

А.В. Москалев¹, Б.Ю. Гумилевский¹,
В.Я. Апчел^{1,2}, В.Н. Цыган¹

Методы изучения генетических модификаций

¹Военно-медицинская академия им. С.М. Кирова, Санкт-Петербург

²Российский государственный педагогический университет им. А.И. Герцена, Санкт-Петербург

Резюме. Рассмотрены объекты и современные методы редактирования генома. Охарактеризована иммунная система прокариот и их защитные механизмы, препятствующие целенаправленному редактированию генома в интересах исследователя. Таким механизмом у прокариот являются кластерные регулятивные межпространственные короткие палиндромные повторы. Число таких повторов у различных объектов отличается, что в итоге не позволяет получить идеальную стандартную модель. В настоящее время идентифицированы три типа таких систем, имеющих свой механизм для генерации белков. Охарактеризованы белки, которые в настоящее время чаще всего используются для редактирования генома и выявления участков протоспейсерных соседних мотивов. Дана подробная характеристика организации иммунной системы прокариот и фаз ее активности. На сегодняшний день идентифицированы три типа систем с короткими палиндромными повторами, расположенными группами или ассоциированным (ферментным) белком Cas9. Каждая система использует свой механизм для генерации белков, катализирующих расщепление нуклеиновых кислот. Чаще всего используется система с короткими палиндромными повторами II типа, лучше адаптированная для редактирования генома из-за своей простоты. Установлено, что систему с короткими палиндромными повторами-Cas9 можно использовать для точечного редактирования генома и у эукариот. Это осуществляется либо посредством негомологичного присоединения конца, либо путем гомологически направленной репарации. Перспективным вариантом генетического моделирования является использование фермента-эндонуклеазы Cpf1, являющейся эффекторным белком систем с короткими палиндромными повторами-Cas V типа. Cpf1 мельче, чем ферментный белок Cas9, и для функционирования системы нужны только спейсеры рибонуклеиновой кислоты без дополнительной рибонуклеиновой кислоты. В отличие от Cas9, который разрезает обе цепи дезоксирибонуклеиновой кислоты в одном и том же месте, Cpf1 генерирует разрез, создавая «липкие» концы, которые можно использовать для вставки интересующих последовательностей путем комплементации и лигирования. Вероятно, что система с использованием фермента-эндонуклеазы Cpf1 будет удобнее системы, где используется белок – Cas9, так как расширяется диапазон редактирования управляемого генома рибонуклеиновой кислоты для осуществления необходимых правок.

Ключевые слова: стволовая клетка, факторы транскрипции, гены, фенотип, клеточная дифференцировка, модификации, мутации, хромосома, плазмиды, нуклеиновые кислоты, транслокация.

В настоящее время для редактирования генов эукариот используются адаптированные иммунные системы бактерий и архей. При этом необходимо учитывать, что прокариоты наделены защитными механизмами против внедрения чужеродных вирусных и плазмидных агентов, как и многоклеточные организмы. Одним из таких защитных механизмов является адаптивная иммунная система, обнаруженная у многих бактерий и большинства архей, характеризующаяся кластерными регулятивными межпространственными короткими палиндромными повторами (clustered regularly interspaced short palindromic repeats – CRISPR) или ассоциированными белками Cas (ферментный белок Cas9 разрезает двойную спираль вирусной дезоксирибонуклеиновой кислоты (ДНК), то есть уничтожает ее, так клетка защищается от инфекции). При интегрировании последовательностей ДНК в геном бактерии и архей генерируют иммунологическую память о предыдущих чужеродных агентах. Эти приобретенные механизмы позволяют бактериям, археям распознавать вирусные и плазмидные агенты как чужеродные, что приводит в итоге к деградациии инвазивных последовательностей и функционированию в качестве адаптивной иммунной системы для прокариот [1, 4].

Система CRISPR-Cas – 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 [6]. 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 последовательностей нуклеотидов), а также число и размер участков некодирующей ДНК, расположенных между тандемно повторяющимися генами (спейсеров) (от 17 до 84 последовательностей нуклеотидов). При этом длина повторов и спейсеров внутри одной кассеты неизменна, а последовательности повторов практически идентичны [11]. Иммунитет CRISPR характеризуется различными фазами. В фазе адаптации бактерии, археи приобретают иммунологическую память о внедряющихся вирусе или плазмиде. Короткие последовательности вирусных или плазмидных геномов интегрированы в локус CRISPR бактериального или археального генома (рис. 1а).

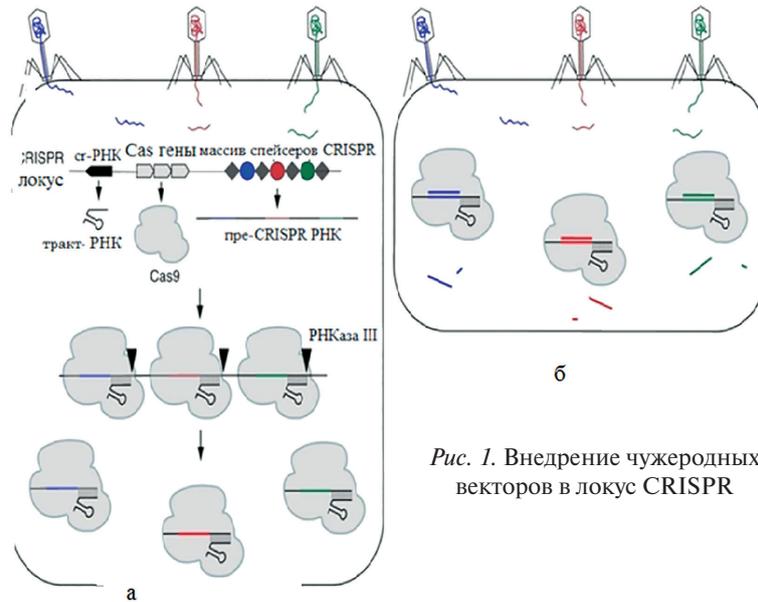


Рис. 1. Внедрение чужеродных векторов в locus CRISPR

Впервые локусы CRISPR были идентифицированы при изучении ферментов промышленного назначения, где прокариоты необходимы для производства ферментированных продуктов. Посредством сравнительного геномного анализа различных штаммов *S. thermophilus* (микробы, используемые при производстве йогурта) был установлен высоко-изменчивый locus в геноме этих бактерий. Этот locus имеет две особенности: множество несмежных повторов, разделенных переменными последовательностями, спейсерами. Впоследствии было установлено, что спейсерные последовательности совпадают с последовательностями, обнаруженными в геномах [15]. При сравнении устойчивых и чувствительных к фагам *S. thermophilus* оказалось, что устойчивые к фагам бактерии имели спейсерные последовательности, которые соответствовали участкам генома этого фага [18]. Так как содержание спейсера коррелирует с устойчивостью к фагам, это приводит к формированию модели, в которой короткие области внедряемого генома интегрируются в локусы CRISPR в виде спейсеров, разделенных повторяющимися последовательностями, что приводит к формированию клеточной памяти о перенесенных инфекциях. Устойчивые к фагам бактерии имели спейсерные последовательности, которые соответствовали областям генома этого фага [19].

