

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ИНСТИТУТ БИОХИМИИ И ФИЗИОЛОГИИ МИКРООРГАНИЗМОВ
ИМЕНИ Г.К.СКРЯБИНА

На правах рукописи
УДК 577.2

СОКОЛОВ АНДРЕЙ СЕРГЕЕВИЧ

**ХОМИНГ-ЭНДОНУКЛЕАЗА SegD БАКТЕРИОФАГА T4: БИОХИМИЧЕСКАЯ И
ФУНКЦИОНАЛЬНАЯ ХАРАКТЕРИСТИКА**

Специальность 03.01.03 – Молекулярная биология

Диссертация
на соискание ученой степени
кандидата биологических наук

Научный руководитель
кандидат биологических наук
И. Э. Грановский

ПУЩИНО – 2018

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	4
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ.....	8
1.1. Семейство Т-четные бактериофаги.....	8
1.1.2. Особенности структуры и развития бактериофага Т4.....	9
1.1.3. Явление частичного исключения генетических маркеров при скрещивании фагов Т4-типа.....	12
1.2. Хоминг-эндонуклеазы.....	14
1.2.1. Описание хоминг-эндонуклеаз.....	17
1.2.2. Описание различных групп хоминг-эндонуклеаз.....	21
1.2.3. Хоминг, как один из вариантов мобильности ДНК.....	28
1.2.4. Биологическая роль хоминг-эндонуклеаз.....	31
1.3. Хоминг-эндонуклеаз бактериофага Т4 и родственных фагов.....	32
1.3.1 Хоминг-эндонуклеаз бактериофага Т4, закодированных в составе интронов ..	32
1.3.2 Свободностоящие хоминг-эндонуклеазы бактериофага Т4, сходства и основные отличия от интрон-кодируемых эндонуклеаз.....	33
1.3.2.1 Сайт-специфические эндонуклеазы Mob-группы.....	36
1.3.3.2 Сайт-специфические эндонуклеазы Seg-группы.....	37
ГЛАВА 2. МЕТОДЫ ИССЛЕДОВАНИЯ И МАТЕРИАЛЫ	
2.1. Материалы.....	42
2.1.1. Бактериофаги и бактериальные штаммы	42
2.1.2. Вектора и полученные в ходе работы плазмиды.....	43
2.1.3. Олигонуклеотиды.....	44
2.1.4. Растворы, буферы и ростовые среды.....	45
2.1.5. Реактивы и материалы	47
2.2. Методы исследования.....	488
2.2.1. Нарботка плазмид и определение концентрации ДНК.....	48
2.2.2. Нарботка ДНК Т-четных бактериофагов.....	48
2.2.3. Трансформация и приготовление компетентных клеток <i>E. coli</i>	48
2.2.4. Полимеразная цепная реакция	49
2.2.5. Обработка ДНК эндонуклеазами и лигирование ДНК-фрагментов.....	49
2.2.6. Получение рекомбинантных плазмид.....	49
2.2.7. Получение фагов Т4ΔsegD, D2, B20dm, D2insZ и D2ets9.....	50
2.2.8. Определение нуклеотидной последовательности ДНК.....	51

2.2.9. Скрининг рекомбинантных плазмид.....	51
2.2.10. Разделение фрагментов ДНК в электрическом поле	51
2.2.11. ДНК гибридизация по методу Саузерна	52
2.2.12. Электрофоретическое разделение белков в полиакриламидном геле	54
2.2.13. Получение TEV-протеазы	54
2.2.14. Очистка TEV-протеазы.....	54
2.2.15. Экспрессия ОРС <i>segD</i>	55
2.2.16. Очистка эндонуклеазы SegD	55
2.2.17. Определение оптимальных условий работы эндонуклеазы SegD <i>in vitro</i>	56
2.2.18. Установление точек расщепления и границ узнаваемой последовательности ДНКэндонуклеазой SegD	57
2.2.19. Анализ эндонуклеазной активности SegD на ДНК бактериофагов	58
2.2.20. Определение эндонуклеазной активности SegD в бесклеточных экстрактах <i>E.coli</i> , зараженных бактериофагом Т4.....	58
2.2.21. Получение Т-четных бактериофагов и определение титра фагов.....	59
2.2.22. Скрещивание бактериофагов и анализ потомства	60
2.2.23. Программы, использованные в работе.....	60
ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.....	61
3.1. Структурная организация области генов 23-24 у Т4-родственных бактериофагов.....	61
3.2. Очистка белка SegD	63
3.3. Исследование биохимических свойств эндонуклеазы SegD.....	65
3.4. Картирование сайта гидролиза эндонуклеазы SegD.....	71
3.5. Сайт гидролиза эндонуклеазы SegD находится на 3'- конце гена 23.....	72
3.6. Исследование влияния модификации ДНК Т-четных бактериофагов на эффективность и специфичность гидролиза SegD.....	76
3.7. Ген <i>segD</i> экспрессируется в клетках <i>E. coli</i> , инфицированных фагом Т4....	79
3.8. Анализ наследования гена <i>segD</i> при скрещивании <i>segD</i> ⁺ и <i>segD</i> ⁻ фагов.....	81
3.9. Анализ SegD-инициируемой рекомбинации в <i>rII</i> -локусе фага Т4.....	85
ЗАКЛЮЧЕНИЕ.....	87
ВЫВОДЫ.....	88
СПИСОК СОКРАЩЕНИЙ.....	89
СПИСОК ЛИТЕРАТУРЫ.....	90

ВВЕДЕНИЕ

Хоминг-эндонуклеазы являются особым классом сайт-специфических эндодезоксирибонуклеаз. Эти ферменты, как правило, вносят разрыв в ДНК родственного аллеля, лишённого гена хоминг-эндонуклеазы, и инициируют тем самым перенос собственного гена и фланкирующих областей в вариант аллеля, не содержащий ген нуклеазы. Данный процесс получил название “хоминг” (от англ. *homing* – возвращение домой).

Ферменты этой группы подразделяются на шесть основных семейств. Объединение в семейства обусловлено присутствием в белках характерных аминокислотных мотивов: “His-Cys box”, “LAGLIDADG”, “EDxHD”, “H-N-H”, “PD-(D/E)xK” и “GIY-YIG” (Guhan and Muniyappa, 2003; Stoddard, 2005, Stoddard, 2011, Taylor, 2011, Stoddard, 2014).

Хоминг-эндонуклеазы, в отличие от эндонуклеаз рестрикции, узнают протяженные (до 40 п.н.), часто вырожденные последовательности ДНК. Это свойство обеспечивается уникальной стратегией узнавания сайтов гидролиза хоминг-эндонуклеазами. В связи с этим изучение ферментов данной группы расширяет границы наших знаний о ДНК-белковых взаимодействиях, и закладывает основу для конструирования белков с заданной специфичностью.

Гены хоминг-эндонуклеаз найдены у: бактериофагов, зубактерий, архей, а также у низших эукариот (Belfort et al., 2005). ОРС, кодирующие данный класс белков, обнаружены в интронах, интеинах и в межгенных областях. Геномы фагов, как предполагает ряд исследователей, наиболее богаты генами хоминг-эндонуклеаз (Edgell et al., 2000). Среди 162 предполагаемых ОРС бактериофага T5, найдено 9 генов хоминг-эндонуклеаз (асс. по AY543070, Диссертация Акуленко Н. В.), а в геноме бактериофага T4 найдено 15 (уже доказанных и предполагаемых) генов хоминг-эндонуклеаз. Они занимают около 15 т.п.о., что составляет 11 % фагового генома (Belfort et al., 2005, Edgell et al., 2010). Высокая плотность данных генов в геномах вышеперечисленных фагов ставит вопросы о роли хоминг-эндонуклеаз в эволюции фагов. Показано, что данные ферменты, помимо направленного переноса собственных генов, могут вызывать и перенос примыкающих к гену хоминг-эндонуклеазы участков ДНК и, тем самым, инициировать обмен фрагментами ДНК. В связи с этим интересна роль хоминг-эндонуклеаз в функционировании геномов крупных бактериофагов, таких как T4 и T5.

Группа исследователей, используя разработанные алгоритмы поиска родственных генов, в геноме T4 нашла 7 ОРС, продукты которых проявляют сходство с интронкодируемой хоминг-эндонуклеазой I-TevI. На N-конце белков обнаружен вырожденный вариант мотива GIY-10 a.o.-YIG, характерный для одноименного семейства хоминг-эндонуклеаз. Предсказанные белки являются близкими по размеру основными белками. Сходство этих белков с интронкодируемыми хоминг-эндонуклеазами позволило предположить, что данные белки

также являются эндонуклеазами. Для белков данной группы была предложена аббревиатура Seg (similar to endonucleases of group I introns).

К настоящему моменту показано, что SegA, B, C, E, F и G являются Mg^{2+} -зависимыми сайт-специфическими эндодезоксирибонуклеазами (Brok-Volchanskaya et al., 2008; Liu et al., 2003; Belle et al., 2002; Kadyrov et al., 1997; Sharma et al., 1994). Для данных белков, за исключением SegA, была показана способность вызывать генетическую рекомбинацию среди Т-четных фагов. В литературе отсутствует информация о функции и свойствах белка SegD. Также нет данных об экспрессии гена *segD* и его мобильности. Данная работа посвящена изучению предполагаемой эндонуклеазы SegD бактериофага Т4.

Цель и задачи исследования. Цель работы - изучить, действительно ли ген *segD* кодирует сайт-специфическую эндонуклеазу, способную инициировать хоминг собственного гена. В соответствии с целью работы были поставлены следующие задачи:

- В соответствии с поставленной целью работы были сформулированы следующие задачи:
1. Исследовать распространенность гена *segD* среди Т-четных фагов.
 2. Разработать схему очистки эндонуклеазы SegD, выделить белок, охарактеризовать его ферментативные свойства и идентифицировать его сайт гидролиза в геноме Т-четных фагов.
 3. Изучить способность эндонуклеазы SegD инициировать хоминг собственного гена.
 4. Сконструировать модель для изучения SegD-инициируемой рекомбинации в *rII*-локусе бактериофага Т4. Проанализировать в модельной системе зависимость SegD-инициируемой рекомбинации в данном локусе от уровня SegD в клетках.

Научная новизна работы.

В данной работе была определена нуклеотидная последовательность области генов 23 - 24 у 14 Т-четных бактериофагов. Было установлено, что ген *segD*, помимо фага Т4, присутствует у фагов RB55, RB59 и Т6.

Установлено, что SegD является Mg^{2+} -зависимой сайт-специфической эндодезоксирибонуклеазой. Данная эндонуклеаза способна гидролизовать природную ДНК Т-четных фагов, содержащую гликозилированные остатки гидроксиметилцитозина. Последовательность ДНК, узнаваемая и гидролизуемая эндонуклеазой SegD обнаружена в геномах всех проанализированных фагов. Эта ситуация является крайне необычной, поскольку, как правило, предпочтительный сайт гидролиза хоминг-эндонуклеазы располагается в фаге, не содержащем ген хоминг-эндонуклеазы.

Изучена способность эндонуклеазы SegD инициировать мобильность собственного гена. Было показано, что в условиях природного уровня экспрессии эндонуклеаза SegD не способна

инициировать мобильность собственного гена. При искусственном увеличении количества эндонуклеазы SegD в клетке, она способна инициировать мобильность собственного гена, а также генетическую конверсию в удаленном локусе по искусственно введенному сайту данной эндонуклеазы.

В этом исследовании описана новая сайт-специфическая эндонуклеаза бактериофага T4. Полученные в настоящей работе данные о свойствах эндонуклеазы SegD дополняют информацию о хоминг-эндонуклеазах и, в частности, о ферментах, относящихся к GIY-YIG-семейству белков.

Научно-практическая ценность.

Результаты, полученные в работе, имеют как фундаментальное, так и прикладное значение. Часть результатов может быть использована для практического применения. Изученная в данной работе эндонуклеаза SegD может быть использована для создания нуклеаз с заданной специфичностью.

Данные о мобильности гена *segD* фага T4 могут быть полезны при исследовании механизмов сайт-специфической рекомбинации и конструировании новых механизмов направленной передачи и замещения генетической информации. Бактерия-продуцент сайт-специфической эндодезоксирибонуклеазы SegD и разработанная схема очистки могут быть использованы для получения препаративных количеств белка, необходимых, например, для определения пространственной структуры данной сайт-специфической эндонуклеазы.

Положения, выносимые на защиту:

1. Ген *segD* фага T4 экспрессируется в ходе инфекционного цикла T4; он не является существенным для его развития и отсутствует в геномах большинства T4-родственных фагов.
2. Ген *segD* кодирует Mg^{2+} -зависимую сайт-специфическую эндонуклеазу, подобную хоминг-нуклеазам семейства GIY-YIG. Эндонуклеаза SegD узнает протяженную ассиметричную последовательность ДНК и нечувствительна к природным модификациям ДНК T-четных бактериофагов.
3. В геномах T4-родственных бактериофагов, содержащих ген *segD*, имеется сайт гидролиза кодируемой им эндонуклеазы.
4. Эндонуклеаза SegD неспособна инициировать хоминг собственного гена при природном уровне экспрессии данного фермента в клетке.

Степень достоверности и апробация результатов работы

Результаты работы опубликованы в рецензируемых научных журналах. Материалы диссертации были представлены на международных конференциях “38th FEBS Congress” (Saint

Petersburg, 2013), “Molecular Genetics of Bacteria and Phages” (Madison, 2007), а также 10-ой Пущинской школе-конференции молодых учёных (Пущино, 2006) и ежегодной сессии Ученого совета ИБФМ РАН, посвященной подведению итогов конкурса научных работ (Пущино, 2005).

Апробация диссертационной работы была проведена на заседании секции «Молекулярная биология» Ученого совета НИЦ "Курчатовский институт"-ГосНИИГенетика, а также на совместном семинаре лаборатории молекулярной микробиологии, биологии плазмид и биохимии клеточной поверхности микроорганизмов Учреждения Российской Академии наук Института биохимии и физиологии микроорганизмов имени Г.К. Скрыбина РАН.

Публикации. Основное содержание диссертации изложено в 3-х статьях в рецензируемых научных журналах

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Семейство Т-четные бактериофаги

Т-четные фаги были описаны в 1945 году (Demerec and Fano, 1945) и относятся к семейству *Myoviridae*. Их хозяевами служат представители групп *Enterobacteria*, *Acinetobacter*, *Aeromonas*, *Pseudomonas* и *Vibrio* (Ackermann and Krisch, 1997). К наиболее известным представителям данной группы относятся бактериофаги Т6, Т4 и Т2. К настоящему моменту известна полная нуклеотидная последовательность только генома Т4 (находится в свободном доступе). Тем не менее проведенные исследования по ДНК-ДНК гибридизации позволили установить родственные отношения среди Т-четных фагов. Оказалось, что гомология между ДНК фагов Т2 и Т6 ближе, чем у Т2 и Т4 или Т6 и Т4, соответственно (Cowie et al., 1971). Эти выводы были подтверждены и в исследованиях по электронной микроскопии гетеродуплексов ДНК Т-четных фагов проведенные Ким и Дэвидсон (Kim and Davidson, 1974). Установленные на основе этих экспериментов размеры фаговых ДНК (исходя из наблюдаемых длин кольцевых дуплексов) составили: 160 ± 2 , 166 ± 2 и 164 ± 2 т.п.н. для фагов Т2, Т4 и Т6. Из соответствия этих данных с современными, полученными в ходе секвенирования, согласно которым размер генома фага Т4 составляет 168903 п.н., можно сделать выводы об их высокой достоверности.

Сравнительная генетика Т-четных фагов была подробно описана Расселом в серии экспериментов по взаимной комплементации амбер мутаций у бактериофагов Т2, Т4 и Т6 (Russel, 1974).

К настоящему времени в литературе описаны около 160 фагов, структура капсида которых похожа на структуру капсида бактериофага Т4 (Ackermann and Krisch, 1997). Криш с соавторами (Tetart et al., 2001; Desplats and Krisch, 2003), основываясь на степени идентичности нуклеотидной последовательности гена 23 (структурного белка капсида), а также генов 18 и 19 (структурные белки отростка), предложили классификацию, согласно которой эти бактериофаги делятся на 4 подгруппы с возрастающим расхождением от генома фага Т4: Т-четные, псевдо-Т-четные, шизо-Т-четные и экзо-Т-четные.

У семейства Т-четных фагов наиболее консервативными являются белки системы репликации ДНК и структурные белки (от 35 до 70% идентичности). Они кодируются 20% ОРС, в то время как функции остальных ОРС на данный момент до конца детально не охарактеризованы (Miller et al., 2003).

Несмотря на сходство генетических карт и тесное иммунологическое родство (способность давать перекрестные иммунологические реакции) (Schwarz et al., 1983), в ДНК этих бактериофагами имеются существенные различия. Возрастающее разнообразие подгрупп

фагов Т4-типа коррелирует также с удаленностью филогенетического родства их хозяев от бактерии *E. coli*.

1.1.2 Особенности структуры и развития бактериофага Т4

Структура вириона фага Т4 принципиально не отличается от других представителей семейства *Myoviridae*. Т4 является относительно крупным фагом, имеет диаметр около 90 нм и длину около 200 нм. Вирус использует литический цикл развития, из одной клетки образуется 100-150 фаговых частиц. Геном Т4 представлен одной двухцепочечной молекулой ДНК. Структура ДНК фага Т4 имеет ряд отличительных особенностей: вместо канонического цитозина она содержит 5-гидроксиметилцитозин (ОМЦ) (Wyatt and Cohen, 1953). Кроме того, все гидроксильные группы 5-гидроксиметилцитозиновых оснований модифицированы в результате ковалентного связывания с глюкозой. При этом остатки глюкозы обращены в сторону большого желобка. Интересно, что картина гликозилирования у разных представителей семейства Т-четных фагов различается (Lehman and Pratt, 1960). Так, у бактериофага Т4, почти все остатки ОМЦ моногликозилированы, и лишь незначительная часть дигликозилирована, либо негликозилирована (менее 1%). Что касается фагов Т2 и Т6, то у них приблизительно четверть остатков ОМЦ негликозилирована, а 70% - моногликозилирована (Т2) или дигликозилирована (Т6). При этом у фага Т4 70% связей - α -гликозидные, а 30% - β -гликозидные. Такая модифицированная ДНК устойчива к действию большинства нуклеаз рестрикции бактерии-хозяина (Warren, 1980) и эндонуклеаз фага, участвующих в деградации хозяйской ДНК. Линейная двухцепочечная молекула ДНК длиной около 172 т.п.н. фага Т4 является одним из крупнейших вирусных геномов и содержит около 1-3% концевой избыточности. Это означает, что примерно около 3 т.п.н. на обоих концах молекулы ДНК повторяются и обуславливают характерные для фага Т4 кольцевые перестановки (Wood and Revel, 1976). При картировании геномной ДНК фага Т4 получена кольцевая генетическая карта (рис. 1), а размер уникальной ДНК составляет примерно 169 т.п.н. При этом содержание ГЦ-пар в геноме Т4 составляет всего лишь 34,5%, в то время как у его хозяина – *E. coli* – около 50%. Повидимому, АТ-богатый геном фага Т4 дает некоторые преимущества вирусу, поскольку структура его ДНК отличается от В-формы ДНК клетки-хозяина (Nelson et al., 1987) и наиболее близка к D-форме ДНК. Эта структурная разновидность ДНК представляет собой полимер (dA-dT) с восемью основаниями на один оборот спирали, с неглубоким и широким большим желобком, более глубоким и узким малым желобком (Leslie et al., 1980).

В настоящее время основные этапы жизненного цикла фага Т4 хорошо охарактеризованы и накоплена обширная информация, посвященная этому вопросу, которая наиболее полно отражена в книге «Молекулярная биология бактериофага Т4» под редакцией Карамы (Karam,

1994). Развитие бактериофага Т4 считается примером временного контроля экспрессии генов. Данный тип регуляции осуществляется факторами транскрипции фага, которые последовательно активируют различные классы промоторов. По времени экспрессии фаговые гены можно подразделить на предранние, ранние (средние) и поздние. Сразу же после инфекции немодифицированная РНК-полимераза *E.coli* начинает транскрипцию предранних генов с промоторов, которые неотличимы от промоторов *E.coli*. Один из новосинтезированных фаговых белков - продукт гена *alc*, связывается с РНК-полимеразой, что приводит к снижению транскрипции цитозин-содержащей ДНК (Snustad et al., 1986). Кроме того, хозяйская РНК-полимераза рибозилируется фаговой АДФ-рибозилтрансферазой *gpAlt*, что приводит к преимущественной транскрипции с ранних фаговых промоторов, богатых А-Т основаниями. Такая ковалентная модификация РНК-полимеразы *E.coli* приводит к изменению как ДНК-белкового, так и белок-белкового взаимодействия. Дальнейшее рибозилирование полимеразы продуктами фаговых генов *modA* и *modB* переориентирует транскрипцию на средние фаговые промоторы. На стадии транскрипции со средних промоторов действуют два фаговых транскрипционных фактора: *AsiA* и *MotA*. Первый из них, связываясь с σ -фактором РНК-полимеразы, активирует ее для узнавания средних фаговых промоторов, а второй является ДНК-связывающим транскрипционным активатором, взаимодействует с промоторами фага Т4. На стадии поздней транскрипции, наряду с другими транскрипционными факторами фага, появляется σ^{55} , который при помощи коактиваторов способствует инициации транскрипции с поздних промоторов фага Т4 (Miller et al., 2003).

Продукты других предранних генов Т-четных фагов модифицируют аппарат трансляции клетки: один из них гидролизует лейциновую тРНК, чьи кодоны редки в фаговых генах. При этом в ходе инфекционного цикла синтезируются несколько фаговых тРНК, кодоны которых с высокой частотой представлены в фаговых генах.

Несколько ранних фаговых белков взаимодействуют вместе и полностью разрушают хромосому хозяйской клетки. Продукт гена *ndd* участвует в разрушении хозяйского нуклеоида, а ингибирование клеточной транскрипции продуктом гена *alc* вызывает разворачивание независимо свернутых доменов хозяйского нуклеоида. Нуклеазы фага Т4 деградируют ДНК *E.coli* до мононуклеотидов, используя присутствие цитозиновых остатков при распознавании хозяйской ДНК. За деградацию цитозин-содержащей ДНК ответственны эндонуклеаза II (продукт гена *denA*) и эндонуклеаза IV (продукт гена *denB*) (Warner, 1971; Vetter, Sadowski, 1974).

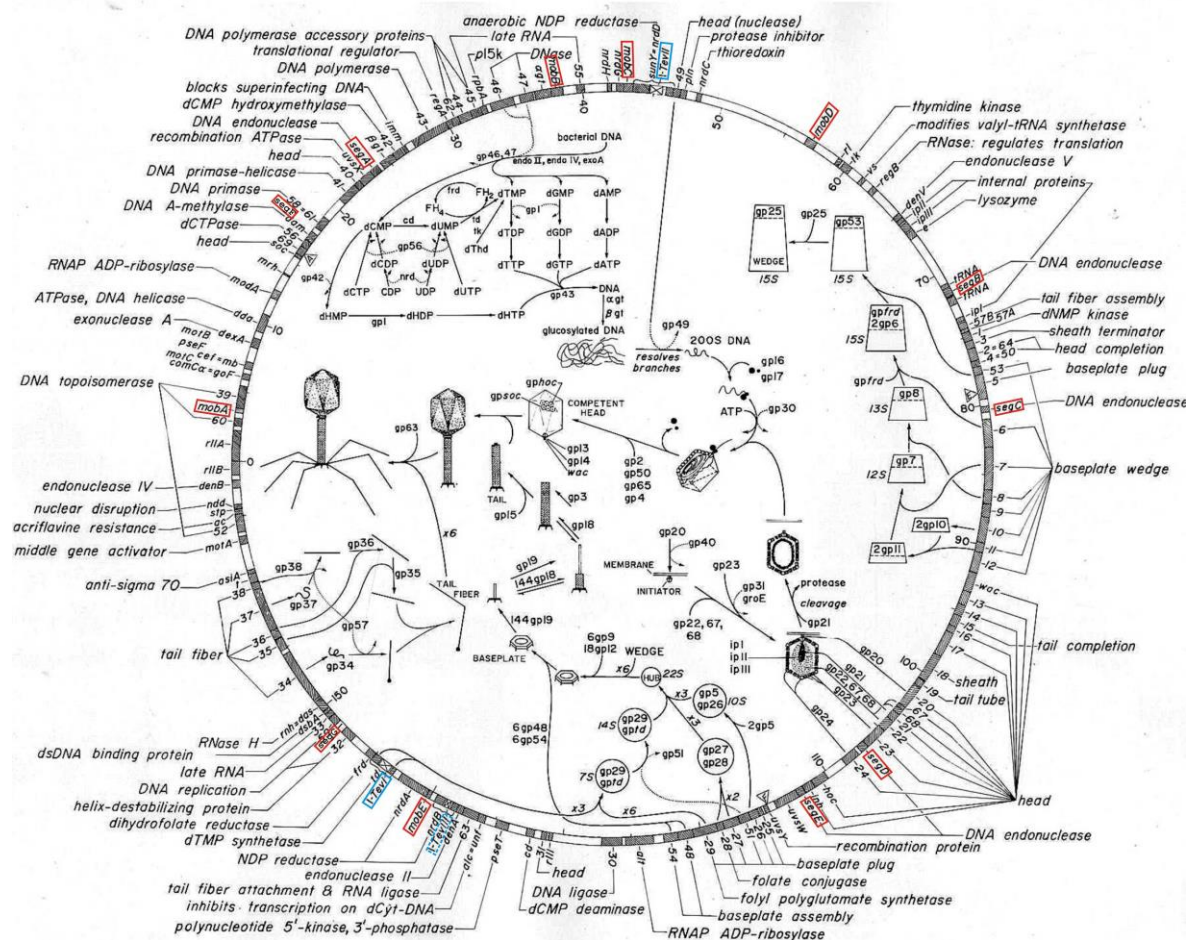


Рис. 1. Генетическая карта фага Т4. Рамками красного цвета выделены гены свободно-стоящих сайт-специфических эндонуклеаз семейства Seg и предполагаемых сайт-специфических эндонуклеаз семейства Mob. Рамками синего цвета выделены гены хоминг-эндонуклеаз, кодируемые интронами.

Большая группа фаговых белков, синтезируемых впервые несколько минут инфекции, представлена белками, которые реплицируют фаговую ДНК: ДНК-полимераза (продукт гена 43), вспомогательные белки ДНК-полимеразы (продукты генов 45 и 44/62), ДНК-связывающий белок (продукт гена 32), топоизомераза (продукты генов 39, 52 и 60), ДНК-лигаза (продукт гена 30), РНКазы H и несколько хеликаз. Взаимодействие белков в репликационной вилке хорошо изучено благодаря работам, выполненным в лабораториях Альбертса и Носсал. Синтез ДНК фага Т4 является уникальным, в связи с использованием двух различных механизмов инициации репликации: ранним - ориджин-зависимым и более поздним - рекомбинационно-зависимым. Рекомбинационно-зависимый процесс инициации репликации может происходить как с участием продукта гена 49, так и независимо от этой эндонуклеазы, расщепляющей ДНК в структурах Холлидея. По-видимому, различные механизмы инициации репликации необходимы для поддержания высокой скорости синтеза фаговой ДНК. Ориджин-зависимая инициация репликации происходит в нескольких промоторных сайтах с вовлечением активности РНК-полимеразы, которая синтезирует РНК-затравки. Рекомбинационно-зависимая инициация требует активности нескольких фаговых белков, которые также участвуют в репарации и рекомбинации фаговой ДНК: *UvsX*, *UvsY*, *UvsW*. Модель рекомбинационной репликации ДНК фага Т4 была предложена Мозиг и основана на предположении, что нереплицированные 3'-концы линейной ДНК могут внедряться в дуплексные регионы соседних молекул и служить праймерами для дальнейшего синтеза ДНК. В результате такого необычайно высокого уровня множественной рекомбинации между хромосомными концами и гомологичными последовательностями в соседних молекулах происходит образование сильно разветвленных мультимеров, содержащих от 20 до 100 эквивалентов фаговых молекул ДНК (Mosig, 1994).

1.1.3 Явление частичного исключения генетических маркеров при скрещивании фагов Т4-типа

Дельбрюк с соавторами, а позже и ряд других авторов, на примере генов хозяйской специфичности показали, что при совместной инфекции клеток *E.coli* бактериофагами Т4 и Т2, полученное потомство фагов преимущественно наследует гены бактериофага Т4. (Delbruck, 1945; Streisinger and Weigle, 1956). Они выяснили, что данное явление носит более сложный характер. Рассел и Хаски (Russell and Huskey, 1974) проанализировали наследование более 70 генетических маркеров и установили, что маркеры фага Т4 преобладают в фаговом потомстве, полученном при совместных инфекциях Т4 и Т2 (рис.2). Этот результат был весьма неожиданным, если учесть тот факт, что фаги Т4 и Т2 являются близкими Т-четными фагами.

Первоначально предполагалось, что генетические маркеры, принадлежащие этим близким фагам, будут наследоваться равновероятно. Расселл и Хаски назвали этот процесс частичным исключением генетических маркеров. При данном процессе гены фага T2 наследуются с разной вероятностью (от 10% до 40%), в зависимости от конкретного генетического локуса, в котором располагаются анализируемые гены (рис.2). Кроме того, в геноме фага T2 были обнаружены две области, расположенные около генов 32 и 56 (рис. 2), в которых происходит практически полное исключение маркеров фага T2. В отсутствии какой-либо информации о последовательности ДНК данных областей и о сайт-специфических ДНК эндонуклеазах Расселл и Хаски предположили, что за явление локального исключения генетических маркеров ответственны белки, способные узнавать определенные последовательности ДНК, данные белки, по всей видимости, и обладают “исключающей” активностью.

На данный момент известно, что области, расположенные около генов 32 и 56 бактериофага T2, соответствуют местам встраивания генов, *segG* и *segF* (Miller et al. 2003). Было показано, что продукты данных генов являются хоминг-эндонуклеазами. Сайты гидролиза SegG и SegF расположены в генах 32 и 56 бактериофага T2, соответственно (Belle et al. 2002; Liu et al. 2003). Было установлено, что двухцепочечный разрыв вызывает генетическую конверсию участков, прилегающих к области гидролиза. Данное событие и приводит к недопредставленности генетических маркеров фага T2 в фаговом потомстве (Liu et al. 2003). В подтверждение данного предположения были проведены работы по выключению генов *segG* и *segF*. Элиминация соответствующих эндонуклеаз приводила к уменьшению исключения генетических маркеров фага T2 фаговым потомством смешанных инфекций. Эти данные подтверждают роль выше упомянутых хоминг-эндонуклеаз в феномене частичного исключения генетических маркеров.

Кроме того, для хоминг-эндонуклеазы SegB, ген которой находится в кластере tPHK, также было показано участие в мобильности собственного гена и генетической конверсии прилегающих областей (Brok-Volchanskaya et al. 2008). Это свидетельствует в пользу того, что и данная хоминг-эндонуклеаза также может быть задействована в феномене исключения генетических маркеров фага T2 при смешанной инфекции. Аналогичные данные были получены и для хоминг-эндонуклеазы SegE, но при скрещивании фагов T4 и RB30 (Kadyrov et al. 1997). Таким образом, учитывая вышеизложенные факты, в научном сообществе сложилось мнение, что за феномен исключения генетических маркеров могут быть ответственны хоминг-эндонуклеазы, открытыми рамками считывания которых богат геном фага T4. По-всей видимости, при протекании процесса хоминга происходит и перенос генетических областей, прилегающих к генам хоминг-эндонуклеаз, что и влияет на представленность прилегающих участков в фаговом потомстве (Belford. M 2005).

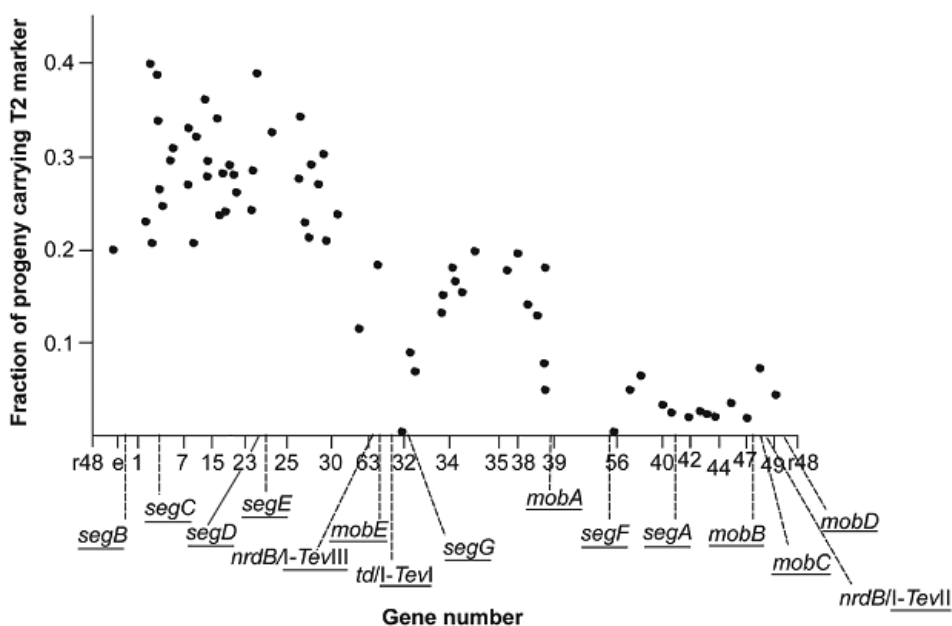


Рис. 2. Представленность генетических маркеров T2 в фаговом потомстве, полученном после смешанных инфекций T4 и T2. На рисунке показана доля наследования генетических маркеров T2 в различных генетических локусах. Цифрами показано расположение соответствующих генов, также показано расположение генов хоминг-эндонуклеаз. Точка отражает величину наследования фаговым потомством соответствующего генетического маркера T2. Рисунок адаптирован из книги *Homing endonucleases and inteins*. 2005.

1.2. Хоминг-эндонуклеазы

Хоминг-эндонуклеазы относятся к особому классу сайт-специфических эндонуклеаз. Характерным свойством ферментов этого класса является их способность узнавать вырожденные, протяженные (14-40 п.о.), часто асимметричные, последовательности ДНК. Данный класс эндонуклеаз широко распространен и встречается у бактериофагов, эубактерий, архей, а также у простейших эукариотических организмов, в частности у дрожжей и одноклеточных водорослей (Guhan and Muniyappa, 2003, Belfort et al., 2014).

Гены хоминг-эндонуклеаз были найдены в составе интронов -I, -II группы, архейных интронов и интеинов, а также в виде свободностоящих генов. Основной функцией хоминг-эндонуклеаз является участие в мобильности собственных генов. Хоминг-эндонуклеазы вносят разрыв в ДНК мишень (родственный аллель, не содержащий ген хоминг-эндонуклеазы) и запускают тем самым процесс переноса генетической информации. Этот процесс получил название хоминга (от английского слова “homing” – возвращение домой). Таким образом, хоминг – это вариант генетической конверсии, результатом которого является перенос в

родственный аллель мобильной последовательности ДНК, кодирующей ген хоминг-эндонуклеазы (Stoddard, B., 2005).

На основании ключевых этапов переноса гена хоминг-эндонуклеазы, принято выделять две разновидности хоминга (хоминг и ретрохоминг).

