

(19)



Евразийское
патентное
ведомство

(21) 202291908 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.11.03

(51) Int. Cl. C12N 15/90 (2006.01)
C12N 15/10 (2006.01)

(22) Дата подачи заявки
2021.01.29

(54) ЛОКУСЫ "БЕЗОПАСНОЙ ГАВАНИ"

(31) 20154496.2

(32) 2020.01.30

(33) EP

(86) PCT/EP2021/052101

(87) WO 2021/152086 2021.08.05

(71) Заявитель:
БИТ БАЙО ЛИМИТЕД (GB)

(72) Изобретатель:

Джуд Джулиан, Васильев Анатолий,
Крейчи Адам, Бюркштюммер
Тильманн, Лоретгер Анке, Бардерчер
Лукас (AT)

(74) Представитель:

Поликарпов А.В., Соколова М.В.,
Путинцев А.И., Черкас Д.А., Игнатьев
А.В., Билык А.В., Дмитриев А.В.,
Бучака С.М., Бельтюкова М.В. (RU)

(57) Выделенная клетка-хозяин млекопитающего, содержащая гетерологичный ген, представляющий интерес (GOI), интегрированный в хромосому в целевом сайте в пределах межгенной области между парой соседних незаменимых генов.

202291908

A1

A1

202291908

ЛОКУСЫ «БЕЗОПАСНОЙ ГАВАНИ»

ОБЛАСТЬ ИЗОБРЕТЕНИЯ

Изобретение относится к рекомбинантной клетке-хозяину и к способам экспрессии гена, представляющего интерес (GOI), из клетки-хозяина. В частности, изобретение относится к способам улучшения способности клетки-хозяина стабильно экспрессировать GOI в течение длительного периода времени и генной терапии с помощью рекомбинантной клетки-хозяина.

ПРЕДШЕСТВУЮЩИЙ УРОВЕНЬ ТЕХНИКИ

Клеточная и генная терапия основана на внедрении трансгенов в клетку-мишень. Внедрение может происходить различными способами, например, с помощью транспозонов или лентивирусных векторов. Внедрение трансгена также может происходить посредством гомологичной репарации при содействии программируемых нуклеаз, таких как нуклеазы с цинковыми пальцами (ZFNs), нуклеазы на основе эффектора, подобного активатору транскрипции (TALENs), коротких палиндромных повторов, регулярно расположенных группами (CRISPRs), хоуминг-эндонуклеаз и мегануклеаз. Популярные трансгены включают гены с нарушенной функцией, а также химерные антигенные рецепторы (для CAR-T терапии), Т-клеточные рецепторы (для TCR терапии) или кДНК, функция которых нарушена у пациентов.

Стабильная экспрессия в эукариотических клеточных линиях находит применение в многочисленных проектах. Например, она поддерживает продукцию рекомбинантного белка в течение длительного периода в клеточной культуре или она способствует оптимизации продуцирования белка в трансгенных эукариотах с высоким выходом. Поддержание стабильной экспрессии трансгенов важно для обеспечения функциональности. Стабильная экспрессия подразумевает встраивание в геном.

Встроенные случайным образом гены подвержены эффекту положения и сайленсингу, что делает их экспрессию ненадежной и непредсказуемой. Определенный прогресс достигается благодаря целенаправленной вставке в сайты, являющиеся «безопасной гаванью». Для этого можно избирать в качестве мишени на так называемые локусы «безопасной гавани», которые активно транскрибируются,

например, локус *AAVS1* у человека (также известный, как *PPP1R12C*), хорошо подтвержденную “безопасную гавань” в геноме человека, или *ROSA26* у мыши и человека.

Сайты, являющиеся «безопасной гаванью», определяются по их расположению относительно смежных кодирующих генов, микроРНК и ультраконсервативных областей и, например, отвечают следующим критериям: (1) расстояние по меньшей мере 50 тпн (тысяч пар нуклеотидов) от 5' конца любого гена, (2) расстояние по меньшей мере 300 тпн от любого гена, связанного с раком, (3) расстояние по меньшей мере 300 тпн от любой микроРНК (miRNA), (4) расположение за пределами транскрипционной единицы и (5) расположение за пределами ультраконсервативных областей (UCR) человеческого генома (Bejerano G, et al.. Science. 2004;304:1321–1325; Pellenz S, et al., Hum Gene Ther. 2019;30(7):814–828).

Browning Jill et al. (Molecular Biology Reports 2019, 47(2):1491-1498) описывает CRISPR, мишенью которых является межгенная область *Hipp11* у мышей, для поддержания индуцибельной экспрессии трансгена.

Keiichiro Suzuki et al. (Nature 2016, 540(7631):144-149) описывает редактирование генома посредством CRISPR/Cas9-опосредованной независимой от гомологии направленной интеграции.

Dahodwala Hussain et al. (Current Opinion in Biotechnology 2019, 60: 128-137) описывает стратегии предсказания нестабильности линии клеток CHO.

Тем не менее, сайленсинг трансгена остается важной проблемой, которую непросто разрешить и которая затрудняет реализацию вышеупомянутых подходов. Сайленсинг представляет собой явление, которое реализуется путем установления репрессивных меток на ДНК или гистонах, например, триметилирования гистона H3 по лизину 9 или 27. Когда происходит сайленсинг, эти репрессивные хроматиновые метки распространяются по некоторой области, обычно по меньшей мере несколько тысяч пар нуклеотидов последовательности ДНК (Mlambo et al. Nucleic Acids Res. 2018; 46(9):4456-4468) и могут влиять не только на сам трансген, но и на смежные области.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

Задача заключается в устранении сайленсинга трансгена путем выбора в геноме локусов «безопасной гавани», которые не могут подвергаться сайленсингу. Другой задачей изобретения является предоставление трансгенных клеток-хозяев с

улучшенными способностями экспрессировать ген в течение длительного периода времени.

Решение данной задачи составляет сущность заявленного изобретения.

Предложена выделенная клетка-хозяин млекопитающего, содержащая гетерологичный ген, представляющий интерес (GOI), интегрированный в хромосому в целевом сайте в пределах межгенной области между парой соседних незаменимых генов.

В изобретении предложена выделенная клетка-хозяин млекопитающего, содержащая гетерологичный ген, представляющий интерес (GOI), интегрированный в хромосому в целевом сайте в пределах межгенной области между парой соседних незаменимых генов, где каждый из соседних незаменимых генов указанной пары имеет сайт начала транскрипции (TSS), и расстояние между TSS меньше, чем любое из 20000, 15000, 10000, 5000, 2500, 1000, 900, 800, 700, 600, 500 или 400 нуклеотидов (нт).

В частности, расстояние между соответствующими TSS представляет собой по меньшей мере любое из 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330 или 340 нт.

Интеграция в хромосому достигается, в частности, благодаря одной или более генетическим модификациям хромосомы клетки-хозяина в пределах межгенной области и/или целевом сайте.

В частности, указанная одна или более генетических модификаций включает вставку или нокин GOI или экспрессионной конструкции, содержащей GOI.

В частности, клетка-хозяин представляет собой диплоидную клетку (данный термин в данном документе охватывает околодиплоидную клетку), такую как содержащую две копии каждой хромосомы.

В частности, выделенная клетка-хозяин предложена в культуре клеток-хозяев или предложена в виде фармацевтического препарата или донорской клетки для клеточной терапии.

В частности, соседние незаменимые гены представляют собой смежные кодирующие гены, разделенные только межгенной областью и/или расположенные таким образом, что межгенная область между двумя соседними незаменимыми генами представляет собой смежную не кодирующую нуклеотидную последовательность.

В частности, описанная межгенная область содержит или представляет собой геномную «безопасную гавань». В частности, «безопасная гавань» содержит целевой

сайт или локус для встраивания GOI. Целевой сайт может быть сайтом в пределах межгенной области, выбранным случайным образом, или заранее заданным сайтом.

Таким образом, описанную в данном документе межгенную область понимают как полинуклеотидную последовательность между соседними незаменимыми генами до встраивания трансгена (например, GOI или экспрессирующей конструкции, содержащей GOI)) в целевой сайт. Другими словами, межгенная область содержит некую смежную последовательность, расположенную между соседними незаменимыми генами, не содержащую трансген. Когда клетку модифицируют, чтобы она содержала трансген в пределах межгенной области, понимают, что межгенная область содержит или состоит из полинуклеотидной последовательности, которая определяет область между соседними незаменимыми генами без трансгена, т.е. содержит или состоит из двух частей межгенной области, фланкирующих 5'-конец и 3'-конец трансгена, в противном случае эти части слиты вместе, с получением непрерывной последовательности, содержащей или представляющей собой нуклеотидную последовательность фланкирующих частей.

В частности, длина межгенной области меньше, чем любое из 20000, 15000, 10000, 5000, 2500, 1000, 900, 800, 700, 600, 500 или 400 нт.

В частности, межгенная область не содержит ген или часть гена.

В частности, межгенная область является некодирующей.

В частности, по меньшей мере любое из 60%, 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% межгенной области состоит из полинуклеотида человеческого происхождения, такого как человеческий полинуклеотид. В частности, межгенная область содержит непрерывную часть любой из SEQ ID NO:1-21, которая представляет собой по меньшей мере любое из 60%, 70%, 80%, 85%, 90%, 91%, 90%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% от полной длины соответствующей последовательности SEQ ID NO:1-21.

В частности, межгенная область содержит или состоит из нуклеотидной последовательности, которая обладает любой идентичностью из по меньшей мере 90%, 91%, 92%, 93%, 94%, 95%, 90%, 97%, 98%, 99%, 96%, 97%, 98%, 99% или 100% с любой из SEQ ID NO:1-21.

Согласно определенному воплощению, межгенная область содержит или состоит из полинуклеотида, имеющего человеческое происхождение, такого как человеческий полинуклеотид, который на 100% идентичен любой из SEQ ID NO:1-21.

В частности, межгенная область содержит или состоит из полинуклеотида, который представляет собой естественный вариант, происходящий от млекопитающего или человека, с любой из SEQ ID NO:1-21, например, такой как полинуклеотид, определенный в любой из SEQ ID NO:1-21, где полинуклеотидная последовательность модифицирована одним или более однонуклеотидными полиморфизмами, которые возникают естественным образом в полинуклеотидной последовательности человеческого генома.

В частности, каждый из незаменимых генов имеет человеческое происхождение и/или имеет последовательность, обладающую любой идентичностью из по меньшей мере 85%, 90%, 95%, 96%, 97%, 98%, 99%, 93%, 94%, 95%, 96%, 97%, 98%, 99% с соответствующей последовательностью гена человека или на 100% соответствующей последовательности гена человека.

В частности, незаменимые гены представляют собой человеческие гены или их природные варианты, происходящие от различных людей, или их аналоги, происходящие от различных видов млекопитающих, таких как мышь, хомяк, крыса, собака или свинья. Последовательность незаменимого гена обладает большой консервативностью, идентичность последовательности соответствующих генов у животных, не являющихся людьми, по сравнению с людьми составляет по меньшей мере 80% или 85% или 90%. Например, нуклеотидная последовательность MED20 человека по меньшей мере на 93% идентична соответствующему гену у хомяка (CHO), на 90% идентична соответствующему гену у мыши, на 99% идентична соответствующему гену у собаки и на 98% идентична соответствующему гену у свиньи.

В частности, известно, что каждый из незаменимых генов указанной пары незаменимых генов экспрессирует белки с высоким уровнем экспрессии. Незаменимые гены выполняют ключевые функции в клетке и зачастую высоко консервативны среди биологических видов. Они определяются как гены, чья инактивация или делеция приводит к гибели клетки.

В частности, пара незаменимых генов представляет собой пару, состоящую из первого и второго гена, где

1) первый ген представляет собой NFS1, а второй ген представляет собой ROMO1; или

2) первый ген представляет собой MED22, а второй ген представляет собой RPLA7; или

3) первый ген представляет собой DDX51, а второй ген представляет собой NOC4L; или

4) первый ген представляет собой CENPK, а второй ген представляет собой PPWD1; или

5) первый ген представляет собой ORC1, а второй ген представляет собой PRPF38A; или

6) первый ген представляет собой POLR3K, а второй ген представляет собой SNRNP25; или

7) первый ген представляет собой COPE, а второй ген представляет собой DDX49; или

8) первый ген представляет собой FTSJ3, а второй ген представляет собой PSMC5; или

9) первый ген представляет собой MED20, а второй ген представляет собой BYSL; или

10) первый ген представляет собой AURKA, а второй ген представляет собой CSTF1; или

11) первый ген представляет собой NUP88, а второй ген представляет собой RPAIN; или

12) первый ген представляет собой ATP6V1D, а второй ген представляет собой EIF2S1; или

13) первый ген представляет собой POLR2I, а второй ген представляет собой TBCB; или

14) первый ген представляет собой UFD1, а второй ген представляет собой CDC45; или

15) первый ген представляет собой CCDC115, а второй ген представляет собой IMP4; или

16) первый ген представляет собой NAA50, а второй ген представляет собой ATP6V1A; или

17) первый ген представляет собой SART3, а второй ген представляет собой ISCU; или

18) первый ген представляет собой C1orf109, а второй ген представляет собой CDCA8; или

19) первый ген представляет собой POLR3A, а второй ген представляет собой RPS24; или

20) первый ген представляет собой RPS16, а второй ген представляет собой SUPT5H; или

21) первый ген представляет собой RPS29, а второй ген представляет собой LRR1.

Каждый из упомянутых в данном документе незаменимых генов характеризуется идентификационным номером в базе Entrez Gene. Entrez Gene представляет собой базу данных Национального центра биотехнологической информации (NCBI) для информации о генах, генерирующую уникальные целые числа (GeneID) в качестве постоянных идентификаторов генов (Maglott et al. Nucleic Acids Res. 2011; 39 (Database issue):D52–D57). Каждый ген имеет сайт начала транскрипции (TSS), который обозначает старт транскрипции, инициируемой промотором соответствующего гена. TSS в нуклеотидной последовательности можно идентифицировать при помощи доступных баз данных, таких как RefSeq или FANTOM.

В частности, пара незаменимых генов и соответствующая межгенная область характеризуется, как показано в Таблице 1.

Таблица 1: Примеры пар незаменимых генов и межгенных областей, содержащих целевой сайт.

X/c	ген1	ID гена1	TSS гена1	ген2	ID гена2	TSS гена2	Межгенная обл, нт*	Начало межгенной области	Конец межгенной области	SEQ ID NO
x/c20	NFS1	9054	35699348	ROMO1	140823	35699401	52	35699349	35699400	1
x/c9	MED22	6837	133348090	RPL7A	6130	133348210	119	133348091	133348209	2
x/c12	DDX51	317781	132144324	NOC4L	79050	132144450	125	132144325	132144449	3
x/c5	CENPK	64105	65563146	PPWD1	23398	65563288	141	65563147	65563287	4
x/c1	ORC1	4998	52404410	PRPF38A	84950	52404606	195	52404411	52404605	5
x/c16	POLR3K	51728	53618	SNRNP25	79622	53858	239	53619	53857	6
x/c19	COPE	11316	18919374	DDX49	54555	18919716	341	18919375	18919715	7
x/c17	FTSJ3	117246	63827092	PSMC5	5705	63827436	343	63827093	63827435	8
x/c6	MED20	9477	41921128	BYSL	705	41921500	371	41921129	41921499	9
x/c20	AURKA	6790	56392204	CSTF1	1477	56392646	441	56392205	56392645	10
x/c17	NUP88	4927	5419662	RPAIN	84268	5420190	527	5419663	5420189	11
x/c14	ATP6V1D	51382	67359791	EIF2S1	1965	67360338	546	67359792	67360337	12

x/c19	POLR2I	5438	36114890	TBCB	1155	36115502	611	36114891	36115501	13
x/c22	UFD1	7353	19479184	CDC45	8318	19479921	736	19479185	19479920	14
x/c2	CCDC115	84317	130342152	IMP4	92856	130342890	737	130342153	130342889	15
x/c3	NAA50	80218	113746236	ATP6V1A	523	113747050	813	113746237	113747049	16
x/c12	SART3	9733	108561169	ISCU	23479	108562603	1433	108561170	108562602	17
x/c1	C1orf109	54955	37690515	CDCA8	55143	37692536	2020	37690516	37692535	18
x/c10	POLR3A	11128	78029506	RPS24	6229	78033862	4357	78029507	78033863	19
x/c19	RPS16	6217	39435944	SUPT5H	6829	39445593	9648	39435945	39445592	20
x/c14	RPS29	6235	49586376	LRR1	122769	49598938	12561	49586377	49598937	21

* Между TSS гена 1 и TSS гена 2

Согласно конкретному аспекту, используют регуляторные последовательности, выбранные из контролирующих экспрессию последовательностей, таких как промотор, последовательность оператора, последовательность энхансера транскрипции или последовательность сайленсера транскрипции; или нуклеотидные последовательности, кодирующие факторы экспрессии, такие как энхансеры или ингибиторы. В частности, одна или более регуляторных последовательностей, например, промотор, расположены на расстоянии от 2000 до 1000 пн (пар нуклеотидов) в направлении против хода транскрипции или на расстоянии 100 пн в направлении по ходу транскрипции от TSS указанного GOI, в некоторых случаях небольшой межгенной области, например, на любом расстоянии от 1000, 5000, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 150, 100 до 50 пн против хода транскрипции и любом расстоянии от 50 до 100 пн по ходу транскрипции от TSS.

В частности, транскрипция пары незаменимых генов осуществляется в противоположной ориентации.

В частности, каждый ген из пары незаменимых генов находится под транскрипционным контролем соответствующего промотора незаменимого гена (EGP), предпочтительно, EGP естественного происхождения соответствующего гена, как конститутивного, так и индуцибельного.

В частности, гены указанной пары незаменимых генов ориентированы таким образом, что их EGP обращены в противоположных направлениях.

В частности, EGP примыкают друг к другу.

Более конкретно, EGP примыкают друг к другу и обращены в противоположных направлениях.

Соответствующие промоторные области, функционально связанные с первым и вторым генами пары незаменимых генов, предпочтительно, примыкают друг к другу

или находятся в тесной близости, так что расстояние между соответствующими TSS меньше любого из 20000, 15000, 10000, 5000, 2500, 1000, 900, 800, 700, 600, 500 или 400 нт.

В частности, межгенная область, содержащая целевой сайт, содержит или состоит из области между соответствующими TSS. Предпочтительно, целевой сайт содержится в такой области между TSS, где два незаменимых гена кодируются противоположными цепями, и TSS находятся в тесной близости.

Согласно конкретному воплощению, целевой сайт по меньшей мере частично или полностью содержится в составе промоторной области. Например, трансген встраивают в область, которая считается промотором одного из двух незаменимых генов, в частности, когда доказано, что встраивание трансгена не влияет на функцию незаменимого гена.

В частности, межгенная область представляет собой область, расположенную между сайтами начала транскрипции EGP и соединяющую их.

В частности, незаменимые гены представляют собой так называемые «ключевые незаменимые гены», то есть они являются незаменимыми для многих или для всех типов клеток или состояний клеток, известных на настоящий момент. Следовательно, необходимо поддерживать их функцию во всех типах клеток и состояниях клеток определенных видов. Например, человеческие ключевые незаменимые гены считаются незаменимыми во всех человеческих клеточных линиях. Ключевые незаменимые гены можно определить посредством полногеномного скрининга нокаутов с помощью CRISPR/SpCas9, как описано Hart et al. (G3: Genes, Genomes, Genetics August 1, 2017 vol. 7 no. 8 2719-2727) или Wang et al. (Science. 2014; 343(6166):80–84). Примеры человеческих основных незаменимых генов приведены в Таблице 1 ниже (как ген 1 и/или ген 2).

Согласно конкретному аспекту, GOI находится под транскрипционным контролем гетерологичного промотора. GOI может быть интегрирован в целевой сайт в прямой или обратной ориентации.

В частности, GOI содержится в гетерологичной экспрессионной кассете, интегрированной в целевой сайт. Экспрессионная кассета может содержать по меньшей мере один контролирующий элемент, промотор или регуляторный элемент, который функционально связан с GOI. В некоторых воплощениях контролирующий элемент,

промотор или регуляторный элемент, функционально связанные с GOI, являются индуцибельными.

В частности, GOI представляет собой гетерологичный ген или трансген.

В частности, GOI выбран из любого из: нуклеиновой кислоты, ингибитора, кДНК или гена, кодирующих пептид или полипептид, и кДНК или гена, кодирующих антитело или фрагмент антитела, слитый белок, антиген, антагонист, агонист, молекулу RNAi или miRNA. В некоторых воплощениях кДНК или ген кодирует один или более транскрипционных факторов.

Целевой сайт, в частности, понимают как сайт встраивания, для встраивания GOI или экспрессионной кассеты, содержащей GOI.

Целевой сайт может содержать по меньшей мере один или два сайта (то есть по меньшей мере одно или два местоположения) для встраивания одного или более GOI (или экспрессирующих конструкций, содержащих GOI) в целевом сайте в пределах межгенной области. Межгенная область может иметь природное происхождение или происходить из последовательности природного происхождения и быть дополнительно модифицированной, например, для встраивания распознаваемого целевого сайта или любого другого гетерологичного средства для подготовки нокина трансгена.

Помимо встраивания GOI (или экспрессионной кассеты, содержащей GOI) в целевой сайт межгенная область может быть дополнительно модифицирована посредством одной или более мутаций, таких как делеция, вставка или замена одного или более нуклеотидов, которых может быть всего несколько, например, до 20 нт или до 15 нт или до 10 нт, но может также изменять нуклеотидную последовательность, так что модифицированная межгенная область (за исключением последовательности GOI или гетерологичной экспрессионной кассеты) обладает любой идентичностью последовательностей из по меньшей мере 50%, 60%, 70%, 80%, 85%, 90%, 95%, 98% или 99% или на 100% идентична межгенной области природного происхождения.

Примеры межгенных областей, характеризующихся любой из SEQ ID NO:1-21, могут нести или не нести один или более однонуклеотидных полиморфизмов, которые, в частности, имеют природное происхождение. В частности, межгенная область представляет собой межгенную область природного происхождения, которая может отличаться от любой из SEQ ID NO:1-21 одним или более однонуклеотидными полиморфизмами, или которая идентична любой из SEQ ID NO:1-21 и не имеет какого-либо однонуклеотидного полиморфизма.

Примерами клеток-хозяев являются стволовые клетки, включая, например, гемопоэтическую стволовую клетку, эмбриональную стволовую клетку, плюрипотентную стволовую клетку, индуцированную плюрипотентную стволовую клетку или эндотелиальную клетку.

Согласно определенному аспекту, клетка-хозяин представляет собой первичную клетку или иммунную клетку, такую как из типа клеток, выбранного из группы, состоящей из клетки-натурального киллера (NK), клетки микроглии, макрофага или Т-клетки, такой как цитотоксический Т-лимфоцит (CTL), регуляторная Т-клетка или хелперная Т-клетка. Такая клетка-хозяин, в частности, применяется в фармацевтическом препарате и/или в клеточной терапии и/или в генной терапии. Например, модифицированные Т-клетки могут быть модифицированы для экспрессии Т-клеточного рецептора (TCR), в частности, гетерологичного TCR, или химерного антигенного рецептора (CAR).

В частности, клетка-хозяин представляет собой иммунную эффекторную клетку, содержащую химерный антигенный рецептор (CAR), где GOI кодирует CAR, экспрессируемый указанной иммунной эффекторной клеткой.

В частности, клетка-хозяин представляет собой иммунную эффекторную клетку, содержащую химерный антигенный рецептор, такую как CAR-Т клетка или CAR макрофаг, а GOI содержит нуклеотидную последовательность, кодирующую CAR или часть CAR, такую как любой один или более доменов CAR, которые могут экспрессироваться указанной клеткой-хозяином на своей поверхности.

Согласно частному аспекту, клетка-хозяин имеет человеческое происхождение, но может происходить из животного, не являющегося человеком, или млекопитающего, такого как мышь, крыса, хомяк, собака, свинья или крупный рогатый скот. В частности, клетка-хозяин не является мышинной клеткой-хозяином.

Согласно другому частному аспекту, клетка-хозяин является продуцирующей клеткой-хозяином, поскольку применяется для продуцирования белка, представляющего интерес, в культуре клетки-хозяина *in vitro*.

В частности, клетка млекопитающего представляет собой клетку или клеточную линию или штамм клеток человека, грызуна или быка. Частными примерами клеток млекопитающих, подходящих в качестве клеток-хозяев, описанных в данном документе, являются линии клеток миеломы мыши (NSO), линии клеток яичника китайского хомячка (CHO), HT1080, H9, HepG2, MCF7, MDBK Jurkat, MDCK, NIH3T3,

PC12, ВНК (клетки почки детеныша хомяка), VERO, SP2/0, YB2/0, Y0, C127, L клетки, COS, например, COS1 и COS7, QC1-3, НЕК-293, VERO, PER.C6, HeLA, EB1, EB2, EB3, линии онколитических или гибридных клеток.

В изобретении дополнительно предложен выделенный клон клетки-хозяина, описанной в данном документе, или *in vitro* культура клеток такого клона.

Согласно частному аспекту, в изобретении предложена фармацевтическая композиция, содержащая клетку-хозяина или популяцию клеток-хозяев, описанных в данном документе, и фармацевтически приемлемый носитель.

В изобретении дополнительно предложен препарат, содержащий популяцию клеток-хозяев, в частности, популяцию или репертуар различных (например, поликлональных) клеток, где любое относительное количество клеток из по меньшей мере 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% являются раскрытыми в данном документе клетками-хозяевами с хромосомной идентичностью или с идентичным генотипом или идентичным хромосомотипом (т.е. идентичностью на основе хромосомных признаков).

Частные примеры относятся к фармацевтическому препарату, содержащему популяцию CAR-T клеток, где по меньшей мере любое относительное количество клеток из по меньшей мере 1%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% являются описанными в данном документе клетками-хозяевами с хромосомной идентичностью или с идентичным генотипом или идентичным хромосомотипом, указанная популяция стабильно экспрессирует GOI, чтобы обеспечить экспрессию в течение длительного периода времени.

В изобретении также предложена *in vitro* культура клеток-хозяев раскрытого здесь клона или препарата, которую поддерживают в условиях для экспрессии указанного GOI, в частности, для стабильной экспрессии указанного GOI, тем самым либо получая соответствующий полипептид или белок, представляющий интерес, например, связанный с поверхностью клетки-хозяина, или секретируемый в среду культивирования клетки-хозяина.

В изобретении дополнительно предложено животное, не являющееся человеком, содержащее клетку-хозяина, описанную в данном документе. В частности, животное, не являющееся человеком, представляет собой генетически модифицированное

животное (также обозначаемое трансгенным животным), содержащим донорскую последовательность, содержащую GOI, донорская последовательность встроена в заданный сайт встраивания на хромосоме животного, где заданный сайт встраивания представляет собой целевой сайт, дополнительно описанный в данном документе. В частности, описанные здесь способы не включают модификацию генетической идентичности зародышевой линии человека.

Приведенный в качестве примера способ получения включает

а) создание клетки с донорской последовательностью, встроенной в заданный сайт встраивания; и

б) внедрение клетки, созданной на стадии а), в животное-носителя для получения генетически модифицированного животного.

В частности, клетка является зиготой или плюрипотентной стволовой клеткой.

В изобретении также предложен шаблон клетки-хозяина, происходящей из млекопитающего, содержащий гетерологичное средство для сайт-направленной интеграции в хромосому экзогенного гена, представляющего интерес (GOI), в частности, гетерологичного GOI, в целевой сайт в пределах межгенной области между парой соседних незаменимых генов. Такая клетка-хозяин дикого типа предпочтительно представляет собой диплоидную клетку-хозяина.

Приведенные в качестве примера средства могут представлять собой одну или более нуклеиновокислотных модификаций для встраивания рестрикционного сайта клонирования в межгенную область или для встраивания одного или более целевых сайтов для одной или более нуклеаз. Соответствующие нуклеазы могут быть выбраны из нуклеазы с цинковыми пальцами (ZFN), нуклеазы с эффекторным доменом TAL (TALEN) или системы CRISPR/Cas. Специфический целевой сайт может быть для гидовой РНК (gRNA), такой как для специфической к последовательности нуклеазы, выбранной из любой из: нуклеазы TAL, нуклеазы с цинковыми пальцами (ZFN), мегануклеазы, мегаTAL, или эндонуклеазы, управляемой при помощи гидовой РНК (например, Cas9, Cpf1, нуклеазы Cas9).

Специфические средства могут включать одну или более нуклеиновокислотных модификаций в межгенной последовательности естественного происхождения в целевом сайте или примыкающих к целевому сайту.

Согласно частному аспекту, средства выбирают из группы, состоящей из гетерологичных целевых сайтов, распознаваемых рекомбиназой, для опосредования

нокина трансгена в соответствующий целевой сайт. Например, распознаваемый целевой сайт может быть встроен в пределах межгенной области, как описано в данном документе, для модификации целевого сайта, и для интеграции GOI в хромосому используют соответствующую рекомбиназу, распознающую такие целевые сайты. Можно использовать любую обычно используемую технологию сайт-направленной рекомбинации, например, требующую подготовительных средств для получения шаблона клетки-хозяина, как описано в данном документе. Приведенными в качестве примера технологиями сайт-направленной рекомбинации являются рекомбинация Flp-FRT, рекомбинация Cre-Lox или сайт-специфическая рекомбинация у фага лямбда.

В изобретении также предложен способ получения рекомбинантных клеток-хозяев, экспрессирующих гетерологичный GOI, где GOI встроен в целевой сайт, содержащийся в межгенной области между парой соседних незаменимых генов.

В изобретении дополнительно предложен способ получения генного продукта, кодируемого геном, представляющим интерес (GOI), путем трансформации клетки-хозяина для интеграции GOI в хромосому в пределах межгенной области между парой соседних незаменимых генов и культивирования трансформированной клетки-хозяина в условиях для экспрессии указанного GOI, предпочтительно, где каждый из соседних незаменимых генов указанной пары имеет сайт начала транскрипции (TSS), и расстояние между TSS представляет собой по меньшей мере любое из 80, 90, 100 или 110 нт и является меньше любого из 20000, 15000, 10000, 5000, 2500 или 1000 нт.

В изобретении дополнительно предложен способ модификации клетки млекопитающего посредством сайт-направленной интеграции в хромосому экзогенного, в частности, гетерологичного GOI, гена, представляющего интерес (GOI), в частности, гетерологичного GOI, в пределах межгенной области между парой соседних незаменимых генов, предпочтительно, где каждый из соседних незаменимых генов указанной пары имеет сайт начала транскрипции (TSS), и расстояние между TSS представляет собой по меньшей мере любое из 80, 90, 100 или 110 нт и является меньше любого из 20000, 15000, 10000, 5000, 2500 или 1000 нт.

Согласно частному воплощению GOI (или экспрессионную кассету, содержащую GOI) встраивают способом, в котором используется нуклеаза. Нуклеаза может включать ДНК-связывающий домен для связывания целевого сайта и домен расщепления нуклеазы. В частности, нуклеаза применяется для создания двуцепочечного разрыва цепи (DSB) в целевом сайте или в межгенной области.

Двухцепочечный разрыв цепи может способствовать встраиванию трансгена посредством негомологичного соединения концов или в присутствии донорской последовательности, несущей гомологичные области, посредством гомологичной репарации.

В частности, в способе используют любую программируемую нуклеазу, такую как модифицированная эндонуклеаза, предпочтительно, нуклеаза с цинковыми пальцами (ZFN), нуклеаза на основе эффектора, подобного активатору транскрипции (TALEN), эндонуклеаза CRISPR или белок Аргонавт (Argonaute) для интеграции в хромосому. В частности, в способе используют белок, редактирующий гены, выбранный из CRISPR/Cas9, TALEN и нуклеазы с цинковыми пальцами.

В частности, в способе используют нуклеиновокислотную последовательность редактирования гена. Нуклеиновокислотная последовательность редактирования гена может кодировать молекулу редактирования гена, выбранную из группы, состоящей из: специфичной к последовательности нуклеазы, одной или более гидовой РНК, CRISPR/Cas, MAD7, рибонуклеопротеина (RNP), или дезактивированной CAS для систем CRISPRi или CRISPRa или любой их комбинации. Эндонуклеазную активность белка можно оценивать способами, известными специалистам в области техники.

В частности, предпочтительными клетками-хозяевами являются трансдуцированные системой CRISPR-Cas.

Также можно использовать белки семейства Аргонавт, как хорошо известно в области техники. Например, было показано, что TtAgo, PfAgo и NgAgo вызывают направленные двухцепочечные разрывы ДНК, которые можно использовать для облегчения встраивания трансгена путем гомологичной репарации (Swarts et al. Nature. 2014;507(7491):258–261).

В частности, в способе используют средства или стадии способа технологии редактирования генома для встраивания GOI в целевой сайт.

Согласно частному аспекту, межгенную область модифицируют для встраивания, по меньшей мере одного или более из следующего:

- (1) последовательности нуклеиновой кислоты редактирования гена;
- (2) целевого сайта для одной или более нуклеаз;
- (3) GOI или
- (4) сайта распознавания гидовой РНК (gRNA) для ДНК-эндонуклеазы, направляемой РНК.

Согласно частному воплощению, GOI встроен в геном клетки посредством гомологичной рекомбинации. Например, донорская последовательность, содержащая GOI, может быть встроена в хромосому в целевом сайте или рядом с ним посредством гомологичной рекомбинации. В частности, для нокина желаемой последовательности нуклеиновой кислоты может применяться вектор для редактирования гена.

Для этого обычно используется вектор, например, вектор на основе вируса, такой как вектор на основе аденоассоциированного вируса, содержащий 5' гомологический участок целевого сайта, экспрессионную кассету, содержащую GOI, и 3' гомологический участок целевого сайта, где 5' гомологический участок и 3' гомологический участок связываются с целевым сайтом, расположенным в межгенной области, в частности, внутри межгенной области, и где 5' и 3' гомологические участки направляют гомологичную рекомбинацию в локус (в частности, локус «безопасной гавани»), расположенный внутри межгенной области. Приведенные в качестве примера 5' и 3' гомологические участки имеют длину 30-2000 пн. В частности, 5' и 3' гомологические участки связываются с целевыми сайтами, которые являются пространственно различными нуклеиновокислотными последовательностями в одной и той же межгенной области. Обычно используют определенную степень гомологии, например, любой из гомологических участков или оба гомологических участка могут быть комплементарны целевой последовательности в составе межгенной области с любой гомологией из по меньшей мере 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% для гибридизации с такой комплементарной целевой последовательностью или быть на 100% гомологичными и комплементарными целевой последовательности.

