

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
НАУЧНОЕ УЧРЕЖДЕНИЕ
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
ИНСТИТУТ ЦИТОЛОГИИ И ГЕНЕТИКИ
СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ НАУК»

На правах рукописи

НЕМУДРЫЙ АРТЕМ АЛЕКСАНДРОВИЧ

ИСПРАВЛЕНИЕ МУТАЦИИ В ГЕНЕ АРГИНИН-ВАЗОПРЕССИНА КРЫС
ЛИНИИ BRATTLEBORO *IN VITRO*

03.02.07 – генетика

Диссертация на соискание ученой степени
кандидата биологических наук

Научный руководитель
д.б.н., проф. Закиян С.М.

Новосибирск – 2017

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ.....	9
1.1 Редактирование генов и геномов с помощью CRISPR/Cas.....	9
1.1.1 Общий принцип организации системы CRISPR/Cas у бактерий	9
1.1.2 Организация и механизм действия систем CRISPR/Cas типа II-A	11
1.1.3 Адаптация системы CRISPR/Cas9 для редактирования геномов эукариот	18
1.1.4 Спектр возможных модификаций при использовании системы CRISPR/Cas9 для редактирования геномов	19
1.1.5 Нецелевые эффекты системы CRISPR/Cas9	24
1.1.6 Современные концепции применения системы CRISPR/Cas для терапии наследственных заболеваний	30
1.2 Использование индуцированных плюрипотентных стволовых клеток для терапии наследственных заболеваний	33
1.2.1 Исправление мутаций в пациент-специфичных ИПСК	35
1.2.2 Перспективы использования ИПСК для терапии наследственных заболеваний	40
1.3 Перспективы использования крыс линии Brattleboro для разработки клеточной терапии наследственных заболеваний	42
1.3.1 Структура локусов <i>Avr</i> и <i>Oxt</i> у крыс	42
1.3.2 Синтез аргинин-вазопрессина	43
1.3.3 Крысы Brattleboro – модель наследственного несахарного гипоталамического диабета.....	44
1.4 Заключение.....	47
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ.....	48
2.1 Конструирование векторов для экспрессии системы CRISPR/Cas9	48
2.2 Методы работы с клеточными культурами	48
2.3 Методы доставки ДНК <i>in vitro</i>	49
2.4 Молекулярно-генетические методы анализа	50

2.5 Выявление потенциальных нецелевых сайтов действия CRISPR/Cas9	54
ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ	55
3.1 Создание донорного вектора для гомологичной рекомбинации в локусе <i>Avr</i>	55
3.2 Выбор последовательностей протоспейсеров для действия CRISPR/Cas9 в локусе <i>Avr</i>	56
3.3 Сравнение активности CRISPR/Cas9 в выбранных сайтах гена <i>Avr</i>	58
3.3.1 Анализ с использованием эндонуклеазы I фага T7	58
3.3.2 Рестрикционный анализ	59
3.4 Биоинформатический анализ потенциальных нецелевых эффектов	61
3.5 Определение эффективности внесения двуцепочечных разрывов системой CRISPR/Cas9 в протоспейсер CR-7	66
3.6 Получение и анализ смешанных популяций в результате позитивно-негативной селекции	75
3.7 Получение и анализ клональных линий	76
3.8 Обсуждение эффективности гомологичной рекомбинации	83
ЗАКЛЮЧЕНИЕ	88
ВЫВОДЫ	90
СПИСОК ИСПОЛЬЗОВАННЫХ СОКРАЩЕНИЙ	91
СПИСОК ЛИТЕРАТУРЫ	92

ВВЕДЕНИЕ

Актуальность проблемы

Наследственные заболевания, вызванные разного рода мутациями (точковые мутации в генах, хромосомные перестройки и т.д.), в настоящее время неизлечимы, поскольку их терапия сводится к устранению симптомов, не влияя на первопричину болезни (мутацию). В настоящее время, благодаря появлению системы CRISPR/Cas, стало возможным эффективное внесение модификаций в геном плюрипотентных стволовых клеток (Cong et al., 2013, Mali et al., 2013). Одна из современных концепций лечения наследственных заболеваний предполагает следующий подход: 1) получение соматических клеток пациента, 2) репрограммирование их к плюрипотентному состоянию, 3) исправление мутации, вызывающей заболевание, с помощью современных методов редактирования геномов, 4) трансплантация клеток с исправленным геномом обратно в организм пациента (Cox et al., 2015). Данная концепция предполагает использование индуцированных плюрипотентных стволовых клеток (ИПСК) – стволовых клеток, которые могут быть получены практически из любых клеток пациента (клетки крови, клетки жировой ткани, фибробласты кожи) в результате репрограммирования (Okita et al., 2013, Takahashi et al., 2007). При дифференцировке из этого типа клеток возможно направленно получить любые типы клеток, из которых состоит взрослый организм. После трансплантации клетки с исправленной мутацией замещают поврежденные в результате патогенеза наследственного заболевания клетки и, соответственно, восполняют утраченные функции (Hanna et al., 2007).

Закономерным этапом перед внедрением подобных технологий является их доскональное изучение на модельных системах. Ряд вопросов, касающихся безопасности такого подхода, вопросов технического характера в настоящее время являются открытым. Именно это обуславливает потребность в адекватных модельных системах и проведении научных исследований с их использованием. В данной работе в качестве модельной системы для разработки терапии наследственных заболеваний с помощью редактирования геномов используются лабораторные крысы линии Brattleboro, гомозиготные носители мутантного аллеля *di* гена аргинин-вазопрессина. У гомозиготных носителей мутация вызывает наследственную форму несахарного гипоталамического диабета. Данная работа

направлена на исправление мутации в гене аргинин-вазопрессина крыс Brattleboro *in vitro* с помощью новейших технологий редактирования генома.

Цель и задачи работы

Цель работы – создать систему для гомологичной рекомбинации в локусе гена аргинин-вазопрессина и получить клетки крыс линии Brattleboro с исправленной мутацией в гене аргинин-вазопрессина.

Задачи:

1. Создать набор генетических конструкций, экспрессирующих элементы системы CRISPR/Cas9, для внесения двунитевых разрывов в различные участки локуса гена аргинин-вазопрессина крыс линии Brattleboro.
2. Провести качественный анализ активности системы CRISPR/Cas9 в целевых сайтах гена аргинин-вазопрессина эмбриональных фибробластов крыс Brattleboro.
3. Провести анализ нецелевых эффектов в паралогичном гене окситоцина и определить с помощью биоинформатических методов потенциальные сайты нецелевого действия CRISPR/Cas9 в геноме.
4. Определить эффективность направленного внесения двунитевых разрывов в выбранном сайте гена аргинин-вазопрессина и спектр модификаций, возникающих при репарации данных разрывов, в эмбриональных фибробластах крыс линии Brattleboro.
5. С помощью гомологичной рекомбинации произвести исправление мутации в гене аргинин-вазопрессина эмбриональных фибробластов крыс линии Brattleboro.

Научная новизна работы

В данной научно-исследовательской работе впервые предлагается использовать крыс лабораторной линии Brattleboro для изучения возможности терапии наследственных заболеваний с помощью редактирования геномов *ex vivo*. Впервые были получены эмбриональные фибробласты крыс линии Brattleboro с исправленной мутацией в гене аргинин-вазопрессина. Предложенная в данной работе линия крыс Brattleboro является адекватной моделью моногенного наследственного заболевания аутосомно-рецессивного типа. Моногенная природа заболевания в данном случае упрощает применение методов редактирования

геномов для исправления мутации и позволяет на более «простом» объекте проводить исследования перед переходом к более сложным случаям при терапии мультигенных заболеваний. При этом заболевание, носителем которого являются крысы Brattleboro – несахарный диабет – связано с нарушением секреции гормона аргинин-вазопрессина в конкретном типе клеток – нейронах гипоталамуса. Дальнейшее использование этих животных позволит моделировать клеточную терапию заболевания, связанного с нарушением в нейронах определенного типа – это такие заболевания, как болезнь Гентингтона (нейроны стриатума), болезнь Паркинсона (нейроны черной субстанции) и другие.

Научно-практическая значимость работы

Научные результаты и предложенные в данной работе подходы и стратегии могут быть применены в биомедицине при разработке технологий терапии наследственных заболеваний человека с помощью редактирования генома *ex vivo*. Проведение дальнейших работ на предложенном модельном объекте позволит: 1) определить возможность клеточной терапии наследственных заболеваний с использованием ИПСК с исправленным генотипом, 2) определить возможность клеточной терапии заболеваний, связанных с гибелью нейронов в структурах головного мозга, 3) всесторонне изучить весь процесс клеточной терапии на всех её этапах, разработать методики и подходы.

Положения, выносимые на защиту:

1. Система CRISPR/Cas9 направленно вносит модификации во второй экзон мутантного гена аргинин-вазопрессина крыс линии Brattleboro, не вызывая при этом нецелевых эффектов в паралогичном гене окситоцина.
2. При репарации двуцепочечного разрыва, направленно внесенного системой CRISPR/Cas9, происходит исправление мутации в гене аргинин-вазопрессина крыс линии Brattleboro в результате гомологичной рекомбинации с донорным плазмидным вектором.

Вклад автора

Основные результаты получены автором самостоятельно. Донорный плазмидный вектор для гомологичной рекомбинации в гене *Avp* был сконструирован к.б.н. С. П. Медведевым. Качественный анализ действия CRISPR/Cas в

эмбриональных фибробластах и рестрикционный анализ проводился совместно с Т.Б. Маланхановой.

Апробация результатов работы

Список работ по теме диссертации, опубликованных в рецензируемых журналах:

1. **Немудрый А.А.**, Валетдинова К.Р., Медведев С.П., Закиян С.М. Системы редактирования геномов TALEN и CRISPR/Cas - инструменты открытий // Acta Naturae. – 2014. – Т. 6. – № 3 (22) . – С. 20-42
2. Vaskova E. A., Medvedev S. P., Sorokina A. E., **Nemudryy A. A.**, Elisaphenko E. A., Zakharova I. S., Shevchenko A. I., Kizilova E. A., Zhelezova A. I., Evshin I. S., Sharipov R. N., Minina J. M., Zhdanova N. S., Khegay, II, Kolpakov F. A., Sukhikh G. T., Pokushalov E. A., Karaskov A. M., Vlasov V. V., Ivanova L. N., Zakian S. M. Transcriptome Characteristics and X-Chromosome Inactivation Status in Cultured Rat Pluripotent Stem Cells // Stem Cells and Development. – 2015. – V. 24. – P. 2912-2924.
3. **Немудрый А.А.**, Маланханова Т.Б., Малахова А.А., Медведев С.П., Закиян С.М. Стратегии редактирования паралогичных генов с помощью CRISPR/CAS9 // Гены и Клетки. – 2016. – Т. XI. – №2. – С. 87-94

Тезисы докладов:

1. **Немудрый А.А.**, Васькова А.А., Стекленева А.Е., Медведев С.П. Разработка технологии исправления генетических мутаций, обуславливающих развитие наследственных заболеваний // Сборник тезисов научной конференции «Фундаментальные науки — медицине: Актуальные проблемы молекулярной медицины», посвященной 10-летию Медицинского факультета НГУ 16-20 сентября 2013 г., с. 137-138
2. **Немудрый А.А.**, Стекленева А.Е., Медведев С.П., Васькова Е.А., Иванова Л.Н., Покушалов Е.А., Закиян С.М. Применение систем TALEN и CRISPR/Cas9 для исправления генетических мутаций в плюрипотентных клетках крыс Brattleboro // Сборник тезисов V Всероссийской научно-практической конференции «Стволовые клетки и регенеративная медицина», 18-21 ноября 2013 г. – М.: МАКС Пресс. с. 51.
3. Васькова Е.А., Медведев С.П., Стекленева А.Е., **Немудрый А.А.**, Елисафенко Е.А., Евшин И.С., Шарипов Р.Н., Сайфутдинова С.Г., Шевченко А.И., Кизилова

Е.А., Железова А.И., Иванова Л.Н., Покушалов Е.А., Сухих Г.Т., Закиян С.М. Создание системы для исправления генетических мутаций на основе плюрипотентных клеток крыс линии Brattleboro // Материалы «I Национального конгресса по регенеративной медицине», 4-6 декабря 2013 г., г. Москва, с. 46-47.

4. **Немудрый А.А.**, Стеклёнова А.Е., Васькова Е.А., Медведев С.П., Иванова Л.Н., Закиян С.М. Исправление мутантного генотипа в клетках крыс линии Brattleboro с помощью современных методов геномной инженерии // VI СЪЕЗД ВАВИЛОВСКОГО ОБЩЕСТВА ГЕНЕТИКОВ И СЕЛЕКЦИОНЕРОВ (ВОГиС) И АССОЦИИРОВАННЫЕ ГЕНЕТИЧЕСКИЕ СИМПОЗИУМЫ, г. Ростов-на-Дону, 15-20 июня 2014г., Тезисы докладов, с.68-69

5. Немудрый А.А., Маланханова Т.Б., Васькова Е.А., Медведев С.П., Закиян С.М. Редактирование последовательности гена *Aur* крыс линии Brattleboro с помощью системы CRISPR/Cas // Материалы форума «Биомедицина-2016», Новосибирск, 26 июня – 1 июля, 2016.

Структура и объем работы

Диссертация состоит из введения, 3-х глав, выводов и списка литературы. Работа изложена на 104 страницах, содержит 26 рисунков и 12 таблиц.

Благодарности

Работа выполнена в лаборатории эпигенетики развития ИЦиГ СО РАН. Автор выражает глубокую благодарность своему научному руководителю С. М. Закияну за поддержку и помощь при выполнении всех этапов настоящего исследования. Автор искренне благодарит С.П. Медведева и Е.А. Васькову за помощь в проведении экспериментов, анализа их результатов, вдохновение и ценные советы на протяжении всей работы. Также автор благодарит Т.Б. Маланханову за помощь в проведении экспериментов по качественному анализу действия системы CRISPR/Cas. Автор благодарен всему коллективу лаборатории за дружеское участие и практическую помощь в работе.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1 Редактирование генов и геномов с помощью CRISPR/Cas

Система CRISPR/Cas, относительно недавно адаптированная для редактирования геномов эукариот, – первые работы были опубликованы в 2013 году – перспективный инструмент, который обладает большим потенциалом применения практически в любой области биологии, в том числе и в биомедицине при терапии заболеваний, имеющих генетическую основу (Cong et al., 2013). Рассмотрим для начала кратко организацию этой системы у бактерий, механизм её действия, каким именно образом она была оптимизирована для редактирования геномов эукариот, и способы её применения в биомедицине.

1.1.1 Общий принцип организации системы CRISPR/Cas у бактерий

Система CRISPR/Cas бактерий и архей выполняет функцию адаптивного иммунитета и обеспечивает уничтожение чужеродной ДНК при инфекции фаговыми частицами. Деградацию чужеродной ДНК обеспечивает комплекс некодирующих РНК и белков, в котором за узнавание мишени отвечает «направляющая» РНК, по принципу комплементарности связывающая целевую ДНК. Кодирована данная система локусом CRISPR, экспрессирующим некодирующую РНК, и генами *cas* (CRISPR-associated), продуктом которых являются белки Cas, обеспечивающие, в том числе и гидролиз целевой ДНК за счет своей эндонуклеазной активности. Локус CRISPR (**C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeats) представляет собой кластер коротких палиндромных повторов (30-40 п.н.), разделенных участками уникальной ДНК – спейсерами (20-80 п.н.) (Рисунок 1). Спейсеры локуса CRISPR гомологичны последовательностям ДНК различных фагов и различных плазмид (Bolotin et al., 2005) и обеспечивают специфичность действия системы (Barrangou et al., 2007).

У бактерий и архей присутствует значительное разнообразие генов *cas*: при сравнении более 200 геномов бактерий и архей различных видов было выявлено 45 семейств генов *cas*, среди которых только гены *cas1* и *cas2* обнаруживаются повсеместно (Haft et al., 2005). Современная классификация включает пять типов систем CRISPR/Cas (I-V), объединенных в два класса, и больше десяти подтипов и основана на выявленных характерных элементах организации оперонов *cas*.

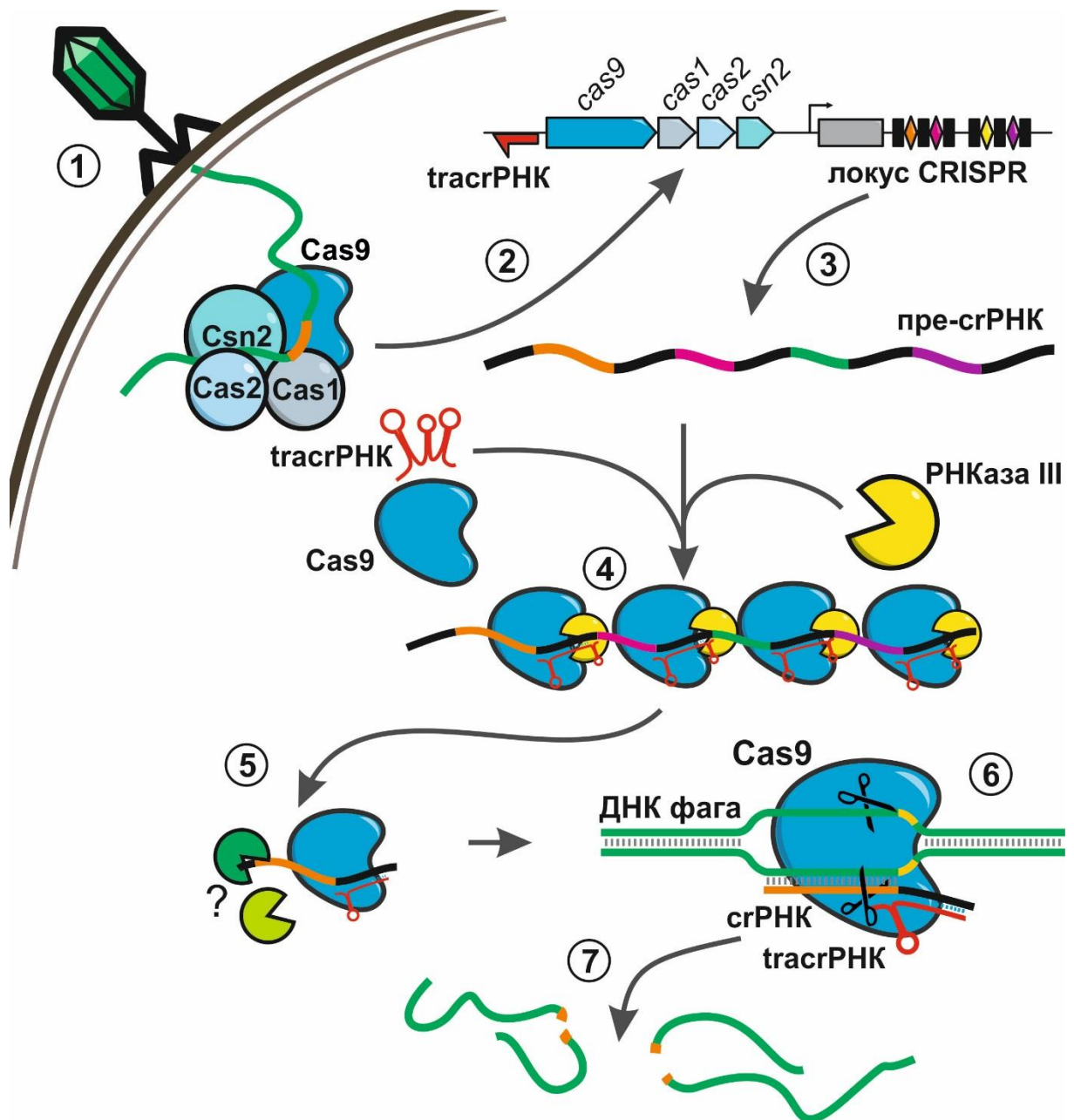


Рисунок 1. Схема действия системы CRISPR/Cas типа II-A. 1) Фаговая инфекция, проникновение чужеродной ДНК. 2) Адаптация – встраивание нового спейсера в локус CRISPR комплексом Cas9-Cas1-Cas2-Csn2 3) Экспрессия длинной некодирующей пре-crRNA. 4-5) Два этапа созревания crRNA. Образуется комплекс пре-crRNA-tracrRNA-Cas9, который разрезается РНКазой III. Образованные более короткие фрагменты, содержащие спейсер и повтор, затем еще укорачиваются с помощью пока неизвестного механизма (5). 6) Образованная зрелая crRNA, содержащая спейсер, направляет комплекс tracrRNA-crRNA-Cas9 к протоспейсеру в чужеродной ДНК. Обязательное условие для систем типа II-A – наличие с 3'-конца от протоспейсера PAM (обозначен желтым). 7) Комплекс tracrRNA-crRNA-Cas9 вносит двуцепочечный разрыв, происходит деградация чужеродной ДНК. На схеме локуса CRISPR прямые повторы обозначены черными прямоугольниками, спейсеры – разноцветными ромбами. Серым обозначена лидерная последовательность. (Nemudryi et al., 2014).

Характерными для систем I, II и III типа являются гены *cas3*, *cas9* и *cas10*, соответственно (Makarova et al., 2011, Makarova et al., 2015). В настоящее время наиболее активно в редактировании геномов используется система CRISPR/Cas типа II-A, поэтому рассмотрим подробнее её организацию и на её примере – механизм адаптивного иммунитета CRISPR/Cas.

1.1.2 Организация и механизм действия систем CRISPR/Cas типа II-A

Отличительной особенностью систем CRISPR/Cas типа II является относительная простота их организации по сравнению с другими типами. Системы этого типа включают всего четыре гена *cas* (три гена в II-C) – общие для подтипов гены *cas1*, *cas2*, *cas9* и гены *csn2* (подтип II-A), *cas4* (подтип II-B) (Chylinski et al., 2014). Важным элементом системы CRISPR/Cas типа II-A является также некодирующая транскрибирующая CRISPR-ПНК (*tracrPНК*), ген которой расположен вблизи *cas*-оперона (Рисунок 1). Преимущество использования именно этой системы для редактирования геномов эукариот заключается в том, что внесение двуцепочечных разрывов (ДЦР) в целевой сайт обеспечивается комплексом некодирующих ПНК и всего одного крупного белка – Cas9. В механизме действия систем CRISPR/Cas бактерий выделяют три отдельных этапа: адаптация, экспрессия и созревание CRISPR-ПНК (*crPНК*), интерференция (деградация чужеродной ДНК). Белок Cas9 в комплексе с *tracrPНК* участвует во всех перечисленных этапах, как в комплексе с другими белками (адаптация), так и независимо (созревание и интерференция).

1.1.2.1 Адаптация

1.1.2.1.1 Выбор протоспейсера

Первый этап – адаптация – приобретение бактериями устойчивости к патогенам, с которыми они встречаются впервые. Этот этап заключается в встраивании чужеродной ДНК – протоспейсера – в эндогенный локус CRISPR уже в виде спейсера, разделяющего две повторенные последовательности. В процессе адаптации выделяют две важные стадии: выбор будущего спейсера и его встраивание в локус CRISPR. Выбор протоспейсера важен для предотвращения аутоиммунного действия системы CRISPR/Cas – деградации собственной ДНК бактерий. Отличие между протоспейсером в чужеродной ДНК и спейсером эндогенном локусе CRISPR

закljučается в наличии короткого мотива, прилежащего к протоспейсеру – PAM. Например, для системы CRISPR/Cas II-A *Streptococcus pyogenes* характерен PAM – NGG (реже –NGA, присутствует вырожденность), прилежащий к 3'-концу протоспейсера в чужеродной ДНК. Как именно происходит выбор протоспейсера (непременно с PAM) и встраивание нового спейсера в системах CRISPR/Cas до сих пор до конца не ясно. Наиболее изучен этот процесс для системы CRISPR/Cas *Escherichia coli* (система I-E): показано, что в адаптации участвуют консервативные среди всех систем CRISPR/Cas белки Cas1 и Cas2, комплекса которых достаточно для выбора и встраивания нового спейсера (Yosef et al., 2012). Белок Cas1 обладает неспецифичной эндонуклеазной активностью, которая является ключевой при приобретении нового спейсера системой CRISPR/Cas. В свою очередь, белок Cas2 так же обладает ДНКазной и РНКазной активностью, которые, однако, не существенны в процессе адаптации – мутации в гене *cas2*, приводящие к нарушению каталитической активности белка, не влияют на адаптацию у *E. coli* (Nunez et al., 2014). В экспериментах по сверхэкспрессии генов *cas*-оперона II-A типа в *E. coli* показано, что совместно с Cas1 преципитируются белки Cas2, Csn2 и Cas9, также были выделены белковые комплексы Cas9-Cas1-Cas2-Csn2. Вдобавок, при сверхэкспрессии только генов *cas1*, *cas2* и *csn2* в штамме *S. aureus* RN4220, геном которого не содержит CRISPR/Cas, встраивания новых спейсеров не происходит, что указывает на важную роль Cas9 в процессе адаптации (Heler et al., 2015). Более того, показано, что в данном процессе нуклеазная активность этого белка не играет роли (в отличие от Cas1): мутации в доменах RuvC и HNH не влияют на адаптацию (Wei et al., 2015). Показано, что при наличии мутаций в доменах Cas9, связывающих PAM, встраивание новых спейсеров происходит, однако соответствующие им протоспейсеры не содержат PAM. Таким образом, по-видимому, в системах CRISPR/Cas типа II-A в адаптации участвует комплекс из четырех белков Cas9-Cas1-Cas2-Csn2, в котором именно Cas9 участвует в выборе будущего спейсера за счет своей способности связывать мотив PAM. При этом в отсутствие *tracrP*НК адаптация нарушается, что говорит о том, что в данном процессе участвует конформационный вариант Cas9 в комплексе с *tracrP*НК (Heler et al., 2015).

Присутствие PAM – важный аспект при выборе протоспейсера, однако он не дает ответ на вопрос: как при адаптации система CRISPR/Cas отличает чужеродную

ДНК от эндогенной? Мотив –NGG, конечно, встречается и в собственной ДНК бактерий, и, если такой спейсер (из собственной ДНК) будет встроен в локус CRISPR, бактерия с большой вероятностью погибнет. Накопление случайных мутаций в протоспейсерах, спейсерах или PAM могут позволить бактерии избежать гибели, с другой стороны было показано, что существует механизм обеспечивающий избирательность CRISPR/Cas в отношении источника встраиваемых спейсеров (Levy et al., 2015, Stern et al., 2010). В работах по изучению адаптации в системе CRISPR типа I-E, показано, что при сверхэкспрессии генов *cas1* и *cas2*, отвечающих за адаптацию, но не генов, участвующих в интерференции (риска аутоиммунного действия нет), происходит накопление предпочтительно спейсеров из чужеродной ДНК (Yosef et al., 2012). На основе результатов полученных при исследованиях на *E. coli* (CRISPR I-E) была предложена модель, которая объясняет предпочтительный выбор в протоспейсеров из экзогенной ДНК. Согласно этой модели, материалом для встраивания в локус CRISPR могут служить фрагменты ДНК, образующиеся при репарации комплексом RecBCD ДЦР, которые возникают при репликации ДНК.

Элементы комплекса RecBCD связываются с линейной двухцепочечной ДНК, образующейся при разрывах (Dillingham, Kowalczykowski, 2008), после чего происходит деградация концов, которая останавливается, когда комплекс достигает Chi-сайта – специфичной последовательности ДНК из восьми нуклеотидов (5'-GCTGGTGG-3'). При сверхэкспрессии в *E. coli* штамма BL21-AI (удалены гены *cas*, участвующие в интерференции) генов *cas1* и *cas2* с помощью глубоко секвенирования новоприобретенных спейсеров показано, что протоспейсеры чаще всего расположены между местами возникновения ДЦР и Chi-сайтами. При этом в штаммах, из генома которых были удалены гены *recB*, *recC* или *recD*, снижается количество новоприобретенных спейсеров, снижается способность к адаптации (при этом доля спейсеров из эндогенной ДНК возрастает 10-кратно) (Levy et al., 2015). Интересно, что при удалении генов *recB* или *recC*, комплекс полностью неактивен, в отличии от ситуации, когда в отсутствии *recD* образуется комплекс RecBC, имеющий геликазную, но не нуклеазную активность. Так как удаление гена *recD* приводит к нарушению адаптации ясно, что в этом механизме важную роль играет именно нуклеазная активность RecBCD. Тот факт, что количество

новоприобретенных спейсеров снижается, но не полностью, предполагает, что фрагменты ДНК, образующиеся при функционировании RecBCD, не единственный источник потенциальных спейсеров. Отличие между эндогенной ДНК *E. coli* и чужеродной может заключаться в частоте расположения Chi-сайтов – в геноме *E. coli* такие сайты встречаются в среднем каждые 4,6 т.п.н., примерно в 14 раз чаще, чем можно было бы ожидать, если бы эти сайты возникали, как случайная комбинация нуклеотидов (Dillingham, Kowalczykowski, 2008). Соответственно, при связывании комплекса RecBCD с линейной чужеродной ДНК (например, ДНК фага при инфекции) участок до Chi-сайта, который будет деградирован, значительно больше, чем при репарации эндогенной ДНК, следовательно, образуется больше фрагментов, среди которых будут выбраны протоспейсеры. Эта модель также объясняет предпочтение к выбору протоспейсеров на высококопийных плазмидах, хоть эти молекулы ДНК и кольцевые. Предполагается, что в клетке присутствует большое число копий плазмиды, соответственно, число репликативных вилок в совокупности на всех копиях плазмиды больше, чем на хромосомной ДНК бактерии. Соответственно, частота ошибок, приводящих к ДЦР, при репликации плазмидной ДНК выше, как следствие (так как на плазмиде вовсе может не встречаться Chi-сайт) число фрагментов, образованных при деградации плазмидной ДНК комплексом RecBCD, является подавляющим. Таким образом, вероятность выбора протоспейсера из плазмидной ДНК значительно выше, чем из эндогенной (Levy et al., 2015).

Описанный выше механизм был обнаружен только в системах CRISPR/Cas типа I-E *E. coli*, и неизвестно, насколько он консервативен среди систем других типов. При исследованиях на *S. thermophilus* (II-A) показано, что при отсутствии нуклеазной активности Cas9 в процессе адаптации происходит предпочтительный выбор протоспейсеров из эндогенной ДНК (Wei et al., 2015). Этот факт говорит о возможном наличии альтернативного механизма защиты эндогенной ДНК, в котором, по-видимому, играет роль нуклеазная активность Cas9. С другой стороны, так как Cas9 участвует и в интерференции, её подавление позволяет исключить аутоиммунное действие и, как следствие, выжить клеткам, встроившим протоспейсеры из собственного генома (может альтернативного механизма нет в

принципе). В этой же ситуации, несмотря на подавление интерференции, в системах типа I-E предпочтение к экзогенной ДНК сохраняется (Amitai, Sorek, 2016).

1.1.2.1.2 Встраивание протоспейсера в локус CRISPR

После выбора протоспейсера происходит его встраивание в локус CRISPR (протоспейсер становится спейсером). Что происходит после выбора спейсера и до его встраивания в геном до сих пор неизвестно. В настоящее время для системы II-A данных не получено, с другой стороны при исследовании этого процесса у *E. coli* была предложена модель встраивания нового спейсера комплексом Cas1-Cas2 (Рисунок 2) со стороны лидерной АТ-богатой последовательности, что сопровождается дупликацией прямого повтора, расположенного на 5'-конце локуса CRISPR (Arslan et al., 2014). Можно предполагать, ввиду консервативности Cas1 и Cas2, схожий механизм, однако неясно участвуют ли в этом процессе Csn2 и Cas9, или же их участие ограничено выбором будущего спейсера.

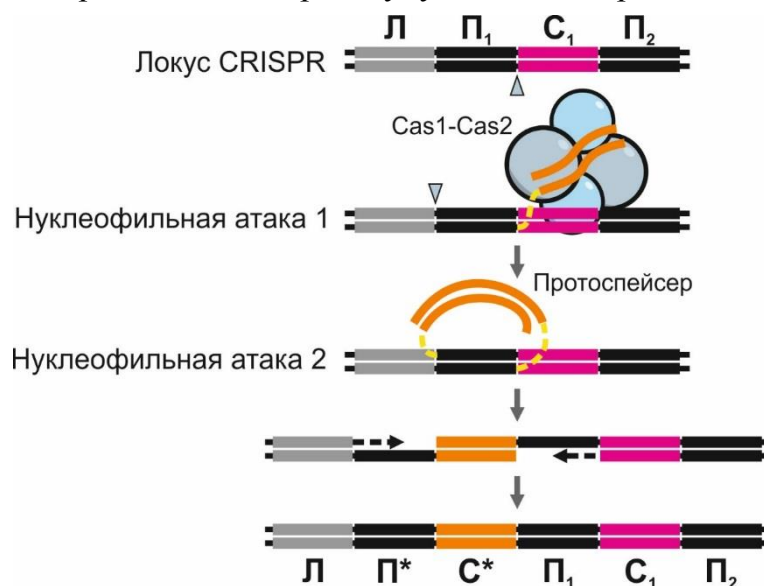


Рисунок 2. Модель встраивания нового спейсера в системах CRISPR типа I-E комплексом Cas1-Cas2. Комплекс Cas1-Cas2 катализирует нуклеофильную атаку 3'-ОН группы одной из цепей протоспейсера на 5'-конец первого повтора в локусе (П₁), при этом образуется промежуточная структура – одна цепь протоспейсера лигирована с одной цепью локуса. Затем 3'-ОН группа другой цепи протоспейсера атакует комплементарную цепь повтора на стыке с лидерной последовательностью (Л). В результате происходит дупликация повтора (П*) и встраивание нового спейсера (С*). Образовавшиеся участки одноцепочечной ДНК достраиваются пока неизвестными ферментами (Arslan et al., 2014, Marraffini, 2015). С – спейсеры, П – повторы, Л – лидерная последовательность. Треугольниками обозначены нуклеофильные атаки.

