

Московский государственный университет имени М.В. Ломоносова

На правах рукописи

Васильев Руслан Алексеевич

**Направленная модификация геномов с помощью новых эндонуклеаз
CRISPR/Cas V типа**

1.5.3 – Молекулярная биология

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата биологических наук

Научный руководитель:
доктор биологических наук
Каменский Петр Андреевич

Москва – 2022

Работа выполнена на кафедре молекулярной биологии биологического факультета ФГБОУ ВО «Московский государственный университет имени М.В. Ломоносова».

Научный руководитель – *Каменский Петр Андреевич, доктор биологических наук*

Официальные оппоненты – *Сергиев Петр Владимирович, доктор химических наук, чл.-корр. РАН, директор института функциональной геномики ФГБОУ ВО «Московский государственный университет имени М.В.Ломоносова»*

Шайтан Алексей Константинович, доктор физико-математических наук, чл.-корр. РАН, доцент ФГБОУ ВО «Московский государственный университет имени М.В.Ломоносова»

Дивашук Михаил Георгиевич, кандидат биологических наук, зав. лабораторией прикладной геномики и частной селекции сельскохозяйственных растений ФГБНУ «Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии»

Защита диссертации состоится «10» ноября 2022 г. в 16 часов на заседании диссертационного совета МГУ.015.4(МГУ.03.01) Московского государственного университета имени М.В.Ломоносова по адресу: 119234, Москва, Ленинские горы, д.1, стр.12, Биологический факультет, ауд. 389.

E-mail: mgu.03.01.dissovet@gmail.com.

С диссертацией можно ознакомиться в отделе диссертаций научной библиотеки МГУ имени М.В.Ломоносова (Ломоносовский просп., д. 27) и на сайте ИАС «ИСТИНА»: <https://istina.msu.ru/dissertations/494119444/>
Автореферат разослан «___» октября 2022 г.

Ученый секретарь
диссертационного совета,
доктор биологических наук



Т.В. Комарова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы и степень ее разработанности

За время, прошедшее с момента открытия коротких палиндромных повторов, регулярно расположенных группами (CRISPR), в 1987 году, до первых попыток применения Cas-нуклеаз в качестве редакторов генома в 2013 году, системы CRISPR/Cas представляли ограниченный интерес для мирового научного сообщества. Была выявлена биологическая роль CRISPR-повторов в геномах многих прокариот, осуществлены попытки классифицировать эти защитные системы, а также исследованы структурные и функциональные особенности систем. Однако попыток к использованию систем CRISPR/Cas в качестве редакторов генома клеток млекопитающих не предпринималось вплоть до 2013 года, когда появилось большое количество работ, исследовавших различные аспекты использования новой технологии для редактирования ядерных геномов различных организмов. Подобный интерес был вызван потенциальными возможностями применения исследуемых ферментов, а также сравнительной легкостью проводимых экспериментов, ставших доступными большинству современных молекулярно-биологических лабораторий. Однако с расширением фундаментальных знаний о структуре и функциях различных систем CRISPR/Cas стало очевидно, что возможности каждой отдельно взятой системы ограничены теми или иными факторами. На сегодняшний день не известно систем CRISPR/Cas, универсально пригодных для решения любых задач практической геномной инженерии. В этой связи поиск новых инструментов редактирования генома и изучение их свойств является актуальной научной задачей, решения которой постепенно снимают барьеры, стоящие на пути исследований в данной области. Одним из подобных барьеров является РАМ-специфичность Cas-нуклеаз, не позволяющая направлять эффекторные комплексы к любому локусу генома. В сфере геномного редактирования растений известен другой фактор, затрудняющий работы по получению мутантов: многие растения, в том числе сельскохозяйственного пользования, плохо переносят температуры выше комнатной в лабораторных условиях, в то время как комплексы CRISPR/Cas часто имеют температурный оптимум работы в районе 37 °С. Возможным путем преодоления этих и других барьеров является изучение нуклеаз, обладающих свойствами, выгодно отличающих их от ортологов, таких как менее строгие требования к РАМ или активность в широком диапазоне температур. Исходя из этого, поиск новых направляемых нуклеаз и определение их биохимических свойств является важной задачей современной молекулярной биологии.

Цель работы

Целью данной работы является получение и характеристика новой эндонуклеазы семейства Cas12a, способной вносить направленные модификации в геном культуры клеток человека.

Задачи работы

1. Биоинформатический поиск нуклеотидной и аминокислотной последовательностей новой нуклеазы Cas подтипа V-A.

2. Получение рекомбинантного белка и определение его активности в составе комплексов с различными направляющими РНК в условиях *in vitro*.

3. Исследование требований к PAM-последовательности новой нуклеазы Cas подтипа V-A.

4. Определение свойств новой нуклеазы Cas подтипа V-A в условиях *in vitro*: кофактор-зависимость, температурный оптимум работы, определение структуры выступающих концов, наличие коллатеральной активности.

5. Исследование возможности осуществления разрезания *in vitro* комплексами различных Cas-нуклеаз подтипа V-A, заряженными двумя молекулами РНК: отдельно спейсерной и скаффолдной частями направляющей РНК.

6. Исследование возможности получения направленной делеции в гене *DNMT1* в геноме культуры клеток НЕК293Т при помощи новой нуклеазы Cas подтипа V-A.

Научная новизна

В работе впервые была показана и изучена активность эффекторной нуклеазы системы CRISPR/Cas подтипа V-A из бактерии *Ruminococcus bromii* sp. (RbCas12a). Продемонстрирована способность нуклеазы RbCas12a в комплексах с направляющими РНК (crРНК) осуществлять гидролиз двуцепочечной ДНК (дцДНК) *in vitro*. Показана активность RbCas12a *in vitro* в отношении ДНК, содержащих PAM-последовательности с консенсусом 5'-YYN-3'. Выявлена независимость эффективности разрезания от последнего нуклеотида PAM в отношении одного и того же протоспейсера; в частности, не наблюдалось дискриминации в отношении тимидинового нуклеотида в последнем положении PAM, свойственной некоторым ортологам Cas12a. Проведен анализ свойств нуклеазы RbCas12a, подтверждающий наибольшую активность фермента при использовании в качестве кофакторов ионов магния или марганца. Проведен сравнительный анализ зависимости эффективности

разрезания матрицы от времени инкубации с комплексами RbCas12a-crРНК и AsCas12a-crРНК, продемонстрировавший высокую эффективность разрезания RbCas12a спустя несколько секунд инкубации. Проведен сравнительный анализ активности комплекса RbCas12a-crРНК при различных температурах, по результатам которого обнаружено, что комплекс способен эффективно функционировать при диапазоне температур от 20 до 42 °С. Впервые показано, что различные ортологи Cas12a, включая RbCas12a, способны формировать эффекторные комплексы с двумя отдельными молекулами РНК, функционирующими соответственно как спейсерная и скаффолдная части crРНК, практически не теряя при этом эффективности в условиях *in vitro*. Продемонстрирована возможность разрезания *in vitro* матрицы комплексами Cas12a только со спейсерной частью crРНК при отсутствии скаффолда. Показано, что RbCas12a проявляет коллатеральную активность в отношении одноцепочечной ДНК (оцДНК). Определена природа выступающих концов при гидролизе ДНК матрицы комплексом RbCas12a-crРНК: формирование 5'-выступающих концов на 1 нуклеотид длиннее, чем при гидролизе ортологом Cas12a из *Francisella novicida* (FnCas12a). Показана возможность внесения направленных делеций в геном культуры клеток человека HEK293T с использованием кодон-оптимизированного RbCas12a.

Научная и практическая значимость

На текущий момент интерес к использованию новых нуклеаз существует в таких отраслях, как биотехнология, медицина, сельское хозяйство, но для их использования в этих сферах необходимо четко и всесторонне охарактеризовать ферменты, выявить оптимальные условия их работы и возможные положительные и отрицательные параметры в приложении к конкретным задачам. Поиск и характеристика новых представителей систем CRISPR/Cas расширяет возможности их применения в различных областях. Обнаружение нуклеаз с различной специфичностью по отношению к РАМ-последовательности позволит редактировать практически любой желаемый локус в геноме. Новые ферменты с уникальными свойствами могут найти свое применение в решении большого количества текущих задач промышленности, медицины и сельского хозяйства. Как показывает международная практика, системы геномного редактирования уже сейчас используются в отдельно взятых государствах. Предложены и внедряются подходы к терапии некоторых генетических заболеваний. Рассматриваются варианты редактирования генома с целью снижения рисков инфекционных заболеваний и других социально-значимых патологий. Также применение культур с отредактированным геномом в

сельском хозяйстве за последние десятилетия выросло во всем мире и продолжает расти, что в условиях повышения плотности населения и возможного глобального потепления будет играть ведущую роль в поддержании темпов урожайности сельскохозяйственных растений. Получение отечественных ферментов для редактирования генома, в частности CRISPR/Cas-нуклеаз, является стратегически-важным направлением развития биотехнологической отрасли.

