

(19)



**Евразийское  
патентное  
ведомство**

(21) **202191068** (13) **A1**

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

(43) Дата публикации заявки  
2021.08.13

(51) Int. Cl. *C12N 15/113* (2010.01)  
*A61K 48/00* (2006.01)  
*C12N 15/90* (2006.01)

(22) Дата подачи заявки  
2019.10.18

---

(54) **КОМПОЗИЦИИ И СПОСОБЫ ДЛЯ ЭКСПРЕССИИ ТРАНСГЕНА ИЗ ЛОКУСА  
АЛЬБУМИНА**

---

(31) 62/747,402; 62/840,346

(32) 2018.10.18; 2019.04.29

(33) US

(86) PCT/US2019/057086

(87) WO 2020/082042 2020.04.23

(88) 2020.07.23

(71) Заявитель:

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

(72) Изобретатель:

**Финн Джонатан Дуглас, Хуан Хон-  
Рен, Рой Монтри, Лай КехДих,  
Саттлер Рэйчел, Киратсоус Кростос,  
Ван Чэн (US)**

(74) Представитель:

**Медведев В.Н. (RU)**

---

(57) Предложены способы редактирования, например внесения гетерологичного трансгена в ген альбумина человека (например, интрон 1).

**202191068**  
**A1**

**202191068**

**A1**

## ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-568574EA/050

### КОМПОЗИЦИИ И СПОСОБЫ ДЛЯ ЭКСПРЕССИИ ТРАНСГЕНА ИЗ ЛОКУСА АЛЬБУМИНА

Данная заявка испрашивает приоритет по предварительной заявке США № 62/747402, поданной 18 октября 2018 г., и предварительной заявке США № 62/840346, 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 может обеспечить потенциал для эффективного излечения пациентов. Однако клинические применения подходов генной терапии все еще нуждаются в усовершенствовании некоторых аспектов.

В данном документе предложены композиции и способы для вставки и экспрессии гетерологичного (экзогенного) гена в геномном локусе, таком как безопасный приемный сайт, клетки-хозяина. Были описаны несколько безопасных приемных сайтов, включая CCR5, HPRT, AAVS1, Rosa и альбумин. Как описано в данном документе, нацеливание и вставка экзогенного гена в локус альбумина (например, в интрон 1) позволяет использовать эндогенный промотор альбумина для управления устойчивой экспрессией экзогенного гена. Настоящее изобретение частично основано на идентификации гидовых РНК, которые специфически нацелены на сайты в гене альбумина, например, интроне 1 гена альбумина, и которые обеспечивают эффективную вставку и/или экспрессию экзогенного гена. Предложены следующие варианты осуществления.

В одном аспекте в настоящем изобретении предложен способ вставки нуклеиновой кислоты, кодирующей гетерологичный полипептид, в локус альбумина клетки-хозяина или популяции клеток, включающий введение: i) гРНК, которая содержит последовательность, выбранную из: а) последовательности, которая является по меньшей мере на 95%, 90%, 85%, 80% или 75% идентичной последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33; б) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33; в) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 и 97; д) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33; е) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности,

выбранной из группы, состоящей из SEQ ID NO: 2-33; f) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34-97; g) последовательности, комплементарной 15 последовательным нуклеотидам +/- 10 нуклеотидам геномных координат, приведенных для SEQ ID NO: 2-33; h) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 98-119; i) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 98-119; и j) последовательности, выбранной из группы, состоящей из SEQ ID NO: 120-163; ii) РНК-направляемого ДНК-связывающего агента; и iii) конструкции, содержащей нуклеиновую кислоту, кодирующую гетерологичный полипептид, тем самым осуществляя вставку нуклеиновой кислоты, кодирующей гетерологичный полипептид, в локус альбумина клетки-хозяина или популяции клеток.

В другом аспекте в настоящем изобретении предложен способ экспрессии гетерологичного полипептида из локуса альбумина клетки-хозяина или популяции клеток, включающий введение: i) гРНК, которая содержит последовательность, выбранную из: а) последовательности, которая является по меньшей мере на 95%, 90%, 85%, 80% или 75% идентичной последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33; b) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33; c) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 и 97; d) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33; e) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33; f) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34-97; и g) последовательности, которая содержит 15 последовательных нуклеотидов +/- 10 нуклеотидов геномных координат, приведенных для SEQ ID NO: 2-33; ii) РНК-направляемого ДНК-связывающего агента; и iii) конструкции, содержащей кодирующую последовательность гетерологичного полипептида, тем самым осуществляя экспрессию гетерологичного полипептида в клетке-хозяине или популяции клеток.

В другом аспекте в настоящем изобретении предложен способ экспрессии терапевтического агента в клетке неделящегося типа или популяции клеток, включающий введение: i) гРНК, которая содержит последовательность, выбранную из: а) последовательности, которая является по меньшей мере на 95%, 90%, 85%, 80% или 75% идентичной последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33; b) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33; c) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 и 97; d) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности,

выбранной из группы, состоящей из SEQ ID NO: 2-33; е) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33; ф) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34-97; и г) последовательности, которая содержит 15 последовательных нуклеотидов +/- 10 нуклеотидов геномных координат, приведенных для SEQ ID NO: 2-33; ii) РНК-направляемого ДНК-связывающего агента; и iii) конструкции, содержащей кодирующую последовательность гетерологичного полипептида, тем самым осуществляя экспрессию терапевтического агента в клетке неделящегося типа или популяции клеток.

В некоторых вариантах осуществления гРНК содержит направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32 и SEQ ID NO: 33.

В некоторых вариантах осуществления способ осуществляют *in vivo*. В некоторых вариантах осуществления способ осуществляют *in vitro*.

В некоторых вариантах осуществления гРНК связывается с областью, расположенной выше мотива, примыкающего к пропоспейсеру (PAM, англ. «*protospacer adjacent motif*»). В некоторых вариантах осуществления PAM выбран из NGG, NNGRRT, NNGRR(N), NNAGAAW, NNNNG(A/C)TT и NNNNRYAC.

В некоторых вариантах осуществления гРНК представляет собой двойную гРНК (дгРНК). В некоторых вариантах осуществления гРНК представляет собой одинарную гРНК (огРНК). В некоторых вариантах осуществления огРНК содержит один или более модифицированных нуклеозидов. В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cas-нуклеазу 2 класса. В некоторых вариантах осуществления Cas-нуклеаза выбрана из группы, состоящей из нуклеазы *S. pyogenes*, нуклеазы *S. aureus*, нуклеазы *C. jejuni*, нуклеазы *S. thermophilus*, нуклеазы *N. meningitidis* и их вариантов. В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cas9. В некоторых вариантах осуществления Cas-нуклеаза представляет собой Cas10.

В некоторых вариантах осуществления конструкция представляет собой двунаправленную конструкцию нуклеиновой кислоты. В некоторых вариантах осуществления конструкция содержит i. первый сегмент, содержащий кодирующую последовательность гетерологичного полипептида; и ii. второй сегмент, содержащий обратную комплементарную последовательность кодирующей последовательности гетерологичного полипептида. В некоторых вариантах осуществления конструкция содержит последовательность сигнала полиаденилирования. В некоторых вариантах осуществления конструкция содержит сайт акцептора сплайсинга. В некоторых вариантах осуществления конструкция не содержит плечо гомологии.

В некоторых вариантах осуществления гРНК вводят в векторе и/или липидной наночастице. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент вводят в векторе и/или липидной наночастице. В некоторых вариантах осуществления конструкцию, содержащую гетерологичный ген, вводят в векторе и/или липидной наночастице. В некоторых вариантах осуществления вектор представляет собой вирусный вектор. В некоторых вариантах осуществления вирусный вектор выбран из группы, состоящей из аденоассоциированного вирусного вектора (AAV), аденовирусного вектора, ретровирусного вектора и лентивирусного вектора. В некоторых вариантах осуществления 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.

В некоторых вариантах осуществления гРНК, РНК-направляемый ДНК-связывающий агент и конструкцию, содержащую кодирующую последовательность гетерологичного полипептида, по отдельности или в любой комбинации, вводят одновременно. В некоторых вариантах осуществления гРНК, РНК-направляемый ДНК-связывающий агент и конструкцию, содержащую кодирующую последовательность гетерологичного полипептида, вводят последовательно, в любом порядке и/или в любой комбинации. В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент или РНК-направляемый ДНК-связывающий агент и гРНК в комбинации вводят до введения конструкции. В некоторых вариантах осуществления конструкцию, содержащую кодирующую последовательность гетерологичного полипептида, вводят до гРНК и/или РНК-направляемого ДНК-связывающего агента.

В некоторых вариантах осуществления гетерологичный полипептид представляет собой секретлируемый полипептид. В некоторых вариантах осуществления гетерологичный полипептид представляет собой внутриклеточный полипептид.

В некоторых вариантах осуществления клетка представляет собой клетку печени. В некоторых вариантах осуществления клетка печени представляет собой гепатоцит.

В некоторых вариантах осуществления экспрессия гетерологичного полипептида в клетке-хозяине повышена по меньшей мере на около 2%, около 3%, около 4%, около 5%, около 6%, около 7%, около 8%, около 9%, около 10% или более относительно уровня в клетке до введения гРНК, РНК-направляемого ДНК-связывающего агента и конструкции, содержащей кодирующую последовательность гетерологичного полипептида.

В некоторых вариантах осуществления гРНК содержит SEQ ID NO: 301.

В некоторых вариантах осуществления гРНК опосредует мишень-специфическое разрезание РНК-направляемым ДНК-связывающим агентом, что приводит к вставке кодирующей последовательности гетерологичного полипептида в интрон 1 гена альбумина. В некоторых вариантах осуществления разрезание приводит к частоте вставки гетерологичной нуклеиновой кислоты в популяции клеток, составляющей по меньшей мере около 10%. В некоторых вариантах осуществления разрезание приводит к частоте

вставки кодирующей последовательности гетерологичного полипептида между около 30 и 35%, около 35 и 40%, около 40 и 45%, около 45 и 50%, около 50 и 55%, около 55 и 60%, около 60 и 65%, около 65 и 70%, около 70 и 75%, около 75 и 80%, около 80 и 85%, около 85 и 90%, около 90 и 95% или около 95 и 99%.

В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий белок представляет собой Cas9-нуклеазу *S. pyogenes*. В некоторых вариантах осуществления нуклеаза представляет собой кливазу или никазу.

В некоторых вариантах осуществления способ дополнительно включает введение ЛНЧ, содержащей гРНК. В некоторых вариантах осуществления способ дополнительно включает введение ЛНЧ, содержащей мРНК, которая кодирует РНК-направляемый ДНК-связывающий агент. В некоторых вариантах осуществления ЛНЧ содержит гРНК и мРНК, которая кодирует РНК-направляемый ДНК-связывающий агент. В некоторых вариантах осуществления гРНК и РНК-направляемый ДНК-связывающий белок вводят в виде РНП. В некоторых вариантах осуществления конструкцию вводят посредством вектора.

В одном аспекте в настоящем изобретении предложена клетка-хозяин, полученная любым одним или более из вышеописанных способов.

В одном аспекте в настоящем изобретении предложена клетка, содержащая двунаправленную конструкцию нуклеиновой кислоты, кодирующую гетерологичный полипептид, интегрированную с интроном 1 локуса альбумина клетки-хозяина. В некоторых вариантах осуществления клетка-хозяин представляет собой клетку печени. В некоторых вариантах осуществления клетка печени представляет собой гепатоцит.

#### КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

На Фиг. 1 проиллюстрированы форматы конструкций, представленных в геномах AAV. AC=акцептор сплайсинга; пА=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 экспрессии hFIX в делящихся клетках. 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 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 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 (т. е. образование инделей в конкретном целевом сайте) не обязательно приводят к более эффективной вставке или экспрессии трансгена.  $\text{cAC}$ =акцептор сплайсинга F9 человека;  $\text{mAC}$ =акцептор сплайсинга альбумина мыши;  $\text{HiBit=tag}$  для обнаружения на основе люциферазы;  $\text{pA}$ =сигнальная последовательность поли-А;  $\text{Nluc}$ =репортер нанолуциферазы;  $\text{ЗФБ}$ =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 образования инделей не обязательно приводят к высоким уровням экспрессии трансгена.

На Фиг. 7А-7D проиллюстрированы результаты *in vivo* скрининга двунаправленных конструкций среди 20 целевых сайтов с использованием ssAAV, полученного из P00147. На Фиг. 7А проиллюстрировано, что для каждой из групп обработки для каждой исследуемой комбинации ЛНЧ/вектор были обнаружены различные уровни редактирования по данным измерения образования инделей. На Фиг. 7В приведены соответствующие данные по направленной вставке. Результаты демонстрируют слабую корреляцию между образованием инделей и вставкой или экспрессией двунаправленных конструкций (Фиг. 7В и Фиг. 7D), а также положительную корреляцию между результатами *in vitro* и *in vivo* (Фиг. 7С).

На Фиг. 8А и 8В 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. 8А). Уровни циркулирующего hFIX коррелировали с числом клеток, положительных в отношении гибридного транскрипта (Фиг. 8В).

На Фиг. 9 проиллюстрировано влияние изменения времени между доставкой ssAAV, содержащего двунаправленную конструкцию hFIX, и ЛНЧ на направленную вставку.

На Фиг. 10 проиллюстрировано влияние на направленную вставку повторного

введения доз (например, 1, 2 или 3 доз) ЛНЧ после доставки двунаправленной конструкции hFIX.

На Фиг. 11А проиллюстрирована продолжительность экспрессии hFIX *in vivo*. На Фиг. 11В продемонстрировано, что экспрессия из интрона 1 альбумина была продолжительной.

На Фиг. 12А и 12В проиллюстрировано, что изменение дозы AAV или ЛНЧ может модулировать степень экспрессии hFIX из интрона 1 гена альбумина *in vivo*.

На Фиг. 13А-13С проиллюстрированы результаты скрининга двунаправленных конструкций среди целевых сайтов в первичных гепатоцитах яванских макаков. На Фиг. 13А проиллюстрированы различные уровни редактирования для каждого из образцов по данным измерения образования инделей. На Фиг. 13В и Фиг. 13С проиллюстрировано, что значительные уровни образования инделей не были прогностическими в отношении вставки или экспрессии двунаправленных конструкций в интрон 1 гена альбумина.

На Фиг. 14А-14D проиллюстрированы результаты скрининга двунаправленных конструкций среди целевых сайтов в первичных гепатоцитах человека. На Фиг. 14А проиллюстрировано редактирование для каждого из образцов по данным измерения образования инделей. Фиг. 14В, Фиг. 14С и 14D проиллюстрировано, что значительные уровни образования инделей не были прогностическими в отношении вставки или экспрессии двунаправленных конструкций в интрон 1 гена альбумина.

На Фиг. 15 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.

На Фиг. 16А и 16В проиллюстрированы уровни экспрессии человеческого фактора IX в образцах плазмы на 6 неделе после инъекции.

**На Фиг. 17 проиллюстрированы сывороточные уровни на 7 неделе и % положительных клеток среди множественных долей для каждого животного.**

#### 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.

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. Хотя представленные идеи описаны в сочетании с различными вариантами осуществления, следует понимать, что это не подразумевает, что представленные идеи

ограничены такими вариантами осуществления. Наоборот, подразумевается, что представленные идеи охватывают все альтернативные варианты, модификации и эквиваленты, которые могут быть предположены специалистами в данной области техники.

#### 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, N<sup>4</sup>-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, O<sup>6</sup>-метилгуанин, 4-тиопиримидины, 4-аминопиримидины, 4-диметилгидразинпиримидины и O<sup>4</sup>-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., 11<sup>th</sup> 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. *Spy 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 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), НураCas9 (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-кливаза, Cas-никаза, Cas9-кливаза 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](http://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 содержит открытую рамку считывания, которая может транслироваться в полипептид (т. е. может служить субстратом для трансляции рибосомой и аминоацилированными тРНК). мРНК будет преимущественно содержать РНК или модифицированную РНК и 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 «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 термин «гетерологичный ген» относится к гену, который был внесен в качестве экзогенного источника в сайт в геноме клетки-хозяина (например, 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 некоторых вариантах осуществления гетерологичный ген или трансген содержит экзогенную последовательность нуклеиновой кислоты, например, последовательность нуклеиновой кислоты, которая в природе не встречается в клетке-реципиенте. Например, гетерологичный ген может быть гетерологичным в отношении сайта его вставки и в отношении клетки-реципиента.

Гетерологичный ген можно вставлять в безопасный приемный локус в геноме 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 сверхэкспрессию экзогенного гена 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. В некоторых вариантах осуществления необходимый безопасный приемный локус может представлять собой локус, в котором экспрессия вставленной генной последовательности не нарушается сквозной экспрессией из соседних генов. В

некоторых вариантах осуществления безопасный приемный локус обеспечивает экспрессию 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 клеткой или популяцией клеток.

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 экзоном 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.

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.

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 трансген. 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.

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 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.

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, 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.

#### Композиции

#### Композиции, содержащие гидовую РНК (гРНК)

В данном документе предложены композиции и способы для вставки и экспрессии гетерологичного (экзогенного) гена в геномном локусе, таком как безопасный приемный сайт, клетки-хозяина. В частности, как проиллюстрировано в данном документе, нацеливание и вставка экзогенного гена в локус альбумина (например, в интрон 1) позволяет использовать эндогенный промотор альбумина для управления устойчивой экспрессией экзогенного гена. Настоящее изобретение частично основано на идентификации гидовых РНК, которые специфически нацелены на сайты в интроне 1 гена альбумина и которые обеспечивают эффективную вставку и экспрессию экзогенного гена. Как показано в примерах и дополнительно описано в данном документе, способность идентифицированных гРНК опосредовать высокие уровни редактирования, измеряемые по активности образования инделей, неожиданно не обязательно коррелирует с использованием тех же самых гРНК для опосредования эффективной вставки трансгенов, определяемой, *например*, по экспрессии трансгена. Таким образом, определенные гРНК, которые способны обеспечивать существенный уровень образования инделей, не обязательно способны опосредовать эффективную вставку, и, наоборот, некоторые гРНК, которые, как было показано, обеспечивают низкий уровень образования инделей, могут опосредовать эффективную вставку и экспрессию трансгена. В частности, данные из примеров показывают, что гРНК, которые эффективно опосредуют образование инделей (также называемое % редактирования), не обладают активностью редактирования инделей, которая коррелирует с активностью редактирования вставки.

В некоторых вариантах осуществления в данном документе предложены композиции и способы, применимые для вставки и экспрессии экзогенного гена в интрон 1 гена альбумина в клетке-хозяине. В некоторых вариантах осуществления в данном документе предложены композиции и способы, применимые для введения или вставки гетерологичной нуклеиновой кислоты в локус альбумина клетки-хозяина, например, с использованием описанных в данном документе гидовой РНК с РНК-направляемым ДНК-связывающим агентом и конструкции, содержащей гетерологичную нуклеиновую кислоту («трансген»). В некоторых вариантах осуществления в данном документе предложены композиции и способы, применимые для экспрессии гетерологичного полипептида в локусе альбумина клетки-хозяина, например, с использованием описанных в данном документе гидовой РНК с РНК-направляемым ДНК-связывающим агентом и конструкции, содержащей гетерологичную нуклеиновую кислоту («трансген»). В некоторых вариантах осуществления в данном документе предложены композиции и способы, применимые для индукции разрыва (например, двухцепочечного разрыва (ДЦР))

или одноцепочечного разрыва (ника)) в гене альбумина клетки-хозяина, например, с использованием описанной в данном документе гидовой РНК с РНК-направляемым ДНК-связывающим агентом (например, системой CRISPR/Cas). Композиции и способы можно применять *in vitro* или *in vivo*, например, в терапевтических целях.

В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая связывается или способна связываться с интроном локуса альбумина. В некоторых вариантах осуществления описанные в данном документе гидовые РНК связываются в области интрона 1 гена альбумина человека (SEQ ID NO: 1). Следует понимать, что не каждое основание направляющей последовательности должно связываться в указанных областях. Например, в некоторых вариантах осуществления 15, 16, 17, 18, 19, 20 или более оснований последовательности гидовой РНК связываются с указанными областями. Например, в некоторых вариантах осуществления 15, 16, 17, 18, 19, 20 или более смежных оснований последовательности гидовой РНК связываются с указанными областями.

В некоторых вариантах осуществления описанные в данном документе гидовые РНК опосредуют целенаправленное разрезание РНК-направляемым ДНК-связывающим агентом (например, Cas-нуклеазой) в сайте внутри интрона 1 альбумина человека (SEQ ID NO: 1). Следует понимать, что в некоторых вариантах осуществления гидовые РНК содержат направляющие последовательности, которые связываются с указанными областями или способны связываться с областью в SEQ ID NO: 1.

В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 164-196. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 98-119. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из таблицы 1. гРНК может содержать одну или более направляющих последовательностей, приведенных в таблице 1. гРНК может содержать одну или более последовательностей, приведенных в таблицах 1, 7 и 9. гРНК может содержать одну или более последовательностей, приведенных в таблицах 2, 8 и 10. Гидовая РНК может содержать одну или более SEQ ID NO: 2-33. гРНК может содержать одну или более SEQ ID NO: 164-196. гРНК может содержать одну или более SEQ ID NO: 98-119.

В некоторых вариантах осуществления описанные в данном документе гидовые

РНК содержат направляющую последовательность, содержащую по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, содержащую по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 164-196. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, содержащую по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 98-119. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, содержащую по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из таблицы 1.

В некоторых вариантах осуществления гидовая РНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32 и SEQ ID NO: 33.

В некоторых вариантах осуществления гидовая РНК (гРНК) альбумина содержит последовательность, выбранную из а) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33; б) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33; в) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 и 97; д) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33; е) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33; ф) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34-97; и г) последовательности, комплементарной 15 последовательным нуклеотидам +/- 10 нуклеотидам геномных координат, приведенных для SEQ ID NO: 2-33. В некоторых вариантах осуществления гидовая РНК альбумина содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления гидовая РНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 4, 13, 17, 19, 27, 28, 30 и 31.

В некоторых вариантах осуществления описанные в данном документе гидовые

РНК связываются с областью, расположенной выше мотива, примыкающего к пропоспейсеру (PAM, англ. «prospacer 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 PAM представляет собой NGG.

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

В некоторых вариантах осуществления последовательность гидовой РНК содержит последовательность, комплементарную последовательности в геномной области, выбранной из таблицы 1, 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 таблицы 1. 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 таблицы 1.

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).

В некоторых вариантах осуществления описанные в данном документе гидовые РНК опосредуют целенаправленное разрезание РНК-направляемым ДНК-связывающим агентом (например, Cas-нуклеазой), что приводит к вставке гетерологичной нуклеиновой кислоты в интрон 1 гена альбумина. В некоторых вариантах осуществления гидовая РНК и/или разрезание в сайте разрезания приводят к частоте вставки гетерологичного гена между 30 и 35%, 35 и 40%, 40 и 45%, 45 и 50%, 50 и 55%, 55 и 60%, 60 и 65%, 65 и 70%, 70 и 75%, 75 и 80%, 80 и 85%, 85 и 90%, 90 и 95% или 95 и 99%. В некоторых вариантах осуществления гидовая РНК и/или разрезание приводит к вставке гетерологичной нуклеиновой кислоты, составляющей по меньшей мере 20%, по меньшей мере 30%, по меньшей мере 40%, по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 70%, по меньшей мере 80%, по меньшей мере 90%. Частоту вставки можно определить *in vitro* или *in vivo*. Например, в некоторых вариантах осуществления частоту вставки можно определить путем обнаружения и измерения вставленной нуклеиновой кислоты в

популяции клеток и расчета процентной доли популяции, которая содержит вставленную нуклеиновую кислоту. Методы определения уровня вставки известны и доступны в данной области техники. В некоторых вариантах осуществления гидовая РНК обеспечивает повышение экспрессии гетерологичного гена на значение между 5 и 10%, 10 и 15%, 15 и 20%, 20 и 25%, 25 и 30%, 30 и 35%, 35 и 40%, 40 и 45%, 45 и 50%, 50 и 55%, 55 и 60%, 60 и 65%, 65 и 70%, 70 и 75%, 75 и 80%, 80 и 85%, 85 и 90%, 90 и 95%, 95 и 99% или более. Повышение экспрессии гетерологичного гена можно измерить *in vitro* или *in vivo*. Например, в некоторых вариантах осуществления повышение экспрессии можно определить путем обнаружения и измерения уровня гетерологичного полипептида и сравнения этого уровня с уровнем полипептида до, например, обработки клеток или введения субъекту. В некоторых вариантах осуществления повышение экспрессии можно определить путем обнаружения и измерения уровня гетерологичного полипептида и сравнения этого уровня с известным уровнем полипептида, нормальным уровнем полипептида у здорового субъекта.

Каждая из направляющих последовательностей, приведенных в таблице 1 может дополнительно содержать дополнительные нуклеотиды с образованием сгРНК, например, со следующей типовой нуклеотидной последовательностью после направляющей последовательности в 3'-конце: GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 300) в ориентации от 5' к 3'. Геномные координаты соответствуют референсному геному человека hg38. В случае оРНК, вышеприведенные направляющие последовательности могут дополнительно содержать дополнительные нуклеотиды с образованием оРНК, например, со следующей типовой нуклеотидной последовательностью, следующей за 3' концом направляющей последовательности: GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGA AAAAGUGGCACCGAGUCGGUGCUUUU (SEQ ID NO: 301) или GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGA AAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 302) в ориентации от 5' к 3'.

**Таблица 1: Последовательности гидовой РНК человека и хромосомные координаты**

<b>ID направляющей</b>	<b>Направляющая последовательность</b>	<b>Геномные координаты</b>	<b>SEQ ID NO:</b>
G009844	GAGCAACCUCACUCUUGUCU	xp4:73405113-73405133	2
G009851	AUGCAUUUGUUUCAAAAUAU	xp4:73405000-73405020	3
G009852	UGCAUUUGUUUCAAAAUAUU	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	UAAUAAAAUUCAAACAUCU	xp4:73404561-73404581	13
G012747	GCAUCUUUAAAAGAAUUUUU	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	UUCAUUUAGUCUGUCUUCU	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	ACCUUUUUUUUUUUUUACCU	xp4:73404581-73404601	27
G012761	AGUGCAAUGGAUAGGUCUUU	xp4:73404714-73404734	28
G012762	UGAUUCCUACAGAAAAACUC	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

Гидовая РНК может дополнительно содержать тРНК. В каждом описанном в данном документе варианте осуществления композиции и способа сРНК и тРНК могут быть связаны в виде одной РНК (огРНК) или могут находиться в отдельных РНК (дгРНК). В контексте огРНК компоненты сРНК и тРНК 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 сРНК, содержащую, например, направляющую последовательность,



приведенную в таблице 1, 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 crРНК и trРНК.

В каждом из описанных в данном документе вариантов осуществления композиции, применения и способа гидовая РНК может содержать одну молекулу РНК в виде «одинарной гидовой РНК» или «огРНК». ogРНК может содержать crРНК (или ее часть), содержащую направляющую последовательность, приведенную в таблице 1, ковалентно связанную с trРНК. ogРНК может содержать 15, 16, 17, 18, 19 или 20 смежных нуклеотидов направляющей последовательности, приведенной в таблице 1. В некоторых вариантах осуществления crРНК и trРНК ковалентно связаны посредством линкера. В некоторых вариантах осуществления ogРНК образует структуру типа «стебель - петля» посредством спаривания оснований между частями crРНК и trРНК. В некоторых вариантах осуществления crРНК и trРНК ковалентно связаны посредством одной или более связей, которые не являются фосфодиэфирными связями.

В некоторых вариантах осуществления trРНК может содержать всю или часть последовательности trРНК, полученной из системы CRISPR/Cas природного происхождения. В некоторых вариантах осуществления trРНК содержит усеченную или модифицированную trРНК дикого типа. Длина trРНК зависит от используемой системы CRISPR/Cas. В некоторых вариантах осуществления trРНК содержит или состоит из 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 или более 100 нуклеотидов. В некоторых вариантах осуществления trРНК может содержать определенные вторичные структуры, такие как, например, одна или более структур типа шпильки или стебель - петля, или одна или более структур с выпетливанием.

В некоторых вариантах осуществления целевая последовательность или область в интроне 1 локуса альбумина человека (например, нуклеотидные последовательности, соответствующие области в SEQ ID NO: 1) может быть комплементарной с направляющей последовательностью гидовой РНК. В некоторых вариантах осуществления степень комплементарности или идентичности между направляющей последовательностью гидовой РНК и соответствующей ей целевой последовательностью может составлять по меньшей мере 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% или 100%. В некоторых вариантах осуществления целевая последовательность и направляющая последовательность гРНК могут быть на 100% комплементарными или идентичными. В других вариантах осуществления целевая последовательность и направляющая последовательность гРНК могут содержать по меньшей мере одно несовпадение. Например, целевая последовательность и направляющая последовательность гРНК могут содержать 1, 2, 3, 4 или 5 несовпадений, при этом общая длина гидовой последовательности составляет около 20 или 20. В некоторых вариантах осуществления целевая последовательность и направляющая последовательность гРНК могут содержать 1-4 несовпадения, при этом длина гидовой последовательности составляет около 20 или 20 нуклеотидов.

Как описано и проиллюстрировано в данном документе, гидовые РНК альбумина можно использовать для вставки и экспрессии гетерологичного гена (например, трансгена) в интроне 1 гена альбумина. Таким образом, в некоторых вариантах осуществления настоящее изобретение включает композиции, содержащие одну или более гидовых РНК (гРНК), содержащих направляющие последовательности, которые направляют РНК-направляемый ДНК-связывающий агент (например, Cas9) к целевой последовательности ДНК в гене альбумина.

В некоторых вариантах осуществления описанные в данном документе композиции или состав содержат мРНК, содержащую открытую рамку считывания (ОРС), кодирующую РНК-направляемый ДНК-связывающий агент, такой как Cas-нуклеаза, как описано в данном документе. 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 предложена, применяется или вводится мРНК, содержащая ОРС, кодирующую РНК-направляемый ДНК-связывающий агент, такой как Cas-нуклеаза.

#### Модифицированные гРНК и мРНК

В некоторых вариантах осуществления гРНК химически модифицирована. гРНК, содержащая один или более модифицированных нуклеозидов или нуклеотидов, называется «модифицированной» гРНК или «химически модифицированной» гРНК для описания наличия одного или более неприродных и/или природных компонентов или

конфигураций, используемых вместо или в дополнение к каноническим остаткам А, G, С и U. В некоторых вариантах осуществления модифицированную гРНК синтезируют с неканоническим нуклеозидом или нуклеотидом, что в данном документе называется «модифицированным». Модифицированные нуклеозиды и нуклеотиды могут включать одно или более из: (i) изменения, например, замены одного или обоих несвязывающих фосфатных атомов кислорода и/или одного или более связывающих фосфатных атомов кислорода в связи фосфодиэфирного остова (типичная модификация остова); (ii) изменения, например, замены составляющего компонента сахара рибозы, например 2'-гидроксила в сахаре рибозы (типичная модификация сахара); (iii) полной замены фосфатного фрагмента «дефосфо»-линкерами (типичная модификация остова); (iv) модификации или замены нуклеос основания природного происхождения, в том числе неканоническим нуклеос основанием (типичная модификация основания); (v) замены или модификации рибозо-фосфатного остова (типичная модификация остова); (vi) модификации 3' конца или 5' конца олигонуклеотида, например, удаление, модификацию или замену концевой фосфатной группы или конъюгацию фрагмента, кэпа или линкера (такие модификации 3' или 5' кэпа могут включать модификацию сахара и/или остова); и (vii) модификации или замены сахара (типичная модификация сахара).

Химические модификации, такие как перечисленные выше, можно комбинировать, чтобы получить модифицированные гРНК и/или мРНК, содержащих нуклеозиды и нуклеотиды (вместе «остатки»), которые могут иметь две, три, четыре или более модификаций. Например, модифицированный остаток может содержать модифицированный сахар и модифицированное нуклеос основание. В некоторых вариантах осуществления каждое основание гРНК модифицировано, например, все основания имеют модифицированную фосфатную группу, такую как тиофосфатная группа. В некоторых вариантах осуществления все или практически все фосфатные группы молекулы гРНК заменены тиофосфатными группами. В некоторых вариантах осуществления модифицированные гРНК содержат по меньшей мере один модифицированный остаток в 5' конце РНК или вблизи него. В некоторых вариантах осуществления модифицированные гРНК содержат по меньшей мере один модифицированный остаток в 3' конце РНК или вблизи него. Определенные гРНК содержат по меньшей мере один модифицированный остаток в 5' конце и 3' конце РНК или вблизи них.

В некоторых вариантах осуществления гРНК содержит один, два, три или более модифицированных остатков. В некоторых вариантах осуществления по меньшей мере 5% (например, по меньшей мере 5%, по меньшей мере 10%, по меньшей мере 15%, по меньшей мере 20%, по меньшей мере 25%, по меньшей мере 30%, по меньшей мере 35%, по меньшей мере 40%, по меньшей мере 45%, по меньшей мере 50%, по меньшей мере 55%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95% или 100%) позиций в модифицированной гРНК представлены модифицированными нуклеозидами или нуклеотидами.

Немодифицированные нуклеиновые кислоты могут быть подвержены деградации, например, внутриклеточными нуклеазами или сывороточными нуклеазами. Например, нуклеазы могут гидролизовать фосфодиэфирные связи нуклеиновых кислот. Соответственно, в одном аспекте описанные в данном документе гРНК могут содержать один или более модифицированных нуклеозидов или нуклеотидов, например, для придания стабильности по отношению к внутриклеточным нуклеазам или сывороточным нуклеазам. В некоторых вариантах осуществления описанные в данном документе модифицированные молекулы гРНК могут демонстрировать сниженный врожденный иммунный ответ при внесении в популяцию клеток как *in vivo*, *in vitro* и *ex vivo*. Термин «врожденный иммунный ответ» включает клеточный ответ на экзогенные нуклеиновые кислоты, включая одноцепочечные нуклеиновые кислоты, который включает индукцию экспрессии и высвобождения цитокинов, в частности интерферонов, и гибель клеток.

В некоторых вариантах осуществления модификации остова фосфатная группа модифицированного остатка может быть модифицирована путем замены одного или более атомов кислорода другим заместителем. Кроме того, модифицированный остаток, например, модифицированный остаток, присутствующий в модифицированной нуклеиновой кислоте, может характеризоваться полной заменой немодифицированного фосфатного фрагмента модифицированной фосфатной группой, как описано в данном документе. В некоторых вариантах осуществления модификация основной цепи фосфатной основной цепи может включать изменения, которые приводят к получению незаряженного линкера или заряженного линкера с несимметричным распределением заряда.

Примеры модифицированных фосфатных групп включают тиофосфат, фосфороселенаты, боранофосфаты, сложные эфиры боранофосфатов, гидрофосфонаты, фосфороамидаты, алкил- или арилфосфонаты и фосфотриэфиры. Атом фосфора в немодифицированной фосфатной группе является ахиральным. Однако замещение одного из мостиковых атомов кислорода одним из вышеперечисленных атомов или групп атомов может сделать атом фосфора хиральным. Стереогенный атом фосфора может обладать «R»-конфигурацией (в данном документе - Rp) или «S»-конфигурацией (в данном документе - Sp). Остов также можно модифицировать путем замены мостикового кислорода (т. е. кислорода, который связывает фосфат с нуклеозидом) на азот (мостиковые фосфорамиидаты), серу (мостиковые тиофосфаты) и углерод (мостиковые метиленфосфонаты). Замена может происходить либо в связывающем атоме кислорода, либо в обоих связывающих атомах кислорода.

В определенных модификациях остова фосфатная группа может быть заменена не содержащими фосфор соединителями. В некоторых вариантах осуществления заряженная фосфатная группа может быть заменена нейтральным фрагментом. Примеры фрагментов, которые могут заменять фосфатную группу, могут включать, без ограничения, например, метилфосфонат, гидроксиламино, силосан, карбонат, карбоксиметил, карбамат, амид, тиоэфир, линкер на основе этиленоксида, сульфонат, сульфонамид, тиоформацеталь,

формацеталь, оксим, метиленимино, метиленметиличино, метиленгидразо, метилендиметилгидразо и метиленоксиметиличино.

Также можно конструировать каркасы, которые могут имитировать нуклеиновые кислоты, в которых фосфатный линкер и сахар рибозы заменены нуклеозидными или нуклеотидными суррогатами, устойчивыми к нуклеазам. Такие модификации могут включать модификации остова и сахара. В некоторых вариантах осуществления нуклеооснования могут быть плотно связаны посредством суррогатного остова. Примеры могут включать, без ограничения, суррогаты нуклеозидов морфолино, циклобутила, пирролидина и пептидной нуклеиновой кислоты (ПНК).

Модифицированные нуклеозиды и модифицированные нуклеотиды могут содержать одну или более модификаций группы сахара, т. е. модификации сахара. Например, 2'-гидроксильная группа (ОН) может быть модифицирована, например, заменена рядом разных «окси»- или «дезокси»-заместителей. В некоторых вариантах осуществления модификации 2'-гидроксильной группы могут повысить стабильность нуклеиновой кислоты, поскольку гидроксил уже невозможно депротонировать с образованием 2'-алкоксидного иона.

Примеры модификаций 2'-гидроксильных групп могут включать алкокси или арилокси (OR, где «R» может представлять собой, например, алкил, циклоалкил, арил, аралкил, гетероарил или сахар); полиэтиленгликоли (ПЭГ),  $O(CH_2CH_2O)_nCH_2CH_2OR$ , где R может представлять собой, например, H или необязательно замещенный алкил, а n может представлять собой целое число от 0 до 20 (например, от 0 до 4, от 0 до 8, от 0 до 10, от 0 до 16, от 1 до 4, от 1 до 8, от 1 до 10, от 1 до 16, от 1 до 20, от 2 до 4, от 2 до 8, от 2 до 10, от 2 до 16, от 2 до 20, от 4 до 8, от 4 до 10, от 4 до 16 и от 4 до 20). В некоторых вариантах осуществления модификация 2'-гидроксильной группы может представлять собой 2'-O-Me. В некоторых вариантах осуществления модификация 2'-гидроксильной группы может представлять собой 2'-фтор-модификацию, которая состоит в замене 2'-гидроксильной группы фторидом. В некоторых вариантах осуществления модификация 2'-гидроксильной группы может представлять собой 2'-H, которая состоит в замене 2'-гидроксильной группы водородом. В некоторых вариантах осуществления модификация 2'-гидроксильной группы может включать «запертые» нуклеиновые кислоты (ЗНК), в которых 2'-гидроксил может быть соединен, например, посредством  $C_{1-6}$  алкиленового или  $C_{1-6}$  гетероалкиленового мостика, с 4'-атомом углерода того же рибозного сахара, где типовые мостики могут включать метиленовые, пропиленовые, эфирные или аминные мостики; O-амино (где амино может представлять собой, например,  $NH_2$ ; алкиламино, диалкиламино, гетероцикл, ариламино, диариламино, гетероариламино или дигетероариламино, этилендиамин или полиамино) и аминококси,  $O(CH_2)_n$ -амино (где амино может представлять собой, например,  $NH_2$ ; алкиламино, диалкиламино, гетероцикл, ариламино, диариламино, гетероариламино или дигетероариламино, этилендиамин или полиамино). В некоторых вариантах осуществления модификация 2'-гидроксильной группы может включать «открытые» нуклеиновые кислоты (ОНК), в

которых в рибозном кольце отсутствует связь C2'-C3'. В некоторых вариантах осуществления модификация 2'-гидроксильной группы может включать метоксиэтильную группу (МОЭ), (OCH<sub>2</sub>CH<sub>2</sub>OCH<sub>3</sub>, например, производное ПЭГ).

«Дезокси»-2'-модификации могут включать водород (т. е. сахара дезоксирибозы, например, в выступающих частях частично дцРНК); галоген (например, бром, хлор, фтор или йод); amino (где amino может представлять собой, например, NH<sub>2</sub>; алкиламино, диалкиламино, гетероцикл, ариламино, диариламино, гетероариламино, дигетероариламино или аминокислоту; NH(CH<sub>2</sub>CH<sub>2</sub>NH)<sub>n</sub>CH<sub>2</sub>CH<sub>2</sub>- amino (где amino может быть, например таким, как описано в данном документе), -NHC(O)R (где R может представлять собой, например, алкил, циклоалкил, арил, аралкил, гетероарил или сахар), циано; меркапто; алкилтиоалкил; тиоалкокси; и алкил, циклоалкил, арил, алкенил и алкинил, которые могут быть необязательно замещены, например, amino, как описано в данном документе.

Модификация сахара может включать сахарную группу, которая также может содержать один или более атомов углерода, которые обладают стереохимической конфигурацией, противоположной конфигурации соответствующего атома углерода в рибозе. Таким образом, модифицированная нуклеиновая кислота может содержать нуклеотиды, содержащие, например, арабинозу в качестве сахара. Модифицированные нуклеиновые кислоты могут также содержать абазические сахара. Эти абазические сахара также могут быть дополнительно модифицированы в одном или более составляющих сахар атомах. Модифицированные нуклеиновые кислоты могут также содержать один или более сахаров в L-форме, например, L-нуклеозиды.

Описанные в данном документе модифицированные нуклеозиды и модифицированные нуклеотиды, которые могут быть включены в модифицированную нуклеиновую кислоту, могут содержать модифицированное основание, также называемое нуклеосооснованием. Примеры нуклеосооснований включают, но не ограничиваются этим, аденин (A), гуанин (G), цитозин (C) и урацил (U). Эти нуклеосооснования могут быть модифицированы или полностью заменены для получения модифицированных остатков, которые могут быть включены в модифицированные нуклеиновые кислоты. Нуклеосооснование нуклеотида может быть независимо выбрано из пурина, пиримидина, аналога пурина или аналога пиримидина. В некоторых вариантах осуществления нуклеосооснование может включать, например, встречающиеся в природе и синтетические производные основания.

В вариантах осуществления, в которых используется двойная гидовая РНК, каждая из сгРНК и trасг-РНК может содержать модификации. Такие модификации могут находиться в одном или обоих концах сгРНК и/или trасгРНК. В вариантах осуществления, включающих огРНК, один или более остатков в одном или обоих концах огРНК могут быть химически модифицированы, и/или внутренние нуклеозиды могут быть модифицированы, и/или вся огРНК может быть химически модифицирована. Некоторые варианты осуществления включают модификацию 5' конца. Некоторые варианты

осуществления включают модификацию 3' конца.

В некоторых вариантах осуществления гидовые РНК, описанные в данном документе, имеют один из профилей модификации, описанных в WO2018/107028 A1, поданной 8 декабря 2017 г., под названием «Chemically Modified Guide RNAs» («Химически модифицированные гидовые РНК»), содержание которой в полном объеме включено в данный документ посредством ссылки. В некоторых вариантах осуществления гидовые РНК, описанные в данном документе, имеют один из профилей структуры/модификации, описанных в US20170114334, содержание которой в полном объеме включено в данный документ посредством ссылки. В некоторых вариантах осуществления гидовые РНК, описанные в данном документе, имеют один из профилей структуры/модификации, описанных в WO2017/136794, содержание которой в полном объеме включено в данный документ посредством ссылки.

В некоторых вариантах осуществления оРНК по настоящему изобретению имеют профили модификации, приведенные ниже в таблице 2. «Полная последовательность» с таблице 2 относится к последовательности оРНК для каждой из направляющих, перечисленных в таблице 1. «Модифицированная полная последовательность» демонстрирует профиль модификации для каждой оРНК.

