

(19)



**Евразийское
патентное
ведомство**

(21) **202191065** (13) **A1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки
2021.08.16

(51) Int. Cl. *C12N 15/11* (2006.01)
C12N 15/10 (2006.01)
C12N 15/86 (2006.01)

(22) Дата подачи заявки
2019.10.18

(54) **КОНСТРУКЦИИ НУКЛЕИНОВЫХ КИСЛОТ И СПОСОБЫ ИХ ПРИМЕНЕНИЯ**

(31) **62/747,393; 62/840,343**

(32) **2018.10.18; 2019.04.29**

(33) **US**

(86) **PCT/US2019/057084**

(87) **WO 2020/082041 2020.04.23**

(71) Заявитель:

**ИНТЕЛЛИА ТЕРАПЬЮТИКС,
ИНК.; РИДЖЕНЕРОН
ФАРМАСЬЮТИКАЛЗ, ИНК. (US)**

(72) Изобретатель:

**Финн Джонатан Дуглас, Хуан Хон-Рен
(US)**

(74) Представитель:

Медведев В.Н. (RU)

(57) Предложены конструкции нуклеиновых кислот, которые обеспечивают возможность вставки и/или экспрессии представляющей интерес последовательности, такой как трансген. Также предложены композиции и способы применения таких конструкций для экспрессии, например, полипептида или терапевтического агента.

A1

202191065

202191065

A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-568571EA/019

КОНСТРУКЦИИ НУКЛЕИНОВЫХ КИСЛОТ И СПОСОБЫ ИХ ПРИМЕНЕНИЯ

Данная заявка испрашивает приоритет по предварительной заявке США № 62/747393, поданной 18 октября 2018 г., и предварительной заявке США № 62/840343, поданной 29 апреля 2019 г. Описание каждой из вышеуказанных заявок в полном объеме включено в данный документ посредством ссылки.

Редактирование генома в подходах генной терапии основано на идее, что экзогенное внесение отсутствующего или, в иных случаях, дефицитного генетического материала может исправить генетическое заболевание. Генная терапия давно известна благодаря своему огромному потенциалу в практических подходах и лечении заболеваний человека. Вместо того, чтобы полагаться на лекарственные препараты или хирургическое вмешательство, пациентов с первопричинными генетическими факторами можно лечить посредством прямого нацеливания на первопричину. Кроме того, за счет нацеливания на первопричину генная терапия обладает потенциалом для эффективного излечения пациентов. При этом клинические применения существующих подходов все еще нуждаются в усовершенствовании некоторых аспектов.

В настоящем изобретении предложены двунаправленные конструкции нуклеиновых кислот, которые обеспечивают возможность более эффективной вставки и экспрессии представляющей интерес последовательности нуклеиновой кислоты, например, кодирующей терапевтический агент, такой как полипептид. В контексте данного документа двунаправленные конструкции содержат по меньшей мере два сегмента нуклеиновой кислоты, причем один сегмент (первый сегмент) содержит кодирующую последовательность, которая кодирует представляющий интерес агент (кодирующая последовательность может называться в данном документе «трансгеном» или первым трансгеном), тогда как другой сегмент (второй сегмент) содержит последовательность, в которой комплементарная последовательность кодирует представляющий интерес агент, или второй трансген. В некоторых вариантах осуществления конструкции содержат по меньшей мере два сегмента нуклеиновой кислоты, при этом один сегмент (первый сегмент) содержит последовательность, которая кодирует представляющий интерес полипептид, тогда как другой сегмент (второй сегмент) содержит последовательность, причем комплементарная последовательность кодирует представляющий интерес полипептид. При применении в комбинации с системой редактирования генов двунаправленность конструкций нуклеиновых кислот позволяет вставлять конструкции в любом направлении (без ограничения вставкой в одном направлении) в целевой сайт вставки, обеспечивая экспрессию представляющего интерес полипептида из а) кодирующей последовательности одного сегмента или 2) комплементарной последовательности другого (второго) сегмента, тем самым повышая эффективность вставки и экспрессии, как проиллюстрировано в данном документе.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

На Фиг. 1 проиллюстрированы форматы конструкций, представленных в геномах AAV. AC=акцептор сплайсинга; pA=сигнальная последовательность поли-А; ПГ=плечо гомологии; ЛПГ=левое плечо гомологии; ППГ=правое плечо гомологии.

На Фиг. 2 проиллюстрировано, что векторы без плеч гомологии неэффективны в иммортализованной линии клеток печени (Hepal-6). scAAV, полученный из плазмиды P00204, содержащей 200 п. о. плечи гомологии, приводил к обнаруживаемой экспрессии hFIX в этой линии клеток. Использование AAV-векторов, полученных из P00123 (scAAV с отсутствием плеч гомологии) и P00147 (двунаправленная конструкция ssAAV с отсутствием плеч гомологии), не приводило к обнаруживаемой экспрессии hFIX.

На Фиг. 3А и 3В проиллюстрированы результаты *in vivo* исследования вставочных матриц с плечами гомологии и без них с использованием векторов, полученных из P00123, P00147 или P00204. На Фиг. 3А проиллюстрировано, что в каждой группе животных, обработанных ЛНЧ, содержащими компоненты системы CRISPR/Cas9, были обнаружены уровни редактирования в печени, составляющие ~60% по данным измерений инделей. На Фиг. 3В проиллюстрировано, что животные, получавшие ssAAV-векторы без плеч гомологии (полученные из P00147) в комбинации с обработкой ЛНЧ, имели наибольший уровень экспрессии hFIX в сыворотке.

На Фиг. 4А и 4В проиллюстрированы результаты *in vivo* исследования вставочных матриц ssAAV с плечами гомологии и без них. На Фиг. 4А проиллюстрировано сравнение направленной вставки с помощью векторов, полученных из плазмид P00350, P00356, P00362 (имеющих, как проиллюстрировано, асимметричные плечи гомологии) и P00147 (двунаправленной конструкции, проиллюстрированной на Фиг. 4В). На Фиг. 4В проиллюстрировано сравнение вставки во второй сайт, на который были нацелены векторы, полученные из плазмид P00353, P00354 (имеющие, как проиллюстрировано, симметричные плечи гомологии) и P00147.

На Фиг. 5А-5D проиллюстрированы результаты направленной вставки трех двунаправленных конструкций среди 20 целевых сайтов в первичных гепатоцитах мышей. На Фиг. 5А приведены схематические изображения каждого из исследуемых векторов. На Фиг. 5В проиллюстрировано редактирование по данным измерения образования инделей для каждой из групп обработки для каждой исследуемой комбинации. На Фиг. 5С и Фиг. 5D проиллюстрировано, что значительные уровни редактирования (в конкретном целевом сайте) не обязательно приводят к более эффективной вставке или экспрессии трансгенов. В этом исследовании нацеленной вставки тестируемые конструкции эффективно приводили к экспрессии трансгена. cAC=акцептор сплайсинга F9 человека; mAC=акцептор сплайсинга альбумина мыши; HiBit=тег для обнаружения на основе люциферазы; pA= последовательность сигнала поли-А; Nluc=репортер нанолуциферазы; ЗФБ=зеленый флуоресцентный белок.

На Фиг. 6 проиллюстрированы результаты *in vivo* скрининга направленной вставки с помощью двунаправленных конструкций среди 10 целевых сайтов с использованием ssAAV, полученного из P00147. Как проиллюстрировано, значительные уровни

редактирования не обязательно приводят к высоким уровням экспрессии трансгена.

На Фиг. 7A-7D проиллюстрированы результаты *in vivo* скрининга двунаправленных конструкций среди 20 целевых сайтов с использованием ssAAV, полученного из P00147. На Фиг. 7A проиллюстрировано редактирование, обнаруженное для каждой из групп обработки для каждой исследуемой комбинации ЛНЧ/вектор. На Фиг. 7B приведены соответствующие данные по направленной вставке. Результаты демонстрируют слабую корреляцию между редактированием и вставкой/экспрессией двунаправленных конструкций (Фиг. 7B и Фиг. 7D), а также положительную корреляцию между результатами *in vitro* и *in vivo* (Фиг. 7C).

На Фиг. 8A и 8B проиллюстрирована введение двунаправленной конструкции на клеточном уровне методом гибридизации *in situ* с использованием зондов, которые могут обнаруживать соединения между трансгеном hFIX и последовательностью экзона 1 мышинового альбумина (Фиг. 8A). Уровни циркулирующего hFIX коррелировали с числом клеток, положительных в отношении гибридного транскрипта (Фиг. 8B).

На Фиг. 9a проиллюстрирована продолжительность экспрессии hFIX *in vivo*. На Фиг. 9b продемонстрировано, что экспрессия из интрона 1 альбумина была продолжительной.

На Фиг. 10A-10B проиллюстрировано, что изменение дозы AAV или ЛНЧ может модулировать степень экспрессии hFIX из интрона 1 гена альбумина *in vivo*.

На Фиг. 11A-11C проиллюстрированы результаты скрининга двунаправленных конструкций среди целевых сайтов в первичных гепатоцитах яванских макаков. На Фиг. 11A проиллюстрированы различные уровни редактирования для каждого из образцов по данным измерения образования инделей. Фиг. 11B и Фиг. 11C проиллюстрировано, что значительные уровни образования инделей не были прогностическими в отношении вставки или экспрессии двунаправленных конструкций в интрон 1 альбумина.

На Фиг. 12A-12C проиллюстрированы результаты скрининга двунаправленных конструкций среди целевых сайтов в первичных гепатоцитах человека. На Фиг. 12A проиллюстрировано редактирование для каждого из образцов по данным измерения образования инделей. На Фиг. 12B, Фиг. 12C и Фиг. 12D проиллюстрировано, что значительные уровни образования инделей не были прогностическими в отношении вставки или экспрессии двунаправленных конструкций в интрон 1 гена альбумина.

На Фиг. 13 проиллюстрированы результаты *in vivo* исследований, в которых отличным от человека приматам вводили дозу ЛНЧ вместе с двунаправленной вставочной матрицей hFIX (полученной из P00147). Системные уровни hFIX были достигнуты только у животных, обработанных как ЛНЧ, так и AAV, причем при применении AAV или ЛНЧ по отдельности hFIX не подлежал обнаружению.

ПОДРОБНОЕ ОПИСАНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

Далее будет приведено подробное описание определенных вариантов осуществления изобретения, примеры которых проиллюстрированы на прилагаемых графических материалах. Хотя представленные идеи описаны в сочетании с различными

вариантами осуществления, следует понимать, что это не подразумевает ограничения представленных идей данными вариантами осуществления. Наоборот, подразумевается, что представленные идеи охватывают все альтернативные варианты, модификации и эквиваленты, которые могут быть предположены специалистами в данной области техники.

Перед подробным описанием представленных идей следует понимать, что изобретение не ограничено конкретными композициями или этапами процессов, поскольку они могут варьироваться. Следует отметить, что, в контексте данного описания и прилагаемой формулы изобретения форма единственного числа включает отсылку на множественное число, если из контекста не следует иное. Таким образом, например, ссылка на «конъюгат» включает множество конъюгатов, а ссылка на «клетку» включает множество клеток и т. п. В контексте данного документа термин «включать» и его грамматические варианты не подразумевают ограничения, поэтому перечисление элементов в списке не исключает другие подобные элементы, которыми можно заменять или которые можно добавлять к перечисленным элементам.

Числовые диапазоны включают числа, определяющие диапазон. Подразумевается, что измеренные и измеряемые значения являются приблизительными с учетом значимых знаков и погрешности, связанной с измерением. Также использование терминов «содержать», «содержит», «содержащий», «включать», «включает» и «включающий» не подразумевает ограничения. Следует понимать, что вышеприведенное общее описание и нижеприведенное подробное описание являются лишь иллюстративными и пояснительными и не ограничивают идеи изобретения.

Если специально не указано в описании, варианты осуществления в описании, которые указаны как «содержащие» различные компоненты, также рассматриваются как «состоящие из» или «состоящие преимущественно из» перечисленных компонентов; варианты осуществления в описании, которые указаны как «состоящие из» различных компонентов, также рассматриваются как «содержащие» или «состоящие преимущественно из» перечисленных компонентов; и варианты осуществления в описании, которые указаны как «состоящие преимущественно из» различных компонентов, также рассматриваются как «состоящие из» или «содержащие» перечисленные компоненты (эта взаимозаменяемость не применима к использованию этих терминов в формуле изобретения). Термин «или» используется во включительном смысле, т. е. он эквивалентен «и/или», если из контекста четко не следует иное. Термин «около», используемый перед перечнем, модифицирует каждый элемент перечня. Термин «около» или «приблизительно» означает допустимую ошибку для конкретного значения, определяемую специалистом в данной области техники, которая частично зависит от того, как данное значение измеряют или определяют.

Заголовки разделов, используемые в данном документе, приведены исключительно в организационных целях, и их не следует толковать как ограничивающие каким-либо образом предмет изобретения. В случае, если какой-либо материал, включенный

посредством ссылки, противоречит любому термину, определенному в этом описании, или любому другому явному содержанию этого описания, приоритет имеет это описание.

Определения

Если не указано иное, в контексте данного документа следующие термины и выражения имеют следующие значения:

В контексте данного документа «полинуклеотид» и «нуклеиновая кислота» обозначают мультимерное соединение, содержащее нуклеозиды или аналоги нуклеозидов, которые состоят из азотистых гетероциклических оснований или аналогов оснований, связанных вместе вдоль остова, включая традиционные РНК, ДНК, смешанные РНК-ДНК и полимеры, которые являются их аналогами. «Остов» нуклеиновой кислоты может состоять из ряда связей, включая одну или более сахарно-фосфодиэфирных связей, связей пептид - нуклеиновая кислота («пептидные нуклеиновые кислоты» или ПНК; РСТ № WO 95/32305), тиофосфатных связей, метилфосфонатных связей или их комбинаций. Сахарные фрагменты нуклеиновой кислоты могут представлять собой рибозу, дезоксирибозу или аналогичные соединения с необязательными заменами, например, 2'-метокси- или 2'-галогенидными заменами. Азотистые основания могут представлять собой традиционные основания (A, G, C, T, U), их аналоги (например, модифицированные уридины, такие как 5-метоксиуридин, псевдоуридин или N1-метилпсевдоуридин или другие); инозин; производные пуринов или пиримидинов (например, N⁴-метилдезоксигуанозин, деаза- или азапурины, деаза- или азапиримидины, пиримидиновые основания с группами заместителей в позиции 5 или 6 (например, 5-метилцитозин), пуриновые основания с заместителем в позициях 2, 6 или 8, 2-амино-6-метиламинопурин, O⁶-метилгуанин, 4-тиопиримидины, 4-аминопиримидины, 4-диметилгидразинпиримидины и O⁴-алкилпиримидины; патент США № 5378825 и РСТ № WO 93/13121). Общее обсуждение смотрите в *The Biochemistry of the Nucleic Acids* 5-36, Adams et al., ed., 11th ed., 1992). Нуклеиновые кислоты могут содержать один или более «абазических» остатков, остов которых не содержит азотистых оснований для позиции(й) полимера (патент США № 5585481). Нуклеиновая кислота может содержать только традиционные сахара, основания и связи РНК или ДНК или может включать как традиционные компоненты, так и замены (например, традиционные нуклеозиды с 2'-метокси-заместителями или полимеры, содержащие как традиционные нуклеозиды, так и один или более нуклеотидных аналогов). Нуклеиновая кислота включает «закрытую нуклеиновую кислоту» (ЗНК), аналог, содержащий один или более нуклеотидных мономеров ЗНК с бициклическим фуранозным звеном, замкнутым в имитирующую РНК сахарную конформацию, что повышает аффинность гибридизации к комплементарным последовательностям РНК и ДНК (Vester and Wengel, 2004, *Biochemistry* 43(42):13233-41). РНК и ДНК имеют разные сахарные фрагменты и могут отличаться наличием урацила или его аналогов в РНК и тимина или его аналогов в ДНК.

Термины «гидовая (направляющая) РНК», «гРНК» и просто «направляющая» взаимозаменяемо используют в данном документе для обозначения направляющей,

которая содержит направляющую последовательность, например, crРНК (также известную как CRISPR РНК), или комбинацию crРНК и trРНК (также известной как tracrРНК). crРНК и trРНК могут быть связаны в виде одной молекулы РНК (одинарная гидовая РНК, ogРНК) или, например, в виде двух отдельных молекул РНК (двойная гидовая РНК, dgРНК). Термин «гидовая РНК» или «гРНК» относится к каждому типу. trРНК может представлять собой последовательность природного происхождения или последовательность trРНК с модификациями или вариациями по сравнению с последовательностями природного происхождения. Гидовые РНК, такие как ogРНК или dgРНК, могут включать модифицированные РНК, как описано в данном документе.

В контексте данного документа «направляющая последовательность» относится к последовательности в гидовой РНК, которая является комплементарной целевой последовательности, а ее функция состоит в направлении гидовой РНК к целевой последовательности для связывания или модификации (например, расщепления) РНК-направляемым ДНК-связывающим агентом. «Направляющая последовательность» также может называться «нацеливающей последовательностью» или «спейсерной последовательностью». Направляющая последовательность может иметь длину в 20 пар оснований, например, в случае *Streptococcus pyogenes* (т. е. Spv Cas9) и родственных гомологов/ортологов Cas9. Также в качестве направляющих можно использовать более короткие или более длинные последовательности, например, длиной 15, 16, 17, 18, 19, 21, 22, 23, 24 или 25 нуклеотидов. В некоторых вариантах осуществления целевая последовательность находится, например, в гене или хромосоме и является комплементарной направляющей последовательности. В некоторых вариантах осуществления степень комплементарности или идентичности между направляющей последовательностью и соответствующей ей целевой последовательностью может составлять по меньшей мере около 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% или 100%. В некоторых вариантах осуществления направляющая последовательность и целевая область могут быть на 100% комплементарными или идентичными. В других вариантах осуществления направляющая последовательность и целевая область могут содержать по меньшей мере одно несовпадение. Например, направляющая последовательность и целевая последовательность могут содержать 1, 2, 3 или 4 несовпадения, при этом общая длина целевой последовательности составляет по меньшей мере 17, 18, 19, 20 или более пар оснований. В некоторых вариантах осуществления направляющая последовательность и целевая последовательность могут содержать 1-4 несовпадения, при этом направляющая последовательность содержит по меньшей мере 17, 18, 19, 20 или более нуклеотидов. В некоторых вариантах осуществления направляющая последовательность и целевая область могут содержать 1, 2, 3 или 4 несовпадения, при этом направляющая последовательность содержит 20 нуклеотидов.

Целевые последовательности для РНК-направляемых ДНК-связывающих агентов включают как положительные, так и отрицательные цепи геномной ДНК (т. е. заданную последовательность и обратную комплементарную последовательность), поскольку

субстратом нуклеиновой кислоты для РНК-направляемого ДНК-связывающего агента является двухцепочечная нуклеиновая кислота. Соответственно, когда говорят, что направляющая последовательность «является комплементарной целевой последовательности», следует понимать, что направляющая последовательность может направлять гидовую РНК для связывания со смысловой или антисмысловой цепью (например, обратной комплементарной последовательностью) целевой последовательности. Таким образом, в некоторых вариантах осуществления, в которых направляющая последовательность связывает обратную комплементарную последовательность целевой последовательности, направляющая последовательность идентична определенным нуклеотидам целевой последовательности (например, целевой последовательности, не содержащей РАМ), за исключением замены U на T в гидовой последовательности.

В контексте данного документа «РНК-направляемый ДНК-связывающий агент» означает полипептид или комплекс полипептидов, обладающих активностью связывания РНК и ДНК, или ДНК-связывающую субъединицу такого комплекса, причем ДНК-связывающая активность является специфической для последовательности и зависит от последовательности РНК. Термин «РНК-направляемый ДНК-связывающий агент» также включает нуклеиновые кислоты, кодирующие такие полипептиды. Типовые РНК-направляемые ДНК-связывающие агенты включают кливазы/никазы Cas. Типовые РНК-направляемые ДНК-связывающие агенты могут включать инактивированные формы («ДНК-связывающие агенты dCas»), например, если эти агенты модифицированы с возможностью расщепления ДНК, например, посредством слияния с доменом расщепления FokI. В контексте данного документа термин «Cas-нуклеаза» включает Cas-кливазы и Cas-никазы. Cas-кливазы и Cas-никазы включают комплекс Csm или Cmr системы CRISPR типа III, его субъединицу Cas10, Csm1 или Cmr2, комплекс Cascade системы CRISPR типа I, ее субъединицу Cas3 и Cas-нуклеазы 2 класса. В контексте данного документа «Cas-нуклеаза 2 класса» представляет собой одноцепочечный полипептид с РНК-направляемой ДНК-связывающей активностью. Cas-нуклеазы 2 класса включают кливазы/никазы Cas 2 класса (например, варианты H840A, D10A или N863A), которые дополнительно обладают активностью РНК-направляемых кливаз или никаз ДНК и ДНК-связывающих агентов dCas класса 2, в которых активность кливазы/никазы инактивирована), если эти агенты модифицированы с возможностью расщепления ДНК. Cas-нуклеазы 2 класса включают, например, белки Cas9, Cpf1, C2c1, C2c2, C2c3, HF Cas9 (например, варианты N497A, R661A, Q695A, Q926A), НупаCas9 (например, варианты N692A, M694A, Q695A, H698A), eSPCas9(1,0) (например, варианты K810A, K1003A, R1060A) и eSPCas9(1,1) (например, например K848A, K1003A, R1060A) и их модификации. Белок Cpf1, Zetsche et al., Cell, 163: 1-13 (2015), также содержит RuvC-подобный нуклеазный домен. Последовательности Cpf1 Zetsche в полном объеме включены посредством ссылки. Смотрите, например, Zetsche, таблицы S1 и S3. Смотрите, например, Makarova et al., Nat Rev Microbiol, 13(11): 722-36 (2015); Shmakov et al.,

Molecular Cell, 60:385-397 (2015). В контексте данного документа доставка РНК-направляемого ДНК-связывающего агента (например, Cas-нуклеазы, Cas9-нуклеазы или Cas9-нуклеазы *S. pyogenes*) включает доставку полипептида или мРНК.

В контексте данного документа термин «рибонуклеопротеин» (РНП) или «комплекс РНП» относится к гидовой РНК вместе с РНК-направляемым ДНК-связывающим агентом, таким как Cas-нуклеаза, например, Cas-квиваза, Cas-никаза, Cas9-квиваза или Cas9-никаза. В некоторых вариантах осуществления гидовая РНК направляет РНК-направляемый ДНК-связывающий агент, такой как Cas9, к целевой последовательности, агент связывается с целевой последовательностью, а гидовая РНК гибридизуется с ней; а за связыванием может следовать расщепление или создание одноцепочечного разрыва.

В контексте данного документа считается, что первая последовательность «содержит последовательность с по меньшей мере X % идентичности со» второй последовательностью, если выравнивание первой последовательности относительно второй последовательностью показывает, что X % или более позиций второй последовательности совпадают с первой последовательностью. Например, последовательность AAGA содержит последовательность со 100% идентичности с последовательностью AAG, поскольку выравнивание даст 100% идентичности в том смысле, что присутствуют совпадения со всеми тремя позициями второй последовательности. Различия между РНК и ДНК (в общем случае замена уридина на тимидин или наоборот) и наличие аналогов нуклеозидов, таких как модифицированные уридины, не влияют на различия в идентичности или комплементарности среди полинуклеотидов при условии, что соответствующие нуклеотиды (такие как тимидин, уридин или модифицированный уридин) имеют один и тот же комплемент (например, аденозин для всех из тимидина, уридина или модифицированного уридина; другим примером является цитозин и 5-метилцитозин, которые оба содержат гуанозин или модифицированный гуанозин в качестве комплемента). Таким образом, например, последовательность 5'-AXG, где X представляет собой любой модифицированный уридин, такой как псевдоуридин, N1-метил-псевдоуридин или 5-метоксиуридин, считается на 100% идентичной AUG в том смысле, что они обе полностью комплементарны одной и той же последовательности (5'-CAU). Типовыми алгоритмами выравнивания являются алгоритмы Смита - Уотермана и Нидлмана - Вунша, которые хорошо известны в данной области техники. Специалист в данной области поймет, какой выбор алгоритма и настроек параметров подходит для заданной пары последовательностей, подлежащих выравниванию; в случае последовательностей в целом одинаковой длины и с ожидаемой идентичностью > 50% для аминокислот или > 75% для нуклеотидов в общем случае подходит алгоритм Нидлмана - Вунша с настройками интерфейса алгоритма Нидлмана - Вунша по умолчанию, предоставленный EBI на веб-сервере www.ebi.ac.uk.

В контексте данного документа первая последовательность считается «на X %

комплементарной» второй последовательности, если X % оснований первой последовательности спариваются со второй последовательностью. Например, первая последовательность 5'ААГА3' является на 100% комплементарной второй последовательности 3'ТТСТ5', а вторая последовательность является на 100% комплементарной первой последовательности. В некоторых вариантах осуществления первая последовательность 5'ААГА3' является на 100% комплементарной второй последовательности 3'ТТСТГТГА5', тогда как вторая последовательность является на 50% комплементарной первой последовательности.

В контексте данного документа термин «мРНК» относится к полинуклеотиду, который полностью или преимущественно представляет собой РНК или модифицированную РНК и содержит открытую рамку считывания, которая может транслироваться в полипептид (т. е. может служить субстратом для трансляции рибосомой и аминокислотами тРНК). мРНК может содержать фосфатно-сахарный остов, включая остатки рибозы или их аналоги, например, остатки 2'-метоксирибозы. В некоторых вариантах осуществления сахара фосфатно-сахарного остова мРНК состоят преимущественно из остатков рибозы, остатков 2'-метоксирибозы или их комбинации. Основания мРНК могут представлять собой модифицированные основания, такие как псевдоуридин, N-1-метилпсевдоуридин или другие основания природного и не природного происхождения.

Типовые направляющие последовательности, применимые в описанных в данном документе композициях и способах, приведены в таблице 1 и по всему объему заявки.

В контексте данного документа «индели» относятся к инсерционным/делеционным мутациям, состоящим из ряда нуклеотидов, которые вставлены или удалены в сайте двухцепочечных разрывов (ДЦР) в целевой нуклеиновой кислоте.

В контексте данного документа термин «полипептид» относится к белку дикого типа или вариантному белку (например, мутанту, фрагменту, слиянию или их комбинациям). Вариантный полипептид может обладать по меньшей мере или около 5%, 10%, 15%, 20%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% или 100% функциональной активности полипептида дикого типа. В некоторых вариантах осуществления вариант является по меньшей мере на 70%, 75%, 80%, 85%, 90%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентичным последовательности полипептида дикого типа. В некоторых вариантах осуществления вариантный полипептид может представлять собой гиперактивный вариант. В определенных случаях вариант обладает от около 80% до около 120%, 140%, 160%, 180%, 200%, 300%, 400%, 500% или более функциональной активности полипептида дикого типа.

В контексте данного документа термин «гетерологичный ген» относится к гену, который был внесен в качестве экзогенного источника в сайт в геноме клетки-хозяина (например, в геномный локус, такой как безопасный приемный локус, включая сайт интрона 1 альбумина). Это означает, что внесенный ген является гетерологичным по отношению к сайту вставки. Полипептид, экспрессируемый из такого гетерологичного

гена, называется «гетерологичным полипептидом». Гетерологичный ген может иметь природное происхождение или быть сконструированным, и может быть дикого типа или вариантным. Гетерологичный ген может содержать нуклеотидные последовательности, отличные от последовательности, кодирующей гетерологичный полипептид (например, сайт внутренней посадки рибосомы). Гетерологичный ген может представлять собой ген, который встречается в природе в геноме хозяина в виде гена дикого типа или варианта (например, мутантного). Например, хотя клетка-хозяин содержит представляющий интерес ген (как ген дикого типа или вариант), тот же ген или его вариант можно вносить в качестве экзогенного источника, например, для экспрессии в локусе с высокой экспрессией. Гетерологичный ген также может представлять собой ген, который не встречается в природе в геноме хозяина, или который экспрессирует гетерологичный полипептид, который в природе не встречается в геноме хозяина. Термины «гетерологичный ген», «экзогенный ген» и «трансген» используют взаимозаменяемо. В некоторых вариантах осуществления гетерологичный ген или трансген содержит экзогенную последовательность нуклеиновой кислоты, например, последовательность нуклеиновой кислоты не является эндогенной для клетки-реципиента. В определенных вариантах осуществления гетерологичный ген не встречается в природе в клетке-реципиенте. Например, гетерологичный ген может быть гетерологичным в отношении сайта его вставки и в отношении клетки-реципиента.

В контексте данного документа «целевая последовательность» относится к последовательности нуклеиновой кислоты в целевом гене, которая комплементарна направляющей последовательности гРНК. Взаимодействие целевой последовательности и направляющей последовательности заставляет РНК-направляемый ДНК-связывающий агент связываться и потенциально создавать одноцепочечный разрыв или расщепление (в зависимости от активности агента) внутри целевой последовательности.

В контексте данного документа «двунаправленная конструкция нуклеиновой кислоты» (взаимозаменяемо называемая в данном документе «двунаправленной конструкцией») содержит по меньшей мере два сегмента нуклеиновой кислоты, причем один сегмент (первый сегмент) содержит кодирующую последовательность, которая кодирует представляющий интерес агент (кодирующая последовательность может называться в данном документе «трансгеном» или первым трансгеном), тогда как другой сегмент (второй сегмент) содержит последовательность, в которой комплементарная последовательность кодирует представляющий интерес агент, или второй трансген. Агент может представлять собой терапевтический агент, такой как полипептид, функциональная РНК, мРНК и т. п. Трансген может кодировать агент, такой как полипептид, функциональная РНК или мРНК. В некоторых вариантах осуществления двунаправленная конструкция нуклеиновой кислоты содержит по меньшей мере два сегмента нуклеиновой кислоты, при этом один сегмент (первый сегмент) содержит последовательность, которая кодирует представляющий интерес полипептид, тогда как другой сегмент (второй сегмент) содержит последовательность, причем комплементарная последовательность

кодирует представляющий интерес полипептид, или второй трансген. Это означает, что по меньшей мере два сегмента могут кодировать идентичные или разные полипептиды или идентичные или разные агенты. Когда два сегмента кодируют идентичный полипептид, кодирующая последовательность первого сегмента не обязательно должна быть идентична комплементарной последовательности второго сегмента. В некоторых вариантах осуществления последовательность второго сегмента является обратной комплементарной последовательностью кодирующей последовательности первого сегмента. Двухнаправленная конструкция может быть одноцепочечной или двухцепочечной. Описанная в данном документе двухнаправленная конструкция включает конструкцию, которая способна экспрессировать любой представляющий интерес полипептид. Двухнаправленные конструкции применимы для геномной вставки трансгенных последовательностей, в частности, нацеленной вставки трансгена.

В контексте данного документа термин «обратная комплементарная последовательность» относится к последовательности, которая является комплементарной последовательностью референсной последовательности, при этом комплементарная последовательность записана в обратной ориентации. Например, для гипотетическое последовательности 5' CTGGACCGA 3' (SEQ ID NO: 500) «идеальной» комплементарной последовательностью является 3' GACCTGGCT 5' (SEQ ID NO: 501), а «идеальная» обратная комплементарная последовательность записывается 5' TCGGTCCAG 3' (SEQ ID NO: 502). Обратная комплементарная последовательность не обязательно должна быть «идеальной» и при этом может кодировать тот же полипептид или сходный полипептид, что и референсная последовательность. Вследствие избыточной частоты использования кодонов обратная комплементарная последовательность может отличаться от референсной последовательности, кодирующей тот же полипептид. В контексте данного документа термин «обратная комплементарная последовательность» также включает последовательности, которые являются, например, по меньшей мере на 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичными обратной комплементарной последовательности референсной последовательности.

В некоторых вариантах осуществления двухнаправленная конструкция нуклеиновой кислоты содержит первый сегмент, который содержит кодирующую последовательность, кодирующую первый полипептид (первый трансген), и второй сегмент, который содержит последовательность, в которой комплементарная последовательность кодирует второй полипептид (второй трансген). В некоторых вариантах осуществления первый и второй полипептиды являются по меньшей мере на 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичными. В некоторых вариантах осуществления первый и второй полипептиды содержат аминокислотную последовательность, которая является по меньшей мере на 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичной, например, по 50, 100, 200, 500, 1000 или более аминокислотным остаткам.

Двунаправленная конструкция нуклеиновой кислоты

В данном документе описаны двунаправленные конструкции нуклеиновых кислот, которые облегчают улучшенную вставку, например, улучшают продуктивную вставку и экспрессию представляющего интерес гена. Вкратце, различные двунаправленные конструкции, описанные в данном документе, содержат по меньшей мере два сегмента нуклеиновой кислоты, причем один сегмент (первый сегмент) содержит кодирующую последовательность, которая кодирует представляющий интерес агент, например, гетерологичный ген (кодирующая последовательность может называться в данном документе «трансгеном» или первым трансгеном), тогда как другой сегмент (второй сегмент) содержит последовательность, в которой комплементарная последовательность кодирует представляющий интерес агент, например, гетерологичный ген, или второй трансген. Агент может представлять собой терапевтический агент, такой как полипептид, функциональная РНК, мРНК и т. п. Трансген может кодировать агент, такой как полипептид, функциональная РНК, мРНК или фактор транскрипции. В некоторых вариантах осуществления кодирующая последовательность кодирует терапевтический агент, такой как полипептид или функциональная РНК. По меньшей мере два сегмента могут кодировать идентичные или разные полипептиды или идентичные или разные агенты. В некоторых вариантах осуществления двунаправленные конструкции, описанные в данном документе, содержат по меньшей мере два сегмента нуклеиновой кислоты, при этом один сегмент (первый сегмент) содержит последовательность, которая кодирует представляющий интерес полипептид, тогда как другой сегмент (второй сегмент) содержит последовательность, причем комплементарная последовательность кодирует представляющий интерес полипептид.

В одном варианте осуществления двунаправленные конструкции содержат по меньшей мере два сегмента нуклеиновой кислоты в *цис*-ориентации, при этом один сегмент (первый сегмент) содержит кодирующую последовательность (иногда взаимозаменяемо называемую в данном документе «трансгеном»), а другой сегмент (второй сегмент) содержит последовательность, в которой комплементарная последовательность кодирует фактор IX. Первый трансген и второй трансген могут быть одинаковыми или разными. Двунаправленные конструкции могут содержать по меньшей мере два сегмента нуклеиновой кислоты в *цис*-ориентации, причем один сегмент (первый сегмент) содержит кодирующую последовательность, которая кодирует гетерологичный ген в одной ориентации, тогда как другой сегмент (второй сегмент) содержит последовательность, в которой его комплементарная последовательность кодирует гетерологичный ген в другой ориентации. Это означает, что первый сегмент является комплементарным ко второму сегменту (не обязательно идеально комплементарным); комплементарная последовательность второго сегмента является обратной комплементарной последовательностью первого сегмента (не обязательно идеальной обратной комплементарной последовательностью, хотя обе кодируют один и тот же гетерологичный белок). Двунаправленная конструкция может содержать первую

кодирующую последовательность, которая кодирует гетерологичный ген, связанный с акцептором сплайсинга, и вторую кодирующую последовательность, в которой комплементарная последовательность кодирует гетерологичный ген в другой ориентации, также связанный с акцептором сплайсинга.

В контексте данного документа такую конструкцию иногда называют «донорной конструкцией/матрицей». В некоторых вариантах осуществления конструкция представляет собой ДНК-конструкцию. Способы разработки и осуществления различных функциональных/структурных модификаций донорных конструкций известны в данной области техники. В некоторых вариантах осуществления конструкция может содержать одно или более из хвостовой последовательности полиаденилирования, последовательности сигнала полиаденилирования, сайта акцептора сплайсинга или селективного маркера. В некоторых вариантах осуществления хвостовая последовательность полиаденилирования кодируется, например, в виде участка «поли-А» в 3' конце кодирующей последовательности.

При использовании в сочетании с системой редактирования генов, описанной в данном документе, двунаправленность конструкций нуклеиновых кислот позволяет осуществлять вставку конструкции в любом направлении (не ограничивается вставкой в одном направлении) в целевой сайт вставки, обеспечивая экспрессию представляющего интерес полипептида из а) кодирующей последовательности одного сегмента (например, левого сегмента, кодирующего «F9 человека», верхняя левая конструкция ssAAV на Фиг. 1) или б) комплементарной последовательности другого сегмента (например, комплементарной последовательности правого сегмента, кодирующей «F9 человека», как проиллюстрировано в перевернутом виде в верхней левой конструкции ssAAV на Фиг. 1), тем самым улучшая эффективность вставки и экспрессии, как проиллюстрировано в данном документе. Нацеленное расщепление системой редактирования генов может облегчать интеграцию конструкции и/или экспрессию трансгена. При практической реализации настоящего изобретения можно использовать различные известные системы редактирования генов, включая, например, сайт-специфические системы расщепления ДНК, включая систему CRISPR/Cas; систему цинк-пальцевой нуклеазы (ZFN); систему эффекторной нуклеазы, подобной активатору транскрипции (TALEN).

В некоторых вариантах осуществления двунаправленная конструкция нуклеиновой кислоты не содержит промотор, который управляет экспрессией агента или полипептида. Например, экспрессия полипептида находится под управлением промотора клетки-хозяина (например, промотора эндогенного альбумина, когда трансген интегрирован в локус альбумина клетки-хозяина).

В некоторых вариантах осуществления двунаправленная конструкция нуклеиновой кислоты содержит первый сегмент, содержащий кодирующую последовательность полипептида, и второй сегмент, содержащий обратную комплементарную последовательность кодирующей последовательности полипептида. То же самое справедливо и для неполипептидных агентов. Таким образом, кодирующая

последовательность в первом сегменте способна экспрессировать полипептид, в то время как комплементарная последовательность обратной комплементарной последовательности во втором сегменте также способна экспрессировать полипептид. В контексте данного документа термин «кодирующая последовательность» при описании второго сегмента, содержащего обратную комплементарную последовательность, относится к комплементарной (кодирующей) цепи второго сегмента (т. е. комплементарной кодирующей последовательности обратной комплементарной последовательности во втором сегменте).

В некоторых вариантах осуществления кодирующая последовательность, которая кодирует полипептид А в первом сегменте, является менее чем на 100% комплементарной обратной комплементарной последовательности кодирующей последовательности, которая также кодирует полипептид А. Это означает, что в некоторых вариантах осуществления первый сегмент содержит кодирующую последовательность (1) полипептида А, а второй сегмент представляет собой обратную комплементарную последовательность кодирующей последовательности (2) полипептида А, причем кодирующая последовательность (1) не идентична с кодирующей последовательностью (2). Например, в кодирующей последовательности (1) и/или кодирующей последовательности (2), которая кодирует полипептид А, могут использоваться разные кодоны. В некоторых вариантах осуществления одна или обе последовательности могут быть кодон-оптимизированными так, чтобы кодирующая последовательность (1) и обратная комплементарная последовательность кодирующей последовательности (2) обладали 100% или менее чем на 100% комплементарности. В некоторых вариантах осуществления кодирующая последовательность второго сегмента кодирует полипептид с использованием одного или более альтернативных кодонов для одной или более аминокислот того же самого полипептида, кодируемого кодирующей последовательностью в первом сегменте. В контексте данного документа термин «альтернативный кодон» относится к вариациям в частоте использования кодонов для заданной аминокислоты и может быть или не быть предпочтительным или оптимизированным кодоном (кодон-оптимизированным) для заданной экспрессионной системы. Предпочтительное использование кодонов или кодоны, которые наиболее допустимы в заданной экспрессионной системе, известны в данной области техники.

В некоторых вариантах осуществления второй сегмент содержит обратную комплементарную последовательность, для которой характерно использование кодонов, отличное от кодирующей последовательности первого сегмента, чтобы снизить вероятность образования шпильки. Такая обратная комплементарная последовательность спаривается по основаниям не со всеми нуклеотидами кодирующей последовательности в первом сегменте, но, необязательно, она кодирует тот же самый полипептид. В таких случаях кодирующая последовательность, например, для полипептида А, первого сегмента может быть гомологичной, но не идентичной кодирующей последовательности, например, для полипептида А, второй половины двунаправленной конструкции. В

некоторых вариантах осуществления второй сегмент содержит обратную комплементарную последовательность, которая не является по существу комплементарной (например, не более чем на 70% комплементарной) кодирующей последовательности в первом сегменте. В некоторых вариантах осуществления второй сегмент содержит обратную комплементарную последовательность, которая не является существенно комплементарной (например, по меньшей мере на 90% комплементарной) кодирующей последовательности в первом сегменте. В некоторых вариантах осуществления второй сегмент содержит обратную комплементарную последовательность, имеющую по меньшей мере около 30%, около 35%, около 40%, около 45%, около 50%, около 55%, около 60%, около 65%, около 70%, около 75%, около 80%, около 85%, около 90%, около 95%, около 97% или около 99% комплементарности с кодирующей последовательностью в первом сегменте.

В некоторых вариантах осуществления второй сегмент содержит обратную комплементарную последовательность, имеющую 100% комплементарности с кодирующей последовательностью в первом сегменте. Это означает, что последовательность второго сегмента является идеальной обратной комплементарной последовательностью кодирующей последовательности в первом сегменте. В качестве примера, первый сегмент содержит гипотетическую последовательность 5' CTGGACCGA 3' (SEQ ID NO: 500), а второй сегмент содержит обратную комплементарную последовательность SEQ ID NO: 1, т. е. 5' TCGGTCCAG 3' (SEQ ID NO: 502).

В некоторых вариантах осуществления двунаправленная конструкция нуклеиновой кислоты содержит первый сегмент, содержащий кодирующую последовательность полипептида или агента (например, первого полипептида), и второй сегмент, содержащий обратную комплементарную последовательность кодирующей последовательности полипептида или агента (например, второго полипептида). В некоторых вариантах осуществления первый полипептид и второй полипептид являются одинаковыми, как описано выше. В некоторых вариантах осуществления первый терапевтический агент и второй терапевтический агент являются одинаковыми, как описано выше. В некоторых вариантах осуществления первый полипептид и второй полипептид являются разными. В некоторых вариантах осуществления первый терапевтический агент и второй терапевтический агент являются разными. Например, первый полипептид представляет собой полипептид А, а второй полипептид представляет собой полипептид В. В качестве дополнительного примера первый полипептид представляет собой полипептид А, а второй полипептид представляет собой вариант (например, фрагмент (такой как функциональный фрагмент), мутант, слияние (включая добавление хотя бы одной аминокислоты в конце полипептида) или их комбинации) полипептида А. Кодирующая последовательность, которая кодирует полипептид, может необязательно содержать одну или более дополнительных последовательностей, таких как последовательности, кодирующие amino- или карбоки-концевые аминокислотные последовательности, такие как сигнальная последовательность, последовательность метки (например, HiBit) или

гетерологичная функциональная последовательность (например, последовательность ядерной локализации (СЯЛ) или саморасщепляющаяся последовательность), связанных с полипептидом. Кодированная последовательность, которая кодирует полипептид, может, необязательно, содержать последовательности, кодирующие одну или более аминоконцевых сигнальных пептидных последовательностей. Каждая из этих дополнительных последовательностей может быть одинаковой или разной в первом и втором сегментах конструкции.

Описанную в данном документе двунаправленную конструкцию можно использовать для экспрессии любого полипептида в соответствии с описанными в данном документе способами. В некоторых вариантах осуществления полипептид представляет собой секреторируемый полипептид. В некоторых вариантах осуществления полипептид является таким, чья функция обычно осуществляется (например, который является функционально активным) в виде секреторируемого полипептида. В контексте данного документа «секреторируемый полипептид» относится к белку, который секреторируется клеткой и/или является функционально активным в виде растворимого внеклеточного белка.

В некоторых вариантах осуществления полипептид представляет собой внутриклеточный полипептид. В некоторых вариантах осуществления полипептид является таким, чья функция обычно осуществляется (например, который является функционально активным) внутри клетки. В контексте данного документа «внутриклеточный полипептид» относится к белку, который не секреторируется клеткой, включая растворимые цитозольные полипептиды.

В некоторых вариантах осуществления полипептид представляет собой полипептид дикого типа.

В некоторых вариантах осуществления полипептид представляет собой печеночный белок или его вариант. В контексте данного документа «печеночный белок» представляет собой белок, который, например, эндогенно вырабатывается в печени и/или является функционально активным в печени. В некоторых вариантах осуществления печеночный белок представляет собой циркулирующий белок, вырабатываемый печенью, или его вариант. В некоторых вариантах осуществления печеночный белок представляет собой белок, который является функционально активным в печени, или его вариант. В некоторых вариантах осуществления печеночный белок демонстрирует повышенную экспрессию в печени по сравнению с одним или более типами ткани. В некоторых вариантах осуществления полипептид представляет собой не печеночный белок. В некоторых вариантах осуществления полипептид включает, но не ограничивается этим, фактор IX и его варианты.

В некоторых вариантах осуществления двунаправленная конструкция нуклеиновой кислоты является линейной. Например, первый и второй сегменты соединены линейным образом посредством линкерной последовательности. В некоторых вариантах осуществления 5' конец второго сегмента, который содержит обратную комплементарную

последовательность, связан с 3' концом первого сегмента. В некоторых вариантах осуществления 5' конец первого сегмента связан с 3' концом второго сегмента, который содержит обратную комплементарную последовательность. В некоторых вариантах осуществления длина линкерной последовательности составляет около 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 500, 1000, 1500, 2000 или более нуклеотидов. Как понятно специалистам в данной области техники, между первым и вторым сегментами могут быть вставлены другие структурные элементы в дополнение к линкерной последовательности или вместо нее.

Описанные в данном документе конструкции могут быть модифицированы для включения любого подходящего структурного элемента, необходимого для любого конкретного применения и/или который обеспечивает одну или более необходимых функций. В некоторых вариантах осуществления описанная в данном документе двунаправленная конструкция нуклеиновой кислоты не содержит плечи гомологии. В некоторых вариантах осуществления описанная в данном документе двунаправленная конструкция нуклеиновой кислоты представляет собой независимую от гомологии донорную конструкцию. В некоторых вариантах осуществления, частично благодаря двунаправленной функции конструкции нуклеиновой кислоты, двунаправленную конструкцию можно вставлять в геномный локус в любом направлении (ориентации), как описано в данном документе, чтобы обеспечить эффективную вставку и/или экспрессию представляющего интерес полипептида. В некоторых вариантах осуществления двунаправленная конструкция нуклеиновой кислоты содержит первый сегмент и второй сегмент, каждый из которых имеет акцептор сплайсинга, расположенный выше транскена. В некоторых вариантах осуществления акцептор сплайсинга совместим с донорной последовательностью сплайсинга безопасного приемного сайта клетки-хозяина, например, с донором сплайсинга интрона 1 гена альбумина человека.

В некоторых вариантах осуществления описанная в данном документе композиция содержит один или более сайтов внутренней посадки рибосомы (IRES). Впервые обнаруженный как элемент РНК пикорнавируса, IRES играет важную роль в инициации синтеза белка в отсутствие структуры 5'-кэпа. IRES может действовать как единственный сайт связывания рибосомы или может служить одним из нескольких сайтов связывания рибосомы полинуклеотидов. Конструкции, содержащие более одного функционального сайта связывания рибосомы, могут кодировать несколько пептидов или полипептидов, которые независимо транслируются рибосомами («мультицистронные молекулы нуклеиновых кислот»). В альтернативном варианте конструкции могут содержать IRES с целью экспрессии гетерологичного белка, который не слит с эндогенным полипептидом (т. е. сигнальным пептидом альбумина). Примеры применимых последовательностей IRES, включают, без ограничения, последовательности из пикорнавирусов (например, FMDV), пестивирусов (CFFV), полиовирусов (PV), вирусов энцефаломиокардита (ECMV), вирусов болезни рук, ног и рта (FMDV), вирусов гепатита С (HCV), вирусов классической

чумы свиней (CSFV), вируса лейкоза мышей (MLV), вирусов иммунодефицита обезьян (SIV) или вирусов паралича сверчка (CrPV).

В некоторых вариантах осуществления конструкция нуклеиновой кислоты содержит последовательность, кодирующую саморасщепляющийся пептид, такую как последовательность 2A или 2A-подобная последовательность. В некоторых вариантах осуществления саморасщепляющийся пептид расположен выше представляющего интерес полипептида. В одном варианте осуществления, последовательность, кодирующую пептид 2A, можно использовать для разделения кодирующей области двух или более представляющих интерес полипептидов. В другом варианте осуществления, последовательность можно использовать для разделения кодирующей последовательности из конструкции и кодирующей последовательности из эндогенного локуса (т. е. эндогенной сигнальной последовательности альбумина). В качестве неограничивающего примера последовательность, кодирующая пептид 2A, может быть расположена между областью А и областью В (А-2А-В). Наличие пептида 2A приведет к расщеплению одного длинного белка на белок А, белок В и пептид 2A. Белок А и белок В могут представлять собой одинаковые или разные представляющие интерес полипептиды.

В некоторых вариантах осуществления один или оба из первого и второго сегментов содержат хвостовую последовательность полиаденилирования и/или сигнальную последовательность полиаденилирования ниже открытой рамки считывания. В некоторых вариантах осуществления хвостовая последовательность полиаденилирования кодируется, например, в виде участка «поли-А» в 3' конце первого и/или второго сегмента. В некоторых вариантах осуществления хвостовую последовательность полиаденилирования обеспечивают котранскрипционно как результат сигнальной последовательности полиаденилирования, которая кодируется в 3' конце первого и/или второго сегмента или вблизи него. В некоторых вариантах осуществления поли-А-хвост содержит по меньшей мере 20, 30, 40, 50, 60, 70, 80, 90 или 100 оснований аденина, необязательно, до 300 оснований аденина. В некоторых вариантах осуществления поли-А-хвост содержит по меньшей мере 95, 96, 97, 98, 99 или 100 нуклеотидов аденина. Способы конструирования подходящей хвостовой последовательности полиаденилирования и/или последовательности сигнала полиаденилирования хорошо известны в данной области техники. Подходящие акцепторные последовательности сплайсинга описаны и проиллюстрированы в данном документе, включая мышинный альбумин и человеческие акцепторные сайты FIX. В некоторых вариантах осуществления сигнальную последовательность полиаденилирования AAUAAA (SEQ ID NO: 800) обычно используют в системах млекопитающих, хотя были идентифицированы такие варианты, как UAUAAA (SEQ ID NO: 801) или AU/GUAAA (SEQ ID NO: 802). Смотрите, например, NJ Proudfoot, *Genes & Dev.* 25(17):1770-82, 2011. В некоторых вариантах осуществления включена хвостовая последовательность поли-А.

В некоторых вариантах осуществления описанные в данном документе

конструкции могут представлять собой ДНК или РНК, быть одноцепочечными, двухцепочечными или частично одно- и частично двухцепочечными, и могут быть внесены в клетку-хозяина в линейной или кольцевой (например, мини-кольцевой) форме. Смотрите, например, патентные публикации США №№ 2010/0047805, 2011/0281361, 2011/0207221. При введении в линейной форме концы донорной последовательности можно защитить (например, от экзонуклеолитической дегградации) способами, известными специалистам в данной области техники. Например, к 3' концу линейной молекулы добавляют один или более дидезоксинуклеотидных остатков и/или к одному или обоим концам лигируют самокомплементарные олигонуклеотиды. Смотрите, например, Chang et al. (1987) Proc. Natl. Acad. Sci. USA 84:4959-4963; Nehls et al. (1996) Science 272:886-889. Дополнительные способы защиты экзогенных полинуклеотидов от дегградации включают, но не ограничиваются этим, добавление концевой(ых) аминогруппы (аминогрупп) и использование модифицированных межнуклеотидных связей, таких как, например, тиофосфаты, фосфорамидаты и остатки O-метилрибозы или дезоксирибозы.