После приобретения спейсерами рибонуклеиновой кислоты (PHK) PHK CRISPR (crRNA) генерируется из спейсеров в локусе CRISPR и внедряется в белок Cas. crRNA направляет белок Cas для распознавания чужеродных последовательностей и расщепления внедренной фаговой или плазмидной ДНК (см. рис. 1б). В настоящее время идентифицированы три типа систем CRISPR-Cas9. Каждая система использует свой механизм для генерации белков crRNA и Cas, которые катализируют расщепление нуклеиновой кислоты [19]. На практике чаще всего используется система CRISPR типа II, которая гораздо лучше адап-

тирована для редактирования генома из-за своей простоты. Для генерации crPHK locus CRISPR транскрибируется, генерируя длинную молекулу PHK с последовательностями, гомологичными предыдущим агентам. Эта молекула PHK называется пре-crRNA (см. рис. 1а). Вторая PHK из геномного локуса выше локуса CRISPR также транскрибируется. Эта PHK называется транс-активирующей PHK CRISPR (tracrRNA) [11]. TracrRNA имеет область, которая комплементарна области повторов локуса CRISPR, и связывается с недавно транскрибированной пре-crRNA, создавая двухцепочечную PHK, которая расщепляется PHK-sellIII (ферментом, который распознает и разрезает двухцепочечную PHK), что приводит к формированию комплекса crRNA : tracrRNA, содержащего только одну спейсерную последовательность (см. рис. 1б). Этот PHK-комплекс затем ассоциируется с одним белком Cas9 (эндонуклеазой), создавая активный рибонуклеопротеиновый (RNP) комплекс [16].

Как только crRNA : tracrRNA связывается с Cas9, он активируется и может расщеплять инвазивные последовательности нуклеиновых кислот (интерференция) (см. рис. 1б). Cas9 – эндонуклеаза, управляемая PHK, – расщепляет ДНК в последовательностях, которые связываются с crRNA. Поиск в инвазивной ДНК последовательностей, комплементарных crRNA, происходит посредством связывания Cas9 с последовательностями в инвазивном вирусном или плазмидном геноме, называемом прото-спейсерными соседними мотивами (protospacer adjacent motif – PAMs) [7]. Белки Cas9 разных видов бактерий или архей распознают сайты PAM. На сегодняшний день для редактирования генома наиболее часто используется Cas9 *S. pyogenes* (SpCas9), распознающий PAM 50-NGG-30 (см. рис. 2а).

После изучения микробной иммунной системы CRISPR-Cas9 было установлено, что ее можно использовать для точечного редактирования генома у

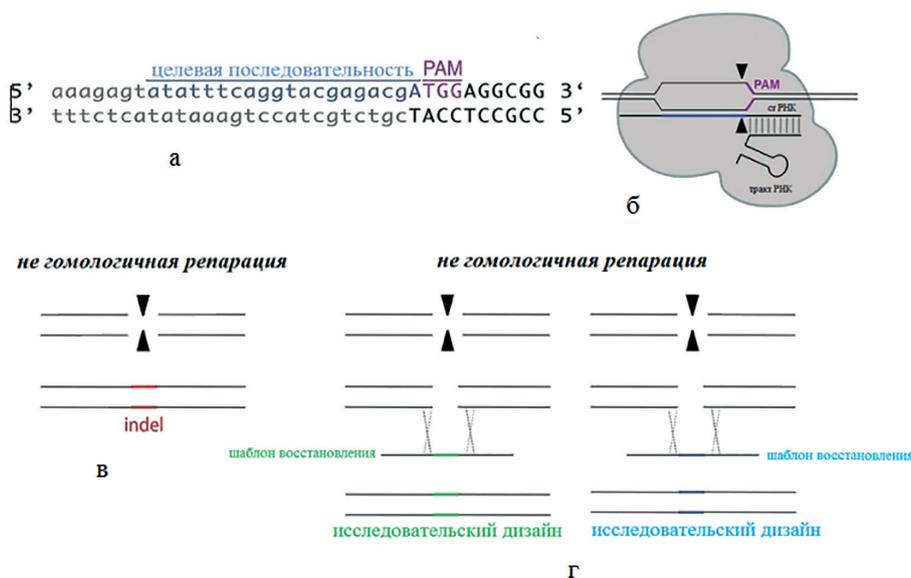


Рис. 2. Редактирование генома белком Cas9 *S. pyogenes*

эукариот. В ответ на вызванные Cas9 двухцепочечные разрывы клетки используют один из двух путей для восстановления повреждения ДНК. Это осуществляется либо посредством негомологичного присоединения конца (non-homologous end joining – NHEJ), либо путем гомологически направленной репарации (homology directed repair – HDR) (рис. 2) [3]. NHEJ может осуществляться через канонический NHEJ (C-NHEJ), который лигирует или «склеивает» поврежденные концы. Кроме того, существует альтернативный путь присоединения конца (alt-NHEJ), в котором одна нить ДНК по обе стороны от разрыва резецируется для восстановления очага поражения [12]. Однако обоим методам восстановления присущи ошибки, приводящие к неполному восстановлению дефектов (см. рис. 2б). Но если рядом находится молекула ДНК гомологичной области вокруг двухцепочечного разрыва, то она может использоваться в качестве матрицы для восстановления разрыва путем HDR. Обычно этот репарационный механизм происходит до деления клетки после репликации ДНК. Поэтому разрыв можно восстановить с помощью недавно реплицированной сестринской хроматиды без каких-либо мутаций [5]. Эта форма восстановления может быть использована для точных правок или введения больших вставок или удалений путем введения донорского шаблона для восстановления (см. рис. 2г).

