

УДК 577.21

TALE-НУКЛЕАЗЫ – НОВЫЙ ИНСТРУМЕНТ ДЛЯ РЕДАКТИРОВАНИЯ ГЕНОМА

© 2014 г. Д. В. Глазкова*, Г. А. Шипулин

Центральный научно-исследовательский институт эпидемиологии Роспотребнадзора, Москва, 111123

Поступила в редакцию 07.11.2013 г.

Принята к печати 23.12.2013 г.

Внесение целенаправленных изменений в геном живых клеток или целых организмов позволяет решать множество задач фундаментальной науки, биотехнологии и медицины. “Выключение” генов-мишеней в зиготах дает возможность изучать функции этих генов у различных организмов, а замена отдельных нуклеотидов в ДНК – исправлять мутации в генах и таким образом лечить наследственные заболевания. Введение гена в нужную область генома можно использовать для получения клеток-продуцентов или организмов с определенными свойствами. Подобные манипуляции с геномами стали возможными после появления технологии, получившей название “редактирование генома”. Эта технология основана на внесении разрыва в определенный участок хромосомной ДНК с помощью эндонуклеазы, распознающей уникальные последовательности, и последующего восстановления целостности ДНК за счет клеточных механизмов репарации. Для редактирования генома необходимы эндонуклеазы с изменяемой специфичностью, способные избирательно расщеплять выбранную последовательность-мишень. Появление нового типа программируемых эндонуклеаз, сконструированных на базе бактериальных белков – TAL-эффекторов (Transcription Activators Like Effector, TALE), стало важным этапом в развитии технологии редактирования и способствовало широкому распространению ее методов. В настоящем обзоре рассмотрена история открытия TAL-эффекторов и создания TALE-нуклеаз, описаны их преимущества перед полученными ранее эндонуклеазами с цинковыми пальцами. Отдельный большой раздел посвящен описанию генетических модификаций, которые можно осуществлять, используя разные способы редактирования генома.

Ключевые слова: редактирование генома, TALE-нуклеазы, нуклеазы с цинковыми пальцами, гомологичная рекомбинация, нехомологичное соединение концов ДНК.

TALE NUCLEASES – NEW TOOL FOR GENOME EDITING, by D. V. Glazkova*, G. A. Shipulin (Central Research Institute for Epidemiology, Moscow, 111123 Russia; *e-mail: glazkova@pcr.ru). The ability to introduce targeted changes in the genome of living cells or entire organisms enables researchers to meet the challenges of basic life sciences, biotechnology and medicine. Knockdown of target genes in the zygotes gives the opportunity to investigate the functions of these genes in different organisms. Replacement of single nucleotide in the DNA sequence allows to correct mutations in genes and thus to cure hereditary diseases. Adding transgene to specific genomic loci can be used in biotechnology for generation of organisms with certain properties or cell lines for biopharmaceutical production. Such manipulations of gene sequences in their natural chromosomal context became possible after the emergence of the technology called “genome editing”. This technology is based on the induction of a double-strand break in a specific genomic target DNA using endonucleases that recognize the unique sequences in the genome and on subsequent recovery of DNA integrity through the use of cellular repair mechanisms. A necessary tool for the genome editing is a custom-designed endonuclease which is able to recognize selected sequences. The emergence of a new type of programmable endonucleases, which were constructed on the basis of bacterial proteins – TAL-effectors (Transcription activators like effector), has become an important stage in the development of technology and promoted wide spread of the genome editing. This article reviews the history of the discovery of TAL effectors and creation of TALE nucleases, and describes their advantages over zinc finger endonucleases that appeared earlier. A large section is devoted to description of genetic modifications that can be performed using the genome editing

Keywords: gene editing, TALE nuclease, zinc finger nuclease, homologous recombination, non-homologous end joining.

DOI: 10.7868/S00026898414030057

Принятые сокращения: TALE – TAL-эффекторы, эффекторы, похожие на активаторы транскрипции; ZF-нуклеазы – нуклеазы с цинковыми пальцами.

* Эл. почта: glazkova@pcr.ru

ВВЕДЕНИЕ

Возможность получать клетки или организмы, в которых целенаправленно изменены отдельные гены или другие элементы генома, позволяет решать многие задачи фундаментальной науки, биотехнологии и медицины. В 2007 г. Нобелевской премией было отмечено “открытие принципов введения специфических генных модификаций в организм мышей посредством эмбриональных стволовых клеток”. Результатом этого открытия стал метод получения мышей с избирательно инактивированными (нокаутными) генами. Этот метод открыл новую страницу в генетике млекопитающих и стал мощным инструментом в изучении функций генов, однако подобные исследования долгое время были ограничены только одним видом животных — мышами. Возможность манипулировать с геномом мышей, но не других организмов, стала одной из причин того, что именно мышей наиболее часто используют для моделирования наследственных болезней человека.

В настоящее время активно разрабатываются методы генной терапии наследственных заболеваний человека. В первых клинических испытаниях в одну из хромосом добавляли исправную копию дефектного гена, которую встраивали произвольным образом с помощью вирусных векторов. Недостаток такой стратегии состоит в том, что, с одной стороны, введенная экзогенная ДНК может нарушать работу окружающих генов, что вызовет тяжелые побочные эффекты. С другой стороны, попадание гена в неактивный участок хромосомы ослабляет его экспрессию и приводит к неэффективности терапии. Идеальное решение такой задачи — коррекция гена, т.е. внесение необходимых исправлений непосредственно в область мутантного гена, в этом случае сохраняется физиологическая регуляция его экспрессии с помощью эндогенного промотора.

Модификация генов в любых организмах или клетках, создание различных модельных животных, коррекция генов при наследственных заболеваниях — эти и многие другие возможности открывает технология, получившая название редактирование генома. В ее основе лежит внесение разрыва в определенный участок геномной ДНК, репарация которого с помощью клеточных механизмов гомологичной или нехомологичной рекомбинации приводит к желаемым перестройкам. Последние научные достижения делают метод все более эффективным и доступным, что открывает новые перспективы его применения в биотехнологии для создания организмов с заданными свойствами и в медицине — в генотерапии наследственных заболеваний. Одно из таких достижений — создание TALE-нуклеаз, которым посвящена наша статья.

ВОЗНИКНОВЕНИЕ ТЕХНОЛОГИИ РЕДАКТИРОВАНИЯ ГЕНОМА И ЕЕ РАЗВИТИЕ

Впервые возможность внесения направленных изменений в геном эукариот была показана на клетках дрожжей [1]. Примененный с этой целью метод был основан на гомологичной рекомбинации между эндогенным локусом-мишенью и введенным экзогенным фрагментом ДНК. В конце 80-х данный способ успешно адаптировали для клеток млекопитающих [2, 3], и до сегодняшнего дня он используется при конструировании перевиваемых клеточных линий или для получения “нокаутных” животных [4]. Главное ограничение этого метода — низкая частота естественных рекомбинационных событий: в лучшем случае желаемая модификация возникает лишь в 1 из 10^5 клеток, получивших экзогенную ДНК [5], что делает его неприемлемым для работы с большинством первичных клеточных линий или с целыми организмами.

Дальнейшее развитие технологии стало возможным после того как, во-первых, было обнаружено, что образование двухцепочечного разрыва в ДНК в несколько тысяч раз увеличивает вероятность рекомбинации [6, 7], а, во-вторых, были открыты мегануклеазы — ферменты, распознающие и расщепляющие ДНК протяженностью более 14 п.н. Столь высокая избирательность нуклеаз позволяет вносить двухцепочечные разрывы в уникальный участок хромосомы, который далее подвергается модификации в процессе репарации. Чтобы метод стал универсальным, необходимо было сконструировать нуклеазы, специфичность которых можно изменять в соответствии с нуклеотидной последовательностью мишени. До недавнего времени работы велись в двух направлениях: перепрограммирование так называемых хоуминг-эндо-нуклеаз, которые кодируются мобильными элементами и встречаются во многих микроорганизмах [8], и конструирование искусственных нуклеаз с цинковыми пальцами (Zink Finger nuclease, ZF-нуклеаза) [9].

Наибольшее число хоуминг-эндо-нуклеаз с различной специфичностью создано на базе эндо-нуклеазы I-CreI, кодируемой хлоропластной ДНК сине-зеленой водоросли *Chlamydomonas reinhardtii*, способной селективно расщеплять последовательности длиной 19–24 п.н. [10]. Путем направленного изменения аминокислотных остатков I-CreI, участвующих во взаимодействии с ДНК, получили коллекцию, включающую несколько сотен вариантов нуклеаз, распознающих новые нуклеотидные последовательности [8, 11]. Возможность комбинировать субдомены различных белков полученной библиотеки позволила еще более расширить число потенциальных сайтов распознавания [11, 12]. Удалось отобрать нуклеазы, избирательно рас-

щепляющие два гена человека — *XPC* и *RAG1* [11, 12], а также уникальную последовательность в геноме кукурузы [13]. Однако несмотря на значительный прогресс, достигнутый в области перепрограммирования I-SceI, процесс конструирования нуклеаз с заданной специфичностью остается трудоемким, а выбор мишеней весьма ограниченным.

При конструировании ZF-нуклеаз объединили домен с неспецифической нуклеазной активностью и ДНК-связывающий домен [9]. Нуклеазный домен заимствовали из рестриктазы FokI, относящейся к ферментам рестрикции типа IIS. В качестве ДНК-связывающего домена использовали белковый домен, включающий несколько так называемых цинковых пальцев Cys₂-His₂ (ZF-домен), который входит в состав многих факторов транскрипции эукариот и является вторым по частоте встречаемости белковым доменом в геноме человека. Каждый цинковый палец состоит примерно из 30 аминокислотных остатков и образует пространственную конфигурацию ββα, которая координируется ионом цинка (рис. 1а). Несколько аминокислот на поверхности α-спирали контактируют с участком из 3 п.н. Именно они определяют специфичность домена. При создании искусственных нуклеаз последовательно объединяют от трех до шести пальцев, которые связывают 9–18 п.н. соответственно (рис. 1б).

