

УДК 577.29

БИОЛОГИЧЕСКАЯ РОЛЬ И ЭВОЛЮЦИОННАЯ ИСТОРИЯ СИСТЕМ CRISPR-Cas ТИПА III

Обзор

© 2021 М.В. Колесник^{1*}, Я.В. Федорова^{2,3}, К.О. Корнеева¹,
Д.Н. Артамонова¹, К.В. Северинов^{1,3,4*}

¹ Сколковский институт науки и технологий, Центр наук о жизни, 121205 Москва, Россия; электронная почта: matveykolesnik@gmail.com; severik@waksman.rutgers.edu

² Санкт-Петербургский политехнический университет Петра Великого, 195251 Санкт-Петербург, Россия

³ Институт биологии гена РАН, Центр высокоточного редактирования и генетических технологий для биомедицины, 119334 Москва, Россия

⁴ Институт микробиологии Ваксмана, 08854 Нью-Джерси, США

Поступила в редакцию 01.06.2021

После доработки 24.08.2021

Принята к публикации 30.08.2021

Возникновение паразитических генетических элементов свойственно всем живым системам. В ходе эволюции у клеточных организмов появилось множество систем защиты, которые препятствуют распространению паразитических элементов. Широко распространённые у прокариот системы CRISPR-Cas позволяют специфически распознавать и уничтожать чужеродные нуклеиновые кислоты. Механизм работы систем CRISPR-Cas основывается на приобретении и хранении фрагментов чужеродных нуклеиновых кислот и их использовании для разрушения чужеродного комплементарного генетического материала эффекторными рибонуклеопротеиновыми комплексами. Системы CRISPR-Cas активно изучаются, поскольку некоторые из них используются для геномного редактирования. Системы CRISPR-Cas типа III широко распространены, но мало изучены из-за повышенной в сравнении с системами других типов сложности механизма их работы. Эффекторные комплексы систем типа III специфически распознают и расщепляют молекулы РНК. Распознавание РНК-мишени активирует субъединицу эффектора – так называемую «CRISPR-полимеразу», которая расщепляет ДНК и производит небольшие циклические олигонуклеотиды. Последние активируют ряд вспомогательных эффекторов, в частности, неспецифические рибонуклеазы. В этом обзоре рассмотрена история исследований систем CRISPR-Cas типа III, а также суммированы текущие данные о структуре и активности компонентов этих систем, их биологической роли и эволюционной истории. С помощью структурного моделирования с использованием программы AlphaFold2 показано, что консервативный ранее не охарактеризованный белок HRAMP систем архей является вырожденным гомологом белка Cas10 CRISPR-Cas типа III, что позволяет предположить, что системы HRAMP и системы CRISPR-Cas типа III имеют общую эволюционную историю.

КЛЮЧЕВЫЕ СЛОВА: системы CRISPR-Cas типа III, защитные системы прокариот, сигнальные олигонуклеотиды, эволюция систем CRISPR-Cas, системы HRAMP.

DOI: 10.31857/S0320972521100134

ВВЕДЕНИЕ

Все известные формы жизни на Земле обладают универсальным свойством: практически все их биологические признаки определены

последовательностями нуклеиновых кислот, воспроизводимыми по матричному принципу. Хотя изменения в последовательностях нуклеиновых кислот могут оказывать очень сильное влияние на биологические признаки (например, на структуры белков, которые они кодируют), сами эти изменения, как правило, не оказывают влияния на химические и физические свойства нуклеиновых кислот, что делает возможным существование молекул нуклеиновых кислот с практически любыми возможными последовательностями. Поскольку повышение точности репликации может быть достигнуто

Принятые сокращения: OAC – олигоденилатсинтетазы; цГАС – синтаза циклического ГМФ-АМФ; цОА – циклические олигоденилаты; CARF – CRISPR-Cas Associated Rossmann Fold; crPНК – CRISPR-PНК; HEPN – Higher Eukaryotes and Prokaryotes Nucleotide-binding; PAM – Protospacer Adjacent Motifs; RAMP – Repair-Associated Mysterious Proteins; RT – обратная транскриптаза.

* Адресат для корреспонденции.

только за счёт повышения затрат энергии [1], процесс репликации всегда подвержен ошибкам, из-за которых возникает изменчивость биологических признаков, кодируемых нуклеиновыми кислотами. Эта изменчивость является объектом действия естественного отбора и/или генетического дрейфа и, таким образом, обеспечивает эволюционный процесс.

В любом сообществе эволюционирующих самовоспроизводящихся систем неизбежно появляются, с одной стороны, генетические паразиты, а с другой – защитные системы для борьбы с ними [2]. Поэтому «гонка вооружений» между эгоистичными генетическими элементами и их хозяевами, по-видимому, началась с момента возникновения жизни. В ходе этой гонки у хозяев развиваются системы защиты, которые нацелены на различные механизмы, существенные для жизненного цикла паразита. В свою очередь, паразиты вырабатывают механизмы, позволяющие им избежать действия защитных систем хозяина. Полностью избавиться от генетических паразитов, по всей видимости, невозможно, поскольку это снижает интенсивность горизонтального переноса генов, необходимого для поддержания стабильности генома и эволюционного процесса [2]. Эволюция защитных систем подчас идёт очень причудливыми путями и включает «перетасовку» компонентов разных защитных систем, а также захват генов самих генетических паразитов. С этой точки зрения системы CRISPR-Cas особенно интересны, поскольку они представляют собой сложную химеру, состоящую из «прирученных» транспозонов, токсин-антитоксиновых систем и других компонентов неясного происхождения. В этом обзоре мы рассмотрим структурные и механистические особенности систем CRISPR-Cas типа III, которые являются самыми сложными из известных на текущий момент защитных систем прокариот.

Разнообразие систем CRISPR-Cas и механизмов опосредованного ими иммунитета были подробно описаны во множестве обзоров (см., например, [3, 4]). Большинство локусов CRISPR-Cas содержат кассеты CRISPR, которые состоят из идентичных повторов, разделённых уникальными спейсерами, а также ассоциированные с кассетами CRISPR кластеры белок-кодирующих генов *cas*. Механизм CRISPR-Cas-опосредованного иммунитета можно разделить на три стадии: а) адаптация, б) экспрессия и в) интерференция (на рис. 1 приведена схема этих стадий для систем CRISPR-Cas типа III). На стадии адаптации короткие фрагменты ДНК встраиваются в CRISPR-кассету с формированием нового спейсера. Встраивание новых

спейсеров производится комплексом белков Cas1–Cas2. Этот комплекс наиболее консервативен и присутствует во всех полных системах CRISPR-Cas, однако детали механизма адаптации зависят от типа системы. На стадии экспрессии CRISPR-кассеты транскрибируются с образованием молекул пре-CRISPR-РНК, которые процессируются с образованием зрелых CRISPR-РНК (сгРНК). Этап созревания сгРНК опосредуется различными механизмами, которые зависят от типа системы CRISPR-Cas. На стадии интерференции сгРНК взаимодействуют с Cas-белками, формируя эффекторные комплексы, которые распознают и расщепляют нуклеиновые кислоты, комплементарные сгРНК. Системы CRISPR-Cas закодированы в ~90% последовательностей геномов архей и ~40% геномов эубактерий [5]. По составу эффекторных комплексов системы CRISPR-Cas можно разделить на два класса. Эффекторы систем класса 1 представляют собой комплексы, состоящие из многих белков; эффекторы класса 2 состоят из одного большого многодоменного белка в комплексе с сгРНК. На основе анализа состава и организации *cas*-локусов системы CRISPR-Cas были подразделены на 6 типов и несколько подтипов; типы I, III и IV относятся к классу 1, а типы II, V и VI относятся к классу 2 [4]. В этом обзоре будут подробно рассмотрены только системы типа III.

ОБЗОР ИСТОРИИ ИССЛЕДОВАНИЙ СИСТЕМ CRISPR-CAS ТИПА III

Системы CRISPR-Cas типа III широко распространены в геномах бактерий и архей, составляя 34% и 25% полных локусов CRISPR-*cas*, обнаруженных в геномах [5]. Впервые последовательности *cas*-генов, относящихся к системам типа III, были обнаружены в процессе поиска консервативных кластеров генов в отсеквенированных на тот момент геномах гипертермофильных архей. В то время было выдвинуто предположение, что эти гены принадлежат к ранее неизвестной системе репарации ДНК [6]. Среди генов, содержащихся в обнаруженных кластерах, Макарова et al. [6] обнаружили группу генов, кодирующих большие консервативные белки, содержащие Palm-домены. Домены Palm входят в состав некоторых ферментов, осуществляющих матричный синтез полинуклеотидов, а также в состав различных циклаз и нуклеотидилтрансфераз [7, 8]. Поэтому было сделано предположение, что эти, содержащие Palm-домены, белки также являются полимеразами или циклазами. Также было обнаружено несколько се-