В частных способах гомологичной рекомбинации, описанных в данном документе, может использоваться вектор, содержащий:

а) экспрессионную кассету, содержащую промотор и по меньшей мере один GOI, или содержащую промотор, функционально связанный по меньшей мере с одним GOI; и

б) две самокомплементарных последовательности, например, асимметричные или симметричные, фланкирующие указанную экспрессионную кассету, которые представляют собой 5' гомологический участок, гомологичный нуклеотидной последовательности, расположенной в направлении против хода транскрипции от сайта расщепления нуклеазой, и 3' гомологический участок, гомологичный нуклеотидной

последовательности, расположенной в направлении по ходу транскрипции от сайта расщепления нуклеазой.

В изобретении дополнительно предложен набор средств для генной инженерии клетки-хозяина, содержащий:

- а) описанный в данном документе шаблон клетки-хозяина и
- б) вектор и, возможно, вспомогательные агенты, адаптированные для указанных средств для сайт-направленной интеграции в хромосому экспрессионной кассеты, экспрессирующей GOI после сайт-направленной интеграции.

В частности, GOI обеспечивается или содержится в экспрессионной кассете, внедряемой в целевой сайт. Экспрессионная кассета может содержать один или более регуляторных элементов, в частности, промотор, и по меньшей мере один GOI. В частности, экспрессионная кассета является гетерологичной.

В изобретении дополнительно предложено вещество, композиция или материал, выбранные из клетки-хозяина, клона, препарата, культуры клеток, шаблона клетки-хозяина и набора инструментов, описанных в данном документе, для применения в медицине, в частности, для клеточной терапии (например, с использованием CAR-T клеток или CAR макрофагальных клеток), T клеточных рецепторов (для TCR терапии) или кДНК для генной терапии.

Частное применение генной терапии включает лечение расстройств, которые либо вызваны недостатком секретируемого генного продукта, либо которые поддаются лечению секрецией белка терапевтического назначения. Возможным решением при таких расстройствах может стать доставка GOI терапевтического назначения в ряд клеток, при условии, что каждая клетка-реципиент экспрессирует высокий уровень GOI терапевтического назначения. Такие применения обычно требуют стабильной, безопасной экспрессии трансгена на высоком уровне.

GOI терапевтического назначения может представлять собой кДНК, которая при интеграции в хромосому клетки-мишени экспрессируется для реализации терапевтической активности, в частности, для генной терапии.

Приведенные в качестве примера способы генной терапии используют GOI терапевтического назначения, экспрессирующий фактор свертывания, например фактор VIII или фактор IX, для применения в способе лечения гемофилии, такой как гемофилия А или В.

В частности, вещество, композиция или материал используются, как описано в данном документе, когда субъекта лечат путем экспрессии указанного GOI *in vivo*, например, у трансгенного животного, такого как человек или животное, не являющееся человеком. В частности, применение *in vivo* предназначено для генной терапии.

В изобретении дополнительно предложен способ лечения субъекта, включающий введение эффективного количества вещества, композиции или материала, выбранных из клетки-хозяина, клона, препарата, культуры клеток, шаблона клетки-хозяина и набора инструментов, описанных в данном документе, для экспрессии указанного GOI *in vivo*.

Согласно частному аспекту, медицинское применение включает лечение онкологического пациента способом, включающим:

а) выделение клетки у субъекта-млекопитающего, предпочтительно, аутологичной или аллогенной для онкологического пациента, где указанная клетка представляет собой иммунную клетку или гемопоэтическую клетку;

б) модификацию клетки одной или более генетическими модификациями для интеграции в хромосому гетерологичного гена, представляющего интерес (GOI) в целевой сайт в пределах межгенной области между парой соседних незаменимых генов, где GOI кодирует химерный антигенный рецептор (CAR); и

в) введение указанной клетки субъекту.

В частности, иммунная клетка представляет собой Т-клетку, микроглию или макрофаг.

В изобретении дополнительно предложен способ получения Т-клетки с химерным антигенным рецептором (CAR), включающий:

(а) получение Т-клетки от субъекта-млекопитающего, предпочтительно, человеческой Т-клетки, и

б) модификацию Т-клетки одной или более генетическими модификациями для интеграции в хромосому гетерологичного гена, представляющего интерес (GOI), в целевой сайт в пределах межгенной области между парой соседних незаменимых генов, где GOI кодирует химерный антигенный рецептор (CAR).

В изобретении дополнительно предложен способ получения или продуцирования микроглии или макрофага с химерным антигенным рецептором (CAR), включающий:

(а) получение микроглии или макрофага от субъекта-млекопитающего, предпочтительно, человеческой микроглии или макрофага, и

(б) модификацию микроглии или макрофага одной или более генетическими модификациями для интеграции в хромосому гетерологичного гена, представляющего интерес (GOI) в целевой сайт в пределах межгенной области между парой соседних незаменимых генов, где GOI кодирует химерный антигенный рецептор (CAR).

В частности, модифицированная микроглия или макрофаг характеризуются признаками клетки-хозяина, как описано далее.

В частности, межгенная область между парой соседних незаменимых генов характеризуется, как описано далее, например, характеризуется одним или более из следующих предпочтительных признаков:

а) межгенная область расположена между парой соседних незаменимых генов, где каждый из соседних незаменимых генов указанной пары имеет сайт начала транскрипции (TSS), и расстояние между TSS составляет менее 20000 нт и, предпочтительно, по меньшей мере 80 нт;

б) длина межгенной области составляет менее 20000 нт;

в) межгенная область содержит или состоит из последовательности нуклеиновой кислоты, которая по меньшей мере на 90% идентична любой из SEQ ID NO:1-21;

г) незаменимые гены указанной пары ориентированы таким образом, что их промоторы обращены в противоположных направлениях; или

д) GOI содержится в гетерологичной экспрессионной кассете, интегрированной в целевой сайт. В изобретении также предложено применение *ex vivo* вещества, композиции или материала, выбранных из клетки-хозяина, клона, препарата, культуры клеток, шаблона клетки-хозяина и набора инструментов, описанных в данном документе, для инженерии *in vitro* трансгенных клеток или рекомбинантных клеток-хозяев. Такие клетки могут быть предложены для экспрессии GOI *in vitro*, например, в культуре клеток, или экспрессии GOI *in vivo*, например, при введении трансгенных клеток субъекту, нуждающемуся в таком клеточном лечении.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

Фиг. 1: Экспрессия трансгена в указанном локусе не подвергается сайленсингу. Трансгены встраивают в локусе, находящемся между парой незаменимых генов, где промоторы незаменимых генов расположены в тесной близости. Эпигенетический сайленсинг трансгена распространяется на один из соседних незаменимых генов (или

оба). Это будет приводить к гибели клетки. Следовательно, все выжившие клетки поддерживают экспрессию трансгена на высоком уровне.

Фиг. 2: Экспериментальный подход для точного картирования области, позволяющей встраивать трансген. Встраивание трансгена может оказать негативное влияние на жизнеспособность клетки. Для картирования точного расположения, где встраивание трансгена допустимо, область между двумя незаменимыми генами покрывают массивом гидовых РНК. Для сравнения, подходящий массив гидовых РНК также был нацелен на соседние незаменимые гены. Гидовые РНК, активность которых хорошо переносится клеткой, выявляют в ходе последующего CRISPR-скрининга.

Фиг. 3: Экспериментальный подход для оценки влияния сайленсинга на специфический локус. На локус между двумя незаменимыми генами нацеливают экспрессионную кассету, несущую ген-репортер (eGFP), активность которого регулируется подходящим конститутивным промотором (SFFV). Рядом включают массив сайтов TetO. При встраивании трансгена клетки становятся положительными по GFP, поскольку поддерживают экспрессию трансгена. Затем популяцию GFP-положительных клеток подвергают воздействию TetR-KRAB. TetR-KRAB осуществляет сайленсинг сайтов TetO, что приводит к подавлению экспрессии GFP. Сайленсинг распространяется на соседние незаменимые гены, клетки, в которых произошел сайленсинг GFP, погибают. Отрицательные по GFP выжившие клетки не выделяются.

Фиг. 4: CRISPR-скрининг для выявления подходящих локусов для межгенного нацеливания. В данном скрининге на незаменимые гены, межгенные области и заменимые гены направленно воздействовали массивом sgRNA (как описано в Примерах). Клетки собирали на 3 день и 21 день, выделяли геномную ДНК и амплифицировали гидовые РНК при помощи ПЦР. На данном рисунке отмечена кратность изменения между 21 и 3 днем для каждой sgRNA, использованной в данном эксперименте.

Фиг. 5: CRISPR-скрининг для выявления подходящих локусов для межгенного нацеливания. В данном скрининге на незаменимые гены, межгенные области и заменимые гены нацеливали массив sgRNA (как описано в Примерах). Клетки собирали на 3 день и 21 день, выделяли геномную ДНК и амплифицировали гидовые РНК при помощи ПЦР. На данном графике нанесена совокупная кратность изменения между 21 и 3 днем на уровне каждого гена/локуса.

Фиг. 6: Схема встраивания трансгена в различные локусы верификации посредством ПЦР. Осуществляли нацеливание на три локуса (AAVS1, MED, FTSJ3), как описано в Примерах. Присутствие трансгенной кассеты (SV40-mKate2-polyA) верифицировали посредством ПЦР, как указано.

Фиг. 7: Подтверждение экспрессии трансгена посредством ПЦР. Выделяли геномную ДНК из клеток, несущих встроенный трансген между двумя незаменимыми генами (подробно описано в Примерах). Затем проводили ПЦР, в которой один праймер был нацелен на трансгенную кассету, а другой праймер был нацелен на область за пределами гомологических участков (см. схему на Фиг. 6).

Фиг. 8: Экспрессию трансгена определяли при помощи проточной цитометрии. Клетки, несущие трансген SV40-mKate2-polyA в двух локусах (MED20, FTSJ3), анализировали при помощи проточной цитометрии. Экспрессию mKate2 регистрировали в канале PE.

Фиг. 9: Последовательности, описанные в данном документе.

SEQ ID NO:1-21: межгенные области у человека

SEQ ID NO:26:

1 донор MED20 (сайты TetO показаны жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом).

SEQ ID NO:27:

2 донор MED20 (сайты TetO показаны жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом).

SEQ ID NO:28:

1 донор FTSJ3 (сайты TetO показаны жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом).

SEQ ID NO:29:

2 донор FTSJ3 (сайты TetO показаны жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом).

SEQ ID NO:30:

TetR-KRAB

SEQ ID NO:35:

Донор гомологии MED20 (левый гомологический участок показан жирным шрифтом, правый гомологический участок выделен курсивом)

SEQ ID NO:36:

Донор гомологии FTSJ3 (левый гомологический участок показан жирным шрифтом, правый гомологический участок выделен курсивом)

SEQ ID NO:37:

Экспрессирующий вектор Cas9 (с интегрированной sgRNA AAVS1)

SEQ ID NO:38:

Донор гомологии AAVS1 (левый гомологический участок показан жирным шрифтом, правый гомологический участок выделен курсивом, трансгенная кассета обозначена заглавными буквами, промотор SV40 подчеркнут, mKate2 подчеркнут и выделен жирным шрифтом, сигнал полиаденилирования выделен жирным шрифтом и курсивом).

SEQ ID NO:39 (sgRNA MED20): GGGCGTGTCTCGGCACCCCT

SEQ IDNO:40 (sgRNA FTSJ3) GAATTCCGGGTCAATGGGCG

SEQ IDNO:41 (sgRNA AAVS1): gtcctagtggccccactgtg

СВЕДЕНИЯ, ПОДТВЕРЖДАЮЩИЕ ВОЗМОЖНОСТЬ ОСУЩЕСТВЛЕНИЯ ИЗОБРЕТЕНИЯ

Если не указано или не определено иное, все термины, используемые в данном документе, имеют общепринятое в области техники значение, понятное специалисту в области техники. Например, упоминаются стандартные руководства, такие как Sambrook et al, "Molecular Cloning: A Laboratory Manual" (2nd Ed.), Vols. 1 -3, Cold Spring Harbor Laboratory Press (1989); Lewin, "Genes IV", Oxford University Press, New York, (1990) и Janeway et al, "Immunobiology" (5е издание или более поздние издания), Garland Science, New York, 2001.

Объекты формулы изобретения, в частности, относятся к продуктам искусственного происхождения или способам, в которых применяют или получают такие продукты искусственного происхождения, которые могут представлять собой варианты нативных продуктов (дикого типа). Несмотря на возможную некоторую степень идентичности последовательностей с нативной структурой, очевидно, что материалы, способы и применения по изобретению, например, в частности, относящиеся к последовательностям выделенных нуклеиновых кислот, последовательностям аминокислот, слитым конструкциям, экспрессионным конструкциям, трансформированным клеткам-хозяевам и модифицированным белкам, являются «рукотворными», или синтетическими, и следовательно, не считаются следствием «законов природы».

Специфические термины, используемые в описании, имеют следующее значение.

Термин «клетка-хозяин», используемый в данном документе, понимают как любой тип клеток, который восприимчив к трансформации, трансфекции, трансдукции или тому подобному нуклеиновокислотными конструкциями или экспрессирующими векторами, содержащими один или более полинуклеотидов, кодирующих продукты экспрессии, описанные в данном документе, или восприимчивые к внедрению иным образом любого или каждого из компонентов комплекса CRISPR, описанного в данном документе. Термин «клетка-хозяин» охватывает любое потомство родительской клетки, которое не идентично родительской клетке вследствие модификаций, например, способом, описанным в данном документе, или которое возникает в ходе репликации и, в частности, будет относиться к единственной клетке, единственному клеточному клону, популяции клеток, такой как популяция, содержащая поликлональные клетки, или клеточной линии клетки-хозяина.

Термин «клеточная линия», используемый в данном документе, относится к установленному клону конкретного типа клеток, который приобрел способность пролиферировать в течение длительного периода времени. Клеточная линия обычно используется для экспрессии эндогенного или рекомбинантного гена или продукта метаболического пути для получения полипептидов или клеточных метаболитов, опосредованных такими полипептидами. «Продуцирующую линию клеток-хозяев» или «продуцирующую клеточную линию», используемые в данном документе для продуцирования ROI, обычно понимают как линию клеток, готовых для применения для культивирования клеток в биореакторе для получения продукта способа продуцирования. Рекомбинантный ROI можно продуцировать с помощью клетки-хозяина и соответствующей линии клеток, описанной в данном документе, путем культивирования в соответствующей среде, выделения продукта экспрессии или метаболита из культуры и, возможно, его выделения подходящим способом.

В данном документе термин «экспрессия» относится к способу, которым полинуклеотид транскрибируется с матрицы ДНК (например, в мРНК или другой РНК-транскрипт) и/или способом, которым транскрибируемая мРНК впоследствии транслируется в пептиды, полипептиды или белки. Транскрипты и кодируемые полипептиды могут совместно обозначаться продуктом экспрессии гена («генным продуктом»), таким как продуцированный в результате экспрессии GOI. Если

полинуклеотид является производным геномной ДНК, экспрессия может включать сплайсинг мРНК в эукариотической клетке.

Термин «экспрессионная кассета», используемый в данном документе, относится к нуклеотидной последовательности или соответствующего полинуклеотида или молекуле нуклеиновой кислоты, содержащей требуемую кодирующую последовательность и регуляторную или контролируемую последовательность в функциональной связи, так чтобы клетки-хозяева, трансформированные или трансфицированные такими последовательностями, были способны продуцировать кодируемые белки или метаболиты клетки-хозяина. Примеры контролирующих экспрессию последовательностей могут включать любой промотор, сайт связывания рибосомы, последовательности начала и окончания транскрипции или трансляции либо последовательность энхансера или активатора.

Экспрессионная кассета может содержать по меньшей мере один интрон. Обычно интроны располагаются на 5' конце открытой рамки считывания, но могут также располагаться на 3' конце. Указанный интрон может находиться между промотором и/или элементами промотора/энхансера на 5' конце открытой рамки считывания полинуклеотида, кодирующего продукт, представляющий интерес для экспрессии. В области техники известно несколько подходящих интронов, которые могут применяться в данном изобретении.

Для осуществления трансформации экспрессионная система может быть включена в экспрессионную конструкцию, например, в форме «вектора» или экспрессионных кассет, интегрированных в хромосому клетки-хозяина. Экспрессионные кассеты обычно представляют собой последовательности ДНК, которые необходимы для транскрипции клонированных рекомбинантных нуклеотидных последовательностей, то есть рекомбинантных генов и трансляции их мРНК в подходящем организме-хозяине. Экспрессирующие векторы обычно содержат один или более ориджинов для автономной репликации или локус для интеграции в геном клетки-хозяина, маркеры для селекции (например, ген синтеза аминокислот или ген, придающий резистентность к антибиотикам, таким как зеоцин, канамицин, G418 или гигромицин, нourseотрицин), ряд сайтов расщепления рестрикционными ферментами и регуляторные последовательности, например, любые одну или более подходящих последовательностей промотора, оператора, энхансера, сайта связывания рибосомы и последовательностей, которые контролируют инициацию и терминацию

транскрипции и трансляции. Регуляторные последовательности обычно функционально связаны с последовательностью ДНК, которую следует экспрессировать.

«Вектор» в данном документе понимают как способный к переносу нуклеиновокислотных последовательностей в клетки-мишени (например, векторы на основе вируса, невирусные векторы, корпускулярные носители и липосомы). Обычно вектор понимают как «экспрессирующий вектор» или «вектор, переносящий ген», что означает любую нуклеиновокислотную конструкцию, способную направлять экспрессию нуклеиновой кислоты, представляющей интерес, и способную переносить нуклеиновокислотные последовательности в клетки-мишени. Таким образом, термин охватывает клонирующие и экспрессирующие носители, а также векторы на основе вируса.

Специфические векторы включают автономно реплицирующиеся нуклеотидные последовательности (например, плазмиды), а также интегрирующиеся в геном нуклеотидные последовательности, такие как искусственные хромосомы, например, искусственную хромосому дрожжей (YAC). Экспрессирующие векторы могут включать клонирующие векторы, модифицированные клонирующие векторы и специально сконструированные плазмиды, без ограничения. Предпочтительные экспрессирующие векторы, описанные в данном документе, представляют собой экспрессирующие векторы, подходящие для экспрессии рекомбинантного гена в эукариотической клетке-хозяине и отбираются в зависимости от организма-хозяина.

Соответствующая ДНК может быть интегрирована в хромосому клетки-хозяина. Экспрессия клеткой-хозяином может относиться к секретируемым или несекретируемым продуктам экспрессии, включая полипептиды или метаболиты. Для обеспечения экспрессии рекомбинантной нуклеотидной последовательности в клетке-хозяине экспрессионная кассета или вектор, описанные в данном документе, обычно содержат промотор, функционально связанный с GOI, например, нуклеотидную последовательность промотора, примыкающую к 5' концу кодирующей последовательности или примыкающую к гену, представляющему интерес (GOI), в направлении против хода транскрипции, или, если применяется сигнальная или лидерная последовательность, примыкающую к указанной сигнальной или лидерной последовательности, соответственно, в направлении против хода транскрипции, для облегчения экспрессии GOI. Последовательность промотора обычно регулирует и

инициирует транскрипцию расположенной по ходу транскрипции последовательности, с которой она функционально связана, в частности, включая GOI.

В частности, промотор представляет собой гетерологичный промотор, в частности, гетерологичный для клетки-хозяина и/или не ассоциированный нативным образом с GOI.

В частных воплощениях можно применять векторы для множественного клонирования, которые представляют собой векторы, имеющие сайт для множественного клонирования. В частности, необходимый гетерологичный ген может быть интегрирован или внедрен в сайт множественного клонирования для получения экспрессирующего вектора. В случае векторов для множественного клонирования промотор обычно размещается в направлении против хода транскрипции от сайта множественного клонирования.

В частности, в предпочтительных воплощениях системы CRISPR используется система доставки, содержащая один или более векторов, возможно, когда векторы содержат один или более векторов на основе вирусов. В частности, могут применяться один или более векторов на основе вирусов, выбранных из группы, состоящей из лентивируса, ретровируса, аденовируса, аденоассоциированного вируса или вируса простого герпеса, лентивирусных, аденовирусных или аденоассоциированных вирусных (AAV) векторов. Частные примеры выбраны из группы, состоящей из лентивирусов на основе ВИЧ. Лентивирусные векторы могут нести определенные признаки безопасности, например, они могут быть основаны на нескольких пакующих плаزمиде или они могут иметь укороченные длинные концевые повторы. Все эти признаки считают понижающими вероятность получения способного к репликации вируса, то есть обычно эти вирусы могут проходить только один цикл инфицирования.

В некоторых альтернативных воплощениях элементы системы CRISPR доставляют через липосомы, частицы, проникающие в клетку пептиды, экзосомы, микровезикулы или генную пушку либо посредством электропорации клетки-мишени.

Уровень экспрессии можно определять количественно или качественно путем измерения уровня мРНК или белка по меньшей мере одного указанного гена, в частности, уровня продукта экспрессии целевого гена. В частности, уровень экспрессии можно определять путем определения транскриптома клетки-хозяина или путем оценки уровней белков на поверхности клетки при помощи проточной цитометрии.

В некоторых воплощениях уровень экспрессии определяют способом, выбранным из группы, состоящей из количественной RT-PCR (qPCR), секвенирования РНК, гибридизации с микрочипом и молекулярной маркировки.

Термин «экспрессия гена» конкретно относится к преобразованию информации, содержащейся в гене, в продукт гена. Экспрессия гена охватывает по меньшей мере одну стадию, выбранную из группы, состоящей из транскрипции ДНК в мРНК, процессинга мРНК, созревания мРНК, экспорта мРНК, трансляции, фолдинга белка и/или транспорта белка.

В данном описании термин «ген», например, используемый в термине «ген, представляющий интерес», или «GOI» охватывает область ДНК, кодирующую генный продукт и, возможно, одну или более областей ДНК (или все области ДНК), которые регулируют продукцию генного продукта, независимо от того, примыкают ли такие регуляторные последовательности к кодирующим и/или транскрибируемым последовательностям. Соответственно, ген можно понимать, как включающий в себя последовательности промотора, терминатора, регулирующие трансляцию последовательности, такие как сайты связывания рибосомы и сайты внутренней посадки рибосомы, энхансеры, сайленсеры, инсуляторы, граничные элементы, ориджины репликации, сайты прикрепления к матриксу и регуляторные области локусов.

Согласно следующим частным примерам, GOI, используемый в контексте настоящего изобретения, представляет собой кДНК, и/или кодирует биологический продукт, лиганд (такой как антитело) или рецептор, такой как химерный антигенный рецептор (CAR) или Т-клеточный рецептор.

Подходящий CAR может включать антиген-связывающую часть или домен (в качестве внеклеточного связывающего участка), трансмембранный домен и внутриклеточный сигнальный домен и, возможно, дополнительно включать одну или более линкерных последовательностей между различными доменами.

В частности, CAR включает антиген-связывающий участок, который связывается с целевым антигеном, представляющим интерес, например, с конкретным антигеном на поверхности клетки-мишени. Например, антиген-связывающий участок, который может содержаться в CAR, может включать антитело, рецептор (например, вариабельный Т-клеточный рецептор или рецептор лимфоцита), фрагмент рецептора

(например, фрагмент Fc рецептора), лиганд, цитокин, дарпин, аднектин, наноантитело и пептид.

Антитело, входящее в антигенсвязывающий участок, предпочтительно способно специфически распознавать антиген, имеющий клиническую значимость, и может быть выбрано из группы, состоящей из полноразмерного антитела, фрагмента антитела, содержащего по меньшей мере один переменный домен антитела или антигенсвязывающий сайт (предпочтительно, одноцепочечный переменный фрагмент, scFv).

Трансмембранный домен специфически объединяет внеклеточный связывающий участок и внутриклеточный сигнальный домен и закрепляет CAR в плазматической мембране клетки. В области техники известны различные подходящие трансмембранные домены.

Согласно частному аспекту, используется по меньшей мере один внутриклеточный сигнальный домен CAR.

Внутриклеточный сигнальный домен CAR, в частности, является частью CAR, участвующей в передаче сигнала от связывания CAR с клеткой-мишенью внутри иммунной эффекторной клетки для запуска эффекторной функции клетки, например, активации, продукции цитокина, пролиферации и/или цитотоксической активности, включая высвобождение цитотоксических факторов в связанную с CAR клетку-мишень, или другие клеточные ответы, вызванные связыванием мишени с внеклеточным доменом CAR.

Внутриклеточный сигнальный домен CAR отвечает за активацию по меньшей мере одной из обычных эффекторных функций иммунной клетки, в которую помещен CAR. Эффекторная функция Т-клетки, например, может представлять собой цитолитическую активность или хелперную активность, включая секрецию цитокинов. В частности, термин «внутриклеточный сигнальный домен» относится к части белка, которая передает сигналы эффекторной функции и направляет клетку на выполнение специализированной функции.

Хотя обычно можно использовать весь внутриклеточный сигнальный домен, во многих случаях нет необходимости использовать всю цепь. В случае, если используется укороченная часть внутриклеточного сигнального домена, такая укороченная часть может использоваться вместо интактной цепи, при условии, что она преобразует сигнал в эффекторную функцию. Следовательно, термин

«внутриклеточный сигнальный домен» охватывает любую укороченную часть внутриклеточного сигнального домена, достаточную для преобразования сигнала в эффекторную функцию.

Предпочтительные примеры внутриклеточных сигнальных доменов для применения в CAR включают цитоплазматические последовательности Т-клеточного рецептора (TCR) и ко-рецепторы, которые действуют согласованно для инициации преобразования сигнала после захвата антигена рецептором, а также любое производное или вариант этих последовательностей и любую синтетическую последовательность, которая обладает такими же функциональными способностями.

Активация Т-клеток может быть опосредована двумя различными классами цитоплазматических сигнальных последовательностей: таких, которые иницируют антиген-зависимую первичную активацию через TCR (первичные цитоплазматические сигнальные последовательности) и такие, которые действуют антиген-независимым образом для обеспечения вторичного или ко-стимулирующего сигнала (вторичные цитоплазматические сигнальные последовательности).

Первичные цитоплазматические сигнальные последовательности регулируют первичную активацию комплекса TCR либо стимулирующим образом, либо ингибирующим образом. Первичные цитоплазматические сигнальные последовательности, действующие стимулирующим образом, могут содержать сигнальные мотивы, которые известны как мотивы активации иммунорецептора на основе тирозина, или ITAM.

Примеры ITAM, содержащие первичные цитоплазматические сигнальные последовательности, которые в частности используются в CAR, раскрытых в данном документе, включают производные TCR дзета (CD3 дзета), FcR гамма, FcR бета, CD3 гамма, CD3 дельта, CD3 эпсилон, CD5, CD22, CD79a, CD79b и CD66d. Не являющиеся исчерпывающими частные примеры ITAM включают пептиды, имеющие последовательности аминокислот CD3.дзета с номерами от 51 до 164. (NCBI RefSeq: NP.sub.--932170.1), аминокислот Fc.эпсилон.RI.гамма с номерами от 45 до 86. (NCBI RefSeq: NP.sub.--004097.1), аминокислот Fc.эпсилон.RI.бета с номерами от 201 до 244. (NCBI RefSeq: NP.sub.--000130.1), аминокислот CD3.гамма с номерами от 139 до 182. (NCBI RefSeq: NP.sub.--000064.1), аминокислот CD3.дельта с номерами от 128 до 171. (NCBI RefSeq: NP.sub.--000723.1), аминокислот CD3.эпсилон с номерами от 153 до 207. (NCBI RefSeq: NP.sub.--000724.1), аминокислот CD5 с номерами от 402 до 495 (NCBI

RefSeq: NP.sub.--055022.2), аминокислот 0022 с номерами от 707 до 847 (NCBI RefSeq: NP.sub.--001762.2), аминокислот CD79a с номерами от 166 до 226 (NCBI RefSeq: NP.sub.--001774.1), аминокислот CD79b с номерами от 182 до 229 (NCBI RefSeq: NP.sub.--000617.1) и аминокислот CD66d с номерами от 177 до 252 (NCBI RefSeq: NP.sub.--001806.2), а также их варианты, обладающие такой же функцией, как эти пептиды. Номера аминокислот, основанные на информации об аминокислотных последовательностях, взятой из NCBI RefSeq ID или GenBank, описанные в данном документе, приведены для полноразмерного предшественника (содержащего последовательность сигнального пептида) каждого белка.

В одном воплощении цитоплазматическая сигнальная молекула в составе CAR содержит цитоплазматическую сигнальную последовательность, происходящую из CD3 дзета.

В частном воплощении внутриклеточный домен CAR может быть сконструирован так, чтобы содержать сигнальный домен CD3-дзета сам по себе или в комбинации с любым необходимым цитоплазматическим доменом, который может применяться в составе CAR. Например, внутриклеточный домен CAR может содержать часть цепи CD3 дзета и костимулирующую сигнальную область. Костимулирующая сигнальная область относится к части CAR, содержащей внутриклеточный домен костимулирующей молекулы. Костимулирующая молекула представляет собой молекулу клеточной поверхности, отличную от антигенного рецептора или их лигандов, которая необходима для эффективного ответа лимфоцитов на антиген. Примеры таких костимулирующих молекул включают CD27, CD28, 4-1BB (CD137), OX40, CD30, CD40, PD-1, ICOS, связанный с функцией лимфоцитов антиген-1 (LFA-1), CD2, CD7, LIGHT, NKG2C, B7-H3 и лиганд, специфически связывающийся с CD83, а также им подобные.

Частные, не являющиеся исчерпывающими, примеры таких костимулирующих молекул включают пептиды, имеющие последовательности аминокислот CD2 с номерами от 236 до 351 (NCBI RefSeq: NP.sub.--001758.2), аминокислот CD4 с номерами от 421 до 458 (NCBI RefSeq: NP.sub.--000607.1), аминокислот CD5 с номерами от 402 до 495 (NCBI RefSeq: NP.sub.--055022.2), аминокислот CD8.альфа с номерами от 207 до 235 (NCBI RefSeq: NP.sub.--000607.1), аминокислот CD83 с номерами от 196 до 210 (GenBank: AAA35664.1), аминокислот CD28 с номерами от 181 до 220 (NCBI RefSeq: NP.sub.--006130.1), аминокислот CD137 с номерами от 214 до 255

(4-1BB, NCBI RefSeq: NP.sub.--001552.2), аминокислот CD134 с номерами от 241 до 277 (OX40, NCBI RefSeq: NP.sub.--003318.1) и аминокислот ICOS с номерами от 166 до 199 (NCBI RefSeq: NP.sub.--036224.1), а также их варианты, обладающие такой же функцией, как эти пептиды. Таким образом, хотя данное изобретение проиллюстрировано в основном на примере 4-1BB в качестве костимулирующего сигнального элемента, в объем изобретения входят другие костимулирующие элементы.

Цитоплазматические сигнальные последовательности в составе цитоплазматического сигнального участка CAR могут быть связаны друг с другом в случайном или определенном порядке. Возможно, связь может быть образована коротким олиго- или полипептидным линкером, предпочтительно, длиной от 2 до 10 аминокислот. Наиболее подходящим линкером является тандем глицин-серин.

В частном воплощении внутриклеточный домен сконструирован таким образом, чтобы содержать сигнальный домен CD3-дзета и сигнальный домен CD28. В другом воплощении внутриклеточный домен сконструирован таким образом, чтобы содержать сигнальный домен CD3-дзета и сигнальный домен 4-1BB. В еще одном воплощении внутриклеточный домен сконструирован таким образом, чтобы содержать сигнальный домен CD3-дзета и сигнальный домен CD28 и 4-1BB.

Частные предпочтительные воплощения включают одно или более из следующего:

- (а) внутриклеточный сигнальный домен, содержащий внутриклеточный домен CD3-дзета,
- (б) внутриклеточный сигнальный домен, содержащий костимулирующий домен, первичный сигнальный домен или любую их комбинацию,
- (в) внутриклеточный сигнальный домен, содержащий костимулирующий домен, содержащий функциональный сигнальный домен белка, выбранного из группы, состоящей из: OX40, CD70, CD27, CD28, CD5, ICAM-1, LFA-1 (CD11a/CD18), ICOS (CD278), DAP10, DAP12 и 4-1BB (CD137).

Согласно частному аспекту, CAR специфически распознает и связывается с молекулой на поверхности клетки-мишени. Согласно частным примерам, клетка-мишень представляет собой опухолевую или раковую клетку, а CAR специфически распознает опухолеассоциированный антиген (ТАА), предпочтительно, любой из

CD19, CD20, CD22, B7-H3 (CD276), CD133, GD2, EGFRvIII, BMSA, MSLN, CEA и HER2.

Согласно дополнительным частным примерам, GOI, используемый в контексте настоящего изобретения, выбран из группы, состоящей из ДНК-вакцин или молекул ДНК терапевтического назначения, например, молекулы ДНК терапевтического назначения, которая

(а) экспрессирует функциональный ген у субъекта, имеющего генетическое нарушение, вызванное дисфункциональной версией указанного функционального гена (например, гена при мышечной дистрофии Дюшена, муковисцидозе, болезни Гоше и дефиците аденозиндезаминазы (АДА), воспалительных заболеваниях, аутоиммунных, хронических и инфекционных заболеваниях, СПИДе, раке, неврологических заболеваниях, сердечно-сосудистых заболеваниях, гиперхолестеринемии, различных гематологических нарушениях (включая различные анемии, талассемию и гемофилию, а также эмфизему) и солидных опухолях);

(б) кодирует токсичный пептид (то есть такие химиотерапевтические агенты, как ризин, дифтерийный токсин и фактор из яда кобры), гены-супрессоры опухолей (такие как p53), гены, кодирующие последовательности мРНК, которые являются бессмысленными по отношению к трансформирующим онкогенам, противоопухолевые пептиды, такие как фактор некроза опухоли (TNF) и другие цитокины, или транс-доминантные негативные мутанты трансформирующих онкогенов;

(в) кодирует активную форму РНК (например, малую интерферирующую РНК (siRNA, miRNA, shRNA) или малую активирующую РНК (saRNA); или

(г) кодирует компонент CRISPR/Cas (такой как фермент Cas9 или sgRNA).