**Таблица 2: оРНК и профили модификации для оРНК направляющих последовательностей альбумина человека**

<b>ID направляющей</b>	<b>Полная последовательность</b>	<b>SEQ ID NO:</b>	<b>Модифицированная полная последовательность</b>	<b>SEQ ID NO:</b>
G009844	GAGCAACCUCACUCUUGUCU GUUUUAGAGCUAGAAAUAG CAAGUUA AAAUAAGGCUAG UCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	34	mG*mA*mG*CAACCUCAC UCUUGUCUGUUUAGAm GmCmUmAmGmAmAmAm UmAmGmCAAGUUA AAAU AAGGCUAGUCCGUUAUC AmAmCmUmUmGmAmAm AmAmAmGmUmGmGmCm AmCmCmGmAmGmUmCmG mGmUmGmCmU*mU*mU* mU	66
G009851	AUGCAUUUGUUUCA AAAUA UGUUUUAGAGCUAGAAUA GCAAGUUA AAAUAAGGCUA GUCCGUUAUCAACUUGAAAA	35	mA*mU*mG*CAUUUGUUU CA AAAUAUGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA	67

	AGUGGCACCGAGUCGGUGCU UUU		AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G00985 2	UGCAUUUGUUUCAAAAUAU UGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	36	mU*mG*mC*AUUUGUUUC AAAAUAUUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	68
G00985 7	AUUUAUGAGAUCAACAGCAC GUUUUAGAGCUAGAAAUA CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAA GUGGCACCGAGUCGGUGCUU UU	37	mA*mU*mU*UAUGAGAUC AACAGCACGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	69
G00985 8	GAUCAACAGCACAGGUUUUG GUUUUAGAGCUAGAAAUA CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAA GUGGCACCGAGUCGGUGCUU UU	38	mG*mA*mU*CAACAGCAC AGGUUUUGGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	70
G00985 9	UUAAAUAAGCAUAGUGCA AGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUA	39	mU*mU*mA*AAUAAAGCA UAGUGCAAGUUUUAGAm GmCmUmAmGmAmAmAmU	71



	GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU		mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G00986 0	UAAAGCAUAGUGCAAUGGA UGUUUUAGAGCUAGAAAUA GCAAGUUA AAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	40	mU*mA*mA*AGCAUAGUG CAAUGGAUGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	72
G00986 1	UAGUGCAAUGGAUAGGUCU UGUUUUAGAGCUAGAAAUA GCAAGUUA AAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	41	mU*mA*mG*UGCAAUGGA UAGGUCUUGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	73
G00986 6	UACUAAAACUUUAUUUUAC UGUUUUAGAGCUAGAAAUA GCAAGUUA AAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	42	mU*mA*mC*UAAAACUUU AUUUUACUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	74
G00986 7	AAAGUUGAACAAUAGAAAA AGUUUUAGAGCUAGAAAUA	43	mA*mA*mA*GUUGAACAA UAGAAAAGUUUUAGAm	75

	GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU		GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G00986 8	AAUGCAUAAUCUAAGUCA AGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	44	mA*mA*mU*GCAUAAUCU AAGUCAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	76
G00987 4	UAAUAAAAUCAAACAUCU GUUUUAGAGCUAGAAUAG CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAA GUGGCACCGAGUCGGUGCU UU	45	mU*mA*mA*UAAAAUCA AACAUCCUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	77
G01274 7	GCAUCUUUAAAGAAUUAU UGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	46	mG*mC*mA*UCUUUAAAG AAUUAUUUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	78
G01274	UUUGGCAUUUAUUUCUAAA	47	mU*mU*mU*GGCAUUUAU	79

8	AGUUUUAGAGCUAGAAAUA GCAAGUUA AAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU		UUCUAAAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G01274 9	UGUAUUUGUGAAGUCUAC AGUUUUAGAGCUAGAAAUA GCAAGUUA AAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	48	mU*mG*mU*AUUUGUGAA GUCUACAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	80
G01275 0	UCCUAGGUAAAAAAAAAAAA AGUUUUAGAGCUAGAAAUA GCAAGUUA AAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	49	mU*mC*mC*UAGGUAAAA AAAAAAAAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	81
G01275 1	UAAUUUUCUUUUGCGCACUA GUUUUAGAGCUAGAAUAG CAAGUUA AAAUAAGGCUAG UCCGUUAUCAACUUGAAAA GUGGCACCGAGUCGGUGCUU UU	50	mU*mA*mA*UUUUCUUUU GCGCACUAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	82

G01275 2	UGACUGAAACUUCACAGAAU GUUUUAGAGCUAGAAAUAG CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUU UU	51	mU*mG*mA*CUGAAACUU CACAGAAUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	83
G01275 3	GACUGAAACUUCACAGAAUA GUUUUAGAGCUAGAAAUAG CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUU UU	52	mG*mA*mC*UGAAACUUC ACAGAAUAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	84
G01275 4	UUCAUUUUAGUCUGUCUUCU GUUUUAGAGCUAGAAAUAG CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUU UU	53	mU*mU*mC*AUUUUAGUC UGUCUUCUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	85
G01275 5	AUUAUCUAAGUUUGAAUAU AGUUUUAGAGCUAGAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	54	mA*mU*mU*AUCUAAGUU UGAAUAUAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm	86

			UmGmCmU*mU*mU*mU	
G01275 6	AAUUUUUAAAAUAGUAUUC UGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	55	mA*mA*mU*UUUUAAAAU AGUAUUCUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	87
G01275 7	UGAAUUAUUCUUCUGUUUA AGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	56	mU*mG*mA*AUUAUUCUU CUGUUUAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	88
G01275 8	AUCAUCCUGAGUUUUUCUGU GUUUUAGAGCUAGAAAUA CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAA GUGGCACCGAGUCGGUGCUU UU	57	mA*mU*mC*AUCCUGAGU UUUUCUGUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	89
G01275 9	UUACUAAAACUUUAUUUUA CGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	58	mU*mU*mA*CUAAAACUU UAUUUUACGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC	90

			mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G01276 0	ACCUUUUUUUUUUUUUUUACCU GUUUUAGAGCUAGAAAUAG CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUU UU	59	mA*mC*mC*UUUUUUUUU UUUUACCUGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	91
G01276 1	AGUGCAAUGGAUAGGUCUU UGUUUUAGAGCUAGAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	60	mA*mG*mU*GCAAUGGAU AGGUCUUUGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	92
G01276 2	UGAUUCCUACAGAAAAACUC GUUUUAGAGCUAGAAAUAG CAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUU UU	61	mU*mG*mA*UCCUACAG AAAAACUCGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	93
G01276 3	UGGGCAAGGGAAGAAAAA AGUUUUAGAGCUAGAAUA GCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCU UUU	62	mU*mG*mG*GCAAGGGAA GAAAAAAGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm	94

			AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G01276 4	CCUCACUCUUGUCUGGGCAA GUUUUAGAGCUAGAAAUAG CAAGUUA AAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUU UU	63	mC*mC*mU*CACUCUUGU CUGGGCAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	95
G01276 5	ACCUCACUCUUGUCUGGGCA GUUUUAGAGCUAGAAAUAG CAAGUUA AAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUU UU	64	mA*mC*mC*UCACUCUUG UCUGGGCAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	96
G01276 6	UGAGCAACCUCACUCUUGUC GUUUUAGAGCUAGAAAUAG CAAGUUA AAAUAAGGCUAG UCCGUUAUCAACUUGAAAAA GUGGCACCGAGUCGGUGCUU UU	65	mU*mG*mA*GCAACCUCA CUCUUGUCGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	97

В некоторых вариантах осуществления модифицированная оГРНК содержит следующую последовательность:

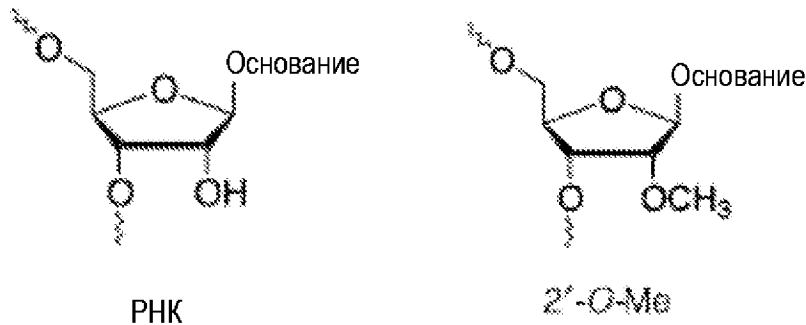
mN\*mN\*mN\*NNNNNNNNNNNNNNNNNGUUUAGAmGmCmUmAmGmAmAmAmUm  
AmGmCAAGUUA AAAUAAGGCUAGUCCGUUAUCAmAmCmUmUmGmAmAmAmAmA  
mGmUmGmGmCmAmCmCmGmAmGmUmCmGmGmUmGmCmU\*mU\*mU\*mU (SEQ ID

NO: 350), где «N» может представлять собой любой природный или неприродный нуклеотид, а совокупность N составляет направляющую последовательность интрона 1 альбумина, описанную в таблице 1. Например, в данный документ включена SEQ ID NO: 350, в которой отсутствуют N из SEQ ID NO: 350, но которая содержит модифицированную консервативную часть гРНК.

Любая из модификаций, описанных ниже, может присутствовать в гРНК и мРНК, описанных в данном документе.

Термины «mA», «mC», «mU» или «mG» можно использовать для обозначения нуклеотида, который был модифицирован с помощью 2'-O-Me.

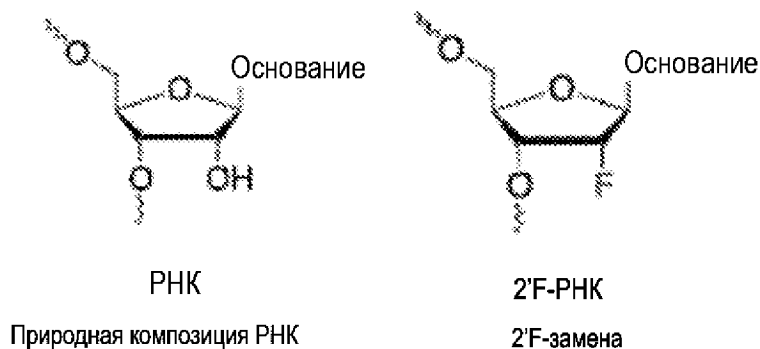
Модификацию 2'-O-метила можно проиллюстрировать следующим образом:



Другой химической модификацией, которая, как было показано, влияет на сахарные кольца нуклеотидов, является замещение галогеном. Например, замена 2'-фтор (2'-F) в сахарных кольцах нуклеотидов может повысить аффинность связывания олигонуклеотидов и стойкость к нуклеазам.

В данной заявке термины «fA», «fC», «fU» или «fG» можно использовать для обозначения нуклеотида, который был замещен 2'-F.

Замену 2'-F можно проиллюстрировать следующим образом:



Тиофосфатная (PS) связь относится к связи, в которой атом серы замещен одним немостиковым фосфатным атомом кислорода в фосфодиэфирной связи, например, в связях между нуклеотидными основаниями. Когда тиофосфаты используют для создания олигонуклеотидов, модифицированные олигонуклеотиды также можно называть S-олигонуклеотидами.

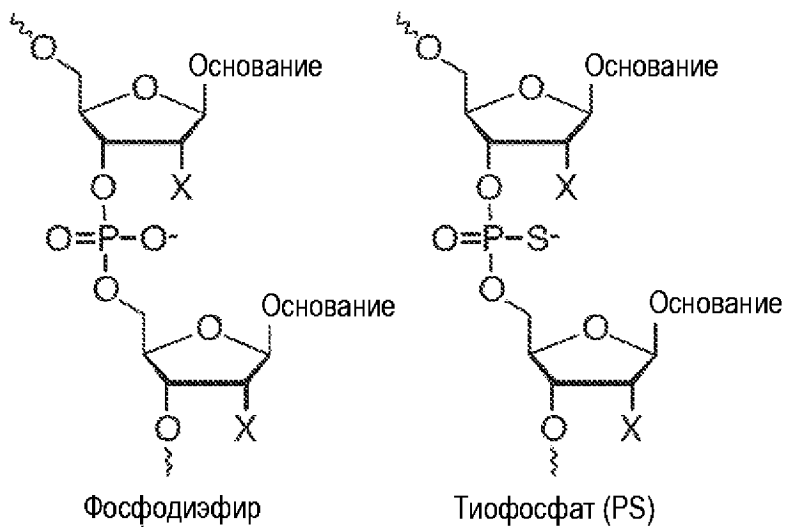
«\*» можно использовать для обозначения модификации PS. В этой заявке термины



A\*, C\*, U\* или G\* можно использовать для обозначения нуклеотида, связанного со следующим (например, 3') нуклеотидом посредством связи PS.

В этой заявке термины «mA\*», «mC\*», «mU\*» или «mG\*» можно использовать для обозначения нуклеотида, замещенного 2'-O-Me и связанного со следующим (например, 3') нуклеотидом посредством связи PS.

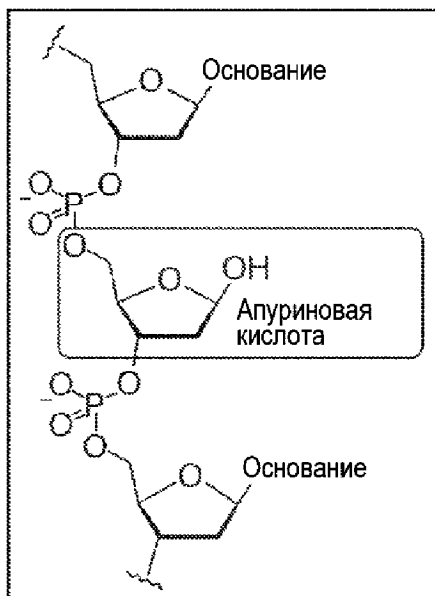
На диаграмме ниже проиллюстрировано замещение S- на немостиновый фосфатный атом кислорода с образованием связи PS вместо фосфодиэфирной связи:



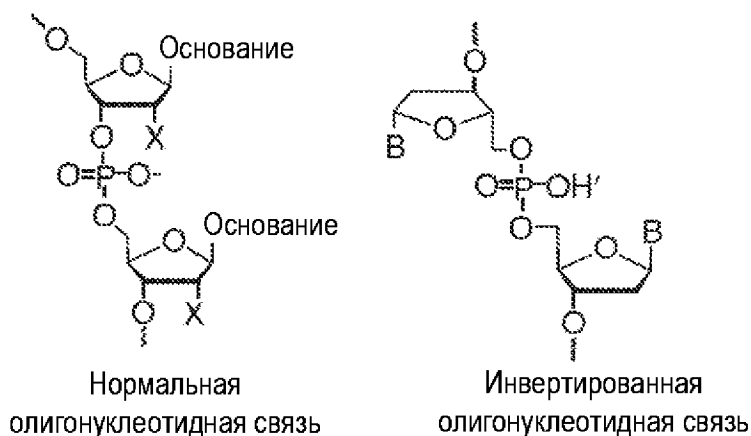
Природная фосфодиэфирная  
связь РНК

Модифицированная  
тиофосфатная (PS) связь

Абазические нуклеотиды относятся к нуклеотидам, в которых отсутствуют азотистые основания. На фигуре ниже изображен олигонуклеотид с абазическим (также известным как апуриновый) сайтом, в котором отсутствует основание:



Инвертированные основания относятся к тем, связи которых инвертированы относительно нормальной связи 5' с 3' (т. е. связь 5' с 5' или связь 3' с 3'). Например:



Абазический нуклеотид может быть присоединен посредством инвертированной связи. Например, абазический нуклеотид может быть присоединен к концевому 5' нуклеотиду посредством связи 5' с 5', или абазический нуклеотид может быть присоединен к концевому 3' нуклеотиду посредством связи 3' с 3'. Инвертированный абазический нуклеотид в концевом 5' или 3' нуклеотиде также может называться инвертированным абазическим концевым кэпом.

В некоторых вариантах осуществления модифицированы один или более из первых трех, четырех или пяти нуклеотидов в 5' конце и один или более из последних трех, четырех или пяти нуклеотидов в 3' конце. В некоторых вариантах осуществления модификация представляет собой 2'-O-Me, 2'-F, инвертированный абазический нуклеотид, связь PS или другую нуклеотидную модификацию, хорошо известную в данной области техники, для повышения стабильности и/или эффективности.

В некоторых вариантах осуществления первые четыре нуклеотида в 5' конце и последние четыре нуклеотида на 3' конце связаны тиофосфатными (PS) связями.

В некоторых вариантах осуществления первые три нуклеотида в 5' конце и последние три нуклеотида в 3' конце содержат модифицированный 2'-O-метил (2'-O-Me) нуклеотид. В некоторых вариантах осуществления первые три нуклеотида в 5' конце и последние три нуклеотида в 3' конце содержат модифицированный 2'-фтором (2'-F) нуклеотид. В некоторых вариантах осуществления первые три нуклеотида в 5' конце и последние три нуклеотида в 3' конце содержат инвертированный абазический нуклеотид.

В некоторых вариантах осуществления гидовая РНК содержит модифицированную оРНК. В некоторых вариантах осуществления оРНК имеет профиль модификации, приведенный в SEQ ID NO: 350, где N представляет собой любой природный или неприродный нуклеотид, и где совокупность N составляет направляющую последовательность, которая направляет нуклеазу к целевой последовательности в интроне 1 альбумина человека, например, как показано в таблице 1.

В некоторых вариантах осуществления гидовая РНК содержит оРНК, приведенную в любой из SEQ ID NO: 34-97 или 120-163. В некоторых вариантах осуществления гидовая РНК содержит оРНК, приведенную в любой из SEQ ID NO: 197-229. В некоторых вариантах осуществления гидовая РНК содержит оРНК, содержащую

любую из направляющих последовательностей SEQ ID NO: 2-33 или 98-119 и нуклеотиды SEQ ID NO: 301, причем нуклеотиды SEQ ID NO: 301 расположены в 3'-конце направляющей последовательности, а огРНК может быть модифицированной, как показано в таблице 2 или SEQ ID NO: 350. В некоторых вариантах осуществления гидовая РНК содержит огРНК, содержащую любую из направляющих последовательностей SEQ ID NO: 2-33 или 197-229 и нуклеотиды SEQ ID NO: 301, причем нуклеотиды SEQ ID NO: 301 расположены в 3'-конце направляющей последовательности, а огРНК может быть модифицированной, как показано в таблице 2 или SEQ ID NO: 350.

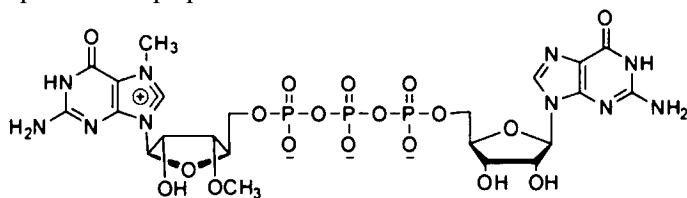
Как отмечено выше, в некоторых вариантах осуществления описанные в данном документе композиция или состав содержат мРНК, содержащую открытую рамку считывания (ОРС), кодирующую РНК-направляемый ДНК-связывающий агент, такой как Cas-нуклеаза, как описано в данном документе. В некоторых вариантах осуществления предложена, применяется или вводится мРНК, содержащая ОРС, кодирующую РНК-направляемый ДНК-связывающий агент, такой как Cas-нуклеаза. 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.

В некоторых вариантах осуществления модифицированная ОРС 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.

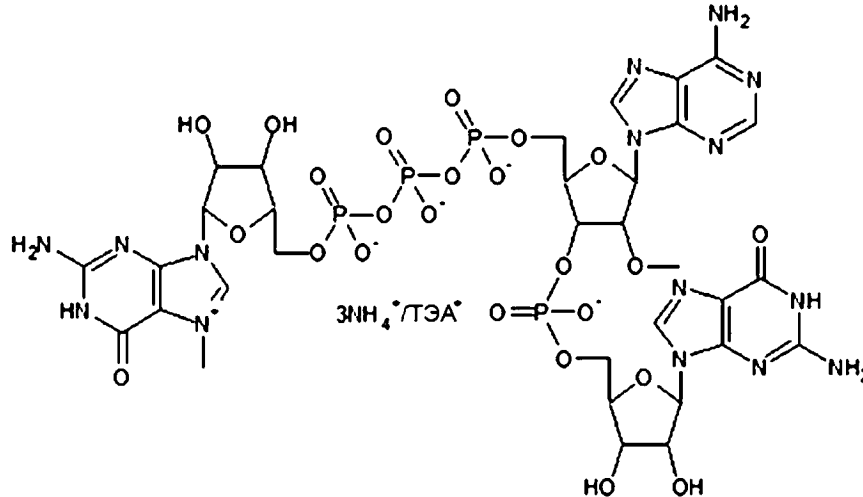
В некоторых вариантах осуществления описанная в данном документе мРНК содержит 5'-кэп, такой как Cap0, Cap1 или Cap2. 5'-кэп обычно представляет собой рибонуклеотид 7-метилгуанина (который может быть дополнительно модифицирован, как обсуждается ниже, например, в отношении ARCA), связанный посредством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

#### РНК-направляемый ДНК-связывающий агент

Как описано в данном документе, гидовые РНК по настоящему изобретению применяют в сочетании с РНК-направляемым ДНК-связывающим агентом для вставки и экспрессии гетерологичного (экзогенного) гена в геномном локусе, таком как безопасный приемный сайт, клетки-хозяина. РНК-направляемый ДНК-связывающий агент может представлять собой белок или нуклеиновую кислоту, кодирующую белок, такую как мРНК. В некоторых вариантах осуществления способы по настоящему изобретению включают применение композиции, которая содержит гидовую РНК, содержащую направляющую последовательность из таблицы 1, и РНК-направляемый ДНК-связывающий агент, например, нуклеазу, такую как Cas-нуклеаза (например, Cas9), с образованием рибонуклеопротеинового комплекса.

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 нуклеазы, 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 beccsii*, *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 *Acaryochloris 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 macacae*. 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 (РНКП). В некоторых вариантах осуществления РНК-направляемый ДНК-связывающий агент представляет собой Cas-нуклеазу. В некоторых вариантах осуществления гРНК вместе с Cas-нуклеазой называется Cas-РНКП. В некоторых вариантах осуществления РНКП 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-РНКП.

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-никаза 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 C-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 C-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, PKKKRKV (SEQ ID NO: 600) или PKKKRRV (SEQ ID NO: 601). В некоторых вариантах осуществления СЯЛ может быть двухкомпонентной последовательностью, такой как СЯЛ нуклеоплазмина, KRPAATKKGAGQAKKKK (SEQ ID NO: 602). В конкретном варианте осуществления одна PKKKRKV (SEQ ID NO: 600) СЯЛ может быть присоединена в C-конце РНК-направляемого ДНК-связывающего агента. В сайт слияния, необязательно, включены один или более линкеров.

#### Донорные конструкции/последовательности

Композиции и способы, описанные в данном документе, включают применение конструкции нуклеиновой кислоты, которая содержит последовательность, кодирующую гетерологичный ген, для вставки в сайт надреза, созданный гидовой РНК по настоящему изобретению и РНК-направляемым ДНК-связывающим агентом. В контексте данного документа такую конструкцию иногда называют «донорной конструкцией/матрицей». Конструкции могут кодировать любую экспрессируемую нуклеиновую кислоту (т. е. нуклеиновую кислоту, которая может быть экспрессирована), например, ДНК, матричную РНК (мРНК), функциональную РНК, малую интерферирующую РНК (миРНК), микроРНК (микро-РНК), одноцепочечную РНК (оцРНК), длинные некодирующие РНК или

антисмысловые олигонуклеотиды.

Описанные в данном документе композиции и способы включают применение не двунаправленной или однонаправленной конструкции, например, кодирующей один трансген, кодирующей два трансгена в *цис*-ориентации и т. д. Однонаправленная конструкция может содержать кодирующую последовательность, связанную с акцептором сплайсинга.

Описанные в данном документе композиции и способы включают применение двунаправленной конструкции, описанной в данном документе, содержащей по меньшей мере два сегмента нуклеиновой кислоты в *цис*-ориентации, при этом один сегмент (первый сегмент) содержит кодирующую последовательность или трансген, а другой сегмент (второй сегмент) содержит последовательность, в которой комплементарная последовательность кодирует трансген. 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 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 трансгенов. 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, Burset, et al., 2001, *Nucleic Acids Res.*, 29, 255-259.

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. 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. Длина конструкции может варьироваться в зависимости от размера вставляемого гена и может составлять, например, от 200 пар оснований (п. о.) до около 5000 п. о., например, от около 200 п. о. до около 2000 п. о., например, от около 500 п. о. до около 1500 п. о. В некоторых вариантах осуществления длина донорной матрицы ДНК составляет около 200 п. о., или около 500 п. о., или около 800 п. о., или около 1000 пар оснований, или около 1500 пар оснований. В других вариантах осуществления длина донорной матрицы составляет по меньшей мере 200 п. о., или по меньшей мере 500 п. о., или по меньшей мере 800 п. о., или по меньшей мере 1000 п. о., или по меньшей мере 1500 п. о.

Конструкция может представлять собой ДНК или РНК, быть одноцепочечной, двухцепочечной или частично одно- и частично двухцепочечной, и может быть внесена в клетку-хозяин 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 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 экзогенных полинуклеотидов от деградации включают, но не ограничиваются этим, добавление концевой(ых) аминокислотной группы (аминогрупп) и использование модифицированных межнуклеотидных связей, таких как, например, тиофосфаты, фосфорамидаты и остатки О-метилрибозы или дезоксирибозы. Конструкцию можно вносить в клетку как часть векторной молекулы, имеющей дополнительные последовательности, такие как, например, точки начала репликации, промоторы и гены, кодирующие устойчивость к антибиотикам. В конструкции могут отсутствовать вирусные элементы. Более того, донорные конструкции можно вносить в виде оголенной нуклеиновой кислоты, в виде комплекса нуклеиновой кислоты с агентом, таким как липосома или полксамер, или можно доставлять с помощью вирусов (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.

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, расположенную ниже сигнальной последовательности, кодирующей сигнальный пептид, например, сигнальный пептид альбумина, сигнальный пептид из гепатоцитарного секретируемого белка, и функционально связанную с ней. Конструкция может содержать последовательность, кодирующую гетерологичный белок, расположенную ниже сигнальной последовательности, кодирующей сигнальный пептид из гетерологичного белка, и функционально связанную с ней. В некоторых вариантах осуществления конструкция нуклеиновой кислоты работает при независимой от гомологии вставке нуклеиновой кислоты, которая кодирует трансгенный белок. 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.

Конструкция может представлять собой двунаправленную конструкцию нуклеиновой кислоты, содержащую 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 кодирующая последовательность кодирует терапевтический агент, такой как полипептид, функциональную РНК или энхансер. 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. 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 2) 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 и второй сегмент, каждый из которых имеет акцептор сплайсинга, расположенный выше трансгена. 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 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 саморасщепляющийся пептид), связанных с полипептидом. Кодирующая 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 секретлируемый полипептид. В некоторых вариантах осуществления полипептид является таким, чья функция обычно осуществляется (например, который является функционально активным) в виде секретлируемого полипептида. В контексте данного документа «секретлируемый полипептид» относится к белку, который секретируется 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 секретируется 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 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 конструкции, например, двунаправленные конструкции нуклеиновых кислот, способны к вставке в геномный локус посредством негомологичного соединения концов (НГСК). В некоторых вариантах осуществления описанные в данном документе конструкции представляют собой независимые от гомологии донорные конструкции. 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 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. Саморасщепляющийся

пептид может представлять собой пептид P2A, пептид T2A и т. п. 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 A и областью B (A-2A-B). Наличие пептида 2A приведет к расщеплению одного длинного белка на белок A, белок B и пептид 2A. Белок A и белок B 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-A» в 3' конце первого и/или второго сегмента.

#### Способы доставки

Описанные в данном документе гидовые РНК можно доставлять в клетку-хозяина или организм субъекта *in vivo* или *ex vivo*, используя различные известные и подходящие способы, доступные в данной области техники. Гидовые РНК можно доставлять вместе (отдельно или в комбинации) с РНК-направляемым ДНК-связывающим агентом, таким как Cas или нуклеиновая кислота, кодирующая Cas9 (например, Cas9 или нуклеиновая кислота, кодирующая Cas9), и конструкцией, которая содержит последовательность, кодирующую гетерологичный ген, подлежащий вставке в сайт разрезания гидовой РНК по настоящему изобретению, как описано в данном документе.

Для внесения описанной в данном документе гидовой РНК, а также РНК-направляемого ДНК-связывающего агента и донорной конструкции в клетки (например, клетки млекопитающих) и целевые ткани можно использовать обычные способы доставки генов на вирусной и невирусной основе. В данном документе дополнительно предложены нуклеиновые кислоты систем доставки на основе невирусных векторов, таких как невирусные векторы, плазмидные векторы и, например, оголенная нуклеиновая кислота, и нуклеиновая кислота в комплексе с носителем для доставки, таким как липосома, липидная наночастица (ЛПЧ) 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 гидовые РНК, РНК-направляемый ДНК-связывающий агент и донорную конструкцию, отдельно или в комбинации, 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 некоторых вариантах осуществления описанные в данном документе композиции гидовых РНК, одних или кодируемых в одном или более векторах, составляют или вводят с помощью липидной наночастицы; смотрите, например, РСТ/US2017/024973, содержание которой в полном объеме включено в данный документ посредством ссылки. Любой состав липидных наночастиц (ЛНП), известный специалистам в данной области техники как способный доставлять нуклеотиды в организм субъекта, можно использовать с описанными в данном документе гидовыми РНК, а также мРНК, кодирующей РНК-направляемый ДНК-связывающий агент, такой как Cas или Cas9, или РНК-направляемым ДНК-связывающим агентом, таким как сам белок Cas или Cas9.

В некоторых вариантах осуществления описанные в данном документе гидовые РНК можно доставлять в клетку-хозяина (*in vitro* или *in vivo*) посредством ЛНЧ. В некоторых вариантах осуществления гРНК/ЛНЧ также связаны с РНК-направляемым ДНК-связывающим агентом, таким как Cas9, или мРНК, кодирующей РНК-направляемый ДНК-связывающий агент, такой как Cas9. В некоторых вариантах осуществления гРНК/ЛНЧ также связаны с донорной конструкцией, описанной в данном документе.

В некоторых вариантах осуществления настоящее изобретение включает способ доставки описанных в данном документе гРНК в клетку *in vitro*, причем доставка гРНК происходит с помощью ЛНЧ. В некоторых вариантах осуществления гРНК доставляют не связанными с ЛНЧ способами, например, посредством AAV-системы, а РНК-

направляемый ДНК-связывающий агент (например, Cas9) или мРНК, кодирующую РНК-направляемый ДНК-связывающий агент (например, Cas9), и/или донорную конструкцию доставляют с помощью ЛНЧ.

В некоторых вариантах осуществления в настоящем изобретении предложена композиция, содержащая любую из описанных в данном документе гРНК и ЛНЧ. В некоторых вариантах осуществления композиция дополнительно содержит Cas9 или мРНК, кодирующую Cas9, или другой РНК-направляемый ДНК-связывающий агент, описанный в данном документе. В некоторых вариантах осуществления композиция дополнительно содержит описанную в данном документе донорную конструкцию.

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)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.

В некоторых вариантах осуществления любые из описанных в данном документе гидовых РНК, РНК-направляемых ДНК-связывающих агентов и/или донорных конструкций (например, двунаправленных конструкций), отдельно или в комбинации, оголенные 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.

Электропорация также является хорошо известным средством доставки груза, и любые методы электропорации можно использовать для доставки любой из описанных в данном документе гРНК. В некоторых вариантах осуществления электропорацию можно использовать для доставки любой из описанных в данном документе гРНК, необязательно, с РНК-направляемым ДНК-связывающим агентом, таким как Cas9, или мРНК, кодирующей РНК-направляемый ДНК-связывающий агент, такой как Cas9, доставляемыми одним или разными способами. В некоторых вариантах осуществления электропорацию можно использовать для доставки любой из описанных в данном документе гРНК и описанной в данном документе донорной конструкции.

В определенных вариантах осуществления в настоящем изобретении предложены ДНК- или РНК-векторы, кодирующие любую из гидовых РНК, содержащих любые одну или более направляющих последовательностей, описанных в данном документе. В некоторых вариантах осуществления изобретение включает ДНК- или РНК-векторы кодирующие любую одну или более из направляющих последовательностей, описанных в

данном документе. В некоторых вариантах осуществления помимо последовательностей РНК векторы дополнительно содержат нуклеиновые кислоты, которые не кодируют гидовые РНК. Нуклеиновые кислоты, которые не кодируют гидовые РНК, включают, но не ограничиваются этим, промоторы, энхансеры, регуляторные последовательности, нуклеиновые кислоты, кодирующие РНК-направляемый ДНК-связывающий агент, который может представлять собой нуклеазу, такую как Cas9, и донорную конструкцию, содержащую гетерологичный ген. В некоторых вариантах осуществления вектор содержит одну или более нуклеотидных последовательностей, кодирующих sgРНК, trРНК или crРНК и trРНК, как описано в данном документе.

В некоторых вариантах осуществления вектор содержит одну или более нуклеотидных последовательностей, кодирующих ogРНК и mРНК, кодирующие РНК-направляемый ДНК-связывающий агент, который может представлять собой белок Cas, такой как Cas9 или Cpf1. В некоторых вариантах осуществления вектор содержит одну или более нуклеотидных последовательностей, кодирующих sgРНК, trРНК и mРНК, кодирующие РНК-направляемый ДНК-связывающий агент, который может представлять собой белок Cas, такой как Cas9 или Cpf1. В одном варианте осуществления Cas9 получен из *Streptococcus pyogenes* (т. е. Cas9 Spy). В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая sgРНК, trРНК или crРНК и trРНК (которая может представлять собой ogРНК), содержит или состоит из направляющей последовательности, фланкируемой всей или частью повторяющейся последовательности из CRISPR/Cas природного происхождения. Нуклеиновая кислота, содержащая sgРНК, trРНК или crРНК и trРНК или состоящая из них, может дополнительно содержать векторную последовательность, причем векторная последовательность содержит или состоит из нуклеиновых кислот, которые в природе не встречаются вместе с sgРНК, trРНК или crРНК и trРНК.

В некоторых вариантах осуществления sgРНК и trРНК кодируются не непрерывными нуклеиновыми кислотами в одном векторе. В других вариантах осуществления sgРНК и trРНК могут кодироваться непрерывной нуклеиновой кислотой. В некоторых вариантах осуществления sgРНК и trРНК кодируются противоположными цепями одной нуклеиновой кислоты. В некоторых вариантах осуществления sgРНК и trРНК кодируются одной цепью одной нуклеиновой кислоты.

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 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 паковать 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.

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. В некоторых вариантах осуществления вирусный вектор может представлять собой AAV-вектор. В некоторых вариантах осуществления вирусный вектор может представлять собой лентивирусный вектор.

В некоторых вариантах осуществления «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 (трансген) 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 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 (Ψ), чтобы увеличить его упаковочную 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.

В некоторых вариантах осуществления вектор может быть способен управлять экспрессией одной или более кодирующих последовательностей 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 может содержать нуклеотидную последовательность, кодирующую описанный в данном документе РНК-направляемый ДНК-связывающий агент, такой как белок Cas (например, Cas9). В некоторых вариантах осуществления нуклеаза, кодируемая вектором, может представлять собой белок Cas. В некоторых вариантах осуществления векторная система может содержать одну копию нуклеотидной последовательности, кодирующей нуклеазу. В других вариантах осуществления векторная система может содержать более одной копии нуклеотидной последовательности, кодирующей нуклеазу. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая нуклеазу, может быть функционально связана по меньшей мере с одной транскрипционной или трансляционной регуляторной последовательностью. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая нуклеазу, может быть функционально связана по меньшей мере с одним промотором.

В некоторых вариантах осуществления вектор может содержать любую одну или более конструкций, содержащих описанный в данном документе гетерологичный ген. В некоторых вариантах осуществления гетерологичный ген может быть функционально связан по меньшей мере с одной транскрипционной или трансляционной регуляторной последовательностью. В некоторых вариантах осуществления гетерологичный ген может быть функционально связан по меньшей мере с одним промотором. В некоторых вариантах осуществления гетерологичный ген не связан с промотором, который управляет экспрессией гетерологичного гена.

В некоторых вариантах осуществления промотор может быть конститутивным,



индуцибельным или тканеспецифическим. В некоторых вариантах осуществления промотор может представлять собой конститутивный промотор. Неограничивающие типовые конститутивные промоторы включают немедленно ранний промотор цитомегаловируса (CMV), промотор вируса обезьян (SV40), главный поздний промотор аденовируса (MLP), промотор вируса саркомы Рауса (RSV), промотор вируса опухоли молочной железы мышей (MMTV), промотор фосфоглицераткиназы (PGK), промотор фактора элонгации-альфа (EF1a), промоторы убиквитина, промоторы актина, промоторы тубулина, промоторы иммуноглобулина, их функциональный фрагмент или комбинацию любых из вышеперечисленных промоторов. В некоторых вариантах осуществления промотор может представлять собой промотор CMV. В некоторых вариантах осуществления промотор может представлять собой усеченный промотор CMV. В других вариантах осуществления промотор может представлять собой промотор EF1a. В некоторых вариантах осуществления промотор может представлять собой индуцибельный промотор. Неограничивающие типовые индуцибельные промоторы включают те, которые индуцируются тепловым шоком, светом, химическими веществами, пептидами, металлами, стероидами, антибиотиками или спиртом. В некоторых вариантах осуществления индуцибельный промотор может представлять собой промотор, который имеет низкий базовый (неиндуцированный) уровень экспрессии, такой как, например, промотор Tet-On<sup>®</sup> (Clontech).

В некоторых вариантах осуществления промотор может представлять собой тканеспецифический промотор, например, промотор, специфический в отношении экспрессии в печени.

Вектор может дополнительно содержать нуклеотидную последовательность, кодирующую описанную в данном документе гидовую РНК. В некоторых вариантах осуществления вектор содержит одну копию гидовой РНК. В других вариантах осуществления вектор содержит более одной копии гидовой РНК. В вариантах осуществления с более чем одной гидовой РНК гидовые РНК могут быть неидентичными в том смысле, что они нацелены на разные целевые последовательности, или могут быть идентичными в том смысле, что они нацелены на одну и ту же целевую последовательность. В некоторых вариантах осуществления, в которых векторы содержат более одной гидовой РНК, каждая гидовая РНК может обладать другими отличными свойствами, такими как активность или стабильность в комплексе с РНК-направляемой ДНК-нуклеазой, например, комплексе РНП Cas. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая гидовую РНК, может быть функционально связана по меньшей мере с одной транскрипционной или трансляционной регуляторной последовательностью, такой как промотор, 3' НТО или 5' НТО. В одном варианте осуществления промотор может представлять собой промотор тРНК, например, тРНК<sup>Lys3</sup> или химерой тРНК. Смотрите Mefferd et al., RNA. 2015 21:1683-9; Scherer et al., Nucleic Acids Res. 2007 35: 2620-2628. В некоторых вариантах осуществления промотор может распознаваться РНК-полимеразой III (Pol III). Неограничивающие примеры

промоторов Pol III включают промоторы U6 и H1. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая гидовую РНК, может быть функционально связана с промотором U6 мыши или человека. В других вариантах осуществления нуклеотидная последовательность, кодирующая гидовую РНК, может быть функционально связана с промотором H1 мыши или человека. В вариантах осуществления с более чем одной гидовой РНК промоторы, используемые для управления экспрессией, могут быть одинаковыми или разными. В некоторых вариантах осуществления нуклеотид, кодирующий sgРНК гидовой РНК, и нуклеотид, кодирующий trРНК гидовой РНК, могут находиться в одном векторе. В некоторых вариантах осуществления нуклеотид, кодирующий sgРНК, и нуклеотид, кодирующий trРНК, могут находиться под управлением одного промотора. В некоторых вариантах осуществления sgРНК и trРНК могут транскрибироваться в один транскрипт. Например, sgРНК и trРНК могут процессироваться из одного транскрипта с образованием гидовой РНК из двух молекул. В альтернативном варианте sgРНК и trРНК могут транскрибироваться в гидовую РНК из одной молекулы (огРНК). В других вариантах осуществления sgРНК и trРНК могут находиться под управлением соответствующих промоторов в одном векторе. В других вариантах осуществления sgРНК и trРНК могут кодироваться разными векторами.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая гидовую РНК, может быть расположена в том же векторе, который содержит нуклеотидную последовательность, кодирующую РНК-направляемый ДНК-связывающий агент, такой как белок Cas. В некоторых вариантах осуществления экспрессия гидовой РНК и РНК-направляемого ДНК-связывающего агента, такого как белок Cas, может находиться под управлением их собственных соответствующих промоторов. В некоторых вариантах осуществления экспрессия гидовой РНК может находиться под управлением того же промотора, который управляет экспрессией РНК-направляемого ДНК-связывающего агента, такого как белок Cas. В некоторых вариантах осуществления гидовая РНК и транскрипт РНК-направляемого ДНК-связывающего агента, такого как белок Cas, могут содержаться в рамках одного транскрипта. Например, гидовая РНК может находиться в нетранслируемой области (НТО) РНК-направляемого ДНК-связывающего агента, такого как транскрипт белка Cas. В некоторых вариантах осуществления гидовая РНК может находиться в 5' НТО транскрипта. В других вариантах осуществления гидовая РНК может находиться в 3' НТО транскрипта. В некоторых вариантах осуществления внутриклеточный период полужизни транскрипта может быть уменьшен за счет содержания гидовой РНК в 3' НТО и, таким образом, сокращения длины 3' НТО. В дополнительных вариантах осуществления гидовая РНК может находиться в интроне транскрипта. В некоторых вариантах осуществления подходящие сайты сплайсинга можно добавлять к интрону, в котором расположена гидовая РНК, чтобы обеспечить правильный сплайсинг гидовой РНК из транскрипта.

В некоторых вариантах осуществления композиции содержат векторную систему. В некоторых вариантах осуществления векторная система может содержать один вектор.

В других вариантах осуществления векторная система может содержать два вектора. В дополнительных вариантах осуществления векторная система может содержать три вектора. Когда для мультиплексирования используют разные гидовые РНК или когда используют несколько копий гидовой РНК, векторная система может содержать более трех векторов. В некоторых вариантах осуществления векторная система может дополнительно содержать описанную в данном документе донорную конструкцию. В некоторых вариантах осуществления векторная система может дополнительно содержать нуклеиновые кислоты, которые кодируют нуклеазу. В некоторых вариантах осуществления векторная система может дополнительно содержать нуклеиновые кислоты, которые кодируют гидовые 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 донорные 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 донорной конструкции (например, двунаправленной конструкции). В некоторых вариантах осуществления векторная система содержит одну или более нуклеотидных последовательностей, кодирующих сРНК, тРНК или сРНК и тРНК. В некоторых вариантах осуществления векторная система содержит одну или более нуклеотидных последовательностей, кодирующих оРНК, и мРНК, кодирующую 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 сРНК, тРНК и мРНК, 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 сРНК, тРНК или сРНК и тРНК (которая может представлять собой оРНК), содержит или состоит из направляющей последовательности, фланкируемой всей или частью повторяющейся последовательности из CRISPR/Cas природного происхождения. Векторная система может содержать нуклеиновую кислоту, содержащую сРНК, тРНК или сРНК и тРНК, причем векторная система содержит или состоит из нуклеиновых кислот, которые в природе не встречаются вместе с сРНК, тРНК или сРНК и тРНК.

