

(19)



**Евразийское
патентное
ведомство**

(21) **202490990** (13) **A1**

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

(43) Дата публикации заявки
2024.07.25

(22) Дата подачи заявки
2022.10.18

(51) Int. Cl. *C12N 15/90* (2006.01)
C12N 5/10 (2006.01)
C12N 15/10 (2006.01)
C12N 15/11 (2006.01)

(54) **КЛЕТКИ МЛЕКОПИТАЮЩИХ, СОДЕРЖАЩИЕ ИНТЕГРИРОВАННЫЕ ГЕНЫ CAS9 ДЛЯ ПРОИЗВОДСТВА СТАБИЛЬНЫХ САЙТОВ ИНТЕГРАЦИИ, И КЛЕТКИ МЛЕКОПИТАЮЩИХ, СОДЕРЖАЩИЕ СТАБИЛЬНЫЕ САЙТЫ ИНТЕГРАЦИИ И ДРУГИЕ САЙТЫ**

(31) **63/256,675**

(32) **2021.10.18**

(33) **US**

(86) **PCT/US2022/078275**

(87) **WO 2023/069931 2023.04.27**

(71) Заявитель:

**РИДЖЕНЕРОН
ФАРМАСЬЮТИКАЛЗ, ИНК. (US)**

(72) Изобретатель:

**Горен Майкл, Бураков Дарья, Чэнь
Ган, Чжао Юй, Дешпанде Дипали (US)**

(74) Представитель:

Медведев В.Н. (RU)

(57) Данные изобретения относятся к клеткам млекопитающих, которые содержат множество стабильных сайтов интеграции. Данные изобретения предлагают сайты, геномно внедренные в геномную безопасную гавань и геномно внедренные за пределы этой конкретной геномной безопасной гавани, включая, помимо прочего, другую геномную безопасную гавань. Представляющие интерес полинуклеотиды, которые кодируют представляющие интерес полипептиды или РНК, могут быть вставлены в стабильные сайты интеграции, предусмотренные согласно данным изобретениям. Клетки и способы по данным изобретениям могут применяться для получения с высоким выходом любого белка, включая вирусные белки. Кроме того, клетки и способы по данным изобретениям применимы для производства вирусных векторов, таких как AAV, антител и других белков.

A1

202490990

202490990

A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-581094EA/061

КЛЕТКИ МЛЕКОПИТАЮЩИХ, СОДЕРЖАЩИЕ ИНТЕГРИРОВАННЫЕ ГЕНЫ CAS9 ДЛЯ ПРОИЗВОДСТВА СТАБИЛЬНЫХ САЙТОВ ИНТЕГРАЦИИ, И КЛЕТКИ МЛЕКОПИТАЮЩИХ, СОДЕРЖАЩИЕ СТАБИЛЬНЫЕ САЙТЫ ИНТЕГРАЦИИ И ДРУГИЕ САЙТЫ

Данная заявка испрашивает приоритет по отношению к заявке США с серийным номером 63/256675, поданной 18 октября 2021 г., которая включена в данный документ посредством ссылки во всей своей полноте.

ОБЛАСТЬ ИЗОБРЕТЕНИЯ

[0001] Данное изобретение относится к клеткам млекопитающих (включая клеточные линии), включая клетки человека и грызунов (включая клеточные линии), которые содержат множество стабильных сайтов интеграции (SIS), которые можно получить с использованием интегрированных генов Cas9. Данные изобретения предлагают стабильные сайты интеграции (1), геномно внедренные в геномную безопасную гавань (GSH), например, AAVS1 (сайт интеграции адено-ассоциированного вируса 1) и подобные AAVS1, и (2) геномно внедренные за пределы этой конкретной геномной безопасной гавани, например, другой геномной безопасной гавани или другого сайта, который не является геномной безопасной гаванью. Представляющие интерес полидезоксирибонуклеотиды, которые кодируют представляющие интерес полипептиды или РНК, могут быть вставлены в стабильные сайты интеграции, предложенные данным изобретением.

ССЫЛКА НА ЭЛЕКТРОННЫЙ ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

[0002] Данная заявка содержит перечень последовательностей, представленный в электронном виде в формате XML и настоящим включенный в данный документ посредством ссылки в полном объеме. Указанная копия XML, созданная 7 октября 2022 года, называется «135975-97402.xml» и имеет размер 709.205 байт. Перечень последовательностей, содержащийся в этом файле XML, является частью спецификации и включен в данный документ посредством ссылки во всей своей полноте.

УРОВЕНЬ ТЕХНИКИ

[0003] Линии клеток млекопитающих являются предпочтительным подходом для производства коммерческих количеств терапевтических белков, таких как антитела. Однако сообщалось, что модифицированные клетки млекопитающих часто демонстрируют снижение продукции из-за генетической и эпигенетической нестабильности. Hilliard and Lee, *Biotech. Bioeng.* 118: 659-75 (2021).

[0004] Интеграция полинуклеотидов является предпочтительным подходом для создания и поддержания трансформированных клеток. Интеграция конкретных последовательностей в AAVS1 человека обсуждается в Liu *et al.*, *BMC Research Note*, 7: 626 (2014) и Ramachandra *et al.*, *Nucl. Acids Res.* 39: e107 (2011). AAVS1 человека известен как геномная безопасная гавань. Parapetrou *et al.*, *Molecular Therapy* 24: 678-84 (2016).

Gaidukov *et al.*, *Nucl. Acids Res.* 46: 4072-86 (2018) раскрыли сайты для интеграции ДНК в посадочные сайты.

[0005] Клетки яичника китайского хомячка (СНО) и клетки почек детеныша хомячка (ВНК) используются в производстве терапевтических белков, а геномы хомяков тщательно изучались. Namaker и Lee сообщили о хромосомных локусах СНО как о потенциальных местах стабильной интеграции и назвали их «геномными горячими точками». *Curr. Op. Chem. Eng.* 22: 152-60 (2018) at 153. В Таблице 1 Namaker и Lee идентифицируют 30 локусов горячих точек, из которых 17 идентифицированы по генам и 13 не аннотированы. *Curr. Op. Chem. Eng.* 22: 152-60 (2018) at 154. За этой работой последовали Hilliard и Lee, которые стремились определить безопасные регионы в СНО с помощью анализа эпигенома. Hilliard and Lee, *Biotech. Bioeng.* 118: 659-75 (2021). Авторы установили, что 10,9% генома СНО содержат структуры хроматина с повышенной генетической и эпигенетической стабильностью. Авторы также определили, что из 30 горячих точек, определенных Namaker и Lee в Таблице 1, пять из них перекрываются со стабильными областями, определяемыми высокопроизводительным захватом конформации хромосом (Hi-C). Ближайшими к областям генами были *ALDH5A1*, *SMAD6* и *CLCN3*, а две другие области не были аннотированы. Hilliard and Lee, *Biotech. Bioeng.* 118: 659-75 (2021) в дополнительной Таблице 3 (S3). Gaidukov *et al.*, *Nucl. Acids Res.* 46: 4072-86 (2018) в Таблице 1 также идентифицирует локусы для интеграции в клетки СНО. Lee *et al.*, *Scientific Reps.* 5: 8572 (2015) идентифицирует локус *COSMC*.

[0006] В данных изобретениях преимущественно используется интегрированный ген Cas9 для эффективного создания промежуточных продуктов клеток млекопитающих, которые дополнительно модифицируются для получения клеток млекопитающих, имеющих множественные стабильные сайты интеграции для стабильной интеграции множества ДНК-кассет и других представляющих интерес полидезоксирибонуклеотидов. Согласно данным изобретениям, стабильный сайт интеграции может быть расположен в геномной безопасной гавани или других сайтах, включая недавно выявленные геномные безопасные гавани.

КРАТКОЕ ОПИСАНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

[0007] Данные изобретения относятся к клеткам млекопитающих, причем любая их клетка может содержать первый стабильный сайт интеграции, расположенный в геномной безопасной гавани, и второй стабильный сайт интеграции, который не расположен в геномной безопасной гавани, причем первый стабильный сайт интеграции содержит первый репортерный ген, кодирующий первый репортерный белок, и второй стабильный сайт интеграции содержит второй репортерный ген, кодирующий второй репортерный белок, при этом первый репортерный белок и второй репортерный белок различны. Первый и второй стабильные сайты интеграции могут содержать сайты узнавания для рекомбиназы (RRS). Первый и второй репортерные гены могут находиться под контролем промоторов SV40. Первый и второй репортерные гены могут представлять собой флуоресцентные белки. Клетки могут дополнительно содержать полинуклеотид,

кодирующий белок-репрессор под контролем промотора CMV. Клетки могут представлять собой клетки амниотического эпителия человека, НЕК 293, СНО или клетки ВНК. Полинуклеотид, кодирующий представляющий интерес белок, может быть вставлен в первый стабильный сайт интеграции или второй стабильный сайт интеграции. Второй стабильный сайт интеграции может быть расположен во второй геномной безопасной гавани, которая отличается от первой геномной безопасной гавани, или в области, которая не является геномной безопасной гаванью.

[0008] Данные изобретения также относятся к клеткам млекопитающих, причем любая их клетка может содержать первый стабильный сайт интеграции, расположенный в геномной безопасной гавани, и второй стабильный сайт интеграции, который не расположен в первой геномной безопасной гавани, отличающаяся тем, что первый стабильный сайт интеграции содержит первый полинуклеотид, кодирующий первый белок, и второй стабильный сайт интеграции содержит второй полинуклеотид, кодирующий второй белок. Первый и второй белки могут быть вирусными белками, такими как аденовирус-ассоциированный вирусный белок или аденовирусный белок. Например, клетки млекопитающих могут содержать полинуклеотид, кодирующий аденоассоциированный вирусный белок, и полинуклеотид, кодирующий аденовирусный белок. Другие полинуклеотиды, кодирующие белки, включают, помимо прочего, гены антител. Клетки могут иметь второй стабильный сайт интеграции, расположенный во второй геномной безопасной гавани, отличной от первой геномной безопасной гавани, в которой расположен первый стабильный сайт интеграции, или в области, которая не является геномной безопасной гаванью.

[0009] Данные изобретения также относятся к клеткам млекопитающих, причем любая их клетка может содержать первый стабильный сайт интеграции, расположенный в геномной безопасной гавани, и второй стабильный сайт интеграции, который не расположен в геномной безопасной гавани, отличающаяся тем, что первый стабильный сайт интеграции содержит полинуклеотид, кодирующий первый репортерный ген, кодирующий первый репортерный белок и второй стабильный сайт интеграции содержат полинуклеотид, кодирующий Cas9, и полинуклеотид, кодирующий второй репортерный ген, кодирующий второй репортерный белок, при этом первый репортерный белок и второй репортерный белок различны. Второй стабильный сайт интеграции может дополнительно содержать ген-маркер селекции и внутренний сайт входа в рибосому (IRES). Первый и второй стабильные сайты интеграции могут содержать сайты узнавания для рекомбиназы. Первый и второй репортерные гены могут находиться под контролем промоторов SV40. Первый и второй репортерные гены могут представлять собой флуоресцентные белки. Клетка может дополнительно содержать полинуклеотид, кодирующий репрессор (например, TetR) под контролем промотора (например, CMV). Клетка может представлять собой амниотическую эпителиальную клетку человека, НЕК293, СНО или клетку ВНК. Полинуклеотид, кодирующий представляющий интерес белок, может быть вставлен в первый стабильный сайт интеграции или второй

стабильный сайт интеграции. Белок-маркер селекции может придавать устойчивость к лекарственным средствам. Второй репортерный ген, ген-маркер селекции, IRES и промотор SV40 могут быть расположены на ДНК-кассете. Клетка может дополнительно содержать полинуклеотид, кодирующий белок-репрессор под контролем промотора (например, CMV). Второй стабильный сайт интеграции может быть расположенный во второй геномной безопасной гавани, отличной от первой геномной безопасной гавани, в которой расположен первый стабильный сайт интеграции, или в области, которая не является геномной безопасной гаванью. Первый репортерный ген может быть фланкирован 5'-концевым ответвлением геномной гомологичности безопасной гавани и 3'-концевым плечом гомологической геномной безопасной гавани. 5'-плечо гомологии геномной безопасной гавани может содержать целевой сайт sgPHK CRISPR, а 3'-плечо гомологии геномной безопасной гавани может содержать целевой сайт sgPHK CRISPR.

[0010] Данные изобретения также предлагают способы получения по меньшей мере одного представляющего интерес белка, причем любой из них может включать: (a) предоставление клеток млекопитающих, содержащих первый стабильный сайт интеграции, расположенный в геномной безопасной гавани, и второй стабильный сайт интеграции, который не расположен в первой геномной безопасной гавани, причем первый стабильный сайт интеграции содержит первый репортерный ген, кодирующий первый репортерный белок, а второй стабильный сайт интеграции содержит второй репортерный ген, кодирующий второй репортерный белок, причем первый репортерный белок и второй репортерный белок отличны, и при этом первый и второй стабильные сайты интеграции содержат сайты узнавания рекомбиназы; (b) введение полинуклеотида, кодирующего представляющий интерес белок, в стабильный сайт интеграции путем опосредованной рекомбиназой замены кассеты и (c) культивирование клетки млекопитающего в условиях, которые позволяют экспрессировать полинуклеотид, кодирующий представляющий интерес полинуклеотид. Первый и второй репортерные гены могут находиться под контролем промоторов SV40. Первый и второй репортерные гены могут представлять собой флуоресцентные белки. Клетка может дополнительно содержать полинуклеотид, кодирующий белок-репрессор под контролем промотора CMV. Клетка может представлять собой клетки амниотического эпителия человека, НЕК 293, СНО или клетки ВНК. Полинуклеотид, кодирующий представляющий интерес белок, может быть вставлен в первый стабильный сайт интеграции или второй стабильный сайт интеграции. Второй стабильный сайт интеграции может быть расположенный во второй геномной безопасной гавани, отличной от первой геномной безопасной гавани, в которой расположен первый стабильный сайт интеграции, или в области, которая не является геномной безопасной гаванью. Первый стабильный сайт интеграции содержит первый полинуклеотид, кодирующий первый белок, и второй стабильный сайт интеграции содержит второй полинуклеотид, кодирующий второй белок. Первый и второй белки могут быть вирусными белками, такими как аденовирус-ассоциированный вирусный белок или аденовирусный белок. Например, клетка млекопитающих может содержать

полинуклеотид, кодирующий аденоассоциированный вирусный белок, и полинуклеотид, кодирующий аденовирусный белок. Другие полинуклеотиды, кодирующие белки, включают, помимо прочего, гены антител. Второй стабильный сайт интеграции также может быть расположен в области, которая не является геномной безопасной гаванью.

[0011] В данных изобретениях также предложены способы создания клеток млекопитающих с множеством стабильных сайтов интеграции, причем любой из них может включать: (А) предоставление клетки млекопитающего, содержащей первую ДНК-кассету, содержащую в порядке от 5' до 3' полинуклеотид, кодирующий первый сайт lox, промотор, ген маркера селекции, кодирующий белок маркера селекции, IRES, первый репортерный ген, кодирующий первый репортерный белок, промотор, функционально связанный с оператором, ген Cas9 и второй lox-сайт; (В) интеграцию второй ДНК-касеты, содержащей в порядке от 5' к 3' полинуклеотид, содержащий первое плечо гомологии геномной безопасной гавани, содержащее целевой сайт sgPНК CRISPR, третий сайт lox, второй репортерный ген, кодирующий второй репортерный белок, четвертый сайт lox и второе плечо гомологии геномной безопасной гавани, содержащее целевой сайт sgPНК CRISPR, при этом первый сайт lox, второй сайт lox, третий сайт lox и четвертый сайт lox различны, при этом первое и второе направляющие плечи могут содержать область с изменениями (если необходимо, чтобы избежать воссоздания целевого сайта), в которой второй репортерный белок отличается от первого репортерного белка; (С) замену первой ДНК-касеты на третью ДНК-кассету, причем третья ДНК-кассета содержит в порядке от 5' к 3' полинуклеотид, кодирующий первый сайт lox, промотор, третий репортерный ген, кодирующий третий репортерный белок, и второй сайт lox, при этом третий репортерный белок отличается от второго репортерного белка, тем самым обеспечивая клетке млекопитающего множество стабильных сайтов интеграции. Клетки млекопитающих могут представлять собой амниотические эпителиальные клетки человека, клетки НЕК 293, клетки СНО или клетки ВНК. Репортерные гены для использования могут флуоресцировать белки. Клетка стадии (А) может дополнительно содержать полинуклеотид, кодирующий репрессор (например, TetR) под контролем промотора (например, CMV). Клетка стадии (В) может дополнительно содержать полинуклеотид, кодирующий репрессор (например, TetR) под контролем промотора (например, CMV). Клетка стадии (С) может дополнительно содержать полинуклеотид, кодирующий репрессор (например, TetR) под контролем промотора (например, CMV). Белок-маркер селекции может придавать устойчивость к лекарственным средствам. Lox-сайты являются наиболее часто используемым типом RRS; однако также можно использовать и другие RRS.

[0012] Данные изобретения также обеспечивают способы создания клетки млекопитающего с множеством сайтов обмена касет, опосредованных рекомбиназой, причем любой из этих способов может включать: (А) случайную интеграцию промотора и полинуклеотида, кодирующего репрессор, в геном клетки, при этом репрессор может связываться с лигандом; (В) случайную интеграцию в клеточный геном первой ДНК-

кассеты, содержащей в порядке от 5' к 3' полинуклеотид, кодирующий первый сайт lox, промотор и, необязательно, оператор, первый репортерный ген, кодирующий первый репортерный белок, IRES, первый ген маркера селекции, кодирующий первый белок-мейкер селекции и второй сайт lox, причем первый сайт lox и второй сайт lox различны; (C) замену первой ДНК-кассеты на вторую ДНК-кассету, при этом вторая ДНК-кассета содержит в порядке от 5' до 3' полинуклеотид, кодирующий первый сайт lox, промотор, второй ген маркера селекции, кодирующий белок второго маркера селекции, IRES, второй репортерный ген, кодирующий второй репортерный белок, промотор и необязательный оператор, ген Cas9 и второй сайт lox, причем первый и второй маркерные белки селекции различны, а первый и второй репортерные белки различны; (D) интеграция третьей ДНК-кассеты, содержащей в порядке от 5' до 3' полинуклеотид, содержащий первое плечо гомологии геномной безопасной гавани (GSH), содержащее целевой сайт sgPНК (единственная направляющая РНК), третий сайт lox, третий репортерный ген, кодирующий третий репортерный белок, четвертый сайт lox и второе плечо гомологии GSH, содержащее целевой сайт sgPНК, при этом первый сайт lox, второй сайт lox, третий сайт lox и четвертый сайт lox различны, при этом первый и второй направляющий сайт плечи могут содержать по меньшей мере одну область с изменениями (если это необходимо, чтобы избежать повторного создания целевого сайта), и при этом третий репортерный белок отличается от второго репортерного белка и может быть таким же или отличаться от первого репортерного белка; и (E) замену второй ДНК-кассеты на четвертую ДНК-кассету, причем четвертая ДНК-кассета содержит в порядке от 5' к 3' полинуклеотид, кодирующий первый сайт lox, промотор, четвертый репортерный ген, кодирующий четвертый репортерный белок, и второй сайт lox, при этом четвертый репортерный белок отличается от третьего репортерного белка и второго репортерного белка, и предпочтительно отличается от первого репортерного белка, тем самым обеспечивая клетке множество стабильных сайтов интеграции. Lox-сайты являются наиболее часто используемым типом RRS; однако также можно использовать и другие RRS.

[0013] Данные изобретения также относятся к клеткам млекопитающих, содержащим модифицированные геномы, причем данный геном модифицирован путем вставки по меньшей мере трех ДНК-кассет в различные области генома, при этом модифицированный геном содержит (1) первую последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 1 и 2 до модификации; (2) вторую последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной последовательности, выбранной из группы,

состоящей из SEQ ID NO: от 5 до 10, до модификации; и (3) третью последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации, при этом первая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки первой ДНК-кассеты, вторая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки второй ДНК-кассеты, и третья последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки третьей ДНК-кассеты. Каждая из клеток млекопитающих может иметь (а) первую ДНК-кассету, содержащую промотор и по меньшей мере один, выбранный из группы, состоящей из селективируемого маркерного гена и репортерного гена; (b) вторую ДНК-кассету, содержащую промотор и по меньшей мере один, выбранный из группы, состоящей из селективируемого маркерного гена и репортерного гена; и (с) третью ДНК-кассету, содержащую промотор и по меньшей мере один, выбранный из группы, состоящей из селективируемого маркерного гена и репортерного гена. Более того, каждая из клеток млекопитающих может иметь (а) первую ДНК-кассету, содержащую промотор, селективируемый маркерный ген и репортерный ген; (b) вторую ДНК-кассету, содержащую промотор, ген селективируемого маркера и ген-репортер; и (с) третью ДНК-кассету, содержащую промотор, селективируемый маркерный ген и репортерный ген. Первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции и представляющий интерес ген, встроенный в него. Представляющий интерес ген может кодировать представляющий интерес полипептид, выбранный из группы, состоящей из антител, цепей антител, рецепторов, Fc-содержащих белков, белков-ловушек, ферментов, факторов, репрессоров, активаторов, лигандов, репортерных белков, белков селекции, белковых гормонов, белковых токсинов, структурных белков, белков хранения, транспортных белков, нейротрансмиттеров и белков сокращения. Клетки млекопитающих могут представлять собой клетки человека, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична SEQ ID NO: 1. Альтернативно, клетка млекопитающего может представлять собой клетку СНО, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична SEQ ID NO: 2. Первая последовательность дезоксирибонуклеиновой кислоты может содержать стабильный сайт интеграции, полученный с использованием направляющей последовательности, выбранной из группы, состоящей из SEQ ID NO: от 13 до 419. Кроме того, первая последовательность дезоксирибонуклеиновой кислоты может содержать стабильный сайт интеграции, полученный с использованием направляющей последовательности, которая

комплементарна целевым последовательностям в SEQ ID NO:2 в диапазонах положений нуклеотидов, выбранных из группы, состоящей из: (a) от 1 до 2000; (b) от 2001 до 4000; (c) от 4001 до 6000; (d) от 6001 до 8000; (e) от 8001 до 10000; (f) от 10001 до 12000; (g) от 12001 до 14000; (h) от 14001 до 16000; (i) от 16001 до 18000; (j) от 18001 до 20000; (k) от 20001 до 22000; (l) от 22001 до 24000; (m) от 24001 до 26000; (n) от 26001 до 28000; (o) от 28001 до 30000; (p) от 30001 до 32000; (q) от 32001 до 34000; (r) от 34001 до 36000; (s) от 36001 до 38000; (t) от 38001 до 40000; (u) от 40001 до 42000; и (v) от 42001 до конечной точки (44232).

[0014] Кроме того, предложены клетки млекопитающих, содержащие модифицированные геномы, причем модифицированный геном содержит последовательность дезоксирибонуклеиновой кислоты, содержащую AAVS1-подобную область, модифицированную вставкой по меньшей мере одной ДНК-кассеты, и при этом направляющая последовательность выбрана из группы, состоящей из SEQ ID NO: от 13 до 419, которая связывается со смысловой или антисмысловой цепи AAVS1-подобной области и/или комплементарна ей. Клетки млекопитающих могут дополнительно содержать вторую последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и третью последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации, при этом первая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки первой ДНК-кассеты, вторая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки второй ДНК-кассеты, и третья последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки третьей ДНК-кассеты. Вторая последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной, выбранной из группы, состоящей из одной, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и третью последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации. Первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции, полученный с использованием направляющей последовательности, которая связывается и/или комплементарна по меньшей мере одной целевой последовательности,

которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична SEQ ID NO:2 в положениях нуклеотидов: (a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

[0015] Также предложены клетки млекопитающих, содержащие модифицированный геном, причем модифицированный геном содержит стабильный сайт интеграции в AAVS1-подобной области, причем стабильный сайт интеграции создается с использованием направляющей последовательности, которая связывается и/или комплементарна по меньшей мере одной целевой последовательности, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична SEQ ID NO:2 в положениях нуклеотидов: (a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

[0016] Кроме того, предложены клетки млекопитающих в соответствии с предыдущим абзацем, дополнительно содержащие вторую последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и третью последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации, при этом первая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки первой ДНК-кассеты, вторая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки второй ДНК-кассеты, и третья последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки третьей ДНК-кассеты. Клетка млекопитающего может иметь вторую последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-

99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной, выбранной из группы, состоящей из одной, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и третью последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 50%-99%, 75%-99%, 85%-99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации.

[0017] Кроме того, предложены способы получения представляющих интерес белков, причем способ включает стадии: (1) культивирование вышеуказанных клеток млекопитающих; и (2) сбор интересующего белка. Также предложены клетки, изготовленные любым из вышеуказанных способов, а также способы использования раскрытых клеток.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

[0018] На фигурах ниже показано типовое развитие и создание промежуточных клеток, полезных для создания клеток, имеющих стабильные сайты интеграции в различных областях генома, и последующее создание клеток, имеющих стабильные сайты интеграции в различных областях генома. Данные фигуры иллюстрируют варианты реализации изобретения и никоим образом не ограничивают данные изобретения.

[0019] На **Фиг. 1** схематически изображена модификация клетки, которая имеет полинуклеотид, кодирующий белок-репрессор, и сигнал полиаденилирования под транскрипционным контролем промотора, при этом полинуклеотид случайно вставлен в геном клетки.

[0020] На **Фиг. 2** схематически изображена модификация клетки, показанной на **Фиг. 1**, после того, как ДНК-кассета (1) случайным образом или сайт-специфически вставлена в геном клетки. ДНК-кассета (1) содержит фланкирующие lox-сайты (1 и 2), промотор, репортерный ген (1), IRES, ген маркера селекции (1) и сигнал полиаденилирования. Вместо lox-сайтов можно использовать и другие RRS.

[0021] На **Фиг. 3** схематически изображена модификация клетки, показанной на **Фиг. 2**, в которой ДНК-кассета (1) заменена в результате опосредованного рекомбиназой обмена кассеты на ДНК-кассету (2). ДНК-кассета (2) содержит фланкирующие сайты lox (1 и 2), промотор, ген маркера селекции (2), IRES и репортерный ген (2), и сигнал полиаденилирования, а также ген Cas9 со вторым сигналом полиаденилирования под контролем второй промотор (оператор не является обязательным). Вместо lox-сайтов можно использовать и другие RRS.

[0022] На **Фиг. 4** схематически изображена модификация клетки, показанной на **Фиг. 3**, которая имеет ДНК-кассету (3), содержащую фланкирующие плечи гомологии геномной безопасной гавани (GSH), сайты lox (3 и 4) и репортерный ген (3) с сигналом полиаденилирования под действием контроля промотора, введенного в геномную безопасную гавань. Вставка представляет собой интеграцию для конкретного сайта и

создает стабильный сайт интеграции между Lox3 и Lox4. Вместо lox-сайтов можно использовать и другие RRS.

[0023] На **Фиг. 5** схематически изображена модификация клетки, показанной на **Фиг. 4**, в которой ДНК-кассета (2) заменена в результате опосредованного рекомбиназой обмена кассеты на ДНК-кассету (4). ДНК-кассета (4) содержит фланкирующие lox-сайты (1 и 2), репортерный ген (4) и сигнал полиаденилирования под контролем промотора. Этот обмен удаляет ген Cas9. Вместо lox-сайтов можно использовать и другие RRS.

[0024] На **Фиг. 6** схематически изображена плазида sgPНК, использованная в Примере 6.

[0025] На **Фиг. 7** представлены графики Примера 6, показывающие популяции, положительные по зеленому флуоресцентному белку (Q1), для матрицы без HDR (контроль), 104-мерной матрицы HDR, 401-мерной матрицы HDR и 1030-мерной матрицы HDR. Положительный GFP представляет собой вертикальную ось, а положительный CFP представляет собой горизонтальную ось.

[0026] На **Фиг. 8** схематически изображена клетка млекопитающего (например, НЕК293) со стабильно интегрированным геном Cas9, фланкированным Lox-сайтами 3 и 4. Ген Cas9 находится под контролем по меньшей мере промотора (не показан). AAVS1 также схематически изображен. Вместо lox-сайтов можно использовать и другие RRS. Промоторы присутствуют на 5'-конце генов, но они не изображены.

[0027] На **Фиг. 9А** и **Фиг. 9В** схематически изображены нацеливающие плазмиды, содержащие целевой сайт sgPНК, левое плечо гомологии (в данном документе плечо гомологии GSH) для вставки в область, такую как геномная безопасная гавань (в данном документе AAVS1), сайт Lox 1, репортерный ген (цвет 1), сайт Lox 2, правое плечо гомологии (в данном документе плечо гомологии GSH) для вставки в область, такую как геномная безопасная гавань (в данном документе AAVS1). На 3'-конце на **Фиг. 9А** схематически изображен репортерный ген (цвет 2), а на **Фиг. 9В** схематически изображен на 3'-конце ген отрицательной селекции (негативный отбор 1). Промоторы и, необязательно, другие фрагменты (такие как операторы) представлены стрелками, направленными в направлении от 5' к 3'. Обе плазмиды вставляют цвет 1 в область, такую как геномная безопасная гавань. Вместо lox-сайтов можно использовать и другие RRS.

[0028] На **Фиг. 10** схематически показаны результаты после опосредованной Cas9 интеграции в геномную безопасную гавань (AAVS1) клетки млекопитающих (например, НЕК293). Цвет 1 окружен Lox 1 и Lox 2. Интересующий ген может заменить цвет 1 посредством RMCE. Когда нацеливающая плазида согласно **Фиг. 9А** правильно интегрирована, клетка будет цвет 1 положительной и цвет 2 отрицательной. Когда нацеливающая плазида согласно **Фиг. 9В** правильно интегрирована, клетка будет цвет 1-положительной и сможет размножаться, поскольку отрицательный ген селекции удален. Вместо lox-сайтов можно использовать и другие RRS. Промоторы и, необязательно, другие фрагменты (такие как операторы) представлены стрелками, направленными в направлении от 5' к 3'.

[0029] На **Фиг. 11** схематически более подробно изображена вставка **Фиг. 10**. Клеточный геном, включая AAVS1, окружает вставку и 5'- и 3'-концы. Цвет 1 окружен Lox 1 и Lox 2. На **Фиг. 11** слева показано расположение 5'-праймера генома и 3'-праймера для вставки, используемых в ПЦР с 5'-соединением. На **Фиг. 11** справа показано расположение 5'-инсерционного праймера и 3'-геномного праймера, используемых в ПЦР с 3'-соединением. Вместо lox-сайтов можно использовать и другие RRS. 5'-промотор гена цвета 1 обозначен 5'-стрелкой.

[0030] На **Фиг. 12** показано, что фрагменты правильного размера амплифицируются в клетках НЕК 293 с помощью ПЦР соединения, схематически изображенной на **Фиг. 11**. Получают и обнаруживают стабильные клетки НЕК293, нацеленные на Cas9, а также 5'- и 3'-соединение, что обеспечивает правильную вставку.

[0031] На **Фиг. 13** показано, что фрагменты правильного размера амплифицируются в клетках СНО с помощью ПЦР соединения, схематически изображенной на **Фиг. 11**. Получают и обнаруживают стабильные клетки СНО, нацеленные на Cas9, а также 5'- и 3'-соединение, что обеспечивает правильную вставку. Вместо lox-сайтов можно использовать и другие RRS.

[0032] На **Фиг. 14** схематически изображена типовая клетка, содержащая три кассеты, интегрированные в области генома с фланкирующими RRS (в данном документе lox 1 и lox 2). В зависимости от типа клеток каждая из трех кассет может быть интегрирована в разные стабильные сайты интеграции (например, AAVS1-подобные), схематически изображенные в позиции А, и другие доступные сайты (такие как стабильный сайт 1 и стабильный сайт 2), схематически изображенные в положениях В и С. Репортерные гены могут быть одинаковыми или разными. Гены негативной селекции могут быть одинаковыми или разными, но предпочтительно одинаковыми. Клетка может содержать дополнительные стабильные сайты интеграции и интегрированные кассеты в соответствии с идеями, содержащимися в данном документе. Промоторы присутствуют на 5'-конце генов, но они не изображены.

[0033] На **Фиг. 15** схематически изображена модификация клетки, показанной на **Фиг. 14**, в схематически изображенных положениях А, В и С. Каждая из трех кассет содержит фланкирующие RRS (в данном документе lox 1 и lox 2), представляющий интерес ген, ген-маркер положительной селекции и репортерный* ген. Гены-маркеры положительной селекции могут быть одинаковыми или разными, но предпочтительно одинаковыми. Гены-репортеры* могут быть одинаковыми или разными, но каждый из них должен отличаться от любого из генов-репортеров в клетке, показанной на **Фиг. 14**. Гены, представляющие интерес, могут быть одинаковыми или разными. Кассеты на **Фиг. 14** заменены кассетами, показанными на **Фиг. 15**, с помощью RMCE. Клетка может содержать дополнительные стабильные сайты интеграции и интегрированные кассеты в соответствии с идеями, содержащимися в данном документе. Промоторы присутствуют на 5'-конце генов, но они не изображены.

[0034] На **Фиг. 16** представлена гистограмма, сравнивающая белок,

продуцируемой клеткой CHO-K1 с тремя сайтами (A, B и C), по сравнению с клеткой CHO-K1 с двумя сайтами (B и C).

[0035] На **Фиг. 17** схематически изображена типовая клетка, содержащая четыре кассеты, интегрированные в области генома с фланкирующими RRS (в данном документе lox 1 и lox 2 или lox 3 и lox 4). В зависимости от типа клеток каждая из четырех кассет может быть интегрирована в разные стабильные сайты интеграции и другие доступные сайты (такие как стабильный сайт 1 и стабильный сайт 2), и схематически изображенные как позиции A и B (SIS) и C и D (стабильные сайты 1 и 2). Репортерные гены могут быть одинаковыми или разными. Гены негативной селекции могут быть одинаковыми или разными, но предпочтительно одинаковыми. Клетка может содержать дополнительные стабильные сайты интеграции и интегрированные кассеты в соответствии с идеями, содержащимися в данном документе. Промоторы присутствуют на 5'-конце генов, но они не изображены.

[0036] На **Фиг. 18** схематически изображена модификация клетки, показанной на **Фиг. 17**, в схематически изображенных положениях A, B, C и D. Каждая из четырех кассет содержит фланкирующие RRS (в данном документе lox 1 и lox 2, или lox 3 и lox 4), представляющий интерес ген, ген-маркер положительной селекции и ген-репортер*. Гены-маркеры положительной селекции могут быть одинаковыми или разными, но предпочтительно одинаковыми. Гены-репортеры* могут быть одинаковыми или разными, но каждый из них должен отличаться от любого из генов-репортеров в клетке, показанной на **Фиг. 17**. Гены, представляющие интерес, могут быть одинаковыми или разными. На данной фигуре показаны две копии интересующего гена 1 и две копии интересующего гена 2. Кассеты на **Фиг. 17** заменены кассетами, показанными на **Фиг. 18**, с помощью RMCE. Клетка может содержать дополнительные стабильные сайты интеграции и интегрированные кассеты в соответствии с идеями, содержащимися в данном документе. Промоторы присутствуют на 5'-конце генов, но они не изображены.

ПОДРОБНОЕ ОПИСАНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

Определения

[0037] Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют те же значения, которые обычно понятны специалисту в области техники, к которой относится это изобретение.

[0038] Термин «около» в контексте числовых значений и диапазонов относится к значениям или диапазонам, которые приближаются или близки к указанным значениям или диапазонам, так что данное изобретение может работать, например, иметь искомую скорость, количество, степень, увеличение, уменьшение или степень выражения, концентрация или время, как очевидно из идей, содержащихся в данном документе. Таким образом, этот термин охватывает значения, выходящие за рамки тех, которые просто возникают в результате систематической ошибки. Например, «около» может обозначать значения выше или ниже указанного значения в диапазоне прикл. +/- 10% или больше или меньше в зависимости от работоспособности.

[0039] «AAVS1» может быть геномной безопасной гаванью и относится к сайту 1 интеграции аденоассоциированного вируса, и, как сообщается, в природе он расположен на хромосоме 19 человека и содержит около 4,7 тыс. нуклеотидов. Локус AAVS1 можно использовать согласно данным изобретениям.

[0040] «AAVS1-подобный» относится к гомологу AAVS1, обнаруженному в клетках CHO, и раскрыт в данном документе. AAVS1-подобная область, содержащая AAVS1-подобную геномную безопасную гавань (GSH), может быть использована согласно данным изобретениям. SEQ ID NO:2 представляет собой пример AAVS1-подобной области.

[0041] «ДНК-кассета» или «кассета» представляет собой тип фрагмента нуклеиновой кислоты, который содержит по меньшей мере промотор, по меньшей мере одну открытую рамку считывания и необязательно сигнал полиаденилирования, например, сигнал полиаденилирования SV40. Другие фрагменты нуклеиновой кислоты, такие как операторы, также являются необязательными. Таким образом, кассета ДНК представляет собой полинуклеотид, который содержит два или более коротких полинуклеотидов. Кассета может содержать один или более генов и промоторов, энхансеров, операторов, репрессоров, сигналов терминации транскрипции, сайтов входа в рибосомы, интронов и сигналов полиаденилирования.

[0042] Сообщается, что «COSMC» был обнаружен в клетках хомяка. Гомологи частичного или полного локуса COSMC являются кандидатами для применения согласно данному изобретению.

[0043] «CCR5» относится к гену хемокинового рецептора CC типа 5 и, как сообщается, был обнаружен в клетках человека, мыши и крысы. Гомологи частичного или полного локуса CCR5 являются кандидатами для применения согласно данному изобретению.

[0044] «Геномные безопасные гавани» или «GSH» относятся к участкам в клеточном геноме, которые могут вмещать вставки полинуклеотидов, такие как ДНК-касеты, и позволяют вставленному полинуклеотиду функционировать, не создавая чрезмерной нагрузки на трансформированную клетку. Соответственно, «геномные безопасные гавани» являются идеальными местами для создания стабильных площадок интеграции для вставки ДНК-кассет посредством применения данных изобретений. Геномные безопасные гавани, которые могут применяться в данном документе, включают, помимо прочего, AAVS1 и AAVS1-подобные. Сообщаемые локусы-кандидаты включают, помимо прочего, CCR5, COSMC и Rosa26.

[0045] «Плечо гомологии геномной безопасной гавани» или «плечо гомологии GSH» получено из геномной безопасной гавани и имеет гомологию с геномной безопасной гаванью. Предпочтительно, плечо гомологии геномной безопасной гавани содержит от около 100 до 2000 оснований, более предпочтительно от около 300 до 1800 оснований, более предпочтительно от около 400 до 1600 оснований, более предпочтительно от около 500 до 1500 оснований, более предпочтительно от около 500 до

1300 оснований, более предпочтительно от около 500 до 1100 оснований, более предпочтительно от около 500 до 1000 оснований, более предпочтительно от около 600 до 1000 оснований, более предпочтительно от около 700 до 1000 оснований, более предпочтительно от около 800 до 1000 оснований и еще более предпочтительно от около 900 до 1000 оснований. Обычно полинуклеотид, который будет вставлен в геномную безопасную гавань, будет окружен 5'-плечом гомологии GSH и 3'-плечом гомологии GSH. Например, см. Фиг. 4 и 5, показывающие ДНК-кассету, фланкированную lox-сайтом, которая дополнительно фланкирована плечом гомологии GSH.

[0046] «hRosa26» относится к человеческому гомологу мышинового локуса Rosa26 («акцептор сплайсинга обратной ориентации»). «Rosa26» относится к частичному или полному локусу Rosa26 и, как сообщается, был обнаружен в клетках хомяка в дополнение к клеткам мыши и человека. Гомологи частичного или полного локуса Rosa26 являются кандидатами для применения согласно данному изобретению.

[0047] «Интрон» представляет собой участок ДНК, расположенный между экзонами. Интрон удаляется с образованием зрелой информационной РНК. Предпочтительными интронами являются те, которые могут влиять на начальную точку трансляции, и примерами являются интрон hCMV-IE (белок немедленной ранней стадии цитомегаловируса человека) и интрон ящура (вирус ящура).

[0048] «Фрагмент нуклеиновой кислоты» включает любое расположение одноцепочечных или двухцепочечных нуклеотидных последовательностей. Фрагменты нуклеиновой кислоты могут включать, помимо прочего, полинуклеотиды, промоторы, энхансеры, операторы, репрессоры, сигналы терминации транскрипции, сайты входа в рибосомы и сигналы полиаденилирования.

[0049] «Функционально связанный» относится к одной или нескольким нуклеотидным последовательностям, находящимся в функциональных отношениях с одной или более другими нуклеотидными последовательностями. Такие функциональные отношения могут прямо или косвенно контролировать, вызывать, регулировать, усиливать, облегчать, разрешать, ослаблять, подавлять или блокировать действие, или деятельность в соответствии с выбранным дизайном. Примеры включают фрагменты одноцепочечной или двухцепочечной нуклеиновой кислоты и могут содержать две или более нуклеотидные последовательности, расположенные внутри данного фрагмента таким образом, что последовательность(и) может оказывать по меньшей мере одно функциональное воздействие на другой(ие) фрагмент(ы). Например, промотор, функционально связанный с кодирующей областью полинуклеотидной последовательности ДНК, может способствовать транскрипции кодирующей области. Другие элементы, такие как энхансеры, операторы, репрессоры, сигналы терминации транскрипции, сайты входа в рибосомы и сигналы полиаденилирования, также могут быть функционально связаны с представляющим интерес полинуклеотидом для контроля его экспрессии. Расположение и расстояние для достижения работоспособных связей можно определить с помощью подходов, доступных специалисту в данной области техники,

таких как скрининг с использованием вестерн-блоттинга и RT-PCR.

[0050] «Оператор» обозначает последовательность ДНК, которая введена в полинуклеотидную последовательность или рядом с ней таким образом, что полинуклеотидная последовательность может регулироваться взаимодействием молекулы, способной связываться с оператором и, в результате, предотвращать или обеспечивать транскрипцию полинуклеотидную последовательность, в зависимости от обстоятельств. Специалист в данной области техники поймет, что оператор должен располагаться достаточно близко к промотору, чтобы он был способен контролировать или влиять на транскрипцию с помощью промотора, что можно рассматривать как тип работоспособной связи. Оператор может быть расположен либо ниже, либо выше промотора. К ним относятся, помимо прочего, операторная область гена *Lex A E. coli*, которая связывает пептид *Lex A* и лактозу, и 45 операторов триптофана, которые связывают репрессорные белки, кодируемые генами *Lad* и *trpR E. coli*. Бактериофаги-операторы из лямбда *Pi* и фага *P22 Mnt*, и *Arc*. Предпочтительными операторами являются оператор *Tet* (тетрациклин) (*TetO* или *TO*) и оператор *Arc* (*ArcO* или *AO*). Операторы могут иметь нативную последовательность или мутантную последовательность. Например, мутантные последовательности оператора *Tet* раскрыты в *Wissmann et al., Nucleic Acids Res.* 14: 4253-4266 (1986).

[0051] Оператор *Tet* является предпочтительным и может использоваться для контроля транскрипции с помощью репрессора, такого как репрессор тетрациклина (*TetR*). Подходящими лигандами для репрессора являются тетрациклин (*tet*), доксициклин (*dox*) и их производные. Когда лиганд связывается с *TetR*, сродство репрессора *Tet* к оператору *Tet* уменьшается, и репрессор *Tet* отделяется от оператора, и, таким образом, оператор становится пермиссивным для транскрипции. Другие репрессоры могут быть объединены в пару для использования со своими соответствующими операторами.

[0052] Фразы «процент идентичности» или «% идентичности» в их различных грамматических формах при описании последовательности подразумевают включение гомологичных последовательностей, которые демонстрируют заявленную идентичность вдоль областей смежной гомологии, но наличие пробелов, делеций или вставок, которые не имеющие гомолога в сравниваемой последовательности, не учитываются при расчете процента идентичности. В данном документе определение «процента идентичности» или «% идентичности» между гомологами не будет включать сравнение последовательностей, когда гомолог не имеет гомологичной последовательности для сравнения при выравнивании. Таким образом, «процент идентичности» и «% идентичности» не включают штрафы за пропуски, удаления и вставки.

[0053] «Гомологичная последовательность» в ее различных грамматических формах в контексте последовательностей нуклеиновой кислоты относится к последовательности, которая по существу гомологична эталонной последовательности нуклеиновой кислоты. В некоторых вариантах реализации две последовательности считаются по существу гомологичными, если по меньшей мере 50%-99%, 75%-99%, 85%-

99%, 90%-99%, 95%-98%, 98%-99%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более соответствующих им нуклеотидов идентичны на соответствующем участке остатков. В некоторых вариантах реализации соответствующий участок представляет собой полную (т.е. полную) последовательность.

[0054] «Полинуклеотид» включает последовательность ковалентно соединенных нуклеотидов и включает РНК, и ДНК. Олигонуклеотиды считаются более короткими полинуклеотидами. Гены представляют собой полинуклеотиды ДНК (полидезоксирибонуклеиновая кислота), которые в конечном итоге кодируют полипептиды, которые транслируются с РНК (полирибонуклеиновой кислоты), которая обычно транскрибируется с ДНК. Полинуклеотиды ДНК также могут кодировать полинуклеотиды РНК, которые не транслируются, а скорее функционируют как «продукты» РНК. Тип полинуклеотида (то есть ДНК или РНК) очевиден из контекста использования этого термина. Полинуклеотид, упомянутый или идентифицированный полипептидом, который он кодирует, представляет собой и охватывает все подходящие последовательности в соответствии с вырожденностью кодонов. Полинуклеотиды, включая те, которые раскрыты в настоящем документе, включают последовательности с процентной идентичностью и гомологичные последовательности, если это указано.

[0055] «Полипептид» и «пептид» относится к последовательности(ям) ковалентно соединенных аминокислот. Полипептиды включают природные, полусинтетические и синтетические белки и белковые фрагменты. «Полипептид» и «белок» могут использоваться взаимозаменяемо. Олигопептиды считаются более короткими полипептидами.

[0056] «Промотор» обозначает последовательность ДНК, которая вызывает транскрипцию последовательности ДНК, с которой она функционально связана, т.е. связана таким образом, чтобы обеспечить транскрипцию интересующей нуклеотидной последовательности, когда присутствуют соответствующие сигналы и отсутствуют репрессоры. Экспрессию интересующего полинуклеотида можно поставить под контроль любого промотора или энхансерного элемента, известного в данной области техники. Эукариотический промотор может быть функционально связан с ТАТА-боксом. ТАТА-бокс обычно расположен выше места начала транскрипции.

[0057] Полезные промоторы, которые можно использовать, включают, помимо прочего, область раннего промотора SV40, промотор SV40 E/L (ранний поздний), промотор, содержащийся в длинном 3'-концевом повторе вируса саркомы Рауса, регуляторные последовательности ген металлотioneина, главный промотор цитомегаловируса мыши или человека с немедленным ранним развитием (CMV-MIE) и другие промоторы CMV, включая промоторы CMVmin. Векторы экспрессии растений, содержащие промоторную область нопалинсинтетазы, промотор 35S РНК вируса мозаики цветной капусты и промотор фотосинтетического фермента рибулозобифосфаткарбоксилазы; промоторные элементы дрожжей или других грибов, такие как промотор Gal 4, промотор ADC (алкогольдегидрогеназы), промотор PGK

(фосфоглицеринкиназы), промотор щелочной фосфатазы и следующие области контроля транскрипции животных, которые проявляют тканеспецифичность и используются в трансгенные животные: эластаза I; инсулин; иммуноглобулин; вирус опухоли молочной железы мышей; альбумин; С.-фетопротеин; С.1-антитрипсин; 3-глобин и легкая цепь миозина-2. Согласно данным изобретениям можно использовать различные формы промотора CMV.

[0058] Минимальные промоторы, такие как промоторы CMVmin, могут быть усеченными промоторами или коровыми промоторами, и являются предпочтительными для использования в системах контролируемой экспрессии. Минимальные промоторы и подходы к разработке широко известны и раскрыты, например, в Saxena *et al.*, *Methods Molec. Biol.* 1651:263-73 (2017); Ede *et al.*, *ACS Synth Biol.* 5:395-404 (2016); Brown *et al.*, *Biotech Bioeng.* 111:1638-47 (2014); Morita *et al.*, *Biotechniques* 0:1-5 (2012); Lagrange *et al.*, *Genes Dev.* 12:34-44 (1998). В данной области техники описано множество промоторов CMVmin.

[0059] «Интересующий белок» или «интересующий полипептид» могут иметь любую аминокислотную последовательность и включают любой белок, полипептид или пептид, а также их производные, компоненты, домены, цепи и фрагменты. Включены, помимо прочего, вирусные белки, бактериальные белки, грибковые белки, растительные белки и животные (включая белки человека) белки. Типы белков могут включать, помимо прочего, антитела, биспецифические антитела, мультиспецифические антитела, цепи антител (включая тяжелые и легкие), фрагменты антител, фрагменты Fv, фрагменты Fc, Fc-содержащие белки, слитые с Fc белки, рецепторные Fc-слитые белки, рецепторы, рецепторные домены, белки-ловушки и мини-ловушки, ферменты, факторы, репрессоры, активаторы, лиганды, репортерные белки, селективные белки, белковые гормоны, белковые токсины, структурные белки, запасные белки, транспортные белки, нейротрансмиттеры и сократительные белки. Производные, компоненты, цепи и фрагменты вышеперечисленного также включены. Последовательности могут быть природными, полусинтетическими или синтетическими. Представляющие интерес белки и интересующие полипептиды кодируются «представляющими интерес генами», которые также можно называть «представляющими интерес полинуклеотидами». Если интегрировано несколько генов (одинаковых или разных), их можно называть «первым», «вторым», «третьим», «четвертым», «пятым», «шестым», «седьмым», «восьмым», «девятым», «десятым» и т. д., как видно из контекста применения.

[0060] «Сайты узнавания рекомбиназы» (RRS), также известные как «сайты гетероспецифической рекомбинации», используются при опосредованной рекомбиназой замены кассеты (RMCE). Подходящими системами RRS являются, например, Cre/Lox, Dre/Rox, Vre/Vlox, SCre/Slox и Flp/Frt. Подходящие RRS для использования согласно данным изобретениям включают Lox P, Lox 66, Lox 71, Lox 511, Lox 2272, Lox 2372, Lox 5171, Lox M2, Lox M3, Lox M7 и Lox M11. Эти сайты можно обозначать в общем как первый (1), второй (2), третий (3), четвертый (4), пятый (5), шестой (6), седьмой (7),

восьмой (8), девятый (9), десятый (10) и т. д., как видно из контекста употребления. Cte/Lox является наиболее часто используемым RRS, но вместо Cte/Lox согласно данным изобретениям, можно использовать и другие RRS.

[0061] Термин «репортерные белки», в контексте данного документа, относится к любому белку, способному прямо или косвенно генерировать обнаружимый сигнал. Репортерные белки обычно флуоресцируют или катализируют колориметрическую, или флуоресцентную реакцию, и их часто называют «флуоресцентными белками» или «цветными белками». Однако репортерный белок также может быть неферментативным и нефлуоресцентным, если его можно обнаружить с помощью другого белка или фрагмента, такого как белок клеточной поверхности, обнаруженный с помощью флуоресцентного лиганда. Репортерный белок также может представлять собой неактивный белок, который становится функциональным за счет взаимодействия с другим белком, который является флуоресцентным или катализирует реакцию. Соответственно, можно использовать любой подходящий репортерный белок, как это понимает специалист в данной области техники. В некоторых аспектах репортерный белок может быть выбран из флуоресцентного белка, люциферазы, щелочной фосфатазы, β -галактозидазы, β -лактамазы, дигидрофолатредуктазы, убиквитина и их вариантов. Флуоресцентные белки полезны для распознавания кассет генов, которые были или не были успешно вставлены и/или заменены, в зависимости от обстоятельств. Для обнаружения подходят жидкостная цитометрия и сортировка клеток, активируемая флуоресценцией. Примеры флуоресцентных белков хорошо известны в данной области техники, включая, помимо прочего, коралл *Discosoma* (DsRed), зеленый флуоресцентный белок (GFP), усиленный зеленый флуоресцентный белок (eGFP), цианофлуоресцентный белок (CFP), усиленный цианофлуоресцентный белок (eCFP), желтый флуоресцентный белок (YFP), усиленный желтый флуоресцентный белок (eYFP) и дальнекрасный флуоресцентный белок (например, mKate, mKate2, mPlum, mRaspberry или E2-crimson. См., например, патенты США № 9816110. Репортерные белки кодируются полинуклеотидами и называются в данном документе «генами-репортерами» или «генами репортерных белков». Репортерные гены и белки можно обозначать в общем как первый (1), второй (2), третий (3), четвертый (4), пятый (5), шестой (6), седьмой (7), восьмой (8), девятый (9), десятый (10) и т. д., как видно из контекста употребления. Репортеров можно рассматривать как своего рода маркер. «Цвет» или «флуоресцентный» в их различных грамматических формах также могут использоваться, более конкретно, для обозначения репортерного белка или гена.