Согласно дополнительным частным примерам, GOI, используемый в рамках настоящего изобретения, экспрессирует белок, представляющий интерес (POI), например, в культуре клеток.

В частности, POI является гетерологичным для видов клетки-хозяина.

В частности, POI представляет собой секретируемый пептид, полипептид или белок, то есть секретируется из клетки-хозяина в культуральную надосадочную жидкость.

В частности, POI представляет собой эукариотический белок, предпочтительно, происходящий из млекопитающего или родственного белок, такой как человеческий

белок или белок, содержащий последовательность человеческого белка, либо бактериальный белок или белок, происходящий из бактерий.

Предпочтительно, POI представляет собой белок терапевтического назначения, функционирующий у млекопитающих.

В частных случаях POI представляет собой мультимерный белок, в частности, димер или тетрамер.

Согласно частному аспекту, POI представляет собой пептид или белок, выбранный из группы, состоящей из антигенсвязывающего белка, терапевтического белка, фермента, пептида, белкового антибиотика, слитого с токсином белка, конъюгата углевод-белок, структурного белка, регуляторного белка, одного или более факторов транскрипции, вакцинного антигена, фактора роста, гормона, цитокина, фермента промышленного назначения, и метаболического фермента.

В частности, антигенсвязывающий белок выбран из группы, состоящей из

а) антител или фрагментов антител, таких как любое из химерных антител, гуманизированных антител, биспецифических антител, Fab, Fd, scFv, диател, триател, тетрамеров Fv, миниантител, однодоменных антител типа VH, VHH, IgNAR или V-NAR;

б) миметиков антител, таких как аднектины, аффитела, аффилины, аффимеры, аффитины, альфатела, антикарины, авимеры, дарпины, финомеры, пептиды с доменом Куница, монотела или NanoCLAMPS; или

в) слитых белков, содержащих один или более доменов с укладкой, характерной для иммуноглобулинов, доменов антител или миметиков антител.

Конкретный POI представляет собой антигенсвязывающую молекулу, такую как антитело или его фрагмент, в частности, фрагмент антитела, содержащий антигенсвязывающий домен. Среди конкретных POI присутствуют антитела, такие как моноклональные антитела (mAb), иммуноглобулины (Ig) или иммуноглобулины класса G (IgG), антитела из тяжелых цепей (HcAb) или их фрагменты, такие как антигенсвязывающий фрагмент (Fab), Fd, одноцепочечный вариабельный фрагмент (scFv) или их генномодифицированные варианты, например, такие как димеры Fv (диатела), тримеры Fv (триатела), тетрамеры Fv, или миниантитела и однодоменные антитела, такие как VH, VHH, IgNAR или V-NAR, или любой белок, содержащий домен с укладкой, характерной для иммуноглобулина. Дополнительные антигенсвязывающие молекулы могут быть выбраны из миметиков антител или (в

качестве альтернативы) каркасных белков, например, таких как генномодифицированные домены Куница, аднектины, аффитела, аффилин, антикалины или дарпины.

Термин «эндогенный», используемый в данном документе, охватывает такие молекулы и последовательности, в частности, эндогенные гены или белки, которые присутствуют в клетке-хозяине дикого типа природного происхождения (нативной). В частности, молекула эндогенной нуклеиновой кислоты (например, ген) или белок, который возникает в конкретной клетке-хозяине (и может быть получен из нее), которая встречается в природе, считается «эндогенной для клетки-хозяина» или «эндогенной по отношению к клетке-хозяину». Более того, клетка, «эндогенно экспрессирующая» нуклеиновую кислоту или белок, экспрессирует такие нуклеиновую кислоту или белок, как и хозяин того же типа, встречающийся в природе. Более того, клетка, «эндогенно продуцирующая» или которая «эндогенно продуцирует» нуклеиновую кислоту, белок или другое соединение, продуцирует такие нуклеиновую кислоту, белок или соединение, как и клетка-хозяин того же типа, которая встречается в природе.

Термин «гетерологичный», используемый в данном документе применительно к нуклеотидной последовательности, такой конструкции, как GOI, промотор, экспрессионная кассета, к аминокислотной последовательности или белку, относится к соединению, которое либо является чужеродным для заданной клетки-хозяина, т.е. «экзогенным», таким, которое не встречается в природе в указанной клетке-хозяине; или которое встречается в природе в заданной клетке-хозяине, например, является «эндогенным», однако нахождение в гетерологичной конструкции или интеграция в такую гетерологичную конструкцию, например, с использованием гетерологичной нуклеиновой кислоты, слитой или комбинированной с эндогенной нуклеиновой кислотой, делает конструкцию гетерологичной. Гетерологичная нуклеотидная последовательность, встречающаяся эндогенно, может также продуцироваться в клетке в количестве, не встречающемся в природе, например, в большем, чем ожидается, или в большем, чем встречается в природе. Гетерологичная нуклеотидная последовательность или нуклеиновая кислота, содержащая гетерологичную нуклеотидную последовательность, возможно отличается по последовательности от эндогенной нуклеотидной последовательности, но кодирует такой же белок, который встречается эндогенно. В частности, гетерологичные нуклеотидные

последовательности не встречаются в такой же взаимосвязи с клеткой-хозяином в природе. Любая рекомбинантная или искусственная нуклеотидная последовательность считается гетерологичной. Примером гетерологичного полинуклеотида является нуклеотидная последовательность, которая не связана нативным образом с промотором, например, для получения гибридного промотора, или функционально связана с кодирующей последовательностью, как описано в данном документе. В результате может быть получен гибридный или химерный полинуклеотид. Следующим примером гетерологичного соединения является кодирующий POI полинуклеотид, функционально связанный с элементом, контролирующим транскрипцию, например с промотором, с которым эндогенная последовательность, кодирующая POI природного происхождения, обычно не связана функционально.

Термин «выделенный» или «выделение», используемый в данном документе, относится к такому соединению, которое достаточно отделено от окружения, с которым оно ассоциировано естественным образом, так чтобы существовать в «очищенной» или «по существу очищенной» форме. Термин «выделенный» может относиться к материалу, который является свободным, достаточно свободным или по существу свободным в различной степени от компонентов, которые обычно сопровождают его в нативном состоянии, встречающемся в природе. «Изолят» также обозначает степень отделения от исходного источника или окружения. При этом «выделенный» не обязательно означает исключение искусственных или синтетических смесей с другими соединениями или материалами или присутствие примесей, которые не интерферируют с фундаментальной активностью и которые могут присутствовать, например, вследствие неполной очистки. Дополнительно, можно готовить составы с выделенными соединениями с получением их препаратов, и при этом, исходя из практических соображений при использовании в диагностике или терапии, выделенные клетки-хозяева или POI, например, могут быть смешаны с фармацевтически приемлемыми носителями или эксципиентами.

В данном описании термин «выделенная клетка» относится к клетке, которая отделена от молекулярных и/или клеточных компонентов, которые естественным образом сопровождают клетку.

Термин «функционально связанный», используемый в данном документе, относится к ассоциации нуклеотидных последовательностей в единой молекуле нуклеиновой кислоты, например, векторе или экспрессионной кассете, таким образом,

что функция одной или более нуклеотидных последовательностей находится под влиянием по меньшей мере одной иной нуклеотидной последовательности, присутствующей в указанной молекуле нуклеиновой кислоты. Посредством функционального связывания нуклеиновокислотная последовательность помещается в функциональную взаимосвязь с другой нуклеиновокислотной последовательностью в одной молекуле нуклеиновой кислоты. Например, промотор функционально связан с кодирующей последовательностью рекомбинантного гена, когда он способен влиять на экспрессию этой кодирующей последовательности. В качестве следующего примера, нуклеиновая кислота, кодирующая сигнальный пептид, функционально связана с GOI или нуклеиновокислотной последовательностью, кодирующей POI, когда она способна экспрессировать белок в секретируемой форме, такой как преформа зрелого белка или зрелый белок. В частности, такие нуклеиновые кислоты, функционально связанные друг с другом, могут быть связаны непосредственно, то есть без дополнительных элементов или нуклеиновокислотных последовательностей между нуклеиновой кислотой, кодирующей сигнальный пептид, и нуклеиновокислотной последовательностью, кодирующей POI.

Термин «рекомбинантный», используемый в данном документе, означает полученный посредством или в результате генетической инженерии. Рекомбинантный хозяин может подвергаться инженерии для встраивания одного или более нуклеотидов, полинуклеотидов или нуклеотидных последовательностей и может, в частности, содержать экспрессирующий вектор или клонирующий вектор, содержащий рекомбинантную нуклеиновокислотную последовательность, в частности, с использованием нуклеотидной последовательности, чужеродной для хозяина. Рекомбинантная клетка-хозяин может быть получена при помощи генетической инженерии, то есть при вмешательстве человека, чтобы внедрить GOI в определенный хромосомный локус. Когда осуществляют инженерию клетки-хозяина для внедрения GOI в локус «безопасной гавани» для стабильной экспрессии GOI, клетку-хозяина обрабатывают таким образом, что клетка-хозяин обладает способностью экспрессировать такой ген на определенном уровне в течение длительного периода культивирования (например, *in situ*, *in vivo*, *ex vivo* или *in vitro*), который превышает уровень экспрессии в клетке-хозяине в тех же условиях до обработки, или по сравнению с клетками-хозяевами, не подвергавшимися инженерии для экспрессии GOI.

Генетическую инженерию удобно осуществлять посредством редактирования генома, гомологичной рекомбинации или других технологий сайт-направленной рекомбинации, таких как рекомбинация Flp-FRT, рекомбинация Cre-Lox или сайт-специфическая рекомбинация у фага лямбда.

Рекомбинацию Flp-FRT понимают как технологию сайт-направленной рекомбинации, и она включает рекомбинацию последовательностей между короткими сайтами-мишенями флиппазы (FRT), распознаваемыми рекомбиназой флиппазой (Flp), происходящей из плазмиды 2 μ пекарских дрожжей *Saccharomyces cerevisiae*.

Рекомбинацию Cre-Lox понимают как сайт-специфическую рекомбиназную технологию. Система состоит из единственного фермента, рекомбиназы Cre, которая осуществляет рекомбинацию пары коротких последовательностей-мишеней, называемых последовательностями *Lox*.

Сайт-специфическая рекомбинация у фага лямбда использует топоизомеразную активность белка Int бактериофага лямбда, внедряющую одноцепочечные разрывы в дуплекс ДНК в распознаваемых последовательностях.

Термин «рекомбинантный», используемый в данном документе в отношении клетки, GOI или POI, включает клетку, молекулу или продукт, представляющие интерес, которые получают, экспрессируют, создают или выделяют рекомбинантными способами, такие как клетка, подвергшаяся инженерии для экспрессии GOI, как описано в данном документе, GOI, интегрированный в хромосому клетки-хозяина, как описано в данном документе, или POI, продуцированный клеткой-хозяином, трансформированной, чтобы экспрессировать POI, как описано в данном документе. В данном изобретении могут применяться стандартные методы молекулярной биологии, микробиологии и рекомбинации ДНК, известные специалистам в области техники. Такие методы полностью описаны в литературе. См., например, Maniatis, Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, (1982).

Термин «идентичность последовательностей» варианта, гомолога или ортолога по сравнению с родительской нуклеотидной или аминокислотной последовательностями указывает на степень идентичности двух или более последовательностей. Две или более аминокислотных последовательности могут иметь в соответствующем положении до некоторой степени, вплоть до 100% одинаковые или консервативные аминокислотные остатки. Две или более нуклеотидных

последовательности могут иметь в соответствующем положении до некоторой степени, вплоть до 100% одинаковые или консервативные пары оснований.

Поиск сходства последовательностей является эффективной и надежной стратегией для выявления гомологов с повышенной (например, по меньшей мере 50%) идентичностью последовательностей. Часто используют такие инструменты поиска последовательностей, как, например, BLAST, FASTA и HMMER.

Поиск сходства последовательностей может выявить такие гомологичные белки или гены путем выявления повышенного сходства и статистически значимого сходства, отражающего наличие общего предшественника. Гомологи могут охватывать ортологи, под которыми понимают в данном документе, один и тот же белок у различных организмов, например, варианты такого белка у различных организмов или биологических видов.

Для определения % комплементарности двух комплементарных последовательностей перед тем, как рассчитать % комплементарности как % идентичности между первой последовательностью и второй конвертированной последовательностью с помощью вышеупомянутого алгоритма одну из двух последовательностей следует конвертировать в комплементарную ей последовательность.

«Процент (%) идентичности аминокислот» в отношении аминокислотной последовательности, гомологов и ортологов, описанных в данном документе, определяют как процент аминокислотных остатков в кандидатной последовательности, которые идентичны аминокислотным остаткам в конкретной полипептидной последовательности, после выравнивания последовательности и введения гэпов, при необходимости, для достижения максимального процента идентичности последовательностей, не считая какие-либо консервативные замены частью идентичности последовательностей. Специалисты в области техники могут определить соответствующие параметры для выравнивания, включая любые алгоритмы, необходимые для максимального выравнивания по всей длине сравниваемых последовательностей.

Для целей, описанных в данном документе, идентичность последовательностей между двумя аминокислотными последовательностями определяют с помощью программы NCBI BLAST версии BLASTP 2.8.1 с параметрами, приведенными в качестве примера ниже: Программа: blastp, Длина сегмента: 6, Ожидаемое значение: 10,

Размер списка совпадений: 100, Цена гэта: 11,1, Матрица: BLOSUM62, Строка фильтра: F, Композиционная корректировка: Условная композиционная корректировка матрицы замен.

«Процент (%) идентичности» в отношении нуклеотидной последовательности, например, молекулы нуклеиновой кислоты или ее части, в частности, кодирующей последовательности ДНК, определяют как процент нуклеотидов в кандидатной последовательности ДНК, которые идентичны нуклеотидам в последовательности ДНК, после выравнивания последовательности и, при необходимости, введения гэпов для достижения максимального процента идентичности последовательностей, не считая какие-либо консервативные замены частью идентичности последовательностей. Выравнивание в целях определения процента идентичности нуклеотидных последовательностей может достигаться различными способами, которые известны в области техники, например, с помощью программного обеспечения, имеющегося в открытом доступе. Специалисты в области техники могут определить соответствующие параметры для выравнивания, включая любые алгоритмы, необходимые для максимального выравнивания по всей длине сравниваемых последовательностей.

Оптимальное выравнивание можно найти с помощью любого подходящего алгоритма для выравнивания последовательностей, не являющиеся исчерпывающими примеры которого включают алгоритм Смита-Ватермана, алгоритм Нидлмана-Вунша, алгоритмы на основе преобразования Барроуза – Уилера (например, программа Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies; доступно по адресу: novocraft.com), ELAND (Illumina, San Diego, CA), SOAP (доступно по адресу: soap.genomics.org.cn) и Maq (доступно по адресу: maq.sourceforge.net).

Термин «стабильная», используемый в данном документе применительно к экспрессии, экспрессирующим или экспрессионным конструкциям, означает, что клетка-хозяин способна корректно экспрессировать GOI в течение длительного периода времени. Стабильно экспрессирующая, в частности, относится к клетке-хозяину, поддерживающей генетические свойства, в частности, поддерживающей экспрессию GOI на высоком и/или почти постоянном уровне, даже в 20 поколениях, предпочтительно по меньшей мере в 30 поколениях, более предпочтительно, по меньшей мере в 40 поколениях, наиболее предпочтительно, по меньшей мере в 50 поколениях. Частные воплощения относятся к стабильной экспрессии по меньшей мере на протяжении 5 или 10, или 20, или 30, или 40, или 50 удвоений или пассажей.

Согласно частному аспекту, предложена популяция клеток-хозяев, где указанные клетки содержат стабильно интегрированный в их хромосому гетерологичный GOI, где в среднем по меньшей мере 20%, по меньшей мере 30%, по меньшей мере 40%, по меньшей мере 50%, по меньшей мере 60% клеток, происходящих из указанной популяции, не теряют более 70%, предпочтительно не более 50% титра экспрессии их генного продукта в течение периода времени по меньшей мере 4 недели, предпочтительно 8 недель, предпочтительно 10 недель, более предпочтительно в течение периода времени 12 недель. Экспрессию можно отслеживать *ex-vivo* (в культуре клеток) или *in vivo* (после трансплантации). Как показано в Примерах, после трансфекции и выявления стабильно трансфицированных клеток количество клеток, не демонстрирующих постепенных потерь продуктивности в ходе длительного культивирования, увеличивается при использовании клеток, описанных в данном документе, то есть из популяции отобранных клеток получают более стабильные клоны клеток. Свойство стабильности можно исследовать путем культивирования отдельных клеток из указанной популяции как клонов клеток и определения титра в течение указанного периода времени. Экспрессию GOI клеткой-хозяином можно определять различными способами, например, при помощи ELISA, Вестерн-блоттинга, радиоиммуноанализа, иммунопреципитации, исследования биологической активности продукта гена или путем иммуноокрашивания с последующим анализом методом FACS. В частном аспекте определяют экспрессию GOI при помощи Вестерн-блоттинга.

Стабильность можно исследовать по отсутствию сайленсинга гена, например, как описано в Примере 3. Ее можно также исследовать путем культивирования клеток в течение длительного периода времени и отслеживания уровней экспрессии трансгена *in vitro* или *in vivo* способами, известными в области техники. Поскольку встраивание трансгена может происходить *in vivo*, экспрессию трансгена можно отслеживать *in vivo*, например, беря образцы у пациента (например, крови или биопсий соответствующей ткани). Как правило, экспрессию считают стабильной, если экспрессия трансгена остается приблизительно такой же ($\pm 50\%$ или $\pm 40\%$ или $\pm 30\%$ или $\pm 20\%$ или $\pm 10\%$) в течение периода наблюдения.

Степень стабильности может варьировать от проекта к проекту, в зависимости от экспрессируемого продукта гена. Однако, численность клеток с характеристиками стабильной экспрессии может быть существенно повышена в популяции успешно

трансдуцированных клеток-хозяев. Таким образом, можно существенно снизить риск того, что для терапии или другого промышленного применения будет выбран нестабильный клон, который постепенно теряет продуктивность в течение длительного культивирования.

В частности, предложена стабильная линия рекомбинантных клеток, что считается большим преимуществом при применении в продуцировании в промышленном масштабе.

Стабильную экспрессию можно осуществлять после трансдукции, например, путем редактирования генома или других подходов для направленного встраивания трансгена, известных в области техники, тем самым интегрируя экспрессионную конструкцию в геном клетки-хозяина, когда направленное встраивание осуществляется в заранее заданной области хромосомы, которая считается «тихой гаванью», такой как межгенная область, как описано далее. Для создания клеток-хозяев с высокой экспрессией или клонов для экспрессии GOI и получения генного продукта, или POI *in vivo* или *in vitro*, например, в промышленном масштабе, стабильная трансдукция предпочтительна по сравнению с транзиторной. Популяция стабильно экспрессирующих клеток включает относительно высокую долю клонов с высоким и стабильным продуцированием. Стабильность, в частности, повышается при включении GOI в локус «безопасной гавани», такой как межгенная область, описанная в данном документе, что позволяет снизить или избежать транскрипционный сайленсинг генов.

В контексте настоящего изобретения стабильная экспрессия гетерологичного GOI достигается благодаря обеспечению встраивания соответствующей экспрессионной конструкции в межгенную область между двумя соседними незаменимыми генами, как описано в данном документе. Такую межгенную область понимают как область «горячей точки», что означает, что она поддерживает высокий уровень транскрипции внедренных генов, и что такая транскрипция стабильна с течением времени и воспроизводится для различных генов и различных условий культивирования.

Термин «незаменимый ген» в настоящем документе понимают как охватывающий такие гены клетки-хозяина, которые кодируют незаменимый полипептид и, предпочтительно, это ген, для которого не была показана несущественность в клетке-хозяине. В одном воплощении незаменимый ген представляет собой ген, дефицит которого делает клетку-хозяина нежизнеспособной в

определенных условиях культивирования по причине низкой экспрессии или отсутствия экспрессии незаменимого полипептида в клетке-хозяине. Незаменимый полипептид может представлять собой полипептид, способный продуцировать питательное вещество, необходимое для жизнеспособности клетки (например, когда незаменимый полипептид представляет собой фермент, способный продуцировать питательное вещество, необходимое для жизнеспособности клетки) или полипептид, задействованный в продуцировании питательного вещества, которое является необходимым для жизнеспособности клетки (например, когда незаменимый полипептид представляет собой фермент, задействованный в метаболическом пути, приводящем к продуцированию питательного вещества, необходимого для жизнеспособности клетки). При этом, незаменимый ген, кодирующий незаменимый полипептид, может представлять собой ген, дефицит которого делает клетку-хозяина нежизнеспособной во всех условиях и в питательной среде любого типа. Незаменимый полипептид может представлять собой полипептид, у которого снижение или отсутствие экспрессии в клетке-хозяине делает клетку-хозяина нежизнеспособной во всех условиях и в питательной среде любого типа, такой как минимальная или комплексная среда. Предпочтительно, незаменимый ген, кодирующий незаменимый полипептид, представляет собой незаменимый ген эукариот, кодирующий полипептид, необходимый для эукариот. Подходящие примеры классов незаменимых генов включают, без ограничения, гены, задействованные в синтезе и модификации ДНК, синтезе и модификации РНК, синтезе и модификации белков, функционировании протеасомы, секреторном пути, биогенезе клеточной стенки и клеточном делении. Хорошо известно, что многие гены, кодирующие соединения, задействованные в первичном метаболизме и метаболических путях, могут представлять собой незаменимые гены, дефицит которых делает клетку-хозяина нежизнеспособной.

Термин «шаблонная клетка-хозяин» означает клетку-хозяина, полученную путем генной инженерии, чтобы содержать искусственную или гетерологичную нуклеотидную последовательность для сайт-направленного встраивания GOI, такую, как описано в данном документе, и готовую к встраиванию GOI. Шаблонная клетка также считается «пустой» клеткой-хозяином (то есть рекомбинантной клеткой-хозяином без трансгена), которую можно применять, например, в качестве клеточной линии для клонирования с помощью рекомбинантных технологий. Соответствующая

клеточная линия может быть трансфицирована гетерологичным GOI, например, с помощью соответствующего экспрессирующего вектора.

Таким образом, шаблонная клеточная линия представляет собой линию клеток-хозяев, которая характеризуется своей способностью экспрессировать любой требуемый GOI. Это соответствует инновационной стратегии «шаблонов» для создания продуцирующих клеточных линий, например, с помощью замены кассет, опосредованной сайт-специфической рекомбиназой, или гомологичной рекомбинации. Такая новая линия клеток облегчает клонирование GOI за несколько дней, например, в заранее заданные «горячие точки» геномной экспрессии, для получения воспроизводимых высокоэффективных продуцирующих клеточных линий.

Таким образом, в данном изобретении предложено новое решение, позволяющее преодолеть сайленсинг трансгенов, благодаря выбору в геноме локусов «безопасной гавани», которые не могут подвергаться сайленсингу. Сайленсинг представляет собой явление, которое реализуется посредством установления репрессивных меток на ДНК или гистоны (например, триметилирование гистона H3 по лизину 9 или 27). В данном документе описаны целевые области, каждая из которых окружена двумя незаменимыми генами. Предпочтительно, два незаменимых гена ориентированы таким образом, что их промоторы расположены вплотную к сайту «безопасной гавани» и обращены в противоположные направления. И предпочтительно, указанные два промотора расположены очень близко друг к другу, что гарантирует, что сайленсинг трансгена, если он произойдет, осуществит сайленсинг расположенных по соседству незаменимых генов. Следовательно, сайленсинг трансгена приведет к клеточной гибели, тем самым обеспечивая механизм, благодаря которому можно отбирать типы клеток и состояния клеток, в которых сайленсинг трансгена не происходит.

Предпочтительно, незаменимые гены являются незаменимыми для множества типов клеток и поэтому относятся к «набору ключевых незаменимых генов», относящегося к генам, которые (1) являются в высшей степени незаменимыми и поэтому абсолютно необходимыми и которые (2) являются незаменимыми в каждой клетке высшего организма, такого как человеческое тело, что обеспечивает универсальность стратегии. Подходящие пары незаменимых генов в идеале расположены вблизи друг друга (предпочтительно, со своими промоторными областями), так, чтобы позволить внедрение трансгена между ними. В частности, в

качестве примера приведено около 20 подходящих локусов, фланкированных двумя в высшей степени незаменимыми генами.

Трансген может быть встроен в описанный локус «безопасной гавани» несколькими способами. Эти способы включают применение нуклеаз с цинковыми пальцами, где массив белков с цинковыми пальцами опосредует связывание со специфической геномной последовательностью, а нуклеаза FokI вызывает двуцепочечный разрыв ДНК вблизи сайта присоединения цинкового пальца. В качестве альтернативы в равной мере подходит нуклеаза на основе эффектора, подобного активатору транскрипции (TALEN), состоящая из массива эффекторов растительного происхождения, подобных активатору транскрипции, и нуклеазы FokI. Кроме того, могут применяться ДНК- или РНК-направляемые нуклеазы. Они включают популярные ферменты CRISPR Cas9, Cas12a/ Cpf1 или Cas12b, CasX, а также более сложные системы CRISPR, такие как комплекс Cascade. Сюда также могут входить белки-аргонавты бактерий, которые используют для направленного внедрения двуцепочечных разрывов ДНК. Трансген также может быть доставлен комбинацией программируемой эндонуклеазы с полученным из AAV донором ДНК, которую часто выбирают для повышения скорости гомологичной рекомбинации.

Экспрессия трансгенов в локусах «безопасной гавани», в частности, важна при создании человеческих или других клеточных линий (например, CHO), экспрессирующих высокие уровни генных продуктов, таких как антитела или цитокины, применяющихся в качестве продуцирующих клеточных линий.

Она также может применяться для генной и клеточной терапии. Например, она может применяться для внедрения химерного антигенного рецептора (CAR) в Т-клетку направленным образом для обеспечения того, что CAR будет стабильно экспрессироваться и не подвергаться сайленсингу в ходе размножения Т-клеток или реперфузии. Аналогично, она может применяться в терапии с использованием Т-клеточного рецептора (TCR), где эндогенный TCR заменяют экзогенным TCR для перепрограммирования специфичности Т-клетки и модификации ее таким образом, чтобы она распознавала новую клетку-мишень (например, опухолевую клетку).

Описанный здесь подход также можно применять для создания стабильной линии индуцированных плюрипотентных стволовых клеток (iPS), экспрессирующих трансген на высоком уровне, где дифференцировка клеток iPS не влияет на экспрессию трансгена. Последнее остается важной проблемой в биологии стволовых клеток.

Экспрессию трансгена можно отслеживать на уровне мРНК или белка различными способами. Уровни мРНК можно определять количественно при помощи количественной RT-PCR или секвенирования РНК. Уровни белка можно оценивать при помощи Вестерн-блоттинга или ELISA. На уровне одиночных клеток экспрессию белка можно определять при помощи проточной цитометрии. Она отображает не только уровень белка в отдельных клетках, но также отображает долю экспрессирующих трансген клеток в популяции. Когда трансген хорошо экспрессируется, его экспрессию можно отслеживать в одиночных клетках при помощи проточной цитометрии, таким образом отображая, какая часть общей популяции экспрессирует этот трансген. На уровне ДНК долю клеток, несущих трансген, можно оценить при помощи цифровой капельной ПЦР. В качестве альтернативы, можно рассортировывать одиночные клетки по отдельным лункам 96-луночного планшета и определять долю положительных по трансгену клеток, используя подход на основе ПЦР.

Проблему сайленсинга генов можно решать двумя способами:

Можно использовать человеческие iPS и экспрессионную кассету, внедренную в кандидатный локус «безопасной гавани» с помощью CRISPR/Cas9. Экспрессионная кассета содержит сильный промотор (например, EF1A), контролирующий eGFP или mCherry, таким образом обеспечивая сильную экспрессию трансгена. После выделения стабильного клона, в котором встраивание трансгена верифицировано при помощи ПЦР, iPS клетки дифференцируются в разные клоны (например, в нейроны или кардиомиоциты). На протяжении процесса дифференцировки отслеживают экспрессию трансгена при помощи FACS, чтобы удостовериться, что экспрессия стабильно поддерживается и не подвергается сайленсингу.

Для внедрения экспрессионной кассеты в кандидатный локус «безопасной гавани» с помощью CRISPR/Cas9 используют линию клеток человека (например, HEK293, Jurkat, U937, Hela или человеческие iPS клетки). Экспрессионная кассета содержит eGFP под контролем промотора (например, SFFV, CMV или EF1A), фланкированного последовательностями, обеспечивающими устойчивость к тетрациклину (tetO). В отсутствие TetR-KRAB, eGFP экспрессируется в клетках-реципиентах на высоком уровне и можно получить стабильную популяцию или клональную линию клеток при помощи FACS-сортирования. Внедрение TetR-KRAB приводит к сайленсингу eGFP. Однако указанный сайленсинг будет распространяться на соседние промоторы, контролируемые два незаменимых гена, что приведет к

гибели клеток, в которых будет иметь место эффективный сайленсинг. Следовательно, можно ожидать, что клетки, утратившие экспрессию eGFP, погибнут. Выживание значительной доли отрицательных по eGFP клеток будет указывать на то, что вопреки таким ожиданиям, сайленсинг данного локуса может действительно иметь место.

Нижеследующие примеры позволяют лучше понять вышеизложенное описание. Однако, эти примеры являются лишь репрезентативными примерами способов реализации одного или более воплощений настоящего изобретения, и их не следует считать ограничивающими объем изобретения.

ПРИМЕРЫ

Пример 1. CRISPR-скрининг для точного картирования двойных локусов незаменимых генов

Встраивание трансгена между двумя незаменимыми генами может быть пагубным для клетки-реципиента, поскольку может изменить экспрессию по меньшей мере одного из этих двух незаменимых генов. Следовательно, оценивают, какую гидовую РНК можно использовать для встраивания трансгена без ухудшения приспособляемости клетки. Для этого, всю область между сайтами начала транскрипции каждой пары незаменимых генов покрывают всеми возможными гидовыми РНК (Фиг. 2) и оценивают фенотип их приспособляемости по выбыванию в результате скрининга. В качестве эталона включают гидовые РНК, нацеленные на кодирующие области рассматриваемых незаменимых генов, тем самым подтверждая, что эти гены действительно являются незаменимыми. Также включают отрицательные контроли, то есть гидовые РНК, нацеленные на несущественные гены или межгенные области.

Пример 2. Нацеливание на локусы «безопасной гавани» и экспериментальное исследование подверженности сайленсингу

Идентифицировали две пары человеческих генов и осуществляли нацеливание на соответствующие межгенные области. Две пары генов представляли собой: MED20/BYSL и FTSJ3/PSMC5.

Первая пара незаменимых генов: первый ген представляет собой FTSJ3 (GeneID: 117246), а второй ген представляет собой PSMC5 (GeneID: 5705), межгенная область: SEQ ID NO:8.

Вторая пара незаменимых генов: первый ген представляет собой MED20 (GeneID: 9477), а второй ген представляет собой BYSL (GeneID: 705), межгенная область: SEQ ID NO:9.

Для каждой из них нацеливание на межгенную область осуществляли с помощью следующих гидовых РНК:

MED20 гидовая РНК 1 (SEQ ID NO:22):

GGGCGUGUCUCGGCACCCUGUUUCAGAGCUAUGCUGGAAACAGCAUAGCAAGUUGAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGCACCGAGUCGGUGCUUUUU

MED20 гидовая РНК 2 (SEQ ID NO:23):

GAGCUCCCGGGUUCGGAGCGUUUCAGAGCUAUGCUGGAAACAGCAUAGCAAGUUGAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGCACCGAGUCGGUGCUUUUU

FTSJ3 гидовая РНК 1 (SEQ ID NO:24):

GGGGCGGCUACUCGAGUUCAGUUUCAGAGCUAUGCUGGAAACAGCAUAGCAAGUUGAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGCACCGAGUCGGUGCUUUUU

FTSJ3 гидовая РНК 2 (SEQ ID NO:25):

GAAUUCGGGUCAUUGGGCGGUUCAGAGCUAUGCUGGAAACAGCAUAGCAAGUUGAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGGCACCGAGUCGGUGCUUUUU

Экспрессионную кассету (Фиг. 3) внедряли в каждый из указанных локусов посредством гомологичной репарации. Экспрессионная кассета несет eGFP, связанный с устойчивостью к пурамицину через пептид 2A, экспрессия которого контролируется промотором SFFV. Промотор SFFV фланкирован массивом из семи сайтов tetO, как указано в соответствующих донорских последовательностях, Фиг. 9:

SEQ ID NO:26, >MED20 Donor1 (сайты TetO показаны жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом).

SEQ ID NO:27: >MED20 Donor 2 (сайты TetO показаны жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом).

SEQ ID NO:28: >FTSJ3 Donor 1 (сайты TetO показаны жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом).

SEQ ID NO:29: >FTSJ3 Donor 2 (сайты TetO показаны жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом).

Клетки HEK293 трансдуцировали экспрессионной плазмидой с Cas9, наряду со следующей комбинацией экспрессионной плазмиды с гидовой РНК и донорской плазмиды:

гидовая РНК 1 MED20 с донором 1 MED20

гидовая РНК 2 MED20 с донором 2 MED20

гидовая РНК 1 FTSJ3 с донором 1 FTSJ3

гидовая РНК 2 FTSJ3 с донором 2 FTSJ3

После экспрессии Cas9 и гидовой РНК донор встраивается в целевой сайт посредством гомологичной репарации. Клетки, несущие соответствующую вставку кассеты с трансгеном, можно отбирать с помощью пуромицина и затем обогащать при помощи FACS сортирования. Наконец, при помощи предельного разведения получали клоны, несущие трансгенную вставку, и генотипировали их при помощи ПЦР.