В некоторых вариантах осуществления векторная система может содержать индуцибельные промоторы, чтобы экспрессия начиналась только после ее доставки в целевую клетку. Неограничивающие типовые индуцибельные промоторы включают те,

которые индуцируются тепловым шоком, светом, химическими веществами, пептидами, металлами, стероидами, антибиотиками или спиртом. В некоторых вариантах осуществления индуцибельный промотор может представлять собой промотор, который имеет низкий базовый (неиндуцированный) уровень экспрессии, такой как, например, промотор Tet-On<sup>®</sup> (Clontech).

В дополнительных вариантах осуществления векторная система может содержать тканеспецифические промоторы, чтобы экспрессия начиналась только после ее доставки в конкретную ткань.

Вектор или векторную систему можно доставлять с помощью липосомы, наночастицы, экзосомы или микровезикулы. Вектор также можно доставлять с помощью липидной наночастицы (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 и составы ЛНЧ подходят для доставки гидовых РНК, Cas-нуклеазы (или мРНК, кодирующей Cas-нуклеазу), их комбинаций и/или конструкции, содержащей гетерологичный ген. 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, кодирующую Cas-нуклеазу. В некоторых случаях липидный компонент включает биоразлагаемый ионизируемый липид, холестерин, ДСФХ и ПЭГ-ДМГ.

Будет очевидно, что описанные в данном документе гидовая РНК, РНК-направляемый ДНК-связывающий агент (например, Cas-нуклеаза или нуклеиновая кислота, кодирующая Cas-нуклеазу) и донорная конструкция могут быть доставлены с использованием одной или разных систем. Например, гидовую РНК, Cas-нуклеазу и конструкцию может нести один вектор (например, AAV). В альтернативном варианте Cas-нуклеазу (в виде белка или мРНК) и/или гРНК могут нести плазида или ЛНЧ, тогда как конструкцию может нести вектор. Кроме того, разные системы доставки можно вводить одним или разными путями.

В некоторых вариантах осуществления способ включает введение гидовой РНК и РНК-направляемого ДНК-связывающего агента (такого как мРНК, кодирующая Cas9-нуклеазу) в ЛНЧ. В дополнительных вариантах осуществления способ включает введение конструкции нуклеиновой кислоты AAV, кодирующей трансгенный белок, такой как двунаправленная конструкция. CRISPR/Cas9 ЛНЧ, содержащую гидовую РНК и мРНК, кодирующую Cas9, можно вводить внутривенно. Донорную конструкцию AAV можно вводить внутривенно.

Разные системы доставки можно доставлять *in vitro* или *in vivo* одновременно или в любом последовательном порядке. В некоторых вариантах осуществления донорную конструкцию, гидовую РНК и Cas-нуклеазу 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 донорную 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 Cas-нуклеазы 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 однонедельными интервалами, например, на 1 неделе, 2 неделе и 3 неделе и т. д. В качестве дополнительного примера, гидовую РНК и Cas-нуклеазу в виде вектора и/или связанные 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 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 однонедельными 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

гРНК и связанные с ними способы и композиции, описанные в данном документе, применимы для эффективной вставки гетерологичного (экзогенного) гена в интрон 1 локуса альбумина человека клетки-хозяина. В некоторых вариантах осуществления в настоящем изобретении предложен способ вставки гетерологичного гена в интрон 1 локуса альбумина человека клетки-хозяина, включающий введение в клетку-хозяина (*in vivo* или *in vitro*) описанной в данном документе гидовой РНК (любой из SEQ ID NO: 2-33), РНК-направляемого ДНК-связывающего агента (например, описанной в данном документе Cas-нуклеазы) и донорной конструкции, которая содержит последовательность, кодирующую представляющий интерес гетерологичный полипептид.

гРНК и связанные с ними способы и композиции, описанные в данном документе,

применимы для экспрессии гетерологичного (экзогенного) гена в интроне 1 локуса альбумина человека клетки-хозяина. В некоторых вариантах осуществления в настоящем изобретении предложен способ экспрессии гетерологичного гена в интроне 1 локуса альбумина человека клетки-хозяина, включающий введение в клетку-хозяина (*in vivo* или *in vitro*) описанной в данном документе гидовой РНК (любой из SEQ ID NO: 2-33), РНК-направляемого ДНК-связывающего агента (например, описанной в данном документе Cas-нуклеазы) и донорной конструкции, которая содержит последовательность, кодирующую представляющий интерес гетерологичный полипептид.

гРНК и связанные с ними способы и композиции, описанные в данном документе, применимы для лечения связанных с печенью нарушений у субъекта, как описано в данном документе. В некоторых вариантах осуществления в настоящем изобретении предложен способ лечения связанного с печенью нарушения, включающий введение в клетку-хозяина (*in vivo* или *in vitro*) описанной в данном документе гидовой РНК (любой из SEQ ID NO: 2-33), РНК-направляемого ДНК-связывающего агента (например, описанной в данном документе Cas-нуклеазы) и донорной конструкции, которая содержит последовательность, кодирующую представляющий интерес полипептид.

Композиции и способы по настоящему изобретению можно использовать и применять с рядом клеток-хозяев. В некоторых вариантах осуществления клетка-хозяин представляет собой клетку печени, нейрональную клетку или мышечную клетку. В некоторых вариантах осуществления клетка-хозяин представляет собой любую подходящую неделяющуюся клетку. 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 мыши, яванского макака или человека В некоторых вариантах осуществления клетка-хозяин представляет собой миоцит, такой как миоцит мыши, 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.

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 фактора IX 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 белка FIX 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 уровней фактора IX после одной дозы в течение по меньшей мере 1, 2, 4 или 6 месяцев или по меньшей мере 1, 2, 3, 4 или 5 лет.

В дополнительных вариантах осуществления, включающих вставку в локус альбумина, циркулирующие уровни альбумина индивида являются нормальными. Способ может включать поддержание циркулирующих уровней альбумина индивида в пределах  $\pm 5\%$ ,  $\pm 10\%$ ,  $\pm 15\%$ ,  $\pm 20\%$  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 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) геномного локуса человека, такого как безопасный приемный сайт, например, 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 фактор IX) и РНК-направляемых ДНК-связывающих агентов (например, Cas-нуклеазы). Вставка в геномный локус, такой как безопасный приемный сайт, такой как безопасный приемный сайт локуса альбумина (например, интрон 1), позволяет обеспечивать сверхэкспрессию гена фактора IX без существенного вредного воздействия на клетку-хозяина или популяцию клеток, таких как гепатоциты или клетки печени. В некоторых вариантах осуществления изобретение включает способ или применение модификации (например, создания двухцепочечного разрыва) интрона 1 локуса 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 гидовая РНК содержит направляющую последовательность, которая содержит по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов, которые связываются в пределах интрона 1 локуса альбумина человека (SEQ ID NO: 1). В некоторых вариантах осуществления гидовая РНК содержит по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления гидовая РНК содержит последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, содержащую по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления



описанные в данном документе гидовые РНК содержат направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая составляет по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая выбрана из группы, состоящей из SEQ ID NO: 34-97. В некоторых вариантах осуществления способ осуществляют *in vitro*. В некоторых вариантах осуществления способ осуществляют *in vivo*. В некоторых вариантах осуществления донорная конструкция представляет собой двунаправленную конструкцию, которая содержит последовательность, кодирующую фактор IX. В некоторых вариантах осуществления клетка-хозяин представляет собой клетку печени, такую как. В дополнительных вариантах осуществления клетка печени представляет собой гепатоцит.

В некоторых вариантах осуществления изобретение включает способ или применение внесения нуклеиновой кислоты фактора IX в клетку-хозяина или популяцию клеток-хозяев, включающие введение или доставку 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 гидовая РНК содержит направляющую последовательность, которая содержит по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов, которые способны связываться с областью в пределах интрона 1 локуса альбумина человека (SEQ ID NO: 1). В некоторых вариантах осуществления гидовая РНК содержит по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления гидовая РНК содержит последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, содержащую по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, выбранную из группы, состоящей из

SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая составляет по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая выбрана из группы, состоящей из SEQ ID NO: 34-97. В некоторых вариантах осуществления способ представляет собой способ *in vitro*. В некоторых вариантах осуществления способ представляет собой способ *in vivo*. В некоторых вариантах осуществления донорная конструкция представляет собой двунаправленную конструкцию, которая содержит последовательность, кодирующую фактор IX. В некоторых вариантах осуществления клетка-хозяин представляет собой клетку печени или популяция клеток-хозяев представляет собой клетки печени, такие как гепатоциты.

В некоторых вариантах осуществления изобретение включает способ или применение экспрессии фактора IX в клетке-хозяине или популяции клеток-хозяев, включающие введение или доставку 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 гидовая РНК содержит направляющую последовательность, которая содержит по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов, которые способны связываться с областью в пределах интрона 1 локуса альбумина человека (SEQ ID NO: 1). В некоторых вариантах осуществления гидовая РНК содержит по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления гидовая РНК содержит последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, содержащую по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. В некоторых вариантах осуществления

описанные в данном документе гидовые РНК содержат направляющую последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая составляет по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая выбрана из группы, состоящей из SEQ ID NO: 34-97. В некоторых вариантах осуществления способ представляет собой способ *in vitro*. В некоторых вариантах осуществления способ представляет собой способ *in vivo*. В некоторых вариантах осуществления донорная конструкция представляет собой двунаправленную конструкцию, которая содержит последовательность, кодирующую фактор IX. В некоторых вариантах осуществления клетка-хозяин представляет собой клетку печени или популяция клеток-хозяев представляет собой клетки печени, такие как гепатоциты.

В некоторых вариантах осуществления изобретение включает способ или применение лечения гемофилии (например, гемофилии А или гемофилии В), включающие введение или доставку 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) нуждающемуся в этом субъекту. В некоторых вариантах осуществления гидовая РНК содержит направляющую последовательность, которая содержит по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов, которые способны связываться с областью в пределах интрона 1 локуса альбумина человека (SEQ ID NO: 1). В некоторых вариантах осуществления гидовая РНК содержит по меньшей мере 15, 16, 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления гидовая РНК содержит последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, содержащую по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32, 33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96, and 97. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую

последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая составляет по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33. В некоторых вариантах осуществления описанные в данном документе гидовые РНК содержат направляющую последовательность, которая выбрана из группы, состоящей из SEQ ID NO: 34-97. В некоторых вариантах осуществления донорная конструкция представляет собой двунаправленную конструкцию, которая содержит последовательность, кодирующую гетерологичный полипептид. В некоторых вариантах осуществления клетка-хозяин представляет собой клетку печени или популяция клеток-хозяев представляет собой клетки печени, такие как гепатоциты.

В контексте данного документа термин «гемофилия» относится к нарушению, вызванному отсутствием или дефектом гена или полипептида фактора IX. Гемофилия относится к нарушению, вызванному отсутствием или дефектом гена или полипептида фактора VIII. Нарушение включает патологические состояния, являющиеся наследственными и/или приобретенными (например, вызванными спонтанной мутацией в гене), и включают гемофилию А и гемофилию В. Гемофилию А вызывает дефицит фактора VIII. Гемофилию В вызывает дефицит фактора IX. В некоторых вариантах осуществления дефективный ген или полипептид фактора IX приводит к снижению уровня фактора IX в плазме и/или снижению коагуляционной активности фактора IX. В контексте данного документа гемофилия включает легкую, умеренную и тяжелую гемофилию. Например, людей с менее чем около 1% активного фактора классифицируют как имеющих тяжелую гемофилию, люди с около 1-5% активного фактора имеют умеренную гемофилию, а люди с легкой гемофилией имеют около 5-40% от нормального уровня активного фактора свертывания крови.

В некоторых вариантах осуществления донорная конструкция содержит последовательность, кодирующую фактор IX, причем последовательность фактора IX представляет собой фактор IX дикого типа. В некоторых вариантах осуществления последовательность кодирует вариант фактора IX. Например, вариант может обладать большей коагуляционной активностью, чем фактор IX дикого типа. Например, вариантный фактор IX может содержать одну или более мутаций, таких как аминокислотная замена в позиции R338 (например, R338L), по сравнению с фактором IX дикого типа. В некоторых вариантах осуществления последовательность кодирует вариант фактора IX, который является на 80%, 85%, 90%, 93%, 95%, 97%, 99% идентичным фактору IX дикого типа и обладает по меньшей мере 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100% или более активности по сравнению с фактором IX дикого типа. В некоторых вариантах осуществления последовательность кодирует фрагмент фактора IX, причем фрагмент обладает по меньшей мере 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100% или более активности по сравнению с фактором IX дикого типа.

В некоторых вариантах осуществления донорная конструкция содержит последовательность, кодирующую вариант фактора IX, причем вариант фактора IX активирует коагуляцию в отсутствие своего кофактора, фактора VIII (экспрессия приводит к терапевтически релевантной имитирующей FVIII активности). Такие варианты фактора IX могут дополнительно поддерживать активность фактора IX дикого типа. Например, такой вариант фактора IX может содержать аминокислотную замену в позиции L6, V181, K265, I383, E185 или их комбинацию по сравнению с фактором IX дикого типа. Например, такой вариант фактора IX может содержать мутацию L6F, мутацию V181I, мутацию K265A, мутацию I383V, мутацию E185D или их комбинацию по сравнению с фактором IX дикого типа.

Композиции и способы по настоящему изобретению применимы для эффективной вставки представляющего интерес гетерологичного гена и безопасно экспрессии гетерологичного полипептида (например, терапевтического полипептида). В некоторых вариантах осуществления полипептид представляет собой секретируемый полипептид. В некоторых вариантах осуществления полипептид является таким, чья функция обычно осуществляется (например, который является функционально активным) в виде секретируемого полипептида. В контексте данного документа «секретируемый полипептид» относится к белку, который секретируется 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 секретируется клеткой, включая растворимые цитозольные полипептиды. Одна или более последовательностей IRES или саморасщепляющегося пептида могут фланкировать внутриклеточный полипептид, например, в конце полипептида, таком как amino-конец полипептида, или вблизи него.

В некоторых вариантах осуществления полипептид представляет собой полипептид дикого типа. В некоторых вариантах осуществления полипептид представляет собой вариантный (например, 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 представляет собой фактор IX или его вариант. 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 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 до введения описанных в данном документе композиций. В дополнительных вариантах осуществления экспрессия гетерологичного полипептида может быть повышена до по меньшей мере выявляемых уровней или терапевтически эффективных уровней.

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 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. Способы выявления и измерения уровней полипептидов в различных образцах хорошо известны в данной области техники.

В некоторых вариантах осуществления композиции и способы по настоящему изобретению применимы для лечения связанного с печенью заболевания. В контексте данного документа «связанное с печенью заболевание» относится к заболеваниям, которые могут непосредственно вызывать повреждение ткани печени, заболеваниям в результате повреждения ткани печени и/или заболеваниям отличных от печени органов или тканей, возникшим в результате дефектов в печени. Примеры связанных с печенью заболевания включают, без ограничения, без ограничения, тирозинемия, болезнь Вильсона, болезнь Тея - Сакса, гипербилирубинемия (болезнь Криглера - Найяра), острую интермиттирующую порфирию, цитруллинемию типа 1, прогрессирующий семейный внутрипеченочный холестаз и болезнь кленового сиропа.

Как описано в данном документе, любую одну или более генов РНК, описанных в данном документе, РНК-направляемый ДНК-связывающий агент и донорную конструкцию, содержащую трансген, можно доставлять, используя любые подходящие систему доставки и способ, известные в данной области техники. Композиции можно доставлять *in vitro* или *in vivo* одновременно или в любом последовательном порядке. В некоторых вариантах осуществления донорную конструкцию, геновую 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 донорную 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 качестве дополнительного примера, гидовую РНК и 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 конструкции в виде вектора и/или связанной с ЛНЧ. В некоторых вариантах осуществления гидовая РНК и РНК-направляемый ДНК-связывающий агент связаны с ЛНЧ и доставляются в клетку-хозяина до доставки донорной конструкции.

В некоторых вариантах осуществления донорная конструкция содержит последовательность, кодирующую фактор IX или его варианты. Например, вариант обладает большей активностью, чем полипептид дикого типа. В некоторых вариантах осуществления последовательность кодирует вариант полипептида, который является на 80%, 85%, 90%, 93%, 95%, 97%, 99% идентичным полипептидной последовательности дикого типа и обладает по меньшей мере 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100% или более активности по сравнению с полипептидом дикого типа. В некоторых вариантах осуществления последовательность кодирует фрагмент полипептида дикого типа, причем фрагмент обладает по меньшей мере 80%, 85%, 90%, 92%, 94%, 96%, 98%, 99%, 100% или более активности по сравнению с полипептидом дикого типа.

В некоторых вариантах осуществления одного введения донорной конструкции, содержащей гетерологичный ген, гидовой РНК и РНК-направляемого ДНК-связывающего агента достаточно для повышения экспрессии представляющего интерес полипептида до необходимого уровня. В других вариантах осуществления может быть необходимо более одного введения композиции, содержащей донорную конструкцию, содержащую гетерологичный ген, гидовую РНК и РНК-направляемый ДНК-связывающий агент, для максимизации терапевтических эффектов.

В некоторых вариантах осуществления гидовые РНК, РНК-направляемый ДНК-связывающий агент и донорную конструкцию вводят отдельно или в любой комбинации внутривенно. В некоторых вариантах осуществления гидовые РНК, РНК-направляемый ДНК-связывающий агент и донорную конструкцию 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 грызуна (например, мышь).

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 AAV2, была синтезирована и клонировали в pUC57-Kan 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 <sup>E.coli</sup> Stb13 TM (Thermo Fisher, кат. № C737303).

##### ***Получение AAV***

Тройную трансфекцию клеток HEK293 использовали для упаковки геномов представляющими интерес конструкциями для получения AAV8 и AAV-DJ, 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 9. Выделенный AAV подвергали диализу в буфере для хранения (ФСБ с 0,001% плуроника 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 линейризованную плазмидную ДНК-матрицу и РНК-полимеразу T7. В общем случае плазмидную ДНК, содержащую промотор T7 и область поли-(A/T) из 100 нуклеотидов, линейризовали путем инкубации при 37°C с XbaI для полного расщепления с последующей тепловой инактивацией XbaI при 65 °C. Линейризованную 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 линейризованной плазмиды; 2 mM каждого из ГТФ, АТФ, ЦТФ и N1-метил псевдо-UTP (Trilink); 10 mM ARCA (Trilink); 5 Е/мкл РНК-полимеразы T7 (NEB); 1 Е/мкл мышиноного ингибитора РНКазы (NEB); 0,004 Е/мкл неорганической пирофосфатазы E. coli (NEB); и 1x реакционного буфера. ДНКазу TURBO (ThermoFisher) добавляли до конечной концентрации 0,01 Е/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 mM цитрате, 100 mM 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°C или -80°C 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 Hera1-6***

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°C 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°C и атмосфере с 5% CO<sub>2</sub>. После инкубации клетки проверяли в отношении образования монослоя, трижды промывали с сохранением гепатоцитов и инкубировали при 37 °C.

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 1е5. 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% плюроник 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 (Abcam, 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 или плазму, как указано. Общие уровни фактора IX человека определяли, используя набор для ELISA для фактора IX человека (Abcam, 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***

BaseScore (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). BaseScore 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 BaseScore 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 для интрона 1 альбумина 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 уровней фактора IX человека, как описано в примере 1.

Клетки Hep1-6 представляют собой иммортализованную линию клеток печени мышей, которая продолжает делиться в культуре. Как проиллюстрировано на Фиг. 2 (72 часа после обработки), вектор (scAAV, полученный из плазмиды P00204), содержащий 200 п. о. плечи гомологии, приводил к обнаруживаемой экспрессии hFIX, например, после вставки в интрон 1 альбумина в делящихся клетках. 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 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 для интрона 1 альбумина 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 C57B1/6 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 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, 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 (мкг/мл)**

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 для интрона 1 альбумина 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 C57B1/6 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 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 интрона 1 альбумина, являющийся мишенью 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 (из плазмид P00353 и P00354, соответственно) для вставки в сайт интрона 1 альбумина, являющийся мишенью 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: hFIX**

<b>Сывороточные уровни AAV</b>	Среднее значение, сывороточный FIX (мкг/мл)	Ст. откл., сывороточный FIX (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 hFIX**

<b>AAV</b>	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 20целевым сайтам в интроне 1 альбумина в первичных гепатоцитах мышей**

Продемонстрировав, что двунаправленные конструкции, в которых отсутствуют плечи гомологии, превосходят по эффективности векторы с другими конфигурациями в отношении вставки в интрон 1 альбумина, в описанном в этом примере эксперименте исследовали эффекты изменения акцепторов сплайсинга. 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  $1 \times 10^5$ . 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. 5C в виде относительных единиц люциферазы («ОЕЛ»). Например, AAV-векторы, содержащие OРС hFIX, содержали пептид HiBit, слитый с их 3' концами, а AAV-вектор, содержащий только репортерные гены, содержал OРС NanoLuc (помимо ЗФБ). Схематические изображения каждого из тестируемых векторов представлены на Фиг. 5A. Протестированные гРНК проиллюстрированы на Фиг. 5B и Фиг. 5C с использованием сокращенных обозначений для тех, которые перечислены в таблице 5 (например, опущены первые нули, например, «G551» соответствует «G000551» в таблице 5).

Как проиллюстрировано на Фиг. 5B 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. 5C. Как проиллюстрировано на Фиг. 5D, существенный уровень образования инделей не обязательно приводит к более эффективной экспрессии трансгенов. Не все направляющие, которые обеспечивали существенное количество инделей, приводили к высоким уровням белков с такой же самой вставочной матрицей по данным измерения относительной активности люциферазы. При применении полученных из P00411 и P00418 матриц значения  $R^2$  составляли 0,54 и 0,37 между инделем и активностью люциферазы, соответственно, при этом направляющие с редактированием менее 10% не включены (Фиг. 5D). Что интересно, несмотря на разные OРС и акцепторы сплайсинга, относительные уровни экспрессии, измеренные в ОЕЛ, были сопоставимыми в трех тестируемых векторах, демонстрируя надежность, воспроизводимость и модульность системы двунаправленной конструкции, например, для применения при вставке представляющих интерес трансгенов в интрон 1 альбумина (смотрите Фиг. 5C и таблицу 17). Акцептор сплайсинга мышинового альбумина и акцептор сплайсинга человеческого FIX приводили к эффективной экспрессии трансгена.

**Таблица 16: % инделей**

ID	P00411	P00418	P00415



направляюще й	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: Экспрессия люциферазы

ID направляюще й	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 интрона 1 альбумина**

SsAAV и ЛНЧ, тестируемые в этом примере, получали и доставляли мышам C57Bl/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 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  $3 \times 10^{11}$  вг/мш и 4 мг/кг (по отношению к общему содержанию РНК-карго), соответственно (n=5 для каждой группы). гРНК, тестируемые в этом эксперименте, проиллюстрированы на Фиг. 6. Как проиллюстрировано на Фиг. 6 и в таблице 18, как наблюдалось in vitro, существенный уровень образования инделей не был прогностическим в отношении вставки или экспрессии трансгенов.

**Таблица 18: Сывороточные уровни hFIX и % инделей**

Направляюща я	Среднее значение,	Ст. откл., индели (%)	Среднее значение,	Ст. откл., люцифераза
------------------	----------------------	--------------------------	----------------------	--------------------------

	индели (%)		люцифераза (ОЕЛ)	(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

В отдельном эксперименте панель из 20 гРНК, нацеленных на 20 разных целевых сайтов, протестированных *in vitro* в примере 5, тестировали *in vivo*. С этой целью составы ЛНЧ, содержащие 20 гРНК, нацеленных на интрон 1 альбумина, доставляли мышам вместе с ssAAV, полученным из P00147. AAV и ЛНЧ доставляли при 3e11 вг/мш и 1 мг/кг (по отношению к общему содержанию РНК-карго), соответственно. гРНК, тестируемые в этом эксперименте, проиллюстрированы на Фиг. 7А и Фиг. 7В.

**Таблица 19. Редактирование 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



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

Как проиллюстрировано на Фиг. 7А и в таблице 19, 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 hFIX, как проиллюстрировано на графике, представленном на Фиг. 7D и в таблице 20. 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, уровни циркулирующего hFIX коррелировали с числом клеток, положительных в отношении гибридного транскрипта. Наконец, в анализе использовали объединенные зонды, которые могут обнаруживать вставку двунаправленной конструкции hFIX 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 интрон 1 альбумина в любой

ориентации, приводя к экспрессии гибридных транскриптов, что коррелирует с экспрессией трансгена на белковом уровне. Эти данные показывают, что достигаемые уровни циркулирующего белка зависят от направляющей, используемой для вставки.

#### **Пример 7 - время доставки AAV и ЛНЧ in vivo**

В этом примере исследовали время между доставкой ssAAV, содержащего двунаправленную конструкцию hFIX, и ЛНЧ в отношении направленной вставки в интрон 1 альбумина.

ssAAV и ЛНЧ, тестируемые в этом примере, получали и доставляли мышам, как описано в примере 1. Состав ЛНЧ содержал G000551, а двунаправленную матрицу доставляли в виде ssAAV, полученного из P00147. AAV и ЛНЧ доставляли при  $3 \times 10^{11}$  вг/мш и 4 мг/кг (по отношению к общему содержанию РНК-карго), соответственно (n=5 для каждой группы). Группа «только матрица» получала только AAV, а группа «ФСБ» не получала AAV или ЛНЧ. Одна группа получала AAV и ЛНЧ последовательно (с интервалом в несколько минут) на 0 сутки («матрица+ЛНЧ сутки 0»); другая группа получала AAV на 0 сутки и ЛНЧ на 1 сутки («матрица+ЛНЧ сутки 1»); и последняя группа получала AAV на 0 сутки и ЛНЧ на 7 сутки («матрица+ЛНЧ сутки 7»). Через 1, 2 и 6 недель собирали плазму для анализа экспрессии hFIX.

Как проиллюстрировано на Фиг. 9, hFIX был обнаружен в каждой группе в каждый анализируемый момент, за исключением временной точки, соответствующей 1 неделе, для группы, которая получала дозу ЛНЧ на те же сутки через 1 неделю после доставки AAV.

#### **Пример 8 - многократное дозирование ЛНЧ после доставки AAV**

В этом примере исследовали эффекты повторного введения ЛНЧ после введения ssAAV для направленной вставки в интрон 1 альбумина.

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 и ЛНЧ доставляли при  $3 \times 10^{11}$  вг/мш и 0,5 мг/кг (по отношению к общему содержанию РНК-карго), соответственно (n=5 для каждой группы). Группа «только матрица» получала только AAV, а группа «ФСБ» не получала AAV или ЛНЧ. Одна группа получала AAV и ЛНЧ последовательно (с интервалом в несколько минут) на 0 сутки без дополнительной обработки («матрица+ЛНЧ (1x)» на Фиг. 10); другая группа получала AAV и ЛНЧ последовательно (с интервалом в несколько минут) на 0 сутки и вторую дозу на 7 сутки («матрица+ЛНЧ (2x)» на Фиг. 10); и последняя группа получала AAV и ЛНЧ последовательно (с интервалом в несколько минут) на 0 сутки, вторую дозу ЛНЧ на 7 сутки и третью дозу ЛНЧ на 14 сутки («матрица+ЛНЧ (3x)» на Фиг. 10). Через 1, 2, 4 и 6 недель после введения AAV собирали плазму для анализа экспрессии hFIX.

Как проиллюстрировано на Фиг. 10, hFIX был обнаружен в каждой группе в каждый анализируемый момент времени, а несколько последующих доз ЛНЧ не приводили к существенному повышению уровня экспрессии hFIX.

#### **Пример 9 - продолжительность экспрессии hFIX in vivo**

В этом примере у обработанных животных оценивали продолжительность экспрессии hFIX по времени после направленной вставки в интрон 1 альбумина. С этой целью измеряли hFIX в сыворотке обработанных животных в рамках однолетнего исследования продолжительности.

ssAAV и ЛНЧ, тестируемые в этом примере, получали и доставляли мышам C57B1/6, как описано в примере 1. Состав ЛНЧ содержал G000551, а ssAAV был получен из P00147. AAV доставляли при  $3 \times 10^{11}$  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. 11А и Фиг. 11В и в таблицах 21-22, экспрессия hFIX из интрона 1 альбумина сохранялась в каждый момент времени, оцениваемый для обеих групп, до 12 недель. Считается, что снижение уровней, наблюдаемое через 8 недель, 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 FIX**

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 10 - 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 после направленной вставки в интрон 1 альбумина оценивали на мышах C57B1/6.

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 и собирали сыворотку для анализа экспрессии hFIX.

Как проиллюстрировано на Фиг. 12А (1 неделя), Фиг. 12В (2 недели) и таблице 23, изменение дозы AAV или ЛНЧ может модулировать степень экспрессии hFIX из интрона 1 альбумина 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
		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	-
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



Момент времени	Доза РНП (мг/кг)	Доза AAV (МЗ)	Среднее значение, FIX (мкг/мл)	СО	N
		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 11 - 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 плазмиды P00415), которая отображена на графиках на Фиг. 13В и Фиг. 14В в виде относительных единиц люциферазы («ОЕЛ»). Данные ОЕЛ, представленные на Фиг. 13В и Фиг. 14В графически, численно отображены в таблице 3 и таблице 4 ниже. Например, AAV-векторы содержали OPC NanoLuc (в дополнение к ЗФБ). Схематические изображения тестируемых векторов представлены на Фиг. 13В и Фиг. 14В. Протестированные гРНК проиллюстрированы на каждой из фигур с использованием сокращенного номера для тех, которые перечислены в таблице 1 и Таблице 3.

Как проиллюстрировано на Фиг. 13А для ПГЯ и на Фиг. 14А 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. 13А и в таблице 3 из-за неудачных экспериментов с определенными парами праймеров, используемых для секвенирования ампликона). Данные по редактированию, представленные на Фиг. 13А и Фиг. 14А графически, численно отображены в таблице 3 и таблице 4 ниже. Однако, как

проиллюстрировано на Фиг. 13В, Фиг. 13С, Фиг. 14В и Фиг. 14С, 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. 13С, составляет 0,13, а значение  $R^2$  на Фиг. 14D составляет 0,22.

**Таблица 3: Данные по редактированию и экспрессии трансгена интрона 1 альбумина для оГРНК, доставленных в первичные гепатоциты яванского макака**

<b>ID направляющей</b>	<b>Средн. % редактирования</b>	<b>Ст. откл. % редактирования</b>	<b>Средн. ОЕЛ</b>	<b>Ст. откл. ОЕЛ</b>
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 4: Данные по редактированию и экспрессии трансгена интрона 1 альбумина для оГРНК, доставленных в первичные гепатоциты человека**

<b>ID направляющей</b>	<b>Средн. % редактирования</b>	<b>Ст. откл. % редактирования</b>	<b>Средн. ОЕЛ</b>	<b>Ст. откл. ОЕЛ</b>
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. 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. 14D и приведена в таблице 24 в виде относительных единиц люциферазы («ОЕЛ»). Например, AAV-векторы содержали OPC NanoLuc (в дополнение к ЗФБ). Схематические изображения тестируемых векторов представлены на Фиг. 13B и Фиг. 14B. Протестированные гРНК проиллюстрированы на Фиг. 14D с использованием сокращенного номера для тех, которые перечислены в таблице 1 и таблице 7.

**Таблица 24: Данные по экспрессии трансгена интрона 1 альбумина для гРНК, доставленных в первичные гепатоциты яванского макака**

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 12 - in vivo тестирование вставки гена фактора 9 человека у отличных от человека приматов**

В этом примере было проведено 8-недельное исследование для оценки вставки гена фактора 9 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 5. После обработки AAV животные получали буфер или ЛНЧ, как описано в таблице 5, 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. № AHIX-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 rAb 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 6 и на Фиг. 15. 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. 15, 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  $390 \pm 28$  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 5: 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 6: 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 13 - in vivo тестирование вставки фактора 9 у отличных от человека приматов**

В этом примере проводили исследование для оценки вставки гена фактора 9 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 12. Уровни белка FIX человека, в > 3 раз превышающие уровни, достигнутые в эксперименте примера 12, 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 показывают, 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 FIX у человека измеряли по завершении исследования тем же методом, а данные представлены в таблице 25.

**Таблица 25: Сывороточные уровни белка фактора IX человека - метод ELISA из примера 13**

	Сутки	Ст.	Сутки	Ст.	Сутки	Ст.	Сутки	Ст.	Сутки	Ст.
	7	Откл.	14	Откл.	28	Откл.	42	Откл.	56	Откл.
	FIX		FIX,		FIX,		FIX,		FIX,	
	нг/мл		нг/мл		нг/мл		нг/мл		нг/мл	
300			2562,		3011,					
1	2532,8	145,6	6	99,0	7	62,7	2936,7	72,4	2748,5	86,0
300			2958,		3350,					
2	2211,4	95,8	5	119,2	2	98,4	3049,7	112,7	3036,7	90,6
300			4433,		3367,					
3	3195,1	475,6	9	238,7	2	157,7	3746,1	95,6	3925,0	157,4

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/м. 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 26.

**Таблица 26: Сывороточные уровни белка фактора IX человека**

	Средняя рассч. конц. (мкг/мл)
--	-------------------------------



	3001		3002		3003	
Момент времени	конц.	CV между анализами	конц.	CV между анализами	конц.	CV между анализами
Сутки 7	7,85	20%	5,63	14%	11,20	26%
Сутки 14	8,65	15%	11,06	18%	14,70	28%
Сутки 28	9,14	7%	14,12	7%	10,85	25%
Сутки 42	9,03	10%	33,12*	0%	13,22	13%
Сутки 56	10,24	13%	16,72	12%	33,84*	4%

#### Пример 14 - 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 гРНК, 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 27. Анализ позволил идентифицировать потенциальные нецелевые сайты для тестируемых гРНК.

**Таблица 27: 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	UGGGCAAGGGAAGAAAAAA (SEQ ID NO: 30)	5443
G009857	Альбумин	AUUUAUGAGAUCAACAGCAC (SEQ ID NO: 5)	131
G009859	Альбумин	UUAAUAAAGCAUAGUGCAA (SEQ ID NO: 7)	91

G009860	Альбумин	UAAAGCAUAGUGCAAUGGAU (SEQ ID NO: 8)	133
G012762	Альбумин	UGAUUCCUACAGAAAACUC (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 этим методом, подтверждают, используя нацеленное секвенирование идентифицированных потенциальных нецелевых сайтов.

**Пример 15 - 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 прорыва 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 28. Кроме того, в 5'-конец пептида можно добавлять остатки (GSG) для повышения эффективности расщепления, как показано в таблице 12.

<b>Таблица 28: Саморасщепляющиеся пептиды для применения в конструкциях</b>	
<b>Пептид</b>	<b>Аминокислотная последовательность</b>
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

Пример 16. Использование гуманизированных в отношении альбумина мышей для скрининга гидовых РНК в отношении вставки F9 человека *in vivo*

Нашей целью было идентифицировать эффективные гидовые РНК для вставки hF9 в локус альбумина человека. С этой целью мы использовали мышей, у которых локус мышинового альбумина был заменен соответствующей геномной последовательностью человеческого альбумина, включая первый интрон (мышь ALB<sup>hu/hu</sup>). Это позволило протестировать эффективность вставки гидовых РНК, нацеленных на первый интрон альбумина человека, в контексте зрелой печени *in vivo*. Проводили два отдельных эксперимента на мышах, используя мышей ALB<sup>hu/hu</sup>, для скрининга всего 11 гидовых РНК, каждая из которых была нацелена на первый интрон локуса альбумина человека. Всех мышей взвешивали и проводили инъекцию через хвостовую вену на 0 сутки эксперимента. Кровь собирали на 1, 3, 4 и 6 неделях путем взятия крови из хвоста и отделяли плазму. Мышей умерщвляли на 7 неделе. Кровь собирали через полую вену и отделяли плазму. Также вырезали печень и селезенку.

В первом эксперименте получали и тестировали, как в примере 1, 6 ЛНЧ, содержащих мРНК Cas9, и следующие направляющие: G009852, G009859, G009860, G009864, G009874 и G012764. ЛНЧ разводили до 0,3 мг/кг (используя среднюю массу 30 граммов) и инъекцировали вместе с AAV8, упакованным двунаправленной вставочной матрицей hF9 в дозе 3E11 вирусных геномов на мыш. В каждой группе инъекцию проводили пяти самцам мышей ALB<sup>hu/hu</sup> возрастом 12-14 недель. Пяти мышам из той же группы инъекцировали AAV8, упакованный промотором CAGG, функционально связанным с hF9, что приводит к эписомальной экспрессии hF9 (при 3E11 вирусных геномов на мыш). Было три группы отрицательного контроля с тремя мышами в группе, которым вводили только буфер, AAV8, упакованный только двунаправленной вставочной матрицей hF9, или только ЛНЧ-G009874.

В этом эксперименте получали и тестировали, как в примере 1, следующие ЛНЧ, содержащие мРНК Cas9, и следующие направляющие: G009860, G012764, G009844, G009857, G012752, G012753 и G012761. Их все разводили до 0,3 мг/кг (используя среднюю массу 40 граммов) и инъекцировали вместе с AAV8, упакованным двунаправленной вставочной матрицей hF9 в дозе 3E11 вирусных геномов на мыш. В каждой группе инъекцию проводили пяти самцам мышей ALB<sup>hu/hu</sup> возрастом 30 недель. Пяти мышам из той же группы инъекцировали AAV8, упакованный промотором CAGG, функционально связанным с hF9, что приводит к эписомальной экспрессии hF9 (при 3E11 вирусных геномов на мыш). Было три группы отрицательного контроля с тремя мышами в группе, которым вводили только буфер, AAV8, упакованный только двунаправленной вставочной матрицей hF9, или только ЛНЧ-G009874.

Для анализа проводили ELISA, чтобы измерить у мышей уровни циркулирующего hFIX в каждый момент времени. С этой целью использовали наборы для ELISA для фактора IX человека (ab188393), а все планшеты исследовали с объединенной человеческой нормальной плазмой от George King Bio-Medical в качестве положительного аналитического контроля. Уровни экспрессии фактора IX человека в образцах плазмы в каждой группе на 6 неделе после инъекции проиллюстрированы на **Фиг. 16А** и **Фиг. 16В**. В соответствии с данными по *in vitro* вставке, при применении гидовой РНК G009852 были обнаружены низкие или отсутствующие сывороточные уровни фактора IX. В соответствии с отсутствием прилегающей последовательности РАМ в человеческом альбумине сывороточные уровни фактора IX не были обнаружены при применении гидовой РНК G009864. Экспрессию фактора IX в сыворотке наблюдали в группах, в которых применяли гидовые РНК G009859, G009860, G009874 и G0012764.

Селезенки и часть левой боковой доли всей печени были представлены для анализа методом секвенирования нового поколения (NGS). NGS использовали для оценки процента клеток печени со вставками/делециями (инделями) в гуманизованном локусе альбумина на 7 неделе после инъекции донора AAV-hF9 и ЛНЧ-CRISPR/Cas9. В соответствии с отсутствием прилегающей последовательности РАМ в человеческом альбумине не было обнаружено редактирования в печени при применении гидовой РНК G009864. Редактирование в печени наблюдали в группах, в которых применяли гидовые РНК G009859, G009860, G009874 и G012764 (данные не приведены).

Оставшуюся печень фиксировали в течение 24 часов в 10% нейтральном забуференном формалине, а затем переносили в 70% этанол. От четырех до пяти образцов из отдельных долей вырезали и отправили в HistoWisz, где их обработали и заключили в парафиновые блоки. Затем из каждого парафинового блока вырезали срезы размером пять микронов, и проводили BASESCOPE™ на Ventana Ultra Discovery (Roche), используя универсальную процедуру BASESCOPE™ и реагенты от Advanced Cell Diagnostics, а также специально разработанный зонд, нацеленный на уникальное соединение мРНК, образуемое между сигнальной последовательностью человеческого альбумина из первого интрона локуса альбумина ALB<sup>hu/hu</sup> и транскриптом hF9 при успешной интеграции и транскрипции. Затем использовали программное обеспечение для визуализации HALO (Indica Labs) для количественной оценки процента положительных клеток в каждом образце. Средний процент положительных клеток во нескольких долях для каждого животного коррелировал с уровнями hFIX в сыворотке на 7 неделе. Результаты приведены на **Фиг. 17** и в **таблице 29**. Сывороточные уровни на 7 неделе и % положительных клеток для мРНК hALB-hFIX сильно коррелируют ( $r=0,89$ ;  $R^2=0,79$ ).

