

Сколковский институт науки и технологий

Центр наук о жизни

Московская школа на Юго-Западе №1543

**Роль комплекса RecBCD в процессе CRISPR-адаптации
у *Escherichia coli***



М. Зубкова

В. Лабор

Научные руководители:

к.б.н. Е.Е. Савицкая

А.О. Шкарута

Москва

2019

Оглавление

Введение.....	3
Литературный обзор.....	4
Материалы и методы.....	9
Результаты.....	13
Обсуждение.....	20
Выводы.....	21
Благодарности.....	22
Ссылки.....	22

Введение

В геноме 90% архей и 50% бактерий обнаружен специфический локус CRISPR и, как правило, ассоциированный с ним блок генов *cas* (от англ. CRISPR-associated). В настоящее время известно, что CRISPR-Cas – это особая иммунная система, направленная на защиту клетки хозяина от чужеродной ДНК бактериофагов и плазмид. CRISPR-локус (от англ. clustered regularly interspaced short palindromic repeats) состоит из чередующихся коротких повторов и уникальных спейсеров. Основой для спейсеров может служить, в первую очередь ДНК бактериофагов или плазмиды, реже собственная геномная ДНК бактерии. В основе механизма CRISPR-Cas иммунитета лежат 2 процесса: CRISPR-адаптация и CRISPR-интерференция (рис. 1). Пополнение CRISPR-кассеты новыми спейсерами происходит за счёт захвата фрагмента ДНК и его встраивания в CRISPR-кассету – этот процесс получил название «CRISPR-адаптация». С CRISPR-кассеты считываются первичные транскрипты, которые процессируются с образованием крРНК (от англ. CRISPR-RNA). Каждая крРНК содержит один спейсер, необходимый для распознавания комплементарных последовательностей в составе плазмиды или генома вируса. Распознавание происходит посредством эффекторного комплекса, состоящего из белков Cas и крРНК, и далее ведет к деградации чужеродной ДНК в процессе CRISPR-интерференции. CRISPR-адаптация может протекать независимо от CRISPR-интерференции. У *Escherichia coli* (*E. coli*), имеющей CRISPR-Cas систему, принадлежащую к типу I-E, для CRISPR-адаптации необходим белковый комплекс, состоящий только из белков Cas1 и Cas2 [1]. Тем не менее, ранее были получены данные о роли комплекса RecBCD, отвечающего за репарацию двухцепочечных разрывов в клетке, в CRISPR-адаптации у *E. coli*, но детали механизма установлены не были [2].

Целью данной работы является изучение того, каким образом комплекс RecBCD влияет на CRISPR-адаптацию у *E. coli*. Для достижения этой цели мы поставили задачи получить штамм с делециями генов *recB*, *recC* и *recD*, трансформировать его плазмидами, кодирующими варианты RecBCD комплекса с измененными активностями, и оценить скорость роста клеток и эффективность CRISPR-адаптации в этих клетках.

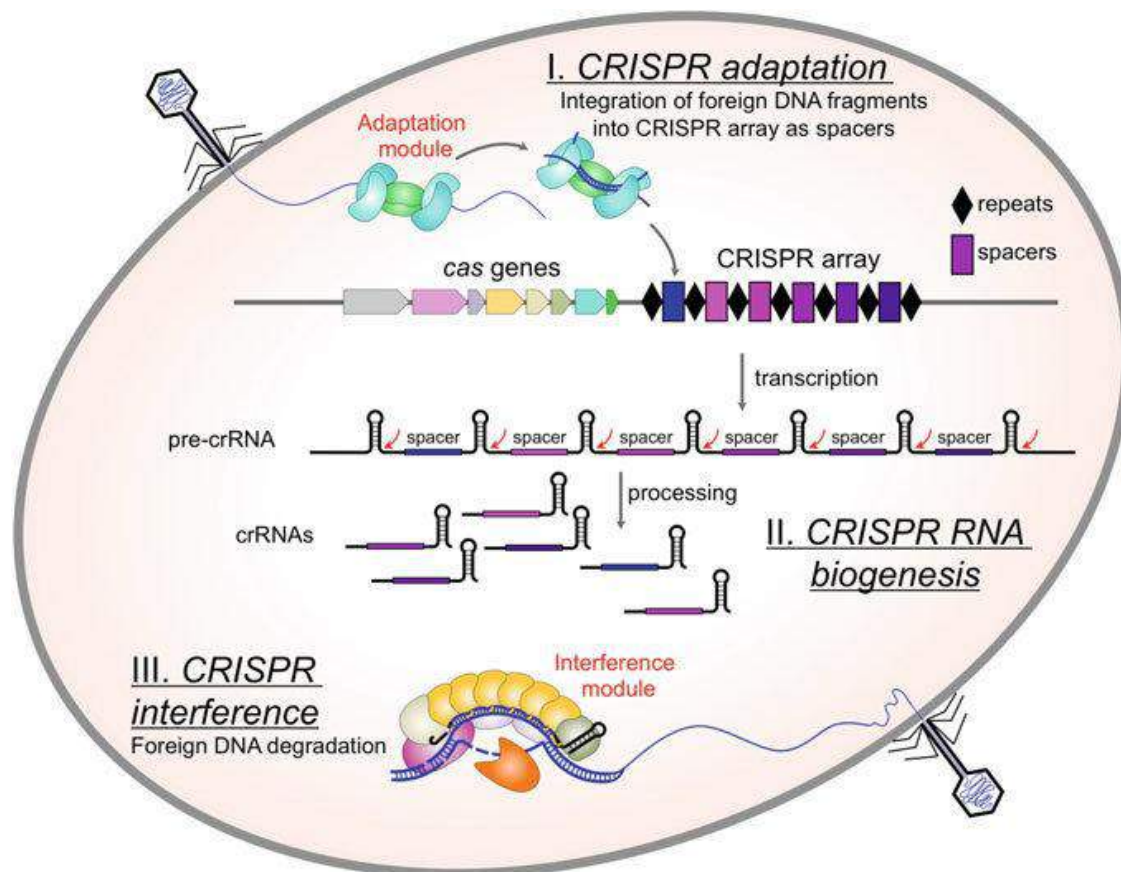


Рисунок 1. (E. Semenova & K. Severinov, 2017 [3]). Схема работы CRISPR-Cas системы.

Литературный обзор

Ещё в 1987 году [4] было замечено, что в геноме у *E. coli* есть короткие повторяющиеся участки, чередующиеся с одинаковыми по длине, но различными по последовательности, спейсерами. Данные локусы позднее получили название CRISPR. Похожие районы были обнаружены также в геноме археи *Haloferax mediterranei*, однако их функция оставалась непонятной [5-6]. После обнаружения аналогичных локусов в геномах различных бактерий было замечено, что некоторые спейсеры комплементарны участкам геномов бактериофагов или плазмидам, поэтому было выдвинуто предположение о защитной функции CRISPR-локусов против чужеродных генетических элементов [7-8]. С CRISPR-локусами часто ассоциированы гены, получившие название *cas*-генов. На основании генов *cas* в настоящее время CRISPR-Cas системы разделяются на 6 типов, которые объединены в 2 класса [9]. Для второго класса (типы II, V и VI) характерно наличие одного многофункционального белка Cas, участвующего в процессе CRISPR-интерференции. В системах первого класса (типы I, III и IV) в CRISPR-интерференции задействован комплекс, состоящий из нескольких Cas-белков. CRISPR-Cas система *E. coli* относится к типу I-E. Для него (как и для всего типа I) характерно наличие белка Cas3. В CRISPR-интерференции у *E. coli* участвуют белки Cse1, Cse2, Cas7, Cas5, Cas6, которые входят в состав комплекса Cascade, и хеликаза-нуклеаза Cas3. На начальном этапе интерференции с CRISPR-кассеты считывается первичный транскрипт, который

после созревания и процессинга дает начало коротким крРНК, содержащим в себе один спейсер, окруженный фрагментами повтора. Комплекс Cascade вместе со зрелой крРНК формирует эффекторный комплекс, который связывается с ДНК-мишенью, комплементарной спейсеру (протоспейсером) и содержащей последовательность PAM (от англ. protospacer adjacent motif). PAM – короткие последовательности, фланкирующие протоспейсер с 5'-конца, которые есть в ДНК вирусов или плазмид, но которых нет в CRISPR-кассете для предотвращения аутоиммунного воздействия. Часто это консервативные последовательности, например, у типа I-E у *E. coli* есть пять возможных PAM-последовательностей: ATG, AAG, AGG, GAG, TAG. Образование комплекса с протоспейсером в составе мишени привлекает белок Cas3, который разрезает мишень, хотя детальный механизм деградации *in vivo* остается не до конца понятным.