1.1.2.2 Созревание crРНК

В результате транскрипции локуса CRISPR образуется длинная некодирующая пре-crРНК, состоящая из повторов и уникальных последовательностей. На первом этапе происходит разрезание пре-crРНК на более короткие фрагменты (Рисунок 1). В этом процессе в системах II-A типа участвуют Cas9, tracrРНК и РНКазы III бактериальной клетки. На длинной пре-crРНК образуются комплексы с tracrРНК, которая содержит участок на 5'-конце (25 нуклеотидов у *S. pyogenes*), комплементарный повторам локуса CRISPR (антиповтор). Анализ вторичной структуры tracrРНК предполагает наличие трех шпилек на 3'-конце, из которых две необходимы для связывания с белком Cas9, который стабилизирует комплекс tracrРНК-пре-crРНК (Deltcheva et al., 2011, Karvelis et al., 2013). РНКазы III разрезают двухцепочечную РНК комплекса tracrРНК-пре-crРНК, с образованием более коротких фрагментов, которые затем еще раз подвергаются процессингу в комплексе с tracrРНК и Cas9 по до сих пор неизвестному механизму. В результате образуются зрелые crРНК, которые у *S. pyogenes* состоят из 20 нуклеотидов спейсера, расположенных с 5'-конца, и на 3'-конце 19-22 нуклеотидов повтора (Deltcheva et al., 2011).

1.1.2.3 Интерференция

Процесс интерференции – деградации чужеродной ДНК, содержащей протоспейсеры, – в системах II типа обеспечивается комплексом crРНК, отвечающей за узнавание мишени, tracrРНК и Cas9. Белок Cas9 при этом отвечает за узнавание PAM и непосредственно внесение ДЦР. Данные кристаллографического анализа комплексов crРНК-tracrРНК-Cas9 и протоспейсер-crРНК-tracrРНК-Cas9 позволили изучить механизм этого процесса (Anders et al., 2014, Jinek et al., 2014). Показано, что белок Cas9 состоит из двух долей: узнающей (recognizing, REC) и нуклеазной (NUC), которые соединены двумя линкерами. Нуклеазная доля состоит из составного (из трех частей) нуклеазного домена RuvC и нуклеазного домена HNH. Была предложена модель, согласно которой апо-Cas9, не связанный с tracrРНК и crРНК, не активен. При образовании комплекса Cas9-tracrРНК-crРНК (связывание происходит с долей REC) происходят конформационные изменения (холо-Cas9): доли поворачиваются относительно друг друга, в результате чего между ними открывается бороздка, которая связывает ДНК. При связывании протоспейсера с

crPHK происходит образование R-петли, что также приводит к конформационным изменениям Cas9: NUC еще поворачивается относительно REC, нуклеазные домены RuvC и HNH приближаются к целевой ДНК и вносят разрыв (Рисунок 3) (Jinek et al., 2014).

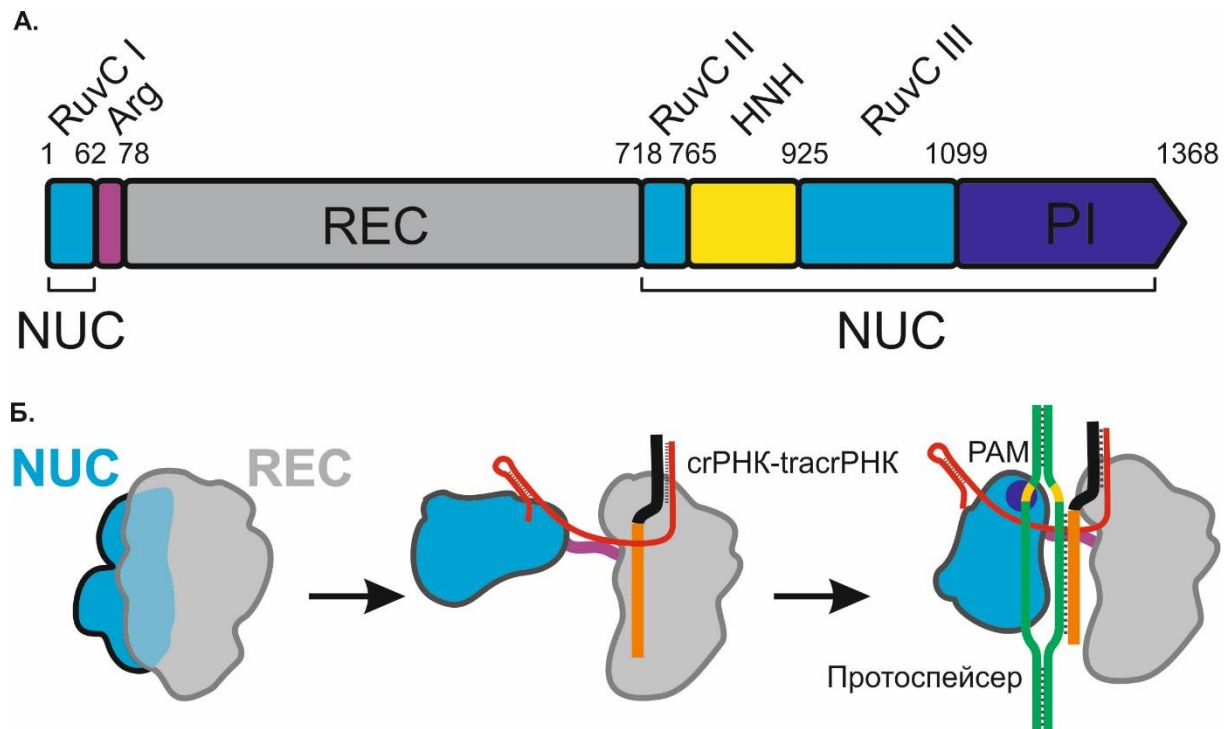


Рисунок 3. А - Доменная организация белка Cas9 *S. pyogenes*. Доля NUC состоит из нуклеазных доменов RuvC и HNH и домена, взаимодействующего с PAM (PI). Доли NUC и REC соединены двумя линкерами: аргинин-богатым участком (Arg) и участком на С-конце доли REC (714-717). Б – схематическое изображение конформационных изменений, приводящих к активации Cas9 (Jinek et al., 2014).

Таким образом, процесс поиска мишени и внесения ДЦР выглядит следующим образом: 1) образуется комплекс Cas9-tracrPHK-crPHK, в результате чего Cas9 активируется, 2) холо-Cas9 «сканирует» ДНК 3) домен PI-Cas9 узнает PAM 4) за счет геликазной активности Cas9 происходит расплетание участка ДНК с 5'-конца от PAM 5) в случае комплементарности с crPHK происходит образование гетеродуплекса протоспейсер-crPHK 6) при успешной гибридизации первых восьми нуклеотидов с 3'-конца протоспейсера с crPHK происходит дальнейшая гибридизация в 5'-направлении 7) При успешном связывании протоспейсер-crPHK нуклеазные домены HNH и RuvC вносят разрывы: RuvC в вытесненную

некомплементарную цепь, HNH в цепь, с которой комплементарно связывается crPНК (Anders et al., 2014, Sternberg et al., 2014).

1.1.3 Адаптация системы CRISPR/Cas9 для редактирования геномов эукариот

Впервые возможность использовать CRISPR/Cas для направленной модификации генома клеток млекопитающих была продемонстрирована в 2013 году в работах коллективов Фенга Чжана (Feng Zhang) и Джорджа Чёрча (George Church), которые были опубликованы в одном номере журнала Science (Cong et al., 2013, Mali et al., 2013). Коллективом Фенга Чжана был разработан ряд плазмидных векторов, размещенных в депозитории Addgene, для экспрессии элементов, необходимых для функционирования системы CRISPR/Cas. Данные векторы содержат следующие элементы: 1) ген *cas9* *S. pyogenes* (*SpCas9*), кодоны которого оптимизированы для экспрессии в клетках человека (*hSpCas9*), 2) химерную единую направляющую РНК (sgPНК). Химерная sgPНК – единая молекула РНК, вторичная структура которой имитирует комплекс зрелой crPНК и tracrPНК за счет структуры стебель-петля (повтор-антиповтор в комплексе crPНК-tracrPНК). Последовательность спейсера при этом клонируют в вектор в ген sgPНК с помощью стандартных методов генной инженерии (Рисунок 4). Исследователь выбирает протоспейсер из 20 нуклеотидов, к которому прилежит PAM –NGG, для действия CRISPR/Cas в целевом гене, после чего синтезируют олигонуклеотиды, содержащие последовательность выбранного протоспейсера, таким образом, что после гибридизации будут образовываться «липкие» концы, комплементарные «липким» концам в векторе pX330 после гидролиза эндонуклеазой рестрикции *BbsI* (Рисунок 4Б). В результате клонирования последовательности 20 нуклеотидов в ген sgPНК в векторе pX330 обеспечивается специфичность узнавания протоспейсера в геноме, к которому sgPНК «направляет» комплекс sgPНК-Cas9 (Cong et al., 2013).

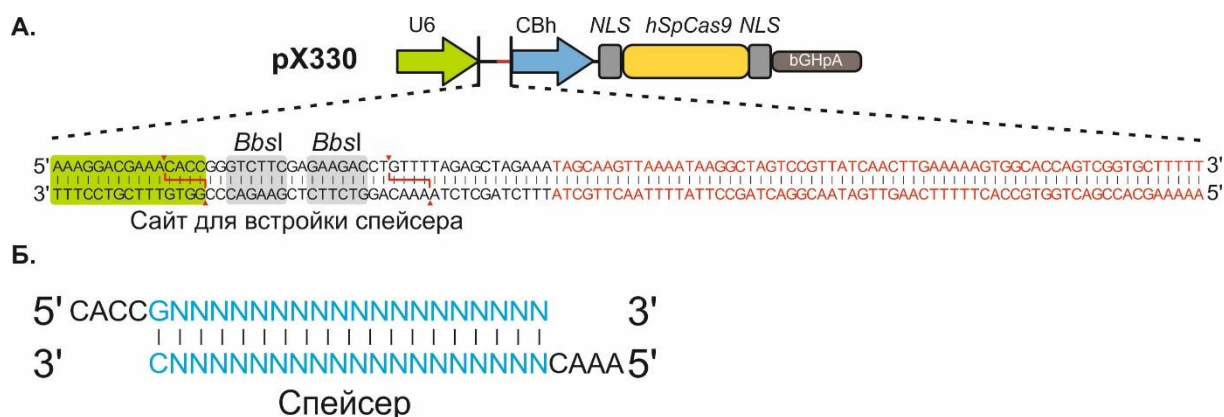


Рисунок 4. А - Схема вектора pX330-U6-Chimeric_BB-CBh-hSpCas9, кодирующего элементы адаптированной системы CRISPR/Cas (Cong et al., 2013). U6 – промотор для транскрипции гена sgPHK (PHK полимеразой III), CBh – конститутивный гибридный промотор на основе промотора β -актина курицы, NLS – сигнал ядерной локализации, hSpCas9 – кодон-оптимизированный ген *cas9* *S. pyogenes*, bGHpA – сигнал полиаденилирования. Б – Последовательность двухцепочечного нуклеотида для встройки в вектор pX330-U6-Chimeric_BB-CBh-hSpCas9, N – любой нуклеотид.

1.1.4 Спектр возможных модификаций при использовании системы CRISPR/Cas9 для редактирования геномов

Система CRISPR/Cas, адаптированная для использования в клетках эукариот, является инструментом, позволяющим вносить ДЦР в выбранные исследователем сайты генома. Как и любой инструмент, CRISPR/Cas9 может быть применена различными способами, зависящими от поставленных целей. От того, каким способом будет репарирован внесенный разрыв, зависит характер модификаций, которые могут быть внесены в геном. Рассмотрим в общих чертах современные представления о путях репарации ДЦР у млекопитающих, и как их результат используется в исследованиях по редактированию генов и геномов с помощью CRISPR/Cas9.

1.1.4.1 Репарация двуцепочечных разрывов в клетках млекопитающих

В настоящее время известны четыре пути репарации двуцепочечных разрывов ДНК у млекопитающих (Рисунок 5): 1) классическое соединение негомολогичных концов (с-NHEJ), 2) гомологичная рекомбинация (HR) 3) альтернативное соединение концов (alt-EJ) 4) отжиг одиночной цепи (SSA). Кратко рассмотрим, как происходит выбор пути репарации ДЦР у млекопитающих на примере человека. Тип

концов ДНК, которые образуются в результате ДЦР, при этом играет определяющую роль. В случае «тупых» концов репарация происходит по пути c-NHEJ, однако, если ДНК после ДЦР будет гидролизована с образованием одноцепочечных выступающих 3'-концов, репарация будет происходить с помощью HR, alt-EJ или SSA (Ceccaldi et al., 2016).

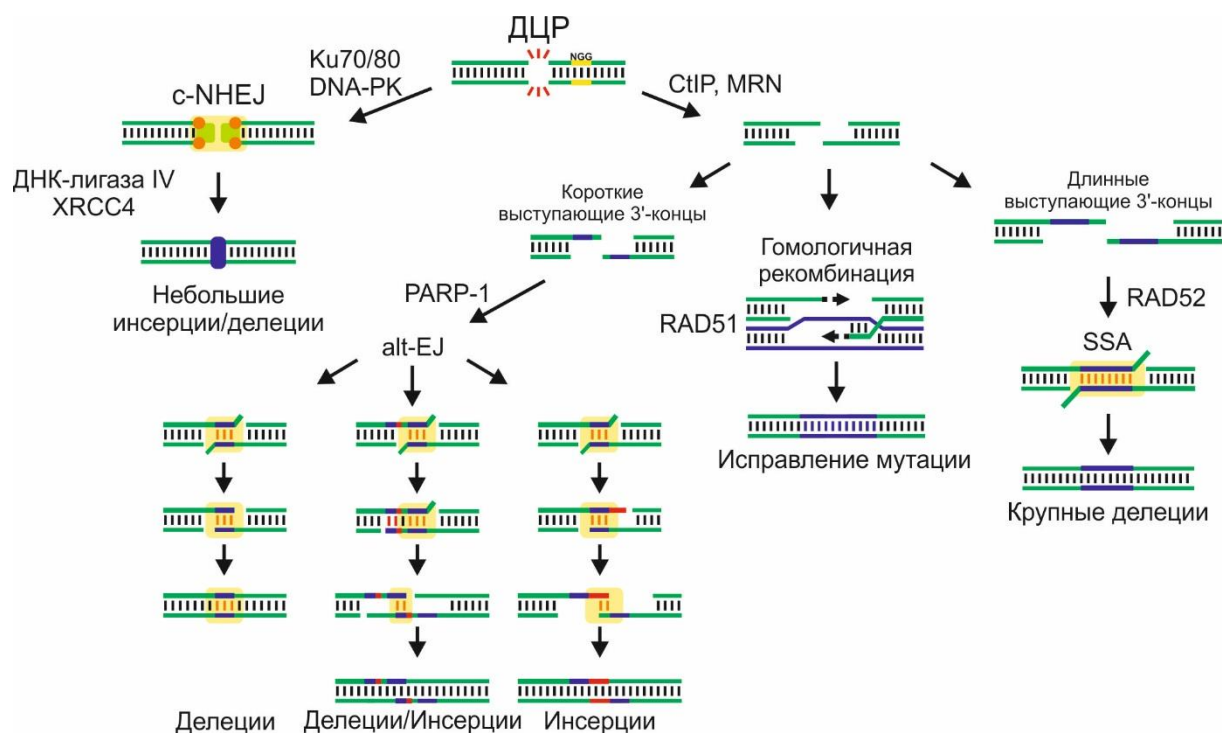


Рисунок 5. Пути репарации ДЦР у млекопитающих. Описание в тексте.

Репарация с помощью c-NHEJ может происходить в любой фазе клеточного цикла (доминантный путь репарации ДЦР в фазах G0/G1 и G2, (Karanam et al., 2012)) и благодаря высокой скорости играет важную роль в сохранении целостности генома и, в частности, предотвращении хромосомных перестроек (Difilippantonio et al., 2000). Гетеродимер Ku70/Ku80 связывается с концами ДНК после ДЦР и защищает их от гидролиза. Далее к разрыву привлекаются ДНК-зависимая протеникиназа (DNA-PK), комплекс ДНК-лигазы IV и XRCC4, фактор Artemis и другие белки репарационного комплекса. В результате происходит соединение концов, при этом могут возникать ошибки: небольшие инсерции или делеции (1-4 нуклеотида), которые являются следствием полимеразной или нуклеазной активностей белков репарационного комплекса (Lieber, 2010).

Если ДЦР не репарируется в течение достаточно длительного времени, к концам ДНК привлекается комплекс MRN (MRE11, RAD50 и NBS1). Происходит связывание комплекса MRN и белка CtIP, что приводит к образованию коротких выступающих 3'-концов (около 20 нуклеотидов у млекопитающих) за счет 3'-5' экзонуклеазной активности MRE11 (Truong et al., 2013). Образование коротких выступающих концов направляет репарацию по путям HR и alt-EJ. Также далее могут образовываться более длинные выступающие концы под действием DNA2, BLM, WRN, CtIP и EXO1, следствием чего может стать репарация по пути SSA. Образование выступающих концов зависит от фазы клеточного цикла: в фазах S/G2 циклин-зависимые протеинкиназы (CDK) фосфорилируют CtIP (Yun, Hiom, 2009) и EXO1 (Robert et al., 2011), что способствует их образованию и, как следствие, выбору пути репарации – репарация выступающих концов с помощью c-NHEJ уже не происходит, и дальше выбор происходит между путями HR, alt-EJ и SSA. Стоит отметить, что зависимость от CDK обуславливает большую частоту HR в активно пролиферирующих клетках.

При HR ключевую роль играет белок RAD51. С новообразованными выступающими одноцепочечными концами связывается белок RPA, защищая их от разрушения, который затем с помощью фактора BRCA2 заменяется на RAD51, что в результате приводит к образованию нуклеопротеинового филамента, инвазии цепи и образованию гетеродуплексного участка с сестринской хромосомой, образованию структур Холлидея и их дальнейшему разрешению. В результате данного процесса происходит безошибочная репарация участка, на котором возник ДЦР, за счет восстановления поврежденной в результате ДЦР последовательности по матрице гомологичной ДНК – участку сестринской хромосомы (Goodarzi, Jeggo, 2013).

Механизмы репарации ДЦР по путям alt-EJ и SSA основаны на гибридизации образованных одноцепочечных концов в случае: 1) микрогомологий между концами – 3-5 нуклеотидов (alt-EJ), 2) крупных участков гомологий – более 20 нуклеотидов, например, в случае повторенных последовательностей (SSA). Предположительно, эти пути, репарация по которым происходит с ошибками, являются «запасными» по отношению к HR. Репарация по пути alt-EJ происходит в S-фазе клеточного цикла и конкурирует с HR (Mateos-Gomez et al., 2015). Ключевым фактором alt-EJ является PARP1, который привлекает остальные белки репарационного комплекса, в

частности, транслезионные ДНК-полимеразы, одна из которых, POLQ, способна ингибировать HR, нарушая формирование нуклеопротеинового филамента (Ceccaldi et al., 2015). В результате репарации по этому пути происходят делеции между сайтами микрогомологии (остается только один сайт), однако возможны и инсерции (Рисунок 5).

В настоящее время известно, что в репарации по пути SSA участвуют белки RAD52 и ERCC4–ERCC10, остальные же факторы пока не изучены (Ceccaldi et al., 2016). Репарация по пути SSA, по-видимому, также конкурирует с HR: например, у дрожжей Rad51 предотвращает гибридизацию выступающих концов под действием Rad52 (Wu et al., 2008). Результатом репарации по пути SSA является образование крупных делеций.

1.1.4.2 Использование системы CRISPR/Cas9 совместно с донорной ДНК

В нормальных условиях в качестве донорной молекулы при репарации ДЦР с помощью HR выступает сестринская хромосома, однако в качестве донорной молекулы для HR репарационным комплексом может быть использована и ДНК, искусственно введенная в клетку, например, плазмидная ДНК, если она содержит участки гомологии с целевым сайтом генома (Cong et al., 2013). При трансфекции в клетку плазмидных векторов они будут конкурировать с сестринской хромосомой, и, более того, так как количество копий плазмидной ДНК выше на многие порядки, они будут получать преимущество. Таким образом, последовательность генома будет репарирована с использованием в качестве «образца» донорной ДНК, внесенной в клетку исследователем. В настоящее время в основном используют два типа донорных молекул: плазмидную ДНК и одноцепочечные олигонуклеотиды (ssODN).

Донорные плазмидные векторы содержат следующие элементы: 1) «плечи гомологии» – участки гомологичные целевому участку генома, которые в случае исправления мутации содержат, в отличие от мутантного гена, последовательность дикого типа, 2) элементы, необходимые для селекции – гены устойчивости к антибиотикам, репортерные гены. Плечи гомологии фланкируют эти дополнительные элементы, и при HR они встраиваются в целевой локус, обеспечивая возможность селекции клеток, в которых произошли целевые события. Показано, что оптимальная длина плеч гомологии составляет 2-3 т.п.н. (Beumer et

al., 2013). Показано, что эффективность гомологичной рекомбинации выше, если ДЦР, индуцируемый CRISPR/Cas, происходит в пределах «левого» (5') плеча гомологии вблизи его 3'-конца (Zhang et al., 2014).

Для другого вида донорных молекул – ssODN – рекомендуются следующие параметры: оптимальная длина составляет 90 нуклеотидов, при этом сайт ДЦР должен быть расположен близко к середине участка гомологичного ssODN (McVey, Lee, 2008). При этом показано, что модификация геномов с антисмысловыми олигонуклеотидами проходит несколько более эффективно (Lieber, 2010). В настоящее время до конца неясно, по какому механизму происходит репарация ДЦР в присутствии ssODN (van Overbeek et al., 2016)

Таким образом, применение CRISPR/Cas для внесения ДЦР в целевой участок генома одновременно с доставкой в клетку донорной ДНК может быть использовано для замены последовательности в геноме (на соответствующую последовательность донорной ДНК) с целью внесения определенных мутаций или, наоборот, исправления мутаций. Также с помощью HR может быть проведена направленная встройка трансгена в определенное место генома, например, чтобы избежать риска мутагенеза, который возможен при случайной интеграции.

Частота HR существенно ниже с-NHEJ, поскольку в клетке это процесс протекает только в поздней S-фазе и на границе S- и G2-фазы, вдобавок NHEJ кинетически протекает быстрее (Goodarzi, Jeggo, 2013). Однако, при использовании различных способов селекции с последующим анализом клональных линий возможно выявить линии, образовавшиеся из клеток, в которых произошла HR.

1.1.4.3 Использование системы CRISPR/Cas9 без донорной ДНК

Репарация ДЦР, внесенного CRISPR/Cas, преимущественно происходит по пути с-NHEJ. При этом с большой частотой происходит внесение в месте разрыва небольших делеций или инсерций. На этом факте основано применение CRISPR/Cas для нокаута генов – небольшие делеции и инсерции приводят к сдвигу рамки считывания. С момента появления этой системы в открытом доступе в 2013 году, CRISPR/Cas была использована для нокаута генов как *in vitro*, так и *in vivo* на различных модельных организмах (Nemudryi et al., 2014).

При внесении двух ДЦР возможно удаление участков генома. Причем это могут быть как сравнительно небольшие участки, например, экзон гена *Dmd* у мыши

(Long et al., 2016, Nelson et al., 2016, Tabebordbar et al., 2016), так и крупные участки – десятки тысяч нуклеотидов и больше (Zhang et al., 2015).

С помощью двух ДЦР возможно вызвать и мутации на хромосомном уровне. В работе по моделированию на мышах рака легкого с помощью CRISPR/Cas *in vivo* с помощью двух ДЦР была индуцирована инверсия участка 17 хромосомы длиной 11 миллионов пар нуклеотидов, в результате чего образовался известный ранее онкоген *Eml4-Alk*, что привело к развитию опухолей (Maddalo et al., 2014).

1.1.5 Нецелевые эффекты системы CRISPR/Cas9

При редактировании геномов в терапевтических целях крайне важно минимизировать риски нецелевых эффектов. При использовании системы CRISPR/Cas9, узнавание мишени в которой достигается за счет связывания по принципу комплементарности между sgРНК и протоспейсером в геноме, возможны нецелевые эффекты в результате связывания с сайтами генома, отличающихся от канонического протоспейсера однонуклеотидными заменами или небольшими инсерциями и делециями (1-2 нуклеотида) (Fu et al., 2013, Ran et al., 2015). Соответственно, чем больше замен в сайте, тем меньше вероятность возникновения нецелевых ДЦР.

От типа исследования и цели, которая в нем поставлена, зависит значимость риска возникновения нецелевых разрывов: при использовании методов редактирования геномов *in vivo* для терапии заболеваний человека, необходимо свести риск к минимуму. С другой стороны, при исследованиях *in vitro* возможные нецелевые эффекты могут быть несущественны. Далее мы рассмотрим, каким образом могут быть минимизированы нецелевые эффекты для различных типов исследований.

1.1.5.1 Выбор целевого сайта в геноме

Важным моментом при дизайне эксперимента по редактированию генома является тщательный выбор целевого сайта для действия искусственных нуклеаз. Например, природа узнавания комплексом sgРНК:Cas9 протоспейсера в геноме позволяет предсказать потенциальные нецелевые сайты. Соответственно, если при выборе целевых сайтов избегать повторенных последовательностей, участков гомологии между генами, то возможно значительно снизить риск

нецелевых эффектов. После «отсева» наиболее очевидных потенциальных нецелевых сайтов (например, в паралогичных генах), возможно определить потенциальные нецелевые сайты в геноме и относительную вероятность внесения в них ДЦР с помощью биоинформатических методов (Таблица 1).

Таблица 1. Сравнение биоинформатических инструментов для предсказания нецелевых сайтов CRISPR/Cas. По материалам (Ishida et al., 2015)

Название	Организм	Тип последовательности	Характер нецелевых сайтов	Ссылка
ZiFiT	<i>Hs, Rn, Mm, Dr, Dm, Ce, Aa, Ec</i>	Целевая последовательность (<1000 п.н.)	Однонуклеотидные замены	(Sander et al., 2007, Sander et al., 2010)
CRISPR Design	>15 видов	Целевая последовательность (<250 п.н.)	Однонуклеотидные замены	(Hsu et al., 2013)
Cas9 design	>10 видов	Целевая последовательность (>10 т.п.н.)	Однонуклеотидные замены	(Ma et al., 2013)
E-CRISP	>15 видов	Целевая последовательность (>10 т.п.н.), название гена	Однонуклеотидные замены	(Heigwer et al., 2014)
CasOT	Любые виды	Целевая последовательность (>10 т.п.н.)	Однонуклеотидные замены	(Xiao et al., 2014)
Cas-OFFinder	>10 видов	sgPHK (10-25 п.н.)	Однонуклеотидные замены, инсерции, делеции	(Bae et al., 2014)
CHOPCHOP	>20 видов	Целевая последовательность (>10 т.п.н.)	Однонуклеотидные замены	(Montague et al., 2014)
GT-Scan	>20 видов	Целевая последовательность (<4000 п.н.)	Однонуклеотидные замены	(O'Brien, Bailey, 2014)
sgRNAcas9	Любые виды	Целевая последовательность (>10 т.п.н.)	Однонуклеотидные замены	(Xie et al., 2014)
CRISPR-P	>20 видов растений	Целевая последовательность (<5 т.п.н.)	Однонуклеотидные замены	(Lei et al., 2014)
COSMID	<i>Hs, Rn, Mm, Ce, Mam, Dr</i>	sgPHK (10-55 п.н.)	Однонуклеотидные замены, инсерции, делеции	(Cradick et al., 2014)
sgRNA Designer	<i>Hs, Mm</i>	Целевая последовательность (<10 т.п.н.), ID гена	-	(Doench et al., 2014)
iGEATs	<i>Hs, Mm</i>	Хромосомный локус, целевая последовательность (<25 т.п.н.), название гена, sgPHK	-	(Li et al., 2015)
CRISPRdirect	>15 видов	Целевая последовательность (<10 т.п.н.)	Однонуклеотидные замены, инсерции, делеции	(Naito et al., 2015)
CRISPR-ERA	<i>Hs, Rn, Mm, Dr, Dm, Ce, Sc, Ec, Bs</i>	Целевая последовательность (<5 т.п.н.)	Однонуклеотидные замены	(Liu et al., 2015)
Protospacer workbench	Любые виды	Целевая последовательность (>10 т.п.н.)	Однонуклеотидные замены	(MacPherson, Scherf, 2015)
<i>Hs: Homo sapiens, Rn: Rattus norvegicus, Mm: Mus musculus, Dr: Danio rerio, Dm: Drosophila melanogaster, Aa: Aedes aegypti, Ec: Escherichia coli, Sc: Saccharomyces cerevisiae, Bs: Bacillus subtilis, Ce: Caenorhabditis elegans, Mam: Macaca mulatta</i>				

Данные инструменты используют разные алгоритмы и при их сравнении было показано, что, хотя определенные инструментами протоспейсеры для CRISPR/Cas9

на заданном участке генома совпадают, количество выявленных потенциальных нецелевых сайтов может различаться (Ishida et al., 2015). Также стоит отметить, что большинство этих инструментов выявляет нецелевые сайты, содержащие однонуклеотидные замены относительно канонического сайта, однако в нескольких работах было показано, что нецелевые ДЦР могут происходить также в сайтах, содержащих инсерции или делеции (Bae et al., 2014, Lin et al., 2014, Ran et al., 2015). Из приведенных инструментов только Cas-OFFinder, COSMID и CRISPRdirect позволяют обнаружить нецелевые сайты, содержащие инсерции и делеции относительно канонического сайта.

Соответственно, снизить токсичность CRISPR/Cas и риски мутагенеза возможно уже на первых этапах эксперимента, выбрав целевой сайт, для которого предсказанная вероятность нецелевых эффектов будет минимальна.

1.1.5.2 Увеличение специфичности Cas9

В клетке комплекс sgРНК:SpCas9 «сканирует» ДНК в поисках целевого сайта. В первую очередь происходит узнавание участком Cas9 (R1333 и R1335 с NGG у SpCas9) последовательности PAM, затем Cas9 разделяет две цепи ДНК, чтобы sgРНК могла связаться с протоспейсером по принципу комплементарности, и, наконец, домены RuvC и HNH Cas9 вносят разрывы в цепи ДНК (Anders et al., 2014, Sternberg et al., 2014). Знания, полученные при исследованиях кристаллической структуры комплексов sgРНК:Cas9:протоспейсер, позволили нескольким группам ученых разработать стратегии для увеличения специфичности внесения ДЦР путем получения мутантных вариантов SpCas9. Одно из направлений – увеличение специфичности узнавания SpCas9 PAM, поскольку известно, что SpCas9 способна вносить ДЦР, хоть и с меньшей вероятностью, в протоспейсеры, PAM которых отличается от канонического –NGG (-NGA, например) (Zhang et al., 2014). В 2015 году, был получен мутантный вариант SpCas9 (D1135E) более специфично узнающий канонический PAM –NGG. Также этим коллективом получены другие мутантные варианты, узнающие измененные последовательности PAM (-NGAN, например) (Kleinstiver et al., 2015). В принципе, создавая мутантные варианты Cas9, которые специфично узнают более длинные последовательности PAM, возможно снизить риск нецелевых эффектов.

С другой стороны к вопросу увеличения специфичности комплекса sgРНК:Cas9 подошли исследователи из коллектива Фенга Чжана (Slaymaker et al., 2016). Анализ кристаллической структуры комплекса sgРНК:Cas9:протоспейсер показал наличие бороздки, содержащей 32 положительно заряженных аминокислоты между доменами RuvC, NHN и доменом PI (Рисунок 6). После того, как в результате геликазной активности Cas9 цепи целевого участка ДНК расплетаются, цепь, содержащая протоспейсер и РАМ, стабилизируется положительно заряженной бороздкой, а комплементарная ей цепь образует гетеродуплекс с sgРНК (Anders et al., 2014, Nishimasu et al., 2014). Авторами была выдвинута гипотеза, что, если снизить суммарный положительный заряд описанной выше бороздки Cas9, то связывание бороздкой нецелевой цепи ДНК станет слабее, и она будет конкурировать с sgРНК при гибридизации. Если sgРНК при этом будет комплементарна не полностью (в случае нецелевых сайтов, содержащих замены, инсерции или делеции), то преимущество получит комплементарная цепь ДНК и нецелевого внесения ДЦР происходить не будет. Для проверки выдвинутой гипотезы авторами был получен набор мутантных вариантов SpCas9, в каждом из которых одна из положительно заряженных аминокислот бороздки была заменена неполярной аминокислотой. Показано, что 5 из 32 одиночных мутаций привели к снижению частоты нецелевых ДЦР в геноме клеток линии НЕК при использовании sgРНК, комплементарной участку гена *EMX1*. Далее авторами были проанализированы все возможные комбинации этих 5 одиночных замен и выявлены те, которые максимально снижают частоту нецелевых ДЦР, при этом не снижая эффективности действия в целевых сайтах. По результатам анализа были выбраны два мутантных варианта – SpCas9 (K855A) и eSpCas9(1.1) (SpCas9 с мутациями K848A/K1003A/R1060A), которые при дальнейшем анализе показали достоверное снижение частоты или полную элиминацию нецелевых эффектов без снижения целевой активности в 24 сайтах генома, расположенных в 10 различных локусах (Slaymaker et al., 2016).