В некоторых вариантах осуществления конструкция можно вставить так, чтобы ее экспрессия находилась под управлением эндогенного промотора в сайте вставки (например, эндогенного промотора альбумина, когда донор интегрирован в локус альбумина клетки-хозяина). В таких случаях в трансгене могут отсутствовать регуляторные элементы (например, промотор и/или энхансер), которые управляют его экспрессией (например, конструкция без промотора). Тем не менее, будет очевидно, что в других случаях конструкция может содержать промотор и/или энхансер, например конститутивный промотор или индуцибельный или тканеспецифический (например, специфический для печени или тромбоцитов) промотор, который управляет экспрессией функционального белка после интеграции. Конструкция может содержать последовательность, кодирующую гетерологичный белок, расположенный ниже сигнальной последовательности, кодирующей сигнальный пептид, и функционально связанную с ней.

В некоторых вариантах осуществления конструкция нуклеиновой кислоты работает при независимой от гомологии вставке нуклеиновой кислоты, которая кодирует гетерологичный полипептид. В некоторых вариантах осуществления конструкция нуклеиновой кислоты работает в неделящихся клетках, например, в клетках, в которых НГСК, а не ГР является основным механизмом, посредством которого происходит репарация двухцепочечных разрывов ДНК. Нуклеиновая кислота может представлять собой независимую от гомологии донорную конструкцию. Например, конструкции могут представлять собой одноцепочечную или двухцепочечную ДНК. В некоторых вариантах осуществления нуклеиновая кислота может быть модифицирована (например, с использованием аналогов нуклеозидов), как описано в данном документе.

В некоторых вариантах осуществления описанные в данном документе конструкции содержат акцепторный сайт сплайсинга в одном или обоих концах

конструкции, например, 5' относительно открытой рамки считывания в первом и/или втором сегментах, или 5' относительно последовательностей одного или обоих трансенов. В некоторых вариантах осуществления акцепторный сайт сплайсинга содержит NAG. В дополнительных вариантах осуществления акцепторный сайт сплайсинга состоит из NAG. В некоторых вариантах осуществления акцептор сплайсинга представляет собой акцептор сплайсинга альбумина, например, акцептор сплайсинга альбумина, используемый при сплайсинге экзонов 1 и 2 альбумина. В некоторых вариантах осуществления акцептор сплайсинга получен из гена альбумина человека. В некоторых вариантах осуществления акцептор сплайсинга получен из гена альбумина мыши. В некоторых вариантах осуществления акцептор сплайсинга представляет собой акцептор сплайсинга F9 (или «FIX»), например, акцептор сплайсинга F9, используемый при сплайсинге экзонов 1 и 2 F9. В некоторых вариантах осуществления акцептор сплайсинга получен из гена F9 человека. В некоторых вариантах осуществления акцептор сплайсинга получен из гена F9 мыши. Дополнительные подходящие акцепторные сайты сплайсинга, используемые у эукариот, включая искусственные акцепторы сплайсинга, известны и могут быть получены из уровня техники. Смотрите, например, Shapiro, et al., 1987, *Nucleic Acids Res.*, 15, 7155-7174, Bursset, et al., 2001, *Nucleic Acids Res.*, 29, 255-259.

В некоторых вариантах осуществления описанные в данном документе конструкции могут быть модифицированы в одном или обоих концах, чтобы включать, при необходимости, один или более подходящих структурных элементов и/или для обеспечения одного или более функциональных преимуществ. Например, структурные модификации могут варьироваться в зависимости от способа(ов), используемого(ых) для доставки описанных в данном документе конструкций в клетку-хозяина, например, доставки на основе вирусного вектора или упаковки в липидные наночастицы для доставки. Такие модификации включают, без ограничения, например, концевые структуры, такие как инвертированные концевые повторы (ИКП), шпильки, петли и другие структуры, такие как тороид. В некоторых вариантах осуществления описанные в данном документе конструкции содержат один, два или три ИКП. В некоторых вариантах осуществления описанные в данном документе конструкции содержат не более двух ИКП. В данной области техники известны различные способы структурных модификаций.

В некоторых вариантах осуществления один или оба конца конструкции могут быть защищены (например, от экзонуклеолитической деградации) способами, известными специалистам в данной области техники. Например, к 3' концу линейной молекулы добавляют один или более дидезоксинуклеотидных остатков и/или к одному или обоим концам лигируют самокомплементарные олигонуклеотиды. Смотрите, например, Chang et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4959-4963; Nehls et al. (1996) *Science* 272:886-889. Дополнительные способы защиты конструкций от деградации включают, но не ограничиваются этим, добавление концевой(ых) аминогруппы (аминогрупп) и использование модифицированных межнуклеотидных связей, таких как, например, тиофосфаты, фосфорамидаты и остатки О-метилрибозы или дезоксирибозы.

В некоторых вариантах осуществления описанные в данном документе конструкции можно вносить в клетку как часть вектора, имеющего дополнительные последовательности, такие как, например, точки начала репликации, промоторы и гены, кодирующие устойчивость к антибиотикам. В конструкции могут отсутствовать вирусные элементы. В некоторых вариантах осуществления конструкции можно вносить в виде оголенной нуклеиновой кислоты, в виде комплекса нуклеиновой кислоты с агентом, таким как липосома, полимер или поллоксамер, или можно доставлять с помощью вирусных векторов (например, аденовируса, AAV, вируса герпеса, ретровируса, лентивируса).

В некоторых вариантах осуществления изобретения, хотя это и не требуется для экспрессии, описанные в данном документе конструкции могут также содержать транскрипционные или трансляционные регуляторные последовательности, например, промоторы, энхансеры, инсуляторы, сайты внутренней посадки рибосомы, последовательности, кодирующие пептиды, и/или сигналы полиаденилирования.

В некоторых вариантах осуществления конструкции, содержащие кодирующую последовательность представляющего интерес полипептида, могут содержать одну или более из следующих модификаций: оптимизацию кодонов (например, в отношении кодонов человека) и/или добавление одного или более сайтов гликозилирования. Смотрите, например, McIntosh et al. (2013) *Blood* (17):3335-44.

Система редактирования генов

При практической реализации настоящего изобретения можно использовать различные известные системы редактирования генов, включая, например, систему CRISPR/Cas; систему цинк-пальцевой нуклеазы (ZFN); и систему эффекторной нуклеазы, подобной активатору транскрипции (TALEN). В общем случае эти способы могут включать использование сконструированных систем расщепления для индукции двухцепочечного разрыва (ДЦР) или ника (например, одноцепочечного разрыва или ОЦР) в целевой последовательности ДНК. Расщепление или создание одноцепочечного разрыва может происходить за счет применения специфических нуклеаз, таких как сконструированные ZFN, TALEN, или применения системы CRISPR/Cas со сконструированной гидовой РНК для управления специфическим расщеплением или созданием одноцепочечного разрыва в целевой последовательности ДНК. Кроме того, уже были разработаны нацеленные нуклеазы, и дополнительные нуклеазы находятся на стадии разработки, например, основе системы Argonaute (например, из *T. thermophilus*, известного как «TtAgo», смотрите Swartz et al (2014) *Nature* 507(7491): 258-261), которые также могут иметь потенциал для применения в редактировании генома и генной терапии.

В некоторых вариантах осуществления систему CRISPR/Cas можно использовать для создания сайта вставки в необходимом локусе в геноме хозяина, в который можно вставлять описанную в данном документе двунаправленную конструкцию для экспрессии одного или более представляющих интерес полипептидов. Способы разработки подходящих гидовых РНК, нацеленных на любой необходимый локус в геноме хозяина для вставки, хорошо известны в данной области техники. Двунаправленная конструкция,

содержащая трансген, может быть гетерологичной по отношению к сайту ее вставки, например, в случае вставки гетерологичного трансгена в «безопасный приемный» локус. Двухнаправленная конструкция, содержащая трансген, может не быть гетерологичной по отношению к сайту ее вставки, например, в случае вставки трансгена дикого типа в его эндогенный локус.

«Безопасный приемный» локус представляет собой локус в геноме, в который экзогенная нуклеиновая кислота может быть вставлена без существенного вредного воздействия на клетку-хозяина, например, гепатоцит, например, не обуславливая апоптоз, некроз и/или старение, или не обуславливая более 5%, 10%, 15%, 20%, 25%, 30% или 40% апоптоза, некроза и/или старения, по сравнению с контрольной клеткой. Смотрите, например, Hsin et al., “Hepatocyte death in liver inflammation, fibrosis, and tumorigenesis,” 2017. В некоторых вариантах осуществления безопасный приемный локус обеспечивает экспрессию экзогенной нуклеиновой кислоты (например, экзогенного гена) без существенного вредного воздействия на клетку-хозяина или популяцию клеток, таких как гепатоциты или клетки печени, например, не обуславливая апоптоз, некроз и/или старение, или не обуславливая более 5%, 10%, 15%, 20%, 25%, 30% или 40% апоптоза, некроза и/или старения, по сравнению с контрольной популяцией клеток. Безопасный приемный локус может находиться в гене альбумина, таком как ген альбумина человека. Безопасный приемный локус может находиться в области интрона 1 альбумина, например, интрона 1 альбумина человека. Безопасный приемный локус может представлять собой безопасный приемный локус человека, например, в случае ткани печени или клетки-хозяина гепатоцита. Неограничивающие примеры безопасных приемных локусов, на которые нацелены нуклеазы, включают CCR5, HPRT, AAVS1, Rosa, альбумин, AAVS1 (PPP1 R12C), Angptl5, ApoC3, ASGR2, FIX (F9), G6PC, Gys2, HGD, Lp(a), Pcsk9, SERPINA1, TF и TTR. Смотрите, например, патенты США №№ 7951925 и 8110379; публикации США №№ 2008/0159996; 2010/00218264; 2012/0017290; 2011/0265198; 2013/0137104; 2013/0122591; 2013/0177983; 2013/0177960; и WO 2017093804. Как проиллюстрировано в данном документе, в некоторых вариантах осуществления можно конструировать гидовые РНК для нацеливания на локус альбумина человека или мыши (например, интрон 1). Примеры гидовых РНК, проиллюстрированные в данном документе, приведены в таблицах 5-10. Следует понимать, что для вставки двухнаправленной конструкции, содержащий трансген, можно проводить нацеливание на любой другой локус в соответствии с представленными способами.

В некоторых вариантах осуществления гетерологичный ген может быть вставлен в безопасный приемный локус и использовать эндогенную сигнальную последовательность безопасного приемного локуса, например, сигнальную последовательность альбумина, кодируемую экзон 1. Например, кодирующая последовательность может быть вставлена в интрон 1 альбумина человека так, чтобы она была расположена ниже и слита с сигнальной последовательностью экзона 1 альбумина человека.

В некоторых вариантах осуществления ген может содержать свою собственную

сигнальную последовательность, может быть вставлен в безопасный приемный локус и может дополнительно использовать эндогенную сигнальную последовательность безопасного приемного локуса. Например, кодирующая последовательность, содержащая нативную сигнальную последовательность, может быть вставлена в интрон 1 альбумина человека так, чтобы она была расположена ниже и слита с сигнальной последовательностью альбумина человека, кодируемой экзоном 1.

В некоторых вариантах осуществления ген может содержать свою собственную сигнальную последовательность и сайт внутренней посадки рибосомы (IRES), может быть вставлен в безопасный приемный локус и может дополнительно использовать эндогенную сигнальную последовательность безопасного приемного локуса. Например, кодирующая последовательность, содержащая нативную сигнальную последовательность и последовательность IRES, может быть вставлена в интрон 1 альбумина человека так, чтобы она была расположена ниже и слита с сигнальной последовательностью альбумина человека, кодируемой экзоном 1.

В некоторых вариантах осуществления ген может содержать свою собственную сигнальную последовательность и IRES, может быть вставлен в безопасный приемный локус и не использовать эндогенную сигнальную последовательность безопасного приемного локуса. Например, кодирующая последовательность, содержащая нативную сигнальную последовательность и последовательность IRES, может быть вставлена в интрон 1 альбумина человека так, чтобы она не была слита с сигнальной последовательностью альбумина человека, кодируемой экзоном 1. В этих вариантах осуществления белок транслируется из сайта IRES и не является химерным (например, сигнальный пептид альбумина, слитый с гетерологичным белком), преимущество чего может состоять в отсутствующей или низкой иммуногенности. В некоторых вариантах осуществления белок не секретируется и/или не транспортируется внеклеточно.

В некоторых вариантах осуществления ген может быть вставлен в безопасный приемный локус и может содержать IRES и не использовать какую-либо сигнальную последовательность. Например, кодирующая последовательность, содержащая последовательность IRES и не содержащая нативную сигнальную последовательность, может быть вставлена в интрон 1 альбумина человека так, чтобы она не была слита с сигнальной последовательностью альбумина человека, кодируемой экзоном 1. В некоторых вариантах осуществления белки транслируются из сайта IRES без какой-либо сигнальной последовательности. В некоторых вариантах осуществления белок не секретируется и/или не транспортируется внеклеточно.

Также следует понимать, что гидовая РНК для Cas-нуклеазы, такой как Cas9-нуклеаза, которую можно использовать в представленных способах, может включать любые из различных известных вариаций и модификаций (например, химических модификаций), включая наличие одного или более не встречающихся в природе и/или встречающихся в природе компонентов или одной или более конфигураций, которые используются вместо канонических остатков А, G, C и U или в дополнение к ним.

Например, каждая из направляющих последовательностей, проиллюстрированных в данном документе (таблицы 5-10), может дополнительно содержать дополнительные нуклеотиды с образованием сгРНК, гидовой РНК и/или огРНК, например, из системы CRISPR/Cas SpyCas9. Например, каждая из направляющих последовательностей, проиллюстрированных в данном документе (таблицы 5-10), может дополнительно содержать дополнительные нуклеотиды с образованием сгРНК или огРНК со следующей типовой нуклеотидной последовательностью после направляющей последовательности в 3'-конце: GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 200) в ориентации от 5' к 3'. В случае огРНК направляющие последовательности, такие как направляющие последовательности, перечисленные в таблицах 5-10, могут дополнительно содержать дополнительные нуклеотиды с образованием огРНК, например, со следующей типовой нуклеотидной последовательностью (направляющая последовательность SpyCas9) после 3'-конца направляющей последовательности: GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGA AAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 201) или GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGA AAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 202) в ориентации от 5' к 3'.

Гидовая РНК может, необязательно, содержать trРНК. В каждом описанном в данном документе варианте осуществления композиции и способа сгРНК и trРНК могут быть связаны в виде одной РНК (огРНК) или могут находиться в отдельных РНК (дгРНК). В контексте огРНК компоненты сгРНК и trРНК могут быть связаны ковалентно, например, посредством фосфодиэфирной связи или другой ковалентной связи. В некоторых вариантах осуществления огРНК содержит одну или более связей между нуклеотидами, которые не являются фосфодиэфирными связями. В каждом из описанных в данном документе вариантов осуществления композиции, применения и способа гидовая РНК может содержать две молекулы РНК в виде «двойной гидовой РНК» или «дгРНК». дгРНК содержит первую молекулу РНК, содержащую сгРНК, содержащую, например, направляющую последовательность, приведенную в любой из таблиц 5-10, и вторую молекулу РНК, содержащую trРНК. Первая и вторая молекулы РНК могут не быть ковалентно связанными, но могут образовывать двойную спираль РНК посредством спаривания оснований между частями сгРНК и trРНК.

В некоторых вариантах осуществления описанные в данном документе гидовые РНК связываются с областью, расположенной выше мотива, примыкающего к пропоспейсеру (PAM, англ. «*protospacer adjacent motif*»). Как будет понятно специалистам в данной области техники, последовательность PAM находится в цепи, противоположной цепи, которая содержит целевую последовательность. Это означает, что последовательность PAM находится в цепи, комплементарной целевой цепи (цепи, которая содержит целевую последовательность, с которой связывается гидовая РНК). В некоторых вариантах осуществления PAM выбран из группы, состоящей из NGG, NNGRRT, NNGRR(N), NNAGAAW, NNNNG(A/C)TT и NNNNRYAC. В некоторых

вариантах осуществления РАР представляет собой NGG.

В некоторых вариантах осуществления предложенные в данном документе последовательности гидовой РНК комплементарны последовательности, смежной с последовательностью РАР.

В некоторых вариантах осуществления последовательность гидовой РНК содержит последовательность, комплементарную последовательности в геномной области, выбранной из таблиц в данном документе, в соответствии с координатами в референсном геноме человека hg38. В некоторых вариантах осуществления последовательность гидовой РНК содержит последовательность, комплементарную последовательности, которая содержит 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 или 25 последовательных нуклеотидов из геномной области, выбранной из таблиц 5-10. В некоторых вариантах осуществления последовательность гидовой РНК содержит последовательность, комплементарную последовательности, которая содержит 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 или 25 последовательных нуклеотидов, охватывающих геномную область, выбранную из таблиц 5-10.

Описанные в данном документе гидовые РНК опосредуют целенаправленное разрезание, приводящее к двухцепочечному разрыву (ДЦР). Описанные в данном документе гидовые РНК опосредуют целенаправленное разрезание, приводящее к одноцепочечному разрыву (ОЦР или нику).

Способы применения различных РНК-направляемых ДНК-связывающих агентов, например, нуклеазы, такой как Cas-нуклеаза, например, Cas9, также хорошо известны в данной области техники. Хотя применение двунаправленной нуклеиновой кислоты с системой CRISPR/Cas проиллюстрировано в данном документе, следует понимать, что также можно использовать подходящие вариации этой системы. Следует понимать, что, в зависимости от контекста, РНК-направляемый ДНК-связывающий агент может быть предоставлен в виде нуклеиновой кислоты (например, ДНК или мРНК) или в виде белка. В некоторых вариантах осуществления представленный способ можно реализовать на практике в клетке-хозяине, которая уже содержит и/или экспрессирует РНК-направляемый ДНК-связывающий агент.

В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент, такой как Cas9-нуклеаза, обладает активностью кливазы, которую также можно назвать двухцепочечной эндонуклеазной активностью. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент, такой как Cas9-нуклеаза, обладает активностью никазы, которую также можно назвать одноцепочечной эндонуклеазной активностью. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент содержит собой Cas-нуклеазу. Примеры Cas-нуклеаз включают системы CRISPR типа II *S. pyogenes*, *S. aureus* и других прокариот (смотрите, например, перечень в следующем параграфе), а также их варианты или мутантные (например, сконструированные, не встречающиеся в природе, встречающиеся в природе или другие варианты) версии. Смотрите, например, US2016/0312198 A1; US 2016/0312199 A1.

Неограничивающие типовые виды, из которых можно получать Cas-нуклеазу, включают *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus sp.*, *Staphylococcus aureus*, *Listeria innocua*, *Lactobacillus gasseri*, *Francisella novicida*, *Wolinella succinogenes*, *Sutterella wadsworthensis*, *Gamma proteobacterium*, *Neisseria meningitidis*, *Campylobacter jejuni*, *Pasteurella multocida*, *Fibrobacter succinogenes*, *Rhodospirillum rubrum*, *Nocardiosis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycooides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Lactobacillus buchneri*, *Treponema denticola*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas sp.*, *Crocospaera watsonii*, *Cyanothece sp.*, *Microcystis aeruginosa*, *Synechococcus sp.*, *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor becscii*, *Candidatus Desulforudis*, *Clostridium botulinum*, *Clostridium difficile*, *Fingoldia magna*, *Natranaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochrochromatium vinosum*, *Marinobacter sp.*, *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium vestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc sp.*, *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira sp.*, *Lyngbya sp.*, *Microcoleus chthonoplastes*, *Oscillatoria sp.*, *Petrogona mobilis*, *Thermosiphon africanus*, *Streptococcus pasteurianus*, *Neisseria cinerea*, *Campylobacter lari*, *Parvibaculum lavamentivorans*, *Corynebacterium diphtheria*, *Acidaminococcus sp.*, бактерию *Lachnospiraceae ND2006* и *Acaaryochloris marina*.

В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cas9-нуклеазу из *Streptococcus pyogenes*. В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cas9-нуклеазу из *Streptococcus thermophilus*. В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cas9-нуклеазу из *Neisseria meningitidis*. В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cas9-нуклеазу из *Staphylococcus aureus*. В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cpf1-нуклеазу из *Francisella novicida*. В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cpf1-нуклеазу из *Acidaminococcus sp.* В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cpf1-нуклеазу из бактерии *Lachnospiraceae ND2006*. В дополнительных вариантах осуществления Cas-нуклеаза представляет собой Cpf1-нуклеазу из *Francisella tularensis*, *Lachnospiraceae bacterium*, *Butyrivibrio proteoclasticus*, *Peregrinibacteria bacterium*, *Parcubacteria bacterium*, *Smithella*, *Acidaminococcus*, *Candidatus Methanoplasma termitum*, *Eubacterium eligens*, *Moraxella bovoculi*, *Leptospira inadai*, *Porphyromonas crevioricanis*, *Prevotella disiens* или *Porphyromonas mascaea*. В определенных вариантах осуществления Cas-нуклеаза представляет собой Cpf1-нуклеазу из *Acidaminococcus* или *Lachnospiraceae*.

В некоторых вариантах осуществления гРНК вместе с РНК-направляемым ДНК-связывающим агентом называется рибонуклеопротеиновым комплексом (РНП). В

некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент представляет собой Cas-нуклеазу. В некоторых вариантах осуществления гРНК вместе с Cas-нуклеазой называется Cas-РНП. В некоторых вариантах осуществления РНП содержит компоненты типа I, типа II или типа III. В некоторых вариантах осуществления Cas-нуклеаза представляет собой белок Cas9 из системы CRISPR/Cas типа II. В некоторых вариантах осуществления гРНК вместе с Cas9 называется Cas9-РНП.

Cas9 дикого типа имеет два нуклеазных домена: RuvC и HNH. Домен RuvC расщепляет нецелевую цепь ДНК, а домен HNH расщепляет целевую цепь ДНК. В некоторых вариантах осуществления белок Cas9 содержит более одного домена RuvC и/или более одного домена HNH. В некоторых вариантах осуществления белок Cas9 представляет собой Cas9 дикого типа. В каждом из вариантов осуществления композиции, применения и способа Cas индуцирует двухцепочечный разрыв в целевой ДНК.

В некоторых вариантах осуществления используют химерные Cas-нуклеазы, в которых один домен или область белка заменены частью другого белка. В некоторых вариантах осуществления домен Cas-нуклеазы может быть заменен доменом другой нуклеазы, такой как FokI. В некоторых вариантах осуществления Cas-нуклеаза может представлять собой модифицированную нуклеазу.

В других вариантах осуществления Cas-нуклеаза может быть получена из системы CRISPR/Cas типа I. В некоторых вариантах осуществления Cas-нуклеаза может быть компонентом комплекса Cascade системы CRISPR/Cas типа I. В некоторых вариантах осуществления Cas-нуклеаза может представлять собой белок Cas3. В некоторых вариантах осуществления Cas-нуклеаза может быть получена из системы CRISPR/Cas типа III. В некоторых вариантах осуществления Cas-нуклеаза может обладать активностью расщепления РНК.

В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент обладает одноцепочечной никазной активностью, т. е. может надрезать одну цепь ДНК с образованием одноцепочечного разрыва, также известного как «ник». В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент содержит собой Cas-никазу. Никаза представляет собой фермент, который создает ник в дцДНК, т. е. надрезает одну цепь двойной спирали ДНК, но не другую. В некоторых вариантах осуществления Cas-никаза представляет собой версию Cas-нуклеазы (например, Cas-нуклеазы, обсуждаемой выше), в которой эндонуклеолитический активный сайт инактивирован, например, за счет одного или более изменений (например, точечных мутаций) в каталитическом домене. Смотрите, например, патент США № 8889356 в отношении обсуждения Cas-никаз и типовых изменений каталитических доменов. В некоторых вариантах осуществления Cas-никаза, такая как Cas9-никаза, имеет инактивированный домен RuvC или HNH.

В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент модифицирован так, чтобы содержать только один функциональный нуклеазный домен. Например, белок агента может быть модифицирован так, чтобы один из доменов

нуклеазы был мутирован или полностью или частично удален, чтобы снизить его активность расщепления нуклеиновой кислоты. В некоторых вариантах осуществления используют никазу, имеющую домен RuvC со сниженной активностью. В некоторых вариантах осуществления используют никазу, имеющую неактивный домен RuvC. В некоторых вариантах осуществления используют никазу, имеющую домен HNH со сниженной активностью. В некоторых вариантах осуществления используют никазу, имеющую неактивный домен HNH.

В некоторых вариантах осуществления консервативная аминокислота в нуклеазном домене белка Cas заменена для снижения или изменения нуклеазной активности. В некоторых вариантах осуществления Cas-нуклеаза может содержать аминокислотную замену в RuvC или RuvC-подобном нуклеазном домене. Примеры аминокислотных замен в RuvC или RuvC-подобном нуклеазном домене включают D10A (на основе белка Cas9 *S. pyogenes*). Смотрите, например, Zetsche et al. (2015) Cell Oct 22:163(3): 759-771. В некоторых вариантах осуществления Cas-нуклеаза может содержать аминокислотную замену в HNH или HNH-подобном нуклеазном домене. Примеры аминокислотных замен в HNH или HNH-подобном нуклеазном домене включают E762A, H840A, N863A, H983A и D986A (на основе белка Cas9 *S. pyogenes*). Смотрите, например, Zetsche et al. (2015). Дополнительные типовые аминокислотные замены включают D917A, E1006A и D1255A (на основе последовательности *Francisella novicida* U112 Cpf1 (FnCpf1) (UniProtKB - A0Q7Q2 (CPF1_FRATN))).

В некоторых вариантах осуществления никаза предоставлена в комбинации с парой гидовых РНК, комплементарных смысловой и антисмысловой цепям целевой последовательности, соответственно. В этом варианте осуществления гидовые РНК направляют никазу к целевой последовательности и вносят ДЦР, создавая ники на противоположных цепях целевой последовательности (т. е. создание двойного ника). В некоторых вариантах осуществления никазу используют с двумя отдельными гидовыми РНК, нацеленными на противоположные цепи ДНК, для создания двойного ника в целевой ДНК. В некоторых вариантах осуществления никазу используют с двумя отдельными гидовыми РНК, выбранными так, чтобы находиться в непосредственной близости, для создания двойного ника в целевой ДНК.

В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент содержит один или более гетерологичных функциональных доменов (например, представляет собой или содержит слитый полипептид).

В некоторых вариантах осуществления гетерологичный функциональный домен может облегчить перенос РНК-направляемого ДНК-связывающего агента в ядро клетки. Например, гетерологичный функциональный домен может представлять собой сигнал ядерной локализации (СЯЛ). В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент может быть слит с 1-10 СЯЛ. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент может быть слит с 1-5 СЯЛ. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент

может быть слит с одним СЯЛ. В случае применения одного СЯЛ, он может быть связан с N-концом или С-концом последовательности РНК-направляемого ДНК-связывающего агента. Он также может быть вставлен в последовательность РНК-направляемого ДНК-связывающего агента. В других вариантах осуществления РНК-направляемый ДНК-связывающий агент может быть слит с более чем одним СЯЛ. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент может быть слит с 2, 3, 4 или 5 СЯЛ. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент может быть слит с двумя СЯЛ. В определенных обстоятельствах два СЯЛ могут быть одинаковыми (например, два СЯЛ SV40) или разными. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент слит с двумя последовательностями СЯЛ SV40, присоединенными в карбокси-конце. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент может быть слит с двумя СЯЛ, одним, присоединенным в N-конце, и другим, присоединенным в С-конце. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент может быть слит с 3 СЯЛ. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент, может не быть слит с СЯЛ. В некоторых вариантах осуществления СЯЛ может быть однокомпонентной последовательностью, такой как, например, СЯЛ SV40, РКККРКV (SEQ ID NO: 600) или РКККRRV (SEQ ID NO: 601). В некоторых вариантах осуществления СЯЛ может быть двухкомпонентной последовательностью, такой как СЯЛ нуклеоплазмина, КРРААТККАGQAKKKK (SEQ ID NO: 602). В конкретном варианте осуществления одна РКККРКV (SEQ ID NO: 600) СЯЛ может быть присоединена в С-конце РНК-направляемого ДНК-связывающего агента. В сайт слияния, необязательно, включены один или более линкеров.

Как указано выше, РНК-направляемый ДНК-связывающий агент может представлять собой нуклеиновую кислоту, кодирующую РНК-направляемые ДНК-связывающие полипептиды. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент содержит мРНК, содержащую открытую рамку считывания (ОРС), кодирующую РНК-направляемый ДНК-связывающий агент, такой как Cas-нуклеаза, как описано в данном документе. В некоторых вариантах осуществления предложена, применяется или вводится мРНК, содержащая ОРС, кодирующую РНК-направляемый ДНК-связывающий агент, такой как Cas-нуклеаза. Как описано ниже, мРНК, содержащая Cas-нуклеазу, может содержать Cas9-нуклеазу, такую как Cas9-нуклеаза *S. pyogenes*, обладающая кливазной, никазной и/или сайт-специфической ДНК-связывающей активностью. В некоторых вариантах осуществления ОРС, кодирующая РНК-направляемую ДНК-нуклеазу, представляет собой «ОРС модифицированного РНК-направляемого ДНК-связывающего агента» или просто «модифицированную ОРС», что используют как сокращение для обозначения того, что ОРС модифицирована.

ОРС Cas9, включая модифицированные ОРС Cas9, предложены в данном документе и известны в данной области техники. В качестве одного примера, ОРС Cas9 может быть кодон-оптимизированной так, чтобы кодирующая последовательность

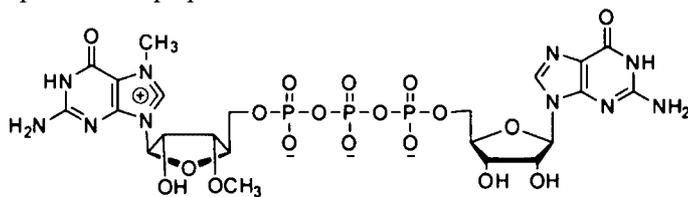
содержала один или более альтернативных кодонов для одной или более аминокислот. В контексте данного документа термин «альтернативный кодон» относится к вариациям в частоте использования кодонов для заданной аминокислоты и может быть или не быть предпочтительным или оптимизированным кодоном (кодон-оптимизированным) для заданной экспрессионной системы. Предпочтительное использование кодонов или кодоны, которые наиболее допустимы в заданной экспрессионной системе, известны в данной области техники. Кодирующие последовательности Cas9, мРНК Cas9 и белковые последовательности Cas9 по WO2013/176772, WO2014/065596, WO2016/106121 и WO2019/067910 включены в данный документ посредством ссылки. В частности, ORF и аминокислотные последовательности Cas9 из таблицы в параграфе [0449] WO2019/067910, а также мРНК и ORF Cas9 из параграфов [0214] - [0234] WO2019/067910 включены сюда посредством ссылки.

В некоторых вариантах осуществления модифицированная OPC может содержать модифицированный уридин по меньшей мере в одном, в нескольких или всех позициях уридина. В некоторых вариантах осуществления модифицированный уридин представляет собой уридин, модифицированный в позиции 5, например, галогеном, метилом или этилом. В некоторых вариантах осуществления модифицированный уридин представляет собой псевдоуридин, модифицированный в позиции 1, например, галогеном, метилом или этилом. Модифицированный уридин может представлять собой, например, псевдоуридин, N1-метил-псевдоуридин, 5-метоксиуридин, 5-йодуридин или их комбинацию. В некоторых вариантах осуществления модифицированный уридин представляет собой 5-метоксиуридин. В некоторых вариантах осуществления модифицированный уридин представляет собой 5-йодуридин. В некоторых вариантах осуществления модифицированный уридин представляет собой псевдоуридин. В некоторых вариантах осуществления модифицированный уридин представляет собой N1-метил-псевдоуридин. В некоторых вариантах осуществления модифицированный уридин представляет собой комбинацию псевдоуридина и N1-метил-псевдоуридина. В некоторых вариантах осуществления модифицированный уридин представляет собой комбинацию псевдоуридина и 5-метоксиуридина. В некоторых вариантах осуществления модифицированный уридин представляет собой комбинацию N1-метил-псевдоуридина и 5-метоксиуридина. В некоторых вариантах осуществления модифицированный уридин представляет собой комбинацию 5-йодуридина и N1-метил-псевдоуридина. В некоторых вариантах осуществления модифицированный уридин представляет собой комбинацию псевдоуридина и 5-йодуридина. В некоторых вариантах осуществления модифицированный уридин представляет собой комбинацию 5-йодуридина и 5-метоксиуридина.

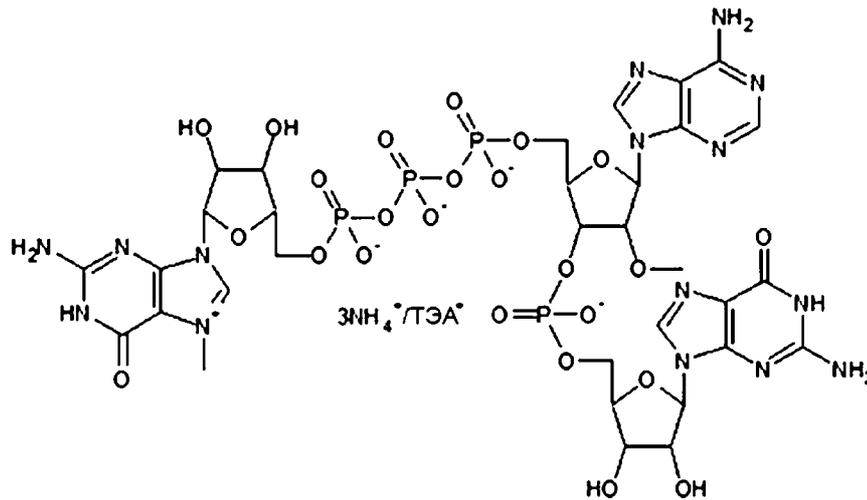
В некоторых вариантах осуществления описанная в данном документе мРНК содержит 5'-кэп, такой как Cap0, Cap1 или Cap2. 5'-кэп обычно представляет собой рибонуклеотид 7-метилгуанина (который может быть дополнительно модифицирован, как обсуждается ниже, например, в отношении ARCA), связанный посредством 5'-трифосфата

с 5'-позицией первого нуклеотида 5'-к-3' цепи мРНК, т. е. первый кэп-проксимальный нуклеотид. В Cap0 рибозы первого и второго кэп-проксимальных нуклеотидов мРНК содержат 2'-гидроксил. В Cap1 рибозы первого и второго транскрибируемых нуклеотидов мРНК содержат 2'-метокси и 2'-гидроксил, соответственно. В Cap2 рибозы первого и второго кэп-проксимальных нуклеотидов мРНК содержат 2'-метокси. Смотрите, например, Katibah et al. (2014) Proc Natl Acad Sci USA 111(33):12025-30; Abbas et al. (2017) Proc Natl Acad Sci USA 114(11):E2106-E2115. Большинство эндогенных мРНК высших эукариот, включая мРНК млекопитающих, такие как мРНК человека, содержат Cap1 или Cap2. Cap0 и другие кэп-структуры, отличные от Cap1 и Cap2, могут быть иммуногенными у млекопитающих, таких как люди, вследствие распознавания их как «чужих» компонентами врожденной иммунной системы, такими как IFIT-1 и IFIT-5, что может привести к повышению уровня цитокинов, включая интерферон I типа. Компоненты врожденной иммунной системы, такие как IFIT-1 и IFIT-5, также могут конкурировать с eIF4E за связывание мРНК с кэпом, отличным от Cap1 или Cap2, потенциально ингибируя трансляцию мРНК.

Кэп может быть включен котранскрипционно. Например, ARCA (англ. «anti-reverse cap analog»; Thermo Fisher Scientific, кат. № AM8045) представляет собой аналог кэпа, содержащий 7-метилгуанин-3'-метокси-5'-трифосфат, связанный с 5'-позицией рибонуклеотида гуанина, который может быть включен в транскрипт *in vitro* при инициации. ARCA приводит к получению кэпа Cap0, в котором 2'-позиция первого кэп-проксимального нуклеотида является гидроксильной. Смотрите, например, Stepinski et al., (2001) “Synthesis and properties of mRNAs containing the novel ‘anti-reverse’ cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl(3'-deoxy)GpppG,” RNA 7: 1486-1495. Структура ARCA проиллюстрирована ниже.



CleanCap™ AG (m7G(5')ppp(5')(2'OMeA)pG; TriLink Biotechnologies, кат. № N-7113) или CleanCap™ GG (m7G(5')ppp(5')(2'OMeG)pG; TriLink Biotechnologies, кат. № N-7133) можно использовать для обеспечения структуры Cap1 котранскрипционно. 3'-О-метилованные версии CleanCap™ AG и CleanCap™ GG также доступны от TriLink Biotechnologies под кат. №№ N-7413 и N-7433, соответственно. Структура CleanCap™ AG проиллюстрирована ниже.



В альтернативном варианте кэп можно добавлять к РНК посттранскрипционно. Например, кэпирующий фермент вируса осповакцины коммерчески доступен (New England Biolabs, кат. № M2080S) и обладает активностью РНК-трифосфатазы и гуанилилтрансферазы, обеспечиваемой субъединицей D1, и активностью гуанинметилтрансферазы, обеспечиваемой субъединицей D12. Таким образом, он может добавлять 7-метилгуанин к РНК с образованием Cap0 в присутствии S-аденозилметионина и ГТФ. Смотрите, например, Guo, P. and Moss, B. (1990) Proc. Natl. Acad. Sci. USA 87, 4023-4027; Mao, X. and Shuman, S. (1994) J. Biol. Chem. 269, 24472-24479.

В некоторых вариантах осуществления мРНК дополнительно содержит полиаденилированный (поли-А) хвост. В некоторых вариантах осуществления поли-А-хвост содержит по меньшей мере 20, 30, 40, 50, 60, 70, 80, 90 или 100 оснований аденина, необязательно, до 300 оснований аденина. В некоторых вариантах осуществления поли-А-хвост содержит по меньшей мере 95, 96, 97, 98, 99 или 100 нуклеотидов аденина.

Способы доставки

Описанные в данном документе конструкции нуклеиновых кислот можно доставлять в клетку-хозяина или организм субъекта *in vivo* или *ex vivo*, используя различные известные и подходящие способы, доступные в данной области техники. Конструкции нуклеиновых кислот можно доставлять вместе с компонентами подходящей системы редактирования генов (например, РНК-направляемым ДНК-связывающим агентом, таким как Cas-нуклеаза с соответствующей гидовой РНК), как описано в данном документе.

Для внесения описанных в данном документе конструкций и компонентов системы редактирования генов в клетки (например, клетки млекопитающих) и целевые ткани можно использовать обычные способы доставки генов на вирусной и невирусной основе. Как дополнительно предложено в данном документе, системы доставки на основе невирусных векторов содержат нуклеиновые кислоты, такие как невирусные векторы, плазмидные векторы и, например, нуклеиновая кислота в комплексе с носителем для доставки, таким как липосома, липидная наночастица (ЛНЧ) или полксамер. Системы доставки на основе вирусных векторов включают ДНК- и РНК-вирусы.

Способы и композиции для невирусной доставки нуклеиновых кислот включают электропорацию, липофекцию, микроинъекцию, биолистику, виросомы, липосомы, иммунолипосомы, ЛНЧ, конъюгаты нуклеиновых кислот с поликатионами или липидами, оголенную нуклеиновую кислоту (например, оголенную ДНК/РНК), искусственные вирионы и усиленное агентом поглощение ДНК. Также для доставки нуклеиновых кислот можно использовать обработку ультразвуком, например, с помощью системы Sonitron 2000 system (Rich-Mar).

Дополнительные типовые системы для доставки нуклеиновых кислот включают предоставляемые AmaxaBiosystems (Cologne, Germany), Maxcyte, Inc. (Rockville, Md.), VTX Molecular Delivery Systems (Holliston, Ma.) и Copernicus Therapeutics Inc., (смотрите, например, патент США № 6008336). Липофекция описана, например, в патентах США №№ 5049386; 4946787; и 4897355), а реагенты для липофекции продаются на коммерческой основе (например, Transfectam™ и Lipofectin™). Получение комплексов липид:нуклеиновая кислота, включая нацеленные липосомы, такие как иммунолипидные комплексы, хорошо известно в данной области техники и описано в данном документе.

Различные системы доставки (например, векторы, липосомы, ЛНЧ), содержащие двунаправленные конструкции и/или компоненты редактирования генов (например, гидовую РНК и Cas), также можно вводить в организм для доставки в клетки *in vivo* или вводить в клетку или клеточную культуру *ex vivo*. Введение осуществляют любым из способов, обычно используемых для итогового приведения молекулы в контакт с кровью, жидкостью или клетками, включая, но не ограничиваясь этим, инъекцию, инфузию, местное нанесение и электропорацию. Подходящие способы введения таких нуклеиновых кислот доступны и хорошо известны специалистам в данной области техники.

В определенных вариантах осуществления в настоящем изобретении предложены векторы, содержащие описанные в данном документе двунаправленные конструкции нуклеиновых кислот для доставки в клетку-хозяина. В определенных вариантах осуществления компоненты системы редактирования генов (например, РНК-направляемый ДНК-связывающий агент и гидовую РНК) также доставляют в клетку-хозяина в виде части вектора. В определенных вариантах осуществления для доставки любого одного или более элемента из двунаправленной конструкции нуклеиновой кислоты, гидовой РНК и/или РНК-направляемого ДНК-связывающего агента в клетку-хозяина можно использовать вирусные векторы.

В некоторых вариантах осуществления в данном документе предложены композиции и способы для доставки описанной в данном документе двунаправленной конструкции нуклеиновой кислоты в клетку-хозяина или организм субъекта, причем конструкция является частью описанной в данном документе векторной системы. В некоторых вариантах осуществления векторная система содержит дополнительные компоненты, такие как компоненты системы редактирования генов (например, гидовую РНК и/или РНК-направляемый ДНК-связывающий агент).

В некоторых вариантах осуществления предложена векторная композиция,

содержащая описанную в данном документе двунаправленную конструкцию нуклеиновой кислоты. В некоторых вариантах осуществления композиция дополнительно содержит компоненты системы редактирования генов (например, гидовую РНК и/или РНК-направляемый ДНК-связывающий агент).

В некоторых вариантах осуществления вектор может быть кольцевым. В других вариантах осуществления вектор может быть линейным. В некоторых вариантах осуществления вектор можно доставлять посредством липидной наночастицы, липосомы, нелипидной наночастицы или вирусного капсида. Неограничивающие типовые векторы включают плазмиды, фагмиды, космиды, искусственные хромосомы, минихромосомы, транспозоны, вирусные векторы и экспрессионные векторы.

В некоторых вариантах осуществления векторная система может быть способна управлять экспрессией одного или более компонентов нуклеазы в клетке. В некоторых вариантах осуществления двунаправленная конструкция, необязательно как часть векторной системы, может содержать промотор, способный управлять экспрессией кодирующей последовательности в клетке. В некоторых вариантах осуществления клетка может быть эукариотической клеткой, такой как, например, клетка дрожжей, растений, насекомых или млекопитающих. В некоторых вариантах осуществления эукариотическая клетка может представлять собой клетку млекопитающего. В некоторых вариантах осуществления эукариотическая клетка может представлять собой клетку грызуна. В некоторых вариантах осуществления эукариотическая клетка может представлять собой клетку человека. Подходящие промоторы для управления экспрессией в разных типах клеток известны в данной области техники. В некоторых вариантах осуществления промотор может быть дикого типа. В других вариантах осуществления промотор может быть модифицирован для более эффективной или действенной экспрессии. В других вариантах осуществления промотор может быть усечен, но при этом сохранять свою функцию. Например, промотор может иметь нормальный или уменьшенный размер, который подходит для надлежащей упаковки вектора в вирус. В некоторых вариантах осуществления вектор не содержит промотор, который управляет экспрессией одной или более кодирующих последовательностей в клетке (например, экспрессией кодирующей последовательности после вставки в целевой эндогенный локус управляет эндогенный промотор).

В некоторых вариантах осуществления вектор может представлять собой вирусный вектор. В некоторых вариантах осуществления вирусный вектор может быть генетически модифицирован по сравнению с его аналогом дикого типа. Например, вирусный вектор может содержать вставку, делецию или замену одного или более нуклеотидов для облегчения клонирования или для изменения одного или более свойств вектора. Такие свойства могут включать паковую способность, эффективность трансдукции, иммуногенность, интеграцию в геном, репликацию, транскрипцию и трансляцию. В некоторых вариантах осуществления часть вирусного генома может быть удалена, чтобы вирус был способен паковать экзогенные последовательности большего размера. В

некоторых вариантах осуществления вирусный вектор может иметь повышенную эффективность трансдукции. В некоторых вариантах осуществления иммунный ответ, индуцированный вирусом у хозяина, может быть снижен. В некоторых вариантах осуществления вирусные гены (такие как, например, интеграза), которые способствуют интеграции вирусной последовательности в геном хозяина, могут быть мутированы так, чтобы вирус становится неинтегрирующимся. В некоторых вариантах осуществления вирусный вектор может быть дефективным по репликации. В некоторых вариантах осуществления вирусный вектор может содержать экзогенные последовательности управления транскрипцией или трансляцией для управления экспрессией кодирующих последовательностей в векторе. В некоторых вариантах осуществления вирус может быть хелпер-зависимым. Например, вирусу могут быть необходимы один или более хелперных вирусов для доставки вирусных компонентов (таких как, например, вирусные белки), необходимых для амплификации и упаковки векторов в вирусные частицы. В таком случае в клетку-хозяина вместе с описанной в данном документе векторной системой можно вносить один или более хелперных компонентов, включая один или более векторов, кодирующих вирусные компоненты. В других вариантах осуществления вирус может быть хелпер-независимым. Например, вирус может быть способен амплифицировать и упаковать векторы без хелперного вируса. В некоторых вариантах осуществления описанная в данном документе векторная система может также кодировать вирусные компоненты, необходимые для амплификации и упаковки вируса.

Преимуществом применения систем на основе РНК- или ДНК-вирусов для доставки нуклеиновых кислот являются хорошо отработанные процессы нацеливания вируса на конкретные клетки в организме и переноса вирусной нагрузки в ядро. Вирусные векторы можно вводить непосредственно субъекту (*in vivo*) или же их можно использовать для обработки клеток *in vitro*. В некоторых вариантах осуществления модифицированные *in vitro* клетки вводят субъекту (например, посредством *ex vivo* обработки клеток, полученных от субъекта или из донорского источника). Неограничивающие типовые вирусные векторы включают вектор на основе аденоассоциированного вируса (AAV), лентивирусные векторы, аденовирусные векторы, хелпер-зависимые аденовирусные векторы (HDAd), векторы на основе вируса простого герпеса (HSV-1), бактериофаг T4, бакуловиральные векторы и ретровирусные векторы. Интеграция в геном хозяина возможна при применении методов переноса генов с использованием ретровирусов, лентивирусов и аденоассоциированных вирусов, что часто приводит к долгосрочной экспрессии вставленного трансгена. Кроме того, во многих разных типах клеток и целевых тканях наблюдали высокую эффективность трансдукции.

Тропизм ретровирусов можно менять путем включения чужеродных белков оболочки, расширяя потенциальную популяцию целевых клеток. Лентивирусные векторы представляют собой ретровирусные векторы, которые могут трансдуцировать или инфицировать неделящиеся клетки и, как правило, дают высокие вирусные титры. Выбор ретровирусной системы переноса генов зависит от целевой ткани. Ретровирусные векторы

состоят из цис-действующих длинных концевых повторов с пакующей способностью до 6-10 т. п. о. чужеродной последовательности. Минимальных цис-действующих ДКП достаточно для репликации и упаковки векторов, которые затем используют для интеграции двунаправленной конструкции, содержащей трансген, в целевую клетку для обеспечения перманентной экспрессии трансгена. Широко используемые ретровирусные векторы включают векторы на основе вируса лейкоза мышей (MuLV), вируса лейкоза гиббонов (GaLV), вируса иммунодефицита обезьян (SIV), вируса иммунодефицита человека (HIV) и их комбинаций (смотрите, например, Buchscher et al., *J. Virol.* 66:2731-2739 (1992); Johann et al., *J. Virol.* 66:1635-1640 (1992); Sommerfelt et al., *Virol.* 176:58-59 (1990); Wilson et al., *J. Virol.* 63:2374-2378 (1989); Miller et al., *J. Virol.* 65:2220-2224 (1991); PCT/US94/05700).

В некоторых вариантах осуществления можно использовать систему на основе аденовируса. Векторы на основе аденовирусов имеют очень высокую эффективность трансдукции во многих типах клеток и не нуждаются в делении клеток. С такими векторами получали высокие титры и высокие уровни экспрессии. Этот вектор можно получать в больших количествах в относительно простой системе. Дефективные по репликации рекомбинантные аденовирусные векторы можно получать при высоком титре и без проблем инфицировать ими ряд разных типов клеток. Большинство аденовирусных векторов конструируют так, чтобы трансген замещал гены Ad E1a, E1b и/или E3; после этого дефективный по репликации вектор размножают в клетках 293 человека, которые обеспечивают функцию удаленного гена в транс-форме. Аденовирусные векторы могут трансдуцировать множество типов тканей *in vivo*, включая неделящиеся, дифференцированные клетки, такие как клетки печени, почки и мышц. Традиционные аденовирусные векторы имеют большую несущую емкость. Пример применения аденовирусного вектора в клиническом исследовании включал полинуклеотидную терапию для противоопухолевой иммунизации посредством внутримышечной инъекции (Sternan et al., *Hum. Gene Ther.* 7:1083-9 (1998)). Дополнительные примеры применения аденовирусных векторов для переноса генов в клинических исследованиях включают Rosenecker et al., *Infection* 24:1 5-10 (1996); Sternan et al., *Hum. Gene Ther.* 9:7 1083-1089 (1998); Welsh et al., *Hum. Gene Ther.* 2:205-18 (1995); Alvarez et al., *Hum. Gene Ther.* 5:597-613 (1997); Topf et al., *Gene Ther.* 5:507-513 (1998); Sternan et al., *Hum. Gene Ther.* 7:1083-1089 (1998).

В некоторых вариантах осуществления для доставки предложенных в данном документе двунаправленных конструкций нуклеиновых кислот используют векторы на основе аденоассоциированного вируса (AAV). AAV-векторы хорошо известны и используются для трансдукции клеток целевыми нуклеиновыми кислотами, например, при *in vitro* получении нуклеиновых кислот и пептидов, и для *in vivo* и *ex vivo* процедур генной терапии (смотрите, например, West et al., *Virology* 160:38-47 (1987); патент США № 4797368; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994). Конструирование рекомбинантных AAV-векторов описано в ряде

публикаций, включая патент США № 5173414; Tratschin et al., *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, et al., *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984); и Samulski et al., *J. Virol.* 63:03822-3828 (1989). В некоторых вариантах осуществления вирусный вектор может представлять собой AAV-вектор. В некоторых вариантах осуществления AAV-вектор представляет собой, например, AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10 или AAVLK03, и также любой новый серотип AAV можно использовать в соответствии с настоящим изобретением. AAV-векторы и рекомбинантные аденоассоциированные вирусные векторы являются перспективной альтернативой системам доставки нуклеиновых кислот, например, векторы на основе дефективного и непатогенного парвовирусного аденоассоциированного вируса типа 2.