Для CRISPR-Cas систем типа I описано две разновидности CRISPR-адаптации: праймированная и наивная. В отличие от наивной, для которой достаточно присутствия комплекса Cas1-Cas2, праймированная адаптация сопряжена с интерференцией и направлена на уже известного чужеродного агента, против которого у клетки уже есть иммунность. При наивной адаптации спейсеры могут встраиваться из любых фрагментов ДНК, случайно образовавшихся в клетке, например, при репарации разрывов или репликации. Cas1 и Cas2 – нуклеазы, но только нуклеазная активность Cas1 важна для адаптации *in vivo*. В составе адаптационного комплекса димер Cas2 заключен между двумя димерами Cas1 (рис.2а). Неизвестно, что является субстратом для CRISPR-адаптации, возможно, это фрагменты ДНК, присутствующие в клетке. При этом важно узнавание и встраивание фрагмента, имеющего фланкирующий мотив PAM, так как только такие спейсеры впоследствии будут активны в процессе CRISPR-интерференции. После обнаружения подходящего фрагмента и его процессинга до участка длиной 33 п.н. без PAM (если он был) (рис.2b) адаптационный комплекс с помощью белка IHF встраивает его между повтором первого спейсера и лидерной, промоторной, областью CRISPR-касеты. Одноцепочечные участки достраиваются, удваивая при этом CRISPR-повтор (рис.2с).

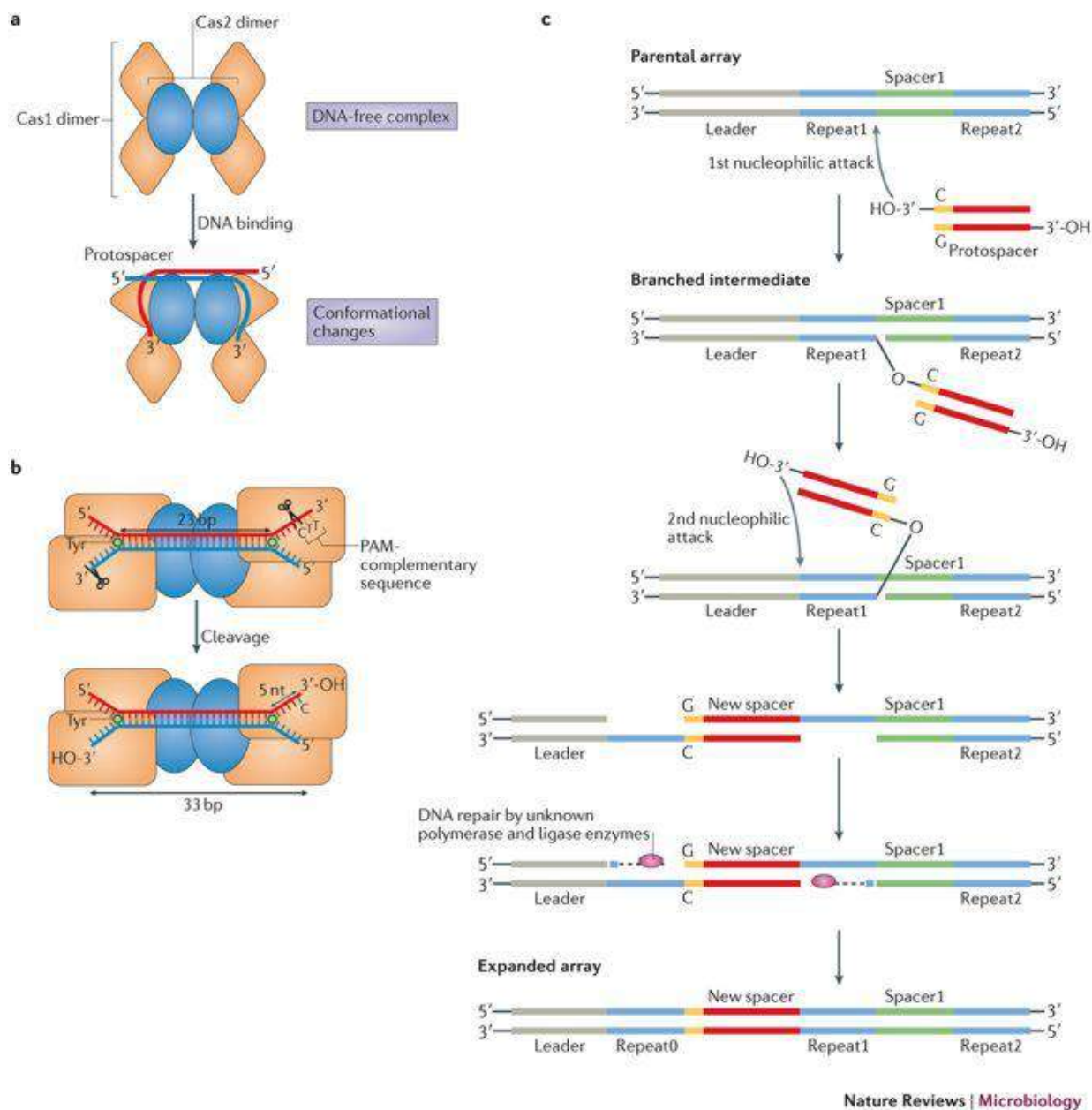


Рисунок 2. (G. Amitai & R. Sorek 2016 [10]). CRISPR-адаптация. а) строение адаптационного комплекса Cas1-Cas2 у *E. coli*, б) предполагаемый механизм процессинга пре-спейсера комплексом Cas1-Cas2, в) встраивание спейсера в CRISPR-кассету с удваиванием повтора.

Комплекс RecBCD отвечает в клетке за репарацию двуцепочечных разрывов и гомологичную рекомбинацию [11]. Он включает в себя три субъединицы: RecB, RecC и RecD (рис.3а). Субъединица RecB состоит из хеликазного и нуклеазного домена, отделённого от первого «мостиком» длиной в 19 аминокислот. RecD – также хеликаза, но RecD движется от 5' к 3'-концу, а RecB, наоборот, от 3' к 5'. RecC может узнавать специфическую последовательность ДНК – 5'-GCTGGTGG-3', названную «Chi-сайт». Во время работы комплекса нуклеазный домен RecB может действовать по-разному в зависимости от концентрации в клетке ионов Mg^{2+} и молекул АТФ (аденозинтрифосфат) (рис.3б). При избытке АТФ комплекс RecBCD садится в месте двуцепочечного разрыва и, двигаясь по цепи,

расплетает её. Когда RecC распознаёт Chi-сайт, то нуклеазный домен RecB делает надрез в 5'-3'-цепи. После комплекс продолжает своё движение с расплетанием цепи ДНК. На получившийся свободный 3'-конец садиться белок RecA, способствующий гомологичной рекомбинации с неповрежденным участком ДНК. При избытке ионов Mg^{2+} нуклеазный домен RecB постоянно активен. Двигаясь по цепи ДНК, RecB разрушает оба конца, но с разной частотой (5'-конец режет). После узнавания субъединицей RecC Chi-сайта комплекс продолжает двигаться, но RecB перестаёт разрушать 3'-конец. На оставшийся свободный 3'-конец, как и в первом варианте, садится RecA. Можно предположить, что получившиеся фрагменты могут быть использованы комплексом Cas1-Cas2, как субстрат для встраивания спейсера [2]. Было показано, что активность комплекса RecBCD существенно стимулирует наивную CRISPR-адаптацию [12], но остаётся непонятно, каким образом это происходит.