Личный вклад автора

Личный вклад автора заключается в работе с литературными источниками, планировании и проведении экспериментов, анализе полученных результатов, подготовке материалов к печати и написании диссертационной работы. Основные результаты диссертации получены лично автором или при его непосредственном участии. Подготовка к публикации полученных результатов проводилась совместно с соавторами, причём вклад соискателя был определяющим. Имена всех соавторов указаны в опубликованных работах. Планирование экспериментов было выполнено совместно с д.б.н. П.А. Каменским, к.б.н. И.О. Мазуниным, к.б.н. М.В. Патрушевым. Выделение белков было выполнено совместно с к.х.н. А.В. Власкиной. Эксперименты по разрезанию *in vitro* комплексами Cas нуклеаз с двумя молекулами РНК были проведены совместно с Р.М. Шебановой, Н.А. Никитчиной, к.б.н. И.О. Мазуниным, д.б.н. К.В. Севериновым. Анализ данных высокопроизводительного секвенирования проводился совместно с к.б.н. С.В. Тоцаковым, А.А. Корженковым.

Методология и методы исследования

Рекомбинантные белки выделены с использованием хроматографических методов с подтвержденной чистотой, пригодной для выполнения ферментативных реакций. Исследования проводились на клеточной культуре НЕК293Т со статусом отсутствия инфицирования микоплазмой, подтвержденным тестами с использованием коммерческого набора MycoReport (Евроген). Анализ данных высокопроизводительного секвенирования был выполнен с использованием различных биоинформатических программ. Визуализация полученных результатов высокопроизводительного секвенирования была выполнена с использованием различных пакетов языка Python3.

Положения, выносимые на защиту

1. Идентифицирована новая нуклеаза Cas12a из *Ruminococcus bromii* sp. (RbCas12a).

2. Рекомбинантный препарат RbCas12a обладает активностью в отношении двуцепочечной ДНК в условиях *in vitro*.

3. Комплексы RbCas12a-crРНК способны направленно гидролизовать ДНК, содержащие последовательность 5'-YYN-3' в качестве PAM.

4. Нуклеаза RbCas12a способна использовать в качестве кофакторов ионы магния и марганца, а также, с меньшей эффективностью, ионы кобальта, цинка и никеля. Температурный оптимум работы RbCas12a находится в диапазоне от 20 до 42 °С. В качестве продуктов гидролиза образуется дцДНК с 5'-выступающими концами. RbCas12a проявляет выраженную коллатеральную активность.

5. Cas-нуклеазы V-A подтипа, в том числе RbCas12a, способны формировать комплексы с двумя молекулами РНК, представляющими собой соответственно спейсерную и скаффолдную части crРНК, и осуществлять в их составе эффективный направленный гидролиз двуцепочечной ДНК в условиях *in vitro*.

6. Нуклеаза RbCas12a активна в клетках человека, что показано путем внесения направленной делеции в ген *DNMT1* культуры клеток НЕК293Т.

Степень достоверности и апробация результатов

Результаты были получены с использованием современных методов молекулярной биологии и биоинформатики. По теме диссертационной работы опубликовано 3 статьи в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ по специальности 1.5.3 – Молекулярная биология.

Результаты работы были представлены на Всероссийской конференции «Синтетическая биология и биофармацевтика» (Новосибирск, 24-28 июля 2022 г.).

Структура и объем диссертации

Диссертация изложена на 116 страницах и состоит из следующих основных разделов: «Введение», «Обзор литературы», «Материалы и методы», «Результаты и обсуждение», «Заключение», «Выводы», «Список использованной литературы», «Дополнительные материалы». Диссертация содержит 2 таблицы и 23 рисунка. Список литературы включает 116 источников. Дополнительные материалы включают 3 приложения.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

1. Получение плазмидных конструкций и рекомбинантного препарата RbCas12a

На первом этапе мы обнаружили последовательность нуклеазы RbCas12a в локальной базе метагеномных данных. После чего последовательности CRISPR-локуса анализировались в ПО CRISPR Target и CCTyper для идентификации CRISPR повторов.

Последовательности гена RbCas12a длиной 3735 п.о. и консенсуса повтора из CRISPR-массива (то есть скаффолда *sgRNA*) далее использовались для молекулярного клонирования векторов. Вектор RbCas12a_pET30a+, кодирующий целевой белок с гексагистиридиновой меткой (6xHis-tag) на С-конце, далее использовался для получения рекомбинантного RbCas12a. Вектор pSQT1313-CMV_RFP-U6_scaffold далее использовался в экспериментах по клонированию спейсеров *sgRNA* для экспрессии в культуре клеток HEK293T.

2. Определение температурного оптимума работы RbCas12a

На начальных этапах определения температурного оптимума работы RbCas12a выяснилось, что при смешивании компонентов реакция успевает пройти на комнатной температуре еще до инкубации смеси на соответствующей температуре уже менее, чем через 1 секунду. Влияние времени инкубации на эффективность разрезания RbCas12a сравнивали с гомологом из *Acidaminococcus sp. BV3L6* (AsCas12a). Как известно, AsCas12a также проявляет активность на комнатной температуре, что позволяет сравнить скорости направленного гидролиза матрицы для двух белков. Полученные данные позволяют сделать вывод о более эффективном разрезании дцДНК *in vitro* в реакциях с нуклеазой RbCas12a, что выгодно отличает ее от гомолога AsCas12a (Рисунок 1).

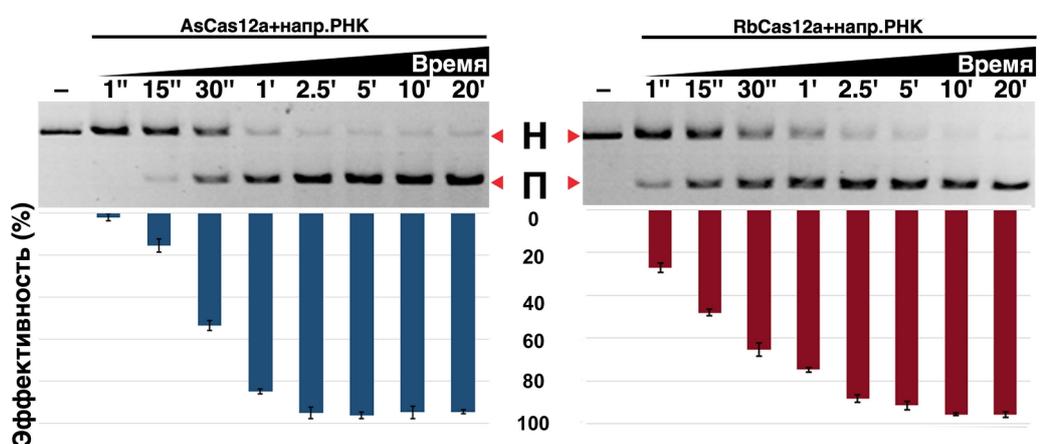


Рисунок 1 - Влияние времени инкубации на эффективность разрезания дцДНК комплексами RbCas12a/crРНК и AsCas12a/crРНК. Разрезание ДНК-мишени проводили рекомбинантными AsCas12a или RbCas12a в присутствии 500 нМ crРНК при соотношении ДНК:Cas12a:РНК 1:3:30. Показаны средние эффективности разрезания и стандартные отклонения, рассчитанные по трем экспериментам. Н., П., напр.РНК — соответственно непорезанные фрагменты, порезанные фрагменты и направляющая (crРНК). Дорожки с реакциями без crРНК обозначены как «-».

Из Рисунка 1 следует, что уже более 20 % матрицы гидролизовано через 1 сек в случае RbCas12a, и менее 1% в случае AsCas12a при эквимольной концентрации компонентов реакции. Через 15 сек более половины всей матрицы разрезано в случае RbCas12a, и менее 20 % в случае AsCas12a. Таким образом, было обнаружено, что определение температурного диапазона работы RbCas12a стандартным способом не представляется возможным. Поэтому было предложено предварительно нагревать по отдельности все компоненты смеси до соответствующих температур, преинкубировать рибонуклеопротеиновый комплекс при данной температуре, а затем добавлять в него ДНК, также предварительно нагретую до соответствующей температуры. Результаты представлены на Рисунке 2.