В контексте данного документа «AAV» относится ко всем серотипам, подтипам и встречающимся в природе AAV, а также к рекомбинантному AAV. Аббревиатуру «AAV» можно использовать для обозначения самого вируса или его производного. Термин «AAV» включает AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV12, rh10 и их гибриды, AAV птиц, AAV бычьих, AAV собачьих, AAV лошадиных, AAV приматов, AAV отличных от приматов животных и AAV овечьих. Геномные последовательности различных серотипов AAV, а также последовательности нативных концевых повторов (КП), Rep-белков и субъединиц капсида известны в данной области техники. Такие последовательности можно найти в литературе или в общедоступных базах данных, таких как GenBank. В контексте данного документа термин «AAV-вектор» относится к AAV-вектору, содержащему гетерологичную последовательность с происхождением, отличным от AAV (т. е. последовательность нуклеиновой кислоты, гетерологичную для AAV), обычно содержащую последовательность, кодирующую представляющий интерес гетерологичный полипептид. Конструкция может содержать капсидную последовательность AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV12, rh10 и их гибридов, AAV птиц, AAV бычьих, AAV собачьих, AAV лошадиных, AAV приматов, AAV отличных от приматов животных и AAV овечьих. В общем случае последовательность гетерологичной нуклеиновой кислоты (трансен) фланкируется по меньшей мере одной и обычно двумя последовательностями инвертированных концевых повторов AAV (ИКП). AAV-вектор может быть одноцепочечным (ssAAV) или самокомплементарным (scAAV).

В некоторых вариантах осуществления вирусный вектор может представлять собой лентивирусный вектор. В некоторых вариантах осуществления лентивирус может быть неинтегрирующимся. В некоторых вариантах осуществления вирусный вектор может представлять собой аденовирусный вектор. В некоторых вариантах осуществления

аденовирус может представлять собой аденовирус с высокой клонирующей способностью или аденовирус «без внутренностей», в котором удалены все кодирующие вирусные области, кроме 5' и 3' инвертированных концевых повторов (ИКП) и сигнала упаковки (Г), чтобы увеличить его пакующую способность. В некоторых вариантах осуществления вирусный вектор может представлять собой вектор HSV-1. В некоторых вариантах осуществления вектор на основе HSV-1 является хелпер-зависимым, а в других вариантах он является хелпер-независимым. Например, для ампликонного вектора, в котором сохранена только последовательность упаковки, необходим хелперный вирус со структурными компонентами для упаковки, тогда как для вектора HSV-1 с удалением 30 т. п. о., что устраняет второстепенные вирусные функции, нет необходимости в хелперном вирусе. В некоторых вариантах осуществления вирусный вектор может представлять собой бактериофаг T4. В некоторых вариантах осуществления бактериофаг T4 может быть способен упаковывать любые линейные или кольцевые молекулы ДНК или РНК при опустошении головки вируса. В дополнительных вариантах осуществления вирусный вектор может представлять собой бакуловирусный вектор. В дополнительных вариантах осуществления вирусный вектор может представлять собой ретровирусный вектор. В вариантах осуществления с использованием AAV или лентивирусных векторов, которые обладают меньшей клонирующей способностью, может быть необходимо использовать более одного вектора для доставки всех компонентов векторной системы, как описано в данном документе. Например, один AAV-вектор может содержать последовательности, кодирующие РНК-направляемый ДНК-связывающий агент, такой как белок Cas (например, Cas9), тогда как второй AAV-вектор может содержать одну или более направляющих последовательностей.

Пакующие клетки используют для формирования вирусных частиц, способных инфицировать клетку-хозяина. Такие клетки включают клетки 293, которые могут упаковывать аденовирусы и AAV, и клетки ψ 2 или клетки PA317, которые пакует ретровирусы. Вирусные векторы, используемые в генной терапии, обычно создают с помощью вырабатывающей линии клеток, которая упаковывает вектор нуклеиновой кислоты в вирусную частицу. Векторы, как правило, содержат минимальные вирусные последовательности, необходимые для упаковки, другие вирусные последовательности замещены последовательностями, кодирующими подлежащий экспрессии белок. Отсутствующие вирусные функции обеспечивает в транс-форме пакующая линия клеток. Например, используемые в генной терапии AAV-векторы содержат только последовательности инвертированных концевых повторов (ИКП) из генома AAV, которые необходимы для упаковки. Упаковка вирусной ДНК происходит в линии клеток, которые содержат хелперную плазмиду, кодирующую другие гены AAV, а именно гер и сар, но в которой отсутствуют последовательности ИКП. Линию клеток также можно инфицировать аденовирусом в качестве хелпера. Хелперный вирус способствует репликации AAV-вектора и экспрессии генов AAV из хелперной плазмиды.

Векторы для генной терапии можно доставлять *in vivo* путем введения отдельному

пациенту, как правило, посредством системного введения (например, внутривенной, внутривентрикулярной, внутримышечной, подкожной или внутримозговой инфузии) или местного введения, как описано ниже. В альтернативном варианте векторы можно доставлять в клетки *ex vivo*, например в клетки, полученные от отдельного пациента (например, лимфоциты, аспират костного мозга, тканевая биопсия), или универсальные донорные гемопоэтические стволовые клетки, с последующей повторной имплантацией клеток пациенту, обычно после отбора клеток, в которые был включен вектор.

В некоторых вариантах осуществления, в дополнение к описанной в данном документе донорной конструкции векторная система может дополнительно содержать нуклеиновые кислоты, которые кодируют нуклеазу. В некоторых вариантах осуществления, в дополнение к описанной в данном документе донорной конструкции векторная система может дополнительно содержать нуклеиновые кислоты, которые кодируют гидовую РНК, и/или нуклеиновую кислоту, кодирующую РНК-направляемый ДНК-связывающий агент, который может представлять собой белок Cas, такой как Cas9. В некоторых вариантах осуществления нуклеиновая кислота, кодирующая гидовую РНК, и/или нуклеиновая кислота, кодирующая РНК-направляемый ДНК-связывающий агент или нуклеазу, каждая или обе, находятся в векторе, отличном от вектора, который содержит описанные в данном документе двунаправленные конструкции. В любом из вариантов осуществления векторная система может содержать другие последовательности, которые включают, но не ограничиваются этим, промоторы, энхансеры, регуляторные последовательности, как описано в данном документе. В некоторых вариантах осуществления промотор в векторной системе не управляет экспрессией трансгена двунаправленной конструкции. В некоторых вариантах осуществления векторная система содержит одну или более нуклеотидных последовательностей, кодирующих *cr*РНК, *tr*РНК или *cr*РНК и *tr*РНК. В некоторых вариантах осуществления вектор содержит одну или более нуклеотидных последовательностей, кодирующих *og*РНК и *m*РНК, кодирующие РНК-направляемый ДНК-связывающий агент, который может представлять собой Cas-нуклеазу (например, Cas9). В некоторых вариантах осуществления векторная система содержит одну или более нуклеотидных последовательностей, кодирующих *cr*РНК, *tr*РНК и *m*РНК, кодирующие РНК-направляемый ДНК-связывающий агент, который может представлять собой Cas-нуклеазу, такую как Cas9. В некоторых вариантах осуществления Cas9 получен из *Streptococcus pyogenes* (т. е. Cas9 Spy). В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая *cr*РНК, *tr*РНК или *cr*РНК и *tr*РНК (которая может представлять собой *og*РНК), содержит или состоит из направляющей последовательности, фланкируемой всей или частью повторяющейся последовательности из CRISPR/Cas природного происхождения. Векторная система может содержать нуклеиновую кислоту, содержащую *cr*РНК, *tr*РНК или *cr*РНК и *tr*РНК, причем векторная система содержит или состоит из нуклеиновых кислот, которые в природе не встречаются вместе с *cr*РНК, *tr*РНК или *cr*РНК и *tr*РНК. Любые из описанных в данном документе

векторов можно доставлять с помощью липосомы, наночастицы, экзосомы, микровезикулы и/или липидных наночастиц (ЛНЧ). Одна или более гидовых РНК, РНК-направляемых ДНК-связывающих агентов (например, мРНК) или донорных конструкций, содержащих последовательность, кодирующую гетерологичный белок, по отдельности или в любой комбинации, могут быть доставлены с помощью липосомы, наночастицы, экзосомы или микровезикулы. Одна или более гидовых РНК, РНК-направляемых ДНК-связывающих агентов (например, мРНК) или донорных конструкций, содержащих последовательность, кодирующую гетерологичный белок, по отдельности или в любой комбинации, могут быть доставлены с помощью ЛНЧ. Любые описанные в данном документе ЛНЧ или составы ЛНЧ подходят для доставки направляющих.

Липидные наночастицы (ЛНЧ) являются хорошо известным средством доставки нуклеотидного и белкового груза и могут использоваться для доставки двунаправленных конструкций нуклеиновых кислот, описанных в данном документе. В некоторых вариантах осуществления ЛНЧ можно использовать для доставки компонентов системы редактирования генов. В некоторых вариантах осуществления ЛНЧ доставляют нуклеиновые кислоты (например, ДНК или РНК), белок (например, РНК-направляемый ДНК-связывающий агент) или нуклеиновую кислоту вместе с белком.

В некоторых вариантах осуществления в данном документе предложен способ доставки описанной в данном документе двунаправленной конструкции нуклеиновой кислоты в клетку-хозяина или организм субъекта, причем доставка конструкции происходит с помощью ЛНЧ. В некоторых вариантах осуществления в данном документе предложен способ доставки описанной в данном документе двунаправленной конструкции нуклеиновой кислоты в клетку-хозяина или организм субъекта, причем доставка одного или более компонентов системы редактирования генов, такой как нуклеазная система CRISPR/Cas, происходит с помощью ЛНЧ. В некоторых вариантах осуществления ЛНЧ содержат двунаправленную конструкцию и/или один или более компонентов системы редактирования генов (например, гидовую РНК и/или РНК-направляемый ДНК-связывающий агент, или мРНК, кодирующую РНК-направляемый ДНК-связывающий агент).

В некоторых вариантах осуществления в данном документе предложена композиция, содержащая описанную в данном документе двунаправленную конструкцию нуклеиновой кислоты и ЛНЧ. В некоторых вариантах осуществления композиция дополнительно содержит компоненты системы редактирования генов (например, гидовую РНК и/или РНК-направляемый ДНК-связывающий агент, такой как Cas9, или векторную систему, способную их кодировать). В некоторых вариантах осуществления в данном документе предложена композиция, содержащая описанную в данном документе двунаправленную конструкцию нуклеиновой кислоты и ЛНЧ, содержащую гидовую РНК и/или мРНК, кодирующую РНК-направляемый ДНК-связывающий агент, такой как Cas9.

В некоторых вариантах осуществления ЛНЧ содержат биоразлагаемые ионизируемые липиды. В некоторых вариантах осуществления ЛНЧ содержат (9Z,12Z)-3-

((4,4-бис(октилокси)бутаноил)окси)-2-(((3-(диэтиламино)пропокси)карбонил)окси)метил)пропил(9Z,12Z)-октадека-9,12-диеноат, также называемый 3-(((4,4-бис(октилокси)бутаноил)окси)-2-(((3-(диэтиламино)пропокси)карбонил)окси)метил)пропил(9Z,12Z)-октадека-9,12-диеноатом), или другой ионизируемый липид. Смотрите, например, липиды из PCT/US2018/053559 (поданной 28 сентября, 2018 г.), WO/2017/173054, WO2015/095340 и WO2014/136086, а также ссылки, приведенные в них. В некоторых вариантах осуществления термины «катионный» и «ионизируемый» в контексте ЛНЧ-липидов являются взаимозаменяемыми, например, когда ионизируемые липиды являются катионными в зависимости от pH.

Электропорация является хорошо известным средством доставки груза, и любые методы электропорации можно использовать для доставки описанной в данном документе двунаправленной конструкции. В некоторых вариантах осуществления электропорацию можно использовать для доставки описанной в данном документе двунаправленной конструкции, необязательно, при этом гидовая РНК и/или РНК-направляемый ДНК-связывающий агент (например, Cas9), или мРНК, кодирующая РНК-направляемый ДНК-связывающий агент (например, Cas9), доставляются одним или разными способами.

В некоторых вариантах осуществления настоящее изобретение включает способ доставки описанной в данном документе двунаправленной конструкции в клетку *in vitro*, причем доставка двунаправленной конструкции происходит с помощью ЛНЧ. В некоторых вариантах осуществления двунаправленную конструкцию доставляют не связанными с ЛНЧ способами, например, посредством AAV-системы, а гидовую РНК и/или РНК-направляемый ДНК-связывающий агент (например, Cas9), или мРНК, кодирующую РНК-направляемый ДНК-связывающий агент (например, Cas9) доставляют с помощью ЛНЧ.

В некоторых вариантах осуществления описанную в данном документе двунаправленную конструкцию, одну или как часть вектора, составляют или вводят с помощью липидной наночастицы; смотрите, например, WO/2017/173054, содержание которой в полном объеме включено в данный документ посредством ссылки.

Любые из описанных в данном документе векторов можно доставлять с помощью ЛНЧ. Любые описанные в данном документе ЛНЧ и составы ЛНЧ подходят для доставки гРНК, Cas-нуклеазы или мРНК, кодирующей Cas-нуклеазу, из комбинаций и/или описанной в данном документе двунаправленной конструкции. В некоторых вариантах осуществления включена композиция ЛНЧ, содержащая компонент РНК и липидный компонент, причем липидный компонент включает аминный липид, такой как биоразлагаемый ионизируемый липид; а компонент РНК включает гидовую РНК и/или мРНК, кодирующей Cas-нуклеазу.

В некоторых случаях липидный компонент включает биоразлагаемый ионизируемый липид, холестерин, ДСФХ и ПЭГ-ДМГ.

Следует понимать, что компоненты системы редактирования генов (например, гидовую РНК и/или РНК-направляемый ДНК-связывающий агент) и двунаправленные

конструкции можно доставлять, используя одинаковые или разные системы. Например, гидовая РНК, последовательность РНК-направляемого ДНК-связывающего агента и двунаправленная конструкция могут содержаться в одном векторе (например, AAV-векторе) или быть составлены в одной или более композициях ЛНЧ. В альтернативном варианте РНК-направляемый ДНК-связывающий агент (в виде белка или мРНК) и/или гРНК могут содержаться в ЛНЧ или быть связанными с ней, тогда как двунаправленные конструкции могут содержаться в векторе, или наоборот. Кроме того, разные системы доставки можно вводить одним или разными путями.

Разные системы доставки можно доставлять *in vitro* или *in vivo* одновременно или в любом последовательном порядке. В некоторых вариантах осуществления двунаправленную конструкцию, гидовую РНК и РНК-направляемый ДНК-связывающий агент можно доставлять *in vitro* или *in vivo* одновременно, например, в одном векторе, двух векторах, отдельных векторах, одной ЛНЧ, двух ЛНЧ, отдельных ЛНЧ или их комбинации. В некоторых вариантах осуществления двунаправленную конструкцию можно доставлять *in vivo* или *in vitro* в виде вектора и/или связанной с ЛНЧ до (например, за около 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 или более суток) доставки гидовой РНК и/или РНК-направляемого ДНК-связывающего агента в виде вектора и/или связанных с ЛНЧ, по отдельности или вместе в виде рибонуклеопротеина (РНП). В некоторых вариантах осуществления донорную конструкцию можно доставлять многократно, например, каждые сутки, каждые двое суток, каждые трое суток, каждые четверо суток, каждую неделю, каждые две недели, каждые три недели или каждые четыре недели. В некоторых вариантах осуществления донорную конструкцию можно доставлять с однедельными интервалами, например, на 1 неделе, 2 неделе и 3 неделе и т. д. В качестве дополнительного примера, гидовую РНК и/или РНК-направляемый ДНК-связывающий агент в виде вектора и/или связанных с ЛНЧ, по отдельности или вместе в виде рибонуклеопротеина (РНП), можно доставлять *in vivo* или *in vitro* до (например, за около 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 или более суток) доставки двунаправленной конструкции в виде вектора и/или связанной с ЛНЧ. В некоторых вариантах осуществления гидовую РНК альбумина можно доставлять многократно, например, каждые сутки, каждые двое суток, каждые трое суток, каждые четверо суток, каждую неделю, каждые две недели, каждые три недели или каждые четыре недели. В некоторых вариантах осуществления гидовую РНК альбумина можно доставлять с однедельными интервалами, например, на 1 неделе, 2 неделе и 3 неделе и т. д. В некоторых вариантах осуществления Cas-нуклеазу можно доставлять за несколько введений, например, можно доставлять каждые сутки, каждые двое суток, каждые трое суток, каждые четверо суток, каждую неделю, каждые две недели, каждые три недели или каждые четыре недели. В некоторых вариантах осуществления Cas-нуклеазу можно доставлять с интервалами в одну неделю, например, на 1 неделю, 2 неделю, 3 неделю и т. д.

Способы применения

В настоящем изобретении предложены способы использования описанной в

данном документе двунаправленной конструкции нуклеиновой кислоты в различных применениях. В некоторых вариантах осуществления способы использования описанной в данном документе двунаправленной конструкции нуклеиновой кислоты в различных применениях включают применение системы редактирования генов, такой как система CRISPR/Cas, описанная в данном документе.

В некоторых вариантах осуществления в данном документе предложен *in vitro* или *in vivo* способ модификации целевого локуса (например, вставки трансгена в целевой сайт в локусе), включающий введение или доставку в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты, гидовой РНК и РНК-направляемого ДНК-связывающего агента, описанного в данном документе (например, Cas-нуклеазы, такой как Cas9). В некоторых вариантах осуществления в данном документе предложен *in vitro* или *in vivo* способ модификации целевого локуса, включающий расщепление целевой последовательности в клетке-хозяине и вставку описанной в данном документе двунаправленной конструкции нуклеиновой кислоты, необязательно, с использованием гидовой РНК и РНК-направляемого ДНК-связывающего агента, описанного в данном документе (например, Cas-нуклеазы, такой как Cas9) на этапе расщепления.

В некоторых вариантах осуществления в данном документе предложен *in vitro* или *in vivo* способ внесения конструкции в клетку-хозяина, включающий введение или доставку в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты, гидовой РНК и РНК-направляемого ДНК-связывающего агента, описанного в данном документе (например, Cas-нуклеазы, такой как Cas9). В некоторых вариантах осуществления в данном документе предложен *in vitro* или *in vivo* способ внесения конструкции в клетку-хозяина, включающий введение или доставку в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты и системы редактирования генов, такой как система ZFN, TALEN, или CRISPR/Cas9.

В некоторых вариантах осуществления в данном документе предложен *in vitro* или *in vivo* способ повышения экспрессии полипептида в клетке-хозяине, включающий введение или доставку в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты, гидовой РНК и РНК-направляемого ДНК-связывающего агента, описанного в данном документе (например, Cas-нуклеазы, такой как Cas9). В некоторых вариантах осуществления в данном документе предложен *in vitro* или *in vivo* способ повышения экспрессии полипептида в клетке-хозяине, включающий введение или доставку в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты и системы редактирования генов, такой как система ZFN, TALEN, или CRISPR/Cas9. Полипептид может быть внеклеточным.

Двунаправленную конструкцию можно вводить с помощью вектора, такого как вектор нуклеиновой кислоты. Гидовую РНК и РНК-направляемый ДНК-связывающий агент можно вводить отдельно или в любой комбинации, например, с помощью ЛНЧ,

содержащей гидовую РНК и мРНК, кодирующую РНК-направляемый ДНК-связывающий агент. Введение и доставку в клетку-хозяина можно осуществлять любым из способов доставки, описанных в данном документе.

В некоторых вариантах осуществления в данном документе предложен *in vitro* или *in vivo* способ экспрессии полипептида, кодируемого трансгеном в целевом локусе, включающий введение или доставку в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты, гидовой РНК и РНК-направляемого ДНК-связывающего агента, описанного в данном документе (например, Cas-нуклеазы, такой как Cas9). В некоторых вариантах осуществления в данном документе предложен *in vitro* или *in vivo* способ экспрессии полипептида, кодируемого трансгеном в целевом локусе, включающий введение или доставку в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты и системы редактирования генов, такой как система ZFN, TALEN, или CRISPR/Cas9. В некоторых вариантах осуществления способ создания клетки-хозяина для экспрессии полипептида включает введение или доставку в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты и системы редактирования генов, такой как система ZFN, TALEN, или CRISPR/Cas9.

Двунаправленную конструкцию, гидовую РНК и РНК-направляемый ДНК-связывающий агент, например, можно доставлять отдельно или в любой комбинации, как описано в данном документе. В некоторых вариантах осуществления двунаправленную конструкцию, гидовую РНК и РНК-направляемый ДНК-связывающий агент можно доставлять одновременно или последовательно, например, в одном векторе, двух векторах, отдельных векторах, одной ЛНЧ, двух ЛНЧ, отдельных ЛНЧ или их комбинации. Введение и доставку в клетку-хозяина можно осуществлять любым из способов доставки, описанных в данном документе.

Кроме того, в некоторых вариантах осуществления способы включают вставку в локус альбумина, такой как интрон 1 альбумина, например, с использованием гидовой РНК, содержащей последовательность, выбранную из любой из таблиц 5, 6, 7, 8, 9 и 10. В определенных вариантах осуществления, включающих вставку в локус альбумина, циркулирующие уровни альбумина индивида являются нормальными. Способ может включать поддержание циркулирующих уровней альбумина индивида в пределах ± 5 , ± 10 , ± 15 , ± 20 или $\pm 50\%$ от нормальных циркулирующих уровней альбумина. В некоторых вариантах осуществления уровни альбумина у индивида остаются неизменными по сравнению с уровнями альбумина у индивидов, не получающих лечения, по меньшей мере на 4 неделе, 8 неделе, 12 неделе или 20 неделе. В некоторых вариантах осуществления уровни альбумина индивида временно падают, а затем возвращаются к нормальным уровням. В частности, способы могут включать обнаружение отсутствия значительных изменений плазменных уровней альбумина.

В некоторых вариантах осуществления изобретение включает способ или применение модификации (например, создания двухцепочечного разрыва) гена

альбумина, такого как ген альбумина человека, включающие введение или доставку в клетку-хозяина или популяцию клеток-хозяев любых описанных в данном документе одной или более гРНК, донорной конструкции (например, двунаправленной конструкции, содержащей последовательность, кодирующую фактор IX) и РНК-направляемых ДНК-связывающих агентов (например, Cas-нуклеазы). В некоторых вариантах осуществления изобретение включает способ или применение модификации (например, создания двухцепочечного разрыва) области интрона 1 альбумина, такого как интрон 1 альбумина человека, включающие введение или доставку в клетку-хозяина или популяцию клеток-хозяев любых описанных в данном документе одной или более гРНК, донорной конструкции (например, двунаправленной конструкции, содержащей нуклеиновую кислоту, кодирующую гетерологичный полипептид) и РНК-направляемых ДНК-связывающих агентов (например, Cas-нуклеазы или нуклеиновой кислоты, кодирующей Cas-нуклеазу). В некоторых вариантах осуществления изобретение включает способ или применение модификации (например, создания двухцепочечного разрыва) безопасного приемного локуса человека, такого как ткань печени или гепатоцитарная клетка-хозяин, включающие введение или доставку в клетку-хозяина или популяцию клеток-хозяев любых описанных в данном документе одной или более гРНК, донорной конструкции (например, двунаправленной конструкции, содержащей последовательность, кодирующую гетерологичный полипептид) и РНК-направляемых ДНК-связывающих агентов (например, Cas-нуклеазы или нуклеиновой кислоты, кодирующей Cas-нуклеазу).

Вставка и/или экспрессия трансгена может происходить в его когнатном локусе (например, вставка трансгена дикого типа в эндогенном локусе) или в некогнатном локусе (например, безопасном приемном локусе, таком как локус альбумина), как описано в данном документе.

В некоторых вариантах осуществления клетка-хозяин принадлежит типу непролиферирующих клеток. В контексте данного документа термин «неделяющаяся клетка» относится к клеткам, которые окончательно дифференцированы и не делятся, а также к покоящимся клеткам, которые не делятся, но сохраняют способность повторно начинать клеточное деление и пролиферацию. Клетки печени, например, сохраняют способность делиться (например, при повреждении или резекции), но, как правило, не делятся. При митотическом делении клеток гомологичная рекомбинация представляет собой механизм, посредством которого происходит защита генома и репарация двухцепочечных разрывов. В некоторых вариантах осуществления «неделяющаяся» клетка относится к клетке, в которой гомологичная рекомбинация (ГР) не является основным механизмом, посредством которого происходит репарация двухцепочечных разрывов ДНК в клетке, например, по сравнению с контрольной делящейся клеткой. В некоторых вариантах осуществления «неделяющаяся» клетка относится к клетке, в которой негомологичное соединение концов (НГСК) является основным механизмом, посредством которого происходит репарация двухцепочечных разрывов ДНК в клетке, например, по сравнению с контрольной делящейся клеткой. Неделяющиеся типы клеток были описаны в

литературе, например, в отношении активных механизмов НГСК для репарации двухцепочечных разрывов ДНК. Смотрите, например, Iyama, DNA Repair (Amst.) 2013, 12(8): 620-636. В некоторых вариантах осуществления клетка-хозяин включает, но не ограничивается этим, клетку печени, мышечную клетку или нейрональную клетку. В некоторых вариантах осуществления клетка-хозяин представляет собой гепатоцит, такой как гепатоцит мышцы, яванского макака или человека. В некоторых вариантах осуществления клетка-хозяин представляет собой миоцит, такой как миоцит мышцы, яванского макака или человека. В некоторых вариантах осуществления в данном документе предложена клетка-хозяин, описанная выше, которая содержит описанную в данном документе двунаправленную конструкцию. В некоторых вариантах осуществления клетка-хозяин экспрессирует трансгенный полипептид, кодируемый описанной в данном документе двунаправленной конструкцией. В некоторых вариантах осуществления в данном документе предложена клетка-хозяин, созданная описанным в данном документе способом. В определенных вариантах осуществления клетка-хозяин создана путем введения или доставки в клетку-хозяина описанной в данном документе двунаправленной конструкции нуклеиновой кислоты и системы редактирования генов, такой как система ZFN, TALEN, или CRISPR/Cas9.

Также предложен способ экспрессии полипептида из описанной в данном документе двунаправленной конструкции. Аналогично, клетка-хозяин, содержащая описанную в данном документе двунаправленную конструкцию, может экспрессировать полипептид, кодируемый конструкцией. В некоторых вариантах осуществления полипептид представляет собой секретлируемый полипептид. В некоторых вариантах осуществления полипептид является таким, чья функция обычно осуществляется (например, который является функционально активным) как у секретлируемого полипептида. В контексте данного документа «секретлируемый полипептид» относится к белку, который секретлируется клеткой. В некоторых вариантах осуществления полипептид представляет собой внутриклеточный полипептид. В некоторых вариантах осуществления полипептид является таким, чья функция обычно осуществляется (например, который является функционально активным) внутри клетки. В контексте данного документа «внутриклеточный полипептид» относится к белку, который не секретлируется клеткой, включая растворимые цитозольные полипептиды. В некоторых вариантах осуществления полипептид представляет собой полипептид дикого типа. В некоторых вариантах осуществления полипептид представляет собой мутантный полипептид (например, гиперактивный мутант полипептида дикого типа). В некоторых вариантах осуществления полипептид представляет собой печеночный белок. В некоторых вариантах осуществления полипептид представляет собой не печеночный белок. В некоторых вариантах осуществления полипептид включает, но не ограничивается этим, фактор IX и его варианты. В некоторых вариантах осуществления печеночный полипептид представляет собой, например, полипептид для разрешения печеночного нарушения, такого как, без ограничения, тирозинемия, болезнь Вильсона,

болезнь Тея - Сакса, гипербилирубинемия (болезнь Криглера - Найяра), острая интермиттирующая порфирия, цитруллинемия типа 1, прогрессирующий семейный внутрипеченочный холестаз или болезнь кленового сиропа.

В некоторых вариантах осуществления способ дополнительно включает достижение продолжительного эффекта, длящегося, например, по меньшей мере 1 месяц, 2 месяца, 6 месяцев, 1 год или 2 года. В некоторых вариантах осуществления способ дополнительно включает достижение продолжительного и устойчивого терапевтического эффекта, длящегося, например, по меньшей мере 1 месяц, 2 месяца, 6 месяцев, 1 год или 2 года. В некоторых вариантах осуществления уровень активности и/или уровень циркулирующего гетерологичного полипептида является стабильным в течение по меньшей мере 1 месяца, 2 месяцев, 6 месяцев, 1 года или более. В некоторых вариантах осуществления достижение стабильной активности и/или уровня полипептида происходит по меньшей мере через 7 суток, по меньшей мере 14 суток или по меньшей мере 28 суток. В дополнительных вариантах осуществления способ включает поддержание активности и/или белковых уровней гетерологичного полипептида после одной дозы двунаправленной конструкции в течение по меньшей мере 1, 2, 4 или 6 месяцев или по меньшей мере 1, 2, 3, 4 или 5 лет.

В некоторых вариантах осуществления экспрессия полипептида клеткой-хозяином (*in vitro* или *in vivo*) повышена по меньшей мере на 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70% или более относительно уровня, экспрессируемого контрольной клеткой-хозяином, в которую не вводили конструкцию, содержащую трансген. В некоторых вариантах осуществления экспрессия полипептида клеткой-хозяином (*in vitro* или *in vivo*) повышена по меньшей мере на 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75% или более относительно известного нормального уровня (например, уровня полипептида у здорового субъекта). В некоторых вариантах осуществления экспрессия полипептида клеткой-хозяином (*in vitro* или *in vivo*) повышена по меньшей мере до около 1 мкг/мл, 2 мкг/мл, 3 мкг/мл, 4 мкг/мл, 5 мкг/мл, 6 мкг/мл, 7 мкг/мл, 8 мкг/мл, 9 мкг/мл, 10 мкг/мл, 15 мкг/мл, 20 мкг/мл, 25 мкг/мл, 30 мкг/мл, 35 мкг/мл, 40 мкг/мл, 45 мкг/мл, 50 мкг/мл, 55 мкг/мл, 60 мкг/мл, 65 мкг/мл, 70 мкг/мл, 75 мкг/мл, 80 мкг/мл, 85 мкг/мл, 90 мкг/мл, 95 мкг/мл, 100 мкг/мл, 120 мкг/мл, 140 мкг/мл, 160 мкг/мл, 180 мкг/мл, 200 мкг/мл, 225 мкг/мл, 250 мкг/мл, 275 мкг/мл, 300 мкг/мл, 325 мкг/мл, 350 мкг/мл, 400 мкг/мл, 450 мкг/мл, 500 мкг/мл, 550 мкг/мл, 600 мкг/мл, 650 мкг/мл, 700 мкг/мл, 750 мкг/мл, 800 мкг/мл, 850 мкг/мл, 900 мкг/мл, 1000 мкг/мл, 1100 мкг/мл, 1200 мкг/мл, 1300 мкг/мл, 1400 мкг/мл, 1500 мкг/мл, 1600 мкг/мл, 1700 мкг/мл, 1800 мкг/мл, 1900 мкг/мл, 2000 мкг/мл или более по определению, например, в клетке, плазме и/или сыворотке субъекта.

В некоторых вариантах осуществления в данном документе предложен способ лечения связанного с печенью нарушения в соответствии с описанными в данном документе способами. В контексте данного документа «связанное с печенью нарушение»

относится к нарушениям, которые могут непосредственно вызывать повреждение ткани печени, нарушений в результате повреждения ткани печени и/или нарушений отличных от печени органов или тканей, возникших в результате дефектов в печени.

В некоторых вариантах осуществления двунаправленную конструкцию, гидовую РНК и РНК-направляемый ДНК-связывающий агент вводят отдельно или в любой комбинации, местно или системно, например, внутривенно. В некоторых вариантах осуществления двунаправленную конструкцию, гидовую РНК и РНК-направляемый ДНК-связывающий агент вводят отдельно или в любой комбинации в печеночное кровообращение.

В некоторых вариантах осуществления хозяин или субъект представляет собой млекопитающее. В некоторых вариантах осуществления хозяин или субъект представляет собой человека. В некоторых вариантах осуществления хозяин или субъект представляет собой примата. В некоторых вариантах осуществления хозяин или субъект представляет собой грызуна (например, мышь или крысу), корову, свинью, обезьяну, овцу, собаку, кошку, рыбу или домашнюю птицу.

Это описание и типовые варианты осуществления не следует рассматривать как ограничивающие. В целях данного описания и прилагаемой формулы изобретения, если не указано иное, все числа, выражающие количества, проценты или доли, и другие числовые значения, используемые в описании и формуле изобретения, следует понимать как модифицируемые во всех случаях термином «около» при условии, что они еще не модифицированы таким образом. Соответственно, если не указано иное, числовые параметры, приведенные в нижеследующем описании и прилагаемой формуле изобретения, являются приближениями, которые могут варьироваться в зависимости от необходимых свойств, которые должны быть получены. По меньшей мере и без попытки ограничить применение доктрины эквивалентов объемом формулы изобретения, каждый числовой параметр следует толковать по меньшей мере в контексте количества приведенных значимых цифр и с применением обычных методов округления.

ПРИМЕРЫ

Нижеприведенные примеры представлены для иллюстрации определенных описанных вариантов осуществления и их не следует воспринимать как ограничивающие каким-либо образом объем данного изобретения.

Пример 1- Материалы и методы

Клонирование и получение плазмиды

Двунаправленная вставочная конструкция, фланкируемая ИКП, была синтезирована и клонировали в pUC57-Кап коммерческим поставщиком. Полученную в результате конструкцию (P00147) использовали в качестве исходного клонирующего вектора для других векторов. Другие вставочные конструкции (без ИКП) также были коммерчески синтезированы и клонированы в pUC57. Очищенную плазмиду расщепляли рестрикционным ферментом BglII (New England BioLabs, кат. № R0144S), и клонировали вставочные конструкции в исходный вектор. Плазмиду размножали в химически

компетентных ^{E.coli} Stb13 TM (Thermo Fisher, кат. № C737303).

Получение AAV

Тройную трансфекцию клеток HEK293 использовали для упаковки геномов представляющими интерес конструкциями для получения AAV8 и AAVDJ, а полученные в результате векторы очищали как из лизированных клеток, так и из культуральной среды методом ультрацентрифугирования в градиенте йодиксанола (смотрите, например, Lock et al., Hum Gene Ther. 2010 Oct;21(10):1259-71). Плазмиды, используемые в тройной трансфекции, которые содержали геном с представляющими интерес конструкциями, обозначены в примерах по номеру «PXXXX», также смотрите, например, таблицу 11. Выделенный AAV подвергали диализу в буфере для хранения (ФСБ с 0,001% плюроники F68). Титр AAV определяли с помощью кПЦР, используя праймеры/зонд, локализованные в области ИКП.

In vitro транскрипция («IVT») мРНК нуклеазы

Кэпированную и полиаденилированную мРНК Cas9 Streptococcus pyogenes («Spy»), содержащую N1-метил псевдо-U, создавали посредством транскрипции *in vitro* транскрипции, используя линейаризованную плазмидную ДНК-матрицу и РНК-полимеразу T7. В общем случае плазмидную ДНК, содержащую промотор T7 и область поли-(A/T) из 100 нуклеотидов, линейаризовали путем инкубации при 37°C с XbaI для полного расщепления с последующей тепловой инактивацией XbaI при 65 °C. Линейаризованную плазмиду очищали от фермента и буферных солей. Реакцию IVT для создания модифицированной мРНК Cas9 инкубировали при 37°C в течение 4 часов в следующих условиях: 50 нг/мкл линейаризованной плазмиды; 2 mM каждого из ГТФ, АТФ, ЦТФ и N1-метил псевдо-UTP (Trilink); 10 mM ARCA (Trilink); 5 E/мкл РНК-полимеразы T7 (NEB); 1 E/мкл мышинового ингибитора РНКазы (NEB); 0,004 E/мкл неорганической пирофосфатазы E. coli (NEB); и 1x реакционного буфера. ДНКазу TURBO (ThermoFisher) добавляли до конечной концентрации 0,01 E/мкл и инкубировали реакцию еще в течение 30 минут, чтобы удалить ДНК-матрицу. мРНК Cas9 очищали, используя набор для очистки MegaClear Transcription Clean-up в соответствии с протоколом производителя (ThermoFisher). В альтернативном варианте мРНК Cas9 очищали, используя осаждение LiCl, осаждение ацетатом аммония и осаждение ацетатом натрия или используя метод осаждения LiCl с последующей очисткой путем фильтрации с тангенциальным потоком. Концентрацию транскрипта определяли путем измерения поглощения света на 260 нм (Nanodrop), а транскрипт анализировали методом капиллярного электрофореза с помощью Bioanalyzer (Agilent).

Нижеприведенные мРНК Cas9 содержат OPC Cas9 SEQ ID NO: 703 или SEQ ID NO: 704 или последовательность из Таблицы 24 в PCT/US2019/053423 (которая включена в данный документ посредством ссылки).

Липидные составы для доставки мРНК Cas9 и гРНК

мРНК Cas9 и гРНК доставляли в клетки и организм животных, используя липидные составы, содержащие ионизируемый липид (9Z,12Z)-3-((4,4-

бис(октилокси)бутаноил)окси)-2-(((3-(диэтиламино)пропокси)карбонил)окси)метил)пропил)октадека-9,12-диеноат, также называемый 3-(((4,4-бис(октилокси)бутаноил)окси)-2-(((3-(диэтиламино)пропокси)карбонил)окси)метил)пропил(9Z,12Z) -октадека-9,12-диеноатом), холестерин, ДСФХ и ПЭГ2т-ДМГ.

Для экспериментов с использованием предварительно смешанных липидных составов (называемых в данном документе «липидными пакетами») компоненты восстанавливали в 100% этаноле в молярном соотношении ионизируемый липид:холестерин:ДСФХ:ПЭГ2т-ДМГ 50:38:9:3 перед смешиванием с нагрузкой РНК (например, мРНК Cas9 и гРНК) в молярном соотношении липидного амина к РНК-фосфату (N:P) около 6,0, как дополнительно описано в данном документе.

Для экспериментов с использованием компонентов, составленных в виде липидных наночастиц (ЛНЧ), компоненты растворяли в 100% этаноле в различных молярных соотношениях. РНК-карго (например, мРНК Cas9 и гРНК) растворяли в 25 мМ цитрате, 100 мМ NaCl, pH 5,0, что приводило к получению концентрации РНК-карго приблизительно 0,45 мг/мл.

Для экспериментов, описанных в примере 2, ЛНЧ формировали путем микрофлюидного смешивания растворов липидов и РНК, используя настольный прибор Precision Nanosystems NanoAssemblr™ Benchtop в соответствии с протоколом производителя. Во время смешивания поддерживали соотношение водного растворителя и органического растворителя 2:1, используя дифференциальные скорости потока. После смешивания ЛНЧ собирали, разводили водой (приблизительно 1:1 об./об.), выдерживали в течение 1 часа при комнатной температуре и дополнительно разводили водой (приблизительно 1:1 об./об.) перед окончательной заменой буфера. Окончательную замену буфера на 50 мМ Трис, 45 мМ NaCl, 5% (масс./об.) сахарозы, pH 7,5 (TSS), проводили на колонках для обессоливания PD-10 (GE). При необходимости составы концентрировали путем центрифугирования на 100 кДа центробежных фильтрах Amicon (Millipore). Потому полученную в результате смесь фильтровали, используя 0,2 мкм стерильный фильтр. Конечный ЛНЧ хранили при -80°C до дальнейшего использования. ЛНЧ составляли в молярном соотношении ионизируемый липид:холестерин:ДСФХ:ПЭГ2т-ДМГ, равном 45:44:9:2, с молярным отношением липидного амина к РНК-фосфату (N:P) около 4,5 и соотношением гРНК к мРНК 1:1 по массе.

Для экспериментов, описанных в других примерах, ЛНЧ получали, используя метод поперечного потока с использованием смешивания с ударной струей липида в этаноле с двумя объемами растворов РНК и одним объемом воды. Липид в этаноле смешивали с помощью смешивающей крестовины с двумя объемами раствора РНК. Четвертый поток воды смешивали с выходным потоком крестовины через встроенный тройник (смотрите WO2016010840, Фиг. 2). ЛНЧ выдерживали в течение 1 часа при комнатной температуре, а затем разводили водой (приблизительно 1:1 об./об.).

Разведенные ЛНЧ концентрировали, используя фильтрацию с тангенциальным потоком на плоском картридже (Sartorius, НОММ 100 кДа), а потом проводили замену буфера посредством диафильтрации в 50 мМ Трис, 45 мМ NaCl, 5% (масс./об.) сахарозы, pH 7,5 (TSS). В альтернативном варианте окончательную замену буфера на TSS проводили с помощью обессоливающих колонок PD-10 (GE). При необходимости составы концентрировали путем центрифугирования на 100 кДа центробежных фильтрах Amicon (Millipore). Потому полученную в результате смесь фильтровали, используя 0,2 мкм стерильный фильтр. Конечные ЛНЧ хранили при 4 °С или -80°С до дальнейшего использования. ЛНЧ составляли в молярном соотношении ионизируемый липид:холестерин:ДСФХ:ПЭГ2т-ДМГ, равном 50:38:9:3, с молярным отношением липидного амина к РНК-фосфату (N:P) около 6,0 и соотношением гРНК к мРНК 1:1 по массе.

Культура клеток и in vitro доставка мРНК Cas9, гРНК и вставочных конструкций

Клетки Нера1-6

Клетки Нера 1-6 высевали при плотности 10000 клеток/лунка в 96-луночные планшеты. Через 24 часа клетки обрабатывали ЛНЧ и AAV. Перед обработкой среду из лунок отсасывали. ЛНЧ разводили до 4 нг/мкл в среде DMEM+10% ФБС и дополнительно разводили до 2 нг/мкл в 10% ФБС (в DMEM) и инкубировали при 37°С в течение 10 мин (при конечной концентрации 5% ФБС). Целевая МЗ AAV составляла 1еб с разведением в среде DMEM+10% ФБС. 50 мкл вышеуказанных разбавленных ЛНЧ в 2 нг/мкл добавляли к клеткам (с доставкой в общей сложности 100 нг РНК-карго) с последующим добавлением 50 мкл AAV. Обработку ЛНЧ и AAV проводили с разницей в несколько минут. Общий объем среды в ячейках составлял 100 мкл. Через 72 часа после обработки и через 30 суток после обработки супернатант из этих обработанных клеток собирали для ELISA-анализа в отношении FIX человека, как описано ниже.

Первичные гепатоциты

Первичные гепатоциты мыши (ПГМ), первичные гепатоциты яванского макака (ПГЯ) и первичные гепатоциты человека (ПГЧ) размораживали и ресуспендировали в среде для размораживания гепатоцитов с добавками (ThermoFisher) с последующим центрифугированием. Супернатант сливали, а осажденные клетки ресуспендировали в среде для посева гепатоцитов с набором добавок (ThermoFisher). Клетки подсчитывали и высевали в покрытые коллагеном I Bio-coat 96-луночные планшеты при плотности 33000 клеток/лунка для ПГЧ, 50 000 клеток/лунка для ПГЯ и 15000 клеток/лунка для ПГМ. Высеянные клетки оставляли для осаждения и прикрепления на 5 часов в инкубаторе для тканевого культивирования при 37 °С и атмосфере с 5% CO₂. После инкубации клетки проверяли в отношении образования монослоя, трижды промывали с сохранением гепатоцитов и инкубировали при 37 °С.

Для экспериментов с использованием липидной пакетной доставки мРНК Cas9 и гРНК, каждую, отдельно разводили в 2 мг/мл в поддерживающей среде и добавляли по 2,9

мкл каждой в лунки (в 96-луночном планшете Эппендорфа), содержащие 12,5 мкл 50 мМ цитрата натрия, 200 мМ хлорида натрия при pH 5 и 6,9 мкл воды. После этого добавляли 12,5 мкл липидного пакетного состава с последующим добавлением 12,5 мкл воды и 150 мкл TSS. Каждую лунку разводили до 20 нг/мкл (по отношению к общему содержанию РНК), используя поддерживающую среду для гепатоцитов, а затем разводили до 10 нг/мкл (относительно общего содержания РНК) 6% свежей мышинной сывороткой. Среду отсасывали из клеток перед трансфекцией и добавляли к клеткам 40 мкл смесей липидный пакет/РНК с последующим добавлением AAV (разведенного в поддерживающей среде) при МЗ 1e5. Среду собирали через 72 часа после обработки для анализа, а клетки собирали для дополнительного анализа, как описано в данном документе.

Анализ с люциферазой

Для экспериментов, включающих обнаружение NanoLuc в клеточной среде, один объем субстрата для анализа люциферазы Nano-Glo® объединяли с 50 объемами буфера для анализа люциферазы Nano-Glo®. Анализ проводили на оборудовании Promega Glomax со временем интеграции 0,5 с, используя 1:10 разведение образцов (50 мкл реагента+40 мкл воды+10 мкл клеточной среды).

Для экспериментов, включающих обнаружение тега HiBit в клеточной среде, белок LgBiT и внеклеточный субстрат Nano-GloR HiBiT разводили 1:100 и 1:50, соответственно, во внеклеточном буфере Nano-GloR HiBiT при комнатной температуре. Анализ проводили на оборудовании Promega Glomax со временем интеграции 1,0 с, используя 1:10 разведение образцов (50 мкл реагента+40 мкл воды+10 мкл клеточной среды).

In vivo доставка ЛНЧ и/или AAV

Мышам вводили дозу AAV, ЛНЧ, как AAV, так и ЛНЧ, или носителей (ФСБ+0,001% плуроник для носителя AAV, TSS для носителя ЛНЧ) через боковую хвостовую вену. AAV вводили в объеме 0,1 мл на животное в количествах (векторных геномов/мышь, «вг/мш»), описанных в данном документе. ЛНЧ разводили в TSS и вводили в количествах, указанных в данном документе, в дозе около 5 мкл/грамм массы тела. Как правило, мышам сначала вводили AAV, а затем, если это было нужно, ЛНЧ. В различные моменты времени после обработки брали сыворотку и/или ткань печени для определенных анализов, как дополнительно описано ниже.

ELISA-анализ фактора IX человека (hFIX)

Для *in vitro* исследований определяли общие уровни фактора IX человека, секретируемого в клеточную среду, используя набор для ELISA для фактора IX человека (Abscam, кат. № ab188393) в соответствии с протоколом производителя. Уровни секретируемого hFIX определяли количественно по стандартной кривой, используя 4-параметрическую логистическую аппроксимацию, и выражали в нг/мл среды.

Для *in vivo* исследований собирали кровь и выделяли сыворотку, как указано. Общие сывороточные уровни фактора IX человека определяли, используя набор для ELISA для фактора IX человека (Abscam, кат. № ab188393) в соответствии с протоколом производителя. Сывороточные или плазменные уровни hFIX определяли количественно

по стандартной кривой, используя 4-параметрическую логистическую аппроксимацию, и выражали в мкг/мл сыворотки.

Секвенирование нового поколения («NGS») и анализ эффективности целевого расщепления

Для выявления наличия вставок и делеций, внесенных за счет генного редактирования, например, в интроне 1 альбумина, использовали глубинное секвенирование. ПЦР-праймеры конструировали вокруг целевого сайта и амплифицировали представляющую интерес область генома. Конструирование последовательности праймеров выполняли как принято в данной области.

Дополнительную ПЦР проводили в соответствии с протоколами производителя (Illumina) для добавления химии для секвенирования. Ампликоны секвенировали на приборе Illumina MiSeq. Чтения выравнивали с референсным геномом после исключения тех, которые имели низкие оценки качества. Полученные в результате файлы, содержащие чтения, картировали на референсный геном (файлы BAM), где выбирали чтения, которые перекрывали представляющую интерес целевую область, и рассчитывали число чтений дикого типа в сравнении с числом чтений, содержащих вставку или делецию («индель»).

Процент редактирования (например, «эффективность редактирования» или «процент редактирования») определяется как общее количество чтений последовательностей со вставками или делециями («инделями») по сравнению с общим числом чтений последовательностей, включая дикий тип.

Анализ гибридизации *in situ*

BaseScope (ACDbio, Ньюарк, Калифорния) представляет специализированную технологию гибридизации РНК *in situ*, которая может обеспечить специфическое обнаружение соединений экзонов, например, в гибридном транскрипте мРНК, который содержит вставочный трансген (hFIX) и кодирующую последовательность из сайта вставки (экзон 1 альбумина). BaseScope использовался для измерения процента клеток печени, экспрессирующих гибридную мРНК.

Чтобы обнаружить гибридную мРНК, ACDbio (Newark, CA) были разработаны два зонда к гибридным мРНК, которые могут появляться после вставки двунаправленной конструкции. Один из зондов был разработан для обнаружения гибридной мРНК, возникающей в результате вставки конструкции в одной ориентации, тогда как другой зонд был разработан для обнаружения гибридной мРНК, возникающей в результате вставки конструкции в другой ориентации. Получали печени от разных групп мышей и делали свежемороженые срезы. Анализ BaseScope с использованием одного зонда или объединенных зондов проводили в соответствии с протоколом производителя. Слайды сканировали и анализировали с помощью программного обеспечения HALO. Фон (группа, обработанная солевым раствором) в этом анализе составлял 0,58%.

Пример 2 - *in vitro* тестирование вставочных матриц с плечами гомологии и без них

В этом примере клетки Hep1-6 культивировали и обрабатывали AAV, несущими

вставочные матрицы различных форм (например, имеющие одноцепочечный геном («ssAAV») или самокомплементарный геном («scAAV»)), в присутствии или в отсутствие ЛНЧ, доставляющих мРНК Cas9 и G000551, например, как описано в примере 1 (n=3). AAV и ЛНЧ получали, как описано в примере 1. После обработки собирали среду для оценки экспрессии трансгена (например, уровней фактора IX человека), как описано в примере 1.

Клетки Нера1-6 представляют собой иммортализованную линию клеток печени мышей, которая продолжает делиться в культуре. Как проиллюстрировано на Фиг. 2 (72 часа после обработки), только вектор (scAAV, полученный из плазмиды P00204), содержащий 200 п. о. плечи гомологии, приводил к обнаруживаемой экспрессии hFIX. Использование AAV-векторов, полученных из P00123 (scAAV с отсутствием плеч гомологии) и P00147 (двунаправленная конструкция ssAAV с отсутствием плеч гомологии), не приводило к какой-либо обнаруживаемой экспрессии hFIX в этом эксперименте. Клетки поддерживали в культуре, и эти результаты были подтверждены при повторном анализе через 30 суток после обработки (данные не приведены).

Пример 3 - in vivo тестирование вставочных матриц с плечами гомологии и без них

В этом примере мышей обрабатывали AAV, полученным из тех же плазмид (P00123, P00204 и P00147), которые тестировали in vitro в примере 2. Дозируемые материалы получали и вводили, как описано в примере 1. Мышам C57Bl/6 вводили вводили (n=5 для каждой группы) 3e11 векторных геномов каждой (вг/мш), после этого - ЛНЧ, содержащие G000551 («G551»), в дозе 4 мг/кг (по отношению к общему содержанию РНК-карго). Через четыре недели после введения дозы животных умерщвляли и собирали ткань печени и сыворотку для анализа редактирования и экспрессии трансгена (например, hFIX), соответственно.

Как проиллюстрировано на Фиг. 3А и в таблице 12, уровни редактирования в печени ~ 60% были обнаружены в каждой группе животных, получавших ЛНЧ, содержащую гРНК, нацеленную на интрон 1 мышинового альбумина. Однако, несмотря на устойчивые и согласующиеся уровни редактирования в каждой группе обработки, животные, получавшие вектор ssAAV без плеч гомологии (вектор ssAAV, полученный из P00147) в комбинации с обработкой ЛНЧ, демонстрировали наибольший уровень экспрессии hFIX в сыворотке (Фиг. 3В и таблица 13).