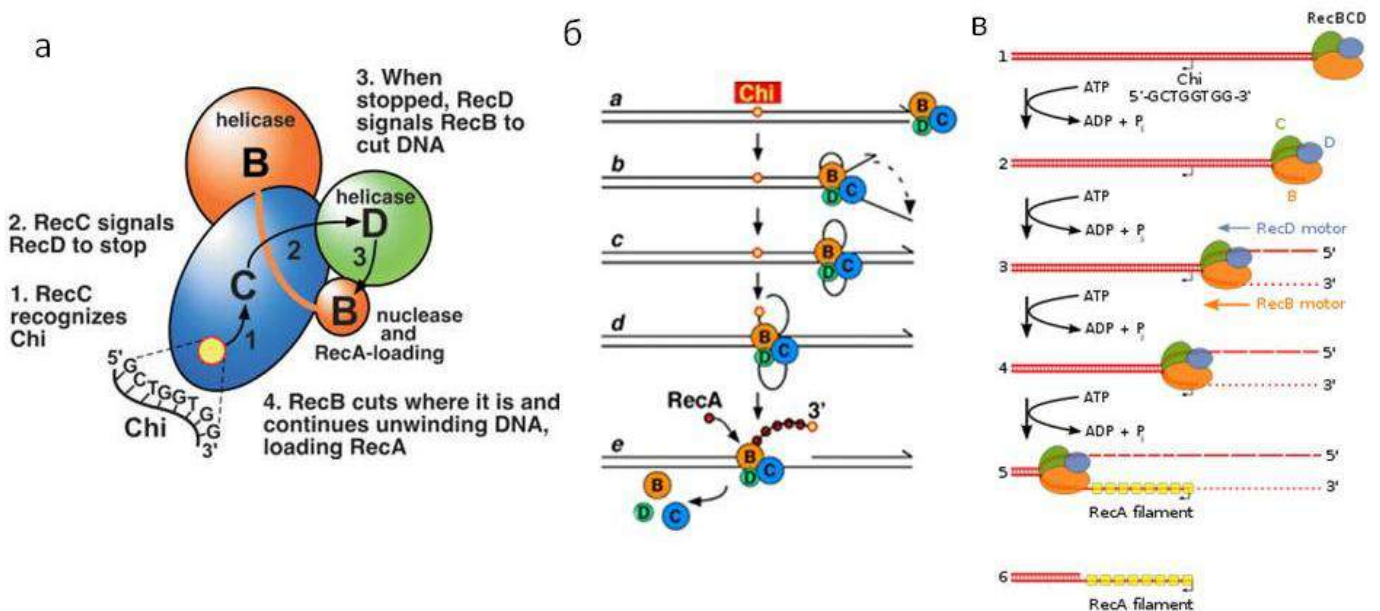
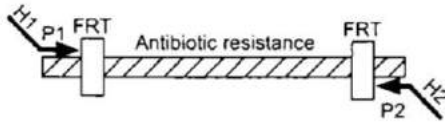


Рисунок 3. (а) Amundsen et al. [13], б) Amundsen, Sharp & Smith [14]). а) Строение комплекса RecBCD, б) механизм действия комплекса RecBCD при избытке молекул АТФ, в) при избытке ионов Mg^{2+} .

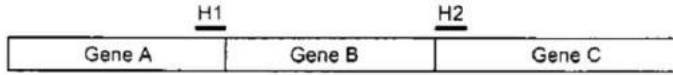
λ -Red-рекомбинация

Одним из самых распространенных методов внесения направленных изменений в геном *E. coli* является λ -Red-рекомбинация [14-15] (рис. 4).

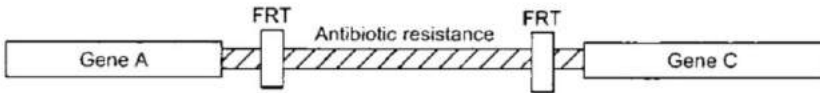
Step 1. PCR amplify FRT-flanked resistance gene



Step 2. Transform strain expressing λ Red recombinase



Step 3. Select antibiotic-resistant transformants



Step 4. Eliminate resistance cassette using a FLP expression plasmid

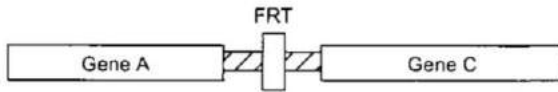


Рисунок 4. (К. А. Datsenko & В. L. Wanner [15]) Схема λ -Red-рекомбинации.

λ -Red-рекомбинация основана на использовании белков Gam (γ), Beta (β) и Eho (*exo*), относящихся к системе Red-рекомбинации фага λ . Gam поддерживает ДНК с двуцепочечным разрывом и защищает ее концы от разрушения. Beta способствует отжигу двух комплементарных участков ДНК, а Eho обладает экзонуклеазной активностью и создаёт «липкие» концы для рекомбинации. Для введения модификаций в геном нужно трансформировать клетки бактерий плазмидой, содержащей гены γ , β и *exo*. Также в клетки вводится ДНК-фрагмент, который должен включать в себя ген резистентности к антибиотику, окруженный участками, гомологичными участкам места генома, в которое надо внести изменение (рис. 4). При индукции экспрессии генов λ -Red-рекомбинирования запускается процесс рекомбинации: Eho садится на ДНК-фрагмент и, двигаясь от 5'-конца к 3'-концу, разрушает одну из цепей ДНК, на образовавшиеся «липкие» концы садится Beta, которая меняет местами гены между ДНК фрагментом и участком генома по участкам гомологии. Всё это время Gam предотвращает разрушение ДНК фрагмента. В итоге на месте заменяемого участка появляется ген устойчивости к антибиотику, который позволяет провести отбор нужных клонов. Часто дополнительно ген устойчивости к антибиотику окружают FRT-сайтами (от англ. *flippase (FLP) recognition target*). С помощью плазмиды с геном FLP-рекомбиназы, можно вырезать селективный ген, что позволит в дальнейшем провести в полученном штамме новый раунд рекомбинации.

Данный метод был использован в данной работе для делеции генов *recB* и *recD*.

Материалы и методы

Штаммы и плазмиды

Клетки выращивались в 10 мл жидкой среды LB (1% Bacto tryptone, 0,5% Bacto yeast extract, 1% NaCl) или на чашках Петри с 15-20 мл среды LA (LB+1,4% агара) с добавлением необходимых антибиотиков в финальных концентрациях: канамицин – 50 мкг/мл, ампициллин – 50 мкг/мл, стрептомицин – 50 мкг/мл. Для активации экспрессии генов изопропил- β -D-1-тиогалактопиранозид (IPTG) добавляли до финальной концентрации 0,1 мМ. Питательная среда SOC для посттрансформационных клеток готовилась путем добавления 0,4% глюкозы к SOB (2% Bacto tryptone, 0,5% Bacto yeast extract, 10 мМ NaCl, 2,5 мМ KCl, 10 мМ MgCl₂, 10 мМ MgSO₄).

Плазмида pKD4, несущая ген резистентности к канамицину, фланкированный FRT-сайтами (FLP-recognition target), использовалась как матрица для получения ПЦР-продукта для λ -Red-рекомбинации.

Плазмида pKD46 (GeneBank Accession number AY048746) несет гены λ -Red-системы γ , β и *exo*, кодирующие белки Gam, Beta и Exo, соответственно, под индуцируемым L-арабинозой промотором P_{araB}, а также ген резистентности к ампициллину. pKD46 имеет температурочувствительный ориджин репликации для легкого удаления плазмиды из полученных мутантов путем выращивания при температуре выше 30°C.

Плазмида pCas1+2 – производная вектора pCDF (GeneBank Accession number KM229384) с генами *cas1,2* под промотором T7 полимеразы (экспрессия которой в свою очередь индуцируется IPTG), и геном резистентности к стрептомицину.

Плазмиды pRecBCD: pJS24, pSA21, pSA618, pSA620, pSA622, pSA673 (подробнее в таблице 1) несут варианты генов *recB*, *recC*, *recD* и ген резистентности к ампициллину. Для контроля использовалась плазмида pSA607, несущая гены *recB*, *recC*, *recD* дикого типа.

Для получения мутанта с делецией генов *recB*, *recC*, *recD* использовался штамм с делецией гена *recC*, из которого был удален ген резистентности к канамицину путем FLP-рекомбинации [в больших деталях – 14]. Полученный штамм – производная BL21 (DE3) $\Delta recB$, $\Delta recC$, $\Delta recD$ KmR – был назван ATS1.

<u>Плазмида</u>	<u>Варианты <i>recBCD</i></u>	<u>Изменения в работе комплекса RecBCD</u>
pJS24	RecBC ^{S39E} D	Аминокислотная замена в участке туннеля RecC. Мутант теряет способность

		распознавать Chi-сайты [19].
pSA21	RecB ²¹ CD	Нефункциональный RecB. Мутантный комплекс теряет активность по отношению к Chi-сайтам и почти полностью способность к рекомбинации.
pSA673	RecB ^{Δ19} CD	Делеция 19 аминокислот, образующих связку между нуклеазным и хеликазным доменом RecB. Мутант теряет способность к рекомбинации и активность по отношению к Chi-сайтам, но нуклеазная активность остается незатронутой [16].
pSA618	RecB ^{K29Q} CD	Лизин в 29 позиции заменен на глутамин в консервативном АТФ-связывающем мотиве. Мутантный RecB теряет АТФ-зависимую моторную и частично нуклеазную активность [17,18]. Комплекс в целом сохраняет способность продвигаться и расплетать ДНК за счет субъединицы RecD, но вместо одноцепочечного хвоста на 3' конце образуется петля [18].
pSA620	RecBCD ^{K177Q}	Аналогичная замена лизина в 177 позиции на глутамин в АТФ-связывающем мотиве RecD. Хеликазная активность комплекса снижается на 80% [18].

pSA622	RecB ^{K29Q} CD ^{K177Q}	Мутант с обеими заменами в RecB и RecD, описанными выше. Комплекс полностью теряет АТФ-азную и хеликазную активность.
pSA607	RecBCD WT	Гены RecBCD дикого типа.

