информации, занимая новую экологическую, начинает приобретать гены из генетического пула экологической ниши, среди которых имеются гены синтеза неканонических

оснований ДНК и гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК. В популяции появляются линии с различным набором генов из пула ниши, как существенных, так и не существенных для приспособления к данной экологической нише, среди которых имеются гены синтеза неканонических оснований ДНК и гены, способствующих воспроизведению фагов с неканоническими основаниями ДНК. Накопление генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, приводит к уменьшению возможностей горизонтального переноса генов и изменению спектра возможных бактерий-хозяев. В результате, в процессе отбора образуются популяции фагов, наиболее адаптированных к данной экологической нише, за счет наличия наибольшего числа генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, что позволяет лучше сохранять гены, существенные для развития популяции в данной экологической нише. Такие популяции специализируются на экспансии одной экологической ниши. Также образуются популяции, наиболее приспособленные к распространению по различным нишам, за счет отсутствия генов, уменьшающих возможности горизонтального переноса генов, в результате чего им проще приобретать и терять гены, необходимые для приспособления к нишам, однако им сложнее их сохранять. Такие популяции специализируются на распространении в различных экологических нишах. На рисунке: «отбор» – естественный отбор; «существенные гены ниши» – гены из генетического пула ниши, необходимые для приспособления к нише; «несущественные гены ниши» – гены из генетического пула ниши, которые не влияют на приспособление фагов к нише; «возможности генетического обмена» – число стрелок означает большие или меньшие возможности горизонтального переноса генов; «гены/модули генов, ассоциированные с неканоническими основаниями» – гены/модули генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК.

Стратегия предков *Tevenvirinae*, ставящая обмен генетической информацией в зависимость от набора неканонических оснований и генов, чьи функции ассоциированы с ними, была одним из основных факторов, способствующих

распространению этих фагов в окружающей среде, о чем свидетельствует их частое появление в метагеномах, выделенных из различных источников.

Стратегия эволюции предков *Tevenvirinae* была сопоставлена с тремя эколого-эволюционными моделями: «Красная Королева», «убить победителя», «выезжай на спине победителя». Модель «выезжай на спине победителя» в данном случае не может быть использована, так как она предполагает наличие у бактериофагов лизогенного цикла развития, тогда как *Tevenvirinae* являются строго литическими. Модель Красной Королевы предполагает, что коэволюция бактерий и бактериофагов происходит за счет контрадаптаций, в результате которой бактерии и их вирусы приобретают новые механизмы ингибирования систем защиты друг друга. Действительно, неканонические основания в ДНК и гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, у *Tevenvirinae* участвуют в защите от действия CRISPR-Cas и RM-систем, в результате чего у бактерий могли, например, появляться RM-системы IV типа. Однако, данные гены также способствуют адаптации к экологической нише, за счет уменьшения частоты ГПГ и увеличения возможности сохранения генов адаптации. Таким образом модель Красной Королевы не отражает полную картину данной стратегии. В отличие от модели Красной Королевы, модель «убить победителя» предполагает, что в экологической нише идет стабилизация популяций бактерий за счет наличия достаточного разнообразия популяций бактериофагов, не акцентируя внимания на контрадаптациях. В результате, для рассмотрения стратегии 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, приспосабливать 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 (пример контрадаптации), так и генами метаболизма (пример адаптации). Во втором случае, например, бактериофаг может нести необходимый ему для развития метаболический ген, который отсутствует у бактерии.

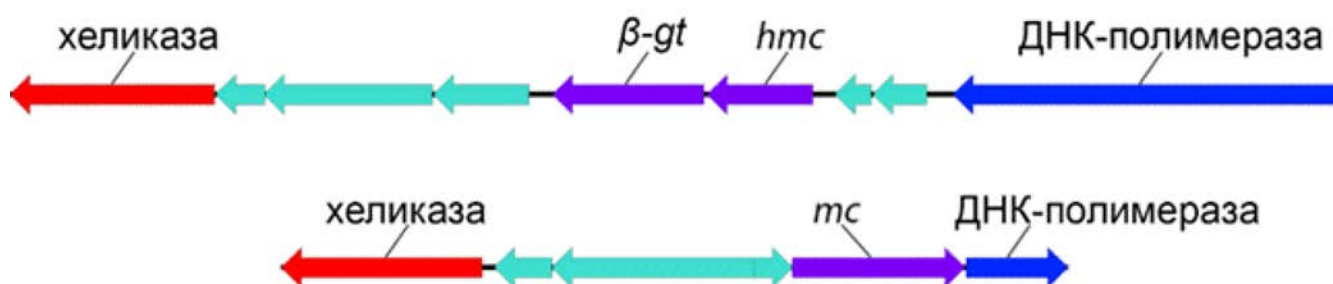
Бактериофаги из других таксономических групп, которые имеют неканонические основания ДНК, также могут использовать стратегию предков *Tevenvirinae*. Если в базах данных зарегистрировано достаточное количество фагов определенных таксонов, методы пангеномного анализа могут быть использованы для выявления эволюционных стратегий этих фагов. Для эффективного решения этой задачи также потребуются применение недавно разработанных методов секвенирования. Комплексное изучение метаболизма вирусов с неканоническими основаниями ДНК с использованием, наряду с другими подходами, методов сравнительной геномики будет способствовать лучшему пониманию эволюционных механизмов, лежащих в основе развития таких вирусов.

В настоящее время *Tevenvirinae* и более широкая группа Т4-родственных фагов Приблизено-Т4-родственные (Near-T4) и Отдаленно-Т4-родственные (Far-T4) представляют собой оптимальную модель для изучения влияния неканонических оснований, их модификаций и генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, на эволюцию и экологию вирусов.

3.2.5. 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 [157]. 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). Кроме того, 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. На рисунке 44 продемонстрированы данные регионы у фагов *Enterobacteria phage T4* и *Saccharomonospora phage PIS 136*.

Enterobacteria phage T4



Saccharomonospora phage PIS 136

Рисунок 44 – регион геномной ДНК между генами хеликазы и ДНК-полимеразы фагов Enterobacteria phage T4 (сверху) и Saccharomonospora phage PIS 136 (снизу). Красным цветом отмечена ОРС гена хеликазы, синим – ДНК-полимеразы, фиолетовым – ОРС генов, участвующих в синтезе и модификации оснований ДНК, голубым – другие гены; « β -gt» – ген β -глюкозилтрансферазы, «hmc» – ген гидроксиметилцитозинтрансферазы, «mc» – ген цитозин-специфичной метилтрансферазы.

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, учитывая, что также и у других вирусов гены синтеза и модификаций пиримидинов могут располагаться в схожих регионах [61]. 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, участвующих в репарации ДНК у T4-родственных бактериофагов – DenV, было предположено, что его действие связано

с неканоническими основаниями ДНК [158]. При рассмотрении его гомологов у других вирусов, была построена двудольная сеть, демонстрирующая взаимосвязи между ними (рисунок 45).

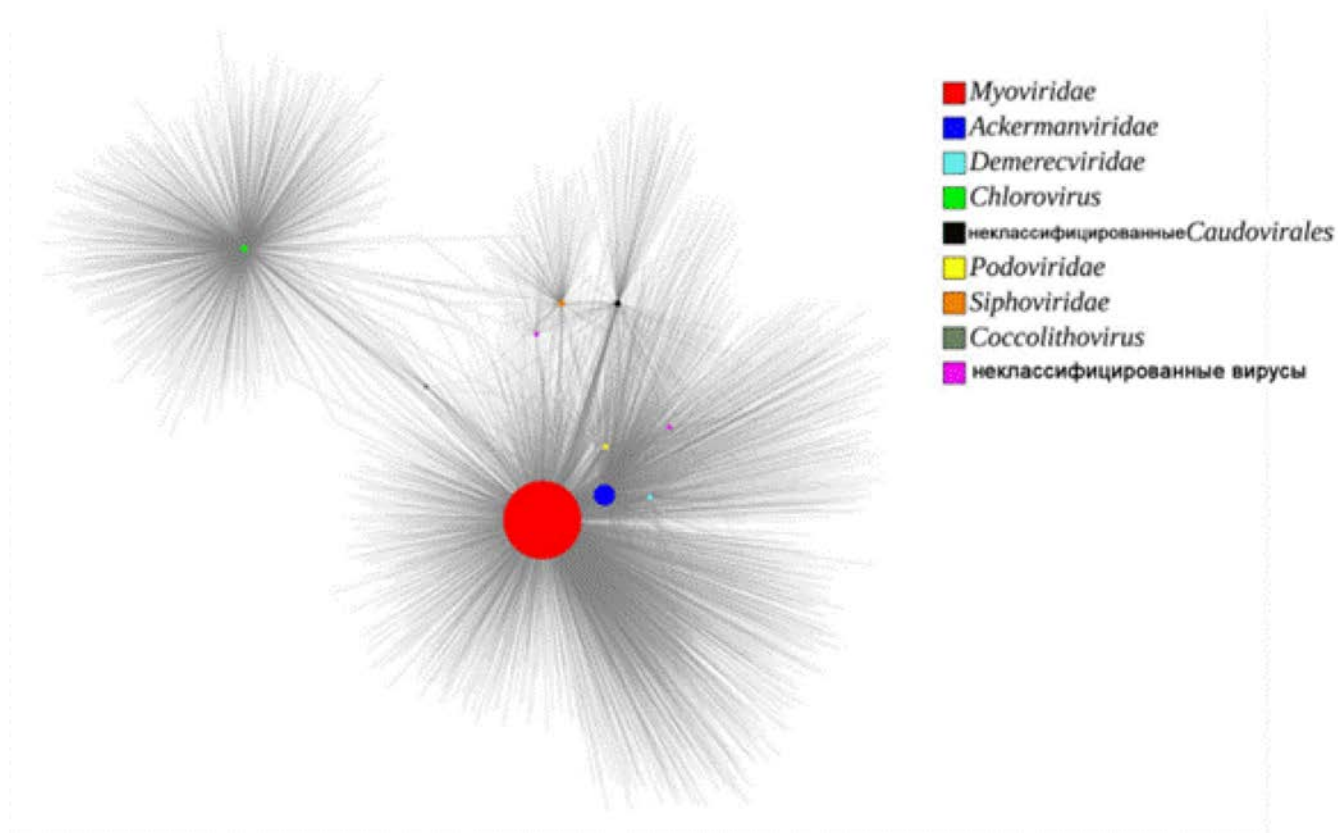


Рисунок 45 – двудольная сеть вирусов, обладающих гомологами DenV. Цветными кругами отмечены коллапсированные узлы геномов семейств/родов/неклассифицированных вирусов. Адаптировано из [158].

Сеть была построена на основе геномной кластеризации использованных *Tevenvirinae*, путем добавления остальных геномов вирусов, имеющих DenV. В результате, отмечалось наличие связей 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.

3.2.6. Связь неканонических оснований с классификацией T4-родственных бактериофагов

Классификация вирусов меняется и совершенствуется достаточно быстро, так же, как и принципы классификации. Как было отмечено в обзоре литературы на данный момент действуют принципы, представленные на рисунке 4. 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. На рисунке 46 представлена эта зависимость.

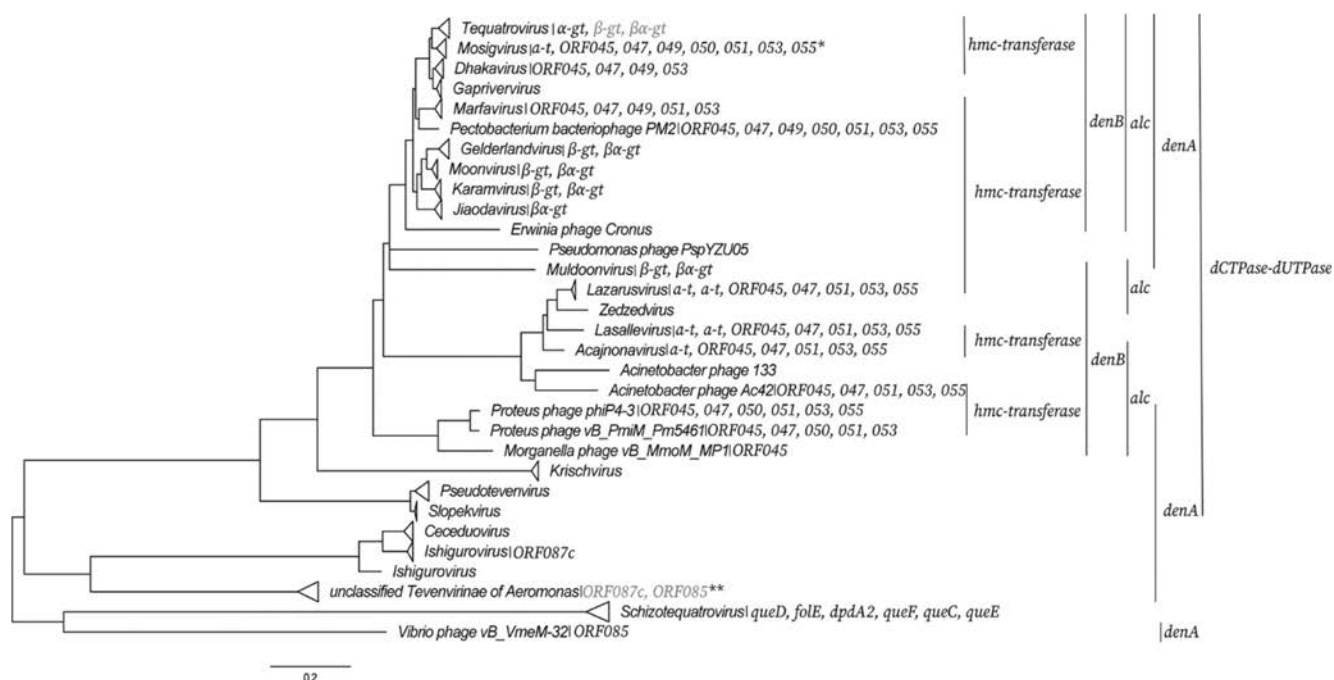


Рисунок 46 – пангеномное дерево исследованных геномов Т4-родственных бактериофагов из базы данных NCBI Nucleotide Collection, являющихся на 2019 год представителями *Tevenvirinae*. Выделены рода, известные на 2021 год. Отмечены гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК: α -gt – альфа-глюкозилтрансфераза; β -gt – β -глюкозилтрансфераза; $\beta\alpha$ -gt – β -глюкозил-1,6- α -глюкозилтрансфераза; a-t – арабинозилтрансфераза; ОРС* – гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК; ОРС** – гены, предположительно способствующие воспроизведению фагов с неканоническими основаниями ДНК. Серым отмечены гены, присутствующие у части определенного рода (группы).

Таким образом, гены, способствующие воспроизведению фагов с неканоническими основаниями ДНК, также могут быть использованы в качестве генетических маркеров родов Т4-родственных бактериофагов.

3.3. Перспективы использования Т4-родственных бактериофагов как агентов фаговой терапии

Для Т4-родственных бактериофагов известен широкий спектр исследований по использованию их в качестве агентов фаговой терапии, созданию препаратов на основе этих вирусов и испытаний этих препаратов на животных, клинических испытаний препаратов и клиническому использованию.

Известно, что ряд представителей Т4-родственных бактериофагов способны инфицировать патогенные штаммы бактерий. Примером таких вирусов могут служить фаги, инфицирующие патогенные штаммы *E.coli*. Enterobacteria phage AR1 является представителем рода *Tequatrovirus*. Он способен инфицировать представителей энтерогемморагической серогруппы *Escherichia coli* O157:H7, при этом не инфицируя ряд других *E.coli* [159]. Enterobacteria phage GEC-3S инфицирует *E. coli* O104:H4 – штамм, способный вызывать гемолитико-уремический синдром и кровавую диарею [160]. Enterobacteria phage vB_EcoM_ACG-C40 способен лизировать антибиотикорезистентные UPEC (уропатогенные *E. coli*) штаммы, образующие биопленки [161]. Ряд других представителей Т4-родственных бактериофагов способны к инфекции O157:H7 [159, 162–170], АРЕС (патогенные для птиц *Escherichia coli*) [171, 172] и другие патогенные *E.coli* [173–175]. Таким образом, представители Т4-подобных вирусов обладают широким спектром возможностей в инфицировании патогенных *E.coli* и других бактерий, относящихся к самым разным группам патогенности, что позволяет использовать их как агенты фаговой терапии.

На Т4-родственных бактериофагах также исследовались различные аспекты получения фаговых препаратов: получение биомассы фагов, очистка и концентрация биомассы, получение различных форм препаратов. Исследование получения биомассы производили в ферментере, в больших колбах Эрленмейера, одноразовом мешочном биореакторе (wave bag disposable system) на культуре *E.coli*. В результате удалось получить лизат с фагом в относительно больших объемах (2, 5, 8 л и более) и относительно высокой концентрации фаговых частиц (примерно 10^9 БОЕ/мл). Выяснилось, что основными параметрами, влияющие на концентрацию фаговых частиц при выращивании в относительно больших объемах являются фаза роста *E.coli* (наибольшие концентрации фагов получались при заражении на ранней экспоненциальной фазе), и множественность заражения (MOI) [9]. На Т4-родственных бактериофагах исследовались стандартные методы концентрирования и очистки лизатов: ультрацентрифугирование, центрифугирование при средней скорости, осаждение PEG, ультрафильтрация,

хроматография. Наименьшие потери при концентрировании были обнаружены при использовании ультрафильтрации, хроматографии [9]. Часто используемая форма для фаговой терапии является раствор очищенного фаголизата. Однако такая форма имеет существенный недостаток – снижение концентрации частиц при использовании раствора оральным путем при прохождении желудка. Одним из подходов для увеличения проходимости фагов в желудке, испытанным на Т4-родственных фагах, является использование разбавленного раствора в минеральной воде (буфер против желудочной кислотности) [176]. Другим, более перспективным подходом, исследованным на Т4-родственных бактериофагах, является инкапсуляция фагов [177]. Исследовались инкапсуляция в альгинат-полимеры методом экструзии [178–180], метакрилатный полимер [181] для введения препарата оральным путем. Показана стабильность данных способов инкапсулирования, стабильность получаемых препаратов в течение продолжительного времени, устойчивость к кислотным условиям, возможность высвобождения. Таким образом, имеются детальные исследования по различным этапам конструирования фаговых препаратов на основе Т4-родственных бактериофагов, что позволяет использовать их для испытаний на животных и клинических испытаний.