Нуклеазный домен FokI функционирует только в виде димера, поэтому для расщепления заданного участка ДНК необходимо иметь пару ZF-нуклеаз. Одна нуклеаза связывается слева, а другая справа от области разрыва, при этом нуклеазные домены образуют димер и вносят двухцепочечный разрыв (рис. 1б). Таким образом, специфичность определяется как минимум 18 н. (если левый и правый ZF-домены состоят из трех пальцев), что дает возможность распознавать уникальные геномные последовательности. В первых нуклеазах использовали домены FokI дикого типа, которые образовывали гомодимеры. Позже было предложено использовать модифицированные домены, которые проявляли каталитическую активность только в виде гетеродимеров. Это позволило избежать образования активных димеров, состоящих из двух левых или двух правых нуклеаз, способных вносить разрыв в другие области генома.

ZF-нуклеазы сконструированы примерно для 100 генов, принадлежащих разным видам животных и растений, среди которых гены человека, мыши, крысы, свиньи, *Drosophila*, нематоды, рыбы *Danio rerio*, табака, сои и кукурузы [14–16]. С целью разработки подходов к генотерапии ряда наследственных заболеваний созданы ZF-нуклеазы, направленные на гены человека, связанные с этими заболеваниями — IL-2Rγ, VEGF-A, HoxB13, CFTR, β-глобина [17–19]. Еще один ген челове-

ка, ген рецептора CCR5, стал мишенью для ZF-нуклеазы в связи с поиском способов терапии при ВИЧ-инфекции [20]. Результаты разработки оказались настолько успешными, что возможность использования нуклеазы, избирательно расщепляющей ген *CCR5*, проходит проверку в двух клинических испытаниях [NCT01044654, NCT01252641].

Однако в целом создание ZF-нуклеаз, специфичных к заданной последовательности, оказалось непростой задачей. Надежды на предлагаемый модульный подход, при котором последовательно собранные цинковые пальцы должны распознавать соответствующие им триплеты ДНК, не оправдались. Наблюдалось перекрестное взаимодействие цинковых пальцев между собой и с соседними нуклеотидами, в результате полученные нуклеазы во многих случаях были неактивными или обладали неспецифической активностью [21]. Существуют другие алгоритмы сборки цинковых пальцев, в которых учитывается влияние соседних белковых повторов и нуклеотидов или проводится дополнительная селекция [22]. Однако все алгоритмы накладывают ограничения на выбор мишени, а некоторые из них весьма трудоемкие. Кроме того, нуклеазы с выявленной в пробирке ферментативной активностью в отношении специфической ДНК часто не работают при их использовании для редактирования эндогенного локуса в живых клетках.

Революция в создании нуклеаз для редактирования генома произошла в 2010 году, когда был предложен альтернативный инструмент для внесения специфического двухцепочечного разрыва — TALE-нуклеазы. Их появление связано с изучением TALE (Transcription Activators Like Effectors), белков с особой структурой ДНК-связывающего домена.

ИСТОРИЯ ИЗУЧЕНИЯ TALE

TALE — белки, секретируемые патогенами растений, бактериями рода *Xanthomonas*. Разные представители этого рода поражают многие сельскохозяйственные и декоративные растения, и TALE во многих случаях определяют вирулентность этих бактерий [23]. Однако в процессе эволюции некоторые растения выработали защитные механизмы, индуцируемые в ответ на присутствие TALE и помогающие противостоять патогену. В этом случае TALE обеспечивают устойчивость растений к патогенам, т.е. выступают в роли факторов авирулентности [23].

Впервые ген, кодирующий TALE, выделили из *X. campestris* pv. *vesicatoria*, которые паразитируют на растениях перца [24]. Бактерии, содержащие данный ген, вызывали реакцию гиперчувствительности у растений, обладающих геном устой-

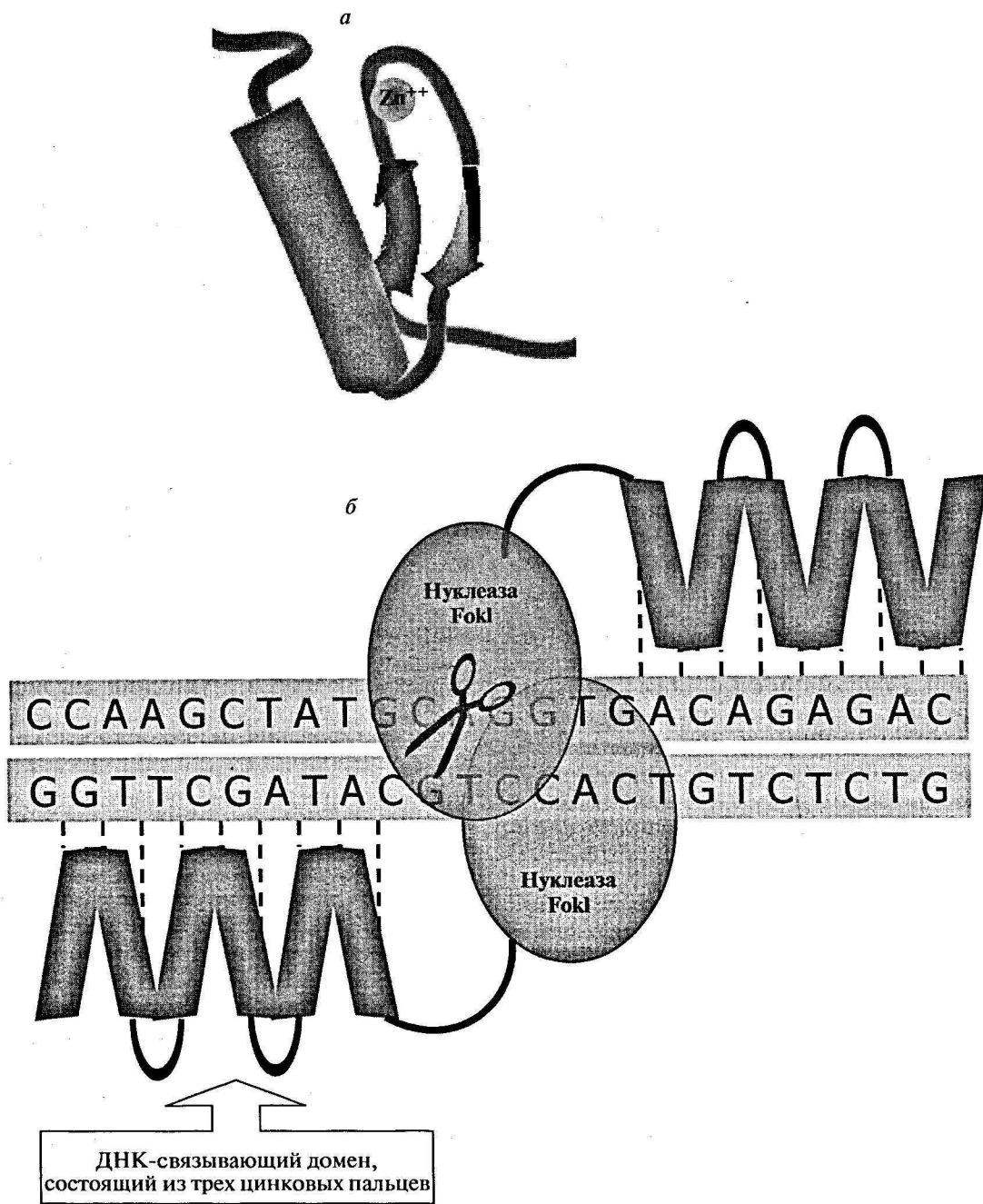


Рис. 1. Пространственная структура цинкового пальца (а) и схема связывания ZF-нуклеаз с ДНК (б).

чивости *Bs3*. Гиперчувствительность приводила к быстрому апоптозу клеток в зоне инфицирования, что предотвращало распространение возбудителя и предохраняло растение от гибели. Поэтому неизвестный тогда бактериальный ген был охарактеризован как ген фактора авирулентности и назван *avrBs3*, его нуклеотидная последовательность была прочитана и опубликована в 1989 году [24]. Се-

квенирование *avrBs3* было сопряжено с определенными трудностями, отчасти из-за его GC-богатого состава, но, главным образом, из-за 17 прямых повторов длиной 102 п.н. в открытой рамке считывания, гомология между которыми составляла от 91 до 100%. Кроме того, ген оказался необычно большим для бактерий, он состоял из 3495 п.н. и кодировал белок размером 122 кДа [25].

