

На правах рукописи

Пугач Ксения Сергеевна

АДАПТИВНАЯ ИММУННОСТЬ К БАКТЕРИОФАГАМ ЗА СЧЕТ
ДЕЙСТВИЯ CRISPR СИСТЕМЫ *ESCHERICHIA COLI*.

Специальность: 03.01.03 – молекулярная биология

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата биологических наук

Москва - 2011

Работа выполнена в Лаборатории молекулярной генетики микроорганизмов Учреждения Российской академии наук Института биологии гена РАН, а также в Лаборатории регуляции экспрессии генов мобильных элементов прокариот Отдела молекулярной генетики клетки Учреждения Российской академии наук Института молекулярной генетики РАН, Лаборатории К. Северинова в Институте Ваксмана (Пискатавэй, США), Кафедре молекулярной биологии биологического факультета МГУ им. М.В. Ломоносова

Научный руководитель: доктор биологических наук,
профессор К.В. Северинов

Официальные оппоненты: доктор биологических наук
А.В. Кульбачинский

кандидат биологических наук
И.И. Артамонова

Ведущая организация: Учреждение Российской академии наук
Институт проблем передачи информации
им. А.А. Харкевича РАН

Защита диссертации состоится « 1 » декабря 2011 года в 12 часов на заседании Диссертационного совета Д002.037.01 при Учреждении Российской академии наук Института биологии гена РАН по адресу: 119334, Москва, ул. Вавилова, д. 34/5.

С диссертацией можно ознакомиться в библиотеке Учреждения Российской академии наук Института молекулярной биологии им. В.А. Энгельгардта РАН по адресу: 119991, г. Москва, ул. Вавилова, д. 32.

Автореферат разослан «26» октября 2011 г.

Ученый секретарь диссертационного совета,
канд. фарм. наук

Л.С. Грабовская

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы.

Значительная часть эубактериальных и архейных геномов представлена генами, приобретенными за счет горизонтального переноса генов (ГПГ). Основными механизмами, обеспечивающими ГПГ, считаются поглощение ДНК из внешней среды и перенос чужеродного генетического материала мобильными генетическими элементами, такими как плазмиды или бактериофаги. Встраивание чужеродной ДНК в геном может приводить к нарушению функций клеточных генов, поэтому у прокариот существуют защитные механизмы, препятствующие проникновению и/или встраиванию чужеродной ДНК. Одним из таких механизмов являются CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) системы. CRISPR локус (часто называемый «CRISPR-кассета») представляет собой набор коротких палиндромных повторов ДНК, разделенных спейсерами - участками ДНК одинаковой длины, но различающимися по последовательности. В непосредственной близости от большинства CRISPR кассет находятся *cas*-гены (CRISPR-associated genes; гены, ассоциированные с CRISPR кассетой). продукты которых обеспечивают функционирование CRISPR локусов.

CRISPR системы способны придавать клетке устойчивость к бактериофагам, геномные последовательности которых содержат участки, полностью совпадающие с последовательностями хотя бы одного из спейсеров CRISPR кассеты. Но самым неожиданным свойством CRISPR оказалось то, что иммунитет, который они обеспечивают, адаптивен: клетка получает информацию о том, от какой чужеродной ДНК (бактериофаги, плазмиды) ей необходимо защищаться лишь после проникновения этой ДНК в клетку. Было обнаружено, что в ходе инфекции *Streptococcus thermophilus* бактериофагом в CRISPR кассеты этого микроорганизма встраиваются новые спейсеры, идентичные участкам генома фага. Клетки, несущие новые вставки, становятся устойчивыми к данному бактериофагу (Barrangou *et al.*, 2007). Однако механизм встраивания новых спейсеров остается совершенно

неизученным, в связи с чем, особенно важной задачей является создание удобной экспериментальной модели для исследования этого процесса.

К сожалению, CRISPR система *E. coli*, наиболее изученной бактерии, по-видимому, не обеспечивает иммунитета, несмотря на то, что даже лабораторные штаммы имеют протяженные CRISPR-кассеты и полный набор *cas*-генов. В отличие от других бактерий и архей лишь очень малый процент CRISPR спейсеров *E. coli* соответствует известным фагам или другим мобильным генетическим элементам. Таким образом, приходится предположить наличие очень большого количества пока еще неизвестных мобильных элементов *E. coli*, иммунность к которым могла бы объяснить наблюдаемое разнообразие спейсеров. Однако, такое предположение представляется маловероятным, учитывая степень изученности самой *E. coli* и фагов этой бактерии.

Отсутствие наблюдаемой функции CRISPR локуса *E. coli* может также объясняться низким уровнем экспрессии в стандартных лабораторных условиях. В единственной работе, где удалось показать возникновение антивирусного иммунитета у *E. coli* за счет действия CRISPR системы, авторы гиперэкспрессировали как Cas белки, так и CRISPR кассету (Brouns *et al.* 2008). Отсутствие видимой функции CRISPR системы у *E. coli* в значительной степени ограничивает прогресс в понимании механизмов функционирования CRISPR-систем вообще, т.к. для *E. coli* имеется не только большое количество специализированных биохимических, молекулярно-генетических методов и уникальных генетических коллекций, но и набор тщательно изученных бактериофагов. Выяснение транскрипционной организации CRISPR-локуса *E. coli* и набора Cas белков, осуществляющих функционирование этого локуса, позволило бы лучше понять причины инертности CRISPR системы *E. coli*, а также условия, которые могут повышать уровень экспрессии каждой из транскрипционных единиц локуса.

Все вышесказанное определяет актуальность темы настоящей диссертационной работы, а исследования, проведенные в рамках этой темы, позволят изучить транскрипционную активность CRISPR локусов *E. coli* K12,

а также выявить механизм встраивания новых спейсеров в CRISPR кассеты этого микроорганизма.

Цель работы и задачи исследования.

Целью данной работы было изучение транскрипционной активности CRISPR локусов *Escherichia coli* K12, а также деталей развития адаптивного иммунитета данного микроорганизма к бактериофагам.

Для реализации поставленной цели были определены следующие экспериментальные задачи:

1. Выявить полноразмерный и процессированные продукты всех CRISPR кассет *E. coli* K12.
2. Сравнить количества транскриптов в исходном штамме и штаммах, несущих делеции индивидуальных *cas*-генов.
3. Определить условия, необходимые для встраивания новых спейсеров в CRISPR кассеты *E. coli* K12, охарактеризовать эти спейсеры и соответствующие им протоспейсеры в геномах бактериофагов.
4. Выявить ключевые *cas*-гены, продукты которых задействованы в процессе встраивания новых спейсеров.