После получения стабильной клеточной линии (моноклональной или поликлональной), клетки трансфицировали экспрессионной кассетой, несущей конструкцию TetR-KRAB, показанную на Фиг. 8, SEQ ID NO:30.

После внедрения в клетку-реципиента конструкция TetR-KRAB экспрессируется и связывается с массивом сайтов tetO. Связывание TetR-KRAB вызывает сайленсинг промотора SFFV, что приводит к эффективному сайленсингу трансгена eGFP. Однако, сайленсинг распространяется и на примыкающие промоторы, контролируемые два незаменимых гена (MED20/ BYSL или FTSJ3/ PSMC5). Следовательно, клетки, в которых возник сайленсинг eGFP, погибают.

Выжившие клетки оценивают при помощи проточной цитометрии на позитивность по GFP. Доказательством, что новые локусы «безопасной гавани», расположенные между двумя незаменимыми генами, являются функциональными, будет отсутствие выживших клеток, отрицательных по GFP, и сохранение экспрессии GFP у всех выживших клеток.

Пример 3. Оценка сайленсинга трансгена в человеческих iPS клетках

Использовали линии человеческих iPS клеток, такие как BOB-C или KOLF2, которые получали в консорциуме HipSci. Репортер GFP, находящийся под контролем подходящего конститутивного промотора (например, CMV, SFFV, EF1As, EF1A) внедряли в представляющий интерес локус «безопасной гавани». Две приведенных в качестве иллюстрации комбинации гидовой РНК и донорского вектора представлены в Примере 2. В отличие от Примера 2 донорский вектор не несет никаких сайтов TetO.

После внедрения с помощью CRISPR-Cas9, как показано в Примере 2, получали стабильную популяцию клеток (поликлональных или моноклональных), несущих трансген в желаемом локусе. Осуществляли дифференцировку данной популяции в различные линии с помощью протоколов, известных в области техники (Volpato V et al.

Stem Cell Reports. 2018;11(4):897–911; Mummery CL et al. Circ Res. 2012;111(3):344–358). Оценивали экспрессию трансгена перед дифференцировкой и на разных стадиях дифференцировки (ранней, поздней, финальной). В зависимости от типа клетки, дифференцировка занимала от 15 до 120 дней, и за дифференцировкой наблюдали при помощи подходящих маркеров, известных в области техники. Доказательством отсутствия сайленсинга служила стабильная экспрессия трансгена (например, по данным проточной цитометрии) на протяжении всего эксперимента. Клоны, в которых происходил сайленсинг трансгена, погибали, поскольку сайленсинг неизбежно распространяется по меньшей мере на один из соседних незаменимых генов (Фиг. 1).

Пример 4. CRISPR-скрининг

Незаменимые гены представляют собой такие гены, разрушение которых приводит к потере приспособляемости или жизнеспособности. В данном примере оценивали, какие из 21 пары генов, перечисленных в Таблице 1, являются незаменимыми. Кроме того, осуществляли нацеливание на область между двумя соответствующими незаменимыми генами. Если нацеливание CRISPR/ Cas9 на указанные межгенные области осуществимо, можно ожидать, что гидовые РНК будут вести себя нейтрально, то есть клетки, несущие указанные гидовые РНК, не будут иметь фенотип приспособляемости.

Для проверки этого на десяти парах генов, упомянутых выше, создавали подходящую специализированную библиотеку sgRNA. В этой библиотеке на каждый незаменимый ген было нацелено приблизительно 10 независимых sgRNA. Отбирали sgRNA, нацеленные на кодирующие экзоны соответствующих генов. Кроме того, на каждую межгенную область нацеливали как можно больше sgRNA, то есть отбирали каждую sgRNA, начинающуюся с G и несущую NGG PAM, примыкающий к ее последовательности. В итоге, библиотека имела следующий состав:

Таблица 2

ID	Ген 1	Ген 2	К-во sgRNA, нацеленных на ген 1	К-во sgRNA, нацеленных на ген 2	К-во sgRNA, нацеленных на межгенную область
NO 1	NFS1	ROMO1	10	10	4
NO 2	MED22	RPL7A	10	10	11
NO 3	DDX51	NOC4L	10	10	9
NO 4	CENPK	PPWD1	9	10	13
NO 5	ORC1	PRPF38A	10	10	24
NO 6	POLR3K	SNRNP25	10	10	28

NO 7	COPE	DDX49	10	10	31
NO 8	FTSJ3	PSMC5	10	10	30
NO 9	MED20	BYSL	10	10	49
NO 10	AURKA	CSTF1	10	10	42
NO 11	NUP88	RPAIN	10	10	71
NO 12	ATP6V1D	EIF2S1	10	10	47
NO 13	POLR2I	TBCB	10	10	63
NO 14	UFD1	CDC45	10	10	29
NO 15	CCDC115	IMP4	10	10	71
NO 16	NAA50	ATP6V1A	9	10	71
NO 17	SART3	ISCU	10	10	94
NO 18	C1orf109	CDCA8	10	10	215
NO 19	POLR3A	RPS24	10	4	114
NO 20	RPS16	SUPT5H	9	10	402
NO 21	RPS29	LRR1	3	10	67

Библиотека sgRNA также была нацелена на группу из 10 несущественных генов, на каждый из которых приходилось по 10 sgRNA.

Клонирование sgRNA осуществляли, как известно в области техники. Вкратце, пулированную олигонуклеотидную библиотеку, содержащую 1989 sgRNA, клонировали в реципиентный экспрессирующий вектор sgRNA. Лентивирус получали с использованием клеток Lenti-X (на основе клеток 293T) и инфицировали при низкой множественности заражения клетки RKO, несущие Cas9. Через 2 дня контролировали качество клеток при помощи проточной цитометрии с получением популяции клеток, которая была положительной по GFP на 10,2%. Обогащение клетками, несущими sgRNA, осуществляли путем отбора с помощью пурамицина (в концентрации 0,5 мкг/мл).

Клетки собирали на 3 день или 21 день и выделяли геномную ДНК. Амплифицировали кассету sgRNA в двух раундах ПЦР с использованием следующей пары праймеров:

1 раунд:

PCR1_1_FW, SEQ ID NO:31:

ACACGACGCTCTCCGATCTACATAACGGTGTGGAAAGGACGAAACACCG

(для амплификации gDNA на 3 день)

Или

PCR1_2_FW, SEQ ID NO:32:

ACACGACGCTCTCCGATCTTAGTTACGGTGTGGAAAGGACGAAACACCG

(для амплификации gDNA на 21 день)

PCR1_REV (NEW SEQ ID NO:49):

GGTCTAACCAGAGAGAGCCAG

2 раунд:

PCR2_FW, SEQ ID NO:33:

AATGATACGGCGACCAACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGA
TCT

PCR2_REV, SEQ ID NO:34:

CAAGCAGAAGACGGCATACGAGATGGTCTAACCAGAGAGAGCCAG

Продукты ПЦР подвергали секвенированию нового поколения и проводили количественное определение каждой sgRNA путем подсчета логарифма кратности изменений (день 21/день 3; Фиг. 4). На графике показано сильное истощение клеток, несущих sgRNA, нацеленные на незаменимые гены (выделены). Это позволяет предположить, что Cas9 в исследованных здесь клетках был функционален, и что использование CRISPR/Cas9 приводило к эффективному истощению генов, отмеченных как незаменимые.

Затем для каждого гена/ локуса рассчитывали совокупный показатель/ логарифмированное значение кратности изменений. Совокупный логарифмированный показатель кратности изменений для каждого гена/ локуса рассчитывали по индивидуальным значениям sgRNA с использованием программного обеспечения MAGECK (версия 0.5.9.2).

Нанесение кратности изменений для всех генов/ локусов на график (Фиг. 5), позволило заметить явное истощение незаменимых генов по сравнению с исходной библиотекой, тогда как несущественные гены распределялись среди межгенных локусов произвольным образом. Это позволяет предположить, что большинство отобранных межгенных локусов допускают расщепление Cas9 и последующую репарацию благодаря эндогенному аппарату NHEJ. Это также подразумевает, что многие из отобранных здесь локусов допускают по меньшей мере небольшие вставки или делеции, такие как обусловленные NHEJ. Важно, что одна межгенная область не допускала расщепление CRISPR/Cas9 и примечательно, что данная область SEQ ID NO:1 относится к паре генов, где сайты начала транскрипции двух незаменимых генов расположены очень близко друг к другу (52 пн), оставляя очень мало пространства для вставок/ делеций. И наоборот, это позволяет предположить, что все другие локусы (от SEQ ID NO:2 до SEQ ID NO:21) могут допускать расщепление CRISPR/Cas9. Это

является неожиданным, особенно поскольку данный эксперимент позволил установить, что все отобранные здесь гены явно являются незаменимыми в клетках RKO.

Пример 5. Вставка трансгена в двух выбранных локусах

Для оценки возможности встраивания трансгена между двумя незаменимыми генами использовали два из отобранных локусов: Первый содержал межгенную область между генами FTSJ3 и PSMC5 (SEQ ID NO:8), второй содержал межгенную область между генами MED20 и BYSL (SEQ ID NO:9).

Решили использовать клетки Lenti-X 293T (Takara Bio), поскольку их легко трансфицировать. Вкратце, утром осуществляли посев клеток Lenti-X 293T 6-луночные планшеты по 1,2 миллиона на лунку. После полудня клетки совместно трансфицировали 500 нг плазмиды, содержащей флуоресцентный маркер (SV40-mKate2-pA; полученный путем генного синтеза, фланкированный соответствующими гомологическими участками для интеграции посредством HDR в различные локусы «безопасной гавани» (SEQ ID NO:35 и SEQ ID NO:36)), а также 1500 нг плазмиды, кодирующей SpCas9 и sgRNA, нацеленной на специфическую геномную область (SEQ ID NO:37). Включали донора, нацеленного локус AAVS1 (SEQ ID NO:38), являющийся обычно используемым локусом «безопасной гавани». sgRNA, использованные для нацеливания CRISPR/ Cas9, показаны как SEQ ID NO:39, SEQ ID NO:40 и SEQ ID NO:41 (последовательности представлены на Фиг. 9).

Через 10 дней культивирования и размножения отсортировывали положительные по mKate2 клетки с помощью системы BD FACSAria III и осуществляли посев во флаконы T25. Культивировали и размножали клетки в течение 3 недель, после чего выделяли ДНК из 5 миллионов клеток для каждого условия с помощью набора для очистки геномной ДНК Monarch (#T3010; NEB) согласно инструкциям производителя. Осуществляли генотипирование посредством ПЦР с использованием ДНК-полимеразы Phusion High-Fidelity (#M0530; NEB) с указанными парами праймеров.

Проведенные ПЦР

Название ПЦР	Прямой праймер	Обратный праймер	Размер ампликона (пн)
AAVS1-mKate2	AAVS1-genomic	mKate2-insert1	1539
MED20-mKate2	MED20-genomic1	mKate2-insert1	1422
mKate2-MED20	mKate2-insert2	MED20-genomic2	1481

FTSJ3-mKate2	FTSJ3-genomic1	mKate2-insert1	1310
mKate2-FTSJ3	mKate2-insert2	FTSJ3-genomic2	1276

Праймеры, использованные для генотипирования

>AAVS1-genomic, SEQ ID NO:42: GTCTGGTCTATCTGCCTGGC
 >MED20-genomic1, SEQ ID NO:43: AAACAGACACAAGCGGGTCT
 >MED20-genomic2, SEQ ID NO:44: TCCTGCTGCAATCGGAGAAG
 >FTSJ3-genomic1, SEQ ID NO:45: GTTACGAACCATCCCCCTGG
 >FTSJ3-genomic2, SEQ ID NO:46: ACCCTTCCTAGCTCCCTCTG
 > mKate2-insert1, SEQ ID NO:47: GGTAGCCAGGATGTCTGAAGG
 > mKate2-insert2, SEQ ID NO:48: CCGGCGTCTACTATGTGGAC

ПЦР реакция

Компонент	Количество
вода без нуклеазной активности	20 мкл
5X буфер Phusion HF	4 мкл
дНТФ, 10 мМ	0,4 мкл
прямой праймер, 10 мкМ	1 мкл
обратный праймер, 10 мкМ	1 мкл
геномная ДНК	50 нг
DMSO	0,6 мкл
ДНК-полимераза Phusion	0,2 мкл

Схематическое изображение модифицированных локусов вместе с расположением праймеров, использованных для генотипирования, представлено на Фиг. 6. Для каждого локуса, представляющего интерес (MED20/ BYSL или FTSJ3/ PSMC5), целевую интеграцию подтверждали с помощью двух разных пар праймеров. Один из них нацелен на 5'конец встроенного трансгена, а другой нацелен на 3'конец встроенного трансгена.

На Фиг. 7 показаны продукты ПЦР, полученные в реакциях генотипирования, описанных выше. Все ПЦР проводили на всех линиях клеток, таким образом устанавливая специфичность проведенных ПЦР. ПЦР подтверждали встраивание трансгена SV40-mKate2-pA во все исследованные локусы, включая AAVS1 (как референтную точку), MED20/ BYSL и FTSJ3/ PSMC5. Таким образом, установили, что

нацеливание на описанные здесь локусы технически осуществимо. Важно, что кассета с трансгеном также содержала активные регуляторные элементы генов, такие как промотор и сигнал полиаденилирования, встраивание которых не влияло на экспрессию соседних незаменимых генов. Это позволяет предположить, что похожие кассеты могут быть встроены в указанные здесь локусы без влияния на клеточную приспособленность.

Наконец, экспрессию трансгена mKate2 также оценивали в клетках, в которых трансген был встроен при помощи CRISPR/Cas9 (см. выше). Для этого 10000 клеток анализировали посредством проточной цитометрии на проточном цитометре BD LSR Fortessa™ с настройками по умолчанию для фикоэритрина (PE) для оценки экспрессии mKate2 (Фиг. 8). В клетках, в которых трансген был встроен в локус MED20/ BYSL, экспрессия mKate2 определялась в 6,5% клеток. В клетках, несущих трансген в межгенной области между FTSJ3 и PSMC5, mKate2 определялась даже в 18,5% клеток. Важно, что эти данные регистрировали на 28 день после трансдукции, что исключало возможность того, что экспрессия трансгена обусловлена эписомальной плазмидой.

В целом, данный пример показывает, что встраивание трансгена в выбранные здесь локусы, осуществимо, и что трансген может экспрессироваться стабильно с течением времени и на детектируемых посредством проточной цитометрии уровнях.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Выделенная клетка-хозяин млекопитающего, содержащая гетерологичный ген, представляющий интерес (GOI), интегрированный в хромосому в целевом сайте в пределах межгенной области между парой соседних незаменимых генов, где каждый из соседних незаменимых генов указанной пары имеет сайт начала транскрипции (TSS), и расстояние между TSS составляет менее 20000 нуклеотидов (нт).

2. Клетка-хозяин по п. 1, где расстояние между указанными TSS составляет по меньшей мере 80 нт.

3. Клетка-хозяин по п. 1 или п. 2, где длина указанной межгенной области составляет менее 20000 нт.

4. Клетка-хозяин по любому из пп. 1-3, где указанная межгенная область содержит или состоит из нуклеиновокислотной последовательности, которая по меньшей мере на 90% идентична любой из SEQ ID NO: 1-21.

5. Клетка-хозяин по любому из пп. 1-4, где незаменимые гены указанной пары ориентированы таким образом, что их промоторы обращены в противоположные направления.

6. Клетка-хозяин по любому из пп. 1-5, где указанный GOI содержится в гетерологичной экспрессионной кассете, интегрированной в указанный целевой сайт.

7. Клетка-хозяин по любому из пп. 1-6, которая представляет собой

а) гемопоэтическую стволовую клетку, эмбриональную стволовую клетку, плюрипотентную стволовую клетку, индуцированную плюрипотентную стволовую клетку или эндотелиальную клетку эндотелия; или

б) клетку типа, выбранного из группы, состоящей из клетки-натурального киллера (NK), клетки микроглии, макрофага или Т-клетки, такой как цитотоксический Т-лимфоцит (CTL), регуляторная Т-клетка или хелперная Т-клетка.

8. Клетка по п. 7, представляющая собой иммунную эффекторную клетку, содержащую химерный антигенный рецептор (CAR), где GOI кодирует CAR, экспрессируемый указанной иммунной эффекторной клеткой.

9. Препарат, содержащий популяцию клеток-хозяев млекопитающего, где по меньшей мере 1% указанных клеток являются клетками-хозяевами по любому из пп. 1-8, обладающими хромосомной идентичностью.

10. Культура *in vitro* клеток-хозяев клетки-хозяина по любому из пп. 1-9, которая поддерживается в условиях для экспрессии указанного GOI.

11. Способ получения генного продукта, кодируемого геном, представляющим интерес (GOI), путем трансформации клетки-хозяина для интеграции GOI в хромосому в пределах межгенной области между парой соседних незаменимых генов, и культивирования указанной трансформированной клетки-хозяина в условиях для экспрессии указанного GOI, предпочтительно, где каждый из соседних незаменимых генов указанной пары имеет сайт начала транскрипции (TSS), и расстояние между TSS составляет по меньшей мере 80 нт и менее 20000 нт.

12. Способ модификации клетки млекопитающего путем сайт-направленной интеграции гетерологичного гена, представляющего интерес (GOI), в хромосому в пределах межгенной области между парой соседних незаменимых генов, предпочтительно, где каждый из соседних незаменимых генов указанной пары имеет сайт начала транскрипции (TSS), и расстояние между TSS составляет по меньшей мере 80 нт и менее 20000 нт.

13. Способ по п. 11 или п. 12, в котором для интеграции в хромосому используется любая из программируемых нуклеаз, предпочтительно, нуклеаза с цинковыми пальцами (ZFN), нуклеаза на основе эффектора, подобного активатору транскрипции (TALEN), эндонуклеаза CRISPR или белок Аргонавт.

14. Клетка-хозяин по любому из пп. 1-8 или препарат по п. 9 для применения в медицине.

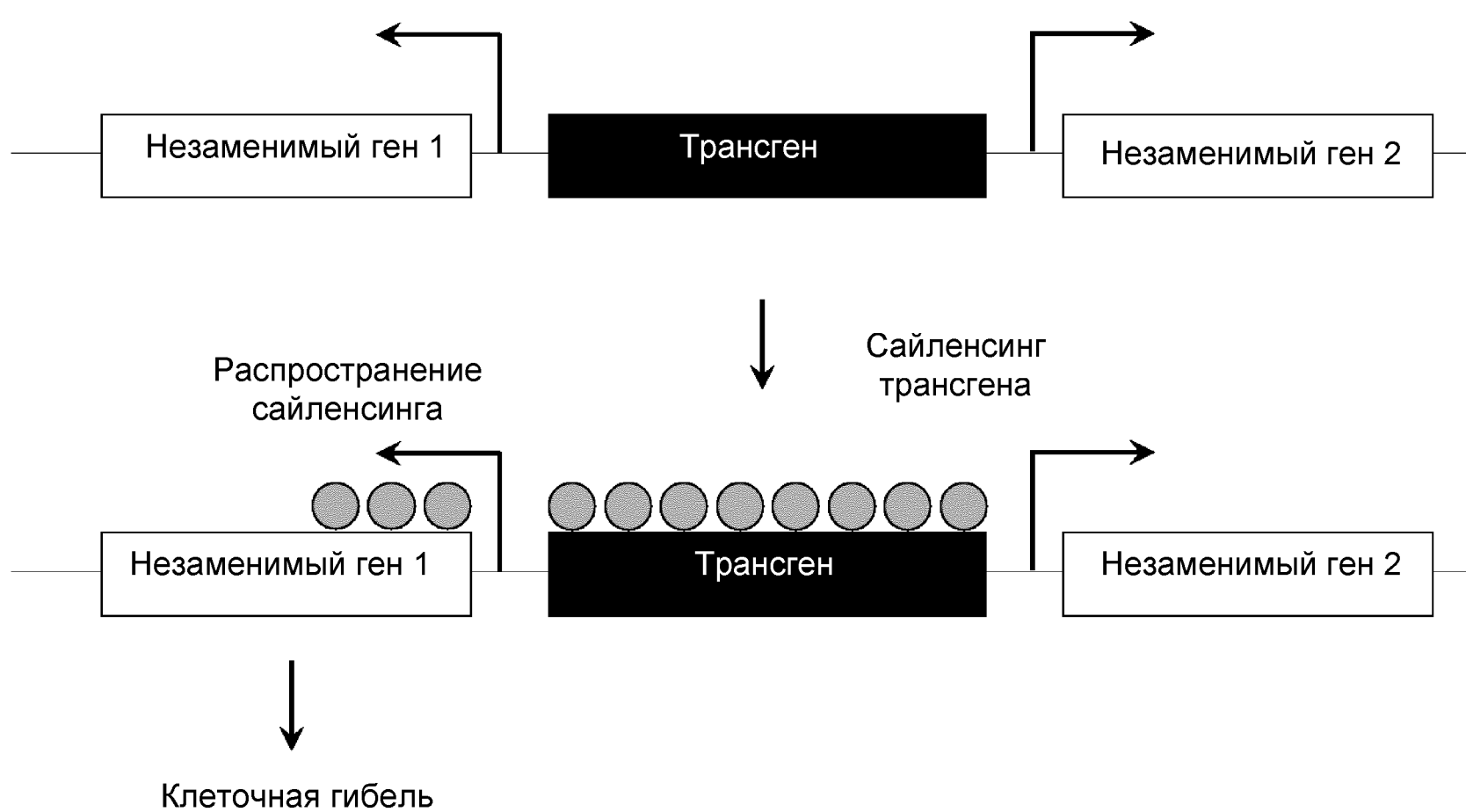
15. Клетка-хозяин по любому из пп. 1-8 или препарат по п. 9 для применения по п. 14, где субъекта лечат посредством экспрессии указанного GOI *in vivo*.

16. Применение *ex vivo* клетки-хозяина по любому из пп. 1-8 в способе получения индуцированных плюрипотентных стволовых клеток, получения антитела и/или получения белка, представляющего интерес.

17. Способ получения микроглии с химерным антигенным рецептором (CAR) или макрофага с химерным антигенным рецептором (CAR), включающий:

(а) получение микроглии или макрофага от субъекта-млекопитающего, предпочтительно, микроглии или макрофага человека, и

(б) модификацию указанной микроглии или указанного макрофага посредством одной или более генетических модификаций для интеграции в хромосому гетерологичного гена, представляющего интерес (GOI), в целевой сайт в пределах межгенной области между парой соседних незаменимых генов, где GOI кодирует химерный антигенный рецептор (CAR).