[0062] «Белок-репрессор», также называемый «репрессор», представляет собой белок, который может связываться с ДНК для подавления транскрипции и кодируется полинуклеотидом, также называемым в данном документе «геном-репрессором» или «ген белка-репрессора». Репрессоры имеют эукариотическое и прокариотическое происхождение. Прокариотические репрессоры являются предпочтительными. Примеры семейств репрессоров включают: Семейства TetR, LysR, LacI, ArsR, IclR, MerR, AsnC,

MarR, DeoR, GntR и Csp. Белки-репрессоры семейства TetR включают: ArcR, ActII, AmeR, AmrR, ArpR, BpeR, EnvR, EthR, HemR, HydR, IfeR, LanK, LfrR, LmrA, MtrR, Pip, PqrA, QacR, RifQ, RmrR, SimReg2, SmeT, SrpR, TcmR, TetR, TtgR, TrgW, UrdK, VarR YdeS, ArpA., BarA, Aur1B, CalR, CprB, FarA, JadR*, JadR2, MphB, NonG, PhIF, TylQ, VanT, TarA, TylP, BM1P1, Bm3R1, ButR, CampR, CamR, DhaR, KstR, LexA-like, AcnR, PaaRR, PslI, Th1R, UidR, YDH1, BetI, McbR, MphR, PhaD, Q9ZF45, TtK, Yhgd, YixD, CasR, IcaR, LitR, LuxR, LuxT, OpaR, Orf2, SmcR, NapR, Ef0113, HlyIIR, BarB, ScbR, MmfR, AmtR, PsrA и YjdC белки См. Ramos *et al.*, *Microbiol. Mol. Biol. Rev.*, 69: 326-56 (2005). Другие репрессоры включают PurR, LacR, MetJ и PadR.

[0063] «Селектируемые» или «селективные» маркерные белки включают белки, придающие определенные признаки, включая, помимо прочего, устойчивость к лекарственным средствам или другие селективные преимущества. Селективные маркеры могут придать клетке, получающей ген селектируемого маркера, устойчивость к определенному токсину, лекарству, антибиотику или другому соединению и позволить клетке продуцировать белок и размножаться в присутствии токсина, лекарства, антибиотика или другого соединения, и их часто называют «положительными селектируемыми маркерами». Подходящие примеры маркеров устойчивости к антибиотикам включают, помимо прочего, белки, придающие устойчивость к различным антибиотикам, такие как канамицин, спектиномицин, неомицин, гентамицин (G418), ампициллин, тетрациклин, хлорамфеникол, пурамицин, гигромицин, зеоцин и/или бластицидин. Существуют и другие селектируемые маркеры, часто называемые «негативными селектируемыми маркерами», которые заставляют клетку прекращать размножение, останавливать выработку белка и/или являются летальными для клетки в присутствии отрицательных селектируемых маркерных белков. Тимидинкиназа и некоторые слитые белки могут служить отрицательными селектируемыми маркерами, включая, помимо прочего, GyrB-PKR. См. White *et al.*, *Biotechniques*, 50: 303-309 (May 2011). Селектируемые маркерные белки и соответствующие гены (гены селектируемых маркеров) в общем можно обозначать как первый (1), второй (2), третий (3), четвертый (4), пятый (5), шестой (6), седьмой (7), восьмой (8), девятый (9), десятый (10) и т. д., как видно из контекста употребления. На фигурах селектируемые маркеры являются положительными селектируемыми маркерами, если не указано иное как отрицательный (отриц.) маркер.

[0064] «Одна направляющая РНК» или «sgРНК» используется для нацеливания Cas9 на сайт и обычно имеет длину 17-24 нуклеотида.

[0065] «Стабильный сайт интеграции» или «SIS» представляет собой область сайт-специфической интеграции представляющих интерес полинуклеотидов ДНК, включая кассеты, которые содержат гены и/или другие открытые рамки считывания, промоторы и, необязательно, другие элементы. Стабильные сайты интеграции содержат ДНК-кассету экзогенного происхождения и могут быть созданы в соответствии со способами данных изобретений, описанными и изображенными в данном документе, предпочтительно в

GSH. Конструкции могут быть вставлены в SIS различными способами. Множественные сайты стабильной интеграции могут быть созданы и расположены на разных хромосомах, в разных областях одной и той же хромосомы или в разных положениях в одной и той же области хромосомы.

[0066] «Элемент ответа на тетрациклин» или «TRE» содержит семь копий 19-нуклеотидного TetO, разделенных спейсерами, содержащими 17-18 нуклеотидов, и коммерчески доступны. Последовательности TetO могут варьироваться, и известны нуклеотидные замены. Например, измененные последовательности на основе оператора Tet раскрыты в Wissmann *et al.*, *Nucleic Acids Res.* 14: 4253-66 (1986). Спейсеры не являются специфичными для последовательности. Спейсеры могут быть одинаковыми, но не все должны быть идентичными. TRE считается типом оператора, используемого в данном документе.

[0067] Все числовые пределы и диапазоны, изложенные в данном документе, включают все числа или значения около них или между числами диапазона или предела. Описанные в данном документе диапазоны и пределы явно обозначают и устанавливают все целые, десятичные и дробные значения, определенные и охватываемые диапазоном или пределом. Описанные в данном документе диапазоны и пределы явно обозначают и устанавливают все целые, десятичные и дробные значения, определенные и охватываемые диапазоном или пределом. Таким образом, перечисление диапазонов значений в данном документе просто предназначено для использования в качестве сокращенного метода индивидуальной ссылки на каждое отдельное значение, попадающее в этот диапазон, если иное не указано в данном документе, и каждое отдельное значение включено в спецификацию, как если бы оно было отдельно изложено в данном документе.

Подробное описание сущности изобретения

[0068] Данные изобретения предлагают клетки млекопитающих с множеством стабильных сайтов интеграции и подходят для продукции представляющих интерес белков, включая вирусные белки, и продукции вирусных векторов, включая аденоассоциированные вирусные векторы (AAV). Один или более стабильных сайтов интеграции могут находиться в пределах «геномной безопасной гавани», а один или более стабильных сайтов интеграции могут находиться за пределами конкретной «геномной безопасной гавани». Множественные сайты стабильной интеграции могут быть созданы и расположены на разных хромосомах, в разных областях одной и той же хромосомы или в разных положениях в одной и той же области хромосомы.

[0069] Геномные безопасные гавани обсуждаются в Pellenz *et al.*, *Hum. Gene Therapy* 30: 814-28 (2019); Parapetrou *et al.*, *Molecular Therapy* 24: 678-84 (2016).

[0070] Предпочтительно, чтобы сайты стабильной интеграции содержали сайты узнавания, обеспечивающие возможность обмена кассет, опосредованного рекомбиназой (RMCE). Стабильную модификацию клеточных геномов можно осуществить с помощью известных подходов, использующих сайты гетероспецифичной рекомбинации (также известные как RRS), таких как Cre/Lox, Flp/Frt, эффлекторная нуклеаза, подобная

активатору транскрипции (TALEN), слитый белок эффекторного домена TAL, нуклеаза с цинковым пальцем (ZFN), димер ZFN или эндонуклеазная система ДНК, управляемая РНК, такая как CRISPR/Cas9. См. патент США № 9816110 в столбцах 17-18; Sajgo *et al.*, *PLoS ONE* 9: e91435 (2014); Suzuki *et al.*, *Nucl. Acids. Res.* 39: e49 (2011). Также может быть осуществлена интеграция с использованием интегразы Vxb1 в клетки человека, мыши и крысы. Russell *et al.*, *Biotechniques* 40: 460-64 (2006).

[0071] Сайты распознавания рекомбиназы, также известные как сайты гетероспецифической рекомбинации, в общем как первый (1), второй (2), третий (3), четвертый (4), пятый (5), шестой (6), седьмой (7), восьмой (8), девятый (9), десятый (10) и т. д., как видно из контекста употребления. Подходящие сайты Lox для применения согласно данным изобретениям включают, помимо прочего, Lox P, Lox 66, Lox 71, Lox 511, Lox 2272, Lox 2372, Lox 5171, Lox M2, Lox M3, lox M7 и Lox M11. Другие RRS также могут быть использованы. Lox-сайты являются наиболее часто используемым типом RRS; однако также можно использовать и другие RRS.

[0072] Гомологические плечи предпочтительно начинаются в пределах от около 10 до 20 оснований, более предпочтительно от 10 до 15 оснований от места разреза. Можно использовать и большее расстояние, но с меньшей эффективностью. Чтобы гарантировать, что ДНК-кассета(ы), вставленная в геномную безопасную гавань(и), сохраняет стабильность в случае, если восстановление гомологии может воссоздать целевой сайт, как это определено квалифицированным специалистом, область направляющего плеча ДНК-касеты можно сконструировать таким образом, чтобы она содержала изменения (например, несоответствия оснований), которые нарушают функцию целевого сайта CRISPR. Существует два подхода, которые можно использовать независимо или вместе. Первый подход заключается во вставке замен оснований для создания несовпадений оснований в целевом сайте CRISPR с двадцатью основаниями или в мотиве, примыкающем к протоспейсеру (PAM), который обычно составляет от 2 до 6 оснований. Второй подход заключается в создании донорской плазмиды, вставка которой разделяет целевой сайт CRISPR или отделяет целевой сайт CRISPR от PAM.

[0073] Линии клеток человека включают амниотические клетки (такие как амниотические эпителиальные клетки человека), клетки Hela, клетки Per.C6 и клетки HEK 293. Примеры клеток HEK 293 включают, помимо прочего, HEK 293, HEK 293A, HEK 293E, HEK 293F, HEK 293FT, HEK 293FTM, HEK 293H, HEK 293MSR, HEK 293S, HEK 293SG, HEK 293SGGD, HEK 293T и мутанты, и их варианты. Линии клеток грызунов, такие как клетки Sp2/0, клетки ВНК и клетки CHO, а также их мутанты и варианты, также можно использовать согласно изобретениям. Клетки CHO включают, помимо прочего, CHO-ori, CHO-K1, CHO-s, CHO-DHB11, CHO-DXB11, CHO-K1SV, и их мутанты и варианты.

[0074] Клетки млекопитающих по изобретению получают путем преимущественного производства и использования клеточного промежуточного продукта, который имеет кассету, содержащую ген эндонуклеазы Cas9, фланкированный сайтами

узнавания для рекомбиназы и интегрированный в геном посредством RCME. Не привязываясь к какой-либо теории, изобретательное использование интегрированного гена Cas9 при экспрессии, по-видимому, увеличивает эффективность интеграции гомологических плеч в геномные безопасные гавани за счет увеличения количества разрезов в геномной ДНК, вызванных эндонуклеазой Cas9. Использование стабильно интегрированного гена Cas9 по изобретению обеспечивает в 10 , 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 или 10^{10} большую эффективность HDR, чем HDR без стабильно интегрированного Cas9 гена. В конечном итоге эту промежуточную клетку можно дополнительно подвергнуть RMCE для удаления кассеты, содержащей ген Cas9.

[0075] В качестве отправной точки для инженерии клеток представляющие интерес полинуклеотидные последовательности, а также функционально связанный промотор и необязательные операторы могут быть введены в клетку путем трансфекции плазмиды, содержащей указанные полинуклеотидные последовательности и элементы. Соответственно, изобретения включают получение клеток, как описано.

[0076] Подходящие плазмидные конструкции могут быть созданы специалистами в данной области техники. Полезные регуляторные элементы, описанные ранее или известные в данной области техники, также могут быть включены в плазмидные конструкции, используемые для трансфекции клеток. Некоторые неограничивающие примеры полезных регуляторных элементов включают, помимо прочего, промоторы, энхансеры, последовательности, кодирующие подходящие сайты связывания рибосом мРНК, и последовательности, которые контролируют терминацию транскрипции и трансляции. Подходящие плазмидные конструкции также могут содержать нетранскрибируемые элементы, такие как точка начала репликации, другие 5'- или 3'-фланкирующие нетранскрибируемые последовательности и 5'- или 3'-нетранслируемые последовательности, такие как донорные и акцепторные сайты сплайсинга. Также могут быть включены один или более селективируемых маркерных генов. Полезные селективируемые маркерные белки и репортерные белки для использования в данном изобретении известны и могут быть легко идентифицированы специалистами в данной области техники.

[0077] Плазмидная конструкция, кодирующая представляющий интерес ген, может быть доставлена в клетку с использованием вирусного вектора или невирусного метода переноса.

[0078] Невирусные методы переноса нуклеиновой кислоты включают голую нуклеиновую кислоту, липосомы и конъюгаты белок/нуклеиновая кислота. Плазмидная конструкция, которую вводят в клетку, может быть линейной или кольцевой, может быть одноцепочечной или двухцепочечной и может представлять собой ДНК, РНК или любую их модификацию или комбинацию.

[0079] Плазмидную конструкцию можно ввести в клетку путем трансфекции. Специалисты в данной области знают множество различных протоколов трансфекции и могут выбрать подходящую систему для использования при трансфекции клеток. Как

правило, методы трансфекции включают, помимо прочего, вирусную трансдукцию, катионную трансфекцию, трансфекцию липосомами, трансфекцию дендримерами, электропорацию, тепловой шок, трансфекцию нуклеофекцией, магнитофекцию, наночастицы, доставку биолисточескими частицами (генная пушка) и запатентованные реагенты для трансфекции, такие как Липофектамин, Dojindo Hilymax, Fugene, jetPEI, Effectene или DreamFect.

[0080] Изобретения дополнительно описываются следующими примерами, которые иллюстрируют множество вариантов реализации и аспекты изобретения, но никоим образом не ограничивают изобретения. В примерах селективируемые маркеры являются положительными селективируемыми маркерами, если не указано иное как отрицательный (отриц.) маркер. **ПРИМЕР 1**

[0081] Этот пример касается создания клеток млекопитающих, содержащих репрессор, такой как TetR, под контролем промотора, такого как промотор CMV. См. Фиг. 1. Клетка трансфицируется полинуклеотидом, содержащим промотор и ген-репрессор. Полинуклеотид случайным образом вставляется в геном клетки. Вестерн-блоттинг и метод Такмана можно использовать в пуле клеток для идентификации трансформантов и определения среднего числа копий. Интеграция репрессора, такого как TetR, позволяет контролировать транскрипцию полинуклеотидов, находящихся под контролем промотора и оператора. **ПРИМЕР 2**

[0082] Этот пример касается дальнейшей инженерии клеток Примера 1. ДНК-кассета 1 схематически изображена на Фиг. 2 и содержит фланкирующие lox-сайты (1 и 2) и дополнительно содержит в порядке от 5' к 3' промотор, репортерный ген (1), кодирующий репортерный белок (1), IRES и ген маркера селекции (1) кодирующей белок маркера селекции (1) и сигнал полиаденилирования. ДНК-кассета (1) необязательно может включать оператор, функционально связанный с промотором. ДНК-кассета (1) случайным образом или сайт-специфично вставляется в геном клетки. Первый lox-сайт и второй lox-сайт на ДНК-кассете (1) различны.

[0083] Если в ДНК-кассете (1) используется оператор tet, несколько раундов сортировки -лиганд/+лиганд и сортировка отдельных клеток позволят идентифицировать стабильные по сайту Lox клетки для экспрессии, регулируемой dox. Таким образом, когда присутствует лиганд, такой как доксициклин или тетрациклин, TetR не будет связываться с оператором, и, таким образом, условия являются благоприятными для транскрипции репортерного гена (1) и полинуклеотида маркера селекции (1).

ПРИМЕР 3

[0084] В этом примере RMCE выполняют для замены ДНК-касеты (1) ДНК-кассетой (2) в клетках Примера 2. Как схематически показано на Фиг. 3, ДНК-кассета (2) содержит фланкирующие lox-сайты (1 и 2) и дополнительно содержит в порядке от 5' к 3' промотор, ген маркера селекции (2), кодирующий белок-маркер селекции (2), IRES и репортерный ген (2), кодирующий репортерный белок (2), и ген Cas9 под контролем второго промотора (необязательно функционально связанного с оператором).

[0085] В одном из вариантов реализации промотор CMV функционально связан с оператором tet для контроля транскрипции гена Cas9. Когда клетки находятся в присутствии доксициклина или тетрациклина, TetR больше не может связываться с оператором tet и, таким образом, обеспечивать возможность транскрипции гена Cas9. Репортерный белок (1) отличается от репортерного белка (2), а маркерный белок селекции (1) отличается от маркерного белка селекции (2).

ПРИМЕР 4

[0086] Этот пример касается интеграции ДНК-кассеты (3) в геномную безопасную гавань. См. Фиг. 4. ДНК-кассета (3) содержит в порядке от 5' до 3' полинуклеотид, содержащий первое плечо гомологии геномной безопасной гавани, содержащее сайт-мишень sgРНК, сайт lox (3), промотор, функционально связанный с репортерным геном (3), кодирующим репортерный белок (3), сигнал полиаденилирования, сайт lox (4) и второе плечо гомологии геномной безопасной гавани, содержащее целевой сайт sgРНК, причем каждый из первого и второго целевых сайтов направляющего плеча может содержать область с изменениями, если это необходимо, чтобы избежать повторного создания сайта, на который можно нацеливать. Лох-сайт (1), lox-сайт (2), lox-сайт (3) и lox-сайт (4) отличаются друг от друга. Репортерный белок (3) отличается от репортерного белка (2). Репортерный белок (3) и репортерный белок (1) могут быть одинаковыми или разными. В этом примере используются плечи гомологии длиной около 1000 оснований.

[0087] Когда экспрессируется эндонуклеаза Cas9, эффективность интеграции ДНК-кассеты 3 увеличивается. Не привязываясь к какой-либо теории, можно сказать, что изобретательное использование интегрированного гена Cas9, по-видимому, увеличивает эффективность интеграции за счет увеличения количества разрезов в геномной ДНК, вызванных эндонуклеазой Cas9. Использование стабильно интегрированного гена Cas9 по изобретению обеспечивает в 10 , 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 или 10^{10} большую эффективность HDR, чем HDR без стабильно интегрированного Cas9 гена.

[0088] При необходимости изменения в первом и втором плечах гомологии геномной безопасной гавани приводят к тому, что ДНК-кассета (3) остается интегрированной, избегая воссоздания целевого сайта. Меньшая кассета в ней, а именно область между lox-сайтом (3) и lox-сайтом (4), доступна для RMCE и называется стабильным сайтом интеграции. **ПРИМЕР 5**

[0089] Этот пример касается окончательной формы клеточной линии и схематически изображен на Фиг. 5. Чтобы обеспечить стабильность клеточной линии во времени, предпочтительно удалить ген Cas9. Соответственно, ДНК-кассета (2) заменяется RMCE на кассету ДНК (4) и удаляет ген Cas9. ДНК-кассета (4) содержит фланкирующие lox-сайты (1 и 2) и репортерный ген (4), кодирующий репортерный белок (4), под контролем промотора. Репортерный белок (4) отличается от репортерного белка (2) и репортерного белка (3) и предпочтительно отличается от репортерного белка (1).

[0090] Полученные клетки будут иметь два сайта интеграции в геноме, один сайт интеграции в геномной безопасной гавани (например, стабильный сайт интеграции) и

один сайт интеграции за пределами этой конкретной геномной безопасной гавани. Можно создать дополнительные сайты интеграции, применяя описанные выше подходы, включая использование интегрированного гена Cas9 и использование дополнительных и различных плеч гомологии GSH. **ПРИМЕР 6**

[0091] Этот пример представляет собой сравнение эффективности использования Cas9 с репарацией, направленной на гомологию (HDR), как описано в данном документе, по сравнению с обычным HDR. Как сообщается в литературе, HDR является точным, но желаемые рекомбинационные события происходят нечасто: 1 из 10^6 - 10^9 клеток (от 0,0001% до 0,0000001%). Hsu *et al.*, *Cell* 157: 1262-78 (2014).

[0092] Чтобы оценить преимущества стабильно интегрированного гена Cas9, была модифицирована клетка CHO, имеющая сайты, раскрытые в патентах США № 7771997 («Стабильный сайт 1») и 9816110 («Стабильный сайт 2»). Regeneron предоставляет набор товаров и услуг под названием EESYR®. Клетки CHO с интегрированными последовательностями в стабильном сайте 1 и стабильном сайте 2 раскрыты в патенте США 2019/0233544 A1, и каждый из них упоминается в нем как «локус усиленной экспрессии». Последовательности, представленные в этих патентах и примерах 11 и 12, можно использовать согласно изобретениям, описанным и изображенным в данном документе.

[0093] Клетка CHO была модифицирована для включения репортерного гена цианофлуоресцентного белка под контролем промотора в стабильном сайте 1, а также гена-маркера селекции и репортерного гена желтого флуоресцентного белка под контролем того же промотора в стабильном сайте 2. Кроме того, в стабильный сайт 2 также был встроены ген Cas9 под контролем второго промотора с оператором. Ген Cas9 может быть в конечном итоге удален в соответствии с идеями, содержащимися в данном документе.

[0094] Цианофлуоресцентный белок можно изменить на флуоресцентный зеленый, заменив остаток тирозина в положении 66 на триптофан. Плазмида доставки sgРНК содержит маркер селекции (резистентность к ампициллину), промотор POL III (промотор РНК-полимеразы III), целевую последовательность и каркас гРНК, терминатор POL III и сайты переваривания 1 и 2. Промоторы Pol III включают H1 и U6.

[0095] Как показано на Фиг. 6, были сконструированы плазмиды для доставки sgРНК, содержащие матрицы HDR: 104-мерную вставку (имеющую плечо длиной 57 пар оснований и плечо длиной 45 пар оснований), 401-мерную вставку (имеющую плечо длиной 198 пар оснований и плечо 201 пар оснований) или 1030-мерную (имеющую плечо длиной 524 пар оснований и плечо длиной 504 пар оснований), вставку, содержащую гомологичные плечи и последовательность, вызывающую изменение от циано к зеленому, которая в этом примере состояла из 2 нуклеотидов («восстанавливающие нуклеотиды»). Шаблоны HDR были вставлены в сайты расщепления (например, NotI и/или другие соответствующие сайты) плазмиды доставки sgРНК для формирования целевой плазмиды sgРНК. Плазмиду для доставки sgРНК без вставки (без матрицы HDR) использовали в

качестве контроля.

[0096] На Фиг. 7 показано, что контрольная группа не показала зеленых положительных результатов в Q1. Клетки с шаблоном HDR показали зеленый положительный результат в Q1, а популяция зеленого положительного результата в Q1 постоянно увеличивалась с увеличением размера шаблона HDR (слева направо). Клетки с 1030-мерными HDR-шаблонами показали наибольшую эффективность восстановления, которая составила около 6,5 процента.

[0097] Клетки этого примера обладают стабильным сайтом 1 и стабильным сайтом 2, а также SIS, созданным в GSH согласно изобретениям. Таким образом, эта клетка обладает тремя сайтами стабильной интеграции интересующих генов.

ПРИМЕР 7 - Создание промежуточной клетки человека, содержащей стабильный сайт интеграции в геномной безопасной гавани (AAVS1)

[0098] В этом примере отправной точкой является клетка HEK293 со стабильно интегрированным геном Cas9, фланкированным сайтами Lox 3 и 4. Ген Cas9 находится под контролем по меньшей мере промотора (не показан). AAVS1 также схематически изображен. См. Фиг. 8. Эту клетку можно создать согласно Примерам 1-4 и Фиг. 1-4.

[0099] Нацеливание на плазмиды, содержащие целевой сайт sgPНК, левое плечо гомологии (в данном документе плечо гомологии GSH) для вставки в область, такую как геномная безопасная гавань (в данном документе AAVS1), сайт Lox 1, репортерный ген (цвет 1), сайт Lox 2, правое плечо гомологии (в данном документе плечо гомологии GSH) для вставки в область, такую как геномная безопасная гавань (в данном документе AAVS1). См. Фиг. 9А и 9В для альтернативных плазмид нацеливания. На 3'-конце одной нацеливающей плазмиды имеется репортерный ген (цвет 2), см. Фиг. 9А. Другая нацеливающая плазида имеет на 3'-конце ген отрицательной селекции (негативная селекция 1). См. Фиг. 9В. Промоторы и, необязательно, другие фрагменты (такие как операторы) представлены стрелками, направленными в направлении от 5' к 3' на Фиг. 9А и 9В. Обе плазмиды вставляют цвет 1 в область, такую как геномная безопасная гавань (в данном документе AAVS1).

[00100] Опосредованная Cas9 интеграция целевой плазмиды (например, Фиг. 9А или Фиг. 9В) в геномную безопасную гавань (AAVS1) клетки HEK293 схематически изображены на Фиг. 10. Цвет 1 окружен Lox 1 и Lox 2. Интересующий ген может заменить цвет 1 посредством RMCE.

[00101] Когда нацеливающая плазида согласно Фиг. 9А правильно интегрирована, клетка будет цвет 1 положительной и цвет 2 отрицательной. Когда нацеливающая плазида согласно Фиг. 9В правильно интегрирована, клетка будет цвет 1-положительной и сможет размножаться, поскольку отрицательный ген селекции удален. Эта клетка считается промежуточной. В конечном итоге клетку можно дополнительно подвергнуть RMCE в сайтах lox 3 и 4 для удаления кассеты, содержащей ген Cas9, как показано на Фиг. 8. См., например, Пример 5.

[00102] Точность этой изобретательской методики показана на Фиг. 10 и 11. На

Фиг. 11 более подробно изображена вставка Фиг. 10. Клеточный геном, включая AAVS1, окружает вставку и 5'- и 3'-концы. Цвет 1 окружен Lox 1 и Lox 2. На Фиг. 11, слева показано расположение 5'-праймера генома и 3'-праймера для вставки, используемых в ПЦР с 5'-соединением. На Фиг. 11, справа, указано расположение 5'-праймера для вставки и 3'-праймера генома, используемых в ПЦР с 3'-соединением.

[00103] Соединительная ПЦР показывает, что фрагменты правильного размера амплифицируются и помечаются как «стабильные клетки, нацеленные на Cas9». См. Фиг. 12 и 13. Получают и обнаруживают стабильные клетки-мишени Cas9, а также 5'-соединение и 3'-соединение, что обеспечивает правильную вставку. Положительные и отрицательные контроли находятся в правой колонке каждого геля. **ПРИМЕР 8 - Регионы и последовательности СНО**

[00104] Для клеток СНО можно использовать последовательности, представленные в патентах США № 7771997 (стабильный сайт 1) и 9816110 (стабильный сайт 2). Последовательности и гомологичные последовательности в пределах значений процента идентичности патентов США № 7771997 и 9816110 включены в данное описание посредством ссылки. AAVS1-подобная область, раскрытая в данном документе, может быть использована для создания стабильных сайтов интеграции согласно изобретениям.

[00105] Локусы-кандидаты для использования согласно изобретениям описаны в литературе. Namaker and Lee, *Curr. Op. Chem. Eng.* 22: 152-60 (2018) идентифицируют 30 локусов горячих точек. Hilliard and Lee, *Biotech. Bioeng.* 118: 659-75 (2021) стремились идентифицировать области «безопасной гавани» в СНО, используя эпигеномный анализ стабильных областей Hi-C, и обнаружили перекрытие с 5 из 30 областей, выявленных Namaker и Lee. См. Дополнительную Таблицу 3 Hilliard and Lee. Gaidukov *et al.*, *Nucl. Acids Res.* 46: 4072-86 (2018) также идентифицирует локусы для интеграции в клетки СНО, включая предполагаемый Rosa26. Lee *et al.*, *Scientific Reps.* 5: 8572 (2015) сообщили о локусе COSMC в клетках хомяка. В целом, в этих статьях идентифицировано несколько неаннотированных областей и областей генов в СНО, и области генов представлены ниже:

<i>BMP5</i>	<i>SSBP2</i>	<i>TRMT6</i>	<i>CLCC1</i>	<i>FAM114A1</i> (<i>NOXP20</i>)
<i>LRBA</i>	<i>DCN</i>	<i>CEP128</i>	<i>AACS</i>	<i>ALDH5A1</i>
<i>SMAD6</i>	<i>PTPRQ</i>	<i>ROSA26</i>	<i>ADGRL4</i>	<i>GPM6A</i>
<i>KIAA1551 (C12ORF35)</i>	<i>HPRT</i>	<i>CLCN3</i>	<i>FER1L4</i>	<i>COSMC</i>

ПРИМЕР 9 - Клетки СНО с тремя или более сайтами вставки

[00106] Клетки СНО, содержащие множественные сайты вставки, с использованием клеток, раскрытых в патенте США 2019/0233544 A1. Стабильный сайт 1 и стабильный сайт 2 можно использовать первоначально в соответствии с идеями, содержащимися в данном документе, в которых используется интегрированный ген Cas9. После создания одного или более стабильных сайтов интеграции в геномных безопасных гаванях, таких как AAVS1-подобная область (см., например, SEQ ID NO:2) и

соответствующие гидовые последовательности (см., например, SEQ ID NO: от 13 до 419). Направляющие последовательности могут связываться с целевыми последовательностями в SEQ ID NO:2 в диапазонах положений нуклеотидов, выбранных из группы, состоящей из: (a) от 1 до 2000; (b) от 2001 до 4000; (c) от 4001 до 6000; (d) от 6001 до 8000; (e) от 8001 до 10000; (f) от 10001 до 12000; (g) от 12001 до 14000; (h) от 14001 до 16000; (i) от 16001 до 18000; (j) от 18001 до 20000; (k) от 20001 до 22000; (l) от 22001 до 24000; (m) от 24001 до 26000; (n) от 26001 до 28000; (o) от 28001 до 30000; (p) от 30001 до 32000; (q) от 32001 до 34000; (r) от 34001 до 36000; (s) от 36001 до 38000; (t) от 38001 до 40000; (u) от 40001 до 42000; и (v) от 42001 до 44232.

[00107] Стабильный сайт 1 и стабильный сайт 2 патентов США №№ 7771997 и 9816110 можно использовать для экспрессии представляющих интерес генов для кодирования представляющих интерес белков. Клетки с SIS в конечном итоге могут иметь 3, 4, 5, 6, 7, 8, 9, 10 или более сайтов для экспрессии интересующих генов.

[00108] Предпочтительно клетка СНО, содержащая стабильные сайты 1 и 2, модифицируется для создания третьего сайта в геномной безопасной гавани, а именно стабильного сайта интеграции. Предпочтительные геномные безопасные зоны для создания такой клетки СНО находятся в AAVS1-подобной области. Другие типы клеток СНО можно использовать для создания множества сайтов в соответствии с идеями, содержащимися в данном документе.

[00109] Фиг. 14 схематически изображает типовую клетку, содержащую три кассеты, интегрированные в области генома с фланкирующими RRS (в данном документе lox 1 и lox 2). В зависимости от типа клеток каждая из трех кассет может быть интегрирована в разные стабильные сайты интеграции и другие доступные сайты (такие как стабильный сайт 1 и стабильный сайт 2), схематически изображенные как позиции А, В и С. Репортерные гены могут представлять собой одинаковые или разные. Гены отрицательной селекции могут быть одинаковыми или разными, но предпочтительно одинаковыми. Клетка может содержать дополнительные стабильные сайты интеграции и интегрированные кассеты в соответствии с идеями, содержащимися в данном документе.

[00110] Фиг. 15 схематически изображает модификацию клетки с Фиг. 14 в схематически изображенных положениях А, В и С. Каждая из трех кассет содержит фланкирующие RRS (в данном документе lox 1 и lox 2), представляющий интерес ген, маркерный ген положительной селекции и репортерный ген. Маркерные гены положительной селекции могут быть одинаковыми или разными, но предпочтительно одинаковыми. Репортерные гены* могут быть одинаковыми или разными, но каждый из них должен отличаться от любого из репортерных генов в клетке, показанной на Фиг. 14. Гены, представляющие интерес, могут быть одинаковыми или разными. Кассеты на Фиг. 14 заменены кассетами, показанными на Фиг. 15, с помощью RMCE. Клетка может содержать дополнительные стабильные сайты интеграции и интегрированные кассеты в соответствии с идеями, содержащимися в данном документе.

[00111] Сочетание отрицательной и положительной селекции обеспечивает

выделение клеток, подвергшихся рекомбинации во всех сайтах. Если интересующий ген один и тот же в каждой из трех кассет, клетка может экспрессировать белок с высоким выходом. Например, возможно производство 7, 8, 9, 10 или более граммов белка на литр (г/л).

[00112] На Фиг. 16 показаны результаты пяти различных антител IgG человека, которые были стабильно интегрированы с использованием рекомбинации Cre-lox в клетки-хозяева, полученные из CHO K1, сконструированных либо с 2 сайтами интеграции (стабильный сайт 1 и 2), либо с 3 сайтами интеграции (стабильный сайт 1, стабильный сайт 2 и AAVS1-подобный сайт (см. SEQ ID NO:2)). Изогенные клеточные линии (ICL) выделяли с помощью проточной цитометрии. Серийное производство ICL с подпиткой инокулировали в производственные среды с определенным химическим составом и производственные культуры проводили в течение 13 дней. Титр антител в кондиционированной среде определяли с использованием способа, основанного на ВЭЖХ с белком А, и каждая клетка с тремя сайтами, экспрессирующая данное антитело (1, 2, 3, 4 или 5), экспрессировала большее количество белка, чем клетка сравнения с двумя сайтами. Трехсайтная клетка может обеспечить прибавки на 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150% или более по сравнению с двухсайтной клеткой.

[00113] Альтернативно, в кассетах можно использовать различные представляющие интерес гены. Например, последовательности тяжелой цепи и легкой цепи антитела могут представлять собой представляющий интерес ген.

[00114] Что касается клетки с четырьмя сайтами, предпочтительно клетка CHO, содержащая стабильный сайт 1 и 2, модифицируется для создания третьего и четвертого сайта в геномной безопасной гавани, а именно стабильного сайта интеграции. Предпочтительные геномные безопасные зоны для создания такой клетки CHO находятся в AAVS1-подобной области, которая может быть третьим сайтом. Четвертый сайт может быть создан в других локусах, включая, помимо прочего:

<i>BMP5</i>	<i>SSBP2</i>	<i>TRMT6</i>	<i>CLCC1</i>	<i>FAM114A1</i> (<i>NOXP20</i>)
<i>LRBA</i>	<i>DCN</i>	<i>CEP128</i>	<i>AACS</i>	<i>ALDH5A1</i>
<i>SMAD6</i>	<i>PTPRQ</i>	<i>ROSA26</i>	<i>ADGRL4</i>	<i>GPM6A</i>
<i>KIAA1551 (C12ORF35)</i>	<i>HPRT</i>	<i>CLCN3</i>	<i>FER1L4</i>	<i>COSMC</i>

[00115] Другие типы клеток CHO можно использовать для создания множества сайтов в соответствии с идеями, содержащимися в данном документе.

[00116] Фиг. 17 схематически изображает типовую клетку, содержащую четыре кассеты, интегрированные в области генома с фланкирующими RRS (в данном документе lox 1 и lox 2 или lox 3 и lox 4). В зависимости от типа клеток каждая из четырех кассет может быть интегрирована в разные стабильные сайты интеграции и другие доступные сайты (такие как стабильный сайт 1 и стабильный сайт 2), схематически изображенные

как позиции А, В, С и D. Репортерные гены могут быть одинаковыми или разными. Гены отрицательной селекции могут быть одинаковыми или разными, но предпочтительно одинаковыми. Клетка может содержать дополнительные стабильные сайты интеграции и интегрированные кассеты в соответствии с идеями, содержащимися в данном документе.

[00117] Фиг. 18 схематически изображает модификацию клетки с Фиг. 17 в схематически изображенных положениях А, В, С и D. Каждая из четырех кассет содержит фланкирующие RRS (в данном документе lox 1 и lox 2 или lox 3 и lox 4), представляющий интерес ген, маркерный ген положительной селекции и репортерный ген*. Маркерные гены положительной селекции могут быть одинаковыми или разными, но предпочтительно одинаковыми. Репортерные гены* могут быть одинаковыми или разными, но каждый из них должен отличаться от любого из генов-репортеров в клетке, показанной на Фиг. 17. Гены, представляющие интерес, могут быть одинаковыми или разными. На этой Фиг. показаны две копии интересующего гена 1 и две копии интересующего гена 2. Кассеты на Фиг. 17 заменены кассетами, показанными на Фиг. 18, с помощью RMCE. Клетка может содержать дополнительные стабильные сайты интеграции и интегрированные кассеты в соответствии с идеями, содержащимися в данном документе.

[00118] Сочетание отрицательной и положительной селекции обеспечивает выделение клеток, подвергшихся рекомбинации во всех сайтах. Клетка с четырьмя сайтами полезна для создания биспецифических антител, в которых две отдельные плазмиды тяжелой цепи/легкой цепи могут быть нацелены на разные сайты.

ПРИМЕР 10 - Последовательности геномных безопасных гаваней

[00119] В данном документе описаны последовательности геномных безопасных гаваней и т.п., многие из них описаны в литературе и общедоступны. Примерные последовательностей представлены ниже. **Последовательность AAVS1 человека**

AAVS1 человека (*указан* нативный RBS и сайт направляющей РНК для вставки в безопасную гавань)

(SEQ ID NO:1)

GAATTCCTAACTGCCCCGGGGCAGTCTGCTATTCATCCCCTTTACGCGGTGCT
 ACACACACTTGCTAGTATGCCGTGGGGACCCCTCCGGCCTGTAGACTCCATTTCCCA
 GCATTCCCCGGAGGAGGCCCTCATCTGGCGATTTCCACTGGGGGCCTCGGAGCTGCG
 GACTTCCCAGTGTGCATCGGGGCACAGCGACTCCTGGAAGTGGCCACTTCTGCTAAT
 GGACTCCATTTCCCAGGCTCCCGCTACCTGCCCAGCACACCCTGGGGCATCCGTGAC
 GTCAGCAAGCCGGGCGGGGACCGGAGATCCTTGGGGCGGTGGGGGGCCAGCGGCA
 GTTCCCAGGCGGCCCGGGGCGGGCGGGCGGGCGGGTGGTGGCGGGCGGTTGGGGC
TCCGGGCGCGTTCGCTCGCTCGCTCGCTGGGGCGGGCGGGCGGTGCGATGTCCGGAGA
 GGATGGCCGGCGGCTGGCCCGGGGGCGGCGGCGGGCTGCCCGGGAGCGGGCGACG
 GGAGCAGCTGCGGCAGTGGGGCGCGGGCGGGCGCCGAGCCTGGCCCCGGAGAGCG
 CCGCGCCCGCACCGTCCGCTTCGAGCGCGCCGCGAGTTCCTGGCGGCCTGTGCGGG
 CGGCGACCTGGACGAGGCGCGTCTGATGCTGCGCGCCGCGACCCTGGCCCCGGCG

CCGGAGCTCGACCCCGCCGGCCGCCGCCGCCGCCGCCGCGCCGTGCTGGACTCCACCAAC
GCCGACGGTATCAGCGCCCTGCACCAGGTCAGCGCCCCCGCGGGCGTCTCCCGGGG
CCAGGTCCACCCTCTGCGCCACCTGGGGCATCCTCCTTCCCCGTTGCCAGTCTCGATC
CGCCCCGTCGTTACTGGCCCTGGGTTTNCACCCTATGCTGACACCCCGTTCCAGTCCC
CTTACCATTCCCTTCGACCACCCCACTTCCGAATTGGAGCGCTTCAACTGGCTGGGC
TAGCACTCTGTGTGACACTCTGAAGCTCTACATTCCCTTCGACCTACTCTCTTCGATT
GGAGTCGCTTTAACTGGCCCTGGCTTTGGCAGCCTGTGCTGACCCATCGAGTCCTCC
TTACCATCCCTCCCTCGACTTCCCCTCTTCCGATGTTGAGCCCCCTCCAGCCGGTCCTG
GACTTTGTCTCCTTCCCTGCCCTGCCCTCTCCTGAACCTGAGCCAGCTCCCATAGCTC
AGGTCTGGTCTATCTGCCTGGCCCTGGCCATTGTCACCTTTGCGCTGCCCTCCTCTCGC
CCCCGAGTGCCCTTGCTGTGCCGCCGGAACCTCTGCCCTCTAACGCTGCCGTGCCGTC
TCTCTCCTGAGTCCGGACCACTTTGAGCTCTACTGGCTTCTGCGCGCCTCTGGCCCAC
TGTTTCCCCTTCCCAGGCAGGTCCTGCTTTCTCTGACCAGCATTCTCTCCCCTGGGCC
TGTGCCGCTTTCTGTCTGCAGCTTGTGGCCTGGGTACCTCTACGGCTGGCCCAAGA
TCCTTCCCTGCCGCCTCCTTCAGGTTCCGTCTTCCCTCCACTCCCTCTTCCCCTTGCTCT
CTGCTGTGTTGCTGCCCAAGGATGCTCTTTCCGGAGCACTTCCCTTCTCGGCGCTGCAC
CACGTGATGTCCTCTGAGCGGATCCTCCCCGTGTCTGGGTCCCTCTCCGGGCATCTCTC
CTCCCTCACCCAACCCCATGCCGTGTTCACTCGCTGGGTCCCTTTTCTTCTCCTTCT
GGGGCCTGTGCCATCTCTCGTTTCTTAGGATGGCCTTCTCCGACGGATGTCTCCCTTG
CGTCCCGCCTCCCCTTCTTGTAGGCCTGCATCATCACCGTTTTTCTGGACAACCCCAA
AGTACCCCGTCTCCCTGGCTTAGCACCTCTCCATCCTCTTGCTTTCTTTGCCTGGACA
CCCCGTTCTCCTGTGGATTCCGGGTACCTCTCACTCCTTTCAATTTGGGCAGCTCCCCT
ACCCCCCTTACCTCTCTAGTCTGTGCTAGCTCTTCCAGCCCCCTGTCATGGCATCTTC
CAGGGGTCCGAGAGCTCAGCTAGTCTTCTTCCCTCCAACCCGGGCCCTATGTCCACTT
CAGGACAGCATGTTTGCTGCCTCCAGGGATCCTGTGTCCCCGAGCTGGGACCACCTT
ATATTCCCAGGGCCGGTTAATGTGGCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCC
ACCCACAGTGGGGCCACTAGGGACAGGATTGGTGACAGAAAAGCCCCCATCCTTA
GGCCTCCTCCTTCTAGTCTCCTGATATTCGTCTAACCCCCACCTCCTGTTAGGCAGA
TTCCTTATCTGGTGACACACCCCATTTCCCTGGAGCCATCTCTCTCCTTGCCAGAACC
TCTAAGGTTTGCTTACGATGGAGCCAGAGAGGATCCTGGGAGGGAGACTTGGCAGG
GGGTGGGAGGGAAGGGGGGGATGCGTGACCTGCCCGGTTCTCAGTGGCCACCCTGC
GCTACCCTCTCCCAGAACCTGAGCTGCTCTGACGCGGCTGTCTGGTGCGTTTCACTG
ATCCTGGTGCTGCAGCTTCCCTTACACTTCCCAAGAGGAGAAGCAGTTTGGA AAAACA
AAATCAGAATAAGTTGGTCCTGAGTTCTAACTTTGGCTCTTACCTTTCTAGNCCCCA
ATTTATATTGTTCCCTCCGTGCGTCAGTTTTACCTGTGAGATAAGGCCAGTAGCCACCC
CCGTCCTGGCAGGGCTGTGGTGAGGAGGGGGGTGTCCGTGTGGAAAACCTCCCTTTGT
GAGAATGGTGCGTCCTAGGTGTTACCAGGTCGTGGCCGCCTCTACTCCCTTTCTCTT
TCTCCATCCATCCTTCTTTCCTTAAAGAGCCCCCAGTGCTATCTGGACATATTCCCTCC
GCCAGAGCAGGGTCCGCTTCCCTAAGGCCCTGCTCTGGGCTTCTGGGTTTGGAGTCC
TTGCAAGCCCAGGAGAGCGCTAGCTTCCCTGTCCCCCTTCCCTCGTCCACCATCTCAT

GCCCTGGCTCTCCTGCCCTTCTACAGGGGTTCTGGCTCTGCTCTTCAGACTGAGC
 CCCGTTCCCCTGCATCCCCGTTCCCCTGCATCCCCCTTCCCCTGCATCCCCCAGAGCC
 CCAGGCCACCTACTTGGCCTGGAACCCACGAGAGGCCACCCAGCCCTGTCTACCA
 GGCTGACCTTTTGGGTGATTCTCCTCCAAGTGTGGGGTACTGCTTGGGCAAACCTCA
 CTCTTCGGGGTATCCAGGAGGCCTGGAGCATTGGGGTGGGCTGGGGTTCAGAGAG
 GAGGGATTCCCTCCAGGTTACGTGGCCAAGAAGCAGGGGAGCTGGGTTTGGGTCAG
 GCTGGGTGTGGGGTGACCAGCTTATGCTGTTTGCCAGGACAGCCTAGTTTTAGCGC
 TGAACCCTCAGTCCTAGGAAAACAGGGATGGTTGGTCACTGTCTCTGGGTGACTCT
 TGATTCCCGGCCAGTTTCTCCACCTGGGGCTGTGTTTCTCGTCCTGCATCCTTCTCCA
 GGCAGGTCCCAAGCATCGCCCCCTGGCTGTTCCCAAGTTCTTAGGTACCCACGT
 GGGTTTATGAACCACTTGGTGAGGCTGGTACCCTGCCCCATTCCTGCACCCCAATT
 GCCTTAGTGGCTAGGGGGTTGGGGGCTAGAGTAGGAGGGGCTGGAGCCAGGATTCT
 TAGGGCTGAACAGAGCCGAGCTGGGGGCTGGGCTCCTGGGTTTGGAGAGAGGAGGG
 GCTGGGGCCTGGACTCCTGGGTCCGAGGGAGGAGGGGCTGGGGCCTGGACTCCTGG
 GTCTGAGGGTGGAGGGACTGGGGCCTGGACTCCTGGGTCCGAGGGAGGAGGGGCT
 GGGCCTGGACTCGTGGGTCTGAGGGAGGAGGGGTGGGGCCTGGACTTCTGGGT
 CTTAGGGAGGCGGGGCTGGGCCTGGACCCTGGGTCTGAATGGGGAGAGGCTGGGG
 GCCTGGACTCCTTCATCTGAGGGCGGAAGGGCTGGGGCCTGGCCTCCTGGGTTGAAT
 GGGGAGGGGTTGGGCCTGGACTCTGGAGTCCCTGGTGCCAGGCCTCAGGCATCTTT
 CACAGGGATGCCTGTAC

Последовательность областей, подобных CHO AAVS1

(Гидовые последовательности для вставки показаны ниже в Примере 13)

(SEQ ID NO:2)

CCAGCACCCACATGGTGGCTCACAAGTCCGTAAGTCCAGTTCAGAGGAT
 CTGATGCCCTCTTCTGTCTCCCGCGAGCACCTGGCACACACGTGATGCACACTTAAA
 CACATGCAAGCAAACCATCAGACACATAACTTTTTTTTCCAATTTTTTAAAGATTTA
 GTTATTATTATTTACTTAATAAATATTTATTATATTTATTACATATACAGTTTCTGCCT
 ACATGCCAGCAGAGGGCACCAGATTGAATTGTAGATGGTTGTGAGCCACCATGTGG
 TTGCTGGGAATTGAACTCAGGACCCTGGAAGAGCAGTCAGTGCTCTTAACCTCTGA
 GCCATCTCTCCAGCCCCTCCATTTTTTTTTTTTTTAAATAAAGAAATGTAATGTCCTAA
 GTGGGGCTTAGAGAGTGGAAGCAGATAAAGAAAGATGGAGTTAAGAATTTAAGA
 AGCCAGTTGGCGGTTGTGCATGCCAGCACTCAGGAGGCAGAGGCAGGTGGATGGAT
 CTCTATGAGTTCGAGGCCAGCCTGGTCTACAGAGAGAGAGTTCCAGGACAGACTTCT
 CCAAAGCTACAGAGAAACCCTGTCTGAACCCACCACGACCACCACAAAGAAAAAAA
 GGATTTCAAGAGGAGAGCCAGGTTTATAGCAAGAGAGAAAGTTGTGAACTAATGCC
 CAGGGCTTAGTGTGGCCTACCTCTGGGCTGGGTCTCTCTCTGAACACAGGGTGGAGC
 TGCCCCGGGAGGAAGAAGCGGCTCCGTACAGTCCCGAATTCTACAGTGGCTGGGAG
 CCTCCCGCCACTGACCCGCAGGGCCGCGCCTGGGAGGACCCGGTGGAAAAACAGCT
 ACAGCATGAGAAGAGGCGCAGGCAGGTGAGGCAGGGTTGCCGGGGGAGCACTGGG
 CTCCCCGTTTCTGCACAACATGGGCGAGCAGGACGTCTGAGGTCTAGCCTGCCTGAC

CCCAAGCTCTCTCTCTTCCCGCAGCAAAGCGCCCCCAGATCGCTGTCAATGGGTGA
GTGACCGCTGCAGGGTGGCCAGGGATGGGGTTGGGAGGACTGAGTCCCGGGGTAC
CCCGGCTCTGACTCCGACCCTCCCCCTTTTTTCTTGTCTTTTTTTTTTTTTTTTT
TTTAAACCTCTGCCTTCCCGGCTCTTTGCAGGTGGGTGAGGTGGTGAGGAGGCGGGG
CTGGGGTGGGGGTGGGGGAGGAGCCAGGAGGGAGGGGGGGAGGAGCCAGAACTC
TGGGTCCAAGGGAAGAGGGAAAGGAGGCTTAGTTTGCTGAAGCTATGAGAGTTAGG
GGCTGAAAGTGGGTGGGTCTAAAGGCTTGGACCCACACCCCCACCCCGGCATCC
TCAAAGATTGAAAAGGTGCAGTTTGGTGTCTAGGACCTGGGAGAGCACCATGCT
TGAGTCCCCAGAGCACAGAGCACTGGGTGTCAGAGAAAAAAAAAAAAAAAAATGGAGACC
AAAAAGCAGGGTTGGGACTTCCGAGGATTCAGGGACAAGTTTGAGGAAACGTGAGA
AAGTGTGGCATCCCTGGACCACTAACTGAGGTGGGACTTCCGGCTTCCTAATGCGC
AAAGGAATAGCACGTA CTGAGCAA CTGGAATGCTCCCAGGGCTGAAAGAATGGA
GGAAATTGAAGGTCAAGGCACGGACTCCTGCCTAGGTCCCTGGGAAGGAAAGA ACT
AGGGACCTAAATTTACAGTTCTACCAA ACTATGGAAGCTGAGGGCTGCAGGTCCAG
GTGAGGAAGTGATGGAGAGGGGGT CACAGCCCTAGGATCCTTGGGGAAATAGGGG
CCAGGAGTGGAGGGCGTGGATGTGGCTTGAGAACAAAATGATAGACTTGGAGGAG
AGGAATTGGGGGCCTAGGTGAGAGCCCCAGCAGAGGGTCTCAGCAGGGACGGCAT
ACTGGGAGCTGTCAGTCCCACACATGGGGCGCCGAGGCCCTGAAGAGTCCCCTCCT
CCCTTCCACAGGTAGGCCTGATCCGGGATGAGGTCTCTCTTGCTGGGGGCGCCAGAG
CTAATCGTCCCCCAGGCTGCCTGGTGTGCAGGGCCCTCTTGTCTGTCTGTCTGCTTC
TGAATCTTGGGCTCAGCACCTGCAAGCTGTTTACTCGCCTTCTCTGGCTGTAATTTCT
TTGCCTGGAAGGGTGAGGACTCTCTGGCGCTGTAAGGGGCTTGCAAAGAGCTCAGT
GCCGTGACTCAGCCTGAGTTCAAATCCAGCTGCATGAAGAACAGTACAGAGTGACC
CTGACAAGGGCAGCCTAGGGCCAGCTCAGTCACACCTTTCTCTTTCTTGTGCACTGG
CCGTTACTACAGTATCCCTCGGTTCCCTTCATATAGAAAGAGAAATAGTGAGCCGGGC
AGTGGTGGCGCACACCTTTAATCCCAGCACTTGGGAGGCAAAGGCAGGTGGACCTC
TGTGAGTTCAAGACCAGCCTGGTCTACAAGAGCTAGTTCCAGGATAGTCTCCAAAGC
CACAGAGAAACCCTGTCTCGAAAAACCAAAAAAGAAAAAGAAAGAAAGAGAAAT
AGTGAGACCGGCAGTGGTGGTGCACGTCTT TAGTCCCAGCACTGGGGAGGCAGAGG
CAGCCGGATTTCTGTGAGTTCAAGGATAGACTGGTCTACAGAGTGAGTTCCAAGAC
AGCCAGA ACTAAACAGTGAAACCCTGTCTTGAAAAA AAAAAAAGTGAAATAATG
GCCATATTCTGGTGATGGTGTAGGCCTGTGGTCCCAGCTACTCAGAGACATGAAGCA
GGAGAATAAAAAATCAAGGCCTGCTTTGACTACAAAGTGAGCTTCAAAGGCCAGCCT
GGGCAAAGCAACAAGGCCTTGCCTCAAAATGAAAAAATAAAAAATAAAAGAGGCTG
GAGAAATGGCTTAGTGGTTAAGAGTACTGGCCGCTCTTCCAGGGGACCAGGGTTCA
ATTCCCAGCACCCAGACATACAGCAGCTCACA ACTCCAGTTTCAGGGAATCCGGTGT
TCTCTCTGGTCTCTGTAGGCACCAGGCACTCAAGTTGTGCAGACATAAAATAACACA
GAGGGCTGGGCTGGGGCTCAGTGGCAGGCATTTGCCCAGAATCCCCCAGTAAAGAC
ATAGCTCAGTGAATCCAGAGCTGAGGGGCTGGGCGTATATTAATGGTGGAAATCCTT
GCCTAGAATTCAACCAGCGAAGGGCTGTGGCCGTGGCTCGGCTGTAGAACCCTGTC