Научная новизна работы. Впервые обнаружены и охарактеризованы короткие РНК, соответствующие всем спейсерам CRISPR кассет *E. coli* K12. Детектирован полноразмерный транскрипт (предшественник) CRISPR I кассеты. Показано, что процессинг предшественника в клетках дикого типа идет неэффективно из-за недостаточной экспрессии гена *casE*, кодирующего нуклеазу. Созданы две оригинальные модельные системы для изучения встраивания новых спейсеров в CRISPR кассеты *E. coli* K12. С использованием этих модельных систем получены и охарактеризованы новые CRISPR вставки, а также установлено, что встраивание новых спейсеров в процессе фаговой инфекции происходит с высокой эффективностью только в клетках, CRISPR кассеты которых уже содержат спейсер, идентичный участку генома инфицирующего фага. Экспериментально доказано участие белков Cas1 и Cas2 в процессе встраивания новых спейсеров. Получено

доказательство того, что нуклеаза Cas3 разрезает чужеродную ДНК *in vivo*. Результаты работы вносят существенный вклад в развитие представлений о функционировании CRISPR систем иммунитета бактерий и раскрывают детали механизма встраивания новых спейсеров в CRISPR кассеты.

Научно-практическая значимость работы. Представленная работа является оригинальным исследованием и позволяет дополнить новыми данными серию работ, проводимых в современной биологической науке и направленных на изучение CRISPR систем иммунитета бактерий. Предложенный в работе метод получения вставок в CRISPR кассеты может быть использован в биотехнологии и пищевой промышленности при создании устойчивых к бактериофагам штаммов-продуцентов.

Апробация работы. Результаты диссертации были представлены на международных конференциях: Regulatory RNA in prokaryotes, June, 3-6, 2009, Берлин, Германия; 2nd, 3rd, 4th CRISPR research meeting, June-July, 2009, 2010, 2011, Беркли, США; CRISPR: Mechanisms and Applications, October, 21-22, 2010, Вагенинген, Нидерланды; 111th ASM General Meeting, May, 21-24, 2011, Новый Орлеан, США.

Публикации. По материалам диссертации опубликовано 4 печатных работы: 1 статья в рейтинговом международном журнале и 3 тезисов международных конференций.

Структура и объем диссертации. Диссертационная работы изложена на 103 страницах машинописного текста и состоит из следующих разделов: введение, обзор литературы, материалы и методы, результаты и обсуждение, выводы и список цитируемой литературы. Список литературы включает в себя 114 источников. Работа содержит 21 рисунок и 4 таблицы.

СОДЕРЖАНИЕ РАБОТЫ

Детекция коротких РНК, соответствующих кассетам CRISPR I, CRISPR II и III *Escherichia coli*.

В геноме *Escherichia coli* K12 закодированы Cas белки CasA, CasB, CasC, CasD, CasE, Cas1, Cas2 и Cas3; а также три кассеты – CRISPR I (содержащая 13 спейсеров и 14 повторов), CRISPR II (6 спейсеров и 7 повторов) и CRISPR III (2 спейсера и 3 повтора). Белки CasA, CasB, CasC, CasD и CasE образуют комплекс, получивший название Cascade (CRISPR-activated complex for antiviral defence; комплекс защиты от вирусов, активируемый CRISPR).

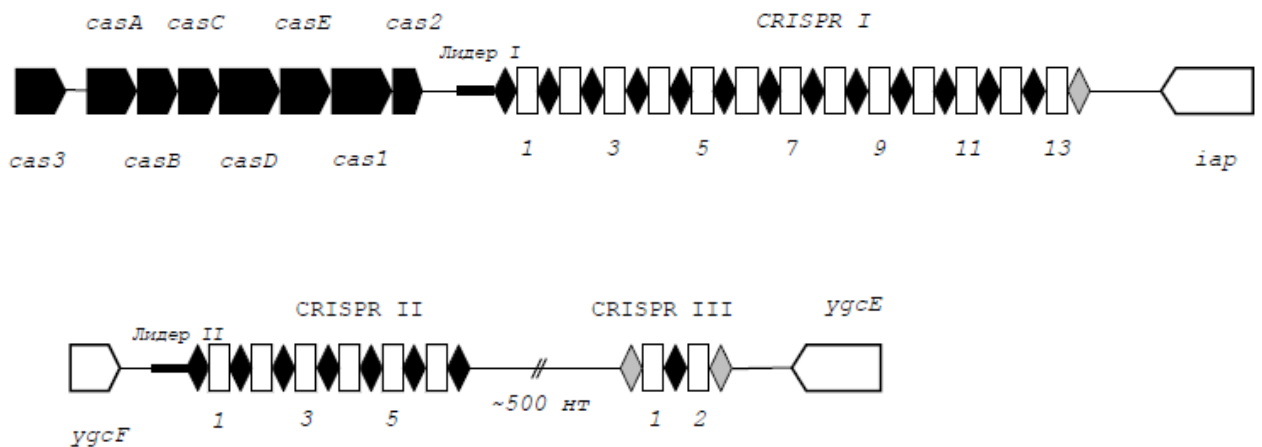


Рис. 1. Структура CRISPR-локусов *Escherichia coli* K12 MG1655. *cas*-гены показаны в виде черных пятиугольников; пятиугольники, изображающие гены, не принадлежащие семейству *cas*, не закрашены. Черными ромбами отмечены CRISPR-повторы, белыми прямоугольниками – спейсеры. Несовершенные (содержащие замены) повторы закрашены серым. Отмечены также последовательности лидеров кассет CRISPR I и II.

В 2008 г. Brouns и сотр. детектировали короткую РНК, соответствующую спейсеру №4 CRISPR I кассеты *E. coli* K12. Количество этой РНК в клетках дикого типа было чрезвычайно мало, но в изогенных штаммах из KEIO-коллекции делеционных мутантов *E. coli* K12, не содержавших гены *casA*, *casB* или *casC*, возрастало на несколько порядков (Brouns *et al.*, 2008 и Рис. 2).

Нами были получены данные, существенно дополняющие результаты Brouns и сотр. (Рис. 2, Рис. 3). С помощью нозерн-блот анализа были выявлены короткие

РНК (длиной около 60 нт), соответствующие каждому спейсеру всех трех CRISPR локусов *E. coli* K12 за исключением последнего (№13) спейсера CRISPR I.