**Таблица 29. Данные по hFIX и BASESCOPE™ на 7 неделе.**

Мышь	Направляющая	hFIX (мкг/мл) (7 неделя)	% мРНК зонда (4-5 срезов)	СТД % мРНК зонда	Общее число клеток
ь	я				



TCATTTTAGTCTGTCTTCTTGGTTGCTGTTGATAGACACTAAAAGAGTATTAGATATT  
 ATCTAAGTTTGAATATAAGGCTATAAATATTTAATAATTTTTAAAATAGTATTCTTGG  
 TAATTGAATTATTCTTCTGTTTAAAGGCAGAAGAAATAATTGAACATCATCCTGAGT  
 TTTTCTGTAGGAATCAGAGCCCAATATTTTGAAACAAATGCATAATCTAAGTCAAAT  
 GGAAAGAAATATAAAAAGTAACATTATTACTTCTTGTTTTCTTCAGTATTTAACAAT  
 CCTTTTTTTCTTCCCTTGCCCAG

Таблица 7. Гидовая РНК мышиноного альбумина

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	UUAUAUUAUUGAUUAUUU U	xp5:90461174-90461194	111
G011726	GCACAGAUUAAAACACUAAA	xp5:90461480-90461500	112
G011727	CACAGAUUAAAACACUAAAC	xp5:90461481-90461501	113
G011728	GGUUUUAAAAUAAUAAUG U	xp5:90461502-90461522	114
G011729	UCAGAUUUUCCUGUAACGAU	xp5:90461572-90461592	115
G011730	CAGAUUUUCCUGUAACGAUC	xp5:90461573-90461593	116
G011731	CAAUGGUAAAUAAGAAAUA A	xp5:90461408-90461428	117
G013018	GGAAAUCUGAAGGUGGCAA	xp5:90461563-90461583	118
G013019	GGCGAUCUCACUCUUGUCUG	xp5:90461717-90461737	119

Таблица 8. огРНК мышиноного альбумина и профиль модификации

ID направ- ляющей	Полная последовательность	SEQ ID NO:	Модифицированная полная последовательность	SEQ ID NO:
G000551	AUUUGCAUCUGAGAACC CUUGUUUUAGAGCUAGA AAUAGCAAGUAAAAUA AGGCUAGUCCGUUAUCA ACUUGAAAAAGUGGCAC CGAGUCGGUGCUUUU	120	mA*mU*mU*UGCAUCUGAG AACCCUUGUUUUAGAmGm CmUmAmGmAmAmUmA mGmCAAGUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	142
G000552	AUCGGGAACUGGCAUCU UCA GUUUUAGAGCUAGAAAU AGC AAGUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	121	mA*mU*mC*GGGAACUGGC AUCUUCAGUUUUAGAmGm CmUmAmGmAmAmUmA mGmCAAGUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	143
G000553	GUUACAGGAAAAUCUGA AGG GUUUUAGAGCUAGAAAU AGC AAGUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	122	mG*mU*mU*ACAGGAAAAU CUGAAGGGUUUUAGAmGm CmUmAmGmAmAmUmA mGmCAAGUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	144
G000554	GAUCGGGAACUGGCAUC UUC	123	mG*mA*mU*CGGGAACUGG CAUCUUCGUUUUAGAmGm	145



	<p>GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>		<p>CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	
G000555	<p>UGCAUCUGAGAACCCUU AGG GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>	124	<p>mU*mG*mC*AUCUGAGAAC CCUUAGGGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	146
G000666	<p>CACUCUUGUCUGUGGAA ACA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>	125	<p>mC*mA*mC*UCUUGUCUGU GGAAACAGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	147
G000667	<p>AUCGUUACAGGAAAAUC UGA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA</p>	126	<p>mA*mU*mC*GUUACAGGAA AAUCUGAGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG</p>	148

	AGU GGCACCGAGUCGGUGCU UUU		mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	
G000668	GCAUCUUCAGGGAGUAG CUU GUUUUAGAGCUAGAAAU AGC AAGUUAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	127	mG*mC*mA*UCUUCAGGGA GUAGCUUGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	149
G000669	CAAUCUUUAAAUAUGUU GUG GUUUUAGAGCUAGAAAU AGC AAGUUAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	128	mC*mA*mA*UCUUUAAAUA UGUUGUGGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	150
G000670	UCACUCUUGUCUGUGGA AAC GUUUUAGAGCUAGAAAU AGC AAGUUAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	129	mU*mC*mA*CUCUUGUCUG UGGAAACGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	151
G011722	UGCUGUAUUUUUCUAG UAA	130	mU*mG*mC*UUGUAUUUUU CUAGUAAGUUUUAGAmGm	152

	<p>GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>		<p>CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	
G011723	<p>GUAAAUAUUCUACUAAGA CAA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>	131	<p>mG*mU*mA*AAUAUCUACU AAGACAAGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	153
G011724	<p>UUUUUCUAGUAAUGGAA GCC GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>	132	<p>mU*mU*mU*UUCUAGUAAU GGAAGCCGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	154
G011725	<p>UUAUAUUUAUUGAUUAU UUU GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA</p>	133	<p>mU*mU*mA*UAUUAUUGAU AUUUUUUGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG</p>	155

	AGU GGCACCGAGUCGGUGCU UUU		mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	
G011726	GCACAGAUUA AACACU UAA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	134	mG*mC*mA*CAGAUUAAA CACUUAAGUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	156
G011727	CACAGAUUA AACACUU AAC GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	135	mC*mA*mC*AGAUUAAAC ACUUAACGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	157
G011728	GGUUUUAAAAAUAUAUA UGU GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	136	mG*mG*mU*UUUAAAAUA AUA AUGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	158
G011729	UCAGAUUUUCCUGUAAC GAU	137	mU*mC*mA*GAUUUCCUG UAACGAUGUUUUAGAmGm	159

	<p>GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>		<p>CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	
G011730	<p>CAGAUUUUCCUGUAACG AUC GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>	138	<p>mC*mA*mG*AUUUUCCUGU AACGAUCGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	160
G011731	<p>CAAUGGUAAAUAAGAAA UAA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU</p>	139	<p>mC*mA*mA*UGGUAAAUA GAAUAAGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU</p>	161
G013018	<p>GGAAAUCUGAAGGUGG CAA GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA</p>	140	<p>mG*mG*mA*AAAUCUGAAG GUGGCAAGUUUUAGAmGm CmUmAmGmAmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAmAm GmUmGmGmCmAmCmCmG</p>	162

	AGU GGCACCGAGUCGGUGCU UUU		mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	
G013019	GGCGAUCUCACUCUUGUC UG GUUUUAGAGCUAGAAAU AGC AAGUUAAAAUAAGGCUA GUC CGUUAUCAACUUGAAAA AGU GGCACCGAGUCGGUGCU UUU	141	mG*mG*mC*GAUCUCACUC UUGUCUGGUUUUAGAmGm CmUmAmGmAmAmUmA mGmCAAGUUAAAAUAAGG CUAGUCCGUUAUCAmAmC mUmUmGmAmAmAmAm GmUmGmGmCmAmCmCmG mAmGmUmCmGmGmUmGm CmU*mU*mU*mU	163

Таблица 9. Гидовая РНК альбумина яванского макака

ID направля- ющей	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	UGCAUUUGUUUCAAAAUAUU	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	AUUUAUGAGAUCAACAGCAC	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	AAAGUUGAACAAUAGAAAAA	xp5:61199401-61199421	187
G009868	AAUGCAUAAUCUAAGUCAAA	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	UAAUAAAUAUCAACAUCCU	xp5:61199258-61199278	194
G009875	UUGUCAUGUAUUUCUAAAAU	xp5:61199322-61199342	195
G009876	UUUGUCAUGUAUUUCUAAAA	xp5:61199323-61199343	196

**Таблица 10: орПНК альбумина яванского макака и профили модификации**

<b>ID направ- ляющей</b>	<b>Полная последовательность</b>	<b>SEQ ID NO:</b>	<b>Модифицированная полная последовательность</b>	<b>SEQ ID NO:</b>
G009844	GAGCAACCUCACUCUUGU CU GUUUUAGAGCUAGAAUA GC AAGUUA AAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	197	mG*mA*mG*CAACCUCAC UCUUGUCUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	230
G009845	AGCAACCUCACUCUUGUC UG GUUUUAGAGCUAGAAUA GC	198	mA*mG*mC*AACCUCACU CUUGUCUGGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA	231

	AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU		AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G009846	ACCUCACUCUUGUCUGGG GA GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	199	mA*mC*mC*UCACUCUUG UCUGGGGAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAA GUUAAAUAAGGCUAGU CCGUUAUCAmAmCmUmU mGmAmAmAmAmAmGmU mGmGmCmAmCm CmGmAmGmUmCmGmGmU mGmCmU*mU*mU*mU	232
G009847	CCUCACUCUUGUCUGGGG AA GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	200	mC*mC*mU*CACUCUUGU CUGGGGAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGU UAAAUAAGGCUAGUCC GUUAUCAmAmCmUmUmG mAmAmAmAmAmGmUmG mGmCmAmCmCm GmAmGmUmCmGmGmUmG mCmU*mU*mU*mU	233
G009848	CUCACUCUUGUCUGGGGA AG GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU	201	mC*mU*mC*ACUCUUGUC UGGGGAAGGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAA GUUAAAUAAGGCUAGU CCGUUAUCAmAmCmUmU mGmAmAmAmAmAmGmU mGmGmCmAmCm CmGmAmGmUmCmGmGmU	234



	UU		mGmCmU*mU*mU*mU	
G009849	GGGAAGGGGAGAAAAAA AA GUUUUAGAGCUAGAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	202	mG*mG*mG*GAAGGGGAG AAAAAAAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	235
G009850	GGGAAGGGGAGAAAAAA AA GUUUUAGAGCUAGAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	203	mG*mG*mG*AAGGGGAGA AAAAAAAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	236
G009851	AUGCAUUUGUUUCAAAAU AU GUUUUAGAGCUAGAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	204	mA*mU*mG*CAUUUGUUU CAAAUAUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	237
G009852	UGCAUUUGUUUCAAAUA UU GUUUUAGAGCUAGAAUA GC	205	mU*mG*mC*AUUUGUUUC AAAAUAUUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA	238

	AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU		AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G009853	UGAUUCCUACAGAAAAG UC GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU	206	mU*mG*mA*UUCCUACAG AAAAAGUCGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	239
G009854	UACAGAAAAGUCAGGAU AA GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU	207	mU*mA*mC*AGAAAAGU CAGGAUAAGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	240
G009855	UUUCUUCUGCCUUAAAC AG GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU	208	mU*mU*mU*CUUCUGCCU UUAACAGGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	241

	UU			
G009856	UUAUAGUUUUUAUUAUCAA AC GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	209	mU*mU*mA*UAGUUUUUAU AUUCAACGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	242
G009857	AUUUAUGAGAUCAACAGC AC GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	210	mA*mU*mU*UAUGAGAUC AACAGCACGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	243
G009858	GAUCAACAGCACAGGUUU UG GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	211	mG*mA*mU*CAACAGCAC AGGUUUUGGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	244
G009859	UUAAAUAAGCAUAGUGC AA GUUUUAGAGCUAGAAAUA GC	212	mU*mU*mA*AAUAAAGCA UAGUGCAAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA	245

	AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU		AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G009860	UAAAGCAUAGUGCAAUGG AU GUUUUAGAGCUAGAAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU	213	mU*mA*mA*AGCAUAGUG CAAUGGAUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	246
G009861	UAGUGCAAUGGAUAGGUC UU GUUUUAGAGCUAGAAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU	214	mU*mA*mG*UGCAAUGGA UAGGUCUUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	247
G009862	AGUGCAAUGGAUAGGUCU UA GUUUUAGAGCUAGAAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU	215	mA*mG*mU*GCAAUGGAU AGGUCUAGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	248

	UU			
G009863	UUACUUUGCACUUUCCUU AG GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	216	mU*mU*mA*CUUUGCACU UCCUUAGGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	249
G009864	UACUUUGCACUUUCCUUA GU GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	217	mU*mA*mC*UUUGCACUU UCCUUAGUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	250
G009865	UCUGACCUUUUAUUUUAC CU GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	218	mU*mC*mU*GACCUUUUA UUUUACCUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	251
G009866	UACUAAAACUUUAUUUUA CU GUUUUAGAGCUAGAAAUA GC	219	mU*mA*mC*UAAAACUUU AUUUUACUGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA	252

	AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU		AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G009867	AAAGUUGAACAAUAGAAA AA GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU	220	mA*mA*mA*GUUGAACAA UAGAAAAAGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	253
G009868	AAUGCAUAAUCUAAGUCA AA GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU	221	mA*mA*mU*GCAUAAUCU AAGUCAAGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	254
G009869	AUUAUCCUGACUUUUUCU GU GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU	222	mA*mU*mU*AUCCUGACU UUUUCUGUGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	255

	UU			
G009870	UGAAUUAUUCCUCUGUUU AA GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	223	mU*mG*mA*AUUAUCCU CUGUUUAAGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	256
G009871	UAAUUUUCUUUUGCCCAC UA GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	224	mU*mA*mA*UUUUCUUUU GCCACUAGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	257
G009872	AAAAGGUCAGAAUUGUUU AG GUUUUAGAGCUAGAAAUA GC AAGUUAAAAUAAGGCUAG UC CGUUAUCAACUUGAAAAA GU GGCACCGAGUCGGUGCUU UU	225	mA*mA*mA*AGGUCAGAA UUGUUUAGGUUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	258
G009873	AACAUCCUAGGUAAAAUA AA GUUUUAGAGCUAGAAAUA GC	226	mA*mA*mC*AUCCUAGGU AAAAUAAAGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUAAAAUA	259

	AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU		AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	
G009874	UAAUAAAUAUCAACAUC CU GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU	227	mU*mA*mA*UAAAUAUCA ACAUCCUGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	260
G009875	UUGUCAUGUAUUUCUAAA AU GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU UU	228	mU*mU*mG*UCAUGUAUU UCUAAAUGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	261
G009876	UUUGUCAUGUAUUUCUAA AA GUUUUAGAGCUAGAAUA GC AAGUUAAAUAAGGCUAG UC CGUUAUCAACUUGAAAA GU GGCACCGAGUCGGUGCUU	229	mU*mU*mU*GUCAUGUAU UUCUAAAAGUUUAGAm GmCmUmAmGmAmAmAmU mAmGmCAAGUUA AAAUA AGGCUAGUCCGUUAUCA mAmCmUmUmGmAmAmAm AmAmGmUmGmGmCmAmC mCmGmAmGmUmCmGmGm UmGmCmU*mU*mU*mU	262



UU

Таблица 11: Компоненты и последовательности векторов

ID плазмиды	5' ИКП	Акцептор сплайсинга (1-ая ориентация)	Трансген (1-ая ориентация)	Поли-А (1-ая ориентация)	Поли-А (2-ая ориентация)	Трансген (2-ая ориентация)	Акцептор сплайсинга (2-ая ориентация)	3' ИКП
P00147	(SEQ ID NO: 263)	Акцептор сплайсинга альбумина мыши (SEQ ID NO: 264)	Фактор IX человека (R338L) (SEQ ID NO: 265)	SEQ ID NO: 266	SEQ ID NO: 267	Фактор IX человека (R338L) (SEQ ID NO: 268)	Акцептор сплайсинга альбумина мыши (SEQ ID NO: 269)	(SEQ ID NO: 270)
P00411	(SEQ ID NO: 263)	Акцептор сплайсинга фактора IX человека (SEQ ID NO: 271)	Фактор IX человека (R338L)-HiBit (SEQ ID NO: 272)	SEQ ID NO: 266	SEQ ID NO: 267	Фактор IX человека (R338L)-HiBit (SEQ ID NO: 273)	Акцептор сплайсинга фактора IX человека (SEQ ID NO: 274)	(SEQ ID NO: 270)
P00415	(SEQ ID NO: 263)	Акцептор сплайсинга альбумина мыши	Nluc-P2A-GFP (SEQ ID NO: 275)	SEQ ID NO: 266	SEQ ID NO: 267	Nluc-P2A-GFP (SEQ ID NO: 276)	Акцептор сплайсинга альбумина мыши	(SEQ ID NO: 270)

		(SEQ ID NO: 264)					(SEQ ID NO: 269)	
P0041 8	(SEQ ID NO: 263)	Акцептор сплайсинга альбумина мыши (SEQ ID NO: 264)	Фактор IX человека (R338L)-HiBit (SEQ ID NO: 272)	SEQ ID NO: 266	SEQ ID NO: 267	Фактор IX человека (R338L)-HiBit (SEQ ID NO: 273)	Акцептор сплайсинга альбумина мыши (SEQ ID NO: 269)	(SEQ ID NO: 270)

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA  
GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT

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

TAGGTCAGTGAAGAGAAGAACA AAAAGCAGCATATTACAGTTAGTTGTCTTC  
ATCAATCTTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTCCACAG

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

TTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATA  
ATTCAGGTA AATTGGAAGAGTTTGTTC AAGGGAACCTTGAGAGAGAATGTATGGAA  
GAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGA AAAACACTGAAAGAACAAC  
TGAATTTTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTAAA  
TGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATT  
TGAAGGAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCG  
AGCAGTTTTGTAAA AATAGTGCTGATAACAAGGTGGTTTTGCTCCTG TACTGAGGGAT  
ATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGA  
AGAGTTTTCTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATG  
TGGACTATGTA AATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCA  
CCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTC  
AATCCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTA  
TCGTTAATGAAA AATGGATTGTA AACTGCTGCCCACTGTGTTGAAACTGGTGTAAAA  
TTACAGTTGTTCG CAGGTGAACATAATATTGAGGAGACAGAACATACAGAGCAAAAAG  
CGAAATGTGATTCGAATTATTCCTCACCACA AACTACAATGCAGCTATTAATAAGTAC  
AACCATGACATTGCCCTTCTGGAACTGGACGAACCCTTAGTGCTAAACAGCTACGTT  
ACACSTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCSTCAAATTTGGATCT  
GGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACA AAGGGAGATCAGCTTTAGTTCTT  
CAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTC

ACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGT  
 CAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGACCAGTTTCTTAAC  
 TGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATA  
 CCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAGCTCACTTAA

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

CCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGC  
 CTTCCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTTCTAATAAAAATGAGGAAA  
 TTGCATCGCATTGTCTGAGTAGGTGTCAATTCTATTCTGGGGGGTGGGGTGGGGCAGG  
 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGG  
 CTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCC

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

AAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAAATGAATGCAATT  
 GTTGTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATC  
 ACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAAC  
 TCATCAATGTATCTTATCATGTCTG

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

TTAGGTGAGCTTAGTCTTTTCTTTTATCCAATTCACGTAGCGAGAGACCTTCG  
 TATAGATGCCATATTTCCCCTTCATCGCACATTCCTCCCCCAACTTATTATCCCGGT  
 CAAGAAACTTGTTCCTTCGACTTCAGTGACGTGTGGTCCACCTGAATCACCTGGCA  
 TGAGTCGCGACCGCCCTCGTGAAACCCAGCACAAAACATGTTATTGTAATCGTAA  
 ATTTTCGTGGACAGAAGACAGGTCGCTCTATCGACCAACGGGACGCGCAAATATTGC  
 AGAACGAGGGCTGATCGACSTTTGTGGAAGACCCGCCCCCACCCTCACATATCC  
 GCTCCCAAATTTCAAGAAGATATTTGTATATTCTTTATCGGCTATACAAATCGGGGT  
 AACATAGGAGTTAAGTACGAGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGT  
 TGACTTGTTTATAGCGGCATTATAATTGTGATGGGGTATGATCCTGATAACATTCCT  
 TTTCTGTTCAGTATGCTCAGTTTCTTCAATGTTGTGTTTCGCCAGCCACGACCGTAATC  
 TTAACCCCCGTCTCGACACAGTGTGCGGCCGTTACAATCCACTTTTCATTGACTATG  
 GAGCCCCACAAAACGCGTCGACTTTTCCGTTGAGCACCCCTGCCATGGAAATTGG  
 CCAGGTTTAGCGTCCTCGCCCCGACAACCCTAGTAAAGTCATTAATGACTGTGTG  
 GATTGTGTTATATTATCAAGAATCGTTTCGGCTTCAGTAGAGTTAACGTAGTCCACA  
 TCGGGAAAAACTGTCTCGGCCCTTGTCAACTTTGATGTCTGGGACACACTTACCCGA  
 CCGCACGGGAAGGGCACCGCCGGTTCACAGCTCTTTTGATTCTCAGCGAGCCGGTAG  
 CCCTCAGTGCAACTACACACAACSTTTGTTGTGCGGCGGAATTTTTACAGAATTGCTCG  
 SATCGTCCATTTTTAATGTTGCAGGTGACGTCCAACCTCGCAGTTTTTTTCTTCAAAAC  
 CAAAAGGGCACCAACACTCGTAGGAATTTATATCGTCTTACAACCTCCCCCATTC  
 GACATGGATTAGATTCGCATTGGTCCCCATCGACATATTGCTTCCAGAACTCAGTGG  
 TCCGTTCTGTATTCTCAAACACCTCGCGCGCTTCTTCAAAACTGCATTTTTCTCCAT  
 AACTCTCGCTCCAAGTTCCTTGCACGAATTCTTCAAGCTTTCCTGAGTTATACSTT  
 TTAGGCCGGTTAAGTATCTTATTCGCGTTTTTCGTGGTCCAGAAA

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

CTGTGGAAACAGGGAGAGAAAAACCACACAACATATTTAAAGATTGATGAA  
GACAACTAAGTGTAAATATGCTGCTTTTTGTTCTTCTCTTCACTGACCTA

3' ИКП последовательность (SEQ ID NO:270):

AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTC  
ACTGAGGCCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTTCGCCCGGCCTC  
AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

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

GATTATTTGGATTAAAAACAAAGACTTTCTTAAGAGATGTAATAATTTTCATGA  
TGTTTTCTTTTTTGCTAAAACATAAAGAAATTTCTTTTACATTTTCAG

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

TTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATA  
ATTCAGGTAAATTGGAAGAGTTTGTTCAGGGAAACCTTGAGAGAGAATGTATGGAA  
GAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAACAACTGAAAGAACAAC  
TGAATTTTGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAA  
TGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATT  
TGAAGGAAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCG  
AGCAGTTTTGTAAAATAGTGCTGATAACAAGGTGGTTTTGCTCCTGTACTGAGGGAT  
ATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGA  
AGAGTTTCTGTTTACAAAATTCTAAGCTCACCCGTGCTGAGACTGTTTTTCTGATG  
TGGACTATGTAATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCA  
CCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTC  
AATTCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTA  
TCGTTAATGAAAAATGGATTGTAAGTCTGCCCCTGTGTTGAAACTGGTGTAAAA  
TTACAGTTGTTCGAGGTGAACATAATTTGAGGAGACAGAACATACAGAGCAAAAAG  
CGAAATGTGATTCGAATTAATTCCTCACCACAACATCAATGCAGCTATTAATAAGTAC  
AACCATGACATTGCCCTTCTGGAAGTGGACGAACCCTTAGTGCTAAACAGCTACGTT  
ACACCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCT  
GGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTT  
CAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTC  
ACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGT  
CAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAAGGGACCAGTTTCTTAAC  
TGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATA  
CCAAGGTCTCCCGGTATGTCAACTGGATTAAGGAAAAACAAAGCTCACTGTCAGC  
GGATGGAGACTGTTCAAGAAGATCAGCTAA

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

TTAGGAAATCTTCTTAAACAGCCGCCAGCCGCTCACGGTGAGCTTAGTCTTTT  
CTTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATATTTCCCTTCAT  
CGCACATTCCTCCCCCAACTTATTATCCCGGTCAAGAACTTGTTCCCTTCGACTTCA  
GTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCCGCTCGTGAAAC  
CCAGCACAAAACATGTTATTGTAATCGTAAATTTTCGTGGACAGAAGACAGGTTCG

TCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGATCGACCTTTGT  
 GGAAGACCCGCCCCACCCACTCACATATCCGCTCCCAAATTTCAAGAAGATATTTG  
 TATATTCTTTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAGTACGAGTGGCT  
 CGTCCAGCTCCAGGAGGGCTATATCATGGTTGTACTIONTGTATAGCGGCATTATAAT  
 TGTGATGGGGTATGATCCTGATAACATTCCTTTTCTGTTTACAGTATGCTCAGTTTCTTC  
 AATGTTGTGTTTCGCCAGCCACGACCGTAATCTTAACCCCCGTCTCGACACAGTGTGC  
 GGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCACAAAACGCGTCGACTTT  
 TCCGTTGAGCACCACCTGCCATGGAAATTGGCCAGGTTTAGCGTCCTCGCCCCGAC  
 AACCTAGTAAAGTCATTAATGACTGTGTGGATTGTGTTATATTATCAAGAATCGT  
 TTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAACTGTCTCGGCCCTTGT  
 CAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGGCACCGCCGTT  
 CACAGCTCTTTGATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTACACACAACTT  
 TGTTGTCGGCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTTAATGTTGCAGGT  
 GACGTCCAACCTCGCAGTTTTTTTCTTCAAACCAAAGGGCACCAACACTCGTAGGA  
 ATTTATATCGTCTTTACAACCTCCCCCATTCAGACATGGATTAGATTTCGATTGGTCC  
 CCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAAACACCTCGC  
 GCGCTTCTTCAAACCTGCATTTTTCTCCATACACTCTCGTCCAAGTTCCTTGCAC  
 GAATTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATCTTATTCGCG  
 TTTTCGTGGTCCAGAAA

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

CTGAAATGTAAAAGAATAATTCTTTAGTTTTAGCAAAAAAGAAAACATCATG  
 AAAATTTTACATCTCTTAAGAAAGTCTTTGTTTTTAATCCAAATAATC

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

TTTCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATA  
 ATTCAGGTAAATTGGAAGAGTTTGTTCAGGGAACTTGAGAGAGAATGTATGGAA  
 GAAAAGTGTAGTTTTGAAGAAGCAGTATTCACCTTTGGAGGACTTTGTTCGGTACTGG  
 AGGCAAACCGCTGGTTATAATCTCGACCAAGTACTGGAACAGGGCGGGGTAAGTTC  
 CCTCTTTCAGAATTTGGGTGTAAGCGTCACACCAATCCAGCGGATTGTGTTGTCTGG  
 AGAGAACGGACTCAAATTTGACATCCATGTTATCATTCCATATGAAGGTCTCAGTGG  
 AGACCAAATGGGGCAGATCGAGAAGATTTTCAAGGTAGTTTACCCAGTCGACGATC  
 ACCACTTCAAAGTCATTCTCCACTATGGCACACTTGTTATCGACGGAGTAACTCCTA  
 ATATGATTGATTACTTTGGTTCGCCCCGTATGAGGGCATCGCAGTGTGTTGATGGCAAAA  
 AGATCACCGTAACAGGAACGTTGTGGAATGGGAACAAGATAATCGACGAGAGATTG  
 ATAAATCCAGACGGGTCACCTCTGTTTCAAGGTTACAATTAACGGCGTCACAGGATG  
 GAGACTCTGTGAACGAATACTGGCCACAAATTTTTCACTCCTGAAGCAGGCCGGAG  
 ACGTGGAGGAAAACCCAGGGCCCGTGAGCAAGGGCGAGGAGCTGTTACCCGGGGT  
 GGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGT  
 CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGC  
 ACCACCGGCAAGCTGCCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGC  
 GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCC

GCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAA  
 CTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGGCGACACCCTGGTGAACCGCATCG  
 AGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAG  
 TACAАCTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCAT  
 CAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCG  
 ACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAAC  
 CACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCA  
 CATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCT  
 GTACAAGGGAGGAGGAAGCCCGAAGAAGAAGAGAAAGGTCTAA

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

TTACACCTTCCTCTTCTTCTTGGGGCTGCCGCCGCCCTTGTACAGCTCGTCCAT  
 GCCCAGGGTGATGCCGGCGGGCGGTCACGAACTCCAGCAGCACCATGTGGTCCCTCTT  
 CTCGTTGGGGTCCTTGCTCAGGGCGCTCTGGGTGCTCAGGTAGTGGTTGTCGGGCAG  
 CAGCACGGGGCCGTCGCCGATGGGGGTGTTCTGCTGGTAGTGGTCGGCCAGCTGCA  
 CGCTGCCGTCCCTCGATGTTGTGCCTGATCTTGAAGTTCACCTTGATGCCGTTCTTCTG  
 CTTGTCGGCCATGATGTACACGTTGTGGCTGTTGTAGTTGTAАCTCCAGCTTGTGGCCC  
 AGGATGTTGCCGTCCCTCCTTGAAGTCGATGCCCTTCAGCTCGATCCTGTTACCAGG  
 GTGTCGCCCTCGAACTTCACCTCGGCCCTGGTCTTGTAGTTGCCGTCGTCCTTGAAGA  
 AGATGGTCCTCTCCTGCACGTAGCCCTCGGGCATGGCGCTCTTGAAGAAGTCGTGCT  
 GCTTCATGTGGTCGGGGTACCTGCTGAAGCACTGCACGCCGTAGGTTCAGGGTGGTCA  
 CCAGGGTGGGCCAGGGCACGGGCAGCTTGCCGGTGGTGCAGATGAACTTCAGGGTC  
 AGCTTGCCGTAGGTGGCGTCGCCCTCGCCCTCGCCGCTCACGCTGAACTTGTGGCCG  
 TTCACGTGCCGTCCAGCTCCACCAGGATGGGCACCACGCCGGTGAACAGCTCCTCG  
 CCCTTGCTCACGGGGCCGGGGTTCTCCTCCACGTCGCCGGCCTGCTTCAGCAGGCTG  
 AAGTTGGTGGCCAGGATCCTCTCGCACAGCCTCCAGCCGGTCACGCCGTTGATGGTC  
 ACCCTGAACAGCAGGCTGCCGTCGGGGTTGATCAGCCTCTCGTCGATGATCTTGTTG  
 CCGTTCCACAGGGTGCCGGTCACGGTGATCTTCTTGCCGTCGAACACGGCGATGCC  
 TCGTAGGGCCTGCCGAAGTAGTCGATCATGTTGGGGGTCACGCCGTCGATCACCAG  
 GGTGCCGTAGTGCAGGATCACCTTGAAGTGGTGGTCGTCCACGGGGTACACCACCTT  
 GAAAATCTTCTCGATCTGGCCATCTGGTCGCCGCTCAGGCCCTCGTAGGGGATGAT  
 CACGTGGATGTCGATCTTCAGGCCGTTCTCGCCGCTCAGCACGATCCTCTGGATGGG  
 GGTACGCTCACGCCAGGTTCTGGAACAGGCTGCTCACGCCGCCCTGCTCCAGCAC  
 CTGGTCCAGGTTGTAGCCGGCGGTCTGCCTCCAGTCGCCACGAAGTCCTCCAGGGT  
 GAACACGGCCTCCTCGAAGCTGCACTTCTCCTCCATGCACTCCCTCTCCAGGTTGCC  
 CTGCACGAACTCCTCCAGCTTGCCGCTGTTGTACCTCTTGGGCCTGTTACAGGATCTTG  
 TTGGCGTTCTCGTGGTCCAGGAA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
 GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCCTCAGTGAGCGAGCGAGCGCGCA  
 GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTAGATCTCTTAGGTCAGTGAA

GAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATG  
TTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACA  
AAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAA  
GGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAG  
AAGTTTTTGAAAACACTGAAAGAACAACCTGAATTTTGAAGCAGTATGTTGATGGA  
GATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAAT  
TCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAAGTGTGAATTAGATGTA  
ACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAA  
CAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTG  
TGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAACTTCTAAGCT  
CACCCGTGCTGAGACTGTTTTTCTGATGTGGACTATGTAATTTCTACTGAAGCTGA  
AACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGT  
TGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTTTTGAATGG  
TAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAAGTGC  
TGCCCCTGTGTTGAAACTGGTGTAAAATTACAGTTGTTCGCAGGTGAACATAATAT  
TGAGGAGACAGAACATACAGAGCAAAAGCGAAATGTGATTTCGAATTATTCCTCACC  
ACAACACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAAGTGG  
ACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAAT  
ACACGAACATCTTCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGGAAGAGTCT  
TCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACC  
GAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTG  
GCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTT  
ACTGAAGTGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTG  
TGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGA  
TTAAGGAAAAACAAAGCTCACTTAACTCGACTGTGCCTTCTAGTTGCCAGCCATC  
TGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTC  
CTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATT  
CTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCA  
GGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGG  
GGCTCTAGGGGGTATCCCCAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATA  
AAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATA  
AAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTG  
TGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTTAGGTGAGCTTAGTCTTT  
TCTTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATATTTCCCCTTCA  
TCGCACATTCCTCCCCCAACTTATTATCCCGGTCAAGAACTTGTTCCCTTCGACTTC  
AGTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCGCCCTCGTGAAA  
CCCAGCACAAAACATGTTATTGTAAATCGTAAATTTTCGTGGACAGAAGACAGGTGCG  
CTCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGATCGACCTTTG  
TGGAAGACCCGCCCCCACTCACATATCCGCTCCCAAATTTCAAGAAGATATTT  
GTATATTCTTTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAGTACGAGTGGC

TCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTA CT TGT TTTATAGCGGCATTATAA  
 TTGTGATGGGGTATGATCCTGATAACATTCCTTTTCTGTT CAGTATGCTCAGTTTCTT  
 CAATGTTGTGTT CGCCAGCCACGACCGTAATCTTAACCCCGTCTCGACACAGTGTG  
 CGGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCCACA AAAACGCGTCGACTT  
 TTCCGTTGAGCACCACCTGCCATGGAAATTGGCCAGGTTTAGCGTCTCTCGCCCCGA  
 CAACCCTAGTAAAGTCATTA AATGACTGTGTGGATTGTGTTATATTATCAAGAATCG  
 TTTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAAACTGTCTCGGCCCTTG  
 TCAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGGCACC GCCGGTT  
 CACAGCTCTTTTGATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTACACACA ACTT  
 TGTTGTCGGCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTTAATGTTGCAGGT  
 GACGTCCA ACTCGCAGTTTTTTTCTTCAA AACCAAAGGGCACCAACTCGTAGGA  
 ATTTATATCGTCTTTACA ACTCCCCCATTCAGACATGGATTAGATTTCGCATTGGTCC  
 CCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAAACACCTCGC  
 GCGCTTCTTCAA AACTGCATTTTTCTCCATACACTCTCGCTCCAAGTTCCCTTGAC  
 GAATTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATCTTATTCGCG  
 TTTTCGTGGTCCAGAAAAACTGTGGAAACAGGGAGAGAAAAACCACACAACATATT  
 TAAAGATTGATGAAGACA ACTA ACTGTAATATGCTGCTTTTTGTTCTTCTTCACTG  
 ACCTAAGAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCT  
 CGCTCGCTCACTGAGGCCGCCGGGCAAAGCCCGGGCGTCTGGGCGACCTTTGGTCTG  
 CCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
 GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCCTCAGTGAGCGAGCGAGCGCGCA  
 GAGAGGGAGTGGCCA ACTCCATCACTAGGGGTTCCCTAGATCTCTGATTATTTGGATT  
 AAAAACA AAGACTTTCTTAAGAGATGTA AAATTTTCATGATGTTTTCTTTTTTGCTAA  
 AACTAAAGAATTATTCTTTTACATTT CAGTTTTTCTTGATCATGAAAACGCCAACA  
 AATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAAG  
 GGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGA  
 AGTTTTTGAAAACACTGAAAGAACA ACTGAATTTTGGAAGCAGTATGTTGATGGAG  
 ATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATT  
 CCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGA ACTGTGAATTAGATGTAA  
 CATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAAC  
 AAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCCTGT  
 GAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAACTTCTAAGCTC  
 ACCCGTGCTGAGACTGTTTTTCTGATGTGGACTATGTAAATTCTACTGAAGCTGAA  
 ACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTT  
 GTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTTTTGAATGGT  
 AAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTA ACTGCT  
 GCCCACTGTGTTGAAACTGGTGTAAAATTACAGTTGTCTCGCAGGTGAACATAATATT  
 GAGGAGACAGAACATACAGAGCAA AAGCGAAATGTGATT CGAATTATTCTCACCA



CAACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAAGTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTTA CTGAAGTGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTCTCCCGGTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTGTCAGCGGATGGAGACTGTTCAAGAAGATCAGCTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAAATGAATGCAATTGTTGTTGTTA ACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCA CAAATAAAGCATTTTTTTTCCTGCACTTCTAGTTGTGGTTTTGTCCAACTCATCAATGT ATCTTATCATGTCTGTTAGGAAATCTTCTTAAACAGCCGCCAGCCGCTCACGGTGAGCTTAGTCTTTTCTTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATATTTCCCCTTCATCGCACATTCCCTCCCCCAACTTATTATCCCGGTCAAGAACTTGTTCTTCGACTTCAGTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCCGCCTCGTGAAACCCAGCACAAAACATGTTATTGTAAATCGTAAATTTTCGTGGACAGAA GACAGGTCGCTCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGATCGACCTTTGTGGAAGACCCGCCCCACCCACTCACATATCCGCTCCCAAATTTCAAG AAGATATTTGTATATTCTTTATCGGCTATACAAATCGGGGTAAACATAGGAGTTAAGTACGAGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTACTTGTTTATAGCGGCATTATAATTGTGATGGGGTATGATCCTGATAACATTCCTTTTTCTGTTCAGTATGCTCAGTTTCTTCAATGTTGTGTTGCGCCAGCCACGACCGTAATCTTAACCCCCGTCTCGACACAGTGTGCGGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCACAAAACGCGTCGACTTTTCCGTTGAGCACCACTGCCATGGAAATTGGCCAGGTTTAGCGTCCTCGCCCCGACAACCCTAGTAAAGTCATTAAATGACTGTGTGGATTGTGTTATATTATCAAGAATCGTTTTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAACTGTCTC GGCCCTTGTCAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGGCACCGCCGGTTCACAGCTCTTTTGATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTACACAACTTTGTTGTGCGGCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTTAATGTTGCAGGTGACGTCCAACCTCGCAGTTTTTTTCTTCAAACCAAAGGGCACCAACACTCGTAGGAATTTATATCGTCTTACAACCTCCCCCATTCAGACATGGATTAGATTCGCATTGGTCCCCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAAA CACCTCGCGCGCTTCTTCAAACCTGCATTTTTCTCCATACACTCTCGCTCCAAGTTCCTTGCACGAATTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATCT

TATTCGCGTTTTTCGTGGTCCAGAAAACTGAAATGTAAAAGAATAATTCTTTAGTTT  
TAGCAAAAAAGAAAACATCATGAAAATTTTACATCTCTTAAGAAAGTCTTTGTTTTT  
AATCCAAATAATCAGAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTC  
TGC GCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCCGGGCGTCCGGGCGACC  
TTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
GGTCGCCCCGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA  
GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCCTAGATCTCTTAGGTCAGTGAA  
GAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATG  
TTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACA  
AAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTCAA  
GGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCAGTATT  
CACTTTGGAGGACTTTGTGCGGTGACTGGAGGCAAACCGCTGGTTATAATCTCGACCA  
AGTACTGGAACAGGGCGGGGTAAAGTTCCTCTTTCAGAATTTGGGTGTAAGCGTCAC  
ACCAATCCAGCGGATTGTGTTGTCTGGAGAGAACGGACTCAAATTGACATCCATGT  
TATCATTCCATATGAAGGTCTCAGTGGAGACCAAATGGGGCAGATCGAGAAGATTT  
TCAAGGTAGTTTACCCAGTCGACGATCACCCTTCAAAGTCATTCTCCACTATGGCA  
CACTTGTATCGACGGAGTAACTCCTAATATGATTGATTACTTTGGTCGCCCCGTATG  
AGGGCATCGCAGTGTTTGATGGCAAAAAGATCACCGTAACAGGAACGTTGTGGAAT  
GGGAACAAGATAATCGACGAGAGATTGATAAATCCAGACGGGTCCTCCTGTTTCAG  
GGTTACAATTAACGGCGTCACAGGATGGAGACTCTGTGAACGAATACTGGCCACAA  
ATTTTTCACTCCTGAAGCAGGCCGGAGACGTGGAGGAAAACCCAGGGCCCCGTGAGC  
AAGGGCGAGGAGCTGTTACCCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGA  
CGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACG  
GCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTGGCCCA  
CCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACA  
TGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGC  
ACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGA  
GGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACG  
GCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATC  
ATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACAT  
CGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCG  
ACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCA  
AAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCC  
GGGATCACTCTCGGCATGGACGAGCTGTACAAGGGAGGAGGAAGCCCGAAGAAGA  
AGAGAAAGGTCTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCC  
CTCCCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCTAATAA  
AATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGG  
GTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGG

ATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGG  
TATCCCCAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCA  
ATTGTTGTTGTTAACTTGTATTATGCAGCTTATAATGGTTACAAATAAAGCAATAGC  
ATCACAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCA  
AACTCATCAATGTATCTTATCATGTCTGTTACACCTTCTCTTCTTGGGGCTGCC  
GCCGCCCTTGACAGCTCGTCCATGCCCAGGGTGATGCCGGCGGGCGGTCACGAACTC  
CAGCAGCACCATGTGGTCCCTCTTCTCGTTGGGGTCCCTTGCTCAGGGCGCTCTGGGT  
GCTCAGGTAGTGGTTGTCTGGGCAGCAGCACGGGGCCGTCGCCGATGGGGGTGTTCT  
GCTGGTAGTGGTCTGGCCAGCTGCACGCTGCCGTCCTCGATGTTGTGCCTGATCTTGA  
AGTTCACCTTGATGCCGTTCTTCTGCTTGTCTGGCCATGATGTACACGTTGTGGCTGT  
GTAGTTGTACTCCAGCTTGTGGCCAGGATGTTGCCGTCCTCCTTGAAGTCGATGCC  
CTTCAGCTCGATCCTGTTACCAGGGTGTCGCCCTCGAACTTCACCTCGGCCCTGGTC  
TTGTAGTTGCCGTCGTCCTTGAAGAAGATGGTCCTCTCCTGCACGTAGCCCTCGGGC  
ATGGCGCTCTTGAAGAAGTCGTGCTGCTTCATGTGGTCTGGGGTACCTGCTGAAGCAC  
TGCACGCCGTAGGTACAGGGTGGTCACCAGGGTGGGCCAGGGCACGGGCAGCTTGCC  
GGTGGTGCAGATGAACTTCAGGGTACAGCTTGCCGTAGGTGGCGTCGCCCTCGCCCTC  
GCCGCTCACGCTGAACTTGTGGCCGTTACGTCGCCGTCAGCTCCACCAGGATGGG  
CACCACGCCGGTGAACAGCTCCTCGCCCTTGCTCACGGGGCCGGGGTTCTCCTCCAC  
GTCGCCGGCCTGCTTCAGCAGGCTGAAGTTGGTGGCCAGGATCCTCTCGCACAGCCT  
CCAGCCGGTACGCCGTTGATGGTCACCCTGAACAGCAGGCTGCCGTCGGGGTTGAT  
CAGCCTCTCGTCGATGATCTTGTGGCCGTTCCACAGGGTGCCGGTACGGTGATCTT  
CTTGCCGTCGAACACGGCGATGCCCTCGTAGGGCCTGCCGAAGTAGTCGATCATGTT  
GGGGGTACGCCGTCGATCACCAGGGTGGCGTAGTGCAGGATCACCTTGAAGTGGT  
GGTCGTCCACGGGGTACACCACCTTGAATACTTCTCGATCTGGCCCATCTGGTCGC  
CGCTCAGGCCCTCGTAGGGGATGATCACGTGGATGTCGATCTTCAGGCCGTTCTCGC  
CGCTCAGCACGATCCTCTGGATGGGGGTACGCTCACGCCCAGGTTCTGGAACAGG  
CTGCTCACGCCGCCCTGCTCCAGCACCTGGTCCAGGTTGTAGCCGGCGGTCTGCCTC  
CAGTCGCCACGAAGTCTCAGGGTGAACACGGCCTCCTCGAAGCTGCACTTCTCC  
TCCATGCACTCCCTCTCAGGTTGCCCTGCACGAACTCCTCAGCTTGCCGCTGTTGT  
ACCTCTTGGGCCTGTTACAGGATCTTGTGGCGTTCTCGTGGTCCAGGAA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA  
GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTAGATCTCTTAGGTCAGTGAA  
GAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATG  
TTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACA  
AAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAATTTGGAAGAGTTTGTTCAA  
GGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAG  
AAGTTTTTGAACAACACTGAAAGAACAACCTGAATTTTTGGAAGCAGTATGTTGATGGA  
GATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAAT

TCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAATTAGATGTA  
ACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAA  
CAAGGTGGTTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTG  
TGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAACTTCTAAGCT  
CACCCGTGCTGAGACTGTTTTTCTGATGTGGACTATGTAAATTCTACTGAAGCTGA  
AACCATTTTGGATAACATCACTCAAAGCACCCAATCATTAAATGACTTCACTCGGGT  
TGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTTTTGAATGG  
TAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTAAGTGC  
TGCCCCTGTGTTGAAACTGGTGTTAAAATTACAGTTGTTCGCAGGTGAACATAATAT  
TGAGGAGACAGAACATAACAGAGCAAAAGCGAAATGTGATTTCGAATTATTCCTCACC  
ACAACACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAAGTGG  
ACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAAT  
ACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGGAAGAGTCT  
TCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACC  
GAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTG  
GCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTT  
ACTGAAGTGGAAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTG  
TGCAATGAAAGGCAAATATGGAATATATACCAAGGTCTCCCGGTATGTCAACTGGA  
TTAAGGAAAAACAAAGCTCACTGTCAGCGGATGGAGACTGTTCAAGAAGATCAGC  
TAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGTGCC  
TTCCTTGACCCTGGAAGGTGCCACTCCACTGTCCTTTCCTAATAAAATGAGGAAAT  
TGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGA  
CAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGC  
TCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCAAA  
AACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTT  
AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTC  
ACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATG  
TATCTTATCATGTCTGTTAGGAAATCTTCTTAAACAGCCGCCAGCCGCTCACGGTGA  
GCTTAGTCTTTTTCTTTATCCAATTCACGTAGCGAGAGACCTTCGTATAGATGCCATA  
TTTCCCCTTCATCGCACATTCCTCCCCCAACTTATTATCCCGGTCAAGAACTTGTT  
CCTTCGACTTCAGTGACGTGTGGTCCACCTGAATCACCTTGGCATGAGTCGCGACCG  
CCCTCGTGAAACCCAGCACAAAACATGTTATTGTAATCGTAAATTCGTGGACAGA  
AGACAGGTGCTCTATCGACCAACGGGACGCGCAAATATTGCAGAACGAGGGCTGA  
TCGACCTTTGTGGAAGACCCGCCCCACCCACTCACATATCCGCTCCCAAATTTCAA  
GAAGATATTTGTATATTCTTTATCGGCTATACAAATCGGGGTAACATAGGAGTTAAG  
TACGAGTGGCTCGTCCAGCTCCAGGAGGGCTATATCATGGTTGTACTTGTTTATAGC  
GGCATTATAATTGTGATGGGGTATGATCCTGATAACATTCCTTTTCTGTTTCAGTATGC  
TCAGTTTCTTCAATGTTGTGTTCCGCCAGCCACGACCGTAATCTTAAACCCCGTCTCGA  
CACAGTGTGCGGCCGTTACAATCCACTTTTCATTGACTATGGAGCCCCACAAAACG  
CGTCGACTTTTCCGTTGAGCACCACTGCCATGGAAATTGGCCAGGTTTAGCGTCCT