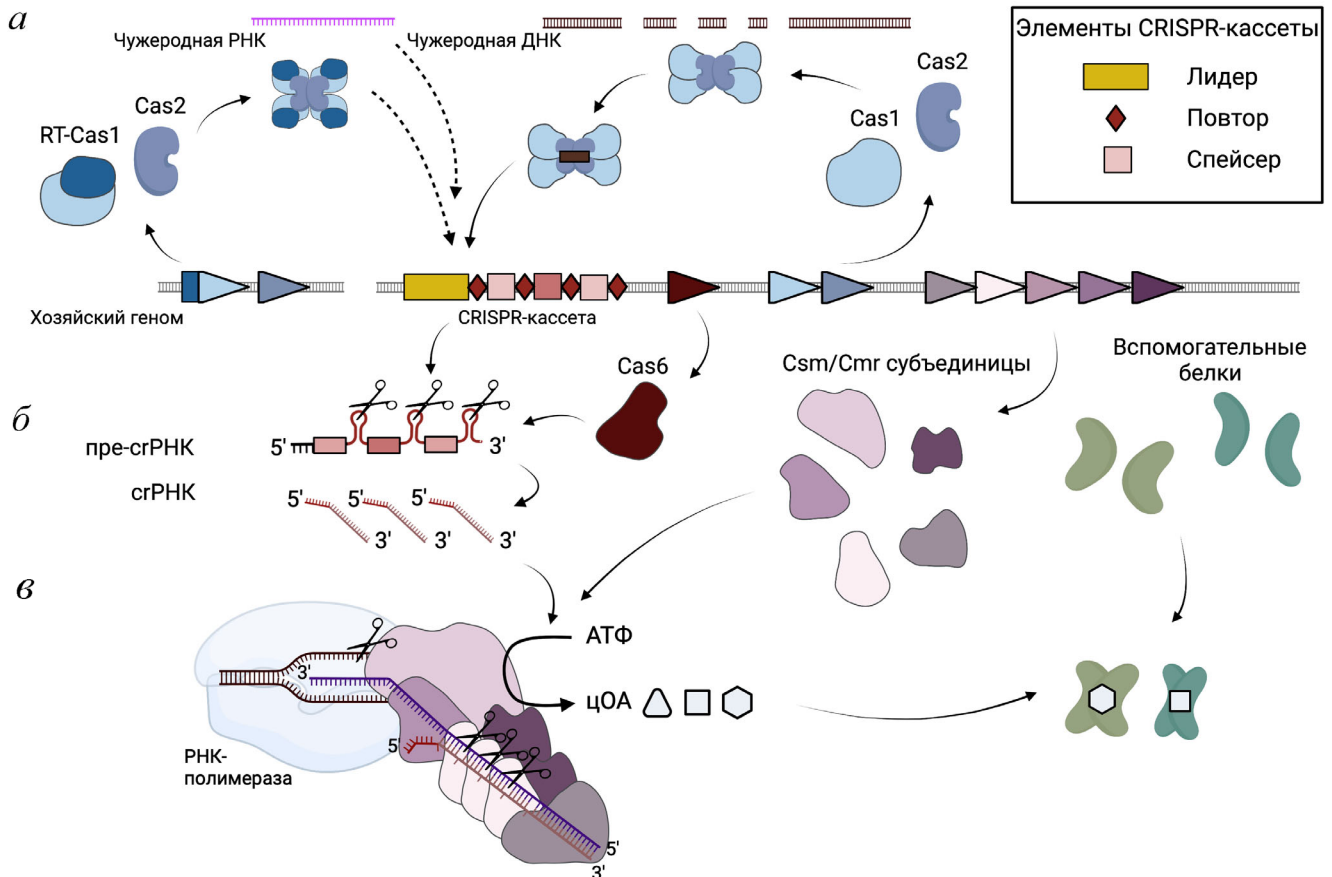


Рис. 1. Схема иммунного ответа в системах CRISPR-Cas типа III. *a* – Адаптация – встраивание небольших фрагментов нуклеиновых кислот мобильного генетического элемента в CRISPR-кассету хозяина. В некоторых системах спейсеры могут приобретаться из РНК за счёт активности домена обратной транскриптазы (RT), слитого с белком Cas1. *б* – Экспрессия; созревание малых crRNA и формирование эффекторных комплексов. *в* – Интерференция; специфическое распознавание чужеродной РНК запускает иммунный ответ

мейств других белок-кодирующих генов, ассоциированных с генами предполагаемых полимераз/циклаз, но роль этих белков оставалась неясной. Предполагая, что обнаруженные локусы относятся к новой системе репарации ДНК, Makarova et al. [6] назвали эти белки Repair-Associated Mysterious Proteins, или RAMP. Кроме локусов, кодирующих полимеразу/циклазу и RAMP, Makarova et al. [6] обнаружили ещё одну группу консервативных кластеров генов, которые сейчас известны как системы CRISPR-Cas типа I. Тем не менее ассоциация обнаруженных кластеров с CRISPR-кассетами на тот момент не была установлена. Несколько лет спустя Haft et al. [9] провели масштабный анализ белок-кодирующих генов, ассоциированных с CRISPR-кассетами, и попытались классифицировать эти гены по организации в локусы. Среди обнаруженных локусов CRISPR-Cas Haft et al. [9] описали локусы типа Csm (CRISPR-Cas Subtype *Mycobacterium tuberculosis*)

и локусы типа Cmr (CRISPR-Cas Module RAMP). Между некоторыми белками, закодированными в локусах *csm* и *cmr*, было обнаружено сходство – оба локуса кодируют белки, содержащие Palm-домены (*csm1* и *cmr2*), и как минимум два белка семейства RAMP (*csm3* и *cmr4*). Позже локусы *csm* и *cmr* стали известны как системы типа III-A и типа III-B соответственно [10]. Haft et al. [9] также отметили, что локус *cmr* никогда не бывает единственной системой CRISPR-Cas в прокариотическом геноме. Позже было показано, что в системах CRISPR-Cas типа III-B, как правило, отсутствует модуль адаптации, и поэтому они используют адаптационные модули, закодированные другими системами CRISPR-Cas [5].

Системы CRISPR-Cas типа III независимо изучались в нескольких научных группах, и поначалу результаты этих исследований были противоречивыми. Marraffini et al. [11], исходя из результатов экспериментов *in vivo*, охарактери-

зовали систему CRISPR-Cas типа III-A из *Staphylococcus epidermidis* как специфичную к ДНК. С другой стороны, было показано, что эффекторный комплекс типа III *in vitro* специфически распознают и расщепляют РНК [12, 13]. Хотя сейчас известно, что выводы, сделанные на основе результатов *in vivo*, были неверными, подробный разбор этих исследований важен не только в рамках изучения систем CRISPR-Cas, но и с точки зрения научной методологии. Marraffini et al. [11] для своего эксперимента использовали два штамма *S. epidermidis* и конъюгативную плазмиду, способную реплицироваться в клетках этих бактерий. Один из штаммов содержал систему CRISPR-Cas типа III-A со спейсером, совпадающим по последовательности с участком плазмидного гена *nes*, кодирующего никазу, которая является частью механизма конъюгации. Было показано, что система типа III-A подавляет конъюгативный перенос плазмиды, несущей ген *nes*. Так как последовательность спейсера совпадала с последовательностью кодирующей цепи гена *nes* (т.е. сгРНК не могла распознать смысловой транскрипт за счёт комплементарных взаимодействий), был сделан вывод, что мишенью для системы типа III-A является ДНК, а не РНК. Тем не менее транскрипт *nes* может распознаваться сгРНК, транскрибированными с противоположной цепи CRISPR-кассеты, однако таких сгРНК не обнаружили. Также известно, что для конъюгативного переноса плазмиды экспрессия никазы необходима лишь в клетках донора, и расщепление транскрипта *nes* в клетках реципиента не должно снижать эффективность конъюгации, что дополнительно подтверждает предположение о специфичности систем типа III-A к ДНК. Чтобы окончательно подтвердить ДНК-специфичность системы типа III-A, авторы провели изящный эксперимент, внедрив в последовательность протоспейсера самосплайсирующийся интрон и показав, что такая плазида избегает иммунного ответа, опосредованного системой типа III-A. Однако они не рассмотрели возможность того, что 1) ген *nes* в клетке транскрибируется и с антисмысловой цепи, 2) иммунитет, опосредованный системами CRISPR-Cas типа III, может запускаться, когда эффектор узнает новосинтезирующиеся РНК. Несколько лет спустя в той же лаборатории показали, что протоспейсеры, закодированные в исследуемых плазидах, транскрибируются в обоих направлениях, и специфичность систем типа III-A к РНК *in vivo* была окончательно подтверждена [14].

Несмотря на специфичность к молекулам РНК, системы CRISPR-Cas типа III защищают

клетки от вирусов с ДНК-геномами и подавляют трансформацию плазидами, если ДНК вируса или плазмиды транскрибируется с образованием РНК, комплементарной сгРНК, даже если эти транскрипты-мишени не важны в жизненном цикле вируса или плазмиды [14, 15]. Изучение активностей эффекторных комплексов Csm и Cmr *in vitro* позволило отчасти объяснить этот странный феномен. Большие субъединицы эффекторных комплексов типа III (белки Csm1 и Cmr2, впоследствии названные Cas10), кроме Palm-доменов, содержат HD-домены (названные по присутствующим в них консервативным остаткам гистидина и аспартата) [5]. Белки с HD-доменами присутствуют и в других CRISPR-Cas-системах. Например, белок Cas3 систем типа I содержит HD-домен, обладающий нуклеазной активностью в отношении к одноцепочечной ДНК (оцДНК); в ходе иммунного ответа за счёт активности HD-домена белок Cas3 разрушает ДНК-мишени [16]. В эффекторных комплексах типов III-A [17] и III-B [18, 19] связывание РНК-мишени активирует HD-домены субъединиц Cas10, и такие активированные комплексы способны расщеплять одноцепочечную ДНК *in vitro*. С учётом этих наблюдений была предложена модель «котранскрипционного расщепления ДНК». Согласно этой модели, эффекторы систем типа III распознают новосинтезированные транскрипты. При этом происходит активация HD-доменов субъединиц Cas10, за счёт которых происходит расщепление однострессовой ДНК в области «котранскрипционного пузыря» (рис. 2, а) [17]. Котранскрипционное расщепление ДНК было обнаружено в экспериментах *in vitro* группой Marraffini [20], однако, согласно более поздним данным, наблюдаемое *in vitro* котранскрипционное расщепление ДНК на деле может быть следствием ошибочной интерпретации результатов эксперимента [21]. Вразрез с моделью котранскрипционного расщепления ДНК было показано, что мутации каталитических остатков в HD-доме Cas10 не влияют на подавление трансформации клеток плазидами, в то время как инактивация Palm белка Cas10 значительно ослабляет антиплазмидный иммунный ответ [22]. Из этих результатов ясно, что, несмотря на все изящество и правдоподобность модели котранскрипционного расщепления ДНК, иммунный ответ систем типа III не может быть сведён только лишь к ней.

Локусы CRISPR-Cas часто содержат гены белков, напрямую не участвующих в процессах приобретения новых спейсеров, созревания сгРНК или формирования эффекторных комплексов. Роль большинства этих вспомогатель-