При помощи Т4-родственных вирусов на модельных животных (мышях) исследовались инфицирование нормофлоры фагами, распределение частиц по ЖКТ, иммуногенность белков вириона, иммуногенность фаговых частиц, терапевтический эффект фагов при моделировании инфекции патогенными штаммами бактерий. В [182] было показано, что введенные орально Т4-родственные бактериофаги, инфицирующие *E.coli*, ассоциированные с диареей, не оказывали значительного влияния на нормофлору мышей, однако они значительно лизировали свежевведенные клетки *E.coli*. В [183] исследовался транзит фага Т4 по кишечнику аксенических мышей, колонизированных чувствительным к фагу штаммом *E.coli*, показаны периоды времени, через которые фаги обнаруживались в различных отделах ЖКТ и фекалиях. Также для Т4-родственных бактериофагов на мышах показано, что наибольшей иммуногенностью среди компонентов

вириона обладает белок Нос [122]. При исследовании влияния Т4 фага на стимуляцию продукции антител в крови и кишечнике мыши при помощи введения Т4, разбавленным в питьевой воде, показано, что значительные повышения IgG и IgM происходит лишь при длительном введении фага (2 недели для IgG и 2 месяца для IgM) при высоких концентрациях частиц ($2 \cdot 10^{10}$ БОЕ в день). При исследовании влияния Т4 и его родственных ему фагов на организм животных, их распределение в организме были получены данные об иммуногенности данных вирусов и влияния отдельных компонентов на иммуногенность, скорости прохождения по отделам ЖКТ, что позволяет путем сравнения этих результатов определить возможное влияние Т4-родственных фагов на человеческий организм и выбрать необходимые испытания (необходимый набор доклинических исследований) при подборе конкретных штаммов вирусов для терапии.

Для Т4-родственных бактериофагов известны использование в клинической практике, а также клинические испытания I и II фазы. В клинических испытаниях I фазы [184] препараты Т4-родственных бактериофагов вводились оральным путем 15 взрослым людям и анализировались наличие фагов, появление специфических антител, уровни аланинаминотрансферазы и аспартатаминотрансферазы в сыворотке крови. В результате исследуемые фаги обнаруживались в фекалиях через день при использовании концентрации 10^5 БОЕ/мл, через неделю не обнаруживались, повышений в исследуемых параметрах не обнаружилось. В [176] проводилась II фаза клинических испытаний фагового коктейля, состоящего из 9 представителей Т4-родственных. 15 взрослым людям давали данный препарат в дозах $3 \cdot 10^9$ и $3 \cdot 10^7$ БОЕ/мл и плацебо, после чего проверялось наличие фагов в стуле, затем оценивали изменения в функциях почек и печени, а также гематологический статус и опрашивали субъектов на предмет жалоб со стороны ЖКТ. Фаги были обнаружены в исследованных стулах в большей мере при введении в большей концентрации (в 64%), в меньшей степени – в низкой концентрации (30%) и плацебо (28%). Из-за отсутствия в исходном стуле у субъектов *E.coli*, увеличение концентрации фагов не наблюдалось. Отсутствовало значительное влияние фагового препарата на микробиом субъектов,

нежелательных явлений на функции печени, почек и гематологии не наблюдались. В [185] описан случай клинического использования фагового коктейля, в котором 3 из 4 фагов являлись Т4-родственными. Данный препарат использовался на 56 мужчине вместе с эртапенемом, перенесшим трансплантацию печени и имевшего рецидивирующие инфекции предстательной железы и мочевыводящих путей, вызванных *E. coli*, продуцирующей бета-лактамазы расширенного спектра действия (ESBL). По окончании лечения, пациент не имел симптомов инфекции уринарного тракта, что свидетельствовало об успешной утилизации целевой ESBL, однако образовалась бессимптомная бактериурия, вызванная схожим с целевым штаммом ESBL. Таким образом, для Т4-родственных бактериофагов имеются данные о применениях их в клинической практике.

По рассмотренным литературным данным, Т4-родственные бактериофаги представляют собой одних из самых перспективных агентов фаговой терапии, за счет наличия достаточного количества исследований о получении и использовании препаратов на их основе. Однако, как показывает данное исследование, следует также оценивать возможность осуществления трансдукции Т4-родственными бактериофагами при подборе фагов для терапии.

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 Т4-родственные бактериофаги могут быть использованы для фаговой терапии. Например, представители родов *Krischvirus* и *Pseudotevenvirus*, обладают относительно высокой частотой трансдукции [154, 186] и у них имеются только гены дУТФазы-дЦТФазы и DenA,

которые хотя и могут снижать частоту трансдукции, но только в совокупности с другими генами, способствующими воспроизведению фагов с неканоническими основаниями ДНК, они снижают частоту трансдукции эффективно.

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* числа генов, способствующих воспроизведению фагов с неканоническими основаниями ДНК, а следовательно, у них отсутствует процесс трансдукции; способны инфицировать патогенные бактерии рода *Escherichia*, *Shigella*, *Yersinia*.

Также, в ходе анализа геномов фагов *Tevenvirinae* из базы данных NCBI Nucleotide Collection было отмечено, что фаги в подавляющем большинстве внутри родов обладают одинаковым набором генов, которые способствуют воспроизведению фагов с неканоническими основаниями ДНК, таким образом эти гены 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 Т4-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 Т4-родственных бактериофагах исследовались способы получения и производства фаговых препаратов для терапии.

В ходе проведенного исследования были рассмотрены способы скрининга бактериофагов для отбора Т4-родственных с неканоническими основаниями. В результате при помощи спот-теста на штаммах с RM-системами, ПЦР, гель-электрофореза частиц удалось выделить Т4-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 ⁵hm dC у T4-родственных бактериофагов располагаются в геномной области между двумя кор-генами – генами ДНК-полимеразы и хеликазы, а также наличие в этой области у представителей T4-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 T4-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 ЛИТЕРАТУРЫ

1. Koonin E. V. Are viruses alive? The replicator paradigm sheds decisive light on an old but misguided question / E. V. Koonin, P. Starokadomskyy // *Studies in history and philosophy of biological and biomedical sciences*. – 2016. – Vol. 59. – P. 125-134.
2. Ehrlich M. A novel, highly modified, bacteriophage DNA in which thymine is partly replaced by a phosphoglucuronate moiety covalently bound to 5-(4',5'-dihydroxypentyl)uracil / M. Ehrlich, K. C. Ehrlich // *The Journal of Biological Chemistry*. – 1981. – Vol. 256. – № 19. – P. 9966-9972.
3. Weigele P. Biosynthesis and Function of Modified Bases in Bacteria and Their Viruses / P. Weigele, E. A. Raleigh // *Chemical Reviews*. – 2016. – Vol. 116. – № 20. – P. 12655-12687.
4. 7-Deazaguanine modifications protect phage DNA from host restriction systems / G. Hutinet, W. Kot, L. Cui [et al.] // *Nature Communications*. – 2019. – Vol. 10. – № 1. – P. 5442.
5. Bacteriophage DNA glucosylation impairs target DNA binding by type I and II but not by type V CRISPR-Cas effector complexes / M. Vlot, J. Houkes, S. J. A. Lochs [et al.] // *Nucleic Acids Research*. – 2018. – Vol. 46. – № 2. – P. 873-885.
6. Genomes of the T4-related bacteriophages as windows on microbial genome evolution / V. M. Petrov, S. Ratnayaka, J. M. Nolan [et al.] // *Virology Journal*. – 2010. – Vol. 7. – № 1. – P. 292.
7. Abedon S. T. The murky origin of Snow White and her T-even dwarfs. / S. T. Abedon // *Genetics*. – 2000. – Vol. 155. – № 2. – P. 481-486.
8. The Odd “RB” Phage-Identification of Arabinosylation as a New Epigenetic Modification of DNA in T4-Like Phage RB69 / J. A. Thomas, J. Orwenyo, L.-X. Wang, L. W. Black // *Viruses*. – 2018. – Vol. 10. – № 6. – P. 313.
9. Amplification and purification of T4-like escherichia coli phages for phage therapy: from laboratory to pilot scale / G. Bourdin, B. Schmitt, L. Marvin Guy [et al.] // *Applied and Environmental Microbiology*. – 2014. – Vol. 80. – № 4. – P. 1469-1476.
10. Young K. K. Genetic analysis of bacteriophage T4 transducing bacteriophages. / K. K. Young, G. J. Edlin, G. G. Wilson // *Journal of Virology*. – 1982. – Vol. 41. – № 1. – P. 345-347.
11. Karam J. D. *Molecular Biology of Bacteriophage T4* / J. D. Karam, J. W. Drake. – American Society for Microbiology, 1994. – 648 p.
12. Bacteriophage T4 Genome / E. S. Miller, E. Kutter, G. Mosig [et al.] // *Microbiology and Molecular Biology Reviews*. – 2003. – Vol. 67. – № 1. – P. 86-156.

13. Delbrück M. Interference Between Bacterial Viruses / M. Delbrück // *Journal of Bacteriology*. – 1945. – Vol. 50. – № 2. – P. 151-170.
14. Delbrück M. The Burst Size Distribution in the Growth of Bacterial Viruses (Bacteriophages) / M. Delbrück // *Journal of Bacteriology*. – 1945. – Vol. 50. – № 2. – P. 131-135.
15. Luria S. E. Electron Microscope Studies of Bacterial Viruses / S. E. Luria, M. Delbrück, T. F. Anderson // *Journal of Bacteriology*. – 1943. – Vol. 46. – № 1. – P. 57-77.
16. Hershey A. D. Mutation of Bacteriophage with Respect to Type of Plaque / A. D. Hershey // *Genetics*. – 1946. – Vol. 31. – № 6. – P. 620-640.
17. Hershey A. D. Coordinate Effects of Electrolyte and Antibody on the Infectivity of Bacteriophage / A. D. Hershey, G. M. Kalmanson, J. Bronfenbrenner // *The Journal of Immunology*. – 1944. – Vol. 48. – № 4. – P. 221-239.
18. Kalmanson G. Studies on the purification of bacteriophage / G. Kalmanson, J. Bronfenbrenner // *The Journal of General Physiology*. – 1939. – Vol. 23. – № 2. – P. 203-228.
19. Demerec M. Bacteriophage-Resistant Mutants in *Escherichia Coli* / M. Demerec, U. Fano // *Genetics*. – 1945. – Vol. 30. – № 2. – P. 119-136.
20. Benzer S. Fine structure of a genetic region in bacteriophage / S. Benzer // *Proceedings of the National Academy of Sciences of the United States of America*. – 1955. – Vol. 41. – № 6. – P. 344-354.
21. Doermann A. H. Genetic Structure of Bacteriophage T4 as Described by Recombination Studies of Factors Influencing Plaque Morphology / A. H. Doermann, M. B. Hill // *Genetics*. – 1953. – Vol. 38. – № 1. – P. 79-90.
22. Tracing ancestors and relatives of *Escherichia coli* B, and the derivation of B strains REL606 and BL21(DE3) / P. Daegelen, F. W. Studier, R. E. Lenski [et al.] // *Journal of Molecular Biology*. – 2009. – Vol. 394. – № 4. – P. 634-643.
23. Hershey A. D. Independent functions of viral protein and nucleic acid in growth of bacteriophage / A. D. Hershey, M. Chase // *The Journal of General Physiology*. – 1952. – Vol. 36. – № 1. – P. 39-56.
24. Edgar B. The Genome of Bacteriophage T4: An Archeological Dig / B. Edgar // *Genetics*. – 2004. – Vol. 168. – № 2. – P. 575-582.
25. Delbruck M. Bacterial viruses or bacteriophages / M. Delbruck // *Biological Reviews of the Cambridge Philosophical Society*. – 1946. – Vol. 21. – P. 30-40.

26. Adams M. H. Classification of Bacterial Viruses : Characteristics of the T5 Species and of the T2, C16 Species / M. H. Adams // *Journal of Bacteriology*. – 1952. – Vol. 64. – № 3. – P. 387-396.
27. Burnet F. M. The classification of dysentery-Coli bacteriophages. III. A correlation of the serological classification with certain biochemical tests / F. M. Burnet // *The Journal of Pathology and Bacteriology*. – 1933. – Vol. 37. – № 2. – P. 179-184.
28. Ackermann H. W. A catalogue of T4-type bacteriophages / H. W. Ackermann, H. M. Krisch // *Archives of Virology*. – 1997. – Vol. 142. – № 12. – P. 2329-2345.
29. Tailed phages of *Pseudomonas* and related bacteria / A. Liss, H. W. Ackermann, L. W. Mayer, C. H. Zierdt // *Intervirology*. – 1981. – Vol. 15. – № 2. – P. 71-81.
30. Degrees of relatedness of T-even type *E. coli* phages using different or the same receptors and topology of serologically cross-reacting sites / H. Schwarz, I. Riede, I. Sonntag, U. Henning // *The EMBO journal*. – 1983. – Vol. 2. – № 3. – P. 375-380.
31. The genome of the pseudo T-even bacteriophages, a diverse group that resembles T4 / C. Monod, F. Repoila, M. Kutateladze [et al.] // *Journal of Molecular Biology*. – 1997. – Vol. 267. – № 2. – P. 237-249.
32. Russell R. L. Speciation among the T-even bacteriophages : phd / R. L. Russell. – California Institute of Technology, 1967. – 208 p.
33. Phylogeny of the major head and tail genes of the wide-ranging T4-type bacteriophages / F. Tétart, C. Desplats, M. Kutateladze [et al.] // *Journal of Bacteriology*. – 2001. – Vol. 183. – № 1. – P. 358-366.
34. Desplats C. The diversity and evolution of the T4-type bacteriophages / C. Desplats, H. M. Krisch // *Research in Microbiology*. – 2003. – Vol. 154. – № 4. – P. 259-267.
35. Evolution of T4-related phages / E. Kutter, K. Gachechiladze, A. Poglazov [et al.] // *Virus Genes*. – 1995. – Vol. 11. – № 2-3. – P. 285-297.
36. Karam J. D. Bacteriophage T4 and its relatives / J. D. Karam, E. S. Miller // *Virology Journal*. – 2010. – Vol. 7. – P. 293.
37. Comeau A. M. The Capsid of the T4 Phage Superfamily: The Evolution, Diversity, and Structure of Some of the Most Prevalent Proteins in the Biosphere / A. M. Comeau, H. M. Krisch // *Molecular Biology and Evolution*. – 2008. – Vol. 25. – № 7. – P. 1321-1332.
38. Genomic characteristics and environmental distributions of the uncultivated Far-T4 phages / S. Roux, F. Enault, V. Ravet [et al.] // *Frontiers in Microbiology*. – 2015. – Vol. 6. – P. 199.

39. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV) / E. J. Lefkowitz, D. M. Dempsey, R. C. Hendrickson [et al.] // *Nucleic Acids Research*. – 2018. – Vol. 46. – № D1. – P. D708-D717.
40. Kuhn J. H. Virus Taxonomy / J. H. Kuhn // *Encyclopedia of Virology*. – 2021. – P. 28-37.
41. Taxon Details | ICTV. – URL: https://ictv.global/taxonomy/taxondetails?taxnode_id=202200332 (дата обращения: 18.04.2023). – Текст : электронный.
42. Four principles to establish a universal virus taxonomy / P. Simmonds, E. M. Adriaenssens, F. M. Zerbini [et al.] // *PLOS Biology*. – 2023. – Vol. 21. – № 2. – P. e3001922.
43. Comparative genomics of marine cyanomyoviruses reveals the widespread occurrence of *Synechococcus* host genes localized to a hyperplastic region: implications for mechanisms of cyanophage evolution / A. D. Millard, K. Zwirgmaier, M. J. Downey [et al.] // *Environmental Microbiology*. – 2009. – Vol. 11. – № 9. – P. 2370-2387.
44. Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments / M. B. Sullivan, K. H. Huang, J. C. Ignacio-Espinoza [et al.] // *Environmental Microbiology*. – 2010. – Vol. 12. – № 11. – P. 3035-3056.
45. Ignacio-Espinoza J. C. Phylogenomics of T4 cyanophages: lateral gene transfer in the “core” and origins of host genes / J. C. Ignacio-Espinoza, M. B. Sullivan // *Environmental Microbiology*. – 2012. – Vol. 14. – № 8. – P. 2113-2126.
46. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome” / H. Tettelin, V. Masignani, M. J. Cieslewicz [et al.] // *Proceedings of the National Academy of Sciences*. – 2005. – Vol. 102. – № 39. – P. 13950-13955.
47. Comparative genomics: the bacterial pan-genome : Antimicrobials/Genomics / H. Tettelin, D. Riley, C. Cattuto, D. Medini // *Current Opinion in Microbiology*. – 2008. – Vol. 11. – № 5. – P. 472-477.
48. Koonin E. V. Orthologs, paralogs, and evolutionary genomics / E. V. Koonin // *Annual Review of Genetics*. – 2005. – Vol. 39. – P. 309-338.
49. Contreras-Moreira B. GET_HOMOLOGUES, a Versatile Software Package for Scalable and Robust Microbial Pangenome Analysis / B. Contreras-Moreira, P. Vinuesa // *Applied and Environmental Microbiology*. – 2013. – Vol. 79. – № 24. – P. 7696-7701.
50. The microbial pan-genome / D. Medini, C. Donati, H. Tettelin [et al.] // *Current Opinion in Genetics & Development*. – 2005. – Vol. 15. – № 6. – P. 589-594.

51. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes / R. S. Kaas, C. Friis, D. W. Ussery, F. M. Aarestrup // BMC genomics. – 2012. – Vol. 13. – P. 577.
52. Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer / Y. I. Wolf, K. S. Makarova, N. Yutin, E. V. Koonin // Biology Direct. – 2012. – Vol. 7. – P. 46.
53. Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery / A. M. Comeau, C. Bertrand, A. Letarov [et al.] // Virology. – 2007. – Vol. 362. – № 2. – P. 384-396.
54. Ten years of pan-genome analyses / G. Vernikos, D. Medini, D. R. Riley, H. Tettelin // Current Opinion in Microbiology. – 2015. – Vol. 23. – P. 148-154.
55. Complete genome sequence of the broad-host-range vibriophage KVP40: comparative genomics of a T4-related bacteriophage / E. S. Miller, J. F. Heidelberg, J. A. Eisen [et al.] // Journal of Bacteriology. – 2003. – Vol. 185. – № 17. – P. 5220-5233.
56. Plasticity of the gene functions for DNA replication in the T4-like phages / V. M. Petrov, J. M. Nolan, C. Bertrand [et al.] // Journal of Molecular Biology. – 2006. – Vol. 361. – № 1. – P. 46-68.
57. Eddy S. R. Artificial mobile DNA element constructed from the EcoRI endonuclease gene / S. R. Eddy, L. Gold // Proceedings of the National Academy of Sciences of the United States of America. – 1992. – Vol. 89. – № 5. – P. 1544-1547.
58. Zeng Q. A free-standing homing endonuclease targets an intron insertion site in the *psbA* gene of cyanophages / Q. Zeng, R. P. Bonocora, D. A. Shub // Current biology: CB. – 2009. – Vol. 19. – № 3. – P. 218-222.
59. Mobile regulatory cassettes mediate modular shuffling in T4-type phage genomes / C. Arbiol, A. M. Comeau, M. Kutateladze [et al.] // Genome Biology and Evolution. – 2010. – Vol. 2. – P. 140-152.
60. Comeau A. M. Composite conserved promoter-terminator motifs (PeSLs) that mediate modular shuffling in the diverse T4-like myoviruses / A. M. Comeau, C. Arbiol, H. M. Krisch // Genome Biology and Evolution. – 2014. – Vol. 6. – № 7. – P. 1611-1619.
61. Identification and biosynthesis of thymidine hypermodifications in the genomic DNA of widespread bacterial viruses / Y.-J. Lee, N. Dai, S. E. Walsh [et al.] // Proceedings of the National Academy of Sciences of the United States of America. – 2018. – Vol. 115. – № 14. – P. E3116-E3125.
62. Lederberg E. M. Genetic Studies of Lysogenicity in *Escherichia Coli* / E. M. Lederberg, J. Lederberg // Genetics. – 1953. – Vol. 38. – № 1. – P. 51-64.

