(19) Евразийское патентное ведомство

(21) 201991647 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2020.04.21

(22) Дата подачи заявки
2018.01.05

(51) Int. Cl. *C12N 15/113* (2010.01)
C12N 15/864 (2006.01)
C12N 9/22 (2006.01)

(54) ОПТИМИЗИРОВАННАЯ СТРАТЕГИЯ МОДИФИКАЦИЙ, ВЫЗЫВАЮЩИХ ПРОПУСК ЭКЗОНА, С ПРИМЕНЕНИЕМ CRISPR/Cas9 С ТРОЙНЫМИ НАПРАВЛЯЮЩИМИ ПОСЛЕДОВАТЕЛЬНОСТЯМИ

(31) **62/442,606; 62/544,449; 62/596,298**
(32) **2017.01.05; 2017.08.11; 2017.12.08**
(33) **US**
(86) **PCT/US2018/012558**
(87) **WO 2018/129296 2018.07.12**

(71) Заявитель:
ДЗЕ БОРД ОФ РИДЖЕНТС ОФ ДЗЕ ЮНИВЕРСИТИ ОФ ТЕХАС СИСТЕМ (US)

(72) Изобретатель:
Амоасии Леонела, Олсон Эрик (US)

(74) Представитель:
Медведев В.Н. (RU)

(57) CRISPR/Cas9-опосредованное редактирование генома обладает клиническим потенциалом для лечения генетических заболеваний, таких как миодистрофия Дюшенна (МДД), которая вызвана мутациями в гене дистрофина. В настоящем изобретении, при использовании трех промоторов для направления экспрессии одной и той же DMD направляющей РНК, была достигнута более эффективная и безопасная форма редактирования генома в модели МДД на гуманизированных мышах с делецией в экзоне 50 и на собаке ΔEx50-MD.

201991647 A1

201991647 A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-557740EA/025

ОПТИМИЗИРОВАННАЯ СТРАТЕГИЯ МОДИФИКАЦИЙ, ВЫЗЫВАЮЩИХ ПРОПУСК ЭКЗОНА, С ПРИМЕНЕНИЕМ CRISPR/CAS9 С ТРОЙНЫМИ НАПРАВЛЯЮЩИМИ ПОСЛЕДОВАТЕЛЬНОСТЯМИ

ПРИТЯЗАНИЕ НА ПРИОРИТЕТ

Настоящая заявка испрашивает приоритет предварительной заявки на патент США 62/596,298, поданной 8 декабря 2017 года, предварительной заявки на патент США 62/544,499, поданной 11 августа 2017 года и предварительной заявки на патент США 62/442,606, поданной 5 января 2017 года, все содержание которых настоящим полностью включено посредством отсылки.

ПОЛОЖЕНИЕ О ПОДДЕРЖКЕ ИЗ СРЕДСТВ ФЕДЕРАЛЬНОГО БЮДЖЕТА

Настоящее изобретение было сделано с правительственной поддержкой по гранту U54 HD 087351, выданному Национальными институтами здравоохранения США. Правительство обладает некоторыми правами на изобретение.

СПИСОК ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Настоящая заявка содержит Список последовательностей, который был представлен в электронном виде в формате ASCII и настоящим полностью включен посредством отсылки. Указанная копия формата ASCII, созданная 5 января 2018 года, называется UTFD_P3178WO.txt и имеет размер 1316974 байтов.

ОБЛАСТЬ ТЕХНИКИ, К КОТОРОЙ ОТНОСИТСЯ ИЗОБРЕТЕНИЕ

Настоящее описание относится к областям молекулярной биологии, медицины и генетики. В частности, описание относится к композициям и их применению для редактирования генома с целью коррекции мутаций *in vivo* при использовании метода пропуска экзонов.

УРОВЕНЬ ТЕХНИКИ

Миодистрофии (MD) – группа больше чем из 30 генетических заболеваний, характеризующихся прогрессирующей слабостью и дегенерацией скелетных мышц, которые контролируют движение. Миодистрофия Дюшена (DMD или МДД) является одной из самых тяжелых форм MD, которая поражает приблизительно одного из 5000 мальчиков и характеризуется прогрессирующей мышечной слабостью и

преждевременной смертью. Кардиомиопатия и сердечная недостаточность являются распространенными, неизлечимыми и летальными признаками DMD. Заболевание вызвано мутациями в гене, кодирующем дистрофин (DMD), большой внутриклеточный белок, который связывает комплекс дистрогликана на поверхности клетки с внутренним цитоскелетом, поддерживая таким образом целостность мембраны мышечной клетки во время сокращения. Мутации в гене дистрофина приводят к потере экспрессии дистрофина, вызывая ломкость мышечной мембраны и прогрессирующее истощение мышц.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

Несмотря на интенсивные попытки поиска средств лечения при использовании различных подходов, в том числе переноса миобластов, вирусной доставки и пропуска экзонов, опосредованного олигонуклеотидами, лечения миодистрофии до сих пор не существует. Авторы настоящего изобретения недавно использовали метод опосредованного регулярно расположенными кластерами короткими палиндромными повторами/Cas9 (CRISPR/Cas9) редактирования генома для коррекции мутации гена дистрофина (DMD) у постнатальных мышей mdx, модели DMD. AAV-опосредованная доставка *in vivo* компонентов редактирования генов успешно удаляла мутантную геномную последовательность путем пропуска экзонов в клетках сердечных и скелетных мышц мышей mdx. Используя различные способы доставки AAV9, авторы изобретения восстановили экспрессию белка дистрофина в сердечной и скелетной мышцах у мышей mdx. Модель на мышах mdx и коррекция экзона 23 с использованием AAV доставки миоредактирующего аппарата были полезны для демонстрации подтверждения концепции пропуска экзонов с использованием нескольких разрезов в геномной области, охватывающей мутацию, *in vivo*. Тем не менее, более эффективные и безопасные подходы к редактированию генома и DMD обеспечат более мощные средства вмешательства при этой болезни.

В некоторых вариантах осуществления нуклеиновая кислота включает последовательность, кодирующую первую DMD направляющую РНК, направленно взаимодействующую с первой геномной последовательностью-мишенью, последовательность, кодирующую вторую DMD направляющую РНК, направленно взаимодействующую со

второй геномной последовательностью-мишенью, последовательность, кодирующую первый промотор, где первый промотор направляет экспрессию последовательности, кодирующей первую DMD направляющую РНК, и последовательность, кодирующую второй промотор, где первый промотор направляет экспрессию последовательности, кодирующей вторую DMD направляющую РНК, где каждая первая геномная последовательность-мишень и вторая геномная последовательность-мишень включают акцепторный сайт сплайсинга дистрофина. В некоторых вариантах осуществления последовательность, кодирующая первый промотор, и последовательность, кодирующая второй промотор, являются идентичными. В некоторых вариантах осуществления последовательность, кодирующая первый промотор, и последовательность, кодирующая второй промотор, не являются идентичными. В некоторых вариантах осуществления последовательность, кодирующая первый промотор, и последовательность, кодирующая второй промотор, обладают по меньшей мере 50%, 60%, 70%, 80%, 90% или 95% идентичностью последовательности. В некоторых вариантах осуществления первая геномная последовательность-мишень и вторая геномная последовательность-мишень являются идентичными. В некоторых вариантах осуществления первая геномная последовательность-мишень и вторая геномная последовательность-мишень не являются идентичными. В некоторых вариантах осуществления первая геномная последовательность-мишень и вторая геномная последовательность-мишень обладают по меньшей мере 50%, 60%, 70%, 80%, 90% или 95% идентичностью последовательности. В некоторых вариантах осуществления первая геномная последовательность-мишень и вторая геномная последовательность-мишень являются комплементарными. В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает последовательность, кодирующую третью DMD направляющую РНК, направленно взаимодействующую с третьей геномной последовательностью-мишенью, и последовательность, кодирующую третий промотор, где третий промотор направляет экспрессию последовательности, кодирующей третью DMD направляющую РНК, и где третья геномная последовательность-

мишень включает акцепторный сайт сплайсинга дистрофина. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, и последовательности, кодирующей третий промотор, являются идентичными. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, и последовательности, кодирующей третий промотор, не являются идентичными. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, и последовательности, кодирующей третий промотор, обладают по меньшей мере 50%, 60%, 70%, 80%, 90% или 95% идентичностью последовательности. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени и третьей геномной последовательности-мишени являются идентичными. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени и третьей геномной последовательности-мишени не являются идентичными. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени и третьей геномной последовательности-мишени обладают по меньшей мере 50%, 60%, 70%, 80%, 90% или 95% идентичностью последовательности. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени и третьей геномной последовательности-мишени являются комплементарными. В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает последовательность, кодирующую четвертую DMD направляющую РНК, направленно взаимодействующую с четвертой геномной последовательностью-мишенью и последовательностью, кодирующей четвертый промотор, где четвертый промотор направляет экспрессию четвертой последовательности, кодирующей DMD направляющую РНК,

где четвертая геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор и последовательности, кодирующей четвертый промотор, являются идентичными. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, и последовательности, кодирующей четвертый промотор, не являются идентичными. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, и последовательности, кодирующей четвертый промотор, обладают по меньшей мере 50%, 60%, 70%, 80%, 90% или 95% идентичностью последовательности. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени и четвертой геномной последовательности-мишени являются идентичными. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени и четвертой геномной последовательности-мишени не являются идентичными. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени и четвертой геномной последовательности-мишени обладают по меньшей мере 50%, 60%, 70%, 80%, 90% или 95% идентичностью последовательности. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени и четвертой геномной последовательности-мишени являются комплементарными. В некоторых

вариантах осуществления нуклеиновая кислота дополнительно включает последовательность, кодирующую пятую DMD направляющую РНК, направленно взаимодействующую с пятой геномной последовательностью-мишенью, и последовательность, кодирующую пятый промотор, где пятый промотор направляет экспрессию последовательности, кодирующей пятую DMD направляющую РНК, где пятая геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, являются идентичными. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, не являются идентичными. В некоторых вариантах осуществления по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, обладают по меньшей мере 50%, 60%, 70%, 80%, 90% или 95% идентичностью последовательности. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени, четвертой геномной последовательности-мишени и пятой геномной последовательности-мишени являются идентичными. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени, четвертой геномной последовательности-мишени и пятой геномной последовательности-

мишени не являются идентичными. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени, четвертой геномной последовательности-мишени и пятой геномной последовательности-мишени обладают по меньшей мере 50%, 60%, 70%, 80%, 90% или 95% идентичностью последовательности. В некоторых вариантах осуществления по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени, четвертой геномной последовательности-мишени и пятой геномной последовательности-мишени являются комплементарными. В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает по меньшей мере одну последовательность, кодирующую дополнительную DMD направляющую РНК, направленно взаимодействующую с геномной последовательностью-мишенью, и по меньшей мере один дополнительный промотор, где дополнительный промотор направляет экспрессию последовательности, кодирующей дополнительную DMD направляющую РНК, где дополнительная геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина. В некоторых вариантах осуществления акцепторный сайт сплайсинга дистрофина включает 5'-акцепторный сайт сплайсинга экзона 51. В некоторых вариантах осуществления последовательность, кодирующая первый промотор, или последовательность, кодирующая второй промотор, включает последовательность, кодирующую конститутивный промотор. В некоторых вариантах осуществления первый промотор или второй промотор включает конститутивный промотор. В некоторых вариантах осуществления по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает последовательность, кодирующую конститутивный промотор. В некоторых вариантах осуществления по меньшей мере один из первого промотора, второго промотора, третьего промотора,

четвертого промотора и пятого промотора включает конститутивный промотор. В некоторых вариантах осуществления последовательность, кодирующая первый промотор, или последовательность, кодирующая второй промотор, включает последовательность, кодирующую индуцируемый промотор. В некоторых вариантах осуществления по меньшей мере один из первого промотора, второго промотора, третьего промотора, четвертого промотора и пятого промотора включает индуцируемый промотор. В некоторых вариантах осуществления последовательность, кодирующая первый промотор, или последовательность, кодирующая второй промотор, включает последовательность, кодирующую специфический для типа клеток промотор. В некоторых вариантах осуществления по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает специфический для типа клеток промотор. В некоторых вариантах осуществления специфический для типа клеток промотор включает мышечно-специфический промотор. В некоторых вариантах осуществления последовательность, кодирующая первый промотор, или последовательность, кодирующая второй промотор, включает последовательность, кодирующую промотор U6, промотор H1 или промотор 7SK. В некоторых вариантах осуществления по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает промотор U6, промотор H1 или промотор 7SK. В некоторых вариантах осуществления по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает промотор

U6. В некоторых вариантах осуществления по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор и последовательности, кодирующей пятый промотор, включает промотор H1. В некоторых вариантах осуществления по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает промотор 7SK. В некоторых вариантах осуществления последовательность, кодирующая первую DMD направляющую РНК, последовательность, кодирующая вторую DMD направляющую РНК, и последовательность, кодирующую третью DMD направляющую РНК, являются идентичными, и 5'-акцепторный сайт сплайсинга включает 5'-акцепторный сайт сплайсинга экзона 51. В некоторых вариантах осуществления последовательность, кодирующая первый промотор, включает последовательность, кодирующую промотор U6, последовательность, кодирующая второй промотор, включает последовательность, кодирующую промотор H1, и последовательность, кодирующая третий промотор, включает промотор 7SK. В некоторых вариантах осуществления нуклеиновая кислота включает последовательность ДНК. В некоторых вариантах осуществления нуклеиновая кислота включает последовательность РНК. В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает одну или более последовательностей, кодирующих инвертированный концевой повтор (ITR). В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает последовательность, кодирующую 5' инвертированный концевой повтор (ITR), и последовательность, кодирующую 3' ITR. В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV). В некоторых вариантах

осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV) серотипа 2 (AAV2). В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), и последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из AAV2. В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV) серотипа 4 (AAV4). В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), и последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из AAV4. В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 145 нуклеотидов. В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 115 нуклеотидов. В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 141 нуклеотида. В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает полиаденозиновую (поли-A) последовательность. В некоторых вариантах осуществления поли-A последовательность является мини-поли-A-последовательностью. В некоторых вариантах осуществления последовательность, кодирующая первую DMD направляющую РНК, или последовательность, кодирующая вторую DMD направляющую РНК, включает последовательность любого из SEQ ID NO. 60-382, 706-708 и 712-789. В некоторых вариантах осуществления последовательность, кодирующая первую DMD направляющую РНК, или последовательность, кодирующая вторую DMD направляющую РНК, включает последовательность SEQ ID NO. 714.

Также предложен вектор, включающий нуклеиновую кислоту, включающую последовательность, кодирующую первую DMD направляющую РНК, направленно взаимодействующую с первой геномной последовательностью-мишенью, последовательность, кодирующую вторую DMD направляющую РНК, направленно взаимодействующую со второй геномной последовательностью-мишенью, последовательность, кодирующую первый промотор, где первый промотор направляет экспрессию последовательности, кодирующей первую DMD направляющую РНК, и последовательность, кодирующую второй промотор, где первый промотор направляет экспрессию последовательности, кодирующей вторую DMD направляющую РНК, где каждая первая геномная последовательность-мишень и вторая геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина. В некоторых вариантах осуществления вектор дополнительно включает последовательность, кодирующую инвертированный концевой повтор транспозируемого элемента. В некоторых вариантах осуществления транспозируемый элемент является транспозоном. В некоторых вариантах осуществления транспозон является транспозоном Tn7. В некоторых вариантах осуществления вектор является невирусным вектором. В некоторых вариантах осуществления невирусный вектор является плазмидой. В некоторых вариантах осуществления вектор является вирусным вектором. В некоторых вариантах осуществления вирусный вектор является вектором на основе аденоассоциированного вируса (AAV). В некоторых вариантах осуществления AAV вектор является репликационно-дефектным или условно репликационно-дефектным. В некоторых вариантах осуществления AAV вектор является рекомбинантным вектором AAV. В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 1 (AAV1), 2 (AAV2), 3 (AAV3), 4 (AAV4), 5 (AAV5), 6 (AAV6), 7 (AAV7), 8 (AAV8), 9 (AAV9), 10 (AAV10), 11 (AAV11) или их любой комбинации. В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 9 (AAV9). В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV вектора

серотипа 2 (AAV2). В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV2, и последовательность, выделенную или полученную из AAV9. В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV4, и последовательность, выделенную или полученную из AAV9. В некоторых вариантах осуществления вектор оптимизирован для экспрессии в клетках млекопитающих. В некоторых вариантах осуществления вектор оптимизирован для экспрессии в клетках человека. В некоторых вариантах осуществления вектор включает последовательность SEQ ID NO. 914, SEQ ID NO. 915, SEQ ID NO. 916 или SEQ ID NO. 917.

Также предложена нуклеиновая кислота, включающая последовательность, кодирующую промотор, и последовательность, кодирующую Cas9 или ее нуклеазный домен, где последовательность, кодирующая промотор, включает последовательность, кодирующую мышечно-специфический промотор. В некоторых вариантах осуществления последовательность, кодирующая мышечно-специфический промотор, включает последовательность, кодирующую промотор CK8. В некоторых вариантах осуществления последовательность, кодирующая мышечно-специфический промотор, включает последовательность, кодирующую промотор CK8e. В некоторых вариантах осуществления последовательность, кодирующая Cas9 или ее нуклеазный домен, выделена или получена из последовательности, кодирующей Cas9 *S. pyogenes* или ее нуклеазный домен. В некоторых вариантах осуществления последовательность, кодирующая Cas9 или ее нуклеазный домен, выделена или получена из последовательности, кодирующей Cas9 *S. aureus* или ее нуклеазный домен. В некоторых вариантах осуществления последовательность, кодирующая Cas9 или ее нуклеазный домен, является кодон-оптимизированной для экспрессии у млекопитающего. В некоторых вариантах осуществления последовательность, кодирующая Cas9 или ее нуклеазный домен, является кодон-оптимизированной для экспрессии у человека. В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает последовательность поли(A). В некоторых

вариантах осуществления поли(A)-последовательность является мини-поли(A)-последовательностью. В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает одну или более последовательностей, кодирующих инвертированный концевой повтор (ITR). В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает последовательность, кодирующую 5' инвертированный концевой повтор (ITR), и последовательность, кодирующую 3' ITR. В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV). В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV) серотипа 2 (AAV2). В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), и последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из AAV2. В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV) серотипа 4 (AAV4). В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), и последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из AAV4. В некоторых вариантах осуществления последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 145 нуклеотидов, 115 нуклеотидов или 141 нуклеотида. В некоторых вариантах осуществления нуклеиновая кислота дополнительно включает сигнал ядерной локализации. В некоторых вариантах осуществления нуклеиновая кислота оптимизирована для экспрессии в клетках млекопитающих. В некоторых вариантах осуществления нуклеиновая кислота

оптимизирована для экспрессии в клетках человека.

Также предложен вектор, включающий нуклеиновую кислоту, включающую последовательность, кодирующую промотор, и последовательность, кодирующую Cas9 или ее нуклеазный домен, где последовательность, кодирующая промотор, включает последовательность, кодирующую мышечно-специфический промотор, такой как промотор CK8 или CK8e. В некоторых вариантах осуществления вектор дополнительно включает последовательность, кодирующую инвертированный концевой повтор (ITR) транспозируемого элемента. В некоторых вариантах осуществления транспозируемый элемент является транспозоном. В некоторых вариантах осуществления транспозон является транспозоном Tn7. В некоторых вариантах осуществления вектор дополнительно включает последовательность, кодирующую 5' ITR транспозона T7, и последовательность, кодирующую 3' ITR транспозона T7. В некоторых вариантах осуществления вектор является невирусным вектором. В некоторых вариантах осуществления невирусный вектор является плазмидой. В некоторых вариантах осуществления вектор является вирусным вектором. В некоторых вариантах осуществления вирусный вектор является вектором на основе аденоассоциированного вируса (AAV). В некоторых вариантах осуществления AAV вектор является репликационно-дефектным или условно репликационно-дефектным. В некоторых вариантах осуществления AAV вектор является рекомбинантным вектором AAV. В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 1 (AAV1), 2 (AAV2), 3 (AAV3), 4 (AAV4), 5 (AAV5), 6 (AAV6), 7 (AAV7), 8 (AAV8), 9 (AAV9), 10 (AAV10), 11 (AAV11) или их любой комбинации. В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 9 (AAV9). В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 2 (AAV2). В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV2, и последовательность, выделенную или полученную из AAV9. В некоторых вариантах

осуществления AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 4 (AAV4). В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV4, и последовательность, выделенную или полученную из AAV9. В некоторых вариантах осуществления вектор оптимизирован для экспрессии в клетках млекопитающих. В некоторых вариантах осуществления вектор оптимизирован для экспрессии в клетках человека. В некоторых вариантах осуществления вектор включает последовательность нуклеиновой кислоты SEQ ID NO. 899, SEQ ID NO. 900, SEQ ID NO. 901 или SEQ ID NO. 902.

Также предложена клетка, включающая одну или более нуклеиновых кислот согласно настоящему описанию. В некоторых вариантах осуществления клетка является клеткой человека. В некоторых вариантах осуществления клетка является мышечной клеткой или сателлитной клеткой. В некоторых вариантах осуществления клетка является индуцированной плюрипотентной стволовой (иПС) клеткой.

Также предложена композиция, включающая одну или более нуклеиновых кислот согласно настоящему описанию. В некоторых вариантах осуществления композиция дополнительно включает фармацевтически приемлемый носитель.

Также предложена композиция, включающая клетку, включающую одну или более нуклеиновых кислот согласно настоящему описанию. В некоторых вариантах осуществления композиция дополнительно включает фармацевтически приемлемый носитель.

Также предложена композиция, включающая один или более векторов согласно настоящему описанию. В некоторых вариантах осуществления композиция дополнительно включает фармацевтически приемлемый носитель.

Также предложена клетка, включающая композицию, включающую один или более векторов согласно настоящему описанию. В некоторых вариантах осуществления клетка является человеческой клеткой. В некоторых вариантах осуществления клетка является мышечной клетка или сателлитной клеткой. В некоторых вариантах осуществления клетка является индуцированной плюрипотентной

стволовой (иПС) клеткой.

В некоторых вариантах осуществления композиция включает: (i) первую последовательность нуклеиновой кислоты, включающую последовательность, кодирующую первую DMD направляющую РНК, направленно взаимодействующую с первой геномной последовательностью-мишенью, последовательность, кодирующую вторую DMD направляющую РНК, направленно взаимодействующую со второй геномной последовательностью-мишенью, последовательность, кодирующую первый промотор, где первый промотор направляет экспрессию последовательности, кодирующей первую DMD направляющую РНК, и последовательность, кодирующую второй промотор, где первый промотор направляет экспрессию последовательности, кодирующей вторую DMD направляющую РНК, где каждая первая геномная последовательность-мишень и вторая геномная последовательность-мишень включает акцептор сплайсинга дистрофина, и (ii) вторую последовательность нуклеиновой кислоты, включающую последовательность, кодирующую промотор, и последовательность, кодирующую Cas9 или ее нуклеазный домен, где последовательность, кодирующая промотор, включает последовательность, кодирующую мышечно-специфический промотор, такой как промотор CK8 или CK8e. В некоторых вариантах осуществления композиция дополнительно включает фармацевтически приемлемый носитель.

В некоторых вариантах осуществления композиция включает: (i) первый вектор, включающий последовательность нуклеиновой кислоты, включающую последовательность, кодирующую первую DMD направляющую РНК, направленно взаимодействующую с первой геномной последовательностью-мишенью, последовательность, кодирующую вторую DMD направляющую РНК, направленно взаимодействующую со второй геномной последовательностью-мишенью, последовательность, кодирующую первый промотор, где первый промотор направляет экспрессию последовательности, кодирующей первую DMD направляющую РНК, и последовательность, кодирующую второй промотор, где первый промотор направляет экспрессию последовательности, кодирующей вторую DMD направляющую РНК, где каждая первая геномная последовательность-

мишень и вторая геномная последовательность-мишень включает акцептор сплайсинга дистрофина, и (ii) второй вектор, включающий последовательность нуклеиновой кислоты, включающую последовательность, кодирующую промотор, и последовательность, кодирующую Cas9 или ее нуклеазный домен, где последовательность, кодирующая промотор, включает последовательность, кодирующую мышечно-специфический промотор, такой как промотор CK8 или CK8e. В некоторых вариантах осуществления по меньшей мере одним из первого вектора и второго вектора является AAV. В некоторых вариантах осуществления композиция дополнительно включает фармацевтически приемлемый носитель.

Также предложен способ коррекции нарушения дистрофина, включающий контакт клетки с одной или более композициями согласно настоящему описанию при условиях, подходящих для экспрессии первой DMD направляющей РНК, второй DMD направляющей РНК и белка Cas9 или его нуклеазного домена, где по меньшей мере одна из первой DMD направляющей РНК или второй DMD направляющей РНК образует комплекс с белком Cas9 или его нуклеазным доменом, с образованием по меньшей мере одного комплекса DMD направляющей РНК-Cas9, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 разрушает сайт сплайсинга дистрофина и вызывает селективный пропуск экзона DMD.

Также предложен способ коррекции нарушения дистрофина, включающий контакт клетки с одной или более композициями согласно настоящему описанию при условиях, подходящих для экспрессии первой DMD направляющей РНК, второй DMD направляющей РНК и белка Cas9 или его нуклеазного домена, где по меньшей мере одна из первой DMD направляющей РНК или второй DMD направляющей РНК образует комплекс с белком Cas9 или его нуклеазным доменом, с образованием по меньшей мере одного комплекса DMD направляющей РНК-Cas9, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 вызывает восстановление рамки считывания дистрофина. В некоторых вариантах осуществления восстановление рамки считывания дистрофина вызывает вставку. В некоторых вариантах осуществления вставка включает или состоит из одного нуклеотида аденозина.

Также предложен способ индукции селективного пропуска экзона DMD, включающий контакт клетки с одной или более композициями согласно настоящему описанию при условиях, подходящих для экспрессии первой DMD направляющей РНК, второй DMD направляющей РНК и белка Cas9 или его нуклеазного домена, где по меньшей мере одна из первой DMD направляющей РНК или второй DMD направляющей РНК образует комплекс с белком Cas9 или его нуклеазным доменом, с образованием по меньшей мере одного комплекса DMD направляющей РНК-Cas9, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 разрушает сайт сплайсинга дистрофина и вызывает селективный пропуск экзона DMD.

Также предложен способ индукции события восстановления в рамке считывания дистрофина, включающий контакт клетки с одной или более композициями согласно настоящему описанию при условиях, подходящих для экспрессии первой DMD направляющей РНК, второй DMD направляющей РНК и белка Cas9 или его нуклеазного домена, где по меньшей мере одна из первой DMD направляющей РНК или второй DMD направляющей РНК образует комплекс с белком Cas9 или его нуклеазным доменом, с образованием по меньшей мере одного комплекса DMD направляющей РНК-Cas9, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 разрушает сайт сплайсинга дистрофина и вызывает селективный пропуск экзона DMD. В некоторых вариантах осуществления по меньшей мере один комплекс DMD направляющей РНК-Cas9 разрушает сайт сплайсинга дистрофина и вызывает селективный пропуск экзона 51 гена DMD человека.

Также предложен способ лечения миодистрофии у нуждающегося в этом субъекта, включающий введение субъекту терапевтически эффективного количества одной или более композиций согласно настоящему описанию. В некоторых вариантах осуществления композицию вводят местно. В некоторых вариантах осуществления композицию вводят непосредственно в мышечную ткань. В некоторых вариантах осуществления композицию вводят путем внутримышечной инфузии или инъекции. В некоторых вариантах осуществления мышечная ткань включает ткань передней большеберцовой мышцы, ткань четырехглавой мышцы, ткань камбаловидной мышцы, ткань

диафрагмы или ткань сердца. В некоторых вариантах осуществления композицию вводят внутрисердечной инъекцией. В некоторых вариантах осуществления композицию вводят системно. В некоторых вариантах осуществления композицию вводят путем внутривенной инфузии или инъекции. В некоторых вариантах осуществления после введения композиции субъект демонстрирует нормальные дистрофин-положительные мышечные волокна и мозаичные дистрофин-положительные мышечные волокна, содержащие централизованные ядра, или комбинацию этого. В некоторых вариантах осуществления после введения композиции субъект демонстрирует появление или увеличение уровня распространенности нормальных дистрофин-положительных мышечных волокон по сравнению с отсутствием или уровнем распространенности нормальных дистрофин-положительных мышечных волокон до введения композиции. В некоторых вариантах осуществления после введения композиции субъект демонстрирует появление или увеличение уровня распространенности мозаичных дистрофин-положительных мышечных волокон, содержащих централизованные ядра, по сравнению с отсутствием или уровнем распространенности мозаичных дистрофин-положительных мышечных волокон, содержащих централизованные ядра, до введения композиции. В некоторых вариантах осуществления после введения композиции субъект демонстрирует пониженный уровень КК сыворотки по сравнению с уровнем КК сыворотки до введения композиции. В некотором варианте осуществления после введения композиции субъект демонстрирует улучшение силы хвата по сравнению с силой хвата до введения композиции. В некоторых вариантах осуществления субъект является новорожденным, младенецем, ребенком, молодым совершеннолетним или взрослым. В некоторых вариантах осуществления субъект имеет миодистрофию. В некоторых вариантах осуществления субъект является носителем гена миодистрофии. В некоторых вариантах осуществления субъект является субъектом мужского пола. В некоторых вариантах осуществления субъект является субъектом женского пола. В некоторых вариантах осуществления субъект представляется бессимптомным, и при этом генетический диагноз обнаруживает мутацию в одной или обеих копиях гена DMD, которая нарушает

функцию продукта гена DMD. В некоторых вариантах осуществления субъект демонстрирует ранний признак или симптом миодистрофии. В некоторых вариантах осуществления ранний признак или симптом миодистрофии включает потерю мышечной массы или слабость проксимальных мышц. В некоторых вариантах осуществления потеря мышечной массы или слабость проксимальных мышц наблюдается в одной или обеих ногах и/или тазе, за которыми следует одна или более мышц верхней части тела. В некоторых вариантах осуществления ранний признак или симптом миодистрофии дополнительно включает псевдогипертрофию, низкую выносливость, затруднение при стоянии, затруднение при ходьбе, затруднение при подъеме по лестнице или их комбинацию. В некоторых вариантах осуществления субъект демонстрирует прогрессирующий признак или симптом миодистрофии. В некоторых вариантах осуществления прогрессирующий признак или симптом миодистрофии включает истощение мышечной ткани, замещение мышечной ткани жиром или замещение мышечной ткани фиброзной тканью. В некоторых вариантах осуществления субъект демонстрирует поздний признак или симптом миодистрофии. В некоторых вариантах осуществления поздний признак или симптом миодистрофии включает патологическое развитие костей, искривление позвоночника, потерю двигательных функций и паралич. В некоторых вариантах осуществления субъект демонстрирует неврологический признак или симптом миодистрофии. В некоторых вариантах осуществления неврологический признак или симптом миодистрофии включает нарушение умственной деятельности и паралич. В некоторых вариантах осуществления введение композиции происходит до проявления у субъекта одного или более прогрессирующих, поздних или неврологических признаков или симптомов миодистрофии. В некоторых вариантах субъект имеет возраст меньше 10 лет, меньше 5 лет или меньше 2 лет.

Также предложено применение терапевтически эффективного количества одной или более композиций согласно настоящему описанию для лечения миодистрофии у нуждающегося в этом субъекта.

При использовании в настоящем описании "a" или "an" могут означать один или более. При использовании в настоящем документе

в формуле изобретения, в случае использования вместе со словом "включающий", слова "a" или "an" могут означать один или больше чем один.

Использование термина "или" в формуле изобретения используется для обозначения "и/или", если прямо не указано, что оно относится только к альтернативам, или альтернативы являются взаимоисключающими, хотя описание поддерживает определение, которое относится только к альтернативам и "и/или". При использовании в настоящем документе "другой" может означать по меньшей мере второй или более.

По всему тексту настоящей заявки термин "приблизительно" используется для указания того, что значение включает в себя внутреннюю погрешность для устройства, для способа, используемого для определения значения, или вариацию, которая существует среди участников исследования. Такая внутренняя вариация может быть вариацией, составляющей ±10% от указанного значения.

По всему тексту настоящей заявки нуклеотидные последовательности перечислены в направлении от 5' к 3', а аминокислотные последовательности перечислены в направлении от N-конца к C-концу, если не указано иное.

Другие объекты, признаки и преимущества настоящего изобретения станут очевидными из следующего подробного описания. Однако следует понимать, что подробное описание и конкретные примеры, в которых хотя и указаны предпочтительные варианты осуществления изобретения, приведены только в качестве иллюстрации, поскольку различные изменения и модификации в пределах сущности и объема изобретения станут очевидными для специалистов в данной области из этого подробного описания.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

Файл патента или заявки содержит как минимум один чертеж, выполненный в цвете. Копии этого патента или публикации заявки на патент с цветными чертежами будут предоставлены Патентным ведомством по запросу и при уплате необходимой пошлины.

Следующие чертежи составляют часть настоящего описания и включены для дополнительной демонстрации некоторых аспектов

настоящего изобретения. Изобретение может быть лучше понято при обращении к одному или нескольким из этих чертежей в сочетании с подробным описанием конкретных вариантов осуществления, представленных в настоящем документе.

ФИГ. 1A–F. Модель на "гуманизированных"–ΔEx50 мышах. (ФИГ. 1A) Стратегия, демонстрирующая метод CRISPR/Cas9-опосредованного редактирования генома для создания модели на гуманизированных мышах. (ФИГ. 1B) анализ ОТ-ПЦР для подтверждения делеции экзона 50 (ΔEx50). (ФИГ. 1C) Последовательность продукта ОТ-ПЦР для подтверждения делеции экзона 50 и получения последовательности вне рамки считывания (последовательность нуклеиновой кислоты=SEQ ID NO: 1; аминокислотная последовательность=SEQ ID NO: 2). (ФИГ. 1E) Вестерн-блоттинг экспрессии дистрофина и винкулина в тканях передней большеберцовой мышцы и сердца. (ФИГ. 1F) Уровни КК сыворотки, маркера миодистрофии, который отражает повреждение мышцы и проницаемость мембраны, измеряли в мышей дикого типа (WT), ΔEx50 и мышей mdx. (ФИГ. 1D) Гистохимия сердечной и скелетной мышцы при окрашивании гематоксилином и эозином (Г-Э) и иммуногистохимия с использованием антитела к дистрофину. Масштабная линейка: 50 мкм.

ФИГ. 2A–B. Пропуск экзона 51. (ФИГ. 2A) ОТ-ПЦР РНК мышей ΔEx50 через 3 недели после внутримышечной инъекции указывает на делецию экзона 51 (обозначена ΔEx50-51, нижняя полоса). (ФИГ. 2B) Последовательность продуктов ОТ-ПЦР полосы ΔEx50-51 подтвердила, что экзон 49 соединен непосредственно с экзоном 52, с исключением экзона 51 (последовательность нуклеиновой кислоты=SEQ ID NO: 3; аминокислотная последовательность=SEQ ID NO: 4).

ФИГ. 3A–G. Коррекция экспрессии дистрофина с применением стратегии тройной gРНК через 3 недели после внутримышечной инъекции. (ФИГ. 3A) Стратегия, демонстрирующая метод CRISPR/Cas9-опосредованного редактирования генома для коррекции рамки считывания в модели на мышах ΔEx50. (ФИГ. 3B) sgРНК, направленно взаимодействующая с последовательностью акцепторного сайта сплайсинга (sgRNA-ex51-SA) и схематическая иллюстрация положения связывания sgРНК. На ФИГ. 3B раскрыты SEQ ID NO: 954-

957, соответственно, в порядке появления. (ФИГ. 3C) Схематическая иллюстрация AAV плазмид для инъекции и стратегия. Стратегия с двойным гидом с использованием плазмиды rAAV9-sgRNA для sgRNA-ex51-SA и sgRNA-ex51-SD. Тройная конструкция с использованием плазмиды rAAV9-sgRNA, содержащей 3 копии sgRNA-ex51-SA. Промотор мышечной креатинкиназы 8 (CK8) для экспрессии SpCas9. U6, H1 и 7SK промотор для РНК-полимеразы III для экспрессии sgРНК. (ФИГ. 3D) иммуногистохимическое окрашивание на дистрофин передней большеберцовой мышцы. (ФИГ. 3E) количественное определение дистрофин-положительных волокон, нормализованное по площади. (ФИГ. 3F) Вестерн-блоттинг экспрессии дистрофина (DMD) и винкулина (VCL) через 3 недели после внутримышечной инъекции. (ФИГ. 3G) количественное определение экспрессии дистрофина после нормализации по винкулину. Данные представлены как среднее ±SEM. **P<0,005. Масштабная линейка: 50 мкм

ФИГ. 4A-B. Гистологическое улучшение мышцы через 3 недели после инъекции. (ФИГ. 4A) Гистохимия передней большеберцовой мышцы при окрашивании гематоксилином и эозином (Г-Э). (ФИГ. 4B) Количественное определение размера волокок и процентной частоты. Данные представлены как среднее ±SEM. Масштабная линейка: 50 мкм. Каждый набор из четырех точек данных для данного размера волокон соответствует Δ50-CTL, WT-CTL, Δ50-AAV-TriSA и Δ50-AAV-SA+SD, слева направо.

ФИГ. 5. Скрининг sgРНК в человеческих клетках 293 и мышиных клетках 10T. Нерасщепленные продукты ПЦР (верхняя панель) и T7E1 перевар (нижняя панель) на 2% агарозном геле. M обозначает дорожку маркера молекулярного веса. пн обозначает длину полос маркера.

ФИГ. 6. Иммуногистохимическое окрашивание на дистрофин всей передней большеберцовой мышцы.

ФИГ. 7A-B. Стратегия CRISPR/Cas9-опосредованного редактирования генома у мышей ΔEx50. (ФИГ. 7A) Стратегия, демонстрирующая метод CRISPR/Cas9-опосредованного редактирования генома для коррекции рамки считывания в модели на мышах ΔEx50. (ФИГ. 7B) sgРНК, направленно взаимодействующая с

последовательностью акцепторного сайта сплайсинга (sgRNA-ex51-SA2), и схематическая иллюстрация положения связывания sgРНК. На ФИГ. 7В раскрыты SEQ ID NO: 958-959, соответственно, в порядке появления.

ФИГ. 8A-F. sgРНК геномный анализ в мышиных и человеческих клетках. (ФИГ. 8A) Cas9 экспрессировалась в присутствии или отсутствии мышиной sgRNA-sgRNA-51-SA2 в клетках 10T1/2, и редактирование гена контролировали с помощью анализа T7E1 в отсортированных на основе флуоресценции клетках (FACS) (+) и неотсортированных клетках (-). GFP использовали в качестве контроля. Нерасщепленные продукты ПЦР (верхняя панель) и перевар T7E1 (нижняя панель) показаны на 2% агарозном геле. Черная стрелка указывает нерасщепленную полосу ПЦР 771 пн. Зеленые стрелки на нижней панели указывают разрезанные полосы в T7E1 анализе. M обозначает маркер молекулярного веса. пн указывает длину полос маркера. (ФИГ. 8B) Геномный анализ методом глубокого секвенирования ПЦР ампликонов, полученных по целевому участку экзона 51 в клетках 10T1/2. Последовательность репрезентативных инделов, выровненная с последовательностью sgРНК (выделенной синим), содержащая вставки (выделенные зеленым) и делеции (выделенные красным) (SEQ ID NO: 960-966, соответственно, в порядке появления). Линия указывает на предсказанную последовательность экзонных энхансеров сплайсинга (ESE), расположенную в сайте sgРНК. Черная стрелка указывает сайт расщепления. (ФИГ. 8C) Последовательность экзона 51 мыши (SEQ ID NO: 967) с предсказанными экзонными энхансерами сплайсинга (ESE), расположенными в сайте sgРНК, указана красным цветом. Последовательность экзона 51 человека (SEQ ID NO: 968) с предсказанными экзонными энхансерами сплайсинга (ESE), расположенными в сайте sgРНК, указана красным цветом. (ФИГ. 8D) Мышиные и человеческие сайты ESE экзона 51 предсказаны с использованием ESEfinder3. (ФИГ. 8E) Cas9 экспрессировалась в присутствии или отсутствии мышиной sgRNA-sgRNA-51-SA2 в человеческих T-клетках 293, и редактирование гена контролировали с помощью T7E1 анализа в отсортированных на основе флуоресценции клетках (FACS) (+) и неотсортированных клетках (-). GFP

использовали в качестве контроля. Нерасщепленные продукты ПЦР (верхняя панель) и перевар T7E1 (нижняя панель) показаны на 2% агарозном геле. Черная стрелка указывает на нерасщепленную полосу ПЦР 771 пн. Зеленые стрелки на нижней панели указывают на разрезанные полосы в T7E1 анализе. M обозначает маркер молекулярного веса. пн указывает длину полос маркера. (ФИГ. 8F) Последовательность репрезентативных инделов, выровненная с последовательностью sgРНК (указана синим цветом), содержит делеции и вставки (SEQ ID NO: 969-978, соответственно, в порядке появления). Черная стрелка указывает сайт расщепления.

ФИГ. 9A-B. Схематическая иллюстрация плазмид AAV для инъекции и стратегия. (ФИГ. 9A) Промотор мышечной креатинкиназы 8 (CK8) для экспрессии SpCas9. (ФИГ. 9B) Тройная конструкция с использованием плазмиды rAAV9-sgRNA, содержащей 3 копии sgRNA-ex51-SA2. U6, H1 и 7SK промотор для РНК-полимеразы III для экспрессии sgРНК.

ФИГ. 10A-B. Редактирование гена *Dmd in vivo*. (ФИГ. 10A) Нерасщепленные продукты ПЦР (верхняя панель) и перевар T7E1 (нижняя панель) показаны на 2% агарозном геле образцов TA (передней большеберцовой мышцы) мышцы мышей WT и мышей ΔEx50 через 3 недели после внутримышечной инъекции векторов экспрессии AAV9-sgRNA-SA2 и AAV9-Cas9. Контролям вводили только AAV9-Cas9, но не AAV9-sgRNA-SA2. Черная стрелка на верхней панели указывает 771 пн полосу ПЦР. Зеленые стрелки на нижней панели указывают разрезанные полосы в T7E1 анализе. M обозначает маркер молекулярного веса. пн указывает длину полос маркера. N=4. (ФИГ. 10B) Геномный анализ методом глубокого секвенирования ПЦР ампликонов, полученных через целевой участок экзона 51 мышей ΔEx50 после инъекции AAV9-sgRNA-51-SA2 и AAV9-Cas9. Последовательность репрезентативных инделов, выровненная с последовательностью sgРНК (указана синим цветом) содержит вставки (выделены зеленым) и делеции (выделены красным). (SEQ ID NO: 979-1014, соответственно, в порядке появления). Черные стрелки указывают сайт расщепления. n=3.

ФИГ. 11A-D. ОТ-ПЦР анализ коррекции рамки считывания. (ФИГ. 11A) ОТ-ПЦР РНК из передней большеберцовой мышцы мышей дикого

типа (WT) и мышей ΔEx50 через 3 недели после внутримышечной инъекции векторов экспрессии sgRNA-51-SA2 и Cas9. Нижние полосы дистрофина указывают на делецию экзона 51. Указаны положения праймеров в экзонах 48 и 53 (Fw, Rv). (ФИГ. 11B) Процент событий, обнаруженных в экзоне 51 после обработки AAV9-sgRNA-51-SA2 с помощью анализа последовательностей ОТ-ПЦР полученных клонов TOPO-TA. Для каждого из 4 различных образцов были получены и секвенированы 40 клонов. Продукты ОТ-ПЦР разделяли на 4 группы: нередактированные (NE), с пропуском экзона 51 (SK), с восстановленной рамкой считывания (RF) и вне рамки считывания (OF). (ФИГ. 11C) Последовательность продуктов ОТ-ПЦР нижней полосы дистрофина мыши ΔEx50-51 подтвердила, что экзон 49 был соединен непосредственно с экзоном 52, с исключением экзона 51. Последовательность продуктов ОТ-ПЦР ΔEx50 с восстановленной рамкой считывания (ΔEx50-RF). На ФИГ. 11C раскрыты SEQ ID NO: 1015-1022, соответственно, в порядке появления. (ФИГ. 11D) Анализ методом глубокого секвенирования продуктов ОТ-ПЦР из верхней полосы, содержащей ΔEx50 нередактированный (NE) и ΔEx50-RF. Последовательность продуктов ОТ-ПЦР, содержащих вставки (выделены зеленым) и делеции (выделены красным). n=4. Данные представлены как среднее ±SEM. На ФИГ. 11D раскрыты SEQ ID NO: 1023-1026, соответственно, в порядке появления.

ФИГ. 12A-D. Внутримышечная инъекция AAV9-Cas9 и AAV9-sgRNA-51-SA2 корректирует экспрессию дистрофина. (ФИГ. 12A) В передние большеберцовые мышцы мышей ΔEx50 вводили вектор AAV9, кодирующий sgRNA и Cas9, и проводили анализ через 3 недели с помощью иммуноокрашивания на дистрофин. Контрольным мышам дикого типа (WT-CTL) и контрольным мышам ΔEx50 (ΔEx50-CTL) вводили только AAV9-Cas9 без sgRNA. Проценты дистрофин-положительных мышечных волокон у мышей ΔEx50-CTL и у обработанных мышей ΔEx50 (ΔEx50-AAV9-sgRNA-51-SA2 и AAV9-Cas9) в сравнении с WT-CTL указаны в каждой панели. (ФИГ. 12B) Окрашивание гематоксилином и эозином (Г-Э) передних большеберцовых мышц. (ФИГ. 12C) Вестерн-блоттинг экспрессии дистрофина (DMD) и винкулина (VCL) в передних большеберцовых мышцах через 3 недели после внутримышечной инъекции. (ФИГ. 12D) Количественное определение экспрессии

дистрофина на блотах после нормализации по винкулину. Звездочка указывает неспецифичные иммунореактивные полосы. n=5 для AAV9-sgRNA-51-SA2. Масштабная линейка: 50μm

ФИГ. 13. Восстановление экспрессии дистрофина после внутримышечных инъекций AAV9-Cas9 и AAV9-sgRNA-51-SA2 в модели на мышах ΔEx50. Иммуногистохимия дистрофина всей передней большеберцовой мышцы. Мышам CTL вводили только AAV9-Cas9 без AAV9-sgRNA-51-SA2. n=5

ФИГ. 14. Гистологическое улучшение мышцы через 3 недели после инъекции. Гистохимия передней большеберцовой мышцы при окрашивании гематоксилином и эозином (Г-Э).

ФИГ. 15A-B. Восстановление экспрессии дистрофина после внутримышечных инъекций AAV9-Cas9 в сочетании с различными AAV9, экспрессирующими одиночную копию или тройную копию sgРНК в модели на мышах ΔEx50. (ФИГ. 15A) Промоторы U6, H1 и 7SK для РНК-полимеразы III использовали индивидуально для экспрессии sgРНК в одиночной копии (AAV9-U6-sgRNA-51-SA2; AAV9-H1-sgRNA-51-SA2; AAV9-7SK-sgRNA-51-SA2) или тройной копии. (ФИГ. 15B) Иммуногистохимия дистрофина всей передней большеберцовой мышцы. Контрольным (CTL) мышам вводили только AAV9-Cas9 без AAV9-sgRNA-51-SA2.

ФИГ. 16A-B. Восстановление экспрессии дистрофина через 4 недели после системной доставки AAV9-Cas9 и AAV9-sgRNA-51-SA2 мышам ΔEx50. (ФИГ. 16A) Иммуноокрашивание на дистрофин передней большеберцовой мышцы (TA), трехглавой мышцы, диафрагмы и сердечных мышц через 8 недель после системной инъекции AAV9-sgRNA-51. (ФИГ. 16B) Вестерн-блоттинг экспрессии дистрофина (DMD) и винкулина (VCL) в TA, трехглавой мышце, мышцах диафрагмы и сердце. n=5 для каждой группы. Масштабная линейка: 50 мкм.

ФИГ. 17A-B. Восстановление экспрессии дистрофина через 8 недель после системной доставки AAV9-Cas9 и AAV9-sgRNA-51-SA2 мышам ΔEx50. (ФИГ. 17A) Иммунокрашивание на дистрофин передней большеберцовой мышцы (TA), трехглавой мышцы, диафрагмы и сердечных мышц через 8 недель после системной инъекции AAV9-sgRNA-51. (ФИГ. 17B) Вестерн-блоттинг экспрессии дистрофина (DMD) и винкулина (VCL) в TA, трехглавой мышце, мышцах диафрагмы

и сердце. n=5 для каждой группы. Масштабная линейка: 50 мкм.

ФИГ. 18А–В. Функциональное улучшение через 4 недели после системной доставки AAV9–Cas9 и AAV9–sgRNA–51–SA2 мышам ΔEx50. (ФИГ. 18А) Мышей дикого типа (WT), контрольных мышей ΔEx50 и мышей ΔEx50, обработанных AAV9–sgRNA–51 (ΔEx50–AAV9–sgRNA–51), подвергали тестированию силы хвата с целью измерения мышечной работоспособности (грамм-сил). (ФИГ. 18В) Креатинкиназу (СК) сыворотки измеряли у мышей WT, ΔEx50 и ΔEx50 AAV9–sgRNA–51. n=5. Звездочка указывает неспецифические иммунореактивные полосы. Данные представлены как среднее ±SEM.

ФИГ. 19. Коррекция экспрессии дистрофина через 6 недель после внутримышечной инъекции у собаки ΔEx50–MD. Иммуногистохимическое окрашивание на дистрофин краниальной большеберцовой мышцы собаки дикого типа (Nathan), необработанной собаки ΔEx50–MD и двух собак ΔEx50–MD (Norman и Newton) без контралатеральной инъекции и после инъекции AAV9–Cas9–sgRNA в краниальную большеберцовую мышцу. Масштабная линейка: 50 мкм.

ФИГИ. 20А–В. Коррекция экспрессии дистрофина через 6 недель после внутримышечной инъекции у собаки ΔEx50–MD. (ФИГ. 20А) Вестерн-блоттинг дистрофина (DMD) и винкулина (VCL) в краниальной большеберцовой мышце собаки дикого типа (Nathan), необработанной собаки ΔEx50–MD и двух собак ΔEx50–MD (Norman и Newton) без контралатеральной инъекции и после инъекции AAV9–Cas9–sgRNA в краниальную большеберцовую мышцу. (ФИГ. 20В) Количественное определение экспрессии дистрофина в двух независимых блотах после нормализации по винкулину.

ФИГ. 21. Гистологическое улучшение мышцы через 6 недель после инъекции у собаки ΔEx50–MD. Гистохимия с окрашиванием гематоксилином и эозином (Г–Э) краниальной большеберцовой мышцы собаки дикого типа (Nathan), собаки ΔEx50–MD (Norman и Newton) без контралатеральной инъекции и после инъекции AAV9–Cas9–sgRNA в краниальную большеберцовую мышцу. Масштабная линейка: 50 мкм.

ПОДРОБНОЕ ОПИСАНИЕ

МДД является новым мутационным синдромом более чем с 4000 независимых мутаций, которые были обнаружены у людей (сайт в интернете по адресу dmd.nl). Большинство мутаций у пациентов

включают делеции, которые группируются в горячей точке, и, таким образом, терапевтический метод пропуска некоторых экзонов применяют к большой группе пациентов. Обоснование метода пропуска экзонов основано на генетическом различии между пациентами с МДД и миодистрофией Беккера (МДБ). У пациентов с МДД рамка считывания мРНК дистрофина прерывается, что приводит к образованию усеченных нефункциональных белков дистрофина. У пациентов с МДБ наблюдаются мутации в гене DMD, при которых рамка считывания сохраняется, позволяя образовываться содержащим внутреннюю делецию, но частично функциональным дистрофинам, что приводит к гораздо более легким симптомам заболевания по сравнению с больными МДД.

В настоящем документе предложены композиции и способы лечения МДД. В некоторых вариантах осуществления предложена конструкция AAV, где конструкция AAV включает нуклеиновую кислоту, кодирующую три промотора, каждый из которых регулирует экспрессию DMD направляющей РНК. С применением композиции и способов, раскрытых в настоящем документе, была создана более эффективная и безопасная форма редактирования генома в модели МДД на гуманизированных мышах с делецией в экзоне 50 и на собаке ΔEx50-MD. Эти и другие аспекты согласно настоящему описанию представлены ниже.

I. Миодистрофия Дюшена

A. Общие сведения

Миодистрофия Дюшена (МДД) – рецессивная Х-сцепленная форма миодистрофии, поражающая приблизительно одного из 5000 мальчиков, которая приводит к дегенерации мышц и преждевременной смерти. Нарушение вызвано мутацией в гене дистрофина (см. регистрационный номер в GenBank NC_000023.11), расположенном на Х-хромосоме человека, который кодирует белок дистрофин (регистрационный номер в GenBank AAA53189; SEQ ID NO: 5), последовательность которого приведена ниже:

1 mlwweevedc yeredvqkkt ftkwvnaqfs kfgkqhienl fsdlqdgrrl ldllegltgq

61 klpkekgstr vhalnnvnka lrvlqnnnvd lvnigstdiv dgnhkltlgl

iwniilhwqv
121 knvmknimag lqqtnsekil lswvrqstrn ypqvnvinft tswsdglaln
alihshrpdl
181 fdwnsvvcqq satqrlehaf niaryqlgie klldpedvdt typdkksilm
yitslfqvlp
241 qqvsieaiqe vemlprppkv tkeehfqlhh qmhysqqitv slaqgyerts
spkprfksya
301 ytqaayvtts dptrspfpsq hleapedksf gsslmesevn ldryqtalee
vlswllsaed
361 tlqaqgeisn dvevvkdqfh thegymmdlt ahqgrvgnil qlgskligtg
klsedeetev
421 qeqmnllnsr weclrvasme kqsnlhrvlm dlqnqklkel ndwltkteer
trkmeeeplg
481 pdledlkrqv qqhkvlqedl eqeqvrvnsl thmvvvvdes sgdhataale
eqlkvlgdrw
541 anicrwtedr wvllqdillk wqrlteeqcl fsawlseked avnkihttgf
kdqnemlssl
601 qklavlkadl ekkkqsmgkl yslkqdllst lknksvtqkt eawldnfarc
wdnlvqklek
661 staqisqavt ttqpsltqtt vmetvttvtt reqilvkhaq eelpppppqk
krqitvdsei
721 rkrldvdite lhswitrsea vlqspefaif rkegnfsdlk ekvnaierek
aekfrklqda
781 srsaqalveq mvnegvnads ikqaseqlns rwiefcqlls erlnwleyqn
niiafynqlq
841 qleqmtttae nwlkiqpttp septaiksql kickdevnrl sglqpqierl
kiqsialkek
901 gqgpmfldad fvaftnhfkq vfsdvqarek elqtifdtlp pmryqetmsa
irtwvqqset
961 klsipqlsvt dyeimeqrlg elqalqsslq eqqsglyyls ttvkemskka
pseisrkyqs
1021 efeeiegrwk klssqlvehc qkleeqmnkl rkiqnhiqtl kkwmaevdvf
lkeewpalgd
1081 seilkkqlkq crllvsdiqt iqpslnsvne ggqkikneae pefasrlete
lkelntqwdh

1141 mcqqvyarke alkgglektv slqkdlsemh ewmtqaeeey lerdfeyktp
delqkaveem
1201 krakeeaqqk eakvkllltes vnsviaqapp vaqealkkel etlttnyqwl
ctrlngkckt
1261 leevwacwhe llsylekank wlnevefklk ttenipggae eisevldsle
nlmrhsednp
1321 nqirilaqtl tdggvmdeli neeletfnsr wrelheeavr rqklleqsiq
saqetekslh
1381 liqesltfid kqlaayiadk vdaaqmpqea qkiqsdltsh eisleemkkh
nqgkeaaqrv
1441 lsqidvaqkk lqdvsmkfrl fqkpanfelr lqeskmilde vkmhlpalet
ksveqevvqs
1501 qlnhcvnlyk slsevkseve mviktgrqiv qkkqtenpke ldervtalkl
hynelgakvt
1561 erkqqlekcl klsrkmrkem nvltewlaat dmeltkrsav egmpsnldse
vawgkatqke
1621 iekqkvhlks itevgealkt vlgkketlve dklsllnsnw iavtsraeew
lnllleyqkh
1681 metfdqnvdh itkwiiqadt lldesekkkp qqkedvlkrl kaelndirpk
vdstrdqaan
1741 lmanrgdhcr klvepqisel nhrfaaishr iktgkasipl keleqfnsdi
qkllepleae
1801 iqqgvnlkee dfnkdmnedn egtvkellqr gdnlqqritd erkreeikik
qqllqtkhna
1861 lkdlrsqrrk kaleishqwy qykrqaddll kclddiekkl aslpeprder
kikeidrelq
1921 kkkeelnavr rqaeglsedg aamaveptqi qlskrwreie skfaqfrrln
faqihtvree
1981 tmmvmtedmp leisyvpsty lteithvsqa lleveqllna pdlcakdfed
lfkqeeslkn
2041 ikdslqqssg ridiihskkt aalqsatpve rvklqealsq ldfqwekvnk
mykdrqgrfd
2101 rsvekwrrfh ydikifnqwl teaeqflrkt qipenwehak ykwylkelqd
gigqrqtvvr
2161 tlnatgeeii qqssktdasi lqeklgslnl rwqevckqls drkkrleeqk

nilsefqrdl
2221 nefvlwleea dniasiplep gkeqqlkekl eqvkllveel plrqgilkql
netggpvlvs
2281 apispeeqdk lenklkqtnl qwikvsralp ekqgeieaqi kdlgqlekkl
edleeqlnhl
2341 llwlspirnq leiynqpnqe gpfdvqetei avqakqpdve eilskgqhly
kekpatqpvk
2401 rkledlssew kavnrllqel rakqpdlapg lttigasptq tvtlvtqpvv
tketaiskle
2461 mpsslmlevp aladfnrawt eltdwlslld qviksqrvmv gdledinemi
ikqkatmqdl
2521 eqrrpqleel itaaqnlknk tsnqeartii tdrieriqnq wdevqehlqn
rrqqlnemlk
2581 dstqwleake eaeqvlgqar akleswkegp ytvdaiqkki tetkqlakdl
rqwqtnvdva
2641 ndlalkllrd ysaddtrkvh miteninasw rsihkrvser eaaleethrl
lqqfpldlek
2701 flawlteaet tanvlqdatr kerlledskg vkelmkqwqd lqgeieahtd
vyhnldensq
2761 kilrslegsd davllqrrld nmnfkwselr kkslnirshl eassdqwkrl
hlslqellvw
2821 lqlkddelsr qapiggdfpa vqkqndvhra fkrelktkep vimstletvr
iflteqpleg
2881 leklyqepre lppeeraqnv trllrkqaee vnteweklnl hsadwqrkid
etlerlqelq
2941 eatdeldlkl rqaevikgsw qpvgdllids lqdhlekvka lrgeiaplke
nvshvndlar
3001 qlttlgiqls pynlstledl ntrwkllqva vedrvrqlhe ahrdfgpasq
hflstsvqgp
3061 weraispnkv pyyinhetqt tcwdhpkmte lyqsladlnn vrfsayrtam
klrrlqkalc
3121 ldllslsaac daldqhnlkq ndqpmdilqi inclttiydr leqehnnlvn
vplcvdmcln
3181 wllnvydtgr tgrirvlsfk tgiislckah ledkyrylfk qvasstgfcd
qrrlglllhd

3241 siqiprqlge vasfggsnie psvrscfqfa nnkpeieaal fldwmrlepq
smvwlpvlhr
3301 vaaaetakhq akcnickecp iigfryrslk hfnydicqsc ffsgrvakgh
kmhypmveyc
3361 tpttsgedvr dfakvlknkf rtkryfakhp rmgylpvqtv legdnmetpv
tlinfwpvds
3421 apasspqlsh ddthsriehy asrlaemens ngsylndsis pnesiddehl
liqhycqsln
3481 qdsplsqprs paqilisles eergeleril adleeenrnl qaeydrlkqq
hehkglsplp
3541 sppemmptsp qsprdaelia eakllrqhkg rlearmqile dhnkqlesql
hrlrqlleqp
3601 qaeakvngtt vsspstslqr sdssqpmllr vvgsqtsdsm geedllsppq
dtstgleevm
3661 eqlnnsfpss rgrntpgkpm redtm.

У людей, мРНК дистрофина содержит 79 экзонов. мРНК дистрофина, как известно, подвергается альтернативному сплайсингу, что приводит к образованию различных изоформ. Примеры изоформ дистрофина перечислены в Таблице 1.

Таблица 1: Изоформы дистрофина

Название последовательности	Рег. номер нуклеиновой кислоты	SEQ ID NO: нуклеиновой кислоты	Рег. номер белка	SEQ ID NO: белка	Описание
Геномная последовательность DMD	NC_000023.11 (положения 31119219-33339609)	Нет	Нет	Нет	Последовательность из X хромосомы человека (положения Xp21.2 – p21.1) из версии GRCh38.p7 (GCF_000001405.33)
Изоформа дистрофина Dp427c	NM_000109.3	6	NP_000100.2	7	Вариант транскрипта: транскрипт Dp427c экспрессируется преимущественно в нейронах коры головного мозга и CA областях гиппокампа. Он использует уникальный промотор/экзон 1, расположенный приблизительно за 130 тпн перед промотором транскрипта Dp427m. Транскрипт включает общий экзон 2 транскрипта Dp427m и имеет такую же длину, 14 тпн. Изоформа Dp427c содержит уникальную N-концевую последовательность MED вместо последовательности MLWWEEVEDCY (SEQ ID NO: 949) изоформы Dp427m. Остальная часть изоформы Dp427c идентична изоформе Dp427m.

Изоформа дистрофина Dp427m	NM_004006.2	8	NP_003997.1	9	Вариант транскрипта: транскрипт Dp427m кодирует основной белок дистрофин, присутствующий в мышце. В результате использования альтернативного промотора, экзон 1 кодирует уникальную N-концевую MLWWEEVEDCY (SEQ ID NO: 949) ак последовательность.
Изоформа дистрофина Dp427p1	NM_004009.3	10	NP_004000.1	11	Вариант транскрипта: транскрипт Dp427p1 начинается с уникального промотора/экзона 1, расположенного на участке, который соотсветствует первому интрону транскрипта Dp427m. Транскрипт добавляет общий экзон 2 Dp427m и имеет такую же длину (14 тпн). Изоформа Dp427p1 заменяет MLWWEEVEDCY (SEQ ID NO: 949) в начаеле Dp427m уникальной N-концевой ак последовательностью MSEVSSD (SEQ ID NO: 950).
Изоформа дистрофина Dp260-1	NM_004011.3	12	NP_004002.2	13	Вариант транскрипта: транскрипт Dp260-1 использует экзоны 30-79 и начинается с последовательности промотора/экзона 1, расположенной в интроне 29 гена дистрофина. В результате Dp260-1 содержит

					95 тпн экзон 1, кодирующий уникальную N-концевую 16 ак последовательность MTEIILLIFFPAYFLN (SEQ ID NO: 951), которая заменяет амминокислоты 1-1357 полноразмерного продукта дистрофина (изоформа Dp427m).
Изоформа дистрофина Dp260-2	NM_004012.3	14	NP_004003.1	15	Вариант транскрипта: транскрипт Dp260-2 использует экзоны 30-79, начинающиеся с последовательности промотора/экзона 1, расположенной в интроне 29 гена дистрофина, т.е. подвергнут альтернативному сплайсингу и не содержит N-концевые аминокислоты 1-1357 полноразмерного дистрофина (изоформа Dp427m). Транскрипт Dp260-2 кодирует а уникальную N-концевую последовательность MSARKLRNLSYKK (SEQ ID NO: 952).
Изоформа дистрофина Dp140	NM_004013.2	16	NP_004004.1	17	Вариант транскрипта: Dp140 транскрипты используют экзоны 45-79, начинающиеся с промотора/экзона 1, расположенных в интроне 44. Транскрипты Dp140 имеют длинную (1 тпн) 5'-UTR, поскольку трансляция

					инициируется в экзоне 51 (соответствующую 2461 дистрофин). В дополнение к альтернативному промотору и экзону 1, дифференцированный сплайсинг экзонов 71-74 и 78 дает по меньшей мере пять Dp140 изоформ. Кроме того, данный транскрипт (Dp140) содержит все экзоны.
Изоформа дистрофина Dp116	NM_004014.2	18	NP_004005.1	19	Вариант транскрипта: транскрипт Dp116 используют экзоны 56-79, начинающиеся с промотора/экзона 1 в интроне 55. В результате изоформа Dp116 содержит уникальную N-концевую ак последовательность MLHRKTYHVK (SEQ ID NO: 953) вместо ак 1-2739 дистрофина. Дифференциальный сплайсинг дает несколько субтипов Dp116. Изоформа Dp116 также известна как S-дистрофин или апо-дистрофин-2.
Изоформа дистрофина Dp71	NM_004015.2	20	NP_004006.1	21	Вариант транскрипта: транскрипты Dp71 используют экзоны 63-79 с новым 80-100 нт экзоном, содержащим сайт инициации ATG для новой кодирующей последовательности из 17 нт. Короткая кодирующая последовательность находится в рамке считывания с

					последующей последовательностью дистрофина из экзона 63. Дифференциальный сплайсинг экзонов 71 и 78 дает по меньшей мере четыре изоформы Dp71. Из них данный транскрипт (Dp71) включает оба экзона 71 и 78.
Изоформа дистрофина Dp71b	NM_004016.2	22	NP_004007.1	23	Вариант транскрипта: транскрипты Dp71 используют экзоны 63-79 с новым 80-100 нт экзоном, содержащим сайт инициации ATG для новой кодирующей последовательности из 17 нт. Короткая кодирующая последовательность находится в рамке считывания с последующей последовательностью дистрофина из экзона 63. Дифференциальный сплайсинг экзонов 71 и 78 дает по меньшей мере четыре изоформы Dp71. Из них данный транскрипт (Dp71b) не содержит экзон 78 и кодирует белок с другим С-концом по сравнению с изоформами Dp71 и Dp71a.
Изоформа дистрофина Dp71a	NM_004017.2	24	NP_004008.1	25	Вариант транскрипта: транскрипты Dp71 используют экзоны 63-79 с новым 80-100 нт экзоном, содержащим сайт

					инициации ATG для новой кодирующей последовательности из 17 нт. Короткая кодирующая последовательность находится в рамке считывания с последующей последовательностью дистрофина из экзона 63. Дифференциальный сплайсинг экзонов 71 и 78 дает по меньшей мере четыре изоформы Dp71. Из них данный транскрипт (Dp71a) не содержит экзон 71.
Изоформа дистрофина Dp71ab	NM_004018.2	26	NP_004009.1	27	Вариант транскрипта: транскрипты Dp71 используют экзоны 63-79 с новым 80-100 нт экзоном, содержащим сайт инициации ATG для новой кодирующей последовательности из 17 нт. Короткая кодирующая последовательность находится в рамке считывания с последующей последовательностью дистрофина из экзона 63. Дифференциальный сплайсинг экзонов 71 и 78 дает по меньшей мере четыре изоформы Dp71. Из них данный транскрипт (Dp71ab) не содержит оба экзона 71 и 78 и кодирует белок с С-концом,

					как у изоформы Dp71b.
Изоформа дистрофина Dp40	NM_004019.2	28	NP_004010.1	29	Вариант транскрипта: транскрипт Dp40 использует экзоны 63-70. 5' UTR и первые кодируемые 7 ак идентичны транскрипту Dp71, однако стоп-кодон расположен в точке сплайсинга экзона/интрона 70. 3' UTR включает нт из интрона 70, который включает альтернативный сайт полиаденилирования. Изоформа Dp40 не имеет нормальной С-концевой области полноразмерного дистрофина (ак 3409-3685).
Изоформа дистрофина Dp140c	NM_004020.3	30	NP_004011.2	31	Вариант транскрипта: транскрипты Dp140 используют экзоны 45-79, начинающиеся с промотора/экзона 1, расположенных в интроне 44. Транскрипты Dp140 имеют длинную (1 тпн) 5' UTR, так как трансляция начинается с экзона 51 (соответствующего ак 2461 дистрофина). В дополнение к альтернативному промотору и экзону 1, дифференциальный сплайсинг экзонов 71-74 и 78 дает по меньшей мере пять изоформ Dp140. Из них данный транскрипт (Dp140c) не содержит экзоны 71-74.