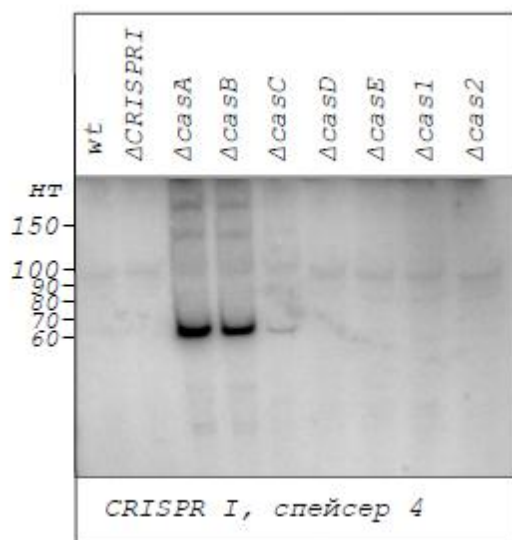


Рис. 2. Нозерн-блот, демонстрирующий количество короткого транскрипта спейсера №4, CRISPR I в штаммах с делециями различных *cas*-генов. Направление транскрипции от лидера к спейсеру №13 (см. Рис. 1).

Специфические сигналы гибридизации были получены только для одной из двух возможных ориентаций CRISPR транскрипта – от лидерной последовательности к концу кассеты (то есть транскрипция CRISPR локуса сонаправлена с транскрипцией *cas*-генов).

Отсутствие короткого транскрипта спейсера №13, CRISPR I может объясняться тем, что последовательность повтора, следующего за этим спейсером (см. Рис. 1) содержит несколько нуклеотидных замен, что, возможно, предотвращает процессинг.

Чтобы оценить количество коротких транскриптов каждого спейсера кассеты CRISPR I, метод нозерн-блот анализа был модифицирован следующим образом. На ПААГ параллельно с тотальной РНК, выделенной из клеток дикого типа или KE10-мутанта *ΔcasA*, наносились известные количества дезоксиолигонуклеотида, комплементарного радиоактивно-меченому зонду. Количество короткой РНК, соответствующей тому или иному спейсеру, определялось путем сравнения интенсивности гибридизации зонда с тотальной РНК и с олигонуклеотидом, количество которого было известно (Рис. 3). Используемый полуколичественный метод выявил наличие 200-2000 молекул коротких CRISPR-специфических РНК на

клетку в $\Delta casA$ штамме, что примерно в 100 раз больше оцененного количества этих РНК в клетках дикого типа.

Для *Pyrococcus furiosus* было показано, что чем дальше спейсер находится от последовательности лидера, тем меньше в клетке обнаруживается соответствующей короткой РНК. Образование такого градиента, возможно обусловлено терминацией транскрипции на последовательностях повторов, которые могут образовывать протяженные тугоплавкие шпильки. Нами было показано, что в *E. coli* К12 такой градиент отсутствует (Рис. 3).

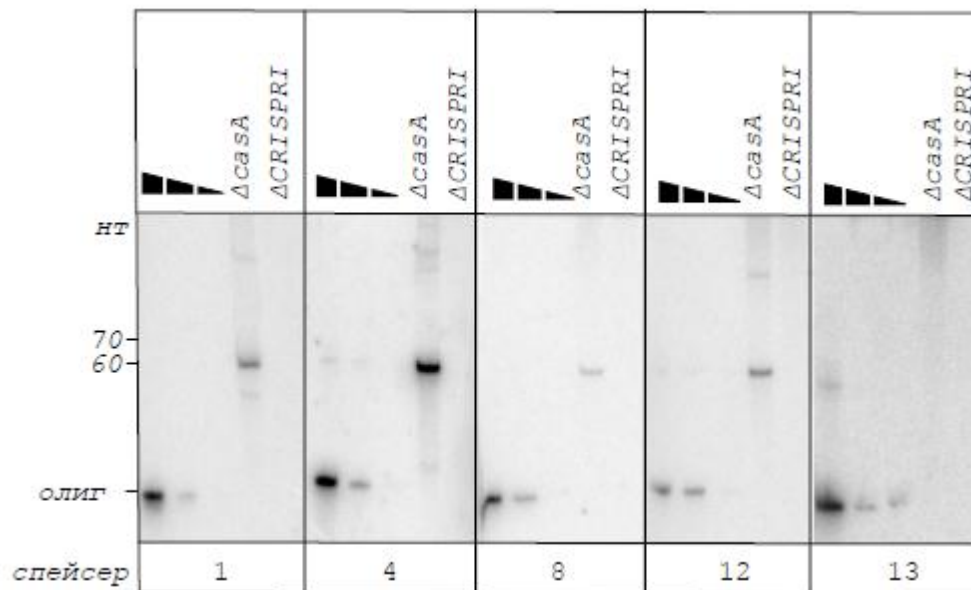


Рис. 3. Нозерн-блот с тотальной РНК, выделенной из $\Delta casA$ мутанта KEIO-коллекции. Представлена гибридизация с радиоактивно мечеными олигонуклеотидами, соответствующими выборочным спейсерам (№1, №4 и т.д.). На параллельные дорожки были нанесены различные количества немеченого олигонуклеотида, соответствующего последовательности спейсера. Тотальная РНК, выделенная из клеток несущих делецию кассеты CRISPR I, выступала в качестве отрицательного контроля.

Увеличение количества коротких РНК в KEIO-мутантах связано с повышенной экспрессией белка CasE.

В нашей работе было проведено подробное изучение причины наблюдаемого резкого увеличения количества коротких CRISPR РНК в некоторых *cas*-мутантах, показанного в работе Brouns *et al.*, 2008. Были проведены эксперименты по

комплементации эффекта увеличения количества коротких CRISPR РНК в делеционных мутантах *casA*, *casB* и *casC* из KEIO-коллекции. Добавление в клетку экстрахромосомных копий *casA*, *casB* и *casC* не приводило к уменьшению количества короткой РНК спейсера №4 кассеты CRISPR I в соответствующих делеционных мутантах (Рис. 6). Нами было сделано предположение о том, что эффект увеличения количества коротких CRISPR РНК обусловлен не собственно отсутствием в клетке белков CasA, CasB и CasC (что предположили Brouns и сотр. в своей работе), а увеличением экспрессии нижележащих генов *casD*, *casE*, *casI* или *cas2*. Описанное увеличение экспрессии могло происходить за счет дополнительной транскрипции этих генов с промотора гена канамициновой устойчивости, встроенного на место генов *casA*, *casB* и *casC* при получении KEIO-мутантов.