Хоминг

Процесс переноса интронов I группы, интеинов, а также свободностоящих ОРС инициируется посредством эндонуклеазной активности соответствующей хоминг-эндонуклеазы. Завершается процесс переноса генетической информации, ферментами метаболизма ДНК клетки. Данные белки восстанавливают целостность поврежденной ДНК, посредством гомологичной рекомбинации, используя в качестве матрицы родственный аллель, содержащий ген хоминг-эндонуклеазы (рис. 3А) (Colleaux et al. 1986; Dujon, 1989; Belfort and Perlman, 1995).

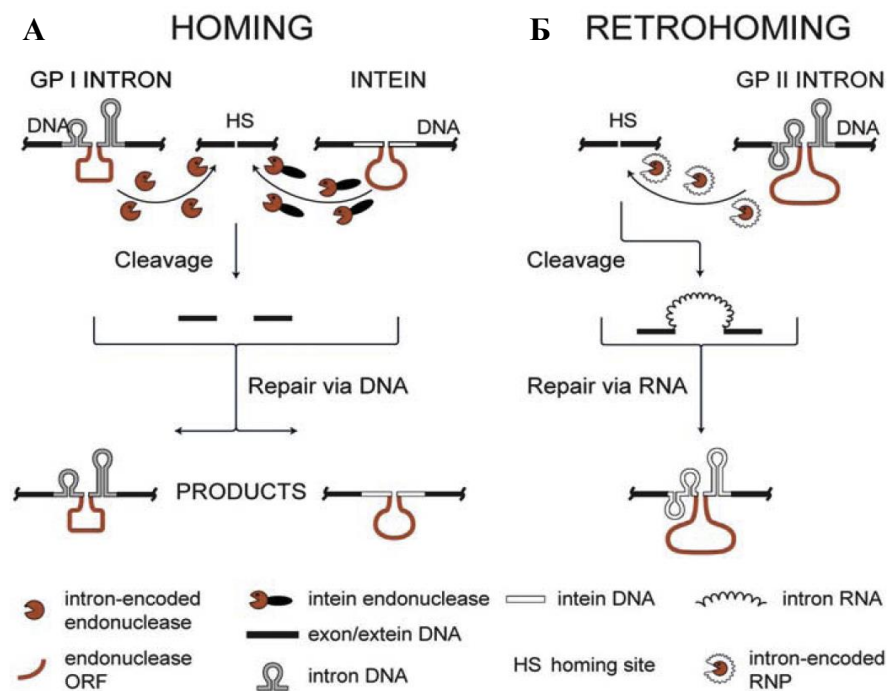


Рис. 3. Основные этапы перемещения мобильных интронов и интеинов. А. Основные этапы переноса интронов I группы, интеинов, а также свободностоящих ОРС. Процесс переноса генетической информации инициируется хоминг-эндонуклеазой. Завершается процесс ферментами репарации ДНК. **Б.** Основные этапы ретрохоминга. Одна из цепей ДНК, содержащая сайт узнавания, гидролизуется РНК компонентом комплекса, вторая цепь ДНК гидролизуется белковым компонентом белок-рибонуклеинового комплекса. Далее происходит встраивание интронной РНК, образование кДНК и синтез второй цепи ДНК. Рисунок адаптирован из статьи Lambowitz et al., 2005.

Ретрохоминг

Хоминг в случае мобильных интронов II группы требует иной последовательности реакций. Данный процесс включает в себя: расщепление ДНК реципиентного аллеля, реакцию обратного сплайсинга, генерирующую ДНК-РНК интермедиат (РНК кодирует последовательность хоминг-эндонуклеазы и интрона), и, в завершении, реакцию обратной транскрипции с использованием встроенной РНК (Stoddard, 2005). Эта цепь последовательных реакций катализируется единственным многофункциональным белком, открытая рамка считывания которого входит в состав мобильного интрона II группы. Данный тип перемещения был назван ретрохомингом, поскольку на одном из этапов процесса используется молекула РНК и реакция обратной транскрипции (рис. 3Б).

Таким образом, гены хоминг эндонуклеа, либо независимо, либо в составе мобильных последовательностей интронов, копируются в последовательность ДНК мишеней (Chevalier and Stoddard, 2001).

Основы номенклатуры, ранее разработанные для эндонуклеаз рестрикции, были переработаны и адаптированы для хоминг-эндонуклеаз. Хоминг-эндонуклеазы обозначают трехбуквенным названием, заглавная первая буква обозначает название рода, две оставшиеся являются производными от названия вида. Римские цифры, следующие после трехбуквенного названия хоминг-эндонуклеазы, используются для разграничения ферментов, найденных в одном организме, и присваиваются в хронологическом порядке. В название хоминг-эндонуклеаз также входят префиксы I-, PI- или F- которые указывают на расположение гена х-э в составе интрона (Intron), интеина (Protein Insert) либо виде свободностоящей ОРС (Freestanding), соответственно. В качестве примера можно привести названия трех хоминг-эндонуклеаз *Saccharomyces cerevisiae* I-SceI, PI-SceI, and F-SceII располагающихся в составе интрона, интеина и в виде свободностоящей ОРС, соответственно. Кроме того, в отношении предполагаемых хоминг-эндонуклеаз, нуклеазная активность которых еще не была установлена, используется суффикс P (Putative). После того как, будет показана эндонуклеазная активность суффикс (P) из названия удаляется (Roberts et al., 2003; Guhan and Muniyappa, 2003).

Открытые рамки считывания, кодирующие хоминг-эндонуклеазы, в настоящий момент, относят к гетерогенной группе мобильных элементов. При недавнем пересмотре номенклатуры мобильных генетических элементов было выдвинуто предложение все мобильные элементы обозначать аббревиатурой Tn и соответствующим номером. По всей видимости, в ближайшее время известным генам хоминг-эндонуклеаз будет, помимо названия, приписан номер в создаваемой общей базе мобильных генетических элементов (Roberts et al.2008).

Однако, помимо общепринятых названий, удовлетворяющих правилам номенклатуры, существуют и широко используются, исторически сложившиеся. У T-четных фагов была

найдена группа свободно стоящих генов, кодирующих белки, содержащие GIY-YIG-мотив. Данный мотив ранее был обнаружен у интрон-кодируемой хоминг-эндонуклеазы I-TevI. В связи с этим они были названы Seg белками (*s*imilar to *e*ndonucleases of *g*roup I *i*ntron) и обозначены буквами A, B, C, D, E, F, G и H. Данная группа белков характеризуется гомологией N-концевых фрагментов молекул, и наличием вырожденного мотива GIY-YIG (Sharma et al. 1992; Sandegren et al. 2005).

Кроме того, у бактериофага T4, помимо представителей GIY-YIG-семейства хоминг-эндонуклеаз, были обнаружены и представители HNH-семейства. Данная группа белков, также вопреки номенклатурным правилам, была названа Mob белками (от слова *mobility*) (Miller et al. 2003)

Хоминг-эндонуклеазы в соответствии с присутствием в молекуле белка определенных аминокислотных мотивов делятся на шесть семейств H-N-H, LAGLIDADG, His-Cys, GIY-YIG PD-(D/E)xK, и EDxHD (Guhan and Muniyappa, 2003; Stoddard, 2005, Taylor, 2011, Stoddard, 2014). Кроме того, высказываются предложения объединить семейства His-Cys и H-N-H в единое суперсемейство $\beta\beta\alpha$ -Me на основании общей укладки каталитического центра (Kuhlmann et al., 1999; Guhan and Muniyappa, 2003). Для всех семейств показано, что аминокислоты, входящие в консервативный аминокислотный мотив, выполняют структурную или функциональную роль (связывание двухвалентных катионов, гидролиз фосфодиэфирной связи) (Stoddard., 2005). Таким образом, в пределах этих белковых семейств, отдельные представители хоминг-эндонуклеаз могут быть разграничены в соответствии с присутствием легко различимых консервативных аминокислотных последовательностей.

Не так давно были найдены две группы белков с аминокислотными мотивами PD-(D/E)xK и EDxHD (Stoddard, 2014, Taylor, 2011). Так в интроне гена $tRNA^{fMet}$ у *Synechocystis* PCC6803 была обнаружена хоминг-эндонуклеаза с характерным для эндонуклеаз рестрикции аминокислотным мотивом PD-(D/E)xK. Эти данные, могут свидетельствовать в пользу существования некоторой эволюционной связи между хоминг-эндонуклеазами и эндонуклеазами рестрикции (Zhao et al., 2007; Orłowski et al., 2007).

1.2.1. Описание хоминг-эндонуклеаз

Хоминг-эндонуклеазы схожи с эндонуклеазами рестрикции, по способности вносить разрывы в молекулу ДНК-мишени. Поэтому целесообразно вести сравнение хоминг-эндонуклеаз с хорошо изученными эндонуклеазами рестрикции. Так оказалось, что эти группы белков существенно различаются по ряду ключевых аспектов, в частности: по функции,

выполняемой в клетке, по особенностям узнавания ДНК-мишени, а также по локализации генов.

Особенности узнавание сайта гидролиза хоминг-эндонуклеазами и эндонуклеазами рестрикции.

Хоминг-эндонуклеазы, в отличие от эндонуклеаз рестрикции II типа, способны специфически узнавать протяженные часто вырожденные последовательности ДНК. Данная особенность обеспечивается уникальной стратегией узнавания сайтов гидролиза хоминг-эндонуклеазами. Специфичность узнавания последовательностей ДНК ферментом определяется количеством специфических связей, которые способен сформировать фермент с основаниями ДНК напрямую, либо посредством молекул воды (Chevalier et al., 2003). Отдельная нуклеотидная пара сайта узнавания способна образовать до 4 водородных связей с аминокислотными остатками фермента, хотя для распознавания нуклеотидной пары достаточно специфическое формирование двух водородных связей (Chevalier et al., 2003).

Сайты узнавания рестриктаз представляют собой полиндромные нуклеотидные последовательности длиной от 4 до 8 п.н.о. Эндонуклеазы рестрикции формируют максимально возможное количество связей с узнаваемым фрагментом дезоксирибонуклеиновой кислоты. Эндонуклеаза рестрикции EcoRV образует контакты с фрагментом ДНК из 6 п. н., формируя в общем 22 водородные связи, в данном случае на одну пару нуклеотидов приходится в среднем 3,5 связи (Kostrewa and Winkler, 1995; Winkler et al., 1993;). Такая прочная связь фермента с каждой парой оснований лежит в основе того, что сайты узнавания эндонуклеаз рестрикции не вырождены, и замена любой нуклеотидной пары в сайте узнавания приводит к утрате значимого количества специфических связей между сайтом узнавания и эндонуклеазой. Это приводит к неспособности эндонуклеаз взаимодействовать с измененными сайтами узнавания.

В отличие от эндонуклеаз рестрикции, хоминг-эндонуклеазы, несмотря на их, относительно малые размеры (< 40 кДа), взаимодействуют с протяженными участками ДНК длиной от 14 до 40 п.н. При этом используется, как правило, лишь часть связей, которые потенциально могут быть сформированы с этим участком дезоксирибонуклеиновой кислоты. Так, например, I-CreI взаимодействует с фрагментом ДНК длиной 24 п. н., посредством образования 24 специфических контактов, или в среднем 1 связь на пару оснований (Chevalier and Stoddard, 2001; Акуленко Н. В. кандидатская диссертация). I-HmuI формирует 14 водородных контактов для распознавания фрагмента ДНК длиной 25 п.н. (Shen et al., 2004). Наименьшее количество водородных связей образует I-TevI: только 10 связей на фрагмент ДНК длиной 20 п. н. (Van Roey et al., 2001). Данная стратегия формирования водородных связей с сайтом узнавания объясняет феномен их вырожденности для хоминг-эндонуклеаз. Другими словами, большинство хоминг-эндонуклеаз не чувствительны к замене одного или более

нуклеотидов в узнаваемой ими последовательности ДНК, поскольку при этом не происходит существенного уменьшения количества специфических контактов. Толерантность отдельных хоминг-эндонуклеаз к изменениям сайта узнавания различна. Так, эндонуклеаза I-SceI узнает последовательность ДНК длиной 18 нуклеотидных остатков, 14 из которых являются консервативными. Для эндонуклеазы I-ProI показано, что только 4 из 15 нуклеотидных остатков в последовательности сайта узнавания являются строго консервативными (Argast et al., 1998; Акуленко Н. В. кандидатская диссертация). В тоже время, эндонуклеаза I-TevI фара Т4 толерантна к замене практически любого нуклеотида в сайте узнавания (Bryk, 1993). Кроме того, для большинства ферментов, при существовании нескольких сайтов гидролиза, характерна иерархия в узнавании последовательностей ДНК. Это позволяет говорить о наличии предпочтительных сайтов гидролиза для хоминг-эндонуклеаз (Brok-Volchanskaya et al., 2008).

Следует остановиться еще на одной уникальной особенности характерной для хоминг-эндонуклеаз, а именно, расположении сайтов гидролиза данных ферментов. Зачастую хоминг-эндонуклеазы узнают и гидролизуют последовательности ДНК, расположенную в существенных генах. Это свойство позволяет максимально расширить ареал распространения генов хоминг-эндонуклеаз как среди родственных, так и относительно далеко отстоящих эволюционно геномов. Такая стратегия выбора мишени гидролиза, по-видимому, эволюционно является более выигрышной, поскольку такие последовательности ДНК, как правило, более консервативны. Так, значительное количество генов хоминг-эндонуклеаз было обнаружено в составе интронов, ассоциированных с генами рибосомальных РНК, а также в составе генов тРНК (Lambowitz and Belfort, 1993). И это вполне понятно, поскольку гены рибосомальных РНК и тРНК консервативны и встречаются у всех представителей живой материи.

Функции эндонуклеаз рестрикции и хоминг-эндонуклеаз, выполняемые в клетке.

Особенности взаимодействия сайт-специфических эндонуклеаз с ДНК-мишенью связаны с выполняемой биологической ролью данных ферментов. Узнавание эндонуклеазами рестрикции коротких последовательностей ДНК необходимо для защиты клетки от проникновения посторонней дезоксирибонуклеиновой кислоты. Поэтому для увеличения вероятности встречи сайта гидролиза в “чужой” ДНК необходима минимизация сайта узнавания. Однако уменьшение длины сайта узнавания чревато уменьшением количества специфических контактов между эндонуклеазой и сайтом узнавания. А это может привести к уменьшению специфичности, а как следствие увеличению вероятности неспецифического гидролиза хозяйского генома.

Хоминг-эндонуклеазы, участвующие в латеральном переносе последовательностей интронов, интеинов и свободностоящих собственных ОРС, балансируют между двух полярных

качеств. С одной стороны, им необходимо обладать устойчивостью к небольшим заменам в сайтах узнавания расположенных в родственных геномах. С другой стороны, хоминг-эндонуклеазы не должны расщеплять ДНК хозяина, поэтому вынуждены быть специфичны. Сочетание этих особенностей возможно только в случае узнавания хоминг-эндонуклеазами протяженных вырожденных последовательностей ДНК (Guhan and Muniyappa, 2003).

Локализация генов хоминг-эндонуклеаз и эндонуклеаз рестрикции.

Филогенетическое распространение двух классов эндонуклеаз, различно. Хоминг-эндонуклеазы найдены у архей, бактерий, и эукариот, тогда как эндонуклеазы рестрикции не встречаются в эукариотических геномах. Кроме того, локализация ОРС хоминг-эндонуклеаз внутри генома более разнообразна. Гены хоминг-эндонуклеаз были обнаружены внутри интронов (-I, -II и -III группы, архейных интронов) и интеинов, а также в виде свободностоящих ОРС. Так, по некоторым предварительным оценкам 25-30% интронов I и II группы содержат в своем составе ОРС, значительная часть из которых кодирует хоминг-эндонуклеаз, и, следовательно, эти интроны обладают потенциалом к перемещению в родственные геномы (Dujon, 1989; Belfort & Perlman, 1995; Curcio & Belfort, 1996; Chevalier & Stoddard, 2001; Stoddard 2005).

Причины ассоциации генов хоминг-эндонуклеаз с самосплайсирующимися элементами (интронами и интеинами) вполне очевидны. Самосплайсирующие элементы, при ассоциации с генами хоминг-эндонуклеаз, приобретают потенциал к распространению среди родственных геномов, в результате процесса хоминга, вместе с генами хоминг-эндонуклеаз. Преимущества, которые получают хоминг-эндонуклеаз в такой связке (ген хоминг-эндонуклеазы + самосплайсирующий элемент) также очевидны, поскольку при такого рода взаимодействии, ОРС, кодирующие хоминг-эндонуклеазы, могут непосредственно встраиваться в последовательность хозяйских генов без нарушения функции гена-мишени. В случае интронов, последовательность, кодирующая функциональный ген мишень, восстанавливается на уровне РНК, в процессе сплайсинга. В случае интеинов процесс сплайсинга проходит на уровне белка, и в конечном итоге также приводит к сохранению функциональной активности гена мишени (Belfort, 2005, Belfort 2014).

В отличие от хоминг-эндонуклеаз, гены эндонуклеаз рестрикции являются свободностоящими и почти всегда находятся рядом с генами, кодирующими родственные ферменты модификации.

1.2.2. Описание различных групп хоминг-эндонуклеаз.

Как упоминалось ранее, хоминг-эндонуклеазы классифицируются по признаку наличия определенной консервативной аминокислотной последовательности в структуре молекулы. По этому признаку различают четыре семейства хоминг-эндонуклеаз, названия которых соответствуют однобуквенным обозначениям аминокислот, присутствующих в консервативном участке белка. Это такие семейства как HNH, LAGLIDADG, His-Cys box, и GIY-YIG, соответственно (рис. 4) (Chevalier, Stoddard, 2001; Stoddard 2005).

A

A LAGLIDADG MOTIF

P1			P2		
L A G L I D A D G			L A G L I D A D G		
[35]	A G I G L I L G D A Y I	[89]	S L A Y W F M D D G G K	[87]	I-SceI
[86]	W L A G L I D G D G Y F	[94]	W F V G F F D A D G T I	[112]	I-SceII
[30]	Y L A G L I E G D G S I	[120]	W L A I L T D A D G N F	[160]	I-SceIII
[101]	I M T G I L L T D G W I	[98]	S L A H M I M C D G S F	[83]	I-SceIV
[42]	I A V G L L L S D A H A	[91]	A L A Y W I A G D G C W	[80]	I-CsmI
[11]	Y L A G F V D G D G S I	[140]			I-CreI
[9]	L I F G S L L G D G N L	[99]	A L A Y F Y I D D G A L	[86]	I-ChuI
[57]	F L A G F L E G E A S L	[149]			I-CeuI
[12]	Y L L G L I I G D G G L	[84]	F I K G L Y V A E G D K	[74]	I-DmoI
[5]	Y V K A L S A F D G Y V	[67]	L L A A A I D A E G N V	[78]	I-PorI
[209]	Y L L G L W I G D G L S	[96]	F L A G L I D S D G Y V	[125]	PI-SceI
[146]	E L V G L I V G D G N W	[86]	F L R G L F S A D G T V	[134]	PI-T7II
[280]	K L L G Y Y V S E G Y A	[83]	F L E A Y F T G D G D I	[151]	PI-T7III
[279]	K L L G Y Y V S E G S A	[83]	F L E G Y F I G D G D V	[151]	PI-PspI
[213]	W M L G L W L G D G T T	[99]	F L A G L I D S D G Y V	[250]	H0 Endo
[202]	Y L S G L I E G D G Y I	[105]	W L A G F T A A D G S F	[145]	Endo.SceI
	(1) g(I)G g(D)g		1 (D)G		Consensus

B GIY-YIG MOTIF

G I Y			Y I G		
[2]	S G I Y Q I K N T L N N K V Y V G S	-S F N K H G N V . F E C S I L E	[183]	I-TevI	
[48]	T G V Y I W I N K E N G K R Y V G S	-A L A K Y G F E S F R L V L I L	[253]	Am atpase 6	
[9]	N Y T Y V I T N L V N N K I Y Y G T	-A K K K Y G K K N F N L S I L G	[158]	SegA	
	g Y i N n n K Y v G s a K y G F i s i (L)			Consensus-3	
	g(1)Y n Y(1)G s (1) k g f (1)I			Consensus-17	

C H-N-H MOTIF

H H			N			H H		
[163]	V G G R R S F E L H H D K P I S Q D G G V Y D M D N L R I T T P K . R H I D I H	[3]	Eco CoE8					
[163]	V G G R K V Y E L H H D K P I S Q G G E V Y D M D N I R V T T P K . R H I D I H	[3]	Eco CoE9					
[539]	V G G R E R F E L H H D K P I S Q D G G V Y D M N N I R V T T P K . R H I D I H	[3]	Eco CoE2					
[206]	Q E N Q [14] L E V H H V I P L S G G A D T T . D N C V A L Q P N . Q H R E L H	[21]	Eco Mcr					
[64]	E G Y E E G L V V D H K D G N K . . D N N L S . T N L R W V T Q K . I N V E N Q	[74]	I-HmuI					
[16]	D S D G R T D E I H H K D G N R . . E N N D L . D N L M C L S I O . E H Y D I H	[217]	I-TevIII					
[520]	Q S H Q [9] I E I D H I I P K S Q G G K D V Y . D N L Q A L Q R H . Q H D V K T	[19]	CpcI gpII					
[]	Q S E Q [9] M E V H H I D Q N R . . G N N K L . S N L T L V Q R H . Q H D I I H	[]	Cpc2 gpII					
[]	Q P M Q [9] W N I H H I I K R H M G G G D E L . D N L V L L Q P N . Q H R Q L H	[]	Avi gpII					
[524]	Q G L Q [10] Y E L H H I L P K R F G G K D T P . N N M V L L Q K S P Q H Q L V S	[37]	Sob gpII					
[723]	Q Q I Q [5] L E V H H V R T L N N A [15] M N R [4] I Q K T . Q H F K V H	[9]	Sce gpII					
	e(1)h H g N(1) (1) h (1)h		Consensus					

D His-Cys BOX

H C			C			H C		
[169]	H S S H L C K G D G S C M E L K H T L R V P A Q T N L A D H E L C P	[58]	I-DirI					
[176]	T I S H L C . G N G G C A R P G H . L R I E K K T V N D E R T H C H	[37]	I-NaaI					
[116]	T A S H L C . H N T R C H N P L H . L C W E S L D D N K G R N W C P	[12]	I-PpoI					
	t S H L C g n g C p H L r e t n r C p		Consensus					

Б.

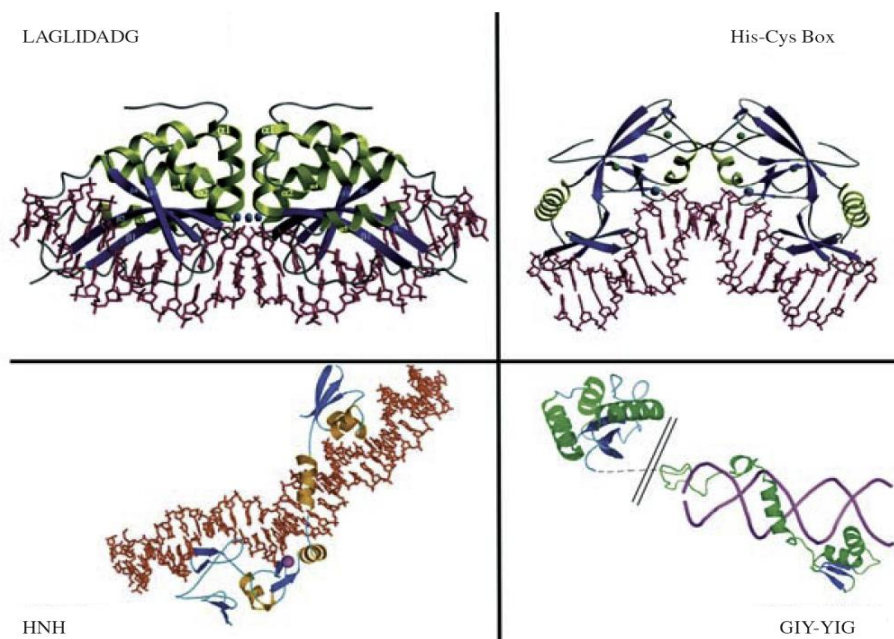


Рис. 4. На рисунке А представлено аминокислотное выравнивание ряда представителей основных семейств хоминг-эндонуклеаз. Также на рисунке выделены ключевые аминокислотные остатки. На рисунке Б представлена пространственная структура четырех представителей хоминг-эндонуклеаз. Показана структура I-CreI (семейство LAGLIDADG гомодимер; сайт узнавания 22-п.о.), I-PpoI (семейство His-Cys гомодимер; сайт узнавания 14-п.о.), I-HmuI (фаговая HNH хоминг-эндонуклеаза; сайт узнавания 26-п.о.) и I-TevI (хоннг-эндонуклеаза семейства GIY-YIG; сайт узнавания 37-п.о.). Взято из обзора Stoddard B., 2005. Рис. А взят из статьи Belfort M., 1995

Семейство LAGLIDADG

Хоминг-эндонуклеазы семейства LAGLIDADG являются наиболее изученными к настоящему моменту. Они содержат один или два аминокислотных мотива LAGLIDADG на молекулу белка, данные аминокислотные мотивы находятся непосредственно в активном центре и необходимы для гидролиза ДНК. Трехмерная структура ряда представителей данного семейства, в частности эндонуклеаз I-CreI, PI-SceI, I-DmoI и PI-PfuI, к настоящему моменту определена с использованием методов рентгеноструктурного анализа (Heath et al., 1997; Duan et al., 1997; Silva et al., 1999; Ichiyanagi et al., 2000). К ферментам семейства LAGLIDADG относится, также, и хорошо изученная *HO*-эндонуклеаза, участвующая в переключении типа спаривания у дрожжей.

Другим хорошо изученным представителем данного семейства является хоминг-эндонуклеаза I-CreI. Она содержит один мотив LAGLIDADG на субъединицу белка и активна в виде гомодимера (Wang et al., 1997). При этом LAGLIDADG мотивы образуют два

каталитических центра, которые расположены в непосредственной близости друг от друга. Каждый центр содержит ион Mg^{2+} , который связан консервативным остатком аспарагиновой кислоты (Asp20) и фосфатом, по которому происходит разрыв цепи ДНК. Третий ион Mg^{2+} соединяет оба каталитических центра между собой (Chevalier et al., 2001). При селекции сайтов узнавания для эндонуклеазы I-CreI была обнаружена одна интересная особенность – связывание фермента с ДНК-мишенью было столь прочным, что эффективное высвобождение продуктов гидролиза происходило только в присутствии SDS.

Другим хорошо охарактеризованным ферментом, относящимся к этому семейству, является эндонуклеаза PI-SceI, которая была обнаружена в составе интеина (белкового интрона). Аминокислотная последовательность интеина является составной частью белка-предшественника дрожжевой вакуолярной АТФазы и вырезается из его состава в процессе белкового сплайсинга (Colston and Davis, 1994). Интеины не только имеют сайт-специфическую эндонуклеазную активность, обеспечивающую мобильность кодирующей последовательности интеина (Gimble and Thorner, 1992), но и участвуют в белковом сплайсинге, с помощью которого они выщепляются из белка-предшественника. PI-SceI также является бифункциональным белком, который, наряду с эндонуклеазной активностью, способен осуществлять сплайсинг белка. Фермент имеет двухдоменную организацию, которая характерна для интеин-кодируемых хоминг-эндонуклеаз: домен I отвечает за сплайсинг, домен II обладает эндонуклеазной активностью. Примечательным является тот факт, что две α -спирали, содержащие мотивы LAGLIDADG в активном центре PI-SceI, формируют структуру схожую с обнаруженной в димере I-CreI (Heath et al., 1997). Как было показано, эндонуклеаза PI-SceI гидролизует каждую из цепей ДНК последовательно и с различной скоростью: фермент предварительно расщепляет только одну из цепей, а затем вторую, при этом скорость гидролиза второй цепи значительно выше (Christ et al., 1999).

Семейство His-Cys

His-Cys мотив, характерный для одноименного семейства хоминг-эндонуклеаз, представляет собой последовательность размером около 30 аминокислот, которая содержит два консервативных остатка гистидина и три консервативных остатка цистеина. Эндонуклеаза I-PpoI – наиболее хорошо охарактеризованный представитель His-Cys семейства. Это относительно небольшой белок размером 18 kDa, работающий как глобулярный гомодимер (Wittmayer and Raines, 1996). Определение пространственной структуры I-PpoI показало, что His-Cys участок является важным функциональным и структурным элементом. Было установлено, что данная аминокислотная последовательность участвует в формировании каталитического центра и двух цинк-связывающих доменов, состоящих из остатков цистеина и

гистидина. Как было показано, *I-PpoI* способна специфично связываться с ДНК-мишенью в отсутствие двухвалентных катионов, при этом эндонуклеазная активность проявлялась в присутствии ионов Mg^{2+} . Узнавание ДНК-мишени эндонуклеазой *I-PpoI* происходит по малой и большой бороздкам. Связывание фермента с субстратом приводит к изгибанию молекулы ДНК и, как следствие, к значительным деформациям малой и большой бороздок в области разрываемых фосфатных групп (Galburt et al., 2000).

Семейство H-N-H

H-N-H эндонуклеазы содержат консенсусную последовательность протяженностью от 30 до 33 аминокислотных остатков, в составе которой находятся две пары консервативных гистидинов, фланкирующих консервативный аспарагин. Этот участок полипептидной цепи формирует структуру, подобную цинковым пальцам. H-N-H семейство объединяет в себя ферменты с различными свойствами. Так, например, эндонуклеаза *I-TevIII* фага RB3, как и большинство хоминг-эндонуклеаз, вносит двухнитевые разрывы в ДНК (Robbins et al. 2007). В то время как ферменты *I-HmuI* и *I-HmuII* фагов *Bacillus subtilis* гидролизуют только одну цепь ДНК (Goodrich-Blair and Shub, 1996). Еще более интересным примером являются эндонуклеазы *I-SceV* и *I-SceVI* митохондриальных интронов группы II. Данные хоминг-эндонуклеазы образуют комплексы с собственной интронной РНК, при этом в гидролиз ДНК-субстрата вносят вклад оба компонента – белок и РНК. Циммерли с соавторами удалось “разделить” эти две эндонуклеазные активности путем введения мутации в H-N-H мотив *I-SceV*. Введение мутации приводило к нарушению опосредованного белком гидролиза одной из цепей ДНК. При этом гидролиз другой цепи ДНК опосредованный РНК-компонентом происходил нормально (Zimmerly et al., 1995).

Кулман с соавторами провели сравнительно-структурный анализ трех эндонуклеаз – колицина E9, содержащего H-N-H мотив, а также *I-PpoI* и неспецифической нуклеазы из *Serratia*, содержащих His-Cys мотив (Kuhlmann et al., 1999). По результатам сравнения третичных структур была выявлена структурная схожесть активных центров этих ферментов. Исходя из этого, авторы предложили внести изменения в классификацию хоминг-эндонуклеаз и объединить между собой H-N-H и His-Cys семейства.

Семейство GIY-YIG

Группа GIY-YIG хоминг-эндонуклеаз занимает второе место по количеству известных на сегодняшний день представителей. Данная группа белков получила своё название благодаря присутствию характерного аминокислотного мотива GIY- $X_{10/11}$ -YIG. Детальный компьютерный анализ позволил выделить характерные дополнительные аминокислотные мотивы (Kowalski et

al.1999). Было выделено пять мотивов длиной от 7 до 17 а.о. с определенными консервативными аминокислотными остатками (рис. 5). Данные аминокислотные мотивы формируют уникальный каталитический модуль, состоящий из 70-100 а.о. Мотив А включает GIY и YIG элемент, который дал название данной группе белков. Мотив В и С включает консервативные Arg и Glu, задействованные в катализе. Мотив D и E включают консервативные Glu и Asp, соответственно (рис.5).

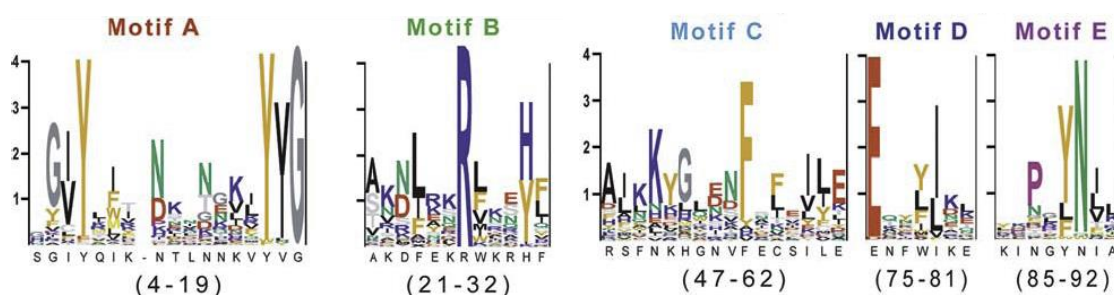


Рис. 5. Представлены основные аминокислотные мотивы, формирующие GIY-YIG каталитический домен. Последовательность аминокислотных мотивов соотнесена с последовательностью каталитического домена I-TevI (внизу приведена аминокислотная последовательность каталитического домена I-TevI). Рисунок адаптирован из статьи Kowalski et al.1999.

По всей видимости, GIY-YIG - домен оказался удобным каталитическим модулем и поэтому обнаруживается у различных белков общим, для которых является расщепление ДНК. В результате комбинаторных событий GIY-YIG каталитический домен может быть связан с различными ДНК-связывающими доменами тем самым, формируя белки с разной специфичностью (Derbyshire et al. 1997). Подтверждением этого является обнаружение GIY-YIG мотива в составе значительного числа ферментов задействованных в метаболизме ДНК. К их числу относятся: хоминг-эндонуклеазы, белки ретроэлементов семейства Пенелопа, ряд белков системы рестрикции-модификации (например, R.Eco29kI (Ibryashkina et al. 2007)), а также белки семейств Cho и UvrC задействованные в репарационных процессах и др. (Dunin-Horkawicz et al. 2006).

На сегодняшний день, члены семейства GIY-YIG хоминг-эндонуклеаз были обнаружены как в составе интронов I группы, так и в виде свободностоящих открытых рамок считывания. Белки SegA-G бактериофага T4 кодируются OРС, рассеянными по геному и расположенными в межгенных областях (Kadyrov et al 1994; Sharma and Hinton 1994; Kadyrov et al 1997; Liu et al. 2003; Belle et al., 2002; Brok-Volchanskaya et al., 2008). У данных белков GIY-YIG аминокислотный мотив был обнаружен на N-конце. Помимо этого белки проявляют гомологию

90-100 N-концевых аминокислотных остатков (Sharma et al., 1992; Liu et al. 2003), на основании чего они были отнесены к одноименному семейству хоминг-эндонуклеаз (рис. 4).

Наиболее изученным представителем семейства GIY-YIG хоминг-эндонуклеаз является эндонуклеаза I-TevI (рис. 6).