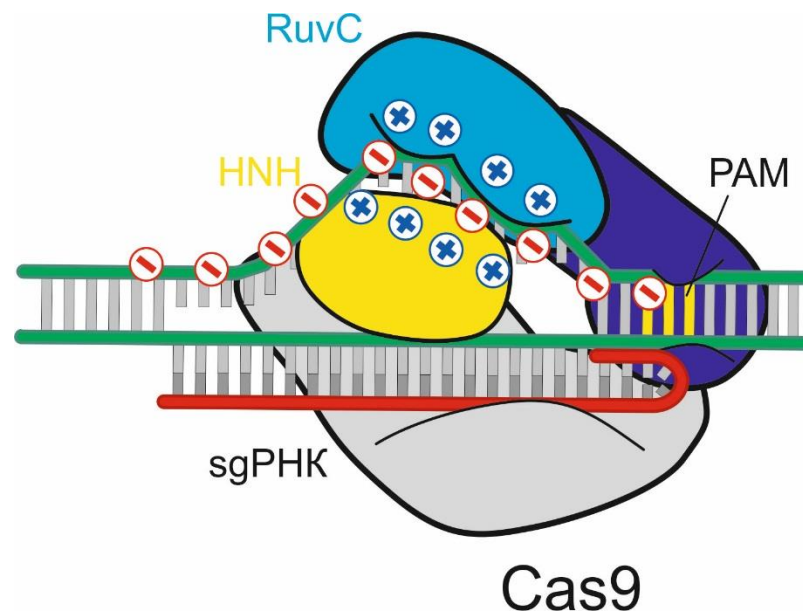


Рисунок 6. Схематичное изображение комплекса sgРНК:SpCas9:протоспейсер (Slaymaker et al., 2016).

Другим коллективом авторов был получен еще один мутантный вариант Cas9 с повышенной специфичностью действия, однако был использован несколько другой аспект механизма действия этой системы. Если в описанной выше работе авторы дестабилизировали связывание белком SpCas9 цепи ДНК, содержащей протоспейсер и PAM («нецелевой»), то в данной работе, наоборот, нарушали связывание между SpCas9 и «целевой» цепью (Kleinstiver et al., 2016). При исследованиях кристаллической структуры комплекса sgРНК:SpCas9:протоспейсер было установлено, что 4 аминокислоты в SpCas9 (N497, R661, Q695, Q926) образуют водородные связи с сахарофосфатным остовом «целевой» цепи ДНК (Anders et al., 2014, Nishimasu et al., 2014). Была выдвинута гипотеза, что замена одной или нескольких из этих аминокислот увеличит специфичность действия SpCas9. В результате из всех комбинаций вариант SpCas9-HF1, в котором все 4 аминокислоты были заменены, показал наименьшую частоту нецелевых эффектов (ниже чувствительности использованного метода), сохранив при этом целевую активность.

1.1.5.3 Снижение нецелевых эффектов путем уменьшения времени нахождения sgРНК:Cas9 в целевых клетках

Другим аспектом, который увеличивает риск неспецифичного действия CRISPR/Cas9, является избыток элементов этой системы в целевых клетках.

Показано, что при доставке CRISPR/Cas9 в виде плазмидной ДНК, кодирующей элементы системы, повышение концентрации приводит к увеличению частоты нецелевых эффектов. Показано, что при трансфекции клеток линии НЕК 293FT плазмидным вектором pX330-U6-Chimeric_BB-CBh-hSpCas9 от 18 пмоль/клетку до 710 пмоль/клетку с увеличением концентрации вектора возрастает частота нецелевых эффектов, и, соответственно, снижается специфичность действия (Hsu et al., 2013). Предполагается, что избыток комплексов sgРНК:Cas9 приводит к увеличению частоты нецелевых эффектов, также при доставке CRISPR/Cas9 в виде ДНК элементы системы экспрессируются в течение достаточно долгого периода времени, следовательно возрастает вероятность того, что присутствующие в клетке комплексы sgРНК:Cas9 внесут ДЦР в нецелевых сайтах.

Сократить время нахождения комплексов sgРНК:Cas9 в клетках возможно, исключив доставку в виде ДНК. Так как время полужизни РНК и белков в клетке относительно невелико, то доставка элементов системы CRISPR/Cas9 *in vivo* в виде комплекса sgРНК:Cas9 позволит сократить время их нахождения в клетке и, таким образом, снизить риск нецелевых эффектов. В экспериментах *in vitro* показано, что при доставке комплексов sgРНК:Cas9 редактирование генома клеток человека проходит более специфично, чем при использовании плазмидной ДНК, при этом в белок Cas9 не детектируется в клетках уже через 24 часа, а при использовании плазмидных векторов экспрессия сохраняется как минимум в течение 72 часов (Kim et al., 2014).

Если все-таки доставка в клетки осуществляется в виде ДНК, то сократить время нахождения sgРНК:Cas9 в клетках возможно, подавив экспрессию кодирующих их генов. Этого можно достичь, если вектор будет экспрессировать sgРНК, мишенью которой является сам вектор (Moore et al., 2015). В этом случае после достижения определенного уровня комплексов sgРНК:Cas9 в клетке все копии вектора будут гидролизованы, и, как следствие, экспрессия прекратится.

Риск внесения нецелевых ДЦР при редактировании геномов *in vivo* является серьезным недостатком этого подхода, препятствием на пути к клиническому применению. Однако, следует понимать, насколько эти риски существенны в различных типах исследований. При этом значимость риска нецелевых эффектов зависит напрямую от поставленных в исследовании целей. В фундаментальных

исследованиях *in vitro* по нокауту или нокину генов для минимизации рисков достаточно тщательного подбора целевого сайта, также нецелевые эффекты могут не приниматься в расчет, если выводы основаны на анализе нескольких клонов или при использовании нескольких sgРНК. Также в таких работах допустимо и использование ленти- и ретровирусных векторов. В случае же редактирования геномов *in vivo*, когда нет возможности выбрать клеточные линии для дальнейшей работы (как при редактировании геномов *ex vivo*), необходимо использовать дополнительные методы для минимизации нецелевых эффектов.

1.1.6 Современные концепции применения системы CRISPR/Cas для терапии наследственных заболеваний

В настоящее время активно разрабатываются подходы по применению системы CRISPR/Cas для редактирования геномов с целью терапии заболеваний, имеющих генетическую основу. От эффекта мутации (потеря функции или изменение функции гена) зависит, каким образом методы редактирования генома могут быть применены для коррекции генотипа (Рисунок 7):

1. Мутантные аллели, продукты которых приобрели патологическую функцию, могут быть нокаутированы. Если при этом аллель дикого типа остается интактным (при гетерозиготной мутации), то будет экспрессироваться только нормальный продукт, что может привести к коррекции мутантного фенотипа в ряде случаев.
2. В случае экспансии повторов (как, например, при болезни Гентингтона или атаксии Фридрейха) с помощью двух ДЦР возможно удалить «лишние» повторы. Также в случае мышечной дистрофии Дюшенна удаление экзонов гена *Dmd*, содержащих нонсенс мутации, позволяет восстановить рамку считывания, в результате чего при экспрессии модифицированного аллеля образуется функциональный белок (Long et al., 2016, Nelson et al., 2016, Tabebordbar et al., 2016)
3. В случае точечных мутаций при потере (например, при болезни Тея-Сакса) или изменении (например, мутация G93A в гене SOD1 приводит к развитию амиотрофического бокового склероза) функции продуктов мутантных аллелей с помощью гомологичной рекомбинации возможно

заменить последовательность ДНК в геноме на последовательность дикого типа.

4. Обнаружены мутации, при которых продукт гена теряет свою функцию, предотвращающие развитие заболеваний. Например, обнаружено, что нокаут гена CCR5 предотвращает заражение ВИЧ (Genovese et al., 2014, Liu et al., 1996). Также мутации, при которой изменяется функция продукта гена, могут предотвращать заболевания, как, например, мутация A673T в гене APP предотвращает развитие болезни Альцгеймера (Jonsson et al., 2012). В первом случае возможно с помощью методов геномной инженерии нокаутировать целевые гены, во втором – с помощью гомологичной рекомбинации внести целевые замены в геном (Cox et al., 2015)
5. Наконец, возможно направленно ввести в геном терапевтический трансген дикого типа или аллель с измененной функцией (как в примере выше).

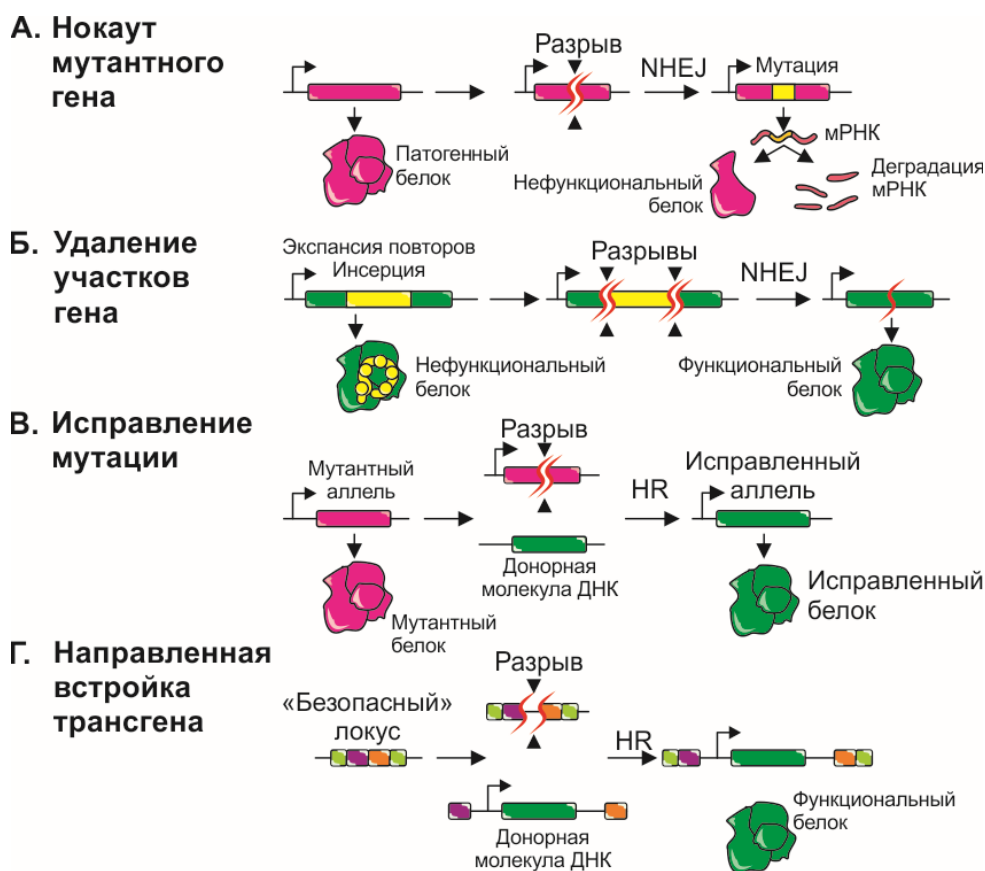


Рисунок 7. Варианты использования методов редактирования геномов при терапии наследственных заболеваний (Cox et al., 2015).

В настоящее время существуют две концепции использования методов редактирования геномов для терапии наследственных заболеваний: редактирование геномов *ex vivo* и *in vivo* (Рисунок 8). В первом случае предполагается извлекать клеточный материал из организма пациента, редактировать геном культивируемых клеток *in vitro*, а затем трансплантировать клетки с измененным геномом обратно в организм пациента. При этом клетки служат своеобразным вектором для доставки измененного генома в организм, и, соответственно, в зависимости от характера внесенных изменений могут либо корректировать мутантный фенотип, например, замещая погибающие при патологии клетки, либо предотвращать развитие заболевания. Преимущества такого подхода следующие: во-первых, методы редактирования геномов *in vitro* уже достаточно хорошо разработаны, существует множество способов доставки элементов систем редактирования геномов (в виде ДНК, мРНК, белков) в клетки, при этом возможно контролировать их дозу на клетку. Во-вторых, в результате клонирования и отбора возможно получить гомогенные линии клеток с измененным геномом и отобрать те линии, в геноме которых, например, не присутствуют нецелевые мутации.

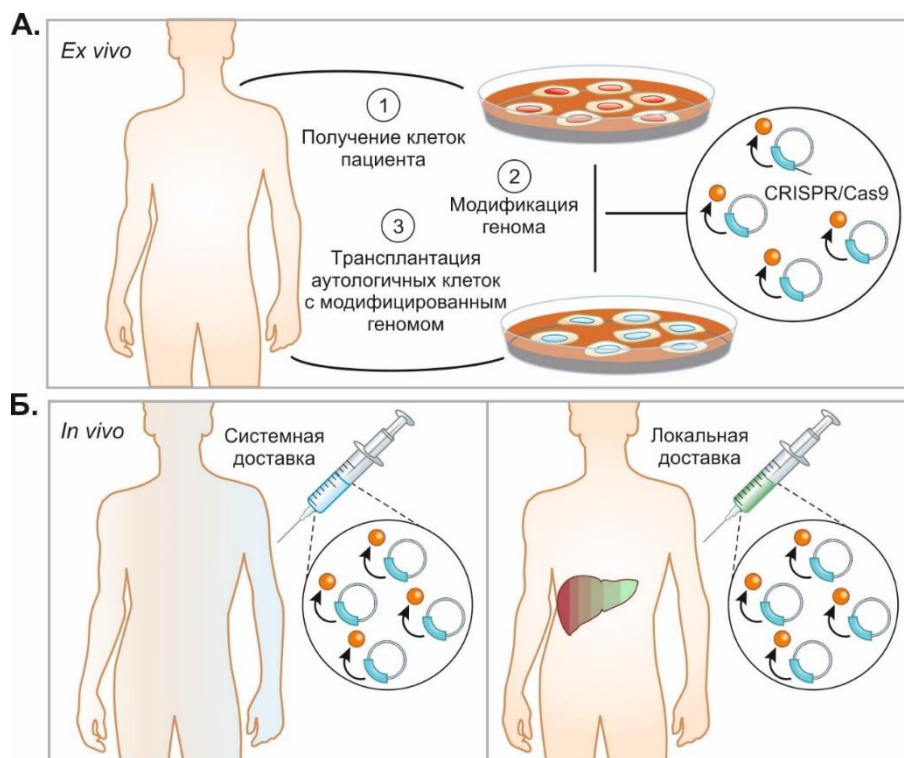


Рисунок 8. Схематическое изображение двух концепций терапии наследственных заболеваний: редактирование геномов *ex vivo* (А) и *in vivo* (Б) (Cox et al., 2015).

Однако у описанного выше подхода есть ряд проблем. Клетки, которые извлекаются из организма пациента должны быть способны «пережить» все манипуляции, проводимые при редактировании геномов, вне организма пациента. При этом, если заболевание поражает определенный тип специализированных клеток (например, моторные нейроны или нейроны головного мозга), зачастую нет возможности выделить такой тип клеток, либо они не способны к длительному культивированию *in vitro*. Второй важный вопрос, требующий тщательного изучения – способность введенных клеток приживаться в организме пациента после трансплантации и замещать поврежденные при патогенезе клетки (Cox et al., 2015).

Редактирование геномов *in vivo* в свою очередь предполагает непосредственно модификацию геномов клеток организма пациента. Эта концепция решает описанные выше проблемы подходов к терапии заболеваний, использующих редактирование геномов *ex vivo*, за счет отсутствия этапа, связанного с культивированием клеток *in vitro*. При этом редактирование генома может быть произведено в различных органах и тканях одновременно при системной доставке. Таким образом, спектр заболеваний, для которых может быть проведена терапия, при редактировании геномов *in vivo* больше. Однако, этот подход на данный момент менее развит, остается ряд вопросов, касающихся способов доставки, безопасности применения, спектра вносимых модификаций, которые ещё не решены окончательно.

1.2 Использование индуцированных плюрипотентных стволовых клеток для терапии наследственных заболеваний

Использование индуцированных плюрипотентных стволовых клеток (ИПСК), может решить проблемы, возникающие при терапии наследственных заболеваний путем редактирования геномов *ex vivo*. Этот тип плюрипотентных клеток впервые был получен в 2006 году и с тех пор активно используется в качестве перспективного объекта для биомедицинских исследований (Takahashi, Yamanaka, 2006). Этот тип плюрипотентных клеток может быть получен из практически любых клеток пациента, в том числе клеток крови (Okita et al., 2011, Okita et al., 2013), в результате репрограммирования с помощью сверхэкспрессии определенных генов (*Oct4*, *Sox2*, *Klf4* и *c-Myc* - OSKM), способен поддерживаться в культуре *in vitro* неограниченное

количество времени за счет самообновления. Также ИПСК теоретически могут быть дифференцированы в любые типы клеток. Учитывая, что ИПСК могут быть получены из собственных клеток организма пациента, то они являются аутологичными, и их использование позволит избежать иммунного ответа на вводимый при клеточной терапии материал. Таким образом, сочетание двух активно развивающихся направлений – редактирования геномов и использования ИПСК в биомедицине – является перспективным в рамках разработки подходов по клеточной терапии наследственных заболеваний. Рассмотрим далее примеры успешного исправления мутаций в геномах ИПСК человека с помощью CRISPR/Cas, а также современное состояние разработок по клеточной терапии наследственных заболеваний с помощью этого типа клеток.

Впервые возможность использования аутологичных ИПСК для клеточной терапии наследственных заболеваний была продемонстрирована в группе Рудольфа Йениша (Rudolf Jaenisch) в 2007 году на гуманизированной мышинной модели серповидноклеточной анемии (Hanna et al., 2007). В геноме этой линии мышей эндогенные ген α -глобина заменен на ген α -глобина человека, а ген β -глобина на ген β^S -глобина человека, вызывающий серповидноклеточную анемию. В данной работе из соматических клеток мышей с помощью репрограммирования были получены ИПСК, в которых с помощью гомологичной рекомбинации с донорным плазмидным вектором была исправлена мутация в гене β -глобина (без использования искусственных нуклеаз). Эффективность исправления мутации составила 1,4% – 1 клон с исправленной мутацией из 72. В результате дифференцировки из ИПСК с исправленной мутацией были получены гематopoэтические клетки, которые были трансплантированы в организм облученных мышей. Введенные клетки не только прижились в организме, но и дифференцировались в клетки эритропоэтического ряда, в результате чего произошла коррекция мутантного фенотипа заболевания.

Через год после получения ИПСК мыши в результате репрограммирования фибробластов кожи впервые были получены ИПСК человека (Takahashi et al., 2007). Однако, в отличие от клеток мыши плюрипотентные стволовые клетки человека трудно поддаются методам геномной инженерии – в том числе частота спонтанной гомологичной рекомбинации в этих клетках достаточно низкая (Zwaka, Thomson, 2003). Развитие системы CRISPR/Cas стало толчком к развитию направления по

получению ИПСК человека с модифицированным геномом, благодаря своей простоте и доступности. В первых же работах по адаптации CRISPR/Cas (см. раздел 1.1.3) были внесены мутации (в результате NHEJ) в локус *AAVS1* в геноме ИПСК человека, эффективность при этом составила 2-4% (10-25% в клетках 293T) (Mali et al., 2013). В этой же работе с помощью HR в локус *AAVS1* был встроен ген устойчивости к пурамицину и репортерный ген *GFP*.

1.2.1 Исправление мутаций в пациент-специфичных ИПСК

С тех пор, как в 2013 была показана возможность эффективного редактирования геномов ИПСК с помощью CRISPR/Cas, был опубликован ряд работ, в которых было проведено исправление мутаций, вызывающих наследственные заболевания, в ИПСК, полученных из клеток пациентов. Одной из тенденций в этом направлении является получение «безопасных» ИПСК, геном которых не содержит экзогенной ДНК. Для этого ИПСК человека получают в результате репрограммирования без интеграции факторов OSKM в геном. В коллективе под руководством Синъи Яманаки для получения ИПСК проводили сверхэкспрессию факторов репрограммирования *OCT4*, *SOX2*, *KLF-4* и *c-MYC* в соматических клетках с помощью ретровирусных векторов, которые встраиваются в геном (Takahashi et al., 2007). Опасность этого подхода заключается в риске реактивации трансгенов после трансплантации ИПСК или их производных в организм, что может привести к образованию тератом. В настоящее время при работе с клетками человека в основном используют векторы, не встраивающиеся в геном (Fusaki et al., 2009, Yu et al., 2009).

Другой аспект безопасности при получении ИПСК с исправленным генотипом – дополнительные гены (репортерные гены, гены устойчивости к антибиотикам), которые попадают в геном при гомологичной рекомбинации в целевом локусе. В настоящее время используют два типа донорных молекул для замены целевых последовательностей в геноме. Более «традиционный» вариант – использование в качестве донорной молекулы двухцепочечной плазмидной ДНК. Использование плазмидных векторов позволяет задействовать дополнительные инструменты для облегчения поиска клонов, в которых произошла HR. Одним из таких способов является позитивно-негативная селекция, в результате которой возможно получить клетки, в которых произошли целевые события. После

получения устойчивых к антибиотику или экспрессирующих репортерный ген клональных клеточных линий среди них отбирают те, в которых произошла HR в нужном количестве копий целевого гена, и те, в которых с помощью, например, Саузерн-блот гибридизации показано отсутствие случайных или нецелевых встроек (Xie et al., 2014). Такой метод и позволяет снизить трудозатраты, однако его недостатком является интеграция, хоть и направленная, дополнительных трансгенов. Эта проблема в рамках данного подхода решается в настоящее время, например, использованием Cre-рекомбиназы. Внедренные в геном трансгены фланкированы *loxP* сайтами, которые узнает Cre-рекомбиназа, и после её экспрессии в отобранных клонах удаляются из генома, в результате чего в целевом локусе остается небольшой «шрам» или «шов», представляющий собой один сайт *loxP*. Другой пример удаления трансгенов после отбора клонов – использование системы транспозона *piggyBac*. Трансгены в донорном векторе при этом фланкированы инвертированными концевыми повторами (ITR). После отбора клонов, в которых произошла гомологичная рекомбинация, в результате экспрессии транспозазы трансгены удаляются из генома. При этом в целевом локусе «шрама» не остается – последовательность трансгенов удаляется полностью вместе с ITR (Xie et al., 2014).

Альтернативой использования в качестве донорной молекулы для HR плазмидных векторов с последующим удалением трансгенов в настоящее время является использование в качестве донорных молекул одноцепочечных олигонуклеотидов (ssODN). Показано, что для эффективного прохождения гомологичной рекомбинации с одноцепочечными олигонуклеотидами оптимальный размер донорной молекулы – 90 нуклеотидов (Yang et al., 2013). При этом для увеличения эффективности гомологичной рекомбинации ДЦР, вносимые системой CRISPR/Cas, должны возникать в пределах участка гомологии между донорной молекулой и целевым участком генома. При использовании данного подхода нет возможности проводить селекцию клеток, в которых произошла гомологичная рекомбинация, однако для обогащения популяции возможно использовать репортерные гены (или гены устойчивости) в векторах, кодирующих элементы системы CRISPR/Cas.

С использованием пациент-специфичных ИПСК в настоящее время проведены работы по исправлению мутаций, вызывающих наследственные заболевания, и

изучению, соответственно, влияния на фенотип заболевания, в том числе в релевантных типах клеток после направленной дифференцировки (Таблица 2).

Таблица 2. Примеры применения CRISPR/Cas9 для редактирования геномов пациент-специфичных ИПСК (Prakash et al., 2016).

Заболевание	Ген	Тип модификации/тип донорной молекулы	Эффективность	Тип ИПСК	Ссылка
β-талассемия	<i>HBB</i>	HR/плазмидный вектор	23,5%	Без интеграции	(Xie et al., 2014)
			12,3%	Без интеграции	(Xu et al., 2015)
Х-сцепленный пигментный ретинит	<i>RPGR</i>	HR/ssODN*	13%	Ретровирусная интеграция	(Bassuk et al., 2016)
Наследственная глухота	<i>MYO15A</i>	HR/ssODN	16,7%	Ретровирусная интеграция	(Chen et al., 2016)
		HR/плазмидный вектор	10,4%		
Мышечная дистрофия Дюшенна	<i>DMD</i>	NHEJ, пропуск экзона	4,4%	Без интеграции	(Li et al., 2015)
		NHEJ, восстановление рамки считывания	4,4%		
		HR/плазмидный вектор	33,3%		
Муковисцидоз	<i>CFTR</i>	HR/плазмидный вектор	16,7%	Ретровирусная интеграция/без интеграции	(Firth et al., 2015)
Гемофилия А	<i>F8</i>	Два разрыва, инверсии 140 и 600 т.п.н.	6,7%	Без интеграции	(Park et al., 2015)
Тяжелый комбинированный иммунодефицит	<i>JAK3</i>	HR/плазмидный вектор	2,56%**	Лентивирусная интеграция	(Chang et al., 2015)
Хроническая гранулематозная болезнь	<i>CYBB</i>	HR/плазмидный вектор	17%	Лентивирусная интеграция	(Flynn et al., 2015)
*ssODN - одноцепочечные олигонуклеотиды					
** - исправление мутации в двух копиях гена сразу					

Например, в 2016 году была исследована возможность получения пациент-специфичных ИПСК, в геноме которых исправлена мутация, вызывающая глухоту. В данной работе были получены ИПСК из клеток пациента, носителя мутаций 4642G>A и 8374G>A в разных копиях гена *MYO15A*, который кодирует структурный белок волосковых клеток внутреннего уха. В гетерозиготном состоянии эти мутации не проявляются фенотипически, однако в совокупности приводят к глухоте. С помощью системы CRISPR/Cas было проведено исправление мутации с использованием в качестве донорной молекулы как одноцепочечных олигонуклеотидов, так и донорных плазмидных векторов. В данной работе авторы использовали вектор pX330, разработанный коллективом Фенга Чжана (см. раздел 1.1.3), в который был клонирован репортерный ген *GFP*. Эффективность трансфекции пациент-специфичных ИПСК была низкая (5-7%), поэтому для обогащения популяции был проведен сортиг GFP+ клеток. В результате были

получены линии, в которых была исправлена только мутация 4642G>A – гетерозиготные по мутации в гене *MYO15A*. Из полученных ИПСК в результате направленной дифференцировки были получены клетки, по своим свойствам соответствующие волосковым клеткам внутреннего уха. Показано, что полученные клетки с одним исправленным аллелем по своим свойствам соответствуют клеткам, полученным из ИПСК гетерозиготным по мутации и гомозиготным по аллелю дикого типа, которые были получены от родственников пациента, не страдающих глухотой (Chen et al., 2016). Авторами этой статьи обсуждается факт, что трансплантация стволовых клеток во внутреннее ухо возможна в рамках терапии наследственной глухоты, и, более того, введенные клетки приживаются в опытах на морских свинках (Watada et al., 2015). Соответственно, аутологичные пациент-специфичные ИПСК с исправленной мутацией (или их дифференцированные производные) могут использоваться для клеточной терапии наследственной глухоты.

Одним из наиболее изученных типов заболеваний в рассматриваемой нами области являются наследственные заболевания системы крови. Возможно, это связано с тем, что клеточная терапия для заболеваний данной системы достаточно давно применяется и хорошо изучена – трансплантация костного мозга и переливание крови являются неотъемлемыми процедурами современной медицины. Однако не всегда возможно найти донора с подходящим HLA гаплотипом, поэтому получение аутологичных ИПСК, которые могут быть дифференцированы в клетки гематопоэтического ряда, с исправленной мутацией является одним из возможных решений этой проблемы при лечении наследственных заболеваний крови. Одним из заболеваний, возможность терапии которого аутологичными ИПСК активно изучается, является β -талассемия. Причиной этого заболевания являются мутации в гене *HBB*, кодирующем субъединицу гемоглобина. Причем, в настоящее время известны более 200 мутаций в локусе *HBB*, приводящих к развитию β -талассемии. В работах двух различных коллективов было проведено исправление мутаций в гене *HBB* (Xie et al., 2014, Xu et al., 2015). В работе, опубликованной в 2014 году, были получены ИПСК из фибробластов пациента, который является двойной гетерозиготой по мутациям в гене *HBB*: замене 28A>G в промоторной области и делеции четырех нуклеотидов, приводящей к сдвигу рамке считывания, во втором

экзоне. Для исправления мутации с помощью HR был сконструирован донорный плазмидный вектор, который содержал плечи гомологии с локусом *HBB*, не содержащие мутаций. При этом «левое» (5') плечо гомологии содержало последовательность промоторной области гена и первый экзон, а «правое» (3') – второй экзон гена. Данные плечи фланкировали трансгены, необходимые для позитивно-негативной селекции, которые в результате HR встраивались в интрон. Таким образом, данный донорный вектор использовался для таргетинга обеих мутаций в гене *HBB*. Для последующего удаления трансгенов была использована система транспозона *piggyBac*. В результате анализа с помощью ПЦР и Саузерн-блот гибридизации полученных после селекции клонов были выбраны клоны без случайных встроек, в которых произошла HR в целевом локусе. Из 12 клонов, в которых произошла гомологичная рекомбинация, в 4 клонках была исправлена мутация 28A>G и в 5 клонках делеция четырех нуклеотидов (Xie et al., 2014). После отбора клонов трансгены были удалены с помощью транспозазы. Чтобы избежать встройки транспозона в другом сайте генома использовали негативную селекцию. Чтобы проверить будет ли проходить экспрессия исправленной копии *HBB*, ИПСК дифференцировали в эритробласты. С помощью ПЦР в реальном времени с использованием в качестве матрицы кДНК было показано, что в полученных эритроблках произошло 16-кратное увеличение уровня мРНК гена *HBB* (Xie et al., 2014).

Другим коллективом авторов были получены пациент-специфичные ИПСК, несущие гомозиготную мутацию IVS2-654(C>T), вызывающую возникновение аберрантного сайта сплайсинга во втором интроне пре-мРНК гена *HBB*. Для исправления мутации в пациент-специфичных ИПСК был использован донорный плазмидный вектор и система CRISPR/Cas. Сайт ДЦР располагался в пределах «левого» плеча гомологии вблизи его 3'-конца. Для последующего удаления трансгенов так же, как и в описанной выше работе, использовали системы транспозона *piggyBac*. После селекции клеток, устойчивых к антибиотику, были выбраны гомозиготные клоны, в которых произошло исправление мутации. В данной работе показано, что отобранные клоны, в которых было проведено исправление мутации, сохранили все свойства плюрипотентных стволовых клеток и

при дифференцировке показывают восстановление экспрессии гена *HBV* (ПЦР в реальном времени с обратной транскрипцией) (Xu et al., 2015).

Выше мы рассмотрели общие принципы редактирования геномов, и как они могут быть применены для клеточной терапии наследственных заболеваний. Каждое заболевание уникально и при его изучении необходимо ответить на ряд вопросов: как именно может быть устранено генетическое нарушение, вызывающее болезнь? Будут ли сохранять ИПСК свои свойства после всех требуемых манипуляций? Будет ли происходить коррекция мутантного фенотипа в релевантных типах клеток с модифицированным геномом? Будут ли приживаться клетки в организме после трансплантации? Какова должна быть доля клеток с модифицированным геномом в ткани/органе/организме, чтобы, как минимум смягчить, заболевание? Ответы на первые три вопроса для некоторых заболеваний, перечисленных в таблице 2, уже получены при проведении исследований *in vitro*. Для ряда заболеваний (Таблица 2) показана возможность исправления мутаций, показано, что такие клетки сохраняют плюрипотентное состояние, и при дифференцировке в релевантные типы клеток наблюдается коррекция мутантного фенотипа. Однако, для ответа на остальные вопросы необходимы эксперименты по трансплантации клеток *in vivo*. В настоящее время уже проводятся клинические испытания с использованием ЭСК человека для клеточной терапии макулодистрофии сетчатки (Schwartz et al., 2015), однако клинические испытания с использованием ИПСК, которые были запущены в сентябре 2014 года научном институте RIKEN в Японии, научным сообществом осуждаются и пока считаются преждевременными (Garber, 2015). Таким образом, актуальной задачей на настоящий момент является проведение доклинических исследований клеточной терапии с использованием ИПСК на модельных животных. Рассмотрим примеры модельных систем, которые в настоящее время используются, и какие данные были получены при их использовании.

1.2.2 Перспективы использования ИПСК для терапии наследственных заболеваний

Система CRISPR/Cas появилась относительно недавно, поэтому примеров её применения для изучения возможности терапии наследственных заболеваний с помощью редактирования геномов *ex vivo* не так много (по сравнению с TALEN и ZFN), однако в нескольких работах в виде функционального теста ИПСК с

исправленной мутацией была проведена трансплантация в организм мышей. Например, были получены 2 линии пациент-специфичных ИПСК, содержащих разные мутации в гене *F8*, которые приводят к развитию гемофилии А. Эти мутация – инверсии различных участков гена (140 т.п.н. и 600 т.п.н.) – были исправлены с помощью внесения двух ДЦР по краям инвертированных участков. Таким образом, была произведена «реверсия». После оценки коррекции фенотипа *in vitro*, ИПСК полученных линий были трансплантированы в организм модельной линии мышей (модель гемофилии А, нокаут *F8*). В тесте *in vivo* показано, что полученные клетки способны приводить к коррекции заболевания (Park et al., 2015).