CGCCCCGACAACCCTAGTAAAGTCATTAATGACTGTGTGGATTGTGTTATATTAT  
CAAGAATCGTTTCGGCTTCAGTAGAGTTAACGTAGTCCACATCGGGAAAAACTGTCT  
CGGCCCTTGTCAACTTTGATGTCTGGGACACACTTACCCGACCGCACGGGAAGGGCA  
CCGCCGGTTCACAGCTCTTTTGATTCTCAGCGAGCCGGTAGCCCTCAGTGCAACTAC  
ACACAACCTTTGTTGTCTGGCGGAATTTTTACAGAATTGCTCGCATCGTCCATTTTAAAT  
GTTGCAGGTGACGTCCAACCTCGCAGTTTTTTCCTTCAAAAACAAAAGGGCACCAACA  
CTCGTAGGAATTTATATCGTCTTTACAACCTCCCCCATTTCAGACATGGATTAGATTTCG  
CATTGGTCCCCATCGACATATTGCTTCCAGAACTCAGTGGTCCGTTCTGTATTCTCAA  
ACACCTCGCGCGCTTCTTCAAAACTGCATTTTTCTCCATACACTCTCGCTCCAAGTT  
CCCTTGCACGAATTCTTCAAGCTTTCCTGAGTTATACCTTTTAGGCCGGTTAAGTATC  
TTATTCGCGTTTTTCGTGGTCCAGAAAAACTGTGGAAACAGGGAGAGAAAAACCACA  
CAACATATTTAAAGATTGATGAAGACAACCTAACTGTAATATGCTGCTTTTTGTTCTTC  
TCTTCACTGACCTAAGAGATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCT  
CTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCTGGGGCGAC  
CTTTGGTTCGCCCGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGGAGTGGCCAA

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

GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAG  
GTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAG  
AGAGGGGAGTGGCCAACTCCATCACTAGGGGGTTCCTGGAGGGGTGGAGTCGTGATAG  
GTCAGTGAAGAGAAGAACA AAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATC  
TTTAAATATGTTGTGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAA  
ACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAG  
TTTGTTC AAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGA  
AGCACGAGAAGTTTTTGA AAACACTGAAAGAACA ACTGAATTTTGG AAGCAGTATG  
TTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATG  
ACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAAT  
TAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGT  
GCTGATAACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCA  
GAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTCACAAAC  
TTCTAAGCTCACCCGTGCTGAGACTGTTTTTCTGATGTGGACTATGTAAATTCTACT  
GAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACTTC  
ACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTT  
TTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATT  
GTA ACTGCTGCCACTGTGTTGAAACTGGTGTAAAATTACAGTTGTCGCAGGTGAA  
CATAATATTGAGGAGACAGAACATAACAGAGCAAAAAGCGAAATGTGATTTCGAATTAT  
TCCTCACCACA ACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCT  
GGA ACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCTATTTGCATTGCTGA  
CAAGGAATACACGAACATCTTCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGG  
AAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACT  
TGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTT

CTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGAC  
CCCATGTTACTGAAGTGGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTG  
AAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTATGTC  
AACTGGATTAAGGAAAAAACAAGCTCACTTAACCTCGACTGTGCCTTCTAGTTGCC  
AGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCC  
CACTGTCCTTTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCA  
TTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGGAGGATTGGGAAGAC  
AATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAAC  
CAGCTGGGGCTCTAGGGGGTATCCCCACTAGTCCACTCCCTCTCTGCGCGCTCGTC  
GCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGG  
CCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA

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

GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAG  
GTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAG  
AGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTTGAGGGGTGGAGTCGTGACCT  
AGGTCGTCTCCGGCTCTGCTTTTTCCAGGGGTGTGTTTCGCCGAGAAGCACGTAAGA  
GTTTTATGTTTTTTCATCTCTGCTTGTATTTTTCTAGTAATGGAAGCCTGGTATTTTAA  
AATAGTTAAATTTTCTTTAGTGCTGATTTCTAGATTATTATTACTGTTGTTGTTGTTA  
TTATTGTCATTATTTGCATCTGAGAACTAGGTCAGTGAAGAGAAGAACA AAAAAGCA  
GCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGGTTTTTCTCTCC  
CTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACA AAAATTCTGAATCGGCCAA  
AGAGGTATAATTCAGGTA AATTGGAAGAGTTTGTTC AAGGGAACCTTGAGAGAGAA  
TGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGA AAACACTGA  
AAGAACA ACTGAATTTTGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATC  
CATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTC  
CCTTTGGATTTGAAGGAAAGA ACTGTGAATTAGATGTAACATGTAACATTAAGAAT  
GGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTTGCTCCTGT  
ACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATT  
TCCATGTGGAAGAGTTTCTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGT  
TTTTCTGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTGGATAACAT  
CACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGC  
CAAACCAGGTCAATTCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTG  
TGGAGGCTCTATCGTTAATGAAA AATGGATTGTA ACTGCTGCCACTGTGTTGAAAC  
TGGTGTTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGAGACAGAACATA  
CAGAGCAAAAGCGAAATGTGATTCGAATTATTCCTCACCACA ACTACAATGCAGCT  
ATTAATAAGTACAACCATGACATTGCCCTTCTGGA ACTGGACGAACCCTTAGTGCTA  
AACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTC  
AAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATC  
AGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCTTCTA  
TCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGT

AGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAGGGAC  
CAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAAT  
ATGGAATATATACCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAACAAAG  
CTCACTTAACCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCC  
CGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGA  
GGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGG  
GCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCG  
GTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCC  
CCCTTAGGTGGTTATATTATTGATATATTTTTGGTATCTTTGATGACAATAATGGGGG  
ATTTTGAAAGCTTAGCTTTAAATTTCTTTTAATTAATAAAAAAATGCTAGGCAGAATG  
ACTCAAATTACGTTGGATACAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCT  
CGACCATGCTATACTAAAAATTAAGGTGTAAGTACTAGTCCACTCCCTCTCTGCGCGCTC  
GCTCGCTCACTGAGGCCGGGCGACCAAAGGTGCCCCGACGCCCGGGCTTTGCCCGG  
GCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA  
GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCCTAGATCTGATTTTGAAAGCTTA  
GCTTTAAATTTCTTTTAATTAATAAAAAAATGCTAGGCAGAATGACTCAAATTACGTT  
GGATACAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATAC  
TAAAAATTAAGGTGTGTGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACC  
TTTATTTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAA  
ATCTTGAGTTTGAATGCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGT  
TGGTGAAAAAATATAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGAT  
TTTCTGTAAACGATCGGGAACCTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAG  
AAGAACA AAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTG  
TGTGGTTTTTCTCTCCCTGTTTCCACAGTTTTTCTTGATCATGAAAACGCCAACAAAA  
TTCTGAATCGGCCAAAGAGGTTTTCTTGATCATGAAAACGCCAACAAAATTCTGAATC  
GGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGTTTGTTC AAGGGAACCTTGAG  
AGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAA  
CACTGAAAGAACA ACTGAATTTTGAAGCAGTATGTTGATGGAGATCAGTGTGAGT  
CCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATGACATTAATTCCTATGAATGTT  
GGTGTCCCTTTGGATTTGAAGGAAAGA ACTGTGAATTAGATGTAACATGTAACATTA  
AGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGTGCTGATAACAAGGTGGTTTTGC  
TCCTGTACTGAGGGATATCGACTTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGT  
GCCATTTCCATGTGGAAGAGTTTCTGTTTCACAACTTCTAAGCTCACCCGTGCTGA  
GACTGTTTTTCTGATGTGGACTATGTAAATTCTACTGAAGCTGAAACCATTTTTGGAT  
AACATCACTCAAAGCACCCAATCATTTAATGACTTCACTCGGGTTGTTGGTGGAGAA  
GATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCA  
TTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATTGTA ACTGCTGCCCACTGTGTT

GAAACTGGTGTAAAATTACAGTTGTCGCAGGTGAACATAATATTGAGGAGACAGA  
 ACATACAGAGCAAAAGCGAAATGTGATTTCGAATTATTCCTCACCACAACACTACAATG  
 CAGCTATTAATAAGTACAACCATGACATTGCCCTTCTGGAAGTGGACGAACCCCTTAG  
 TGCTAAACAGCTACGTTACACCTATTTGCATTGCTGACAAGGAATACACGAACATCT  
 TCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGA  
 GATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACTTGTTGACCGAGCCACATGTCT  
 TCTATCTACAAAGTTCACCATCTATAACAACATGTTCTGTGCTGGCTTCCATGAAGG  
 AGGTAGAGATTCATGTCAAGGAGATAGTGGGGGACCCCATGTTACTGAAGTGGAAG  
 GGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGC  
 AAATATGGAATATATAACCAAGGTATCCCGGTATGTCAACTGGATTAAGGAAAAAAC  
 AAAGCTCACTTAACTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCT  
 CCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCCTTTCCTAATAAAA  
 TGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGT  
 GGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAT  
 GCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTA  
 TCCCCGTGAGATCGCCATCGGTATAATGATTTGGGAGAACAACATTTCAAAGGCCT  
 GTAAGTTATAATGCTGAAAGCCCACTTAATATTTCTGGTAGTATTAGTTAAAGTTTT  
 AAAACACCTTTTTCCACCTTGAGTGTGAGAATTGTAGAGCAGTGCTGTCCAGTAGAA  
 ATGTGTGCATTGACAGAAAGACTGTGGATCTGTGCTGAGCAATGTGGCAGCCAGAG  
 ATCACAAGGCTATCAAGCACTTTGCACATGGCAAGTGTAAGTGAAGCACACATT  
 CAAATAATAGTTAATTTTAATTGAATGTATCTAGCCATGTGTGGCTAGTAGCTCCTTT  
 CCTGGAGAGAGAATCTGGAGCCACATCTAACTTGTTAAGTCTGGAATCTTATTTTT  
 TATTTCTGGAAAGGTCTATGAACTATAGTTTTGGGGGCAGCTCACTTACTAACTTTTA  
 ATGCAATAAGAATCTCATGGTATCTTGAGAACATTATTTTGTCTCTTTGTAGATCTAG  
 GAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAG  
 GCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAG  
 CGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
 GGTCGCCCCGACGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA  
 GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCCTAGATCTTAGCCTCTGGCAA  
 ATGAAGTGGGTAACCTTTCTCCTCCTCCTCTTCGTCTCCGGCTCTGCTTTTTCCAGGG  
 GTGTGTTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTT  
 TTCTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTT  
 CTAGATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACCCT  
 TAGGTGGTTATATTATTGATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTT  
 TGAAAGCTTAGCTTTAAATTTCTTTTAAATTAATAAAAAAATGCTAGGCAGAATGACTC  
 AAATTACGTTGGATACAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGAC  
 CATGCTATACTAAAAATTAAGTGTGTGTTACTAATTTTATAAATGGAGTTTCCAT  
 TTATATTTACCTTTATTTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGAC



AGAAACACTAAATCTTGAGTTTGAATGCACAGATATAAACACTTAACGGGTTTTAA  
AATAATAATGTTGGTGAAAAATATAACTTTGAGTGTAGCAGAGAGGAACCATTGC  
CACCTTCAGATTTTCCTGTAACGATCGGGAACCTGGCATCTTCAGGGAGTAGCTTAGG  
TCAGTGAAGAGAAGAACAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTT  
TAAATATGTTGTGTGGTTTTTCTCTCCCTGTTCCACAGTTTTTCTTGATCATGAAAA  
CGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAGGTAAATTGGAAGAGT  
TTGTTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAGTGTAGTTTTGAAGAA  
GCACGAGAAGTTTTTGAAAACACTGAAAGAACAACCTGAATTTTGAAGCAGTATGT  
TGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGGCAGTTGCAAGGATG  
ACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGGAAAGAACTGTGAAT  
TAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGTTTTGTAAAAATAGT  
GCTGATAACAAGGTGGTTTGCTCCTGTACTGAGGGATATCGACTTGCAGAAAACCA  
GAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTTTCTGTTTCACAAAC  
TTCTAAGCTCACCCGTGCTGAGACTGTTTTTCTGATGTGGACTATGTAAATTCTACT  
GAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATCATTTAATGACTTC  
ACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCTTGGCAGGTTGTT  
TTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAATGAAAAATGGATT  
GTAAGTCTGCCACTGTGTTGAAACTGGTGTAAAATTACAGTTGTCGCAGGTGAA  
CATAATATTGAGGAGACAGAACATAACAGAGCAAAAGCGAAATGTGATTTCGAATTAT  
TCCTCACCAACTACAATGCAGCTATTAATAAGTACAACCATGACATTGCCCTTCT  
GGAAGTGGACGAACCCTTAGTGCTAACAGCTACGTTACACCTATTTGCATTGCTGA  
CAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTATGTAAGTGGCTGGGG  
AAGAGTCTTCACAAAGGGAGATCAGCTTTAGTTCTTCAGTACCTTAGAGTTCCACT  
TGTTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCTATAACAACATGTT  
CTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGAGATAGTGGGGGAC  
CCCATGTTACTGAAGTGAAGGGACCAGTTTCTTAACTGGAATTATTAGCTGGGGTG  
AAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGGTATCCCGGTATGTC  
AACTGGATTAAGGAAAAACAAAGCTCACTTAACCTCGACTGTGCCTTCTAGTTGCC  
AGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCCTTGACCCTGGAAGGTGCCACTCC  
CACTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCA  
TTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC  
AATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAAC  
CAGCTGGGGCTCTAGGGGGTATCCCCGTGAGATCGCCCATCGGTATAATGATTTGGG  
AGAACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCACTTAATATTTCT  
GGTAGTATTAGTTAAAGTTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAATTGTA  
GAGCAGTGTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGTGC  
TGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCACATGGCAAG  
TGTAAGTGAAGCACACATTCAAATAATAGTTAATTTTAATTGAATGTATCTAGCC  
ATGTGTGGCTAGTAGCTCCTTTCCTGGAGAGAGAATCTGGAGCCCACATCTAACTTG  
TTAAGTCTGGAATCTTATTTTTTATTTCTGGAAAGGTCTATGAACTATAGTTTTGGGG

GCAGCTCACTTACTAACTTTTAATGCAATAAGAATCTCATGGTATCTTGAGAACATT  
ATTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAGTAAGGGGTCTATACTTA  
AGTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAAAAAATGAGTAATTAATT  
TATTTTTAGAAGGTCAATAGTATCATGTATTCCAAATAACAGAGGTATATGGTTAGA  
AAAGAAACAATTCAAAGGACTTATATAATATCTAGCCTTGACAATGAATAAATTTA  
GAGAGTAGTTTGCCTGTTTGCCTCATGTTCATAAATCTATTGACACATATGTGCATCT  
GCACTTCAGCATGGTAGAAGTCCATATTCAGATCTAGGAACCCCTAGTGATGGAGTT  
GGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCCGGGCAAAGCCCGG  
GCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGG  
GAGTGGCCAA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA  
GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTAGATCTAAGTATATTAGAGC  
GAGTCTTTCTGCACACAGATCACCTTTCCTATCAACCCCACTAGCCTCTGGCAAAT  
GAAGTGGGTAACTTTCTCCTCCTCCTCTTCGTCTCCGGCTCTGCTTTTTCCAGGGGT  
GTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTTTT  
CTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTTCT  
AGATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACCTTTT  
TCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAG  
GTAAATTGGAAGAGTTTGTCAAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAG  
TGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACCTGAATT  
TTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGG  
CAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGG  
AAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGT  
TTTGTA AAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTG TACTGAGGGATATCGAC  
TTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTT  
CTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTCCTGATGTGGACTA  
TGTA AATTCTACTGAAGCTGAAACCATTTTGATAACATCACTCAAAGCACCCAATC  
ATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCC  
TTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAA  
TGAAA AATGGATTGTA ACTGCTGCCCACTGTGTTGAAACTGGTGT TAAAATTACAGT  
TGTCGCAGGTGAACATAATATTGAGGAGACAGAACATAACAGAGCAAAAGCGAAAT  
GTGATTCGAATTATTCCTCACCACA ACTACAATGCAGCTATTAATAAGTACAACCAT  
GACATTGCCCTTCTGGA ACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCT  
ATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTAT  
GTAAGTGGCTGGGGAAGAGTCTTCCACA AAGGGAGATCAGCTTTAGTTCTTCAGTAC  
CTTAGAGTTCCACTTGTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCT  
ATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGA  
GATAGTGGGGGACCCCATGTTACTGAAGTGG AAGGGACCAGTTTCTTAACTGGAAT

TATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGG  
TATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAGCTCACTTAACCTCGACTGT  
GCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTG  
GAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGT  
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA  
GGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTG  
AGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTA  
TTGATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTT  
AAATTTCTTTTAATTAAAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATA  
CAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAA  
ATTAAGAGTGTGTGTTACTAATTTATAAATGGAGTTTCCATTTATATTTACCTTTAT  
TTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAAGAT  
CTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCAC  
TGAGGCCGCCCGGGCAAAGCCCAGGCGTGGGCGACCTTTGGTCGCCCGGCCTCAG  
TGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA  
GAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCTAGATCTAAGTATATTAGAGC  
GAGTCTTTCTGCACACAGATCACCTTTCCTATCAACCCCACTAGCCTCTGGCAAAT  
GAAGTGGGTAACCTTTCCTCCTCCTCTTCGTCTCCGGCTCTGCTTTTTCCAGGGGT  
GTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTTTT  
CTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTTCT  
AGATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACCTTTT  
TCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAG  
GTAAATTGGAAGAGTTTGTTCAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAG  
TGTAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACAACCTGAATT  
TTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGG  
CAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGG  
AAAGAACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGT  
TTTGTA AAAAATAGTGCTGATAACAAGGTGGTTTGCTCCTGTA CTGAGGGATATCGAC  
TTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTT  
CTGTTTCACAAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTTCCTGATGTGGACTA  
TGTA AATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATC  
ATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTTCC  
TTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAA  
TGAAA AATGGATTGTA ACTGCTGCCACTGTGTTGAAACTGGTGT TAAAATTACAGT  
TGTCGCAGGTGAACATAATATTGAGGAGACAGAACATAACAGAGCAAAAAGCGAAAT  
GTGATTCGAATTATTCCTCACCACA ACTACAATGCAGCTATTAATAAGTACAACCAT  
GACATTGCCCTTCTGGA ACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCT

ATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTAT  
GTAAGTGGCTGGGGAAGAGTCTTCCACAAAGGGAGATCAGCTTTAGTTCTTCAGTAC  
CTTAGAGTTCCACTTGTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCT  
ATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGA  
GATAGTGGGGGACCCCATGTTACTGAAGTGGAAAGGGACCAGTTTCTTAACTGGAAT  
TATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGG  
TATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTTAACTCGACTGT  
GCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTG  
GAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGT  
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA  
GGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTG  
AGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTA  
TTGATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTT  
AAATTTCTTTTAATTAATAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATA  
CAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAA  
ATTAAGTGTGTGTTACTAATTTTATAAATGGAGTTTCCATTTATATTTACCTTTAT  
TTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAATCTT  
GAGTTTGAATGCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGTTGGTG  
AAAAATATAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGATTTTCCT  
GTAACGATCGGGAACCTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAGAAGAA  
CAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGG  
TTTTTCTCTCCCTGTTTCCACAGACAAGAGTGAGATCGCCCATCGGTATAATGATTTG  
GGAGAACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCACTTAATATTT  
CTGGTAGTATTAGTTAAAGTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAATTGT  
AGAGCAGTGCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGTG  
CTGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCACATGGCAA  
GTGTAACCTGAGAAGCACACATTCAAATAATAGTTAATTTTAATTGAATGTATCTAGC  
CATGTGTGGCTAGTAGCTCCTTTCCTGGAGAGAGAATCTGGAGCCCACATCTAACTT  
GTAAAGTCTGGAATCTTATTTTTTATTTCTGGAAAGGTCTATGAACTATAGTTTTGGG  
GGCAGCTCACTTACTAACTTTTAATGCAATAAGATCCATGGTATCTTGAGAACATTA  
TTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAGTAAGGGGTCTATACTTAA  
GTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAAAAAATGAGTAATTAATTT  
ATTTTTAGAAGGTCAATAGTATCATGTATTCCAAATAACAGAGGTATATGGTTAGAA  
AAGAAACAATTCAAAGGACTTATATAATATCTAGCCTTGACAATGAATAAATTTAG  
AGAGTAGTTTGCCTGTTTGCCTCATGTTTCATAAATCTATTGACACATATGTGCATCTG  
CACTTCAGCATGGTAGAAGTCCATATTCCTTTGCTTGGAAAGGCAGGTGTTCCCAT  
ACGCCTCAGAGAATAGCTGACGGGAAGAGGCTTTCTAGATAGTTGTATGAAAGATA  
TACAAAATCTCGCAGGTATACACAGGCATGATTTGCTGGTTGGGAGAGCCACTTGCC  
TCATACTGAGGTTTTTGTGTCTGCTTTTCAGAGTCCTGATTGCCTTTTCCAGTATCTC  
CAGAAATGCTCATACGATGAGCATGCCAAATTAGTGCAGGAAGTAACAGACTTTGC

AAAGACGTGTGTTGCCGATGAGTCTGCCGCAACTGTGACAAATCCCTTGTGAGTAC  
CTTCTGATTTTGTGGATCTACTTTCCTGCTTTCTGGA ACTCTGTTTCAAAGCCAATCA  
TGACTCCATCACTTAAGGCCCCGGGAACACTGTGGCAGAGGGCAGCAGAGAGATTG  
ATAAAGCCAGGGTGATGGGAATTTTCTGTGGGACTCCATTTTCATAGTAATTGCAGAA  
GCTACAATACTCAAAAAGTCTCACCATGACTGCCCAAATGGGAGCTTGACAG  
TGACAGTGACAGTAGATATGCCAAAGTGGATGAGGGAAAGACCACAAGAGCTAAA  
CCCTGTAAAAAGAACTGTAGGCAACTAAGGAATGCAGAGAGAAAAGATCTAGGAAC  
CCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCG  
CCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAG  
CGAGCGCGCAGAGAGGGAGTGGCCAA

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

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA  
GGTCGCCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCA  
GAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTAGATCTAAGTATATTAGAGC  
GAGTCTTTCTGCACACAGATCACCTTTCCTATCAACCCCACTAGCCTCTGGCAAAT  
GAAGTGGGTAACTTTCTCCTCCTCCTCTTCGTCTCCGGCTCTGCTTTTTCCAGGGGT  
GTGTTTCGCCGAGAAGCACGTAAGAGTTTTATGTTTTTTCATCTCTGCTTGTATTTTT  
CTAGTAATGGAAGCCTGGTATTTTAAAATAGTTAAATTTTCCTTTAGTGCTGATTTCT  
AGATTATTACTGTTGTTGTTGTTATTATTGTCATTATTTGCATCTGAGAACCTTTT  
TCTTGATCATGAAAACGCCAACAAAATTCTGAATCGGCCAAAGAGGTATAATTCAG  
GTAAATTGGAAGAGTTTGTTC AAGGGAACCTTGAGAGAGAATGTATGGAAGAAAAG  
TG TAGTTTTGAAGAAGCACGAGAAGTTTTTGAAAACACTGAAAGAACA ACTGAATT  
TTGGAAGCAGTATGTTGATGGAGATCAGTGTGAGTCCAATCCATGTTTAAATGGCGG  
CAGTTGCAAGGATGACATTAATTCCTATGAATGTTGGTGTCCCTTTGGATTTGAAGG  
AAAGA ACTGTGAATTAGATGTAACATGTAACATTAAGAATGGCAGATGCGAGCAGT  
TTTGTA AAAATAGTGCTGATAACAAGGTGGTTTGCTCCTG TACTGAGGGATATCGAC  
TTGCAGAAAACCAGAAGTCCTGTGAACCAGCAGTGCCATTTCCATGTGGAAGAGTTT  
CTGTTTCACAACTTCTAAGCTCACCCGTGCTGAGACTGTTTTTTCCTGATGTGGACTA  
TGTA AATTCTACTGAAGCTGAAACCATTTTGGATAACATCACTCAAAGCACCCAATC  
ATTTAATGACTTCACTCGGGTTGTTGGTGGAGAAGATGCCAAACCAGGTCAATTCCC  
TTGGCAGGTTGTTTTGAATGGTAAAGTTGATGCATTCTGTGGAGGCTCTATCGTTAA  
TGAAA AATGGATTGTA ACTGCTGCCCACTGTGTTGAAACTGGTGT TAAAATTACAGT  
TGTCGCAGGTGAACATAATATTGAGGAGACAGAACATAACAGAGCAAAGCGAAAT  
GTGATTCGAATTATTCCTCACCACA ACTACAATGCAGCTATTAATAAGTACAACCAT  
GACATTGCCCTTCTGGA ACTGGACGAACCCTTAGTGCTAAACAGCTACGTTACACCT  
ATTTGCATTGCTGACAAGGAATACACGAACATCTTCCTCAAATTTGGATCTGGCTAT  
GTAAGTGGCTGGGGAAGAGTCTTCCACA AAGGGAGATCAGCTTTAGTTCTTCAGTAC  
CTTAGAGTTCCACTTGTGACCGAGCCACATGTCTTCTATCTACAAAGTTCACCATCT  
ATAACAACATGTTCTGTGCTGGCTTCCATGAAGGAGGTAGAGATTCATGTCAAGGA  
GATAGTGGGGGACCCCATGTTACTGAAGTGG AAGGGACCAGTTTCTTAACTGGAAT

TATTAGCTGGGGTGAAGAGTGTGCAATGAAAGGCAAATATGGAATATATACCAAGG  
TATCCCGGTATGTCAACTGGATTAAGGAAAAAACAAAGCTCACTTAACCTCGACTGT  
GCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTG  
GAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGT  
CTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA  
GGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTG  
AGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCCTTAGGTGGTTATATTA  
TTGATATATTTTTGGTATCTTTGATGACAATAATGGGGGATTTTGAAAGCTTAGCTTT  
AAATTTCTTTTAATTAATAAAAAAATGCTAGGCAGAATGACTCAAATTACGTTGGATA  
CAGTTGAATTTATTACGGTCTCATAGGGCCTGCCTGCTCGACCATGCTATACTAAAA  
ATTAAGTGTGTGTTACTAATTTATAAATGGAGTTTCCATTTATATTTACCTTTAT  
TTCTTATTTACCATTGTCTTAGTAGATATTTACAAACATGACAGAAACACTAAATCTT  
GAGTTTGAATGCACAGATATAAACACTTAACGGGTTTTAAAAATAATAATGTTGGTG  
AAAAATATAACTTTGAGTGTAGCAGAGAGGAACCATTGCCACCTTCAGATTTTCCT  
GTAACGATCGGGAACCTGGCATCTTCAGGGAGTAGCTTAGGTCAGTGAAGAGAAGAA  
CAAAAAGCAGCATATTACAGTTAGTTGTCTTCATCAATCTTTAAATATGTTGTGTGG  
TTTTCTCTCCCTGTTTCCACAGACAAGAGTGAGATCGCCCATCGGTATAATGATTTG  
GGAGAACAACATTTCAAAGGCCTGTAAGTTATAATGCTGAAAGCCCACTTAATATTT  
CTGGTAGTATTAGTTAAAGTTTTAAAACACCTTTTTCCACCTTGAGTGTGAGAATTGT  
AGAGCAGTGCTGTCCAGTAGAAATGTGTGCATTGACAGAAAGACTGTGGATCTGTG  
CTGAGCAATGTGGCAGCCAGAGATCACAAGGCTATCAAGCACTTTGCACATGGCAA  
GTGTAAGTGAAGCACACATTCAAATAATAGTTAATTTAATTGAATGTATCTAGC  
CATGTGTGGCTAGTAGCTCCTTTCCTGGAGAGAGAATCTGGAGCCACATCTAACTT  
GTTAAGTCTGGAATCTTATTTTTTATTTCTGGAAAGGTCTATGAACTATAGTTTTGGG  
GGCAGCTCACTTACTAACTTTAATGCAATAAGATCCATGGTATCTTGAGAACATTA  
TTTTGTCTCTTTGTAGTACTGAAACCTTATACATGTGAAGTAAGGGGTCTATACTTAA  
GTCACATCTCCAACCTTAGTAATGTTTTAATGTAGTAAAAAATGAGTAATTAATTT  
ATTTTTAGAAGGTCAATAGTATCATGTATTCCAAATAACAGAGGTATATGGTTAGAA  
AAGAAACAATTCAAAGGACTTATATAATATCTAGCCTTGACAATGAATAAATTTAG  
AGAGTAGTTTGCCTGTTTGCCTCATGTTCATAAATCTATTGACACATATGTGCATCTG  
CACTTCAGCATGGTAGAAGTCCATATTCCTTTGCTTGGAAGGCAGGTGTTCCCAT  
ACGCCTCAGAGAATAGCTGACGGGAAGAGGCTTTCTAGATAGTTGTATGAAAGATA  
TACAAAATCTCGCAGGTATACACAGGCATGATTTGCTGGTTGGGAGAGCCACTTAG  
ATCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTC  
ACTGAGGCCCGCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTC  
AGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAA

OPC Cas9 (SEQ ID NO: 703)

ATGGATAAGAAGTACTCAATCGGGCTGGATATCGGAACTAATTCCGTGGGTT  
GGGCAGTGATCACGGATGAATACAAAGTGCCGTCCAAGAAGTTCAAGGTCCTGGGG  
AACACCGATAGACACAGCATCAAGAAAAATCTCATCGGAGCCCTGCTGTTTGACTC

CGGCGAAACCGCAGAAGCGACCCGGCTCAAACGTACCGCGAGGGCGACGCTACACCC  
GGCGGAAGAATCGCATCTGCTATCTGCAAGAGATCTTTTCGAACGAAATGGCAAAG  
GTCGACGACAGCTTCTTCCACCGCCTGGAAGAATCTTTCCTGGTGGAGGAGGACAA  
GAAGCATGAACGGCATCCTATCTTTGGAAACATCGTCGACGAAGTGGCGTACCACG  
AAAAGTACCCGACCATCTACCATCTGCGGAAGAAGTTGGTTGACTCAACTGACAAG  
GCCGACCTCAGATTGATCTACTTGGCCCTCGCCCATATGATCAAATTCCGCGGACAC  
TTCCTGATCGAAGGCGATCTGAACCCTGATAACTCCGACGTGGATAAGCTTTTCATT  
CAACTGGTGCAGACCTACAACCAACTGTTCGAAGAAAACCCAATCAATGCTAGCGG  
CGTCGATGCCAAGGCCATCCTGTCCGCCCGGCTGTCGAAGTCGCGGGCGCCTCGAAA  
ACCTGATCGCACAGCTGCCGGGAGAGAAAAAGAACGGACTTTTCGGCAACTTGATC  
GCTCTCTCACTGGGACTCACTCCCAATTTCAAGTCCAATTTTGACCTGGCCGAGGAC  
GCGAAGCTGCAACTCTCAAAGGACACCTACGACGACGACTTGGACAATTTGCTGGC  
ACAAATTGGCGATCAGTACGCGGATCTGTTCCCTTGCCGCTAAGAACCTTTCGGACGC  
AATCTTGCTGTCCGATATCCTGCGCGTGAACACCGAAATAACCAAAGCGCCGCTTAG  
CGCCTCGATGATTAAGCGGTACGACGAGCATCACCAGGATCTCACGCTGCTCAAAG  
CGCTCGTGAGACAGCAACTGCCTGAAAAGTACAAGGAGATCTTCTTCGACCAGTCC  
AAGAATGGGTACGCAGGGTACATCGATGGAGGGCGCTAGCCAGGAAGAGTTCTATAA  
GTTTCATCAAGCCAATCCTGGAAAAGATGGACGGAACCGAAGAAGTCTGGTCAAGC  
TGAACAGGGAGGATCTGCTCCGGAAACAGAGAACCTTTGACAACGGATCCATTCCC  
CACCAGATCCATCTGGGTGAGCTGCACGCCATCTTGCGGGCGCCAGGAGGACTTTTAC  
CCATTCCTCAAGGACAACCGGGAAAAGATCGAGAAAATTCTGACGTTCCGCATCCC  
GTATTACGTGGGCCCACTGGCGCGCGGCAATTCGCGCTTCGCGTGGATGACTAGAA  
AATCAGAGGAAACCATCACTCCTTGGAATTTTCGAGGAAGTTGTGGATAAGGGAGCT  
TCGGCACAAAGCTTCATCGAACGAATGACCAACTTCGACAAGAATCTCCCAAACGA  
GAAGGTGCTTCTAAGCACAGCCTCCTTTACGAATACTTCACTGTCTACAACGAACT  
GACTAAAGTGAAATACGTTACTGAAGGAATGAGGAAGCCGGCCTTTCTGTCCGGAG  
AACAGAAGAAAGCAATTGTTCGATCTGCTGTTCAAGACCAACCGCAAGGTGACCGTC  
AAGCAGCTTAAAGAGGACTACTTCAAGAAGATCGAGTGTTTCGACTCAGTGGAAAT  
CAGCGGGGTGGAGGACAGATTCAACGCTTCGCTGGGAACCTATCATGATCTCCTGA  
AGATCATCAAGGACAAGGACTTCCTTGACAACGAGGAGAACGAGGACATCCTGGAA  
GATATCGTCCTGACCTTGACCCTTTTCGAGGATCGCGAGATGATCGAGGAGAGGCTT  
AAGACCTACGCTCATCTCTTCGACGATAAGGTCATGAAACAACCTCAAGCGCCGCCG  
GTACACTGGTTGGGGCCGCCTCTCCCGCAAGCTGATCAACGGTATTTCGCGATAAACA  
GAGCGGTAAAACCTATCCTGGATTTCTCAAATCGGATGGCTTCGCTAATCGTAACTT  
CATGCAATTGATCCACGACGACAGCCTGACCTTTAAGGAGGACATCCAAAAGCAC  
AAGTGTCCGGACAGGGAGACTCACTCCATGAACACATCGCGAATCTGGCCGGTTTCG  
CCGGCGATTAAGAAGGGAATTCTGCAAACCTGTGAAGGTGGTTCGACGAGCTGGTGAA  
GGTCATGGGACGGCACAAACCGGAGAATATCGTGATTGAAATGGCCCGAGAAAACC  
AGACTACCCAGAAGGGCCAGAAAAACTCCCGCGAAAGGATGAAGCGGATCGAAGA  
AGGAATCAAGGAGCTGGGCAGCCAGATCCTGAAAGAGCACCCGGTGGAAAACACG

CAGCTGCAGAACGAGAAGCTCTACCTGTACTATTTGCAAAATGGACGGGACATGTA  
 CGTGGACCAAGAGCTGGACATCAATCGGTTGTCTGATTACGACGTGGACCACATCGT  
 TCCACAGTCCTTTCTGAAGGATGACTCGATCGATAACAAGGTGTTGACTCGCAGCGA  
 CAAGAACAGAGGGGAAGTCAGATAATGTGCCATCGGAGGAGGTTCGTGAAGAAGATG  
 AAGAATTACTGGCGGCAGCTCCTGAATGCGAAGCTGATTACCCAGAGAAAGTTTGA  
 CAATCTCACTAAAGCCGAGCGCGGCGGACTCTCAGAGCTGGATAAGGCTGGATTCA  
 TCAAACGGCAGCTGGTCGAGACTCGGCAGATTACCAAGCACGTGGCGCAGATCTTG  
 GACTCCCGCATGAACACTAAATACGACGAGAACGATAAGCTCATCCGGGAAGTGAA  
 GGTGATTACCTTGAAAAGCAAACCTTGTGTCTGGACTTTCGGAAGGACTTTCAGTTTAA  
 CAAAGTGAGAGAAATCAACAACCTACCATCACGCGCATGACGCATACCTCAACGCTG  
 TGGTCGGTACCGCCCTGATCAAAAAGTACCCTAAACTTGAATCGGAGTTTGTGTACG  
 GAGACTACAAGGTCTACGACGTGAGGAAGATGATAGCCAAGTCCGAACAGGAAATC  
 GGGAAAGCAAACCTGCGAAATACTTCTTTTACTCAAACATCATGAACTTTTTCAAGACT  
 GAAATTACGCTGGCCAATGGAGAAATCAGGAAGAGGCCACTGATCGAAACTAACGG  
 AGAAACGGGCGAAATCGTGTGGGACAAGGGCAGGGACTTCGCAACTGTTTCGAAAG  
 TGCTCTCTATGCCGCAAGTCAATATTGTGAAGAAAACCGAAGTGCAAACCGGCGGA  
 TTTTCAAAGGAATCGATCCTCCCAAAGAGAAATAGCGACAAGCTCATTGCACGCAA  
 GAAAGACTGGGACCCGAAGAAGTACGGAGGATTCGATTCGCCGACTGTTCGCATACT  
 CCGTCCTCGTGGTGGCCAAGGTGGAGAAGGGAAAGAGCAAAAAGCTCAAATCCGTC  
 AAAGAGCTGCTGGGGATTACCATCATGGAACGATCCTCGTTCGAGAAGAACCCGAT  
 TGATTTCTCGAGGGCGAAGGGTTACAAGGAGGTGAAGAAGGATCTGATCATCAAAC  
 TCCCAAGTACTCACTGTTTCGAACTGGAAAATGGTCGGAAGCGCATGCTGGCTTCGG  
 CCGGAGAACTCCAAAAGGAAATGAGCTGGCCTTGCTTAGCAAGTACGTCAACTTC  
 CTCTATCTTGCTTCGCACTACGAAAACTCAAAGGGTCACCGGAAGATAACGAACA  
 GAAGCAGCTTTTCGTGGAGCAGCACAAGCATTATCTGGATGAAATCATCGAACAAA  
 TCTCCGAGTTTTCAAAGCGCGTGATCCTCGCCGACGCCAACCTCGACAAAGTCCTGT  
 CGGCCTACAATAAGCATAGAGATAAGCCGATCAGAGAACAGGCCGAGAACATTATC  
 CACTTGTTACCCTGACTAACCTGGGAGCCCCAGCCGCCTTCAAGTACTTCGATACT  
 ACTATCGATCGCAAAAGATACACGTCCACCAAGGAAGTTCTGGACGCGACCCTGAT  
 CCACCAAAGCATCACTGGACTCTACGAACTAGGATCGATCTGTTCGCAGCTGGGTG  
 GCGAT

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

ATGGACAAGAAGTACAGCATCGGACTGGACATCGGAACAAACAGCGTCGGA  
 TGGGCAGTCATCACAGACGAATACAAGGTCCCGAGCAAGAAGTTCAAGGTCCTGGG  
 AAACACAGACAGACACAGCATCAAGAAGAACCTGATCGGAGCACTGCTGTTTCGACA  
 GCGGAGAAACAGCAGAAGCAACAAGACTGAAGAGAACAGCAAGAAGAAGATACAC  
 AAGAAGAAAGAACAGAATCTGCTACCTGCAGGAAATCTTCAGCAACGAAATGGCAA  
 AGGTCGACGACAGCTTCTTCCACAGACTGGAAGAAAGCTTCTGGTCGAAGAAGAC  
 AAGAAGCACGAAAGACACCCGATCTTCGGAAACATCGTCGACGAAGTCGCATACCA  
 CGAAAAGTACCCGACAATCTACCACCTGAGAAAGAAGCTGGTCGACAGCACAGACA



AGGCAGACCTGAGACTGATCTACCTGGCACTGGCACACATGATCAAGTTCAGAGGA  
CACTTCCTGATCGAAGGAGACCTGAACCCGGACAACAGCGACGTCGACAAGCTGTT  
CATCCAGCTGGTCCAGACATAACAACCAGCTGTTTCGAAGAAAACCCGATCAACGCAA  
GCGGAGTCGACGCAAAGGCAATCCTGAGCGCAAGACTGAGCAAGAGCAGAAGACT  
GGAAAACCTGATCGCACAGCTGCCGGGAGAAAAGAAGAACGGACTGTTTCGGAAAC  
CTGATCGCACTGAGCCTGGGACTGACACCGAACTTCAAGAGCAACTTCGACCTGGC  
AGAAGACGCAAAGCTGCAGCTGAGCAAGGACACATACGACGACGACCTGGACAAC  
CTGCTGGCACAGATCGGAGACCAGTACGCAGACCTGTTTCCTGGCAGCAAAGAACCT  
GAGCGACGCAATCCTGCTGAGCGACATCCTGAGAGTCAACACAGAAATCACAAAGG  
CACCGCTGAGCGCAAGCATGATCAAGAGATACGACGAACACCACCAGGACCTGACA  
CTGCTGAAGGCACTGGTCAGACAGCAGCTGCCGGAAAAGTACAAGGAAATCTTCTT  
CGACCAGAGCAAGAACGGATACGCAGGATACATCGACGGAGGAGCAAGCCAGGAA  
GAATTCTACAAGTTCATCAAGCCGATCCTGGAAAAGATGGACGGAACAGAAGAACT  
GCTGGTCAAGCTGAACAGAGAAGACCTGCTGAGAAAGCAGAGAACATTTCGACAAC  
GGAAGCATCCCGCACCAAGATCCACCTGGGAGAAGTGCACGCAATCCTGAGAAGACA  
GGAAGACTTCTACCCGTTCTGAAGGACAACAGAGAAAAGATCGAAAAGATCCTGA  
CATTGAGAATCCCGTACTACGTCGGACCGCTGGCAAGAGGAAACAGCAGATTTCGCA  
TGGATGACAAGAAAGAGCGAAGAAACAATCACACCGTGGAACTTCGAAGAAGTCG  
TCGACAAGGGAGCAAGCGCACAGAGCTTCATCGAAAGAATGACAAACTTCGACAAG  
AACCTGCCGAACGAAAAGGTCCTGCCGAAGCACAGCCTGCTGTACGAATACTTCAC  
AGTCTACAACGAACTGACAAAGGTCAAGTACGTCACAGAAGGAATGAGAAAGCCG  
GCATTCTGAGCGGAGAACAGAAGAAGGCAATCGTCGACCTGCTGTTCAAGACAAA  
CAGAAAGGTCACAGTCAAGCAGCTGAAGGAAGACTACTTCAAGAAGATCGAATGCT  
TCGACAGCGTCGAAATCAGCGGAGTCGAAGACAGATTCAACGCAAGCCTGGGAACA  
TACCACGACCTGCTGAAGATCATCAAGGACAAGGACTTCCTGGACAACGAAGAAAA  
CGAAGACATCCTGGAAGACATCGTCCTGACACTGACACTGTTTCGAAGACAGAGAAA  
TGATCGAAGAAAGACTGAAGACATACGCACACCTGTTTCGACGACAAGGTCATGAAG  
CAGCTGAAGAGAAGAAGATACACAGGATGGGGAAGACTGAGCAGAAAGCTGATCA  
ACGGAATCAGAGACAAGCAGAGCGGAAAGACAATCCTGGACTTCCTGAAGAGCGA  
CGGATTCGCAAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACATTCA  
AGGAAGACATCCAGAAGGCACAGGTCAGCGGACAGGGAGACAGCCTGCACGAACA  
CATCGCAAACCTGGCAGGAAGCCCGGCAATCAAGAAGGGAATCCTGCAGACAGTCA  
AGGTCGTCGACGAACTGGTCAAGGTCATGGGAAGACACAAGCCGGAAAACATCGTC  
ATCGAAATGGCAAGAGAAAACCAGACAACACAGAAGGGACAGAAGAACAGCAGAG  
AAAGAATGAAGAGAATCGAAGAAGGAATCAAGGAACTGGGAAGCCAGATCCTGAA  
GGAACACCCGGTCGAAAACACACAGCTGCAGAACGAAAAGCTGTACCTGTACTACC  
TGCAGAACGGAAGAGACATGTACGTCGACCAGGAACTGGACATCAACAGACTGAGC  
GACTACGACGTCGACCACATCGTCCCGCAGAGCTTCCTGAAGGACGACAGCATCGA  
CAACAAGGTCCTGACAAGAAGCGACAAGAACAGAGGAAAGAGCGACAACGTCCCG  
AGCGAAGAAGTCGTCAAGAAGATGAAGAAGTACTGGAGACAGCTGCTGAACGCAA