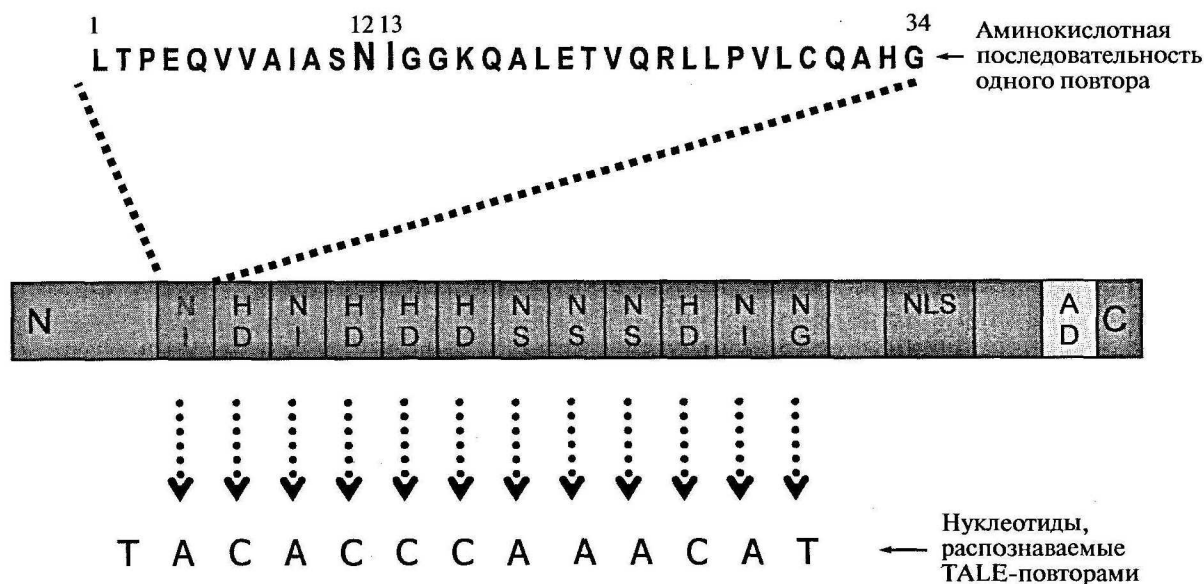


Рис. 2. Структура TALE. N и C – N и C концевые области белка, NLS – сигнал ядерной локализации, AD – домен, отвечающий за активацию транскрипции. В центральной части TALE находятся повторы, которые обеспечивают специфическое связывание белка с нуклеотидами.

В ходе дальнейших исследований гомологи AvrBs3 обнаружили у разных видов *Xanthomonas*, одни из них также играли роль факторов авирулентности, другие, наоборот, были необходимы для того, чтобы поражать чувствительные к патогену растения, а третьи могли играть ту или иную роль в зависимости от наличия у растения гена резистентности [26–28]. Белки этого семейства найдены и у отдельных изолятов бактерий *Ralstonia solanacearum* [29]. Число повторов в разных TALE варьировало от пяти до более 30 и было равно в среднем 17 [30]. Почти все они состояли из 34 аминокислотных остатков и имели гомологию не менее 90%, причем наиболее вариабельными были аминокислотные остатки в положениях 12 и 13, названные Repeat Variable Diresidue, или сокращенно RVD [30]. Необычное строение этих белков вызывало много вопросов, но окончательные ответы на них удалось получить только в 2009 году, через два десятилетия после прочтения последовательности гена *avrBs3*.

Первым шагом в понимании функции бактериального белка AvrBs3 стало неожиданное обнаружение в его составе сигнала ядерной локализации [31], а также локализации этого белка в ядре клеток растения во время инфекции [32]. Затем показали, что C-концевой домен AvrBs3 отвечает за активацию транскрипции, а сам белок способен связывать ДНК [33, 34]. По совокупности этих трех свойств – локализации в ядре, способности связываться с ДНК и способности активировать транскрипцию, присущих активаторам транскрипции, белок получил название TALE, от английского

Transcription Activator Like Effector, т.е. эффектор, похожий на активатор транскрипции [35]. В дальнейшем установили, что TALE действительно активируют экспрессию генов растений, взаимодействуя с фактором транскрипции TFIIA [36].

Долгое время неясным оставалось назначение большого количества практически идентичных повторов в этом и других белках из семейства TALE. Оказалось, что удаление части повторов из AvrBs3 приводило к потере активности белка в растениях с геном *Bs3*, но в некоторых случаях варианты с делециями становились факторами авирулентности для других сортов или видов растений [31]. То, что именно повторы, расположенные в определенном порядке, определяют специфичность связывания с ДНК, подтверждено экспериментально [32, 38]. В 2007 году обнаружили гены, экспрессию которых регулировал фактор AvrBs3, и в промоторе одного из них, *UPA20*, выявили последовательность, с которой связывается данный фактор [39]. Такие же последовательности найдены и в других генах-мишенях AvrBs3. Определение мишеней нескольких других TALE и стало ключевым фактором, который позволил понять, каким образом домен, состоящий из повторов, связывается с ДНК.

В 2009 году две группы ученых одновременно опубликовали в декабрьском номере журнала *Science* статьи, раскрывающие секрет TALE [40, 41]. Исследователи обратили внимание на тот факт, что число прямых повторов в TALE соответствует числу нуклеотидов в последовательности, с которой он взаимодействует, отсюда предположили, что один повтор распознает один нуклеотид.

а		
RVD	Нуклеотид	
HD	C	
NG	T; 5metC	
NI	A	
NN	G > A	
NS	A, C, G > T	
N-*	C, T; 5metC	
HG*	C, T	

б		
RVD	Нуклеотид	Взаимодействие
NN	G > A	Сильное
HN	G > A	Сильное
NK	G	Слабое
NH	G	Среднее

Рис. 3. а – Семь наиболее часто встречающихся RVD [41] и связываемые ими нуклеотиды. *Данные о специфичности этих RVD отличаются [40, 41, 49, 50]. б – RVD, распознающие гуанин, и сила взаимодействия этих RVD и нуклеотидов [50]. Знак > указывает на большую специфичность распознавания первых нуклеотидов. Жирным выделены RVD, используемые для конструирования TALEN.

Классификация повторов в соответствии с RVD позволила выявить корреляцию между типом повтора и нуклеотидом (рис. 2). В результате был написан простой код, в котором выбор нуклеотида диктуется двумя гипервариабельными остатками в положениях 12 и 13. Этот код оказался частично вырожденным: повторы некоторых типов могли связываться с разными нуклеотидами, в то время как другие в подавляющем большинстве случаев связывались с определенной буквой ДНК (рис. 3а, б) [42]. Последний повтор в ДНК-связывающем домене всегда был укороченным, он состоял из 20 аминокислотных остатков. Кроме того, сразу перед первым нуклеотидом последовательности-мишени обязательно находился тимин. Полученный код, с одной стороны, позволил успешно предсказывать последовательности-мишени уже известных TALE, с другой стороны, его использовали для конструирования новых искусственных TALE, взаимодействующих с заданной нуклеотидной последовательностью.

В 2012 году удалось провести рентгеноструктурный анализ двух TALE в комплексе с ДНК [43, 44]. Было обнаружено, что вторичная структура каждого повтора представляет собой две спирали, образующие при соединении V-образную форму, а гипервариабельные остатки находятся в середине петли, соединяющей эти спирали. Аминокислотные остатки в позициях 8 и 12 одного повтора

взаимодействуют друг с другом, стабилизируя структуру домена, а остаток 13 образует контакт со специфичным нуклеотидом. Повторы собраны в правозакрученную суперспираль, которая обвивается вокруг ДНК, следуя вдоль большой канавки таким образом, что петли с RVD обращены к ДНК (рис. 4).

СОЗДАНИЕ TALE-НУКЛЕАЗ

Открытие кода TALE привело к появлению нового инструмента – уникального ДНК-связывающего домена, специфичность которого легко изменять. По аналогии с ZF-нуклеазами этот домен был объединен с нуклеазным доменом FokI для получения TALE-нуклеаз, сокращенно TALEN (рис. 5). Уже в 2010 году была опубликована первая статья, описывающая проектирование и сборку TALEN, мишенями которых стали ген *ADH1* в геноме *Arabidopsis thaliana* и ген *gridlock* в геноме *D. rerio* [45].

Дальнейшие исследования позволили оптимизировать структуру искусственных TALEN. Оказалось, что для получения активного фермента в состав белка необходимо включать не только повторы, но и фланкирующие их N- и C-концевые области [46]. N-концевой домен должен состоять не меньше чем из 127 аминокислотных остатков. Эти остатки образуют четыре сильно



Рис. 4. Пространственная структура ДНК-связывающего домена TALE.

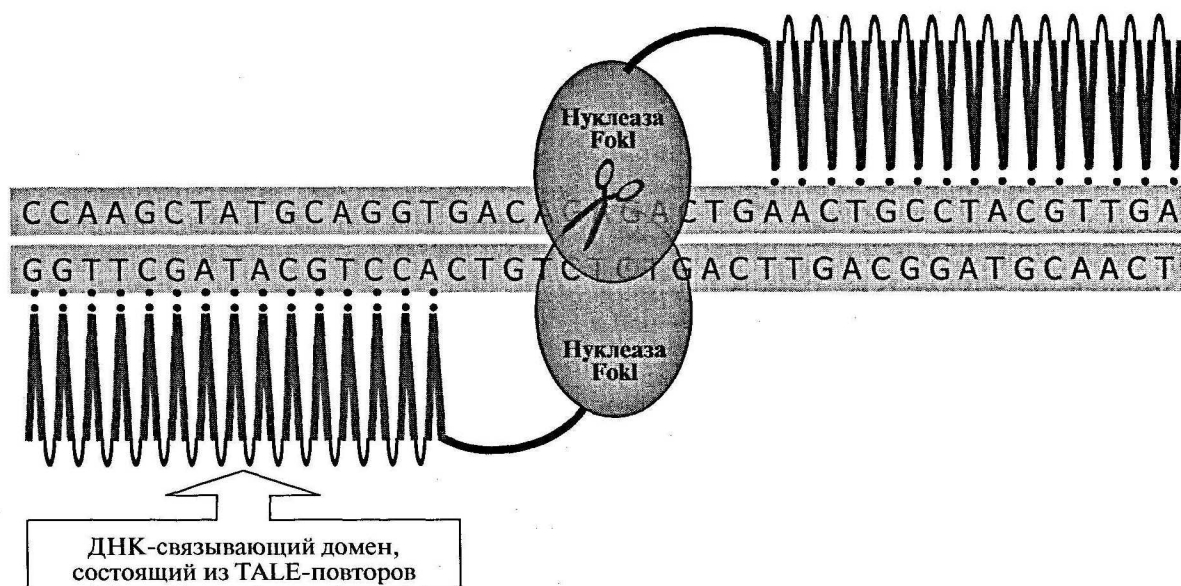


Рис. 5. Схема связывания TALEN с ДНК.