CTGGTATCTACCATGAAGGGCTGGGACATGGCTCAGAGATAAAACACTTGCCTAGA
CTCTACCGCTGAGAGCCTGGGGTGTGGATCAGTGGACAGTGCCCGCCTAGCATGCA
CAAGGCCCTGGGTTCAATCCCCTGTACCACAAAAAAAAGGGGGGGTGGAGGGAGG
GTAAGAGTGAGATCTCAGGAGAAGGAAGGAACCAAATTCATGGA ACTACAAGGGA
ACTCCAGGAGAATCGAAGCGTTTCTGGCGTACGTTGCTGTGTAAGCACAAAGGGTCG
GCTATTTTTGCACCCTGTTCAATTATCCTAGCGGGTGATGGGAATAGATCTGCTGTCTC
TAGCCGATTCCTCATGATCCTCACTGATGAAAATGCAGGTGAGGGGCTGGAGAGAT
TAAGA AACTGTCTGCTCTGGCACTGGACCTAGGTTCAATCAGCTCCCCACAGCACA
TGGTGGCCCAAAATATCTGTA ACTCCAGCTCTAAGA ACCCAGGTCTAGGACACCCT
CTCCTGGACTCTGTGGCTACTGCACACAGGTGATGCACATACACACACATGCATGCA
GGCAACACACACACACACACACACACACACACACACACACAATGCATGTGAA
CGACTGGGGATGAAGCTCGGAAGCTAAGCACTTCCCTGGCATGCACGGGCCCTGGG
TTCAATCCCCAGCACCCATAATGAATTAATCGTTATCATGATACGGTGTGTTTAC
TGCATGGTGCCAGGCAAGGAAATGAGCTAACTCCATTCAAGCTGTGACTCCAGTGTC
AAGCCTGTATTAACATATTAACCTGGGCCTCTGCTCTGACCCCTGCTTGGCTCTAAC
CCCACCTCACACCTTAGAGTCCAGACCAGCAGGGCTGGCTACCTCCTAATCTCCTGC
TGGTTTCTTTCTCCCCAGTCATCAAGATCCAGACCTGGAAGCCGCCGAGCTAGAAGA
GAGAGCCAGAAAGTGGGTTCTGTGTA ACTATGACTTCCAGGCCCGAAATAGCAGCG
AGCTGTCTGTCAAGCACGGAGATGTGTTGGAGGTTAGCGGTGTGGGGGGCCTGAGA
CCCTGAAATTGGTCAATTTAGCCCTAGGTATAGA ACCGGAGCGTGAATTCTCTCCTT
ATACGCCACCTAGGTCCTGGATGACAGGCGCAAGTGGTGGAAAGGTTCCGGGACCATC
AGGGACAGGAGGGTTATGTACCCTATAACATCCTGACACCCACCCTGGACCTCAG
GTGCACCGCAGCCAAAGTCCTGCAGGAAACCTAGTAAGTCGGCGTGTTCTTGCTTCT
TCGGGGAGAAAGGGGGGCAAGATCCTAGGTCCTGGGGATGAGGACAGAGAAAATC
AGGTGTGAAGGTTGCTGTTTGGAAAGGGGGGGGGTGGTCAGATGTTTATTGGGAA
AGGAGCTGGAAGCCTCTCTTCATTCCCTTCCAGGAGACGAGTACTCCTCCTCCCCA
CCCGCACCAGCTCCAGCCCCTGCTCAGGTGCGACCCCACTGGGACAGTTGCGACAGT
CTCAACAATTTGGACCCCAAGCGAGAAGGGTGAGTGGTGGAGCGTCACTCTGGGAAG
TGATCCTTGTCTTCGCTTTTCAGGCTCCACCCTGGGCACCCTAGCGGCTCCCAGCCCC
CTGACCCCAAGAACCCCTGAGCGCGCACTCCCCTCCGCCCCCCCCCTCACGGTTTCG
CTTCTGCAGAGAAATTCTCCCAGATGCTCAGTGTCAATGAGGAGCTGCAGGCGCGCC
TTGCGCAGGGCCGTTCCGGTCCCAGCCGGGTAGCCCCGGGACCCCGCGCCCCGGAG
CCTCAGCTCAGCCCGCGCTCTGAGGCCCTCGGTGGTCCGTGCCTGGCTGCAGACCAAG
GGCTTTAGCTCGGGGTGAGTGGGGCTCCCCCGGGGCTAGTCTGAAGAGACCTGTG
CTTGA ACTGAAAGGCGAGGTTCCCAATTGGTCCAGGGGTGGGGGCGTGGA AACTGTG
GAGCAGGCCCAAATTGCAACGCCCAATGCCCAGGGACAGGCTCCAAACGGAGGCCA
CAGGAAAGGAAGTCCCATCCCCTTTCCGAAGCCCCAAATCTCCAAGAGTTTGAACAT
CCCCCCTCCCCCAGCTTCCTTGTTTGAGAACTCTGATTGCACAAGCAGCTAGGTA
GGTGTGGCGTGATTGGTGGAGGGCCGAGGGAGCTTGATGAGCTGTGATGGCCCCTG
CTGCCTCGCTCAGGACTGTGGACGCGCTCGGCGTGCTGACCGGAGCACAGCTCTTCT

CGCTGCAAAAGGAAGAGTTGCGGGCGGTGTGCCCCGAGGAAGGGGCGCGGGTGTA
CAGCCAAGTCACCGTGCAGCGCGCGCTGCTGGAGGTGAGCGAATCCTTGGGGCCGG
ACAAGGCGACGGAGGGTAGGGTGGGGATGGGGGACCTGGGGGGAGGGGGTTCGTCC
AGGGTTCACATACTAAGATCTTGATTTCTACCCCGCTCTGCAGGACAGAGAAAAAGT
GTCGGAGCTGGAGCCGTGATGGAGAAGCAAAAGAAAAAAGTGGAAAGGCGAGACCA
AAACAGAAGTTATTTGATCCTTCCTGACTCGGTACAAAACGTGATGGCATGGCGGG
GCTCCCAGCGCCCCCTAGGACAACAGTCGCCAGACTCCTCCCCGTGACCGGGGACA
GTAGATGTCCCGAAGGATCGCCACCCTCATCTCCCGGCTCACTCGCTCGCTCGCTC
TCCTGGCGGGCAGGCTGCGCTGACAGTGCCGGCTGGAATCCTTCCGGGGGACCTCA
GACTGACGGGGACGGGGACGGGGACGGGGACGGAGCATAACAGACTACTAC
CAGAGAGGCACGCCAAGAGGCGCACGGAGGGAGGGCCCTGGGCGTTCGTGACGTG
CTATAAACAGCCTCCTTTCTAGACCATGCGTGTACCTGCTGTCCCCTTCTCTCGCCG
GCTACCCAGGAGCCAGGAATCTGAGAGATGCCCCACGCTTCCTCCCCATAAACCTG
GAGAGTCCAGCCCAGGCTTCCTAATCACCAGTCTATCCTCGCACTGGCCCCATCTAC
ATCCCTTCTCCTGTTCAAACCCCTCGCCTGGCTGGCTCCTCGTTGTTCTCAGTCCTGT
CTCCTGGTGTTTAAGGCCTGGGCTTTTCTCATTGTCTCCGCCACCCTGCATTTCCGGC
CCAGCCGCTCCAGACCACAAGCGGTTTGCACCTAACGCTTCTGAGGGTTGGAGCGGC
CCCCATCACCCCTGGCTCGGCTCTCCTAGCCACACCGTGGACACCCGTGTCCAGCCTC
TAAGGACCGGCCATGCAGATCTGGACGCTCCCGGGGCATGCCACGGGCTCTTGGTTC
TTCCTGGCCCCCTCAACAACCTTTCTCCCTGCCAAGCCCTGCAACTTGTCCAGGTTATGC
AGGTGGATGGTAAGAGCCGGTTTTCTCATCCGCGCTAGGTTTATCTAAGGCCTTTCT
TTTCCCTGCATCCTTGGAAACTCCCAAGAGTCCCACCGTTGCAGTCGGCCTCTGCTC
CCCGCGCAGCTCAGTCCTTACCTGGGCCACCAGGTGGCGCACCTCGAATCTGACCCA
GGAGGGCCAGCCTTGGGCTGACTTCACTAAGCCCCCTTTCCCTTCTGGAACACTGTAG
CGTTCAGTAAGCCTTTAGTGTCCATTCCCTTGGTTTTCTCCTGGTACATGAGATAAAA
CCTAACTCCAGCATGACAGCCGATGGCCTGTGACCCCTATGGGCTCAGGTCGCCCTT
CCTCTCTGTTCCGGGACTCCAGGCACTGGTCCATGCTGTTGGTTCCTGTTGGGATGTCTT
GGCTCCATGGTGTCTTATCACTGCCTGGGGCGTCATTTCTTATGTCGCGCTTGGTTGG
TTTGTGGAGGCGTCTGGGTACAGCCCCAAACTCTCGGTCCTCCAGTTTCAGTTTCC
TGCATGTGGGGATATTGGCAGGCGCCCTGCTGCCACCCTCTTTTCTAATCGAGAAAC
CAAAAGTACAAGCAGTTGCCCAAGCTGTTTTGATTCCGGCAGTGAGGTCCCAGACTA
CAGACTGAAATGCCAGCAGGAGCCATCTGGCTTGCTGGGACATCAGGTGATCAGGT
GCCTGTGGCTGGCTCTCTGTGGTTTTGGAGTCTGACCTTTTCATCCTGACTTGACCCTC
TGTCGATCACTTTGTCCATCCATCACTCCCCAAGTCTACATCCAGCCAGGGGCACCT
GTCAGAGCTCAAGCCGGATGGTAACCTGGTGGTCAGGCCTCCCAGCTCAGGTGGAG
CTCAAGTTCTTAACAGAGCCATGATCACACACAAAGCCATCACCTCAGCGCCACAG
CACGCCAGGCCTGCTCTACCCACGCTGCACACGGTTCTCATCATCATGCAAAAGGT
GCTTCCTTCAGATACAGGGCTCACCGTCACCTTCTAGCATCTGTCTGTGCAGCTTGTG
ATGGGGCCTACTTTTGACTGTCATAAACACCACACACGCACATATATATACACACCA
GATACACACACACCACACACATGCCCAATACACTGTGCATGCGCACACACAAACAC

ACACACATACCTCATAACACCATAACACCCTATAACCCACACCAGCCATAACCACACACC
ACATATACACAGTTCACCTCAGACAGCATGGCACACCACACACACACACACACACA
CACACACGCGCGCGCGCGCACACACACACACACACACACACACTCCGCACTC
TCCCCTTCTCCACAGCACTGTAGCTGAAATCCACACAGTGGCAACCTTCCTCAGTGT
ACTGGCTGCTGGACCAAGCTGTTCACTCCTGTGACGCCAGCTGGCAGAACAGCCCAT
TCCTGACTGTCAGGATGGAGGAGGCACCACGCGATCCATCTCAAGACTGATTCCTGG
CTCTGCCCCAGTCACTGTGGCCACGAAGGACTACTTACCATCACCTACTCCTTTCCC
AGAAAACCTAGACTTGCGGTTTCCTATGTTGGCCATCCTACCTTTTCAATGTTAAGCC
ACTGACTCCGCTCACTTCCAAAGCACTGAGGGTCAATGTGAGCACCCGGATCAGGTC
ACAGGCTTCCTTCTGACCCCCCTACCTCACCTGGGGCTCTTTCTCTCCAGCTGCTCA
CTCGAGCAAGCTCCCCTCCCCACACCTGTGAGCAAGCTCCCAGCCACCCACTGGCCC
TCATCCAAATGGATGAGCGGTTTCAGTCAGATACACAGGCTGAGTATAACAAGCAGG
AACCAGTGCCCCACACCAGGGGGAGACAAGTCACTGAGTGGCAATGTCACGACTT
TATTTGTGGTGCCTGTGCTTTGTCTCAAAAATACCTTCTCCCCCTCCCCAGACAATGG
GTGGGAAGGAGGCAGCAAAAATAGAAGACAACCCTCCCTATTGCACACGGACCCTA
TATACAGGCCACCTGGCAGAGGCCAGTGGGGCTCTTGGCACATTCTGGATCCCTG
CTGGGGAGGGGAAGGGATACTGGGTAGCATCACACGTGAGGTGGGCCCGGGGCAGC
CACTCTGCTCCTGGATACTGATCCTGGCTTCCTTGGTCCTTGCTTCCTTCTGGTCCC
ATCTCTGGTGCCTGCCACTCTCGGCAACATTTCCCTACCTGGCTCAGCCTCCCACCT
CCACCCTGGTTCTGGGGACTCTGTGCTTTCTCCGGGTTCTGAGGTCCCGAGAGGAG
GTTATGGCTTCTCAACAACCTTCCCCCGGAGCCCTGTCACTCATGTTCACTCGGGGGA
AGGGGTGCGTGTGTCAAAGCAGCTGTATAAATACGGTGCGGGAGCCCCTCCAGAG
TCACTTGGAGAGCTTGCTAATGACGCGGATCAGTGCTGCATTCTCATCCTTGAGCCG
CTGGTTGTCAGCGCGGAGGTCGGACAGGGCCTAGGGGGCAGGGTGGAGTCAGCTGG
GCAGGGCGGGGCAGGGTGGGCTCTGGCCACCGCCCTTACAAGCTCGTTACCTTCA
GCTCCTCCTCCAGCTCTGCGGCCTTGCCTCCAGGGCCCTGCGCTCCTGCAGGAATG
GGCTGGGCTCAGAAGCAGGGTAAGGGCAGGGGACAGGGCAAGGGCGGGACACCAC
CCCAGCGGCCCAAACCTCACGAATCTCTCCAGCTCCAGGAGGGCGGGCCTCTCGGCA
AAGCGTTCCTGCCGCTGGTAAGGGCAGAGAAGACTGGGCGTCAGGAGCTGCTTCTT
ACCCCTAGGACATCAGAGCCCTGCCCCCCCCCCCCGAGTGGGGGACCTCCAACCTCC
CAGCCACGGCCAGGCCCTTGCCACTGGGGCTCTGACTCCCACTGCCCCAACAGCTG
GTTCTTAGGTCTCAGTATCTGCACCTGCGTGGCCCGCTCAAGCTCCACCTTGAGCTGT
GCCAGCCGCAGGGTGGTCTCTGTGAGGGCCTCACGAAGCCGCTCGTTCTCCCTCCGA
AGCTCCATGTACAGCTAGGGACACAGAGGAAGCAGGCAGGCTCAGAAGGGCCCGG
GAAGGGGCCAGGACAGGGTGGGGTGGGGCAGGAGGTAGCATGCGGCACCTTCCGG
AAGCTTCCATCGGGTTCTTCCCTGTTCCCTGCTTGGATTCTGGATTGAGGTCTCTCTGCA
AACGCTGTCTACGGGCAGTGGAGCCGCCATCCACGGTGTGACAGAAATTCAGGC
CTTAGGGCCCAGGCCCTGCCCCAGGGGTGCCCCAGCCCCACGCATGACCCGGCCT
ACCTGCACTCCAGGCTCCGTTCTGCCGGCCCCGCTCCTCCCCCTGCAGAAGAGCCC
TGAGAGTTCAGTCTCCATGCAACGTCTCCCTCCAGCCCCGCCCGGCCTCCACACAGC

ATCCTCACCTCCGCGGCCCTCCCTCCAGCCCGCCCGGCCTCCACACAGCATCCTCA
CCTCCGCGGGCCCCTCCCTCCAGCCCGCCCGGCCTCCACACAGCATCCTCACCTCCG
CGGCCCTCCCTCCAGCCCGCCCGGCCTCCACACAGCATCCTCACCTCCGCGGGCCCC
TCCCTCCAGCCCGCCCGGCCTCCACACAGCATCCTCACCTCCGCGGGCCCCCTCCCTC
CAGCCCGCCCGGCCTCCACACAGCATCCTCACCTCCGCTGCCCTCCCTCCAGCCC
GCCCCGCCTCCACACAGCATCCTCACCTCCGCGGGCCCTCCCTCCAGCCCGCCCGGC
CTCCACACAGCATCCTCACCTCCGCGGGCCCTCCCTCCAGCCCGCCCGGCCTCCACA
CAGCAGCATCCTCACCTCCGCGGGCCCTCCCTCCAGCCCGCCCGGCCTCCACACAGC
ATCCTCACCTCCGCGGGCCCCTCCCTCCAGCCCGCCCGGCCTCCACACAGCATCCTCA
CCTCCGCGGGCCCCTCCCTCCAGCCCGCCCCGCCTCCACACAGCATCCTCACCTCCG
CGGCCCTCCCTCCAGCCCGCCCGGCCTCCACACAGCATCCTCACCTCCGCGGGCCCC
TCCCTCCAGCCCGCCCGCCTCCACACAGCATCCTGACCTCCGCGGGCCCTCCCTCCA
GCCCGCCCGGCCTCCACACAGCATCCTCACCTCCGCGGGCCCTCCCTCCAGCCCCC
CCCCCGCCTCCACACAGCAGCATCCTCACCTCCGCGGGCCCCTCCCTCCAGCCCGC
CCCCGGCCTCCACACAGCAGCATCCTCACCTCCGCGGGCCCCCTCCCTCCAGCTCCGC
CCGGCCTCCACACAGCATCCTCACCTCCTCGGCCCTCCCTCCAGCCGCCCCCCCCC
CGCCTCCACACAGCATCCTCACCTCCGCGGGCCCTCCCTCCAGCCCGCCCGGCCTCCA
CACAGCAGCATCCTCACCTCCGCGGGCCCTCCCTCCAGCCCGCCCGGCCTCCACACAG
CATCCTCACCTCCGCGGGCCCCTCCCTCCAGCCCGCCCGGCCTCCACACAGCATCC
TCACCTCCGCGGGCCCCTCCCTCCCCGCCC GGCTCCACACAGCATCCTCACCTCCGC
GGGCCCTCCCTCCAGCCCGCCCGGCCTCCACACAGCATCCTCACCTCCGCGGGCCCCT
CCCTCCAGCCCGGCCCCCCCCCCCCACCCCCGCCTCCACACAGCATCCTCACCTC
CGCGGGCCCTCCCTCCAGCCCGCCCGCCTCCACACAGCATCCTCACCTCCGCGGGCC
CTCCCTCCAGCCCGCCCGCCTCCACACAGCAGCATCTCACCTCCGCGGGCCCCTCCC
TCCAGCCCGCCCGGCCTCCACACAGCAGCATCCTCACCTCCGCGGGCCCCTCCCTCC
AGCCCGCCCGCCTCCACACAGCAGCATCCTCACCTCTGCGGGCCCCTCCCTCCAGCC
CGCCCGGCCTCCACACAGCATCCTCACCTCCGCGGGCCCCTCCCTCCAGCCCGCCCCG
GCCTCCACACAGCATCCTCACCTCTGCGGGCCCCTCCCTCCAGCCCGCCCGGCCTCCA
CACAGCATCCTCACCTCCGCGGGCCCCCTTCGCTCGTGGCCGACCTTTCGATGCTCC
CTGGCCGCCTGTGGTCCCTGGCCCTGCCCGTCGGGCGCCTCTGCTGGGGAACCAGTG
GGAATCAGCTCAGACACCACCATAGGGGCCCTGTCTACTGTGCAGGGAACCTGAC
TTAGCCCCAGTGAACAAAGACACTTTATGGGGAGACAGGATGGCTCCCTGGGGAG
CGACTTCCCAGAAAGCCGACCTCACCTCTCTGGGCAGGGCCCTCAGCGTTCTCCACG
CCAGGGACCCGAGGTCTCCGAGAAGGGTCCTGCCGGGAGGAGCACAGTCAGAAA
CAGGGAGACGGGTCCACCCGCCCAAGTTCACCTGACCCTGCTCACCAGGCTGGGCA
GGGCAGGCTGCTCTGGCTCTGGAGCCTTCCCTGCCACCTTCTCTGCTTCTTCAAGTC
TGTC AAGGTCACGCCCTAGTGTAGAGAACTCGGTCAAGGAAAAAGGGCCTAGACTT
CCACAGGACTCAAGTTCAAGACCCCGGCCCTCCTCCCTCAGACCCGGGAGCACAGC
CCCAGCCCCATCCCACACCTGTGTAGACCTCCGAGACTGGCGCATAAGGCGGGAGC
GAGCCTTCCGCTGAGACTCGGACTTTCATCACGCACAGGCATCTGGTAGGACCTGA

GTGGAGAATGTCCCCTTGGGTGCTGTGCGCACTGTGAATGACACCTTAGGGGAGTGCA
CATTCTGGCAGAGAACGTGTCAACTGGGCAAACAGGACCCCAGGAGCCTACCCAGA
GCCCCAGAGACCCCTAAACACTGTCTTCCCCTAGCCTCTTTACCTCCGCGGATCCCTG
GAGTCAGGTAGGGCTGCTGCAGAGGCTGCGAGGGCATTGGCTTCACTGGGGATCC
AGGCTCCCTCCTGGAGGATGGGGCGGAGTGATCCGAAGAAGGAGGCATGGCAGCCT
CACACCTGTATGGATTCATTCATTCATCAGCAAATATTCCTCAAGCCCGCATTCTGTG
TCAGGCATAGGAGAGACCACAGAGAAGGAGCCAATCATGGCTGCTGATGAGCCATT
TCTGGGCAAAACAGATAAAACAAACAGCAGCCAAAGAGACCAGTGTGGAGCTTGG
GGAGAAAAGGTGCTTGGAAAAATAAAGAGAATAAGCAATTATTTGATGCACCCTA
AGGGCTTTCTCAGATCTCAAATGCCAGGATGGCACCAGACCTGTCCCCTTGCCCCAG
CCACTGGTACTTACAGGGTAGAGGGCTCTGGCACTTTCTGGGCAGGGGTAGGGGTT
ATTCTGGCAAGACGGGGTTCCCTGGCCTGTGGAGACAGGAGAGAAGCAAAGGAGGC
ACTGTCTGCCCCAAGGCAGGAGCCTGTACCCACACACTTCACGGCACCTACCTGAG
AGGAGGCCTTTTCTAGGAGGGAGGAGGAGGCTGAGCGCTGCAGACCGAGAACCCCC
TCTGCACCCCTCCTCTCTGAGGGACCCAGGGCACCAGAGCTTCTGTCTTCTGGAGA
CCGCCGCGCCTGGAGAAGGGAGCCTCTTCTGGCGGCTGGGAGAGGAAGAAGGTCTT
CATTACTGAGCAAAGCAATGACCCTTCTCCTCAGAGCCTACGCGTGTA ACTCCAGGG
GAATTACAGTAAACCACAGCCAAAGCAATGACCCTCCTCCTCAGAGCCTACGCGTG
TA ACTCCAGGAGAATCACAGTAAACCACAGCCAAAGCAATGGCCCTTCTCCTCAGA
GCCTACGCGTG TAGCTCCAGGGGAATCACAGTAAACCACAGCCAAAGCAATGACCC
TTCTCCTCAGAGCCTACGCGTG TAGCTCCAGGGGAATCACAGTAAACCACAGCCAA
AGCAATGGCCCTTCTCCTCAGAGCCTACGCGTG TAGCTCCAGGGGAATTACAGTAAA
CCACAGCCAAAGCAATGACCCTTCTCCTCAGAGCCTACGCGTGTA ACTCCAGGAGA
ATCACAGTAAACCACAGCCAAAGCAATGGCCCTTCTCCTCAGAGCCTACGCGTGTA
ACTCCAGGAGAATCACAGTAAACCACAGCCAAAGCAATGGCCCTTCTCCTCAGAGC
CTACGCGTGTA ACTCCAGGAGAATCACAGTAAACCACAGCCAAAGCAATGGCCCTT
CTCCTCAGAGCCTACGCGTG TAGCTCCAGGGGAATTACAGTAAACCACAGCCAAAG
CAATGACCCTTCTCCTCAGAGCCTACGCGTGTA ACTCCAGGAGAATCACAGTAAACC
ACAGCCAAAGCAATGACCCTTCTCTTCAGAGCCTAAGAGTGTA ACTCCAGGAGAAT
CACAGTAAACCACAGCCCAGGCAGGTGCCACCAAAAAAAAAAAAAAAAAAAAAACAT
TACTTCTTGGTCCACAAGGACCTAAGAACCAAGTCAAAAAGCCACTTTCCTCAGCGG
AAGCAGAAGTATTTACCGTATCCCACCCGCTGCCCCAACCTCACATCTGCTCAGGG
CGCTCAGGCTCACACAGGGCTCTTGGGGCTGGAGGACACAGGAGAAGACACGCCA
TTGAGGGCTCTTGGTTGCACAGGAGGGTGATCTGTGTGCAGGAACAGGAGAGGGGG
GTCACAGGAGAGGCCGGCCGCTCTGAGATTGGGGACCCACAAGTCCAGCTCCTTC
CTCAGACCCAGGGTCCAGCATCCCTACCAGCTGCCTCTTCTTCTCCCTCATCCTCATC
CCCAAGAGAGGGGGCCCGCGGCCCCACCAGGCCGGCGCTCCTTCGACAGATCCTGCA
GGGAGATCTTCTCACGGCTGCTCAAACGACACACGGAGCTCCTAGGAGGACAGGGT
GTCCGTGTCCAAGTCTGGGGGCGAGTCCGACCCACCCAGGCCTAGGCATCTCTTAC
CTTCTGTGCTTGCTGTTGGAGGGCACCTGTGGCTCTTGGCCTCGGCTCTGAGAGGCTT

CCTTTTGGTTCCGAAGCTACAAGGATGGAAGGGGGCAACTGGGGAGGGGCAGAGAG
CACAAGCCCTCCAGGGTCTCCTGGCCGCCCCCTCTGTGCCACCTCTCCACCTCGAGG
GCCATCACGCATAACTGGGCTAGTCACACTTTATGCAGGGTCTTGCAAACATGGGG
GACTCAGTAACCCGGCAGCACACTGGCTCTGGGGCTTATTCAGGCTCTCCCAGGCTT
GGCCTGGTCCAGCTGTCACTGCCTCCAGCCTCATTCCCAGGGGGATTCTGTCTTCTTCC
CAGGAGCGAGCACCTTGCTCAGACTTCCCCCTACCCTCCAGCACATCCAGGGCAGG
ACAGGGCAGGTGGCTCTTTCTGGTTATCACAGGCCAGCTCTCAGCTCAAGGACAACG
GCCACCGTCCCATACTAAGCAGTCTGGTGTGCGTAACCCCAGGAACACCTCTTGCCCA
TGCCCTCCTTGCACTCCAGTGTGCCACGGGACTCCTCTCTGGACAATGTTCCCGATG
GTTCCACGAGGCCCGGGCCACCTCACTAAATAATGGAATTGCAGCCATGCCGTCTGC
TTGGGGCCACACCCATGATGCCTCACTCTCCACTTTCCTAGCAAAAGTGCTAACTAG
AGTGGGGGGGGGGTAGATACAGGTTCAACCTGTGTACACACAGCTGTCTTCCCAA
GCGAGCAGGCAGGAACTCTGGGCATAGCCTCAAGTCCTCCAGATATGGAGGTGCC
TCTGTTCTTAGCCCTCCACCAGAGCTGGGCTGACAGGTGGGAATAGCGGGTCTCAGT
ACTGAGGGTGTCAAGGGACAAAGACTGTCAGCCCTCCCGGTTACTGTTACCTCCTCA
GAGCTGCCAAGTAAAGAGGCAAACTAGAGTCGAGACTCACGTCCTCCTGTTTCTGG
GCCAGTTCCTCCAAAAGGTTTCATCACTTCCTCATCAGCCAGGTCACAGGGCCGCTGC
CCCTGAGTAGGAGAAGGAGGCAGATGACGGTGTGGTGGTGGTGTAGTAGGGGCTC
CCCCGCCACCCTGCCCCACCATCTGAGATGGCCCTTACCGCATGGGTGAGCGAATCC
ATGCCCCACCGTGCTCAGCCAGGAGACGGCAGGCGTCCTCCACACCCCAGTGGGC
TGCTGCGTGCAACGGTGTCCAGCCATCTCCATCCCGGAGCTCTGTGTGCTAGCCAGC
TTGGAGTAGCAGCCTAAGGGCCAGGGAGGCTTGGGTGAGATGGCAAGCTAGGCCAA
TGGCTGATCTCAACTTCTGTTCTGTGGCCACAGGACTACTGATCAATAACCCAAGCGT
TACTAGTTTTACCAGCAACCAGCCCCACCCCAAGCTCAACTGAGCCCTCCCTTGGAC
CAGCAGCTACTAATGAAAAAGCTCCCTCATAACCACAGGGATCCCCTCCTCAGGCCC
CAGGGTAAAGGGTTAGGGCAGTGGTGAGGCGATGAGGTGGATGCAGGACTCCCCAC
TAACGCAAGCCCATGGAGAGGATGGACCCTGAAGGGGCTGTGATGCTGGAACCACT
GGAACCACGCGGTTTTAGGACACGGATCCTCAACAGTGTCAAGCAGCTCTCACACC
CTCTCTACAACCTGGAGACATCACCCTAGAATCCTAACTTACGGGTACAAGCAGGA
AGCACCAGTGTGTGGGAGCTGGAGAGGCTGCTCAACCCCCTCCCACGCACAGGACA
GCCCTACCACAGCACGGTAAGACCCCAAACATCACAGTGCCGGAGGAGAGCGAGCC
TGGCTCAGCCTTCCAGAAGGTAACAACCTGGAGCTCTCAAAACTCAGCATGGCAGC
AGGCGAGGCCTCTTTTGAAGCAGTGTGATGAGGTCCTGTGTGCTGAGGAAGGCT
TCAAGCCCAGGGAGGCAGAGGTACAAGGCACAAGGTGCTGTGTGGCCCTGGGACCC
TCCTCCCTCACACTTCCCAAGATTCCCCTGTCCCCTTGCAGCAGGGCACGCTGGGCTT
CTTGTTACATTCCCACATGCCAGGGTCTCTAGCCAGCTGTGCGCTCCTTCTGGTCAGT
ATCCTAGGAGCCTGAAGCGTGCCACCCAGCCACACCCCCTAGTCCATCAGCACTTCC
TCACCTGGCAGTTTCTTACCACCATCTCTGCCAGGGGGCCTCCCTACTGCCCACTA
GTTATAGCCTCCCAAGGCCAAGGTTTTCTTTGTATAAGCTTAGTGTTATTTACCATTA
GTGTGTGTGTCTGT

GTGTGTGTCTGTGTGTGTCTGTGTGTGTCTATGTGTGTGTGTGTCTGTGTGTGTCTGT
GTGTGTGTGTGTCTGTGTGTGTGTGTGTGTGTGTCTATGTGTGTGTGTGTCTGCGTGTCT
GTGTTTGTGTGTGTGTGTGTGTGTCTGTGTCTGTGTGTTGTGCATAAATGCCAACACACA
TGCCCCAGTATGAAGATCATGGATGAAGATCAGAGGACATATTCAGGATTCACCTTC
TCCTTCCACCACCGGTTCCAGGACCTAACACAAGTCACCAGGCTCTTGTGTGGCCAA
CACTTTTACCTCTGAGCTATCTCACTGGTCTAGAAGCCAACGTTTGCAGCTGGACCC
TGCTACTCCCCAGAGGACCTGTGGCAATGTCTACAGTCATCACACA ACTGGGTCAGA
GGTGCTGCAATGGACTGGACAGCCATCAGAATAGAATGACCCAGCCCATCAAGTCT
CTCATTGGCTACGGTGGGTACACATCTGAAACACCACGACCAGCCCAGGAGGCTAG
CCCCTAACAGACACCAATATTTACCTGTACTTCAATGAGTACAATCATAGAAGACTT
TTAATACAGTCAGAAACAATAGATAACTATAAATTCAGTGAACAGGAGTCTAAACG
CAAACCTCACACAAAGGGGGCCATCACAAAATTACAAAATTCAGTATGATGGCTCAC
ACCTGCAATCCCAGAACACAGAAGCTGAGGCAGGAGGACAGCTGTGAGTGCAAGG
CCAACCTAGGCTATCTATCCAGTACCAGGCTAGTCAGGACTACATAGCAAGACCTTG
TCTCCATTAGAAAAGAAAGAAGCCAGAGGGGAGGGAGGCAAGCATGGTGGCTCTC
ACCTCTATCCCACAGGAAGGTGAAGGAACAAAGAGTAGAAATTCAAGACCAGTGAA
CTAGAGGCGATCATGACCGACATGAGCTATTTATGGAAGAGGCCAAATAAACAAC
ACAAAAGTTGTCATCAGTGCATTTTTTTTTTCAGGGCTGGGACTGGAACCCAGAACG
CTAGGCAAGTGCTCTATCCCTGAGGCACCCCCCTTCCCTCACGGGTAGACACCAG
GGAAGCATCTATCTACCTATGGCCTGCGACCACAGCCCAGTGCTTCAGTTCTGGGAC
AAGTATTGGCTCACTTTCTCTACTAACTAGCCCCCGGACCTATGCAGGTGACACCG
GGGAAAGCATTTAAGCACAAAGACAGGAAGGAGTTCTGATCACCAGAATCCACTTA
AAAACCTCAGTGGATAGCTGTTATAAAAAAATGACATCAGGGTGGAGAGAGATAGAT
GGCTCTGCTCTTCCAGAGACCCGGGTTCAATCCCAGCACCCACACGGCAGCTCCAG
GGGTTCTGACCCCTCACACTGACATAACACAGACAGGCAAAGCACTAATTAATGCA
CATTAAAAAAAATAACATCATGAAATCTGCAGGCAAATGGATGGAACTGGAAAAAA
AAAAAAAACATCCTGGGTGAGGTAACCCAGCCCAGAAAGACAAACA
TGGTGTGTACTCATTTACAAGTGCACATTAGCTGTTCAAGTGAAGGACAATCGTGCTA
CAATCCACAGACCCAGAGAGGCTAGGTAACAAGGAGGGCTCCGGGGAGGGGACGGT
GCACGGATGCCCCAGGGAAAGGGAAAGAGAAAAGACTTTGCAGATGGAATGGGCA
GGTAGGGATGGAAACAGGAGAGGTGGGGAGAGGGAGTGGAGGGGAAATACTGGGG
GGGGTGGCTGCAATGGGGGCTCACTTTGGGGGTGTTAAGGAAACCCAGCACAGTGG
GAACTCCTGGACTCTGCAAGGGTGGACCTAGCCAAGTAACGAGGGACACAGAGTCT
GAACCGGCTACTTTGGGTAACAGGCAAGGCTCCCAGCAGTGGGACATCAACCCGGC
CACAAAACCTTTTGACCTACGATGTGCCCTGCCTGCAAGGTGTGCTGAGGTAATGGTG
GCGCAGAGCTTGTGGGAGTGGCCAACCAATGACAGGTCCAGCTTGAGGTCCATGCC
ACAAGAGGGAGCCACGCCTGACACAGCCTTGATGGCCAGGAGCCTGGATAGCCCCG
AGACCTGGGGTAGAACCAAAATACAATTGGGGGAAAAGAAAAAAGGCAAGAAACA
ATTCTTAATGATATTCTGCTGTTCTCATGGATCTGTGGCTAGCCCAACTGTCGTCAGA
GAGCTTTTTCCAGCAGTTGACGGGAGCAGATGCAGAGACCCACAGCTCAGGGAACC

CCACAGGAAGGATTATGGGGGGGGGGGGCGCGAGGACACCAGGAGAACAAGCCC
ACAGAATCAACTAAGCAGGGCTCCTTGGGGCTCATGGAGACTGAAGGAGCTAGCCA
TCAGGACCTGTATGGGTCTGCGCTGGGTCCCTCGCCTGGTGCTCTTGCGGGACTCCTT
AACACTGGGACTGGAGCTGTCGCTGACTCTTGTGCCTGTTTGGGGACCCAGACAAGC
ATAACTGGTTACGCTGTGCTTGGCTGTCATCTCTGAGATGCCTGTTCTTTTCTGAAGG
GAAACAGAGGACTGGATCTGGAGGAGGGGTGGAGGGGAACAGGGCAGAGGGGAG
GGAGGAATGTAATATGAGAGGAAAAACAACAATAATTATTGAGTGGACATGG
CAGCCCATCTGCAGAGACAGGCCACCCTCAGACGGAGATGGCAGCTAAACTTGCCA
AAAAGGCAAGCTGAGGGATCGGCCAGAGGCCCTGCCTCAATATTAGAGTGGAGAGC
AACCAGAGAAAGTACTACATGCCAACACACACACGAGTGTGAACACACACACACAC
ACAAGTCATACCCATACACATGCACACGCGCGCGCACACACACACACACACCAC
AACCGTTAACAGACATATAGTTGTGTGGAAACAAACCTAGTTTTCTTGCAACTAG
GACTGGCCAATGGTGAGAAGTGGGTTAATGGAACACAGATATTAATATGCACACT
TCTGGAATGTTCTCCTGAAAAGGAATAGACATTCGCTCCCTTTGCCTCTGCTTCCCAC
CAACTTGAGATATAGACGCAAAGGCAGGTGAGGCAAGTCACCCTCAAGTGAGAGGC
ACCGCTAGAGCAGGGCGCAAGCTCTGCACTCGGAGATTTAGGGCATCCTGTCCCCC
AAAAGGAATGGGCTCAGAGCGCACTGGGACTCATGCTGTAACACTACAGAGACTGATG
CCCCTCCCCCAGGAGCACAACCTATGCAGGCAGGCTGTAAGTCTGGGGGTGGCACGA
GGTCTTAAATCCTGCTGGAGAAAACCTGCCTGCAACCTTACCAGTATGAAAAGCAG
AGAGGTTTCACTTAATTCAATTTGGGTCTTTGTTTTTTTTGTTGTTTTTTTTTACAACAG
GATCCCTCTATAAAGCACTAGCCTCACACTCAGTATATAGACAAATCTATCCTGGAA
TTCCAGTAATCCTCCTGCCTCTGATTCTCAAGTGTAATTATAGACATATAACACCGTA
TCAAGCAAGCAAGTGCACACACGCACGCACACGCTCTTGTTACATAGCCTGGGCTA
GCCTACAACCTCACAGCAATCCTGCCTCGACCTCCCAAGTGAGGAAATTAAGCGT
ATACCACCATGCCTGGCTTAATGCCATTTTTTTTAGGTTGGTATTATTTTTATGCGTAT
ATGTTTTGCCTACATGTATGTATGCATACAAATACACACAGACACAGAGATAAATAA
ATGTAATTTTTAAACCTCTTTGGCTTTAGGTATGTAAACCAGGAGAAGAAAAGGACA
AGAGCCCCGAAAAGCTTCCAGACACAAAACAATCACTCTGGCCTCGCTCACCTCATC
ACCTCGATGTAGCCCTTGGCGGCAGCCACATGCAGGGCAGAGGCCCGGTCCGGGG
GTGGCGGGCCTCTGGCATGGCACCCCCATTAGCCAGCACCTTGTGTCATGAAGCAG
CAGTTCTTCTCAGCCCGCTTGGCTGCCTCGACATCCACACCTGGGAGAATGAGAGG
TGACAGGTGGACTCACACAGGGTGGCCTAGGAAACCCCGGCTGCGGTCTCAACTAG
TCACAGCCCGGCCCGTGACTCATCAAGTCTCTGGACCACTCAGGAGACCGGGACT
GCCCCAGTGTTTCCCAACTGTGCTCCCTGAAGACCTGGGCACCACCGAGGGGGCCA
AGACAGGCCAGGAATGGAAACCACAGGTCCTGACCCCTGTGGGTGAGTATCCTCTTT
ATGTTTTTCTAATAGAAAACCCACACCGGATTCCATCTAGGTTTTCTACCCCTCCA
GCTATAAGCTAAAGCCAGCGCCTTACACAATGTCACTGCTGGTTCTTCTCCCTTTG
AAGTACGATAGGCCAAACAAAACCTTCACTACGGCGTTGTACGTGGTGGCTCCGGCC
TCTATTCCAGATCTCAGCCTTGGCAGGATGACCGGTGGCCTCGAATCTGAGAACAGC
CTGAGCTACATACATGGTGTCAAGCCAACCAGGACTAGAGAGACAGACTCTGTCTT

AGACAACAGTAAAACTAAAACTCAAAGCTTCTGGGCGGTGGTGCACACCATTAA
TCCCAGCACTCGGGAGGCAGAGGCAGGCGGATCTCTGTGAGTTCGAGACCAGCCTG
GTCTCCAGAGTGAGTGCCAGGATAGGCTCCAAAGCTACACAGAGAAACCCTGTCTC
GAGAAAAAAAAAAAAAAAAAAAAAAAAAAGCTATTTCCCAAACCTATTTGCATGCATAG
TTTCATTCTTGCCCAGATGTCCAGGCATTTGACACCTCGCTGGCCACGACAGAAGT
GAGAAGTGAGTGAAGTGCCTTGGCACTTTGTGCTTATGCGGGTATGCTGCATGCCTGT
GACCCCAACACAGGCAAGAGGCAAGAGACCAGCAGGGCTCCACAGAGACCCTGAG
TCAAAGACAAACAGAGGGGGAGGGGCTGGAGAGATGGTTTAGAGGATGAAGTGC
CAAGCCCGATGACCAAGTTCAATCCTGGGAACCTCATGAGGCAAAGGAAAGAATCA
AGTTGCACAAGGGGTTTCCCTTTGTGAACCCAGCTTGGCCTCAAACCTCACAGCAATC
CTCAGTCTCTGGAAAGCTGAGATTAAGAGGGGGTTTTTTTTGTTGTTTATTTGTTTTT
TTGGTTTTGGTTTTACGAGACAGGGTTTCTCTGTGTATCTTTGGAGCCTATCCTGGCA
CTCGCTCTGGAGACCAGGCTGGTCTCAAACCTCACAGAGATCCACCTGCCTCTGCCTC
CCGAGTGCTGGGATTAAGGCGTGCACCACCAACACCTGGCTAAGATTAAGAGTTA
CACACCACACCTGTACTCCCTGCACTCAAAGAGGCTGAGGCAGGAGGATTGCTCCA
AGTCCAAGGTCAGCCTGGGCGCCAGCATGAGAGCCTGTCTCAAACACCTCAGGGGG
GAAACAGAAAGCCAGGCAGACTAGCTGAGGCTGAAGCACTCCAGCCCTCACTGTGA
CCCTCATCCCTTAAAGCACCCCTAACTCACTGAGACCACAGCAAAATGGCCTCTGCT
GAATAACTTCCTCCTGGGAAGGTTATTACTGCCCATGCTTTTGCAGTTGTGAAACTCT
TGAAGTGGCGAAGTTCCTCAGAGTTGAGCTGTTGTATCCAGTAGCCGGCAGCTATGT
GGAAGTGGCGAGCGAGCACTCGAAACTGAGACATGCTGTGAACGTCAAGTGCATC
GGGATTTCAAACACAGGAAAAGGTAAGGGCTCTCGTGGACAGTTGTCTCTAAACT
GTTTTGTGCACACGTGCATGGCAGCGTGGTCAGGGGGCAATTCTGAGGAGCCGAGT
CCTGCTTCCCACCTTGCTGAGGCTGGGTCTCTGGATTCTGCGGCTGTGCTGTGTAATC
TAGGCTACCCGGCCACAGGGCGTCCCCACAGTTCTCCCACCTTCTCTAGGACCTG
GGAGTACGGACGTGCACCAGTGCCGCGCGGGCTTTTTACAAGGGTTCTAGAGGTG
GTAAGTGGGTGCATCTAACATCTTTACTGGCTGAGCTATCTCCCAGTTCCCCTTAC
TGTGTTGATTGCACCCAACGGAATACTGGGTTTGTTTTTTTGTTCTGGAGCGTGCGTGA
GAG
AGAGAGAGAGATACAGACACAGAATCTTAACAGAGAGACAGAGTCTTACATTATAT
AGTGGAGGCTGAATATCTAGTCTTCTCCTGCCCTCCCAAGTACTATGGGATAGTCTTCT
GTACGCTACGAATATGGACTCTTCTCATTGGTTAATAATAAAGCTGAGTTGGCCAC
AGCCAGGCAGAATAAGGGTAGGGCGGAAAGCCAAACAGAGATACAGGGAGAAAGA
AGGGCAGAGTTGAGTGAGACGTAAGCAGCCACCAGGGAAGCAAGATGCCAGGTGA
CAGGTAAAGCCACGAGCCATGTGGCAAACACAGGCTAATAGAAATGGGTTGATTT
AAGTTGTAAGAGCTAGTTAGTAATAAGCCTGAGCTATAAGCCGAGCATTCCGTAATC
AATACGAGCTCTTGTGTATTTATTTGGGGCCTGGCGATTGGAACCTAAAGGGAAAGCTT
AGACTACGGACTTGCACCTCCATGCTTAGTTTATGGGTTTACAGAGGACCATGCTAATG
GATAAGCACTCTACCAACTAAGCTACACCCCCAGCCTATGGCTTGCAAGTTTCAAAC
TACATCTGTGGCTCATTACAGGATTTCCACTGGGCGTCACTGGCAAAGGCCTTCAGGT

CCCACCTGGAGCGCTGGCTCAGCCATTAGAGCCATTAATGGTAGACTCACAACCTAC
ACAAGAGACAAAAACCCACACAAGGGGTGGAATGCAGAGACTCAACCAGTTCCAA
GCCAGCCGGGACTAACAAAGCAAGATCCTGGCTCATAAACCCAGGAGCAGGGTTTA
GCCCAGTGGTGGTATGCCTGCCTGGAAAGGGATGGCCCCAGGTTTCAGGCCTCTACA
CAGAGGGCTGCTTTCCTCACCACACTCCCTCTTAACCAAGGTGAGCAGCCGCTCCCC
TCAGCACACACATTGTACTGCCACCATAAAGCTTTACATGGGACCCAAGAAACA
GTCCTGAAAGCTGGTTCGGGATGTTCTTTCTCATTGCAAGGCAAGGCCAACTCCATG
CGGACACCGGCTGCAGCTTGGTGCTACCTGGCGGCAGCCGGGTCTAGCTCCTTGTG
TCTCCTGGCCA ACTAGGGTTTCCCTTGTGGTGGCAGAGTTCAAGAATGCATGGCGAA
AGTCCACCCGCAGCACAGTCACAGGGAACAGGGCAGGGAGGGCCAGGCCCGCCCTC
GTCCTCCAGACTCCTGCTTCCTTAAAGGGAGTCTCCACAGTTCCACCTACTGTGGG
GGGAAGGGGAAGGGGAAGGGCGGAGCTCCCTTGCTGTTCTTCAACCACCAGCCAGT
ACTCCCGTGCAGGCTCAAGGGCAGCCTGTGCTCTCCACACAGCCAAGACCTGCTTGC
TTGTTACTCAGTTTTTCTTACACAGGCCGTTAGCTGATTAATTGGGTTTTTATTTTATG
TGTATGGATGTGTTGCCTGTGTGCATGAATGTATACATGTGTGCTGGTGCCCGAAAA
GGCCAGAAGAGGGTGTGAGATTCTCTGGAAATGGAGTTACAGGTGTCATGTGGGT
GCTAGAGTTGAACCCGAGTCTTAACCACCGAGGCATAGCCACTGATTCAGCCAGAC
TGACCAGCCTGCAAACCCAGGGATCCTCTGTATCGGCCTTGCCCCACACCTGCTAG
GATTACAGGTGGTGGTGGGCTTGGCTTTGTGGTTGCTGGGGA ACTGAACTTAAGACC
TCAGTATGTGCACCAAACCCTTCTACTGACTTAGCAACATTCCCGTGGAAGTCCTGA
AAATGAAGGACGGGGGGAAGGATACTGACTCTGCTGTAAGAAAATTCTTACATTTA
TGTTATTGTGTGGATGTGTGTGCACTCAAGCACAAATGAGAGCTAGAGAGGGCCTG
CAGAAGTCACTTCTCTCCTCCCACCAAGCGGATCCCAGGGACTGAGCCCAGGGGTC
AGGCTTGGTGGTGTGAGCGCCTTCGCCC ACTGAAGCACCTCACCAGCCTGAAAATAAA
GCTCTCATGGCACCCAGCCACGTCTGCTTTTTCTCATCCCGTCGCTGGCTCATTTCCC
CCGATAGCAGCTGGTAGAGTCATTATAGCAAGACCGTGCACCCTGTAAGGCCTGAA
ACACATACTGACCAGCCCTCCACAGGTCCCAGCTGACTCCTGCTGGGACCACTGAGT
TATAAATCAGAGCGTCATCTACCGGCTGCAGAGGGCGACAGCTTTTTGGTGTACCAA
CAGCAAACACTGTGCTGTATTCTGTGCACTCACCACCTGTGAGAAAATGCACCAGG
GCAGGAGCTCAGGCCTGCAGCTTCAGCCAGGTAAAGGCCGGCCTGAGCTCTCCTTCC
AGGCACAGCCTTGC ACTACTCTAGCTGGCATCTGTA ACTACCGCAGTCCACTGTGCC
CATCTCTGCATGCTACAGCCCTCACTGTCCTTCTGGATGTCAGTTTCCATGGGAGAC
AGCTTCGCCTTTCTCCAAAGCACATCCTAAGTCTCCTTCTACCCTGTCCCCCAAGG
GGGGCTCCTCCTCCACGGACACTGTGCTTTCAGTCTTTGCCAGGGGTCTGACCTAGG
CCTGGGCCGCACCAACACTGCTAGGACCTGGCAGCACCCACTCTTCTCCTTCAGGA
ACCACGTCCTGACTTTCCTCGCCCACAGGGCTTCAGTGTGACACTTGT CACAAGGTG
ACAAGTCCTCCTGCACACTGGTGGGCTCTGGGACTGACATGAGATCATGTGCCAGTG
TCACACAGAGAGCTGTGGCTCAGCCACTGAGGGGGGTAGGGCTACTGTACCCCACC
TAGGCATTAGCCCTGCTAAGCACACAGGGAGAGACCAGACCCCAACCAGGAGGCCA
AAGCGGCTGGTGGCATTAGGGATCCACTGCCAAAGACTGGA ACTCTAGGGCTTGA

CAGACAGATGGCTCAGTTGGTAAAGTGTTACCACACAAGCTGCAAACCTGAGTTC
AACCCCCAGCACCCATGTAAAATGCCAGGCATGGTGGAGCATGTGTAATCCCAGTA
CTGGGGAGGTAGAGACTGGAGGACAACCGGGGTTCACTGGCCAGCCAACTAGCCC
AATCGTGGAAGCTACTGAGACACCCTGACTCAAAAATCAAGGTAGACGGCTCCTGA
CGAACATTTCGATTGACCTCTGGTCTCCAAACACACCTGTGCACGCACACATGCACAC
ACAAACATATGAAGGACTGAGATTCTGCATACGTTAAACAGCAACTCTCTCCTCCCC
CTTTTTATTGCATTCATTTACTGGTGTGTGTGGCTGAGAACAATCTATGGGGTTCAGTT
CTCTGGTCCACCAAGTGGGTCCAAGGAATCAGACTCAGTTGTAGGCTTGGCAGCAA
GCACCTTGACCCACTGAGCCATCCAGCCAACTCTCTTTTTTGTGGTTAGTTTGTG
GGTCAGAATCTCTCTGTGTAGTCGTGGCTGTCTTGGAAAGCCACTCTGTAGACCAGGC
TGGCCTCAAACCTCAGAGATTCACCTGTCTCTGCCTCCAGTGCTGGGAATAAGGCATG
TGATAGGGTTAAACCCACCACACCCAGCTCTACTCCCAGGGTTGTTAGCGCCTGGAA
ACCACATTGGATGCTCTGTGGCTATGACTTTGGGAGCGCTATTATTTTCATTACGCTT
GGTGTCTCACACAAATGGAAACCTCGCACTTGTCTTTGAGTGGCTGACTTGCTCCAT
GTAGGTCCTTAAGTCTCATCAAATCTGGCAGGCGGTGGTGGCGCACGCCTTTAGTC
CCAGTACTCAGGAGGCAGAGGCAGGCGGAGTTCGAAGCCAGCCTGGTCTACAGAGA
CAGTTCAGGACAGGCTCCAAAGCAATACAGAGAAACCCTGCCTCAAAAAACCAA
AAGAAAGAAAAAGTTTCA
TCCACATCGTTAGACGTGTTGGTTTCCTTCCCTGTAAAGGATGAGCAGCATTCTTTTA
TCTGTATAGACCACATTTTGCTTATCCAGCCAGTGATGGACTTCCTACCCTCCCTGGT
CCTGAGGCTGGAATTCTCATAACCAGACCAAGTGCCACAGGCTGGAGTACTAAGCAA
CACCCAAGTCCTAGGCTACTGTAAATACGCCTGCAGGTGGCTTTCTCTTTGCCAAC
ACCACAAAACATAAAGAAGGGAGCCAGACATAGAGGCAACTACTACAATCCAGAC
CCTCAAGAGGCTGAAGCAGAAGGATCCAAAACCGTACTACAGTCAGTTTCGCCGCA
GCCGAGGCCAACTAACTAATTAATTAACAAAATAATACGTATTGTGGGTGTGCATCT
TGCGATGCTTGTGTTGTGGTGTGGGGATTGAACCTGCCTACGCTCAGCAAACGCTCT
GCCAGTATGAAGTGTCCAGTCCTCTCCACACACACGCGAACACACGCGAACACAC
GGGAACACACGCACG
CACACGTATATTTAAGATCTTTCTCTCTCTCTCTCTCTCTCTCACACACACACAC
ACACACAGGGTTAGTTAAGACCTTATTTGTATTACTTTTAATTGTGTGTGCGTGTGTC
TGTGCAGGGCACGCACACACAGTGTTCAGTACCGGAAGAGGTGTGGGATCCCCAG
GTGCTGGAGCTACAGTGAGCCAGACACTGGTGCTCTGAACTGAACTGAATCCTCTGC
CAAAGCAGAAAGCACTCTCTGGACTCCTGCTTTTGTGGTTTTTGTGGTTTTTGTGGT
TTTTTGTGGTTTTTCTTTGTACTTTTCAACAGATTCTCACTAAACTGTCCAAGTTGGC
TTGAACTCCCTCTGTAGTTCAGGCAGGCCTTGAACCTCACAATTCCTCAGGCCTGGGA
GAGCGGAATCTTTTCGTCAAGAAAACACCCCATTTGAAAACCTGAGGAATGCTGTATC
AGCACAGGAAGGGGGAAGCTCAGGCCTTGTGGCTTAGAGAAGCGGCCTTGTGCCAT
GGGGTTAAGAGCCCCAGGCTGCCCCATCTCGTTGGCTGAGGGAGGCGCTTCCCGTTA
TCTGAGCAGAGCTTCCAGCATCAGCAACATAGTCTCCAAGTGGCTGAGGAATGGAA
GGAGGATGGACTGGAAAGGAGAAAACAGGAAGGGTACTGCCGCCTGTGTGTGGGG