Если рассматривать в целом использование ИПСК в регенеративной медицине, то за 10 лет, прошедших с момента их получения впервые, был достигнут серьезный прогресс. Наиболее впечатляющим примером является ряд работ, в которых были получены многообещающие результаты по использованию клеток пигментного эпителия сетчатки (RPE), полученных из аутологичных ИПСК, для терапии возрастной макулодистрофии (ВМД). При аллогенной трансплантации RPE, полученных из ИПСК человека, в субретинальное пространство глаза у крыс наблюдается позитивный терапевтический эффект на модели дегенерации сетчатки, однако несмотря на иммуносупрессию клетки трансплантата элиминируется из организма уже через 13 недель (Carr et al., 2009). С другой стороны, в 2014 году показано, что через год после трансплантации обезьянам RPE, полученных из аутологичных ИПСК, не наблюдается ни признаков отторжения аутографта, ни формирования тератом (Kamao et al., 2014). В этой работе был согласно стандартам GMP разработан протокол формирования *in vitro* из RPE, полученных из «безопасных» ИПСК человека, своеобразных клеточных пластин или слоев, представляющих собой монослой клеток на базальной мембране, которая по своим свойствам соответствует мембране Бруха сосудистой оболочки глаза. Показано, что такие трансплантаты лучше приживаются по сравнению с трансплантациями клеток в виде суспензии в субретинальное пространство. Успехи, достигнутые в этой работе, позволили запустить первые клинические испытания по клеточной терапии ВМД с использованием аутологичных ИПСК в институте RIKEN, которые были приостановлены в связи изменениями в законодательстве Японии, касающихся регенеративной медицины. Из 6 предполагаемых пациентов была проведена

трансплантация ИПСК только одной пациентке с диагнозом ВМД. Спустя год после трансплантации осенью 2015 года, согласно пресс-релизам RIKEN (<http://www.riken-ibri.jp/AMD/>), не было обнаружено признаков иммунной реакции на трансплантат, а также признаков туморогенеза. При этом падение остроты зрения, наблюдавшееся у пациентки, стабилизировалось.

1.3 Перспективы использования крыс линии Brattleboro для разработки клеточной терапии наследственных заболеваний

Для разработки подходов терапии наследственных заболеваний необходимы модельные объекты – инбредные линии животных, в геноме которых присутствуют мутации, имеющие выраженный фенотипический эффект. В качестве такого модельного объекта могут быть использованы крысы линии Brattleboro, являющиеся гомозиготными носителями мутантного аллеля *di* (diabetes insipidus) гена *Avp*, который кодирует гормон аргинин-вазопрессин (Schmale, Richter, 1984, Valtin et al., 1962). Данная мутация вызывает аутосомно-рецессивное наследственное заболевание – нейрогипофизарный несахарный диабет, для которого у данных животных характерно постоянное мочеизнурение и высокий уровень потребления воды (И.И., 2003).

В следующих разделах рассмотрим структурную организацию локуса гена *Avp*, механизм синтеза и секреции аргинин-вазопрессина в норме, а затем природу мутации в аллеле *di* и её эффект.

1.3.1 Структура локусов *Avp* и *Oxt* у крыс

Гены нейрогипофизарных пептидных гормонов окситоцина и вазопрессина образовались в результате дупликации предкового гена у челюстных животных (у бесчелюстных одна копия гена) (Gwee et al., 2008). У млекопитающих существует два варианта вазопрессина, отличающихся одной аминокислотой в 8 положении нонапептида — аргинин-вазопрессин (человек, грызуны) или лизин-вазопрессин (свиньи). У *Rattus Norvegicus* гены *Avp* и *Oxt* располагаются на третьей хромосоме (3q41) в обратной ориентации на расстоянии 11 т.п.н. друг от друга и отделены блоком повторов LINE длиной 6335 п.н. (Mohr et al., 1988, Schmitz et al., 1991) (Рисунок 9А)

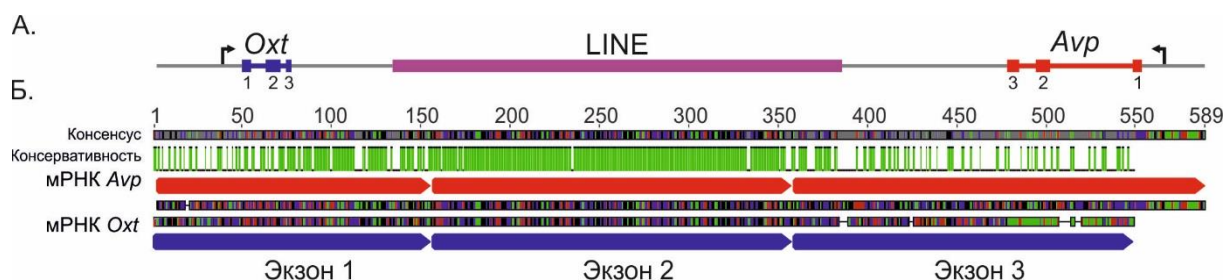


Рисунок 9. А. Структура локусов *Avp* и *Oxt*. Цифрами обозначены экзоны. Б. Выравнивание нуклеотидных последовательностей мРНК генов *Avp* и *Oxt* с помощью алгоритма ClustalW в программном обеспечении Geneious.

Гены *Oxt* и *Avp* содержат три экзона, кодирующие соответствующие препрогормоны, которые после трансляции подвергаются последующему процессингу с образованием соответствующих пептидных гормонов. Интересным является тот факт, что последовательности вторых экзонов генов вазопрессина и окситоцина практически идентичны (95% у крыс), в то время как последовательности первого и третьего экзонов различаются сильнее – у крыс нуклеотидные последовательности совпадают на ~60% и ~36%, соответственно (Рисунок 9Б). Этот факт может быть объяснен либо генной конверсией, либо действием очищающего отбора на нуклеотидную последовательность второго экзона (Gwee et al., 2008).

Первый экзон гена *Avp* кодирует сигнальный пептид, последовательность гормона аргинин-вазопрессина и N-концевую часть белка-переносчика нейрофизина II. Второй экзон кодирует центральную часть нейрофизина II. Последовательность, кодирующая С-конец белка-переносчика и гликопротеин копептин, находится в третьем экзоне гена *Avp* (Рисунок 10).

1.3.2 Синтез аргинин-вазопрессина

Синтез гормона аргинин-вазопрессина преимущественно происходит в магно- и парвоцеллюлярных нейронах супраоптического и паравентрикулярного ядер гипоталамуса. В результате трансляции мРНК гена *Avp* образуется молекула препрогормона, которая благодаря гидрофобному N-концу легко проникает в эндоплазматический ретикулум (ЭПР). В ЭПР с N-конца препрогормона отщепляется сигнальный пептид, образуется гидрофильный прогормон. После упаковки в секреторные гранулы в результате процессинга молекулы прогормона образуется гормон вазопрессин, белок-переносчик нейрофизин II и гликопротеин

копептин (Christensen, Rittig, 2006). В секреторных гранулах нейрофизин II и аргинин-вазопрессин по аксонам вазопрессинэргических нейронов гипоталамуса транспортируются в нейрогипофиз, где происходит стимул-зависимая секреция – активация секреции происходит при водной депривации.

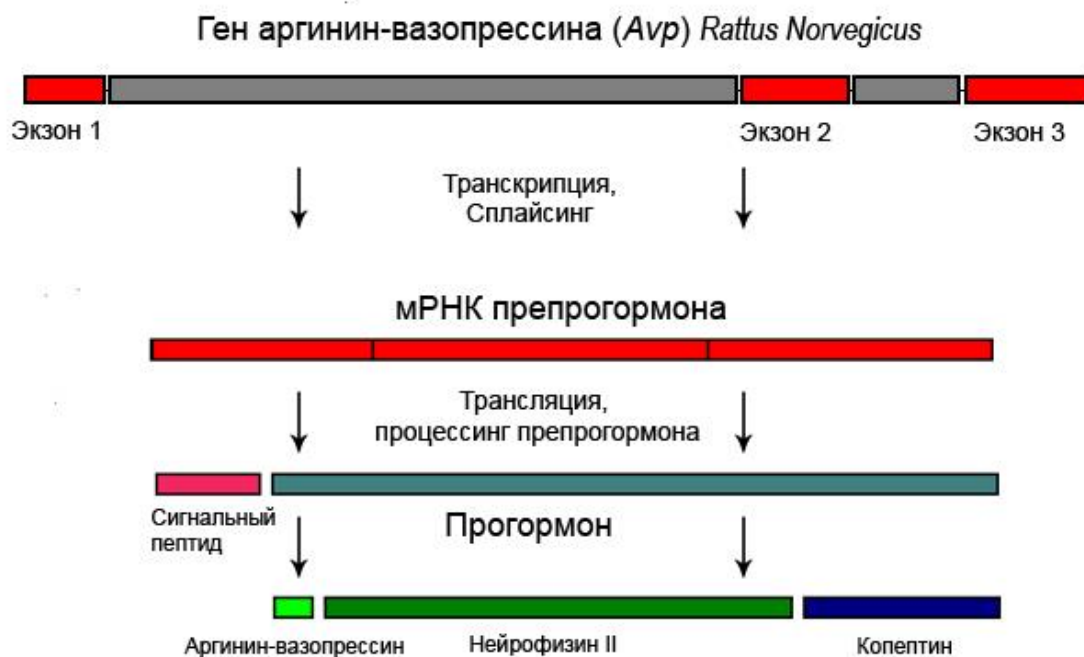


Рисунок 10. Схематическое изображение синтеза гормона аргинин-вазопрессина.

1.3.3 Крысы Brattleboro – модель наследственного несахарного гипоталамического диабета

У крыс линии Brattleboro обнаружена делеция одного азотистого основания – гуанина – во втором экзоне гена *Avp*, кодирующем аргинин-вазопрессин, которая возникла в результате случайной мутации. Было показано, что данная мутация не затрагивает последовательность, кодирующую сам нонапептид аргинин-вазопрессин (расположенную в первом экзоне), однако происходит сдвиг рамки считывания, что приводит к замене более 30 аминокислотных остатков в препрогормоне, большая часть которых относится к консервативному району белка-переносчика нейрофизина II. В результате сдвига рамки считывания пропадает стоп-кодон, что приводит к трансляции полиаденилированного 3'-конца с образованием

полилизинового С-конца. Транслированный полилизинный участок на С-конце несёт положительный электрический заряд, что приводит к изменению конформации мутантного нейрофизина II (Gabreels et al., 1997, Schmale et al., 1987). Таким образом, несмотря на то, что последовательность самого пептидного гормона не изменяется в результате мутации, изменение конформации белка-переносчика приводит к нарушению его функции и как следствие к нарушению секреции аргинин-вазопрессина. Клетки, ответственные за синтез гормона — нейроны супраоптического и паравентрикулярного ядер, гипертрофированы. При этом обнаружены плотные нейросекреторные гранулы, содержащие продукт мутантного гена.

На организменном уровне мутация проявляется у гомозиготных мутантов в форме постоянного мочеизнурения и высокого уровня потребления воды, варьирующего от 20% до 125% в пересчёте на вес тела/сутки (Хегай, 2003). Высокий уровень потребления воды является наиболее ярким, но не единственным симптомом у особей *di/di*. Полное отсутствие у гомозиготных мутантов вазопрессина изменяет показатели фертильности, динамику глюкокортикоидов в крови и ферментов катаболизма нейромедиаторов в мозге (Хегай, 2002).

Также дефицит вазопрессина может вызывать нарушение иммунной функции. У крыс линии Brattleboro наблюдается снижение количества периферических лимфоцитов крови, увеличение числа нейтрофилов, нарушение фагоцитарной активности макрофагов. Также усиливается инволюция тимуса и селезенки. У крыс данной линии усиливаются изменения иммунореактивности организма, связанные со старением. Таким образом, происходит снижение общей и специфической иммунной защиты организма (Khegai et al., 2003).

Еще одним эффектом, возникающем при дефиците вазопрессина, является необычное развитие опухолей у крыс линий Brattleboro, образующихся при введении клеток саркомы линии Walker256. У контрольных животных наблюдается быстрое увеличение размера привитой опухоли. Животные при этом гибнут через 28 дней после введения опухолевых клеток. Однако при введении клеток Walker256 крысам линии Brattleboro наблюдается медленный рост опухоли в течение десяти дней, а затем её полная резорбция. При этом гибридные линии, полученные от скрещивания крыс Brattleboro с крысами, несущими аллель *Avp* дикого типа (лабораторная линия

WAG), по динамике роста опухоли не отличаются от крыс контрольной линии. Введение же экзогенного вазопрессина крысам Brattleboro усиливает рост опухоли, однако не предотвращает её регрессии (Khegay et al., 2010, Zakharova et al., 2011).

Лабораторная линия крыс Brattleboro обладает рядом преимуществ, которые делают этих животных уникальным объектом для разработки подходов по терапии наследственных заболеваний с использованием технологий редактирования геномов. Во-первых, наследственное заболевание, которое проявляется у крыс данной линии, имеет моногенную природу. Мутация, вызывающая у крыс Brattleboro несхарный диабет локализована и детально изучена, что упрощает разработку подходов по исправлению генетических мутаций и коррекции их проявлений. Во-вторых, данная мутация имеет однозначные фенотипические проявления (полиурия и полидипсия), которые позволят судить об эффективности коррекции заболевания при терапии. Показатели суточного потребления воды и объем выделенной за день мочи позволяют отслеживать динамику терапии без вмешательств в организм животного.

1.4 Заключение

Использование современных методов редактирования геномов является перспективным подходом для терапии наследственных заболеваний, которые до недавнего времени являлись неизлечимыми, поскольку не было способа подействовать на саму причину заболевания. С появлением в 2013 году в открытом доступе системы CRISPR/Cas, был совершен существенный скачок в этом направлении. Благодаря исправлению мутации могут быть получены клетки аутологичные по отношению к организму пациента, страдающего наследственным заболеванием, что существенно снижает, если не ликвидирует, риск иммунной реакции на трансплантат. Результаты научных исследований, проведенных за последние 10 лет с момента появления ИПСК, показали, что этот новый тип плюрипотентных стволовых клеток может быть использован в регенеративной медицине. Появление этого типа клеток позволило снять этические ограничения, связанные с использованием эмбриональных стволовых клеток. Использование этого типа клеток позволяет решить проблему отсутствия подходящего донора, например, при пересадках костного мозга – разработаны протоколы эффективной дифференцировки ИПСК в клетки гематопоетического ряда, которые могут быть трансплантированы в организм пациента.

Главный вопрос – какое будущее ожидает это направление? Альтернативой редактированию геномов *ex vivo* является активно развивающаяся область по редактированию геномов *in vivo* (Cox et al., 2015). Не ясно, какой из подходов является более перспективным. Не исключено, что для каждого из заболеваний будет разработан специфический подход, а не универсальный. Поэтому исследования по разработке клеточной терапии с использованием аутологичных клеток с исправленными мутациями на модельных животных являются крайне актуальной задачей современной биомедицины, поскольку результаты этих исследований, разработанные в них концепции применения методов геномной инженерии, клеточной биологии и трансплантологии в дальнейшем могут быть использованы для терапии заболеваний человека.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1 Конструирование векторов для экспрессии системы CRISPR/Cas9

Плазмидные векторы pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene plasmid #42230) и pX335-U6-Chimeric_BB-CBh-hSpCas9n(D10A) (Addgene Plasmid #42335) были приобретены в депозитарии Addgene. Последовательности спейсеров были клонированы в соответствующие векторы как описано ранее (Cong et al., 2013).

CR7_F 5'-CACCGCAGAAGCCTTGCGGAAGCG-3'

CR7_R 5'-AAACCGCTTCCGCAAGGCTTCTGC-3'

CR8_F 5'-CACCGCAGCGGCCTCGCTTCCGCA-3'

CR8_R 5'-AAACTGCGGAAGCGAGGCCGCTGC-3'

CRN_F 5'-CACCGCGATGGTGCGCACAAAGCC

CRN_R 5'-AAACGGCTTTGTGCGCACCATCGC-3'

Пары олигонуклеотидов фосфорилировали с использованием полинуклеотидкиназы фага T4 (New England Biolabs, M0201S). Затем проводили гибридизацию пары олигонуклеотидов при понижении температуры с 95 °C до 25 °C со скоростью 0,3 °C/с, после чего двухцепочечные олигонуклеотиды клонировали в соответствующий плазмидный вектор, гидролизированный с помощью эндонуклеазы рестрикции *BbsI* (New England Biolabs, R0539S). В результате было получено пять плазмидных векторов: pX330_CR7, pX330_CR8, pX330_CRN, pX335_CR7 и pX335_CR8.

2.2 Методы работы с клеточными культурами

2.2.1 Получение культуры эмбриональных фибробластов крыс

Эмбриональные фибробласты получали из кожи 19-тидневных эмбрионов крыс. Кожу механически измельчали и помещали в 0,25% трипсин (Thermo Fisher Scientific, США) на 10 мин при 37 °C. Клетки осаждали при 300g, супернатант убирали. После промывки в фосфатном буфере (PBS) опять осаждали, затем полученный осадок ресуспендировали в культуральной среде (см. ниже) и помещали в культуральную посуду.

2.2.2 Культивирование эмбриональных фибробластов крыс

Клетки культивировали при 37 °C, 5% CO₂ в среде DMEM/F12 с добавлением 10% FBS, 1 mM L-глутамина, 50 ед./мл пенициллина, 50 ед./мл стрептомицина (Life Technologies, США). Для открепления клеток от поверхности культуральной посуды использовали аналог трипсина TrypLE Express (Thermo Fisher Scientific, США).

2.2.3 Позитивная селекция

Для позитивной селекции клеток, содержащих трансген *NeoR*, в культуральную среду добавляли антибиотик G418 (Santa Cruz Biotechnology, sc-29065) в финальной концентрации 250 мкг/мл в течение 7 суток.

Для позитивной селекции клеток, экспрессирующих SpCas9, после трансфекции вектором pX330_CR7_Puro в культуральную среду добавляли антибиотик пурамицин (Santa Cruz Biotechnology, sc-108071) в концентрации 0,9 мкг/мл в течение 2 суток.

2.2.4 Негативная селекция

Для негативной селекции клеток со случайной встройкой в геном донорного вектора pAvp-target8 в культуральную среду добавляли ганцикловир (InvivoGen, sud-gcv) в концентрации 50 мкг/мл в течение 7 суток.

После селекции эмбриональные фибробласты наращивали в культуральной среде с добавлением основного фактора роста фибробластов (FGF-basic, BioLegend, 710304) в финальной концентрации 5 нг/мл.

2.3 Методы доставки ДНК *in vitro*

2.3.1 Электропорация

Электропорацию эмбриональных фибробластов крысы (10⁶ клеток на образец) проводили с помощью набора Mouse/Rat Hepatocyte Nucleofector Kit (Lonza, VPL-1004) по протоколу производителя, используя программу Q-025.

Для качественного анализа активности CRISPR/Cas9 использовали следующие комбинации плазмидных векторов: 1) 6 мкг pX330_CR7, 2) 6 мкг pX330_CR8, 3) 6 мкг pX330_CRN, 4) 3 мкг pX335_CR7 + 3 мкг pX335_CR8. Для количественного анализа активности системы CRISPR/Cas9 использовали 6 мкг

плазмидного вектора pX330_CR7_Puro. В экспериментах по исправлению мутации с помощью гомологичной рекомбинации использовали 3 мкг pX330_CR7_Puro + 3 мкг pAvr-target8. Для определения эффективности электропорации использовали 2 мкг плазмидного вектора pmaxGFP (Lonza).

После электропорации каждый образец высевали отдельно в одну лунку 6-луночного планшета (Thermo Fisher Scientific, США) в культуральную среду. Анализировали клетки через 2 суток.

2.3.3 Определение эффективности трансфекции

Эффективность трансфекции определяли через 2 суток в контрольных образцах с помощью проточной цитофлуорометрии, используя BD FACS Canto II (BD Biosciences, США).

2.4 Молекулярно-генетические методы анализа

2.4.1 Выделение геномной ДНК

Геномную ДНК выделяли с помощью набора Quick-gDNA MiniPrep (Zymo Research, D3025) согласно протоколу производителя.

2.4.2 Полимеразная цепная реакция

1. Taq-полимераза

Состав реакции:

Матрица – 100 нг

Праймеры (10 мМ) – по 0,5 мкл

10хПЦР буфер (100 мМ KCl, 200 мМ Трис-HCl, pH 8,8, 100 мМ (NH₄)₂SO₄; 0,1% Triton X-100) – 2 мкл

Смесь dNTP (10 мМ каждого) – 0,5 мкл

Taq-полимераза (5 ед./мкл) – 0,5 мкл

MilliQ H₂O – до 20 мкл

Амплификацию проводили на термоциклере S1000 (Bio-Rad) с использованием следующей программы – предварительная денатурация 95°C, 5 мин. Далее 35 циклов: 1. Денатурация 95°C, 30 с. 2. Отжиг праймеров, 30 с. 3. Элонгация 72°C, 60 с/1 т.п.н. Элонгация 72°C, 5 мин.

2. Химерная полимеразы Herculanase II

Амплификацию проводили согласно рекомендациям производителя (Agilent Genomics) на термоциклере S1000 (Bio-Rad) с использованием следующей программы – предварительная денатурация 95°C, 5 мин Далее 35 циклов: 1. Денатурация 95°C, 20 с, 2. Отжиг праймеров, 20 с, 3. Элонгация 72°C, 30 с/1 т.п.н. Элонгация 72°C, 5 мин.

3. Выявление рекомбинантных клонов с помощью ПЦР

Для выявления рекомбинантных клонов выделяли геномную ДНК и проводили амплификацию с помощью набора для ПЦР длинных фрагментов БиоМастер LR HS-ПЦР-Color (2x) (MHC040-400, Биоллабмикс) по протоколу производителя. Для амплификации использовали олигонуклеотидные праймеры, подобранные на различные участки локуса *Avr* (Таблица 3).

Таблица 3. Олигонуклеотидные праймеры для выявления рекомбинантных клонов

Название	Структура	Длина продукта ПЦР, п.н.
AVP_HR_L_1F	AGTGGATGTGATTGCCCCAG	3559
AVP_HR_L_1R	TAAAGCGCATGCTCCAGACT	
AVP_HR_L_2F	TGTAGGGTCTCTCTGGGAGC	3689
AVP_HR_L_2R	ACCCGGTAGAATTGACCTGC	
AVP_HR_L_3F	GCTTGGCAAAAACCTCCCAGG	4862
AVP_HR_L_3R	CTCCAGACTGCCTTGGGAAA	
AVP_HR_L_4F	CGGGTAGGGGAGGCG	4228
AVP_HR_L_4R	CTGTGTGCTTGAGGGCTGT	
AVP_HR_R_1F	TGGCACTTGCGCTACACAA	4215
AVP_HR_R_1R	TCCCCAGCCTGAAAATGGTCTA	
AVP_HR_R_2F	GCTGGCACTTGCGCT	4217
AVP_HR_R_2R	TCCCCAGCCTGAAAATGGTCTAA	
AVP_HR_R_3F	TCCTCCCTCTCTCCCTCTCT	4878
AVP_HR_R_3R	CGCCTCCCTACCCG	
AVP_HR_LR_1F	GTAGTGAGTTTGAGTTTAGGGGC	8024
AVP_HR_LR_1R	TGTGCTTGAGGGCTGTGC	
AVP_HR_LR_2F	GAACCTCAGGCTACTGGGGG	8922
AVP_HR_LR_2R	TTCACATTGCTTCGCTGTGTG	

2.4.3 Качественный анализ активности системы CRISPR/Cas9 в целевых сайтах

Анализ с помощью эндонуклеазы I фага T7 проводили, как описано ранее (Reyon et al., 2012). Содержащий протоспейсер участок геномной ДНК, выделенной из образцов после электропорации, амплифицировали с помощью ПЦР с использованием Taq-полимеразы и олигонуклеотидных праймеров:

Avr6f GACCCCTTCATACTACTCACCT

Avp6r GTTAGATTTCCACTCTCGCCCTT

Oxt4f CATCCCCTCACACTTGCC

Oxt4r GCCCCCTCCTGCTCAC

Продукт ПЦР денатурировали при 95 °C, затем снижали температуру до 25 °C со скоростью 0,3 °C/с. Образовавшиеся гетеродуплексы подвергали гидролизу с помощью эндонуклеазы I фага T7 (T7 Endonuclease I, New England Biolabs, M0302S). Результаты анализа визуализировали с помощью электрофореза в 3% агарозном геле.

2.4.4 Рестрикционный анализ

Целевой участок геномной ДНК, выделенной из образцов после электропорации конструкций CRISPR/Cas9, амплифицировали с помощью ПЦР с использованием олигонуклеотидных праймеров (Avp6f и Avp6r). Продукт ПЦР подвергали гидролизу с помощью эндонуклеазы рестрикции *BcgI* (New England Biolabs, R0545S). Результаты анализа визуализировали с помощью электрофореза в 3% агарозном геле.

2.4.5 Количественный анализ активности системы CRISPR/Cas9 в целевых сайтах

После электропорации вектором pX330_CR7_Puro проводили селекцию, добавляя в культуральную среду пурамицин концентрации 0,9 мкг/мл. После селекции клетки наращивали до плотности ~90%, после чего выделяли геномную ДНК. Выделенную геномную ДНК использовали в качестве матрицы для ПЦР с использованием фермента Herculase II с олигонуклеотидными праймерами Avp6f и Avp6r (раздел 2.4.3).

Для биоинформатического анализа эффективности системы CRISPR/Cas9 использовали программное обеспечение TIDE (<https://tide-calculator.nki.nl/>, Brinkman et al., 2014). Для анализа проводили секвенирование продукта ПЦР, полученного с использованием в качестве матрицы геномной ДНК опытных и контрольного образцов.

Для анализа спектра модификаций полученные продукты ПЦР клонировали в вектор pGEM-T Easy (Promega, A1360) по протоколу производителя. Для определения клонов со встройкой продукта ПЦР использовали сине-белую

селекцию. После трансформации *E. coli* (штамм XL-10 Gold) случайным образом выбирали 20 белых колоний, из которых для дальнейшего анализа выделяли плазмидную ДНК. Колонии *E. coli* инокулировали в 3 мл жидкой среды LB и растили при 37 °C 16 ч. Плазмидную ДНК выделяли с помощью щелочного лизиса (Маниатис, 1984). Полученные клоны плазмидных векторов секвенировали с использованием олигонуклеотидного праймера:

M13R AACAGCTATGACCATG

M13F TGTAACACGACGGCCAGT

2.4.6 Гидролиз ДНК комплексами sgРНК:Cas9 *in vitro*

Для реакции использовали 100 нг линейаризованного вектора, очищенного с помощью выделения из агарозного геля. В качестве отрицательного контроля использовали вектор pX552_Avp_Puro, в котором протоспейсер CR-7 содержал 4 однонуклеотидные замены с 3'-конца. Для положительного контроля использовали вектор pGEM-T Easy, в который были клонированы продукты ПЦР с использованием праймеров: 1) Avp6f и Avp6r, 2) Oxt4f и Oxt4r. Векторы линейаризовали следующими эндонуклеазами рестрикции:

pAvp-target8 – *NruI* (New England Biolabs, R0192S)

pX552_Avp_Puro – *BamHI* (New England Biolabs, R0136S)

pGEM-T_Avp6f6r и pGEM-T_Avp4f4r – *BsaI* (New England Biolabs, R0535S)

Проводили амплификацию участка плазмидного вектора pX330_CR7_Puro, содержащего ген sgРНК, с использованием олигонуклеотидных праймеров:

T7_CR7_F AAGCTAATACGACTCACTATAGGCAGAAGCCTTGC

sgRNA_R AAAAGCACCGACTCGGTGCC

Праймер sgRNA_R является универсальным, поскольку комплементарен 3'-концу гена sgРНК. В праймере T7_CR7_F содержится с 5'-конца последовательность промотора РНК-полимеразы бактериофага Т7, а с 3'-конца последовательность комплементарная спейсеру CR-7. sgРНК_CR7 была получена с помощью *in vitro* транскрипции сотрудником лаборатории биотехнологии ИХБФМ СО РАН к.б.н. Степановым Г. А. Использовали белок SpCas9, выделенный и очищенный сотрудницей лаборатории эпигенетики развития ИЦиГ СО РАН Устьянцевой Е. И.

Для гидролиза использовали молярное соотношение вектор:sgРНК:SpCas9 = 1:100:100. Объем реакции составлял 20 мкл. Смешивали sgРНК и белок Cas9 в реакционном буфере (5% глицерин, 20 мМ HEPES pH 7.5, 100 мМ KCl, 2 мМ MgCl₂, 1 мМ ДТТ, 0.5 мМ ЭДТА). Выдерживали при комнатной температуре 10 мин для образования комплексов sgРНК:SpCas9, после чего добавляли 100 нг ДНК линеаризованного вектора и инкубировали 3 часа при 37 °С. Реакцию инактивировали, добавляя 2 мкл ЭДТА 0,5М и 1,6 мкл протеинкиназы К 25мг/мл к 20 мкл реакционной смеси, 20 мин при комнатной температуре. Результаты реакции анализировали с помощью электрофореза в 1% агарозном геле.

2.4.7 Секвенирующая реакция

Продукты ПЦР продукт был очищали с помощью выделения из агарозного геля. Для секвенирующей реакции использовали 20-50 нг очищенного продукта ПЦР. При количественном анализе активности системы CRISPR/Cas9 для секвенирующих реакции использовали 100 нг ДНК плазмидной ДНК.

Секвенирующие реакции проводили с использованием набора реактивов BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Образцы проанализированы в ЦКП "Геномика" ИХБФМ СО РАН.

2.4.8 Выделение фрагментов ДНК из агарозного геля

Для очистки фрагментов ДНК путем выделения из агарозного геля использовали набор Zymoclean Gel DNA Recovery Kit (Zymo Research, D4008). Выделение проводили согласно протоколу производителя.

2.5 Выявление потенциальных нецелевых сайтов действия CRISPR/Cas9

Потенциальные нецелевые сайты выявляли с помощью программного обеспечения CasOT (<http://eendb.zfgenetics.org/casot/>), Cas-OFFinder (<http://www.rgenome.net/cas-offinder/>), CRISPR Design (<http://crispr.mit.edu/>), COSMID (<https://crispr.bme.gatech.edu/>), GT-SCAN (<http://gt-scan.braembl.org.au/gt-scan/>) (Bae et al., 2014, Cradick et al., 2014, Hsu et al., 2013, O'Brien, Bailey, 2014, Xiao et al., 2014). Для анализа использовали аннотированную последовательность генома крысы *Rattus norvegicus* Rnor_5.0.75 (Ensembl).

ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1 Создание донорного вектора для гомологичной рекомбинации в локусе *Avr*

Для исправления мутации в гене *Avr* с помощью методов молекулярного клонирования был сконструирован плазмидный вектор p*Avr*-target8, содержащий два участка гомологичных целевому гену. В качестве основы был использован вектор p079 pPNT6 (Addgene #11072), содержащий элементы необходимые для позитивно-негативной селекции: 1) ген устойчивости к антибиотику неомицину (*NeoR*), 2) ген тимидинкиназы вируса простого герпеса (*HSV-TK1*). Тимидинкиназа вируса простого герпеса фосфорилирует ганцикловир, который затем при фосфорилировании клеточными киназами образует трифосфат и включается в ДНК при синтезе, обрывая его, что приводит к гибели клеток. Так называемые «плечи гомологии» были получены путем амплификации участков гена *Avr* помощью ПЦР с использованием в качестве матрицы геномной ДНК крыс контрольной линии WAG, которые являются гомозиготами по аллелю *Avr* дикого типа. Амплифицированные участки были клонированы в вектор p079 pPNT6, длина плеч гомологии составила ~3.2 т.п.н. и ~2.3 т.п.н., что, согласно литературным данным, является оптимальной длиной для прохождения гомологичной рекомбинации (Beumer et al., 2013).

В полученном векторе плечи гомологии фланкируют ген *NeoR* (Рисунок 11), таким образом при гомологичной рекомбинации этот ген будет встроен в локус *Avr*, и, соответственно, целевые клетки могут быть селектированы добавлением в среду антибиотика неомицина или его аналогов. Также *NeoR* фланкирован сайтами *loxP*, что позволит в случае необходимости при дальнейших исследованиях с помощью экспрессии в рекомбинантных клетках Cre-рекомбиназы удалить этот трансген, в результате чего останется только небольшой «шрам» – один сайт *loxP*. Ген *HSV-TK1* расположен за пределами донорного участка, соответственно, при гомологичной рекомбинации попадать в геном не должен. При добавлении в культуральную среду ганцикловира будет проходить негативная селекция: клетки, в которых произошла случайная встройка вектора, будут погибать.

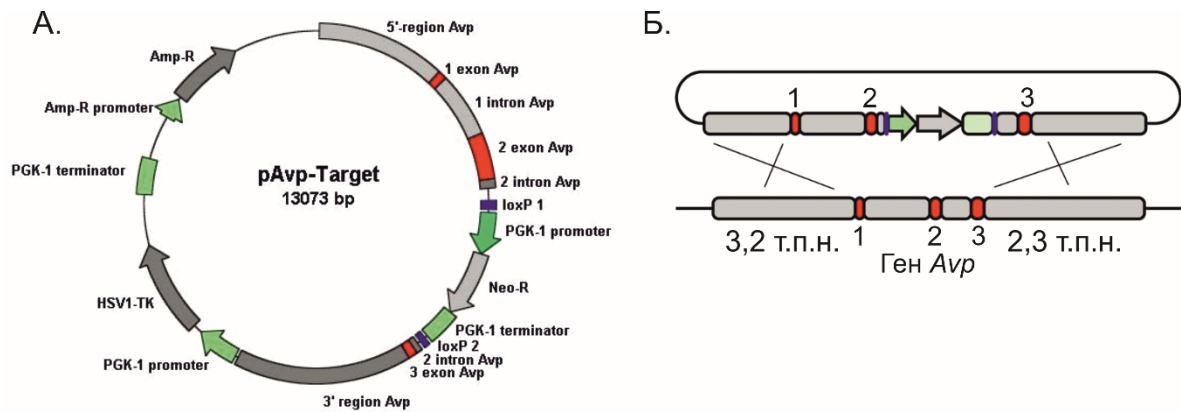


Рисунок 11. А. Карта донорного вектора pAvp-Target8, Б. Схема целевого события – гомологичной рекомбинации локуса *Avp* с донорным вектором. Цифрами обозначены экзоны гена *Avp*.

