

УДК 612.089:612.014:602.9:577.113.5:575.224

Системы редактирования геномов TALEN и CRISPR/Cas – инструменты открытий

А. А. Немудрый^{1,2,3,§}, К. Р. Валетдинова^{1,2,3,4,§}, С. П. Медведев^{1,2,3,4}, С. М. Закиян^{1,2,3,4*}

¹Институт цитологии и генетики СО РАН, 630090, Новосибирск, просп. Акад. Лаврентьева, 10

²Институт химической биологии и фундаментальной медицины СО РАН, 630090, Новосибирск, просп. Акад. Лаврентьева, 8

³Новосибирский научно-исследовательский институт патологии кровообращения им. акад. Е.Н. Мешалкина МЗ РФ, 630055, Новосибирск, ул. Речкуновская, 15

⁴Новосибирский государственный университет, 630090, Новосибирск, ул. Пирогова, 2

*E-mail: zakian@bionet.nsc.ru

§Эти авторы внесли одинаковый вклад в написание данного обзора.

Поступила в редакцию 03.03.2014

РЕФЕРАТ Детальное изучение геномов растений, животных и человека открывает широчайшие возможности применения полученных знаний в биотехнологии и медицине. Однако только данных о нуклеотидных последовательностях геномов недостаточно для понимания функциональных взаимосвязей отдельных элементов геномов и их роли в формировании фенотипических признаков и патогенезе отдельных заболеваний. В постгеномную эпоху активно развиваются методы, позволяющие манипулировать с ДНК в геномах, а также визуализировать и управлять экспрессией генов и работой регуляторных элементов. Тем не менее далеко не все методы отвечают высоким требованиям к их эффективности, безопасности и доступности для широкого круга исследователей. В последние несколько лет появились новейшие методы редактирования геномов – это системы TALEN (Transcription Activator-Like Effector Nucleases) и CRISPR (Clustered Regulatory Interspaced Short Palindromic Repeats)/Cas9. Эти появившиеся относительно недавно системы уже зарекомендовали себя как эффективные и надежные инструменты геномной инженерии. Данный обзор в основном посвящен применению указанных систем для редактирования геномов основных модельных объектов современной биологии, а также для функционального скрининга геномов, создания клеточных моделей наследственных заболеваний человека, изучения эпигеномов и визуализации клеточных процессов. Кроме того, рассмотрены основные методы конструирования подобных систем, проанализировано их действие, обсуждаются некоторые проблемы, которые ожидают исследователей при применении этих инструментов редактирования геномов.

КЛЮЧЕВЫЕ СЛОВА TALEN, CRISPR/Cas9, редактирование генома.

СПИСОК СОКРАЩЕНИЙ TALEN (Transcription Activator-Like Effector Nucleases); CRISPR (Clustered Regulatory Interspaced Short Palindromic Repeats)/Cas9; PAM – мотив, прилежащий к протоспейсеру (Protospacer Adjacent Motive); sgRNA (single guide RNA); crRNA (CRISPR RNA) – CRISPR РНК; tracrRNA – транскрибирующая CRISPR РНК; SpCas9 – Cas9 *Streptococcus pyogenes*; pre-crRNA (poly-spacer precursor crRNA) – полиспейсерный предшественник CRISPR РНК.

ВВЕДЕНИЕ

Генетическая инженерия возникла в 1972 году в лаборатории Пола Берга как технология рекомбинантных ДНК, когда ученые объединили геном кишечной палочки с генами бактериофага и вируса SV40. С того времени эта наука достигла колоссальных успехов, были открыты и детально изучены молекулярно-генетические механизмы и явления, которые теперь можно воспроизводить *in vitro*. Исследования в об-

ласти молекулярной генетики и биохимии бактерий и вирусов позволили разработать методы манипуляции с ДНК, создать различные векторные системы и способы их доставки в клетку. Все это позволило получить не только трансгенные микроорганизмы, но и генетически модифицированные растения и животные. Бурное развитие получила прикладная область генной инженерии, дав толчок прогрессу в селекции и биотехнологии. Однако традиционная

стратегия генной инженерии имеет ряд недостатков и ограничений, одно из которых – сложность манипуляции с большими геномами животных и человека.

С 1990 по 2003 год в рамках Международного проекта «Геном человека» была определена последовательность нуклеотидов ядерной ДНК человека и идентифицировано около 20.5 тысяч генов. Аналогичные проекты осуществляются и в настоящее время, расшифрованы нуклеотидные последовательности геномов всех основных модельных биологических объектов – кишечной палочки, нематоды, дрозофилы, мыши и др. Однако эти проекты позволяют получить данные лишь о последовательности нуклеотидов в ДНК, но не дают какой-либо информации о том, как функционируют отдельные элементы генома, как они связаны между собой в целостную систему. Понимание функциональных взаимоотношений в геноме человека позволит не только выявить причинно-следственные связи в патологии наследственных, а также многофакторных заболеваний, но и найти мишени для их терапии.

В 2003 году Национальный исследовательский институт генома человека организовал новый международный проект ENCODE (Encyclopedia Of DNA Elements), целью которого стало объединение усилий ученых с целью получения полного списка функциональных элементов генома человека, включая элементы, которые действуют на уровне белков и РНК, а также регуляторные элементы, контролирующие фундаментальные генетические процессы (транскрипцию, трансляцию и репликацию). Для установления подобных функциональных взаимоотношений используют две стратегии: выключение гена (нокаут или нокдаун) и усиление работы гена либо его эктопическую экспрессию. Традиционные методы – трансгенез с использованием гомологичной рекомбинации у мышей [1], а также применение различных вирусных, в том числе лентивирусных векторов, не только дороги и довольно трудоемки, они не позволяют внести точные изменения в строго определенный локус генома.

В настоящий момент в распоряжении ученых появилось несколько инструментов, которые позволяют решать задачи высокоточного редактирования генома растений, животных и человека.

Еще в 1996 году было впервые показано, что белковый домен типа «цинковые пальцы», соединенный с FokI-эндонуклеазным доменом, действует как сайт-специфическая нуклеаза, разрезая ДНК *in vitro* в строго определенных участках [2]. Такой химерный белок имеет модульную структуру, поскольку каждый домен «цинковые пальцы» узнает один триплет нуклеотидов (Zinc-finger Nuclease, ZFN). Этот метод стал основой редактирования культиви-

руемых клеток, включая плюрипотентные стволовые клетки, и модельных животных и растений [3–8]. Однако технология, основанная на ZFN, имеет ряд недостатков, включая сложность и высокую стоимость конструирования белковых доменов для каждого конкретного локуса генома, вероятность неточного разрезания ДНК-мишени по причине одонуклеотидных замен или неправильного взаимодействия между доменами. Поэтому продолжались активные поиски новых методов редактирования генома. В последние годы эти поиски привели к созданию новых инструментов редактирования геномов – системы TALEN (Transcription Activator-Like Effector Nucleases; эффекторные нуклеазы, подобные активаторам транскрипции) и CRISPR/Cas (Clustered Regulatory Interspaced Short Palindromic Repeats; короткие палиндромные повторы, расположенные группами, равномерно удаленными друг от друга). Эти системы отличаются относительной простотой конструирования и высокой эффективностью работы в клетках человека, животных и растений. Такие системы, активно применяемые для различных манипуляций с геномами, позволяют решать сложные задачи, включая получение мутантных и трансгенных растений и животных, создание и исследование моделей заболеваний на основе культивируемых плюрипотентных клеток человека. Кроме того, химерные белки на основе ДНК-связывающих доменов TALE и инактивированной нуклеазы Cas9 использовали в экспериментах по регуляции транскрипции генов, для изучения эпигеномов и поведения хромосомных локусов в клеточном цикле.

В представленном обзоре подробно описаны возможности конструирования, применения и анализа действия TALEN и CRISPR/Cas9 на примере различных модельных систем, а также сложности и проблемы, связанные с использованием данных инструментов редактирования геномов.

ЕСТЕСТВЕННЫЕ БАКТЕРИАЛЬНЫЕ СИСТЕМЫ TALE И CRISPR/Cas КАК ОСНОВА ДЛЯ СОЗДАНИЯ НОВЫХ ИНСТРУМЕНТОВ РЕДАКТИРОВАНИЯ ГЕНОМОВ ЭУКАРИОТ

TALEN

В 2011 году методы высокоточного редактирования геномов, среди которых была и система TALEN, названы журналом *Nature Methods* методом года [9]. История разработки этой системы связана с изучением бактерий рода *Xanthomonas*. Эти бактерии являются патогенами таких культурных растений, как рис, перец, томат, они наносят экономически значимый вред сельскому хозяйству, что стало причиной их тщательного изучения. Оказалось, что бакте-

рии секретируют в цитоплазму растительных клеток эффекторный белок (Transcription Activator-Like Effectors, TALE), которые влияют на процессы в растительной клетке и увеличивают ее восприимчивость к патогену. При дальнейшем изучении механизмов действия эффекторных белков было обнаружено, что они способны связываться с ДНК и активировать экспрессию своих генов-мишеней, имитируя факторы транскрипции эукариот.

Белки TALE состоят из центрального домена, ответственного за связывание ДНК, сигнала ядерной локализации и домена, активирующего транскрипцию целевого гена [10]. Впервые способность этих белков связываться с ДНК была описана в 2007 году [11], а всего через год двумя группами ученых был расшифрован код узнавания целевой последовательности ДНК белками TALE [12, 13]. Показано, что ДНК-связывающий домен состоит из мономеров, каждый из которых связывается с одним нуклеотидом в целевой нуклеотидной последовательности. Мономеры представляют собой тандемные повторы из 34 аминокислотных остатков, из которых два, расположенные в позициях 12 и 13, высоко вариабельные (Repeat Variable Diresidue, RVD), и именно они отвечают за узнавание определенного нуклеотида. Этот код является вырожденным, некоторые RVD могут связываться с несколькими нуклеотидами с разной эффективностью. При этом перед 5'-концом последовательности, с которой связываются мономеры TALE, в целевой молекуле ДНК всегда находится один и тот же нуклеотид – тимидин, влияющий на эффективность связывания [14]. Последний тандемный повтор, связывающий нуклеотид на 3'-конце сайта узнавания, состоит из 20 аминокислотных остатков, поэтому он назван полуповтором.

После расшифровки кода узнавания ДНК белками TALE, который привлек внимание ученых всего мира благодаря своей простоте (один мономер – один нуклеотид), были проведены первые работы по созданию химерных нуклеаз TALENs. С этой целью последовательность, кодирующая ДНК-связывающий домен TALE, встроили в плазмидный вектор, использованный ранее при создании ZFN [15]. В результате были получены генетические конструкции, экспрессирующие искусственные химерные нуклеазы, которые содержат ДНК-связывающий домен и каталитический домен эндонуклеазы рестрикции FokI. Данная система позволяет, комбинируя мономеры ДНК-связывающего домена с разными RVD, создать искусственные нуклеазы, мишенью которых будет любая нуклеотидная последовательность. В большинстве работ используют мономеры, содержащие RVD Asn и Ile (NI), Asn и Gly (NG), два Asn (NN), His и Asp (HD), для связывания нуклеотидов А, Т, G, С соответ-

ственно. При этом RVD NN может связывать как G, так и A. В ряде работ показано, что использование мономеров NH или NK для более специфичного связывания гуанина снижает риск нецелевых эффектов [16, 17]. Показано, что первый аминокислотный остаток в RVD (H или N) не участвует непосредственно в связывании нуклеотида, а отвечает за стабилизацию пространственной конформации. Второй аминокислотный остаток взаимодействует с нуклеотидом, при этом природа этого взаимодействия различна: D и N образуют водородные связи с азотистыми основаниями, а I и G связываются с целевым нуклеотидом за счет сил Ван-дер-Ваальса [18].

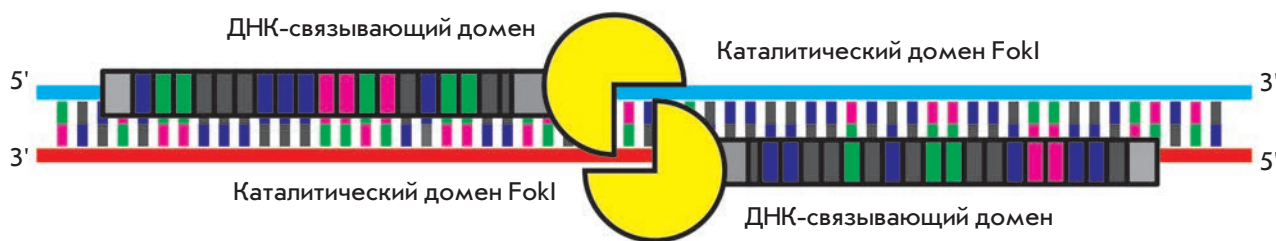
Искусственный ДНК-связывающий домен встраивают в генетическую конструкцию, содержащую сигнал ядерной локализации, полуповтор, N-концевой домен и каталитический домен FokI. Целевые сайты для искусственных нуклеаз выбирают таким образом, чтобы они находились на разных цепях ДНК и были разделены небольшим участком (12–25 п.н.) – спейсерной последовательностью. При попадании в ядро искусственные нуклеазы связываются с целевыми сайтами, в результате чего расположенные на C-концах химерного белка домены FokI димеризуются и вносят двухцепочечный разрыв в спейсерную последовательность (рис. 1).

Теоретически с помощью искусственных нуклеаз TALENs двухцепочечный разрыв можно внести в любой участок генома, с известными сайтами узнавания ДНК-связывающих доменов. Единственное ограничение по выбору сайтов нуклеаз TALEN заключается в необходимости присутствия T перед 5'-концом целевой последовательности. Тем не менее, варьируя длину спейсерной последовательности, в подавляющем большинстве случаев можно осуществить выбор сайтов. Показано, что остаток W232 в составе N-концевой области ДНК-связывающего домена взаимодействует с 5'-T, влияя на эффективность связывания TALEN с целевым сайтом [19]. Однако и это ограничение удается преодолеть в результате селекции мутантных вариантов N-концевого домена TALEN, способных связываться с A, G или C [14].

CRISPR/Cas

Примерно через два года после открытия системы химерных белков TALEN получила развитие и стала активно применяться другая система редактирования геномов – CRISPR, элементами которой являются некодирующие РНК и белки Cas (CRISPR-associated). В отличие от химерных белков TALENs, узнавание системой CRISPR/Cas осуществляется за счет комплементарного взаимодействия между некодирующей РНК и ДНК целевых сайтов. При этом образуется комплекс из некодирующих РНК и бел-

Целевой локус генома



Пара химерных белков TALEN

Код узнавания белковыми доменами нуклеотидов

NI = A

NG = T

NN = G

HD = C

Рис. 1. Схема внесения двуникового разрыва с помощью химерных белков TALEN. Один мономер ДНК-связывающего домена белка узнает один нуклеотид целевой последовательности ДНК. За связывание отвечают два аминокислотных остатка в мономере, приведен код узнавания (аминокислотные остатки в однобуквенных обозначениях). Сайты узнавания расположены на разных цепях ДНК на расстоянии, достаточном для димеризации каталитических доменов FokI. В виде димера FokI вносит двуникового разрыв в ДНК

ков Cas, которые обладают нуклеазной активностью. Еще в 1987 году в некоторых бактериальных генах были обнаружены таинственные повторы [20], функции которых в течение почти 20 лет оставались неизвестными. Секвенирование бактериальных геномов позволило обнаружить в геноме многих микроорганизмов аналогичные нуклеотидные последовательности, обладающие характерной структурой: короткие участки уникальной ДНК – спейсеры – отделены друг от друга короткими палиндромными повторами (рис. 2). Благодаря этой особенности они и получили свое название – CRISPR. Кроме того, такие CRISPR-кассеты находятся в непосредственной близости от *cas*-генов (CRISPR-associated – ассоциированные с CRISPR), белковые продукты которых обладают хеликазной и нуклеазной активностью [21]. В 2005 году три независимые группы биоинформатиков сообщили о том, что спейсерная ДНК зачастую гомологична ДНК многих фагов и плазмид [22–24]. А в 2007 году было показано, что клетки *Streptococcus thermophilus*, несущие в локусе CRISPR спейсер, комплементарный участку геномной ДНК бактериофага, становятся устойчивыми к этому фагу [25]. Таким образом, стало очевидным, что CRISPR/Cas-система – это уникальный механизм, обеспечивающий защиту микроорганизмов от проникновения чужеродной ДНК и действующий, наряду с системой рестрикции-модификации, как ограничитель горизонтального переноса генетической информации.

CRISPR-системы широко распространены среди прокариот: они обнаружены у 87% архей и 48% эубактерий [26]. Поэтому у разных видов широко варьирует как количество самих CRISPR-кассет в геноме (1–18), так и количество (в среднем 60) и величина повторов (в среднем 23–37 п.н.), а также число и размер спейсеров (17–84 п.н.). При этом длина повторов и спейсеров внутри одной кассеты неизменна, а последовательности повторов практически идентичны [27].