AAAGGGGCAGAGCTGGACAAAACAGTAAAGGCGTCTATTTAAAGTGTGCGATTCCA
TCTGCACGAAATGTCCGCAACAGACAGATCCCTAATGAGAGAAACCTTAGAGGCTT
GCCCAGGGATGGGACAGGGGACTAGACTTTTGGAGGGTGACTTTAAAATGCTCTGAA
ATCAATGTGGCATCGACTCCGACAACCTCTGTGGACCACAGCACAACCAGGAAGTGC
ACTTTGGATGGGCAAACCTTCTGGGCATATTAATTACAGCCCCAAAGGCTGCTTTGTT
ATAAAAAGCACTGGTGGGCTGGCCGGCTAGCTCAGCAGGTCAGGCGCTTGCCACCA
ACAACCTGAGTCCCAGACCAAGGCCACATGGTGGAAAGGAGAGGCCTGCCAGAAATT
GCCCTCTGACCCCCACAGTGCCGAGCACTCACACTCATGATAAACTAAATAAATCTA
AAAAACAAAACAAGACTTCAAAGCAGCAGATGGAGCGCTGACACAGACACTCGG
GAGGCTTGGGCAGGACGGCTCCAAATTCAAGGCCAGCCTGATCTACGCAGTGAGTA
CCAAGCCAGCCAGGACTTCGTAGCAAGACCCTGTCTCAAAGACATAAACAGGGCTA
AAGGGCTGCCTCAGTGGTTAACAGCGCTGGATGCTCTTCCTGAAGACCCAGGTTCAA
TTCCCAGCACCCCGGGCAGGCAGCTCACAACCATCTGTAACACTCCAGGGATC
CAGTGCCATCTTCTAGCCTCCGCAGACACCAGGCACACATGGAACAAAATACCGGG
ACATAAAGAACACACTGTGTGGTGAAGCCAGGGAAGGATCTGTGGTGGCTGTCAC
AGTACCGAGGCGACTCTTCTGAGTTTGAATCAGGGGACGGGAAGGAGAGCTCAGCT
CAACCGCTGCTACCTGTGGCTCCTGACCACTGCCCTTCAGCTCTTGGTGCCCACTGG
CTACCAAGCATTCCCAAGTGACTCGCAGTCACCTGAAATTCAATATGCCAACATGGT
GAACCCACTGTCTCTCCATCCTGCGTAGCAACACGCAAGGACGGGGAGCCAAGACT
ATGCCTCCCATGAACTATCTGTCCTCTGTCCCCGCTTATCTCCTAACTGGACAGTCCC
CAGTCTGGAACCTGGTGCCTTATGTTCTGGAGAGCCTGCAAAGCTGCCTGTTTGCTG
ATCCCTTTCCTTCCAGACCCTGCACTACAGAGCTGAGAGCCACCCAGCTATAACCCA
GTGTTTCGTTTGTAGCTGACAGGGACTCACAGAGCCCAGGCTGCTCACAACTTACT
ATGTGGGGAAGCCTGACCTAACTCCTGATCTTCCTGCCCTGCCTCCCAAGGCTGGC
TGGGATTACAGGCCTGTGCCGGGACACCTGGCCGGGACACTAGCTTGTGAGGCAGG
CAGAGAGGGCTCTCAAGCCCTGTAAAGAACTTGCTATTGGGAACACACACGCCCCA
CCCAGGAAAATGAATAGGACCCAACATGGAGTTTCAAGGGGCATGATGGGAGCTCA
GGAGAGAATCCTCTGCATGCTCCAGTGCCTCCTAACACGAGCTGGGTCTAGCCATCT
TGCTGCTTACTCCTCGACAGGCCCTTGCTGACAGCACCTCCCTCCTCAGTTCCTCAG
AACTCACAGCAGTTGGGGCTCTTACTCTGTGTCTGGCAGTGTCTCACTAGACCCTT
GGCAACCCACCCTGGGGACACGTACCACCCCCACTTCACAGGGAAGGAAACTGAGG
CACAAAGAGCAAGAGTACAAGGAAATGGGCTGGGCCTTTGAGCCCAGACTCCCAGA
CGCCAAAGCTCTCGATCCCACAGGCCACCTCGGCGGGCGATCTCCGCCTTCAGCAG
CCCCTCCATGGCATCCGACTCAGCCAGGTCCAAAGACAGGTCTCCATCACTGTTGAC
GGCGGCGATGTTGGCCCCATGGCTCAGGAGGTACCTAGGGGCAGGGGAAGGTCAGA
GCCACCAGGCCTGGACCTAACGCCTAACCCAAGCCCTGCCCTTCAACCCCAGCCTCA
CCTGGCAATGTCCAAGTACCCACAGGAGGCTGCCACATGCAGCGGCGTCCAGCCCT
CGTTGTCCGCCTGGTTCACAGTAGCACCTGCTCCACCAGGAAGCGCACCCACCTCCA
GGTTCTCGTCTATGCAGGCCTGGGGACGGGGACAGGCCATCAGCTCCCGGCCGGG
CCAATGAGAGGTGTGGAAAGCAACGCCGATGGGCTGCAGCACAGATTCCAGGGGCC

CTCTGGTCAGTGGCCGCCTAAAATATGCCTCGTTACCCATGCTTGGGTAATCTATGC
ATGCAGAGCTCATGGAGACTAGAGCAGGCTCCAAAAGGCAGATTGAAAAGGCGAC
CAGGGAAGAGGCGGAGCTGCCATCCCTGCATGTGACTGCTGAACATAACCCTATGAG
GCAGAGGAACCCCAAGAGCCCAGCCATGTTCTTCCAAGGGGCAGGGCAAGGCTAGGT
TGAGGCAAAACGCTCACCTAGCCCTGGGTTCCATCCACAACACAGGAAAAAGAGAG
ATCACCACAAAGAGACACACGCACATCCCAGAGTTGAGGGCTTGGGGCACAGTTCC
CAAAGGGATGAGTAGGCTATGTTCCCGGTGTCCAGGGATCCAAGCAGACCAAGCT
CTGGGTCACTGAGGGCCTACCGTGCACAGGTCTCCTCAGAACTTCTTTTCTAAACAC
CCCACACCACACTAATCCCCACCTCCTCACCCTTCGAACCAACAGCTCAGGAGAGG
AAGGCCTCACCTAGCCAGCACAGCACCCAGGGCCACAAGAGAGCTGGTCTCAATC
CCTGATGACACTGCTTGGACACTGGAACCACTGAGTCTAGGGAGGGGGTTAGGTCC
GGACTCCTGGGTCCTCATGAAAATTAACCCCTTCTACAAGCGCGACCATCTGGAGA
AAGAGGGAAGGAAGCTACGAGGGCCAAGTGCATGAAGTCATGGAAATTTAGGCTG
GGGGGGGGGGCACGTGCCCTGGGAACGGGATGAACTCTGGGCTTCACTCTGGGCTC
AGTTTATTTCCACCCTGTTGTCATGGTGATGGGAGGGGGGGCAAGGAGGCAGATGG
GCCTTTCCCTTTCAAGGACCTGGCCGGGTACGGGCATCCATGTGAAAGATGCCTGAG
GCTGGGCACTGGGGACCCAAGAATCCTCCTCCTCAGATGTAGA ACTCTAGCCAATC
CTCTTTCCTTAGACCCAGGGATCCAGACTTGGCCCTCCTCCCTCAGGCCCAGGTGCT
AGGGCTCCCCATCTCTCCCTGCTCAAACCTAGGACTCTTAACTCCAGCCCTACCT
ACTCCAGACCCA ACTCATAGCCATTATTGGACAAGGCAATTATTGGACAAGGGAAA
GAGGAAGGAATGTCCCTGCCTTGCTAAGGCAGAGGCTGGGGCTTAGGAAATGTCAT
TGCAGGAGGCTGATGCCCAAGGAGGGTCTAGAACCGGAAACACTAAAAAGTCTGA
GGTGTAGAAATCACCACAGACTGGGTGGCTCAATGCCCTGCTTTCCTGGGACTGAA
ACTAGTTTCAGGAGTTTTCACTGCTGAAGCCAGGGCAGTGGTACTAGGAGGTGATGC
TACGTACGCACCACTCCAAACCCAGCCCCCTCTGCGTTCTGGCCCTGAAAGCCAAA
TGATCTCACTGAATCTGATCTCCAGTCTCCAAGCCTCCTGCAAAGGCCTGAAGAGT
CAGGTCACCAAGGTGTCTGCATGGCGGGAGGAGAGTCCCACCTGGAAGGCTGACAC
GTCAGGCCTGAGGTCACAGGTTCTGTGTCAGAGAGGATGCTCTAGGGACCTCCAGCA
GATGCAGAGGAAGGGGATGCAGTTGGGAGGGAACTCTTGGGAGGGCCAGGGACTTT
GGTGATCATGTGAGCAGCCTGAGCTGATCTCCTGGACTGGTCAAAGACGCTGACAC
CCTGAGTGTGGCCTCGGGAAAACAGGACCCTGCTATATATAGAGGACGATGTCCCA
CACAGCTCACGCCGGCCCCATAAAGGAAGTTTTCCACAGGACGCCTCTCACCATAG
AGTCCCTCTGGGGACAGGGGTGACCACTGGTCCCATTCTACAGGTAAAAAACTAA
GGCGCACCGAGAAAAGACACTCAAGATACAAGACACAAGAAGCAGACTGACACAA
AAAGTCAGCACGGACTTTTTTTGTTTGGTCAAGATTTTGCAACTGGGTCTCATGTGG
GCTATCAAGATGGCCTCAAAGTCACTGTGTAGATGAAGGTGACACTGAATTCCAAC
CCTCCTGCCTCTACTTTCCAAGCACCTGATTTCTGTGGTATTGGGGTTGGAACCTGAG
GCTTCTGCACTCTAGGCAAGCACTCTGTCTAAAAGACAGCCCAGCCCAGGACAGA
CGGATTCTGTTTTTCTCTGCCTGGATGAGTGAAACACTGAACCTTTATTCCCCACCT
CCACGAGCATCCTAGCAAGAGGACGACAACCCAGGAGATGGAAGTTGCCATGAAA

GACTGAAAGTGAACCAACACTGTGGCCAGGAGGAAGAAGAACGGGGATGGGGGCT
CTGCTGTGACTAATCTTGTCCCTGACAATGCCAGCTTTTGGATGACGGGGAGATAAA
AGCATCCCGAATCCAGAAGGATCCCGGCAATACAAGATGGTCCTCACTCTCGGGGC
ACAGATCACTGGAAAAAGATAATCACAGTGTCTGAGTCGCCCAGGGTCCTGGTGGG
GTAGGTTCTAGAAGGTGACAGGGTGGAAATCTAAGAGACAGGGCATAAGTTTTTAAA
GCAGGATGCTGCCCAAATATAGTCCATGGGGTGGTAGGTGGAGTGGGCATGCCTGT
AATCCCACAAATGAGGGGGTCAGAGACACAGGACCTGTATTTGAGGTTGGCCTGGG
CTACACAGGGGAAAAAACAAAAACAACAACAAAAACAAAACCAGAGGAGAGAG
AAATAGGGCTTGGAAAGACCGAATGCTTAGAAGTTTCCAGAAAGGCAAATCAATGTG
GACACAGAGAGAGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA
GAGAGAGAGAGAGAGAGAGAGTTGAGGAGAAACCTATGGGGGGTGGGGGTGGGTGGC
AAGATCACCAAAGGGGATGAGCCGAGAATTGAATTAATAGGACATGGGGAGGGG
AGGAAGGTTATGGGATGGGGCCCAACAGACGGTGGAGCGCCTCTTCTCCAGGGGAA
CAAAGGGGTACACTGCCTTGGAGGGGCAAAGGACCCTCCTGAGGCCACAGCGGACA
GCACGGGTACAGGAAGTGGGGTAGGGAACAAGGTGGACCCCCAAAAGAAGTGA
CACTGAAGGGCCTGGGCCTGGCTAGCCTCAGAGGAGGGAGTGGGGGATTGGGGGG
GGGGGCGTCAAGTCAGAGCTGGGCCCTGGAAGCCTGCGGCACAGCCAGGGCAGCCA
CCAGCCTGGAAAGGCACGGGGTGTAAGCCATCCGTGTGCGGAAGACGCCGCCGGGG
AGAGCGGTGACAGCGCGGATGACAGGGGCGAGGCGGCCCTGCAGGGCAGGAGGC
GGGGAGGGAGGAGGGGTGGCTCGGGGGGCCCCGGGGAGGGAGGATGCTCGGGGGC
CGCTGACCTGGTGCAGGGCGCTGATGCCGTGCGCGTTGGTGGAGTCGAGCACCGCT
CGGGCAGGCGGCGGGACGCCGGCGTCCGATTCCCCGCTGGCGCCGGGATCGGGGCC
CCCCGGGTCCCCGGGATCCCCGGGGTCCGGCTGCGCGAAGCATCAGGCGAGCCTCAT
CCAGATCGCCGCCCGCACAGGCCGCCAGGAACTCGGCGGCGCGCTCGAAGCGCACG
GTGCGAGCACGACGCTCCCCGGGGCCCCGGCTCAGCCCCCGCCGCGCCCCCACCG
CCGCAGCTGCTCCTGCCGCCGCTCGCGGGCTGCCGCCGCCGCCGAAGACGACGACG
ACGACGACGACGACGACGACGCCGCCCGGGGCCGTCTCTGCCCCGACATCGCGCCC
CACACCGGGCCGCTCGCCCGCTACCCACCGAGCGAGCGAGCGAGCGAGCTGAGCG
AGCGCCCGCCCGAAGGCCGGCCGGCGACGAACAGCCGCCACCCGCCCGCTCGCTCG
CTCGCCCGCCCGCCCGCCAGCCCCGGGGGCCCGCCGGGAACCGCCGCCGCCGCCGCC
GCCGCCGCCACAAGCACCGCCCCGAGGCTCAGGCTGGGCCCCACCCCTCCCCCAC
GGACGGGCGTTGACGTCACGACGCTGCCCCACAGCCCTCTGGGAAATGGAGTCCTC
CGTTGAGAAGCCCGCAGGGTTTTTTTTCAGCAGACTCGCTAACTGCTGAGGGAACGGTC
GGGGTGGCACGGAAGCCGCCAGCAGGCGCGCCTACAGCCCCAGCACCTGAGAGGC
AACTGCTCTCTCGAGTTCGAGGTCAGCCTGGGCTACAGAGGCAGTGCCAGGGTAG
CACCAATTGCCTAAAGCAGGACACGCCCCCCCCGGGAATGCTGGAAATCTGAGTT
TAGAGGCGGGACGGGATGCCCGGGGGGATGCTGGGAGATGTAGTTTTTTTTGGTAAA
GCGGCGCAAAGGATGGCGCGTGGGAAATGATGGCGTGTAGCGGAACCCGAGAGAC
GCAGAAATAAGACTCGCGTACTTTCAGTTGTGTTTTTGTGTGAGATGGGTTTGGCC
TCGAGCTCGCTGTGTGACTGCGATTGTCTGTTTTAACTCCCGACCTTCTGCCTCCG

AGCAAGGGCTCTACCACTGAGCCACATCCCCAGCCCCTCACTGGGGGATTCCAGGC
AGGGGCTCTACCACTGAGCCACGCCCCCAGCCCCTCCTCACTGGGGGGATTCTAGGT
AGGGGCTCCACCACTGAGCCACACCCCCAGCCCCTCCTCACTGGGGGGATTCTAGGC
AGGGGCTCCACCACTGAGCCACGCCCCCAGCCCCTCCTCACTGGGGGGACTCTAGG
CAGGGGCTCTACCACTGAGCCACGCCCCCAGCCCCTCCTCACTGGGGGGACTCTAGG
CAGGGGCTCTACCACTGAGCCACGCCCCCAGCCCCTCACTGGGGGATTCTAGGCAG
GGGCTCCACCACTGAGCCACACCCAGCCCCTCACTGGGGGATTCTAGGCAGGGGCT
ATACCACTGAGCCACACCCCCAGCCCCTCACTGGGGGATTCTAGGCAGGGGCTCCA
CCACTGAGCCACGCCCCCAGCCCCTCACTGGGGGATTCTAGGCAGGGGCTCTACCAC
TGAGCCACACCCCCAGCCCCTCACTGGGGGATTCTAGGCAGGGGCTCCACCACTGA
GCCACGCCCCCATAAGGGCATCTCTTTTCAGATAATTCCCAGTAGGGGGTTGGTGGC
CATGTTGGAGTTGACTTTCTTGGGTTAGTTTCGGAGAACACATGCAAATTTATGAGTA
AGGGGCCTGAGGGAGAAGGAAGGGTGAGCTGGAGTTGGTGACTTGCATGCAACAAT
GTTGAGTGAGGCTGGAACAGTACAGAAAATGCTAGAAAAAGGCAGAGACTGAGCA
GTGAGCGGCCTGGATACGGTGGAGCACATCGGTAATCTCTGCACTCTGAAGGGGAT
GAGGCAGGAGGATCACCGAGAGTTTGAGGACAACCTGGGCTATATAGCAAGAGCCT
GACTCAAATGAAAACAACAACAACAGCAAAAAAGTCGGGTATGATGGCTCTGTAAT
CCCTGAACTTGGGAAGCAGAGGCAGGAAAGTGTGAGGAGTTCAAGGACACCCTCAA
CTACAAATGGAGTTCAAGGTCATTCACGCTTACAGGAGACCTTGTCTTAAAGCAAGA
AATAGAAGGAAAAGGGGCAGGAAGTGGACAGACAGATGGAGAAGGGGGGAGGGG
GGAAAGAAAGGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA
GAGAGAAAGAAAGGCAGACAGAGGGGGGCACTGAGATGGCTCAGCAGGTAAAGGA
GCTTGCAGCCAAGCCTAGGCCCTGAGTTTCAACTCTGGGACCCACATGATAGAAGG
AGAAAACCGACTTGTTTCGAGTCATCCTTCGGCCACATCTGCACCATAACAGCACACA
CACACACACACACACACACACACACACACACACACGCACGCACGCACGCACGCACGCA
CGCGCGCGCACACACACACACACACTATGCGGTGTGATATGATACAAAAAAAGTG
TAAAAGAAAATGTACTCAGAAAGAAAGGGTTGGAGGGAGGCAGAGAGGCGGGGAG
ATTGAGACCAAAGAGTTGATAAAGAGAAGCAAGAGATTGGAGTGCAGGCCAATAA
ACACAGCACTCAGCAGGCTGAAGCCAGGGGACCAGGAGGAGTTCAAGGTCAGCCTC
AGCTACCTAGTGAGACTGGGCTGCATGAAACCTTGCCTTAAAAATAAATAGACAGA
GCCGGGCAGTGATGCGCACGCCTTTAATCCCAGCACTTGGGAGGCAGAGGCAGGCG
GATCTCTGTGAGTTCGAGACCAGCCTGGTCTACAGAGCTAGTTCCAGGACAGCCTCC
AAAGCCACAGAGAAACCCTGTCTCGAAAAAACCAAAAAATAAATGAATAAATAAA
TAAATAAATAAATAGACATACCAAAAAAAAAAAAAAAAAAACAGGAACAGTGAGTCA
TGCCAATCATCCCCACATACATGGGATTAAGCAAGAGGATCTCCTACAGGTTCAA
AGAAAGCCTGGTCTACATAGTGAGTACCAGGCCAGCCTGGGCTACAAAGTAAGACT
TCCTCGAAATAATAAACAACAACTAAACAACAAACAACAAATAAATAAACAACACC
GAGAGAACAGATACAGAAAGGATGTCTCAGGGAGCAAGGAACAAGACATATAAG
ATGCCAAAAGGAGGGCTGGAGAGATGGCTCAGCAGTTAAGAGCACTGGCTGCTGTT
CTGGAGGTCCTGAGTTCAATTCCCAGCAACCATATGGTGGCTCACAGCCATCTATAA

TGAGATCTGGTGCCCTCTTCTGGCCTGCAGGCAGGCATACAAGCTGGCAGAACACTG
CATAATAAATAAATAAATCAAAAAAAGATAACACTTTAAAGAAAATGATACTTTG
AGAATTCTATGTATAGAGCCAGGCGGTGGTGGCGCACGCCTTTAATCTAGGTCCTCA
GGAGGCAGAGGCAGGAGGATCTCTGTGAGTTCGAGGCCAGCCTGGTCTACTGAGCA
AGTTCAGGACAGGCTCCAAAGCTACAGAGAAACCCTGTCTCAGAAGAAGAAAAAA
TAAAAGATCCCAAAGGGCAGTGGTATGCAGAAGACAGGGAGGAAGGGAGGGAGGG
AGGGACAGAGGGAGGGACAGAGGGAGGACAGCAGGCCTTTTGTGGAAGCAGCACT
TACAATTTCTGGGCATGGCTGATTCGGTTGTACAGCACATTGATCTGTAGAACGAGA
AGCCAGGCTAGGTGCAGATGTCCAACCAAGCCCTGCCCTGCCTATCACCCCTGTCA
CCCCAGCCTGGACCCCAACAGAGGCAGGTCCCACCTCGTATTTCTGTTGCTTCAGTT
TCTCCATCAGATCAAATTTCTCTGACTCGAGCTGGTGGATCCATTCCGACAGCTCCT
GGGCCTTCTCCCTGGCAGAATGAGAATAACTGGGATGCAGCGAGACTATGTTCTGG
GCCCAGAAAGGTTGAGACACCTACCCCAAGCCTCAAGGCAAGTCTCTCTGAGCTTG
ATACAGTTGGTATGCTAAGCCACTATGAGCCTGATACAGTTGGTATGTTGTGGAATG
TTACTTTAACTATGTAAAGATGCCTTACATTTGTTTACTTTGTGGAATGTTACTTTAA
CTATGTAAAGATGCGTTACATTTGTTTACTTTGTGGAATGTTACTTTAACTATGTAAA
GATGAGTTACATTTGTTTACGCTGTGGAATGTTACTATAACTATGTAAAGATGCGTT
ACATTTGTTTATATTGTGGAATGTTACTTTAACTATGTAAAGATGCGTTACATTTGTT
TACGTTGTGGAATGTTACTATAACTATGTAAAGATGCGTTACATTTGTTTAGGTTGTG
GAATGTTACTTTAACTATGTAAAGATGCGTTACATTTGTTTACTTTGTGGAATGTTAC
TCTAACTATGGGAAGGTGTGTTGCATTTGTTTCTGCTGCATTTGTTGAGTTGGATAAA
GGTGTGTTGCTGTTTCACCTTGCCTGCCTAAGGCACCTGATTGGTCTAATAAAAAGC
CGAACAGCCAATAGCTAGGCAGGAGAGGGATAAGCGGGGCTGGCAGGCAGAGAGA
ATAAGTAGGAGGAGGAATCTAGGATGCAGGGAGGGAGACCAAGGGAGAAAGAGAG
GGAGATGCCTGGAGCCAACCAGACATGGAGTAGTCAAATGCAGATGAAGAGAAA
CAGGTTAATTTAAGTTATAAGAGCTGGTAGGACAAGCATAAGCTAAGGCCAAGCTT
TCATAACTAAATTATCTCTCCACGTCTTGATTTGCGAACCGGTTGGTGGCCCCGAAAG
GAAGGCAGCTACACTGGTAAATCTTGTTAAGATTGGAGGCTGAAGTTTTAAAATAT
GGCCACTTTCTGAAACGGAGGTCTCCCAAAGGAGAGAAGGAAGTGATGACTGGGAG
CCCCAGAATCGGAAAGATGTTTGGTTTTTATTTATTGCTTATGTAGTGTGTGTGTG
TGTGCGTGTGTGTGTGTTAATTTTATTCTGAGACAAGGTCTCGTGTAACCCAAGC
TTCAATATATAGGAGAGGATGACCCAGAACTTCTGATCCTCCTGCCTCCCCCTCCTG
AGTGCTAAGATTCCACCTAAATGAGGGAGAGAGCACTGTCAATGCCAGACTCCACA
GGCCATGCCAGGCACAGGACAGATTGCACCCGAGTGACATTTTGAACAGAGAAAGC
AACAGTGGCCCAGAAAAGGAATGTCACTTGCCTAAAGTGACACAGCACCAAGGCTC
ACACCTGGAACATCACCCACGGAAACCTTAGGAGAGCCACAGGTGTTGCTGTAGTT
CCTTGTCCACCAGGTCCCTTCGTCTTCTCAGCCCCAAATAAACACACAGATACTT
ATATTAATTATAAAACTGTTGGCTGATGGCTAGGGCTTCTTATTGGCCAGCTCTGTCT
TAATTGACCCATTTCTATAACTCTATGTATCTCCACGTGGTCTTGGCTTACCGGAGAA
TGGCCGGACCTGTTACTCCTTTTGGCAGCTACATGGTGTCTTCCCTGTGGCCCCCTCTC

TACCTACCTTTCCCAGAATCCTCCTCGTCTCCTAGCCCCGCCTATCTTGCTGCCTTTA
TTGGCCAAGCAGTGTTTCATTCATCAACCAATAAGAGAAACACATATACAGAAAGG
CATCCCCCATCACACAGGCACTTACCGGAGTTGGTCCTCCCCCATGTAGTCGATGTT
CAAGGGTTTTTTCTCTCAGACAGGATCCTGAGCTTCATCTCCCGGCCGGTCTGCCG
CTTCCCACGTTTCTGCTCAGCCTGGAGGGGGAGGAATCCCAAGCCAGGGATGGGAC
CTGGAGGCCAAATCCACTCAGGGTCCTACAGTCATGGCCAGGGCCTCCACCACTTGC
AAGGGGCCAGGCTGCAGCCCTCCTCCCCAGACCCAGGAATCCAAGCTCCATGCC
TCCTCCATCAGACACGGGAGTACAGGCCCACTCTTCCCTTGCACCCAGGCACCACCA
GAATGTTATCTAAAGCGACAGCTGCCCCCCCTTACCCAAACCCCAACGGGATCCAG
GCTCAGTGCCCTTCTGAGGACCAGCATTCTGCAGCTCCCCTTCCCTCCTCCTTGTA
ACGCAGACACCCCCCCCCAGGGCTAGACTCACCTTGACCAGGTAGCCCCCAAAT
GAGCCCCCATGTTGGAGAGAACCTTCTTCTTCTTGGCATCGTCCTCGGCTCGCTTCTT
GGCTTCTCCTCCTCTTTGCGCATCTTCTTCTTCTGTGAAAACAGAGGGGTTCCCTCC
ATGTGGCCCTACTAAGGAAGGCACGAGCCTGGGTAGTGCATGGCTAGGCTCCATAG
ACGGGGCCGCAGAGGGATCCACTCCTAACTTGATATATAGTTCACCGAGCCCTCA
GGTTGGCAGAAATCCTCCTGCCTCAGCTTCTCCAGGGCTGGGATCACAGGTCTGAGC
CCACCGCAGGCAGCAAAGCTCAGTCATTGTGCTATGGTTTTGAGTCTCTTGATCCAG
CCATGCCTGAACTTTTTAAAATTTATATGCAGTTCATCCTTTCCTAGGAAAAACAA
AAAAACAAAAACAAAAACAAAAAACCTTGCCTGTACATTACAGGGACTTTCTGAG
AACCACAGGGGACATCAACGGCAGGAAAGGAAAGTATTCGTCTCCAAGAAAGGGG
CAGACACCTACCGCCAGCTTGGCCTGCCGCTCCCGCTCCTTTTCGGTTCTGAATCGCT
GCTGCTCAGCTCTCTCTGCACGACGCCTCTCCTGGGGTGGGGTTGGGGATGGAAGAG
AGGAAGATAGCGGAGGGGGATTGGAGGCTACCCTCTCCCCAGTTTATCCCTCCCTC
CACCAGTAGACAGCAAGAGAGGTAAGTGCGGTTTCTTTTGTTTTCTTTTCCTTTTTTG
AGATAGTCTCAGAGAGCCCAGGCTGGTCTCAAATTCGTGGCAATCTACCTTCCTTGT
CCCTCTAAGAGCACGCCTGGCCGAAATACGCATTGGAGTGCAACCACCCGGTGTCC
CCATTGTTTCTCCCGGAGACCCCAAAGCACTTCACGGAGGAAGGGACGAGGAGGGA
CCGAAGGACCCGGCCTCCCCGGCCTCCCCGGCCTCCCAGGCTCACGATGCGATCTTT
CAGCGCAATGAGCTCCTCTTCCCTCCTTTTTGCGCTGCTCGAAGTGCACGTCAATCAG
AGTCTGCAGCTCCAGTAAGTCTTTCTCCATGCGCTTCCGGTGGATGTCCTGCAGAGG
GAGCCCGTGAGGCAGAGGGACCAGACCCTAGAGCCGCCCTCTCCGGCCCCAGGGG
ATGATGATTGACAGCCAGCTGGGACGGCTTCCAGCAGAGGTCAGCAAAGCATTCTC
GGCTGGCAACAGGGCACCGAGGTAATGCAGGCATTTTCAAGAAAGGAGCAAAGA
GGGGCGCCTATCAGAATGGGCTCAAGGCGCTGAGGAGCCAGCAAGTATGTCTGGGG
TGGGACACCTGTCGCTTACATCAAAGTCCACACGCTCCCCTTCTGGGATCTTGGGGG
GGATCAAAGGAGGCACCACAGGACGGCTGTCGGGACAGAAATGGGAAGAGATCAT
TAGCAGGCTGGCCTCCTCATCCCCCTCACACAGACAGTTCAAAGTGACAGCTGCC
CCTGAGTCTAAGCAGAGACCAGTCAGAAAGTACACCGTCTGAGCATGTTGGGTTTA
ATAAAACGTGTAAGACGGTGTGTTTTGCCTGTTTTGACATACTTGTAAAAAAAAAA
AAAAATGGACAGAGCCTGATGGCACAAACCTGTCACCCGATCTACTTGGGAATCAG

AGACAAGTTCCGGGCCTGCCTGATCTACAGTAAAGCTAAAGCCATCCCTGGAAACTT
 GTGAGACCCCGGCCTCAAACТАААААGТАААААCAGGACTGGGACTGTGGCTCAGGG
 GTAGAGCCCCTGCCTAGAATCCCCCAGTGAGGGGGCTGGGGTGTGGCTCAGTGGTGG
 AGCCCCTGCCTAGAATCCCCCAGTGAGGGGGCTCGGGTGTGGCTCAGTGGTAGAGCC
 CCTGCCTAGAATCCCCCAGTGAGGGGGCTGGGGTGTGGCTCAAGACGGAGCCCTTGT
 GTCAAACATATAAGGCCACAGGTTGGACACTGGACATТААААAGGG

Последовательность CCR5

(указана предполагаемая гидовая последовательность для вставки сайта интеграции) (SEQ ID NO:3)

GTAAACAGAGTCCTGTAATGCAAGGTCCGGCCTTGGCAGCCCCAGCCTGGAG
 CCACAGTGAGATGTGAGCCGAGGGTTATGCTGGGAAAAACCTCTCCCTCCCAGCAC
 CTGAAAGGCTCTGCAGGCCAGCAGCTCAGCAAGCAAGGGTAAGGGCATGGACTAA
 CATCTTATTTCATACTATCCCTTATAACACATCCTAATGTAATCAGCTCACAATATGA
 AATTATTTCAATTTCTCTCCAGTCATTGTTTCAATGGGGCCTTAGGGTTGACTGGATTC
 TGGAGGGCCCTGCCTAGAGGAGGGGGTGCATTCTGTCCCTATGTCCCCTCCTGCTCC
 ATCCTCCACAGCACGTGCCTAGTGGTCTACCTTGTGGGGAATTCTTGTACCTCCCTCT
 TCTAGGCATGGACTAGCATTGAGAAGTGGGAGAGGAGTGTTAGGAAAAAGGGCAA
 ATATAGACATACCTTGTCTTATTGTGCTTTACAGATATTGTTTTTGTGTTGTTGTTGT
 TGTTTACAAATTGAAGGTTTGTGGCAACCCTGCCTCGAGCAAGTCTATTGGTGCTGT
 TTTTCCAACAGCATGTGCTTGTTTTACATCTCTGTGTACATTTTGGTAATTCTCCA
 ATATTTCAAACCTTTGTCATTATTTCTATATCTGTTATGGTAATCTGTGATCAGTGATC
 TTTGATGTCACTATTGTAGTTGTTTTGGGGCACCATGAAGTGCACCCATGTAAGATG
 GCAAACAATCAATAAATGTTGTGTGTGTTCTGACTGCTCCATGGACTGCCTGTTCCCT
 GAGACACAATAATGTATATATAACAATTATATATATATATATATTTATAACAATTATAT
 ATATATATATATATATTTTTTTTTTTGAGGCAGAGTCGCACTCTGATTGCCCAGGCTGG
 AGTGCAATGATGTGATTTTCACTGCAACCTCTGCCTCCCCAGGCTCAGGTGAT
 TCTCCACTTCAGCCTCCCAAGCTGGGACTACAGGTGTGCACCATCACACCCGGCTA
 ATTTTTTTTTTGTATTTTTAGGAGAGACAGGGTTTTGCCATGTTGCCCAGGCTGGCCT
 TAAACTCCTAGACTCAAACAATCCACCTGCCTCAGCTTCCCAAAGGGGCTGGGATTAC
 AGGCATGAGCCACTGTGCCCAGCCCAAGACACAATAAATTGAAATTAAGCCAATT
 AATAACCCTACAATGGCCTCTAAGTGTTCAAGTGAAGGGAAAAGTCCCACGTCTCTC
 ACTTTAAATCAAAATCTAGAAATGATTAAGCTTAGTAAGGAGGACATATTGAAAGT
 CAAGGCCAAAAGCTCACCTCTGCACCAGTTAGCCAAATTGCGACTTCACAGGAAAA
 GTTCTTGAAGGATATTTAAGCTCTACTCCAGGGAACATGCAAATGAAGAGAAAAACA
 AAGCAGCCATATTGCTAATATGGAGAAAGTTTGTGAGTGGTCTGGAGAAAAGATCCAA
 CCAGCCACAACATTTCTTAAGTCAAAGCCTAATCCAGAGCAAGACTCTAACTCTCT
 TCAATGCTATGAAGGCGGAGAGAGGTTGAGGAAGCTGCAGAAGAAAAGTTTGAAGC
 TAGCGGAGGTTGGTTTGTGAGGTTTAAATGAAAGACAACATCTCCATAACATAAAAA
 TGCAAGATGAAGCAGCAAGTGCAAAGGGAGAAGCTGTGGCAAGTTATCCAGAAAA
 TCTAGATAAGATAATTGATGAAAGTGTCTACACGAAACAACAGATTTTCAGTGTAG

ACAAAACAGTCTTATGTTGGAAGAAGATGCCATCCAGGACTTTCACAGCTAGAGAG
GAGATGTCAAGGCAAGCTGCAAAGCTCCACAGGACAGGCTGACTCTCTTTTAGAG
GTGAATGCAGCTGATGACTTTAAGTTGAAGTAAATGTTCACTTACTATTTTGTAAT
CCTGGTGTCAATTAAGAATTATGCGAAATCTACTCTATCTGTGCTCCATAAATGGAAC
AATAAAGCCTGGATGACAACACATCTGTTTACAGCATGGTTTACTGAATATTTCAAG
CCCCTATTGAGA ACTATTGCTCAGAAAAAAGATTCCTTTCAAATATTACTGCTC
TGCACCATGTGATCAAGAGCTGTGTTGGAGATGTACGAGAATATTCATGTTGTTTT
CATCCCTGCTAACACAAACATCCATTCTGCAGTCCATGGACCAAGACTTTCAAGTCT
TATTAAGAAATATATTTTCATAAGGCTATTAAGAAATAGCTATATATATATATAGC
CTTATATAGTTTATATAGCTACCATTGATAGTGATTCCATTGATGGATCTGAGCAA
GCAAATTGAAAAGCTTCTGGAAAGTAGTCATTATTCTAGATGCCATTAGGAACATTT
GTAATTCATGGGAGGAGGTCAAATAACCAACATTAACAGGAGTGTGAAAGACATTG
ATTCCAACCCCATAGATGACTTTCAGGGGTTACAGTCTTCAGTGGAGGAAGTCGCT
GTAGATGTGGTGGAAACAGCAAGAGAACTAGA ACTAGAAGTGGAGCCTGAAGTTGT
GACTGAATTGCCGCACTCTCATGATCAA ACTTGAACAGATGAAGAGTTGCTTCTTAC
ATATGAGCAGTGAAAGTGGTCTCTTGAGATGGAATCTCCTCCTGGTGAAGATGCTGT
GAACACGGTTAAAATGACAACAATCGATTTAGAATATTACATAAATTTAGTTAATAA
AGCAGTGGCAGGGTTTGAGAGGATTGACTCCAATTTTGAAAGAAGTGGGTAAAATG
CTATCAAATAGCATCACATGGTATGGAGAAATCTTTTGTGAAGGGAAGAGTCGACC
AAGGTGGCAAATTGCATTGTCATCTTATTTTAAGAAATTGCCACAGCCACCCCCAGC
TTAGCAACCACCACCCTGATCAGTAAGCAGCCATCAACATCAAACAAGACCGCC
ATCCTCTTCAGCAAAAACACTATGACTTGCTGAAGGCTCAGATGATGGTTAGCATTT
TTAGCAATACAATATTTTAAATTAAGGTATGCACATTGGTTTTTCTGACATAATACTA
TTGCATACTTAATAGACTACAGTATAGGATAAACACA ACTTTTATATGCACTGGGAA
ACCAAAAAGGTTATTTTGTAGATATTTGCTTTACTGTGGTGGTCTGAAGCTGAACTC
ACAATCTCACCAAGGTGTGCCTGAACCTCTTTAGCTAACTGGCCACTGCCACAGTCC
ACTCTGTGTTGGTCAAGATGCCCCAGAGTGGCAGGCACACTGTGTGGTCACATCCAA
GGGCCTAGATATGGTGGGGGCTCCAAATGGATCTAGATATGTGAGATCTCTCTTTGA
TTTGACTTCTTCCAACCCACCATTTTCTGGGTGCTGGGCTCATCTCACCCAGAAAGTA
GGACCCAATGTGACAGTTCCTGCCAGTTCCTCCTGTGGTAGCCACTTGACCCAGG
GGCACTCTTGATCCTTGACGCTCACTTACACACCCTATCTCTACCCCTATTA ACTCT
CTCCAATCCCCACTCCCCCTGCTCAGCTTGTCTGCTGCCAGTGGGGGCCCCACCCA
TGCTGGCCTCTCCTTTTGCAAGTCCCCATTCCTCATATGGTTTCTTCAGAGCCCCTTT
CTTTGGCTTTGAGGAGAGATGCCCTCACTCGCTTCCCCACCAATCCTGCCACTTCTA
CAATCCATTCAATTATCCTAATTGCCTCCGTATACAGACTGGAGTGAGAGGAGTTGAT
GTGATGGGTGTGGATACAGGGCTGGTGTGTCATCTTCTAGTAAGCCCTGGGAGAG
GTGTCTGAGCCCAGGTGTGAGTGGTTTTCTTTGGA ACTGTGAGTGCATAACACTTCTT
TGCCTTCAGCCTTAGGCCATAGTTGCTAGTTCTGGGACAACCAGAAAAGCCCTACAT
AATCTCGTGTTATGTGCAGAGCTGAGTATAGAGCTCCAGGTATGATCTGACTCACTT
AAGATCACAGTGAGTCTATTGTATTGTTGAACTGTTAGCTTAGACATCTGTTACTGT

ACCTACATGGCACTAGCCTCACGCCTAGACACCGATCTGAAAGAAATCCCCTAAAT
GCATAGAGAAGACTTCTCAGCTGAGCTAAGGGGCTCCCACCAGGTTTGAGCCTATCT
AATGAATCCATGAGGTAGACAGCCTGCACATGTCCACTTGGTTTGATGAATTGCACA
AATCCCCTATGGGGGATGTGGTTCATGGGCTGGGAAGTGGGTTACCCTGGGAAAGGT
CTACAGGACAGAGGCAGGGATGGAGACAACAGCATGGTGAGTTCCCAACCCACCCA
CGATGATAGGTGTCTGAGGCAGAAGGTAAAGAGGCTGTCACCTGGTGGGTGTCATA
AGACTCAAGTGTCAATTGTTGAGGCACATGGGTAACAAAGCGTGGCACTGGATGGGG
GTAGATTCTTCCTATTTCTGTGAGGATCAGGGGGACTCCCTGGCTCTCCTGCTAAAG
GTGGCTCTAGGGACAGGAAGAGTGTACTTCTTGACAGGGATGTCAGAGCACTGATG
GTGACAATCAGTGTGACACTGCTCACATGACTGAACAACCGAGAAGAGCCCGACTG
TCTACTGAACAACGGGAAGAGCCCGACTGTCAATGACGGAGCTCTGTTAAATATAG
TTAAGGCTATTTTGTTGAATGAATGAAGCCAGACAGGAAAGAGGACAGTATCTTTA
ATCCATTTATAGAAGTTAAAGACAGGCTTATTTAATCTCTATGAAGACAGAGTGGCC
CTTACCTCTGGGTGGAGCAAAAGGCACCTTCTGAAGTGATAGGGATGTTCCCTTATCA
TCTTGATCCGGAGTGGTAGTTACATGCATGTGTGCATATCAAACTCACCAAGCTGT
ACCACTAAGTGTGTTCTTCCTCAATAAAAATAATAAAGAACTACACTTATAAAGAAT
TTTTTAATAATATAGGAAAATGTCTACACTATAATCTTTAGCTAAAAAAAAAAAAAAAA
AAGAAGCCGCCTACAGAATGGTATATGCATGAGAACAATTAATCGAAAAGTGCATG
GGAAAAGTCAGGATTGAAACATCATGTTTTAAAGACATTGTTTTGATACTGTGAGA
ATGTACCTAAGTTTTTCTTTTTTCTGTTTTTCCCAATTTTATACAATGAGCATGTGTT
GGTTTTATAATTAGACATTTTGTGTTGTTGGTTTTGGTTTTGAGACACAGCTTGCTGTC
ACCCAGGTTGGAGTGCAATGGCCCAATCTTGGTTCCTGCAACCTCCATCTCCTGGG
TTCAAGAGATTCTCCACTTCAGCCTCCTGAGTAGCTGGGACTATAGGGGGCGCACCA
CCACATCCAGCTAATTTTGTGTATTTTATAGTAGAGATGGGGTTTCACCATGCTGGCC
AGGTTGGTCTCAAACCTGACCTCAAGTTATCCACTCGCCTTGGCTTCCCAAAGTG
CTGGGATTATAGGCATGAGCCACCGCACTTGGCCTAGACATTTGTTTTTAAAAATAA
AAGATTCATTTGCTCTTTTTACAGCCCGTCTCACTGTTGACTGATATTGACCAGGAGT
CAACTCAGGCCCCAGGGATTTTCAACAACAGCTGCTGTATGGCAGGGTTTCTGCTCAC
TGTGCTCATGTAGTTGGCCCTTGACCCAAAGTGAATAATTAACATTCTCCCCATCCT
GTTGACGATGCTCTGAAAATATGGTCCAGAAATGGTGTGAGCAAGGAGACAGCAAA
GCAATGCTTGGAACATAGGTGCAGTGACTAGACATGGGGCAGCTGTTTAAAGACAA
AAAGGCCCCAAAAGGAGGGATGGCACGAAACACCCTCCAATATGGGCATGGAGT
CTAGAGTGACAAAGTGATCAAAAGTTCATTTCTATGGGGTGTCCGAATGTACTTAA
TAATAAAAAGAGAACAAGAGCCATGCAAACCTGAGAGGGACAAAGTAGAAAGAGTA
GCAGACACCAAGCAACTAAGTCACAGCATGATAAGCTGCTAGCTTGTGTCATTATT
GTATCCAGAACAACATTTCAATTAATGCTGAAGAATTTCCCATGGGTCCCCACTTT
CTTGTGAATCCTTGGGCTGAACCCCTGTCTGAGTGGTTACTAGAACACACCTCT
GGACCAGAAACACAAAAGTGGAGTAACGCACACTGCAAAGCTGTGCTTCCTTGTTT
CAGCCTGTGAATCCTCACCTTGTTCCTATCTAGCCTATATTTTTCAAACCTAAGTGG
CCATAGAATCATGTAGTATTTAGGGTGGAAAGCTGCCCCAGGTCTAGCACGTCATTTA

ACAGATGAGGAAATGGAAGCTTGGGCAGTGGAAGTATCTTGCCGAGGTCACACAGC
AAGTCAGCAGCACAGCGTGTGTGACTCCGAGCCTGCTCCGCTAGCCCACATTGCCCT
CTGGGGGTGAGTATGTCTTCACATCCTCCAATACCCTAATGACAGACAAACAGAAC
ATGGCAAAGCCTCAGCTCTGCATGGTGAAAGTAAGAACCAGCAATTGCCACAAACA
GAAATACAGTGTGGTCCGGCAGCCTCCGGGGGTCTGCACAAGTGGATTACCAGT
GAATACAAGGCTATCTATCTTTCGAAAAACCAAAGTTGTATTTATGCTATCTATTTTC
TATAAAATTTTATATTAATTTATTTGTTACCTATTTTTGAACTCTTTCAAAAGCACAC
TTTATATTTCCCTGCTTAAACAGTCCCCCGAGGGTGGGTGCCCAAAGGCTCTACAC
TTGTTATCATTCCCTCTCCACCACAGGCATATTGAGTAAGTTTGTATTTGGGTTTTTT
TAAAACCTCCACTCTACAGTTAAGAAAATAAGGCACAGAGCTTCAATAATTTGGTC
AGAGCCAAGTAGCAGTAATGAAGCTGGAGGTTAAACCCAGCAGCATGACTGCAGTT
CTTAATCAATGCCTTTTGAATTGCACATATGGGATGAACTAGAACATTTTCTCGATG
ATTCGCTGTCCTTGTTATGATTATGTTACTGAGCTCTGTTGTAGCACAGACATATGTC
CCTATATGGGGCGGGGTGGGGGTGTCTTGATCGCTGGGCTATTTCTATACTGTTCT
GGCTTTTCCCAAGCAGTCATTTCTTTCTATTCTCCAAGCACCAGCAATTAGCTTTACC
TTTTCAGCTTCTAGTTTGCTGAAACTAATCTGCTATAGACAGAGACTCCGGTGAACC
AATTTTATTAGGATTTGATCAAATAAACTCTCTCTGACAAAGGACTGCTGAAAGAGT
AACTAAGAGTTTGATGTTTACTGAGTGCATAGTATGTGCTAGATGCTGGCCGTGGAT
GCCTCATAGAATCCTCCCAACAACCTCATGAAATGACTACTGTCATTCAGCCCAATAC
CCAGACGAGAAAGCTGAGGGTAAGACAGGTTTCAAGCTTGGCAGTCTGACTACAGA
GGCCACTGGCTTAGCCCCTGGGTTAGTCTGCCTCTGTAGGATTGGGGGCACGTAATT
TTGCTGTTTGGGGTCTCATTTCCTTCTTAGAGATCACAAGCCAAAGCTTTTTATTCT
AGAGCCAAGGTCACGGAAGCCCAGAGGGCATCTTGTGGCTCGGGAGTAGCTCTCTG
CTGTCTTCTCAGCTCTGCTGACAATACTTGAGATTTTCAGATGTCACCAACCGCCAA
GAGAGCTTGATATGACTGTATATAGTATAGTCATAAAGAACCTGAACTTGACCATAT
ACTTATGTCATGTGGAAAATTTCTCATAGCTTCAGATAGATTATATCTGGAGTGAAG
AATCCTGCCACCTATGTATCTGGCATAGTGTGAGTCCTCATAAATGCTTACTGGTTTG
AAGGGCAACAAAATAGTGAACAGAGTGAAAATCCCCACTAAGATCCTGGGTCCAGA
AAAAGATGGGAAACCTGTTTAGCTCACCCGTGAGCCCATAGTTAAAACCTTTTAGAC
AACAGGTTGTTTCCGTTTACAGAGAACAATAATATTGGGTGGTGAGCATCTGTGTGG
GGGTTGGGGTGGGATAGGGGATACGGGGAGAGTGGAGAAAAAGGGGACACAGGGT
TAATGTGAAGTCCAGGATCCCCCTCTACATTTAAAGTTGGTTAAAGTTGGCTTTAATT
AATAGCAACTCTTAAGATAATCAGAATTTTCTTAACCTTTTAGCCTTACTGTTGAAA
AGCCCTGTGATCTTGTACAAATCATTGCTTCTTGGATAGTAATTTCTTTTACTAAAA
TGTGGGCTTTTACTAGATGAATGTAAATGTTCTTCTAGCTCTGATATCCTTTATTCT
TTATATTTTCTAACAGATTCTGTGTAGTGGGATGAGCAGAGAACAAAAACAAAATA
ATCCAGTGAGAAAAGCCCGTAAATAAACCTTCAGACCAGAGATCTATTCTCTAGCTT
ATTTAAGCTCAACTTAAAAAGAAGAAGTGTCTCTGATTCTTTTCGCCTTCAATACA
CTTAATGATTTAACTCCACCCTCCTTCAAAGAAACAGCATTTCCTACTTTTATACTG
TCTATATGATTGATTTGCACAGCTCATCTGGCCAGAAGAGCTGAGACATCCGTTCCC

CTACAAGAACTCTCCCCGGTAAGTAACCTCTCAGCTGCTTGGCCTGTTAGTTAGCT
TCTGAGATGAGTAAAAGACTTTACAGGAAACCCATAGAAGACATTTGGCAAACACC
AAGTGCTCATACAATTATCTTAAAATATAATCTTTAAGATAAGGAAAGGGTCACAGT
TTGGAATGAGTTTCAGACGGTTATAACATCAAAGATACAAAACATGATTGTGAGTG
AAAGACTTTAAAGGGAGCAATAGTATTTTAATAACTAACAATCCTTACCTCTCAAAA
GAAAGATTTGCAGAGAGATGAGTCTTAGCTGAAATCTTGAAATCTTATCTTCTGCTA
AGGAGAACTAAACCCTCTCCAGTGAGATGCCTTCTGAATATGTGCCCAAGAAGTT
GTGTCTAAGTCTGGTTCTCTTTTTCTTTTTCTCCAGACAAGAGGGAAGCCTAAAAA
TGGTCAAAATTAATATTAATTAACAAACGCCAAATAAAATTTTCTCTAATATATCA
GTTTCATGGCACAGTTAGTATATAATTCTTTATGGTTCAAAATTAATAATGAGCTTTT
CTAGGGGCTTCTCTCAGCTGCCTAGTCTAAGGTGCAGGGAGTTTGAGACTCACAGGG
TTTAATAAGAGAAAATTCTCAGCTAGAGCAGCTGAACTTAAATAGACTAGGCAAGA
CAGCTGGTTATAAGACTAAACTACCCAGAATGCATGACATTCATCTGTGGTGGCAGA
CGAAACATTTTTTATTATATTATTTCTTGGGTATGTATGACAACTCTTAATTGTGGCA
ACTCAGAACTACAAACACAACTTCACAGAAAATGTGAGGATTTTACAATTGGCT
GTTGTCATCTATGACCTTCCCTGGGACTTGGGCACCCGGCCATTTCACTCTGACTACA
TCATGTCACCAAACATCTGATGGTCTTGCCTTTTAATTCTCTTTTCGAGGACTGAGAG
GGAGGGTAGCATGGTAGTTAAGAGTGCAGGCTTCCCGCATTCAAATCGGTTGCTTA
CTAGCTGTGTGGCTTTGAGCAAGTACTCACCTCTCTGTGCTTCAAGGTCCTTGTCT
GCAAAATGTGAAAAATATTTCTGCCTCATAAGGTTGCCCTAAGGATTAATGAATG
AATGGGTATGATGCTTAGAACAGTGATTGGCATCCAGTATGTGCCCTCGAGGCCTCT
TAATTATTACTGGCTTGCTCATAGTGCATGTTCTTTGTGGGCTAACTCTAGCGTCAAT
AAAAATGTTAAGACTGAGTTGCAGCCGGGCATGGTGGCTCATGCCTGTAATCCCAG
CATTCTAGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTTCGAGACCAGCC
TGGGCAACATAGTGTGATCTTGTATCTATAAAAATAAACAAAATTAGCTTGGTGTGG
TGGCGCCTGTAGTCCCAGCCACTTGGAGGGGTGAGGTGAGAGGATTGCTTGAGCC
CGGGATGGTCCAGGCTGCAGTGAGCCATGATCGTGCCACTGCACTCCAGCCTGGGC
GACAGAGTGAGACCCTGTCTCACAACAACAACAACAACAAAAGGCTGAGCTG
CACCATGCTTGACCCAGTTTCTTAAAATTGTTGTCAAAGCTTCATTCACTCCATGGTG
CTATAGAGCACAAGATTTTATTTGGTGAGATGGTGCTTTCATGAATCCCCAACAG
AGCCAAGCTCTCCATCTAGTGGACAGGGAAGCTAGCAGCAAACCTTCCCTTCACTAC
AAAACCTCATTGCTTGGCCAAAAGAGAGTTAATTCAATGTAGACATCTATGTAGGC
AATTAAAAACCTATTGATGTATAAACAGTTTGCATTCATGGAGGGCAACTAAATAC
ATTCTAGGACTTTATAAAAGATCACTTTTTATTTATGCACAGGGTGGAAACAAGATGG
ATTATCAAGTGTCAAGTCCAATCTATGACATCAATTATTATACATCGGAGCCCTGCC
AAAAAATCAATGTGAAGCAAATCGCAGCCCGCCTCCTGCCTCCGCTCTACTCACTGG
TGTTTCATCTTTGGTTTTGTGGGCAACATGCTGGTCATCCTCATCCTGATAAACTGCAA
AAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTT
TTTCTTCTTACTGTCCCCTTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGA
AATACAATGTGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCT

TCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCGTCCATGCTGTGTTTGC
TTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAGTGTGATCACTTGGGTGGT
GGCTGTGTTTGCCTCTCTCCAGGAATCATCTTTACCAGATCTCAAAAAGAAGGTCT
TCATTACACCTGCAGCTCTCATTTC*CATACAGTCAGTATCAATTCTGGA*AAGAATTC
CAGACATTAAGATAGTCATCTTGGGGCTGGTCCTGCCGCTGCTTGTATGGTCATC
TGCTACTCGGGAATCCTAAAACTCTGCTTCGGTGTGCGAAATGAGAAGAAGAGGCA
CAGGGCTGTGAGGCTTATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCC
TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAATAATTGCA
GTAGCTCTAACAGGTTGGACCAAGCTATGCAGGTGACAGAGACTCTTGGGATGACG
CACTGCTGCATCAACCCCATCATCTATGCCTTTGTGCGGGGAGAAGTTCAGAACTAC
CTCTTAGTCTTCTTCCAAAAGCACATTGCCAAACGCTTCTGCAAATGCTGTTCTATTT
TCCAGCAAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACCCGATCCACTGGGGAG
CAGGAAATATCTGTGGGCTTGTGACACGGACTCAAGTGGGCTGGTGACCCAGTCAG
AGTTGTGCACATGGCTTAGTTTTATACACAGCCTGGGCTGGGGGTGGGGTGGGAG
AGGTCTTTTTTAAAAGGAAGTTACTGTTATAGAGGGTCTAAGATTCATCCATTTATTT
GGCATCTGTTTAAAGTAGATTAGATCTTTTAAAGCCCATCAATTATAGAAAGCCAAAT
CAAAATATGTTGATGAAAAATAGCAACCTTTTTATCTCCCCTTCACATGCATCAAGT
TATTGACAACTCTCCCTTCACTCCGAAAGTTCCTTATGTATATTTAAAAGAAAGCC
TCAGAGAATTGCTGATTCTTGAGTTTAGTGATCTGAACAGAAATACCAAATTTATTT
CAGAAATGTACAACTTTTTACCTAGTACAAGGCAACATATAGGTTGTAAATGTGTTT
AAAACAGGTCTTTGTCTTGCTATGGGGAGAAAAGACATGAATATGATTAGTAAAGA
AATGACACTTTTCATGTGTGATTTCCCCTCCAAGGTATGGTTAATAAGTTTCACTGAC
TTAGAACCAGGCGAGAGACTTGTGGCCTGGGAGAGCTGGGGAAGCTTCTTAAATGA
GAAGGAATTTGAGTTGGATCATCTATTGCTGGCAAAGACAGAAGCCTCACTGCAAG
CACTGCATGGGCAAGCTTGGCTGTAGAAGGAGACAGAGCTGGTTGGGAAGACATGG
GGAGGAAGGACAAGGCTAGATCATGAAGAACCTTGACGGCATTGCTCCGTCTAAGT
CATGAGCTGAGCAGGGAGATCCTGGTTGGTGTGTCAGAAGGTTTACTCTGTGGCCAA
AGGAGGGTCAGGAAGGATGAGCATTTAGGGCAAGGAGACCACCAACAGCCCTCAG
GTCAGGGTGAGGATGGCCTCTGCTAAGCTCAAGGCGTGAGGATGGGAAGGAGGGA
GGTATTCGTAAGGATGGGAAGGAGGGAGGTATTCGTGCAGCATATGAGGATGCAGA
GTCAGCAGA ACTGGGGTGGATTTGGGTTGGAAGTGAGGGTCAGAGAGGAGTCAGAG
AGAATCCCTAGTCTTCAAGCAGATTGGAGAAACCCTTGAAAAGACATCAAGCACAG
AAGGAGGAGGAGGAGGTTTAGGTCAAGAAGAAGATGGATTGGTGTAAGGATGG
GTCTGGTTTGCAGAGCTTGAACACAGTCTCACCCAGACTCCAGGCTGTCTTCACTG
AATGCTTCTGACTTCATAGATTTCCCTTCCCATCCCAGCTGAAATACTGAGGGGTCTCC
AGGAGGAGACTAGATTTATGAATACACGAGGTATGAGGTCTAGGAACATACTTCAG
CTCACACATGAGATCTAGGTGAGGATTGATTACCTAGTAGTCATTTTCATGGGTTGTT
GGGAGGATTCTATGAGGCAACCACAGGCAGCATTTAGCACATACTACACATTCAAT
AAGCATCAA ACTCTTAGTTACTCATT CAGGGATAGCACTGAGCAAAGCATTGAGCA
AAGGGGTCCCATAGAGGTGAGGGAAGCCTGAAAACTAAGATGCTGCCTGCCAGT

GCACACAAGTGTAGGTATCATTCTGCAATTAACCGTCAATAGGCAAAGGGGGGA
AGGGACATATTCATTTGGAAATAAGCTGCCTTGAGCCTTAAAACCCACAAAAGTAC
AATTTACCAGCCTCCGTATTTGAGACTGAATGGGGGTGGGGGGGGCGCCTTAGGTAC
TTATTCCAGATGCCTTCTCCAGACAAACCAGAAGCAACAGAAAAAATCGTCTCTCCC
TCCCTTTGAAATGAATATACCCCTTAGTGTTTGGGTATATTCATTTCAAAGGGAGAG
AGAGAGGTTTTTTCTGTTCTGTCTCATATGATTGTGCACATACTTGAGACTGTTTTG
AATTTGGGGGATGGCTAAAACCATCATAGTACAGGTAAGGTGAGGGAATAGTAAGT
GGTGAGAACTACTCAGGGAATGAAGGTGTCAGAATAATAAGAGGTGCTACTGACTT
TCTCAGCCTCTGAATATGAACGGTGAGCATTGTGGCTGTCAGCAGGAAGCAACGAA
GGGAAATGTCTTTCCTTTTGCTCTTAAGTTGTGGAGAGTGCAACAGTAGCATAGGAC
CCTACCCTCTGGGCCAAGTCAAAGACATTCTGACATCTTAGTATTTGCATATTCTTAT
GTATGTGAAAGTTACAAATTGCTTGAAAGAAAATATGCATCTAATAAAAAACACCT
TCTAAAATAATTCATTATATTCTTGCTCTTTCAGTCAAGTGTACATTTAGAGAATAGC
ACATAAACTGCCAGAGCATTATAAGCAGCTGTTTTCTTCTTAGTGTGTGTGCAT
GTGTGTGTGATGTATACAAAGAGAGAGATAATTGTATTTTTGTATTTTCTTTAAATA
ATTTTTAAAATTGACCCTTTTCTGAGACAAATTGCCAGAATAGTTTGTATTTAGAG
ATGGTACCTCTAAGAGTAAGGTTGCTGGTTGCTGAGCAATTGACTTGAAAACCTTTTA
AAATTCAAATTTAATTCCACTACTCAAAGAATTGCCATGTTTTAAAAAAGAGAAT
TGGTGCCATAAGTTAGTTGTCTATGTTTAAAATGAAGAAGATATGCAACGTCATGG
CCTGGTCACTTACCCGCAGCCCTGAGTTGTAGGCACATCATATGTGAGAATGAGGAT
GCTTTTCTTTCATTTAAAATCCCTCCCCAAAACCTTGGCTCTAATTGCAGTCATGACAA
TCATGTACATTTGGATTTATGTGCACGAGTCTCTTACCCTGAGAGAGGACAGGTGCT
ACAGGTGGAGGGGACCCGTCTGGGTACGTTACATTTTGAACATGCTGGTTTTTCAG
TCACTGCACACTCATCTCCCAGCACAGGTCATGGGCAGCAGATGCAAAAGCTGCCC
GTGGTCCTATTTGGAGGTGCATGAAATGAGCAGAAGACAGAACAGCTTGATCTGAC
TAGAAGGGCAGCTTGTCCCTACCAAGACTTGAAGGATTGCCTTTCATCTGTTAGGGT
AAAAGGTAGAATGAACCAAGGAAGGGCAGGAGGGGGCTGGGGTTAGGGTAGAAGG
AAGGGGCCATGGAGAAGGGAGATCCATCCCATAGGAGGAAGGCAGTGCGGCAGGG
AGGTTTGAAGGTATCAGCTTTTGTGGCTGACATACATGCAGTCATGTCAATTGCTCG
TTTTTCTTTTCCATCTTATTAATGTCTTCCAACGTTAGCACGAAGAAAAGCTATTT
GCAGTGTTGCCAGCCTTTCAGAGCCCGTCCCCATTACCTCCCCAGGCCCATGCCTTT
ACTCCTTGGAGTTTCAACTCACGACCTTCAGGATCTGACTTTATTCACCAACTCTGGG
GTGAACGTACCTTCTGTCTCCACCCAGAGGTCTCTATCAAAGAGGAGATTGCATGCC
ATGGATAAAGTCAAAGTAGAGGTGACTGTCCTTAGGAAGAGTAATGTGAAAATTCA
TAACTGGGATTCTGTTTACATTTTGTACTCCAGGGGTTCTTAGTTTAAATCGCTCTG
AATAAATTAAGATGCAATGGCATTTCAACTGTTATGATTAAATTTACAAATCATTTA
TTTTCTATCACGGGGAGAGATAGAGCTCCAAATGCAAACATAACTGCTCAAGTGTTA
ACACTTATAATGAAAACATAAGAATTACCACCAACTACCCTGGGGGCTAGAAGCAG
AAATGTGAACCAGAAAACAATCATGAACTTTCCTTTTTTTTTTTGAGATGGAGTCT
CGCTCTGTTGCCAGGCTGGAGTGCAATGGTGCGATCTCGGCTCACTGCAACCACTG

CCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGG
CATGCACCACCACGCCTGGGTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACC
GTATTAGCCAGGATGCTCTCGATCTCCTGACCTCGTGATCTGCCC GCCTCGGCCTCCC
ACCGAAGTGCTGGGATTACAGGCATGAGCCACTGTGCCCCGGCCAACAAATCATGAA
CTTTCTAACTGCAGTTCCTTGTAGCTTGTTAACACATCCACTTACTTATTGTCAGAGT
ACGTGGAGATTTTCCACAACCCTCGGGGATAAGGCTGAACAGAAGAGGGCAAAAACG
TGAAAACATTTTCGATAGCTCCTATACTTTGAAATAAAATTCACTGTAAAAGTTGCTT
GTATTTTTCCAAAACAGAGTCAACCCTTAATATTTAAGATTCTGTATACAAATACAT
ATTTTTATATAATTAATATATATTGTCATATGACATATATCTTTATATTAATATGCAT
GCATATAATATATATTTCTTCCTAATTTTCTATAAGCAATTTTACAAGACTGACTTC
TATTTGCCTCCTTATTGTTACTACGTGGTTTTGATAATCCGTTTTGTGTCATTGTGATTC
TGTCATGTTTTGGGGACTTATTTTTGTTTCTCTGGGTGGTCACTAGTTTTTTTTAAAGC
ATTCATGGAAGAGTGTGAATCTTTTACAAGCTAGGAAGCCATGGCAAGCCTTGGGTC
ATACTGCCCCCGCGAGGCCACATTGGCAAACCAGCAAGGGTGTCAACTTCCAGAC
TTGGCCATGGAGAAGACACACGAGGAGGCTTTTCACATTCAGCTCTTTAATGTTTTGT
CTCTGCCGGCACCATCCCAGTTGTGAAAAAGAGGTATTTCCACAGCGGCTCAGGGTA
GGTAGTGCACAGCTCACATTCATCATTCTGAAAACCGAGAGGAGTCTCCATTCTG
GGGTACAGGTTGATGCCTGTCGTGGAATGAAGGTTCCAACACCAGACCAATCTCTG
CAGTGTGCTGCTCTCATGAGCTTGCAACAAGATCAGAAAATGTTTTGTGACTAAGCA
TTTTTCATATTGCATAAAATGCTTCAAGCTCCTCCCTTGTTTTCTCTCTATAATCCTGTA
TATCTGATGATTGTGGGTACCAAGTGTGAAATAATCAAATGTGATTTGATGTTGG
TAAATTTCTTTTTTTTTTTTTTTTTTTTACTTCTATTTTTTTTTATTATACTTTAAGTTTTAGG
GTACATGTGCACATTGTGCAGGTTAGTTACATATGTATAACATGTGCCATGCTGGTGC
GCTGCACCCACTAACTCGTCATCTAGCATTAGGTATATCTCCCAATGCTATCCCTCCC
CCCTCCCCCACCACAGTCCCCAAAGTGTGATATTCCCCTTCCTATGTCCATG
TGATCTCATTGTTCAATCCCACCTATGAGTGAGAATATGCGGTGTTTTGGTTTTTTGT
TCTTGCGATAGTTTACTGAGAATGATGGTTTCCAATTTTCATCCATGTCCCTACAAAG
GACATGAACATAGCAAAGACTTGGAACCAACCCAAATGTCCAACAATGATAGACTG
GATTAAGAAAATGTGGCACATATACCCATGGTAAATTTCTTTATCATTTCGCACTCT
CCTTTCTCTATTATTGTTATTGTAACCTGAACCGCAGATTAGTCACTCATTGCTTGCAG
AATCCAATTAACAAGAGCGAGGTCAGATATAAAGAAAATGATTTATTCCAAACCTC
CTTCAGGGAAGAGGTGCAGCCTCCTGCCTCTAAATGCACTGCTTCGCCAGGCGTGGT
GGCTCACACCTGTAATCCCAGCACTTTGGGAGACCGAGGAGGGCAGATCACTTAAG
GTCAGGAGTTCAAGACCGGCCTGGCCAATATAGTGAAACCCCTGCCTCTACTAAAA
ATACAAAAAATTAGCCAGACGTGGTGGCGGGTGCTTGTAATCCCAGCTACTCGGGA
GGCTGAGGCAGGAGAATCGCTTGAACCTGGGAGGTGGAAGTTGCAGTGAGCTGACA
TCTAGCCACTGCACTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAATAAATA
AATAAATAAATAAATAAATAAATAAATAAATAGTAAATGCACTGCTTTGCTTTTGGGA
GCAGAAAGCAGGCACTTTGAAAAGGCAGGGGAGGAAGTGAGCAAGGGCAGGGGGT
CTGCACACTGGCATGGTGCCTGATCTATCCAGGCAGTTGAATTGGCACTTTCATAGG

CAGAAATAAGTTGAAAAAGTGGCCTAAAACCTCTCTAGGTGGGAGTGGATAGTGGGC
ATGCCTTCAACCTGCCTTTCTGGAGGGTGAGTTCCATGGCAACCCCCTGAAGGGTGA
GAGTTCCATGGAGATCATGCTTTGGTCTGTAAATCAGCTGTAACTCTCTAGAAAGT
TCTGTCTTGGAGCATATAGTTAGATGAACTTGCCCTGTAAAGAATGTCTGGTGAAGG
GGAAGTAAAAGGTGAGATTTGCATTTCTAAAGGGCTAAGTAGAAAGTGGGGTACAA
GAGGAAAGGAGAAAAGAGAAAATAATTTAAAAAATAATTGTAACCTATTCCCTTTT
ACTTAGAAAAAAGGGAATACTCAGTTACATTATCACCTCGTTTACATCAAACCCTCT
TATGGAATCCTATGGTTTGAAAACAAAAAGGTTGTTGAGGACCAGTGAGCCCAACC
CCTTTGCTTTATAAATGAAGAGCATTGCCTGCCCTAAGCCCCAGAGACTCTGATGTC
GTGGGTCTGGAGTGGGCTCCAACAGCGGCATGTTTTGATGGTGCTTCCCAGTGGCAC
GCCAGCGATGAGCCTTTGAGTAGGGAAAGTAGGAGCACTCGTGACTCCCTTCACGA
TCAGCACCTGTGTGCTAATAAATTCACAAAAGCCAACATATTGGAGTCACTCAGGG
AGTTTTACAAATAGTGAGGTTAAATCCAACCTCAAATAGTTCTGATTCGATCTGCCT
GCATTGCTGCCCTGTGGTTCCCCACTGTAGAAGCTCCCCAGGTGATTCTAAGTGTAG
CCAAGTCTGAGAAATACTGCCTAAAGCCTGTTGGACTGACAGCAAGGGCTGTTGTCT
GAGCAAGACTTTGCCTGGCCTGGGGTGGCATGTGCACCAGGAAGAGTCTCAACTTTC
ATAACAGAACATTCCCCAAGCTGGTTTTTTTTAAAGCATGTGAATCTAGACTTCATTG
GCAATACCAAAGATCTGTATTTGAGGCTCCAAGTATTTCACTTTCATTTTTGGTTTTG
GGTTATGTTTTACCCCTTCTTTCCAAGTGAAAAGTAAACAGAAGTGGGATGTCTGG
CGCCCATGCTGAGCTTGGCAACTTCAAATTCAATAGAGAAGAAGTCTCTTGTATAGA
AAAGGGCCTGTCTGAGATGTTTCTCAAATAAATATAGATTTTGCTTATGTGGCTAAA
GGATTCTTCTCCCCCATTTCCTTATCCCTGCAGTGAGCCATCCTTCTTAACTCTTTCC
ATGAAAGCATTATTCCTGAAGAACTGGGAACCTCATGCCAGCCCTGATCAGGCAATG
ATAATTCTGCAGAGAATTAGAATTTAGATTTAAATTGTCAACTCTTATACATCCTGG
CATATGGTTTTAAACACATGTACACACACACAAACACCTCCTACTATTTACTGAAGAG
CAGATATCTGATAACTTAATCTTTTTGGTTTTGAGTCAAGACAATTCCTCCTTTTGAA
ACTGCATACCGCTGAATATAATAAAATGTAATTAAGATTAATAAATAAGAACTAAT
GGGAGAATTTCAATATTGTCTATGTTCACTTTAAAATTCCTCTACTTAGGTTTACTGC
CATTACCAAAGACTATTCAAAAATCCTTTTTAGGAGAATCCTAATGGTTTCTTGACA
TATAATCAAATAAGGACTCTGTTGATTGGCTAACTCAATCTTCTGTGCCAAAAAGC
AGAGCCAGCAGAGAAGAGGGCAGGGACTTGAAAGTCAGACTGACTCGAGTTCCA
GCCTTGGGGCTGTGGGAGCTTGGGCAAGTGACTTAACGTCTCTGGCTCTCAGGATCT
AAAAGGATTTCCAGTAGTAATTTGGGGTGTACTGATACAGGAGCTAAAAAGAAAT
TATTTAGGTGGTTAGTGAGGGTCAGAGAGTCCTCGGTAAGATTTGCCTTTTAAACAAA
AAGCAGCCCCAAAATCATTTGTTTGCTAACAAAGAGAAGCCTGTAAAATTGAGCTG
CAGACATAGATAAGCAAGCTGGAAGCTTGACGGGTGAATGCCGGCAGCTGTGCCA
ATAGGAAAAGGCTATCTGGGGGCCAGGCATGTTCAACATGGATTCTCCATCTTCCCT
TTTCTTTGTCAACCAAGTGTACAGTAAAGGAACAGGCAACATGGCACGGGGCCAGGT
AGAGAACCCTTCTGCATAATAAAAGATTAGGGTGAGATGGCCAGCTTCTTCCCGTGC
TATGTAAATGGCATACTGGTCCAACCAGTCTTTTGGGCCCTGTGTAAATCAGACAC

CGCCTCCTCAAGTTAGTCTATAAAAACCCCATGCATTTTACCGTGAAACTGGGAGATC
 CACTCGGAACCCCCTCCTGCACGAGAGACCTTTTCTCTTTTGCCTATTACACTTCCGC
 TCTTAAACTCACTGCTCATGTGTTAGCATCCTTGATTTCCCTTGGCATGAGGCAACGA
 ACCTTGTGTATTACCCCATACAAATGATGCTGCTTCATTAATAAGCAACCTGACA
 GGGTTGTGTTGGGGTATAAATTATCTAGACCAGGGAGATCCAATATAATTTTTTTGT
 AATGACGGGAATGCTTTGTATCTGCATCATCCAAAATGGTAGCCACCAGGCCAGGG
 TGAAATGTGGCCAGTGTGACTGAGGAACTGAATGTTTTCCATGATTTAATTTAAATG
 TGGCCAATGGCTACTGTAGGAGACAGTGTGAGTCTGGCATATTATAAATAATAAAT
 ATTAATATAATTTGAACTTTGGCATCAGTGTTCCTAGATTTGAATTACTATGCAAGT
 TGCTTACTGTTTCCAAGCCTCAGCTTTCTAATCTGTAATTGGGGCTAATAATAGTATC
 TGCCTTACAGGTTTGTTCAGAGGATAAATGAGAAATTGCATGTTGAGGGCTTAACAC
 AGTGCCTGGCACATAAAAGCTCTGGTAACAGTTAGCCACTTTAATAATTTGCTAATA
 ATGGCTATTTCTTCTTCAGATTAGGATGTGCTCCCCCAAACAGTGCCTTAGACATA
 GCGGGCAATCCAGCTCACTCTCTGCAGTGAGAGAGAAGCACTGGCCGACCAGAGTC
 AGCCAGGGGCTCATGGGTATGAAATCAACAGCATGATTTTGTAAGTAATGGATGGA
 AAGGGCCTCACAACCTTTATGGCACTGTGTTCAATTTGCTTGGTCTTCTGTAGCTCCTT
 TTGAAAGCCTTTTAGGGTGGATTAACCTGCTACCAATAATTCTGGTCAGATGTAGAC
 TCCATAGCTCAAAGCAAACCTGAGAGAGTGAGGGCAGCAGGCCAATTCCCCACCCCT
 TCCTTCTGGACTCTGACAGAAGCTTACACTCAAGGAAGAGCAAGTAGGAATTAACG
 TGTTAAGAGCTAGGTAAGCAAAAACCCAATGAGAAGTTCTGGCAAAGCCCCATGGGC
 AGGGGTGGCTTAGGCACAGGAAACAAGTAGGATTTCATACCACGCGCCTCAGTCTA
 CTTCCGGGGCCCTCATCCTCAGCTGTGCCTATGCAAAGGAGAGCAACCAATAAACCC
 CACCGCCACTCTCCTACTGTGGAGGCCAGGGATGGCCAGGGGTAAAGAGAGGGATGG
 GAAGTGTTTCCTCCAGCCGTCCTCTGAGAAGGAGAGGAAACTGGGCAGAGCTTCTG
 TCCTCCTTCAAGCAGAAACAGAAACAAAAGAAACCCCTAAGGGGGTTCTTACTTCC
 CCTCTAGTTCAGTTGTGCACTAACCATCTGCAGCTCAACATTCAGCATTCAATTCATTG
 ATTCAGCAAACATTGAAGGAGGGCCAGCTATGTGCCAGATGCCAACTCATGCCATG
 AAAGAGAGTCCCTGTCCTTATGAAATTCACTATTTAGAGAGAAAAGCAAGCAAAAA
 GGCAAAGTTTGAAAAGTACTGTTGAAGTGGCATCATTGTCTGGGGTGAATACCTGA
 GGTTTGTGGTCTCACGCCAAGGGAATCAAGGACTCAGGCACACAAGAAGTGAGTTT
 AAGAGCAGAGGTTTAATAGGCAAAAGAAAGAGAAAAAAGAATAGCTCTCTTGCCTG
 CACAGAGAGAGGGGCACCTGAGTGGATCTTCTGTTTTGTGGTGAATGCAAGGCA
 TTTTATAGACGAGCTTGAGGAAGTGGTGTCTGATTTACTTAGGACCCGAGAGATTGG
 TCAGACCAGGTGTGATGTTTACATAGCATACAAAGAAGCTGGCCATCCCATCCTAAT
 CTTTTATTACGCAGACGGGGTCTATACCTGGCTGGTGCCATGTTGTCTGTTCTTACT
 GTACACGTGGTTGACAAAGAAAAGGGAAGATGAAGAATCCATGTTGAACATGCCTG
 GCCCCCAGATAGTCTTTTCTTATTGGCACAGCTGCCGGCATTCACTCTTGCAAGCTTC
 CAGATTGCTTATCTATGTCTGCAGCCCAATTTTACAGGTTGCTCTTTGCTAGAAAAG
 AAATGATTTGGGGGCTGCTTTTCATTA

(SEQ ID NO:4)

(указаны предполагаемые гидовые последовательности для вставки сайта интеграции)

ACCATTTAAACCTCAAATTAAGCAACCCACAGAACCAGGAAGTTCAAGGACC
ATGTCTGTTTTACACGATGTCTTTCCCCACCCCCCACCCTCCACCCCCCA
СТАAGGGCAGGGTATTGTATCTGCCAGACTGGGTATTTGTTGAACAAGCGAGTATTT
TCGCCTATTAGCTTAGTTTTTAAGGAAATCATTTTTTACTTGATTCATCATAGCTTTA
ATTCTATTACATACTACAATAAAAAATTTGACAAGACTGATACAAATATGTAGTGGGC
AATAGTTTGCCGTCTTCTCCCTAGTATGGTGTTTTTCAATCTGGTGACTAGAATAGG
CAGTGGGCTATAAGCAGGATTCATAAGGCCTGGAGCTGAGTTATATGTGACACTGC
CACCTATTCATTGTGTGACCTTGGTTTTAACCTTCAAAGTGGGTCTCCTGGACTAAAA
GAATGTGAAAAGATGGGGAAATAAATCTGTAATCTGAACATGGAATGACTTAGTTA
CAGACCAGACATATTGTTACTGGGAATGAAAAAGTCAATATATTTGAGGGGAAAAA
AATGTAATAAATATTGAGAAAGATTTTACAAATCTAATTAGGGGAAGATAGTTAT
CTCCCAATACTAGAGGGTACCAGAGTGCTTTTAAGGGGAACATTTGGTTACCCTTAT
TTCTTTAAAAAATGGCAGTTTAGGAAATCTGCCCTAACTGTAGTCCCAATGTCAGA
TAGGACTCAGGTCTCCACTGCAAGGACCAAAATGTTAAGTTGAAGACTGAAAATGG
GAAAATTTGGAAATGTCTTTGGAACCTCAAGTACATAAAAGCCTGTAAGTGCTTCAT
ACTCATTAACAACATAGGCATAGAAAAAAGATATCCTTATTCTCAAGCATAGCCTTT
TCTAATAAGTTCATGTTAGATGTCATGAAGTTTAGTGAGGAGTGAAAATCTATGAGG
AAAAACATGAACCATTCTACTCTGGCAAAGTTCAGGACAAACACCATAGGCCTGT
ATACCAAATTTTAAACCATGTTGAATAATGTGAAAAAAGCATCACATTGCTTATGA
AAGGCTTTCCTGTCGCCCTTAATACTTCTGTCTCAGGCTAACATGTTTGTTAATGAG
TTACAGTGGTGAAGTTAAGGAAATCTGCTTCCTGTCCTAGCATGCCATTATCCCAG
CCATACAGATTTAATAACCAGGAGTCACTTTAACTCCATGAAGTCATTCAACAGGTAC
TTGAGTATTTACTATGTGCGTTTGTGCTAGAGTAGCCATTTCTTAACTTTGTGGCCT
CAGAGAATCCCTTCACATTCTTAAAGATTGAGGGCCCCAAATCACTTTACTAGTATC
TTACCATATTAGAAAGTAAAATATTTAAATATTATAGCAACAAATGCATATTTTTAA
AAAAATAGGTGCATTGTTTTACATTTTGCCAATTGTTTCCCTAAATGTCTGACTTAAA
TAGAAGACAACCTGAATTGTTTCTTTTGCATTCAATTGGTTACAATGTTACTAGCAGTT
TTCTATAATCTCACGTATTAGTCATTAGGAAAATATAGCTTCACTGAGTTATGTCAAT
CTCCCAAATGTTGGCCATTTTACTGTATACTACCTAAAATATCATTGGCCGGGCAC
AGTGGCTTAAACCTGTAATCCTAGCACTTTGGGAGGCCAAGGTGGATCACCTGAGGT
CAGGAGTTCAAGACCACCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATA
CAAAAATTAGCCGGGTACAGTGGTACACACCTGTAGTCCCAGTTATGCAGGAGGCT
GAGGCAGGAGAATTGCTTGAACCCAGGAGACAGAGGTTGCAGTCAGCCAGATGTCC
CAAAAAAATAAATAAATTTTCAATATCACTATCTCATGAAGTATTGGGAAGCT
GTCAAGTTCCTGATATTAGACACAAGTTTTCCCGAAATTCTGATTTTACTCAACGTT
TGGATTTTATCATTGAAAACAATTGCTGTCAGTTGTTAGGCTCCAAGGAAATAGCAG
ATAATTCAGCTTACATGAGTGCTTTTTCTTGAGACAACCATTTCAAAAGTTATGTA

CTGTAGGGTTTAAAGATTTAACAAAATGTCACTGCTTTCACAAGGACATTCTTGAGTG
AAACTGGTTTTTTTTCTTTTGTGGGGGTTACACCACAAAATGCATGGCAGTGAAAAAT
AACTTAGAGTTTGATGCCACTGCCACAGTTTGTGCCAAGGTGCCTGAAATTTTACTT
TTACCTACTGTTGCCTTATCACCCTCTTATGTCAACATATAGTTTAGCATAAACCAT
GAGATTTTAAAAAGTTATTCTACACTTGCATTATTTTCAGGACATGTGTTTGTGCCAA
GCTTTCACGTAAGAGTATCTTTAACTAGTTGGTGCTGATGCCTGGCAAATACAAGCC
AAGTAACAAGTCCAGCCATGTTTATGCACGCATCCATTGATAACGTATTAGCACAGT
CAGCCCTCCATATCCTCAGGTTCTGCATCTGCAGATTCAACCAAATGTGCTTGAAAA
TATTTGAGAATATTTAATACAATAGTACAAATATAAGTACACAATGTAACAACCTATT
TATATAGCATTACATTGTATGAAGTATAAGAAATCTGGAAATGATTTAAAGTATAT
GGGAGGATGTGTGTAAGTTGTATGCAGATACAGCGCCATTTTATAAAAGGGAATTC
AACATCCTTGGGTTTTGGTGTCTTGGCAAATGGCCAGCAAGGGTGGGGGAGTGCTG
TCCAGAAACCAATCCCAAAGCAAGGGACAAATATATACTTCTACAGATGAGATATT
AACTCAGAATCCATGTCTGCACACACATCTTTAATGACAAGTTGCTTTATCACTCAA
CAGCGGCACTACTATGATCTTTTACTTACACTTCAGCCAGCAGAGACATTCAGCAGT
GTGAGATTTTTTAAGTTTTTCAGCTATTGTGTATGTTTCTAGTGTATAATAAAGTAAC
TTATCCTTTAAGATACTTAAGTAGCTTTTTCATTTCTAGCTTTAAAACCTGTTTTTTTT
TTTTCCAGTAGTGGCATACTGCATTAATAAATAATGCCTTACCAAAAAAAGCACT
CTGAATGATTGGTTTTCAAATGATGCCACAACATAGGTGGCACCAACACTATTCAAG
ATCATTCCATTCCCATCTCTAAAAAATTTTTGGCTGGGTATGGTGGCTCACGCCTAT
TAACTCAACATTTTGAGAGGCCCAAAGCAAGATCACTCAGGGCTAGGAGTTGAAGA
CTAGCCTGAGCAACATGGCAAGATCCTGTCTCAAATTCTTTTTAAAATTTTTTTAAA
AGCCCAGGCTTGGTGGCGCATGCCTATAGTTCCAGCTACTCAGGAGGCTGAGGCAG
AAGGATCTCTTGAACCCAGGAGTTTCAGGCTGTAGTTCACTATGATGGCAGCTGTGA
ATAGCCTGGGCAACACAGCAAGACCTCATCTCCAAAAAAGACAAAAGAATAAAT
TATTCTACTGCAGAACATGATTAGGTAAATATCTCCAAAGCAGAAAGACAGGTTTCA
TATTTTCGTTAGTTTGAGTCAGTCCTTCAAATCAAATCTTGTTTTTTATTAGTATAC
AGATGGTATAGCCAGTAAGTAAATGAGAAGCAGTCTTTTTAAGCCGATCCATTCTTA
AATGAAAAAATATATAAATATTTTAGAATAAATTTATTAATTTCTAAAGTTGTAGAA
TTTTTAAATTTGGATATTTTGGGAAAATATTTAAACCACTATTGCAAACAAAACAAC
AAAATGTACTTATGTTTATACTTAGGCACAAAGAAAACCTACAGTATTTTAAAGTAAC
CATTACACAATATTGAGGTTGCAAAGATTACTGAAGGCATAACCTAAAAAATGAGT
TGATTTCTAAAAATGGGAAAAAGGAAAAAATAATTTCTAAAAACAAGTATGCATA
CCTAAACCTACCTAATGACACCTTAGAAAATTCAAGTATAGCACCATTCAATTAACAT
CAATGAGGATGTCATCACACATCATGTAGCCTCTGCACACCGTGAGAATAAATGAA
AAAGACAGGCATCTTGCTATCATGACAATAGTTTTGACCTCGCAGACCTCTCTGTGC
TTACGCAACGGATAAAGCCATAAGAAGTGTCTGCCCTCAAGGAGCAACCTAAAGT
AGGAAAAAACAATAAATTACACAATTATTATTTACAATTGTGAGAAGAGCTCTT
GACAACATTCAAATGGAGGATACAGTGTAGTAAGGGTGAGGGTATCAAGGCTTCT
TGAGAAGTGATGTTTTGAGGCCATTCTTTCTTCTCAATAACTGGTATTTGGTTCCTGA

ATCCTTTAACTTCCTTACCATTGTCACTCCTAAGCCAAATCTCATTACGTCATGTCTA
GACTACTGTTAAGAGAACCACTTAAGTGGTCTCTGCAGCCCTCAATTTATTGGTGTT
ATCTATGGGAAAATTGCTCAAACCTTAAGCCTTAGTTTCCTACCCTATAAAAATGGGG
TTTTATATAACAAGGAACATACTAAATACACAGGTATACCTCAGAAACACGGCAGGC
TCAATTCCAGAGCACTACAATAAAGCGAATCTCATGAATTTGTTGGTTTCCCAGTGC
ATAAATTATGGTTACACTATACCATAGTCTATTAAGAAGTGTACAATAGCATTATGT
ATAATTGATAAATACATCACTGCTAAAAAAATGCTAACAATCTTTTTTGCTGGTGGAG
GGTCTTACGCCAACGTTAACGATGGCTACTGACTCATCAGGGTGGTGGTGGTTGAAG
ATTACGATAGCTGTGGCAATTTCTTAAAAGACAATGAAGTTTGCCACACTGACATCC
TTTACAAGACTTCTCTGTAGCATGTGATGCTGTTTGATAGCATGATAGCATTTTACC
CACAGTAGAACTTTTTTTTTCTTTTTCTTTTTTTTTTTTTTTTTAAACGCAAGGTCTCACT
CTGTCACCCAGGCTGGAGTGCAGGGGCGCCATCTCGGCTTACTGCAACCTCCTCCTC
CCTGGTTCAAGAGATTCTCCTGCCTCAGCTTCCAAGTAGCTGGGACTACAGGTGTG
CACCACACCTGGCTAATTTGGTAGAGGTGGGGTTTACCATGTTGGCAAGGCTGGTC
TTGAACTCCTGACCTCAAATGATCTACCAGTCTCGGCCTCCTAAAGTACTGGGATTG
CAGGTGTGAGCCACCACACCCAGCCAGTGGAACCTATTTCAAATTGAAGTCAACTC
TCTCACACCCTGGCACTGCTTTATCAACCAGGTTTCTGTAATTCCTAAATCCTTTATT
GGCATTTTAACAATGTTACAGCAACTTCACCAGTAGATTCCATCTCGAGAAACCAC
TTTCTTTGCTCATCCCTAAAAAGCAACTCCTCATCCATTCAAATTTGATCATGAGATT
GCAGCAATTCAGTCACATCTTCAATGCTTTACTTCCAGTTCTAGTTCTCTTCTGTTT
CCACACCTGCAGTACACAAAAAGCATTCAATAACTATTACTTCATTTCTTCTACCTAT
GTTTCCATTAGCTTTTGCCTATAGTACGCACTAGAGTATGTTACCATTATTTGTTATA
AGTAGTACCTCATTATTACACTATTCGTAAGCAATACCTCAAGGTCTAAGATTAGAT
TTTAAATCAAGGTCAGTAAAAATAGAAAAGGCTGTGAAGACTGTTGACTGACTTTA
CCAGAATCCATACTAGAGGTGAGATTAGTTAGGTGATGAAATAACCATTCTATA
AACATGATCTGAAACTCTGTTACTGTTGTCAGCAGGAAAAGCCAATGTTACATATGT
TTAAAAAAGAAAAAACCACAAAACCAGAAAACAAAAGGTGACAAAGTATCAA
GACAAAAGGTCACTGATGACTGATCTCTAGGAAAAGCTGGAAAGCAGGATTATTA
ATGTAACCACGACTAAGATAAAAATCAGAGACAGAAAAGTCTTTGTCACCAAGAAG
ATATACTCCATGAGAGAGCAGAAACAATTCATCAGGTTTAAACCCTGCTCTAGATAAA
ATAAACTATCTGATTCAATACTCACACTTCTCTAATAATCCAATACATTATCCCATC
TCAAGAAGAGAGAGTACAGATAAGAAAAAAGGCTTCTTGAGAAGTATGTGCTC
TAATATAAACTAATATGCCACTAAGAAAGCAACCTGCAAAGTCCAGTACCAGACTT
CTGGATTTGTGACCTAACAAGGTGCTCTACAATTAACCTAACAGTCAAACCAGAGTG
TTGTA AAAAGAGAATTATGTAATTATGCCAAACCTCCACTCACAAAAAATATATGGA
AGTAACCTAAGTTTACATTTTGCAAATCTCACACACACACTAGCCCTGACAAAAGTT
TCACCAGCTTTCTCATCCAAGTACAAGCGTGTAATATACTTAATAAATTTGTCTTATA
AGGGTAAGAAATAGTATGTAACCTACTTGAAAAGGAGATAGGTAGCTGGTTAATTTA
AACAAAAAGCCCAAGGAAGTAAGGTGCAGGAAAAGGATAACTGCAATGATTAGTA
CAGGAAACCCAAAGAAGAAGTGAATGGTGGGATAGATGTACTCAGAGACCATGAG

GCATCAGTTTCCTCTATGAATAGAATATTAGGAGATGTAGGTTAAATGGGACCCTGA
AGTCTCTCCCAAAAAGCCTTGTTTATATGTTTTCTGAGCTTAACTATTACTTGAGAAT
CAATTCACGTATAAACCAACAAAACAACTAACATTTATTGAGCTTCCAGCTCTGTGCTT
AGGCACTGAAAAATCACTTTCCTTAAGGATTGCAATTAAGCAGGAGAAACACAAAT
AAGGTGAACTTCTCTTGTTTCGAAAGAATATATTTCAACATTCCTTTTTAAAAGGAAAA
CCTGACCTGCAAGTTTCCAAAAATATTAATTACTATTCCTCTTTGCCTCTCAAAATTC
CCATTCTGTTATTTTTTAGGAGGAGGAAAAAACAGTTCATTTGAGGAAAAATTGAGG
GTCACATACTATAACAATTGAGAAGAGTTTCTCTGAAACTGTAATCATTTTTGGCAGG
TAAATAGGCATATCCGAGTCAGCAAATGAACTTGAAGATACTGAGTTATACTGCCTG
CCCTGTGGGGTTCCACCTTCCCCAAAAGAATTCAGAATTTTTGGGTGATCTGAGAAT
CTACATTAAGACAACCTGTCTCCACACACAGGAGGCCTGAAGATCGCTGACATAAGG
GTCTTTTTAAAAGTATATTTAATGGCCTAGGGCGGTGGCTCACACCTGTAATCCCA
GGACTTTGGGAAGCTTAGGGCAGGAAGATCACTTGAGCCCAGGAGTTCTAACCTGT
GCAGCACAGCAAAAACCCATCTCTACAAAAAAAAAAAAACAAAAAAAAATTAGCTGG
GCATGGAAGCGTGTGCCTGTAGTTCAGCTACTCAGGAGGCTGAGGCAGGAGGATC
ACTTGAGCCCAGGAAGTCAAGGCTGCGTGAGCCATGATCATGCCATTGCAATCCAG
TATGTGACACTAAGACTCCGTCTCAAAAAAAAAAAAAAGATAATTA AAAATGTGTAA
GATACTGTATTAGCAATATAAAAAGCATTGGTGTTAAAATGTTGGTATTATAATTC
CTCAGGATAAAAACCTACTTTGTGATTGTTTTCTATAACTCAAGATATGATGCTTAGA
GCTCCTCCAATCAAGTGTTTCCAGGAAGTGAAAACCTGTAGGACAGAAATTTAGGCT
GGGTTCAATTTGTATCACACAGACCTATTCTTCATTCAAGTTCTGATATATTTAACTAT
GTAGCTCCTGTAACAGTTTAATGGAATCTCACCTCCCTAAAATTCATTATGCATTTTT
TTTTGAAATCCAACTCATTAACGCTTGCTTTCAGTGTGTTCCAAGGCAGGCACATCT
TTAAAATGGTTTGTGGACTTAGCTTTCAGCTAAATATATAATAAATAAAAACAAAA
CAAGCAGTTAAATGAAATGTAATGGGCCAGAGAGCTTCAGCTTTTATTTCTTACTG
CTCAGTAAAAAGAGAAAACCATCAATGTCCACGTATTCTGTAATCCACAGAACAAG
TCCGGGGCTACAGCTATACTGTCCACAGTTGCAATTCAAATTAGATAAAAAATAAA
AATTCAGTTCTTTAGTCATAACCAGCCACTTTTCCAATGCTCAAGATTAATAAAAATGTC
AAACCATAAAGACATTTACATGTCGCTCACTCCATTTACTTAAAGTTGGCTAGACAT
CAGAGTATACTAGGAGCTCAGGAGTACAAGACACTATTCCTTCAAAAAGCTCAGAA
TAGTTAAGGTAATTTAAATCAGCAATGACAACAACCCCAAGAAATTACTATGACCCAC
GCAGTACAACTGCTCAGGAGTCAGAAGAAAACCTGCTTTTTTAAAAGGGCAGTTTG
GGTCATAGAACAACAGACCATGGAAGGCATGACCAAAGGGGAGATGACATTTGAAT
CTGCAGGATTA AAAAGCAGCAAGGGTAGCATTCCAAAAAGAACCACCCACAAAAGAT
ATATGACGTCTCTATGATTTGGGTAACCTGCAATTCATTCCATGTGACTTCAGGAGAG
AGGTCATATTTGTGTGTGTAGTATGTGGAAAATAGTGAAAATGAAAAGCTGTTA
AATTGAGGAAAGTCTATCCAGGGACCTTATGCATCACATTCACGAGAACAAGAAATC
ATCCTGTAAACCAGGGGTGTCCAATCTTTCGGCTTCCCTGGGCCACACTGCAAGAAC
TGTCTTGGGCCACATATAAAGGACAGCTGATGAGCAAAAAAAAAAAAAACAGACAAC
AACAAACAAAAAAAAACACCCCGCAAAAAAAAAACTCCTAAAACCTTAAGAAAGTTTACG

AATTTGTGTTGGGTCGCATTCAAAGCTGTCCTGGGTCCCATGCGGCCCGCGGGTTAG
ACAACCTGCTGTAAACAGTACAAGCCAGTAATGGAGTTTCACCTGTCATTTTCATGC
TCTATCTTCCTTTAGGACAATCATCCTAACAAAGATGTAAGATGGATCAAAAGATAAC
ACTAAAGACAGAGACAGCAATTTGGAAGCTATCACACAGGCATCTGAGATCAGTTA
CTAACTGGTAAGAACAGAAATGAGAGGTATTTAGAGGAAGAAAAAGGGAGATGTT
GCCTAACCTCAGATCCAATTCTCTGTAAAGCAGTAGTCAAGATCACCTGGACTGTGA
AGACGGTCAGGGACAGAATCCCAGCTAAGGAAAAAGGATAAAAATGAAAATCAAGA
TAAACATTTAAGAACGTGAACTAGGGAGGAATAAAAGCACTGCTGGGTAAAGAGTCA
AGCCCCAGCTCAAGCCTTAATTTGTGGTGGAAACCAATCTGTCTGGTTTCGCGAGACA
CCAGGCTACCCAAGATCAAGAGAGGGAGAAAGCTAGTGCTATGTCTGAATACTAGA
GGAGCAAGTACAACAAATGGAAAATGGGATCAAGTATGAGTGAGAGTTGCTAAGAT
GCCTGGTAGGGATGCAAAGGGGTAGAGAGCCTGGGGAGAGAGGGTGAGGGAGGGA
AGCACTGGTTTCTCAAGCAAAAAGCTAAAATTTTTCTATTAAGATTTAACCTGATGCT
ACACTTTGGTGGTGCAGCAAGGGTCTCAAATGGTATAAACTCAGGTGATCATGCTT
TATGTCTGTCTCTAGAAAAATGCTCCAAAAATGATAAGTAGTGATAATCCGCAGTCT
CGTTGCATAAAATCAGCCCCAGGTGAATGACTAAGCTCCATTTCCCTACCCACCCCT
TATTACAATAACCTCGACACCAACTCTAGTCCGTGGGAAGATAAACTAATCGGAGTC
GCCCCTCAAATCTTACAGCTGCTCACTCCCCTGCAGGGCAACGCCAGGGACCAAGT
TAGCCCCTTAAGCCTAGGCAAAAGAATCCCGCCATAATCGAGAAGCGACTCGACAT
GGAGGCGATGACGAGATCACGCGAGGAGGAAAGGAGGGAGGGCTTCTTCCAGGCCC
AGGGCGGTCTTACAAGACGGGAGGCAGCAGAGAACTCCATAAAGGTATTGCGGC
ACTCCCCTCCCCCTGCCAGAAGGGTGCGGCCTTCTCTCCACCTCCTCCACCGCAGC
TCCCTCAGGATTGCAGCTCGCGCCGGTTTTTGGAGAACAAGCGCCTCCACCCACAA
ACCAGCCGGACCGACCCCGCTCCTCCCCACCCCCACGAGTGCCTGTAGCAGGTCG
GGCTTGTCTCGCCCTCAGGCGGTGGGAACCCGGGGCGGAGCCGCGGCCGCCCA
TCCAGAAGTCTCGGCCGGCAGCCCGCCCCCGCCTCCAGCGCGCGCTTCTGCCACGT
TGCGCAGGGGCGCGGGGCCAGACACTGCGGCGCTCGGCCTCGGGGAGGACCGTACC
AACGCCC GCCTCCCCGCCACCCCGCGCCCCGCGCAGTGGTTTCGCTCATGTGAGAC
TCGAGCCAGTAGCAAGGGCCCCGGTCCCACAGCTTCGACAGCCAATCAGGTGTGAA
GACAAGCAGGCGGCGGGTAAACCGACTCCCCGAAGGAAGGGGAGGGTGGGAGGA
CGCCCGCGCCAGAGCCGATTTCACTGACCCTCCCCTCCCGCCGAGGAGGCCGGCCG
CGCCCGCACACCCAGCATCTCTACACCCACCTACCTACCCGCCCCACCCAGGGGGC
AACGCGAGAGTCGCTAAGCGGCTGCGTACTCCCGACGGCGTAACTGACAGGAGCTT
TACTCCAACCAGAATACGCCATTTGTGTTTTTCACACACGGCGGGAGGAGAAACGGC
CAATCGGCGACAAGAGGCTAGCCGGAAGCGCTCCTCCCTCTGCGAGAGCAATGGCT
CCGTCCGGTTTCGAGCATTTTCCGCTCCCTTCTCCCTCCCCCTCCGGTTGCCGCAGGG
CGGGCCTCCCTCCCGCCTGCATCCAGCCACCCCTTTCCCTCCCAACGTAACAAACAT
TATGTTCCCGACTTCCCACGGGAAAGGCAACCCCGCAAGCCACCAGACGGCCCCC
CTAGCCACCCATCCCCCAGTGTACCGCACCTCCCCTCCCACCAGAGTTCCGCTCCC
CTACCTAGCCGAGGCTCTCTGAGGAGCCGGAGCGCCGAAGCACAGCCTCTTCTCTAG

GCGGCCCCGGCGGCTTCCGCTGATTGGCGGCGAGTGGGCCAATGGGTGCGGGGCGG
TGGGCGGAGAGGCCAATGGCGCGGCGGGAGGGGGCGTGTCCCGGGTGCCCTGGCG
CCGGCGCTGGGAATCCCCGTGCGGTACGTGGCGTTTCCGCTCGGGCAGCGGGCTGA
GTGAGCTGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCTGCCGGGGGAGGGG
CGGCCGCCGCCGCCCTGCGCTCAGAGACTCACGCAGCCCCAGTCCCGCCAGTCCGCC
AACACAGTAGTGCCGGCCCCCTCTTTCCCTGGCCCTGCCCCCTCCCCGCCTTTGG
CTCGCTCCGCCTTTCTGCCCCCACCCCCACCTCACGGGTACGGGGCCATTCGCGCC
AGGAAACGCCGTGGCGCCGCGTTGGGCCTAACTCGAGTCCTGCCGCCTCCCGGGAG
TGCCGTGCGCCGACGCCGGGCCAGGCCCGGCAGCGCCTGGGACAAGGTAAGGG
TCCGACAGAAAAGAGACCGAACCTCACGATCGGGCCCCAGGGGAGGGAAGGGTCA
CCTCCTCCGTCTCCCCGCGCTCGCTCTCCTTGGGTGCTGGGCCTGGCCCTCCCCAAGC
TCTTAGGAGGATGCTGCCACTTCTACCCCCCTCGCCGCCTTGCACACACCGTTGCA
ACACCCCATTTTCCAGGGAGAGATCCCCCTCTAATCTAGGCGACCCAACTCCCC
CTTTCATGTTTTTCTGGGTACAGGACGCTTCCCCTCCCCAACGCCTCTTACCCCCCT
TTCCTGGGAACTGCCTACTCCACGTTTACCTTTCCCTTGAGGAGAGGCCTCTTGCTGC
CCTCCGCTCGAAATACACAGGCATACTTTTTTCTCTCCCCGATCCCCACTCCCTAC
CCCCGTTCTCGCGGCCTTGTGACAGACAACTCTGATCGCTCTGGGGGCGCGATCTC
CCCTCCGTAATCTTCCTGGACGCCTTCCCTCTCGTTTTTCTGGCTTCCCACCTCAGATG
GCTGCTTCCCAAAGGCATTACCTTCGCCACCCCCACCACACGTTCTCTGGCTCCCCGT
GGCGTGTGCCACAGCGTGTCTGAGATAGCCTCGTTGAATGTGTAGGGTTCGAGCCTG
GAGTTGAGCCAGATTGTGTCGTTTTACTTGCCTTGGGCGTGGAGAACGATCTTGTGA
GAATATCTTCAAAGGCAGAAAAATATCCCTTTATGAATTCTCTTCCCTCTGCGTGT
AAGTCGGGAATGTGAAGAGGAGTGTAGGAAAGAGCCCTGGTTCAAGTAGGTAATC
GCATGAGAGGGAAAGTTAAACTGTTGGGAAAGCCCCTTCTATGCTAATTGATTCTAT
AGAGTCCTTGCTTGTCTCACTTCTTGGGCGTCAGTGGTCTTTCTCTTGGATATGGATG
CTGCAGTCAGCTCTGCTGGTCTGGGTCAGGGGTGCGTGTATGACCTGCATTTTCTGC
TTTCTCATGTTACTTGTGCAATGTATTCACCGGTAACTCATTTCTTTCCAGACCTCT
GGGTCCACTGGGCTTTGTCTATATTTAAGTTCATTTCTCCAGTTTCCCTCCTGCACA
TAGGTACTGAACGAATCCCCAAGTTCTGTGCTAATTACCTTCATCAGTTGACTAAAC
AAGTTTTTAGATGACATATTTGTGACCAAGGTCATATTTACATTTCTTTGTTGGACAG
ATGTTACATAGCTATACTTGTGATTGGGGAGGATCCAGCTGAGTGGAGTGTGCTGAG
CTTTTTAGGAGAGTGTGTACTCCCTATTTGAAATTATTTTTTGGTTGTTAATTTTATAT
TATTAATGTTTTTLAGGTCACAGAAAGTTCTAAGTGGTAATTTTAGATGTGTGGGATC
TGAGCTAGGACTAAAGCAGAGAATACCCACGTAATCAGAGGTTTCTGGGCTCCATA
GAGGACGTAGGGCTTTTTTTTTTCTATTGGATTTCTTCCAGTTTTCTCAGGATCATA
GTTCTCTTCTGTAGCCAAAAATTCTGGCCTGTTATGGGATTAGAGTCTTTAAGGTTTA
CTCAGACTGTCATTATGTGTAGAAAAATGAATTATGCCCTTTGGTAGGACATGACAC
AAGGCTCTGTTTCTAGCTGCAAATTTAAATTAGATTGTAGAGTGCTTGGGAAATTGG
CTTCAAAGACCAAAGCTTAATCTTCACTCCTAAACTGCTGGCTTAATTAATAATGG
ATATTTAGAATTTGGTAAATGTTGATTTTTCTAATAAAAGGCCTTGGTTTAAAAGGG