AGCTGATCACACAGAGAAAGTTTCGACAACCTGACAAAGGCAGAGAGAGGAGGACT  
GAGCGAACTGGACAAGGCAGGATTCATCAAGAGACAGCTGGTTCGAAACAAGACAG  
ATCACAAAGCACGTCGCACAGATCCTGGACAGCAGAATGAACACAAAGTACGACGA  
AAACGACAAGCTGATCAGAGAAGTCAAGGTCATCACACTGAAGAGCAAGCTGGTCA  
GCGACTTCAGAAAGGACTTCCAGTTCTACAAGGTCAGAGAAATCAACAACCTACCAC  
CACGCACACGACGCATACCTGAACGCAGTCGTTCGGAACAGCACTGATCAAGAAGTA  
CCCGAAGCTGGAAAGCGAATTTCGTCTACGGAGACTACAAGGTCTACGACGTCAGAA  
AGATGATCGCAAAGAGCGAACAGGAAATCGGAAAGGCAACAGCAAAGTACTTCTTC  
TACAGCAACATCATGAACTTCTTCAAGACAGAAATCACACTGGCAAACGGAGAAAT  
CAGAAAGAGACCGCTGATCGAAACAAACGGAGAAACAGGAGAAATCGTCTGGGAC  
AAGGGAAGAGACTTCGCAACAGTCAGAAAGGTCCTGAGCATGCCGCAGGTCAACAT  
CGTCAAGAAGACAGAAGTCCAGACAGGAGGATTCAGCAAGGAAAGCATCCTGCCG  
AAGAGAAACAGCGACAAGCTGATCGCAAGAAAGAAGGACTGGGACCCGAAGAAGT  
ACGGAGGATTCGACAGCCCGACAGTCGCATACAGCGTCCTGGTCGTCGCAAAGGTC  
GAAAAGGGAAAGAGCAAGAAGCTGAAGAGCGTCAAGGAACTGCTGGGAATCACAA  
TCATGGAAAGAAGCAGCTTCGAAAAGAACCCGATCGACTTCTGGAAGCAAAGGGA  
TACAAGGAAGTCAAGAAGGACCTGATCATCAAGCTGCCGAAGTACAGCCTGTTCGA  
ACTGGAAAACGGAAGAAAGAGAATGCTGGCAAGCGCAGGAGAACTGCAGAAGGGA  
AACGAACTGGCACTGCCGAGCAAGTACGTCAACTTCCTGTACCTGGCAAGCCACTA  
CGAAAAGCTGAAGGGAAGCCCGGAAGACAACGAACAGAAGCAGCTGTTCGTCGAA  
CAGCACAAGCACTACCTGGACGAAATCATCGAACAGATCAGCGAATTCAGCAAGAG  
AGTCATCCTGGCAGACGCAAACCTGGACAAGGTCCTGAGCGCATAACAAGCACA  
GAGACAAGCCGATCAGAGAACAGGCAGAAAACATCATCCACCTGTTCACACTGACA  
AACCTGGGAGCACCGGCAGCATTCAAGTACTTCGACACAACAATCGACAGAAAGAG  
ATACACAAGCACAAAGGAAGTCCTGGACGCAACACTGATCCACCAGAGCATCACAG  
GACTGTACGAAACAAGAATCGACCTGAGCCAGCTGGGAGGAGACGGAGGAGGAAG  
CCCGAAGAAGAAGAGAAAGGTCTAG

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

GGGUCCCGCAGUCGGCGUCCAGCGGCUCUGCUUGUUCGUGUGUGUGUCGU  
UGCAGGCCUUAUUCGGAUCCGCCACCAUGGACAAGAAGUACAGCAUCGGACUGGA  
CAUCGGAACAAACAGCGUCGGAUUGGGCAGUCAUCACAGACGAAUACAAGGUCCCG  
AGCAAGAAGUUCAAGGUCCUGGGAAACACAGACAGACACAGCAUCAAGAAGAAC  
CUGAUCGGAGCACUGCUGUUCGACAGCGGAGAAACAGCAGAAGCAACAAGACUG  
AAGAGAACAGCAAGAAGAAGAUACACAAGAAGAAAGAACAGAAUCUGCUACCUG  
CAGGAAAUCUUCAGCAACGAAAUGGCAAAGGUCGACGACAGCUUCUCCACAGAC  
UGGAAGAAAGCUUCCUGGUCGAAGAAGACAAGAAGCACGAAAGACACCCGAUCU  
UCGGAACAUCGUCGACGAAGUCGCAUACCACGAAAAGUACCCGACAUCUACCA  
CCUGAGAAAGAAGCUGGUCGACAGCACAGACAAGGCAGACCUGAGACUGAUCUAC  
CUGGCACUGGCACACAUGAUCAGUUCAGAGGACACUCCUGAUCGAAGGAGACC  
UGAACCCGGACAACAGCGACGUCGACAAGCUGUUCAUCCAGCUGGUCCAGACAUA

CAACCAGCUGUUCGAAGAAAACCCGAUCAACGCAAGCGGAGUCGACGCAAAGGCA  
AUCCUGAGCGCAAGACUGAGCAAGAGCAGAAGACUGGAAAACCUGAUCGCACAGC  
UGCCGGGAGAAAAGAAGAACGGACUGUUCGGAAACCUGAUCGCACUGAGCCUGG  
GACUGACACCGAACUUCAAGAGCAACUUCGACCUGGCAGAAGACGCAAAGCUGCA  
GCUGAGCAAGGACACAUAACGACGACGACCUGGACAACCUGGCUGGCACAGAUCGGA  
GACCAGUACGCAGACCUGUUCUGGCAGCAAAGAACCUGAGCGACGCAAUCCUGC  
UGAGCGACAUCUGAGAGUCAACACAGAAAUCACAAAGGCACCGCUGAGCGCAAAG  
CAUGAUCAAGAGAUACGACGAACACCACCAGGACCUGACACUGCUGAAGGCACUG  
GUCAGACAGCAGCUGCCGGAAAAGUACAAGGAAAUCUUCUUCGACCAGAGCAAAG  
AACGGAUACGCAGGAUACAUCGACGGAGGAGCAAGCCAGGAAGAAUUCUACAAG  
UUCAUCAAGCCGAUCCUGGAAAAGAUGGACGGAAACAGAAGAACUGCUGGUCAAG  
CUGAACAGAGAAGACCUGCUGAGAAAGCAGAGAACAUUCGACAACGGAAGCAUC  
CCGCACCAGAUCCACCUGGGAGAACUGCACGCAAUCCUGAGAAGACAGGAAGACU  
UCUACCCGUUCCUGAAGGACAACAGAGAAAAGAUCGAAAAGAUCUGACAUCU  
GAAUCCCGUACUACGUCGGACCGCUGGCAAGAGGAAACAGCAGAUUCGCAUGGAU  
GACAAGAAAGAGCGAAGAAACAUCACACCGUGGAACUUCGAAGAAGUCGUCGA  
CAAGGGAGCAAGCGCACAGAGCUUCAUCGAAAGAAUGACAAACUUCGACAAGAA  
CCUGCCGAACGAAAAGGUCCUGCCGAAGCACAGCCUGCUGUACGAAUACUUCACA  
GUCUACAACGAACUGACAAAGGUCAAGUACGUCACAGAAGGAAUGAGAAAGCCG  
GCAUUCUGAGCGGAGAACAGAAGAAGGCAAUCGUCGACCUGCUGUUCAAGACA  
AACAGAAAGGUCACAGUCAAGCAGCUGAAGGAAGACUACUUCAAGAAGAUUCGAA  
UGCUCGACAGCGUCGAAAUCAGCGGAGUCGAAGACAGAUUCAACGCAAGCCUGG  
GAACAUACCACGACCUGCUGAAGAUCAUCAAGGACAAGGACUUCUGGACAACGA  
AGAAAACGAAGACAUCUGGAAGACAUCGUCCUGACACUGACACUGUUCGAAGAC  
AGAGAAAUGAUUCGAAGAAAGACUGAAGACAUCGCACACCUGUUCGACGACAAG  
GUCAUGAAGCAGCUGAAGAGAAGAAGAUACACAGGAUGGGGAAGACUGAGCAGA  
AAGCUGAUCAACGGAAUCAGAGACAAGCAGAGCGGAAAGACAAUCCUGGACUUC  
CUGAAGAGCGACGGAUUCGCAAACAGAAACUUCAUGCAGCUGAUCCACGACGACA  
GCCUGACAUCUAAAGGAAGACAUCAGAAAGGCACAGGUCAGCGGACAGGGAGACA  
GCCUGCACGAACACAUCGCAAACCUGGCAGGAAGCCCGGCAAUCAAGAAGGGAAU  
CCUGCAGACAGUCAAGGUCGUCGACGAACUGGUCAAGGUCAUGGGAAGACACAA  
GCCGGAAAACAUCGUCAUCGAAAUGGCAAGAGAAAACCAGACAACACAGAAGGG  
ACAGAAGAACAGCAGAGAAAGAAUGAAGAGAAUCGAAGAAGGAAUCAAGGAACU  
GGGAAGCCAGAUCCUGAAGGAACACCCGGUCGAAAACACACAGCUGCAGAACGAA  
AAGCUGUACCUGUACUACCUGCAGAACGGAAGAGACAUGUACGUCGACCAGGAAC  
UGGACAUCAACAGACUGAGCGACUACGACGUCGACCACAUCGUCCCGCAGAGCUU  
CCUGAAGGACGACAGCAUCGACAACAAGGUCCUGACAAGAAGCGACAAGAACAGA  
GGAAAGAGCGACAACGUCCCGAGCGAAGAAGUCGUCAAGAAGAUGAAGAACUAC  
UGGAGACAGCUGCUGAACGCAAAGCUGAUCACACAGAGAAAGUUCGACAACCUG  
ACAAAGGCAGAGAGAGGAGGACUGAGCGAACUGGACAAGGCAGGAUUCAUCAAG



**ФОРМУЛА ИЗОБРЕТЕНИЯ**

1. Способ вставки нуклеиновой кислоты, кодирующей гетерологичный полипептид, в локус альбумина клетки-хозяина или популяции клеток, включающий введение

- i) гРНК, которая содержит последовательность, выбранную из
  - a) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32 и 33;
  - b) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32 и 33;
  - c) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 и 97;
  - d) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33;
  - e) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33;
  - f) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34-97;
  - g) последовательности, комплементарной 15 последовательным нуклеотидам +/- 10 нуклеотидам геномных координат, приведенных для SEQ ID NO: 2-33;
  - h) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 98-119;
  - i) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 98-119; и
  - j) последовательности, выбранной из группы, состоящей из SEQ ID NO: 120-163;
  - ii) РНК-направляемого ДНК-связывающего агента; и
  - iii) конструкции, содержащей нуклеиновую кислоту, кодирующую гетерологичный полипептид,

с осуществлением, таким образом, вставки нуклеиновой кислоты, кодирующей гетерологичный полипептид, в локус альбумина клетки-хозяина или популяции клеток.

2. Способ экспрессии гетерологичного полипептида из локуса альбумина клетки-хозяина или популяции клеток, включающий введение

- i) гРНК, которая содержит последовательность, выбранную из
  - a) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32 и 33;
  - b) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32 и 33;
  - c) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 и 97;
  - d) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33;

e) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33;

f) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34-97; и

g) последовательности, которая содержит 15 последовательных нуклеотидов +/- 10 нуклеотидов геномных координат, приведенных для SEQ ID NO: 2-33;

ii) РНК-направляемого ДНК-связывающего агента; и

iii) конструкции, содержащей кодирующую последовательность гетерологичного полипептида,

с осуществлением, таким образом, экспрессии гетерологичного полипептида в клетке-хозяине или популяции клеток.

3. Способ экспрессии терапевтического агента в клетке или популяции клеток неделящегося типа, включающий введение

i) гРНК, которая содержит последовательность, выбранную из

a) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32 и 33;

b) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2, 8, 13, 19, 28, 29, 31, 32 и 33;

c) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34, 40, 45, 51, 60, 61, 63, 64, 65, 66, 72, 77, 83, 92, 93, 95, 96 и 97;

d) последовательности, которая по меньшей мере на 95%, 90%, 85%, 80% или 75% идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33;

e) по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33;

f) последовательности, выбранной из группы, состоящей из SEQ ID NO: 34-97; и

g) последовательности, которая содержит 15 последовательных нуклеотидов +/- 10 нуклеотидов геномных координат, приведенных для SEQ ID NO: 2-33;

ii) РНК-направляемого ДНК-связывающего агента; и

iii) конструкции, содержащей кодирующую последовательность гетерологичного полипептида,

с осуществлением, таким образом, экспрессии терапевтического агента в клетке или популяции клеток неделящегося типа.

4. Способ по любому одному из пп. 1-3, отличающийся тем, что гРНК содержит направляющую последовательность, выбранную из группы, состоящей из SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32 и SEQ ID NO: 33.

5. Способ по любому одному из пп. 1-4, отличающийся тем, что способ осуществляют *in vivo*.

6. Способ по любому одному из пп. 1-4, отличающийся тем, что способ осуществляют *in vitro*.

7. Способ по любому одному из пп. 1-6, отличающийся тем, что гРНК связывает область выше мотива, примыкающего к протоспейсеру (PAM - англ.: protospacer adjacent motif).

8. Способ по п. 7, отличающийся тем, что PAM выбран из NGG, NNGRRT, NNGRR(N), NNAGAAW, NNNNG(A/C)TT и NNNNRYAC.

9. Способ по любому одному из пп. 1-8, отличающийся тем, что гРНК представляет собой двойную гРНК (дгРНК).

10. Способ по любому из пп. 1-8, отличающийся тем, что гРНК представляет собой одинарную гРНК (огРНК).

11. Способ по п. 10, отличающийся тем, что огРНК содержит один или более модифицированных нуклеозидов.

12. Способ по любому одному из пп. 1-11, отличающийся тем, что РНК-направляемый ДНК-связывающий агент представляет собой Cas9 или нуклеиновую кислоту, кодирующую Cas9.

13. Способ по любому одному из пп. 1-12, отличающийся тем, что РНК-направляемый ДНК-связывающий агент представляет собой нуклеиновую кислоту, кодирующую РНК-направляемый ДНК-связывающий агент.

14. Способ по п. 13, отличающийся тем, что нуклеиновая кислота, кодирующая РНК-направляемый ДНК-связывающий агент, представляет собой мРНК.

15. Способ по п. 14, отличающийся тем, что мРНК представляет собой модифицированную мРНК.

16. Способ по любому одному из пп. 1-15, отличающийся тем, что РНК-направляемый ДНК-связывающий агент представляет собой Cas-нуклеазу или нуклеиновую кислоту, кодирующую Cas-нуклеазу.

17. Способ по п. 16, отличающийся тем, что Cas-нуклеаза представляет собой Cas-нуклеазу класса 2.

18. Способ по п. 16 или п. 17, отличающийся тем, что Cas-нуклеаза выбрана из группы, состоящей из нуклеазы *S. pyogenes*, нуклеазы *S. aureus*, нуклеазы *C. jejuni*, нуклеазы *S. thermophilus*, нуклеазы *N. meningitidis* и их вариантов.

19. Способ по любому одному из пп. 16-18, отличающийся тем, что Cas-нуклеаза представляет собой Cas9.

20. Способ по п. 19, отличающийся тем, что Cas-нуклеаза представляет собой Cas9-нуклеазу *S. pyogenes*.

21. Способ по любому одному из пп. 16-20, отличающийся тем, что Cas-нуклеаза обладает сайт-специфической ДНК-связывающей активностью.

22. Способ по любому одному из пп. 16-21, отличающийся тем, что Cas-нуклеаза

представляет собой никазу.

23. Способ по любому одному из пп. 16-21, отличающийся тем, что Cas-нуклеаза представляет собой кливазу (cleavase).

24. Способ по любому одному из пп. 19-21, отличающийся тем, что Cas-нуклеаза не обладает активностью никазы или кливазы.

25. Способ по любому одному из пп. 19-24, отличающийся тем, что конструкция нуклеиновой кислоты представляет собой независимую от гомологии донорную конструкцию.

26. Способ по любому одному из пп. 1-25, отличающийся тем, что конструкция представляет собой двунаправленную конструкцию нуклеиновой кислоты.

27. Способ по п. 26, отличающийся тем, что конструкция содержит  
i. первый сегмент, содержащий кодирующую последовательность гетерологичного полипептида; и  
ii. второй сегмент, содержащий обратную комплементарную последовательность кодирующей последовательности гетерологичного полипептида.

28. Способ по любому одному из пп. 1-27, отличающийся тем, что конструкция содержит сигнал полиаденилирования.

29. Способ по любому одному из пп. 1-28, отличающийся тем, что конструкция содержит сайт акцептора сплайсинга.

30. Способ по любому одному из пп. 1-29, отличающийся тем, что конструкция не содержит плечо гомологии.

31. Способ по любому одному из пп. 1-30, отличающийся тем, что гРНК вводят в векторе и/или липидной наночастице.

32. Способ по любому одному из пп. 1-31, отличающийся тем, что РНК-направляемый ДНК-связывающий агент вводят в векторе и/или липидной наночастице.

33. Способ по любому одному из пп. 1-32, отличающийся тем, что конструкцию, содержащую гетерологичный ген, вводят в векторе и/или липидной наночастице.

34. Способ по любому одному из пп. 31-33, отличающийся тем, что вектор представляет собой вирусный вектор.

35. Способ по п. 34, отличающийся тем, что вирусный вектор выбран из группы, состоящей из аденоассоциированного вирусного (AAV) вектора, аденовирусного вектора, ретровирусного вектора и лентивирусного вектора.

36. Вектор по 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.

37. Способ по любому одному из пп. 1-36, отличающийся тем, что гРНК, РНК-направляемый ДНК-связывающий агент и конструкцию, содержащую кодирующую последовательность гетерологичного полипептида, отдельно или в любой комбинации, вводят одновременно.



38. Способ по любому одному из пп. 1-36, отличающийся тем, что гРНК, РНК-направляемый ДНК-связывающий агент и конструкцию, содержащую кодирующую последовательность гетерологичного полипептида, вводят последовательно, в любом порядке и/или в любой комбинации.

39. Способ по любому одному из пп. 1-36 и 38, отличающийся тем, что РНК-направляемый ДНК-связывающий агент или РНК-направляемый ДНК-связывающий агент и гРНК в комбинации вводят до введения конструкции.

40. Способ по любому одному из пп. 1-36 и 38, отличающийся тем, что конструкцию, содержащую кодирующую последовательность гетерологичного полипептида, вводят до гРНК и/или РНК-направляемого ДНК-связывающего агента.

41. Способ по любому одному из пп. 1-40, отличающийся тем, что двунаправленную конструкцию нуклеиновой кислоты, РНК-направляемый ДНК-связывающий агент и гРНК, в любой комбинации, вводят с интервалом не более часа.

42. Способ по любому одному из пп. 1-41, отличающийся тем, что гетерологичный полипептид представляет собой секретлируемый полипептид.

43. Способ по любому одному из пп. 1-41, отличающийся тем, что гетерологичный полипептид представляет собой внутриклеточный полипептид.

44. Способ по любому одному из пп. 1-43, отличающийся тем, что клетка представляет собой клетку печени.

45. Способ по п. 44, отличающийся тем, что клетка печени представляет собой гепатоцит.

46. Способ по любому одному из пп. 1-45, отличающийся тем, что экспрессия гетерологичного полипептида в клетке-хозяине повышена по меньшей мере на около 2%, около 3%, около 4%, около 5%, около 6%, около 7%, около 8%, около 9%, около 10% или более относительно уровня в клетке до введения гРНК, РНК-направляемого ДНК-связывающего агента и конструкции, содержащей кодирующую последовательность гетерологичного полипептида.

47. Способ по любому из пп. 1-46, отличающийся тем, что гРНК содержит SEQ ID NO: 301 или SEQ ID NO: 302.

48. Способ по любому одному из пп. 1-47, отличающийся тем, что гРНК опосредует мишень-специфическое разрезание РНК-направляемым ДНК-связывающим агентом, что приводит к вставке кодирующей последовательности гетерологичного полипептида в интрон I гена альбумина.

49. Способ по любому предшествующему пункту, отличающийся тем, что разрезание приводит к частоте вставки гетерологичной нуклеиновой кислоты в популяции клеток, составляющей по меньшей мере около 2%, около 5% или около 10%.

50. Способ по п. 49, отличающийся тем, что разрезание приводит к частоте вставки кодирующей последовательности гетерологичного полипептида между около 30 и 35%, около 35 и 40%, около 40 и 45%, около 45 и 50%, около 50 и 55%, около 55 и 60%, около 60 и 65%, около 65 и 70%, около 70 и 75%, около 75 и 80%, около 80 и 85%, около 85 и

90%, около 90 и 95% или около 95 и 99%.

51. Способ по любому предшествующему пункту, отличающийся тем, что РНК-направляемый ДНК-связывающий белок представляет собой Cas9-нуклеазу *S. pyogenes*.

52. Способ по п. 51, отличающийся тем, что нуклеаза представляет собой кливазу или никазу.

53. Способ по любому одному из пп. 1-52, дополнительно включающий введение ЛНЧ, содержащей гРНК.

54. Способ по любому одному из пп. 1-53, дополнительно включающий введение ЛНЧ, содержащей мРНК, которая кодирует РНК-направляемый ДНК-связывающий агент.

55. Способ по п. 54, отличающийся тем, что ЛНЧ содержит гРНК и мРНК, которая кодирует РНК-направляемый ДНК-связывающий агент.

56. Способ по любому одному из пп. 1-55, отличающийся тем, что гРНК и РНК-направляемый ДНК-связывающий белок вводят в виде РНП.

57. Способ по любому одному из пп. 53-56, отличающийся тем, что конструкцию вводят с помощью вектора.

58. Клетка-хозяин, полученная способом по любому одному из пп. 1-57.

59. Клетка-хозяин, содержащая двунаправленную конструкцию нуклеиновой кислоты, кодирующую гетерологичный полипептид, интегрированную с интроном 1 локуса альбумина клетки-хозяина.

60. Клетка-хозяин по п. 58 или п. 59, отличающаяся тем, что клетка-хозяин представляет собой клетку печени.

61. Клетка-хозяин по любому одному из пп. 58-60, отличающаяся тем, что клетка печени представляет собой гепатоцит.

62. Способ или клетка-хозяин по предшествующему пункту, отличающиеся тем, что РНК-направляемый ДНК-связывающий агент представляет собой нуклеиновую кислоту, кодирующую РНК-направляемый ДНК-связывающий агент.

63. Способ или клетка-хозяин по любому предшествующему пункту, отличающиеся тем, что РНК-направляемый ДНК-связывающий агент представляет собой Cas-нуклеазу.

64. Способ или клетка-хозяин по любому предшествующему пункту, отличающиеся тем, что РНК-направляемый ДНК-связывающий агент представляет собой нуклеиновую кислоту, кодирующую Cas-нуклеазу.

65. Способ или клетка-хозяин по любому предшествующему пункту, отличающиеся тем, что РНК-направляемый ДНК-связывающий агент представляет собой мРНК, которая кодирует Cas-нуклеазу.

66. Способ или клетка-хозяин по п. 65, отличающиеся тем, что мРНК представляет собой модифицированную мРНК.

67. Способ или клетка-хозяин по любому одному из пп. 63-66, отличающиеся тем, что Cas-нуклеаза представляет собой Cas-нуклеазу класса 2.

68. Способ или композиция по любому одному из пп. 63-67, отличающиеся тем, что Cas-нуклеаза представляет собой Cas9.

69. Способ или композиция по любому одному из пп. 63-68, отличающиеся тем, что Cas-нуклеаза выбрана из группы, состоящей из нуклеазы *S. pyogenes*, нуклеазы *S. aureus*, нуклеазы *C. jejuni*, нуклеазы *S. thermophilus*, нуклеазы *N. meningitidis* и их вариантов.

70. Способ или композиция по п. 69, отличающиеся тем, что Cas-нуклеаза представляет собой Cas9-нуклеазу *S. pyogenes* или ее вариант.

71. Способ или композиция по любому одному из пп. 63-70, отличающиеся тем, что Cas-нуклеаза обладает сайт-специфической ДНК-связывающей активностью.

72. Способ или композиция по любому одному из пп. 63-71, отличающиеся тем, что Cas-нуклеаза представляет собой никазу.

73. Способ или композиция по любому одному из пп. 63-72, отличающиеся тем, что Cas-нуклеаза представляет собой кливазу.

74. Способ по любому одному из пп. 63-73, отличающийся тем, что Cas-нуклеаза не обладает активностью никазы или кливазы.

75. Способ по любому одному из пп. 1-74, дополнительно включающий достижение активности гетерологичного полипептида или уровней гетерологичного полипептида, составляющих по меньшей мере около 1% от нормы, например, по меньшей мере около 5% от нормы.

76. Способ по любому одному из пп. 1-75, отличающийся тем, что активность гетерологичного полипептида или уровень гетерологичного полипептида составляют менее чем около 500% от нормы.

77. Способ по любому одному из пп. 1-76, дополнительно включающий достижение активности гетерологичного полипептида или уровней гетерологичного полипептида, составляющих по меньшей мере от около 1% до 300% от нормы.

78. Способ *in vivo* по любому одному из пп. 5-77, дополнительно включающий достижение у индивида продолжительного эффекта, например, эффекта, длящегося по меньшей мере 1 месяц, 2 месяца, 6 месяцев, 1 год или 2 года.

79. Способ *in vivo* по любому одному из пп. 5-78, отличающийся тем, что уровни циркулирующего альбумина у индивида являются нормальными по меньшей мере через 1 месяц, 2 месяца, 6 месяцев или 1 год после введения конструкции нуклеиновой кислоты.

80. Способ *in vivo* по любому одному из пп. 5-79, отличающийся тем, что уровни циркулирующего альбумина у индивида сохраняются через 4 недели после введения двунаправленной конструкции нуклеиновой кислоты.

81. Способ *in vivo* по любому одному из пп. 5-80, отличающийся тем, что уровни циркулирующего альбумина у индивида временно падают, а потом возвращаются к норме.

82. Способ по любому одному из пп. 1-81, отличающийся тем, что гидовая РНК содержит по меньшей мере 17, 18, 19 или 20 смежных нуклеотидов последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33.

83. Способ по любому одному из пп. 1-82, отличающийся тем, что гидовая РНК содержит последовательность, которая по меньшей мере на 95%, 90%, 85%, 80% или 75%

идентична последовательности, выбранной из группы, состоящей из SEQ ID NO: 2-33.

84.Способ по любому одному из пп. 1-83, отличающийся тем, что гидовая РНК содержит последовательность, выбранную из группы, состоящей из SEQ ID NO: 2-33.

85.Способ или клетка-хозяин по любому предшествующему пункту, отличающиеся тем, что гидовая РНК содержит SEQ ID NO: 301 или 302.

86.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 2.

87.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 3.

88.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 4.

89.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 5.

90.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 6.

91.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 7.

92.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 8.

93.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 9.

94.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 10.

95.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 11.

96.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 12.

97.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты



111.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 27.

112.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 28.

113.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 29.

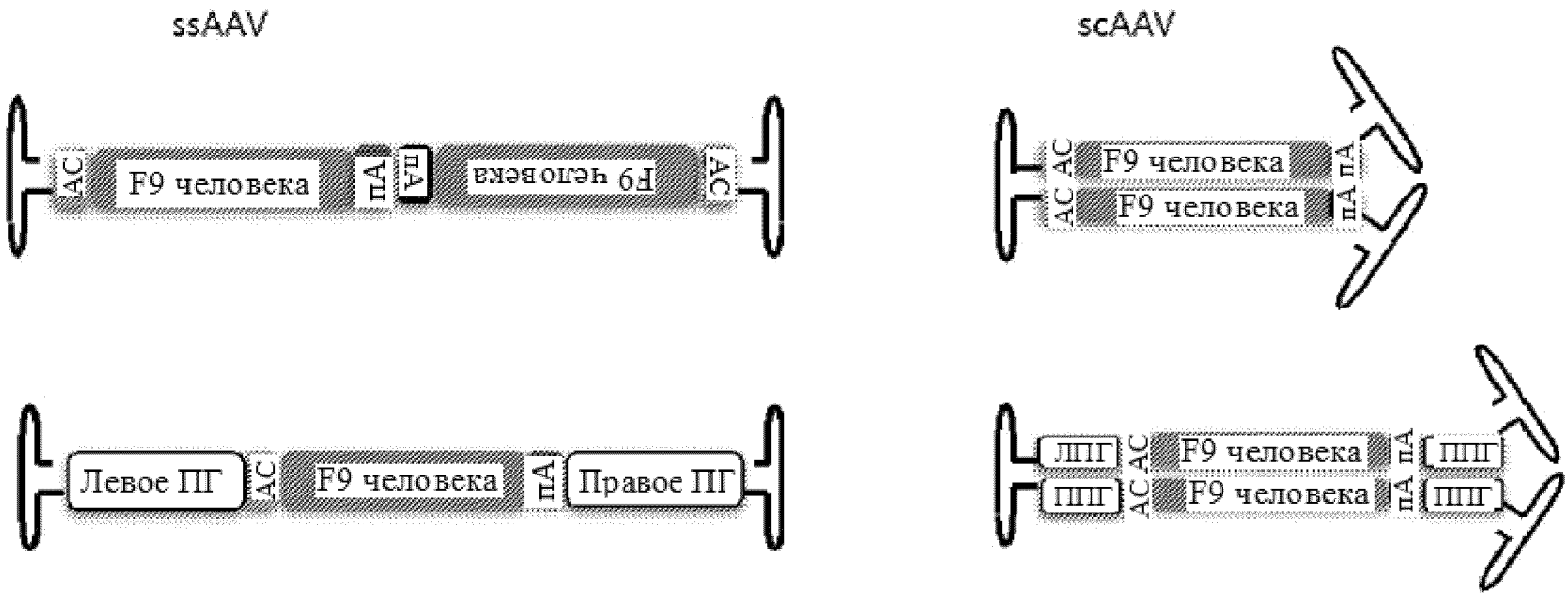
114.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 30.

115.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 31.

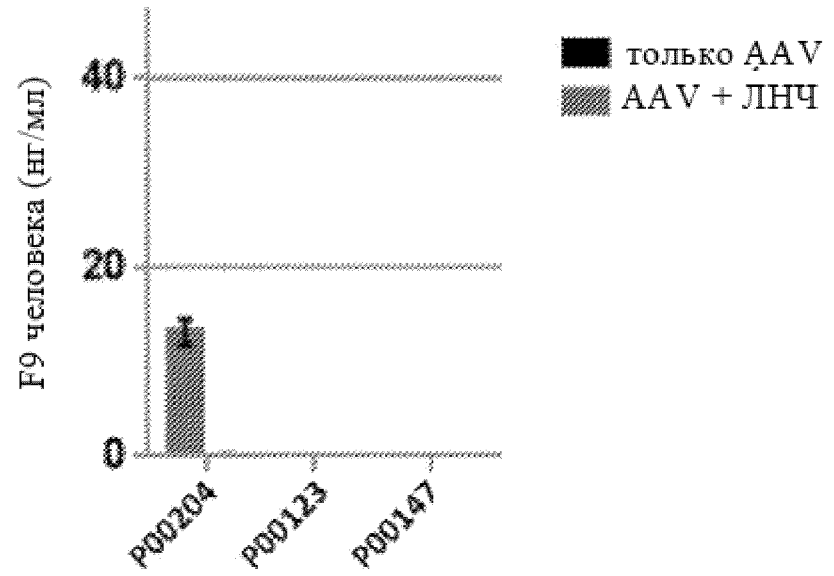
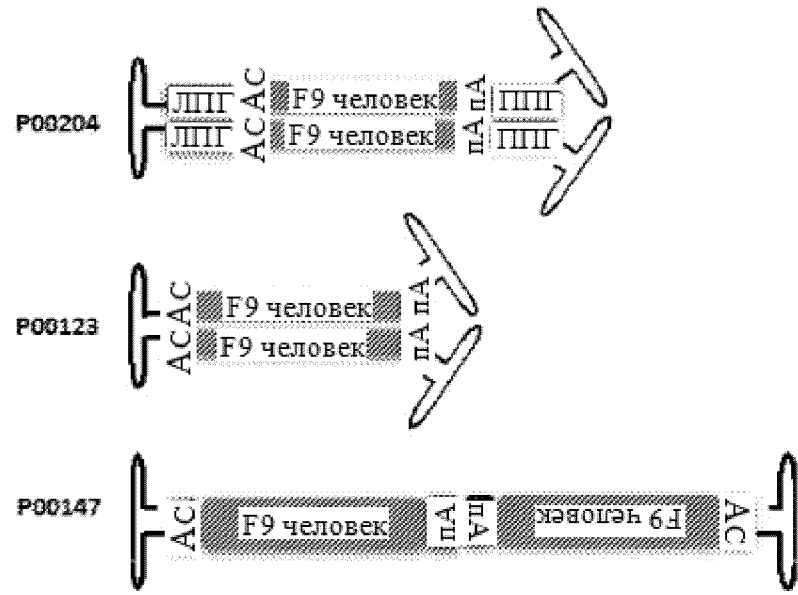
116.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 32.

117.Способ и композиция по любому одному из предшествующих пунктов, отличающиеся тем, что гидовая РНК содержит последовательность нуклеиновой кислоты SEQ ID NO: 33.

По доверенности

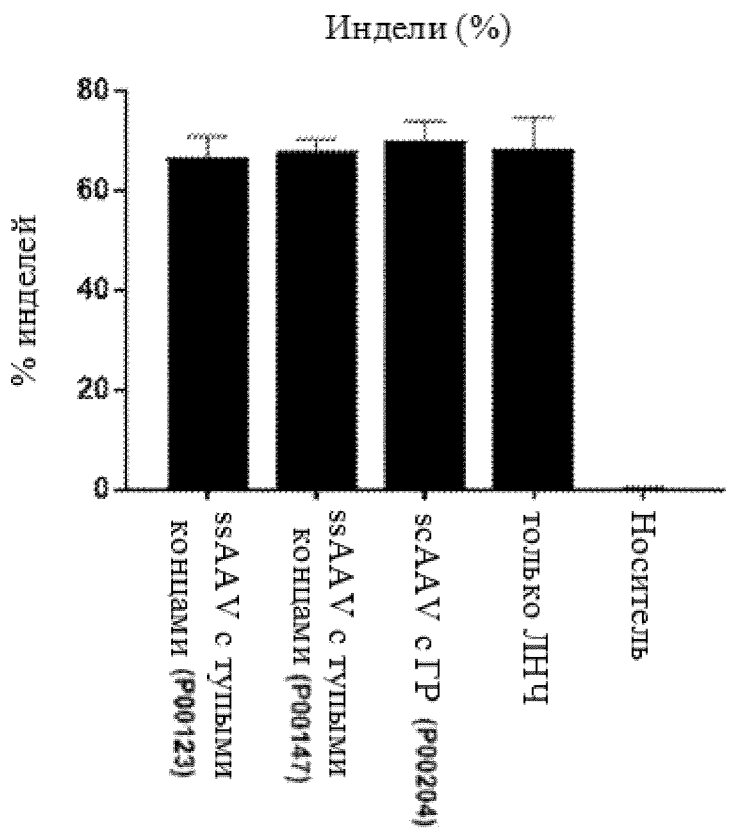


Фиг. 1

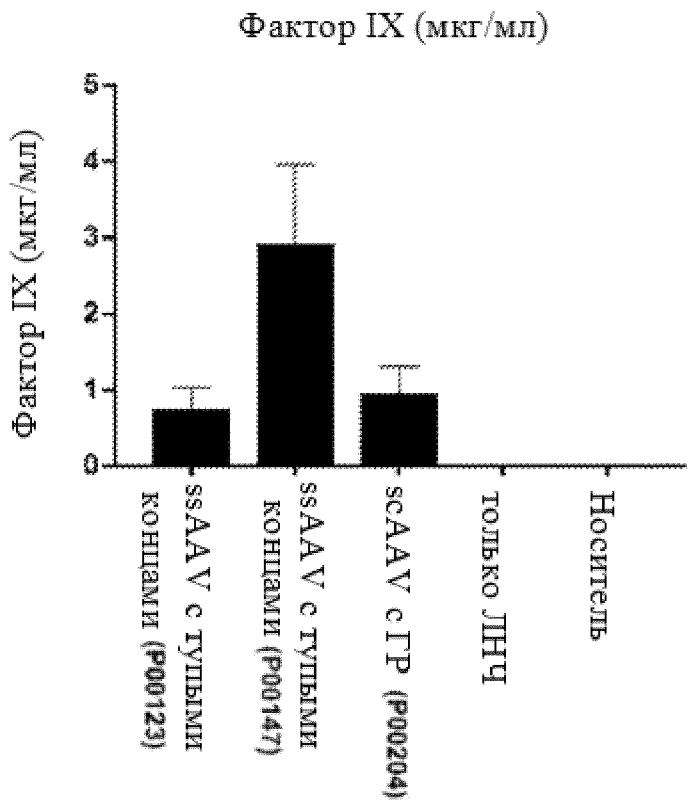


Фиг. 2

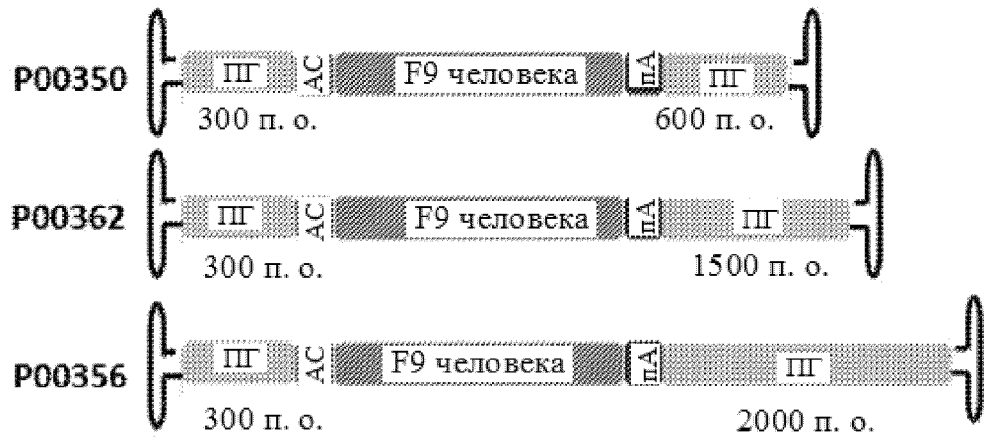
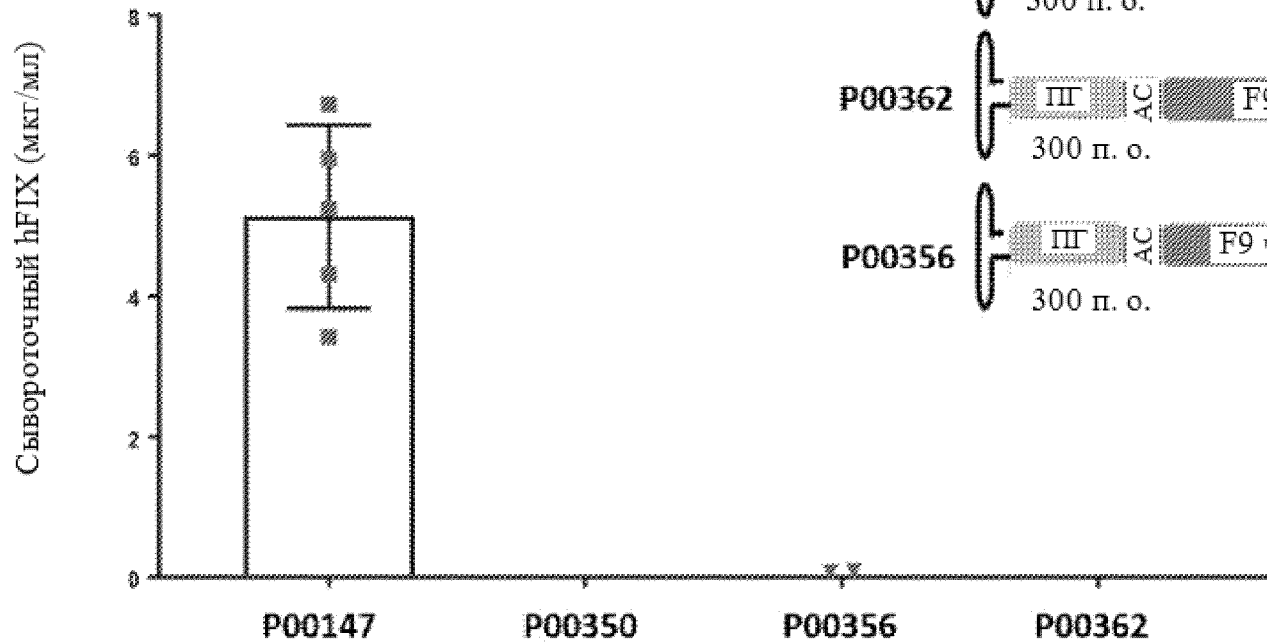