63. A widespread pathway for substitution of adenine by diaminopurine in phage genomes / Y. Zhou, X. Xu, Y. Wei [et al.] // *Science* (New York, N.Y.). – 2021. – Vol. 372. – № 6541. – P. 512-516.
64. 2-aminoadenine is an adenine substituting for a base in S-2L cyanophage DNA / M. D. Kirnos, I. Y. Khudyakov, N. I. Alexandrushkina, B. F. Vanyushin // *Nature*. – 1977. – Vol. 270. – № 5635. – P. 369-370.
65. A third purine biosynthetic pathway encoded by aminoadenine-based viral DNA genomes / D. Sleiman, P. S. Garcia, M. Lagune [et al.] // *Science* (New York, N.Y.). – 2021. – Vol. 372. – № 6541. – P. 516-520.
66. Hypermodified DNA in Viruses of *E. coli* and *Salmonella* / G. Hutinet, Y.-J. Lee, V. de Crécy-Lagard, P. R. Weigele // *EcoSal Plus*. – 2021. – Vol. 9. – № 2. – P. eESP00282019.
67. Speedpera. English: Chemical structure of N6-carbamoyl methyladenine. – URL: <https://commons.wikimedia.org/wiki/File:N6-Carbamoylmethyladenin.svg> (date accessed: 25.04.2023). – Text : electronic.
68. Purification and characterization of the unusual deoxynucleoside, alpha-N-(9-beta-D-2'-deoxyribofuranosylpurin-6-yl)glycinamide, specified by the phage Mu modification function. / D. Swinton, S. Hattman, P. F. Crain [et al.] // *Proceedings of the National Academy of Sciences of the United States of America*. – 1983. – Vol. 80. – № 24. – P. 7400-7404.
69. Modification of DNA by a viral enzyme and charged tRNA / R. M. B. Silva, A. Slyvka, Y.-J. Lee [et al.] // *bioRxiv*. – 2023. – P. 2023.03.24.534169.
70. Specificity and functions of guanine methylase of *Shigella sonnei* DDVI phage / I. I. Nikolskaya, M. I. Tediashvili, N. G. Lopatina [et al.] // *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*. – 1979. – Vol. 561. – № 1. – P. 232-239.
71. 7-Methylguanine // Page Version ID: 1047868421. – URL: <https://en.wikipedia.org/w/index.php?title=7-Methylguanine&oldid=1047868421> (date accessed: 25.04.2023). – Text : electronic.
72. Novel mechanism of post-transcriptional modification of tRNA. Insertion of bases of Q precursors into tRNA by a specific tRNA transglycosylase reaction / N. Okada, S. Noguchi, H. Kasai [et al.] // *The Journal of Biological Chemistry*. – 1979. – Vol. 254. – № 8. – P. 3067-3073.
73. Structure of the archaeal transfer RNA nucleoside G*-15 (2-amino-4,7-dihydro-4-oxo-7-beta-D-ribofuranosyl-1H-pyrrolo[2,3-d]pyrimidine-5-carboximide (archaeosine)) / J. M. Gregson, P. F. Crain, C. G. Edmonds [et al.] // *The Journal of Biological Chemistry*. – 1993. – Vol. 268. – № 14. – P. 10076-10086.

74. Genomic Sequencing and Biological Characteristics of a Novel Escherichia Coli Bacteriophage 9g, a Putative Representative of a New Siphoviridae Genus / E. E. Kulikov, A. K. Golomidova, M. A. Letarova [et al.] // Viruses. – 2014. – Vol. 6. – № 12. – P. 5077-5092.
75. Novel genomic island modifies DNA with 7-deazaguanine derivatives / J. J. Thiaville, S. M. Kellner, Y. Yuan [et al.] // Proceedings of the National Academy of Sciences. – 2016. – Vol. 113. – № 11. – P. E1452-E1459.
76. Deoxyinosine and 7-Deaza-2-Deoxyguanosine as Carriers of Genetic Information in the DNA of Campylobacter Viruses / C. S. Crippen, Y.-J. Lee, G. Hutinet [et al.] // Journal of Virology. – 2019. – Vol. 93. – № 23. – P. e01111-19.
77. Katz G. E. Bacteriophage PBS2-induced inhibition of uracil-containing DNA degradation / G. E. Katz, A. R. Price, M. J. Pomerantz // Journal of Virology. – 1976. – Vol. 20. – № 2. – P. 535-538.
78. Replacement of the deoxycytidine residues in Rhizobium bacteriophage RL38JI DNA / D. Swinton, S. Hattman, R. Benzinger [et al.] // FEBS letters. – 1985. – Vol. 184. – № 2. – P. 294-298.
79. 5-Hydroxycytosine // Page Version ID: 1102530242. – URL: <https://en.wikipedia.org/w/index.php?title=5-Hydroxycytosine&oldid=1102530242> (date accessed: 27.04.2023). – Text : electronic.
80. The Sequence of Two Bacteriophages with Hypermodified Bases Reveals Novel Phage-Host Interactions / A. M. Kropinski, D. Turner, J. H. E. Nash [et al.] // Viruses. – 2018. – Vol. 10. – № 5. – P. 217.
81. Production and expression of dTMP-enriched DNA of bacteriophage SP15. / E. Casella, O. Markewych, M. Dosmar, H. Witmer // Journal of Virology. – 1978. – Vol. 28. – № 3. – P. 753-766.
82. Biochemical characterization of recombinant β -glucosyltransferase and analysis of global 5-hydroxymethylcytosine in unique genomes / J. Terragni, J. Bitinaite, Y. Zheng, S. Pradhan // Biochemistry. – 2012. – Vol. 51. – № 5. – P. 1009-1019.
83. Pathways of thymidine hypermodification / Y.-J. Lee, N. Dai, S. I. Müller [et al.] // Nucleic Acids Research. – 2022. – Vol. 50. – № 6. – P. 3001-3017.
84. Witmer H. Synthesis of deoxythymidylate and the unusual deoxynucleotide in mature DNA of Bacillus subtilis bacteriophage SP10 occurs by postreplicational modification of 5-hydroxymethyldeoxyuridylate / H. Witmer // Journal of Virology. – 1981. – Vol. 39. – № 2. – P. 536-547.

85. Kropinski A. M. 5-(4-Aminobutylaminomethyl)uracil, an unusual pyrimidine from the deoxyribonucleic acid of bacteriophage phiW-14 / A. M. Kropinski, R. J. Bose, R. A. Warren // *Biochemistry*. – 1973. – Vol. 12. – № 1. – P. 151-157.
86. Gold L. M. Synthesis of phage-specific alpha- and beta-glucosyl transferases directed by T-even DNA in vitro / L. M. Gold, M. Schweiger // *Proceedings of the National Academy of Sciences of the United States of America*. – 1969. – Vol. 62. – № 3. – P. 892-898.
87. Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases / M. P. Callahan, K. E. Smith, H. J. Cleaves [et al.] // *Proceedings of the National Academy of Sciences*. – 2011. – Vol. 108. – № 34. – P. 13995-13998.
88. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes / R. J. Roberts, M. Belfort, T. Bestor [et al.] // *Nucleic Acids Research*. – 2003. – Vol. 31. – № 7. – P. 1805-1812.
89. Loenen W. A. M. The other face of restriction: modification-dependent enzymes / W. A. M. Loenen, E. A. Raleigh // *Nucleic Acids Research*. – 2014. – Vol. 42. – № 1. – P. 56-69.
90. Comparative characterization of the PvuRts1I family of restriction enzymes and their application in mapping genomic 5-hydroxymethylcytosine / H. Wang, S. Guan, A. Quimby [et al.] // *Nucleic Acids Research*. – 2011. – Vol. 39. – № 21. – P. 9294-9305.
91. Bair C. L. A Type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. / C. L. Bair, L. W. Black // *Journal of molecular biology*. – 2007. – Vol. 366. – № 3. – P. 768-778.
92. The T4 Phage DNA Mimic Protein Arn Inhibits the DNA Binding Activity of the Bacterial Histone-like Protein H-NS / C.-H. Ho, H.-C. Wang, T.-P. Ko [et al.] // *The Journal of Biological Chemistry*. – 2014. – Vol. 289. – № 39. – P. 27046-27054.
93. Restriction endonuclease inhibitor IPI* of bacteriophage T4: a novel structure for a dedicated target / D. Rifat, N. T. Wright, K. M. Varney [et al.] // *Journal of Molecular Biology*. – 2008. – Vol. 375. – № 3. – P. 720-734.
94. Lagerbäck P. Amino Acid Residues in the GIY-YIG Endonuclease II of Phage T4 Affecting Sequence Recognition and Binding as Well as Catalysis / P. Lagerbäck, K. Carlson // *Journal of Bacteriology*. – 2008. – Vol. 190. – № 16. – P. 5533-5544.
95. Carlson K. Bacteriophage T4 endonuclease II: concerted single-strand nicks yield double-strand cleavage / K. Carlson, P. Lagerbäck, A.-C. Nyström // *Molecular Microbiology*. – 2004. – Vol. 52. – № 5. – P. 1403-1411.
96. Ohshima H. A hexanucleotide sequence (dC1–dC6 tract) restricts the dC-specific cleavage of single-stranded DNA by endonuclease IV of bacteriophage T4 / H. Ohshima,

N. Hirano, H. Takahashi // *Nucleic Acids Research*. – 2007. – Vol. 35. – № 19. – P. 6681-6689.

97. The production of undegraded cytosine-containing DNA by bacteriophage T4 in the absence of dCTPase and endonucleases II and IV, and its effects on T4-directed protein synthesis / E. Kutter, A. Beug, R. Sluss [et al.] // *Journal of Molecular Biology*. – 1975. – Vol. 99. – № 4. – P. 591-607.

98. Bacteriophage T4 Alc protein: a transcription termination factor sensing local modification of DNA / M. Kashlev, E. Nudler, A. Goldfarb [et al.] // *Cell*. – 1993. – Vol. 75. – № 1. – P. 147-154.

99. The phage T4 MotA transcription factor contains a novel DNA binding motif that specifically recognizes modified DNA / M. G. Cuypers, R. M. Robertson, L. Knipling [et al.] // *Nucleic Acids Research*. – 2018. – Vol. 46. – № 10. – P. 5308-5318.

100. High-frequency generalised transduction by bacteriophage T4 / G. G. Wilson, K. Y. Young, G. J. Edlin, W. Konigsberg // *Nature*. – 1979. – Vol. 280. – № 5717. – P. 80-82.

101. Selick H. E. The bacteriophage T4 insertion/substitution vector system. A method for introducing site-specific mutations into the virus chromosome / H. E. Selick, K. N. Kreuzer, B. M. Alberts // *The Journal of Biological Chemistry*. – 1988. – Vol. 263. – № 23. – P. 11336-11347.

102. Bacteriophage-Bacteria Interactions in the Gut: From Invertebrates to Mammals / J. M. Kirsch, R. S. Brzozowski, D. Faith [et al.] // *Annual Review of Virology*. – 2021. – Vol. 8. – № 1. – P. 95-113.

103. Stern A. The phage-host arms race: shaping the evolution of microbes / A. Stern, R. Sorek // *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*. – 2011. – Vol. 33. – № 1. – P. 43-51.

104. Lytic to temperate switching of viral communities / B. Knowles, C. B. Silveira, B. A. Bailey [et al.] // *Nature*. – 2016. – Vol. 531. – № 7595. – P. 466-470.

105. Rapid evolution of the human gut virome / S. Minot, A. Bryson, C. Chehoud [et al.] // *Proceedings of the National Academy of Sciences of the United States of America*. – 2013. – Vol. 110. – № 30. – P. 12450-12455.

106. Virulent coliphages in 1-year-old children fecal samples are fewer, but more infectious than temperate coliphages / A. Mathieu, M. Dion, L. Deng [et al.] // *Nature Communications*. – 2020. – Vol. 11. – № 1. – P. 378.

107. Prophages mediate defense against phage infection through diverse mechanisms / J. Bondy-Denomy, J. Qian, E. R. Westra [et al.] // *The ISME journal*. – 2016. – Vol. 10. – № 12. – P. 2854-2866.

108. Maslov S. Population cycles and species diversity in dynamic Kill-the-Winner model of microbial ecosystems / S. Maslov, K. Sneppen // *Scientific Reports*. – 2017. – Vol. 7. – P. 39642.
109. Mizuno C. M. Evidence for metaviromic islands in marine phages / C. M. Mizuno, R. Ghai, F. Rodriguez-Valera // *Frontiers in Microbiology*. – 2014. – Vol. 5. – P. 27.
110. Explaining microbial population genomics through phage predation / F. Rodriguez-Valera, A.-B. Martin-Cuadrado, B. Rodriguez-Brito [et al.] // *Nature Reviews Microbiology*. – 2009. – Vol. 7. – № 11. – P. 828-836.
111. Luria S. E. The Identification and Characterization of Bacteriophages with the Electron Microscope / S. E. Luria, T. F. Anderson // *Proceedings of the National Academy of Sciences of the United States of America*. – 1942. – Vol. 28. – № 4. – P. 127-130.1.
112. Studier F. W. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes / F. W. Studier, B. A. Moffatt // *Journal of Molecular Biology*. – 1986. – Vol. 189. – № 1. – P. 113-130.
113. Chromosomal mutations that increase the production of a plasmid-encoded haemolysin in *Escherichia coli* / N. Godessart, F. J. Muñoa, M. Regue, A. Juárez // *Journal of General Microbiology*. – 1988. – Vol. 134. – № 10. – P. 2779-2787.
114. Stephanie H. Genetic map of coliphage 186 : phd / H. Stephanie. – University of Adelaide, 1977. – 278 p.
115. Flow cytometry studies of recombinant *Escherichia coli* in batch and continuous cultures: DNA and RNA contents; light-scatter parameters / P. Fouchet, C. Manin, H. Richard [et al.] // *Applied Microbiology and Biotechnology*. – 1994. – Vol. 41. – № 5. – P. 584-590.
116. Properties of the replicator region of the natural plasmid pLG13 containing genes of the EcoRV restriction-modification system / M. V. Zakharova, A. N. Kravets, N. V. Klimashina, A. S. Solonin // *Molekuliarnaiia Biologiya*. – 1991. – Vol. 25. – № 6. – P. 1615-1625.
117. Schneider S. S. Identification by immunobinding assay of the polypeptide coded by the DNA polymerase gene of bacteriophage T5 and its amber mutants and the direction of transcription of the gene / S. S. Schneider, B. C. Roop, R. K. Fujimura // *Journal of Virology*. – 1985. – Vol. 56. – № 1. – P. 245-249.
118. Hayes S. Initiation of coliphage lambda replication, lit, oop RNA synthesis, and effect of gene dosage on transcription from promoters PL, PR, and PR / S. Hayes // *Virology*. – 1979. – Vol. 97. – № 2. – P. 415-438.
119. ImageJ. – URL: <https://imagej.nih.gov/ij/> (date accessed: 25.05.2023). – Text : electronic.

120. Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere / J. Filée, F. Tétart, C. A. Suttle, H. M. Krisch // *Proceedings of the National Academy of Sciences of the United States of America*. – 2005. – Vol. 102. – № 35. – P. 12471-12476.
121. Functional Analysis of the Highly Antigenic Outer Capsid Protein, Hoc, a Virus Decoration Protein from T4-like Bacteriophages / T. Sathaliyawala, M. Z. Islam, Q. Li [et al.] // *Molecular microbiology*. – 2010. – Vol. 77. – № 2. – P. 444-455.
122. Immunogenicity studies of proteins forming the T4 phage head surface / K. Dąbrowska, P. Miernikiewicz, A. Piotrowicz [et al.] // *Journal of Virology*. – 2014. – Vol. 88. – № 21. – P. 12551-12557.
123. Yamaguchi Y. Head shell protein hoc alters the surface charge of bacteriophage T4. Composite slab gel electrophoresis of phage T4 and related particles / Y. Yamaguchi, M. Yanagida // *Journal of Molecular Biology*. – 1980. – Vol. 141. – № 2. – P. 175-193.
124. Chauleau M. Kinetic mechanism of nick sealing by T4 RNA ligase 2 and effects of 3'-OH base mispairs and damaged base lesions / M. Chauleau, S. Shuman // *RNA*. – 2013. – Vol. 19. – № 12. – P. 1840-1847.
125. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction / J. Ye, G. Coulouris, I. Zaretskaya [et al.] // *BMC Bioinformatics*. – 2012. – Vol. 13. – P. 134.
126. Serwer P. In-Gel Isolation and Characterization of Large (and Other) Phages / P. Serwer, E. T. Wright // *Viruses*. – 2020. – Vol. 12. – № 4. – P. 410.
127. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads / R. R. Wick, L. M. Judd, C. L. Gorrie, K. E. Holt // *PLoS computational biology*. – 2017. – Vol. 13. – № 6. – P. e1005595.
128. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes / T. Brettin, J. J. Davis, T. Disz [et al.] // *Scientific Reports*. – 2015. – Vol. 5. – P. 8365.
129. PATRIC, the bacterial bioinformatics database and analysis resource / A. R. Wattam, D. Abraham, O. Dalay [et al.] // *Nucleic Acids Research*. – 2014. – Vol. 42. – № Database issue. – P. D581-591.
130. Grant J. R. The CGView Server: a comparative genomics tool for circular genomes / J. R. Grant, P. Stothard // *Nucleic Acids Research*. – 2008. – Vol. 36. – № Web Server issue. – P. W181-W184.
131. Vinuesa P. GET_PHYLOMARKERS, a Software Package to Select Optimal Orthologous Clusters for Phylogenomics and Inferring Pan-Genome Phylogenies, Used for a Critical Geno-Taxonomic Revision of the Genus *Stenotrophomonas* / P. Vinuesa, L.

E. Ochoa-Sánchez, B. Contreras-Moreira // *Frontiers in Microbiology*. – 2018. – Vol. 9. – P. 771.

132. Edgar R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity / R. C. Edgar // *BMC bioinformatics*. – 2004. – Vol. 5. – P. 113.

133. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis / J. Trifinopoulos, L.-T. Nguyen, A. von Haeseler, B. Q. Minh // *Nucleic Acids Research*. – 2016. – Vol. 44. – № W1. – P. W232-235.

134. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms / S. Kumar, G. Stecher, M. Li [et al.] // *Molecular Biology and Evolution*. – 2018. – Vol. 35. – № 6. – P. 1547-1549.

135. Jones D. T. The rapid generation of mutation data matrices from protein sequences / D. T. Jones, W. R. Taylor, J. M. Thornton // *Computer applications in the biosciences: CABIOS*. – 1992. – Vol. 8. – № 3. – P. 275-282.