Механизм защиты включает три основных стадии (рис. 2). На первой стадии – адаптации – небольшой фрагмент чужеродной ДНК, проникшей в бактериальную клетку, встраивается в CRISPR-локус генома хозяина, формируя новый спейсер. В вирусном геноме этот фрагмент присутствует в качестве протоспейсера, комплементарного спейсеру и фланкированного короткой (2–5 п.н.) консервативной последовательностью, называемой PAM (Protospacer Adjacent Motif; мотив, прилежащий к протоспейсеру) [28, 29]. Новый спейсер всегда встраивается со стороны АТ-богатой лидерной последовательности, находящейся перед CRISPR-кассетой, в ней же находятся промоторные элементы и сайты посадки регуляторных белков [30, 31]. По всей видимости, именно таким образом формируются мишени большинства CRISPR/Cas-систем.

На второй стадии – транскрипции – весь CRISPR-локус транскрибируется в длинную pre-crRNA

(poly-spacer precursor crRNA; полиспейсерный предшественник CRISPR РНК) (рис. 2). Процессинг незрелого транскрипта в зрелые crRNA в большинстве CRISPR/Cas-систем осуществляет эндонуклеаза Cas6 [32–36]. Короткие crRNA (CRISPR РНК) длиной 39–45 нуклеотидов содержат одну спейсерную последовательность, а на концах находятся повторы, которые участвуют в образовании структуры стержень–петля: восемь последних нуклеотидов повтора с гидроксильной группой на 5'-конце образуют стержень, а шпильчатая структура с 2',3'-циклическим фосфатом формирует петлю на 3'-конце [37, 38].

Третья стадия – интерференция чужеродной ДНК или РНК – обеспечивается за счет взаимодействия crRNA и комплекса Cas-белков; crRNA комплементарно узнает последовательность протоспейсера, а Cas-белки обеспечивают ее разрушение (рис. 2).

Для деградации ДНК-мишени эффекторным комплексом необходимо, чтобы не возникали комплементарные взаимодействия между нуклеотидами crRNA и ДНК-мишени в позициях –2, –3, –4 (если за +1 принять первое основание протоспейсера) [39]. По-видимому, комплементарные взаимодействия crRNA и ДНК-мишени по этим позициям нарушают образование эффекторного комплекса, что предотвращает разрезание геномной ДНК и ее последующую деградацию.

Длительная коэволюция вирусов и их хозяев привела к появлению у вирусов защитных механизмов против CRISPR-интерференции [40], что объясняет большое разнообразие CRISPR/Cas-систем у бактерий и архей. Биоинформатические исследования подразделяют все CRISPR/Cas-системы на три основных типа (I–III) и, по меньшей мере, на 10 подтипов [21, 27, 41]. Среди них в геномной инженерии в настоящее время наиболее активно используется CRISPR/Cas-система типа II-A, выделенная из патогена *S. pyogenes*. У этой бактерии обнаружен минимальный набор генов *cas* [27, 41]. Один полифункциональный белок Cas9 осуществляет и процессинг pre-crRNA, и интерференцию чужеродной ДНК [42]. Процессинг crRNA зависит также от малой некодирующей РНК – tracrRNA (trans-activating crRNA; трансактивирующая crРНК). Молекулы tracrRNA комплементарно связываются с последовательностями повторов в pre-crRNA, формируя дуплекс, а одна из рибонуклеаз клетки-хозяина – РНКаза III, в присутствии Cas9 разрезает дуплекс с образованием зрелой crRNA, которая содержит 20-нуклеотидную спейсерную последовательность на 5'-конце. Cas9 вносит двухцепочечный разрыв в целевой locus в присутствии ионов Mg^{2+} , причем нуклеазный домен HNH этого фермента разрезает нить ДНК,

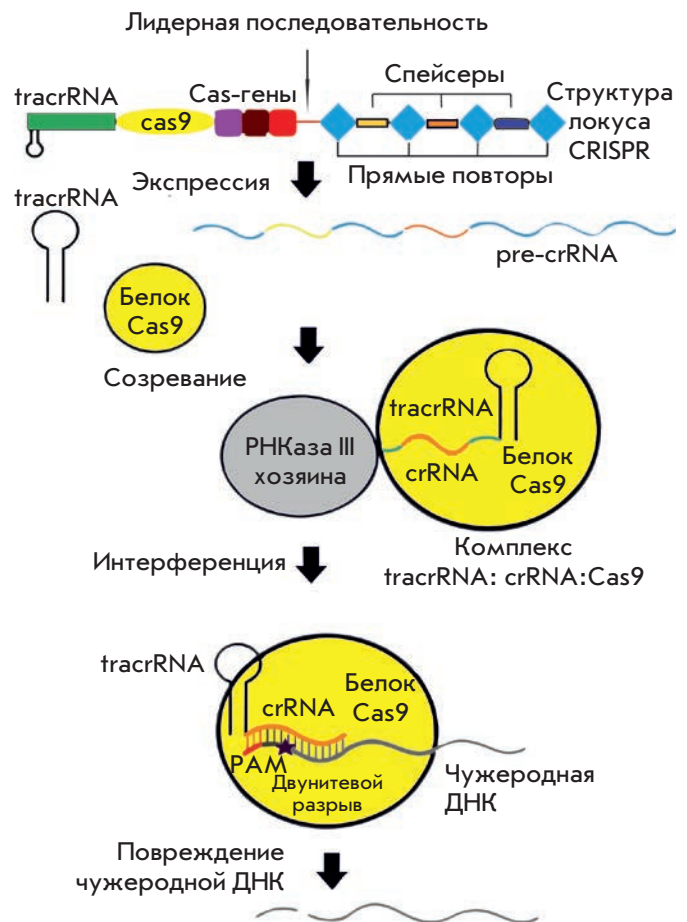


Рис. 2. Механизм действия CRISPR/Cas9 в клетках бактерий (описание в тексте статьи)

комплементарную crRNA, а RuvC-домен режет не-комплементарную нить [43]. ДНК-мишень для Cas9 *S. pyogenes* должна непременно содержать 5'-NGG-3' PAM [43, 44], через три нуклеотида от которого и происходит разрезание. У *S. thermophilus* и *Neisseria meningitidis* мишени для Cas9 типа II имеют другой консенсус – 5'-NGGNG-3' и 5'-NNNNGATT-3' соответственно.

ГЕНОМНАЯ ИНЖЕНЕРИЯ С ИСПОЛЬЗОВАНИЕМ TALENs и CRISPR/Cas9

Общая стратегия геномной инженерии с помощью сайт-специфических нуклеаз включает четыре основных этапа (рис. 3):

- Выбор целевой нуклеотидной последовательности в геноме.
- Создание нуклеазной конструкции, направленной на выбранную мишень.
- Доставка этой конструкции в клеточное ядро.
- Анализ полученных мутаций.

Биоинформатический анализ

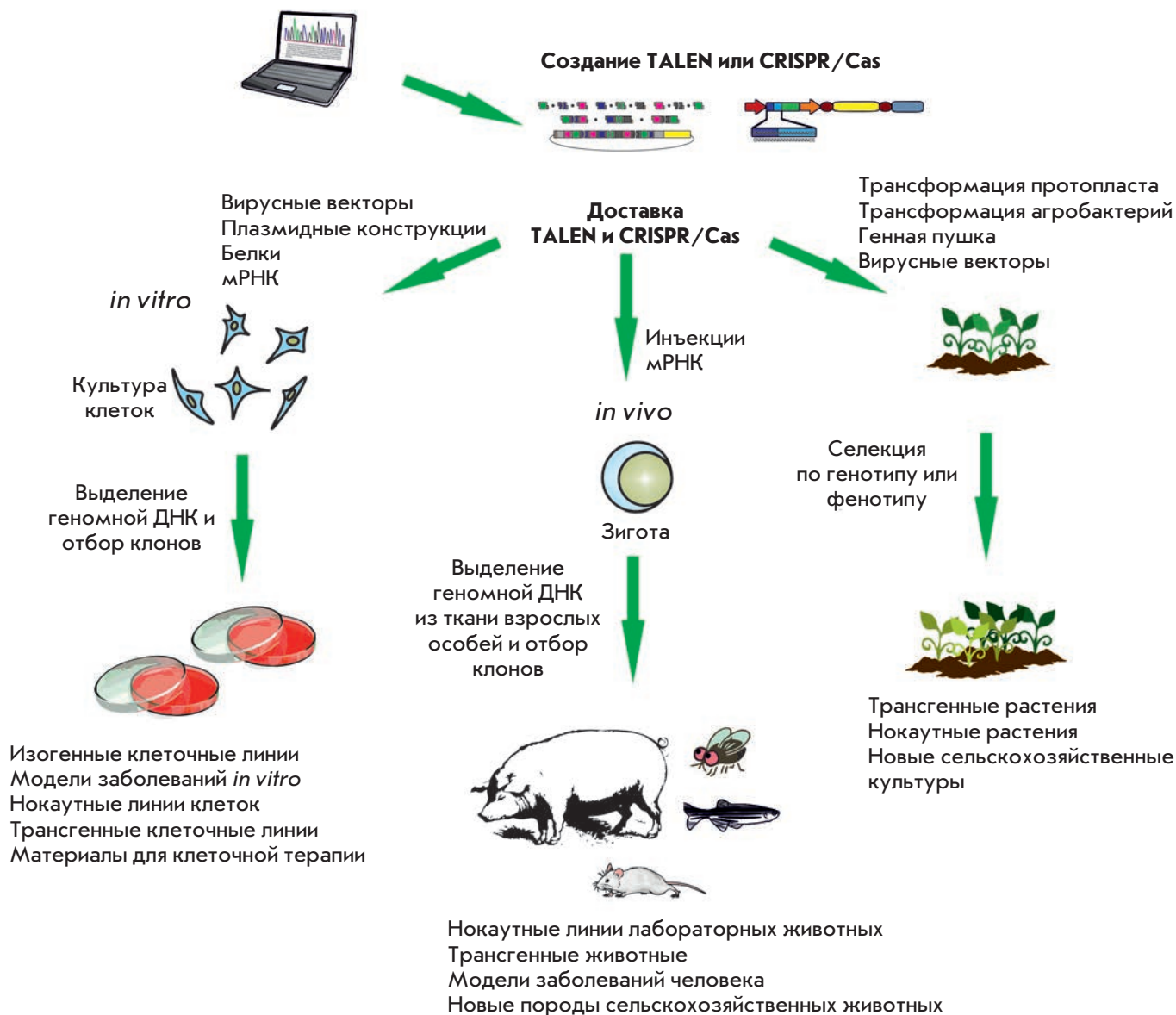


Рис. 3. Общая схема стратегии применения систем TALEN и CRISPR/Cas в геномной инженерии

Выбор целевой нуклеотидной последовательности в геноме

Важным моментом при работе с системами TALENs и CRISPR/Cas9 является тщательный подбор сайтов для специфичного внесения двухцепочечного разрыва. Необходимость предварительного биоинформатического анализа объясняется возможностью нецелевых эффектов – внесения неспецифичных двухцепочечных разрывов в геном. При выборе нужных сайтов следует избегать участков повторенных последовательностей, а также участков,

имеющих высокую гомологию с другими районами генома.

Нецелевые эффекты при использовании системы химерных белков TALENs возникают по нескольким причинам. Во-первых, это различия в эффективности связывания RVD и специфических нуклеотидов. Мономеры HD и NN образуют сильные водородные связи с нуклеотидами, в то время как NG и NI – слабые. Это обуславливает возможное связывание ДНК-узнающего домена с сайтами, отличающимися на несколько нуклеотидов от целевых. Во-вторых, из-за

вырожденности кода связывания нуклеотидов мономерами возможно, например, взаимодействие NG и A. В-третьих, возможна димеризация доменов FokI двух нуклеаз с одинаковыми ДНК-связывающими доменами (образование гомодимеров). Этот вопрос решен в ряде работ благодаря получению TALENs, которые содержат домены FokI, работающие в качестве облигатных гетеродимеров. И наконец, возможные нецелевые эффекты могут возникать в результате того, что размер спейсерной ДНК между сайтами узнавания нуклеаз не фиксирован. Это свойство делает возможным внесение двухцепочечных разрывов при связывании нуклеаз с нецелевыми сайтами, расположенными на расстоянии, достаточном для димеризации доменов FokI [45].

Поскольку нуклеаза Cas9 *S. pyogenes* нуждается в обязательном присутствии PAM с консенсусом 5'-NGG-3', то это хоть и не значительно, но ограничивает выбор мишени. В частности, в геноме человека целевые сайты расположены через каждые 8–12 п.н. [46, 47]. Один из основных недостатков системы CRISPR/Cas9 – сравнительно высокая вероятность возникновения нецелевых мутаций. В ряде работ, выполненных *in vitro* [43], на бактериях [48] и на клетках человека [46], показано, что некоторые однонуклеотидные замены в 20-нуклеотидном спейсерном участке sgRNA (single guide RNA) могут привести к существенному снижению активности CRISPR/Cas9, особенно если эти замены локализованы в последних 10–12 нуклеотидах 3'-конца данного района sgRNA [49]. В то же время замены на 5'-конце sgRNA практически не влияют на функционирование системы [43, 46, 48]. Однако известны случаи, когда некоторые одно- и динуклеотидные замены на 3'-конце sgRNA не влияют на активность системы CRISPR/Cas9 и, наоборот, препятствуют ее действию, если находятся на 5'-конце [49]. В целом же, нецелевой эффект определяется положением замен – 8–12 п.н. на 3'-конце направляющей последовательности менее важны для Cas9, чем нуклеотиды 5'-конца; количеством замен, которых не должно быть больше трех; особенностями самого сайта-мишени; концентрацией вводимых Cas9 и sgRNA [46–49]. Преодолеть указанные недостатки позволит поиск и разработка методов, основанных на использовании ортологов Cas9, для активности которых необходимы PAM, имеющие более сложную консенсусную последовательность. Так, например, CRISPR/Cas типа II *N. meningitidis* распознает PAM с консенсусом 5'-NNNNGATT-3', что, безусловно, ограничивает возможности выбора мишени, однако может повысить специфичность.

С целью повышения специфичности редактирования генома с помощью CRISPR/Cas-системы также используют две никазы Cas9 с парой sgRNA

[50, 51] по аналогии с парами ZFNs и TALENs, которые вносят разрывы в ДНК только при действии двух независимых белков с доменами FokI. Мутации в одном из каталитически активных доменов (D10A в HNH и H840A в RuvC) превращают нуклеазу Cas9 в ДНК-никазу [43, 46, 52]. Если разрезание обеих цепей ДНК парой никаза Cas9 приводит к образованию сайт-специфических двухцепочечных разрывов, которые репарируются с помощью негомологичного сшивания концов ДНК (NHEJ – non-homologous end joining), то отдельные одноцепочечные повреждения репарируются преимущественно посредством высокоточной эксцизии оснований (BER – base excision repair) [53]. Показано, что применение двух никазов Cas9 с парой sgRNA обеспечивает значительное снижение образования нецелевых мутаций, при этом выход целевых мутаций в целом соответствует выходу при использовании нуклеазы [50, 51].

Перечисленные свойства узнавания целевых сайтов системами CRISPR/Cas9 и TALENs были учтены при разработке компьютерных алгоритмов поиска таких сайтов. В настоящее время доступно on-line программное обеспечение, разработанное разными коллективами и предназначенное для выбора потенциальных сайтов систем TALENs [54–59] и CRISPR/Cas9 [47, 60–62], а также определения возможных нецелевых эффектов.