TGACCTTAGGATTGTTTCTTTCTTAAAAGCATAATTCCAGCCCTTCTGGCATGGAGCA
 CTGGTCCAAAAAAGTGTGTGTAAGGAGTGGGGGTGGGGTAAAGA
 GAAGTTGTTCCCTTTGGGTTGGATCACAGGGGTGAGTATACAAGGCAGCAGCAGCT
 GCTGGCTCTGGAGCTCTGGTTGCTACGTGAGAAGCTTGAGTAGTGCTGGCTGCTGTC
 TCCAGGGAAGGACAGCAGTGCAGCGTCCATTAATGCTGCTGGCTGCAGGGAGCAGC
 ACTTAGGCGATGGCTGCTTCAGGACTAAGAAGAAACCTTGCTTTTCTGGGAATTTTC
 ACTGCTGAGCTGGTTTGTCTTTTATTGGTGGGGAGATGGGAATTAGTAATTCATAAT
 CTCCTACCCATTTATGGATATTGGCATCTGGAAACTGGATCATGGTTAAAGCCTTTCT
 TTTTTGTTTGTGTTGATTTGATTTTTGTTTTTGGCAGATTTTTGTTTTTATCTAGACA
 TTTGTGCTTGGATAGGACTAAAAGTTCATTAGAGTTTTAATTTTTCAATCAGTTTAA
 AAACCCAAGTAATAATTTAAGAATCTTTCTGATAACCACAATAGGAAGAAAATAA
 CAGGAATTTTTCTGCAGCTCACATATCATGCCTTCCTCCATCTCTTTAATCATAGA
 ATCAATTCTTATTATTTGTTATGTGTCTCCATCCTTTCGATTAGACCACATTTACCTT
 ATAGACGATTTGCTAAACATTTTACTAAGCTTGAACCTTAAACTCTAAAAAGGTGC
 CATTTTGGAGTGGTTTCTAAATAAATATTTTTAATTTGTATATTAGTAATAAACTTCT
 CCAGATTAGATATTTCTTTGGAGTTTGACTTATAAGATTGATTCATTATATACATGT
 TGGATATAGCCTTCTGACATCACAAATATATGTCTTTGGCCATAATCCATCTGAAAT
 GTAGGACAGACCAGAAGAAATATGCAGAAATCGAATAAGTCTAGTTCAGGATACTG
 AGAAGATGGCCTCTGAGCCCCTTAGGTGATCTCCCCTCCCCACAACCTCCTGAACAT
 TAGGATGATCTCTGATTAAGCAAACAGTCTGAGCGTGGAAAACTTGAAGGAGAA
 CCACCACCACCAATTATATGCAATACTGGACATATTCCTGTGTGCTGTTTTTCTTCCC
 CAAGACTCGTGTATCCTATACTTTTTTCTCTCAGAATTTTGATTTGTTTCAATTTTCGTGT
 AAATGTACTTAAATCTCACAAACATCTATAATTTGTAGTATCACTCTGGCATTGTG
 GCAGAGAACCAAAAAGAATGGAAATGAGTTTTGTCATTCACAAATGTGGCTCACAT
 TGTTTTCCAGTAATAAAAGCAGACCAATGAAACAGAACCTTTAATGGATACTATTT
 TAGGAGGTTCCAATTCTTATTACTATCACATAGATAAGATGCAATAGCAGATAAATA
 TGATTCATGTATACTGGCTGTTTGACATACTTAGGGTTAAGATAAAAATGTTTGTA
 GTTTTTACTCTGTGGCTTAAGTTGCTATATAAAATAATTGCTTTTACACTCGAATTT
 CCTGTTGTTTGGAACTTTTGTGCTCTTGATATTATCATTTTTTAGAGGATCATAACAG
 GCCCTTTTCATAGAAGGATTTACTTAAGTTATACCCTTGAAAACCTTTTTATATCTTT
 TGATACTGTTTTGTGTCCAGGAACTGACTTTCTGAAATTATTCTGGCTTTTCTGGGGA
 GAATGACTATTTCAATTTTACCTTTGAATGGGGAAATAATAAAGTGCAAAGTACAGA
 TTTGCAGATAATTACTTTTGTCTTATCCTCTCCATGTTGAAATAACTTATGAAAAATT
 AGGCCATAGTTAACAGCAGTCAATGACTATTGGATACATTTTATCAGAGGGGAACT
 GGATCATGAATAAAATAAAATTTTAAAAATAATTTTTGGCTGAACTCTGGTGATTCA
 TCAGTTTAATTTGAAGTCAGAAGGTCTAGCAGTGAATTTTATTTATAAAAATTGTAT
 TTCAAGTGTTGAAAACCTGAACTTCTTGACCAGTATATTTTGTGTTGAGGCATCAAAC
 TTTGCAAAATGTGCATCGTATATTTAGTGATATAACTGGTAGTCATTTGTAATTTAAA
 GTATTCTTTCAAAGGCACTCTTTAGAAAGTAATGTAGTGTACCCGTGATGGGCAGGG
 ATTGGTACCATTCTTACTGCCAAAAATTCCAAAATATGTGGCAAAATGATTGATTT

ATCTTGTGGGTGGGATTCTGGGAAGTTCATGAAAGGTGGAGAGAATATAGTTTCCTT
CACTTGTCTATATACATTTTGTAAATAAGTCTTAGGAAAAGTGTTTTATTGTATCTT
TAATTATGAATTGCGTAAAAGATAACCCAGTAACTTTGGGGGGAGGTGCTGTTAGAA
AGCATTACATTGGAGAGAATTCCCCTACCCTGGGACAAAATGCATTCTGTCTTTAAT
ACTTAGCGAAGGGAACTATGGGATAAAAATAACAATGAAGGTAAGCTCAGTCTGCT
TTATATGTGCCCTCACTGAGCAAGGAATTTGTAATCGCATCGTGCCTCATTTCGTTTAT
ACCATCATATTGATTTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
GTCACATTAAGTATGTTTGATATGAAGACAGGGAAAGGAATTTTCTGCTTCTTGGAG
TAATGTCTTAGTATTTTTAAAACACTTAAGTTTTTACATCAGGCCAGTTTTGCCTGAT
GCTCATGTCTGTTGCTTTGGTTGGGCTGCTGCTTTCTCTTCTGTGTTCTTATGGGTTCCG
TTGTGGTATAAAGGATCCCACAGCTTTTCATGGCAGTATGAAGTAATGAGAAGCATTG
CCTTAGCCATGTTAGTTACATGTATACTTTTGGCCTATGTTATGAATCACAAAAGC
GGTAGCTATAGGAATGTATACAAAATAGATTTCTGTCTGGGGAATCAAGTTTTTGTAT
TTGTGCTACCTAATGGAGGGGAAAATGCTGAATTTCTTGCTGCTCTGTTTGAGAAAT
AGATGGAAGCATGGGAGGAGCCAGAGACCTCTGCAGCAGGATTTGGTCTAAGTAGA
AAAGGAAGATTTTTGTTTCAAATTGCCAGCTGCTTATGTCAGACTGACTCCCTTATTA
TGCCTCCAGTAGGCCTGTCAATATGGCCAAACAGCTAGATAAGTGCGGGGCAGGAC
AAAGGGCTCTTTGCACAGCAGGGAGGCAATGTTGGTGGGGGAGGGGCAGGAGGTA
GGAAAGGCAAGAGGAGGAGGTTCTTTCCCTGGGAGATTATTCAGTTTGGCATAACA
ATTAAGAAATCATTTTTAGTTCCCACTCAAGCATTGAATTTTTGCCAACACATACT
ATTAACCCCAAATTTGATACATTTCAGAATATCTTGTAGGGATCCATTCTCGCCAAG
GAAAAATAAAAAATAAATAAAGCTCTGTATAGGTTAAAATAAATAAATCCCACA
CTCTGCACCCTCCTAGGTGCAAGTCACCTCCCGAGGAGACCCGTTCTAGAGCTGAAT
TCTCATTAAAGAAATGGAAAAGAATACTCTATCTGAATAAAAACACATTGTAATACA
ATGTGTTTATTTGGGTTGGGATTGGACCTGAACATGTAGAATAATTTGTTTCCCTTTA
TGAAATAGTTGCTCGTAGTTGTCTACAATTTTATTTCATTAAAGATAGGTAGCACATTA
CAGCTTTCATGTGTTGGGTTGCCATATGTAAAATGCTAACTGAAGAAAGGCTACTTT
TTAATTTAGCCTCATCCTTAGTTCCTGGAGAACCTGATATTTCTGGAGATTACTCC
CTCCCCACCTTTTAGTTTAGGCAACCTCTTTTGATACATTTGTGTTTCTGCTCGCATA
CAAGTGGGATAGTTGCATCCAGTTTATTAAGACTTAGTATGAATCATAGAGTTGGAA
AAGATCTGTTGGTTATCTGGTCCCTTAAACCAAAAATCATAATGAAATATTTTGAAAT
TTGGGTCCCTATTGAAGTTTTTATTAAATGTTAAAGGATCGGTGTTCTGAACAACA
TTTTTAGTTACTTTTAAAATAAATGTTTTGCGTCAGTTCTTTTTTTAAAATAAAGAA
TTTCATTTATAGGCAAATTAGCTGGCAATTATTTGAATTGTGATAGGATTTCTCTTTT
ATGAAGGAATATATGACAAGGTTTTTCAAATGCTTAATATATTTTTAAAAGACTTTA
ATTTTTAGAAATAATTGGTTTGAACAGTTTTCCAAGAGCACATTTGTTGCTTGGGTTG
AGGTACCACCTATATTGCAATGTTACTAACTAGCCTTAAAGTTTTCCCTTCTGTCTA
TACTGCATGCAACAATAAAGGGAACTGGAATGTTAATTTCCATTTATGGATTAGCAG
AGGAGATGTTTTAACCGATTAATAACCAAAAAACTGCCTTTCGTACACGTAATATTA
AGCAAGCCTGACCAAGTTTTGTGTTATTTCTCTCTGTAAAGAAAAGTGGATGTGTT

ACTACTTAACATTATATTGTTATTTAATGGTCTTGGCAGTAATGATATAATATTTTCGA
 CAAAAGAAATTTTGAGTAATTAATTATTATTGTAATTAGTTGGAAGTTTCTCATCA
 GTAAAATAGCAACAGCATTAAACACAAAATCTAGTGAGCTATATTTTATATTACTACA
 GAAATTTAGGGTAGTCATTTCTTTCTTTATAATTTATTCACATGGATTATTTCCATAA
 ATTTGTGGGACTAAAATAGAAGCCATCTAGTCAAGCACCAGTCTCCATAACCAGACA
 GTTTTCTCTGCATGTGCTATGACCCACATTGCCAGTATTAACATCCTTTACACCCTC
 CCCCTTCCCAGATAATTAGAAATCTCTTCAGGGTAGCTTCCATTGCTCCTATTACCTG
 GATCTTGCTAGAGGCTCTAAGAAGTTCCTGGTAAAAGTGAGACAGTAAGGGACCAC
 ATTTTGATTCCAAAGGTTTTGATAACTGTTAGGGCTCCCCAAACAGCTAATCTCATT
 TCACCAAGACTTAGCCAGCAGAGGGCTGGAATGGAGGTGAAACACAAGCACTGTAC
 CTCATCTTGCTGTGCAGCTGCTCCACCTTATTTCTGCTATTATTATCTCACAAACGC
 CTCCTCCCATCAAAAAGAACTAGGACAAAGGGGGAAAATTGGATGGGCTAATGTG
 ATTTTTATTATGCTAGGTTGTGGGCTTGTTTATATGTACTTAAATACAAAGCTAATTT
 GCCCCATTCTTAAAAGTCTTTAGTGATAGAGATTTTGTAACCTTCTGTATCTTCTACTT
 TCTTTCTTGATAAACCATTTTCAGATTCTCAGCCTTACAGAAAGAAAGGTTTTAAGCA
 TACTTAATTTTCGTTGGCCGTTACAGTCATTATTACCACCAGATGCCACTGTATTAT
 TAGCTTGAAGAAAGGTGGGCTCTCTTCTGTACATAATATCTGCAATTTGTTTTGGAA
 AATACTAATTTGTATAAATCTGATTTATGACTAAAATAAGGTAAAAATTAGACCTC
 TATGTATGTTTACCCTATTACCTTAGTGGGGGTGAAATTAATTAGCTCTTTGAACATA
 AATTTTTCATGTCTTAGAGTTCTTTTTTCAAGCTGCATAATTTATGTTCTTCAAGCCAT
 TTTTATCCCATACCACCCCCACAAAGGGGGAAAATTTTATTTTTTATCATTTTTTATTGT
 CTTTCAATGGTGAGATTTTCGCCACCCCACTCCTGAAATGTGAAGACTCAAATAAAA
 CTGAGTAATCTAATAAGGTATATGCGTTGCTGAATGTAGTAAGATGATTGTTTCATC
 ATTCTTAGATATTATGATCTAGTTTGAATCTGGTTTCCAGTATCATGTTAGCATATTT
 AACTGTGATATGTTAATTTTAATACATGCCAGGTGGATCTCCTTGCTTTCTATT
 TGTGCCCTTGTTTGTGTTTTGTATGAAGGGGGTTTTTGTGTTGGATTTTCTTCCCC
 ATCTCTGTGTCCTGTTATGTTCTTTGGCTTATGTTTCAAAAATTCTGTTTCTTACCACC
 AACCTCTGTACATGCCACAACACATACAATTTGTACTTTACAGTTTCTGTGAAGTA
 GGATGATCTGCAGTTAATAATCAACTGTTTGGGCATTCTTGGTATCCAAGGAAGGTT
 TTACTIONAGAAGGAAGAACCTGGAAGGACCTGTTGGCAATTAGACTACTTCTGCGTTT
 ATTTTACATTTTCCCTTATTAACGTAGGCTGTTGAGAGTTGACTTGTTTTATAAGAGA
 AACAGATTGACAGAGAAGACCCCAATCAGATAGAGTTATTTTAAAAATAAATGT
 GTTTATTATGGTAACATTTGGGGTAGAATCTAAAGGGCATATTTTAAAAAAACTTT
 TAGTTCTAAAGACAAAAGAGTTTAAACCTAAAACAGAACAAAGAGAAGGGCCTTTGA
 AGCAGTATGATTGATTATAT

ПРИМЕР 11 - Последовательности стабильного сайта 1 СНО и мыши - Патент США № 7771997

211> 6473

<212> ДНК

<213> *Cricetulus griseus*

<400> 1

(SEQ ID NO:5)

tctagaaaca aaacaaaaa tattaagtca ggcttgctt caggtgctgg ggtggagtgc 60
 tgacaaaaat acacaaatc ctggctttct aaggctttt cggggattca ggtattgggt 120
 gatggtagaa taaaaatctg aacataggt gatgtatctg ccatactgca tgggtgtgta 180
 tgtgtgtgta tgtgtgtctg tgtgtgtgcc cagacagaaa taccatgaag gaaaaaaca 240
 cttcaagac aggagagaag agtgacctgg gaaggactcc ccaatgagat gagaactgag 300
 cacatgccag aggaggtgag gactgaacca ttcaacacaa gtggtgaata gtccctgcaga 360
 cacagagagg gccagaagca ctcagaactc cagggggctca ggagtggttc tctggaggct 420
 tctgcccttg gaggttctg aggaggaggc ttccatattg aaaatgtagt tagtggccgt 480
 ttccattagt acagtgacta gagagagctg aggaccact ggactgaggc ctatgctc 540
 agtcagatgg ccatgaaagc ctagacaagc acttccgggt ggaaagaaa cagcaggtgt 600
 gaggggctcag gggcaagtta gtgggagagg tcttccagat gaagtagcag gaacggagac 660
 gcactggatg gccccactg tcaaccagca aaagcttggga tcttgttcta agaggccagg 720
 gacatgacaa ggggtatctc ggtttttaa aggctttgtg ttacctaac acttctatta 780
 gtcagatact ttgtaacaca aatgagtact tggcctgtat tttagaaact tctgggatcc 840
 tgaaaaaaca caatgacatt ctggctgcaa cacctggaga ctcccagcca ggcctggac 900
 ccgggtccat tcatgcaaat actcaggac agattcttca ctaggactg atgagctgc 960
 ttggatgcaa atgtggcctc ttcatttac tacaagtcac catgagtcag gaggtgctgt 1020
 ttgcacagtg tgactaagtg atggagtgtt gactgcagcc attcccggcc ccagcttctg 1080
 agagagatcc ttttaaatg aaagtaagct caaagttacc acgaagccac acatgtataa 1140
 actgtgtgaa taatctgtgc acatacaca accatgtgaa taatctgtgt acatgtataa 1200
 actgtgtgaa taatctgtgt gcagccttc cttactact acctccagt gatcaggttt 1260
 ggactgectg tgtgctactg gacctgaat gtccccacc ctgtcccctg tcttttacga 1320
 ttctgacatt ttaataaat tcagcggctt cccctctgct ctgtgcctag ctatacctg 1380
 gtactctgca ttttggttc tgtgacattt ctctgtgact ctgctacatt ctcagatgac 1440
 atgtgacaca gaaggtgttc cctctggaga catgtgatgt cctgtcatt agtggaatca 1500
 gatgccccca aactgttgc cagtgtttgg gaaagtgaca cgtgaaggag gatcaggaaa 1560
 agaggggtgg aaatcaagat gtgtctgagt atctcatgct cctgagtggc ccaggctgct 1620
 gacttcactc cccaagtga gggaggccat ggtgagtaca cacacctcac acatactata 1680
 tcaacacac acacacacac acacacacac acgcacgac gcacgcacgc acgcacacat 1740
 gcacacacac gaactacatt tcacaaacca catagcata ttacaccca aacgtatcac 1800
 ctatacatac cacacataca caccctcca cacatcacac acataccaca cccacacaca 1860
 gcacacacat acataggeac acattcacac accacacata tacatttctg tatgcataca 1920
 tgcatacaca cacaggcaca cagacaccac acacatgcat tgtgtacgca cacatgcata 1980
 cacacacata ggcacacatt gagcacacac atacatttctg gtacgcacac tacatagaca 2040
 tatatgcatt tgtatagca cacatgatg cacacataca taggcacaca tagagcacac 2100
 acatacattt gtgtatgca acatgcacac accaatcaca tgggaagact caggttctc 2160
 actaaggctc acatgaactt agcagttcct ggttatctcg tgaaacttgg aagattgctg 2220
 tggagaagag gaagcgttgg cttgagccct ggcagcaatt aaccccgcc agaagaagta 2280

ggtttaaaaa tgagagggtc tcaatgtgga acccgcaggg cgccagtca gagaagagac 2340
 ctaccaage caactgagag caaaggcaga gggatgaacc tgggatgtag tttgaacctc 2400
 tgtaccagct gggetcatg ctatfttftt atatctttat taaatattct tttagtftta 2460
 tgtgctgtaa taccttgctt gcataaatgt atgggcaactg tatgtgttct tgggtcccgt 2520
 ggaggccagg agagggcatg gatctccgg agctggcgtt tgagacagtt gtgaccaca 2580
 gtgtggggtc tgggaactgg gtcttagtgt tccgcaagtg cagctggggc tctaacctc 2640
 tgaccatcc ctccagcttc aagaaactta tttcttagg acatggggga agggatccag 2700
 ggctttagc ttgtttgfc agcaaaact ctttctgtgt attttgaatt ttatfttatt 2760
 ttactftttt gggatagaat cacattctgc agctcaggct gggcctgaac tcatcaaat 2820
 cctctgtct cagtctacca ggtgataaga ttactgatgt gagcctggct ttgacaagca 2880
 ctfttagagtc cccagccctt ctggacactt gttccaagta taatataat atatataat 2940
 atataatata atataatata atataatgt tgtgtgtgt tgtgtgtgta tgagacactt 3000
 gctctaaggg tatcatata atcttgatt tgcfttfaat ttatftttta attaaaaatg 3060
 attagctaca tgcacctgt atgctctgt atcatctata tctcttct tcttctctc 3120
 tcttctctc ttcttctct cacccecaag catctatftt caaatcctt tgcagaggag 3180
 atgccaagag tctcgttggg ggagatgggt agggggcgat acaggggaag agcaggagga 3240
 aagggggaca gactgggtgt ggtctttgga gagctcagga gaatagcagc gatcttccct 3300
 gtccttggtg tcaccttta cagccaacac cftttgtgg cctggcagaa gagttgtcaa 3360
 gctggtgca ggtctgccac acaaceccaa tctggccca agaaaaggca cctgtgtgtg 3420
 actctgggggt taaaggcgt gctgtgtct cccagctgg acttgaact cccgttfaat 3480
 aaagagtct gcaaaataat acccgcagag tcacagtcc aggttccct gcttctctga 3540
 agcggcaggc acgggttccc taggaaatgg ggccttgctt gccaaagctc cacggcttc 3600
 cctgcaaac gctgaaatga tctggcactc tgcgttcca ctgggatgaa atggaaaaaa 3660
 gaaaaagaag aagtgtctct ggaagcgggc gcctcacac aaaccgcaa cgattgtgta 3720
 aacactctcc atfgagaatc tggagtgcgg ttgccctta ctggggagct gaagacagct 3780
 agtgggggag gggggaggac cgtctagca tcttccacg gtctctctg gctgtgtgtc 3840
 atccgggaa ccgaaacgcg gaactaaagt caagtcttc tttgttgaa ctgacaatca 3900
 acgaaatcac ttcgattgtt ttctctttt tactggaatt ctggatttg atagatgggg 3960
 gaggatcaga gggggagggg aggggcgggg agacggaggg agggggggag gaggggagga 4020
 ggggaggagg ggaggagggg aagggatgga ggaaaact aactttcta attcaacatg 4080
 acaaagatc ggagaaagt caccgctagt gaccgggagg aggaatgccc tattgggcat 4140
 tatattcct gctctaat ggaatcaaac tctgttcc agcaccaag attctgagcc 4200
 tctctatc aagacagtaa ctacagcca cacggaagag gctataaac tgaagaata 4260
 aaatftcac tttatftcat ttctgtact gcatgtcac atgtagagag ccacctgtgt 4320
 ctaggggctg atgtctggg cagtagagtt ctgagcccgt taactggaac aaccagaac 4380
 tcccaccaca gttagagctt gctgagagag ggaggccctt ggtgagattt cttgtgtat 4440
 ttatfttag agagggtct atactgtagt ccaagctagc ctccagctca cagaaattct 4500
 cctgttccgg tttccaaagt actggagfta tgagtgtgtg ttaattgaa gctaagaatt 4560
 tctgtattga agaaaacctc aagtgggttt ggctaatccc cagacacca gaggtgagg 4620
 caggaggaat gagagaatc aaggttgc agagccacag ggtgagctca atgtggagac 4680

tgtgagggtg agctcaatgt ggagactgtg aggggtgagct caatgtggag actgtgaggg 4740
 tgagctcaat tgggagactg tgagggtgag ctcaatgtgg agactgtgag ggtgagctca 4800
 atgtggagac ctgtatcaag ataataatag tagtagtaac aatgcaggcg aggggtgtgtg 4860
 tgagtgttag agcagttagt tgatttgaca tcttgaggt ctcccgtcc atctgtggcc 4920
 ctgcaacagg aagggagggga ggaagggggg gaacgagaga gaggaaagag agacagaagc 4980
 taagataggg aatgagagag gaaggaagaa acgggaagaa attcagactc ctctctgagt 5040
 tccccaacg cctagtaca tctgtgcaac acctaaagt ggcctttgtg tggcactggc 5100
 ttgggtggtc gggaaaggca tttcagctt gttgcagaac tgccacagta gcatgctggg 5160
 tccgtgaaag tttctcccc ttaacaagaa gtcttacta cttgtgacct caccagtga 5220
 aatttttta attgtctct ggtgttctg gttttgcatt ttttttcta aggatacatt 5280
 cctgggtgat gtcatgaagt ccccaaagac acagtggggc tgtgttgat tgggaaagat 5340
 gatttatctg ggggtcaaa aggaaaagaa gggaaacagg cacttgggaa aatgtctcc 5400
 cgccccccg aattttggt tggcaaccgt ggtggaggag caagaaacac gtggacgttt 5460
 gaggaggcat ggggtcttag gaggacagga agcagaagga gagagctggg ctgacagcct 5520
 gcaggcattg cacagttca gaaggagatt acagcatgac tgagtttta gggatccaac 5580
 agggacctgg gtagagattc tgtggctct gaggcaactt gacctcagcc agatggtatt 5640
 tgaataacct gctcttagag gaaaacaga catagcaaac agagccacgt ttagtatga 5700
 aactctact ttgctgagt catgtgcggc catgccagg ggtcaggctg aactcaact 5760
 caaaaacaag tgagaaattg aagacaatcc gtgggtggcag ctactggaag ggccaccaca 5820
 tccccagaaa gagtggagct gctaaaaagc cattttgat aggcacagtt atcttgaatg 5880
 catggagcag agattacgga aaaatcgaga atgttaatga ggcaacattc gagttgagtc 5940
 attcagtgtg ggaaaccag acgttccat ccctaaaag gaacatttg ctctcagta 6000
 aaatgaaat aaaattggg gcttgaattt ggcaaatgat tcagaactct gtgtaggtat 6060
 tttcacacgc acagtggata atttcatgt tggagttat ttgtgctaaa aggcagaaaa 6120
 gggtaaaaag cacatctaa gagttatgag gttctacgaa taaaataat gttacttaca 6180
 gctattcctt aattagtacc ccttccacc tgggtaatt tctgagata gtcagtgggg 6240
 aaaagatctc tcttctctt cttctctccc ctcccctct cteectctc cctcctctc 6300
 ctccctctc tccctctc ccccttctc tctttcttg ctctctctc tetgectct 6360
 tctccttctc tcttctctt attctaagta gcttttaaca gcacaccaat tacctgtgta 6420
 taacgggaaa acacaggctc aagcagctta gagaagattg atctgtgttc act 6473

<211> 7045

<212> ДНК

<213> *Cricetulus griseus*

<400> 2

(SEQ ID NO:6)

actagcgtgc aattcagagg tgggtgaaga taaaaggcaa acatttgagg ccatttctt 60
 atttggcag gcacttagga agtggacat gcctaactca ctggtttgta ccaccttcc 120
 ctataatgga ctgtttggga agctcctggg caaccgattc tggcatctca ttggtcagag 180

ctctgtgagt tcgagaccag cctggctctat aagagctagt tccaggacag cctccaaagc 2640
 cacagagaaa ccctgtctcg aaaaaccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2700
 gfatggattc taagaaagcc gtaacagctg gagctgtgta cggagttcag cgtggtacta 2760
 gaagaacaga cattcatgat gaaacacccc aggattttta cttagtatct agtttccatt 2820
 gttgttttga gaccggctct tatgtcttcc aggctggcct caaactgctg atcttcccgc 2880
 ctctacctct caagtcctgg gactacttgg ctcataaac agttttgtc gggtccctg 2940
 aagtatggt tgtacaaacc gtgggggtca atatactcac ttgggcagag agagaaggtc 3000
 tgaatcccag acaatgactg catctcagga cagtgggaa gaggacaatg gcagaaggac 3060
 ttgaaaaaga tagactggag ggtggaaaag cagcaggaac agagaaacaa aacaggaagc 3120
 ttgctatcca gggccactct ggagtcctgt ggcaagatgg aagcgggcta ggggaataca 3180
 tttgtctac tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgat caatgcctat 3240
 caatgtttaa ggggaaatat gtataccaca ttgattctgg gagcaattct cagtatctgg 3300
 cctagagaaa ggaatggccc ctgcagaata gacagagtga atgggtccct ttatcatttg 3360
 ctaaagttaa ggagaaataa acatccttcc atagagtffc aggtaaatga accccacagt 3420
 tcactgtgc cgtggtggag gctggccaa cagttaaaaa gattagacac ggacaaagtc 3480
 tgaaggaaac acctcgaata ggaagaggag agccacctca ttctgtaact ttctcaagg 3540
 ggaagatgtt ccaagagtgg gaataaatgg tcaaaggggg gatttttaat taggaaaacg 3600
 atttctgta tcactgtga aactggaggt tgatttggg cataggacaa tagatttgat 3660
 gctttgaaa aagctgttc aaagcagaga aatggaatag agacaattat gtagcgagga 3720
 gggaggggtg ggcaagatg gagacagaga agtggaaagct gactttaggg aagaggaaca 3780
 tagaccacag gggcggggcg gggggcaggg gcggggggcg gggtcaaag gaggcagtgg 3840
 gaacgttct agtfttcga gcgtaagcgt gaatgtgcaa gcgttttgt ggtgtgtgac 3900
 caggagtgc gtggctggct tgtgtgtgc ttgtaatccc agtctttgag gtttccacac 3960
 tgttccacag tgggtgtgat ttccctcgg agagcatgag ggctctgctt tccccacatc 4020
 ctccccagcg ttctttgta ttgtttcca agatgttagt gggtagaca aagcctctct 4080
 gttgatttgc ctftaacagg tgacaaaaa agctcaacca ggagacattt ttgccttctt 4140
 ggaaggtaat gctcccatgt agagcaatgg gacctctc taagtgagg ctactcttgc 4200
 agtttgcacc cagctcttct gatgcaggaa ggaagtggg gggcaagcaa gactgttgc 4260
 ttcttgcgat ggacacattc tgacacaaa ggctcaggag gggagaagge tgtttgatgt 4320
 ttgactca ggaaggcccc tgatcatct gtgattagct gctccatct gtggagcaga 4380
 cacggactaa ctaaaaacca gtgtttttaa atgtcaagc ctftaaggtg aggaaattga 4440
 ctattgtgc tggccatac gtagagcaag tgctctgcat tggccaacc cccggctctg 4500
 gtttctaggc accagaatgg ctagaacta actcacaate ctcccattcc aggtctcagg 4560
 tgctagaatg aaccactata ccagctgccc tgctgctta cctgccttcc taaattttaa 4620
 atcatgggga gtaggggaga atacacttat cttagttagg gtttctattg ctgtgaagag 4680
 acacatgag catggcaact ctataaagg aaaacattta gttgggtggc agtttcagag 4740
 gttttagtac atgtcatca tggtgggaa catgatggca tgcagacaga catggtgctg 4800
 gagaaaggga tgagatcct acatcttga ggaacagga cctcagctga gacactggct 4860
 ggtacctga gcataggaaa cctcacagcc caccctaca gtgacatatt tcttcaaca 4920
 aagccatacc tctaatagt gccactcct atgagatgac agggccaatt acattcaaac 4980

tgctataaca ctftaaagta ttttattttt attattgtaa attatgtatg tagctgggtg 5040
 gtggcagccg aggtgcacgc ctftaatccc agcacttggg aggcagagge agatggatct 5100
 ctgtgagttc aagaccagcc tggctataa gagctagtgt caaggaagga tatacaaaga 5160
 acagttctag gatagccttc aaagccacag agaagtgtg tcttgaaaac caaaaattgt 5220
 gctgggacct gtctctgctt tggttgcttc ccactcccc agagctggac tcttggtaa 5280
 cactgaatca gctgcaaaat aaactcctgg attcctctct tgtaacagga gcccgaagtc 5340
 aggcgcccac ttgtcttc gcaggattgc catagacttt ttctgtgtgc ccaccattcc 5400
 agactgaagt agagatggca gtggcagaga ctgggaagge tgcaacgaaa acaggaagtt 5460
 attgcacct gggaatagtc tggaaatgaa gcttcaaac ttgcttcatt ttcagttgta 5520
 cacagactca ctcccagggt gactcacag tgtaaatatt cctgactatg tctgactgc 5580
 ttttatctga tgcctcttc ccaaatgcc aagtgtacaa ggtgagggaa tcaccttgg 5640
 attcagagcc cagggtctc ctcttaacc tggactgtc tttctccggc agcctctgac 5700
 accctcccc ccaftttctc taccagaag tctgagcaga gttggggcac gctcatgctc 5760
 tgataactc cttgtcttc tgaagatcta acttctgacc cagaaagatg gctaaggtgg 5820
 tgaagtgttt gacatgaaga cttggtctta agaactggag caggggaaaa aagtcggatg 5880
 tggcagcatg taccgaaat cccagaactg gggaggtaga gacggatgag tgcccggggc 5940
 tagctggctg ctacccagc ctagtgaat tgccaaatc caactctat tgaaaaacct 6000
 ttaccaaaa aacaacaaa caaataataa caacaacaac aacaacaaac taccatac 6060
 aaggtgggct gctctggct cttgaggaat gactcacca aaccacaaag ttgccacagc 6120
 tgttctctgg cctaaatggg gtgggggtgg ggcagagaca gagacagaga gagacatgac 6180
 ttctgggct gggctgtgtg ctctaggeca ccaggaactt tctgtctt ctctctgt 6240
 ggcacagcca gaccaccagc acccagcagg tgcacacacc tccctccgtg cttctgagc 6300
 aaacacaggt gccttggct gtctattgaa cggagtaag ttctgcaga tgtatgcatg 6360
 gaaacaacat tctctgggt ttattctac tttgtgata aaaaccgggg aactccagga 6420
 agcagctgag gcagaggcaa atgcaaggaa tctgctctc tagcttctc cccatggctt 6480
 gccgggctg cttctgcaa gccctctct cccattggc atgctgaca tgaacagcgt 6540
 ttgaaatgt ccaaatgtc acttcaaaag aagcttctc tgatcttct aactaaatca 6600
 gaccatgttt caccgtgcat tatctctct ctgtctgt gtctgtct ctgtctatct 6660
 gtctatcacc tacaatcat ctatctatct atctctatt tatctaccta tcaatcaate 6720
 atctatctc taactagtta tcaftattt atttgttac ttacttttt tatttgagac 6780
 agtattctc tgagtacag cttggctgt cctggaacc attctgtaac caggctgtcc 6840
 tcaaacacac agagatcaa ctgctctgc ctctctggtg ctggggttaa agactgcaac 6900
 caccacgccc ccctctatc atctattat gtaactatta ttcagctatt atctatctc 6960
 taactatca tcatctgt atccatcacc tatctatcta tctatctatc tatctatcta 7020
 tctatcacc atctataatc aattg 7045

<211> 6473

<212> ДНК

<213> *Cricetulus griseus*

<400> 3

(SEQ ID NO: 7)

agtgaacaca gatcaatctt ctctaagctg cttgagcctg tgttttcccg ttatacacag 60
 gtaattggtg tgctgttaaa agctacttag aataaatgaa gaagaaaggg agaaggaggc 120
 agaggagaag gagcaaagaa agaaggaaag ggggagggag ggagaggagg gagggaggga 180
 gggagggagg gagaggagg gagggggaga aagaagagaa ggagagatct tttcccact 240
 gactatctca gaaatfacc acaggtgaa ggggtacta attaaggaat agctgtaagt 300
 aacattatft ttattcgtag aacctataa ctctaagat gtgctttta ccttttctg 360
 ctttttagca caaataaact ccaacatgaa aattatccac tgtgcgtgtg aaaataccta 420
 cacagagtc tgaatcatft gccaaatca agcccattt tttattcca ttttgactga 480
 gagcaagatg ttcttttag gggatggaag cgtctgggtt tcccactg aatgactca 540
 ctgaatgtt gcctcattaa cattctgat ttttcgtaa tctctgctc atgcattca 600
 gataactgtg cctatcacia atggctttt agcagctcca ctcttctgg ggatgtggtg 660
 gccctccag tagctgccac cacggattgt cttcaattc tcaactgitt ttgagttgag 720
 tgtcagcctg accctgggc atggccgac atgactcagg caaagtgaga gtttcatcac 780
 taaacgtggt tctgtttgt atgtctgtt tcccttaag agcaggttat tcaatacca 840
 tctggctgag gtcaagttgc ctgagagccc acagaatctc taccaggtc cctgttgat 900
 cctaaaaac tcagtcatgc tgtaatctc ttctgaaact gtgcaatgcc tgcaggtctg 960
 cagcccagct ctctctctt gcttctgct ctctaggac cccatgctc ctcaaactg 1020
 cacgtgttc ttgctctcc accacggtg ccaagccaaa attcgggtgg gcgggaggac 1080
 atttcccaa gtgctgtt ccttcttt cttttgaca cccagataa atcatcttc 1140
 ccaatccaac acagcccac tgtgtttg gggacttcat gatcaccac agaatgtat 1200
 ccttagaac aaaaatgca aaccagaac accaggagac aattaaagaa atttcactg 1260
 gtgaggtcac aagtagtaga gactctgt taacgggag aaactttcac ggaccagca 1320
 tgctactgtg gcagttctg aacaagctga aaatgcctt cccgaccacc caagccagt 1380
 ccacacaaag gccacctag ggtgtgcaca gtagtcaact aggcgttggc ggaactcagg 1440
 aaggagtctg aatttctcc cgtttctcc ttctctctc atccctate ttacttctg 1500
 tctctctc ctctctctg tccccctt tctctctc ctctctgtg caggccaca 1560
 gatggaccgg gagacctca gcatgcaaa tcaactaact gctctaccac tcaaccacac 1620
 cctgcctgc atgttacta ctactattat tatcttgata caggtctcca cattgagctc 1680
 acctcagag tctccactt gagctcacc tcacagtct cacattgagc taccctcac 1740
 agtctccaca ttgagctcac cctcagctc tccacattga gtcaccctc acagtctca 1800
 cattgagctc acctgtggtc tctggcaaac cttgaattct ctactctc ctgctcagc 1860
 ctctggggtc gtggggatta gccaaacca cttgaggtt tcttcaata gcaattctt 1920
 agcgttcaat taacacacac tcaactctc agtactttg aaaccggaac aggagaattt 1980
 ctgtgagctg gaggttagct tggactacag tatgagacc tctctctaaa taaatacaca 2040
 aagaaatctc accaagggcc tccctctctc agcaagctct aactgtggtg ggagttctgg 2100
 gttgttccag ttaacgggtc cagaactctc ctgccagca catcagcccc tagacacagg 2160
 tgctctctc catgtgaaca tgcagtcaca gaaatgaaat aaagtgaaa ttttattct 2220
 tcagttgat agcctctcc gtgtgggtg tagttactgt cttgaatagg ataggctcag 2280
 aatccttggg gctggaacca agatftgat tccattagac gacagggaa ataatgcca 2340

atagggcatt cctctccccg gtcactagcg gtgcacttcc tccgaatctt tgcatgttg 2400
 aattagaaaa gttagtattt tectccatcc ctccccctcc tccccctc cctctctccc 2460
 ctctccccct cctccccctg tctccccgcc cctccccctc cctctctgac cccccccatc 2520
 tatcaaatcc aagaattcca gtaaaaagag gaaaacaatc gaagtgattt cgttgattgt 2580
 cagttccacc aaagcaagac ttgactttag ttccgcgttt cggttccccg catgcaccac 2640
 agccagcgag caccgtggaa ggatgctagc acggctctcc ccccccccc actagctgtc 2700
 ttcagctccc cagtagaggg caaccgcaat ccagattctc aatggagagt gtttacacaa 2760
 tctgtgcggg tttgtgtgag cgcgccccgt tccagagaca ctctctctt ttctttttc 2820
 catttcatcc cagtggcaac gcagagtgcc agatcattca ggccgtttgc agggcaagcc 2880
 gtgggagctt ggcaagcaag gccccatttc ctagggaacc cgtgcctggc gcttcaggaa 2940
 agcacgggaa cctggcactg tgactctcgc ggtattattt tgcagaactc tttattaaac 3000
 gggagtttca agtccagctg gagacgacca ggcagcgcct ttaaccccag agtcacacac 3060
 aggtgccttt tcttggggcc agattggggg tgtgtggcag acctgcgacc agcttgacaa 3120
 ctctctgcc aggccacaaa atggtgttgg ctgtaagagg tgacaccagg gacagggaa 3180
 atcgtgcta ttctctgag ctctccaaag acccacacca gtctgtcccc ctctctctc 3240
 gctcttcccc tgtatgccc cctcaccatc tcccccaacg agactcttgg catctctctg 3300
 gcacaaggat ttgaaaatag atgcttgggg gtgagaagaa gaagagagaa agagagagaa 3360
 ggaaggaagg atatatagat gatacagacg catacaggtg acatgtagct aatcattttt 3420
 aattaaaaaa taaattaa gcaaatcaag gatatatatg atacccttag agcaagtgtc 3480
 tcatacacac acaaacacac acacacaata tatatatata tatatatata tatatatata 3540
 tatatatata ttacttgg aacaagtgtc cagaagggtt ggggactcta aagtgttgt 3600
 caaagccagg ctacatcag taatcttacc acctggtaga ctgagacagg aggttttga 3660
 tgagttcagg cccagcctga gctgcagaat gtgattctat cccaaaaaag taaaataaaa 3720
 taaaattcaa aatacagaa aagagtattt gctgaacaaa caagcctaaa gcctggatc 3780
 cttccccca tctctaaaga aaataagttt ctggaagctg gagggatggc tcagaggtta 3840
 agagccccag ctgcacttgc ggaactacta gaccagttc ccagaccca cactgtgggt 3900
 cacaactgtc tcaaagcca gctccggagg atccatgccc tctctggcc tccaccgca 3960
 ccaagaacac atacagtgcc catacattta tgcaagcaag gtattcacgc acataaaact 4020
 aaaagaatat ttaataaaga tataacaaaa tagcatgaag cccagctggt acagagttc 4080
 aaactacatc ccagttcat cctctgcct ttgctctcag ttgcttggg taggtctctt 4140
 ctctgaactg gcgccccg ggttccatc tgagaccctc tcaattttaa acctactct 4200
 tctggcgggg gtttaattgt gccagggctc aagccaacgc ttctctctc ccacagcaat 4260
 ctccaagtt tcacagata accaggaact gctaagtcca tgtgaacctt agtgaagaac 4320
 ctgagtcttc ccattgatt ggtgtgtgca tgtgtgcata cacaaatgta tgtgtgtgt 4380
 ctatgtgtgc ctatgtatgt gtgcattcat gtgtcatat acaaatgcat atatgtctat 4440
 gtatgtgtgc tacacaaatg tatgtgtgtg ctcaatgtgt gcctatgtgt gtgtatgcat 4500
 gtgtgcgtac acaatgcatg tgtgtgtgt ctgtgtgct gtgtgtgtat gcatgtatgc 4560
 atacacaaat gtatatgtgt ggtgtgtgaa tgtgtgcta tgtatgtgt tgctgtgtgt 4620
 ggggtgtgta tgtgtgtgat gtgtggagg gtgtgtatgt gtggtatgta taggtgatac 4680
 gtttgggtg taatatgct atgtgtttg tgaatgtag ttcgtgtgtg tgcatgtgtg 4740

cgtgcgtgcg tgcgtgctg cgtgtgtgtg tgtgtgtgtg tgtgtgtgtt ggatatagta 4800
 tgtgtgaggt gtgtgtactc accatggcct cctcacttg ggggagtgaa gtcagcagcc 4860
 tggaccactc agggacatga gatactcaga cacatcttga ttccacccc tctttctctg 4920
 atcctccttc acgtgtcact ttcccaaaaca ctggacaaca gtttgggggc atctgattec 4980
 actaatgaca gggacatcac atgtctccag agggaacacc ttctgtgtca catgtcatct 5040
 gagaatgtag cagagtcaca gagaaatgtc acagaaacca aatgcagag taccaaggtg 5100
 tagctagcca cagagcagag gggaagccgc tgaatttatt aaaaatgtca gaatcgtaaa 5160
 agacagggga cagcgggtggg gacattcagg gtccagtgc acacaggcag tccaaacctg 5220
 atcactggaa ggtagtaggt aaggaaaggc tgcacacaga ttattcacac agtttataca 5280
 tgtacacaga ttattccat ggtttgtgta tgtgcacaga ttattcacac agtttataca 5340
 tgtgtggctt cgtggtaact ttgagcttac ttcaattta aaaggatctc tctcacaagc 5400
 tggggccggg aatggctgca gtcaacactc catcacttag tcactctgt caaacagcac 5460
 ctctgactc atggtgactt gtagtaaaat gaagaggcca catttgcate caagacagct 5520
 cateagtacc tagtgaagaa tctgtccctg agtatttga tgaatggacc cgggtccagg 5580
 gcttggtctg gactctccag gtgttcagc cagaatgtca ttgtgtttt tcaggatccc 5640
 agaagttct aaaatacagg ccaagtactc atttgtgta caaagtatct gactaataga 5700
 agtgattag taacacaaag ccttttaaaa accgagatca ccttctcat gtcctggcc 5760
 tcttagaaca agatccaagc ttttctggt tgacaagtgg ggccatccag tgcgtctccg 5820
 ttctgctac ttcactgga agactctcc cactaactg ccctgacct ctcacacctg 5880
 ctgttctt tccaccgga agtgettgc taggtttca tggccatctg actgagcate 5940
 taggcctcag tccagtgtc cctcagctct ctctagtcac tgtactaatg gaaacggcca 6000
 ctaactacat ttcaatatg gaagcctct cctcaggaac ctccaagggc agaagcctcc 6060
 agagaaccac tctgacccc ctggagtct gactgcttct ggccctctct gtgtctgcag 6120
 gactattcac cactgtgtt gaatgttca gtctcacct cctctggcat gtgtcagtt 6180
 ctcatctcat tggggagtec ttccaggtc actctctct cctgtctttg aagtgtttt 6240
 ttcttcatg gtattctgt ctgggcacac acacagacac acatacacac acatacacac 6300
 ccatgcagta tggcagatac atacctatg ttccagattt ttattctacc atcacciaat 6360
 acctgaatec ccgaaaaagc cttagaaagc caggaaattg tgfattttg tcagcactcc 6420
 acccagcac ctgaagccaa gctgactta atattttgg tttgtttct aga 6473

<211> 7045

<212> ДНК

<213> *Cricetulus griseus*

<400> 4

(SEQ ID NO: 8)

caattgatta tagatggatg atagatagat agatagatag atagatagat agatagatga 60
 tggatagaca gatgatgat agttagagga tagataatga ctgaataata agtacataaa 120
 tagatgatag agcggggcgt tgggtgtgca cgtctttaa cccagcacca gagaggcaga 180
 ggcagtggga tctctgtgag tttgaggaca gctgtgttac agaatgggtt ccaggacagc 240
 caaggctgtc actcagagaa atactgtctc aaataaaaaa agtaagtaaa caataaata 300

aatgataact agttagaaga tagatgattg aatgataggt agataaatag aagatagata 360
gatagatgat tgatagatga tagacagata gacagacaga cagacagaca gacagcagaa 420
agataatgca cgggtgaaaca tggctctgatt tagttagcaa gatcagagaa gccttctttg 480
aaagtgcacat ttgagagcat ttcaaacgct gttcatgtca ggcatgccaa tggggagaga 540
agggcttgca gaaagcagge cgggcaagcc atggggagca agctaggagg cagcattcct 600
tgcatttgcc tctgcctcag ctgcttctg gagttccccg gttttatca caacagtaga 660
aataaaacca ggacaatgft gtttccatgc atacatctgc aagaacttac tccggttcaa 720
tagacagacc aaggcacctg tgtttgctca agaagcacgg agggaggtgt gtgcacctgc 780
tgggtgctgg tctctggct gtgccagaca gagagcaaga caggaaagt cctggtggcc 840
tagagcacac agcccagccc aggaagtcac gtctctctct gtctctgtct ctgccccacc 900
cccaccccat ttaggccaga gaacagctgt ggcaagcttt gggtttgggt gagtcattcc 960
tcaagagcca agagccgccc acctgtatg gggtagttg ttgttgtgt tgtttatt 1020
atttgtttg ttgtttgtt ggtaaaggtt ttcaatagg agttggaatt tggcaattca 1080
gctaggctgg ctgagcagcc agctagcccc gggcactcat cctctctac ctccccagt 1140
ctgggatttc ggttacatgc tgccacatcc gactttttc cctgctcca gttcttaaga 1200
ccaagtctc atgtcaaaca cttcaccacc ttaccatct ttctgggtca gaagttagat 1260
cttcaggaag acaaggagtg taccagaca tgagcgtgcc ccaactctgc tcagacctc 1320
tgatagagaa aatgggggga ggggtgtcag aggctgccgg agaaagacaa gtccaggta 1380
aggaggacga cctgggctc tgaatccaag ggtgattccc tcaccttga cactggcat 1440
tttgggaagg aagcatcaga taaaagcagt gcagacatag tcaggaatat ttacacgtgt 1500
gagtaacct gggagtgagt ctgtgtacaa ctgaacatga agcaagtttt gaagcttcat 1560
ttccagacta ttcccagggt gcaataact cctgttttcg ttgcagcctt cccagtctct 1620
gccactgcca tctctacttc agtctggaat ggtgggcaca cagaaaaagt ctatggcaat 1680
cctgcgagaa gacaagtggg cgcctgactt cgggctctg ttacaagaga ggaatccagg 1740
agttttttt gcagctgatt cagtgttgac caagagtcca gctctggggg agtgggaagc 1800
aaccaaagca gagacagtc ccagcacaat ttttggttt caagacagca cttctctgtg 1860
gctttgaagg ctatctaga actgttttt gtatatcctt ccttgaact agctcttata 1920
gaccaggctg gtcttgaact cacagagatc catctgcctc tgctcccaa gtctgggat 1980
taaaggcgtg cacctggct gccaccacc agctacatac ataatttaca ataataaaaa 2040
taaaatactt taaagtgtta tagcagttt aatgtaattg gccctgtcat ctcattagga 2100
gtggcactat taggaggtat ggtttgttg aaggaaatat gtcactgtga ggtgggctg 2160
tgaggtttc tatgtcagg gtaccageca gtgtctcagc tgaggctctg ttgctgcaa 2220
gatgtaggac tctatcctt ttctccagea ccatgtctgt ctgcatgcca tcatgtccc 2280
agccatgatg acaatgtact aaaacctctg aaactgccac ccaactaaat gtttcttt 2340
ataagagttg ccatgtctat ggtgtctct cacagcaata gaaacctaa ctaagataag 2400
tgtattctcc cctactcccc atgattttaa atttaggaag gcaggtaggc aggcaggcag 2460
gctggtatag tggttcattc tagcacctga gacctggaat gggaggattg tgagttagt 2520
ctaggccatt ctggtgccta gaaaccagag cggggggtg gcccaatgca gagcactgc 2580
tctacgtatg gccagcaca ataagtcaat ttctcact taaaggctg acaattttaa 2640
aacactggtt tttagttagt cctgtctgc tccacagatg gagacagcta atcacagatg 2700

catcaggggc ctctctgagt gctaaacatc aaacagcctt cteccctct gagcctttgt 2760
 gtgcagaatg tgcctatcgc aagaagcaaa cagtcttget tgcccaccaa ctctctctct 2820
 gcatcagaag agctgggtgc aaactgcaag agtagcctca ccttagagat gggcccatt 2880
 gctctacatg ggagcattac ctccaagaa ggcaaaaatg tctctgggt gagcttttt 2940
 tgtcacctgt taaaggcaaa tcaacagaga ggctttgtct cacccactaa catcttgaa 3000
 acaaatacca acgaacgctg gggaggatgt ggggaaagea gagccctcat gctctccgag 3060
 ggaaaatcac acccactgtg gaacagtgtg gaaacctcaa agactgggat tacaagcagc 3120
 acacaagcca gccacgctac tctggctac acaccacaaa gacgcttga cttcacgct 3180
 tacgctgca acactagcaa cgtcccact gctctcttg agccccccc cccgccctg 3240
 cccccccc cgccccgtg gtctatgttc ctctcccta aagtcagctt ccattctct 3300
 gtctccatct tcgcccacc ctctctctc gctacataat tgtctctatt ccattctct 3360
 gctttgaaac agctttttgc aaagcatcaa atctattgtc ctatgcccc aatcaacctc 3420
 cagtttaca agtgatacag gaaatcgttt tctaattaa aaatecccc ttgaccatt 3480
 tatteccact ctggaacat ctcccctg aggaaagtta cagaatgagg tggctctct 3540
 ctctctatc gagtggttc ctccagctt tgcctgtc taatctttt aactgttgc 3600
 caggcctcca ccacggcaca gatgaactgt ggggttcatt tacctgaaac tctatggaag 3660
 gatgtttatt tctcttca tttageaaat gataaagggc accattcact ctgtctatc 3720
 tgcaggggcc attcctttct ctaggccaga tactgagaat tgcctccaga atcaatgtg 3780
 tatacatatt tcccctcaa cattgatagg cattgateac acacacacac acacacacac 3840
 acacacacac acacagtage acaaatgtat tcccctagcc cgttccate ttgccacag 3900
 actccagagt ggccctggat agcaagcttc ctgtttgtt tctctgttc tctgtttt 3960
 ccacctcca gtctatctt tctaagctct tctgcaatg tctcttccc aactgtctg 4020
 agatgcagtc attgtctggg attcagacct tctctctctg cccaagtgag tatattgacc 4080
 cccacggttt gtacaacct aactcaggg agcccacaaa aaactgttt atgagccaag 4140
 tagtcccagg acttgagagg tagaggcggg aagatcagca gtttgaggcc agcctggaga 4200
 gcataagagc cggctcaaaa acaacaatgg aaactagata ctaagtaaaa atctgggg 4260
 gtttcatcat gaatgtctgt tctctagta ccacgtgaa ctccgtacac agctccagct 4320
 gttacggctt tctagaatc catactctt tttttttt tttttttt tttttttg 4380
 ttttfcgaga cagggtttct ctgtggctt ggaggctgc ctggaactag ctcttataga 4440
 ccaggtggt ctggaactca cagatcca cctgctctg cctccagagt gctgggatta 4500
 aaggctgctg ccaccaacac ccggcagaat ccatactctt tttaaaaaa gatttatcaa 4560
 tttactatgt atacagctt ctgctgcat gtatccatgc atgcagaag atggcaccag 4620
 gtcgcattac agatggtgt gagccacct gtggtgctg ggaattgaac tcagaatgc 4680
 tagaagagca accagttct ttaacctctg agccatctct ccggccccca gaaatccata 4740
 ttttgagga tttttacac cccccacc aaaagacgta tatctaaatt ttaatgtgag 4800
 aattcacatt tcttaagag ttgaacatag atttagagga aatcagatc ccacatgatt 4860
 acaaaagcat gctgtgggc aggtctgcta ccaagagtg ggccgtagct tctagctcag 4920
 acaaaactac tcccctctc gtggctctt cgcctcaag tcagaaactc acctgtgat 4980
 tctgccccag aagtgtctat agagcacagt gcatcttcc gtcttactc tgtggttga 5040
 attgttcca tctctatga ttacaacccc tcacagagca tctaactgg ttttttga 5100

tgctatggg cactectcca ttctagaaca ccttgccat caatactatg aaaggagggg 5160
 tggaggagga agagcaggaa gaggaggggg aagcgaggga agaggaagac acggtggca 5220
 atgaggaggg gggagcacc aagtcctccc tggatgagag tctactggg agacttaata 5280
 ttaattataa atgcttggc agcagctggg caggataagg ttaggcagga gaaccagact 5340
 aaggactctg ggaagcagaa gggcagagtc agacaaggag aggaaacagg aagtacaagg 5400
 taaagtcacg tggcagaatg tagataatg aatgggttc atttaagttg gaagagttg 5460
 ctagtaacaa gctgagcta tcagccgagc atttataatt aatattgagc ctccatattg 5520
 gttatctggg aattggcggg cagaaaaaaa aaagtctgcc tacaagtcaa tgcatgtag 5580
 ctccaaagc caaggtacct ttgttcagtg cttgactgag ccagcattat aaattttctc 5640
 cagatgtacc gaatcacatt tcatagcaac atgcagacat caagttttcc ctgaagctct 5700
 aaccagctgg ttgcatgctg tccggagtct cagctataac ccagaagtga cctgggtcgg 5760
 ggaagaggtg gtaacttggc ttctttgcac tctctgtgtt gectcacca ttcagettca 5820
 agcaatgtga ctgctgacc ctgagggcgt ttacaacgcc tgaccacag accacaagtc 5880
 aaccagctgg tgtgctcag atacctagtc tgaacctag cctgctccc accctgctc 5940
 catctccacc cttttctac tgctcatcac agctggctag caaagactgc ctgagactg 6000
 agcacaggct cactccaca gccgtgactg ttcgagccac ttaaatcaaa gagcgttgt 6060
 ctccgctca gtaaactct cctcagctca ctgatgactg tgactttctc tagacagcac 6120
 atttgggtt aagacactgc tacttgagct cttcattcag ttctcagaa tacctcattt 6180
 gggtcagatt cccaaagagg aagatagggt tctggcaga cagacatgc tcattccttt 6240
 gaaatcttc agagaaatgc agtgactatg gcacctctt aaaaagcaca cacacaaata 6300
 acacacacac acacacacac acacacacac acacacacac atateccct cactgctc 6360
 cttgatatgt atatgatata tataaaatca ttgtttata ctgtgataat tgattatgaa 6420
 taaatftac taaatgaac aattaaaatt atgggggggg ctggagagat ggctcatcag 6480
 ttaagagaac agttgctgct cttgcagaac acgagagttc agttcccagc accacatca 6540
 ggcagctcat aacctgtgt ggtgctagtt ccaggagatc tgggtccctc ttctggctc 6600
 ctccagcacc tgctacatgt ggtcacaca cacacacaca cacacacaca cacacacaca 6660
 cacacacaca caaataaata taaagattat tttttcaaa actgagttaa aataggttc 6720
 tatctgattc atactaagc tttcacagt ggtaagtct attagatag tctagccata 6780
 tctttctcc cttttctt gaggagagc ttttaaagct acaagttaca gcctttttg 6840
 caaataagag taccatttaa caggcctctg accaatgaga tgccagaatc ggtgcccag 6900
 gagcttccca aacgtccat tataggaaa ggtggtacaa accagtagat taggcatgtt 6960
 ccaactccta agtgcctgc caaataagga aatggctca aatgtttgcc ttttatctc 7020
 accacctct gaattgcacg ctagt 7045