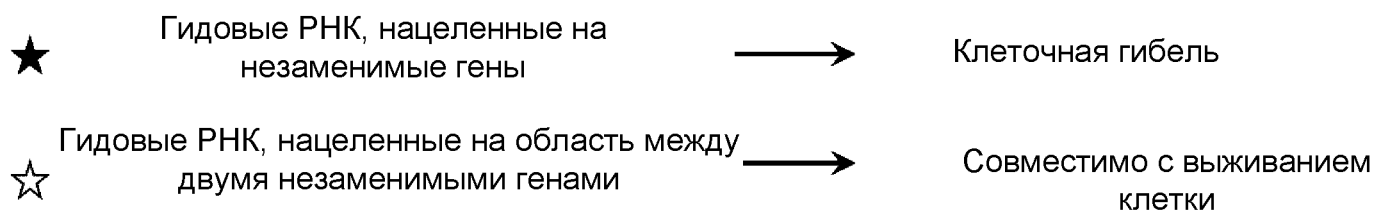


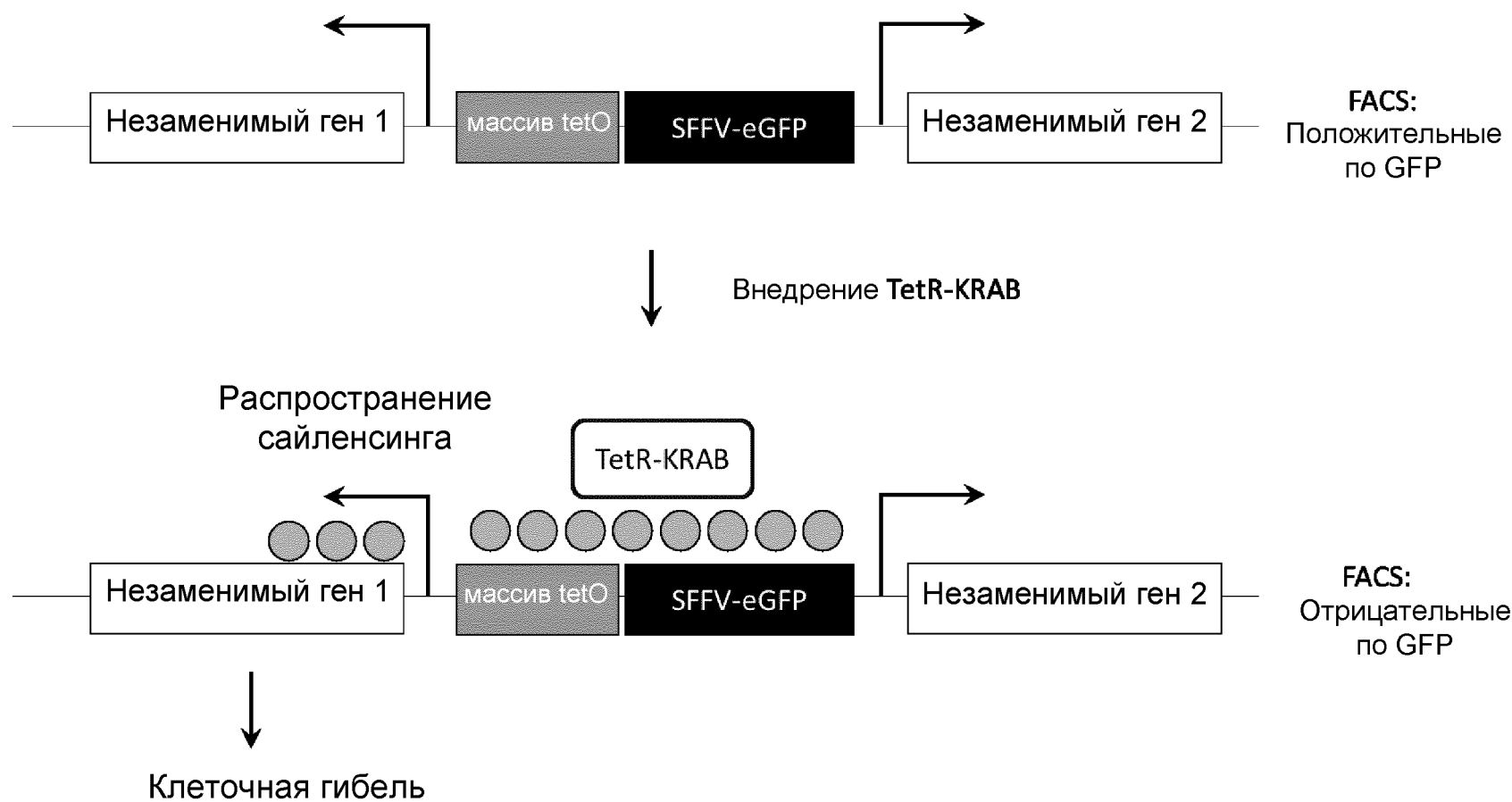
Фиг. 1

Ф



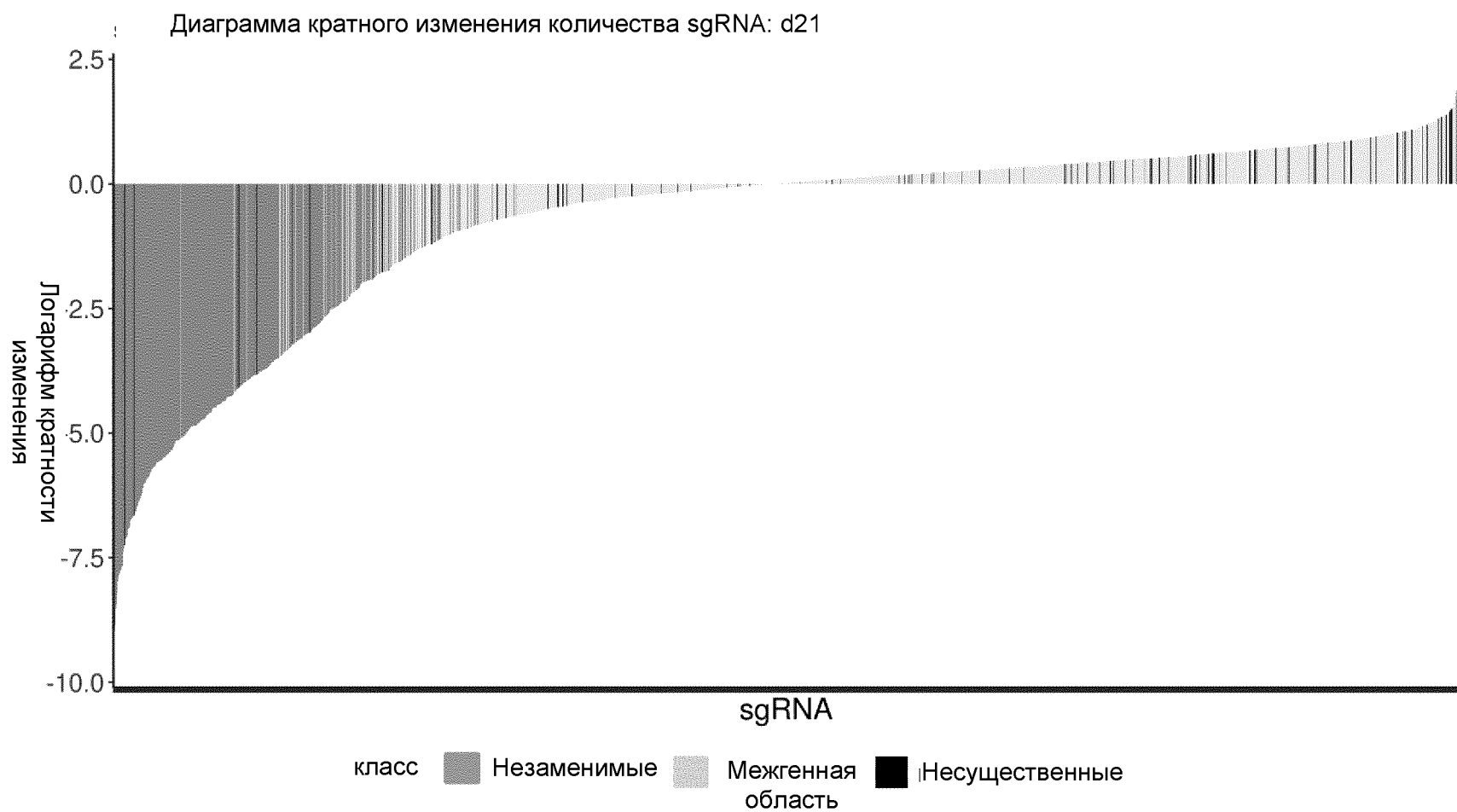
Фиг. 2



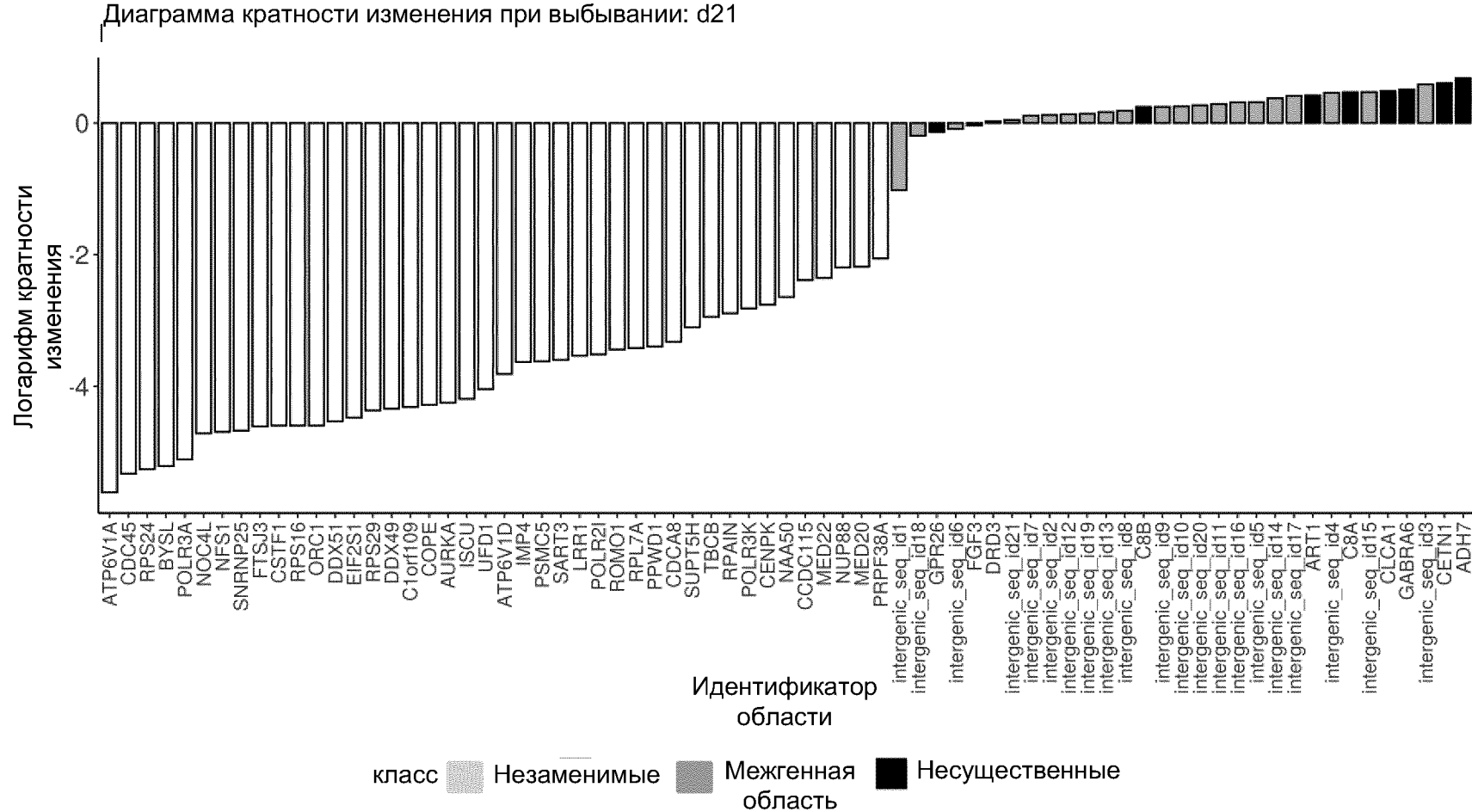


Фиг. 3

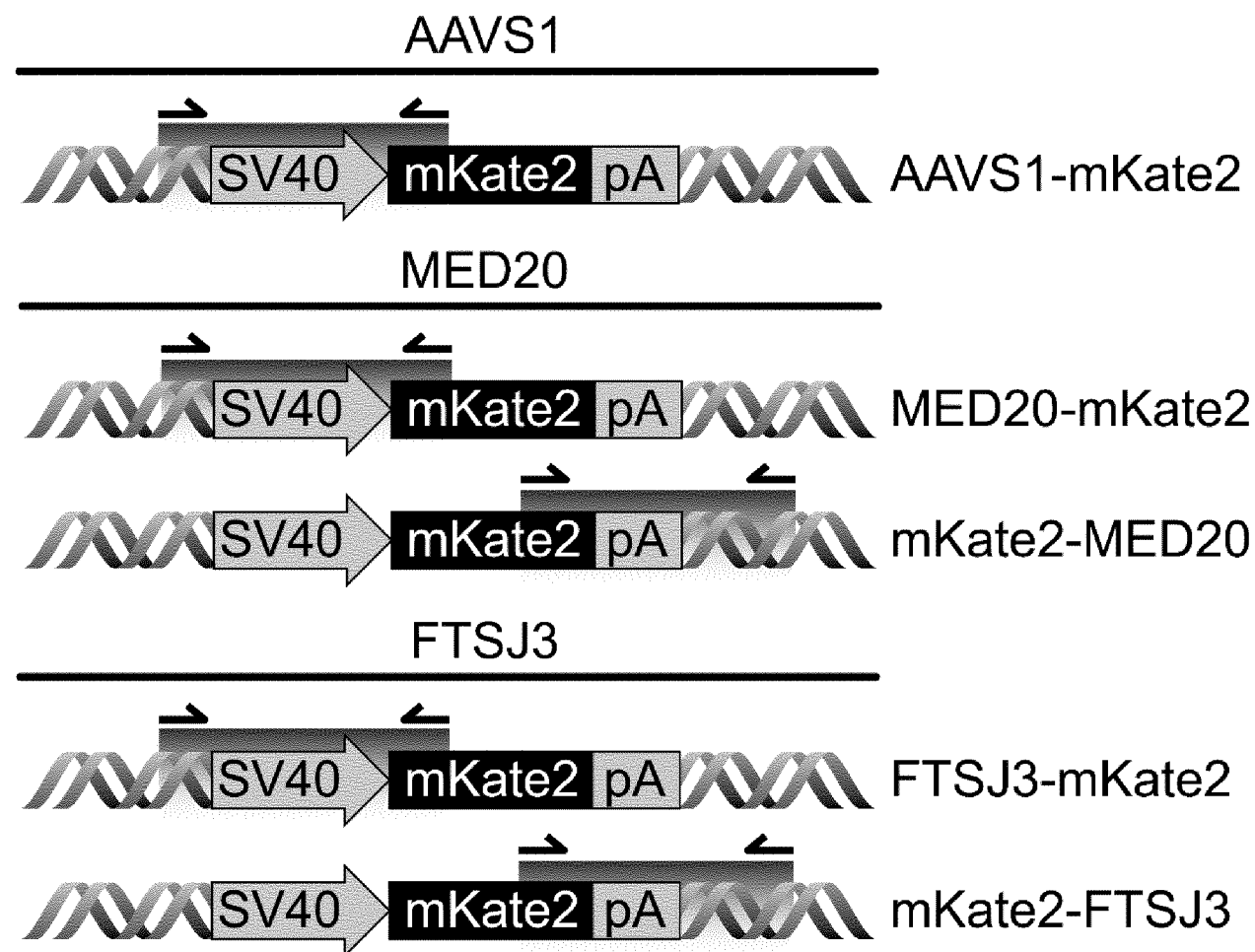
Фиг. 4



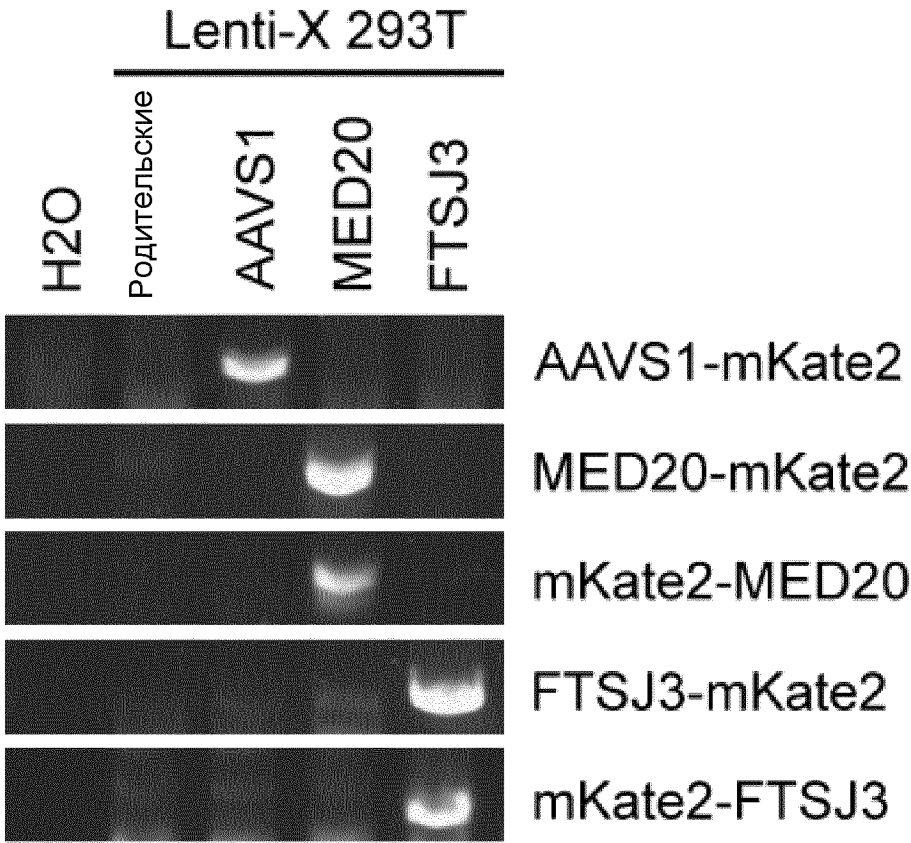
Фиг. 5

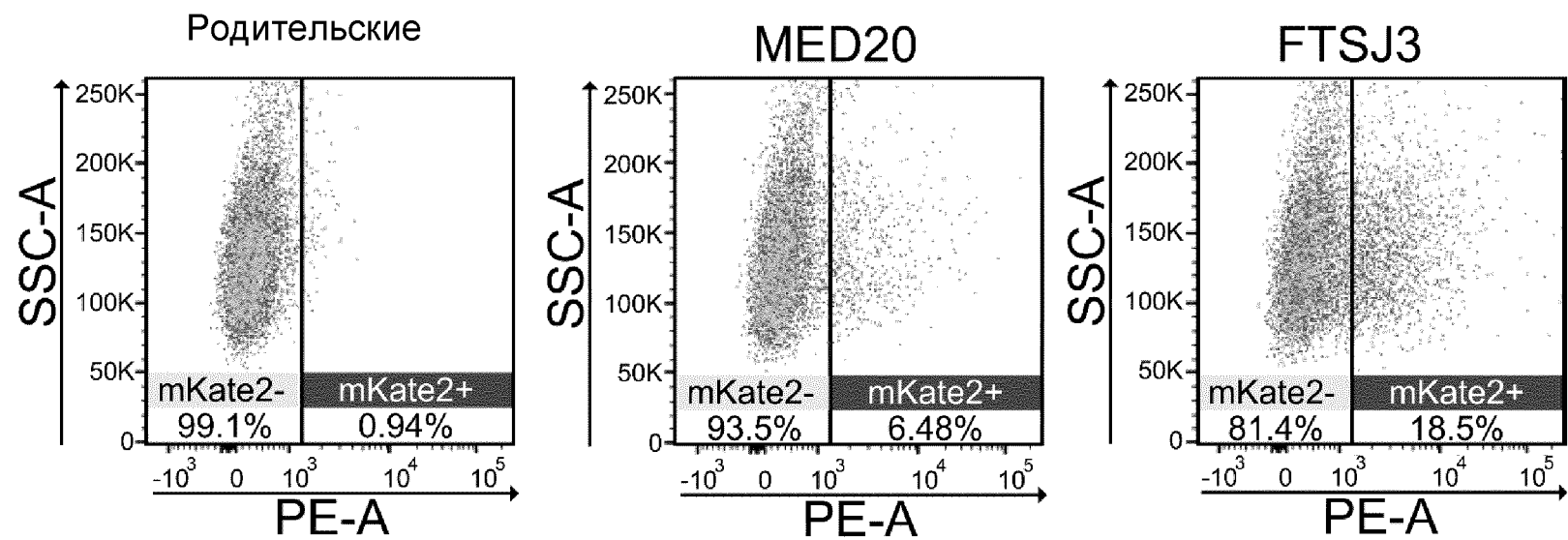


Фиг. 6



Фиг. 7





Фиг. 8

Фиг. 9

SEQ ID NO:1

AAGTGCTGCCCCGGCGCTCCGGAAGCGATTTCATCGGGCCGCGAGCGCCCTCCC

SEQ ID NO:2

CAGCCTCTCGGCCCCGCCTCGATTTTCTAGCTTTATAGGAATGCTGTTGCTTTAAATCCGAAATCCCGTGCCGGTATCAAC
TCTCGCGATCTCCGAGGCCGCATACATATTACCCACAAT

SEQ ID NO:3

TGGCGCGCTGCGATGACGTGCGCGGCACGCCTGCGACTCGGGCTCCGCGCAAAAGATGGGGTTGGGGTACGGCGCGTAGA
GATGACGTGCGGTTCTACGCGCAGTGGTGACGTACGGGAGCGCC

SEQ ID NO:4

CTTGCCAGCCCCGGAAGTCTGCGCGCGCTGCATGCCATTGGATGTGCTCATTCCCACCCAGCCAATCCCGAAACTCGC
TCGGACGCTGACAGAAGACTGCGCCGCTGCTTTGGGATTGGTAGCTGAGTTTGTGTCGCG

SEQ ID NO:5

AGAAGGTGGAAGGAGAAAAGAAAAGTTCGCGCCAATCGGCGTGGCCCCGCCCCAGACCGTCCCTTCGTTGCGACACCAA
CTAGTCTGCCCCGCCCCTCGTACACCTAAGAGGGGCGCATGACGTACATCCGCTAACATGCAGCCGCCATCATTGATT
CGATAGGCGACGCGGGGAGCGGAAGCCCTTTACAC

SEQ ID NO:6

AGCTCCCACTGCCGCTCAGCGCCGATGCGCGCCCGCCTCGAGCTCACATTGGTCCTGGCAGCCTTCCCGCCACACCAAC
CAACCAATAGACAGGGCGATTCTGCGCTCCCGGCCCTGCTGCAGGCTGTCTCGCACTTGTCAATTGGTCACTGCAGCCGCC
CCACCCCCCCCCGGCGGCCAGTGGCTGGGCGGCCCTCGCTGGGGCGGGCCGCAAGTTCCTGCGCGTGCGCGCTTGGCCTCC

SEQ ID NO:7

TCCTGAAAGACACGTACGCCGAAGCAAGACACGGGCACGCTAGGAAATGTAGTTTACTTTTTCCGAGCGGCGGCGAAGT
AAGCCAATGGGAAGCTCCACAATAATATGCAAGGGGCGGAAGGCGTAAGTGCCTCACGGAGAGCATCTCGGGAATTGTA
GTGTGTTTCTGCAAGCCAATGGGATCCGGGGAACATGATAAGGATCGGGTTCGAGGCTGGGAATAGGCTGAGGGAGCTGG
AGAGGGTGGGACCACGCAGGAGGCAGATCCAATAAACAGGAAGATTTTCTCTGTCGTCGTCGCGCGCGCCGGAAGCG
CGGATCACACGGGCCCTACA

SEQ ID NO:8

GTATGCAGCTAACTACTTCCGCTTCCGCTTGCCTCCCTGCGTCCCTGGCTCGAGGTTTCCGCCATTTTGAGTGTGGCC
CTAGATTGTTCTCAATACTCTGTCTTGGGTTTTGTAAAGTTGAAAGTTGAGACTAGGAAGGCATATCACAAGAACTATAA
ACAGAGTTTCAATGAGGCAACACTGAGGAGGAGTTCTACTCCTTCTCATCCCCCTCCAAAGCCCCGAACTCGAGTAGC
CGCCCCCTCCCATCATGGTTCCCTTAGTGTTGGTCTCGCCGCACACCCCGCCATTGACCCGGAATTCTTCTGCTGCTGGT
TTCAGGTCCCAATCCTCCGCTTC

SEQ ID NO:9

TCCGGAAATACGTAAAAACAGAGCGCCGGGTCCACGCCTTCCCCAAGGAATGGGGCGTGTCTCGGCACCCCTCGGTGACG
CCTCCACTGACATCATCTGCGGATTCCCGCCCCGAGGTTTCTAAATCCAGACATTCCCGTTTGGCTGAGCACTCTAGGCC
TACATCCATGAAGCTAGGAGAGCGACACTCAAAGACTGCACTATTGAGAGAAGCTAACGTTAAAGGCAGTGAATATATTC
GGGAGTCCAGCTCCGGAACCCGGGAGCTCTTTTAGTGGGAGGGGCGCGCTGATGGCGCTTCTGGCCTCCGAATGCTAGG
GGGCGCTGTGATCGCCGGGCGGCTCTTGGGCGCTGGGAGTCCACCGCGCA

Фиг. 9 (продолж)

SEQ ID NO:10

CGTTAGAATTCAAAGGTTCTTCTTTTAAATAGGAGTTAAACCCCTAGCTAGAAAGCCGATTGGCCTGCGCTCGAGCGA
GGCTGCGCTTGTGATTGGTTGACAACAAACCCGACGCCTTGAGTCGCGTTTAAGGGCGCACACCGCGCGCAGGCGCACTTG
CTCCCTAAGAACCCGGAAGTGGTGGAGCCAACGGCCAGCCAGACCTGTGGCAGTCCCAGCCCCAGGGTGCTCTCCGCGGT
CGGCCATTTATCGGTGAAGCAACAGCCCCCTTGACGTCAAGCTTGACGCATTGGAGATAAAGTCCAAGGAGGTGAACATAAC
CTTTGAACGTGCCAGTAAGTAAGATGGCGAGAAAAGCAAGAGAGAGTGGGACCGATCGATAGCGCAGCGGTTCGGCTTGGC
GCCCTTTCAGCGTGCGCAGTGAACGTGCGCTCGGAGCGGTA

SEQ ID NO:11

GCTCCCCCTCACTCCAGTTGCTCAGCCCGCGCTGAGCACAGCCAATCCGCGGCCGTCCACTCACTGTTTCGGCCAATCAC
CGAGCAGACGGCCCCGAAGCTCGAGAGGCTCGTGCCAGGTGCGGGGGCGAAGGCGAGGCCAGCCACCAAGCAGGGGCGT
CACCAAGGGAAAGCCAATCACCATTATTTCTGAAGCCCAAGGGGGCGGGACGAGAAGTCAACATAATGGTGGTGCTGCTG
GAAATCTAGCACCTGGGGAGAGCCTGAGACTTGTGTATCTAATTGCCTATTGGACACCTCCATCTGGGTAAACCCAGGCA
TCCGAACCTTTCCCTCAAGACTTACCTCCCTCAGTGCAGGTGCGGCTGGCCTCACGCATTTTCGCTTGGATGCCTAAATCT
GTCTCCACGTGGGGCGGTCGATAAAACAACACTTCCAGAATACGTGCGATCCATTGCGGGTCTGCACCCGAAAGTA
CACCTCCAGGGGTCTGTGCGCGCCTCCAGGGAACGGGTCTTGTGGC

SEQ ID NO:12

GTCAGCTGGTTGTGTCACTTGACCCCTCTTGTTGCTAAGAGGCAGCAGGGACACCCATTTTCACTTCCGCTTCCGTTGT
CACAGGCGACGAAGGGGAGGACAGAACAGAGGCGTTGCCCGGATGTCGAGGCGCGCGCACCGCGGAACCAACCCCTCAAT
TCCGGTTTCGCTCCCAACCCCTCCCATCCAGCTTCCCTGCGGCTGCGCAATACGCGGCACGCGCCAGGCCAAAGACTAG
GAGCGAAGCGTGACGCGCTTGTCTTTCTCCACCTTTCTGCTTTTGGTCACAGCGCCATCTCATGGTTAGCTGAAACTTTGA
AAGAGGATTCATTTCTCAGAGGCAAGGAATCGGAGTCTATGTCTTTCTCGCGCCTTGAAGGTTACTTAGCTAAAGCTAA
GATAGCGAGAAACAGTACTTTTCAGGTCCTTCCATGATCTACATTCGTCCTTGTTCCTCGGAGGACCCAGACTCTATGGTT
AATGATGAAGAGGGGCTCTATTGTTTCAGCGCTTGCGCATGCGAGGAGGTTCCGCATGCGCGGTG

SEQ ID NO:13

CCGCGCTTGCTCCGCCGCGCTACAGGTCGCGGGAGGGGCCAGGCCTAGGTGCTTCGCCAAAAGAGGGCGGAGCTGCGCAC
ACGTGCTGGCTCCCGCCGCGCTCGGGCTCGGCTCTCTCCGCGCAGAGGTTTCATGGCTTGCTTCAATCTTTTGATATCTT
AGAGTAATTGCTTCCAATTTTGGGCATCGCCGCTCAGTTTACAGAATAGTATGGATCTGGTGCCAGAGCCGGGAAAAC
GATGCGCATCGTAAACCGCAAAACACTGAGCATCCAGGACAGAAGCCAAGGCGCTTGCTTGGTTGAAGGCCAGTGCGCA
TGCGCGAGTTTACCTTTCGCGGACTGGGCGAGCTTGGCACTCCTACCTCCGGGCTTCAGTCTTCGCCGCTGTTCTGG
TCCTTTCCCGGGCCGCTATTGGAAGAACCGTGAACATGGAAGGCCCTCTGGATTGGCTGGGCTGGCCCAAGAGTAAGG
AGAGCCCTCTTCTTGGCGGTGGGGAAGGGACGGCGGCGATTGGCGGACGCTCCTGGCAGGAGAGCGCGGATTGGTCAGGC
ACGGAGCAGGAGGCGGGGCTGATAGCCAGCAGCAGCAGCGGCGGCGCGG

SEQ ID NO:14

AGCCGGTACGCCCCAGAGGCTCACCGGAAGTGCCGGACCCGAGACCCGAGGGGGGGTGGCCGGACAAAGCGTCGGCTG
CAAAAGAGAAAAGGCCAGGCGAAGACGCTAGGCTCTAAAACACCCCTCAGTAGAAGCATTTAAATGCCCTTTTCGTGATTTA
TTTCCCTTTTATAGGTATTTTTAAACGTTATTTTAAATTACGTGTTGACAGTATTTCCCTCCAGACGAATTTAGTAGGG
CCCATTTGTCCTTACGACTACACTACCCACAATGCATACCCAGACTGCCTCTGCTTCAGCCATCGAGGACTCGGGCGGAA
CTAAGCTACAAATCCTTCTTCAAACCACACCAAAAACCTATTGCAACGTAAATGCTGCCCCATTGGGTATGTGACTGAG
TTTAATGTATGGTTAGAAGTATACTTTCCCTGAGGGTCCCAGTGGGAGGCAGTAGGCTTCTCTTCCCAAAGTAAAGATG
TCCGCCCAGCGTCGCTTTTCTGGGTAAAAGCGAGTCAGAGTTGAAGGGGGCAACAGTGTTTGCCTTCGCCCTCAGAGGTG
ACGCTTCTTTGGGGCGGGCTGGAGTTTCTGGCCGTGAATGGCAGAGCGCTAATGGCCGCTCCAATGTGGCCCAATCA
GAAAGAAAGGAAGGCTGGGAACTAAGAAGCTATTGGTTGGTGATCCCTGGACCAATCGGAGGAGCCGTGATTTGGCGGGA
GTCTTGACCGCCGCGG

Фиг. 9 (продолж)

SEQ ID NO:15

TTCTTCCTTGTACCCCTCCGCAGTTTGGCAGGACAGCCAATAGGCACACGAGATCCTCCATCAAGGGGCGTTCCCATCTG
 GGGATCCCCACGCCGGCAGCACCAAGGCCTCTGGGTGAGCTTCTTTAAGGGCAAGCGCCCTTAGAGAAGCAGCACGGAAG
 CAACCAATCAACTGGTTTCTCCAATTAGCCCCCAAGTGGATCAGCCTGCTGAAGTGTGCGCTAGGGTTATAGGTGGTTCT
 GAGTCCTTCCGGAAGCTGCCGAATCTTAGAAGTCGCGCCACTGTTGTGGCGATGGAAGAAGGCGAAGCTCTACAGAAACC
 ATTTTGGTTGAGGGAAGAGTGGTTTAAAGTTAATTTGGAAAGGAGCCGGCAAGCAACCTTCAGTCGGTGTGCGGTTTTT
 ACAGTTGAACTGTTTCCAGCCGCTAATTATGATTGTTTTCAGCGGTCGTCATCTGAAACTGCATTCTGCACGGCGGTGG
 TGGGGGAGGACGAGGATGGAGAAGCCTGGGCGGACATATTTATTAGCTGCGGTCCTAAGTCACGTGTGACATTCTGTAT
 CCAGTCTATCGGTTATCCCAAGGAGAACGCCCTGCCGAATGGCCAATCGCCTTGAGATGTCCCGCCCTATTCGCCATC
 TTAACATAGGGCATTCACACAGCTGCAGGCACAGTGAAGCGTCTAAGGGACCATCTTGGTTGAGGGCGGAAGTGAGGA
 CTGTGTCGCGCTGGAGG

SEQ ID NO:16

GCTGTGACCTTTACCCCCGCCCTCGCGCGCCGGCGCGCCGCCGCTCACGCGCAGCCGCACTGCTCGCTCCGCGCACC
 CTGCCCCCTCCACTCCGGAACCTCCATCTTTCTCTGGACTGGGCCCTTCTGGGGTCATGCTCCTCCCTCCACCCGCTTA
 CGACACTGCTTGTGGGCTGGCTCGGTAGGCCCTGGCGGTAATAATCCAGATTTTCCACCCCTCCCTCTCTTGTCCATA
 TCCGCAAGATAGTTTTTCTCAGCCCCGTCTTCCCAGTTTGTGGTCTCCAGAGTGACCTTCAAACGCTGTTGCCTCCTAG
 TAGGGTTTTATAAAGTTAGTTCTTCTAAATGTAAAGAAATCTGTGGAAGCTCAAGTTTAAAGTTAGGTGTTTTAAACAA
 ACGTAGTTAACAATCGGCCAGATAGATTGGCTACCTCCAGCACGCCCTCAAGTTTTAAGATCACGTGATTGGCCCAAACC
 TCAGAGCCTGGAATTTTACCGGCGGACTCCAGAGGGACGGAACGCCACCGCATTTCTGTCCTCGATTGTGGCGTGAA
 CGTCACTTCTTAGCAATGCGCTTTAACTCCTTGCTGCCACATTTTCCACATTGTACTTTATAACATACTCTTTGTGAAA
 ATTTATCTTTTATTTCTTTAATCCAAGAAAATTAATATTCAGGGATCCGTAAAAATTCAAGGCCCCCGAGATGTAGAGC
 GCTGGGATGGAGGAGGAGGCGGAGCTCGACGTCTGTGACTGGTTGGGCCCGCCCCCCCCCCCCCCCCAACCTGTGC
 AGCTGAGGACTTC

SEQ ID NO:17

ACTCTCGGGTCTTCCCGCGGCCGCCGAAGACCCGCTGAGGCAGAAGCTGCCCGCGGCGCGCCTTATCAAAAACCCGTT
 TTCTTGCCATTGAAAGAGAAACTGGGAACCTAGATAACAGGAAAATCTCCACATACACTAAATTCACATTATTTTATACC
 CTGTAAGGCTCCCAACAGATATAAACTCATTATAATGGCATGCGTTCCGTCACAACTTCGGACAGAAGTTGGGTCTT
 TGGGCAGGAAGTGGAAGCCATAGAGGAAGTCATGGAGCGAGAGTGATGTCTGTAGTGAAGAATTTGGCGCCATCTTGAGA
 AGGTCAAAGTGATTCTTCGATTAAATGCCTTACTTATCGCCTAGTGTGACCTGGCGCTTGGGTTTACTCTCCCCAGCAAC
 TGGCCTCAAACGCAAAGAAACTGCAGGTCTCGGGGAAGGATCGATTTACCATTCAGAGAAAAATCCATTTCCAG
 ATTCCTCCCTTTTATGTAACGCTCTGGTACCTAGTAGACGTTAAATACACGTGAGTTTCCGGTTTTATAAGATCACATA
 CTGCAAACCAACAGTAAATTACATTAATACGCATTTCTTTAGAAATTCTTAATTGCTAGGCACTGACATATAAATTGAGT
 TTGCTGTTGTTTTAGAGAGAAAAATAGACATATTCATTTTTTTTCAATCAGGACATGGTATTCGTTTCCTCTTAAGCTT
 CCCATTTTTGAGGTAGGAAAGTATTGGTGTGTGCTCTATCTCTACCTTGCTCCCTCATAGCAGGCTTAATGGGCAGGA
 ATTATGTCTGTTTTGCTTTCCGTTGGATCCCCGGGTCTAGTAGAGCGCCTTGACACATAGTAGGTACTCGACACAGATGC
 TGGCTTCATGATGAAGCAATTGCTCTCGAATTAATACAAAAGGAAGATTGAGACTCAGCGAGCCGGTACCATTCTGCTA
 GGGACACACAGCCAGGAAGGAACCGAACCAGGATTCTAACCCAGGATTACCCAAGTTCGAATTTCTAATTGAACATAAGTC
 CTTAGCGACCTTGGACGGACTTACTAACTGCCCTGAATCTCGGCTTCCTTAGCTGCAAAATAGGGTTAGGAAAAGTGGA
 CTCGGTGCTCCTCTGCGCGCATTCCTCGCAGCGGGATTGTGAAAGCGAACTTGAGCCCCACACTGCAGAAGCCCTGAGG
 TCCTGTCTCAACCAGAATCCAGCACCCAGGACAAGCCCTCCAGCGACTTCAGACCTTCCCGCTCCCGACCCCGCCGCCC
 CCGAGAGGCGTGGTCAAGGCGGGGCCAGAAGGCGGACGCGTGACGGAGCGGTAAAGCGCAGGCGCAAAGCTTCGGTGAC
 GTCAGAGAGGCGCGCTCCAGCTCGGAGCCGACTCGCAGACGCGCCCCCGCCCTCGGCGTCTGCTGACTGG

Фиг. 9 (продолж)

SEQ ID NO:18

GTCTCTCTGCGCATGCTCACACGCGACCTCCCCACAAAGCTACTCCGCGGGCCTGCCTGGCTACTGAGTCGCGAAAATAT
ATCGTAGTGCCGCAATGACTTCTGGGTACTGTAGTCCAACATGTGGACTGGCCACTCGGCTTGAAGCTTCCTTGGGATTG
GATGTTGGGACCGTCCAAACACGGCGACTTCTCATTGTGTCAGCCTGCGGCAGTTTAAATTGGGTGGCGGTGGTGCGATGT
CCTGTTAGCTGTGGAGTCTTTTCTCTTATCTTGGCCTCCGGTTGTGCCTTAATGACACTTTTGACATTGTGTATCACA
ACCAACTGGCAAAACTAGTAAATCACGGGCTCCTGACGTGGCAGCTCTGGCTAAAAATGCAGGTGGACGTCCTAGGTGCC
CAGGGAGTTTCGGGCTCACTGGGGAGTAGGGTGAAGGAAACCTGAAACCGAGAGGCTGGAAGCCCGACTTGCAGCAGCATA
AAAACACCTGCAGGGGGAAGCCTACATGAGGTACGGGTGTGCATCCCGTCTCATCTTGGCCATCTGACTGATGGAGATC
CTCAGGCACCCAGCAGACTGGGGTGTCAACCCCACTACCCACTGCCCAGGGCTGACCTCGGGCGGAAGAGCGACATTCAA
ATATTTGTTGCATATGGAGCAACAGACCAGGGTGGAGAGAGAGAAAAGGACTTTCCAGGAAGGTCCAGGCTTGCCAGGAC
TTCTATGCTTGGACTGCTCTGTCCCCAAGCTCGCTGGTGGCCATCAGGAGCCTGCTTGGCACCTTTGCTTATTTCTTCGA
TTGTTCTCAACTCAGAGGGGGAGAGGCAAGATTATCACTCCCATAGCCTATAACATGCACGCTTAAACACAACCCCTCT
GACATGGCTTCAGACCGCATCCACAGGTGAAAAGCACCAAATCCAAACAATCAGGCCACAAACCCGTCAGAGCTCGCCG
GTCTGGCCGTCCAGTCGCCGCCCCGCTTGAGACTCAGCCTGAGCTACAGCTTCGAAACTAGCCTCTAGCCTTCTGGGTCT
CAGCACAACCTCTTAGCAATCCAACCTGTGGCCGTGGGCAGATACCAAGATGCACAGTTAGCACCTCCGACAGGACCGC
GGCTCGGCGGGGCTCGCCGTCTACACGCCCATTCACAGGCCACGGCCACGACACCCAGAGCTGCCTCCTGCCCTGGGCT
CCCTGGAGCCTGGTGGCCGCGGCTCATGCTGCAGTCACCGTGTGGCCAGCTGCCAAGCGCCACTACTTGACCTGAAGTC
GCAGCCCCGCGCTGCGCACCCAGTCGAGGCGCCAACACGCGAACTGGGCGGGAGTGGCCGGCCCGCGGGGGAGCGAGGAA
CGACAGGGTTCCCCCACTCGCTGCCAGACTGGCCGCGGGATAGCGCGCCCCCCCCAAAGCCTGCACCCCTGCCCTGCT
CCTCTATCATTCGGTCCCCCTTGCTCTGTCAAACCTGCGGGGGCTGCTCTGTCTCTGCCGACCCGCCACTTGCCCGAAG
TTCGGCCTGAGGAGCTTGGGGCCCTGAGGAGCGCCGACGTAGGCAGCCAGCCAGTCAATCAGTCAGAGGGTCACGGCGGC
CCCTCGGCTACTCAGTGGCAGCCAGGCCAGAAAGAAGCAGCCGCCAACCTGGAGCAGAATCCTACAGCCGACCTGACCT
AGTAGGCCCCCAGCCCTCCAGAGCTTGGGGTATTGCAGAGCCGCCACAAGGCCTGCCGGGAGTAGTAGTTGTATAGTTC
GCAGCCGGCCAGGGCGGAGCCGGGTGGGTGGTGGGGAGGGCGGGGAAACAAGCCTGGACCACAACCTCCAGGAGTACCC
GCGACGGCGGCGGCTCGCTGTGCACTCAGGCTATAGGGGGCGCATTTGGCGGAAGACCCTGGGCCGGCAGGAACCTTGC
TTGTGATTGGATGTTGTGGGACGGCGCTTCTCATTCGTCAGCCTGTGACTGTGGAGTTTGAATTGGGTGGCGGTTGACT
GTAGAGCCGCTCTCTCTCAC

Фиг. 9 (продолж)