3.2 Выбор последовательностей протоспейсеров для действия CRISPR/Cas9 в локусе *Avp*

Для увеличения частоты прохождения гомологичной рекомбинации с донорным плазмидным вектором и, как следствие, замены последовательности ДНК в геноме и введения трансгена в целевой сайт необходимо, чтобы ДЦР возникал в пределах плеч гомологии в непосредственной близости от последовательности, которую необходимо заменить исследователю (Yang et al., 2013). Соответственно, для увеличения частоты гомологичной рекомбинации во втором экзоне гена *Avp* крыс линии Brattleboro на участке вблизи мутации (в пределах 100 нуклеотидов) были выявлены потенциальные протоспейсеры (содержащие с 3'-конца мотив PAM SpCas9 – NGG) CRISPR/Cas9. Всего в данной области гена обнаружено 18 потенциальных протоспейсеров для действия CRISPR/Cas9, состоящих из 20 нуклеотидов и содержащих на 3'-конце мотив –NGG (PAM SpCas9).

При внесении модификаций в ген аргинин-вазопрессина потенциальной мишенью нецелевого действия CRISPR/Cas9 в рамках предложенной нами модельной системы является ген гормона окситоцина (*Oxt*) – паралог целевого гена с практически идентичной последовательностью ДНК. Для предотвращения нецелевого эффекта в гене *Oxt* были разработаны три стратегии выбора сайтов действия CRISPR/Cas9, использующие различия между двумя генами (*Avp* и *Oxt*): 1) использование самой мутации в качестве отличия между генами, 2) выбор

протоспейсера на негомологичном участке – в интроне гена *Avp*, 3) использование двух нуклеаз Cas9 для внесения ДЦР, (Рисунок 12А, Б, В).

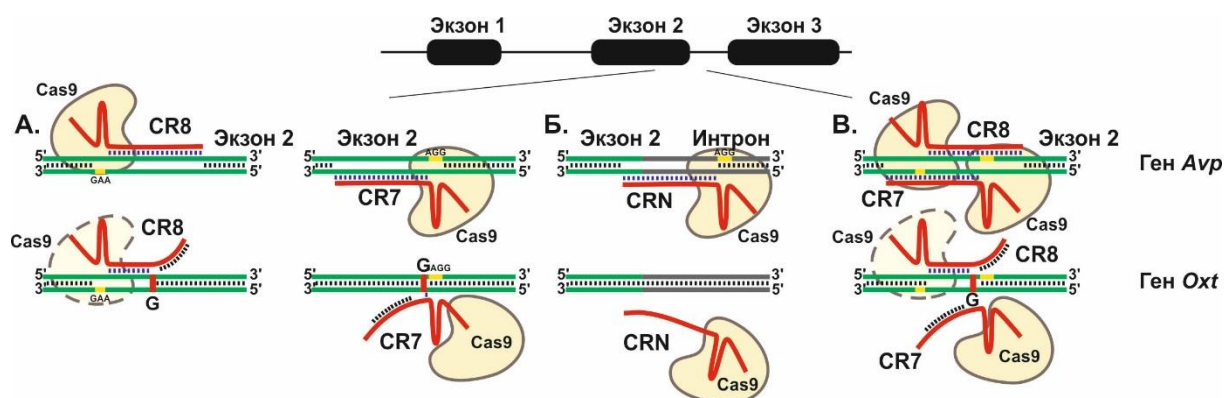


Рисунок 12. Разработка стратегий для направленного внесения ДЦР в ген гормона аргинин-вазопрессина в эмбриональных фибробластах крыс линии Brattleboro. А, Б, В – Схема расположения сайтов для действия системы CRISPR/Cas9. Зеленым изображена кодирующая часть гена, серым – интронная часть гена. Красным изображены sgРНК, узнающие соответствующий протоспейсер. Желтым обозначены последовательности мотивов, прилежащих к протоспейсеру (РАМ). *Avp* – ген гормона аргинин-вазопрессина, *Oxt* – ген гормона окситоцина.

В рамках первой стратегии было выбрано два протоспейсера для действия комплекса sgРНК + нуклеаза Cas9: CR-7 и CR-8 (Таблица 4). В мутантной последовательности гена *Avp* крыс линии Brattleboro РАМ -AGG непосредственно прилежит с 3'-конца к протоспейсеру CR-7, а в последовательности дикого типа в гене *Oxt* между протоспейсером и РАМ находится один нуклеотид (G), в результате чего спейсер сдвигается на один нуклеотид относительно протоспейсера. В случае протоспейсера CR-8 в гене *Oxt* первые 10 нуклеотидов с 3'-конца спейсера совпадают с последовательностью протоспейсера, а затем происходит сдвиг на один нуклеотид. По литературным данным наиболее «важными» для узнавания протоспейсера системой CRISPR/Cas9 являются 8-10 нуклеотидов на 3'-конце, при этом эта система с определенной частотой способна вносить нецелевые ДЦР в сайты, отличающиеся от целевого протоспейсера заменами на 5'-конце (Fu et al., 2013). Исходя из этих соображений, мы предполагаем теоретическую возможность внесения ДЦР в последовательность гена *Oxt* CRISPR/Cas9 со спейсером CR-8.

В рамках второй стратегии был использован тот факт, что делеция у крыс Brattleboro произошла близко к стыку экзон-интрон (37 нуклеотидов). Выбран протоспейсер CR-N, большая часть которого и PAM находится в интроне гена *Avr* (Таблица 4). Так как только кодирующие части генов *Avr* и *Oxt* гомологичны, то в локусе гена *Oxt* протоспейсера CR-N нет. Для внесения ДЦР в сайты CR-7, CR-8 и CR-N были сконструированы плазмидные векторы, кодирующие соответствующую sgRNA и нуклеазу Cas9 – pX330_CR-7, pX330_CR-8, pX330_CR-N.

Таблица 4. Последовательности целевых сайтов в гене *Avr* крыс линии Brattleboro

Название	Последовательность протоспейсера	PAM*
CR-7	CCAGAAGCCTTGCGGAAGCG	-AGG
CR-8	GCAGCGGCCTCGCTTCCGCA	-AGG
CR-N	GCGATGGTGCGCACAAAGCC	-AGG
* - мотив, прилежащий к протоспейсеру		

Наконец, в третьей стратегии для снижения потенциальных нецелевых эффектов был реализован подход с использованием двух никак – мутатных нуклеаз Cas9n(D10A), которые вносят разрыв только в одну цепь ДНК. Два одноцепочечных разрыва на расстоянии до 100 нуклеотидов друг от друга приводят к двухцепочечному разрыву. Использование двух никак уменьшает риск случайных нецелевых эффектов (Ran et al., 2013). Поэтому для снижения вероятности возникновения нецелевых эффектов, в том числе и в гене *Oxt*, были сконструированы парные векторы, кодирующие sgRNA и никакы Cas9n(D10A) – pX335_CR-7 и pX335_CR-8.

3.3 Сравнение активности CRISPR/Cas9 в выбранных сайтах гена *Avr*

3.3.1 Анализ с использованием эндонуклеазы I фага T7

Для определения способности элементов системы CRISPR/Cas9 вносить ДЦР в целевом сайте гена *Avr* проводили электропорацию сконструированных плазмидных векторов в эмбриональные фибробласты крыс линии Brattleboro с помощью электропорации. Эффективность электропорации составила ~95% (Рисунок 13А). Через двое суток анализировали продукты ПЦР с помощью эндонуклеазы I фага T7. Мы обнаружили, что ДЦР происходят во всех целевых протоспейсерах, причем при использовании двух никак также детектируются следы

ДЦР (Рисунок 13Б). При этом следов ДЦР на участке второго экзона гена *Oxt* обнаружено не было, следовательно, все разработанные стратегии позволяют избежать нецелевого эффекта в паралогичном гене.

3.3.2 Рестрикционный анализ

В результате делеции во втором экзоне гена *Avp* крыс линии Brattleboro возникает сайт узнавания эндонуклеазы рестрикции *BcgI* (Рисунок 13В). В данной работе мы использовали этот факт, чтобы отличить мутантную последовательность от последовательности дикого типа. С помощью электропорации в эмбриональные фибробласты крыс линии Brattleboro доставляли плазмидные векторы pX330_CR-7, pX330_CR-8 или pX335_CR-7+ pX335_CR-8. Через двое суток выделяли геномную ДНК и проводили рестрикционный анализ продуктов ПЦР после амплификации участка второго экзона гена *Avp*, используя эндонуклеазу рестрикции *BcgI*. В результате этого анализа показано, что продукт ПЦР, полученный при использовании в качестве матрицы геномной ДНК опытных образцов, гидролизует не полностью (Рисунок 13Г). Следовательно, в опытных образцах пропадает сайт узнавания эндонуклеазы рестрикции *BcgI*, что возможно является результатом репарации ДЦР, которые вносит CRISPR/Cas9.

С помощью описанных выше методик было показано, что все разработанные стратегии по внесению ДЦР в ген *Avp* соответствуют заданному нами критерию – система CRISPR/Cas9 вносит ДЦР в целевой ген *Avp*, не вызывая нецелевых разрывов в гене-паралоге *Oxt*.

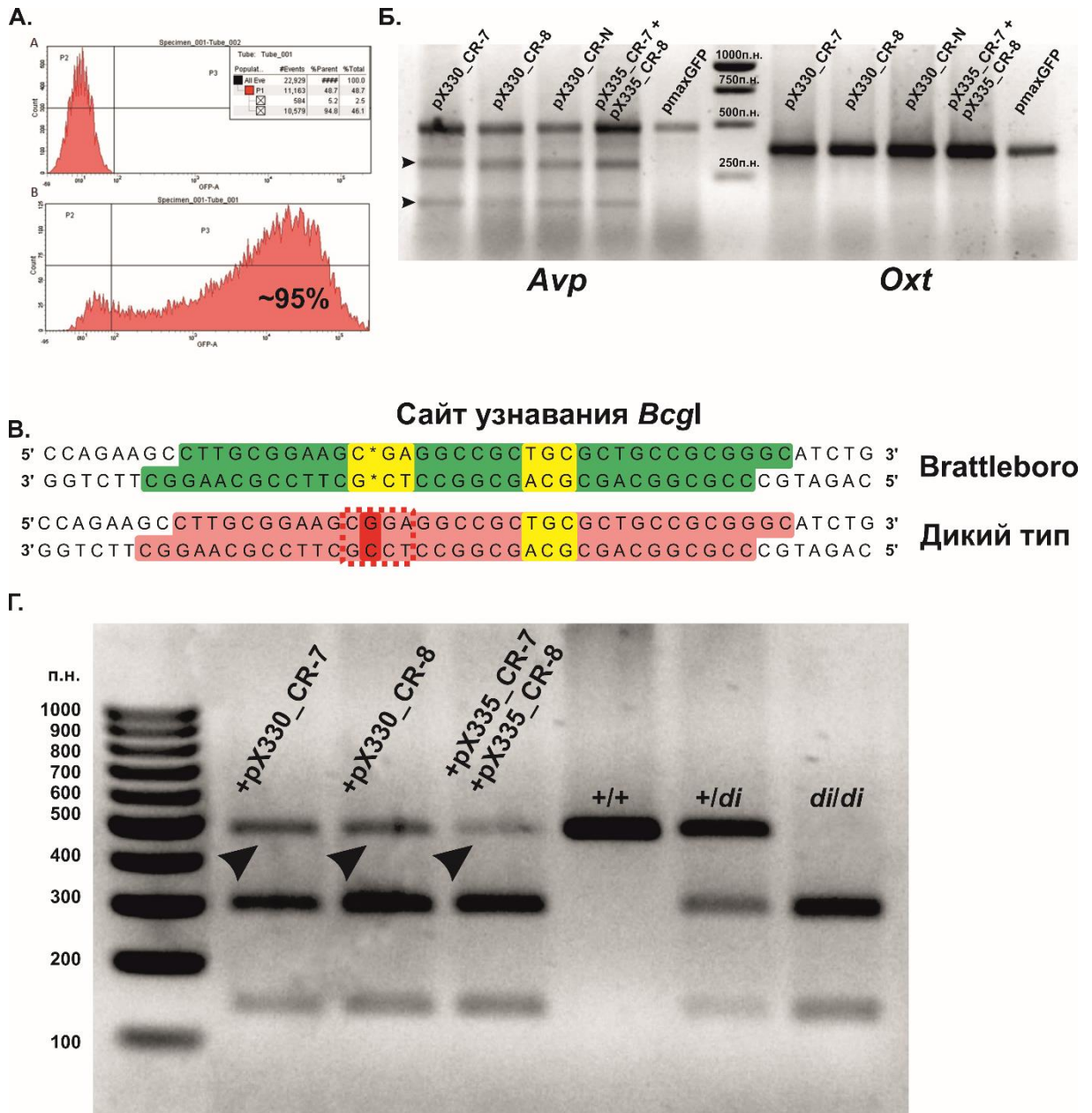


Рисунок 13. А – Оценка эффективности электропорации. Использовали 2мкг контрольного вектора pmaxGFP (Lonza), 10^6 клеток. Б – результаты анализа с помощью эндонуклеазы I фага T7. Стрелками обозначены продукты гидролиза. Наличие продуктов гидролиза свидетельствует о том, что в целевой ген были внесены ДЦР. В – сайт узнавания эндонуклеазы рестрикции *BcgI* в гене *Avp*. Желтым обозначены нуклеотиды, которые узнает *BcgI*. Зеленым – участок, вырезаемый *BcgI*. Знаком * обозначена делеция. Красным обозначены соответствующие последовательности в аллеле гена *Avp* дикого типа. Г – результаты рестрикционного анализа. +/+ – образец дикого типа. +/di – гетерозиготный образец. Стрелками обозначены не гидролизованные продукты в опытных образцах.

3.4 Биоинформатический анализ потенциальных нецелевых эффектов

Были выбраны несколько протоспейсеров в гене *Avr* (CR-7, CR-8 и CR-N) и показано, что в них происходят ДЦР под действием системы CRISPR/Cas9, и не детектируются следы разрывов в гене *Oxt*. Однако, кроме гена *Oxt* нецелевые разрывы могут возникать и в других сайтах генома, имеющих сходную последовательность с целевым. Что определить, какой из выбранных протоспейсеров имеет наименьшее число нецелевых сайтов в геноме, был проведен биоинформатический анализ.

Современные биоинформатические инструменты для определения потенциальных нецелевых сайтов действия системы CRISPR/Cas9 используют разные алгоритмы и при их сравнении было показано, что, хотя определенные различными инструментами протоспейсеры для CRISPR/Cas9 на заданном участке генома совпадают, количество выявленных потенциальных нецелевых сайтов может различаться (Ishida et al., 2015). Также стоит отметить, что большинство этих инструментов выявляет нецелевые сайты, содержащие однонуклеотидные замены относительно канонического сайта, однако в нескольких работах было показано, что нецелевые ДЦР могут происходить также в сайтах, содержащих инсерции или делеции (Bae et al., 2014, Lin et al., 2014, Ran et al., 2015). В данной работе ввиду приведенных выше факторов для выявления потенциальных нецелевых сайтов были использованы несколько доступных в настоящее время биоинформатических инструментов: CasOT, Cas-OFFinder, CRISPR Design, COSMID и GT-Scan. В качестве входных данных для программного обеспечения использовали последовательности спейсеров CRISPR/Cas9 CR-7, CR-8, CR-N и полную последовательность генома крысы.

3.4.1 Программное обеспечение CasOT

Для определения потенциальных сайтов нецелевого действия системы CRISPR/Cas9 в геноме (кроме гена *Oxt*) был проведен анализ с помощью программного обеспечения CasOT для каждого из выбранных спейсеров. Вероятность внесения ДЦР в нецелевом сайте, отличающегося от целевого однонуклеотидными заменами, зависит от количества и характера распределения этих замен (Fu et al., 2013).

Таблица 5. Пять наиболее вероятных нецелевых сайтов для выбранных спейсеров, выявленных с помощью CasOT*.

Спейсер	Последовательность сайта в геноме	PAM	Число замен	Ранг
CR-7	ctgagccg_CTTGCGGAAGCG	TGG	8	A08
	aCAGgAGC_CTTGCGcAAGCG	CGG	3	A12
	GgAccAGC_CTTGgGGAAGCG	CGG	4	A13
	aCAGtgGC_CTTGtGGAAGCG	AGG	4	A13
	GaAGAAat_CTTGCaGAAGCG	TGG	4	A13
Общее число сайтов	3625			
CR-8	GCtGCacC_CTCGCTTCCGCA	GGG	3	A03
	Gatcaca_CTCGCTTCCGCA	CGG	7	A07
	GgAGgaGC_CTgGCTTCCGCA	GGG	4	A13
	GaAcCttC_CTcCCTTCCGCA	TGG	5	A14
	cCgGCtaC_CTcCCTTCCGCA	GGG	5	A14
Общее число сайтов	3935			
CR-N	GtagTGGT_GCGCACAAAGCC-CGGG	CGG	3	A03
	GCcccGGa_GCGCACAAAGCC-TGGA	TGG	4	A04
	aCtggGag_GCGCACAAAGCC-TGGA	TGG	6	A06
	GacccGcc_GCGCACAAAGCC-AGGG	AGG	6	A06
	caccccT_GCGCACAAAGCC-TGGT	TGG	7	A07
Общее число сайтов	6795			
*Замены в нецелевых сайтах относительно канонического обозначены строчными буквами				

Выявленные нецелевые сайты были ранжированы, после чего определяли: 1) общее число потенциальных нецелевых сайтов 2) количество замен в сайте, для которого наиболее вероятно внесение нецелевого ДЦР (первый в рейтинге) (Таблица 5). Программное обеспечение CasOT присваивает нецелевым сайтам на основании числа и характера отличий от целевого сайта ранг, представляющий собой двузначное число, где первая цифра – число замен в пределах первых 8 нуклеотидов от 3' конца (наиболее важный для узнавания участок), вторая цифра – число замен на 5'-конце сайта. Соответственно, чем больше присвоенный ранг, тем больше замен в нецелевом сайте и тем меньше вероятность внесения нецелевого ДЦР. Данная система позволяет при ранжировании учитывать характер замен – оказывающие наибольшее влияние на вероятность нецелевого эффекта замены в первых 8 нуклеотидах дают вклад во второй разряд двузначного числа ранга в то время, как замены на 5'-конце влияют на первый разряд (Xiao et al., 2014). И наоборот, чем меньше ранг, тем меньше различия между целевым и нецелевым сайтом и тем больше вероятность нецелевого эффекта. Что также важно инструмент CasOT

позволяет оценивать нецелевые сайты внесения ДЦР при использовании двух никаз. Для этого определяется ранги нецелевых сайтов каждого спейсера, расположенных в пределах 100 нуклеотидов друг от друга. Соответственно, выявляли нецелевые сайты для пары никаз pX335_CR-7 + pX335_CR-8 (Таблица 6).

Таблица 6. Наиболее вероятные нецелевые сайты при использовании двух никаз, выявленных с помощью CasOT*.

pX335_CR-7			pX335_CR-8			Расстояние между сайтами, п.н.
Последовательность сайта в геноме	РАМ	Ранг	Последовательность сайта в геноме	РАМ	Ранг	
tCctgctC_CcTGCGGAAGCG	AGG	A16	GCtGCacC_CTCGCTTCCGCA	GGG	A03	18
cCAGtctC_CTCGCTTCCaCA	AGG	A14	aCAGtgGC_CTTGtGGAAGCG	AGG	A13	18
aCgtacaC_CTCGtTCCGCA	AGG	A16	GgAaggaC_CTTGCGGAaCG	AGG	A15	18
tCcactGC_CTTcCGGAAGCG	AGG	A15	GgcctccC_CTCGCTTCCGgA	AGG	A16	18
taAcaGcC_CTGCTTCCGCA	AGG	A15	agcGcctC_CTTGCGGAAGCc	AGG	A16	18
*Замены в нецелевых сайтах относительно канонического обозначены строчными буквами						

3.4.2 Программное обеспечение Cas-OFFinder

Программное обеспечение Cas-OFFinder позволяет обнаружить нецелевые сайты, содержащие не только содержащие замены относительно канонического сайта, но также и инсерции и делеции (Bae et al., 2014). Однако, в отличие от предыдущего инструмента алгоритм Cas-OFFinder не позволяет ранжировать выявленные нецелевые сайты по характеру замен, демонстрируя лишь их общее число (Bae et al., 2014). Это программное обеспечение больше подходит для использования в качестве инструмента для быстрого выявления и сравнения общего количества нецелевых сайтов и замен в них, что также позволяет оценивать специфичность выбранного спейсера CRISPR/Cas9 (Ishida et al., 2015). С помощью данного инструмента было определено общее количество нецелевых сайтов для протоспейсеров CR-7, CR-8 и CR-N. Если не учитывать характер замен в нецелевых сайтах, то, согласно результатам данного анализа, более безопасно использовать спейсер CR-N, так как для него общее число нецелевых сайтов меньше, чем для спейсеров CR-7 и CR-8 (Таблица 7).

Таблица 7. Общее число нецелевых сайтов в геноме для выбранных спейсеров.

Спейсер	Минимальное число замен в нецелевом сайте	Общее число нецелевых сайтов
CR-7	3	280704
CR-8	3	216071
CR-N	3	202761

3.4.3 Программное обеспечение CRISPR Design

Программное обеспечение CRISPR Design разработано на основе результата научной работы, проведенной коллективом Фенга Чжана, по исследованию специфичности действия CRISPR/Cas9 (Hsu et al., 2013). Данный инструмент каждому нецелевому сайту присваивает ранг, который зависит от числа и характера распределения замен относительно канонического сайта. Чем выше ранг, тем выше, соответственно, вероятность внесения ДЦР в нецелевом сайте.

Согласно результатам, проведенного с помощью данного инструмента анализа, менее вероятно возникновение нецелевых эффектов при использовании спейсера CR-7 – ранг наиболее вероятного нецелевого сайта для этого спейсера ниже, чем для других спейсеров (Таблица 8).

Таблица 8. Пять наиболее вероятных нецелевых сайтов для выбранных спейсеров, выявленных с помощью CRISPR Design.

Спейсер	Нецелевой сайт	Ранг	Число и положение замен	Идентификатор гена UCSC	Позиция в геноме
CR-7	ACAGGAGCCTTGCGCAAGCGCGG	0,8	3, [1:5:15]	—	chr7:-121340821
	GAAGAATCCTTGCGGAAGAGAAG	0,7	3, [2:7:19]	—	chr2:-159642441
	AAACAAGCCTTGCGGAAGCCCGG	0,7	4, [1:2:4:20]	—	chr9:-2996143
	GGACCAGCCTTGCGGAAGCGCGG	0,6	4, [2:4:5:13]	—	chr2:-231000840
	GCAGCAGCCTGCGGAAGTGC GG	0,6	3, [5:11:19]	—	chr3:+178218659
Общее число сайтов - 161 (21 в пределах известных генов)					
CR-8	GCTGCACCCTCGCTTCCGCAGGG	1,0	3, [3:6:7]	NM_001005885	chr1:-171719487
	CCTTCCGCCTCGCTTCCGCACAG	0,8	4, [1:3:4:6]	NM_017231	chr10:+64272977
	GGAACGGGCTGGCTTCCGCATAG	0,8	4, [2:4:8:11]	—	chr14:-91673445
	GCAGCCGCTTGTCTCCGCAGAG	0,5	3, [6:9:11]	—	chr3:-86502084
	GCTGCACCCTCGCTTCCGCAGGG	1	3, [3:6:7]	NM_001005885	chr1:-171719487
Общее число сайтов - 75 (18 в пределах известных генов)					
CR-N	GGGATGGAGAGCACAAAGCCAAG	2,5	3, [2:8:10]	—	chr17:-88791213
	CCTATGGTGCTCACAAAGCCGGG	1,5	3, [1:3:11]	—	chr5:-170207584
	GAAATGGTGACACAAAGCCAG	1,5	3, [2:3:11]	—	chr8:+11906425
	GCAATAGTGAGCACAAAGCCTGG	1,4	3, [3:6:10]	—	chr16:-36654651
	GGCCCGGAGCGCACAAAGCCTGG	1,3	4, [3:4:5:8]	—	chr9:+46273498
Общее число - 78 (9 в пределах известных генов)					

3.4.4 Программное обеспечение COSMID

Программное обеспечение COSMID было использовано ввиду возможности поиска нецелевых сайтов, отличающихся от канонического делециями и инсерциями (Cradick et al., 2014). Этот алгоритм поиска и анализа нецелевых сайтов также присваивает каждому сайту ранг – чем меньше ранг, тем выше вероятность возникновения нецелевого эффекта (у целевого сайта ранг 0,00). В результате анализа был получен список наиболее вероятных нецелевых сайтов в геноме, содержащих не более 3 замен и/или инсерций или делеций не более 2 нуклеотидов (Таблица 9). Согласно данному алгоритму ранжирования нецелевых сайтов менее вероятны нецелевые эффекты при использовании спейсера CR-7 (ранг наиболее вероятного нецелевого сайта – 2,21).

Таблица 9. Наиболее вероятные нецелевые сайты для выбранных спейсеров, выявленные с помощью COSMID*.

Спейсер	Нецелевой сайт	Инсерции /делеции	Число замен	Позиция в геноме	Ранг
CR-7	aCAGgAGCCTTGCGcAAGCGCGG	нет	3	Chr7:-121350831	2,21
	GCAGAAAGCtgTGCGGAAGgGAGG	нет	3	Chr20:+25574841	5,85
	GCAGcAGCCTgGCGGAAGtGCGG	нет	3	Chr3:+172166905	5,89
	GCAcAAGCCTTGCaGAAGCtGGG	нет	3	Chr16:+21175918	7,47
	GCAGtAGCCTTGCaGAAGCcAGG	нет	3	Chr17:+5545082	7,49
Общее число	17				
CR-8	GCtGCacCCTCGCTTCCGCAGGG	нет	3	Chr1:-165518246	0,59
	GCAGCaGCCcCGCTTCCGCAGGC	нет	3	Chr1:-227117881	20,71
Общее число	2				
CR-N	GtagTGGTGCGCACAAAGCCCGG	нет	3	Chr8:+115110696	0,45
	GCaATaGTGcGCACAAAGCCTGG	нет	3	Chr16:-36849242	0,86
	cCgATGGTGctCACAAAGCCGGG	нет	3	Chr5:-166561440	0,97
	GCGATGGTctaCACAAAGCCTGG	нет	3	Chr9:-49084890	1,55
	GCGATGGTGCGCACAAgGCCAGG	нет	1	Chr3:-123117921	3
Общее число	12				
*Замены в нецелевых сайтах относительно канонического обозначены строчными буквами					

3.4.5 Программное обеспечение GT-SCAN

Наконец, последний использованный алгоритм – GT-SCAN – хотя и не присваивает нецелевым сайтам ранги, все же сортирует их из характера замен, которые они содержат: сперва по числу замен в области с высокой чувствительностью (10 нуклеотидов с 3'-конца протоспейсера), затем по числу

замен в области с низкой (на 5'-конце протоспейсера) (O'Brien, Bailey, 2014). Соответственно, получен список нецелевых сайтов, содержащих не более трех замен относительно канонического сайта. Для спейсера CR-7 предсказанная вероятность нецелевых эффектов ниже относительно других выбранных спейсеров – наиболее вероятный сайт нецелевого эффекта содержит 1 замену в чувствительной области по сравнению с отсутствием замен в чувствительной области для спейсеров CR-8 и CR-N (Таблица 10).

Таблица 10. Наиболее вероятные нецелевые сайты для выбранных спейсеров, выявленные с помощью GT-SCAN*.

Спейсер	Нецелевой сайт	Всего замен	Число замен в области с высокой чувствительностью	Число замен в области с низкой чувствительностью
CR-7	GCAGAAGC tg TGCGGAAG g GAGG	3	1	2
	a CAG g AGCCTTGCG c AAGCGCGG	3	1	2
	GCAG t AGCCTTG Ca GAAG Cc AGG	3	2	1
	GCAG g AGCCTTGCGGAAG Ct CaG	3	2	1
CR-8	G Ct GC ac CCTCGCTTCCGCAGGG	3	0	3
CR-N	G tag TGGTGCGCACAAAGCCCGG	3	0	3
	G Ca AT a GT Ga GCACAAAGCCTGG	3	0	3
	GCGATGGT tca CACAAAGCCTGG	3	1	2
	c CtATGGTG Ct CACAAAGCCGGG	3	1	2
	G t GATGGTG Cc CACACAGCCTGG	3	2	1
	GCGATGGT tCa CACAA At CCTGG	3	2	1
*Замены в нецелевых сайтах относительно канонического обозначены строчными буквами				

На основании проведенного биоинформатического анализа с использованием различных алгоритмов были предсказаны потенциальные нецелевые сайты действия системы CRISPR/Cas9 с выбранными в рамках разработанных стратегий спейсерами CR-7, CR-8 и CR-N. На основании полученных в ходе анализа результатов для дальнейшего использования был выбран спейсер CR-7, поскольку согласно всем алгоритмам, за исключением Cas-OFFinder, предсказанная вероятность нецелевых эффектов ниже, чем для спейсеров CR-8 и CR-N.

3.5 Определение эффективности внесения двуцепочечных разрывов системой CRISPR/Cas9 в протоспейсер CR-7

Чтобы определить эффективность внесения ДЦР системой CRISPR/Cas9 в протоспейсер CR-7, выбранный на основании биоинформатического анализа, а

также изучить спектр модификаций, возникающий при репарации данных разрывов, в исходный вектор pX330_CR-7 был добавлен ген устойчивости к антибиотику пурамицину (Pu_{ro}R) (Рисунок 14А). В результате позитивной селекции может быть получена популяция клеток, в которые в результате электропорации была успешно доставлена ДНК CRISPR/Cas9 – что, соответственно, позволяет точнее определить эффективность внесения ДЦР (не будет вклада не трансфецированных клеток), а также за счет обогащения популяции изучить каким образом в данном протоспейсере происходит репарация ДЦР.

Из вектора pX330_CR-7 с помощью эндонуклеаз рестрикции *Age*I и *Eco*RI вырезали последовательность гена SpCas9 и вместо неё клонировали в вектор короткий двуцепочечный олигонуклеотид. Данный двухцепочечный олигонуклеотид был получен путем отжига комплементарных одноцепочечных олигонуклеотидов, подобранных таким образом, что в результате отжига образовывались липкие концы (Рисунок 14Б). Гидролиз вектора эндонуклеазами рестрикции и лигирование ДНК проводили в одной циклической реакции (37 °C 5 мин, 20 °C 5 мин, 15 циклов). Липкие концы олигонуклеотида комплементарны липким концам вектора, которые образуются после гидролиза *Age*I и *Eco*RI, однако после лигирования ДНК сайты узнавания эндонуклеаз рестрикции не восстанавливались – копия вектора, в которую был успешно клонирован олигонуклеотид, в следующем цикле реакции гидролизу уже не подвергалась. Таким образом в результате циклической реакции происходило постепенное накопление копий векторов с целевой встройкой. В полученный промежуточный вектор pX330_CR-7_Ann клонировали фрагмент ДНК, который содержал последовательности генов SpCas9 и Pu_{ro}R, разделенные последовательностью 2A-пептида. Данный фрагмент ДНК был получен при гидролизе одновременно четырьмя эндонуклеазами рестрикции (*Mlu*I, *Xba*I, *Kpn*I, *Xma*I) плазмидного вектора lentiCRISPR_v2 (Addgene #52961) и очищен путем выделения ДНК из агарозного геля (Рисунок 14А).