Таблица 1. Плазмиды с мутациями в recB, recC и recD, заменявшие отсутствующие гены у штамма AT51 с делецией соответствующего локуса.

Гель-электрофорез ДНК фрагментов

Для электрофореза использовался 1% или 2% (для проверки адаптации) агарозный гель и TBE-буфер (Трис-борат-ЭДТА) при 110 V. Образцы смешивались с 6-кратной краской TriTrack Loading Dye (Thermo Scientific) и разгонялись на геле в течение 30-75 мин в зависимости от плотности геля и длины разделяемых фрагментов.

ПЦР, ПЦР на колониях

Скрининг колоний мутантов проводился путем перенесения части колонии стерильным носиком для пипетки в смесь для ПЦР в 0,2 мл пробирках (общий объем 20 мкл: 2 мкл буфера 10X, 0,4 мкл dNTPs 10мМ, 0,8-1 мкл каждого праймера 0,5 мкМ, 0,4-0,5 мкл Taq-полимеразы). В амплификаторе использовалась следующая программа: 95°C/5' 30 циклов: | 95°C/15" 59°C/15" 72°C/90" | 72°C/5' 12°C/∞ .

Аmplифицируя CRISPR-кассету, можно детектировать встраивание новых спейсеров по увеличению длины ПЦР-продукта на длину спейсера и одного повтора. Для проверки адаптации ночная культура разбавлялась в 10 раз, затем 1 мкл каждого образца добавлялся в смесь для ПЦР, приготовленную так же, как описано выше. Использовались праймеры на CRISPR-кассету: F: ATTTTGCGTTTCGTTTCAGGT, R: TGGATGTGTTGTTTGTGTGATACTA. В амплификаторе использовалась следующая программа: 95°C/3' 34 цикла: | 95°C/30", 55°C/30", 72°C/40" | 72°C/5' 12°C/∞ .

Химическая трансформация

Химически компетентные клетки готовились следующим образом: из ночной культуры готовилась субкультура путем добавления 100 мкл клеток в 10 мл LB. Субкультура выращивалась в шейкере

при 37°C 3-5 часов. При достижении $OD_{600} \approx 0.4-0.55$ клетки центрифугировались на 3000 gcf 10 мин при 4°C, супернатант сливался, и остатки среды удалялись из пробирки. Клетки ресуспендировали в 10 мл ледяного раствора 0,1 М CaCl₂, инкубировали во льду в течение 10-20 мин, затем центрифугировали снова при тех же условиях. Супернатант сливался, и клетки были ресуспендированы в 2-5 мл ледяного раствора 0,1 М CaCl₂ + 15% глицерола, распределены по пробиркам по 300 мкл и хранились при -80°C.

Для химической трансформации 10 нг рKD46 добавляли к 100 мкл компетентных клеток во льду. Клетки с рKD46 инкубировались во льду в течение 30 мин, затем были перенесены в термостат, нагретый до 42°C ровно на 1,5 мин. Из термостата трансформированные клетки были немедленно перенесены в лед, выдержаны 2-10 мин во льду, и к ним было добавлено 0,9 мл SOC, после чего клетки инкубировались в шейкере при 30°C (чтобы обеспечить репликацию плазмиды рKD46 с температурочувствительным ориджином) в течение 40 мин – 1 часа, а затем были центрифугированы на 4500 об/мин 10 мин при 4°C, ресуспендированы в 1/10 объема и высеяны на чашки с ампициллином. Чашки инкубировались также при 30°C.

Электропорация

Для получения электрокомпетентных клеток культура реципиентов при $OD_{600} \approx 0,6-0,8$ была отцентрифугирована на 4500 об/мин 10 мин при 4°C, затем три раза промыта в 1 мл ледяного 10% глицерола и ресуспендирована в 500 мкл 10% глицерола. Полученные электрокомпетентные клетки были трансформированы в предварительно охлажденных 0,1 см кюветках с помощью MicroPulser (Bio-Rad) со стандартными параметрами (1.8 kV, $t \approx 5,5-6$ мс) в соотношении 40 мкл клеток и 10-100 нг ПЦР-продукта или плазмиды. Трансформанты были немедленно перенесены в 1 мл SOC и инкубировались при 37°C (30°C, если клетки реципиентов несут рKD46) в течение ~1 часа, а затем были высеяны на чашки с необходимыми антибиотиками для отбора успешных трансформантов.

Индукция CRISPR-адаптации

Штамм AT51 ($\Delta RecBCD Km^R$) при помощи электропорации был трансформирован одной из семи плазмид с вариантами генов *recB*, *recC* и *recD* и плазмидой рCDF Cas1+2 или плазмидой рCDF без вставки с использованием SOC, предварительно нагретой до 37°C. Трансформанты инкубировались 1 ч 50 мин на 37°C, а затем были высажены на чашки с канамицином, ампициллином и стрептомицином для отбора успешных трансформантов по резистентности к антибиотикам и инкубировались на 37°C до появления колоний.

Детекция плотности культуры клеток в реальном времени

Образцы каждой реплики были загружены в лунки плашки для Plate Reader в трех технических репликах с добавлением индукторов (L-арабиноза и IPTG) и без добавления индукторов. В Plate Reader клетки росли в течение 20 часов при 37°C, измерение оптической плотности производилось каждые 10 мин, на основе средних значений оптической плотности по трем репликам образцов были построены кривые роста в Microsoft Excel.

По окончании роста одна из технических реплик каждого образца использовались для проверки адаптации при помощи ПЦР, как описано выше.

Результаты

Нокаут генов *recB* и *recD* методом λ -Red -рекомбинации (рис.5)

С помощью метода λ -Red -рекомбинации мы получили мутанта с нокаутом генов, кодирующих комплекс RecBCD. λ -Red -рекомбинация проводилась так, как было описано ранее [15], то есть включала в себя пять этапов:

1. Получение ПЦР-продукта, содержащего участок, на который заменяются целевые гены *recB* и *recD*, – ген резистентности к канамицину между FRT-сайтами – и гомологичные участки, соответствующие фланговым участкам генов *recB* и *recD*.
2. Химическая трансформация исходного штамма плазмидой pKD46, несущей гены λ -Red-системы.
3. Индукция генов λ -Red –рекомбинации с плазмиды pKD46.
4. Электропорация клеток с плазмидой pKD46 полученным ПЦР-продуктом.
5. Селекция и проверка мутантов. Для дизайна ПЦР-продукта использовалась программа SnapGene Viewer (<https://www.snapgene.com>). ПЦР-продукт длиной ~1,5 килобазы был амплифицирован с плазмиды pKD4 в качестве матрицы с использованием праймеров, содержащих гомологичные участки ко фланкам целевых генов длиной ~50 нуклеотидов

и участки отжига длиной ~20 нуклеотидов (F: CAAACAATGCCCTGATGAGTGAAAAGAATGAGTGATGTCGCCGAGACACTAGATCCTgtgtaggctggagctgctt; R: CGGTGCAAAGTGCTGGCATCTTCCGGAATGCGTTTCTTTGTTTCATCGGTCAGCGcatatgaatcctccttagtt, программа: 95°C/5' 25 циклов: | 95°C/15" 55°C/15" 72°C/60" | 72°C/5' 12°C/ ∞). ПЦР-продукт был выделен из геля при помощи GeneJET Gel Extraction Kit (Thermo Scientific) согласно рекомендациям производителя.

Клетки реципиентов были трансформированы pKD46 при помощи химической трансформации, как описано в материалах и методах. Успешные трансформанты были отобраны на среде с ампициллином и далее инкубировались на 30°C для обеспечения репликации pKD46. С ночной культуры (10 мл LB + Amp) выросших трансформантов была поставлена субкультура, в которую при достижении $OD_{600} \approx 0,1-0,15$ добавлялась L-арабиноза для индукции генов γ , β и *exo* с pKD46. Из

культуры с индуктором и ампициллином были получены электрокомпетентные клетки, которые затем были трансформированы полученным ПЦР-продуктом путем электропорации (подробнее в Материалах и методах) и инкубировались на 30°C. Полученные мутанты были отобраны по резистентности к канамицину. Для скрининга мутантов использовалась ПЦР на колониях с использованием праймеров *check1*: TTGTTTGTGCTGACCACGCG и *check2*: AACCAAGTACCAACAGCTCTG на фланки генов *recD* и *recC* (см. рис.5). Генотип мутантов был подтвержден при помощи секвенирования по Сэнгеру в компании Евроген. В результате был получен мутантный штамм ATS1 с нокаутом генов *recB*, *recC* и *recD*, который использовался в дальнейшей работе.