Таким образом, при выполнении разреза в определенном локусе и использовании преимуществ путей репарации клеточной ДНК, существует возможность для генерации целевых мутаций и вставки последовательностей, представляющих интерес. Однако создание двухцепочечных разрывов в конкретных участках генома – это непростая задача из-за сложности направления нуклеаз ДНК в специфические последовательности. Cas9 может быть легко нацелен уникальной crRNA для разрезания в любом желаемом сайте. Поскольку сайт PAM необходим для связывания

Cas9, мишень должна быть выше сайта 50-NGG-30 (в случае SpCas9) (см. рис. 2а). Таким образом, до тех пор, пока известна последовательность целевого гена, Cas9 может быть нацелен практически на любой сайт с учетом присутствия ближайшего PAM (50-NGG-30).

Два критических остатка аргинина в SpCas9, Arg1333 и Arg1335 взаимодействуют с гуаниновыми нуклеосооснованиями PAM на некомплементарной цепи [17]. Это взаимодействие между гуанинами PAM и аргининами в SpCas9 обеспечивает фосфат основной цепи ДНК 5' в PAM, чтобы взаимодействовать с петлей фосфат-блокировки в Cas9, и облегчит разматывание нити ДНК [16]. Если ДНК дополняет направляющую РНК, образуется гибрид РНК – ДНК, называемый R-петлей, в результате чего происходит расщепление. Расщепление ДНК осуществляется в результате действия двух разных нуклеазных доменов Cas9. Домен HNH пересекает цепь ДНК, которая комплементарна crRNA, а RuvC-подобный домен пересекает цепь, которая некомплементарна crRNA [8] (см. рис. 3а). Cas9 расщепляет пары оснований ДНК 3' перед PAM. Расщепление ДНК вредно для проникающей плазмиды или вируса, что приводит к их деградации и устойчивости к этим чужеродным агентам.

Чтобы иметь прецизионную точность, Cas9 должен делать точные изменения без непреднамеренных сокращений. Одна из стратегий повышения специфичности Cas9 заключается в соединении двух разных ферментов Cas9 таким образом, чтобы каждый фермент разрезал только одну цепь ДНК. Это может быть легко достигнуто, поскольку Cas9 имеет два различных каталитических домена для разрезания каждой цепи ДНК (см. рис. 3а). Так, каталитический домен HNH разрезает только одну цепь ДНК, комплементарную однонаправленной РНК, – single guide RNA (sgRNA), а мутирующий каталитический домен RuvC с единственной аминокислотной мутацией инак-

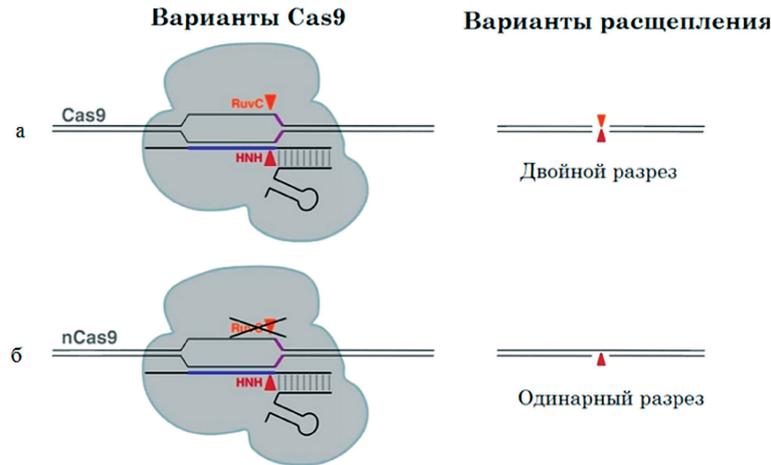


Рис. 3. Расщепление ДНК разными доменами Cas9

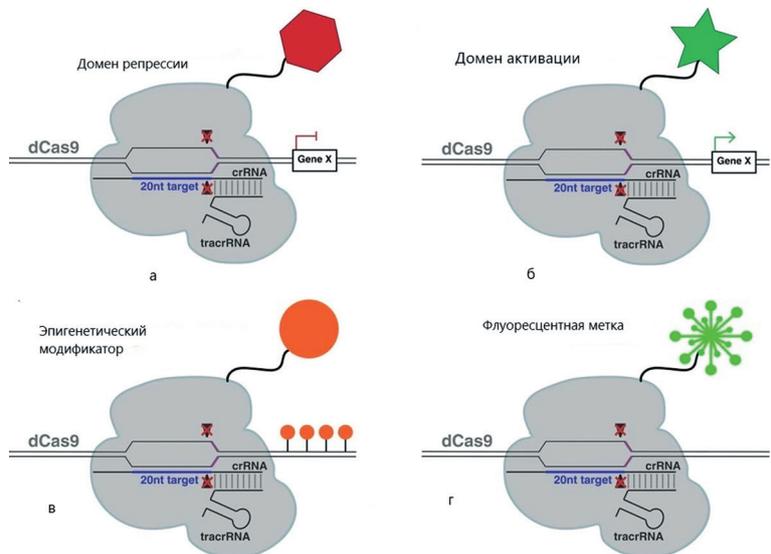


Рис. 4. Роль системы CRISPR-dCas9 в репрессии и активации генов

тивирует домен D10A [13]. Этот вариант называется никазой Cas9 или nCas9-белок, у которого активна только одна из его нуклеаз, а другая инактивирована. Такой фермент разрезает только одну из двух цепей ДНК (см. рис. 3б). Двухцепочечный разрыв может быть получен путем присоединения двух nCas9 к соседним PAM в выбранном локусе с двумя разными sgRNA [14].

Для Cas9 характерны две ключевые особенности: он может генерировать двухцепочечные разрывы, местоположение которых определяется направляющими РНК. Мутация каждого из каталитических доменов Cas9 приводит к образованию белка dCas9 (он ничего не режет, зато систему CRISPR-dCas9 можно использовать для репрессии набора генов или как платформу для конструирования более сложных регуляторных и модифицирующих комплексов). Например, если к ней привязать активирующий домен, то экспрессия целевых генов активируется, а сам он может быть нацелен на желаемый сайт в геноме с помощью sgRNA без разрезания ДНК. Белок dCas9 применяют

для инактивации генов на стадии транскрипции – он блокирует продвижение РНК-полимеразы, фермента, синтезирующего РНК на матрице ДНК. Этот способ называется CRISPR-интерференцией, или CRISPRi. Аналогичным способом можно частично выключать транскрипцию и избирательно менять ее активность. Для этого к dCas9 добавляют домен белка – фактора транскрипции, который увеличивает или подавляет активность генов, а затем снабжают его sgРНК, которая доставит эту конструкцию на регуляторный участок – промотор нужного гена [12] (рис. 4).