Изоформа дистрофина Dp140b	NM_004021.2	32	NP_004012.1	33	Вариант транскрипта: транскрипты Dp140 используют экзоны 45-79, начинающиеся с промотора/экзона 1, расположенного в интроне 44. Транскрипты Dp140 имеют длинную (1 тпн) 5' UTR, так как трансляция начинается с экзона 51 (соответствующего ак 2461 дистрофина). В дополнение к альтернативному промотору и экзону 1, дифференциальный сплайсинг экзонов 71-74 и 78 дает по меньшей мере пять изоформ Dp140. Из них данный транскрипт (Dp140b) не содержит экзон 78 и кодирует белок с уникальным С-концом.
Изоформа дистрофина Dp140ab	NM_004022.2	34	NP_004013.1	35	Вариант транскрипта: транскрипты Dp140 используют экзоны 45-79, начинающиеся с промотора/экзона 1, расположенных в интроне 44. Транскрипты Dp140 имеют длинную (1 тпн) 5' UTR, так как трансляция начинается с экзона 51 (соответствующего ак 2461 дистрофина). В дополнение к альтернативному промотору и экзону 1, дифференциальный сплайсинг экзонов 71-74 и 78 дает по меньшей мере пять изоформ Dp140. Из них данный

					транскрипт (Dp140ab) не содержит экзоны 71 и 78 и кодирует белок с уникальным С-концом.
Изоформа дистрофина Dp140bc	NM_004023.2	36	NP_004014.1	37	Вариант транскрипта: транскрипты Dp140 используют экзоны 45-79, начинающиеся с промотора/экзона 1, расположенных в интроне 44. Транскрипты Dp140 имеют длинную (1 тпн) 5' UTR, так как трансляция начинается с экзона 51 (соответствующего ак 2461 дистрофина). В дополнение к альтернативному промотору и экзону 1, дифференциальный сплайсинг экзонов 71-74 и 78 дает по меньшей мере пять изоформ Dp140. Из них данный транскрипт (Dp140bc) не содержит экзоны 71-74 и 78 и кодирует белок с уникальным С-концом.
Изоформа дистрофина X2	XM_006724469.3	38	XP_006724532.1	39	
Изоформа дистрофина X5	XM_011545467.1	40	XP_011543769.1	41	
Изоформа дистрофина X6	XM_006724473.2	42	XP_006724536.1	43	
Изоформа дистрофина X8	XM_006724475.2	44	XP_006724538.1	45	
Изоформа дистрофина X4	XM_017029328.1	46	XP_016884817.1	47	

Изоформа дистрофина X1	XM_006724468.2	48	XP_006724531.1	49	
Изоформа дистрофина X13	XM_017029331.1	50	XP_016884820.1	51	
Изоформа дистрофина X3	XM_006724470.3	52	XP_006724533.1	53	
Изоформа дистрофина X7	XM_006724474.3	54	XP_006724537.1	55	
Изоформа дистрофина X9	XM_011545468.2	56	XP_011543770.1	57	
Изоформа дистрофина X11	XM_017029330.1	58	XP_016884819.1	59	
Изоформа дистрофина X10	XM_017029329.1	865	XP_016884818.1	866	
Изоформа дистрофина X12	XM_011545469.1	867	XP_011543771.1	868	

Мышиный белок дистрофин имеет следующую аминокислотную последовательность (регистрационный номер в Uniprot P11531, SEQ. ID NO: 869):

1 MWWVDCYRDV KKTTKWNASK GKHDNSDDGK RDGTGKKKGS TRVHANNVNK ARVKNNVDVN

61 GSTDVDGNHK TGWNHWVKNV MKTMAGTNSK SWVRSTRNYV NVNTSSWSDG ANAHSHRDDW

121 NSVVSHSATR HANAKCGKDD VATTYDKKSM YTSVVSAVMR TSSKVTRHHH MHYSTVSAGY

181 TSSSKRKSYA TAAYVATSDS TSYSHARDKS DSSMTVNDSY TAVSWSADTR AGSNDVVKHA

241 HGMMDTSHGV GNVGSVGKGK SDAVMNNSRW CRVASMKSKH KVMDNKKDDW TKTRTKKMGD

301 DKCVHKVDVR VNSTHMVVVV DSSGDHATAA KVGDRWANCR WTDRWVDKWH TCSTWSKDAM

361 KNTSGKDNMM SSHKSTKDKK KTMKSSNDSA KNKSVTKMWM NARWDNTKKS SASAVTTTST

421 TTVMTVTMVT TRMVKHAKKR TVDSRKRDVD THSWTRSAVS SAVYRKGNSD KVNAARKAKR

481 KDASRSAAVM ANGVNASRAS NSRWTCSRVN WYTNTYNMTT TANKTSTTST AKSKCKDVNR

541 SAKSKKGGMD ADVATNHNHD GVRAKKTDTM RYTMSSRTWS SKSVYSVTYM RGKASSKNGN

601 YSDTVKMAKK ASCKYSGHWK KSSVSCKHMN KRKNHKTKWM AVDVKWAGDA KKKCRVGDTS

661 NSVNGGKKSA ASRTRNTWDH CRVYTRKAKA GDKTVSKDSM HWMTAYRDYK TDTAVMKRAK

721 AKTKVKTTVN SVAHASAAKK TTTNYWCTRN GKCKTVWACW HSYKANKWNV KKTMNVAGTV

781 SNMHHSNNRA TTDGGVMDNT NSRWRHAVRK KSSAKSHSDK AAYTDKVDAA MAKSDTSHSM

841 KKHNGKDANR VSDVAKKDVS MKRKANRSKM DVKMHATKSV VSSHCVNYKS SVKSVMVKTG

901 RVKKTNKDRV TAKHYNGAKV TRKKCKSRKM RKMNVTWAAT DTTKRSAVGM SNDSVAWGKA

961 TKKKAHKSVT GSKMVGKKTV DKSNSNWAVT SRVWNYKHMT DNTKWHADDS

KKKKDKRKAM

1021 NDMRKVDSTR DAAKMANRGD HCRKVVSNRR AASHRKTGKA SKNSDKAGVN KDNKDMSDNG

1081 TVNRGDNRTD RKRKKTKHNA KDRSRRKKAS HWYYKRADDK CDKKASRDRK KDRKKKNAVR

1141 RAGSNGAAMA VTSKRWRSNA RRNAHTHTMV VTTDMDVSYV STYTSHASVD HNTCAKDDKS

1201 KNKDNSGRDH KKKTAASATS MKVKVAVAMD GKHRMYKRGR DRSVKWRHHY DMKVNWNVKK

1261 TNNWHAKYKW YKDGGRAVVR TNATGSSKTD VNKGSSRWHD CKARRKRKNV SRDNVWADNA

1321 TGDKVKARGK NTGGAVVSAR DKKKKTNWKV SRAKGVHKDR DHWSRNYNSA GDKVTVHGKA

1381 DVRSKGHYKK STVKRKDRSW AVNHRRTKDR AGSTTGASAS TVTVTSVVTK TVSKMSSVAA

1441 DNRAWTTDWS DRVKSRVMVG DDNMKKATDR RTAANKNKTS NARTTDRRWD VNRRNMKDST

1501 WAKAVGVRGK DSWKGHTVDA KKTTKAKDRR SVDVANDAKR DYSADDTRKV HMTNNTSWGN

1561 HKRVSAATHR DKSWTATTAN VDASRKKDSR GVRMKWDGTH TDYHNDNGKR SGSDARRDNM

1621 NKWSKKSNRS HASSDWKRHS VWKDDSRAGG DAVKNDHRAK RKTKVMSTTV RTGKYRRANV

1681 TRRKAVNAWD KNRSADWRKD ARAADDKRAV KGSWVGDDSD HKVKARGAKN VNRVNDAHTT

1741 GSYNSTDNTR WRVAVDRVRH AHRDGASHST SVGWRASNKV YYNHTTTCWD HKMTYSADNN

1801 VRSAYRTAMK RRKACDSSAA CDADHNKNDM DNCTTYDRHN NVNVCVDMCN WNVYDTGRTG

1861 RRVSKTGSCK AHDKYRYKVA SSTGCDRRGH DSRGVASGGS NSVRSCANNK AADWMRSMVW

1921 VHRVAAATAK HAKCNCKCGR YRSKHNYDCS CSGRVAKGHK MHYMVYCTTT SGDVRDAKVK

1981 NKRTKRYAKH RMGYVTVGDN MTVTNWVDSA ASSSHDDTHS RHYASRAMNS NGSYNDSSNS

2041 DDHHYCSNDS SRSASSRGRA DNRNAYDRKH HKGSSMMTSS RDAAAKRHKG

RARMDHNKSH

2101 RRAAKVNGTT VSSSTSRSDS SMRVVGSTSS MGDSDTSTGV MNNSSSRGRN AGKMRDTM

Дистрофин является важным компонентом в мышечной ткани, который придает структурную стабильность дистрогликановому комплексу (DGC) клеточной мембраны. Хотя мутацию могут нести представители обоих полов, скелетно-мышечная форма болезни редко поражает женщин.

Мутации различаются по природе и частоте. Большие генетические делеции обнаруживаются примерно в 60-70% случаев, большие дупликации обнаруживаются примерно в 10% случаев, а точечные мутанты или другие небольшие изменения составляют примерно 15-30% случаев. Bladen с соавт. (2015), который исследовал около 7000 мутаций, зарегистрировал в общей сложности 5682 крупных мутации (80% от общего числа мутаций), из которых 4894 (86%) были делециями (1 экзон или более), а 784 (14%) были дупликациями (1 экзон или более). Было 1445 небольших мутаций (меньше 1 экзона, 20% всех мутаций), из которых 358 (25%) были небольшими делециями и 132 (9%) небольшими вставками, тогда как 199 (14%) затрагивали сайты сплайсинга. Точечные мутации в общей сложности составляли 756 (52% малых мутаций), из которых 726 (50%) были нонсенс-мутациями, а 30 (2%) - миссенс-мутациями. Наконец, наблюдали 22 (0,3%) внутриинтронных мутации. Кроме того, в базе данных были идентифицированы мутации, при которых потенциально можно было бы эффективно применять новые генетические терапии для лечения МДД, включая терапии на основе сквозного прочтения стоп-кодонов (10% от общего числа мутаций) и терапию с пропуском экзонов (80% делеций и 55% от общего числа мутаций).

B. Симптомы

Симптомы обычно появляются у мальчиков в возрасте от 2 до 3 лет и могут быть заметны в грудном возрасте. Несмотря на то, что симптомы появляются только в первый год жизни, лабораторные исследования позволяют выявить детей с активной мутацией при рождении. Сначала наблюдается прогрессирующая слабость проксимальных мышц ног и таза, ассоциированная с потерей

мышечной массы. Со временем эта слабость распространяется на руки, шею и другие области. Ранние признаки могут включать псевдогипертрофию (увеличение икроножных и дельтовидных мышц), низкую выносливость и затруднения при стоянии без посторонней помощи или неспособность подниматься по лестнице. По мере прогрессирования состояния мышечная ткань истощается и в конечном итоге замещается жировой и фиброзной тканью (фиброз). К 10 годам могут потребоваться ортопедические корсеты, помогающие при ходьбе, однако большинство пациентов становятся зависимыми от инвалидной коляски к 12 годам. Более поздние симптомы могут включать аномальное развитие костей, которое приводит к деформации скелета, включая искривление позвоночника. Из-за прогрессирующего ухудшения функции мышц наблюдается потеря способности к передвижению, что в конечном итоге приводит к параличу. Интеллектуальные нарушения могут присутствовать или нет, но если они присутствуют, их прогрессирование не наблюдается по мере взросления ребенка. Средняя продолжительность жизни мужчин, страдающих МДД, составляет около 25 лет.

Основным симптомом миодистрофии Дюшенна, прогрессирующего нервномышечного расстройства, является мышечная слабость, ассоциированная с истощением мышц, при этом сначала поражаются произвольно сокращающиеся мышцы, в особенности мышцы бедра, тазовой области, бедер, плеч и икр. Мышечная слабость позже также появляется в руках, шее и других областях. Икры часто увеличены. Симптомы обычно появляются в возрасте до 6 лет и могут появляться в первый год жизни. Другие физические симптомы являются следующими:

1. Неуклюжая манера ходьбы, шага или бега (пациенты склонны идти на передней части стопы из-за повышенного тонуса икроножных мышц. Кроме того, ходьба на пальцах ног является компенсаторной адаптацией к слабости разгибающих мышц колена.)

2. Частые падения.

3. Усталость.

4. Трудность с моторными навыками (бег, прыжки на одной ноге, прыжки на двух ногах).

5. Люмбальный гиперлордоз, который может приводить к укорачиванию мышц-сгибателей бедра. Это влияет на общую осанку и манеру ходьбы, шага или бега.

6. Контрактуры мышц ахиллова сухожилия и подколенных сухожилий нарушают функциональность, потому что мышечные волокна укорачиваются и превращаются в фиброзную соединительную ткань.

7. Прогрессирующие затруднения при ходьбе.

8. Деформации мышечных волокон.

9. Псевдогипертрофия (увеличение) мышц языка и икроножных мышц. Мышечная ткань, в конечном счете, замещается жировой и соединительной тканью, отсюда термин псевдогипертрофия.

10. Более высокий риск нейроповеденческих нарушений (например, СДВГ), нарушения способности к обучению (дислексия) и непрогрессирующая слабость некоторых когнитивных навыков (в частности, кратковременной вербальной памяти), которые, как полагают, являются результатом отсутствия, или присутствия дисфункционального, дистрофина в головном мозге.

11. Возможная потеря способности ходить (обычно к возрасту 12).

12. Скелетные деформации (включая сколиоз в некоторых случаях).

13. Затруднение при поднятии из положения лежа или сидя.

Состояние часто можно наблюдать клинически с момента, когда пациент делает свои первые шаги, при этом способность ходить обычно полностью утрачивается в возрасте мальчика от 9 до 12 лет. Большинство мужчин, страдающих МДД, в возрасте 21 года по существу "парализованы ниже шеи". Мышечные истощения начинаются в ногах и тазу, затем прогрессируют до мышц плеч и шеи с последующей потерей функции мышц рук и дыхательных мышц. Увеличение икроножных мышц (псевдогипертрофия) является совершенно очевидным. Особенно распространена кардиомиопатия (дилатационная кардиомиопатия), однако развитие застойной сердечной недостаточности или аритмии (нерегулярного сокращения сердца) является случайным.

Положительный симптом Говерса отражает более тяжелое поражение мышц нижних конечностей. Ребенок помогает себе

подняться верхними конечностями: сначала поднимается, чтобы встать на руки и колени, а затем "перемещает" руки по ногам, чтобы встать вертикально. Больные дети обычно быстрее устают и в целом менее сильные, чем их сверстники. Уровень креатинкиназы (СРК-ММ) в крови сильно повышен. Электромиография (ЭМГ) показывает, что слабость вызвана разрушением мышечной ткани, а не повреждением нервов. Генетическое исследование может выявить генетические дефекты в гене Хр21. Биопсия мышц (иммуногистохимия или иммуноблоттинг) или генетическое исследование (анализ крови) подтверждают отсутствие дистрофина, хотя улучшения в генетическом тестировании часто делают это ненужным.

Больные МДД могут страдать от следующего:

1. Патология миокарда (кардиомиопатия).

2. Застойная сердечная недостаточность или нарушения сердечного ритма (аритмия).

3. Деформация груди и спины (сколиоз).

4. Увеличенные мышцы икр, ягодиц и плечей (в возрасте 4 или 5 лет). Эти мышцы в конечном счете замещаются жировой и соединительной тканью (псевдогипертрофия).

5. Потеря мышечной массы (атрофия).

6. Контрактуры мышц в пятках, ногах.

7. Деформации мышц.

8. Нарушения дыхания, включая пневмонию и попадание пищи или жидкости при глотании в легкие (на поздних стадиях болезни).

С. Этиология

Миодистрофия Дюшенна (МДД) вызвана мутацией гена дистрофина в локусе Хр21, расположенном на коротком плече Х-хромосомы. Дистрофин отвечает за соединение цитоскелета каждого мышечного волокна с подлежащей базальной пластинкой (внеклеточным матриксом) через белковый комплекс, содержащий множество субъединиц. Отсутствие дистрофина позволяет проникать избытку кальция в сарколемму (клеточную мембрану). Изменения в уровне кальция и сигнальных путях приводят к наполнению водой митохондрий, которые затем разрываются.

При дистрофии скелетных мышц митохондриальная дисфункция приводит к усилению индуцированных стрессом цитозольных сигналов

кальция и повышению образования индуцированных стрессом активных форм кислорода (АФК). В этом сложном каскадном процессе, который включает в себя несколько путей и до сих пор полностью не изучен, нарастающий окислительный стресс внутри клетки повреждает сарколемму и, в конечном счете, приводит к гибели клетки. Мышечные волокна подвергаются некрозу и, в итоге, замещаются жировой и соединительной тканью.

МДД наследуется по Х-сцепленному рецессивному типу. Женщины, как правило, являются носителями этой болезни, тогда как поражаются мужчины. Как правило, женщина-носитель не знает, что несет мутацию, пока у нее не рождается больной сын. Сын матери-носителя имеет 50% шанс унаследовать дефектный ген от своей матери. Дочь матери-носителя имеет 50% шанс быть носителем и 50% шанс иметь две нормальные копии гена. Во всех случаях здоровый отец передаст нормальную Y-хромосому своему сыну или нормальную Х-хромосому своей дочери. У женщин-носителей Х-сцепленной рецессивной патологии, такой как МДД, симптомы могут проявляться в зависимости от характера их инактивации.

Делеции экзонов, предшествующих экзону 51, в гене DMD человека, которые нарушают открытую рамку считывания (ORF) с соединением экзонов вне рамки считывания, представляет наиболее распространенный тип мутации DMD у человека. Пропуск экзона 51 может, в принципе, восстанавливать ORF DMD у 13% больных МДД с делециями экзона.

Миодистрофия Дюшена встречается у 1 из 5000 новорожденных мальчиков. Мутации в гене дистрофина могут либо наследоваться, либо возникать спонтанно при наследовании зародышевой линии. Таблица примерных, но не ограничивающих мутаций и соответствующих моделей представлена ниже:

Таблица 2: Мутации дистрофина и соответствующие модели на мышах

Делеция, небольшая вставка и нонсенс-мутации	**Название модели на мышах**
Экзон 44	ΔEx44
Экзон 52	ΔEx52
Экзон 43	ΔEx43

D. Диагноз

Консультация генетика рекомендуется лицам со случаями нарушения у прямых родственников. Миодистрофия Дюшенна может быть обнаружена с точностью около 95% с помощью генетических исследований, проводимых во время беременности.

Анализ ДНК. Мышечно-специфическая изоформа гена дистрофина состоит из 79 экзонов, и исследование и анализ ДНК обычно позволяют идентифицировать специфический тип мутации экзона или экзонов, которые были затронуты. Исследование ДНК подтверждает диагноз в большинстве случаев.

Биопсия мышц. Если исследование ДНК не позволяет обнаружить мутацию, может быть проведено исследование с биопсией мышц. Небольшой образец мышечной ткани извлекают (обычно с помощью скальпеля вместо иглы) и наносят краситель, который обнаруживает присутствие дистрофина. Полное отсутствие белка указывает на наличие заболевания.

За последние несколько лет были разработаны методы анализа ДНК, которые позволяют обнаруживать большее число из множества мутаций, которые вызывают данное заболевание, при этом биопсия мышц не требуется так часто для подтверждения наличия миодистрофии Дюшена.

Пренатальные исследования. Мутацию при МДД несет Х-сцепленный рецессивный ген. Мужчины имеют только одну Х-хромосому, поэтому одна копия мутантного гена послужит причиной МДД. Отцы не могут передавать Х-сцепленные признаки своим сыновьям, поэтому мутация передается от матери.

Если мать является носителем, и, следовательно, одна из ее двух Х-хромосом несет мутацию DMD, существует 50% вероятность того, что девочка унаследует такую мутацию на одной из ее двух Х-хромосом и станет носителем. Существует 50% вероятность, что мальчик унаследует эту мутацию в одной своей Х-хромосоме, и поэтому будет иметь МДД.

Пренатальные исследования могут показать, присутствуют ли у плода наиболее распространенные мутации. Существует множество мутаций, ответственных за МДД, причем некоторые пока не были определены, поэтому генетическое исследование работает только в

том случае, если у членов семьи с МДД есть мутация, которая уже была определена.

До инвазивного тестирования важно определить пол плода; тогда как мужчины иногда страдают этим Х-сцепленным заболеванием, у женщин МДД встречается крайне редко. Это может быть выполнено с помощью ультразвукового исследования на 16 неделе или позже путем исследования свободной ДНК плода. Проба ворсинчатого хориона (CVS) может быть выполнена на 11-14 неделе, при этом риск выкидыша составляет 1%. Амниоцентез может быть проведен после 15 недель, при этом риск выкидыша составляет 0,5%. Забор крови плода может быть сделан примерно на 18 неделе. Другим методом в случае неясных результатов генетического исследования является биопсия мышц плода.

E. Лечение

В настоящее время лечения МДД не существует, и регулирующие органы признают сохраняющуюся потребность медицины. Исследования Фазы 1-2а лечения с пропуском экзонов при некоторых мутациях показали остановку ухудшения и небольшие клинические улучшения при ходьбе. Лечение, как правило, направлено на контроль появления симптомов, чтобы максимально повысить качество жизни, и может включать в себя следующее:

1. Кортикостероиды, такие как преднизолон и дефлазакорт, повышают энергию и силу и отсрочивают нарастание тяжести некоторых симптомов.

2. Рандомизированные контрольные исследования показали, что бета-2-агонисты повышают мышечную силу, но не модифицируют прогрессирование болезни. Длительность последующих наблюдений для большей части РКИ на бета-2-агонистах составляет всего лишь приблизительно 12 месяцев и, следовательно, результаты нельзя экстраполировать с выходом за этот промежуток времени.

3. Рекомендуется легкая физическая активность, например, плавание. Отсутствие активности (например, постельный режим) может ухудшить заболевание мышц.

4. Физиотерапия полезна для поддержания силы, гибкости и функции мышц.

5. Ортопедические приспособления (такие как корсеты и

инвалидные коляски) могут улучшить подвижность и способность к самообслуживанию. Облегающие съемные ортезы, которые фиксируют лодыжку во время сна, могут отсрочить развитие контрактур.

6. Важна надлежащая респираторная поддержка по мере прогрессирования болезни.

Центрами по контролю и профилактике заболеваний США (CDC) были разработаны всесторонние междисциплинарные стандарты/руководства по лечению МДД, которые доступны на сайте www.treat-nmd.eu/dmd/care/diagnosis-management-DMD.

Как правило, МДД прогрессирует через пять стадий, как описано в публикациях Bushby et al., Lancet Neurol., 9(1):77-93 (2010) и Bushby et al., Lancet Neurol., 9(2):177-198 (2010), полностью включенных посредством отсылки. В течение предсимптомной стадии у пациентов обычно наблюдается задержка развития, но отсутствует нарушение походки. На ранней амбулаторной стадии пациенты, как правило, демонстрируют симптом Говерса, "утиную" походку и ходьбу на пальцах ног. На поздней амбулаторной стадии пациенты обычно демонстрируют возрастающие затруднения при походке и начинают терять способность подниматься по лестнице и с пола. На ранней неамбулаторной стадии пациенты, как правило, в течение некоторого времени могут самостоятельно двигаться, сохранять положение тела, при этом у больных может развиваться сколиоз. На поздней неамбулаторной стадии функция верхних конечностей и способность сохранять положение тела все более ограничены.

В некоторых вариантах осуществления лечение начинают на предсимптоматической стадии болезни. В некоторых вариантах осуществления лечение начинают на ранней амбулаторной стадии. В некоторых вариантах осуществления лечение начинают на поздней амбулаторной стадии. В вариантах осуществления лечение начинают в течение ранней неамбулаторной стадии. В вариантах осуществления лечение начинают в течение поздней неамбулаторной стадии.

1. Физиотерапия

Физиотерапевты стремятся предоставить пациентам возможность достичь своего максимального физического потенциала. Их цель

состоит в следующем:

1. свести к минимуму развитие контрактур и деформации посредством разработки программы растяжек и физических упражнений в соответствующих случаях,

2. предвидеть и минимизировать другие вторичные осложнения физического характера посредством рекомендации ортопедических устройств и надежного медицинского оборудования, и

3. контролировать дыхательную функцию и давать консультации по методикам выполнения дыхательных упражнений и методам очистки от выделений.

2. Поддержка дыхания

Современные "аппараты ИВЛ/респираторы с управляемым объемом", которые подают регулируемый объем (количество) воздуха пациенту при каждом вдохе, удобны при лечении лиц с нарушениями дыхания, связанными с миодистрофией. Аппарат ИВЛ может требовать инвазивной эндотрахеальной или трахеотомической трубки, через которую непосредственно подается воздух, однако для некоторых пациентов достаточна неинвазивная подача воздуха через дыхательную маску или ротовой мундштук. Аппараты поддержания положительного давления в дыхательных путях, особенно двухуровневые, иногда используются таким последним способом. Дыхательное оборудование может быть легко установлено в держателе для аппарата ИВЛ в нижней или задней части кресла-коляски с электроприводом с внешней батареей для мобильности.

Поддержку аппаратами ИВЛ можно начинать у подростков серднего или старшего возраста, когда может начинать нарушаться функция дыхательных мышц. Если жизненная емкость легких снизилась ниже 40% от нормы, аппарат ИВЛ/респиратор с управляемым объемом можно использовать во время сна, в то время, когда больной с наибольшей вероятностью подвергается недостаточной вентиляции легких ("гиповентиляции"). Гиповентиляция во время сна определяется подробным анамнезом нарушений сна с оксиметрическим исследованием и анализом газового состава капиллярной крови (см. Исследование функции легких). Устройство для облегчения кашля может помочь при удалении избытка слизи в легких путем гиперинфляции легких при

положительном давлении воздуха, а затем при отрицательном давлении, чтобы поднять слизь. Если жизненная емкость продолжает снижаться до уровня меньше 30 процентов от нормы, аппарат ИВЛ/респиратор с управляемым объемом может также потребоваться в течение дня для обеспечения дополнительной поддержки. Пациент постепенно будет увеличивать продолжительность применения аппарата ИВЛ/респиратора в течение дня по мере необходимости.

F. Прогноз

Миодистрофия Дюшенна – прогрессирующее заболевание, которое в конечном итоге поражает все произвольно сокращающиеся мышцы, и на более поздних стадиях вовлекает миокард и дыхательные мышцы. Ожидаемая продолжительность жизни в настоящее время составляет около 25 лет, однако она изменяется в зависимости от конкретного пациента. Последние достижения в области медицины продлевают продолжительность жизни больных. Muscular Dystrophy Campaign, ведущая британская благотворительная организация, занимающаяся всеми заболеваниями мышц, утверждает, что "при высоких стандартах медицинского обслуживания молодые мужчины с миодистрофией Дюшенна часто благополучно доживают до возраста 30 лет".

В редких случаях больные МДД доживали до сорока или пятидесяти лет при правильном расположении в инвалидных колясках и кроватях, поддержке аппаратами ИВЛ (с трахеостомией или мундштуком), очистке дыхательных путей и применении сердечных препаратов в случае необходимости. Раннее планирование требуемых программ поддержки для последующего ухода показало более длительную продолжительность жизни у людей, живущих с МДД.

Любопытно, что в модели миодистрофии Дюшенна на мышах mdx недостаток дистрофина связан с повышенным уровнем кальция и мионекрозом скелетных мышц. Внутренние мышцы гортани (ВМГ) защищены и не подвергаются мионекрозу. ВМГ имеют такие характеристики системы регуляции кальция, которые предполагают лучшую способность выдерживать изменения кальция по сравнению с другими мышцами, и это может давать механистическое объяснение их уникальных патофизиологических свойств. ВМГ могут способствовать разработке новых стратегий по профилактике и

лечению мышечного истощения при различных клинических сценариях.

II. Другие заболевания, при которых может быть эффективным редактирование генома

В дополнение к применению при лечении МДД, предложенный авторами настоящего изобретения метод с применением множества гидов может быть с преимуществами применен при других патологических состояниях, когда подобное вмешательство может привести к коррекции первопричинных генетических нарушений. Некоторые из них коротко описаны ниже.

Поясно-конечностная миодистрофия. Поясно-конечностная миодистрофия (ПКМД) или миодистрофия Эрба представляет собой генетически и клинически гетерогенную группу редких мышечных дистрофий. Она характеризуется прогрессирующим истощением мышц, которое затрагивает преимущественно мышцы бедра и плеча. ПКМД имеет аутосомный тип наследования и в настоящее время не имеет известного лечения. Симптомы больного поясно-конечностной миодистрофией (ПКМД) обычно включают затруднения при ходьбе, движении вверх и вниз по лестнице и при подъеме со стула. Также характерны затруднения при наклоне и регулярные падения. Затруднения при поднятии некоторых объектов также являются обычным проявлением ПКМД, а также затруднения при удерживании рук по сторонам или над головой. В конечном счете, нарушается способность ходить/бегать.

Заболевание неизбежно становится более тяжелым с течением времени, хотя у некоторых пациентов прогрессирование более быстрое, чем у других. В конечном счете, заболевание может поражать другие мышцы, такие как мышцы лица. Болезнь обычно приводит к зависимости от инвалидного кресла в течение нескольких лет после появления симптомов, однако в зависимости от конкретного пациента присутствует высокая вариабельность, и некоторые пациенты сохраняют двигательную функцию.

Слабость мышц обычно симметричная, проксимальная и прогрессирует медленно. В большинстве случаев боль при ПКМД отсутствует, и психическая функция не затронута. ПКМД может начинаться в детстве, подростковом, юношеском возрасте или даже позже, возраст начала заболевания обычно составляет от 10 до 30

лет. Оба пола поражаются одинаково, когда поясно-конечностная миодистрофия начинается в детстском возрасте, прогрессирование оказывается более быстрым, при этом заболевание приводит к более тяжелой инвалидизации. Если нарушение проявляется в подростковом или взрослом возрасте, заболевание, как правило, не такое тяжелое и прогрессирует медленнее. При первичном обследовании сенсорная нейропатия, вегетативные или висцеральные нарушения отсутствуют.

С точки зрения генетики ПКМД является наследственным заболеванием, хотя оно может наследоваться как доминантное или рецессивное генетическое нарушение. Результатом нарушения является то, что мышцы не могут правильно формировать некоторые белки, необходимые для нормальной функции мышц. Может быть затронуто несколько различных белков, при этом специфический белок, который отсутствует или является дефектным, определяет конкретный тип миодистрофии. Белки, затронутые при ПКМД, включают саркогликаны α, β, γ и δ. Саркогликанопатии могут поддаваться генотерапии.

Дисферлинопатия. Дисферлинопатия – аутосомно-рецессивное нервно-мышечное нарушение, вызванное дефицитом функционального белка дисферлина из-за мутаций в гене дисферлина. Дисферлинопатия характеризуется прогрессирующим истощением мышц и чаще всего клинически диагностируется как поясно-конечностная миодистрофия типа 2В (ПКМД2В) или миодистрофия Миоши 1 (МДМ1; тип дистальной мышечной дистрофии), в зависимости от начального профиля мышечного поражения на момент постановки диагноза. Дисферлинопатия – редкое заболевание, точная частота которого еще не определена.

Симптомы дисферлинопатии обычно проявляются в старшем юношеском возрасте между 16 и 25 годами и преимущественно затрагивают скелетные мышцы конечностей и тазового пояса (бедра и плечей), при этом жизненно важные мышцы, такие как сердце и диафрагма, в основном остаются незатронутыми. Большинство больных дисферлинопатией утрачивают способность передвигаться в течение 10-20 лет после постановки диагноза, однако продолжительность жизни нормальная. Возраст начала и

прогрессирование заболевания сильно различаются.

Хотя ПКМД2В и МДМ1 вызваны дефицитом дисферлина, диагноз ПКМД2В ставят, когда слабость первоначально присутствует в проксимальных мышцах (бедрах и плечах), тогда как диагноз МДМ1 ставят, когда слабость первоначально присутствует в дистальных мышцах (икрах и предплечьях). В обоих случаях слабость, в конечном счете, прогрессирует, затрагивая дистальные и проксимальные мышцы. ПКМД2В и МДМ1 очень сложно диагностировать, и пациентам часто много раз неправильно ставят диагноз, прежде чем у них успешно диагностируют дисферлинопатию.

Тогда как "дисферлинопатия" просто относится к отсутствию белка дисферлина, ПКМД2В и миопатия Миоши 1 становятся подтвержденным диагнозом, только когда могут быть обнаружены мутации в гене дисферлина. Важно, чтобы пациенты имели генетический диагноз для участия в клинических исследованиях и исследованиях непосредственно дисферлинопатии. Jain Foundation – некоммерческая семейная организация (www.jain-foundation.org), которая помогает пациентам с поясно-конечностной миодистрофией в получении генетического диагноза путем организации исследований и покрытия расходов.

Титинопатия. Титин, также известный как коннектин, представляет собой белок, который у людей кодируется геном *TTN*. Титин – гигантский белок длиной больше 1 мкм, который функционирует как молекулярная пружина, которая отвечает за пассивную эластичность мышцы. Он состоит из 244 отдельно свернутых белковых доменов, связанных неструктурированными пептидными последовательностями. Эти домены разворачиваются, когда белок растягивается, и снова сворачиваются, когда напряжение снимается.

Титин играет важную роль при сокращении волокон поперечнополосатых мышц. Он соединяет линию Z с линией M в саркомере. Белок способствует передаче усилия на линии Z и напряжению покоя в области I-диска. Он ограничивает амплитуду движения саркомера при растяжении, способствуя тем самым пассивной жесткости мышц. Изменения в последовательности титина в мышцах различных типов (например, сердечной или скелетной)

коррелировали с различиями в механических свойствах этих мышц.

Титин – большой распространенный белок поперечнополосатых мышц. Основные функции титина заключаются в стабилизации толстой нити, расположении ее по центру между тонкими нитями, в предотвращении чрезмерного растяжения саркомера и отводе саркомера, как пружины, после его растяжения. N-концевая область Z-диска и C-концевая область M-линии связываются с Z-линией и M-линией саркомера соответственно, при этом одна молекула титина охватывает половину длины саркомера. Титин также содержит участки связывания мышечно-ассоциированных белков, при этом он служит в качестве матрицы адгезии для сборки сократительного аппарата в мышечных клетках. Он также был идентифицирован как структурный белок для хромосом. Значительная вариабельность присутствует в областях I-диска, M-линии и Z-диска титина. Вариабельность в области I-диска обуславливает различия в эластичности различных изоформ титина и, следовательно, различия в эластичности различных типов мышц. Из многих идентифицированных вариантов титина пять вариантов описаны с полной доступной информацией по транскриптам.

Мутации в любом участке необычно длинной последовательности этого гена могут вызывать появление преждевременных стоп-кодонов или другие дефекты. Мутации титина ассоциированы с наследственной миопатией с ранней дыхательной недостаточностью, ранней миопатией с фатальной кардиомиопатией, миопатией центрального стержня с заболеванием сердца, центронуклеарной миопатией, поясно-конечностной миодистрофией типа 2J, наследственной дилатационной кардиомиопатией 9, гипертрофической кардиомиопатией и тибиальной миодистрофией. Другие исследования также указывают, что ни одну генетически сцепленную форму какой-либо дистрофии или миопатии, вызванную мутацией в гене TTN, нельзя безопасно исключить. Мутации с усечениями у пациентов с дилатационной кардиомиопатией чаще всего обнаруживаются в A-области; хотя можно предположить, что усечения перед I-областью полностью блокируют трансляцию A-области, альтернативный сплайсинг создает некоторые транскрипты, в которых не встречается преждевременный стоп-кодон, уменьшая его

воздействие. Аутоантитела к титину вырабатываются у пациентов с аутоиммунным заболеванием склеродермией.

Врожденная миопатия, "болезнь центрального стержня", вызванная мутацией рианодинового рецептора (мутацией с псевдоэкзонами). Болезнь центрального стержня (БЦС), также известная как миопатия центрального стержня, представляет собой аутосомно-доминантную врожденную миопатию (врожденное мышечное нарушение). Впервые она была описана Шаем и Мейджи в 1956 году. Она характеризуется появлением миофибрилл, заметных под микроскопом.

Симптомы БЦС могут быть разными, но обычно включают гипотонус (сниженный мышечный тонус) при рождении, легкую задержку развития ребенка (крайне изменчивую в зависимости от случая), слабость лицевых мышц и скелетные нарушения, такие как сколиоз и врожденный вывих бедра.

Симптомы могут присутствовать при рождении или могут появляться на любой стадии жизни. По-видимому, растет число людей, у которых не проявляется симптоматика до наступления зрелости или среднего возраста. Тогда как обычно прогрессирование отсутствует, опять же, по-видимому, растет число людей, которые действительно испытывают медленное, клинически значимое прогрессирование симптоматологии. Эти случаи могут быть гипотетически обусловлены большим количеством мутаций гена при нарушении функции рианодинового рецептора, и при продолжении исследования могут быть действительно обнаружены клинические варианты.

Диагноз ставят по комбинации типичных симптомов и признаков на биопсии (образце ткани) из мышцы. Такое название это заболевание получило из-за характерного вида биопсии при оптической микроскопии, где мышечные клетки имеют центральные области, в которых нет митохондрий и некоторых ферментов. Дыхательная недостаточность развивается в малом проценте случаев. Креатинкиназа и электромиография (ЭМГ) обычно соответствуют нормальным показателям.

Болезнь центрального стержня наследуется по аутосомно-доминантному типу. Большинство случаев характеризуется

присутствием явных мутаций в гене рианодинового рецептора 1 типа (RYR1), которые часто являются мутациями *de novo* (развившимися впервые). Люди с БЦС подвергаются риску злокачественной гипертермии (ЗГ) при получении общего наркоза.

Специального лечения не существует, однако избегают применения анестетиков, а родственники проходят скрининг на мутации RYR1, поскольку они могут сделать их восприимчивыми к ЗГ.

Миотоническая дистрофия. Миотоническая дистрофия – наследственное заболевание, проявляющееся в зрелом возрасте, которое влияет на функцию мышц. Симптомы включают постепенную потерю мышечной массы и нарастающую слабость. Мышцы часто сокращаются и не могут расслабиться. Другие симптомы могут включать катаракту, умственную отсталость и проблемы с сердечной проводимостью. У мужчин может наблюдаться раннее облысение и неспособность иметь детей.

Миотоническая дистрофия является аутосомно-доминантным нарушением, которое обычно наследуется от родителей. Существует два основных типа: 1 тип (DM1), обусловленный мутациями в гене DMPK, и 2 тип (DM2), вызванный мутациями в гене CNBP. Данное заболевание обычно прогрессирует с каждым поколением. Тип DM1 может быть заметен при рождении. DM2, как правило, менее выраженный. Это типы мышечной дистрофии. Диагноз подтверждается при генетическом исследовании.

Лечения не существует. Уход может включать ортопедические устройства или инвалидные коляски, кардиостимуляторы и неинвазивную вентиляцию с положительным давлением. Иногда могут применяться такие лекарственные средства как мексилетин или карбамазепин. Если возникает боль, ее можно лечить с помощью трициклических антидепрессантов и нестероидных противовоспалительных средств (НПВС).

Миотоническая дистрофия поражает более 1 на 8000 человек во всем мире. Хотя миотоническая дистрофия может возникать в любом возрасте, ее начало обычно наступает в 20–30 лет. Это наиболее распространенная форма мышечной дистрофии, которая начинается в зрелом возрасте. Впервые она была описана в 1909 году, при этом

первопричину 1 типа определили в 1992 году.

Появление симптомов и признаков значительно варьирует в зависимости от формы (DM1/DM2), по степени тяжести и даже необычным фенотипам DM2. Симптомы DM1 для DM2 включают нарушения исполнительной функции (например, организации, концентрации, подбора слов) и гиперсомнию. Нарушения проводимости чаще встречаются при DM1, чем при DM2, однако всем пациентам рекомендуется ежегодно проводить ЭКГ. Оба типа также связаны с инсулинорезистентностью. Миотоническая дистрофия может включать кортикальную катаракту с появлением голубоватого пятна или заднюю субкапсулярную катаракту.

DM2, как правило, менее тяжелая, чем DM1, и обычно меньше пациентов с DM2 нуждаются во вспомогательных устройствах, чем с DM1. Кроме того, тяжелая врожденная форма, которая поражает детей при DM1, при DM2 не обнаружена, при этом в медицинской литературе редко отмечают раннее появление симптомов у молодых людей.

Симптомы могут появляться в любое время, от грудного возраста до зрелости. МД вызывает общую слабость, обычно начинающуюся в мышцах рук, ног, шеи или лица. Она медленно прогрессирует, вовлекая другие группы мышц, в том числе сердце. МД также поражает ряд других систем органов.

Миотоническая дистрофия – наследственное заболевание, которое наследуется по аутосомно-доминантному типу и, таким образом, в среднем будет передано 50% потомков носителя. Миотоническая дистрофия является одним из нескольких известных нарушений с тринуклеотидными повторами. Некоторые области ДНК содержат повторяющиеся последовательности из двух или трех нуклеотидов.

Миотоническая дистрофия (МД) является наследственным заболеванием. Тяжелая форма МД, врожденная миотоническая дистрофия, может появиться у новорожденных детей матерей, которые имеют МД. Врожденная миотоническая дистрофия также может наследоваться через отцовский ген, хотя считается, что это происходит относительно редко. "Врожденная" означает, что заболевание присутствует с рождения.

При DM1 затронутый ген называется *DMPK* и кодирует протеинкиназу миотонической дистрофии, белок, экспрессируемый преимущественно в скелетной мышце. Ген расположен на длинном плече хромосомы 19.

При DM1 в гене *DMPK* присутствует последовательность триплетных повторов цитозина-гуанина-тимина (CTG). Число от 5 до 37 повторов считается нормальным, при этом считается, что у лиц с числом повторов от 38 до 49 присутствует премутация, и они подвергаются риску рождения детей с еще большим числом повторов и поэтому имеющих симптоматическое заболевание. Люди с числом повторов больше 50 почти всегда демонстрируют клинические симптомы за редкими известными исключениями. Более протяженные повторы обычно связаны с ранним проявлением и более тяжелым заболеванием.

Аллели DMPK, содержащие больше 37 повторов, являются нестабильными, и дополнительные тринуклеотидные повторы могут встраиваться во время деления клеток при митозе и мейозе. Как следствие, дети лиц с премутациями или мутациями наследуют аллели DMPK, которые длиннее, чем у их родителей, и, следовательно, с большей вероятностью будут иметь заболевание или демонстрировать более раннее проявление и большую тяжесть состояния, явление, известное как антиципация. Примечательно, что передача этого состояния по отцовской линии крайне редкая, вероятно, из-за давления отбора против сперматозоидов с большим числом повторов, при этом антиципация обычно менее тяжелая, чем в случаях материнского наследования.

РНК из области с увеличенным числом тринуклеотидных повторов образует внутринуклеоплазматические шпилечные петли из-за большого количества водородных связей между парами оснований C-G, при этом было продемонстрировано, что они приводят к секвестрации регулятора сплайсинга MBNL1 с образованием характерных очагов, при его мечении GFP и олигонуклеотидным зондом с красным флуоресцентным красителем Cyanine5 (Cy5).

DM2 вызвана дефектом гена ZNF9 на хромосоме 3. Данный дефект представляет собой повтор тетрануклеотида цитозин-цитозин-тимин-гуанозин (CCTG) в гене ZNF9. Поскольку это

заболевание включает повтор из четырех нуклеотидов, это не нарушение с тринуклеотидными повторами, а скорее нарушение с тетрануклеотидными повторами.

Увеличение числа повторов при DM2 намного больше, чем при DM1, в пределах от 75 до более чем 11000 повторений. В отличие от DM1, размер содержащего повторы участка ДНК при DM2, по-видимому, не влияет на возраст проявления или тяжесть заболевания. Антиципация, по-видимому, менее значима при DM2, и в последних обзорах сообщают лишь о легкой антиципации, характерной для DM2.

Атаксия Фридрейха. Атаксия Фридрейха – аутосомно-рецессивное наследственное заболевание, которое вызывает прогрессирующее поражение нервной системы. Она проявляется в начальных симптомах плохой координации, таких как нарушение походки; она также может привести к сколиозу, заболеваниям сердца и диабету, но не влияет на когнитивную функцию. Заболевание прогрессирует, пока для обеспечения подвижности не требуется инвалидное кресло. Число случаев в общей совокупности населения составляет примерно 1 на 50000.

Определенная генетическая мутация (увеличение числа триплетных GAA-повторов в интроне гена FXN) приводит к снижению экспрессии митохондриального белка фратаксина. С течением времени это нарушение вызывает вышеуказанное повреждение, а также частую усталость из-за воздействия на клеточный метаболизм.

Атаксия при атаксии Фридрейха является результатом дегенерации нервной ткани в спинном мозге, в частности сенсорных нейронов, необходимых (через контакты с мозжечком) для направления движения мышц рук и ног. Спинной мозг истончается, и нервные клетки теряют часть своей миелиновой оболочки (изолирующего покрытия на некоторых нервных клетках, которое помогает проводить нервные импульсы).

Симптомы обычно начинают проявляться в возрасте от 5 до 15 лет, но при поздней АФ могут появляться в 20–30 лет. Симптомы включают любую комбинацию, но не обязательно все, мышечной слабости рук и ног, потери координации, нарушения зрения,

нарушения слуха, невнятной речи, искривления позвоночника, высокого вогнутого свода стопы (деформации стопы), диабета (приблизительно у 20% больных атаксией Фридрейха развивается непереносимость углеводов, а у 10% развивается сахарный диабет) и нарушения сердечного ритма (например, фибрилляция предсердий и, как результат, тахикардия (учащенное сердцебиение) и гипертрофическая кардиомиопатия).

Заболевание проявляется в возрасте до 22 лет в прогрессирующей шаткой или неуверенной походке и частых падениях. Нижние конечности поражены более тяжело. Симптомы медленные и прогрессирующие. Длительное наблюдение показывает, что многие пациенты достигают плато в симптомах в старшем юношеском возрасте. В среднем после 10-15 лет заболевания, пациенты, как правило, прикованы к инвалидному креслу и нуждаются в помощи в повседневной деятельности.

При физикальном обследовании могут быть обнаружены следующие физические признаки: мозжечковые признаки, такие как нистагм, быстрые саккадические движения глаз, атаксия ствола, дизартрия, дисметрия; поражения нижних мотонейронов, такие как отсутствие глубоких сухожильных рефлексов; обычно встречаются пирамидальные признаки, такие как подошвенные разгибательные рефлексы и дистальная слабость; наблюдаются аспекты заднего столба, такие как потеря вибрационных и проприоцептивных ощущений. Нарушение сердечной деятельности наблюдаются у 91% пациентов, включая кардиомегалию (вплоть до дилатационной кардиомиопатии), симметричную гипертрофию, шумы в сердце и нарушения проводимости. Средний возраст смерти составляет 35 лет, при этом женщины имеют лучший прогноз со 100% 20-летней выживаемостью по сравнению с 63% у мужчин. В двадцати процентах случаев присутствует сахарный диабет.

Атаксия Фридрейха является аутосомно-рецессивным нарушением, которое возникает, когда ген FXN содержит амплифицированные интронные GAA-повторы (пример увеличения числа тринуклеотидных повторов). Ген FXN кодирует белок фратаксин. Повтор GAA вызывает снижение уровня фратаксина. Фратаксин является железосвязывающим белком, ответственным за образование

железосерных кластеров. Одним из результатов дефицита фратаксина является перегрузка митохондриального железа, которая может привести к повреждению многих белков. Точная роль фратаксина в нормальной физиологии остается невыясненной. Ген расположен на хромосоме 9.

Мутантный ген содержит увеличенное число повторов триплета GAA в первом интроне; в гинеалогических схемах нескольких семей были обнаружены точечные мутации. Поскольку дефект находится в интроне (который удаляется из мРНК транскрипта между транскрипцией и трансляцией), эта мутация не приводит к образованию аномальных форм белка фратаксина. Вместо этого мутация вызывает сайленсинг гена (то есть мутация снижает транскрипцию гена) в результате индукции структуры гетерохроматина аналогично эффекту положения мозаичного типа.

Помимо снижения экспрессии фратаксина, длинные участки GAA-повторов вызывают разрывы хромосом в исследованиях *in vivo* на дрожжах.

Болезнь Хантингтона. Болезнь Хантингтона (БХ), также известная как хорея Хантингтона, является наследственным заболеванием, которое приводит к гибели клеток головного мозга. Самыми ранними симптомами часто являются молозаметные нарушения настроения или умственных способностей. Часто наблюдается общее нарушение координации и неустойчивая походка. По мере прогрессирования болезни, нескоординированные, резкие движения тела становятся все более заметными. Физические способности постепенно ухудшаются, пока координированное движение не становится сложным, и пациент не теряет способность говорить. Умственные способности обычно снижаются с наступлением деменции. Специфические симптомы у разных людей несколько различаются. Симптомы обычно проявляются в возрасте от 30 до 50 лет, но могут начинаться в любом возрасте. Заболевание может развиваться во все более раннем возрасте в каждом последующем поколении. Около 8% случаев начинаются в возрасте до 20 лет и обычно имеют симптомы, более похожие на болезнь Паркинсона. Люди с БХ часто недооценивают степень своих проблем.

БХ обычно наследуется от родителей пациента в 10% случаев

вследствие новой мутации. Заболевание вызвано аутосомно-доминантной мутацией в любой из двух копий гена хантингтина у пациента. Это означает, что ребенок больного человека, как правило, имеет 50% вероятность наследования заболевания. Ген хантингтина содержит генетическую информацию белка, который также называется "хантингтин". Удлинение последовательности триплетных повторов CAG (цитозин-аденин-гуанин) в гене, кодирующем белок Хантингтин, приводит к аномальному белку, который постепенно повреждает клетки головного мозга посредством механизмов, которые не полностью ясны. Диагноз ставят по результатам генетического исследования, которое можно проводить в любое время независимо от присутствия симптомов. Этот факт вызывает некоторые этические споры: возраст, в котором человек считается достаточно зрелым, чтобы выбрать исследование; имеют ли родители право на исследование у своих детей; и решение о сохранении конфиденциальности и раскрытии результатов исследования.

Лечения БХ не существует. На поздних стадиях заболевания требуется постоянный уход. Лечение может облегчить некоторые симптомы, а в некоторой степени повысить качество жизни. Лучшим средством для устранения симптомов двигательных нарушений является тетрабеназин. БХ поражает около 4-15 человек на 100000 человек европейского происхождения. Это заболевание редко встречается среди японцев и с неизвестной частотой встречается в Африке. Болезнь поражает мужчин и женщин в одинаковой степени. Такие осложнения, как пневмония, болезни сердца и физические травмы в результате падений, сокращают продолжительность жизни. Самоубийство является причиной смерти примерно в 9% случаев. Смерть обычно наступает через пятнадцать-двадцать лет после первого обнаружения болезни.

Симптомы болезни Хантингтона чаще всего становятся заметными в возрасте от 35 до 44 лет, но они могут начинаться в любом возрасте, от младенчества до старости. На ранних стадиях происходят незначительные изменения личности, когнитивной функции и физических навыков. Обычно первыми становятся заметны физические симптомы, поскольку когнитивные и поведенческие

симптомы, как правило, недостаточно серьезны, чтобы их можно было самостоятельно распознать на более ранних стадиях. Почти у каждого пациента с болезнью Хантингтона в конечном итоге проявляются сходные физические симптомы, но начало, прогрессирование и степень когнитивных и поведенческих симптомов у разных людей значительно различаются.

Наиболее характерными начальными физическими симптомами являются резкие, случайные и неконтролируемые движения, называемые хореей. Первоначально хорея может проявляться как общее беспокойство, небольшие, непреднамеренно начатые или незавершенные движения, отсутствие координации или замедленные саккадические движения глаз. Эти незначительные моторные нарушения обычно предшествуют более выраженным признакам моторной дисфункции в течение по меньшей мере трех лет. Четкое проявление таких симптомов, как ригидность, корчащиеся движения или ненормальное положение тела, проявляется по мере прогрессирования нарушения. Это признаки поражения системы головного мозга, которая отвечает за движение. Психомоторные функции становятся все более нарушенными, при этом любое действие, которое требует мышечного контроля, нарушается. Общими последствиями этого является физическая нестабильность, ненормальное выражение лица и трудности при жевании, глотании, а также нарушения речи. Нарушение питания обычно вызывает потерю веса и может привести к голоданию. Нарушения сна также являются связанными симптомами. Ювенильная БХ отличается от этих симптомов тем, что оно обычно прогрессирует быстрее, при этом хорея кратковременная, если вообще присутствует, а ригидность является преобладающим симптомом. Приступы также являются частым симптомом этой формы БХ.

Когнитивные способности постепенно ухудшаются. Особенно затронуты исполнительные функции, которые включают планирование, когнитивную гибкость, абстрактное мышление, усвоение правил, инициирование соответствующих действий и запрещение неадекватных действий. По мере прогрессирования заболевания появляются нарушения памяти. Сообщаемые нарушения варьируют от нарушения кратковременной памяти до нарушения долговременной памяти, в том

числе эпизодической (памяти о своей жизни), процедурной (памяти о выполнении различных действий) и рабочей памяти. Когнитивные нарушения нарастают с течением времени и, в конечном итоге, приводят к деменции. Этот тип нарушений был назван синдромом подкорковой деменции, чтобы отличать его от типичных эффектов кортикальных деменций, например, болезни Альцгеймера.

Сообщаемыми нервно-психиатрическими проявлениями являются беспокойство, депрессия, сниженное проявление эмоций (уплощенный аффект), эгоцентризм, агрессия и компульсивное поведение, причем последнее может вызывать или усугублять зависимости, в том числе алкоголизм, азартные игры и гиперсексуальность. Также отмечали трудности в распознавании негативных выражений других людей. Частота этих симптомов в разных исследованиях сильно различается, с оцениваемыми процентными показателями проявлений психических нарушений в течение жизни от 33% до 76%. Для многих пациентов и членов их семей эти симптомы являются одними из наиболее тяжелых аспектов заболевания, часто влияющими на повседневную деятельность и являющимися причиной для помещения в специализированные учреждения. Суицидальные мысли и попытки самоубийства встречаются чаще, чем в общей совокупности населения. Часто у больных снижается критическое отношение к хорее, когнитивным и эмоциональным нарушениям.

Мутантный хантингтин экспрессируется во всем организме и ассоциирован с патологиями в периферических тканях, которые непосредственно вызваны такой экспрессией вне головного мозга. Такие патологии включают атрофию мышц, сердечную недостаточность, нарушение толерантности к глюкозе, потерю веса, остеопороз и тестикулярную атрофию.

У всех людей есть две копии гена хантингтина (HTT), который кодирует белок хантингтин (HTT). Ген также называют HD и IT15, что означает 'интересующий транскрипт 15'. Часть этого гена представляет собой повторяющийся участок, называемый тринуклеотидным повтором, который варьирует по длине у различных людей и может изменяться по длине у разных поколений. Если повтор присутствует в нормальном гене, динамическая мутация может увеличить количество повторов и привести к дефектному

гену. Когда длина этого содержащего повторы участка достигает определенного порога, образуется измененная форма белка, называемого мутантным белком хантингтином (mHTT). Отличающиеся функции этих белков являются причиной патологических изменений, которые в свою очередь вызывают симптомы заболевания. Мутация при болезни Хантингтона генетически доминантная и почти полностью пенетрантная: заболевание вызывает мутация в любом из аллелей HTT человека. Она не наследуется в соответствии с полом, однако длина содержащего повторы участка гена и, следовательно, тяжесть заболевания может находиться под влиянием пола затронутого родителя.

БХ является одним из нескольких нарушений с тринуклеотидными повторами, которые вызваны превышением границ диапазона нормальных значений длиной содержащего повторы участка гена. Ген HTT расположен на коротком плече хромосомы 4 в локусе 4p16.3. HTT содержит последовательность из трех оснований ДНК – цитозина-аденина-гуанина (CAG) – повторяющуюся много раз (т.е. ...CAGCAGCAG...), известную как тринуклеотидный повтор. CAG – 3-буквенный генетический код (кодон) аминокислоты глутамина, таким образом, серия таких повторов приводит к образованию глутаминовой цепи, известной как полиглутаминовый тракт (или поли-Q тракт), и повторяющейся части гена, поли-Q области.

Как правило, люди имеют меньше 36 повторяющихся глутаминов в поли-Q области, что приводит к продукции цитоплазматического белка хантингтина. Однако последовательность из 36 или более глутаминов приводит к получению белка, который имеет другие свойства. Эта измененная форма, называемая мутантным хантингтином (mHTT), увеличивает скорость распада некоторых типов нейронов. Области мозга имеют различное количество и зависимость от этих типов нейронов и поражаются в зависимости от этого. Как правило, количество повторов CAG связано с тем, насколько сильно затронут этот процесс, и ответственно приблизительно за 60% возрастной вариации проявления симптомов. Остальная вариация связана с окружением и другими генами, которые модифицируют механизм БХ. 36-39 повторов приводят к форме болезни с пониженной пенетрантностью, с гораздо более

поздним началом и более медленным прогрессированием симптомов. В некоторых случаях начало может быть настолько поздним, что симптомы остаются незамеченными. При очень большом количестве повторов БХ имеет полную пенетрантность и может проявляться в возрасте до 20 лет, когда ее называют ювенильной БХ, акинетико-ригидной формой или Вестфальским вариантом БХ. Эта форма составляет примерно 7% носителей БХ.

Атаксия телеангиэктазия. Атаксия-телеангиэктазия (АТ или А-Т), также называемая синдромом атаксии-телеангиэктазии или синдромом Луи-Бар, является редким, нейродегенеративным, аутосомно-рецессивным заболеванием, вызывающим тяжелую инвалидность. Атаксия относится к плохой координации, а телеангиэктазия к расширению мелких кровеносных сосудов, и являются характерными признаками заболевания.

При А-Т страдают различные части организма. Она поражает некоторые области головного мозга, включая мозжечок, вызывая затруднения при движении и координации. Она ослабляет иммунную систему, вызывая предрасположенность к инфекции. Она препятствует репарации поврежденной ДНК, повышая риск развития рака.

Симптомы чаще всего проявляются впервые в раннем детстве (с года до трех лет), когда дети начинают ходить. Хотя они обычно начинают ходить в нормальном возрасте, они колеблются или покачиваются при ходьбе, стоя или сидя, и могут выглядеть почти так же, как будто они пьяные. В старшем дошкольном и младшем школьном возрасте у них возникают трудности с естественным перемещением глаз из одного участка в другой (глазодвигательная апраксия). У них развивается невнятная или искаженная речь и затруднения при глотании. У некоторых повышено число случаев инфекций дыхательных путей (ушных инфекций, синусита, бронхита и пневмонии). Поскольку не все дети развиваются одинаково или с одинаковой скоростью, может пройти несколько лет, прежде чем будет правильно установлен диагноз А-Т. Большинство детей с А-Т имеют стабильные неврологические симптомы в течение первых 4-5 лет жизни, но в начальной школе начинают проявляться все более серьезные нарушения.

А-Т вызван дефектом в гене АТМ, который отвечает за управление реакцией клеток на различные формы стресса, включая разрывы двухцепочечной ДНК. Проще говоря, белок, продуцируемый геном АТМ, распознает разрыв ДНК, рекрутирует другие белки, чтобы устранить разрыв, и останавливает в клетке синтез новой ДНК до завершения репарации.

Существует значительное разнообразие тяжести признаков А-Т у пораженных лиц, и в различном возрасте. Следующие симптомы или нарушения являются частыми или важными признаками А-Т

Атаксия (затруднение при контроле движения), которая проявляется в раннем возрасте, но усугубляется в школе в предподростковом возрасте.

Глазодвигательная апраксия (затруднение при координации движений головы и глаз в момент перевода взгляда с одного места в другое).

Непроизвольные движения.

Телеангиэктазия (расширенные кровеносные сосуды) на белках (склерах) глаз, напоминающее воспаление. Она не выражена в грудном возрасте и может появиться в 5-8 лет. Телеангиэктазия также может появиться на открытых для солнечного света участках кожи.

Частые инфекции, в особенночти ушей, носовых пазух и легких.

Увеличенная частота случаев рака (прежде всего, но не только, лимфом и лейкозов).

Позднее начало или неполное пубертатное развитие и очень ранняя менопауза.

Замедленное развитие (вес и/или рост).

Слюнотечение, особенно у маленьких детей, когда они устают или концентрируются на каких-либо действиях.

Дизартрия (невнятные, медленные или искаженные звуки речи).

Диабет в пубертатном периоде или позже.

Преждевременные изменения волос и кожи.

У многих детей первоначально неправильно диагностируют атаксический церебральный паралич. Диагноз А-Т может быть поставлен только в дошкольном возрасте, когда появляются или

усиливаются неврологические симптомы нарушения походки, координации рук, речи и движения глаз, а также телеангиэктазия. Поскольку А-Т встречается очень редко, врачи могут быть не знакомы с симптомами или методами постановки диагноза. Позднее проявление телеангиэктазии может стать препятствием для постановки диагноза. Может потребоваться некоторое время, прежде чем врачи начнут допускать вероятность А-Т из-за ранней стабильности симптомов и признаков.

А-Т вызван мутациями в гене АТМ (АТМ серин/треониновая киназа или атаксия телеангиэктазия мутантном), который был клонирован в 1995 году. АТМ расположен на хромосоме 11 (11q22.3) человека и состоит из 69 экзонов, расположенных на 150 тпн сементе геномной ДНК.

Тип наследования А-Т аутосомно-рецессивный. Каждый родитель является носителем, что означает то, что у них есть одна нормальная копия гена А-Т (АТМ) и одна копия, которая содержит мутацию. А-Т возникает, если ребенок наследует мутантный ген А-Т от каждого родителя, поэтому в семье с двумя родителями-носителями существует 1 шанс из 4, что ребенок, рожденный от родителей, будет иметь это нарушение. Пренатальная диагностика (и выявление носителей) могут проводить в семьях, если были обнаружены ошибки (мутации) в двух генах АТМ у больного ребенка. Процесс такой диагностики может быть сложным и, поскольку он требует много времени, должен быть спланирован до зачатия.

Поиск мутаций в гене АТМ не являющегося родственником лица (например, супруга известного носителя А-Т) представляет значительную сложность. Гены часто имеют разные варианты последовательности (полиморфизмы), которые не влияют на функцию. В таком большом гене, как АТМ, такие варианты часто могут возникать, и врачи не всегда могут предсказать, будет ли конкретный вариант вызывать заболевание или не вызывать его. Консультации генетика могут помочь членам семьи пациента с А-Т понять, что можно или нельзя исследовать, и как следует интерпретировать результаты исследования.

Носители А-Т, такие как родители пациента с А-Т, имеют одну мутантную копию гена АТМ и одну нормальную копию. В целом они

здоровы, но у женщин повышен риск развития рака молочной железы. Эти данные были подтверждены различными способами и являются предметом текущих исследований. Рекомендуется стандартное наблюдение (включая ежемесячные самообследования молочных желез и маммографию согласно стандартному для возраста графику), если только не показаны дополнительные тесты, поскольку пациент имеет другие факторы риска (например, случаи рака молочной железы у прямых родственников).

Нейрофиброматоз 1 и 2 типа. Нейрофиброматоз (НФ) – это группа из трех состояний, при которых опухоли растут в нервной системе. Существует три основных типа: нейрофиброматоз 1-го типа (НФ1), нейрофиброматоз 2-го типа (НФ2) и шванноматоз. Симптомы НФ1 включают светло-коричневые пятна на коже, веснушки в подмышечной впадине и паху, небольшие выпуклости в нервах и сколиоз. При НФ2 возможны потеря слуха, катаракта в раннем возрасте, нарушения равновесия, пигментные пятна на коже и истощение мышц. Опухоли, как правило, являются доброкачественными.

Причиной является генетическая мутация в некоторых генах. В половине случаев они наследуются от родителей пациента, тогда как в остальных случаях они возникают на ранних этапах развития. Опухоли вовлекают поддерживающие клетки в нервной системе, а не нейроны. При НФ1 опухоли представляют собой нейрофибромы (опухоли периферических нервов), тогда как при NF2 и шванноматозе чаще встречаются опухоли из шванновских клеток. Диагноз обычно основан на признаках и симптомах и иногда подтверждается генетическим исследованием.

Известных методов профилактики или лечения не существует. Может быть проведена операция для удаления опухолей, которые вызывают проблемы или стали злокачественными. Лучевая терапия и химиотерапия также могут применяться в случае возникновения рака. Кохлеарный имплантат или слуховой стволомозговой имплантат может помочь пациентам с потерей слуха.

В США приблизительно 1 из 3500 человек имеет НФ1 и 1 из 25000 человек имеет НФ2. Частота возникновения заболевания у мужчин и женщины одинакова. При НФ1 симптомы часто присутствуют

при рождении, а в ином случае развиваются до 10-летнего возраста. Хотя состояние обычно ухудшается с течением времени, большинство людей с НФ1 имеют нормальную продолжительность жизни. При НФ2 симптомы могут не проявляться до старшего юношеского возраста. НФ2 повышает риск преждевременной смерти. Описания состояния известны еще с 1-го века.

Нейрофиброматоз (НФ1) в раннем возрасте может вызвать проблемы с обучением и поведением - около 60% детей с НФ1 имеют легкую форму затруднений в школе. Что касается признаков заболевания, пациент может иметь следующее: шесть или более светло-коричневых пятен на коже ("пятна цвета кофе с молоком"), по меньшей мере две нейрофибромы, по меньшей мере два нароста на радужной оболочке глаза и аномалии позвоночника (сколиоз).

Нейрофиброматоз является аутосомно-доминантным нарушением, что означает, что для его развития требуется только одна копия поврежденного гена. Поэтому, если только один из родителей имеет нейрофиброматоз, его или ее дети также имеют 50%-ную вероятность развития данного заболевания. У больного ребенка может быть легкая форма НФ1, даже если он унаследован от родителей с тяжелой формой нарушения. Существует два типа нейрофиброматоза. Нейрофиброматоз I типа характеризуется разрастанием нервной ткани с образованием опухолей (нейрофибром), которые могут быть доброкачественными и могут приводить к тяжелому повреждению в результате сдавливания нервов и других тканей. Нейрофиброматоз II типа проявляется наличием двусторонних акустических невром (опухолей вестибулокохлеарного нерва или черепного нерва 8 (CN VIII), также известных как шванномы), часто приводящих к потере слуха.

Шванноматоз, при котором болезненные шванномы развиваются на позвоночных и периферических нервах.

Мутация CEP290 при врожденном амаврозе Лебера. Центросомальный 290 кДа белок представляет собой белок, который у человека кодируется геном CEP290. CEP290 расположен на плече Q хромосомы 12. Ген CEP290 является центросомальным белком, который играет важную роль в развитии центросом и ресничек. Этот ген жизненно важен для формирования первичной реснички,

маленького антенноподобного выступа клеточной мембраны, который играет важную роль в фоторецепторах в задней части сетчатки (которые чувствительны к свету и цвету), а также в почках, головном мозге и многих других органах тела. Нокдаун транскрипта гена CEP290 приводил к существенному подавлению цилиогенеза в клетках пигментного эпителия сетчатки в культуре, что доказывает, насколько важен CEP290 для формирования ресничек.

На молекулярном уровне было показано, что CEP290 играет важную регуляторную и структурную роль в формировании первичных ресничек. В недавних исследованиях предположили, что CEP290 является связывающим микротрубочки и мембраны белком, который может служить в качестве структурной связи между ядром микротрубочек реснички и налегающей мембраной реснички. Было показано, что разрушение домена связывания микротрубочек CEP290 в модели заболевания CEP290 на мышах rd16 приводит к быстрой и значительной дегенерации сетчатки, что демонстрирует важность связывания CEP290 с микротрубочками при заболевании. Функция CEP290 при цилиогенезе ингибируется как ауторегуляторными доменами, обнаруженными на обоих концах белка CEP290, так и при взаимодействии CEP290 с ингибирующим белком CP110.

Открытие гена CEP290 привело исследователей к обнаружению другого гена, важного для функции сетчатки, LCA5. Клинические исследования, включающие замену этих двух генов, были начаты в Филадельфии, где исследователи надеются, что врожденный амавроз Лебера когда-нибудь удастся вылечить.

Этот ген кодирует белок с 13 предполагаемыми суперспиральными доменами, областью гомологии с АТФазами хромосомной сегрегации SMC, шестью мотивами KID, тремя доменами гомологии с тропомиозином и мотивом А сайта связывания АТФ/ГТФ. Белок локализуется в центросоме и ресничках и содержит сайты N-гликозилирования, сульфатирования тирозина, фосфорилирования, N-миристоилирования и амидирования.

Мутации в этом гене были связаны с синдромом Жубер и нефронофтизом, а недавно также с частой формой врожденного амавроза Лебера, называемой LCA10. Присутствие антител против этого белка связано с несколькими формами рака.

Мутация в этом гене приводит к слепоте у младенцев и детей, заболеванию, известному как врожденный амавроз Лебера. На сегодняшний день 35 различных мутаций в CEP290 ответственны за возникновение LCA. Другие мутации в CEP290 также были выявлены при возникновении синдрома Меккеля и синдрома Жубер, лишь некоторых из многих других синдромов. Дефектный ген CEP290 обычно является причиной этих нарушений из-за аномалий ресничек. Неизвестно, как одна мутация в гене может вызывать так много различных типов синдромов, многие из которых влияют на центральную нервную систему.

III. Системы CRISPR

A. CRISPR

CRISPR (регулярно расположенные кластерами короткие палиндромные повторы) являются локусами ДНК, содержащими короткие повторы последовательностей оснований. После каждого повтора расположены короткие сегменты "спейсерной ДНК", остающиеся после предыдущих контактов с вирусом. CRISPR обнаружены приблизительно в 40% секвенированных геномов эубактерий и 90% секвенированных археев. CRISPR часто ассоциированы с генами Cas, которые кодируют белки, связанные с CRISPR. Система CRISPR/Cas представляет собой прокариотическую иммунную систему, которая придает устойчивость к чужеродным генетическим элементам, таким как плазмиды и бактериофаги, и обеспечивает форму приобретенного иммунитета. Спейсеры CRISPR распознают и вызывают сайленсинг таких экзогенных генетических элементов, как РНКи в эукариотических организмах.