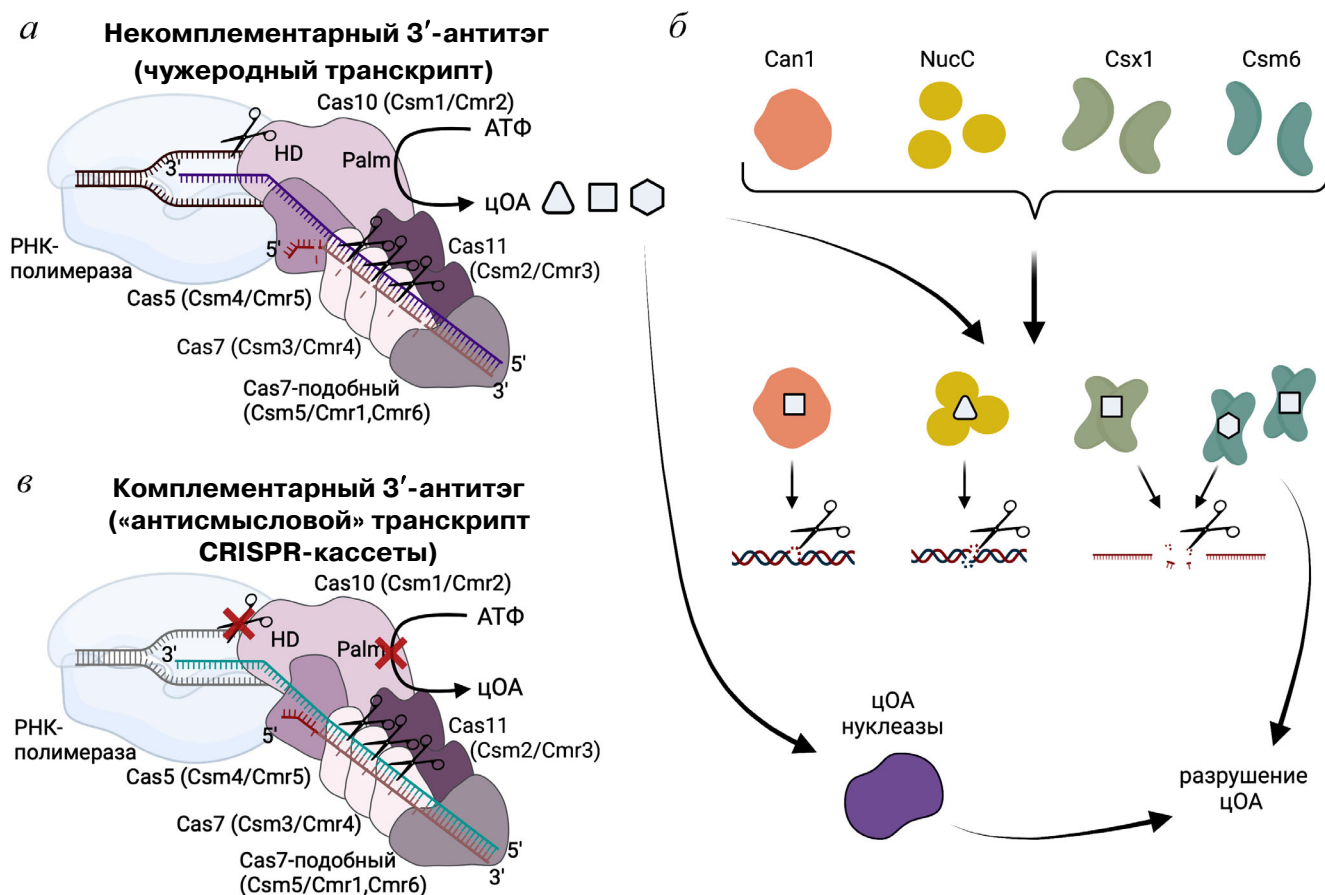


Рис. 2. Модель котранскрипционного расщепления ДНК эффекторами типа III с активацией вспомогательных нуклеаз, запускаемых при распознавании мишени. *а* – Распознавание транскрипта, комплементарного crРНК, активирует нуклеазную активность домена HD и полимеразную активность домена Palm субъединицы Cas10. *б* – Циклические олигоаденилаты (цОА), синтезированные за счёт активности Palm-домена, активируют вспомогательные эффекторы, неспецифически расщепляющие ДНК или РНК; активность вспомогательных эффекторов регулируется за счёт деградации цОА специализированными цОА-нуклеазами либо самими вспомогательными эффекторами. *в* – Предотвращение аутоиммунного ответа в системах CRISPR-Cas типа III: при наличии комплементарных взаимодействий между РНК-мишенью и «5'-тэгом» crРНК домены HD и Palm субъединицы Cas10 остаются неактивными

ных генов до сих пор неясна. Некоторые из таких генов кодируют белки, содержащие домен с особым вариантом укладки Россмана, называемым CARF (CRISPR-Cas Associated Rossmann Fold). В этих белках CARF-домены часто слиты с другими доменами с предсказанной РНКазной, ДНКазной или ДНК-связывающей активностями. Поскольку укладка Россмана часто встречается в нуклеотид-связывающих белках, было предположено, что белки с CARF-доменами могут распознавать некие лиганды нуклеотидной природы, активирующие эффекторные домены этих белков [23]. Вспомогательные белки важны для работы систем типа III: было показано, что повреждение генов *csx1* и *csm6*, которые кодируют CARF-домены и часто ассоциированы с оперонами *cas*-систем типа III, значительно снижает способность этих систем подавлять трансформацию плазмидами [15, 22].

В белках Csx1 и Csm6 домены CARF слиты с доменами HEPN (Higher Eukaryotes and Prokaryotes Nucleotide-binding) [23]. Белки с HEPN-доменами обычно обладают РНКазной активностью и часто входят в состав различных защитных систем [24]. Некоторые ортологи Csx1 и Csm6 способны расщеплять одноцепочечные РНК *in vitro* за счёт активности HEPN-доменов. Однако обнаруженная РНКазная активность была довольно слабой, поэтому было предположено, что во время иммунного ответа *in vivo* может быть задействована какая-то активация [25, 26]. Впоследствии это предположение полностью подтвердилось – было показано, что при распознавании РНК-мишени эффекторные комплексы типа III катализируют синтез циклических олигоаденилатов (цОА) за счёт полимеразной активности Palm-доменов Cas10 [27, 28]. Эти олигоаденилаты играют роль вторичных

мессенджеров, активируя неспецифическую РНКазную активность белков CARF–HEPN (рис. 2, а, б) [27–29]. Роль неспецифической РНКазной активности белков Csx1/Csm6 в иммунном ответе до сих пор не вполне ясна; предполагается, что результатами такой неспецифической дегградации РНК может быть подавление экспрессии вирусных транскриптов, прекращение роста/деления или даже гибель клетки [30, 31]. В последнем случае можно предположить, что подавление инфекции ДНК-вирусами может происходить по механизму abortивной инфекции без привлечения ДНКазной активности эффекторных комплексов систем типа III. Однако, если бы системы типа III функционировали по принципу систем abortивной инфекции, клетки, получившие спейсеры в процессе адаптации, не смогли бы дать потомства. Это соображение подтверждается экспериментами – иммунный ответ, опосредованный системами типа III, позволяет клеткам выжить после инфекции литическими бактериофагами [32].

Поскольку иммунный ответ систем типа III позволяет клеткам выжить после инфекции, активация вспомогательных РНКаз, вызванная цОА, должна быть временной. Существует два наиболее очевидных способа контролировать активность вспомогательных эффекторов – за счёт регуляции синтеза цОА и за счёт их расщепления. Оба пути регуляции были экспериментально подтверждены. Во-первых, синтез цОА, осуществляемый Cas10, активируется при связывании с РНК-мишенью и прекращается при её расщеплении [33, 34]. Во-вторых, были охарактеризованы несколько нуклеаз, способных расщеплять цОА. В некоторых организмах дегградация цОА осуществляется специализированными цОА-нуклеазами [35]; в других – белки с доменами CARF–HEPN способны сами дегградировать цОА, которые являются их собственными лигандами (рис. 2, б) [36, 37]. Кроме цОА-нуклеаз, задействованных в регуляции иммунного ответа, известна высокоактивная цОА-нуклеаза, кодируемая вирусом, заражающим клетки археи *Sulfolobus*; эта нуклеаза позволяет вирусу обойти иммунный ответ клетки-хозяина, опосредованный системой типа III [38].

Интересно, что цОА-зависимый сигнальный путь прокариотических иммунных систем схож с одним из путей врождённого иммунитета млекопитающих. Присутствие молекул двухнитевой РНК в цитоплазме клеток млекопитающих активирует олигоденилатсинтазу (ОАС), которая превращает молекулы АТФ в 2'-5'-олигоденилаты, активирующие РНКазу L, которая неспецифически расщепляет РНК в цитоплазме [39]. Сходство этих механизмов оказывается

ещё более примечательным, если учесть, что каталитический центр ОАС схож с Palm-доменом [40], а РНКазы L содержит в своём составе вариант HEPN-домена, который и обуславливает рибонуклеазную активность [24]. В последнее время этот путь врождённого иммунитета привлёк к себе много внимания, так как было обнаружено, что некоторые полиморфизмы в локусе, кодирующем ОАС, ассоциированы с дыхательной недостаточностью, вызванной COVID-19; судя по этим данным, вероятно, этот сигнальный путь играет важную роль в иммунном ответе против SARS-Cov-2 [41].

Если роль цОА-пути в иммунном ответе систем CRISPR-Cas типа III относительно понятна, функция HD-домена Cas10 остаётся неясной. В ранних исследованиях было показано, что мутации в каталитических остатках HD-домена белка Cas10 не влияют на опосредованное системой III-A подавление конъюгативного переноса плазмид между клетками *S. epidermidis* [22]. В этой работе эффективность конъюгации отслеживали по числу колоний трансконъюгантов, то есть колоний, образованных клетками, получившими плазмиду в ходе конъюгативного переноса. Но позже было показано, что система типа III-A с инактивированным HD-доменом Cas10 не препятствовала формированию колоний трансконъюгантов, но сильно замедляла их рост, так что такие колонии можно было обнаружить только спустя длительное время инкубации [31]. В случае, когда оба домена (HD и Palm) были инактивированы, интерференция не наблюдалась, и количество и внешний вид колоний трансконъюгантов, несущих таргетируемую и нетаргетируемую плазмиды, были одинаковыми [31]. Так, было установлено, что для эффективного подавления конъюгации требуется совместная активность обоих доменов белка Cas10. На деле, необходимость в цОА-зависимом пути интерференции типа III-A может зависеть от количества РНК-мишени: когда транскрипта, распознаваемого эффектором типа III-A, образуется много, активации HD-домена оказывается достаточно для эффективной интерференции; когда уровень транскрипции протоспейсера низок, необходимы активности обоих доменов белка Cas10. Вероятно, роль неспецифических цОА-активируемых РНКаз в системах типа III состоит в подавлении репликации мобильных генетических элементов, за счёт чего более медленная активность HD-домена Cas10 успевает разрушить их ДНК [31]. Этот механизм иммунного ответа предполагает ко-транскрипционное расщепление ДНК, однако чётких экспериментальных данных, подтверждающих существование этого явления, на теку-

ший момент нет. Котранскрипционное расщепление ДНК было описано лишь в экспериментах *in vitro* всего в одной статье [20]. В этих экспериментах субстраты для *in vitro* транскрипции были получены путём отжига двух комплементарных дезоксиолигонуклеотидов — матричного и нематричного. Однако более поздние исследования показали, что процесс, принятый за котранскрипционное расщепление ДНК, может происходить только в том случае, если олигонуклеотид, соответствующий нематричной, «таргетируемой» цепи ДНК, находится в избытке по отношению к олигонуклеотиду «не таргетируемой» цепи ДНК, комплементарной РНК-транскрипту [21]. Согласно этим исследованиям, в экспериментах Samai et al. [20] эффекторы типа III расщепляли не ДНК внутри «транскрипционного пузыря», а свободную однонитевую ДНК, присутствующую в реакционной смеси [21]. В качестве альтернативной гипотезы, объясняющей способность систем типа III защищать клетки от мобильных генетических элементов с ДНК-геномами так, чтобы клетки при этом остались в живых, можно предположить, что HD-домен Cas10 может расщеплять одноцепочечную ДНК, которая формируется на промежуточных стадиях репликации некоторых вирусов и плазмид [21]. Таким образом, можно заключить, что, хотя активность HD-домена белка Cas10 необходима для интерференции в системах типа III, роль этой активности остаётся неясной из-за отсутствия данных о мишенях ДНКазного домена Cas10 *in vivo*. В недавнем исследовании было показано, что активность HD-домена Cas10 повышает частоту возникновения мутаций в геноме бактерии, кодирующей систему CRISPR-Cas типа III, что косвенно свидетельствует о том, что HD-домен белка Cas10 способен неспецифически расщеплять клеточную ДНК [42].

Согласно представленным выше результатам, опосредованная системами типа III защита против чужеродной ДНК подразумевает слаженную работу HD- и Palm-доменов белка Cas10 вместе с цОА-активируемыми РНКазами. Однако во многих системах III-типа нередко отсутствуют какие-либо из этих компонентов. Вероятно, такие «неполные» системы либо задействуют какие-то механизмы, компенсирующие отсутствующие компоненты, либо функционируют иным путём.