Привязывая другие белки к dCas9, можно передавать специфические молекулярные активности геномным областям «интереса». Например, путем нацеливания dCas9, слитого с белковым доменом, который репрессирует транскрипцию в открытой рамке считывания или к промотору гена через sgRNA. Этот химерный белок эффективно нейтрализует гены-мишени (см. рис. 4а) [2, 14]. Альтернативн, путем слияния активатора транскрипции, белка, который

рекрутирует РНК-полимеразу с dCas9 и нацеливания этой конструкции на промотор желаемого гена через sgRNA, транскрипция может быть активирована (см. рис. 4б) [14]. Используя те же принципы, что и с белком dCas9, система CRISPR-dCas9 может регулировать направления ферментативной активности в конкретное место генома (см. рис. 4в). Эпигенетические модификаторы могут быть нацелены на интересующий сайт генома путем присоединения белков, таких как ДНК-метилирующие или гистон-модифицирующие ферменты [6], модулирующие транскрипцию в желаемом месте генома. dCas9 также можно использовать для визуализации конкретной геномной последовательности в ядре, например, прикрепляя зеленый флуоресцентный белок (GFP) из медузы к dCas9 [14]. Местоположение в ядре геномной последовательности, комплементарной экспрессируемой sgRNA, можно идентифицировать в живых клетках (см. рис. 4г) [10]. Привлекая белки и ферменты в желаемые места, dCas9 предоставляет новые возможности для изучения молекулярной биологии ядра в живых клетках.

Еще одним из перспективных ферментов, которые распознают альтернативные PAM, является Cpf1 – эндонуклеаза, являющаяся эффекторным белком систем CRISPR-Cas V типа [14, 18]. Она мельче, чем Cas9, а для её функционирования нужна только crPНК, без tracrPНК. В связи с этим, возможно, в некоторых случаях система CRISPR-Cpf1 будет удобнее системы CRISPR-Cas9, что упрощает редактирование [7, 19]. Cpf1 распознает тимидин (Т)-обогащенные последовательности прото-спейсера PAM, богатые тимидином, расширяя диапазон редактирования РНК-управляемого генома – Т-обогащенный PAM, что необходимо для осуществления необходимых правок. В отличие от Cas9, который разрезает обе цепи ДНК в одном и том же месте, Cpf1 генерирует разрез, создавая «липкие» концы, которые можно использовать для вставки интересующих последовательностей путем комплементации и лигирования. Продолжающиеся исследования и характеристика микробной иммунной системы обещают выяснить и другие потенциальные ферменты редактирования генома [7, 19].

Методы изучения генетических модификаций с использованием системы CRISPR-Cas9 были осуществлены и усовершенствованы с помощью трансгенных мышей. Оригинальный метод получения трансгенных мышей заключался во введении ДНК в пронуклеусы (гаплоидные ядра гамет в составе зиготы). В процессе оплодотворения в яйцеклетке формируется два клеточных ядра – мужское и женское, но так как это может давать случайные вставки соединенных повторов, то используется редко. Стандартный метод в течение многих лет заключался в том, чтобы получить генетические изменения эмбриональных стволовых клеток (ЭСК) мышей и вводить эти модифицированные СК в ранние эмбрионы мыши, где они включаются во внутреннюю массу клеток и могут способствовать росту эмбриона вместе с клетками хозяина (рис. 5а).

Трансгенные мыши должны быть протестирова-

ны на наличие донорских клеток (химеризм). и те, которые содержат донорские клетки в зародышевой линии, отбираются, а донорские клетки доводят до необходимой концентрации. Этот метод используется очень широко. Было получено около 25000 генетически модифицированных мышей, в том числе нокаутированных, содержащих примерно около половины белково-кодирующих генов в геноме мыши. Сегодня можно модифицировать зиготы непосредственно с помощью CRISPR-Cas9, что значительно сокращает общее время, необходимое для создания нового вида мышей [8].

Для получения нового вида мышей необходимо модифицированные зиготы или предимплантационные эмбрионы внедрить в репродуктивный тракт «псевдобеременной» самки. «Псевдобеременных» самок получают путем спаривания с вазэктомизированными самцами, которые создают им необходимый гормональный статус для стельности без супоросности. Затем модифицированные эмбрионы внедряются в яйцеклетку, а модифицированные бластоцисты в матку для обеспечения нормального маршрута предимплантационного эмбриона по репродуктивному тракту [10].

При рождении мышата тестируются с помощью полимеразной цепной реакции (ПЦР). Для этого обычно используется материал от уха или хвоста. Если модифицированные ЭСК были внедрены в бластоцисту, то кожные покровы диких и химерных штаммов мышей будут отличаться по цвету, что позволяет осуществлять грубую предварительную идентификацию химеризма, но анализ ДНК с помощью ПЦР все равно необходим. Поколение трансгенных мышей будет отличаться в зависимости от используемого метода. А полученных экспериментальным путем мышат называют основателями [10, 14].

Введение трансгенных ДНК обычно осуществляется в конкретный локус, согласно правилам Менделя. Полученные таким образом трансгенные химеры содержат клетки донора и клетки хозяина. Химеры не используются, если у них отсутствуют донорские клетки в зародышевой линии, которые позволяют получить генетически модифицированные гаметы. Обычно в зависимости от уровня зародышевой линии химеризма удается получить до 50% потомства первого поколения (F1), содержащего генную модификацию. F1 животные, несущие модификацию, являются гетерозиготными. Они могут быть спарены вместе, чтобы генерировать до 25% гомозиготных мышей в поколении F2. Использование CRISPR-Cas9 в зиготах в принципе позволяет получить мышат-основателей, имеющих гомозиготную модификацию [10, 14].