Повторы CRISPR имеют размер от 24 до 48 пар оснований. Они обычно показывают некоторую двойную симметрию, предполагающую образование вторичной структуры, такой как шпилька, но по существу не являются палиндромными. Повторы разделены спейсерами одинаковой длины. Некоторые спейсерные последовательности CRISPR полностью совпадают с последовательностями из плазмид и фагов, хотя некоторые спейсеры соответствуют геному прокариота (самонаправленные спейсеры). Новые спейсеры могут быстро добавляться в ответ на заражение фагом.

B. Нуклеазы Cas

CRISPR-ассоциированные (*cas*) гены часто ассоциированы с комплексами CRISPR повторов-спейсеров. По состоянию на 2013 год были описаны больше чем сорок различных семейств белков Cas. Из этих семейств белков наиболее распространенным среди различных систем CRISPR/Cas, по-видимому, является Cas1. Конкретные комбинации генов *cas* и повторяющихся структур использовали для определения 8 подтипов CRISPR (Ecoli, Ypest, Nmeni, Dvulg, Tneap, Hmari, Apern и Mtube), некоторые из которых связаны с дополнительным генным модулем, кодирующим ассоциированные с повторами загадочные белки (RAMP). В одном геноме могут присутствовать больше одного подтипа CRISPR. Спорадическое распределение подтипов CRISPR/Cas дает основание предположить, что система подвергается горизонтальному переносу генов в ходе эволюции микроорганизмов.

По-видимому, экзогенная ДНК процессируется белками, кодируемыми генами Cas, с образованием малых элементов (длиной ~30 пар оснований), которые затем каким-то образом встраиваются в локус CRISPR рядом с лидерной последовательностью. Молекулы РНК из локусов CRISPR конститутивно экспрессируются и процессируются белками Cas с получением малых РНК, состоящих из отдельных элементов последовательности экзогенного происхождения с фланкирующей последовательностью, содержащей повторы. РНК направляют другие белки Cas, вызывая сайленсинг экзогенных генетических элементов на уровне РНК или ДНК. Данные свидетельствуют о функциональном разнообразии подтипов CRISPR. Белки Cse (Cas субъединица *Ecoli*) (называемые CasA-E в *E.coli*) образуют функциональный комплекс, Cascade, который процессирует РНК транскрипты CRISPR с образованием спейсер-повторяющихся единиц, в которых сохраняется Cascade. В других прокариотах Cas6 процессирует транскрипты CRISPR. Интересно, что инактивация фага на основе CRISPR в *E.coli* требует Cascade и Cas3, но не Cas1 и Cas2. Белки Cmr (Cas RAMP модуль), обнаруженные у *Pyrococcus furiosus* и других прокариот, образуют функциональный комплекс с малыми РНК CRISPR, который распознает и расщепляет комплементарные РНК-мишени. РНК-направляемые ферменты CRISPR относятся к классу рестриктаз V типа.

Cas9 является нуклеазой, ферментом, специализированным для разрезания ДНК, с двумя активными разрезающими сайтами, по одному на каждой цепи двойной спирали. Команда продемонстрировала, что они могут удалить один или оба сайта, сохранив при этом способность Cas9 определять свою ДНК-мишень. Jinek et al. (2012) объединили tracrРНК и спейсерную РНК в молекулу "единой направляющей РНК", которая вместе с Cas9 может находить и разрезать требуемые ДНК-мишени, и такие синтетические направляющие РНК используются для редактирования генов.

Белки Cas9 распространены у патогенных и условно-патогенных бактерий. CRISPR/Cas-опосредованная регуляция генов может способствовать регуляции эндогенных бактериальных генов, особенно во время взаимодействия бактерий с эукариотическими организмами-хозяевами. Например, Cas-белок Cas9 из *Francisella novicida* использует уникальную, малую CRISPR/Cas-ассоциированную РНК (scaРНК) для репрессии эндогенного транскрипта, кодирующего бактериальный липопротеин, который важен для *F. novicida*, чтобы ослабить ответ хозяина и повысить вирулентность. Ванг с соавт. (Wang et al. (2013)) показали, что коинъекция мРНК и sgРНК Cas9 в зародышевую линию (зиготы) приводит к получению мышей с мутациями. Также предусмотрена доставка ДНК последовательностей Cas9.

Системы CRISPR/Cas подразделяются на три класса. Класс 1 использует несколько белков Cas вместе с молекулами CRISPR РНК (crРНК) для сборки функциональной эндонуклеазы. Системы CRISPR 2 класса используют один белок Cas с кРНК. Недавно Cpf1 был идентифицирован как система CRISPR/Cas V типа, II класса, содержащая белок из 1300 аминокислот. См. также публикацию патента США 2014/0068797, которая полностью включена посредством отсылки.

В некоторых вариантах осуществления композиции согласно настоящему изобретению включают малый вариант Cas9 из бактерии *Staphylococcus aureus* (рег. номер UniProt J7RUA5). Малый вариант Cas9 обеспечивает преимущества по сравнению с Cas9 дикого типа или полноразмерной Cas9. В некоторых вариантах осуществления Cas9 представляет собой spCas9 (AddGene).

C. Нуклеазы Cpf1

Регулярно расположенные кластерами короткие палиндромные повторы из *Prevotella* и *Francisella* 1 или CRISPR/Cpf1 представляютт собой технологию редактирования ДНК, которая обладает некоторыми сходствами с системой CRISPR/Cas9. Cpf1 является РНК-направляемой эндонуклеазой системы CRISPR/Cas II класса. Этот приобретенный иммунный механизм обнаружен у бактерий *Prevotella* и *Francisella*. Он предотвращает генетическое повреждение, вызываемое вирусами. Гены Cpf1 ассоциированы с локусом CRISPR, кодирующим эндонуклеазу, которая использует направляющую РНК для поиска и расщепления вирусной ДНК. Cpf1 является менее крупной и более простой эндонуклеазой, чем Cas9, преодолевая некоторые ограничения системы CRISPR/Cas9.

Cpf1 присутствует во многих видах бактерий. Окончательный вариант эндонуклеазы Cpf1, который был разработан в качестве инструмента для редактирования генома, был получен из одного из первых 16 видов, которые, как известно, несли ее.

В вариантах осуществления Cpf1 является ферментом Cpf1 из *Acidaminococcus* (вид BV3L6, рег. номер UniProt U2UMQ6; SEQ ID NO: 870), имеющим последовательность, представленную ниже:

1 MTQFEGFTNL YQVSKTLRFE LIPQGKTLKH IQEQGFIEED KARNDHYKEL KPIIDRIYKT

61 YADQCLQLVQ LDWENLSAAI DSYRKEKTEE TRNALIEEQA TYRNAIHDYF IGRTDNLTDA

121 INKRHAEIYK GLFKAELFNG KVLKQLGTVT TTEHENALLR SFDKFTTYFS GFYENRKNVF

181 SAEDISTAIP HRIVQDNFPK FKENCHIFTR LITAVPSLRE HFENVKKAIG IFVSTSIEEV

241 FSFPFYNQLL TQTQIDLYNQ LLGGISREAG TEKIKGLNEV LNLAIQKNDE TAHIIASLPH

301 RFIPLFKQIL SDRNTLSFIL EEFKSDEEVI QSFCKYKTLL RNENVLETAE ALFNELNSID

361 LTHIFISHKK LETISSALCD HWDTLRNALY ERRISELTGK ITKSAKEKVQ RSLKHEDINL

421 QEIISAAGKE LSEAFKQKTS EILSHAHAAL DQPLPTTLKK QEEKEILKSQ LDSLLGLYHL

481 LDWFAVDESN EVDPEFSARL TGIKLEMEPS LSFYNKARNY ATKKPYSVEK FKLNFQMPTL

541 ASGWDVNKEK NNGAILFVKN GLYYLGIMPK QKGRYKALSF EPTEKTSEGF DKMYYDYFPD

601 AAKMIPKCST QLKAVTAHFQ THTTPILLSN NFIEPLEITK EIYDLNNPEK EPKKFQTAYA

661 KKTGDQKGYR EALCKWIDFT RDFLSKYTKT TSIDLSSLRP SSQYKDLGEY YAELNPLLYH

721 ISFQRIAEKE IMDAVETGKL YLFQIYNKDF AKGHHGKPNL HTLYWTGLFS PENLAKTSIK

781 LNGQAELFYR PKSRMKRMAH RLGEKMLNKK LKDQKTPIPD TLYQELYDYV NHRLSHDLSD

841 EARALLPNVI TKEVSHEIIK DRRFTSDKFF FHVPITLNYQ AANSPSKFNQ RVNAYLKEHP

901 ETPIIGIDRG ERNLIYITVI DSTGKILEQR SLNTIQQFDY QKKLDNREKE RVAARQAWSV

961 VGTIKDLKQG YLSQVIHEIV DLMIHYQAVV VLENLNFGFK SKRTGIAEKA VYQQFEKMLI

1021 DKLNCLVLKD YPAEKVGGVL NPYQLTDQFT SFAKMGTQSG FLFYVPAPYT SKIDPLTGFV

1081 DPFVWKTIKN HESRKHFLEG FDFLHYDVKT GDFILHFKMN RNLSFQRGLP GFMPAWDIVF

1141 EKNETQFDAK GTPFIAGKRI VPVIENHRFT GRYRDLYPAN ELIALLEEKG IVFRDGSNIL

1201 PKLLENDDSH AIDTMVALIR SVLQMRNSNA ATGEDYINSP VRDLNGVCFD SRFQNPEWPM

1261 DADANGAYHI ALKGQLLLNH LKESKDLKLQ NGISNQDWLA YIQELRN

В некоторых вариантах осуществления Cpf1 является ферментом Cpf1 из *Lachnospiraceae* (вид ND2006, рег. номер UniProt A0A182DWE3; SEQ ID NO: 871), имеющим последовательность, представленную ниже:

1 AASKLEKFTN CYSLSKTLRF KAIPVGKTQE NIDNKRLLVE DEKRAEDYKG VKKLLDRYYL

61 SFINDVLHSI KLKNLNNYIS LFRKKTRTEK ENKELENLEI NLRKEIAKAF KGAAGYKSLF

121 KKDIIETILP EAADDKDEIA LVNSFNGFTT AFTGFFDNRE NMFSEEAKST

SIAFRCINEN

181 LTRYISNMDI FEKVDAIFDK HEVQEIKEKI LNSDYDVEDF FEGEFFNFVL TQEGIDVYNA

241 IIGGFVTESG EKIKGLNEYI NLYNAKTKQA LPKFKPLYKQ VLSDRESLSF YGEGYTSDEE

301 VLEVFRNTLN KNSEIFSSIK KLEKLFKNFD EYSSAGIFVK NGPAISTISK DIFGEWNLIR

361 DKWNAEYDDI HLKKKAVVTE KYEDDRRKSF KKIGSFSLEQ LQEYADADLS VVEKLKEIII

421 QKVDEIYKVY GSSEKLFDAD FVLEKSLKKN DAVVAIMKDL LDSVKSFENY IKAFFGEGKE

481 TNRDESFYGD FVLAYDILLK VDHIYDAIRN YVTQKPYSKD KFKLYFQNPQ FMGGWDKDKE

541 TDYRATILRY GSKYYLAIMD KKYAKCLQKI DKDDVNGNYE KINYKLLPGP NKMLPKVFFS

601 KKWMAYYNPS EDIQKIYKNG TFKKGDMFNL NDCHKLIDFF KDSISRYPKW SNAYDFNFSE

661 TEKYKDIAGF YREVEEQGYK VSFESASKKE VDKLVEEGKL YMFQIYNKDF SDKSHGTPNL

721 HTMYFKLLFD ENNHGQIRLS GGAELFMRRA SLKKEELVVH PANSPIANKN PDNPKKTTTL

781 SYDVYKDKRF SEDQYELHIP IAINKCPKNI FKINTEVRVL LKHDDNPYVI GIDRGERNLL

841 YIVVVDGKGN IVEQYSLNEI INNFNGIRIK TDYHSLLDKK EKERFEARQN WTSIENIKEL

901 KAGYISQVVH KICELVEKYD AVIALEDLNS GFKNSRVKVE KQVYQKFEKM LIDKLNYMVD

961 KKSNPCATGG ALKGYQITNK FESFKSMSTQ NGFIFYIPAW LTSKIDPSTG FVNLLKTKYT

1021 SIADSKKFIS SFDRIMYVPE EDLFEFALDY KNFSRTDADY IKKWKLYSYG NRIRIFAAAK

1081 KNNVFAWEEV CLTSAYKELF NKYGINYQQG DIRALLCEQS DKAFYSSFMA LMSLMLQMRN

1141 SITGRTDVDF LISPVKNSDG IFYDSRNYEA QENAILPKNA DANGAYNIAR KVLWAIGQFK

1201 KAEDEKLDKV KIAISNKEWL EYAQTSVK

В некоторых вариантах осуществления Cpf1 является кодон-оптимизированной для экспрессии в клетках млекопитающих. В некоторых вариантах осуществления Cpf1 является кодон-оптимизированной для экспрессии в человеческих клетках или мышиных клетках.

Локус Cpf1 содержит смешанный альфа/бета домен, RuvC-I, после которого расположена спиральная область, RuvC-II, и домен, подобный цинковым пальцам. Белок Cpf1 имеет RuvC-подобный эндонуклеазный домен, который подобен домену RuvC Cas9. Кроме того, Cpf1 не имеет HNH эндонуклеазного домена, и N-концевая область Cpf1 не имеет альфа-спирального распознающего сегмента Cas9.

Архитектура домена Cpf1 CRISPR-Cas показывает, что Cpf1 является функционально уникальным, относящимся к системе CRISPR V типа, 2 класса. Локусы Cpf1 кодируют белки Cas1, Cas2 и Cas4, более сходные с I и III типами, чем с системами II типа. Поиск в базе данных указывает на распространение белков семейства Cpf1 во многих видах бактерий.

Функциональный Cpf1 не требует tracrРНК, поэтому требуется только crРНК. Это удобно для редактирования генома, поскольку Cpf1 не только меньше, чем Cas9, но также имеет менее крупную молекулу sgРНК (примерно в два раза меньше нуклеотидов, чем в Cas9).

Комплекс Cpf1-crРНК расщепляет целевую ДНК или РНК при идентификации примыкающего к протоспейсеру мотива 5'-YTN-3' (где "Y" - пиримидин, и "N" является любым нуклеиновым основанием) или 5'-TTN-3', в отличие от G-богатого PAM, с которым направленно взаимодействует Cas9. После идентификации PAM, Cpf1 создает двухцепочечный разрыв ДНК с подобными липкими концами с 4 или 5 выступающими нуклеотидами.

Система CRISPR/Cpf1 состоит из фермента Cpf1 и направляющей РНК, которая находит и помещает комплекс в нужный участок на двойной спирали, чтобы произвести расщепление ДНК-мишени. Активность систем CRISPR/Cpf1 включает три стадии:

Адаптация, в ходе которой белки Cas1 и Cas2 способствуют адаптации малых фрагментов ДНК в комплексе CRISPR;

Образование молекул crРНК: процессинг молекул пре-cr-Рнк с получением зрелых crРНК, требуемых для направления белка Cas; и

Интерференция, при которой Cpf1 связывается с crРНК, образуя бинарный комплекс, который обнаруживает и расщепляет целевую последовательность ДНК.

D. gРНК

Будучи РНК-направляемым белком, Cas9 требует короткой РНК для направления распознавания ДНК-мишеней. Хотя Cas9 преимущественно взаимодействует с последовательностями ДНК, содержащими PAM последовательность NGG, она может связываться и без протоспейсерной мишени. Однако комплекс Cas9-gРНК требует близкого совпадения с gРНК для создания двухцепочечного разрыва. Последовательности CRISPR в бактериях экспрессируются в виде нескольких РНК и затем процессируются с образованием направляющих цепей для РНК. Поскольку в эукариотических системах отсутствуют некоторые белки, необходимые для процессинга молекул CRISPR РНК, была создана синтетическая конструкция gРНК, объединяющая важные фрагменты РНК для направления Cas9 в одной РНК, экспрессируемой с промотора РНК-полимеразы III типа, U6. Синтетические gРНК имеют минимальную длину немногим более 100 п.н. и содержат участок, который направленно взаимодействует с 20 нуклеотидами протоспейсера, непосредственно предшествующего PAM последовательности NGG; gРНК не содержат PAM последовательность.

В некоторых вариантах осуществления gРНК направленно взаимодействует с сайтом в гене дистрофина дикого типа. В некоторых вариантах осуществления gРНК направленно взаимодействует с сайтом в мутантном гене дистрофина. В некоторых вариантах осуществления gРНК направленно взаимодействует с интроном дистрофина. В некоторых вариантах осуществления gРНК направленно взаимодействует с экзоном дистрофина. В некоторых вариантах осуществления gРНК направленно взаимодействует с сайтом в экзоне дистрофина, который экспрессируется и присутствует в одной или более изоформах дистрофина, показанных в Таблице 1. В вариантах осуществления gРНК направленно взаимодействует с сайтом сплайсинга дистрофина.

В некоторых вариантах осуществления gРНК направленно взаимодействует с донорным сайтом сплайсинга на гене дистрофина. В вариантах осуществления gРНК направленно взаимодействует с акцепторным сайтом сплайсинга на гене дистрофина.

В вариантах осуществления Направляющая РНК направленно взаимодействует с мутантным экзоном DMD. В некоторых вариантах осуществления мутантный экзон является экзоном 23 или 51. В некоторых вариантах осуществления направляющая РНК направленно взаимодействует по меньшей мере с одним из экзонов 1, 23, 41, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54 или 55 гена дистрофина. В вариантах осуществления направляющая РНК направленно взаимодействует по меньшей мере с одним из интронов 44, 45, 50, 51, 52, 53, 54 или 55 гена дистрофина. В предпочтительных вариантах осуществления направляющая РНК сконструирована так, чтобы она вызывала пропуск экзона 51 или экзона 23. В вариантах осуществления gРНК направленно взаимодействует с акцепторным сайтом сплайсинга экзона 51 или экзона 23.

Подходящие gРНК для применения в различных композициях и способах, раскрытых в настоящем документе, представлены как SEQ ID NO: 383-705, 709-711, 715-717, 790-862, 864. (Таблицы 7, 9, 11, 13 и 15). В предпочтительных вариантах осуществления gРНК выбрана из любого из SEQ ID NO: 790 - SEQ ID NO: 862.

В некоторых вариантах осуществления, gРНК согласно настоящему описанию включают последовательность, которая комплементарна последовательности-мишени в кодирующей последовательности или некодирующей последовательности, соответствующей гену *DMD*, и, поэтому, гибридизуется с целевой последовательностью-мишенью. В некоторых вариантах осуществления, gРНК для Cpf1 включают одну crРНК, содержащую прямую каркасную последовательность с повторами, после которой расположены 24 нуклеотида направляющей последовательности. В некоторых вариантах осуществления "направляющая" последовательность crРНК включает последовательность gРНК, которая комплементарна последовательности-мишени. В некоторых вариантах осуществления crРНК согласно настоящему описанию включает последовательность gРНК, которая не комплементарна

последовательности-мишени. "Каркасные" последовательности согласно настоящему описанию соединяют gРНК с полипептидом Cpf1. "Каркасные" последовательности согласно настоящему описанию не эквивалентны последовательности tracrРНК в конструкции gРНК-Cas9.

В некоторых вариантах осуществления нуклеиновая кислота может включить одну или более последовательностей, кодирующих gРНК. В некоторых вариантах осуществления нуклеиновая кислота может включать 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 последовательностей, кодирующих gРНК. В некоторых вариантах осуществления все последовательности кодируют одинаковую gРНК. В некоторых вариантах осуществления все последовательности кодируют разные gРНК. В некоторых вариантах осуществления по меньшей мере 2 последовательности кодируют одинаковую gРНК, например по меньшей мере 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 или 20 последовательностей кодируют одинаковую gРНК.

E. Cas9 в сравнении с Cpf1

Cas9 требует две молекулы РНК для разрезания ДНК, тогда как для Cpf1 нужна одна. Белки также разрезают ДНК в разных местах, предоставляя исследователям больше возможностей при выборе сайта редактирования. Cas9 разрезает обе цепи молекулы ДНК в одном и том же положении, оставляя после разреза "тупые" концы. Cpf1 оставляет одну цепь длиннее другой, образуя "липкие" концы, с которыми легче работать. Вероятно Cpf1 в большей степени способна встраивать новые последовательности в сайт разреза по сравнению с Cas9. Несмотря на то, что система CRISPR/Cas9 может эффективно удалять гены, гены сложно вставлять или создавать нокин. Cpf1 не имеет tracrРНК, использует Т-богатый PAM и расщепляет ДНК с образованием двухцепочечного разрыва (DSB) ДНК с выступающими концами.

Подводя итог вышесказанному, важные различия между системами Cpf1 и Cas9 заключаются в том, что Cpf1 распознает разные PAM, обеспечивая новые возможности направленного воздействия, создает липкие концы длиной 4-5 нт вместо тупых концов, которые дает Cas9, что повышает эффективность

генетических вставок и специфичность в ходе NHEJ или HDR, и разрезает ДНК-мишень дальше от PAM, дальше от сайта расщепления Cas9, что обеспечивает новые возможности для расщепления ДНК.

Таблица 3: Различия между Cas9 и Cpf1

Свойства	Cas9	Cpf1
Структура	Требуются две РНК (или 1 слитый транскрипт (crРНК+tracrРНК=gРНК))	Требуется одна РНК
Механизм разреза	Разрез с тупыми концами	Липкие концы
Сайт расщепления	Расположен вблизи от сайта узнавания	Удален от сайта узнавания
Сайты-мишени	G-богатый PAM	T-богатый PAM

F. CRISPR-опосредованное редактирование генов

Первый этап в редактировании гена DMD с применением CRISPR/Cpf1 или CRISPR/Cas9 заключается в идентификации геномной последовательности-мишени. Геномной мишенью для gРНК согласно изобретению может быть любая последовательность ДНК из ~24 нуклеотидов, при условии, что эта последовательность является уникальной по сравнению с остальной частью генома. В некоторых вариантах осуществления геномная последовательность-мишень соответствует последовательности в экзоне 51, экзоне 45, экзоне 44, экзоне 53, экзоне 46, экзоне 52, экзоне 50, экзоне 43, экзоне 6, экзоне 7, экзоне 8 и/или экзоне 55 гена дистрофина человека. В некоторых вариантах осуществления геномная последовательность-мишень представляет собой 5' или 3' сайт сплайсинга экзона 51, экзона 45, экзона 44, экзона 53, экзона 46, экзона 52, экзона 50, экзона 43, экзона 6, экзона 7, экзона 8 и/или экзона 55 гена дистрофина человека. В некоторых вариантах осуществления геномная последовательность-мишень соответствует последовательности в интроне непосредственно перед или после экзона 51, экзона 45, экзона 44, экзона 53, экзона 46, экзона 52, экзона 50, экзона 43, экзона 6, экзона 7, экзона 8 и/или экзона 55 гена дистрофина человека. Примеры геномных последовательностей-мишеней можно найти в Таблицах 6, 8, 10, 12 и 14.

Следующим этапом в редактировании гена DMD является идентификация всех последовательностей примыкающих к

протоспейсеру мотивов (PAM) в генетической области, которая служит мишенью. Последовательность-мишень должна находиться непосредственно перед PAM. После того, как все возможные последовательности PAM и предполагаемые сайты-мишени были идентифицированы, следующим этапом является выбор того сайта, который может обеспечить наиболее эффективное расщепление на мишени. Направляющая последовательность gРНК должна совпадать с последовательностью-мишенью, при этом направляющая последовательность gРНК не должна совпадать с дополнительными сайтами в геноме. В предпочтительных вариантах осуществления направляющая последовательность gРНК обладает полной гомологией с мишенью без гомологии с каким-либо другим участом в геноме. В некоторых вариантах осуществления данная направляющая последовательность gРНК будет иметь дополнительные сайты по всему геному, в которых будет присутствовать частичная гомология. Эти сайты называются "нецелевыми" и должны учитываться при создании gРНК. Как правило, нецелевые сайты не расщепляются так же эффективно, если некомплементарные нуклеотиды присутствуют вблизи последовательности PAM, поэтому молекулы gРНК без гомологии или с некомплементарными нуклеотидами, расположенными вблизи от последовательности PAM, будут обладать наиболее высокой специфичностью. В дополнение к "нецелевой активности" необходимо учитывать факторы, которые обеспечивают максимальное расщепление нужной последовательности-мишени ("целевая активность"). Специалистам в данной области известно, что две направляющих последовательности gРНК, каждая из которых обладает 100% гомологией с ДНК-мишенью, могут не обеспечивать эквивалентную эффективность расщепления. Фактически, эффективность расщепления может увеличиваться или уменьшаться в зависимости от конкретных нуклеотидов в выбранной последовательности-мишени. Тщательное изучение предполагаемой целевой и нецелевой активности для каждой потенциальной направляющей последовательности gРНК необходимо для создания лучшей gРНК. Были созданы несколько программ для конструирования gРНК, которые способны определять положение потенциальных PAM и последовательностей-мишеней и ранжировать связанные с ними gРНК

на основе их прогнозируемой целевой и нецелевой активности (например, CRISPRdirect, доступная на сайте www.crispr.dbcls.jp).

Следующий этап состоит в синтезе и клонировании нужных gРНК. Направляющие олиги могут быть синтезированы, отожжены и встроены в плазмиды, содержащие gРНК каркас, при использовании стандартного клонирования с рестрикцией-лигированием. Впрочем, точная стратегия клонирования будет зависеть от выбранного gРНК вектора. Молекулы gРНК для Cpf1 заметно более простые, чем gРНК для Cas9, и состоят только из одной crРНК, содержащей прямую повторяющуюся каркасную последовательность, после которой расположена направляющая последовательность из ~24 нуклеотидов.

Затем каждую gРНК нужно проверить на одной или нескольких линиях клеток-мишеней. Например, после того, как Cas9 или Cpf1 и gРНК доставлены в клетку, геномная область-мишень может быть амплифицирована с помощью ПЦР и секвенирована в соответствии со способами, известными специалистам в данной области.

В некоторых вариантах осуществления редактирование генов может быть выполнено *in vitro* или *ex vivo*. В некоторых вариантах осуществления клетки подвергают контакту *in vitro* или *ex vivo* с Cas9 или Cpf1 и gРНК, которая направлена на сайт сплайсинга дистрофина. В некоторых вариантах осуществления клетки подвергают контакту с одной или более нуклеиновыми кислотами, кодирующими Cas9 или Cpf1 и направляющую РНК. В некоторых вариантах осуществления одну или более нуклеиновых кислот вводят в клетки при использовании, например, липофекции или электропорации. Редактирование генов также можно проводить на зиготах. В вариантах осуществления в зиготы можно вводить одну или более нуклеиновых кислот, кодирующих Cas9 или Cpf1 и gРНК, которая направлена на сайт сплайсинга дистрофина. Затем зиготы могут быть введены реципиенту.

В вариантах осуществления Cas9 или Cpf1 представлены на векторе. В вариантах осуществления вектор содержит Cas9, полученный из *S. pyogenes* (SpCas9, SEQ ID NO. 872). В вариантах осуществления вектор содержит Cas9, полученный из *S. aureus* (SaCas9, SEQ ID NO. 873). В вариантах осуществления вектор

содержит последовательность Cpf1, полученную из бактерии *Lachnospiraceae*. См., например, рег. номер Uniprot A0A182DWE3; SEQ ID NO. 871. В вариантах осуществления вектор содержит последовательность Cpf1, полученную из бактерии *Acidaminococcus*. См., например, рег. номер Uniprot U2UMQ6; SEQ ID NO. 870. В некоторых вариантах осуществления последовательность Cas9 или Cpf1 является кодон-оптимизированной для экспрессии в клетках человека или мыши. В некоторых вариантах осуществления вектор дополнительно содержит последовательность, кодирующую флуоресцентный белок, такой как GFP, который позволяет сортировать клетки, экспрессирующие Cas 9 или Cpf1, при использовании сортировки клеток с активированной флуоресценцией (FACS). В некоторых вариантах осуществления вектор является вирусным вектором, таким как аденоассоциированный вирусный вектор.

В вариантах осуществления gРНК представлена на векторе. В некоторых вариантах осуществления вектор является вирусным вектором, таким как аденоассоциированный вирусный вектор. В вариантах осуществления Cas9 или Cpf1 и направляющая РНК представлены на одном векторе. В вариантах осуществления Cas9 или Cpf1 и направляющая РНК представлены на разных векторах.

В некоторых вариантах осуществления клетки дополнительно подвергают контакту с одноцепочечным олигонуклеотидом DMD для проведения репарации, направленной гомологией. В некоторых вариантах осуществления небольшие инделы восстанавливают рамку считывания белка дистрофина (стратегия "восстановления рамки"). Когда используется стратегия восстановления рамки, клетки могут контактировать с одной gРНК. В вариантах осуществления донорный или акцепторный сайт сплайсинга прерываются, что приводит к пропуску экзона и восстановлению рамки считывания белка (стратегия "пропуска экзона"). Когда используется стратегия пропуска экзонов, клетки могут контактировать с двумя или более gРНК.

Эффективность *in vitro* или *ex vivo* Cas9 или Cpf1-опосредованного расщепления ДНК можно оценивать при использовании методов, известных специалистам, таких как анализ

T7 E1. Восстановление экспрессии DMD может быть подтверждено с помощью методов, известных специалистам, таких как ОТ-ПЦР, вестерн-блоттинг и иммуноцитохимия.

В некоторых вариантах осуществления редактирование генов *in vitro* или *ex vivo* проводят в мышечной или сателлитной клетке. В некоторых вариантах осуществления редактирование генов проводят в клетках иПСК или вКМ. В вариантах осуществления клетки иПСК дифференцируются после редактирования генов. Например, клетки иПСК могут быть дифференцированы в мышечную клетку или сателлитную клетку после редактирования. В вариантах осуществления клетки иПСК дифференцируются в клетки сердечной мышцы, клетки скелетных мышц или клетки гладких мышц. В вариантах осуществления клетки иПСК дифференцируются в кардиомиоциты. Клетки иПСК могут быть индуцированы для дифференцировки согласно способам, известным специалистам в данной области.

В некоторых вариантах осуществления контакт клетки с Cas9 или Cpf1 и gРНК восстанавливает экспрессию дистрофина. В вариантах осуществления клетки, которые были отредактированы *in vitro* или *ex vivo*, или клетки, полученные из них, демонстрируют уровень белка дистрофина, который сопоставим с клетками дикого типа. В вариантах осуществления отредактированные клетки или клетки, полученные из них, экспрессируют дистрофин на уровне, который составляет 50%, 60%, 70%, 80%, 90%, 95%, или любой процент между указанными значениями, от уровней экспрессии дистрофина дикого типа. В вариантах осуществления клетки, которые были отредактированы *in vitro* или *ex vivo*, или клетки, полученные из них, содержат количество митохондрий, сравнимое с количеством митохондрий в клетках дикого типа. В вариантах осуществления отредактированные клетки или полученные из них клетки содержат 50%, 60%, 70%, 80%, 90%, 95%, или любой процент между указанными значениями, от количества митохондрий в клетках дикого типа. В вариантах осуществления отредактированные клетки или клетки, полученные из них, демонстрируют увеличение скорости потребления кислорода (OCR) по сравнению с нередактированными клетками в исходном состоянии.

G. РНК-полимераза III и промоторы Pol III

В эукариотах РНК-полимераза III (также называемая Pol III) транскрибирует ДНК, синтезируя рибосомную 5S рРНК, тРНК и другие малые РНК. Гены, транскрибируемые РНК Pol III, относятся к категории генов "домашнего хозяйства", экспрессия которых требуется во всех типах клеток и большей части условий окружающей среды. Поэтому регуляция транскрипции Pol III, прежде всего, связана с регуляцией роста клеток и клеточного цикла, что требует меньшего количества регуляторных белков, чем РНК-полимераза II. В условиях стресса, тем не менее, белок Maf1 блокирует активность Pol III.

В процессе транскрипции (любой полимеразой) присутствуют три основных стадии: (i) инициация, требующая сборки комплекса РНК-полимеразы на промоторе гена; (ii) элонгация, синтез РНК транскрипта; и (iii) терминация, окончание транскрипции РНК и разборка комплекса РНК-полимеразы.

Промоторы под контролем РНК Pol III включают промоторы рибосомной 5S рРНК, тРНК и некоторых других малых РНК, таких как сплайсосомная РНК U6, РНК РНКазы P и РНКазы MRP, РНК 7SL (РНК компонент сигнал-распознающих частиц), Vault РНК, Y РНК, SINE (короткие рассеянные повторяющиеся элементы), РНК 7SK, двк микроРНК, несколько малых ядрышковых РНК и некоторые регуляторные антисмысловые РНК

IV. Доставка нуклеиновой кислоты

Как обсуждалось выше, в некоторых вариантах осуществления экспрессионные кассеты используются для экспрессии продукта фактора транскрипции, либо для последующей очистки и доставки в клетку/субъекту, либо для применения непосредственно в методе генетической доставки. В настоящем документе предложены векторы экспрессии, которые содержат одну или более нуклеиновых кислот, кодирующих Cas9 или Cpf1 и по меньшей мере одну направляющую РНК DMD, которая направленно взаимодействует с сайтом сплайсинга дистрофина. В некоторых вариантах осуществления нуклеиновая кислота, кодирующая Cas9 или Cpf1, и нуклеиновая кислота, кодирующая по меньшей мере одну направляющую РНК, представлены в одном векторе. В других вариантах осуществления нуклеиновая

кислота, кодирующая Cas9 или Cpf1, и нуклеиновая кислота, кодирующая по меньшей мере одну направляющую РНК, представлены в отдельных векторах.

Экспрессия требует, чтобы соответствующие сигналы были представлены в векторах и включали различные регуляторные элементы, такие как энхансеры/промоторы как из вирусных источников, так и из млекопитающих, которые направляют экспрессию представляющих интерес генов в клетках. Также определены элементы, предназначенные для оптимизации стабильности и способности трансляции матричной РНК в клетках-хозяевах. Также предусмотрены условия применения ряда доминантных селективных маркеров, обеспечивающих отбор с использованием лекарственных средств, для создания постоянных стабильных клонов клеток, экспрессирующих продукты, а также элемент, который связывает экспрессию таких селективных маркеров с экспрессией полипептида.

А. Регуляторные элементы

По всему тексту настоящей заявке термин "кассета экспрессии" служит для включения любого типа генетической конструкции, содержащей нуклеиновую кислоту, кодирующую продукт гена, в которой часть или вся последовательность, кодирующая нуклеиновую кислоту, способна транскрибироваться и транслироваться, т.е. находится под контролем промотора. "Промотор" относится к последовательности ДНК, распознаваемой синтетическим аппаратом клетки или введенным синтетическим аппаратом, необходимой для инициации специфической транскрипции гена. Фраза "под транскрипционным контролем" означает, что промотор находится в правильном положении и ориентации по отношению к нуклеиновой кислоте для контроля инициации РНК-полимеразы и экспрессии гена. Под "вектором экспрессии" подразумеваются кассеты экспрессии, содержащиеся в генетической конструкции, которая способна к репликации, и, таким образом, включает одну или более точек начала репликации, сигналов терминации транскрипции, поли-А областей, селективных маркеров и многоцелевых сайтов клонирования.

Термин промотор будет использоваться в настоящем документе

для обозначения группы модулей транскрипционного контроля, которые сгруппированы вокруг сайта инициации РНК-полимеразы II. Большая часть размышлений о том, как организованы промоторы, основана на анализе нескольких вирусных промоторов, в том числе промоторов тимидинкиназы (tk) HSV и ранних единиц транскрипции SV40. Эти исследования, дополненные более поздней работой, показали, что промоторы состоят из дискретных функциональных модулей, каждый из которых содержит приблизительно 7-20 пн ДНК и содержит один или более сайтов распознавания для белков-активаторов или репрессоров транскрипции.

По меньшей мере один модуль в каждом промоторе функционирует так, чтобы определять положение сайта начала транскрипции для синтеза РНК. Наиболее известным примером этого является ТАТА-бокс, но в некоторых промоторах, не имеющих ТАТА-бокса, таких как промотор гена терминальной деоксинуклеотидилтрансферазы у млекопитающих и промотор поздних генов SV40, дискретный элемент, перекрывающий сам сайт начала транскрипции, помогает фиксировать участок инициации.

В некоторых вариантах осуществления конструкции Cas9 или Cpf1 согласно настоящему описанию экспрессируются специфическим промотором мышечных клеток. Такой специфический промотор мышечных клеток может быть конститутивно активным или может быть индуцируемым промотором.

Дополнительные элементы промоторов регулируют частоту инициации транскрипции. Как правило, они расположены в области 30-110 п.н. перед сайтом начала транскрипции, хотя недавно было показано, что ряд промоторов также содержат функциональные элементы после сайта начала транскрипции. Интервал между элементами промоторов часто является переменным, при этом функция промотора сохраняется, когда элементы инвертированы или перемещены друг относительно друга. В промоторе tk расстояние между промоторными элементами может быть увеличено до 50 п.н., прежде чем активность начнет снижаться. В зависимости от промотора оказывается, что отдельные элементы могут функционировать совместно или независимо для активации транскрипции.

В некоторых вариантах осуществления вирусные промоторы, такие как промотор предранних генов цитомегаловируса человека (ЦМВ), ранний промотор SV40, длинный концевой повтор вируса саркомы Рауса, промотор инсулина крысы и глицеральдегид-3-фосфатдегидрогеназы, можно использовать для получения высокого уровня экспрессии кодирующей последовательности, представляющей интерес. Также предусмотрено применение других вирусных промоторов или клеточных промоторов млекопитающих, или бактериальных фаговых промоторов, которые хорошо известны в данной области для обеспечения экспрессии представляющей интерес кодирующей последовательности, при условии, что уровни экспрессии достаточны для данной цели. При использовании промотора с хорошо известными свойствами можно оптимизировать уровень и профиль экспрессии представляющего интерес белка после трансфекции или трансформации. Кроме того, выбор промотора, который регулируется в ответ на конкретные физиологические сигналы, может обеспечить индуцируемую экспрессию продукта гена.

Энхансеры представляют собой генетические элементы, которые увеличивают транскрипцию с промотора, расположенного в удаленном положении на той же молекуле ДНК. Энхансеры очень похожи на промоторы. То есть они состоят из множества отдельных элементов, каждый из которых связывается с одним или несколькими транскрипционными белками. Основное различие между энхансерами и промоторами является функциональным. Область энхансера в целом должна быть способна стимулировать транскрипцию на расстоянии; это не обязательно должно относиться к области промотора или его составным элементам. С другой стороны, промотор должен иметь один или более элементов, которые направляют инициацию синтеза РНК на определенном сайте и в определенной ориентации, тогда как энхансеры не обладают такой специфичностью. Промоторы и энхансеры часто перекрываются и примыкают друг к другу, часто имея, по-видимому, очень сходную модульную организацию.

Ниже приведен список промоторов/энхансеров и индуцируемых промоторов/энхансеров, которые могут использоваться в комбинации с нуклеиновой кислотой, кодирующей интересующий ген в экспрессионной конструкции. Кроме того, любая комбинация

промотора/энхансера (согласно базе данных эукариотических промоторов EPDB) также может использоваться для направления экспрессии гена. Эукариотические клетки могут поддерживать цитоплазматическую транскрипцию с некоторых бактериальных промоторов, если предоставлена соответствующая бактериальная полимераза, либо в составе комплекса доставки, либо в качестве дополнительной генетической экспрессирующей конструкции.

Промотором и/или энхансером может быть, например, легкая цепь иммуноглобулина, тяжелая цепь иммуноглобулина, Т-клеточный рецептор, HLA DQ α и/или DQ β, β-интерферон, интерлейкин-2, рецептор интерлейкина-2, МНС класса II 5, МНС класса II HLA-Dra, β-актин, мышечная креатинкиназа (MCK), преальбумин (транстиретин), эластаза I, металлотионеин (MTII), коллагеназа, альбумин, α-фетопротеин, t-глобин, β-глобин, c-fos, c-HA-*ras*, инсулин, молекула адгезии нервных клеток (NCAM), α_1-антитрипаин, гистон H2B (TH2B), мышиный коллаген и/или коллаген I типа, регулируемые глюкозой белки (GRP94 и GRP78), гормон роста крысы, человеческий сывороточный амилоид А (SAA), тропонин I (TN I), тромбоцитарный фактор роста (PDGF), миодистрофия Дюшенна, SV40, полиома, ретровирусы, вирус папилломы, вирус гепатита В, вирус иммунодефицита человека, цитомегаловирус (ЦМВ) и вирус лейкоза гиббонов.

В некоторых вариантах осуществления могут использоваться индуцируемые элементы. В некоторых вариантах осуществления индуцируемым элементом является, например, MTII, MMTV (вирус опухоли молочной железы мышей), β-интерферон, аденовирус 5 E2, коллагеназа, стромелизин, SV40, ген MX мыши, ген GRP78, α-2-макроглобулин, виментин, ген H-2kb МНС класса I, HSP70, пролиферин, фактор некроза опухоли и/или ген тиреотропного гормона α. В некоторых вариантах осуществления индуктором является форболовый эфир (TFA), тяжелые металлы, глюкокортикоиды, поли(rI)x, поли(rc), ElA, форболовый эфир (TPA), интерферон, вирус болезни Ньюкасла, A23187, IL-6, сыворотка, интерферон, большой Т-антиген SV40, PMA и/или гормон щитовидной железы. Любой из индуцируемых элементов, описанных в настоящем документе, может использоваться с любым из индукторов,

описанных в настоящем документе.

Особый интерес представляют мышечно-специфические промоторы. К ним относятся промотор легкой цепи-2 миозина, промотор α-актина, промотор тропонина 1; промотор Na^+/Ca^{2+}-обменника, промотор дистрофина, промотор α7-интегрина, промотор натрийуретического пептида головного мозга и промотор αB-кристаллина/малого белка теплового шока, промотор тяжелой цепи α-миозина и промотор ANF. В некоторых вариантах осуществления мышечно-специфический промотор является промотором CK8. Промотор CK8 имеет следующую последовательность (SEQ ID NO: 874):

1 CTAGACTAGC ATGCTGCCCA TGTAAGGAGG CAAGGCCTGG GGACACCCGA GATGCCTGGT

61 TATAATTAAC CCAGACATGT GGCTGCCCCC CCCCCCCCAA CACCTGCTGC CTCTAAAAAT

121 AACCCTGCAT GCCATGTTCC CGGCGAAGGG CCAGCTGTCC CCCGCCAGCT AGACTCAGCA

181 CTTAGTTTAG GAACCAGTGA GCAAGTCAGC CCTTGGGGCA GCCCATACAA GGCCATGGGG

241 CTGGGCAAGC TGCACGCCTG GGTCCGGGGT GGGCACGGTG CCCGGGCAAC GAGCTGAAAG

301 CTCATCTGCT CTCAGGGGCC CCTCCCTGGG GACAGCCCCT CCTGGCTAGT CACACCCTGT

361 AGGCTCCTCT ATATAACCCA GGGGCACAGG GGCTGCCCTC ATTCTACCAC CACCTCCACA

421 GCACAGACAG ACACTCAGGA GCCAGCCAGC

В некоторых вариантах осуществления клеточно-специфический промотор мышечных клеток является вариантом промотора CK8, называемым CK8e. Промотор CK8e имеет следующую последовательность (SEQ ID NO. 875):

1 TGCCCATGTA AGGAGGCAAG GCCTGGGGAC ACCCGAGATG CCTGGTTATA ATTAACCCAG

61 ACATGTGGCT GCCCCCCCCC CCCCAACACC TGCTGCCTCT AAAAATAACC CTGCATGCCA

121 TGTTCCCGGC GAAGGGCCAG CTGTCCCCCG CCAGCTAGAC TCAGCACTTA GTTTAGGAAC

181 CAGTGAGCAA GTCAGCCCTT GGGGCAGCCC ATACAAGGCC ATGGGGCTGG

GCAAGCTGCA

241 CGCCTGGGTC CGGGGTGGGC ACGGTGCCCG GGCAACGAGC TGAAAGCTCA TCTGCTCTCA

301 GGGGCCCCTC CCTGGGGACA GCCCCTCCTG GCTAGTCACA CCCTGTAGGC TCCTCTATAT

361 AACCCAGGGG CACAGGGGCT GCCCTCATTC TACCACCACC TCCACAGCAC AGACAGACAC

421 TCAGGAGCCA GCCAGC

Когда используется вставка кДНК, обычно желательно включать сигнал полиаденилирования для обеспечения правильного полиаденилирования транскрипта гена. Могут использоваться любые последовательности полиаденилирования, такие как сигналы полиаденилирования человеческого гормона роста и SV40. Также в качестве элемента кассеты экспрессии рассматривается терминатор. Эти элементы могут служить для повышения уровня мРНК и сведения к минимуму сквозного прочтения кассеты с получением других последовательностей.

B. 2A пептид

В некоторых вариантах осуществления используется 2A-подобный саморасщепляющийся домен из вируса насекомых *Thosea asigna* (2A пептид TaV) (SEQ ID NO. 876, EGRGSLLTCGDVEENPGP). Такие 2A-подобные домены, как было показано, функционировали в эукариотах и вызывали котрансляционное расщепление аминокислот внутри 2A-подобного пептидного домена. Поэтому включение 2A пептида TaV дает возможность экспрессии нескольких белков с одного мРНК транскрипта. Важно отметить, что домен TaV при тестировании в эукариотических системах показал активность расщепления более 99%. Другие подходящие 2A-подобные пептиды включают, без ограничения перечисленными, 2A пептид вируса ринита A (ERAV) лошадей (SEQ ID NO: 877; QCTNYALLKLAGDVESNPGP), 2A пептид тешовируса-1 (PTV1) свиней (SEQ ID NO: 878; ATNFSLLKQAGDVEENPGP) и 2A пептид вируса ящура (FMDV) (SEQ ID NO: 879; PVKQLLNFDLLKLAGDVESNPGP) или их модифицированные варианты.

В некоторых вариантах осуществления 2A пептид используется для экспрессии репортера и Cas9 или Cpf1 одновременно. Репортером может быть, например, GFP.

Другие саморасщепляющиеся пептиды, которые могут использоваться, включают, без ограничения перечисленными, протеазу белка внутриядерного включения а (Nia), P1 протеазу, 3C протеазу, L протеазу, 3C-подобную протеазу или их модифицированные варианты.

C. Доставка векторов экспрессии

Существует несколько способов, которыми векторы экспрессии могут быть введены в клетки. В некоторых вариантах осуществления экспрессионная конструкция содержит вирус или сконструированную конструкцию, полученную из вирусного генома. Способность некоторых вирусов проникать в клетки путем рецептор-опосредованного эндоцитоза, интегрироваться в геном клетки-хозяина и стабильно и эффективно экспрессировать вирусные гены сделали их привлекательными кандидатами для переноса чужеродных генов в клетки млекопитающих. Они обладают относительно низкой емкостью для включения чужеродных последовательностей ДНК и имеют ограниченный спектр хозяев. Кроме того, их онкогенный потенциал и цитопатические эффекты в пермиссивных клетках вызывают опасения за безопасность. Они могут вмещать только до 8 тпн чужеродного генетического материала, но могут быть легко введены в различные клеточные линии и лабораторных животных.

Один из предпочтительных способов доставки *in vivo* включает применение аденовирусного вектора экспрессии. Под "аденовирусным вектором экспрессии" подразумеваются такие конструкции, которые содержат последовательности аденовируса, достаточные для (a) обеспечения упаковки конструкции и (b) для экспрессии антисмыслового полинуклеотида, который был клонирован в нем. В этом контексте экспрессия не требует синтеза продукта гена.

Вектор экспрессии включает генно-инженерно модифицированную форму аденовируса. Знание генетической организации аденовируса, линейного двухцепочечного ДНК-вируса, размером 36 тпн, позволяет заменять большие фрагменты аденовирусной ДНК чужеродными последовательностями до 7 тпн. В отличие от ретровируса, аденовирусная инфекция в клетках-хозяевах не приводит к хромосомной интеграции, потому что аденовирусная ДНК может реплицироваться эписомно без потенциальной генотоксичности.

Кроме того, аденовирусы являются структурно стабильными, и после масштабной амплификации не было обнаружено никакой перестройки генома. Аденовирус может инфицировать практически все эпителиальные клетки независимо от стадии их клеточного цикла. До настоящего времени аденовирусная инфекция, по-видимому, была связана только с легкими заболеваниями, такими как острые респираторные заболевания у людей.

Аденовирус особенно подходит для применения в качестве вектора для переноса генов из-за своего генома среднего размера, простоты работы, высокого титра, широкого диапазона клеток-мишеней и высокой инфекционности. Оба конца вирусного генома содержат инвертированные повторы (ITR) длиной 100-200 пар оснований, которые являются цис-элементами, необходимыми для репликации и упаковки вирусной ДНК. Ранняя (E) и поздняя (L) области генома содержат различные единицы транскрипции, которые подразделены по началу репликации вирусной ДНК. Область E1 (E1A и E1B) кодирует белки, ответственные за регуляцию транскрипции вирусного генома и нескольких клеточных генов. Экспрессия области E2 (E2A и E2B) приводит к синтезу белков для репликации вирусной ДНК. Эти белки участвуют в репликации ДНК, экспрессии поздних генов и отключении собственной экспрессии клетки-хозяина. Продукты поздних генов, включая большинство вирусных капсидных белков, экспрессируются только после значительного процессинга единого первичного транскрипта, синтезируемого с главного позднего промотора (MLP). MLP (расположенный на расстоянии 16,8 m.u.) особенно эффективен во время поздней фазы инфекции, при этом все мРНК, синтезируемые с этого промотора, имеют 5'-трехчленную лидерную (TPL) последовательность, что делает их предпочтительными мРНК для трансляции. В одной системе рекомбинантный аденовирус образуется в результате гомологичной рекомбинации между челночным вектором и провирусным вектором. Из-за возможной рекомбинации между двумя провирусными векторами в этом процессе может образоваться аденовирус дикого типа. Таким образом, очень важно выделить одиночный клон вируса из отдельной бляшки и исследовать его геномную структуру.

Создание и размножение современных аденовирусных векторов,

которые являются дефектными по репликации, зависят от уникальной линии хелперных клеток, обозначенной 293, которая была трансформирована из эмбриональных клеток почки человека фрагментами ДНК Ad5 и конститутивно экспрессирует белки E1. Поскольку область E3 не зависит от генома аденовируса, современные аденовирусные векторы с помощью клеток 293 несут чужеродную ДНК либо в области E1, либо D3, или в обеих областях. В природе аденовирус может упаковывать приблизительно 105% генома дикого типа, обеспечивая емкость для примерно 2 дополнительных тпн ДНК. В сочетании с приблизительно 5,5 тпн ДНК, которая может быть заменена в областях E1 и E3, максимальная емкость существующего аденовирусного вектора составляет меньше 7,5 тпн или около 15% от общей длины вектора. Более 80% вирусного генома аденовируса остается в основной последовательности вектора и является источником трансмиссивной цитотоксичности. Кроме того, дефицит репликации вируса с делецией E1 является неполным.

Линии хелперных клеток могут быть получены из клеток человека, таких как эмбриональные клетки почки человека, мышечные клетки, гемопоэтические клетки или другие мезенхимальные или эпителиальные клетки человека. В альтернативе хелперные клетки могут быть получены из клеток других видов млекопитающих, которые являются пермиссивными для аденовируса человека. Такие клетки включают, например, клетки Vero или другие мезенхимальные или эпителиальные клетки обезьян. Как указано выше, предпочтительной линией хелперных клеток является 293.

В уровне техники известны улучшенные способы культивирования клеток 293 и размножения аденовируса. В одном формате натуральные клеточные агрегаты выращивают путем инокуляции отдельных клеток в силиконизированные флаконы с перемешивающим элементом емкостью 1 литр (Techne, Cambridge, UK), содержащие 100-200 мл среды. После перемешивания при 40 об/мин жизнеспособность клеток оценивают с помощью трипанового синего. В другом формате микроносители Fibra-Cel (Bibby Sterlin, Stone, UK) (5 г/л) используют следующим образом. Инокулят

клеток, ресуспендированных в 5 мл среды, добавляют к носителю (50 мл) в колбу Эрленмейера на 250 мл и оставляют в неподвижном состоянии с периодическим перемешиванием на 1-4 часа. Затем среду заменяют 50 мл свежей среды и начинают встряхивание. Для получения вируса клеткам дают расти примерно до 80% конфлюэнтности, после чего среду заменяют (до 25% от конечного объема) и добавляют аденовирус с MOI 0,05. Культуры оставляют на ночь, после чего объем увеличивают до 100% и начинают встряхивание в течение еще 72 часов.

Аденовирусы согласно настоящему описанию являются репликационно-дефектными или, по меньшей мере, условно репликационно-дефектными. Аденовирус может относиться к любому из 42 различных известных серотипов или подгрупп A-F. Аденовирус типа 5 подгруппы C является предпочтительным исходным материалом для получения условно репликационно-дефектного аденовирусного вектора для применения в настоящем описании.

Как указано выше, типичный вектор согласно настоящему описанию является репликационно-дефектным и не содержит область E1 аденовируса. Таким образом, будет наиболее удобно вводить полинуклеотид, кодирующий представляющий интерес ген, в положение, из которого были удалены последовательности, кодирующие E1. Однако положение вставки конструкции в последовательности аденовируса не является столь важным. Полинуклеотид, кодирующий представляющий интерес ген, также может быть вставлен вместо удаленной области E3 в векторах для замены E3 или в область E4, где линия хелперных клеток или хелперный вирус восполняют дефект E4.

Аденовирус легко выращивать и работать с ним, при этом он демонстрирует широкий спектр хозяев *in vitro* и *in vivo*. Эта группа вирусов может быть получена с высокими титрами, например, 10^9–10^{12} бляшкообразующих единиц на мл, при этом они являются высокоинфекционными. Жизненный цикл аденовируса не требует интеграции в геном клетки-хозяина. Чужеродные гены, доставляемые с помощью аденовирусных векторов, являются эписомными и, следовательно, обладают низкой генотоксичностью для клеток-хозяев. Никаких побочных эффектов в исследованиях вакцинации с

применением аденовируса дикого типа не было обнаружено, что демонстрирует их безопасность и терапевтический потенциал в качестве векторов для переноса генов *in vivo*.

Аденовирусные векторы использовались для экспрессии эукариотических генов и разработки вакцин. Исследования на животных показали, что рекомбинантный аденовирус может использоваться для генотерапии. Исследования по введению рекомбинантного аденовируса в различные ткани включают инстилляцию трахеи, внутримышечную инъекцию, периферические внутривенные инъекции и стереотаксическую инокуляцию в головной мозг.

Ретровирусы являются группой одноцепочечных РНК-вирусов, характеризующихся способностью превращать свою РНК в двухцепочечную ДНК в инфицированных клетках с помощью процесса обратной транскрипции. Полученная ДНК затем стабильно интегрируется в клеточные хромосомы в виде провируса и направляет синтез вирусных белков. Интеграция приводит к сохранению последовательностей вирусных генов в реципиентной клетке и ее потомстве. Ретровирусный геном содержит три гена, gag, pol и env, которые кодируют капсидные белки, фермент полимеразу и компоненты оболочки соответственно. Последовательность, обнаруженная перед геном gag, содержит сигнал для упаковки генома в вирионы. Две последовательности длинного концевого повтора (LTR) присутствуют на 5'и 3'-концах вирусного генома. Они содержат последовательности сильного промотора и энхансера, а также необходимы для интеграции в геном клетки-хозяина.

Для создания ретровирусного вектора нуклеиновую кислоту, кодирующую представляющий интерес ген, встраивают в вирусный геном вместо определенных вирусных последовательностей с получением вируса, дефектного по репликации. Для получения вирионов конструируют пакующую клеточную линию, содержащую гены gag, pol и env, но без LTR и пакующих компонентов. При введении рекомбинантной плазмиды, содержащей кДНК, вместе с ретровирусными LTR и пакующими последовательностями, в эту клеточную линию (например, при осаждении фосфатом кальция),

пакующая последовательность позволяет упаковывать РНК-транскрипт рекомбинантной плазмиды в вирусные частицы, которые затем секретируются в среду культивирования. Затем среду, содержащую рекомбинантные ретровирусы, собирают, необязательно концентрируют и используют для переноса гена. Ретровирусные векторы способны инфицировать самые различные типы клеток. Однако интеграция и стабильная экспрессия требуют деления клеток-хозяев.

Недавно на основе химической модификации ретровируса путем химического добавления остатков лактозы в вирусную оболочку был разработан новый подход, позволяющий обеспечить специфическое направление ретровирусных векторов. Эта модификация может обеспечивать специфическое заражение гепатоцитов посредством сиалогликопротеиновых рецепторов.

Может использоваться другой метод направления рекомбинантных ретровирусов, в котором используются биотинилированные антитела против белка ретровирусной оболочки и против специфического клеточного рецептора. Антитела связываются посредством биотиновых компонентов при использовании стрептавидина. При использовании антител против антигенов главного комплекса гистосовместимости I и II класса было продемонстрировано инфицирование различных клеток человека, несущих эти поверхностные антигены, экотропным вирусом *in vitro* (Roux et al., 1989).

Существуют некоторые ограничения по применению ретровирусных векторов во всех аспектах настоящего описания. Например, ретровирусные векторы обычно интегрируются в случайные сайты в геноме клетки. Это может приводить к инсерционному мутагенезу в результате прерывания хозяйских генов или вставки вирусных регуляторных последовательностей, которые могут нарушать функцию фланкирующих генов. Другая проблема, связанная с применением дефектных ретровирусных векторов, заключается в потенциальном появлении репликационно-компетентного вируса дикого типа в пакующих клетках. Это может быть результатом событий рекомбинации, в результате которых интактная последовательность рекомбинантного вируса встраивается перед

последовательностью gag, pol, env, интегрированной в геном клетки-хозяина. Однако в настоящее время доступны новые пакующие клеточные линии, которые должны в значительной степени снизить вероятность рекомбинации (см., например, Markowitz et al., 1988; Hersdorffer et al., 1990).

Другие вирусные векторы могут использоваться в качестве экспрессионных конструкций в настоящем описании. Могут использоваться векторы, полученные из вирусов, таких как вирус осповакцины, аденоассоциированный вирус (AAV) и герпесвирусы. Они обеспечивают несколько привлекательных свойств для различных клеток млекопитающих.

В вариантах осуществления AAV вектор является репликационно-дефектным или условно репликационно-дефектным. В вариантах осуществления AAV вектор является рекомбинантным AAV вектором. В некоторых вариантах осуществления AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 или их любой комбинации.

В некоторых вариантах осуществления один вирусный вектор используют для доставки нуклеиновой кислоты, кодирующей Cas9 или Cpf1 и по меньшей мере одну gРНК, в клетку. В некоторых вариантах осуществления Cas9 или Cpf1 доставляют в клетку при использовании первого вирусного вектора, а по меньшей мере одну gРНК доставляют в клетку при использовании второго вирусного вектора.

В некоторых вариантах осуществления один вирусный вектор используют для доставки нуклеиновой кислоты, кодирующей Cas9 или Cpf1 и по меньшей мере одну gРНК, в клетку. В некоторых вариантах осуществления Cas9 или Cpf1 доставляют в клетку при использовании первого вирусного вектора, а по меньшей мере одну gРНК доставляют в клетку при использовании второго вирусного вектора. Чтобы осуществить экспрессию смысловых или антисмысловых генных конструкций, экспрессионную конструкцию нужно доставить в клетку. Клетка может быть мышечной клеткой, сателлитной клеткой, мезангиобластом, клеткой из костного мозга, стромальной клеткой или мезенхимальной стволовой клеткой. В

вариантах осуществления клетка является клеткой миокарда, клеткой скелетной мышцы или клеткой гладкой мышцы. В вариантах осуществления клетка является клеткой передней большеберцовой мышцы, четырехглавой мышцы, камбаловидной мышцы, диафрагмы или сердца. В некоторых вариантах осуществления клетка является индуцированной плюрипотентной стволовой клеткой (иПСК) или клеткой внутренней клеточной массы (вКМ). В других вариантах осуществления клетка является человеческой иПСК или человеческой вКМ. В некоторых вариантах осуществления человеческая иПСК или человеческая вКМ согласно изобретению могут быть получены из линии культивируемых стволовых клеток, взрослой стволовой клетки, плацентарной стволовой клетки или из другого источника взрослых или эмбриональных стволовых клеток, которые не требуют уничтожения человеческого эмбриона. Доставка в клетку может быть осуществлена *in vitro*, как при лабораторных процедурах трансформации клеточных линий, или *in vivo* или *ex vivo*, как при лечении некоторых патологических состояний. Одним из механизмов доставки является вирусная инфекция, когда экспрессирующая конструкция инкапсулирована в инфекционной вирусной частице.

В настоящем изобретении также предусмотрено несколько невирусных способов переноса экспрессионных конструкций в культивируемые клетки млекопитающих. К ним относятся осаждение фосфатом кальция, DEAE-декстран, электропорация, прямая микроинъекция, ДНК-нагруженные липосомы и комплексы липофектамина-ДНК, обработка клеток ультразвуком, генная бомбардировка с использованием высокоскоростных микрочастиц и рецептор-опосредованная трансфекция. Некоторые из этих методов могут быть успешно адаптированы для применения *in vivo* или *ex vivo*.

После доставки экспрессионной конструкции в клетку нуклеиновая кислота, кодирующая представляющий интерес ген, может быть помещена и экспрессирована в других сайтах. В некоторых вариантах осуществления нуклеиновая кислота, кодирующая ген, может стабильно интегрироваться в геном клетки. Эта интеграция может происходить в родственном положении и ориентации посредством гомологичной рекомбинации (замены гена),

или она может интегрироваться в случайное, неспецифическое положение (увеличение гена). В других вариантах осуществления нуклеиновая кислота может стабильно поддерживаться в клетке в виде отдельного эписомного сегмента ДНК. Такие сегменты нуклеиновой кислоты или "эписомы" кодируют последовательности, достаточные для поддержания и репликации независимо от цикла клетки-хозяина или синхронно с ним. То, каким образом экспрессионная конструкция доставляется в клетку, и где в клетке остается нуклеиновая кислота, зависит от типа используемой экспрессионной конструкции.

В еще одном варианте осуществления экспрессионная конструкция может состоять просто из голой рекомбинантной ДНК или плазмид. Перенос конструкции может быть осуществлен любым из указанных выше способов, которые физически или химически пермеабилизируют клеточную мембрану. Это особенно применимо при переносе *in vitro*, но также может применяться и для использования *in vivo*. Dubensky с соавт. (1984) успешно ввели ДНК полиомавируса в осажденной фосфатом кальция форме в печень и селезенку взрослых и новорожденных мышей, продемонстрировав активную репликацию вируса и острую инфекцию. Benvenisty and Neshif (1986) также продемонстрировали, что прямое внутрибрюшинное введение плазмид, осажденных фосфатом кальция, приводит к экспрессии трансфицированных генов. ДНК, кодирующая представляющий интерес ген, также может быть перенесена аналогичным способом *in vivo* и может экспрессировать продукт гена.

В еще одном варианте осуществления перенос экспрессионной конструкции в виде голой ДНК в клетки может включать бомбардировку частицами. Этот метод зависит от способности ускорять покрытые ДНК микрочастицы до высокой скорости, что позволяет им проникать через клеточные мембраны и попадать в клетки, не убивая их. Были разработаны несколько устройств для ускорения мелких частиц. Одно такое устройство основано на разряде высокого напряжения с генерацией электрического тока, который, в свою очередь, обеспечивает движущую силу. Используемые микрочастицы состояли из биологически инертных

веществ, таких как вольфрамовые или золотые сферы.

В некоторых вариантах осуществления экспрессионную конструкцию доставляют непосредственно в печень, кожу и/или мышечную ткань субъекта. Это может потребовать обеспечения хирургического доступа к ткани или клеткам, устранения любой промежуточной ткани между генной пушкой и целевым органом, т.е. лечения *ex vivo*. Опять же, ДНК, кодирующая определенный ген, может быть доставлена с помощью этого метода и при этом будет включена в настоящее изобретение.

В дополнительном варианте осуществления экспрессионная конструкция может быть заключена в липосому. Липосомы представляют собой везикулярные структуры, характеризующиеся наличием фосфолипидной бислойной мембраны и внутренней водной среды. Мультиламеллярные липосомы имеют несколько липидных слоев, разделенных водной средой. Они образуются спонтанно, когда фосфолипиды суспендируют в избытке водного раствора. Липидные компоненты подвергаются самоорганизации перед образованием замкнутых структур и захватывают воду и растворенные вещества между липидными бислоями. Также рассматриваются комплексы липофектамина-ДНК.

Опосредованная липосомами доставка нуклеиновой кислоты и экспрессия чужеродной ДНК *in vitro* были очень успешными. Реактив, известный как липофектамин (Lipofectamine 2000™), широко используется и доступен в продаже.

В некоторых вариантах осуществления липосома может образовывать комплекс с гемагглютинирующим вирусом Сендай (HVJ), облегчающий слияние с клеточной мембраной и способствующий проникновению в клетку ДНК, инкапсулированной в липосомах. В других вариантах осуществления липосома может быть связана в комплекс или может использоваться в сочетании с ядерными негистоновыми хромосомными белками (HMG-1). В других вариантах осуществления липосома может быть связана в комплекс или может использоваться в сочетании как с HVJ, так и с HMG-1. Поскольку такие экспрессионные конструкции успешно использовались для переноса и экспрессии нуклеиновой кислоты *in vitro* и *in vivo*, то они применимы для настоящего изобретения. Если в ДНК конструкции

используется бактериальный промотор, также предпочтительно включить в липосому соответствующую бактериальную полимеразу.

Другими экспрессионными конструкциями, которые могут применяться для доставки в клетки нуклеиновой кислоты, кодирующей конкретный ген, являются носители для рецептор-опосредованной доставки. В них используются преимущества селективного захвата макромолекул рецептор-опосредованным эндоцитозом почти во всех эукариотических клетках. Из-за клеточно-специфического распределения различных рецепторов доставка может быть высокоспецифичной.

Носители для рецептор-опосредованной адресной доставки гена обычно состоят из двух компонентов: лиганда, специфичного в отношении клеточного рецептора, и ДНК-связывающего средства. Несколько лигандов использовались для рецептор-опосредованного переноса генов. Наиболее широко описанными лигандами являются асиалооросомукоид (ASOR) и трансферрин. Синтетический неогликопротеин, который распознает тот же рецептор, что и ASOR, использовали в качестве носителя для доставки генов, а эпидермальный фактор роста (EGF) также использовали для доставки генов в клетки плоскоклеточной карциномы.

D. AAV-Cas9 векторы

В некоторых вариантах осуществления Cas9 может быть упакован в AAV вектор. В некоторых вариантах осуществления AAV вектор является AAV вектором дикого типа. В некоторых вариантах осуществления AAV вектор содержит одну или более мутаций. В некоторых вариантах осуществления AAV вектор выделен или получен из AAV вектора серотипа AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 или их любой комбинации.

Примеры AAV-Cas9 векторов содержат две последовательности ITR (инвертированных концевых повторов), которые фланкируют центральную область последовательности, включающую последовательность Cas9. В некоторых вариантах осуществления ITR выделены или получены из AAV вектора серотипа AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 или их любой комбинации. В некоторых вариантах осуществления ITR включают или состоят из полноразмерных последовательностей и/или

последовательностей дикого типа серотипа AAV. В некоторых вариантах осуществления ITR включают или состоят из усеченных последовательностей серотипа AAV. В некоторых вариантах осуществления ITR включают или состоят из удлиненных последовательностей серотипа AAV. В некоторых вариантах осуществления ITR включают или состоят из последовательностей, включающих вариант последовательности по сравнению с последовательностью дикого типа такого же серотипа AAV. В некоторых вариантах осуществления вариант последовательности включает одно или более из замены, делеции, вставки, инверсии или транспозиции. В некоторых вариантах осуществления ITR включают или состоят из по меньшей мере 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149 или 150 пар оснований. В некоторых вариантах осуществления ITR включают или состоят из 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149 или 150 пар оснований. В некоторых вариантах осуществления ITR имеют длину 110±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 120±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 130±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 140±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 150±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 115, 145 или 141 пару оснований. В некоторых вариантах осуществления ITR имеют последовательность, выбранную из SEQ ID NO: 880, SEQ ID NO: 881, SEQ ID NO: 882, SEQ ID NO: 883 и SEQ ID SEQ NO: 946.

В некоторых вариантах осуществления AAV-Cas9 вектор может содержать один или более сигналов ядерной локализации (NLS). В некоторых вариантах осуществления AAV-Cas9 вектор содержит 1, 2, 3, 4 или 5 сигналов ядерной локализации. Примеры NLS включают

NLS c-myc (SEQ ID NO: 884), NLS SV40 (SEQ ID NO: 885), NLS hnRNPAI M9 (SEQ ID NO: 886), NLS нуклеоплазмина (SEQ ID NO: 887), последовательность IBB домена из импортина-альфа: RMRKFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 888), последовательности Т-белка миомы: VSRKRPRP (SEQ ID NO: 889) и PPKKARED (SEQ ID NO: 890), последовательность p53 человека: PQPKKKPL (SEQ ID NO: 891), последовательность c-abl IV мыши: SALIKKKKKMAP (SEQ ID NO: 892), последовательности вируса гриппа NS1: DRLRR (SEQ ID NO: 893) и PKQKKRK (SEQ ID NO: 894), последовательность дельта-антигена вируса гепатита: RKLKKKIKKL (SEQ ID NO: 895), и последовательность белка Mx1 мыши: REKKKFLKRR (SEQ ID NO: 896). Другие подходящие сигналы ядерной локализации включают двучленные последовательности ядерной локализации, такие как последовательность KRKGDEVDGVDEVAKKKSKK (SEQ ID NO: 897) человеческой поли-(АДФ-рибоза) полимеразы или последовательность RKCLQAGMNLEARKTKK (SEQ ID NO: 898) рецепторов (человеческих) стероидных гормонов глюкокортикоидов.