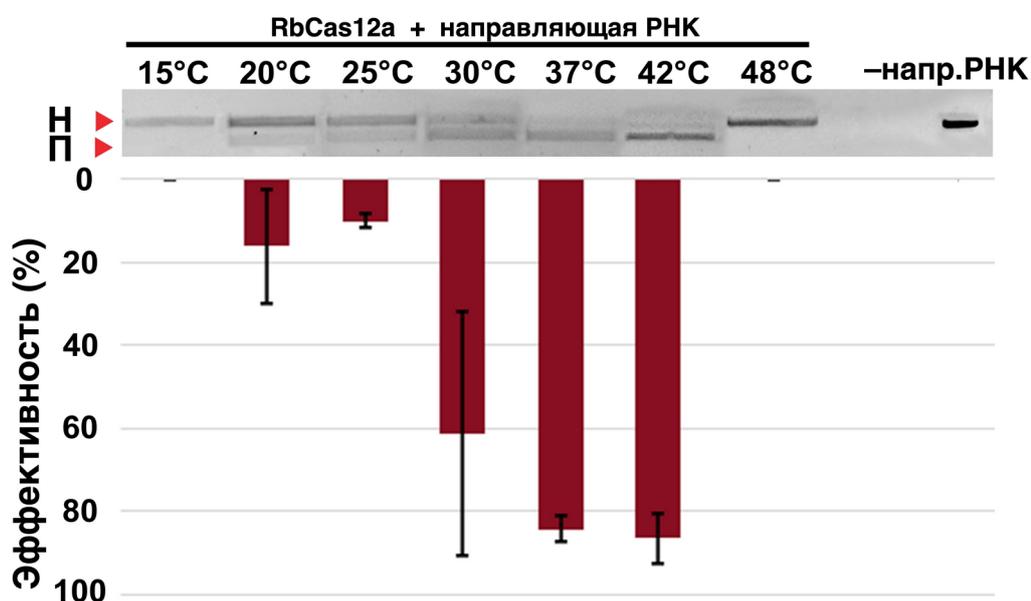


Рисунок 2 - Влияние температуры на эффективность разрезания *in vitro* RbCas12a. Разрезание дцДНК-мишени с помощью RbCas12a, в присутствии 500 нМ crРНК при соотношении ДНК:RbCas12a:РНК 1:3:30. Показаны средние эффективности разрезания и стандартные отклонения, рассчитанные по трем экспериментам. Н., П., напр.РНК — соответственно непорезанные фрагменты, порезанные фрагменты и направляющая (crРНК).

Из Рисунка 2 следует, что наибольшая активность наблюдается при 42 °С, но она не сильно отличается от таковой при 37 °С. При 48 °С активность комплекса не детектируется, что соотносится с полученными данными о точке плавления белка, равной 45.5 °С, полученными в предыдущих экспериментах. При 30 °С активность комплекса снижается в некоторых репликах, в результате по полученным данным составляя 61.3 ± 29.2 процента разрезания. Интересно, что небольшая активность комплекса RbCas12a/crРНК сохраняется и при 20 °С, однако не детектируется при 15 °С. Вероятнее всего, при 20 °С комплекс становится нестабильным, что отражает большой разброс эффективностей в различных репликах эксперимента — от 5.4 % до 35.5 % разрезания. Крайне важным является тот факт, что комплекс RbCas12a/crРНК сохраняет относительно высокую активность при температурах ниже 37 °С, что говорит о потенциальной возможности использования данного фермента в экспериментах по геномному редактированию видов, чувствительных к температурам. В частности, известно, что для эффективного редактирования растений нужны температуры ниже 37 °С. Таким образом, фермент RbCas12a является потенциальным редактором сельскохозяйственных культур растений. Однако стоит иметь в виду, что текущие данные иллюстрируют лишь результаты экспериментов в конкретных условиях: при использовании конкретной crРНК, конкретной последовательности РАМ и конкретного соотношения ДНК:РНК:белок при отсутствии других компонентов клетки в системе *in vitro*. Для более детального изучения этого вопроса требуются дополнительные эксперименты.

4. Определение кофактора RbCas12a

Поскольку разные нуклеазы подтипа V-A обладают различными предпочтениями к кофакторам, было решено определить, какие катионы могут связываться с RbCas12a для координации гидролиза дцДНК. Катионы Ca^{2+} , Ni^{2+} , Mn^{2+} , Mg^{2+} , Zn^{2+} , Cu^{2+} и Co^{2+} тестировали как возможные кофакторы RbCas12a в концентрации 10 мМ при 37 °С. Результаты представлены на Рисунке 3.

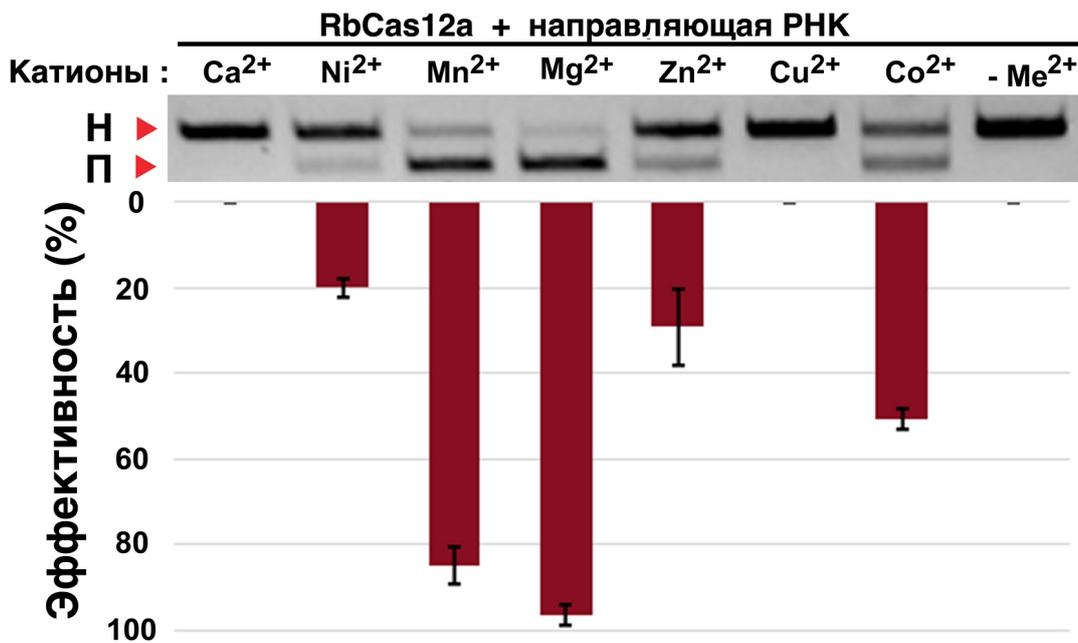


Рисунок 3 - Эффект двухвалентных катионов на разрезание RbCas12a ДНК матрицы *in vitro*. Показаны средние значения относительной эффективности разрезания в % и стандартные отклонения по результатам трех экспериментов. Последняя дорожка (-Me²⁺) - разрезание в том же буфере, но при полном отсутствии двухвалентных катионов. Н., П. — соответственно непорезанные и порезанные фрагменты ДНК.

Как видно из Рисунка 3, Mn²⁺ и Mg²⁺ оказывают схожий эффект на активность RbCas12a, лишь с небольшим преимуществом Mg²⁺, в то время как присутствие вместо них ионов Ni²⁺ или Zn²⁺ практически ингибирует реакции. Хотя RbCas12a и сохраняет активность в присутствии ионов кобальта, только половина молекул ДНК-матрицы гидролизовалась за это время в трех повторных сериях эксперимента. Ионы Ca²⁺ и Cu²⁺ полностью ингибировали реакции разрезания при отсутствии других катионов. Также, как и ожидалось, не наблюдалось разрезания и при полном отсутствии двухвалентных металлов в реакциях.

5. Определение последовательности РАМ

Схема определения последовательности РАМ и результаты соответствующих экспериментов приведены на Рисунке 4.