Традиционный метод получения генетически модифицированных ЭСК включает внедрение целевых конструкций в клетки, а затем использование «положительно-отрицательного» метода отбора с двумя антибиотиками для увеличения вероятности получения гомологичных рекомбинаций между целевой конструкцией и геномной ДНК (см. рис. 5б). Это должно

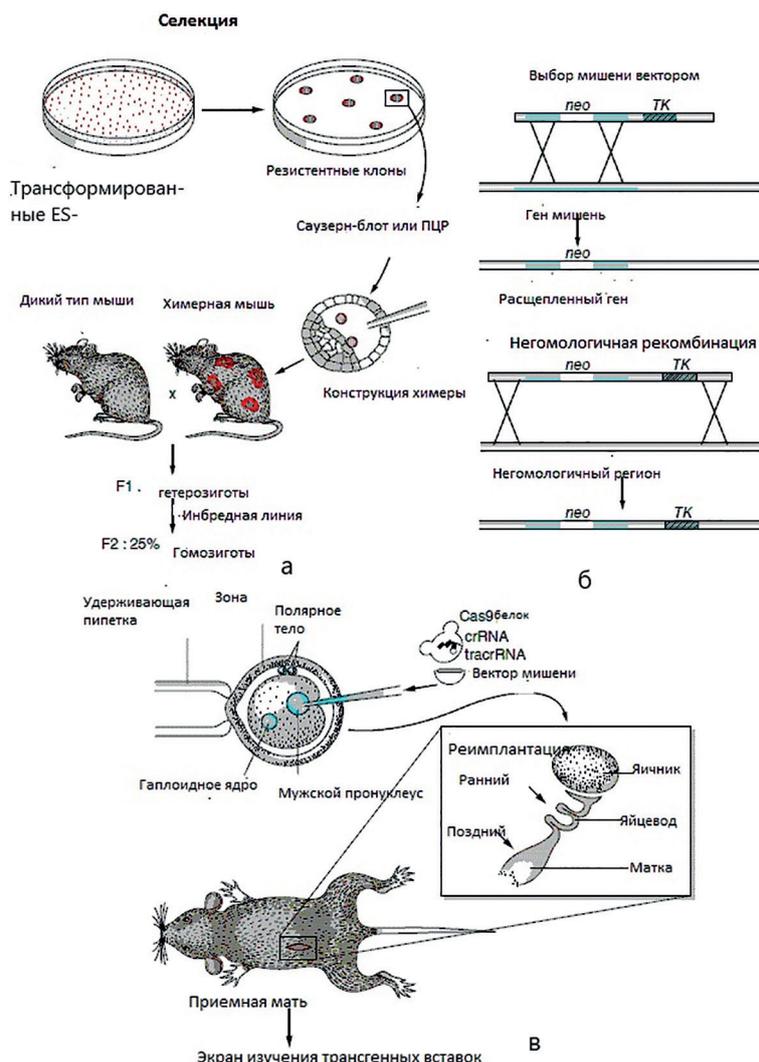


Рис. 5. Создание трансгенных мышей из ЭСК: а – клетки модифицируются путем введения рекомбинантных клонов, отобранных и протестированных. Они вводятся в бластоцисты матки мышей-реципиентов; б – механизм позитивной и негативной селекции (neo – неомицин-резистентные гены. ТК – ген тимидин киназы); в – один из методов использования CRISPR-Cas9 для трансгенной модификации зигот. Здесь компоненты вводятся в пронуклеусы

гарантировать, что нужная рекомбинация пройдет в соответствующем сайте, а не в каком-то случайном. Клоны клеток ES, прошедшие отбор антибиотиками, затем тестируются ПЦР, чтобы убедиться в соответствии прошедшей генетической модификации [14].

Маркированный антибиотиком локус извлекается с помощью рекомбиназы Cre для получения фланкированных аллелей – локуса кроссинговера (loxP-сайты). После получения желаемых клонов модифицированные ЭСК используются для получения трансгенных мышей путем их введения в бластоцисты. Использование CRISPR-Cas9 позволяет повысить вероятность гомологичной рекомбинации в ЭСК из-за специфичности, введенной РНК [3, 14].

Кроме того, используя различные RNAs, можно осуществлять модификации более чем в одном локусе одновременно. Однако независимо от того, какой метод используется, необходимо сохранять выбранные антибиотики для получения чистых рекомбинантных

клонов, которые по-прежнему нуждаются в анализе ДНК. Это необходимо для контроля правильности проведенной модификации и отсутствия нецелевого эффекта от Cas9 в другом месте генома. Система CRISPR-Cas9, кроме проведения генетических изменений в ЭСК, также используется для генетических манипуляций непосредственно с зиготами [5, 10].

Однако представленные методы по-прежнему остаются неидеальными. Фермент Cas9 может быть введен в клетку как белок или как РНК. РНК в одной части клетки используется для распознавания цели, а в другой – для связывания Cas9, как это происходит в естественной среде с иммунной системой бактерий. Эти структуры могут быть введены в пронуклеус или в цитоплазму или внедрены электропорацией. Для улучшения гомологической рекомбинации обычно подавляется доминирующий процесс конечного присоединения путем включения ингибитора ДНК-лигазы IV [14].

Потенциальные преимущества этого прямого метода заключаются в том, что он гораздо более быстрый в получении нового вида мышей, так как уже с самого начала можно получить гомозиготную модификацию. Метод позволяет избежать спонтанных мутаций, связанных с ЭСК, которые могут накапливаться в клеточных культурах и присутствовать в трансгенных мышах. Кроме того, аналогичный метод может быть применен к сельскохозяйственным животным, для которых технология ЭСК в настоящее время пока недоступна. Однако всегда существует риск нецелевого воздействия Cas9 [9, 14].

По мере совершенствования технологий чистота трансгенных мышей неуклонно возрастала. Первоначально гены были сверхэкспрессированы, так как использовались либо сильные универсальные, либо мощные тканевые промоторы. Это приводило к отклонениям развития и было недостаточно информативно относительно состояния и функций экзогенного гена. Затем гены у мышей-мутантов (нокаут-мыши) удалялись, чтобы вызвать полную потерю функции. Для получения необходимого фенотипа нокаут-мыши объединялись в одну группу, что в итоге позволяло вывести нужные виды мышей [14].

В целом даже удаление целого семейства генов, например паралоговой группы генов *Hox*, не приводило к изменению фенотипа. Так что удаление отдельного гена не приводило к нежелательным последствиям, так как биологические функции отдельных генов заменялись другими, что приводило к другим фенотипическим изменениям. Иногда нокауты приводили к летальному исходу уже на ранней стадии эмбрионального развития. Это в первую очередь связано с тем, что ген сохранял свою функцию в плаценте [10].

Для того чтобы обойти это раннее плацентарное требование и исследовать более поздние функции того же гена в эмбрионе, гомозиготные мутантные ЭСК могут быть введены в тетраплоидные бластоцисты, т. е. те, которые состоят из клеток с четырьмя наборами хромосом вместо двух. Это приводит к тому, что плацента формируется из тетраплоида, в то время как сам эмбрион происходит от ЭСК и будет отображать фенотип, соответствующий первой эмбриональной функции удаленного гена. Под воздействием импульсного электрического тока тетраплоидный эмбрион хозяина формируется путем слияния двух клеток в тетраплоидную одноклеточную клетку. Сами по себе тетраплоидные эмбрионы могут образовывать нормальную плаценту, но не жизнеспособный плод [14].