В некоторых вариантах осуществления AAV-Cas9 вектор может включать дополнительные элементы, облегчающие упаковку вектора и экспрессию Cas9. В некоторых вариантах осуществления AAV-Cas9 вектор может включать поли(А) последовательность. В некоторых вариантах осуществления поли(А) последовательность может быть мини-поли(А) последовательностью. В некоторых вариантах осуществления AAV-CAs9 вектор может включать транспозируемый элемент. В некоторых вариантах осуществления AAV-Cas9 вектор может включать регуляторный элемент. В некоторых вариантах осуществления регуляторный элемент является активатором или репрессором.

В некоторых вариантах осуществления AAV-Cas9 может содержать один или более промоторов. В некоторых вариантах осуществления один или более промоторов направляют экспрессию Cas9. В некоторых вариантах осуществления один или более промоторов являются мышечно-специфическими промоторами. Примеры мышечно-специфических промоторов включают промотор легкой цепи 2 миозина, промотор α-актина, промотор тропонина 1, промотор Na^+/Ca^{2+}-обменника, промотор дистрофина, промотор α7 интегрина,

промотор натрийуретического пептида головного мозга, промотор αB-кристаллина/малого белка теплового шока, промотор тяжелой цепи α-миозина, промотор ANF, промотор CK8 и промотор CK8e.

В некоторых вариантах осуществления AAV-Cas9 вектор может быть оптимизирован для продукции в дрожжах, бактериях, клетках насекомых или клетках млекопитающих. В некоторых вариантах осуществления AAV-Cas9 вектор может быть оптимизирован для экспрессии в клетках человека. В некоторых вариантах осуществления AAV-Cas9 вектор может быть оптимизирован для экспрессии в системе бакуловирусной экспрессии.

В некоторых вариантах осуществления AAV-Cas9 вектор включает последовательность, выбранную из SEQ ID NO: 899, SEQ ID NO: 900, SEQ ID NO: 901 или SEQ ID NO: 902, как показано в Таблице 4.

Таблица 4: Примеры конструкций для редактирования генов (от ITR до ITR для доставки с помощью AAV вектора)

SEQ ID NO:	Последовательность
899	GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCC GACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTG GCCAACTCCATCACTAGGGGTTCCTTGTAGTTAATGATTAACCCGCCATGCTACTTATC TACGTAGCCATGCTCTAGATTAATTAATGCCCATGTAAGGAGGCAAGGCCTGGGGACAC CCGAGATGCCTGGTTATAATTAACCCAGACATGTGGCTGCCCCCCCCCCCCCAACACCT GCTGCCTCTAAAAATAACCCTGCATGCCATGTTCCCGGCGAAGGGCCAGCTGTCCCCCG CCAGCTAGACTCAGCACTTAGTTTAGGAACCAGTGAGCAAGTCAGCCCTTGGGGCAGCC CATACAAGGCCATGGGGCTGGGCAAGCTGCACGCCTGGGTCCGGGGTGGGCACGGTGCC CGGGCAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCCCCTCCCTGGGGACAGCCCCTC CTGGCTAGTCACACCCTGTAGGCTCCTCTATATAACCCAGGGGCACAGGGGCTGCCCTC ATTCTACCACCACCTCCACAGCACAGACAGACACTCAGGAGCCAGCCAGCGCTAGCGCC ACCATGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGA CAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCA CCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCAC AGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGC CACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCT ATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGA CTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGG CAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAA AGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCC CACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAG CGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAA ACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAG AGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTT CGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACC TGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAAC

	CTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTC CGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCC TGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAA GCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAA GAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCA TCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGA GAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCA CCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGG ACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCT CTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCAC CCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGC GGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTG CTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGG AATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGT TCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATC GAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGG CACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAA ACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATG ATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCT GAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCC GGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAAC AGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAA AGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCA GCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAA GTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGAC CACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCA AAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAAC GAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACT GGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTTTCTGA AGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGC GACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCT GAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCG GCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAG ATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAA TGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATT TCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCCCAC GACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGA AAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGA GCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAAC TTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGA GACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGC GGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGC GGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAA GAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTG TGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAG CTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCT GGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACT CCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAG AAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCA CTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAAC AGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTG ATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAA GCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAG

	CCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACC AAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACG GATCGACCTGTCTCAGCTGGGAGGCGACAAGCGTCCTGCTGCTACTAAGAAAGCTGGTC AAGCTAAGAAAAAGAAATGAATCGATTAGCAATAAAGGATCGTTTATTTTCATTGGAAG CGTGTGTTGGTTTTTTGATCAGGCGCGGGTACCAGGTCGCGGCCGCTCTAGAGCATGGC TACGTAGATAAGTAGCATGGCGGGTTAATCATTAACTACAAGGAACCCCTAGTGATGGA GTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCG CCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA
900	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGT CGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGC CAACTCCATCACTAGGGGTTCCTGTGGCCAACTCCATCACTAGGGGTTCCTTGTAGTTA ATGATTAACCCGCCATGCTACTTATCTACGTAGCCATGCTCTAGATTAATTAATGCCCA TGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATTAACCCAGACATG TGGCTGCCCCCCCCCCCCCAACACCTGCTGCCTCTAAAAATAACCCTGCATGCCATGTT CCCGGCGAAGGGCCAGCTGTCCCCCGCCAGCTAGACTCAGCACTTAGTTTAGGAACCAG TGAGCAAGTCAGCCCTTGGGGCAGCCCATACAAGGCCATGGGGCTGGGCAAGCTGCACG CCTGGGTCCGGGGTGGGCACGGTGCCCGGGCAACGAGCTGAAAGCTCATCTGCTCTCAG GGGCCCCTCCCTGGGGACAGCCCCTCCTGGCTAGTCACACCCTGTAGGCTCCTCTATAT AACCCAGGGGCACAGGGGCTGCCCTCATTCTACCACCACCTCCACAGCACAGACAGACA CTCAGGAGCCAGCCAGCGCTAGCGCCACCATGATGGCCCCAAAGAAGAAGCGGAAGGTC GGTATCCACGGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCAC CAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCA AGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTG TTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATA CACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCA AGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAG AAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAA GTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACC TGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATC GAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCA GACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGG CCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTG CCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGAC CCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGG ACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGAC CTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGT GAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGC ACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGC CAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCG AGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGAC AACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCA GGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCT TCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATG ACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGG CGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACG AGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTG ACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCA GAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGC TGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTG GAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGA CAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCC TGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTG

	TTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCT GAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATT TCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGC CTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCA CGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAG TGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTG ATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAG AATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACC CCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGG CGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGA CCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCA GAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAG ATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGA CAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCA AGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCC CGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCAC CCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCG AGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCC CTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTA CGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGT ACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGC GAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGA TAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCG TGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGG AACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTT CGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGT CCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGC TTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGA CCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAA TGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATAT GTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAA TGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGC AGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTG TCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCA CCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCA TCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAG AGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAAGCG TCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAATGAATCGATTAGCAAT AAAGGATCGTTTATTTTCATTGGAAGCGTGTGTTGGTTTTTTGATCAGGCGCGGGTACC AGGTCGCGGCCGCTCTAGAGCATGGCTACGTAGATAAGTAGCATGGCGGGTTAATCATT AACTACAAGGAACCCCTAGTGATGGAGTTAGGAACCCCTAGTGATGGAGTTGGCCACTC CCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCG GGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG
901	AAGATGACGGTTTGTCACATGGAGTTGGCAGGATGTTTGATTAAAAACATAACAGGAAG AAAAATGCCCCGCTGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAG TTTTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGT AAGCTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCACTTATCTGG TTCGGATCTCCTAGGCGATATCAGTGATCAcggatctcgaccaattgacattattgaag caACTAGTATCGATtttatcagggttattgtctcagaCCTGCAGGCAGCTGCGCGCTCG CTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGG CCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCC TGTGGCCAACTCCATCACTAGGGGTTCCTTGTAGTTAATGATTAACCCGCCATGCTACT TATCTACGTAGCCATGCTCTAGATTAATTAATGCCCATGTAAGGAGGCAAGGCCTGGGG

	ACACCCGAGATGCCTGGTTATAATTAACCCAGACATGTGGCTGCCCCCCCCCCCCCAAC ACCTGCTGCCTCTAAAAATAACCCTGCATGCCATGTTCCCGGCGAAGGGCCAGCTGTCC CCCGCCAGCTAGACTCAGCACTTAGTTTAGGAACCAGTGAGCAAGTCAGCCCTTGGGGC AGCCCATACAAGGCCATGGGGCTGGGCAAGCTGCACGCCTGGGTCCGGGGTGGGCACGG TGCCCGGGCAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCCCCTCCCTGGGGACAGCC CCTCCTGGCTAGTCACACCCTGTAGGCTCCTCTATATAACCCAGGGGCACAGGGGCTGC CCTCATTCTACCACCACCTCCACAGCACAGACAGACACTCAGGAGCCAGCCAGCGCTAG CGCCACCATGATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAG CCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTG ATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCG GCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCG AGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATC TGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCA CAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCT TCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTG AGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCT GGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACA ACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAG GAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAG CAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCC TGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTC GACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGA CAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACC TGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCC CCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCT GAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGA GCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAG TTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAA CAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGA TCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTG AAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGG CCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCA TCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATC GAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAG CCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCG AGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTG CTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAA AATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCC TGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAG GAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGA GATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGC AGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGC ATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGC CAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCC AGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCC GGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGT GAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACC AGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGC ATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCA GAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGG AACTGGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTTT CTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAA GAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGC

	TGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGA GGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCG GCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACG AGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCC GATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGC CCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGC TGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCC AAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCAT GAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGA TCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACC GTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGAC AGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCA GAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTAT TCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAA AGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACT TTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAG TACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACT GCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCA GCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTG GAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAG AGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGG ATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTG GGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAG CACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGA CACGGATCGACCTGTCTCAGCTGGGAGGCGACAAGCGTCCTGCTGCTACTAAGAAAGCT GGTCAAGCTAAGAAAAAGAAATGAATCGATTAGCAATAAAGGATCGTTTATTTTCATTG GAAGCGTGTGTTGGTTTTTTGATCAGGCGCGGGTACCAGGTCGCGGCCGCTCTAGAGCA TGGCTACGTAGATAAGTAGCATGGCGGGTTAATCATTAACTACAAGGAACCCCTAGTGA TGGAGTTaggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgct cactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcag tgagcgagcgagcgcgcagctgcctgcaggCATGCAAGCTGTAGCCAACCACTAGAACT ATAGCTAGAGTCCTGGGCGAACAAACGATGCTCGCCTTCCAGAAAACCGAGGATGCGAA CCACTTCATCCGGGGTCAGCACCACCGGCAAGCGCCGCGACGGCCGAGGTCTTCCGATC TCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGCCGC CAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCAGCCAGGACAGAA ATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACGGATG AAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCG TATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAG TGGCGGTTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGTCTATGCCTCGGGCATCC AAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGT TACGCAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGGC TCAAGTATGGGCATCATTCGCACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCG GGCTGCTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCCCAACATC AGCCGGACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCATCGCGCTTGCT GCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGA GCAGCCGCGTAGTGAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGC AGGGCATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTGGTGCT TATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTCTATACAAA GTTGGGCATACGGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCCACCTAAC AATTCGTTCAAGCCGAGATCGGCTTCCCGGCCGCGGAGTTGTTCGGTAAATTGTCACAA CGCCGCGAATATAGTCTTTACCATGCCCTTGGCCACGCCCCTCTTTAATACGACGGGCA ATTTGCACTTCAGAAAATGAAGAGTTTGCTTTAGCCATAACAAAAGTCCAGTATGCTTT TTCACAGCATAACTGGACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAAGACTTTA

	TTGTCATAGTTTAGATCTATTTTGTTCAGTTTAAGACTTTATTGTCCGCCCACACCCGC TTACGC
902	AAGATGACGGTTTGTCACATGGAGTTGGCAGGATGTTTGATTAAAAACATAACAGGAAG AAAAATGCCCCGCTGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAG TTTTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGT AAGCTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCACTTATCTGG TTCGGATCTCCTAGGCGATATCAGTGATCAcggatctcgaccaattgacattattgaag caACTAGTATCGATtttatcagggttattgtctcagaGGCCACTCCCTCTCTGCGCGCT CGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGC GGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTT CCTTGTAGTTAATGATTAACCCGCCATGCTACTTATCTACGTAGCCATGCTCTAGATTA ATTAATGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATTA ACCCAGACATGTGGCTGCCCCCCCCCCCCCAACACCTGCTGCCTCTAAAAATAACCCTG CATGCCATGTTCCCGGCGAAGGGCCAGCTGTCCCCCGCCAGCTAGACTCAGCACTTAGT TTAGGAACCAGTGAGCAAGTCAGCCCTTGGGGCAGCCCATACAAGGCCATGGGGCTGGG CAAGCTGCACGCCTGGGTCCGGGGTGGGCACGGTGCCCGGGCAACGAGCTGAAAGCTCA TCTGCTCTCAGGGGCCCCTCCCTGGGGACAGCCCCTCCTGGCTAGTCACACCCTGTAGG CTCCTCTATATAACCCAGGGGCACAGGGGCTGCCCTCATTCTACCACCACCTCCACAGC ACAGACAGACACTCAGGAGCCAGCCAGCGCTAGCGCCACCATGATGGCCCCAAAGAAGA AGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTG GACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAG CAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCG GAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCC AGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAA CGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGG AAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCC TACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGA CAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCC ACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATC CAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGT GGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGA TCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGCCCTGAGC CTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCA GCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACC AGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGAC ATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAG ATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGC CTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATT GACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGAT GGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGC GGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATT CTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAA GATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGAT TCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTG GTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAA CCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGT ATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTG AGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGAC CGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAA TCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAA ATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATAT CGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCT ATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGC

	TGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGAC AATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCC ACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGC GATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCAT CCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCG AGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAAC AGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCT GAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACC TGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGAC TACGATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAA GGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGG TCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAG AGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGC CGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGA TCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTG AAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTA CAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCG TGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGAC TACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGC TACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCC TGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAG ATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCA AGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCC TGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAG TACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGA AAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGG AAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAA GTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGG CCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGC CCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCC CCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGA GATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGG ACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAG AATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTT TGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCC TGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGA GGCGACAAGCGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAATGAAT CGATTAGCAATAAAGGATCGTTTATTTTCATTGGAAGCGTGTGTTGGTTTTTTGATCAG GCGCGGGTACCAGGTCGCGGCCGCTCTAGAGCATGGCTACGTAGATAAGTAGCATGGCG GGTTAATCATTAACTACAAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGC GCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCG GGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGACATGCAAGCTGTAGCCAACCAC TAGAACTATAGCTAGAGTCCTGGGCGAACAAACGATGCTCGCCTTCCAGAAAACCGAGG ATGCGAACCACTTCATCCGGGGTCAGCACCACCGGCAAGCGCCGCGACGGCCGAGGTCT TCCGATCTCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGTAGAAGAACAGCA AGGCCGCCAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCAGCCAG GACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAA ACGGATGAAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAATGCA AGTAGCGTATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAAC GGCGCAGTGGCGGTTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGTCTATGCCTCG GGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCA ACGATGTTACGCAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTT AGGTGGCTCAAGTATGGGCATCATTCGCACATGTAGGCTCGGCCCTGACCAAGTCAAAT

	CCATGCGGGCTGCTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCC CAACATCAGCCGGACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCATCGC GCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACGTTCTGCCCA AGTTTGAGCAGCCGCGTAGTGAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCAC CGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCT TGGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTCT ATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCC ACCTAACAATTCGTTCAAGCCGAGATCGGCTTCCCGGCCGCGGAGTTGTTCGGTAAATT GTCACAACGCCGCGAATATAGTCTTTACCATGCCCTTGGCCACGCCCCTCTTTAATACG ACGGGCAATTTGCACTTCAGAAAATGAAGAGTTTGCTTTAGCCATAACAAAAGTCCAGT ATGCTTTTTCACAGCATAACTGGACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAA GACTTTATTGTCATAGTTTAGATCTATTTTGTTCAGTTTAAGACTTTATTGTCCGCCCA CACCCGCTTACGC

В некоторых вариантах осуществления конструкций для редактирования генов согласно настоящему изобретени, включая варианты осуществления, охватывающие SEQ ID NO: 899-902, конструкция включает или состоит из промотора и нуклеазы. В некоторых вариантах осуществления конструкция включает или состоит из промотора CK8e и нуклеазы Cas9. В некоторых вариантах осуществления конструкция включает или состоит из промотора CK8e и нуклеазы Cas9, выделенной или полученной из *Staphylococcus pyogenes* ("SpCas9"). В некоторых вариантах осуществления промотор CK8e включает или состоит из нуклеотидной последовательности:

TGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATTAACCCAGACAT
GTGGCTGCCCCCCCCCCCCCAACACCTGCTGCCTCTAAAAATAACCCTGCATGCCATGTTCCCG
GCGAAGGGCCAGCTGTCCCCCGCCAGCTAGACTCAGCACTTAGTTTAGGAACCAGTGAGCAAGT
CAGCCCTTGGGGCAGCCCATACAAGGCCATGGGGCTGGGCAAGCTGCACGCCTGGGTCCGGGGT
GGGCACGGTGCCCGGGCAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCCCCTCCCTGGGGACA
GCCCCTCCTGGCTAGTCACACCCTGTAGGCTCCTCTATATAACCCAGGGGCACAGGGGCTGCCC
TCATTCTACCACCACCTCCACAGCACAGACAGACACTCAGGAGCCAGCCAGC (SEQ ID NO: 875). В некоторых вариантах осуществления нуклеаза SpCas9 включает или состоит из нуклеотидной последовательности:

GACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCG
ACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAA
GAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAG
AGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCA
GCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGA
AGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCAC
GAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACC

TGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGG
CGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAAC
CAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCA
GACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGG
CCTGTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGAC
CTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGC
TGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCAT
CCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATG
ATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGC
TGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGA
CGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGC
ACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACA
ACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGA
TTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCC
TACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGG
AAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCAT
CGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTG
CTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGA
GAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAA
CCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCC
GTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGA
AAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGT
GCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCAC
CTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGA
GCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAA
GTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAA
GAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATC
TGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGT
GAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACC
ACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGC
TGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTA
CCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTG
TCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACA
AGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGT
GAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTC
GACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGA

GACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAA
CACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAG
CTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACC
ACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCT
GGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGC
GAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCA
AGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGA
AACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATG
CCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCC
TGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGG
CGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAG
TCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCG
AGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCAT
CAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCC
GGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGG
CCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGA
ACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATC
CTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCA
GAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTT
CAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCC
ACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAG
GCGAC (ID NO:872 SEQ). В некоторых вариантах осуществления конструкция, включающая промотор и нуклеазу, дополнительно включает по меньшей мере две последовательности инвертированных концевых повторов (ITR). В некоторых вариантах осуществления конструкция, включающая промотор и нуклеазу, дополнительно включает по меньшей мере две последовательности ITR, выделенных или полученных из AAV серотипа 2 (AAV2). В некоторых вариантах осуществления конструкция, включающая промотор и нуклеазу, дополнительно включает по меньшей мере две последовательности ITR, каждая из которых включает или состоит из нуклеотидной последовательности
GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGC
CCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA (ID NO:880 SEQ). В некоторых вариантах осуществления конструкция, включающая промотор и нуклеазу, дополнительно включает по

меньшей мере две последовательности ITR, где первая последовательность ITR включает или состоит из нуклеотидной последовательности CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGC GACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATC ACTAGGGGTTCCT (SEQ ID NO: 881), и вторая последовательность ITR включает или состоит из нуклеотидной последовательности AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC AGCTGCCTGCAGG (SEQ ID NO: 882). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей нуклеазу, и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR AAV2, последовательности, кодирующей промотор CK8e, последовательности, кодирующей нуклеазу SpCas9, и второго ITR AAV2. В некоторых вариантах осуществления, конструкция, включающая или состоящая из, от 5' к 3', первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей нуклеазу, и второго ITR, дополнительно включает поли(A) последовательность. В некоторых вариантах осуществления поли(A) последовательность включает или состоит из мини-поли(A) последовательности. Примеры мини-поли(A) последовательностей согласно настоящему описанию включают или состоят из нуклеотидной последовательности TAGCAATAAAGGATCGTTTATTTTCATTGGAAGCGTGTGTTGGTTTTTTGATCAGGCGCG (SEQ ID NO: 903). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей нуклеазу, поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей нуклеазу, мини-поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR AAV2, последовательности, кодирующей

промотор CK8e, последовательности, кодирующей нуклеазу SpCas9, мини-поли(A) последовательности и второго ITR AAV2. В некоторых вариантах осуществления конструкция, включающая, от 5' к 3', первый ITR, последовательность, кодирующую промотор, последовательность, кодирующую нуклеазу, поли(A) последовательность и второй ITR, дополнительно включает по меньшей мере один сигнал ядерной локализации. В некоторых вариантах осуществления, конструкция, включающая, от 5' к 3', первый ITR, последовательность, кодирующую промотор, последовательность, кодирующую нуклеазу, поли(A) последовательность и второй ITR, дополнительно включает по меньшей мере два сигнала ядерной локализации. Примеры сигналов ядерной локализации согласно настоящему описанию включают или состоят из нуклеотидной последовательности AAGCGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAA (SEQ ID NO. 887) или нуклеотидной последовательности ATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCC (SEQ ID NO. 885). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей первый сигнал ядерной локализации, последовательности, кодирующей нуклеазу, поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей первый сигнал ядерной локализации, последовательности, кодирующей нуклеазу, последовательности, кодирующей второй сигнал ядерной локализации, поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR AAV2, последовательности, кодирующей промотор CK8e, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 885, последовательности, кодирующей нуклеазу SpCas9, мини-поли(A) последовательности и второго ITR AAV2. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR AAV2, последовательности, кодирующей промотор CK8e,

последовательности, кодирующей нуклеазу SpCas9, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 887, мини-поли(A) последовательности и второго ITR AAV2. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR AAV2, последовательности, кодирующей промотор CK8e, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 885, последовательности, кодирующей нуклеазу SpCas9, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 887, мини-поли(A) последовательности и второго ITR AAV2. В некоторых вариантах осуществления конструкция, включающая, от 5' к 3', первый ITR, последовательность, кодирующую промотор, последовательность, кодирующую первый сигнал ядерной локализации, последовательность, кодирующую нуклеазу, последовательность, кодирующую второй сигнал ядерной локализации, поли(A) последовательность и второй ITR, дополнительно включает стоп-кодон. Стоп-кодон может иметь последовательность TAG (SEQ ID NO: 904), TAA (SEQ ID NO: 905) или TGA (SEQ ID NO: 906). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей первый сигнал ядерной локализации, последовательности, кодирующей нуклеазу, последовательности, кодирующей второй сигнал ядерной локализации, стоп-кодона, поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR AAV2, последовательности, кодирующей промотор CK8e, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 885, последовательности, кодирующей нуклеазу SpCas9, стоп-кодона, мини-поли(A) последовательности и второго ITR AAV2. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR AAV2, последовательности, кодирующей промотор CK8e, последовательности, кодирующей нуклеазу SpCas9, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 887, стоп-кодона, мини-поли(A) последовательности и второго ITR AAV2. В некоторых вариантах осуществления конструкция

включает или состоит из, от 5' к 3', первого ITR AAV2, последовательности, кодирующей промотор CK8e, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 885, последовательности, кодирующей нуклеазу SpCas9, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 887, стоп-кодона, мини-поли(A) последовательности и второго ITR AAV2. В некоторых вариантах осуществления конструкция, включающая или состоящая из, от 5' к 3', первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей первый сигнал ядерной локализации, последовательности, кодирующей нуклеазу, последовательности, кодирующей второй сигнал ядерной локализации, стоп-кодона, поли(A) последовательности и второго ITR, дополнительно включает инвертированные повторы транспозируемого элемента. Примеры инвертированных повторов транспозируемого элемента согласно настоящему описанию включают или состоят из нуклеотидной последовательности TGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAGTTTTTAAGGATTATTTAG GGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGTAAGCTAATACGAAAATTAAAAAT GACAAAATAGTTTGGAACTAGATTTCACTTATCTGGTT (SEQ ID NO. 907) и/или нуклеотидной последовательности GAATATAGTCTTTACCATGCCCTTGGCCACGCCCCTCTTTAATACGACGGGCAATTTGCACTTC AGAAAATGAAGAGTTTGCTTTAGCCATAACAAAAGTCCAGTATGCTTTTTCACAGCATAACTGG ACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGATCTATT TTGTTCAGTTTAAGACTTTATTGTCCGCCCACA (SEQ ID NO. 908). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей первый сигнал ядерной локализации, последовательности, кодирующей нуклеазу, последовательности, кодирующей второй сигнал ядерной локализации, стоп-кодона, поли(A) последовательности, второго ITR и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, имеющего последовательность SEQ ID NO: 907, первого ITR AAV2, последовательности, кодирующей

промотор CK8e, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 885, последовательности, кодирующей нуклеазу SpCas9, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 887, стоп-кодона, мини-поли(A) последовательности, второго ITR AAV2 и второго инвертированного повтора транспозируемого элемента, имеющего последовательность SEQ ID NO. 908. В некоторых вариантах осуществления, конструкция, включающая или состоящая из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей первый сигнал ядерной локализации, последовательности, кодирующей нуклеазу, последовательности, кодирующей второй сигнал ядерной локализации, стоп-кодона, поли(A) последовательности, второго ITR и второго инвертированного повтора транспозируемого элемента, дополнительно включают регуляторную последовательность. Примеры регуляторных последовательностей согласно настоящему описанию включают или состоят из нуклеотидной последовательности

CATGCAAGCTGTAGCCAACCACTAGAACTATAGCTAGAGTCCTGGGCGAACAAACGATGCTCGC
CTTCCAGAAAACCGAGGATGCGAACCACTTCATCCGGGGTCAGCACCACCGGCAAGCGCCGCGA
CGGCCGAGGTCTTCCGATCTCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGTAGAAG
AACAGCAAGGCCGCCAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCAGCC
AGGACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACG
GATGAAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCGT
ATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGG
TTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGC
GCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGCAACGA
TGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATTCGC
ACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTG
AGTTCGGAGACGTAGCCACCTACTCCCAACATCAGCCGGACTCCGATTACCTCGGGAACTTGCT
CCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTCGCG
GCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGTGAGATCTATATCTATGATCTCGCAGTCT
CCGGCGAGCACCGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAA
CGCGCTTGGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTC
TATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCCACCT

AACAATTCGTTCAAGCCGAGATCGGCTTCCCGGCCGCGGAGTTGTTCGGTAAATTGTCACAACG CCG (SEQ ID NO. 909). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, последовательности, кодирующей промотор, последовательности, кодирующей первый сигнал ядерной локализации, последовательности, кодирующей нуклеазу, последовательности, кодирующей второй сигнал ядерной локализации, стоп-кодона, поли(A) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, имеющего последовательность SEQ ID NO: 907, первого ITR AAV2, последовательности, кодирующей промотор CK8e, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 885, последовательности, кодирующей нуклеазу SpCas9, сигнала ядерной локализации, имеющего последовательность SEQ ID NO: 887, стоп-кодона, мини-поли(A) последовательности, второго ITR AAV2, регуляторной последовательности, имеющей последовательность SEQ ID NO: 909, и второго инвертированного повтора транспозируемого элемента, имеющего последовательность SEQ ID NO: 908. В некоторых вариантах осуществления конструкция может дополнительно включать одну или более спейсерных последовательностей. Примерные спейсерные последовательности согласно настоящему описанию имеют длину от 1-1500 нуклеотидов, включая все промежуточные диапазоны. В некоторых вариантах осуществления спейсерные последовательности могут быть расположены 5' или 3' относительно ITR, промотора, последовательности ядерной локализации, нуклеазы, стоп-кодона, поли(A) последовательности, инвертированного повтора транспозируемого элемента и/или регуляторного элемента. В некоторых вариантах осуществления конструкция может иметь последовательность, включающую или состоящую из SEQ ID NO: 899, SEQ ID NO: 900, SEQ ID NO: 901 или SEQ ID NO: 902.

E. AAV-sgРНК векторы

В некоторых вариантах осуществления по меньшей мере первая последовательность, кодирующая gРНК, и вторая последовательность, кодирующая gРНК, могут быть упакованы в AAV вектор. В некоторых вариантах осуществления по меньшей мере первая последовательность, кодирующая gРНК, вторая последовательность, кодирующая gРНК, и третья последовательность, кодирующая gРНК, могут быть упакованы в AAV вектор. В некоторых вариантах осуществления по меньшей мере первая последовательность, кодирующая gРНК, вторая последовательность, кодирующая gРНК, третья последовательность, кодирующая gРНК, и четвертая последовательность, кодирующая gРНК, могут быть упакованы в AAV вектор. В некоторых вариантах осуществления по меньшей мере первая последовательность, кодирующая gРНК, вторая последовательность, кодирующая gРНК, третья последовательность, кодирующая gРНК, четвертая последовательность, кодирующая gРНК, и пятая последовательность, кодирующая gРНК, могут быть упакованв в AAV вектор. В некоторых вариантах осуществления множество последовательностей, кодирующих gРНК, упаковано в AAV вектор. Например, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 или 20 последовательностей, кодирующих gРНК, могут быть упакованы в AAV вектор. В некоторых вариантах осуществления каждая последовательность, кодирующая gРНК, отличается. В некоторых вариантах осуществления по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 или 20 из последовательностей, кодирующих gРНК, являются одинаковыми. В некоторых вариантах осуществления все последовательности, кодирующие gРНК, являются одинаковыми.

В некоторых вариантах осуществления AAV вектор является AAV вектором дикого типа. В некоторых вариантах осуществления AAV вектор содержит одну или более мутаций. В некоторых вариантах осуществления AAV вектор выделен или получен из AAV вектора серотипа AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 или их любой комбинации.

Примерные AAV-sgРНК векторы содержат две последовательности ITR (инвертированных концевых повторов), которые фланкируют

центральную область последовательности, включающую последовательности sgРНК. В некоторых вариантах осуществления ITR выделены или получены из AAV вектора серотипа AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 или их любой комбинации. В некоторых вариантах осуществления ITR выделены или получены из AAV вектора первого серотипа, а последовательность, кодирующая капсидный белок AAV-sgРНК вектора, выделена или получена из AAV вектора второго серотипа. В некоторых вариантах осуществления первый серотип и второй серотип являются одинаковыми. В некоторых вариантах осуществления первый серотип и второй серотип не являются одинаковыми. В некоторых вариантах осуществления первым серотипом является AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 или AAV11. В некоторых вариантах осуществления вторым серотипом является AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 или AAV11. В некоторых вариантах осуществления первым серотипом является AAV2, а вторым серотипом является AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 или AAV11. В некоторых вариантах осуществления первым серотипом является AAV2, а вторым серотипом является AAV9.

Примерные AAV-sgРНК векторы содержат две последовательности ITR (инвертированных концевых повторов), которые фланкируют центральную область последовательности, включающую последовательности gРНК. В некоторых вариантах осуществления ITR выделены или получены из AAV вектора серотипа AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 или их любой комбинации. В некоторых вариантах осуществления первый ITR выделен или получен из AAV вектора первого серотипа, второй ITR выделен или получен из AAV вектора второго серотипа, и последовательность, кодирующая капсидный белок AAV-sgРНК вектора, выделена или получена из AAV вектора третьего серотипа. В некоторых вариантах осуществления первый серотип и второй серотип являются одинаковыми. В некоторых вариантах осуществления первый серотип и второй серотип не являются одинаковыми. В некоторых вариантах осуществления первый серотип, второй серотип и третий серотип являются одинаковыми. В

некоторых вариантах осуществления первый серотип, второй серотип и третий серотип не являются одинаковыми. В некоторых вариантах осуществления первым серотипом являются AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 или AAV11. В некоторых вариантах осуществления вторым серотипом являются AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 или AAV11. В некоторых вариантах осуществления третьим серотипом являются AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 или AAV11. В некоторых вариантах осуществления первым серотипом является AAV2, вторым серотипом является AAV4, и третьим серотипом являются AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 или AAV11. В некоторых вариантах осуществления первым серотипом является AAV2, вторым серотипом является AAV4, и третьим серотипом является AAV9. Примерные AAV-sgРНК векторы содержат две последовательности ITR (инвертированных концевых повторов), которые фланкируют центральную область последовательности, включающую последовательности sgРНК. В некоторых вариантах осуществления ITR выделены или получены из AAV вектора серотипа AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 или их любой комбинации. В некоторых вариантах осуществления ITR включают или состоят из полноразмерных последовательностей и/или последовательностей дикого типа серотипа AAV. В некоторых вариантах осуществления ITR включают или состоят из усеченных последовательностей серотипа AAV. В некоторых вариантах осуществления ITR включают или состоят из удлиненных последовательностей серотипа AAV. В некоторых вариантах осуществления ITR включают или состоят из последовательностей, включающих вариант последовательности по сравнению с последовательностью дикого типа такого же серотипа AAV. В некоторых вариантах осуществления вариант последовательности включает одно или более из замены, делеции, вставки, инверсии или транспозиции. В некоторых вариантах осуществления ITR включают или состоят из по меньшей мере 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139,

140, 141, 142, 143, 144, 145, 146, 147, 148, 149 или 150 пар оснований. В некоторых вариантах осуществления ITR включают или состоят из 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149 или 150 пар оснований. В некоторых вариантах осуществления ITR имеют длину 110±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 120±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 130±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 140±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 150±10 пар оснований. В некоторых вариантах осуществления ITR имеют длину 115, 145 или 141 пару оснований. В некоторых вариантах осуществления ITR имеют последовательность, выбранную из SEQ ID NO: 880, SEQ ID NO: 881, SEQ ID NO: 882 или SEQ ID NO: 883.

В некоторых вариантах осуществления AAV-sgРНК вектор может включать дополнительные элементы, которые облегчают упаковку вектора и экспрессию sgРНК. В некоторых вариантах осуществления AAV-sgРНК вектор может включать транспозируемый элемент. В некоторых вариантах осуществления AAV-sgРНК вектор может включать регуляторный элемент. В некоторых вариантах осуществления регуляторный элемент включает активатор или репрессор. В некоторых вариантах осуществления последовательность AAV-sgРНК может включать нефункциональную последовательность или "спейсерную" последовательность. Примерные спейсерные последовательности согласно настоящему описанию могут обладать некоторой (отличным от нуля процентом) идентичностью или гомологией с геномной последовательностью млекопитающего (включая человека). В альтернативе примерные спейсерные последовательности согласно настоящему описанию могут не обладать идентичностью или гомологией с геномной последовательностью млекопитающего (включая человека). Примерные спейсерные последовательности согласно настоящему описанию могут включать или состоять из природных некодирующих

последовательностей или последовательностей, которые ни транскрибируются, ни транслируются после введения AAV вектора субъекту.

В некоторых вариантах осуществления AAV-sgРНК вектор может быть оптимизирован для продукции в дрожжах, бактериях, клетках насекомых или клетках млекопитающих. В некоторых вариантах осуществления AAV-sgРНК вектор может быть оптимизирован для экспрессии в клетках человека. В некоторых вариантах осуществления AAV-Cas9 вектор может быть оптимизирован для экспрессии в системе бакуловирусной экспрессии.

В некоторых вариантах осуществления AAV-sgРНК вектор включает по меньшей мере один промотор. В некоторых вариантах осуществления AAV-sgРНК вектор включает по меньшей мере два промотора. В некоторых вариантах осуществления AAV-sgРНК вектор включает по меньшей мере три промотора. В некоторых вариантах осуществления AAV-sgРНК вектор включает по меньшей мере четыре промотора. В некоторых вариантах осуществления AAV-sgРНК вектор включает по меньшей мере пять промоторов. Примеры промоторов включают, например, легкую цепь иммуноглобулина, тяжелую цепь иммуноглобулина, Т-клеточный рецептор, HLA DQ α и/или DQ β, β-интерферон, интерлейкин-2, рецептор интерлейкина-2, МНС класса II 5, МНС класса II HLA-Dra, β-актин, мышечную креатинкиназу (MCK), преальбумин (транстиретин), эластазу I, металлотионеин (MTII), коллагеназу, альбумин, α-фетопротеин, t-глобин, β-глобин, c-fos, c-HA-ras, инсулин, молекулу адгезии нервных клеток (NCAM), α_1-антитрипаин, гистон H2B (TH2B), коллаген мыши и/или I типа, регулируемые глюкозой белки (GRP94 и GRP78), гормон роста крысы, человеческий сывороточный амилоид А (SAA), тропонин I (TN I), тромбоцитарный фактор роста (PDGF), миодистрофию Дюшенна, SV40, полиому, ретровирусы, вирус папилломы, вирус гепатита В, вирус иммунодефицита человека, цитомегаловирус (ЦМВ) и вирус лейкоза гиббонов. Другие примеры промоторов включают промотор U6, промотор H1 и промотор 7SK.

В некоторых вариантах осуществления последовательность, кодирующая gРНК, включает последовательность, выбранную из SEQ ID NO: 383-705, 709-711, 715-717, 790-862 и 864.

В некоторых вариантах осуществления AAV вектор включает первую последовательность, кодирующую gРНК, и вторую последовательность, кодирующую gРНК, первый промотор направляет экспрессию первой последовательности, кодирующей gРНК, а второй промотор направляет экспрессию второй последовательности, кодирующей gРНК. В некоторых вариантах осуществления первый и второй промоторы являются одинаковыми. В некоторых вариантах осуществления первый и второй промоторы являются разными. В некоторых вариантах осуществления первый и второй промоторы выбраны из промотора H1, промотора U6 и промотора 7SK. В некоторых вариантах осуществления первая последовательность, кодирующая gРНК, и вторая последовательность, кодирующая gРНК, являются идентичными. В некоторых вариантах осуществления первая последовательность, кодирующая gРНК, и вторая последовательность, кодирующая gРНК, не являются идентичными.

В некоторых вариантах осуществления AAV вектор включает первую последовательность, кодирующую gРНК, вторую последовательность, кодирующую gРНК, и третью последовательность, кодирующую gРНК, первый промотор направляет экспрессию первой последовательности, кодирующей gРНК, второй промотор направляет экспрессию второй последовательности, кодирующей gРНК, и третий промотор направляет экспрессию третьей последовательности, кодирующей gРНК. В некоторых вариантах осуществления по меньшей мере два из первого, второго и третьего промоторов являются одинаковыми. В некоторых вариантах осуществления каждый из первого, второго и третьего промоторов является различным. В некоторых вариантах осуществления первый, второй и третий промоторы выбраны от промотора H1, промотора U6 и промотора 7SK. В некоторых вариантах осуществления первый промотор является промотором U6. В некоторых вариантах осуществления второй промотор является промотором H1. В некоторых вариантах осуществления третий промотор является промотором 7SK. В некоторых вариантах осуществления первый промотор является промотором U6, второй промотор является промотором H1, и третий промотор является промотором 7SK. В некоторых вариантах осуществления первая последовательность,

кодирующая gРНК, вторая последовательность, кодирующая gРНК, и третья последовательность, кодирующая gРНК, являются идентичными. В некоторых вариантах осуществления первая последовательность, кодирующая gРНК, вторая последовательность, кодирующая gРНК, и третья последовательность, кодирующая gРНК, не являются идентичными.

В некоторых вариантах осуществления AAV вектор включает первую последовательность, кодирующую gРНК, вторую последовательность, кодирующую gРНК, третью последовательность, кодирующую gРНК, и четвертую последовательность, кодирующую gРНК, первый промотор направляет экспрессию первой последовательности, кодирующей gРНК, второй промотор направляет экспрессию второй последовательности, кодирующей gРНК, третий промотор направляет экспрессию третьей последовательности, кодирующей gРНК, и четвертый промотор направляет экспрессию четвертой последовательности, кодирующей gРНК. В некоторых вариантах осуществления по меньшей мере два из первого, второго, третьего и четвертого промоторов являются одинаковыми. В некоторых вариантах осуществления каждый из первого, второго, третьего и четвертого промоторов является различным. В некоторых вариантах осуществления каждый из первого, второго, третьего и четвертого промоторов выбран от промотора H1, промотора U6 и промотора 7SK. В некоторых вариантах осуществления первая последовательность, кодирующая gРНК, вторая последовательность, кодирующая gРНК, третья последовательность, кодирующая gРНК, и четвертая последовательность, кодирующая gРНК, являются идентичными. В некоторых вариантах осуществления первая последовательность, кодирующая gРНК, вторая последовательность, кодирующая gРНК, третья последовательность, кодирующая gРНК, и четвертая последовательность, кодирующая gРНК, не являются идентичными.

В некоторых вариантах осуществления AAV вектор включает первую последовательность, кодирующую gРНК, вторую последовательность, кодирующую gРНК, третью последовательность, кодирующую gРНК, четвертую последовательность, кодирующую gРНК, и пятую последовательность, кодирующую gРНК, первый промотор

направляет экспрессию первой последовательности, кодирующей gРНК, второй промотор направляет экспрессию второй последовательности, кодирующей gРНК, третий промотор направляет экспрессию третьей последовательности, кодирующей gРНК, четвертый промотор направляет экспрессию четвертой последовательности, кодирующей gРНК, и пятый промотор направляет экспрессию пятой последовательности, кодирующей gРНК. В некоторых вариантах осуществления по меньшей мере два из первого, второго, третьего, четвертого и пятого промоторов являются одинаковыми. В некоторых вариантах осуществления каждый из первого, второго, третьего, четвертого и пятого промоторов является различным. В некоторых вариантах осуществления каждый из первого, второго, третьего и четвертого промоторов является различным. В некоторых вариантах осуществления каждый из первого, второго, третьего, четвертого и пятого промоторов выбран от промотора H1, промотора U6 и промотора 7SK. В некоторых вариантах осуществления первая последовательность, кодирующая gРНК, вторая последовательность, кодирующая gРНК, третья последовательность, кодирующая gРНК, четвертая последовательность, кодирующая gРНК, и пятая последовательность, кодирующая gРНК, являются идентичными. В некоторых вариантах осуществления первая последовательность, кодирующая gРНК, вторая последовательность, кодирующая gРНК, третья последовательность, кодирующая gРНК, четвертая последовательность, кодирующая gРНК, и пятая последовательность, кодирующая gРНК, не являются идентичными.

В некоторых вариантах осуществления AAV-sgРНК вектор включает последовательность, выбранную из SEQ ID NO: 910, SEQ ID NO: 911, SEQ ID NO: 912 или SEQ ID NO: 913. В некоторых вариантах осуществления AAV-sgРНК вектор включает последовательность, выбранную из SEQ ID NO: 914, SEQ ID NO: 915, SEQ ID NO: 916 или SEQ ID NO: 917. В некоторых вариантах осуществления AAV-sgРНК вектор включает последовательность, выбранную из SEQ ID NO: 918, SEQ ID NO: 919, SEQ ID NO: 920 или SEQ ID NO: 921. Примеры AAV-sgРНК векторов представлены в Таблице 5.

Таблица 5. Примеры AAV-sgРНК векторов.

SEQ ID NO.	Последовательность
910	CCACTCCCTCTATGCGCGCTCGCTCACTCACTCGGCCCTGGAGACCAAAGGTCTCCAGAC TGCCGGCCTCTGGCCGGCAGGGCCGAGTGAGTGAGCGAGCGCGCATAGAGGGAGTGGGTA CCTCCATCATCTAGGTTTGCCAGATCTGATATCGGCGCGCCCCTGggcgcgccCGAGTCC AACACCCGTGGGAATCCCATGGGCACCATGGCCCCTCGCTCCAAAAATGCTTTCGCGTCG CGCAGACACTGCTCGGTAGTTTCGGGGATCAGCGTTTGAGTAAGAGCCCGCGTCTGAACC CTCCGCGCCGCCCCGGCCCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAG CGTGACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGAT TCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGAC TGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTA GTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAA GTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCACTAGAGTA ACAGTCTGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGC GCCCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTG GATTTGGGAATCTTATAAGTTCTGTATGAGACCACGGTACACCGCACTAGAGTAACAGTC TGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCTGACGGCGCGCCCTG CAGTATTTAGCATGCCCCACCCATCTGCAAGGCATTCTGGATAGTGTCAAAACAGCCGGA AATCAAGTCCGTTTATCTCAAACTTTAGCATTTTGGGAATAAATGATATTTGCTATGCTG GTTAAATTAGATTTTAGTTAAATTTCCTGCTGAAGCTCTAGTACGATAAGTAACTTGACC TAAGTGTAAAGTTGAGATTTCCTTCAGGTTTATATAGCTTGTGCGCCGCCTGGGTACACC GCACTAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCC GCGCTTGCGGCCGCCTCGAGTGATCAAAAAAACCAACACACGCTTCCAATGAAAATAAAC GATCCTTTATTGCTAGCCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGG CGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGG CGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAACACGGGGCCGTCGCCGATGGGGG TGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCT TGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCTGT TGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCT TCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGT AGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGG ACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGT AGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGA ACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACT TGTGGCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCT CCTCGCCCTTGCTCACCATGGTGGCGACCGGTggatcccgggcccgcggggtggctttac caacagtacccggaatgccaagcttacttagatcgcagtctcgacGCTGGCTGGCTCCTG AGTGTCTGTCTGTGCTGTGGAGGTGGTGGTAGAATGAGGGCAGCCCCTGTGCCCCTGGGT TATATAGAGGAGCCTACAGGGTGTGACTAGCCAGGAGGGGCTGTCCCCAGGGAGGGGCCC CTGAGAGCAGATGAGCTTTCAGCTCGTTGCCCGGGCACCGTGCCCACCCCGGACCCAGGC GTGCAGCTTGCCCAGCCCCATGGCCTTGTATGGGCTGCCCCAAGGGCTGACTTGCTCACT GGTTCCTAAACTAAGTGCTGAGTCTAGCTGGCGGGGGACAGCTGGCCCTTCGCCGGGAAC ATGGCATGCAGGGTTATTTTTAGAGGCAGCAGGTGTTGGGGGGGGGGGGCAGCCACATG TCTGGGTTAATTATAACCAGGCATCTCGGGTGTCCCCAGGCCTTGCCTCCTTACATGGGC ACgtcgacGATATCAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCG CGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCG CCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGG

918	CCACTCCCTCTATGCGCGCTCGCTCACTCACTCGGCCCTGGAGACCAAAGGTCTCCAGAC TGCCGGCCTCTGGCCGGCAGGGCCGAGTGAGTGAGCGAGCGCGCATAGAGGGAGTGGGTA CCTCCATCATCTAGGTTTGCCAGATCTGATATCGGCGCGCCCCTGggcgcgccCGAGTCC AACACCCGTGGGAATCCCATGGGCACCATGGCCCCTCGCTCCAAAAATGCTTTCGCGTCG CGCAGACACTGCTCGGTAGTTTCGGGGATCAGCGTTTGAGTAAGAGCCCGCGTCTGAACC CTCCGCGCCGCCCCGGCCCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAG CGTGACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGAT TCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGAC TGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTA GTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAA GTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCACCAGAGTA ACAGTCTGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGC GCCCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTG GATTTGGGAATCTTATAAGTTCTGTATGAGACCACGGTACACCGCACCAGAGTAACAGTC TGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCTGACGGCGCGCCCTG CAGTATTTAGCATGCCCCACCCATCTGCAAGGCATTCTGGATAGTGTCAAAACAGCCGGA AATCAAGTCCGTTTATCTCAAACTTTAGCATTTTGGGAATAAATGATATTTGCTATGCTG GTTAAATTAGATTTTAGTTAAATTTCCTGCTGAAGCTCTAGTACGATAAGTAACTTGACC TAAGTGTAAAGTTGAGATTTCCTTCAGGTTTATATAGCTTGTGCGCCGCCTGGGTACACC GCACCAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCC GCGCTGTTTAAACcaacaaaatcagcagctaatgaaggcaagtcagcaggtcactcatca ttttccacttcggcaatgcagtgggattattccaacagaggtttttcacagcattccttc agttaactggagatcgaatcttgattttcacagatatacttggcaaggtccgccctgtca tcagcacattcaagcagatctccatggcagcattccgtgtggacttaggtaagatctgtc actaacttggaaacttctgcaaactcagcttagggaaatctctggctcaggcgagctact gcctcagcttagaaagctctttctccaaattattggagactggcacacttaagtccctgt taggcagacgaagccttcccttcatcccgaagttcatcgagctttggcaacaggcaggca gctttatcagcagcttggcaacattctgtaaaagcagctttatacctttaagcaaagaaa aggagttccggggcataaaagtaaggatgtcttctggcaatttataataagtatttttc aaaaatgtctcttcattgtcatgaaaagcagtgcatcacacatcaacctctggtctcacc aatcgggggaggtttgggttgtttacttagtgttgcaagaattattttattctctcaggt tcttgttttgcacagcagtcagctcattcaccataggtttcacgaagagttgcTGCGGCC GCCTCGATAATACGACTCACTATAGGGTCGACGATATCAGATCTAGGAACCCCTAGTGAT GGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCC CGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGG AGTGG
914	CCACTCCCTCTATGCGCGCTCGCTCACTCACTCGGCCCTGGAGACCAAAGGTCTCCAGAC TGCCGGCCTCTGGCCGGCAGGGCCGAGTGAGTGAGCGAGCGCGCATAGAGGGAGTGGGTA CCTCCATCATCTAGGTTTGCCAGATCTGATATCGGCGCGCCCCTGggcgcgccCGAGTCC AACACCCGTGGGAATCCCATGGGCACCATGGCCCCTCGCTCCAAAAATGCTTTCGCGTCG CGCAGACACTGCTCGGTAGTTTCGGGGATCAGCGTTTGAGTAAGAGCCCGCGTCTGAACC CTCCGCGCCGCCCCGGCCCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAG CGTGACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGAT TCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGAC TGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTA GTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAA GTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCACCAGAGTA ACAGTCTGAGGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGC GCCCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTG

	GATTTGGGAATCTTATAAGTTCTGTATGAGACCACGGTACACCGCACCAGAGTAACAGTC TGAGGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCTGACGGCGCGCCCTG CAGTATTTAGCATGCCCCACCCATCTGCAAGGCATTCTGGATAGTGTCAAAACAGCCGGA AATCAAGTCCGTTTATCTCAAACTTTAGCATTTTGGGAATAAATGATATTTGCTATGCTG GTTAAATTAGATTTTAGTTAAATTTCCTGCTGAAGCTCTAGTACGATAAGTAACTTGACC TAAGTGTAAAGTTGAGATTTCCTTCAGGTTTATATAGCTTGTGCGCCGCCTGGGTACACC GCACCAGAGTAACAGTCTGAGGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCC GCGCTGTTTAAACcaacaaaatcagcagctaatgaaggcaagtcagcaggtcactcatca ttttccacttcggcaatgcagtgggattattccaacagaggtttttcacagcattccttc agttaactggagatcgaatcttgattttcacagatatacttggcaaggtccgccctgtca tcagcacattcaagcagatctccatggcagcattccgtgtggacttaggtaagatctgtc actaacttggaaacttctgcaaactcagcttagggaaatctctggctcaggcgagctact gcctcagcttagaaagctctttctccaaattattggagactggcacacttaagtccctgt taggcagacgaagccttcccttcatcccgaagttcatcgagctttggcaacaggcaggca gctttatcagcagcttggcaacattctgtaaaagcagctttatacctttaagcaaagaaa aggagttccggggcataaaagtaaggatgtcttctggcaatttataataagtatttttc aaaaatgtctcttcattgtcatgaaaagcagtgcatcacacatcaacctctggtctcacc aatcgggggaggtttgggttgtttacttagtgttgcaagaattattttattctctcaggt tcttgttttgcacagcagtcagctcattcaccataggtttcacgaagagttgcTGCGGCC GCCTCGATAATACGACTCACTATAGGGTCGACGATATCAGATCTAGGAACCCCTAGTGAT GGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCC CGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGG AGTGG
911	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTC GGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTCCTAGATCTGATATCGGCGCGCCCCTGGGCGCGCCCGAGTCC AACACCCGTGGGAATCCCATGGGCACCATGGCCCCTCGCTCCAAAAATGCTTTCGCGTCG CGCAGACACTGCTCGGTAGTTTCGGGGATCAGCGTTTGAGTAAGAGCCCGCGTCTGAACC CTCCGCGCCGCCCCGGCCCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAG CGTGACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGAT TCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGAC TGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTA GTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAA GTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCACTAGAGTA ACAGTCTGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGC GCCCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTG GATTTGGGAATCTTATAAGTTCTGTATGAGACCACGGTACACCGCACTAGAGTAACAGTC TGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCTGACGGCGCGCCCTG CAGTATTTAGCATGCCCCACCCATCTGCAAGGCATTCTGGATAGTGTCAAAACAGCCGGA AATCAAGTCCGTTTATCTCAAACTTTAGCATTTTGGGAATAAATGATATTTGCTATGCTG GTTAAATTAGATTTTAGTTAAATTTCCTGCTGAAGCTCTAGTACGATAAGTAACTTGACC TAAGTGTAAAGTTGAGATTTCCTTCAGGTTTATATAGCTTGTGCGCCGCCTGGGTACACC GCACTAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCC GCGCTTGCGGCCGCCTCGAGTGATCAAAAAAACCAACACACGCTTCCAATGAAAATAAAC GATCCTTTATTGCTAGCCTTTACTTGTACAGCTCGTCCATGCCGAGAGTGATCCCGGCGG CGGTCACGAACTCCAGCAGGACCATGTGATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGG CGGACTGGGTGCTCAGGTAGTGGTTGTCGGGCAGCAACACGGGGCCGTCGCCGATGGGGG TGTTCTGCTGGTAGTGGTCGGCGAGCTGCACGCTGCCGTCCTCGATGTTGTGGCGGATCT

	TGAAGTTCACCTTGATGCCGTTCTTCTGCTTGTCGGCCATGATATAGACGTTGTGGCTGT TGTAGTTGTACTCCAGCTTGTGCCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCT TCAGCTCGATGCGGTTCACCAGGGTGTCGCCCTCGAACTTCACCTCGGCGCGGGTCTTGT AGTTGCCGTCGTCCTTGAAGAAGATGGTGCGCTCCTGGACGTAGCCTTCGGGCATGGCGG ACTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTAGCGGCTGAAGCACTGCACGCCGT AGGTCAGGGTGGTCACGAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGA ACTTCAGGGTCAGCTTGCCGTAGGTGGCATCGCCCTCGCCCTCGCCGGACACGCTGAACT TGTGGCCGTTTACGTCGCCGTCCAGCTCGACCAGGATGGGCACCACCCCGGTGAACAGCT CCTCGCCCTTGCTCACCATGGTGGCGACCGGTGGATCCCGGGCCCGCGGGGTGGCTTTAC CAACAGTACCCGGAATGCCAAGCTTACTTAGATCGCAGTCTCGACGCTGGCTGGCTCCTG AGTGTCTGTCTGTGCTGTGGAGGTGGTGGTAGAATGAGGGCAGCCCCTGTGCCCCTGGGT TATATAGAGGAGCCTACAGGGTGTGACTAGCCAGGAGGGGCTGTCCCCAGGGAGGGGCCC CTGAGAGCAGATGAGCTTTCAGCTCGTTGCCCGGGCACCGTGCCCACCCCGGACCCAGGC GTGCAGCTTGCCCAGCCCCATGGCCTTGTATGGGCTGCCCCAAGGGCTGACTTGCTCACT GGTTCCTAAACTAAGTGCTGAGTCTAGCTGGCGGGGGACAGCTGGCCCTTCGCCGGGAAC ATGGCATGCAGGGTTATTTTTAGAGGCAGCAGGTGTTGGGGGGGGGGGGGCAGCCACATG TCTGGGTTAATTATAACCAGGCATCTCGGGTGTCCCCAGGCCTTGCCTCCTTACATGGGC ACGTCGACGATATCAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCG CGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCG GGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG
919	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTC GGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTCCTAGATCTGATATCGGCGCGCCCCTGggcgcgccCGAGTCC AACACCCGTGGGAATCCCATGGGCACCATGGCCCCTCGCTCCAAAAATGCTTTCGCGTCG CGCAGACACTGCTCGGTAGTTTCGGGGATCAGCGTTTGAGTAAGAGCCCGCGTCTGAACC CTCCGCGCCGCCCCGGCCCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAG CGTGACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGAT TCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGAC TGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTA GTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAA GTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCACCAGAGTA ACAGTCTGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGC GCCCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTG GATTTGGGAATCTTATAAGTTCTGTATGAGACCACGGTACACCGCACCAGAGTAACAGTC TGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCTGACGGCGCGCCCTG CAGTATTTAGCATGCCCCACCCATCTGCAAGGCATTCTGGATAGTGTCAAAACAGCCGGA AATCAAGTCCGTTTATCTCAAACTTTAGCATTTTGGGAATAAATGATATTTGCTATGCTG GTTAAATTAGATTTTAGTTAAATTTCCTGCTGAAGCTCTAGTACGATAAGTAACTTGACC TAAGTGTAAAGTTGAGATTTCCTTCAGGTTTATATAGCTTGTGCGCCGCCTGGGTACACC GCACCAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCC GCGCTGTTTAAACcaacaaaatcagcagctaatgaaggcaagtcagcaggtcactcatca ttttccacttcggcaatgcagtgggattattccaacagaggtttttcacagcattccttc agttaactggagatcgaatcttgattttcacagatatacttggcaaggtccgccctgtca tcagcacattcaagcagatctccatggcagcattccgtgtggacttaggtaagatctgtc actaacttggaaacttctgcaaactcagcttagggaaatctctggctcaggcgagctact gcctcagcttagaaagctctttctccaaattattggagactggcacacttaagtccctgt taggcagacgaagccttcccttcatcccgaagttcatcgagctttggcaacaggcaggca gctttatcagcagcttggcaacattctgtaaaagcagctttatacctttaagcaaagaaa aggagttccggggcataaaagtaaggatgtcttctggcaatttataataagtatttttc aaaaatgtctcttcattgtcatgaaaagcagtgcatcacacatcaacctctggtctcacc

	aatcgggggaggtttgggttgtttacttagtgttgcaagaattattttattctctcaggt tcttgttttgcacagcagtcagctcattcaccataggtttcacgaagagttgcTGCGGCC GCCTCGATAATACGACTCACTATAGGGTCGACGATATCAGATCTAGGAACCCCTAGTGAT GGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGT CGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCT GCAGG
915	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTC GGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTCCTAGATCTGATATCGGCGCGCCCCTGggcgcgccCGAGTCC AACACCCGTGGGAATCCCATGGGCACCATGGCCCCTCGCTCCAAAAATGCTTTCGCGTCG CGCAGACACTGCTCGGTAGTTTCGGGGATCAGCGTTTGAGTAAGAGCCCGCGTCTGAACC CTCCGCGCCGCCCCGGCCCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAG CGTGACCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGAT TCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGAC TGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTA GTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAA GTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCACCAGAGTA ACAGTCTGAGGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGC GCCCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGTCTTTG GATTTGGGAATCTTATAAGTTCTGTATGAGACCACGGTACACCGCACCAGAGTAACAGTC TGAGGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCTGACGGCGCGCCCTG CAGTATTTAGCATGCCCCACCCATCTGCAAGGCATTCTGGATAGTGTCAAAACAGCCGGA AATCAAGTCCGTTTATCTCAAACTTTAGCATTTTGGGAATAAATGATATTTGCTATGCTG GTTAAATTAGATTTTAGTTAAATTTCCTGCTGAAGCTCTAGTACGATAAGTAACTTGACC TAAGTGTAAAGTTGAGATTTCCTTCAGGTTTATATAGCTTGTGCGCCGCCTGGGTACACC GCACCAGAGTAACAGTCTGAGGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCC GCGCTGTTTAAACcaacaaaatcagcagctaatgaaggcaagtcagcaggtcactcatca ttttccacttcggcaatgcagtgggattattccaacagaggtttttcacagcattccttc agttaactggagatcgaatcttgattttcacagatatacttggcaaggtccgccctgtca tcagcacattcaagcagatctccatggcagcattccgtgtggacttaggtaagatctgtc actaacttggaaacttctgcaaactcagcttagggaaatctctggctcaggcgagctact gcctcagcttagaaagctctttctccaaattattggagactggcacacttaagtccctgt taggcagacgaagccttcccttcatcccgaagttcatcgagctttggcaacaggcaggca gctttatcagcagcttggcaacattctgtaaaagcagctttatacctttaagcaaagaaa aggagttccggggcataaaagtaaggatgtcttctggcaatttataataagtatttttc aaaaatgtctcttcattgtcatgaaaagcagtgcatcacacatcaacctctggtctcacc aatcgggggaggtttgggttgtttacttagtgttgcaagaattattttattctctcaggt tcttgttttgcacagcagtcagctcattcaccataggtttcacgaagagttgcTGCGGCC GCCTCGATAATACGACTCACTATAGGGTCGACGATATCAGATCTAGGAACCCCTAGTGAT GGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGT CGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCT GCAGG
912	AAGATGACGGTTTGTCACATGGAGTTGGCAGGATGTTTGATTAAAAACATAACAGGAAGA AAAATGCCCCGCTGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAGTT TTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGTAAG CTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCACTTATCTGGTTCG GATCTCCTAGGCGATATCAGTGATCAcggatctcgaccaattgacattattgaagcaACT AGTATCGATtttatcagggttattgtctcagaCCTGCAGGCAGCTGCGCGCTCGCTCGCT CACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGT GAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTG

	ATATCGGCGCGCCCCTGGGCGCGCCCGAGTCCAACACCCGTGGGAATCCCATGGGCACCA TGGCCCCTCGCTCCAAAAATGCTTTCGCGTCGCGCAGACACTGCTCGGTAGTTTCGGGGA TCAGCGTTTGAGTAAGAGCCCGCGTCTGAACCCTCCGCGCCGCCCCGGCCCCAGTGGAAA GACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGACCGCGCGCCGAGCGCGCGCCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAG GCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAAT ACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAA ATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATC TTGTGGAAAGGACGAAACACCGCACTAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTTTTTGCGGCCGCGCTTCGGCGCGCCCATATTTGCATGTCGCTATGTGTT CTGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTAT GAGACCACGGTACACCGCACTAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGGAAACA GCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTGCGGCCGCTGACGGCGCGCCCTGCAGTATTTAGCATGCCCCACCCATCTG CAAGGCATTCTGGATAGTGTCAAAACAGCCGGAAATCAAGTCCGTTTATCTCAAACTTTA GCATTTTGGGAATAAATGATATTTGCTATGCTGGTTAAATTAGATTTTAGTTAAATTTCC TGCTGAAGCTCTAGTACGATAAGTAACTTGACCTAAGTGTAAAGTTGAGATTTCCTTCAG GTTTATATAGCTTGTGCGCCGCCTGGGTACACCGCACTAGAGTAACAGTCTGACGTTTAA GAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAA AAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTGCGGCCGCCTCGAGTGATCAAA AAAACCAACACACGCTTCCAATGAAAATAAACGATCCTTTATTGCTAGCCTTTACTTGTA CAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTG ATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTC GGGCAGCAACACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTG CACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTG CTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAG GATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTC GCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGT GCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTG GTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCA GGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGC ATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTC GACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGAC CGGTGGATCCCGGGCCCGCGGGGTGGCTTTACCAACAGTACCCGGAATGCCAAGCTTACT TAGATCGCAGTCTCGACGCTGGCTGGCTCCTGAGTGTCTGTCTGTGCTGTGGAGGTGGTG GTAGAATGAGGGCAGCCCCTGTGCCCCTGGGTTATATAGAGGAGCCTACAGGGTGTGACT AGCCAGGAGGGGCTGTCCCCAGGGAGGGGCCCCTGAGAGCAGATGAGCTTTCAGCTCGTT GCCCGGGCACCGTGCCCACCCCGGACCCAGGCGTGCAGCTTGCCCAGCCCCATGGCCTTG TATGGGCTGCCCCAAGGGCTGACTTGCTCACTGGTTCCTAAACTAAGTGCTGAGTCTAGC TGGCGGGGGACAGCTGGCCCTTCGCCGGGAACATGGCATGCAGGGTTATTTTTAGAGGCA GCAGGTGTTGGGGGGGGGGGGCAGCCACATGTCTGGGTTAATTATAACCAGGCATCTCG GGTGTCCCCAGGCCTTGCCTCCTTACATGGGCACGTCGACGATATCAGATCTaggaaccc ctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgggcga ccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcgc agctgcctgcaggCATGCAAGCTGTAGCCAACCACTAGAACTATAGCTAGAGTCCTGGGC GAACAAACGATGCTCGCCTTCCAGAAAACCGAGGATGCGAACCACTTCATCCGGGGTCAG CACCACCGGCAAGCGCCGCGACGGCCGAGGTCTTCCGATCTCCTGAAGCCAGGGCAGATC CGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGCCGCCAATGCCTGACGATGCGTGGA GACCGAAACCTTGCGCTCGTTCGCCAGCCAGGACAGAAATGCCTCGACTTCGCTGCTGCC CAAGGTTGCCGGGTGACGCACACCGTGGAAACGGATGAAGGCACGAACCCAGTTGACATA AGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAG AACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTCATGGCTTGTTATGA

	CTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGG TCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGCAACGATGTTACGCAGC AGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATTCGCACATGTA GGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTGAGT TCGGAGACGTAGCCACCTACTCCCAACATCAGCCGGACTCCGATTACCTCGGGAACTTGC TCCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTC TCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGTGAGATCTATATCTATGATC TCGCAGTCTCCGGCGAGCACCGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCCTCA AGCATGAGGCCAACGCGCTTGGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACG ATCCCGCAGTGGCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATA TCGACCCAAGTACCGCCACCTAACAATTCGTTCAAGCCGAGATCGGCTTCCCGGCCGCGG AGTTGTTCGGTAAATTGTCACAACGCCGCGAATATAGTCTTTACCATGCCCTTGGCCACG CCCCTCTTTAATACGACGGGCAATTTGCACTTCAGAAAATGAAGAGTTTGCTTTAGCCAT AACAAAAGTCCAGTATGCTTTTTCACAGCATAACTGGACTGATTTCAGTTTACAACTATT CTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGATCTATTTTGTTCAGTTTAAGACTTT ATTGTCCGCCCACACCCGCTTACGC
920	AAGATGACGGTTTGTCACATGGAGTTGGCAGGATGTTTGATTAAAAACATAACAGGAAGA AAAATGCCCCGCTGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAGTT TTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGTAAG CTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCACTTATCTGGTTCG GATCTCCTAGGCGATATCAGTGATCAcggatctcgaccaattgacattattgaagcaACT AGTATCGATttatcagggttattgtctcagaCCTGCAGGCAGCTGCGCGCTCGCTCGCT CACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGT GAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTG ATATCGGCGCGCCCCTGGGCGCGCCCGAGTCCAACACCCGTGGGAATCCCATGGGCACCA TGGCCCCTCGCTCCAAAAATGCTTTCGCGTCGCGCAGACACTGCTCGGTAGTTTCGGGGA TCAGCGTTTGAGTAAGAGCCCGCGTCTGAACCCTCCGCGCCGCCCCGGCCCCAGTGGAAA GACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGACCGCGCGCCGAGCGCGCGCCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAG GCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAAT ACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAA ATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATC TTGTGGAAAGGACGAAACACCGCACCAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGCGCCCATATTTGCATGTCGCTATGTGTTC TGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATG AGACCACGGTACACCGCACCAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGGAAACAG CATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTGCGGCCGCTGACGGCGCGCCCTGCAGTATTTAGCATGCCCCACCCATCTGC AAGGCATTCTGGATAGTGTCAAAACAGCCGGAAATCAAGTCCGTTTATCTCAAACTTTAG CATTTTGGGAATAAATGATATTTGCTATGCTGGTTAAATTAGATTTTAGTTAAATTTCCT GCTGAAGCTCTAGTACGATAAGTAACTTGACCTAAGTGTAAAGTTGAGATTTCCTTCAGG TTTATATAGCTTGTGCGCCGCCTGGGTACACCGCACCAGAGTAACAGTCTGACGTTTAAG AGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAA AGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTGTTTAAACcaacaaaatcagcag ctaatgaaggcaagtcagcaggtcactcatcattttccacttcggcaatgcagtgggatt attccaacagaggtttttcacagcattccttcagttaactggagatcgaatcttgatttt cacagatatacttggcaaggtccgccctgtcatcagcacattcaagcagatctccatggc agcattccgtgtggacttaggtaagatctgtcactaacttggaaacttctgcaaactcag cttagggaaatctctggctcaggcgagctactgcctcagcttagaaagctctttctccaa attattggagactggcacacttaagtccctgttaggcagacgaagccttcccttcatccc gaagttcatcgagctttggcaacaggcaggcagctttatcagcagcttggcaacattctg taaaagcagctttatacctttaagcaaagaaaaggagttccggggcataaaagtaaggat

	gtcttctggcaatttataataagtatttttcaaaaatgtctcttcattgtcatgaaaag cagtgcatcacacatcaacctctggtctcaccaatcgggggaggtttgggttgtttactt agtgttgcaagaattattttattctctcaggttcttgttttgcacagcagtcagctcatt caccataggtttcacgaagagttgcTGCGGCCGCCTCGATAATACGACTCACTATAGGGT CGACGATATCAGATCTaggaacccctagtgatggagttggccactccctctctgcgcgct cgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcg gcctcagtgagcgagcgagcgcgcagctgcctgcaggCATGCAAGCTGTAGCCAACCACT AGAACTATAGCTAGAGTCCTGGGCGAACAAACGATGCTCGCCTTCCAGAAAACCGAGGAT GCGAACCACTTCATCCGGGGTCAGCACCACCGGCAAGCGCCGCGACGGCCGAGGTCTTCC GATCTCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGC CGCCAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCAGCCAGGACAG AAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACGGAT GAAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCG TATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGT GGCGGTTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAA GCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTAC GCAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAA GTATGGGCATCATTCGCACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTG CTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCCCAACATCAGCCGG ACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCG ACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGC GTAGTGAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGCAGGGCATTG CCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTGGTGCTTATGTGATCT ACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTCTATACAAAGTTGGGCATAC GGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCCACCTAACAATTCGTTCAAG CCGAGATCGGCTTCCCGGCCGCGGAGTTGTTCGGTAAATTGTCACAACGCCGCGAATATA GTCTTTACCATGCCCTTGGCCACGCCCCTCTTTAATACGACGGGCAATTTGCACTTCAGA AAATGAAGAGTTTGCTTTAGCCATAACAAAAGTCCAGTATGCTTTTTCACAGCATAACTG GACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGAT CTATTTTGTTCAGTTTAAGACTTTATTGTCCGCCCACACCCGCTTACGC
916	AAGATGACGGTTTGTCACATGGAGTTGGCAGGATGTTTGATTAAAAACATAACAGGAAGA AAAATGCCCCGCTGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAGTT TTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGTAAG CTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCACTTATCTGGTTCG GATCTCCTAGGCGATATCAGTGATCAcggatctcgaccaattgacattattgaagcaACT AGTATCGATtttatcagggttattgtctcagaCCTGCAGGCAGCTGCGCGCTCGCTCGCT CACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGT GAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTAGATCTG ATATCGGCGCGCCCCTGGGCGCGCCCGAGTCCAACACCCGTGGGAATCCCATGGGCACCA TGGCCCCTCGCTCCAAAAATGCTTTCGCGTCGCGCAGACACTGCTCGGTAGTTTCGGGGA TCAGCGTTTGAGTAAGAGCCCGCGTCTGAACCCTCCGCGCCGCCCCGGCCCCAGTGGAAA GACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGACCGCGCGCCGAGCGCGCGCCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAG GCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAAT ACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAA ATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATC TTGTGGAAAGGACGAAACACCGCACCAGAGTAACAGTCTGAGGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGCGCCCATATTTGCATGTCGCTATGTGTTC TGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATG AGACCACGGTACACCGCACCAGAGTAACAGTCTGAGGTTTAAGAGCTATGCTGGAAACAG CATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTGCGGCCGCTGACGGCGCGCCCTGCAGTATTTAGCATGCCCCACCCATCTGC

	AAGGCATTCTGGATAGTGTCAAAACAGCCGGAAATCAAGTCCGTTTATCTCAAACTTTAG CATTTTGGGAATAAATGATATTTGCTATGCTGGTTAAATTAGATTTTAGTTAAATTTCCT GCTGAAGCTCTAGTACGATAAGTAACTTGACCTAAGTGTAAAGTTGAGATTTCCTTCAGG TTTATATAGCTTGTGCGCCGCCTGGGTACACCGCACCAGAGTAACAGTCTGAGGTTTAAG AGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAA AGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTGTTTAAACcaacaaaatcagcag ctaatgaaggcaagtcagcaggtcactcatcattttccacttcggcaatgcagtgggatt attccaacagaggtttttcacagcattccttcagttaactggagatcgaatcttgatttt cacagatatacttggcaaggtccgccctgtcatcagcacattcaagcagatctccatggc agcattccgtgtggacttaggtaagatctgtcactaacttggaaacttctgcaaactcag cttagggaaatctctggctcaggcgagctactgcctcagcttagaaagctctttctccaa attattggagactggcacacttaagtccctgttaggcagacgaagccttcccttcatccc gaagttcatcgagctttggcaacaggcaggcagctttatcagcagcttggcaacattctg taaaagcagctttatacctttaagcaaagaaaaggagttccggggcataaaagtaaggat gtcttctggcaatttataataagtatttttcaaaaatgtctcttcattgtcatgaaaag cagtgcatcacacatcaacctctggtctcaccaatcgggggaggtttgggttgtttactt agtgttgcaagaattattttattctctcaggttcttgttttgcacagcagtcagctcatt caccataggtttcacgaagagttgcTGCGGCCGCCTCGATAATACGACTCACTATAGGGT CGACGATATCAGATCTaggaacccctagtgatggagttggccactccctctctgcgcgct cgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcg gcctcagtgagcgagcgagcgcgcagctgcctgcaggCATGCAAGCTGTAGCCAACCACT AGAACTATAGCTAGAGTCCTGGGCGAACAAACGATGCTCGCCTTCCAGAAAACCGAGGAT GCGAACCACTTCATCCGGGGTCAGCACCACCGGCAAGCGCCGCGACGGCCGAGGTCTTCC GATCTCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGC CGCCAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCAGCCAGGACAG AAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACGGAT GAAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCG TATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGT GGCGGTTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAA GCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTAC GCAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAA GTATGGGCATCATTCGCACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTG CTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCCCAACATCAGCCGG ACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCG ACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGC GTAGTGAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGCAGGGCATTG CCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTGGTGCTTATGTGATCT ACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTCTATACAAAGTTGGGCATAC GGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCCACCTAACAATTCGTTCAAG CCGAGATCGGCTTCCCGGCCGCGGAGTTGTTCGGTAAATTGTCACAACGCCGCGAATATA GTCTTTACCATGCCCTTGGCCACGCCCCTCTTTAATACGACGGGCAATTTGCACTTCAGA AAATGAAGAGTTTGCTTTAGCCATAACAAAAGTCCAGTATGCTTTTTCACAGCATAACTG GACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGAT CTATTTTGTTCAGTTTAAGACTTTATTGTCCGCCCACACCCGCTTACGC
913	AAGATGACGGTTTGTCACATGGAGTTGGCAGGATGTTTGATTAAAAACATAACAGGAAGA AAAATGCCCCGCTGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAGTT TTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGTAAG CTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCACTTATCTGGTTCG GATCTCCTAGGCGATATCAGTGATCAcggatctcgaccaattgacattattgaagcaACT AGTATCGATtttatcagggttattgtctcagaCCACTCCCTCTATGCGCGCTCGCTCACT CACTCGGCCCTGGAGACCAAAGGTCTCCAGACTGCCGGCCTCTGGCCGGCAGGGCCGAGT GAGTGAGCGAGCGCGCATAGAGGGAGTGGGTACCTCCATCATCTAGGTTTGCCAGATCTG ATATCGGCGCGCCCCTGGGCGCGCCCGAGTCCAACACCCGTGGGAATCCCATGGGCACCA

	TGGCCCCTCGCTCCAAAAATGCTTTCGCGTCGCGCAGACACTGCTCGGTAGTTTCGGGGA TCAGCGTTTGAGTAAGAGCCCGCGTCTGAACCCTCCGCGCCGCCCCGGCCCCAGTGGAAA GACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGACCGCGCGCCGAGCGCGCGCCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAG GCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAAT ACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAA ATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATC TTGTGGAAAGGACGAAACACCGCACTAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTTTTTGCGGCCGCGCTTCGGCGCGCCCATATTTGCATGTCGCTATGTGTT CTGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTAT GAGACCACGGTACACCGCACTAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGGAAACA GCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGG TGCTTTTTTTGCGGCCGCTGACGGCGCGCCCTGCAGTATTTAGCATGCCCCACCCATCTG CAAGGCATTCTGGATAGTGTCAAAACAGCCGGAAATCAAGTCCGTTTATCTCAAACTTTA GCATTTTGGGAATAAATGATATTTGCTATGCTGGTTAAATTAGATTTTAGTTAAATTTCC TGCTGAAGCTCTAGTACGATAAGTAACTTGACCTAAGTGTAAAGTTGAGATTTCCTTCAG GTTTATATAGCTTGTGCGCCGCCTGGGTACACCGCACTAGAGTAACAGTCTGACGTTTAA GAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAA AAGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTGCGGCCGCCTCGAGTGATCAAA AAAACCAACACACGCTTCCAATGAAAATAAACGATCCTTTATTGCTAGCCTTTACTTGTA CAGCTCGTCCATGCCGAGAGTGATCCCGGCGGCGGTCACGAACTCCAGCAGGACCATGTG ATCGCGCTTCTCGTTGGGGTCTTTGCTCAGGGCGGACTGGGTGCTCAGGTAGTGGTTGTC GGGCAGCAACACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCGAGCTG CACGCTGCCGTCCTCGATGTTGTGGCGGATCTTGAAGTTCACCTTGATGCCGTTCTTCTG CTTGTCGGCCATGATATAGACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGCCCCAG GATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATGCGGTTCACCAGGGTGTC GCCCTCGAACTTCACCTCGGCGCGGGTCTTGTAGTTGCCGTCGTCCTTGAAGAAGATGGT GCGCTCCTGGACGTAGCCTTCGGGCATGGCGGACTTGAAGAAGTCGTGCTGCTTCATGTG GTCGGGGTAGCGGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCACGAGGGTGGGCCA GGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGC ATCGCCCTCGCCCTCGCCGGACACGCTGAACTTGTGGCCGTTTACGTCGCCGTCCAGCTC GACCAGGATGGGCACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGAC CGGTGGATCCCGGGCCCGCGGGGTGGCTTTACCAACAGTACCCGGAATGCCAAGCTTACT TAGATCGCAGTCTCGACGCTGGCTGGCTCCTGAGTGTCTGTCTGTGCTGTGGAGGTGGTG GTAGAATGAGGGCAGCCCCTGTGCCCCTGGGTTATATAGAGGAGCCTACAGGGTGTGACT AGCCAGGAGGGGCTGTCCCCAGGGAGGGGCCCCTGAGAGCAGATGAGCTTTCAGCTCGTT GCCCGGGCACCGTGCCCACCCCGGACCCAGGCGTGCAGCTTGCCCAGCCCCATGGCCTTG TATGGGCTGCCCCAAGGGCTGACTTGCTCACTGGTTCCTAAACTAAGTGCTGAGTCTAGC TGGCGGGGGACAGCTGGCCCTTCGCCGGGAACATGGCATGCAGGGTTATTTTTAGAGGCA GCAGGTGTTGGGGGGGGGGGGCAGCCACATGTCTGGGTTAATTATAACCAGGCATCTCG GGTGTCCCCAGGCCTTGCCTCCTTACATGGGCACGTCGACGATATCAGATCTAGGAACCC CTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGG GCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC AGAGAGGGAGTGGCATGCAAGCTGTAGCCAACCACTAGAACTATAGCTAGAGTCCTGGGC GAACAAACGATGCTCGCCTTCCAGAAAACCGAGGATGCGAACCACTTCATCCGGGGTCAG CACCACCGGCAAGCGCCGCGACGGCCGAGGTCTTCCGATCTCCTGAAGCCAGGGCAGATC CGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGCCGCCAATGCCTGACGATGCGTGGA GACCGAAACCTTGCGCTCGTTCGCCAGCCAGGACAGAAATGCCTCGACTTCGCTGCTGCC CAAGGTTGCCGGGTGACGCACACCGTGGAAACGGATGAAGGCACGAACCCAGTTGACATA AGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCGTATGCGCTCACGCAACTGGTCCAG AACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGGTTTTCATGGCTTGTTATGA CTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGG

	TCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGCAACGATGTTACGCAGC AGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATTCGCACATGTA GGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTGAGT TCGGAGACGTAGCCACCTACTCCCAACATCAGCCGGACTCCGATTACCTCGGGAACTTGC TCCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTC TCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGTGAGATCTATATCTATGATC TCGCAGTCTCCGGCGAGCACCGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCCTCA AGCATGAGGCCAACGCGCTTGGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACG ATCCCGCAGTGGCTCTCTATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATA TCGACCCAAGTACCGCCACCTAACAATTCGTTCAAGCCGAGATCGGCTTCCCGGCCGCGG AGTTGTTCGGTAAATTGTCACAACGCCGCGAATATAGTCTTTACCATGCCCTTGGCCACG CCCCTCTTTAATACGACGGGCAATTTGCACTTCAGAAAATGAAGAGTTTGCTTTAGCCAT AACAAAAGTCCAGTATGCTTTTTCACAGCATAACTGGACTGATTTCAGTTTACAACTATT CTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGATCTATTTTGTTCAGTTTAAGACTTT ATTGTCCGCCCACACCCGCTTACGC
921	AAGATGACGGTTTGTCACATGGAGTTGGCAGGATGTTTGATTAAAAACATAACAGGAAGA AAAATGCCCCGCTGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAGTT TTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGTAAG CTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCACTTATCTGGTTCG GATCTCCTAGGCGATATCAGTGATCAcggatctcgaccaattgacattattgaagcaACT AGTATCGATtttatcagggttattgtctcagaCCACTCCCTCTATGCGCGCTCGCTCACT CACTCGGCCCTGGAGACCAAAGGTCTCCAGACTGCCGGCCTCTGGCCGGCAGGGCCGAGT GAGTGAGCGAGCGCGCATAGAGGGAGTGGGTACCTCCATCATCTAGGTTTGCCAGATCTG ATATCGGCGCGCCCCTGGGCGCGCCCGAGTCCAACACCCGTGGGAATCCCATGGGCACCA TGGCCCCTCGCTCCAAAAATGCTTTCGCGTCGCGCAGACACTGCTCGGTAGTTTCGGGGA TCAGCGTTTGAGTAAGAGCCCGCGTCTGAACCCTCCGCGCCGCCCCGGCCCCAGTGGAAA GACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGACCGCGCGCCGAGCGCGCGCCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAG GCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAAT ACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAA ATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATC TTGTGGAAAGGACGAAACACCGCACCAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGCGCCCATATTTGCATGTCGCTATGTGTTC TGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATG AGACCACGGTACACCGCACCAGAGTAACAGTCTGACGTTTAAGAGCTATGCTGGAAACAG CATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTGCGGCCGCTGACGGCGCGCCCTGCAGTATTTAGCATGCCCCACCCATCTGC AAGGCATTCTGGATAGTGTCAAAACAGCCGGAAATCAAGTCCGTTTATCTCAAACTTTAG CATTTTGGGAATAAATGATATTTGCTATGCTGGTTAAATTAGATTTTAGTTAAATTTCCT GCTGAAGCTCTAGTACGATAAGTAACTTGACCTAAGTGTAAAGTTGAGATTTCCTTCAGG TTTATATAGCTTGTGCGCCGCCTGGGTACACCGCACCAGAGTAACAGTCTGACGTTTAAG AGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAA AGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTGTTTAAACcaacaaaatcagcag ctaatgaaggcaagtcagcaggtcactcatcattttccacttcggcaatgcagtgggatt attccaacagaggtttttcacagcattccttcagttaactggagatcgaatcttgatttt cacagatatacttggcaaggtccgccctgtcatcagcacattcaagcagatctccatggc agcattccgtgtggacttaggtaagatctgtcactaacttggaaacttctgcaaactcag cttagggaaatctctggctcaggcgagctactgcctcagcttagaaagctctttctccaa attattggagactggcacacttaagtccctgttaggcagacgaagccttcccttcatccc gaagttcatcgagctttggcaacaggcaggcagctttatcagcagcttggcaacattctg taaaagcagctttatacctttaagcaaagaaaaggagttccggggcataaaagtaaggat gtcttctggcaatttataataagtatttttcaaaaatgtctcttcattgtcatgaaaag

	cagtgcatcacacatcaacctctggtctcaccaatcgggggaggtttgggttgtttactt agtgttgcaagaattattttattctctcaggttcttgttttgcacagcagtcagctcatt caccataggtttcacgaagagttgcTGCGGCCGCCTCGATAATACGACTCACTATAGGGT CGACGATATCAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCT CGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCG GCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCATGCAAGCTGTAGCCAACCACT AGAACTATAGCTAGAGTCCTGGGCGAACAAACGATGCTCGCCTTCCAGAAAACCGAGGAT GCGAACCACTTCATCCGGGGTCAGCACCACCGGCAAGCGCCGCGACGGCCGAGGTCTTCC GATCTCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGC CGCCAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCAGCCAGGACAG AAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACGGAT GAAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCG TATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGT GGCGGTTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAA GCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTAC GCAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAA GTATGGGCATCATTCGCACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTG CTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCCCAACATCAGCCGG ACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCG ACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGC GTAGTGAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGCAGGGCATTG CCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTGGTGCTTATGTGATCT ACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTCTATACAAAGTTGGGCATAC GGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCCACCTAACAATTCGTTCAAG CCGAGATCGGCTTCCCGGCCGCGGAGTTGTTCGGTAAATTGTCACAACGCCGCGAATATA GTCTTTACCATGCCCTTGGCCACGCCCCTCTTTAATACGACGGGCAATTTGCACTTCAGA AAATGAAGAGTTTGCTTTAGCCATAACAAAAGTCCAGTATGCTTTTTCACAGCATAACTG GACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGAT CTATTTTGTTCAGTTTAAGACTTTATTGTCCGCCCACACCCGCTTACGC
917	AAGATGACGGTTTGTCACATGGAGTTGGCAGGATGTTTGATTAAAAACATAACAGGAAGA AAAATGCCCCGCTGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAGTT TTTAAGGATTATTTAGGGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGTAAG CTAATACGAAAATTAAAAATGACAAAATAGTTTGGAACTAGATTTCACTTATCTGGTTCG GATCTCCTAGGCGATATCAGTGATCAcggatctcgaccaattgacattattgaagcaACT AGTATCGATttatcagggttattgtctcagaCCACTCCCTCTATGCGCGCTCGCTCACT CACTCGGCCCTGGAGACCAAAGGTCTCCAGACTGCCGGCCTCTGGCCGGCAGGGCCGAGT GAGTGAGCGAGCGCGCATAGAGGGAGTGGGTACCTCCATCATCTAGGTTTGCCAGATCTG ATATCGGCGCGCCCCTGGGCGCGCCCGAGTCCAACACCCGTGGGAATCCCATGGGCACCA TGGCCCCTCGCTCCAAAAATGCTTTCGCGTCGCGCAGACACTGCTCGGTAGTTTCGGGGA TCAGCGTTTGAGTAAGAGCCCGCGTCTGAACCCTCCGCGCCGCCCCGGCCCCAGTGGAAA GACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGACCGCGCGCCGAGCGCGCGCCAAG GTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAG GCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAAT ACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAA ATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATC TTGTGGAAAGGACGAAACACCGCACCAGAGTAACAGTCTGAGGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTTTTTGCGGCCGCGCTCGGCGCGCCCATATTTGCATGTCGCTATGTGTTC TGGGAAATCACCATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATG AGACCACGGTACACCGCACCAGAGTAACAGTCTGAGGTTTAAGAGCTATGCTGGAAACAG CATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGT GCTTTTTTTGCGGCCGCTGACGGCGCGCCCTGCAGTATTTAGCATGCCCCACCCATCTGC AAGGCATTCTGGATAGTGTCAAAACAGCCGGAAATCAAGTCCGTTTATCTCAAACTTTAG

	CATTTTGGGAATAAATGATATTTGCTATGCTGGTTAAATTAGATTTTAGTTAAATTTCCT GCTGAAGCTCTAGTACGATAAGTAACTTGACCTAAGTGTAAAGTTGAGATTTCCTTCAGG TTTATATAGCTTGTGCGCCGCCTGGGTACACCGCACCAGAGTAACAGTCTGAGGTTTAAG AGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAA AGTGGCACCGAGTCGGTGCTTTTTTTGCGGCCGCGCTGTTTAAACcaacaaaatcagcag ctaatgaaggcaagtcagcaggtcactcatcattttccacttcggcaatgcagtgggatt attccaacagaggtttttcacagcattccttcagttaactggagatcgaatcttgatttt cacagatatacttggcaaggtccgccctgtcatcagcacattcaagcagatctccatggc agcattccgtgtggacttaggtaagatctgtcactaacttggaaacttctgcaaactcag cttagggaaatctctggctcaggcgagctactgcctcagcttagaaagctctttctccaa attattggagactggcacacttaagtccctgttaggcagacgaagccttcccttcatccc gaagttcatcgagctttggcaacaggcaggcagctttatcagcagcttggcaacattctg taaaagcagctttatacctttaagcaaagaaaaggagttccggggcataaaagtaaggat gtcttctggcaatttataataagtatttttcaaaaatgtctcttcattgtcatgaaaag cagtgcatcacacatcaacctctggtctcaccaatcgggggaggtttgggttgtttactt agtgttgcaagaattattttattctctcaggttcttgttttgcacagcagtcagctcatt caccataggtttcacgaagagttgcTGCGGCCGCCTCGATAATACGACTCACTATAGGGT CGACGATATCAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCT CGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCG GCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCATGCAAGCTGTAGCCAACCACT AGAACTATAGCTAGAGTCCTGGGCGAACAAACGATGCTCGCCTTCCAGAAAACCGAGGAT GCGAACCACTTCATCCGGGGTCAGCACCACCGGCAAGCGCCGCGACGGCCGAGGTCTTCC GATCTCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGTAGAAGAACAGCAAGGC CGCCAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCAGCCAGGACAG AAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACGGAT GAAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCG TATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGT GGCGGTTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAA GCAGCAAGCGCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTAC GCAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAA GTATGGGCATCATTCGCACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTG CTCTTGATCTTTTCGGTCGTGAGTTCGGAGACGTAGCCACCTACTCCCAACATCAGCCGG ACTCCGATTACCTCGGGAACTTGCTCCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCG ACCAAGAAGCGGTTGTTGGCGCTCTCGCGGCTTACGTTCTGCCCAAGTTTGAGCAGCCGC GTAGTGAGATCTATATCTATGATCTCGCAGTCTCCGGCGAGCACCGGAGGCAGGGCATTG CCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAACGCGCTTGGTGCTTATGTGATCT ACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTCTATACAAAGTTGGGCATAC GGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCCACCTAACAATTCGTTCAAG CCGAGATCGGCTTCCCGGCCGCGGAGTTGTTCGGTAAATTGTCACAACGCCGCGAATATA GTCTTTACCATGCCCTTGGCCACGCCCCTCTTTAATACGACGGGCAATTTGCACTTCAGA AAATGAAGAGTTTGCTTTAGCCATAACAAAAGTCCAGTATGCTTTTTCACAGCATAACTG GACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGAT CTATTTTGTTCAGTTTAAGACTTTATTGTCCGCCCACACCCGCTTACGC

В некоторых вариантах осуществления конструкций для редактирования генов согласно настоящему изобретения, включая варианты осуществления, охватывающие SEQ ID NO: 910-921, конструкция включает или состоит из первого промотора, первой последовательности, кодирующей gРНК, второго промотора и второй последовательности, кодирующей gРНК, третьего промотора и третьей последовательности, кодирующей gРНК. Примерами

последовательностей, кодирующих gРНК согласно настоящему изобретению, являются SEQ ID NO. 383-705, 709-711, 715-717, 790-862, 864. В некоторых вариантах осуществления последовательностью, кодирующей gРНК, является CACTAGAGTAACAGTCTGAC (SEQ ID NO. 708). В некоторых вариантах осуществления последовательностью, кодирующей gРНК, является CACCAGAGTAACAGTCTGAG (SEQ ID NO. 714). В некоторых вариантах осуществления последовательностью, кодирующей gРНК, является CACCAGAGTAACAGTCTGAC (SEQ ID NO. 863). В некоторых вариантах осуществления конструкция включает, от 5' к 3', первый промотор, последовательность, кодирующую gРНК SEQ ID NO: 708, второй промотор, вторую последовательность, кодирующую gРНК SEQ ID NO: 708, третий промотор и третью последовательность, кодирующую gРНК SEQ ID NO: 708. В некоторых вариантах осуществления конструкция включает, от 5' к 3', первый промотор, последовательность, кодирующую gРНК SEQ ID NO: 714, второй промотор, вторую последовательность, кодирующую gРНК SEQ ID NO: 714, третий промотор и третью последовательность, кодирующую gРНК SEQ ID NO: 714. В некоторых вариантах осуществления конструкция включает, от 5' к 3', первый промотор, последовательность, кодирующую gРНК SEQ ID NO: 863, второй промотор, вторую последовательность, кодирующую gРНК SEQ ID NO: 863, третий промотор и третью последовательность, кодирующую gРНК SEQ ID NO: 863. Примеры промоторов согласно настоящему описанию включают промотор U6, имеющий последовательность CGAGTCCAACACCCGTGGGAATCCCATGGGCACCATGGCCCCTCGCTCCAAAAATGCTTTCGCG TCGCGCAGACACTGCTCGGTAGTTTCGGGGATCAGCGTTTGAGTAAGAGCCCGCGTCTGAACCC TCCGCGCCGCCCCGGCCCCAGTGGAAAGACGCGCAGGCAAAACGCACCACGTGACGGAGCGTGA CCGCGCGCCGAGCGCGCGCCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATA TTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAG ATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAT TATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTT ATATATCTTGTGGAAAGGACGAAA (SEQ ID NO: 922), промотор H1, имеющий последовательность GCTCGGCGCGCCCATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTGAAATGT CTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACGGTA (SEQ ID NO: 923), и

промотор 7SK, имеющий последовательность TGACGGCGCGCCCTGCAGTATTTAGCATGCCCCACCCATCTGCAAGGCATTCTGGATAGTGTCA AAACAGCCGGAAATCAAGTCCGTTTATCTCAAACTTTAGCATTTTGGGAATAAATGATATTTGC TATGCTGGTTAAATTAGATTTTAGTTAAATTTCCTGCTGAAGCTCTAGTACGATAAGTAACTTG ACCTAAGTGTAAAGTTGAGATTTCCTTCAGGTTTATATAGCTTGTGCGCCGCCTGGGTA (SEQ ID NO: 924). В некоторых вариантах осуществления каждый первый, второй и третий промоторы индивидуально выбраны из промотора U6 (SEQ ID NO: 922), промотора H1 (SEQ ID NO: 923) и промотора 7SK (SEQ ID NO: 924). В некоторых вариантах осуществления каждый первый, второй и третий промоторы индивидуально выбраны из промотора U6 (SEQ ID NO: 922) и промотора H1 (SEQ ID NO: 923). В некоторых вариантах осуществления конструкция включает, от 5' к 3', промотор U6, первую последовательность, кодирующую gРНК, промотора H1, вторую последовательность, кодирующую gРНК, промотор 7SK и третью последовательность, кодирующую gРНК. В некоторых вариантах осуществления конструкция включает, от 5' к 3', промотор U6, последовательность, кодирующую gРНК SEQ ID NO: 708, промотор H1, вторую последовательность, кодирующую gРНК SEQ ID NO: 708, промотор 7SK и третью последовательность, кодирующую gРНК SEQ ID NO: 708. В некоторых вариантах осуществления конструкция включает, от 5' к 3', промотор U6, последовательность, кодирующую gРНК SEQ ID NO: 714, промотор H1, вторую последовательность, кодирующую gРНК SEQ ID NO: 714, промотор 7SK и третью последовательность, кодирующую gРНК SEQ ID NO: 714. В некоторых вариантах осуществления конструкция включает, от 5' к 3', промотор U6, последовательность, кодирующую gРНК SEQ ID NO: 863, промотор H1, вторую последовательность, кодирующую gРНК SEQ ID NO: 863, промотор 7SK и третью последовательность, кодирующую gРНК SEQ ID NO: 863. В некоторых вариантах осуществления, конструкция, включающая первый промотор, первую последовательность, кодирующую gРНК, второй промотор и вторую последовательность, кодирующую gРНК, третий промотор и третью последовательность, кодирующую gРНК, дополнительно включает по меньшей мере две последовательности инвертированных концевых повторов (ITR). В некоторых вариантах осуществления конструкция, включающая первый промотор, первую

последовательность, кодирующую gРНК, второй промотор и вторую последовательность, кодирующую gРНК, третий промотор и третью последовательность, кодирующую gРНК, дополнительно включает по меньшей мере две последовательности ITR, выделенные или полученные из AAV серотипа 2 (AAV2). В некоторых вариантах осуществления конструкция, включающая первый промотор, первую последовательность, кодирующую gРНК, второй промотор и вторую последовательность, кодирующую gРНК, третий промотор и третью последовательность, кодирующую gРНК, дополнительно включает по меньшей мере две последовательности ITR, где первая последовательность ITR выделена или получена из AAV серотипа 4 (AAV4), и вторая последовательность ITR выделена или получена из AAV серотипа 2 (AAV2). Примерами последовательностей ITR являются:
CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGC GACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATC ACTAGGGGTTCCT (SEQ ID NO: 881), CCACTCCCTCTATGCGCGCTCGCTCACTCACTCGGCCCTGGAGACCAAAGGTCTCCAGACTGCC GGCCTCTGGCCGGCAGGGCCGAGTGAGTGAGCGAGCGCGCATAGAGGGAGTGGGTACCTCCATC ATCTAGGTTTGCC (SEQ ID NO: 883), AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGC CCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC AGAGAGGGAGTGG (SEQ ID NO: 946) и AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGG GCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC AGCTGCCTGCAGG (ID 882 SEQ). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, первой последовательности, кодирующей gРНК, второго промотора и второй последовательности, кодирующей gРНК, третьего промотора, третьей последовательности, кодирующей gРНК и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, промотора U6, первой последовательности, кодирующей gРНК, промотора H1 и второй последовательности, кодирующей gРНК, промотора 7SK, третьей последовательности, кодирующей gРНК, и второго ITR. В некоторых вариантах осуществления конструкция включает или

состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, второго промотора и последовательности, кодирующей gРНК SEQ ID NO: 708, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, второго промотора и последовательности, кодирующей gРНК SEQ ID NO: 714, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, второго промотора и последовательности, кодирующей gРНК SEQ ID NO: 863, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO.: 863, и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, U6, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора H1 и последовательности, кодирующей gРНК SEQ ID NO: 708, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 708, и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, U6, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора H1 и последовательности, кодирующей gРНК SEQ ID NO: 714, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 714, и второго ITR. В некоторых вариантах осуществления конструкция, включающая, от 5' к 3', первый ITR, первый промотор, первую последовательность, кодирующую gРНК, второй промотор и вторую последовательность, кодирующую gРНК, третий промотор, третью последовательность, кодирующую gРНК, и второй ITR, дополнительно включает поли(A) последовательность. В некоторых вариантах осуществления поли(A) последовательность включает или состоит из мини-поли(A) последовательности. Примерные мини-поли(A) последовательности согласно настоящему описанию включают или состоят из нуклеотидной последовательности TAGCAATAAAGGATCGTTTATTTTCATTGGAAGCGTGTGTTGGTTTTTTGATCAGGCGCG

(SEQ ID NO: 903). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, первой последовательности, кодирующей gРНК, второго промотора, и вторую последовательности, кодирующей s gРНК, третьего промотора, третьей последовательности, кодирующей gРНК, мини-поли(А) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, промотора U6, первой последовательности, кодирующей gРНК, промотора H1, второй последовательности, кодирующей gРНК, промотора 7SK, третьей последовательности, кодирующей gРНК, мини-поли(А) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, мини-поли(А) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, мини-поли(А) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, мини-поли(А) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 708, мини-поли(А) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от

5' к 3', первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 714, мини-поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция, включающая, от 5' к 3', первый ITR, первый промотор, первую последовательность, кодирующую gРНК, второй промотор и вторую последовательность, кодирующую gРНК, третий промотор, третью последовательность, кодирующую gРНК, мини-поли(A) последовательность и второй ITR, дополнительно включает инвертированные повторы транспозируемого элемента. Примерные инвертированные повторы транспозируемого элемента согласно настоящему описанию включают или состоят из нуклеотидной последовательности TGTGGGCGGACAAAATAGTTGGGAACTGGGAGGGGTGGAAATGGAGTTTTTAAGGATTATTTAG GGAAGAGTGACAAAATAGATGGGAACTGGGTGTAGCGTCGTAAGCTAATACGAAAATTAAAAAT GACAAAATAGTTTGGAACTAGATTTCACTTATCTGGTT (SEQ ID NO: 907) и/или нуклеотидной последовательности GAATATAGTCTTTACCATGCCCTTGGCCACGCCCCTCTTTAATACGACGGGCAATTTGCACTTC AGAAAATGAAGAGTTTGCTTTAGCCATAACAAAAGTCCAGTATGCTTTTTCACAGCATAACTGG ACTGATTTCAGTTTACAACTATTCTGTCTAGTTTAAGACTTTATTGTCATAGTTTAGATCTATT TTGTTCAGTTTAAGACTTTATTGTCCGCCCACA (SEQ ID NO: 908). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, первой последовательности, кодирующей gРНК, второго промотора, второй последовательности, кодирующей gРНК, третьего промотора, третьей последовательности, кодирующей gРНК, мини-поли(A) последовательности, второго ITR и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, первой последовательности, кодирующей gРНК, промотора H1, второй последовательности, кодирующей gРНК, промотора 7SK, третьей последовательности, кодирующей gРНК, мини-поли(A) последовательности, второго ITR и второго

инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, мини-поли(А) последовательности, второго ITR и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, мини-поли(А) последовательности, второго ITR и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, мини-поли(А) последовательности, второго ITR и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 708, мини-поли(А) последовательностм, второго ITR и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, последовательности,

кодирующей gРНК SEQ ID NO: 714, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 714, мини-поли(A) последовательности, второго ITR и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция, включающая первый инвертированный повтор транспозируемого элемента, первый ITR, первый промотор, первую последовательность, кодирующую gРНК, второй промотор, вторую последовательность, кодирующую gРНК, третий промотор, третью последовательность, кодирующую gРНК, мини-поли(A) последовательность, второй ITR, второй инвертированный повтор транспозируемого элемента, дополнительно включает регуляторную последовательность. Примерные регуляторные последовательности согласно настоящему описанию включают или состоят из нуклеотидной последовательности CATGCAAGCTGTAGCCAACCACTAGAACTATAGCTAGAGTCCTGGGCGAACAAACGATGCTCGC CTTCCAGAAAACCGAGGATGCGAACCACTTCATCCGGGGTCAGCACCACCGGCAAGCGCCGCGA CGGCCGAGGTCTTCCGATCTCCTGAAGCCAGGGCAGATCCGTGCACAGCACCTTGCCGTAGAAG AACAGCAAGGCCGCCAATGCCTGACGATGCGTGGAGACCGAAACCTTGCGCTCGTTCGCCAGCC AGGACAGAAATGCCTCGACTTCGCTGCTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACG GATGAAGGCACGAACCCAGTTGACATAAGCCTGTTCGGTTCGTAAACTGTAATGCAAGTAGCGT ATGCGCTCACGCAACTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTGGCGG TTTTCATGGCTTGTTATGACTGTTTTTTTGTACAGTCTATGCCTCGGGCATCCAAGCAGCAAGC GCGTTACGCCGTGGGTCGATGTTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGCAACGA TGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATTCGC ACATGTAGGCTCGGCCCTGACCAAGTCAAATCCATGCGGGCTGCTCTTGATCTTTTCGGTCGTG AGTTCGGAGACGTAGCCACCTACTCCCAACATCAGCCGGACTCCGATTACCTCGGGAACTTGCT CCGTAGTAAGACATTCATCGCGCTTGCTGCCTTCGACCAAGAAGCGGTTGTTGGCGCTCTCGCG GCTTACGTTCTGCCCAAGTTTGAGCAGCCGCGTAGTGAGATCTATATCTATGATCTCGCAGTCT CCGGCGAGCACCGGAGGCAGGGCATTGCCACCGCGCTCATCAATCTCCTCAAGCATGAGGCCAA CGCGCTTGGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACGATCCCGCAGTGGCTCTC TATACAAAGTTGGGCATACGGGAAGAAGTGATGCACTTTGATATCGACCCAAGTACCGCCACCT AACAATTCGTTCAAGCCGAGATCGGCTTCCCGGCCGCGGAGTTGTTCGGTAAATTGTCACAACG CCG (SEQ ID NO: 909). В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR,

первого промотора, первой последовательности, кодирующей gРНК, второго промотора, второй последовательности, кодирующей gРНК, третьего промотора, третьей последовательности, кодирующей gРНК, мини-поли(А) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, первой последовательности, кодирующей gРНК, промотора H1, второй последовательности, кодирующей gРНК, промотора 7SK, третьей последовательности, кодирующей gРНК, мини-поли(А) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, мини-поли(А) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, мини-поли(А) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, третьего промотора, последовательности,

кодирующей gРНК SEQ ID NO: 863, мини-поли(А) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 708, мини-поли(А) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 714, мини-поли(А) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента В некоторых вариантах осуществления конструкция, включающая первый ITR, первый промотор, первую последовательность, кодирующую gРНК, второй промотор и вторую последовательность, кодирующую gРНК, третий промотор, третью последовательность, кодирующую gРНК, мини-поли(А) последовательность и второй ITR, дополнительно включают спейсерную последовательность. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3' первого ITR, промотора U6, первой последовательности, кодирующей gРНК, промотора H1, второй последовательности, кодирующей gРНК, промотора 7SK, третьей последовательности, кодирующей gРНК, спейсерной последовательности, мини-поли(А) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, третьего промотора, последовательности, кодирующей gРНК

SEQ ID NO: 708, спейсерной последовательности, мини-поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, спейсерной последовательности, мини-поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, спейсерной последовательности, мини-поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3' первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 708, спейсерной последовательности, мини-поли(A) последовательности и второго ITR. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 714, спейсерной последовательности, мини-поли(A) последовательности и второго ITR. В некоторых вариантах осуществления, конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, первой последовательности, кодирующей gРНК, второго промотора, второй последовательности, кодирующей gРНК, третьего промотора, третьей последовательности, кодирующей gРНК, спейсерной последовательности, мини-поли(A) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых

вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, первой последовательности, кодирующей gРНК, промотора H1, второй последовательности, кодирующей gРНК, промотора 7SK, третьей последовательности, кодирующей gРНК, спейсерной последовательности, мини-поли(A) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 708, спейсерной последовательности, мини-поли(A) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 714, спейсерной последовательности, мини-поли(A) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, первого промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, второго промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, третьего промотора, последовательности, кодирующей gРНК SEQ ID NO: 863, спейсерной последовательности, мини-поли(A) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или

состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 708, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 708, спейсерной последовательности, мини-поли(A) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция включает или состоит из, от 5' к 3', первого инвертированного повтора транспозируемого элемента, первого ITR, промотора U6, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора H1, последовательности, кодирующей gРНК SEQ ID NO: 714, промотора 7SK, последовательности, кодирующей gРНК SEQ ID NO: 714, спейсерной последовательности, мини-поли(A) последовательности, второго ITR, регуляторной последовательности и второго инвертированного повтора транспозируемого элемента. В некоторых вариантах осуществления конструкция может дополнительно включать одну или более спейсерных последовательностей. Примерные спейсерные последовательности согласно настоящему описанию имеют длину от 1-1500 нуклеотидов, включая все промежуточные диапазоны. В некоторых вариантах осуществления спейсерные последовательности могут быть расположены в положении 5' или 3' относительно ITR, промотора, последовательности, кодирующей gРНК, поли(A) последовательности, инвертированного повтора транспозируемого элемента, спейсерной последовательности и/или регуляторного элемента.

V. Фармацевтические композиции и способы доставки

Для клинического применения фармацевтические композиции изготавливают в форме, подходящей для предполагаемого применения. Как правило, это включает изготовление композиций, которые по существу не содержат пирогенов, а также других примесей, которые могут быть вредными для людей или животных.

Подходящие соли и буферы используются для придания стабильности лекарственным средствам, белкам или векторам для доставки, а также обеспечивает захват клетками-мишенями. Водные

композиции согласно настоящему изобретению содержат эффективное количество лекарственного средства, вектора или белков, растворенных или диспергированных в фармацевтически приемлемом носителе или водной среде. Фраза "фармацевтически или фармакологически приемлемый" относится к химическим соединениям и композициям, которые не вызывают нежелательных, аллергических или других неблагоприятных реакций при введении животному или человеку. При использовании в настоящем документе термин "фармацевтически приемлемый носитель" включает растворители, буферы, растворы, дисперсионные среды, покрытия, противобактериальные и противогрибковые средства, изотонические вещества и вещества, замедляющие абсорбцию, и т.п., подходящие для применения в производстве лекарственных препаратов, таких как лекарственные препараты, подходящие для введения людям. Использование таких сред и средств для фармацевтически активных веществ хорошо известно в данной области. Любая стандартная среда или средство, которые не являются несовместимыми с активными веществами согласно настоящему изобретению, могут применяться в терапевтических композициях. Дополнительные активные вещества также могут быть включены в композиции при условии, что они не инактивируют векторы или клетки композиций.

В некоторых вариантах осуществления активные композиции согласно настоящему изобретению могут включать классические лекарственные препараты. Введение таких композиций согласно настоящему изобретению может быть осуществлено любым стандартным путем, при условии, что целевая ткань доступна при использовании этого пути, но обычно включает системное введение. Оно включает пероральное, назальное или буккальное. В альтернативе введение может быть внутрикожным, подкожным, внутримышечным, внутрибрюшинным или внутривенным введением или осуществляться путем прямой инъекции в мышечную ткань. Такие композиции обычно вводят в виде фармацевтически приемлемых композиций, как описано выше.

Активные соединения также могут вводить парентерально или внутрибрюшинно. В качестве иллюстрации растворы активных соединений в виде свободного основания или фармакологически

приемлемых солей могут быть приготовлены в воде, надлежащим образом смешанной с поверхностно-активным веществом, таким как гидроксипропилцеллюлоза. Дисперсии также могут быть приготовлены в глицерине, жидких полиэтиленгликолях и их смесях, а также в маслах. В обычных условиях хранения и применения эти препараты обычно содержат консервант для предотвращения роста микроорганизмов.

Лекарственные формы, подходящие для инъекционного применения, включают, например, стерильные водные растворы или дисперсии и стерильные порошки для экстемпорального приготовления стерильных растворов или дисперсий для инъекций. Как правило, такие препараты являются стерильными и жидкими в такой степени, чтобы их можно было легко водить путем инъекции. Препараты должны быть стабильными в условиях производства и хранения и должны быть защищены от контаминирующего воздействия микроорганизмов, таких как бактерии и грибы. Подходящие растворители или дисперсионные среды могут содержать, например, воду, этанол, полиол (например, глицерин, пропиленгликоль и жидкий полиэтиленгликоль и т.п.), их подходящие смеси, а также растительные масла. Подходящую текучесть можно поддерживать, например, путем применения покрытия, такого как лецитин, путем поддержания требуемого размера частиц в случае дисперсии и путем применения поверхностно-активных веществ. Предотвращение воздействия микроорганизмов можно обепечить различными противобактериальными и противогрибковыми средствами, например, парабенами, хлорбутанолом, фенолом, сорбиновой кислотой, тимеросалом и т.п. Во многих случаях предпочтительно включать изотонические вещества, например, сахара или хлорид натрия. Пролонгированная абсорбция композиций для инъекций может быть достигнута при использовании в композициях веществ, задерживающих абсорбцию, например, моностеарата алюминия и желатина.

Стерильные растворы для инъекций могут быть приготовлены путем включения активных соединений в подходящем количестве в растворитель вместе с любыми другими компонентами (например, как указано выше), при необходимости, с последующей стерилизацией

фильтрованием. Обычно дисперсии приготавливают путем включения различных стерильных активных веществ в стерильный носитель, который содержит основную дисперсионную среду и другие требуемые компоненты, например, как перечисленные выше. В случае стерильных порошков для приготовления стерильных растворов для инъекций предпочтительные способы приготовления включают в себя методы вакуумной сушки и лиофильной сушки, которые дают порошок активного(ых) вещества(в) плюс любой дополнительный требуемый компонент из их предварительно простерилизованного фильтрованием раствора.

В некоторых вариантах осуществления композиции согласно настоящему изобретению изготовлены в нейтральной или солевой форме. Фармацевтически приемлемые соли включают, например, соли присоединия кислот (образованные со свободными аминогруппами белка), полученные из неорганических кислот (например, соляной или фосфорной кислот) или из органических кислот (например, уксусной, щавелевой, винной, миндальной), и т.п. Соли, образованные со свободными карбоксильными группами белка, также могут быть получены из неорганических оснований (например, гидроксидов натрия, калия, аммония, кальция или железа) или из органических оснований (например, изопропиламина, триметиламина, гистидина, прокаина), и т.п.

После приготовления растворы предпочтительно вводят способом, совместимым с лекарственной формой, и в таком количестве, которое является терапевтически эффективным. Композиции можно легко вводить в различных лекарственных формах, таких как растворы для инъекций, капсулы для высвобождения лекарственного средства и т.п. Например, для парентерального введения в водном растворе раствор обычно содержит буфер, и жидкий разбавитель сначала делают изотоническим, например, при добавлении достаточного количества раствора хлорида натрия или глюкозы. Такие водные растворы можно использовать, например, для внутривенного, внутримышечного, подкожного и внутрибрюшинного введения. Предпочтительно используют стерильные водные среды, как известно специалистам в данной области, особенно в рамках настоящего описания. В качестве иллюстрации, одна доза может

быть растворена в 1 мл изотонического раствора NaCl и либо добавлена к 1000 мл жидкости для гиподермоклизиса, либо введена в предлагаемый участок инфузии (см., например, "Remington's Pharmaceutical Sciences" 15-е издание, стр. 1035-1038 и 1570-1580). Некоторое изменение дозы обязательно производят в зависимости от состояния субъекта, который подвергается лечению. Лицо, ответственное за введение, в любом случае определяет подходящую дозу для отдельного субъекта. Кроме того, для введения человеку препараты должны соответствовать стандартам стерильности, пирогенности, общей безопасности и чистоты, согласно требованиям стандарта Центра исследований биопрепаратов FDA.

В некоторых вариантах осуществления Cas9 или Cpf1 и gРНК, описанные в настоящем документе, могут быть доставлены пациенту с помощью адоптивного переноса клеток (ACT). В адоптивном переносе клеток одну или более экспрессионных конструкций вводят *ex vivo* в клетки, которые были получены от (аутологичного) пациента или от одного или более лиц, отличных от пациента (аллогенных). Затем клетки вводят или повторно вводят пациенту. Таким образом, в некоторых вариантах осуществления одна или более нуклеиновых кислот, кодирующих Cas9 или Cpf1 и направляющую РНК, которая направленно взаимодействует с сайтом сплайсинга дистрофина, вводят в клетку *ex vivo*, после чего клетку вводят или повторно вводят пациенту.

В следующих таблицах представлены примеры последовательностей праймеров и геномных последовательностей-мишеней для применения с композициями и способами, раскрытыми в настоящем документе.

ТАБЛИЦА 6 – Геномные последовательности-мишени

Целевой экзон gРНК	Номер гида	Цепь	Геномная последовательность-мишень*	PAM	SEQ ID NO.
Человек-Экзон 51	4	1	tctttttcttctttttcctttt	tttt	60
Человек-Экзон 51	5	1	ctttttcttctttttcctttttG	tttt	61
Человек-Экзон 51	6	1	tttttcttctttttcctttttGC	tttc	62
Человек-Экзон	7	1	tcttctttttcctttttGCAAAA	tttt	63

51					
Человек-Экзон 51	8	1	cttctttttcctttttGCAAAAA	tttt	64
Человек-Экзон 51	9	1	ttctttttcctttttGCAAAAAC	tttc	65
Человек-Экзон 51	10	1	ttcctttttGCAAAAACCCAAAAT	tttt	66
Человек-Экзон 51	11	1	tcctttttGCAAAAACCCAAAATA	tttt	67
Человек-Экзон 51	12	1	ccttttttGCAAAAACCCAAAATAT	tttt	68
Человек-Экзон 51	13	1	ctttttGCAAAAACCCAAAATATT	tttc	69
Человек-Экзон 51	14	1	tGCAAAAACCCAAAATATTTTAGC	tttt	70
Человек-Экзон 51	15	1	GCAAAAACCCAAAATATTTTAGCT	tttt	71
Человек-Экзон 51	16	1	CAAAAACCCAAAATATTTTAGCTC	tttG	72
Человек-Экзон 51	17	1	AGCTCCTACTCAGACTGTTACTCT	TTTT	73
Человек-Экзон 51	18	1	GCTCCTACTCAGACTGTTACTCTG	TTTA	74
Человек-Экзон 51	19	-1	CTTAGTAACCACAGGTTGTGTCAC	TTTC	75
Человек-Экзон 51	20	-1	GAGATGGCAGTTTCCTTAGTAACC	TTTG	76
Человек-Экзон 51	21	-1	TAGTTTGGAGATGGCAGTTTCCTT	TTTC	77
Человек-Экзон 51	22	-1	TTCTCATACCTTCTGCTTGATGAT	TTTT	78
Человек-Экзон 51	23	-1	TCATTTTTTCTCATACCTTCTGCT	TTTA	79
Человек-Экзон 51	24	-1	ATCATTTTTTCTCATACCTTCTGC	TTTT	80
Человек-Экзон 51	25	-1	AAGAAAAACTTCTGCCAACTTTTA	TTTA	81
Человек-Экзон 51	26	-1	AAAGAAAAACTTCTGCCAACTTTT	TTTT	82
Человек-Экзон 51	27	1	TCTTTAAAATGAAGATTTTCCACC	TTTT	83
Человек-Экзон 51	28	1	CTTTAAAATGAAGATTTTCCACCA	TTTT	84
Человек-Экзон 51	29	1	TTTAAAATGAAGATTTTCCACCAA	TTTC	85
Человек-Экзон 51	30	1	AAATGAAGATTTTCCACCAATCAC	TTTA	86
Человек-Экзон 51	31	1	CCACCAATCACTTTACTCTCCTAG	TTTT	87
Человек-Экзон 51	32	1	CACCAATCACTTTACTCTCCTAGA	TTTC	88
Человек-Экзон	33	1	CTCTCCTAGACCATTTCCCACCAG	TTTA	89

51					
Человек-Экзон 45	1	-1	agaaaagattaaacagtgtgctac	tttg	90
Человек-Экзон 45	2	-1	tttgagaaaagattaaacagtgtg	TTTa	91
Человек-Экзон 45	3	-1	atttgagaaaagattaaacagtgt	TTTT	92
Человек-Экзон 45	4	-1	Tatttgagaaaagattaaacagtg	TTTT	93
Человек-Экзон 45	5	1	atcttttctcaaatAAAAAGACAT	ttta	94
Человек-Экзон 45	6	1	ctcaaatAAAAAGACATGGGGCTT	tttt	95
Человек-Экзон 45	7	1	tcaaatAAAAAGACATGGGGCTTC	tttc	96
Человек-Экзон 45	8	1	TGTTTTGCCTTTTTGGTATCTTAC	TTTT	97
Человек-Экзон 45	9	1	GTTTTGCCTTTTTGGTATCTTACA	TTTT	98
Человек-Экзон 45	10	1	TTTTGCCTTTTTGGTATCTTACAG	TTTG	99
Человек-Экзон 45	11	1	GCCTTTTTGGTATCTTACAGGAAC	TTTT	100
Человек-Экзон 45	12	1	CCTTTTTGGTATCTTACAGGAACT	TTTG	101
Человек-Экзон 45	13	1	TGGTATCTTACAGGAACTCCAGGA	TTTT	102
Человек-Экзон 45	14	1	GGTATCTTACAGGAACTCCAGGAT	TTTT	103
Человек-Экзон 45	15	-1	AGGATTGCTGAATTATTTCTTCCC	TTTG	104
Человек-Экзон 45	16	-1	GAGGATTGCTGAATTATTTCTTCC	TTTT	105
Человек-Экзон 45	17	-1	TGAGGATTGCTGAATTATTTCTTC	TTTT	106
Человек-Экзон 45	18	-1	CTGTAGAATACTGGCATCTGTTTT	TTTC	107
Человек-Экзон 45	19	-1	CCTGTAGAATACTGGCATCTGTTT	TTTT	108
Человек-Экзон 45	20	-1	TCCTGTAGAATACTGGCATCTGTT	TTTT	109
Человек-Экзон 45	21	-1	CAGACCTCCTGCCACCGCAGATTC	TTTG	110
Человек-Экзон 45	22	-1	TGTCTGACAGCTGTTTGCAGACCT	TTTC	111
Человек-Экзон 45	23	-1	CTGTCTGACAGCTGTTTGCAGACC	TTTT	112
Человек-Экзон 45	24	-1	TCTGTCTGACAGCTGTTTGCAGAC	TTTT	113
Человек-Экзон 45	25	-1	TTCTGTCTGACAGCTGTTTGCAGA	TTTT	114
Человек-Экзон	26	-1	ATTCCTATTAGATCTGTCGCCCTA	TTTC	115

45					
Человек-Экзон 45	27	-1	CATTCCTATTAGATCTGTCGCCCT	TTTT	116
Человек-Экзон 45	28	1	AGCAGACTTTTTAAGCTTTCTTTA	TTTT	117
Человек-Экзон 45	29	1	GCAGACTTTTTAAGCTTTCTTTAG	TTTA	118
Человек-Экзон 45	30	1	TAAGCTTTCTTTAGAAGAATATTT	TTTT	119
Человек-Экзон 45	31	1	AAGCTTTCTTTAGAAGAATATTTC	TTTT	120
Человек-Экзон 45	32	1	AGCTTTCTTTAGAAGAATATTTCA	TTTA	121
Человек-Экзон 45	33	1	TTTAGAAGAATATTTCATGAGAGA	TTTC	122
Человек-Экзон 45	34	1	GAAGAATATTTCATGAGAGATTAT	TTTA	123
Человек-Экзон 44	1	1	TCAGTATAACCAAAAAATATACGC	TTTG	124
Человек-Экзон 44	2	1	acataatccatctatttttcttga	tttt	125
Человек-Экзон 44	3	1	cataatccatctatttttcttgat	ttta	126
Человек-Экзон 44	4	1	tcttgatccatatgcttttACCTG	tttt	127
Человек-Экзон 44	5	1	cttgatccatatgcttttACCTGC	tttt	128
Человек-Экзон 44	6	1	ttgatccatatgcttttACCTGCA	tttc	129
Человек-Экзон 44	7	-1	TCAACAGATCTGTCAAATCGCCTG	TTTC	130
Человек-Экзон 44	8	1	ACCTGCAGGCGATTTGACAGATCT	tttt	131
Человек-Экзон 44	9	1	CCTGCAGGCGATTTGACAGATCTG	tttA	132
Человек-Экзон 44	10	1	ACAGATCTGTTGAGAAATGGCGGC	TTTG	133
Человек-Экзон 44	11	-1	TATCATAATGAAAACGCCGCCATT	TTTA	134
Человек-Экзон 44	12	1	CATTATGATATAAAGATATTTAAT	TTTT	135
Человек-Экзон 44	13	-1	TATTTAGCATGTTCCCAATTCTCA	TTTG	136
Человек-Экзон 44	14	-1	GAAAAAACAAATCAAAGACTTACC	TTTC	137
Человек-Экзон 44	15	1	ATTTGTTTTTTCGAAATTGTATTT	TTTG	138
Человек-Экзон 44	16	1	TTTTTTCGAAATTGTATTTATCTT	TTTG	139
Человек-Экзон 44	17	1	TTCGAAATTGTATTTATCTTCAGC	TTTT	140
Человек-Экзон	18	1	TCGAAATTGTATTTATCTTCAGCA	TTTT	141

44					
Человек-Экзон 44	19	1	CGAAATTGTATTTATCTTCAGCAC	TTTT	142
Человек-Экзон 44	20	1	GAAATTGTATTTATCTTCAGCACA	TTTC	143
Человек-Экзон 44	21	-1	AGAAGTTAAAGAGTCCAGATGTGC	TTTA	144
Человек-Экзон 44	22	1	TCTTCAGCACATCTGGACTCTTTA	TTTA	145
Человек-Экзон 44	23	-1	CATCACCCTTCAGAACCTGATCTT	TTTC	146
Человек-Экзон 44	24	1	ACTTCTTAAAGATCAGGTTCTGAA	TTTA	147
Человек-Экзон 44	25	1	GACTGTTGTTGTCATCATTATATT	TTTT	148
Человек-Экзон 44	26	1	ACTGTTGTTGTCATCATTATATTA	TTTG	149
Человек-Экзон 53	1	-1	AACTAGAATAAAAGGAAAAATAAA	TTTC	150
Человек-Экзон 53	2	1	CTACTATATATTTATTTTTCCTTT	TTTA	151
Человек-Экзон 53	3	1	TTTTTCCTTTTATTCTAGTTGAAA	TTTA	152
Человек-Экзон 53	4	1	TCCTTTTATTCTAGTTGAAAGAAT	TTTT	153
Человек-Экзон 53	5	1	CCTTTTATTCTAGTTGAAAGAATT	TTTT	154
Человек-Экзон 53	6	1	CTTTTATTCTAGTTGAAAGAATTC	TTTC	155
Человек-Экзон 53	7	1	ATTCTAGTTGAAAGAATTCAGAAT	TTTT	156
Человек-Экзон 53	8	1	TTCTAGTTGAAAGAATTCAGAATC	TTTA	157
Человек-Экзон 53	9	-1	ATTCAACTGTTGCCTCCGGTTCTG	TTTC	158
Человек-Экзон 53	10	-1	ACATTTCATTCAACTGTTGCCTCC	TTTA	159
Человек-Экзон 53	11	-1	CTTTTGGATTGCATCTACTGTATA	TTTT	160
Человек-Экзон 53	12	-1	TGTGATTTTCTTTTGGATTGCATC	TTTC	161
Человек-Экзон 53	13	-1	ATACTAACCTTGGTTTCTGTGATT	TTTG	162
Человек-Экзон 53	14	-1	AAAAGGTATCTTTGATACTAACCT	TTTA	163
Человек-Экзон 53	15	-1	AAAAAGGTATCTTTGATACTAACC	TTTT	164
Человек-Экзон 53	16	-1	TTTTAAAAAGGTATCTTTGATACT	TTTA	165
Человек-Экзон 53	17	-1	ATTTTAAAAAGGTATCTTTGATAC	TTTT	166
Человек-Экзон	1	-1	TTAATGCAAACTGGGACACAAACA	TTTG	167

46					
Человек-Экзон 46	2	1	TAAATTGCCATGTTTGTGTCCCAG	TTTT	168
Человек-Экзон 46	3	1	AAATTGCCATGTTTGTGTCCCAGT	TTTT	169
Человек-Экзон 46	4	1	AATTGCCATGTTTGTGTCCCAGTT	TTTA	170
Человек-Экзон 46	5	1	TGTCCCAGTTTGCATTAACAAATA	TTTG	171
Человек-Экзон 46	6	-1	CAACATAGTTCTCAAACTATTTGT	tttC	172
Человек-Экзон 46	7	-1	CCAACATAGTTCTCAAACTATTTG	tttt	173
Человек-Экзон 46	8	-1	tCCAACATAGTTCTCAAACTATTT	tttt	174
Человек-Экзон 46	9	-1	tttCCAACATAGTTCTCAAACTAT	tttt	175
Человек-Экзон 46	10	-1	ttttCCAACATAGTTCTCAAACTA	tttt	176
Человек-Экзон 46	11	-1	tttttCCAACATAGTTCTCAAACT	tttt	177
Человек-Экзон 46	12	1	CATTAACAAATAGTTTGAGAACTA	TTTG	178
Человек-Экзон 46	13	1	AGAACTATGTTGGaaaaaaaaaTA	TTTG	179
Человек-Экзон 46	14	-1	GTTCTTCTAGCCTGGAGAAAGAAG	TTTT	180
Человек-Экзон 46	15	1	ATTCTTCTTTCTCCAGGCTAGAAG	TTTT	181
Человек-Экзон 46	16	1	TTCTTCTTTCTCCAGGCTAGAAGA	TTTA	182
Человек-Экзон 46	17	1	TCCAGGCTAGAAGAACAAAAGAAT	TTTC	183
Человек-Экзон 46	18	-1	AAATTCTGACAAGATATTCTTTTG	TTTG	184
Человек-Экзон 46	19	-1	CTTTTAGTTGCTGCTCTTTTCCAG	TTTT	185
Человек-Экзон 46	20	-1	AGAAAATAAAATTACCTTGACTTG	TTTG	186
Человек-Экзон 46	21	-1	TGCAAGCAGGCCCTGGGGGATTTG	TTTA	187
Человек-Экзон 46	22	1	ATTTTCTCAAATCCCCCAGGGCCT	TTTT	188
Человек-Экзон 46	23	1	TTTTCTCAAATCCCCCAGGGCCTG	TTTA	189
Человек-Экзон 46	24	1	CTCAAATCCCCCAGGGCCTGCTTG	TTTT	190
Человек-Экзон 46	25	1	TCAAATCCCCCAGGGCCTGCTTGC	TTTC	191
Человек-Экзон 46	26	1	TTAATTCAATCATTGGTTTTCTGC	TTTT	192
Человек-Экзон	27	1	TAATTCAATCATTGGTTTTCTGCC	TTTT	193

46					
Человек-Экзон 46	28	1	AATTCAATCATTGGTTTTCTGCCC	TTTT	194
Человек-Экзон 46	29	1	ATTCAATCATTGGTTTTCTGCCCA	TTTA	195
Человек-Экзон 46	30	-1	GCAAGGAACTATGAATAACCTAAT	TTTA	196
Человек-Экзон 46	31	1	CTGCCCATTAGGTTATTCATAGTT	TTTT	197
Человек-Экзон 46	32	1	TGCCCATTAGGTTATTCATAGTTC	TTTC	198
Человек-Экзон 52	1	-1	TAGAAAACAATTTAACAGGAAATA	TTTA	199
Человек-Экзон 52	2	1	CTGTTAAATTGTTTTCTATAAACC	TTTC	200
Человек-Экзон 52	3	-1	GAAATAAAAAAGATGTTACTGTAT	TTTA	201
Человек-Экзон 52	4	-1	AGAAATAAAAAAGATGTTACTGTA	TTTT	202
Человек-Экзон 52	5	1	CTATAAACCCTTATACAGTAACAT	TTTT	203
Человек-Экзон 52	6	1	TATAAACCCTTATACAGTAACATC	TTTC	204
Человек-Экзон 52	7	1	TTATTTCTAAAAGTGTTTTGGCTG	TTTT	205
Человек-Экзон 52	8	1	TATTTCTAAAAGTGTTTTGGCTGG	TTTT	206
Человек-Экзон 52	9	1	ATTTCTAAAAGTGTTTTGGCTGGT	TTTT	207
Человек-Экзон 52	10	1	TTTCTAAAAGTGTTTTGGCTGGTC	TTTA	208
Человек-Экзон 52	11	1	TAAAAGTGTTTTGGCTGGTCTCAC	TTTC	209
Человек-Экзон 52	12	-1	CATAATACAAAGTAAAGTACAATT	TTTA	210
Человек-Экзон 52	13	-1	ACATAATACAAAGTAAAGTACAAT	TTTT	211
Человек-Экзон 52	14	1	GGCTGGTCTCACAATTGTACTTTA	TTTT	212
Человек-Экзон 52	15	1	GCTGGTCTCACAATTGTACTTTAC	TTTG	213
Человек-Экзон 52	16	1	CTTTGTATTATGTAAAAGGAATAC	TTTA	214
Человек-Экзон 52	17	1	TATTATGTAAAAGGAATACACAAC	TTTG	215
Человек-Экзон 52	18	1	TTCTTACAGGCAACAATGCAGGAT	TTTG	216
Человек-Экзон 52	19	1	GAACAGAGGCGTCCCCAGTTGGAA	TTTG	217
Человек-Экзон 52	20	-1	GGCAGCGGTAATGAGTTCTTCCAA	TTTG	218
Человек-Экзон	21	-1	TCAAATTTTGGGCAGCGGTAATGA	TTTT	219

52					
Человек-Экзон 52	22	1	AAAAACAAGACCAGCAATCAAGAG	TTTG	220
Человек-Экзон 52	23	-1	TGTGTCCCATGCTTGTTAAAAAAC	TTTG	221
Человек-Экзон 52	24	1	TTAACAAGCATGGGACACACAAAG	TTTT	222
Человек-Экзон 52	25	1	TAACAAGCATGGGACACACAAAGC	TTTT	223
Человек-Экзон 52	26	1	AACAAGCATGGGACACACAAAGCA	TTTT	224
Человек-Экзон 52	27	1	ACAAGCATGGGACACACAAAGCAA	TTTA	225
Человек-Экзон 52	28	-1	TTGAAACTTGTCATGCATCTTGCT	TTTA	226
Человек-Экзон 52	29	-1	ATTGAAACTTGTCATGCATCTTGC	TTTT	227
Человек-Экзон 52	30	-1	TATTGAAACTTGTCATGCATCTTG	TTTT	228
Человек-Экзон 52	31	1	AATAAAAACTTAAGTTCATATATC	TTTC	229
Человек-Экзон 50	1	-1	GTGAATATATTATTGGATTTCTAT	TTTG	230
Человек-Экзон 50	2	-1	AAGATAATTCATGAACATCTTAAT	TTTG	231
Человек-Экзон 50	3	-1	ACAGAAAAGCATACACATTACTTA	TTTA	232
Человек-Экзон 50	4	1	CTGTTAAAGAGGAAGTTAGAAGAT	TTTT	233
Человек-Экзон 50	5	1	TGTTAAAGAGGAAGTTAGAAGATC	TTTC	234
Человек-Экзон 50	6	-1	CCGCCTTCCACTCAGAGCTCAGAT	TTTA	235
Человек-Экзон 50	7	-1	CCCTCAGCTCTTGAAGTAAACGGT	TTTG	236
Человек-Экзон 50	8	1	CTTCAAGAGCTGAGGGCAAAGCAG	TTTA	237
Человек-Экзон 50	9	-1	AACAAATAGCTAGAGCCAAAGAGA	TTTG	238
Человек-Экзон 50	10	-1	GAACAAATAGCTAGAGCCAAAGAG	TTTT	239
Человек-Экзон 50	11	1	GCTCTAGCTATTTGTTCAAAAGTG	TTTG	240
Человек-Экзон 50	12	1	TTCAAAAGTGCAACTATGAAGTGA	TTTG	241
Человек-Экзон 50	13	-1	TCTCTCACCCAGTCATCACTTCAT	TTTC	242
Человек-Экзон 50	14	-1	CTCTCTCACCCAGTCATCACTTCA	TTTT	243
Человек-Экзон 43	1	1	tatatatatatatatTTTTCTCTT	TTTG	244
Человек-Экзон	2	1	TCTCTTTCTATAGACAGCTAATTC	tTTT	245

43					
Человек-Экзон 43	3	1	CTCTTTCTATAGACAGCTAATTCA	TTTT	246
Человек-Экзон 43	4	-1	AAACAGTAAAAAAATGAATTAGCT	TTTA	247
Человек-Экзон 43	5	1	TCTTTCTATAGACAGCTAATTCAT	TTTC	248
Человек-Экзон 43	6	-1	AAAACAGTAAAAAAATGAATTAGC	TTTT	249
Человек-Экзон 43	7	1	TATAGACAGCTAATTCATTTTTTT	TTTC	250
Человек-Экзон 43	8	-1	TATTCTGTAATATAAAAATTTTAA	TTTA	251
Человек-Экзон 43	9	-1	ATATTCTGTAATATAAAAATTTTA	TTTT	252
Человек-Экзон 43	10	1	TTTACTGTTTTAAAATTTTTATAT	TTTT	253
Человек-Экзон 43	11	1	TTACTGTTTTAAAATTTTTATATT	TTTT	254
Человек-Экзон 43	12	1	TACTGTTTTAAAATTTTTATATTA	TTTT	255
Человек-Экзон 43	13	1	ACTGTTTTAAAATTTTTATATTAC	TTTT	256
Человек-Экзон 43	14	1	CTGTTTTAAAATTTTTATATTACA	TTTA	257
Человек-Экзон 43	15	1	AAAATTTTTATATTACAGAATATA	TTTT	258
Человек-Экзон 43	16	1	AAATTTTTATATTACAGAATATAA	TTTA	259
Человек-Экзон 43	17	-1	TTGTAGACTATCTTTTATATTCTG	TTTG	260
Человек-Экзон 43	18	1	TATATTACAGAATATAAAAGATAG	TTTT	261
Человек-Экзон 43	19	1	ATATTACAGAATATAAAAGATAGT	TTTT	262
Человек-Экзон 43	20	1	TATTACAGAATATAAAAGATAGTC	TTTA	263
Человек-Экзон 43	21	-1	CAATGCTGCTGTCTTCTTGCTATG	TTTG	264
Человек-Экзон 43	22	1	CAATGGGAAAAAGTTAACAAAATG	TTTC	265
Человек-Экзон 43	23	-1	TGCAAGTATCAAGAAAAATATATG	TTTC	266
Человек-Экзон 43	24	1	TCTTGATACTTGCAGAAATGATTT	TTTT	267
Человек-Экзон 43	25	1	CTTGATACTTGCAGAAATGATTTG	TTTT	268
Человек-Экзон 43	26	1	TTGATACTTGCAGAAATGATTTGT	TTTC	269
Человек-Экзон 43	27	1	TTTTCAGGGAACTGTAGAATTTAT	TTTG	270
Человек-Экзон	28	-1	CATGGAGGGTACTGAAATAAATTC	TTTC	271