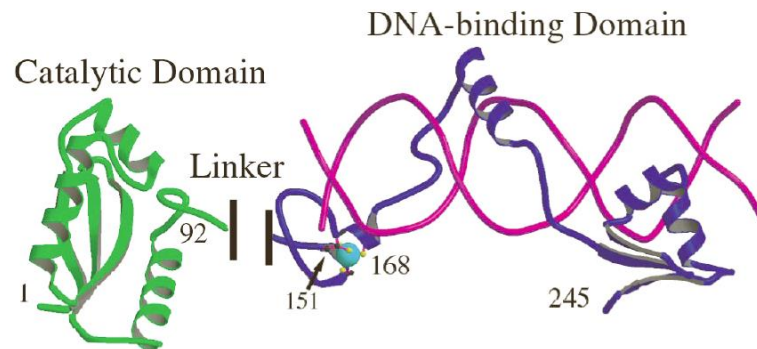


Рис. 6. Структурная организация комплекса I-TevI с ДНК. В структуре белка можно выделить каталитический и ДНК-связывающий домен. Рисунок адаптирован из статьи Van Roey et al., 2002

Данная эндонуклеаза обладает рядом уникальных свойств, в частности, является небольшим мономерным белком (28кДа), I-TevI узнает последовательность ДНК длиной 37 п.о. (Mueller et al. 1995). Хоминг-эндонуклеаза I-TevI нечувствительна к замене любого нуклеотида в сайте узнавания (Bryk et al. 1993). Сайт узнавания можно условно разбить на два фрагмента разделенных сегментом 5 п.о. Основной сайт 20 п.о. включает точку инсерции *td* интрона. Второй фрагмент длиной 12 п.о. включает сайт расщепления, таким образом, сайт расщепления отстоит от точки инсерции на расстоянии 23-25 п.о. Другой интересной особенностью I-TevI является нечувствительность к значительным вставкам и делециям в сайте узнавания. Толерантность к вставкам и делециям обеспечивается структурными особенностями организации белка. В экспериментах, с использованием направленного мутагенеза и ограниченного протеолиза, было установлено, что I-TevI состоит из двух функциональных доменов, N-концевого каталитического и C-концевого ДНК-связывающего (рис. 6) (Debyshire et al. 1997). Было также установлено, что ДНК-связывающий домен взаимодействует с сайтом узнавания с той же эффективностью, как и целый белок. Этот факт позволил сделать предположение о том, что за основной вклад в узнавание и связывание субстрата отвечает C-концевой домен. В то время как N-концевой, каталитический домен, по всей видимости, участвует в определении точки гидролиза (Bryk et al., 1995; Edgell et al., 2004).

Трехмерную структуру целой молекулы, к сожалению, установить не удалось, однако удалось закристаллизовать отдельно N- и C- домены данного белка. Так было выяснено, что

первые 97 а.о. формируют глобулярный домен с уникальной трехмерной структурой, состоящей из трех антипараллельных β -листов, фланкированных двумя и одной α -спиралями с каждой стороны. Характерные аминокислотные последовательности GIY и YIG расположены на первом и втором β -листах (рис. 7). Первый GIY элемент, по всей видимости, выполняет структурную роль, поскольку данная аминокислотная последовательность была обнаружена в ядре N-концевого домена. Вторая часть последовательности, сегмент YIG (в случае I-TevI аминокислотные остатки YVG 17-19), задействован как в формировании структуры каталитического домена, так и в катализе. Кроме того, методом направленного мутагенеза удалось установить и другие аминокислотные остатки, в частности Arg27 и Glu75, задействованные в катализе (Derbyshire et al. 1997; Kowalski et al. 1999). Показано также, что GIY-YIG-домен I-TevI помимо катализа, по-видимому, вовлечен в дополнительное узнавание в области точки гидролиза (Derbyshire, et al., 1997; Edgell et al., 2003).

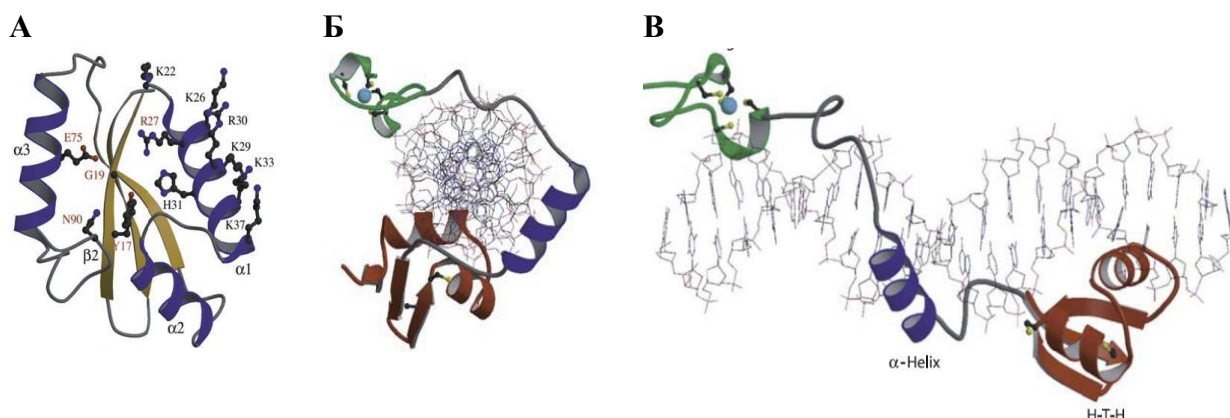


Рис. 7. Структурная организация: **А** – каталитического домена I-TevI, **Б** и **В** ДНК-связывающего домена в комплексе с ДНК субстратом (связывающий домен приведен в двух разных ориентациях). На рисунке **В** отмечены субдомены (ДНК распознающие мотивы) формирующие ДНК-связывающий домен. Рисунок адаптирован из книги Homing endonucleases and inteins. 2005.

Структура С-концевого ДНК-связывающего домена I-TevI была определена в комплексе с фрагментом ДНК, включающим сайт связывания фермента (Van Roey et al. 2001). Удалось установить, что данный домен представляет собой вытянутую структуру, окружающую ДНК субстрат (рис. 7). В структуре С-концевого домена можно выделить три элемента, разделенных неструктурированными сегментами: цинковый палец, α -спираль и субдомен спираль-поворот-спираль (Н-Т-Н). Как было установлено, взаимодействие с ДНК происходит преимущественно по малой бороздке, за исключением домена (Н-Т-Н), где одна из α -спиралей входит в большую бороздку ДНК. Выяснение структуры ДНК-связывающего домена позволило

объяснить ранее непонятные факты. Так, удалось понять, как такой относительно небольшой белок может узнавать такой протяженный сайт связывания. Кроме того, объяснить толерантность I-TevI к природной модификации цитозина в случае Т-четных фагов.

При изучении регуляции экспрессии интрон-кодируемой хоминг-эндонуклеаз I-TevI было установлено, что экспрессия I-TevI подвержена строгому количественному и временному контролю. Так, было показано, что I-TevI экспрессируется на поздних этапах фаговой инфекции с собственной мРНК. Хотя ОРС, кодирующая I-TevI, находится в составе интрон-содержащей пре-мРНК, транскрибируемой с раннего промотора, расположенного перед геном *td*. Однако трансляция гена I-TevI с ранней пре-мРНК отсутствует. Этот факт объясняется тем, что в пре-мРНК последовательность Шайна-Дальгарно, необходимая для инициации трансляции, закрыта в стабильной шпильке. В то время как у поздних транскриптов I-TevI, иницируемых с собственного позднего промотора, блокирующая шпилька отсутствует, поэтому они эффективно транслируются (Gott et al., 1988).

Кроме того, была установлена гомология между сайтом узнавания данной хоминг-эндонуклеазы и последовательностью ДНК, частично перекрывающей промотор I-TevI. Так было показано, что ДНК-связывающий домен хоминг-эндонуклеазы I-TevI эффективно связывается с операторным участком собственного гена. Основываясь на данном факте, авторы выдвинули идею авторегуляции количества в клетке молекул I-TevI. Эта идея была подтверждена в экспериментах *in vivo*, когда было установлено, что I-TevI подавляет собственную экспрессию на уровне транскрипции, связываясь (но не гидролизуя) с операторным участком ДНК (Edgell et al. 2004). Становится очевидным, что время появления и количество хоминг-эндонуклеазы в инфицированной клетке подвержено строгому многоуровневому контролю.

1.2.3. Хоминг, как один из вариантов мобильности ДНК

Мобильные генетические элементы образуют в высшей степени разнородную группу последовательностей ДНК. Общим, для данных последовательностей ДНК является способность внедряться (перемещаться) в один или несколько сайтов, которые могут располагаться как в одном, так и в нескольких геномах (Roberts et al. 2008). Размер мобильных элементов варьирует, и может составлять от нескольких сотен до нескольких десятков тысяч п.о. Помимо фрагментов ДНК, задействованных в транспозиции, мобильные элементы могут содержать в своем составе последовательности ДНК, кодирующие факторы антибиотикорезистентности, вирулентности и другие дополнительные фрагменты ДНК.

С точки зрения классификации, гены хоминг-эндонуклеаз входят в разнородную группу мобильных генетических элементов (Roberts et al. 2008). Для понимания хоминга, как варианта мобильности ДНК, следует вкратце разобраться с имеющимися вариантами мобильных элементов.

Мобильные элементы встречаются практически у всех изученных к настоящему моменту организмов - от бактерий до человека и растений. Обычно мобильные элементы подразделяются на две основные категории, в зависимости от механизма их транспозиции: ретротранспозоны (класс 1) и ДНК-транспозоны (класс 2).

Ретротранспозоны кодируют обратную транскриптазу и перемещаются с использованием РНК-интермедиата, т.е. используют кодируемую ими РНК в качестве матрицы для синтеза кДНК, на основании которой строится вторая цепь ДНК. В дальнейшем эта ДНК и включается в хромосому с помощью интегразы, кодируемой обычно тем же самым элементом (Евгеньев, 2007).

ДНК-транспозоны, относящиеся ко второму классу мобильных элементов, используют иной механизм перемещения. Данные элементы кодируют собственную транспозазу, узнающую короткие инвертированные повторы, расположенные обычно на концах мобильного генетического элемента. С помощью транспазы обеспечивается перемещение повторов по механизму “вырезание-встраивание”.

В отличие от описанных выше способов перемещения мобильных генетических элементов, процесс хоминга существенно отличается от используемых выше схем транспозиции. Хоминг это репликативно-рекомбинационно зависимый процесс перемещения гена хоминг-эндонуклеазы и примыкающих к нему нуклеотидных последовательностей в родственный гомологичный геном. Механизмы и основные этапы лежащие в основе этого процесса описаны на примере интрона тимидилатсинтазного гена фага T4 (Mueller et al., 1996; Parker et al., 1999; Huang et al., 1999).

Предложенные модели хоминга включают следующие этапы: 1) введение двухцепочечного разрыва хоминг-эндонуклеазой в ДНК безинтронного аллеля (на рис. 8 обозначен как “recipient”); 2) деградация ДНК в месте разрыва экзонуклеазами, приводящая к образованию двухцепочечной “брешы” и 3'-выступающих концов; 3) рекомбиназа (*RecA* у *E.coli*, *UvsX* у фага T4) взаимодействует с 3'- концом реципиентной ДНК и внедряет его в интрон-содержащий донорный дуплекс, что вызывает “плавление” последнего (на рис. 8 обозначена как “donor”); 4) внедрившийся 3'-конец функционирует как праймер, а интактная молекула дезоксирибонуклеиновой кислоты выступает в роли ДНК-матрицы.

Предложено несколько схем описывающих репарационный процесс. К одной из наиболее изученных моделей можно отнести репарацию двух-цепочечного разрыва “DSBR”

(double-strand-break repair). В ходе репарационного процесса образуется две структуры Холлидея, для расщепления данных структур необходимо участие разолвазы. Тем не менее, в группе Белфорт, было показано, что хоминг тимидилатсинтазного интрона может проходить и в отсутствии резолвазы (Mueller et al., 1996).

Поэтому была предложена альтернативная модель (“SDSA”), описывающая процесс хоминга. При этом не образуется структур Холлидея и кроссоверных продуктов. В случае бактериофага T4, по-видимому, возможна вариация отдельных этапов хоминга (Mueller et al., 1996).

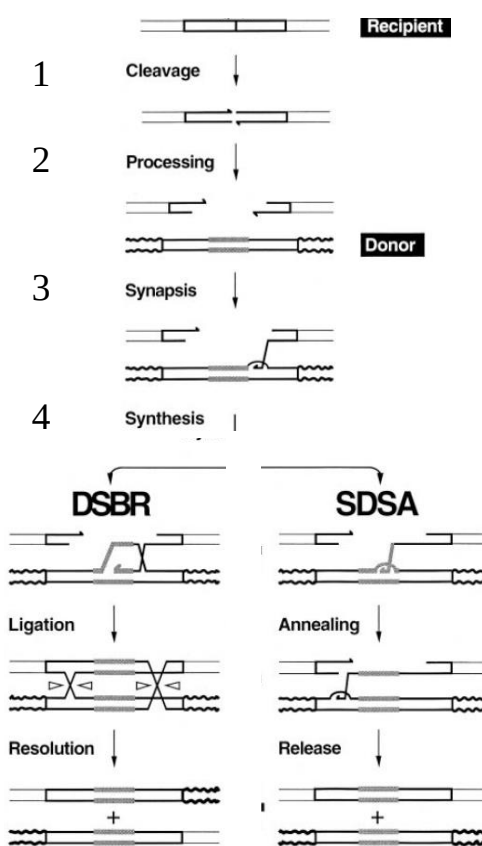


Рис. 8. Предполагаемые механизмы хоминга интрона *td*. Рисунок адаптирован из статьи Паркера и др. (Parker et al., 1999).

Таким образом, хоминг-эндонуклеазы представляют собой уникальный класс мобильных элементов. Механизм, лежащий в основе их мобильности (перемещения) отличается от механизма перемещения ретротранспозонов (класс 1) и ДНК-транспозонов (класс 2).

1.2.4. Биологическая роль хоминг-эндонуклеаз

Помимо участия в мобильности собственных генов для ряда хоминг-эндонуклеаз также было показано участие и в других клеточных процессах (Burt and Kouforanou, 2004).

Хоминг-эндонуклеаза I-DmoI (семейство LAGLIDADG), помимо участия в процессе хоминга собственного гена, ответственна за феномен подовления роста клеток, не содержащих ген хоминг-эндонуклеазы *I-DmoI*. Данное явление, заключающееся в том, что в смешанной культуре клеток *Desulfurococcus mobilis*, HEG⁺-клетки растут быстрее HEG⁻-клеток. Причины вызывающие данный феномен до сих пор не понятны, однако Агард (Aagaard et al. 1995) предположил, что хоминг-эндонуклеаза I-DmoI может проникать в клетки без проникновения ДНК, в которой закодирована информация о данном белке. Следовательно, при проникновении и гидролизе ДНК HEG⁻-клеток хоминг-эндонуклеазой I-DmoI, не возможна рекомбинационно-зависимая репарация с участием HEG⁺-аллеля в качестве матрицы. В таком случае, хоминг-эндонуклеаза выступает в качестве «токсина». Данный «токсин» действует на клетки, не содержащие ген хоминг-эндонуклеазы I-DmoI. Этот механизм действия, в конечном итоге, также приводит к распространению гена хоминг-эндонуклеазы в популяции клеток, но не за счет хоминга, а за счет элиминации HEG⁻-клеток. Таким образом, в данном случае HEG⁺-клетки обладают некоторым преимуществом в сравнении с HEG⁻-клетками.

Существуют примеры эффективного использования хоминг-эндонуклеаз для нужд клетки. Так было показано, что регулятором экспрессии гена глутатионового транспортера (the high-affinity glutathione transporter) *GSH11* у *S. cerevisiae* (Miyake et al., 2003) является белок VDE (эндонуклеаза LAGLIDADG семейства). Было установлено, что данный белок связывается с регуляторным цис-действующим элементом гена *GSH11* (CCGCCACAC) и стимулирует транскрипцию. Показано, что делеция гена *VDE* приводит к полной блокаде синтеза данного транспортера, которая снимается при восстановлении функции VDE. В данном случае, эндонуклеаза VDE выступает в роли фактора транскрипции.

Еще одним примером хоминг-эндонуклеазы, приспособленной для нужд клетки, является HO эндонуклеаза. Она задействована в переключении типа спаривания у дрожжей. “Пол” дрожжевой клетки определяет генетическая кассета, расположенная в *MAT* локусе. Существуют две генетические кассеты HML α и HMR α (Butler et al., 2004). HO эндонуклеаза вносит двухцепочечный разрыв в *MAT* локус и запускает процесс генетической конверсии, приводящий к смене кассет. Ген HO эндонуклеазы представляет собой свободностоящую ОРС. Предполагают, что в ходе эволюции произошло изменение специфичности продукта данного гена, и HO эндонуклеаза приобрела способность гидролизовать *MAT* локус и, тем самым, участвовать в переключении типа спаривания у дрожжей.

Еще одним интересным процессом, в котором, как предполагают, могут быть задействованы хоминг-эндонуклеазы, является исключение генетических маркеров в случае совместных инфекций Т4 и родственных бактериофагов. Более детально о феномене исключения генетических маркеров и роли хоминг-эндонуклеазы в данном явлении описан выше.

1.3. Хоминг-эндонуклеаз бактериофага Т4 и родственных фагов

1.3.1 Хоминг-эндонуклеаз бактериофага Т4, закодированных в составе интронов

Одним из первых интронов, обнаруженным у прокариот, оказался интрон в гене тимидилатсинтазы (*td*) бактериофага Т4 (Chu et al., 1984). На сегодняшний момент известно, что в геноме бактериофага Т4 имеется три самосплайсирующихся интрона I группы, которые расположены в генах *td*, *nrdB* и *nrdD*. Последние два гена кодируют субъединицы аэробной и анаэробной рибонуклеотидредуктазы, соответственно. По своей структуре данные интроны относятся к группе IA2 и используют ГТФ для реакции трансэтирификации (Shub et al., 1988).

Все три интрона помимо структур, необходимых для катализа реакции сплайсинга, в своем составе имеют открытые рамки считывания. Данные ОРС кодируют ДНК-эндонуклеазы, вовлеченные в процесс хоминга соответствующих мобильных генетических элементов. Продукты выше описанных генов, в соответствии с правилами номенклатуры, были названы I-TevI, I-TevII, и I-TevIII. Показано, что данные хоминг-эндонуклеазы узнают и вносят разрыв безинтронный аллель, тем самым, стимулируют рекомбинационные процессы, приводящие к распространению данных интронов. Кроме того, ферменты итроной локализации проявляют гомологию с их свободностоящими аналогами, относящиеся к Seg- и Mob- группам белков. Хоминг-эндонуклеаза I-TevI на сегодняшний момент является наиболее изученным представителем GIY-YIG семейства хоминг-эндонуклеаз. Более подробно о структуре и особенностях работы данной хоминг-эндонуклеазы написано выше.

Хоминг-эндонуклеаза I-TevII кодируется интроном *sunY* бактериофага Т4. Установлено, что данная эндонуклеаза расщепляет безинтронный вариант *sunY*, в области соединения экзонов и, тем самым, запускает процесс переноса интрона *sunY*, в безинтронный вариант аллеля. Для данного фермента, как и большинства других хоминг-эндонуклеаз, показано, что гидролиз ДНК-субстрата происходит с образованием 3'-выступающих ОН концов длиной 2 н.о. Основные контакты между ДНК-субстратом и эндонуклеазой были картированы в малом желобке ДНК. Интересно еще и то, что после гидролиза фермент остается связанным с одним из концов субстрата. Высказывается мнение, что связь эндонуклеазы с одним из концов

субстрата может влиять на использование данного фрагмента ДНК на последующих стадиях процесса переноса интрона (Loizon et al., 1996).

В случае фага T4 ген *I-TevIII* содержит делецию, в то время как у бактериофага RB3 (родственнен T4) была обнаружена полноразмерная ОРС, продукт которой обладал эндонуклеазной активностью (Eddy and Gold, 1991). Было показано, что *nrdB* интрон, содержащий в своем составе ОРС данной хоминг-эндонуклеазы, является мобильным. *I-TevIII* гидролизует ДНК-мишень с образованием 3'-выступающих концов длиной 2 н.о. Как и для многих HNH эндонуклеаз, для *I-TevIII* было показано наличие каталитического и ДНК-связывающего домена. Наиболее интересным является то, что данная хоминг-эндонуклеаза работает как димер (Robbins et al., 2007), в то время как сайт узнавания не имеет симметрии второго порядка.

1.3.2. Свободностоящие хоминг-эндонуклеазы бактериофага T4, сходства и основные отличия от интрон-кодируемых эндонуклеаз

Гены хоминг-эндонуклеаз обнаружены не только в составе мобильных самосплайсирующихся элементов (интронов и интеинов), они также были найдены в виде свободностоящих ОРС. Для ряда генов бактериофагов было показано, что они сами по себе являются мобильными генетическими элементами. Такими ОРС чрезвычайно насыщен геном бактериофага T4 (Belle et al., 2002; Kadyrov et al., 1997; Liu et al., 2003; Brok-Volchanskaya et al., 2008). Поскольку процесс распространения данных мобильных генетических элементов схож с процессом распространения интронов и интеинов, его также называют хомингом.

Белковые продукты свободностоящих ОРС бактериофага T4 во многом схожи с ферментами, закодированными внутри интронов. Однако существует ряд особенностей, отличающих эти группы белков. Одна из таких особенностей - расстояние между точкой гидролиза и местом встраивания мобильной последовательности ДНК. Для генов, кодирующих свободностоящие хоминг-эндонуклеазы это расстояние может достигать до нескольких сотен п.о. (рис.9А). В то время как для ОРС интроной локализации данное расстояние не превышает 25 п.о. (Chevalier and Stoddard 2001). При таком близком расположении, узнаваемый участок ДНК включает и сайт инсерции интрона. Таким образом, при встраивании мобильной последовательности в сайт инсерции происходит одновременно и нарушение сайта узнавания. Данная стратегия близкого расположения участка узнавания и точки инсерции обеспечивает устойчивость HEG⁺-генома к действию интронкодируемых хоминг-эндонуклеаз, поскольку одновременно со встраиванием мобильной последовательности происходит и элиминация сайта узнавания.

Свободностоящие хоминг-эндонуклеазы бактериофага Т4 также могли бы воспользоваться стратегией перекрывания сайта узнавания и точки инсерции (рис. 9Б). Однако, в таком случае сайт узнавания должен был бы располагаться в межгенной области. Данная стратегия, в случае свободностоящих хоминг-эндонуклеаз, ограничена высокой вариабельностью межгенных областей и, как следствие, меньшей встречаемости данных последовательностей у родственных геномов. Картирование сайтов гидролиза хоминг-эндонуклеаз SegE, F, G и B (Belle et al., 2002; Kadyrov et al., 1997; Liu et al., 2003; Brok-Volchanskaya et al., 2008) показало, что у всех приведенных выше ферментов сайты узнавания находятся в составе генов, а не в межгенных областях. Эта стратегия выбора сайта гидролиза позволяет, как и в случае интрон кодируемых ферментов, увеличить потенциальную представленность сайта гидролиза среди родственных геномов (потенциальных новых хозяев данной мобильной последовательности ДНК).

Однако из любого правила существуют, и исключения. Так, в случае хоминг-эндонуклеазы SegC последовательность ДНК, узнаваемая эндонуклеазой, включает и сайт инсерции, как и в случае интрон кодируемых ферментов. При встраивании мобильной последовательности в сайт инсерции, расположенный в области генов 5-6 бактериофага Т2, происходит нарушение сайта узнавания, как и в случае интрон кодируемых хоминг-эндонуклеаз (Диссертация Грановского И. Э.). Становится очевидным, что в случае SegC и интрон кодируемых хоминг-эндонуклеаз встраивание мобильной последовательности приводит к элиминации сайта гидролиза, что предотвращает в дальнейшем гидролиз хозяйского генома (рис. 9).

Свободностоящие хоминг-эндонуклеазы используют другие механизмы для предотвращения гидролиза хозяйского генома, поскольку сайт узнавания данных ферментов не элиминируется инсерцией эндонуклеазного гена.

По всей видимости, полиморфизм отдельных нуклеотидов в сайте узнавания и определяет устойчивость собственного генома к действию своих хоминг-эндонуклеаз (рис. 9Б). Так, предпочтительный сайт хоминг-эндонуклеазы SegE из фага RB30 длиной 19-п.н. отличается от менее предпочтительного Т4-сайта по трем п.н. (Kryukov et al., 1996), а сайт SegB в гене tPHK^{Asn} фага Т2L длиной 27-п.н. содержит 7 замен в сравнении с сайтом, обнаруженным в гене tPHK^{Ser} Т4 (Brok-Volchanskaya et al., 2008). При анализе фаговой ДНК, выделенной из клеток, инфицированных фагами Т4 и Т2L, авторы не детектировали появление SegF- или SegG-специфических разрывов в ДНК Т4 (Belle et al., 2002; Liu et al., 2003). Тем не менее *in vitro*, при избытке фермента в реакционной смеси, помимо гидролиза безнуклеазного аллеля наблюдается и гидролиз ДНК фага Т4 (Kadyrov et al., 1997; Liu et al., 2003; Belle et al., 2002; Brok-Volchanskaya et al., 2008).

Таким образом, отдельные нуклеотидные замены в сайте гидролиза, расположенного в безнуклеазном аллеле родственного генома, делают его более предпочтительным для гидролиза и тем самым определяют направление переноса генетической информации (ОРС хоминг-эндонуклеазы) от HEG^+ в HEG^- геном.

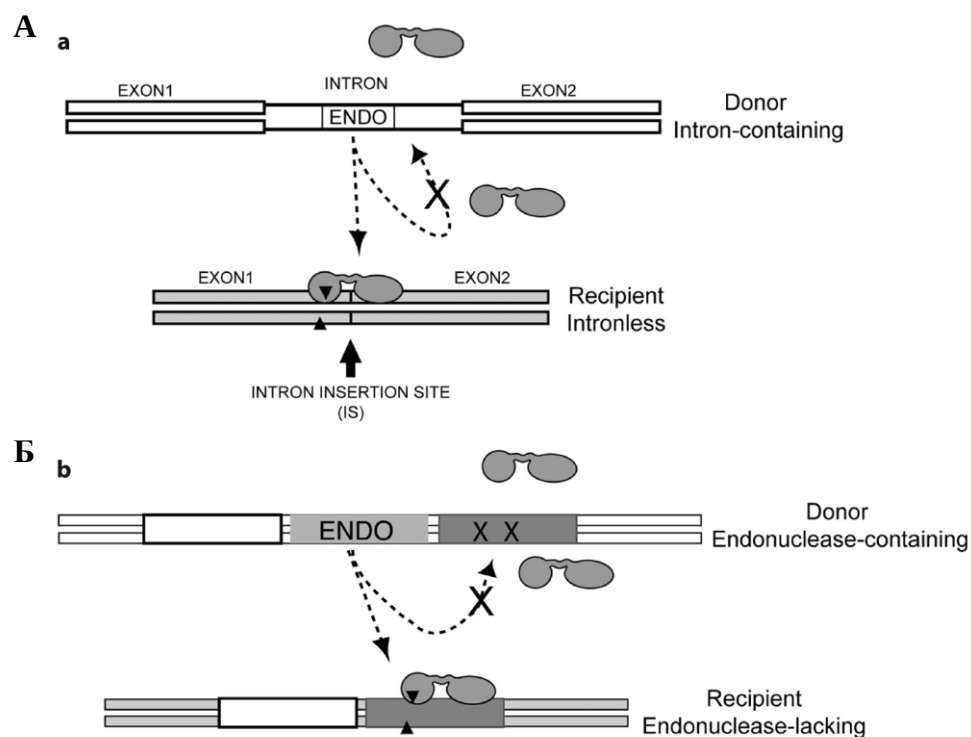


Рис. 9. Различные механизмы, используемые хоминг-эндонуклеазами, обеспечивающие минимизацию гидролиза хозяйского генома. (А) Механизм характерный для интрон ассоциированных и (Б) для свободностоящих хоминг-эндонуклеаз. Черными треугольниками обозначены точки гидролиза. X обозначены отдельные нуклеотидные замены в сайте узнавания, которые могут предотвращать гидролиз собственного генома, ситуация (Б). Рисунок адаптирован из книги Homing endonucleases and inteins. 2005

1.3.2.1 Сайт специфические эндонуклеазы Mob-группы

Как упоминалось ранее, в геноме бактериофага Т4 обнаружено три интрон-ассоциированных и 12 свободностоящих генов хоминг-эндонуклеаз (рис. 10). Семь из свободностоящих генов кодируют белки семейства GIY-YIG хоминг-эндонуклеаз. Пять оставшихся ОРС, кодируют предполагаемые эндонуклеазы семейства H-N-H. Данные гены были названы *mobA-E* генами (от слова *mobile* endonucleases) (Sandegren, L. et al., 2005). На сегодняшний момент существует мало информации относительно данных ОРС и кодируемых ими белков.

Ген *mobA* является псевдогеном, поскольку основная его часть делетирована. Данный псевдоген располагается в области генов 39-60, которые кодируют субъединицы ДНК топоизомеразы (E. Thomas, F. Zucker, and E. Kutter, не опубликованные данные, Miller et al., 2003).

Ген *mobB*, в отличие от предыдущего примера, является, по-видимому, полноразмерным. Он располагается между геном, кодирующим α -гликозилтрансферазу, и геном 55, кодирующим σ -субъединицу РНК-полимеразы, ответственную за транскрипцию поздних промоторов. Хоминг для данного гена показан не был, как и не была показана нуклеазная активность кодирующего им белка (Belfort, M. et al., 2002, Miller et al., 2003).

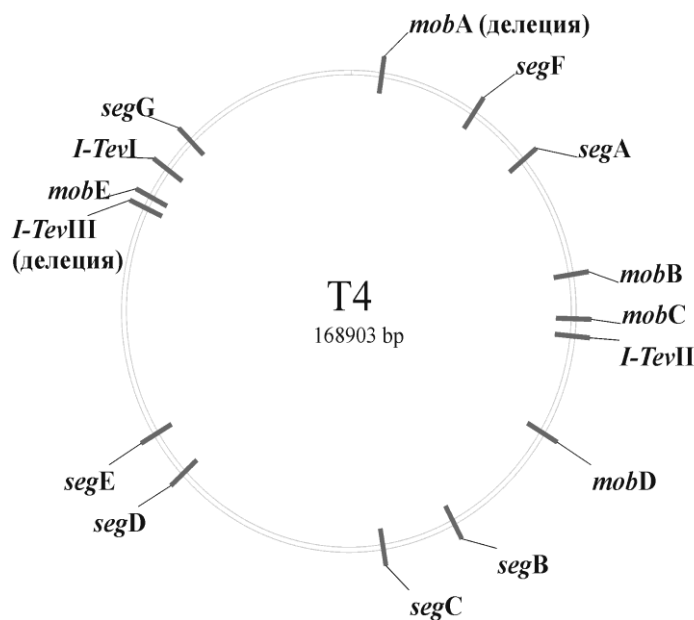


Рис. 10. Генетическая карта бактериофага Т4 с отмечены на ней известными и предполагаемыми генами хоминг-эндонуклеазами.

Ген *mobC* бактериофага Т4 располагается между генами *nrdD* и *nrdG*, кодирующими анаэробную рибонуклеотидредуктазу и ее активатор, соответственно. На основании анализа различных представителей семейства Т4-родственных бактериофагов было установлено, что данный ген имеется только у Т4, однако, в этой межгенной области у фагов Т6, RB3 и LZ2 обнаружен ген, продукт которого проявляет гомологию с Seg-белками. Ген *mobC* бактериофага Т4 был экспрессирован в бесклеточной системе трансляции, однако эндонуклеазной активности *in vitro* обнаружено не было (Sandegren, L. et al., 2005).

Ген *mobD* располагается в межгенной области *nrdC-rI*. Известно, что данный ген был клонирован и успешно экспрессирован (Kutter, не опубликованные данные, Miller et al., 2003) однако результаты данного исследования на сегодняшний день остаются неопубликованными.

ОРС *mobE* располагается между генами α - и β -субъединиц рибонуклеотид редуктазы. Ген *mobE*, помимо Т4, обнаружен у таких фагов как Т6, RB2, RB3, RB15, LZ7 и у фага Aeh1, относящегося к шизо-Т-четным бактериофагом (Sandegren, L. et al., 2005, Thomas et al., не опубликованные данные, Miller et al., 2003). Как было установлено, ОРС *mobE* не имеет собственного промотора, а транскрипция гена зависит от активности вышестоящих промоторов. Внутри ОРС *mobE* был обнаружен терминатор, ограничивающий транскрипцию данного участка. Было показано, что MobE является сайт-специфической нуклеазой, сайт гидролиза фермента был обнаружен на 5'-конце гена *nrdB* фага Т2. Оказалось, что никаза MobE помимо перемещения собственного гена инициирует мобилизацию близлежащего интрона (расстояние между генами - около 700 п.н.), содержащего в своем составе делеционный вариант гена *I-TevIII*. Это обеспечивает высокую частоту наследования данного интрона фаговым потомством (Wilson et al., 2009). Таким образом, ген *mobE* - единственный представитель *mob*-генов, для которого была показана мобильность и установлено, что кодируемый им белок обладает нуклеазной активностью.

У фага U5 на месте *mobE* был обнаружен ген, который был назван *hef* (*homing endonuclease-like function*). Для продукта данного гена была показана эндонуклеазная активность и способность вызывать хоминг собственного гена. Интересно то, что данный белок не относится ни к одному из основных известных семейств хоминг-эндонуклеаз (Sandegren, L. et al., 2005).

1.3.3.2 Сайт-специфические эндонуклеазы Seg-группы

Белки SegA-G фага Т4 кодируются ОРС, рассеянными по геному и расположенными в межгенных областях (Sharma et al., 199; Belle et al., 2002; Kadyrov et al., 1997; Liu et al., 2003; Brok-Volchanskaya et al., 2008). В начале белков обнаружен GIY-YIG мотива, на основании чего

они были отнесены к одноименному семейству хоминг-эндонуклеаз. Белки SegA-E проявляют высокую степень подобия 90-100 N-концевых аминокислот (Sharma et al., 1992; Liu et al., 2003; Sandegren et al., 2005) (рис. 11.).

Для ряда членов группы Seg белков *in vitro*, было показано наличие одного либо нескольких сайтов гидролиза в собственном геноме. Эти сайты являются, как правило, менее предпочтительными в сравнении с сайтами гидролиза расположенными в *seg*-несодержащих аллелях родственных геномов. Это свойство отличает их от хоминг-эндонуклеаз кодируемых интронами I-группы в частности от I-TevI и др.

Ген *segA* расположен между генами *βgt* и *UvsX* бактериофага T4. Установлено что SegA *in vitro* гидролизует ДНК фага T4 по ряду сайтов, причем наиболее близкий располагается на расстоянии ~600 п.о. от собственной OPC. SegA гидролизует ДНК в гене *UvsX* с образованием 3' выступающего конца длиной 1 н.о. Хоминг для данного гена показан не был (Sharma and Hinton, 1994).

В то время как для OPC *segB*, располагающейся в кластере генов тРНК, был показан хоминг в случае скрещивания фагов T4 и T2L. Показано, что ген *segB* кодирует сайт-специфическую эндонуклеазу, предпочтительный сайт гидролиза которой имеет длину 27 п.н. и локализуется в гене аспаргиновой тРНК бактериофага T2L. SegB вносит двуцепочечные разрывы в ДНК с образованием 3' выступающих концов длиной 2 н.о. Было показано, что основным мотивом узнавания SegB является фрагмент GGTTTCRA, кодирующий консервативный элемент TψC-стебля тРНК. При анализе *in vitro*, во всех генах тРНК фагов T4 и T2L были обнаружены сайты хоминг-эндонуклеазы SegB, которые гидролизировались с различной эффективностью (Brok-Volchanskaya et al., 2008). Это свидетельствует о высокой вырожденности сайта узнавания SegB.