Один из примеров такой неполной системы — система CRISPR-Cas типа III-B бактерии *Thermus thermophilus*. В этой системе в белке Cas10 отсутствует HD-домен [43]. Однако, поскольку в этом белке присутствует интактный Palm-домен, он все ещё способен активировать

вспомогательные эффекторы за счёт цОА-пути [44]. В локусе типа III-B *T. thermophilus*, помимо РНКазы Csm6, закодирована цОА-активируемая ДНК-никаза Can1, которая, возможно, осуществляет деградацию чужеродной ДНК, компенсируя отсутствие HD-домена Cas10 [45].

В системах CRISPR-Cas типа III-F Palm-домен Cas10 содержит мутации в каталитическом центре и, вероятно, неактивен. В соответствии с этим в локусах типа III-F отсутствуют гены, кодирующие белки с CARF-доменами [4]. Можно предположить, что за счёт активности HD-домена Cas10 системы типа III-F могут обеспечивать по крайней мере частичную защиту от мобильных генетических элементов. Так, было показано, что некоторые «полные» CRISPR-Cas-системы типа III могут защищать клетки от инфекции вирусами даже при удалении из генома хозяина генов, кодирующих цОА-активируемые РНКазы [30].

Системы типа III-E ещё более необычны, поскольку в них полностью отсутствуют гены, кодирующие субъединицы Cas10. Что ещё любопытнее, несколько субъединиц Cas7, которые составляют многосубъединичный сгРНК-связывающий филамент в других системах типа III, в этих системах слиты и формируют один большой многодоменный белок. В локусах типа III-E отсутствуют гены, кодирующие CARF-доменные белки, но в некоторых локусах предсказаны гены, кодирующие каспазоподобные эффекторы [4]. Механизмы действия этих систем до сих пор неясны.

Системы HRAMP (Halobacterial RAMP), биоинформатически предсказанные в геномах архей класса Halobacteria, представляют собой ещё один пример «неполных» локусов, сходных с системами типа III. В состав локусов HRAMP входят так называемый «сигнатурный ген систем HRAMP», гены группы *cas7* и *cas5*. Также эти локусы часто кодируют нуклеазы, содержащие домены DEDDy и HNH. CRISPR-кассеты в этих системах отсутствуют. Функции систем HRAMP на текущий момент неизвестны, однако, исходя из их состава, предполагается, что они могут функционировать как иммунные системы, расщепляющие двухцепочечную ДНК. Гены «сигнатурных белков HRAMP» входят в состав всех известных локусов HRAMP, однако при анализе аминокислотных последовательностей не удавалось обнаружить какие-либо гомологи, не входящие в локусы HRAMP, и функция этих белков остаётся неизвестной [46].

Используя недавно разработанную систему AlphaFold2, предсказывающую пространственные структуры белков, мы обнаружили структурное сходство между сигнатурными бел-

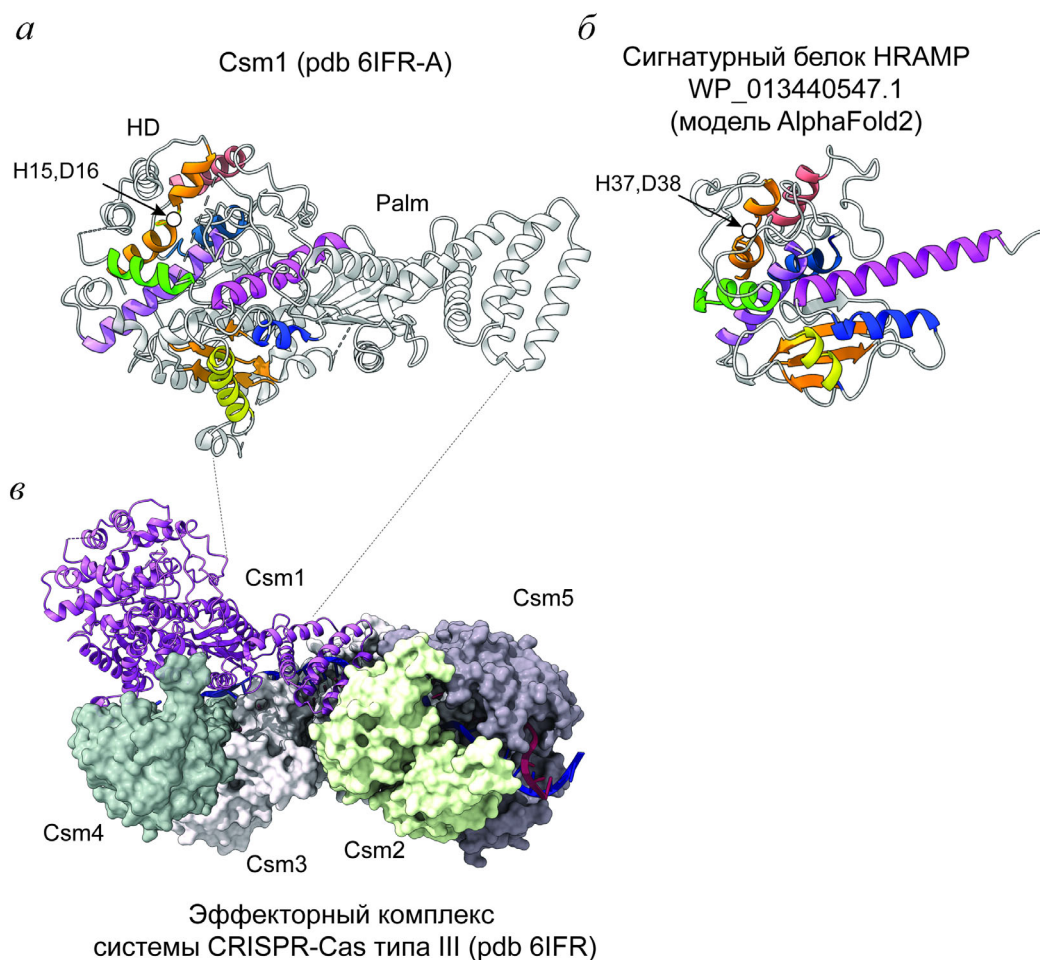


Рис. 3. Структурная гомология между сигнаурным белком HRAMP и белком Cas10. Структура Csm1 из *Streptococcus thermophilus* (а) и модель сигнаурного белка HRAMP WP_013440547.1 из *Halogeometricum borinquense*, предсказанная AlphaFold2 (б), показаны в виде ленточных диаграмм. Элементы со структурным сходством выделены соответствующими цветами. Домены, сходство между которыми отсутствует, показаны светло-серым цветом. Активные центры His–Asp указаны стрелками. в – Белок Csm1 из *S. thermophilus* в составе эффекторного комплекса типа III-A

ками HRAMP и белком Csm1 (белок Cas10 систем типа III-A) (моделирование было проведено с помощью ColabFold AlphaFold2_mmseqs2 со стандартными параметрами, поиск структурных гомологов был осуществлён на сервере Dali [47–49]; значения Z-коэффициента Dali для предсказанной структуры HRAMP WP_013440547.1 из *Halogeometricum borinquense* и для структуры Csm1 из *Streptococcus thermophilus* (PDB 6IFR-A) – 10.6). Согласно результатам предсказания, в сигнаурном белке HRAMP присутствует HD-домен, очень схожий с HD-доменом в Csm1, содержащий характерный для этого домена активный центр с остатками гистидина и аспартата (рис. 3, а, б). Сигнаурный белок HRAMP в два раза меньше белка Csm1, и в нём нет доменов, необходимых для взаимодействия с белками Csm2 и Csm4, которые входят в системы CRISPR-Cas типа III-A, но отсут-

ствуют в системах HRAMP (рис. 3, а–в). Сигнаурный белок HRAMP также содержит домен, схожий с Palm, но этот домен значительно меньше Palm-домена Csm1 и, вероятно, неактивен. В соответствии с этим предположением, в локусах систем HRAMP отсутствуют гены белков с CARF-доменами; тем не менее эта гипотеза все же требует экспериментальной проверки.

Поскольку ближайшими гомологами белков Cas7 систем HRAMP являются белки Cas7 из систем CRISPR-Cas типа III, было сделано предположение о родстве этих систем [46]. Структурная гомология между сигнаурными белками HRAMP и белками Cas10 типа III также подтверждает это родство. Возможно, системы HRAMP произошли от систем типа III, схожих с современными, а CRISPR-кассеты и модули адаптации были утрачены в ходе эволюции. В качестве альтернативной гипотезы мож-

но предположить, что системы HRAMP произошли от предковых систем типа III, которые существовали ещё до появления CRISPR-кассет и до приобретения адаптационных модулей. В последнем случае системы HRAMP могут быть молекулярными реликтами, а их изучение может прояснить функции древних прототипических систем типа III.

Хотя наиболее распространёнными вспомогательными эффекторами в системах типа III являются CARF-доменные РНКазы семейства HEPN, активируемые циклическими олигоаденилатами [4], в некоторых системах типа III используются и другие цОА-активируемые эффекторы.

Недавно был охарактеризован один из таких необычных вспомогательных эффекторов, ассоциированных с системами типа III. Активируемый циклическим тетрааденилатом белок Card1, закодированный в локусе типа III-A бактерии *Treponema succinifaciens*, обладает нуклеазной активностью по отношению к одноцепочечным ДНК и РНК *in vitro*. Однако, поскольку активация Card1 в клетках *Staphylococcus aureus* не приводит к заметным изменениям в транскрипте [50], РНКазная активность этого белка либо подавлена *in vivo*, либо является высокоспецифичной и не вызывает массовой деградации транскриптов. Экспрессия Card1 компенсирует делецию гена *csmb* в системе типа III-A из *S. epidermidis*, восстанавливая способность клеток противостоять фаговой инфекции. Однако штаммы с каталитически неактивным HD-доменом Cas10 не могли выжить после инфекции даже при наличии Card1 [50], что говорит о том, что ДНКазной активности Card1 недостаточно для эффективного иммунного ответа. Было показано, что для того, чтобы клетки полностью избавились от плазмиды, содержащей транскрибируемый протоспейсер, требуется активность как HD-домена Cas10, так и Card1, что косвенно подтверждает специфичность нуклеазной активности HD-домена Cas10 к ДНК, несущей протоспейсер [50].

ПРИБРЕТЕНИЕ СПЕЙСЕРОВ В СИСТЕМАХ CRISPR-Cas ТИПА III

Практически во всех системах CRISPR-Cas, включая тип III, интеграция новых спейсеров в CRISPR-кассету производится комплексом, состоящим из консервативных белков Cas1 и Cas2 (рис. 1, а) [4]. Интеграционные комплексы Cas1-Cas2 способны захватывать короткие фрагменты ДНК и встраивать их в CRISPR-кассету [51]. Происхождение этих фрагментов

ДНК остаётся малоизученным, однако для систем типов I и II было показано, что такие фрагменты могут образовываться в ходе деградации ДНК комплексами RecBCD, или AddAB [52, 53], или, в случае «праймированной адаптации» в системах типа I, процессивной нуклеазой Cas3 [54]. Иммунный ответ систем типа III против плазмид и вирусов с ДНК-геномами требует транскрипции протоспейсера, поэтому для клеток выгодны механизмы, обеспечивающие преимущественное приобретение спейсеров из транскрипционно активных участков мобильных генетических элементов. Модули адаптации некоторых систем типа III содержат белки Cas1, слитые с доменами обратной транскриптазы (RT), которые могут приобретать новые спейсеры напрямую из РНК (рис. 1, а) [55, 56]. Было показано, что при оверэкспрессии компонентов адаптации *cas1::RT-cas2* происходит приобретение спейсеров из РНК, предпочтительно из высокопредставленных транскриптов [57, 58]. Однако приобретённые спейсеры встраиваются в CRISPR-кассету в обеих ориентациях, так что, вероятно, лишь половина приобретённых из РНК спейсеров может обеспечить эффективный иммунный ответ.