43					
Человек-Экзон 43	29	-1	CCATGGAGGGTACTGAAATAAATT	TTTT	272
Человек-Экзон 43	30	1	CAGGGAACTGTAGAATTTATTTCA	TTTT	273
Человек-Экзон 43	31	-1	TCCATGGAGGGTACTGAAATAAAT	TTTT	274
Человек-Экзон 43	32	1	AGGGAACTGTAGAATTTATTTCAG	TTTC	275
Человек-Экзон 43	33	-1	TTCCATGGAGGGTACTGAAATAAA	TTTT	276
Человек-Экзон 43	34	-1	CCTGTCTTTTTTCCATGGAGGGTA	TTTC	277
Человек-Экзон 43	35	-1	CCCTGTCTTTTTTCCATGGAGGGT	TTTT	278
Человек-Экзон 43	36	-1	TCCCTGTCTTTTTTCCATGGAGGG	TTTT	279
Человек-Экзон 43	37	1	TTTCAGTACCCTCCATGGAAAAAA	TTTA	280
Человек-Экзон 43	38	1	AGTACCCTCCATGGAAAAAAGACA	TTTC	281
Человек-Экзон 6	1	1	AGTTTGCATGGTTCTTGCTCAAGG	TTTA	282
Человек-Экзон 6	2	-1	ATAAGAAAATGCATTCCTTGAGCA	TTTC	283
Человек-Экзон 6	3	-1	CATAAGAAAATGCATTCCTTGAGC	TTTT	284
Человек-Экзон 6	4	1	CATGGTTCTTGCTCAAGGAATGCA	TTTG	285
Человек-Экзон 6	5	-1	ACCTACATGTGGAAATAAATTTTC	TTTG	286
Человек-Экзон 6	6	-1	GACCTACATGTGGAAATAAATTTT	TTTT	287
Человек-Экзон 6	7	-1	TGACCTACATGTGGAAATAAATTT	TTTT	288
Человек-Экзон 6	8	1	CTTATGAAAATTTATTTCCACATG	TTTT	289
Человек-Экзон 6	9	1	TTATGAAAATTTATTTCCACATGT	TTTC	290
Человек-Экзон 6	10	-1	ATTACATTTTTGACCTACATGTGG	TTTC	291
Человек-Экзон 6	11	-1	CATTACATTTTTGACCTACATGTG	TTTT	292
Человек-Экзон 6	12	-1	TCATTACATTTTTGACCTACATGT	TTTT	293
Человек-Экзон 6	13	1	TTTCCACATGTAGGTCAAAAATGT	TTTA	294
Человек-Экзон 6	14	1	CACATGTAGGTCAAAAATGTAATG	TTTC	295
Человек-Экзон 6	15	-1	TTGCAATCCAGCCATGATATTTTT	TTTG	296
Человек-Экзон	16	-1	ACTGTTGGTTTGTTGCAATCCAGC	TTTC	297

6					
Человек-Экзон 6	17	-1	CACTGTTGGTTTGTTGCAATCCAG	TTTT	298
Человек-Экзон 6	18	1	AATGCTCTCATCCATAGTCATAGG	TTTG	299
Человек-Экзон 6	19	-1	ATGTCTCAGTAATCTTCTTACCTA	TTTA	300
Человек-Экзон 6	20	-1	CAAGTTATTTAATGTCTCAGTAAT	TTTA	301
Человек-Экзон 6	21	-1	ACAAGTTATTTAATGTCTCAGTAA	TTTT	302
Человек-Экзон 6	22	1	GACTCTGATGACATATTTTTCCCC	TTTA	303
Человек-Экзон 6	23	1	TCCCCAGTATGGTTCCAGATCATG	TTTT	304
Человек-Экзон 6	24	1	CCCCAGTATGGTTCCAGATCATGT	TTTT	305
Человек-Экзон 6	25	1	CCCAGTATGGTTCCAGATCATGTC	TTTC	306
Человек-Экзон 7	1	1	TATTTGTCTTtgtgtatgtgtgta	TTTA	307
Человек-Экзон 7	2	1	TCTTtgtgtatgtgtgtatgtgta	TTTG	308
Человек-Экзон 7	3	1	tgtatgtgtgtatgtgtatgtgtt	TTtg	309
Человек-Экзон 7	4	1	AGGCCAGACCTATTTGACTGGAAT	ttTT	310
Человек-Экзон 7	5	1	GGCCAGACCTATTTGACTGGAATA	tTTA	311
Человек-Экзон 7	6	1	ACTGGAATAGTGTGGTTTGCCAGC	TTTG	312
Человек-Экзон 7	7	1	CCAGCAGTCAGCCACACAACGACT	TTTG	313
Человек-Экзон 7	8	-1	TCTATGCCTAATTGATATCTGGCG	TTTC	314
Человек-Экзон 7	9	-1	CCAACCTTCAGGATCGAGTAGTTT	TTTA	315
Человек-Экзон 7	10	1	TGGACTACCACTGCTTTTAGTATG	TTTC	316
Человек-Экзон 7	11	1	AGTATGGTAGAGTTTAATGTTTTC	TTTT	317
Человек-Экзон 7	12	1	GTATGGTAGAGTTTAATGTTTTCA	TTTA	318
Человек-Экзон 8	1	-1	AGACTCTAAAAGGATAATGAACAA	TTTG	319
Человек-Экзон 8	2	1	ACTTTGATTTGTTCATTATCCTTT	TTTA	320
Человек-Экзон 8	3	-1	TATATTTGAGACTCTAAAAGGATA	TTTC	321
Человек-Экзон 8	4	1	ATTTGTTCATTATCCTTTTAGAGT	TTTG	322
Человек-Экзон	5	-1	GTTTCTATATTTGAGACTCTAAAA	TTTG	323

8					
Человек-Экзон 8	6	-1	GGTTTCTATATTTGAGACTCTAAA	TTTT	324
Человек-Экзон 8	7	-1	TGGTTTCTATATTTGAGACTCTAA	TTTT	325
Человек-Экзон 8	8	1	TTCATTATCCTTTTAGAGTCTCAA	TTTG	326
Человек-Экзон 8	9	1	AGAGTCTCAAATATAGAAACCAAA	TTTT	327
Человек-Экзон 8	10	1	GAGTCTCAAATATAGAAACCAAAA	TTTA	328
Человек-Экзон 8	11	-1	CACTTCCTGGATGGCTTCAATGCT	TTTC	329
Человек-Экзон 8	12	1	GCCTCAACAAGTGAGCATTGAAGC	TTTT	330
Человек-Экзон 8	13	1	CCTCAACAAGTGAGCATTGAAGCC	TTTG	331
Человек-Экзон 8	14	-1	GGTGGCCTTGGCAACATTTCCACT	TTTA	332
Человек-Экзон 8	15	-1	GTCACTTTAGGTGGCCTTGGCAAC	TTTA	333
Человек-Экзон 8	16	-1	ATGATGTAACTGAAAATGTTCTTC	TTTG	334
Человек-Экзон 8	17	-1	CCTGTTGAGAATAGTGCATTTGAT	TTTA	335
Человек-Экзон 8	18	1	CAGTTACATCATCAAATGCACTAT	TTTT	336
Человек-Экзон 8	19	1	AGTTACATCATCAAATGCACTATT	TTTC	337
Человек-Экзон 8	20	-1	CACACTTTACCTGTTGAGAATAGT	TTTA	338
Человек-Экзон 8	21	1	CTGTTTTATATGCATTTTTAGGTA	TTTT	339
Человек-Экзон 8	22	1	TGTTTTATATGCATTTTTAGGTAT	TTTC	340
Человек-Экзон 8	23	1	ATATGCATTTTTAGGTATTACGTG	TTTT	341
Человек-Экзон 8	24	1	TATGCATTTTTAGGTATTACGTGC	TTTA	342
Человек-Экзон 8	25	1	TAGGTATTACGTGCACatatatat	TTTT	343
Человек-Экзон 8	26	1	AGGTATTACGTGCACatatatata	TTTT	344
Человек-Экзон 8	27	1	GGTATTACGTGCACatatatatat	TTTA	345
Человек-Экзон 55	1	-1	AGCAACAACTATAATATTGTGCAG	TTTA	346
Человек-Экзон 55	2	1	GTTCCTCCATCTTTCTCTTTTTAT	TTTA	347
Человек-Экзон 55	3	1	TCTTTTTATGGAGTTCACTAGGTG	TTTC	348
Человек-Экзон	4	1	TATGGAGTTCACTAGGTGCACCAT	TTTT	349

55					
Человек-Экзон 55	5	1	ATGGAGTTCACTAGGTGCACCATT	TTTT	350
Человек-Экзон 55	6	1	TGGAGTTCACTAGGTGCACCATTC	TTTA	351
Человек-Экзон 55	7	1	ATAATTGCATCTGAACATTTGGTC	TTTA	352
Человек-Экзон 55	8	1	GTCCTTTGCAGGGTGAGTGAGCGA	TTTG	353
Человек-Экзон 55	9	-1	TTCCAAAGCAGCCTCTCGCTCACT	TTTC	354
Человек-Экзон 55	10	1	CAGGGTGAGTGAGCGAGAGGCTGC	TTTG	355
Человек-Экзон 55	11	1	GAAGAAACTCATAGATTACTGCAA	TTTG	356
Человек-Экзон 55	12	-1	CAGGTCCAGGGGGAACTGTTGCAG	TTTC	357
Человек-Экзон 55	13	-1	CCAGGTCCAGGGGGAACTGTTGCA	TTTT	358
Человек-Экзон 55	14	-1	AGCTTCTGTAAGCCAGGCAAGAAA	TTTC	359
Человек-Экзон 55	15	1	TTGCCTGGCTTACAGAAGCTGAAA	TTTC	360
Человек-Экзон 55	16	-1	CTTACGGGTAGCATCCTGTAGGAC	TTTC	361
Человек-Экзон 55	17	-1	CTCCCTTGGAGTCTTCTAGGAGCC	TTTA	362
Человек-Экзон 55	18	-1	ACTCCCTTGGAGTCTTCTAGGAGC	TTTT	363
Человек-Экзон 55	19	-1	ATCAGCTCTTTTACTCCCTTGGAG	TTTC	364
Человек-Экзон 55	20	1	CGCTTTAGCACTCTTGTGGATCCA	TTTC	365
Человек-Экзон 55	21	1	GCACTCTTGTGGATCCAATTGAAC	TTTA	366
Человек-Экзон 55	22	-1	TCCCTGGCTTGTCAGTTACAAGTA	TTTG	367
Человек-Экзон 55	23	-1	GTCCCTGGCTTGTCAGTTACAAGT	TTTT	368
Человек-Экзон 55	24	-1	TTTTGTCCCTGGCTTGTCAGTTAC	TTTG	369
Человек-Экзон 55	25	-1	GTTTTGTCCCTGGCTTGTCAGTTA	TTTT	370
Человек-Экзон 55	26	1	TACTTGTAACTGACAAGCCAGGGA	TTTG	371
Человек-G1-экзон51		1	gCTCCTACTCAGACTGTTACTCTG	TTTA	372
Человек-G2-экзон51		1	taccatgtattgctaaacaaagta	TTTC	373
Человек-G3-экзон51		-1	attgaagagtaacaatttgagcca	TTTA	374

мышь-Экзон23-G1		1	aggctctgcaaagttctTTGAAAG	TTTG	375
мышь-Экзон23-G2		1	AAAGAGCAACAAAATGGCttcaac	TTTG	376
мышь-Экзон23-G3		1	AAAGAGCAATAAAATGGCttcaac	TTTG	377
мышь-Экзон23-G4		-1	AAAGAACTTTGCAGAGCctcaaaa	TTTC	378
мышь-Экзон23-G5		-1	ctgaatatctatgcattaataact	TTTA	379
мышь-Экзон23-G6		-1	tattatattacagggcatattata	TTTC	380
мышь-Экзон23-G7		1	Aggtaagccgaggtttggccttta	TTTC	381
мышь-Экзон23-G8		1	cccagagtccttcaaagatattga	TTTA	382

* В этой таблице заглавные буквы обозначают нуклеотиды, которые выравниваются с последовательностью экзона гена. Строчные буквы обозначают нуклеотиды, которые выравниваются с последовательностью интрона гена.

ТАБЛИЦА 7 – последовательности gРНК

Экзон-мишень gРНК	**Номер гида**	**Цепь**	**Последовательность gРНК***	**PAM**	**SEQ ID NO.**
Человек-Экзон 51	4	1	aaaaaggaaaaaagaagaaaaaga	tttt	383
Человек-Экзон 51	5	1	Caaaaaggaaaaaagaagaaaaag	tttt	384
Человек-Экзон 51	6	1	GCaaaaaggaaaaaagaagaaaaa	tttc	385
Человек-Экзон 51	7	1	UUUUGCaaaaaggaaaaaagaaga	tttt	386
Человек-Экзон 51	8	1	UUUUUGCaaaaaggaaaaaagaag	tttt	387
Человек-Экзон 51	9	1	GUUUUUGCaaaaaggaaaaaagaa	tttc	388
Человек-Экзон 51	10	1	AUUUUGGGUUUUUGCaaaaaggaa	tttt	389
Человек-Экзон 51	11	1	UAUUUUGGGUUUUUGCaaaaagga	tttt	390
Человек-Экзон 51	12	1	AUAUUUUGGGUUUUUGCaaaaagg	tttt	391
Человек-Экзон 51	13	1	AAUAUUUUGGGUUUUUGCaaaaag	tttc	392
Человек-Экзон 51	14	1	GCUAAAAUAUUUUGGGUUUUUGCa	tttt	393
Человек-Экзон 51	15	1	AGCUAAAAUAUUUUGGGUUUUUGC	tttt	394
Человек-Экзон 51	16	1	GAGCUAAAAUAUUUUGGGUUUUUG	tttG	395

Человек-Экзон 51	17	1	AGAGUAACAGUCUGAGUAGGAGCU	TTTT	396
Человек-Экзон 51	18	1	CAGAGUAACAGUCUGAGUAGGAGC	TTTA	397
Человек-Экзон 51	19	-1	GUGACACAACCUGUGGUUACUAAG	TTTC	398
Человек-Экзон 51	20	-1	GGUUACUAAGGAAACUGCCAUCU	TTTG	399
Человек-Экзон 51	21	-1	AAGGAAACUGCCAUCUCCAAACUA	TTTC	400
Человек-Экзон 51	22	-1	AUCAUCAAGCAGAAGGUAUGAGAA	TTTT	401
Человек-Экзон 51	23	-1	AGCAGAAGGUAUGAGAAAAAAUGA	TTTA	402
Человек-Экзон 51	24	-1	GCAGAAGGUAUGAGAAAAAAUGAU	TTTT	403
Человек-Экзон 51	25	-1	UAAAAGUUGGCAGAAGUUUUUCUU	TTTA	404
Человек-Экзон 51	26	-1	AAAAGUUGGCAGAAGUUUUUCUUU	TTTT	405
Человек-Экзон 51	27	1	GGUGGAAAAUCUUCAUUUUAAAGA	TTTT	406
Человек-Экзон 51	28	1	UGGUGGAAAAUCUUCAUUUUAAAG	TTTT	407
Человек-Экзон 51	29	1	UUGGUGGAAAAUCUUCAUUUUAAA	TTTC	408
Человек-Экзон 51	30	1	GUGAUUGGUGGAAAAUCUUCAUUU	TTTA	409
Человек-Экзон 51	31	1	CUAGGAGAGUAAAGUGAUUGGUGG	TTTT	410
Человек-Экзон 51	32	1	UCUAGGAGAGUAAAGUGAUUGGUG	TTTC	411
Человек-Экзон 51	33	1	CUGGUGGGAAAUGGUCUAGGAGA	TTTA	412
Человек-Экзон 45	1	-1	guagcacacuguuuaaucuuuucu	tttg	413
Человек-Экзон 45	2	-1	cacacuguuuaaucuuuucucaaa	TTTa	414
Человек-Экзон 45	3	-1	acacuguuuaaucuuuucucaaau	TTTT	415
Человек-Экзон 45	4	-1	cacuguuuaaucuuuucucaaauA	TTTT	416
Человек-Экзон 45	5	1	AUGUCUUUUUauuugagaaaagau	ttta	417
Человек-Экзон 45	6	1	AAGCCCCAUGUCUUUUUauuugag	tttt	418
Человек-Экзон 45	7	1	GAAGCCCCAUGUCUUUUUauuuga	tttc	419
Человек-Экзон 45	8	1	GUAAGAUACCAAAAAGGCAAAACA	TTTT	420
Человек-Экзон 45	9	1	UGUAAGAUACCAAAAAGGCAAAAC	TTTT	421

Человек-Экзон 45	10	1	CUGUAAGAUACCAAAAAGGCAAAA	TTTG	422
Человек-Экзон 45	11	1	GUUCCUGUAAGAUACCAAAAAGGC	TTTT	423
Человек-Экзон 45	12	1	AGUUCCUGUAAGAUACCAAAAAGG	TTTG	424
Человек-Экзон 45	13	1	UCCUGGAGUUCCUGUAAGAUACCA	TTTT	425
Человек-Экзон 45	14	1	AUCCUGGAGUUCCUGUAAGAUACC	TTTT	426
Человек-Экзон 45	15	-1	GGGAAGAAAUAAUUCAGCAAUCCU	TTTG	427
Человек-Экзон 45	16	-1	GGAAGAAAUAAUUCAGCAAUCCUC	TTTT	428
Человек-Экзон 45	17	-1	GAAGAAAUAAUUCAGCAAUCCUCA	TTTT	429
Человек-Экзон 45	18	-1	AAAACAGAUGCCAGUAUUCUACAG	TTTC	430
Человек-Экзон 45	19	-1	AAACAGAUGCCAGUAUUCUACAGG	TTTT	431
Человек-Экзон 45	20	-1	AACAGAUGCCAGUAUUCUACAGGA	TTTT	432
Человек-Экзон 45	21	-1	GAAUCUGCGGUGGCAGGAGGUCUG	TTTG	433
Человек-Экзон 45	22	-1	AGGUCUGCAAACAGCUGUCAGACA	TTTC	434
Человек-Экзон 45	23	-1	GGUCUGCAAACAGCUGUCAGACAG	TTTT	435
Человек-Экзон 45	24	-1	GUCUGCAAACAGCUGUCAGACAGA	TTTT	436
Человек-Экзон 45	25	-1	UCUGCAAACAGCUGUCAGACAGAA	TTTT	437
Человек-Экзон 45	26	-1	UAGGGCGACAGAUCUAAUAGGAAU	TTTC	438
Человек-Экзон 45	27	-1	AGGGCGACAGAUCUAAUAGGAAUG	TTTT	439
Человек-Экзон 45	28	1	UAAAGAAAGCUUAAAAAGUCUGCU	TTTT	440
Человек-Экзон 45	29	1	CUAAAGAAAGCUUAAAAAGUCUGC	TTTA	441
Человек-Экзон 45	30	1	AAAUAUUCUUCUAAAGAAAGCUUA	TTTT	442
Человек-Экзон 45	31	1	GAAAUAUUCUUCUAAAGAAAGCUU	TTTT	443
Человек-Экзон 45	32	1	UGAAAUAUUCUUCUAAAGAAAGCU	TTTA	444
Человек-Экзон 45	33	1	UCUCUCAUGAAAUAUUCUUCUAAA	TTTC	445
Человек-Экзон 45	34	1	AUAAUCUCUCAUGAAAUAUUCUUC	TTTA	446
Человек-Экзон 44	1	1	GCGUAUAUUUUUUGGUUAUACUGA	TTTG	447

Человек-Экзон 44	2	1	ucaagaaaaauagauggauuaugu	tttt	448
Человек-Экзон 44	3	1	aucaagaaaaauagauggauuaug	ttta	449
Человек-Экзон 44	4	1	CAGGUaaaagcauauggaucaaga	tttt	450
Человек-Экзон 44	5	1	GCAGGUaaaagcauauggaucaag	tttt	451
Человек-Экзон 44	6	1	UGCAGGUaaaagcauauggaucaa	tttc	452
Человек-Экзон 44	7	-1	CAGGCGAUUUGACAGAUCUGUUGA	TTTC	453
Человек-Экзон 44	8	1	AGAUCUGUCAAAUCGCCUGCAGGU	tttt	454
Человек-Экзон 44	9	1	CAGAUCUGUCAAAUCGCCUGCAGG	tttA	455
Человек-Экзон 44	10	1	GCCGCCAUUUCUCAACAGAUCUGU	TTTG	456
Человек-Экзон 44	11	-1	AAUGGCGGCGUUUUCAUUAUGAUA	TTTA	457
Человек-Экзон 44	12	1	AUUAAAUAUCUUUAUAUCAUAAUG	TTTT	458
Человек-Экзон 44	13	-1	UGAGAAUUGGGAACAUGCUAAAUA	TTTG	459
Человек-Экзон 44	14	-1	GGUAAGUCUUUGAUUUGUUUUUUC	TTTC	460
Человек-Экзон 44	15	1	AAAUACAAUUUCGAAAAAACAAAU	TTTG	461
Человек-Экзон 44	16	1	AAGAUAAAUACAAUUUCGAAAAAA	TTTG	462
Человек-Экзон 44	17	1	GCUGAAGAUAAAUACAAUUUCGAA	TTTT	463
Человек-Экзон 44	18	1	UGCUGAAGAUAAAUACAAUUUCGA	TTTT	464
Человек-Экзон 44	19	1	GUGCUGAAGAUAAAUACAAUUUCG	TTTT	465
Человек-Экзон 44	20	1	UGUGCUGAAGAUAAAUACAAUUUC	TTTC	466
Человек-Экзон 44	21	-1	GCACAUCUGGACUCUUUAACUUCU	TTTA	467
Человек-Экзон 44	22	1	UAAAGAGUCCAGAUGUGCUGAAGA	TTTA	468
Человек-Экзон 44	23	-1	AAGAUCAGGUUCUGAAGGGUGAUG	TTTC	469
Человек-Экзон 44	24	1	UUCAGAACCUGAUCUUUAAGAAGU	TTTA	470
Человек-Экзон 44	25	1	AAUAUAAUGAUGACAACAACAGUC	TTTT	471
Человек-Экзон 44	26	1	UAAUAUAAUGAUGACAACAACAGU	TTTG	472
Человек-Экзон 53	1	-1	UUUAUUUUUCCUUUUAUUCUAGUU	TTTC	473

Человек-Экзон 53	2	1	AAAGGAAAAAUAAAUAUAUAGUAG	TTTA	474
Человек-Экзон 53	3	1	UUUCAACUAGAAUAAAAGGAAAAA	TTTA	475
Человек-Экзон 53	4	1	AUUCUUUCAACUAGAAUAAAAGGA	TTTT	476
Человек-Экзон 53	5	1	AAUUCUUUCAACUAGAAUAAAAGG	TTTT	477
Человек-Экзон 53	6	1	GAAUUCUUUCAACUAGAAUAAAAG	TTTC	478
Человек-Экзон 53	7	1	AUUCUGAAUUCUUUCAACUAGAAU	TTTT	479
Человек-Экзон 53	8	1	GAUUCUGAAUUCUUUCAACUAGAA	TTTA	480
Человек-Экзон 53	9	-1	CAGAACCGGAGGCAACAGUUGAAU	TTTC	481
Человек-Экзон 53	10	-1	GGAGGCAACAGUUGAAUGAAAUGU	TTTA	482
Человек-Экзон 53	11	-1	UAUACAGUAGAUGCAAUCCAAAAG	TTTT	483
Человек-Экзон 53	12	-1	GAUGCAAUCCAAAAGAAAAUCACA	TTTC	484
Человек-Экзон 53	13	-1	AAUCACAGAAACCAAGGUUAGUAU	TTTG	485
Человек-Экзон 53	14	-1	AGGUUAGUAUCAAAGAUACCUUU	TTTA	486
Человек-Экзон 53	15	-1	GGUUAGUAUCAAAGAUACCUUUUU	TTTT	487
Человек-Экзон 53	16	-1	AGUAUCAAAGAUACCUUUUUAAAA	TTTA	488
Человек-Экзон 53	17	-1	GUAUCAAAGAUACCUUUUUAAAAU	TTTT	489
Человек-Экзон 46	1	-1	UGUUUGUGUCCCAGUUUGCAUUAA	TTTG	490
Человек-Экзон 46	2	1	CUGGGACACAAACAUGGCAAUUUA	TTTT	491
Человек-Экзон 46	3	1	ACUGGGACACAAACAUGGCAAUUU	TTTT	492
Человек-Экзон 46	4	1	AACUGGGACACAAACAUGGCAAUU	TTTA	493
Человек-Экзон 46	5	1	UAUUUGUUAAUGCAAACUGGGACA	TTTG	494
Человек-Экзон 46	6	-1	ACAAAUAGUUUGAGAACUAUGUUG	tttC	495
Человек-Экзон 46	7	-1	CAAAUAGUUUGAGAACUAUGUUGG	tttt	496
Человек-Экзон 46	8	-1	AAAUAGUUUGAGAACUAUGUUGGa	tttt	497
Человек-Экзон 46	9	-1	AUAGUUUGAGAACUAUGUUGGaaa	tttt	498
Человек-Экзон 46	10	-1	UAGUUUGAGAACUAUGUUGGaaaa	tttt	499

Человек-Экзон 46	11	-1	AGUUUGAGAACUAUGUUGGaaaaa	tttt	500
Человек-Экзон 46	12	1	UAGUUCUCAAACUAUUUGUUAAUG	TTTG	501
Человек-Экзон 46	13	1	UAuuuuuuuuuCCAACAUAGUUCU	TTTG	502
Человек-Экзон 46	14	-1	CUUCUUUCUCCAGGCUAGAAGAAC	TTTT	503
Человек-Экзон 46	15	1	CUUCUAGCCUGGAGAAAGAAGAAU	TTTT	504
Человек-Экзон 46	16	1	UCUUCUAGCCUGGAGAAAGAAGAA	TTTA	505
Человек-Экзон 46	17	1	AUUCUUUUGUUCUUCUAGCCUGGA	TTTC	506
Человек-Экзон 46	18	-1	CAAAAGAAUAUCUUGUCAGAAUUU	TTTG	507
Человек-Экзон 46	19	-1	CUGGAAAAGAGCAGCAACUAAAAG	TTTT	508
Человек-Экзон 46	20	-1	CAAGUCAAGGUAAUUUUAUUUUCU	TTTG	509
Человек-Экзон 46	21	-1	CAAAUCCCCCAGGGCCUGCUUGCA	TTTA	510
Человек-Экзон 46	22	1	AGGCCCUGGGGGAUUUGAGAAAAU	TTTT	511
Человек-Экзон 46	23	1	CAGGCCCUGGGGGAUUUGAGAAAA	TTTA	512
Человек-Экзон 46	24	1	CAAGCAGGCCCUGGGGGAUUUGAG	TTTT	513
Человек-Экзон 46	25	1	GCAAGCAGGCCCUGGGGGAUUUGA	TTTC	514
Человек-Экзон 46	26	1	GCAGAAAACCAAUGAUUGAAUUAA	TTTT	515
Человек-Экзон 46	27	1	GGCAGAAAACCAAUGAUUGAAUUA	TTTT	516
Человек-Экзон 46	28	1	GGGCAGAAAACCAAUGAUUGAAUU	TTTT	517
Человек-Экзон 46	29	1	UGGGCAGAAAACCAAUGAUUGAAU	TTTA	518
Человек-Экзон 46	30	-1	AUUAGGUUAUUCAUAGUUCCUUGC	TTTA	519
Человек-Экзон 46	31	1	AACUAUGAAUAACCUAAUGGGCAG	TTTT	520
Человек-Экзон 46	32	1	GAACUAUGAAUAACCUAAUGGGCA	TTTC	521
Человек-Экзон 52	1	-1	UAUUUCCUGUUAAAUUGUUUUCUA	TTTA	522
Человек-Экзон 52	2	1	GGUUUAUAGAAAACAAUUUAACAG	TTTC	523
Человек-Экзон 52	3	-1	AUACAGUAACAUCUUUUUUAUUUC	TTTA	524
Человек-Экзон 52	4	-1	UACAGUAACAUCUUUUUUAUUUCU	TTTT	525

Человек-Экзон 52	5	1	AUGUUACUGUAUAAGGGUUUAUAG	TTTT	526
Человек-Экзон 52	6	1	GAUGUUACUGUAUAAGGGUUUAUA	TTTC	527
Человек-Экзон 52	7	1	CAGCCAAAACACUUUUAGAAAUAA	TTTT	528
Человек-Экзон 52	8	1	CCAGCCAAAACACUUUUAGAAAUA	TTTT	529
Человек-Экзон 52	9	1	ACCAGCCAAAACACUUUUAGAAAU	TTTT	530
Человек-Экзон 52	10	1	GACCAGCCAAAACACUUUUAGAAA	TTTA	531
Человек-Экзон 52	11	1	GUGAGACCAGCCAAAACACUUUUA	TTTC	532
Человек-Экзон 52	12	-1	AAUUGUACUUUACUUUGUAUUAUG	TTTA	533
Человек-Экзон 52	13	-1	AUUGUACUUUACUUUGUAUUAUGU	TTTT	534
Человек-Экзон 52	14	1	UAAAGUACAAUUGUGAGACCAGCC	TTTT	535
Человек-Экзон 52	15	1	GUAAAGUACAAUUGUGAGACCAGC	TTTG	536
Человек-Экзон 52	16	1	GUAUUCCUUUUACAUAAUACAAAG	TTTA	537
Человек-Экзон 52	17	1	GUUGUGUAUUCCUUUUACAUAAUA	TTTG	538
Человек-Экзон 52	18	1	AUCCUGCAUUGUUGCCUGUAAGAA	TTTG	539
Человек-Экзон 52	19	1	UUCCAACUGGGGACGCCUCUGUUC	TTTG	540
Человек-Экзон 52	20	-1	UUGGAAGAACUCAUUACCGCUGCC	TTTG	541
Человек-Экзон 52	21	-1	UCAUUACCGCUGCCCAAAAUUUGA	TTTT	542
Человек-Экзон 52	22	1	CUCUUGAUUGCUGGUCUUGUUUUU	TTTG	543
Человек-Экзон 52	23	-1	GUUUUUUAACAAGCAUGGGACACA	TTTG	544
Человек-Экзон 52	24	1	CUUUGUGUGUCCCAUGCUUGUUAA	TTTT	545
Человек-Экзон 52	25	1	GCUUUGUGUGUCCCAUGCUUGUUA	TTTT	546
Человек-Экзон 52	26	1	UGCUUUGUGUGUCCCAUGCUUGUU	TTTT	547
Человек-Экзон 52	27	1	UUGCUUUGUGUGUCCCAUGCUUGU	TTTA	548
Человек-Экзон 52	28	-1	AGCAAGAUGCAUGACAAGUUUCAA	TTTA	549
Человек-Экзон 52	29	-1	GCAAGAUGCAUGACAAGUUUCAAU	TTTT	550
Человек-Экзон 52	30	-1	CAAGAUGCAUGACAAGUUUCAAUA	TTTT	551

Человек-Экзон 52	31	1	GAUAUAUGAACUUAAGUUUUUAUU	TTTC	552
Человек-Экзон 50	1	-1	AUAGAAAUCCAAUAAUAUAUUCAC	TTTG	553
Человек-Экзон 50	2	-1	AUUAAGAUGUUCAUGAAUUAUCUU	TTTG	554
Человек-Экзон 50	3	-1	UAAGUAAUGUGUAUGCUUUUCUGU	TTTA	555
Человек-Экзон 50	4	1	AUCUUCUAACUUCCUCUUUAACAG	TTTT	556
Человек-Экзон 50	5	1	GAUCUUCUAACUUCCUCUUUAACA	TTTC	557
Человек-Экзон 50	6	-1	AUCUGAGCUCUGAGUGGAAGGCGG	TTTA	558
Человек-Экзон 50	7	-1	ACCGUUUACUUCAAGAGCUGAGGG	TTTG	559
Человек-Экзон 50	8	1	CUGCUUUGCCCUCAGCUCUUGAAG	TTTA	560
Человек-Экзон 50	9	-1	UCUCUUUGGCUCUAGCUAUUUGUU	TTTG	561
Человек-Экзон 50	10	-1	CUCUUUGGCUCUAGCUAUUUGUUC	TTTT	562
Человек-Экзон 50	11	1	CACUUUUGAACAAAUAGCUAGAGC	TTTG	563
Человек-Экзон 50	12	1	UCACUUCAUAGUUGCACUUUUGAA	TTTG	564
Человек-Экзон 50	13	-1	AUGAAGUGAUGACUGGGUGAGAGA	TTTC	565
Человек-Экзон 50	14	-1	UGAAGUGAUGACUGGGUGAGAGAG	TTTT	566
Человек-Экзон 43	1	1	AAGAGAAAAauauauauauauauaua	TTTG	567
Человек-Экзон 43	2	1	GAAUUAGCUGUCUAUAGAAAGAGA	tTTT	568
Человек-Экзон 43	3	1	UGAAUUAGCUGUCUAUAGAAAGAG	TTTT	569
Человек-Экзон 43	4	-1	AGCUAAUUCAUUUUUUUACUGUUU	TTTA	570
Человек-Экзон 43	5	1	AUGAAUUAGCUGUCUAUAGAAAGA	TTTC	571
Человек-Экзон 43	6	-1	GCUAAUUCAUUUUUUUACUGUUUU	TTTT	572
Человек-Экзон 43	7	1	AAAAAAAUGAAUUAGCUGUCUAUA	TTTC	573
Человек-Экзон 43	8	-1	UUAAAAUUUUUAUAUUACAGAAUA	TTTA	574
Человек-Экзон 43	9	-1	UAAAAUUUUUAUAUUACAGAAUAU	TTTT	575
Человек-Экзон 43	10	1	AUAUAAAAAUUUUAAAACAGUAAA	TTTT	576
Человек-Экзон 43	11	1	AAUAUAAAAAUUUUAAAACAGUAA	TTTT	577

Человек-Экзон 43	12	1	UAAUAUAAAAAUUUUAAAACAGUA	TTTT	578
Человек-Экзон 43	13	1	GUAAUAUAAAAAUUUUAAAACAGU	TTTT	579
Человек-Экзон 43	14	1	UGUAAUAUAAAAAUUUUAAAACAG	TTTA	580
Человек-Экзон 43	15	1	UAUAUUCUGUAAUAUAAAAAUUUU	TTTT	581
Человек-Экзон 43	16	1	UUAUAUUCUGUAAUAUAAAAAUUU	TTTA	582
Человек-Экзон 43	17	-1	CAGAAUAUAAAAGAUAGUCUACAA	TTTG	583
Человек-Экзон 43	18	1	CUAUCUUUUAUAUUCUGUAAUAUA	TTTT	584
Человек-Экзон 43	19	1	ACUAUCUUUUAUAUUCUGUAAUAU	TTTT	585
Человек-Экзон 43	20	1	GACUAUCUUUUAUAUUCUGUAAUA	TTTA	586
Человек-Экзон 43	21	-1	CAUAGCAAGAAGACAGCAGCAUUG	TTTG	587
Человек-Экзон 43	22	1	CAUUUUGUUAACUUUUUCCCAUUG	TTTC	588
Человек-Экзон 43	23	-1	CAUAUAUUUUUCUUGAUACUUGCA	TTTC	589
Человек-Экзон 43	24	1	AAAUCAUUUCUGCAAGUAUCAAGA	TTTT	590
Человек-Экзон 43	25	1	CAAAUCAUUUCUGCAAGUAUCAAG	TTTT	591
Человек-Экзон 43	26	1	ACAAAUCAUUUCUGCAAGUAUCAA	TTTC	592
Человек-Экзон 43	27	1	AUAAAUUCUACAGUUCCCUGAAAA	TTTG	593
Человек-Экзон 43	28	-1	GAAUUUAUUUCAGUACCCUCCAUG	TTTC	594
Человек-Экзон 43	29	-1	AAUUUAUUUCAGUACCCUCCAUGG	TTTT	595
Человек-Экзон 43	30	1	UGAAAUAAAUUCUACAGUUCCCUG	TTTT	596
Человек-Экзон 43	31	-1	AUUUAUUUCAGUACCCUCCAUGGA	TTTT	597
Человек-Экзон 43	32	1	CUGAAAUAAAUUCUACAGUUCCCU	TTTC	598
Человек-Экзон 43	33	-1	UUUAUUUCAGUACCCUCCAUGGAA	TTTT	599
Человек-Экзон 43	34	-1	UACCCUCCAUGGAAAAAAGACAGG	TTTC	600
Человек-Экзон 43	35	-1	ACCCUCCAUGGAAAAAAGACAGGG	TTTT	601
Человек-Экзон 43	36	-1	CCCUCCAUGGAAAAAAGACAGGGA	TTTT	602
Человек-Экзон 43	37	1	UUUUUUCCAUGGAGGGUACUGAAA	TTTA	603

Человек-Экзон 43	38	1	UGUCUUUUUUCCAUGGAGGGUACU	TTTC	604
Человек-Экзон 6	1	1	CCUUGAGCAAGAACCAUGCAAACU	TTTA	605
Человек-Экзон 6	2	-1	UGCUCAAGGAAUGCAUUUUCUUAU	TTTC	606
Человек-Экзон 6	3	-1	GCUCAAGGAAUGCAUUUUCUUAUG	TTTT	607
Человек-Экзон 6	4	1	UGCAUUCCUUGAGCAAGAACCAUG	TTTG	608
Человек-Экзон 6	5	-1	GAAAAUUUAUUUCCACAUGUAGGU	TTTG	609
Человек-Экзон 6	6	-1	AAAAUUUAUUUCCACAUGUAGGUC	TTTT	610
Человек-Экзон 6	7	-1	AAAUUUAUUUCCACAUGUAGGUCA	TTTT	611
Человек-Экзон 6	8	1	CAUGUGGAAAUAAAUUUUCAUAAG	TTTT	612
Человек-Экзон 6	9	1	ACAUGUGGAAAUAAAUUUUCAUAA	TTTC	613
Человек-Экзон 6	10	-1	CCACAUGUAGGUCAAAAAUGUAAU	TTTC	614
Человек-Экзон 6	11	-1	CACAUGUAGGUCAAAAAUGUAAUG	TTTT	615
Человек-Экзон 6	12	-1	ACAUGUAGGUCAAAAAUGUAAUGA	TTTT	616
Человек-Экзон 6	13	1	ACAUUUUUGACCUACAUGUGGAAA	TTTA	617
Человек-Экзон 6	14	1	CAUUACAUUUUUGACCUACAUGUG	TTTC	618
Человек-Экзон 6	15	-1	AAAAAUAUCAUGGCUGGAUUGCAA	TTTG	619
Человек-Экзон 6	16	-1	GCUGGAUUGCAACAAACCAACAGU	TTTC	620
Человек-Экзон 6	17	-1	CUGGAUUGCAACAAACCAACAGUG	TTTT	621
Человек-Экзон 6	18	1	CCUAUGACUAUGGAUGAGAGCAUU	TTTG	622
Человек-Экзон 6	19	-1	UAGGUAAGAAGAUUACUGAGACAU	TTTA	623
Человек-Экзон 6	20	-1	AUUACUGAGACAUUAAAUAACUUG	TTTA	624
Человек-Экзон 6	21	-1	UUACUGAGACAUUAAAUAACUUGU	TTTT	625
Человек-Экзон 6	22	1	GGGGAAAAAUAUGUCAUCAGAGUC	TTTA	626
Человек-Экзон 6	23	1	CAUGAUCUGGAACCAUACUGGGGA	TTTT	627
Человек-Экзон 6	24	1	ACAUGAUCUGGAACCAUACUGGGG	TTTT	628
Человек-Экзон 6	25	1	GACAUGAUCUGGAACCAUACUGGG	TTTC	629
Человек-Экзон 7	1	1	uacacacauacacaAAGACAAAUA	TTTA	630
Человек-Экзон 7	2	1	uacacauacacacauacacaAAGA	TTTG	631
Человек-Экзон 7	3	1	aacacauacacauacacacauaca	TTtg	632
Человек-Экзон 7	4	1	AUUCCAGUCAAAUAGGUCUGGCCU	ttTT	633
Человек-Экзон 7	5	1	UAUUCCAGUCAAAUAGGUCUGGCC	tTTA	634
Человек-Экзон 7	6	1	GCUGGCAAACCACACUAUUCCAGU	TTTG	635
Человек-Экзон 7	7	1	AGUCGUUGUGUGGCUGACUGCUGG	TTTG	636
Человек-Экзон 7	8	-1	CGCCAGAUAUCAAUUAGGCAUAGA	TTTC	637
Человек-Экзон 7	9	-1	AAACUACUCGAUCCUGAAGGUUGG	TTTA	638
Человек-Экзон 7	10	1	CAUACUAAAAGCAGUGGUAGUCCA	TTTC	639
Человек-Экзон 7	11	1	GAAAACAUUAAACUCUACCAUACU	TTTT	640
Человек-Экзон 7	12	1	UGAAAACAUUAAACUCUACCAUAC	TTTA	641
Человек-Экзон 8	1	-1	UUGUUCAUUAUCCUUUUAGAGUCU	TTTG	642
Человек-Экзон 8	2	1	AAAGGAUAAUGAACAAAUCAAAGU	TTTA	643
Человек-Экзон 8	3	-1	UAUCCUUUUAGAGUCUCAAAUAUA	TTTC	644
Человек-Экзон 8	4	1	ACUCUAAAAGGAUAAUGAACAAAU	TTTG	645
Человек-Экзон 8	5	-1	UUUUAGAGUCUCAAAUAUAGAAAC	TTTG	646
Человек-Экзон 8	6	-1	UUUAGAGUCUCAAAUAUAGAAACC	TTTT	647
Человек-Экзон 8	7	-1	UUAGAGUCUCAAAUAUAGAAACCA	TTTT	648
Человек-Экзон 8	8	1	UUGAGACUCUAAAAGGAUAAUGAA	TTTG	649
Человек-Экзон 8	9	1	UUUGGUUUCUAUAUUUGAGACUCU	TTTT	650
Человек-Экзон 8	10	1	UUUUGGUUUCUAUAUUUGAGACUC	TTTA	651

Человек-Экзон 8	11	-1	AGCAUUGAAGCCAUCCAGGAAGUG	TTTC	652
Человек-Экзон 8	12	1	GCUUCAAUGCUCACUUGUUGAGGC	TTTT	653
Человек-Экзон 8	13	1	GGCUUCAAUGCUCACUUGUUGAGG	TTTG	654
Человек-Экзон 8	14	-1	AGUGGAAAUGUUGCCAAGGCCACC	TTTA	655
Человек-Экзон 8	15	-1	GUUGCCAAGGCCACCUAAAGUGAC	TTTA	656
Человек-Экзон 8	16	-1	GAAGAACAUUUUCAGUUACAUCAU	TTTG	657
Человек-Экзон 8	17	-1	AUCAAAUGCACUAUUCUCAACAGG	TTTA	658
Человек-Экзон 8	18	1	AUAGUGCAUUUGAUGAUGUAACUG	TTTT	659
Человек-Экзон 8	19	1	AAUAGUGCAUUUGAUGAUGUAACU	TTTC	660
Человек-Экзон 8	20	-1	ACUAUUCUCAACAGGUAAAGUGUG	TTTA	661
Человек-Экзон 8	21	1	UACCUAAAAAUGCAUAUAAAACAG	TTTT	662
Человек-Экзон 8	22	1	AUACCUAAAAAUGCAUAUAAAACA	TTTC	663
Человек-Экзон 8	23	1	CACGUAAUACCUAAAAAUGCAUAU	TTTT	664
Человек-Экзон 8	24	1	GCACGUAAUACCUAAAAAUGCAUA	TTTA	665
Человек-Экзон 8	25	1	auauauauGUGCACGUAAUACCUA	TTTT	666
Человек-Экзон 8	26	1	uauauauauGUGCACGUAAUACCU	TTTT	667
Человек-Экзон 8	27	1	auauauauauGUGCACGUAAUACC	TTTA	668
Человек-Экзон 55	1	-1	CUGCACAAUAUUAUAGUUGUUGCU	TTTA	669
Человек-Экзон 55	2	1	AUAAAAAGAGAAAGAUGGAGGAAC	TTTA	670
Человек-Экзон 55	3	1	CACCUAGUGAACUCCAUAAAAAGA	TTTC	671
Человек-Экзон 55	4	1	AUGGUGCACCUAGUGAACUCCAUA	TTTT	672
Человек-Экзон 55	5	1	AAUGGUGCACCUAGUGAACUCCAU	TTTT	673
Человек-Экзон 55	6	1	GAAUGGUGCACCUAGUGAACUCCA	TTTA	674
Человек-Экзон 55	7	1	GACCAAAUGUUCAGAUGCAAUUAU	TTTA	675
Человек-Экзон 55	8	1	UCGCUCACUCACCCUGCAAAGGAC	TTTG	676
Человек-Экзон 55	9	-1	AGUGAGCGAGAGGCUGCUUUGGAA	TTTC	677
Человек-Экзон 55	10	1	GCAGCCUCUCGCUCACUCACCCUG	TTTG	678
Человек-Экзон 55	11	1	UUGCAGUAAUCUAUGAGUUUCUUC	TTTG	679
Человек-Экзон 55	12	-1	CUGCAACAGUUCCCCCUGGACCUG	TTTC	680
Человек-Экзон 55	13	-1	UGCAACAGUUCCCCCUGGACCUGG	TTTT	681
Человек-Экзон 55	14	-1	UUUCUUGCCUGGCUUACAGAAGCU	TTTC	682
Человек-Экзон 55	15	1	UUUCAGCUUCUGUAAGCCAGGCAA	TTTC	683
Человек-Экзон 55	16	-1	GUCCUACAGGAUGCUACCCGUAAG	TTTC	684
Человек-Экзон 55	17	-1	GGCUCCUAGAAGACUCCAAGGGAG	TTTA	685

Человек-Экзон 55	18	-1	GCUCCUAGAAGACUCCAAGGGAGU	TTTT	686
Человек-Экзон 55	19	-1	CUCCAAGGGAGUAAAAGAGCUGAU	TTTC	687
Человек-Экзон 55	20	1	UGGAUCCACAAGAGUGCUAAAGCG	TTTC	688
Человек-Экзон 55	21	1	GUUCAAUUGGAUCCACAAGAGUGC	TTTA	689
Человек-Экзон 55	22	-1	UACUUGUAACUGACAAGCCAGGGA	TTTG	690
Человек-Экзон 55	23	-1	ACUUGUAACUGACAAGCCAGGGAC	TTTT	691
Человек-Экзон 55	24	-1	GUAACUGACAAGCCAGGGACAAAA	TTTG	692
Человек-Экзон 55	25	-1	UAACUGACAAGCCAGGGACAAAAC	TTTT	693
Человек-Экзон 55	26	1	UCCCUGGCUUGUCAGUUACAAGUA	TTTG	694
Человек-G1-экзон51		1	CAGAGUAACAGUCUGAGUAGGAGc	TTTA	695
Человек-G2-экзон51		1	uacuuuguuuagcaauacauggua	TTTC	696
Человек-G3-экзон51		-1	uggcucaaauuguuacucuucaau	TTTA	697
мышь-Экзон23-G1		1	CUUUCAAagaacuuugcagagccu	TTTG	698
мышь-Экзон23-G2		1	guugaaGCCAUUUUGUUGCUCUUU	TTTG	699
мышь-Экзон23-G3		1	guugaaGCCAUUUUAUUGCUCUUU	TTTG	700
мышь-Экзон23-G4		-1	uuuugagGCUCUGCAAAGUUCUUU	TTTC	701
мышь-Экзон23-G5		-1	aguuauuaaugcauagauauucag	TTTA	702
мышь-Экзон23-G6		-1	uauaauaugcccuguaauauaaua	TTTC	703
мышь-Экзон23-G7		1	uaaaggccaaaccucggcuuaccU	TTTC	704
мышь-Экзон23-G8		1	ucaauaucuuugaaggacucuggg	TTTA	705

* В этой таблице заглавные буквы обозначают нуклеотиды sgРНК, которые выравниваются с последовательностью экзона гена. Строчные буквы обозначают нуклеотиды sgРНК, которые выравниваются с последовательностью интрона гена.

VI. Таблицы последовательностей

Таблица 8: Геномные целевые сайты sgРНК в экзоне 51 *Dmd* мыши

sgРНК	Цепь	Целевой сайт	SEQ ID NO:	PAM
Ex51-SA1	3'	AGAGTAACAGTCTGACTGG	706	CAG
Ex51-SD	5'	GAAATGATCATCAAACAGA	707	AGG
Ex51-SA-2	3'	CACTAGAGTAACAGTCTGAC	708	TGG

Таблица 9: Последовательности gРНК, направленные на экзон

51 *Dmd* мыши

sgРНК	Цепь	Целевой сайт	SEQ ID NO:	PAM
Ex51-SA1	3'	CCAGUCAGACUGUUACUCU	709	CAG
Ex51-SD	5'	UCUGUUUGAUGAUCAUUUC	710	AGG
Ex51-SA-2	3'	GUCAGACUGUUACUCUAGUG	711	TGG

Таблица 10: Геномные последовательности-мишени sgРНК, направленных на экзон 51 *Dmd* человека

sgРНК	Цепь	Целевой сайт	SEQ ID NO:	PAM
Ex51-SA	3'	AGAGTAACAGTCTGAGTAG	712	GAG
Ex51-SD	5'	GAGATGATCATCAAGCAGA	713	AGG
Ex51-SA-2	3'	CACCAGAGTAACAGTCTGAG	714	TAG

Таблица 11: Последовательности sgРНК, направленные на экзон 51 Dmd человека

sgРНК	Цепь	Целевой сайт	SEQ ID NO:	PAM
Ex51-SA	3'	CUACUCAGACUGUUACUCU	715	GAG
Ex51-SD	5'	UCUGCUUGAUGAUCAUCUC	716	AGG
Ex51-SA-2	3'	CUCAGACUGUUACUCUGGUG	717	TAG

Таблица 12: Геномные последовательности-мишени sgРНК, направленных на сайты в различные экзонах *Dmd* человека

sgРНК	Цепь	Целевой сайт	SEQ ID NO:	PAM
Экзон51-#1	3'	CAGAGTAACAGTCTGAGTAG	947	GAG
Экзон51-#2	3'	CACCAGAGTAACAGTCTGAG	718	TAG
Экзон51-#3	3'	TATTTTGGGTTTTTGCAAAA	719	AGG
Экзон51-#4	3'	AGTAGGAGCTAAAATATTTT	720	GGG
Экзон51-#5	3'	GAGTAGGAGCTAAAATATTT	721	TGG
Экзон51-#6	3'	ACCAGAGTAACAGTCTGAGT	722	AGG
Экзон51-#7	5'	TCCTACTCAGACTGTTACTC	723	TGG
Экзон51-#8	5'	TACTCTGGTGACACAACCTG	724	TGG
Экзон51-#9	3'	GCAGTTTCCTTAGTAACCAC	725	AGG
Экзон51-#10	5'	GACACAACCTGTGGTTACTA	726	AGG
Экзон51-#11	3'	TGTCACCAGAGTAACAGTCT	727	GAG
Экзон51-#12	3'	AGGTTGTGTCACCAGAGTAA	728	CAG
Экзон51-#13	3'	AACCACAGGTTGTGTCACCA	729	GAG
Экзон51-#14	3'	GTAACCACAGGTTGTGTCAC	730	CAG
Экзон53-#1	5'	ATTTATTTTTCCTTTTATTC	731	TAG
Экзон53-#2	5'	TTTCCTTTTATTCTAGTTGA	732	AAG
Экзон53-#3	3'	TGATTCTGAATTCTTTCAAC	733	TAG
Экзон53-#4	3'	AATTCTTTCAACTAGAATAA	734	AAG
Экзон53-#6	5'	TTATTCTAGTTGAAAGAATT	735	CAG

Экзон53-#7	5'	TAGTTGAAAGAATTCAGAAT	736	CAG
Экзон53-#8	5'	AATTCAGAATCAGTGGGATG	737	AAG
Экзон53-#9	3'	ATTCTTTCAACTAGAATAAA	738	AGG
Экзон53-#10	5'	TTGAAAGAATTCAGAATCAG	739	TGG
Экзон53-#11	5'	TGAAAGAATTCAGAATCAGT	740	GGG
Экзон53-#12	3'	ACTGTTGCCTCCGGTTCTGA	741	AGG
Экзон44-#1	3'	CAGATCTGTCAAATCGCCTG	742	CAG
Экзон44-#2	3'	AAAACGCCGCCATTTCTCAA	743	CAG
Экзон44-#3	3'	AGATCTGTCAAATCGCCTGC	744	AGG
Экзон44-#4	3'	TATGGATCAAGAAAAATAGA	745	TGG
Экзон44-#5	3'	CGCCTGCAGGTAAAAGCATA	746	TGG
Экзон44-#6	5'	ATCCATATGCTTTTACCTGC	747	AGG
Экзон44-#8	5'	TTGACAGATCTGTTGAGAAA	748	TGG
Экзон44-#9	5'	ACAGATCTGTTGAGAAATGG	749	CGG
Экзон44-#11	5'	GGCGATTTGACAGATCTGTT	750	GAG
Экзон44-#13	5'	GGCGTTTTCATTATGATATA	751	AAG
Экзон44-#14	5'	ATGATATAAAGATATTTAAT	752	CAG
Экзон44-#15	5'	GATATTTAATCAGTGGCTAA	753	CAG
Экзон44-#16	5'	ATTTAATCAGTGGCTAACAG	754	AAG
Экзон44-#17	3'	AGAAACTGTTCAGCTTCTGT	755	TAG
Экзон43-#1	5'	GTTTTAAAATTTTTATATTA	756	CAG
Экзон43-#2	5'	TTTTATATTACAGAATATAA	757	AAG
Экзон43-#3	5'	ATATTACAGAATATAAAAGA	758	TAG
Экзон45-#1	3'	GTTCCTGTAAGATACCAAAA	759	AGG
Экзон45-#2	5'	TTGCCTTTTTGGTATCTTAC	760	AGG
Экзон45-#3	5'	TGGTATCTTACAGGAACTCC	761	AGG
Экзон45-#4	5'	ATCTTACAGGAACTCCAGGA	762	TGG
Экзон45-#5	3'	GCCGCTGCCCAATGCCATCC	763	TGG
Экзон45-#6	5'	CAGGAACTCCAGGATGGCAT	764	TGG
Экзон45-#7	5'	AGGAACTCCAGGATGGCATT	765	GGG
Экзон45-#8	5'	TCCAGGATGGCATTGGGCAG	766	CGG
Экзон45-#9	5'	GTCAGAACATTGAATGCAAC	767	TGG
Экзон45-#10	3'	AGTTCCTGTAAGATACCAAA	768	AAG
Экзон45-#11	3'	TGCCATCCTGGAGTTCCTGT	769	AAG
Экзон45-#12	5'	TTGGTATCTTACAGGAACTC	770	CAG
Экзон45-#13	3'	CGCTGCCCAATGCCATCCTG	771	GAG
Экзон45-#14	5'	AACTCCAGGATGGCATTGGG	772	CAG
Экзон45-#15	5'	GGGCAGCGGCAAACTGTTGT	773	CAG

Экзон52-#1	3'	AGATCTGTCAAATCGCCTGC	774	AGG
Экзон52-#2	3'	AATCCTGCATTGTTGCCTGT	775	AAG
Экзон52-#3	5'	CGCTGAAGAACCCTGATACT	776	AAG
Экзон52-#4	3'	GAACAAATATCCCTTAGTAT	777	CAG
Экзон52-#5	3'	CTGTAAGAACAAATATCCCT	778	TAG
Экзон52-#6	5'	CTAAGGGATATTTGTTCTTA	779	CAG
Экзон52-#8	5'	TGTTCTTACAGGCAACAATG	780	CAG
Экзон52-#9	5'	CAACAATGCAGGATTTGGAA	781	CAG
Экзон52-#10	5'	ACAATGCAGGATTTGGAACA	782	GAG
Экзон52-#11	5'	ATTTGGAACAGAGGCGTCCC	783	CAG
Экзон52-#12	5'	ACAGAGGCGTCCCCAGTTGG	784	AAG
Экзон2-#1	5'	TATTTTTTTATTTTGCATTT	785	TAG
Экзон2-#2	5'	TTATTTTGCATTTTAGATGA	786	AAG
Экзон2-#3	5'	ATTTTGCATTTTAGATGAAA	787	GAG
Экзон2-#4	5'	TTGCATTTTAGATGAAAGAG	788	AAG
Экзон2-#5	5'	ATGAAAGAGAAGATGTTCAA	789	AAG

Таблица 13: Последовательности gРНК, направленных на сайты в различных экзонах *Dmd* человека

sgРНК	Цепь	Целевой сайт	SEQ ID NO:	PAM
Экзон51-#1	3'	CUACUCAGACUGUUACUCUG	790	GAG
Экзон51-#2	3'	CUCAGACUGUUACUCUGGUG	791	TAG
Экзон51-#3	3'	UUUUGCAAAAACCCAAAAUA	792	AGG
Экзон51-#4	3'	AAAAUAUUUUAGCUCCUACU	793	GGG
Экзон51-#5	3'	AAAUAUUUUAGCUCCUACUC	794	TGG
Экзон51-#6	3'	ACUCAGACUGUUACUCUGGU	795	AGG
Экзон51-#7	5'	GAGUAACAGUCUGAGUAGGA	796	TGG
Экзон51-#8	5'	CAGGUUGUGUCACCAGAGUA	797	TGG
Экзон51-#9	3'	GUGGUUACUAAGGAAACUGC	798	AGG
Экзон51-#10	5'	UAGUAACCACAGGUUGUGUC	799	AGG
Экзон51-#11	3'	AGACUGUUACUCUGGUGACA	800	GAG
Экзон51-#12	3'	UUACUCUGGUGACACAACCU	801	CAG
Экзон51-#13	3'	UGGUGACACAACCUGUGGUU	802	GAG
Экзон51-#14	3'	GUGACACAACCUGUGGUUAC	803	CAG
Экзон53-#1	5'	GAAUAAAAGGAAAAAUAAAU	804	TAG
Экзон53-#2	5'	UCAACUAGAAUAAAAGGAAA	805	AAG
Экзон53-#3	3'	GUUGAAAGAAUUCAGAAUCA	806	TAG
Экзон53-#4	3'	UUAUUCUAGUUGAAAGAAUU	807	AAG

Экзон53-#6	5'	AAUUCUUUCAACUAGAAUAA	808	CAG
Экзон53-#7	5'	AUUCUGAAUUCUUUCAACUA	809	CAG
Экзон53-#8	5'	CAUCCCACUGAUUCUGAAUU	810	AAG
Экзон53-#9	3'	UUUAUUCUAGUUGAAAGAAU	811	AGG
Экзон53-#10	5'	CUGAUUCUGAAUUCUUUCAA	812	TGG
Экзон53-#11	5'	ACUGAUUCUGAAUUCUUUCA	813	GGG
Экзон53-#12	3'	UCAGAACCGGAGGCAACAGU	814	AGG
Экзон44-#1	3'	CAGGCGAUUUGACAGAUCUG	815	CAG
Экзон44-#2	3'	UUGAGAAAUGGCGGCGUUUU	816	CAG
Экзон44-#3	3'	GCAGGCGAUUUGACAGAUCU	817	AGG
Экзон44-#4	3'	UCUAUUUUUCUUGAUCCAUA	818	TGG
Экзон44-#5	3'	UAUGCUUUUACCUGCAGGCG	819	TGG
Экзон44-#6	5'	GCAGGUAAAAGCAUAUGGAU	820	AGG
Экзон44-#8	5'	UUUCUCAACAGAUCUGUCAA	821	TGG
Экзон44-#9	5'	CCAUUUCUCAACAGAUCUGU	822	CGG
Экзон44-#11	5'	AACAGAUCUGUCAAAUCGCC	823	GAG
Экзон44-#13	5'	UAUAUCAUAAUGAAAACGCC	824	AAG
Экзон44-#14	5'	AUUAAAUAUCUUUAUAUCAU	825	CAG
Экзон44-#15	5'	UUAGCCACUGAUUAAAUAUC	826	CAG
Экзон44-#16	5'	CUGUUAGCCACUGAUUAAAU	827	AAG
Экзон44-#17	3'	ACAGAAGCUGAACAGUUUCU	828	TAG
Экзон43-#1	5'	UAAUAUAAAAAUUUUAAAAC	829	CAG
Экзон43-#2	5'	UUAUAUUCUGUAAUAUAAAA	830	AAG
Экзон43-#3	5'	UCUUUUAUAUUCUGUAAUAU	831	TAG
Экзон45-#1	3'	UUUUGGUAUCUUACAGGAAC	832	AGG
Экзон45-#2	5'	GUAAGAUACCAAAAAGGCAA	833	AGG
Экзон45-#3	5'	GGAGUUCCUGUAAGAUACCA	834	AGG
Экзон45-#4	5'	UCCUGGAGUUCCUGUAAGAU	835	TGG
Экзон45-#5	3'	GGAUGGCAUUGGGCAGCGGC	836	TGG
Экзон45-#6	5'	AUGCCAUCCUGGAGUUCCUG	837	TGG
Экзон45-#7	5'	AAUGCCAUCCUGGAGUUCCU	838	GGG
Экзон45-#8	5'	CUGCCCAAUGCCAUCCUGGA	839	CGG
Экзон45-#9	5'	GUUGCAUUCAAUGUUCUGAC	840	TGG
Экзон45-#10	3'	UUUGGUAUCUUACAGGAACU	841	AAG
Экзон45-#11	3'	ACAGGAACUCCAGGAUGGCA	842	AAG
Экзон45-#12	5'	GAGUUCCUGUAAGAUACCAA	843	CAG
Экзон45-#13	3'	CAGGAUGGCAUUGGGCAGCG	844	GAG
Экзон45-#14	5'	CCCAAUGCCAUCCUGGAGUU	845	CAG

Экзон45-#15	5'	ACAACAGUUUGCCGCUGCCC	846	CAG
Экзон52-#1	3'	GCAGGCGAUUUGACAGAUCU	847	AGG
Экзон52-#2	3'	ACAGGCAACAAUGCAGGAUU	848	AAG
Экзон52-#3	5'	AGUAUCAGGGUUCUUCAGCG	849	AAG
Экзон52-#4	3'	AUACUAAGGGAUAUUUGUUC	850	CAG
Экзон52-#5	3'	AGGGAUAUUUGUUCUUACAG	851	TAG
Экзон52-#6	5'	UAAGAACAAAUAUCCCUUAG	852	CAG
Экзон52-#8	5'	CAUUGUUGCCUGUAAGAACA	853	CAG
Экзон52-#9	5'	UUCCAAAUCCUGCAUUGUUG	854	CAG
Экзон52-#10	5'	UGUUCCAAAUCCUGCAUUGU	855	GAG
Экзон52-#11	5'	GGGACGCCUCUGUUCCAAAU	856	CAG
Экзон52-#12	5'	CCAACUGGGGACGCCUCUGU	857	AAG
Экзон2-#1	5'	ACAGAGGCGUCCCCAGUUGG	858	TAG
Экзон2-#2	5'	UCAUCUAAAAUGCAAAAUAA	859	AAG
Экзон2-#3	5'	UUUCAUCUAAAAUGCAAAAU	860	GAG
Экзон2-#4	5'	CUCUUUCAUCUAAAAUGCAA	861	AAG
Экзон2-#5	5'	UUGAACAUCUUCUCUUUCAU	862	AAG

Таблица 14: Геномная последовательность-мишень для sgРНК, направленных на экзон 51 *Dmd* собаки

sgРНК	Цепь	Целевой сайт	SEQ ID NO:	PAM
Ex51-SA-2	3'	CACCAGAGTAACAGTCTGAC	863	TGG

Таблица 15: Последовательность gРНК, направленной на экзон 51 *Dmd* собаки

sgРНК	Цепь	Целевой сайт	SEQ ID NO:	PAM
Ex51-SA-2	3'	GUCAGACUGUUACUCUGGUG	864	TGG

VII. Примеры

Следующие примеры включены для демонстрации предпочтительных вариантов осуществления изобретения. Специалистам в данной области будет очевидно, что методики, раскрытые в следующих примерах, представляют собой методики, открытые автором изобретения для эффективного применения при осуществлении изобретения, и, таким образом, могут рассматриваться как составляющие предпочтительные варианты осуществления изобретения. Однако специалистам в данной области с учетом настоящего описании должно быть очевидным, что в конкретных вариантах осуществления, которые раскрыты, может быть

внесено множество изменений, и при этом будет получен подобный или аналогичный результат, без отступления от сущности и объема настоящего изобретения.

ПРИМЕР 1 – Материалы и методы

Одобрение исследования. Все экспериментальные процедуры, включающие применение животных в данном исследовании, были рассмотрены и утверждены Институциональным комитетом по содержанию и использованию животных Юго-западного медицинского центра Техасского университета.

CRISPR/Cas9-опосредованная делеция экзона 50 у мышей. Две единых направляющих РНК (sgРНК), специфических к интронной области, окружающей последовательность экзона 50 локуса *Dmd* мыши, клонировали в вектор px330 с использованием следующих праймеров: Dmd exon 50_F1: 5'-CACCGAAATGATGAGTGAAGTTATAT-3'(SEQ ID NO: 926); Dmd exon 50_R1: 5'-AAACATATAACTTCACTCATCATTTC-3' (SEQ ID NO: 927); Dmd exon 50_F2: 5'-CACCGGTTTGTTCAAAAGCGTGGCT-3' (SEQ ID NO: 928); Dmd exon 50_R2: 5'-AAACAGCCACGCTTTTGAACAAAC-3'(SEQ ID NO: 929).

Для *in vitro* транскрипции sgРНК, последовательность промотора T7 добавляли к sgРНК матрице с помощью ПЦР при использовании следующих праймеров: Dmd exon 50_T7-F1: GAATTGTAATACGACTCACTATAGGAATGATGAGTGAAGTTATAT (SEQ ID NO: 930); Dmd exon 50_T7-F2: GAATTGTAATACGACTCACTATAGGGTTTGTTCAAAAGCGTGGCT (SEQ ID NO: 931); Dmd exon 50_T7-Rv: AAAAGCACCGACTCGGTGCCAC (SEQ ID NO: 932).

Выделенные из геля продукты ПЦР использовали в качестве матрицы для *in vitro* транскрипции с использованием набора MEGAshortscript T7 (Life Technologies). sgРНК очищали с использованием набора MEGAclear (Life Technologies) и элюировали безнуклеазной водой (Ambion). Концентрацию направляющей РНК измеряли с помощью прибора NanoDrop (Thermo Scientific).

Генотипирование мышей ΔEx50. Мышей ΔEx50 генотипировали при использовании праймеров, охватывающих целевую область: Geno50-F: 5'-GGATTGACTGAAATGATGGCCAAGG-3' (SEQ ID NO: 937); Geno50-R: 5'-CTGCCACGATTACTCTGCTTCCAG-3' (SEQ ID NO: 938). Биопсии хвоста растворяли в 100 мкл 25 мМ NaOH, 0,2 мМ ЭДТА (pH 12) в течение

20 мин при 95°C. Хвосты быстро центрифугировали, затем добавляли 100 мкл 40 мМ Tris·HCl (pH 5) и перемешивали для гомогенизации. Два микролитра этой смеси использовали для последующих реакций ПЦР с праймерами, указанными ниже, с последующим гель-электрофорезом.

Плазмиды. Плазмиду pSpCas9 (BB)-2A-GFP (PX458), содержащую человеческий кодон-оптимизированный ген SpCas9 с 2A-EGFP и каркас sgРНК, приобретали в Addgene (Плазмида #48138). Плазмиды AAV TRIPCR были получены от д-ра Дирка Гримма (Клиника при Гейдельбергском университете). Клонирование sgРНК выполняли при использовании сайта BbsI. Плазмида pGL3-CK8e была предоставлена д-ром Стивеном Хаушкой (Отделение биохимии, Вашингтонский университет, Сиэтл, США). Плазмида AAV-miniCMV-Cas9-short-PolyA была получена от д-ра Дирка Гримма (Клиника при Гейдельбергском университете). Для получения окончательной конструкции вектора AAV9-CK8-CRISPR/Cas9, используемого в этой рукописи, AAV-miniCMV-Cas9-shortPolyA расщепляли ферментами PacI и NheI для удаления промотора miniCMV. Промотор CK8 амплифицировали с плазмиды pGL3-CK8e при использовании праймеров, содержащих последовательность сайтов PacI и NheI, и клонировали в расщепленный вектор с получением плазмиды AAV-CK8-Cas9-shortPolyA.