<211> 13515

<212> ДНК

<213> *Cricetulus griseus*

<400> 5

(SEQ ID NO:9)

tctagaaaca aaacaaaaa tattaagtca ggcttgctt caggtgctgg ggtggagtgc 60

tgacaaaaat acacaaatfc ctggctttct aaggcttttt cggggattca ggtattgggt 120
 gatggtagaa faaaaatctg aaacataggt gatgtatctg ccatactgca tgggtgtgta 180
 tgtgtgtgta tgtgtgtctg tgtgtgtgcc cagacagaaa taccatgaag gaaaaaaca 240
 ctcaaaagac aggagagaag agtgacctgg gaaggactcc ccaatgagat gagaactgag 300
 cacatgccag aggaggtgag gactgaacca ttcaacacaa gtggtgaata gtctgcaga 360
 cacagagagg gccagaagca ctcagaactc caggggggtca ggagtggttc tctggaggct 420
 tctgcccttg gaggttctg aggaggaggc ttccatattg aaaatgtagt tagtggccct 480
 ttccattagt acagtgacta gagagagctg aggaccact ggactgaggc ctatgatctc 540
 agtcagatgg ccatgaaagc ctagacaagc acttccgggt ggaaaggaaa cagcaggtgt 600
 gaggggtcag gggcaagtta gtgggagagg tcttcagat gaagtagcag gaacggagac 660
 gcactggatg gccccacttg tcaaccagca aaagcttggg tcttgttcta agaggccagg 720
 gacatgacaa ggggtatctc ggtttttaa aggctttgtg ttacctaac acttctatta 780
 gtcagatact ttgtaacaca aatgagtact tggcctgtat tttagaaact tctgggatcc 840
 tgaaaaaaca caatgacatt ctggctgcaa cacctggaga ctcccagcca ggcctggac 900
 ccgggtccat tcatgcaaat actcaggac agattcttca ctaggactg atgagctgc 960
 ttggatgcaa atgtggcctc ttcaatttac tacaagtcac catgagtcag gaggtgctgt 1020
 ttgcacagtg tgactaagt atggagtgtt gactgcagcc attcccggcc ccagcttctg 1080
 agagagatcc ttttaattg aaagtaagct caaagttacc acgaagccac acatgtataa 1140
 actgtgtgaa taatctgtgc acatacaca accatgtgaa taatctgtgt acatgtataa 1200
 actgtgtgaa taatctgtgt gcagccttc cttacctact acctccagt gatcaggttt 1260
 ggactgectg tgtgtactg gacctgaat gtccccacc ctgtcccctg tcttttacga 1320
 ttctgacatt ttaataaat tcagcggctt cccctctgct ctgtgcctag ctatacctg 1380
 gtactctgca ttttggtttc tgtgacattt ctctgtgact ctgtacatt ctcatgac 1440
 atgtgacaca gaaggtgttc cctctggaga catgtgatgt cctgtcatt agtggaatca 1500
 gatccccca aactgttgc cagtgtttgg gaaagtgaca cgtgaaggag gatcaggaaa 1560
 agaggggtgg aaatcaagat gtgtctgagt atctcatgtc cctgagtggc ccaggctgct 1620
 gacttcactc cccaagtga gggaggccat ggtgagtaca cacacctac acatactata 1680
 tccaacacac acacacacac acacacacac acgcaagcac gcacgcagc acgcacacat 1740
 gcacacacac gaactacatt tcacaaacca ctaacgata ttacaccca aacgtatcac 1800
 ctatacatac cacacataca caccctcca cacatcacac acataccaca cccacacaca 1860
 gcacacacat acataggeac acattcacac accacacata tacatttctg tatgcataca 1920
 tgcatacaca cacaggcaca cagacaccac acacatgcat tgtgtacgca cacatgata 1980
 cacacacata ggcacacatt gagcacacac atacatttct gtacgcacac tacatagaca 2040
 tatatgcatt tgtatagca cacatgcatg cacacataca taggcacaca tagagcacac 2100
 acatacattt gtgtatgca acatgcacac accaatcaca tgggaagact caggttctc 2160
 actaaggttc acatgaactt agcagttcct ggttatctcg tgaaacttg aagattgctg 2220
 tggagaagag gaagcgttgg cttgagccct ggcagcaatt aaccccgcc agaagaagta 2280
 ggtttaaaaa tgagagggtc tcaatgtgga acccgcaggc cgcagttca gagaagagac 2340
 ctaccaagc caactgagag caaaggcaga gggatgaacc tgggatgtag tttgaacctc 2400
 tgtaccagct ggccttcatg ctattttgtt atatctttat taaatattct tttagttta 2460

tgtgcgtgaa taccttgctt gcataaatgt atgggcactg tatgtgttct tgggtcccgt 2520
 ggaggccagg agagggcatg gatcctccgg agctggcgtt tgagacagtt gtagaccaca 2580
 gtgtggggtc tgggaactgg gtcttagtgt tccgcaagtg cagctggggc tcttaacctc 2640
 tgagccatcc ctccagcttc aagaaactta tttcttagg acatggggga agggatccag 2700
 ggctttaggc ttgtttgttc agcaaaact ctttctgtgt attttgaatt ttattttatt 2760
 ttactttttt gggatagaat cacattctgc agctcaggct gggcctgaac tcatcaaaat 2820
 cctctgtct cagtctacca ggtgataaga ttactgatgt gagcctggct ttgacaagca 2880
 ctttagagtc cccagccctt ctggacactt gtccaagta taatataat atatataat 2940
 atatataat atatataat atatattgtg tgtgtgtgtt tgtgtgtgta tgagacactt 3000
 gctctaaggg tatcatatat atccttgatt tgcctttaa ttattttta attaaaaatg 3060
 attagctaca tgcacctgt atgcgtctgt atcatctata taccctctct tcttctctc 3120
 tcttctctc ttcttctct cacccecaag catctatttt caaatccttg tgccgaggag 3180
 atgccaagag tctcgttggg ggagatgggt agggggcgtat acaggggaag agcaggagga 3240
 aagggggaca gactgggtgt ggtctttgga gagctcagga gaatagcagc gatcttccct 3300
 gtccttgggt tcacctctta cagccaacac cttttgtgg cctggcagaa gagttgtcaa 3360
 gctgtgtgca ggtctgccac acaaccccaa tctggccca agaaaaggca cctgtgtgtg 3420
 actctgggggt taaagggcgt gctgtgtct ctccagctgg actgaaact cccgtttaa 3480
 aaagattct gcaaaataat accgcagag tcacagtgc aggttccctg gcttctctga 3540
 agcggcaggc acgggttccc taggaaatgg ggcttgcctt gccaaagctc cacggcttc 3600
 cctgcaaacg gctgaatga tctggcactc tgcgttcca ctgggatgaa atggaaaaaa 3660
 gaaaaagaag aagtgtctct ggaagcgggc gcgtcacac aaaccgcaa cgattgtgta 3720
 aacactctcc atfgagaatc tggagtgcgg ttgccctcta ctggggagct gaagacagct 3780
 agtggggggcg gggggaggac cgtgctagca tcttccacg gtgtcgtctg gctgtgtgtc 3840
 atgccgggaa ccgaaacgcg gaactaaagt caagtcttc tttgtggaa ctgacaatca 3900
 acgaaatcac ttcgattgtt ttctctttt tactggaatt ctggatttg atagatgggg 3960
 gaggatcaga gggggagggg aggggcgggg agacggaggg aggaggggag gaggggagga 4020
 ggggaggagg ggaggagggg aagggatgga ggaaaactt aactttcta attcaacatg 4080
 acaaagatc ggagaaagt caccgctagt gaccgggagg aggaatgccc tattgggcat 4140
 tatattccct gtcgtctaat ggaatcaaac tttgtgttc agcaccaag attctgagcc 4200
 taccctatc aagacagtaa ctacagccca cacggaagag gctatacaac tgaagaaata 4260
 aaatttcaac tttattcat ttctgtact gcatgttca atgtagagag ccacctgtgt 4320
 ctaggggctg atgtgtggg cagtagagtt ctgagcccgt taactggaac aaccagaac 4380
 tccaccaca gttagagctt gctgagagag ggaggccctt ggtgagattt ctttgtgtat 4440
 ttatttagag acagggtctc atactgtagt ccaagctagc ctccagctca cagaaattct 4500
 cctgttccgg tttccaaagt actggagtta tgagtgtgtg ttaattgaac gctaagaatt 4560
 tgetgattga agaaaacctc aagtgggttt ggctaatccc cacgaccca gaggtgagg 4620
 caggaggaat gagagaatc aaggtttgcc agagccacag ggtgagctca atgtggagac 4680
 tgtgagggtg agctcaatgt ggagactgtg aggtgagct caatgtggag actgtgaggg 4740
 tgagctcaat tgggagactg tgagggtgag ctcaatgtgg agactgtgag ggtgagctca 4800
 atgtggagac ctgtatcaag ataataatag tagtagtaac aatgcaggcg aggtgtgtgt 4860

tgagtggtag agcagttagt tgatttgaca tgcttgaggt ctccccgtcc atctgtggcc 4920
 ctgcaacagg aagggaggga ggaagggggg gaacgagaga gaggaaagag agacagaagc 4980
 taagataggg aatgagagag gaaggaagaa acgggaagaa attcagactc ctctctgagt 5040
 tccgccaacg cctagtgaca tectgtgcac accetaaggt ggcctttgtg tggcactggc 5100
 ttgggtggtc gggaaaggca tttcagctt gttgcagaac tgccacagta gcatgctggg 5160
 tccgtgaaag tttctgcccg ttaacaagaa gtcttacta cttgtgacct caccagtga 5220
 aatttttta atgtctct ggtgttctgg gttttgcatt tttgttcta aggatacatt 5280
 cctgggtgat gtcataaggt ccccaaagac acagtggggc tgtgttggat tgggaaagat 5340
 gatttatctg ggggtgcaaa aggaaaagaa gggaaacagg cacttgggaa aatgtcctcc 5400
 cgccccccg aattttggct tggcaaccgt ggtggaggag caagaaacac gtggacgttt 5460
 gaggaggcat ggggtcctag gaggacagga agcagaagga gagagctggg ctgacagcct 5520
 gcaggcattg cacagttca gaaggagatt acagcatgac tgagtttta gggatccaac 5580
 agggacctgg gtagagattc tgtggctct gaggcaactt gacctagcc agatggtatt 5640
 tgaataacct gctcttagag ggaaaacaga catagcaaac agagccactg ttagtgatga 5700
 aactctact ttgctgagt catgtgcggc catgcccagg ggtcaggctg aactcaact 5760
 caaaaacaag tgagaaattg aagacaatcc gtggtggcag ctactggaag ggccaccaca 5820
 tccccagaaa gagtggagct gctaaaaagc cattttgat aggcacagtt atcttgaatg 5880
 catggagcag agattacgga aaaatcgaga atgtaatga ggcaacattc gagttgagtc 5940
 attcagtggt ggaaaccag acgttccat ccctaaaag gaacatctt ctctcagtc 6000
 aaatggaaat aaaattggg gcttgaattt ggcaaatgat tcagaactct gtgtaggtat 6060
 ttccacacgc acagtggata atttcatgt tggagttat ttgtgctaaa aggcagaaaa 6120
 gggtaaaaag cacatctaa gagttatgag gttctacgaa taaaataat gttacttaca 6180
 gctattcctt aattagtacc ccttccacc tgtgtaatt tctgagata gtcagtgggg 6240
 aaaagatctc tcttctctt cttctcccc ctcccctct cteectcct cctcctctc 6300
 ctcccctc tcccctctc ccccttctt tcttttttg ctctctctc tetgectct 6360
 tctcccctt tcttctttt attctaagta gcttttaaca gcacaccaat tacctgtgta 6420
 taacgggaaa acacaggctc aagcagctta gagaagattg atctgtgtc actagcgtgc 6480
 aattcagagg tgggtgaaga taaaaggcaa acatttgagg ccatttctt atttggcagc 6540
 gcaacttagga agtggaaat gcctaacta ctggtttgta ccactttcc ctataatgga 6600
 ctgtttggga agctcctggg caaccgattc tggcatctca ttggtcagag gcctgttaa 6660
 tggctactctt atttgcaaag aaggctgtaa cttgtagctt taaaagcctc tectcaagaa 6720
 agaagggaga aaggatagtg ctagacatat ctaatagact taaccactgt gaaaagcctt 6780
 agtatgaatc agatagaacc tatttttaac tcagttttga aaaaaataat ctttatattt 6840
 atttgtgtg gtgtgtgtg gtgtgtgtg gtgtgtgtg gtgtgtgtg gaaccacatg 6900
 tagcaggtgc tggaggagge cagaagaggg caccagatct cctggaactg acaccacaca 6960
 tggttatgag ctgcctgatg tgggtgctgg gaactgaact ctctgttct gcaagagcag 7020
 caactgttct cttactgat gagccatctc tccagcccc ccataattt taattgttca 7080
 ttttagtaaa ttttattcat aatcaattat cacagtataa acaatgatt ttatatatat 7140
 catatacata tcaaggatga cagtgagggg gataggtgtg tgtgtgtgtg tgtgtgtgtg 7200
 tgtgtgtgtg tgtgttattt gtgtgtgtgc ttttaagaa ggtgcatag tcaactgatt 7260

tctctgaagg atttcaaagg aatgagacat gtctgtctgc caggaaccct atcttctct 7320
ttgggaatct gacccaaatg aggtattctg aggaactgaa tgaagagctc aagtagcagt 7380
gtcttaaacc caaatgtgct gtctagagaa agtcaacgtc atcagtgagc tgaggagaga 7440
tttactgagc ggaagacaag cgctcttga ttaagtggc tgaacagtc acggctgtgg 7500
agtggagcct gtgctcaggt ctgaggcagt ctttgctagc cagctgtgat gagcagtgaa 7560
gaaaggggtg agatggaggc aggggtgggag cagggctatg gttcagacta ggtatcgtga 7620
gcacaccagc tggttgactt gtggtctgtg ggtcaggcgt tgtaaaccgc ctcagggtca 7680
ggcagtcaca ttgcttgaag ctgaatgggt gaggcaacac agagagtgca aagaaggcaa 7740
agtaccacct ctccccgac ccaggtcact tctgggttat agctgagact ccggacagca 7800
tgcaaccagc tggttagagc ttcagggaaa acttgatgct tgcattgtgc tatgaaatgt 7860
gattcggtag atctggagaa aatttataat gctggctcag tcaagcactg acaaaggta 7920
ccttggcttt gggagctaca tgacattgac ttgtaggcag acttttttt ttctgcccgc 7980
caattcccag ataaccaata tggaggctca atatttaata taaatgctcg gctgatagct 8040
caggttgtt actagetaac tctccaact taaatgaacc catttctatt atctacattc 8100
tgccactgca ctttacctg tacttctctg ttctctctct tctctgactc tgcctctctg 8160
cttcccagag tcttagtct ggttctctg cctaaccctt tctgcccag ctgctgacca 8220
agcatttata attaatatta agtctcccag tgagactctc atccaggag gacttgggtg 8280
ctccccctc ctcaatgcca tccgtctctt cctcttccct cgttcccc tctcttct 8340
gctcttctc ctccacct ctttcatag tattgatggc aagggtgttc tagaatggag 8400
gagtgccc atggcatgcaa agaaaccagt taggatgctc tgtgaggggt tgtaacata 8460
agcgtatgac acaattcaag ccacagagtg aagacggaag gatgactgt gctctagagc 8520
aacttctggg gcagaatcac aggggtgagtt tctgactga gggcgaagag gccacgagga 8580
agggagttag tttgtctgag ctagaagcta cggcccact cttggtagca gacctgcca 8640
caagcatgct ttgttaatca tgtgggatct gatttctc taaatctatg ttcaacttt 8700
aagaaaatgt gaattctcac attaaaattt agatatacgt ctttgggtg ggggggtgta 8760
aaaaatctc aagaataggg atttctgggg gccggagaga tggctcagag gtaagagaa 8820
ctggttctc ttctagacat tctgagtca attcccagca accacatggt ggctcacaac 8880
catctgtaat gcgacctggt gccatctt gacatgcatg gatacatgca ggcagaaagc 8940
tgtatacata gtaaattgat aaatctttt ttaaaaagag tatggattct gccgggtgtt 9000
ggtggcgcac gccttfaatc ccagcactct ggaggcagag gcaggtggat ctctgtgagt 9060
tcgagaccag cctggtctat aagagctagt tccaggacag cctccaaagc cacagagaaa 9120
cctgtctcg aaaaacaaa aaaaaaaaa aaaaaaaaa aaaaaaaga gtatggattc 9180
taagaaagcc gtaacagctg gagctgtgta cggagttcag cgtggtacta gaagaacaga 9240
cattcatgat gaaacacccc aggattttta cttagtatct agtttccatt gttgtttga 9300
gaccgctct tatgtctctc aggtggcct caaactgctg atcttcccgc ctctacctct 9360
caagtctgg gactacttgg ctataaaac agttttgtc gggtccctg aagttatggt 9420
tgtacaaacc gtgggggtca atatactcac ttgggcagag agagaaggct tgaatcccag 9480
acaatgactg catctcagga cagttgggaa gaggacaatg gcagaaggac ttagaaaaga 9540
tagactggag ggtggaaaag cagcaggaac agagaaacaa aacaggaagc ttgctatcca 9600
gggccactct ggagtctctg ggcaagatgg aagcgggcta ggggaataca tttgtctac 9660

tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgat caatgcctat caatgttgaa 9720
 ggggaaatat gtataccaca ttgattctgg gagcaattct cagtatctgg cctagagaaa 9780
 ggaatggccc ctgcagaata gacagagtga atgggtccct ttatcatttg ctaaagtga 9840
 ggagaaataa acatcctcc atagagtttc aggtaaatga accccacagt tcatctgtgc 9900
 cgtggtggag gcctggccaa cagttaaaaa gattagacac ggacaaagtc tgaaggaac 9960
 acctgaata ggaagaggag agccacctca ttctgtaact ttctcaagg ggaagatgtt 10020
 ccaagagtgg gaataaatgg tcaaaggggg gatttttaat taggaaaacg atttctgta 10080
 tcaattgtga aactggaggt tgattfgggg cataggacaa tagatttgat gctttgcaa 10140
 aagctgtttc aaagcagaga aatggaatag agacaattat gtagcgagga gggagggtgg 10200
 ggcaagatg gagacagaga agtggaagct gactttaggg aagaggaaca tagaccacag 10260
 gggcggggcg gggggcaggg gcggggggcg gggctcaaag gaggcagtgg gaacgttgc 10320
 agtgttcga gcgtaagcgt gaatgtgcaa gcgtctttgt ggtgtgtgac caggagtagc 10380
 gtggtggct tgtgtgtgc ttgtaatccc agtctttgag gttccacac tgtccacag 10440
 tgggtgtgat ttccctcgg agagcatgag ggctctgctt tccccacac ctecccagcg 10500
 ttctgtgta tttgttcca agatgttagt ggggtgagaca aagcctctct gttgattgc 10560
 cttaacagg tgacaaaaa agctcaacca ggagacattt ttgccttctt ggaaggtaat 10620
 gctcccatgt agagcaatgg gacctatc taaggtgagg ctactcttgc agttgcacc 10680
 cagctctct gatgcaggaa ggaagtggg gggcaagcaa gactgtttgc ttcttgcgat 10740
 ggacacattc tgacacaaa ggctcaggag gggagaaggc tgtttgatgt ttagcactca 10800
 ggaaggcccc tgatgcatct gtgattagct gtctccatct gtggagcaga cacggactaa 10860
 ctaaaaacca gtgttttaa atgtcaagc cttaagggtg aggaaattga ctattgtgc 10920
 tggccatac gtagagcaag tgctctgcat tggccaacc cccggctctg gtttctagc 10980
 accagaatgg cctagaacta actcacaatc ctcccattcc aggtctcagg tgctagaatg 11040
 aacctata ccagctgcc tgectgccta cctgccttcc taaattttaa atcatgggga 11100
 gttagggaga atacacttat cttagttagg gtttctattg ctgtgaagag acaccatgag 11160
 catggcaact ctataaagg aaaacattta gttgggtggc agtttcagag gttttagtag 11220
 atgtcatca tggctgggaa catgatggca tgcagacaga catggtgctg gagaaaggga 11280
 tgagagtctt acatcttga ggcaacagga cctcagctga gacactggct ggtaccctga 11340
 gcataggaaa cctcacagcc caccctcaca tgacatatt tectcaaca aagccatacc 11400
 tctaatagt gccactcct atgagatgac agggccaatt acattcaaac tgetataaca 11460
 cttaaaagta tttattttt attattgtaa attatgatg tagctgggtg gtggcagccg 11520
 aggtgcacgc cttaatccc agcacttggg aggcagagc agatggatct ctgtgagttc 11580
 aagaccagcc tggctataa gagctagttg caaggaaaga tatacaaaga acagttctag 11640
 gatagccttc aaagccacag agaagtgctg tcttgaanaa caaaaattgt gctgggacct 11700
 gtctctgctt tggttgcttc cactcccc agagctggac tcttggtaaa cactgaatca 11760
 gctgcaaaat aaactctggt attctctct tgtaacagga gcccgaaagtc aggcgcccac 11820
 ttgtctctc gcaggattgc catagacttt ttctgtgtgc ccaccattcc agactgaagt 11880
 agagatggca gtggcagaga ctgggaagc tgcaacgaaa acaggaagtt attgcacct 11940
 gggaaatgct tggaaatgaa gctcaaaaac ttgctcatg ttcagttgta cacagactca 12000
 ctcccaggtt gactcacag tgtaaatatt cctgactatg tctgactgc tttatctga 12060

tgettccttc ccaaaatgcc aagtgtacaa ggtgagggaa tcacccttgg attcagagcc 12120
 cagggtcgtc ctcttaacc tggacttgc ttctccggc agcctctgac accctcccc 12180
 ccattttctc tatcagaagg tctgagcaga gttggggcac gctcatgtcc tgatacactc 12240
 ctgtcttcc tgaagatcta acttctgacc cagaaagatg gctaaggtgg tgaagtgttt 12300
 gacatgaaga cttggtctta agaactggag caggggaaaa aagtcggatg tggcagcatg 12360
 taccgaaat ccagaactg gggaggtaga gacggatgag tgcccggggc tagctggctg 12420
 ctcagccagc ctactgtaat tgccaaatc caactctat tgaaaaacct ttaccaaaca 12480
 aacaaacaaa caaataataa caacaacaac aacaacaac taccacatac aaggtgggcg 12540
 gctcttgct cttgaggaat gactcacca aaccaaagc ttgccacagc tgttctctgg 12600
 cctaaatggg gtgggggtgg ggcagagaca gagacagaga gagacatgac ttctgggct 12660
 gggctgtgtg ctctaggcca ccaggaactt tctgtcttg ctctctgtct ggcacagcca 12720
 gagcaccagc acccagcagg tgcacacacc tccctccgtg cttcttgagc aaacacaggt 12780
 gccttggtct gtctattgaa ccggagtaag ttctgcaga tgtatgcatg gaaacaacat 12840
 tgtcttggtt ttattctac tgtgtgata aaaaccgggg aactccagga agcagctgag 12900
 gcagaggcaa atgcaaggaa tgetgectc tagcttctc cccatggctt gccgggctg 12960
 ctttctgaa gcccttctc cccattggc atgectgaca tgaacagcgt ttgaaatgct 13020
 ctcaaagtc acctcaag aaggcttctc tgacttctc aactaatca gaccatgttt 13080
 caccgtgcat tatcttctg ctgtctgtct gtctgtctgt ctgtctatct gtctatcctc 13140
 tatcaatcat ctatctatct atctctatt tatctaccta tcattcaate atctatctc 13200
 taactagtta tcatttattt atttgtttac ttactttttt tatttgagac agtatttctc 13260
 tgagtacag cttggctgt cctggaacce attctgtaac caggctgtcc tcaaactcac 13320
 agagatccaa ctgctctgc ctctctggg ctgggggttaa agactgtcac caccaacgcc 13380
 ccgctctate atctattat gtaactatta ttcagtcatt atctatctc taactatcca 13440
 tcactgtct atccatcctc tatctatcta tctatctc tatctatcta tctatcctc 13500
 atctataatc aattg 13515

<211> 14553

<212> ДНК

<213> Mus musculus

<400> 6

(SEQ ID NO:10)

cttgaagaac acatgtttc caagagggag caccatgtt ggaatgaaa ttagttagt 60
 gctctctcc ttaggttag tctctcttg ctataggtaa gtctctctc cctataggc 120
 agtctctc tctatagg tagtctctc ctctatagg ttagtctcc tctctacag 180
 gttagtctc ctctctcta gttagtctc gctctctat agtacctaga gagctaggc 240
 aaatgggcta ggcccgaagt gcagagacaa acagctatgg aagactgggt aagcactcc 300
 aagctacgaa agagcagtgt gaagggcag ggcttctgca gttagtagg gagatctcc 360
 agttgaagaa acagaagaac tgagagccac tgggtatcat cctctctgc catgcctcc 420
 tggatactgc catgctccca cttgatgat aatggaatga acctctgaac ctgtaagcca 480
 gcccacatga aatattgtt ttatgagagt tgccttggc atgctgtctg ttcacagcag 540

taaaacccta aataaggcag aagttggtac cagtattgct gtgatagacc tgaccatgct 600
 ttcctttgaa agaatgtgga tttggtgact ttggatttgc aacacagtgg aatgctttaa 660
 atggagatta atgggtcadc aattcctagt aggaatatgg aagactttgt tgctgggagt 720
 atttgaactg tttgacctg gcctaagaga tttcaaagga gaagaatttc agaatgtggc 780
 ataaagacag tttttgtggt attttgggtga agaatgtggc tactttttgc ccttgtctga 840
 aaagtctgcc tgagactaaa gtgaagagaa tcagattaat tgcattgaca agggaagttt 900
 gtggctgcgc tatctggaaa cttacagcca gcctcttggga cctcgggtga cttacgcaa 960
 tactcagga cagagatgct tgactctgta ctgatgagtt gtcttggatg caaatatggg 1020
 ctcttcattt gactacatgt cacgatgagt caggagctgc tctctccaga gtgtgacaaa 1080
 gcgaggggat gctgacggta gctgttctag ctttgaaggt aagcctgcac ttatgctaaa 1140
 gtcacacata cacgagccgg gtggagaacc tgtctgtgtg gagacacctt tcattacctg 1200
 tggcatccag cctctcaagc ttggactgcc tgtgtgctcc tggactctgg aggtcccact 1260
 gctctgtcct ctgetctta tgatactgac attttaaaag aatccagtgg tccccctg 1320
 tactcgggtg ctactctac ctggatgttc ctcaattatg ttctgtgaca cttctctgtg 1380
 actctgetgc attcctgggt gacatgtgga caccctgtcc ctttgcagac catgatgca 1440
 ctgtactag tggaaatcaga tgcccagaat gttgtcctgt gtttgggaac gtgacaggca 1500
 gtacagaagc agaagaggaa gggtgaaaac ggaaatgtca cagcagcadc tgatgtgtgc 1560
 ctcagtcaag catgtctgtg attggaacta ctcagcatga gagagggcca tggtaatac 1620
 acaacctat acacactgtg tccatttctc tctctctctt acacagagag agagggagga 1680
 gggggagggg gaggcggagg gggaggggga gggagaggga gtgggagagg gagagggaga 1740
 gggagaggga gagggagagg gagagggaga gggagagttt aatgtctgtg aagagatacc 1800
 atgaccaaag caactcttat aaaggacaac attttaattgg ggctggctta caggttcaga 1860
 aattcagtc attctacca tgggtgggaag catgcaggtg gatgtgtgtc tggaggaacc 1920
 aagagttcta tatectgac tgaaggcagc caggagaaga ctgcctcttc tgcacagggc 1980
 agagctttag catagaacat caaagecctt cccacactt cctccaacaa ggtcatacat 2040
 acttcaacaa agacacacct ctaacgggtg cactcctctg tggaccaacc atttaaacgc 2100
 atgagtctat gagggctaaa gctcttcaaa ccaccacact catgtacaca cacacacaca 2160
 cacacacaca ctctcataca cacacacaca cacactcaca cacacacaca cacacacaca 2220
 cacacacaca ccacacacac acacacacac agagttctat tttgactgt ttaactgtca 2280
 caaggttcta cttatctcag acacactgcc aggaattgtg tgggaagact ttaactttct 2340
 ttgggttcaac atggacttag cagttcttgg tgactctgaa agatttctgc agaaagaagc 2400
 caaagtgttg agcccaaggc ctggccacac attagtctct tctagatgaa caggggttta 2460
 aaaataaggg ggcacaaagg tgaagccagc aggggctgac ttagagagga gaccacca 2520
 agccaactgc tgaagtcaa aagcagatgaa tcccatac cagctgtgcc cggtgtgtc 2580
 ttgctacac ttagtaaat gttcttttag ttgtatgctg atgaatattt tgettgcata 2640
 tatttgtgta caccataggt gttcctaggg cctatggagg ccagaagagg gcatcagac 2700
 ctttggaaact ggaattatag acactgttta cccatagagt agattgtggg aatgagcct 2760
 ttagctctcg agagcggcca gtgctcttaa cctttggtcg tttctccagg tctttgagac 2820
 tttattttct tggacatcag gacaggatcc agggcttga gcttgtttct tcagccagct 2880
 ttttttcat gtatattaaa ttttatgta ttttcttctc ttttccca agacagaac 2940

acactctata tagctcagcg tgggtttgaa ttcagtttcc ctgtctcagt ctaccgggta 3000
 atatgattac agatgtgagt ctgactttgg tatcaaagtc cccagccctt ctggatatgt 3060
 gttttaagga tatcagatat atccttgatt tgccttgaat tttcttttta gttacaacat 3120
 aattagtcc gtgtcacctg aafatgtgta tgtcacctac atagtcttcc ttcttctctt 3180
 ctccctctc ccaccttccc aggtacctgt ctgtcttcat atccttgtgc tgagagtctt 3240
 gttgagggag atgatgaccg agacagagcc actgggggaa ggagatgggc tagtgcaggt 3300
 ctccagagag gagctcgtga atattgtagc ccttttagtc cctggcatgt cctcttgtat 3360
 agccaccgcc atgctgtggc ctggcagaag tgaataagtt gtccagctgt tgacaggcct 3420
 gccctccaga cccagtctga tcccaagaaa gggcatctgt gtctgtctct gaggccgtaa 3480
 gtgctgcctg gttgtctcca gcttgacttg acactcctc ctttaataaga gtaccacaga 3540
 acagggctg cagagtccct gggccaggtc cctgtgtctgt cctggaatgc caggcgtgaa 3600
 tttctgtga agtaggactt tgcctgcca gctcccacgg cttgcccttc agatagccag 3660
 aattatctgg taccctgcat tgccgttcaa tacgcagagt atcaactggaa gcgcgcgcgc 3720
 gcacacacac acacacacac acacacacac acacacacac acacgccac tccatcttta 3780
 aacccacc cccagcaacg gcggtgtaaa cactctccat caggaagctg aaacgcagtt 3840
 gccctctgct ggggagatga aggcagcttg ctgggggcca ggaccgtgct agcaacctc 3900
 cctggtgcac acgggctctg gtgatgacg ggaacggaaa cgcggaacta aagtcagtc 3960
 tgccttttt ttttttttt ttttttttt ttttttttt ttttttttt ggcgttggtg 4020
 gtggactgag tgacaatcag tgaatcact taggtgttt ttctcttctt cgttgggttt 4080
 gatagacggt gggagagggg cagaggagaa ggggagggat ggggagagag ggaggagga 4140
 ggggcggggag gcggggggcg aggaaaacgt gctaacttct ccaatctac aagacaaag 4200
 tttggagaaa gccgactga gtgaccagc agaaggaatc caggaatgtc cgctggaatc 4260
 tgactgttga ttccagccc atgcagagaa tctaggtctg taggaacatt ctttgccta 4320
 tccgacataa taactccaac caacacggaa aagaaaggct atacaagtga aaaaatggca 4380
 ttttcactt catgactata caatcactc caggtagtaa cacgtgtcta gcacagcgtt 4440
 tctcaacctg ggggtcacga tccccactt ttctgcatat cagacattt tacgttgta 4500
 ttcataacag tagcaaaatt gcagctatga agtaacaatg aatgcattt atggtgcgtg 4560
 tgtgtgtgtg tgggggggta tcacctaac atttactgta agaaggttga gaactactgct 4620
 ccagcagcta gtgtgttga cttaggttct gggatatata tttagcaatag ccaaccagaa 4680
 tccccacca ccacagcatt gagccccat gcagggcttg ctgggagagg cactgataag 4740
 acttctttat gtatttattt agagacgaat actcattagg taggccaagc tagcgtcaaa 4800
 ctcatggcaa ttctctctc cagtttctc aagtactgga ctcaggagtg tgttgccatc 4860
 atatacagta aggatttatt gactgaagaa aatctcaagt ggctttggtt aatccctact 4920
 acgccagagg ctgaggcagg aggcgcgcaa ggtcaaggct tcctgggct acatatagag 4980
 tgactcaat tttgacactt ggtgcggtgt tagtagtaat agtaaagatg aaggtgtgca 5040
 tcaggtgggg ccggtgattg gacacacttg gggctctctg gtccatctgc agctgtgcaa 5100
 caggaagagc ggagaatgag aggaaagaga gaaaagacag aatgagagag agggaggaag 5160
 agagaaaaag gaaaagagag aggaaaggaa aaaggaaaat gaggaaagcg agaaagaaga 5220
 aatgagaaaag aggaaaggga gaaagaaatg agagagagaa aagaaaagac agaatgcgag 5280
 agagggagga agagagaaaa aggaaaagag agaggaaagg aaaaaggaaa atgaggaaag 5340

cgagaaagaa gaaatgagaa agaggaaagg gagaaagaaa tgagagagag aaaagaaaag 5400
 acagaatgag agagagggag gaagagagaa aaaggaaaag agagaggaag ggaaaaagga 5460
 aatgaggaa agcgagaaag aagaaatgag aaagaggaaa gggagaaaaga aatgagagag 5520
 agaaaaagaa agacagaatg cgagagaggg aggaagagag aaaaaggaaa agagagagga 5580
 agggaaaaag gaaaatgagg aaagcgagaa agaagaaatg agaaagagga aaggagagaaa 5640
 gaaatgagag agagaaaaga aaagacagaa tgcgagagag ggaggaagag agaaaaagga 5700
 aaagagagag gaagggaaaa tggaaaatga ggaaagcgag aaagaagaaa tgagaaagag 5760
 gaaagggaga aagaaatgag cgagataaaa gacagaatft gagagagggga ggaagaaata 5820
 ggaaaagaga ggaaaggatg gagaaaagag agaaagaaag agagatgaaa gagagaaaag 5880
 agaaatgaaa tgagagagag agagagacac aaagagccag agagagaaga aaaaagggga 5940
 aagagaaaaga gaaagaggaa ggtctctctt ggacacatct tctttatct ttcctgggg 6000
 accgccaag cctggtggca tactgtacat tctgtacact gttcattca aacaggctct 6060
 gtcttaaaga tggctgagc ggtcagaaaa ggttattgtt aactgtttg caaaactgcc 6120
 tcaggagagt gctgagtgcg tgaagttgc tgcccgttaa ggagaagtct ctactactg 6180
 tgatctcacc atcgaaaatt tcttaattg tctctggtg ttctgggttt tgcagttttg 6240
 ttttaagga tacattctg ggtgatgtca caaagtcccc aaagacacgg tggagctgtg 6300
 ttgatgggg aaagacagtc tgctgaggat ttatctggaa ctgtcagaag gaaaagaag 6360
 taaatggggc acttgggaaa gtggcctcta gtttgacttc tggcttagca aagttgtgg 6420
 ggagataagg catacacagt agttagcagg aggcaacagg gtctgggag gacgcgaggc 6480
 agaaggagag gctgggctga cagcatgcaa tcattgcata gtctccaaag gagattgcaa 6540
 catggtgag tttcagagg tctacagag cccgtggtag agattctgtg gttctgaga 6600
 caactgact ttagccagat ggtatttgag taatctggga gagagaaaac agctacagca 6660
 aacaggcca catttagtga cgaactctc actttgactg ttgagtcatt tgcagtgggc 6720
 cctgaggtca ggctggccct cagctcaaaa acaagcgagg aactgaagca attactcaga 6780
 taatccacag ccacagccac tggaaagggc cacatccca gagacagcac agcaggggtg 6840
 ggggtggggc tatgagaaag ttagtgattg tagcagttat ctagaatgtg cggagcagag 6900
 gaggttacac aaaaacctag aatgtcattc aatgtgggaa accgagaggc tccaagecc 6960
 taaaaggaac agtttcttt cagccaaaat ggaaataaaa tttggggctt aaatctggca 7020
 aatgattcag acctctgtg tagtgtctt taaatgcaca gcagattgat tttcatgtt 7080
 gagtttatt gaactaaaag acagaaatgg tgaagagcac acctgaagaa attgagatgc 7140
 tatgaataaa atcattfact tacagctatc acttaattag tactctctc caccttctg 7200
 atttattggg ctagtcaagg aagaaaagat ctccctct cttctctc tctccccct 7260
 ctctctctc tccccctc tcttgact tctctctc ctttctctc ctccccct 7320
 tctctctc accctctc cctctctc ctctgtact ctccccct cctcccaatc 7380
 tcttttct cccccctt ctctctct cctctctc tctctctc tctctctc 7440
 ctctctc ctctctc tctctctc ctctctc tctctctc tctctctc tctctctc 7500
 ctctctct tctctctc acctctct ctttctct tctttatc aaagtagctt 7560
 tgaacagcac tactcggtt agttgtgtat aaaaggaaaa tgcaggtcca agcagcttg 7620
 ggaagattgc ttttctct ctggaggcag atgatgacag tcaagatca tcttttgc 7680
 tcatgtcac aggaaggggg acatgccgaa tctaccagt tgcagccacc tacacagat 7740

ccaccttcac ttctaaggaa atgtttggga agctacctac caaccacttc tggcatctca 7800
 tgggctagag gactcttaa tggcactctt atttgtttaa taaaggaggt tgtgacgtgt 7860
 agttttaaat ccttccaca caacaattgc tactctctga ccaaaaaaga agggagacag 7920
 gatacggcta ggtgtctagt agactttacc actttgaaaa gccttaatat aatcaggtta 7980
 gatacatctt tttacttat tctgttaaag acaaaaacaa aactttatft ttatttgtgt 8040
 gtatgcttgt gtgtgtgtgc ctgtgtgtat accacatgtc gctggtgccg gagaacacca 8100
 gaagagggga cctgatctcc tggagctaaa gctatccatg gttctgagct gcctgatgtg 8160
 ggtgctggga acagaactct ggtcttctgc aagagcaaca agcctcctct taactacgaa 8220
 tctctcccc atcccccaa atacatftaa ttattcattt tagcagcttt atttcgtaac 8280
 tacttatac agcataaaac aaggatttta tatatattac atgcaatcga ggataagagt 8340
 tgaggggaga tgcgtgtgct ccttctgggt gtctgtgctt ttgaagaatg taagcagtgc 8400
 acaagggacc gaggcgtgcc tgtctgccag gagctgtctt ctcccttgg actctgagct 8460
 gagtgcagtgc ctccgaagaa gtaaaagacg acctcatgaa gcaatgtctt caacccaaac 8520
 atgctgtcca gacaaagtc agcttcatta gtgctctgag gagagactta ctgagcctca 8580
 ggaaagcccc cctcagcatg gcgaaagtc actttgattg aagtactcgc aaagccatgg 8640
 cagtgcggcg gcggccgcgt ggagcttctg ctcgagtcgg aagcggcacc tttgtcagcg 8700
 ggctgtgatt agcacgggga ggcaggactg gagtgaagga agagtgggg gcggggctta 8760
 gcgctctggt ctctaagct gtagtcagcg cctcaagatt tgtaacctgc cttctgctt 8820
 cccagccagg cagtcaagtg gctccaagct gaagactgca aagtgccctt aaccttttg 8880
 ttatagcgag gctgaagaca cctgtctctt tcatgaaagc cggatgtctg aaatccgatt 8940
 tgataaatat ggataaaacg tataacgctc gatcaatcga atcgaaggag ctacagattg 9000
 gcaccacggc tttggggaca acagagtact gactcgttgg gaggacttgg atactcccc 9060
 tctcttcca tctctcccc ttctctact tctctctct tctcttcca tttctctct 9120
 ctctactgtt tctactatt ttacaaaag atttattta tttatttatt tatttattta 9180
 tttatttatt tatttattta tttattta gtagcagat aactgtagc tgtctcaga 9240
 cacaccagaa gagggcgtca agttccatta gagatggtt cgagccacca tgtggttct 9300
 ggggcctctg gaaggaccgc cagtgtctt aacctctgag ccatttctcc agtaccctc 9360
 teaccgttcc tctcaatct tctctctt cctctccac ttctctgct ttcttggtt 9420
 cattatctt ctctcttct tctctctc cctctctc tctctactg tagtttct 9480
 tctctctct ttctctgct cctctctct cctctctat tctctctct ctctctct 9540
 tctctctct ctctctct tctctctc cctctctc tctctctct cctctctc 9600
 ctctctct ctctctct ttctctcc ctctctcag agtatcaatg gcaaggtgt 9660
 tctagaatgg aggagtctc ctaggcaact aacgaaagcc agttaggatg ctctgagacg 9720
 ggtacaatc agggagggcc gtggggatgg aaggttctg ctgcgattca ttctggagca 9780
 accccagcg agaatcatga ggttggctcc ggtctcag gcacaatc agaagaggaa 9840
 ggttctagga aggacagtt tctctgagat aggagttaca tctgatgtct tggcagcaga 9900
 gccactgtac aagcgtgctt tattaaccac gtgggattaa atctctttt aaatttatt 9960
 tcaactctta aggaaacgtg aacttcaca tcaaatfta gacttgcagc tcttatgggg 10020
 aaaaaaggg gatcttaaga atattaagca taggcgctg gagagatggc tcagcggtta 10080
 agagcactct ctgctctccc agaggtctg agttcaatc ctgcaacca cataatagt 10140

aacaacagtc tftaatgaat tetaatgcc tctctggtg tgtctgaaga cagttacagt 10200
gtactcatat aaataaata aagaaatfta aaaaaatgaa tattagggcat agattcctgg 10260
atcctaagaa agccatcaga gctggagcca tgtgtgggat cctgcttggg gctggagggg 10320
cagagtfcac gccccgggg ttttactta ttatcacatt ttcacgttg tttgaaaca 10380
gggtcttggg tggccaggc tggcctfgaa ctcactttc agcctctacc tcacaggtc 10440
tgggattact tggftctaa aagtatctc gteaagctcc ctggtgttat ggctgtgcca 10500
accaggaggg tctatacact cgctcaggta gagggagaag atccgaatct ctgacaggga 10560
ctgctgcctc tggggcaaaa tggagtgaag gacagcggca gaaggattta ggaaagatgg 10620
acgggagagt ggaatgctg cagaagccag aaaacaaagc aggaagcctg ctgtccagtg 10680
gggctcaaga gcggagggat gcgagggggc tgcgcaggaa catttagcgt ctgcgtctat 10740
gggggtaggg gcggggtgcc agcacctagt cacctgaagg ggaaatgctt gccagggag 10800
caggtctcag tagctgacct agagaaagga gcggccccta cagaggagac acgggtcact 10860
gtttgtaaa gtgaaggaga aataaatatt ctttcaaga atcttaggtg agcccagttc 10920
atctgcctg tggaggcctg gggaacagtt aaaaagacc tgacacacac ccaaggcaaa 10980
caagcaaac acggtcctt ccgtaagggt ceatgattct ctgaagaatc agccccgaa 11040
tcagccccgg aatcaggtag tccgtaaaca caatgagtgt tttactctgc agaagtcag 11100
cctgctggcg tctcccatta ccaaataga gggatagca cgtgagctca ccggtcgtat 11160
ttaaggcacg tggtttcca gggtagatga gcttggctt ctggaacct tatggggcac 11220
gaaggatgga gccagattt tttttttt tttttttt tattagcaat tgatttgett 11280
gggcttggct ggactgccc agttcttagg cccagtctt ttaactgcc atctgaagtc 11340
tgtcatggag tcagcctagc cttctcactt ccttcagct cgaataggaa gaggaggtgc 11400
acaccagatg gtctgagagc agggataaat ggtgtgcctt tgtcttcag tatttcgta 11460
ttttaagtag gaagatgctt ttctgtatta cattgcttgt gaaaccggaa gttgattcgg 11520
ggcacaggac aatggattg gtgtttgca aggactgtt cagaagagag aggagtggaa 11580
gggtggttag agtgaggagt ggggtgggac gggatggggg aagagaagga agggccagac 11640
aggctaggta gggctgagag gaggcgggtg gaactcttg agttagcgc gcagtaaact 11700
tggatgtgcg tctatcttg tgatatatga cccggagccg ttagctggc tccgatagta 11760
ctgctaafgt cagtgtcggg gggggggggg cccatactgt tccacagggg ctgcacattc 11820
ccatcgagag caggagggt cctctctcca tacactctg ccagcattcc ttgtgtttc 11880
tgtgatgaca ggggtggga tgaatctct ctgttggtt gagagaccgt gaagaagctc 11940
aaccacagga catttgcag tcttggaaag cagtgcctcc atgtggagcc gtggagccca 12000
tctctgagtc caggtcactc ttgagttcg cactcagctc ttcagatgca ggagagacgt 12060
tgggtggaaa gcaagattg ttgcttgtg agatagacac attctccaca caaaggctca 12120
cgtggggcaa aggctgattg acgtacagc ttcaggaac cctgtggtag agctatgatt 12180
agctgtctcc atctatgaag cagacaaaga gttataaaaa aatcaatgt tttcaaatg 12240
tcaactttt aaccgacag caagcctct gtcctgggc taatccctag cctggtttc 12300
ttgagatggg gtcttttg cactagactg gcctagaact cacgatctta gtgtccage 12360
ctcccagctg ctgggatgag ccgtataac cagtctgctt gccttctaa attttaagt 12420
atgggaagtg ggggagaata cagtttaaag tatgcagatc tgagagcagg aacctggca 12480
agccaagggg ccggagtac aggcggctaa catgggtgct gggaactgac ccaggtcctt 12540

gagaggagca gtgtgtactc ttgaccaaac aggtccgtct ctccagtccc cgtagtatta 12600
 aaaataghta ctacgggcat ggtggtgcac acctttaate ccagcactag ggaggcagag 12660
 gcaggtggat ttctgagttt gaggccagcc tggctacaaa aatgagttcc aggacagcca 12720
 cggctataca gagaaacct gtcttgaaaa caaaacaaca acaaaatagg tactacaaaag 12780
 cgatgtaatt gtgtcaaac atgcaaaccg aggggactgt atgcataaga aagagaaaaga 12840
 cggccacact ggttctatct gggtgacagg aatcagtat ttttatttt cacattcatt 12900
 tttttgtgt tgttgtgac acagtgattt ttctatcaaa aacattattt cttttatagt 12960
 tcccctgagg agctgttttt aaagccgtgc ttgaaaaaac cattgaagga gcagaggcag 13020
 ggagactcct gtgtggcagt cggggaagca ggcctctgc aggcaggctg gcctggact 13080
 tgggagtctc tttccctccc tctgtgctc aaatagcaaa tgcaggctt caatgtagct 13140
 agaaggttct agaatgatta agttccaag gctgaagagc ttcctgttt gcctttcact 13200
 tcctggaga ggtcgttgtg tgttccggag tctgcaaggt gcctttggtg atcggggtgg 13260
 ttcctctcgg gagattccgc ctggaggacc caagtcaag cctgcctga gctacagagt 13320
 gactttcagg tcttctcgc aattcagtg gaccagctc acaataaaa agtaaaaaga 13380
 aggctgtgga tggaaactcg tggtagagtt ctgggtttac tcctagagg aggggagaag 13440
 gaggaggagg gaggaggaag aggaagaaag aagaagaga gggaagagga gaaggaagg 13500
 agggaagggg ctgacaagaa gagagaagag ggaggagggg gagggaaagg aaggggaaag 13560
 gaagggaggg aaggggctga caagaagaga gaagaggag ggaggggagg gaaaggaag 13620
 ggaaagaaga gaaggtaag aagaaactgt tccaatggc tggccacag agtgatggc 13680
 ttttgggtg atcagctgta atccttgatt tgacacaacc tagaatctgg gaagcgagtt 13740
 tctgtgaag agcattcaca ctggctggcc tgtggcgctg catgtgggag actgtcataa 13800
 ttaggttcat taatacagga agtcccagcc cactacaaat ggcttcgttc cataccaag 13860
 agatgetaac ttagacaggt tggagaaagc aagcaagctg tggatacccc acgtctttc 13920
 acctcggctc ctggggggtg ggtgactgt gtctcttgg attttaaagt cctgcctga 13980
 cgtccctget gtgacagact gtaactggaa ttgtgagctt tagtcttta gttttctac 14040
 ttggttttc tcaggatatt ttatcgagt aacagaaaca agaccaggac actgatctc 14100
 ctctgateaa cactgaagag ttacaaaaca ggtgaggaa acaactttc ttctcctct 14160
 cccccctg tccctccc tcttctcgc tccctcctt gccccctc tccctgtctc 14220
 tgtctctg tctgtctg tctctctc tctctctc tctctctc tctctctc 14280
 ctctctct ctctctct ctctctct ctctctct ctctctct ctttctctc 14340
 tatctcctaa atggctggag gccatgctag ctcaatgtt aactttgaac acgtatttag 14400
 gaaatctttg ttctaacag ttctgaagt ctgaagtgg ggtttagtct ctggcctga 14460
 caagctcact tctctcact ctgtctaat gaccaaatc gccattccc taaaacaga 14520
 caggtccag ctccaggtg ctccggagcg gag 14553