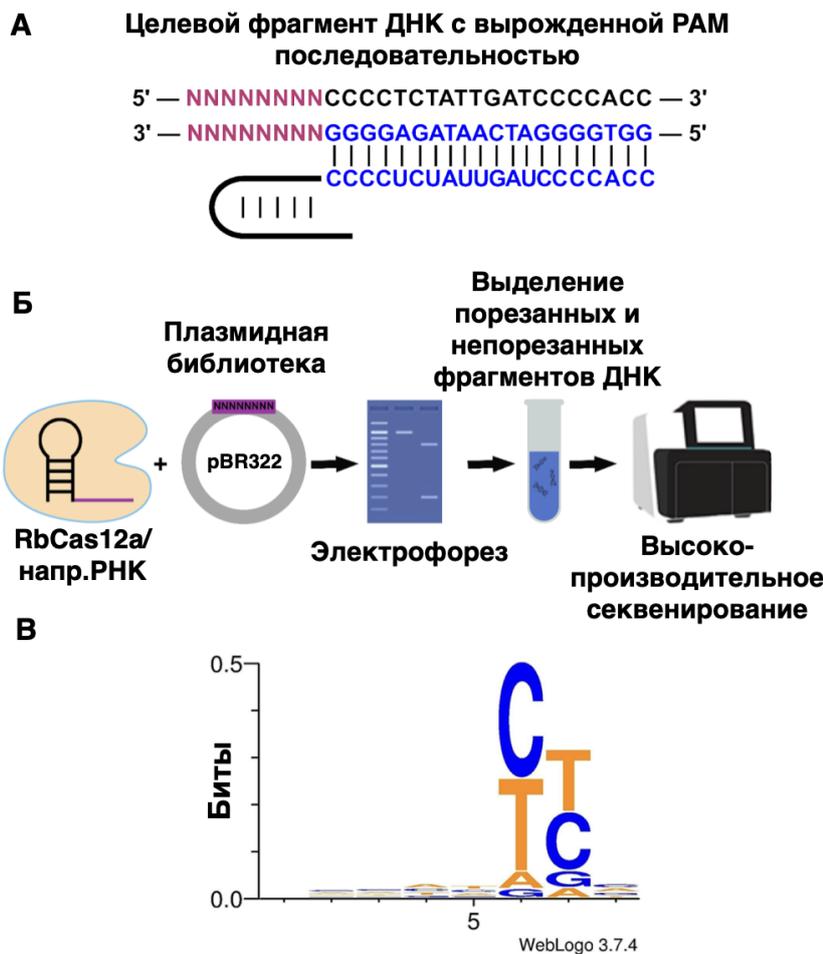


Рисунок 4 - Определение последовательности PAM системы CRISPR–Cas типа V *Ruminococcus bromii* при секвенировании линейаризованных PAM библиотек. (А) Комплекс crРНК с фрагментом двуцепочечной PAM-библиотеки, содержащей 8 вырожденных нуклеотидов на месте предполагаемого PAM. Красным цветом показаны вырожденные нуклеотиды, синим цветом — спейсерная часть crРНК и протоспейсер на таргетной цепи ДНК. (Б) Общая схема эксперимента, используемого для идентификации последовательности PAM RbCas12a. Линейаризованную плазмидную библиотеку, содержащую 8 вырожденных нуклеотидов на месте предполагаемого PAM, гидролизовали в реакциях разрезания *in vitro* с помощью комплекса RbCas12a/crРНК. Далее фрагменты ДНК разделялись в геле электрофоретически и очищались. Из очищенных гидролизованных и негидролизованных фрагментов ДНК готовили библиотеки для высокопроизводительного секвенирования и секвенировали. (В) WebLogo график полученной при анализе данных высокопроизводительного секвенирования PAM-последовательности RbCas12a - 5'-YYN-3'. Нумерация нуклеотидов по оси абсцисс соответствует 5'-3' направлению, начиная с первого вырожденного нуклеотида по нетаргетной цепи PAM-библиотеки (он же нуклеотид -8). Напр.РНК — направляющая (crРНК).

Как видно из лого-графика (Рисунок 4В), для RbCas12a последовательность РАМ была определена как трехнуклеотидный мотив 5'-YYN-3', где Y - пиримидиновый нуклеотид (либо С, либо Т), а N - любой нуклеотид.

Для подтверждения полученных данных, представленных на лого-графике, была проведена серия реакций *in vitro* разрезания ДНК-матриц с различными РАМ-последовательностями из множества 5'-YYN-3'. Поскольку наиболее распространенным консенсусом РАМ нуклеаз подтипа V-A является 5'-TTTN-3' или в некоторых случаях 5'-TTN-3', в первую очередь было решено проверить, влияет ли последний 3'-концевой нуклеотид РАМ на эффективность протекания *in vitro* разрезов матриц с одним и тем же протоспейсером, но разным РАМ (5'-TTN-3' в случае RbCas12a). Как видно из Рисунка 5А, существенных различий по результатам трех независимых экспериментов между 5'-TTC-3', 5'-TTG-3', 5'-TTA-3', 5'-TTT-3' РАМ-последовательностями не обнаружено. В качестве контроля использовалась матрица с 5'-TTG-3' РАМ-последовательностью без добавления crРНК (–напр.РНК). Это доказывает, что положение -1 в РАМ не принципиально для эффективного гидролиза RbCas12a, в то время как наличие тимидина в этом положении РАМ негативно сказывается на работе некоторых других ортологов, например, нуклеаз AsCas12a и LbCas12a, предпочитающих РАМ вида 5'-TTTV-3' (где V представляет собой А, С или G) и некоторые другие субоптимальные РАМ.

Чтобы изучить влияние нуклеотидов в положениях -2 и -3 на эффективность разрезания, мы протестировали активность комплекса RbCas12a/crРНК со всеми оставшимися возможными комбинациями 5'-YYN-3'. Результаты показали, что эффективность RbCas12a-опосредованного разрезания ДНК с РАМ вида 5'-TCB-3' и 5'-CTS-3' составила ~80 % (где B представляет собой С, Т или G, а S представляет собой G или С), что соизмеримо с эффективностями разрезов матриц с РАМ вида 5'-TTN-3' (Рисунок 5Б). Гораздо менее эффективно гидролиз проходил в случае РАМ вида 5'-CCT-3' и 5'-TCA-3' (эффективность ~ 27%), тогда как количество продуктов разрезания матриц с РАМ вида 5'-CCA-3', 5'-CCG-3', 5'-CTT-3', составляли только ~ 10 % от количества интактных фрагментов.

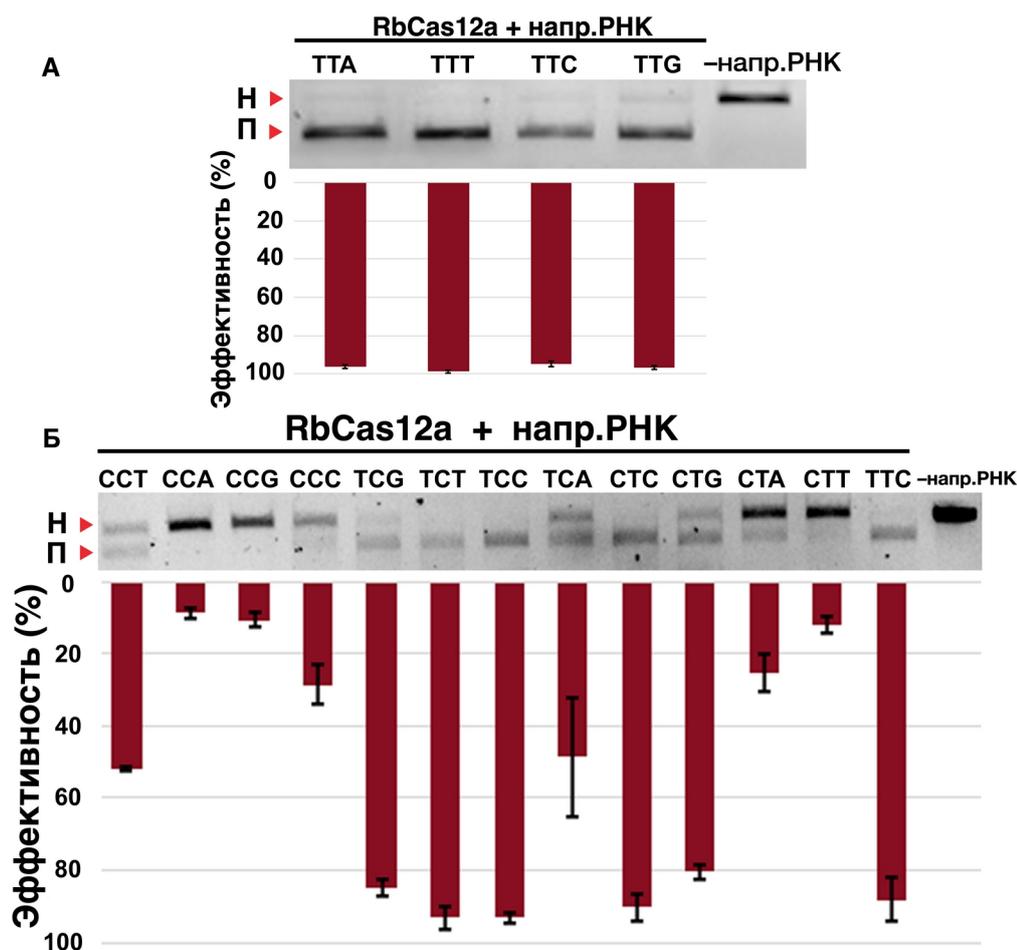


Рисунок 5 - Изучение PAM-специфичности нуклеазы RbCas12a в условиях *in vitro*. (А) Разрезания матрицы с одним и тем же протоспейсером, но разными нуклеотидами в последней 3'-концевой позиции PAM (позиция -1). (Б) Разрезания ДНК с использованием различных crРНК на протоспейсеры с различными последовательностями PAM вида 5'-YYN-3', на последней дорожке контрольная матрица без добавления crРНК. Приведены средние значения относительных эффективностей разрезания в % по трем экспериментам со стандартным отклонением. Н., П., напр.РНК — соответственно непорезанные, порезанные фрагменты ДНК и направляющая (crРНК).

В совокупности эти результаты ясно демонстрируют, что RbCas12a способен эффективно распознавать PAM, определяемый как 5'-YYN-3', для осуществления направленного гидролиза дцДНК-матриц, по крайней мере, *in vitro*. Следует отметить, что в случае 5'-CCV-3' и 5'-CTW-3' (где V представляет собой С, А или G, а W представляет собой А или Т) PAM, RbCas12a не способен разрезать ДНК-матрицы с эффективностью более 25% в условиях описанного эксперимента.