Идентификация sgРНК и клонирование для пропуска экзона 51. Направляющие РНК к Экзону 51 *Dmd* определяли при использовании crispr.mit.edu. Направляющие последовательности клонировали в плазмиду Addgene #42230 (6), переданную в дар Фен Чжаном (Feng Zhang), с использованием следующих праймеров: *Dmd* exon 51_F1: 5'-CACCGAGAGTAACAGTCTGACTGG-3' (SEQ ID NO: 942); *Dmd* exon 51_R1: 5'-AAACGTCAGACTGTTACTCTAGTGC-3' (SEQ ID NO: 943); *Dmd* exon 51_F2: 5'-CACCGCACTAGAGTAACAGTCTGAC-3' (SEQ ID NO: 944); *Dmd* exon 51_R2: 5'-AAACCCAGTCAGACTGTTACTCTC-3' (SEQ ID NO: 945). Направляющие последовательности тестировали в культуре при использовании клеток 10T1/2 перед клонированием в каркас AAV.

Сборка тройной sgРНК в каркасе AAV при использовании системы Golden Gate. Сборка системы клонирования в каркасе AAV TRISPR основана на двух последовательных этапах клонирования

Golden Gate Assembly. Первый этап – сборка gРНК в донорной плазмиде. Отжиг олигонуклеотидов проводили при нагревании реакции, содержащей 2,5 мкл каждого олига (0,5 мкМ), 5 мкл NEBuffer 2 (NEB) и 40 мкл ddH_2O, до 95°С в течение 5 минут при использовании нагревательного блока. Для реакции сборки в донорной плазмиде, смешивали 40 фмоль (~100 нг) целевого каркаса, 1 мкл отожженных, разбавленных олигов, 0,75 мкл Esp3I, 1 мкл буфера tango (оба производства Thermo Scientific), 1 мкл ДНК-лигазы Т4 (400 Ед/мкл) (NEB), а также АТФ и DTT в конечной концентрации 1 мМ в общем объеме 10 мкл. При использовании амплификатора проводили от 25 до 50 циклов при 37°С/3 мин, затем 20°С/5 минут. Рестриктазу и лигазу денатурировали нагреванием до 80°С в течение 20 минут. 3 мкл этой реакционной смеси использовали для трансформации хемо-компетентных бактерий, подращивали в SOC (37°С, 800 об/мин, 40 мин) и сеяли на чашках с LB-агаром, содержащим хлорамфеникол (25 мкг/мл). Отожженные олигонуклеотиды, кодирующие sgРНК, клонировали в донорные плазмиды, которые несут маркер отрицательной селекции ccdB (для уменьшения фона при клонировании), а также ген устойчивости к хлорамфениколу. Для проверки правильности сборки плазмиду секвенировали с использованием праймера Dono-R-5'-GTATGTTGTGTGGAATTGTGAG-3'(SEQ ID NO: 948). Второй этап состоит в том, что три из этих донорных плазмид, направляющих экспрессию одной sgРНК под транскрипционным контролем промотора U6, H1 или 7SK, объединяли в пул во второй сборке Golden Gate вместе с реципиентной плазмидой, которая несет повторы ITR AAV. Реакция сборки будет содержать все четыре плазмиды: донорную плазмиду – #1-U6-sgRNA, донорную плазмиду – #2-H1-sgRNA, донорную плазмиду-#3-7SK-sgRNA и реципиентную плазмиду, содержащую ITR. При расщеплении BbsI будут получены уникальные липкие концы для каждого фрагмента (U6, H1, 7SK, реципиентный каркас). Во время процедуры лигирования эти липкие концы отжигаются с получением только кольцевой плазмиды, когда три кассеты соответствуют друг другу.

Измерение сывороточной креатинкиназы (СК). Мышиную СК сыворотки измеряли в Центральной лаборатории метаболического

фенотипирования в Юго-западном медицинском центре Техасского университета. Кровь забирали из подчелюстной вены и измеряли уровень СК сыворотки с помощью слайдов для анализа СК VITROS Chemistry 7 Products, для количественного измерения активности СК при использовании системы VITROS 250 Chemistry System.

Исследования *in vivo* на собаках. Внутримышечные (в/м) инъекции делали 1-месячным собакам ΔE50 MD под общим ингаляционным наркозом. Одну инъекцию делали в 4 разных участка левой краниальной большеберцовой мышцы. Все собаки получали премедикацию опиатом (бупренорфин), периоперационную антибиотикопрофилактику и 4 дня послеоперационной анальгезии (карпрофен). Две обработанных собаки получали rAAV в смеси с миоредактирующими компонентами, AAV-Cas9 и AAV-sgRNA-ex51, которые были приготовлены в лактированном растворе Рингера с конечной концентрацией 1E13vg на инъецируемую мышцу.

ПРИМЕР 2 – Результаты

Гуманизированная модель DMD.

Областью горячей точки с наиболее частыми мутациями у пациентов с МДД является область между 45 и 51 экзонами, и пропуск экзона 51 можно использовать для лечения самой большой группы (13-14%) пациентов. Для исследования CRISPR/Cas9-опосредованного пропуска экзона 51 *in vivo*, авторы изобретения создали модель на мышах, которая имитирует область "горячей точки" человека путем делеции экзона 50 с использованием системы CRISPR/Cas9, направляемой 2-мя sgРНК (ФИГ. 1A). Делецию экзона 50 подтверждали секвенированием ДНК (ФИГ. 1B). Делеция экзона 50 помещала ген дистрофина вне рамки считывания, что приводило к отсутствию белка дистрофина в скелетных мышцах и сердце (ФИГ. 1C-E). Мыши, у которых отсутствовал экзон 50, демонстрировали выраженные дистрофические мышечные изменения к возрасту 2 месяца (ФИГ. 1E). Анализ сыворотки у дельта-экзон 50 мышей показал значительное повышение уровней креатинкиназы (СК), свидетельствующее о повреждении мышц (ФИГ. 1F). В совокупности, экспрессия белка дистрофина, гистологическое исследование мышц и уровни СК в сыворотке подтверждали дистрофический фенотип модели на мышах ΔEx50.

Восстановление экспрессии дистрофина с применением стратегии одиночного разреза с пропуском экзона 51. Cas9 *S. pyogenes* требует NAG/NGG в качестве последовательности PAM для создания двухцепочечного разрыва ДНК. Интересно то, что универсальные акцепторные и донорные сайты сплайсинга экзонов содержат NAG или NGG (ФИГ. 3B). Поэтому для коррекции рамки считывания и экспрессии дистрофина в модели на мышах ΔEx50 авторы изобретения создали sgРНК, направленную на акцепторные и донорные сайты сплайсинга экзона 51 для его удаления, с восстановлением рамки считывания белка дистрофина. Для проверки, способны ли sgРНК гиды эффективно создавать разрезы, авторы изобретения сначала оценили их эффективность в линиях клеток мыши и человека (ФИГ. 5). Для определения наиболее эффективного способа коррекции рамки считывания DMD, авторы изобретения сравнили 2 разные стратегии: (1) стратегию двойного гида, в которой одна копия первой sgРНК, направленной на акцепторный сайт сплайсинга (sgRNA-SA), и одна копия второй sgРНК, направленной на донорный акцепторный сайт (sgRNA-SD), клонировали в вектор rAAV9-sgRNA; (2) тройную стратегию, в которой авторы изобретения клонировали 3 копии одинаковой sgРНК (sgRNA-SA) в вектор rAAV9-sgRNA (ФИГ. 3C). Экспрессию каждой копии sgRNA-SA направлял определенный РНК промотор (U6, H1 и 7SK). Авторы изобретения создали AAV-Cas9 с использованием вектора AAV-Cas9 (CK8-Cas9-shortPolyA), в котором используется промотор CK8 для направления экспрессии гуманизированной SpCas9, в частности, в скелетных мышцах и тканях сердца. После внутримышечной (в/м) инъекции различных AAV мышам P12 исследовали мышечные ткани. ОТ-ПЦР РНК мышей ΔEx50, которым вводили AAV-Tri-SA и AAV-SA+SD, показала, что делеция экзона 51 (ΔEx50-51) обеспечивала сплайсинг от экзона 49 до 52 (ФИГ. 2A, нижняя полоса). Секвенирование продуктов ОТ-ПЦР полосы ΔEx50-51 подтвердило, что экзон 49 был соединен с экзоном 52 (ФИГ. 2B). Неожиданно анализ ОТ-ПЦР показал, что стратегия одиночного разреза с использованием тройного варианта sgRNA-SA (AAV-Tri-SA) столь же эффективна, как и применение двух sgRNA – sgRNA-SA и sgRNA-SD (AAV-SA+SD).

Чтобы дополнительно оценить эффективность стратегии редактирования AAV-Tri-SA, авторы изобретения провели гистологическое исследование мышцы после инъекции для оценки количества волокон, которые экспрессируют дистрофин в полных мышечных срезах. Примечательно то, что иммуноокрашивание дистрофина в криосрезах мышц у мышей ΔEx50, которым вводили AAV-Tri-SA, обнаружило значительно более высокое количество дистрофин-положительных волокон (в среднем 43±0,9%) по сравнению с мышцами мышей ΔEx50, которым вводили AAV-SA+SD (в среднем 31±0,1%) (ФИГ. 3D-E, ФИГ. 6). Вестерн-блоттинг подтвердил восстановление экспрессии дистрофина в скелетных мышцах. (ФИГ. 3F-G). Окрашивание мышц гематоксилином и эозином (Г-Э) показало, что патогистологические признаки мышечной дистрофии, такие как некротические миофибрилы, уменьшились в TA мышцах через 3 недели после доставки AAV (ФИГ. 4A). Количественный анализ распределения областей мышечных волокон показал явное увеличение размера волокон как мышц, обработанных AAV-Tri-SA, так и AAV-SA-SD, по сравнению с мышцами ΔEx50 (ФИГ. 4B). Тем не менее, мышцы, обработанные AAV-Tri-SA, демонстрировали большее снижение частоты мелких волокон (<500 мкм) по сравнению с мышцами, обработанными AAV-SA+SD. В совокупности эти результаты демонстрируют, что направленное воздействие на акцепторный сайт сплайсинга экзона 51 посредством одного одиночного разреза с приенением AAV-Tri-SA является высокоэффективным для восстановления экспрессии дистрофина при МДД. Этот подход может применяться при многих нарушениях, которые можно корректировать путем пропуска экзон.

Оптимизация стратегии редактирования генома с одиночным разрезом ДНК. Cas9 *S. pyogenes*, направляемая молекулами sgРНК, связывается с целевым геномным локусом вблизи PAM и создает двухцепочечный разрыв ДНК (DSB) за 3 нуклеотида до последовательности PAM. Для дальнейшей оценки эффективности способа путем направленного воздействия на акцепторный сайт сплайсинга авторы изобретения разработали вторую тройную направляющую конструкцию sgРНК (sgRNA-ex51-SA2), направленную на область, примыкающую к акцепторному сайту сплайсинга экзона 51.

Эта gРНК использует последовательность PAM, расположенную через 3 нуклеотида в экзоне, чтобы создавать DSB вблизи акцепторного сайта сплайсинга для экзона 51 (ФИГ. 7A-ФИГ. 7B). Разрез поблизости от акцепторной области сплайсинга и в последовательности экзона приводил к событиям восстановления рамки считывания и событиям пропуска экзонов. Кроме того, конструирование sgРНК в последовательности экзона, которая демонстрирует более высокую консервативность, чем последовательность интронов у разных видов, облегчает трансляцию sgРНК в других видах.

ДНК-разрезающую активность Cas9 в сочетании с sgRNA-ex51-SA2 оценивали в фибробластах мышей 10T1/2 при использовании анализа со специфичной к некомплементарным нуклеотидам эндонуклеазой I T7 (T7E1) (ФИГ. 8A). Для исследования типа мутаций, создаваемых Cas9 в сочетании с sgRNA-51-SA2, проводили геномный анализ методом глубокого секвенирования. Секвенирование показало, что 9,3% мутаций содержали одиночную вставку аденозина (A), расположенную за 3 нуклеотида 3' от последовательности PAM. Кроме того, 7,3% мутаций содержали делеции, охватывающие акцепторный сайт сплайсинга и высоко предсказуемый экзонный энхансерный сайт сплайсинга для экзона 51 (ФИГ. 8B). sgRNA-ex51-SA2 соответствует высококонсервативной области гена *Dmd* (ФИГ. 8C-D), и авторы изобретения проверили способность Cas9 и sgRNA-51 человека разрезать человечееский локус *Dmd* в клетках 293T. Анализ T7E1 выявил явное расщепление в предсказанном сайте (ФИГ. 8E). Аналогичным образом, анализ последовательности показал, что Cas9 в сочетании с sgRNA-ex51-SA2 человека создает такую же вставку аденозина (A) и ряд различных делеций около сайта расщепления (ФИГ. 8F).

Для *in vivo* доставки Cas9 и sgRNA-ex51-SA2 в скелетные мышцы и сердечную ткань использовали аденоассоциированный вирус 9 (AAV9), который демонстрирует предпочтительный тропизм в отношении этих тканей. Для дальнейшего повышения мышечно-специфической экспрессии использовали вектор AAV9-Cas9 (CK8e-Cas9-shortPolyA), который содержит мышечно-специфическую CK регуляторную кассету, называемую промотором CK8e, который

высокоспецифичен для экспрессии в мышце и сердце (ФИГ. 9А). Вместе эта 436 пн мышечно-специфическая кассета и 4101 пн кДНК Cas9 не превышают лимита упаковки AAV9. Экспрессия каждой sgРНК направлялась одним из трех промоторов РНК-полимеразы III (U6, H1 и 7SK) (ФИГ. 9В).

Коррекция рамки считывания дистрофина у мышей ΔEx50 путем одиночного разреза ДНК. sgRNA-ex51-SA2 доставляли мышам в тройной копии (AAV-Tri-SA2) вместе с Cas9 (AAV-Cas9) путем внутримышечной (в/м) инъекции. После инъекции исследовали мышечные ткани. Эффективность таргетинга *in vivo* оценивали с помощью ОТ-ПЦР с праймерами к последовательностям в экзонах 48 и 53 и анализа T7E1 для целевых областей генома. С целью исследования, было ли достигнуто эффективное расщепление мишени, авторы изобретения амплифицировали область длиной 771 пн, охватывающую сайт-мишень, и исследовали ее с помощью анализа T7EI (фигура 10А). Активность SpCas9 с соответствующей sgРНК исследовали на сайте-мишени. Анализы T7EI выявили мутагенез локуса *Dmd* после доставки AAV-Cas9 и AAV9-sgRNA-51-SA2 (фиг. 10А). Чтобы исследовать тип мутаций, полученных у мышей ΔEx50, которым вводили Cas9 и AAV9, экспрессирующие sgРНК, геномные продукты амплификации ПЦР, охватывающие сайт-мишень, анализировали методом глубокого секвенирования ампликона. Глубокое секвенирование области-мишени показало, что 27,9% от общего числа чтений содержали изменения в целевом геномном сайте (ФИГ. 10В). В среднем 15% идентифицированных мутаций содержали одну и ту же инсерцию А, наблюдаемую *in vitro* в клетках мыши 10T1/2 и человеческих клетках 293. Делеции, идентифицированные с помощью этого метода, охватывают высоко предсказуемый сайт экзонного энхансерного сайта сплайсинга экзона 51 (ФИГ. 10В).

Продукты ОТ-ПЦР с РНК из мышцы мышей ΔEx50, которым внутримышечно вводили AAV9-Cas9 и AAV9-sgRNA-51, показали, что делеция экзона 51 (ΔEx50-51) вызывала сплайсинг от экзона 49 до 52 (ФИГ. 11А, нижняя полоса). При секвенировании продуктов ОТ-ПЦР полосы ΔEx50-51 было подтверждено, что экзон 49 был сплайсирован с экзоном 52. Чтобы дополнительно определить мутации, введенные с помощью предложенной авторами изобретения

стратегии редактирования генов, продукты амплификации ОТ-ПЦР из 4 образцов непосредственно подвергали клонированию с заменой тимидина на аденозин на основе топоизомеразы (TOPO-TA), без очистки из геля, а затем секвенированию. Неожиданно анализ последовательности 40 клонов из каждого образца показал, что в дополнение к продуктам кДНК с пропуском экзона 51 (ΔEx50-51), идентифицированным в 15% секвенированных клонов, мыши ΔEx50, которым вводили AAV9-Cas9 и AAV9-sgRNA-51, показали высокую частоту событий восстановления рамки считывания. Из всех секвенированных клонов 63% содержали одиночную нуклеотидную вставку в последовательности экзона 51 (ФИГ. 11 B-C). Наиболее доминирующей из наблюдаемых инсерционных мутаций была инсерция A.

На гелях инсерция A была неотличима по размеру от нередактированных продуктов кДНК, поэтому проводили глубокое секвенирование для определения представленности этой инсерции по сравнению с другими небольшими инсерциями. Глубокое секвенирование верхней полосы, содержащей нередактированный продукт кДНК и продукты кДНК с восстановленной рамкой считывания, показало, что 69,22% от общего числа чтений содержали продукты кДНК с восстановленной рамкой считывания с инсерцией A, 17,71% содержали нередактированный продукт кДНК, а остальные содержали небольшие делеции и вставки (ФИГ. 11D). Анализ методом глубокого секвенирования мышей ΔEx50, не получавших инъеций, подтвердил, что вставка A является результатом редактирования, выполненного Cas9. Эти результаты глубокого секвенирования ампликона подтвердили результаты клонирования TOPO-TA и секвенирования. В совокупности анализ ОТ-ПЦР показал, что мыши ΔEx50, которым вводили AAV9-Cas9 и AAV9-sgRNA-51-SA2, показали высокую частоту событий восстановления рамки считывания, при этом продукты кДНК содержали вставку A в последовательности экзона 51 в дополнение к событиям пропуска экзона 51 в результате делеции в высококонсервативной экзонной энхансерной области сплайсинга.

Восстановление экспрессии дистрофина после внутримышечной AAV9 доставки Cas9 и sgRNA-51-SA2. Примечательно то, что

иммуноокрашивание дистрофина в криосрезах мышц мышей ΔEx50, которым вводили AAV-Tri-SA2, выявило значительно более высокое количество дистрофин-положительных волокон со средним 99% восстановлением нормальных волокон (ФИГ. 12A, ФИГ. 13). Вестерн-блоттинг подтвердил восстановление экспрессии дистрофина в скелетных мышцах (ФИГ. 12C, ФИГ. 12D). Окрашивание мышц гематоксилином и эозином (Г-Э) показало, что патогистологические признаки мышечной дистрофии, такие как некротические мышечные волокна, были скорректированы в TA мышцах через 3 недели после доставки AAV (ФИГ. 12B и ФИГ. 14). Этот способ, в котором используется особая конструкция sgРНК, представляет большой успех в повышении эффективности коррекции МДД с возможностью непосредственного применения у пациентов с наиболее распространенными мутациями дистрофина.

Восстановление экспрессии дистрофина после внутримышечных инъекций AAV9-Cas9 в сочетании с различными AAV9, экспрессирующими одиночную копию или тройную копию sgРНК. Для оценки преимуществ экспрессии sgRNA-ex51-SA2 с тремя промоторами *in vivo* изучали различные конструкции, где экспрессию sgРНК контролировал один промотор РНК-полимеразы III (U6 или H1, или 7SK) и, отдельно, три промотора РНК-полимеразы III (U6, H1 и 7SK) (ФИГ. 15A). Авторы изобретения доставляли sgRNA-ex51-SA2 в одной копии под контролем, в отдельности, промотора U6 (AAV9-U6-sgRNA-51-SA2), промотора H1 (AAV9-H1-sgRNA-51-SA2), промотора 7SK (AAV9-7SK-sgRNA-51-SA2), и тройной копии (AAV9-Triple-sgRNA-51-SA2). После внутримышечной (в/м) инъекции различных AAV9 мышам P12 исследовали мышечные ткани. Неожиданно иммуноокрашивание дистрофина в криосрезах мышц мышей ΔEx50, которым вводили AAV9-Triple-sgRNA-51-SA2, показало значительно более высокое количество дистрофин-положительных волокон со средним 95% восстановлением нормальных волокон по сравнению с мышами ΔEx50, которым вводили AAV9-U6-sgRNA-51-SA2, AAV-H1-sgRNA-51-SA2 и AAV-7SK-sgRNA-51-SA2, в среднем показавшие 70%; 40% и 50% восстановления нормальных волокон соответственно (ФИГ. 15B).

Восстановление структуры и функции мышц после системной

доставки AAV9-Cas9 и AAV9-sgRNA-51-SA2. Системная доставка AAV9-Cas9 и AAV9-sgRNA-51-SA2 мышам P4 ΔEx50 вызвала обширную экспрессию дистрофина в сердце, трехглавой мышце, передней большеберцовой мышце (TA) и диафрагме у подвергнутых редактированию гена мышей ΔEx50 через 4 и 8 недель после инъекции (ФИГ. 16A и ФИГ. 17A). Вестерн-блоттинг подтвердил восстановление экспрессии дистрофина в скелетных и сердечных мышцах (ФИГ. 16B и ФИГ. 17B). Сила хвата, которую также тестировали, показала значительное увеличение мышечной силы мышей ΔEx50 через 4 недели постле внутрибрюшинной инъекции AAV9 по сравнению с контрольными мышами ΔEx50 (контроль дикого типа, 92,6±1,63; контроль ΔEx50 50,5±1,85; ΔEx50-AAV9-sgRNA-51 79,7±2,63) (ФИГ. 18A). Соответственно, AAV9-sgRNA-51-SA2 генoредактированные мыши ΔEx50 также показали значительное снижение концентраций СК сыворотки по сравнению с контрольными мышами ΔEx50 (ФИГ. 18B).

Коррекция экспрессии дистрофина в модели миодистрофии Дюшена на собаках. Для дополнительной оценки эффективности и терапевтического потенциала этого нового подхода, авторы изобретения исследовали коррекцию вызывающей заболевание мутации в модели МДД на собаках, а именно собаке ΔE50-MD, которая несет миссенс-мутацию в 5' донорском сайте сплайсинга экзона 50, которая приводит к делеции экзона 50 (Walmsley et al., 2010). Собака ΔE50-MD является идеальной моделью на собаках для исследования редактирования гена как подхода к постоянной коррекции наиболее распространенных мутаций при МДД у людей.

Авторы изобретения использовали sgРНК (sgRNA-ex51-SA2), направленную на тот же геномный локус акцепторной области сплайсинга. Последовательность sgRNA-ex51-SA2 является высоко консервативной. Последовательности-мишени sgRNA-ex51-SA2 мыши и sgRNA-ex51-SA2 собаки (sgRNA-D-ex51-SA2) отличаются только одним нуклеотидом (Таблица 4). Авторы изобретения доставили sgRNA-D-ex51-SA2 в трех копиях (AAV-D-Tri-SA2). После внутримышечной (в/м) инъекции AAV собакам возрастом 1 месяц исследовали мышечные ткани. Примечательно то, что иммуноокрашивание дистрофина в криосрезах мышц мышей ΔEx50, которым вводили AAV-

Tri-SA2, выявило значительно более высокое количество дистрофин-положительных волокон (ФИГ. 19). Вестерн-блоттинг подтвердил восстановление экспрессии дистрофина в скелетных мышцах (ФИГ. 20). Окрашивание мышц гематоксилином и эозином (Г-Э) показало, что патогистологические признаки мышечной дистрофии, такие как некротические мышечные волокна, были уменьшены в краниальной большеберцовой мышце через 6 недель после доставки AAV (ФИГ. 21). Этот способ, в котором применяется особая конструкция sgРНК, представляет собой большой успех в повышении эффективности коррекции МДД с возможностью непосредственного применения у пациентов с наиболее распространенными мутациями дистрофина.

Все композиции и/или способы, раскрытые и заявленные в настоящем документе, могут быть изготовлены и осуществлены без чрезмерных экспериментов с учетом настоящего описания. Хотя композиции и способы согласно настоящему изобретению были описаны в отношении предпочтительных вариантов осуществления, специалистам в данной области техники будет очевидно, что изменения могут быть внесены в композиции и/или способы и на этапах или в последовательности этапов способа, описанного в настоящем документе, без отступления от концепции, сущности и объема настоящего описания. В частности, будет очевидно, что некоторые средства, которые являются как химически, так и физиологически родственными, могут заменять средства, описанные в настоящем документе, и при этом будут достигнуты такие же или аналогичные результаты. Все подобные аналогичные заменители и модификации, очевидные для специалистов в данной области, считаются включенными в сущность, объем и концепцию настоящего изобретения, определенные в прилагаемой формуле изобретения.

VII. Использованные источники

Следующие источники, в той степени, в которой они предоставляют примерные процедурные или иные подробности, которые дополняют изложенные в настоящем документе, прямо включены в настоящий документ посредством отсылки.

Angel et al., Mol. Cell. Biol., 7:2256, 1987a.

Angel et al., Cell, 49:729, 1987b.

Aartsma-Rus et al., Hum. Mutat. 30, 293-299, 2009.

Baichwal and Sugden, In: Gene Transfer, Kucherlapati (Ed), NY, Plenum Press, 117-148, 1986.

Banerji et al., Cell, 27(2 Pt 1):299-308, 1981.

Banerji et al., Cell, 33(3):729-740, 1983.

Barnes et al., J. Biol. Chem., 272(17):11510-7, 1997.

Baskin et al., EMBO Mol Med 6, 1610-1621, 2014.

Benvenisty and Neshif, Proc. Natl. Acad. Sci. USA, 83:9551-9555, 1986.

Berkhout et al., Cell, 59:273-282, 1989.

Bhavsar et al., Genomics, 35(1):11-23, 1996.

Bikard et al., Nucleic Acids Res. 41(15): 7429-7437, 2013.

Blanar et al., EMBO J., 8:1139, 1989.

Bodine and Ley, EMBO J., 6:2997, 1987.

Boshart et al., Cell, 41:521, 1985.

Bostick et al., Mol Ther 19, 1826-1832, 2011.

Bosze et al., EMBO J., 5(7):1615-1623, 1986.

Braddock et al., Cell, 58:269, 1989.

Brinster et al., Proc. Natl. Acad. Sci. USA, 82(13):4438-4442, 1985.

Bulla and Siddiqui, J. Virol., 62:1437, 1986.

Burridge et al., Nat. Methods 11, 855-860, 2014.

Bushby et al., Lancet Neurol., 9(1): 77-93 (2010).

Bushby et al., Lancet Neurol., 9(2): 177-198 (2010).

Campbell & Kahl, Nature 338, 259-262, 1989.

Campbell and Villarreal, Mol. Cell. Biol., 8:1993, 1988.

Campere and Tilghman, Genes and Dev., 3:537, 1989.

Campo et al., Nature, 303:77, 1983.

Celander and Haseltine, J. Virology, 61:269, 1987.

Celander et al., J. Virology, 62:1314, 1988.

Chandler et al., Cell, 33:489, 1983.

Chang et al., Mol. Cell. Biol., 9:2153, 1989.

Chang et al., Stem Cells., 27:1042-1049, 2009.

Chatterjee et al., Proc. Natl. Acad. Sci. USA, 86:9114, 1989.

Chen and Okayama, Mol. Cell Biol., 7:2745-2752, 1987.

Cho et al., Nat. Biotechnol. 31(3): 230-232, 2013.

Choi et al., Cell, 53:519, 1988.

Cirak et al., The Lancet 378, 595-605, 2011.

Coffin, In: Virology, Fields et al. (Eds.), Raven Press, NY, 1437-1500, 1990.

Cohen et al., J. Cell. Physiol., 5:75, 1987.

Costa et al., Mol. Cell. Biol., 8:81, 1988.

Couch et al., Am. Rev. Resp. Dis., 88:394-403, 1963.

Coupar et al., Gene, 68:1-10, 1988.

Cripe et al., EMBO J., 6:3745, 1987.

Culotta and Hamer, Mol. Cell. Biol., 9:1376, 1989.

Dandolo et al., J. Virology, 47:55-64, 1983.

De Villiers et al., Nature, 312(5991):242-246, 1984.

Deschamps et al., Science, 230:1174-1177, 1985.

DeWitt et al., Sci Transl Med 8, 360ra134-360ra134, 2016.

Donnelly et al., J. Gen. Virol. 82, 1027-1041, 2001.

Dubensky et al., Proc. Natl. Acad. Sci. USA, 81:7529-7533, 1984.

Edbrooke et al., Mol. Cell. Biol., 9:1908, 1989.

Edlund et al., Science, 230:912-916, 1985.

EP 0273085

Fechheimer et al., Proc Natl. Acad. Sci. USA, 84:8463-8467, 1987.

Feng and Holland, Nature, 334:6178, 1988.

Ferkol et al., FASEB J., 7:1081-1091, 1993.

Firak et al., Mol. Cell. Biol., 6:3667, 1986.

Foecking et al., Gene, 45(1):101-105, 1986.

Fonfara et al., Nature 532, 517-521, 2016.

Fraley et al., Proc Natl. Acad. Sci. USA, 76:3348-3352, 1979

Franz et al., Cardoscience, 5(4):235-43, 1994.

Friedmann, Science, 244:1275-1281, 1989.

Fujita et al., Cell, 49:357, 1987.

Ghosh and Bachhawat, In: Liver Diseases, Targeted Diagnosis and Therapy Using Specific Receptors and Ligands, Wu et al. (Eds.), Marcel Dekker, NY, 87-104, 1991.

Ghosh Choudhury et al., EMBO J., 6:1733 1739, 1987.

Gilles et al., Cell, 33:717, 1983.

Gloss et al., EMBO J., 6:3735, 1987.

Godbout et al., Mol. Cell. Biol., 8:1169, 1988.

Gomez Foix et al., J. Biol. Chem., 267:25129 25134, 1992.

Goncalves et al., Mol Ther, 19(7): 1331-1341 (2011).

Goodbourn and Maniatis, Proc. Natl. Acad. Sci. USA, 85:1447, 1988.

Goodbourn et al., Cell, 45:601, 1986.

Gopal, Mol. Cell Biol., 5:1188-1190, 1985.

Gopal-Srivastava et al., J. Mol. Cell. Biol., 15(12):7081-90, 1995.

Graham and Prevec, In: Methods in Molecular Biology: Gene Transfer and Expression Protocol, Murray (Ed.), Humana Press, Clifton, NJ, 7:109 128, 1991.

Graham and van der Eb, Virology, 52:456-467, 1973.

Graham et al., J. Gen. Virol., 36:59-72, 1977.

Greene et al., Immunology Today, 10:272, 1989

Grosschedl and Baltimore, Cell, 41:885, 1985.

Grunhaus and Horwitz, Seminar in Virology, 3:237-252, 1992.

Harland and Weintraub, J. Cell Biol., 101:1094-1099, 1985.

Haslinger and Karin, Proc. Natl. Acad. Sci. USA, 82:8572, 1985.

Hauber and Cullen, J. Virology, 62:673, 1988.

Hen et al., Nature, 321:249, 1986.

Hensel et al., Lymphokine Res., 8:347, 1989.

Hermonat and Muzycska, Proc. Nat'l Acad. Sci. USA, 81:6466-6470, 1984.

Herr and Clarke, Cell, 45:461, 1986.

Hersdorffer et al., DNA Cell Biol., 9:713-723, 1990.

Herz and Gerard, Proc. Nat'l. Acad. Sci. USA 90:2812 2816, 1993.

Hirochika et al., J. Virol., 61:2599, 1987.

Holbrook et al., Virology, 157:211, 1987.

Hollinger & Chamberlain, Current Opinion in Neurology 28, 522-527, 2015.

Horlick and Benfield, Mol. Cell. Biol., 9:2396, 1989.

Horwich et al., J. Virol., 64:642-650, 1990.

Hsu et al., Natl Biotechnol. 31:827-832, 2013

Huang et al., Cell, 27:245, 1981.

Hug et al., Mol. Cell. Biol., 8:3065, 1988.

Hwang et al., Mol. Cell. Biol., 10:585, 1990.

Imagawa et al., Cell, 51:251, 1987.

Imbra and Karin, Nature, 323:555, 1986.

Imler et al., Mol. Cell. Biol., 7:2558, 1987.

Imperiale and Nevins, Mol. Cell. Biol., 4:875, 1984.

Jakobovits et al., Mol. Cell. Biol., 8:2555, 1988.

Jameel and Siddiqui, Mol. Cell. Biol., 6:710, 1986.

Jaynes et al., Mol. Cell. Biol., 8:62, 1988.

Jinek et al., Science 337, 816-821, 2012.

Johnson et al., Mol. Cell. Biol., 9:3393, 1989.

Jones and Shenk, Cell, 13:181 188, 1978.

Kadesch and Berg, Mol. Cell. Biol., 6:2593, 1986.

Kaneda et al., Science, 243:375-378, 1989.

Karin et al., Mol. Cell. Biol., 7:606, 1987.

Karlsson et al., EMBO J., 5:2377 2385, 1986.

Katinka et al., Cell, 20:393, 1980.

Kato et al., J Biol Chem., 266(6):3361-3364, 1991.

Kawamoto et al., Mol. Cell. Biol., 8:267, 1988.

Kelly et al., J. Cell Biol., 129(2):383-96, 1995.

Kiledjian et al., Mol. Cell. Biol., 8:145, 1988.

Kim et al., Nature Biotechnology, 1-2, 2016.

Kim et al., Nature Biotechnology 34, 876-881, 2016.

Kimura et al., Dev. Growth Differ., 39(3):257-65, 1997.

Klamut et al., Mol. Cell. Biol., 10:193, 1990.

Klein et al., Nature, 327:70-73, 1987.

Koch et al., Mol. Cell. Biol., 9:303, 1989.

Kriegler and Botchan, In: Eukaryotic Viral Vectors, Gluzman (Ed.), Cold Spring Harbor: Cold Spring Harbor Laboratory, NY, 1982.

Kriegler and Botchan, Mol. Cell. Biol., 3:325, 1983.

Kriegler et al., Cell, 38:483, 1984.

Kriegler et al., Cell, 53:45, 1988.
Kuhl et al., Cell, 50:1057, 1987.
Kunz et al., Nucl. Acids Res., 17:1121, 1989.
LaPointe et al., Hypertension, 27(3):715-22, 1996
LaPointe et al., J. Biol. Chem., 263(19):9075-8, 1988.
Larsen et al., Proc. Natl. Acad. Sci. USA., 83:8283, 1986.
Laspia et al., Cell, 59:283, 1989.
Latimer et al., Mol. Cell. Biol., 10:760, 1990.
Le Gal La Salle et al., Science, 259:988 990, 1993.
Lee et al., Nature, 294:228, 1981.
Levinson et al., Nature, 295:79, 1982.
Levrero et al., Gene, 101:195 202, 1991.
Lin et al., Mol. Cell. Biol., 10:850, 1990.
Long et al., Science 345: 1184-1188, 2014.
Long et al., Science 351, 400-403, 2016.
Pinello et al., Nature Biotechnol. 34, 695-697, 2016.
Long et al., JAMA Neurol., 3388, 2016.
Long et al., Science 351, 400-403, 2016.
Luria et al., EMBO J., 6:3307, 1987.
Lusky and Botchan, Proc. Natl. Acad. Sci. USA, 83:3609, 1986.
Lusky et al., Mol. Cell. Biol., 3:1108, 1983.
Majors and Varmus, Proc. Natl. Acad. Sci. USA, 80:5866, 1983.
Mali et al., Science 339, 823-826, 2013a.
Mali et al., Nat Methods 10, 957-963, 2013b.
Mali et al., Nat. Biotechnol. 31:833-838, 2013c.
Mann et al., Cell, 33:153-159, 1983.
Maresca et al., Genome Research 23, 539-546, 2013.
Markowitz et al., J. Virol., 62:1120 1124, 1988.
McNeall et al., Gene, 76:81, 1989.
Miksicek et al., Cell, 46:203, 1986.
Millay et al., Nat. Med. 14, 442-447, 2008.
Mojica et al., J. Mol. Evol. 60, 174-182, 2005.
Mordacq and Linzer, Genes and Dev., 3:760, 1989.
Moreau et al., Nucl. Acids Res., 9:6047, 1981.

Moss et al., J. Gen. Physiol., 108(6):473-84, 1996.

Muesing et al., Cell, 48:691, 1987.

Nelson et al., Science 351, 403-407, 2016.

Nicolas and Rubinstein, In: Vectors: A survey of molecular cloning vectors and their uses, Rodriguez and Denhardt (Eds.), Stoneham: Butterworth, 494-513, 1988.

Nicolau and Sene, Biochim. Biophys. Acta, 721:185-190, 1982.

Nicolau et al., Methods Enzymol., 149:157-176, 1987.

Ondek et al., EMBO J., 6:1017, 1987.

Ornitz et al., Mol. Cell. Biol., 7:3466, 1987.

Padgett, R.A., Trends Genet. 28, 147-154, 2012.

Palmiter et al., Cell, 29:701, 1982a.

Palmiter et al., Nature, 300:611, 1982b.

Paskind et al., Virology, 67:242-248, 1975.

Pech et al., Mol. Cell. Biol., 9:396, 1989.

Perales et al., Proc. Natl. Acad. Sci. USA, 91(9):4086-4090, 1994.

Perez-Stable and Constantini, Mol. Cell. Biol., 10:1116, 1990.

Picard et al., Nature, 307:83, 1984.

Pinkert et al., Genes and Dev., 1:268, 1987.

Ponta et al., Proc. Natl. Acad. Sci. USA, 82:1020, 1985.

Porton et al., Mol. Cell. Biol., 10:1076, 1990.

Potter et al., Proc. Natl. Acad. Sci. USA, 81:7161-7165, 1984.

Queen and Baltimore, Cell, 35:741, 1983.

Quinn et al., Mol. Cell. Biol., 9:4713, 1989.

Racher et al., Biotech. Techniques, 9:169 174, 1995.

Ragot et al., Nature, 361:647 650, 1993.

Ran et al., Nature 520, 186-191, 2015.

Redondo et al., Science, 247:1225, 1990.

Reisman and Rotter, Mol. Cell. Biol., 9:3571, 1989.

Renan, Radiother. Oncol., 19:197-218, 1990.

Resendez Jr. et al., Mol. Cell. Biol., 8:4579, 1988.

Rich et al., Hum. Gene Ther., 4:461 476, 1993.

Ridgeway, In: Vectors: A survey of molecular cloning vectors and their uses, Stoneham: Butterworth, 467 492, 1988.

Ripe et al., Mol. Cell. Biol., 9:2224, 1989.

Rippe et al., Mol. Cell Biol., 10:689-695, 1990.

Rittling et al., Nuc. Acids Res., 17:1619, 1989.

Rosen et al., Cell, 41:813, 1988.

Rosenfeld et al., Cell, 68:143-155,1992.

Rosenfeld et al., Science, 252:431-434,1991.

Roux et al., Proc. Natl. Acad. Sci. USA, 86:9079-9083, 1989.

Sambrook and Russell, Molecular Cloning: A Laboratory Manual 3rd Ed., Cold Spring Harbor Laboratory Press, 2001.

Sakai et al., Genes and Dev., 2:1144, 1988.

Satake et al., J. Virology, 62:970, 1988.

Schaffner et al., J. Mol. Biol., 201:81, 1988.

Searle et al., Mol. Cell. Biol., 5:1480, 1985.

Sharp et al., Cell, 59:229, 1989.

Shaul and Ben-Levy, EMBO J., 6:1913, 1987.

Sherman et al., Mol. Cell. Biol., 9:50, 1989.

Sleigh et al., J. EMBO, 4:3831, 1985.

Shimizu-Motohashi et al., Am J Transl Res 8, 2471-89, 2016.

Spalholz et al., Cell, 42:183, 1985.

Spandau and Lee, J. Virology, 62:427, 1988.

Spandidos and Wilkie, EMBO J., 2:1193, 1983.

Stephens and Hentschel, Biochem. J., 248:1, 1987.

Stratford Perricaudet and Perricaudet, In: Human Gene Transfer, Cohen Haguenauer and Boiron (Eds.), John Libbey Eurotext, France, 51 61, 1991.

Stratford-Perricaudet et al., Hum. Gene. Ther., 1:241-256, 1990.

Stuart et al., Nature, 317:828, 1985.

Sullivan and Peterlin, Mol. Cell. Biol., 7:3315, 1987.

Swartzendruber and Lehman, J. Cell. Physiology, 85:179, 1975.

Tabebordbar et al., Science 351, 407-411, 2016.

Takebe et al., Mol. Cell. Biol., 8:466, 1988.

Tavernier et al., Nature, 301:634, 1983.

Taylor and Kingston, Mol. Cell. Biol., 10:165, 1990a.

Taylor and Kingston, Mol. Cell. Biol., 10:176, 1990b.

Taylor et al., J. Biol. Chem., 264:15160, 1989.

Temin, In: Gene Transfer, Kucherlapati (Ed.), NY, Plenum Press, 149-188, 1986.

Thiesen et al., J. Virology, 62:614, 1988.

Top et al., J. Infect. Dis., 124:155-160, 1971.

Toth et al., Biology Direct, 1-14, 2016.

Tronche et al., Mol. Biol. Med., 7:173, 1990.

Trudel and Constantini, Genes and Dev. 6:954, 1987.

Tsai et al., Nature Biotechnology 34, 882-887, 2016.

Tur-Kaspa et al., Mol. Cell Biol., 6:716-718, 1986.

Tyndell et al., Nuc. Acids. Res., 9:6231, 1981.

Vannice and Levinson, J. Virology, 62:1305, 1988.

Varmus et al., Cell, 25:23-36, 1981.

Vasseur et al., Proc Natl. Acad. Sci. U.S.A., 77:1068, 1980.

Wagner et al., Proc. Natl. Acad. Sci. USA, 87(9):3410-3414, 1990.

Wang and Calame, Cell, 47:241, 1986.

Wang et al., Cell, 153:910-910, 2013.

Weber et al., Cell, 36:983, 1984.

Weinberger et al. Mol. Cell. Biol., 8:988, 1984.

Winoto et al., Cell, 59:649, 1989.

Wong et al., Gene, 10:87-94, 1980.

Wu and Wu, Adv. Drug Delivery Rev., 12:159 167, 1993.

Wu and Wu, Biochemistry, 27:887-892, 1988.

Wu and Wu, J. Biol. Chem., 262:4429-4432, 1987.

Wu et al., Cell Stem Cell 13, 659-662, 2013.

Wu et al., Nat Biotechnol 32, 670-676, 2014.

Xu et al., Mol Ther 24, 564-569, 2016.

Yamauchi-Takihara et al., Proc. Natl. Acad. Sci. USA, 86(10):3504-3508, 1989.

Yang et al., Proc. Natl. Acad. Sci. USA, 87:9568-9572, 1990.

Yin et al., Nat Biotechnol 32, 551-553, 2014.

Yin et al., Physiol Rev 93, 23-67, 2013.

Young et al., Cell Stem Cell 18, 533-540, 2016.

Yutzey et al. Mol. Cell. Biol., 9:1397, 1989.

Zechner et al., Cell Metabolism 12, 633-642, 2010.

Zelenin et al., FEBS Lett., 280:94-96, 1991.

Zetsche et al., Cell 163, 759-771, 2015.

Ziober and Kramer, J. Bio. Chem., 271(37):22915-22922, 1996.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Нуклеиновая кислота, включающая:

последовательность, кодирующую первую DMD направляющую РНК, направленно взаимодействующую с первой геномной последовательностью-мишенью,

последовательность, кодирующую вторую DMD направляющую РНК, направленно взаимодействующую со второй геномной последовательностью-мишенью,

последовательность, кодирующую первый промотор, где первый промотор направляет экспрессию последовательности, кодирующей первую DMD направляющую РНК, и

последовательность, кодирующую второй промотор, где второй промотор направляет экспрессию последовательности, кодирующей вторую DMD направляющую РНК,

где каждая первая геномная последовательность-мишень и вторая геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина.

2. Нуклеиновая кислота по п.1, где последовательность, кодирующая первый промотор, и последовательность, кодирующая второй промотор, являются идентичными.

3. Нуклеиновая кислота по п.1, где последовательность, кодирующая первый промотор, и последовательность, кодирующая второй промотор, не являются идентичными.

4. Нуклеиновая кислота по любому из пп.1-3, где первая геномная последовательность-мишень и вторая геномная последовательность-мишень являются идентичными.

5. Нуклеиновая кислота по любому из пп.1-3, где первая геномная последовательность-мишень и вторая геномная последовательность-мишень не являются идентичными.

6. Нуклеиновая кислота по любому из пп.1-5, где нуклеиновая кислота дополнительно включает:

последовательность, кодирующую третью DMD направляющую РНК, направленно взаимодействующую с третьей геномной последовательностью-мишенью, и

последовательность, кодирующую третий промотор, где третий промотор направляет экспрессию последовательности, кодирующей

третью DMD направляющую РНК,

где третья геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина.

7. Нуклеиновая кислота по п.6, где по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, и последовательности, кодирующей третий промотор, являются идентичными.

8. Нуклеиновая кислота по п.6, где по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, и последовательности, кодирующей третий промотор, не являются идентичными.

9. Нуклеиновая кислота по любому из пп.6-8, где по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени и третьей геномной последовательности-мишени являются идентичными.

10. Нуклеиновая кислота по любому из пп.6-8, где по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени и третьей геномной последовательности-мишени не являются идентичными.

11. Нуклеиновая кислота по любому из пп.6-10, где нуклеиновая кислота дополнительно включает:

последовательность, кодирующую четвертую DMD направляющую РНК, направленно взаимодействующую с четвертой геномной последовательностью-мишенью, и

последовательность, кодирующую четвертый промотор, где четвертый промотор направляет экспрессию четвертой последовательности, кодирующей DMD направляющую РНК,

где четвертая геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина.

12. Нуклеиновая кислота по п.11, где по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, и последовательности, кодирующей четвертый промотор, являются

идентичными.

13. Нуклеиновая кислота по п.11, где по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, и последовательности, кодирующей четвертый промотор, не являются идентичными.

14. Нуклеиновая кислота по любому из пп.11-13, где по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени и четвертой геномной последовательности-мишени являются идентичными.

15. Нуклеиновая кислота по любому из пп.11-13, где по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени и четвертой геномной последовательности-мишени не являются идентичными.

16. Нуклеиновая кислота по любому из пп.11-15, где нуклеиновая кислота дополнительно включает:

последовательность, кодирующую пятую DMD направляющую РНК, направленно взаимодействующую с пятой геномной последовательностью-мишенью, и

последовательность, кодирующую пятый промотор, где пятый промотор направляет экспрессию последовательности, кодирующей пятую DMD направляющую РНК,

где пятая геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина.

17. Нуклеиновая кислота по п.16, где по меньшей мере две из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, являются идентичными.

18. Нуклеиновая кислота по п.16, где по меньшей мере две из последовательности, кодирующей первый промотор,

последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, не являются идентичными.

19. Нуклеиновая кислота по любому из пп.16–18, где по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени, четвертой геномной последовательности-мишени и пятой геномной последовательности-мишени являются идентичными.

20. Нуклеиновая кислота по любому из пп.16–18, где по меньшей мере две из первой геномной последовательности-мишени, второй геномной последовательности-мишени, третьей геномной последовательности-мишени, четвертой геномной последовательности-мишени и пятой геномной последовательности-мишени не являются идентичными.

21. Нуклеиновая кислота по любому из пп.16–20, где нуклеиновая кислота дополнительно включает:

по меньшей мере одну последовательность, кодирующую дополнительную DMD направляющую РНК, направленно взаимодействующую с геномной последовательностью-мишенью, и

по меньшей мере один дополнительный промотор, где дополнительный промотор направляет экспрессию последовательности, кодирующей дополнительную DMD направляющую РНК,

где дополнительная геномная последовательность-мишень включает акцепторный сайт сплайсинга дистрофина.

22. Нуклеиновая кислота по любому из пп.1–21, где акцепторный сайт сплайсинга дистрофина является 5' акцепторным сайтом сплайсинга экзона 51.

23. Нуклеиновая кислота по любому из пп.1–23, где последовательность, кодирующая первый промотор, или последовательность, кодирующая второй промотор, включает последовательность, кодирующую конститутивный промотор.

24. Нуклеиновая кислота по любому из пп.1–23, где первый

промотор или второй промотор включают конститутивный промотор.

25. Нуклеиновая кислота по п.20, где по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает последовательность, кодирующую конститутивный промотор.

26. Нуклеиновая кислота по п.20, где по меньшей мере один из первого промотора, второго промотора, третьего промотора, четвертого промотора и пятого промотора включает конститутивный промотор.

27. Нуклеиновая кислота по любому из пп.1-23, где последовательность, кодирующая первый промотор, или последовательность, кодирующая второй промотор, включает последовательность, кодирующую индуцируемый промотор.

28. Нуклеиновая кислота по п.20, где по меньшей мере один из первого промотора, второго промотора, третьего промотора, четвертого промотора и пятого промотора включает индуцируемый промотор.

29. Нуклеиновая кислота по любому из пп.1-23, где последовательность, кодирующая первый промотор или последовательность, кодирующая второй промотор, включает последовательность, кодирующую специфический для типа клеток промотор.

30. Нуклеиновая кислота по п.20, где по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает специфический для типа клеток промотор.

31. Нуклеиновая кислота по п.29 или 30, где последовательность, кодирующая специфический для типа клеток промотор, включает последовательность, кодирующую мышечно-специфический промотор.

32. Нуклеиновая кислота по любому из пп.1–22, где последовательность, кодирующая первый промотор, или последовательность, кодирующая второй промотор, включает последовательность, кодирующую промотор U6, промотор H1 или промотор 7SK.

33. Нуклеиновая кислота по п.20, где по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает последовательность, кодирующую промотор U6, последовательность, кодирующую промотор H1, или последовательность, кодирующую промотор 7SK.

34. Нуклеиновая кислота по п.20, где по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает последовательность, кодирующую промотор U6.

35. Нуклеиновая кислота по п.20, где по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает последовательность, кодирующую промотор H1.

36. Нуклеиновая кислота по п.20, где по меньшей мере одна из последовательности, кодирующей первый промотор, последовательности, кодирующей второй промотор, последовательности, кодирующей третий промотор, последовательности, кодирующей четвертый промотор, и последовательности, кодирующей пятый промотор, включает последовательность, кодирующую промотор 7SK.

37. Нуклеиновая кислота по п.6,

где последовательность, кодирующая первую DMD направляющую РНК, последовательность, кодирующая вторую DMD направляющую РНК и последовательность, кодирующую третью DMD направляющую РНК, являются идентичными,

где последовательность, кодирующая первый промотор, последовательность, кодирующая второй промотор, и последовательность, кодирующая третий промотор, не являются идентичными, и

где 5' акцепторный сайт сплайсинга включает 5' акцепторный сайт сплайсинга экзона 51.

38. Нуклеиновая кислота по п.37, где последовательность, кодирующая первый промотор, включает последовательность, кодирующую промотор U6, последовательность, кодирующая второй промотор, включает последовательность, кодирующую промотор H1, и последовательность, кодирующая третий промотор, включает последовательность, кодирующую промотор 7SK.

39. Нуклеиновая кислота по любому из пп.1–38, где нуклеиновая кислота включает последовательность ДНК.

40. Нуклеиновая кислота по любому из пп.1–38, где нуклеиновая кислота включает последовательность РНК.

41. Нуклеиновая кислота по любому из пп.1–40, где нуклеиновая кислота дополнительно включает одну или более последовательностей, кодирующих инвертированный концевой повтор (ITR).

42. Нуклеиновая кислота по любому из пп.1–40, где нуклеиновая кислота дополнительно включает последовательность, кодирующую 5' инвертированный концевой повтор (ITR) и последовательность, кодирующую 3' ITR.

43. Нуклеиновая кислота по п.42, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR) или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV).

44. Нуклеиновая кислота по п.43, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает

последовательность, выделенную или полученную из аденоассоциированного вируса (AAV) серотипа 2 (AAV2).

45. Нуклеиновая кислота по п.43, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR) и последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из AAV2.

46. Нуклеиновая кислота по п.43, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV) серотипа 4 (AAV4).

47. Нуклеиновая кислота по п.46, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), и последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из AAV4.

48. Нуклеиновая кислота по любому из пп.43-47, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 145 нуклеотидов.

49. Нуклеиновая кислота по любому из пп.43-47, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 115 нуклеотидов.

50. Нуклеиновая кислота по любому из пп.43-47, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 141 нуклеотида

51. Нуклеиновая кислота по любому из пп.1-50, где нуклеиновая кислота дополнительно включает полиаденозиновую (поли(A)) последовательность.

52. Нуклеиновая кислота по п.51, где поли(A) последовательность является мини-поли(A) последовательностью.

53. Нуклеиновая кислота по любому из пп.6-52, где последовательность, кодирующая первую DMD направляющую РНК, последовательность, кодирующая вторую DMD направляющую РНК, или последовательность, кодирующая третью DMD направляющую РНК,

включает последовательность любого из SEQ ID NO: 60-382, 706-708 и 712-719.

54. Нуклеиновая кислота по любому из пп.6-52, где последовательность, кодирующая первую DMD направляющую РНК, последовательность, кодирующая вторую DMD направляющую РНК, или последовательность, кодирующая третью DMD направляющую РНК, включает последовательность SEQ ID NO: 714.

55. Нуклеиновая кислота по любому из пп.6-52, где последовательность, кодирующая первую DMD направляющую РНК, последовательность, кодирующая вторую DMD направляющую РНК, и последовательность, кодирующая третью DMD направляющую РНК, включает последовательность SEQ ID NO: 714.

56. Клетка, включающая нуклеиновую кислоту по любому из пп.1-55.

57. Композиция, включающая нуклеиновую кислоту по любому из пп.1-55.

58. Клетка, включающая композицию по п.56.

59. Композиция, включающая клетку по п.58.

60. Вектор, включающий нуклеиновую кислоту по любому из пп.1-55.

61. Вектор по п.60, где вектор дополнительно включает последовательность, кодирующую инвертированный концевой повтор (ITR) транспозируемого элемента.

62. Вектор по п.61, где транспозируемый элемент является транспозоном.

63. Вектор по п.62, где транспозон является транспозоном Tn7.

64. Вектор по п.63, где вектор дополнительно включает последовательность, кодирующую 5' ITR транспозона T7, и последовательность, кодирующую 3' ITR транспозона T7.

65. Вектор по любому из пп.60-64, где вектор является невирусным вектором.

66. Вектор по п.65, где невирусный вектор является плазмидой.

67. Вектор по любому из пп.60-64, где вектор является вирусным вектором.

68. Вектор по п.67, где вирусный вектор является аденоассоциированным вирусным (AAV) вектором.

69. Вектор по п.69, где AAV вектор является репликационно-дефектным или условно репликационно-дефектным.

70. Вектор по пп.68 или 69, где AAV вектор является рекомбинантным AAV вектором.

71. Вектор по любому из пп.68-70, где AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 1 (AAV1), 2 (AAV2), 3 (AAV3), 4 (AAV4), 5 (AAV5), 6 (AAV6), 7 (AAV7), 8 (AAV8), 9 (AAV9), 10 (AAV10), 11 (AAV11) или их любой комбинации.

72. Вектор по любому из пп.68-71, где AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 9 (AAV9).

73. Вектор по любому из пп.68-72, где AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 2 (AAV2).

74. Вектор по любому из пп.68-73, где AAV вектор включает последовательность, выделенную или полученную из AAV2, и последовательность, выделенную или полученную из AAV9.

75. Вектор по любому из пп.68-72, где AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 4 (AAV4).

76. Вектор по любому из пп.68-72 или 75, где AAV вектор включает последовательность, выделенную или полученную из AAV4, и последовательность, выделенную или полученную из AAV9.

77. Вектор по любому из пп.60-76, где вектор оптимизирован для экспрессии в клетках млекопитающих.

78. Вектор по любому из пп.60-77, где вектор оптимизирован для экспрессии в клетках человека.

79. Вектор по любому из пп.60-78, где вектор включает последовательность нуклеиновой кислоты SEQ ID NO: 914, SEQ ID NO: 915, SEQ ID NO: 916 или SEQ ID NO: 917.

80. Композиция, включающая вектор по любому из пп.60-79.

81. Композиция по п.80, дополнительно включающая фармацевтически приемлемый носитель.

82. Клетка, включающая композицию по п.80 или 81.

83. Клетка по п.82, где клетка является клеткой человека.

84. Клетка по п.82 или 83, где клетка является мышечной клеткой или сателлитной клеткой.

85. Клетка по п.82 или 83, где клетка является индуцированной плюрипотентной стволовой (иПС) клеткой.

86. Композиция, включающая клетку по любому из пп.82-85.

87. Нуклеиновая кислота, включающая последовательность, кодирующую промотор, и последовательность, кодирующую белок Cas9 или его нуклеазный домен,

где последовательность, кодирующая промотор, включает последовательность, кодирующую мышечно-специфический промотор.

88. Нуклеиновая кислота по п.87, где последовательность, кодирующая мышечно-специфический промотор, включает последовательность, кодирующую промотор CK8.

89. Нуклеиновая кислота по п.87 или 88, где последовательность, кодирующая мышечно-специфический промотор, включает последовательность, кодирующую промотор CK8e.

90. Нуклеиновая кислота по любому из пп.87-89, где последовательность, кодирующая белок Cas9 или его нуклеазный домен, выделена или получена из последовательности, кодирующей белок Cas9 *S. pyogenes* или его нуклеазный домен.

91. Белок Cas9 или его нуклеазный домен, выделенный или полученный из последовательности, кодирующей белок Cas9 *S. aureus* или его нуклеазный домен.

92. Нуклеиновая кислота по любому из пп.87-91, где последовательность, кодирующая белок Cas9 или его нуклеазный домен, является кодон-оптимизированной для экспрессии в млекопитающем.

93. Нуклеиновая кислота по любому из пп.87-91, где последовательность, кодирующая белок Cas9 или его нуклеазный домен, является кодон-оптимизированной для экспрессии у человека.

94. Нуклеиновая кислота по любому из пп.87-93, где нуклеиновая кислота дополнительно включает поли(A) последовательность.

95. Нуклеиновая кислота по п.94, где поли(А) последовательность является мини-поли(А) последовательностью.

96. Нуклеиновая кислота по любому из пп.1-94, где нуклеиновая кислота дополнительно включает одну или более последовательностей, кодирующих инвертированный концевой повтор (ITR).

97. Нуклеиновая кислота по любому из пп.1-94, где нуклеиновая кислота дополнительно включает последовательность, кодирующую 5' инвертированный концевой повтор (ITR), и последовательность, кодирующую 3' ITR.

98. Нуклеиновая кислота по п.97, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV).

99. Нуклеиновая кислота по п.98, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV) серотипа 2 (AAV2).

100. Нуклеиновая кислота по п.99, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), и последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из AAV2.

101. Нуклеиновая кислота по п.98, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из аденоассоциированного вируса (AAV) серотипа 4 (AAV4).

102. Нуклеиновая кислота по п.46, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), и последовательность, кодирующая 3' ITR, включает последовательность, выделенную или полученную из AAV4.

103. Нуклеиновая кислота по любому из пп.96-102, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR,

включает или состоит из 145 нуклеотидов.

104. Нуклеиновая кислота по любому из пп.96-102, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 115 нуклеотидов.

105. Нуклеиновая кислота по любому из пп.96-102, где последовательность, кодирующая 5' инвертированный концевой повтор (ITR), или последовательность, кодирующая 3' ITR, включает или состоит из 141 нуклеотида.

106. Нуклеиновая кислота по любому из пп.1-55 или 87-105, где нуклеиновая кислота дополнительно включает сигнал ядерной локализации.

107. Нуклеиновая кислота по любому из пп.87-106, где нуклеиновая кислота оптимизирована для экспрессии в клетках млекопитающих.

108. Нуклеиновая кислота по любому из пп.87-107, где нуклеиновая кислота оптимизирована для экспрессии в клетках человека.

109. Композиция, включающая нуклеиновую кислоту по любому из пп.87-108.

110. Клетка, включающая нуклеиновую кислоту по любому из пп.87-108.

111. Клетка, включающая композицию по п.109.

112. Композиция, включающая клетку по п.110 или 111.

113. Вектор, включающий нуклеиновую кислоту по любому из пп.87-108.

114. Вектор по п.113, где вектор дополнительно включает последовательность, кодирующую инвертированный концевой повтор (ITR) транспозируемого элемента.

115. Вектор по п.114, где транспозируемый элемент является транспозоном.

116. Вектор по п.115, где транспозон является транспозоном Tn7.

117. Вектор по п.116, где вектор дополнительно включает последовательность, кодирующую 5' ITR транспозона T7, и последовательность, кодирующую 3' ITR транспозона T7.

118. Вектор по любому из пп.113-117, где вектор является невирусным вектором.

119. Вектор по п.118, где невирусный вектор является плазмидой.

120. Вектор по любому из пп.113-117, где вектор является вирусным вектором.

121. Вектор по п.120, где вирусный вектор является аденоассоциированным вирусным (AAV) вектором.

122. Вектор по п.121, где AAV вектор является репликационно-дефектным или условно репликационно-дефектным.

123. Вектор по пп.121 или 122, где AAV вектор является рекомбинантным AAV вектором.

124. Вектор по любому из пп.121-123, где AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 1 (AAV1), 2 (AAV2), 3 (AAV3), 4 (AAV4), 5 (AAV5), 6 (AAV6), 7 (AAV7), 8 (AAV8), 9 (AAV9), 10 (AAV10), 11 (AAV11) или их любой комбинации.

125. Вектор по любому из пп.121-124, где AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 9 (AAV9).

126. Вектор по любому из пп.121-125, где AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 2 (AAV2).

127. Вектор по любому из пп.121-126, где AAV вектор включает последовательность, выделенную или полученную из AAV2 и последовательность, выделенную или полученную из AAV9.

128. Вектор по любому из пп.121-125, где AAV вектор включает последовательность, выделенную или полученную из AAV вектора серотипа 4 (AAV4).

129. Вектор по любому из пп.121-125 или 128, где AAV вектор включает последовательность, выделенную или полученную из AAV4, и последовательность, выделенную или полученную из AAV9.

130. Вектор по любому из пп.121-129, где вектор оптимизирован для экспрессии в клетках млекопитающих.

131. Вектор по любому из пп.121-130, где вектор оптимизирован для экспрессии в клетках человека.

132. Вектор по любому из пп.113-132, где вектор включает последовательность нуклеиновой кислоты SEQ ID NO: 899, SEQ ID NO: 900, SEQ ID NO: 901 или SEQ ID NO: 902.

133. Композиция, включающая вектор по любому из пп.113-132.

134. Композиция по п.133, дополнительно включающая фармацевтически приемлемый носитель.

135. Клетка, включающая композицию по п.133 или 134.

136. Клетка, включающая вектор по любому из пп.113-132.

137. Клетка по п.135 или 136, где клетка является клеткой человека.

138. Клетка по любому из пп.135-137, где клетка является мышечной клеткой или сателлитной клеткой.

139. Клетка по любому из пп.135-137, где клетка является иПС или вКМ клеткой.

140. Композиция, включающая клетку по любому из пп.135-139.

141. Композиция, включающая:

первую последовательность нуклеиновой кислоты, включающую последовательность нуклеиновой кислоты по любому из пп.1-55 и

вторую последовательность нуклеиновой кислоты, включающую последовательность нуклеиновой кислоты по любому из пп.87-108.

142. Композиция, включающая:

первый вектор, включающий последовательность нуклеиновой кислоты по любому из пп.1-55, и

второй вектор, включающий последовательность нуклеиновой кислоты по любому из пп.87-108.

143. Композиция, включающая:

вектор по любому из пп.60-79 и

вектор по любому из пп.113-132.

144. Композиция по любому из пп.141-143, где композиция дополнительно включает фармацевтически приемлемый носитель.

145. Способ коррекции дефекта дистрофина, включающий контакт клетки и композиции по любому из пп.141-144 при условиях, подходящих для экспрессии первой DMD направляющей РНК, второй DMD направляющей РНК и белка Cas9 или его нуклеазного домена, где по меньшей мере одна из первой DMD направляющей РНК или второй DMD направляющей РНК образует комплекс с белком Cas9

или его нуклеазным доменом, с образованием по меньшей мере одного комплекса DMD направляющей РНК-Cas9, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 разрушает сайт сплайсинга дистрофина и вызывает селективный пропуск экзона DMD.

146. Способ коррекции дефекта дистрофина, включающий контакт клетки и композиции по любому из пп.141-144 при условиях, подходящих для экспрессии первой DMD направляющей РНК, второй DMD направляющей РНК и белка Cas9 или его нуклеазного домена, где по меньшей мере одна из первой DMD направляющей РНК или второй DMD направляющей РНК образует комплекс с белком Cas9 или его нуклеазным доменом, с образованием по меньшей мере одного комплекса DMD направляющей РНК-Cas9, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 вызывает восстановление рамки считывания дистрофина.

147. Способ по п.146, где восстановление рамки считывания дистрофина вызывает вставку.

148. Способ по п.147, где вставка включает или состоит из одного нуклеотида аденозина.

149. Способ индукции селективного пропуска экзона DMD, включающий контакт клетки и композиции по любому из пп.141-144 при условиях, подходящих для экспрессии первой DMD направляющей РНК, второй DMD направляющей РНК и белка Cas9 или его нуклеазного домена, где по меньшей мере одна из первой DMD направляющей РНК или второй DMD направляющей РНК образует комплекс с белком Cas9 или его нуклеазным доменом, с образованием по меньшей мере одного комплекса DMD направляющей РНК-Cas9, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 разрушает сайт сплайсинга дистрофина и вызывает селективный пропуск экзона DMD.

150. Способ индукции события восстановления в рамке считывания дистрофина, включающий контакт клетки и композиции по любому из пп.141-144 при условиях, подходящих для экспрессии первой DMD направляющей РНК, второй DMD направляющей РНК и белка Cas9 или его нуклеазного домена, где по меньшей мере одна из первой DMD направляющей РНК или второй DMD направляющей РНК образует комплекс с белком Cas9 или его нуклеазным доменом, с

образованием по меньшей мере одного комплекса DMD направляющей РНК-Cas9, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 разрушает сайт сплайсинга дистрофина и вызывает восстановление рамки считывания дистрофина.

151. Способ по любому из пп.145-150, где по меньшей мере один комплекс DMD направляющей РНК-Cas9 разрушает сайт сплайсинга дистрофина и вызывает селективный пропуск экзона 51 гена DMD человека.

152. Клетка, полученная способом по любому из пп.145-151.

153. Способ лечения миодистрофии у нуждающегося в этом субъекта, включающий введение субъекту терапевтически эффективного количества композиции по любому из пп.141-144

154. Способ по п.153, где композицию вводят местно.

155. Способ по п.153 или 154, где композицию вводят непосредственно в мышечную ткань.

156. Способ по любому из пп.153-155, где композицию вводят путем внутримышечной инфузии или инъекции.

157. Способ по п.155 или 156, где мышечная ткань включает ткань передней большеберцовой мышцы, ткань четырехглавой мышцы, ткань камбаловидной мышцы, ткань диафрагмы или ткань сердца.

158. Способ по любому из пп.153-157, где композицию вводят путем внутрисердечной инъекции.

159. Способ по п.153, где композицию вводят системно.

160. Способ по п.159, где композицию вводят путем внутривенной инфузии или инъекции.

161. Способ по любому из пп.153-160, где после введения композиции субъект демонстрирует нормальные дистрофин-положительные мышечные волокна и мозаичные дистрофин-положительные мышечные волокна, содержащие централизованные ядра, или их комбинацию.

162. Способ по любому из пп.153-161, где после введения композиции субъект демонстрирует появление или увеличение уровня распространенности нормальных дистрофин-положительных мышечных волокон по сравнению с отсутствием или уровнем распространенности нормальных дистрофин-положительных мышечных волокон до введения композиции.

163. Способ по любому из пп.153-162, где после введения композиции субъект демонстрирует появление или увеличение уровня распространенности мозаичных дистрофин-положительных мышечных волокон, содержащих централизованные ядра, по сравнению с отсутствием или уровнем распространенности мозаичных дистрофин-положительных мышечных волокон, содержащих централизованные ядра, до введения композиции.

164. Способ по любому из пп.153-163, где после введения композиции субъект демонстрирует пониженный уровень СК сыворотки по сравнению с уровнем СК сыворотки до введения композиции.

165. Способ по любому из пп.153-164, где после введения композиции субъект демонстрирует повышенную силу хвата по сравнению с силой хвата до введения композиции.

166. Способ по любому из пп.153-165, где субъект является новорожденным, младенцем, ребенком, молодым совершеннолетним или взрослым.

167. Способ по любому из пп.153-166, где субъект имеет миодистрофию.

168. Способ по любому из пп.153-167, где субъект является генетическим носителем миодистрофии.

169. Способ по любому из пп.153-168, где субъект является субъектом мужского пола.

170. Способ по любому из пп.153-168, где субъект является субъектом женского пола.

171. Способ по любому из пп.153-170, где субъект представляется бессимптомным, и где генетическая диагностика обнаруживает в одной или обеих копиях гена *DMD* мутацию, которая нарушает функцию продукта гена *DMD*.

172. Способ по любому из пп.153-171, где у субъекта наблюдается раннее проявление или симптом миодистрофии.

173. Способ по п.172, где раннее проявление или симптом миодистрофии включают потерю мышечной массы или слабость проксимальных мышц.

174. Способ по п.173, где потеря мышечной массы или слабость проксимальных мышц появляется в одной или обеих ногах и тазе, а затем в одной или более мышцах верхней части тела.

175. Способ по п.174, где раннее проявление или симптом миодистрофии дополнительно включают псевдогипертрофию, низкую выносливость, затруднения при стоянии, затруднения при ходьбе, затруднения при подъеме по лестнице или их комбинацию.