Ген *segC* расположен в межгеной области 5 и 6 и кодирует Mg²⁺-зависимую сайт-специфическую эндонуклеазу. В отличие от других Seg-белков, SegC не гидролизует ДНК фага T4 *in vitro*, однако вносит двух-цепочечный разрыв в межгеноую область 5-6 фага T2L. SegC инициирует хоминг собственного гена, при совместной инфекции T4 и T2L (Грановский И. Э. не опубликованные данные). Кроме того, для SegC, как и для I-TevI, была показана возможность инициации генетической конверсии в удаленном локусе. Конверсия в локусе *rII* инициировалась посредством внесения двухцепочечного разрыва в собственный сайт гидролиза, который был встроен в ген *rIIB* (Shcherbakov et al., 2002).

В области генов *hoc.1* и *UvsW* бактериофага T4 располагается OPC *segE*. При картировании сайтов гидролиза SegE эндонуклеазы, локализующихся в собственном геноме, было показано, что существует предпочтительный участок расщепления, находящийся в 5'-области гена *UvsW*. При детальном картировании было показано, что расстояние от разрыва до

конца гена *segE* составляет 128 п.о. Кроме того, был обнаружен дополнительный, менее предпочтительный сайт, расположенный в области 3'-конца гена *segE*. Анализ сайтов гидролиза показал, что SegE в обоих случаях гидролизует ДНК с образованием 3'-выступающих липких концов длиной 2 н.о. При этом, разрыв фосфодиэфирной связи происходит с образованием 5'-концевого фосфата. Минимальная последовательность предпочтительного сайта находится в положении от -1 до +17 по отношению к точкам разрыва ДНК и составляет 19 н.о. Секвенирование участков ДНК, соответствующих сайту гидролиза, у родственных фагов RB30 и RB51, показало присутствие трех нуклеотидных замен в узнаваемой последовательности ДНК в позициях -2, +2 и +10. Было установлено, что фрагменты ДНК фагов T4 и RB30, содержащие область 5'-конца гена *ivsW*, гидролизуются SegE (Kadyrov et al., 1996). Причем, в условиях одновременного присутствия в реакции ПЦР-фрагментов ДНК фагов T4 и RB30, преимущественно осуществляется гидролиз ДНК фага RB30. Преимущественный гидролиз аллеля родственного фага, не содержащего *segE*, определяет нерцепируемое наследование *segE* фаговым потомством. На примере ОРС *segE* было впервые показано, что свободностоящие гены хоминг-эндонуклеаз могут являться мобильными генетическими элементами (Kadyrov et al., 1997).

Ген 69, который впоследствии был переименован в *segF*, находится между генами 56 и *soc.1* фага T4. SegF - хоминг-эндонуклеаза, гидролизующая ДНК с образованием 3'-выступающих концов длиной 2 н.о. Сайт гидролиза хоминг-эндонуклеазы SegF расположен в консервативной области гена 56 на расстоянии ~200 п.н. от начала собственной ОРС. Эксперименты *in vitro* показали, что SegF может гидролизовать как собственную ДНК, так и ДНК фага T2, в котором отсутствует ген *segF*. Как и SegE, SegF предпочтительно гидролизует безнуклеазный аллель, в данном случае расположенный в геноме фага T2. Эксперименты *in vivo* подтвердили, что при совместной инфекции T4 и T2 преимущественно гидролизуется сайт расположенный в ДНК T2. Для SegF также была показана способность инициировать хоминг собственного гена (Belle, A., et al., 2002).

Ген *segG* находится между генами 59 и 32 фага T4. SegG - хоминг-эндонуклеаза гидролизующая ДНК с образованием 3'-выступающих концов длиной 2 н.о. Сайт гидролиза хоминг-эндонуклеазы SegG расположен в полиморфной области гена 32 на расстоянии ~330 п.н. от сайта инсерции ОРС. В экспериментах по связыванию ПЦР-фрагментов фагов T2 и T4, содержащих ген 32, было показано, что SegG связывает данные субстраты с равной эффективностью. Однако, как *in vivo*, так и *in vitro* SegG предпочтительно гидролизует сайт, расположенный в ДНК фага T2. Исходя из этого было высказано предположение, что различия в последовательностях сайтов влияют не на связывание фермента с ДНК, а на последующую

	102030405060708090100	
T4segA	MKRHKKEKKYNYTYVITNLVNNKI..VYGTHSTDDL...NDGYMGSGTLAQAQKKYGGKKNFNLSILGFYKDFKSARDAERELVTIDVVNDPM.TYNLKIG	94
T4SegB	MPFYTYKITNKINNKI..YIGVHSTENL...DDGYMGSGCKLKRAQDKYGIENFSKEILEYFDDKESMLEAEKNIIVTEEFVLRNRPD.VYNLKL	87
T4SegC	.MPMQDHKYFYLYSITNKTTTEKI..YVGVHKTSNL...DDGYMGSGVAIKNAIKKYGIDNPFYKHIIKFFESEKAMYDAEAEIVTEEFVKSCK.TYNMKG	93
T4SegD	MFKMKYLIYQITNIINGKI..YIGAHATLDE...NDGYMGSGVNIKKSIKKYGIHNFKEILYSFSSSEEMYKMEALLVNEEFVMRTD.TYNAAG	90
T4SegE	MYHFVYETTNLNGKK..YIGKHSTDDL...NDGYLGSGKAIQQAIAKKYGENNFSRTILKEFKTSEEAYMYEEIIITPELIKSN.YYNMKPG	87
T4SegF	..MDIKQKFYRTYIVKVRTPKGVFWYAGKHESFIVNPYNDKYPGSGKILWNIYRKYG.FNFKIRWSKCHGSREKSYPEVERELISALKRKHDPDTCINISPG	97
T4SegG	MNIVYWFTEPDRVKNKTPPYYYTCSALNCSF.....ENGCIYDSSSGKEYWSSCKQKRFNALMLQKPSVKIIQIDDDLDVTEAER.KYQLEVN	87
LZ2SegH	MHRYYIYKITNIINSKE..YIGAHSTDNL...NDGYMGSGTLIRAALLKYGRKNFNKEILHVFENEKQMYDKAEELVTEEYCLKEN.TYNIKPG	89
	110120130140150160170180190200	
T4segA	GEGGRRIGYRVSSSETKEKISKAQKGKPKHLGFSQVCRKAQLGKKQSE.....ETKAKRKEALLNNPYGYN.....	159
T4SegB	GEGGWDHVNIPIGMLNQKK.DASLKEAKSPKSRFENDILLQEKYRKIG.....SNVFKRLWSTPEYREKFLNNS.....	154
T4SegC	GIGGFPHNTAGAKNG....FYGKS.HSRETRLKIS.....IKSSRK.....RGPRGPR.....	137
T4SegD	GRGNPVIIVHLQDPSYRNMLSIRTKEGMTVEAKRKISKAKTGVKQSD.....TIAKRVKGRNEYKYKTHVHPRK.....	158
T4SegE	GIGGIVMTTDVIAKMKESSAKRFENSPGTVLGKTCYTNGTKNIFIKP.....GELVPEGFVKGMVHPN.....	150
T4SegF	GQGGEGRKWTEQQRLEHKLRLNNPETKTRMKNSQRIAQNRARERKARQ.....SEVMKKFYSNNGGNKKISEGTS.....	165
T4SegG	ARDNPDYFNLVYAGGGFGVSGETHPAKDPEVREHMLANYMNRDDFR.....PWKTSRANIESWKLSHIAYEN.....	155
LZ2SegH	GKGGTAYEQTLEHRHKNSLANKGKK.RTEETKIAISRALKKYKRTPEHQANLNKAVSIAMKRKDVRSKLAGERSEKDKLSISNGVKKYNSLSEEEERELH	188
	210220230240250260270280	
T4segA	RNKPSHKRDPIMWDNIEKIKEIWENSGKSGAIAKLKLAIEAGFPNKSARMLEVFRTGTRTLL.	221
T4SegB	RFLNKHHTPETINKMKESHAKNNHQKGEKNSQFGMMWIHSLDEKVSRIKKTDPPIEGWFKGRKMKF	221
T4SegC	GKTLK..MCGANNPRYGKIAAPNAKSVIINGVLYKSIKIAAKALNINYSTLKGRVKAGYYKCQD....	198
T4SegD	NKPISSHRKAISEKLGGIK.KFIPICIKGVKFDNPDYAAEHFKVSPKTIARNWINADDMPDCYRIK.	223
T4SegE	RKSRKGCKVKPTTTGTFWVNGAINKLIQPDGIIIPDGFIKGRLMKRDSKGKFSKA....	205
T4SegF	RAQRKAPHWHEPLKSEIHNLWVSLGKPKATGPVVKALKGKYDVTSSALKNLIYLFKEDV.	224
T4SegG	YVLLSSNLYGKTPGWRVRVGNINITDITAKSMVKYFNSGWIPLEDPEYCELCQL.....	210
LZ2SegH	RKRIIEGKKNSPMSLESRQKLSKALKGLQVGSKNPMYGKISPTSKTVSVDGIKYDSIKKCSNAIKVSRHLILKRCNSTDEAWKNWILL	276

Рис. 11. Выравнивание аминокислотных последовательностей известных Seg-белков. Эндонуклеазы SegA, B, C, D, E, F, G относятся к эдонуклеазам бактериофага T4, SegH- нуклеаза фага LZ2. Адаптировано из статьи Sandegren et al., 2005.

стадию гидролиза. Для данного гена был показан хоминг в скрещиваниях фагов T4 и T2 (Liu et al., 2003).

У фагов T6, RB3 и LZ2 в области генов *nrdD* и *nrdG* вместо *mobC* был обнаружен ген *segH*. Продукт данного гена проявляет гомологию с Seg-белками. Было установлено, что SegH является эндонуклеазой, гидролизующей ДНК с образованием 3' выступающих концов длиной 2 н.о. Сайт гидролиза хоминг-эндонуклеазы SegH расположен на расстоянии ~80 п.о. от гена *nrdD* в области, в которой наблюдается полиморфизм ДНК между фагами *segH*⁺ и *segH*⁻. Тем не менее, было показано что SegH *in vitro* может гидролизовать как HEG⁺ так и HEG⁻ аллели, как и в случае других Seg-эндонуклеаз бактериофага T4. Для данного гена был показан хоминг, а это предполагает, что *in vivo* фермент предпочтительно гидролизует HEG⁻ аллель (Sandegren et al., 2005).

ОРС *segD* фага T4 располагается между генами 23 и 24, кодирующими мажорные белки-предшественники капсида оболочки, и имеет противоположную к ним ориентацию (Miller et al., 2003). В данной работе показано, что ген *segD* кодирует сайт-специфическую эндонуклеазу, которая имеет общие свойства с другими Seg-нуклеазами фага T4, так и ряд существенных отличий.

ГЛАВА 2. МЕТОДЫ ИССЛЕДОВАНИЯ И МАТЕРИАЛЫ

2.1. Материалы

2.1.1. Бактериофаги и бактериальные штаммы

Штаммы бактерий, бактериофаги	Описание	Ссылка
Штаммы <i>E. coli</i>		
DH5α	F ⁻ , <i>gyrA96</i> , (Nal ^r), <i>recA1</i> , <i>relA1</i> , <i>endA1</i> , <i>thi-1</i> , <i>hsdR17</i> , (<i>r_k⁻m_k⁺</i>) <i>glnV44</i> , <i>deoR</i> , Δ(<i>lacZYA-argF</i>)U169 [Φ80dΔ(<i>lacZ</i>)M15]; реципиент при конструировании плазмид.	Woodcock et al., 1989
JM109	<i>endA1</i> , <i>recA1</i> , <i>gyrA96</i> , <i>thi</i> , Δ(<i>lac-proAB</i>), <i>hsdR17</i> (<i>r_k⁻m_k⁻</i>), <i>supE44</i> , <i>relA1</i> , [F ⁺ , <i>traD36</i> , <i>proAB</i> , <i>lacI^r</i> , ZΔM15]; реципиент при конструировании плазмид, а также при конструировании рекомбинантных фагов	Yanish-Perron et al., 1985
Top10	<i>mcrA</i> , Δ(<i>mrr-hsdRMS-mcrBC</i>), <i>φ80</i> , <i>lacZ</i> ΔM15, Δ <i>lacX74</i> , <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , Δ(<i>ara</i> , <i>leu</i>)7697, <i>galU</i> , <i>galK</i> , <i>rpsL</i> , (Str ^R), <i>endA1</i> , <i>nupG</i> ; реципиент при конструировании плазмид.	Invitrogen
BL21(DE3)/pLysE	F ⁻ , <i>dcm</i> , <i>ompT</i> , <i>hsdS</i> (<i>r_B⁻ m_B⁻</i>), <i>gal</i> , λ(DE3); использован для экспрессии гена <i>segD</i> .	Studier and Moffatt, 1986
B _E (Sup ^o)	Использован для выращивания фагов	Лаб. кол-лекция
B40 (Su I)	<i>argF40</i> (Am), <i>supD60</i> , (Su1, EhoV ⁺); супрессорный штамм, использован для наработки и скрещивания фагов.	Лаб. кол-лекция
B834	F ⁻ , <i>hsdR_B</i> , <i>hsdM_B</i> , <i>met</i> , <i>thi</i> , Sup ^o ; бессупрессорный штамм, использован в качестве хозяина при выращивании бактериофагов T4290 и T4alc7.	Wood, 1966
CR63	<i>supD</i> ; Использовался в экспериментах по исследованию наследования <i>rII</i> -фенотипа.	Shcherbakov et al. 1982
CR63(λ)	<i>supD</i> , λ; Использовался в экспериментах по исследованию наследования <i>rII</i> -фенотипа потомством фагов.	Shcherbakov et al. 1982
Бактериофаги		
T4B, T6, T2L, RB1, RB2, RB3, RB4, RB6, RB7, RB8, RB9, RB10, RB15, RB55, RB59	Использованы для оценки распространенности гена <i>segD</i> . ДНК фагов T2L была использована в качестве доноров сайта гидролиза SegD. ДНК фагов T4B и T2L была использована для изучения влияния природных модификаций ДНК Т-четных фагов на эффективность гидролиза SegD.	В.И. Теляшин
T4amB20	<i>g14am</i> ; использован для получения фага B20Δ <i>segD</i> и при скрещивании с фагом T2L.	В.И. Теляшин
T4290	<i>agt</i> , <i>am8</i> , <i>βgt</i> , <i>am10</i> ; использован для изучения влияния природных модификаций ДНК Т-четных фагов на эффективность гидролиза	В.И. Теляшин
T4alc7	<i>g56amE51</i> , <i>g42amC87</i> , <i>denB</i> , NB5060, <i>alc7</i> ; использован для изучения влияния природных модификаций ДНК Т-четных фагов на эффективность гидролиза SegD.	В.И. Теляшин
T4Δ <i>segD</i>	Δ <i>segD</i> ; использован в экспериментах по изучению хоминга <i>segD</i> и определению активности SegD в клеточных экстрактах <i>E.coli</i> зараженных T4, в качестве отрицательного контроля.	Данная работа
D2	Производный T4, содержит область генов 23-24 фага T2L	Данная работа
B20dm	Производный T4, содержит 3 нуклеотидные замены в	Данная работа

	сайте гидролиза SegD и амбер мутацию в гене 14, использован при анализе наследования <i>segD</i> в случае смешанных инфекциях	
D2ets9	Производный T4, содержит в <i>rII</i> локусе сайт ридролиза SegD, использован при исследовании способности SegD индуцировать генетическую конверсию в <i>rII</i> локусе	Данная работа
D2insZ	Фаг производный T4, содержит в <i>rII</i> локусе фрагмент <i>lacZ</i> , использован в качестве неготивного кантроля при исследовании способности SegD, индуцировать генетическую конверсию в <i>rII</i> локусе	Данная работа

2.1.2. Вектора и полученные в ходе работы плазмиды

Плазмидны		
pUC18(19)	<i>ColE1</i> , Ap^r , <i>lacZ</i> , MCS; pUC18(19) отличаются ориентацией MCS. Использованы в качестве вектора для клонирования.	Vieira et al., 1982;
pBluescriptKS(-)	<i>ColE1</i> , Ap^r , <i>lacZ</i> , <i>lacI</i> , f1(-) <i>origin</i> , MCS.	Stratagene
pET-19mod	<i>lacI</i> , pBR322 <i>oirgin</i> , Ap^r , T7 <i>lac</i> -промотор; за T7 <i>lac</i> -промотором находится His-Tag-последовательность и последовательность ДНК кодирующая сайт Tев протезы; использована для клонирования и экспрессии <i>segD</i> .	Dursina et al., 2005
pET-19Tev	<i>lacI</i> , pBR322 <i>oirgin</i> , Ap^r , T7 <i>lac</i> -промотор; за T7 <i>lac</i> -промотором находится His-Tag-последовательность и последовательность ДНК кодирующая Tев протезу; использована для получения Tев протезы.	Rak et al., 2002
Рекомбинантные плазмиды		
pUT2LsD	Плазида pUT2LsD содержит фрагмент области генов 23-24 фага T2L размером 1.4 т.п.н. который был амплифицирован с праймерами #3 и #4 и клонирован в HindIII и EcoRI сайты pUC18 (Yanisch-Perron, C. et al. 1985)	Лаб. колекция
pSDET-19mod	Плазида pSDET-19mod кодирует эндонуклеазу SegD, слитую на N-конце с His ₆ и сайтом TEV-протеазы. Она была получена клонированием фрагмента ДНК фага T4, полученного при помощи ПЦР с праймерами #7 and #8 и содержащего OPC <i>segD</i> размером 672 п.н., в NdeI и BamHI сайты вектора pET-19mod (Rak, A., et al. 2002).	Данная работа
pBluT4sDs	Для получения плазмиды pBluT4sDs, участок ДНК области генов 23 – 24 фага T4 размером 1859 п.н. был амплифицирован при помощи ПЦР с праймерами #5 и #6. Полученный ПЦР-фрагмент был гидролизован эндонуклеазами SalI и HindIII. Продукт гидролиза размером 1084 п.н., содержащий 3'-концы OPC 23 и <i>segD</i> (105941_107024 bp. T4 NC_000866), был помещен в вектор pBluescriptKS (-) (Stratagene) по тем же сайтам.	Данная работа
pBluT4sDdel	Плазида pBluT4sDdel содержит фрагмент области генов 23-24 фага T4 с делецией 5'-концевой области OPC <i>segD</i> (3-208 н.о.), которая приводит к сдвигу рамки считывания. Она была сконструирована на основе плазмиды pBluT4sDs путем клонирования ПЦР-фрагмента ДНК фага T4 размером 570 п.н. (107230_107799 b.p. T4 NC_000866), амплифицированного с парой праймеров #6 и #9, по сайтам HindIII и PstI.	Данная работа

pBluT4sDs_mut	Плазмида pBluT4sDs_mut является производным плазмиды pBluT4sDs и была получена путем введения замен в последовательность сайта гидролиза SegD. Сайт-направленный мутагенез плазмиды pBluT4sDs проводили при помощи набора QuikChange Site-Directed Mutagenesis Kit (Stratagene) с использованием праймеров #10 и #11. Плазмида pBluT4sDs_mut использовалась при конструировании бактериофага T4 с нарушенным сайтом SegD.	Данная работа
pUrIIBsDs	Плазмида pURIIBsDs содержит фрагмент <i>rII</i> -локуса, в котором в 5'-конец гена <i>rIIB</i> встроен фрагмент ДНК размером 37 п.н. с сайтом SegD. Для получения данной плазмиды, олигонуклеотиды #15 и #16 были отождествлены, построены фрагментом Кленова и клонированы в сайт HincII плазмиды pUrIIB-18 (Shcherbakov, V. et al. 2002). Плазмида pURIIBsDs использовалась для введения в ген <i>rIIB</i> фага T4 сайта SegD.	Данная работа

lacZ содержит участок *lac*-оперона с 5'-концевой частью гена *lacZ* (β - галактозидаза), что позволяет проводить селекцию с помощью сине-белого теста; *ColE1* – ориджин репликации в *E. coli*; *bla*– ген определяющий устойчивость к ампицилину; MCS – полилинкер содержащий сайты гидролиза эндонуклеаз рестрикции; *f1(-) origin* – ориджин репликации, позволяющий получить одноцепочечную ДНК; T7 – промотор фага T7; *lacI* –ген, кодирующий репрессор LacI.

2.1.3. Олигонуклеотиды

Олигонуклеотиды были составлены на основе последовательности ДНК T4 (GenBank accession no. 000866). Положения олигонуклеотидов в последовательности ДНК T4 приведены в скобках. Нуклеотидные остатки, которые не соответствуют последовательности ДНК T4, обозначены строчными буквами.

- #1: 5'-TGGTATTTACTATGCTCCATATGTAGCTCTG-3' (106282-106312),
- #2: 5'-GTGGTAGCGCGAGTCAAAGC-3' (107398-107417),
- #3: 5'-TTAAGcTTACGAGATTAACTGC-3' (107827-107849),
- #4: 5'-TGGaATTCTGGCTACAGAAATTATG-3' (105817-105841),
- #5: 5'- tttgtcGACTTCCAGGACCCAATTGA-3' (105941-105960),
- #6: 5'-tttctgcAGCATCCGCAATTTCACTATC-3' (107779-107799),
- #7: 5'-ttcatATGTTTAAAATGAAATATTTAATAT-3' (107208-107232),
- #8: 5'-ttggatCCTTTTTTCTATTTTATACGATAGC-3' (106553-106577)
- #9: 5'-ttttaagcttCATCACTCCGTTTGTATATGATTATAATATC-3' (107260-107230),
- #10: 5'-CTTTAGACGTGTATAcGTgAAgGGTATCTAATCTCTAACG -3' (106480-106519),
- #11: 5'-AGAGATTAGATACCCtTtACgTATACACGTCTAAAGTAAG -3' (106476-106515),
- #12: 5'- CTTCCATTTTATACATCTCCTCAG -3' (107022-107045)
- #13: 5'- CGTTGATTTTAGCCATTGCG -3' (107319-107338)
- #14: 5'- GGCTTCTCGCATCCAGAGC -3' (106420-106438)

#15: 5' AGCTAAGATTAGATACCTTTAACATATACACGC 3'

#16: 5' AGCTGCGTGTATATGTAAAGGTATCTAATCTT 3'

#17: 5' TTTGAAGCTTTGCGTGATTTATATGTAGAT 3'

#18 : 5' AATATATCTAGAACTACTTTCCAAGCATAGAAGTAACCA 3'

Олигонуклеотидные праймеры были разработаны с помощью программы Oligo 4. Среда, основные буферы и растворы

2.1.4. Растворы, буферы и ростовые среды

Среды

Выращивание бактерий и фагов проводили в жидких и агаризованных средах LB и 2× YT (Sambrook et al., 1989). **LB**: триптон – 10 г/л, дрожжевой экстракт – 5 г/л, NaCl – 10 г/л (при выращивании фагов среда дополнительно содержала казаминовые кислоты – 20 г/л) (pH 7,6). **2× YT**: триптон – 16 г/л, дрожжевой экстракт – 10 г/л, NaCl – 5 г/л (pH 7,6). Твердая агаризованная среда на основе LB и 2× YT содержала 15 г/л агара. Мягкая агаризованная среда сделанная на основе LB содержала 6 г/л агара или 7 г/л агарозы.

При отборе рекомбинантных клонов с помощью сине-белого теста, основанного на α-комплементации β-галактозидазной активности, к расплавленной агаризованной среде добавляли: X-gal до конечной концентрации 30 мкМ, ИПТГ до конечной концентрации 100 мкМ. Конечная концентрация антибиотиков в селективной среде для *E. coli* составляла: Ap – 50-100 мкг/мл, Tc – 12,5 мкг/мл, Km – 25 мкг/мл; Cm – 34 мкг/мл.

Буферы

Для всех буферов приведены конечные концентрации компонентов.

10× TBE буфер для электрофореза ДНК: 0,89 М трис (108 г/л), 0,89 М борная кислота (55 г/л) и 20 мМ ЭДТА (43,84 г/л).

10× трис-глициновый буфер для электрофореза белков: 250 мМ трис (29 г/л), 1,92 М глицина (144 г/л) и 1% ДДС-Na (10 г/л), pH 8,3.

1× буфер TE для растворения и элюирования ДНК: 10 мМ трис-HCl (pH 8,0) и 1 мМ ЭДТА.

1× буфер SM для титрования и хранения фагов: 10 мМ трис-HCl (pH 7,5), 150 мМ NaCl и 10 мМ MgCl₂ (или MgSO₄).

Буфер для приготовления проб для белкового фореа 80 мМ трис-HCl (pH 6,8), 2% ДДС-Na, 25 мМ β-меркаптоэтанол, 10% глицерин и 0,05% бромфеноловый синий.

Буферы для диализа фаговой суспензии:

1. Буфер А: 0,1 М трис-НCl (рН 7,45) и 3 М NaCl.
2. Буфер Б: 0,1 М трис-НCl (рН 7,45) и 0,3 М NaCl.

Для гидролиза ДНК эндонуклеазами рестрикции использовались буферы, рекомендуемые фирмами-поставщиками.

Растворы

Для всех растворов приведены конечные концентрации компонентов.

Растворы для препаративного выделения ДНК плазмид:

Раствор I: 50 мМ глюкозы, 10 мМ ЭДТА и 25 мМ Трис-НCl (рН 8,0)

Раствор II: 0,2 М NaOH и 1% ДДС-Na.

Раствор III: 3 М ацетата К (рН 4,8) (к 600 мл 5 М ацетата К (294,4 гр ацетата К) добавить 115 мл уксусной кислоты и 285 мл воды).

Растворы для получения рубидиевых компетентных клеток:

1. TfbI: 30 мМ ацетата К, 10 мМ CaCl₂, 100 мМ RbCl, 50 мМ MnCl₂ и глицерин 15% об./об. (рН 5,8) (доводить уксусной кислотой).

2. TfbII: 10 мМ MOPS рН 6,8, 10 мМ RbCl, 75 мМ CaCl₂, глицерин 15% об./об.

Раствор “ПЭГ-NaCl” для преципитации фаговых частиц: 30% ПЭГ 6000 и 3 М NaCl.

Растворы для переноса и гибридизации:

1. 20× SSC: 3 М NaCl и 0,3 М цитрат натрия (рН 7,0).

2. 50× раствор Денхардта: 1% фико-400, 1% поливинилпирролидон и 1% БСА.

3. Раствор для предгибридизации: 2× SSC, 0,5% ДДС-Na, 5× раствор Денхардта и 100 мкг/мл ДНК спермы лосося.

4. Раствор для гибридизации: 2× SSC, 0,5% ДДС-Na и 3-7 млн импульсов/мин ³³P-меченой денатурированной ДНК соответствующего “зонда”.

5. Растворы для отмывки фильтров: Раствор I содержит 2× SSC и 0,5% ДДС-Na. Раствор II содержит 0,5× SSC и 0,5% ДДС-Na.

Раствор для нанесения проб ДНК на гель: 0,15% бромфеноловый синий, 1% ДДС-Na, 0,1% ксиленианол, 1 мМ ЭДТА, 50% глицерин об./об.

Стоп-раствор для прерывания реакции гидролиза: раствор для нанесения проб ДНК на гель, в котором содержится 100 мМ ЭДТА.

Раствор Бредфорда: 100 мг/л кумаси бриллиантовый синий G-250, 4,8% этанола, 8,5% фосфорной кислоты.

2.1.5 Реактивы и материалы

В работе использовали следующие материалы и реактивы: бакто-агар, бакто-триптон, бакто-дрожжевой экстракт, (“Difco”, США); акриламид, N,N-бисакриламид, агароза, ДДС-Na, гидроксилпатит HRPT (“Bio-Rad”, США); dNTP, ddNTP (“Amersham”, Великобритания), этидиум бромид, кумасси R-250, ИПТГ, X-gal, ЭДТА, Тритон X-100, хлорамфеникол, глицерин (“Serva”, Германия), трис, CaCl₂, MgCl₂, ампициллин, дитиотреит, АТФ, БСА (5 фракция), лизоцим, рибонуклеаза А (“Sigma”, США); β-меркаптоэтанол; уксусная кислота, KCl, ПЭГ 6000 (“Merck”, Германия); наборы dNTP (“Boehringer Mannheim”, Германия); ингибиторы протеаз “Complete” (“Roche Applied Science”, Великобритания); Ni-NTA агароза (“Qiagen”, США); колонка Mono-S HR 5/5 (“Amersham Biosciences”, Великобритания, Швеция); набор для секвенирования ДНК “Sequenase version 2,0 DNA sequencing kit” (“USB Corporation”, США); набор для элюции ДНК-фрагментов из геля “QIAquick Gel Extraction Kit” (“Qiagen”, США); набор для мечения ПЦР-фрагментов “DECAprime II DNA Labelling Kit” (“Ambion”, США), набор для очистки меченных ПЦР-фрагментов “QIAquick Nucleotide Removal Kit” (“Qiagen”, США); набор для автоматического секвенирования ДНК “DTCS” – “CEQTM Dye Terminator Cycle Sequencing kit” (“Beckman Coulter”, США). В работе использовали нейлоновые фильтры Hybond-N (“Amersham Pharmacia Biotech”, Великобритания). Эндонуклеазы рестрикции и ДНК-модифицирующие ферменты: полинуклеотидкиназа фага Т4, щелочная креветочная фосфатаза (Shrimp), BamHI, BglII, EcoRI, Eco31I, EcoO109I, HindIII, KpnI, NdeI, PstI, PvuI, SalI, SmaI, XmaI, термостабильная ДНК-полимераза Taq, ДНК-лигаза фага Т4, фрагмент Кленова, а также маркеры длин фрагментов ДНК λ/HindIII и “MassRulerTM DNA Ladder” фирмы (“Fermentas”, Литва); SspI (“NEB”, USA), Acc65I, EgeI, MfeI, NruI, Sfr303I, SspI, SmlI (“Сибэнзим”, Россия); ДНК-полимеразы Accusime, Bio-X-Act с 3' редактирующей активностью, маркеры длин фрагментов ДНК “Hyper Ladder I” (“BioLine”, Германия). Все остальные использованные в работе реактивы были отечественного производства марки х.ч. или о.с.ч.

Радиоактивное соединение [α-³²P]dATP было получено из ИБХ РАН, г. Москва. В данной работе использовались электрофорезные камеры “GNA-100” (“Pharmacia Fine Chemicals”, Швеция), “FIGE Mapper Cell” (“BioRad”), “Pharmacia-LKB” (Швеция); ультразвуковой дезинтегратор “MSE” (Великобритания); 1 мм кюветы для электропорации “CE-0001-50” (“Eurogentec”, Бельгия), электропоратор “MicroPulserTM” (“BioRad”, Германия); ПЦР приборы Eppendorf и “Gene Amp PCR System 2400” (“Perkin Elmer”, США); автоматический секвенатор “CEQ2000XL DNA Analysis System” (“Beckman Coulter”, США).

2.2. Методы исследования

2.2.1. Нарботка плазмид и определение концентрации ДНК

Плазмидную ДНК нарабатывали с использованием наборов компании Fermentas, в соответствии с рекомендацией производителя. Концентрация ДНК вычислялась исходя из соотношения $OD_{260}=1$, соответствует 50 мкг/мл двухцепочечной ДНК.

2.2.2. Нарботка ДНК Т-четных бактериофагов

Получение ДНК фагов проводили, как описано Карлсон и Миллер (Carlson and Miller 1994). Для этого к фаговой суспензии добавляли 20 mM ЭДТА (pH 8,0) и 0,5% ДДС-Na, плавно перемешивали, добавляли 100 ед. (или 50 мкг/мл) протеиназы К и инкубировали при 65 °C в течение 1 ч. Удаление белков проводили экстракцией равными объемами фенола, фенол/хлороформа (1: 1) и хлороформа. После каждого добавления смесь плавно перемешивали в течение 1 ч до образования эмульсии и центрифугировали. К отобранной водной фазе добавляли 1/10 объема 3 M ацетата Na и такой же объем изопропанола. После плавного перемешивания для повышения эффективности преципитации ДНК, смесь инкубировали при комнатной температуре в течение 10 – 15 мин. Образовавшиеся нити агрегированной ДНК переносили в чистую пробирку, дважды промывали 75% этанолом, подсушивали 30 – 40 мин, растворяли в 750 мкл буфера TE и хранили при 4 °C.

2.2.3. Трансформация и приготовление компетентных клеток *E. coli*

Rb-Cl-метод получения компетентных клеток E. coli. Ночную культуру, разбавляли в 100 раз и подращивали при интенсивном перемешивании при 37°C до $OD_{590} = 0,5$. Далее все процедуры проводили при 0 °C. Клетки охлаждали на ледяной бане в течение 15 мин и центрифугировали. Осадок ресуспендировали в 0,4 объема буфера TfbI и инкубировали в течение 60 мин. Далее проводили центрифугирование, осадок ресуспендировали в 0,04 объема TfbII. Компетентные клетки фасовали по 50 мкл и хранили при -70 °C. Частота трансформации полученных компетентных клеток находилась в раене 1×10^7 клонов на 1 мкг ДНК.

Трансформация компетентных клеток E. coli. К компетентным клеткам добавляли 1 нг ДНК, перемешивали и инкубировали во льду в течение 60 мин. Далее проводили тепловой шок прогреванием клеток при 42 °C в течение 1,5 мин. Клетки охлаждали во льду в течение 5 мин, добавляли 800 мкл $2 \times$ YT и инкубировали при 37 °C в течение 60 мин. 100 мкл смеси высевали

на свежее залитые чашки Петри с агаризованной LB-средой, содержащей селективный антибиотик. Клетки росли при 37 °C до появления колоний.

2.2.4. Полимеразная цепная реакция

Полимеразную цепную реакцию проводили с использованием *Taq*, *KOD* либо *Pwo* ДНК-полимераз в соответствии с рекомендацией производителя. Температуру отжига олигонуклеотидных праймеров рассчитывалась с использованием программы Oligo 4. При анализе бактериальных клонов на наличие рекомбинантных плазмид, небольшое количество клеточной суспензии добавляли в ПЦР-смесь и проводили реакцию амплификации.

2.2.5. Обработка ДНК эндонуклеазами и лигирование ДНК-фрагментов

Гидролиз плазмидной ДНК, фаговой ДНК, и ПЦР-фрагментов. Гидролиз ДНК эндонуклеазами рестрикции проводили согласно рекомендации производителя. В экспериментах по изучению влияния природных модификаций ДНК Т4-родственных фагов на специфичность и эффективность работы эндонуклеазы SegD 0,5 мкг ДНК соответствующих фагов последовательно гидролизовали эндонуклеазой рестрикции SmaI далее 5 либо 10 ед. эндонуклеазы SegD при 30 °C в течение 30 мин. Для оценки размеров образующихся фрагментов ДНК использовали маркеры Bio-Rad 8-48 kbp и Q-Step4 (YorBio).

Лигирование фрагментов ДНК. Для проведения лигирования использовали ДНК фрагменты, выделенные из агарозного геля при помощи набора QIAquick Gel Extraction Kit (Qiagen). Лигирования фрагментов ДНК проводили согласно рекомендации производителя Т4 ДНК-лигазы (Fermentas).

2.2.6. Получение рекомбинантных плазмид.

Клонирование ДНК проводили при помощи стандартных методов (Sambrook et al., 1989). Секвенирование гена *segD* бактериофага Т4 показало, что размер данного гена составляет 672 п.н. Плазмида pSDET-19mod содержит ген *segD* и кодирует белок SegD слитый на N-конце с с His₆ и сайтом TEV-протеазы. Плазмида была получена клонированием фрагмента ДНК Т4 размером 680 п.н., наработанного при помощи ПЦР с праймерами #7 и #8 , по сайтам *NdeI* – *BamHI* вектора pET-19mod.

Плазмида pUT2LsD содержит фрагмент области генов 23-24 фага Т2L размером 1.4 т.п.н, который был амплифицирован с праймерами #3 и #4 и клонирован в *EcoRI* и *HindIII* сайты pUC18.