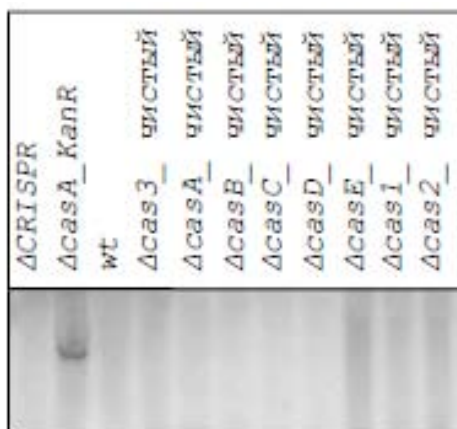


Рис. 4. Нозерн-блот, демонстрирующий количество короткого транскрипта спейсера №4, CRISPR I в штаммах с делециями различных *cas*-генов и удаленной кассетой канамициновой устойчивости.

Для проверки этого предположения ген канамициновой устойчивости и его промотор были удалены из каждого *cas*-мутанта из KEIO-коллекции. Тотальная РНК выделенная из таких штаммов была проанализирована с помощью нозерн-блоттинга. Как оказалось, в полученных нами «чистых» делеционных мутантах увеличения количества коротких CRISPR РНК не детектировалось (Рис. 4).

С помощью ОТ-ПЦР в KEIO-мутантах *ΔcasA*, *ΔcasB* и *ΔcasC* было выявлено значительное усиление транскрипции генов *casD*, *casE*, *casI* и *cas2*, тогда как в «чистых» делеционных мутантах количество транскриптов этих генов оставалось неизменным (см. Рис.5).

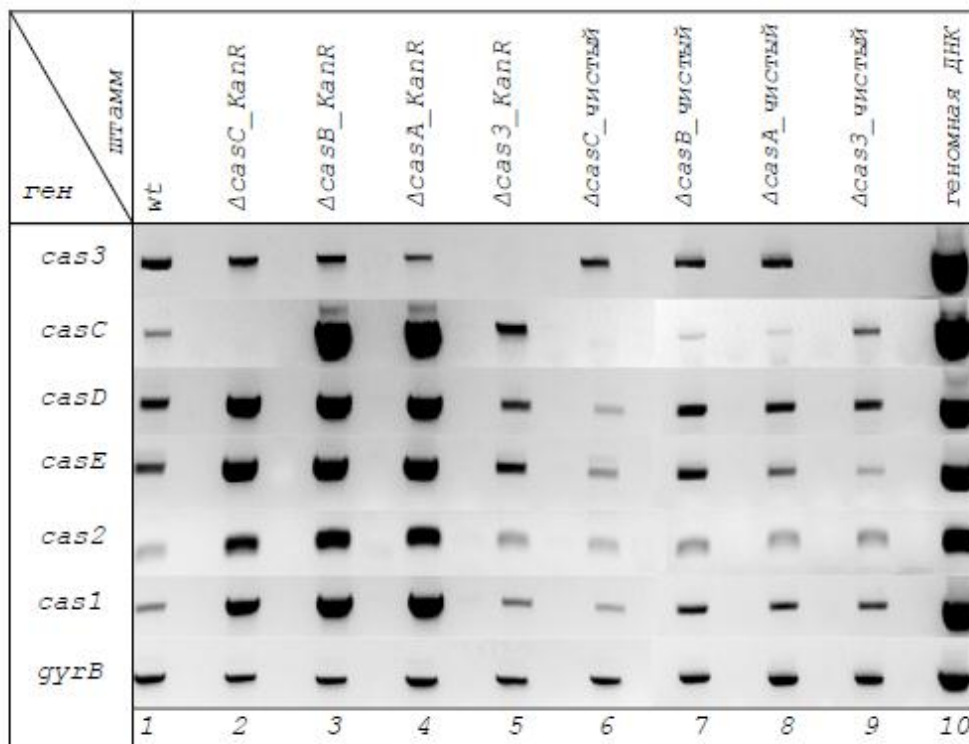
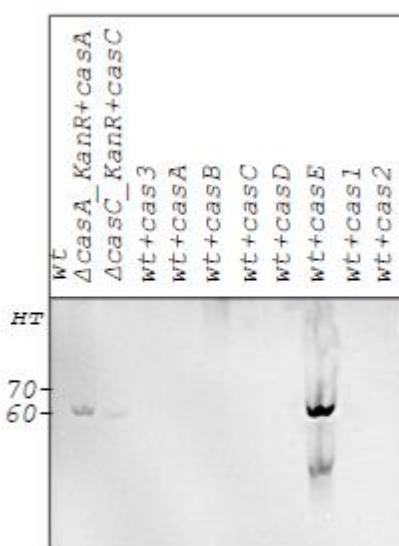


Рис. 5. Электрофореграмма, демонстрирующая количества транскриптов различных *cas*-генов в делеционных мутантах из KEIO-коллекции (KanR) и «чистых» (без гена устойчивости к канамицину) делеционных мутантах. Данные получены методом ОТ-ПЦР со специфическими к каждому гену праймерами. Ген домашнего хозяйства *gyrB* использован в качестве контроля на одинаковое количество РНК в пробах.



Чтобы выяснить, повышенная экспрессия какого *cas*-гена ответственна за увеличение количества короткой CRISPR РНК, в клетки дикого типа были введены плазмиды, несущие копии различных *cas*-генов под контролем индуцируемого промотора *T5-lac*. Оказалось, что увеличенное количество коротких CRISPR РНК обнаруживалось лишь в индуцированных клетках, содержащих плазмиду с геном *casE* (Рис. 6).

Рис. 6. Нозерн-блот, демонстрирующий количество РНК спейсера №4, CRISPR I в клетках дикого типа и указанных *cas*-мутантах из KEIO-коллекции, в которые были введены плазмиды, экспрессирующие различные *cas*-гены.

Повышение содержания процессированных CRISPR РНК в клетках, экспрессирующих большое количество CasE, может быть объяснено: 1) увеличением стабильности коротких РНК; 2) увеличением уровня транскрипции CRISPR кассет (числа молекул предшественника коротких РНК); 3) увеличением эффективности процессинга предшественника и образованием большого числа коротких РНК.

Ингибирование транскрипции антибиотиком рифампицином показало, что как в KEIO штамме *ΔcasA*, так и в штамме дикого типа короткие транскрипты стабильны в течение не менее 25 мин (Рис. 7). Таким образом, изменение стабильности коротких CRISPR РНК скорее всего не является причиной их накопления в штаммах с повышенной экспрессией гена *casE*.