136. Meier-Kolthoff J. P. VICTOR: genome-based phylogeny and classification of prokaryotic viruses / J. P. Meier-Kolthoff, M. Göker // *Bioinformatics (Oxford, England)*. – 2017. – Vol. 33. – № 21. – P. 3396-3404.

137. Cytoscape: a software environment for integrated models of biomolecular interaction networks / P. Shannon, A. Markiel, O. Ozier [et al.] // *Genome Research*. – 2003. – Vol. 13. – № 11. – P. 2498-2504.

138. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs / S. F. Altschul, T. L. Madden, A. A. Schäffer [et al.] // *Nucleic Acids Research*. – 1997. – Vol. 25. – № 17. – P. 3389-3402.

139. Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements / A. A. Schäffer, L. Aravind, T. L. Madden [et al.] // *Nucleic Acids Research*. – 2001. – Vol. 29. – № 14. – P. 2994-3005.

140. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core / L. Zimmermann, A. Stephens, S.-Z. Nam [et al.] // *Journal of Molecular Biology*. – 2018. – Vol. 430. – № 15. – P. 2237-2243.

141. Sullivan M. J. Easyfig: a genome comparison visualizer / M. J. Sullivan, N. K. Petty, S. A. Beatson // *Bioinformatics (Oxford, England)*. – 2011. – Vol. 27. – № 7. – P. 1009-1010.

142. Никулин Н. А. Бактериофаговая флора очистных сооружений города Пущино Московской области / Н. А. Никулин, А. А. Зимин // в книге: Биохимия, физиология и биосферная роль микроорганизмов. Сборник тезисов III Пущинской школы-конференции. Под редакцией Т.А. Решетиловой. 2016. С. 143-146.

143. Abedon S. T. Look Who's Talking: T-Even Phage Lysis Inhibition, the Granddaddy of Virus-Virus Intercellular Communication Research / S. T. Abedon // *Viruses*. – 2019. – Vol. 11. – № 10. – P. 951.
144. Classification of Myoviridae bacteriophages using protein sequence similarity / R. Lavigne, P. Darius, E. J. Summer [et al.] // *BMC Microbiology*. – 2009. – Vol. 9. – № 1. – P. 224.
145. Serwer P. Gated Ethidium- and Bleomycin-Loading in Phage T4 That Is Subsequently Purified Leak-Free / P. Serwer, E. T. Wright // *Biophysica*. – 2022. – Vol. 2. – № 4. – P. 366-380.
146. Basic local alignment search tool / S. F. Altschul, W. Gish, W. Miller [et al.] // *Journal of Molecular Biology*. – 1990. – Vol. 215. – № 3. – P. 403-410.
147. Adriaenssens E. M. How to Name and Classify Your Phage: An Informal Guide / E. M. Adriaenssens, J. R. Brister // *Viruses*. – 2017. – Vol. 9. – № 4. – P. 70.
148. 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.
149. PPanGGOLiN: Depicting microbial diversity via a partitioned pangenome graph / G. Gautreau, A. Bazin, M. Gachet [et al.] // *PLoS computational biology*. – 2020. – Vol. 16. – № 3. – P. e1007732.
150. 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. – № 5. – P. 88-108.
151. Comparative Functional Genomic Analysis of Two Vibrio Phages Reveals Complex Metabolic Interactions with the Host Cell / D. Skliros, P. G. Kalatzis, P. Katharios, E. Flemetakis // *Frontiers in Microbiology*. – 2016. – Vol. 7. – P. 1807.
152. Genetic diversity among five T4-like bacteriophages / J. M. Nolan, V. Petrov, C. Bertrand [et al.] // *Virology Journal*. – 2006. – Vol. 3. – P. 30.
153. Computational identification of novel biochemical systems involved in oxidation, glycosylation and other complex modifications of bases in DNA / L. M. Iyer, D. Zhang, A. M. Burroughs, L. Aravind // *Nucleic Acids Research*. – 2013. – Vol. 41. – № 16. – P. 7635-7655.
154. Transduction of Plasmid Antibiotic Resistance Determinants with PseudoT-Even Bacteriophages / V. I. Tanyashin, A. A. Zimin, M. G. Shlyapnikov, A. M. Boronin // *Russian Journal of Genetics*. – 2003. – Vol. 39. – № 7. – P. 761-772.

155. Bair C. L. Exclusion of glucosyl-hydroxymethylcytosine DNA containing bacteriophages is overcome by the injected protein inhibitor IPI* / C. L. Bair, D. Rifat, L. W. Black // *Journal of Molecular Biology*. – 2007. – Vol. 366. – № 3. – P. 779-789.
156. Nucleotide sequence and revised map location of the *arn* gene from bacteriophage T4 / B. C. Kim, K. Kim, E. H. Park, C. J. Lim // *Molecules and Cells*. – 1997. – Vol. 7. – № 5. – P. 694-696.
157. 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.
158. 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.
159. Ronner A. B. Isolation and Characterization of a Coliphage Specific for *Escherichia coli* O157:H7 / A. B. Ronner, D. O. Cliver // *Journal of Food Protection*. – 1990. – Vol. 53. – № 11. – P. 944-947.
160. Selection and characterization of a candidate therapeutic bacteriophage that lyses the *Escherichia coli* O104:H4 strain from the 2011 outbreak in Germany / M. Merabishvili, D. De Vos, G. Verbeke [et al.] // *PloS One*. – 2012. – Vol. 7. – № 12. – P. e52709.
161. Bacteriophages with the ability to degrade uropathogenic *Escherichia coli* biofilms / A. Chibeu, E. J. Lingohr, L. Masson [et al.] // *Viruses*. – 2012. – Vol. 4. – № 4. – P. 471-487.
162. Complete Genome Sequence of vB_EcoM_112, a T-Even-Type Bacteriophage Specific for *Escherichia coli* O157:H7 / B. Coffey, R. P. Ross, G. O'Flynn [et al.] // *Genome Announcements*. – 2014. – Vol. 2. – № 6. – P. e00393-14.
163. Bacteriophage cocktail significantly reduces *Escherichia coli* O157: H7 contamination of lettuce and beef, but does not protect against recontamination / C. D. Carter, A. Parks, T. Abuladze [et al.] // *Bacteriophage*. – 2012. – Vol. 2. – № 3. – P. 178-185.
164. Characterization and Genomic Study of the Novel Bacteriophage HY01 Infecting Both *Escherichia coli* O157:H7 and *Shigella flexneri*: Potential as a Biocontrol Agent in Food / H. Lee, H.-J. Ku, D.-H. Lee [et al.] // *PloS One*. – 2016. – Vol. 11. – № 12. – P. e0168985.
165. Analysis of whole genome sequencing for the *Escherichia coli* O157:H7 typing phages / L. A. Cowley, S. J. Beckett, M. Chase-Topping [et al.] // *BMC genomics*. – 2015. – Vol. 16. – № 1. – P. 271.

166. Isolation, Characterisation and Complete Genome Sequence of a Tequatrovirus Phage, Escherichia phage KIT03, Which Simultaneously Infects Escherichia coli O157:H7 and Salmonella enterica / N. H. Pham-Khanh, H. Sunahara, H. Yamadeya [et al.] // Current Microbiology. – 2019. – Vol. 76. – № 10. – P. 1130-1137.
167. Characterization of a virulent bacteriophage specific for Escherichia coli O157:H7 and analysis of its cellular receptor and two tail fiber genes / M. Morita, Y. Tanji, K. Mizoguchi [et al.] // FEMS microbiology letters. – 2002. – Vol. 211. – № 1. – P. 77-83.
168. Escherichia coli O157:H7 typing phage V7 is a T4-like virus / A. M. Kropinski, E. J. Lingohr, D. M. Moyles [et al.] // Journal of Virology. – 2012. – Vol. 86. – № 18. – P. 10246.
169. Whole-Genome Sequences of Phages p000v and p000y, Which Infect the Bacterial Pathogen Shiga-Toxigenic Escherichia coli / C. Howard-Varona, D. R. Vik, N. E. Solonenko [et al.] // Microbiology Resource Announcements. – 2018. – Vol. 7. – № 20. – P. e01400-18.
170. Genome Sequence of a T4-like Phage, Escherichia Phage vB_EcoM-Sa45lw, Infecting Shiga Toxin-Producing Escherichia coli Strains / Y.-T. Liao, Y. Zhang, A. Salvador, V. C. H. Wu // Microbiology Resource Announcements. – 2019. – Vol. 8. – № 32. – P. e00804-19.
171. T4-like phage Bp7, a potential antimicrobial agent for controlling drug-resistant Escherichia coli in chickens / C. Zhang, W. Li, W. Liu [et al.] // Applied and Environmental Microbiology. – 2013. – Vol. 79. – № 18. – P. 5559-5565.
172. Isolation and Characterization of Lytic Phage vB_EcoM_JS09 against Clinically Isolated Antibiotic-Resistant Avian Pathogenic Escherichia coli and Enterotoxigenic Escherichia coli / Y. Zhou, H. Bao, H. Zhang, R. Wang // Intervirology. – 2015. – Vol. 58. – № 4. – P. 218-231.
173. Genome Sequence of the Enterohemorrhagic Escherichia coli Bacteriophage UFV-AREG1 / M. E. S. Lopez, L. S. Batalha, P. M. P. Vidigal [et al.] // Genome Announcements. – 2016. – Vol. 4. – № 5. – P. e00412-16.
174. Lytic bacteriophages against multidrug-resistant Staphylococcus aureus, Enterococcus faecalis and Escherichia coli isolates from orthopaedic implant-associated infections / J. Barros, L. D. R. Melo, P. Poeta [et al.] // International Journal of Antimicrobial Agents. – 2019. – Vol. 54. – № 3. – P. 329-337.
175. Bacteriophages from ExPEC Reservoirs Kill Pandemic Multidrug-Resistant Strains of Clonal Group ST131 in Animal Models of Bacteremia / S. I. Green, J. T. Kaelber, L. Ma [et al.] // Scientific Reports. – 2017. – Vol. 7. – P. 46151.

176. Oral T4-like phage cocktail application to healthy adult volunteers from Bangladesh / S. A. Sarker, S. McCallin, C. Barretto [et al.] // *Virology*. – 2012. – Vol. 434. – № 2. – P. 222-232.
177. Encapsulation and Delivery of Therapeutic Phages / B. Loh, V. S. Gondil, P. Manohar [et al.] // *Applied and Environmental Microbiology*. – 2021. – Vol. 87. – № 5. – P. e01979-20, AEM.01979-20.
178. Moghtader F. Phages in modified alginate beads / F. Moghtader, S. Eđri, E. Piskin // *Artificial Cells, Nanomedicine, and Biotechnology*. – 2017. – Vol. 45. – № 2. – P. 357-363.
179. Encapsulation in alginate-polymers improves stability and allows controlled release of the UFV-AREG1 bacteriophage / L. Silva Batalha, M. T. Pardini Gontijo, A. Vianna Novaes de Carvalho Teixeira [et al.] // *Food Research International* (Ottawa, Ont.). – 2021. – Vol. 139. – P. 109947.
180. Encapsulation of bacteriophage T4 in mannitol-alginate dry macrospheres and survival in simulated gastrointestinal conditions / P. Śliwka, P. Mituła, A. Mituła [et al.] // *LWT*. – 2019. – Vol. 99. – P. 238-243.
181. Oral delivery systems for encapsulated bacteriophages targeted at *Escherichia coli* O157:H7 in feedlot cattle / K. Stanford, T. A. McAllister, Y. D. Niu [et al.] // *Journal of Food Protection*. – 2010. – Vol. 73. – № 7. – P. 1304-1312.
182. In vitro and in vivo bacteriolytic activities of *Escherichia coli* phages: implications for phage therapy / S. Chibani-Chennoufi, J. Sidoti, A. Bruttin [et al.] // *Antimicrobial Agents and Chemotherapy*. – 2004. – Vol. 48. – № 7. – P. 2558-2569.
183. In vivo replication of T4 and T7 bacteriophages in germ-free mice colonized with *Escherichia coli* / M. Weiss, E. Denou, A. Bruttin [et al.] // *Virology*. – 2009. – Vol. 393. – № 1. – P. 16-23.
184. Bruttin A. Human volunteers receiving *Escherichia coli* phage T4 orally: a safety test of phage therapy / A. Bruttin, H. Brüssow // *Antimicrobial Agents and Chemotherapy*. – 2005. – Vol. 49. – № 7. – P. 2874-2878.
185. Phage Therapy Related Microbial Succession Associated with Successful Clinical Outcome for a Recurrent Urinary Tract Infection / A. Terwilliger, J. Clark, M. Karris [et al.] // *Viruses*. – 2021. – Vol. 13. – № 10. – P. 2049.
186. Tanyashin V. I. The Cotransduction of pET System Plasmids by Mutants of T4 and RB43 Bacteriophages / V. I. Tanyashin, A. A. Zimin, A. M. Boronin // *Microbiology*. – 2003. – Vol. 72. – № 6. – P. 694-700.

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. Кощев, 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

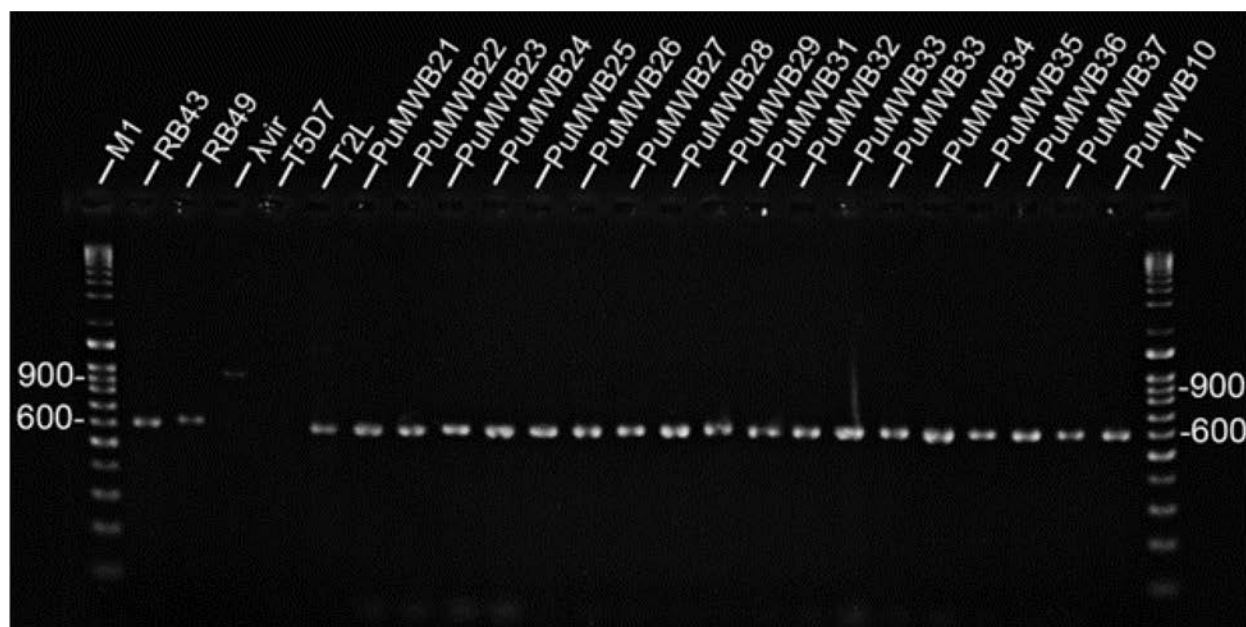
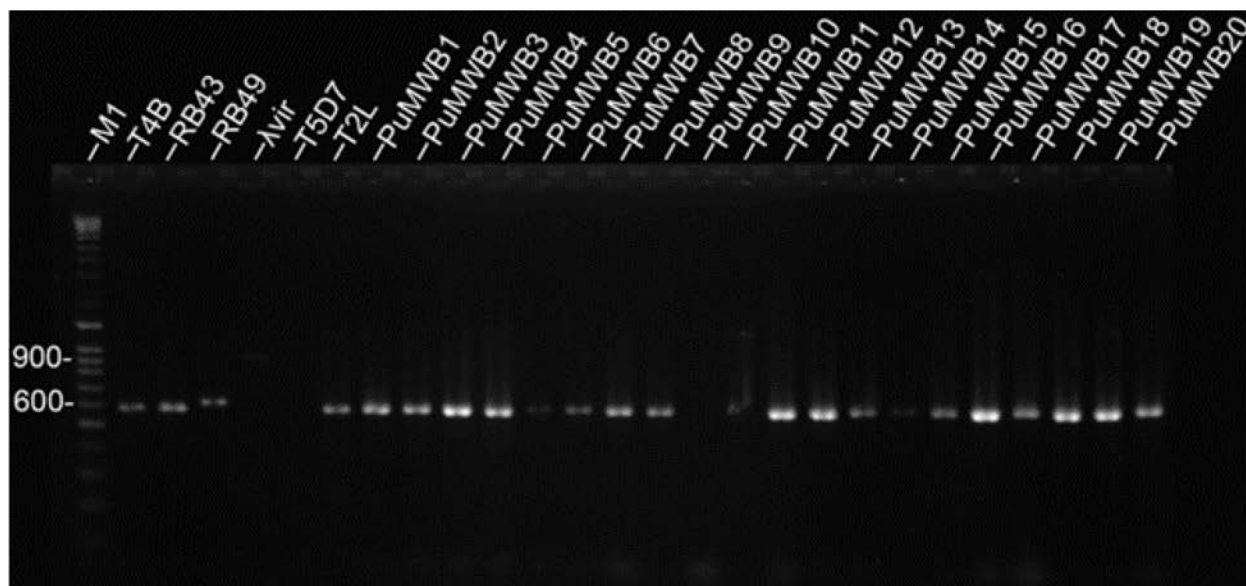
Публикации в материалах научных конференций:

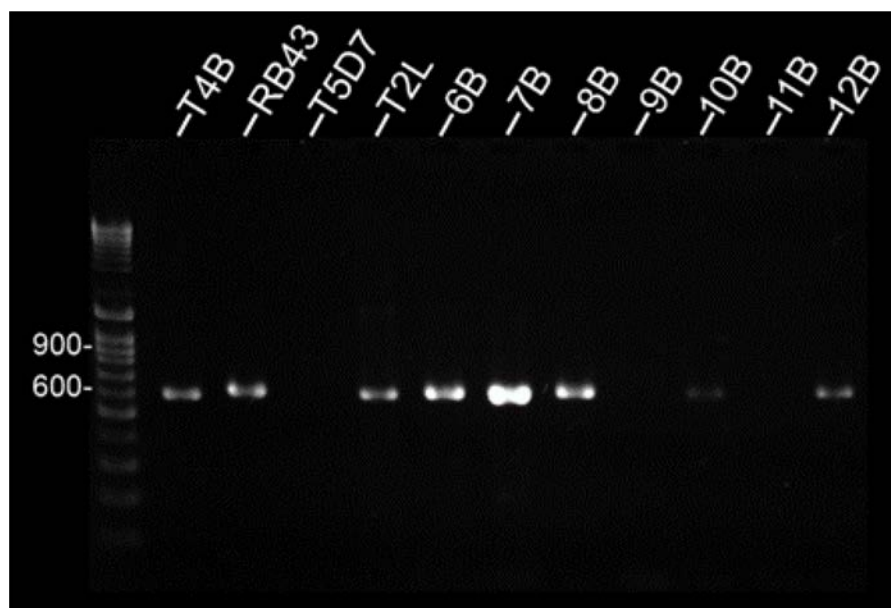
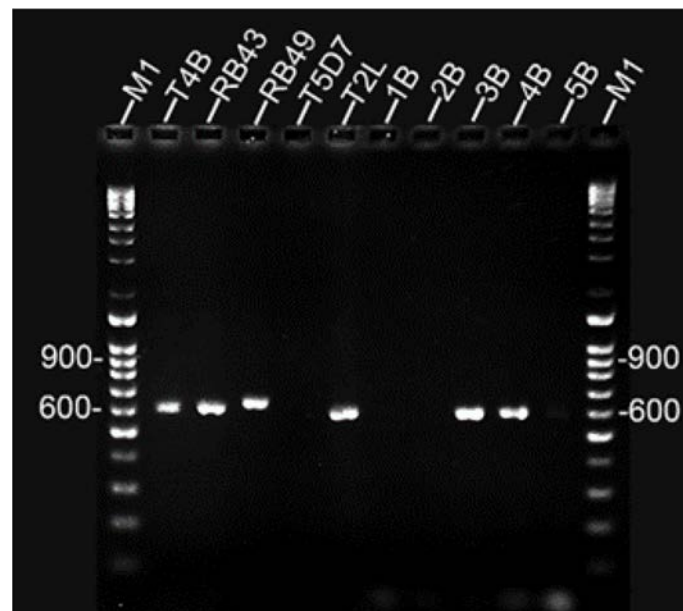
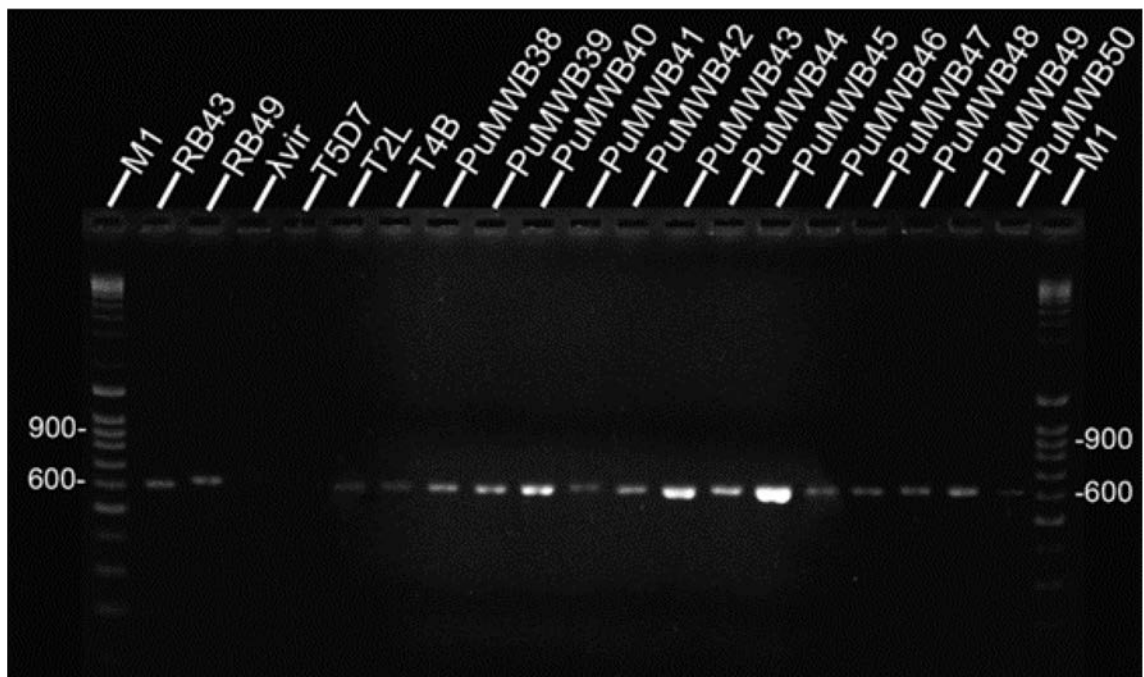
1. Никулин Н. А. Бактериофаговая флора очистных сооружений города Пущино Московской области / Н. А. Никулин, А. А. Зимин // в книге: Биохимия,

- физиология и биосферная роль микроорганизмов. Сборник тезисов III Пущинской 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, DH1, B, BL21 / 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.

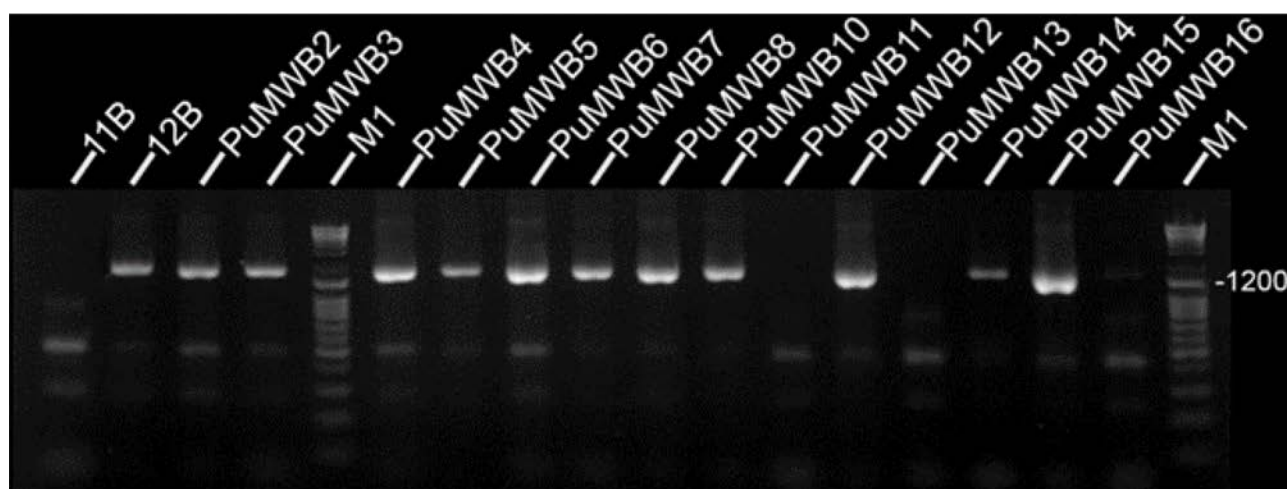
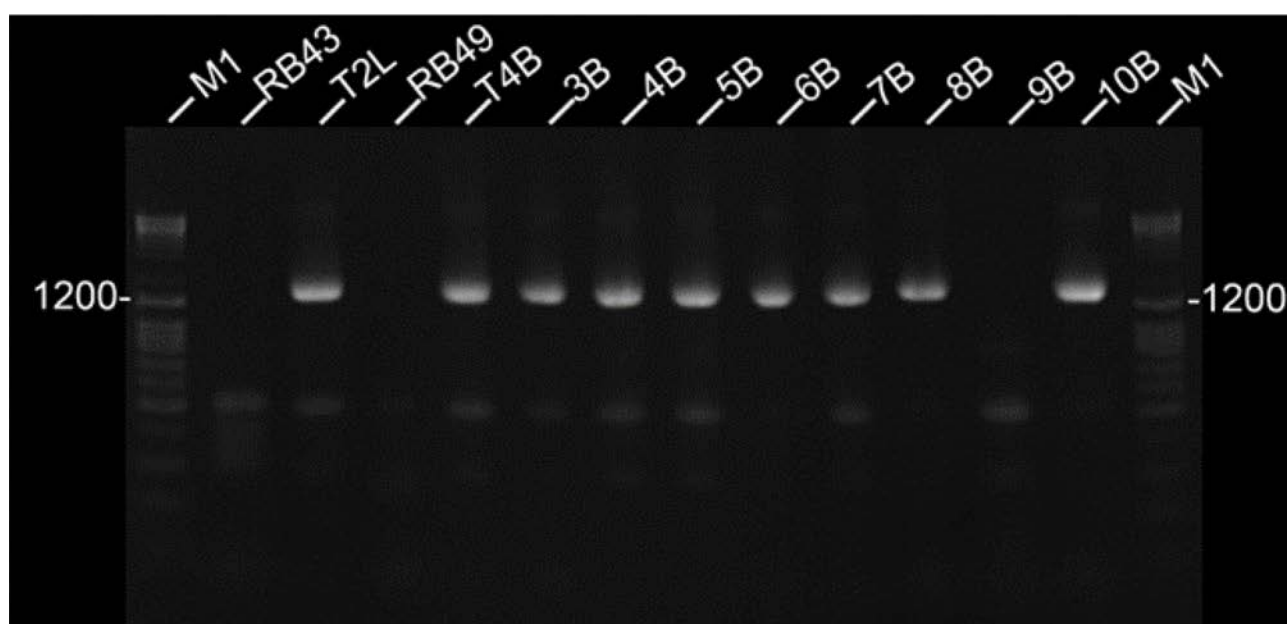
ПРИЛОЖЕНИЕ 1

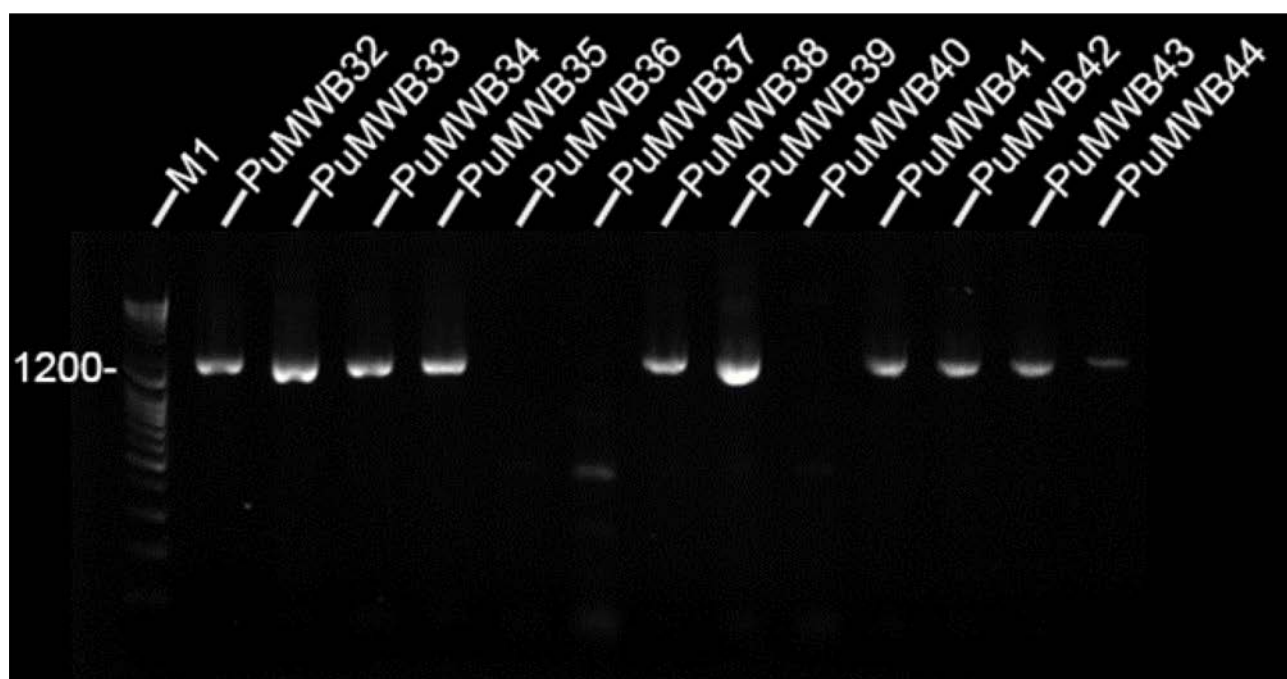
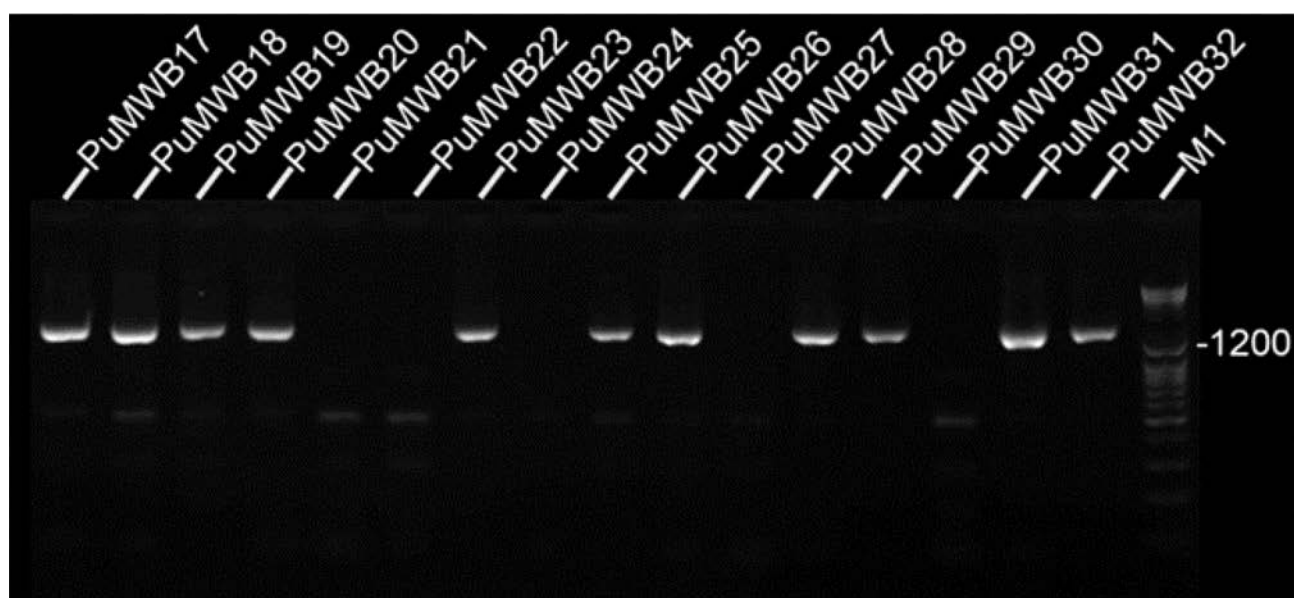
Электрофореграммы результатов ПЦР с праймерами MZIA1bis и MZIA6, которые отжигаются на последовательностях гена главного белка капсида Т4-родственных бактериофагов





T4B, T2L, RB43, RB49 – лабораторные штаммы T4-родственных бактериофагов; T5D7, λ vir – лабораторные штаммы фагов, не принадлежащие к T4-родственным; nB – бактериофаги рабочей коллекции, выделенные из фекалий зубров, где n – порядковый номер; PuMWBn – бактериофаги рабочей коллекции, выделенные из сточных вод, где n – порядковый номер. M1 – ДНК-маркер; 600, 900 – размеры фрагментов ДНК в пар оснований. 600 п.о. – размер целевого ампликона. Электрофореграммы результатов ПЦР с праймерами *hocfR* и *inhrc*, которые специфически отжигаются на консервативной последовательности фрагмента генома *Tequatrovirus T4*





T4В, T2L – лабораторные штаммы фагов, близкородственных *Tequatrovirus* T4; RB43, RB49 – лабораторные штаммы T4-родственных бактериофагов, не являющихся близкородственными *Tequatrovirus* T4. M1 – ДНК-маркер; 1200 – размер фрагмента ДНК в пар оснований. 1276 п.о. – размер целевого ампликона.

ПРИЛОЖЕНИЕ 2

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

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» – рост на изогенных штаммах с и без RM-системы одинаковый; «ограничен» – рост на изогенных штаммах с и без RM-системы отличается; «отсутствует» – рост на штамме с RM-системой отсутствует, без – присутствует; «неизвестно» – рост на изогенных штаммах с и без RM-системы отсутствует.

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
λ vir	-	-	отсутствует	отсутствует
1B	-	-	отсутствует	отсутствует
2B	-	-	отсутствует	отсутствует
3B	+	+	не ограничен	не ограничен
4B	+	+	не ограничен	не ограничен
5B	+	+	не ограничен	не ограничен
6B	+	+	не ограничен	не ограничен
7B	+	+	не ограничен	не ограничен
8B	+	+	не ограничен	не ограничен
9B	-	-	отсутствует	отсутствует
10B	+	+	не ограничен	не ограничен
11B	-	-	отсутствует	отсутствует
12B	+	+	не ограничен	не ограничен
PuMWB1	+	+	не ограничен	не ограничен
PuMWB2	+	+	не ограничен	не ограничен
PuMWB3	+	+	не ограничен	не ограничен
PuMWB4	+	+	не ограничен	не ограничен
PuMWB5	+	+	не ограничен	не ограничен
PuMWB6	+	+	не ограничен	не ограничен
PuMWB7	+	+	не ограничен	не ограничен

[illegible]

ПРИЛОЖЕНИЕ 3

Таблица использованных геномов Т4-родственных бактериофагов из баз данных

Номер доступа	Название фага	Чувствительные бактерии	Известные данные по источнику выделения фага
NC_019398.1	Cronobacter phage vB_CsaM_GAP161	<i>Cr. sakazakii</i>	сточные воды, Гвельф, Канада
KX431559.1	Cronobacter phage vB_CsaM_leB	<i>Cr. sakazakii</i>	навозная жижа, животноводческая ферма в Клонакилти, Ирландия
KX431560.1	Cronobacter phage vB_CsaM_leN	<i>Cr. sakazakii</i>	навозная жижа, животноводческая ферма в Клонакилти, Ирландия
MG999954.1	Enterobacter phage myPSH1140	<i>E. cloacae</i>	очистные сооружения сточных вод, Индия
NC_023561.1	Enterobacter phage PG7	<i>E. cloacae</i>	вода из пруда, Китай
NC_041980.1	Enterobacter phage phiEap-3	<i>E. aerogenes</i>	сточные воды, Военно-морской госпиталь, Китай
KT184308.1	Enterobacteria phage Aplg8	<i>E.coli</i> O121:H19	сточные воды, США
NC_027983.1	Enterobacteria phage AR1 DNA	<i>E.coli</i> O157:H7	свежий навоз, молочные фермы в округе Санта-Крус, США
KT184309.1	Enterobacteria phage ATK47	<i>E.coli</i> O26:H11	сточные воды, США
KT184310.1	Enterobacteria phage ATK48	<i>E.coli</i> O26:H12	сточные воды, США
NC_019500.1	Enterobacteria phage Bp7	<i>E.coli</i> O78.A	куриные фекалии, провинция Шаньдун, Китай
NC_014662.1	Enterobacteria phage CC31	<i>E.coli</i> B strain S/6/4	сточные воды, Новый Орлеан, США
NC_025425.1	Enterobacteria phage GEC-3S	<i>E.coli</i> O104:H4	библиотека терапевтических фагов, Институт бактериофагов Элиавы, Грузия
KT184311.1	Enterobacteria phage GiZh	<i>E.coli</i> O111:NM 96-3166	сточные воды, США
NC_018855.1	Enterobacteria phage HX01	<i>E.coli</i> DM01	утиные фекалии, утиная фабрика, утиные фекалии, утиная фабрика, Китай
NC_014260.1	Enterobacteria phage IME08	<i>E.coli</i> strain 8099	больничные сточные воды, Китай
NC_012741.1	Enterobacteria phage JS10	<i>E.coli</i> K12	фекалии пациента с диареей, Бангладеш
NC_010105.1	Enterobacteria phage JS98	<i>E.coli</i>	фекалии пациента с диареей, Бангладеш
NC_012740.1	Enterobacteria phage JSE	<i>E.coli</i> K12	Швейцария
KT184312.1	Enterobacteria phage Kha5h	<i>E.coli</i> O103:H2 90-3128	сточные воды, США
NC_009821.1	Enterobacteria phage Phi1	<i>E.coli</i> K-12 (F+)	библиотека терапевтических фагов, Институт бактериофагов Элиавы, Грузия
NC_028847.1	Enterobacteria phage QL01	<i>E.coli</i> DE205B	утиные фекалии, птичий рынок в Нанкине, Китай
KM606999.1	Enterobacteria phage RB10	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США