При введении ЭСК формируют плод, поддержанный тетраплоидной плацентой. Более сложный метод нокаута – это тканеспецифический нокаут с использованием системы Cre. В этом случае тканеспецифический активатор используется для получения нокаута системой Cre, а целевой – замещается локусом кроссинговера. Обычно только одна из копий флуксированного гена с другой аллелью может быть мутантом. Однако даже этот метод имеет выраженные

недостатки, связанные с ранней фазой деятельности промотора над достаточно большой областью, вызывая повреждение или удаление генов во всей этой области [5, 14].

Это может быть исправлено с помощью системы CreER. Активность Cre индуцируется путем инъекции тамоксифена беременной самке, и поэтому нокаут возникает только в тех клетках эмбриона, где промотор активен и совпадает по времени с индукцией тамоксифена. Точно такой же метод гомологичной рекомбинации может быть использован для замены гена с измененной версией. Эти виды мышей (штаммы) называются «knock-ins» – заданный ген не удаляется из организма, а заменяется другим. Это значимо отличается по сравнению со случайной вставкой трансгенов, управляемых тканевыми активаторами, протекающими в хромосомной среде с генетической регуляцией модифицированного локуса. Все эти методы генетической модификации мышей используются в биологии СК [3, 14].

Так как клетки в организме возникают в результате митотического деления, каждая клетка имеет свое собственное генеалогическое древо, или родословную, ведущую обратно к оплодотворенной яйцеклетке. Установление клеточной линии имеет важное значение в биологии развития, потому что история любой отдельной клетки указывает на то, какие типы клеток-предков были ее предками. Линия клеток также очень важна в биологии СК, потому что СК является предшественником для всех типов клеток собственной ткани. Таким образом, возможность наблюдать и отслеживать линию клеток имеет центральное значение для определения существования и свойств различных типов СК [2, 14].

Большинство из них зависят от генетической модификации и применимы только у мышей, хотя некоторые из них зависят от естественно существующих генетических вариаций и могут быть применены к людям. С целью отслеживания линии клеток используются «метки», или «маркеры». Это может быть продукт генов, введенный в клетки с целью отслеживания линии и идентификации типа клеток. Считается, что модели экспрессии генов остаются постоянными во время развития, так что если клетка экспрессирует нитевидные структуры белка виментина, то клетка считается «мезенхимального происхождения» [1, 8, 17].

Конечно, гены могут быть включены и выключены при различных обстоятельствах, но наличие конкретного белкового маркера в клетке еще не доказывает, что он потомок предка, экспрессирующего тот же маркер. Отдельные клеточные линии млекопитающих могут выявляться у некоторых беспозвоночных животных, таких как лягушки, рыбы, что подтверждается гистохимическим обнаружением флуоресцентных белков или ферментов [14].

Необходимо различать три связанных, но перекрывающихся понятия: линия клеток, карта развития клетки и клональный анализ. Клеточная линия пред-

ставляет собой описание истории клетки начиная с оплодотворенной яйцеклетки. То есть можно узнать полную линию клеток для нескольких типов беспозвоночных, которые имеют небольшое общее число клеток и инвариантную структуру деления клеток и миграции. Хорошо известным примером является нематода *Caenorhabditis elegans*, которая часто используется в качестве модели развития организма. Знать линию клетки очень важно, хотя сама по себе клеточная линия не дает никакой пространственной информации о том, где находится каждая клетка и какие другие клетки являются ее соседями на разных стадиях развития. Эта пространственная информация является картой развития, которая отражает развитие как раннего эмбриона, так и его конкретной области. Она показывает, как он движется, изменяет форму, что важно с точки зрения дифференциации клеток. В отличие от нематод, большинство животных, включая всех млекопитающих, имеют важные различия в делении, движении клеток, поэтому клеточные линии имеют существенные отличия у отдельных индивидумов [10, 14].

Необходимо понимать, что карта развития не дает полной информации о предрасположенности развития клеток на конкретных стадиях. Для этого требуются дополнительные данные, такие как перемещение клеток в чужеродную среду или получение моделей экспрессии ключевых генов, участвующих в развитии. Клональный анализ фокусируется на том, что происходит с отдельной клеткой с точки зрения последующей дифференциации. В отличие от карты развития, клональный анализ может дать информацию о предрасположенности к конкретному развитию [3, 8].

Если помеченный клон состоит из двух типов клеток, то его формирование происходит после маркировки этих двух типов клеток. Если весь клон состоит из одного типа клеток, то это может быть связано с тем, что ячейка уже была в форме этого типа во время маркировки или область, в которой находится клетка, впоследствии была выявлена в форме данного типа клеток с помощью внешнего сигнала [3, 10].

Клональный анализ особенно полезен и при анализе поведения СК. Если генетическая метка применяется к одной СК, то метка позже появится как участок, состоящий из самой СК, ее потомков – транзитных и дифференцирующихся клеток. Кроме того, этот участок будет сохраняться в долгосрочной перспективе – до тех пор, пока сохраняется СК. Таким образом, клональная маркировка прямо указывает на две ключевые особенности СК: способность формировать ансамбль типов клеток в ее тканях и долгосрочную устойчивость.

Клональный анализ может быть проведен в пробирке, *in vitro*. Обычно это делается путем внесения отдельных клеток в лунки и наблюдения за изменениями клона. Если один клон может генерировать более одного дифференцированного типа клеток, он считается мультипотентным. Такие данные имеют особенно важное значение при установлении свойств клеток в

гемопоетической системе. Также клон можно расширить, поместив клетки в несколько лунок, заполненных различными средствами, которые могут вызывать отличающиеся эффекты. Эта процедура имеет важное значение для установления многофункциональности мезенхимальных СК [8, 14].