Создание генетических конструкций, экспрессирующих CRISPR и TALEN

TALEN. ДНК-связывающий домен состоит из практически идентичных повторов, поэтому при создании генетических конструкций, экспрессирующих TALENs, существуют определенные трудности технического характера. Предложен ряд методов, которые позволяют создавать ДНК-связывающие домены TALE, состоящие из 20–30 и даже болееномеров. Одна из стратегий основана на стандартном клонировании ДНК с использованием гидролиза ДНК эндонуклеазами рестрикции типа II и лигирования – REAL (REstriction and Ligation [63]). На первом этапе подготавливают библиотеку мономеров, в которые с 5'- и 3'-концов внесены сайты эндонуклеаз рестрикции. После гидролиза ДНК проводят попарное лигирование, в результате чего образуются димеры (N_1N_2 , N_3N_4 , $N_{2k-1}N_{2k}$), которые затем объединяют в тетрамеры и так далее. Правильная последовательность при этом достигается использованием различных эндонуклеаз рестрикции. Эта методика весьма трудна и длительна, поскольку на каждом этапе необходимо очищать продукты реакции, а также подтверждать правильность ориентации. Для ускорения этого процесса создана би-

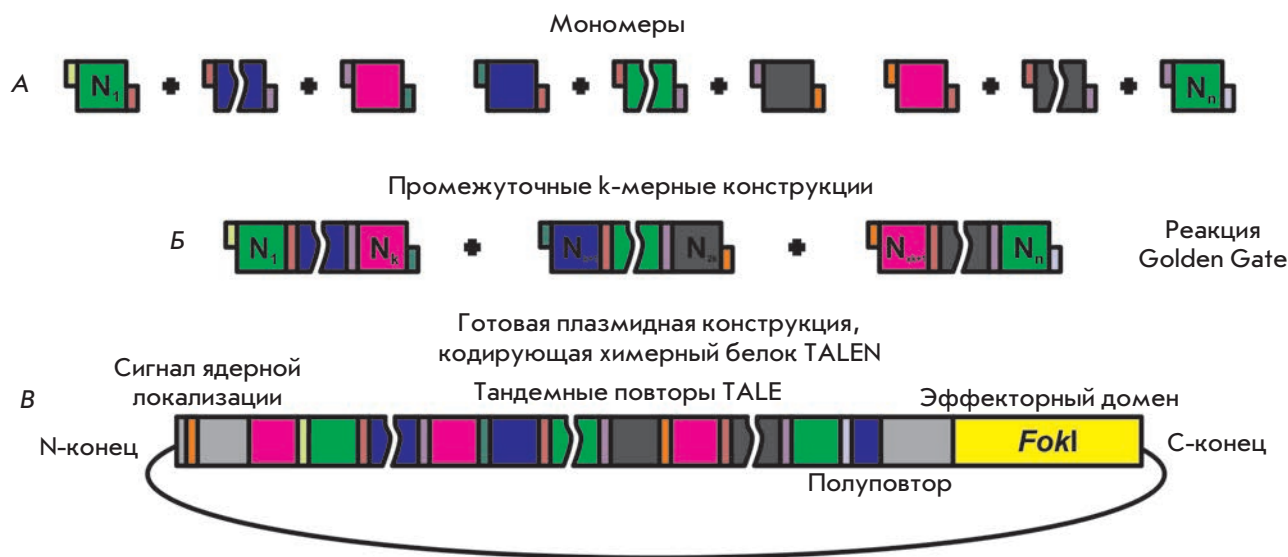


Рис. 4. Схема стратегии модульного иерархического лигирования на основе системы клонирования Golden Gate для создания генетических конструкций, экспрессирующих химерные белки TALEN. А – на первом этапе создают библиотеку мономеров, которая представляет собой своеобразный «конструктор», состоящий из набора деталей. Такие детали – амплифицированные последовательности мономеров с использованием специфических олигонуклеотидных праймеров. Праймеры подбирают таким образом, чтобы в результате гидролиза эндонуклеазами рестрикции типа IIS образовывались липкие концы, которые определяют позицию мономера в готовой конструкции. Б – в одной реакции Golden Gate возможно одновременное лигирование нескольких мономеров, в результате чего получают промежуточные k-мерные конструкции. В – на последнем этапе проводят реакцию Golden Gate, в результате которой происходит рестрикция и лигирование нескольких промежуточных k-мерных конструкций и плазмиды «основы», содержащей остальные элементы TALEN

блиотека из 376 элементов, состоящая из моно-, ди-, три- и тетрамеров (REAL-Fast, [64]).

Для увеличения эффективности и ускорения процесса сборки используют реакцию Golden Gate [65, 66], которая представляет собой одновременное лигирование и гидролиз эндонуклеазами рестрикции в одной реакционной смеси (рис. 4). В реакции Golden Gate используют эндонуклеазы рестрикции типа IIS, которые гидролизуют ДНК на фиксированном расстоянии от сайта узнавания, например, BsmBI или BsaI. Таким образом при лигировании происходит «бесшовная» сборка, поскольку эндонуклеазы рестрикции «отрезают» свой собственный сайт от мономера, и продукт лигирования уже не подвергается рестрикции. Амплифицируя последовательности мономеров (NI, HD, NG и NN) с различными олигонуклеотидными праймерами, создают библиотеку, которая содержит различные варианты всех четырех мономеров, соответствующие различным позициям (например, от 1 до 20) в будущем ДНК-связывающем домене. При обработке таких мономеров эндонуклеазами рестрикции IIS образуются липкие концы, комплементарные липким концам соседних мономеров.

В одной реакции можно лигировать сразу несколько мономеров, например, четыре [67] или шесть [68] одновременно. Далее, опять используя реакцию Golden Gate, можно лигировать несколько тетра- или гексамеров и клонировать уже полную последовательность в плазмидный вектор, содержащий 3'-полуповтор и каталитический домен FokI.

С целью сокращения времени создания генетических конструкций, экспрессирующих TALEN, предложен метод, позволяющий избежать лигирования ДНК и соответственно этапов, связанных с проверкой его результатов. Выбранный ДНК-связывающий домен собирается из мономеров с длинными специфическими одноцепочечными концами (10–30 нуклеотидов). При смешивании нескольких мономеров происходит отжиг комплементарных одноцепочечных концов, в результате чего мономеры выстраиваются в нужной последовательности. Затем клетки *E. coli* трансформируют полученную смесь, и лигирование происходит уже в бактериях с участием их собственных ферментов [69].

Перечисленные методы создания генетических конструкций, экспрессирующих TALENs,

относительно просты и по разным оценкам на их выполнение требуется 1–2 недели при наличии соответствующих реактивов. Помимо простоты и эффективности, данная технология также отличается доступностью – на сегодняшний день через депозитарий Addgene (<http://www.addgene.org/TALEN/>) возможно приобрести наборы для создания TALENs, разработанные разными коллективами авторов [64, 68–71], и использовать их в лабораторных условиях.

Также появились системы для автоматизированного высокоэффективного производства конструкций, экспрессирующих нуклеазы TALENs. Так, например, коммерческая платформа компании Collectis Bioresearch позволяет создавать до 7200 таких конструкций в год. В научной литературе описаны три метода [72–74], основанные на использовании твердофазных поверхностей. Эти методы позволяют избежать анализа промежуточных конструкций, их очистки с помощью выделения из геля и других стадий, что делает эти методы пригодными для автоматизированного производства и ускоряет процесс. Суть этих методов заключается в использовании покрытых стрептавидином магнитных частиц, к которым присоединены биотинилированные двухцепочечные ДНК-адаптеры. В результате последовательного чередования этапов гидролиза ДНК эндонуклеазами рестрикции и лигирования наращивается последовательность мономеров, связанная через адаптер с магнитной частицей. Продукты реакций очищают с помощью отмывочных буферов на магнитной подложке. При этом побочные продукты и компоненты реакций вымываются, а целевой продукт, благодаря притяжению между магнитными частицами и подложкой, остается в пробирке (или лунке). В конце с помощью эндонуклеаз рестрикции расщепляют связи между биотинилированным адаптером и синтезированной последовательностью мономеров ДНК-связывающего домена TALEN, которую затем в результате лигирования ДНК клонируют в плазмидный вектор. С помощью такого метода можно быстро и эффективно параллельно синтезировать генетические конструкции в 96-луночных планшетах, используя мультиканальные автоматические пипетки или роботизированные раскапывающие станции.

CRISPR/Cas9. Показано, что для разрезания ДНК *in vitro* [43, 52] и в бактериальных клетках [42] с использованием CRISPR/Cas9 необходимы и достаточны следующие компоненты: некодирующие РНК *tracrRNA* и *pre-crRNA*, РНКазы III и белок Cas9. Использование этой системы в клетках млекопитающих имеет ряд особенностей.

Во-первых, нуклеаза SpCas9 (Cas9 *S. pyogenes*) должна быть адаптированной для адекватной транс-

крипции в клетках высших эукариот, в частности, оптимизированной по кодонам, а для обеспечения ядерной компарментализации необходимо присоединить сигналы ядерной локализации (NLS – nuclear localization signal); двух NLS достаточно для эффективного направления Cas9 в ядро [46].

Во-вторых, для созревания *pre-crRNA* в клетках эукариот не требуется вносить экзогенную РНКазу III, так как с этой функцией успешно справляются собственные клеточные РНКазы [75–77].

В-третьих, вместо двух некодирующих РНК чаще вводят единую химерную *sgRNA*, в которой зрелая *crRNA* соединена с частью *tracrRNA* с помощью синтетической структуры «петля-стебель» для имитации естественного дуплекса *crRNA-tracrRNA* [43] (рис. 5). Для транскрипции *sgRNA* необходим соответствующий промотор, например, U6-промотор РНК-полимеразы III.

В лаборатории Фенга Занга созданы базовые плазмидные конструкции, содержащие элементы, необходимые для работы CRISPR/Cas9. Плазмиды pX260/pX334 содержат три экспрессионные кассеты: Cas9-нуклеаза/никаза, CRISPR РНК-матрица и *tracrRNA* (рис. 6). Чтобы изменить последовательность-мишень, из такой конструкции необходимо всего

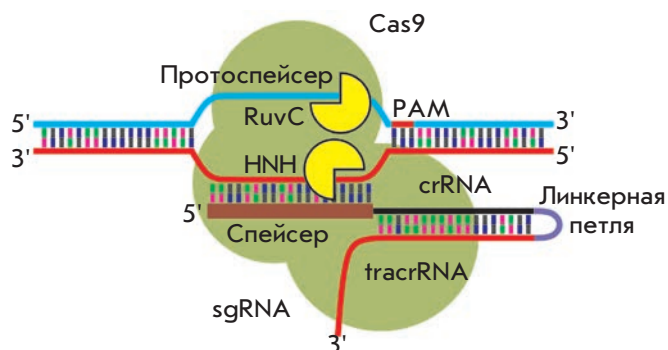


Рис. 5. Единая химерная *sgRNA* для внесения двухцепочечных разрывов в целевых локусах. Комплекс *sgRNA* и Cas9 способен вносить двухцепочечные разрывы в выбранных сайтах ДНК. *SgRNA* – искусственно созданная конструкция, представляющая собой объединенные в одну молекулу РНК элементы системы CRISPR/Cas9: *crRNA* и *tracrRNA*. Протоспейсер – сайт, который узнает система CRISPR/Cas9. Спейсер – последовательность в составе *sgRNA*, которая отвечает за связывание целевого сайта по принципу комплементарного взаимодействия. RuvC и HNH – каталитические домены, которые вносят разрывы в цепи ДНК в целевом сайте. PAM – короткий мотив (NGG в случае CRISPR/Cas9), наличие которого с 3'-конца протоспейсера обязательно для внесения разрыва

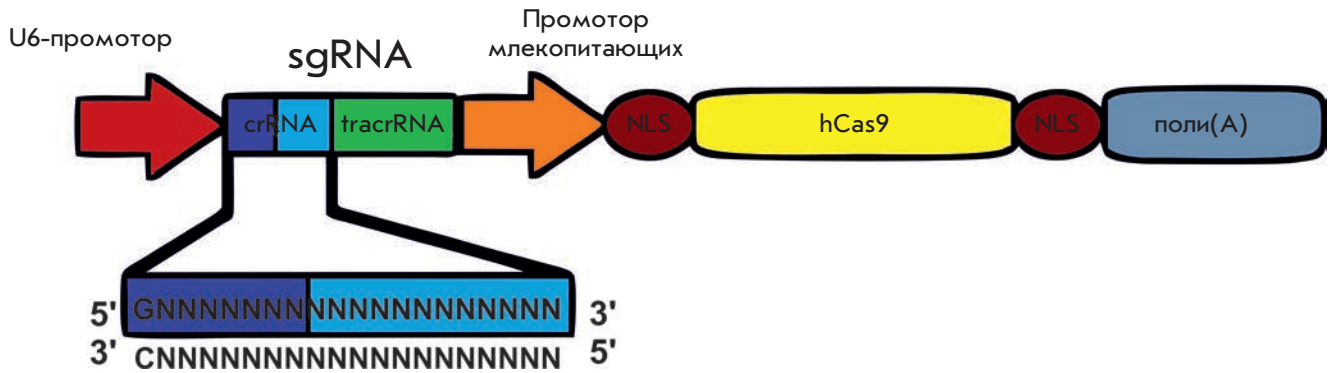


Рис. 6. Схема генетической конструкции, экспрессирующей элементы системы CRISPR/Cas. hCas9 – последовательность белка Cas9, оптимизированная для экспрессии в клетках эукариот. sgRNA – единая химерная РНК, содержащая части crRNA и tracrRNA, необходимые для функционирования. NLS – сигнал ядерной локализации, который обеспечивает попадание конструкций в ядро. Поли(A) – сигнал полиаденилирования

лишь вырезать исходную 30-нуклеотидную направляющую последовательность, фланкированную сайтами BbsI, и заменить ее на искусственно синтезированную. С этой целью 30-членные олигонуклеотиды, комплементарные целевой последовательности и содержащие соответствующие липкие концы, сплавляются вместе и лигируют с плазмидой.

Плазмиды рХ330/рХ335 содержат две экспрессионные кассеты: Cas9-нуклеаза/никаза, химерная sgRNA, включающая 85-нуклеотидную tracrRNA. Принцип смены направляющей последовательности тот же, но длина ее меньше – 20 нуклеотидов, причем 20-м должен быть гуанин, так как используемый в данном случае U6-промотор содержит это основание в точке старта транскрипции. Кроме того, в эти плазмиды могут быть встроены дополнительные элементы, такие, как сайты 2A-GFP или 2A-Puro, для последующей селекции несущих плазмиду клеток.

Доставка конструкций, экспрессирующих компоненты системы CRISPR/Cas9

Для трансформации клеточных культур человека, мыши и других организмов чаще используют плазмиды, обеспечивающие активную продукцию нуклеазы Cas9 и sgRNA *in vitro* [46, 78–80]. В случае трансформации целого организма разработан метод, основанный на микроинъекции мРНК cas9 и sgRNA в одноклеточные эмбрионы [81–83]. Этот метод активно применяют у мышей, полосатого данио (*Danio rerio*) и дрозофилы. Для широкомасштабного охвата генома нокаута с использованием больших библиотек sgRNA используют лентивирусные векторы [84, 85]. У растений, клетки которых имеют плотную клеточную стенку, широко применяется метод плазмидной трансформации протопласта в клеточ-

ных культурах [86, 87], а также агроинфильтрация при помощи *Agrobacterium tumefaciens* [88, 89].

Анализ мутаций, вызванных действием CRISPR/Cas9 и TALEN

В результате активности систем CRISPR/Cas9 или TALENs в эукариотическую ДНК вносится двухцепочечный разрыв в области протоспейсера CRISPR/Cas9 или спейсерной последовательности, разделяющей сайты узнавания TALEN (*рис. 7*). В отсутствие гомологичной донорной ДНК двухцепочечный разрыв репарируется путем нехомологичного сшивания концов. В ходе этого процесса возникают ошибки, и с большой частотой в области соединения образуются небольшие инсерции или делеции [90]. Разработан ряд методик, позволяющих изучать активность искусственных нуклеаз в эукариотических клетках, основой которых является детекция изменений в целевой ДНК (*рис. 7*).