Таблица 12: % инделей

Матрица	Среднее значение, индели (%)	Ст. откл., индели (%)
scAAV с тупым концом (P00123)	66,72	4,09
ssAAV с тупым концом (P00147)	68,10	2,27

ssAAV с ГР (P00204)	70,16	3,68
Только ЛНЧ	68,24	6,47
Носитель	0,28	0,08

Таблица 13: Уровни фактора IX

Матрица	Среднее значение, фактор IX (мкг/мл)	Ст. откл., фактор IX (мкг/мл)
scAAV с тупым концом (P00123)	0,75	0,28
ssAAV с тупым концом (P00147)	2,92	1,04
ssAAV с ГР (P00204)	0,96	0,35
Только ЛНЧ	0	0
Носитель	0	0

Пример 4 - in vivo тестирование вставочных матриц ssAAV с плечами гомологии и без них

Эксперимент, описанный в этом примере был направлен на изучение эффекта включения плеч гомологии в ssAAV-векторы in vivo.

Дозируемые материалы, используемые в этом эксперименте, получали и вводили, как описано в примере 1. Мышам C57Bl/6 вводили вводили (n=5 для каждой группы) 3e11 вг/мш, после этого - ЛНЧ, содержащие G000666 («G666») или G000551 («G551»), в дозе 0,5 мг/кг (по отношению к общему содержанию РНК-карго). Через четыре недели после введения дозы сыворотку животных собирали для анализа экспрессии трансгена (например, hFIX).

Как проиллюстрировано на Фиг. 4А и в таблице 14, использование ssAAV-векторов с асимметричными плечами гомологии (300/600 п. о. плечи, 300/2000 п. о. плечи и 300/1500 п. о. плечи для векторов, полученных из плазмид P00350, P00356 и P00362, соответственно) для вставки в сайт-мишень G551 приводило к уровням циркулирующего hFIX, которые были меньше нижнего предела обнаружения для анализа. Однако использование ssAAV-вектора (полученного из P00147) без плеч гомологии и с двумя открытыми рамками считывания (ОРС) hFIX в двунаправленной ориентации привело к обнаруживаемым уровням циркулирующего hFIX у каждого животного.

Аналогично, использование ssAAV-векторов с симметричными плечами гомологии (500 п. о. плечи и 800 п. о. плечи для векторов, полученных из плазмид P00353 и P00354, соответственно) для вставки в сайт-мишень G666, приводило к более низким, но обнаруживаемым уровням по сравнению с использованием двунаправленного вектор без плеч гомологии (полученного из P00147) (смотрите Фиг. 4В и таблицу 15).

Таблица 14: Сывороточные уровни FIX

AAV	Среднее значение,	Ст. откл., сывороточный FIX
-----	-------------------	-----------------------------

	сывороточный FIX (мкг/мл)	(мкг/мл)
P00147	5,13	1,31
P00350	-0,22	0,08
P00356	-0,23	0,04
P00362	-0,09	0,16

Таблица 15: Сывороточные уровни FIX

AAV	Среднее значение, сывороточный FIX (мкг/мл)	Ст. откл., сывороточный FIX (мкг/мл)
P00147	7,72	4,67
P00353	0,20	0,23
P00354	0,46	0,26

Пример 5 - in vitro скрининг двунаправленных конструкций по целевым сайтам в первичных гепатоцитах мышей

Продемонстрировав, что двунаправленные конструкции с отсутствием плеч гомологии, превосходят векторы с другими конфигурациями, в описанном в этом примере эксперименте исследовали эффекты изменения модулей двунаправленной конструкции, в данном случае - OPC и акцепторов сплайсинга, и изменения гРНК для нацеливания CRISPR/Cas9-опосредованной вставки. Эти разнообразные двунаправленные конструкции тестировали на панели целевых сайтов, используя 20 разных гРНК, нацеленных на интрон 1 мышинового альбумина, в первичных гепатоцитах мышей (ПГМ).

Материалы для доставки ssAAV и липидных пакетов, тестируемых в этом примере, получали и доставляли в ПГМ, как описано в примере 1, с AAV при 1×10^5 . После обработки выделенную геномную ДНК и клеточную среду собирали для анализа редактирования и экспрессии трансгена, соответственно. Каждый из векторов содержал репортер, который можно измерить с помощью регистрации флуоресценции люциферазы, как описано в примере 1, которая отображена на графике на Фиг. 5С в виде относительных единиц люциферазы («ОЕЛ»). Например, AAV-векторы, содержащие OPC hFIX, содержали пептид HiBit, слитый с их 3' концами, а AAV-вектор, содержащий только репортерные гены, содержал OPC NanoLuc (помимо ЗФБ). Схематические изображения каждого из тестируемых векторов представлены на Фиг. 5А. Протестированные гРНК проиллюстрированы на Фиг. 5В и 5С с использованием сокращенных обозначений для тех, которые перечислены в таблице 4 (например, опущены первые нули, например, «G551» соответствует «G000551» в таблице 4).

Как проиллюстрировано на Фиг. 5В и в таблице 16, были обнаружены согласующиеся, но разные уровни редактирования для каждой из групп обработки для каждой тестируемой комбинации. Экспрессия трансгена при применении различных комбинаций матрицы и гидовой РНК проиллюстрирована на Фиг. 5С и в таблице 17. Как проиллюстрировано на Фиг. 5D, существенный уровень образования инделей не

обязательно приводит к более эффективной экспрессии трансгенов. При применении полученных из P00411 и P00418 матриц значения R^2 составляли 0,54 и 0,37, соответственно, при этом направляющие с редактированием менее 10% не включены. Акцептор сплайсинга мышинового альбумина и акцептор сплайсинга человеческого FIX приводили к эффективной экспрессии трансгена. Что интересно, несмотря на разные ОРС и акцепторы сплайсинга, относительные уровни экспрессии, измеренные в ОЕЛ, были сопоставимыми в трех тестируемых векторах, демонстрируя надежность, воспроизводимость и модульность системы двунаправленной конструкции (смотрите Фиг. 5С).

Таблица 16: % инделей

ID направляющей	P00411		P00418		P00415	
	Среднее значение, индели (%)	Ст. откл., индели (%)	Среднее значение, индели (%)	Ст. откл., индели (%)	Среднее значение, индели (%)	Ст. откл., индели (%)
G000551	67,4	1,42	70,67	2,29	66,73	4,90
G000552	90,93	0,15	91,10	2,43	90,37	1,01
G000553	77,80	3,83	77,47	1,87	80,50	0,85
G000554	72,37	6,49	70,53	3,16	70,60	2,91
G000555	35,37	2,63	35,77	9,34	40,47	4,75
G000666	62,47	3,87	50,90	19,41	65,90	3,99
G000667	30,57	2,73	25,30	3,67	31,67	2,29
G000668	63,60	2,02	66,65	4,60	68,30	4,90
G000669	19,10	2,51	19,33	1,53	18,70	1,25
G000670	47,80	3,27	49,10	4,42	51,97	2,06
G011722	4,20	0,72	4,27	1,20	4,20	0,26
G011723	5,63	1,27	6,07	0,15	5,93	0,15
G011724	6,10	1,28	8,50	2,69	7,13	1,27
G011725	1,93	0,29	2,60	0,79	2,53	0,65
G011726	10,73	1,46	11,70	0,50	12,43	1,33
G011727	14,20	1,56	14,80	2,36	16,20	2,69
G011728	10,55	1,20	13,65	0,92	15,50	1,56
G011729	5,00	0,10	5,63	0,25	6,00	1,01
G011730	7,83	0,97	9,13	0,59	7,33	0,59
G011731	23,70	0,66	25,27	1,21	24,87	1,01

Только AAV	0,15	0,07	0,05	0,07	0,10	0,00
------------	------	------	------	------	------	------

Таблица 17: Уровни люциферазы

ID направляющей	P00411		P00418		P00415	
	Среднее значение, люцифера за (ОЕЛ)	Ст. откл., люцифера за (ОЕЛ)	Среднее значение, люцифера за (ОЕЛ)	Ст. откл., люцифера за (ОЕЛ)	Среднее значение, люцифера за (ОЕЛ)	Ст. откл., люцифера за (ОЕЛ)
G000551	58000,00	4331,28	41800,00	2165,64	78633,33	20274,70
G000552	95700,00	10573,08	80866,67	27911,35	205333,33	30664,86
G000553	205333,33	52993,71	177333,33	32929,22	471666,67	134001,00
G000554	125333,33	55949,38	91933,33	19194,10	232666,67	67002,49
G000555	59933,33	11566,04	77733,33	11061,80	155666,67	15947,83
G000666	88500,00	28735,87	93266,67	30861,19	313000,00	15394,80
G000667	75333,33	22653,11	68966,67	27222,11	153000,00	30805,84
G000668	164000,00	56320,51	133400,00	65111,29	429000,00	120751,80
G000669	28933,33	11636,29	22033,33	2413,16	46466,67	6543,19
G000670	162666,67	32959,57	200000,00	33867,39	424666,67	36473,73
G011722	16766,67	3384,28	8583,33	4103,10	24000,00	8915,16
G011723	22733,33	7252,82	17133,33	4905,44	26100,00	8109,87
G011724	17300,00	2400,00	28033,33	9091,94	30933,33	3365,02
G011725	8253,33	1163,20	8890,00	1429,27	20366,67	13955,05
G011726	12223,33	3742,54	11610,00	2490,44	14950,00	8176,03
G011727	35600,00	8128,35	36300,00	12301,22	86700,00	5023,94
G011728	14900,00	5011,99	22466,67	7130,45	38166,67	13829,08
G011729	10460,00	2543,95	11223,33	2220,28	26966,67	16085,50
G011730	14833,33	2307,24	21700,00	8681,59	41233,33	25687,03
G011731	16433,33	3274,65	22566,67	2205,30	20756,67	13096,20
Только AAV	217,00	15,56	215,00	15,56	207,00	1,41

Пример 6 - in vivo скрининг двунаправленных конструкций по целевым сайтам

SsAAV и ЛНЧ, тестируемые в этом примере, получали и доставляли мышам C57B1/6, как описано в примере 1, чтобы оценить эффективность двунаправленных конструкций в целевых сайтах in vivo. Через четыре недели после введения дозы животных умерщвляли и собирали ткань печени и сыворотку для анализа редактирования и экспрессии трансгена (например, hFIX), соответственно.

В первоначальном эксперименте 10 разных составов ЛНЧ, содержащих 10 разных гРНК, нацеленных на интрон 1 альбумина, доставляли мышам вместе с ssAAV, полученным из P00147. AAV и ЛНЧ доставляли при 3×10^{11} вг/мш и 4 мг/кг (по отношению к общему содержанию РНК-карго), соответственно. гРНК, тестируемые в этом эксперименте, проиллюстрированы на Фиг. 6 и в таблице 18. Как проиллюстрировано на Фиг. 6 и как наблюдалось *in vitro*, существенный уровень образования инделей не был прогностическим в отношении вставки или экспрессии трансгенов.

В отдельном эксперименте полную панель из 20 гРНК, нацеленных на 20 разных целевых сайтов, протестированных *in vitro* в примере 5, тестировали *in vivo*. С этой целью составы 20 ЛНЧ, содержащие 20 гРНК, нацеленных на интрон 1 альбумина, доставляли мышам вместе с ssAAV, полученным из P00147. AAV и ЛНЧ доставляли при 3×10^{11} вг/мш и 1 мг/кг (по отношению к общему содержанию РНК-карго), соответственно. гРНК, тестируемые в этом эксперименте, проиллюстрированы на Фиг. 7А и 7В и в таблицах 19 и 20, с использованием сокращенных обозначений для тех, которые перечислены в таблице 4.

Как проиллюстрировано на Фиг. 7А, были обнаружены различные уровни редактирования для каждой из групп обработки для каждой тестируемой комбинации ЛНЧ/вектор. Однако, как проиллюстрировано на Фиг. 7В и в соответствии с данными *in vitro*, описанными в примере 5, более высокие уровни редактирования не обязательно приводят к более высоким уровням экспрессии трансгенов *in vivo*, что указывает на отсутствие корреляции между редактированием и вставкой/экспрессией двунаправленных конструкций. Действительно, существует очень небольшая корреляция между уровнем обеспечиваемого редактирования и уровнем экспрессии трансгена (hFIX), как проиллюстрировано на графике, представленном на Фиг. 7D. В частности, в этом эксперименте для наборов данных по редактированию и экспрессии получено значение R^2 , составляющее всего 0,34, при этом гРНК, демонстрирующее менее 10% редактирования, были удалены из анализа. Интересно, как проиллюстрировано на Фиг. 7С, представлен график корреляции, на котором приведено сравнение уровней экспрессии, измеренных в ОЕЛ, из *in vitro* эксперимента примера 5, с уровнями экспрессии трансгена, обнаруженными *in vivo* в этом эксперименте, со значением R^2 , равным 0,70, что демонстрирует положительную корреляцию между первичным клеточным скринингом и обработкой *in vivo*.

Чтобы оценить вставку двунаправленной конструкции на клеточном уровне, анализировали ткани печени обработанных животных, используя метод гибридизации *in situ* (BaseScore), например, как описано в примере 1. В этом анализе использовали зонды, которые могут обнаруживать соединения между трансгеном hFIX и последовательностью экзона 1 мышинового альбумина в виде гибридного транскрипта. Как проиллюстрировано на Фиг. 8А, клетки, положительные в отношении гибридного транскрипта, были обнаружены у животных, которые получали как AAV, так и ЛНЧ. В частности, когда вводили только AAV, менее 1,0% клеток были положительными в отношении гибридного

транскрипта. При введении ЛНЧ, содержащих G011723, G000551 или G000666, 4,9%, 19,8% или 52,3% клеток были положительными в отношении гибридного транскрипта. Кроме того, как проиллюстрировано на Фиг. 8В и в таблице 14, уровни циркулирующего hFIX коррелировали с числом клеток, положительных в отношении гибридного транскрипта. Наконец, в анализе использовали объединенные зонды, которые могут обнаруживать вставку двунаправленной конструкции в любой ориентации. Однако при применении одного зонда, который может обнаруживать только одну ориентацию, количество клеток, положительных в отношении гибридного транскрипта, составляло около половины от количества, обнаруженного при применении объединенных зондов (в одном примере 4,46% и 9,68%), что позволяет предположить, что двунаправленная конструкция действительно способна к вставке в любой ориентации, приводя к экспрессии гибридных транскриптов, что коррелирует с экспрессией трансгена на белковом уровне.

Таблица 18: Уровни фактора IX и % инделей

Направляющая	Среднее значение, индели (%)	Ст. откл., индели (%)	Среднее значение, люцифераза (ОЕЛ)	Ст. откл., люцифераза (ОЕЛ)
G000551	75,02	1,27	3,82	3,38
G000555	51,18	1,19	32,56	9,05
G000553	62,78	2,64	25,07	4,04
G000667	52,96	4,96	32,03	6,74
G000554	55,24	2,28	29,48	7,34
G000552	67,56	1,73	14,79	5,34
G000668	43,14	5,78	26,72	7,97
G000669	50,68	2,97	10,70	4,43
G000666	64,62	1,34	26,19	5,56
G000670	55,90	1,30	30,96	8,44

Таблица 19: % редактирования в печени

Направляющая	Среднее значение, редактирование в печени (%)	Ст. откл., редактирование в печени (%)
G000551	59,48	4,02
G000555	58,72	3,65
G000553	51,26	2,81
G000554	33,04	8,76
G000555	12,72	4,46

G000666	53,60	4,92
G000667	26,74	4,98
G000668	39,22	3,04
G000669	33,34	4,77
G000670	47,50	5,58
G011722	10,34	1,68
G011723	4,02	0,84
G011724	2,46	0,64
G011725	8,26	1,24
G011726	6,90	1,01
G011727	13,33	6,43
G011728	35,78	9,34
G011729	4,62	1,46
G011730	12,68	3,14
G011731	26,70	1,86

Таблица 20: Уровни FIX

Направляющая	1 неделя		2 неделя		4 неделя	
	Среднее значение, FIX (мкг/мл)	Ст. откл., FIX (мкг/мл)	Среднее значение, FIX (мкг/мл)	Ст. откл., FIX (мкг/мл)	Среднее значение, FIX (мкг/мл)	Ст. откл., FIX (мкг/мл)
G000551	10,88	2,74	10,25	2,51	9,39	3,48
G000555	13,34	2,09	12,00	2,75	12,43	2,57
G000553	17,64	4,34	20,27	6,35	15,31	2,43
G000554	12,79	4,99	14,29	6,09	12,74	4,93
G000555	11,94	5,79	11,99	5,76	8,61	4,02
G000666	21,63	1,32	20,65	1,55	17,23	0,62
G000667	16,77	2,86	12,35	2,85	12,57	5,60
G000668	21,35	1,51	18,20	3,18	17,72	2,25
G000669	5,76	2,10	6,72	2,93	3,39	0,78
G000670	18,18	2,17	19,16	3,05	15,49	3,61
G011722	8,07	1,74	7,74	2,41	8,07	1,74
G011723	2,11	0,28	1,65	0,28	2,11	0,28
G011724	0,92	0,43	0,60	0,30	0,92	0,43

G011725	1,75	0,77	1,14	0,67	1,75	0,77
G011726	0,59	0,30	1,01	0,64	0,59	0,30
G011727	6,71	2,80	6,90	3,68	6,71	2,80
G011728	11,77	3,12	12,29	3,43	11,77	3,12
G011729	0,94	0,35	0,89	0,29	0,94	0,35
G011730	5,93	1,77	6,33	1,73	5,93	1,77
G011731	3,56	0,87	3,78	0,50	3,56	0,87
Только AAV	0,00	0,00	0,00	0,00	0,00	0,00
Носитель	0,00	0,00	0,00	0,00	0,00	0,00
Сыворотка человека	3,63	0,32	3,61	0,35	3,28	0,03

Пример 7 - продолжительность экспрессии hFIX in vivo

В этом примере у обработанных животных оценивали продолжительность экспрессии hFIX по времени. С этой целью измеряли hFIX в сыворотке обработанных животных после введения дозы в рамках однолетнего исследования продолжительности.

ssAAV и ЛНЧ, тестируемые в этом примере, получали и доставляли мышам C57Bl/6, как описано в примере 1. Состав ЛНЧ содержал G000551, а ssAAV был получен из P00147. AAV доставляли при 3e11 вг/мш, а ЛНЧ доставляли при 0,25 или 1,0 мг/кг (по отношению к общему содержанию РНК-карго) (n=5 для каждой группы).

Как проиллюстрировано на Фиг. 9А и 9В и в таблицах 21 и 22, экспрессия hFIX сохранялась в каждый момент времени, оцениваемый для обеих групп, до 41 или 52 недели, соответственно. Считается, что снижение уровней, наблюдаемое через 8 недель на Фиг. 9А, связано с вариабельностью анализа ELISA. Уровни сывороточного альбумина измеряли с помощью ELISA на 2 и 41 неделе, что показало, что уровни циркулирующего альбумина поддерживаются на протяжении всего исследования.

Таблица 21: Уровни hFIX

Неделя	Доза			
	0,25 мг/кг ЛНЧ		1 мг/кг ЛНЧ	
	Среднее значение, hFIX (мкг/мл)	Ст. откл., hFIX (мкг/мл)	Среднее значение, hFIX (мкг/мл)	Ст. откл., hFIX (мкг/мл)
2	0,48	0,21	2,24	1,12
4	0,55	0,18	2,82	1,67
8	0,40	0,17	1,72	0,77
12	0,48	0,20	2,85	1,34
20	0,48	0,27	2,45	1,26
41	0,79	0,49	4,63	0,95

Таблица 22: Уровни hFIX

Неделя	Доза			
	0,25 мг/кг ЛНЧ		1 мг/кг ЛНЧ	
	Среднее значение, hFIX (мкг/мл)	Ст. откл., hFIX (мкг/мл)	Среднее значение, hFIX (мкг/мл)	Ст. откл., hFIX (мкг/мл)
2	0,87	0,15	4,02	1,75
8	0,99	0,15	4,11	1,41
12	0,93	0,14	4,15	1,35
20	0,83	0,22	4,27	1,54
41	0,83	0,37	4,76	1,62
52	0,82	0,25	4,72	1,54

Пример 8 - эффекты различных доз AAV и ЛНЧ для модуляции экспрессии hFIX in vivo

В этом примере эффекты изменения дозы как AAV, так и ЛНЧ для модуляции экспрессии hFIX оценивали на мышах C57Bl/6.

ssAAV и ЛНЧ, тестируемые в этом примере, получали и доставляли мышам, как описано в примере 1. Состав ЛНЧ содержал G000553, а ssAAV был получен из P00147. AAV доставляли при $1e11$, $3e11$, $1e12$ или $3e12$ вг/мш, а ЛНЧ доставляли при 0,1, 0,3 или 1,0 мг/кг (по отношению к общему содержанию РНК-карго) ($n=5$ для каждой группы). Через две недели после введения дозы животных умерщвляли. Сыворотку собирали в два момента времени для анализа экспрессии hFIX.

Как проиллюстрировано на Фиг. 10А (1 неделя), Фиг. 10В (2 недели) и таблице 23, изменение дозы AAV или ЛНЧ может модулировать степень экспрессии hFIX in vivo.

Таблица 23: Сывороточный hFIX

Момент времени	Доза РНП (мг/кг)	Доза AAV (МЗ)	Среднее значение, FIX (мкг/мл)	СО	N
1 неделя	0,1	1E+11	0,08	0,02	2
		3E+11	0,11	0,04	5
		1E+12	0,41	0,15	5
		3E+12	0,61	0,17	5
	0,3	1E+11	0,36	0,14	5
		3E+11	0,67	0,26	5
		1E+12	1,76	0,14	5
		3E+12	4,70	2,40	5
1,0	1E+11	3,71	0,31	4	

Момент времени	Доза РНП (мг/кг)	Доза AAV (МЗ)	Среднее значение, FIX (мкг/мл)	СО	N
		3E+11	8,00	0,51	5
		1E+12	14,17	1,38	5
		3E+12	20,70	2,79	5
		Сыворотка человека 1:1000	6,62	-	1
2 неделя	0,1	1E+11	0,12	0,01	2
		3E+11	0,26	0,07	5
		1E+12	0,83	0,24	5
		3E+12	1,48	0,35	5
	0,3	1E+11	0,70	0,26	4
		3E+11	1,42	0,37	5
		1E+12	3,53	0,49	5
		3E+12	8,94	4,39	5
	1,0	1E+11	5,40	0,47	4
		3E+11	12,31	2,45	5
		1E+12	17,89	1,95	5
		3E+12	25,52	3,62	5
		Сыворотка человека 1:1000	4,47	-	1

Пример 9 - in vitro скрининг двунаправленных конструкций по целевым сайтам в первичных гепатоцитах яванского макака и человека

В этом примере ssAAV-векторы, содержащие двунаправленную конструкцию, тестировали на панели целевых сайтов, используя гРНК, нацеленные на интрон 1 альбумина яванского макака («яв. м.») и человека в первичных гепатоцитах яванского макака (ПГЯ) и первичных гепатоцитах человека (ПГЧ), соответственно.

Материалы для доставки ssAAV и липидных пакетов, тестируемых в этом примере, получали и доставляли в ПГЯ и ПГЧ, как описано в примере 1. После обработки выделенную геномную ДНК и клеточную среду собирали для анализа редактирования и экспрессии трансгена, соответственно. Каждый из векторов содержал репортер, который можно измерить с помощью регистрации флуоресценции люциферазы, как описано в примере 1 (полученных из P00415), которая отображена на графиках на Фиг. 11В и 12В в виде относительных единиц люциферазы («ОЕЛ»). Например, AAV-векторы содержали ОРС NanoLuc (в дополнение к ЗФБ). Схематические изображения тестируемых векторов представлены на Фиг. 11В и 12В. Протестированные гРНК проиллюстрированы на каждой из фигур с использованием сокращенного номера для тех, которые перечислены в

таблице 1 и таблице 7.

Как проиллюстрировано на Фиг. 11А для ПГЯ и на Фиг. 12А для ПГЧ, были обнаружены различные уровни редактирования для каждой из тестируемых комбинаций (данные по редактированию для некоторых комбинаций, которые тестировали в эксперименте ПГЯ, не представлены на Фиг. 11А и в таблице 1 из-за неудачных экспериментов с определенными парами праймеров, используемых для секвенирования ампликона). Данные по редактированию, представленные на Фиг. 11А и 12А графически, численно отображены в таблице 1 и таблице 2 ниже. Однако, как проиллюстрировано на Фиг. 11В, 11С и Фиг. 12В и 12С, существенный уровень образования инделей не был прогностическим в отношении вставки или экспрессии трансгенов, что указывает на слабую корреляцию между редактированием и вставкой/экспрессией двунаправленных конструкций в ПГЯ и ПГЧ, соответственно. В качестве одного из показателей, значение R^2 , рассчитанное на Фиг. 11С, составляет 0,13, а значение R^2 на Фиг. 12D составляет 0,22.

Таблица 1: Данные по редактированию и экспрессии трансгена интрона 1 альбумина для оГРНК, доставленных в первичные гепатоциты яванского макака

ИД направляющей	Средн. % редактирования	Ст. откл. % редактирования	Средн. ОЕЛ	Ст. откл. ОЕЛ
G009867	25,05	0,21	10650,67	1455,97
G009866	18,7	3,96	75556,67	12182,98
G009876	14,85	4,88	27463,33	10833,53
G009875	12,85	2,33	51660,00	6362,36
G009874	28,25	6,01	270433,30	133734,10
G009873	42,65	5,59	178600,00	87607,25
G009865	59,15	0,21	301666,70	18610,03
G009872	48,15	3,46	320233,30	63517,43
G009871	46,5	5,23	211966,70	65852,44
G009864	33,2	8,34	210033,30	61201,33
G009863	54,8	12,45	69853,33	15216,92
G009862	44,6	7,21	508666,70	119876,30
G009861	28,65	0,21	178666,70	15821,93
G009860	33,2	7,07	571333,30	52728,87
G009859	0,05	0,07	258333,30	79052,73
G009858	14,65	1,77	402333,30	25579,94
G009857	23	0,99	312333,30	73036,52
G009856	14,8	0,99	95900,00	21128,42

G009851	1,5	0,42	105766,70	27048,91
G009868	12,15	2,47	43033,33	9141,85
G009850	63,45	13,93	228200,00	101542,10
G009849	57,55	8,27	225400,00	46001,30
G009848	33	5,37	156333,30	20647,84
G009847	66,75	7	100866,70	22159,72
G009846	61,85	5,02	31766,67	10107,59
G009845	54,4	7,5	43020,00	11582,23
G009844	47,15	2,05	110466,70	32031,44

Таблица 2: Данные по редактированию и экспрессии трансгена интрона 1 альбумина для оГРНК, доставленных в первичные гепатоциты человека

ID направляющей	Средн. % редактирования	Ст. откл. % редактирования	Средн. ОЕЛ	Ст. откл. ОЕЛ
G009844	19,07	2,07	268333,30	80432,17
G009851	0,43	0,35	18033,33	2145,54
G009852	47,20	3,96	18400,00	2251,67
G009857	0,10	0,14	71100,00	14609,24
G009858	8,63	9,16	32000,00	18366,55
G009859	3,07	3,50	59500,00	16014,99
G009860	18,80	4,90	190333,30	54307,76
G009861	10,27	2,51	62233,33	9865,26
G009866	13,60	13,55	96200,00	46573,81
G009867	12,97	3,04	3916,67	1682,03
G009868	0,63	0,32	10176,67	2037,80
G009874	49,13	0,60	318000,00	114118,40
G012747	3,83	0,23	51000,00	6161,17
G012748	1,30	0,35	17433,33	2709,86
G012749	9,77	1,50	75066,67	11809,04
G012750	42,73	4,58	5346,67	2977,35
G012751	7,77	1,16	32066,67	18537,62
G012752	32,93	2,27	402000,00	83144,45
G012753	21,20	2,95	71800,00	32055,73
G012754	0,60	0,10	16933,33	4254,80
G012755	1,10	0,10	13833,33	3685,56

G012756	2,17	0,40	35600,00	6055,58
G012757	1,07	0,25	13993,33	6745,08
G012758	0,90	0,10	34900,00	15308,82
G012759	2,60	0,35	30566,67	15287,36
G012760	39,10	6,58	6596,67	2133,13
G012761	36,17	2,43	467666,70	210965,20
G012762	8,50	0,57	217000,00	13000,00
G012763	47,07	3,07	142333,30	37581,02
G012764	44,57	5,83	1423333,0 0	261023,60
G012765	19,90	1,68	179666,70	57011,69
G012766	8,50	0,28	243333,30	17473,79

Кроме того, ssAAV-векторы, содержащие двунаправленную конструкцию, тестировали на панели целевых сайтов, используя одинарные гидовые РНК, нацеленные на интрон 1 альбумина человека, в первичных гепатоцитах человека (ПГЧ).

Материалы ssAAV и ЛНЧ получали и доставляли в ПГЧ, как описано в примере 1. После обработки выделенную геномную ДНК и клеточную среду собирали для анализа редактирования и экспрессии трансгена, соответственно. Как описано выше, каждый из векторов (полученных из P00415) содержал репортер, который можно измерить с помощью регистрации флуоресценции люциферазы, как описано в примере 1, которая отображена на графике на Фиг. 12С и приведена в таблице 23 в виде относительных единиц люциферазы («ОЕЛ»). Например, AAV-векторы содержали ОРС NanoLuc (в дополнение к ЗФБ). Схематические изображения тестируемых векторов представлены на Фиг. 11В и 12В. Протестированные гРНК проиллюстрированы на Фиг. 12С с использованием сокращенного номера для тех, которые перечислены в таблице 1 и таблице 7.

Таблица 23: Данные по экспрессии трансгена интрона 1 альбумина для огРНК, доставленных в первичные гепатоциты человека

Направляющая	Среднее значение, люцифераза (ОЕЛ)	Ст. откл., люцифераза (ОЕЛ)
G009844	3 700 000	509 117
G009852	281 000	69 296
G009857	1 550 000	127 279
G009858	551 000	108 894

G009859	1 425 000	77 782
G009860	2 240 000	183 848
G009861	663 500	238 295
G009866	274 000	11 314
G009867	44 700	566
G009874	2 865 000	431 335
G012747	651 000	59 397
G012749	867 000	93 338
G012752	4 130 000	268 701
G012753	1 145 000	162 635
G012757	579 000	257 387
G012760	129 000	36 770
G012761	4 045 000	728 320
G012762	2 220 000	127 279
G012763	1 155 000	205 061
G012764	11 900 000	1 555 635
G012765	1 935 000	134 350
G012766	2 050 000	169 706
ЛНЧ	8 430	212

Пример 10 - in vivo тестирование экспрессии фактора IX из альтернативного безопасного приемного локуса

В этом примере оценивали вставку ssAAV, содержащего двунаправленную конструкцию hFIX, в альтернативный безопасный приемный локус. Чтобы протестировать вставку в альтернативный безопасный приемный локус, получали AAV, как описано выше. Мышам вводили AAV в дозе 3e11 вг/мышь, за чем сразу следовало введение ЛНЧ,

составленных с мРНК Cas9 и гидовыми РНК в дозе 0,3 мг/кг. Через 4 недели после введения дозы животных умерщвляли и собирали образцы печени и крови. Редактирование в образцах печени определяли методом NGS. Уровни hFIX человека в сыворотке определяли методом ELISA. Данные NGS и ELISA показали эффективную вставку и экспрессию hFIX в альтернативном безопасном приемном локусе.

Пример 11 - in vivo тестирование вставки гена фактора IX человека у отличных от человека приматов

В этом примере было проведено 8-недельное исследование для оценки вставки гена фактора IX человека и экспрессии белка hFIX у яванских макаков посредством введения аденоассоциированного вируса (AAV) и/или липидных наночастиц (ЛНЧ) с различными направляющими. Это исследование проводили с составами ЛНЧ и составами AAV, полученными, как описано выше. Каждый состав ЛНЧ содержал мРНК Cas9 и гидовую РНК (гРНК) с соотношением мРНК:гРНК 2:1 по массе. ssAAV был получен из P00147.

Самцов яванских макаков обрабатывали в группах по n=3. Животным вводили дозу AAV путем медленной болюсной инъекции или инфузии в дозах, описанных в таблице 3. После обработки AAV животные получали буфер или ЛНЧ, как описано в таблице 3, посредством медленного болюса или инфузии.

Через две недели после введения дозы собирали образцы печени с помощью однократной чрескожной биопсии под ультразвуковым контролем. Каждый образец биопсии мгновенно замораживали в жидком азоте и хранили при температуре от -86 до -60 °C. Анализ редактирования образцов печени проводили с помощью секвенирования NGS, как описано ранее.

Для ELISA-анализа фактора IX у животных брали образцы крови на 7, 14, 28 и 56 сутки после введения дозы. Образцы крови собирали, перерабатывали в плазму после взятия крови и хранили при температуре от -86 до -60°C до анализа.

Общие уровни фактора IX человека определяли из образцов плазмы методом ELISA. Вкратце, 96-луночный микропланшет Reacti-Bind (VWR, кат. № PI15041) покрывали захватывающим антителом (мышинное mAB к фактору IX человека (HTI, кат. № ANIX-5041)) в концентрации 1 мкг/мл, затем блокировали, используя 1х ФСБ с 5% альбумина бычьей сыворотки. Затем тестируемые образцы или стандарты очищенного белка фактора IX человека (ERL, кат. № HFIX 1009, № партии HFIX4840), разведенного в плазме яванского макака, инкубировали в отдельных лунках. Антитела для обнаружения (поликлональные овечьи антитела к фактору 9 человека, Abscam, кат. № ab128048) адсорбировали в концентрации 100 нг/мл. Вторичное антитело (ослиное анти-овечьё IgG pAb с HRP, Abscam, кат. № ab97125) использовали в концентрации 100 нг/мл. Набор реагентов с ТМБ-субстратом (BD OptEIA кат. № 555214) использовали для проявления планшета. Оптическую плотность оценивали спектрофотометрически на 450 нм на микропланшет-ридере (система Molecular Devices i3) и анализировали с помощью SoftMax pro 6.4.

Было обнаружено образование инделей, подтверждающее, что редактирование произошло. Данные NGS показали эффективное образование инделей. Экспрессию hFIX из локуса альбумина в ОЧП измеряли методом ELISA, а результаты приведены в таблице 4 и на Фиг. 13. Уровни hFIX в плазме достигли уровней, которые были ранее описаны как терапевтически эффективные (George, et al., NEJM 377(23), 2215-27, 2017).

По данным измерения уровни циркулирующего белка hFIX сохранялись на протяжении восьминедельного исследования (смотрите Фиг. 13, на которой проиллюстрированы средние уровни на 7, 14, 28 и 56 сутки, составляющие ~135, ~140, ~150 и ~110 нг/мл, соответственно) с достижением уровня белка от ~75 нг/мл до ~250 нг/мл. Плазменные уровни hFIX рассчитывали, используя удельную активность, которая в ~8 раз выше для гиперфункционального варианта hFIX R338L ((Simioni et al., NEJM 361(17), 1671-75, 2009) (в которой сообщается о белок-специфической активности hFIX-R338L, составляющей 390 ± 28 Е на миллиграмм, и белок-специфической активности для фактора IX дикого типа, составляющей $45 \pm 2,4$ Е на миллиграмм). Вычисления функционально нормализованной активности фактора IX для гиперфункционального варианта фактора IX, тестируемого в этом примере, показали, что эксперимент позволил достичь стабильных уровней белка фактора IX человека в ОЧП в течение 8-недельного исследования, что соответствует около 20-40% активности фактора IX дикого типа (диапазон составляет 12-67% активности фактора IX дикого типа).

Таблица 3: Редактирование в печени

ID животного	ID направляющей	F9-AAV (вг/кг)	Объем F9-AAV (мл/кг)	ЛПЧ (мг/кг)	Объем ЛПЧ (мл/кг)
4001	G009860	3E+13	1	3	2
4002	G009860	3E+13	1	3	2
4003	G009860	3E+13	1	3	2
5001	TSS	3E+13	1	0	0
5002	TSS	3E+13	1	0	0
5003	TSS	3E+13	1	0	0
6001	G009862	0	0	3	2
6002	G009862	0	0	3	2
6003	G009862	0	0	3	2

Таблица 4: Экспрессия hFIX

ID животного	Сутки 7 Фактор IX (нг/мл)	Сутки 14 Фактор IX (нг/мл)	Сутки 28 Фактор IX (нг/мл)	Сутки 56 Фактор IX (нг/мл)
4001	122,84/+2,85	94,93/+0,56	105,65/+1,94	97,31/+1,49
4002	149,77/+13,5	222,92/+9,61	252,49/+6,46	152,05/+7,46
4003	134,06/+6,17	107,04/+6,46	95,30/+3,18	74,23/+3,53
5001	Н/Д	Н/Д	Н/Д	Н/Д
5002	Н/Д	Н/Д	Н/Д	Н/Д
5003	Н/Д	Н/Д	Н/Д	Н/Д
6001	Н/Д	Н/Д	Н/Д	Н/Д
6002	Н/Д	Н/Д	Н/Д	Н/Д
6003	Н/Д	Н/Д	Н/Д	Н/Д

Пример 12 - in vivo тестирование вставки фактора IX у отличных от человека приматов

В этом примере проводили исследование для оценки вставки гена фактора IX и экспрессии белка hFIX у яванских макаков после введения ssAAV, полученного из липидных наночастиц P00147 и/или CRISPR/Cas9, с различными направляющими, содержащими G009860 и различные компоненты ЛНЧ.

Образование инделей определяли с помощью NGS, подтверждая, что редактирование произошло. Общие уровни фактора IX человека определяли в образцах плазмы методом ELISA, используя мышинные mAB к фактору IX человека (HTI, кат. № ANIX-5041), поликлональные овечьи антитела к фактору 9 человека (Abscam, кат. № ab128048) и ослиные антиовечьи IgG pAb с HRP (Abscam, кат. № ab97125), как описано в примере 11. Уровни белка FIX человека, в > 3 раз превышающие уровни, достигнутые в эксперименте примера 13, были получены из двунаправленной матрицы с использованием альтернативных ЛНЧ CRISPR/Cas9. В этом исследовании результаты анализа ELISA показывают, что уровни циркулирующего белка hFIX, находящиеся на уровне или превышающие нормальный диапазон уровней FIX человека (3-5 мкг/мл; Amiral et al., Clin. Chem., 30(9), 1512-16, 1984), были достигнуты при применении G009860 в ОЧП как минимум на 14 и 28 сутки. Первоначальные данные показали, что уровни циркулирующего белка FIX человека составляют ~ 3-4 мкг/мл на 14 сутки после однократной дозы, при этом уровни сохранялись в течение первых 28 суток (~ 3-5 мкг/мл) исследования.

Уровни циркулирующего альбумина измеряли методом ELISA и показали, что исходные уровни альбумина сохранялись через 28 суток. Исследуемые уровни альбумина у необработанных животных в исследовании варьировались $\pm \sim 15\%$. У обработанных животных уровни циркулирующего альбумина изменялись минимально и не выходили за

пределы нормального диапазона, при этом уровни восстанавливались до исходного уровня в течение одного месяца.

Уровни циркулирующего белка FIX человека также определяли методом сэндвич-иммуноанализа с большим динамическим диапазоном. Вкратце, 96-луночный планшет MSD GOLD со стрептавидином SECTOR (Meso Scale Diagnostics, кат. L15SA-1) блокировали 1% блокирующим агентом ECL (Sigma, GERPN2125). После удаления блокирующего раствора на планшете иммобилизовали биотинилированное захватывающее антитело (Sino Biological, 11503-R044). Рекомбинантный белок FIX человека (Enzyme Research Laboratories, HFIX 1009) использовали для получения калибровочного стандарта в 0,5% блокирующем агенте ECL. После промывки в планшет добавляли калибровочные стандарты и образцы плазмы и инкубировали. После промывки в лунки добавляли антитело для обнаружения (Haematologic Technologies, ANIX-5041), конъюгированное с сульфо-тегом, и инкубировали. После вымывания всех несвязанных антител для обнаружения на лунки наносили буфер для считывания T. без какой-либо дополнительной инкубации планшет визуализировали на приборе MSD Quick Plex SQ120 и анализировали данные с помощью пакета программного обеспечения Discovery Workbench 4,0 (Meso Scale Discovery). Концентрации выражены как средние рассчитанные концентрации в мкг/мл. Для образцов N=3, если не указано звездочкой, в этом случае N=2. Экспрессия hFIX из локуса альбумина в обработанной исследуемой группе, измеренная с помощью MSD ELISA, приведена в таблице 24.

Таблица 24: Сывороточные уровни белка фактора IX человека

	Средняя расщ. конц. (мкг/мл)		
	3001	3002	3003
Сутки 7	7,85	5,63	11,20
Сутки 14	8,65	11,06	14,70
Сутки 28	9,14	14,12	10,85
Сутки 42	9,03	33,12*	13,22
Сутки 56	10,24	16,72	33,84*

Пример 13 - Нецелевой анализ направляющих альбумина человека

Использовали биохимический метод (смотрите, например, Cameron et al., Nature Methods. 6, 600-606; 2017) для определения потенциальных нецелевых геномных сайтов, расщепляемых Cas9, нацеленной на альбумин. В этом эксперименте проводили скрининг 13 огРНК, нацеленных на человеческий альбумин, и две контрольные направляющие с известными нецелевыми профилями, используя выделенную геномную ДНК НЕК293. Число потенциальных нецелевых сайтов, обнаруженных с использованием концентрации направляющих 16 нМ в биохимическом анализе, приведено в таблице 26. Анализ позволил идентифицировать потенциальные нецелевые сайты для тестируемых огРНК.

Таблица 25: Нецелевой анализ

ID гРНК	Мишень	Направляющая последовательность (SEQ ID NO:)	Число нецелевых сайтов
G012753	Альбумин	GACUGAAACUUCACAGAAUA (SEQ ID NO: 20)	62
G012761	Альбумин	AGUGCAAUGGAUAGGUCUUU (SEQ ID NO: 28)	75
G012752	Альбумин	UGACUGAAACUUCACAGAAU (SEQ ID NO: 19)	223
G012764	Альбумин	CCUCACUCUUGUCUGGGCAA (SEQ ID NO: 31)	3985
G012763	Альбумин	UGGGCAAGGGAAGAAAAAAA (SEQ ID NO: 30)	5443
G009857	Альбумин	AUUUAUGAGAUCAACAGCAC (SEQ ID NO: 5)	131
G009859	Альбумин	UUAAAUAAGCAUAGUGCAA (SEQ ID NO: 7)	91
G009860	Альбумин	UAAAGCAUAGUGCAAUGGAU (SEQ ID NO: 8)	133
G012762	Альбумин	UGAUUCCUACAGAAAAACUC (SEQ ID NO: 29)	68
G009844	Альбумин	GAGCAACCUCACUCUUGUCU (SEQ ID NO: 2)	107
G012765	Альбумин	ACCUCACUCUUGUCUGGGCA (SEQ ID NO: 32)	41
G012766	Альбумин	UGAGCAACCUCACUCUUGUC (SEQ ID NO: 33)	78
G009874	Альбумин	UAAUAAAUAUCAACAUCU (SEQ ID NO: 13)	53
G000644	EMX1	GAGUCCGAGCAGAAGAAGAA (SEQ ID NO: 1129)	304
G000645	VEGFA	GACCCCUCCACCCCGCCUC (SEQ ID NO: 1130)	1641

В известных анализах обнаружения нецелевых объектов, таких как биохимический метод, который использовали выше, большое количество потенциальных нецелевых

сайтов обычно восстанавливают при конструировании, чтобы «создать широкую сеть» для потенциальных сайтов, которые можно проверить в другом контексте, например, в представляющей интерес первичной клетке. Например, в биохимическом методе, как правило, число потенциальных нецелевых сайтов представлено в избытке, поскольку в анализе используется очищенная высокомолекулярная геномная ДНК, не содержащая клеточное окружение, а анализ зависит от используемой дозы РНП Cas9. Соответственно, потенциальные нецелевые сайты, идентифицированные этими методами, можно подтверждать, используя нацеленное секвенирование идентифицированных потенциальных нецелевых сайтов.

Пример 14 - Конструирование конструкций для экспрессии секреторных или несекреторных белков

Конструкции, такие как двунаправленные конструкции, можно конструировать так, чтобы они экспрессировали секреторные или несекреторные белки. Для выработки секреторного белка конструкция может содержать сигнальную последовательность, которая помогает при транслокации полипептида в просвет ЭР. В альтернативном варианте в конструкции может использоваться эндогенная сигнальная последовательность клетки-хозяина (например, эндогенная сигнальная последовательность альбумина, когда трансген интегрирован в локус альбумина клетки-хозяина).

В противоположность этому, конструкции для экспрессии несекреторных белков могут быть сконструированы так, чтобы они не содержали сигнальную последовательность и, следовательно, не использовали эндогенную сигнальную последовательность клетки-хозяина. Некоторые методы, посредством которых это можно обеспечить, включают внесение в конструкцию последовательности сайта внутренней посадки рибосомы (IRES). Последовательности IRES, такие как EMCV IRES, позволяют инициировать трансляцию из любой позиции мРНК непосредственно ниже расположения IRES. Это делает возможной экспрессию белка, в котором отсутствует эндогенная сигнальная последовательность клетки-хозяина, из сайта вставки, который содержит сигнальную последовательность выше (например, сигнальная последовательность экзона 1 локуса альбумина не будет включена в экспрессируемый белок). В отсутствие сигнальной последовательности белок секретироваться не будет. Примеры последовательностей IRES, которые можно использовать в конструкции, включают последовательности из пикорнавирусов (например, FMDV), пестивирусов (CFFV), полиовирусов (PV), вирусов энцефаломиокардита (EMCV), вирусов болезни рук, ног и рта (FMDV), вирусов гепатита С (HCV), вирусов классической чумы свиней (CSFV), вируса лейкоза мышей (MLV), вирусов иммунодефицита обезьян (SIV) или вирусов паралича сверчка (CrPV).

Альтернативным подходом для экспрессии несекреторных белков является включение в конструкцию одного или более саморасщепляющихся пептидов выше представляющего интерес полипептида. Саморасщепляющийся пептид, такой как 2A или 2A-подобные последовательности, служит сигналом проскока рибосомы для получения

множества отдельных белков из одного мРНК-транскрипта. Как показано в плазмиде с ID P00415 из таблицы 11, можно использовать саморасщепляющийся пептид (например, P2A) для создания бицистронного вектора, который экспрессирует два трансгена (например, нанолуциферазы и ЗФБ). В альтернативном варианте саморасщепляющийся пептид можно использовать для экспрессии белка, в котором отсутствует эндогенная сигнальная последовательность клетки-хозяина (например, последовательность 2A, расположенная выше представляющего интерес белка, приведет к расщеплению между эндогенной сигнальной последовательностью альбумина и представляющим интерес белком). Типовые пептиды 2A, которые можно использовать, приведены в таблице 12. Кроме того, в 5'-конец пептида можно добавлять остатки (GSG) для повышения эффективности расщепления, как показано в таблице 12.