вырожденных повтора, которые отвечают за связывание тимина в позиции -1, а также обеспечивают неспецифическое связывание с ДНК. Длина С-концевого домена может варьировать от 18 до 207 аминокислотных остатков, но его размер влияет на оптимальное число нуклеотидов, находящихся между местами посадки левой и правой частей нуклеазы. Наибольшую активность *in vivo* проявляли нуклеазы GoldyTALEN, которые содержат 136 аминокислот в N-концевой области, 63 – в С-концевой области, а также характеризуются некоторыми особенностями структурных элементов [47, 48].

Получена дополнительная информация о разнообразии повторов и о специфичности кода. Первоначально было описано 15 различных RVD, среди которых семь встречаются часто (рис. 3а), а остальные в единичных случаях (таблица). Наиболее строгое соответствие выявлено между следующими парами аминокислотных остатков и нуклеотидами: NN–G/A, NI–A, HD–C, NG–T, которые и используются для конструирования TALEN (рис. 3а). Поскольку повтор NN, распознающий G, недостаточно специфичен и может образовывать связь с аденином, был проведен поиск новых повторов и тестирование их специфичности. В результате было выявлено 13 неизвестных ранее RVD (таблица) и дополнительно охарактеризована специфичность всех повторов. Обнаружено, что повтор, содержащий NK, способен связывать гуанин более специфично [47], однако TALEN, использующие эти повторы, обладают значительно меньшей активностью [46, 49, 50]. Причиной этого оказалось более слабое

взаимодействие данного повтора с соответствующим нуклеотидом. В дальнейшем удалось найти другой редко встречающийся в природе повтор NH, который распознает гуанин с большей селективностью, чем NN, активность нуклеазы, содержащей NH-повторы, при этом меняется незначительно (рис. 3б) [49, 50]. Показано также, что активность TALEN определяется не только специфичностью, но и таким важным фактором, как сила взаимодействия повторов и нуклеотидов [50].

В одной из работ показано, что повторы NG и HG, селективно связывающие тимин, могут взаимодействовать с 5'-метилцитозином, что позволяет конструировать TALEN для метилированных участков, а также с помощью нуклеаз отличать ДНК-мишени, содержащие метилированные CpG-участки, от немодифицированных участков [43].

Список RVD, встречающихся в природе

Первые 15 RVD, обнаруженные в бактериях рода *Xanthomonas* [42]

HD, NG, NI, NN, NS, N–, HG – часто встречающиеся
HA, ND, NK, H–, HI, HN, IG, NA – редкие

Дополнительные RVD, обнаруженные в бактериях рода *Xanthomonas* [49]

NC, NQ, SS, SN, S–, NV, NH, YG

Дополнительные RVD, обнаруженные только в бактериях *Ralstonia solanacearum* [50]

NT, SH, NP, SN, IS

СБОРКА TALEN: СУЩЕСТВУЮЩИЕ ПЛАТФОРМЫ

Получение конструкций, кодирующих TALEN, представляет определенные трудности, так как необходимо объединить большое число практически идентичных повторяющихся последовательностей. На сегодняшний день предложено несколько технологических платформ для сборки таких конструкций, из которых наиболее известны безлигазный метод сборки плазмид (LIC, Ligation-Independent Cloning) [51], клонирование методом Golden Gate [52], метод REAL (Restriction Enzyme And Ligation) [53] и FLASH (Fast Ligation-based Automatable solid-phase High-throughput) [54]. Во всех случаях сборка проводится из уже готовых последовательностей, кодирующих один повтор или комбинацию из нескольких повторов.

Платформы отличаются технологиями клонирования, количеством необходимых исходных плазмид и праймеров, гибкостью в отношении получаемого числа повторов, временем сборки, возможностью масштабировать синтез, доступностью реагентов и протоколов, а также наличием программ для дизайна нуклеаз. Усилиями нескольких лабораторий многие наборы для сборки доступны через некоммерческую систему распространения плазмид Addgene. Создан отдельный сайт со ссылками на протоколы, реагенты и другую информацию о технологии TALE (TALengineering.org). Более того, коммерческие услуги по сборке TALEN и валидации их активности предоставляют несколько компаний — “Cellestis bioresearch” (Франция), “Labomics S.A.” (Бельгия), “Life Technologies” (США), “Transposagen Biopharmaceuticals, Inc.” (США), “Vigene Biotechnology” (США).

СПОСОБЫ РЕДАКТИРОВАНИЯ ГЕНОМА ПРИ ПОМОЩИ НУКЛЕАЗ

Двухцепочечный разрыв, созданный нуклеазой в последовательности-мишени, активирует процесс репарации ДНК в клетке. Существуют два основных пути восстановления целостности ДНК: негомологичное соединение концов и гомологичная рекомбинация. Ниже рассмотрены способы внесения различных изменений в геном с помощью высокоспецифичных нуклеаз и негомологичного соединения концов ДНК или гомологичной рекомбинации. Получить полное представление о возможностях редактирования генома помогут данные, полученные при использовании как TALEN, так и ZF-нуклеаз. Первые ZF-нуклеазы были созданы намного раньше, чем TALEN, поэтому многие геномные модификации впервые получили именно с их помощью. Появление TALEN позволило проводить манипуляции с другими генами или организмами во многих неспе-

циализированных лабораториях. Следует отметить, что в большинстве случаев тип нуклеазы не влияет на результат редактирования, а лишь обуславливает эффективность и специфичность процесса и доступность мишени, используемой в конкретной работе. Преимущества TALEN, а также некоторые возможности, связанные с особенностями данных нуклеаз, будут описаны в следующих разделах.

Редактирование генома с помощью негомологичного соединения концов ДНК

Негомологичное соединение концов ДНК сопровождается образованием небольших делеций или вставок в месте разрыва, которые в большинстве случаев нарушают функцию гена [55] (рис. 6). Мутагенез посредством негомологичного соединения концов очень удобен для направленного “выключения” генов. Если внести двухцепочечный разрыв в экзон, то теоретически можно ожидать, что 66% случайных вставок-делеций будут приводить к мутациям за счет сдвига рамки считывания. Наблюдаемая частота мутаций еще выше, поскольку соединению концов предшествует достройка 5'-выступающих одноцепочечных концов ДНК размером 4–5 н., которые образуются после расщепления нуклеазой FokI [56].

На сегодняшний день в подавляющем большинстве работ высокоспецифичные нуклеазы применялись именно для создания “нокаутных мутаций” за счет негомологичного соединения концов ДНК. В зависимости от нуклеазы, типа клеток и статуса хроматина гена-мишени количество клеток, которые удалось модифицировать, варьировало от менее 1% (в частности для эмбриональных стволовых или индуцированных плюрипотентных клеток) [57, 58] до практически 100% [48]. У 33% модифицированных клеток мутации обнаружены в обоих аллелях [18]. Столь высокая эффективность упростила получение одноклеточных эмбрионов или линий эмбриональных клеток с инактивированными генами-мишенями, что сделало возможным создание животных и растений с мутациями в заданных генах. Учитывая, что последние 20 лет единственным объектом, доступным для направленного изучения функций генов *in vivo*, были “нокаутные” мыши, технология редактирования генома открыла новый этап в биологии и генетике, когда выключение отдельных генов и изучение их фенотипических проявлений стало возможным и в других сложных организмах. С помощью этой технологии уже созданы крысы, рыбы, растения кукурузы и табака, в геноме которых инактивированы отдельные гены [16].

Выключение генов с помощью негомологичного соединения концов ДНК уже нашло свое применение в медицине. С 2009 года проводятся кли-

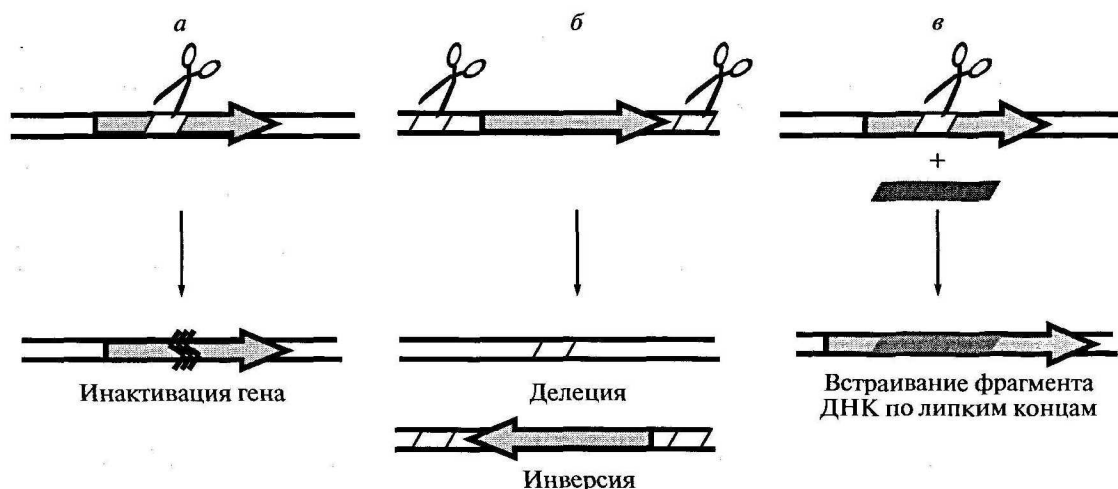


Рис. 6. Редактирование генома с помощью нехомологичного соединения концов ДНК.