Система CreER стала стандартным методом отслеживания линии клеток у млекопитающих. Ее различные варианты могут быть использованы либо для оценки развития, либо для клонального анализа. В сущности, метод дает постоянный маркер клеткам, которые имеют конкретный активный промоутер. Если такие клетки образуют когерентный участок ткани, как это часто бывает в раннем эмбрионе, то маркировка отражает карту развития для мечения, показывая, где участок заканчивается и какие типы клеток он производит. Если помечены несколько клеток, то отдельные клоны могут быть визуализированы и проанализированы. Для надежности этого лучше всего использовать гены «knock-ins». Для достижения этой цели в качестве метки можно использовать флуоресцентный белок (GFP) или фермент «галактозидаза», легко обнаруживаемый гистохимическим методом. Маркировочный ген белка контролируется активатором Rosa 26, чья нормальная функция заключается в том, чтобы управлять экспрессией транспортной непереведенной РНК [7, 18]. В настоящее время существует доступный набор видов мышей-репортеров, которые имеют различные гены, вставленные в сайт локуса Rosa 26. Между активатором и областью кодирования находится транскрипционная стоп-последовательность, окруженная сайтами loxP. Фермент CreER вырабатывается клетками, в которых активен его промоутер. Дозированная маркировка животных осуществляется тамоксифеном (или 5-гидрокситамоксифеном). Это активирует CreER, позволяет вырезать стоп-последовательности и способствует транскрипции репортера. Поскольку конечным результатом является модификация ДНК, она является постоянной, и маркировка поддерживается независимо от последующих делений или дифференциации клеток. Доля маркированных клеток зависит от общей дозы тамоксифена. Достаточно нескольких инъекций тамоксифена, чтобы получить полную маркировку всех клеток. Низкая разовая доза позволяет маркировать лишь небольшую часть клеток, и, пока меченые клетки хорошо дифференцируются, можно проводить клональный анализ. Применительно к СК он позволяет выявить область ткани, содержащей хоть одну СК. Более сложные версии клональной маркировки связаны с одновременным мечением нескольких клонов разными цветами. Эти методы общеизвестны как методы «Brainbow». Они зависят от использования вариантов сайта loxP таким образом, что альтернативные события вырезания могут происходить в локусе репортера, что приводит к экспрессии различных цветных белков в зависимости от конкретного вырезания. Метод маркировки линии CreER имеет огромное значение в исследованиях СК. Но, как и у других методов, у него

также есть свои проблемы. Самая большая – зависимость от стабильности промотора [11, 14, 18].

Тем не менее, с помощью технологии, основанной на CRISPR-Cas9, созданы изогенные СК человека [7], разрабатываются методы исправления мутантного фенотипа клеток [8], проводятся работы по регуляции экспрессии генов [10], изучению функциональных взаимоотношений между большими группами генов [15] и визуализации функционирующих районов генома в живых клетках [19].

Создание панелей изогенных плюрипотентных СК 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 иммунодефицита, нестабильности центромерных районов хромосом и лицевых аномалий (immunodeficiency, centromeric region instability and facial anomalies syndrome – ICF) на индуцированных плюрипотентных СК человека. Были получены гомозиготные мутации в гене *DNMT3B* с частотой 63%, при этом клетки имели фенотип центромерной нестабильности [10, 14]. 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) [7, 11, 15, 18].

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 СК (линия HUES62) [13]. Показано, что в формировании резистентности к вемурафенибу (PLX), который является BRAF-ингибитором протеинкиназы при меланоме, участвуют не только гены *NF1* и *MED12*, но и ген *CUL3*, а также комплекс гистон-специфических ацетилтрансфераз STAGA: *TADA1* и *TADA2* [15, 19]. При использовании лентивирусной библиотеки, со-

державшей порядка 73000 sgRNA, на примере линий опухолевых клеток KBM7 и HL60 изучены гены, участвующие в пролиферации и клеточном цикле [19]. 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 [10].

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 СК. В частности, применение систем редактирования геномов позволяет исправлять точечные мутации в клетках, полученных от больных. Объектом исследования в данном случае могут быть индуцированные плюрипотентные СК и региональные СК. При этом донорными молекулами могут служить как сложные генетические конструкции, так и одноцепочечные ДНК-олигонуклеотиды [15]. Интересным примером подобного подхода представляется работа, в которой осуществлена коррекция локуса муковисцидозного регулятора трансмембранной проводимости (cystic fibrosis transmembrane conductor regulator – *CFTR*) в культивируемых интестинальных СК, полученных от больных муковисцидозом (cystic fibrosis – CF) [11]. 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 [14].

Следовательно, система CRISPR-Cas может эффективно применяться в функциональной геномике клеток для создания клеточных моделей заболеваний человека и клеточной терапии.

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, лабораторные мыши и крысы. На этих и ряде других модельных организмов активно проводятся эксперименты по геномной инженерии с помощью системы CRISPR-Cas9 [10, 16].

Различные варианты применения CRISPR-Cas9 и модификации технологии редактирования генома у нематоды *Caenorhabditis elegans* представлены в работах S. Sterberg [15], J.B. Zabriskie [18]. С помощью инъекции мРНК-белка Cas9 и продуцированной *in vitro/in vivo* sgRNA в клетки зародышевой линии у взрослых животных в следующем поколении получали стабильные целевые модификации генома, включая небольшие инсерции/делеции, более крупные хромосомные делеции и перестройки [4, 14], внедрение трансгена путем гомологичной рекомбинации с донорскими молекулами [11]. Такой метод активно используется для изучения процессов дозовой компенсации у нематоды, сравнения функций генов у родственных видов *C. elegans* и *C. briggsae* [16].

Плодовая мушка *Drosophila melanogaster* относится к наиболее изученным модельным объектам. Однако получение новых мутантных аллелей посредством гомологичной рекомбинации по-прежнему остается очень трудоемкой процедурой [16]. 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. 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 [15]. Эти мутации стабильно передаются из поколения в поколение, что обеспечивает возможность создания новых линий мух [11]. В 2015 г. разработано приложение, 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* [16].

Посредством инъекции компонентов CRISPR-Cas9 в эмбрионы полосатого данио получены разнообразные целевые модификации, стабильно передающиеся по наследству. В 2011 году был открыт международный проект по созданию мутантных аллелей (The Zebrafish Mutation Project) в каждом белок-кодирующем гене полосатого данио. На июнь 2013 г. получены мутантные модели 46% всех белок-кодирующих 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 СК [19], а также посредством инсерционного мутагенеза [19]. 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. Такие разрывы репарируются по механизму негомологичного сшивания концов, что приводит к появлению мутантных крыс и мышей, несущих делеции либо инсерции в разрезанном сайте. При добавлении донорной плазмиды или олигонуклеотида разрывы могут репарироваться с помощью высокоточного механизма гомологичной рекомбинации, что позволяет получить животных, несущих целевые вставки ДНК. Редактирование генома с использованием CRISPR-Cas9 обеспечивает внесение мутаций как в один ген, так и в несколько генов сразу. Так, использование CRISPR-Cas9 с высокой эффективностью приводит к появлению в ЭСК мыши мутаций в пяти генах одновременно, а введение мРНК Cas9 и sgRNAs, направленных на гены *Tet1* и *Tet2*, в зиготы мыши позволяет в 80% случаев получить животных с биаллельными мутациями обоих генов [6, 11, 15].