Пептид	Аминокислотная последовательность
T2A (SEQ ID NO: 1131)	EGRGSLTTCGDVEENPGP
P2A (SEQ ID NO: 1132)	ATNFSLLKQAGDVEENPGP
E2A (SEQ ID NO: 1133)	QCTNYALLKLAGDVESNPGP
F2A (SEQ ID NO: 1134)	VKQTLNFDLLKLAGDVESNPGP
T2A с остатками GSG (SEQ ID NO: 1135)	GSGEGRGSLTTCGDVEENPGP
P2A с остатками GSG (SEQ ID NO: 1136)	GSGATNFSLLKQAGDVEENPGP
E2A с остатками GSG (SEQ ID NO: 1137)	GSGQCTNYALLKLAGDVESNPGP
F2A с остатками GSG (SEQ ID NO: 1138)	GSGVKQTLNFDLLKLAGDVESNPGP

Таблица 5: Последовательности гидовой РНК человека

ID направляющей	Направляющая последовательность	Геномные координаты	SEQ ID NO:
G009844	GAGCAACCUCACUCUUGUCU	xp4:73405113-73405133	2
G009851	AUGCAUUUGUUUCAAAAUAU	xp4:73405000-73405020	3
G009852	UGCAUUUGUUUCAAAAUAAU	xp4:73404999-73405019	4
G009857	AUUUAUGAGAUCAACAGCAC	xp4:73404761-73404781	5
G009858	GAUCAACAGCACAGGUUUUG	xp4:73404753-73404773	6
G009859	UUAAAUAAAGCAUAGUGCAA	xp4:73404727-73404747	7
G009860	UAAAGCAUAGUGCAAUGGAU	xp4:73404722-73404742	8
G009861	UAGUGCAAUGGAUAGGUCUU	xp4:73404715-73404735	9
G009866	UACUAAAACUUUAUUUUACU	xp4:73404452-73404472	10
G009867	AAAGUUGAACAAUAGAAAAA	xp4:73404418-73404438	11
G009868	AAUGCAUAAUCUAAGUCAAA	xp4:73405013-73405033	12

G009874	UAAUAAAAUUCAACAUCU	xp4:73404561-73404581	13
G012747	GCAUCUUUAAAGAAUUAUUU	xp4:73404478-73404498	14
G012748	UUUGGCAUUUAUUUCUAAAA	xp4:73404496-73404516	15
G012749	UGUAUUUGUGAAGUCUUACA	xp4:73404529-73404549	16
G012750	UCCUAGGUAAAAAAAAAAAA	xp4:73404577-73404597	17
G012751	UAAUUUUCUUUUGCGCACUA	xp4:73404620-73404640	18
G012752	UGACUGAAACUUCACAGAAU	xp4:73404664-73404684	19
G012753	GACUGAAACUUCACAGAAUA	xp4:73404665-73404685	20
G012754	UUCAUUUUAGUCUGUCUUCU	xp4:73404803-73404823	21
G012755	AUUAUCUAAGUUUGAAUAUA	xp4:73404859-73404879	22
G012756	AAUUUUUAAAAUAGUAUUCU	xp4:73404897-73404917	23
G012757	UGAAUUAUUCUUCUGUUUAA	xp4:73404924-73404944	24
G012758	AUCAUCCUGAGUUUUUCUGU	xp4:73404965-73404985	25
G012759	UUACUAAAACUUUAUUUUAC	xp4:73404453-73404473	26
G012760	ACCUUUUUUUUUUUUUUACCU	xp4:73404581-73404601	27
G012761	AGUGCAAUGGAUAGGUCUUU	xp4:73404714-73404734	28
G012762	UGAUUCCUACAGAAAACUC	xp4:73404973-73404993	29
G012763	UGGGCAAGGGAAGAAAAAAA	xp4:73405094-73405114	30
G012764	CCUCACUCUUGUCUGGGCAA	xp4:73405107-73405127	31
G012765	ACCUCACUCUUGUCUGGGCA	xp4:73405108-73405128	32
G012766	UGAGCAACCUCACUCUUGUC	xp4:73405114-73405134	33

Таблица 6: Последовательности гидовой РНК мыши

ID направляющей	Направляющая последовательность	Геномные координаты	SEQ ID NO:
G000551	AUUUGCAUCUGAGAACCCUU	xp5:90461148-90461168	98
G000552	AUCGGGAACUGGCAUCUUCA	xp5:90461590-90461610	99
G000553	GUUACAGGAAAUCUGAAGG	xp5:90461569-90461589	100
G000554	GAUCGGGAACUGGCAUCUUC	xp5:90461589-90461609	101
G000555	UGCAUCUGAGAACCCUUAGG	xp5:90461151-90461171	102
G000666	CACUCUUGUCUGUGGAAACA	xp5:90461709-90461729	103
G000667	AUCGUUACAGGAAAUCUGA	xp5:90461572-90461592	104
G000668	GCAUCUUCAGGGAGUAGCUU	xp5:90461601-90461621	105
G000669	CAAUCUUUAAAUAUGUUGUG	xp5:90461674-90461694	106
G000670	UCACUCUUGUCUGUGGAAAC	xp5:90461710-90461730	107

G011722	UGCUGUAUUUUUCUAGUAA	xp5:90461039-90461059	108
G011723	GUAAAUAUCUACUAAGACAA	xp5:90461425-90461445	109
G011724	UUUUUCUAGUAAUGGAAGCC	xp5:90461047-90461067	110
G011725	UUAUAUUUAUGAUUAUUUU	xp5:90461174-90461194	111
G011726	GCACAGAUUAUAAACACUAAA	xp5:90461480-90461500	112
G011727	CACAGAUUAUAAACACUUAAC	xp5:90461481-90461501	113
G011728	GGUUUUAAAAAUAAUAUGU	xp5:90461502-90461522	114
G011729	UCAGAUUUUCCUGUAACGAU	xp5:90461572-90461592	115
G011730	CAGAUUUUCCUGUAACGAUC	xp5:90461573-90461593	116
G011731	CAAUGGUAAAUAAGAAAUA	xp5:90461408-90461428	117
G013018	GGAAAUCUGAAGGUGGCAA	xp5:90461563-90461583	118
G013019	GGCGAUCUCACUCUUGUCUG	xp5:90461717-90461737	119

Таблица 7: Последовательности гидовой РНК яванского макака

ID направляющей	Направляющая последовательность	Геномные координаты	SEQ ID NO:
G009844	GAGCAACCUCACUCUUGUCU	xp5:61198711-61198731	164
G009845	AGCAACCUCACUCUUGUCUG	xp5:61198712-61198732	165
G009846	ACCUCACUCUUGUCUGGGGA	xp5:61198716-61198736	166
G009847	CCUCACUCUUGUCUGGGGAA	xp5:61198717-61198737	167
G009848	CUCACUCUUGUCUGGGGAAG	xp5:61198718-61198738	168
G009849	GGGGAAGGGGAGAAAAAAAAA	xp5:61198731-61198751	169
G009850	GGGAAGGGGAGAAAAAAAAA	xp5:61198732-61198752	170
G009851	AUGCAUUUGUUUCAAAAUAU	xp5:61198825-61198845	171
G009852	UGCAUUUGUUUCAAAAUUU	xp5:61198826-61198846	172
G009853	UGAUUCCUACAGAAAAGUC	xp5:61198852-61198872	173
G009854	UACAGAAAAGUCAGGAUAA	xp5:61198859-61198879	174
G009855	UUUCUUCUGCCUUAAACAG	xp5:61198889-61198909	175
G009856	UUAUAGUUUUAUUAUCAAAC	xp5:61198957-61198977	176
G009857	AUUUAUGAGAUAACAGCAC	xp5:61199062-61199082	177
G009858	GAUCAACAGCACAGGUUUUG	xp5:61199070-61199090	178
G009859	UUAAAUAAGCAUAGUGCAA	xp5:61199096-61199116	179
G009860	UAAAGCAUAGUGCAAUGGAU	xp5:61199101-61199121	180
G009861	UAGUGCAAUGGAUAGGUCUU	xp5:61199108-61199128	181
G009862	AGUGCAAUGGAUAGGUCUUA	xp5:61199109-61199129	182

G009863	UUACUUUGCACUUUCCUUAG	xp5:61199186-61199206	183
G009864	UACUUUGCACUUUCCUUAGU	xp5:61199187-61199207	184
G009865	UCUGACCUUUUAUUUUACCU	xp5:61199238-61199258	185
G009866	UACUAAAACUUUAUUUUACU	xp5:61199367-61199387	186
G009867	AAAGUUGAACAAUAGAAAA	xp5:61199401-61199421	187
G009868	AAUGCAUAAUCUAAGUCAA	xp5:61198812-61198832	188
G009869	AUUAUCCUGACUUUUUCUGU	xp5:61198860-61198880	189
G009870	UGAAUUAUUCUCUGUUUAA	xp5:61198901-61198921	190
G009871	UAAUUUUCUUUUGCCCACUA	xp5:61199203-61199223	191
G009872	AAAAGGUCAGAAUUGUUUAG	xp5:61199229-61199249	192
G009873	AACAUCCUAGGUAAAAUAAA	xp5:61199246-61199266	193
G009874	UAAUAAAUAUCAACAUCU	xp5:61199258-61199278	194
G009875	UUGUCAUGUAUUUCUAAAAU	xp5:61199322-61199342	195
G009876	UUUGUCAUGUAUUUCUAAAA	xp5:61199323-61199343	196

Таблица 8. орPHK альбумина человека и профили модификации

ID направляющей	Полная последовательность	SEQ ID NO:	Модифицированная полная последовательность	SEQ ID NO:
G009844	GAGCAACCUCA CUCUUGUCUGU UUU AGAGCUAGAAA UAGCAAGUUA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	34	mG*mA*mG*CAACCUCACUC UUGUCUGU UUUAGAmGmCmUmAmGmA mAmAmUm AmGmCAAGUUAAAAUAAGG CUAGUCC GUUAUCAmAmCmUmUmGmA mAmAmAm AmGmUmGmGmCmAmCmCm GmAmGmUm CmGmGmUmGmCmU*mU*mU *mU	66
G009851	AUGCAUUUGUU UCAAAAUAUGU UUU AGAGCUAGAAA	35	mA*mU*mG*CAUUUGUUUCA AAUAUG UUUUAGAmGmCmUmAmGmA mAmAmUm	67

	UAGCAAGUUA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU		AmGmCAAGUUAAAAUAAGG CUAGUCCG UUAUCAmAmCmUmUmGmAm AmAmAmAm GmUmGmGmCmAmCmCmGmA mGmUmCm GmGmUmGmCmU*mU*mU*m U	
G009852	UGCAUUUGUUU CAAAUAUUGU UUU AGAGCUAGAAA UAGCAAGUUA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	36	mU*mG*mC*AUUUGUUUCA AAUAUUGU UUUAGAmGmCmUmAmGmAm AmAmUmAm GmCAAGUUAAAAUAAGGCU AGUCCGUUA UCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGmU mGmCmU*mU*mU*mU	68
G009857	AUUUAUGAGAU CAACAGCACGU UUU AGAGCUAGAAA UAGCAAGUUA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	37	mA*mU*mU*UAUGAGAUCAA CAGCACGU UUUAGAmGmCmUmAmGmAm AmAmUmAm GmCAAGUUAAAAUAAGGCU AGUCCGUUA UCAmAmCmUmUmGmAmAmA mAmAmGm UmGmGmCmAmCmCmGmAmG mUmCmGmGmUmGmCmU*mU *mU*mU	69
G009858	GAUCAACAGCA CAGGUUUUGGU UUU	38	mG*mA*mU*CAACAGCACAG GUUUUGGU UUUAGAmGmCmUmAmGmAm	70

	AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU		AmAmUmAm GmCAAGUUAAAAUAAGGCU AGUCCGUUA UCAmAmCmUmUmGmAmAmA mAmAmGm UmGmGmCmAmCmCmGmAmG mUmCmGm GmUmGmCmU*mU*mU*mU	
G009859	UUAAAUAAGC AUAGUGCAAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	39	mU*mU*mA*AAUAAAGCAUA GUGCAAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	71
G009860	UAAAGCAUAGU GCAAUGGAUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	40	mU*mA*mA*AGCAUAGUGCA AUGGAUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	72
G009861	UAGUGCAAUGG AUAGGUCUUGU	41	mU*mA*mG*UGCAAUGGAUA GGUCUUGUUUUAGAmGmCm	73

	UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU		UmAmGmAmAmAmUmAmGm CAAGUUA AAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	
G009866	UACUAAAACUU UAUUUUACUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	42	mU*mA*mC*UAAAACUUUAU UUUACUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUA AAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	74
G009867	AAAGUUGAACA AUAGAAAAAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	43	mA*mA*mA*GUUGAACAAUA GAAAAAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUA AAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	75
G009868	AAUGCAUAAUC	44	mA*mA*mU*GCAUAAUCUAA	76

	UAAGUCAAAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU		GUCAAAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	
G009874	UAAUAAAAUUC AAACAUCCUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	45	mU*mA*mA*UAAAAUUCAAA CAUCCUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	77
G012747	GCAUCUUUAAA GAAUUAUUUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	46	mG*mC*mA*UCUUUAAAGAA UUAUUUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	78

G012748	UUUGGCAUUUA UUUCUAAAAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	47	mU*mU*mU*GGCAUUUAUUU CUAAAAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	79
G012749	UGUAUUUGUGA AGUCUUACAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	48	mU*mG*mU*AUUUGUGAAGU CUUACAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	80
G012750	UCCUAGGUAAA AAAAAAAAAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU	49	mU*mC*mC*UAGGUAAAAAA AAAAAAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	81

	UUU			
G012751	UAAUUUUCUUU UGCGCACUAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	50	mU*mA*mA*UUUUCUUUUGC GCACUAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	82
G012752	UGACUGAAACU UCACAGAAUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	51	mU*mG*mA*CUGAAACUUCA CAGAAUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	83
G012753	GACUGAAACUU CACAGAAUAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC	52	mG*mA*mC*UGAAACUUCAC AGAAUAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	84

	GAGUCGGUGCU UUU			
G012754	UUCAUUUUAGU CUGUCUUCUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	53	mU*mU*mC*AUUUUAGUCUG UCUUCUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	85
G012755	AUUAUCUAAGU UUGAAUAUAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	54	mA*mU*mU*AUCUAAGUUUG AAUAUAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	86
G012756	AAUUUUUAAAA UAGUAUUCUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA	55	mA*mA*mU*UUUUAAAAUAG UAUUCUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	87

	AAAGUGGCACC GAGUCGGUGCU UUU			
G012757	UGAAUUAUUCU UCUGUUUAAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	56	mU*mG*mA*AUUAUUCUUCU GUUUUAGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	88
G012758	AUCAUCCUGAG UUUUUCUGUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	57	mA*mU*mC*AUCCUGAGUUU UUCUGUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	89
G012759	UUACUAAAACU UUAUUUUACGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU	58	mU*mU*mA*CUAAAACUUUA UUUUACGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU*	90

	GAA AAAGUGGCACC GAGUCGGUGCU UUU		mU	
G012760	ACCUUUUUUUU UUUUUACCUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	59	mA*mC*mC*UUUUUUUUUU UUACCUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	91
G012761	AGUGCAAUGGA UAGGUCUUUGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	60	mA*mG*mU*GCAAUGGAUAG GUCUUUGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	92
G012762	UGAUUCCUACA GAAAAACUCGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC	61	mU*mG*mA*UCCUACAGAA AAACUCGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC	93

	GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU		mGmGmUmGmCmU*mU*mU* mU	
G012763	UGGGCAAGGGA AGAAAAAAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	62	mU*mG*mG*GCAAGGGAAGA AAAAAAGUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	94
G012764	CCUCACUCUUG UCUGGGCAAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	63	mC*mC*mU*CACUCUUGUCU GGGCAAGUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	95
G012765	ACCUCACUCUU GUCUGGGCAGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU	64	mA*mC*mC*UCACUCUUGUC UGGGCAGUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm	96

	AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU		GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	
G012766	UGAGCAACCUC ACUCUUGUCGU UUU AGAGCUAGAAA UAGCAAGUAAA AAU AAGGCUAGUCC GUUAUCAACUU GAA AAAGUGGCACC GAGUCGGUGCU UUU	65	mU*mG*mA*GCAACCUCACU CUUGUCGUUUUAGAmGmCm UmAmGmAmAmAmUmAmGm CAAGUAAAAUAAGGCUAG UCCGUUAUCAmAmCmUmUm GmAmAmAmAmAmGmUmGm GmCmAmCmCmGmAmGmUmC mGmGmUmGmCmU*mU*mU* mU	97

Таблица 9. Гидовая оРНК альбумина мыши и профиль модификации

ID направляющей	Полная последовательность	SEQ ID NO:	Модифицированная последовательность	полная SEQ ID NO:
G000551	AUUUGCAUCUG AGAACCCUUGU UUUAGAGCUAG AAAUAGCAAGU UAAAAUAAGGC UAGUCCGUUAU CAACUUGAAAA AGUGGCACCGA GUCGGUGCUUU U	120	mA*mU*mU*UGCAUCUGAGAA CCCUUGUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	142
G000552	AUCGGGAACUG GCAUCUUCA GUUUUAGAGCU	121	mA*mU*mC*GGGAACUGGCAU CUUCAGUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG	143

	AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU		UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	
G000553	GUUACAGGAAA AUCUGAAGG GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	122	mG*mU*mU*ACAGGAAAAUCU GAAGGGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	144
G000554	GAUCGGGAACU GGCAUCUUC GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	123	mG*mA*mU*CGGGAACUGGCA UCUUCGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	145
G000555	UGCAUCUGAGA ACCCUUAAGG GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU	124	mU*mG*mC*AUCUGAGAACCCU UAGGGUUUUAGAmGmCmUmA mGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUA UCAmAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmCmCm GmAmGmUmCmGmGmUmGmCm U*mU*mU*mU	146

	GGCACCGAGUC GGUGCUUUU			
G000666	CACUCUUGUCU GUGGAAACA GUUUUAGAGCU AGAAAUAGC AAGUUAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAGU GGCACCGAGUC GGUGCUUUU	125	mC*mA*mC*UCUUGUCUGUGG AAACAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	147
G000667	AUCGUUACAGG AAAUCUGA GUUUUAGAGCU AGAAAUAGC AAGUUAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAGU GGCACCGAGUC GGUGCUUUU	126	mA*mU*mC*GUUACAGGAAA UCUGAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	148
G000668	GCAUCUUCAGG GAGUAGCUU GUUUUAGAGCU AGAAAUAGC AAGUUAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAGU GGCACCGAGUC GGUGCUUUU	127	mG*mC*mA*UCUUCAGGGAGU AGCUUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	149
G000669	CAAUCUUUAAA UAUGUUGUG GUUUUAGAGCU	128	mC*mA*mA*UCUUUAAAUAUG UUGUGGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG	150

	AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU		UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	
G000670	UCACUCUUGUC UGUGGAAAC GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	129	mU*mC*mA*CUCUUGUCUGUG GAAACGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	151
G011722	UGCUGUAUUU UUCUAGUAA GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	130	mU*mG*mC*UUGUAUUUUUCU AGUAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	152
G011723	GUAAAU AUCUA CUAAGACAA GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU	131	mG*mU*mA*AAUAUCUACUAA GACAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	153

	GGCACCGAGUC GGUGCUUUU			
G011724	UUUUUCUAGUA AUGGAAGCC GUUUUAGAGCU AGAAAUAGC AAGUUAAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	132	mU*mU*mU*UUCUAGUAAUGG AAGCCGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	154
G011725	UUAUAUUAUUG AUAUAUUUU GUUUUAGAGCU AGAAAUAGC AAGUUAAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	133	mU*mU*mA*UAUUAUUGAUAU AUUUUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	155
G011726	GCACAGAUUA AACACUUA GUUUUAGAGCU AGAAAUAGC AAGUUAAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	134	mG*mC*mA*CAGAUUAAACA CUUAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	156
G011727	CACAGAUUAA ACACUUAAC GUUUUAGAGCU	135	mC*mA*mC*AGAUUAAACAC UUAACGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG	157

	AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU		UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	
G011728	GGUUUUA AAAA UAAUAAUGU GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	136	mG*mG*mU*UUUAAAAUAAU AAUGUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	158
G011729	UCAGAUUUUCC UGUAACGAU GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	137	mU*mC*mA*GAUUUCCUGUA ACGAUGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	159
G011730	CAGAUUUUCCU GUAACGAUC GUUUUAGAGCU AGAAAUAGC AAGUUA AAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU	138	mC*mA*mG*AUUUUCCUGUAA CGAUCGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	160

	GGCACCGAGUC GGUGCUUUU			
G011731	CAAUGGUAUUU AAGAAAUA GUUUUAGAGCU AGAAAUAGC AAGUUAAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	139	mC*mA*mA*UGGUAAAUAAGA AAUAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	161
G013018	GGAAAAUCUGA AGGUGGCAA GUUUUAGAGCU AGAAAUAGC AAGUUAAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	140	mG*mG*mA*AAAUCUGAAGGU GGCAAGUUUUAGAmGmCmUm AmGmAmAmAmUmAmGmCAAG UUAAAAUAAGGCUAGUCCGUU AUCAmAmCmUmUmGmAmAmA mAmAmGmUmGmGmCmAmCmC mGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	162
G013019	GGCGAUCUCAC UCUUGUCUG GUUUUAGAGCU AGAAAUAGC AAGUUAAAAUA AGGCUAGUC CGUUAUCAACU UGAAAAAGU GGCACCGAGUC GGUGCUUUU	141	mG*mG*mC*GAUCUCACUCUUG UCUGGUUUUAGAmGmCmUmA mGmAmAmAmUmAmGmCAAGU UAAAAUAAGGCUAGUCCGUUA UCAmAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmCmCm GmAmGmUmCmGmGmUmGmCm U*mU*mU*mU	163

Таблица 10: орПНК альбумина яванского макака и профили модификации

ID направл	Полная последовательность	SEQ ID	Модифицированная последовательность	полная SEQ ID
---------------	------------------------------	-----------	--	---------------------

яющей		NO:		NO:
G009844	GAGCAACCUCACU CUUGUCU GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	197	mG*mA*mG*CAACCUCACUCUUG UCUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAAU AAGGCUAGUCCGUUAUCA mAmCmUmUm GmAmAmAmAmAmGmUmGmGmC mAmCmCmGm AmGmUmCmGmGmUmGmCmU*mU *mU*mU	230
G009845	AGCAACCUCACUC UUGUCUG GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	198	mA*mG*mC*AACCUCACUCUUGU CUGGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAAU AAGGCUAGUCCGUUAUCA mAmCmUmUm GmAmAmAmAmAmGmUmGmGmC mAmCmCmGm AmGmUmCmGmGmUmGmCmU*mU *mU*mU	231
G009846	ACCUCACUCUUGU CUGGGGA GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	199	mA*mC*mC*UCACUCUUGUCUGG GGAGUUUU AGAmGmCmUmAmGmAmAmUm mAmGmCAA GUUAAAAUAAGGCUAGUCCGUU AUCAmAmCm UmUmGmAmAmAmAmAmGmUmG mGmCmAmCm CmGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	232
G009847	CCUCACUCUUGUC UGGGGAA GUUUUAGAGCUAG AAAUAGC	200	mC*mC*mU*CACUCUUGUCUGGG GAAGUUUUA GAmGmCmUmAmGmAmAmUm AmGmCAAGU	233

	AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUIUUU		UAAAUAAGGCUAGUCCGUUAU CmAmCmUm UmGmAmAmAmAmGmUmGmG mCmAmCmCm GmAmGmUmCmGmGmUmGmCmU* mU*mU*mU	
G009848	CUCACUCUUGUCU GGGGAAG GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUIUUU	201	mC*mU*mC*ACUCUUGUCUGGGG AAGGUUUU AGAmGmCmUmAmGmAmAmAmU mAmGmCAA GUUAAAUAAGGCUAGUCCGUU AUCAmAmCm UmUmGmAmAmAmAmGmUmG mGmCmAmCm CmGmAmGmUmCmGmGmUmGmC mU*mU*mU*mU	234
G009849	GGGGAAGGGGAGA AAAAAAA GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUIUUU	202	mG*mG*mG*GAAGGGGAGAAAA AAAGUUUUAG AmGmCmUmAmGmAmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	235
G009850	GGGAAGGGGAGAA AAAAAAA GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG	203	mG*mG*mG*AAGGGGAGAAAA AAAGUUUUAG AmGmCmUmAmGmAmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*	236

	UGCUUUU		mU	
G009851	AUGCAUUUGUUUC AAAAUAU GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	204	mA*mU*mG*CAUUUGUUUCAAAA UAUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	237
G009852	UGCAUUUGUUUCA AAAUUU GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	205	mU*mG*mC*AUUUGUUUCAAAAU AUUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	238
G009853	UGAUUCCUACAGA AAAAGUC GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	206	mU*mG*mA*UUCCUACAGAAAAA GUCGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	239
G009854	UACAGAAAAGUC AGGAUAA GUUUUAGAGCUAG AAAUAGC	207	mU*mA*mC*AGAAAAGUCAGGA UAAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA	240

	AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU		AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	
G009855	UUUCUUCUGCCUU UAAACAG GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	208	mU*mU*mU*CUUCUGCCUUUAAA CAGGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	241
G009856	UUAUAGUUUAUA UUCAAC GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	209	mU*mU*mA*UAGUUUAUAUUCA AACGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	242
G009857	AUUUAUGAGAUCA ACAGCAC GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG	210	mA*mU*mU*UAUGAGAUCAACAG CACGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*	243

	UGCUUUU		mU	
G009858	GAUCAACAGCACA GGUUUUG GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	211	mG*mA*mU*CAACAGCACAGGUU UUGGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	244
G009859	UUAAAUAAGCAU AGUGCAA GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	212	mU*mU*mA*AAUAAAGCAUAGUG CAAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	245
G009860	UAAAGCAUAGUGC AAUGGAU GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	213	mU*mA*mA*AGCAUAGUGCAAUG GAUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	246
G009861	UAGUGCAAUGGAU AGGUCUU GUUUUAGAGCUAG AAAUAGC	214	mU*mA*mG*UGCAAUGGAUAGGU CUUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA	247

	AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU		AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	
G009862	AGUGCAAUGGAUA GGUCUUA GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	215	mA*mG*mU*GCAAUGGAUAGGUC UUAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	248
G009863	UUACUUUGCACUU UCCUUAG GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	216	mU*mU*mA*CUUUGCACUUCCU UAGGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	249
G009864	UACUUUGCACUUU CCUUAGU GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG	217	mU*mA*mC*UUUGCACUUCCUU AGUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU*	250

	UGCUUUU		mU	
G009865	UCUGACCUUUUAU UUUACCU GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	218	mU*mC*mU*GACCUUUUAUUUA CCUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	251
G009866	UACUAAAACUUUA UUUUACU GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	219	mU*mA*mC*UAAAACUUUAUUUU ACUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	252
G009867	AAAGUUGAACAAU AGAAAA GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	220	mA*mA*mA*GUUGAACAAUAGAA AAAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	253
G009868	AAUGCAUAAUCUA AGUCAAA GUUUUAGAGCUAG AAAUAGC	221	mA*mA*mU*GCAUAAUCUAAGUC AAAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA	254

	AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUIUUU		AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	
G009869	AUUAUCCUGACUU UUUCUGU GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUIUUU	222	mA*mU*mU*AUCCUGACUUUUUC UGUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	255
G009870	UGAAUUAUCCUC UGUUUAA GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUIUUU	223	mU*mG*mA*AUUAUCCUCUGUU UAAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	256
G009871	UAAUUUUCUUUUG CCCACUA GUUUUAGAGCUAG AAAUAGC AAGUUAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG	224	mU*mA*mA*UUUUCUUUUGCCCA CUAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUm GmAmAmAmAmAmGmUmGmGmC mAmCmCmGm AmGmUmCmGmGmUmGmCmU*mU	257

	UGCUUUU		*mU*mU	
G009872	AAAAGGUCAGAAU UGUUUAG GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	225	mA*mA*mA*AGGUCAGAAUUGUU UAGGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	258
G009873	AACAUCCUAGGUA AAAUAAA GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	226	mA*mA*mC*AUCCUAGGUAAAAU AAAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	259
G009874	UAAUAAAAUUCAA ACAUCCU GUUUUAGAGCUAG AAAUAGC AAGUUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	227	mU*mA*mA*UAAAAUCAAACA CCUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	260
G009875	UUGUCAUGUAUUU CUAAAAU GUUUUAGAGCUAG AAAUAGC	228	mU*mU*mG*UCAUGUAUUUCUAA AAUGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA	261

	AAGUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU		AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	
G009876	UUUGUCAUGUAUU UCUAAAA GUUUUAGAGCUAG AAAUAGC AAGUAAAAUAAG GCUAGUC CGUUAUCAACUUG AAAAAGU GGCACCGAGUCGG UGCUUUU	229	mU*mU*mU*GUCAUGUAUUUCUA AAAGUUUUAG AmGmCmUmAmGmAmAmUmA mGmCAAGUUA AAUAAGGCUAGUCCGUUAUCAm AmCmUmUmGm AmAmAmAmAmGmUmGmGmCmA mCmCmGmAmGm UmCmGmGmUmGmCmU*mU*mU* mU	262

Таблица 11: Компоненты и последовательности векторов

I D плазми ды	ИК П	А кцептор сплайсин га (1-ая ориентац ия)	Т рансген (1-ая ориентац ия)	П оли-А (1- ая ориентац ия)	П оли-А (2- ая ориентац ия)	Т рансген (2-ая ориентац ия)	А кцептор сплайсин га (2-ая ориентац ия)	ИК П
00147	SE Q ID NO : 263)	А кцептор сплайсин га альбуми на мыши (SEQ ID NO: 264)	Ф актор IX человека (R338L) (SEQ ID NO: 265)	S EQ ID NO: 266	S EQ ID NO: 267	Ф актор IX человека (R338L) (SEQ ID NO: 268)	А кцептор сплайсин га альбуми на мыши (SEQ ID NO: 269)	SE Q ID NO : 270)
00411	SE Q ID	А кцептор сплайсин га	Ф актор IX человека (R338L)-	S EQ ID NO: 266	S EQ ID NO: 267	Ф актор IX человека (R338L)-	А кцептор сплайсин га	SE Q ID

	NO : 263)	фактора IX человека (SEQ ID NO: 271)	HiBit (SEQ ID NO: 272)			HiBit (SEQ ID NO: 273)	фактора IX человека (SEQ ID NO: 274)	NO : 270)
00415	SE Q P ID NO : 263)	А кцептор сплайсин га альбуми на мыши (SEQ ID NO: 264)	NI uc-P2A- GFP (SEQ ID NO: 275)	S EQ ID NO: 266	S EQ ID NO: 267	NI uc-P2A- GFP (SEQ ID NO: 276)	А кцептор сплайсин га альбуми на мыши (SEQ ID NO: 269)	SE Q ID NO : 270)
00418	SE Q P ID NO : 263)	А кцептор сплайсин га альбуми на мыши (SEQ ID NO: 264)	Ф актор IX человека (R338L)- HiBit (SEQ ID NO: 272)	S EQ ID NO: 266	S EQ ID NO: 267	Ф актор IX человека (R338L)- HiBit (SEQ ID NO: 273)	А кцептор сплайсин га альбуми на мыши (SEQ ID NO: 269)	SE Q ID NO : 270)

Интрон 1 альбумина человека: (SEQ ID NO: 1)

GTAAGAAATCCATTTTTCTATTGTTCAACTTTTATTCTATTTTCCCAGTAAAAT
 AAAGTTTTAGTAAACTCTGCATCTTTAAAGAATTATTTTGGCATTATTTCTAAAATG
 GCATAGTATTTTGTATTTGTGAAGTCTTACAAGGTTATCTTATTAATAAAAATTCAAAC
 ATCCTAGGTAAAAAAGGTCAGAATTGTTTAGTGACTGTAATTTTCTTTTG
 CGCACTAAGGAAAGTGCAAAGTAACTTAGAGTGACTGAACTTCACAGAATAGGGT
 TGAAGATTGAATTCATAACTATCCCAAAGACCTATCCATTGCACTATGCTTTATTTA
 AAAACCACAAAACCTGTGCTGTTGATCTCATAAATAGAAGTTGTATTTATATTTATTT
 TCATTTTAGTCTGTCTTCTTGGTTGCTGTTGATAGACACTAAAAGAGTATTAGATATT
 ATCTAAGTTTGAATATAAGGCTATAAATATTTAATAATTTTAAAATAGTATTCTTGG
 TAATTGAATTATTCTTCTGTTTAAAGGCAGAAGAAATAATTGAACATCATCCTGAGT
 TTTTCTGTAGGAATCAGAGCCCAATATTTTGAACAAATGCATAATCTAAGTCAAAT
 GGAAAGAAATATAAAAAGTAACATTATTAATTCTTGTTTTCTTCAGTATTTAACAAT
 CCTTTTTTTTCTTCCCTTGCCCAG

5' ИКП последовательность (SEQ ID NO: 263):

TTGGCCACTCCCTCTCTGCGGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA

GGTCGCCCCGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
GAGAGGGGAGTGGCCAACCTCCATCACTAGGGGGTTCCCT

Акцептор сплайсинга альбумина мыши (1-ая ориентация) (SEQ ID NO: 264):

TAGGTCAAGTGAAGAGAAGAACAACAAAAGCAGCATATTACAGTTAGTTGTCTTC
ATCAATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAG

Фактор IX человека (R338L), 1-ая ориентация (SEQ ID NO: 265):

TTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATA
ATTCAGGTAATTTGGAAGAGTTTGTTCAGGGAAACCTTGAGAGAGAATGTATGGAA
GAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAACAACTGAAAGAACAAC
TGAATTTTGGAAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTAAA
TGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATT
TGAAGGAAAGAAGTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCG
AGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTTGCTCCTGTACTGAGGGAT
ATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGA
AGAGTTTCTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATG
TGGACTATGTAATTTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCA
CCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTC
AATTCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTA
TCGTTAATGAAAAATGGATTGTAAGTCTGCCCCTGTGTTGAAACTGGTGTAAAA
TTACAGTTGTTCGCAGGTGAACATAATTTGAGGAGACAGAACATACAGAGCAAAAAG
CGAAATGTGATTCGAATTTATTCCTCACCACAACCTACAATGCAGCTATTAATAAGTAC
AACCATGACATTTGCCCTTCTGGAAGTGGACGAACCCTTAGTGCTAAACAGCTACGTT
ACACCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCT
GGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAGGGAGATCAGCTTTAGTTCTT
CAGTACCTTAGAGTTCCACTTGTGACCGAGCCACATGTCTTCTATCTACAAAGTTC
ACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGT
CAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAAGGGACCAGTTTCTTAAC
TGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATA
CCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAGCTCACTTAA

Поли-А (1-ая ориентация) (SEQ ID NO: 266):

CCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGC
CTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCCTAATAAAAATGAGGAAA
TTGCATCGCATTGTCTGAGTAGGTGTCAATTCTATTCTGGGGGGTGGGGTGGGGCAGG
ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGG
CTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCC

Поли-А (2-ая ориентация) (SEQ ID NO: 267):

AAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAAATGAATGCAATT
GTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATC
ACAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTTGTCCAAAC
TCATCAATGTATCTTATCATGTCTG

Фактор IX человека (R338L), 2-ая ориентация (SEQ ID NO: 268):

TTAGGTGAGCTTAGTCTTTTCTTTTATCCAATTCACGTAGCGAGAGACCTTCG
 TATAGATGCCATATTTCCCTTCATCGCACATTCCTCCCCCAACTTATTATCCCGGT
 CAAGAAACTTGTTCCCTTCGACTTCAGTGACGTGTGGTCCACCTGAATCACCTGGCA
 TGAGTCGCGACCGCCCTCGTGAAACCCAGCACAAAACATGTTATTGTAAATCGTAA
 ATTTCTGTGGACAGAAGACAGGTCGCTCTATCGACCAACGGGACGCGCAAATATTGC
 AGAACGAGGGCTGATCGACCTTTGTGGAAGACCCGCCCCACCCACTCACATATCC
 GCTCCCAAATTTCAAGAAGATATTTGTATATTCTTTATCGGCTATACAAATCGGGGT
 AACATAGGAGTTAAGTACGAGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGT
 TGTACTIONTTTATAGCGGCATTATAATTGTGATGGGGTATGATCCTGATAACATTCCT
 TTTCTGTTTCAAGTATGCTCAGTTTCTTCAATGTTGTGTTTCGCCAGCCACGACCGTAATC
 TTAACCCCGTCTCGACACAGTGTGCGGCCGTTACAATCCACTTTTCATTGACTATG
 GAGCCCCACAAAACGCGTCGACTTTTCCGTTGAGCACCACCTGCCATGGAAATTGG
 CCAGGTTTAGCGTCCTCGCCCCGACAACCCTAGTAAAGTCATTAATGACTGTGTG
 GATTGTGTTATATTATCAAGAATCGTTTCGGCTTCAGTAGAGTTAACGTAGTCCACA
 TCGGGAAAACTGTCTCGGCCCTTGTCAACTTTGATGTCTGGGACACACTTACCCGA
 CCGCACGGGAAGGGCACCGCCGGTTCACAGCTCTTTTGATTCTCAGCGAGCCGGTAG
 CCCTCAGTGCAACTACACACAACCTTTGTTGTCGGCGGAATTTTTACAGAATTGCTCG
 CATCGTCCATTTTTAATGTTGCAGGTGACGTCCAACCTCGCAGTTTTTTCTTCAAAC
 CAAAAGGGCACCAACACTCGTAGGAATTTATATCGTCTTTACAACCTCCCCCATTC
 GACATGGATTAGATTCGCATTGGTCCCCATCGACATATTGCTTCCAGAACTCAGTGG
 TCCGTTCTGTATTCTCAAACACCTCGCGCGCTTCTTCAAACACTGCATTTTTCTCCAT
 AACTCTCGCTCCAAGTTCCTTGCACGAATTCTTCAAGCTTTCCTGAGTTATACCTT
 TTAGGCCGGTTAAGTATCTTATTCGCGTTTTTCGTGGTCCAGAAA

Акцептор сплайсинга альбумина мыши (2-ая ориентация) (SEQ ID NO: 269):

CTGTGGAAACAGGGAGAGAAAAACCACACAACATATTTAAAGATTGATGAA
 GACAACCTAAGTAAATATGCTGCTTTTTGTTCTTCTCTTCACTGACCTA

3' ИКП последовательность (2-ая ориентация) (SEQ ID NO: 270):

AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTC
 ACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTC
 AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Акцептор сплайсинга фактора IX человека (1-ая ориентация) (SEQ ID NO: 271):

GATTATTTGGATTAAAAACAAAGACTTTCTTAAGAGATGTAATAATTTTCATGA
 TGTTTTCTTTTTTGCTAAAACTAAAGAATTATTCTTTTACATTTTCAG

Фактор IX человека (R338L)-HiBit, 1-ая ориентация (SEQ ID NO: 272):

TTTCTTGATCATGAAAACGCCAACAATAATTCTGAATCGGCCAAAGAGGTATA
 ATTCAGGTAAATTGGAAGAGTTTGTTCAGGGAACCTTGAGAGAGAATGTATGGAA
 GAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAACAACTGAAAGAACAAC
 TGAATTTTGGAAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAA
 TGCGCGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATT

TGAAGGAAAGAAGCTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCG
 AGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTTGCTCCTGTACTGAGGGAT
 ATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGA
 AGAGTTTCTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATG
 TGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCA
 CCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTC
 AATCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTA
 TCGTTAATGAAAAATGGATTGTAAGCTGCTGCCACTGTGTTGAAACTGGTGTAAAA
 TTACAGTTGTGCGCAGGTGAACATAATATTGAGGAGACAGAACATACAGAGCAAAAG
 CGAAATGTGATTCTGAATTATTCCTCACCACAACCTACAATGCAGCTATTAATAAGTAC
 AACCATGACATTGCCCTTCTGGAAGTGGACGAACCCTTAGTGCTAAACAGCTACGTT
 ACACCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCT
 GGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTT
 CAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTC
 ACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGT
 CAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGACCAGTTTCTTAAC
 TGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATA
 CCAAGGTCTCCCGGTATGTCAACTGGATTAAGGAAAAACAAGCTCACTGTCAGC
 GGATGGAGACTGTTCAAGAAGATCAGCTAA

Фактор IX человека (R338L)-HiBit (2-ая ориентация) (SEQ ID NO: 273):

TTAGGAAATCTTCTTAAACAGCCGCCAGCCGCTCACGGTGAGCTTAGTCTTTT
 CTTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATATTTCCCCTTCAT
 CGCACATTCCTCCCCCAACTTATTATCCCGGTCAAGAACTTGTTCCTTCGACTTCA
 GTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCGCCCTCGTGAAAC
 CCAGCACAAAACATGTTATTGTAAATCGTAAATTTTCGTGGACAGAAGACAGGTGCG
 TCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGATCGACCTTTGT
 GGAAGACCCGCCCCACCCACTCACATATCCGCTCCCAAATTTCAAGAAGATATTTG
 TATATTCTTTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAGTACGAGTGGCT
 CGTCCAGCTCCAGGAGGGCTATATCATGGTTGTAAGTGTACTTGTATAGCGGCATTATAAT
 TGTGATGGGGTATGATCCTGATAACATTCCTTTTCTGTTTCAAGTATGCTCAGTTTCTTC
 AATGTTGTGTTTCGCCAGCCACGACCGTAATCTTAACCCCGTCTCGACACAGTGTGC
 GGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCACAAAACGCGTCGACTTT
 TCCGTTGAGCACCACCTGCCATGGAAATTGGCCAGGTTTAGCGTCCCTCGCCCCGAC
 AACCTAGTAAAGTCATTAATGACTGTGTGGATTGTGTTATATTATCAAGAATCGT
 TTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAAACTGTCTCGGCCCTTGT
 CAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGGCACCGCCGGTT
 CACAGCTCTTTTGAATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTACACACAACCT
 TGTGTCGGCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTTAATGTTGCAGGT
 GACGTCCAACCTCGCAGTTTTTTTCCCTTCAAACCAAAAAGGGCACCAACACTCGTAGGA
 ATTTATATCGTCTTTACAACCTCCCCCATTCAGACATGGATTAGATTTCGATTGGTCC

CCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAAACACCTCGC
GCGCTTCTTCAAAACTGCATTTTTCTCCATACACTCTCGCTCCAAGTTCCTTGCAC
GAATTCTTCAAGCTTTCCTGAGTTATACSTTTTAGGCCGGTTAAGTATCTTATTCGCG
TTTTCGTGGTCCAGAAA

Акцептор сплайсинга фактора IX человека (2-ая ориентация) (SEQ ID NO: 274):

CTGAAATGTAAGAATAATTCTTTAGTTTTAGCAAAAAAGAAAACATCATG
AAAATTTTACATCTCTTAAGAAAGTCTTTGTTTTTAATCCAAATAATC

Nluc-P2A-GFP (1-ая ориентация) (SEQ ID NO: 275):

TTTCTTGATCATGAAAACGCCAACAATTTCTGAATCGGCCAAAGAGGTATA
ATTCAGGTAAATTGGAAGAGTTTGTTC AAGGGAACCTTGAGAGAGAATGTATGGAA
GAAAAGTGTAGTTTTGAAGAAGCAGTATTCACCTTGGAGGACTTTGTCTGGTACTGG
AGGCAAACCGCTGGTTATAATCTCGACCAAGTACTGGAACAGGGCGGGGTAAGTTC
CCTCTTTCAGAATTTGGGTGTAAGCGTCACACCAATCCAGCGGATTGTGTTGTCTGG
AGAGAACGGA CTCAA AATTGACATCCATGTTATCATTCCATATGAAGGTCTCAGTGG
AGACCAAATGGGGCAGATCGAGAAGATTTTCAAGGTAGTTTACCCAGTCGACGATC
ACCACTTCAAAGTCATTCTCCACTATGGCACACTTGTTATCGACGGAGTAACTCCTA
ATATGATTGATTACTTTGGTCTGCCCCTATGAGGGCATCGCAGTGTTCATGGCAAAA
AGATCACCGTAACAGGAACGTTGTGGAATGGGAACAAGATAATCGACGAGAGATTG
ATAAATCCAGACGGGTCCTCTGTTTCAGGGTTACAATTAACGGCGTCACAGGATG
GAGACTCTGTGAACGAATACTGGCCACAATTTTTCACTCCTGAAGCAGGCCGGAG
ACGTGGAGGAAAACCCAGGGCCCGTGAGCAAGGGCGAGGAGCTGTTACCCGGGGT
GGTGCCCATCCTGGTCTGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGC
ACCACCGGCAAGCTGCCCCTGCCCCTGCCCCACCCTCGTGACCACCCTGACCTACGGC
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCC
GCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAA
CTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCG
AGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAG
TACA ACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCAT
CAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCG
ACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAAC
CACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCA
CATGGTCTGCTGGAGTTCGTGACCGCCCGGGATCACTCTCGGCATGGACGAGCT
GTACAAGGGAGGAGGAAGCCCGAAGAAGAAGAGAAAGGTCTAA

Nluc-P2A-GFP (2-ая ориентация) (SEQ ID NO: 276):

TTACACCTTCTCTTCTTGGGGCTGCCGCCGCCCTTGTACAGCTCGTCCAT
GCCCAGGGTGATGCCGGCGGGTACGAACTCCAGCAGCACCATGTGGTCCCTCTT
CTCGTTGGGGTCTTGTCTCAGGGCGCTCTGGGTGCTCAGGTAGTGGTTGTCTGGGCAG
CAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCTGGCCAGCTGCA
CGCTGCCGTCTCGATGTTGTGCCTGATCTTGAAGTTCACCTTGTATGCCGTTCTTCTG

CTTGTCGGCCATGATGTACACGTTGTGGCTGTTGTAGTTGTACTCCAGCTTGTGGCCC
 AGGATGTTGCCGTCCTCCTTGAAGTCGATGCCCTTCAGCTCGATCCTGTTCAACCAGG
 GTGTGCCCTCGAACTTCACCTCGGCCCTGGTCTTGTAGTTGCCGTCGTCCTTGAAGA
 AGATGGTCCTCTCCTGCACGTAGCCCTCGGGCATGGCGCTCTTGAAGAAGTCGTGCT
 GCTTCATGTGGTCGGGGTACCTGCTGAAGCACTGCACGCCGTAGGTCAGGGTGGTCA
 CCAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTC
 AGCTTGCCGTAGGTGGCGTCGCCCTCGCCCTCGCCGCTCACGCTGAACTTGTGGCCG
 TTCACGTCGCCGTCCAGCTCCACCAGGATGGGCACCACGCCGGTGAACAGCTCCTCG
 CCCTTGCTCACGGGGCCGGGGTTCTCCTCCACGTCGCCGGCCTGCTTCAGCAGGCTG
 AAGTTGGTGGCCAGGATCCTCTCGCACAGCCTCCAGCCGGTCACGCCGTTGATGGTC
 ACCCTGAACAGCAGGCTGCCGTCGGGGTTGATCAGCCTCTCGTCGATGATCTTGTG
 CCGTTCCACAGGGTGGCCGGTACGGTGATCTTCTTGCCGTCGAACACGGCGATGCC
 TCGTAGGGCCTGCCGAAGTAGTCGATCATGTTGGGGGTACGCCGTCGATCACCAG
 GGTGCCGTAGTGCAGGATCACCTTGAAGTGGTGGTCGTCCACGGGGTACACCACCTT
 GAAAATCTTCTCGATCTGGCCATCTGGTCGCCGCTCAGGCCCTCGTAGGGGATGAT
 CACGTGGATGTCGATCTTCAGGCCGTTCTCGCCGCTCAGCACGATCCTCTGGATGGG
 GGTACGCTCACGCCAGGTTCTGGAACAGGCTGCTCACGCCGCCCTGCTCCAGCAC
 CTGGTCCAGGTTGTAGCCGGCGGTCTGCCTCCAGTCGCCACGAAGTCCTCCAGGGT
 GAACACGGCCTCCTCGAAGCTGCACTTCTCCTCCATGCACTCCCTCTCCAGGTTGCC
 CTGCACGAACTCCTCCAGCTTGCCGCTGTTGTACCTCTTGGGCCTGTTCAAGGATCTTG
 TTGGCGTTCTCGTGGTCCAGGAA

Полная последовательность P00147 (от ИКП до ИКП): (SEQ ID NO: 277)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
 GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
 GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCTAGATCTCTTAGGTCAGTGAA
 GAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATG
 TTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACA
 AAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAA
 GGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAG
 AAGTTTTTGAAAACACTGAAAGAACAACCTGAATTTTGAAGCAGTATGTTGATGGA
 GATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAAT
 TCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTA
 ACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTTGTA AAAAATAGTGCTGATAA
 CAAGGTGGTTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTG
 TGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAACTTCTAAGCT
 CACCCGTGCTGAGACTGTTTTTCTGATGTGGACTATGTA AATTCTACTGAAGCTGA
 AACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGT
 TGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTTTTGAATGG
 TAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAATGGATTGTA ACTGC
 TGCCCACTGTGTTGAAACTGGTGTAAAATTACAGTTGTTCGCAGGTGAACATAATAT

TGAGGAGACAGAACATACAGAGCAAAAGCGAAATGTGATTCTGAATTATTCCTCACC
ACAACACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAACTGG
ACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAAT
ACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAAGAGTCT
TCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTCCACTTGTGACC
GAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTG
GCTTCCATGAAGGAGGTAGAGATTGATGTCAAGGAGATAGTGGGGGACCCCATGTT
ACTGAAGTGGAAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTG
TGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGA
TTAAGGAAAAACAAAGCTCACTTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATC
TGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCACTGTG
CTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATT
CTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCA
GGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGG
GGCTCTAGGGGGTATCCCCAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATA
AAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATA
AAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTTCACTGCATTCTAGTTG
TGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTTAGGTGAGCTTAGTCTTT
TCTTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATATTTCCCTTCA
TCGCACATTCTCCCCCAACTTATTATCCCGGTCAAGAACTTGTTCCCTTCGACTTC
AGTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCGCCCTCGTGAAA
CCCAGCACAAAACATGTTATTGTAAATCGTAAATTTTCGTGGACAGAAGACAGGTGCG
CTCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGGCTGATCGACCTTTG
TGGAAGACCCGCCCCCACCCTCACATATCCGCTCCCAAATTTCAAGAAGATATTT
GTATATTCTTTATCGGCTATACAAATCGGGGTAAACATAGGAGTTAAGTACGAGTGGC
TCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTACTTGTTTATAGCGGCATTATAA
TTGTGATGGGGTATGATCCTGATAACATTCCTTTTCTGTTTCAGTATGCTCAGTTTCTT
CAATGTTGTGTTGCGCCAGCCACGACCGTAATCTTAACCCCGTCTCGACACAGTGTG
CGGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCACAAAACGCGTCGACTT
TTCCGTTGAGCACCACTGCCATGGAAATTGGCCAGGTTTAGCGTCCTCGCCCCGA
CAACCCTAGTAAAGTCATTAAATGACTGTGTGGATTGTGTTATATTATCAAGAATCG
TTTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAAACTGTCTCGGCCCTTG
TCAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGGCACCGCCGGTT
CACAGCTCTTTTGATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTACACACAACCTT
TGTTGTCGGCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTTTAATGTTGCAGGT
GACGTCCAACCTCGCAGTTTTTTTCTTCAAACCAAAGGGCACCAACACTCGTAGGA
ATTTATATCGTCTTTACAACCTCCCCCATTCAGACATGGATTAGATTGCGATTGGTCC
CCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAAACACCTCGC
GCGCTTCTTCAAACCTGCATTTTTCTCCATACACTCTCGCTCCAAGTTCCTTGCAC
GAATTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATCTTATTCGCG

TTTTCGTGGTCCAGAAAACTGTGGAAACAGGGAGAGAAAAACCACACAACATATT
TAAAGATTGATGAAGACAACCTAACTGTAATATGCTGCTTTTTGTTCTTCTCTTCACTG
ACCTAAGAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCT
CGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGGCGACSTTTGGTCCG
CCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGGAGTGGCCAA

Полная последовательность P00411 (от ИКП до ИКП): (SEQ ID NO: 278)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
GGTCGCCCCGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCCTAGATCTCTGATTATTTGGATT
AAAAACAAAGACTTTCTTAAGAGATGTAAAATTTTCATGATGTTTTCTTTTTTGCTAA
AACTAAAGAATTATTCTTTTACATTTTCAGTTTTTTCTTGATCATGAAAACGCCAACAA
AATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTC AAG
GGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGA
AGTTTTTGAAAACACTGAAAGAACAACCTGAATTTTGGAAGCAGTATGTTGATGGAG
ATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATT
CCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTAA
CATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAAC
AAGGTGGTTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGT
GAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTTCTGTTTTCAAACTTCTAAGCTC
ACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTATGTAAATTCTACTGAAGCTGAA
ACCATTTTGATAACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTT
GTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGT
AAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAACCTGCT
GCCCACTGTGTTGAAACTGGTGTAAAATTACAGTTGTCGCAGGTGAACATAATATT
GAGGAGACAGAACATACAGAGCAAAAGCGAAATGTGATTTCGAATTATTCCTCACCA
CAACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAACCTGGA
CGAACCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATA
CACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTT
CCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCG
AGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGG
CTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTTA
CTGAAGTGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGT
GCAATGAAAGGCAAATATGGAATATATACCAAGGTCTCCCGGTATGTCAACTGGAT
TAAGGAAAAACAAAGCTCACTGTCAGCGGATGGAGACTGTTCAAGAAGATCAGCT
AACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCT
TCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAAATGAGGAAATT
GCATCGCATTGTCTGAGTAGGTGTCACTTCTATTCTGGGGGGTGGGGTGGGGCAGGAC
AGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCT
CTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCAAAAA
ACCTCCCACACCTCCCCCTGAACCTGAAACATAAAAATGAATGCAATTGTTGTTGTTA

ACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCA
 CAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGT
 ATCTTATCATGTCTGTTAGGAAATCTTCTTAAACAGCCGCCAGCCGCTCACGGTGAG
 CTTAGTCTTTTCTTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATAT
 TTCCCCTTCATCGCACATTCCCTCCCCCAACTTATTATCCCGGTCAAGAACTTGTTCC
 CTTCGACTTCAGTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCCGC
 CCTCGTGAAACCCAGCACAAAACATGTTATTGTAAATCGTAAATTTTCGTGGACAGAA
 GACAGGTCGCTCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGAT
 CGACCTTTGTGGAAGACCCGCCCCCACTCACATATCCGCTCCCAAATTTCAAG
 AAGATATTTGTATATTCTTTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAGT
 ACGAGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTACTTGTTTATAGCG
 GCATTATAATTGTGATGGGGTATGATCCTGATAACATTCCTTTTCTGTTCAAGTATGCT
 CAGTTTCTTCAATGTTGTGTTTCGCCAGCCACGACCGTAATCTTAACCCCGTCTCGAC
 ACAGTGTGCGGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCACAAAACGC
 GTCGACTTTTCCGTTGAGCACCACCTGCCATGGAAATTGGCCAGGTTTAGCGTCCCTC
 GCCCCGACAACCCTAGTAAAGTCATTAATGACTGTGTGGATTGTGTTATATTATC
 AAGAATCGTTTTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAACTGTCTC
 GGCCCTTGTCAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGGCAC
 CGCCGGTTCACAGCTCTTTTGATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTACA
 CAACTTTGTTGTCGGCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTAAATG
 TTGCAGGTGACGTCCAACCTCGCAGTTTTTTCTTCAAACCAAAGGGCACCAACAC
 TCGTAGGAATTTATATCGTCTTTACAACCTCCCCCATTCAGACATGGATTAGATTCGC
 ATTGGTCCCCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAA
 CACTCGCGCGCTTCTTCAAACCTGCATTTTTCTCCATACACTCTCGCTCCAAGTTC
 CCTTGCACGAATTTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATCT
 TATTCGCGTTTTTCGTGGTCCAGAAAACTGAAATGTAAAAGAATAATTCTTTAGTTT
 TAGCAAAAAAGAAAACATCATGAAAATTTTACATCTCTTAAGAAAGTCTTTGTTTTT
 AATCCAAATAATCAGAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTC
 TGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCCGGGCGACC
 TTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Полная последовательность P00415 (от ИКП до ИКП): (SEQ ID NO: 279)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
 GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
 GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCCTAGATCTCTTAGGTCAGTGAA
 GAGAAGAACAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATG
 TTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACA
 AAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAA
 GGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCAGTATT
 CACTTTGGAGGACTTTGTGGTACTGGAGGCAAACCGCTGGTTATAATCTCGACCA
 AGTACTGGAACAGGGCGGGGTAAGTTCCCTCTTTCAGAATTTGGGTGTAAGCGTCAC