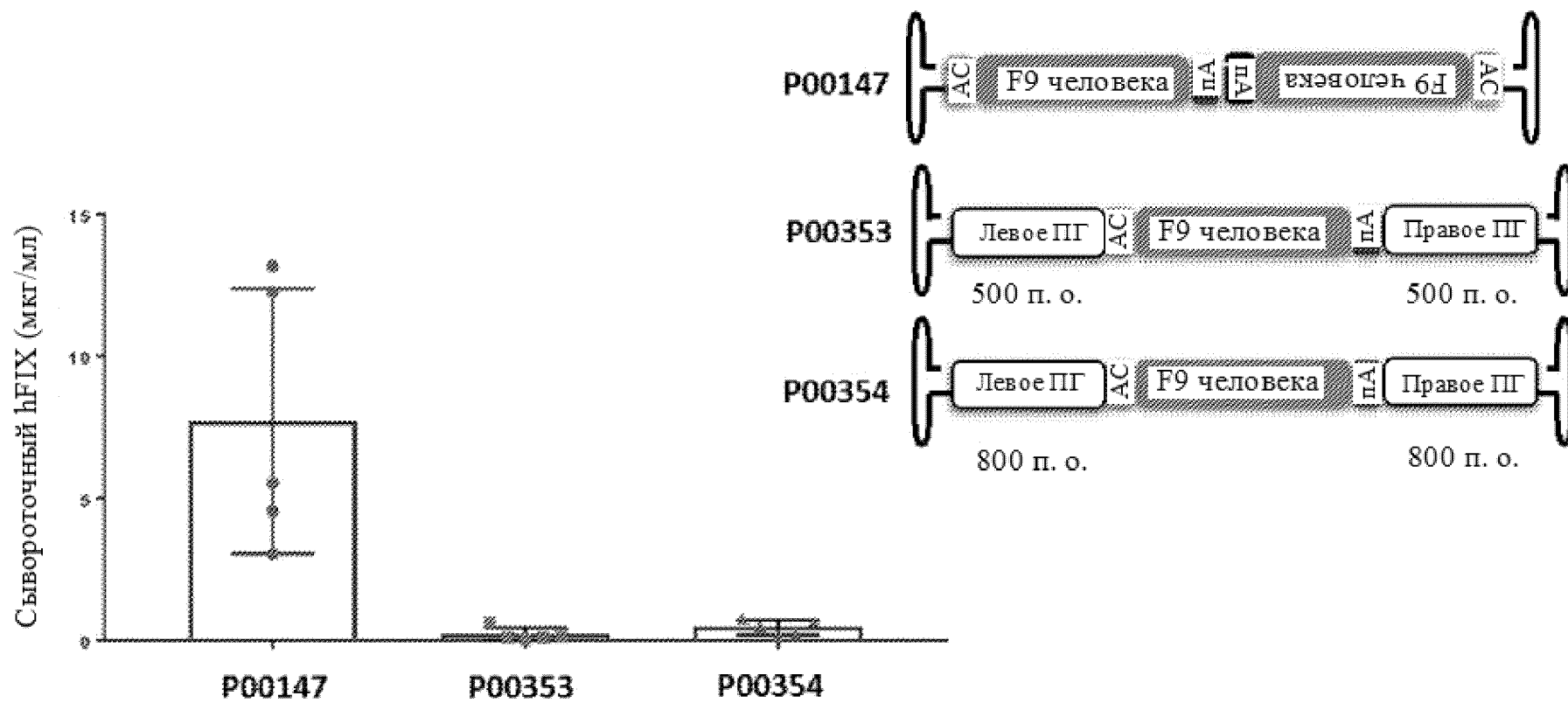
Фиг. 3А



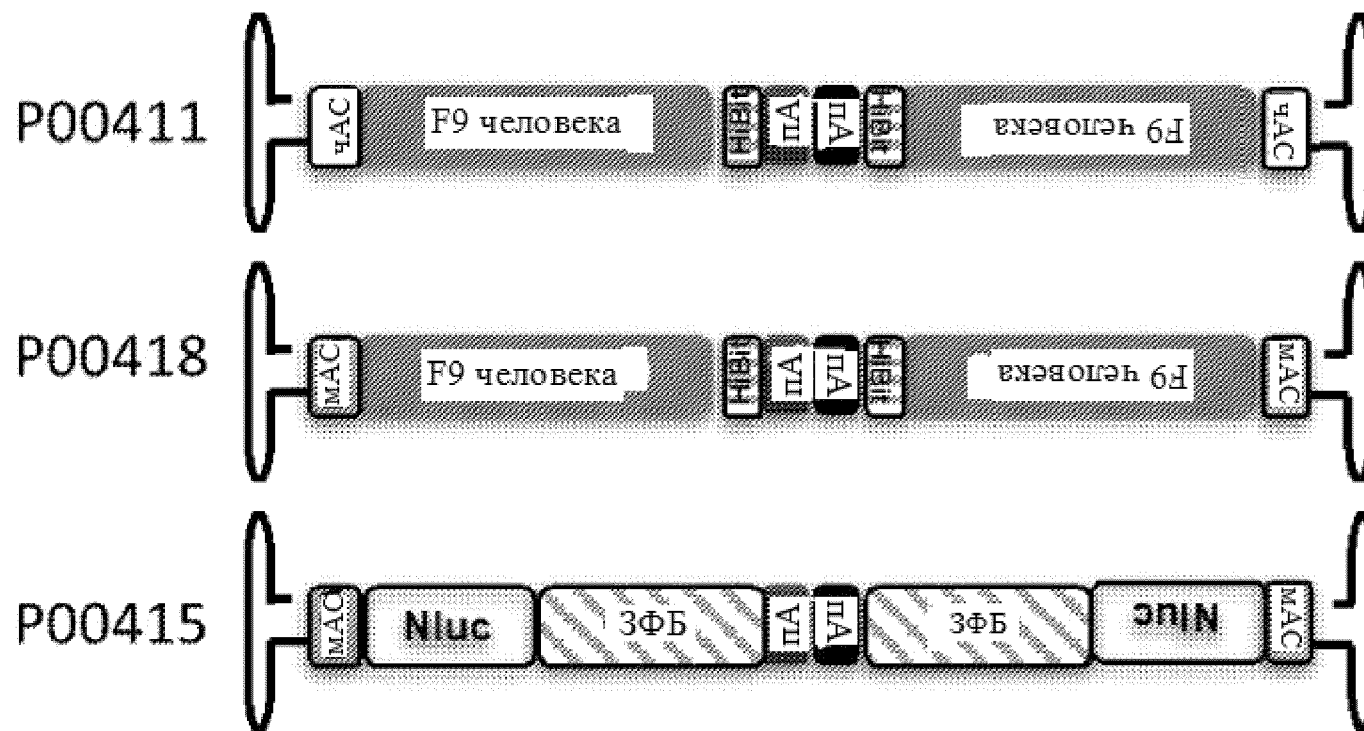
Фиг. 3В



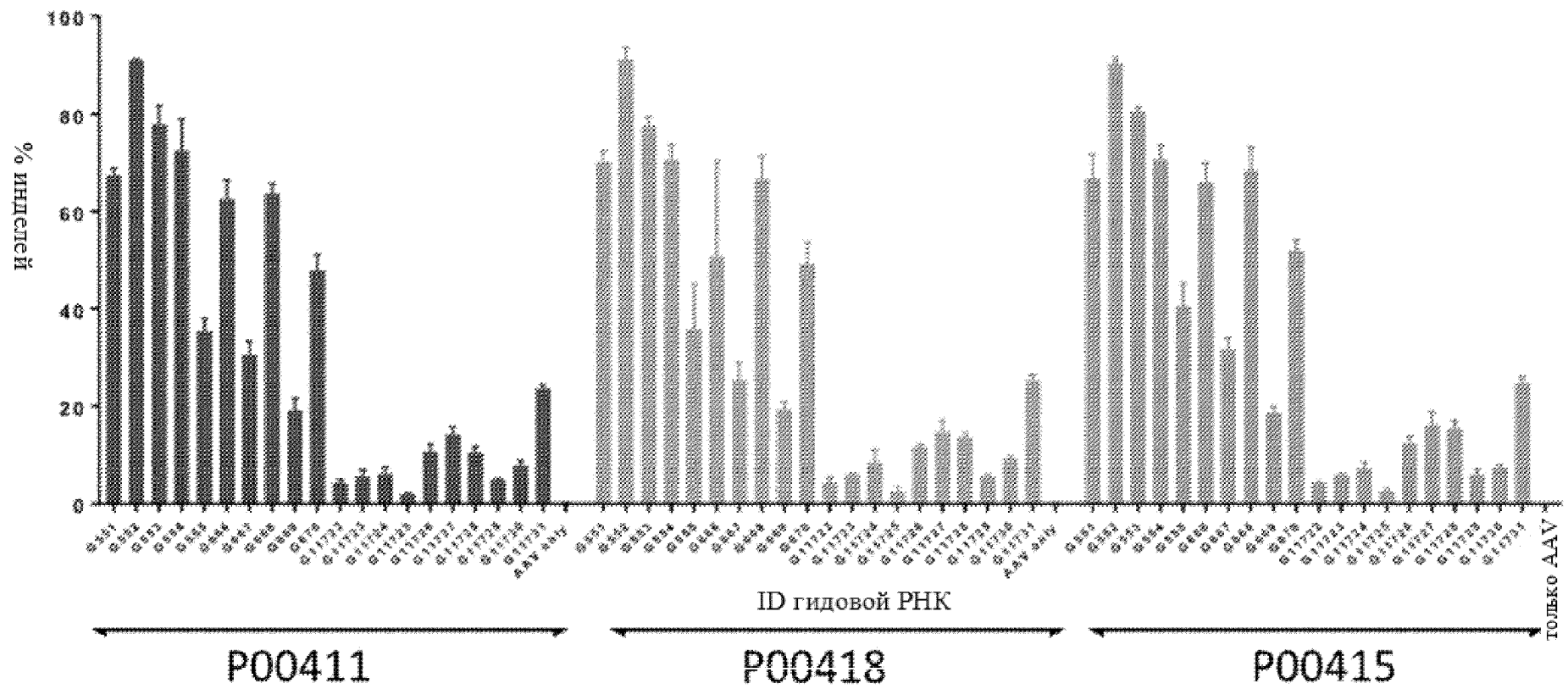
Фиг. 4А



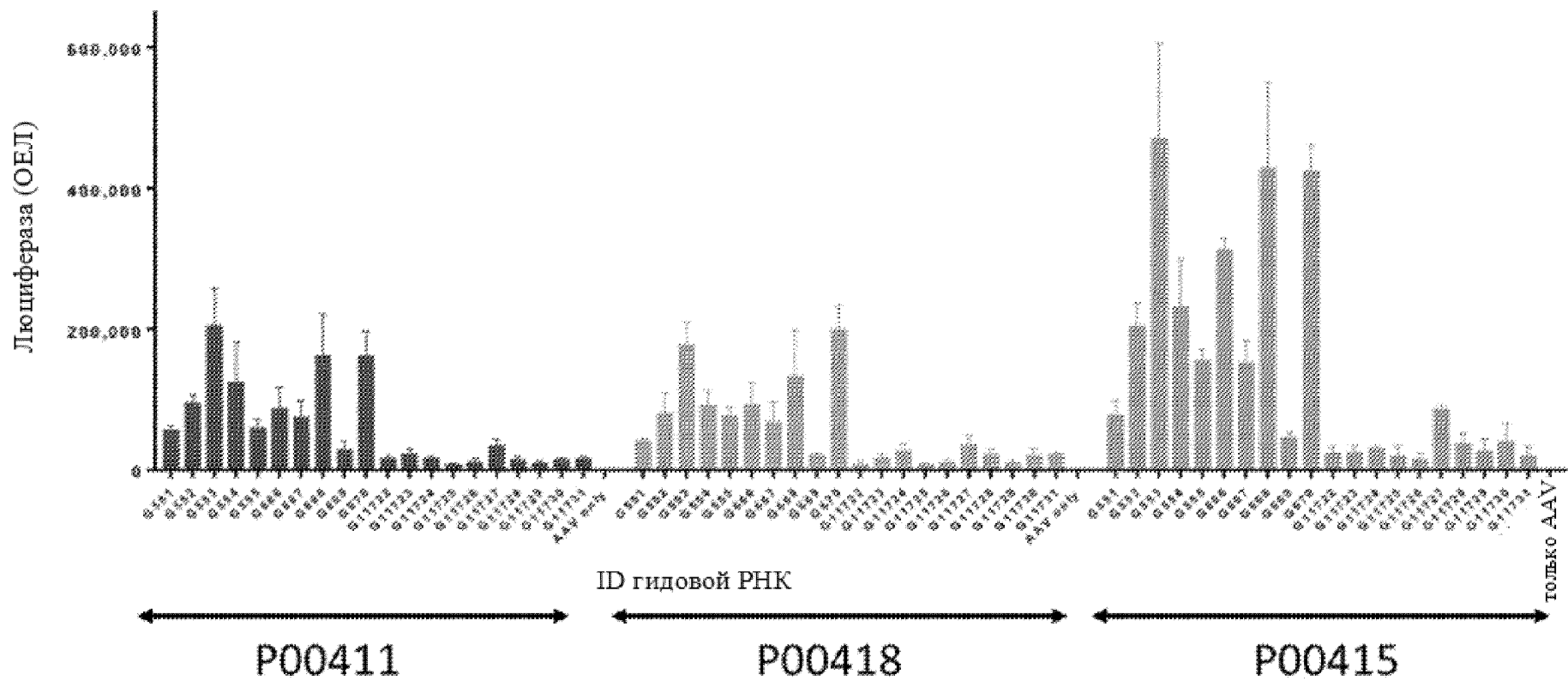
Фиг. 4В



Фиг. 5А

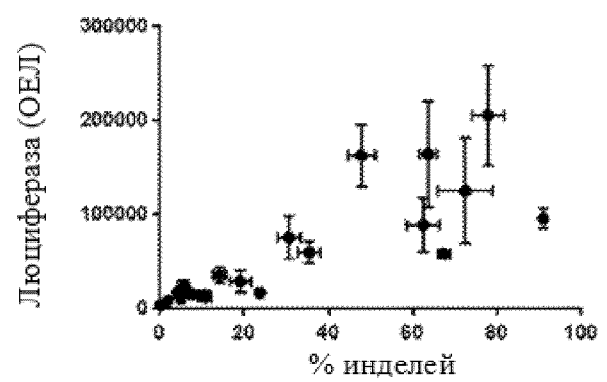


Фиг. 5В



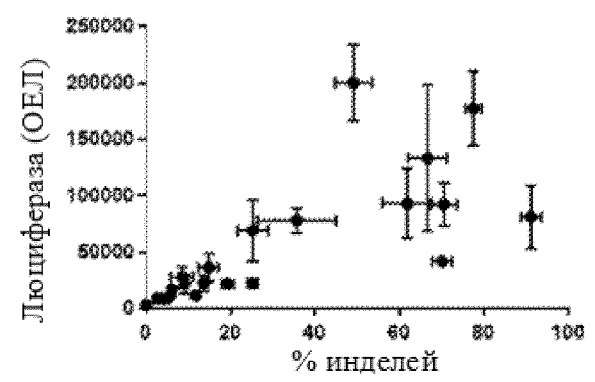
ФИГ. 5С

Значение  $R^2 = 0,54$   
(когда не включены  
направляющие с менее  
чем 10% редактирования)



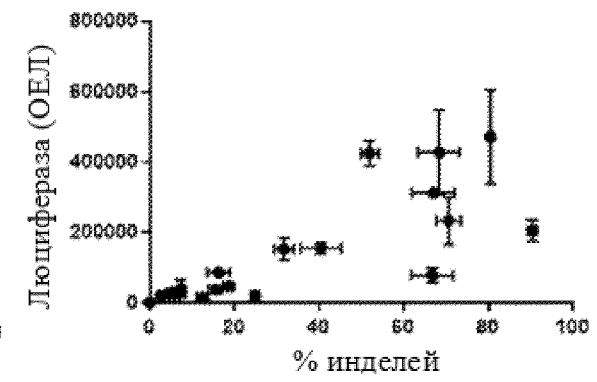
R00411

Значение  $R^2 = 0,37$   
(когда не включены  
направляющие с менее  
чем 10% редактирования)



R00418

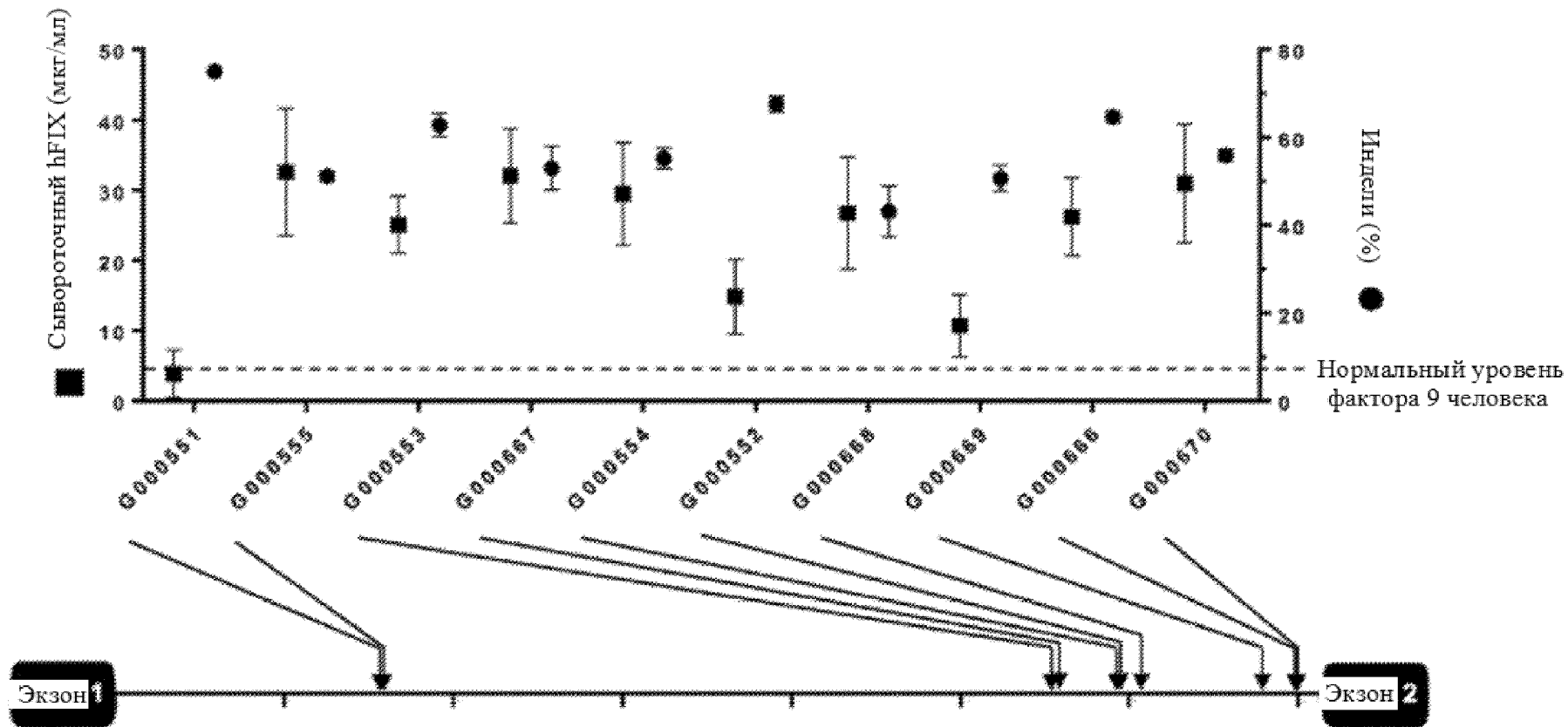
Значение  $R^2 = 0,50$   
(когда не включены  
направляющие с менее  
чем 10% редактирования)



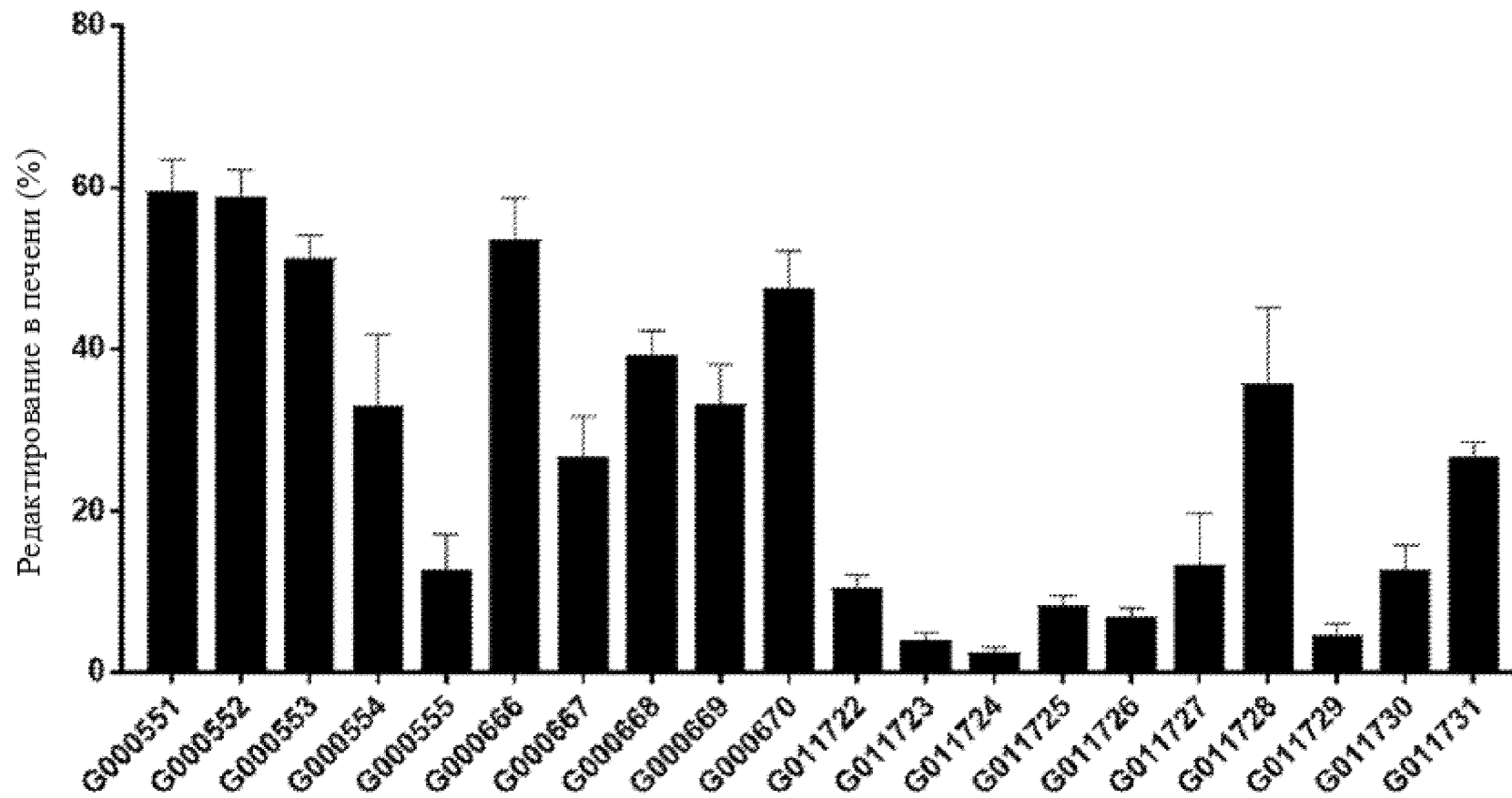
R00415

Фиг. 5D

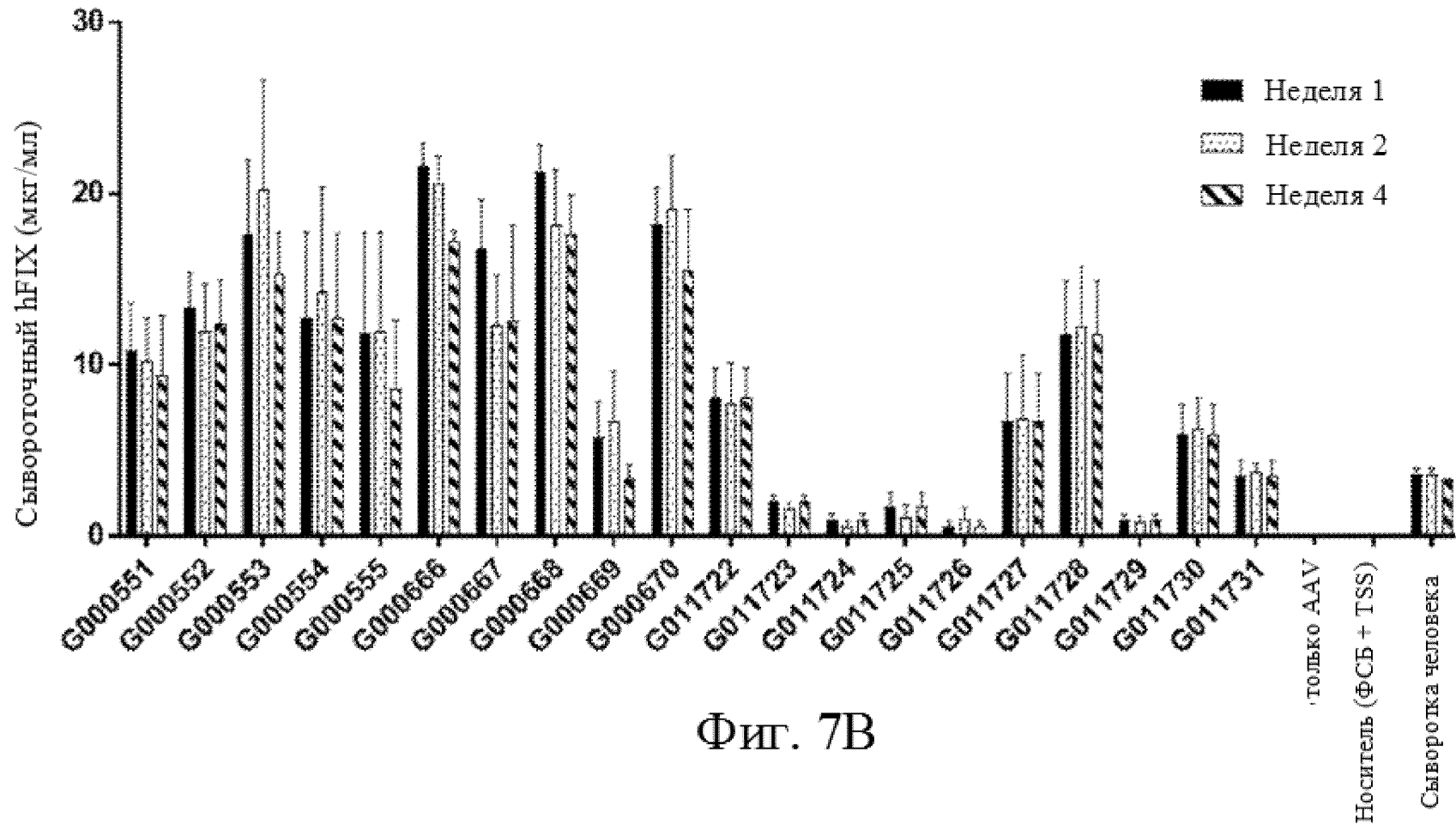
Фиг. 6



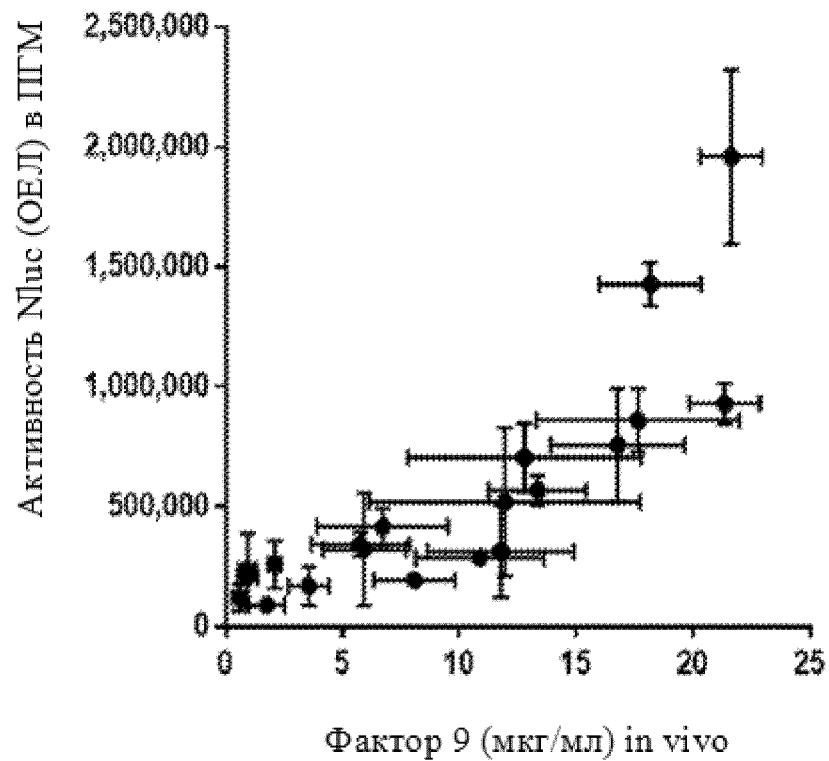




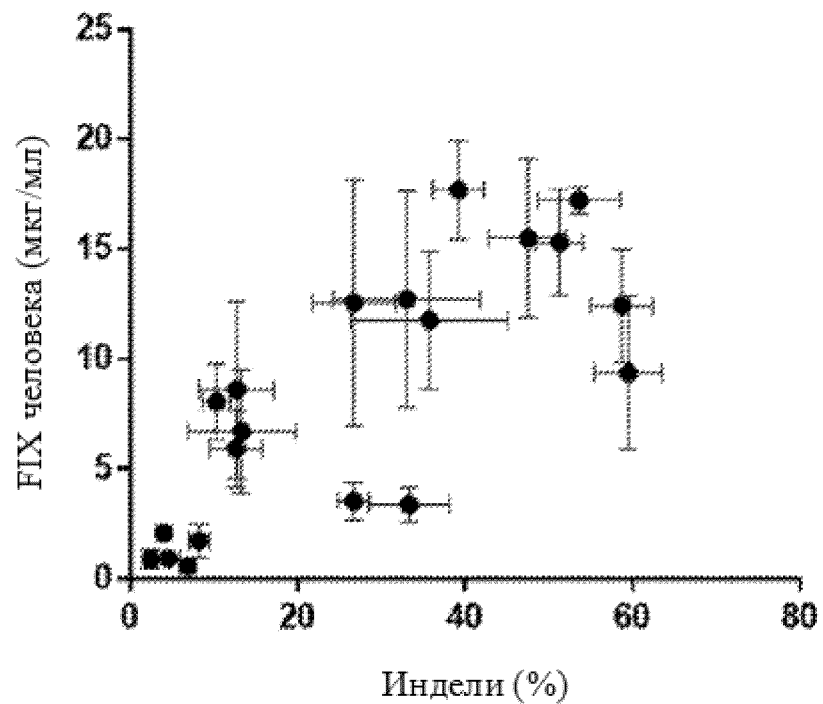
Фиг. 7А



Фиг. 7В

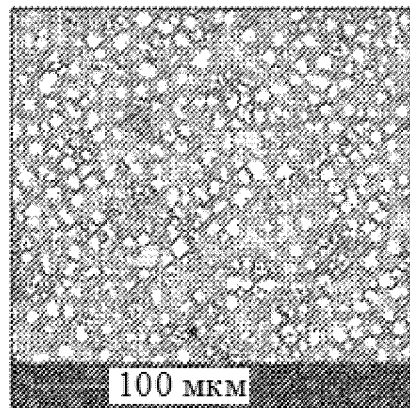


Фиг. 7C



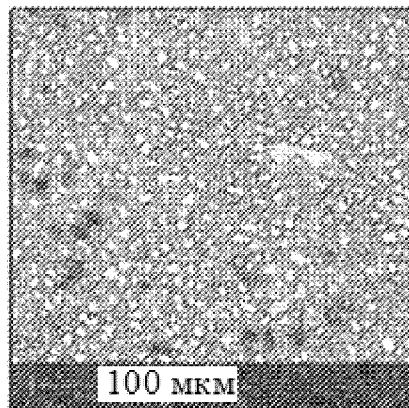
Фиг. 7D

только AAV



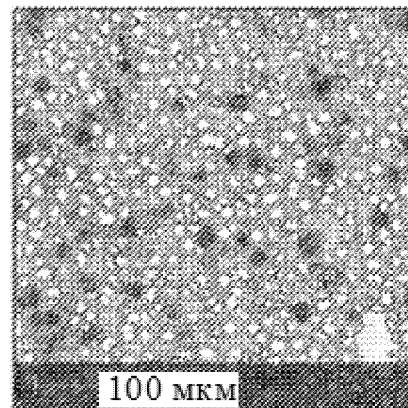
Менее 1,0% клеток,  
положительных в отношении  
гибридного транскрипта

G011723



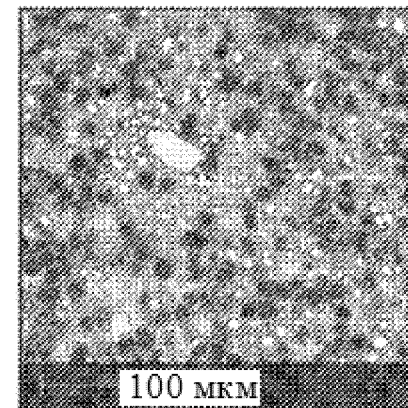
4,9% клеток, положительных  
в отношении гибридного  
транскрипта

G000551



19,8% клеток,  
положительных в  
отношении гибридного  
транскрипта

G000666

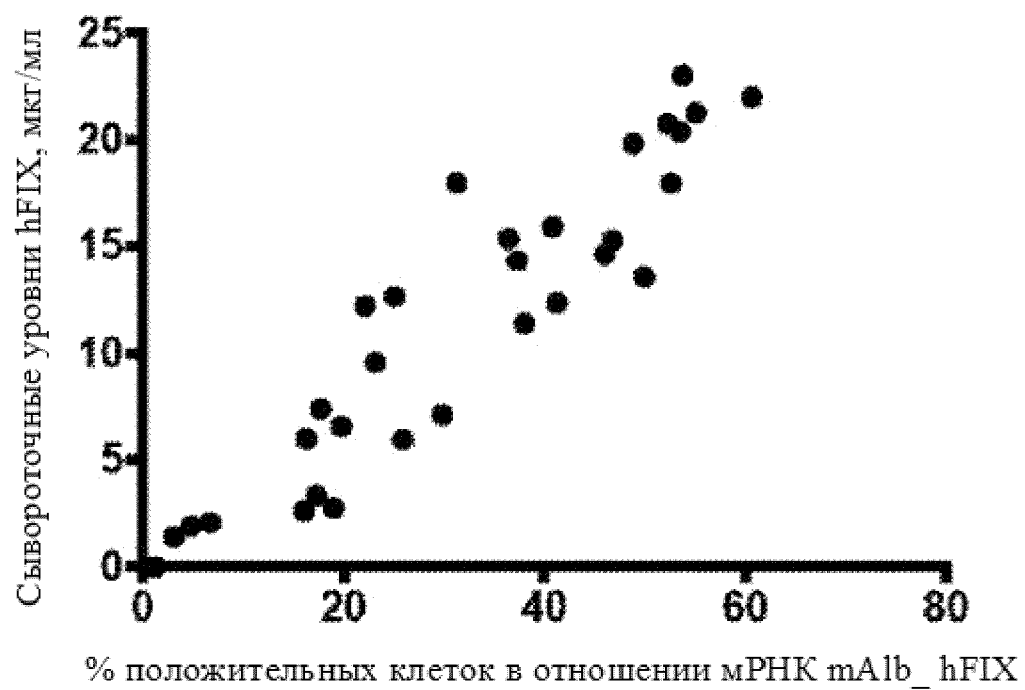


52,3% клеток,  
положительных в  
отношении гибридного  
транскрипта

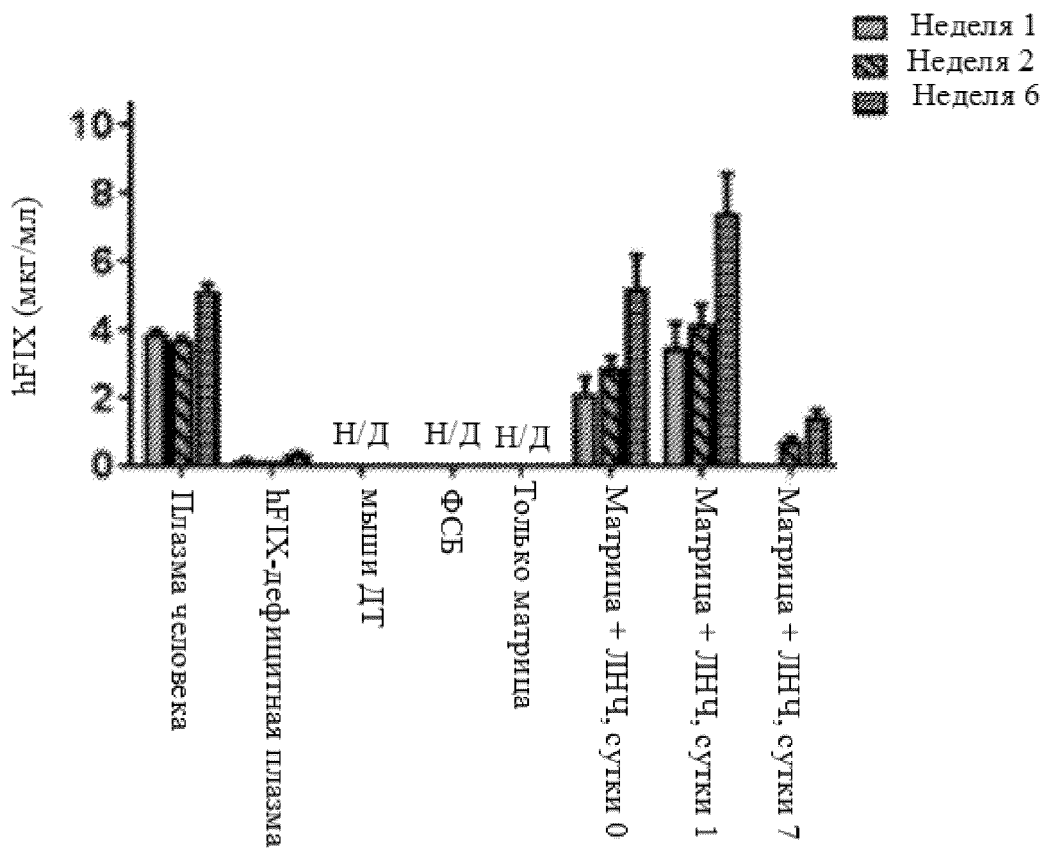
Фиг. 8А

Фиг. 8В

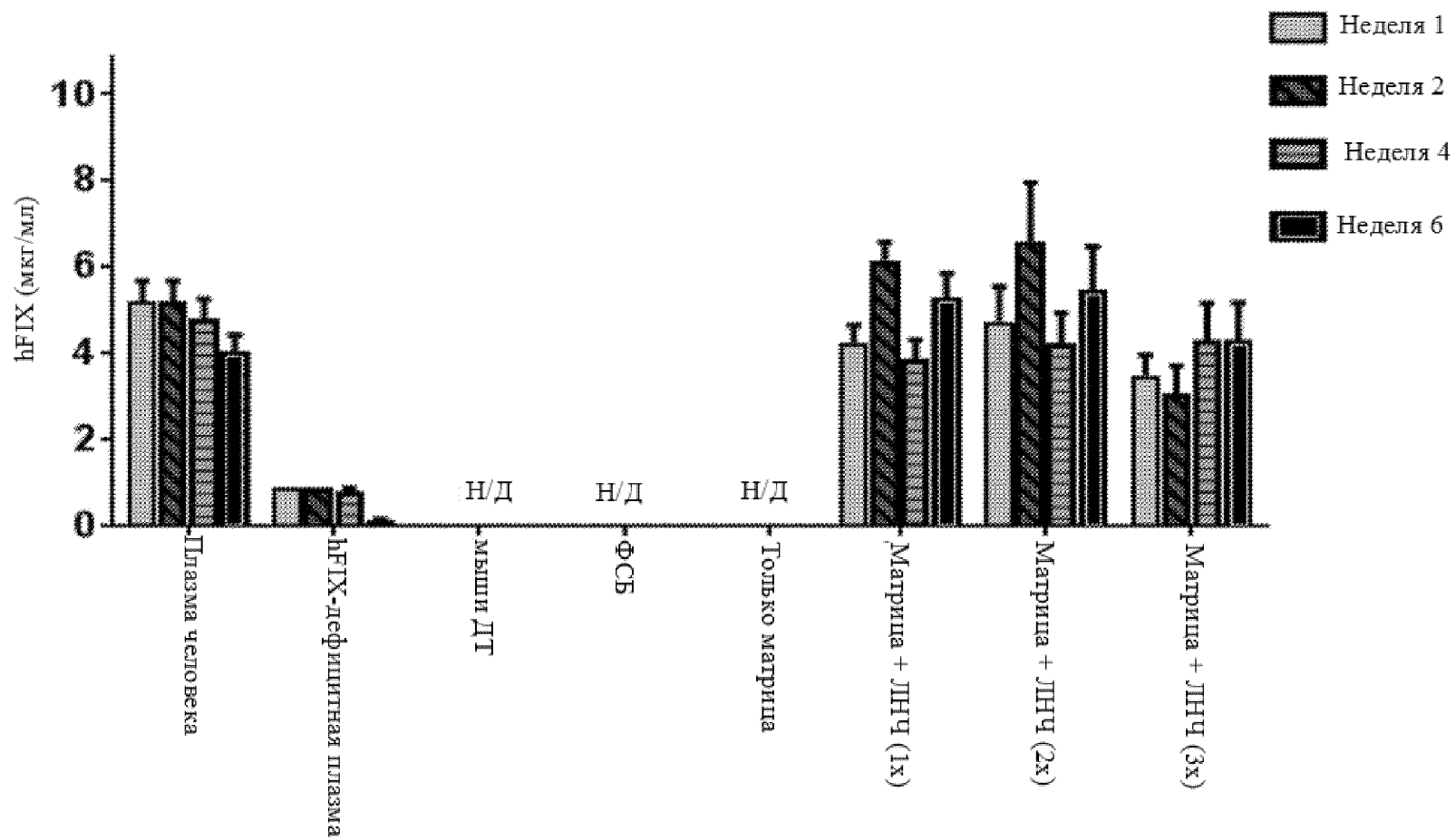
Уровни hFIX, неделя 1 vs. % положительных клеток