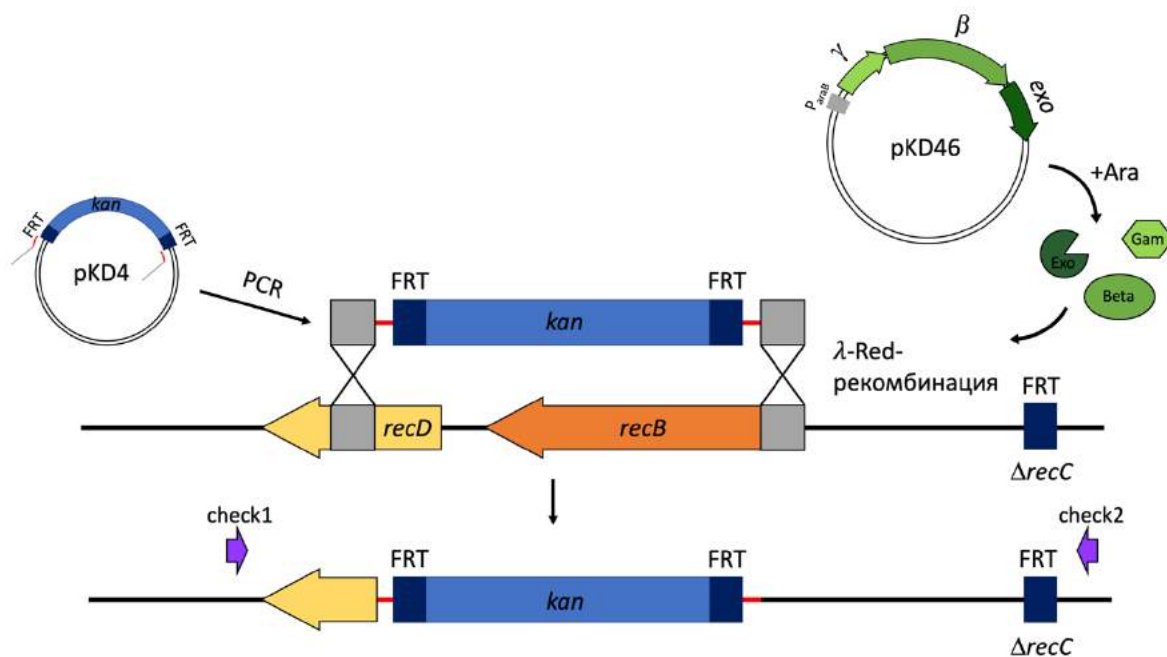


Рисунок 5. Схема нокаута генов *recB* и *recD* методом λ -Red-рекомбинации. Исходный штамм с нокаутом гена *recC* (в локусе, соответствующем *recC*, остался FRT-сайт (темно-синий) после удаления гена резистентности к канамицину при помощи FLP-рекомбиназы) был сначала трансформирован плазмидой, несущей гены λ -Red-системы – γ , β и *exo*. Затем гены на pKD46 индуцировались L-арабинозой (Ara), после чего клетки реципиента электропорировались ПЦР-продуктом, амплифицированным с pKD4, несущим ген резистентности к канамицину между FRT-сайтами (темно-синий). Праймеры состоят из участков отжига длиной 20 нуклеотидов (обозначены красным) и из участков, гомологичных фланговому участку *recB* и участку в середине *recD* (обозначены серым). В результате рекомбинации *recB* нокаутируется полностью, от *recD* остается бессмысленный не транскрибируемый кусок длиной ~ 1100 бп. Генотип мутантов был проверен при помощи ПЦР с использованием праймеров *check1* и *check2* на фланговые участки генов *recD* и *recC* (подробнее в тексте).

Схема эксперимента по индукции CRISPR-адаптации в клетках *E. coli*, несущих измененные варианты генов, кодирующих комплекс RecBCD (рис.6)

Для исследования роли комплекса RecBCD в клетки полученного штамма ATS1 вносились плазмиды с вариантами генов *recB*, *recC* и *recD*, по-разному влияющих на работу отдельных субъединиц и всего комплекса в целом (таблица 1). Клетки также трансформировали плазмидой pCDF Cas1+2, кодирующей белки адаптационного комплекса Cas1-Cas2 или пустой контрольной плазмидой pCDF. Эксперимент проводился далее по нижеописанной схеме в двух биологических повторах.

Были получены ночные культуры соответствующих трансформантов. После контрольной проверки количества клеток путем измерения OD₆₀₀ и

CRISPR-адаптации при росте в ночной культуре клетки

пересевали в 96-ти луночные плашки в трех технических повторах. Проводили выращивание клеток в отсутствие или в присутствии индуктора, обеспечивающего транскрипцию генов *cas1,2* с плазмиды pCDF Cas1+2, на приборе Plate Reader для отслеживания роста клеток в режиме реального времени.

После 20 часов культивирования образцы снова использовались для проверки адаптации методом ПЦР с последующим электрофорезом в 2% агарозном геле.

Для правильной интерпретации экспериментов в начале нами были оценены оптимальные условия экспрессии генов *cas1,2*, достаточные для детекции адаптационных событий, но не токсичные для клетки, а также сравнение скорости роста клеток, несущих различные варианты генов комплекса RecBCD, таким образом, чтобы определить временную точку, в которой плотности культур будут сопоставимы и, следовательно, не будут влиять на результаты анализа CRISPR-адаптации.

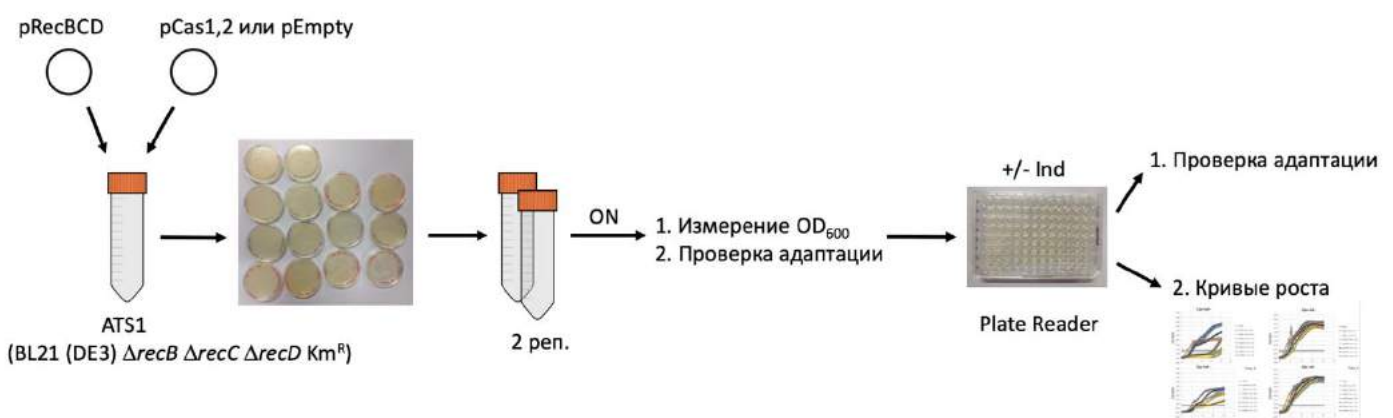


Рисунок 6. Схема эксперимента. Клетки штамма ATS1 с делецией генов *recB*, *recD*, *recC* были трансформированы плазмидой с набором генов, кодирующих варианты комплекса RecBCD с измененными активностями, и плазмидой с генами или без генов *cas1,2*. Трансформанты были отобраны на чашках с необходимыми антибиотиками (подробнее в Материалах и методах). С чашек были поставлены жидкие культуры в двух репликах, которые выращивались ночь, а затем

использовались для измерения OD₆₀₀ и проверки адаптации. Затем образцы ночной культуры были разбавлены и загружены в Plate Reader с добавлением и без добавления индуктора и выращивались 20 часов, после чего были использованы для проверки адаптации.