6. RbCas12a проявляет коллатеральную активность

Способность осуществлять неспецифический гидролиз оцДНК после направленного разрезания дцДНК-матрицы, известная как коллатеральная активность, свойственна всем известным ортологам Cas12a. Чтобы показать, что RbCas12a также проявляет эту активность, проводили реакции разрезания *in vitro* в присутствии ДНК-зонда, меченного флуорофором и гасителем на его 5'- и 3'-концах, соответственно. Результаты флуоресцентной детекции неспецифического гидролиза зонда в присутствии целевой матрицы и при ее отсутствии по итогам трех независимых экспериментов представлены на Рисунке 6.

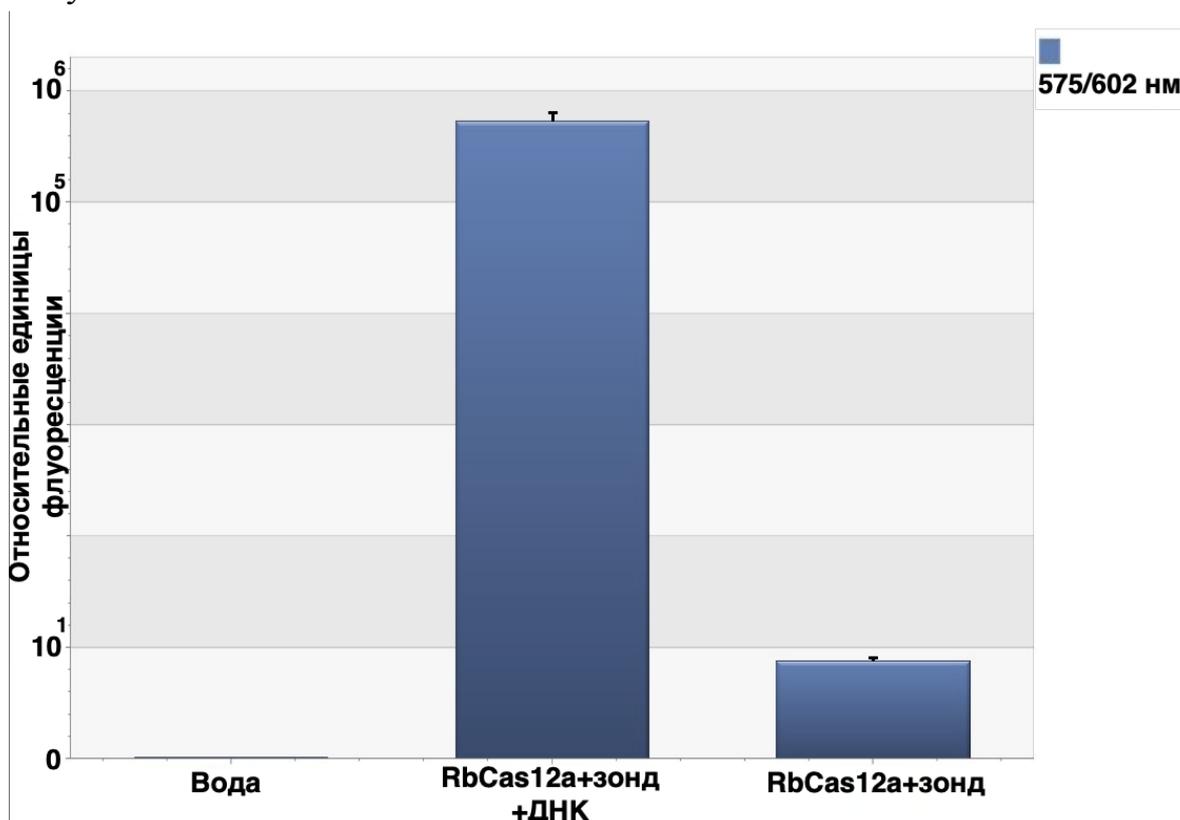


Рисунок 6 - Коллатеральная активность RbCas12a. Данные показывают среднее значение в относительных единицах флуоресценции по результатам трех экспериментов со стандартным отклонением. Левый столбик - вода, средний - реакция разрезания *in vitro* RbCas12a целевой ДНК в присутствии crPНК и зонда с ROX. Правый столбик - контроль реакции разрезания *in vitro* RbCas12a без ДНК матрицы, но в присутствии crPНК и зонда с ROX.

Заметно значительное увеличение уровня флуоресценции в реакциях с целевой ДНК-матрицей по сравнению с отрицательным контролем без матрицы, что говорит о способности RbCas12a нуклеазы проявлять коллатеральную активность по отношению к оцДНК. Следовательно, RbCas12a может быть в

будущем использован в отечественных разработках диагностических тест-систем на основе CRISPR/Cas.

7. RbCas12a формирует выступающие 5'-концы

Все известные ортологи Cas12a формируют при гидролизе выступающие 5'-концы дцДНК, однако средние длины этих концов варьируют для разных белков данного семейства. Чтобы определить природу и длины формируемых RbCas12a концов, продукты гидролиза *in vitro* подвергали «затуплению», клонировали в составе плазмидного вектора и секвенировали по Сэнгеру. Среди семейства Cas12a наиболее подробно природа выступающих концов продуктов гидролиза изучена для ортолога из *Francisella novicida* (FnCas12a). Для сравнения с RbCas12a длины концов в реакциях разрезания *in vitro* использовали коммерческий FnCas12a (Applied Biological Materials, США). По результатам секвенирования по Сэнгеру выяснилось, что выступающий 5'-конец ДНК в случае гидролиза RbCas12a нуклеазой оказался длиннее на 1 нуклеотид, чем тот же 5'-конец, образовавшийся после гидролиза FnCas12a. Стоит отметить, что длина выступающих концов несколько варьирует от клона к клону, что объясняется нативной природой их образования — триммингом таргетной цепи. Таким образом, полученные данные отражают лишь наиболее часто встретившиеся варианты. Более подробную статистику удалось набрать при анализе данных высокопроизводительного секвенирования после редактирования генома клеток млекопитающих, о чем будет сказано ниже.

8. Сплит направляющей РНК

В ходе работ было изучено влияние укорочения crРНК и расщепление ее на две молекулы на *in vitro* разрезание ДНК различными нуклеазами Cas12a. Для выяснения, способен ли RbCas12a осуществлять направленный гидролиз ДНК, связывая сплит-версию crРНК или одну только спейсерную часть crРНК, проводили разрезания *in vitro* в присутствии указанных компонентов реакций. Как показано на Рисунке 7А, в реакциях, содержащих как полноразмерные, так и сплит crРНК, в условиях эксперимента целевая ДНК гидролизовалась на ~ 90 %. Данные, представленные на Рисунке 7Б, показывают, что увеличение концентрации спейсерной РНК от 0.5 до 5 мкМ вызывает дозозависимое расщепление ДНК с помощью RbCas12a. Достаточно эффективное разрезание ДНК (~ 50 % после 20-минутной инкубации) удалось наблюдать при соотношении RbCas12a:crРНК 1:500.

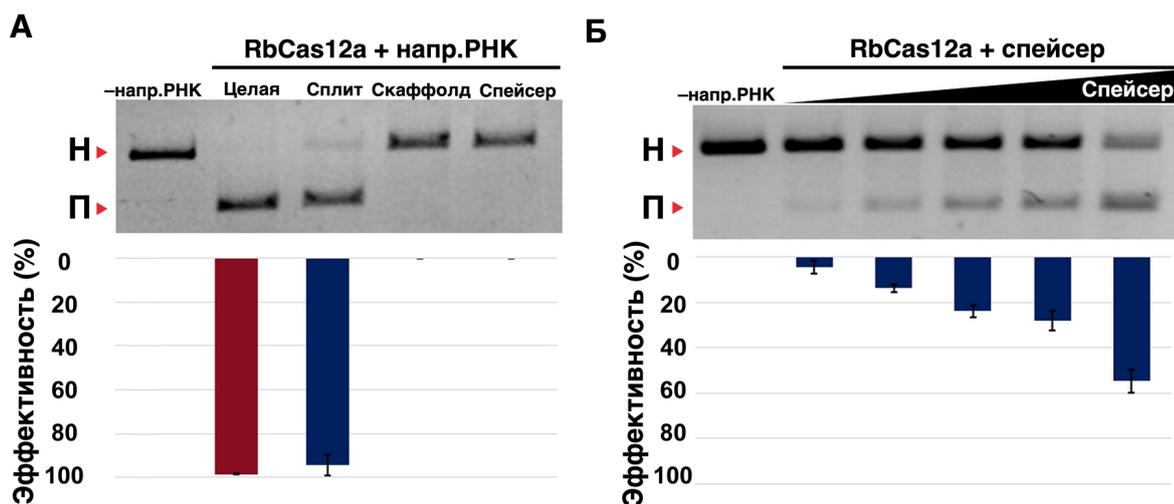


Рисунок 7 - Влияние сплит crРНК и отдельно спейсера crРНК на активность разрезания RbCas12a дцДНК матрицы. (А) Разрезание ДНК с помощью RbCas12a, заряженного 0.3 мкМ целых, сплит или отдельных скаффолдных и спейсерных частей crRNA при соотношении ДНК:RbCas12a:РНК 1:3:30. **(Б)** Разрезание ДНК с помощью RbCas12a, заряженного с увеличением концентрации (0.5, 1, 2, 2.5 и 5 мкМ) спейсерной РНК, что соответствует соотношениям 1:5:50, 1:5:100, 1:5:200, 1:5:250 и 1:5:500. Показана средняя эффективность разрезания в % по трем независимым экспериментам со стандартными отклонениями. Н., П., напр.РНК — соответственно непорезанные, порезанные фрагменты ДНК и направляющая (crРНК).