Механизмы приобретения спейсеров из РНК на текущий момент плохо изучены. Комплекс белков Cas1::RT-Cas2 из *Marinomonas mediterranea* в условиях *in vitro* способен присоединять 3'-концы молекул РНК к 5'-концам повторов в молекулах ДНК, содержащих CRISPR-кассету. Образующиеся гибридные молекулы РНК-ДНК могли служить матрицами для обратной транскрипции, катализируемой RT-доменом белка Cas1::RT [57]. В условиях этих экспериментов молекулы РНК могут присоединяться к обеим сторонам первого повтора CRISPR-касеты [57], что согласуется с наблюдаемым в экспериментах *in vivo* отсутствием предпочтения к ориентации спейсеров относительно цепи транскрипта. С другой стороны, белковый комплекс Cas1::RT-Cas2 из *Thiomicrospira* способен присоединять молекулы РНК к содержащим CRISPR-кассету фрагментам ДНК *in vitro*, но не может преобразовывать такие промежуточные продукты в удлинённые CRISPR-касеты. Это может означать, что *in vivo* для такого преобразования либо требуются дополнительные факторы, либо синтез кДНК происходит до интеграции пре-спейсеров в CRISPR-кассету [59].

Домен обратной транскриптазы, входящий в состав некоторых адаптационных модулей, вероятно, играет скорее вспомогательную роль – белок Cas1-RT с инактивированным RT-доменом способен встраивать спейсеры из

ДНК подобно обычным белкам Cas1 [57]. Более того, большинство локусов типа III кодируют стандартные белки Cas1 [57] и не кодируют белки с RT-доменами. В таких системах новые спейсеры должны приобретаться из ДНК, и этот процесс не должен зависеть от транскрипции протоспейсера, если только в этих системах не используются другие, ещё не описанные механизмы, позволяющие приобретать спейсеры из транскрипционно активных участков. Модуль адаптации систем типа III бактерии *T. thermophilus* содержит обычные гены *cas1-cas2*, а в геноме бактерии не обнаружено никаких белков с RT-доменами. Тем не менее при адаптации спейсеров в процессе инфекции литическим бактериофагом наблюдается крайне неравномерное распределение приобретённых спейсеров: практически все выжившие клетки экспрессируют сгРНК комплементарные вирусным транскриптам. Однако такой состав приобретённых спейсеров является результатом отбора сгРНК на их способность защищать клетки от инфекции фагом и не связан с какой-либо специфичностью механизмов адаптации [32].

ПРЕДОТВРАЩЕНИЕ АУТОИММУННОГО ОТВЕТА В СИСТЕМАХ CRISPR-Cas ТИПА III

Эффекторные комплексы CRISPR-Cas распознают нуклеиновые кислоты, комплементарные спейсерной части сгРНК. Поскольку сгРНК образуются из CRISPR-кассет, мишенью эффекторов также может оказаться сама геномная ДНК (или, в случае систем с РНК-мишенями, «антисмысловые» транскрипты CRISPR-кассет). Такие «аутоиммунные» взаимодействия вредны для клетки, и потому должны существовать механизмы, позволяющие их предотвратить. Для распознавания мишени ДНК-специфичными системами CRISPR-Cas (к которым относятся системы типов I, II и V) необходимы короткие (порядка нескольких нуклеотидов) вырожденные последовательности, расположенные в ДНК мишени рядом с протоспейсерами, но отсутствующие вблизи спейсеров в CRISPR-кассетах — так называемые PAM (Protospacer Adjacent Motifs) [60]. Эффекторные комплексы «сканируют» двухцепочечные молекулы ДНК и после распознавания PAM инициируют плавление дуплекса ДНК с последующим внедрением сгРНК и формированием R-петли [61–63]. В системах типа III механизм предотвращения аутоиммунного ответа устроен совершенно иначе. Эффекторы типа III распознают и расщепляют любые молекулы РНК, комп-

лементарные спейсерной части сгРНК, но для активации HD- и Palm-доменов субъединицы Cas10 требуется отсутствие комплементарных взаимодействий между мишенью и происходящим из повтора 5'-концевым участком («5'-тэгом») сгРНК (рис. 2, в) [17, 64, 65]. Любопытно, что такой же принцип избегания аутоиммунного ответа используется в РНК-специфичных CRISPR-Cas-системах типа VI [66, 67].

Можно задаться вопросом, почему все ДНК-специфичные системы CRISPR-Cas избегают аутоиммунного ответа за счёт распознавания PAM, в то время как РНК-специфичные системы для этой цели используют другие механизмы. Теоретически, стратегия «тэг-антитэг» должна подходить и для ДНК-специфичных систем (вообще говоря, механизм «тэг-антитэг» был обнаружен в системах типа III, когда предполагалось, что их мишенью является ДНК [64]). Однако в ДНК-специфичных системах используются лишь PAM-зависимые механизмы, которые независимо возникли как минимум трижды [68]. Кроме того, согласно текущим данным, системы типа III являются предшественниками всех известных систем класса I [68]. Если исходить из этого эволюционного сценария, в системах типа I предковая стратегия избегания аутоиммунности, основанная на взаимодействиях «тэг-антитэг», должна была смениться на PAM-зависимый механизм.

Такое предпочтение к PAM-зависимым механизмам в ДНК-специфичных системах может объясняться кинетикой распознавания мишени эффекторными комплексами. В одноцепочечных молекулах РНК нуклеотидные основания экспонированы, и молекулы сгРНК в составе эффекторных комплексов могут взаимодействовать с РНК-мишенями напрямую. Напротив, в двухцепочечной ДНК комплементарное взаимодействие с сгРНК невозможно, по крайней мере, на начальном этапе. В этом случае для осуществления комплементарных взаимодействий требуется локальное плавление дуплекса ДНК, что существенно замедляет процесс поиска мишени. Движущая сила для начала плавления обеспечивается энергией связывания белковых компонентов эффектора и дополнительно взаимодействием с двухцепочечной PAM-последовательностью. Поскольку плавление дуплекса ДНК и формирование R-петли происходит только после распознавания PAM, это ограничивает количество потенциальных мишеней по крайней мере на порядок, ускоряя тем самым поиск протоспейсера. Таким образом, можно предположить, что основная причина использования PAM заключается не столько в избегании аутоиммунных реакций, сколько в

ускорении поиска протоспейсеров в двухцепочечной ДНК, столь важном для эффективного иммунного ответа.

ПРОИСХОЖДЕНИЕ СИСТЕМ CRISPR-Cas ТИПА III

Иммунный ответ в системах CRISPR-Cas типа III требует скоординированного взаимодействия нескольких сложных механизмов — приобретения новых спейсеров, созревания *cr*РНК, сборки эффекторных комплексов, распознавания мишени, синтеза сигнальных соединений, активации вспомогательных эффекторов и регуляции всех вышеперечисленных процессов. При рассмотрении таких сложных систем любой человек рано или поздно задаётся вопросом: «Как же все это смогло возникнуть?»

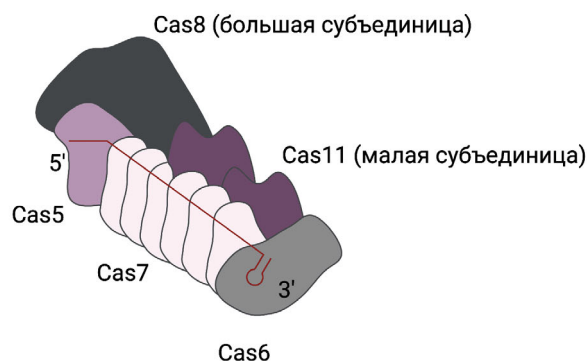
В геномах прокариот можно обнаружить гомологи некоторых генов *cas*, не ассоциированные с системами CRISPR-Cas, и анализ таких генов отчасти прояснил эволюционную историю CRISPR-Cas-систем. Оказалось, что некоторые не ассоциированные с системами CRISPR-Cas гены *casI* принадлежат к ранее неизвестному семейству мобильных генетических элементов, названных каспозонами, которые используют белки Cas1 для собственной интеграции/эксцизии [69, 70]. Механизм интеграции каспозонов схож с механизмом интеграции новых спейсеров в CRISPR-касеты. По всей видимости, CRISPR-ассоциированные белки Cas1 произошли от интеграз каспозонов, а повторы CRISPR могли произойти от фланкирующих каспозоны концевых повторов [68, 70]. Наряду с белками Cas1 каспозоны также кодируют гомологи белков Cas4 [69], участвующие в процессе

приобретения новых спейсеров в ряде систем CRISPR-Cas [51].

Другим важным компонентом комплекса интеграции спейсеров является белок Cas2, который гомологичен нуклеазным токсинам семейства *VapD*, и, вероятно, происходит от соответствующих токсин-антитоксиновых систем [71]. В комплексе интеграции белок Cas2 выполняет структурные функции, скрепляя друг с другом димеры Cas1, однако во многих системах белки Cas2 сохраняют нуклеазную активность, хотя она не является необходимой для адаптации [72].