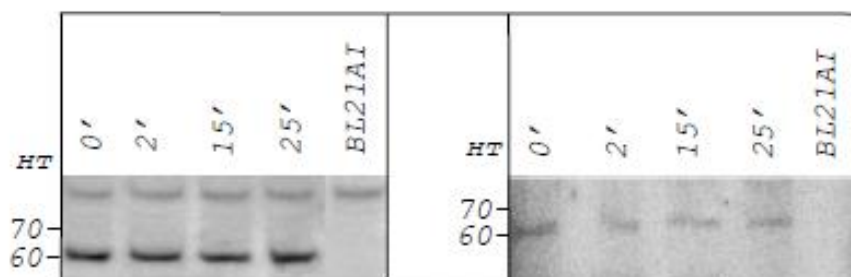


Рис. 7. Нозерн-блот, демонстрирующий количества короткой РНК спейсера №4, CRISPR I в клетках дикого типа (справа) и *ΔcasA* мутанте из KEIO-коллекции (слева). После обработки рифампицином равные объемы культуры отбирались через указанные промежутки времени (0 минут, 2 минуты и т.д.) для выделения и анализа РНК.

Нозерн-блот анализ полноразмерных продуктов CRISPR I, CRISPR II-III кассет со специфическим зондом, комплементарным последовательности CRISPR повтора, выявил наличие в клетках *E. coli* K12 дикого типа сигнала, соответствующего РНК размером около 800 нт. Эта РНК являлась транскриптом CRISPR локуса, так как соответствующая полоса отсутствовала на дорожке с РНК, выделенной из *E. coli* BL21AI, не содержащих CRISPR-локусов. В сконструированном нами штамме *E. coli* K12 Δ CRISPR I, несущем делецию кассеты CRISPR I, интенсивность этой полосы была существенно снижена. Таким образом, большая часть полученного на нозерн-блоте сигнала соответствует полноразмерному транскрипту кассеты CRISPR I. Остаточный 800-нуклеотидный

продукт, детектируемый в Δ CRISPR1 клетках, возможно, соответствует частично процессированному транскрипту объединенного CRISPR II-III локуса. Полноразмерный продукт этого локуса по данным ОТ-ПЦР имеет размер не менее 1100 нт и, возможно, перекрывается с полосой 16S рРНК, что препятствует его обнаружению.

Полоса, соответствующая полноразмерному транскрипту CRISPR I, отсутствовала в клетках, гиперэкспрессирующих белок CasE (KEIO-мутант Δ casA). В этом штамме наблюдались различные продукты процессинга полноразмерного транскрипта длиной от 600-700 нт и меньше.

Резкое уменьшение количества полноразмерного транскрипта происходило также в штамме с делецией гена, кодирующего белок H-NS. Как было показано ранее (Pul *et al.*, 2010) этот белок является негативным регулятором оперона *cas*-генов, и в Δ hns клетках наблюдается увеличение количества всех Cas белков, в том числе и CasE.

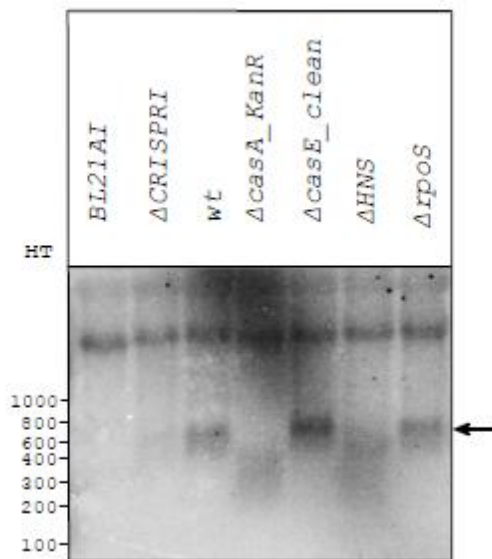


Рис. 8. Нозерн-блот, демонстрирующий размер и количества полноразмерных CRISPR транскриптов (отмечено стрелкой) в диком штамме и штаммах, несущих делеции различных генов. В качестве зонда использован меченый олигонуклеотид, комплементарный последовательности CRISPR повтора. Штамм *E. coli* BL21AI не имеет геномных CRISPR кассет и выступает в качестве отрицательного контроля.

В клетках, несущих делецию гена σ^s субъединицы РНК-полимеразы, необходимой для транскрипции генов на стационарной фазе роста (*ArpO*S), не было обнаружено уменьшения количества полноразмерного транскрипта. По-видимому, σ^s не участвует в транскрипции CRISPR каскеты.

В целом, полученные данные указывают на то, что увеличение количества коротких CRISPR РНК в клетках с повышенной экспрессией *casE* не связано с увеличением уровня транскрипции CRISPR касет, а обусловлено увеличенной скоростью процессинга. По-видимому, процессинг полноразмерного CRISPR транскрипта в клетках дикого типа происходит неэффективно, чем и объясняются низкие концентрации коротких CRISPR РНК и высокие концентрации предшественника. В клетках с повышенной экспрессией гена *casE* полноразмерная CRISPR РНК подвергается процессингу и перестает детектироваться, а количество короткой РНК (продукта процессинга) резко увеличивается.

Полученные нами данные были подтверждены в работе Westra и сопр., которые показали, что введение в клетку экстрахромосомной копии гена *leuO* (продукт которого является антагонистом H-NS) активирует транскрипцию *cas*-генов, в том числе и гена *casE*, что приводит к увеличению в клетке количества коротких РНК (Westra *et al.*, 2010).

Экспериментальная система для изучения встраивания новых CRISPR спейсеров.

Barrangou и соавторы в 2007 г. получили мутантных *S. thermophilus*, приобретших устойчивость к бактериофагу в ходе инфекции (Barrangou *et al.*, 2007). Ими было показано, что CRISPR касеты многих мутантов содержали вставки новых спейсеров, идентичные фрагментам генома фага, к которому приобреталась устойчивость. Удаление этих вставок приводило к возврату исходного чувствительного к фагу фенотипа. Введение спейсера в CRISPR касету чувствительных клеток генно-инженерными методами также приводило к возникновению устойчивости к фагу. Таким образом, наличие CRISPR спейсера, идентичного фрагменту генома бактериофага, являлось достаточным условием для возникновения устойчивости.