176. Способ по любому из пп.153-175, где у субъекта наблюдается прогрессирующее проявление или симптом миодистрофии.

177. Способ по п.176, где прогрессирующее проявление или симптом миодистрофии включают истощение мышечной ткани, замещение мышечной ткани жировой тканью или замещение мышечной ткани фиброзной тканью.

178. Способ по любому из пп.153-177, где у субъекта наблюдается позднее проявление или симптом миодистрофии.

179. Способ по п.178, где позднее проявление или симптом миодистрофии включают нарушение развития костей, искривление позвоночника, потерю двигательных функций и паралич.

180. Способ по любому из пп.153-179, где у субъекта наблюдается неврологическое проявление или симптом миодистрофии.

181. Способ по п.180, где неврологическое проявление или симптом миодистрофии включают нарушение интеллекта и паралич.

182. Способ по любому из пп.153-181, где введение композиции производят до наблюдения у субъекта одного или более прогрессирующих, поздних или неврологических проявлений или симптомов миодистрофии.

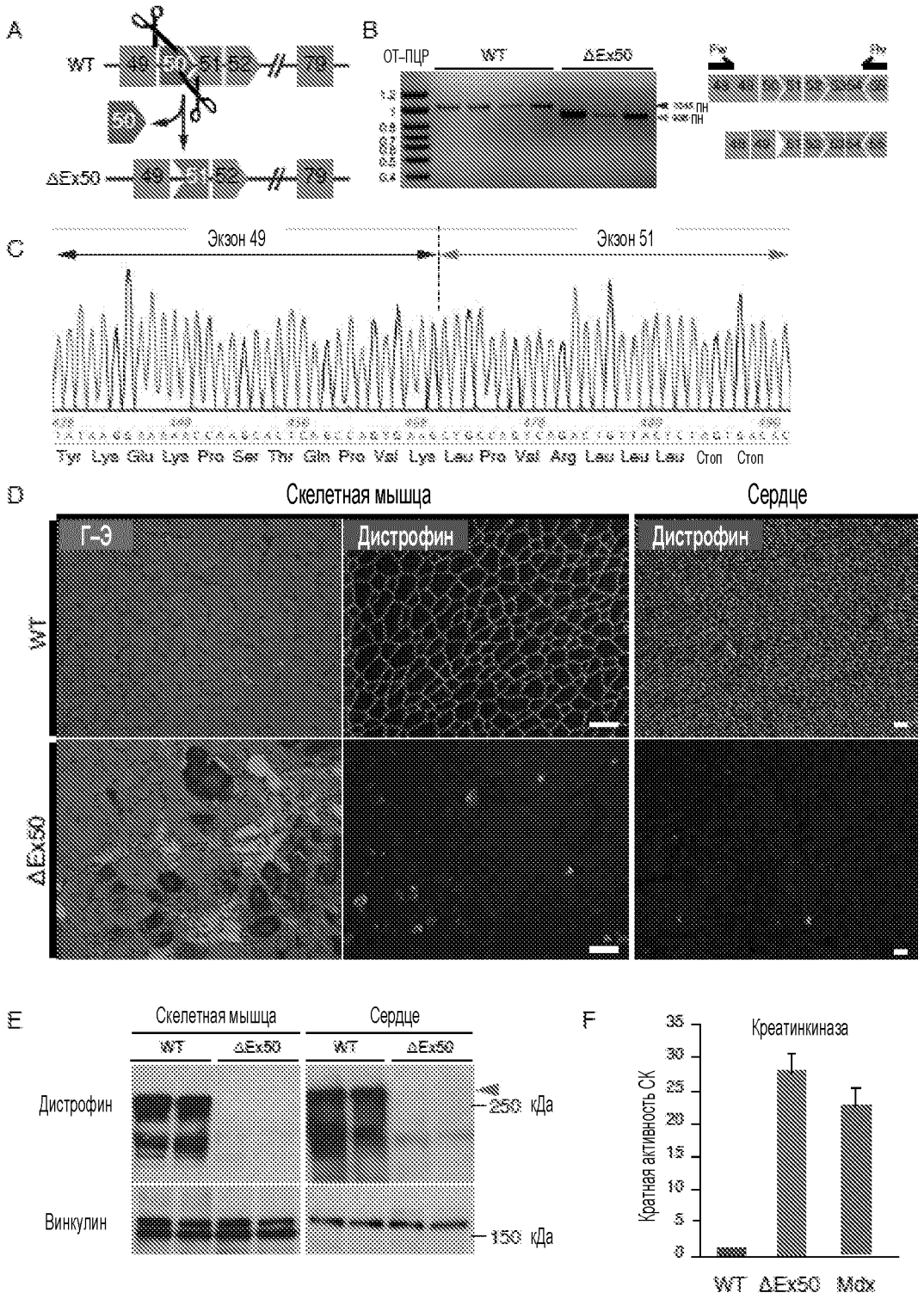
183. Способ по любому из пп.153-182, где возраст субъекта составляет меньше 10 лет.

184. Способ по п.183, где возраст субъекта составляет меньше 5 лет.

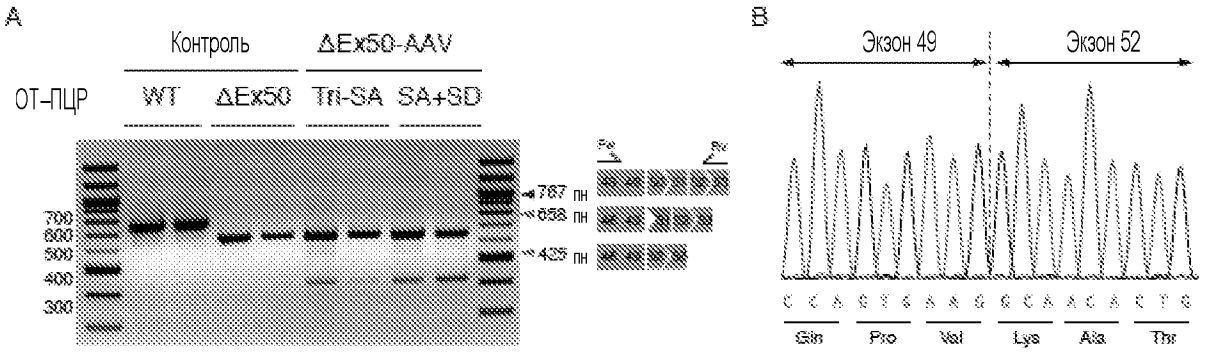
185. Способ по п.184, где возраст субъекта составляет меньше 2 лет.

186. Применение терапевтически эффективного количества композиции по любому из пп.141-144 для лечения миодистрофии у нуждающегося в этом субъекта.

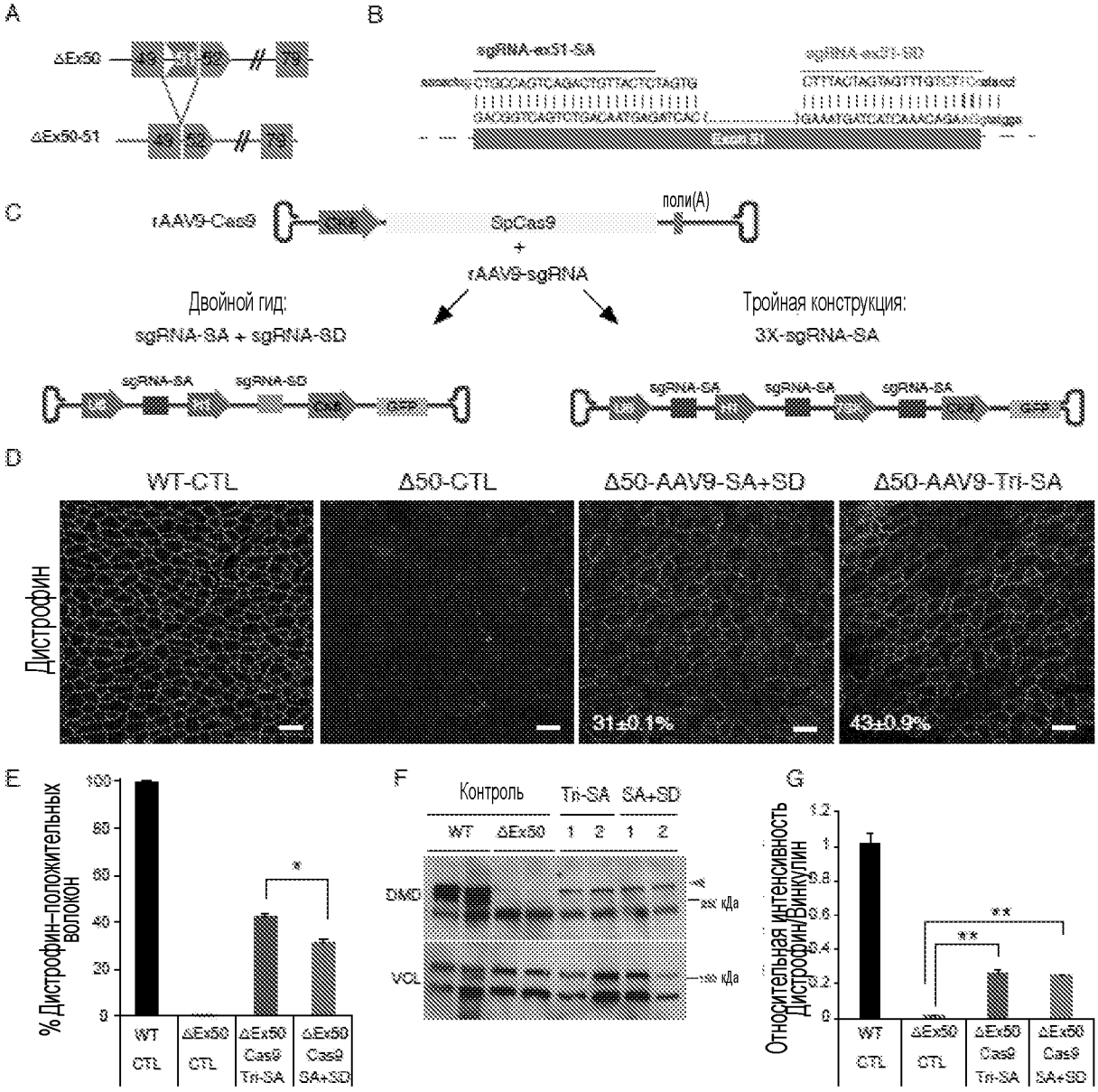
По доверенности


A
WT
49
50
51
52
79
50
ΔEx50
49
51
52
79
B
ОТ-ПЦР
WT
ΔEx50
ПН
ПН
C
Экзон 49
Экзон 51
Tyr Lys Glu Lys Pro Ser Thr Gln Pro Val Lys Leu Pro Val Arg Leu Leu Leu Стоп Стоп
D
Скелетная мышца
Сердце
Г-Э
Дистрофин
Дистрофин
WT
ΔEx50
E
Скелетная мышца
Сердце
WT
ΔEx50
WT
ΔEx50
Дистрофин
Винкулин
250 кДа
150 кДа
F
Креатинкиназа
Кратная активность CK
35
30
25
20
15
10
5
0
WT
ΔEx50
Mdx

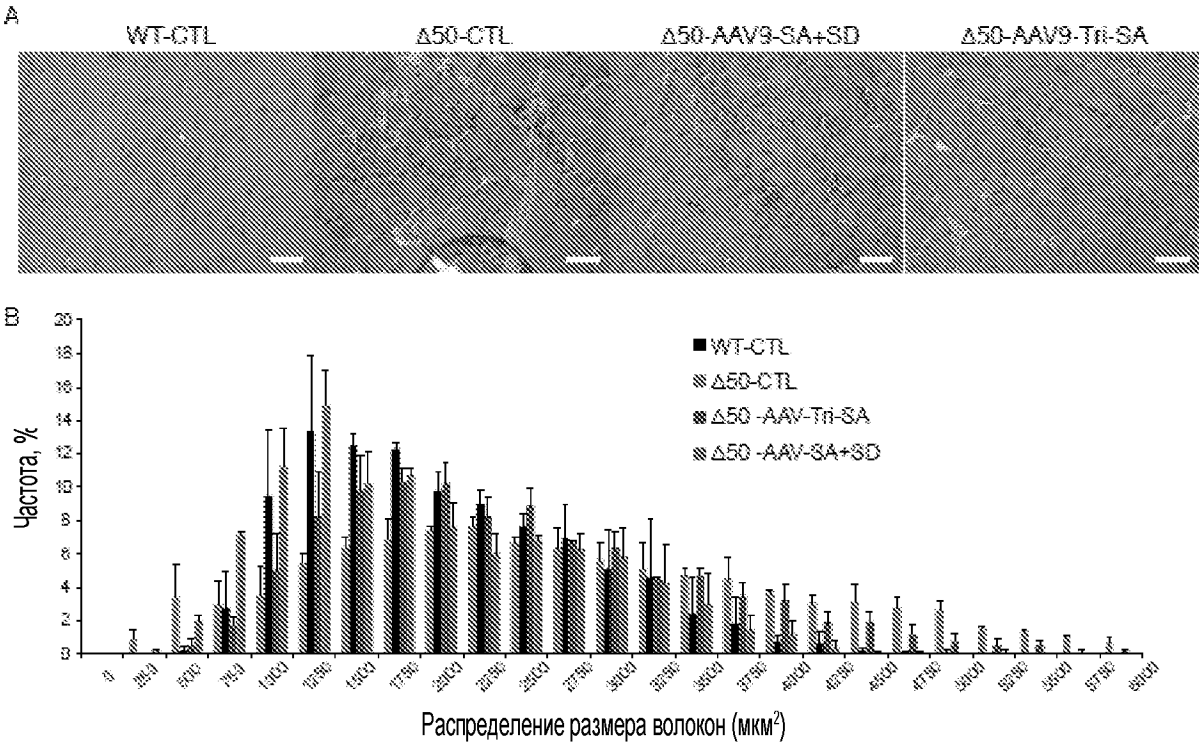
ФИГ. 1A-F


A
Контроль
ΔEx50-AAV
ОТ-ПЦР
WT
ΔEx50
Tri-SA
SA+SD
700
600
500
400
300
787 пн
658 пн
435 пн
B
Экзон 49
Экзон 52
C C A G T G A A G G C A A C A C T G
Gln
Pro
Val
Lys
Ala
Thr

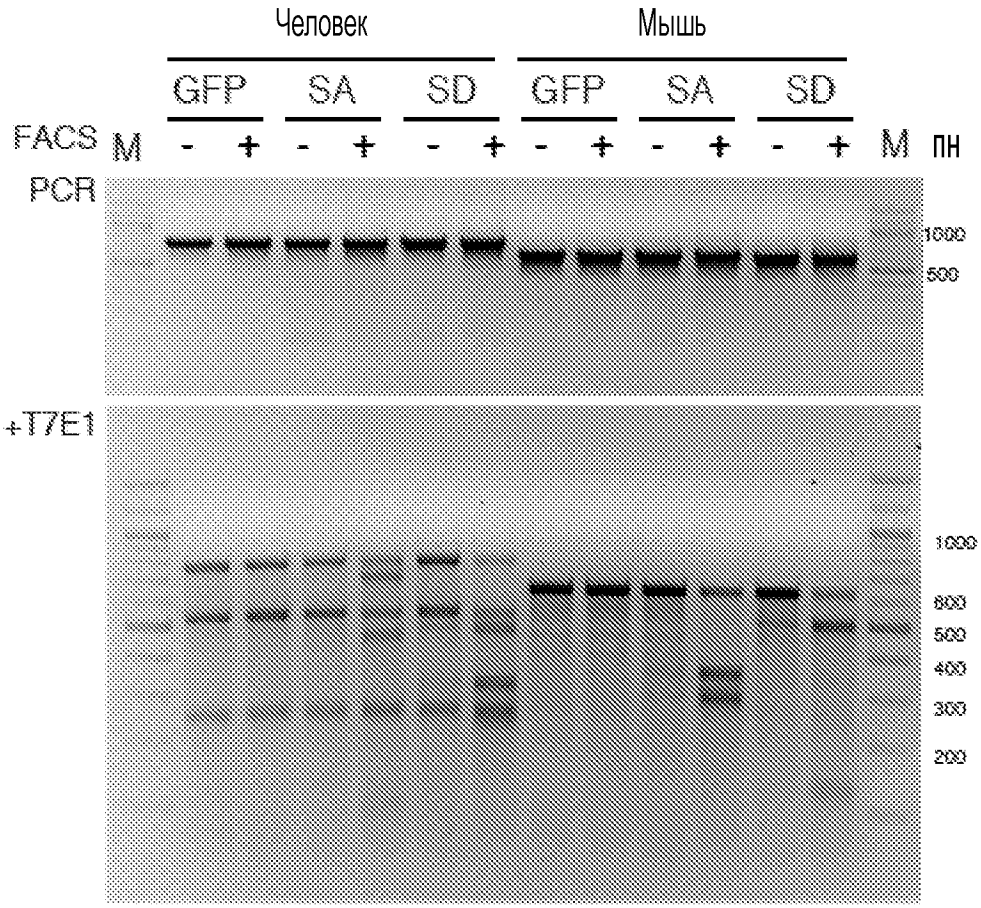
ФИГ. 2A-B


A
ΔEx50
ΔEx50-51
B
sgRNA-exS1-SA
sgRNA-exS1-SD
C
rAAV9-Cas9
SpCas9
поли(A)
+
rAAV9-sgRNA
Двойной гид:
sgRNA-SA + sgRNA-SD
Тройная конструкция:
3X-sgRNA-SA
sgRNA-SA
sgRNA-SD
D
WT-CTL
Δ50-CTL
Δ50-AAV9-SA+SD
Δ50-AAV9-Tri-SA
Дистрофин
31±0.1%
43±0.9%
E
% Дистрофин–положительных волокон
WT CTL
ΔEx50 CTL
ΔEx50 Cas9 Tri-SA
ΔEx50 Cas9 SA+SD
F
Контроль
Tri-SA
SA+SD
WT
ΔEx50
DMD
VCL
кДа
G
Относительная интенсивность Дистрофин/Винкулин

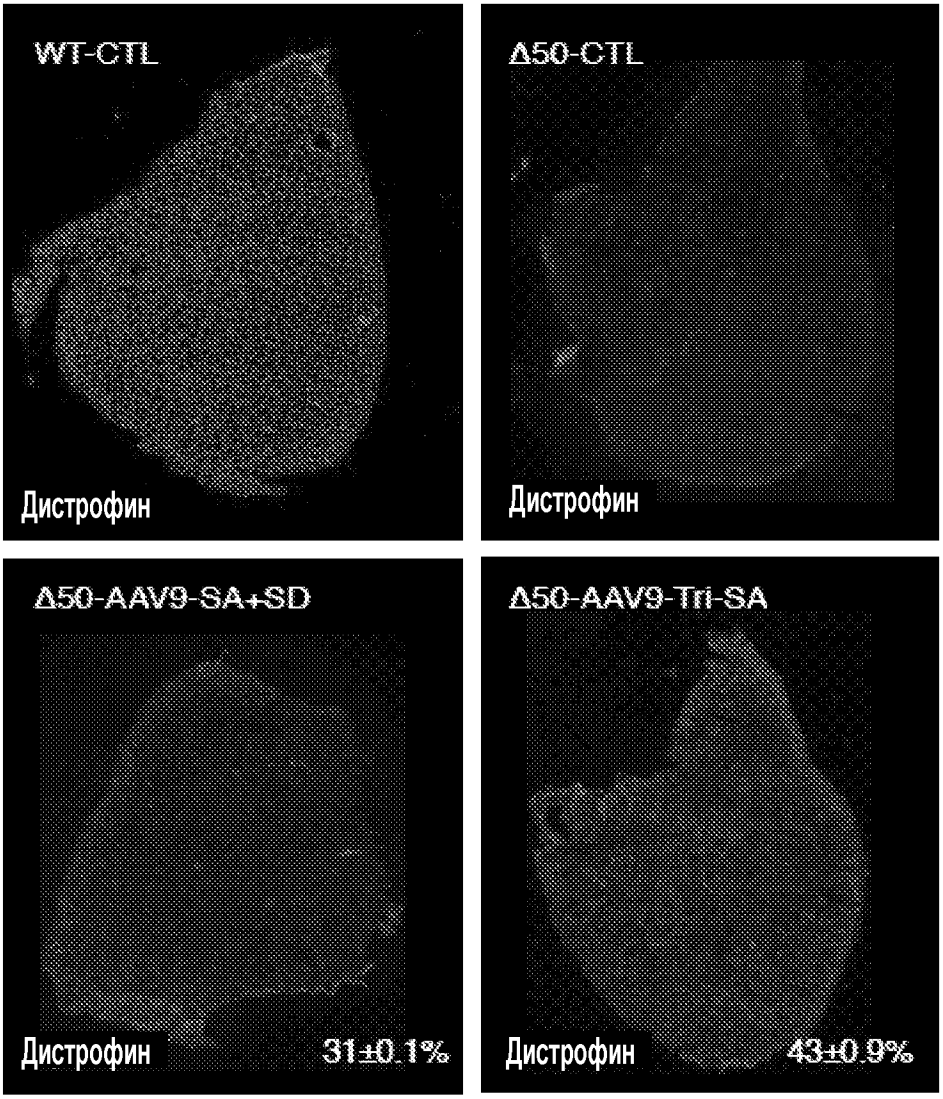
ФИГ. 3A-G


A
WT-CTL
Δ50-CTL
Δ50-AAV9-SA+SD
Δ50-AAV9-Tri-SA
B
Частота, %
WT-CTL
Δ50-CTL
Δ50 -AAV-Tri-SA
Δ50 -AAV-SA+SD
Распределение размера волокон (мкм²)

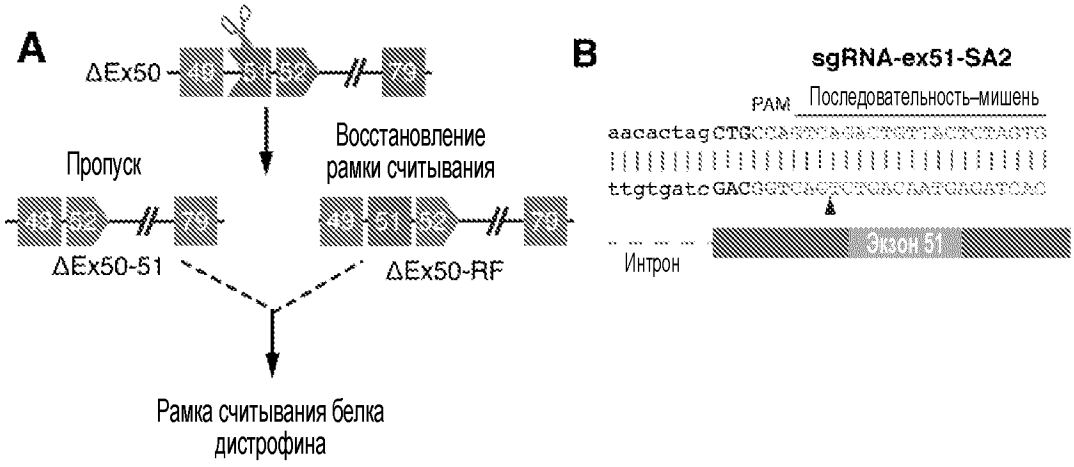
ФИГ. 4А-В


Человек
Мышь
GFP
SA
SD
GFP
SA
SD
FACS
M
- + - + - + - + - + - +
M
пн
PCR
1000
500
+T7E1
1000
800
500
400
300
200

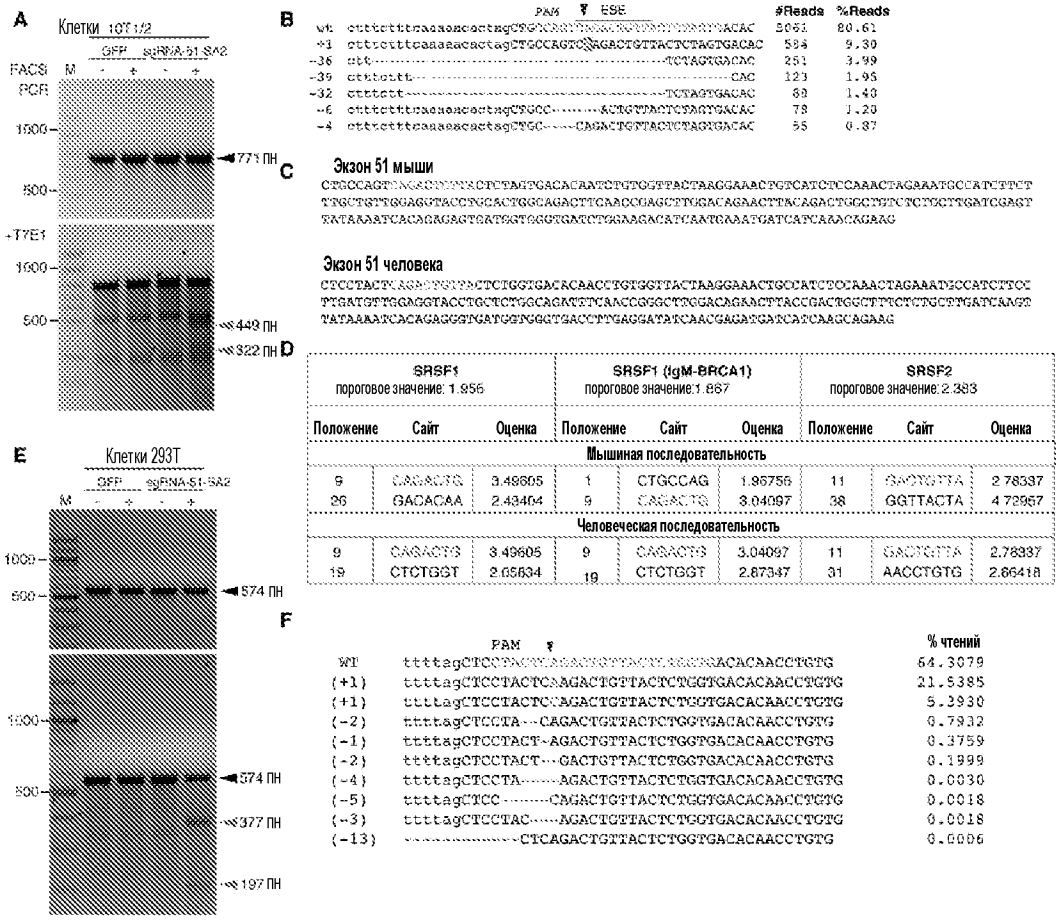
ФИГ. 5


WT-CTL
Дистрофин
Δ50-CTL
Дистрофин
Δ50-AAV9-SA+SD
Дистрофин
31±0.1%
Δ50-AAV9-Tri-SA
Дистрофин
43±0.9%

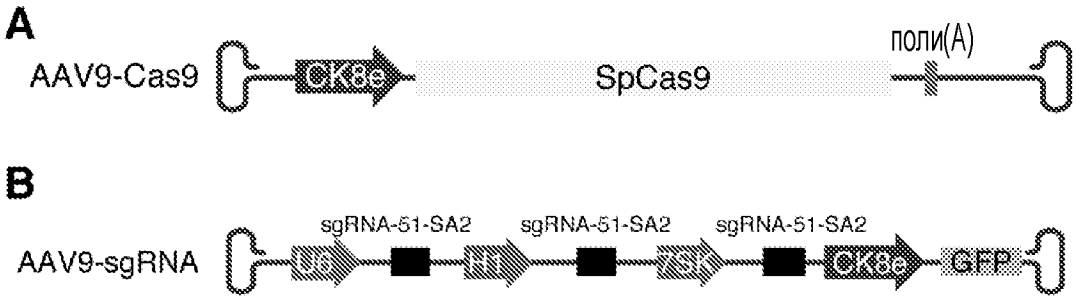
ФИГ. 6


A
ΔEx50
49
51
52
79
Пропуск
Восстановление
рамки считывания
49
52
79
ΔEx50-51
49
51
52
79
ΔEx50-RF
Рамка считывания белка
дистрофина
B
sgRNA-ex51-SA2
PAM
Последовательность–мишень
aacactag
ttgtgatc
Интрон
Экзон 51

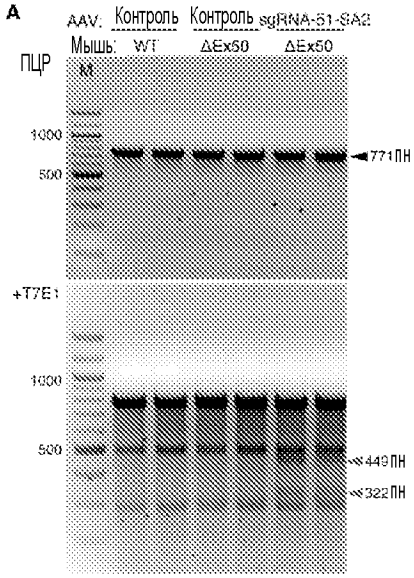
ФИГ. 7A-B


A
Клетки 107 1/2
GFP
sgRNA-51-SA2
FACS
M
PCR
1000
500
773 ПН
+T7E1
1000
500
449 ПН
322 ПН
B
PAM
ESE
#Reads
%Reads
WT
-1
-36
-39
-32
-6
-4
5061
80.61
584
9.30
251
3.99
123
1.95
88
1.40
75
1.20
55
0.87
C
Экзон 51 мыши
Экзон 51 человека
D
SRSF1
пороговое значение: 1.956
SRSF1 (IgM-BRCA1)
пороговое значение:1.867
SRSF2
пороговое значение:2.383
Положение
Сайт
Оценка
Мышиная последовательность
9 CAGACTG 3.49605
26 GACACAA 2.43404
1 CTGCCAG 1.86756
9 CAGACTG 3.04097
11 GACTGTTA 2.78337
38 GGTTACTA 4.72967
Человеческая последовательность
9 CAGACTG 3.49605
19 CTCTGGT 2.05634
9 CAGACTG 3.04097
19 CTCTGGT 2.87347
11 GACTGTTA 2.78337
31 AACCTGTG 2.66418
E
Клетки 293T
GFP
sgRNA-51-SA2
M
1000
500
574 ПН
1000
500
574 ПН
377 ПН
197 ПН
F
PAM
% чтений
WT
(+1)
(+1)
(-2)
(-1)
(-2)
(-4)
(-5)
(-3)
(-13)
64.3079
21.5385
5.3930
0.7932
0.3759
0.1999
0.0030
0.0018
0.0018
0.0006

ФИГ. 8A-F


A
AAV9-Cas9
CK8e
SpCas9
поли(A)
B
AAV9-sgRNA
sgRNA-51-SA2
sgRNA-51-SA2
sgRNA-51-SA2
U6
H1
7SK
CK8e
GFP

ФИГ. 9A-B



B **ΔEx50-AAV9-sgRNA-51-SA2-TA-#1**

	PAM ↓	% чтений	Кол-во чтений
(NE)	acactagCTGCCAGTCAAGACTGTTACTCTAGTGACA	67.17	36253
(+1)	acactagCTGCCAGTCAAGACTGTTACTCTAGTGACA	18.24	9846
(S1)	acactagCTGCCAGTCATACTGTTACTCTAGTGACA	1.17	631
(+2)	acactagCTGCCAGTCAAAGACTGTTACTCTAGTGACA	0.64	345
(S1)	acactagCTGCCAGTCAGACTTTTACTCTAGTGACA	0.57	309
(+2)	acactagCTGCCAGTCAAAGACTGTTACTCTAGTGACA	0.49	263
(-2)	acactagCTGCCAGTC----ACTGTTACTCTAGTGACA	0.36	194
(+2)	acactagCTGCCAGTCAAGACTGTTACTCTAGTGACA	0.30	164
(-1)	acactagCTGCCAGTC-GACTGTTACTCTAGTGACA	0.17	92
(-4)	acactagCTGCCAGT-----CTGTTACTCTAGTGACA	0.09	51
(-6)	acactagCTG--------AGACTGTTACTCTAGTGACA	0.09	51
(-4)	acactagCTGCCAGTCA-----GTTACTCTAGTGACA	0.09	50

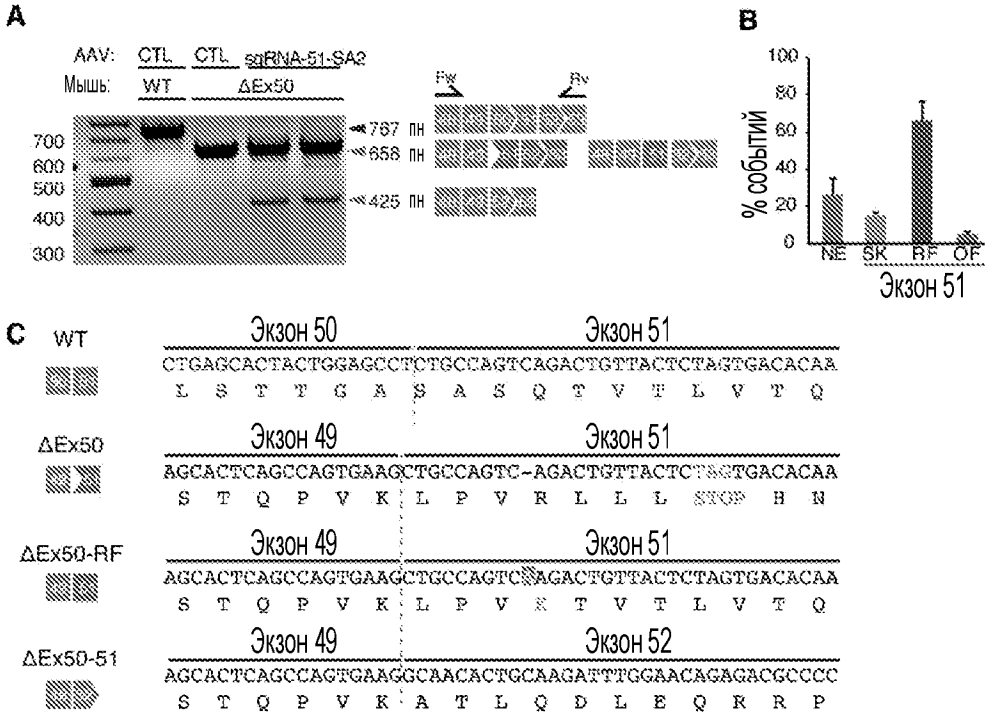
ΔEx50-AAV9-sgRNA-51-SA2-TA-#2

	PAM ↓	% чтений	Кол-во чтений
(NE)	acactagCTGCCAGTCAAGACTGTTACTCTAGTGACA	72.79	39665
(+1)	acactagCTGCCAGTCAAGACTGTTACTCTAGTGACA	15.43	8407
(S1)	acactagCTGCCAGTCATACTGTTACTCTAGTGACA	1.02	556
(-2)	acactagCTGCCAGTC---ACTGTTACTCTAGTGACA	0.34	183
(+2)	acactagCTGCCAGTCAAAGACTGTTACTCTAGTGACA	0.31	171
(+2)	acactagCTGCCAGTCAAGACTGTTACTCTAGTGACA	0.30	162
(+2)	acactagCTGCCAGTCAAAGACTGTTACTCTAGTGACA	0.25	135
(S1)	acactagCTGCCAGTCAAGACTTTTACTCTAGTGACA	0.27	147
(S1)	acactagCTGCCAGTCAGGCTGTTACTCTAGTGACA	0.21	113
(-1)	acactagCTGCCAGTC-GACTGTTACTCTAGTGACA	0.10	54
(-1)	acactagCTGCCAGTCA-TGTCACTAGAGTAACAGT	0.04	21
(-4)	acactagCTGCCAGTCA-----TGTTACTCTAGTGACA	0.04	20

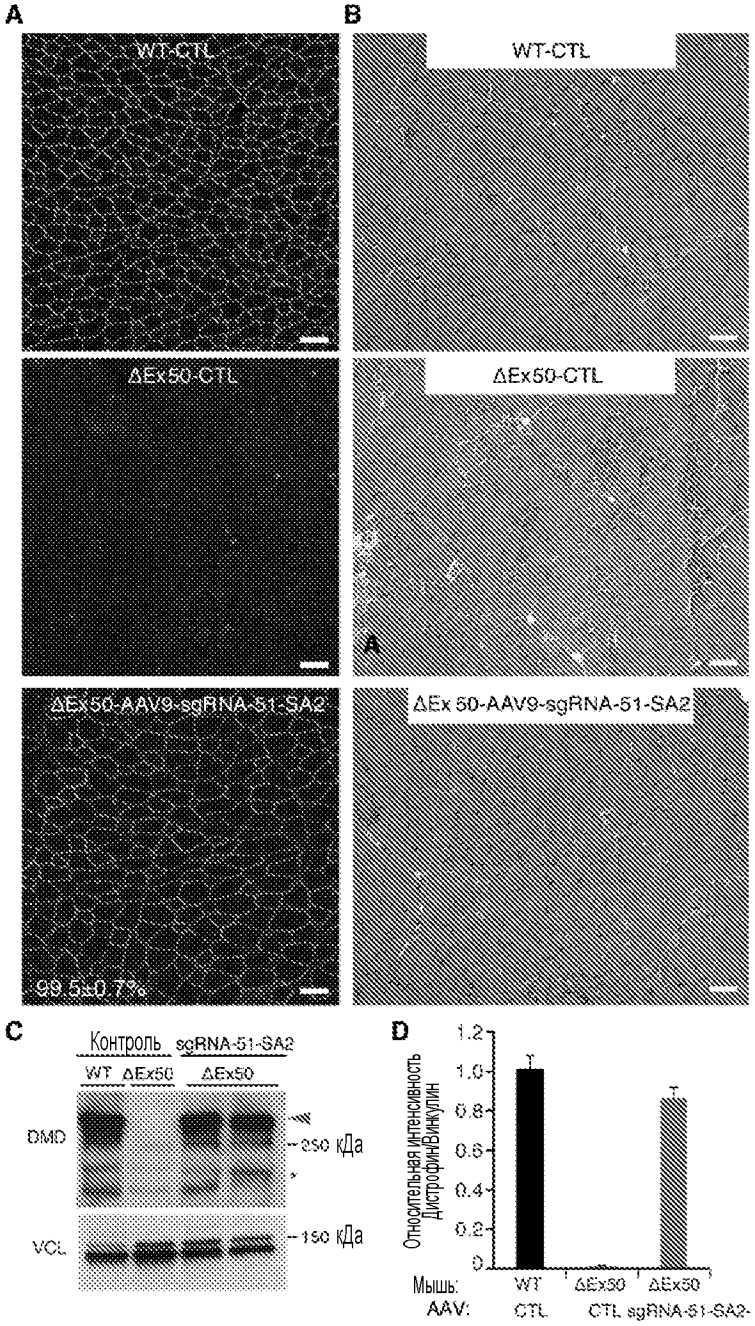
ΔEx50-AAV9-sgRNA-51--SA2-TA-#3

	PAM ↓	% чтений	Кол-во чтений
(NE)	acactagCTGCCAGTCAAGACTGTTACTCTAGTGACA	76.19	40992
(+1)	acactagCTGCCAGTCAAGACTGTTACTCTAGTGACA	12.38	6659
(S1)	acactagCTGCCAGTCATACTGTTACTCTAGTGACA	1.12	603
(+2)	acactagCTGCCAGTCAAAGACTGTTACTCTAGTGACA	0.33	175
(+1S1)	acactagCTGCCAGTCAAGACTTTTACTCTAGTGACA	0.24	131
(-1)	acactagCTGCCAGTCA-TGTCACTAGAGTAACAGT	0.18	96
(S1)	acactagCTGCCAGTCAAAACTGTTACTCTAGTGACA	0.15	79
(-2)	acactagCTGCCAGTC---TGTCACTAGAGTAACAGT	0.12	67
(S1)	acactagCTGCCAGTCCGACTGTTACTCTAGTGACA	0.11	58
(S1)	acactagCTGCCAGTTAGACTGTTACTCTAGTGACA	0.08	45
(-4)	acactagCTGCCAGT------TGTCACTAGAGTAACAG	0.05	25
(-6)	acactagCTGCCAGTCA---------TGTCACTAGAGTA	0.04	22

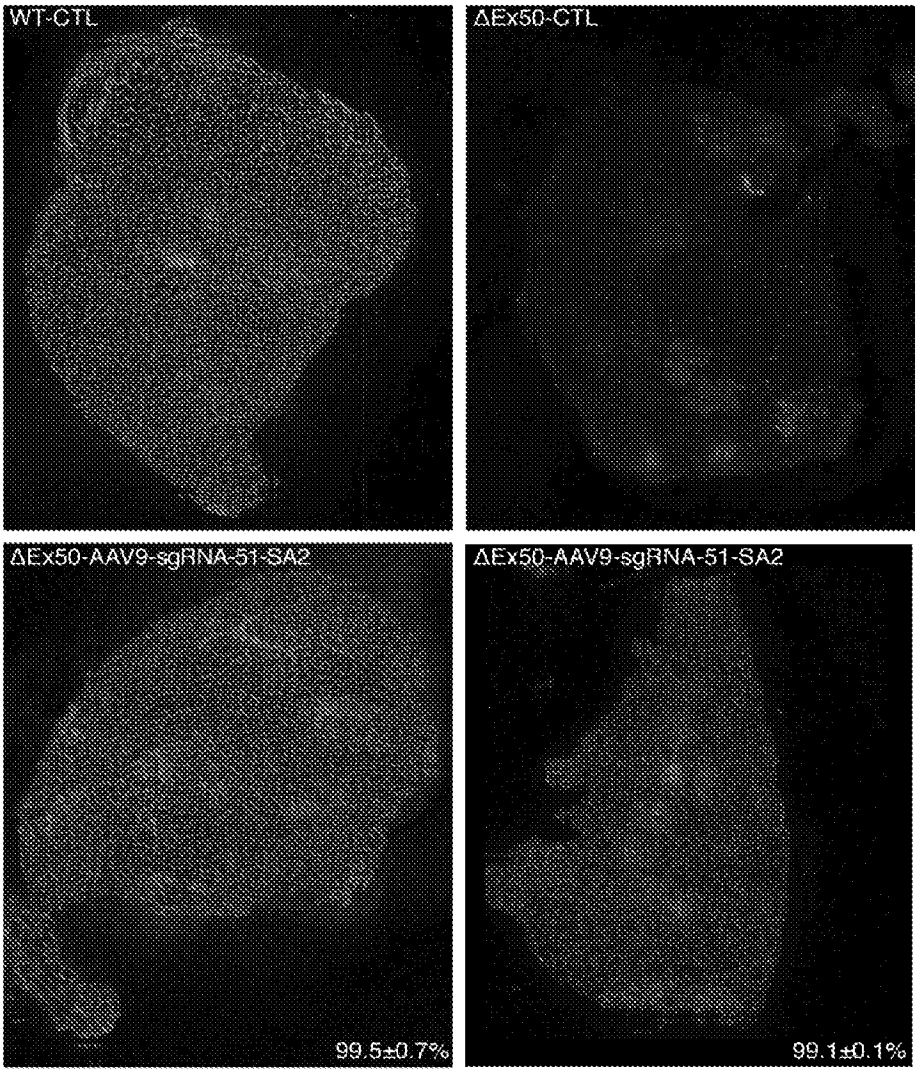
ФИГ. 10A-B


A
AAV: CTL CTL sgRNA-51-SA2
Мышь: WT ΔEx50
700
600
500
400
300
767 пн
658 пн
425 пн
Fw
Rv
B
% событий
100
80
60
40
20
0
NE SK RF OF
Экзон 51
C
WT
Экзон 50
Экзон 51
CTGAGCACTACTGGAGCCTCTGCCAGTCAGACTGTTACTCTAGTGACACAA
L S T T G A S A S Q T V T L V T Q
ΔEx50
Экзон 49
Экзон 51
AGCACTCAGCCAGTGAAGCTGCCAGTC-AGACTGTTACTCTAGTGACACAA
S T Q P V K L P V R L L L STOP H N
ΔEx50-RF
Экзон 49
Экзон 51
AGCACTCAGCCAGTGAAGCTGCCAGTCAGACTGTTACTCTAGTGACACAA
S T Q P V K L P V R T V T L V T Q
ΔEx50-51
Экзон 49
Экзон 52
AGCACTCAGCCAGTGAAGGCAACACTGCAAGATTTGGAACAGAGACGCCCC
S T Q P V K A T L Q D L E Q R R P
D
ΔEx50–без инъекции
Экзон 51
Средний % чтений
(RF+1) AAGCTGCCAGTCAGACTGTTACTCTAGTGACACAATCTGTGGTTA 0
(NE) AAGCTGCCAGTCAGACTGTTACTCTAGTGACACAATCTGTGGTTA 97.94±0.25
ΔEx50-AAV9-sgRNA-51-SA2
Экзон 51
(RF+1) AAGCTGCCAGTCAGACTGTTACTCTAGTGACACAATCTGTGGTTA 69.22±6.59
(NE) AAGCTGCCAGTCAGACTGTTACTCTAGTGACACAATCTGTGGTTA 17.71±4.49

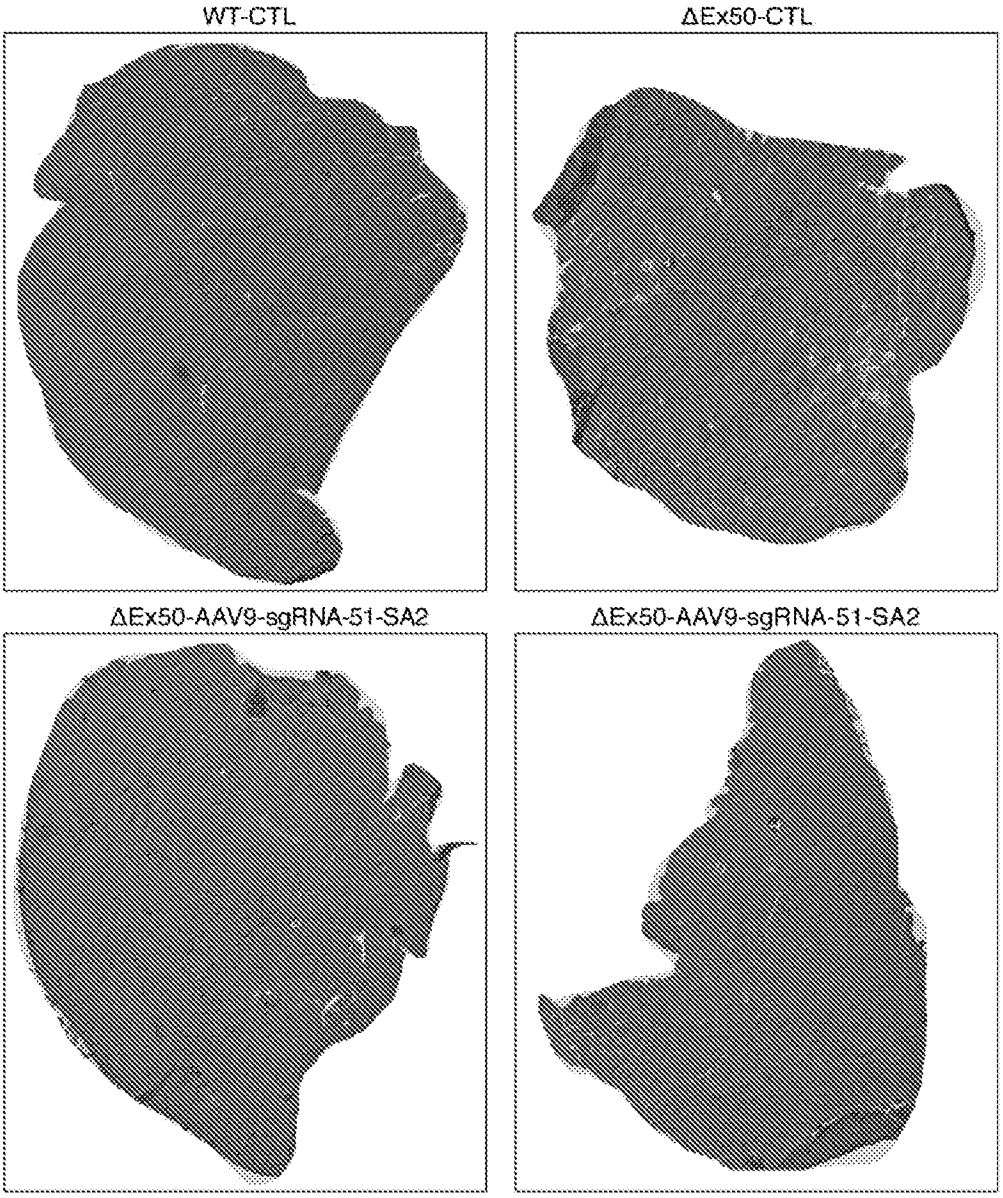
ФИГ. 11A-D


A
B
WT-CTL
ΔEx50-CTL
ΔEx50-AAV9-sgRNA-51-SA2
99.5±0.7%
WT-CTL
ΔEx50-CTL
ΔEx 50-AAV9-sgRNA-51-SA2
C
Контроль
sgRNA-51-SA2
WT
ΔEx50
ΔEx50
DMD
VCL
250 кДа
150 кДа
D
Относительная интенсивность Дистрофин/Винкулин
1.2
1.0
0.8
0.6
0.4
0.2
0
Мышь:
AAV:
WT
CTL
ΔEx50
CTL
ΔEx50
sgRNA-51-SA2-

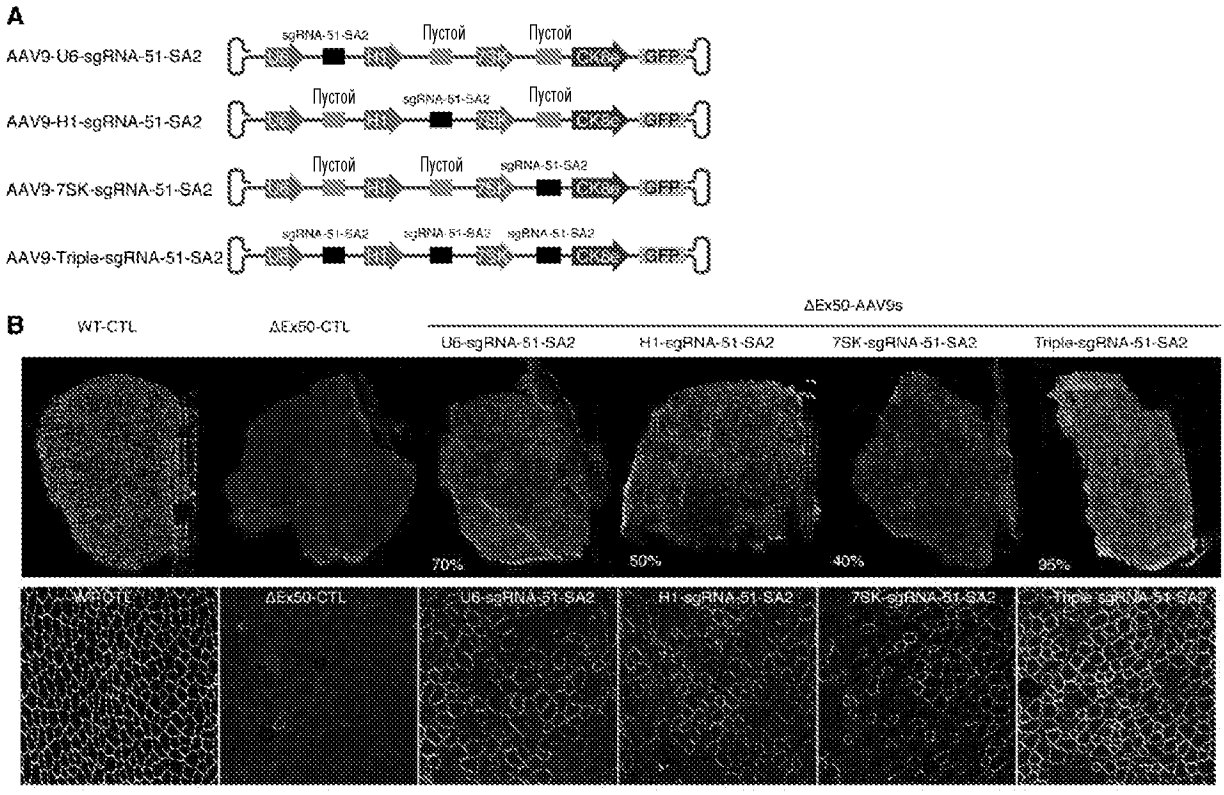
ФИГ. 12A-D


WT-CTL
ΔEx50-CTL
ΔEx50-AAV9-sgRNA-51-SA2
99.5±0.7%
ΔEx50-AAV9-sgRNA-51-SA2
99.1±0.1%

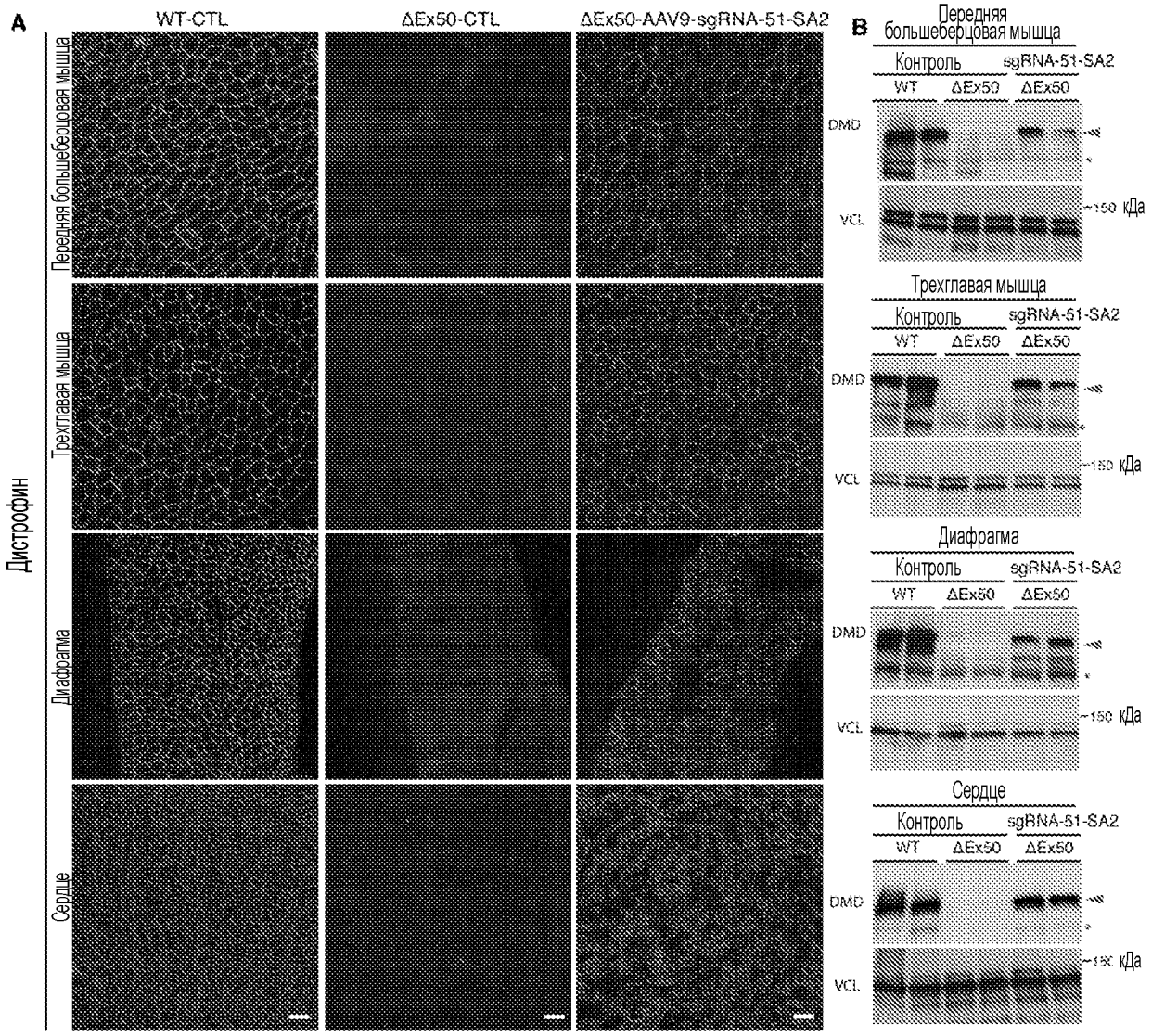
ФИГ. 13


WT-CTL
ΔEx50-CTL
ΔEx50-AAV9-sgRNA-51-SA2
ΔEx50-AAV9-sgRNA-51-SA2

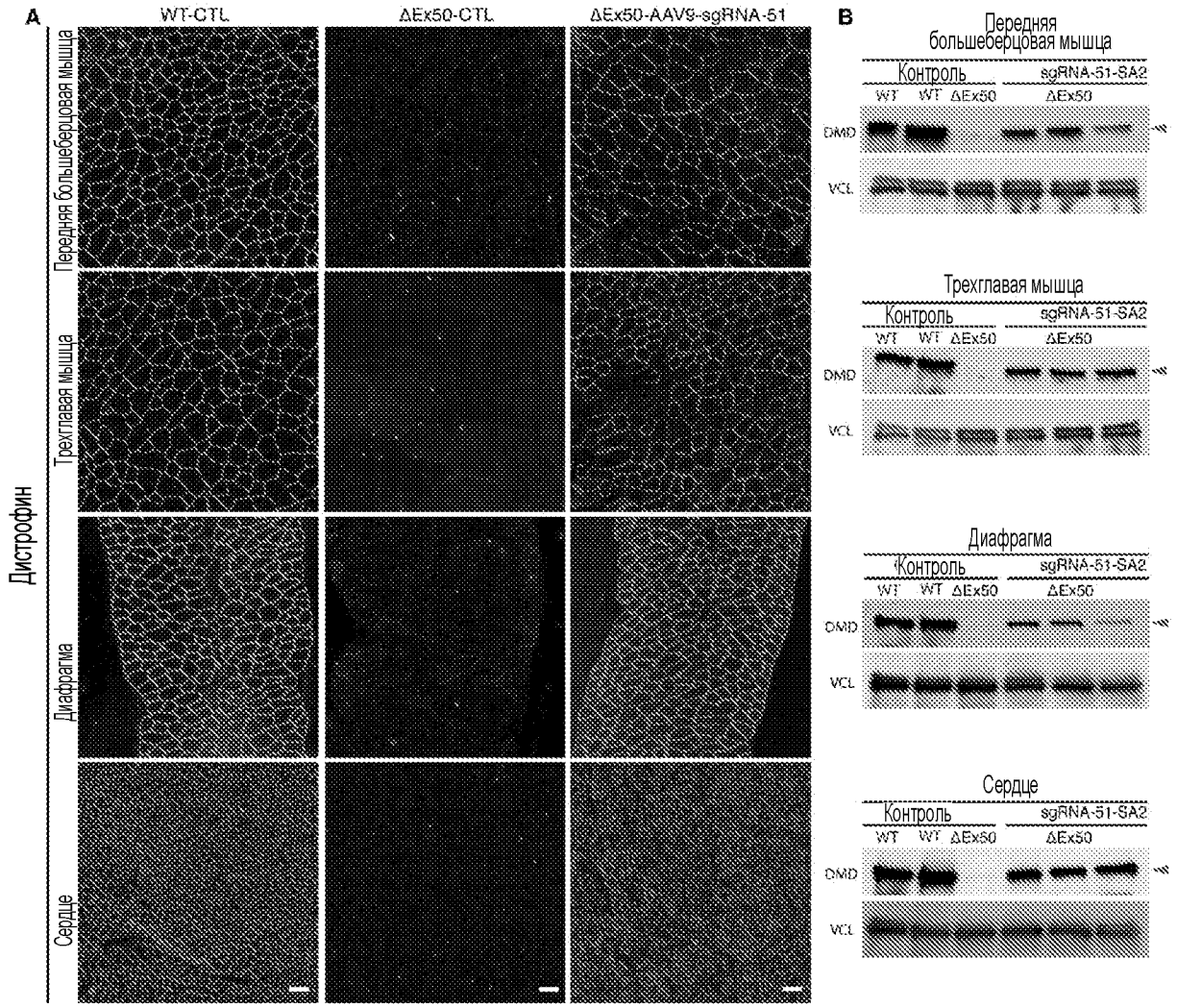
ФИГ. 14


A
AAV9-U6-sgRNA-51-SA2
sgRNA-51-SA2
Пустой
Пустой
AAV9-H1-sgRNA-51-SA2
Пустой
sgRNA-51-SA2
Пустой
AAV9-7SK-sgRNA-51-SA2
Пустой
Пустой
sgRNA-51-SA2
AAV9-Triple-sgRNA-51-SA2
sgRNA-51-SA2
sgRNA-51-SA2
sgRNA-51-SA2
B
WT-CTL
ΔEx50-CTL
ΔEx50-AAV9s
U6-sgRNA-51-SA2
H1-sgRNA-51-SA2
7SK-sgRNA-51-SA2
Triple-sgRNA-51-SA2
70%
50%
40%
95%
WT-CTL
ΔEx50-CTL
U6-sgRNA-51-SA2
H1-sgRNA-51-SA2
7SK-sgRNA-51-SA2
Triple-sgRNA-51-SA2

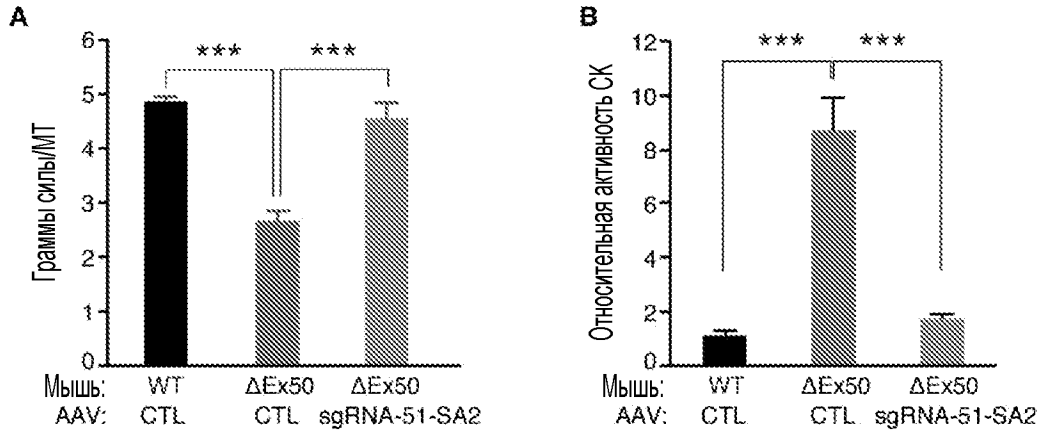
ФИГ. 15A-B


A
WT-CTL
ΔEx50-CTL
ΔEx50-AAV9-sgRNA-51-SA2
Дистрофин
Передняя большеберцовая мышца
Трехглавая мышца
Диафрагма
Сердце
B
Передняя большеберцовая мышца
Контроль
sgRNA-51-SA2
WT
ΔEx50
ΔEx50
DMD
VCL
~150 кДа
Трехглавая мышца
Диафрагма
Сердце

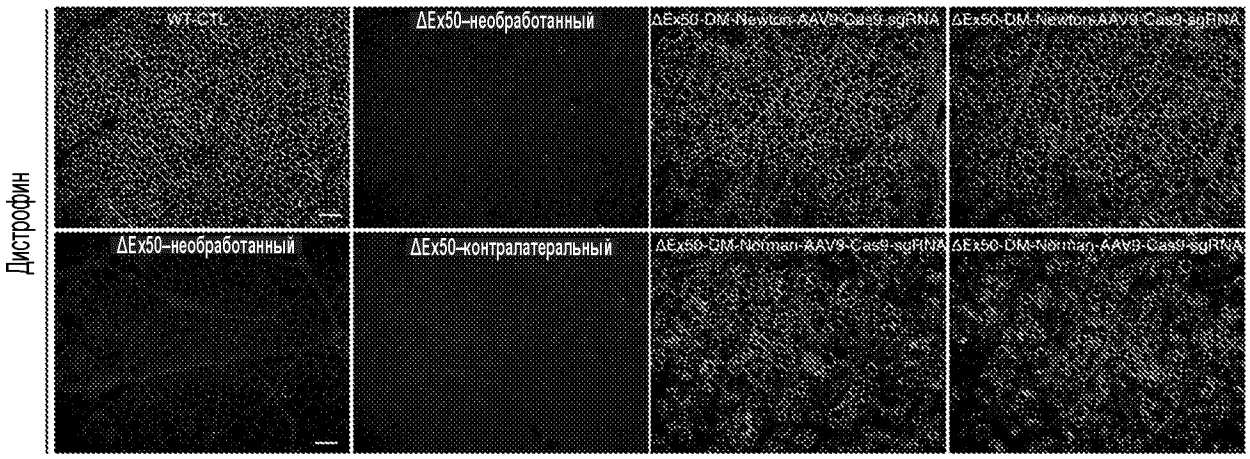
ФИГ. 16A-B


A
WT-CTL
ΔEx50-CTL
ΔEx50-AAV9-sgRNA-51
Дистрофин
Передняя большеберцовая мышца
Трехглавая мышца
Диафрагма
Сердце
B
Передняя большеберцовая мышца
Контроль
sgRNA-51-SA2
WT
WT
ΔEx50
ΔEx50
DMD
VCL
Трехглавая мышца
Диафрагма
Сердце

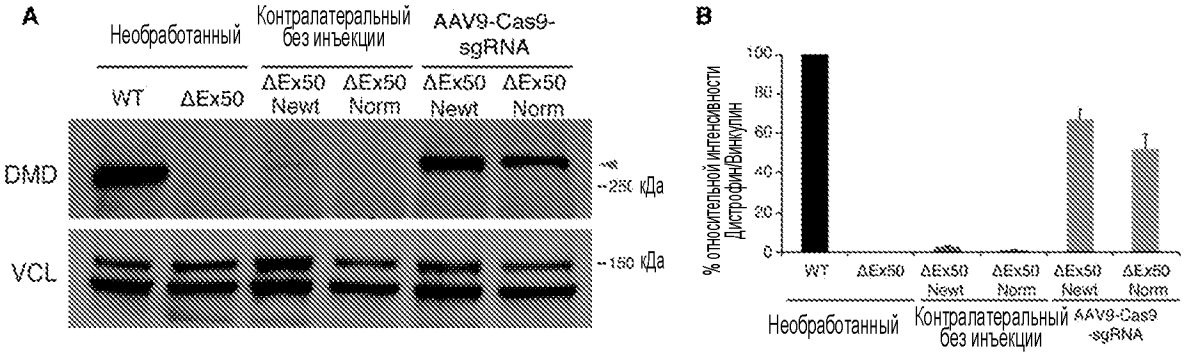
ФИГ. 17A-B



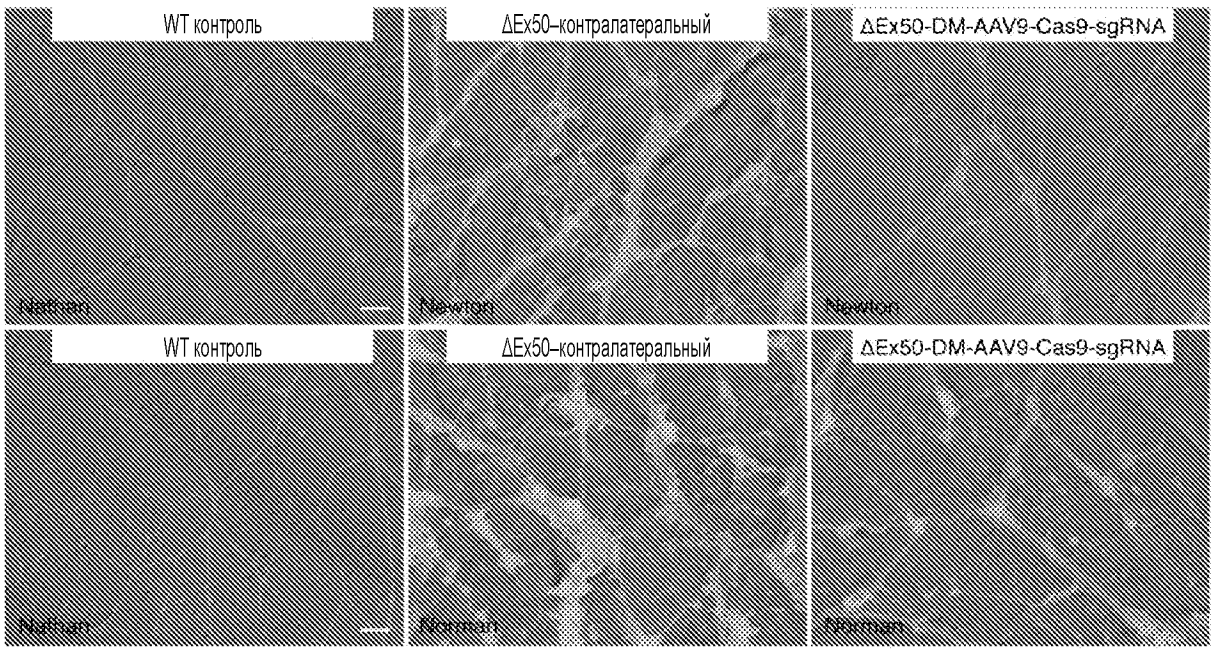
ФИГ. 18A-B



ФИГ. 19



ФИГ. 20A-B



ФИГ. 21