Изучить нуклеотидные последовательности мутантных аллелей, возникающих в результате нехомологичного сшивания концов ДНК, а также наиболее точно количественно определить эффективность работы искусственных нуклеаз позволяет метод, основанный на ТОРО-клонировании (*рис. 7*). Из клеток эукариот, обработанных искусственными нуклеазами, выделяют геномную ДНК, и методом ПЦР амплифицируют участок ДНК, содержащий сайт узнавания нуклеазы. Продукты ПЦР клонируют в плазмидном векторе с последующим секвенированием клонов, полученных после трансформации клеток *E. coli* [72]. Таким образом определяют разнообразие возникающих мутаций и их частоту. Более того, если из клеток, обработанных искусственными нуклеазами, получить клональные популяции, то после секвенирования

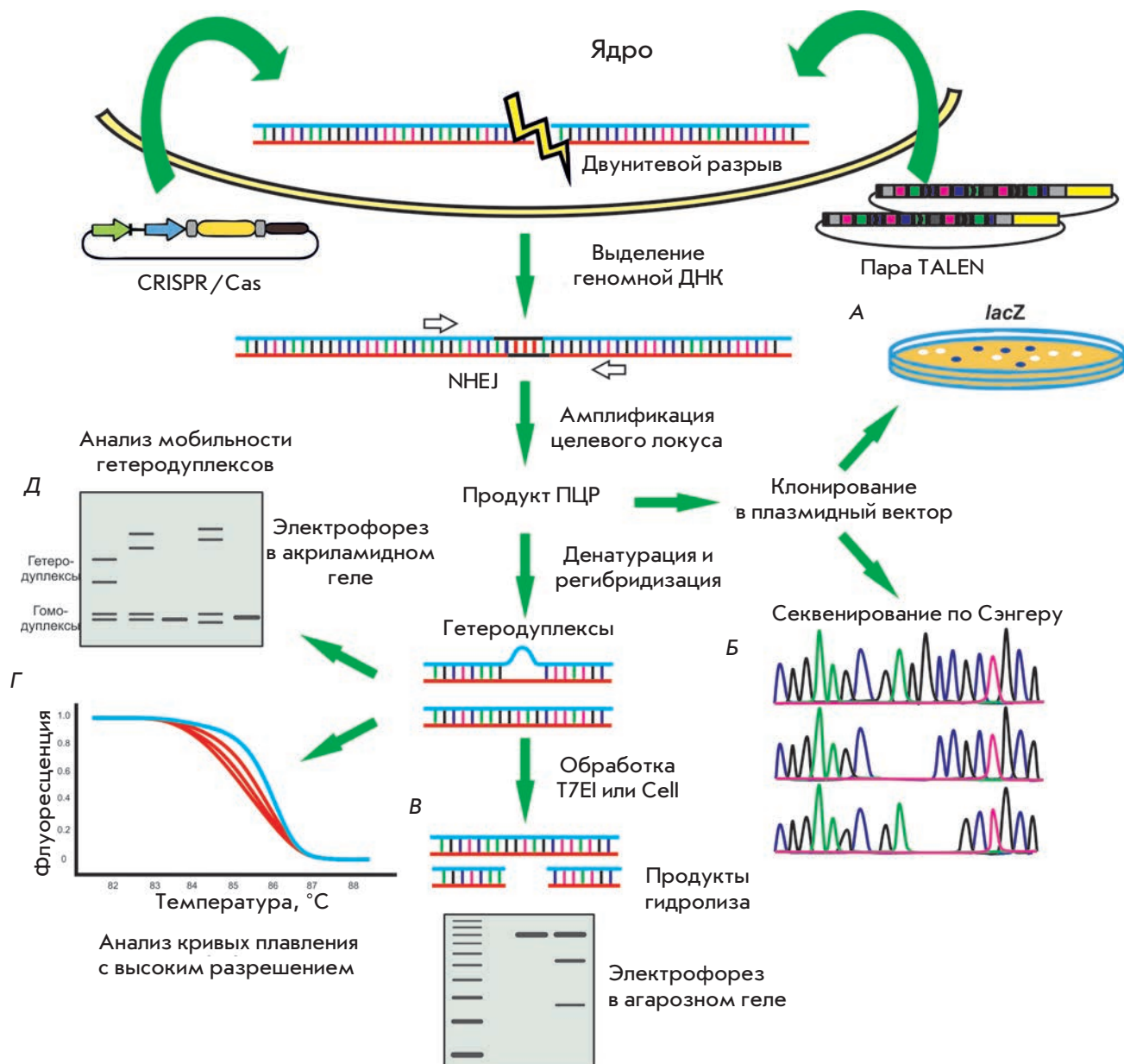


Рис. 7. Схема различных анализов для выявления и определения эффективности внесения двухцепочечного разрыва системами TALEN и CRISPR/Cas. Сначала конструкции, кодирующие CRISPR/Cas9 или TALEN, доставляют в клетки. В клетках происходят двухцепочечные разрывы в целевых локусах, которые репарируются с помощью негомлогичного соединения концов (NHEJ). В результате образуются инсерции или делеции. Далее амплифицируют целевой локус с помощью ПЦР. Продукты ПЦР анализируют с помощью следующих методов. А – клонируют целевой участок в плазмидный вектор. Из-за инсерций или делеций происходит нарушение или наоборот восстановление рамки считывания гена *lacZ*. В результате подсчета синих и белых колоний после трансформации *E. coli* определяют эффективность работы систем CRISPR/Cas9 или TALEN; Б – после клонирования в плазмидный вектор и трансформации *E. coli* производят секвенирование по Сэнгеру. Подсчитывают клоны, содержащие инсерции/делеции, определяют эффективность; В – после денатурации и регибридации продукта ПЦР образуются гетеродуплексы ДНК – например, одна цепь «дикого типа», другая содержит делецию. После обработки ферментами, разрезающими ДНК в неспаренных участках, наносят образцы на гель и проводят электрофорез. Продукты гидролиза означают, что проба содержала гетеродуплексы, следовательно, в геномной ДНК возник двухцепочечный разрыв под действием CRISPR/Cas9 или TALEN; Г – в результате анализа кривых плавления с высоким разрешением также можно детектировать гетеродуплексы. Синий – контрольные образцы, красный – образцы, содержащие гетеродуплексы; Д – неспаренные участки ДНК снижают мобильность гетеродуплексов в 15% полиакриламидном геле. После гель-электрофореза можно наблюдать полосы, соответствующие гомо- и гетеродуплексам

можно отобрать линии, несущие определенные мутации. Так, например, в результате отбора клонов с делецией определенного размера были получены клеточные линии, в которых удалось восстановить рамку считывания, нарушенную мутацией, вызывающей мышечную дистрофию Дюшенна [91].

Активность искусственных нуклеаз анализируют при помощи ферментов, которые расщепляют фосфодиэфирные связи на неспаренных участках ДНК (рис. 7). При амплификации участка, выбранного в качестве мишени для искусственных нуклеаз, получается смесь из молекул ДНК, нуклеотидные последовательности которых отличаются из-за возникших при негомологичном соединении концов инсерций или делеций. При денатурации, а затем регибридизации продукта ПЦР образуются гетеродуплексы, содержащие петли на неспаренных участках. Регибридизованные продукты ПЦР обрабатывают ферментами, например, эндонуклеазой I фага T7 [92] или нуклеазами семейства *CelI* [93], а затем полученные фрагменты разделяют электрофоретически. Обнаружение продуктов гидролиза свидетельствует о присутствии в смеси продуктов ПЦР фрагментов, содержащих инсерции или делеции, возникшие при негомологичном соединении концов. По соотношению интенсивности основного продукта и фрагментов, возникших при гидролизе, можно оценить эффективность работы искусственных нуклеаз, однако это неточная оценка [92].

Образовавшиеся гетеродуплексы по своим свойствам отличаются от гомодуплексов. Одно из таких отличий – изменение профиля кривой плавления, которое можно обнаружить путем анализа кривых плавления с высоким разрешением (HRMA) (рис. 7). С помощью ПЦР в реальном времени с использованием интеркалирующих флуоресцентных красителей амплифицируют небольшой участок (100–300 п.н.), содержащий сайт двухцепочечного разрыва. Затем после денатурации и регибридизации проводят HRMA. В результате сравнения контрольных и опытных образцов можно определить изменения профиля кривой плавления, а следовательно, и изменения в нуклеотидных последовательностях, возникших в результате негомологичного сшивания концов [94]. Этот анализ отличается чувствительностью и простотой, однако это качественный метод, который не позволяет точно оценить эффективность работы искусственных нуклеаз, а также характер изменений ДНК.

Еще один метод, с помощью которого можно определить был ли внесен в целевой сайт двухцепочечный разрыв, это анализ электрофоретической подвижности гетеродуплексов. Неспаренные участки одноцепочечной ДНК, образующие петли в гетероду-

плексах, снижают их подвижность в 15% полиакриламидном геле по сравнению с гомодуплексами. Благодаря этому свойству можно не только установить, произошел ли двухцепочечный разрыв, но и оценить разнообразие возникающих мутаций, а также генотипировать различные клоны, поскольку различные по размеру делеции или инсерции по-разному меняют подвижность гетеродуплексов. При этом профиль подвижности для линий, содержащих одинаковые мутации, также будет одинаковым [95].

Количественно оценить и сравнить эффективность искусственных нуклеаз позволяют методы, в которых применяют репортерные генетические конструкции, содержащие гены люминесцентных белков. При этом используется явление одноцепочечного отжига (Single-strand annealing, SSA), одного из способов репарации двухцепочечных разрывов в геноме эукариот. Если двухцепочечный разрыв происходит между двумя прямыми повторами, то в результате SSA происходит отжиг комплементарных последовательностей, фланкирующих разрыв. Затем негомологичные участки подвергаются гидролизу специфическими нуклеазами, а в одноцепочечных участках происходит синтез и лигирование новой ДНК. Последовательность между прямыми повторами, где произошел двухцепочечный разрыв, всегда делетируется, а вместо двух повторенных последовательностей остается один. Этот процесс используют для восстановления репортерного гена, например гена люциферазы. После двухцепочечного разрыва, внесенного в целевую последовательность, клонированную в плазмидный вектор между двумя повторенными элементами гена-репортера, в результате SSA происходит восстановление его (репортера) функции. Таким образом, по уровню люминесценции можно количественно оценить эффективность работы искусственных нуклеаз. При этом репортерные конструкции трансфицируют в эукариотические клетки, например в линии НЕК293 или некоторые линии дрожжей. Недостатком этого метода является то, что он не учитывает геномного окружения, в котором находится сайт разрезания, поэтому его результаты могут не коррелировать с результатами, полученными при работе с целевыми сайтами в геноме [96].

Учеными из Японии разработан метод анализа, основанный на нарушении/восстановлении функции гена *lacZα* (рис. 7). С этой целью сайт для внесения двухцепочечного разрыва клонируют в ген *lacZα*. Олигонуклеотидные праймеры при этом выбирают таким образом, чтобы целевой сайт дикого типа нарушал (1) или сохранял (2) рамку считывания. Если в сайте произошел двухцепочечный разрыв, репарированный с помощью негомологичного сшивания концов, то в первом случае после клонирования в ре-

зультате делеций или инсерций в 1/3 конструкций рамка считывания будет восстановлена. Соответственно после трансформации клеток *E. coli* полученными конструкциями часть колоний будет иметь синюю окраску. Во втором случае в 2/3 конструкций из-за мутаций, вызванных искусственными нуклеазами, рамка считывания будет нарушена. Колонии, в которые попадут такие генетические конструкции, будут иметь белую окраску. Эффективность искусственных нуклеаз можно тогда определить путем простого подсчета доли синих или белых колоний в первом и втором случаях соответственно [97].

ПРИМЕНЕНИЕ СИСТЕМ CRISPR/Cas9 И TALEN

Нуклеаза делает в сайте-мишени двухцепочечные разрывы, которые репарируются в клетке по одному из двух возможных механизмов:

Негомологичное соединение концов, при котором возникают ошибки, что приводит в результате к появлению в целевом локусе мутаций по типу инделей (инсерций, делеций).

Гомологичная рекомбинация, при которой неповрежденный гомолог служит матрицей для восстановления исходной структуры ДНК; это событие происходит в клетке довольно редко, но использование CRISPR/Cas9 и TALENs позволяет повысить вероятность прохождения гомологичной рекомбинации на несколько порядков. Если добавить к компонентам CRISPR/Cas9 искусственно синтезированную молекулу ДНК, имеющую гомологию с последовательностью нуклеотидов в месте разрыва, то она может служить матрицей для другого способа репарации – HDR (homology-directed repair), при котором небольшой фрагмент искусственной матрицы встраивается в целевой локус. В качестве такой матрицы чаще всего используют два типа конструкций: одноцепочечные олигонуклеотиды или плазмидные векторы. В первом случае искусственно синтезируют олигонуклеотиды, гомологичные сайту внесения двухцепочечного разрыва, оптимальная длина которых составляет около 90 нуклеотидов [98]. Такие олигонуклеотиды могут содержать небольшие отличия от целевого сайта. При использовании плазмидных векторов в качестве донорных молекул для рекомбинации в них клонируют достаточно длинные плечи гомологии (от 500 до нескольких тысяч пар нуклеотидов). Такие плечи гомологии могут фланкировать дополнительные элементы, например, репортерные гены, гены устойчивости к антибиотикам и так далее. С помощью HDR в целевой локус можно поместить сайт рестрикции, маркерную метку или же нуклеотиды для исправления «неправильных» ДНК. Однако HDR активно происходит только в делящихся клетках, и ее эффективность очень сильно зависит

от типа клетки, стадии жизнедеятельности, а также целевого локуса генома и самой матрицы [99].

Таким образом, с помощью сайт-специфических нуклеаз можно получить следующие мутации:

- негомологичное соединение концов в отсутствие донорной плазмиды опосредует делеции или инсерции нескольких нуклеотидов сайта-мишени и, как один из результатов, генный нокаут из-за мутаций рамки считывания и образования стоп-кодонов [100];
- в присутствии двухцепочечных олигонуклеотидов или донорной плазмиды фрагменты ДНК длиной свыше 14 т.п.н. могут встраиваться посредством лигирования, опосредованного негомологичным сшиванием концов [101, 102];
- одновременное внесение нескольких двухцепочечных разрывов может приводить к делециям, инверсиям или транслокациям участков ДНК, расположенных между этими разрывами [46, 103];
- гомологичная рекомбинация в присутствии донорной плазмиды с плечами гомологии, фланкирующими встраиваемый фрагмент [104], линейной донорной последовательности с гомологией менее 50 п.н. [105] или олигонуклеотида [103] приводит к внедрению одного или нескольких трансгенов для коррекции или замены существующих генов.

На сегодняшний день описанные выше методы активно применяются в фундаментальных и прикладных исследованиях. При этом редактирование генома возможно как *in vitro* при доставке элементов систем TALEN или CRISPR/Cas в культуры клеток, так и *in vivo* с помощью инъекций мРНК в зиготы (рис. 3).

Редактирование генома *in vitro*

Клеточные линии HEK293T/HEK293FT, легко трансформируемые плазмидами и относительно простые в поддержании, наиболее часто используются для проверки эффективности работы систем TALENs и CRISPR/Cas в человеческой модели *in vitro* [46, 50, 68, 78, 106]. По данным разных авторов уровень целевых мутаций, а также гомологичной рекомбинации с донорными плазмидами/олигонуклеотидами колеблется в широких пределах, что, вероятно, зависит не только от метода, но и от клеточной линии и самой геномной мишени (таблица). Культивируемые линии индуцированных плюрипотентных стволовых клеток и эмбриональных стволовых клеток человека представляют особый интерес для регенеративной медицины, изучения структуры и функционирования сложных генных сетей, создания систем для поиска лекарственных средств и множества других фундаментальных и биомедицинских исследований.

С использованием системы TALEN Динг и соавт. [71] внесли двухцепочечные разрывы и получили линии стволовых клеток человека с мутациями в раз-