Попытки получить мутантов *E. coli* K12, устойчивых к бактериофагу за счет встраивания в CRISPR новых спейсеров, не принесли результата. Нами было сделано предположение, что CRISPR система *E. coli* неактивна из-за пониженного уровня экспрессии Cas белков. Для проверки этого предположения и разработки

метода анализа встраивания спейсеров у *E. coli*, нами была адаптирована система гиперэкспрессии CRISPR/Cas системы, использованная в работе Brouns *et al.*, 2008. Эта система состоит из трех совместимых индуцируемых плазмид. На двух плазмидах закодирован весь набор Cas белков *E. coli* K12 под контролем промотора РНК-полимеразы бактериофага T7 (*casABCDE12* оперон на плазмидном векторе pCDF (Novagen), и *cas3* ген – на pRSF (Novagen)), а на третьей плазмиде, на основе вектора pACYC-Duet (Novagen), - искусственная CRISPR кассета из семи спейсеров и восьми повторов. В качестве одного из спейсеров нами был введен участок ДНК, идентичный фрагменту гена 8 бактериофага M13 (этот спейсер получил название G8). Клетки *E. coli*, содержащие полный набор плазмид (*cas*-гены и CRISPR кассета с G8 спейсером) были высоко устойчивы к фагу M13. Замена спейсера G8 на спейсер, последовательность которого не встречается в геноме M13 (NT, non-targeting), приводила к возвращению чувствительного к фагу фенотипа. Клетки, содержащие полный набор *cas*-генов и CRISPR плазмиду с G8 или NT спейсером, заражались бактериофагом M13, инкубировались в жидкой среде в течение 8 часов, а затем рассеивались на твердую среду. Полученные индивидуальные колонии тестировались с помощью ПЦР с праймерами, специфическими к CRISPR кассетам I, II и искусственной CRISPR кассете, находящейся на плазмиде. Встраивание нового спейсера (и, соответственно, повтора) должно сопровождаться удлинением кассеты на ≈ 60 пн, что легко детектируется с помощью электрофореза в агарозном геле. Полученные результаты представлены в Табл. 1.

Тип клеток	CRISPR I	CRISPR II	CRISPR плаزمида
NT	0/100	0/100	0/100
G8	4/100	22/100	0/100

Табл. 1. Количество новых вставок (на 100 проверенных клонов) в геномные кассеты CRISPR I и CRISPR II, а также искусственную CRISPR кассету, расположенную на плазмиде. Наличие новых вставок проверялось в клетках G8 (содержащих спейсер, идентичный участку гена 8 фага M13) или NT клетках (содержащих спейсер, последовательность которого не встречается в геноме M13).

Встраивание новых спейсеров с частотой до 25% происходило *только* в клетках, содержащих плазмиду со спейсером G8 (клетки G8). 85% новых вставок обнаруживалось в CRISPR II кассете, 15% - в CRISPR I кассете; в одном из клонов вставки присутствовали в обеих кассетах. Нам не удалось обнаружить вставки в искусственную CRISPR кассету, расположенную на CRISPR плазмиде. Последовательности всех вставок были идентичны участкам генома фага M13 (Рис. 9). Встраивание происходило в ближний к лидерной последовательности конец кассеты. Большинство клонов содержало только один новый спейсер (и соответствующий повтор), хотя в одном случае было показано встраивание сразу трех новых спейсеров (и соответствующего числа повторов).

В результате нескольких повторных экспериментов нами было получено 75 клонов, содержащих новые спейсеры в CRISPR кассетах; распределение соответствующих протоспейсеров на карте генома фага M13 представлено на Рис. 9. Последовательности большинства спейсеров происходят из кодирующей цепи.

Чтобы показать, что встраивание нового спейсера приводит к возникновению устойчивости к фагу, несколько клонов, содержащих новые спейсеры, были заражены мутантными фагами M13. Мутантные фаги несут однонуклеотидные замены в протоспейсере G8, что позволяет этим фагам заражать G8 клетки, являющиеся устойчивыми к дикому типу фага M13 (Semenova *et al.*, 2011). Клетки, несущие новые спейсеры, идентичные различным участкам генома M13, как в транскрибируемой, так и в нетранскрибируемой цепях репликативной формы фага, оказались устойчивыми к мутантным фагам.

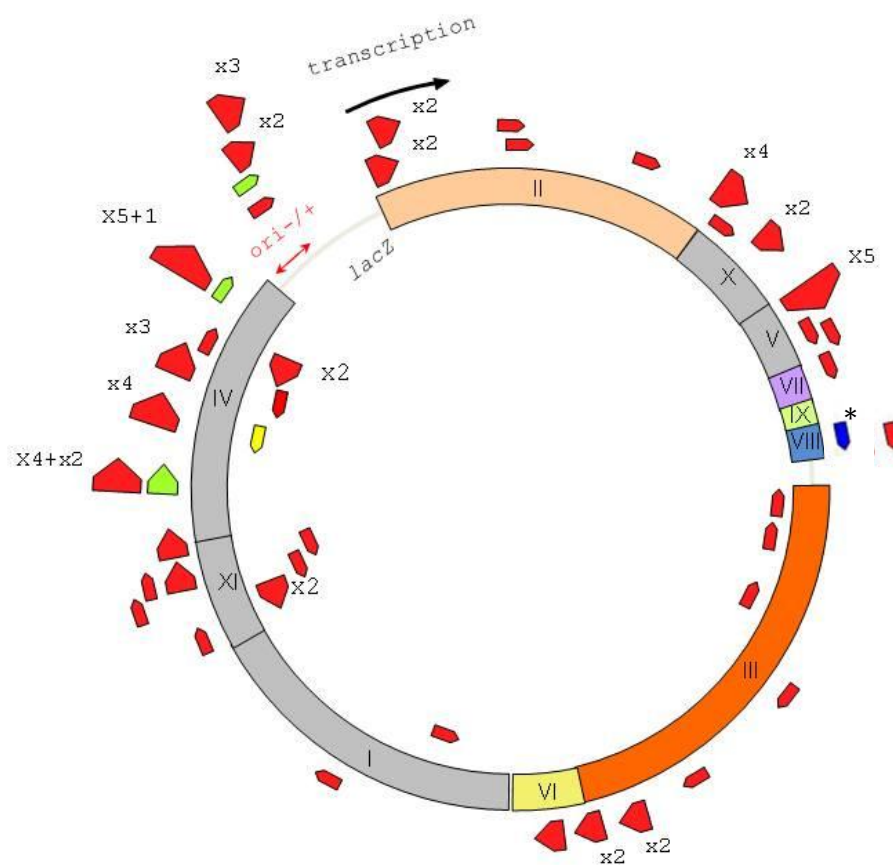


Рис. 9. Карта генома фага M13 с отмеченными на ней протоспейсерами, идентичными найденным новым спейсерам. Ширина стрелки отражает встречаемость данного спейсера в выборке из 75 клонов. Спейсеры, соответствующие серым протоспейсерам, встроились в CRISPR II кассету, светло-серым – в CRISPR I. Звездочкой обозначен протоспейсер, соответствующий исходному – G8 – спейсеру. Также указано направление транскрипции генома (transcription) и расположение ориджина (ori) репликации.