ACCAATCCAGCGGATTGTGTTGTCTGGAGAGAACGGACTCAA AATTGACATCCATGT
TATCATTCCATATGAAGGTCTCAGTGGAGACCAAATGGGGCAGATCGAGAAGATTT
TCAAGGTAGTTTACCCAGTCGACGATCACC ACTTCAAAGTCATTCTCCACTATGGCA
CACTTGTATCGACGGAGTAACTCCTAATATGATTGATTACTTTGGTCGCCCCGTATG
AGGGCATCGCAGTGT TTGATGGCAAAAAGATCACCGTAACAGGAACGTTGTGGAAT
GGGAACAAGATAATCGACGAGAGATTGATAAATCCAGACGGGTCACTCCTGTT CAG
GGTTACAATTAACGGCGTCACAGGATGGAGACTCTGTGAACGAATACTGGCCACAA
ATTTTTCACTCCTGAAGCAGGCCGGAGACGTGGAGGAAAACCCAGGGCCCCGTGAGC
AAGGGCGAGGAGCTGTTACCCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGA
CGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACG
GCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTGGCCCA
CCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACA
TGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGC
ACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGA
GGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACG
GCAACATCCTGGGGCACAAGCTGGAGTACA ACTACAACAGCCACAACGTCTATATC
ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACAT
CGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCG
ACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCA
AAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCC
GGGATCACTCTCGGCATGGACGAGCTGTACAAGGGAGGAGGAAGCCCGAAGAAGA
AGAGAAAGGTCTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCC
CTCCCCCGTGCCCTTCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCTAATAA
AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGG
GTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGG
ATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGG
TATCCCCAAAAACCTCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCA
ATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC
ATCACAAATTTACAAATAAAGCATT TTTTTTCACTGCATTCTAGTTGTGGTTTGTCCA
AACTCATCAATGTATCTTATCATGTCTGTTACACCTTCTCTTCTTGGGGCTGCC
GCCGCCCTTGTACAGCTCGTCCATGCCAGGGTGATGCCGGCGGGCGGTCACGAACTC
CAGCAGCACCATGTGGTCCCTCTTCTCGTTGGGGTCCTTGCTCAGGGCGCTCTGGGT
GCTCAGGTAGTGGTTGTGCGGCAGCAGCACGGGGCCGTGCGCGATGGGGGTGTTCT
GCTGGTAGTGGTCGGCCAGCTGCACGCTGCCGTCTCGATGTTGTGCCTGATCTTGA
AGTTCACCTTGATGCCGTTCTTCTGCTTGTGCGGCCATGATGTACACGTTGTGGCTGTT
GTAGTTGTA CTCCAGCTTGTGGCCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCC
CTTCAGCTCGATCCTGTTCAACAGGGTGTGCGCCCTCGAACTTCACCTCGGCCCTGGTC
TTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTCCTCTCCTGCACGTAGCCCTCGGGC
ATGGCGCTCTTGAAGAAGTCGTGCTGCTTCATGTGGTCGGGGTACCTGCTGAAGCAC
TGCACGCCGTAGGT CAGGGTGGTCACCAGGGTGGGCCAGGGCACGGGCAGCTTGCC

GGTGGTGCAGATGAACTTCAGGGTCAGCTTGCCGTAGGTGGCGTCGCCCTCGCCCTC
GCCGCTCACGCTGAACTTGTGGCCGTTACGTCGCCGTCCAGCTCCACCAGGATGGG
CACCACGCCGGTGAACAGCTCCTCGCCCTTGCTCACGGGGCCGGGGTTCTCCTCCAC
GTCGCCGGCCTGCTTCAGCAGGCTGAAGTTGGTGGCCAGGATCCTCTCGCACAGCCT
CCAGCCGGTCACGCCGTTGATGGTCACCCTGAACAGCAGGCTGCCGTCCGGGGTTGAT
CAGCCTCTCGTCGATGATCTTGTTGCCGTTCCACAGGGTGCCGGTCACGGTGATCTT
CTTGCCGTCGAACACGGCGATGCCCTCGTAGGGCCTGCCGAAGTAGTCGATCATGTT
GGGGGTCACGCCGTCGATCACCAGGGTGCCGTAGTGCAGGATCACCTTGAAGTGGT
GGTCGTCCACGGGGTACACCACCTTGA AAAATCTTCTCGATCTGGCCCATCTGGTCGC
CGCTCAGGCCCTCGTAGGGGATGATCACGTGGATGTCGATCTTCAGGCCGTTCTCGC
CGCTCAGCACGATCCTCTGGATGGGGGTCACGCTCACGCCAGGTTCTGGAACAGG
CTGCTCACGCCGCCCTGCTCCAGCACCTGGTCCAGGTTGTAGCCGGCGGTCTGCCTC
CAGTCGCCACGAAGTCTCCAGGGTGAACACGGCCTCCTCGAAGCTGCACTTCTCC
TCCATGCACTCCCTCTCCAGGTTGCCCTGCACGA ACTCCTCCAGCTTGCCGCTGTTGT
ACCTCTTGGGCCTGTTTCAGGATCTTGTTGGCGTTCTCGTGGTCCAGGAA

Полная последовательность P00418 (от ИКП до ИКП): (SEQ ID NO: 280)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
GAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTCTAGATCTCTTAGGTCAGTGAA
GAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATG
TTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACA
AAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAA
GGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAGAAGCACGAG
AAGTTTTTGAAAACACTGAAAGAACA ACTGAATTTTGAAGCAGTATGTTGATGGA
GATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAAT
TCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTA
ACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAA
CAAGGTGGTTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTG
TGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAACTTCTAAGCT
CACCCGTGCTGAGACTGTTTTTCTGATGTGGACTATGTAAATTCTACTGAAGCTGA
AACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGT
TGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTTTTGAATGG
TAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAATGGATTGTA ACTGC
TGCCCACTGTGTTGAACTGGTGTAAAATTACAGTTGTTCGAGGTGAACATAATAT
TGAGGAGACAGAACATACAGAGCAAAAAGCGAAATGTGATTCGAATTATTCCTCACC
ACA ACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAACTGG
ACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAAT
ACACGAACATCTTCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAAGAGTCT
TCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACC
GAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTG

GCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTT
ACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTG
TGCAATGAAAGGCAAATATGGAATATATACCAAGGTCTCCCGGTATGTCAACTGGA
TTAAGGAAAAACAAAGCTCACTGTCAGCGGATGGAGACTGTTCAAGAAGATCAGC
TAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGTGCC
TTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAAATGAGGAAAT
TGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGA
CAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGC
TCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCAAA
AACCTCCCACACCTCCCCCTGAACCTGAAACATAAAAATGAATGCAATTGTTGTTGTT
AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTC
ACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATG
TATCTTATCATGTCTGTTAGGAAATCTTCTTAAACAGCCGCCAGCCGCTCACGGTGA
GCTTAGTCTTTTCTTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATA
TTTCCCCTTCATCGCACATTCCTCCCCCAACTTATTATCCCGGTCAAGAACTTGTT
CCTTCGACTTCAGTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCG
CCCTCGTGAAACCCAGCACAAAACATGTTATTGTAAATCGTAAATTCGTGGACAGA
AGACAGGTCGCTCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGA
TCGACCTTGTGGAAGACCCGCCCCCACTCACATATCCGCTCCCAAATTTCAA
GAAGATATTTGTATATTCTTTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAG
TACGAGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTACTTGTTTATAGC
GGCATTATAATTGTGATGGGGTATGATCCTGATAACATTCTTTTCTGTTTCAGTATGC
TCAGTTTCTTCAATGTTGTGTTTCGCCAGCCACGACCGTAATCTTAAACCCCGTCTCGA
CACAGTGTGCGGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCACAAAACG
CGTCGACTTTTCCGTTGAGCACCACTGCCATGGAAATTGGCCAGGTTTAGCGTCCT
CGCCCCGACAACCCTAGTAAAGTCATTAATGACTGTGTGGATTGTGTTATATTAT
CAAGAATCGTTTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAAACTGTCT
CGGCCCTTGTCAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGGCA
CCGCCGGTTCACAGCTCTTTTGATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTAC
ACACAACCTTTGTTGTCGGCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTAAAT
GTTGCAGGTGACGTCCAACCTCGCAGTTTTTTCCTTCAAACCAAAGGGCACCAACA
CTCGTAGGAATTTATATCGTCTTTACAACCTCCCCCATTTCAGACATGGATTAGATTCTG
CATTGGTCCCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAA
ACACCTCGCGCGCTTCTTCAAACCTGCATTTTTCTCCATACACTCTCGCTCCAAGTT
CCCTTGACGAATTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATC
TTATTCGCGTTTTCTGTGGTCCAGAAAACTGTGGAAACAGGGAGAGAAAAACCACA
CAACATATTTAAAGATTGATGAAGACAACCTAAGTAAATATGCTGCTTTTTGTTCTTC
TCTTCACTGACCTAAGAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCT
CTGCGCGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAGCCCGGGCGTGGGGCGAC
CTTTGGTCCGCCCGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Полная последовательность P00123 (от ИКП до ИКП): (SEQ ID NO: 281)

GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAG
GTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAG
AGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGGAGGGGTGGAGTCGTGATAG
GTCAGTGAAGAGAAGAACA AAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATC
TTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAA
ACGCCAACA AAATTTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAG
TTTGTTC AAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGA
AGCACGAGAAGTTTTTGA AAACACTGAAAGAACA AACTGAATTTTTGGAAGCAGTATG
TTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATG
ACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGA AACTGTGAAT
TAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGT
GCTGATAACAAGGTGGTTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCA
GAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAAC
TTCTAAGCTCACCCGTGCTGAGACTGTTTTTTCCTGATGTGGACTATGTAAATTCTACT
GAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACTTC
ACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTT
TTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATT
GTA ACTGCTGCCACTGTGTTGAAACTGGTGTAAAATTACAGTTGTCGCAGGTGAA
CATAATATTGAGGAGACAGAACATAACAGAGCAAAAAGCGAAATGTGATTTCGAATTAT
TCCTCACCACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCT
GGA ACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGA
CAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGG
AAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACT
TGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTT
CTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGAC
CCCATGTTACTGAAGTGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTG
AAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTATGTC
AACTGGATTAAGGAAAAAACA AAGCTCACTTAACTCGACTGTGCCTTCTAGTTGCC
AGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCC
CACTGTCCTTTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCA
TTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC
AATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAAC
CAGCTGGGGCTCTAGGGGGTATCCCCACTAGTCCACTCCCTCTCTGCGCGCTCGCTC
GCTCACTGAGGCCGGGCGACCAAAGGTGCGCCCGACGCCCGGGCTTTGCCCGGGCGG
CCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA

Полная последовательность P00204 (от ИКП до ИКП): (SEQ ID NO: 282)

GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAG
GTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAG
AGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTGGAGGGGTGGAGTCGTGACCT

AGGTCGTCTCCGGCTCTGCTTTTTCCAGGGGTGTGTTTCGCCGAGAAGCACGTAAGA
GTTTTATGTTTTTTCATCTCTGCTTGTATTTTTCTAGTAATGGAAGCCTGGTATTTTAA
AATAGTTAAATTTTCCTTTAGTGCTGATTTCTAGATTATTACTGTTGTTGTTGTTA
TTATTGTCATTATTTGCATCTGAGAACTAGGTCAGTGAAGAGAAGAACAAAAAGCA
GCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGGTTTTTCTCTCC
CTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAA
AGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAGGGAACCTTGAGAGAGAA
TGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGA
AAGAACAACCTGAATTTTGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATC
CATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTC
CCTTTGGATTTGAAGGAAAGAAGTGTGAATTAGATGTAACATGTAACATTAAGAAT
GGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGT
ACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATT
TCCATGTGGAAGAGTTTCTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGT
TTTTCTGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATAACAT
CACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGC
CAAACCAGGTCAATTCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTG
TGGAGGCTCTATCGTTAATGAAAAATGGATTGTAAGTCTGCCCCTGTGTTGAAAC
TGGTGTTAAAATTACAGTTGTCTGCAGGTGAACATAATATTGAGGAGACAGAACATA
CAGAGCAAAAGCGAAATGTGATTCTGAATTATTCCTCACCACAACCTACAATGCAGCT
ATTAATAAGTACAACCATGACATTGCCCTTCTGGAAGTGGACGAACCCTTAGTGCTA
AACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTC
AAATTTGGATCTGGCTATGTAAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATC
AGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTA
TCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGT
AGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAAGGGAC
CAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAAT
ATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAG
CTCACTTAACTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCC
CGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAATGA
GGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGG
GCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCG
GTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCC
CCCTTAGGTGGTTATATTATTGATATATTTTTGGTATCTTTGATGACAATAATGGGGG
ATTTTGAAAGCTTAGCTTTAAATTTCTTTTAAATTAATAAAAAAATGCTAGGCAGAATG
ACTCAAATTACGTTGGATACAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCT
CGACCATGCTATACTAAAAATTAAGTGTACTAGTCCACTCCCTCTCTGCGCGCTC
GCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCCGG
GCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA

Полная последовательность P00353 (от ИКП до ИКП): (SEQ ID NO: 283)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAA
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTAGATCTGATTTTGAAAGCTTA
GCTTTAAATTTCTTTTAATTAATAAAAAAAAAATGCTAGGCAGAATGACTCAAATTACGTT
GGATACAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATAC
TAAAAATTAAGTGTGTGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACC
TTTATTTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAA
ATCTTGAGTTTGAATGCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGT
TGGTGAAAAAATAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGAT
TTTCTGTAAACGATCGGGAACCTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAG
AAGAACA AAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTG
TGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACAAAA
TTCTGAATCGGCCAAAGAGGTTTTCTTGATCATGAAAACGCCAACAAAAATTCTGAATC
GGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTC AAGGGAACCTTGAG
AGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAA
CACTGAAAGAACA ACTGAATTTTGAAGCAGTATGTTGATGGAGATCAGTGTGAGT
CCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTT
GGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTAACATGTAACATTA
AGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTTGC
TCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGT
GCCATTTCCATGTGGAAGAGTTTTCTGTTTCACAACTTCTAAGCTCACCCGTGCTGA
GACTGTTTTTCTGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGAT
AACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAA
GATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCA
TTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTA ACTGCTGCCCACTGTGTT
GAAACTGGTGTAAAATTACAGTTGTTCGCAGGTGAACATAATATTGAGGAGACAGA
ACATACAGAGCAAAAGCGAAATGTGATTTCGAATTATTCCTCACCACA ACTACAATG
CAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGA ACTGGACGAACCCTTAG
TGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGAACATCT
TCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGA
GATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCT
TCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGG
AGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGGAAG
GGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGC
AAATATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAAC
AAAGCTCACTTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCT
CCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCACTGTCTTTTCTAATAAAA
TGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGT
GGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT
GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTA

TCCCCGTGAGATCGCCCATCGGTATAATGATTTGGGAGAACAACATTTCAAAGGCCT
GTAAGTTATAATGCTGAAAGCCCACTTAATATTTCTGGTAGTATTAGTTAAAGTTTT
AAAACACCTTTTTCCACCTTGAGTGTGAGAATTGTAGAGCAGTGCTGTCCAGTAGAA
ATGTGTGCATTGACAGAAAGACTGTGGATCTGTGCTGAGCAATGTGGCAGCCAGAG
ATCACAAGGCTATCAAGCACTTTGCACATGGCAAGTGTAAGTGAAGAAGCACACATT
CAAATAATAGTTAATTTTAATTGAATGTATCTAGCCATGTGTGGCTAGTAGCTCCTTT
CCTGGAGAGAGAATCTGGAGCCACATCTAACTTGTTAAGTCTGGAATCTTATTTTT
TATTTCTGGAAAGGTCTATGAACTATAGTTTTGGGGGCAGCTCACTTACTAACTTTTA
ATGCAATAAGAATCTCATGGTATCTTGAGAACATTATTTTGTCTCTTTGTAGATCTAG
GAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAG
GCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAG
CGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

Полная последовательность P00354 (от ИКП до ИКП): (SEQ ID NO: 284)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCCTAGATCTTAGCCTCTGGCAA
ATGAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGCTCTGCTTTTTCCAGGG
GTGTGTTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTT
TTCTAGTAATGGAAGCCTGGTATTTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTT
CTAGATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACCCT
TAGGTGGTTATATTATTGATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTT
TGAAAGCTTAGCTTTAAATTTCTTTTAATTAATAAAAAAAAAATGCTAGGCAGAATGACTC
AAATTACGTTGGATACAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGAC
CATGCTATACTAAAAATTAAGTGTGTGTTACTAATTTTATAAATGGAGTTTCCAT
TTATATTTACCTTTATTTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGAC
AGAAACACTAAATCTTGAGTTTGAATGCACAGATATAAACACTTAACGGGTTTTAAA
AATAATAATGTTGGTGAAAAAATATAACTTTGAGTGTAGCAGAGAGGAACCATTGC
CACCTTCAGATTTTCTGTAAACGATCGGGAAGTGGCATCTTCAGGGAGTAGCTTAGG
TCAGTGAAGAGAAGAACAATAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTT
TAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAA
CGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGT
TTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAA
GCACGAGAAGTTTTTGAAAACACTGAAAGAACAACCTGAATTTTGAAGCAGTATGT
TGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATG
ACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAAGTGTGAAT
TAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTTGTAATAATAGT
GCTGATAACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCA
GAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAAC
TTCTAAGCTCACCCGTGCTGAGACTGTTTTTCTGATGTGGACTATGTAAATTTCTACT
GAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACTTC

ACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTT
TTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATT
GTAAGTGTGCTGCCACTGTGTTGAAACTGGTGTAAAATTACAGTTGTCGCAGGTGAA
CATAATATTGAGGAGACAGAACATACAGAGCAAAAGCGAAATGTGATTTCGAATTAT
TCCTCACCACAACACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCT
GGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGA
CAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGG
AAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACT
TGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTT
CTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGAC
CCCATGTTACTGAAGTGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTG
AAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTATGTC
AACTGGATTAAGGAAAAACAAAGCTCACTTAACTCGACTGTGCCTTCTAGTTGCC
AGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCC
CACTGTCCTTTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCA
TTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC
AATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAAC
CAGCTGGGGCTCTAGGGGGTATCCCCGTGAGATCGCCATCGGTATAATGATTTGGG
AGAACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCCTTAATATTTCT
GGTAGTATTAGTTAAAGTTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAATTGTA
GAGCAGTGCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGTGC
TGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCACATGGCAAG
TGTAAGTGAAGCACACATTCAAATAATAGTTAATTTTAATTGAATGTATCTAGCC
ATGTGTGGCTAGTAGCTCCTTTCTGGAGAGAGAATCTGGAGCCCACATCTAACTTG
TTAAGTCTGGAATCTTATTTTTTATTTCTGGAAAGGTCTATGAACTATAGTTTTGGGG
GCAGCTCACTTACTAACTTTTAAATGCAATAAGAATCTCATGGTATCTTGAGAACATT
ATTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAGTAAGGGGTCTATACTTA
AGTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAAAAAATGAGTAATTAATT
TATTTTTAGAAGGTCAATAGTATCATGTATTCAAATAACAGAGGTATATGGTTAGA
AAAGAAACAATTCAAAGGACTTATATAATATCTAGCCTTGACAATGAATAAATTTA
GAGAGTAGTTTGCCTGTTTGCCTCATGTTCATAAATCTATTGACACATATGTGCATCT
GCACTTCAGCATGGTAGAAGTCCATATTCAGATCTAGGAACCCCTAGTGATGGAGTT
GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGG
GCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGG
GAGTGGCCAA

P00350: 300/600 п. о. конструкция НА F9 (для G551) (SEQ ID NO: 285)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
GGTCGCCCCGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCTAGATCTAAGTATATTAGAGC
GAGTCTTTCTGCACACAGATCACCTTTCCTATCAACCCCACTAGCCTCTGGCAAAT

GAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGCTCTGCTTTTTCCAGGGGT
GTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTTTT
CTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTTCT
AGATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACCTTTT
TCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAAATTCAG
GTAAATTGGAAGAGTTTGTTC AAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAG
TGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACCTGAATT
TTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGG
CAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGG
AAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGT
TTTGTA AAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGAC
TTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTT
CTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTTCCTGATGTGGACTA
TGTA AATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATC
ATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCC
TTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAA
TGAAA AATGGATTGTA ACTGCTGCCACTGTGTTGAAACTGGTGT TAAAATTACAGT
TGTCGCAGGTGAACATAATATTGAGGAGACAGAACATAACAGAGCAAAGCGAAAT
GTGATT CGAATTATTCCTCACCACA ACTACAATGCAGCTATTAATAAGTACAACCAT
GACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCT
ATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTAT
GTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTAC
CTTAGAGTTCCACTTGTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCT
ATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGA
GATAGTGGGGGACCCCATGTTACTGAAGTGAAGGGACCAGTTTCTTAACTGGAAT
TATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGG
TATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAGCTCACTTAACCTCGACTGT
GCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTG
GAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGT
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA
GGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTG
AGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTA
TTGATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTT
AAATTTCTTTTAATTA AAAAAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATA
CAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAA
ATTA AAAAGTGTGTGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTAT
TTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAAGAT
CTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAC
TGAGGCCGCCCCGGGCAAAGCCCCGGGCGTCCGGGCGACCTTTGGTCCGCCGGCCTCAG
TGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

P00356: 300/2000 п. о. конструкция НА F9 (для G551) (SEQ ID NO: 286)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
GGTCGCCCCGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCTAGATCTAAGTATATTAGAGC
GAGTCTTTCTGCACACAGATCACCTTTCCSTATCAACCCCACTAGCCTCTGGCAAAAT
GAAGTGGGTAAACCTTTCTCCTCCTCCTCTTCGTCTCCGGCTCTGCTTTTTCCAGGGGT
GTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTTTT
CTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTTCT
AGATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACCTTTT
TCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTGAG
GTAAATTGGAAGAGTTTTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAG
TGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACCTGAATT
TTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGG
CAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGG
AAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGT
TTTGTA AAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTA CTGAGGGATATCGAC
TTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTT
CTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTA
TGTA AATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATC
ATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCC
TTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAA
TGAAA AATGGATTGTA ACTGCTGCCACTGTGTTGAAACTGGTGT TAAAATTACAGT
TGTCGCAGGTGAACATAATATTGAGGAGACAGAACATAACAGAGCAAAAGCGAAAT
GTGATTCGAATTATTCCTCACCACA ACTACAATGCAGCTATTAATAAGTACAACCAT
GACATTGCCCTTCTGGA ACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCT
ATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTAT
GTAAGTGGCTGGGGAAGAGTCTTCCACA AAGGGAGATCAGCTTTAGTTCTTCAGTAC
CTTAGAGTTCCACTTGTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCT
ATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGA
GATAGTGGGGGACCCCATGTTACTGAAGTGG AAGGGACCAGTTTCTTAACTGGAAT
TATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGG
TATCCCGGTATGTCAACTGGATTAAGGAAAAAACA AAGCTCACTTAACCTCGACTGT
GCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTG
GAAGGTGCCACTCCC ACTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGT
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA
GGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTG
AGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTA
TTGATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTT
AAATTTCTTTTAAATTA AAAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATA
CAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAA

ATTAAAAGTGTGTGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTAT
TTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAATCTT
GAGTTTGAATGCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGTTGGTG
AAAAAATATAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGATTTTCCT
GTAACGATCGGGAACCTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAGAAGAA
CAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGG
TTTTTCTCTCCCTGTTTCCACAGACAAGAGTGAGATCGCCCATCGGTATAATGATTTG
GGAGAACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCACTTAATATTT
CTGGTAGTATTAGTTAAAGTTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAATTGT
AGAGCAGTGCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGTG
CTGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCACATGGCAA
GTGTAACCTGAGAAGCACACATTCAAATAATAGTTAATTTTAATTGAATGTATCTAGC
CATGTGTGGCTAGTAGCTCCTTTCCCTGGAGAGAGAATCTGGAGCCCACATCTAACTT
GTTAAGTCTGGAATCTTATTTTTTATTTCTGGAAAGGTCTATGAACTATAGTTTTGGG
GGCAGCTCACTACTAACTTTTTAATGCAATAAGATCCATGGTATCTTGAGAACATTA
TTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAGTAAGGGGTCTATACTTAA
GTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAAAAAATGAGTAATTAATTT
ATTTTTAGAAGGTCAATAGTATCATGTATTCCAAATAACAGAGGTATATGGTTAGAA
AAGAAACAATTCAAAGGACTTATATAATATCTAGCCTTGACAATGAATAAATTTAG
AGAGTAGTTTGCCTGTTTGCCTCATGTTCATAAATCTATTGACACATATGTGCATCTG
CACTTCAGCATGGTAGAAGTCCATATTCCTTTGCTTGGAAGGCAGGTGTTCCCAT
ACGCCTCAGAGAATAGCTGACGGGAAGAGGCTTTCTAGATAGTTGTATGAAAGATA
TACAAAATCTCGCAGGTATACACAGGCATGATTTGCTGGTTGGGAGAGCCACTTGCC
TCATACTGAGGTTTTTTGTGTCTGCTTTTCAGAGTCCTGATTGCCTTTTCCCAGTATCTC
CAGAAATGCTCATAACGATGAGCATGCCAAATTAGTGCAGGAAGTAACAGACTTTGC
AAAGACGTGTGTTGCCGATGAGTCTGCCGCCAACTGTGACAAATCCCTTGTGAGTAC
CTTCTGATTTTGTGGATCTACTTTCCTGCTTTCTGGAACCTCTGTTTCAAAGCCAATCA
TGACTCCATCACTTAAGGCCCCGGGAACACTGTGGCAGAGGGCAGCAGAGAGATTG
ATAAAGCCAGGGTGATGGGAATTTTCTGTGGGACTCCATTTCATAGTAATTGCAGAA
GCTACAATACTCAAAAAGTCTCACCACATGACTGCCCAAATGGGAGCTTGACAG
TGACAGTGACAGTAGATATGCCAAAGTGGATGAGGGAAAGACCACAAGAGCTAAA
CCCTGTAAAAAGAACTGTAGGCAACTAAGGAATGCAGAGAGAAAGATCTAGGAAC
CCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCG
CCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAG
CGAGCGCGCAGAGAGGGAGTGGCCAA

P00362: 300/1500 п. о. конструкция НА F9 (для G551) (SEQ ID NO: 287)

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA
GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCCTAGATCTAAGTATATTAGAGC
GAGTCTTTCTGCACACAGATCACCTTTCCTATCAACCCCACTAGCCTCTGGCAAAT

GAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGCTCTGCTTTTTCCAGGGGT
GTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTTTT
CTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTTCT
AGATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACCTTTT
TCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAAATTCAG
GTAAATTGGAAGAGTTTGTTC AAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAG
TGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACCTGAATT
TTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGG
CAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGG
AAAGA ACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGT
TTTGTA AAAATAGTGCTGATAACAAGGTGGTTTTGCTCCTGTACTGAGGGATATCGAC
TTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTT
CTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTTCCTGATGTGGACTA
TGTA AATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATC
ATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCC
TTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAA
TGAAA AATGGATTGTA ACTGCTGCCACTGTGTTGAAACTGGTGTAAAATTACAGT
TGTCGCAGGTGAACATAATATTGAGGAGACAGAACATAACAGAGCAAAGCGAAAT
GTGATTCGAATTATTCCTCACCACA ACTACAATGCAGCTATTAATAAGTACAACCAT
GACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCT
ATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTAT
GTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTAC
CTTAGAGTTCCACTTGTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCT
ATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGA
GATAGTGGGGGACCCCATGTTACTGAAGTGG AAGGGACCAGTTTCTTAACTGGAAT
TATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGG
TATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAGCTCACTTAACCTCGACTGT
GCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTG
GAAGGTGCCACTCCC ACTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGT
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA
GGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTG
AGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTA
TTGATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTT
AAATTTCTTTTAATTA AAAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATA
CAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAA
ATTA AAAAGTGTGTGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTAT
TTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAATCTT
GAGTTTGAATGCACAGATATAAACACTTAACGGGTTTTTAAAAATAATAATGTTGGTG
AAAAAATATAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGATTTTCCT
GTAACGATCGGGA ACTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAGAAGAA

CAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGG
 TTTTCTCTCCCTGTTTCCACAGACAAGAGTGAGATCGCCCATCGGTATAATGATTTG
 GGAGAACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCACTTAATATTT
 CTGGTAGTATTAGTTAAAGTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAATTGT
 AGAGCAGTGCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGTG
 CTGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCACATGGCAA
 GTGTAAGTGAAGAAGCACACATTCAAATAATAGTTAATTTTAATTGAATGTATCTAGC
 CATGTGTGGCTAGTAGCTCCTTTCCCTGGAGAGAGAATCTGGAGCCCACATCTAACTT
 GTTAAGTCTGGAATCTTATTTTTTTATTTCTGGAAAGGTCTATGAACTATAGTTTTGGG
 GGCAGCTCACTTACTAACTTTTAATGCAATAAGATCCATGGTATCTTGAGAACATTA
 TTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAGTAAGGGGTCTATACTTAA
 GTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAAAAAATGAGTAATTAATTT
 ATTTTTAGAAGGTCAATAGTATCATGTATTCCAAATAACAGAGGTATATGGTTAGAA
 AAGAAACAATTCAAAGGACTTATATAATATCTAGCCTTGACAATGAATAAATTTAG
 AGAGTAGTTTGCCTGTTTGCCTCATGTTTCATAAATCTATTGACACATATGTGCATCTG
 CACTTCAGCATGGTAGAAGTCCATATTCCTTTGCTTGGAAGGCAGGTGTTCCCAT
 ACGCCTCAGAGAATAGCTGACGGGAAGAGGCTTTCTAGATAGTTGTATGAAAGATA
 TACAAAATCTCGCAGGTATACACAGGCATGATTTGCTGGTTGGGAGAGCCACTTAG
 ATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTC
 ACTGAGGCCGCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTC
 AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

OPC Cas9 (SEQ ID NO: 703)

ATGGATAAGAAGTACTCAATCGGGCTGGATATCGGAACTAATTCCGTGGGTT
 GGGCAGTGATCACGGATGAATACAAAGTGCCGTCCAAGAAGTTCAAGGTCCTGGGG
 AACACCGATAGACACAGCATCAAGAAAAATCTCATCGGAGCCCTGCTGTTTGACTC
 CGGCGAAACCGCAGAAGCGACCCGGCTCAAACGTACCGCGAGGCGACGCTACACCC
 GGCGGAAGAATCGCATCTGCTATCTGCAAGAGATCTTTTCGAACGAAATGGCAAAG
 GTCGACGACAGCTTCTTCCACCGCCTGGAAGAATCTTTCCTGGTGGAGGAGGACAA
 GAAGCATGAACGGCATCCTATCTTTGGAAACATCGTCGACGAAGTGGCGTACCACG
 AAAAGTACCCGACCATCTACCATCTGCGGAAGAAGTTGGTTGACTCAACTGACAAG
 GCCGACCTCAGATTGATCTACTTGGCCCTCGCCCATATGATCAAATTCGCGGACAC
 TTCCTGATCGAAGGCGATCTGAACCCTGATAACTCCGACGTGGATAAGCTTTTCATT
 CAACTGGTGCAGACCTACAACCAACTGTTTCGAAGAAAACCAATCAATGCTAGCGG
 CGTCGATGCCAAGGCCATCCTGTCCGCCCGGCTGTGCAAGTCGCGGCGCCTCGAAA
 ACCTGATCGCACAGCTGCCGGGAGAGAAAAAGAACGGACTTTTCGGCAACTTGATC
 GCTCTCTCACTGGGACTCACTCCCAATTTCAAGTCCAATTTTGACCTGGCCGAGGAC
 GCGAAGCTGCAACTCTCAAAGGACACCTACGACGACGACTTGGACAATTTGCTGGC
 ACAAATTGGCGATCAGTACGCGGATCTGTTCCCTTGCCGCTAAGAACCTTTCGGACGC
 AATCTTGCTGTCCGATATCCTGCGCGTGAACACCGAAATAACCAAAGCGCCGCTTAG
 CGCCTCGATGATTAAGCGGTACGACGAGCATCACCAGGATCTCACGCTGCTCAAAG

CGCTCGTGAGACAGCAACTGCCTGAAAAGTACAAGGAGATCTTCTTCGACCAGTCC
AAGAATGGGTACGCAGGGTACATCGATGGAGGCGCTAGCCAGGAAGAGTTCTATAA
GTTCATCAAGCCAATCCTGAAAAGATGGACGGAACCGAAGAAGTCTGCTGGTCAAGC
TGAACAGGGAGGATCTGCTCCGGAAACAGAGAACCTTTGACAACGGATCCATTCCC
CACCAGATCCATCTGGGTGAGCTGCACGCCATCTTGCGGCGCCAGGAGGACTTTTAC
CCATTCTCAAGGACAACCGGGAAAAGATCGAGAAAATTCTGACGTTCCGCATCCC
GTATTACGTGGGCCCACTGGCGCGCGGCAATTCGCGCTTCGCGTGGATGACTAGAA
AATCAGAGGAAACCATCACTCCTTGGAAATTCGAGGAAGTTGTGGATAAGGGAGCT
TCGGCACAAAGCTTCATCGAACGAATGACCAACTTCGACAAGAATCTCCCAAACGA
GAAGGTGCTTCTAAGCACAGCCTCCTTTACGAATACTTCACTGTCTACAACGAACT
GACTAAAGTGAAATACGTTACTGAAGGAATGAGGAAGCCGGCCTTTCTGTCCGGAG
AACAGAAGAAAGCAATTGTTCGATCTGCTGTTCAAGACCAACCGCAAGGTGACCGTC
AAGCAGCTTAAAGAGGACTACTTCAAGAAGATCGAGTGTTCGACTCAGTGGAAAT
CAGCGGGGTGGAGGACAGATTCAACGCTTCGCTGGGAACCTATCATGATCTCCTGA
AGATCATCAAGGACAAGGACTTCCTTGACAACGAGGAGAACGAGGACATCCTGGAA
GATATCGTCCTGACCTTGACCCTTTTCGAGGATCGCGAGATGATCGAGGAGAGGCTT
AAGACCTACGCTCATCTCTTCGACGATAAGGTCATGAAACAACCTCAAGCGCCGCCG
GTACACTGGTTGGGGCCGCCTCTCCCGCAAGCTGATCAACGGTATTCGCGATAAACA
GAGCGGTAAAACCTATCCTGGATTTCTCAAATCGGATGGCTTCGCTAATCGTAACTT
CATGCAATTGATCCACGACGACAGCCTGACCTTTAAGGAGGACATCCAAAAGCAC
AAGTGTCCGGACAGGGAGACTCACTCCATGAACACATCGCGAATCTGGCCGGTTTCG
CCGGCGATTAAGAAGGGAATTCTGCAAACCTGTGAAGGTGGTCGACGAGCTGGTGAA
GGTCATGGGACGGCACAAACCGGAGAATATCGTGATTGAAATGGCCCGAGAAAACC
AGACTACCCAGAAGGGCCAGAAAAACTCCCGCGAAAGGATGAAGCGGATCGAAGA
AGGAATCAAGGAGCTGGGCAGCCAGATCCTGAAAGAGCACCCGGTGGAAAACACG
CAGCTGCAGAACGAGAAGCTCTACCTGTACTATTTGCAAATGGACGGGACATGTA
CGTGGACCAAGAGCTGGACATCAATCGGTTGTCTGATTACGACGTGGACCACATCGT
TCCACAGTCCTTTCTGAAGGATGACTCGATCGATAACAAGGTGTTGACTCGCAGCGA
CAAGAACAGAGGGAAGTCAGATAATGTGCCATCGGAGGAGGTTCGTGAAGAAGATG
AAGAATTACTGGCGGCAGCTCCTGAATGCGAAGCTGATTACCCAGAGAAAGTTTGA
CAATCTCACTAAAGCCGAGCGCGGCGGACTCTCAGAGCTGGATAAGGCTGGATTCA
TCAAACGGCAGCTGGTCGAGACTCGGCAGATTACCAAGCACGTGGCGCAGATCTTG
GACTCCCGCATGAACACTAAATACGACGAGAACGATAAGCTCATCCGGGAAGTGAA
GGTGATTACCCTGAAAAGCAAACCTTGTGTTCGGACTTTCGGAAGGACTTTCAGTTTTA
CAAAGTGAGAGAAATCAACAACCTACCATCACGCGCATGACGCATACCTCAACGCTG
TGGTCGGTACCGCCCTGATCAAAAAGTACCCTAAACTTGAATCGGAGTTTGTGTACG
GAGACTACAAGGTCTACGACGTGAGGAAGATGATAGCCAAGTCCGAACAGGAAATC
GGGAAAGCAAACCTGCGAAATACTTCTTTTACTCAAACATCATGAACTTTTTCAAGACT
GAAATTACGCTGGCCAATGGAGAAATCAGGAAGAGGCCACTGATCGAAACTAACGG
AGAAACGGGCGAAATCGTGTGGGACAAGGGCAGGGACTTCGCAACTGTTTCGCAAAG

TGCTCTCTATGCCGCAAGTCAATATTGTGAAGAAAACCGAAGTGCAAACCGGCGGA
TTTTCAAAGGAATCGATCCTCCCAAAGAGAAATAGCGACAAGCTCATTGCACGCAA
GAAAGACTGGGACCCGAAGAAGTACGGAGGATTTCGATTTCGCCGACTGTTCGATACT
CCGTCCTCGTGGTGGCCAAGGTGGAGAAGGGAAAGAGCAAAAAGCTCAAATCCGTC
AAAGAGCTGCTGGGGATTACCATCATGGAACGATCCTCGTTTCGAGAAGAACCCGAT
TGATTTCTCGAGGCGAAGGGTTACAAGGAGGTGAAGAAGGATCTGATCATCAAAC
TCCCAAGTACTCACTGTTTCGAACTGGAAAATGGTCGGAAGCGCATGCTGGCTTCGG
CCGGAGAACTCCAAAAAGGAAATGAGCTGGCCTTGCTAGCAAGTACGTCAACTTC
CTCTATCTTGCTTCGCACTACGAAAACTCAAAGGGTCACCGGAAGATAACGAACA
GAAGCAGCTTTTCGTGGAGCAGCACAAGCATTATCTGGATGAAATCATCGAACAAA
TCTCCGAGTTTTCAAAGCGCGTGATCCTCGCCGACGCCAACCTCGACAAAGTCCTGT
CGGCCTACAATAAGCATAGAGATAAGCCGATCAGAGAACAGGCCGAGAACATTATC
CACTTGTTACCCTGACTAACCTGGGAGCCCCAGCCGCCTTCAAGTACTTCGATACT
ACTATCGATCGCAAAAGATACACGTCCACCAAGGAAGTTCTGGACGCGACCCTGAT
CCACCAAAGCATCACTGGACTCTACGAAACTAGGATCGATCTGTTCGACGCTGGGTG
GCGAT

U-3aB. OPC Cas9 (SEQ ID NO: 704)

ATGGACAAGAAGTACAGCATCGGACTGGACATCGGAACAAACAGCGTCGGA
TGGGCAGTCATCACAGACGAATACAAGGTCCCGAGCAAGAAGTTCAAGGTCCTGGG
AAACACAGACAGACACAGCATCAAGAAGAACCTGATCGGAGCACTGCTGTTTCGACA
GCGGAGAAACAGCAGAAGCAACAAGACTGAAGAGAACAGCAAGAAGAAGATACAC
AAGAAGAAAGAACAGAATCTGCTACCTGCAGGAAATCTTCAGCAACGAAATGGCAA
AGGTCGACGACAGCTTCTTCCACAGACTGGAAGAAAGCTTCTGCTCGAAGAAGAC
AAGAAGCACGAAAGACACCCGATCTTCGGAAACATCGTCGACGAAGTCGCATACCA
CGAAAAGTACCCGACAATCTACCACCTGAGAAAGAAGCTGGTCGACAGCACAGACA
AGGCAGACCTGAGACTGATCTACCTGGCACTGGCACACATGATCAAGTTCAGAGGA
CACTTCTGATCGAAGGAGACCTGAACCCGGACAACAGCGACGTCGACAAGCTGTT
CATCCAGCTGGTCCAGACATAACAACCAGCTGTTTCGAAGAAAACCCGATCAACGCAA
GCGGAGTCGACGCAAAGGCAATCCTGAGCGCAAGACTGAGCAAGAGCAGAAGACT
GGAAAACCTGATCGCACAGCTGCCGGGAGAAAAGAAGAACGGACTGTTTCGGAAAC
CTGATCGCACTGAGCCTGGGACTGACACCGAACTTCAAGAGCAACTTCGACCTGGC
AGAAGACGCAAAGCTGCAGCTGAGCAAGGACACATACGACGACGACCTGGACAAC
CTGCTGGCACAGATCGGAGACCAGTACGCAGACCTGTTCTGTCAGCAAAGAACCT
GAGCGACGCAATCCTGCTGAGCGACATCCTGAGAGTCAACACAGAAATCACAAAGG
CACCGCTGAGCGCAAGCATGATCAAGAGATACGACGAACACCACCAGGACCTGACA
CTGCTGAAGGCACTGGTCAGACAGCAGCTGCCGGAAAAGTACAAGGAAATCTTCTT
CGACCAGAGCAAGAACGGATACGCAGGATACATCGACGGAGGAGCAAGCCAGGAA
GAATTCTACAAGTTCATCAAGCCGATCCTGGAAAAGATGGACGGAACAGAAGA
GCTGGTCAAGCTGAACAGAGAAGACCTGCTGAGAAAGCAGAGAACATTTCGACAAC
GGAAGCATCCCGCACCAAGATCCACCTGGGAGAACTGCACGCAATCCTGAGAAGACA

GGAAGACTTCTACCCGTTCTGAAAGGACAACAGAGAAAAGATCGAAAAGATCCTGACATT
CATTGAGAAATCCCGTACTACGTCGGACCGCTGGCAAGAGGAAACAGCAGATTCGCA
TGGATGACAAGAAAGAGCGAAGAAACAATCACACCGTGGAACTTCGAAGAAGTCG
TCGACAAGGGAGCAAGCGCACAGAGCTTCATCGAAAGAATGACAACTTCGACAAG
AACCTGCCGAACGAAAAGGTCTGCGCAAGCACAGCCTGCTGTACGAATACTTCAC
AGTCTACAACGAACTGACAAAGGTCAAGTACGTCACAGAAGGAATGAGAAAGCCG
GCATTCCTGAGCGGAGAACAGAAGAAGGCAATCGTCGACCTGCTGTTCAAGACAAA
CAGAAAGGTCACAGTCAAGCAGCTGAAGGAAGACTACTTCAAGAAGATCGAATGCT
TCGACAGCGTCGAAATCAGCGGAGTCGAAGACAGATTCAACGCAAGCCTGGGAACA
TACCACGACCTGCTGAAGATCATCAAGGACAAGGACTTCCTGGACAACGAAGAAAA
CGAAGACATCCTGGAAGACATCGTCCTGACACTGACACTGTTTCGAAGACAGAGAAA
TGATCGAAGAAAGACTGAAGACATACGCACACCTGTTTCGACGACAAGGTCATGAAG
CAGCTGAAGAGAAGAAGATACACAGGATGGGGAAGACTGAGCAGAAAGCTGATCA
ACGGAATCAGAGACAAGCAGAGCGGAAAGACAATCCTGGACTTCCTGAAGAGCGA
CGGATTCGCAAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACATTCA
AGGAAGACATCCAGAAGGCACAGGTCAGCGGACAGGGAGACAGCCTGCACGAACA
CATCGCAAACCTGGCAGGAAGCCCGGCAATCAAGAAGGGAATCCTGCAGACAGTCA
AGGTCGTCGACGAACTGGTCAAGGTCATGGGAAGACACAAGCCGGAAAACATCGTC
ATCGAAATGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAG
AAAGAATGAAGAGAATCGAAGAAGGAATCAAGGAACTGGGAAGCCAGATCCTGAA
GGAACACCCGGTCGAAAACACACAGCTGCAGAACGAAAAGCTGTACCTGTACTACC
TGCAGAACGGAAGAGACATGTACGTCGACCAGGAACTGGACATCAACAGACTGAGC
GACTACGACGTCGACCACATCGTCCCGCAGAGCTTCCTGAAGGACGACAGCATCGA
CAACAAGGTCCTGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGTCCCG
AGCGAAGAAGTCGTCAAGAAGATGAAGAATACTGGAGACAGCTGCTGAACGCAA
AGCTGATCACACAGAGAAAGTTCGACAACCTGACAAAGGCAGAGAGAGGAGGACT
GAGCGAACTGGACAAGGCAGGATTCATCAAGAGACAGCTGGTTCGAAACAAGACAG
ATCACAAAGCACGTCGCACAGATCCTGGACAGCAGAATGAACACAAAGTACGACGA
AAACGACAAGCTGATCAGAGAAGTCAAGGTCATCACACTGAAGAGCAAGCTGGTCA
GCGACTTCAGAAAGGACTTCCAGTTCTACAAGGTCAGAGAAATCAACAACCTACCAC
CACGCACACGACGCATACCTGAACGCAGTCGTCGGAACAGCACTGATCAAGAAGTA
CCCGAAGCTGGAAAGCGAATTCGTCTACGGAGACTACAAGGTCTACGACGTCAGAA
AGATGATCGCAAAGAGCGAACAGGAAATCGGAAAGGCAACAGCAAAGTACTTCTTC
TACAGCAACATCATGAACTTCTTCAAGACAGAAATCACACTGGCAAACGGAGAAAT
CAGAAAGAGACCGCTGATCGAAACAAACGGAGAAACAGGAGAAATCGTCTGGGAC
AAGGGAAGAGACTTCGCAACAGTCAGAAAGGTCCTGAGCATGCCGCAGGTCAACAT
CGTCAAGAAGACAGAAGTCCAGACAGGAGGATTCAGCAAGGAAAGCATCCTGCCG
AAGAGAAACAGCGACAAGCTGATCGCAAGAAAGAAGGACTGGGACCCGAAGAAGT
ACGGAGGATTCGACAGCCCGACAGTCGCATACAGCGTCCTGGTCGTCGCAAAGGTC
GAAAAGGGAAAGAGCAAGAAGCTGAAGAGCGTCAAGGAACTGCTGGGAATCACAA

TCATGGAAAGAAGCAGCTTCGAAAAGAACCCGATCGACTTCCTGGAAGCAAAGGGA
TACAAGGAAGTCAAGAAGGACCTGATCATCAAGCTGCCGAAGTACAGCCTGTTCGA
ACTGGAAAACGGAAGAAAGAGAATGCTGGCAAGCGCAGGAGAAGTGCAGAAGGGA
AACGAACTGGCACTGCCGAGCAAGTACGTCAACTTCCTGTACCTGGCAAGCCACTA
CGAAAAGCTGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCTGTTCGTCGAA
CAGCACAAGCACTACCTGGACGAAATCATCGAACAGATCAGCGAATTCAGCAAGAG
AGTCATCCTGGCAGACGCAAACCTGGACAAGGTCCTGAGCGCATAACAACAAGCACA
GAGACAAGCCGATCAGAGAACAGGCAGAAAACATCATCCACCTGTTCACTGACA
AACCTGGGAGCACCGGCAGCATTCAAGTACTTCGACACAACAATCGACAGAAAGAG
ATACACAAGCACAAAGGAAGTCCTGGACGCAACACTGATCCACCAGAGCATCACAG
GACTGTACGAAAACAAGAATCGACCTGAGCCAGCTGGGAGGAGACGGAGGAGGAAG
CCCGAAGAAGAAGAGAAAGGTCTAG

мРНК, содержащая U-зав. Cas9 (SEQ ID NO: 705)

GGGUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGU
UGCAGGCCUUAUUCGGAUCCGCCACCAUGGACAAGAAGUACAGCAUCGGACUGGA
CAUCGGAACAAACAGCGUCGGAUUGGCAGUCAUCACAGACGAAUACAAGGUCCCG
AGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAAC
CUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG
AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUG
CAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUCCACAGAC
UGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCU
UCGGAAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAAUCUACCA
CCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUAC
CUGGCACUGGCACACAUGAUCAAGUUCAGAGGACACUCCUGAUCGAAGGAGACC
UGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUA
CAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCA
AUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGC
UGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGG
GACUGACACCCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCA
GCUGAGCAAGGACACAUACGACGACGACCUGGACAACCUGCUGGCACAGAUCGGA
GACCAGUACGCAGACCUGUUCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGC
UGAGCGACAUCCUGAGAGUCAACACAGAAAUCACAAAGGCACCCGUCGAGCGCAAG
CAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG
GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAG
AACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAG
UUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAAACAGAAGAACUGCUGGUCAAG
CUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUCGACAACGGAAGCAUC
CCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACU
UCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCUGACAUCU
GAAUCCCGUACUACGUCGGACCCGUCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAU

GACAAGAAAGAGCGAAGAAACAAUCACACCGUGGAACUUCGAAGAAGUCGUCGA
CAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAA
CCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACA
GUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCG
GCAUUCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUAAGACA
AACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUAAGAAGAUCGAA
UGCUUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGG
GAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCUGGACAACGA
AGAAAACGAAGACAUCCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGAC
AGAGAAAUGAUCGAAGAAAGACUGAAGACAUCGCACACCUGUUCGACGACAAG
GUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGA
AAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUC
CUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACA
GCCUGACAUUAAGGAAGACAUCCAGAAGGCACAGGUCAGCGGACAGGGAGACA
GCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGCAAUCAAGAAGGGAAU
CCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAA
GCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGG
ACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACU
GGGAAGCCAGAUCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAA
AAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAAC
UGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUU
CCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGA
GGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUAC
UGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUG
ACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAG
AGACAGCUGGUCGAAACAAGACAGAUCAACAAGCACGUCGCACAGAUCCUGGACA
GCAGAAUGAACACAAAGUACGACGAAAACGACAAGCUGAUCAGAGAAGUCAAGG
UCAUCACACUGAAGAGCAAGCUGGUCAGCGACUUCAGAAAGGACUUCAGUUCUA
CAAGGUCAGAGAAAUCAACAACUACCACCACGCACACGACGCAUACCUGAACGCA
GUCGUCGGAACAGCACUGAUCAAGAAGUACCCGAAGCUGGAAAGCGAAUUCGUC
UACGGAGACUACAAGGUCUACGACGUCAGAAAGAUGAUCGCAAAGAGCGAACAG
GAAAUCGGAAAGGCAACAGCAAAGUACUUCUUCUACAGCAACAUCAUGAACUUC
UUCAAGACAGAAAUCACACUGGCAAACGGAGAAAUCAGAAAGAGACCCGUCGAUC
GAAACAAACGGAGAAACAGGAGAAAUCGUCUGGGACAAGGGAAGAGACUUCGCA
ACAGUCAGAAAGGUCCUGAGCAUGCCGCAGGUCAACAUCGUCAAGAAGACAGAA
GUCCAGACAGGAGGAUUCAGCAAGGAAAGCAUCCUGCCGAAGAGAAACAGCGAC
AAGCUGAUCGCAAGAAAGAAGGACUGGGACCCGAAGAAGUACGGAGGAUUCGAC
AGCCCGACAGUCGCAUACAGCGUCCUGGUCGUCGCAAAGGUCGAAAAGGGAAAGA
GCAAGAAGCUGAAGAGCGUCAAGGAACUGCUGGGAAUCACAAUCAUGGAAAGAA
GCAGCUUCGAAAAGAACCCGAUCGACUUCUGGAAGCAAAGGGAUACAAGGAAG

UCAAGAAGGACCUGAUCAUCAAGCUGCCGAAGUACAGCCUGUUCGAACUGGAAA
ACGGAAGAAAGAGAAUGCUGGCAAGCGCAGGAGAACUGCAGAAGGGAAACGAAC
UGGCACUGCCGAGCAAGUACGUCAACUUCUGUACCUGGCAAGCCACUACGAAAA
GCUGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCUGUUCGUCGAACAGCAC
AAGCACUACCUGGACGAAAUCAUCGAACAGAUCAGCGAAUUCAGCAAGAGAGUC
AUCCUGGCAGACGCAAACCUGGACAAGGUCCUGAGCGCAUACAACAAGCACAGAG
ACAAGCCGAUCAGAGAACAGGCAGAAAACAUCAUCCACCUGUUCACACUGACAAA
CCUGGGAGCACCCGGCAGCAUUCAAGUACUUCGACACAACAUCGACAGAAAGAGA
UACACAAGCACAAAGGAAGUCCUGGACGCAACACUGAUCCACCAGAGCAUCACAG
GACUGUACGAAACAAGAAUCGACCUGAGCCAGCUGGGAGGAGACGGAGGAGGAA
GCCCGAAGAAGAAGAGAAAGGUCUAGCUAGCCAUCACAUUUAAAAGCAUCUCAG
CCUACCAUGAGAAUAAGAGAAAGAAAAUGAAGAUCAAUAGCUUAUUCAUCUCUU
UUUCUUUUUCGUUGGUGUAAAGCCAACACCCUGUCUAAAAACAUAUUUCUU
UAAUCAUUUUGCCUCUUUCUCUGUGCUUCAAUUAAUAAAAAUGGAAAGAACC
UCGAGAA
AA

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Двухнаправленная конструкция нуклеиновой кислоты, содержащая
 - а) первый сегмент, содержащий кодирующую последовательность агента, такого как полипептид; и
 - б) второй сегмент, содержащий обратную комплементарную последовательность кодирующей последовательности агента,
причем конструкция не содержит промотор, который управляет экспрессией агента, такого как полипептид.
2. Двухнаправленная конструкция нуклеиновой кислоты, содержащая
 - а) первый сегмент, содержащий кодирующую последовательность первого агента, такого как первый полипептид; и
 - б) второй сегмент, содержащий обратную комплементарную последовательность кодирующей последовательности второго агента, такого как второй полипептид,
причем конструкция не содержит промотор, который управляет экспрессией агента(ов).
3. Двухнаправленная конструкция нуклеиновой кислоты по п. 1 или п. 2, отличающаяся тем, что второй сегмент расположен 3' относительно первого сегмента.
4. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-3, отличающаяся тем, что кодирующая последовательность обратной комплементарной последовательности во втором сегменте имеет частоту использования кодонов, отличную от первой кодирующей последовательности первого сегмента.
5. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-4, отличающаяся тем, что второй сегмент содержит нуклеотидную последовательность, имеющую по меньшей мере около 30%, около 35%, около 40%, около 45%, около 50%, около 55%, около 60%, около 65%, около 70%, около 75%, около 80%, около 85%, около 90%, около 95%, около 97% или около 99% комплементарности с кодирующей последовательностью в первом сегменте.
6. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-5, отличающаяся тем, что кодирующая последовательность второго сегмента кодирует полипептид с использованием одного или более альтернативных кодонов для одной или более аминокислот, кодируемых кодирующей последовательностью в первом сегменте.
7. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-6, отличающаяся тем, что второй сегмент содержит обратную комплементарную последовательность кодирующей последовательности первого сегмента или ее фрагмент.
8. Двухнаправленная конструкция нуклеиновой кислоты по п. 7, отличающаяся тем, что обратная комплементарная последовательность является:
 - а. не в значительной степени комплементарной кодирующей последовательности первого сегмента;
 - б. не в значительной степени комплементарной фрагменту кодирующей последовательности первого сегмента;

- c. высококомплементарной кодирующей последовательности первого сегмента;
- d. высококомплементарной фрагменту кодирующей последовательности первого сегмента;
- e. по меньшей мере на 60% идентичной обратной комплементарной последовательности кодирующей последовательности первого сегмента;
- f. по меньшей мере на 70% идентичной обратной комплементарной последовательности кодирующей последовательности первого сегмента;
- f. по меньшей мере на 90% идентичной обратной комплементарной последовательности кодирующей последовательности первого сегмента;
- g. на 50-80% идентичной обратной комплементарной последовательности кодирующей последовательности первого сегмента; и/или
- h. на 60-100% идентичной обратной комплементарной последовательности кодирующей последовательности первого сегмента.

9. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-8, отличающаяся тем, что конструкция не содержит плечо гомологии.

10. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-9, отличающаяся тем, что первый сегмент связан со вторым сегментом линкером.

11. Двухнаправленная конструкция нуклеиновой кислоты по п. 10, отличающаяся тем, что длина линкера составляет 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 500, 1000, 1500, 2000 нуклеотидов.

12. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-11, отличающаяся тем, что каждый из первого и второго сегментов содержит последовательность полиаденилирования, такую как сигнальная последовательность полиаденилирования или хвостовая последовательность полиаденилирования.

13. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-12, отличающаяся тем, что конструкция содержит сайт акцептора сплайсинга.

14. Двухнаправленная конструкция нуклеиновой кислоты по п. 13, отличающаяся тем, что конструкция содержит первый сайт акцептора сплайсинга, расположенный 5' относительно первого сегмента, и второй сайт акцептора сплайсинга, расположенный 3' относительно второго сегмента.

15. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-14, отличающаяся тем, что конструкция является двухцепочечной, необязательно, двухцепочечной ДНК.

16. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-15, отличающаяся тем, что конструкция является одноцепочечной, необязательно, одноцепочечной ДНК.

17. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-16, отличающаяся тем, что последовательность, кодирующая полипептид, является кодон-оптимизированной.

18. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп.

1-17, отличающаяся тем, что конструкция содержит одну или более из следующих концевых структур: шпильку, петли, инвертированные концевые повторы (ИКП) или тороид.

19. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-18, отличающаяся тем, что конструкция содержит один, два или три инвертированных концевых повтора (ИКП).

20. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-19, отличающаяся тем, что конструкция содержит не более двух ИКП.

21. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-20, отличающаяся тем, что полипептид представляет собой секретлируемый полипептид.

22. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-20, отличающаяся тем, что полипептид представляет собой внутриклеточный полипептид.

23. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 2-22, отличающаяся тем, что первый полипептид и второй полипептид являются разными.

24. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-23, отличающаяся тем, что полипептид представляет собой вариантный полипептид.

25. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-24, отличающаяся тем, что полипептид представляет собой печеночный белок.

26. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-24, отличающаяся тем, что полипептид представляет собой не печеночный белок.

27. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-26, отличающаяся тем, что конструкция представляет собой независимую от гомологии конструкцию.

28. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-27, отличающаяся тем, что при экспрессии полипептид содержит гетерологичный сигнальный пептид.

29. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-28, отличающаяся тем, что нуклеиновая кислота кодирует гетерологичный сигнальный пептид.

30. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-27, отличающаяся тем, что нуклеиновая кислота не кодирует сигнальный пептид.

31. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-27, отличающаяся тем, что при экспрессии полипептид содержит собственный сигнальный пептид.

32. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-31, отличающаяся тем, что нуклеиновая кислота кодирует гетерологичный пептид.

33. Двухнаправленная конструкция нуклеиновой кислоты по п. 32, отличающаяся тем, что гетерологичный пептид представляет собой 2А.

34. Вектор, содержащий конструкцию по любому одному из пп. 1-33.

35. Вектор по п. 34, отличающийся тем, что вектор представляет собой вектор на основе аденоассоциированного вируса (AAV).

36. Вектор по п. 34 или п. 35, отличающийся тем, что AAV содержит одноцепочечный геном (ssAAV) или самокомплементарный геном (scAAV).

37. Вектор по п. 34 или п. 35, отличающийся тем, что AAV-вектор выбран из группы, состоящей из AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV12, rh10 и их гибридов.

38. Вирусный вектор, содержащий самокомплементарную (или двухцепочечную) конструкцию нуклеиновой кислоты, которая содержит нуклеотидную последовательность, кодирующую полипептид, при этом вектор не содержит промотор, управляющий экспрессией полипептида.

39. Вектор по п. 38, отличающийся тем, что вектор не содержит плечо гомологии.

40. Липидная наночастица, содержащая конструкцию по любому одному из пп. 1-33.

41. Клетка-хозяин, содержащая конструкцию по любому одному из пп. 1-33.

42. Клетка-хозяин по п. 41, отличающаяся тем, что клетка-хозяин представляет собой клетку печени, мышечную клетку или нейрональную клетку.

43. Клетка-хозяин по п. 41, отличающаяся тем, что клетка-хозяин относится к типу неделящихся клеток.

44. Клетка-хозяин по любому одному из пп. 41-43, отличающаяся тем, что клетка-хозяин экспрессирует полипептид, кодируемый конструкцией.

45. Клетка-хозяин по любому одному из пп. 41-44, отличающаяся тем, что клетка-хозяин представляет собой гепатоцит.

46. Способ модификации целевого локуса, включающий введение в клетку конструкции по любому одному из пп. 1-33, вектора по любому одному из пп. 34-39 или ЛНЧ по п. 40.

47. Способ внесения конструкции в клетку, включающий введение в клетку конструкции по любому одному из пп. 1-33, вектора по любому одному из пп. 34-39 или ЛНЧ по п. 40.

48. Способ экспрессии полипептида в клетке, включающий введение в клетку конструкции по любому одному из пп. 1-33, вектора по любому одному из пп. 34-39 или ЛНЧ по п. 40.

49. Способ повышения экспрессии полипептида в клетке, включающий введение в клетку конструкции по любому одному из пп. 1-33, вектора по любому одному из пп. 34-39 или ЛНЧ по п. 40.

50. Способ по любому одному из пп. 46-49, отличающийся тем, что конструкцию, вектор или ЛНЧ вводят *in vivo*.

51. Способ по любому одному из пп. 46-49, отличающийся тем, что конструкцию, вектор или ЛНЧ вводят *in vitro*.

52. Способ *ex vivo* модификации целевого локуса, включающий введение в клетку конструкции по любому одному из пп. 1-33, вектора по любому одному из пп. 34-39 или ЛНЧ по п. 40.

53. Способ по любому одному из пп. 46-52, отличающийся тем, что клетка относится к типу неделящихся клеток.

54. Способ по любому одному из пп. 46-53, дополнительно включающий введение в клетку нуклеазы, причем нуклеаза способна создавать сайт-специфический разрыв в оцДНК или дцДНК.

55. Способ по любому одному из пп. 46-54, включающий введение в клетку РНК-направляемого ДНК-связывающего агента и гидовой РНК (гРНК).

56. Способ по п. 55, отличающийся тем, что гРНК представляет собой одиночную гРНК (огРНК).

57. Способ по п. 55 или п. 56, отличающийся тем, что конструкцию, РНК-направляемый ДНК-связывающий агент и гРНК вводят в клетку одновременно.

58. Способ по п. 55 или п. 56, отличающийся тем, что конструкцию, РНК-направляемый ДНК-связывающий агент и гРНК вводят в клетку последовательно в любом порядке.

59. Способ по п. 58, отличающийся тем, что конструкцию вводят в клетку до введения гРНК и РНК-направляемого ДНК-связывающего агента.

60. Способ по п. 58, отличающийся тем, что РНК-направляемый ДНК-связывающий агент или РНК-направляемый ДНК-связывающий агент и гРНК в комбинации вводят в клетку до введения конструкции.

61. Способ по любому одному из пп. 54-60, отличающийся тем, что конструкцию вводят в векторе.

62. Способ по любому одному из пп. 54-60, отличающийся тем, что конструкцию вводят в липидной наночастице.

63. Способ по любому одному из пп. 54-62, отличающийся тем, что РНК-направляемый ДНК-связывающий агент вводят в виде нуклеиновой кислоты, кодирующей РНК-направляемый ДНК-связывающий агент.

64. Способ по любому одному из пп. 54-62, отличающийся тем, что РНК-направляемый ДНК-связывающий агент вводят в виде фермента Cas или мРНК, кодирующей РНК-направляемый ДНК-связывающий агент.

65. Способ по любому одному из пп. 54-62, отличающийся тем, что РНК-направляемый ДНК-связывающий агент представляет собой Cas-нуклеазу, такую как Cas9-нуклеаза.

66. Способ по п. 65, отличающийся тем, что Cas-нуклеаза представляет собой Cas-нуклеазу класса 2 или ее вариант.

67. Способ по любому одному из пп. 54-66, отличающийся тем, что Cas-нуклеаза представляет собой Cas9 *S. pyogenes* или ее вариант.

68. Способ по любому одному из пп. 54-67, отличающийся тем, что Cas-нуклеаза

представляет собой никазу.

69. Способ по любому одному из пп. 54-68, отличающийся тем, что локус представляет собой локус альбумина.

70. Способ по любому одному из пп. 54-69, отличающийся тем, что клетка относится к типу неделящихся клеток.

71. Способ по любому одному из пп. 54-70, отличающийся тем, что клетка представляет собой клетку печени, нейрональную клетку или мышечную клетку.

72. Клетка, полученная способом по любому одному из пп. 46-71.

73. Клетка по п. 72, отличающаяся тем, что клетка представляет собой клетку-хозяина.

74. Клетка-хозяин по п. 73, отличающаяся тем, что клетка-хозяин представляет собой клетку печени, мышечную клетку или нейрональную клетку.

75. Способ по любому одному из пп. 72-74, отличающийся тем, что клетка-хозяин относится к типу неделящихся клеток.

76. Клетка по любому одному из пп. 72-75, отличающаяся тем, что клетка-хозяин экспрессирует полипептид, кодируемый конструкцией.

77. Клетка по любому одному из пп. 72-76, отличающаяся тем, что клетка-хозяин представляет собой гепатоцит.

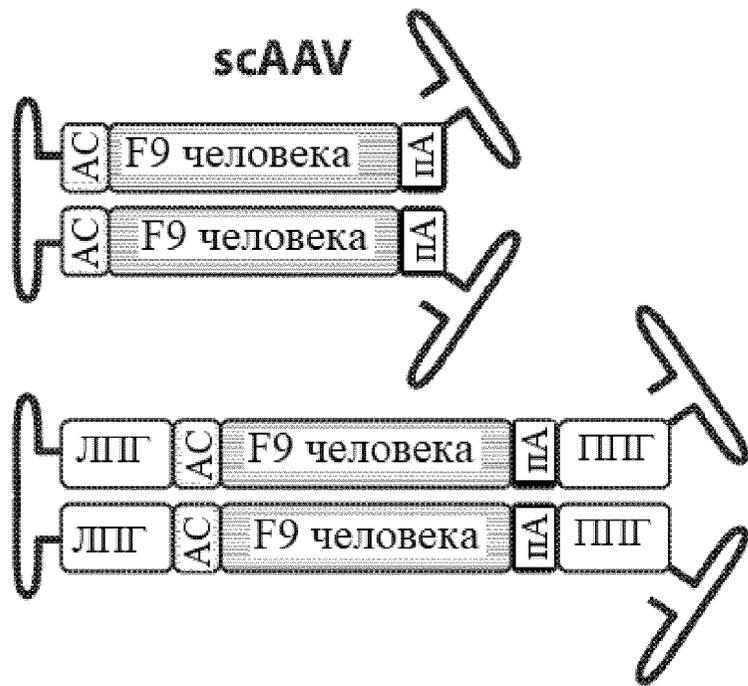
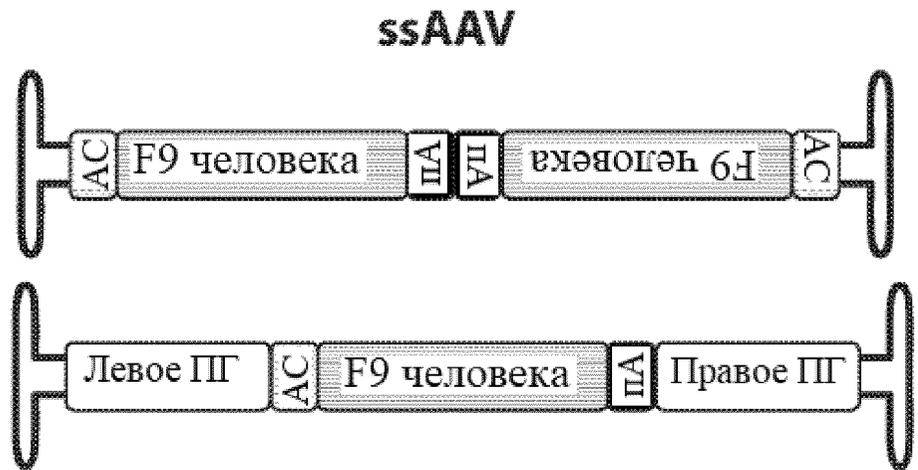
78. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-33, отличающаяся тем, что агент представляет собой терапевтический агент.

79. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-33, отличающаяся тем, что агент представляет собой одно или более из функциональной РНК, мРНК или полипептида.

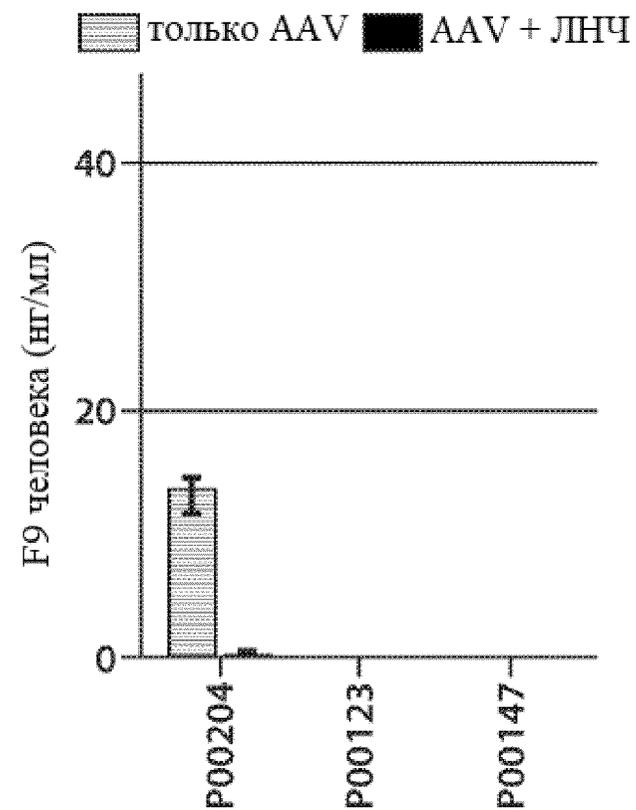
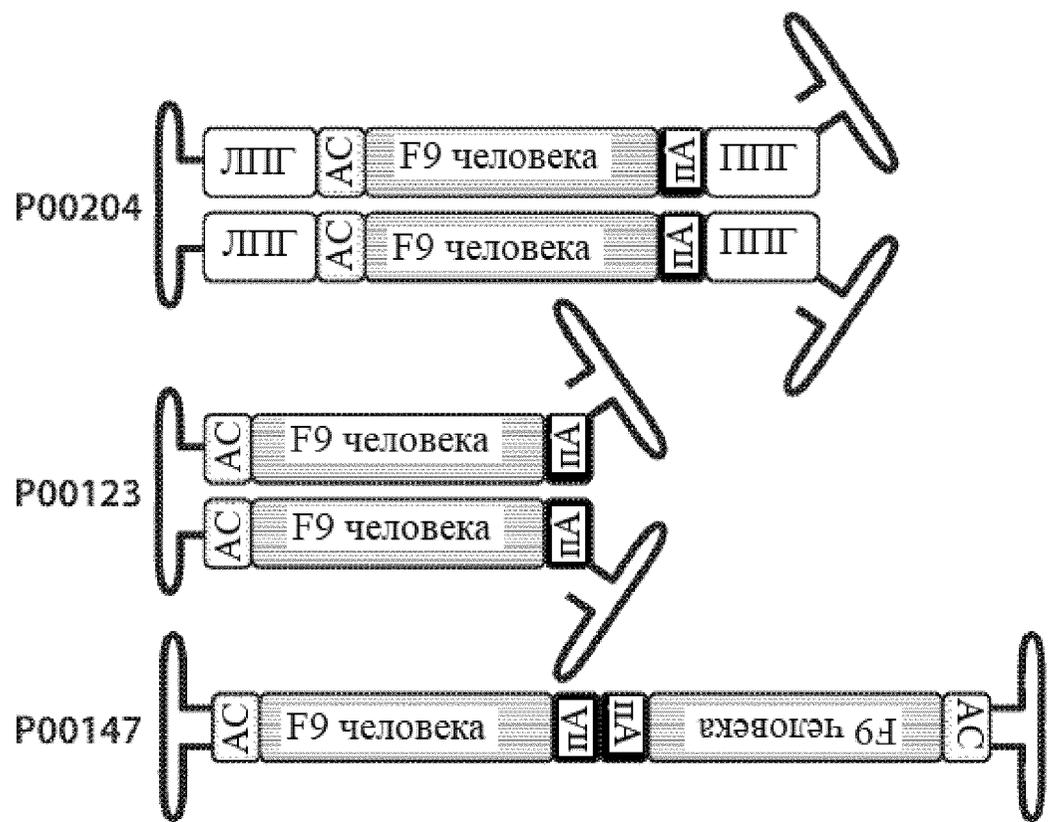
80. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-33, отличающаяся тем, что агент представляет собой полипептид.

81. Двухнаправленная конструкция нуклеиновой кислоты по любому одному из пп. 1-33, отличающаяся тем, что агент представляет собой терапевтический полипептид.