Плазмида pBluT4sDs содержит фрагмент области генов 23-*segD* фага T4 размером 1084 п.н. (105941_107024 bp. T4 NC_000866). Плазмида была сконструирована на основе вектора pBluescriptKS (-) (Stratagene) путем клонирования фрагмента ДНК фага T4 размером 1084 п.н., амплифицированного с парой праймеров #5 и #6, и гидролизованного по сайтам *Sall* – *HindIII*.

Плазмида pBluT4sDdel содержит фрагмент области генов 23-24 фага T4 с делецией 5'-концевой области ОРС *segD* (3-208 н.о.), которая приводит к сдвигу рамки считывания. Она была сконструирована на основе плазмиды pBluT4sDs путем клонирования ПЦР-фрагмента ДНК фага T4 размером 570 п.н. (107230_107799 b.p. T4 NC_000866), амплифицированного с парой праймеров #6 и #9, по сайтам *HindIII* и *PstI*.

Плазмида pBluT4sDs_mut является производным плазмиды pBluT4sDs и была получена путем введения замен в последовательность сайта гидролиза *SegD*. Сайт-направленный мутагенез плазмиды pBluT4sDs проводили при помощи набора QuikChange Site-Directed Mutagenesis Kit (Stratagene) с использованием праймеров #10 и #11. Плазмида pBluT4sDs_mut использовалась при конструировании бактериофага T4 с нарушенным сайтом *SegD*.

Плазмида pURIIBsDs содержит фрагмент *rII*-локуса, в котором в 5'-конец гена *rIIB* встроен фрагмент ДНК размером 37 п.н. с сайтом *SegD*. Для получения данной плазмиды, олигонуклеотиды #15 и #16 были отождествлены, достроены фрагментом Кленова и клонированы в сайт *HincII* плазмиды pUrIIB-18 (Shcherbakov et. al., 2002). Плазмида pURIIBsDs использовалась для введения в ген *rIIB* фага T4 сайта *SegD*.

2.2.7. Получение фагов T4Δ*segD*, D2, B20dm, D2insZ и D2ets9

Бактериофаг T4Δ*segD* содержит делецию с 3 по 208 н.о. в ОРС *segD*. Данный фаг был получен путем скрещивания фага T4 и плазмиды pBluT4sDdel. Скрещивание между плазмидой и фагом проводили, как описано ранее (Brok-Volchanskaya et al., 2008). Селекцию рекомбинантных фагов проводили при помощи гибридизации с ³²P-меченным ДНК-зондом к области делеции, который нарабатывали при помощи ПЦР с праймерами #7 и #12. Были отобраны фаги, которые давали отрицательный сигнал на гибридизацию. Наличие делеции в ОРС *segD* подтверждали при помощи ПЦР с праймерами #1 и #2.

Бактериофаг D2 является производным фага T4, у которого область генов 23 – 24, замещена на гомологичную область фага T2L. Данный фаг был получен путем скрещивания фага T4Δ*segD* с плазмидой pUT2LsD. Фаговое потомство гибридизовали с ³²P-меченным ДНК-зондом к ОРС *segD*, который был получен при помощи ПЦР с праймерами #7 and #8. Были отобраны фаги, дающие отрицательный сигнал на гибридизацию. Полученные рекомбинантные фаги дополнительно проверяли при помощи ПЦР с праймерами #1 и #2.

Бактериофаг B20dm содержит амбер-мутацию в гене 14, а также замены 3 н.о. в последовательности сайта SegD. Эти замены не приводят к изменению аминокислотной последовательности gp23. Данный фаг был получен путем скрещивания фага T4amB20 (B20) и плазмиды pBluT4sDs_mut. Отбор рекомбинантных фагов проводили при помощи гибридизации с ^{32}P -меченным олигонуклеотидом #11, специфичным к мутантному сайту SegD. Наличие замен в последовательности сайта SegD подтверждали секвенированием.

Фаги D2ets9 и D2insZ являются производными фага D2, которые содержат на расстоянии 108 н.о. от начала гена *rIIB* сайт SegD (*ets9*) или фрагмент гена *lacZ* (*insZ*), соответственно. Данные фаги имеют *rII*-фенотип. Они были получены путем скрещивания фага D2 с плазмидой pURIIBsDs (D2ets9), и pUlacZ (D2insZ) (Shcherbakov et al., 2002). Фаговый урожай высевали в разведении на агаризованную среду с газонной культурой *E. coli* В. Отдельные фаговые проверяли на наличие *rII* фенотипа высевая на *E.coli* CR63(λ) и CR63. Далее, отобранные фаги анализировали при помощи ПЦР к данной области, сиквенса, а также, последующего гидролиз ПЦР-фрагментов эндонуклеазой SegD.

2.2.8. Определение нуклеотидной последовательности ДНК

Для всех полученных в ходе работы рекомбинантных плазмид была определена нуклеотидная последовательность ДНК. Секвенирование ДНК проводилось на автоматическом секвенаторе “CEQ2000XL DNA Analysis System” в соответствии с рекомендациями производителя.

2.2.9. Скрининг рекомбинантных плазмид

Единичные клоны пересевали по секторам на LB-агар, содержащий соответствующий антибиотик, и выращивали в течение 8 – 14 ч при 37 °С. Бактериальный мазок ресуспендировали в 15 мкл буфера I для выделения плазмидой ДНК, добавляли 15 мкл раствора II для скрининга колоний и прогревали в течение 30 мин при 75 °С. Наличие вставки оценивали по замедлению подвижности анализируемых плазмид в 1% агарозном геле по сравнению с подвижностью векторной плазмиды, полученной аналогичным методом.

2.2.10 Разделение фрагментов ДНК в электрическом поле

Разделение фрагментов ДНК в агарозном геле. Электрофоретическое разделение фрагментов ДНК проводили в 0,8-1% агарозе, содержащей 5 мкг/мл бромистого этидия, в 0,5× TBE при напряжении электрического поля 110-140 В. Образцы наносили на гель в “растворе для нанесения проб на гель”. После электрофореза гель визуализировали в УФ-свете при длине волны 254 или 360 нм.

Разделение фрагментов ДНК в денатурирующем ПААГ. Электрофоретическое разделение фрагментов ДНК проводили в 6% ПААГ (для приготовления 6% ПААГ к 50 мл денатурирующего геля, содержащего 7 М мочевины, акриламид:бисакриламид в соотношении 19 : 1 и 1× TBE, добавляли 50 мкл ТЕМЕД и 500 мкл 10% ПСА). Перед заливкой геля одно из плат обрабатывали диметилдихлорсиланом, а другое – смесью из 10 мл ацетона, 40 мкл ледяной уксусной кислоты и 40 мкл “Bind Silan”. Обработанные поверхности промывали водой и подсушивали. Преэлектрофорез проводили в 1× TBE при напряжении 2000 В в течение 30 мин, а электрофорез в 1× TBE при напряжении 2000 В в течение 2,5 ч. После электрофореза гель выдерживали в течение 15 – 20 мин в 10% уксусной кислоте и 20 мин в проточной воде. Затем гель промывали дистиллированной водой и сушили в течение 30 мин с помощью фена. Радиоавтографию проводили в течение 1 – 2 дней.

Разделение фрагментов ДНК в инверсионных полях. Разделение фрагментов ДНК, полученных после гидролиза эндонуклеазой SegD ДНК бактериофагов, проводили в камере “FIGE Mapper Cell” (Bio-Rad) с расстоянием между электродами 29 см. Скорость рециркуляции буфера составляла 600 мл/мин. Для создания инверсионного поля устанавливали напряжение прямого тока – 180 В и обратного – 120 В. Электрофорез проводили в течение 14 ч с линейным изменением продолжительности прямого и обратного пульса от 0,5 до 0,1 сек. (значения напряженности электрического поля в электрофорезной камере составляли при прямом токе – 6,2 В/см, при обратном – 4,1 В/см). Для определения размера фрагментов ДНК применялись маркеры Q-Step4 (YorBio) (10-0.1 т.п.н.), Bio-Rad 8-48 т.п.н. После электрофореза фрагменты ДНК визуализировали окрашиванием бромистым этидием в течение 40 мин, затем отмывали в дистиллированной воде в течение 1 – 1,5 ч и фотографировали в проходящем УФ-свете.

2.2.11 ДНК гибридизация по методу Саузерна

Гибридизацию по Саузерну проводили на незаряженных нейлоновых фильтрах Hybond-N щелочным методом как описано Сэмбруком.

Получение меченых α -³²P- фрагментов ДНК для проведения гибридизации. Для получения меченных фрагментов ДНК использовали ПЦР-фрагменты. Выделение ПЦР-фрагментов проводили из 1% агарозного геля с использованием набора для выделения ДНК из агарозного геля (Quagen) и метили их методом “рассеянной затравки” при помощи набора “DECAprime II DNA Labelling Kit” согласно рекомендациям производителя. Денатурацию 25 нг ПЦР-фрагмента проводили в 10 мкл реакции, содержащей 2,5 мкл декаолигонуклеотидов и 0,06 мМ ЭДТА при 100 °С в течение 5 мин. Затем реакционную смесь быстро охлаждали во льду, добавляли 5 мкл 5× буфера, содержащего смесь из трех dNTP (за исключением dATP), α -³²P-

меченный dATP (4000 Ки/мМ), 5 ед. фрагмента Кленова exo^- , доводили объем до 25 мкл и инкубировали при 37 °С в течение 30 мин. Реакцию останавливали добавлением 16 мМ ЭДТА и прогреванием при 65 °С в течение 20 мин. Очистку меченных ПЦР-фрагментов от невключившейся метки проводили при помощи “QIAquick Nucleotide Removal Kit”. Зонд растворяли в 100 мкл деионизованной воды. Включение радиоактивной метки в ДНК (импульсы/мин) высчитывали по Черенкову как отношение количества единиц активности, содержащейся в 1 мкл раствора зонда, за 1 мин к эффективности счета 0,3 и умноженное на общий объем смеси, содержащей зонд (в данном случае – 100).

Гибридизация ДНК-фрагментов и ДНК негативных колоний бактериофагов.

Перенос ДНК. а) *Перенос ДНК негативных колоний бактериофагов с чашек Петри на нейлоновые фильтры:* Разведениями фагового потомства инфицировали 0,1 мл газонной культуры *E. coli* и выращивали бактерии в слое верхнего агара (для приготовления мягкого агара использовали агарозу) при 37 °С в течение ночи, так, чтобы на одной чашке Петри выросло 200 – 800 негативных фаговых колоний размером 1 – 2 мм. Чашки Петри охлаждали при 4 °С в течение 1 ч. Нейлоновые фильтры Hybond-N накладывали на поверхность чашек Петри, маркировали их ориентацию и переносили негативные фаговые колонии в течение 1 мин. ДНК фаговых колоний на фильтрах денатурировали в 0,5 М NaOH, 1,5 М NaCl в течение 5 мин. Затем фильтры последовательно промывали в течение 5 мин в 0,5 М трис-HCl (pH 7,5), 1,5 М NaCl, 1× SSC и в течение 1 мин в 96% этаноле, подсушивали и “пришивали” ДНК к нейлону в течение 2 мин облучением УФ-светом при длине волны 254 нм.

б) *Перенос ДНК из агарозного геля на нейлоновый фильтр:* Агарозный гель после электрофореза выдерживали в 200 мл 0,25 М HCl в течение 25 мин. Этот этап необходим для частичной депуринизации ДНК, что приводит к разрыву цепей и способствует эффективному переносу молекул размером более 5 т.п.н. После промывания в дистиллированной воде гель выдерживали в 200 мл 0,25 М NaOH в течение 20 мин. Перенос ДНК на фильтр проводили в течение 2 ч в 300 мл раствора, содержащего 0,25 М NaOH и 1,5 М NaCl. Затем фильтр обрабатывали 2× SSC с целью нейтрализации, высушивали на воздухе и ДНК на нейлоне иммобилизовали облучением УФ-светом при длине волны 254 нм.

Гибридизация. Фильтры выдерживали в течение 2 ч при 65 °С в 25 мл “раствора для предгибридизации”, высушивали и смачивали в прогретом до 65 °С “растворе для гибридизации” из расчета 1 мл раствора на фильтр диаметром 90 мм. Непосредственно перед добавлением к гибридизационному раствору, ДНК-зонд денатурировали прогреванием в кипящей бане в течение 5 мин и охлаждали при 4 °С. После смачивания гибридизационным раствором фильтры немедленно погружали в сосуд с вазелиновым маслом и выдерживали в течение 16 ч при 65 °С.

Отмывка. После гибридизации фильтры, содержащие ДНК фагов, последовательно отмывали по два раза в течение 15 мин при 65 °С в 250 мл “растворов для отмывки” I и II, а фильтры, содержащие ДНК-фрагменты, дважды отмывали в 200 мл раствора I в течение 15 – 20 мин при постоянном покачивании до фона приблизительно 25 – 50 cpm на дорожку. Затем фильтры высушивали, маркировали радиоактивными чернилами и проводили радиоавтографию.

2.2.12 Электрофоретическое разделение белков в полиакриламидном геле

Электрофорез белков проводили в ПААГ по методу Лэммли (Laemmli, 1970). Соотношение акриламид/бисакриламид в концентрирующем и разделяющий гель составляло 29:1.

Анализируемые образцы кипятили в течение 5 мин в “буфере для приготовления проб для белкового фореа”. Электрофоретическое разделение проводили в трис-глициновом буфере при начальном напряжении 80 В с последующем его увеличением до 150 В. После электрофореза гель окрашивали при 100 °С раствором, содержащим кумасси бриллиантового синего R-250, а затем 2 – 3 раза отмывали в дистиллированной воде при 100 °С.

2.2.13. Получение TEV-протеазы

Экспрессию проводили по методу, предложенному Штудьером (Studier et al., 1990), с некоторыми модификациями. В качестве хозяина использовали штамм *E.coli* BL21(DE3)/pLysE. Клетки трансформировали плазмидой pET-19Tev и высевали на 3 чашки с агаризованной средой LB+0.2% глюкозы, с ампицилином (80 мкг/мл) и хлорамфениколом (30 мкг/мл) из расчета 300-400 колоний на одну чашку. Клетки растили в течение ночи при 37 °С. Затем клетки суспендировали в 30 мл LB, растирая биомассу шпателем Дригальского, полученной суспензией инокулировали 0,4 л среды 2xYT (2 колб по 200 мл среды) с ампицилином (100 мкг/мл) и хлорамфениколом (30 мкг/мл). Клетки подращивали на качалке при 37 °С до $OD_{590}=0.9$, далее индуцировали добавлением ИПТГ до концентрации 0.05 мМ/мл. После чего культура продолжала расти в течение 18 часов при 20 °С. Затем клетки осаждали на центрифуге J2-21 в роторе JA-10, 5000 об/мин, 4 °С в течение 10 минут и ресуспендировали в 30 мл буфера для озвучивания.

2.2.14. Очистка TEV-протеазы

Буферы для хроматографической очистки Tев-протеазы:

Буфер для разрушения: 300 мМ NaCl, 50 мМ K/Na- фосфатный буфер, pH 7.5, 5 мМ имидазол, pH 7.5, 5 мМ 2-меркаптоэтанол;

Промывочный буфер: 300 мМ NaCl, 50 мМ K/Na- фосфатный буфер, pH 7.5, 50 мМ имидазол, pH 7.5, 5 мМ 2-меркаптоэтанол, 0.01% Triton X-100;

Элюирующий буфер: 50 мМ K/Na- фосфатный буфер, pH 7.3, 250 мМ имидазол pH 7.5, 5 мМ 2-меркаптоэтанол;

Буфер для диализа: 50 мМ Tris-HCl pH 7.5, 1мМ ЭДТА, 10 мМ 2-меркаптоэтанол, 0.01% Triton X-100

Выделение TEV-протеазы проводили при 4⁰С. Клетки разрушали в присутствии ингибитора протеаз PMSF, конечная концентрация которого составляла 0.5 мМ. Озвучивание клеток и получение грубого экстракта проводилось при тех же условиях что и для SegD. После озвучивания бесклеточный экстракт наносили на колонку с 1.2 мл Ni-NTA агарозы (Qiagen), уравновешенной буфером для разрушения. После посадки образца колонку промывали буфером для промывки (15 мл). Белки элюировали 5 мл буфера для элюции. Фракцию очищенного белка последовательно диализовали против буфера для диализа и буфера для диализа, содержащего 50% глицерина. Фермент хранили при -70⁰С.

2.2.15. Экспрессия OPC *segD*

Экспрессию проводили по методу, предложенному Штудьером (Studier et al., 1990), с некоторыми модификациями. В качестве хозяина использовали штамм *E.coli* BL21(DE3)/pLysE. Клетки трансформировали плазмидой pSDET-19mod. Трансформированные клетки высевали на 3 чашки с агаризованной средой LB+0.2% глюкозы, с ампицилином (70 мкг/мл) и хлорамфениколом (30 мкг/мл) из расчета 300-400 колоний на одну чашку и растили в течение ночи при 37⁰С. Затем клетки суспендировали в 30 мл LB, растирая биомассу шпателем Дригальского, и полученной суспензией инокулировали 0,8 л среды 2хYT (4 колб по 200 мл среды) с ампицилином (100 мкг/мл) и хлорамфениколом (30 мкг/мл). Клетки подращивали на напольной качалке при 37⁰С и 110 об/мин до ОД₅₉₀ =0.5 (2х10⁸ кл./мл) и индуцировали добавлением ИПТГ до концентрации 0.5 мМ/мл, а также ампицилина до 150 мкг/мл. После чего культуру продолжали растить при температуре 26 ⁰С в течении 3.5 часов. Затем клетки осаждали на центрифуге J2-21 в роторе JA-10, 5000 об/мин, 4⁰С в течение 10 минут и ресуспендировали в 160 мл буфера для озвучивания.

2.2.16. Очистка эндонуклеазы SegD

Буферы для хроматографической очистки эндонуклеазы SegD:

1. Буфер для разрушения: 20 mM Tris-HCl pH 8, 300 mM NaCl, 5 mM имидазол, pH 7.5, 5 mM 2-меркаптоэтанол;
2. Промывочный буфер: 20 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM имидазол, pH 7.5, 5 mM 2-меркаптоэтанол;
3. Элюирующий буфер: 20 mM Tris-HCl pH 8, 300 mM NaCl, 250 mM имидазол, pH 7.5, 5 mM 2-меркаптоэтанол;
4. Буфер для диализа: 20 mM Tris-HCl pH 7.45, 200 mM NaCl, 0,1mM ЭДТА, 10 mM 2-меркаптоэтанол.

Выделение и очистку SegD проводили при 4⁰С. На всех этапах хроматографической очистки фракции, содержащие фермент, отбирали по специфической эндонуклеазной активности и соотносили с данными белкового электрофореза. Клетки разрушали в присутствии ингибитора протеаз PMSF на установке MSE (Великобритания) при частоте 20кГц: 9 раз по 30 сек. с интервалом в 1 мин. для охлаждения. Бесклеточный экстракт, полученный центрифугированием на J2-21 в роторе JA-20, 18 000 об/мин, 4⁰С в течение 4 часов, наносили на колонку с 1.5 мл Ni-NTA агарозы (Qiagen), уравновешенной буфером для разрушения. Затем колонку промывали буфером для промывки (15 мл). Белки элюировали 6 мл буфера для элюции. К активной фракции добавляли TEV-протеазу и диализовали против диализного буфера. Количество добавленной TEV-протеазы определялась молярным соотношением 1/50-1/25 фермент / субстрат. Дальнейшим этапом было удаление TEV-протеазы. Образец после диализа и отщепления His-tag наносили на колонку с 0.5 мл Ni-NTA агарозы (Qiagen). Фракцию находящуюся в проскоке далее диализовали и переводили в буфер содержащий глицерин. Фермент хранили при – 20 ⁰С. *Определение концентрации белка.* Концентрацию белка определяли по поглощению белка при 280 нм.

2.2.17. Определение оптимальных условий работы эндонуклеазы SegD *in vitro*

Реакция гидролиза включала в себя следующие компоненты: 0.2 мкг ДНК плазмиды pUT2LsD, линеаризованной по SspI, 50 mM Tris-HCl pH 8, 10 mM MgCl₂, 10 mM β-меркаптоэтанол, 0.02% Тритон X-100 и 0.1 мг/мл БСА, а также 1 ед. эндонуклеазы SegD в 10 мкл. Обычно реакция проходила при 30⁰С в течение 30 мин.. Реакция гидролиза останавливалась охлаждением и добавлением ЭДТА до концентрации 50 mM. За одну единицу активности SegD было принято количество фермента, достаточное для 50 % расщепления 0,2 мкг линеаризованной плазмидной ДНК pUT2LsD.

При изучении влияния на гидролиз различных факторов таких как: pH, ионная сила, температура, концентрация Mg²⁺, БСА и ТХ-100 варьировались. При определении оптимума

pH, гидролиз в диапазоне pH 5.5 и 6.5 проводился в 50 mM MES буферной системы, в диапазоне pH 7.0 – 9.0 - в 50 mM Tris-HCl буфере и при pH 10.0 - в 50 mM Tris-Glycine буфере.

В экспериментах по проверке влияния BSA и TX-100 на стабильность SegD исследуемый белок разводился в буфере (20 mM Tris-HCl pH 7.5, 100 mM KCl, 0,1 mM ЭДТА, 10 mM β -меркаптоэтанола) в отсутствии либо присутствии BSA и TX-100, инкубировался при 30°C. Далее отбирались временные точки 0, 10, 30, 60, 120 мин. в которых определялась остаточная эндонуклеазная активность в стандартной реакции гидролиза.

Продукты реакции разделяли в 1% агарозном геле, окрашенном бромистым этидием. Процент гидролизованного субстрата определялся с использованием системы гель документирования Kodak Electrophoresis and Analysis System (EDAS290) с последующим обсчетом данных в программе OptiQuant (Packard Instrumental Co). Представленные данные были получены в трех независимых экспериментах.

2.2.18. Установление точек расщепления и границ узнаваемой последовательности ДНК эндонуклеазой SegD

Установление точек расщепления проводили на ДНК, селективно меченной по одной цепи. Для этого 2.5 мкг предварительно денатурированной плазмиды pUT2LsD отжигали с 1 пмоль соответствующего олигонуклеотидного праймера, и проводили реакцию мечения ДНК с [α - 32 P]dATP в соответствии с протоколом к набору для секвенирования ДНК “Sequenase version 2.0 DNA sequencing kit” (Amersham Biosciences). Для удлинения отожденного праймера к 8 мкл смеси добавляли 1 мкл 2 mM дНТФ и продолжали реакцию в течение 10 мин. «Секвеназу» инактивировали прогреванием при 65°C в течение 20 мин., затем 2 мкл полученной смеси обрабатывали эндонуклеазой SegD в 10 мкл реакционного буфера содержащего: 10 mM Трис-HCl (pH 8,5), 50 mM KCl, 2 mM MgCl₂, 10 mM β -меркаптоэтанол, 0.1 мг/мл БСА и 0.02% Тритон X-100 при 30°C в течение 1 ч. Количество фермента в реакционной смеси составляло 2.5 ед. Определение точек расщепления в гене 23 проводили независимо для каждой цепи ДНК с использованием праймера #13 и #14.

Границы сайта узнавания SegD устанавливали по методу, предложенному Венцлау с соавт. (Wenzlau e al., 1989), суть которого заключается в определении минимальной длины фрагмента ДНК, необходимого для расщепления. Для получения фрагментов, отличающихся по длине на 1 нуклеотид, проводили стандартную реакцию секвенирования ДНК согласно протоколу фирмы USB (“Sequenase version 2.0 DNA sequencing kit”), за тем исключением, что реакцию терминации останавливали прогреванием при 65°C в течение 20 мин. для инактивации «Секвеназы». Затем к 6 мкл каждой терминирующей реакции добавляли 1 мкл БСА (1 мг/мл), 1

мкл 0.2% Тритона X-100 и 2.5 ед. SegD, и инкубировали при 30°C в течение 2 ч. Реакцию останавливали добавлением 6 мкл стоп-раствора. Продукты реакций разделяли электрофоретически в 6%-ном ПААГ в денатурирующих условиях. В качестве маркеров длин фрагментов использовали продукты реакций секвенирования плазмиды pUT2LsD с соответствующими праймерами. Для определения границ сайта узнавания в верхней и нижней цепях использовали олигонуклеотиды 7 и 1, соответственно.

2.2.19. Анализ эндонуклеазной активности SegD на ДНК бактериофагов

Получение в высоком титре, очистку фагов в градиенте хлористого цезия и выделение ДНК фагов проводили, как описано (Carlson and Miller, 1994). 0.5 мкг гидролизованной эндонуклеазой рестрикции *SmiI* (СибЭнзим) ДНК бактериофагов обрабатывали 5 – 10 ед. SegD в буфере содержащем: 50 mM KCl, 10 mM Трис-HCl (pH 8,5), 2 mM MgCl₂, 10 mM β-меркаптоэтанол, 0.02% Тритон X-100 и 0.1 мг/мл БСА. Продукты гидролиза разделяли в 1% агарозном геле при помощи инверсионного пульс-электрофореза в 0.5× TBE буфере. После электрофореза фрагменты ДНК визуализировали окрашиванием бромистым этидием. Для идентификации фрагментов ДНК, содержащих гены 23 и 24, проводили гибридизацию по Саузерну, как описано Сэмбруком с соавт. В качестве зондов для фага T2L и вариантов T4 использовали меченный щелочной фосфатазой ПЦР-фрагмент ДНК T4, наработанный с парой праймеров #5 и #6. Мечение ПЦР-фрагмента проводили в соответствии с протоколом производителя Amersham Biosciences (Gene Images AlkPhos Direct Labelling and Detection System RPN3680).

2.2.20. Определение эндонуклеазной активности SegD в бесклеточных экстрактах *E.coli*, зараженных бактериофагом T4

Для определения эндонуклеазной активности SegD, клетки *E. coli* BL21 выращивали в LB при 37 °C до плотности 5×10^8 клеток/мл. К культуре клеток *E.coli* добавляли фаги T4B или T4Δ*segD*, соответственно, из расчета множественности заражения равной 10 и выдерживали 5 мин. при комнатной температуре. Инфицированные клетки далее инкубировали еще 10 мин. при 37°C и 150 rpm. Далее отбирали 5 мл инфицированной клеточной культуры и добавляли равный объем охлажденного до 0 °C раствора EDTA-Saline (150 mM NaCl, 50 mM EDTA pH 8). После чего клетки концентрировали и ресуспендировали в 500 мкл раствора для озвучивания (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM β-мер, 1 mM EDTA). Озвучивание проводили в течение 10 сек. 2 раза при частоте 20 кГц с 2 мин. перерывом для охлаждения на установке

MSE (UK). Клеточный дебрис удаляли центрифугированием. Далее бесклеточный экстракт анализировался на наличие сайт-специфической эндонуклеазной активности, соответствующей эндонуклеазе SegD. В качестве субстрата использовалась линейаризованная плаزمиды рUT2LsD/SspI. Реакционная смесь содержала 10 mM Tris-HCl pH 8.5, 50 mM KCl, 1 mM DTT, 2 mM MgCl₂, 0.35 мкг ДНК плазмиды рUT2LsD, а также 1 мкл бесклеточного экстракта с A₂₈₀ равным 3, реакция гидролиза проводилась в 10 мкл. Реакцию инкубировали при 27°C в течение 30 мин. и останавливали охлаждением и добавлением ЭДТА до концентрации 50 mM. Продукты реакции разделяли в 1% агарозном геле, содержащем бромистый этидий.

2.2.21. Получение Т-четных бактериофагов и определение титра фагов

Получение фагов в высоком титре и очистку в градиенте хлористого цезия проводили, как описано Карлсон и Миллер (Carlson and Miller 1994).

Для получения первичного лизата, негативную фаговую колонию, выращенную на супрессорном штамме B40suI *E. coli*, вырезали из мягкого агара стеклянным капилляром, ресуспендировали в 100 мкл фагового буфера и инкубировали в течение 30 мин при комнатной температуре. Ночную культуру B40 *E. coli* разводили в 100 раз и растили в 5 – 10 мл LB с казаминовыми кислотами на роторной качалке при 37 °C до ОД₅₉₀ = 0,2. К культуре добавляли 10 мкл экстракта фага и инкубировали при 37 °C до полного лизиса, затем добавляли 1/100 объема хлороформа и продолжали инкубировать в течение 20 мин. Клеточный дебрис удаляли центрифугированием (K-23, 5000 об/мин, 4 °C, 15 мин) и осветленный лизат хранили при 4 °C. Концентрацию инфекционноспособных частиц бактериофага определяли через три дня титрованием на культуре клеток штамма B40 suI *E. coli* (см. ниже).

Для получения вторичного лизата, по 2 мл ночной культуры клеток B834 *E. coli* вносили в 4 колбы по 200 мл LB, содержащей казаминовые кислоты, растили при интенсивной аэрации до ОД₅₉₀ = 0,2 в течение 3 ч при 37 °C и заражали фаговыми частицами из полученных ранее лизатов с множественностью инфекции 1. Для равномерной адсорбции фаговых частиц на клетках, фаги вносили в 10 мл LB. Через 2 мин проводили инфекцию тем же количеством фагов. Далее проводили инкубацию при 37 °C и интенсивной аэрации (примерно 9 – 10 ч) до полного лизиса бактериальных клеток. Затем в каждую колбу добавляли 1/100 объема хлороформа и продолжали инкубировать в течение 30 мин при тех же условиях. Клеточный дебрис удаляли центрифугированием. Далее к супернатанту была добавлена панкреатическая ДНКазу и РНКазу А до конечной концентрации 1 мкг/мл и инкубировали в течение 30 мин при 37°C. Для преципитации фаговых частиц к каждому 200 мл фаговой суспензии добавляли по 100 мл раствора “ПЭГ-NaCl” и инкубировали при 4 °C в течение ночи. Фаговые частицы осаждали

(J2-21, 11000 об/мин, 4 °C, 1 ч), осадок ресуспендировали в 30 мл фагового буфера и инкубировали в течение ночи при 4 °C. После прогревания фаговой суспензии при 43 °C в течение 15 мин, проводили осаждение остатков клеточного дебриса (K-23, 6000 об/мин, 4 °C, 30 мин). В супернатант добавляли несколько капель хлороформа и хранили при 4 °C.

Определение содержания фаговых частиц в лизате (титрование). Для определения титра фага, разведениями фагового потомства инфицировали 0,2 мл культуры *E. coli* с $OD_{590} = 0,7 - 0,8$, затем ее смешивали с мягким агаром и выливали на твердую агаризованную среду LB. Выращивание в слое верхнего агара проводили инкубированием при 37 °C в течение 16 ч до появления фаговых бляшек.

2.2.22. Скрещивание бактериофагов и анализ потомства

Для скрещивания бактериофагов ночную культуру *E. coli* разводили в 100 раз средой LB и растили при 37 °C на роторной качалке до $OD_{590} = 0,8$ (5×10^8 кл/мл). Инфекцию проводили путем добавления 1 мл культуры к 1 мл среды LB, содержащей по $2,5 \times 10^9$ фаговых частиц/мл каждого из бактериофагов. Для абсорбции фагов смесь выдерживали в течение 5 мин при комнатной температуре, а затем инкубировали при 37 °C в течении 90 мин. После этого добавляли 1/100 объема хлороформа и продолжали инкубировать в течение 20 мин. Клеточный дебрис удаляли центрифугированием (K-23, 5000 об/мин, 4 °C, 15 мин).

Для анализа доли *amber*-мутации, потомство, полученное после скрещиваний, высевали в разведении на штаммы В_E (Sup⁰) и B40 (Su I) *E. coli*. Долю потомства с *amber*-мутацией определяли как отношение разницы титров на бессупрессорном и супрессорном штаммах к титру на супрессорном, умноженное на 100. В экспериментах по анализу наследования *segD* потомством фагов, полученном после скрещиваний, долю фагов, содержащих ОРС *segD*, определяли как отношение числа фагов, давших положительный сигнал на гибридизацию с ДНК-зондом к ОРС *segD*, к общему числу проанализированных фагов, умноженное на 100.

2.2.23. Программы, использованные в работе

Построение графиков проводилось в программе “SigmaPlot Graph v.9.0”. Программа “OptiQuant v.3.0” использовалась для анализа изображений. Анализ последовательностей ДНК, а также выбор стратегии клонирования осуществлялся с использованием программного пакета Vector NTI Suite 9. Разработка праймеров осуществлялась с использованием программы Oligo4.

ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1. Структурная организация области генов 23-24 у Т4-родственных бактериофагов

Гены хоминг-эндонуклеаз неравномерно распространены среди Т-четных фагов (Sharma and Hinton, 1994; Kadyrov et al., 1996, Sandegren et al., 2005). Поэтому первым этапом работы был анализ распределения гена *segD* среди фагов, родственных Т4. Ген *segD* бактериофага Т4 располагается между генами 23 и 24, кодирующих белки капсида оболочки.

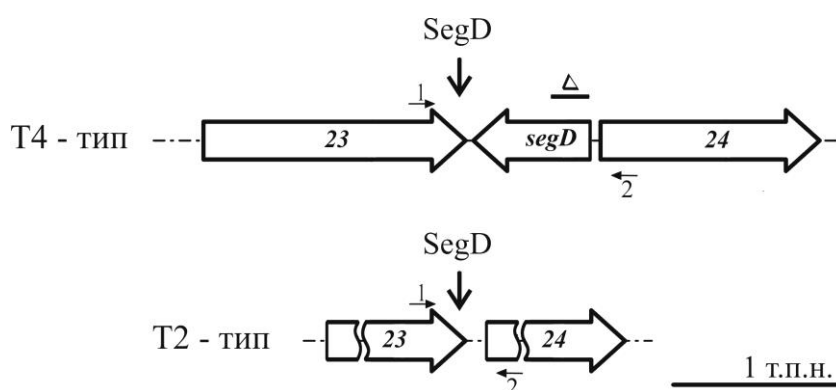


Рис. 12. Генетические карты области генов 23 - 24 Т-четных бактериофагов. Внизу приведен относительный размер в т.п.н. Вертикальными стрелками обозначено положение сайта расщепления эндонуклеазы SegD. Положение праймеров указано горизонтальными стрелками над и под последовательностями генов. Δ - область делеции, внесенная в ОРС *segD* Т4.

С целью изучения встречаемости гена *segD* среди других представителей семейства Т-четных фагов, был проведен ПЦР-анализ области генов 23 - 24 фагов, родственных Т4. Были проанализированы следующие фаги: Т2L, Т6, RB1, RB2, RB3, RB4, RB6, RB7, RB8, RB9, RB10, RB15, RB55 и RB59. Поскольку исследуемые фаги являются близкородственными Т4 и проявляют высокую гомологию последовательностей ДНК, для амплификации исследуемых областей анализируемых фагов, были использованы праймеры 1 и 2 (рис. 12), комплементарные последовательности ДНК генов 23 и 24 фага Т4 (NC_000866). После амплификации соответствующей области генов, проводили сравнение размеров полученных ПЦР-фрагментов между исследуемыми фагами и фагом Т4 (рис. 13).

Оказалось, что одинаковыми по размеру с Т4 являются ПЦР-фрагменты, наработанные с ДНК фагов Т6, RB55 и RB59 (~1100 п.н.), что косвенно указывает на сходную структуру данной области у этих фагов.