Влияние индукции экспрессии генов *cas1* и *cas2* на рост клеток и на процесс CRISPR-адаптации

В плазмиде pCDF Cas1+2 гены *cas1* и *cas2* находятся под контролем T7 полимеразы. Штамм является производным штамма BL21 (DE3), в геноме которого присутствует ген, кодирующий T7 полимеразу. Хотя для экспрессии T7 полимеразы необходим индуктор, в его отсутствие часто наблюдается «подтекание» промотора, обеспечивающее экспрессию генов. С другой стороны, индуцированная экспрессия может приводить к повышенным концентрациям белка интереса, что может быть токсичным для клетки. Поэтому в начале мы решили определить, какие условия индукции будут обеспечивать оптимальное количество продуктов генов *cas1* и *cas2*, достаточное для CRISPR-адаптации, но не токсичное для клетки.

При выращивании ночной культуры индукторы не добавлялись, что предполагает отсутствие экспрессии этих генов или их низкий уровень экспрессии в результате «подтекания» (ослабленного контроля) промотора. Измерения оптической плотности образцов ночных культур не показали различий в плотностях культур клеток, несущих на плазмиде гены, кодирующие адаптационный комплекс Cas1-Cas2, и без них, то есть с пустой плазмидой (рис.7), что было верно для образцов со всеми вариантами плазмид, кодирующих комплекс RecBCD. Таким образом, возможный низкий уровень экспрессии генов не влияет на рост клеток. Проверка уровня CRISPR-адаптации в клетках ночных культур методом ПЦР не выявила удлиненных вариантов CRISPR-кассет (данные не показаны).

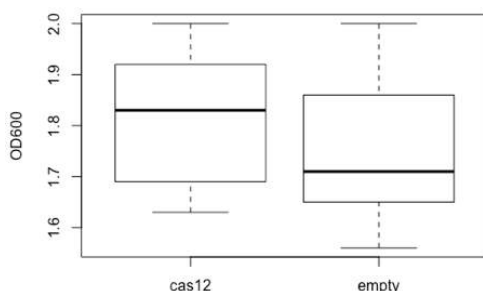


Рисунок 7. Измерения оптической плотности ночных культур, выросших без индукторов в присутствии генов *cas1,2* (*cas12*) и без них (*empty*).

На второй стадии эксперимента клетки выращивались в присутствии или отсутствии индукторов экспрессии генов *cas1* и *cas2* в течение 20 часов в приборе Plate Reader (рис.8). В отсутствие индуктора скорость роста клеток в присутствии плазмиды, кодирующей гены *cas1* и *cas2*, и в

присутствии пустой контрольной плазмиды были сопоставимы для всех образцов (светло- и темно-зеленые кривые роста на графиках). В присутствии индуктора в большинстве случаев существенной разницы также не наблюдается, за исключением клеток, несущих плазмиды pSA21 и pSA618 (рис.8 б,г). Однако наблюдается сильное торможение роста клеток в присутствии индукторов, не связанное с экспрессией генов *cas1* и *cas2*. Таким образом, на основании измерений оптической плотности в Plate Reader видно, что значительно сильнее на выживаемость клеток влияет присутствие индуктора, нежели наличие генов *cas1* и *cas2* (рис.8).

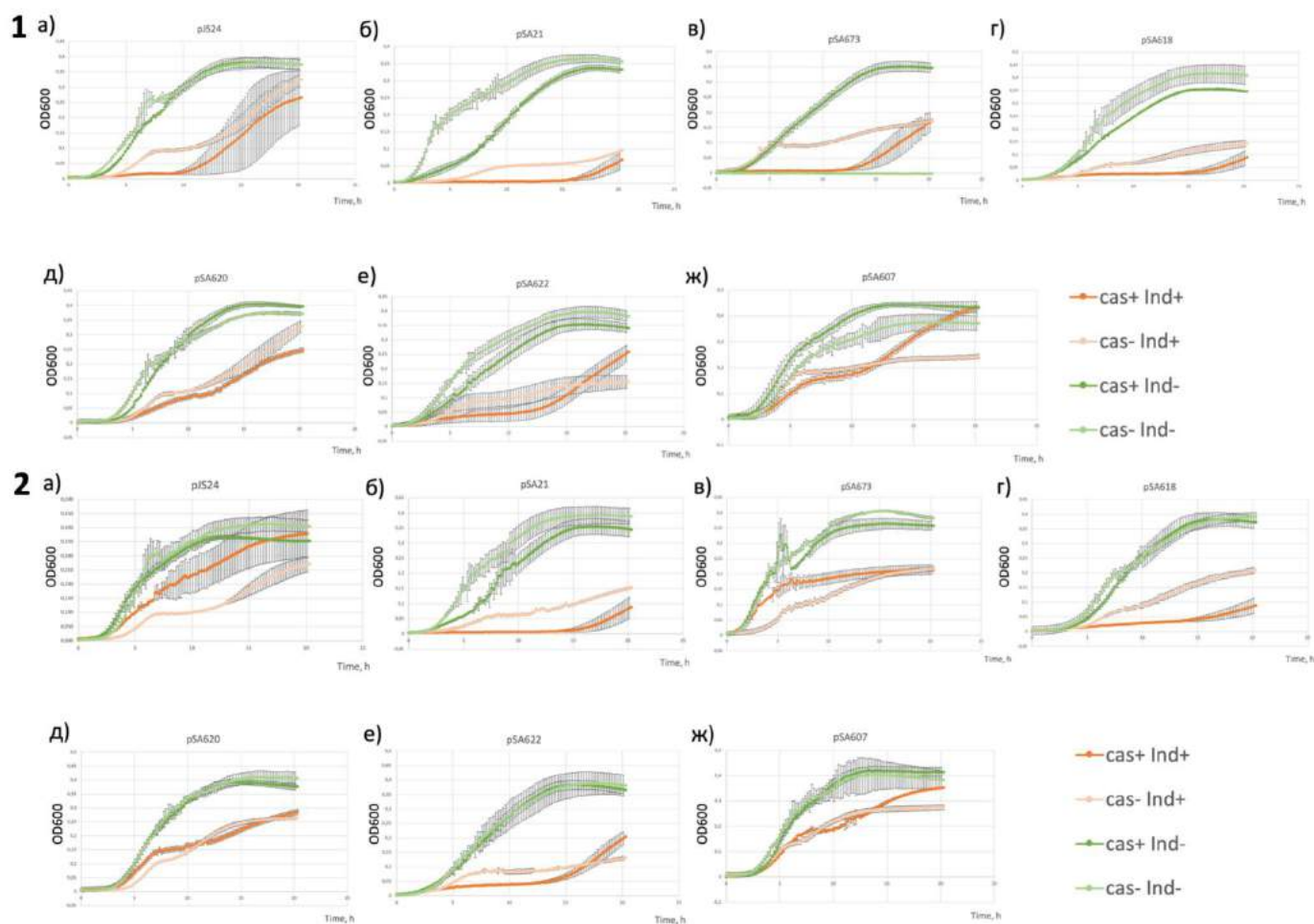


Рисунок 8. Кривые роста, построенные в Microsoft Excel на основе измерений оптической плотности образцов в Plate Reader для первой (1) и второй (2) реплик. Буквы а, б, в, г, д, е, ж обозначают кривые роста образцов клеток, трансформированных плазмидами с разными вариантами генов, кодирующих комплекс RecBCD: pJ524, pSA21, pSA673, pSA618, pSA620, pSA622, pSA607, соответственно. Оранжевым цветом обозначены кривые роста образцов в присутствии индуктора (Ind+): ярким – несущих плазмиду с генами *cas1* и *cas2* (*cas+*), бледным – несущих пустую плазмиду (*cas-*). Аналогично обозначены кривые роста образцов в отсутствии индуктора (Ind-) ярким (*cas+*) и бледным (*cas-*) зелеными цветами. Каждая точка соответствует

среднему по наблюдениям трех технических реплик. Планки погрешности соответствуют стандартному отклонению по трем техническим репликам.

Анализ CRISPR-адаптации в клетках после 20 часов инкубации в Plate Reader показал, что встраивание спейсеров идет одинаково хорошо в присутствии и в отсутствие индукторов (результат в присутствии индукторов не показан, результат в отсутствие индукторов см ниже (рис.10)). Этот результат свидетельствует о том, что индуцируемый промотор на плазмиде pCDF, несущей *cas*-гены, сильно подтекает. Так или иначе, в отсутствие индуктора гены *cas1* и *cas2* экспрессируются в количестве, достаточном для адаптации, но не токсичном для клетки. Так как в присутствии индукторов скорость клеток значительно уменьшалась, дальнейший анализ решили проводить только для образцов в отсутствие индукторов.