Определение последовательности PAM для *E. coli* K12.

В работах Н. Deveau и Р. Horwath было показано, что встраивание нового спейсера в CRISPR кассету *S. thermophilus* происходит из участков фагового генома, прилегающих к последовательности AGAAW (Deveau *et al.*, 2008; Horwath *et al.*, 2008). Мутантные фаги, имеющие точечные замены в этой последовательности, способны заражать клетки, CRISPR кассеты которых содержат спейсеры, в точности соответствующие протоспейсеру фага, то есть клетки устойчивые к дикому типу фага. Последовательности, примыкающие к

протоспейсерам, получили название PAM (protospacer adjacent motif). Для разных CRISPR систем последовательности PAM отличаются. На основании биоинформатического анализа было сделано предположение, что PAM мотивом у *E. coli* является последовательность AWG, однако полученные нами данные выявили преобладание (96%) мотива AAG, несмотря на то, что частоты встречаемости тринуклеотидов AAG и ATG в геноме фага M13 примерно одинаковые. Спейсеры, которым предшествуют как AAG так и ATG мотивы одинаково хорошо обеспечивают защиту клетки от фага, и предпочтение AAG на стадии адаптации, говорит о том, что требования к PAM на стадиях адаптации и интерференции отличаются.

Частота встречаемости тринуклеотида AAG на кодирующей цепи примерно в два раза меньше, чем на некодирующей (1:2). Это значит, что происхождение новых спейсеров преимущественно из кодирующей цепи не может быть объяснено большим по сравнению с некодирующей цепью числом AAG мотивов в кодирующей цепи.

Короткая CRISPR РНК, соответствующая спейсеру G8, комплементарна кодирующей цепи ДНК фага, и, возможно, именно ориентация короткой РНК определяет выбор той цепи, из которой произойдет новый CRISPR спейсер. Чтобы проверить это предположение была создана CRISPR плаزمид, в которую спейсер G8 был клонирован в обратной ориентации (то есть короткая РНК G8 в данном случае была комплементарна некодирующей цепи фага M13). Встраивание новых спейсеров, как и в случае прямой ориентации (см. выше), происходило только в клетках G8, и эффективность его была также около 20%. Однако встроенные спейсеры происходили преимущественно из некодирующей цепи.

Определение набора Cas белков, необходимых для встраивания новых спейсеров.

Для определения Cas белков, необходимых для встраивания новых спейсеров в CRISPR кассеты, был создан набор плазмид на основе исходных *casABCDE12* и *cas3* плазмид, но содержащих мутации со сдвигом рамки в индивидуальных *cas*-генах. Был также создан штамм с делецией всех геномных копий *cas*-генов (Δcas

штамм), наличие продуктов которых могло бы повлиять на результат эксперимента.

Δcas клетки, содержащие набор *cas*-плазмид с мутацией в том или ином *cas* гене, и плазмиду со спейсером G8 в искусственной CRISPR кассете, заражались бактериофагом M13 и выжившие в результате инфекции клетки проверялись на наличие новых вставок в CRISPR II кассету. Выбор кассеты CRISPR II был обусловлен тем, что встраивание спейсеров в нее идет с большей эффективностью, чем в CRISPR I.

Набор <i>cas</i> генов	CRISPR II	Устойчивость к фагу
<i>casABCDE123</i> (wt)	16/100	+
<i>casABCDE123</i>	0/100	-
<i>casBCDE123</i>	0/100	-
<i>casABCCDE123</i>	0/100	-
<i>casABCDE123</i>	0/100	-
<i>casABCDEE123</i>	0/100	-
<i>casABCDE123</i>	0/100	+
<i>casABCDE123</i>	0/100	+
<i>casABCDE123</i>	0/100	+
<i>casABCDE123</i>	1/100	-

Табл. 2. Количество новых вставок в кассету CRISPR II (на 100 проверенных клонов) в штаммах, содержащих мутации со сдвигом рамки в определенных *cas*-генах (мутантный ген обозначен полужирным шрифтом). Также указано наличие/отсутствие (+/-) у штамма устойчивости к фагу M13.

В работе Brouns и сотр. было показано, что из всех Cas белков только CasA, CasB, CasC, CasD, CasE и Cas3 необходимы для защиты клеток от фага, что было

также подтверждено в нашей работе. Однако для встраивания новых спейсеров, как показали полученные нами результаты, необходимы все вышеперечисленные Cas белки, а также Cas1 и Cas2 (см. Табл. 2).

На клетках G8, не несущих ген *cas3* ($\Delta cas3$), был получен только один клон, содержащий новую вставку. Последовательность нового спейсера была идентична гену *exoV* генома *E. coli* K12. Таким образом, в отсутствие *cas3* произошло встраивание геномного участка в CRISPR кассету, что должно сопровождаться появлением аутоиммунитета. Для проверки этого предположения, мы трансформировали клетки $\Delta cas3$, содержащие *exoV* спейсер, экспрессионной плазмидой, несущей ген *cas3*. Как видно из Рис. 10, нам не удалось получить соответствующих трансформантов, не смотря на то, что клетки $\Delta cas3$, не содержащие спейсера, трансформировались *cas3* плазмидой с высокой эффективностью. Таким образом, Cas3 не необходим для встраивания нового спейсера (так как встраивание происходило и в его отсутствие). Однако Cas3, по-видимому, участвует в разрезании чужеродной ДНК в процессе CRISPR иммунного ответа. Встраивание в CRISPR кассету спейсера, нацеленного на собственную ДНК клетки, при наличии всех Cas белков приводит к «самоубийству» клетки. Однако в отсутствие Cas3 этого не происходит и клетки, несущие такой спейсер, выживают. В подтверждение этого предположения недавно было показано, что Cas3 является нуклеазой *in vitro* (Sinkunas *et al.*, 2011).