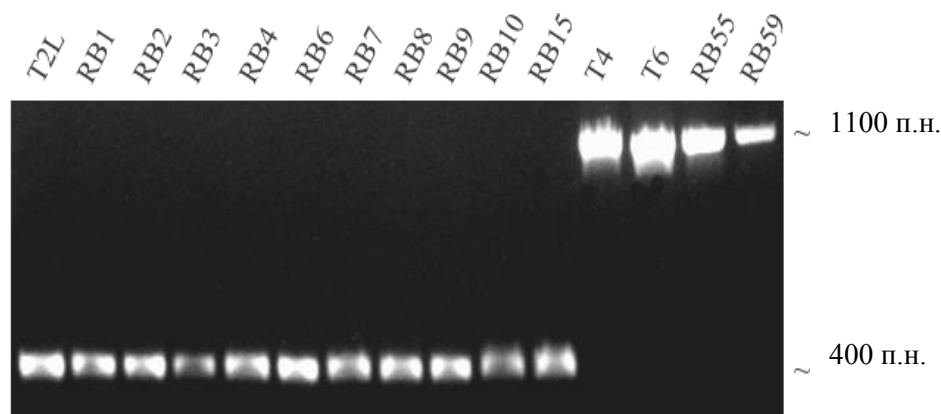


Рис. 13. ПЦР-анализ области генов 23 - 24 родственных Т-четным бактериофагов. Справа указан размер амплифицированных фрагментов ДНК.

В случае фагов Т2L, RB1, RB2, RB3, RB4, RB6, RB7, RB8, RB9, RB10 и RB15, размер ПЦР-фрагмента составил ~400 п.н, что однозначно указывало на то, что в геномах данных фагов в области генов 23 - 24 отсутствует ОРС *segD*, размер которой составляет 672 п.н.

Далее была определена нуклеотидная последовательность данной области у исследуемых фагов. Определенные нуклеотидные последовательности были зарегистрированы в банке данных GenBank: GU187339.1 - GU187352. Было установлено, что генетическая организация области генов 23-24 у фагов Т2L, RB1, RB2, RB3, RB4, RB6, RB7, RB8, RB9, RB10 и RB15 одинакова. Структура правого и левого флангов секвенированных фрагментов у различных фагов сходна с бактериофагом Т4, и включает фрагменты генов 23 и 24. У бактериофагов Т6, RB55 и RB59, как и предполагалось, присутствует ОРС *segD*, причем нуклеотидная последовательность ОРС *segD* у фагов RB55 и RB59 полностью идентична фагу Т4, а в случае фага Т6 существует одна нуклеотидная замена в гене *segD*, не приводящая к изменению аминокислотной последовательности белка. Таким образом, всю проанализированную фаговую коллекцию можно разбить, по структуре организации межгенной области 23-24, на две группы, отличающиеся наличием, либо отсутствием ОРС *segD*.

3.2. Очистка белка SegD

С целью изучения свойств фермента SegD был получен очищенный препарат белка. Учитывая токсичность изученных Seg-эндонуклеаз (Kadyrov et al., 1996, Sandegren et al., 2005) для клеток *E. coli*. Было принято решение использовать индуцибельную систему основанную на высокоэффективной и специфичной РНК-полимеразе фага Т7. Чтобы предотвратить возможный токсичный эффект, ген *segD* клонировали под контроль индуцибельного *T7lac*-промотора. Гибридный промотор *T7lac* состоит из двух функциональных элементов: первый – промотор гена *10* фага Т7, который специфически узнается РНК-полимеразой фага Т7, а второй – операторный участок *lac*-оперона *E. coli*, с которым связывается репрессор LacI. Появление в клетках РНК-полимеразы фага Т7 обеспечивалось наличием ее гена под контролем *lac*-промотора в хромосоме штамма BL21(DE3)/pLysE *E. coli*, в котором экспрессировали *segD*. Фоновые количества Т7 РНК-полимеразы в клетках до индукции ингибировал фаговый лизоцим, ген которого находился на плазмиде pLysE (Studier and Moffatt, 1986). Использование данной системы позволило избежать существенного токсического эффекта, связанного с преждевременной транскрипцией гена *segD*.

Использованная в данной работе конструкция pSDET-19mod была создана на основе вектора pET-19mod путем клонирования ОРС *segD*. Особенностью данного вектора является то, что он позволяет нарабатывать в *E. coli* гибридные белки, на N-конец которых добавлена аминокислотная последовательность из шести остатков гистидина и сайт гидролиза TEV протеазы.

Клетки *E. coli* BL21(DE3)/pLysE были трансформированы рекомбинантной плазмидой pSDET-19mod. Не смотря на то, что в ряде похожих работ были отмечены трудности с получением плазмидонесущих клонов, как, например, в случае с SegE, мы особых проблем с получением рекомбинантных клонов при трансформации клеток BL21(DE3)/pLysE плазмидой pSDET-19mod не наблюдали.

В предварительных экспериментах, было показано, что при температуре выше 28 °С практически весь синтезированный белок находится в тельцах включения. Подобная локализация чужеродных и токсичных белков встречается достаточно часто и усложняет получение белков в нативной форме. Поэтому, для получения белка в растворимой форме, индукцию культуры и наработку биомассы проводили при температуре 26 °С.

Результаты экспрессии гена *segD* в клетках *E. coli* BL21(DE3)/pLysE/pSDET-19mod представлены на рис. 14. Контролем служил лизат клеток BL21(DE3)/pLysE/pET-19mod в условиях отсутствия индукции. Электрофоретический анализ показал появление в клетках видимых количеств белка, отсутствующего в контроле, с молекулярной массой около 30 кДа.

Данное значение несколько отличается от его вычисленной массы, равной 27,792 кДа. Аномальная подвижность таких белков уже отмечалась (Williams and Gratzer, 1971) и показана для SegE (Kadyrov et al., 1996). По всей видимости, это связано с тем, что SegD является щелочным белком, вычисленный pI которого равен 10,1.

На первом этапе выделения проводилась хроматография на Ni-NTA агарозе. Чистота препарата, оцененная по данным электрофореза, была около 95% (рис. 14). Присутствие дополнительных белковых примесей можно объяснить наличием у данных белковых примесей сил неспецифического сродства к сорбенту, на котором проводилось выделение.

Вторым этапом было отщепление N-концевого гистидинового довеска. В предварительных экспериментах было установлено, что препарат TEV-протеазы может эффективно гидролизовать субстрат и при 4 °C. Поэтому процесс отщепления His-tag было решено совместить с диализом после первой хроматографии, с этой целью к целевому белковому препарату была добавлена TEV-протеаза. Рекомбинантный препарат TEV-протеазы был получен, как описано в материалах и методах. Как видно из рисунка 14 после отщепления N-концевого довеска электрофоретическая подвижность SegD увеличивается, что связано с уменьшением его молекулярной массы. Так расчетные данные показывают, что протеаза должна отщепить N-концевой пептид массой 1.9 кДа, что согласуется с данными электрофореза (рис.14).

Третьим этапом очистки было удаление TEV-протеазы. Отличительной особенностью данного белка помимо его протеолитической специфичности является присутствие на N-конце гистидинового довеска. Это свойство позволило нам легко провести последующее удаление протеазы из препарата белка. С этой целью была проведена дополнительная хроматография на Ni-NTA агарозе. Целевой белок находился в проскоке т.к. больше не имел N-концевого гистидинового довеска. Кроме того, вторая хроматография позволила избавиться от белковых примесей т.к. они обладали сродством к сорбенту, и основная их часть задержалась на колонке. Использование данной схемы позволило получить электрофоретически чистый препарат фермента (рис. 14).

Таким образом, после обработки протеазой конечный препарат SegD имеет лишь два дополнительных аминокислотных остатка на N-конце оставшиеся от сайта гидролиза TEV-протеазы. Данными аминокислотами являются глицин и гистидин. Разработанная схема очистки позволила использовать преимущество высокоэффективной металл-хелатной хроматографии и получить препарат SegD, максимально соответствующий белку дикого типа (рис. 14).

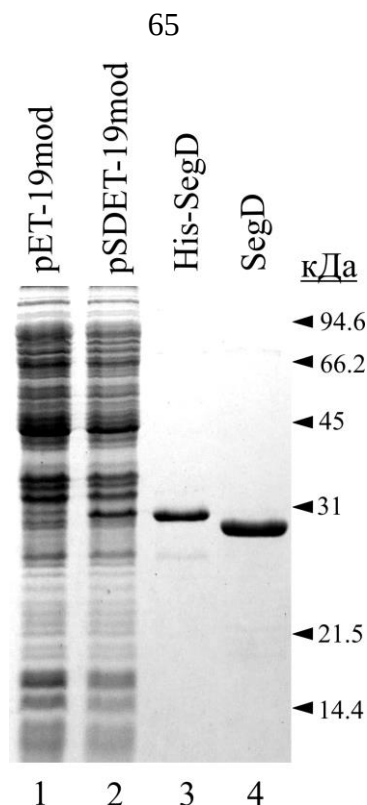


Рис. 14. Продукция в *E. coli* и очистка эндонуклеазы SegD.

Представлены: экстракты клеток BL21(DE3)/pLysE, трансформированные вектором pET-19mod (1) или рекомбинантной плазмидой pSDET-19mod (2), содержащей ОРС *segD*; очищенный препарат эндонуклеазы SegD до (3) и после (4) обработки TEV-протеазой. Справа приведены маркеры молекулярных весов. Белки были разделены в 12% ДСН-ПААГ.

3.3. Исследование биохимических свойств эндонуклеазы SegD

Нуклеаза SegD является небольшим по размеру щелочным белком, состоящим из 224 аминокислотных остатков. Для SegD, как и для других Seg-эндонуклеаз, отмечено расхождение расчетной молекулярной массы и массы, установленной экспериментально на основании измерения электрофоретической подвижности в полиакриломидном геле в присутствии ДДС. Расчетная масса SegD равна 25,618 кДа, в то время как измеренная составляет ~28 кДа. Схожие данные были получены для хоминг-эндонуклеазы SegE. Так, расчетная масса для данного белка составляет 22,966 кДа, а определенная на основании его электрофоретической подвижности – 27,7 кДа (диссертация, Кадыров, 1996). Расхождения в молекулярных массах Seg-эндонуклеаз полученных различными методами связывают с высокой *pI* данных белков, которая и влияет на электрофоретическую подвижность, так вычисленная *pI* для SegD и SegE составляет 10,1 и 9,7, соответственно.

Для изучения основных биохимических свойств данного фермента был проведен анализ эндонуклеазной активности SegD при различных условиях *in vitro*. Выбор субстрата для проведения реакции гидролиза был обусловлен тем что, как правило, сайт гидролиза хоминг-эндонуклеазы располагается в безнуклеазном аллеле. Ранее, предпочтительные сайты гидролиза эндонуклеаз SegB, C, F и G были обнаружены в геноме бактериофага T2. Поэтому в качестве субстрата для анализа использовали плазмиду pUT2LsD, содержащую область генов 23 - 24 фага T2L, линейаризованную по сайту эндонуклеазы рестрикции SspI.

Для эффективного гидролиза большинству эндонуклеаз в качестве кофактора требуются ионы металлов (Cowan, 1998; Cowan, 2002). Поэтому было исследовано влияние двухвалентных (Ca^{2+} , Mg^{2+} , Mn^{2+} , Zn^{2+}) и одновалентных (Na^+ , K^+) катионов на гидролиз SegD (рис. 15А, 5В, 5Г

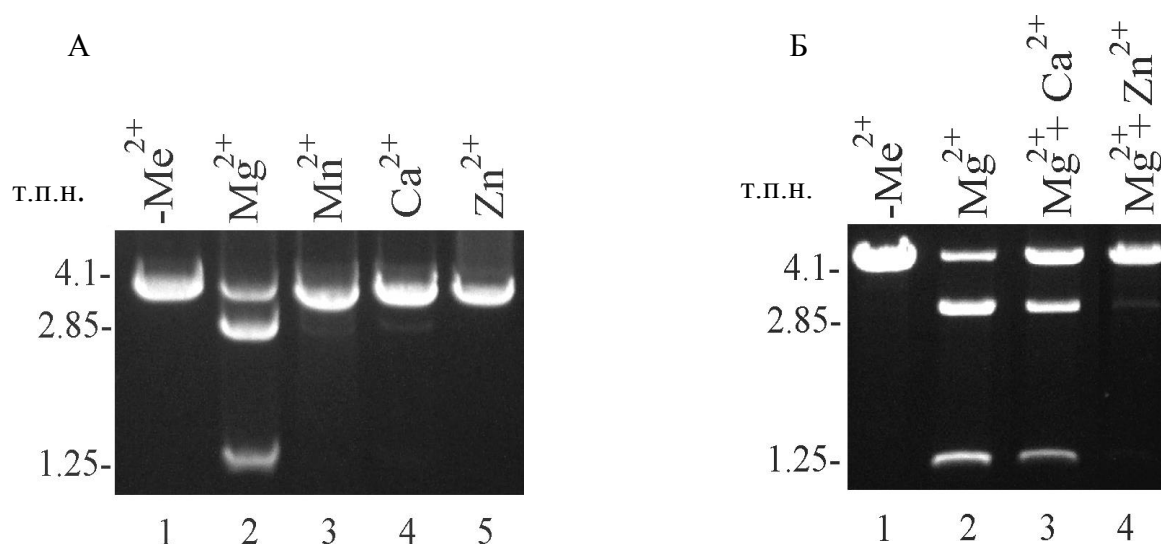


Рис. 15. Влияние двухвалентных катионов на гидролиз SegD. Субстратом выступала плазмида pUT2LsD, содержащая область генов 23-24 фага T2L, линейаризованная по SspI. - Me²⁺ – реакция в отсутствии двухвалентных катионов. В реакционной смеси концентрация катионов составляла 10 мМ. В случае присутствия двух катионов, концентрация каждого из них составляла 5 мМ.

В отсутствии в реакционной смеси двухвалентных катионов (Me²⁺) активность SegD не обнаруживается (рис. 15А, дорожка 1). Сайт-специфическая эндонуклеазная активность проявляется в присутствии катионов Mg²⁺. При замене ионов Mg²⁺ на Mn²⁺, Ca²⁺ или Zn²⁺ SegD не способен эффективно гидролизовать субстрат (рис. 15А, дорожки 3 - 5), что, по-видимому, обусловлено несоответствием размера данных катионов активному центру фермента. Оптимальная концентрация ионов Mg²⁺ находится в диапазоне от 2 до 5 мМ (рис. 19Г).

Кроме того, было проверено влияние на активность SegD катионов Ca^{2+} и Zn^{2+} в присутствии ионов Mg^{2+} . Наличие ионов Ca^{2+} в реакционной смеси снижает эффективность расщепления субстрата, из чего можно сделать вывод о возможной конкуренции за активный центр между катионами Mg^{2+} и Ca^{2+} (рис. 15Б, дорожка 3). В присутствии катионов Zn^{2+} в реакционной смеси наблюдалось полное подавление Mg^{2+} -зависимой эндонуклеазной активности SegD (рис. 15Б, дорожка 4).

Для SegD обнаружено полное ингибирование Mg^{2+} -зависимого расщепления ДНК ионами Zn^{2+} , что, по-видимому, связано с конкуренцией ионов Zn^{2+} и Mg^{2+} за активный центр фермента. Для ряда эндонуклеаз было показано, что каталитически неактивные ионы металлов в низких концентрациях могут стимулировать прохождение реакции гидролиза. Подобный эффект расценивается как показатель прохождения реакции гидролиза по пути катализа “два-металла” (Vipond et al., 1995, диссертация Брок-Волчанская В.С., 2008). Поскольку, в случае SegD, как и в случае SegB, катионы Zn^{2+} и Ca^{2+} ингибировали протекание гидролиза, опосредованное катионами Mg^{2+} , это может косвенно указывать на то, что эти белки используют механизм расщепления “один-металл” (диссертация Брок-Волчанская В.С., 2008).

Для каталитического домена белка UvrC, участвующего в эксцизионной репарации нуклеотидов у эубактерий, также был предложен механизм расщепления “один-металл” (Truglio et al., 2005). У данного белка, как и у SegD, обнаружен вырожденный GIY-YIG-мотив, аминокислоты которого, как показано, важны для катализа. Сходство N-концевых фрагментов UvrC и SegD может являться еще одним косвенным доказательством, что SegD осуществляет катализ по механизму расщепления “один-металл”.

В случае SegD было показано, что продолжительное инкубирование субстратной ДНК с ферментом, или увеличение концентрации эндонуклеазы SegD в реакционной смеси, может приводить к появлению дополнительных участков расщепления. (рис. 16). Эта особенность так же характерна и для родственных эндонуклеаз SegA, SegB и SegE (Sharma and Hinton, 1994; Kadyrov et al., 1994, Brok-Volchanskaya et al., 2008). По-видимому, в отличие от эндонуклеаз рестрикции, Seg-эндонуклеазы толерантны к небольшим заменам нуклеотидных остатков в узнаваемой последовательности ДНК.

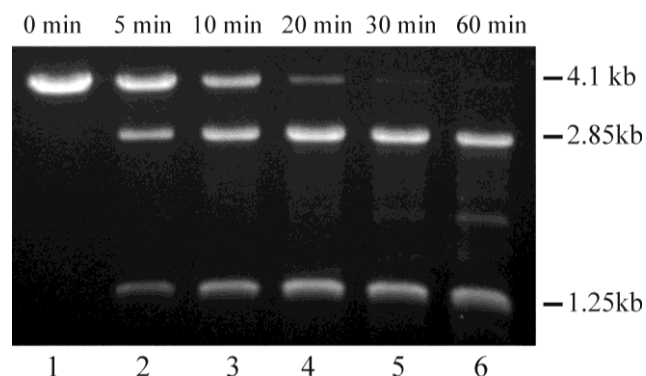


Рис. 16. Кинетика гидролиза плазмиды pUT2LsD эндонуклеазой SegD. Количество ДНК в реакции составило 0,2 пмоль, SegD – 30 фмоль. В качестве субстрата была использована линейаризованная по SspI pUT2LsD. Реакция гидролиза проводилась в присутствии 1mM Mg^{2+} .

Показано, что для проявления ферментативной активности эндонуклеазе SegD не требуется наличие моновалентных катионов в реакционной смеси, тем не менее, небольшие концентрации K^+ , до 50 mM, стимулируют работу фермента. Увеличение концентрации NaCl и, в меньшей степени, KCl приводит к снижению общей активности фермента (рис. 19B).

Было проанализировано влияние на активность SegD таких факторов как pH и температура. Зависимость активности SegD от pH оценивали в разных буферных системах в диапазоне от 5,5 до 10,5. Установлено, что эндонуклеаза SegD способна гидролизовать субстратную ДНК в диапазоне pH от 6,5 до 9,5 (рис. 19A). При pH 5,5 гидролиз практически отсутствовал. Наибольший процент гидролизованного субстрата наблюдался при pH 7,0 – 8,0.

На рис. 19Б показано влияние температуры на активность SegD. Наибольшее количество гидролизованного субстрата наблюдалась при температуре 30 °C. Данный температурный оптимум совпадает с оптимумами ближайших родственных белков, таких как SegB и SegE.

При проведении экспериментов было замечено, что при разведении исходного препарата белка, эндонуклеаза SegD быстро теряет активность. Поэтому было проверено влияние некоторых стабилизаторов, таких как БСА и Тритон X-100, на стабильность SegD. Оказалось, что присутствие 0,1 мг/мл БСА и Тритон X-100 в буфере для разведения существенно стабилизирует фермент (рис. 17). Кроме того, присутствие в реакционной смеси 0,02% Тритон X-100 и БСА также приводило к увеличению количества гидролизованного субстрата (рис. 18).

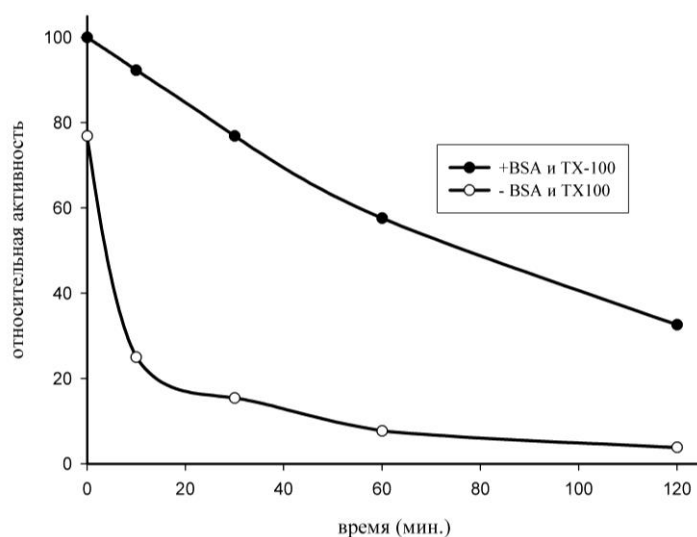


Рис. 17. Влияние БСА и TX-100 на стабильность SegD. Черными кружками обозначено присутствие, белыми – отсутствие БСА и TX-100 в буфере для хранения. Количество БСА и TX-100 в буфере для хранения составляло 0,1 мг/мл и 0,02%, соответственно.

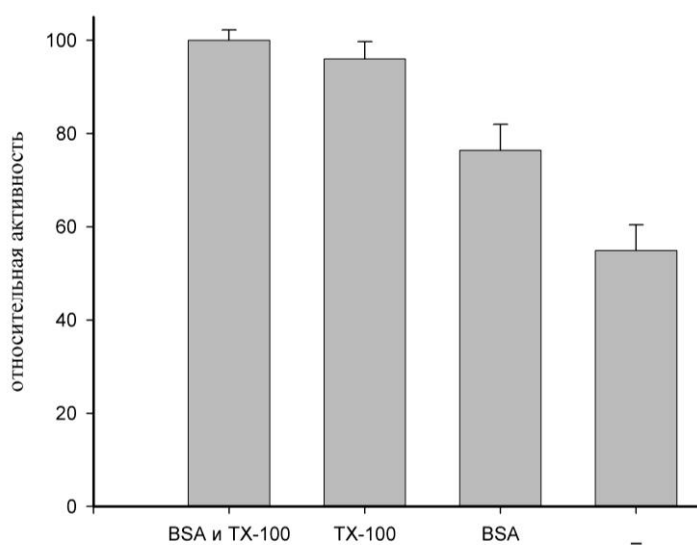


Рис. 18. Влияние БСА и Тритон X-100 на активность SegD. Количество БСА и Тритон X-100 в реакционной смеси составляло 0,1 мг/мл и 0,02%, соответственно.

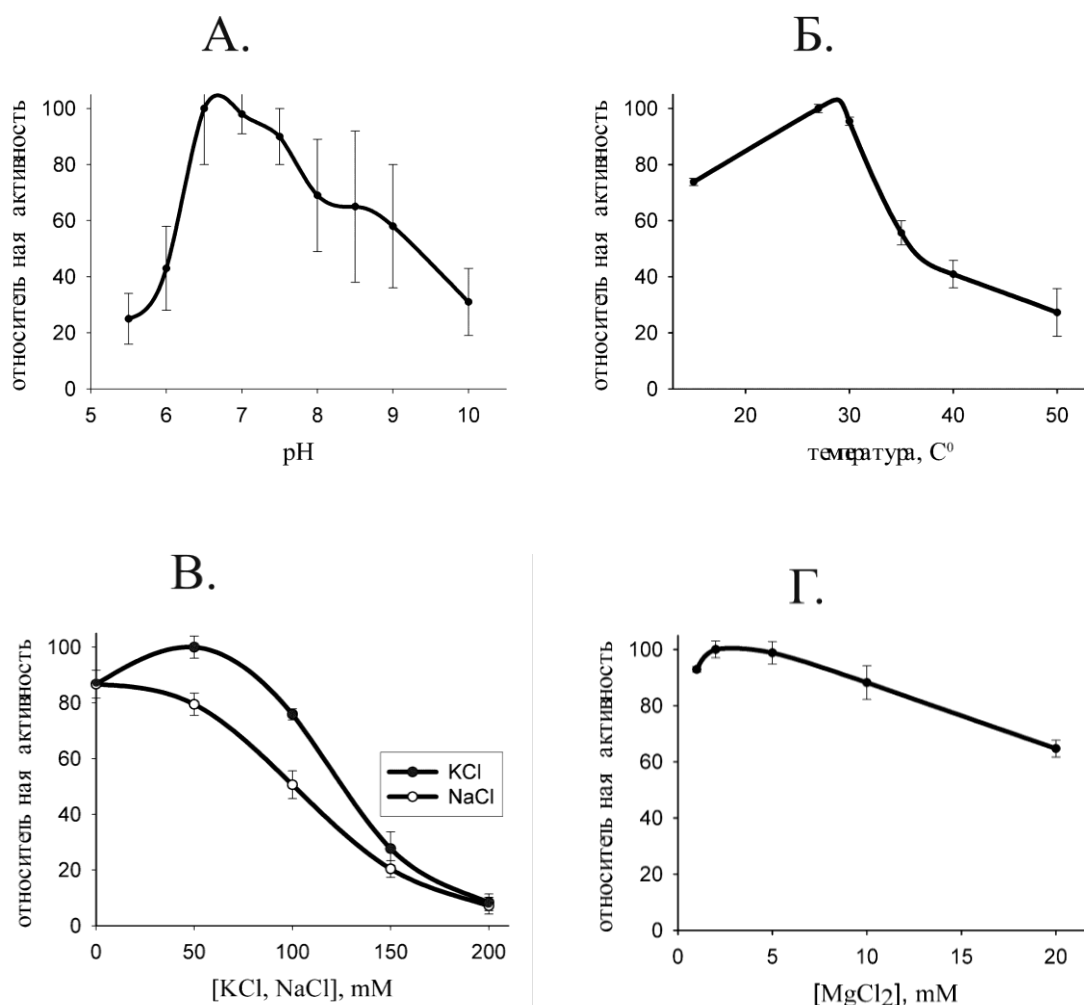


Рис. 19. Анализ биохимических свойств эндонуклеазы SegD. Влияние на активность эндонуклеазы SegD: pH (А), температуры (Б), моновалентных катионов Na⁺ и K⁺ (В), катионов Mg²⁺ (Г).

Была исследована структура 5'- и 3'-концов, образующихся после разрыва фосфодиэфирной связи ДНК эндонуклеазой SegD. Для этого продукты гидролиза инкубировали с ДНК-лигазой фага Т4 (рис. 20). Известно, что необходимым условием для проявления активности этого фермента является наличие 5'-фосфатной группы и 3'-гидрооксильной на концах сшиваемых молекул ДНК. В результате реакции продуктов гидролиза с ДНК-лигазой Т4, наблюдалось появление фрагмента большего размера (рис. 20, проба 3), и поэтому был сделан вывод, что эндонуклеаза SegD расщепляет ДНК-субстрат с образованием 5'-Р и 3'-ОН концов, как и подавляющее большинство дезоксирибонуклеаз.

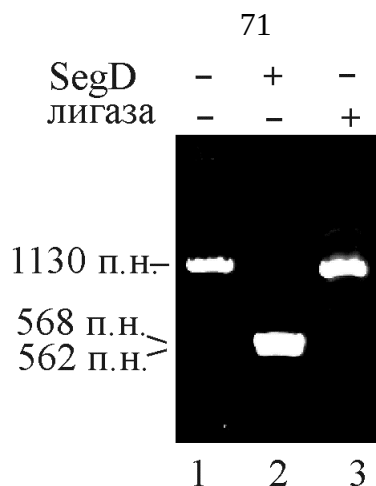


Рис. 20. Лигирование продуктов гидролиза эндонуклеазы SegD. ПЦР-фрагмент фага T2L (дорожка 1) был обработан эндонуклеазой SegD (дорожка 2). Фермент инактивировали прогреванием при 65 °C в течение 20 мин. Затем, продукты гидролиза инкубировали с 1 ед. T4 ДНК-лигазы при 10 °C в течение 16 ч (дорожка 3). Слева указаны размеры фрагментов ДНК.

3.4. Картирование сайта гидролиза эндонуклеазы SegD

Как установлено ранее хоминг-эндонуклеазы, SegB, C, E, F, и G инициируют мобильность собственных ОРС. Данная форма конверсии генов требует наличия разрыва в последовательности ДНК родственного аллеля, не кодирующего нуклеазу. По аналогии с другими Seg-эндонуклеазами было предположено, что в геноме фагов, у которых не был обнаружен ген *segD*, будет присутствовать сайт расщепления для данной эндонуклеазы. В связи с этим, в области генов 23 – 24 фага T2L был проведен поиск участков расщепления для эндонуклеазы SegD. Фрагмент ДНК был генов 23 – 24 фага T2L предварительно клонирован в плазмидный вектор pUC18 по сайтам EcoRI и HindIII, как это описано в главе материалы и методы исследования. В ходе рестрикционного картирования было установлено, что сайт гидролиз эндонуклеазы SegD находится в области 3'-конца гена 23 (рис. 21).

Помимо этого был проведен поиск сайтов гидролиза эндонуклеазы SegD в имеющейся коллекции фагов. В качестве субстратой ДНК выступали ПЦР-фрагменты области генов 23 - 24 фагов, структура которых была определена ранее. Анализ показал, что эндонуклеаза SegD гидролизует ДНК фагов, как не содержащих, так и содержащих ген *segD*, с близкой эффективностью (данные не приведены).

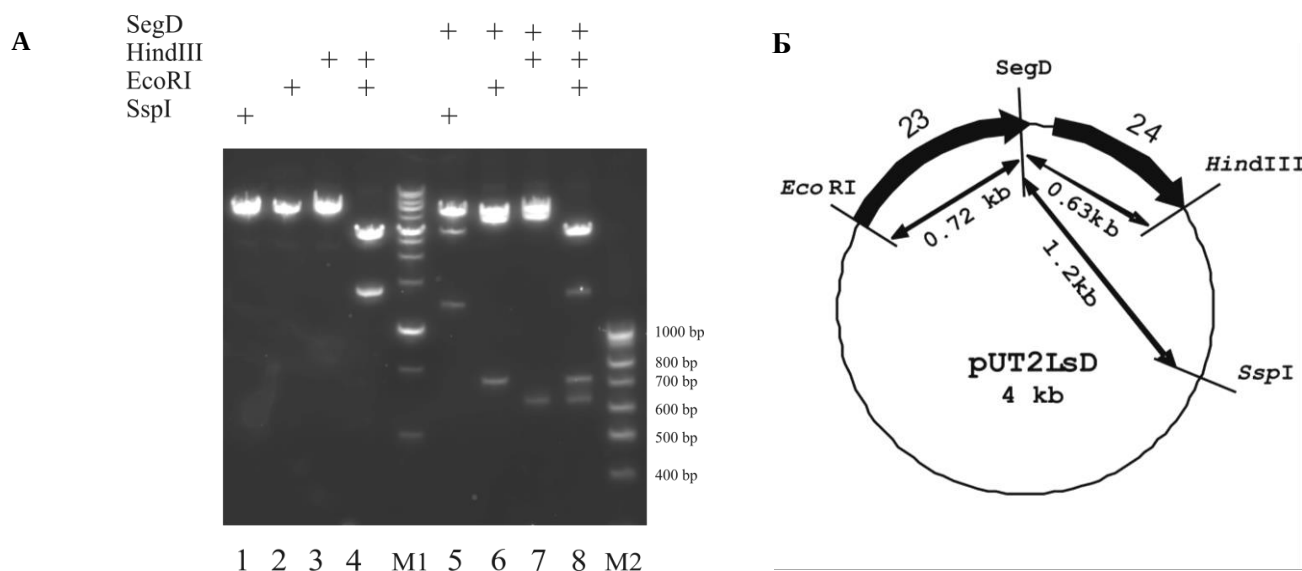


Рис. 21. Картирование сайта гидролиза. Плазмида pUT2LsD содержит фрагмент ДНК 1.35 т.п.н. области генов 23 – 24 бактериофага T2L. Данный фрагмент был клонирован по сайтам *EcoRI* – *HindIII* в плазмиду pUC18. (**А**) Плазмида pUT2LsD была линейаризованна эндонуклеазами рестрикции *SspI*, *EcoRI* или *HindIII* (дорожки 1 – 3, соответственно) далее соответствующие линейаризованные фрагменты плазмидной ДНК были гидролизованны SegD (дорожки 5 – 7). На дорожке 4 представлен гидролиз плазмиды одновременно двумя эндонуклеазами *EcoRI* and *HindIII*. На дорожке 8 представлен гидролиз образца, приведенного на дорожке 4 эндонуклеазой SegD. Сайт гидролиза эндонуклеазы SegD был локализован на фрагменте ДНК T2L (1.35 т.п.н., дорожка 8). Установлено сайт гидролиза SegD находится на расстоянии 0.72 т.п.н. от сайта *EcoRI* (дорожка 6) и на расстоянии 0.63 т.п.н от сайта *HindIII* (дорожка 7). Сайт гидролиза был картирован на 3'-конце гена 23. Трек M1 - 1 kb DNA ladder, M12 SibEnzyme (10, 8, 6, 5, 4, 3, 2.5, 2, 1.5, 1, 0.75, 0.5, 0.25 т.п.н.), Трек M2 - 100 bp DNA ladder, M16 SibEnzyme (1000, 900, 800, 700, 600, 500, 400, 300, 200, 100 т.п.н.). Продукты гидролиза разделялись в 1% агарозном геле. На рисунке (**Б**) приведена карта pUT2LsD на которой обобщены данные рестрикционного картирования приведенного в части (**А**).

3.5. Сайт гидролиза эндонуклеазы SegD находится на 3'-конце гена 23

Для того чтобы охарактеризовать последовательность ДНК, узнаваемую SegD, были установлены точки гидролиза и границы сайта узнавания в области гена 23. При определении границ сайта узнавания SegD был использован метод, предложенный Венцлау с соавт. (Wenzlau et al., 1989). Для этого, продукты четырех секвенирующих реакций, полученные в присутствии

соответствующих терминаторов с праймерами, комплементарными последовательности выше и ниже сайта расщепления, гидролизовали SegD. Продукты секвенирующих реакций, которые не содержат сайт узнавания, не расщепляются эндонуклеазой, что и определяет минимальную последовательность, необходимую для расщепления. Так же были локализованы точки гидролиза. (рис. 22).

Установлено, что SegD вносит двухцепочечный разрыв в последовательность ДНК, расположенную в конце 23 гена бактериофага Т4 (рис. 12). В результате гидролиза ДНК SegD образуются 3'-выступающие концы длиной три нуклеотидных остатка. Минимальная последовательность ДНК, узнаваемая эндонуклеазой SegD, составляет 16 нуклеотидных остатков. Разрыв, вносимый в молекулу ДНК SegD, смещен относительно центра узнаваемой белком последовательности. Кроме того, сайт гидролиза ДНК SegD не обладает симметрией второго порядка, как в случае эндонуклеаз рестрикции второго типа.

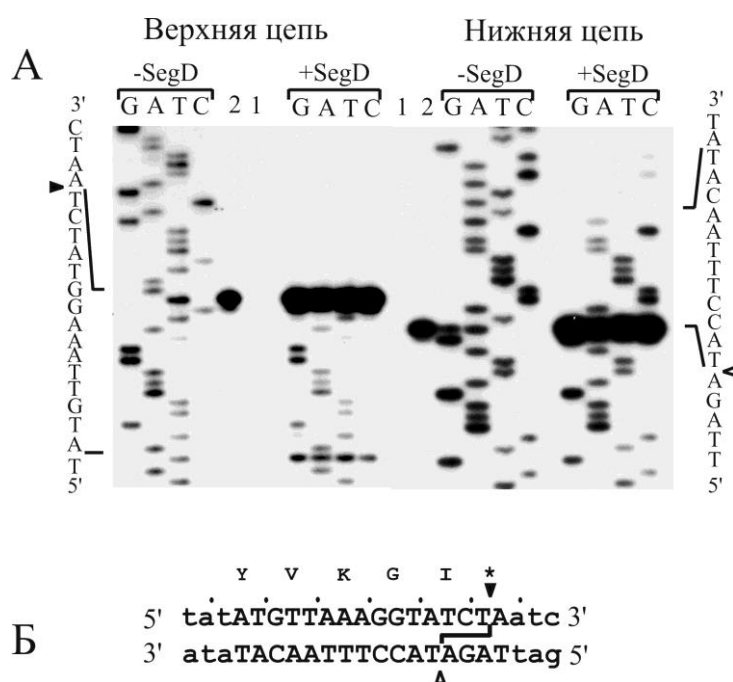


Рис. 22. Определение точек гидролиза и границ сайта узнавания SegD.