SEQ ID NO:19

CACATCTGCGGCATCCTCTCGGCCACCGTAGAGTGTCCACGTGTCCGCTTCTCTTCCGCTACTCGCCCCGCGCTCTGT
 CTGTCAAGAGAGACTTCCGGCATGCCCTTTTTTTTTTTTTTTTTTCCCTTTCCCGCACCGTCACTGGTGCGCTTACTG
 GAAGCCGCACCGAGGCGTGTGCTGTATTTTAAAAATCCCATTTCTGGTCAATTTCTCCACGTAATCAGACACCCGAGAG
 AACTGCATATAACTCTTTGCCTCAAGGCTTTTAGACAAGCAGAGGGCCAAAACCTGCATAAAATGGAAAATCATCATCA
 TTAACATCAGTACCATCTAAATTAAGAGCTTCTAATTGTTTTAAGGGAAGGTGGAAGCTAAAATGAACCGGCCATCCAA
 CAGAGGGCACTATCTGTGGAGGACCAACAAAGGAAAACAGAGCGGAAAATATGGGCGCACATGTGCCTCCAGTTTCATGAA
 GTTATTTGGAGGGAGGGGGGAAGAAATTAACCTTACATTATTGATTACAAAACAGTTGAAAATCTGTTAGAGTGAAAACA
 GCCATGGTAATCTACAGAAGAAAAGGAGAATGGAAGTAAAGATCATTCATGTCTTGAAATAAGTTGTCTACAGTCTGT
 GAGAACTTGTGGGAGGTGATCCTAGTGTGATATATTAGTGATGAGCCTAGCAAGGAAATCAGATCACCACAGGGCTTAG
 CGCTTAGAATTAGAACTAAGTCCTCATTTATTGTGAGAACTCAAAACAGAACTTAACAAAAGTGTTTAACAACCTTCAA
 TTTTAAAGAGGGACGCTATCTCATTTTCAGACGCCATGGATCCTCACTAGCAAAACACAATGTCACAGACAAACGAGGC
 ATATCAGGTTTTTTCTAACTCATTTTCGCTACTCCATCACTTAATCATGAAATAATGTTGACTATATATATATATATATA
 TATATATATAGTCAACATTATGGGTTTTTTTTTGTTTTTTTTTTTTTTTTTTTTGTAGACGGAGTCTCGCTCTGTGCGCT
 AGACTGGAGTGCAGTGGCGAAATCTCGGCTCACCACAACCTCCGCTCCCGGGTTCAAGCGATTCTCCTGGCTCAGCCTC
 CCGAGTAGCTGGGACTACAAGCGCCCCGCCACCACGCCAGCTAATTTTGTATTTTGTAGTAGAGACGGGGTTTTACTATG
 TTGGCCAGGCTGGTCTCGAATCCCTGACCTCGTGATCCGCCCGCCTCAGTCTCCCAAAGTGATAGGATTACAGGTGTGAG
 CCACCACACCCGGCTCTGATGTTTACAATATTATACCAACAAGGGTTGGATCGTGTGTGACTTATCTAGAGGTTAAGAGC
 AAAAGTTATGAAATCAGAGTCCCATCATGGGAGTCTAGCTTTACCATTAAAGTTGTGTAGGCAGCCAGGCGTGGTGGC
 TCACGCTGTAATCCTAGCACTTTGGGAGGCCAAGGTGGGCGGATCACCTGAGGTGAGGAGTTCCGGACAGCCTGGTCA
 ACATGGTGAAACCCCATCTCTACTAAAAATACAAAATTAGCTGGGGTTGGTGGTGTGCGCTGTAATCCAGGTACTCG
 GGAGGCTGAGGCAGGAGAATCGCTTGAAGCTGGGAGGTGGAGGTTGAGTGAGCCGAGACCATGCCATTGCACTGCAGCC
 TGGGCGACAAGAGTGAACCTCCATCTCAAAAATAAAAAAAATAAAAAATAAAGTTGTGCAGGCCGGGCGCGGTGGCTCA
 CGCTGTAATCTCAGCACTTTGGGAGGCCGAGGTGGGCAGATCATGAGGTGAGGAATTTGAGACCAGCCTGACCAATATG
 GTGAAACTCCGTCTCTACTAAAAATACAAAATTAGCTGGGTGGTGGTGGCGCATTCCTGTAATTCAGCAACTCAGGAG
 GCTAAGGCAGGAGAATCGCTTAAATCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCCACTGCACTCCAGCCTGGG
 CAACAGAGCAAGACTCTGCCTCAAAAAAAAAAAAAAAAAAAT
 TGTGTATATATACATACGTGTATATATACAGTATATAAAGTGTGAACCTGGGAAAGTTATTTAATCTTATTGTGTGT
 TAGTTTCCCATTTTATAACATGGGTGAGTAATATGTATCCTATTTCTTAGAGTCAATGGGAAGGTTAAATAGCAAAATGA
 AGGTAAAAGTAATTAGAATGTAACAGGCACATTACAATCACTCAATAAATATTCTTCTGCCTCAGCCTCCCGAGTAGCTG
 GGAATACAGGTGCGCACCAACCATGCCTGGCTAATTTTTGTATTTTGTAGTAGAGACGGGGTTTACCATATTGGCTAGGCT
 GGTCTTGAACCTCCTGACCTCGTGATCTGCCTGCCTCGGCTTCCAGAGTGCTGGGATTACAGGCATGAGCCACCGTGCCC
 AGCCGGCTGACTATTTCTTAAGAATAAGGATAGTTGGTCACCTCGGAGGCTGAGGCAGAGAATCGCTTGAACCTGGG
 AGGCGGATGTTGCAGTGACCCGAGATCCCGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGATGCGGTCTCAAAAAA
 AAAAAAACAACAAACGAAAAAAGAAAAATAGTCGGCGGGGACGGTGGCTCAGCCTATAATCCAGCACTTTGAG
 AGGCCGAGGCGGGGCGGATCACTTGCCTGAGGAGTTGGAGACAGCCTGGCCAACATGGTGAAACCCCTGCTCTCTCTTA
 AAAATACAAAATTAGCTGGGCATTGTGGCGCAGGCTGTAATTTTAGTTACTCTGGAAGCTGAGCCAGGAGAATCGCTT
 GAATCCGGGAGGTCGAGGTTGCAGTGAGCCGCGATGAAGCCATTGCTCTCCAACCTGGGTGACAAGAGCTAGACTCTGTC
 TCAGGAAAAAAAAAAGAAAAATAGGAAAGCTGCCCGCAAAGCCTGCACCTTGGACACCTCGTCTGCACTAACATTTGG
 GAAATTTCTTATTATTTCTCCGGAACATATATATTTACATTCTTTACATGTCTGTTTACGTTAAAGAAAATATTTCT
 GGAGCAATCAATTTTTTTTCTTTTTTTTTTTTTTTCTGAGACGGAGTTTCGCTCTTATTGCCCAGGCGGAGTGCAA
 TGGTGGCATCTCGGCCACTGCAACCTCCGCTCCCGGTTCAAGCGATTCTCCTGTCTCAGCCTCCAGATAGCTGGGA
 TTAGAGGCATGCACCACACGCCCCGCCAATTTTTTTTTTTTGTAGTAGAGACGGGGTTTCCGCAATTTTGGCCAGGCTG
 GTATCGAACTTCTGACCTCAGGTCACTACCTGCTTCGGCCTCCCAAAGTGTTGGGATTATAGGCGTGAGCCACTGCCCC
 TGGCTTCCAATTTTTTTTCTTTTTTTCTTTTTTTTTTTTTTTTTTCTGAGACGGATTTTCGCTCCTGTTGCCCAGG
 CTGGAGTGCAATGGTGTGATCTCGGCCACTGCAACCTTGTCTCCCGGTTCTGGCGATTCTCCTGCCTCAGCCTCCCG
 AGTAGCTGGGATTACAGGCATGAGCCACACGCCCGCTAATTTGTATTTGTATTTATTTATTTATTTATTTATTTCTGTAG
 ATACGTGGTTTTCGCATGTTGGCCAGGCTGGTATCGAACTCCTGACCTCAGGTGATCCGCCCGCCTTGCCCTCCCAAAGT
 GCTGGGATTACAGCGGTGGCCATCGCGCCCGGCCAAATATTTTTTCAAAATCTAATTTTGCATAATGACCAAGATTGTCAG
 AAAAAACAAATTAAGAGATCTTTTAAAGTAAAGATTCTGCATAAAAGTCACAGAAATAATACTCATGAATAATAAATT
 AGAATGGGCAATGTTAAGTACTTGAATAACAAAGGTTTLAGAGTCAGGGTTTCTGCAGTTTCCCTAGCCATATAACGAG
 GAGACTGTTAGAGAGGATATCCTAAAATGCTTTCCAGTTTACCACCTCTAACTCGGAAGAAGGCAATTTAAACTAATT
 TGGTTTACCGGAATTCAAATACGTACAATTCAACATAATGAATTGAAGTGAACACCTGCTTGGCTAAACCTAGT
 CGGCTCGAGTTGTTCTGGTCTGGATTGGTCAGTCTAGAGTTCCAGGGGTGTGGCCAAAAGTCTGTGAGGCGGGGCGGTA
 GGCAGCCTAAATCAGGCATCGGCGCGGTGAGCCTCGTGCGCGCCACGCCCCACGCCGGCTCTTCCGGGGTCTTCC
 GTGCGCGTTGATATGATTGGCCGGCGAATCGTGGTTT

Фиг. 9 (продолж)

SEQ ID NO:20

AAAGGAAAGCTAGGGGCCACCTTGCCGCTTTTCAGGGTCTGCGCAGGCGCCTTGAGCACTGCGTCGCGACGGGAGAGAG
 CTTTACGGCTTCTGTGCGGCTAACGGGCGGGGCGTCTGGTCAGCGCGGGAAGGCCCTTAGGCGCTCAGAGCGGGCGGG
 CCCTGTGAGGCGGAAGTGCGGGCTCTGTGGGCGGAAGGGGCGGGGCTGGGAGTGTCTGTGCGGACTGAAGGGAAGGC
 TGGCGGAGTGTGCTTCGCGTCTCTGTAGTCGACTACCGCAGCTCTTCTTTGCATTCTGGCTCTTTAATAATCTCCCTTC
 GCCCTTCTTTAAATCAGTTCCACCTCCCAACACTGTAACTTTCTCCCTCTGGACTCCCTCCTCCCGAATCCCC
 CGCATCAGAATCTTCTTTGGAATCTACTGCAAGAAATCTGGCTGGATTCTCAGATTAAGGACTTCCAGAATCCTATCCT
 CCTAGAATCCCCCTCCTTGCGGATCCTCTCCAGAAATCCTTTTCCCTCAGAGCCCCCTCCCTCAGAATCTCCCTTTC
 ACCAGAATGTCTCTTCTTTCAGTTTCCCTCTTCCAGAAATGTTTCCCATCTTCTCTCAGATGGCCAACCCAAATGTCCC
 TCTTAAGAGTGTCTCACTTAGACCCCCCTTTCCCTCAATCTCCCTCCATTGCGTCTCGTCCCTCAGTCACTCCCTTCCC
 AGAATCTTTCTCTTGTGAGGACCCAGTACCCCTGGAATCTCCCTCTCCAGAATCTCCCTTTCCGGCAAAATCCCCCTTT
 CCCGGCAAAATCCCCCTTTCCAGAACACCCCGCCCCGCAATCCACAAAACTTTTCTCTCAGAGACCCGTCCCTCAAAA
 TCTACCTTTCTCCTGAGTTTATTTCCCTTCCGAGACTCCTCTTCCAGCAACTGCCTTCCCTGGAGTTTCCCTCCGTTA
 GAATCCCCCTCTCCTAAGAGACCCCTCCCCAAGAATTAACCTCTTCCCGTCCCCACCTAGAACTCTCTCTTTTTTTT
 GAGACGGAGTTTCGCTCGTTGCCCTGGCTAGAGTGCAGTGGTGCATCTCGTCTCACTGCAACCTCCGCTCCCGGGTTC
 AAGAGATTCTCCTGCCTTAGGCTCCCGAGTAGCTGGAATTACAGGCGTCTACCACCACGCCCTGGTCTTTTTTGTATTT
 TTAGTAGAAATGGGTTTACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATGTGATCCGCTCAGCCT
 CCCAGAGTCTGGGATTACAGGCGTGAGCCACCACGCCCGCCGAATCTCTCTGTTTTTGTTTTTTTTTTTCTTTTTT
 TTGAGACAGAGTCTGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCATGATCTCTGCTCACTGCAACCTTGCCCTCCCGG
 TTAAGAGCATTTCTCTGCCTCAGCCTCTTGAGTAGCTGAGATTACAGGCATGCACCACCACGCCCGGCTAATTTTTGTAT
 TTGTAGTAGAGACGGGTTTACCATATTGGCCAGACTGGTCTCGAACTCCTGACATCAAATGATCCACTCATCTCAGCC
 TCCCAAAGTGTGGATTACAGGCGTGAGCCACCGCGCCAGCCCTAGAATCTCTTTCTCCAGAATCTCCAGTCACTCT
 TTTTCTCTCTCCCCCAAATCACCTTCACTTAAGCCTCTCTCCCCACAGAATTCTCCATCCCCAGCCCCCAACCCACCTG
 AGGAACCCAGGATTGGGTTGTCTCTGGTTCCCGGCGAGATGTCTCTGCGTCTCACCCCATCAAGAGTCTCCCCAGCGA
 AGAGTTCTCTCTCAGAGCTAGAGGGAGCCACAGTTGCCTCAGTTTCTCTGGCAGTACAGGGTCTGGGATCCAGGGTT
 CAAAGGGGCGTCTGGTGACCTTGCCAAAGAGGGTGGAGACAGGGAGACACTGGAGGGCAGGTCTGGGGCCAGAGGGAA
 AAGCCAGCCCGGAGTCCAGCTCTAAAGGGTGAGGGGGGAGGGGGCTGTGGGAGGGAGGAAGGATACCCAGAGGCTG
 ACCGGGAGGGGACACAGACACAGCCCAAAGAAATGTAGGCACAAGTGGAGAAGAAGAGAAAGATGAAAGGCCAAGGGAGAA
 GAGGAGAGAAAAGGAGAGACACAGGACAGTCAAGAGATGGACCCCAAACAAGCCAGTGACAGATTAGAGGCAGAG
 AGGCCAACCTCAGACAAAAAGAGATAAAGGCAGGCATGATGTTTCCAAAATACTTGAGACAGAGGCCAGGTGCAGTGG
 CTCACACCTATAATCCAGCACTTTGGGAGGCCAAGGCAGCAGGATCACTCGAGCCAGGGTTCAAGACCAGCCTGGCCA
 ACATGGTGAAACCCCGCTTTACTAAAAATACAAAAATCAGCTGGGCATGGTGGCAGGTGCTGTATCCAGCCACAAGG
 GAGGCTGAGGCAGGAGAACCCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATGTGCGCACTGCACCCTAGGCT
 GGGCGCCAGAGCAAGACTCTGTCTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
 AATAGAGCAGAGACCTCAAATATCACTTGACAGACAGATAAAGGCTAGCACAGGTGTTTTCCAATGAAGACACAGGCA
 GATGAGAGGATGCACAGATTCACTGTACTTTTTCTTTTTTTTTTTTTTAAATTGAGATAGAGTCTGTCTCACCAGGCT
 GGAATGCAGTGATGCAATCTGAGCTCACTGCAACCTCCACCTCCCGGTTCAAGTGATTGTCTCTCAGACTCTGGAG
 TAGCTGAGATTACAGATTCTCTGCTTCAGACTCCTGAGTAGCCAGATCACGGGCTCCACCCTATGCTTGACTAATA
 ATTTTTCTTTTTTTTTTTTGGAGATGGAGTATCGCTCTGTGCGCCAGACTGGAGGGCAGTGGCTTGATCTCAGCTCACTGC
 AACGTCAGCTTCTCGGGTTCAAGCAATTCTCCTGCATCAGCCTTCCAAGTAGCTGGGATTAAGGCATGCACCACCATGC
 CCAGCTAATTTTTGTATTTATTAGTAGAGACAGGGTTTCACTATGCTGGGCAGACTAGTCTTGAACCTCTGACCTCAAGT
 GATCCACTTGCTTCCCTCCGCTCCCAAGTGTGAGGATGAGCCACTGTGCTGGCCACTAATTTTTGTATTTTTT
 TAGTAGAGACAGGGTTCCATCGTGTGGCCAGGCTGTCTTGAACCTCTTACCTCAAGTGATCCGCCACCTCGGCCCTCC
 CAAGGTTCTGGGATTACAGGTGTGAGCCACAGCGGTTGGCCTAGCCTGCACTGTCTTAAGGGCTGTGGGCAGGGGAGT
 TCAGGTTCTCTCGTCTCTCCAGATTGATCTTACACATTTGCTAAGTCTCGTATCGTCCCTGAGCTTGGGAGTCAGGGAG
 CGCTCTGACCTAAGCCCTCTGACTTTGGGGTACCATCTTTCCCGAGCCTCTATGGGCACACTCCCTGAATATCTGTG
 GCTTGAGTGAGCAAGCAGGGGATGGGATGAGGGGTGGCGGGGAGTTCCCTCATCCACTTCAGCCTTCCAGTCTGCTC
 CCCCTTCTGTACTCTCTCTTCCAGAACCCAGGCCAAGCACCAGAGAAGGCTGGGAGTGAGCAAGGTGCCCATGTTG
 CTTCTGTGGATCATCTCTCTGTGTACAGCCCTGGGAGGGCCTAGGAACGCCCGCCCTCTGTGAGGGGAGCCTTTTT
 ACTTCCTTGACGCCCTCTGAAGATACCAGCGAGTGGGAGCAGAGCCAGCAGACGGGGTGGGCCAATTGATGGGGCGTGG
 TAGGGGGTGTGTTTGGGAAGGGTCAGCCTCTTTCCCTCCAAGATCTGATTCTCTCTCTTACCCCTGAACCTGTAGGAAA
 CAAAATGATGACACCCTCTACACCCAGATGATCTCAGTGAGTATTGCATTGCGTCTTTCTGCGAGTTGCAGGTTTT
 GAGAGAATTCAAGTGTCTCCATGTGTGAGGGCTGTGTGTGACCATATGAATATGTTGCGGGGAGGAAGTAGTGGGACTA
 TCTCTGACCCAGCCCTAGCCTTGACAGATATATTTTGTCTACTGCTGTCTCCAACACCTAGAATAGGGAGGTATATAA
 TAACAAATGATGATGATGACTACAGCTGCATGTAATAGTGTCTTACTCTGTGCCAGGTATTGGTTTAAATGCTTTAAGTA
 TGTAGCTTATTTTACTCTGTACACCAATCCTATAAGATAGATAGGACCACCATGTATGGATGTGAATGTTGTACTACGCA
 CAAGGATCCACATCTACAAAGACAGTACTATATTCACAGGCTGGATATTGTAGATTGTCATATTGCAACTTGCGGCA
 GATGGCAGTAAGTATCTTGTCTAATAAAATCAAGACTTAGAAGGTAGAAGGAACCTTAAGGAAGGTCTTTTTATTTTCT
 TTTTTTTGAGACAGGGTCTTGCTCTGTCCCTACGCTGGAGTGCAGTGGCACAATCACAGCTCACTGCAGTCTTGACCTT
 TTAGGCTCAAGCAATCCTCCACCTCAGCCTCCAGAGTAGCTGGGACCACATACACGTGACACCACAAATAGCTAATGGC
 TAATTTTTTTTTTTTTTTTGGAGACAGTGTGTCTGTGTGCCAGCTGGAGTGCAGTGGCGTGATCTGTCTCACTGC
 AACCTCTGCGCTCGGTTGAAGCAATTCTTCTGCTCAGCCTCCGAGTAGCTGGGACTACAGGCATGTGCCACCATGC
 CCAGCTAATTTTTGTATTTTGTAGTAGACAGGGTTTACCATGTTGGCCAGGATGTTCTCGATCTCTTGACCTCGTGAT
 CCACCTGCCTCAACCTCCAAAGTGTGGGATTACAGGCGTGAGCCACCGTGCCCGACCTTTTTTTTTTTTTTTTTTTT
 TTGTAGATATGGGGTGTGCTGTGTTGCCAGGCTGTTTTGGACACCTGGGCTCAAGTGATCTTCTGCTTGCCCTCT

CAAAGAGCTGGGATTACAGGCATGAGCCACTGTGCCAGCTGAGTGTCTTTTTCTAGCTTGACACATAAGTACAGCATGGG
 ATAGTGTGACGGCCTTACGGTTGGGATTATCCAGAGCCCTATTTACAGATGAGGAACTGAGGCCAGAGAACTGAAGGGG
 TTTGTTCCAGGACACAGATAAGAAATCTTGGGAGTGGGGACTTGGACTCATGCCCTTACCCATAGATACCAGGTCACCTT
 GTCCCTACTGTGTCTCTGCTGAGACTCTAGGCTGCTCCAGGTTGGAGCTGTCTGTGATTGGGTTTGAGGAGGGACCATC
 TATGGGAACCTGCTGTGTTCCAGCTACAGTACCTGCTGGATCTCTGGGGTCCCGGCTGTGGGTGGCCAGCCAGGCCCTT
 GCCTGTCTGTGAAGGACATCCCCAGAGACCTGTCCCCCTCTTGGAGTTACTACAGGGAGCTTGTGCTTGGCACCTTACC
 ACAGCCTGAGCTTTGGAGGACATGGGCTCCCAGGTACACCCAGAGAGGCTGGGTAGAGAGAGGGAGGGAATAGTCTTCAG
 CTCCAACCTGCCAAACCTAAGCCTTTCTTCCCTCTAGGCCAGAGGACCCACCTGGAGCCCTAGTAGCCAAGGGGAAGA
 GTTAAGAAGGAGGAAAGTAAGGTGCGATCTGAAGTCTCTGACCTCTTGGGCAGGGAGTTGGATAAATTGCCCCACCTGCC
 CCCATTCTCACCTCTCTTAGATACTCTTCTCTCTTCTTCCAGGAAGGGACAGAGAAGTGCACAGGTATTCTCTAACCTC
 ATCCCTATACCTTGAGAACAGGGAGGGTCTGTGGTTCACATCCCTACATCCCTGGCTACAAAAAATTACCACTTGAAGA
 ACTTTCTTTTCTTTTCTTTTCTGTTTTGAGGCAGAGTCTCACTCTGTTGCCAGGCTAGAGTGCAGTGGCATGGTCTCA
 GCTCGCTGCAACCTCCACCTCCTGGGTCAAGTAATTCTCCTGCCTCAGCCTCCTGAGTAACTGGGATTACAGGCACGCG
 CCACTACACCTGACTAATTTTTGTATTTTTAGTAGAGATGGGGTTTACAATGTTGGTCACGCTGGTCTCGAACGCTGA
 CCTGTGATCTGCCACCTCAGCCTCCCAAACCTGCTGGGATTACAGGCGTGAGCCACCGTGCCAGCCACCACTTGGAG
 AACTTTCAAGTAATACTGATGGCCAGACCCAGCCAGATAATTTGATAATTGGTCTCTGGTGAAGCCTGAGCATCTATA
 TTATTTTTATTATTATTAGAGACAGGGTTTCACTATGCTGCACAGGCTGTAACCTGGCATCTGCCCATATATTTTTTGG
 AGACAGGGTCTCAGCTCTGTACCCAGGCTGGAGTGCACTGGCATGATCAAAACTCATTTGCAGGCTCGATCTTTCAGGCTC
 AAGTGACTCTCCACCTCAGTCCCCCAAGTAGCTAGCGCTATAGGCGCATGCCACCACACTCAGCTAATTTTTGTATTTTT
 TTGTAGAGATGGGGTTACCTTTTATTGCCAGGCTAATCTTGAACCTCCTGGCTCAAGCGATACTCCCCCTTGGCCTCCC
 AAAGTGCTGGGATTATAGGCATGAGCCACCACACCTGACATTTTTTTTTTTTTTTTTTGAGACAGAGTCTCCCTTTGTACC
 CAGGCTGGAGTGACCCGGCTGGTGGACTGGATGGCCACATTTCCATCACCCCAAAGCATTCCTTCTCTCTCTTCCCAGTC
 AATTTCCACCTCCCTCTCCAGCCCCAGACATCCACTAATCTGCATTGTCTGTATAGGTTGGTTTTGCTTGTCTGGAGTT
 TCATATGGAATGGAAATGACTGTGCATCTCCTGATGTCTGGCTTCTTCACTCAGCATATGTTTTGGAGCATTCAGTCA
 GGGTGATTGGTAGTTGTTACTATTCTGTTGCATGGATGGATGTACCAAAATTTGTTTTCTTTTGAGACGAGTCTCACT
 CTGTTGCCAGGCTGGAGTGCAGTGGTGCATCTCTGCTCACTGCAACCTTCACCTCCCTGGTTCAAGTGATTCTCCTGC
 CTCAGCCTCCCAAGTAGCTGGGATTACAGGCGTGCGCCACCACACCTGGCTATTTTTTTTTTTTTTTTTTTGTATTTTT
 AGTAGAGATGGGGTTTACCATATTGGCCAGGCTGGTCTTGAACCTCCTGACTTCGTGATCCGCGCTGCCCTTGGCCTTCCAA
 AGTGCTGGGATTACAGGCGTGAGTCACTGCGCCCGGCCCTAATTTTTGTGTTTCCAGTTTAGGGCTATTATGCAACTCTT
 TCTTTCTAAAATATTCTCAAGTATTCTGATGTGTACCAGAGTTGAGTTCAGAACCACGGATCCACCAAGGAAGAACTT
 CCATATGGCTGGGAAATGACTGTGCCCTCTTTTTTCTGCAGGGAATGGCTTTCTGCAGCAGTGTCTGTAGCGATTTAATAT
 GTAGAATCCTATTTTCTTGTGTTTTTTTTTTTGGAGACGGAGTCTTGCTCTTTTGTACAGCAACCTCCACCTCTCAAGT
 TCAAGAGATTCTCTGCCACAGCCTCCCGAGTAGCTGGGATTACGGGTGTGCGCCACTACGCCACCTGGCTAATTTTTGT
 ATTTTTTAGTAGAGACAGGGTTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGACCTCAGGTGATCCGCGCTGCCCTG
 GCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGTGGAGAATCCTTTCTTTTTTTGAGATGGAGTCTAG
 CTCTGTTGCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCACTGCAAGCTCGGCCCTCCCGGGTTACGCCATTCTCCT
 GCCTCAGCCTTACCTGTAGCTGGGACTACAGGCGCCGACCCAGCCAGCTAATTTTTTTTTTTTGTATTTTAGAGAGA
 TGGGGTTTACCGTGTATGCCAGGATGGTCTCGGTCTCCTGACCTCGTGATCCACCCGCTCGGCCCTCCAAAGTGCTGG
 GATTACAGACGTGAGCCACCGCGCCCGGCCAGGTGTGTCATACTTATGTGAGAAATTTTACCAATGTGTCTGAGGCTT
 CTTTTTTTTTTTTTTTTTTTTTGGAGCTGAGTCTCACTCTCGCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGC
 AAGCTCCGCGCTCCTGGGTTACGCCATTCTCCTGCCCTCAGCCTCCCGAGTAGCTAGAATCACAGGCACATGCCACCACGC
 CCGGCTAATTTTTTTGTATTTTTTAGTAGAGACGGGGTTTACCATGTTGGCCAGACTGGTTTCGAACCTCCTGACCTCAGG
 CAATCTGCCCCGCCCGGCTCCCAATGTGTTGGGATTACAGGCATGAGCCACCGCGACTGACATAGACTTCTGATTTTTA
 AAATTTCCCTATCAGGTTAATTTTTCAATTTTTTAAAAATTTAAGTTTAAATTTTTTTTTTTTTTTTTTTTAGAGACAG
 GGTCTCTATTACCCAGTCTGGAGTGCAGTGGTGCATCATAGCTCACTGAAGCCTCAACATGGGTTCAAGTGATCCTCCC
 ACCTTAGCCTTTGGAGTAGCAGGGACCACAGGCACACGCCACCACACCATTAACAAAAACACAAAAATTGGATACATG
 GGGTCTCAATTTGTTGCCCATGCTAGTCTCAAATACCTGGCCTCAGGTGATCCTCCGCGCTCAGCCTCCCAAAGCGCTGG
 GATTACAGGCGAAGGTTCAATTTTTCATCTCTCATTACACCTGTCCACTCTGCCATCACTCGTCAGCCGTCATTACTAAT
 AGATAAATTCGCCTTTAAATGAACCTTAGGAAGCCTGGAAAGAAAAATAGGGACTTGGGCCGGGCGCGGAGGTTATGCCT
 GTGATCTCAGCACTTTGGCAGGACGAGGCCGGAGGATGGCTTTTTAGTTTGAAGTTAGAGGAGTTTGAAGACCGCGGGC
 AACAGCCCGACTCTACCAAAAAAAGTACAAAAAGAAAGAAAAAGGAAAAACAGGGCTCAGCGTCACTGAGTCACT
 CAGACTTGGCGAGGAGGCTTGGGTCTGCCGCTTTCCACTGAGTTTTTCTGAGCTTCAGCGTTCTCTCCAAAACGGGGATA
 CGGATCTCCTGGGGATGCTTTAGAGGGCTAGGAGATAAGCCTTGTCTTTTATAATTTGCTATTTCGCTGATTTGCC
 GCCTCCTCGCAGCTTGGCCTACCGCGCTAGTGTCCGACTAACGAACAAACGGCTCACTAAATTCAAACATGAGTTTCA
 GCCTGGTTCGCCGAGCCGGCGGGTAGAAAAACCTTTCCCTCCTAACCTCCGAGATGTGCGCCCCGCACTCCGCCCCAG
 GCCAGCCCTGCCTCTCTCAAGATGGCGACCTGGAAGTGCCTGCCCCGTGCCAGCCCCGAAATCTCGTGTAGGCGAGGT
 CAACGGGTAGGTTCTCGCGAGAGGACCCGTCAGCCCCAGTCAAGCGTC

Фиг. 9 (продолж)

SEQ ID NO:21

[illegible]