Геномная инженерия с помощью TALEN и CRISPR/Cas

Нуклеаза	Объект	Ген	Назначение	Источник
TALEN	Клетки человека (<i>Homo sapiens</i>)	<i>ccr5, akt2, e17k, angptl3, apob, atgl, c6orf106, celsr2, cftr, ciita, foxo1, foxo3, gli1, glut4, hbb, hdac1, hdac2, hdac6, hmga2, hoxa13, hoxa9, hoxc13, hpvt, il2rg, jak2, kras, linc00116, maoa, map2k4, mdm2, met, mlh1, msh2, mutyh, myc, mycl1, mycn, nbn, ncor1, ncor2, nlr5, ntf3, pdgfra, pdgfrb, phf8, plin1, pms2, ppp1r12c (aavs1), ptch1, pten, rara, rbbp5, recql4, ret, runx1, sdhb, sdhc, sdhd, setdb1, sirt6, smad2, sort1, sox2, klf4ss18, suz12, tfe3, tp53, trib1, tsc2, ttn, vhl, xpa, xpc, abl1, alk, apc, atm, axin2, bax, bcl6, bmpr1a, brca1, brca2, cbx3, cbx8, ccnd1, cdc73, cdk4, cdh4, chd7, ctmb1, cylld, ddb2, ercc2, ewsr1, ext1, ext2, ezh2, fanca, fancf, fancg, fes, fgfr1, fh, flcn, flt4, mstn, aavs2, oct4, pitx3</i>	Нокаут, встройка	[67, 68, 70–72, 74, 92, 176–179, 180]
	Дрожжи (<i>Saccharomyces cerevisiae</i>)	<i>URA3, ADE2, LYS3</i>	Нокаут, встройка	[181]
	Нематода (<i>Caenorhabditis elegans</i>)	<i>ben-1, tex-1, sdc-2</i>	Нокаут	[182]
	Дрозофила (<i>Drosophila melanogaster</i>)	<i>yellow, crhdr1, ponzr1, bmil, cdh5, dip2a, elmo1, epas1b, fh, golden, gria3, hey2, hif1ab, ikzf1, jak3, moesina, myod, phf6, ppp1cab, ryr1a, ryr3, scl6a3, tbx6, tnkb, th, fam46c, smad5</i>	Нокаут, встройка	[94, 183–187]
	Шелкопряд (<i>Bombyx mori</i>)	<i>blos2</i>	Нокаут	[188]
	Сверчок (<i>Gryllus bimaculatus</i>)	<i>lac2</i>	Нокаут	[189]
	Шпорцевая лягушка (<i>Xenopus tropicalis</i>)	<i>ets1, foxd3, grp78/bip, hhex, noggin, ptf1a/p48, sox9, vpp1</i>	Нокаут	[190]
	Мышь (<i>Mus musculus</i>)	<i>c9orf72, fus, lepr, pak1ip1, gpr55, rprm, fbx6b, smurf1, tmem74, wdr20a, dcaf13, fam73a, mlkl, mstn, pibf1, sepw1, rab38, zic2</i>	Нокаут, встройка	[179, 191–196]
	Крыса (<i>Rattus norvegicus</i>)	<i>bmpr2, IqM</i>	Нокаут	[197, 198]
	Свинья (<i>Sus scrofa</i>)	<i>amely, dmd, gdf8, ggta, ghdrhdr, il2rg, ldlr, rag2, rela (p65), sry</i>	Нокаут	[199]
	Корова (<i>Bos taurus</i>)	<i>acan, gdf8, ggta, mstn, prnp</i>	Нокаут	[179, 199]
	Арабидопсис (<i>Arabidopsis thaliana</i>)	<i>adh1</i>	Нокаут	[70]
	Табак (<i>Nicotiana benthamiana</i>)	<i>surA, surB, hax3</i>	Нокаут, встройка	[156, 157]
	Коротконожка (<i>Brachypodium distachyon</i>)	<i>aba1, cxx2, coi1, hta1, rht, sbp, smc6, spl</i>	Нокаут	[154]
Рис (<i>Oryza sativa</i>)	<i>avrxa7, pthxo3, badh2, cxx2, dep1, sd1</i>	Нокаут	[154, 155]	
CRISPR/Cas	Дрожжи (<i>Saccharomyces cerevisiae</i>)	<i>CAN1, ADE2</i>	Нокаут, встройка	[200]
	Клетки человека (<i>Homo sapiens</i>)	<i>dnmt3b-tdTomato, pou5f1(oct4), emx1, dyrk1a, grin2b, egfp, ccr5, c4bpb, pvalb, aavs, akt2, celsr2, ciita, glut4, linc00116, sort1, ldlr</i>	Нокаут, встройка	[46, 51, 78, 80, 201, 202]
	Нематода (<i>Caenorhabditis elegans</i>)	<i>dpy-11, unc-4, ben-1, unc-36, daf-2, klp-12, lab-1, egfp, dpy-11, lin-5, rol-1, dpy-3, unc-1, dpy-13, unc-119, klp-12</i>	Нокаут, встройка	[119–124, 126]
	Дрозофила (<i>Drosophila melanogaster</i>)	<i>yellow, white, rosy, cg14251 (k81), cg3708cg17629 (kl-3), light</i>	Нокаут, встройка	[130–133]
	Полосатый данио (<i>Danio rerio</i>)	<i>etsrp, gata5, etsrp, gsk3b, апоea, fh, fh1, th1, rgs4, tia1l, tph1a, drd3, egfp, tyr, gol, mitfa, ddx19, sema3fb, dre-mir-126a, dre-mir-126b, dre-mir-17a-1-dre-mir-92a-1, dre-mir-17a-2-dre-mir-92a-2, fgd5, ensdarg00000070653, ensdarg00000076787, psmf1, dre-mir-126a, dre-mir-17a-2, dre-mir-92a-2, tardbp, tardbpl, c13h9orf72</i>	Нокаут, встройка, хромосомные перестройки	[81, 82, 203–206]
	Лягушка (<i>Xenopus tropicalis</i>)	<i>tyr, six3</i>	Нокаут	[207]
	Свинья (<i>Sus scrofa</i>)	<i>gdf8, p65</i>	Нокаут, встройка	[208]
	Мышь (<i>Mus musculus</i>)	<i>tet1, tet2, tet3, sry, uty, rosa26, hpvt, egfp, th, rheb, uhrf2</i>	Нокаут, встройка	[83, 144, 209, 210]
	Крыса (<i>Rattus norvegicus</i>)	<i>dnmt1, dnmt3a, dnmt3b, tet1, tet2, tet3, mc3r, mc4r</i>	Нокаут, встройка	[144, 145, 211]
	Арабидопсис (<i>Arabidopsis thaliana</i>)	<i>pds33, fls2, bri1, jaz1, gaj, chl, chl2, 5g13930</i>	Нокаут, встройка	[87, 88, 149]
	Табак (<i>Nicotiana benthamiana</i>)	<i>pds</i>	Нокаут, встройка	[88, 89]
	Рис (<i>Oryza sativa</i>)	<i>ods, badh2, mrk2, 02g23823, roc5, spp, ysa, myb1, cao1, lazy1, sweet11, sweet14</i>	Нокаут, встройка	[86, 150, 152]
	Пшеница (<i>Triticum aestivum</i>)	<i>mlo</i>	Нокаут	[86]

ных генах, вызывающих заболевания. Всего мутации внесли в 15 генов и провели подробный анализ фенотипа дифференцированных производных стволовых клеток с мутациями в четырех из них – *APOB*, *SORT1*, *AKT2* и *PLIN1*. Благодаря таким клеточным моделям были получены новые данные о роли этих генов в патогенезе заболеваний. Например, показано, что продукт гена *APOB* необходим для репликации вируса гепатита С в гепатоцитах человека. В клетках с гомозиготной мутацией в этом гене репликация вируса существенно снижена. А мутация E17K в гене *AKT2* приводит к снижению синтеза глюкозы в гепатоцитах человека и увеличению уровня триглицеридов в адипоцитах.

Помимо создания моделей, необходимых для разработки подходов к лечению заболеваний, искусственные нуклеазы могут использоваться непосредственно в терапевтических целях. Одно из таких направлений – терапия хронических вирусных инфекций. Могут быть сконструированы нуклеазы TALENs, позволяющие вносить мутации в открытые рамки считывания таких вирусов, как ВИЧ, вирус гепатита В, герпесвирус, которые могут находиться в организме в латентном состоянии и не поддаваться терапии, направленной на реплицирующиеся вирусы [107, 108]. Так с помощью TALENs можно модифицировать ген С-С рецептора хемокина 5 Т-лимфоцитов, мутации в котором делают человека невосприимчивым к ВИЧ [100, 109].

С помощью технологии, основанной на CRISPR/Cas9, созданы изогенные стволовые клетки человека [110], разрабатываются методы исправления мутантного фенотипа клеток [111], проводятся работы по регуляции экспрессии генов [112–116], изучению функциональных взаимоотношений между большими группами генов [84, 85] и визуализации функционирующих районов генома в живых клетках [117].

Создание панелей изогенных плюрипотентных стволовых клеток человека позволит осуществить моделирование наследственных и многофакторных заболеваний, скрининг больших библиотек лекарственных средств, а также поиск новых мутаций, вовлеченных в патологический процесс. В настоящее время активно ведутся работы по всем этим направлениям. Так, CRISPR/Cas9-систему эффективно использовали для создания модели синдрома ICF (ICF – immunodeficiency, centromeric region instability and facial anomalies syndrome; иммунодефицит, нестабильность центромерных районов хромосом и лицевые аномалии) на индуцированных плюрипотентных стволовых клетках человека. Были получены гомозиготные мутации в гене *DNMT3B* с частотой 63%, при этом клетки имели фенотип центромерной нестабильности [110]. Особенно актуальным представляется

изучение тяжелых нейродегенеративных заболеваний, таких, как болезнь Альцгеймера, Паркинсона, различные мышечные атрофии.

Поскольку Cas9 узнает конкретную мишень в геноме при участии короткой направляющей последовательности в sgRNA, то в современных условиях относительно просто создать достаточно большую библиотеку олигонуклеотидов и соответственно sgRNA, охватывающую масштабы целого генома. А использование в качестве вектора для доставки компонентов CRISPR/Cas9 лентивирусов, которые стабильно поддерживаются в геноме и реплицируются вместе с геномной ДНК, позволило разработать новую технологию GeCKO – CRISPR-Cas9-нокаут в масштабе генома (Genome-scale CRISPR/Cas9 knockout) [84]. Большая библиотека sgRNA позволяет выключить транскрипцию многих генов одновременно и установить тем самым функциональные взаимоотношения между ними, роль в тех или иных процессах жизнедеятельности или вовлеченность в патологический процесс. Так, при использовании лентивирусной библиотеки, охватывающей 18080 генов (три-четыре sgRNA на каждый ген), выявлены гены, необходимые для жизнедеятельности раковых клеток (A375 клеточная линия меланомы человека) и плюрипотентных стволовых клеток (линия HUES62) [84]. Показано, что в формировании резистентности к вемурафенибу (PLX), который является BRAF-ингибитором протеинкиназы при меланоме, участвуют не только гены *NF1* и *MED12*, но ген *CUL3*, а также комплекс гистон-специфических ацетилтрансфераз STAGA: *TADA1* и *TADA2* [84]. При использовании лентивирусной библиотеки, содержащей порядка 73000 sgRNA, на примере линий опухолевых клеток KBM7 и HL60 изучены гены, участвующие в пролиферации и клеточном цикле [83]. Показано, что мутации, приводящие к формированию нефункциональных продуктов четырех генов репарации однонуклеотидных замен в ДНК (MMR) – *MSH2*, *MSH6*, *MLH1*, *PMS2*, обуславливают устойчивость к нуклеотидному аналогу 6-тиогуанину и, следовательно, обеспечивают пролиферацию клеток. Изучена также работа генов *TOP2A*, *CDK6*, *BCR*, *ABL1* и генов, кодирующих рибосомные белки.

Таким образом, использование библиотек CRISPR/Cas9 позволяет осуществлять функциональный скрининг геномов, который может дать важнейшие сведения о физиологии и биохимии клеток разного типа, поможет раскрыть молекулярные механизмы развития заболеваний и выявить потенциальные мишени для лекарственной и генной терапии.

Методы, основанные на использовании системы CRISPR/Cas9, могут эффективно применяться для редактирования геномов культивируемых

стволовых клеток. В частности, применение систем редактирования геномов позволяет исправлять точечные мутации в клетках, полученных от больных. Объектом исследования в данном случае могут быть индуцированные плюрипотентные стволовые клетки и региональные стволовые клетки. При этом донорными молекулами могут служить как сложные генетические конструкции, так и одноцепочечные ДНК-олигонуклеотиды [98].

Интересным примером подобного подхода представляется работа, в которой осуществлена коррекция локуса *CFTR* (cystic fibrosis transmembrane conductor regulator, муковисцидозный регулятор трансмембранной проводимости) в культивируемых кишечных стволовых клетках, полученных от больных муковисцидозом (CF – cystic fibrosis, муковисцидоз) [111]. Этот подход позволяет получать так называемые органоиды – функциональные многоклеточные образования с исправленным геномом, аутологичные по отношению к донору клеток, которые могут быть введены обратно в организм больного. Безусловно, данное направление открывает большие перспективы для клеточной терапии заболеваний человека.

В случае функциональной коррекции генетических аномалий, связанных с делецией генов или нарушениями экспрессии, которые выражаются в существенном снижении уровня продуктов гена (белка или РНК), можно использовать контролируемое внесение трансгенов в геном. Существуют участки генома, внесение трансгенов в которые считается безопасным. Это такие сайты, как AAVS1, которые обеспечивают стабильную экспрессию введенного трансгена [118]. Таким образом, системы TALENs и CRISPR/Cas могут эффективно применяться в функциональной геномике клеток, для создания клеточных моделей заболеваний человека и клеточной терапии.

Редактирование генома *in vivo*

В генетике за многие годы ее существования сформировался ряд модельных объектов, изученных наиболее подробно и используемых в большинстве фундаментальных и прикладных исследований. К модельным организмам относятся, например, дрожжи, нематода, дрозофила, арабидопсис, полосатый данио, лабораторная мышь и крыса. На этих и ряде других модельных организмов активно проводятся эксперименты по геномной инженерии с помощью систем CRISPR/Cas9 и TALENs.

Различные варианты применения CRISPR/Cas и модификации технологии редактирования генома у нематоды *Caenorhabditis elegans* представлены в целом ряде работ [119–126]. С помощью инъекции мРНК/белка Cas9 и продуцированной *in vitro/in vivo* sgRNA в клетки зародышевой линии у взрослых жи-

вотных в следующем поколении получали стабильные целевые модификации генома, включая небольшие инсерции/делеции, более крупные хромосомные делеции и перестройки [119], внедрение трансгена путем гомологичной рекомбинации с донорными молекулами [121, 123]. Такой метод активно используется для изучения процессов дозовой компенсации у нематоды, сравнения функций генов у родственных видов *C. elegans* и *C. briggsae* [122].

Плодовая мушка *Drosophila melanogaster* относится к наиболее изученным модельным объектам. Однако получение новых мутантных аллелей посредством гомологичной рекомбинации по-прежнему остается очень трудоемкой процедурой [127–129]. Инъекция мРНК *cas9* и sgRNA в эмбрионы дрозофилы обеспечивает получение двухцепочечных разрывов в целевых локусах генома, репарация которых приводит к формированию мутаций по типу инсерций/делеций на довольно высоком уровне (*таблица*). Эмбриональная инъекция позволяет получить мутации в обоих аллелях целевого гена и во всех клетках развивающегося впоследствии взрослого насекомого, однако при этом появляется определенный процент мозаиков [130–132]. Эти мутации стабильно передаются из поколения в поколение, что обеспечивает возможность создания новых линий мух [133]. Недавно было разработано приложение (<http://www.flyrnai.org/crispr>), с помощью которого можно наиболее эффективно планировать эксперименты по редактированию генома у дрозофилы. Таким образом, технология CRISPR/Cas9 позволяет быстро и эффективно получать мутации с целью дальнейшего изучения функционирования генов у *Drosophila*.

Полосатый данио является на сегодняшний день очень популярным объектом не только для фундаментальных исследований структурно-функциональных взаимоотношений в геноме, но и для моделирования метаболических и нейродегенеративных заболеваний человека *in vivo* [134]. Посредством инъекции компонентов CRISPR/Cas9 в эмбрионы полосатого данио получены разнообразные целевые модификации, стабильно передающиеся по наследству (*таблица*). В 2011 году был открыт международный проект – The Zebrafish Mutation Project – по созданию мутантных аллелей в каждом белоккодирующем гене полосатого данио. Все данные анализируются на веб-сайте http://www.sanger.ac.uk/Projects/D_erio/zmp. На июнь 2013 года получены мутантные модели 46% всех белоккодирующих генов зебрафиш.

Такие лабораторные животные, как мышь и крыса, считаются важнейшими модельными объектами для изучения заболеваний человека, фундаментальных исследований структуры и функции генов

и регуляции их экспрессии, а также в фармакологии и токсикологии. Ранее линии мышей с нокаутом определенных генов получали с помощью гомологичной рекомбинации в эмбриональных стволовых клетках [1, 83], а также посредством инсерционного мутагенеза [135, 136]. Это очень длительные и трудоемкие эксперименты, а получение животных с двойным нокаутом представляет собой еще более сложную задачу. Технология редактирования генома, основанная на CRISPR-Cas9, – это более быстрый и менее трудоемкий способ, позволяющий делать эту работу за один этап. Направленная инъекция сайт-специфических нуклеаз в зиготу на стадии одной клетки обеспечивает возникновение двухцепочечных разрывов в ДНК локуса-мишени [137–139]. Такие разрывы репарируются по механизму негомологичного сшивания концов, что приводит к появлению мутантных крыс и мышей, несущих делеции либо инсерции в разрезанном сайте [140, 141]. При добавлении донорной плазмиды или олигонуклеотида разрывы могут репарироваться с помощью высокоточного механизма гомологичной рекомбинации, что позволяет получить животных, несущих целевые вставки ДНК [83, 142, 143]. Редактирование генома с использованием CRISPR/Cas9 обеспечивает внесение мутаций как в один ген, так и в несколько генов сразу. Так, показано, что CRISPR-Cas9 с высокой эффективностью приводит к появлению в эмбриональных стволовых клетках мыши мутаций в пяти генах одновременно, а введение мРНК *cas9* и sgRNAs, направленных на гены *Tet1* и *Tet2*, в зиготы мыши позволяет получить животных с биаллельными мутациями обоих генов с эффективностью 80% [83]. Аналогичные результаты получены в экспериментах на крысах, причем и мыши, и крысы стабильно наследовали выявляемые мутации [144, 145]. Кроме того, произведена эффективная коррекция мутации в гене *Crygc* у мышей с доминантной формой катаракты, вызванной этой мутацией [146]. Создание модельных грызунов, несущих специфические мутации в нескольких локусах, делает возможным анализ функций генов, входящих в состав генных семейств с избыточными функциями, а также эпистатических взаимодействий генов. Данные, объединяющие информацию по нокауту того или иного гена мыши, собраны на сайте международного проекта IMPC (International mouse phenotyping consortium, <https://www.mousephenotype.org/>).