Эволюционная история компонентов эффекторных комплексов остаётся куда менее ясной. В этом обзоре будет рассмотрен только сценарий происхождения систем CRISPR-Cas класса I, включающего типы I, III и IV. Эффекторные комплексы систем класса I имеют сходную архитектуру. Кроме того, между некоторыми субъединицами этих комплексов прослеживается гомология, что позволяет предположить их общее происхождение (рис. 4). Центральная часть эффекторов типов I и III представляет собой филамент, состоящий из молекулы *cr*РНК, покрытой несколькими субъединицами Cas7, находящимися в комплексе с малыми субъединицами Cas11; 5'-конец *cr*РНК связан с белком семейства Cas5. Большие субъединицы (Cas8 в системе типа I и Cas10 в системе типа III) располагаются вблизи 5'-конца *cr*РНК [73, 74]. Созревание *cr*РНК в системах типов I и III проходит при участии белков Cas6, которые распознают вторичные структуры, образованные элементами повторов в пре-*cr*РНК [75]. Хотя эффекторные комплексы систем типа IV на текущий момент плохо исследованы, известно, что в их основе также лежит филамент из нескольких субъединиц Cas7, связанных с *cr*РНК [76]; в локусах

Эффектор систем типа I



Эффектор систем типа III

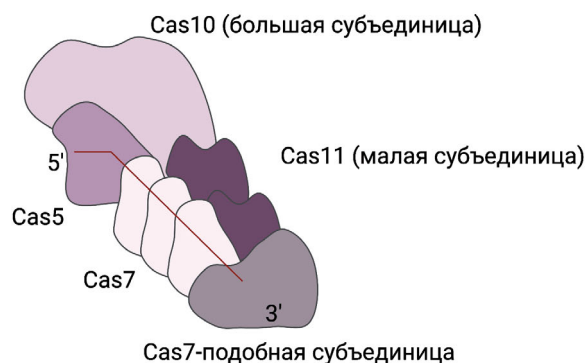


Рис. 4. Сравнение структур эффекторов типа I и типа III; гомологичные белки изображены одинаковыми цветами

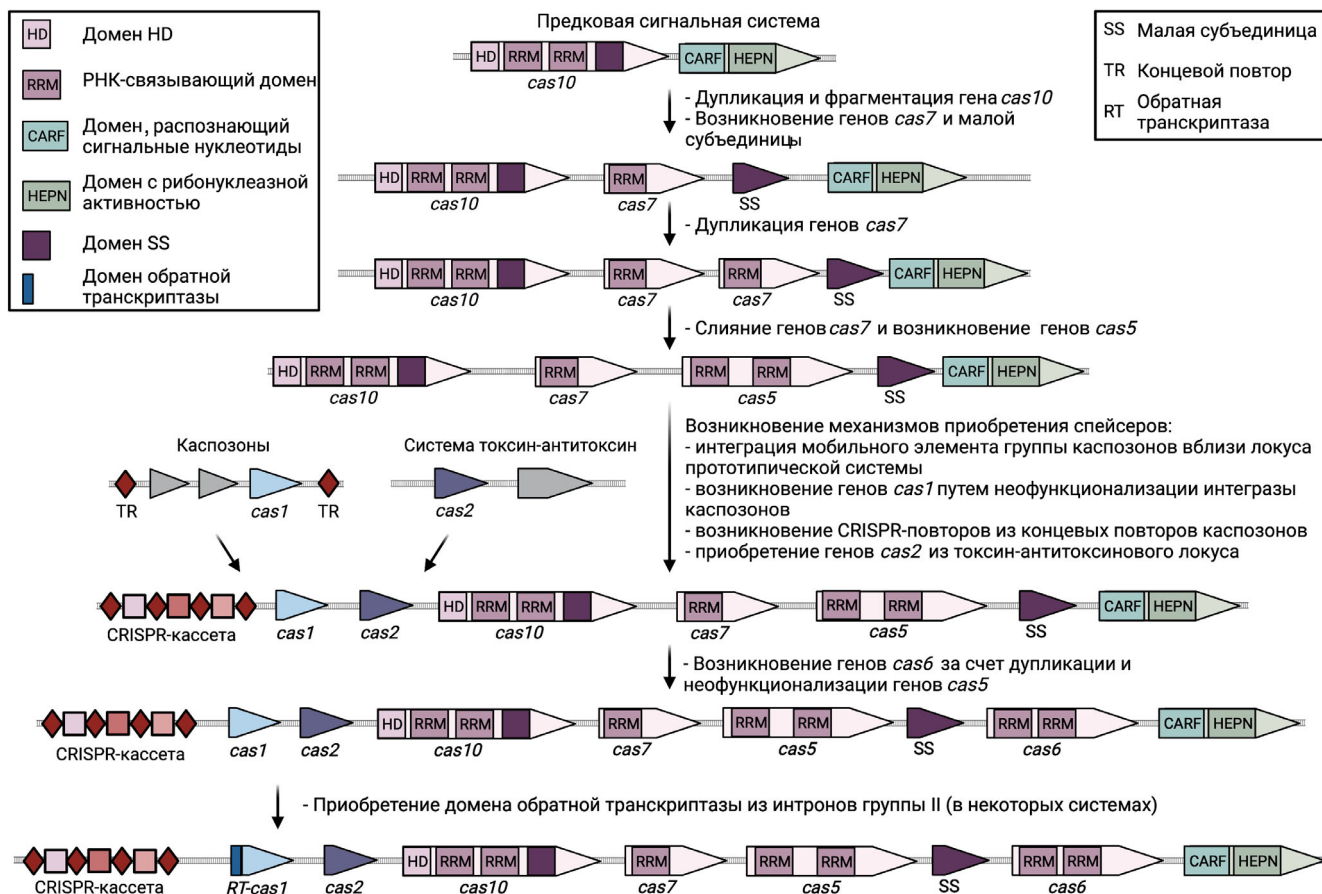


Рис. 5. Предлагаемая модель происхождения систем CRISPR-Cas III типа. Адаптировано из статьи Koonin и Makarova [68]

типа IV также закодированы гомологи белков Cas5 и Cas6 [4]. Главные компоненты эффекторов типов I и III – белки Cas10, Cas5, Cas7 и Cas6 обладают структурным сходством и содержат домены с RRM-укладкой [68, 77]. Кроме того, белки малых субъединиц Cas11 систем типа I имеют структурное сходство с C-концевым доменом Cas10 систем типа III [77].

Исходя из перечисленных сведений, Koonin и Makarova [68] предложили сценарий происхождения систем CRISPR-Cas класса I (рис. 5). Согласно этому сценарию системы типа III, несмотря на всю свою сложность, являются предковыми по отношению ко всем известным системам класса I [4, 68, 77]. Предполагается, что эффекторы класса I произошли от гипотетической сигнальной системы, состоящей из Cas10-подобной полимеразы/циклазы и эффекторного белка CARF–HEPN. Эта сигнальная система могла производить вторичные мессенджеры нуклеотидной природы (вероятно, циклические олигоаденилаты) в ответ на стресс или какие-то другие сигналы окружающей среды с

последующей активацией РНКазной активности белка CARF–HEPN [4]. Подобные локусы, кодирующие гомологи Cas10, слитые с доменами CARF–HEPN, были предсказаны биоинформатически [78], но на текущий момент ни одна из таких систем не была экспериментально охарактеризована. Дупликация предкового гена Cas10-подобного белка с последующим разбиением могла привести к появлению предковых белков Cas7 и Cas11. Гены, кодирующие Cas5 и Cas6, вероятно, возникли в результате слияния двух *cas7*-подобных генов, поскольку все эти гены имеют специфический структурный мотив, отсутствующий в белках Cas10 [68]. Наконец, приобретение компонентов адаптационного модуля от каспозонов и токсин-антитоксинового систем привело к появлению функциональной адаптивной иммунной системы [68].

О первоначальных функциях предковых сигнальных систем, состоящих из белка Cas10 и эффекторов CARF–HEPN, и о том, какие стимулы активировали полимеразную/циклазную активность прототипического белка Cas10, мож-

но лишь предполагать. Функции прототипических систем типа III, существовавших до приобретения модуля адаптации, также неизвестны. Тем не менее известны отдельные варианты систем класса I, которые не ассоциированы с CRISPR-касетами [46]; исследование этих систем, вероятно, может помочь выявить функции предковых эффекторов класса I. Известно, что эффекторные комплексы систем типа IV при гетерологичной экспрессии в клетках *Escherichia coli* преимущественно связывают малые РНК, транскрибируемые с плазмид, несмотря на присутствие функциональных компонентов созревания сгРНК и транскрибируемой CRISPR-касеты [79]. Можно предположить, что древние прототипические эффекторы типа III также могли связывать транскрипты мобильных генетических элементов и использовать их для распознавания мишеней, функционируя как примитивная и малоэффективная иммунная система, сходная с современными прокариотическими белками-аргонатами [80, 81]. Кроме того, Palm-домен Cas10 имеет сходство с каталитическим доменом фермента Hng1 – необычной 3'-5' РНК-полимеразой, необходимой для созревания гистидиновой тРНК [8]. Учитывая этот факт и предполагая, что РНК-связывающие белки Cas7 произошли от RRM-домена Cas10, можно предположить, что предковый белок Cas10 обладал РНК-связывающей и РНК-полимеразной активностью и мог быть частью системы, участвующей в синтезе, репарации и/или созревании молекул РНК. Функция этой Cas10-РНК-полимеразы со временем могла специализироваться для синтеза сигнальных соединений. Лиганды, способные активировать полимеразную активность древнего белка Cas10, не известны, однако, учитывая, что древний Cas10 мог обладать РНК-связывающей активностью, можно предположить, что такими лигандами должны быть молекулы РНК. Хотя достаточно трудно вообразить, какие молекулы РНК могли активировать иммунный ответ в такой системе, стоит отметить, что многие мобильные элементы, в том числе и некоторые каспозоны, используют гены тРНК в качестве мишеней для интеграции [69, 82]. Поскольку эффекторы типа IV связывают и тРНК [83], можно предположить, что до приобретения адаптационных модулей предковые системы типа III могли распознавать транскрипты генов тРНК, повреждённые вставками мобильных генетических элементов.

Защитные системы, которые синтезируют сигнальные вещества нуклеотидной природы в ответ на определенные стимулы, широко распространены среди всех клеточных организмов.

Помимо описанного цОА-пути в системах типа III и эукариотической системы ОАС-РНКазы L, стоит упомянуть большую группу цГАС/DncV-подобных нуклеотидилтрансфераз. У животных синтаза циклического ГМФ-АМФ (цГАС) производит молекулы циклического ГМФ-АМФ в ответ на присутствие ДНК в цитозоле. Циклический ГМФ-АМФ активирует сигнальный каскад, стимулирующий экспрессию ряда генов, вовлечённых в иммунный ответ [84]. В геномах прокариот обнаружены многочисленные гомологи цГАС, многие из них ассоциированы с генами различных эффекторов, такими как нуклеазы, фосфолипазы и трансмембранные белки [85–88]. В таких защитных системах передача сигнала между сенсорами (циклазами) и эффекторами осуществляется малыми диффундирующими молекулами, что позволяет существование большого количества пар сенсор/эффектор и взаимосвязей между различными системами. Любопытно, что некоторые из цГАС/DncV-подобных циклаз производят циклические олигоаденилаты и активируют цОА-чувствительные эффекторы [88]; кроме того, некоторые из таких эффекторов активируются и в ходе иммунного ответа в системах CRISPR-Cas типа III [89, 90].

ЗАКЛЮЧЕНИЕ

Системы III типа, будучи самыми сложными из всех известных систем CRISPR-Cas, по-видимому, являются предковыми по отношению к системе класса I. Если учесть, что системы класса I являются самыми распространёнными среди прокариот, системы типа III, вероятно, также являются самыми древними системами CRISPR-Cas. В течение долгого времени механизм работы систем типа III оставался малоизвестным. Хотя открытие сигнального пути, связывающего эффекторные комплексы и вспомогательные нуклеазы, во многом прояснило механизмы иммунного ответа в системах типа III, некоторые аспекты работы этих систем все ещё неясны. Хотя ДНКазная активность HD-домена Cas10 необходима для иммунного ответа, до сих пор неизвестно, какие мишени она расщепляет в клетке. Вспомогательные РНКазы Csm6 и Csx1 были охарактеризованы как неспецифические в экспериментах *in vitro*, но до сих пор нет данных об их специфичности *in vivo*. Существует множество генов, связанных с системами CRISPR-Cas типа III, но многие из них до сих пор не охарактеризованы экспериментально. Помимо полных систем типа III, в ходе биоинформатического анализа была предсказана

группа упрощённых RAMP-содержащих локусов; эти системы представляют интерес сами по себе, но они также могут представлять собой промежуточные звенья, отражающие эволюционную историю систем CRISPR-Cas типа III. В частности, обнаруженная нами структурная гомология между сигнатурными белками HRAMP и белками Cas10 систем типа III позволяет предположить, что системы HRAMP произошли от существующих систем типа III или их предков. Наконец, ещё предстоит охарактеризовать активности и биологические функции современных сигнальных систем Cas10-CARF-HEPN. Таким образом, существенная часть разнообразия систем CRISPR-Cas типа III и механизмов работы этих систем остаётся неизученной, и многое из этого ещё предстоит открыть.