Для подтверждения протоспейсер-независимого характера разрезания ДНК со сплит crРНК проводили разрезания *in vitro* на трех различных мишенях дцДНК. В качестве контроля использовали полноразмерные crРНК (Рисунок 8).

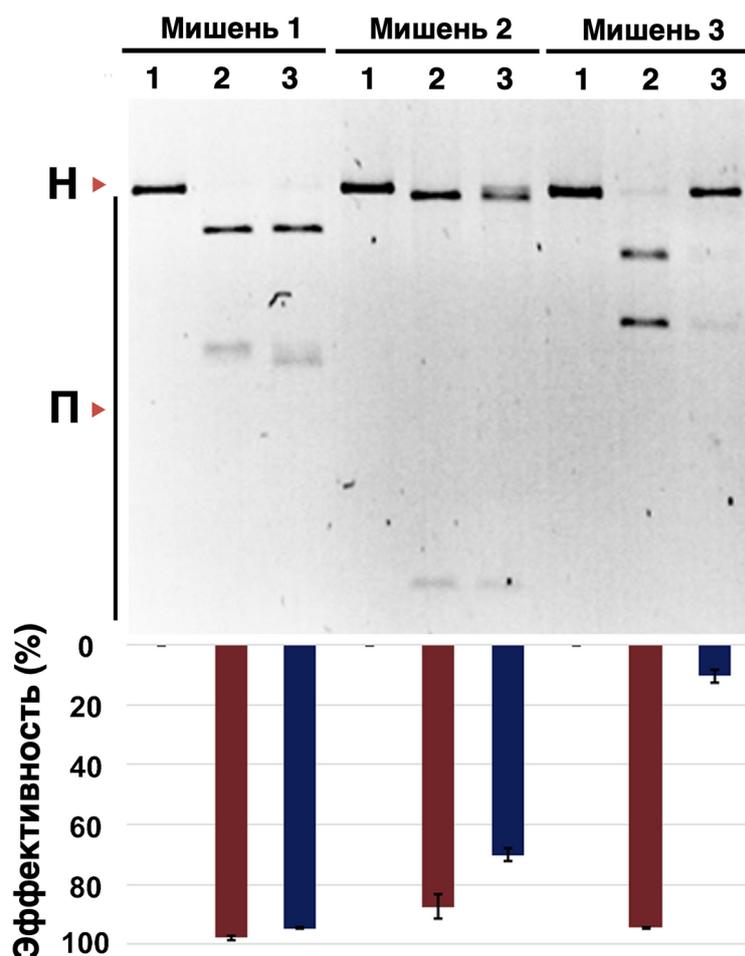


Рисунок 8 - Эффект сплит-версии crRNA на *in vitro* разрезание с помощью RbCas12a. Верх: электрофореграмма продуктов разрезания с использованием трех различных crRNA, их полноразмерной и сплит-версии. В качестве отрицательного контроля проводили реакции без добавления crRNA. Низ: относительные эффективности разрезания в % и стандартные отклонения по трем независимым экспериментам. Н., П.— соответственно непорезанные, порезанные фрагменты ДНК. 1, 2, 3 – соответственно разрезания без crRNA, с использованием полноразмерной crRNA и сплит-версии crRNA.

Далее проверяли, способен ли другой белок из семейства Cas12a осуществлять гидролиз матрицы, при укорочении длины crRNA. Выяснилось, что нуклеазная активность ортолога AsCas12a сохраняется после значительного уменьшения длины скаффолдной части crRNA и может быть обнаружена даже при полном удалении скаффолда, то есть в присутствии 20-нуклеотидной РНК, соответствующей только спейсерной части. Кроме того, остаточная нуклеазная активность AsCas12a, заряженного спейсерной РНК, может быть доведена до уровней, типичных для эффекторов, заряженных полноразмерной crRNA, путем добавления *in trans* скаффолдной части РНК из 20 оснований (Рисунок 9).

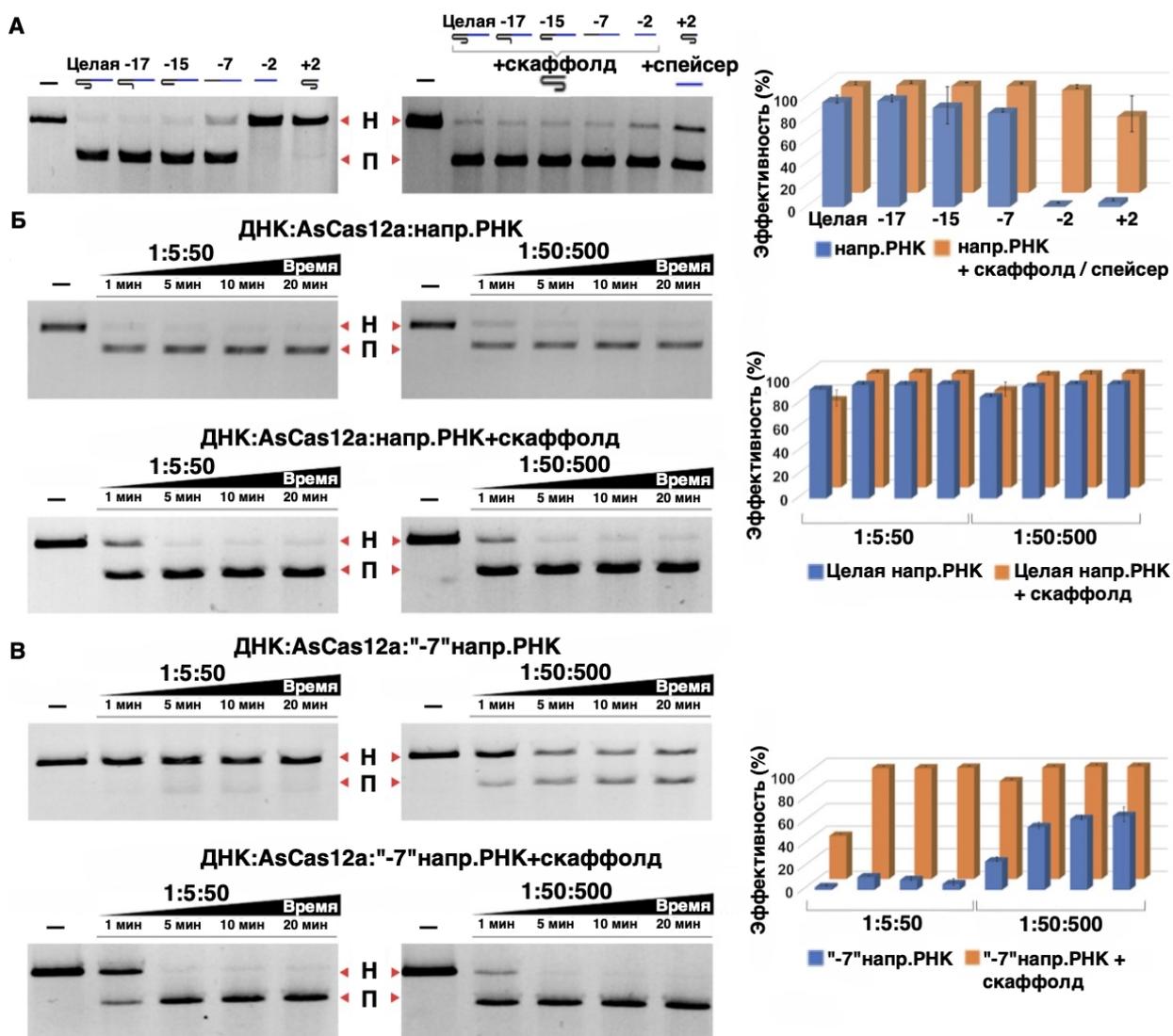


Рисунок 9 - Результаты анализа разрезания ДНК *in vitro* с помощью AsCas12a, запрограммированного укороченными crРНК. (А) Реакции содержали 10 нМ ДНК, 500 нМ AsCas12a и 5 мкМ указанных crРНК (соотношение 1:50:500). Показаны только неразрезанная ДНК и один из продуктов разрезания (2776 п.н.). На правом геле в реакции добавляли 5 мкМ скаффолдной или спейсерной РНК. Справа показаны средние эффективности разрезания по четырем независимым экспериментам и стандартные отклонения. (Б-В) Кинетика и концентрационная зависимость разрезания ДНК с помощью AsCas12a, заряженного полноразмерными (целыми) (Б) или «-7» (В) crРНК. На гелях отражены результаты анализов разрезов, проведенных в присутствии эквимольных количеств скаффолдной части crРНК. Соотношения ДНК:AsCas12a:crРНК подписаны над каждым гелем. Средние значения эффективности разрезания и стандартные отклонения по трем независимым экспериментам показаны справа. Н., П., напр.РНК — соответственно непорезанные, порезанные фрагменты ДНК и направляющая (crРНК). Дорожки с реакциями без добавления crРНК в качестве контроля обозначены как «—».