Номер доступа	Название фага	Чувствительные бактерии	Известные данные по источнику выделения фага
NC_012638.1	Enterobacteria phage RB14	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_014467.1	Enterobacteria phage RB16	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
MH553563.1	Enterobacteria phage RB18	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_025448.1	Enterobacteria phage RB27	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_025419.1	Enterobacteria phage RB3	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
KM607001.1	Enterobacteria phage RB33	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_007023.1	Enterobacteria phage RB43	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_005066.1	Enterobacteria phage RB49	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
KM606995.1	Enterobacteria phage RB5	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_012635.1	Enterobacteria phage RB51	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
KM607002.1	Enterobacteria phage RB55	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
KM607003.1	Enterobacteria phage RB59	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
KM606996.1	Enterobacteria phage RB6	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_027979.1	Enterobacteria phage RB68	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_004928.1	Enterobacteria phage RB69	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
KM606997.1	Enterobacteria phage RB7	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
KM606998.1	Enterobacteria phage RB9	<i>E.coli</i>	очистные сооружения, Лонг-Айленд, США
NC_000866.4	Enterobacteria phage T4	<i>E.coli</i>	сточные воды или фекалии
KJ477684.1	Enterobacteria phage T4 strain wild	<i>E.coli</i>	сточные воды или фекалии
HM137666.1	Enterobacteria phage T4T	<i>E.coli</i>	сточные воды или фекалии
MH550421.1	Enterobacteria phage T6	<i>E.coli</i>	сточные воды или фекалии
NC_019399.1	Enterobacteria phage vB_EcoM_ACG-C40	<i>E.coli</i>	сточные воды, водоочистная станция Гуэлфа, Канада
MH051915.1	Enterobacteria phage vB_EcoM_IME339	<i>E.coli</i> BL21(DE3)	сточные воды, Китай
MH051916.1	Enterobacteria phage vB_EcoM_IME340	<i>E.coli</i> BL21(DE3)	сточные воды, Китай
NC_028894.1	Enterobacteria phage vB_EcoM_VR20	<i>E.coli</i> Be	муниципальные сточные воды, Литва
NC_028925.1	Enterobacteria phage vB_EcoM_VR25	<i>E.coli</i> MH1	муниципальные сточные воды, Литва
NC_028957.1	Enterobacteria phage vB_EcoM_VR26	<i>E.coli</i> MH1	муниципальные сточные воды, Литва
NC_028881.1	Enterobacteria phage vB_EcoM_VR5	<i>E.coli</i> Be	сточные воды, Литва
NC_014792.1	Enterobacteria phage vB_EcoM-VR7	<i>E.coli</i> Be	сточные воды, Литва
MH059636.2	Erwinia phage Cronus	<i>Er. amylovora</i> DSM17948	органические отходы, Дания

Номер доступа	Название фага	Чувствительные бактерии	Известные данные по источнику выделения фага
NC_041863.1	<i>Escherichia coli</i> O157 typing phage 3	<i>E.coli</i> O157:H7	Национальная лаборатория микробиологии, Виннипег
NC_041864.1	<i>Escherichia coli</i> O157 typing phage 6	<i>E.coli</i> O157:H7	Национальная лаборатория микробиологии, Виннипег
MK234886.1	Escherichia phage AnYang	<i>E.coli</i> O157	почва, Китай
NC_029091.1	Escherichia phage APCEc01	<i>E.coli</i> DPC6051	человеческие фекалии, Ирландия
NC_041919.1	Escherichia phage CF2	<i>E.coli</i> CP9	куриные фекалии, частная ферма в Техасе, США
MK327929.1	Escherichia phage D5505	<i>E.coli</i> DSM101114	поверхностные воды, Брауншвейг, Германия
NC_024125.2	Escherichia phage e11/2	<i>E.coli</i> O157:H7	бычьи фекалии, Ирландия
NC_041936.1	Escherichia phage ECD7	<i>E.coli</i> O104:H4	куриные фекалии, Московская область, Россия
NC_025449.1	Escherichia phage ECML-134	<i>E.coli</i> O157:H7	пресная и соленая вода, эстуарий, Грузия
MH791409.1	Escherichia phage EcWhh-1	<i>E.coli</i>	чистая вода
NC_041920.1	Escherichia phage HP3	<i>E.coli</i> CP9	утиные/гусиные фекалии, парки Хьюстона, США
NC_027349.1	Escherichia phage HY01	<i>E.coli</i> O157:H7	свиные фекалии, ферма Сеульского национального университета в Сувоне, Южная Корея
NC_031047.1	Escherichia phage HY03	<i>Escherichia coli</i> O157:H7	-
AP018932.1	Escherichia phage KIT03	<i>Escherichia coli</i> O157:H7	почва, Киотская птицефабрика, Япония
NC_021344.2	Escherichia phage Lw1	<i>Escherichia coli</i> BL21(DE3)	загрязнение лизированных клеток
NC_031934.1	Escherichia phage MX01	<i>Escherichia coli</i> DE217	утиные фекалии
MH992122.1	Escherichia phage OLB35	<i>Escherichia coli</i>	сточные воды
MK047717.1	Escherichia phage p000v	<i>Escherichia coli</i> O157:H7	сточные воды, Сан-Франциско, Калифорния
MK047718.1	Escherichia phage p000y	<i>Escherichia coli</i> O157:H7	сточные воды, Сан-Франциско, Калифорния
KU925172.1	Escherichia phage PE37	<i>Escherichia coli</i> O157:H7	бычий кишечник, розничные магазины в Фукуоке, Япония
LC348379.1	Escherichia phage PP01	<i>Escherichia coli</i> O157:H7	свиные фекалии, Япония
MH359124.1	Escherichia phage SF	<i>Escherichia coli</i> O157:H7	сточные воды, кампус Научного университета Малайзии в Пенанге, Малайзия
NC_028927.1	Escherichia phage slur02	<i>Escherichia coli</i> MG1655	навозная жижа, резервуар для навозной жижи молочной фермы в Ист-Мидлендсе, Великобритания
NC_042129.1	Escherichia phage slur03	<i>Escherichia coli</i> MG1655	навозная жижа, резервуар для навозной жижи молочной фермы в Ист-Мидлендсе, Великобритания
NC_042130.1	Escherichia phage slur04	<i>Escherichia coli</i> MG1655	навозная жижа, резервуар для навозной жижи молочной фермы в Ист-Мидлендсе, Великобритания
NC_028780.1	Escherichia phage slur07	<i>Escherichia coli</i> MG1655	навозная жижа, резервуар для навозной жижи молочной фермы в Ист-Мидлендсе, Великобритания

Номер доступа	Название фага	Чувствительные бактерии	Известные данные по источнику выделения фага
LN881733.1	Escherichia phage slur08	<i>Escherichia coli</i> MG1655	навозная жижа, резервуар для навозной жижи молочной фермы в Ист-Мидлендсе, Великобритания
LN881734.1	Escherichia phage slur11	<i>Escherichia coli</i> MG1655	навозная жижа, резервуар для навозной жижи молочной фермы в Ист-Мидлендсе, Великобритания
LN881737.1	Escherichia phage slur13	<i>Escherichia coli</i> MG1655	навозная жижа, резервуар для навозной жижи молочной фермы в Ист-Мидлендсе, Великобритания
NC_028448.1	Escherichia phage slur14	<i>Escherichia coli</i> MG1655	навозная жижа, резервуар для навозной жижи молочной фермы в Ист-Мидлендсе, Великобритания
NC_041990.1	Escherichia phage ST0	<i>Escherichia coli</i> H8	сточные воды, Пекинская станция очистки сточных вод, Китай
LC348380.1	Escherichia phage T2	<i>E.coli</i>	сточные воды или фекалии
NC_031030.1	Escherichia phage UFV-AREG1	<i>Escherichia coli</i> O157	сточные воды коровника, Бразилия
MH992510.1	Escherichia phage vB_EcoM_DalCa	<i>E.coli</i>	неочищенные сточные воды, очистные сооружения Фроссос, Португалия
MK327937.1	Escherichia phage vB_EcoM_G10400	<i>E.coli</i> DSM 103266	свиной навоз, Хамельн, Германия
MK327932.1	Escherichia phage vB_EcoM_G2248	<i>E.coli</i> DSM 103255	свиной навоз, Хамельн, Германия
MK327933.1	Escherichia phage vB_EcoM_G2285	<i>E.coli</i> DSM 103256	свиной навоз, Хамельн, Германия
MK327934.1	Escherichia phage vB_EcoM_G2469	<i>E.coli</i> DSM 103258	свиной навоз, Хамельн, Германия
MK327935.1	Escherichia phage vB_EcoM_G2494	<i>E.coli</i> DSM 103259	свиной навоз, Хамельн, Германия
MK327936.1	Escherichia phage vB_EcoM_G2540	<i>E.coli</i> DSM 103260	свиной навоз, Хамельн, Германия
MK327944.1	Escherichia phage vB_EcoM_G2540-3	<i>E.coli</i> DSM 103260	свиной навоз, Хамельн, Германия
MK327940.1	Escherichia phage vB_EcoM_G29	<i>E.coli</i> DSM 103247	свиной навоз, Хамельн, Германия
MK327941.1	Escherichia phage vB_EcoM_G37-3	<i>E.coli</i> DSM 103248	свиной навоз, Хамельн, Германия
MK327939.1	Escherichia phage vB_EcoM_G4498	<i>E.coli</i> DSM 103261	свиной навоз, Хамельн, Германия
MK327945.1	Escherichia phage vB_EcoM_G4500	<i>E.coli</i> DSM 103262	свиной навоз, Хамельн, Германия
MK327946.1	Escherichia phage vB_EcoM_G4507	<i>E.coli</i> DSM 103263	свиной навоз, Хамельн, Германия
MK327942.1	Escherichia phage vB_EcoM_G50	<i>E.coli</i> DSM 103250	свиной навоз, Хамельн, Германия
MK327947.1	Escherichia phage vB_EcoM_G5211	<i>E.coli</i> DSM 103264	свиной навоз, Хамельн, Германия
MK327943.1	Escherichia phage vB_EcoM_G53	<i>E.coli</i> DSM 103251	свиной навоз, Хамельн, Германия
MK373787.1	Escherichia phage vB_EcoM_G8	<i>E.coli</i> DSM 103243	свиной навоз, Хамельн, Германия
MK373779.1	Escherichia phage vB_EcoM_G9062	<i>E.coli</i> DSM 103265	свиной навоз, Хамельн, Германия
MH355584.1	Escherichia phage vB_EcoM_JB75	<i>E.coli</i>	неочищенные сточные воды, очистные сооружения Фроссос, Португалия

Номер доступа	Название фага	Чувствительные бактерии	Известные данные по источнику выделения фага
NC_024124.2	Escherichia phage vB_EcoM_JS09	<i>Escherichia coli</i> (APEC)	сточные воды свинофермы в провинции Цзянсу, Китай
MK373781.1	Escherichia phage vB_EcoM_KAW1E185	<i>E.coli</i> DSM 103253	клинические сточные воды, Брауншвейг, Германия
MK373782.1	Escherichia phage vB_EcoM_KAW3E185	<i>E.coli</i> DSM 103253	клинические сточные воды, Брауншвейг, Германия
MK373784.1	Escherichia phage vB_EcoM_MM02	<i>E.coli</i> DSM 498	утиные фекалии, Брауншвейг, Германия
MK373785.1	Escherichia phage vB_EcoM_OE5505	<i>E.coli</i> DSM 101114	поверхностные воды, Брауншвейг, Германия
NC_024794.1	Escherichia phage vB_EcoM_PhAPEC2	<i>Escherichia coli</i> (APEC)	вода, реки и ручьи в Брюсселе и его окрестностях вблизи птичников, Бельгия
MK373786.1	Escherichia phage vB_EcoM_R5505	<i>E.coli</i> DSM 101114	поверхностные воды, Брауншвейг, Германия
MK373778.1	Escherichia phage vB_EcoM_WFbE185	<i>E.coli</i> DSM 103253	канализация, Вольфенбюттель, Германия
MK373775.1	Escherichia phage vB_EcoM_WFK	<i>E.coli</i> DSM 101124	канализация, Вольфенбюттель, Германия
MK373774.1	Escherichia phage vB_EcoM_WFL6982	<i>E.coli</i> DSM 101124	канализация, Вольфенбюттель, Германия
MG781190.1	Escherichia phage vB_EcoM-fFiEco06	<i>Escherichia coli</i> 123738	муниципальные сточные воды, Финляндия
MG781191.1	Escherichia phage vB_EcoM-fHoEco02	<i>Escherichia coli</i> 123738	больничные сточные воды, Хельсинки, Финляндия
MG867727.1	Escherichia phage vB_EcoM-G28	<i>E.coli</i>	навоз, Германия
MK977694.1	Escherichia phage vB_EcoM- Sa45lw	<i>Escherichia coli</i> ATCC 13706	поверхностные воды, район выращивания сельскохозяйственных культур Салинаса, США
NC_031103.1	Escherichia phage vB_EcoM- UFV13	<i>Escherichia coli</i> UFV30	сточные воды канализации Викоzy, Бразилия
NC_031928.1	Escherichia phage WG01	<i>Escherichia coli</i> DE017	утиные фекалии, птичий рынок в Нанкине, Китай
NC_019505.1	Escherichia phage wV7	<i>Escherichia coli</i> O157:H7	человеческие фекалии/мясо для гамбургеров
MH560358.1	Escherichia virus KFS-EC	<i>Escherichia coli</i> O157:H7	сточные воды, Южная Корея
LR027390.1	Escherichia virus vB_Eco_mar005P1	<i>E.coli</i> MG1655	морская вода, Грейт-Ярмут, Великобритания
LR027387.1	Escherichia virus vB_Eco_mar005P1 strain vB_Eco_mar006P2	<i>E.coli</i> MG1655	морская вода, Грейт-Ярмут, Великобритания
LR027383.1	Escherichia virus vB_Eco_mar005P1 strain vB_Eco_mar007P3	<i>E.coli</i> MG1655	морская вода, Грейт-Ярмут, Великобритания
LR027386.1	Escherichia virus vB_Eco_mar005P1 strain vB_Eco_mar008P4	<i>E.coli</i> MG1655	морская вода, Грейт-Ярмут, Великобритания
LR027391.1	Escherichia virus vB_Eco_mar005P1 strain vB_Eco_mar009P5	<i>E.coli</i> MG1655	морская вода, Грейт-Ярмут, Великобритания
MN022785.1	Escherichia virus VEc20	<i>Escherichia coli</i>	фекалии и сточные воды птицефабрики, Россия
MG751100.1	Klebsiella phage KP1	<i>Klebsiella pneumoniae</i>	сточные воды, Чхунчхон-Пукто, Южная Корея
KY000080.1	Klebsiella phage KPV15	<i>Klebsiella pneumoniae</i>	клинический материал, Россия

Номер доступа	Название фага	Чувствительные бактерии	Известные данные по источнику выделения фага
MN044033.1	Klebsiella phage Marfa	<i>K. pneumoniae</i> 1776c	свиные фекалии, Техас и Мичиган, США
NC_028750.1	Klebsiella phage Matisse	<i>Klebsiella pneumoniae</i>	сточные воды, Колледж-Стейшн, США
MH333064.1	Klebsiella phage Mineola	<i>K. pneumoniae</i>	активный ил, городские сточные сооружения Брайана, США
NC_041981.1	Klebsiella phage Miro	<i>K. pneumoniae</i> A1	сточные воды, Колледж-Стейшн, США
NC_031095.1	Klebsiella phage PKO111	Klebsiella oxytoca ATCC 43863	сточные воды, очистные сооружения Сувоны, Южная Корея
MG746602.1	Klebsiella phage vB_Kpn_F48	<i>K. pneumoniae</i> 12C47	сточные воды, университетская больница Флоренс Кареджи, Италия
NC_031087.1	Klebsiella phage vB_KpnM_KpV477	<i>Klebsiella pneumoniae</i> KPB463	клинический образец, Институт нейрохирургии им. Бурденко, Россия
KX078569.1	Morganella phage vB_MmoM_MP1	<i>M. morganii</i>	сточные воды, очистные сооружения Браги, Португалия
NC_028940.1	Pectobacterium bacteriophage PM2	<i>P. carotovorum</i> subsp. <i>carotovorum</i>	почва, поля пекинской капусты в Пхёнчхане, Южная Корея
MK310184.1	Phage NC-G	<i>Escherichia coli</i> NC101	фекалии человека, лица с диагнозом болезнь Крона, США
MG696114.1	Proteus phage phiP4-3	<i>Proteus penneri</i>	рыба, Китай
NC_028762.1	Proteus phage vB_PmiM_Pm5461	<i>Proteus mirabilis</i> 417	сточные воды очистных сооружений, Португалия
KY971610.1	Pseudomonas phage PspYZU05	<i>Pseudomonas</i> sp.	сточные воды
NC_042044.1	Salmonella phage Melville	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport	сточные воды, Остин, США
NC_020416.1	Salmonella phage S16	<i>S. enterica</i>	сточные воды очистных сооружений Гелдерланда, Нидерланды
NC_027344.1	Salmonella phage STML-198	Salmonella	поверхностные воды, Мэриленд, США
NC_026607.2	Salmonella phage STP4-a	Salmonella typhimurium ATCC 14028	сточные воды, Циндао, Китай
NC_031065.1	Salmonella phage vB_SnwM_CGG4-1	Salmonella enterica subsp. <i>enterica</i> serovar Newport	сточные воды, очистных сооружений Гвельфа, Канада
NC_024121.1	Serratia phage PS2	<i>S. marcescens</i> S2	вода из пруда
NC_025829.1	Shigella phage pSs-1	<i>Shigella sonnei</i> ATCC 25931	вода, ручей Хончечхон, Южная Корея
MF327006.1	Shigella phage Sf20	<i>Shigella boydii</i> ; <i>Shigella flexneri</i>	вода, Ред Сидар Ривер, США
NC_042077.1	Shigella phage Sf21	<i>Shigella flexneri</i> ; <i>Shigella sonnei</i> ; <i>E.coli</i> K12	вода, Ред Сидар Ривер, США
NC_042039.1	Shigella phage Sf22	<i>Shigella flexneri</i> , <i>Shigella boydii</i> , <i>Shigella dysenteriae</i> , <i>Shigella sonnei</i>	биоупленка с буя посреди реки Гранд, США
MF158046.1	Shigella phage Sf23	<i>Shigella flexneri</i> , <i>Shigella boydii</i> , <i>Shigella dysenteriae</i>	биоупленка с буя посреди реки Гранд, США