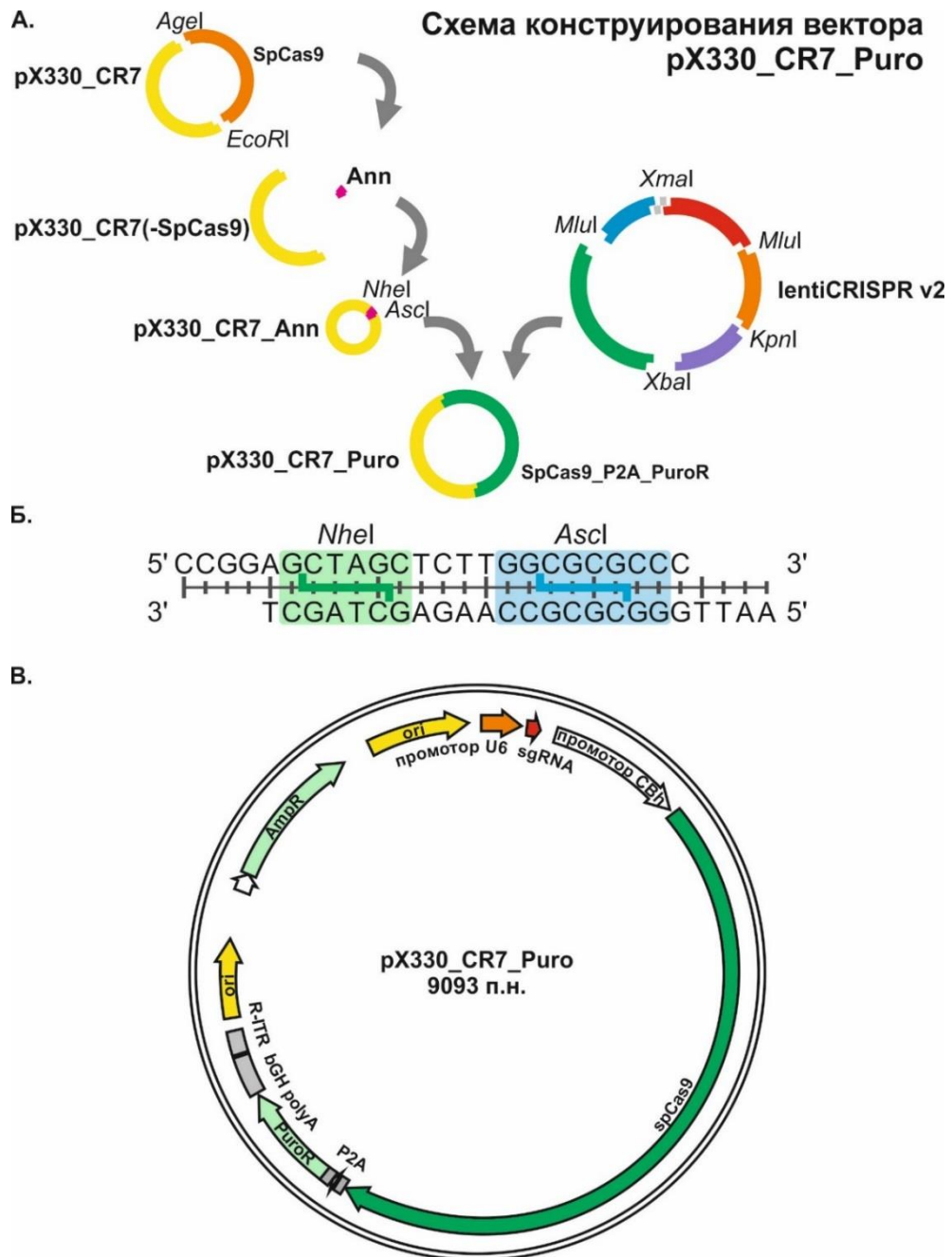


Рисунок 14. А – схема конструирования вектора pX330_CR-7_Puro. Кратко: из вектора pX330_CR-7 с помощью эндонуклеаз рестрикции *AgeI* и *EcoRI* был удален ген *spCas9*, вместо которого в вектор был клонирован двухцепочечный олигонуклеотид (Ann), содержащий сайты узнавания эндонуклеаз рестрикции *NheI* и *AscI*. В полученный вектор pX330_CR-7_Ann был клонирован фрагмент *spCas9_P2A_PuroR*, полученный при гидролизе эндонуклеазами рестрикции *MluI*, *XbaI*, *KpnI*, *XmaI* вектора lentiCRISPR v2. Липкие концы, образующиеся при гидролизе эндонуклеазами *MluI* и *XbaI* комплементарны липким концам, которые образуется при гидролизе *AscI* и *NheI*, соответственно. Б – последовательность двухцепочечного нуклеотида Ann с обозначенными сайтами узнавания эндонуклеаз рестрикции, В – карта плазмидного вектора pX330_CR-7_Puro. п.н. – пары нуклеотидов

Последовательность 2А-пептида обеспечивает одновременную экспрессию генов SpCas9 и PuroR под управлением общего промотора CBh (Рисунок 14В). Очищенный фрагмент ДНК был клонирован в вектор pX330_CR-7_Ann, гидролизированный с помощью эндонуклеаз рестрикции *NheI* и *AscI*, сайты узнавания которых были привнесены в этот вектор ранее при клонировании олигонуклеотида. Таким образом был получен плазмидный вектор pX330_CR-7_Puro.

3.5.1 Определение частоты целевых эффектов в гене *Avr* при использовании спейсера CR-7

Проводили электропорацию эмбриональных фибробластов линии RNFF1 вектором pX330_CR-7_Puro, после селекции устойчивые к пурамицину клетки наращивали до плотности ~90%, после чего анализировали. Всего было получено 3 опытных образца геномной ДНК из трех повторностей при электропорации. С использованием олигонуклеотидных праймеров Avr6f и Avr6r амплифицировали целевой участок гена *Avr*. В качестве контроля использовали геномную ДНК интактных эмбриональных фибробластов RNFF1. Продукты ПЦР очищали с помощью выделения из агарозного геля, после чего проводили секвенирующие реакции с использованием тех же олигонуклеотидных праймеров. Таким образом было получено по 3 опытных образца для каждого из генов и по 1 контрольному. Для анализа эффективности внесения модификаций (результат репарации ДЦР) в целевой участок использовали программное обеспечение TIDE (Tracking of Indels by DEcomposition), доступное онлайн <https://tide-calculator.nki.nl/>. Суть данного алгоритма заключается в сравнении секвенограмм, полученных при анализе продуктов секвенирующих реакций опытного и контрольного образцов, и выявлении факта наличия/отсутствия делеций или инсерций в сайте действия CRISPR/Cas9 (Brinkman et al., 2014).

В результате проведенного биоинформатического анализа было показано, что средняя эффективность (среднее из 6 значений: 3 образца, прямой и обратный праймеры для каждого) внесения модификаций целевого сайта в клетках, в которых экспрессировалась CRISPR/Cas9, составила 78,7%. При этом в ~60% случаев происходит инсерция одного нуклеотида (Таблица 11, Рисунок 15А). Используемое программное обеспечение позволяет оценить вероятность инсерции нуклеотидов в случае инсерции +1 (Рисунок 15Б).

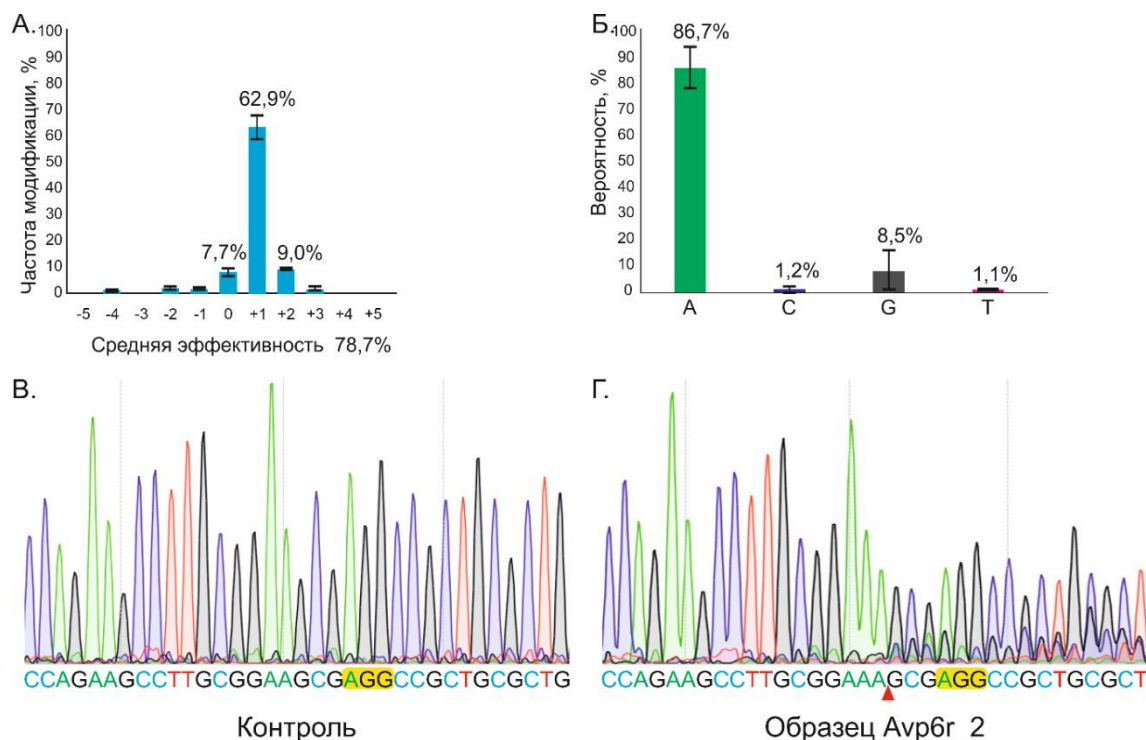


Рисунок 15. Результаты анализа эффективности внесения ДЦР в протоспейсер CR-7 с помощью TIDE (Brinkman et al., 2014). А. Анализ частоты мутаций в протоспейсере CR-7 при репарации ДЦР. По оси абсцисс указан размер делеции/инсерции. Б. Вероятность инсерции нуклеотидов. Указаны средние значения для всех образцов $\pm$ стандартное отклонение. В. Секвенограмма контрольного образца. Г. Секвенограмма опытного образца. Красным указатель - место внесения ДЦР. Желтым обозначен РАМ протоспейсера CR-7.

Таблица 11. Результаты анализа с помощью TIDE частоты модификаций в гене *Avp*

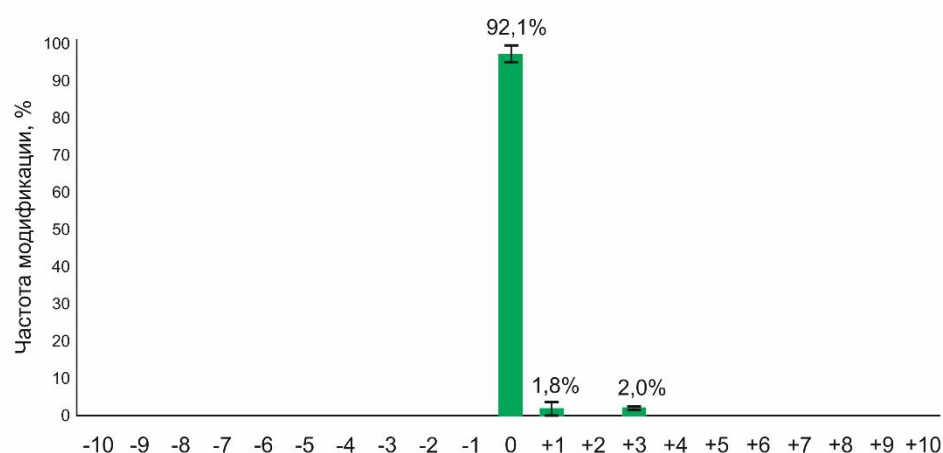
Мутац ия	Avp6f_1		Avp6r_1		Avp6f_2		Avp6r_2		Среднее	SD*
	%	p	%	p	%	p	%	p		
-10	0	1	0	1	0	0.96	0.5	0.78	0,1	0,3
-9	0	1	0	1	0	1	0	1	0,0	0,0
-8	0	1	0	1	0	1	0	1	0,0	0,0
-7	0	1	0	1	0	1	1.5	0.35	0,4	0,8
-6	0	1	0	1	0	1	0	1	0,0	0,0
-5	0	1	0	1	0.1	0.92	0	1	0,0	0,1
-4	1.6	0.11	0	1	1.5	0.16	0	1	0,8	0,9
-3	0.1	0.91	0	1	0.6	0.6	0	1	0,2	0,3
-2	0.7	0.49	1.3	0.25	1.4	0.18	2.5	0.1	1,5	0,8
-1	0.4	0.67	2.3	0.032	0.5	0.66	1.5	0.32	1,2	0,9
0	7.5	2.10E-14	9.8	5.10E-20	6.5	2.90E-10	7.1	1.60E-06	7,7	1,4
1	61.3	0	69	0	57.9	0	63.5	0	62,9	4,7
2	8.6	2.50E-18	8.7	3.10E-16	9.8	2.40E-21	9	2.00E-09	9,0	0,5
3	1	0.32	2.2	0.036	1.1	0.3	2.4	0.12	1,7	0,7
4	0.6	0.56	0.7	0.48	0.3	0.78	0.9	0.56	0,6	0,3
5	0	1	0	1	0.3	0.77	0	1	0,1	0,2
6	0.5	0.6	0	1	0.6	0.56	0	1	0,3	0,3
7	0	1	0	1	0.1	0.9	0	1	0,0	0,1
8	0	1	0	1	0	1	0	1	0,0	0,0
9	0	1	0	1	0	1	0	1	0,0	0,0
10	0	1	0	1	0	1	0	1	0,0	0,0

*SD - Стандартное отклонение, % - частота мутации, p – p-значение

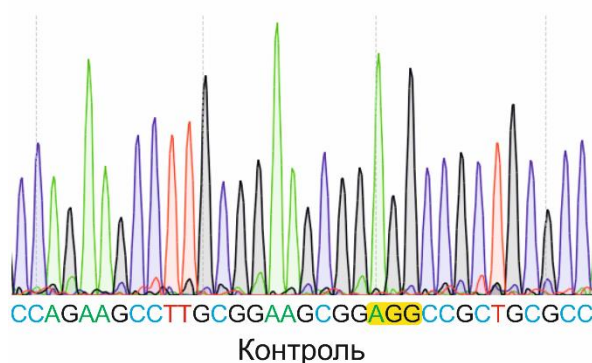
3.5.2 Определение частоты нецелевых эффектов в гене *Oxt* при использовании спейсера CR-7

Используя геномную ДНК, полученную после электропорации рХ330_CR-7_Puro и селекции (см. предыдущий раздел), и олигонуклеотидные праймеры Oxt4f и Oxt4r были получены продукты ПЦР, которые далее использовали в качестве матрицы для секвенирующих реакций. Полученные секвенограммы были проанализированы с помощью программного обеспечения TIDE для определения частоты нецелевых эффектов в участке гена *Oxt*, гомологичном протоспейсеру CR-7 (Рисунок 16).

А.



Б.



В.

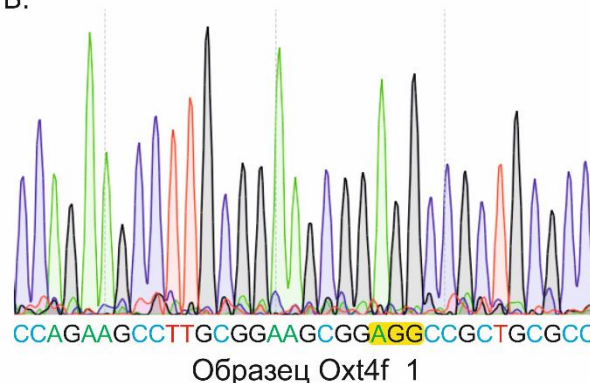


Рисунок 16. Результаты анализа эффективности внесения ДЦР в в гене *Oxt* с помощью программного обеспечения TIDE (Brinkman et al., 2014). А. Анализ частоты модификаций целевой последовательности в результате репарации потенциальных ДЦР. По оси абсцисс указан размер делеции/инсерции. На графике указаны средние значения для всех образцов $\pm$ стандартное отклонение, Б. Секвенограмма контрольного образца. В. Секвенограмма опытного образца. Желтым обозначен РАМ.

Показано, что в среднем 92,1% копий участка гена *Oxt* не содержат инсерций или делеций (Рисунок 16А). Выявленные инсерции +1 и +3 (1,8% и 2%) не являются

статистически достоверными ($p > 0,1$ – по данным TIDE, Таблица 12) и, скорее всего, представляют собой «шум» на секвенограмме. Также между секвенограммами опытного и контрольного образцов не наблюдается значительных отличий (Рисунок 16Б, В). Таким образом, показано, что при использовании системы CRISPR/Cas9 со спейсером CR-7 достоверно не выявляются следы внесения ДЦР в паралогичном гене *Oxt*. Это говорит о том, что ДЦР в нецелевом сайте не происходит, либо их частота ниже детектируемого уровня использованного метода анализа.

Таблица 12. Результаты анализа TIDE частоты модификаций в гене *Oxt*

Мутация	Oxt4f_1		Oxt4f_2		Oxt4f_3		Среднее	SD*
	%	p	%	p	%	p		
-10	0	1	0	1	0	1	0,0	0,0
-9	0	1	0.5	0.82	0	1	0,2	0,3
-8	0	1	0	1	0	1	0,0	0,0
-7	0	1	0	1	0	1	0,0	0,0
-6	0	1	0	1	0	1	0,0	0,0
-5	0	1	0	1	0	1	0,0	0,0
-4	0	1	0	1	0	1	0,0	0,0
-3	0	1	0	1	0	1	0,0	0,0
-2	0	1	0	1	0	1	0,0	0,0
-1	0	1	0.4	0.87	0	1	0,1	0,2
0	96.3	0	91.1	6.60E-279	89	0	92,1	3,8
+1	0	1	0	1	5.5	0.011	1,8	3,2
+2	0	1	0.8	0.76	0	1	0,3	0,5
+3	1.2	0.56	2.3	0.38	2.4	0.29	2,0	0,7
+4	0	1	0.5	0.86	0	1	0,2	0,3
+5	0	1	0.1	0.98	0	1	0,0	0,1
+6	0	1	0	1	0	1	0,0	0,0
+7	0	1	0	1	0	1	0,0	0,0
+8	0	1	0	1	0	1	0,0	0,0
+9	0	1	0	1	0	1	0,0	0,0
+10	0	1	0	1	0	1	0,0	0,0

*SD - Стандартное отклонение, % - частота мутации, p – p-значение

3.5.3 Изучение спектра модификаций, вносимых системой CRISPR/Cas9 в протоспейсер CR-7

Чтобы подтвердить данные, полученные при биоинформатическом анализе, продукт ПЦР целевого участка гена *Avr* (матрица – геномная ДНК, раздел 3.5.1) клонировали в вектор pGEM-T Easy (Promega), после чего трансформировали продуктами лигирования ДНК *E. coli* (штамм XL-10 Gold). Индивидуальные колонии *E. coli* инокулировали в жидкую среду LB и растили в течение 16 часов, после чего выделяли плазмидную ДНК. Полученные образцы плазмидной ДНК, содержащие последовательность целевого участка гена *Avr*, использовали в качестве матрицы в секвенирующих реакциях. Для образца 1 были получены

результаты секвенирующих реакций с плазмидной ДНК 25 индивидуальных колоний, для образца 2 – 33 индивидуальных колоний.

Результаты данного анализа в целом совпадают с результатами анализа TIDE: видна преимущественная репарация с инсерцией А (в среднем 49,2%). Значение частоты инсерций А несколько ниже, чем при анализе TIDE, однако это может быть объяснено тем, что в последнем не были учтены крупные делеции, выявленные при секвенировании плазмидной ДНК клонов *E. coli*. Средняя эффективность внесения ДЦР составила 82,4% – посчитано, как среднее из двух значений “100 - доля интактных последовательностей (wt, Рисунок 17)”.

А. Образец 1

Протоспейсер				PAM				МГ
TGG	CCAGAAGCCTTGCGGAA	-	-	GCG	AGGCC			
TGGCCAGAAGCCTTGCGGAA	A	-	-	GCGAGGCC	56,0%	+A		
TGGCCAGAAGCCTTGCGGAA	AA	-	-	GCGAGGCC	16,0%	+AA		
TGGCCAGAAGCCTTGCGGAA	-	-	-	GCGAGGCC	8,0%	wt		
TGGCCAGAAGCCTTGCGGA	-	-	-	GCGAGGCC	8,0%	ΔA		
-	-	-	-	-	4,0%	Δ123	GGCG	
-	-	-	-	-	4,0%	Δ71	?	
-	-	-	-	-	4,0%	Δ45	?	

Б. Образец 2

Протоспейсер				PAM				МГ
TGG	CCAGAAGCCTTGCGGAA	-	-	GCG	AGGCC			
TGGCCAGAAGCCTTGCGGAA	A	-	-	GCGAGGCC	42,4%	+A		
TGGCCAGAAGCCTTGCGGAA	-	-	-	GCGAGGCC	27.3%	wt		
TGGCCAGAAGCCTTGCGGAA	AA	-	-	GCGAGGCC	9,1%	+AA		
TGGCCAGAAGCCTTGCGGA	-	-	-	GCGAGGCC	3,0%	ΔA		
-	-	-	-	-	3,0%	Δ206	AG	
-	-	-	-	-	3,0%	Δ146	?	
-	-	-	-	-	3,0%	Δ38	TGC	
-	-	-	-	-	3,0%	Δ6	GCG	

Рисунок 17. Спектр мутаций, возникающих в результате репарации ДЦР, вносимых CRISPR/Cas9 в целевом сайте CR-7. А. Образец 1, анализ плазмидной ДНК 25 колоний. Б. Образец 2, анализ плазмидной ДНК 33 колоний. Красной стрелкой обозначено место предполагаемого ДЦР. wt – последовательность ДНК гена *Avr* крыс Brattleboro. МГ – обнаруженные микрогомологии, ? – микрогомология не очевидна.

Таким образом, с помощью секвенирования по Сенгеру подтверждены результаты биоинформатического анализа. Интересным представляется тот факт,

что репарация ДЦР в протоспейсере CR-7 происходит в больше чем половине случаев, по-видимому, по одному пути. Данная инсерция предположительно происходит при репарации по классическому пути соединения негомологичных концов (с-NHEJ), который является активным на протяжении всего клеточного цикла эукариот, и может быть результатом активности ДНК-полимераз μ и λ в составе репаративного комплекса (Lieber, 2010). Данная тенденция наблюдается во всех проанализированных образцах, что говорит о её неслучайности, что соответствует литературным данным, согласно которым, результат репарации ДЦР под действием CRISPR/Cas9, по-видимому, определяется самим генетическим контекстом и воспроизводится при повторных экспериментах (van Overbeek et al., 2016). Причем в некоторых описанных в литературе случаях преимущественным типом модификаций при репарации с-NHEJ разрывов CRISPR/Cas9 являются небольшие делеции – например, в протоспейсерах в локусе *HBB* (Phadke et al., 2016), а в других – инсерции, в частности, +A, причем в данном случае с 5' стороны от ДЦР находилась последовательность из нескольких A, как и в случае протоспейсера CR-7 (van Overbeek et al., 2016). Интересным также является факт выявленных инсерций +AA. Данные инсерции, очевидно, также являются результатом полимеразной активности репаративного комплекса, однако возникают с меньшей частотой чем +A (в среднем 12,55%).

Помимо инсерций нами были обнаружены также и делеции, которые могут возникать за счет различных путей репарации ДЦР. В части случаев наблюдаются микрогомологии, по которым происходила гибридизация одноцепочечных выступающих 3'-концов, которые образуются под действием экзонуклеазной 5'-3' экзонуклеазной активности MRN, привлекаемой к свободным концам ДНК, в случае если репарация разрыва ДНК не происходит достаточно долгий период времени. После гибридизации на участках микрогомологии происходит деградация выступающих концов, заполнение брешей полимеразми и лигирование (Goodarzi, Jeggo, 2013). В данном случае репарация происходит по пути альтернативного соединения концов (alt-EJ) (Ceccaldi et al., 2016, McVey, Lee, 2008). В исследованных образцах наблюдали делеции от 6 до 206 нуклеотидов, для которых были обнаружены сайты микрогомологий (2-4 нуклеотида, рисунок 17). В некоторых случаях микрогомологий выявлено не было, сложно судить о пути, в

результате которого возникли эти делеции (25, 71 и 146 нуклеотидов, рисунок 17). В случае же делеции –А, можно предположить, что она является результатом нуклеазной активности комплекса Artemis и ДНК-зависимой протеинкиназы (DNA-PK) при репарации по пути c-NHEJ. Возможно, конец 5'-GAA-3', образующийся при разрыве, в комплексе с Ku преимущественно привлекает полимеразы (инсерция +А) и в редких случаях нуклеазы (делеция А, рисунок 17).

Предположительно, наблюдаемая картина результатов репарации ДЦР является особенностью выбранного сайта – протоспейсера CR-7, однако без дополнительных исследований на других типах клеток и исследований с использованием, например, ингибиторов DNA-PK или LIG4, входящих в репаративный комплекс c-NHEJ, утверждать это наверняка не стоит.

3.6 Получение и анализ смешанных популяций в результате позитивно-негативной селекции

Чтобы проверить выбранную систему позитивно-негативной селекции, в эмбриональные фибробласты крыс линии Brattleboro RNFF1 с помощью электропорации доставляли смесь ДНК: 3 мкг pX330_CR-7 + 3 мкг pAvp-target8. Сперва в течение 7 суток проводили позитивную селекцию, добавляя в культуральную среду антибиотик G418. Клетки наращивали, рассаживали, после чего проводили ещё один раунд уже позитивно-негативной селекции, добавляя в культуральную среду ганцикловир. Клетки после первого и второго раундов селекции анализировали на наличие в геноме трансгенов *NeoR* и *HSV-TK1* с помощью ПЦР (Рисунок 18).

В результате проведенного анализа видно, что после первого раунда селекции в геноме клеток полученных смешанных популяций присутствуют оба трансгена. После второго раунда селекции трансген *HSV-TK1* не детектируется, что говорит о том, что при селекции 1 выживают как клетки с предположительно целевой встройкой, так и клетки со случайными встройками донорного вектора. Последние при селекции с добавлением ганцикловира гибнут и, соответственно, в смешанных популяциях, полученных в результате селекции 2, детектируется только трансген *NeoR*.

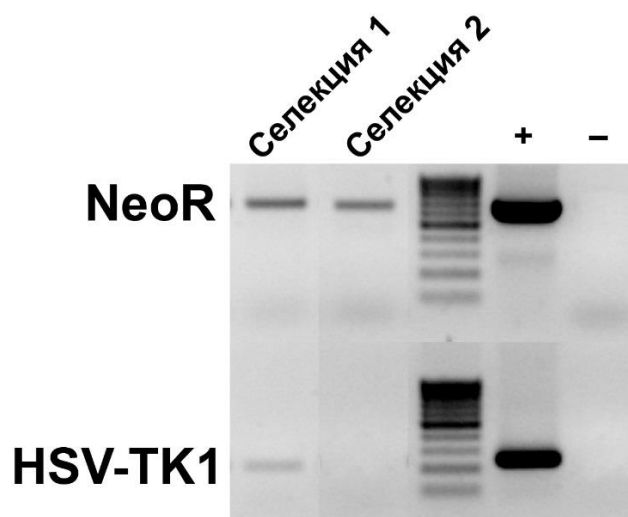


Рисунок 18. Анализ смешанных популяций, полученных в результате позитивно-негативной селекции, с помощью ПЦР. На рисунке приведены результаты ПЦР с использованием олигонуклеотидных праймеров для амплификации участков генов *NeoR* и *HSV-TK1*. В качестве положительного контроля (+) использовали ДНК плазмидного вектора pAvp-target8. Отрицательный контроль – вода.

В результате данного эксперимента было показано, что клетки, в которые была доставлена смесь векторов CRISPR/Cas9 и донорных векторов, после позитивно-негативной селекции содержат предположительно целевые встройки, произошедшие в результате гомологичной рекомбинации в целевом гене, либо случайные встройки участка донорного вектора, содержащего только ген *NeoR*. Соответственно, следующим шагом в данной работе стало проведение экспериментов по получению и анализу клональных гомогенных популяций клеток с последующим отбором тех линий, в геноме которых произошли целевые события.

3.7 Получение и анализ клональных линий

3.7.1 Позитивно-негативная селекция

Для получения клональных линий проводили электропорацию клеток линии RNFF1 (10^6 клеток/образец) смесью плазмидных векторов: 3 мкг pX330_CR-7_Puro + 3 мкг pAvp-target8. Клетки после электропорации пересевали на чашки Петри (диаметр 100 мм) в культуральную среду с добавлением активатора RAD51 7,6 мкМ RS-1 (R9782, Sigma Aldrich). На следующие сутки производили смену культуральной среды без добавления RS-1. Через 4 суток начинали позитивную селекцию – в течение 14 суток культивировали клетки в среде с добавлением 500 мкг/мл G418 – после этого этапа в результате селекции выживают целевые клетки

NeoR⁺/*HSV-TK1*⁻, клетки со случайными встройками донорного вектора *NeoR*⁺/*HSV-TK1*⁺, а также клетки как с целевыми, так и случайными встройками одновременно. После позитивной селекции, в течение 7 суток проводили негативную селекцию, добавления в культуральную среду ганцикловира до конечной концентрации 250 мкМ – после этого этапа должны выживать только клетки *NeoR*⁺/*HSV-TK1*⁻. При этом часть колоний клеток, образовавшихся после позитивной селекции погибла, очевидно, погибшие колонии произошли из клеток, содержащих случайные встройки донорного вектора. После негативной селекции выжившие клетки культивировали в течение ~14 суток в среде с добавлением bFGF. Образовавшиеся из прошедших селекцию клеток колонии механически пересевали (стеклянным капилляром) в индивидуальную культуральную посуду для получения клональных клеточных линий. После механической пересадки колоний оставшиеся на исходной чашке Петри клетки открепляли и пересевали на культуральные флаконы (25см²) для получения смешанных суммарных популяций (Рисунок 19).

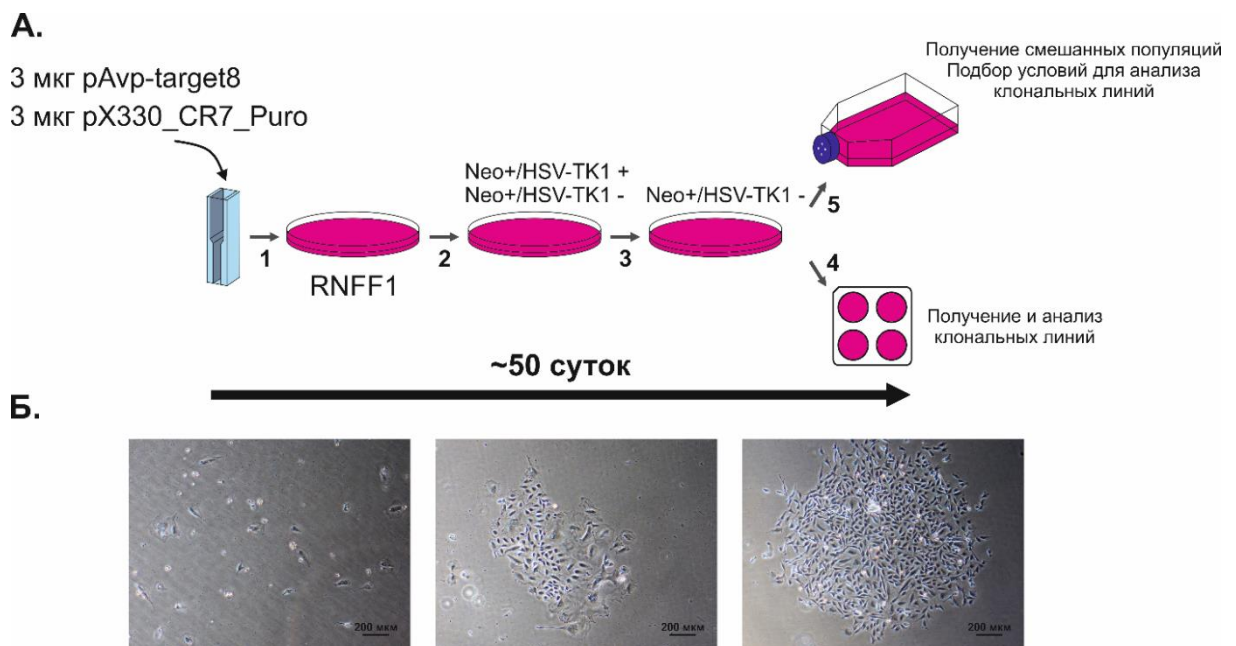


Рисунок 19. А – Схема эксперимента по получению клональных популяций в результате позитивно-негативной селекции, 1. Электропорация смесью векторов pX330_CR-7_Puro + pAvp-target8 2. позитивная селекция с добавлением G418 3. Негативная селекция с добавлением ганцикловира. После негативной селекции клетки культивировали в среде с добавлением bFGF до формирования индивидуальных колоний 4. Механический пересев колоний, получение клональных линий 5. Пересев оставшихся клеток на исходной культуральной поверхности, получение смешанных популяций Б – морфология клеток на различных этапах селекции – формирование колоний.

Электропорация была проведена на двух образцах, соответственно, в результате селекции были получены две смешанные популяции (Рисунок 19) – RNFF1_t1 и RNFF1_t2. Полученные смешанные популяции были использованы при подборе условий ПЦР для выявления клонов, в геноме которых произошла рекомбинация в целевом локусе.

3.7.2 Подбор условий ПЦР для выявления целевых клонов

Для выявления рекомбинантных клонов было подобрано 9 пар олигонуклеотидных праймеров (Рисунок 20А). Для подбора условий ПЦР в качестве матрицы использовали смесь геномной ДНК, выделенной из популяций RNFF1_t1 и RNFF1_t2. В результате проведения ПЦР с градиентом температур отжига для всех подобранных пар была выбрана пара олигонуклеотидных праймеров AVP_HR_L1F и AVP_HR_L1R (Рисунок 20Б). Размер продукта ПЦР с данной парой праймеров соответствовал теоретическому.