нические испытания, в которых с помощью ZF-нуклеазы осуществляют нокаут гена рецептора CCR5 в Т-лимфоцитах *ex vivo* при ВИЧ-инфекции [NCT00842634, NCT01044654, NCT01252641]. Начаты клинические испытания, в которых проверяется эффективность инактивации рецептора глюкокортикоидов в Т-клетках при злокачественной глиоме [NCT01082926]. Предложено использовать ZF-нуклеазу для разрушения эндогенного Т-клеточного рецептора в клеточной иммунотерапии, при которой вводят новый рецептор, обладающий высокой противоопухолевой специфичностью [59, 60].

Показана возможность и более сложных геномных изменений в клетках млекопитающих с использованием механизма нехомологичного соединения концов. Создание с помощью двух пар нуклеаз двух значительно удаленных друг от друга двухцепочечных разрывов в одной хромосоме приводит к образованию протяженных делеций, возникающих после лигирования удаленных концов ДНК. Таким образом удалось получить хромосомные делеции размером до 15 млн.п.н., частота которых в культивируемых клетках человека составила 0.1–10% [61]. Показано также, что с помощью двух двухцепочечных разрывов можно индуцировать дупликации или инверсии хромосом (рис. 6) [62, 63].

Интерес представляет и использование механизма нехомологичного соединения концов для точного встраивания донорской ДНК. В этом случае наряду с высокоспецифичной нуклеазой в клетку вводят ДНК-фрагменты с выступающими концами, комплементарные 5'-выступающим липким концам, которые образуются после расщепления хромосомы нуклеазой FokI (рис. 6). Используя данный способ, в хромосому культивируемых кле-

ток человека удалось с эффективностью 10% интегрировать короткий фрагмент ДНК [56]. Недавно показали возможность встраивания таким же способом более протяженного фрагмента ДНК (15000 п.н.) в заданный локус генома линий клеток человека и мыши [64].

Редактирование генома с помощью гомологичной рекомбинации

Для репарации ДНК путем гомологичной рекомбинации необходима донорская ДНК, обладающая гомологией с последовательностью в области разрыва. В результате гомологичной рекомбинации хромосомная последовательность замещается донорской ДНК [65]. Одно из генетических изменений, которое можно осуществить с помощью гомологичной рекомбинации – добавление последовательностей генов в нужные участки хромосомы (рис. 7). В 2009 году Hockemeyer и соавт. [57] показали возможность встраивания репортерного гена в три различных локуса генома индуцированных плюрипотентных клеток человека (iPSC) с помощью ZF-нуклеазы. Двумя годами позже фрагменты ДНК ввели в те же локусы с помощью TALEN [58]. Одна из трех изученных в этих работах мишеней, локус *AAVS1*, расположен на хромосоме 19 человека и является местом интеграции аденоассоциированного вируса. Этот локус интересен тем, что встраивание экзогенных ДНК в область *AAVS1* не влияет на экспрессию окружающих генов и не вызывает злокачественного перерождения клеток, при этом активная транскрипция транскрипта может поддерживаться длительное время. Области генома человека, обладающие такими свойствами, получили название “безопасной гавани” (safe harbor). Кроме *AAVS1*, к ним относят локус *ROSA26* и ген *CCR5* [16].

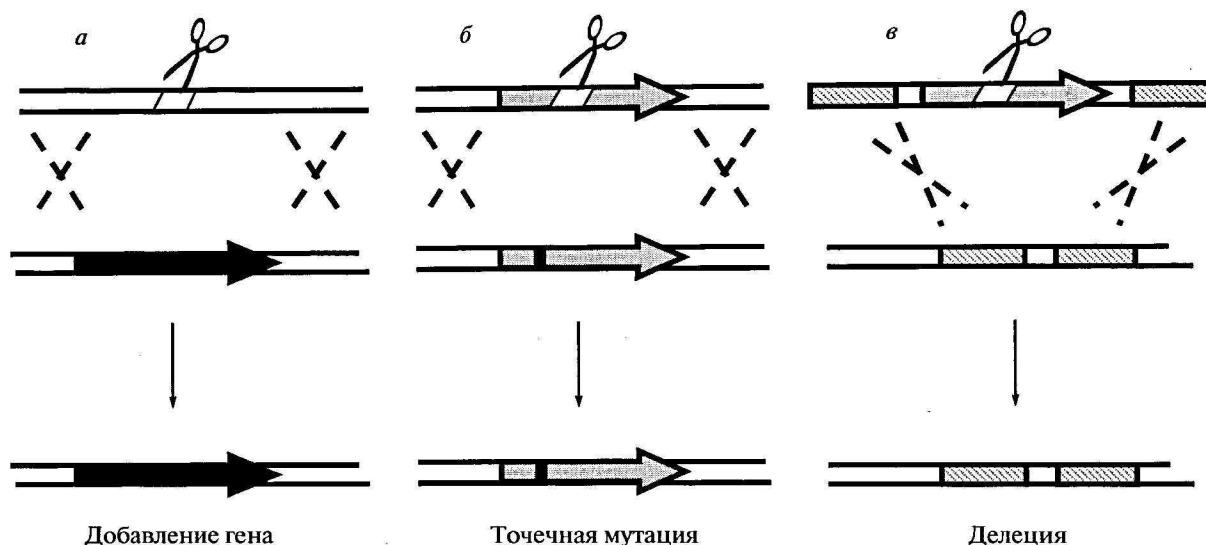


Рис. 7. Редактирование генома с помощью гомологичной рекомбинации.

Направленное встраивание терапевтического гена в “безопасную гавань” позволяет избежать генотоксичности, наблюдаемой в нескольких клинических испытаниях, в которых для интеграции трансгена использовали вирусные векторы. Примером может служить попытка генотерапии X-сцепленной хронической гранулематозной болезни — врожденного иммунодефицита, обусловленного мутацией в гене *gp91(phox)*. Мутация в этом гене приводит к тому, что нейтрофилы и макрофаги не способны генерировать активные формы кислорода и уничтожать грибы и бактерии. В клиническом испытании функциональный ген *gp91(phox)* встраивали в гемопоэтические стволовые клетки с помощью γ -ретровирусного вектора. Последующая пересадка этих клеток трем больным вызвала развитие у них миелоидного лейкоза. Причиной злокачественного перерождения клеток стало нарушение регуляции генов *MDS1-EV11*, *PRDM16* или *SETBP1*, рядом с которыми произошла интеграция трансгена [66]. В настоящее время разрабатывается альтернативный подход, где ген *gp91(phox)* направленно встраивается в локус *AAVS1* путем внесения разрыва с помощью ZF-нуклеазы и последующей гомологичной рекомбинации с донорской ДНК, содержащей ген *gp91(phox)*. В лабораторных опытах модифицированные индуцированные плюрипотентные стволовые клетки, в геном которых встроен данный ген, успешно дифференцировались *in vitro* в зрелые нейтрофилы, при этом не наблюдалось их злокачественной трансформации. В полученных нейтрофилах сохранялась стабильная экспрессия трансгена, что привело к восстановлению в них нормального окислительного метаболизма [67].

Успешное встраивание гена в требуемую область генома различных организмов описано еще в нескольких работах [68, 69]. В ходе разработки препарата против гемофилии показали возможность направленного добавления терапевтического гена *in vivo*. В клетки печени гуманизированных мышей, у которых мутация в гене фактора свертывания крови IX приводила к отсутствию данного белка в крови, с помощью аденоассоциированных векторов доставляли ДНК, кодирующую ZF-нуклеазу, и ДНК, содержащую скорректированные экзоны 2–8 гена *F9*. Встраивание донорской ДНК сразу за экзоном 1 гена *F9*, выявленное в 3–7% клеток, оказалось достаточным для восстановления гемостаза у животных [70].

Гомологичную рекомбинацию можно использовать и для внесения единичных нуклеотидных замен в интересующий ген (рис. 7), например для создания организмов, несущих определенный аллельный вариант или для исправления точечной мутации при наследственном заболевании. Оказалось, что в этих случаях в качестве донорской ДНК можно использовать одноцепочечные олигонуклеотиды. В результате микроинъекции в одноклеточные эмбрионы олигонуклеотидов, соответствующих модифицированной последовательности гена, и мРНК, кодирующей TALEN, были получены мыши, из которых 2% несли заданные изменения в геноме. Сходным образом для модификации генов *ponzr1* и *crhr2* применили микроинъекции олигонуклеотидов TALEN мРНК в протоплазму эмбрионов *D. rerio* [48]. С помощью одноцепочечных олигонуклеотидов можно получать также хромосомные делеции размером до 10000 п.н., в этом случае одна часть олигонуклео-

тида должна быть гомологичной удаленному участку хромосомы [71] (рис. 7).

По современным представлениям 80–90% генетических заболеваний человека обусловлены заменой, вставкой или делецией всего одного нуклеотида [72]. Перспективной стратегией при таких заболеваниях может стать исправление мутаций в индуцированных плюрипотентных клетках, полученных от больных. Далее эти клетки могут дифференцироваться в нужные клетки и использоваться для клеточной терапии. Опубликованы работы, в которых в индуцированных плюрипотентных стволовых клетках, полученных из клеток больных, удалось исправить мутации, вызывающие X-сцепленный тяжелый иммунодефицит, дефицит α -антитрипсина, гемофилию В, болезнь Паркинсона и серповидно-клеточную анемию [19, 73–75]. Основная проблема при внесении изменений в индуцированные плюрипотентные стволовые клетки – относительно низкая частота модификаций, при которой необходимо проводить отбор модифицированных клеток. Например, для отбора клеток, в которых произошла замена мутантного нуклеотида в гене α -синуклеина, проводили скрининг изолированных колоний клеток. Правильный клон удалось найти после проверки 240 колоний [73]. В других работах использовали вектор, содержащий дополнительный ген для селекции. Однако для дальнейшего клинического применения неприемлемо присутствие селективного гена, поэтому после отбора клеток этот ген удаляли с помощью системы рекомбинации Cre/loxP или транспозазы PiggyBack [19, 74, 75]. Манипуляция с помощью системы Cre/loxP оставляла “след” в геноме в виде короткой (34 п.н.) последовательности loxP, которая, как оказалось, может влиять на экспрессию исправленного гена. С помощью транспозазы удалось более аккуратно вырезать фрагмент, содержащий селективный ген, не изменив при этом последовательность геномной ДНК [74].