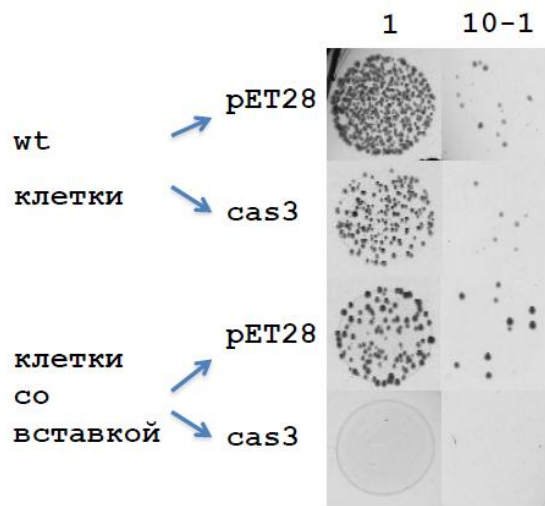


Рис. 10. Результаты трансформации плазмидой, несущий ген *cas3*, клеток дикого типа (wt) и клеток со спейсером, идентичным гену *exoV* *E. coli* K12. В качестве контроля на компетентность клеток был использован «пустой» вектор pET28a+. Цифрами 1 и 10-1 обозначены разведения клеток (неразведенные и разведенные в 10 раз соответственно).

Встраивание новых спейсеров происходит в клетках с делецией гена *hns*

Чтобы выяснить, не являются ли полученные нами данные по встраиванию новых спейсеров артефактом, вызванным гиперэкспрессией Cas белков и CRISPR кассеты, был сконструирован штамм, несущий G8 спейсер непосредственно в клеточном геноме – в CRISPR I кассете (см. Рис. 11) В этом штамме был также делетирован ген *hns*, кодирующий репрессор *cas*-генов. Клетки *G8Δhns* были устойчивы к бактериофагу M13, тогда как клетки, содержащие только спейсер G8 в CRISPR I (клетки G8) или клетки несущие только делецию гена *hns* (клетки *Δhns*), были чувствительны к фагу.

Клетки всех трех типов (*G8*, *Δhns* и *G8Δhns*) заражались бактериофагом M13, инкубировались в жидкой среде в течение 8 часов, а затем рассеивались на твердую среду. Полученные индивидуальные колонии тестировались с помощью ПЦР с праймерами, специфическими к CRISPR кассетам I, II. Полученные нами данные представлены в Табл. 4. Встраивание новых спейсеров в геномные CRISPR кассеты происходило только в клетках *G8Δhns*, хотя эффективность встраивания была существенно ниже, чем для использованной ранее системы, где компоненты CRISPR/Cas системы были закодированы на плаزمиде. Все новые спейсеры также происходили из генома фага M13 и встраивались в ближний к лидеру конец кассеты. Последовательность PAM, как и в случае искусственной системы гиперэкспрессии CRISPR/Cas, представляла собой AAG.

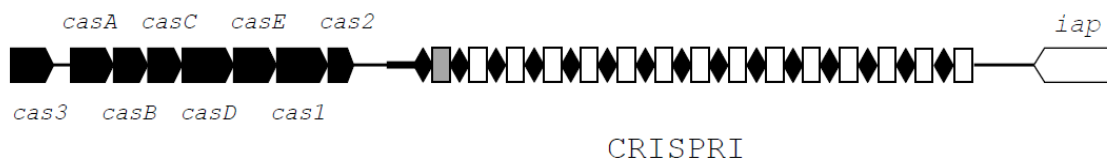


Рис. 11. Схематическое изображение CRISPR I кассеты. Серым закрашен спейсер G8, искусственно встроенный в кассету со стороны лидерной последовательности (указана утолщенной черной линией).

Штамм	CRISPR I	CRISPR II
Δhns	0/100	0/100
G8	0/100	0/100
G8 Δhns	1/100	5/100

Табл. 3. Количество новых вставок в геномные кассеты CRISPR I и CRISPR II в штамме, несущем делецию гена *hns* (Δhns); штамме, несущем геномные спейсер G8 (G8); и в штамме, несущем как делецию гена *hns*, так и геномный спейсер G8 (G8 Δhns).

Таким образом, данные по встраиванию новых спейсеров в клетки G8 Δhns в целом повторяют результаты, полученные для искусственной системы гиперэкспрессии CRISPR/Cas компонентов, из чего можно заключить, что выявленные нами закономерности являются скорее общим правилом для процесса адаптации в *E. coli*, нежели исключением.

ВЫВОДЫ:

1. Впервые показано, что в клетках *Escherichia coli* K12 присутствуют короткие РНК (≈ 60 нт), соответствующие всем спейсерам CRISPR кассет I, II, III. Относительное количество коротких CRISPR РНК не зависит от положения соответствующих спейсеров в кассете.
2. Увеличенный уровень экспрессии белка CasE необходим и достаточен для накопления коротких CRISPR РНК в клетках *E. coli* K12.
3. Встраивание новых спейсеров с высокой эффективностью происходит только в клетках уже несущих CRISPR спейсер, идентичный по последовательности участку генома фага. Для встраивания новых спейсеров в CRISPR кассеты *E. coli* K12 также необходим PAM мотив AAG.
4. Ориентация вновь встраиваемых спейсеров зависит от ориентации предшествующего спейсера, идентичного фрагменту геномной ДНК инфицирующего фага.
5. Встраивание нового спейсера в CRISPR кассету не происходит в отсутствие любого из Cas белков.

Список печатных работ, опубликованных по теме диссертации

Статьи

1. **Pougach, K.**, Semenova, E., Bogdanova, E., Datsenko, K.A., Djordjevic, M., Wanner, B.L., Severinov, K. (2010). Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol Microbiol*, 77(6), p. 1367-1379, 2010

Тезисы конференций

1. **Pougach, K.**, E. Semenova, K. Datsenko, Severinov, K. (2011) Characterization of New CRISPR Spacer Acquisition Process in *Escherichia coli*. 111th General Meeting American Society for Microbiology, May, 21-24, New Orleans, LA, USA, p.114
2. Semenova, E., **Pougach, K.**, Severinov K. (2010) The use of bacteriophage M13 to study CRISPR/Cas function in *E. coli*. International meeting “3rd CRISPR Research Conference”, July, 22-23, Berkley, CA, USA. p.11
3. **Pougach, K.**, Severinov, K. (2009). Systematic analysis of CRISPR-cassette transcription and regulation in *E. coli*. International meeting “Regulatory RNA in prokaryotes”, June, 3-6, 2009, Berlin, Germany, p.147