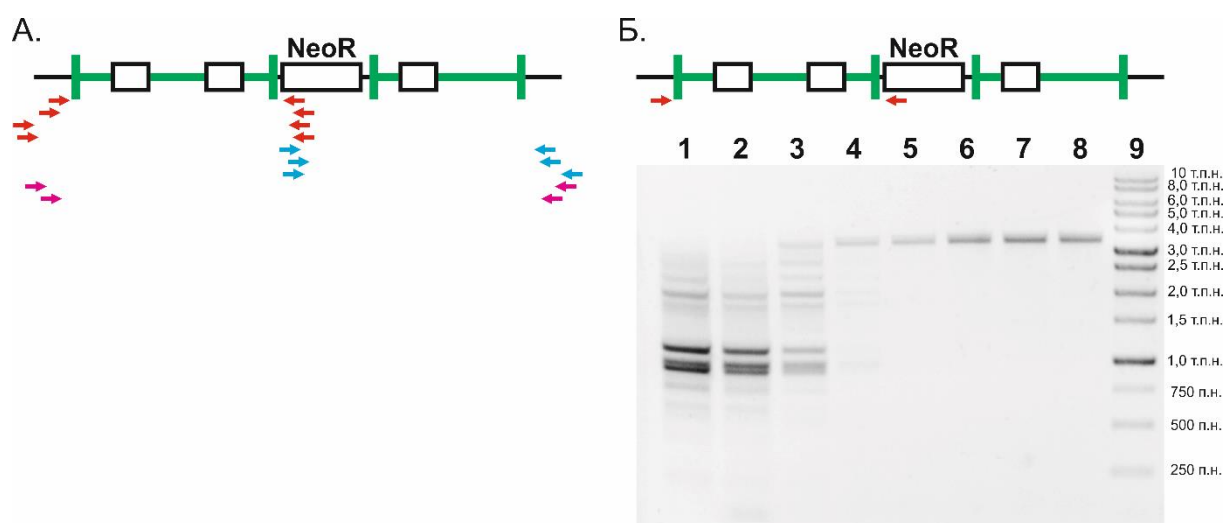


Рисунок 20. А – Схема расположения олигонуклеотидных праймеров в рекомбинантном аллеле *Avr*. Праймеры подобраны таким образом, что: а) один праймер вне участка гомологии между геном *Avr* и донорным вектором (участки гомологии обозначены зеленым цветом), другой в трансгене *NeoR* б) оба праймера за пределами участков гомологии Б – результаты ПЦР с градиентом температур отжига праймеров с использованием пары AVP_HR_L1F + AVP_HR_L1R. Дорожки 1-8 – градиент температур отжига от 53 °С до 63 °С, дорожка 9 – маркер молекулярного веса ДНК.

С помощью ПЦР показано, что в смешанных популяциях присутствуют клетки, в которых произошла гомологичная рекомбинация в целевом сайте гена *Avr*,

соответственно, далее проводили анализ полученных клональных линий для определения наличия среди них рекомбинантных образцов.

3.7.3 Анализ полученных клональных линий

В результате механической пересадки наиболее крупных колоний, прошедших позитивно-негативной селекции, для популяций RNFF1_t1 и RNFF1_t2 было получено 6 и 8 клональных линий, соответственно. Из полученных клональных линий была выделена геномная ДНК и проведен анализ с помощью ПЦР с использованием праймеров AVP_HR_L1F и AVP_HR_L1R. Соответствующий продукт был получен только для клонов RNFF1_t1.4 и RNFF1_t2.2 – в геноме этих клональных линий присутствует трансген *NeoR* в локусе гена *Avp*, что свидетельствует о произошедшей гомологичной рекомбинации с донорным вектором pAvp-target8 (Рисунок 21).

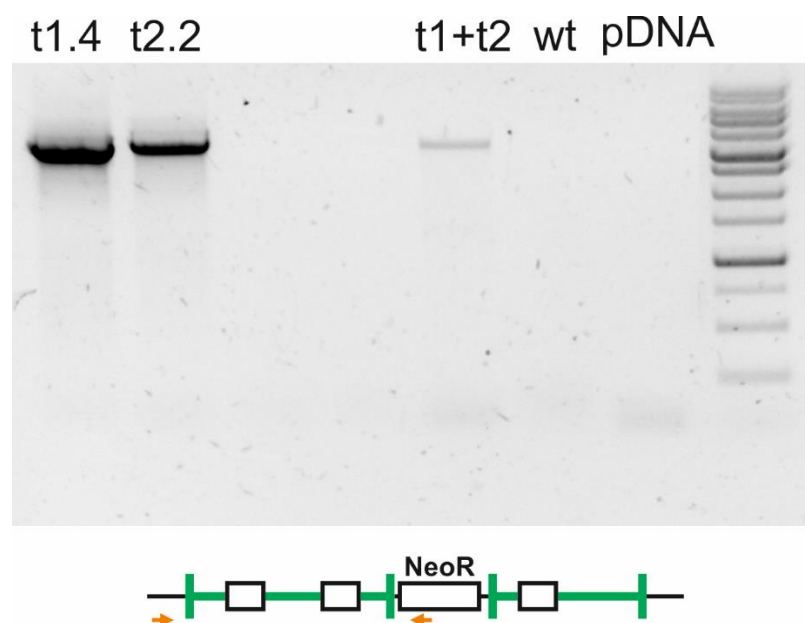


Рисунок 21. Выявление рекомбинантных клонов с помощью ПЦР. В качестве матрицы для ПЦР использовали: t1.4, t2.2 – геномную ДНК клонов RNFF1_t1.4 и RNFF1_t2.2, соответственно, t1+t2 – смесь геномной ДНК популяций RNFF1_t1 и RNFF1_t2, wt – геномная ДНК интактных фибробластов RNFF1, pDNA – донорный вектор pAvp-target8.

Чтобы определить произошло ли в данных клонах исправление мутации в гене *Avp*, продукты ПЦР очищали выделением из агарозного геля и проводили секвенирование по Сенгеру. При анализе результатов секвенирующих реакций было обнаружено, что в клоне RNFF1_t1.4 произошла делеция 205 п.н. (Рисунок 22А). Эти данные согласуются с тем, что при анализе с помощью электрофореза в агарозном

геле продукт ПЦР с использованием геномной ДНК RNFF1_t1.4 обладал большей электрофоретической подвижностью (масса продукта меньше), чем соответствующий продукт для клона RNFF1_t2.2 (Рисунок 21). Мы предполагаем, что в данном клоне уже после рекомбинации с донорным вектором в гене *Avr* система CRISPR/Cas9 повторно внесла ДЦР, который был репарирован с образованием делеции. Предположительно, данный разрыв был репарирован по механизму alt-EJ, при котором происходит образование выступающих одноцепочечных 3'-концов, которые могут гибридизоваться на участках микрогомологий (2-8 нуклеотидов), после чего происходит деградация выступающих концов, заполнение брешей и лигирование (McVey, Lee, 2008). По полученной секвенограмме видно, что в данном случае участком микрогомологии является последовательность 5'-GGG-3' (Рисунок 22Б). Данный клон в дальнейшей работе использован не был.

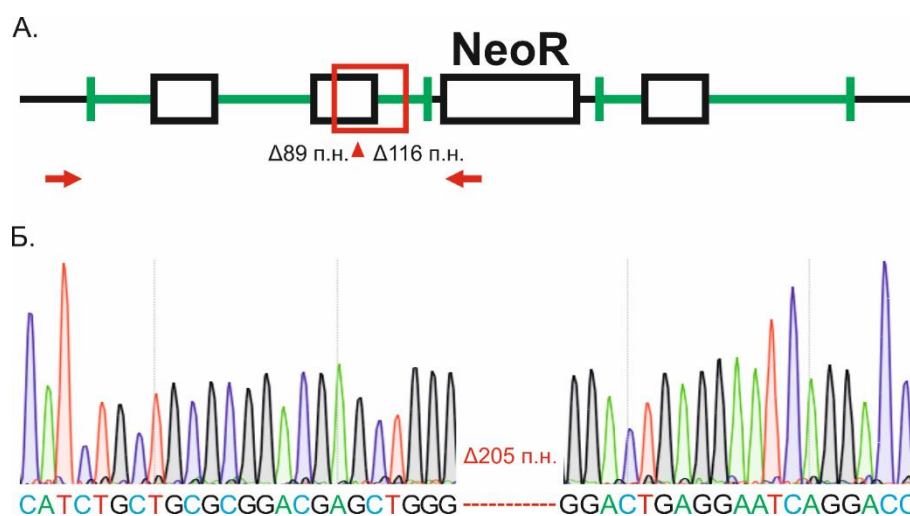


Рисунок 22. А. Схематическое изображение рекомбинантного аллеля гена *Avr* в клоне RNFF1_t1.4. Красной рамкой обозначен участок делеции. Красным указателем место предполагаемого ДЦР. Также указано число делетированных нуклеотидов «слева» и «справа» от места разрыва. Б. Результат секвенирования по Сенгеру рекомбинантного аллеля гена *Avr* в клоне RNFF1_t1.4.

При анализе результатов секвенирующей реакции с использованием в качестве матрицы продукта ПЦР AVP_HR_L1F+AVP_HR_L1R для клона RNFF1_t2.2 обнаружено, что в результате гомологичной рекомбинации произошло исправление мутации в гене *Avr* (Рисунок 23Б). Для того, чтобы определить является ли данный клон гетеро- или гомозиготой по рекомбинантному аллелю была

проведена ПЦР с использованием пары праймеров, лежащей в пределах плеч гомологии. Был получен продукт ПЦР, размер которого соответствовал последовательности, не содержащей трансгена *NeoR* (1124 п.н.). Соответственно, было показано, что клон RNFF1_t2.2 является гетерозиготой по рекомбинантному аллелю. Далее очищенный продукт ПЦР был секвенирован по Сенгеру (Рисунок 23В).

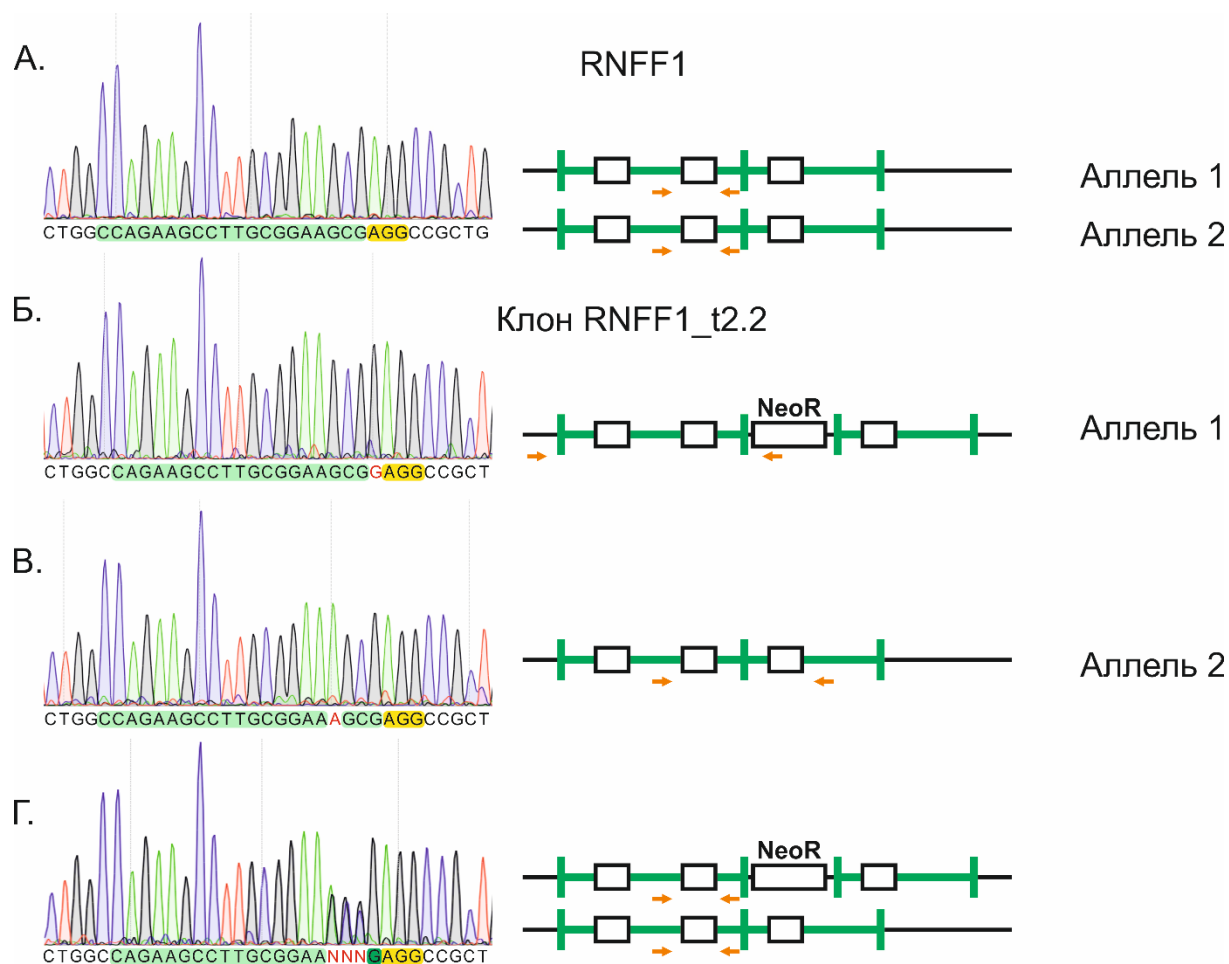


Рисунок 23. А. Секвенограмма контрольного образца интактных фибробластах RNFF1. Б. Результаты секвенирования по Сенгеру продукта ПЦР AVP_HR_L1F + AVP_HR_L1R – рекомбинантного аллеля. В. Результаты секвенирования по Сенгеру продукта ПЦР Avp_LF8 + Avp_RR1 – аллеля в котором не произошла рекомбинация. Г. Результаты секвенирования по Сенгеру продукта ПЦР Avp6f + Avp6r.

Показано, что аллель, в котором не произошла гомологичная рекомбинация с донорным вектором, содержит инсерцию А – мутацию, которая, по нашим данным, (раздел 3.5.1) возникает в большинстве случаев при репарации ДЦР, вносимых CRISPR/Cas9 в протоспейсере CR-7.

Также было проведено секвенирование продукта ПЦР, полученного с использованием праймеров Avr6f и Avr6r (Рисунок 23Г), который амплифицируется с обоих аллелей. На секвенограмме видно, что интенсивность пиков, соответствующих разным аллелям одинакова, что говорит о равных долях обоих аллелей. Наконец, чтобы определить не произошло ли нецелевых разрывов в гене *Oxt*, участок данного локуса был амплифицирован с использованием праймеров *Oxt4f* и *Oxt4r*. Очищенный продукт ПЦР был использован в качестве матрицы для секвенирования по Сенгеру. Показано, что последовательность гена *Oxt* осталась интактной и не содержит следов возникновения ДЦР (Рисунок 24).

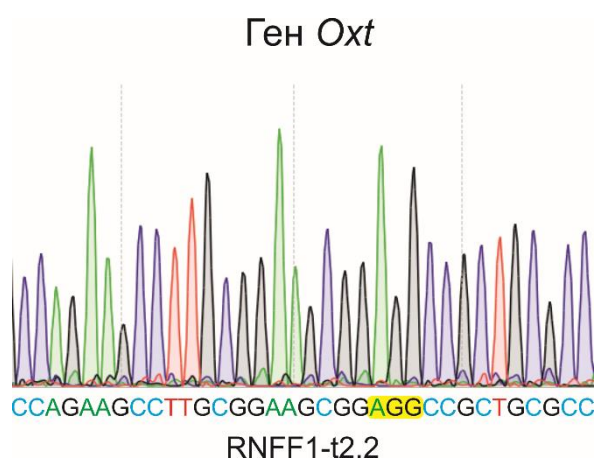


Рисунок 24. Результаты секвенирования по Сенгеру локуса *Oxt* в линии RNFF1_t2.2

Таким образом, был получен клон, содержащий один аллель гена *Avr* с исправленной в результате гомологичной рекомбинации мутацией *di* и аллель, содержащий мутацию, возникшую в результате репарации ДЦР. Однако, стоит отметить, что полученная нами мутация в гене *Avr* – инсерция А – приводит к восстановлению сдвинутой в результате делеции G у Brattleboro рамки считывания, что может привести к восстановлению нормального процессинга прогормона с образованием нейрофизина II (Рисунок 25). Данная мутация приводит к замене двух аминокислот: происходит замена полярного незаряженного (при pH = 7) серина и неполярного глицина на положительно заряженные лизин и аргинин, соответственно. Неясно, как повлияют данные замены на структуру белка

нейрофизина II и, что более важно, на его функцию, как белка переносчика аргинин-вазопрессина.

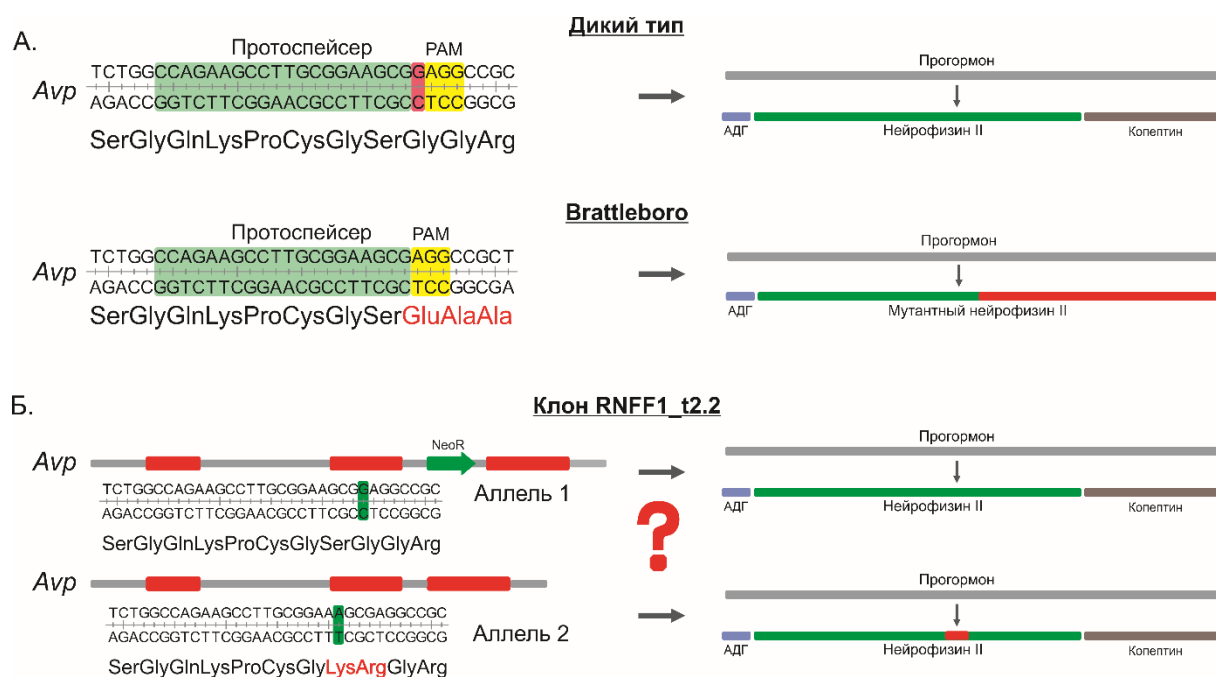


Рисунок 25. А. Сравнение участка последовательности гена *Avp* дикого типа и аллеля *di* крыс Brattleboro и схематичное изображение результатов процессинга прогормона. Б. Последовательности участков аллелей гена *Avp* в клоне RNFF1_t2.2 и соответствующие им последовательности аминокислотных остатков.

3.8 Обсуждение эффективности гомологичной рекомбинации

В результате проведенного эксперимента по получению клональных линий, в локусе *Avp* которых произошла гомологичная рекомбинация с донорным вектором, было получено всего два клон RNFF1_t1.4 и RNFF1_t2.2 – по одному клону с каждой популяции клеток, в которые была проведена электропорация. Учитывая тот факт, что в данном эксперименте эффективность трансфекции составила ~16%, эффективность гомологичной рекомбинации составила: $1/(10^6 \times 0,16) = 6,25 \times 10^{-6}$. Если учесть, что эффективность гомологичной рекомбинации по разным оценкам составляет около 10^{-6} , а применение системы CRISPR/Cas9 увеличивает её на несколько порядков, то ожидаемая нами эффективность HR – 10^{-4} , то есть при эффективности электропорации 16% мы ожидали получить порядка 16 клонов, из которых несколько были бы гомозиготами по рекомбинантному аллелю.

Соответственно, можно утверждать, что в нашей экспериментальной системе наблюдается низкая частота гомологичной рекомбинации с донорным вектором. Второй факт, который вызывает ряд вопросов, – делеция 205 нуклеотидов в рекомбинантном аллеле клона RNFF1_t1.4. По-видимому, данная делеция возникла уже после рекомбинации, что говорит о том, что созданная система CRISPR/Cas9 способна вносить ДЦР в последовательность дикого типа.

Для проверки этой гипотезы был использован выделенный и очищенный белок SpCas9 в комплексе с sgPНК, которая была получена с помощью реакции *in vitro* транскрипции, в которой в качестве матрицы был использован продукт ПЦР, содержащий последовательность sgPНК со спейсером CR-7 и с 5'-конца промотор T7 PНК-полимеразы. В результате данного анализа было показано, что комплексы SpCas9:sgPНК_CR-7 в реакции *in vitro* способны гидролизовать линейаризованный донорный вектор pAvp-target8 (Рисунок 26Б).

В изначальном дизайне эксперимента мы исходили из гипотезы, что, так как в последовательности дикого типа гена *Avp* между PAM –AGG и протоспейсером CR-7 расположен G, разрезания такого сайта происходить не будет, потому что PAM не прилежит непосредственно к протоспейсеру (раздел 3.2, Рисунок 26В). Таким образом, предполагалось защитить донорный вектор от ДЦР, а также избежать нецелевых разрывов в гене *Oxt*. При анализе литературы, касающейся системы CRISPR/Cas9 и, в частности, белка SpCas9, было обнаружено, что для SpCas9 показано внесение разрывов в протоспейсеры с неканоничным PAM –NAG (Zhang et al., 2014). При анализе последовательности *Avp* дикого типа был обнаружен неканоничный PAM –GAG, непосредственно прилежащий к протоспейсеру CR7 (Рисунок 26В). При исследованиях неканоничного –NAG показано, что для него характерна сниженная примерно на порядок эффективность нокаута трансгена *GFP* в клетках НЕК-293 по сравнению с каноничным –NGG (Zhang et al., 2014).

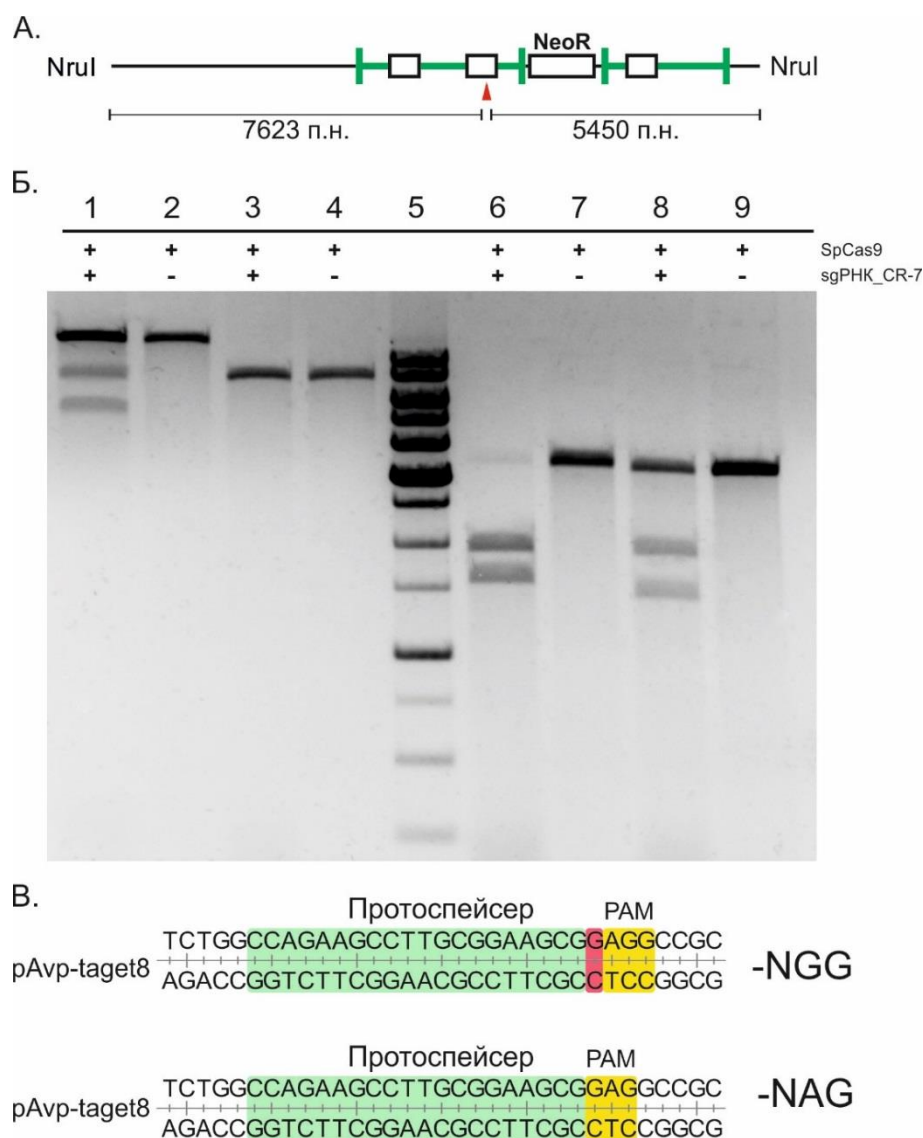


Рисунок 26. А. Схематичное изображение вектора pAvp-target8, линейаризованного с помощью эндонуклеазы рестрикции *NruI*. Красным указателем отмечено место предполагаемого разрыва в протоспейсере CR7. Б. Результаты *in vitro* реакции комплексов SpCas9:sgPHK и ДНК: 1-2 линейаризованный с помощью *NruI* pAvp-target8; 3-4 Вектор pX552_Avp_Puro – отрицательный контроль; 5 – маркер молекулярного веса ДНК; 6-7 положительный контроль – линейаризованный вектор pGEM-T Easy, в который был встроена целевая последовательность гена *Avp*, содержащая протоспейсер CR7; 8-9 линейаризованный вектор pGEM-T Easy, в который был встроена последовательность гена *Oxt*. Векторы pGEM-T Easy линейаризовали с помощью *BsaI*. В. К протоспейсеру CR7 в векторе pAvp-target8 прилежит неканоничный PAM – NAG.

В качестве положительного контроля при проведении реакции *in vitro* комплексов SpCas9:sgPHK_CR7 использовали вектор pGEM-T Easy, в который был клонирован участок последовательности гена *Avp* крыс Brattleboro, содержащий

протоспейсер CR_7 (встраивали продукт ПЦР с праймерами Avr6f и Avr6r). Стоит отметить, что в данном случае эффективность реакции была близка к 100% – вектор гидролизован практически полностью (Рисунок 26Б). Также происходит и гидролиз комплексами SpCas9:sgPHK_CR7 линейаризованного вектора pGEM-T Easy, содержащего соответствующую последовательность гена *Oxt*, однако эффективность при этом заметно ниже (Рисунок 26Б). Этот факт согласуется с приведенными выше литературными данными о том, что эффективность внесения разрывов SpCas9 в протоспейсеры с неканоничным PAM –NAG существенно ниже таковой для каноничных сайтов.

Таким образом, факт делеции в клоне RNFF1_t1.4 и факт низкой эффективности гомологичной рекомбинации объясняется способностью созданной системы CRISPR/Cas9 вносить ДЦР в последовательность ДНК дикого типа: в рекомбинировавший аллель и в донорный вектор (Рисунок 26Б). Мы не наблюдали следов нецелевых разрывов в паралогичном гене *Oxt* эмбриональных фибробластов Brattleboro (Рисунок 14, таблица 12), однако показали способность комплексов SpCas9:sgPHK_CR7 в реакции *in vitro* вносить разрывы в последовательность этого гена, содержащую протоспейсер с неканоничным PAM –NAG (Рисунок 26Б,В). Это противоречие может быть объяснено следующими факторами: 1) сниженной эффективностью внесения ДЦР в сайты с неканоничным PAM, 2) в ряде работ показано, что состояние хроматина может оказывать влияние на эффективность CRISPR/Cas9 – неактивный хроматин препятствует внесению ДЦР комплексами sgPHK:Cas9 (Daer et al., 2017, Wu et al., 2014). Согласно данным о транскрипции гена *Oxt* (<https://www.ncbi.nlm.nih.gov/gene/25504>), его экспрессия наблюдается только в клетках головного мозга и половых органов крыс, соответственно, в эмбриональных фибробластах крыс он репрессирован (предположительно, закрытый хроматин). Мы полагаем, что совокупное действие приведенных выше факторов снижает эффективность ДЦР в гене *Oxt* эмбриональных фибробластов Brattleboro, как минимум, ниже достоверно детектируемого нами уровня. В случае же клона RNFF1_t1.4 можно предположить, что, так как в результате гомологичной рекомбинации в локус попадает трансген *NeoR*, который экспрессируется под управлением конститутивного промотора, состояние хроматина на данном участке

становится открытым (более «доступен»), в следствие чего в локусе произошел повторный разрыв в протоспейсере с PAM –NAG.

Что касается низкой эффективности гомологичной рекомбинации, то она может быть объяснена внесением разрывов в донорный плазмидный вектор, который является более «доступным» (относительно закрытого хроматина в локусе *Oxt*). В проведенной нами реакции *in vitro* эффективность гидролиза комплексами sgPНК:Cas9 донорного вектора pAvp-target8 была невелика по сравнению с положительным контролем (Рисунок 26Б), однако, стоит отметить, что реакция проводилась в течение 60 минут и соотношение вектор:sgPНК:Cas9 составляло 1:100:100, в то время как в клетках после трансфекции, согласно данным литературы, экспрессия элементов системы CRISPR/Cas9 детектируется в течение минимум 72 часов (Kim et al., 2014), а также соотношение комплексов к вектору может быть на порядки выше. Линеаризованный в результате разрыва в протоспейсере CR-7 с неканоничным PAM плазмидный вектор pAvp-target8 не может выступать в качестве донорной молекулы для гомологичной рекомбинации. Таким образом, клетки в которых не произошла гомологичная рекомбинация из-за недостатка донорных молекул или клетки, в которых рекомбинация произошла, но в результате повторного разрыва (при крупной делеции) был нарушен промотор гена *NeoR*, гибнут в результате позитивно-негативной селекции. В таких условиях могут выжить и образовать колонии клетки, в которых, если прошла рекомбинация, повторный разрыв не нарушил экспрессии *NeoR* (клон RNFF1_t1.4), либо клетки, в которых в рекомбинировавшем аллеле не успел произойти повторный ДЦР в течение времени присутствия в клетке элементов системы CRISPR/Cas9.

ЗАКЛЮЧЕНИЕ

В данной работе мы использовали систему CRISPR/Cas для внесения направленных изменений в ген аргинин-вазопрессина крыс линии Brattleboro. Целевой участок гена отличается всего на один нуклеотид от соответствующего участка гена окситоцина – делецией гуанина во втором экзоне гена гормона аргинин-вазопрессина, которая приводит к развитию гипоталамического несахарного диабета у крыс этой линии. Тем не менее с помощью тщательного выбора сайтов для действия CRISPR/Cas9 нам удалось обеспечить узнавание целевого гена и избежать нецелевых эффектов в гене окситоцина. Для этого нами в качестве отличия между паралогичными генами была использована сама мутация.

Разработанные нами стратегии выбора спейсеров CRISPR/Cas9 позволяют направленно вносить модификации в целевые мутантные копии генов, не затрагивая при этом копии дикого типа. Такой подход может быть использован при исправлении мутаций или, наоборот, внесении мутаций (для моделирования заболеваний *in vitro*) в геном гетерозигот по целевому гену, либо в случае, когда целевой ген входит в семейство генов с близкими последовательностями. Также другой точкой приложения этих стратегий является направленный нокаут целевых генов, например, для изучений функций отдельных генов, входящих в кластеры генов.

Нами была создана система для гомологичной рекомбинации в гене *Avp* крыс линии Brattleboro, состоящая из векторов, кодирующих элементы системы CRISPR/Cas, и донорного плазмидного вектора. В результате позитивно-негативной селекции нами были получены две клональные линии, одна из которых содержит делецию в рекомбинировавшем аллеле *Avp*. В геноме линия RNFF1_t2.2 детектировано две различных модификации, которые являются результатами двух различных путей репарации ДЦР – гомологичной рекомбинации и c-NHEJ. В обоих аллелях данного клона при этом восстановлена рамка считывания (в одном случае прецизно, в другом – с заменой двух аминокислот). Особенность репарации ДЦР в протоспейсере CR-7, которую мы выявили – репарация происходит, видимо, преимущественно по одному и тому же пути в большинстве случаев, может быть использована при проведении экспериментов по исправлению мутации *in vivo*. Если замена двух аминокислот, возникающая при восстановлении рамки считывания за

счет инсерции А, не влияет на функциональность белка-переносчика аргинин-вазопрессина, либо, например, не полностью нарушает её, то при экспериментах *in vivo* (например, при доставке вирусных векторов, кодирующих CRISPR/Cas9 со спейсером CR-7) в данном конкретном случае возможно отказаться от использования донорного вектора и от гомологичной рекомбинации в целом, как целевого пути внесения модификаций. Это позволит эффективно восстанавливать рамку считывания в вазопрессинэргических нейронах – по данным, полученным на фибробластах RNFF1 в клетках, в которые попала CRISPR/Cas9, в более чем 50% случаев происходит инсерция А – что потенциально может восстановить секрецию аргинин-вазопрессина. В дальнейшей работе, несомненно, это предположение будет принято во внимание и детально изучено.