Аналогичные результаты получены в экспериментах на крысах, причем и мыши, и крысы стабильно наследовали выявляемые мутации. Кроме того, произведена эффективная коррекция мутации в гене *Crygc* у мышей с доминантной формой катаракты, вызванной этой мутацией. 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 методу молекулярной генетики, при котором из организма удаляют или делают неработоспособными определенные гены (нокаут того или иного гена мыши), собраны на сайте международного проекта – International mouse phenotyping consortium (IMPC).

Литература

1. Москалев, А.В. Общая иммунология с основами клинической иммунологии / А.В. Москалев, В.Б. Сбойчаков, А.С. Рудой. – М.: Гэотар-Медиа, 2015. – 351 с.
2. Ярилин, А.А. Иммунология / А.А. Ярилин. – М.: Гэотар-Медиа, 2010. – 957 с.

3. Abbas, A.K. Cellular and Molecular Immunology. 9-th edition / A.K. Abbas, A.H. Lichtman, S. Pillai. Philadelphia, Pennsylvania: W. B. Saunders Company, 2018. – 565 p.
4. Bhaya, D. CRISPR-Cas systems in bacteria and archaea: Versatile small RNAs for adaptive defense and regulation / D. Bhaya [et al.] // Annu. Rev. Genet. – 2011. – Vol. 45. – P. 273–297.
5. Cai, L. Suppression of hepatocyte growth factor production impairs the ability of adipose-derived stem cells to promote ischemic tissue revascularization / L. Cai [et al.] // Stem Cells. – 2007. – Vol. 25. – P. 3234–3243.
6. Hilton, I.B. Epigenome editing by a CRISPR-Cas9 – based acetyltransferase activates genes from promoters and enhancers / I.B. Hilton [et al.] // Nat. Biotechnol. – 2015. – Vol. 33. – P. 510–517.
7. Gilbert, L.A. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes / L.A. Gilbert [et al.] // Cell. – 2013. – Vol. 154. – P. 442 – 451.
8. Kern, S.E. Comparative analysis of mesenchymal stem cells from bone marrow, umbilical cord blood, or adipose tissue / S.E. Kern [et al.] // Stem Cells. – 2006. – Vol. 24. – P. 1294–1301.
9. Lee, J. Human adipose-derived stem cells display myogenic potential and perturbed function in hypoxic conditions / J. Lee [et al.] // Biochem. Biophys. Res. Commun. – 2006. – Vol. 341. – P. 882–888.
10. Li, B. Adipose tissue stromal cells transplantation in rats of acute myocardial infarction / B. Li [et al.] // Coron Artery Dis. – 2007. – Vol. 18. – P. 221–227.
11. McDonald, J.I. Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation / J.I. McDonald [et al.] // Biol. Open. – 2016. – Vol. 5. – P. 866–874.
12. Olson, K. Contemporary clinical immunology and serology / K. Olson, E. De Nardin. New Jersey: Upper Saddle River, 2013. – 439 p.
13. Rose, N.R. The autoimmune diseases. fifth edition / N.R. Rose, I.R. Mackay. Philadelphia, 2018. 1265 p.
14. Slack, J.M.W. The science of stem cells / J.M.W. Slack. – Wiley, 2018. – 248 p.
15. Sternberg, S.H. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9 / S.H. Sternberg // Nature. – 2014. – Vol. 507. – P. 62–67.
16. Thakore, P.I. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements / P.I. Thakore [et al.] // Nat. Methods. – 2015. – Vol. 12. – P. 1143–1149.
17. Wu, Y. Mesenchymal stem cells enhance wound healing through differentiation and angiogenesis / Y. Wu [et al.] // Stem Cells. – 2007. – Vol. 25. – P. 2648–2659.
18. Zabriskie, J.B. Essential clinical immunology / J.B. Zabriskie – N. Y., 2009. – 362 p.
19. Zetsche, B. Cpf1 is a single RNA – guided endonuclease of a class 2 CRISPR-Cas system / B. Zetsche [et al.] // Cell. – 2016. – Vol. 163. – P. 759–771.

A.V. Moskalev, B.Yu. Gumilevskiy, V.Ya. Apchel, V.N. Tsygan

Methods for studying genetic modifications

Abstract. *Objects and modern methods of genome editing are considered. The immune system of prokaryote and their protective mechanisms that prevent the purposeful editing of the genome for the benefit of the researcher is characterized. This mechanism in prokaryotes are cluster regulatory interspatial short palindrome repetitions. The number of such repetitions varies from object to object, which ultimately makes it impossible to get the perfect standard model. Three types of such systems that have their own mechanism for generating proteins have now been identified. The proteins, which are now most commonly used to edit the genome and identify areas of proto-special adjacent motifs, are described. Detailed characteristics of the organization of the immune system prokaryote and phases of its activity are given. Three types of short-palin re-recurrence systems have now been identified, and the teams are being identified as cluster regulatory interspatial short palindrome repetitions-Cas9. Each system uses its own mechanism to generate proteins that catalyze the fission of nucleic acids. The type II cluster regulatory interspatial short palindrome repetitions system is most commonly used, better adapted to edit the genome because of its simplicity. It has been established that the cluster regulatory interspatial short palindrome repetitions-Cas9 system can be used for point editing of the genome and in eukaryotes. This is done either through non-homological annexation of the end, or by homologically directed reparation. A promising variant of genetic modeling is the use of the enzyme-endonuclease Cpf1, which is the effector protein of the cluster regulatory interspatial short palindrome repetitions-Cas V type systems. Cpf1 is smaller than the enzyme protein Cas9 and for the system to function only require specers of ribonucleic acid, without additional ribonucleic acid. Unlike Cas9, which cuts both chains of deoxyribonucleic acid in the same place, Cpf1 generates an incision, creating «ticky» ends that can be used to insert interesting sequences by complementing and ligation. It is likely that the system using the enzyme-endonuclease Cpf1 will be more convenient than the system where the protein is used – Cas9, as the range of editing of the controlled genome of ribonucleic acid is expanded to make the necessary edits.*

Key words: *stem cells, transcriptional factors, genes, phenotype, cellular differentiation, modification, mutation, chromosome, plasmids, nucleic acids, translocation.*

Контактный телефон: 8-921-989-17-42; e-mail: vmeda-nio@mil.ru