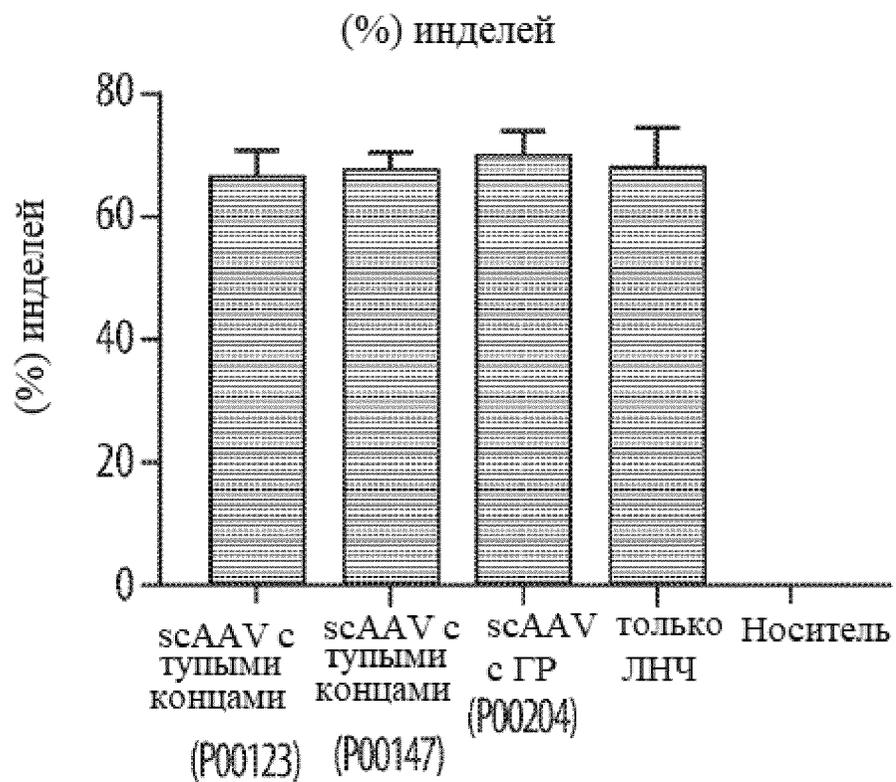
По доверенности



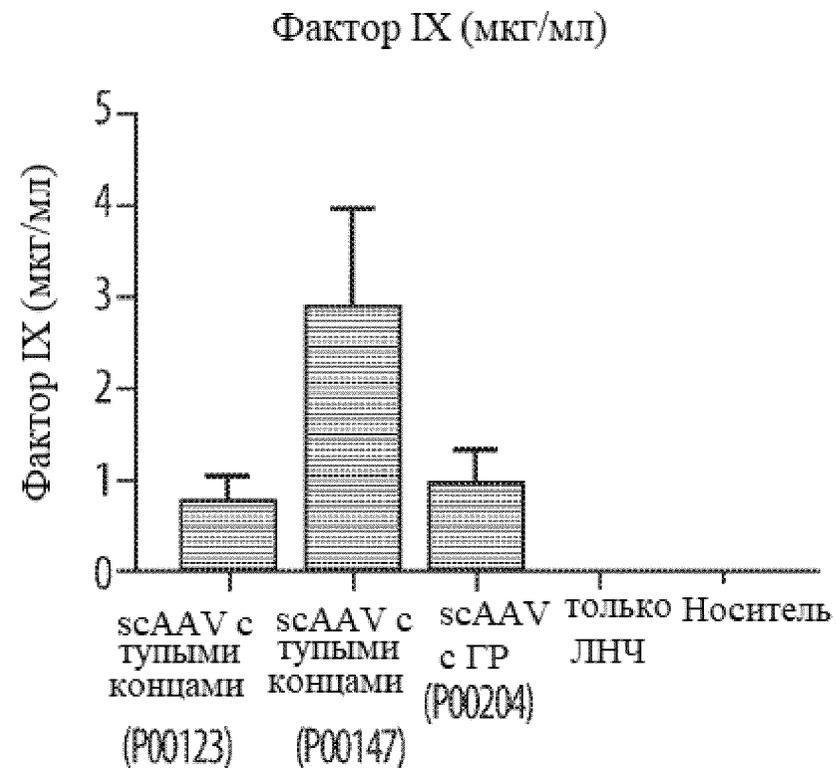
Фиг. 1



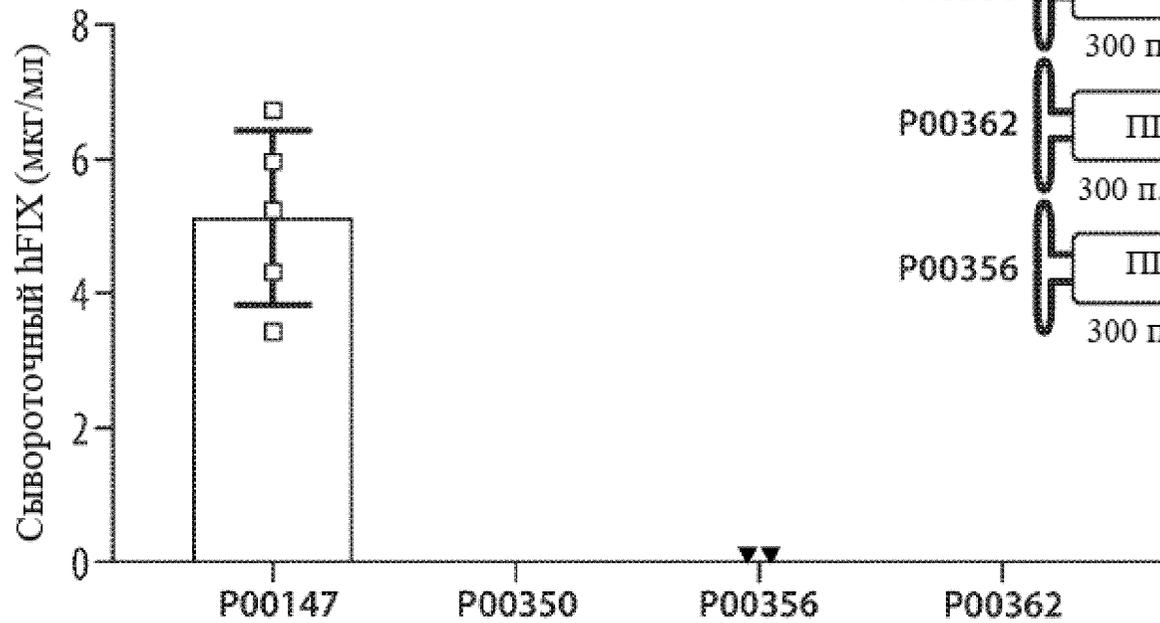
Фиг. 2



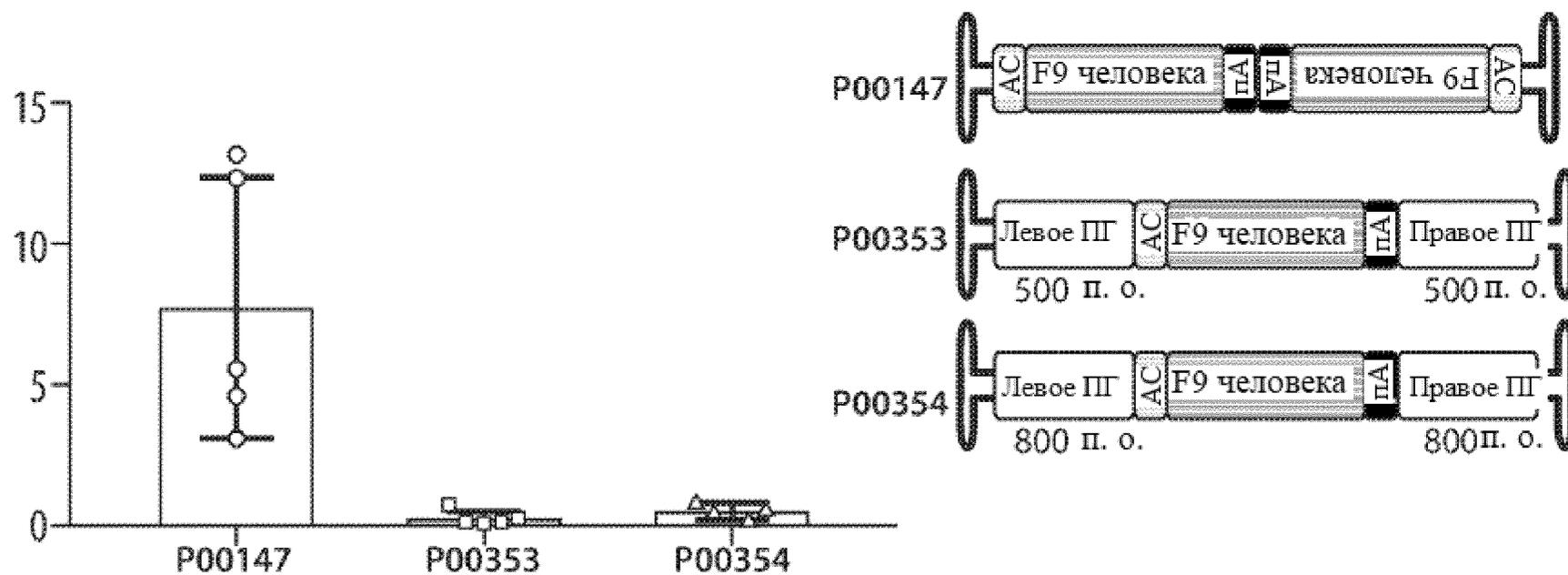
Фиг. 3А



Фиг. 3В

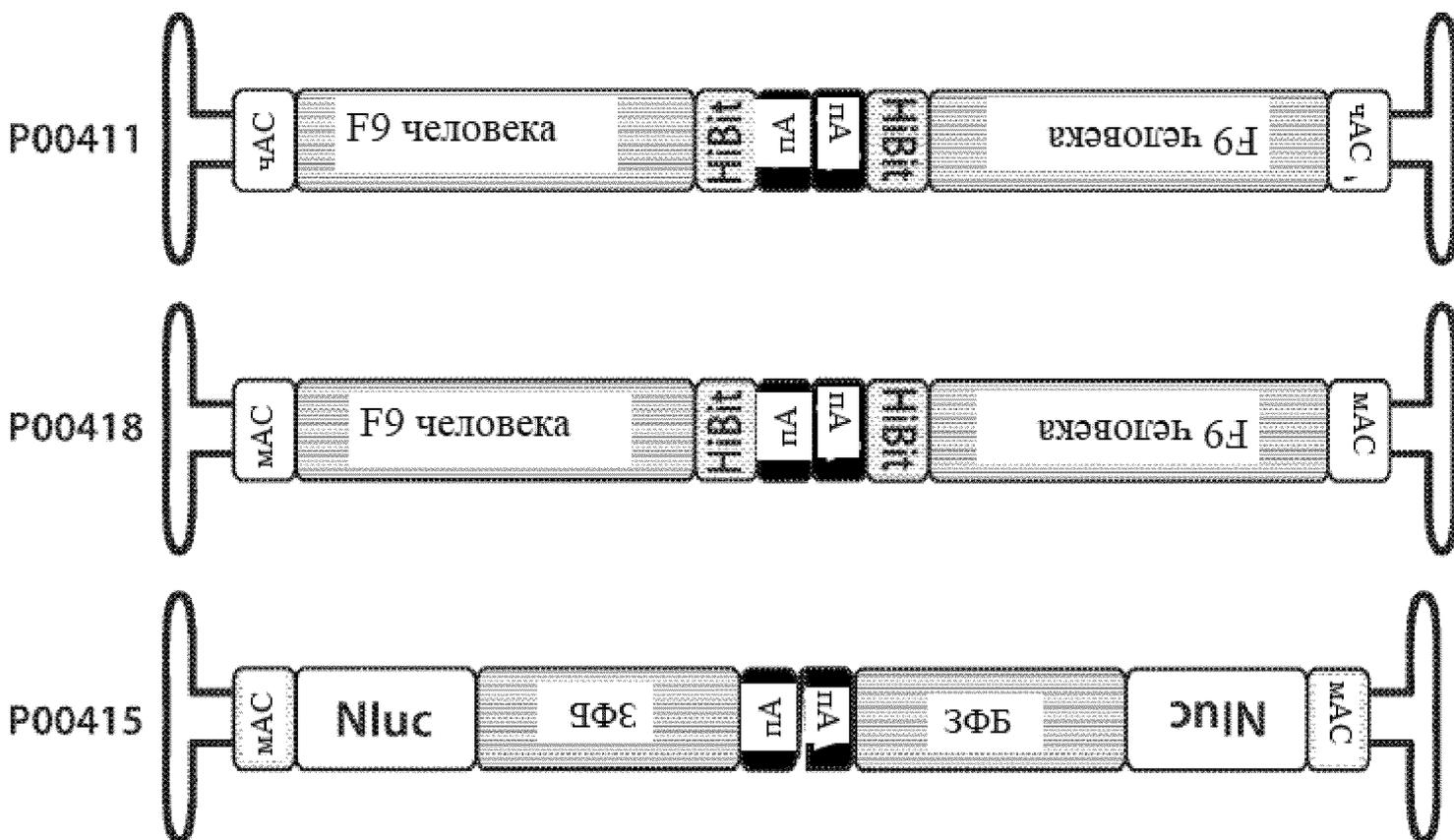


Фиг. 4А

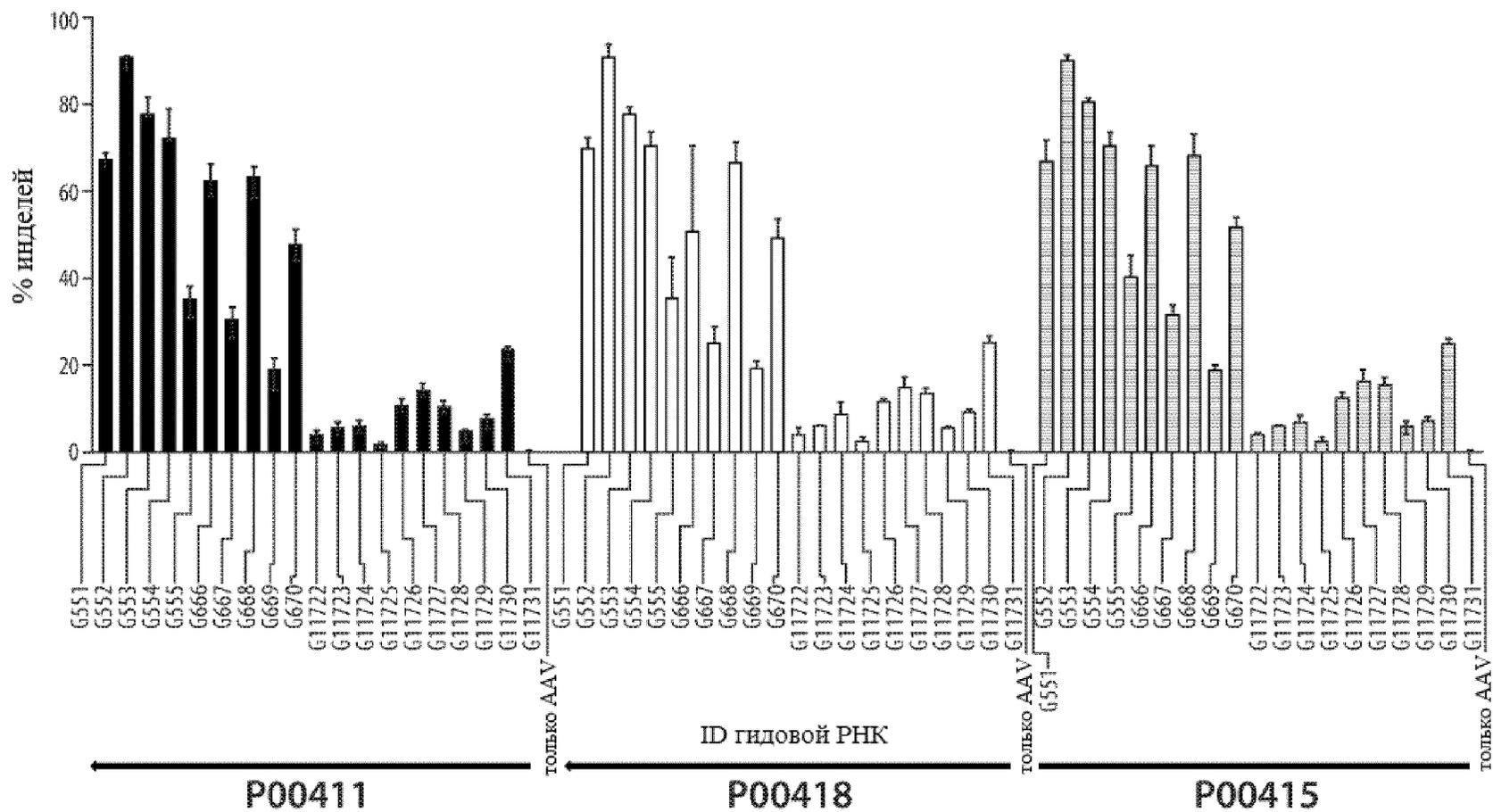


5/27

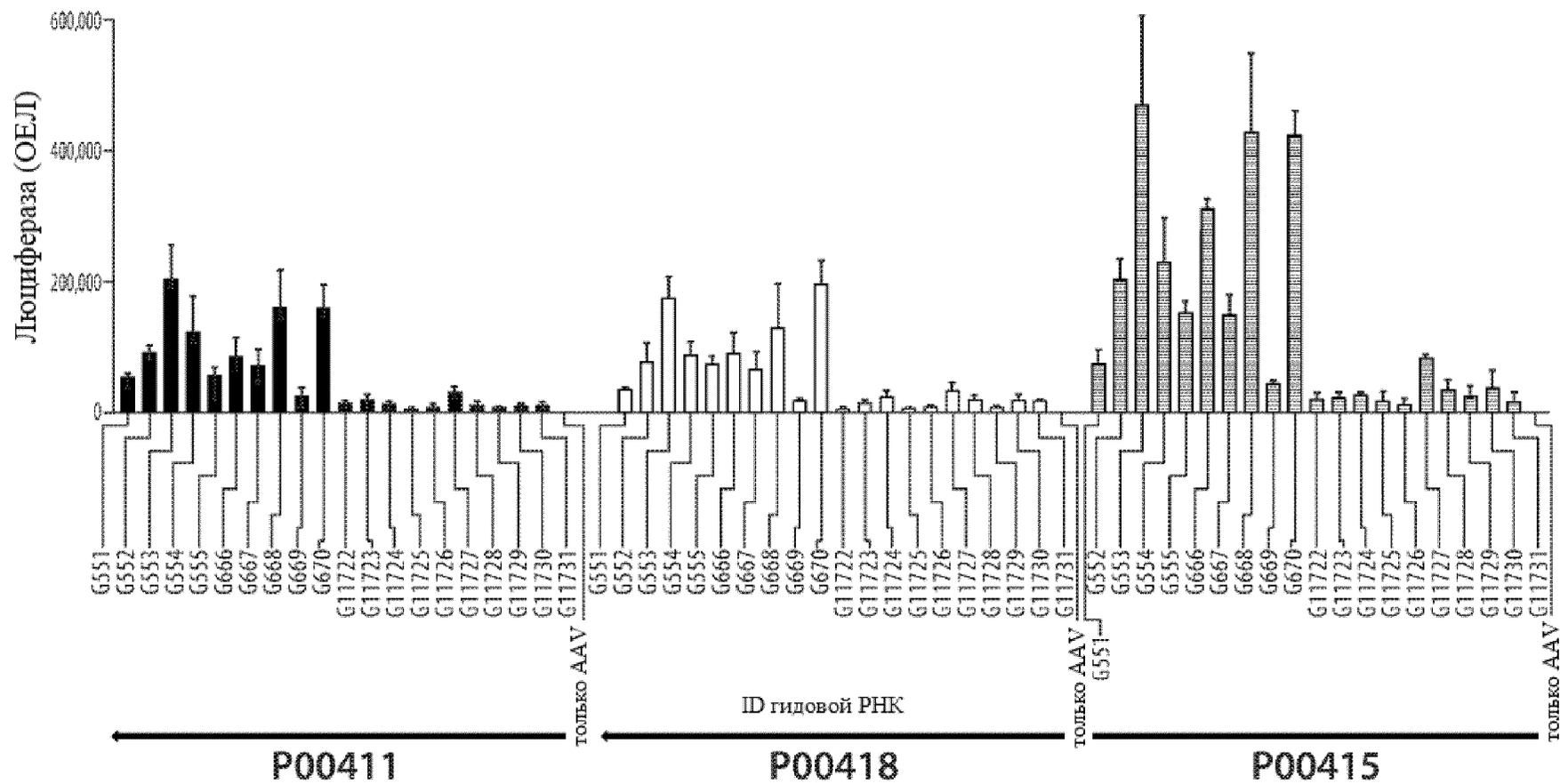
Фиг. 4В



Фиг. 5А

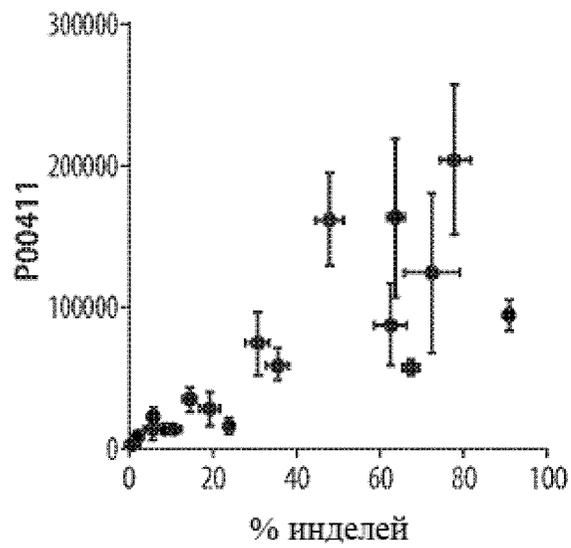


Фиг. 5В



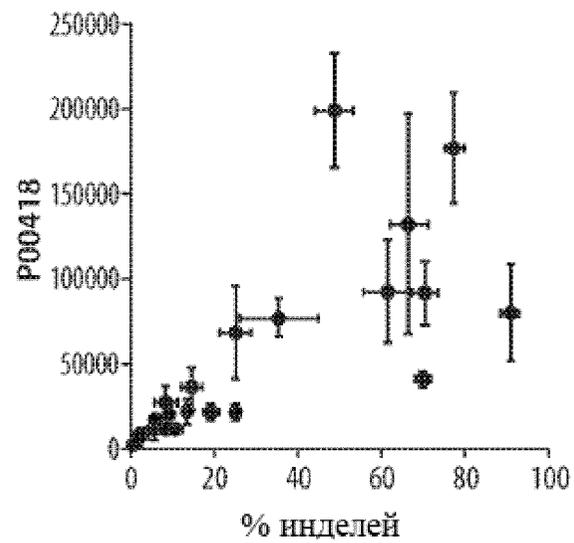
Фиг. 5С

Значение $R^2 = 0,54$



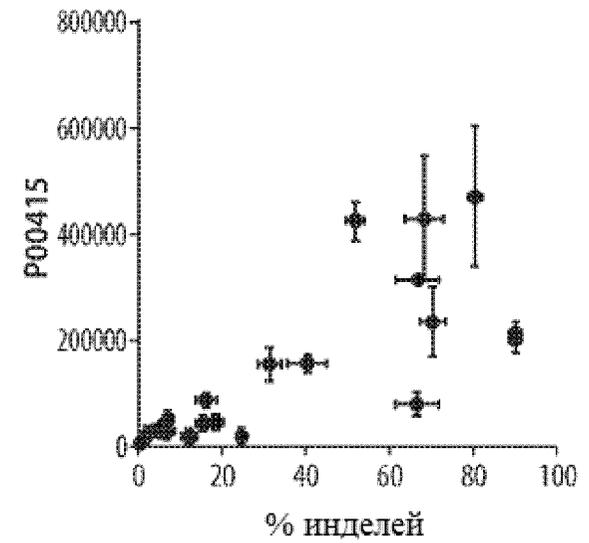
P00411

Значение $R^2 = 0,37$



P00418

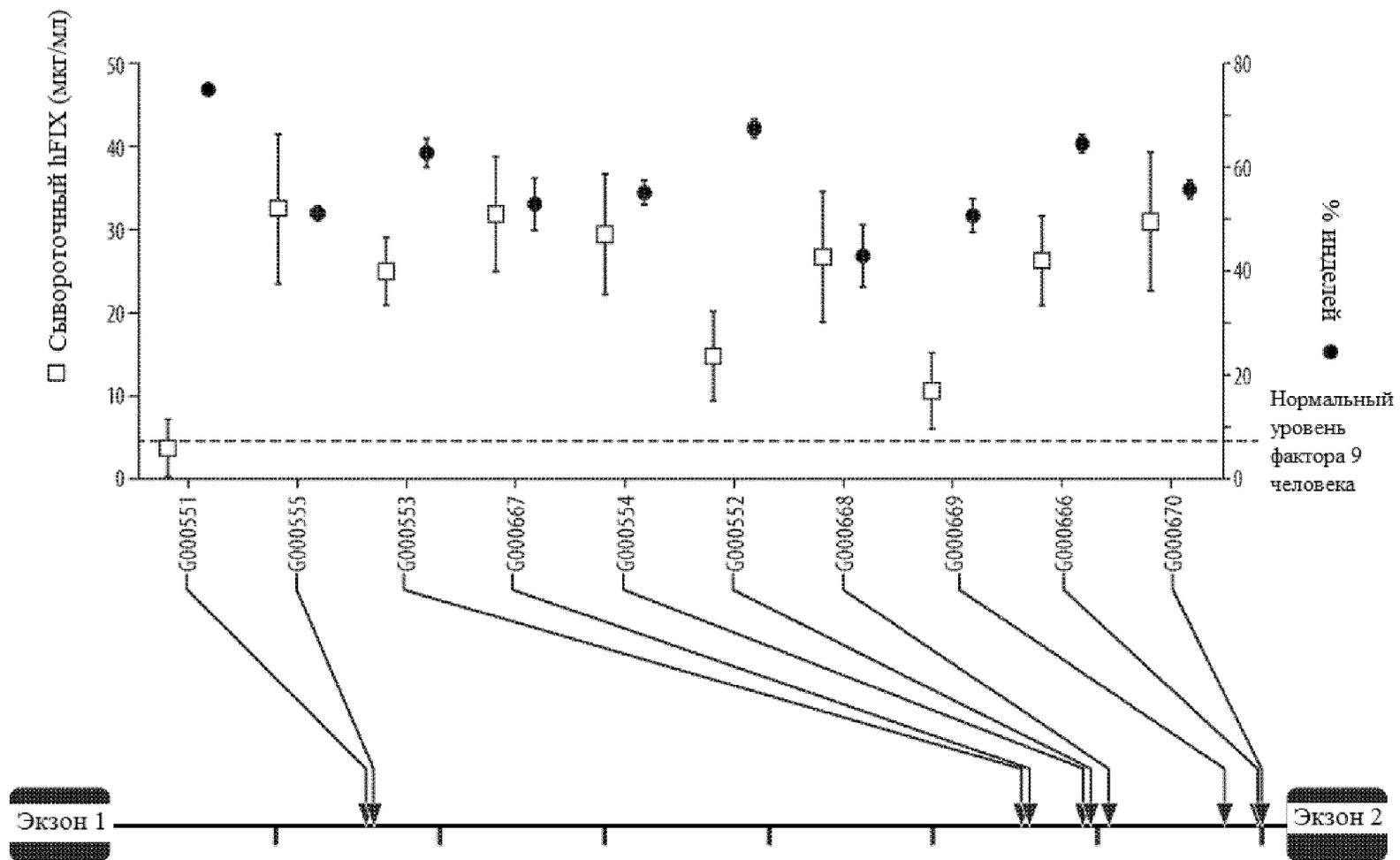
Значение $R^2 = 0,50$



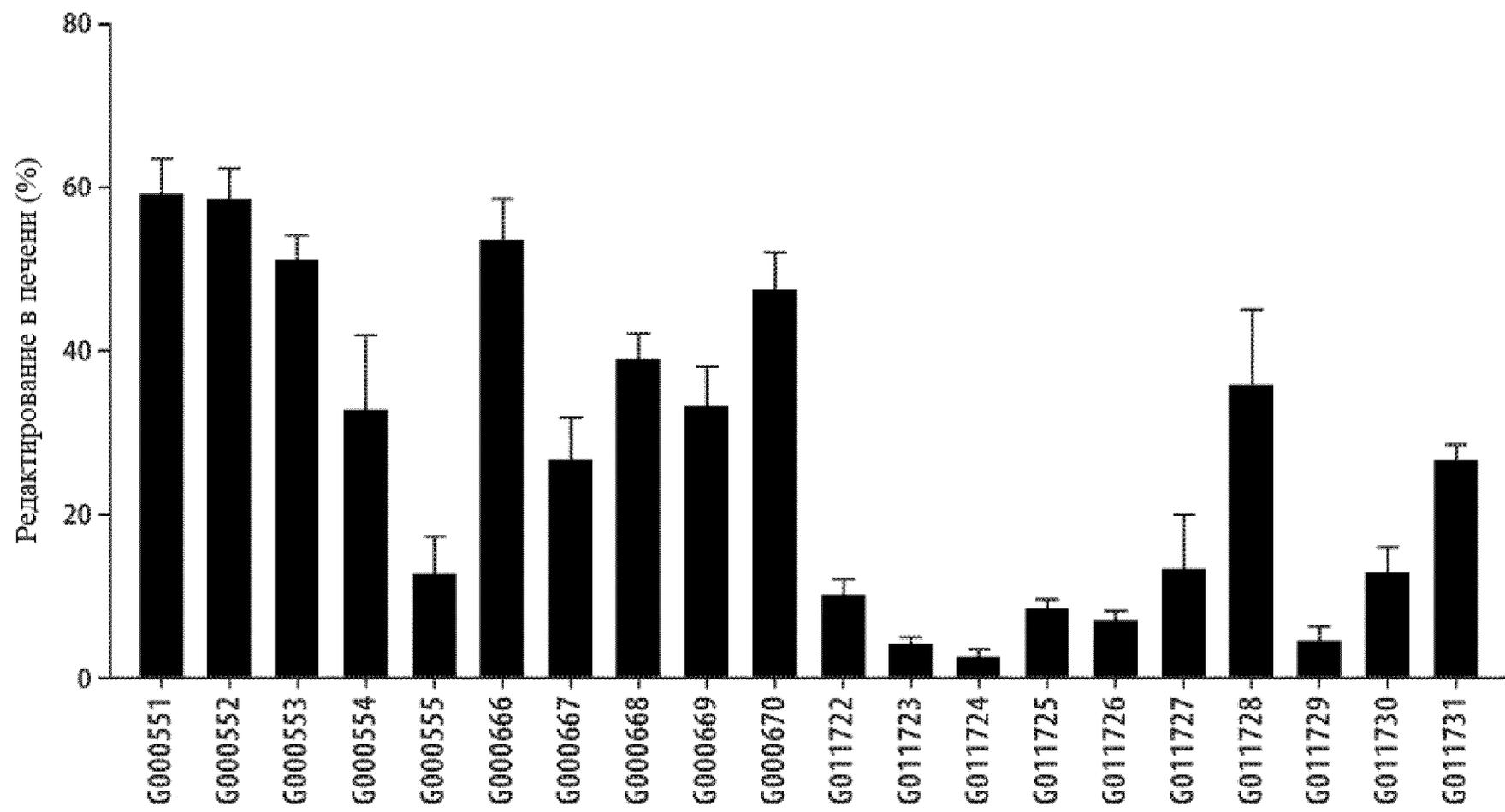
P00415

9/27

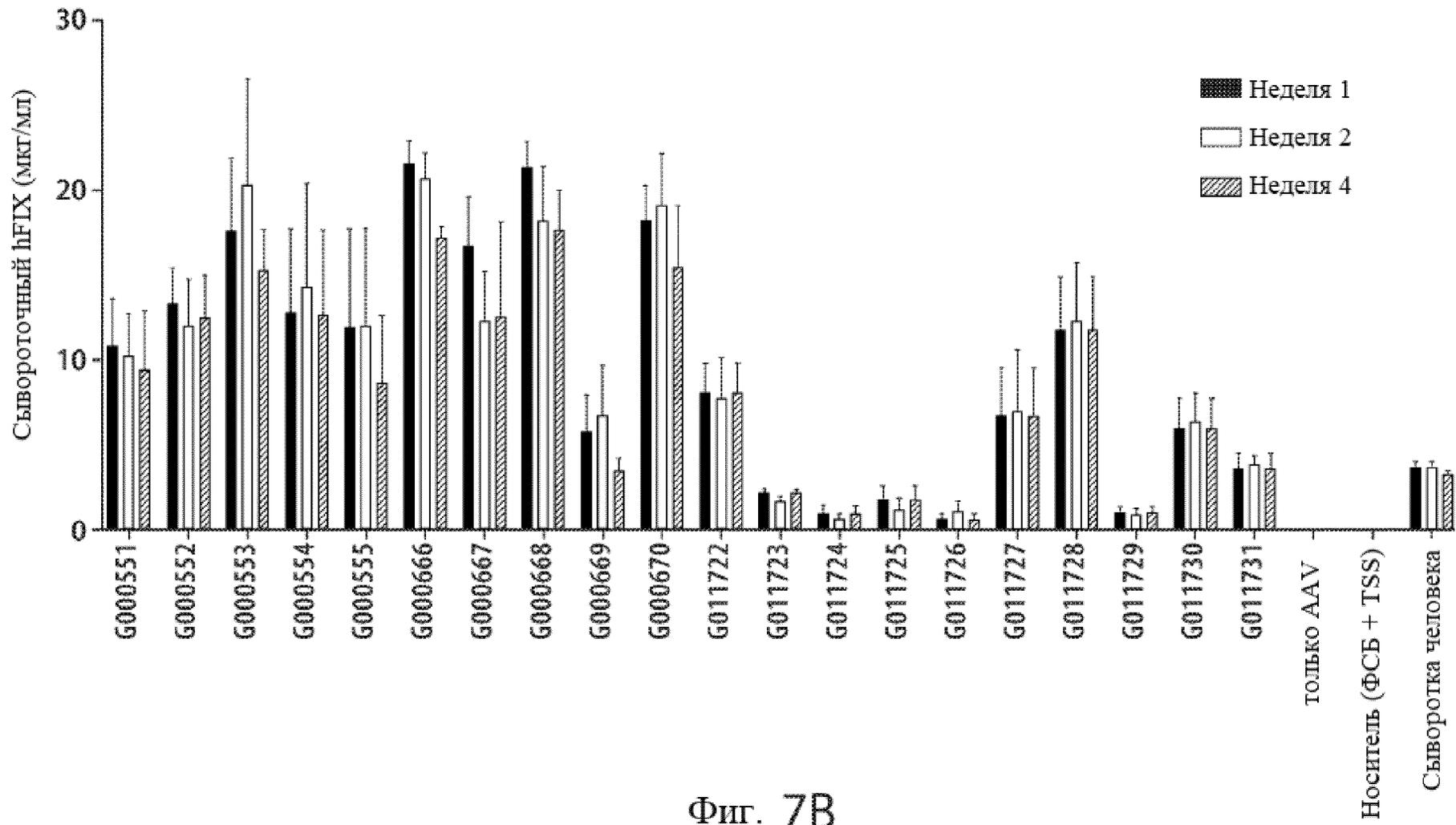
Фиг. 5D



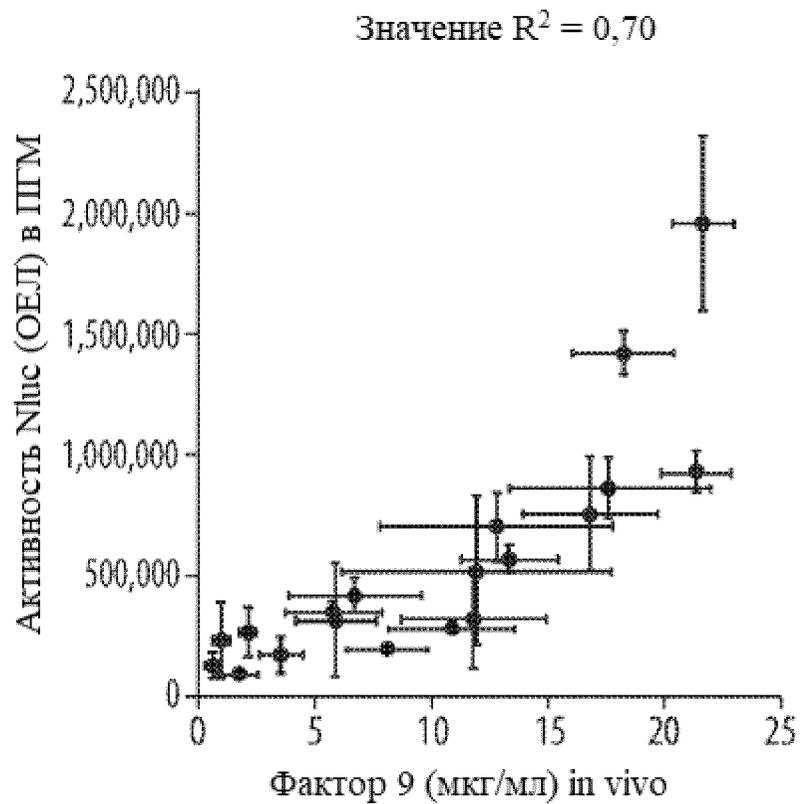
Фиг. 6



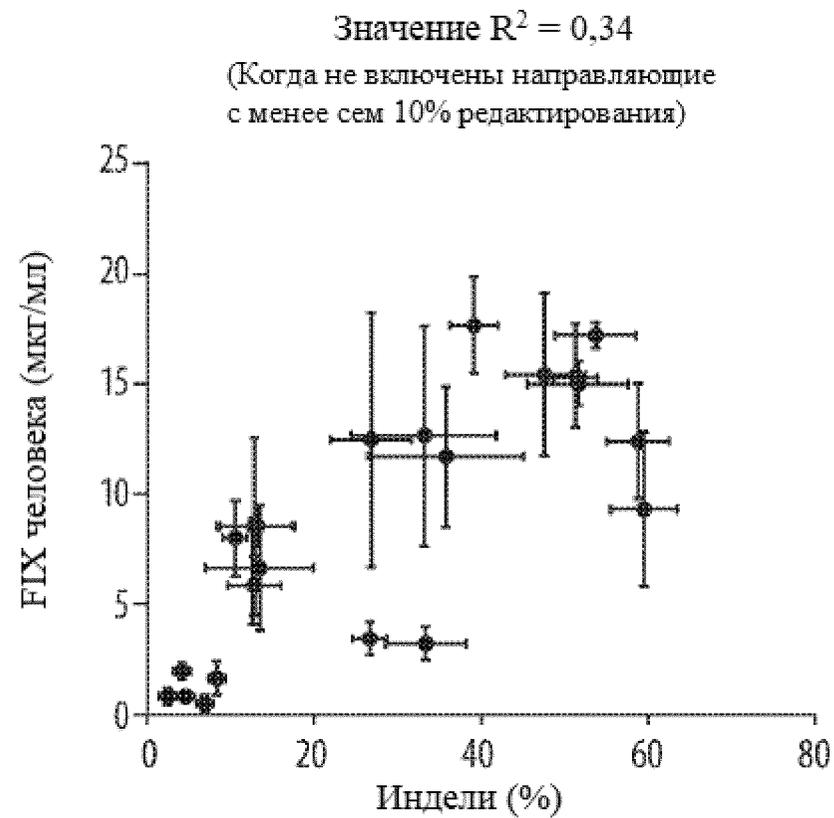
Фиг. 7А



Фиг. 7В

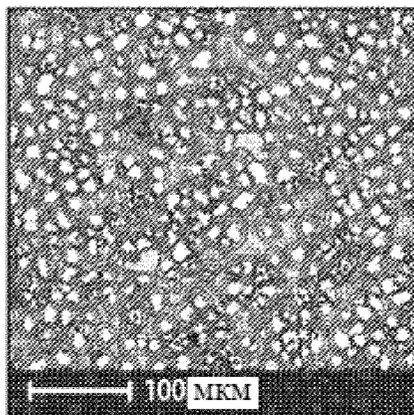


Фиг. 7C



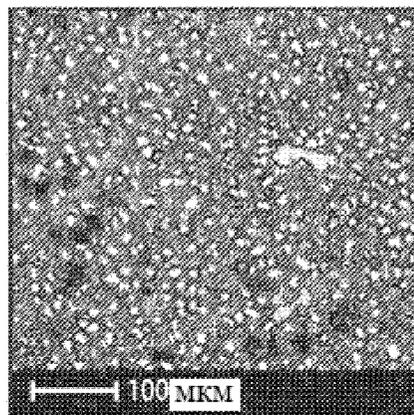
Фиг. 7D

только AAV



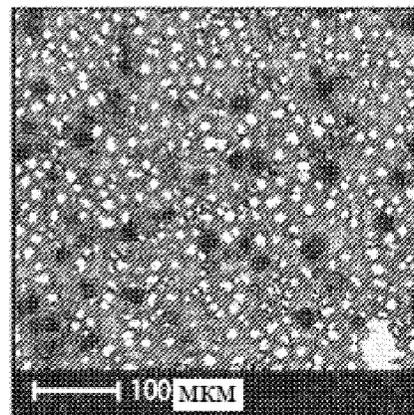
Менее 1,0% клеток, положительных в отношении гибридного транскрипта

G011723



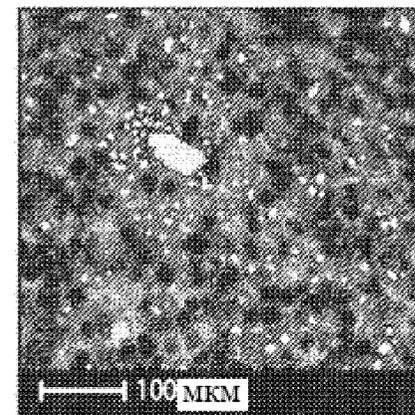
4,9% клеток, положительных в отношении гибридного транскрипта

G000551



19,8% клеток, положительных в отношении гибридного транскрипта

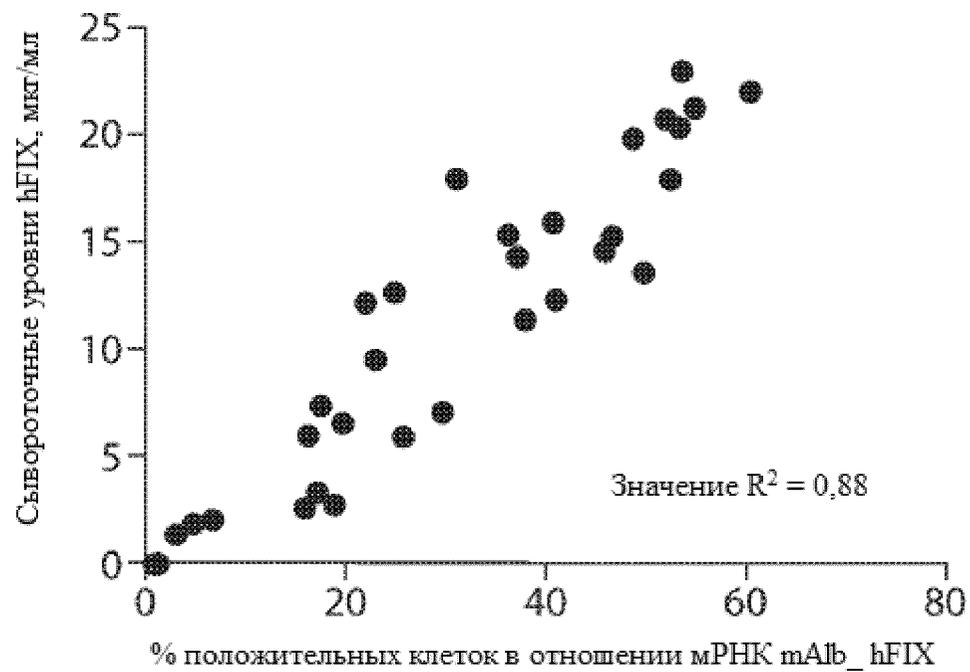
G000666



52,3% клеток, положительных в отношении гибридного транскрипта

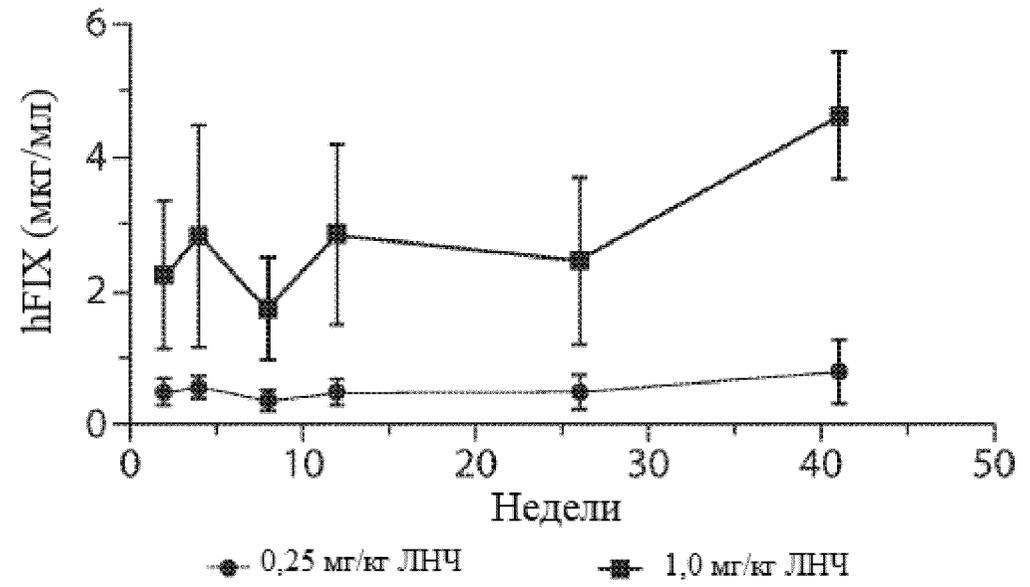
Фиг. 8А

Уровни hFIX, неделя 1 vs. % положительных клеток

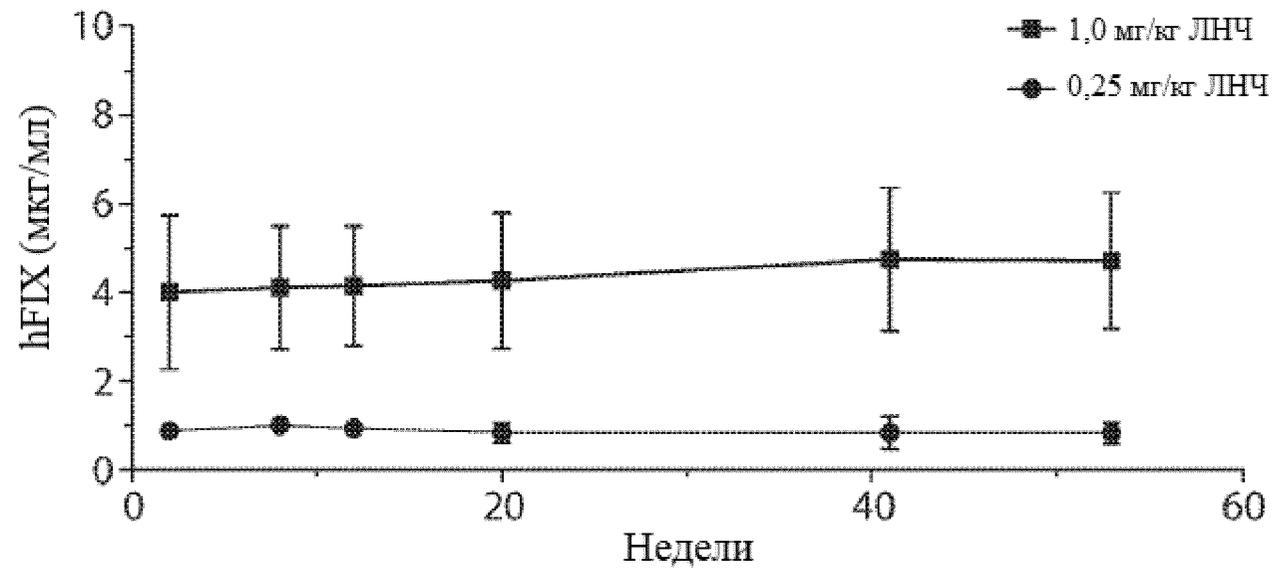


15/27

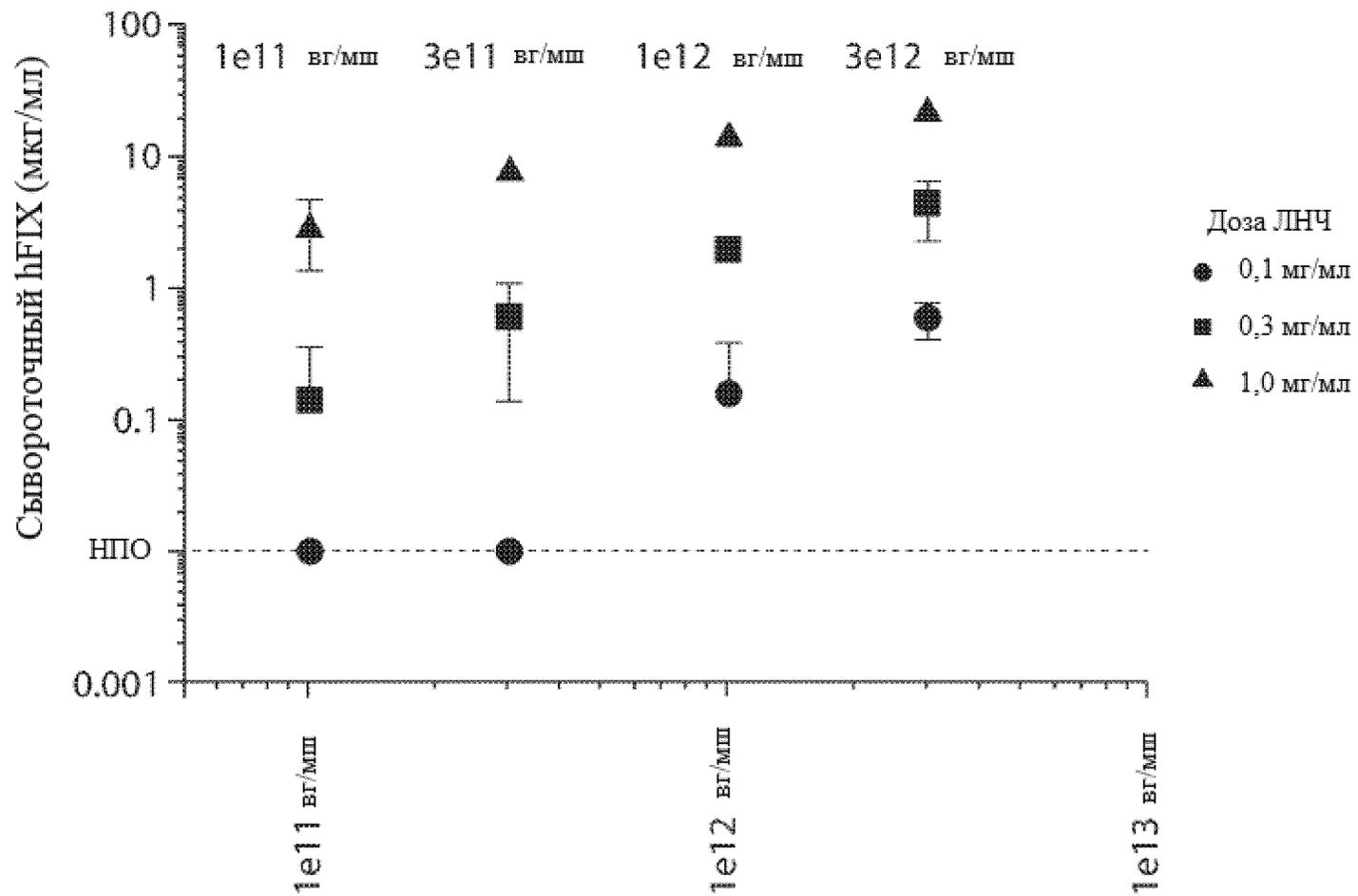
Фиг. 8В



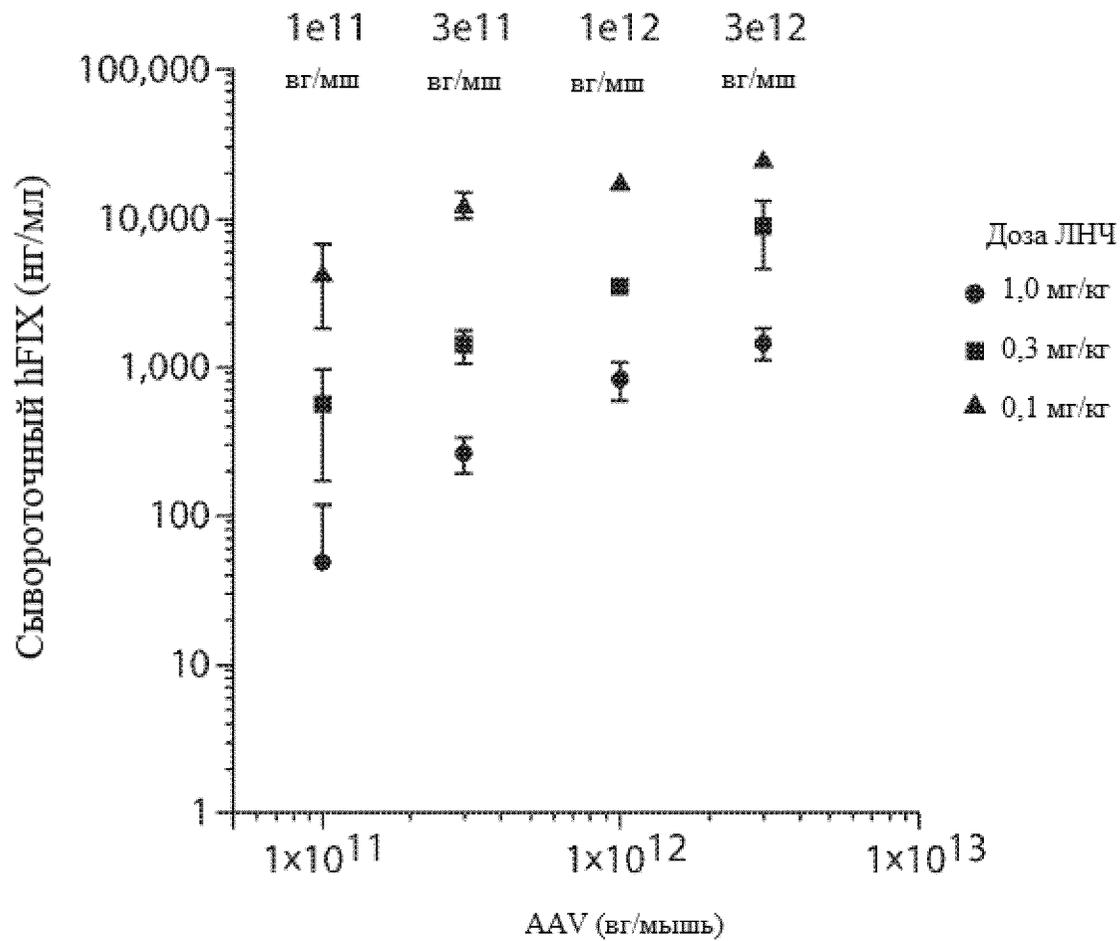
Фиг. 9А



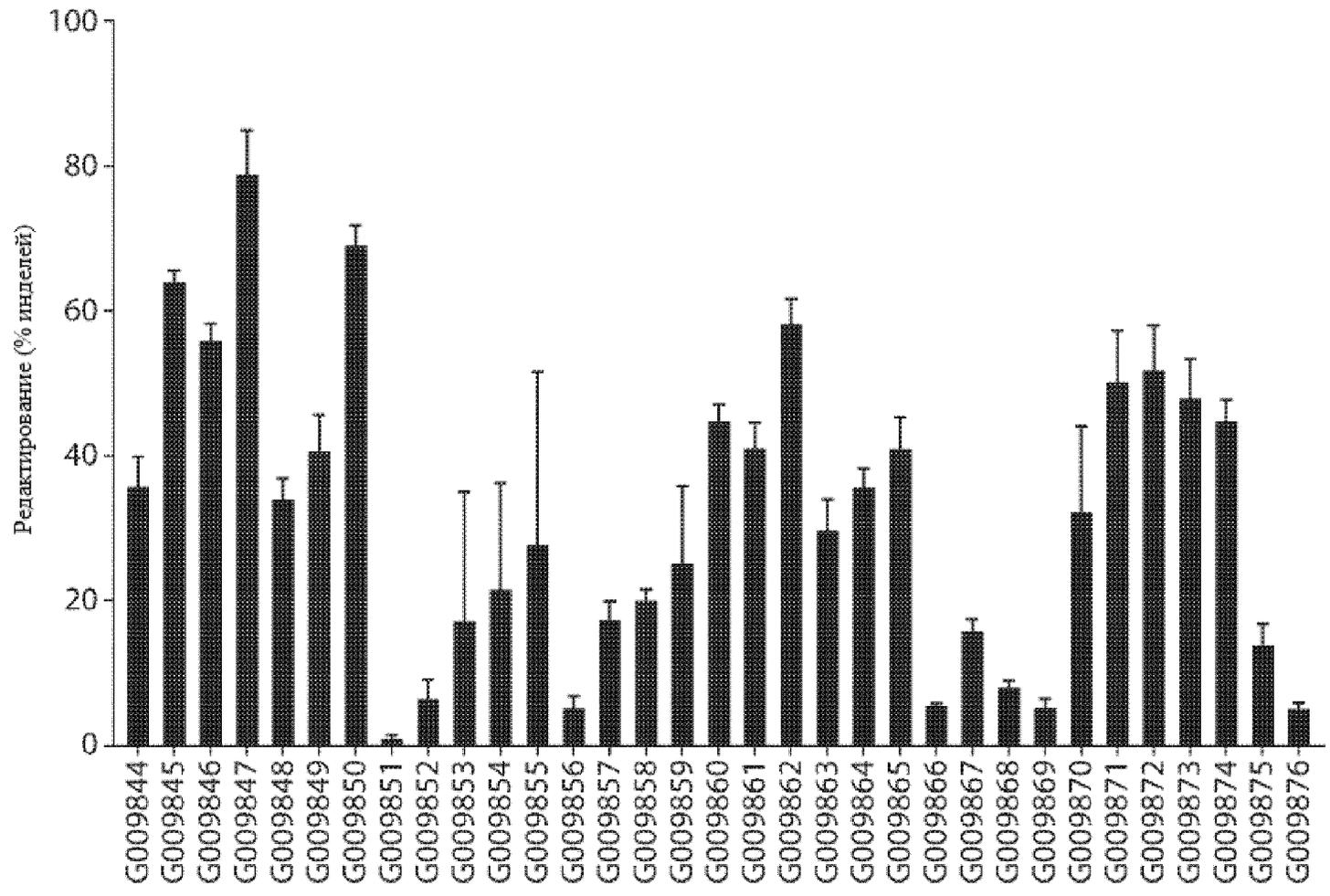
Фиг. 9В



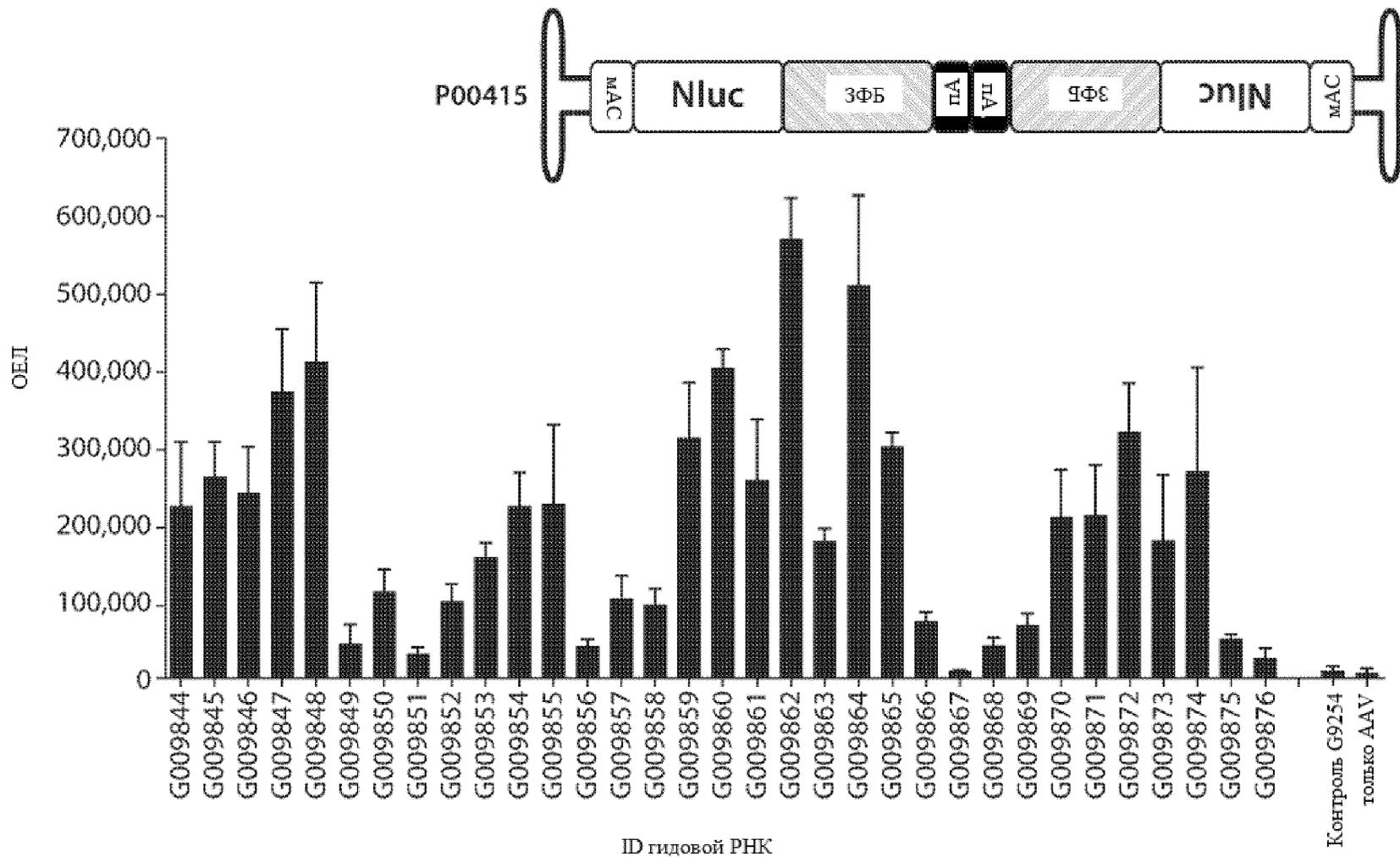
Фиг. 10А



Фиг. 10В

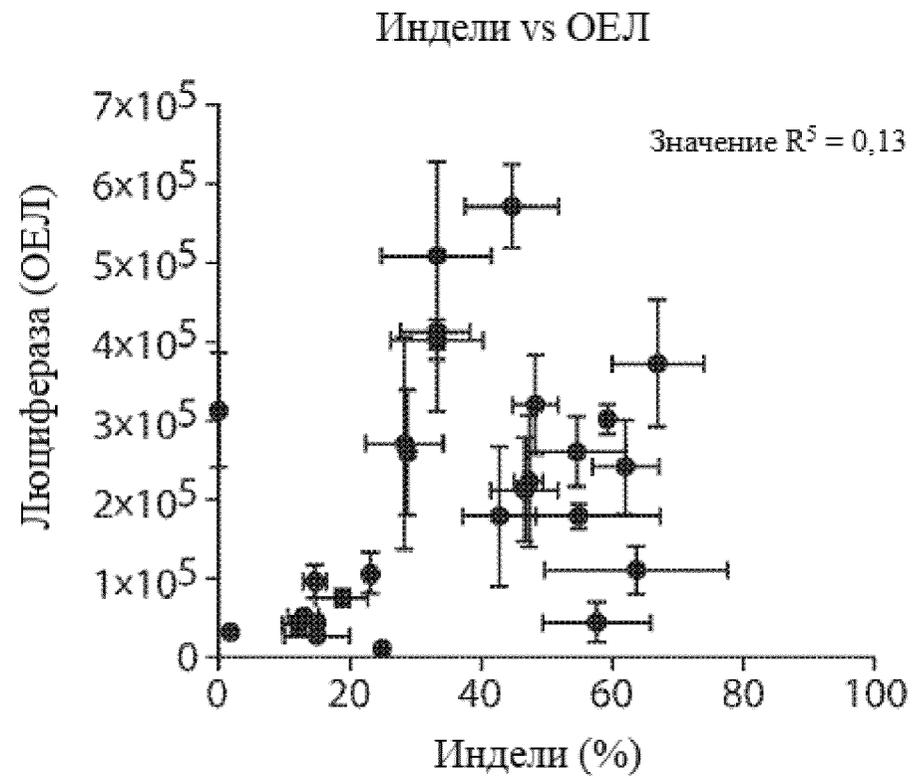


Фиг. 11А

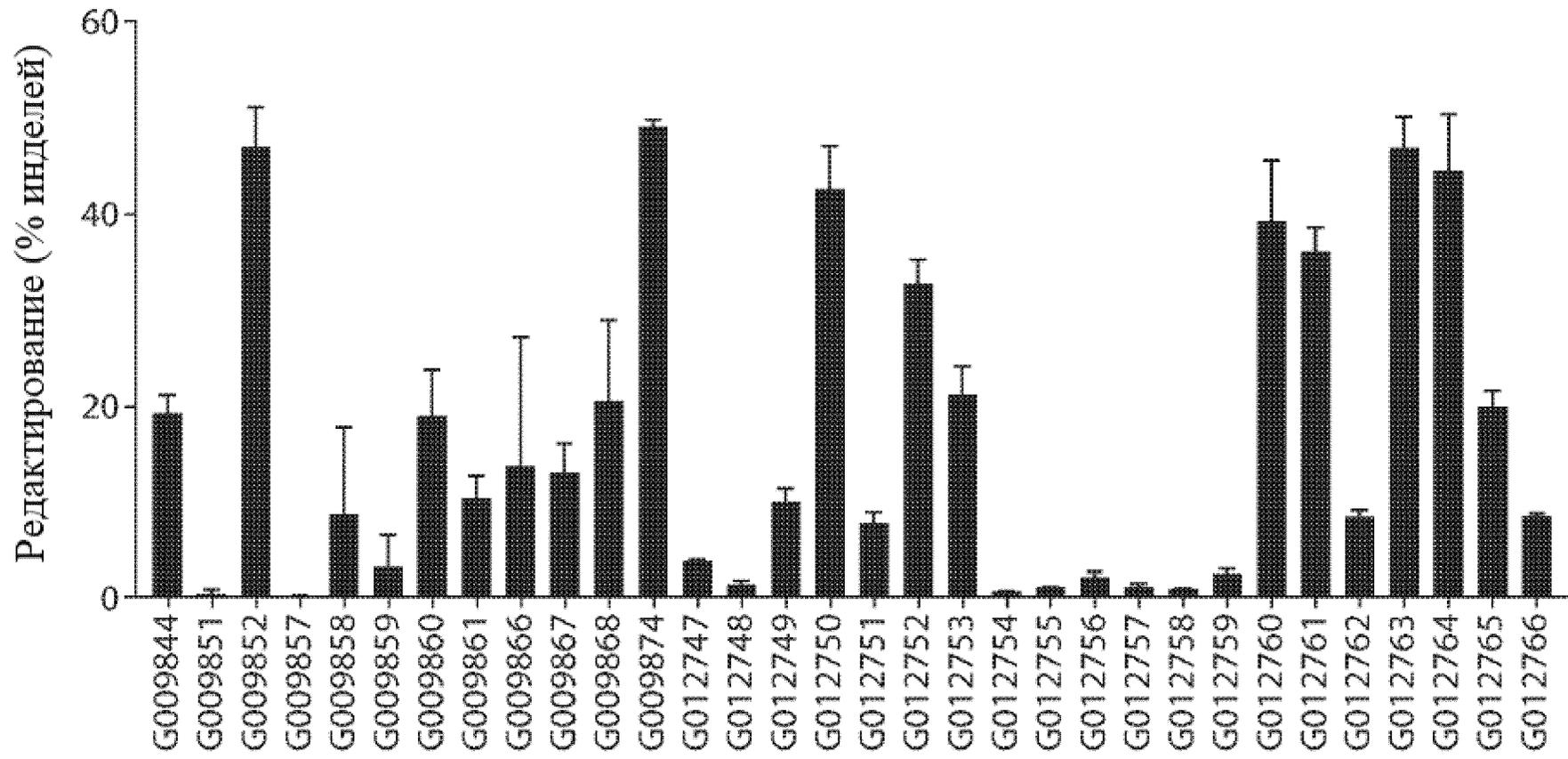


ID гидовой РНК

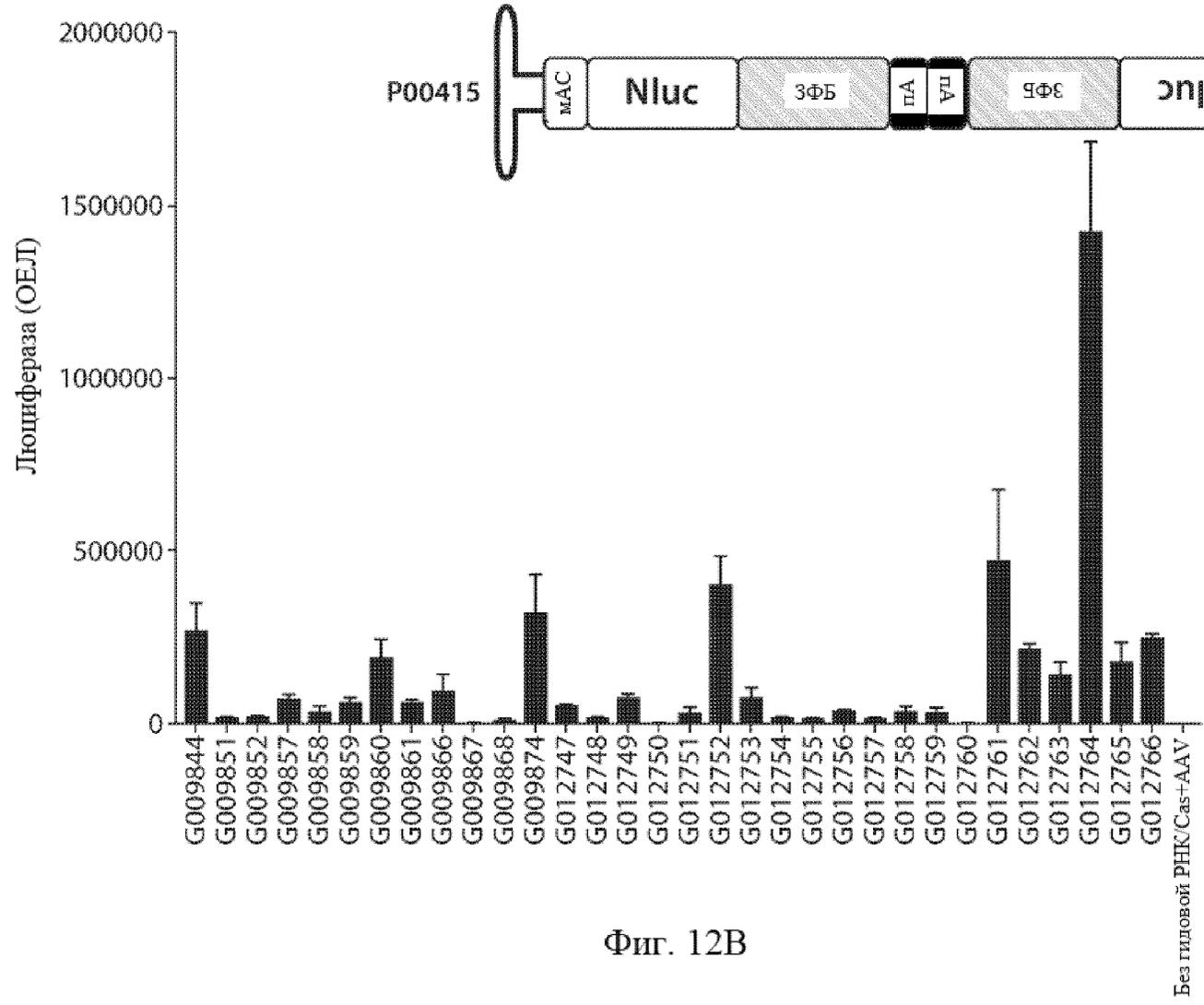
Фиг. 11В



Фиг. 11С

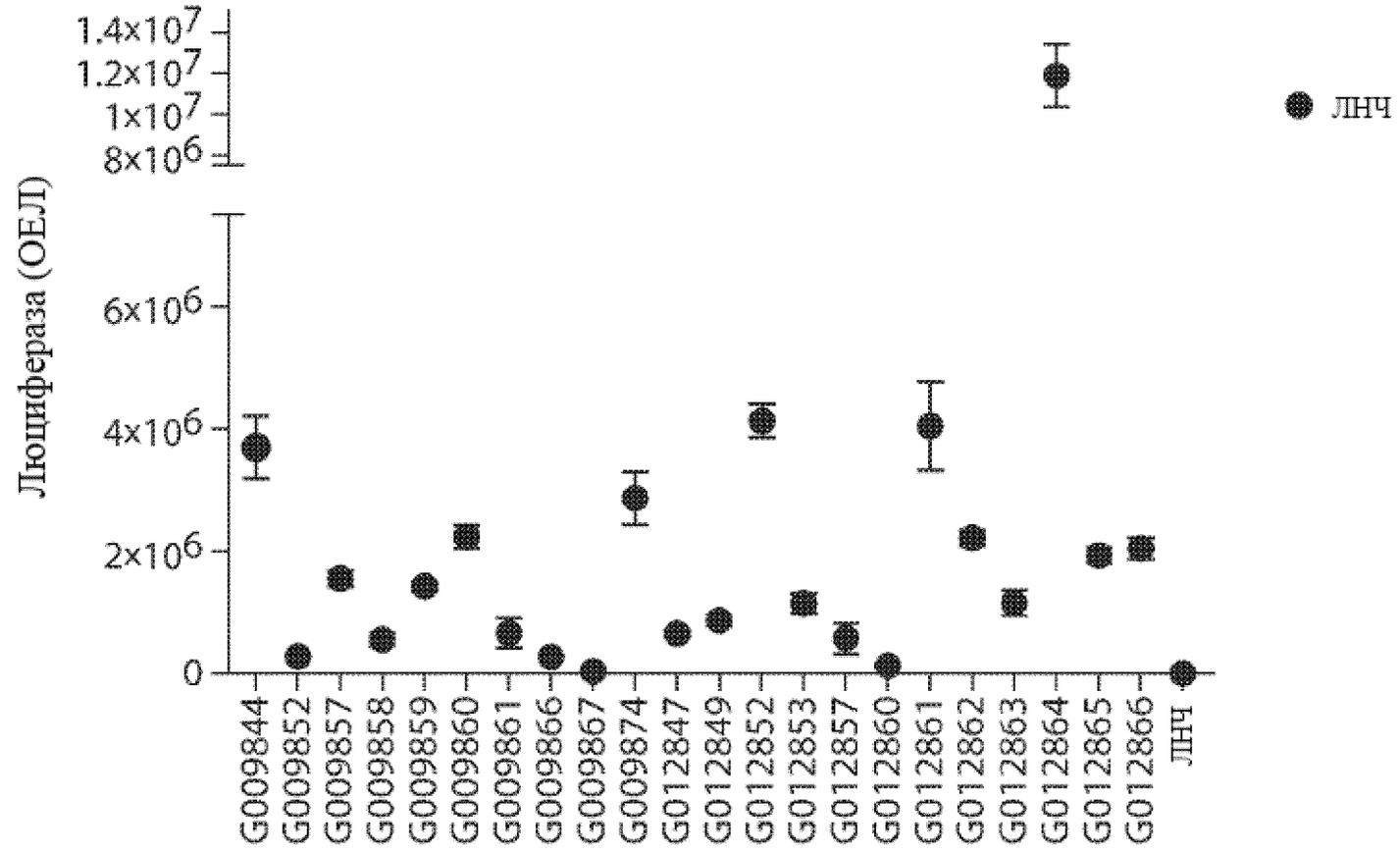


Фиг. 12А

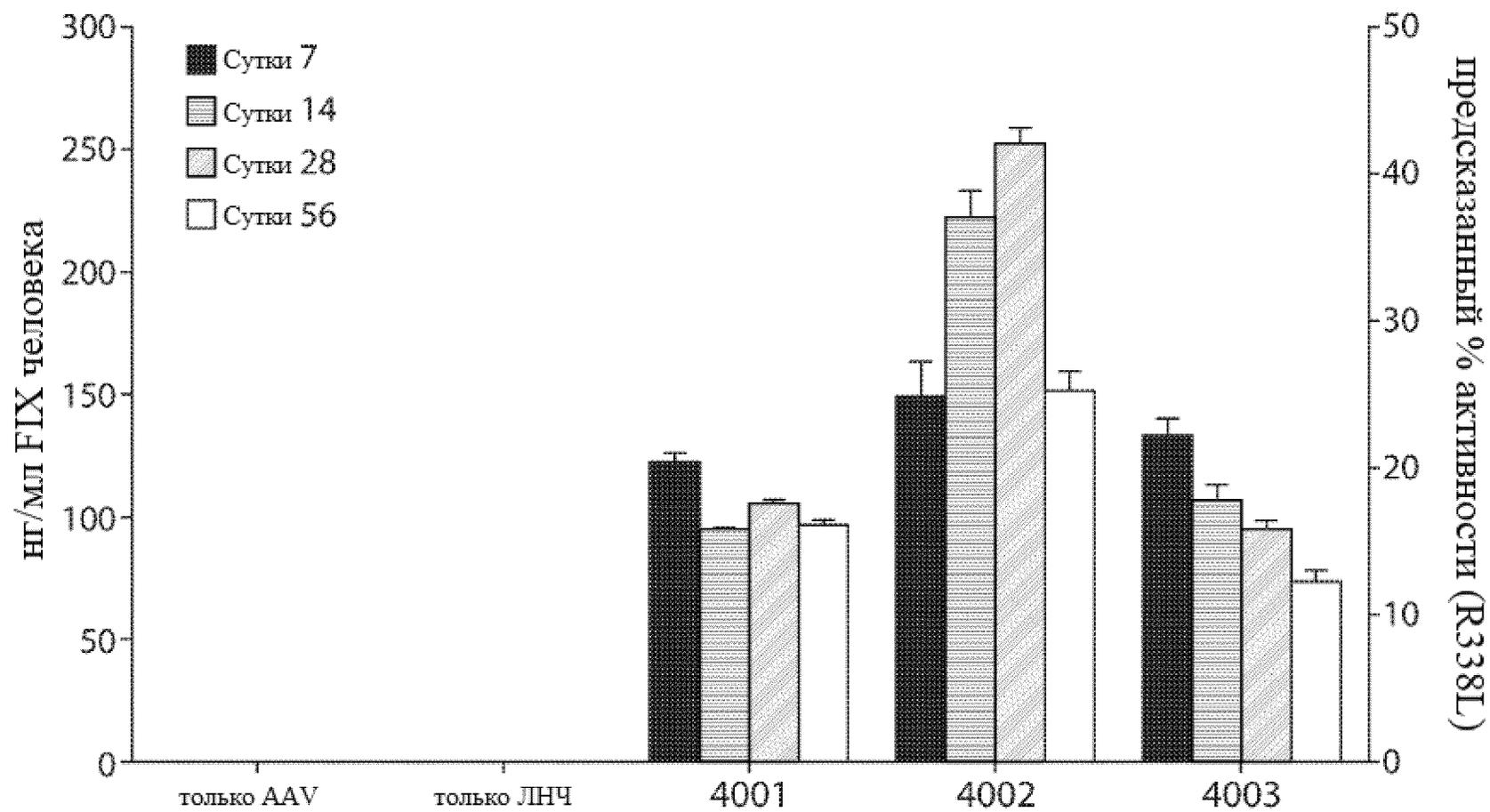


Фиг. 12В

Скрининг гРНК вставки альбумина человека



Фиг. 12С



Фиг. 13