(А) Представлена радиоавтограмма участков секвенирующего геля для верхней и нижней цепей ДНК: продукты реакции элонгации меченного праймера до (1) и после (2) инкубации с SegD; G, A, T и C – продукты секвенирующих реакций, полученные в присутствии соответствующих терминаторов, до и после инкубации с SegD, как указано. Слева и справа приведены участки последовательностей верхней и нижней цепей ДНК, соответственно. Положение точек разрывов ДНК в верхней и нижней цепях указано черными и белыми стрелками, соответственно. Результаты анализа обобщены в части (Б), в которой минимальная последовательность ДНК, необходимая для гидролиза SegD, выделена заглавными буквами.

У всех проанализированных бактериофагов - T4, T6, RB55, RB59, T2L, RB1, RB2, RB4, RB6, RB7, RB8, RB9, RB10 и RB15- нуклеотидные последовательности сайтов гидролиза оказались идентичны.

		SegD ↓									
segD⁺ phages	T4	ttagacg	tgtat	ATGTTAAAGGTATCTA	aatct						
	T4T	ttagacg	tgtat	ATGTTAAAGGTATCTA	aatct						
	T6	ttagacg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB55	ttagacg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB59	ttagacg	tgtat	ATGTTAAAGGTATCTA	aatct						
	ECML-134	ttagacg	tgtat	ATGTTAAAGGTATCTA	aatct						
segD⁻ phages	T2L	ttagacg	tgtat	ATGTTAAAGGTATCTA	aatct						
	ime09	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	vB EcoM	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	ACG-C40	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	phiD1	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB1	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB2	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB3	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB4	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB6	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB7	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB8	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB9	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB10	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB14	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB15	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB32	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	RB51	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	Shf12	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	AR1	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	wV7	tccgccg	tgtat	ATGTTAAAGGTATCTA	aatct						
	JS98	tccg	tcgcgtat	ATGTTAAAGGTATCTA	aat--						
	Bp7	tccg	tcgcgtat	ATGTTAAAGGTATCTA	aagat						

Рис. 23. Сайт SegD идентичен у 27 T4-родственных бактериофагов. Выравнивание последовательностей ДНК бактериофагов в области 3'-конца гена 23. Рамкой выделена область сайта эндонуклеазы SegD. Наличие в геноме OPC *segD* указано в левой части рисунка. Номера доступа последовательностей в базах данных EMBL или GenBank приведены в таблице 1.

Таблица 1. Номера доступа последовательностей в базах данных EMBL или GenBank.

Enterobacteria phage T4	gi_215870_gb_K01765.1_PT4G22G23:1919-2152
Enterobacteria phage T4T	gi_299780356_gb_HM137666.1_:106267-106500
Enterobacteria phage T6	GU187340
Enterobacteria phage RB55	GU187351
Enterobacteria phage RB59	GU187352
Escherichia phage ECML-134	gi_397134020_gb_JX128259.1_:103906-104139
Enterobacteria phage T2L	gi_270272118_gb_GU187339.1_:1-331
Enterobacteria phage ime09	gi_339791204_gb_JN202312.1_:103011-103341
Enterobacteria phage vB EcoM ACG-C40	gi_383396143_gb_JN986846.1_:103911-104241
Yersinia phage phiD1	gi_398313544_emb_HE956711.1_:103854-104184
Enterobacteria phage RB1	GU187341
Enterobacteria phage RB2	GU187342
Enterobacteria phage RB3	GU187343
Enterobacteria phage RB4	GU187344
Enterobacteria phage RB6	GU187345
Enterobacteria phage RB7	GU187346
Enterobacteria phage RB8	GU187347
Enterobacteria phage RB9	GU187348
Enterobacteria phage RB10	GU187349
Enterobacteria phage RB14	gi_227438331_gb_FJ839692.1_:102400-102729
Enterobacteria phage RB15	GU187350
Enterobacteria phage RB32	gi_115343815_gb_DQ904452.1_:103828-104158
Enterobacteria phage RB51	gi_227462430_gb_FJ839693.1_:105756-106086
Shigella phage Shfl2	gi_327397459_gb_HM035025.1_:101952-102282
Enterobacteria phage AR1	gi_291290213_dbj_AP011113.1_:103528-103858
Escherichia phage wV7	gi_343177337_gb_HM997020.1_:103313-103643
Enterobacteria phage JS98	gi_160213687_gb_EF469154.1_:105566-105788
Enterobacteria phage Bp7	HQ829472

Дополнительно, был проведен анализ области гена 23 у 15-ти T4-родственных бактериофагов, последовательности которых представлены в базе данных GenBank. Два фага, T4T и ECML-134, содержат ген *segD*. У остальных 13 фагов – ime09, vB EcoM, ACG-C40, phiD1, RB3, RB14, RB32, RB51, Shfl2, AR1, wV7, JS98 и Bp7 – ген *segD* отсутствует. Установлено, что все фаги содержат сайт SegD, последовательность которого идентична установленной ранее для фага T2L (рис. 23). В случае эндонуклеазы SegD имеется уникальная ситуация, поскольку, как в собственном геноме, так и в родственных геномах, не содержащих нуклеазный ген, имеются идентичные сайты гидролиза.

3.6. Исследование влияния модификации ДНК Т-четных бактериофагов на эффективность и специфичность гидролиза SegD.

Ситуация, когда в геноме фага Т4 одновременно присутствует ген хоминг-эндонуклеазы SegD и ее сайт гидролиза, является необычной. Можно предположить, что отсутствие разрыва в сайте гидролиза в собственном геноме фага Т4, связано с модификацией ДНК фага. ДНК Т-четных бактериофагов имеет необычную структуру: остаток цитозина метилируется и представляет собой 5-гидроксиметилцитозин (ОМЦ); кроме того, все гидроксильные группы 5-гидроксиметилцитозиновых оснований модифицированы в результате ковалентного связывания с глюкозой (Lehman and Pratt, 1960). При этом 70% связей - α -гликозидные, а 30% - β -гликозидные (см. гл. 1). Ввиду того, что во всех ранее проводившихся экспериментах использовались плазмиды, либо синтетические фрагменты ДНК, содержащие немодифицированный цитозин, интерес представляло установление влияния природных модификаций ДНК фага на активность эндонуклеазы SegD.

Для этого, наряду с фагами дикого типа, были выбраны для анализа два мутантных штамма фага Т4, T4290 и T4alc7. Фаг T4290 дефектен по генам α - и β -гликозилтрансфераз, осуществляющих перенос глюкозы на ОМЦ. Бактериофаг T4alc7 дефектен по гену гидроксиметилазы, в результате чего он содержит цитозиновую ДНК. Для детального анализа специфичности SegD на модифицированной ДНК фагов, геномная ДНК была гидролизована эндонуклеазой рестрикции *SmiI* с последующей обработкой SegD. Продукты гидролиза разделяли в 1% агарозном геле при помощи инверсионного пульс-электрофореза (см. материалы и методы). Эндонуклеаза *SmiI* нечувствительна к природной модификации ДНК Т-четных фагов, поскольку не имеет остатков цитозина в последовательности сайта узнавания (5'-АТТТАААТ-3'). Восьминуклеотидный сайт *SmiI* обеспечивает получение сравнительно небольшого количества фрагментов ДНК, что облегчает рестрикционный анализ. На основании информации о полной нуклеотидной последовательности генома Т4 (NCBI, accession number 000866) была построена рестрикционная карта гидролиза ДНК фага Т4 *SmiI* (табл. 2).

Таблица 2. Теоретически рассчитанные фрагменты ДНК, образующиеся при обработке геномной ДНК фага Т4 рестриктазой *SmiI* и хоминг-эндонуклеазой SegD.

№ фрагмента	Размер, п. о.	
	T4B/ <i>SmiI</i>	T4B/ <i>SmiI</i> /SegD
1	18371	18371
2	14382*	14382
3	13896	13896
4	12390	12390
5	10767	11674*

6	10298	10767
7	10044	10298
8	9620	10044
9	9150	9620
10	8746	9150
11	8728	8764
12	7651	8728
13	7313	7313
14	4463	4463
15	4124	4124
16	3368	3368
17	2768	2768
18	2496	2708*
19	2346	2496
20	1994	2346
21	1318	1994
22	1020	1318
23	1012	1020
24	894	1012
25	771	894
26	412	771
27	297	412
28	264	297
29		264

* выделен исходный фрагмент, содержащий область генов 23-24, и фрагменты, образующиеся после обработки SegD.

Ввиду того, что полная последовательность генома T2L не определена, не было возможности построить аналогичную карту для этого фага.

Из ранее проведенных исследований известно, что предпочтительный сайт гидролиза SegD у фага T4 находится на 3'-конце 23 гена, который расположен на *SmiI*-фрагменте размером 14382 п.н. Согласно этому, фрагмент ДНК размером 14382 п.н. должен расщепляться SegD на 2 фрагмента – 11674 и 2708 п.н. Фрагменты такого размера действительно появляются после гидролиза SegD, что видно на электрофореграмме (рис.24А). Мутантные фаги T4290 и T4alc7, содержащие негликозилированную и цитозиную ДНК соответственно, имеют немного отличную от фага T4 картину гидролиза *SmiI*, видимо, вследствие мутационных изменений генома. Однако можно отметить, что при обработке эндонуклеазой SegD ДНК мутантных фагов образуются фрагменты такого же размера, как и фага дикого типа.

Фаг T2L, как и T4, имеет предпочтительный сайт гидролиза SegD на 3'-конце 23 гена. На электрофореграмме видно, что при действии SegD на ДНК фага T4 исчезает полоса размером ~14,4 т.п.н. и появляется полоса размером ~11,7 т.п.н. второй фрагмент, ожидаемый размер которого 2,7 т.п.н., на геле не виден из-за низкой интенсивности свечения и наложения на близкие по размеру фрагменты ДНК (рис 24 А). В случае бактериофага T2 наблюдается исчезновение фрагмента ~13,5 т.п.н. и появление полос ~11,5 т.п.н. и ~2 т.п.н.

Для подтверждения того, что наблюдаемое расщепление ДНК Т-четных фагов эндонуклеазой SegD действительно происходит в предполагаемом участке гена 23, был проведен гибридизационный анализ продуктов гидролиза. После проведения электрофореза и переноса фрагментов ДНК исследуемых бактериофагов из агарозного геля на нейлоновый фильтр была проведена их гибридизация с радиоактивно меченым ДНК-зондом к области генов 23-24 бактериофага Т4. Зонд подбирался таким образом, чтобы он перекрывал область как выше сайта гидролиза SegD, так и ниже, с целью детекции как изначального фрагмента, так и обоих продуктов гидролиза. В качестве зонда использовался ПЦР-продукт фага Т4, отмеченный на рис. 23В. В случае бактериофага Т2L, Т4 и его мутантов, наблюдалось появление на радиоавтограмме полос ожидаемых размеров, что подтвердило результаты электрофоретического анализа (рис. 24Б). Т.к. выбранный нами зонд окрашивает оба продукта реакции гидролиза это свидетельствует о том что SegD вносит разрыв в область генов 23-24.

Оказалось, что ДНК всех исследуемых фагов расщепляются эндонуклеазой SegD со сходной эффективностью и специфичностью (рис. 24). Ранее установлено, что гликозильные остатки, ковалентно связанные с гидроксиметилцитозином, ориентированы в большой желобок (Revel et al., 1983). Поскольку модификации ДНК Т-четных фагов (гликозилирование и гидроксиметилирование) не влияли на эффективность и специфичность гидролиза, можно предположить, что SegD узнает ДНК преимущественно по малому желобку. Подобные результаты были получены и для ближайших родственных эндонуклеаз, таких как SegB и интрон кодируемая хоминг-эндонуклеаза I-TevI, в случае которых как немодифицированная, так и гликозилированная ОМЦ ДНК фагов расщеплялась эффективно и с высокой специфичностью (Brok-Volchanskaya et al., 2008, Bell-Pedersen et al., 1991).

Таким образом, с помощью Саузерн-анализа было показано, что эндонуклеаза SegD вносит разрыв в ДНК фага Т2L, Т4 и его производных в 3'-конец гена 23 (рис. 24В). При этом установлено, что исследуемые модификации ДНК существенно не влияют на узнавание и каталитическую активность SegD.

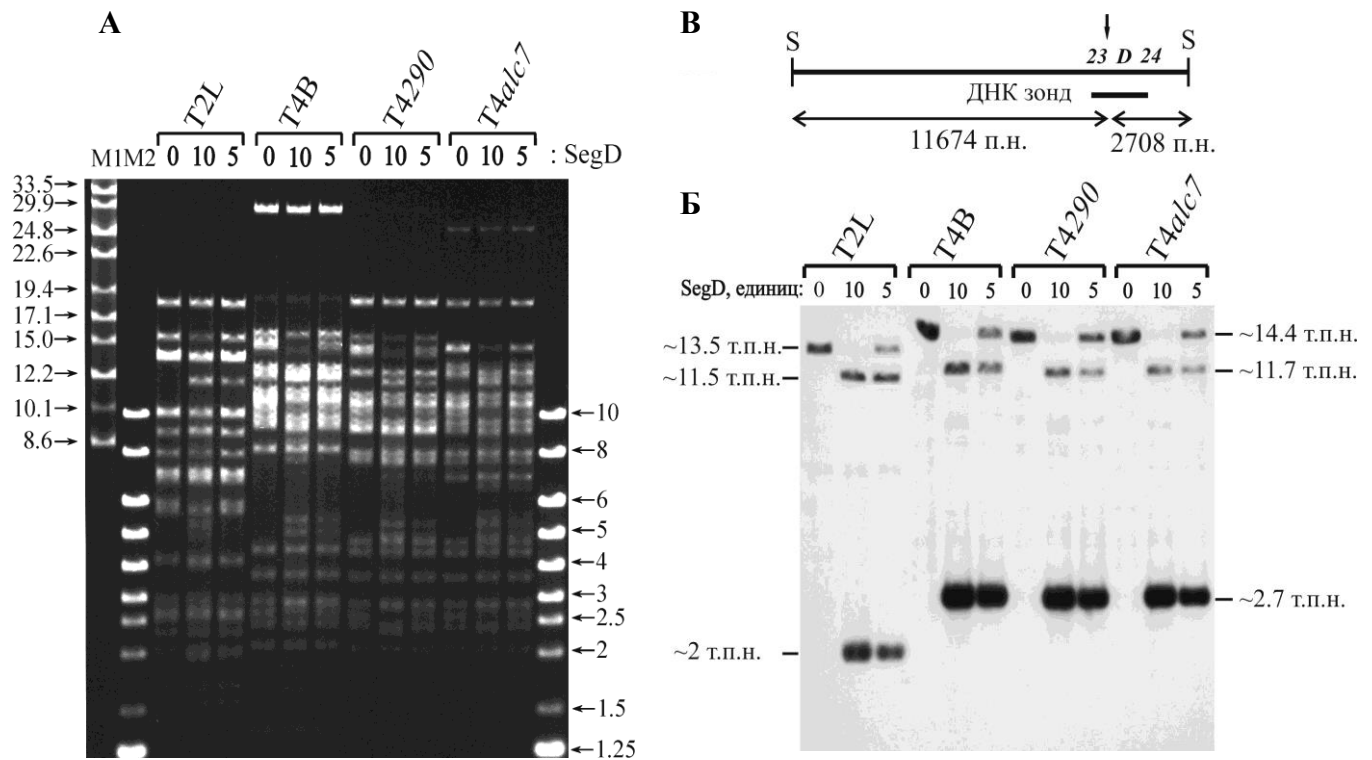


Рис. 24. Влияние природных модификаций ДНК Т-четных фагов на эффективность гидролиза эндонуклеазой SegD. (А) ДНК фагов T2L и T4B, T4290 T4alc7 была гидролизована последовательно SmlI и 5 и 10 ед. эндонуклеазы SegD при стандартных условиях с последующим разделением продуктов гидролиза ДНК в 1% агарозном геле с помощью инверсионного пульс-электрофореза, как описано в главе “Материалы и Методы”. В качестве маркеров длин фрагментов использовали Bio-Rad 8-48 kbp и Q-Step4 (YorBio) 10-0.1 kbp. **(Б)** Саузерн-блот анализ продуктов гидролиза ДНК фагов T2L и T4B, T4290 T4alc7 SmlI и SegD приведенных на рис А. Количество единиц активности SegD в реакции указано над дорожками. **(В)** Схема фрагмента SmlI (S) фага T4, содержащего область генов 23 – 24. Вертикальной стрелкой указан сайт гидролиза эндонуклеазы SegD. Двусторонними стрелками указаны продукты гидролиза SegD, а также приведены их размеры. Горизонтальной линией отмечено положение ДНК-зонда, использованного при гибридизации. Положение генов 23 и 24 указано над схемой.

3.7. Ген *segD* экспрессируется в клетках *E. coli*, инфицированных фагом T4

Сохранение в геноме T4 сайта SegD может быть обусловлено тем, что ген *segD* не экспрессируется в ходе цикла развития бактериофага. Для проверки того, что ген *segD*

экспрессируется в ходе развития бактериофага Т4 и кодирует функционально активную хоминг-эндонуклеазу семейства GIY-YIG, экстракты клеток *E. coli*, инфицированных фагом Т4, были протестированы на наличие нуклеазной активности, характерной для SegD. В качестве т.п.н. тельного контроля был получен фаг Т4ΔsegD с делецией в области 5'-конца гена *segD* (с 3-208 н.о.) приводящей к сдвигу рамки считывания данной ОРС. Схема получения мутантного фага подробно описана в главе материалы и методы исследования. В качестве субстрата для определения эндонуклеазной активности в экстрактах клеток была использована линейаризованная плазмида рUT2LsD содержащая область генов 23-24 фага Т2L, включающая сайт гидролиза SegD. Были получены экстракты клеток *E.coli* BL21, инфицированных фагом Т4 или Т4ΔsegD. Результаты тестирования эндонуклеазной активности в экстрактах клеток *E. coli*, инфицированных каждым из фагов представлены на рис. 25.

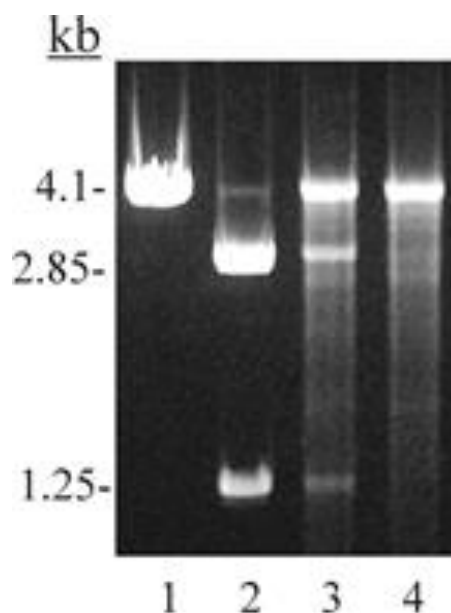


Рис. 25. Активность эндонуклеазы SegD в экстрактах клеток *E.coli*, инфицированных фагом Т4. Линейаризованную плазмиду рUT2LsD (1), содержащую сайт гидролиза эндонуклеазы SegD, инкубировали с экстрактом клеток *E.coli* BL21, инфицированных фагом Т4 (3), либо Т4ΔsegD (4). Продукты реакции разделяли электрофоретически в 1% агарозе. В качестве маркеров использовали фрагменты ДНК, образуемые при гидролизе данной плазмиды очищенным препаратом SegD (2).

Из рисунка видно, что инкубация ДНК-субстрата с экстрактом клеток, инфицированных фагом Т4, приводит к образованию дискретных фрагментов ДНК (дорожка 3), которые совпадают по размеру с продуктами гидролиза ДНК препаратом эндонуклеазы SegD *in vitro*. В

то время как, при инкубации ДНК-субстрата с экстрактом клеток, инфицированных фагом T4 Δ *segD* (дорожка 4), образование этих фрагментов не наблюдается. Результаты анализа свидетельствуют, что сайт-специфическая эндонуклеазная активность, наблюдаемая в экстрактах T4-инфицированных клеток, соответствует активности эндонуклеазы SegD. Эти данные однозначно свидетельствуют о том, что ген *segD* экспрессируется в ходе инфекционного цикла развития T4.

3.8. Анализ наследования гена *segD* при скрещивании *segD*⁺ и *segD*⁻ фагов

Хоминг, как частый случай генной конверсии, обусловлен внесением предпочтительного разрыва в одну из аллелей. Последовательность обнаруженного сайта SegD на 3'-конце гена 23 идентична у всех проанализированных *segD*⁺ и *segD*⁻ фагов. Это позволило предположить, что в скрещиваниях таких фагов не будет происходить хоминг гена *segD*, поскольку эндонуклеаза SegD будет вносить разрывы в обе родительские аллели с равной вероятностью. Для проверки этого предположения было проанализировано наследование ОРС *segD* в потомстве скрещиваний фагов T4B (*segD*⁺) и T2L (*segD*⁻). Ранее Расселлом и Хаски для смешанных инфекций этих фагов было описано так называемое явление частичного исключения (Russell et al., 1974). Данное явление заключается в том, что генетические маркеры T4 предпочтительно наследуются потомством. При этом, исключение маркеров фага T2 в потомстве происходит неравномерно, варьируя от ~ 70% до 100%. В связи с этим можно было ожидать, что соотношение фагов *segD*⁺/*segD*⁻ будет увеличено в потомстве. Для учета явления частичного исключения в анализ был введен дополнительный генетический маркер – амбер мутация в гене 14 (бактериофаг B20). Ген 14 находится на расстоянии ~ 11,5 т.п.н. от ОРС *segD*, и является независимо наследуемым маркером. Если при ко-инфекции фаг B20 будет ограничивать развитие реципиентного фага (T2L), то доля потомства, содержащая амбер-мутацию, будет изменяться. Так же в качестве контроля была проанализирована мобильность генов *segB* и *segC*, для которых ранее была показан хоминг в скрещиваниях T4 и T2. Анализ показал, что при скрещивании фагов B20 и T2L, доля потомства, дающая положительный сигнал на гибридизацию с зондом к ОРС *segD*, составляла около 73%. При этом доля потомства, с амбер-мутацией приближалась к 65 %. В случае генов *segB* и *segC* величина процента наследования генов потомством скрещиваний приближалась к 100% (данные не приведены). Наблюдаемое увеличение в потомстве доли *segD* относительно амбер-мутации в гене 14 не позволяет сделать однозначный вывод, обусловлено ли это хомингом гена *segD*, который происходит с меньшей эффективностью, чем хоминг *segB* и *segC*, либо это связано со сложным характером наследования маркеров в потомстве скрещиваний T4 и T2L.

В связи с этим, для анализа способности SegD инициировать хоминг была сконструирована экспериментальная модель на основе фага T4. В данной модели в качестве одного из родителей был использован бактериофаг D2 (*segD*⁻) - производный T4, у которого область генов 23 – 24 замещена на гомологичную область фага T2L. В качестве второго родителя был использован фаг B20*dm* (*segD*⁺). Данный фаг является производным фага B20, в котором в последовательность сайта SegD внесены замены 3 н.о., существенно снижающие эффективность его гидролиза *in vitro* (рис. 26).

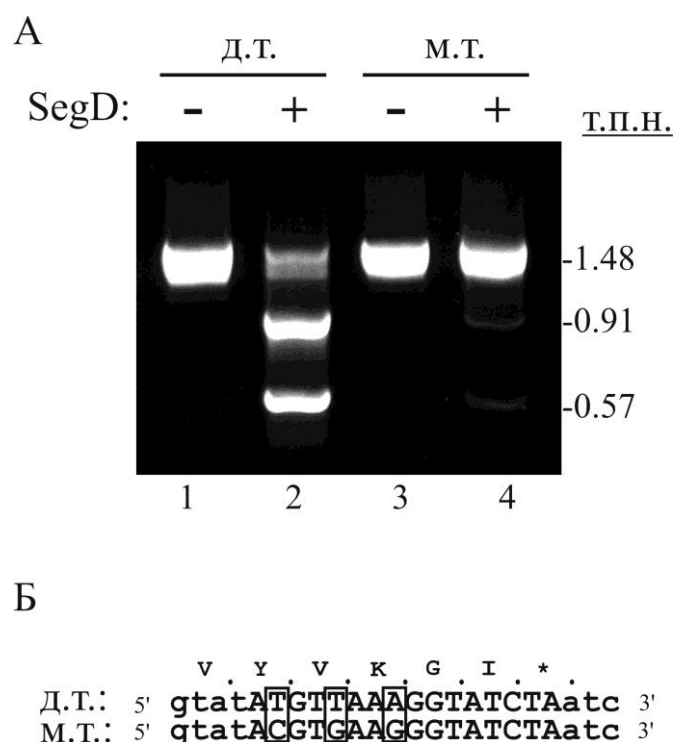


Рис. 26. Влияние нуклеотидных замен в сайте узнавания на эффективность гидролиза сайта, эндонуклеазой SegD *in vitro*. А. Гидролиз ПЦР продуктов фагов T4 (д.т.) и B20*dm* (м.т.), соответственно. 1 и 3 до инкубации с SegD 2 и 4 после инкубации с SegD.

Б. Сравнение природного (д.т.) и мутантного (м.т) сайтов SegD. Измененные нуклеотиды выделены рамкой. Последовательность сайта гидролиза приведена заглавными буквами. Сверху указана аминокислотная последовательность С-конца белка 23.

Были проведены скрещивания фага D2 с B20 и B20*dm* и проанализировано, как ОРС *segD* и *am14* наследуются потомством (табл. 3).

Таблица 3. Наследование гена *segD* и амбер мутации в гене *14* потомством смешанных инфекций *segD*⁺ и *segD*⁻ фагов.

Скрещивания	Доля потомства, содержащего ген <i>segD</i> , %	Доля потомства с <i>amber</i> -мутацией в гене <i>14</i> , %
B20 × D2	50±1,5	48±5,7
B20 ^{dm} × D2	47,8±4,3	45±3,4

Соотношение родительских фагов составляло 1:1. Приведены результаты трех независимых экспериментов для каждого скрещивания. Представлено среднее значение ± стандартное отклонение. Скрещивания фагов проводили на супрессорном штамме *E. coli* B₄₀supI. Процент потомства, содержащий *amber*-мутацию, рассчитан как отношение разницы между титром потомства на супрессорном и бессупрессорном штаммах к титру на супрессорном штамме, умноженное на 100. Процент потомства, содержащего ген *segD*, рассчитан как отношение количества фаговых бляшек, дающих положительный сигнал при гибридизации с ДНК-зондом к данному гену, к общему количеству проанализированных фаговых бляшек, умноженное на 100.

В обоих скрещиваниях доля фагов в потомстве с амбер-мутацией в гене *14* незначительно отличалась от 50%, что в пределах погрешности соответствовало титру родительских фагов. Доля фагов, содержащих ОРС *segD*, в обоих скрещиваниях не изменялась и была также около 50%. Таким образом, в модельных скрещиваниях *segD*⁺ и *segD*⁻ фагов хоминг гена *segD* не происходит независимо от интактности сайта SegD в одном из родителей.

Как было установлено ранее, эндонуклеаза SegD нарабатывается в ходе развития бактериофага T4. Было сделано предположение, что природный уровень эндонуклеазы в клетке не достаточен для инициации хоминга собственного гена. Для проверки данного предположения были проведены скрещивания фагов B20^{dm} и D2 на штамме *E. coli*, продуцирующем эндонуклеазу SegD. В качестве хозяина для скрещиваний был выбран штамм *E. coli* BL21(DE3)/pLysE/pSDet-19mod, который ранее был использован при получении препарата эндонуклеазы SegD. Через 50 мин. после добавления индуктора клетки совместно инфицировали фагами B20 и D2, либо B20^{dm} и D2. Потомство фагов было проанализировано на наследование ОРС *segD* и *am14* (табл. 4). Также экстракты клеток были проверены на

наличие SegD специфической эндонуклеазной активности и показано её существенное увеличение относительно её природного уровня при T4-инфекции. (данные не приведены).

Таблица 4. Наследование гена *segD* и амбер мутации в гене *14* потомством смешанных инфекций B20 × D2 и B20dm × D2 при повышенном содержании эндонуклеазы SegD в клетках.

Скрещивания	Доля потомства, содержащего ген <i>segD</i> , %	Доля потомства с амбер-мутацией в гене <i>14</i> , %
B20 × D2	57±3	55±3
B20dm × D2	94±5	62±6

Соотношение родительских фагов составляло 1:1. Приведены результаты трех независимых экспериментов для каждого скрещивания. Представлено среднее значение ± стандартное отклонение. Скрещивания фагов проводили на штамме *E.coli* BL21(DE3)/pLysE/pSDet-19mod. Процент потомства, содержащий амбер-мутацию, рассчитан как отношение разницы между титром потомства на супрессорном и бессупрессорном штаммах к титру на супрессорном штамме, умноженное на 100. Процент потомства, содержащего ген *segD*, рассчитан как отношение количества фаговых бляшек, дающих положительный сигнал при гибридизации с ДНК-зондом к данному гену, к общему количеству проанализированных фаговых бляшек, умноженное на 100.

Установлено, что доли фагов с амбер-мутацией, а также геном *segD* в потомстве скрещивания B20 и D2 составили 55 и 57%, соответственно. Данное значение в пределах погрешности соответствует титру родительских фагов. В случае скрещивания фагов B20dm и D2, доля потомства, содержащего ОРС *segD*, увеличилась до 94 %. При этом доля маркера *am14* составила 62%. Наблюдаемое увеличение доли гена *segD* в потомстве свидетельствует о том, что происходит хоминг данного гена. По-видимому, в условиях повышенного уровня SegD репарация двухцепочечного разрыва в SegD-сайте фага-реципиента проходит недостаточно эффективно, что приводит к частичному ограничению фага D2. Об этом свидетельствует некоторое увеличение в потомстве маркера *am14* родительского фага B20dm с нарушенным сайтом SegD.

Таким образом, хоминг *segD* может проходить в условиях повышенного содержания фермента в клетке относительно природного уровня и при наличии предпочтительного сайта гидролиза в реципиентом фаге.

3.9. Анализ SegD-инициируемой рекомбинации в *rII*-локусе фага Т4

Гены *segD* и 23 являются поздними генами (Luke et al., 2002). Ген 23 кодирует мажорный белок головки, и его продукция составляет до 30% всего синтезируемого белка на поздней стадии инфекционного цикла (Young et al., 1981). Было предположено, что высокий уровень транскрипционной активности делает сайт недоступным для эндонуклеазы SegD. Для проверки данного предположения была исследована способность SegD инициировать рекомбинацию в *rII*-локусе Т4. Ранее, для хоминг-эндонуклеаз SegC и I-TevI фага Т4, которые также кодируются поздними генами, была показана способность инициировать рекомбинацию в данном локусе после инсерции в него соответствующих сайтов этих эндонуклеаз (Shcherbakov et al., 2002, Stohr et al., 2002). В соответствии с этим был встроен фрагмент ДНК размером 37 п.н., содержащий сайт SegD, в 5'-конец гена *rIIB* фага D2. В полученном фаге D2ets9 данная вставка приводила к нарушению гена *rIIB*. Известно, что *rII*-мутанты фага Т4 не способны расти на штаммах *E. coli*, лизогенных по фагу λ . Этот признак был использован для генетического анализа потомства скрещивания фагов Т4В и D2ets9, которые обладают фенотипом *rII*⁺ и *rII*⁻, соответственно. Как ожидалось, внесение разрыва эндонуклеазой SegD в аллель *ets9* будет приводить к снижению доли маркера *rII*⁺ в потомстве, поскольку матрицей для репарации будет служить интактная аллель дикого типа фага Т4В. В качестве контроля был проведен анализ наследования *rII*-маркеров потомством скрещивания Т4В и D2insZ. Бактериофаг D2insZ содержит негидролизуюмую вставку, фрагмент ОРС *lacZ E. coli*, которая находится в том же положении ОРС *rIIB* и приводит к её нарушению (Shcherbakov et al., 2002). Результаты скрещивания фагов приведены в таблице 5 (колонка 1). Анализ показал, что в потомстве обоих скрещиваний изменение доли *rII*-мутантов не происходило. Таким образом, в отличие от SegC и I-TevI, эндонуклеаза SegD не инициирует генную конверсию в *rII*-локусе. Следовательно, неспособность SegD инициировать хоминг собственного гена не обусловлена специфическим расположением сайта SegD в фаговом геноме.

Дополнительно, была проанализирована SegD-инициируемая рекомбинация в *rII*-локусе при повышенном уровне SegD в клетке. Было проведено скрещивание фагов Т4В и D2ets9 на штамме *E. coli*, экспрессирующем SegD. Очевидно, что уровень белка SegD в клетках будет зависеть от времени, прошедшего с момента индукции синтеза данного белка. Учитывая это, генетическую конверсию в *rII*-локусе анализировали в зависимости от времени прошедшего с момента индукции синтеза эндонуклеазы SegD (таблица 5). В случае пары фагов Т4В и D2insZ,

как и ожидалось, изменения доли маркеров в потомстве не происходит. В скрещиваниях T4B × D2ets9 наблюдалось снижение доли в потомстве фагом, имеющих *rII*-фенотип (маркер *ets9*). При этом, если инфекция проводилась через 10 и 30 минут после индукции, доля *rII*-фагов уменьшалась до ~ 25%, тогда как через 50 минут – до 4%.

Необходимо отметить, что для скрещиваний, проводившихся в штамме *E. coli* с повышенным содержанием SegD в клетках, наблюдалось снижение урожая фагов (данные не приведены). Так, при инфекции через 50 минут индукции, урожай падал в ~ 2 раза, тогда как через 70 минут падение составило ~ 2 порядка. Падение урожая фагов может быть обусловлено снижением жизнеспособности клеток *E. coli* ввиду токсичности эндонуклеазы SegD. Также можно предположить, что избыток эндонуклеазы SegD на ранней стадии инфекции приводит к появлению неспецифических разрывов в фаговой ДНК, которые не способны эффективно репарироваться, что приводит к ограничению развития фага.

Таблица 5. Наследование *rII*-фенотипа потомством смешанных инфекций фагов T4B и D2insZ или D2ets9.

Скрещивания	^а Доля потомства (%) с фенотипом <i>rII</i> ⁻ .	^б Доля потомства (%) с фенотипом <i>rII</i> ⁻ .	^в Доля потомства (%) с фенотипом <i>rII</i> ⁻ .		
			10 мин.	30 мин.	50 мин.
T4B × D2insZ	50±2,8	49±1,8	49±2,5	48±3	48±3
T4B × D2ets9	51±1,3	51±1,2	25±1	20,5±5	6±1

Соотношение родительских фагов составляло 1:1. Приведены результаты двух независимых экспериментов для каждого скрещивания. Представлено среднее значение ± стандартное отклонение. Доля потомства, содержащего *rII*-мутацию, рассчитана как отношение разницы между титрами на пермиссивном и непермиссивном штаммах к титру на пермиссивном штамме, умноженное на 100. Скрещивание проводилось в условиях природного уровня SegD в клетках *E.coli* (**а**) CR-63, (**б**) BL21(DE3)/pLysE и в условиях повышенного уровня SegD (**в**) BL21(DE3)/pLysE/pSDet-19mod. 10 мин., 30 мин., 50 мин – время индукции гена *segD* до заражения фагами.