ПРЕИМУЩЕСТВО TALE-НУКЛЕАЗ

Достаточно быстро стало очевидным, что TALE-нуклеазы превосходят своих предшественников, ZF-нуклеазы, по многим параметрам. Прежде всего, их отличает высокая модульность, позволяющая проектировать нуклеазы, не прибегая к дополнительным алгоритмам для анализа перекрестных взаимодействий. Опубликованные результаты свидетельствуют о том, что почти все TALE-нуклеазы, собранные по схеме один повтор – один нуклеотид, обладали некоторой активностью *in vivo*, расщепляя хромосомы в клетках, причем уровень этой активности в среднем был выше, чем у ZF-нуклеаз. Reyon и соавт. [54] для оценки эффективности конструирования TALEN провели сборку 96 пар нуклеаз, направ-

ленных на различные гены-мишени, используя высокопроизводительную технологию FLASH. Оказалось, что 84 пары этих нуклеаз (88%) обладали активностью в отношении целевых участков хромосом, результат, недостижимый ранее для ZF-нуклеаз, даже при использовании сложных алгоритмов и предварительного отбора вариантов *in vitro*. С TALEN столь высокую эффективность удалось обеспечить без предварительного скрининга. В среднем удалось модифицировать 22.2% клеток, однако подобный результат достигался только при использовании отдельных ZF-нуклеаз. Непосредственное сравнение двух типов нуклеаз показало, что использование TALEN приводило к возникновению мутаций у 20–77% эмбрионов *D. rerio*, в то время как ZF-нуклеаз – только у 1.1–3.3% [76]. В одной из работ использование Golden TALEN позволило довести содержание модифицированных клеток до 100% [48].

В отличие от ZF-нуклеаз TALEN могут быть нацелены практически на любую интересующую последовательность в геноме. Единственное ограничение при выборе мишени для классических TALEN – наличие тимины непосредственно с 5'-конца от мишени, однако уже найдена возможность преодолеть и это ограничение [77].

Дополнительное преимущество TALEN, которое расширяет возможности применения этого фермента, – существование повторов с различной специфичностью. Некоторые повторы высокоспецифичны, в то время как другие могут распознавать все пурины или связываться с любым из нуклеотидов. Такое свойство позволяет конструировать TALEN, которые связываются с выродженными или содержащими SNP последовательностями. Кроме того, как уже упоминалось, TALEN можно использовать для распознавания метилированных участков.

Дизайн TALEN подразумевает конструирование белков, специфичность которых определяется 30–36 н., в то время как у ZF-нуклеаз только 18–24 п.н. Большая специфичность TALEN обуславливает меньшую вероятность расщепления нецелевых последовательностей, что снижает уровень цитотоксичности. В нескольких работах показано, что токсичность TALEN ниже или такая же, как у ZF-нуклеаз [54, 76].

Создание TALEN вызвало огромный интерес. Число работ, посвященных TALE и TALE-нуклеазам, растет экспоненциально (рис. 8). Весьма показательно количество генов и организмов, для которых собраны и опробованы TALE-нуклеазы. Разные группы ученых использовали TALEN для модификации генов дрожжей, плодовой мушки, рыб, лягушек, крыс, свиней, крупного рогатого скота, риса, тутового шелкопряда, сверчка, соматических и плюрипотентных стволовых клеток человека, сконструировав уникальные нуклеазы

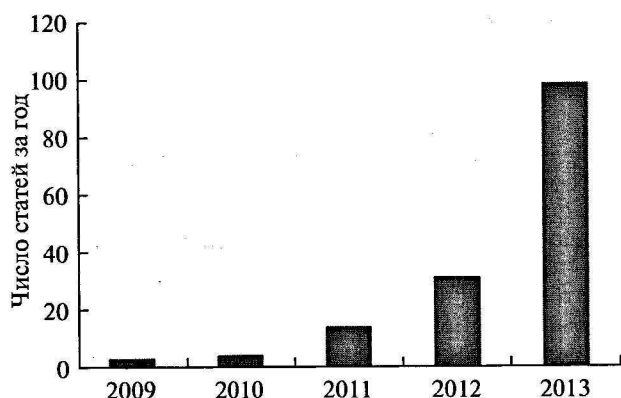


Рис. 8. Рост числа публикаций, посвященных TALE и TALE-нуклеазам.

более чем для 200 генов [16, 46]. Группа корейских ученых создала библиотеку TALEN, включающую около 18000 пар нуклеаз, мишенями для которых стала большая часть генов человека [78].

Интересно, что совсем недавно, в 2012 году, появилась еще одна система, позволяющая направленно вносить двухцепочечный разрыв в ДНК — система CRISPR-Cas. В основе этой системы лежит механизм, который выполняет роль адаптивного иммунитета у бактерий и архей и защищает клетку-хозяина от инвазии нуклеиновых кислот тех вирусов и плазмид, с которыми эти организмы уже встречались. Предложена упрощенная система CRISPR-Cas, состоящая всего из одного белка Cas9 и одной химерной РНК. Эта система позволяет специфически расщеплять ДНК, при этом специфичность определяется 20 н., которые входят в состав химерной РНК длиной примерно 100 н. [79]. Поэтому, чтобы получить нуклеазу с новой специфичностью, достаточно синтезировать новую РНК. Подобная гибкость и простота конструирования высокоспецифичной нуклеазы вызвали огромный интерес. Было высказано предположение, что использование данной системы может оказаться более простым и удобным, чем конструирование TALE-нуклеаз. Однако для того, чтобы полностью охарактеризовать потенциал технологии CRISPR-Cas, необходимо дальнейшее ее изучение. Есть опасения, что существенным недостатком CRISPR-Cas при редактировании сложных геномов может стать невысокий уровень специфичности [80].

ДОПОЛНИТЕЛЬНЫЕ ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ ДНК-СВЯЗЫВАЮЩЕГО ДОМЕНА TALE

Значительный интерес вызывает использование ДНК-связывающего домена TALE с целью направленной доставки белкового домена, кото-

рый обладает не нуклеазной, а другой активностью, способной влиять на функционирование генома. Например, описаны искусственные белки, полученные в результате объединения TALE-повторов и белковых доменов, обладающих рекомбиназной активностью [81, 82]. Такие белки могут стать альтернативой TALEN при проведении сайт-специфической интеграции. В этом случае процесс интеграции проходит без участия клеточных белков репарации, что позволяет добиться более эффективного и регулируемого встраивания гена в геном клеток любого типа.

В нескольких работах ДНК-связывающий домен TALE объединяли с доменом, обладающим активностью репрессора или активатора транскрипции. Такие конструкции позволяли изменять экспрессию целевых генов в клетках растений и человека в 2–30 раз [49, 83–85]. Направленно изменять экспрессию отдельных генов или даже целых геномных локусов можно, присоединяя белки, влияющие на эпигенетический статус ДНК, такие как метилтрансфераза или гистондеацетилаза. Этот подход применили в нескольких работах, в которых ДНК-связывающим доменом служили белковые домены с цинковыми пальцами [86, 87]. В целом, создание гибридных белков может привести к тому, что появится возможность не только вносить изменения в геном, но и влиять на экспрессию отдельных генов в геноме, регулируя транскрипцию или контролируя уровень метилирования ДНК.

ЗАКЛЮЧЕНИЕ

Редактирование генома с помощью высоко-специфичных нуклеаз — активно развивающаяся технология, которая открывает много новых возможностей. С ее помощью можно осуществлять манипуляции со сложными геномами в клетках и живых организмах, добавлять или удалять гены, вносить желаемые замены, делеции и вставки TALE-нуклеазы открыли новый этап в развитии этой технологии. Простота дизайна, отсутствие ограничений на нуклеотидную последовательность и высокий уровень успеха сделали редактирование генома с помощью TALEN доступным и способствовали широкому распространению данной технологии. Результаты сегодняшних исследований говорят о высоком потенциале TALE-нуклеаз, которые могут стать мощным инструментом для решения самых разных задач не только в науке, но также в биотехнологии и медицине.