Далее анализировалась способность AsCas12a осуществлять направленный гидролиз как оц-, так и дцДНК при формировании комплекса со сплит-версией crРНК по сравнению с полноразмерной crРНК. В качестве матрицы использовали ДНК фага PhiX174.

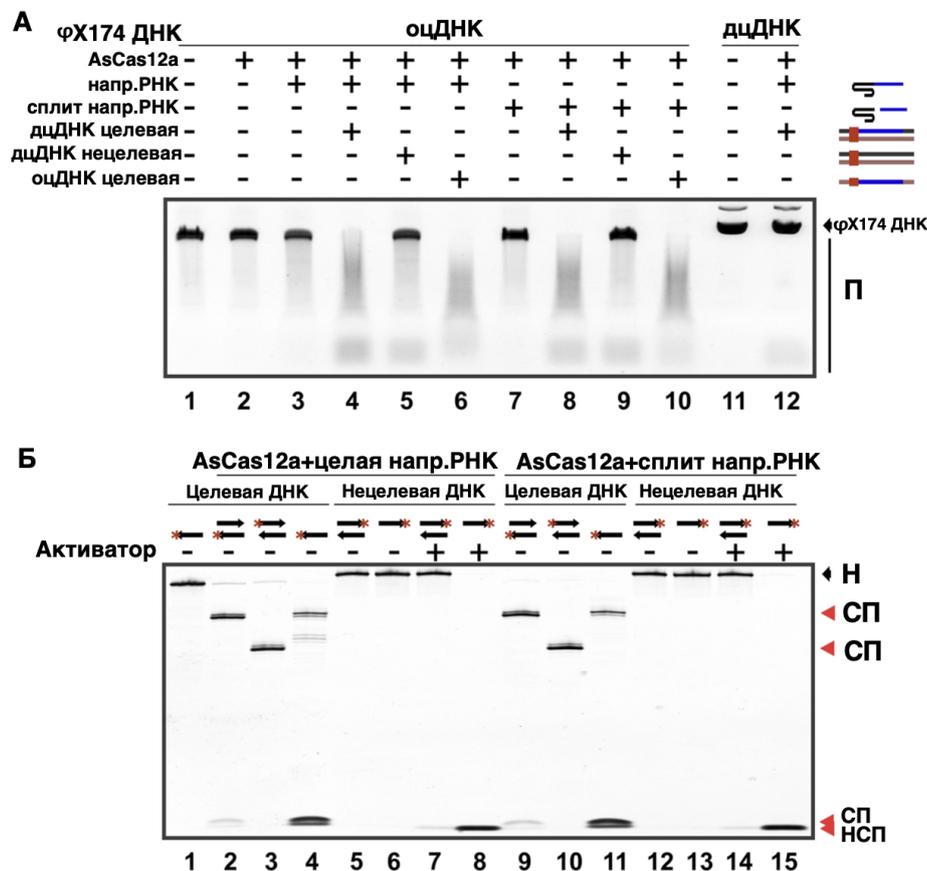


Рисунок 10 - Все известные активности разрезания ДНК с помощью AsCas12a поддерживаются сплит-версией РНК. (А) Электрофореграмма *in vitro* разрезов оц- и дцДНК в присутствии и отсутствии различных компонентов реакции с AsCas12a. **(Б)** Меченные флуоресцеином (красная звездочка) целевые и нецелевые ДНК-субстраты в одноцепочечной или двухцепочечной форме (схематически показаны сверху геля) в реакциях с AsCas12a, заряженного с помощью полноразмерной (целой) или сплит-версии crРНК. «Активатор» представляет собой немеченый одноцепочечный олигонуклеотид, содержащий комплементарную спейсеру РНК и несущий функциональный РАМ. Продукты реакции разделяли в денатурирующем 10 % ПААГ. Указаны непорезанные (Н), порезанные (П), порезанные специфически (СП) и неспецифически (НСП) фрагменты ДНК, напр.РНК — направляющая (crРНК).

Как видно из Рисунка 10, разрезание ДНК PhiX174 было индуцировано распознаванием двуцепочечных и одноцепочечных мишеней нуклеазой

AsCas12a, заряженной контрольной полноразмерной crРНК (дорожки 4 и 6, соответственно) или с помощью сплит-версии crРНК (дорожки 8 и 10). Расщепление PhiX174 не наблюдалось в реакциях без целевой ДНК (дорожки 3 и 7) или в присутствии нецелевой ДНК (дорожки 5 и 9). AsCas12a, заряженный либо полноразмерной, либо сплит-версией crРНК, разрезает двуцепочечные и одноцепочечные мишени в одних и тех же положениях (дорожки 2 и 4 и дорожки 9 и 11, соответственно). Помимо полос, возникающих в результате сайт-специфического разрезания оцДНК, продукты деградации в результате неспецифической активности наблюдались как с полноразмерными crРНК, так и со сплит-заряженными эффекторами. При активации немеченым таргетным субстратом (активатором), AsCas12a, заряженный полноразмерной или сплит РНК, полностью разрушает меченую оцДНК (дорожки 8 и 15), оставляя дцДНК нетронутой (полосы 7 и 14).

Для того, чтобы узнать, является ли способность связывать сплит-версию crРНК общим свойством нуклеаз V-A подтипа, проводили разрезания *in vitro* с использованием FnCas12a и LbCas12a (Рисунок 11).

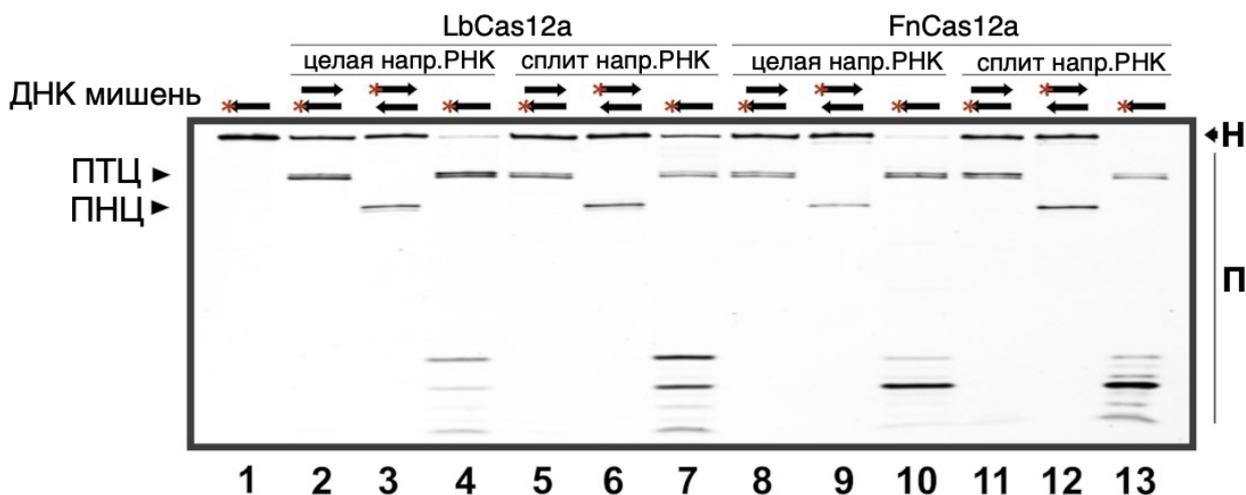


Рисунок 11 - Эффект сплит-версии crРНК на *in vitro* разрезание с помощью LbCas12a и FnCas12a. Красной звездочкой обозначена цепь ДНК, меченная флуоресцеином. Непорезанные (Н) и порезанные (П) фрагменты разделяли в 10 % ПААГ, напр.РНК — направляющая (crРНК). ПТЦ, ПНЦ – соответственно порезанная таргетная и нетаргетная цепи.

Наличие продуктов разрезания на гелях, представленных на Рисунке 8, Рисунке 9, Рисунке 10 и Рисунке 11 свидетельствует о том, что нуклеазная активность при связывании сплит-версии crРНК сохраняется у всех исследованных ортологов Cas12a в условиях *in vitro* и является, по всей вероятности, общим свойством нуклеаз подтипа V-A. Хотя сплит-версии

направляющих РНК работают в более высоких концентрациях, чем полноразмерные crРНК, можно допустить, что использование таких РНК для специфического распознавания и гидролиза ДНК-матриц открывает возможности для будущих механистических исследований и практического применения эффекторов Cas12a, которые требуют мультиплексирования и/или одновременного связывания нескольких целевых молекул.