Редактирование генома с помощью TALENs и системы CRISPR/Cas9 активно используется и у растений. Направленное редактирование геномов растений может применяться для решения задач как фундаментальной – изучение функций генов, так и прикладной науки – получение растений с новыми

свойствами, такими, как устойчивость к патогенам, гербицидам, изменения метаболизма, показателей урожайности и т.д. [147]. При этом для доставки генно-инженерных конструкций в основном используют трансформацию протопласта или *in planta*-экспрессию с *Agrobacterium tumefaciens* (агроинфильтрация) [148]. Таким образом получены генные нокауты и точные модификации у растений арабидопсиса, пшеницы, риса, табака [86, 88, 89, 149–153].

Редактирование геномов растений с помощью системы TALEN к настоящему времени проведено на четырех модельных объектах [70, 154–157]. Примером растения, которое приобрело новые свойства в результате редактирования генома системой TALEN, служит рис, устойчивый к патогену *Xanthomonas oryzae pv.* С помощью искусственных TALENs был внесен двухцепочечный разрыв в сайт узнавания природного TAL-эффектора патогена в локусе гена *Os11N3*. Таким способом были получены растения, устойчивые к заражению *X. oryzae pv* [155].

АЛЬТЕРНАТИВНЫЕ ВАРИАНТЫ ПРИМЕНЕНИЯ TALE И CRISPR/Cas9

Расшифровка кода узнавания белками TALE целевых нуклеотидных последовательностей, а также разработка методов создания искусственных ДНК-связывающих доменов, основанных на этом коде, позволила ученым создавать химерные белки, способные направленно воздействовать на геном. Такие белки состоят из ДНК-связывающего и эффекторного доменов. В качестве эффекторного в основном используются нуклеазные домены, однако в ряде работ получены химерные белки, содержащие помимо ДНК-связывающего домена рекомбиназные, гистон-метилтрансферазные, гистон-деацетилазные домены, а также домены, активирующие или подавляющие экспрессию генов. Такие химерные белки обладают колоссальными перспективами применения как в прикладной, так и в фундаментальной науке. Аналогичным образом модифицируют и систему CRISPR/Cas9: к каталитически неактивному белку Cas9 присоединяют тот или иной эффекторный домен, например, активатор или репрессор транскрипции, флуоресцентный белок GFP и т.д.

Регуляция экспрессии генов с помощью систем TALE и CRISPR/Cas9

Для направленной активации экспрессии генов используются конструкции, содержащие ДНК-связывающий домен TALE и синтетический домен VP64 [158] – TALE-TF. Попадая в ядро, химерный белок связывается с целевой нуклеотидной последовательностью, домен VP64 привлекает эндогенные активаторы экспрессии генов [159]. При этом стати-

стически значимо увеличивается экспрессия целевого гена, что обычно подтверждают с помощью ПЦР в реальном времени. Возможна также активация некодирующих генов, например генов микроРНК [160]. Подавление экспрессии целевого гена может достигаться при использовании химерных белков, содержащих домены KRAB [161] или SRDX [162].

Возможным терапевтическим применением TALE-TF является направленная регуляция экспрессии генов, ассоциированных с заболеваниями человека. Для проверки этого подхода данную стратегию применили для повышения уровня экспрессии гена *FXN*, кодирующего белок фратаксин. Экспансия тринуклеотидных повторов GAA в этом гене приводит к развитию атаксии Фридрейха, при этом структура белка не меняется, однако снижается его экспрессия. Показано, что с помощью TALE-TF можно повысить экспрессию гена *FXN* в фибробластах человека, несмотря на увеличенное число тринуклеотидных повторов [163].

Активация экспрессии эндогенных генов позволяет избежать использования эктопической сверхэкспрессии факторов репрограммирования *Oct4*, *Sox2*, *Klf4* и *c-Myc* (OSKM) при получении индуцированных плюрипотентных стволовых клеток. В результате можно получить индуцированные плюрипотентные стволовые клетки, не содержащие трансгенов, и соответственно снизить риск инсерционного мутагенеза, который возникает при использовании лентивирусных векторов, экспрессирующих OSKM. Так, репрограммирование эмбриональных фибробластов мыши к плюрипотентному состоянию было достигнуто с помощью направленной активации экспрессии генов *Oct4* и *Nanog* под действием TALE-TF, содержащих домен VP64 [164].

Совсем недавно были созданы факторы транскрипции для направленной регуляции экспрессии генов, в ответ на внешний химический стимул. Такие факторы состоят из ДНК-связывающего домена TALE и лигандсвязывающего домена рецептора стероидных гормонов. При попадании в клетку лиганда (экидизона) происходит димеризация лигандсвязывающего домена и соответственно активация экспрессии целевого гена [165].

Сочетанием двух крайне перспективных направлений современной биотехнологии – оптогенетики и геномной инженерии – стала разработанная недавно система индуцируемых светом эффекторов транскрипции (light-inducible transcriptional effectors, LITEs). Эта система состоит из двух частей. Первая представляет собой ДНК-связывающий домен TALE, соединенный со светочувствительным доменом – криптохромом 2 (CRY2), выделенным из *Arabidopsis thaliana*. Вторая – активатор транскрипции VP64,

соединенный с CIB1, который способен взаимодействовать с CRY2. Под действием синего света CRY2 изменяет конформацию и связывается с CIB1, таким образом привлекая VP64 к целевому сайту [166]. В работе Конерманн и соавт. [166], разработавших систему LITE, показано статистически значимое увеличение экспрессии ряда генов как в нейронах мыши *in vitro*, так и в головном мозге *in vivo*. Ими также предложена система, в которой домен VP64 заменен метилтрансферазой или деацетилазой, способными модифицировать гистоны.

Интересным применением направленной регуляции транскрипции с помощью TALE-TF является создание генетических логических контуров внутри клетки, основанных на взаимодействии нескольких TALE-TF с промоторами друг друга и репортерным геном, и промоторов факторов, регулирующих экспрессию. С помощью этого подхода внутри клеток созданы логические контуры НЕ-ИЛИ [167] и И [168].

Каталитически неактивная dCas9 или dCas9, к которой присоединены факторы, регулирующие экспрессию генов, также позволяет активировать или репрессировать транскрипцию в клетках человека, бактерий и дрожжей [112–116]. С этой целью можно использовать омега-субъединицу РНК-полимеразы *E. coli* [113], тандемные копии вирусного белка VP64, KRAB-домен [112, 115]. Так, удалось обеспечить высокоспецифичный сайленсинг генов *CD71* и *CXCR4* (на уровне 60–80%), а также эффективный нокдаун локуса *TEF1* у дрожжей [112]. Более того, получена мультиплексная активация/репрессия промоторов нескольких генов, причем тип регуляции (положительная или отрицательная) определяется положением мишени в промоторе гена [114, 115]. Таким образом, система CRISPR/Cas9 может использоваться как модульная платформа, связывающаяся с заданной нуклеотидной последовательностью и привлекающая к ней белковые факторы, тем самым открывая возможности использования этой системы как основного метода точной регуляции экспрессии генов в клетках эукариот.

Визуализация внутренних локусов генома с помощью систем TALE и CRISPR/Cas9

Известно, что организация и динамика хроматина играют определяющую роль в регуляции функционирования генома. Однако чрезвычайно трудно получить визуальное изображение функциональных локусов генома в живых клетках. Использование систем TALE и CRISPR/Cas9 открывает новые возможности в решении этой проблемы.

С помощью конструкций, содержащих ДНК-связывающий домен TALE и флуоресцентный бе-

лок, удалось визуализировать целевые ДНК в динамике [169–171]. Этот подход позволяет изучать пространственно-временную организацию повторенных элементов генома, в том числе центромерных и теломерных повторов.

С помощью эндонуклеазодефицитного белка Cas9, меченного EGFP, и структурно оптимизированных sgRNA разработан метод визуализации повторяющихся элементов в теломерах и кодирующих генов в живых клетках [117]. На опухолевых клеточных линиях RPE, HeLa и UMUC3 удалось визуализировать повторяющиеся и неповторяющиеся элементы в генах *MUC4* и *MUC1*, ответственных за выработку различных форм муцина, одного из компонентов защитной слизи в различных эпителиальных тканях и важного при малигнизации [117]. Таким образом, появляется возможность проводить мониторинг количества копий генов в живой клетке. С использованием этого метода наблюдали динамику процессов элонгации и разрушения теломер, субъядерную локализацию локусов *MUC4*, когезию реплицированных локусов *MUC4* на сестринских хроматидах и их изменяющееся поведение в течение митоза [117]. Такая стратегия обладает значительным потенциалом для изучения конформации и динамики нативных хромосом в живых клетках человека.

Химерные рекомбиназы и транспозазы как альтернатива TALEN

Рекомбиназы и транспозазы представляют собой альтернативу TALEN при редактировании генома. К их преимуществам относится отсутствие зависимости от внутриклеточных механизмов репарации. Эти ферменты также производят разрезание и лигирование в целевых сайтах, и соответственно при этом не происходит накопления двухцепочечных разрывов, которые могут привести к гибели клеток. Вдобавок рекомбиназы и транспозазы встраивают донорную ДНК в геном, что упрощает детекцию их активности. Недостатком этих химерных ферментов является достаточно высокий уровень нецелевых эффектов [172]. В качестве эффекторного домена используют каталитический домен рекомбиназы Gin [173, 174] или транспозазы piggyBac [175]. Активность TALE-рекомбиназ показана с помощью репортерного гена, промотор которого специфически вырезался рекомбиназой Gin. Возможность редактирования генома с использованием транспозазы показана на примере локуса *CCR5*.

ЗАКЛЮЧЕНИЕ

Разработка систем TALEN и CRISPR/Cas9 является важной ступенью в развитии современной геномной инженерии. Появление этих систем, благодаря

их дешевизне и простоте конструирования, стало мощным толчком для развития как фундаментальной, так и прикладной науки. Перспективы применения этих систем в самых различных областях, начиная с пищевой промышленности и заканчивая персонализированной медициной, действительно невероятны. Однако до сих пор остаются открытыми некоторые вопросы, связанные со специфичностью и безопасностью их использования (из-за возможных нецелевых эффектов), способами доставки при терапевтическом применении, а также нет ответа на вопрос: какая из этих систем сочетает в себе наибольшую эффективность и безопасность?

Использование системы CRISPR/Cas9 имеет ряд преимуществ перед методами, основанными на ZFN и TALEN: ее значительно проще создать, она обладает более высокой эффективностью и подходит для высокопроизводительного и мультиплексного редактирования генома в самых разных клеточных линиях и живых организмах. Для переориентации на новую мишень нужно только поменять 20-нуклеотидную направляющую последовательность sgRNA. Причем Cas9 вносит разрыв строго между 17-м и 18-м нуклеотидами в целевой последовательности (считая от 5'-конца спейсера), т.е. на расстоянии трех нуклеотидов от PAM. А одновременное редактирование нескольких генов значительно упрощается введением комбинации sgRNA. Использование никазы и модификация конструкции sgRNA для более точного узнавания мишени в геноме позволяет избежать нежелательных нецелевых эффектов.

Система TALEN более трудоемкая, она требует больше времени на конструирование, чем CRISPR/Cas9. Однако сейчас существуют методы автоматизированного создания конструкций, экспрессирующих TALEN, что позволяет получать их эффективно в коммерческом масштабе. При этом тот факт, что TALEN вносят разрывы только при димеризации домена FokI, т.е. парами, увеличивает специфичность и снижает риск нецелевых эффектов.

На сегодняшний день не существует однозначного ответа на вопрос, какую из систем использовать. Необходимо детальное сравнение обеих систем, каждая из которых обладает своими особенностями. Не исключено, что универсальный ответ на этот вопрос так и не будет найден, и для каждого конкретного случая нужно будет пробовать различные варианты и выбирать из них наиболее соответствующие поставленным целям. ●

Работа поддержана междисциплинарным интеграционным проектом СО РАН № 55, РФФИ (грант № 12-04-00208-а).

СПИСОК ЛИТЕРАТУРЫ

1. Sapecchi M.R. // *Nat. Rev. Genet.* 2005. V. 6. № 6. P. 507–512.
2. Kim Y.G., Cha J., Chandrasegaran S. // *Proc. Natl. Acad. Sci. USA.* 1996. V. 93. № 3. P. 1156–1160.
3. Bibikova M., Golic M., Golic K.G., Carroll D. // *Genetics.* 2002. V. 161. № 3. P. 1169–1175.
4. Townsend J.A., Wright D.A., Winfrei R.J., Fu F., Maeder M.L., Joung J.K., Voytas D.F. // *Nature.* 2009. V. 459. № 7245. P. 442–445.
5. Zhang F., Maeder M.L., Unger-Wallace E., Hoshaw J.P., Reyon D., Christian M., Li X., Pierick C.J., Dobbs D., Peterson T., et al. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. № 26. P. 12028–12033.
6. Torikai H., Reik A., Liu P.Q., Zhou Y., Zhang L., Maiti S., Huls H., Miller J.C., Kebriaei P., Rabinovitch B., et al. // *Blood.* 2012. V. 119. № 24. P. 5697–5705.
7. Provasi E., Genovese P., Lombardo A., Magnani Z., Liu P.Q., Reik A., Chu V., Paschon D.E., Zhang L., Kuball J., et al. // *Nat. Med.* 2012. V. 18. № 5. P. 807–815.
8. Lombardo A., Cesana D., Genovese P., Di Stefano B., Provasi E., Colombo D.F., Neri M., Magnani Z., Cantore A., Lo Riso P., et al. // *Nat. Methods.* 2011. V. 8. № 10. P. 861–869.
9. Becker M. // *Nat. Methods.* 2012. V. 9. № 1. P. 1.
10. Schornack S., Meyer A., Römer P., Jordan T., Lahaye T. // *J. Plant Physiol.* 2006. V. 163. № 3. P. 256–272.
11. Romer P., Hahn S., Jordan T., Strauss T., Bonas U., Lahaye T. // *Science.* 2007. V. 318. № 5850. P. 645–648.
12. Boch J., Scholze H., Schornack S., Landgraf A., Hahn S., Kay S., Lahaye T., Nickstandt A., Bonas U., et al. // *Science.* 2009. V. 326. № 5959. P. 1509–1512.
13. Moscou M.J., Bogdanove A.J. // *Science.* 2009. V. 326. № 5959. P. 1501.
14. Lamb B.M., Mercer A.C., Barbas C.F. // *Nucleic Acids Res.* 2013. V. 41. № 21. P. 9779–9785.
15. Christian M., Cermak T., Doyle E.L., Schmidt C., Zhang F., Hummel A., Bogdanove A.J., Voytas D.F. // *Genetics.* 2010. V. 186. № 2. P. 757–761.
16. Cong L., Zhou R., Kuo Y.C., Cunniff M.M., Zhang F. // *Nat. Commun.* 2012. V. 3. P. 968.
17. Christian M.L., Demorest Z.L., Starker C.G., Osborn M.J., Nyquist M.D., Zhang Y., Carlson D.F., Bradley P., Bogdanove A.J., Voytas D.F. // *PLoS One.* 2012. V. 7. № 9. e45383.
18. Streubel J., Blucher C., Landgraf A., Boch J. // *Nat. Biotechnol.* 2012. V. 30. № 7. P. 593–595.
19. Mak A.N., Bradley P., Cernadas R.A., Bogdanove A.J., Stoddard B.L. // *Science.* 2012. V. 335. № 6069. P. 716–719.
20. Ishino Y., Shinagawa H., Makino K., Amemura M., Nakata A. // *J. Bacteriol.* 1987. V. 169. № 2. P. 5429–5433.
21. Haft D.H., Selengut J., Mongodin E.F., Nelson K.E. // *PLoS Comput. Biol.* 2005. V. 1. № 6. e60.
22. Bolotin A., Quinguis B., Sorokin A., Ehrlich S.D. // *Microbiology.* 2005. V. 151. P. 2551–2561.
23. Mojica F.J., Diez-Villasenor C., Garcia-Martinez J., Soria E. // *J. Mol. Evol.* 2005. V. 60. № 2. P. 174–182.
24. Pourcel C., Salvignol G., Vergnaud G. // *Microbiology.* 2005. V. 151. P. 653–663.
25. Barrangou R., Flemaux C., Deveau H., Richards M., Boyaval P., Moineau S., Romero D.A., Horvath P. // *Science.* 2007. V. 315. № 5819. P. 1709–1712.
26. Grissa I., Vergnaud G., Pourcel C. // *BMC Bioinformatics.* 2007. V. 8. P. 172.
27. Makarova K.S., Haft D.H., Barrangou R., Brouns S.J., Charpentier E., Horvath P., Moineau S., Mojica F.J., Wolf Y.I., Yakunin A.F., et al. // *Nat. Rev. Microbiol.* 2011. V. 9. № 6. P. 467–477.
28. Mojica F.J., Diez-Villasenor C., Garcia-Martinez J., Almendros C. // *Microbiology.* 2009. V. 155. P. 733–740.
29. Swarts D.C., Mosterd C., van Passel M.W., Brouns S.J. // *PLoS One.* 2012. V. 7. № 4. e35888.
30. Hale C.R., Majumdar S., Elmore J., Pfister N., Compton M., Olson S., Resch A.M., Glover C.V., Graveley B.R., Tern R.M. // *Mol. Cell.* 2012. V. 45. № 3. P. 292–302.
31. Lillestol R.K., Shah S.A., Brugger K., Redder P., Phan H., Christiansen J., Garrett R.A. // *Mol. Microbiol.* 2009. V. 72. № 1. P. 259–272.
32. Carte J., Wang R., Li H., Terns R.M., Terns M.P. // *Genes. Dev.* 2008. V. 22. № 24. P. 3489–3496.
33. Haurwitz R.E., Jinek M., Wiedenheft B., Zhou K., Doudna J.A. // *Science.* 2010. V. 329. № 5997. P. 1355–1358.
34. Gesner E.M., Schellenberg M.J., Garside E.L., George M.M., Macmillan A.M. // *Nat. Struct. Mol. Biol.* 2011. V. 18. № 6. P. 688–692.
35. Richter H., Zoepfel J., Schermuly J., Maticzka D., Backofen R., Randau L. // *Nucleic Acids Res.* 2012. V. 40. № 19. P. 9887–9896.
36. Sashital D.G., Jinek M., Doudna J.A. // *Nat. Struct. Mol. Biol.* 2011. V. 18. № 6. P. 680–687.
37. Hale C.R., Zhao P., Olson S., Duff M.O., Graveley B.R., Wells L., Terns R.M., Terns M.P. // *Cell.* 2009. V. 139. № 5. P. 945–956.
38. Makarova K.S., Grishin N.V., Shabalina S.A., Wolf Y.I., Koonin E.V. // *Biol. Direct.* 2006. V. 1. P. 7.
39. Marraffini L.A., Sontheimer E.J. // *Nature.* 2010. V. 463. № 7280. P. 568–571.
40. Bondy-Denomy J., Pawluk A., Maxwell K.L., Davidson A.R. // *Nature.* 2013. V. 493. № 7432. P. 429–432.
41. Makarova K.S., Aravind L., Wolf Y.I., Koonin E.V. // *Biol. Direct.* 2011. V. 6. P. 38.
42. Sapranaukas R., Gasiunas G., Fremaux C., Barrangou R., Horvath P., Siksnys V. // *Nucleic Acids Res.* 2011. V. 39. № 21. P. 9275–9282.
43. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. // *Science.* 2012. V. 337. № 6096. P. 816–821.
44. Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E. // *Nature.* 2011. V. 471. № 7340. P. 602–607.
45. Mussolino C., Cathomen T. // *Curr. Opin. Biotechnol.* 2012. V. 23. № 5. P. 644–650.
46. Cong L., Ran F.A., Cox D., Lin S., Barretto R., Habib N., Hsu P.D., Wu X., Jiang W., Marraffini L.A., et al. // *Science.* 2013. V. 339. № 6121. P. 819–823.
47. Hsu P.D., Scott D.A., Weinstein J.A., Ran F.A., Konermann S., Agarwala V., Li Y., Fine E.J., Wu X., Shalem O., et al. // *Nat. Biotechnol.* 2013. V. 31. № 9. P. 827–832.
48. Jiang W., Bikard D., Cox D., Zhang F., Marraffini L.A. // *Nat. Biotechnol.* 2013. V. 31. № 3. P. 233–239.
49. Fu Y., Foden J.A., Khayter C., Maeder M.L., Reyon D., Joung J.K., Sander J.D. // *Nat. Biotechnol.* 2013. V. 31. № 9. P. 822–826.
50. Mali P., Aach J., Stranges P.B., Esvelt K.M., Moosburner M., Kosuri S., Yang L., Church G.M. // *Nat. Biotechnol.* 2013. V. 31. № 9. P. 833–838.
51. Ran F.A., Hsu P.D., Lin C., Gootenberg J.S., Konermann S., Trevino A.E., Scott D.A., Inoue A., Matoba S., Zhang Y. // *Cell.* 2013. V. 154. № 6. P. 1380–1389.
52. Gasiunas G., Barrangou R., Horvath P., Siksnys V. // *Proc. Natl. Acad. Sci. USA.* 2012. V. 109. № 39. P. 2579–2586.
53. Dianov G.L., Hubscher U. // *Nucleic Acids Res.* 2013. V. 41. № 6. P. 3483–3490.
54. Doyle E.L., Booher N.J., Standage D.S., Voytas D.F., Brendel V.P., Vandyk J.K., Bogdanove A.J. // *Nucleic Acids Res.* 2012. V. 40. P. W117–122.