СПИСОК ЛИТЕРАТУРЫ

1. Hinnen A., Hicks J.B., Fink G.R. 1978. Transformation of yeast. *Proc. Natl. Acad. Sci. USA*. 75, 1929–1933.

2. Thomas K.R., Folger K.R., Capecchi M.R. 1986. High frequency targeting of genes to specific sites in the mammalian genome. *Cell*. **44**, 419–428.
3. Doetschman T., Gregg R.G., Maeda N., Hooper M.L., Melton D.W., Thompson S.S.O. 1987. Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature*. **330**, 576–578.
4. Capecchi M. 2001. Generating mice with targeted mutations. *Nat. Med.* **7**, 1086–1090.
5. Vasquez K.M., Marburger K., Intody Z., Wilson J.H. 2001. Manipulating the mammalian genome by homologous recombination. *Proc. Natl. Acad. Sci. USA*. **98**, 8403–8410.
6. Donoho G., Jasin M., Berg P. 1998. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol. Cell Biol.* **18**, 4070–4078.
7. Choulika A., Perrin A., Dujon B., Nicolas J.-F. 1995. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **15**, 1968–1973.
8. Arnould S., Chames P., Perez C., Lacroix E., Duclert A., Epinat J.-C., Stricher F., Petit A.-S., Patin A., Guillier S., Rolland S., Prieto J., Blanco F.J., Bravo J., Montoya G., Serrano L., Duchateau P., Pâques F. 2006. Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J. Mol. Biol.* **355**, 443–458.
9. Durai S., Mani M., Kandavelou K., Wu J., Porteus M.H., Chandrasegaran S. 2005. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucl. Acids Res.* **33**, 5978–5990.
10. Dürrenberger F., Rochaix J.D. 1993. Characterization of the cleavage site and the recognition sequence of the I-CreI DNA endonuclease encoded by the chloroplast ribosomal intron of *Chlamydomonas reinhardtii*. *Mol. Gen. Genet.* **236**, 409–414.
11. Smith J., Grizot S., Arnould S., Duclert A., Epinat J.-C., Chames P., Prieto J., Redondo P., Blanco F.J., Bravo J., Montoya G., Pâques F., Duchateau P. 2006. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucl. Acids Res.* **34**, e149.
12. Arnould S., Perez C., Cabaniols J.-P., Smith J., Gouble A., Grizot S., Epinat J.-C., Duclert A., Duchateau P., Pâques F. 2007. Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J. Mol. Biol.* **371**, 49–65.
13. Gao H., Smith J., Yang M., Jones S., Djukanovic V., Nicholson M.G., West A., Bidney D., Falco S.C., Jantz D., Lyznik L.A. 2010. Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J.* **61**, 176–187.
14. Carroll D. 2008. Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Ther.* **15**, 1463–1468.
15. Molina R., Montoya G., Prieto J. 2011. in *Encyclopedia of Life Sciences (ELS)*. Chichester: John Wiley & Sons, Ltd, 1–10.
16. Segal D.J., Meckler J.F. 2013. Genome engineering at the dawn of the golden age. *Annu. Rev. Genomics Hum. Genet.* **14**, 135–158.
17. Maeder M.L., Thibodeau-Beganny S., Osiaik A., Wright D.A., Anthony R.M., Eichinger M., Jiang T., Foley J.E., Winfrey R.J., Townsend J.A., Unger-Wallace E., Sander J.D., Müller-Lerch F., Fu F., Pearlberg J., Göbel C., Dassie J.P., Pruett-Miller S.M., Porteus M.H., Sgroi D.C., Iafrate A.J., Dobbs D., McCray P.B., Cathomen T., Voytas D.F., Joung J.K. 2008. Rapid “open-source” engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell*. **31**, 294–301.
18. Urnov F.D., Miller J.C., Lee Y.-L., Beausejour C.M., Rock J.M., Augustus S., Jamieson A.C., Porteus M.H., Gregory P.D., Holmes M.C. 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. **435**, 646–651.
19. Sebastiano V., Maeder M.L., Angstrom J.F., Haddad B., Khayter C., Yeo D.T., Goodwin M.J., Hawkins J.S., Ramirez C.L., Batista L.F.Z., Artandi S.E., Wernig M., Joung J.K. 2011. *In situ* genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. *Stem Cells*. **29**, 1717–1726.
20. Perez E.E., Wang J., Miller J.C., Jouvenot Y., Kim K.A., Liu O., Wang N., Lee G., Bartsevich V.V., Lee Y.-L., Guschin D.Y., Rupniewski I., Waite A.J., Carpenito C., Carroll R.G., Orange J.S., Urnov F.D., Rebar E.J., Ando D., Gregory P.D., Riley J.L., Holmes M.C., June C.H. 2008. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816.
21. Ramirez C.L., Foley J.E., Wright D.A., Müller-Lerch F., Rahman S.H., Cornu T.I., Winfrey R.J., Sander J.D., Fu F., Townsend J.A., Cathomen T., Voytas D.F., Joung J.K. 2008. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat. Methods*. **5**, 374–375.
22. Joung J.K., Sander J.D. 2013. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* **14**, 49–55.
23. Bogdanove A.J., Schornack S., Lahaye T. 2010. TAL effectors: finding plant genes for disease and defense. *Curr. Opin. Plant Biol.* **13**, 394–401.
24. Bonas U., Stall R.E., Staskawicz B. 1989. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol. Gen. Genet.* **218**, 127–136.
25. Knoop V., Staskawicz B., Bonas U. 1991. Expression of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria* is not under the control of *hrp* genes and is independent of plant factors. *J. Bacteriol.* **173**, 7142–7150.
26. De Feyter R., Yang Y., Gabriel D.W. 1993. Gene-for-genes interactions between cotton *R* genes and *Xanth-*

- omonas campestris* pv. *malvacearum* avr genes. *Mol. Plant Microbe Interact.* **6**, 225–237.
27. Swarup S., De Feyter R., Brlansky R.H., Gabriel D.W. 1991. A pathogenicity locus from *Xanthomonas citri* enables strains from several pathovars of *X. campestris* to elicit cankerlike lesions on citrus. *Phytopathology*. **81**, 802–809.
 28. Swarup S., Yang Y., Kingsley M.T., Gabriel D.W. 1992. An *Xanthomonas citri* pathogenicity gene, *pthA*, pleiotropically encodes gratuitous avirulence on nonhosts. *Mol. Plant Microbe Interact.* **5**, 204–213.
 29. Heuer H., Yin Y.-N., Xue Q.-Y., Smalla K., Guo J.-H. 2007. Repeat domain diversity of *avrBs3*-like genes in *Ralstonia solanacearum* strains and association with host preferences in the field. *Appl. Environ. Microbiol.* **73**, 4379–4384.
 30. Boch J., Bonas U. 2010. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* **48**, 419–436.
 31. Yang Y., Gabriel D.W. 1995. *Xanthomonas* avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. *Mol. Plant Microbe Interact.* **8**, 627–631.
 32. Van den Ackerveken G., Marois E., Bonas U. 1996. Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. *Cell*. **87**, 1307–1316.
 33. Zhu W., Yang B., Chittoor J.M., Johnson L.B., White F.F. 1998. AvrXa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. *Mol. Plant Microbe Interact.* **11**, 824–832.
 34. Yang B., Zhu W., Johnson L.B., White F.F. 2000. The virulence factor AvrXa7 of *Xanthomonas oryzae* pv. *oryzae* is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein. *Proc. Natl. Acad. Sci. USA*. **97**, 9807–9812.
 35. Sugio A., Yang B., Zhu T., White F.F. 2007. Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes *OstFIIAgamma1* and *OstTFXI* during bacterial blight of rice. *Proc. Natl. Acad. Sci. USA*. **104**, 10720–10735.
 36. Iyer A.S., McCouch S.R. 2004. The rice bacterial blight resistance gene *xa5* encodes a novel form of disease resistance. *Mol. Plant Microbe Interact.* **17**, 1348–1354.
 37. Herbers K., Conrads-Strauch J., Bonas U. 1992. Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein. *Nature*. **356**, 172–174.
 38. Yang B., Zhu W., Johnson L.B., White F.F. 2000. The virulence factor AvrXa7 of *Xanthomonas oryzae* pv. *oryzae* is a type III secretion pathway-dependent nuclear-localized double-stranded DNA-binding protein. *Proc. Natl. Acad. Sci. USA*. **97**, 9807–9812.
 39. Kay S., Hahn S., Marois E., Hause G., Bonas U. 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science*. **318**, 648–651.
 40. Boch J., Scholze H., Schornack S., Landgraf A., Hahn S., Kay S., Lahaye T., Nickstadt A., Bonas U. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*. **326**, 1509–1512.
 41. Moscou M.J., Bogdanove A.J. 2009. A simple cipher governs DNA recognition by TAL effectors. *Science*. **326**, 1501.
 42. Bogdanove A.J., Voytas D.F. 2011. TAL effectors: customizable proteins for DNA targeting. *Science*. **333**, 1843–1846.
 43. Deng D., Yin P., Yan C., Pan X., Gong X., Qi S., Xie T., Mahfouz M., Zhu J.-K., Yan N., Shi Y. 2012. Recognition of methylated DNA by TAL effectors. *Cell Res.* **22**, 1502–1504.
 44. Mak A.N., Bradley P., Cernadas R.A., Bogdanove A.J., Stoddard B.L. 2012. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science*. **335**, 716–719.
 45. Christian M., Cermak T., Doyle E.L., Schmidt C., Zhang F., Hummel A., Bogdanove A.J., Voytas D.F. 2010. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*. **186**, 757–761.
 46. Sun N., Zhao H. 2013. Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing. *Biotechnol. Bioeng.* **110**, 1811–1821.
 47. Miller J.C., Tan S., Qiao G., Barlow K.A., Wang J., Xia D.F., Meng X., Paschon D.E., Leung E., Hinkley S.J., Dulay G.P., Hua K.L., Ankoudinova I., Cost G.J., Urnov F.D., Zhang H.S., Holmes M.C., Zhang L., Gregory P.D., Rebar E.J. 2011. A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–150.
 48. Bedell V.M., Wang Y., Campbell J.M., Poshusta T.L., Starker C.G., Krug R.G., Tan W., Penheiter S.G., Ma A.C., Leung A.Y.H., Fahrenkrug S.C., Carlson D.F., Voytas D.F., Clark K.J., Essner J.J., Ekker S.C. 2012. *In vivo* genome editing using a high-efficiency TALEN system. *Nature*. **491**, 114–118.
 49. Cong L., Zhou R., Kuo Y.-C., Cunniff M., Zhang F. 2012. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat. Commun.* **3**, 968.
 50. Streubel J., Blücher C., Landgraf A., Boch J. 2012. TAL effector RVD specificities and efficiencies. *Nat. Biotechnol.* **30**, 593–595.
 51. Schmid-Burgk J.L., Schmidt T., Kaiser V., Höning K., Hornung V. 2013. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat. Biotechnol.* **31**, 76–81.
 52. Ding Q., Lee Y.-K., Schaefer E.A.K., Peters D.T., Veres A., Kim K., Kuperwasser N., Motola D.L., Meissner T.B., Hendriks W.T., Trevisan M., Gupta R.M., Moisan A., Banks E., Friesen M., Schinzel R.T., Xia F., Tang A., Xia Y., Figueroa E., Wann A., Ahfeldt T., Daheron L., Zhang F., Rubin L.L., Peng L.F., Chung R.T., Musunuru K., Cowan C.A. 2013. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell*. **12**, 238–251.
 53. Sander J.D., Cade L., Khayter C., Reyon D., Peterson R.T., Joung J.K., Yeh J.-R.J. 2011. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat. Biotechnol.* **29**, 697–698.