Финансирование. Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 20-34-90147, М.В.К.), Российского научного фонда (гранты №№ 19-74-00124, Д.Н.А и 21-14-00122, Я.В.Ф.), гранта Президента Российской Федерации (грант № МК-3793.2021.1.4, Я.В.Ф.) и Министерства науки и высшего образования Российской Федерации (грант № 075-15-2019-1661).

Благодарность. Мы благодарим Светлану Белухину за помощь в работе над текстом.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм. Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

СПИСОК ЛИТЕРАТУРЫ

- Bennett, C. H. (1979) Dissipation-error tradeoff in proof-reading, *Biosystems*, **11**, 85-91.
- Koonin, E. V., Wolf, Y. I., and Katsnelson, M. I. (2017) Inevitability of the emergence and persistence of genetic parasites caused by evolutionary instability of parasite-free states, *Biol. Direct*, **12**, 31.
- Barrangou, R., and Horvath, P. (2017) A decade of discovery: CRISPR functions and applications, *Nat. Microbiol.*, **2**, 17092.
- Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., et al. (2020) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants, *Nat. Rev. Microbiol.*, **18**, 67-83.
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., et al. (2015) An updated evolutionary classification of CRISPR-Cas systems, *Nat. Rev. Microbiol.*, **13**, 722-736.
- Makarova, K. S., Aravind, L., Grishin, N. V., Rogozin, I. B., and Koonin, E. V. (2002) A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis, *Nucleic Acids Res.*, **30**, 482-496.
- Aravind, L., and Koonin, E. V. (1999) DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history, *Nucleic Acids Res.*, **27**, 1609-1618.
- Anantharaman, V., Iyer, L. M., and Aravind, L. (2010) Presence of a classical RRM-fold palm domain in Thg1-type 3'-5' nucleic acid polymerases and the origin of the GGDEF and CRISPR polymerase domains, *Biol. Direct*, **5**, 43.
- Haft, D. H., Selengut, J., Mongodin, E. F., and Nelson, K. E. (2005) A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes, *PLoS Comput. Biol.*, **1**, e60.
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J. J., Charpentier, E., et al. (2011) Evolution and classification of the CRISPR-Cas systems, *Nat. Rev. Microbiol.*, **9**, 467-477.
- Marraffini, L. A., and Sontheimer, E. J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA, *Science*, **322**, 1843-1845.
- Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., et al. (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex, *Cell*, **139**, 945-956.
- Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., et al. (2012) Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity, *Mol. Cell*, **45**, 303-313.
- Goldberg, G. W., Jiang, W., Bikard, D., and Marraffini, L. A. (2014) Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting, *Nature*, **514**, 633-637.
- Deng, L., Garrett, R. A., Shah, S. A., Peng, X., and She, Q. (2013) A novel interference mechanism by a type IIB CRISPR-Cmr module in *Sulfolobus*, *Mol. Microbiol.*, **87**, 1088-1099.
- Gong, B., Shin, M., Sun, J., Jung, C.-H., Bolt, E. L., et al. (2014) Molecular insights into DNA interference by CRISPR-associated nuclease-helicase Cas3, *Proc. Natl. Acad. Sci. USA*, **111**, 16359-16364.
- Kazlauskienė, M., Tamulaitis, G., Kostiuk, G., Venclovas, Č., and Siksnys, V. (2016) Spatiotemporal control of Type III-A CRISPR-Cas immunity: coupling DNA degradation with the target RNA recognition, *Mol. Cell*, **62**, 295-306.
- Elmore, J. R., Sheppard, N. F., Ramia, N., Deighan, T., Li, H., et al. (2016) Bipartite recognition of target RNAs activates DNA cleavage by the type III-B CRISPR-Cas system, *Genes Dev.*, **30**, 447-459.
- Estrella, M. A., Kuo, F.-T., and Bailey, S. (2016) RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex, *Genes Dev.*, **30**, 460-470.
- Samai, P., Pyenson, N., Jiang, W., Goldberg, G. W., Hatoum-Aslan, A., and Marraffini, L. A. (2015) Co-transcriptional DNA and RNA cleavage during Type III CRISPR-Cas immunity, *Cell*, **161**, 1164-1174.
- Liu, T. Y., Liu, J.-J., Aditham, A. J., Nogales, E., and Doudna, J. A. (2019) Target preference of Type III-A CRISPR-Cas complexes at the transcription bubble, *Nat. Commun.*, **10**, 3001.
- Hatoum-Aslan, A., Maniv, I., Samai, P., and Marraffini, L. A. (2014) Genetic characterization of antiplasmid immunity through a type III-A CRISPR-Cas system, *J. Bacteriol.*, **196**, 310-317.
- Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V., and Aravind, L. (2014) CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems, *Front. Genet.*, **5**, 102.
- Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V., and Aravind, L. (2013) Comprehensive

- analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing, *Biol. Direct*, **8**, 15.
25. Sheppard, N. F., Glover, C. V. C., 3rd, Terns, R. M., and Terns, M. P. (2016) The CRISPR-associated Csx1 protein of *Pyrococcus furiosus* is an adenosine-specific endoribonuclease, *RNA*, **22**, 216-224.
 26. Niewoehner, O., and Jinek, M. (2016) Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6, *RNA*, **22**, 318-329.
 27. Niewoehner, O., Garcia-Doval, C., Rostøl, J. T., Berk, C., Schwede, F., et al. (2017) Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers, *Nature*, **548**, 543-548.
 28. Kazlauskienė, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G., and Siksnys, V. (2017) A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems, *Science*, **357**, 605-609.
 29. Han, W., Stella, S., Zhang, Y., Guo, T., Sulek, K., et al. (2018) A type III-B Cmr effector complex catalyzes the synthesis of cyclic oligoadenylate second messengers by cooperative substrate binding, *Nucleic Acids Res.*, **46**, 10319-10330.
 30. Jiang, W., Samai, P., and Marraffini, L. A. (2016) Degradation of phage transcripts by CRISPR-associated RNases enables type III CRISPR-Cas immunity, *Cell*, **164**, 710-721.
 31. Rostøl, J. T., and Marraffini, L. A. (2019) Non-specific degradation of transcripts promotes plasmid clearance during type III-A CRISPR-Cas immunity, *Nat. Microbiol.*, **4**, 656-662.
 32. Artamonova, D., Karneyeva, K., Medvedeva, S., Klimuk, E., Kolesnik, M., et al. (2020) Spacer acquisition by Type III CRISPR-Cas system during bacteriophage infection of *Thermus thermophilus*, *Nucleic Acids Res.*, **48**, 9787-9803.
 33. Nasef, M., Muffly, M. C., Beckman, A. B., Rowe, S. J., Walker, F. C., et al. (2019) Regulation of cyclic oligoadenylate synthesis by the Cas10-Csm complex, *RNA*, **25**, 948-962.
 34. Rouillon, C., Athukoralage, J. S., Graham, S., Grüşchow, S., and White, M. F. (2018) Control of cyclic oligoadenylate synthesis in a type III CRISPR system, *Elife*, **7**, e36734.
 35. Athukoralage, J. S., Rouillon, C., Graham, S., Grüşchow, S., and White, M. F. (2018) Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate, *Nature*, **562**, 277-280.
 36. Garcia-Doval, C., Schwede, F., Berk, C., Rostøl, J. T., Niewoehner, O., et al. (2020) Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6, *Nat. Commun.*, **11**, 1596.
 37. Athukoralage, J. S., Graham, S., Grüşchow, S., Rouillon, C., and White, M. F. (2019) A type III CRISPR ancillary ribonuclease degrades its cyclic oligoadenylate activator, *J. Mol. Biol.*, **431**, 2894-2899.
 38. Athukoralage, J. S., McMahon, S. A., Zhang, C., Grüşchow, S., Graham, S., et al. (2020) An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity, *Nature*, **577**, 572-575.
 39. Hornung, V., Hartmann, R., Ablasser, A., and Hopfner, K.-P. (2014) OAS proteins and cGAS: unifying concepts in sensing and responding to cytosolic nucleic acids, *Nat. Rev. Immunol.*, **14**, 521-528.
 40. Sarkar, S. N., Ghosh, A., Wang, H. W., Sung, S. S., and Sen, G. C. (1999) The nature of the catalytic domain of 2'-5'-oligoadenylate synthetases, *J. Biol. Chem.*, **274**, 25535-25542.
 41. Pairo-Castineira, E., Clohisey, S., Klaric, L., Bretherick, A. D., Rawlik, K., et al. (2021) Genetic mechanisms of critical illness in COVID-19, *Nature*, **591**, 92-98.
 42. Mo, C. Y., Mathai, J., Rostøl, J. T., Varble, A., Banh, D. V., and Marraffini, L. A. (2021) Type III-A CRISPR immunity promotes mutagenesis of staphylococci, *Nature*, **592**, 611-615.
 43. Staals, R. H. J., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D. W., et al. (2013) Structure and activity of the RNA-targeting Type III-B CRISPR-Cas complex of *Thermus thermophilus*, *Mol. Cell*, **52**, 135-145.
 44. Steens, J. A., Zhu, Y., Taylor, D. W., Bravo, J. P. K., Prinsen, S. H. P., et al. (2021) SCOPE enables type III CRISPR-Cas diagnostics using flexible targeting and stringent CARF ribonuclease activation, *Nat. Commun.*, **12**, 5033.
 45. McMahon, S. A., Zhu, W., Graham, S., Rambo, R., White, M. F., and Gloster, T. M. (2020) Structure and mechanism of a type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate, *Nat. Commun.*, **11**, 500.
 46. Makarova, K. S., Karamycheva, S., Shah, S. A., Vestergaard, G., Garrett, R. A., and Koonin, E. V. (2019) Predicted highly derived class 1 CRISPR-Cas system in Haloarchaea containing diverged Cas5 and Cas7 homologs but no CRISPR array, *FEMS Microbiol. Lett.*, **366**, fnz079, doi: 10.1093/femsle/fnz079.
 47. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., et al. (2021) Highly accurate protein structure prediction with AlphaFold, *Nature*, **596**, 583-589, doi: 10.1038/s41586-021-03819-2.
 48. You, L., Ma, J., Wang, J., Artamonova, D., Wang, M., et al. (2019) Structure studies of the CRISPR-Csm complex reveal mechanism of co-transcriptional interference, *Cell*, **176**, 239-253.e16.
 49. Mirdita, M., Ovchinnikov, S., and Steinegger, M. (2021) ColabFold – Making protein folding accessible to all, *bioRxiv*, doi: 10.1101/2021.08.15.456425.
 50. Rostøl, J. T., Xie, W., Kuryavii, V., Maguin, P., Kao, K., et al. (2021) The Card1 nuclease provides defence during type III CRISPR immunity, *Nature*, **590**, 624-629.
 51. Jackson, S. A., McKenzie, R. E., Fagerlund, R. D., Kieper, S. N., Fineran, P. C., and Brouns, S. J. J. (2017) CRISPR-Cas: adapting to change, *Science*, **356**, eaal5056, doi: 10.1126/science.aal5056.
 52. Levy, A., Goren, M. G., Yosef, I., Auster, O., Manor, M., et al. (2015) CRISPR adaptation biases explain preference for acquisition of foreign DNA, *Nature*, **520**, 505-510.
 53. Modell, J. W., Jiang, W., and Marraffini, L. A. (2017) CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity, *Nature*, **544**, 101-104.
 54. Shiriaeva, A. A., Savitskaya, E., Datsenko, K. A., Vvedenskaya, I. O., Fedorova, I., et al. (2019) Detection of spacer precursors formed *in vivo* during primed CRISPR adaptation, *Nat. Commun.*, **10**, 4603.
 55. Toro, N., Martínez-Abarca, F., Mestre, M. R., and González-Delgado, A. (2019) Multiple origins of reverse transcriptases linked to CRISPR-Cas systems, *RNA Biol.*, **16**, 1486-1493.
 56. Silas, S., Makarova, K. S., Shmakov, S., Páez-Espino, D., Mohr, G., et al. (2017) On the origin of reverse transcriptase-using CRISPR-Cas systems and their hyperdiverse, enigmatic spacer repertoires, *MBio*, **8**, e00897-17.
 57. Silas, S., Mohr, G., Sidote, D. J., Markham, L. M., Sanchez-Amat, A., et al. (2016) Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein, *Science*, **351**, aad4234.
 58. González-Delgado, A., Mestre, M. R., Martínez-Abarca, F., and Toro, N. (2019) Spacer acquisition from