55. Fine E.J., Cradick T.J., Zhao C.L., Lin Y., Bao G. // *Nucleic Acids Res.* 2013. V. 42. № 6. e42.
56. Grau J., Boch J., Posch S. // *Bioinformatics.* 2013. V. 29. № 22. P. 2931–2932.
57. Grau J., Wolf A., Reschke M., Bonas U., Posch S. // *PLoS Comput. Biol.* 2013. V. 9. № 3. e1002962.
58. Heigwer F., Kerr G., Walther N., Glaeser K., Pelz O., Breinig M., Boutros M. // *Nucleic Acids Res.* 2013. V. 41. № 20. e190.
59. Neff K.L., Argue D.P., Ma A.C., Lee H.B., Clark K.J., Ekker S.C. // *BMC Bioinformatics.* 2013. V. 14. P. 1.
60. Grissa I., Vergnaud G., Pourcel C. // *Nucleic Acids Res.* 2007. V. 35. P. W52–57.
61. Rousseau C., Gonnet M., Le Romancer M., Nicolas J. // *Bioinformatics.* 2009. V. 25. № 24. P. 3317–3318.
62. Xiao A., Cheng Z., Kong L., Zhu Z., Lin S., Gao G., Zhang B. // *Bioinformatics.* 2014. V. 30. № 8. P. 1180–1182.
63. Sander J.D., Cade L., Khayter C., Reyon D., Peterson R.T., Joung J.K., Yeh J.R. // *Nat. Biotechnol.* 2011. V. 29. № 8. P. 697–698.
64. Reyon D., Khayter C., Regan M.R., Joung J.K., Sander J.D. // *Curr. Protoc. Mol. Biol.* 2012. C.12. U. 12.15.
65. Weber E., Engler C., Gruetznern R., Werner S., Marillonnet S. // *PLoS One.* 2011. V. 6. № 2. e16765.
66. Engler C., Gruetznern R., Kandzia R., Marillonnet S. // *PLoS One.* 2009. V. 4. № 5. e5553.
67. Zhang F., Cong L., Lodato S., Kosuri S., Church G.M., Arlotta P. // *Nat. Biotechnol.* 2011. V. 29. № 2. P. 149–153.
68. Sanjana N.E., Cong L., Zhou Y., Cunniff M.M., Feng G., Zhang F. // *Nat. Protoc.* 2012. V. 7. № 1. P. 171–192.
69. Schmid-Burgk J.L., Schmidt T., Kaiser V., Honing K., Hornung V. // *Nat. Biotechnol.* 2013. V. 31. № 1. P. 76–81.
70. Cermak T., Doyle E.L., Christian M., Wang L., Zhang Y., Schmidt C., Baller J.A., Somia N.V., Bogdanove A.J., Voytas D.F. // *Nucleic Acids Res.* 2011. V. 39. № 12. e82.
71. Ding Q., Lee Y.K., Schaefer E.A., Peters D.T., Veres A., Kim K., Kuperwasser N., Motola D.L., Meissner T.B., Hendriks W.T. // *Cell Stem Cell.* 2013. V. 12. № 2. P. 238–251.
72. Reyon D., Tsai S.Q., Khayter C., Foden J.A., Sander J.D., Joung J.K. // *Nat. Biotechnol.* 2012. V. 30. № 5. P. 460–465.
73. Briggs A.W., Rios X., Chari R., Yang L., Zhang F., Mali P., Church G.M. // *Nucleic Acids Res.* 2012. V. 40. № 15. e117.
74. Wang Z., Li J., Huang H., Wang G., Jiang M., Yin S., Sun C., Zhang H., Zhuang F., Xi J.J. // *Angew. Chem. Int. Ed. Engl.* 2012. V. 51. № 34. P. 8505–8508.
75. Jinek M., Doudna J.A. // *Nature.* 2009. V. 457. № 7228. P. 405–412.
76. Malone C.D., Hannon G.J. // *Cell.* 2009. V. 136. № 4. P. 656–668.
77. Meister G., Tuschl T. // *Nature.* 2004. V. 431. № 7006. P. 343–349.
78. Cho S.W., Kim S., Kim J.M., Kim J.S. // *Nat. Biotechnol.* 2013. V. 31. № 3. P. 230–232.
79. Jinek M., East A., Cheng A., Lin S., Ma E., Doudna J. // *Elife.* 2013. V. 2. e00471.
80. Mali P., Yang L., Esvelt K.M., Aach J., Guell M., DiCarlo J.E., Norville J.E., Church G.M. // *Science.* 2013. V. 339. № 6121. P. 823–826.
81. Chang N., Sun C., Gao L., Zhu D., Xu X., Zhu X., Xiong J.W., Xi J.J. // *Cell Res.* 2013. V. 23. № 4. P. 465–472.
82. Hwang W.Y., Fu Y., Reyon D., Maeder M.L., Tsai S.Q., Sander J.D., Peterson R.T., Yeh J.R., Joung J.K. // *Nat. Biotechnol.* 2013. V. 31. № 3. P. 227–229.
83. Wang H., Yang H., Shivalia S.C., Dawlaty M.M., Cheng W.A., Zhang F., Jaenisch R. // *Cell.* 2013. V. 153. № 4. P. 910–918.
84. Shalem O., Sanjana N.E., Hartenian E., Shi X., Scott D.A., Mikkelsen T.S., Heckl D., Ebert B.L., Root D.E., Doench J.G. // *Science.* 2014. V. 343. № 6166. P. 84–87.
85. Wang T., Wei J.J., Sabatini D.M., Lander E.S. // *Science.* 2014. V. 343. № 6166. P. 80–84.
86. Shan Q., Wang Y., Li J., Zhang Y., Chen K., Liang Z., Zhang K., Liu J., Xi J.J., Qiu J.L. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 686–688.
87. Mao Y., Zhang H., Xu N., Zhang B., Gao F., Zhu J.K. // *Mol. Plant.* 2013. V. 6. № 6. P. 2008–2011.
88. Li J.F., Norville J.E., Aach J., McCormack M., Zhang D., Bush J., Church G.M., Sheen J. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 688–691.
89. Nekrasov V., Staskawicz B., Weigel D., Jones J.D., Kamoun S. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 691–693.
90. Lieber M.R. // *Annu. Rev. Biochem.* 2010. V. 79. P. 181–211.
91. Ousterout D.G., Perez-Pinera P., Thakore P.I., Kabadi A.M., Brown M.T., Qin X., Fedrigo O., Mouly V., Tremblay J.P., Gersbach C.A. // *Mol. Ther.* 2013. V. 21. № 9. P. 1718–1726.
92. Kim H., Um E., Cho S.R., Jung C., Kim H., Kim J.S. // *Nat. Methods.* 2011. V. 8. № 11. P. 941–943.
93. Guschin D.Y., Waite A.J., Katibah G.E., Miller J.C., Holmes M.C., Rebar E.J. // *Methods Mol. Biol.* 2010. V. 649. P. 247–256.
94. Dahlem T.J., Hoshijima K., Juryne M.J., Gunther D., Starker C.G., Locke A.S., Weis A.M., Voytas D.F., Grunwald D.J. // *PLoS Genet.* 2012. V. 8. № 8. e1002861.
95. Ota S., Hisano Y., Muraki M., Hoshijima K., Dahlem T.J., Grunwald D.J., Kawahara A. // *Genes Cells.* 2013. V. 18. № 6. P. 450–458.
96. Kim H.J., Lee H.J., Kim H., Cho S.W., Kim J.S. // *Genome Res.* 2009. V. 19. № 7. P. 1279–1288.
97. Hisano Y., Ota S., Arakawa K., Muraki M., Kono N., Oshita K., Sakuma T., Tomita M., Yamamoto T., Okada Y. // *Biol. Open.* 2013. V. 2. № 4. P. 363–367.
98. Yang L., Guell M., Byrne S., Yang J.L., De Los Angeles A., Mali P., Aach J., Kim-Kiselak C., Briggs A.W., Rios X. // *Nucleic Acids Res.* 2013. V. 41. № 19. P. 9049–9061.
99. Saleh-Gohari N., Helleday T. // *Nucleic Acids Res.* 2004. V. 32. № 12. P. 3683–3688.
100. 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. // *Nat. Biotechnol.* 2008. V. 26. № 7. P. 808–816.
101. Cristea S., Freyvert Y., Santiago Y., Holmes M.C., Urnov F.D., Gregory P.D., Cost G.J. // *Biotechnol. Bioeng.* 2013. V. 110. № 3. P. 871–880.
102. Maresca M., Lin V.G., Guo N., Yang Y. // *Genome Res.* 2013. V. 23. № 3. P. 539–546.
103. Chen F., Pruett-Miller S.M., Huang Y., Gjoka M., Taunton J., Collinwood T.N., Frodin M., Davis G.D. // *Nat. Methods.* 2011. V. 8. № 9. P. 753–755.
104. Moehle E.A., Rock J.M., Lee Y.L., Jouvenot Y., DeKolver R.C., Gregory P.D., Urnov F.D., Holmes M.C. // *Proc. Natl. Acad. Sci. USA.* 2007. V. 104. № 9. P. 3055–3060.
105. 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. // *Nucleic Acids Res.* 2010. V. 38. № 15. e152.
106. Hu R., Wallace J., Dahlem T.J., Grunwald D.J., O'Connell R.M. // *PLoS One.* 2013. V. 8. № 5. e63074.
107. Bloom K., Ely A., Mussolino C., Cathomen T., Arbuthnot P. // *Mol. Ther.* 2013. V. 21. № 10. P. 1889–1897.
108. Schiffer J.T., Aubert M., Weber N.D., Mintzer E., Stone D., Jerome K.R. // *J. Virol.* 2012. V. 86. № 17. P. 8920–8936.
109. Holt N., Wang J., Kim K., Friedman G., Wang X., Taupin V., Crooks G.M., Kohn D.B., Gregory P.D., Holmes M.C. // *Nat. Biotechnol.* 2010. V. 28. № 8. P. 839–847.
110. Horii T., Tamura D., Morita S., Kimura M., Hatada I. // *Int. J. Mol. Sci.* 2013. V. 14. № 10. P. 19774–19781.