9. Формирование делеций в геноме клеток линии НЕК293

Чтобы проверить активность кодон-оптимизированного RbCas12a в клетках млекопитающих, вносили делеции в ген *DNMT1* культуры клеток НЕК293Т с использованием пары crРНК, закодированных в отдельных плазмидных векторах, и RbCas12a, закодированного в третьем векторе. ДНК-метилтрансфераза 1 (DNMT1) играет основную роль в метилировании ДНК, а также участвует в процессах репликации ДНК и клеточной пролиферации. Мутации в гене *DNMT1* ассоциированы с возникновением аутосомно-доминантной мозжечковой атаксии, глухоты и нарколепсии, а также наследственной сенсорной нейропатией с глухотой и деменцией. Нокаутные линии по гену *DNMT1* являются важными моделями этих заболеваний.

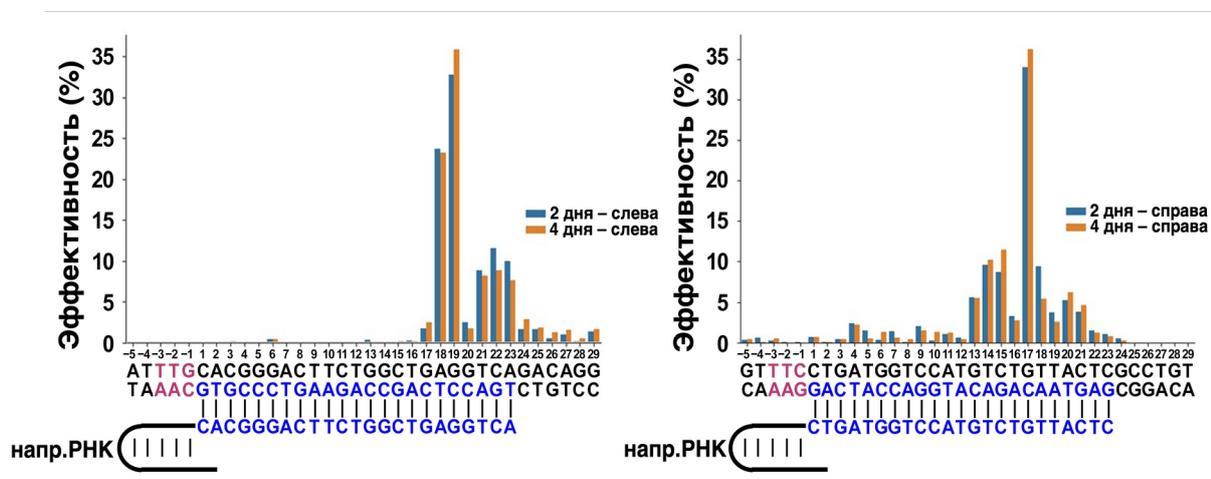


Рисунок 12 - Частоты сайтов разрыва, фланкирующих делецию в *DNMT1*.

Слева: последовательность crРНК снизу, которая в комплексе с RbCas12a формирует разрывы в левом плече. Справа: последовательность crРНК снизу, которая в комплексе с RbCas12a формирует разрывы в правом плече. Последовательности РАМ на обеих цепях выделены розовым цветом. Протоспейсерная часть таргетной цепи ДНК и спейсер crРНК выделены синим цветом. Относительная глубина покрытия в редактируемой области рассчитывалась как отношение глубины покрытия в определенной позиции к средней глубине покрытия за пределами редактируемой области. Эффективность разрыва цепи в определённой позиции была определена как

разность между относительным покрытием в заданной и последующей позициях. Напр. РНК — направляющая (сгРНК).

Клетки линии НЕК293Т снимали через 48 и 96 часов после трансфекции, после чего проверяли наличие делеций методом ПЦР. Далее проводили высокопроизводительное секвенирование ампликонов, содержащих делеции. Поскольку каждое прочтение состоит из двух фрагментов ДНК, фланкирующих делецию, связанных друг с другом, мы называем сайты лигирования «сайтами разрыва». Это место перехода от одного протоспейсера к другому. Как видно из Рисунка 12, наиболее распространенные сайты разрыва находятся между 19-м и 20-м нуклеотидами в области протоспейсера левого плеча делеций и между 17-м и 18-м нуклеотидами в зоне протоспейсера правого плеча делеций. Анализ показал, что эффективность вышеупомянутых сайтов разрыва составляла ~ 35 % как для левой, так и для правой сгРНК в условиях эксперимента. Интересно отметить, что распределение сайтов разрывов внутри протоспейсера для правой направляющей РНК выглядит унимодальным, а для левой - бимодальным. Это может быть следствием разницы в эффективности разрезания матрицы ДНК комплексом RbCas12a в разных сайтах. Другое пиковое значение для сайтов разрыва в протоспейсере, обозначенном слева, составляло ~ 23 % и локализовалось между 18-й и 19-й парой нуклеотидов. Реакции разрезания проходили на ~ 10 % как для «левой», так и для «правой» сгРНК, но в разных сайтах: через 21, 22 и 23 пары нуклеотидов внутри протоспейсера для «левой» направляющей РНК и через 14 и 15 пар нуклеотидов для «правой». Между вторым и четвертым днями эксперимента нет значительной разницы (p -значение критерия хи-квадрат $\gg 0,05$) в эффективности внесенных разрывов. Также нет статистической связи между разрывами в двух протоспейсерах в эксперименте.

Таким образом, исходя из полученных данных, RbCas12a активен в клетках человека и способен формировать значительные делеции длиной как минимум в 471 нуклеотид в геноме культуры клеток человека.

ВЫВОДЫ

1. По результатам биоинформатического поиска в геноме *Ruminococcus bromii* sp. обнаружен ген ранее не описанной в литературе нуклеазы семейства Cas12a (RbCas12a) длиной 3735 п.о.

2. Показана активность рекомбинантного белка RbCas12a в составе комплексов с различными сгРНК в реакциях разрезания *in vitro* на различных ДНК-матрицах, содержащих РАМ-последовательности с консенсусом 5'-YYN.

3. Нуклеаза RbCas12a способна использовать в качестве кофактора ионы магния и марганца, а также, хоть с меньшей эффективностью, ионы кобальта, цинка и никеля. Рабочий температурный диапазон RbCas12a варьирует от 20 до 42 °С, с наибольшей активностью в диапазоне 30-42 °С.

4. Нуклеаза RbCas12a способна проявлять коллатеральную активность в отношении оцДНК.

5. Определена природа выступающих концов двуцепочечных продуктов гидролиза ДНК-матрицы с помощью нуклеазы RbCas12a: продукты имеют 5'-концевые выступы, длина которых зависит от экзонуклеазной активности фермента в отношении таргетной цепи и может составлять 4-6 нуклеотидов.

6. Нуклеазы Cas подтипа V-A (RbCas12a, AsCas12a, FnCas12a, LbCas12a) формируют комплексы с двумя молекулами РНК, представляющими из себя отдельные спейсерную и скаффолдную части crРНК (сплит РНК), и осуществляют в их составе направленный гидролиз ДНК-матриц.

7. С помощью RbCas12a получена направленная делеция в гене *DNMT1* культуры клеток HEK293T.

Список работ, опубликованных по теме диссертации

Соискатель имеет 3 опубликованных работы по теме диссертации в рецензируемых научных изданиях, рекомендованных для защиты в диссертационном совете МГУ по специальности 1.5.3 - Молекулярная биология:

1. **Vasilev, R. A.**, Chernikovich, V. Y., Evteeva, M. A., Sakharov, D. A., & Patrushev, M. V. Synthetic biology: current state and applications // *Molecular Genetics, Microbiology and Virology* – 2021.– V.36. No.1.– P.15-26. IF 0.352 (Web of Science)
2. Shebanova, R., Nikitchina, N., Shebanov, N., Mekler, V., Kuznedelov, K., Ulashchik, E., **Vasilev, R.**, Sharko, O., Shmanai, V., Tarassov, I., Severinov, K., Entelis N. & Mazunin I. Efficient target cleavage by Type V Cas12a effectors programmed with split CRISPR RNA // *Nucleic acids research* – 2022.– V.50. No.2.– P.1162-1173. IF 16.971 (Web of Science)
3. **Vasilev, R.**, Gunitseva, N., Shebanova, R., Korzhenkov, A., Vlaskina, A., Evteeva, M., Polushkina, I., Nikitchina, N., Toshchakov, S., Kamenski, P., Patrushev, M. & Mazunin, I. Targeted modification of mammalian DNA by a novel Type V Cas12a endonuclease from *Ruminococcus bromii* // *International Journal of Molecular Sciences* – 2022.– V.23. No.16.– P.9289. IF 5.923 (Web of Science)