ACACGGAGTTTCCACCATGTTGGCCAGGCTGGTCTTGAACCTCTGACCTCAGGTAATCCACCCGCTCAGCCTCCCAAAGT
GCTGGGATTACAGGGATGAGCCACCATGCTTGGCCCAATTTGATTTTTTTTTTTTTTTTTTTTTTTTGGAGATGGAGTCTTGCTCT
GTCAGCCAGGATGGAGTGCAGTGGCGTGATCTCGGCTCACTGCAACCTCTGCTTCCCGGGTTCAAACGATTCACTACCT
CAGCCTCCCGAGTAGCTGGGATTACAGGCGCCACCACCATGGCTGGCTAAATTTTTTTGTGTTTTTAGTAGAGACAGGGT
TCACCGTGTTAGCCAGGATGGTCTCGATATCTGACCTCATGATCCGCTGCCTCGGCTCCCAAAGTCTGGGATTGCA
GGCATGAGCCACACGCGCCGCCCAATTTGATTTTTTAAAGTCCACTAGCAAGATTCAATTTGATTTTTCTGATGAGAC
AGAGCGTATTTTCTATGTTTTATAAGCCAGACATATATCTTTATCTGGAACTTTTTTTGAGTCCACAGCCCAATTTTTTT
AATCCATTGGTAAGCTAGTTTTATATATTTATAAATACTTATGTATAAGGAAATACACCTTTGACCTTCATATAAGATGCA
AAGTTATTTTTCTTTGCCACCATCCAGATTCTTACTTATTTTTTAAATTTGTTTATGGTATATTTTTCCATGCAAAAAGT
TTTTTTTTTTTATTTTTATTTTTTGCTAGTCAAGTGAAGCAGTGGGAGTAGAGAAGGAACAAATAAATCTGTAAGTGGT
GTAATTAATTAGTTGTAAACTCCACTGCACCTCGGACCAGCACCATGCAAAAAGTTTTTAAAGTCTATGTTTTTTTTTTTT
TTTTTTTTGAGACGGAGTTTTCATCTTGTGGCCAGGCTGAAGTGCAGTGGCGTGATCTTGCCCTCACCTCAACCTCTGT
TCCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACCACCCCAAGCTAATTT
TTGTATTTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGCTGGTCTTGAACCCCGCCGGGCCAAAAGTCTATGTAT
AAAATTAACCTTGCCGGGCGCAGTGGCTCACACCTGTAATCCAGCACTTTGGGAGGCCGAGACGCGTGGGTACCTGAGG
TCAGGAGTTTCAGACCAGCCTGACCAACATGGAGAAACCACGTGTCTAGGAAAAACACAAAATTAGCCCGGCGTGGTGGT
GCATGCCTGTAATCCAGCTACTCAGGAGGCTGAGACAGGAGAATCGTTTGAACCCAGAAGGCGGAGGTTGCAGTGAGCC
GAGATTGCGCCACTGCACCTCCAGCTGGGCGAGCAAGAGCAAGAACCCGATCTCAAAAAAATAAATAAATAAATTAACCTGT
CACACTACTGCCTTAGGTTTTGCATCAAAATCTTTAAAGGTTACATGTGTTTTCTGTGATATTTTTTAAAGTTTTTAT
TTAAATCTTGAACCAACCTAAAATTTATGTCACTGTCTCCCGAGGAATGAAGGAAATTCAGCTTTTTTCTCTCCCAAA
TGGCTACTTATCCCAACACAAGCCGTATTAATTGCATGTAAGGCTGCTGTATTCTAGCACAAATGCATGCTTTCAGAAATA
TTTGCTAAATGAATAAATGAAAAGCAGTATTTGTTTTATCTTTCTACAAATATGCATTGAGTACCGACTGCATATAACCA
CTATGCTAGGCTTTGCATATAACAGTAAATGACTCAGAAGCCAGGAAGCATATCATCTAGCGGTGAAGATAAGCATTAAAT
AAGTAATCATATAGCAATATTTTGAACAGGATAAAATTCCTCTTACAGTCAACACTCAAATGTGCTGGATAAAATATGAC
ATTTTGGCCGGGCGCAGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGGCCAAGGTAGGTGGATCACCTGAGGTAGG
AGTTTGGAGCAGCCTGACTAACATGGTGAACCTCGTCTCTACTAAAAAGTCAAAAGTTAGCCAGGTGATGTGAATCC
CAGCTACTCAGGAGGCTGAGGCAGGAGAATCGTTGAACCTGGGAAGTGAAGTTGCAATGAGCCGAGATCAGGCACTCG
CACTCCAGCCTGGGCGACAGAGCAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGACGTTTTTAAATGCAGAGG
TGAGCTCACAAGAAAGGGGTATACCTGGGGGAGGTGCTAAGTGGGAAGTGAAGAACAGAGGCTGTCTGGGGCAATCTG
TCCTCCTAAGTAATATTACATATAAAGTGGATTCCAAAGGGAGCAAGAGACAAGCCTTTGGGCAACTGTATTTAATCAA
TTGTGATGAGTAGAAATCCATGAAGGTTTTTGGGAGAAATTAATACAGCATCTGAGACTTGAGTTCTAGTCTGTTTAT
CTAACATTAGCTGTCTATGAGGTCATGGTCAAAATAGTGACTCTTACTCTTAATTCATTCCTCTTGTGCAAAATGAG
AGAATTTGAAGTAAATATCTATAACAGCAATTTACGTCTAAACTTGAGGATTCTGATAGGATATCCCTCTACAAAT
TTAATTTGGCTGCAGTATGGATGGTGGATTATTACACAGCCTCCTAACTGATACATCTGCTTTTCAGGCTTGCTCTACCT
CCAAAATCATCTTTTTACTGACCACAAGAACCTTTCTGATCAGGTCTGAATTTTATCCCGCTTAAAATTTTTCAAAGGT
TTTCAGAAGTTTCAGAGCCTGGGGGACAGGGTGTGAACCTGTCTCAAAAAAAAAAAAAATCCTAAGTGTAAATGAATG
AATAAATGTATTGAAGCAAGAAGACTGGTAAATTTAGCAACAGCTCTGTAGCAAAGTAATGCAGCTCTGAATTGATACG
GCAGTGCTAAGAAATGGAAGGTTGGTCATTTAGTAAATAATCGTGGAAACACAATCTATAAGACATTCAAAGAGTGTC
ACTGATGAAAGGCAGAACTTAGAGGAGTGAAGTTGCAGATACAGGATGAATCACTTTTGTAGAACCAACAAATA
GGGCGAGAAAGGCAAAAAGGAAGGCTCATGTTTGTATGTCTGAGATGAAGAACTGTTTCTGAAGACTTTCTAAAACTC
CACAAGAGATCACTTCACATCCTTCACACATGTCTCCTTTGCACAGTTGGCAGGTTTCAGACATATACATATATTTCTA
TGACAAGGCTTATCATTCTTTAGGACTGCAGTAATTTAGATAAGAAGCTCTCAAAGAACACTTGCCAGTAACAGCATC
TCCACCAATGAAGTATGACAGCTCTGGCTCTGAGCCTCTGGAACCAATGAAGTCTGTTTATTGAACAGAACTCTTGCTT
CTGTTTCCAAGCAGCTTTATTTGAAGTCTGCTCTTTTGGCAATAAAAGCTTCCTTTTACCTTCCCTCAGTGGAGTACT
TTTGGCTTGCCATTTCTGTGATCTCTGAATTTCTAATTTCTCTTCCCAATTTCTGAATAAATTCACATATTTGGAGATATT
TTTTCTCAGTCTTTTTTTTCTTTTTTTTGGAGACTGAAAGGCTGATGCTTAGGTGCATATTTTTTAACTATTTAGCTG
TATTGAGTCTTAGATGATGTTTACCTATAAAAAAACTAACTCCAACATCTGGATCAATGTGCTAGGGCAAGCAAGT
CTACACTGTCCATAGATTGTGACCTCGTTCTTTAAGAACTGTATGTGGGTGGGGCATGGTGGCTCATACCTGTAATCCGA
GCACCTTTGGGAGGCTGAGGCAGGAGGACAGCTTGAGGTGAGGAGTTCAAGACCAGACTAGGCAACATAGCAAGAACCC
ATCTCTAGCAAAAAGTAAAAAACAAACAAAAAACACACACATCTGGCCATGGTGGCAGGCACCTGTAGTCTCAGCTA
CTCGGGAGGCTGGGGTAGGAGGATGGCTTGAGCCAGGAGTTTGAGGTTGCAGTGAGCTATGAATGTTGCCACTGCACCT
CAGCCTGGGTGAAACCTCTCCAAAACATAAGCTAAACTAAAAAATAAATAAATCTGATGTGACTGTGATTTTACTAGC
TCAAACATAATGAGCTACAAAGTGAAGCAAGTTTCACTATGACTACAAATGAAAGTTGAAAGAAATCCATTTGCTCGCAGG
CGCGGGTGGCTCAGCCTGTAATCCCTCAGCACTTTGGGAGGCTGGGGCGGTTGGATCACCAGAGGTCAGGAGTCTGAGACC
AGCCTGACCAATATGGTGAACCCCGTCTGTACTAAAAATACAAAAATTAAGTGGGCATGGTGACAGGCACCTGTAATCC
CAGCTACTCAGGAGGCTGAGGCAGGAGAATCGCTTGAGCCGGGGAGGCTGAGGTTGCAGTGAGGCAGATCTGCCACTG
CACTCCAGCCTGGGCAACAGAGCGGAAGGAAGGAAGAAATGAAGGAGGAGGGAGGGAGGGAGGGAGGGAACATTATCT
GTTGAATGCCATGATGTGCCAGAAAGCCAGGGAAATAAATGAAAAGTGCAATTTGGAGCTTGCAAAAAGGAAGGTCAAGA
TGCTAAAGAAGGTTAGCGGAAACTGAACCTGGGATTTTGATCATATTAATTAATGAACAAAAAGAGAAAGAAATGTA
GTTTGTAGGCCAGGTAGAAGAAGAACTAATTTTTATAGAGAAAAACAAATCTAAATGTTTATCTCTTTTATTTCTTGTC
AAAAACTTGTAGTCAATATTTTTTAACTCCTTATGTATTTCCATTTTCACTGTCTGTCTACAGGAAAACTGACTCGAGGT
AATTAAGGATAGAGATCAAGGGGCGAGTTTTATTTGTATGGTGGTAAAGTTCAAAGGGCTCTTTGCTCCACAATCGATCC
ACTAAGACCAAAAACCTTTAGGTGTACTCTAGAGTGTACATCCAAGTATTTAAACCATTAAATTTGGGTGGACAAGGGG
AGGATGTGGAAGGGGAAATGAGATACCAAAACTGATAGATGTTTAAAAAAGAGGTAGTTTGGAAACTACCTTTTTTAAG
ACTGACAATTAGGAAGATACATTTTTTGGCAGTATATACCAGCTGTTTCACTTAAAAATCACCGTTGAAGCTGTTATGTA
TTTATCTATTTTATTTAATTAATGAACTCATGCTTCAACACCCAGCTAAGTAAACACCAATTAAGGACTGAGATTGAG
GCTTTTTTTTTTTTTTTTTTGGAGCAAGGCTCTGCTCTGTGCGCCAGGCTGGAGGCGAGTTCTCTGCTCACTACCTCTGCC
TCCTAGGCTCAAGCTGGGACTACAGACGCGCGCCGCCGCCCACTAATTTTTTTCTTCTCTTTTTTTTTTTTTTTTTTTT

18/27 ЛОКУСЫ «БЕЗОПАСНОЙ ГАВАНИ»

GAGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCACTATCTCGGCTCACTGCAACCTCCGTCTCCCAGGTTCAAG
CGATTCTCCTGCCTCAGCTTCCCGGGTAGCTGGGACTACAGGTGCGCGCCAACACTCTGTCTAATTTTTGTATTTTTAGT
AGAGACGGGGTTTTACCATGCTGGCCAGGCTGGTCTCCAACCTCCCGACCTCAGGCGATTCTCCCGCCTCGGCCCTCCCAA
GTGCTGGGATTACAGTCGTGAGCCACCGCGCCCGGCTAGTTAATTTTTTAATTTAATTTTTTTGTGGAGACGAGGTCTC
ACTATTGCCAGGCTAATCTCGAACTTCTGGGCTAAGCGATCCTCCTGCCTCGAGCCTCCCAAATTTGTGGGCTTACAGG
CGTGAGCCACCGCACCTGGCGGAACATTTCTACTCTAATGAAAATTATAACAGTTTTTACAGTTTTCTAGCTCCCAACT
GTACCTATAAGTTCAGAGTCCATTTAATGTACTTGTAAATAATTCAGGCTGGTGTGTTTTATTATCTTTTCTACATTATTACC
CTCTCAAAGAACAAAAGCAGGTAACCTCTATAAAGTCTAATAGTCTGAAAATTGTAGTGTCGGTTTTCTCTTTAGCCATA
AAGTTTTGCTTAAATGCTGAGTCGCGTTTTAATGCTCTACTAGATGTACATATAAGAATACTAAATGGAGGGAAAGTAAT
TTCTTTTTTTTTTTTGAGACGGGGTCACTCAGGCTGGAGTGCCAGTGCGCGATTACAGCTCACCGGAGCCTCTCGACCTC
CCCAGGCTCAAGTGATCCTGCCACCTCAGCCTCCCCAATAGCTGGGACTACAGACACGCGCCACCACGCCCCGGCTAATTT
TTTTGTATTTTTCTTAGAGACTGGGTTTTGCCATGTTGCCCAAACCTGGTCTTGAACCTCCTGGGCTCAAGCGACCTGCCTG
CCTCAGCCCCGCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGCCATAATTTTATCTTAATCACTTTTTTCCA
GGAATATTACGTTTCAAAGAATGAGGAATTAACAGCAACACTTCAGGTTCAACATGTAAATAGGGACGCTCATTGAG
AATAAATCACATTTTGTACAGTTGTATTGAAAAATCGGTATTTCTGAACGTCAATTTACCATTACGGAAGTATAACTCA
GACACAATAAAATGTTAATATACGGCTAGGATTAGGTTAAGGTGAGAGAGGCGCCTAGAGCACAAATTTAAGGAGGCA
AGACCACGACCCGACCGTGAGCGCTGCTTAACATTTGCTCCCCAGGCGCTTCCCACTCCCCAGCCGGGCCCCGTCCTA
GTTCAATCAATCAAGAAAATCAAGAGTACGAAGGCCTTGTCGGTCTGGACGGTCTTCCACGTCTCAGGTGCAGTTAAG
CCAGTCACACTCCATCCACCTACCCGCCATGAAAAATTATTTGTTTCAGCAGGAATTGCCAAGTCAAATACAGCCATATAA
AGCCCCCTTCGACGTGCCAGGTGTGCCTCCGGAGAGCAGCCACCAAAACCACCAGCAGCCCGTCCGCGCCGGGAAATGC
GCCCTCGCTGGCTTACGGGGCCACTCAGACAGTTCTAAGTGCCTTTCCCTGTAACGCAGAGGCACACCTGCCTTCCCTCG
GGAAACAGCGGCCCCGACGTGCGCCCGCAACGCGTAAAGCGTCAACAGCTGAAACCTTCGGAATCCTCCCCGAACAGAAAC
AACCACCGCTGCCAGCTGCGCGCTCGGGGGAAAAGACGTTGCGCCCCCGCGACTGCCGGTTTCCCGGGCGCGAGCCCGG
ATCCAGGTGGTCAGTCCCGGTACGCAACCACGGCGAGAACC CGCCCTGCTAAGGGAGAAGGGAAGCCGTTTCCCGCGGG
CTTCATACAGACACGGTGCTAAGACAGCCGAGTCGCGGACCAGTCCAACAGGCAGGCAGATCAGTGGGCGAGCCCGCGTG
C

Фиг. 9 (продолж)

SEQ ID NO:26: >Донор 1 MED20 (сайты TetO выделены жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом)

GTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGGGACTACAGGAGAATCGCTTGAACCCAAGAGGCGG
 AGTGTTGCAGTGAGCCAAGATCGCTCCACTGCACTCCAGTCTGGCGACAGAGCGAGACCCCGTCTTAA
 AAAAAACAAAACAAAACAAAAAACAACCTTTGCTTTCCCGCCTTCCAAGAGGAAGGGGCGCATCTCT
 GAGGAAACAAGCCCGGGGAACCCGCGAGCCCGGACGGCATCCTCTCATCTACACAACCCGAGGACATCT
 CCCTCAGTCCCAGAGGACGGCGGACCATGGTCAAGGTCTCTTTCCGGCCTTTACAACTCCAAGCC
 CCGCCCCACAAAACCCCTGCGAGTCTACCAAGTCACTCCCATGGCGTCGGGCCAGGAAGGTGGCAGA
 ATCACACAGTAGAAACGCAGGGTTCCCTGTCCGCCACAGAACTCCTTCAGTTCCCAACACAACCT
 TCTGTCTCAGAAGGGACTCCGGAATACGTAAAAACAGAGCGCCGGGTCCACGCCTTCCCCAAGGAAT
 GGGGCGCGTCTCGGCACCTAGGCTAGCGATCCTCTAGATTGAGTTTAC**TCCCTATCAGTGATAGAGA**
ACGTATGAAGAGTTTACTTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACT**TCCCTATCAGTGATA**
GAGAACGTATAAGGAGTTTACT**TCCCTATCAGTGATAGAGA**ACGTATGACCAGTTTACT**TCCCTATCAGT**
GATAGAGAACGTATCTACAGTTTACT**TCCCTATCAGTGATAGAGA**ACGTATATCCAGTTTACT**TCCCTAT**
CAGTGATAGAGAACGTATAAGCTTTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGCTAGCTGCAG
 TAACGCCATTTTGAAGGCATGGAAAAATACCAAACCAAGAATAGAGAAGTTCAGATCAAGGGCGGGT
 ACATGAAAATAGCTAACGTTGGGCCAAACAGGATATCTGCGGTGAGCAGTTTCGGCCCCGGCCCCGGG
 CCAAGAACAGATGGTCACCGCAGTTTCGGCCCCGGCCCCGAGGCCAAGAACAGATGGTCCCCAGATATG
 GCCCAACCCCTCAGCAGTTTCTTAAGACCCATCAGATGTTTCCAGGCTCCCCAAGGACCTGAAATGAC
 CCTGCGCCTTATTTGAATTAACCAATCAGCCTGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTTCCCGA
 GCTCTATAAAAGAGCTCACAACCCCTCACTCGGCGCGCCAGTCTCCGACTCGAGAGCGCTGCCACCA
 TGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTA
 AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAA
 GTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCCACCTCGTGACCACCTGACCTACGGCG
 TGCACTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAA
 GGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAA
 GTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACA
 TCCTGGGGCACAAGCTGGAGTACAACCTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAG
 AACGGCATCAAGGTGAACCTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCA
 CTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCC
 AGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCC
 GCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGATCCGGAGCGACTAACTTCTCCCTGCTGAA
 GCAGGCCGGCGATGTAGAAGAGAACCCGGGCCCTACCGAGTACAAGCCCACGGTGCGCCTCGCCACCC
 GCGACGACGTCCCCAGGGCCGTACGCACCCCTCGCCGCGCGCTTCGCCGACTACCCCGCCACGCGCCAC
 ACCGTGATCCGGACCGCCACATCGAGCGGGTACCGAGCTGCAAGAACTTCTCCTCACGCGCGTCCGG
 GCTCGACATCGGCAAGGTGTGGGTGCGCGACGACGGCGCGCCGTGGCGGTCTGGACCACGCGGAGA
 GCGTCAAGCGGGGGCGGTGTTGCGCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCTGGCTG
 GCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGACCCGGCCCAAGGAGCCCGCGTGGTTCTGGC
 CACCGTCCGAGTCTCGCCCGACCACAGGGCAAGGGTCTGGGCAGCGCCGTGCTGCTCCCCGGAGTGG
 AGGCGGCCGAGCGCGCCGGGGTGGCCGCTTCTTGAAACCTCCGCGCCCCGCAACCTCCCTTCTAC
 GAGCGGCTCGGCTTACCGTCAACCGCCGACGTGAGGTGCCGAAGGACCGCGCACCTGGTGCATGAC
 CCGCAAGCCCGGTGCTGAGGCACCCCTCGATGACGCTCCACTGACATCATCTGCGGATTCCTCGCCC
 CGAGGTTTCTAAATCCAGACATTCCTGTTTGGCTGAGCACTCTAGGCCTACATCCATGAAGCTAGGAG
 AGCGACACTCAAAGACTGCACTATTGAGAGAAGCTAACGTTAAAGGCAGTGAATATATTTCGGGAGTAC
 AGCTCCGGAACGCGTGAGCTCTTTAGTGGGAGGGGCGGCGCTGATGGCGCTTCTGGCCTCCGAATGC
 TAGGGGGCGCTGTGATCGCCGGGCGGCCTCTTGGGCGCTGGGAGTCCACCGCGCAAGCGCATCCTGGC
 CTTTCTTCAGTCCCCACGTGCGATCCTTCCCGGCAACTTTTTCGAGAAAAATGCCCAAATCAAGGCG
 GCCCGTGGGGTGGGGGGTCAGGAAAAACATGCGCCCCTGGCCGATCAGATCCTGGCTGGGAATGCGGT
 GCGGGCGGGGTCCGGGAGAAGCGGCGGGGTGCGGGGACAGGAGAAGCGGAGGAAGAGTATGTGGGGC
 CCCGGCTGAGCCGACGATTTTGCAGCAAGCACGGCAGCAACAGGAGGAACCTCGAGGCCGAGCATGGG
 ACTGGGGACAAGCCCGCGGCGCCGCGGGAACGCACCACGCGGCTGGGTGAGTGTCTGGGATGAGGTCC
 GAGGAAGACAGTGGCCAGTAGTGGGGCGGGGTGGGCAGCTAAAAAGTGTCTGGCAGGGGAATTGCTTC
 GAGTCAACAAA

Фиг. 9 (продолж)

SEQ ID NO:27: >Донор 2 MED20 (сайты TetO выделены жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом)

GTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGGGACTACAGGAGAATCGCTTGAACCCAAGAGGCGG
 AGTGTTCAGTGAGCCAAGATCGCTCCACTGCACTCCAGTCTGGCGACAGAGCGAGACCCCGTCTTAA
 AAAAAACAAAACAAAACAAAAAACAACCTTTGCTTTCCCGCCTTCCAAGAGGAAGGGGCGCATCTCT
 GAGGAAACAAGCCCGGGGAACCCGACGCCGACGGCATCCTCTCATCTACACAACCCGAGGACATCT
 CCCTCAGTCCCAGAGGACGGCGGACCATGGTCAAGGTCTCTTTCCGGCCTTTCACTCACTCCAAGCC
 CCGCCCCACAAAACCCCTGCAGTCTTACCAAGTCACTCCCATGGCGTCGGGCCAGGAAGGTGGCAGA
 ATCACACAGTAGAAACGCAGGGTTCCCTGTCCGCCACAGAACTCCTTCAGTTCCTCAACACAACCT
 TCTGTCTCAGAAGGGACTCCGGAATACGTAAAAACAGAGCGCCGGGTCCACGCCTTCCCCAAGGAAT
 GGGGCGCGTCTCGGCACCCCTCGATGACGCCTCCACTGACATCATCTGCGGATTCCCGCCCCGAGGTT
 TCTAAATCCAGACATTCCTGTTTGGCTGAGCACTCTAGGCCTACATCCATGAAGCTAGGAGAGCGACA
 CTCAAAGACTGCACTATTGAGAGAAGCTAACGTAAAGGCAGTGAATATATTCGGGAGTACAGCTCCG
 GAACGCGCTAGGCTAGCGATCCTCTAGATTTCAGTTTTAC**TCCCTATCAGTGATAGAGA**ACGTATGAAG
 AGTTTAC**TCCCTATCAGTGATAGAGA**ACGTATGCAGACTTTAC**TCCCTATCAGTGATAGAGA**ACGTAT
 AAGGAGTTTAC**TCCCTATCAGTGATAGAGA**ACGTATGACCAGTTTAC**TCCCTATCAGTGATAGAGA**AC
 GTATCTACAGTTTAC**TCCCTATCAGTGATAGAGA**ACGTATATCCAGTTTAC**TCCCTATCAGTGATAG**
GAACGTATAAGCTTTAGGCGTGTACGGTGGGCGCCTATAAAAGCAGAGCTAGCTGCAGTAACGCCATT
 TTGCAAGGCATGGAAAAATACCAAACCAAGAATAGAGAAGTTCAGATCAAGGGCGGGTACATGAAAT
 AGCTAACGTTGGGCCAAACAGGATATCTGCGGTGAGCAGTTTCGGCCCCGGCCGGGGCCAAGAACAG
 ATGGTCACCGCAGTTTCGGCCCCGGCCGAGGCCAAGAACAGATGGTCCCCAGATATGGCCCCAACCT
 CAGCAGTTTCTTAAGACCCATCAGATGTTTCCAGGCTCCCCCAAGGACCTGAAATGACCTGCGCCTT
 ATTTGAATTAACCAATCAGCCTGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTTCCCGAGCTCTATAAA
 AGAGCTCACAAACCCCTCACTCGGCGCGCCAGTCTCCGACTCGAGAGCGCTGCCACCATGGTGAGCAA
 GGGCGAGGAGCTGTTACCCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGGCCACA
 AGTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTTCATCTGC
 ACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGTGCACTGCTT
 CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCC
 AGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGC
 GACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCA
 CAAGCTGGAGTACAACACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCA
 AGGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAG
 AACACCCCATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCT
 GAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCTGTGACCGCCGCGGGATCA
 CTCTCGGCATGGACGAGCTGTACAAGGGATCCGGAGCGACTA**ACTTCTCCCTGCTGAAGCAGGCCGGC**
GATGTAGAAAGAGAACCCGGGCCCTACCGAGTACAAGCCACGGTGGCGCTCGCCACCCGCGACGACGT
 CCCCAGGGCCGTACGCACCTCGCCGCGCGTTCGCCGACTACCCGCCACGCGCCACACCGTCGATC
 CGGACCGCCACATCGAGCGGGTCAACGAGCTGCAAGAACTCTCCTCACGCGCGTCCGGGCTCGACATC
 GGCAAGGTGTGGGTTCGCGGACGACGGCGCCGCCGTGGCGGTCTGGACCACGCGGAGAGCGTCAAGC
 GGGGGCGGTGTTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGC
 AACAGATGGAAGGCCTCCTGGCGCCGACCGGCCAAGGAGCCCGCGTGGTTCTGGCCACCGTCCGA
 GTCTCGCCCGACCAACAGGGCAAGGGTCTGGGCAGCGCCGTCTGTCTCCCGGAGTGGAGGCGGCCGA
 GCGCGCCGGGTGCCCGCCTTCTTGAAACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCG
 GCTTCAACGTACCCGCCGACGTGAGGTGCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCC
 GGTGCTGAGCGCGTGAAGTCTTTTGTGGAGGGGCGCGCTGATGGCGCTTCTGGCCTCCGAATGC
 TAGGGGGCGCTGTGATCGCCGGGCGGCTCTTGGGCGCTGGGAGTCCACCGCGCAAGCGCATCCTGGC
 CTTTCTTCAGTCCCCACGTGCGATCCTTCCCGGCAACTTTTTCGAGAAAAATGCCCAAATCAAGGCG
 GCCCGTGGGGTGGGGGGTCAGGAAAAACATGCGCCCCTGGCCGATCAGATCCTGGCTGGGAATGCGGT
 GCGGGCGGGGTCCGGGAGAAGCGGCGGGGTGCGGGACAGGAGAAGCGGAGGAAGAGTATGTGGGGC
 CCCGGCTGAGCCGACGATTTTGCAGCAAGCACGGCAGCAACAGGAGGAACCTCGAGGCCGAGCATGGG
 ACTGGGGACAAGCCCGCGGCGCCGCGGGAACGCACCACGCGGCTGGGTGAGTGTCTGGGATGAGGTCC
 GAGGAAGACAGTGGCCAGTAGTGGGGCGGGGTGGGCAGCTAAAAAGTGTCTGGCAGGGGAATTGCTTC
 GAGTCAACAAA

Фиг. 9 (продолж)

SEQ ID NO:28: >Донор 1 FTSJ3 (сайты TetO выделены жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом)

TCCGGCCTCGCAACCGACCACCCCTTCTGCAGGAGAGGTACATAATGGTCGCAAAGAGCAGGCGAACCGGGCAGA
 GATCGCTCGCTCGCTCGCAGACTGAGCGAATCCGGACTCACCCGTCTCCTTCGCCAAGTGATAAACTTGTCTCG
 TCGGCTCTTGCCAACTTTGCCCTTCTTGCCCATGGTGGAAAGGGGAGTATCCCTCAATCTGGGGAGAAAAGTAGA
 AGTTGGGAACGTCTCTTTCCGGAGCACCAGATTCCACTTCCAGTTTCCCACTTCCAGTTTCCCACTTCCCGCAGC
 AATTCACCCCGCGCCCTCTCCGCACACTACCTAGACCCAGAGCCGCTTTCTCCACACTTGGAAACCGCACAAAGT
 ATGCAGCTAACTACTTCCGCTTCCGCTTGGCTCCCTGGCTCGAGGTTTTCCGCCATTTTGAGTGTG
 GCCCTAGATTGTTCTCAATACTCTGTCTTGGGTTTTGTAAGTTGAAAGTTGAGACTAGGAAGGCATATCACAAAG
 AACTATAAACAGAGTTTCAATGAGGCAACACTGAGGAGGAGTTTCTACTCCTTCTCATCCCTTCCAAAGCACCT
 GAACTCTAGCTAGGCTAGCGATCCTCTAGATTCTAGTTTACT**TCCCTATCAGTGATAGAGA**ACGTATGAAGAGTTT
 ACT**TCCCTATCAGTGATAGAGA**ACGTATGCAGACTTTACT**TCCCTATCAGTGATAGAGA**ACGTATAAGGAGTTTACT
CCCTATCAGTGATAGAGAACGTATGACCACTTTACT**TCCCTATCAGTGATAGAGA**ACGTATCTACAGTTTACT**TCC**
TATCAGTGATAGAGAACGTATATCCAGTTTACT**TCCCTATCAGTGATAGAGA**ACGTATAAGCTTTAGGCGTGTACG
 GTGGGCGCCTATAAAAGCAGAGCTAGCTGCAGTAACGCCATTTTGCAAGGCATGGAAAAATACCAAACCAAGAA
 AGAGAAGTTTCAAGGCGGGTACATGAAAATAGCTAACGTTGGGCCAAACAGGATATCTGCGGTGAGCAGT
 TTCGGCCCCGGCCCGGGGCCAAGAACAGATGGTCACCGCAGTTTCGGCCCCGGCCCGAGGCCAAGAACAGATGGT
 CCCCAGATATGGCCCAACCTCAGCAGTTTCTTAAGACCCATCAGATGTTTCCAGGCTCCCCAAGGACCTGAAA
 TGACCCTGCGCCTTATTTGAATTAACCAATCAGCCTGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTTCCCGAGCT
 CTATAAAAGAGCTCACAAACCCCTCACTCGGCGCGCCAGTCCCTCCGACTCGAGAGCGCTGCCACCATGGTGAGCAA
GGGCGAGGAGCTGTTACCGGGGTGGTGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAG
CGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCT
GCCCGTGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCGACCACAT
GAAGCAGCAGCACTTCTTCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA
CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCAT
CGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTATATCAT
GGCCGACAAGCAGAAGAACGGCATCAAGGTGAAGTTCAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT
CGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGGGATCCGGAGCGACTAACTTCTCCCTGCTGAAGCAGGCCGGCGA
TGTAGAAGAGAACCCGGGCCCTACCGAGTACAAGCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCAGGGC
CGTACGCACCCTCGCCGCCGCTTCGCCGACTACCCGCCACGCGCCACACCGTCGATCCGGACCGCCACATCGA
GCGGGTCAACGAGCTGCAAGAACTTCTCTCACGCGCTCGGGCTCGACATCGGCAAGGTGTGGGTTCGCGGACGA
CGGCGCCGCCGTGGCGGTCTGGACCACGCCGAGAGCGTCAAGCGGGGGCGGTGTTTCGCCGAGATCGGCCCCGCG
CATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGACCCGGCCCAA
GGAGCCCGCTGGTTCTTGCCACCGTCGGAGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTGCT
GCTCCCCGAGTGAGGCGGCGGAGCGCGCGGGGTGCCCGCTTCTGGAAACCTCCGCGCCCCGCAACCTCCC
CTTCTACGAGCGGCTCGGCTTACCGTCACCGCCGACGTTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGAC
CCGCAAGCCCGGTGCCTGAGCTAGAAGCCGCCCTCCCATCATGGTTCCCTTAGTGTGGTCTCGCCGCACAACCC
 GCCCATGTACACGGAATTCTTCTCTGCCTGGTCTTCAGGTCCCAATCCTCCGCTTCCGCGCTTGCGCGCCAAGAC
 GGCTCGGATGCCGGCGGTCTCTGCTGAAGAGAGAAGATGGCGCTTGACGGACCAGAGCAGGTATGGCGGGTGCAG
 TGGCGGGCCCGCAGGTTACGGGGCTGGGTGCGGAGCGAGCGTGATCTGAGTGGAGAGCGGGCCGGGGCAGGAGCG
 TCGGGTGGGTTCGAGGTGCCTGGACAGTCCATGCTGGACGCTGCCTGCGACCGGATCTGCTGTGTACGCGCCGC
 CCTCGGTGAGTCAGGGCAAGGCTGGGAGAGGAAGACGCGGTAGAAGCGGAGTGAGTGAGGGAAGCGATGGGCGCG
 GGAATGGCCGGCCACGGGTGCGAGGAGACGGGACGCCAGTCCCTTTGGCTCCGTTCCGCTGGCTCCTTCGTCACT
 ACTGACACCTCGGGCTTGTAGAGCACTTACGCGAGTGCAAAGCGCCCCCGTCTATATCATATCGCCTCTCGGTC
 CTCTAAAAGTCGATGAGGTAGGCGGCGTTCCCATTTCTTTTATTTTAGTCTGGGAAAATGGAGGGTTCCAAGG
 GAGGGAGGGATTAGAACC

Фиг. 9 (продолж)

SEQ ID NO:29: >Донор 2 FTSJ3 (сайты TetO выделены жирным шрифтом, eGFP подчеркнут, 2A-PuroR выделен курсивом)

TCCGGCCTCGCAACCGACCACCCCTTCTGCAGGAGAGGTACATAATGGTTCGCAAAGAGCAGGCGAACC
GGGCAGAGATCGCTCGCTCGCTCGCAGACTGAGCGAATCCGGACTCACCCGTCTCCTTCGCCAAGTGA
TAAAACTTGTCTCGTCGGCTCTTGCCAACTTTGCCCTTCTTGCCCATGGTGGAAGGGGGAGTATCCC
TCAATCTGGGGAGAAAGTAGAAGTTGGGAACGTCTCTTTCCGGAGCACCAGATTCCTACTTCCAGTTTC
CCTACTTCCAGTTTCCCTACTTCCCGCAGCAATTCCACCCCGCGCCCCTCTCCGCACACTACCTAGACCC
AGAGCCGCTTTCTCCACACTTGGAAACCGCACAAAGTATGCAGCTAACTACTTCCGCTTCCGCTTGCCTC
CCCTGCGTCCCTGGCTCGAGGTTTTCCGCCATTTTGAGTGTGGCCCTAGATTGTTCTCAATACTCTGT
CCTTGGGTTTTGTAAAGTTGAAAGTTGAGACTAGGAAGGCATATCACAAGAAGTATAAACAGAGTTTCA
ATGAGGCAACACTGAGGAGGAGTTCTACTCCTTCCTCATCCCCCTTCCAAAGCACCTGAACTCTAGAAG
CCGCCCCCTCCCATCATGGTTCCCTTAGTGTGGTCTCGCCGCACAACCCGCCCATGTACCTAGGCTAGC
GATCCTCTAGATTCTGAGTTTTACT**TCCCTATCAGTGATAGAGA**ACGTATGAAGAGTTTTACT**TCCCTATCAG**
TGATAGAGAACGTATGCAGACTTTACT**TCCCTATCAGTGATAGAGA**ACGTATAAGGAGTTTTACT**TCCCTA**
TCAGTGATAGAGAACGTATGACCAGTTTACT**TCCCTATCAGTGATAGAGA**ACGTATCTACAGTTTACT**TC**
CCTATCAGTGATAGAGAACGTATATCCAGTTTACT**TCCCTATCAGTGATAGAGA**ACGTATAAGCTTTAG
GCGTGTACGGTGGGCGCCTATAAAAGCAGAGCTAGCTGCAGTAACGCCATTTTGCAAGGCATGGAAAA
ATACCAAACCAAGAATAGAGAAGTTCAGATCAAGGGCGGGTACATGAAAATAGCTAACGTTGGGCCAA
ACAGGATATCTGCGGTGAGCAGTTTCGGCCCCGGCCCGGGGCCAAGAACAGATGGTCACCGCAGTTTC
GGCCCCGGCCCGAGGCCAAGAACAGATGGTCCCCAGATATGGCCCAACCCCTCAGCAGTTTCTTAAGAC
CCATCAGATGTTTTCCAGGCTCCCCAAGGACCTGAAATGACCCTGCGCCTTATTTGAATTAACCAATC
AGCCTGCTTCTCGCTTCTGTTTCGCGCGCTTCTGCTTCCCGAGCTCTATAAAAGAGCTCACAAACCCCTC
ACTCGGCGCGCCAGTCCCTCCGACTCGAGAGCGCTGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCA
CCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGC
GAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCC
CGTGCCCTGGCCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACC
ACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC
TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCG
CATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCT
ACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATC
CGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCCGA
CGGCCCGTGTCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACG
AGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAG
CTGTACAAGGATCCGGAGCGACTAACTTCTCCCTGCTGAAGCAGGCGGCGATGTAGAAGAGAACCC
GGGCCCTACCGAGTACAAGCCCACGGTTCGCGCTCGCCACCCGCGACGACGTCCCCAGGGCCGTACGCA
CCCTCGCCGCGCGTTCGCGGACTACCCCGCCACGCGCCACACCGTTCGATCCGGACCGCCACATCGAG
CGGGTCACCGAGCTGCAAGAAGTCTTCTCCTCACGCGCGTTCGGGCTCGACATCGGCAAGGTGTGGGTTCGC
GGACGACGGCGCCGCGCTGGCGGTCTGGACCACGCGGAGAGCGTCAAGCGGGGGCGGTGTTTCGCCG
AGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAGGCCTC
CTGGCGCCCGCACCGGCCCAAGGAGCCCGCGTGGTTCTTGGCCACCGTTCGGAGTCTCGCCCCGACCACCA
GGGCAAGGGTCTGGGACGCGCCGTCTGTCTCCCCGAGTGGAGGCGGCGGAGCGCGCCGGGGTGGCCG
CCTTCTTGGAAACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTACCGTACCGCC
GACGTTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGAGGTACACG
GAATTCTTCTCTGCCTGGTCTTCAGGTCCCAATCCTCCGCTTCCGCGCTTTCGCGCCAAGACGGCTCG
GATGCCGCGCGTCTCTGCTGAAGAGAGAAGATGGCGCTTGACGGACCAGAGCAGGTATGGCGGGTGCA
GTGGCGGCCCGGCAGGTTACGGGGCTGGGTGCGGAGCGAGCGTGATCTGAGTGGAGAGCGGGCCGGGG
CAGGAGCGTCGGGTGGGTTCGGAGGTGCCTGGACCAGTCCATGCTGGACGCTGCCTGCGACCGGATCTG
CTGTGTACGCGCCGCCCTCGGTGAGTCAGGGCAAGGCTGGGAGAGGAAGACGCGGTAGAAGCGGAGTG
AGTGAGGGAAGCGATGGGCGCGGGAATGGCCGGCCACGGGTTCGAGGAGACGGGACGCCAGTCCCTTT
GGCTCCGTTCCGCTGGCTCCTTCGTCAGTACTGACACCTCGGGCTTGTAGAGCACTTACGCGAGTGCA
AAGCGCCCCCGTCTATATCATATCGCCTCTCGGTCTCCTAAAAGTCGTATGAGGTAGGCGGCGTTC
CCATTCTTTTATTTTAGTCTGGGAAAATGGAGGGTTCCAAGGGAGGGAGGGATTAGAACC