Влияние мутаций в генах, кодирующих субъединицы RecBCD, на рост клеток

При сравнении роста клеток, несущих плазмиды, кодирующие вариант комплекса RecBCD дикого типа (серые кривые) и полностью неактивный вариант (желтые кривые), видно, что в отсутствие функционального комплекса скорость роста клеток замедлена (рис.9).

При сравнении кривых роста клеток, несущих плазмиды с разными вариантами генов *recB*, *recC* и *recD*, с кривыми роста клеток, несущих плазмиду с генами комплекса RecBCD дикого типа (pSA607) и несущих плазмиду с геном неработающего RecB (pSA21), можно увидеть, что все мутации снижают выживаемость клеток (рис.9). Больше всего затронута выживаемость клеток, несущих плазмиду, кодирующую субъединицу RecB без моторной и нуклеазной активности (pSA618 – рис.9в) и плазмиду с аналогичной мутацией в гене *recB* и мутацией в гене *recD*, которые вместе приводят к полному отсутствию нуклеазной и хеликазной активности комплекса (pSA622 – рис.9е).

Выживаемость клеток с мутацией, приводящей к потере способности субъединицы RecC распознавать Chi-сайты (pJS24 – рис.9а), затронута не так сильно. Клетки, несущие плазмиду с мутацией в *recD*, снижающей хеликазную активность комплекса на 80%, (pSA620 – рис.9г) растут не так хорошо, как дикий тип, но и не так плохо, как клетки с полностью нефункциональной субъединицей RecB. Результаты для образца с плазмидой pSA673 (рис.9б), несущей небольшую делецию в гене *recB*, не согласуются между двумя биологическими повторами, для определения роли данной мутации потребуются повторы эксперимента.

Однако через 20 часов инкубирования плотности культур отличались друг от друга наименьшим образом, поэтому данная точка была выбрана для анализа CRISPR-адаптации методом ПЦР (см. следующий раздел).

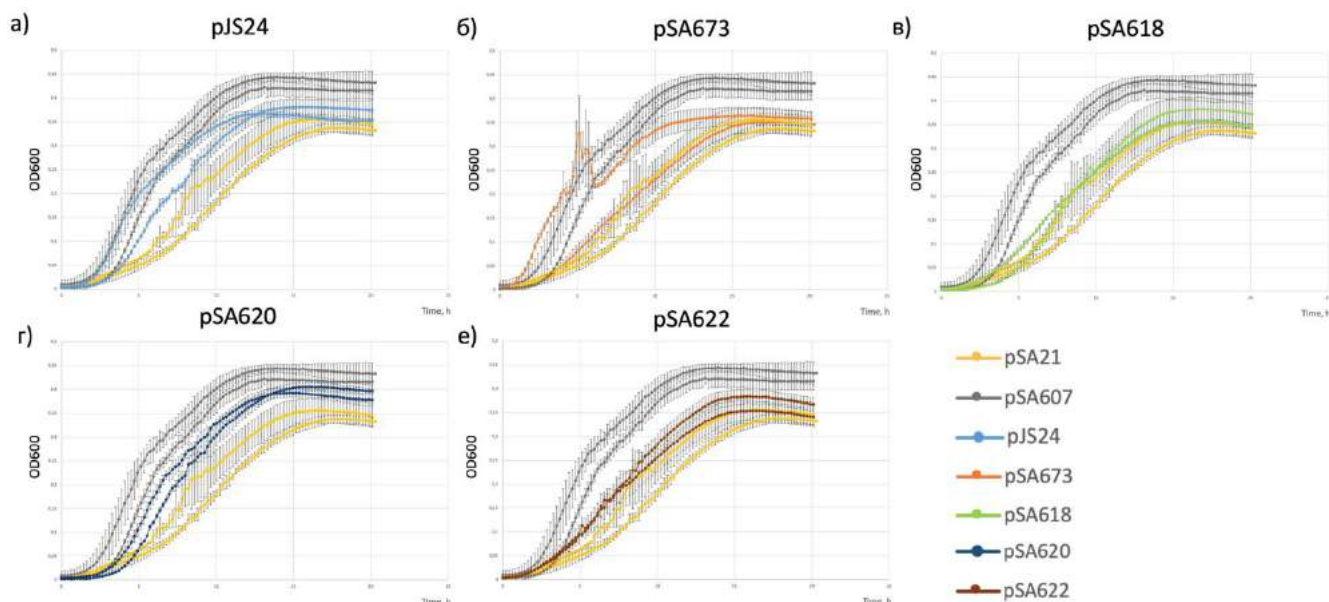


Рисунок 9. Сравнение скорости роста клеток, несущих плазмиды с разными вариантами генов, кодирующих комплекс RecBCD. Серым цветом обозначена кривая роста клеток, несущих плазмиду pSA607 (вариант дикого типа), желтым – плазмиду pSA21 (полностью нефункциональный вариант). Кривые роста клеток, несущих плазмиды pJS21, pSA673, pSA618, pSA620, pSA622 обозначены голубым, оранжевым, зеленым, синим и коричневым, соответственно. Кривые роста образцов культуры клеток, несущих каждую плазмиду, представлены в двух репликах. Взяты кривые роста образцов в отсутствие индуктора. Каждая точка соответствует среднему по наблюдениям трех технических реплик. Планки погрешности соответствуют стандартному отклонению по трем техническим репликам.

Влияние мутаций в генах, кодирующих субъединицы RecBCD, на процесс адаптации

Исследования CRISPR-адаптации проводили после 20 часов инкубации культуры, в данной временной точке количество клеток было сопоставимо во всех анализируемых образцах. Нами был проведен полуколичественный анализ CRISPR-адаптации. Об эффективности CRISPR-адаптации судили исходя из интенсивности полос, соответствующих фрагментам удлиненных кассет на электрофореграмме (рис.10). Исходя из полученных данных, мы не нашли значительных различий в эффективности CRISPR-адаптации в клетках с разными вариантами генов, кодирующих комплекс RecBCD (рис.10).

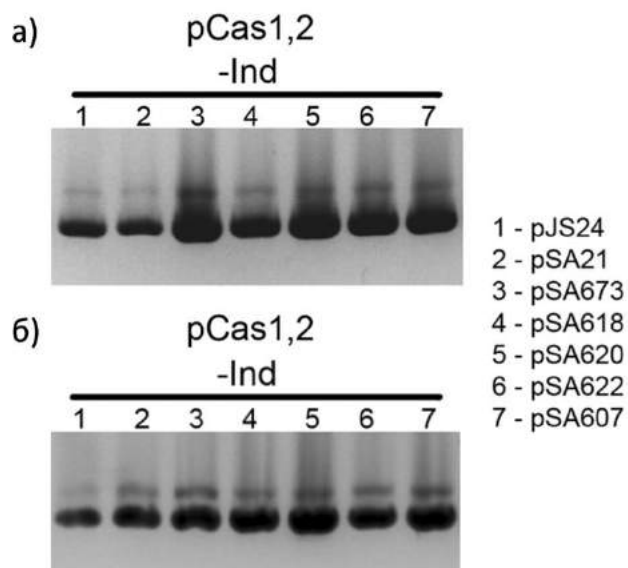


Рисунок 10. Гель-электрофорез ДНК-фрагментов, соответствующих CRISPR-кассетам, амплифицированным при помощи ПЦР после роста в Plate Reader без индуктора. Жирные полосы соответствуют неудлиненным CRISPR-кассетам, бледные полосы выше соответствуют CRISPR-кассетам, удлиненным на один спейсер и один повтор – ~60 нуклеотидов, и свидетельствуют об адаптации. а) Электрофорез ПЦР-продуктов, амплифицированных с образцов первой реплики. б) Электрофорез ПЦР-продуктов, амплифицированных с образцов второй реплики.

Обсуждение

Главной целью данной работы было изучение механизма влияния комплекса RecBCD на CRISPR-адаптацию у *E. coli*. Согласно нашим данным, уровень адаптации в клетках с вариантом комплекса RecBCD дикого типа и в клетках с нефункциональным вариантом комплекса принципиально схожи, более того, нет существенной разницы между уровнем адаптации в клетках с вариантами комплекса RecBCD с измененными активностями. Ранее были получены данные о том, что в отсутствие функционального комплекса RecBCD уровень адаптации существенно падает и становится труднодетектируемым. Также было установлено, что Chi-сайты обуславливают место, откуда берется спейсер, что привело к предположению, что продукты нуклеазной активности RecBCD могут служить потенциальным субстратом для комплекса Cas1-Cas2, встраивающего пре-спейсеры в CRISPR-кассету [2]. Несогласованность ранее полученных и наших данных может быть обусловлена несколькими причинами: 1) различным уровнем экспрессии генов и/или соотношением уровней экспрессии генов, кодирующих адаптационный комплекс и комплекс RecBCD, 2) различным набором генов в штаммах BL21(DE3), который использовался в данной работе, и K12, который использовался в работе других авторов, 3) недостоверностью ранее полученных результатов.