Номер доступа	Название фага	Чувствительные бактерии	Известные данные по источнику выделения фага
NC_042078.1	Shigella phage Sf24	<i>Shigella flexneri</i> ; <i>Shigella sonnei</i> ; <i>E.coli</i> K12	вода, Ред Сидар Ривер, США
MF327009.1	Shigella phage Sf25	<i>Shigella flexneri</i> ; <i>Shigella sonnei</i> ; <i>E.coli</i> K12, <i>Shigella dysenteriae</i>	вода, Ред Сидар Ривер, США
KX828711.1	Shigella phage SH7	<i>Shigella flexneri</i>	вода вади Чотрана, Тунис
NC_031085.1	Shigella phage SHBML-50-1	<i>Shigella sonnei</i>	поверхностные воды, парк Мэриленд, США
NC_015457.1	Shigella phage Shf12	<i>Shigella flexneri</i> 37	сточные воды, Бразилия
NC_030953.1	Shigella phage SHFML-11	<i>Shigella sonnei</i>	библиотека терапевтических фагов, Институт бактериофагов Элиавы, Грузия
NC_031011.1	Shigella phage SHFML-26	<i>Shigella sonnei</i>	библиотека терапевтических фагов, Институт бактериофагов Элиавы, Грузия
NC_031090.1	Shigella phage SHSML-52-1	<i>Shigella sonnei</i>	поверхностные воды, парк Мэриленд, США
NC_014595.1	Shigella phage SP18	<i>Shigella sonnei</i>	вода реки Гэп, Южная Корея
NC_005083.2	Vibrio phage KVP40	<i>Vibrio parahaemolyticus</i>	загрязненная морская вода залива Урадо, Япония
NC_021529.2	Vibrio phage nt-1	<i>Vibrio natriegens</i>	грязь солончака, округ Глостер, США
KT919972.1	Vibrio phage phi-Grn1	<i>Vibrio alginolyticus</i>	прибрежная морская вода, Крит, Греция
KT919973.1	Vibrio phage phi-ST2	<i>Vibrio alginolyticus</i>	прибрежная морская вода, Крит, Греция
MK568540.1	Vibrio phage ValB1_HC	<i>Vibrio alginolyticus</i>	-
NC_028829.1	Vibrio phage ValKK3	<i>Vibrio alginolyticus</i>	морские отложения, Кота-Кинабалу, Малайзия
KU160494.1	Vibrio phage vB_VmeM-32	<i>Vibrio metschnikovii</i> DSM 29715	проба отложений, озеро Эльментейта, Кения
NC_023568.1	Vibrio phage VH7D	<i>Vibrio harveyi</i>	морская вода, ферма по выращиванию морских ушек, Сямэнь, Китай
JN849462.1	Vibriophage phi-pp2	<i>Vibrio parahaemolyticus</i>	вода, водные пути аквакультуры вокруг южного Тайваня
LR215722.1	Yersinia phage fPS-2	<i>Y. pseudotuberculosis</i> O:1a	фекалии свиней, свинофермы Финляндии
NC_027353.1	Yersinia phage phiD1	<i>Yersinia pestis</i>	-
NC_027404.1	Yersinia phage PST	<i>Yersinia pseudotuberculosis</i>	-

ПРИЛОЖЕНИЕ 4

Число компонентов пангенома в исследованных геномах

Фаги	Сокращенное название	Облако	Оболочка	Софткор	Кор
_Acinetobacter_phage_133	133	116	78	75	61
_Acinetobacter_phage_AM101	AM101	20	170	73	61
_Acinetobacter_phage_Ac42	Ac42	116	81	74	61
_Acinetobacter_phage_Acj61	Acj61	70	104	73	61
_Acinetobacter_phage_Acj9	Acj9	69	115	73	61
_Acinetobacter_phage_KARL-1	KARL-1	21	172	73	61
_Acinetobacter_phage_ZZ1	ZZ1	69	111	73	61
_Acinetobacter_phage_vB_ApiM_fHyAci03	fHyAci03	20	173	73	61
_Aeromonas_phage_65.2	65.2	89	282	75	61
_Aeromonas_phage_65	65	81	274	73	61
_Aeromonas_phage_AS-szw	AS-szw	21	335	75	61
_Aeromonas_phage_AS-zj	AS-zj	19	335	75	61
_Aeromonas_phage_Aeh1	Aeh1	46	231	76	61
_Aeromonas_phage_AsFcp_2	AsFcp_2	66	227	76	61
_Aeromonas_phage_Asswx_1	Asswx_1	27	336	75	61
_Aeromonas_phage_Aswh_1	Aswh_1	119	245	75	61
_Aeromonas_phage_CC2	CC2	37	339	75	61
_Aeromonas_phage_PX29	PX29	74	231	76	61
_Bacteriophage_RB32	RB32	1	184	76	61
_Citrobacter_phage_IME-CF2	IME-CF2	25	203	75	61
_Citrobacter_phage_Margaery	Margaery	3	199	76	61
_Citrobacter_phage_Maroon	Maroon	1	197	76	61
_Citrobacter_phage_Merlin	Merlin	36	196	76	61
_Citrobacter_phage_Miller	Miller	6	205	75	61
_Citrobacter_phage_Moon	Moon	33	189	76	61
_Citrobacter_phage_vB_CfrM_CfP1	CfP1	3	193	76	61
_Cronobacter_phage_Pet-CM3-4	Pet-CM3-4	34	211	76	61
_Cronobacter_phage_vB_CsaM_GAP161	GAP161	7	190	76	61
_Cronobacter_phage_vB_CsaM_leB	leB	5	198	76	61
_Cronobacter_phage_vB_CsaM_leN	leN	4	204	76	61
_Enterobacter_phage_PG7	PG7	15	204	76	61
_Enterobacter_phage_myPSH1140	myPSH1140	27	213	76	61
_Enterobacter_phage_phiEap-3	phiEap-3	5	196	76	61
_Enterobacteria_phage_AR1	AR1	10	189	76	61
_Enterobacteria_phage_ATK47	ATK47	1	201	76	61
_Enterobacteria_phage_ATK48	ATK48	2	201	76	61
_Enterobacteria_phage_Aplg8	Aplg8	3	204	76	61
_Enterobacteria_phage_Bp7	Bp7	6	173	76	61

_Enterobacteria_phage_CC31	CC31	20	198	75	61
_Enterobacteria_phage_GEC-3S	GEC-3S	5	197	76	61
_Enterobacteria_phage_GiZh	GiZh	3	202	76	61
_Enterobacteria_phage_HX01	HX01	1	185	76	61
_Enterobacteria_phage_IME08	IME08	1	169	76	61
_Enterobacteria_phage_JS10	JS10	2	181	76	61
_Enterobacteria_phage_JS98	JS98	0	183	76	61
_Enterobacteria_phage_JSE	JSE	7	189	76	61
_Enterobacteria_phage_Kha5h	Kha5h	3	200	76	61
_Enterobacteria_phage_Phi1	Phi1	2	196	76	61
_Enterobacteria_phage_QL01	QL01	2	190	76	61
_Enterobacteria_phage_RB10	RB10	0	185	76	61
_Enterobacteria_phage_RB14	RB14	0	191	76	61
_Enterobacteria_phage_RB16	RB16	10	190	75	61
_Enterobacteria_phage_RB18	RB18	0	190	76	61
_Enterobacteria_phage_RB27	RB27	1	187	76	61
_Enterobacteria_phage_RB33	RB33	2	187	76	61
_Enterobacteria_phage_RB3	RB3	0	186	76	61
_Enterobacteria_phage_RB43	RB43	44	202	75	61
_Enterobacteria_phage_RB49	RB49	6	196	76	61
_Enterobacteria_phage_RB51	RB51	0	188	76	61
_Enterobacteria_phage_RB55	RB55	0	181	76	61
_Enterobacteria_phage_RB59	RB59	1	183	76	61
_Enterobacteria_phage_RB5	RB5	0	184	76	61
_Enterobacteria_phage_RB68	RB68	1	189	76	61
_Enterobacteria_phage_RB69	RB69	12	181	76	61
_Enterobacteria_phage_RB6	RB6	0	184	76	61
_Enterobacteria_phage_RB7	RB7	0	185	76	61
_Enterobacteria_phage_RB9	RB9	0	185	76	61
_Enterobacteria_phage_T4T	T4T	1	183	76	61
_Enterobacteria_phage_T4	T4	6	177	76	61
_Enterobacteria_phage_T4_strain_wild	T4_strain_wild	5	195	76	61
_Enterobacteria_phage_T6	T6	3	189	76	61
_Enterobacteria_phage_vB_EcoM-VR7	VR7	6	204	76	61
_Enterobacteria_phage_vB_EcoM_ACG-C40	ACG-C40	2	193	76	61
_Enterobacteria_phage_vB_EcoM_IME339	IME339	0	173	76	61
_Enterobacteria_phage_vB_EcoM_IME340	IME340	0	178	76	61
_Enterobacteria_phage_vB_EcoM_VR20	VR20	7	201	76	61
_Enterobacteria_phage_vB_EcoM_VR25	VR25	4	209	76	61
_Enterobacteria_phage_vB_EcoM_VR26	VR26	8	203	76	61
_Enterobacteria_phage_vB_EcoM_VR5	VR5	5	186	76	61
_Erwinia_phage_Cronus	Cronus	145	121	76	61

_Escherichia_coli_O157_typing_phage_3	typing_phage_3	2	187	76	61
_Escherichia_coli_O157_typing_phage_6	typing_phage_6	1	169	76	61
_Escherichia_phage_APCEc01	APCEc01	0	192	76	61
_Escherichia_phage_AnYang	AnYang	7	196	76	61
_Escherichia_phage_CF2	CF2	7	197	76	61
_Escherichia_phage_D5505	D5505	1	185	76	61
_Escherichia_phage_ECD7	ECD7	9	196	76	61
_Escherichia_phage_ECML-134	ECML-134	1	184	76	61
_Escherichia_phage_EcWhh-1	EcWhh-1	12	202	76	61
_Escherichia_phage_HP3	HP3	2	199	76	61
_Escherichia_phage_HY01	HY01	2	203	76	61
_Escherichia_phage_HY03	HY03	6	177	76	61
_Escherichia_phage_KIT03	KIT03	1	183	76	61
_Escherichia_phage_Lw1	Lw1	6	189	76	61
_Escherichia_phage_MX01	MX01	9	191	76	61
_Escherichia_phage_OLB35	OLB35	5	186	76	61
_Escherichia_phage_PE37	PE37	3	195	76	61
_Escherichia_phage_PP01	PP01	2	188	76	61
_Escherichia_phage_SF	SF	3	199	76	61
_Escherichia_phage_ST0	ST0	1	186	76	61
_Escherichia_phage_T2	T2	11	178	76	61
_Escherichia_phage_UFV-AREG1	UFV-AREG1	3	186	76	61
_Escherichia_phage_WG01	WG01	3	190	76	61
_Escherichia_phage_e112	e112	4	190	76	61
_Escherichia_phage_p000v	p000v	0	183	76	61
_Escherichia_phage_p000y	p000y	3	200	76	61
_Escherichia_phage_slur02	slur02	3	198	76	61
_Escherichia_phage_slur03	slur03	0	191	76	61
_Escherichia_phage_slur04	slur04	0	183	76	61
_Escherichia_phage_slur07	slur07	3	195	75	61
_Escherichia_phage_slur08	slur08	0	192	76	61
_Escherichia_phage_slur11	slur11	0	183	76	61
_Escherichia_phage_slur13	slur13	0	184	76	61
_Escherichia_phage_slur14	slur14	0	191	76	61
_Escherichia_phage_vB_EcoM-G28	G28	2	176	76	61
_Escherichia_phage_vB_EcoM-Sa45lw	Sa45lw	0	189	76	61
_Escherichia_phage_vB_EcoM-UFV13	UFV13	1	185	76	61
_Escherichia_phage_vB_EcoM-fFiEco06	fFiEco06	0	191	76	61
_Escherichia_phage_vB_EcoM-fHoEco02	fHoEco02	0	189	76	61
_Escherichia_phage_vB_EcoM-DalCa	DalCa	0	187	76	61
_Escherichia_phage_vB_EcoM_G10400	G10400	0	194	76	61
_Escherichia_phage_vB_EcoM_G2248	G2248	7	194	76	61
_Escherichia_phage_vB_EcoM_G2285	G2285	12	190	76	61

_Escherichia_phage_vB_EcoM_G2469	G2469	4	199	76	61
_Escherichia_phage_vB_EcoM_G2494	G2494	17	202	75	61
_Escherichia_phage_vB_EcoM_G2540-3	G2540-3	1	184	76	61
_Escherichia_phage_vB_EcoM_G2540	G2540	13	195	76	61
_Escherichia_phage_vB_EcoM_G29	G29	1	192	76	61
_Escherichia_phage_vB_EcoM_G37-3	G37-3	23	196	76	61
_Escherichia_phage_vB_EcoM_G4498	G4498	5	200	76	61
_Escherichia_phage_vB_EcoM_G4500	G4500	1	195	76	61
_Escherichia_phage_vB_EcoM_G4507	G4507	2	188	76	61
_Escherichia_phage_vB_EcoM_G50	G50	0	183	76	61
_Escherichia_phage_vB_EcoM_G5211	G5211	6	195	76	61
_Escherichia_phage_vB_EcoM_G53	G53	0	190	76	61
_Escherichia_phage_vB_EcoM_G8	G8	1	182	76	61
_Escherichia_phage_vB_EcoM_G9062	G9062	0	192	76	61
_Escherichia_phage_vB_EcoM_JB75	JB75	0	196	76	61
_Escherichia_phage_vB_EcoM_JS09	JS09	2	190	76	61
_Escherichia_phage_vB_EcoM_KAW1E185	KAW1E185	0	192	76	61
_Escherichia_phage_vB_EcoM_KAW3E185	KAW3E185	1	203	75	61
_Escherichia_phage_vB_EcoM_MM02	MM02	1	193	76	61
_Escherichia_phage_vB_EcoM_OE5505	OE5505	2	198	76	61
_Escherichia_phage_vB_EcoM_PhAPEC2	PhAPEC2	1	194	76	61
_Escherichia_phage_vB_EcoM_R5505	R5505	1	196	76	61
_Escherichia_phage_vB_EcoM_WFK	WFK	3	179	76	61
_Escherichia_phage_vB_EcoM_WFL6982	WFL6982	1	180	76	61
_Escherichia_phage_vB_EcoM_WFbE185	WFbE185	1	194	76	61
_Escherichia_phage_wV7	wV7	0	191	76	61
_Escherichia_virus_KFS-EC	KFS-EC	10	199	76	61
_Escherichia_virus_VE20	VE20	1	190	76	61
_Escherichia_virus_vB_Eco_mar005P1	mar005P1	0	187	76	61
_Escherichia_virus_vB_Eco_mar005P1 strain_vB_Eco_mar006P2	mar005P1_0	0	188	76	61
_Escherichia_virus_vB_Eco_mar005P1 strain_vB_Eco_mar007P3	mar005P1_1	0	187	76	61
_Escherichia_virus_vB_Eco_mar005P1 strain_vB_Eco_mar008P4	mar005P1_2	0	187	76	61
_Escherichia_virus_vB_Eco_mar005P1 strain_vB_Eco_mar009P5	mar005P1_3	0	187	76	61
_Klebsiella_phage_KP1	KP1	8	197	76	61
_Klebsiella_phage_KPV15	KPV15	31	191	76	61
_Klebsiella_phage_Marfa	Marfa	62	168	76	61
_Klebsiella_phage_Matisse	Matisse	3	197	76	61
_Klebsiella_phage_Mineola	Mineola	7	193	76	61
_Klebsiella_phage_Miro	Miro	4	193	76	61
_Klebsiella_phage_PKO111	PKO111	31	199	75	61
_Klebsiella_phage_vB_KpnM_KpV477	KpV477	9	188	76	61

_Klebsiella_phage_vB_Kpn_F48	F48	47	163	76	61
_Morganella_phage_vB_MmoM_MP1	MP1	109	100	75	61
_Pectobacterium_bacteriophage_PM2	PM2	54	168	76	61
_Phage_NC-G	NC-G	3	190	76	61
_Proteus_phage_phiP4-3	phiP4-3	113	101	75	61
_Proteus_phage_vB_PmiM_Pm5461	Pm5461	98	106	75	61
_Pseudomonas_phage_PspYZU05	PspYZU05	82	61	72	61
_Salmonella_phage_Melville	Melville	7	175	76	61
_Salmonella_phage_S16	S16	8	175	76	61
_Salmonella_phage_STML-198	STML-198	17	173	76	61
_Salmonella_phage_STP4-a	STP4-a	19	183	75	61
_Salmonella_phage_vB_SnwM_CGG4-1	CGG4-1	25	177	75	61
_Serratia_phage_PS2	PS2	144	78	74	61
_Shigella_phage_SH7	SH7	1	183	76	61
_Shigella_phage_SHBML-50-1	SHBML-50-1	0	184	76	61
_Shigella_phage_SHFML-11	SHFML-11	3	178	76	61
_Shigella_phage_SHSML-52-1	SHSML-52-1	5	201	76	61
_Shigella_phage_SP18	SP18	7	194	76	61
_Shigella_phage_Sf20	Sf20	2	191	76	61
_Shigella_phage_Sf21	Sf21	1	183	76	61
_Shigella_phage_Sf22	Sf22	0	185	76	61
_Shigella_phage_Sf23	Sf23	0	188	76	61
_Shigella_phage_Sf24	Sf24	1	187	76	61
_Shigella_phage_Shfl2	Shfl2	0	183	76	61
_Shigella_phage_pSs-1	pSs-1	2	182	76	61
_Vibrio_phage_KVP40	KVP40	64	302	73	61
_Vibrio_phage_ValB1_HC	ValB1_HC	11	334	73	61
_Vibrio_phage_ValKK3	ValKK3	18	335	73	61
_Vibrio_phage_phi-Grn1	phi-Grn1	27	339	73	61
_Vibrio_phage_phi-ST2	phi-ST2	30	343	73	61
_Vibrio_phage_vB_VmeM-32	32	165	49	71	61
_Vibriophage_phi-pp2	phi-pp2	63	296	73	61
_Yersinia_phage_PST	PST	0	185	76	61
_Yersinia_phage_fPS-2	fPS-2	0	176	76	61
_Yersinia_phage_phiD1	phiD1	6	190	76	61