- RNA mediated by a natural reverse transcriptase-Cas1 fusion protein associated with a type III-D CRISPR-Cas system in *Vibrio vulnificus*, *Nucleic Acids Res.*, **47**, 10202-10211.
59. Wang, J. Y., Hoel, C. M., Al-Shayeb, B., Banfield, J. F., Brohawn, S. G., and Doudna, J. A. (2021) Structural coordination between active sites of a CRISPR reverse transcriptase-integrase complex, *Nat. Commun.*, **12**, 2571.
 60. Gleditzsch, D., Pausch, P., Müller-Esparza, H., Özcan, A., Guo, X., et al. (2019) PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures, *RNA Biol.*, **16**, 504-517.
 61. Xiao, Y., Luo, M., Hayes, R. P., Kim, J., Ng, S., et al. (2017) Structure basis for directional R-loop formation and substrate handover mechanisms in type I CRISPR-Cas system, *Cell*, **170**, 48-60.e11.
 62. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., and Doudna, J. A. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9, *Nature*, **507**, 62-67.
 63. Singh, D., Mallon, J., Poddar, A., Wang, Y., Tippiana, R., et al. (2018) Real-time observation of DNA target interrogation and product release by the RNA-guided endonuclease CRISPR Cpf1 (Cas12a), *Proc. Natl. Acad. Sci. USA*, **115**, 5444-5449.
 64. Marraffini, L. A., and Sontheimer, E. J. (2010) Self versus non-self discrimination during CRISPR RNA-directed immunity, *Nature*, **463**, 568-571.
 65. Tamulaitis, G., Kazlauskienė, M., Manakova, E., Venclovas, Č., Nwokeji, A. O., et al. (2014) Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus*, *Mol. Cell*, **56**, 506-517.
 66. Liu, L., Li, X., Ma, J., Li, Z., You, L., et al. (2017) The molecular architecture for RNA-guided RNA cleavage by Cas13a, *Cell*, **170**, 714-726.e10.
 67. Meeske, A. J., and Marraffini, L. A. (2018) RNA guide complementarity prevents self-targeting in Type VI CRISPR systems, *Mol. Cell*, **71**, 791-801.e3.
 68. Koonin, E. V., and Makarova, K. S. (2019) Origins and evolution of CRISPR-Cas systems, *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **374**, 20180087.
 69. Krupovic, M., Makarova, K. S., Forterre, P., Prangishvili, D., and Koonin, E. V. (2014) Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity, *BMC Biol.*, **12**, 36.
 70. Béguin, P., Charpin, N., Koonin, E. V., Forterre, P., and Krupovic, M. (2016) Casposon integration shows strong target site preference and recapitulates protospacer integration by CRISPR-Cas systems, *Nucleic Acids Res.*, **44**, 10367-10376.
 71. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action, *Biol. Direct*, **1**, 7.
 72. Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014) Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity, *Nat. Struct. Mol. Biol.*, **21**, 528-534.
 73. Jackson, R. N., Golden, S. M., van Erp, P. B. G., Carter, J., Westra, E. R., et al. (2014) Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from *Escherichia coli*, *Science*, **345**, 1473-1479.
 74. Taylor, D. W., Zhu, Y., Staals, R. H. J., Kornfeld, J. E., Shinkai, A., et al. (2015) Structural biology. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning, *Science*, **348**, 581-585.
 75. Charpentier, E., Richter, H., van der Oost, J., and White, M. F. (2015) Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity, *FEMS Microbiol. Rev.*, **39**, 428-441.
 76. Özcan, A., Pausch, P., Linden, A., Wulf, A., Schühle, K., et al. (2019) Type IV CRISPR RNA processing and effector complex formation in *Aromatoleum aromaticum*, *Nat. Microbiol.*, **4**, 89-96.
 77. Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2013) The basic building blocks and evolution of CRISPR-CAS systems, *Biochem. Soc. Trans.*, **41**, 1392-1400.
 78. Burroughs, A. M., Zhang, D., Schäffer, D. E., Iyer, L. M., and Aravind, L. (2015) Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling, *Nucleic Acids Res.*, **43**, 10633-10654.
 79. Zhou, Y., Bravo, J. P. K., Taylor, H. N., Steens, J. A., Jackson, R. N., et al. (2021) Structure of a type IV CRISPR-Cas ribonucleoprotein complex, *iScience*, **24**, 102201.
 80. Swarts, D. C., Jore, M. M., Westra, E. R., Zhu, Y., Janssen, J. H., et al. (2014) DNA-guided DNA interference by a prokaryotic Argonaute, *Nature*, **507**, 258-261.
 81. Kuzmenko, A., Oguienko, A., Esyunina, D., Yudin, D., Petrova, M., et al. (2020) DNA targeting and interference by a bacterial Argonaute nuclease, *Nature*, **587**, 632-637.
 82. Krupovic, M., Béguin, P., and Koonin, E. V. (2017) Casposons: mobile genetic elements that gave rise to the CRISPR-Cas adaptation machinery, *Curr. Opin. Microbiol.*, **38**, 36-43.
 83. Zhou, Y., Bravo, J. P. K., Taylor, H. N., Steens, J., Jackson, R. N., et al. (2020) Structure of a type IV CRISPR-Cas effector complex, *bioRxiv*, doi: 10.1101/2020.07.31.231399.
 84. Hopfner, K.-P., and Hornung, V. (2020) Molecular mechanisms and cellular functions of cGAS-STING signalling, *Nat. Rev. Mol. Cell Biol.*, **21**, 501-521.
 85. Millman, A., Melamed, S., Amitai, G., and Sorek, R. (2020) Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems, *Nat. Microbiol.*, **5**, 1608-1615.
 86. Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., et al. (2019) Cyclic GMP-AMP signalling protects bacteria against viral infection, *Nature*, **574**, 691-695.
 87. Lowey, B., Whiteley, A. T., Keszei, A. F. A., Morehouse, B. R., Mathews, I. T., et al. (2020) CBASS immunity uses CARF-related effectors to sense 3'-5'- and 2'-5'-linked cyclic oligonucleotide signals and protect bacteria from phage infection, *Cell*, **182**, 38-49.e17.
 88. Ye, Q., Lau, R. K., Mathews, I. T., Birkholz, E. A., Watrous, J. D., et al. (2020) HORMA domain proteins and a Trip13-like ATPase regulate bacterial cGAS-like enzymes to mediate bacteriophage immunity, *Mol. Cell*, **77**, 709-722.e7.
 89. Lau, R. K., Ye, Q., Birkholz, E. A., Berg, K. R., Patel, L., et al. (2020) Structure and mechanism of a cyclic trinucleotide-activated bacterial endonuclease mediating bacteriophage immunity, *Mol. Cell*, **77**, 723-733.e6.
 90. Malone, L. M., Warring, S. L., Jackson, S. A., Warnecke, C., Gardner, P. P., et al. (2020) A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA targeting but is vulnerable to type III RNA-based immunity, *Nat. Microbiol.*, **5**, 48-55.

TYPE III CRISPR-Cas SYSTEMS: THE DISSECTION OF THE MOST COMPLEX PROKARYOTIC IMMUNE SYSTEM

Review

**M. V. Kolesnik^{1*}, I. Fedorova^{2,3}, K. A. Karneyeva¹,
D. N. Artamonova¹, and K. V. Severinov^{1,3,4*}**

¹ *Center of Life Science, Skolkovo Institute of Science and Technology, 121205 Moscow, Russia;
e-mail: matveykolesnik@gmail.com; severik@waksman.rutgers.edu*

² *Peter the Great St. Petersburg Polytechnic University, 195251 St. Petersburg, Russia*

³ *Center for Precision Genome Editing and Genetic Technologies for Biomedicine,
Institute of Gene Biology, Russian Academy of Sciences, 119334 Moscow, Russia*

⁴ *Waksman Institute of Microbiology, Piscataway, NJ 08854, USA*

The emergence and persistence of selfish genetic elements is an intrinsic feature of all living systems. Cellular organisms have evolved a plethora of elaborate defense systems that limit the spread of such genetic parasites. CRISPR-Cas are RNA-guided defense systems used by prokaryotes to recognize and destroy foreign nucleic acids. These systems acquire and store fragments of foreign nucleic acids and utilize the stored sequences as guides to recognize and destroy genetic invaders. CRISPR-Cas systems have been extensively studied, as some of them are used in various genome editing technologies. Although Type III CRISPR-Cas systems are among the most common CRISPR-Cas systems, they are also some of the least investigated ones, mostly due to the complexity of their action compared to other CRISPR-Cas system types. Type III effector complexes specifically recognize and cleave RNA molecules. The recognition of the target RNA activates the effector large subunit – the so-called CRISPR polymerase – which cleaves DNA and produces small cyclic oligonucleotides that act as signaling molecules to activate auxiliary effectors, notably non-specific RNases. In this review, we provide a historical overview of the sometimes meandering pathway of the Type III CRISPR research. We also review the current data on the structures and activities of Type III CRISPR-Cas systems components, their biological roles, and evolutionary history. Finally, using structural modeling with AlphaFold2, we show that the archaeal HRAMP signature protein, which heretofore has had no assigned function, is a degenerate relative of Type III CRISPR-Cas signature protein Cas10, suggesting that HRAMP systems have descended from Type III CRISPR-Cas systems or their ancestors.

Keywords: CRISPR-Cas Type III, prokaryotic immunity, RNA-guided defense, cyclic oligonucleotides, CRISPR evolution, HRAMP