Фиг. 9 (продолж)

SEQ ID NO:30: >TetR-KRAB

MARLDKSKVINSALELLNEVGIEGLTTRKLAQKLGV EQPTLYWHVKNKRALLDALAIEMLDRHHTHFC
 PLEGESWQDFLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAFLCQQGFSLENALYALSA
 VGHFTLGCVLEDQEHQVAKEERETPTTDSMPPLLRQAIELFDHQGAEP AFLFGLELIICGLEKQLKCE
 SGSPKKRKVDGGGALSPQHS AVTQGSIIKNKEGMDAKSLTAWSR TLVTFKDFVDFDTREEWKLLDTA
 QQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGE EPWLVEREIHQETHPDSETAFEIKSSV

SEQ ID NO:35 Донор гомологии MED20 (левый гомологический участок выделен жирным шрифтом, правый гомологический участок выделен курсивом)

GTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGGGACTACAGGAGAATCGCTTGAACCCAAGAGGCGG
 AGTGTTCAGTGAGCCAAGATCGCTCCACTGCACTCCAGTCTGGCGACAGAGCGAGACCCCGTCTTAA
 AAAAAACAAAACAAAACAAAAAAAACCTTTGCTTTCCCGCCTTCCAAGAGGAAGGGGCGCATCTCT
 GAGGAAACAAGCCCGGGGAACCCGACGCCCGGACGGCATCCTCTCATCTACACAACCCGAGGACATCT
 CCCTCAGTCTCCAGAGGACGGCGGACCATGGTCAAGGTCTCTTTCCGGCCTTTCACAACCTCCAAGCC
 CCGCCCCACAAAACCCCTGCAGTCTACCAAGTCACTCCCATGGCGTCCGGGCCAGGAAGGTGGCAGA
 ATCACACAGTAGAAACGCAGGGTTCCCTGTCCGCCACAGAACTCCTTCAGTTCCCCAACACAACCT
 TCTGTCTCAGAAGGGACTCCGGAAATACGTAAAAACAGAGCGCCGGGTCCACGCCTTCCCCAAGGAAT
 GGGGCGCGTCTCGCTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCA
 GAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCA
 GGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCAT
 CCCGCCCTAACTCCGCCCAGTTCCGCCCATTTCTCCGCCCATGGCTGACTAATTTTTTTTATTTATG
 CAGAGGCCGAGGCCGCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCTA
 GGCTTTTGCAAAAAGCTGCGATCGCGCCACCATGGTGAGCGAGCTGATTAAGGAGAACATGCACATGA
 AGCTGTACATGGAGGGCACCGTGAACAACCACTTCAAGTGCACATCCGAGGGCGAAGGCAAGCCC
 TACGAGGGCACCCAGACCATGAGAATCAAGGCGGTGAGGGCGGCCCTCTCCCCTTCGCCTTCGACAT
 CCTGGCTACCAGCTTCATGTACGGCAGCAAAACCTTCATCAACCACACCCAGGGCATCCCCGACTTCT
 TTAAGCAGTCTTCCCCGAGGGCTTCACATGGGAGAGAGTCAACACATACGAAGACGGGGCGTGCTG
 ACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCTCATCTACAACGTCAAGATCAGAGGGGTGAA
 CTTCCCATCCAACGGCCCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGAGACCCTGT
 ACCCCGCTGACGGCGGCCTGGAAGGCAGAGCCGACATGGCCCTGAAGCTCGTGGGCGGGGGCCACCTG
 ATCTGCAACTTGAAGACCACATACAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCCCGGCGTCTA
 STATGTGGACAGAAGACTGGAAGAATCAAGGAGGCCGACAAAGAGACCTACGTCGAGCAGCACGAGG
 TGGCTGTGGCCAGATACTGCGACCTCCCTAGCAAACCTGGGGCACAGATGAACCGGTCTGTGCCTTCTA
 GTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACT
 GTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGG
 TGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGG
 GCTCTATGCCCCTCGATGACGCCTCCACTGACATCATCTGCGGATTCCCGCCCCGAGGTTTCTAAATC
 CAGACATTCCCGTTTGGCTGAGCACTCTAGGCCTACATCCATGAAGCTAGGAGAGCGCACTCAAAGA
 CTGCACTATTGAGAGAAGCTAACGTTAAAGGCAGTGAATATATTCTGGGAGTACAGCTCCGGAACGCGT
 GAGCTCTTTTAGTGGGAGGGGCGGCGCTGATGGCGCTTCTGGCCTCCGAATGCTAGGGGGCGCTGTGA
 TCGCCGGGCGGCCTCTTGGGCGCTGGGAGTCCACCGCGCAAGCGCATCCTGGCCTTTCTTCAGTCCCC
 ACGTGCGATCCTTCCCGGCAACTTTTTTCGAGAAAAATGCCCAAATTCAAGGCGGCCCGTGGGGTGGGG
 GGTGAGGAAAAACATGCGCCCTGGCCGATCAGATCCTGGCTGGGAATGCGGTGCGGGCGGGGGTCCG
 GGAGAAGCGGCGGGGTGCGGGGACAGGAGAAGCGGAGGAAGAGTATGTGGGGCCCCGGCTGAGCCGAC
 GGATTTTGCAGCAAGCACGGCAGCAACAGGAGGAACCTCGAGGCGGAGCATGGGACTGGGGACAAGCCC
 GCGGCGCCGCGGGAACGCACACGCGGCTGGGTGAGTGTCTGGGATGAGGTCCGAGGAAGACAGTGGC
 CAGTAGTGGGGCGGGTGGGCAGCTAAAAAGTGTCTGGCAGGGGAATTGCTTCGAGTCAACAA

Фиг. 9 (продолж)

SEQ ID NO:36 Донор гомологии **FTSJ3** (левый гомологический участок выделен жирным шрифтом, правый гомологический участок выделен курсивом)

TCCGGCCTCGCAACCGACCACCCCTTCTGCAGGAGAGGTACATAATGGTCGCAAAGAGCAGGCGAACC
GGGCAGAGATCGCTCGCTCGCTCGCAGACTGAGCGAATCCGGACTCACCCGTCTCCTTCGCCAAGTGA
TAAAACTTGTCTCGTCGGCTCTTGCCAACTTTGCCCTTCTTGCCCATGGTGAAAGGGGAGTATCCC
TCAATCTGGGGAGAAAGTAGAAGTTGGGAACGTCTCTTTCCGGAGCACCAGATTCCACTTCCAGTTTC
CCACTTCCAGTTTCCCACTTCCCGCAGCAATTCCACCCCGCGCCCTCTCCGCACACTACCTAGACCC
AGAGCCGCTTTCTCCACACTTGGAAACCGCACAAAGTATGCAGCTAACTACTTCCGCTTCCGCTTGCGTC
CCCTGCGTCCCTGGCTCGAGGTTTTCGCCCATTTTGAGTGTGGCCCTAGATTGTTCTCAATACTCTGT
CCTTGGGTTTTGTAAAGTTGAAAGTTGAGACTAGGAAGGCATATCACAAGAACTATAAACAGAGTTTTCA
ATGAGGCAACACTGAGGAGGAGTTCTACTCCTTCCTCATCCCCCTTCCAAAGCACCTGAACTCTCTGTG
GAATGTGTGTCAAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAAGTATGCAAAGCATGC
ATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAAGTATGCAAAGC
ATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCCATCCCGCCCTAACTCCGCC
CAGTTCCGCCCATTTCTCCGCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCT
CTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTG
CGATCGCGCCACCATGGTGAGCGAGCTGATTAAGGAGAACATGCACATGAAGCTGTACATGGAGGGCA
CCGTGAACAACCACCACTTCAAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAGACC
ATGAGAATCAAGGCGTTCGAGGGCGGCCCTCTCCCTTCGCTTCGACATCCTGGCTACCAGCTTCAT
GTACGGCAGCAAAACCTTCATCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCCTTCCCCG
AGGGCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGGCGTGCTGACCGCTACCCAGGACACC
AGCCTCCAGGACGGCTGCCTCATCTACAACGTCAAGATCAGAGGGGTGAACTTCCCATCCAACGGCCC
TGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGAGACCCTGTACCCCGCTGACGGCGGCC
TGGAAGGCAGAGCCGACATGGCCCTGAAGCTCGTGGGCGGGGGCCACCTGATCTGCAACTTGAAGACC
ACATACAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCCCCGGCGTCTACTATGTGGACAGAAGACT
GGAAAGAATCAAGGAGGCCGACAAAGAGACCTACGTCGAGCAGCACGAGGTGGCTGTGGCCAGATACT
GCGACCTCCCTAGCAAACCTGGGGCACAGATGAACCGGTCTGTGCCTTCTAGTTGCCAGCCATCTGTTG
TTTGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTTTCTAATAAAAT
GAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCTTCTATTCTGGGGGGTGGGGTGGGGCAGGACAG
CAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAGAAGCCGC
CCCTCCCATCATGGTTCCCTTAGTGTGGTCTCGCCGCACAACCCGCCCATGTACACGGAATTCTTCTC
TGCTTGGTCTTCAGGTCCCAATCCTCCGCTTCCGCGCTTGCGCGCCAAGACGGCTCGGATGCCGGCGG
TCTCTGCTGAAGAGAGAAGATGGCGCTTGACGGACCAGAGCAGGTATGGCGGGTGCAGTGGCGGCCCCG
GCAGGTTACGGGGCTGGGTGCGGAGCGAGCGTGATCTGAGTGGAGAGCGGGCCGGGGCAGGAGCGTCG
GGTGGGTGCGAGGTGCCTGGACCAGTCCATGCTGGACGCTGCCTGCGACCGGATCTGCTGTGTAGCG
CCGCCCTCGGTGAGTCAGGGCAAGGCTGGGAGAGGAAGACGCGGTAGAAGCGGAGTGAGTGAGGGAAG
CGATGGGCGCGGAATGGCCGGCCACGGGTGCGAGGAGACGGGACGCCAGTCCTTTGGCTCCGTTCC
GCTGGCTCCTTCGTCTGACTGACACCTCGGGCTTGTAGAGCACTTACGCAGTGCAAAGCGCCCCC
GTCTATATCATATCGCCTCTCGGTCCTCCTAAAAGTCGTATGAGGTAGGCGGCGTTCCCATTTCTTTTA
TTTTAGTCCTGGGAAAATGGAGGGTTCCAAGGGAGGGAGGGATTAGAACC

Фиг. 9 (продолж)

SEQ ID NO:37 Экспрессирующий вектор с Cas9 (с интегрированной AAVS1 sgRNA)

ttcatgcagctgacccacgacgacagcctgaccttttaagaggacatccagaaagccaggtgtccggccaggggcatagcctgcacg
agcacattgccaatctgtgcccggcagccccgccattaaagaagggcacctctgcagacagtgaaagtggtggacgagctcgtgaaagtgat
gggcccgcacaaagcccagagaacatcgtgatcgaatggccagagagaaccagaccaccagaaagggacagaaacagccgcgagaga
atgaagcgggatcgaagagggtcatcaaagagctgggcagccagatcctgaaagaacaccccgtggaaaacacccagctgcagaaacgaga
agctgtacctgtactacctgcagaatggcggggatagtacgtggaccagggaactggacatcaaccggctgtccgactacgagtgtgga
ccatctcgtgcctcagagcttttctgaaggacgactccatcgacacaaaggtgctgaccagaagcgacaaagacccggggcaagagcgac
aacgtgcccctccgaagaggtcgtgaagaagatgaagaactactggcgccagctgctgaaacgccaagctgattaccagagaaagtctog
acaatctgaccaagggccgagagagggcgccctgagcgaactggataaggccggcttcatcaagagacagctggttgaaaacccggcagat
cacaaagcacgtggcacagatcctggactcccggatgaacactaagtacgacgagaatgacaagctgatccgggaagtgaagtgtatc
accctgaagtccaagctgggtgtccgatttccggaaggtattccagttttacaaagtgcgcgagatcaacaactaccaccacgcccacg
acgcctacctgaacgcgctcgtgggaaccgcctgatcaaaaagtaccctaagctggaaagcgagttcgtgtacggcgactacaaggt
gtacgacgtgcggaagatgatcgccaagagcgagcaggaaatcggaaggctaccgccaagtacttcttctacagcaacatcatgaac
tttttcaagaccgagattaccctggccaacggcgagatccggaagcgccctctgatcgagacaaacggcgaaaacccgggagatcgtgt
gggataagggccgggatttttgccaccgtgcggaaagtgtgagcatgcccgaagtgaatatcgtgaaaaagaccgaggtgcagacagg
cgcttcagcaaaagagctctatctgcccgaaggaacacgcgataaagctgtatcgccagaaagaaggactgggaccctaaagagtacggc
ggcttcgacagccccaccgtggcctattctgtgctggtggtggccaaagtggaaaaggcgcaagtccaagaaactgaagagtgtgaaag
agctgctggggatcacccatcatggaaagaagcagcttcogagaagaatcccatcgactttctggaagccaagggtacaaaagaagtga
aaaggacctgatcatcaagctgcctaagtaactccctgttcogagctggaaaacggccggaagagaatgctggcctctgccggcgaaactg
cagaagggaaacgaactggccctgccctccaaatatgtgaacttctgtacctggccagccactatgagaagctgaagggtctccccg
aggataattgagcaaacagctgtttgtggaacagcaccaagcgtactacgtgacgagatcatcgagcagatcagcagctgttctccaaag
agtgtactctggcgacgctaatctggacaaaagtgtgtccgctacaacaagcacccgggataagcccatcagagagcagggccgagaat
atcatccacctgtttacctgaccaatctgggagccctgcgccttcaagtaactttgacaccaccatcgaccggaagaggtacacca
gcaccaaaagaggtgctggagcgcacccctgatccaccagagcatcacccggctgtacgagacacggatcgacctgtctcagctgggagg
cgacaaaaggccggcgccacgaaaaaggccggccaggcaaaaaagaaaaggatccggagcgactaatttttccctgctgaagcag
cgcgcgatgtagaagaaaacccggccctatggtgagcagcaagcgaggagctgttcacgggggtggtgcccactgttcogagctgg
acggcgacgttaaaccggccacaagtccagcgtgtccggcgaggcgaggggcgatgccacctacggcaagctgacctgaagtccatctg
caccaccggcaagctgcccgtgcccctggcccacctctgtgaccacctgacctacggcgtgcagtgcttcagccgctaccccgaccac
atgaagcagcacgacttcttcaagtcgccatgcccgaaaggtacgtccaggagcgcaccatcttcttcaaggacgacggcaactaca
agaccgcgcggaggtgaagtctcagggcgacacccctggtgaaccgcactcgagctgaagggcacgacttcaaggaggacggcaacat
cctggggcacaactcgtgagataactacaacagccacaacgcgtctatatctggtgcgacaaagcagaagacggcgcacaggtgaactc
aagatccgcacacaactcagaggacggcagcgtgcagctcgcgaccactaccagcagaacacccccatcggcgacggcccgctgctgc
tgccgcacaaaccactacctgagcaccacgtccgcccgtgagcaaaagaccccaacgagaagcgcgacacatggtcctgctggagtctgt
gacccgcgcgggatcactctcggcatggagcagctgtacaagtaaaattcacgcgttaagtgcagctcgactttaagaccaatgact
tacaaggcagctgtagatcttagccacttttttaaaagaaaagggggagctgaagggttaattcactcccaacgaagacaagatctgc
ttttgtctgtactggtgtctctgtttagaccagatctgagcctgggagctctctggttaactagggaacccactcttaagcctca
ataaagcttgcccttagtgcttcaagtagtgtgtgcccgtctgttgtgactctggttaactagagatccctcagaccttttagtca
gtgtgaaaaatctctagcagtagtatagtagttcatgtcatcttattattcagtaatttataaacttgcaaaagaatgaatatcagaga
gtgagaggaaactgtttattgagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcaact
gcattctagtgtgtgttgcacaaactcatcaatgtatcttcatgtctggtcttagctatcccgcccctaacctccgcccaccccgcc
ccctaacttcgcccaggttcgcccattctccgcccactggtccttaaaatcggggctccctttttattttagcagagccgagctccctcgcc
tgagctattccagaagtagtgaggaggttttttgaggcctagggacgtacccaattccgcccctatagtgagtcggtattacggcgct
cactggccgtctgttttacaacgtcgtgactgggaaaacccctggcggttaccacaacttaatcgcttgacgacatccccctttcgccag
ctggcgtaatagcgaagaggcccgacccgatcgcccttcccaacagtgcgcagcctgaatggcgaaatgggacgcgcctgtagcggc
gcattaaagcgcggcggtgtggtggttacgcgcagcgtgacccgtacacttgccagcgccttagcgcgcctctctctctctctctcc
ctctctctctcgccagcttcgcccgtttcccgctcaagctctaaatcggggctcccttttaggttcogatttagtgcgtttttagcga
cctcgaccccaaaaaacttgattagggtgatggttcacgtagtggccatcgccctgatagacggtttttcgcccttgacgttggaag
tccaggtcttttaatagtggactctgttccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggattt
tgccgatttcggcctatttggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaacaaaatatttaacgcttacaattta
ggtggcacttttcggggaaatgtgcgcgggaacccctattgtttattttttaaatacattcaaatatgtatccgctcatgagacaat
aacctgtacaatgcttcaataatattgaaaaggaagtagtagatttcaacatttcogtgcgcttcttccctttttagcgttcgccc
attttgccctctctgtttttgctcaccagaaaacgctggtgaaagttaaagatgctgaagatcagttgggtgcacgagtggttgatc
gaactggatctcaacagcggtaagatccttgagagttttcgcccccgaagaacgttttccaatgatgagcacttttaagttctgtctat
gtggcgcggtattatccggtattgacgcgggcaagagcaactcggctgcgcgcatacactattctcagaatgacttggttgagtagctc
accagtcacagaaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggcc
aacttacttctgacaaacgatcggaggaccgaaggagctaacgccttttttgcaacacatgggggatcgaactgcactctgtcgtt
gggaaccggagctgaatgaagccataccaaaacgacgagcgtgacaccacgatgcctgtagcaatggcaaacgttgccgaacactatt
aactggcgaactacttactctagcttcccggaacaattaatagactggatggaggcgataaagtgcaggaccactctgcgctcg
gccttccggtggtggtttattgctgataaatctggagcgggtgagcgtgggtctcgcggtatcatgacgactggggccagatg
gtaagccctcccgatcgtagtattctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgc
ctcactgattaaagcatttggaatgtcagaccaagtttactcatatatacttttagattgattttaaactctatttttaaaagg
atctaggtgaagatccttttgataatctcatgacaaaatcccttaacgtgagtttctgttccactgagcgtcagaccccgtagaaa
agatcaaaaggtatcttcttgagatccttttttctgcgcgtaactctgctgcttgcaaaacaaaaaaaccaccgctaccagcggtggtttg
tttgccggatcaagagctaccaactctttttccgaaggttaactggttcagcagagcgcagataccaaatactgttctcttagtgtag
ccgtagttagccaccacttcaagaactctgtagcaccgctacatacctcgtctgtctaatctgtttaccagtggtgctgcccagtg
cgataaagtctgtgcttaccgggttgactcaagacgatagttaccggataaaggcgacgggtcggtgacccgggggtctgctgac
acagccagcttgagcgaacgactacaccgaactgagatcactacagcgtgagctatgagaaagcgccacgctcccggaaggaga
aaggcggacaggtatccggttaagcggcagggtcggaacaggagagcgcacgaggagcttccagggggaaacgctggtatctttata
gtcctgtcgggtttcgccacctctgacttgagcgtcgattttgtgatgctcgtcaggggggcgagcctatggaaaaacgccagcaa
cgcgccctttttacggttccctggccttttctggtggttttctgctcactgttcttccctgattctgtggataaccgt
attaccgccttttagtgagctgataccgctcgcgcagccgaacgacgcagcagcagtgctcagtgagcaggaagcggaagagcgcc

caatagcgcaaacccgctctcccgcgcggttgccgattcattaaagcagctggcaacagaggtttcccgactggaaagcggcgagtga
gcgcaacgcaattaatgtgtagttagctcactcatttaggcaccccaggctttacacgtttatgcttccggctcgtatgttgtgtggaatt
gtgagcggataacaatttcacacaggaacagctatgaccatgattacgccaagcgcgcaattaacctcactaaaggggaacaaaagc
tgagctgcaagcttaatgtagtctttagcaatactctttagtcttgcaactggttaacgtagagttagcaacatgccttacaaggga
gagaaaaagcccgctgcatccgattgttggaagttaaggtgtacagatgccttatttaggaaggcaacagcgggtcgacatgga
ttggacgaaccactgaattgcccgcattgcagagatatgtattttaagtgccctagctcgatacataaacgggtctctctggttagacca
gatctgagcctgggagctctctggctaactagggaaacccactgcttaagcctcaataaagcttgcccttagtgcttcaagtagtgtgt
gcccgtctgttgtgtgactctggttaactagagatccctcagacccttttagtcagtggtgaaaaatctctagcagtgccgcgccgaacag
ggacttgtaagcgaagggaaacacagaggagctctctcagcaggactcggttctgctgaagcgcgcagcggcaagaggcgaggggcg
cgactggtgagtagcccaaaaaattttagctagcggaggctagaaggagagagatgggtgcgagacgctcagttatgaagcggggagaa
ttagatcgcgcatgggaaaaaattcggttaaggccagggggaaagaaaaaataataataaacatatagtatgggaagcaggggagc
agaacgattcgcagttaatcctggcctgttagaaacatcagaaggctgtagacaaatactgggacagctacaacccatcccttcagaca
ggatcagaagaacttagatcattatataataacagtagcaaccctctattgtgtgcatacaaggatagagataaaagacaccaaggaag
ctttagacaagatagaggaaagacaaaacaaaagtaagaccacccgcacagcaagcggccgctgatcttcagacactggaggaggagata
tgaggacaatttggaagtgtaattatataataataaagtagtaaaaaatgaaccattaggatagcaccaccaaggaagcagaag
agtggtgcagagagaaaaagagcagtgggaaataggagctttgttccctgggttcttgggagcagcagggaagcactatgggcgcagcg
tcaatgacgctgacggtacagccagacaaattattgtctggtatagtgacgcagcagacaacatttgcctgagggtctattgaggcgcaac
agcatctgttgcaactcacagctctggggcatcaagcagctcaggcaagaatcctggctgtggaagatacctaaaggatcaacagct
cctggggattgggggttgcctctggaaaactcttgaccactgctgtgcttggaaatgcttaggtggaataaatactcttggaacag
atttgaatcacacagctcgtgagtgagtcagacagaaaattaaacattacaagcttaatacactccttaattgagaatccgcaaa
accagcaagaaaagaatgaacaagaattattggaattagataaatgggcaagtttgtggaattggtttaacataacaaaattggctgtg
gtatataaaattattcataatgatagtaggaggcttggtaggtttaagaatagtttttgcctgtactttctatagtgaatagagttagg
cagggtatataccaattatcgtttcagacccacctcccaaccccgaggggacccgatccggggcggggcgaatcttttaaaagaaaa
ggggggattgggggttacagctgcaggggaaagaattagtagacataatagcaacagacatacaaaataaagaattacaaaaacaaatta
caaaaaattcaaaattttgattccggggcgggcgcaatcgattccggggcgggcgcaatcccgagtttcccatgatcttccatattt
gcataatacgatacaaggctgttagagagataaattagaatttaatttgactgtaaacacaaagatatagtagcaaaatcgtgacgtaga
agtaataaatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttcg
atttcttggcttttatatcttgtggaaggacgaacacccgtccctagtggccccactgtggtttcagagctatgctggaacagca
tgcaagttgaaataaggctagtcggttatcaacttgaaaagtggcaccgagctcggtgcttttttAATTAAGggggcccggttgc
gcttttccaaggcacccctgggttttgcgcaggacgcggtgctgtctgcggctggttccgggaaacgcagcgcgcgcacctgggtc
tcgcacattcttcacgtccgttcgcagcgtcaccgggatcttcgcgcgtacccttgtggggcccccggcgcgcttctctgctccgcc
ctaagtcgggaaggttctctgcgggttcgcggcgtgcggagctgacaaaagggaagccgcacctctcactagtaccttcgcagacggac
agcggcaggggagcaatggcagcgcgcgcgacccgatgggctgtggccaatagcggctgctcagcagggcgcgcgcgagagcagcgcgcg
ggaagggcgcggtgcgggagggcggttgcgggcgttagtggtggccctgttctgcgcgcgcgggtgttcgcgatttgcgaacctccg
gagcgcagctcggcagctcggctccctcgttgaccgaataccgcagcctctcccgaggggagcgttacctagaaggaagggcaccac
tggccccaagaaagaaagcgggaaggtcggtatccacggagtcccagcagccgacaaagagtacagcatcgccctggacatcggcacca
ctctgtgggtgggcccgtgatccgcagcagtagacaaggtgccagcaagaattcaaggtgctgggcaacaccgacccgcagacatc
aagaagaacctgatcggagccctgctgttcgacagcgcggaacacgcgagccacccggctgaagagaaccgcgagaagaagataca
ccagacgggaagaaacccgatctgctatgcgaagagatcttcagcaacagagatggccaaggtggacgacagcttctccacagactgga
agatctctctggttggaagagataaagaagcagcagcggcaccacctcttcggcaacctctggacagagtggtctaccacagagaag
taccaccacctctaccacctgagaaaagaaactggtggacagcaccgacaaaggccacctgcgctgatctatctggccctggccaca
tgatcaagttccggggccacttctgatcgagggcgacctgaaccccgacacacagcagctggacaagctgttcacagctggtgca
gacctacaaccagctgttcgaggaaaaccccatcaacgcgcgcgctggacgccaaggccatctgcttgccagactgagcaagagc
agacggctggaaaatctgatccgccagctgcgccgcgagagaagaagtggcctgttcggaaaacctgattgccctgagcctgggacctga
ccccaaactcaagagcaactctgacctgcgcgagtagtccaaactgcagctgagacaaggaacctacgcagcagcactggacaact
gctggcccagatcggcgaccagtacgcgcacctgtttctggccgcgaagaacctgtccgacgccaatcctgctgagcgacatcctgaga
gtgaacaccgagatcaccaaggccccctgagcgcctctatgatcaagagatcgaacgagcaccaccaggacctgacctgctgaaag
ctctcgtgcggcagcagctgctgagaagtacaaagagattttcttcgaccagagaagaacggctacgcgcggtacattgacggcgg
agccacgcaggagaagagttctacaagttctatcaagcccatctgaaaaagctggacggcacagggaactgctcgtgaagctgaacaga
gaggacctgctgcggaagcagcagaccttcgacaaacgcgcagatccccacagatccacctgggagagctgacgcgcatcttcgcg
ggcaggaagatttttaccattctgaaggacaacccgggaaaaagatcgagaagatcctgaccttcgcacccccactactacgtgggccc
tctggccagggggaaacagcagattcgctggatgaccagaaaagagcaggagaaaccatccccctggaacttcgaggaaagtggtggac
aagggcgtctcgcgccagagcttcatcgagcggatgaccaacttcgataagaacctgcccaacgagaaggtgctgcccaagcacagcc
tgctgtacgagtagcttcaacctgtataacgagctgaccaaaagtgaatactgacggaggaatgagaaaagccgcctctctgagcgc
cgagcagaaaagggccatcgtgacctgctgttcaagacaacccgggaaagtgcctggaagcagctgaaagaggactacttcaagaaa
atcgagtgcttcgactcgtggaattctccggcgtggaagatcggttcaacgcctccctgggcacataccacgactgctgaaaatta
tcaaggacaaggacttctggacaatgaggaaaaacaggagacattctggaagatatcgtgctgacctgacactgtttgaggacagaga
gatgctgagggaacggctgaaaacctatgccacctgttcgacgacaaaagtgatgaagcagctgaagcgcgggagatcacccgctgg
ggcagctgagcgggaagctgatcaacggcatccgggacaagcagctccggcaagacaactcctggatttctgaagtccgacggcttcg
ccaacagtaaac

Фиг. 9 (продолж)

>SEQ ID NO:38 Донор гомологии AAVS1 (левый гомологический участок выделен жирным шрифтом, правый гомологический участок выделен курсивом, кассета с трансгеном обозначена заглавными буквами, промотор SV40 подчеркнут, mKate2 выделен жирным шрифтом и подчеркнут, сигнал полиаденилирования выделен жирным шрифтом и курсивом).

tgctttctctgaccagcattctctccccctgggcctgtgcccgtttctgtctgtagcttgtggcctggg
tcacctctacggctggcccagatccttccctgccgcctccttcaggttccgtcttccctccactccctc
tcccccttgctctctgtctgtgtgctgcccgaaggatgctcttccggagcacttccctctcggegtg
caccacgtgatgtcctctgagcggatcctccccgtgtctgggtcctctccgggcatctctcctccctc
acccaaccccatgccgtcttccactcgtggttcccttttcccttctccttctggggcctgtgccatct
ctcgtttcttaggatggccttctccgacggatgtctcccttgcgccccgcctcccccttcttgtaggac
tgcattcatcacggtttttctggacaaccccaagtagcccgctctccctggccttagccacctctccat
cctcttgctttctttgacctggacaccccgcttctcctgtggattcgggtcacctctcactcctttcatt
tgggcagctccccctaccccccttacctctctagctctgtgtagctcttccagccccctgtcatggcat
cttccaggggtccgagagctcagctagctcttcttccctccaacccgggccccctatgtccacttcaggac
agcatgtttgctgcctccaggatcctgtgtccccgagctgggaccacctataattccagggccgggt
taatgtggctctgggttctgggtacttttatctgtccccctccacccacagtggggcaagCTGTGGAAT
GTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCT
CAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGC
ATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACCTCCGCCCATCCCGCCCCCTAACTCCGCCCAGT
TCCGCCCATCTCCGCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCTGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCCTAGGCTTTTGCAAAAAGCTGCGAT
CGCGCCACCATGGTGAGCGAGCTGATTAAAGGAGAACATGCACATGAAGCTGTACATGGAGGGCACCGT
GAACAACCACCACTTCAAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAGACCATGA
GAATCAAGGCGGTCGAGGGCGGCCCTCTCCCTTCGCCTTCGACATCCTGGCTACCAGCTTCATGTAC
GGCAGCAAAACCTTCATCAACCACACCCAGGGCATCCCCGACTTCTTTAAGCAGTCTTCCCGAGGG
CTTCACATGGGAGAGAGTCAACACATACGAAGACGGGGGCGTGTGACCGCTACCCAGGACACCAGCC
TCCAGGACGGCTGCCTCATCTACAACGTCAAGATCAGAGGGGTGAACCTCCCATCCAACGGCCCTGTG
ATGCAGAAGAAAACACTCGGCTGGGAGGCCTCCACCGAGACCCTGTACCCCGCTGACGGCGGCCTGGA
AGGCAGAGCCGACATGGCCCTGAAGCTCGTGGGCGGGGGCCACCTGATCTGCAACTTGAAGACCACAT
ACAGATCCAAGAAACCCGCTAAGAACCTCAAGATGCCCGGCGTCTACTATGTGGACAGAAGACTGGAA
AGAATCAAGGAGGCCGACAAAGAGACCTACGTGAGCAGCACGAGGTGGCTGTGGCCAGATACTGCCA
CCTCCCTAGCAAACCTGGGGCACAGATGAACCGGTCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTG
CCCCCCCCGTGCCCTTCCTTGACCCCTGGAAGGTGCCACTCCCACTGTCTTTCTTAATAAAATGAGG
AAATTGCATCGCATTTGTCTGAGTAGGTGTCAATCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAG
GGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGgacttactagggg
caggattggtgacagaaaagcccccatccttaggcctcctccttccctagctctcctgatattgggtctaa
ccccacctcctgttaggcagattccttatctggtgacacacccccatttccctggagccatctctctc
cttgccagaacctctaaggtttgcttacgatggagccagagaggatcctgggagggagagcttggcag
ggggtgggaggggaaggggggggatgctgtgacctgcccgggttctcagtggccaccctgcgctaccctctc
ccagaacctgagctgctctgacgcggctgtctggtgctgtttcactgatcctggtgctgcagcttccct
acacttcccaagaggagaagcagtttggaataaataacagaataagttggtcctgagttctaactt
tggctcttcacctttctagtccccaatttatattgttccctccgtgcgtcagttttacctgtgagataa
ggccagtagccagccccgtcctggcagggctgtggtgaggaggggggtgtccgtgtggaactccct
ttgtgagaatggtgctcctaggtgttcaccaggtcgtggccgcctctactccctttctctttctcca
tccttctttccttaagagtccccagtgctatctgggacataattcctccgcccagagcaggggtcccg
ttccctaaggccctgctctgggcttctgggtttgagtccttggcaagcccaggagagggcgtcaggct
tccctgtcccccttccctcgtccaccatctcatgcccctggctctcctgccccttccctacaggggttc
ctggctctgctct

>SEQ ID NO:39 (MED20 sgRNA):	GGGCGTGTCTCGGCACCCCT
>SEQ ID NO:40 (FTSJ3 sgRNA):	GAATTCCGGGTCAATGGGCG
>SEQ ID NO:41 (AAVS1 sgRNA):	gtccctagtggccccactgtg