При использовании созданной нами системы для исправления мутации в гене *Avr* крыс Brattleboro мы наблюдали низкую частоту прохождения гомологичной рекомбинации. Было выяснено, что этот факт, а также факт делеции в рекомбинировавшем аллеле клона RNFF1_t1.4, может быть объяснен способностью SpCas9 вносить разрывы в протоспейсеры с неканоничными PAM, что не было учтено при дизайне эксперимента. Стоит отметить, что данный случай, например, обосновывает актуальность дальнейшего исследования механизмов действия системы CRISPR/Cas9 с целью усовершенствования её эффективности и специфичности. Например, уже получены мутантные варианты SpCas9 со сниженной способностью вносить ДЦР в сайты с неканоничными PAM (Kleinstiver et al., 2015).

Полученные нами эмбриональные фибробласты с исправленной мутацией в дальнейших работах будут репрограммированы к плюрипотентному состоянию – будут получены ИПСК крыс Brattleboro с исправленной мутацией. Эти клетки могут быть использованы для разработки клеточной терапии наследственных заболеваний, для исследований возможности восстановления нормального фенотипа в результате исправления мутации на генетическом уровне, в частности, фенотипа релевантных типов клеток для заболевания.

ВЫВОДЫ

1. Создана система для проведения гомологичной рекомбинации в локусе гена аргинин-вазопрессина, состоящая из набора генетических конструкций, экспрессирующих элементы системы CRISPR/Cas9, и донорной генетической конструкции.
2. Разработана стратегия, позволяющая эффективно вносить двуцепочечные разрывы во второй экзон гена аргинин-вазопрессина, не оказывая нецелевого эффекта в паралогичном гене окситоцина.
3. Репарация двуцепочечных разрывов в отсутствие донорной молекулой в целевом участке гена аргинин-вазопрессина преимущественно происходит с инсерцией аденина в точке разрыва, в результате чего происходит восстановление рамки считывания с заменой двух аминокислотных остатков.
4. В результате позитивно-негативной селекции фибробластов крыс линии Brattleboro, трансфицированных смесью донорных плазмид и CRISPR/Cas9, была получена клональная линия RNFF1_t2.2, в одном аллеле гена аргинин-вазопрессина которой произошла гомологичная рекомбинация, а во втором аллеле была восстановлена рамка считывания в результате инсерции аденина.
5. Причиной сниженной эффективности процесса гомологичной рекомбинации является способность созданной системы CRISPR/Cas9 вносить разрывы в донорный вектор из-за присутствия в нем неканонического PAM SpCas9 –NAG.

СПИСОК ИСПОЛЬЗОВАННЫХ СОКРАЩЕНИЙ

ДЦР – двухцепочечный разрыв

ИПСК – индуцированные плюрипотентные стволовые клетки

п.н. – пара нуклеотидов

т.п.н. – тысяча пар нуклеотидов

ЭСК – эмбриональные стволовые клетки

Cas – CRISPR-associated

CRISPR– clustered regularly interspaced short palindromic repeats

crPHK – CRISPR-PHK

FBS – фетальная бычья сыворотка (fetal bovine serum)

c-NHEJ– классическое соединение негомологичных концов

HR – гомологичная рекомбинация

NUC – нуклеазная доля Cas9

PAM – protospacer adjacent motif; мотив, прилежащий к протоспейсеру

PBS – фосфатно-солевой буфер (phosphate-buffered saline)

REC – узнающая доля Cas9

sgPHK – single-guide PHK, единая направляющая PHK

tracrPHK – трансктивирующая CRISPR-PHK

СПИСОК ЛИТЕРАТУРЫ

1. Методы генетической инженерии. Молекулярное клонирование: Пер. с англ./ Маниатис Т., Фрич Э., Сэмбрук Дж. – М.: Мир, 1984. – 480 с.
2. Хегай И.И. Особенности экспрессии гена *di* (diabetes insipidus) в гомологичных инбредных линиях крыс // Генетика. – 2002. – Т. 38. – С. 1677-1681.
3. Хегай И.И. Фенотипическое проявление мутантного гена *diabetes insipidus* у крыс и критерии генотипирования по фенотипу // Генетика. – 2003. – Т. 39. – С. 57-60.
4. Amitai G., Sorek R. CRISPR-Cas adaptation: insights into the mechanism of action // *Nat Rev Microbiol.* – 2016. – V. 14. – № 2. – P. 67-76.
5. Anders C., Niewoehner O., Duerst A., Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease // *Nature.* – 2014. – V. 513. – № 7519. – P. 569-573.
6. Arslan Z., Hermanns V., Wurm R., Wagner R., Pul U. Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system // *Nucleic Acids Res.* – 2014. – V. 42. – № 12. – P. 7884-7893.
7. Bae S., Park J., Kim J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases // *Bioinformatics.* – 2014. – V. 30. – № 10. – P. 1473-1475.
8. Barrangou R., Fremaux C., Deveau H., Richards M., Boyaval P., Moineau S., Romero D. A., Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes // *Science.* – 2007. – V. 315. – № 5819. – P. 1709-1712.
9. Bassuk A. G., Zheng A., Li Y., Tsang S. H., Mahajan V. B. Precision Medicine: Genetic Repair of Retinitis Pigmentosa in Patient-Derived Stem Cells // *Sci Rep.* – 2016. – V. 6. – P. 19969.
10. Beumer K. J., Trautman J. K., Mukherjee K., Carroll D. Donor DNA Utilization during Gene Targeting with Zinc-finger Nucleases // *G3 (Bethesda).* – 2013.10.1534/g3.112.005439.
11. Bolotin A., Quinquis B., Sorokin A., Ehrlich S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin // *Microbiology.* – 2005. – V. 151. – № Pt 8. – P. 2551-2561.
12. Brinkman E. K., Chen T., Amendola M., van Steensel B. Easy quantitative

assessment of genome editing by sequence trace decomposition // *Nucleic Acids Res.* – 2014. – V. 42. – № 22. – P. e168.

13. Carr A. J., Vugler A. A., Hikita S. T., Lawrence J. M., Gias C., Chen L. L., Buchholz D. E., Ahmado A., Semo M., Smart M. J., Hasan S., da Cruz L., Johnson L. V., Clegg D. O., Coffey P. J. Protective effects of human iPS-derived retinal pigment epithelium cell transplantation in the retinal dystrophic rat // *PLoS ONE.* – 2009. – V. 4. – № 12. – P. e8152.

14. Ceccaldi R., Liu J. C., Amunugama R., Hajdu I., Primack B., Petalcorin M. I., O'Connor K. W., Konstantinopoulos P. A., Elledge S. J., Boulton S. J., Yusufzai T., D'Andrea A. D. Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair // *Nature.* – 2015. – V. 518. – № 7538. – P. 258-262.

15. Ceccaldi R., Rondinelli B., D'Andrea A. D. Repair Pathway Choices and Consequences at the Double-Strand Break // *Trends Cell Biol.* – 2016. – V. 26. – № 1. – P. 52-64.

16. Chang C. W., Lai Y. S., Westin E., Khodadadi-Jamayran A., Pawlik K. M., Lamb L. S., Jr., Goldman F. D., Townes T. M. Modeling Human Severe Combined Immunodeficiency and Correction by CRISPR/Cas9-Enhanced Gene Targeting // *Cell Rep.* – 2015. – V. 12. – № 10. – P. 1668-1677.

17. Chen J. R., Tang Z. H., Zheng J., Shi H. S., Ding J., Qian X. D., Zhang C., Chen J. L., Wang C. C., Li L., Chen J. Z., Yin S. K., Shao J. Z., Huang T. S., Chen P., Guan M. X., Wang J. F. Effects of genetic correction on the differentiation of hair cell-like cells from iPSCs with MYO15A mutation // *Cell Death Differ.* – 2016.10.1038/cdd.2016.16.

18. Christensen J. H., Rittig S. Familial neurohypophyseal diabetes insipidus--an update // *Semin Nephrol.* – 2006. – V. 26. – № 3. – P. 209-223.

19. Chylinski K., Makarova K. S., Charpentier E., Koonin E. V. Classification and evolution of type II CRISPR-Cas systems // *Nucleic Acids Res.* – 2014. – V. 42. – № 10. – P. 6091-6105.

20. Cong L., Ran F. A., Cox D., Lin S., Barretto R., Habib N., Hsu P. D., Wu X., Jiang W., Marraffini L. A., Zhang F. Multiplex genome engineering using CRISPR/Cas systems // *Science.* – 2013. – V. 339. – № 6121. – P. 819-823.

21. Cox D. B., Platt R. J., Zhang F. Therapeutic genome editing: prospects and challenges // *Nat Med.* – 2015. – V. 21. – № 2. – P. 121-131.

22. Cradick T. J., Qiu P., Lee C. M., Fine E. J., Bao G. COSMID: A Web-based Tool for Identifying and Validating CRISPR/Cas Off-target Sites // *Mol Ther Nucleic Acids*. – 2014. – V. 3. – P. e214.
23. Daer R. M., Cutts J. P., Brafman D. A., Haynes K. A. The Impact of Chromatin Dynamics on Cas9-Mediated Genome Editing in Human Cells // *ACS Synth Biol*. – 2017. – V. 6. – № 3. – P. 428-438.
24. Deltcheva E., Chylinski K., Sharma C. M., Gonzales K., Chao Y., Pirzada Z. A., Eckert M. R., Vogel J., Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III // *Nature*. – 2011. – V. 471. – № 7340. – P. 602-607.
25. Difilippantonio M. J., Zhu J., Chen H. T., Meffre E., Nussenzweig M. C., Max E. E., Ried T., Nussenzweig A. DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation // *Nature*. – 2000. – V. 404. – № 6777. – P. 510-514.
26. Dillingham M. S., Kowalczykowski S. C. RecBCD enzyme and the repair of double-stranded DNA breaks // *Microbiol Mol Biol Rev*. – 2008. – V. 72. – № 4. – P. 642-671, Table of Contents.
27. Doench J. G., Hartenian E., Graham D. B., Tothova Z., Hegde M., Smith I., Sullender M., Ebert B. L., Xavier R. J., Root D. E. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation // *Nat Biotechnol*. – 2014. – V. 32. – № 12. – P. 1262-1267.
28. Firth A. L., Menon T., Parker G. S., Qualls S. J., Lewis B. M., Ke E., Dargitz C. T., Wright R., Khanna A., Gage F. H., Verma I. M. Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs // *Cell Rep*. – 2015. – V. 12. – № 9. – P. 1385-1390.
29. Flynn R., Grundmann A., Renz P., Hanseler W., James W. S., Cowley S. A., Moore M. D. CRISPR-mediated genotypic and phenotypic correction of a chronic granulomatous disease mutation in human iPS cells // *Exp Hematol*. – 2015. – V. 43. – № 10. – P. 838-848 e833.
30. Fu Y., Foden J. A., Khayter C., Maeder M. L., Reyon D., Joung J. K., Sander J. D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells // *Nat Biotechnol*. – 2013. – V. 31. – № 9. – P. 822-826.
31. Fusaki N., Ban H., Nishiyama A., Saeki K., Hasegawa M. Efficient induction of

transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome // *Proc Jpn Acad Ser B Phys Biol Sci.* – 2009. – V. 85. – № 8. – P. 348-362.

32. Gabreels B. A., Verwer R. W., Sonnemans M. A., Sluiter A. A., Ang C. W., van Leeuwen F. W. Lack of translation of normal 7B2 mRNA levels in hypothalamic mutant vasopressin cells of the homozygous Brattleboro rat // *Neurosci Lett.* – 1997. – V. 239. – № 1. – P. 5-8.

33. Garber K. RIKEN suspends first clinical trial involving induced pluripotent stem cells // *Nat Biotechnol.* – 2015. – V. 33. – № 9. – P. 890-891.

34. Genovese P., Schirotti G., Escobar G., Di Tomaso T., Firrito C., Calabria A., Moi D., Mazzieri R., Bonini C., Holmes M. C., Gregory P. D., van der Burg M., Gentner B., Montini E., Lombardo A., Naldini L. Targeted genome editing in human repopulating haematopoietic stem cells // *Nature.* – 2014. – V. 510. – № 7504. – P. 235-240.

35. Goodarzi A. A., Jeggo P. A. The repair and signaling responses to DNA double-strand breaks // *Adv Genet.* – 2013. – V. 82. – P. 1-45.

36. Gwee P. C., Amemiya C. T., Brenner S., Venkatesh B. Sequence and organization of coelacanth neurohypophysial hormone genes: evolutionary history of the vertebrate neurohypophysial hormone gene locus // *BMC Evol Biol.* – 2008. – V. 8. – P. 93.

37. Haft D. H., Selengut J., Mongodin E. F., Nelson K. E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes // *PLoS Comput Biol.* – 2005. – V. 1. – № 6. – P. e60.

38. Hanna J., Wernig M., Markoulaki S., Sun C. W., Meissner A., Cassady J. P., Beard C., Brambrink T., Wu L. C., Townes T. M., Jaenisch R. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin // *Science.* – 2007. – V. 318. – № 5858. – P. 1920-1923.

39. Heigwer F., Kerr G., Boutros M. E-CRISP: fast CRISPR target site identification // *Nat Methods.* – 2014. – V. 11. – № 2. – P. 122-123.

40. Heler R., Samai P., Modell J. W., Weiner C., Goldberg G. W., Bikard D., Marraffini L. A. Cas9 specifies functional viral targets during CRISPR-Cas adaptation // *Nature.* – 2015. – V. 519. – № 7542. – P. 199-202.

41. 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., Cradick T. J., Marraffini L. A., Bao G., Zhang F. DNA

targeting specificity of RNA-guided Cas9 nucleases // *Nat Biotechnol.* – 2013. – V. 31. – № 9. – P. 827-832.

42. Ishida K., Gee P., Hotta A. Minimizing off-Target Mutagenesis Risks Caused by Programmable Nucleases // *Int J Mol Sci.* – 2015. – V. 16. – № 10. – P. 24751-24771.

43. Jinek M., Jiang F., Taylor D. W., Sternberg S. H., Kaya E., Ma E., Anders C., Hauer M., Zhou K., Lin S., Kaplan M., Iavarone A. T., Charpentier E., Nogales E., Doudna J. A. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation // *Science.* – 2014. – V. 343. – № 6176. – P. 1247997.

44. Jonsson T., Atwal J. K., Steinberg S., Snaedal J., Jonsson P. V., Bjornsson S., Stefansson H., Sulem P., Gudbjartsson D., Maloney J., Hoyte K., Gustafson A., Liu Y., Lu Y., Bhangale T., Graham R. R., Huttenlocher J., Bjornsdottir G., Andreassen O. A., Jonsson E. G., Palotie A., Behrens T. W., Magnusson O. T., Kong A., Thorsteinsdottir U., Watts R. J., Stefansson K. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline // *Nature.* – 2012. – V. 488. – № 7409. – P. 96-99.

45. Kamao H., Mandai M., Okamoto S., Sakai N., Suga A., Sugita S., Kiryu J., Takahashi M. Characterization of human induced pluripotent stem cell-derived retinal pigment epithelium cell sheets aiming for clinical application // *Stem Cell Reports.* – 2014. – V. 2. – № 2. – P. 205-218.

46. Karanam K., Kafri R., Loewer A., Lahav G. Quantitative live cell imaging reveals a gradual shift between DNA repair mechanisms and a maximal use of HR in mid S phase // *Mol Cell.* – 2012. – V. 47. – № 2. – P. 320-329.

47. Karvelis T., Gasiunas G., Miksys A., Barrangou R., Horvath P., Siksnys V. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus* // *RNA Biol.* – 2013. – V. 10. – № 5. – P. 841-851.

48. Khegai, II, Gulyaeva M. A., Popova N. A., Zakharova L. A., Ivanova L. N. Immune system in vasopressin-deficient rats during ontogeny // *Bull Exp Biol Med.* – 2003. – V. 136. – № 5. – P. 448-450.

49. Khegay, II, Popova N. A., Ivanova L. N. Reduced Walker 256 carcinosarcoma growth in vasopressin-deficient Brattleboro rats // *Tumour Biol.* – 2010. – V. 31. – № 6. – P. 569-573.

50. Kim S., Kim D., Cho S. W., Kim J., Kim J. S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins // *Genome Res.* –

2014. – V. 24. – № 6. – P. 1012-1019.

51. Kleinstiver B. P., Prew M. S., Tsai S. Q., Topkar V. V., Nguyen N. T., Zheng Z. L., Gonzales A. P. W., Li Z. Y., Peterson R. T., Yeh J. R. J., Aryee M. J., Joung J. K. Engineered CRISPR-Cas9 nucleases with altered PAM specificities // *Nature*. – 2015. – V. 523. – № 7561. – P. 481-489.

52. Kleinstiver B. P., Pattanayak V., Prew M. S., Tsai S. Q., Nguyen N. T., Zheng Z., Joung J. K. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects // *Nature*. – 2016. – V. 529. – № 7587. – P. 490-495.

53. Lei Y., Lu L., Liu H. Y., Li S., Xing F., Chen L. L. CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants // *Mol Plant*. – 2014. – V. 7. – № 9. – P. 1494-1496.

54. Levy A., Goren M. G., Yosef I., Auster O., Manor M., Amitai G., Edgar R., Qimron U., Sorek R. CRISPR adaptation biases explain preference for acquisition of foreign DNA // *Nature*. – 2015. – V. 520. – № 7548. – P. 505-510.

55. Li H. L., Fujimoto N., Sasakawa N., Shirai S., Ohkame T., Sakuma T., Tanaka M., Amano N., Watanabe A., Sakurai H., Yamamoto T., Yamanaka S., Hotta A. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9 // *Stem Cell Reports*. – 2015. – V. 4. – № 1. – P. 143-154.

56. Lieber M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway // *Annu Rev Biochem*. – 2010. – V. 79. – P. 181-211.

57. Lin Y., Cradick T. J., Brown M. T., Deshmukh H., Ranjan P., Sarode N., Wile B. M., Vertino P. M., Stewart F. J., Bao G. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences // *Nucleic Acids Res*. – 2014. – V. 42. – № 11. – P. 7473-7485.

58. Liu H., Wei Z., Dominguez A., Li Y., Wang X., Qi L. S. CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation // *Bioinformatics*. – 2015. – V. 31. – № 22. – P. 3676-3678.

59. Liu R., Paxton W. A., Choe S., Ceradini D., Martin S. R., Horuk R., MacDonald M. E., Stuhlmann H., Koup R. A., Landau N. R. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection // *Cell*. –

1996. – V. 86. – № 3. – P. 367-377.

60. Long C., Amoasii L., Mireault A. A., McAnally J. R., Li H., Sanchez-Ortiz E., Bhattacharyya S., Shelton J. M., Bassel-Duby R., Olson E. N. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy // *Science*. – 2016. – V. 351. – № 6271. – P. 400-403.

61. Ma M., Ye A. Y., Zheng W., Kong L. A guide RNA sequence design platform for the CRISPR/Cas9 system for model organism genomes // *Biomed Res Int*. – 2013. – V. 2013. – P. 270805.

62. MacPherson C. R., Scherf A. Flexible guide-RNA design for CRISPR applications using Protospacer Workbench // *Nat Biotechnol*. – 2015. – V. 33. – № 8. – P. 805-806.

63. Maddalo D., Manchado E., Concepcion C. P., Bonetti C., Vidigal J. A., Han Y. C., Ogirodowski P., Crippa A., Rekhtman N., de Stanchina E., Lowe S. W., Ventura A. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system // *Nature*. – 2014. – V. 516. – № 7531. – P. 423-427.

64. 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., van der Oost J., Koonin E. V. Evolution and classification of the CRISPR-Cas systems // *Nat Rev Microbiol*. – 2011. – V. 9. – № 6. – P. 467-477.

65. Makarova K. S., Wolf Y. I., Alkhnbashi O. S., Costa F., Shah S. A., Saunders S. J., Barrangou R., Brouns S. J., Charpentier E., Haft D. H., Horvath P., Moineau S., Mojica F. J., Terns R. M., Terns M. P., White M. F., Yakunin A. F., Garrett R. A., van der Oost J., Backofen R., Koonin E. V. An updated evolutionary classification of CRISPR-Cas systems // *Nat Rev Microbiol*. – 2015. – V. 13. – № 11. – P. 722-736.

66. Mali P., Yang L., Esvelt K. M., Aach J., Guell M., DiCarlo J. E., Norville J. E., Church G. M. RNA-guided human genome engineering via Cas9 // *Science*. – 2013. – V. 339. – № 6121. – P. 823-826.

67. Marraffini L. A. CRISPR-Cas immunity in prokaryotes // *Nature*. – 2015. – V. 526. – № 7571. – P. 55-61.

68. Mateos-Gomez P. A., Gong F., Nair N., Miller K. M., Lazzerini-Denchi E., Sfeir A. Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination // *Nature*. – 2015. – V. 518. – № 7538. – P. 254-257.

69. McVey M., Lee S. E. MMEJ repair of double-strand breaks (director's cut): deleted

- sequences and alternative endings // *Trends Genet.* – 2008. – V. 24. – № 11. – P. 529-538.
70. Mohr E., Schmitz E., Richter D. A single rat genomic DNA fragment encodes both the oxytocin and vasopressin genes separated by 11 kilobases and oriented in opposite transcriptional directions // *Biochimie.* – 1988. – V. 70. – № 5. – P. 649-654.
71. Montague T. G., Cruz J. M., Gagnon J. A., Church G. M., Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing // *Nucleic Acids Res.* – 2014. – V. 42. – № Web Server issue. – P. W401-407.
72. Moore R., Spinhirne A., Lai M. J., Preisser S., Li Y., Kang T., Bleris L. CRISPR-based self-cleaving mechanism for controllable gene delivery in human cells // *Nucleic Acids Res.* – 2015. – V. 43. – № 2. – P. 1297-1303.
73. Naito Y., Hino K., Bono H., Ui-Tei K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites // *Bioinformatics.* – 2015. – V. 31. – № 7. – P. 1120-1123.
74. Nelson C. E., Hakim C. H., Ousterout D. G., Thakore P. I., Moreb E. A., Castellanos Rivera R. M., Madhavan S., Pan X., Ran F. A., Yan W. X., Asokan A., Zhang F., Duan D., Gersbach C. A. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy // *Science.* – 2016. – V. 351. – № 6271. – P. 403-407.
75. Nemudryi A. A., Valetdinova K. R., Medvedev S. P., Zakian S. M. TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery // *Acta Naturae.* – 2014. – V. 6. – № 3. – P. 19-40.
76. Nishimasu H., Ran F. A., Hsu P. D., Konermann S., Shehata S. I., Dohmae N., Ishitani R., Zhang F., Nureki O. Crystal structure of Cas9 in complex with guide RNA and target DNA // *Cell.* – 2014. – V. 156. – № 5. – P. 935-949.
77. Nunez J. K., Kranzusch P. J., Noeske J., Wright A. V., Davies C. W., Doudna J. A. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity // *Nat Struct Mol Biol.* – 2014. – V. 21. – № 6. – P. 528-534.
78. O'Brien A., Bailey T. L. GT-Scan: identifying unique genomic targets // *Bioinformatics.* – 2014. – V. 30. – № 18. – P. 2673-2675.
79. Okita K., Matsumura Y., Sato Y., Okada A., Morizane A., Okamoto S., Hong H., Nakagawa M., Tanabe K., Tezuka K., Shibata T., Kunisada T., Takahashi M., Takahashi J., Saji H., Yamanaka S. A more efficient method to generate integration-free human iPS cells // *Nat Methods.* – 2011. – V. 8. – № 5. – P. 409-412.

80. Okita K., Yamakawa T., Matsumura Y., Sato Y., Amano N., Watanabe A., Goshima N., Yamanaka S. An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells // *Stem Cells*. – 2013. – V. 31. – № 3. – P. 458-466.
81. Park C. Y., Kim D. H., Son J. S., Sung J. J., Lee J., Bae S., Kim J. H., Kim D. W., Kim J. S. Functional Correction of Large Factor VIII Gene Chromosomal Inversions in Hemophilia A Patient-Derived iPSCs Using CRISPR-Cas9 // *Cell Stem Cell*. – 2015. – V. 17. – № 2. – P. 213-220.
82. Phadke T., Bothmer A., Lee C., Abdulkerim H., Barrera L., Moss S., Jayaram H., Cotta-Ramusino C. DNA Ends Matter: The Impact of Using CRISPR/Cas9 Variants on DNA Repair Pathway Choices and Editing Profiles at the HBB Locus // *Molecular Therapy*. – 2016. – V. 24. – P. S230-S230.
83. Prakash V., Moore M., Yanez-Munoz R. J. Current Progress in Therapeutic Gene Editing for Monogenic Diseases // *Mol Ther*. – 2016. – V. 24. – № 3. – P. 465-474.
84. Ran F. A., Hsu P. D., Lin C. Y., Gootenberg J. S., Konermann S., Trevino A. E., Scott D. A., Inoue A., Matoba S., Zhang Y., Zhang F. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity // *Cell*. – 2013. – V. 154. – № 6. – P. 1380-1389.
85. Ran F. A., Cong L., Yan W. X., Scott D. A., Gootenberg J. S., Kriz A. J., Zetsche B., Shalem O., Wu X., Makarova K. S., Koonin E. V., Sharp P. A., Zhang F. In vivo genome editing using *Staphylococcus aureus* Cas9 // *Nature*. – 2015. – V. 520. – № 7546. – P. 186-191.
86. Reyon D., Tsai S. Q., Khayter C., Foden J. A., Sander J. D., Joung J. K. FLASH assembly of TALENs for high-throughput genome editing // *Nat Biotechnol*. – 2012. – V. 30. – № 5. – P. 460-465.
87. Robert T., Vanoli F., Chiolo I., Shubassi G., Bernstein K. A., Rothstein R., Botrugno O. A., Parazzoli D., Oldani A., Minucci S., Foiani M. HDACs link the DNA damage response, processing of double-strand breaks and autophagy // *Nature*. – 2011. – V. 471. – № 7336. – P. 74-79.
88. Sander J. D., Zaback P., Joung J. K., Voytas D. F., Dobbs D. Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool // *Nucleic Acids Res*. – 2007. – V. 35. – № Web Server issue. – P. W599-605.

89. Sander J. D., Maeder M. L., Reyon D., Voytas D. F., Joung J. K., Dobbs D. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool // *Nucleic Acids Res.* – 2010. – V. 38. – № Web Server issue. – P. W462-468.
90. Schmale H., Richter D. Single base deletion in the vasopressin gene is the cause of diabetes insipidus in Brattleboro rats // *Nature.* – 1984. – V. 308. – № 5961. – P. 705-709.
91. Schmale H., Fehr S., Richter D. Vasopressin biosynthesis--from gene to peptide hormone // *Kidney Int Suppl.* – 1987. – V. 21. – P. S8-13.
92. Schmitz E., Mohr E., Richter D. Rat vasopressin and oxytocin genes are linked by a long interspersed repeated DNA element (LINE): sequence and transcriptional analysis of LINE // *DNA Cell Biol.* – 1991. – V. 10. – № 2. – P. 81-91.
93. Schwartz S. D., Regillo C. D., Lam B. L., Elliott D., Rosenfeld P. J., Gregori N. Z., Hubschman J. P., Davis J. L., Heilwell G., Spirn M., Maguire J., Gay R., Bateman J., Ostrick R. M., Morris D., Vincent M., Anglade E., Del Priore L. V., Lanza R. Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies // *Lancet.* – 2015. – V. 385. – № 9967. – P. 509-516.
94. Slaymaker I. M., Gao L., Zetsche B., Scott D. A., Yan W. X., Zhang F. Rationally engineered Cas9 nucleases with improved specificity // *Science.* – 2016. – V. 351. – № 6268. – P. 84-88.
95. Stern A., Keren L., Wurtzel O., Amitai G., Sorek R. Self-targeting by CRISPR: gene regulation or autoimmunity? // *Trends Genet.* – 2010. – V. 26. – № 8. – P. 335-340.
96. Sternberg S. H., Redding S., Jinek M., Greene E. C., Doudna J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9 // *Nature.* – 2014. – V. 507. – № 7490. – P. 62-+.
97. Tabebordbar M., Zhu K., Cheng J. K., Chew W. L., Widrick J. J., Yan W. X., Maesner C., Wu E. Y., Xiao R., Ran F. A., Cong L., Zhang F., Vandenberghe L. H., Church G. M., Wagers A. J. In vivo gene editing in dystrophic mouse muscle and muscle stem cells // *Science.* – 2016. – V. 351. – № 6271. – P. 407-411.
98. Takahashi K., Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors // *Cell.* – 2006. – V. 126. – № 4. – P. 663-676.
99. Takahashi K., Tanabe K., Ohnuki M., Narita M., Ichisaka T., Tomoda K.,

Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors // *Cell*. – 2007. – V. 131. – № 5. – P. 861-872.

100. Truong L. N., Li Y., Shi L. Z., Hwang P. Y., He J., Wang H., Razavian N., Berns M. W., Wu X. Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells // *Proc Natl Acad Sci U S A*. – 2013. – V. 110. – № 19. – P. 7720-7725.

101. Valtin H., Schroeder H. A., Benirschke K., Sokol H. W. Familial hypothalamic diabetes insipidus in rats // *Nature*. – 1962. – V. 196. – P. 1109-1110.

102. van Overbeek M., Capurso D., Carter M. M., Thompson M. S., Frias E., Russ C., Reece-Hoyes J. S., Nye C., Gradia S., Vidal B., Zheng J., Hoffman G. R., Fuller C. K., May A. P. DNA Repair Profiling Reveals Nonrandom Outcomes at Cas9-Mediated Breaks // *Mol Cell*. – 2016. – V. 63. – № 4. – P. 633-646.

103. Watada Y., Yamashita D., Toyoda M., Tsuchiya K., Hida N., Tanimoto A., Ogawa K., Kanzaki S., Umezawa A. Magnetic resonance monitoring of superparamagnetic iron oxide (SPIO)-labeled stem cells transplanted into the inner ear // *Neurosci Res*. – 2015. – V. 95. – P. 21-26.

104. Wei Y., Terns R. M., Terns M. P. Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation // *Genes Dev*. – 2015. – V. 29. – № 4. – P. 356-361.

105. Wu X., Scott D. A., Kriz A. J., Chiu A. C., Hsu P. D., Dadon D. B., Cheng A. W., Trevino A. E., Konermann S., Chen S., Jaenisch R., Zhang F., Sharp P. A. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells // *Nat Biotechnol*. – 2014. – V. 32. – № 7. – P. 670-676.

106. Wu Y., Kantake N., Sugiyama T., Kowalczykowski S. C. Rad51 protein controls Rad52-mediated DNA annealing // *J Biol Chem*. – 2008. – V. 283. – № 21. – P. 14883-14892.

107. Xiao A., Cheng Z., Kong L., Zhu Z., Lin S., Gao G., Zhang B. CasOT: a genome-wide Cas9/gRNA off-target searching tool // *Bioinformatics*. – 2014.10.1093/bioinformatics/btt764.

108. Xie F., Ye L., Chang J. C., Beyer A. I., Wang J., Muench M. O., Kan Y. W. Seamless gene correction of beta-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac // *Genome Res*. – 2014. – V. 24. – № 9. – P. 1526-1533.

109. Xie S., Shen B., Zhang C., Huang X., Zhang Y. sgRNACas9: a software package for

designing CRISPR sgRNA and evaluating potential off-target cleavage sites // PLoS ONE. – 2014. – V. 9. – № 6. – P. e100448.

110. Xu P., Tong Y., Liu X. Z., Wang T. T., Cheng L., Wang B. Y., Lv X., Huang Y., Liu D. P. Both TALENs and CRISPR/Cas9 directly target the HBB IVS2-654 (C > T) mutation in beta-thalassemia-derived iPSCs // Sci Rep. – 2015. – V. 5. – P. 12065.

111. Yang H., Wang H., Shivalila C. S., Cheng A. W., Shi L., Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering // Cell. – 2013. – V. 154. – № 6. – P. 1370-1379.

112. 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., Huang P. Y., Daley G., Church G. Optimization of scarless human stem cell genome editing // Nucleic Acids Res. – 2013. – V. 41. – № 19. – P. 9049-9061.

113. Yosef I., Goren M. G., Qimron U. Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli // Nucleic Acids Res. – 2012. – V. 40. – № 12. – P. 5569-5576.

114. Yu J., Hu K., Smuga-Otto K., Tian S., Stewart R., Slukvin, II, Thomson J. A. Human induced pluripotent stem cells free of vector and transgene sequences // Science. – 2009. – V. 324. – № 5928. – P. 797-801.

115. Yun M. H., Hiom K. CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle // Nature. – 2009. – V. 459. – № 7245. – P. 460-463.

116. Zakharova L. A., Khagai, II, Sharova N. P., Melnikova V. I., Karpova Y. D., Astakhova T. M., Popova N. A., Ivanova L. N. Pattern of MHC class I and immune proteasome expression in Walker 256 tumor during growth and regression in Brattleboro rats with the hereditary defect of arginine-vasopressin synthesis // Cell Immunol. – 2011. – V. 271. – № 2. – P. 385-391.

117. Zhang L., Jia R., Palange N. J., Satheka A. C., Togo J., An Y., Humphrey M., Ban L., Ji Y., Jin H., Feng X., Zheng Y. Large genomic fragment deletions and insertions in mouse using CRISPR/Cas9 // PLoS ONE. – 2015. – V. 10. – № 3. – P. e0120396.

118. Zhang Y., Ge X., Yang F., Zhang L., Zheng J., Tan X., Jin Z. B., Qu J., Gu F. Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells // Sci Rep. – 2014. – V. 4. – P. 5405.

119. Zwaka T. P., Thomson J. A. Homologous recombination in human embryonic stem cells // Nat Biotechnol. – 2003. – V. 21. – № 3. – P. 319-321.