Исходя из полученных результатов, можно заключить, что природный уровень экспрессии эндонуклеазы SegD не достаточен для инициации хоминга собственного гена, а также генетической рекомбинации в *rII*-локусе.

ЗАКЛЮЧЕНИЕ

Геном бактериофага Т4 содержит 15 генов хоминг-эндонуклеаз, для большинства из которых показана способность инициировать генетический обмен между Т4 и родственными фагами. В данной работе охарактеризована новая эндонуклеаза SegD фага Т4, относящаяся к семейству GIY-YIG эндонуклеаз. Ген *segD* экспрессируется в ходе развития бактериофага Т4. Эндонуклеаза SegD узнает сайт протяженностью 16 п.н. и гидролизует его с образованием 3'-выступающих концов длиной 3 н.о. Фермент с близкой эффективностью гидролизует как немодифицированную ДНК, так и ДНК Т-четных фагов, содержащую гликозилированные остатки гидроксиметилцитозина. Сайт эндонуклеазы SegD расположен на 3'-конце гена 23 на расстоянии 52 п.н. от сайта инсерции ОРС *segD*. У 29-ти Т4-родственных бактериофагов, проанализированных в данной работе, последовательность сайта узнавания оказалась идентична. Случай с эндонуклеазой SegD является нетипичным для хоминг-эндонуклеаз, поскольку, как в геномах фагов, содержащих ген *segD*, так и родственных фагов, в которых данный ген отсутствует, имеется сайт гидролиза данной эндонуклеазы.

В условиях природного уровня экспрессии эндонуклеаза SegD не способна инициировать мобильность собственной ОРС, даже в случае, если ее сайт узнавания искусственно нарушен в геноме *segD*⁺ фага. При увеличении количества фермента в клетке, эндонуклеаза SegD способна инициировать хоминг, а также генетическую рекомбинацию в сайте SegD, искусственно введенном в *rII*-локус. По-видимому, при природном уровне экспрессии гена *segD*, активности эндонуклеазы SegD недостаточно для внесения разрывов в фаговую ДНК *in vivo*.

ВЫВОДЫ

1. Определена структурная организация области генов 23-24 фагов T2L, T6 и группы из 12-ти RB-фагов. Ген *segD*, помимо фага T4, обнаружен у фагов T6, RB55 и RB59. У остальных фагов в данном локусе ген *segD* отсутствует.
2. Показано, что SegD является Mg^{2+} -зависимой сайт-специфической эндонуклеазой. Определены оптимальные условия работы фермента: максимальная активность фермента наблюдается при значении pH 7,0 – 8,0, концентрации катионов K^+ 50 мМ, Mg^{2+} 2,5 мМ и температуре 30 °С.
3. Эндонуклеаза SegD узнает протяженную ассиметричную последовательность ДНК длиной 16 п.н. и гидролизует её с образованием 3'-выступающих концов длиной три нуклеотидных остатка. Природные модификации ДНК Т-четных бактериофагов не влияют на эффективность гидролиза эндонуклеазой SegD.
4. Сайт гидролиза эндонуклеазы SegD расположен на 3'-конце гена 23, и его последовательность идентична в геномах 29-ти проанализированных T4-родственных бактериофагов.
5. Ген *segD* экспрессируется в ходе развития бактериофага T4: в экстрактах клеток *E.coli*, инфицированных фагом, наблюдается активность, характерная для эндонуклеазы SegD.
6. Эндонуклеаза SegD не способна инициировать хоминг собственного гена при скрещивании Т-четных бактериофагов, а также генетическую рекомбинацию на модели *rII*-локуса. Для инициации данных процессов требуется повышенный, в сравнении с природным, уровень эндонуклеазы SegD в клетке.

СПИСОК СОКРАЩЕНИЙ

ОРС	открытая рамка считывания
Heg	ген хоминг-эндонуклеазы
Heg ^{+/-} (<i>heg</i> ^{+/-})	геном содержащий или не содержащий ген хоминг-эндонуклеазы
ДНК	дезоксирибонуклеиновая кислота
тРНК	транспортная рибонуклеиновая кислота
рРНК	рибосомальная рибонуклеиновая кислота
мРНК	матричная рибонуклеиновая кислота
<i>E. coli</i>	<i>Escherihia coli</i>
A (Ala)	аланин
C (Cys)	цистеин
D (Asp)	аспарагиновая кислота
G (Gly)	глицин
H (His)	гистидин
I (Ile)	изолейцин
L (Leu)	лейцин
M (Met)	метионин
Y (Tyr)	тирозин
T (Thr)	треонин
N (Asn)	аспарагин
R (Arg)	аргинин
S (Ser)	серин
E (Glu)	глутаминовая кислота
F (Phe)	фенилаланин
K (Lys)	лизин
P (Pro)	пролин
Q (Gln)	глутамин
W (Trp)	триптофан
ПЦР	полимеразная цепная реакция
т.п.н.	тысяча пар нуклеотидов
ак	аминокислота
а.о.	аминокислотный остаток
мкл	микролитр
мкМ	микромольная концентрация

mM	миллимолярная концентрация
w/v	вес/объем
v/v	объем/объем
ОД ₂₆₀ , ОД ₅₉₀	оптическая плотность при 260 и 590 нм, соответственно
кДа	килодальтон
Me ²⁺	ионы двухвалентных металлов
Ар	ампициллин
См	хлорамфеникол
(d)dNTP	(ди)дезоксинуклеозидтрифосфат
(d)АТР	(дезокси)аденозин-5'-трифосфат
БСА	бычий сывороточный альбумин
ДТТ	дититреитол
ДСН	додецилсульфат натрия
ПААГ	полиакриламидный гель
ТЕМЕД	N,N,N',N'-тетраэтилендиамид
ЭДТА	этилендиаминтетрауксусная кислота
X-gal	5-бром-4-хлор-3-индолил-Д-галактопиранозид
Трис	трис(гидроксиметил)аминометан
ИПТГ	изопропилтио-β-D-галактозид
ПЭГ 6000	полиэтиленгликоль с молекулярной массой 6000
РНКаза А	рибонуклеаза А
ДНКаза	дезоксирибонуклеаза
ТЕ	трис-ЭДТА
ДМСО	диметилсульфоксид
MOPS	морфолинпропансульфонокислота
ФМСФ	фенилметилсульфонилфлуорид

СПИСОК ЛИТЕРАТУРЫ

- Ackermann, H.-W. and Krisch, H.M. (1997) A catalogue of T4-type bacteriophages Arch. Virol., 1997. 142, 2329-45.
- Aagaard C, Dalgaard JZ, Garrett RA (1995) Intercellular mobility and homing of an archaeal rDNA intron confers a selective advantage over intron cells of *Sulfolobus acidocaldarius*. Proc Natl Acad Sci USA, 92,12285-12289.
- Argast, G. M., Stephens, K. M., Emond, M. J. and Monnat, R.J.(1998) I-PpoI and I-CreI homing site sequence degeneracy determined by random mutagenesis and sequential *in vitro* enrichment. J. Mol. Biol., 280, 345-53.
- Belle, A., Landthaler, M. and Shub, D.A. (2002) Intronless homing: site-specific endonuclease SegF of bacteriophage T4 mediates localized marker exclusion analogous to homing endonucleases of group I introns. Genes Dev., 16, 351-62.
- Belfort, M., Derbyshire, M., Parker, B., Cousineau, and A. M. Lambowitz. 2002. Mobile introns: pathways and proteins, p. 761–783. In N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D.C.).
- Belfort, M. and Bonocora, R. (2014) Homing Endonucleases: From Genetic Anomalies to Programmable Genomic Clippers. Methods Mol Biol. 1123, 1–26
- Burt, A. and Koufopanou, V.. (2004) Homing endonuclease genes: the rise and fall and rise again of a selfish element Current Opinion in Genetics & Development, 14, 609–615
- Butler G, Kenny C, Fagan A, Kurischko C, Gaillardin C, Wolfe KH (2004) Evolution of the MAT locus and its Ho endonuclease in yeast species. Proc Natl Acad Sci USA, 101,1632-1637
- Brok-Volchanskaya, V.S., Kadyrov, F.A., Sivogrivov, D.E., Kolosov, P.M., Sokolov, A.S., Shlyapnikov, M.G., Kryukov, V.M. and Granovsky, I.E. (2008) Phage T4 SegB protein is a homing endonuclease required for the preferred inheritance of T4 tRNA gene region occurring in co-infection with a related phage. Nucleic Acids Res., 36, 2094-105.
- Belfort, M. and Perlman, P. S. (1995). Mechanisms of intron mobility. J. Biol. Chem., 270, 30237–30240.
- Belfort, M., Stoddard, B., Wood, D. and Derbyshire, V.(Eds.) (2005) Homing endonucleases and inteins.
- Bryk, M., Quirk, S.M., Mueller, J.E., Loizos, N., Lawrence, C. and Belfort, M. (1993) The td intron endonuclease I-TevI makes extensive sequence-tolerant contacts across the minor groove of its DNA target. EMBO J., 12, 2141-9.

- Bryk, M., Belisle, M., Mueller, J.E. and Belfort, M. (1995) Selection of a remote cleavage site by I-tevI, the td intron-encoded endonuclease. *J. Mol. Biol.*, 247, 197-210.
- Carlson, K. and Miller, E.S. (1994). Experiments in T4 genetics. In Karam, J.D. (ed.), *Molecular biology of bacteriophage T4*. ASM Press, Washington DC, pp. 427-34.
- Chu, K., Maley, G. F, Maley, F. and M. Belfort. 1984. Intervening sequence in the thymidylate synthase gene of bacteriophage T4. *Proc. Natl. Acad. Sci. USA*, 81, 3049–3053.
- Chevalier, B.S. and Stoddard, B.L. (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.*, 29, 3757-74.
- Chevalier B.S., Monnat R.J. Jr. and Stoddard B.L. (2001) The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites. *Nat. Struct. Biol.*, 8, 312-316.
- Chevalier, B., Turmel, M., Lemieux, C., Monnat R.J. Jr. and Stoddard, B.L. 2003. Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. *J. Mol. Biol.* 329: 253-269.
- Chevalier, B., Sussman, D., Otis, C., Noel, A.J., Turmel, M., Lemieux, C., Stephens, K., Monnat, R.J.Jr. and Stoddard, B.L. (2004) Metal-dependent DNA cleavage mechanism of the I-CreI LAGLIDADG homing endonuclease. *Biochemistry*, 43, 14015-26.
- Cowan, J.A. (1998) Metal Activation of Enzymes in Nucleic Acid Biochemistry. *Chem. Rev.*, 98, 1067-88.
- Cowan, J.A. (2002). Structural and catalytic chemistry of magnesium-dependent enzymes. *Biometals.*, 15, 225-35.
- Colleaux, L., d'Auriol, L., Betermier, M., Cottarel, G., Jacquier, A., Galibert, F. and Dujon, B. (1986) Universal code equivalent of a yeast mitochondrial intron reading frame is expressed into *E. coli* as a specific double strand endonuclease. *Cell*, 44, 521-33.
- Colston M.J., and Davis E.O. (1994) The ins and outs of protein splicing elements. *Mol. Microbiol.*, 12, 359-363.
- Curcio, M.J. and Belfort, M. (1996). Retrohoming: cDNA-mediated mobility of group II introns requires a catalytic RNA. *Cell* 84, 9–12.
- Christ F., Schoettler S., Wende W., Steuer S., Pingoud A. and Pingoud V. (1999) The monomeric homing endonuclease PI-SceI has two catalytic centres for cleavage of the two strands of its DNA substrate. *EMBO J.*, 18, 6908-6916.
- Clyman, J. and Belfort, M. (1992) *Trans* and *cis* requirements for intron mobility in prokaryotic system. *Genes Dev.*, 6, 1269-1279.
- Derbyshire, V., Kowalski, J.C., Dansereau, J.T., Hauer, C.R. and Belfort, M. (1997) Two-domain structure of the td intron-encoded endonuclease I-TevI correlates with the two-domain configuration of the homing site. *J. Mol. Biol.*, 265, 494-506.

- Delahodde, A., Goguel, V., Becam, A.M., Creusot, F., Perea, J., Banroques, J. and Jacq, C. (1989) Site-specific DNA endonuclease and RNA maturase activities of two homologous intron-encoded proteins from yeast mitochondria. *Cell*, 56, 431-441.
- Delbruck, M. (1945) Interference between bacterial viruses, III. The mutual exclusion effect and the depressor effect. *J. Bacteriol*, 50, 151-170.
- Demerec, M. and Fano, U. (1945) Bacteriophages resistant mutants in *E. coli*. *Genetics*, 30, 119.
- Desplats, C. and Krisch, H.M. (2003) The diversity and evolution of the T4-type bacteriophages. *Res. Microbiol.*, 154, 259-67.
- Duan X., Gimble F.S. and Quiococho F.A. (1997) Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity. *Cell*, 89, 555-564.
- Dunin-Horkawicz, S., Feder, M. and Bujnicki, J. B.(2006) Phylogenomic analysis of the GIY-YIG nuclease superfamily. *BMC Genomics*, 7, 98
- Dujon, B. (1989). Group I introns as mobile genetic elements: facts and mechanistic speculations – a review. *Gene* 82, 91–114.
- Dujon, B. (1980) Sequence of the intron and flanking exons of the mitochondrial 21S rRNA gene of yeast strains having different alleles at the omega and rib-1 loci. *Cell*, 20, 185-97.
- Dursina, B-E., Reents, R., Niculaea, A., Veligodsky, A., Breitling, R., Pyatkov, K., Waldmann, H., Goody, R. and Alexandrov, K. (2005) A genetically encodable microtag for chemo-enzymatic derivatization and purification of recombinant proteins. *Protein Expression and Purification*, 39, 71–81.
- Eddy, S. R., and L. Gold. (1991) The phage T4 nrdB intron: a deletion mutant of a version found in the wild. *Genes Dev.*, 5, 1032–1041.
- Edgell, D.R., Stanger, M.J. and Belfort, M. (2003) Importance of a single base pair for discrimination between intron-containing and intronless alleles by endonuclease I-BmoI. *Curr. Biol.*, 13, 973-8.
- Edgell, D.R., Stanger, M.J. and Belfort, M. (2004) Coincidence of cleavage sites of intron endonuclease I-TevI and critical sequences of the host thymidylate synthase gene. *J. Mol. Biol.*, 343, 1231-41.
- Edgell, D.R., Derbyshire, V., Van Roey, P., LaBonne, S., Stanger, M., Li, Z., Boyd, T., Shub, D., and Belfort, M.. (2004) Intron-encoded homing endonuclease I-TevI also functions as a transcriptional autorepressor. *Nature Structural & Molecular Biology*
- Edgell D.R., Gibb E.A. and Belfort M (2010) Mobile DNA elements in T4 and related phages: a review in the series on bacteriophage T4 and its relatives. *Virology*, 7, 290–300.
- Feng, H., Dong, L. and Cao, W. (2006) Catalytic mechanism of endonuclease v: a catalytic and regulatory two-metal model. *Biochemistry*, 45, 10251-9.

- Galburt E.A. and Stoddard B.L. (2002) Catalytic Mechanisms of Restriction and Homing Endonucleases. *J. Biochemistry*, 41, 13851-13660
- Galburt E.A., Chadsey M.S., Jurica M.S., Chevalier B.S., Erho D., Tang W., Monnat R.J. Jr., and Stoddard B.L. (2000) Conformational changes and cleavage by the homing endonuclease I-PpoI: a critical role for a leucine residue in the active site. *J. Mol. Biol.*, 300, 877-887.
- Gimble F.S. and Thorner J. (1992) Homing of a DNA endonuclease gene by meiotic gene conversion in *Saccharomyces cerevisia*. *Nature*, 357, 301-306.
- Guhan, N. and Muniyappa, K. (2003) Structural and functional characteristics of homing endonucleases. *Crit. Rev. Biochem. Mol. Biol.*, 38, 199-248.
- Goodrich-Blair H. and Shub D.A. (1996) Beyond homing: competition between intron endonucleases confers a selective advantage on flanking genetic markers. *Cell*, 84, 211-221.
- Gott, J.M., Zeeh, A., Bell-Pedersen, D., Ehrenman, K., Belfort, M. and Shub, D.A. (1988) Genes within genes: independent expression of phage T4 intron open reading frames and the genes in which they reside. *Genes Dev.*, 2, 1791-9.
- Granovsky, I.E., Kadyrov, F.A., Pyatkov, K.I., Shlyapnikov, M.G., Ivanov, V.A. and Kryukov, V.M. (1998) Role of the T4 SegC gene in genetic exchange between T-even bacteriophages. *Molecular Genetics of bacteria & phages*. Cold Spring Harbor, New York, p. 14.
- Hambly, E., Tetart, F., Desplats, C., Wilson, W.H., Krisch, H.M. and Mann, N.H. (2001) A conserved genetic module that encodes the major virion components in both the coliphage T4 and the marine cyanophage S-PM2. *Proc. Natl. Acad. Sci. USA*, 98, 11411-6.
- Heath P.J., Stephens K.M., Monnat R.J., and Jr.Stoddard B.L. (1997) The structure of I-CreI, a group I intron-encoded homing endonuclease. *Nature Struct. Biol.*, 4, 468-476.
- Heitman, J., Zinder, N.D. and Model, P. (1989) Repair of the *Escherichia coli* chromosome after *in vivo* scission by the EcoRI endonuclease. *Proc. Natl. Acad. Sci. USA*, 86, 2281-5.
- Huang, Y. J., Parker, M. M and Belfort, M. (1999) Role of exonucleolytic degradation in group I intron homing in phage T4. *Genetics*. 153, 1501-1512.
- Ibryashkina, E., Zakharova, M., Baskunov, V., Bogdanova, E., Nagornykh, M., Denmukhamedov, M., Melnik, B., Kolinski, A., Gront, D., Feder, M., Solonin, A., and Bujnicki, J. (2007). Type II restriction endonuclease R.Eco29kI is a member of the GIY-YIG nuclease superfamily. *BMC Structural Biology*, 7, 48.
- Ichiyanagi K., Ishino Y., Ariyoshi M., Komori K. and Morikawa K. (2000) Crystal structure of an archaeal intein-encoded homing endonuclease PI-PfuI. *J. Mol. Biol.*, 300, 889-890.
- Kadyrov, F. A., Kriukov, V. M., Shliapnikov, M. G. and Baev, A .A. (1994) SegE - a new site-specific endodeoxyribonuclease from bacteriophage T4. *Dokl. Akad. Nauk*, 339, 404-406.
- Kadyrov, F.A., Kaliman, A.V. and Kryukov, V.M. (1996) *Mol. Biol. (Moscow)* 30, 1096-1106.

- Kryukov, V.M., Shlyapnikov, M.G. and Kadyrov, F.A. (1996) *Mol. Biol. (Moscow)* 30, 1307-1315
- Kadyrov, F.A., Shlyapnikov, M.G. and Kryukov, V.M. (1997) A phage T4 site-specific endonuclease, SegE, is responsible for a non-reciprocal genetic exchange between T-even-related phages. *FEBS Letters*, 415, 75-80.
- Kadyrov, F.A., Shlyapnikov, M.G. and Kryukov, V.M. (1997) A phage T4 site-specific endonuclease, SegE, is responsible for a non-reciprocal genetic exchange between T-even-related phages. *FEBS Letters*, 415, 75-80.
- Karam J.D. (ed.). (1994) *Molecular biology of bacteriophage T4*. ASM Press. Washington, D.C..
- Kutter, E., Gachechiladze, K., Poglazov, A., Marusich, E., Shneider, M., Aronsson, P., Napuli, A., Porter, D. and Mesyanzhinov, V. (1995) Evolution of T4-related phages. *Virus Genes*, 11, 285-97.
- Kostrewa, D. and Winkler, F. K. (1995) Mg^{2+} binding to the active site of EcoRV endonuclease: a crystallographic study of complexes with substrate and product DNA at 2 Å resolution. *Biochemistry*, 34, 683-696.
- Kowalski, J.C., Belfort, M., Stapleton, M.A., Holpert, M., Dansereau, J.T., Pietrokovski, S., Baxter, S.M. and Derbyshire, V. (1999) Configuration of the catalytic GIY-YIG domain of intron endonuclease I-TevI: coincidence of computational and molecular findings. *Nucleic Acids Res.*, 27, 2115-25.
- Kuhlmann U.C., Moore G.R., James R., Kleanthous C. and Hemmings A.M. (1999) Structural parsimony in endonuclease active sites: should the number of homing endonuclease families be redefined? *FEBS Lett.*, 463, 1-2.
- Kim, J.S. and Davidson, N. (1974) Electron microscope heteroduplex study of sequence relations of T2, T4, and T6 bacteriophage DNAs. *Virology*, 57, 93-111.
- Lehman, I.R. and Pratt, E.A. (1960) On the structure of the glucosylated hydroxymethylcytosine nucleotides of coliphages T2, T4, and T6. *J. Biol. Chem.*, 235, 3254-9.
- Loizos, N., George H. Silva, G., H. and Belfort, M. (1996) Intron-encoded Endonuclease I-TevII Binds Across the Minor Groove and Induces Two Distinct Conformational Changes in its DNA Substrate *J. Mol. Biol.*, 255, 412-424
- Leslie A.G., Arnott S., Chandrasekaran R., and Ratliff R.L. (1980) Polymorphism of DNA double helices. *J. Mol. Biol.*, 143, 49-72.
- Liu, Q., Belle, A., Shub, D.A., Belfort, M. and Edgell, D.R. (2003) SegG endonuclease promotes marker exclusion and mediates co-conversion from a distant cleavage site. *J. Mol. Biol.*, 334, 13-23.
- Lambowitz, A.M. and Belfort, M. (1993) Introns as mobile genetic elements. *Annu. Rev. Biochem.*, 62, 587-622.
- Lambowitz, A.M. and Zimmerly, S. (2004) Mobile group II introns. *Annu. Rev. Genet.*, 38, 1-35.

- Lykke-Andersen, J., Garrett, R.A. and Kjems, J. (1997) Mapping metal ions at the catalytic centres of two intron-encoded endonucleases. *EMBO J.*, 16, 3272-81.
- Miyake T, Hiraishi H, Sammoto H, Ono BI (2003) Involvement of the VDE homing endonuclease and rapamycin in regulation of the *Saccharomyces cerevisiae* GSH11 gene encoding the high affinity glutathione transporter. *J Biol Chem*, 278, 39632-39636.
- Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T. and Ruger W. (2003) Bacteriophage T4 genome. *Microbiol. Mol. Biol. Rev.*, 67, 86-156.
- Minton, K.W. (1994) DNA repair in the extremely radioresistant bacterium *Deinococcus radiodurans*. *Mol. Microbiol.*, 13, 9-15.
- Mosig, G. (1994) Homologous recombination. *Molecular biology of bacteriophage T4* (ed. J.Karam). ASM Press. Washington. DC., p. 54-82.
- Mueller, J.E., Smith, D., Bryk, M. and Belfort, M. (1995) Intron-encoded endonuclease I-TevI binds as a monomer to effect sequential cleavage via conformational changes in the td homing site. *EMBO J.*, 14, 5724-35
- Mueller, J.E., Clyman, J., Huang, Y.-J., Parker, M.M. and Belfort, M. (1996) Intron mobility in phage T4 occurs in the context of recombination-dependent DNA replication by way of multiple pathways. *Genes Dev.*, 10, 351-364.
- Nassif, N., Penney, J., Pal, S., Engels, W.R. and Gloor, G.B. (1994) Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol. Cell. Biol.*, 14, 1613-1625.
- Nelson, H.C., Finch, J.T., Luisi, B.F. and Klug, A. (1987) The structure of an oligo(dA)oligo(dT) tract and its biological implications. *Nature*, 330, 221-6.
- Orlowski, J., Boniecki, M. and Bujnicki, J. (2007) I-Ssp6803I: the first homing endonuclease from the PD-(D/E)XK superfamily exhibits an unusual mode of DNA recognition. *Bioinformatics Discovery Note*, 23, 527-530
- Parker, M., Belisle, M. and Belfort, M. (1999) Intron homing with limited exon homology: illegitimate double-strand-break repair in intron acquisition by phage T4. *Genetics*. 153, 1513-1523.
- Rak, A., Niculae, A., Kalinin, A., Thoma, N., Sidorovitch, V., Goody, R. and Alexandrov, K. (2002) *In Vitro* Assembly, Purification, and Crystallization of the Rab Geranylgeranyl Transferase:Substrate Complex. *Protein Expression and Purification*, 25, 23-30
- Roberts, R.J., Belfort, M., Bestor, T., Bhagwat, A.S., Bickle, T.A., Bitinaite, J., Blumenthal, R.M., Degtyarev, S.Kh., Dryden, D.T., Dybvig, K., Firman, K., Gromova, E.S., Gumport, R.I., Halford, S.E., Hattman, S., Heitman, J., Hornby, D.P., Janulaitis, A., Jeltsch, A., Josephsen, J., Kiss, A., Klaenhammer, T.R., Kobayashi, I., Kong, H., Kruger, D.H., Lacks, S., Marinus, M.G., Miyahara, M., Morgan, R.D., Murray, N.E., Nagaraja, V., Piekarowicz, A., Pingoud, A., Raleigh, E., Rao, D.N.,

- Reich, N., Repin, V.E., Selker, E.U., Shaw, P.C., Stein, D.C., Stoddard, B.L., Szybalski, W., Trautner, T.A., Van Etten, J.L., Vitor, J.M., Wilson, G.G. and Xu, S.Y. (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, 31, 1805-1812.
- Roberts, A. P., Chandler, M., Courvalin, P., Guedon, G., Mullany, P., Pembroke, T., Rood, J., Smith, J., Summers, A., Tsuda, T. and Berg, D. E. (2008) Revised nomenclature for transposable genetic elements. *Plasmid*, 60, 167–173
- Robbins, J.B, Stapleton, M., Stanger, M. J., Smith, D., Dansereau, J. T, Derbyshire, V. and Belfort, M. (2007) Homing endonuclease I-TevIII: dimerization as a means to a double-strand break. *Nucleic Acids Res.*, 1–12
- Russel, R.L. and Huskey, R.J. (1974) Partial exclusion between T-even bacteriophages: an incipient genetic isolation mechanism. *Genetics*, 78, 989-1014.
- Russell, R.L. (1974) Comparative genetics of the T-even bacteriophages. *Genetics*, 78, 967-88.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Ed.2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Shub, D. A., J. M. Gott, M. Q. Xu, B. F. Lang, F. Michel, J. Tomaschewski, J. Pedersen-Lane, and M. Belfort. (1988) Structural conservation among three homologous introns of bacteriophage T4 and the group I introns of eukaryotes. *Proc. Natl. Acad. Sci. USA*, 85, 1151–1155.
- Schwarz, H., Riede, I., Sonntag, I. and Henning, U. (1983) Degrees of relatedness of T-even type E. coli phages using different of the same receptors and topology of serologically cross-reacting sites. *EMBO J.*, 2, 375-80.
- Shcherbakov, V., Granovsky, I., Plugina, L., Shcherbakova, T., Sizova, S., Pyatkov, K., Shlyapnikov, M. and Shubina, O. (2002) Focused genetic recombination of bacteriophage t4 initiated by double-strand breaks. *Genetics*, 162, 543-56.
- Shcherbakov, V. P., L. A. Plugina, L., Kudryashova, E., Efremova, O., Sizova, S (1982) Marker-dependent recombination in T4 bacteriophage. I. Outline of the phenomenon and evidence suggesting mismatch repair mechanism. *Genetics*, 102, 615–625.
- Szostak, J.W., Orr-Weaver, T.L, Rothstein, R.J. and Stahl, F.W. (1983) The double-strand-break repair model for recombination. *Cell* 33, 25-35.
- Snustad, D.P., Haas, N. and Oppenheimer, D.G. (1986) The bacteriophage regulatory protein gp alc/unf binds to DNA in the absense of RNA-polymerase. *J. Virol.*, 60, 1145-7.
- Streisinger G. and Weigle J. (1956) Properties of bacteriophage T2 and T4 with unusual inheritance. *PNAS*, 1956, 42, 504-10.
- Stoddard, B.L. (2005) Homing endonuclease structure and function. *Q. Rev. Biophys.*, 38, 49-95.

- Stoddard, B.L. (2011) Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure*, 19, 7–15.
- Stoddard, B.L. (2014) Homing endonucleases from mobile group I introns: discovery to genome engineering. *Mobile DNA*, 5:7
- Stohr, B.A. and Kreuzer, K.N. (2002) Coordination of DNA ends during double-strand-break repair in bacteriophage T4. *Genetics*, 162, 1019-30.
- Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, 189, 113-30.
- Studier, W., Rosenberg, A., Dunn, J. and Dubendorff, J. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymology*, 185, 60-89.
- Sharma, M., Ellis, R.L. and Hinton D.M. (1992) Identification of a family of bacteriophage T4 genes encoding proteins similar to those present in group I introns of fungi and phages. *Proc. Natl. Acad. Sci. USA*, 89, 6658-62.
- Sharma, M. and Hinton, D. M. (1994) Purification and characterization of the SegA protein of bacteriophage T4, an endonuclease related to proteins encoded by group I introns. *J. Bacteriol.*, 176, 6439-6448
- Shen, B. W., Landthaler, M., Shub, D.A. and Stoddard, B.L. 2004. DNA binding and cleavage by the HNH homing endonuclease I-*HmuI*. *J. Mol. Biol.* 342: 43-56.
- Sandegren, L., Nord, D. and Britt-Marie Sjöberg, B-M. (2005) SegH and Hef: two novel homing endonucleases whose genes replace the *mobC* and *mobE* genes in several T4-related phages. *Nucleic Acids Res.*, 33, 6203–6213
- Silva G.H., Dalgaard J.Z., Belfort M., Van Roey P.(1999) Crystal structure of the thermostable archaeal intron-encoded endonuclease I-*DmoI*. *J. Mol. Biol.*, 286, 1123-1136.
- Taylor G, Heiter D, Pietrokovski S, Stoddard B.(2011) Activity, specificity and structure of I-Bth0305I: a representative of a new homing endonuclease family. *Nucleic Acids Res.*, 30, 9705–9719.
- Tetart, F., Desplats, C., Kutateladze, M., Monod, C., Ackermann, H.W. and Krisch, H.M. (2001) Phylogeny of the major head and tail genes of the wideranging T4-type bacteriophages. *J. Bacteriol.*, 183, 358-6.
- Truglio, J.J., Rhau, B., Croteau, D.L., Wang, L., Skorvaga, M., Karakas, E., DellaVecchia, M.J., Wang, H., Van Houten, B. and Kisker, C. (2005) Structural insights into the first incision reaction during nucleotide excision repair. *EMBO J.*, 24, 885-94.
- Van Roey, P., Waddling, C.A., Fox, K.M., Belfort, M. and Derbyshire, V. (2001) Intertwined structure of the DNA-binding domain of intron endonuclease I-TevI with its substrate. *EMBO J.*, 20, 3631-7.
- Vetter, D. and Sadowski, P.D. (1974) Point mutants in the D2a region of bacteriophage T4 tail to induce T4 endonuclease IV. *J. Virol.*, 14, 207-13.

- Vipond, I.B., Baldwin, G.S. and Halford, S.E. (1995). Divalent metal ions at the active sites of the EcoRV and EcoRI restriction endonucleases. *Biochemistry*, 34, 697-704.
- Vieira J. and Messing J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with syntetic universal primers. *Gene*, 19, 259-68.
- Vipond, I.B., Baldwin, G.S. and Halford, S.E. (1995). Divalent metal ions at the active sites of the EcoRV and EcoRI restriction endonucleases. *Biochemistry*, 34, 697-704.
- Wang J., Kim H.-H., Yuan X. and Herrin D. L. (1997) Purification, biochemical characterization and protein-DNA interactions of the I-CreI endonuclease produced in *Escherichia coli*. *Nucl. Acids Res.*, 25, 3767-3776.
- Wenzlau, J.M., Saldanha, R.J., Butow, R.A. and Perlman, P.S. (1989) A latent intron-encoded maturase is also an endonuclease needed for intron mobility. *Cell*, 56, 421-30.
- Williams J.G. and Gratzner W.B.(1971) Limitations of the detergent polyacrilamide gel electrophoresis method for molecular weight determination. *J. Chromatogr.*, 57, 121-125.
- Wilson, G.W. and Edgel D. R. (2009) Phage T4 promotes trans homing of the defunct homing endonuclease I-TevIII. *Nucleic Acids Res.*, 37, 7110-7123
- Winkler, F.K., Banner, D.W., Oefner, C., Tsernoglou, D., Brown, R.S., Heathman, S. P., Bryan, R.K., Martin, P. D., Petratos, K. and Wilson, K.S. (1993) The crystal structure of *EcoRV* endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.*, 12, 1781-1795.
- Wittmayer P.K. and Raines R.T. (1996) Substrate binding and turnover by the highly specific I-PpoI endonuclease. *Biochem.*, 35, 1076-1083.
- Wyatt, G.R. and Cohen, S.S. (1952) A new pyrimidine base from bacteriophage nucleic acids. *Nature*, 170, 1072-3.
- Warner, H.R. (1971) Partial suppression of bacteriophage T4 ligase mutations by T4 endonuclease II deficiency: role of host DNA ligase. *J. Virol.*, 7, 534-6.
- Warren, R.A.J. (1980) Modified bases in bacteriophages DNAs. *Ann. Rev. Microbiol.*, 34, 137-58.
- Wood, W.B. (1966) Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. *J. Mol. Biol.*, 16, 118-33.
- Wood, W.B. and Revel, H.R. (1976) The genome of bacteriophage T4. *Bacteriological Rev.*, 40, 847-68.
- Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson. S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z. and Graham, M.W. (1989) Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res.*, 17, 3469-78.
- Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, 33, 103-119.

- Zimmerly H.G., Perlman P.S., and Lambowitz A.M. (1995) Group II Intron mobility occurs by target DNA-primed reverse transcription. *Cell*, 82, 545.
- Zhao, L., Bonocora, R., Shub, D. and Stoddard, B. (2007) The restriction fold turns to the dark side: a bacterial homing endonuclease with a PD-(D/E)-XK motif. *The EMBO Journal*, 26, 2432–2442
- Акуленко Н. В. Характеристика нового семейства сайт-специфических эндонуклеаз, кодируемых T5-подобными бактериофагами. Кандидатская диссертация. Пущино. 2006.
- Брок-Волчанская В.С. Свойства эндонуклеазы SegB бактериофага T4 и ее роль в генетическом обмене между T-четными бактериофагами. Кандидатская диссертация. Пущино. 2008.
- Грановский И.Э. Участие хоминг-эндонуклеазы SegC бактериофага T4 в генетической рекомбинации у T-четных бактериофагов. Кандидатская диссертация. Пущино. 2009.
- Кадыров Ф.А. SegE-новая сайт специфическая эндодезоксирибонуклеаза бактериофага T4. Кандидатская диссертация. Пущино. 1996.
- Таняшин В.И. Бактериофаг T4: молекулярное клонирование генов и суперпродукция ферментов метаболизма ДНК. Докторская диссертация. Пущино. 1983.
- Евгеньев М.Б.(2007) Мобильные элементы и эволюция генома. *Молекулярная биология*, 41, 234-245