54. Reyon D., Tsai S.Q., Khayter C., Foden J.A., Sander J.D., Joung J.K. 2012. FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.* **30**, 460–465.
55. Brugnans L., Kanaar R., Essers J. 2007. Analysis of DNA double-strand break repair pathways in mice. *Mutat. Res.* **614**, 95–108.
56. Orlando S.J., Santiago Y., DeKolver R.C., Freyvert Y., Boydston E.A., Moehle E.A., Choi V.M., Gopalan S.M., Lou J.F., Li J., Miller J.C., Holmes M.C., Gregory P.D., Urnov F.D., Cost G.J. 2010. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucl. Acids Res.* **38**, e152.
57. Hockemeyer D., Soldner F., Beard C., Gao Q., Mitalipova M., DeKolver R.C., Katibah G.E., Amora R., Boydston E.A., Zeitler B., Meng X., Miller J.C., Zhang L., Rebar E.J., Gregory P.D., Urnov F.D., Jaenisch R. 2009. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* **27**, 851–857.
58. Hockemeyer D., Wang H., Kiani S., Lai C.S., Gao Q., Cassady J.P., Cost G.J., Zhang L., Santiago Y., Miller J.C., Zeitler B., Cherone J.M., Meng X., Hinkley S.J., Rebar E.J., Gregory P.D., Urnov F.D., Jaenisch R. 2011. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* **29**, 731–734.
59. Torikai H., Reik A., Liu P.-Q., Zhou Y., Zhang L., Maiti S., Huls H., Miller J.C., Kebriaei P., Rabinovitch B., Lee D.A., Champlin R.E., Bonini C., Naldini L., Rebar E.J., Gregory P.D., Holmes M.C., Cooper L.J.N. 2012. A foundation for universal T-cell based immunotherapy: T cells engineered to express a CD19-specific chimeric-antigen-receptor and eliminate expression of endogenous TCR. *Blood*. **119**, 5697–5705.
60. Provasi E., Genovese P., Lombardo A., Magnani Z., Liu P.-Q., Reik A., Chu V., Paschon D.E., Zhang L., Kuball J., Camisa B., Bondanza A., Casorati G., Ponzoni M., Ciceri F., Bordignon C., Greenberg P.D., Holmes M.C., Gregory P.D., Naldini L., Bonini C. 2012. Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer. *Nat. Med.* **18**, 807–815.
61. Lee H.J., Kim E., Kim J.-S. 2010. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res.* **20**, 81–89.
62. Carlson D.F., Tan W., Lillico S.G., Stverakova D., Proudfoot C., Christian M. 2012. Efficient TALEN-mediated gene knockout in livestock. *Proc. Natl. Acad. Sci. USA*. **43**, 17382–17387.
63. Ma S., Zhang S., Wang F., Liu Y., Liu Y., Xu H., Liu C., Lin Y., Zhao P., Xia Q. 2012. Highly efficient and specific genome editing in silkworm using custom TALENs. *PLoS One*. **7**, e45035.
64. Maresca M., Lin V.G., Guo N., Yang Y. 2013. Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome Res.* **23**, 539–546.
65. Rémy S., Tesson L., Ménoret S., Usal C., Scharenberg A.M., Anegón I. 2010. Zinc-finger nucleases: a powerful tool for genetic engineering of animals. *Transgenic Res.* **19**, 363–371.
66. Biasco L., Baricordi C., Aiuti A. 2012. Retroviral integrations in gene therapy trials. *Mol. Ther.* **20**, 709–716.
67. Zou J., Sweeney C.L., Chou B., Choi U., Pan J., Wang H., Sarah N., Cheng L., Malech H.L., Dowey S.N. 2011. Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPSCs: functional correction by zinc finger nuclease-mediated safe harbor targeting. *Blood*. **117**, 5561–5572.
68. Wang H., Hu Y.-C., Markoulaki S., Welstead G.G., Cheng A.W., Shivalila C.S., Pyntikova T., Dadon D.B., Voytas D.F., Bogdanove A.J., Page D.C., Jaenisch R. 2013. TALEN-mediated editing of the mouse Y chromosome. *Nat. Biotechnol.* **31**, 530–532.
69. Zu Y., Tong X., Wang Z., Liu D., Pan R., Li Z., Hu Y., Luo Z., Huang P., Wu Q., Zhu Z., Zhang B., Lin S. 2013. TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat. Methods*. **10**, 329–331.
70. Li H., Haurigot V., Doyon Y., Li T., Wong S.Y., Bhagwat A.S., Malani N., Anguela X.M., Sharma R., Ivanciu L., Murphy S.L., Finn J.D., Khazi F.R., Zhou S., Paschon D.E., Rebar E.J., Bushman F.D., Gregory P.D., Holmes M.C., High K.A. 2011. *In vivo* genome editing restores haemostasis in a mouse model of haemophilia. *Nature*. **475**, 217–221.
71. Chen F., Pruett-Miller S.M., Huang Y., Gjoka M., Duda K., Taunton J., Collingwood T.N., Frodin M., Davis G.D. 2011. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods*. **8**, 753–755.
72. Perez-Pinera P., Ousterout D.G., Gersbach C.A. 2012. Advances in targeted genome editing. *Curr. Opin. Chem. Biol.* **16**, 268–277.
73. Soldner F., Laganière J., Cheng A.W., Hockemeyer D., Gao Q., Alagappan R., Khurana V., Golbe L.I., Myers R.H., Lindquist S., Zhang L., Guschin D., Fong L.K., Vu B.J., Meng X., Urnov F.D., Rebar E.J., Gregory P.D., Zhang H.S., Jaenisch R. 2011. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*. **146**, 318–331.
74. Yusa K., Rashid S.T., Strick-Marchand H., Varela I., Liu P.-Q., Paschon D.E., Miranda E., Ordóñez A., Hannan N.R.F., Rouhani F.J., Darche S., Alexander G., Marciniak S.J., Fusaki N., Hasegawa M., Holmes M.C., Di Santo J.P., Lomas D.A., Bradley A., Vallier L. 2011. Targeted gene correction of α 1-antitrypsin deficiency in induced pluripotent stem cells. *Nature*. **478**, 391–394.
75. Zou J., Maeder M.L., Mali P., Pruett-Miller S.M., Thibodeau-Beganny S., Chou B.-K., Chen G., Ye Z., Park I.-H., Daley G.Q., Porteus M.H., Joung J.K., Cheng L. 2009. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell*. **5**, 97–110.
76. Moore F.E., Reyon D., Sander J.D., Martinez S.A., Blackburn J.S., Khayter C., Ramirez C.L., Joung J.K.,

- Langenau D.M. 2012. Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). *PLoS One*. **7**, e37877.
77. Lamb B.M., Mercer A.C., Barbas C.F. 2013. Directed evolution of the TALE N-terminal domain for recognition of all 5' bases. *Nucl. Acids Res.* **41**, 9779–9785.
 78. Kim Y., Kweon J., Kim A., Chon J.K., Yoo J.Y., Kim H.J., Kim S., Lee C., Jeong E., Chung E., Kim D., Lee M.S., Go E.M., Song H.J., Kim H., Cho N., Bang D., Kim S., Kim J.-S. 2013. A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.* **31**, 251–258.
 79. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. **337**, 816–821.
 80. Fu Y., Foden J.A., Khayter C., Maeder M.L., Reyon D., Joung J.K., Sander J.D. 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826.
 81. Owens J.B., Mauro D., Stoytchev I., Bhakta M.S., Kim M.-S., Segal D.J., Moisyadi S. 2013. Transcription activator like effector (TALE)-directed piggyBac transposition in human cells. *Nucl. Acids Res.* **41**, 9197–9207.
 82. Mercer A.C., Gaj T., Fuller R.P., Barbas C.F. 2012. Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucl. Acids Res.* **40**, 11163–11172.
 83. Zhang F., Cong L., Lodato S., Kosuri S., Church G.M., Arlotta P. 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* **29**, 149–154.
 84. Bultmann S., Morbitzer R., Schmidt C.S., Thanisch K., Spada F., Elsaesser J., Lahaye T., Leonhardt H. 2012. Targeted transcriptional activation of silent *oct4* pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucl. Acids Res.* **40**, 5368–5377.
 85. Geissler R., Scholze H., Hahn S., Streubel J., Bonas U., Behrens S.-E., Boch J. 2011. Transcriptional activators of human genes with programmable DNA-specificity. *PLoS One*. **6**, e19509.
 86. Minczuk M., Papworth M.A., Kolasinska P., Murphy M.P., Klug A. 2006. Sequence-specific modification of mitochondrial DNA using a chimeric zinc finger methylase. *Proc. Natl. Acad. Sci. USA*. **103**, 19689–19694.
 87. Li F., Papworth M., Minczuk M., Rohde C., Zhang Y., Ragozin S., Jeltsch A. 2007. Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes. *Nucl. Acids Res.* **35**, 100–112.