111. Schwank G., Koo B.K., Sasselli V., Dekkers J.F., Heo I., Demircan T., Sasaki N., Boymans S., Cuppen E., van der Ent C.K. // *Cell Stem Cell*. 2013. V. 13. № 6. P. 653–658.
112. Gilbert L.A., Larson M.H., Morsut L., Liu Z., Brar G.A., Torres S.E., Stern-Ginossar N., Brandman O., Whitehead E.H., Doudna J.A. // *Cell*. 2013. V. 154. № 2. P. 442–451.
113. Bikard D., Jiang W., Samai P., Hochschild A., Zhang F., Marraffini L.A. // *Nucleic Acids Res.* 2013. V. 41. № 15. P. 7429–7437.
114. Perez-Pinera P., Kocak D.D., Vockley C.M., Adler A.F., Kabadi A.M., Polstein L.R., Thakore C.A. // *Nat. Methods*. 2013. V. 10. № 10. P. 973–976.
115. Farzadfard F., Perli S.D., Lu T.K. // *ACS Synth. Biol.* 2013. V. 2. № 10. P. 604–613.
116. Heintze J., Luft C., Ketteler R. // *Front Genet.* 2013. V. 4. P. 193.
117. Chen B., Gilbert L.A., Cimini B.A., Schnitzbauer J., Zhang W., Li G.W., Park J., Blackburn E.H., Weissman J.S., Qi L.S., et al. // *Cell*. 2013. V. 155. № 7. P. 1479–1491.
118. Smith J.R., Maguire S., Davis L.A., Alexander M., Yang F., Chandran S., ffrench-Constant C., Pedersen R.A. // *Stem Cells*. 2008. V. 26. № 2. P. 496–504.
119. Chiu H., Schwartz H.T., Antoshechkin I., Sternberg P.W. // *Genetics*. 2013. V. 195. № 3. P. 1167–1171.
120. Cho S.W., Lee J., Carroll D., Kim J., Lee J. // *Genetics*. 2013. V. 195. № 3. P. 1177–1180.
121. Katic I., Grosshans H. // *Genetics*. 2013. V. 195. № 3. P. 1173–1176.
122. Lo T.W., Pickle C.S., Lin S., Ralston E.J., Gurling M., Schartner C.M., Bian Q., Doudna J.A., Meyer B.J. // *Genetics*. 2013. V. 195. № 2. P. 331–348.
123. Tzur Y.B., Friedland A.E., Nadarajan S., Church G.M., Calarco J.A., Colaiácovo M.P. // *Genetics*. 2013. V. 195. № 3. P. 1181–1185.
124. Waaijers S., Portegijs V., Kerver J., Lemmens B.B., Tijsterman M., van den Heuvel S., Boxem M. // *Genetics*. 2013. V. 195. № 3. P. 11887–11891.
125. Dickinson D.J., Ward J.D., Reiner D.J., Goldstein B. // *Nat. Methods*. 2013. V. 10. № 10. P. 1028–1034.
126. Friedland A.E., Tzur Y.B., Esvelt K.M., Colaiácovo M.P., Church G.M., Calarco J.A. // *Nat. Methods*. 2013. V. 10. № 8. P. 741–743.
127. Maggert K.A., Gong W.J., Golic K.G. // *Methods Mol. Biol.* 2008. V. 420. P. 155–174.
128. Rong Y.S., Golic K.G. // *Science*. 2000. V. 288. № 5473. P. 2013–2018.
129. Venken K.J., Bellen H.J. // *Nat. Rev. Genet.* 2005. V. 6. № 3. P. 167–178.
130. Bassett A.R., Tibbit C., Ponting C.P., Liu J. // *Cell Rep.* 2013. V. 4. № 1. P. 220–228.
131. Gratz S.J., Cummings A.M., Nguyen J.N., Hamm D.C., Donohue L.K., Harrison M.M., Wildonger J., O'Connor-Giles K.M. // *Genetics*. 2013. V. 194. № 4. P. 1029–1035.
132. Yu Z., Ren M., Wang Z., Zhang B., Rong Y.S., Jiao R., Gao G. // *Genetics*. 2013. V. 195. № 1. P. 289–291.
133. Ren X., Sun J., Housden B.E., Hu Y., Roesel C., Liu L.P., Yang Z., Mao D., Sun L., Wu Q., et al. // *Proc. Natl. Acad. Sci. USA*. 2013. V. 110. № 47. P. 19012–19017.
134. Seth A., Stemple D.L., Barroso I. // *Dis. Model. Mech.* 2013. V. 6. № 5. P. 1080–1088.
135. Copeland N.G., Jenkins N.A. // *Nat. Rev. Cancer*. 2010. V. 10. № 10. P. 696–706.
136. Kool J., Berns A. // *Nat. Rev. Cancer*. 2009. V. 9. № 6. P. 389–399.
137. Bogdanove A.J., Voytas D.F. // *Science*. 2011. V. 333. № 6051. P. 1843–1846.
138. Carroll D., Beumer K.J., Morton J.J., Bozas A., Trautman J.K. // *Methods Mol. Biol.* 2008. V. 435. P. 63–77.
139. Urnov F.D., Rebar E.J., Holmes M.C., Zhang H.S., Gregory P.D. // *Nat. Rev. Genet.* 2010. V. 11. № 9. P. 636–646.
140. Carbery I.D., Ji D., Harrington A., Brown V., Weinstein E.J., Liaw L., Cui X. // *Genetics*. 2010. V. 186. № 2. P. 451–459.
141. Geurts A.M., Cost G.J., Freyvert Y., Zeitler B., Miller J.C., Choi V.M., Jenkins S.S., Wood A., Cui X., Meng X., et al. // *Science*. 2009. V. 325. № 5939. P. 433.
142. Cui X., Ji D., Fisher D.A., Wu Y., Briner D.M., Weinstein E.J. // *Nat. Biotechnol.* 2011. V. 29. № 1. P. 64–67.
143. Meyer M., de Angelis M.H., Wurst W., Kuhn R. // *Proc. Natl. Acad. Sci. USA*. 2010. V. 107. № 34. P. 15022–15026.
144. Li D., Qiu Z., Shao Y., Guan Y., Liu M., Li Y., Gao N., Wang L., Lu X., Zao Y., et al. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 681–683.
145. Li W., Teng F., Li T., Zhou Q. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 684–686.
146. Wu Y., Liang D., Wang Y., Bai M., Tang W., Bao S., Yan Z., Li D., Li J. // *Cell Stem Cell*. 2013. V. 13. № 6. P. 659–662.
147. Chen K., Gao C. // *J. Genet. Genomics*. 2013. V. 40. № 6. P. 271–279.
148. van der Hoorn R.A., Laurent F., Roth R., De Wit P.J. // *Mol. Plant. Microbe Interact.* 2000. V. 13. № 4. P. 439–446.
149. Feng Z., Zhang B., Ding W., Liu X., Yang D., Wei P., Cao F., Zhu S., Zhang F., Mao Y., et al. // *Cell Res*. 2013. V. 23. № 10. P. 1229–1232.
150. Xie K., Yang Y. // *Mol. Plant*. 2013. V. 6. № 6. P. 1975–1983.
151. Miao J., Guo D., Zhang J., Huang Q., Qin G., Zhang X., Wan J., Gu H., Qu L. // *Cell Res*. 2013. V. 23. № 10. P. 1233–1236.
152. Jiang W., Zhou H., Bi H., Fromm M., Yang B., Weeks D.P. // *Nucleic Acids Res.* 2013. V. 41. № 20. e188.
153. Upadhyay S.K., Kumar J., Alok A., Tuli R. // *G3 (Bethesda)*. 2013. V. 3. № 12. P. 2233–2238.
154. Shan Q., Wang Y., Chen K., Liang Z., Li J., Zhang Y., Zhang K., Liu J., Voytas D.F., Zheng X. // *Mol. Plant*. 2013. V. 6. № 4. P. 1365–1368.
155. Li T., Liu B., Spalding M.H., Weeks D.P., Yang B. // *Nat. Biotechnol.* 2012. V. 30. № 5. P. 390–392.
156. Zhang Y., Zhang F., Li X., Baller J.A., Starker C.G., Bogdanove A.J., Voytas D.F. // *Plant Physiol*. 2013. V. 161. № 1. P. 20–27.
157. Mahfouz M.M., Li L., Shamimuzzaman M., Wibowo A., Fang X., Zhu J.K. // *Proc. Natl. Acad. Sci. USA*. 2011. V. 108. № 6. P. 2623–2628.
158. Beerli R.R., Segal D.J., Dreier B., Barbas C.F. 3rd. // *Proc. Natl. Acad. Sci. USA*. 1998. V. 95. № 25. P. 14628–14633.
159. Perez-Pinera P., Ousterout D.G., Brunger J.M., Farin A.M., Guilak F., Crawford G.E., Hartemink A.J., Gersbach C.A. // *Nat. Methods*. 2013. V. 10. № 3. P. 239–242.
160. Maeder M.L., Linder S.J., Reyon D., Angstman J.F., Fu Y., Sander J.D., Joung J.K. // *Nat. Methods*. 2013. V. 10. № 3. P. 243–245.
161. Li Y., Moore R., Guinn M., Bleris L. // *Sci. Rep.* 2012. V. 2. P. 897.
162. Mahfouz M.M., Li L., Piatek M., Fang X., Mansour H., Bangarusamy D.K., Zhu J.K. // *Plant Mol. Biol.* 2012. V. 78. № 3. P. 311–321.
163. Chapdelaine P., Coulombe Z., Chikh A., Gerard C., Tremblay J.P. // *Mol. Ther. Nucleic Acids*. 2013. V. 2. e119.
164. Gao X., Yang J., Tsang J.C., Ooi J., Wu D., Liu P. // *Stem Cell Reports*. 2013. V. 1. № 2. P. 183–197.
165. Mercer A.C., Gaj T., Sirk S.J., Lamb B.M., Barbas C.F. 3rd. // *ACS Synth. Biol.* 2013.

166. Konermann S., Brigham M.D., Trevino A.E., Hsu P.D., Heidenreich M., Cong L., Platt R.J., Scott D.A., Church G.M., Zhang F. // *Nature*. 2013. V. 500. № 7463. P. 472–476.
167. Gaber R., Lebar T., Majerle A., Ster B., Dobnikar A., Bencina M., Jerala R. // *Nat. Chem. Biol.* 2014. V. 10. № 3. P. 203–208.
168. Lienert F., Torella J.P., Chen J.H., Norsworthy M., Richardson R.R., Silver P.A. // *Nucleic Acids Res.* 2013. V. 41. № 21. P. 9967–9975.
169. Miyanari Y., Ziegler-Birling C., Torres-Padilla M.E. // *Nat. Struct. Mol. Biol.* 2013. V. 20. № 11. P. 1321–1324.
170. Thanisch K., Schneider K., Morbitzer R., Solovei I., Lahaye T., Bultmann S., Leonhardt H. // *Nucleic Acids Res.* 2013. V. 42. № 6. e38.
171. Ma H., Reyes-Gutierrez P., Pederson T. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 52. P. 21048–21053.
172. Owens J.B., Urschitz J., Stoytchev I., Dang N.C., Stoytcheva Z., Belcaid M., Maragathavally K.J., Coates C.J., Segal D.J., Moisyadi S. // *Nucleic Acids Res.* 2012. V. 40. № 14. P. 6978–6991.
173. Mercer A.C., Gaj T., Fuller R.P., Barbas C.F., 3rd. // *Nucleic Acids Res.* 2012. V. 40. № 21. P. 11163–11172.
174. Lamb B.M., Mercer A.C., Barbas C.F., 3rd. // *Nucleic Acids Res.* 2013. V. 41. № 21. P. 9779–9785.
175. Owens J.B., Mauro D., Stoytchev I., Bhakta M.S., Kim M.S., Segal D.J., Moisyadi S. // *Nucleic Acids Res.* 2013. V. 41. № 19. P. 9197–9207.
176. 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., et al. // *Nat. Biotechnol.* 2011. V. 29. № 2. P. 143–148.
177. Mussolino C., Morbitzer R., Lutge F., Dannemann N., Lahaye T., Cathomen T. // *Nucleic Acids Res.* 2011. V. 39. № 21. P. 9283–9293.
178. Sun N., Liang J., Abil Z., Zhao H. // *Mol. Biosyst.* 2012. V. 8. № 4. P. 1255–1263.
179. Xu L., Zhao P., Mariano A., Han R. // *Mol. Ther. Nucleic Acids.* 2013. V. 2. e112.
180. Hockemeyer D., Wang H., Kiani S., Lai C.S., Gao Q., Cassady J.P., Cost G.J., Zhang L., Santiago Y., Miller J.C., et al. // *Nat. Biotechnol.* 2011. V. 29. № 8. P. 731–734.
181. Li T., Huang S., Zhao X., Wright D.A., Carpenter S., Spalding M.H., Weeks D.P., Yang B. // *Nucleic Acids Res.* 2011. V. 39. № 14. P. 6315–6325.
182. Wood A.J., Lo T.W., Zeitler B., Pickle C.S., Ralston E.J., Lee A.H., Amora R., Miller J.C., Leung E., Meng X., et al. // *Science*. 2011. V. 333. № 6040. P. 307.
183. Liu J., Li C., Yu Z., Huang P., Wu H., Wei C., Zhu N., Chen Y., Zhang B., Deng W.M., et al. // *J. Genet. Genomics.* 2012. V. 39. № 5. P. 209–215.
184. Bedell V.M., Wang Y., Campbell J.M., Poshusta T.L., Starker C.G., Krug R.G. 2nd, Tan W., Penheiter S.G., Ma A.C., Leung A.Y., et al. // *Nature*. 2012. V. 491. № 7422. P. 114–118.
185. 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. // *PLoS One*. 2012. P. 7. № 5. e37877.
186. Huang P., Xiao A., Zhou M., Zhu Z., Lin S., Zhang B. // *Nat. Biotechnol.* 2011. V. 29. № 8. P. 699–700.
187. Cade L., Reyon D., Hwang W.Y., Tsai S.Q., Patel S., Khayter C., Joung J.K., Sander J.D., Peterson R.T., Yeh J.R. // *Nucleic Acids Res.* 2012. V. 40. № 16. P. 8001–8010.
188. Ma S., Zhang S., Wang F., Liu Y., Liu Y., Xu H., Liu C., Lin Y., Zhao P., Xia Q. // *PLoS One*. 2012. V. 7. № 9. e45035.
189. Watanabe T., Ochiai H., Sakuma T., Horch H.W., Hamaguchi N., Nakamura T., Bando T., Ohuchi H., Yamamoto T., Noji S., et al. // *Nat. Commun.* 2012. V. 3. P. 1017.
190. Lei Y., Guo X., Liu Y., Cao Y., Deng Y., Chen X., Cheng X., Cheng C.H., Dawid I.B., Chen Y., et al. // *Proc. Natl. Acad. Sci. USA.* 2012. V. 109. № 43. P. 17484–17489.
191. Panda S.K., Wefers B., Ortiz O., Floss T., Schmid B., Haass C., Wurst W., Kuhn R. // *Genetics*. 2013. V. 195. № 3. P. 703–713.
192. Qiu Z., Liu M., Chen Z., Shao Y., Pan H., Wei G., Yu C., Zhang L., Li X., Wang P., et al. // *Nucleic Acids Res.* 2013. V. 41. № 11. e120.
193. Wu J., Huang Z., Ren J., Zhang Z., He P., Li Y., Ma J., Chen W., Zhang Y., Zhou X., et al. // *Cell Res.* 2013. V. 23. № 8. P. 994–1006.
194. Sung Y.H., Baek I.J., Kim D.H., Jeon J., Lee J., Lee K., Jeong D., Kim J.S., Lee H.W. // *Nat. Biotechnol.* 2013. V. 31. № 1. P. 23–24.
195. Wefers B., Meyer M., Ortiz O., Hrabe de Angelis M., Hansen J., Wurst W., Kuhn R. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 10. P. 3782–3787.
196. Davies B., Davies G., Preece C., Puliyadi R., Szumska D., Bhattacharya S. // *PLoS One*. 2013. V. 8. № 3. e60216.
197. Tong C., Huang G., Ashton C., Wu H., Yan H., Ying Q.L. // *J. Genet. Genomics.* 2012. V. 39. № 6. P. 275–280.
198. Tesson L., Usal C., Menoret S., Leung E., Niles B.J., Remy S., Santiago Y., Vincent A.I., Meng X., Zhang L., et al. // *Nat. Biotechnol.* 2011. V. 29. № 8. P. 695–696.
199. Carlson D.F., Tan W., Lillico S.G., Stverakova D., Proudfoot C., Christian M., Voytas D.F., Long C.R., Whitelaw C.B., Fahrenkrug S.C. // *Proc. Natl. Acad. Sci. USA.* 2012. V. 109. № 43. P. 17382–17387.
200. DiCarlo J.E., Norville J.E., Mali P., Rios X., Aach J., Church G.M. // *Nucleic Acids Res.* 2013. V. 41. № 7. P. 4336–4343.
201. Hou Z., Zhang Y., Propson N.E., Howden S.E., Chu L., Sontheimer E.J., Thomson J.A. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 39. P. 15644–15649.
202. Ding Q., Regan S.N., Xia Y., Oostrom L.A., Cowan C.A., Musunuru K. // *Cell Stem Cell.* 2013. V. 12. № 4. P. 393–394.
203. Hwang W.Y., Fu Y., Reyon D., Maeder M.L., Kaini P., Sander J.D., Joung J.K., Peterson R.T., Yeh J.J. // *PLoS One*. 2013. V. 8. № 7. e68708.
204. Jao L.E., Wente S.R., Chen W. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 34. P. 13904–13909.
205. Xiao A., Wang Z., Hu Y., Wu Y., Luo Z., Yang Z., Zu Y., Li W., Huang P., Tong X., et al. // *Nucleic Acids Res.* 2013. V. 41. № 14. e141.
206. Hruscha A., Krawitz P., Rechenberg A., Heinrich V., Hecht J., Haass C., Schmid B. // *Development*. 2013. V. 140. № 24. P. 4982–4987.
207. Nakayama T., Fish M.B., Fisher M., Oomen-Hajagos J., Thomsen G.H., Grainger R.M. // *Genesis*. 2013. V. 51. № 12. P. 835–843.
208. Tan W., Carlson D.F., Lancto C.A., Garbe J.R., Webster D.A., Hackett P.B., Fahrenkrug S.C. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 41. P. 16526–16531.
209. Fujii W., Kawasaki K., Sugiura K., Naito K. // *Nucleic Acids Res.* 2013. V. 41. № 20. e187.
210. Shen B., Zhang J., Wu H., Wang J., Ma K., Li Z., Zhang X., Zhang P., Huang X. // *Cell Res.* 2013. V. 23. № 5. P. 720–723.
211. Ma Y., Zhang X., Shen B., Lu Y., Chen W., Ma J., Bai L., Huang X., Zhang L. // *Cell Res.* 2014. V. 24. № 1. P. 122–125.