Подтвердить или опровергнуть данные предположения можно будет в ходе дальнейших исследований.

Интересно, что мы детектировали разницу в скорости роста клеток (рис. 8), имеющих и не имеющих адаптационный комплекс, на фоне RecBCD-комплекса без нуклеазной активности (pSA21 – RecB²¹CD): в условиях, когда адаптация возможна, клетки росли немного медленнее. Для вариантов RecBCD-комплекса с нарушением других активностей данная задержка в росте не характерна. В целом можно предположить, что нуклеазная активность RecBCD-комплекса необходима для жизнеспособности клетки в условиях CRISPR адаптации. Можно предположить, что RecBCD-комплекс важен для поддержания стабильности генома при встраивании спейсеров в CRISPR-кассету, которое происходит с образованием двуцепочечных разрывов.

Полученные нами данные позволяют также сделать дополнительные выводы. Добавление IPTG к клеткам, производным штамма BL21 (DE3), существенно снижает скорость их роста. Эксперименты, не описанные в данной работе, свидетельствуют о том, что из двух индукторов (L-арабиноза и IPTG) для клеток токсичен IPTG. Возможно, скорость роста клеток снижается в результате того, что вся транскрипция и ресурсы клетки переключаются на обслуживание высокоэффективного T7 промотора. Скорость роста клеток не зависит от количества адаптационного комплекса Cas1-Cas2, так как и клетки с пустой плазмидой демонстрируют сильно замедленный рост, т.е., возможно, клетки способны выдерживать достаточно высокие концентрации адаптационного комплекса, который не является токсичным. Еще одним возможным объяснением является токсичность самого IPTG для клеток. Ранее было показано, что в высоких концентрациях IPTG значительно ухудшает темпы роста клеток *E. coli* [20,21].

Наши результаты также позволяют сделать выводы о жизнеспособности клеток в присутствии вариантов комплекса RecBCD (рис. 8 и 9). Наиболее выраженное снижение выживаемости клеток вызывают мутации в гене субъединицы RecB (pSA21 – RecB²¹CD, pSA673 – RecB^{Δ19}CD, pSA618 – RecB^{K29Q}CD), что говорит о том, что хеликазная и нуклеазная активности этой субъединицы играют более важную роль, чем хеликазная активность RecD (pSA620 – RecBCD^{K177Q}) и активность RecC (pJS24 – RecBC^{S39E}D) и всего комплекса по отношению к Chi-сайтам. В данной работе нам не удалось оценить роль именно нуклеазной активности RecB, это может быть исследовано в дальнейшем.

Выводы

1. Методом λ-Red-рекомбинации мы получили производную штамма *E. coli* BL21 (DE3) с делецией генов *recB*, *recC* и *recD*.
2. Экспрессия генов *cas1* и *cas2* и процесс адаптации в общем случае не влияют на выживаемость клеток.

3. Хеликазная и нуклеазная активности субъединицы RecB комплекса RecBCD больше всего необходимы для роста и выживания клеток.
4. Мутации в генах комплекса RecBCD существенно не влияют на уровень адаптации в клетках. Нуклеазная активность субъединицы RecB RecBCD-комплекса участвует в поддержании жизнеспособности клеток в условиях CRISPR-адаптации.

Благодарности

Мы благодарим Сергея Менделевича Глаголева за организацию практики, Екатерину Евгеньевну Савицкую за научное руководство, Антона Олеговича Шкаруту за помощь в получении и подтверждении мутантного штамма и за неоценимую помощь в ходе работы, Артёма Исаева за помощь в обработке данных, а также всех сотрудников лаборатории.

Ссылки

1. Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.*, 2012, Vol. 40, no. 12, pp. 5569–5576
2. Levy A, et al. CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature*, 2015, Vol. 520, no. 7548, pp. 505-510
3. Semenova E, Severinov K. Interdependencies Between the Adaptation and Interference Modules Guide Efficient CRISPR-Cas Immunity. 2017, *Evolutionary Biology: Self/Nonsel Evolution, Species and Complex Traits Evolution, Methods and Concepts*, pp. 51-62
4. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakamura A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isoenzyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.*, 1987, Vol. 169, no. 12, pp. 5429–5433
5. Mojica FJ, Rodriguez-Valera F. The discovery of CRISPR in archaea and bacteria. *FEBS J.*, 2016, Vol. 283, no. 17, pp. 3162-3169
6. Mojica FJ, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.*, 2000, Vol. 36, no. 1, pp. 244–246
7. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 2007, Vol. 315, no. 5819, pp. 1709-1712

8. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, 2005, Vol. 151 no. 8, pp. 2551-2561
9. Ширяева А, Семёнова Е, Северинов К. Системы CRISPR-Cas бактерий и архей. Как компоненты адаптивной иммунной системы прокариот стали инструментом для модификации геномов и регуляции транскрипции? Редактирование генов и геномов: Том 1 / отв. ред. С. М. Закиян, С. П. Медведев, Е. В. Дементьева, Е. А. Покушалов, В. В. Власов; Рос. акад. наук, Сиб. отд-ние, ФИЦ Ин-т цитологии и генетики [и др.]. — 2-е изд., расширенное и доп. — Новосибирск: Издательство СО РАН, 2018, с. 94-157
10. Amitai G, Sorek R. CRISPR–Cas adaptation: insights into the mechanism of action. *Nat. Rev. Microbiol.*, 2016, Vol. 14, no. 2, pp. 67-76
11. Dillingham MS, Kowalczykowski SC. RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks. *Microbiol. Mol. Biol. Rev.*, 2008, Vol. 72, no. 4, pp. 642-671
12. Ivančić-Baće I, et al. Different genome stability proteins underpin primed and naïve adaptation in *E. coli* CRISPR-Cas immunity. *Nucleic Acids Res.*, 2015, Vol. 43, no. 22, pp. 10821–10830
13. Amundsen SK, Taylor AF, Reddy M, Smith GR. Intersubunit signaling in RecBCD enzyme, a complex protein machine regulated by Chi hot spots. *Genes Dev.*, 2007, Vol. 21, no. 24, pp. 3296–3307
14. Amundsen SK, Sharp JW and Smith GR. RecBCD Enzyme “Chi Recognition” Mutants Recognize Chi Recombination Hotspots in the Right DNA Context. *Genetics*, 2016, Vol. 204, no. 1, pp. 139–152.
15. Datsenko KA, Wanner BL. One step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.*, 2000, Vol. 97, no. 12, pp. 6640-6645
16. Amundsen SK, Smith GR. The RecB helicase-nuclease tether mediates Chi hotspot
17. control of RecBCD enzyme. *Nucleic Acids Res.*, 2019, Vol. 47, no. 1, pp. 197-209
18. Hsieh S, Julin DA. Alteration by site-directed mutagenesis of the conserved lysine residue in the
19. consensus ATP-binding sequence of the RecB protein of *Escherichia coli*. *Nucleic Acids Res.*, 1992, Vol. 20, no. 21, pp. 5647–5653
20. Taylor AF, Smith GR. RecBCD enzyme is a DNA helicase with fast and slow motors of
21. opposite polarity. *Nature*. 2003, Vol. 423, no. 6942, pp. 889-893
22. Amundsen SK, Sharp JW, Smith GR. RecBCD Enzyme "Chi Recognition" Mutants
23. Recognize Chi Recombination Hotspots in the Right DNA Context. *Genetics*. 2016, Vol. 204, no. 1, pp. 139-152
24. Kosinski MJ, Rinas U, Bailey JE. Isopropyl- β -D-thiogalactopyranoside influences the metabolism of *Escherichia coli*. *Appl Microbiol Biotechnol.*, 1992, Vol. 36, no. 6, pp. 782-784

25. Malakar P, Venkatesh KV. Effect of substrate and IPTG concentrations on the burden to growth of Escherichia coli on glycerol due to the expression of Lac proteins. *Appl Microbiol Biotechnol.*, 2012, Vol. 93, no. 6, pp. 2543-2549
26. Dvorak P, et al. Exacerbation of substrate toxicity by IPTG in Escherichia coli BL21(DE3) carrying a synthetic metabolic pathway. *Microb Cell Fact.*, 2015, Vol.14, no. 1, p. 201
27. Sharan SK, Thomason LC, Kuznetsov SG, Court DL. Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.*, 2009, Vol. 4, no. 2, pp. 206-223