ПРИЛОЖЕНИЕ 5

Поддержка к пангеномному дереву на рисунке 37

■ /0.333/15	■ /0.806/39	■ /0.979/93
■ /0.333/58	■ /0.806/48	■ /0.98/74
■ /0.333/84	■ /0.82/48	■ /0.982/100
■ /0.333/90	■ /0.824/70	■ /0.983/100
■ /0.333/92	■ /0.828/63	■ /0.984/100
■ /0.333/94	■ /0.828/95	■ /0.984/86
■ /0.333/98	■ /0.836/38	■ /0.989/100
■ /0.336/52	■ /0.841/88	■ /0.989/82
■ /0.379/31	■ /0.841/97	■ /0.989/91
■ /0.38/49	■ /0.843/91	■ /0.99/31
■ /0.401/79	■ /0.85/96	■ /0.992/100
■ /0.404/84	■ /0.862/91	■ /0.993/70
■ /0.419/27	■ /0.867/31	■ /0.993/80
■ /0.431/62	■ /0.876/91	■ /0.993/97
■ /0.439/98	■ /0.876/93	■ /0.994/54
■ /0.44/91	■ /0.886/40	■ /0.994/59
■ /0.458/80	■ /0.895/63	■ /0.995/100
■ /0.463/94	■ /0.905/100	■ /0.995/62
■ /0.465/87	■ /0.906/99	■ /0.995/76
■ /0.466/55	■ /0.907/88	■ /0.995/88
■ /0.473/55	■ /0.912/49	■ /0.996/63
■ /0.486/94	■ /0.912/81	■ /0.996/95
■ /0.501/97	■ /0.919/89	■ /0.997/100
■ /0.503/87	■ /0.919/97	■ /0.997/76
■ /0.517/76	■ /0.926/58	■ /0.998/28
■ /0.545/47	■ /0.927/99	■ /0.998/69
■ /0.546/83	■ /0.93/63	■ /0.998/91
■ /0.551/49	■ /0.939/97	■ /0.999/100
■ /0.556/99	■ /0.942/93	■ /0.999/56
■ /0.571/40	■ /0.945/53	■ /0.999/76
■ /0.601/99	■ /0.945/98	■ /0.999/93
■ /0.607/94	■ /0.947/86	■ /0.999/96
■ /0.613/98	■ /0.95/72	■ /0.999/97
■ /0.613/99	■ /0.952/35	■ /0.999/98
■ /0.632/99	■ /0.953/92	■ /0.999/99
■ /0.634/55	■ /0.953/98	■ /1/100
■ /0.645/99	■ /0.954/99	■ /1/60
■ /0.657/88	■ /0.957/99	■ /1/66
■ /0.677/58	■ /0.96/65	■ /1/75
■ /0.684/56	■ /0.962/91	■ /1/84
■ /0.687/90	■ /0.963/51	■ /1/89
■ /0.709/99	■ /0.966/87	■ /1/94
■ /0.745/53	■ /0.968/99	■ /1/96
■ /0.748/99	■ /0.969/96	■ /1/97
■ /0.763/65	■ /0.974/86	■ /1/98
■ /0.767/96	■ /0.976/97	■ /1/99
■ /0.777/88	■ /0.978/75	

Левое значение – поддержка aBayes, правое – поддержка ultrafast bootstrap

Поддержка к филогенетическому дереву на рисунке 38

0/0.333/13	72.4/0.423/80	84.1/0.987/65	93.8/1/95
0/0.333/16	72.7/0.936/69	85.4/0.995/95	93/1/100
0/0.333/24	73.5/0.622/100	85.7/0.987/92	94.8/1/100
0/0.333/27	73.9/0.836/99	86.6/0.998/66	94.9/1/82
0/0.333/3	74.3/0.941/99	86.8/0.987/98	94/1/77
0/0.333/31	74.9/0.583/100	86.9/0.999/66	95.3/1/94
0/0.333/34	74.9/0.997/76	86/0.993/64	95.6/1/95
0/0.333/42	74/0.996/90	86/0.993/96	96.1/1/95
0/0.333/47	75.1/0.483/92	87.2/0.999/92	96.9/1/100
0/0.333/49	75.1/0.581/98	87.3/0.996/100	97.9/1/100
0/0.333/50	75.2/0.604/97	87.4/0.996/99	97/1/100
0/0.333/51	75.3/0.897/100	87.4/1/100	98.4/1/100
0/0.333/52	75.8/0.686/98	87.6/0.992/98	98.7/1/100
0/0.333/55	75/0.908/98	87.7/0.999/85	99.2/1/100
0/0.333/56	76.1/0.991/93	87.8/1/83	99.4/1/100
0/0.333/57	76.7/0.964/99	87/1/75	99.5/1/100
0/0.333/61	76.9/0.9/90	88.5/1/97	99.5/1/98
0/0.333/65	77.1/0.718/50	88.6/1/90	99.7/1/100
0/0.333/69	77.3/0.703/92	88.8/1/100	99.7/1/82
0/0.333/8	77.9/0.922/76	89.5/1/75	99.8/1/100
0/0.333/84	77.9/0.958/96	89.6/1/85	99.9/1/100
0/0.333/86	78.7/0.913/99	89.7/0.999/97	99.9/1/99
0/0.333/88	78.7/0.996/96	89.8/1/68	99/1/100
0/0.333/90	79.3/0.917/28	90.2/1/99	
0/0.333/91	79.6/0.938/89	90.6/0.999/87	
0/0.338/96	79.6/0.968/100	90.6/0.999/99	
100/1/100	80.4/0.92/97	90.7/1/100	
23.4/0.605/80	80.4/0.983/100	90.9/0.999/99	
24.3/0.708/78	81.2/0.72/100	91.1/0.996/98	
31.2/0.745/61	81.5/0.988/64	91.4/1/67	
36/0.546/99	81.7/0.908/99	91.9/0.999/100	
42.1/0.863/88	81.9/0.888/78	91.9/1/100	
42.6/0.834/100	82.4/0.928/81	92.1/0.978/94	
54.3/0.981/86	82.5/0.999/94	92.2/1/99	
55.4/0.935/79	82.7/0.993/100	92.5/1/100	
62.9/0.968/69	82.7/0.998/89	92.6/1/98	
62.9/0.979/98	83.5/0.999/84	92.9/1/88	
65.2/0.931/73	83.8/0.979/99	93.2/0.999/73	
69.3/0.883/80	83.9/0.975/99	93.2/1/99	
71.8/0.987/72	83.9/0.996/100	93.6/1/100	

Левое значение – поддержка SH-aLRT, значение по центру – поддержка aBayes, правое значение – поддержка ultrafast bootstrap

**Таблица гомологов РНК-лигазы 2, использованных при получении
филогенетического дерева на рисунке 41**

№	Род	Фар	Номер доступа белка
1	unclassified <i>Tevenvirinae</i>	Acinetobacter phage 133	YP_004300608.1
2		Acinetobacter phage AM101	AWY10415.1
3		Acinetobacter phage Ac42	YP_004009372.1
4		Acinetobacter phage Acj61	YP_004009627.1
5		Acinetobacter phage Acj9	YP_004010147.1
6		Acinetobacter phage KARL-1	AXY82640.1
7		Acinetobacter phage ZZ1	YP_006488989.1
8		Acinetobacter phage vB_ApiM_fHyAci03	AXF40578.1
9		Aeromonas phage 65.2	APU01459.1
10		Aeromonas phage 65	YP_004300909.1
11		Aeromonas phage AS-szw	ATI17438.1
12		Aeromonas phage AS-zj	ASU00157.1
13		Aeromonas virus Aeh1	NP_944126.1
14		Aeromonas phage AsFcp_2	QAX98490.1
15		Aeromonas phage Asswx_1	QAX97878.1
16		Aeromonas phage Aswh_1	QAY01272.1
17		Aeromonas phage CC2	YP_007010325.1
18		Aeromonas phage PX29	YP_009011662.1
19		Citrobacter phage IME-CF2	YP_009218766.1
20		Citrobacter phage Margaery	YP_009195056.1
21		Citrobacter phage Maroon	AYJ73100.1
22		Citrobacter phage Miller	YP_009097842.1
23		Citrobacter phage vB_CfrM_CfP1	YP_009285773.1
24		Enterobacteria phage RB16	YP_003858534.1
25		Enterobacteria phage RB43	YP_239225.1
26		Cronobacter phage vB_CsaM_GAP161	YP_006986511.1
27		Cronobacter phage vB_CsaM_leB	AOG16366.1
28		Cronobacter phage vB_CsaM_leN	AOG16651.1
29		Erwinia phage Cronus	AWD90329.1
30		Escherichia phage Lw1	YP_008060759.1
31		Klebsiella phage Marfa	QDB71837.1
32		Proteus phage phiP4-3	AUM58455.1
33		Proteus phage vB_PmiM_Pm5461	YP_009195583.1
34		Pseudomonas phage PspYZU05	ASD52094.1
35		Klebsiella phage vB_Kpn_F48	AUO78857.1
36		Morganella phage vB_MmoM_MP1	YP_009280031.1
37		Pectobacterium bacteriophage PM2	YP_009211620.1
38		Serratia phage PS2	AHY25432.1
39		Vibrio phage vB_VmeM-32	ALY07226.1
40	 <i>Moonvirus</i>	Citrobacter phage Merlin	YP_009203914.1
41		Citrobacter phage Moon	YP_009146629.1
42	 <i>Karamvirus</i>	Cronobacter phage Pet-CM3-4	SCN45872.1
43		Enterobacter phage PG7	YP_009005458.1
44		Enterobacter phage myPSH1140	AVR55365.1
45	 <i>Slopekvirus</i>	Enterobacteria phage CC31	YP_004010038.1
46		Enterobacter phage phiEap-3	YP_009607157.1
47		Klebsiella phage Matisse	YP_009194487.1
48	 <i>Mosigvirus</i>	Klebsiella phage Miro	YP_009607435.1
49		Enterobacteria phage HX01	*

№	Род	Фар	Номер доступа белка
50	 <i>Mosigvirus</i>	Enterobacteria phage RB69	NP_861881.1
51		Enterobacteria phage ATK47	ANZ51023.1
52		Enterobacteria phage ATK48	ANZ51366.1
53		Escherichia coli O157 typing phage 6	YP_009593176.1
54		Escherichia coli O157 typing phage 3	YP_009592780.1
55		Escherichia phage APCEc01	YP_009225077.1
56		Escherichia phage HP3	*
57		Escherichia phage OLB35	AYR04095.1
58		Escherichia phage SF	AWY07818.1
59		Escherichia phage ST0	YP_009608487.1
60		Escherichia phage p000v	AYN56237.1
61		Escherichia phage p000y	AYN56688.1
62		Escherichia phage vB_EcoM_G2285	QBO62608.1
63		Escherichia phage vB_EcoM_G2469	QBO62878.1
64		Escherichia phage vB_EcoM_G53	QBO65324.1
65		Escherichia phage vB_EcoM_JS09	YP_009037590.1
66		Escherichia phage vB_EcoM_KAW3E185	QBQ78739.1
67		Escherichia phage vB_EcoM_MM02	QBQ79237.1
68		Escherichia phage vB_EcoM_WFK	QBQ77198.1
69		Escherichia phage vB_EcoM_WFL6982	QBQ76937.1
70		Escherichia phage vB_EcoM_WFbE185	QBQ77690.1
71		Escherichia phage vB_EcoM_PhAPEC2	YP_009056764.1
72		Escherichia virus vB_Eco_mar005P1	VCU44444.1
73		Escherichia virus vB_Eco_mar005P1 strain vB_Eco_mar006P2	VCU44449.1
74		Escherichia virus vB_Eco_mar005P1 strain vB_Eco_mar007P3	VCU43253.1
75		Escherichia virus vB_Eco_mar005P1 strain vB_Eco_mar008P4	VCU43961.1
76		Escherichia virus vB_Eco_mar005P1 strain vB_Eco_mar009P5	VCU44820.1
77		Shigella phage SHSML-52-1	YP_009289105.1
78	 <i>Dhakavirus</i>	Enterobacteria phage IME08	YP_003734320.1
79		Enterobacteria phage JS10	YP_002922522.1
80		Escherichia phage JS98	YP_001595305.1
81		Escherichia phage QL01	YP_009202912.1
82		Enterobacteria phage vB_EcoM_VR5	YP_009205871.1
83		Escherichia phage AnYang	QAU03652.1
84		Escherichia phage EcWhh-1	QAX99933.1
85		Escherichia phage MX01	YP_009324073.1
86		Escherichia phage WG01	YP_009323380.1
87		Phage NC-G	QBP35520.1
88		Escherichia phage vB_EcoM_VR7	YP_004063874.1
89		Enterobacteria phage Bp7	YP_007004267.1
90	 <i>Gaprivervirus</i>	Escherichia phage vB_EcoM_VR20	YP_009207369.1
91		Escherichia phage vB_EcoM_VR25	YP_009209935.1
92		Escherichia phage vB_EcoM_VR26	YP_009214027.1
93		Shigella phage SP18	YP_003934814.1
94	 <i>Tequatrovirus</i>	Enterobacteria phage Aplg8	ANZ50774.1
95		Enterobacteria phage GiZh	*
96		Escherichia phage AR1	YP_009167991.1
97		Enterobacteria phage Kha5h	ANZ51799.1
98		Enterobacteria phage RB10	AIT74268.1
99		Escherichia virus RB14	YP_002854510.1
100		Escherichia virus RB32	YP_803117.1
101		Enterobacteria phage RB18	AXF42485.1

№	Род	Фар	Номер доступа белка
102		Enterobacteria phage RB27	YP_009102379.1
103		Enterobacteria phage RB33	AIT74814.1
104		Escherichia phage RB3	YP_009098560.1
105		Enterobacteria phage RB51	ACP31095.1
106		Enterobacteria phage RB55	AIT75086.1
107		Enterobacteria phage RB59	AIT75360.1
108		Enterobacteria phage RB5	AIT73181.1
109		Enterobacteria phage RB68	AIT75637.1
110		Enterobacteria phage RB6	AIT73452.1
111		Enterobacteria phage RB7	AIT73723.1
112		Enterobacteria phage RB9	AIT73995.1
113		Enterobacteria phage T4T	ADJ39898.1
114		Escherichia virus T4	NP_049790.1
115		Enterobacteria phage T4 strain wild	*
116		Enterobacteria phage T6	AXN58211.1
117		Escherichia phage vB_EcoM_ACG-C40	YP_006986729.1
118		Enterobacteria phage vB_EcoM_IME339	AWD91505.1
119		Enterobacteria phage vB_EcoM_IME340	AWD91761.1
120		Escherichia phage CF2	*
121		Escherichia phage D5505	QBO61184.1
122		Escherichia phage ECML-134	YP_009102648.1
123		Escherichia phage HY01	YP_009148617.1
124		Escherichia phage HY03	YP_009284038.1
125		Escherichia phage HY03	YP_009284039.1
126		Escherichia phage KIT03	BBG28692.1
127		Escherichia phage PE37	ANH49691.1
128		Escherichia phage PP01	BBC14517.1
129		Escherichia phage T2	BBC14796.1
130		Escherichia phage UFV-AREG1	YP_009281510.1
131		Escherichia phage vB_EcoM_112	YP_009030783.1
132		Escherichia phage slur02	YP_009210095.1
133		Escherichia phage slur03	YP_009625074.1
134		Escherichia phage slur04	YP_009625392.1
135		Escherichia phage slur07	YP_009197280.1
136		Escherichia phage slur08	CUL02618.1
137		Escherichia phage slur11	CUL02946.1
138		Escherichia phage slur13	CUL03732.1
139		Escherichia phage slur14	YP_009180676.1
140		Escherichia phage vB_EcoM-G28	AVH86005.1
141		Escherichia phage vB_EcoM-Sa45lw	QDF15200.1
142		Escherichia phage vB_EcoM-UFV13	YP_009290443.1
143		Escherichia phage vB_EcoM-fFiEco06	AUV61038.1
144		Escherichia phage vB_EcoM-fHoEco02	AUV61311.1
145		Escherichia phage vB_EcoM_DalCa	AYP69772.1
146		Escherichia phage vB_EcoM_G10400	QBO63692.1
147		Escherichia phage vB_EcoM_G2540-3	QBO65583.1
148		Escherichia phage vB_EcoM_G2540	QBO63414.1
149		Escherichia phage vB_EcoM_G29	QBO64498.1
150		Escherichia phage vB_EcoM_G4498	QBO64219.1
151		Escherichia phage vB_EcoM_G4500	QBO65870.1
152		Escherichia phage vB_EcoM_G4507	QBO66145.1
153		Escherichia phage vB_EcoM_G50	QBO65046.1

№	Род	Фар	Номер доступа белка
154		Escherichia phage vB_EcoM_G8	QBQ80057.1
155		Escherichia phage vB_EcoM_G9062	QBQ77959.1
156		Escherichia phage vB_EcoM_JB75	AXC34018.1
157		Escherichia phage vB_EcoM_KAW1E185	QBQ78462.1
158		Escherichia phage vB_EcoM_OE5505	QBQ79505.1
159		Escherichia phage vB_EcoM_R5505	QBQ79789.1
160		Escherichia phage wV7	YP_007004917.1
161		Escherichia virus VEc20	QDK04303.1
162		Shigella phage SH7	APC45045.1
163		Shigella phage SHBML-50-1	YP_009288541.1
164		Shigella phage SHFML-11	YP_009277549.1
165		Shigella phage Sf21	YP_009618985.1
166		Shigella phage Sf22	YP_009614828.1
167		Shigella phage Sf23	ATE86434.1
168		Shigella phage Sf24	YP_009619109.1
169		Shigella phage Shfl2	YP_004415069.1
170		Shigella phage pSs-1	YP_009110995.1
171		Yersinia phage PST	YP_009153775.1
172		Yersinia phage fPS-2	VEV89788.1
173		Yersinia phage phiD1	YP_009149419.1
174	● <i>Jiaodavirus</i>	Klebsiella phage KP1	AUV57527.1
175		Klebsiella phage KPV15	APD20579.1
176		Klebsiella phage Mineola	AWY07077.1
177		Klebsiella phage PKO111	YP_009289602.1
178		Klebsiella phage vB_KpnM_KpV477	YP_009288865.1
179	● <i>Gelderlandvirus</i>	Salmonella phage Melville	YP_009615654.1
180		Salmonella phage vB_SenMS16	YP_007501209.1
181		Salmonella phage STML-198	YP_009148159.1
182		Salmonella phage STP4-a	YP_009126374.1
183		Salmonella phage vB_SnwM_CGG4-1	YP_009286529.1
184	● <i>Schizotequatrovirus</i>	Vibrio phage KVP40	NP_899381.1
185		Vibrio phage ValB1 HC	QBX05956.1
186		Vibrio phage ValKK3	YP_009201230.1
187		Vibrio phage phi-Grn1	ALP47010.1
188		Vibrio phage phi-ST2	ALP47390.1
189		Vibriophage phi-pp2	AFN37364.1

«*» – белок получен путем аннотирования/реаннотирования генома фага