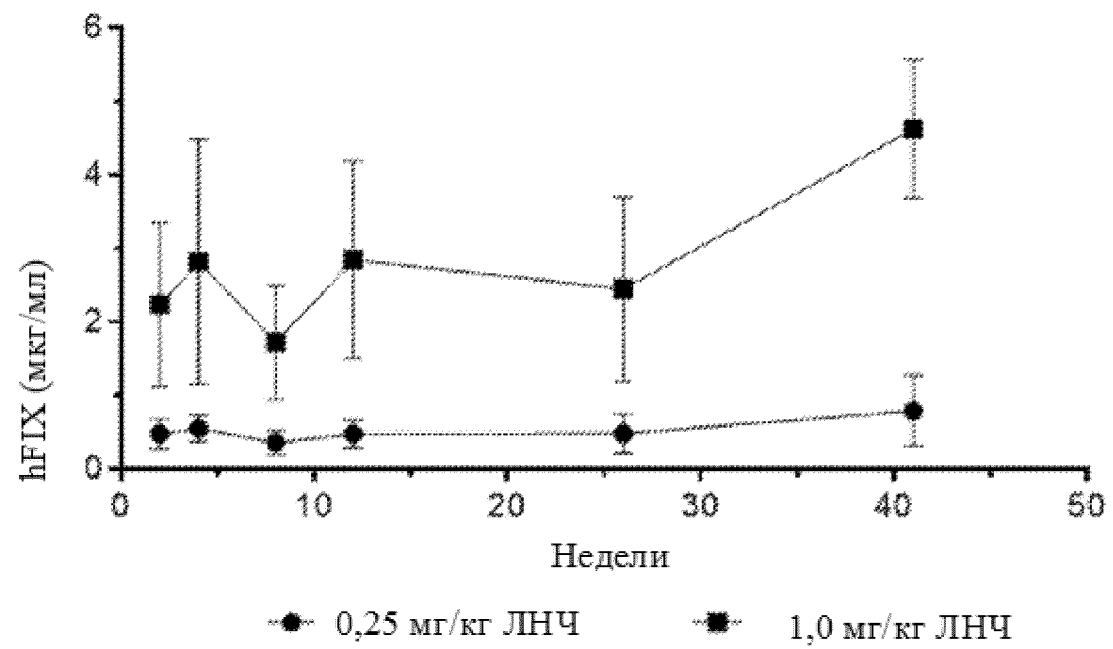
Фиг. 9



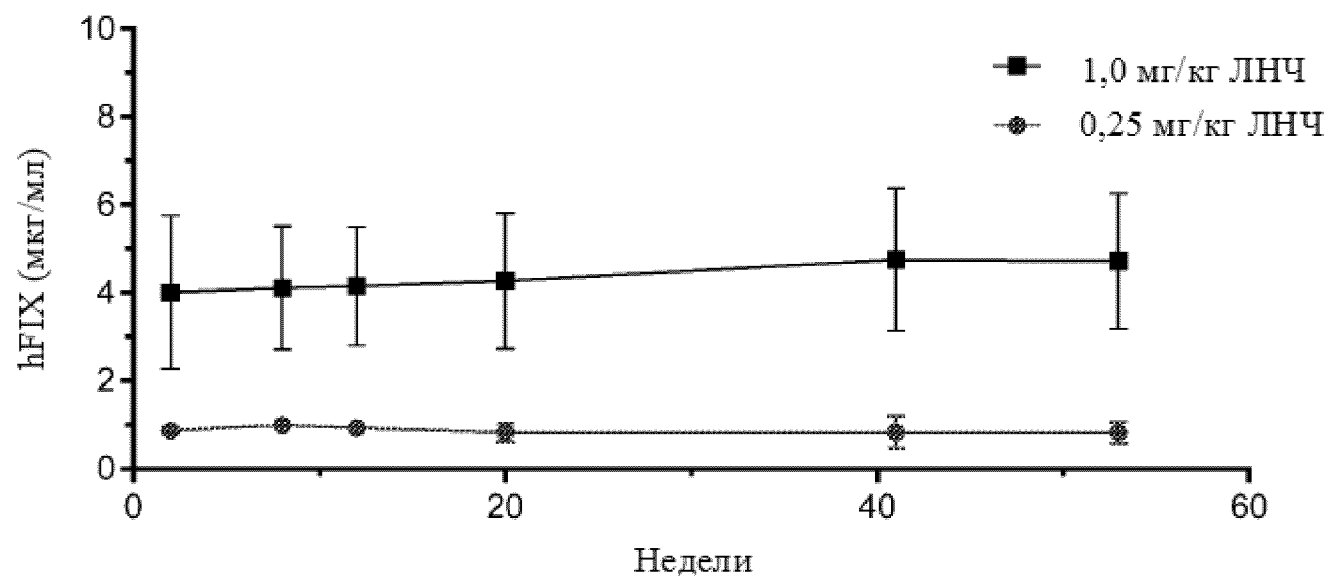
Фиг. 10



Фиг. 11А

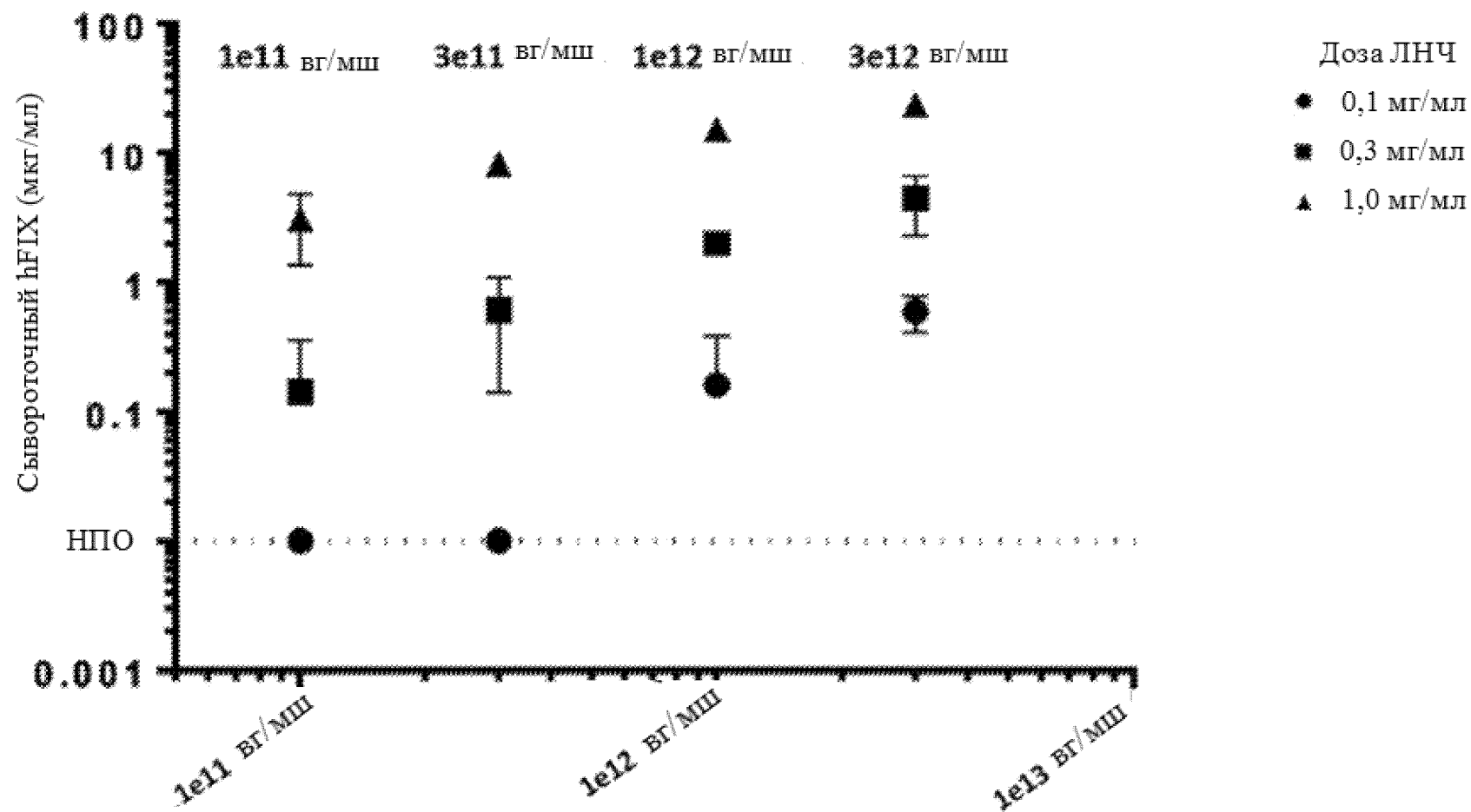




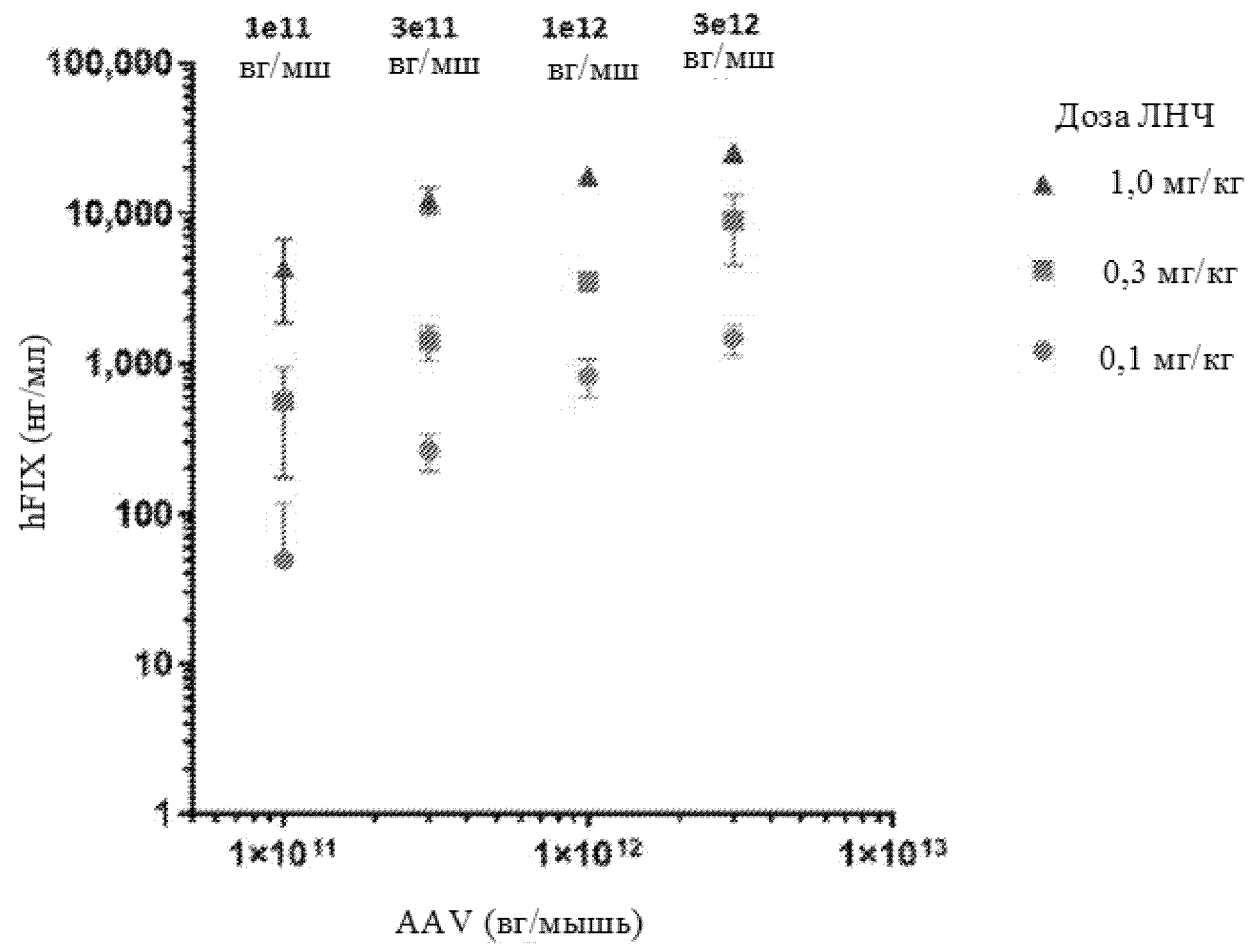


Фиг. 11В

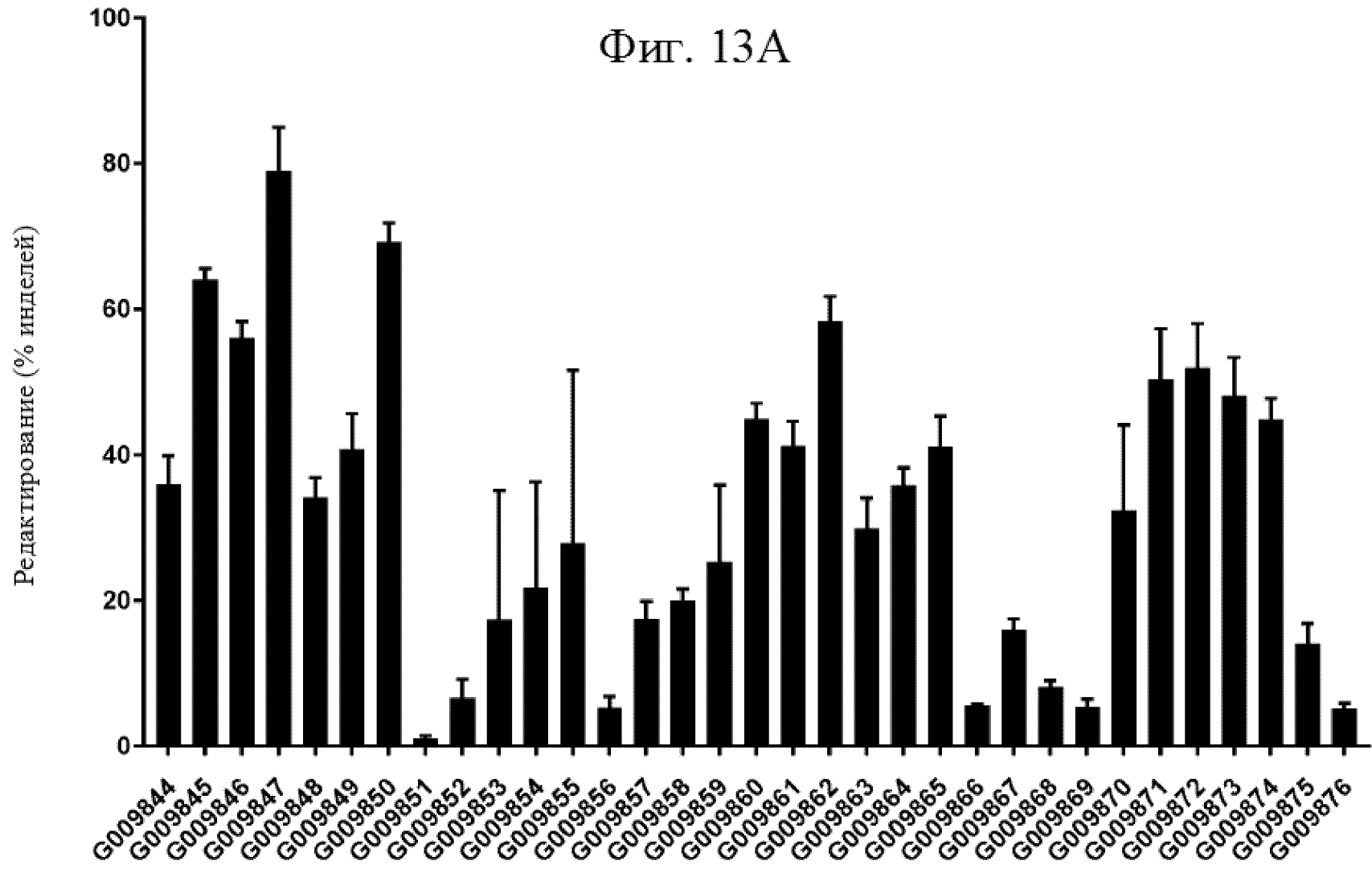
Фиг. 12А

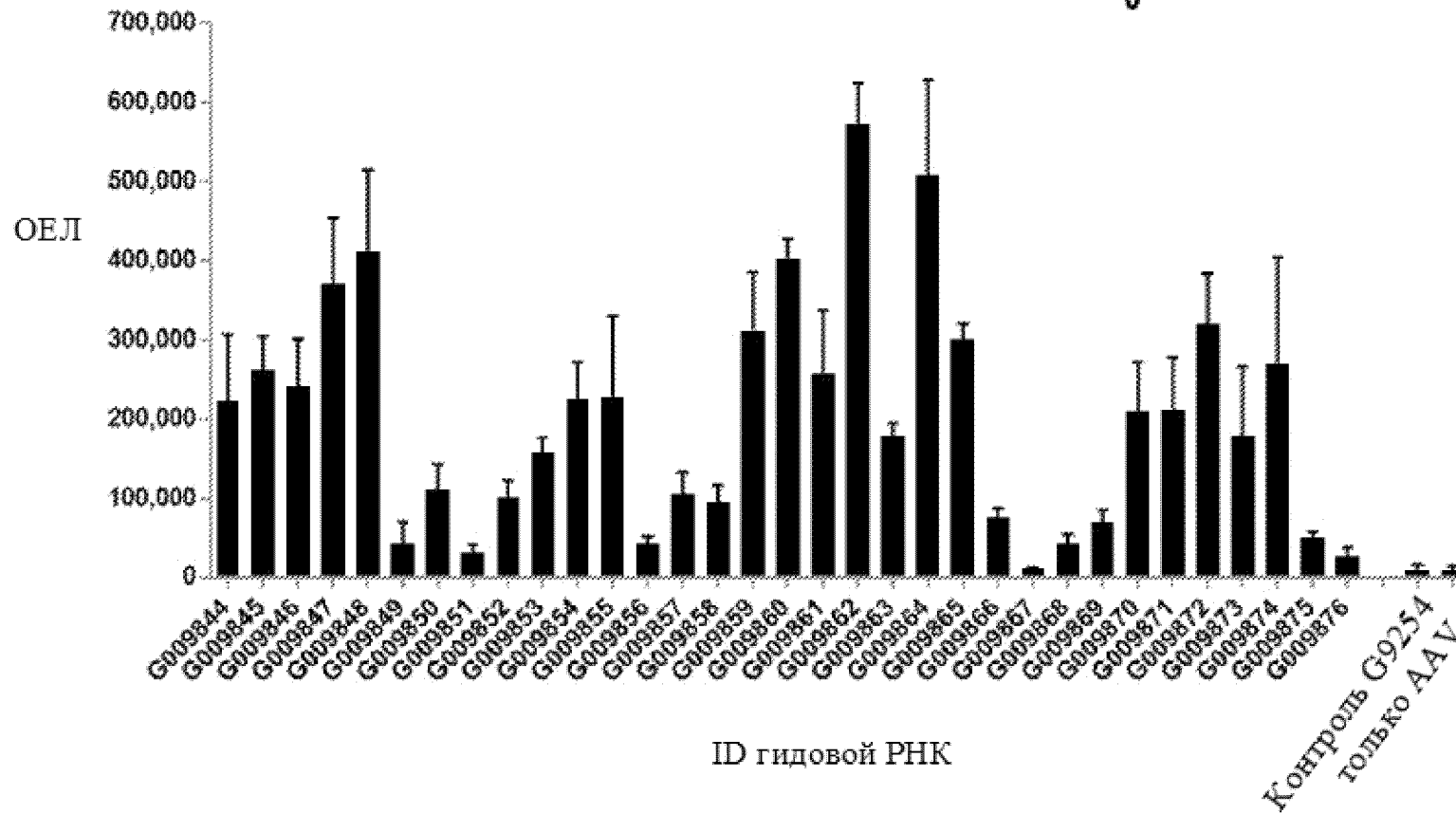
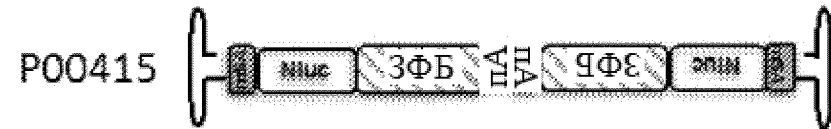


Фиг. 12В

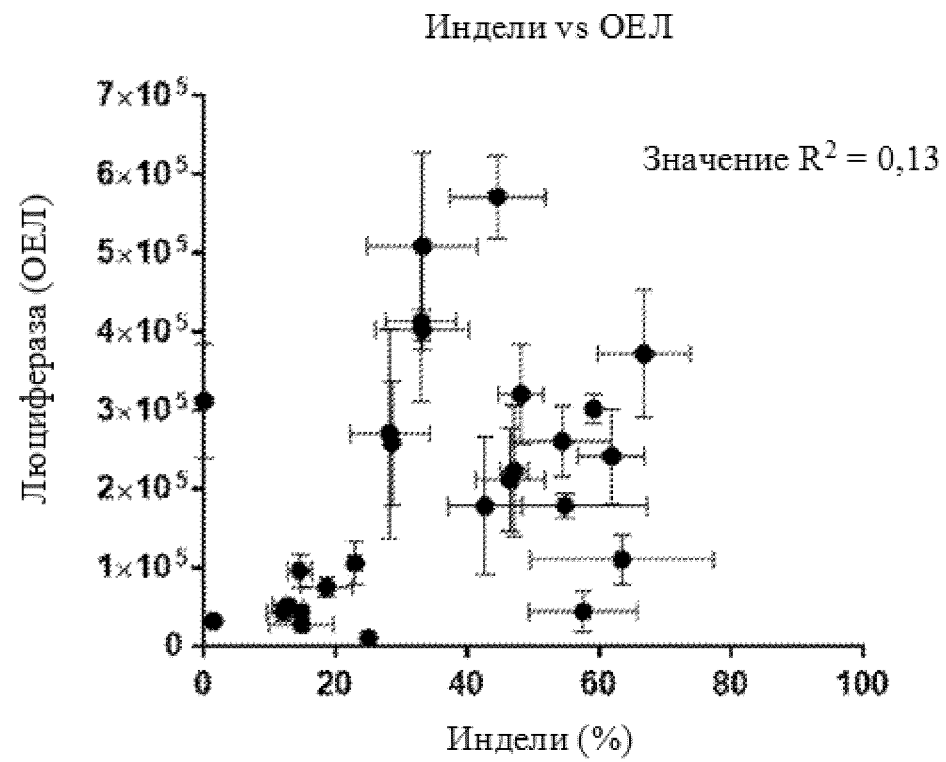


Фиг. 13А

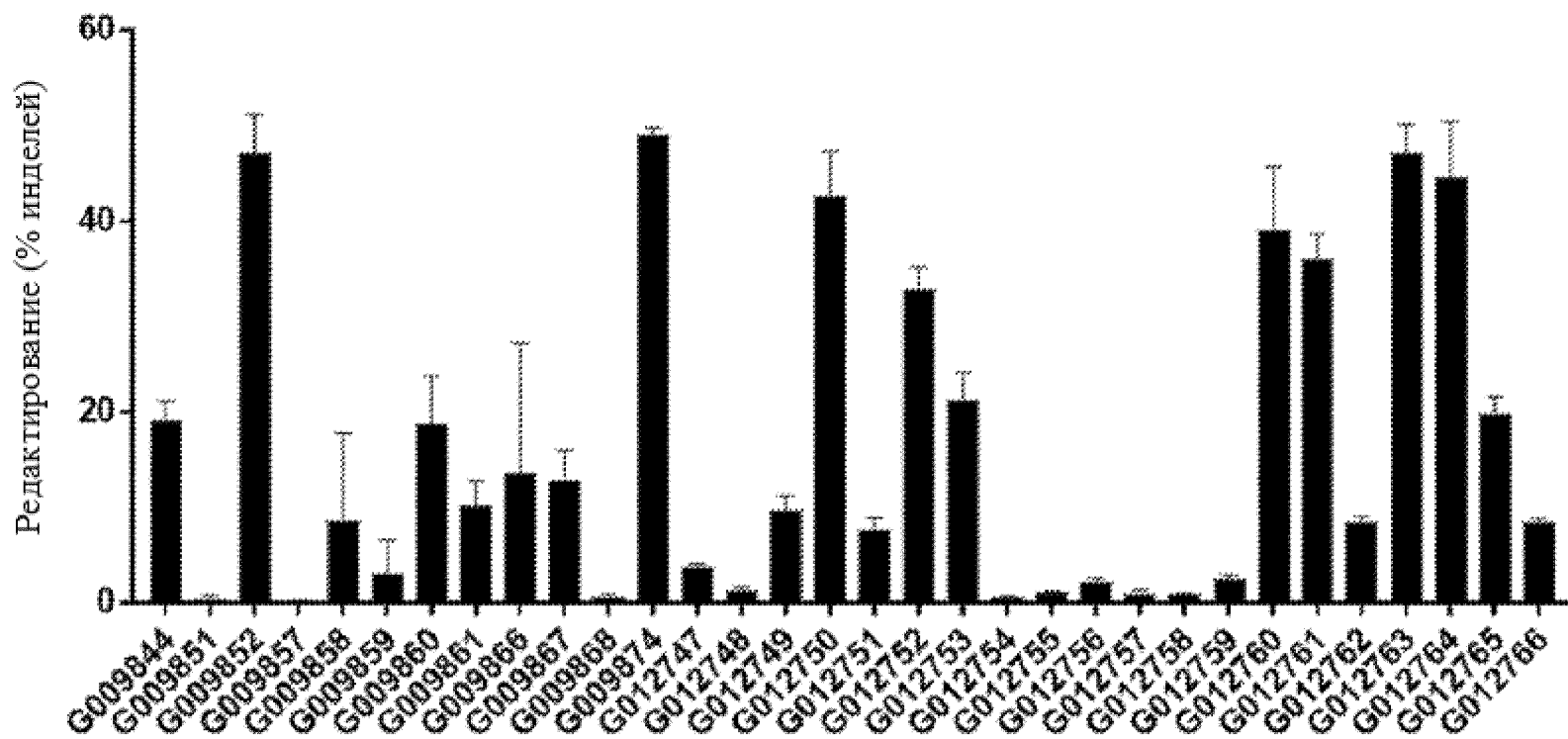




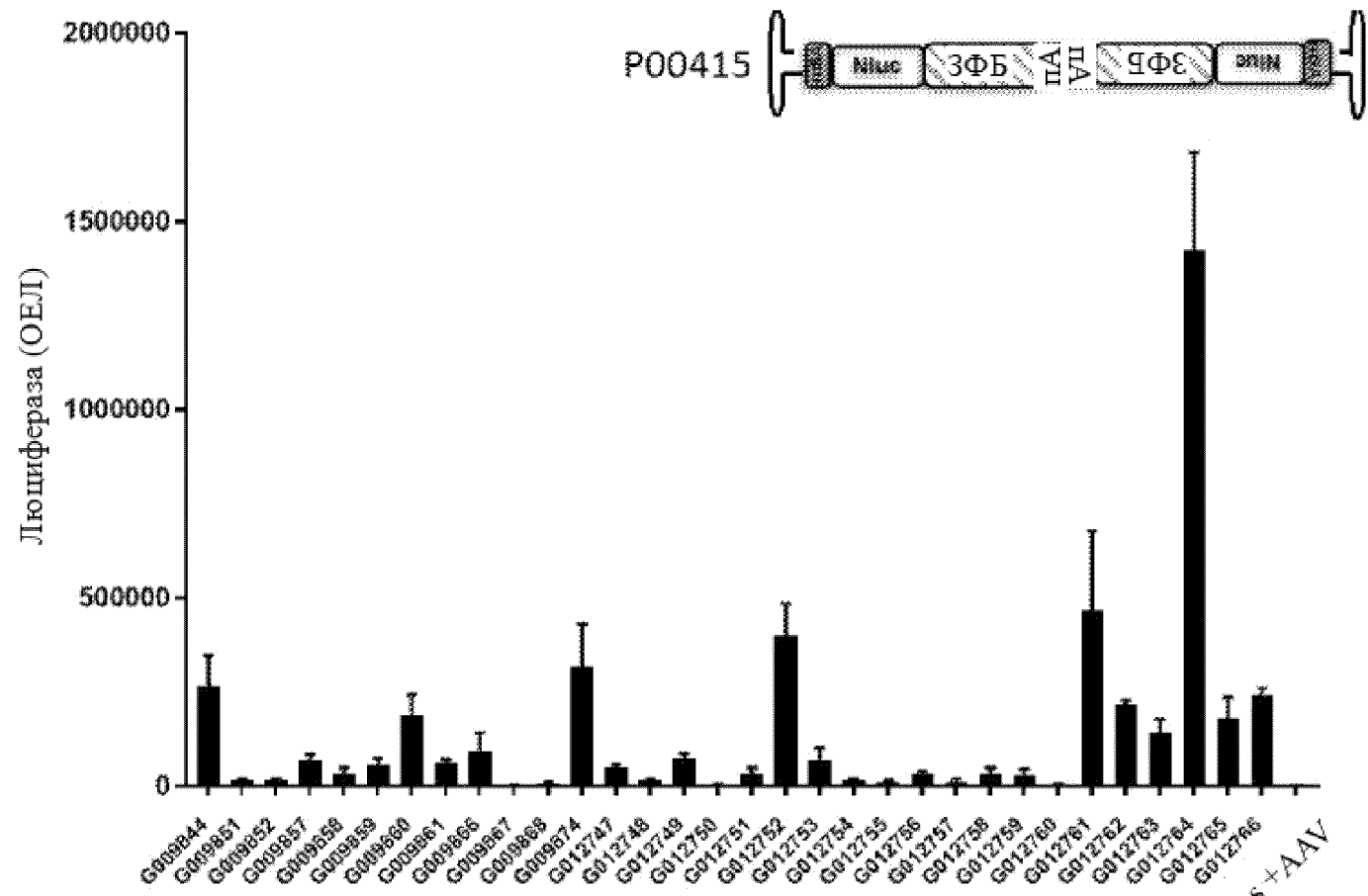
Фиг. 13В



Фиг. 13С



Фиг. 14А

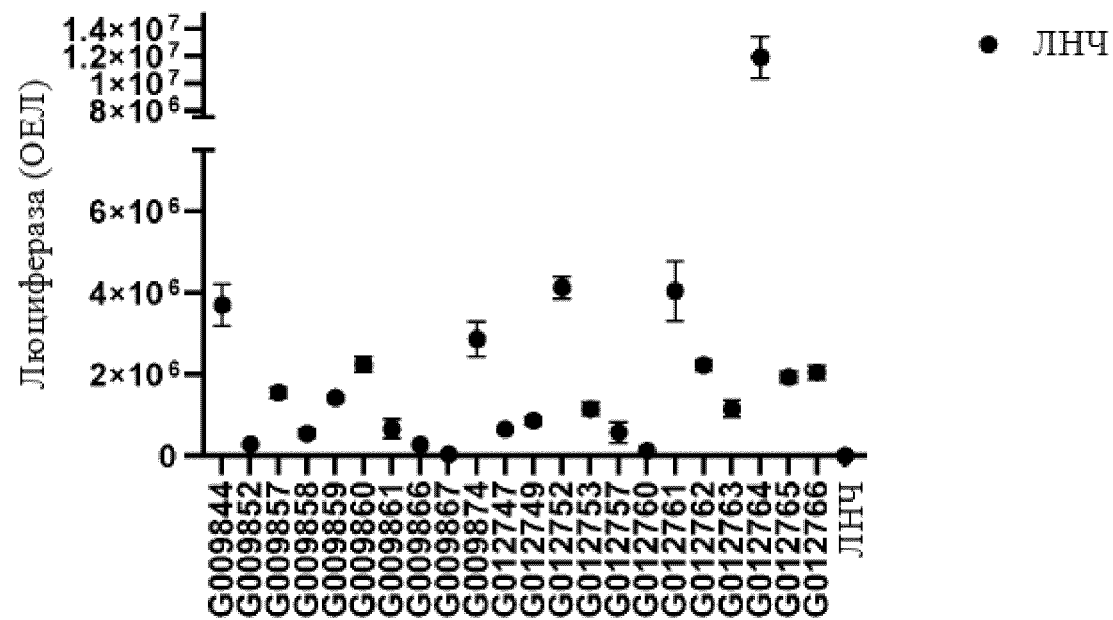


Фиг. 14В

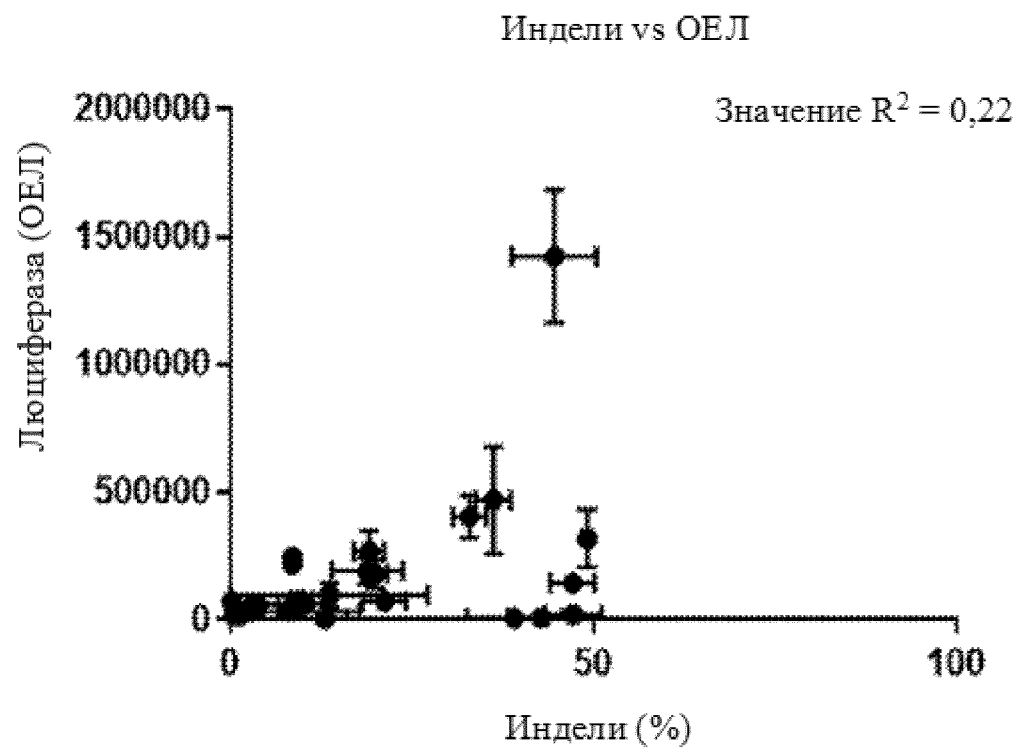
Без геновой РНК/Cas + AAV



Скрининг гРНК вставки альбумина человека

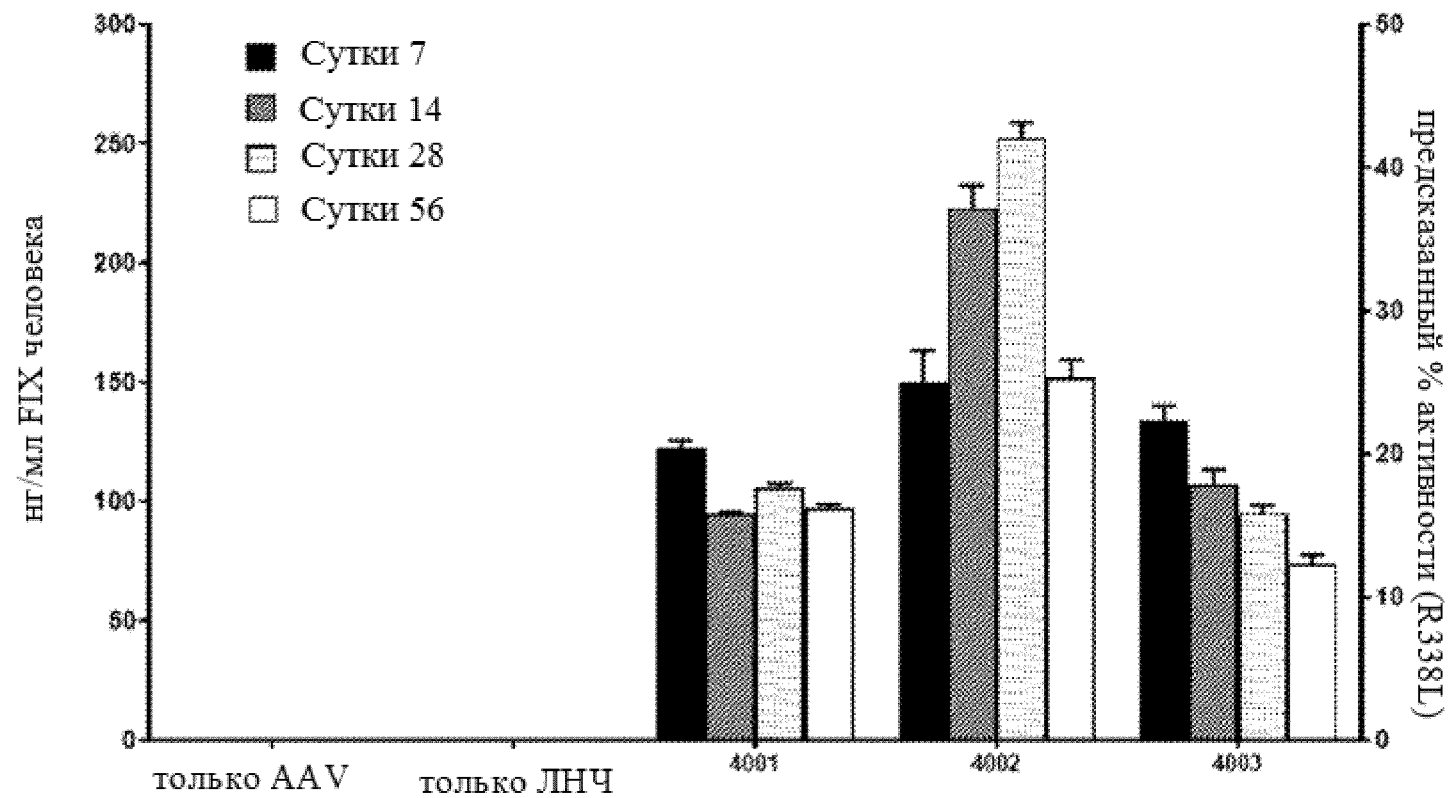


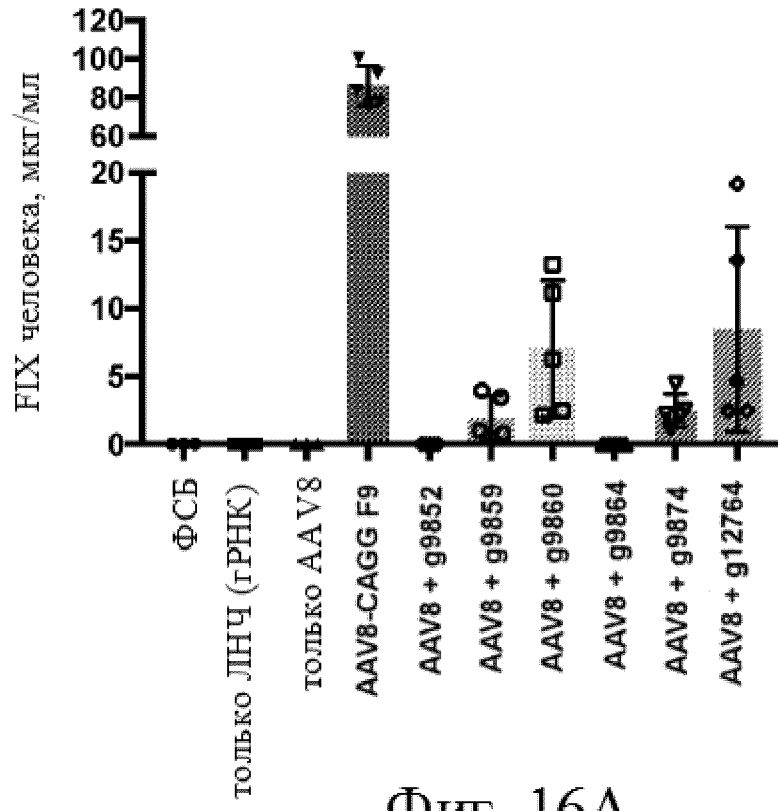
Фиг. 14С



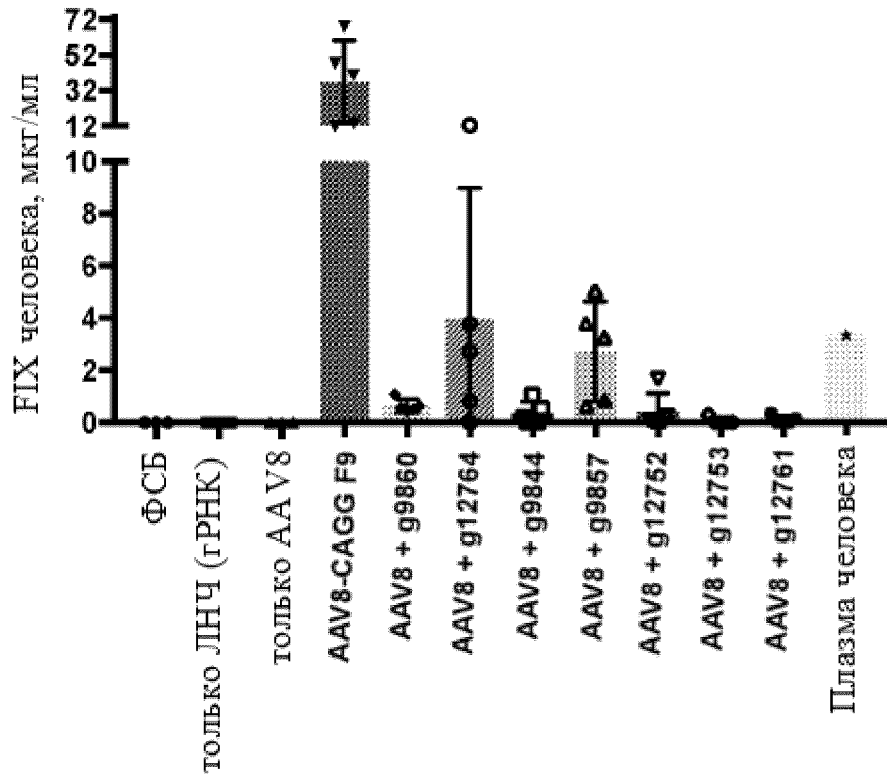
Фиг. 14D

Фиг. 15

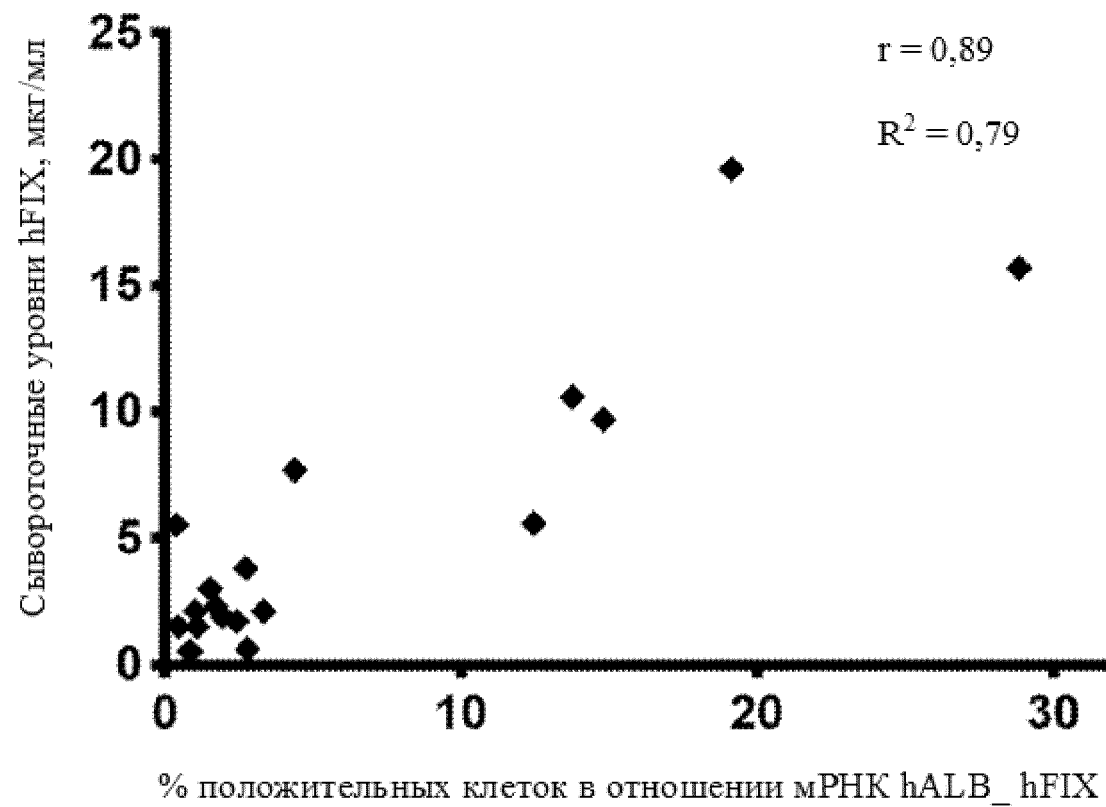




Фиг. 16А



Фиг. 16В



Фиг. 17