ПРИМЕР 12 - Последовательности стабильного сайта 2 СНО - Патент США № 9816110

<211> 4001

<212> ДНК

<213> *Cricetulus griseus*

<400> 1

(SEQ ID NO:11)

ccaagatgcc catcaactga ttaatagatg ataaaattat tgtacatttc agtghtaatat 60
 tattcagttt ttaagaaaaa tgaattatg taataagcat gtaatggat atatcttgaa 120
 acaaccattc cccattatat tacctaaaca ttgaaagtcc aaaatcatat gatcttttta 180
 gtggatctac taatctttg ctatatgtat tttattgaac tacccatgga tgtgagataa 240
 ttgtaacaa cagcacatgg gagagcatgg gatcattcaa ggaagattag agagaatgca 300
 ttttttagga gataatggag gagcaataga aaggattaaa tgaggttact gatgaaagt 360
 atggttagag aaggcaatat gaggaggat aactagcact tagggccttt tgaaaaagac 420
 atagagaaaa tactattgta gaaacttct ataattggtg tatagttata tacaccaaag 480
 agtcagatg gagttacct ataatggaaa tattaactac tttttatcac tgtgataaaa 540
 catctgaac agagcaacat agattgggaa gcatttactt tggcttacag ttctaacggg 600
 ataaaaattc atgatgaaag aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga 660
 agcaggtgag agtcacatc ttgaagtga agaattgagc agagagaaca aactgcaaat 720
 gaccagaaaa tgcttttggg tcagagccca taccctctg actgacttct ccagaaattc 780
 tgaacaaata aaactcccca aacagagcca taactgaagg tccagtgtct gagactacta 840
 ggggtatttc ttattcaaac cactacaatg ggggtggggg agcaatcctc caagtaggca 900
 ctacacacag acaataaaaa actctagtaa ctggaatgga ttgacttatt tgaattactt 960
 gccagtggag ctacatagag cacaattatt gtatttaaat taccctttat gatcttaca 1020
 aacttgacag taagatcata ttgctaaaga aaccacatat ttgaatcagg gaacatggtg 1080
 atatctagtt gttctcaac tggaaacttc atgctttctg ccagcattc atgttgctgg 1140
 aaagagcaat gtacactacc agttagaana ttaaatcacc aatcttaca agatgtggat 1200
 cctataagtt acaataaaaa ttagectgat aagatatecc caccagaaga atattcacat 1260
 aaatgctatg ggagcaacaa gctattttct aaattagctt taatctatt ctacaagaga 1320
 gaatccatat ctagaatagt tatagggatc aagaacccat ggcttgattg gtcataggcc 1380
 caatgggaga tctaatatt atgtttctac aaaatgaaaa taactctaa tgactgttg 1440
 ctgcagtaat aagttagat gttgctcaac tctacaaga gaagttttgt cttacaataa 1500
 atggcaatta aagcagcccc acaagattta tatcataccg atctctcat ggctatgca 1560
 tctagaagct aggaacaaa gaggacccta agagagacat acatggctcc cctggagaag 1620
 ggggaagggg caagacctcc aaagctaatt gggagcatgg gggaggggag agggagttag 1680
 aagaaagaga aggggataaa aggagggaga ggaggacaag agagagaagg aagatctagt 1740
 caagagaaga tagaggagag caagaaaaga gataccatag tagagggagc cttgtatgtt 1800
 taaatagaaa actggcacta gggaattgc caaagatcca caaggtccaa ctaataatct 1860
 aagcaatagt cgagaggcta cctaaaagc ctttctctga taatgagatt gatgactacc 1920
 ttatatacca tctagagcc ttcatecagt agctgatgga agcagaagca gacatctaca 1980
 gctaaacact gagctagttg cagacagggg ggagtgatga gcaaagtcaa gaccaggctg 2040
 gagaacaca cagaaacagc agacctgaaa aaaatgttgc acatggacce cagactgata 2100
 gctgggagtc cagcatagga cttttctaga aacctgaat gaggatatca gtttgaggt 2160
 ctggttaac tatggggaca ctggtagtgg atcaatattt atccctagtt catgactgga 2220
 atttgggtac ccattccaca tggaggaatt ctctgtcagc ctgacacat gggggaggtt 2280
 ctaggtcctg ctccaaataa tgtgttagac tttgaagaac tcccttgaga agactcacc 2340

tcctgggga gcagaaagg gatgggatga gggttggtga gggacaggag aggaggggag 2400
 ggtgaggga ctgggattga caagtaaatg atgcttgttt ctaatttaa tgaataaagg 2460
 aaaagtaaaa gaagaaaaga aaacaggcca aaagattata aaagacagag gtggtgggtg 2520
 actataaaga aactactata tctaaataaa aacatgtcag aagcacacat gaacttatag 2580
 tgtttatgaa agtatgtata ataactacat aatctcaagc caagaaaaaa atatcatctt 2640
 tcagtgatga aggtgatttt atttctccca gaattaaagc caaagaccta atgaaagtaa 2700
 ttatcttcaa aaggttgaaa atacatactt tgcaatacac agatctgcct agaaatctca 2760
 tgttcacaat acacatgatg ctcaattgaa ttccattcaa tgttacagtt tagataaaca 2820
 gttttagat aaactcacia tgtatctttt ctttttattt ttgaccaa cagcttctca 2880
 tctgttattc agaataatc ctgatggca ggatatccat cccaattggg ggaaggggag 2940
 aatttgaaga aaacctagac cacatacata ttgcccattg ggaacaaaag tctaaatga 3000
 tgttgttcac atcttcteta ctagtctct ccccgccca aagaaccttg gtatatgtgc 3060
 ctcatcttac agagagagga aagcaggaac tgagcatecc ttaactgcca tectcaacc 3120
 aaaatttga cattgtctca gctctgcct tctcatatga cagttacaag tcaaggcttc 3180
 caaagtcct ctgtcatggt tgggtgcaat agtttataca gatgacttca tgtctcata 3240
 tctaatgtct tatatagatt aatattaac aatgttattt cttaaccac attttaatt 3300
 aatttaaaaa tccattaatt gtgtctataa aatgcagaca gagtgtgag acacaatata 3360
 agcctgatga tctgaatttg aaactcacac ccaccatg gagaatcaac ttccaaaaat 3420
 ttctctatta ctccacact tacaccattg tacaacaca ataataatga acaaatgaa 3480
 atgaaataaa aaattaagtc tctgtaggta atgctactgt gcagcaaaag taaaaatggc 3540
 agcttaaget tgccttatgg ttacacttta ccatctcca ttaattataa ggacttcaat 3600
 catggcagaa ctatgctggt atgtctcag tgtaacctaa ccaggtgttc cagatgttct 3660
 taatgtggac acctaaacta ttgatattt gggtaagat cttccctct ttcagaagaa 3720
 acctcaggac agaggggaatc ttgtcttta attttgagtc ttagacttt ttccattca 3780
 aatatacatg aaacaagtga tgaagaaat taatcaaaag gtgggaattg caatgatatt 3840
 aggttcaata ttaagcttca atattatcat ggaatgcct gttatactact gagtgtttgg 3900
 caataaggga tttttagaag aaggagtttt tattctcaac aggttctta agtttagctc 3960
 aaataaatct aagcaatcca ctctagaatt aatagtctc c 4001

<211> 14931

<212> ДНК

<213> *Cricetulus griseus*

<220>

<221> прочий_признак

<222> (2176)..(2239)

<223> n представляет собой а, с, г, т или нуклеотид отсутствует

<400> 4

(SEQ ID NO:12)

catgtacct tatgcaagta tgatatggcc caacacagta ttttacacca attttatct 60

ataaaatata catgtacatc aaaaatatt attaataata acatcattat tcttcttctc 120

caagtaataa acacatacac tgaattttg gttcttgtgg ataatttaa tgaacagga 180

aatgcaaatt tatcttagca tgtttacttc actttctttg catagataac cagtaatcac 240
attgatggat catgtagtga aatgtatfff taggtatcta aggaatfff gcttcgtttt 300
gtgcttggtg aactgaatt ctattcctaa caacagtgtg taaggattct gtctgatttc 360
ttttaccagt atttgtccat ttgcattttc tttattatc atggctgctg ttctagaaa 420
tggaaggtag tgtgtcaagt ctgtttaaca tgtttccctg atgatecagtg tcttaacacc 480
tctctgagta catgttggcc aatgtcgttt ctagacccat ctattcttgc ttgacttate 540
ctggtacatg cctgccaaga aatttctctc catcctttct gtctcttcac tgatttactt 600
gatgtgtgga tttcacattg atcatatgga aatagaagat acaatfffct ttattcacag 660
tttggaaagac tttcaatctc atagatcate attatfffft gctactgttc cctatgctat 720
gggtgaaatff ccatttgaat aattgcttaa acaattaaca agaaagaate tattttfact 780
tgcaataact tccattfcag aacattfact aactgttac tatatccaaa aactagtttt 840
atatacatg tgagaaatga ctaattcata atttggccat gacatfffft tcagaaacag 900
aaaaagtgc caatacatc acaatgctat aatattaag acttcagcaa attaaatatt 960
tattcatgat atcacataaa attcattfat tatgttttat ttaaagtgt ttttaaaaca 1020
gtggtatcac taaatattaa gttagatgtg tttatgtgct taatgaatff atatfftaga 1080
atgttataag ttgtatatag tcaaatatgt aataaattff atfffftagg tctttctcat 1140
taaggtatff taatffggg tccctfftc agagtgactc tagctcatga tgagttgaca 1200
taaaaactaa acagtacaaa atgtacattg cattcagtat tgcactgat ctttgactg 1260
aagtttgagt cagttcatac atttagactc tgggaagtac attaaagctaa ctttcattgc 1320
tctggcaaaa tctcagataa gataagagtc tattgtggaa agccatggca gcaggaaagt 1380
aagactgctg atgatgttta atccatagtc aagacgcaga aggagatgaa tctgtgtatc 1440
caacatffff tctgttcat tttctctaga accctagtc ataaagatgt atgacttga 1500
ttcaaatgc gtcccttca gttgttcaac ttttctgtaa atactcttc aggeatgtct 1560
agaagattgt ttcgcaata ctctcaatc cattcaagtt gatagtgcag attaatcact 1620
gcagaataaa agcctgtaac ttggctcag tgcgaaggaa tatgcacact cctgacacat 1680
caataagtaa atcaaatgt agctfftgcc tttaacattg ccagacttat gtaatgtct 1740
gcagttctt cctccatc tttttattct aatgggtfff ccttgacatt gaatcacgct 1800
gtggaagctg cttagaatta acattgaaat ctactgatat atttatgatg cagcaattta 1860
gattfactat tttactfaga atfffftata attgagagaa tataatattf tcacagttat 1920
ctatctgctg taaatagagg atfftaaaaa aaatctctat aactfffft tacaacacac 1980
agtaaaatta agttaaatt taataaagtc actatgttga tttcaaagtg tgctacgccc 2040
acgggtgta cgcaggtgta gcagaagatg ccaactaaggt gggctaagge cgatgggtg 2100
gggtctgcgc tccctggaga tgagccccag gcggttccct ggcaatcagc tgcgatcatg 2160
atgcccgatg agccannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220
nnnnnnnnnn nnnnnnnnnc tgggtgactt tatggaaaga atttgataga tttcatgatg 2280
tagaagaatt ttattagct tattttacag gagactaaga ccttgggacc taaagatate 2340
tgggtctga gaatcaggaa atgggtagag acgtggttga tggatgaga cagattttag 2400
agaactctta gatcatgggc aatgaccgca atctgatgct tagaatagat catctataaa 2460
caattatgct gttctfffft tttctgtgt atgatctgat gatgtagccc ccttggcaag 2520
ttccctgac cccctgcca agttccctga ttgtaacagt atataagcat tgcctgagag 2580

catattcaac tacattgagt gtgtctgtct gtcatttctt cgcggattcc tgatttctcc 2640
 ttgagccitt tccctgttc tccctcggtc ggtggtctcc acgagaggcg gtccgtggca 2700
 aaagtgtata aatgttctaa aacattttaa ctctaaaaca tgcaaatga aaaattaaaa 2760
 taaataaaca tgaaaattaa aatatattag ctgctaaaag ttaacaata ctatataata 2820
 tttgttatt agaattcaaa atcacattag ttggatttaa tttgaacatt gcattcttcc 2880
 aataataatt tcaataaaaa aagtttcccc atgatagtag aaaataataa catatgtatc 2940
 tatctattta tttaactaca catatfatgc atttgttca actaaaataa atgaatgagc 3000
 aaagcaccta agtaattggt gtctattata ttatgaagc caatagtctc aaataaatta 3060
 tcatgcataa ggaggtattg caaatgttaa accttttttg aaacagatat tcccagttac 3120
 agaaattata atttctaate ttctctataa gtagaatgat gataattaat ataggccatt 3180
 tgtaaataat gttcagatta aatattctc ttttctacta gagaagaatg atattaaatg 3240
 tattatattt tatttcccat tttgtttgca cactattct atatecctca gcagtttaaa 3300
 tttgtttcac catatgtgtg tgtgtttgta tcttaaatat ggcaactaaa ttagaataat 3360
 ttaataataa tctttaggag aaaagatatt gaattatttt atgttgatag gaaaatatct 3420
 tftaattgtc caagaatact ttttcttcta ttttaggact gatcagacc aggactaata 3480
 ttttatatgt actaattcta tgtacaaaa tatgttatta tctcatgaat tctgtctcaa 3540
 tattgagga ataaaaatag tccatcatga actttaaaat taaaataatg attaattaat 3600
 ttttattcat atttgtttg tatgaatggt tatacatcac atgtgtgcct ggtgactgtg 3660
 aatgtcagga gaaggtatga aagccactgg aattggaata agagataata ttgagatgt 3720
 tatgtgggtg ctgagaatta gacgcaagcc atctcaaga atagccagca tactatacca 3780
 ctgagtaate cattcatecc tcaataatta tctttgtaga cagtaaatat atttetaaac 3840
 tataaatgac cagaaaaatt aatgtattat taatgaagac atctatctca tgtgacacac 3900
 ttcactgtc taaatcagta acactctctc cactaattaa gattttctaa gtgcatgaca 3960
 cttactattt ctaaagctgt ccaatggggg ccagtcacca gtcagcacc agtgagataa 4020
 tccatgaatg catttatate ttaggaaaa ttcttatcta ttagtatttt agaacatttt 4080
 catgtgaggg gataaacaag gaagcacaga tgccttctga tagaaacttt ctctttaatt 4140
 catctagaaa aaaaaaacct ctcaggaaaa tctctcttgc tctctccca atgctctatt 4200
 cagcatcttc tccctactta attctagatc ttttctcta tgcctcttgc ctgctgcct 4260
 gctggctctg ctctatgctt ccccatgtca ctttttttg ctatctcacc gttacctct 4320
 ctgctcact ctctgcttc ttctctgctt ctacatggc caggctctgg acaattatag 4380
 ttatattgta cattctcata acacatgata tgcacatag tttctctcag gctagggata 4440
 tcacaatgac tggccaatga gcaagtggcc ttgcatgtag ctctaagtgt gtgatggctc 4500
 ccagacagta agtagccatt tggttgaaat ttgaggttgg gtagtcatg aagactgaat 4560
 tttctcaaa ctctggcctt gaaatagtaa aacaacacct atgaaaatga cgacctgat 4620
 ttgtctttag aggcaaccac atattgtctg caggcctgc tttgaattg ctctgaagt 4680
 agcttgtttg tgtaaaagga agaactctat atcagcctga gaaatgtaa atactctagc 4740
 atttcaagtc atcaaaatta tatggagagt ataatcctc ctctgacta ttcattatca 4800
 tatttgtctc caccaagtat aaaacacact accaaagggc tgtggaaaaa atgccataa 4860
 ctgttcttat tagggaggca tagcagtggg acctgaggaa gttacagcaa caaccagtca 4920
 tccagtcaat aacccatggt ctttgcact tggaggtacc caataatgtt tggctttgcc 4980

gagtaggact ccaacaaatt cagagggtca attttfaat gctggtgtc actgctgaac 5040
 agtcccattg ccctctgcat aatccacaa tggaaagctt tttactga ttgccaatca 5100
 ttaaacagcc tactcagcat aaacaggtat gatattatc tgcattttgt tacattacta 5160
 gatgaattcc tatttctcc tacaatagt gaaactgaaa aagatacaca atcactactac 5220
 cccttacta atcttatgac ttatatcatt tcaattttca gaccataatg caaactattg 5280
 accaaaacat gtgaagatga aaaatagaaa tgtagaataa tattacatat aaaaagaaaa 5340
 ggcggactta tttgtttta tttcttagca tgcatagcaa tacatgattt gaggtttata 5400
 taataaaggg acaataaatc ttcaagaaac ttaccctac tgaattaaaa tattaaagaa 5460
 ggtcacacat ttactcaaat atattagact actgggcaaa tagacatgaa aagtagagtt 5520
 aatattgagg taggccttct gtgaaatgtc taaggaaatt atgtttcata cagtgtgtaa 5580
 ccaagtggga atcatatcag aaagcagtc aaagcttata ttacaagtaa cagatgcttg 5640
 gttatatgac ctcccagagc ttgactgtct atacacaaa agtgggtgta ataaaactgt 5700
 aatttgggct atgtttttt aaatggctc accaacatga aaggaaggga atgagcatgt 5760
 catggatgct tagagattat gcttccagca agaagaattg agctttggct cttattacag 5820
 aaacatgaca aggtgtgagt tttatttatt agaaattata taatattta agctggggac 5880
 taaaaattt attgaaacaa acaggcaagg gatagggcatg tactagaagc aaaaatagga 5940
 tgtcaatgct gtaatgttat ttttggacc aaaatagtat ttctataga aatgacaatg 6000
 atcttaggtt attattctc ataaagatga caagttcaca agatataccta gttcattaa 6060
 atcgttttag tcaattata gagtctgtg atagattaca caaaggaaag cacttacgat 6120
 gagaaataat gatatacaca attattttct taattcttag aacattcta ttgttatatc 6180
 tcaatctcag aagccactta ttgctttatt attgaaacat atgaaattgt aagttatata 6240
 ttgtctatgg tgacattca aagaacatgt gacgtacagt gtagcacaga taaagaacat 6300
 aactgcagct gaatcagtaa ctaaaactac atacattaaa tctgcatgt tggcaacagt 6360
 gtgtgcaact ccaaaggatg tactaatgct cagcactc ccctatgca cctttgttc 6420
 atcattacat cataggtcta tttgtttgc tttgaaatc tagaccaagt cttttgtgc 6480
 ttccaagca cagagctcat taattacct catagactg ttaaacttct tctggttcat 6540
 caattgaata gaaataccta ctactaatta tgtgagacc tggcagttacc atagcacatg 6600
 gataatttt acataaaaca tgcatacaag taagattatt cagactgaac atgaattta 6660
 gagaaatcag gaaggagtat atgggagtgg ttggagtgag actagagaaa tgaattaaa 6720
 ctataatctc aatacaaaaga tctactaagc aaaaacatg aacattgtc attcaagtga 6780
 aacatcagtc ttcaaatgg aaagatattt ttactaggaa aatgtctggt agatggttat 6840
 tatctagaaa acacaaaaat tagaaaacgg taaactftaa taaaagaat aatacaatga 6900
 gactacatga aaagtctta actaatgaaa caaatatctt gaaactttt tcttaaaagt 6960
 ttaatatcaa taaccatcat ggaaatcaa attaaaacta ttacatatt acccctgaaa 7020
 taataactaa taccaataa aaataatata aacaaaaat ggcaatgcat gccatcatgg 7080
 atttgggaga gagaatgttc attgcagttc tgaatggata ctgggtccac cacggtgaaa 7140
 atctctgtat agtctctcc aaaagctgaa aatagacata tcacaagacc tgccacacat 7200
 tttcaagca aatacccaaa ggactctacc tgactgcaga gacactttct cataaaat 7260
 tattgtgat ctattcataa tatctgaaa atagaacag ccaagatgcc catcaactga 7320
 ttaatagatg ataaaattat tgtacattc agtgaatat tattcagttt ttaagaaaa 7380

tgaattatg taataagcat gtaaatggat atatcttgaa acaaccattc cccattatat 7440
 tacctaaaca ttgaaagtec aaaatcatat gatcttttta gtggatctac taatcttttg 7500
 ctatatgtat tttattgaac tacccatgga tgtgagataa ttgtaacaa cagcacatgg 7560
 gagagcatgg gatcattcaa ggaagattag agagaatgca ttttttagga gataatggag 7620
 gagcaataga aaggattaaa tgaggttact gatgaaagtg atggttagag aaggcaatat 7680
 gaggagggat aactagcact tagggccttt tgaaaaagac atagagaaaa tactattgta 7740
 gaaacttctc ataattgggtg tatagtata tacaccaaag agctcagatg gagttaccct 7800
 ataatggaaa tattaactac tttttatcac tgtgataaaa catcctgaac agagcaacat 7860
 agattgggaa gcatttactt tggtttacag ttctaaccggg ataaaaatc atgatgaaag 7920
 aatgaatag tcagcaaca gcagtagcaa tggcctgaga agcaggtgag agctcacatc 7980
 ttgaagtga agaattgagc agagagaaca aactgcaaat gaccagaaaa tgcttttga 8040
 tcagagccca taccctctg actgacttct ccagaaatc tgaacaaata aaactcccca 8100
 aacagagcca taactgaagg tccagtgtc gagactacta ggggtatttc ttattcaaac 8160
 cactacaatg ggggtggggg agcaatctc caagtaggca ctacacacag acaataaaaa 8220
 actctagtaa ctggaatgga ttgacttatt tgaattactt gccagtggag ctacatagag 8280
 cacaattatt gtatttaaat taccctttat gatcttaca aacttgacag taagatcata 8340
 ttgctaaaga aaccacatat ttgaatcagg gaacatgggtg atatctagtt gttcttcaac 8400
 tggaaacttc atgctttctg cccagcattc atgttgctgg aaagagcaat gtacactacc 8460
 agttagataa tfaatcatc aatcttaca agatgtggat cctataagtt acaataaaaa 8520
 ttgcctgat aagatatccc caccagaaga atattccat aatgctatg ggagcaacaa 8580
 gctattttct aaattagctt taactctatt ctacaagaga gaatccatat ctagaatagt 8640
 tatagggatc aagaacccat ggcttgattg gtcataggcc caatgggaga tcctaatatt 8700
 attgttctac aaaatgaaaa taactctaa tgactgttg ctgcagtaat aagttagtat 8760
 gttgctcaac tctcaaga gaagttttgt cttacaataa atggcaatta aagcagcccc 8820
 acaagattta tatcataccg atctctcat ggcctatgca tctagaagct aggaaacaaa 8880
 gaggacccta agagagacat acatggctcc cctggagaag gggagggggg caagacctcc 8940
 aaagctaatt gggagcatgg gggagggggag agggagttag aagaaagaga aggggataaa 9000
 aggaggggaga ggaggacaag agagagaagg aagatctagt caagagaaga tagaggagag 9060
 caagaaaaga gataccatag tagagggagc cttgtatgtt taaatagaaa actggcacta 9120
 gggaaattgc caaagatcca caaggtccaa ctaataatct aagcaatagt cgagaggcta 9180
 ctttaaaagc ctttctctga taatgagatt gatgactacc ttatatacca tcttagagcc 9240
 ttcattccagt agctgatgga agcagaagca gacatctaca gctaaacact gagctagttg 9300
 cagacaggga ggagtgatga gcaaagcaa gaccaggctg gagaacaca cagaacagc 9360
 agacctgaaa aaaatgttgc acatggacc cagactgata gctgggagtc cagcatagga 9420
 cttttctaga aacctgaat gaggatatca gtttgaggt ctggttaatc tatggggaca 9480
 ctggtagtgg atcaatatt atccctagtt catgactgga atttgggtac cactccaca 9540
 tggaggaatt ctctgcagc ctagacacat gggggagggt ctaggctctg ctccaaataa 9600
 tgtgttagac ttgaagaac tccctgaga agactcacc tccctgggga gcagaaaggg 9660
 gatgggatga gggfttggtga gggacaggag agggaggggag ggtgagggaa ctgggattga 9720
 caagtaaatg atgctgttt ctaatttaaa tgaataaagg aaaagtaaaa gaagaaaaga 9780

aaacaggcca aaagattata aaagacagag gtggtgggtg actataaaga aacactatta 9840
tctaaataaa aatattgacag aagcacacat gaacttatag tgtttatgaa agtatgtata 9900
ataactacat aatctcaagc caagaaaaaa atatcatctt tcagtgatga aggtgatttt 9960
atftctccca gaattaaagc caaagaccta atgaaagtaa ttatcttcaa aaggttgaaa 10020
atacatactt tgcaatacac agatctgcct agaaatctca tgttcacaat acacatgatg 10080
ctcaattgaa ttccattcaa tgtfacagtt tagataaaca gttttagatg aaactcaca 10140
tgtatcattt ctttttattt ttgacccaaa cagcttctca tctgttattc agaataattc 10200
ctcgatggca ggatatccat cccaattggg ggaaggggag aatttgaaga aaacctagac 10260
cacatacata ttgcccattg ggaacaaaag tctaaatga tgttgttcac atcttctcta 10320
ctagctctct ccccgctcca aagaaccttg gtatatgtgc ctcatcttac agagagagga 10380
aagcaggaac tgagcatecc ttacttgcca tctcaacce aaaatttga ctattgctca 10440
gctctgcct tctcatatga cagttacaag tcaaggcttc caaagtcct ctgctatgtt 10500
tgtgtcaat agtttataca gatgacttca tgtctcata tctaatgtct tatatagatt 10560
aatattaaac aatgttattt cttaaccac attttaaatt aatttaaaaa tccattaatt 10620
gtgtctataa aatgcagaca gagtgtctgag acacaatata agcctgatga tctgaatttg 10680
aaactcacac ccaccacatg gagaatcaac ttccaaaaat ttctctatta ctccacact 10740
tacaccattg tacaacaca ataataatga acaaaatgaa atgaaataaa aaattaagtc 10800
tctgtaggta atgctactgt gcagcaaaag taaaaatggc agcttaagct tgccttatgg 10860
ttacacttta ccacttcca ttaattataa ggacttcaat catggcagaa ctatgctgtt 10920
attgtctcag tgtaacctaa ccaggtgttc cagatgttct taatgtggac acctaaacta 10980
tttgatattt gggftaagat ctttccctct ttcagaagaa acctcaggac agagggaate 11040
ttgtctttta attttgagtc ttagactttt ttccatttca aatatacatg aaacaagtga 11100
tgaagaaaat taatcaaaag gtgggaattg caatgatatt aggttcaata ttaagcttca 11160
atattatcat ggaatgcct gtatatact gagtgtttgg caataaggga tttttagaag 11220
aaggagtttt tattctcaac aggttcttca agtttagctc aaataaatct aagcaatcca 11280
ctctagaatt aaatagtttc ctaagggcac agctatgaat agagctcaat ttacataaa 11340
aattttgttc accatttatg teattccagt ttccattagt acaaggaaaa taaaaaatat 11400
ttagatgtca atatacaagt aatagttcat ctctttttt aatatatc acctaaatca 11460
ccattttctc agaaaaatct ggcctgaagt tctgtctgga acttcaacat gaaaaaatg 11520
cacagcttgc tattataat cctagttgat ttttaagatt catgtctggt gtctgactca 11580
gaggggccag aggctagaca aatatttttt gaacttctcat tctgaagatt tttatgatt 11640
attttaatat aaatacaaaa gatgatggat aatgtaactt tctacagttc atagacgctg 11700
aactactttg tctttaaata gttagttccc taccataat gataggtgat aagtgtatgt 11760
ttaactttt cctctgagc tatattcatg tactagagaa ttattttaaa catgaaaaga 11820
ctgtgtttat agtctcagct cctgagaact ggtccaacct taggcaggtg aatgccagga 11880
gcaacgtttt tctctacag aggatgcttt gctgccaagc aacctgggtg tctggaatg 11940
ttctttttt aatcaagttt aaagggtctt catcatgctg ttgctccaca tattttcagg 12000
ttagagcttg gtccttgag tattatcttt taccagaaaa tcatagtat tctttcaata 12060
actaacaact aaacttttcg ataaaaaaga atfggaattt caattttaaa gctgagtaa 12120
aattctgtg aatcaggata tttattttta agtcttatct tttaaaaagt tattttattt 12180

tftaaaaaat tataatatac tftcataaft tectccttc acttttcttt acaaacactt 12240
 ctatagatca ccatgtgttt tttttttac atttatggcc tctttctgtt cattgttatt 12300
 acatacaaat agtcttgcct atagaagaac accacaatft gttacctgat aacaaattat 12360
 caacccttaa aacctacaaa ctattgatat tactgaaaag actatactta tagatgtaaa 12420
 gatatatgtg tgtgcacata tatagatata catatatgta ggatttttaa ttttagattt 12480
 tagacatcaa aattatftat atgactgaga aactagacac tataaatgag cattcagtat 12540
 tcaacaccgt gattttgat atgtcacia tgacagaaaa tttcttata gaaaatttta 12600
 agttttgtga ttgctctgtg cacttagtga agtctcacag aaaaagaatc atagtatttt 12660
 tagttataa taaaaagtac atataaftaa aatgggtggc acaaaacaac atttgagcat 12720
 ttttctatt tactatcaag tagtatcatt ttgaaataat aatttgacta gtttcaaaaa 12780
 tgaaaacaaa atftaaacta aatgcctaata ctgacctgat aacattttta tgaatgaaat 12840
 tattcaatag tgtatcaat taggggceca aaacttttcc taaaataaaa cttttaattt 12900
 ttttccattt ttattfaaat tagaaacaaa atgttttac atgtaaatca gagtttctc 12960
 accctccct tctcctgtc cctcactaac accctactg tccatacca tttctgctc 13020
 ccagggaggg tgaggccttc catgggggaaa cttcagagtc tgtctatcct ttccgatagg 13080
 gcctaggccc tcaccatft gtctaggcta aggcacaaa agtttactcc tatgctagtg 13140
 ataagtactg atctactaca agagacacca tagatttctc agcttctc actgacaccc 13200
 atgttcatgg ggtctggaac aatcatatgc tagtttcta ggtatcagtc tggggacat 13260
 gagctcccc ttgttcaggt caactgttc tgtgggttc accacctgg tcttgactgc 13320
 tttgctcctc actcctcct ttctgtaact gggttccagt acaattccgt gtttagctgt 13380
 ggggtgtctac ttctacttc atcagcttct gggatggagc ctctaggata gcatacaatt 13440
 agtcatcctc tcattatcag ggaaggcat taaagtgc ctctccattg ttgcttgat 13500
 tgttagttgg tgcactttt gtagatctct ggacatttcc ctagtgccag atatctttt 13560
 aaacctaaa gactacctc attatggtat ctcttttctt gctctctctt attcttccag 13620
 acaaaatctt cctgctcct tatattttcc tctccctcct tcttctccc ttctcattt 13680
 cctagateca tcttccctc cccatgctc ccaagagaga tttgctcag gagatcttgt 13740
 tcttaaccc tttcttggg gatctgtctc tcttaggtt gtcttgttt cctagcttct 13800
 ctggaagtgt ggattgtaag ctggtaatca tttgctccat gctaaaaac catatatgag 13860
 tgatgtttgt cttttgtga ctgggttacc tcaactcaaaa tggtttctc catatgtctg 13920
 tggatttcaa tagcacaac aacatacagt atcttggggc aactaacc aaacaagtga 13980
 aagaccagta tagcaagaac tttgagtta aagaaagaaa ttaaagaaga taccagaaaa 14040
 tggaaagatc tccatgctc tttgataggc agaatacaaa tagtaaaaat ggcaatctt 14100
 ccaaaatcca tctacagact caatgcaatc ccaatfaaat accagcacac ttctcacag 14160
 acctgaaaga ataatactta actttatag gagaacaaa agaccagga taggcaaac 14220
 aacctgtac aatgaaggca ctccagagg catccccatc cctgacttca agctctatta 14280
 tagagtaata atctgaaaa cagcttggta atggcacaaa aatagacagg tagaccaatg 14340
 gaattgagt gaaaacctg atattaaccc acatatctat gaacacctga ctttgacaaa 14400
 gaagctaagg ttatacaatg taagaaagaa agcatcttca acaaatctg ctggcataac 14460
 tggatgctgg catgtagaag actgcagata gatccatgct taatgccatg cacaaaactt 14520
 aagtccaat ggatcaaaaa cctcaacata aatccagcca cactgaacct catagaagag 14580

aaagtgggaa gtatccttga ataaattggt acaggagacc acatcttgaa cttaacacca 14640
 gtagcacaga caatcagatc aataatcaat aatgggacc tctgaaact gagaagcttc 14700
 tgtaaggcaa tggataagtc aacaggacaa aatggcagcc cacggaatgg gaaaagatat 14760
 tcaccaatcc tataatctgac agagggtctc tctctatttg caaagaacac aataagctag 14820
 ttttfaaac accaattaat ccgattataa agttgggtag agaactaat aaagaattgt 14880
 taacagagca atctaactg gcagaagac acataagaaa gtgctcacca t 14931

ПРИМЕР 13 - Гидовые последовательности для последовательностей AAVS1-подобных областей в СНО

(Приведенные ниже гидовые последовательности могут представлять собой смысловые гидовые последовательности или антисмысловые гидовые последовательности)

	SEQ ID NO
CCCCGCTGGCGCCGGGATCGGGG	13
GAGTCGAGCACCGCTCGGGCAGG	14
TTCCCCGCTGGCGCCGGGATCGG	15
GTGTGCGGAAGACGCCCGGGG	16
CGGTGACAGCGCGGATGACAGGG	17
CAGCGCGGATGACAGGGGCGAGG	18
GCCGGCGTCCGATTCCCCGCTGG	19
CGTGTGCGGAAGACGCCCGGGG	20
GAGGCGCTCCACCGTCTGTTGGG	21
GTCCGATTCCCCGCTGGCGCCGG	22
GACCCCGGGGGCCCCGATCCCGG	23
CGGCGTCTTCCGCACACGGATGG	24
TCGAGCACCGCTCGGGCAGGCGG	25
AGCTCACGCCGGCCCCATAAAGG	26
CATCGTCCTCTATATATAGCAGG	27
AGAGGCGCTCCACCGTCTGTTGG	28
GGTCGGCTGCGCGAAGCATCAGG	29
TGCTTCGCGCAGCCGACCCCGGG	30
GGCCCCGATCCCGGCGCCAGCGG	31
TCCGATTCCCCGCTGGCGCCGGG	32
TCCCGGCGCCAGCGGGGAATCGG	33
TGGTGGAGTCGAGCACCGCTCGG	34
CAAGATGGTCCTCACTCTCGGGG	35
GCTTCGCGCAGCCGACCCCGGGG	36

	SEQ ID NO
GTGGAGCGCCTCTTCTCCAGGGG	37
GACGTGTCAGCCTTCCAGGTGGG	38
GCCAGCGGGGAATCGGACGCCGG	39
TCTCCCCGTCATCCAAAAGCTGG	40
GGTGGAGTCGAGCACCGCTCGGG	41
GCTGCCCAAATATAGTCCATGGG	42
ATGCTTCGCGCAGCCGACCCCGG	43
GCGGTGACAGCGCGGATGACAGG	44
ATGCTCGGGGGCCGCTGACCTGG	45
TTGTATTGCCGGGATCCTTCTGG	46
TCCCCGCTGGCGCCGGGATCGGG	47
CTCGACTCCACCAACGCCGACGG	48
GGTGGCAAGATCACCAAAGGGG	49
GGCGCTGATGCCGTCGGCGTTGG	50
TCCACGAGCATCCTAGCAAGAGG	51
AGGCTGACACGTCAGGCCTGAGG	52
GGGTGTAAGCCATCCGTGTGCGG	53
AGGATCCCGGCAATACAAGATGG	54
GGATGGGGCCCAACAGACGGTGG	55
GAGGACCATCTTGTATTGCCGGG	56
AGTCGCCCAGGGTCCTGGTGGGG	57
GGCGGGAGGAGAGTCCACCTGG	58
ACCTACCCCACCAGGACCCTGGG	59
TCAGCGTCTTTGACCAGTCCAGG	60
CGTCCCGCCGCCTGCCCGAGCGG	61
GTCCCCGGGATCCCCGGGGTTCGG	62
AGCACCGCTCGGGCAGGCGGCGG	63
CGGTGGAGCGCCTCTTCTCCAGG	64
GCCCCGATCCCGGCGCCAGCGGG	65
ATCGGGGCCCCCGGGGTCCCCGG	66
ATTCTCGGCTCATCCCCTTTTGG	67
ACCACCCCATGGACTATATTTGG	68
TAGCAAGAGGACGACAACCCAGG	69

	SEQ ID NO
GCTGATGCCGTCGGCGTTGGTGG	70
TACAAGATGGTCCTCACTCTCGG	71
TGCCGGGATCCTTCTGGATTCGG	72
CCCAAATATAGTCCATGGGGTGG	73
TACCTGTAGAATGGGACCAGTGG	74
TCTTGCTAGGATGCTCGTGGAGG	75
ATGCCAGCTTTTGGATGACGGGG	76
ATAGTCCATGGGGTGGTAGGTGG	77
CAGGACCCTGCTATATATAGAGG	78
GGGGCCGGCGTGAGCTGTGTGGG	79
ACCTGGAAGGCTGACACGTCAGG	80
CAGCGGACAGCACGGGTCACAGG	81
CGCCGGGATCGGGGCCCCGGGG	82
GCGCCGGGATCGGGGCCCCGGGG	83
GCCGGGATCCTTCTGGATTCGGG	84
GCTGTCACCGCTCTCCCCGGCGG	85
GGTGACAGCGCGGATGACAGGGG	86
GAAGACGCCGCCGGGGAGAGCGG	87
AAGGGGTACACTGCCTTGGAGGG	88
CTTCTAGAACCTACCCCACCAGG	89
GGGGCCGCTGACCTGGTGCAGGG	90
TCCCGAATCCAGAAGGATCCCGG	91
TCACAGTGTCTGAGTCGCCCAGG	92
GCTAGGATGCTCGTGGAGGTGGG	93
CTCGTGGAGGTGGGGAATAAAGG	94
TGATCTGTGCCCCGAGAGTGAGG	95
TGAATTAATAGGACATGGGGAGG	96
ACTCTCGGGGCACAGATCACTGG	97
GACCACTGGTCCCATTCTACAGG	98
TGAGGACCATCTTGTATTGCCGG	99
AGAATCCGTCTGTCCTGGGCTGG	100
TGACGTGTCAGCCTTCCAGGTGG	101
AAAAGCATCCCGAATCCAGAAGG	102

	SEQ ID NO
TTTTCCCGAGGCCACACTCAGGG	103
AGCGCCCTGCACCAGGTCAGCGG	104
TGAGTCGCCCAGGGTCCTGGTGG	105
CACGTCAGGCCTGAGGTCACAGG	106
GTTTTCCCGAGGCCACACTCAGG	107
GAGTCGCCCAGGGTCCTGGTGGG	108
GAATCCGTCTGTCCTGGGCTGGG	109
CCGGGGAGGGAGGATGCTCGGGG	110
GAGCCGAGAATTGAATTAATAGG	111
GTGACCCGTGCTGTCCGCTGTGG	112
GGGGGCGTCAAGTCAGAGCTGGG	113
CGCGCTGTCACCGCTCTCCCCGG	114
TGCTGCCCAAATATAGTCCATGG	115
AGTTGAGGAGAAACCTATGGGGG	116
ATCGTCCTCTATATATAGCAGGG	117
GAGTTGAGGAGAAACCTATGGGG	118
AGGGGTACACTGCCTTGGAGGGG	119
ATAGAGTCCCTCTGGGGACAGGG	120
CTAGGATGCTCGTGGAGGTGGGG	121
CGCGGATGACAGGGGCGAGGCGG	122
AAAACAGAATCCGTCTGTCCTGG	123
TGCTAGGATGCTCGTGGAGGTGG	124
GAGTGTGGCCTCGGGAAAACAGG	125
CGAGGCGGCCCTGCAGGGCAGG	126
CTTACACCCCGTGCCTTTCCAGG	127
CTGCCCAAATATAGTCCATGGGG	128
AGGGGCAAAGGACCCTCCTGAGG	129
TCTCACCATAGAGTCCCTCTGGG	130
AGTGTACCCCTTTGTTCCCCTGG	131
GTTGAGGAGAAACCTATGGGGGG	132
CACAGTGTCTGAGTCGCCCAGGG	133
TCCTCTTGCTAGGATGCTCGTGG	134
CACATGATCACCAAAGTCCCTGG	135

	SEQ ID NO
AAACAGAATCCGTCTGTCCTGGG	136
GTGCAGGGCGCTGATGCCGTCGG	137
CAGCACGGACTTTTTTTGTTTGG	138
CAGCCGACCCCGGGGATCCCGGG	139
TTTGGTCAAGATTTTGCAACTGG	140
TTGGTCAAGATTTTGCAACTGGG	141
CCTGTCAGAGAGGATGCTCTAGG	142
TGGGTGTCGTCCTCTTGCTAGG	143
CTCACCATAGAGTCCCTCTGGGG	144
AGACGCTGACACCCTGAGTGTGG	145
GGTGGAGCGCCTCTTCTCCAGGG	146
AACAAAGGGGTACACTGCCTTGG	147
CTGTCCCCAGAGGGACTCTATGG	148
GTCTCTGACCCCTCATTTGTGG	149
GTCTGAGTCGCCCAGGGTCCTGG	150
GCCTGACGTGTCAGCCTTCCAGG	151
ATGTCCCACACAGCTCACGCCGG	152
GGCTGTGCCGCAGGCTTCCAGGG	153
CACCCCGTGCCTTTCAGGCTGG	154
TCACCAAGGTGTCTGCATGGCGG	155
GAATTGAATTAATAGGACATGGG	156
AATATAGTCCATGGGGTGGTAGG	157
CAGGGCATAGTTTTTAAAGCAGG	158
CGGCATCAGCGCCCTGCACCAGG	159
GCAGCCGACCCCGGGGATCCCGG	160
ACAGGACCTGTATTTGAGGTTGG	161
GGGGCGAGGCGGCCCTGCAGGG	162
AACCTACCCACCAGGACCCTGG	163
GCACCGCTCGGGCAGGCGGCGGG	164
TGGAAGCCTGCGGCACAGCCAGG	165
GAACCAAACTGTGGCCAGGAGG	166
AGCAGGGTCCTGTTTTCCCGAGG	167
TGGGTGGCAAGATCACCAAAAAGG	168

	SEQ ID NO
AGCCGACCCCGGGGATCCCGGGG	169
CATGGCAACTTCCATCTCCTGGG	170
ACAGCACGGGTCACAGGAAGTGG	171
GGGGGCCGCTGACCTGGTGCAGG	172
AGGGGCGAGGCGGCCCTGCAGG	173
GTGGTCACCCCTGTCCCCAGAGG	174
TGGGGCCGGCGTGAGCTGTGTGG	175
GGAAGCCTGCGGCACAGCCAGGG	176
CCTGAGCTGATCTCCTGGACTGG	177
ATCCAAAAGCTGGCATTGTCAGG	178
CTCTCACCATAGAGTCCCTCTGG	179
GGGGGGCGTCAAGTCAGAGCTGG	180
GGGTGGAAATCTAAGAGACAGGG	181
CGGGGAGAGCGGTGACAGCGCGG	182
AGAATTGAATTAATAGGACATGG	183
CCATAAAGGAAGTTTTCCACAGG	184
AGTGAACCAACACTGTGGCCAGG	185
GTTGGGAGGGAACTCTTGGGAGG	186
CGGGTCACAGGAAGTGGGGTAGG	187
GGCCTGGCTAGCCTCAGAGGAGG	188
GGTAGGTTCTAGAAGGTGACAGG	189
ACAAGATGGTCCTCACTCTCGGG	190
CAAGGTGTCTGCATGGCGGGAGG	191
TGTTTCACTCATCCAGGCAGAGG	192
TGTGGAAAACCTCCTTTATGGGG	193
GAGGACGACAACCCAGGAGATGG	194
GGGTGGCAAGATCACCAAAGGG	195
CTGGTGGGGTAGGTTCTAGAAGG	196
ACTCTTCAGGCCTTTGCAGGAGG	197
CAGAGGGACTCTATGGTGAGAGG	198
AGCACGGGTCACAGGAAGTGGGG	199
CTATGGTGAGAGGCGTCCTGTGG	200
CAGCACGGGTCACAGGAAGTGGG	201

	SEQ ID NO
ATGGGATGGGGCCCAACAGACGG	202
CTCCCGCCATGCAGACACCTTGG	203
GGGGATCCCGGGGACCCCGGGG	204
TGTCCGCTGTGGCCTCAGGAGGG	205
GCAGTTGGGAGGGAACCTTTGGG	206
GCCTGGCTAGCCTCAGAGGAGGG	207
GAGAGTTGAGGAGAAACCTATGG	208
ACCTGTATTTGAGGTTGGCCTGG	209
TCGGGCAGGCGGCGGGACGCCGG	210
TAGAGTCCCTCTGGGGACAGGGG	211
GAAGTGACACTGAAGGGCCTGGG	212
AGCAGCCTGAGCTGATCTCCTGG	213
TCATGGCAACTTCCATCTCCTGG	214
GTCACAGGTTCCCTGTCAGAGAGG	215
CACCAAGGTGTCTGCATGGCGGG	216
GTAGGTTCTAGAAGGTGACAGGG	217
AATTGAATTAATAGGACATGGGG	218
TTCCAAGCACCTGATTTCTGTGG	219
CTGTCAGAGAGGATGCTCTAGGG	220
GTGCTGTCCGCTGTGGCCTCAGG	221
CCCGGGGAGGGAGGATGCTCGGG	222
GGGGTGGCTCGGGGGGCCCCGGG	223
GTCAAGTCAGAGCTGGGCCCTGG	224
GACTACTGAAGGGCCTGGGCCTGG	225
CTGAAAGTGAACCAACACTGTGG	226
AAAGGGGTACACTGCCTTGGAGG	227
TCTGGAAACTTCTAAGCATTCGG	228
AGAGTTGAGGAGAAACCTATGGG	229
CTCTGAGGCTAGCCAGGCCCAGG	230
AGGGGTGGCTCGGGGGGCCCCGG	231
TCCACATTGATTTGCCTTTCTGG	232
AGGGTGGAAATCTAAGAGACAGG	233
CTGTCCGCTGTGGCCTCAGGAGG	234

	SEQ ID NO
AGACACAGGACCTGTATTTGAGG	235
GGGTGGCTCGGGGGGCCCCGGGG	236
TGGCTGTGCCGCAGGCTTCCAGG	237
TCCAGAAAGGCAAATCAATGTGG	238
CACCAGCCTGGAAAGGCACGGGG	239
CGGGGTCCCCGGGATCCCCGGGG	240
AACTGCCTTGAGGGGGCAAAGG	241
TGACAGGAACCTGTGACCTCAGG	242
TCTCATGTGGGCTATCAAGATGG	243
GGGAACTCTTGGGAGGGCCAGGG	244
GAGGCCACAGCGGACAGCACGGG	245
TGTCCTATTAATTCAATTCTCGG	246
AGGGAACTCTTGGGAGGGCCAGG	247
GCACCTGATTTCTGTGGTATTGG	248
GTATCTTGAGTGTCTTTTCTCGG	249
CTGTGGAAAACCTCCTTTATGGG	250
GGGACCTCCAGCAGATGCAGAGG	251
CTGGGGACAGGGGTGACCACTGG	252
TTCAGTGTCACTTCTTTTGGGGG	253
TCCCTCCTCTGAGGCTAGCCAGG	254
GGCGCCGGGATCGGGGCCCCCGG	255
GGTTCTAGAAGGTGACAGGGTGG	256
TTGGGAGGGAACTCTTGGGAGGG	257
ATCAGGTGCTTGGAAAGTAGAGG	258
CGGGGAGGGAGGATGCTCGGGGG	259
TGAGGCCACAGCGGACAGCACGG	260
TGGTCACCCCTGTCCCCAGAGGG	261
TCCCCGCCTCCTGCCCTGCAGGG	262
GGCTGCCCTGGCTGTGCCGCAGG	263
TCCAAAAGCTGGCATTGTCAGGG	264
CAATGCCAGCTTTTGGATGACGG	265
CTGGGCCTGGCTAGCCTCAGAGG	266
GGTTCACCTTCAGTCTTTCATGG	267

	SEQ ID NO
AGGAGAAACCTATGGGGGGTGGG	268
TCTAAAAGACAGCCCAGCCCAGG	269
CGGGGATCCCGGGGACCCCGGGG	270
TCTTCTCCAGGGGAACAAAGGGG	271
ACTGACACAAAAAGTCAGCACGG	272
CCTGAAGAGTCAGGTCACCAAGG	273
GGAGGAGAGTCCCACCTGGAAGG	274
GGGCAGCCACCAGCCTGGAAAGG	275
AGCCCTATTTCTCTCTCTCTGG	276
GCCACCAGCCTGGAAAGGCACGG	277
TGACACCCTGAGTGTGGCCTCGG	278
AATTAATAGGACATGGGGAGGGG	279
GGCTCGGGGGGCCCCGGGGAGGG	280
TAATAGGACATGGGGAGGGGAGG	281
CTCTTCTCCAGGGGAACAAAGGG	282
TCGGGGCCCCCGGGTCCCCGGG	283
TCCCTGACAATGCCAGCTTTTGG	284
GAATTAATAGGACATGGGGAGGG	285
AATGAGGGGGTCAGAGACACAGG	286
GAAAACCTCCTTTATGGGGCCGG	287
CTTGGGAGGGCCAGGGACTTTGG	288
CCCCTGCAGGGCAGGAGGCGGGG	289
TCAGTGTCACTTCTTTTGGGGGG	290
ATCCCCGTTCTTCTTCCTCCTGG	291
CTTCCTCCTGGCCACAGTGTTGG	292
TGCAGTTGGGAGGGAACTCTTGG	293
TGGCTCGGGGGGCCCCGGGGAGG	294
CTGCAAAGGCCTGAAGAGTCAGG	295
CCGTGTGCGGAAGACGCCGCCGG	296
CCCCGGGGAGGGAGGATGCTCGG	297
TTCCAGGCTGGTGGCTGCCCTGG	298
AGGTCACCAAGGTGTCTGCATGG	299
GGCGGCCCTGCAGGGCAGGAGG	300

	SEQ ID NO
TCGGGGGGCCCCGGGGAGGGAGG	301
CCAAAAGAAGTGACACTGAAGGG	302
GGCCAGGAGGAAGAAGAACGGGG	303
GCCCAGGGTCCTGGTGGGGTAGG	304
GCTAGCCTCAGAGGAGGGAGTGG	305
GAGGGTCCTTTGCCCTCCAAGG	306
CCACCAGCCTGGAAAGGCACGGG	307
GATTTCTGTGGTATTGGGGTTGG	308
CTAGCCTCAGAGGAGGGAGTGGG	309
CCCGGGGTCCCCGGGATCCCCGG	310
CATGGGGTGGTAGGTGGAGTGGG	311
AATGCCAGCTTTTGGATGACGGG	312
GCCCCTGCAGGGCAGGAGGCGGG	313
GAGGAGAAACCTATGGGGGGTGG	314
AGAAGTGACACTGAAGGGCCTGG	315
CCTCCAGCAGATGCAGAGGAAGG	316
CCTCTTCTCCAGGGGAACAAAGG	317
CCGGGGTCCCCGGGATCCCCGGG	318
TAGCCTCAGAGGAGGGAGTGGGG	319
CAGAGGAAGGGGATGCAGTTGGG	320
CTCCAGCAGATGCAGAGGAAGGG	321
GATTCTGTTTTTCTCTGCCTGG	322
CTTCAGTGTCACTTCTTTTGGGG	323
CATAGAGTCCCTCTGGGGACAGG	324
GGACCCTCCTGAGGCCACAGCGG	325
CCATGGGGTGGTAGGTGGAGTGG	326
GACACCCTGAGTGTGGCCTCGGG	327
ATGCTTAGAAGTTTCCAGAAAGG	328
AGCTGGGCCCTGGAAGCCTGCGG	329
TACCACAGAAATCAGGTGCTTGG	330
ACCCAATACCACAGAAATCAGG	331
TTCTACAGGTAAAAAACTAAGG	332
GGCCCCTGCAGGGCAGGAGGCGG	333

	SEQ ID NO
CTCCCCGCCTCCTGCCCTGCAGG	334
TCTCTGACCCCCTCATTGTGGG	335
GGAGAAACCTATGGGGGGTGGGG	336
ACAGCCCAGCCCAGGACAGACGG	337
CCTGTATTTGAGGTTGGCCTGGG	338
AGCCAGGGCAGCCACCAGCCTGG	339
AGCCTCAGAGGAGGGAGTGGGGG	340
G TTCAGTGTTTCACTCATCCAGG	341
CTGACTCTTCAGGCCTTGCAGG	342
ATCCCCACTCCCTCCTCTGAGG	343
CCCAAAGAAGTGACACTGAAGG	344
TGGCCAGGAGGAAGAAGAACGGG	345
GGGAGGAAGGTTATGGGATGGGG	346
CCTGAGGCTTCCTGCACTCTAGG	347
TAGTTTTTTTACCTGTAGAATGG	348
AAGTGGGGTAGGGAACAAGGTGG	349
GGGTCACAGGAAGTGGGGTAGGG	350
CACCTGATTTCTGTGGTATTGGG	351
TTTGCAACTGGGTCTCATGTGGG	352
GAGAAACCTATGGGGGGTGGGGG	353
GAGGGAGGAGGGGTGGCTCGGGG	354
GCCTGTAATCCCACAAATGAGGG	355
GCAGAGGAAGGGGATGCAGTTGG	356
ACCTGATTTCTGTGGTATTGGGG	357
AAACCAGAGGAGAGAGAAATAGG	358
AACCAGAGGAGAGAGAAATAGGG	359
GAGGAGAGAGAAATAGGGCTTGG	360
CTGCAGGGCAGGAGGCGGGGAGG	361
AGGAAGGGGATGCAGTTGGGAGG	362
GGTATTGGGGTTGGAACCTGAGG	363
TTTTGCAACTGGGTCTCATGTGG	364
GGGAGGAGGGGTGGCTCGGGGGG	365
TCCCCTTCCTCTGCATCTGCTGG	366

	SEQ ID NO
AGGAAGTGGGGTAGGGAACAAGG	367
AGGGGAGGAAGGTTATGGGATGG	368
CCTGTAATCCCACAAATGAGGGG	369
AGTTTTTTTACCTGTAGAATGGG	370
AGGAAGAAGAACGGGGATGGGGG	371
TGCAGGGCAGGAGGCGGGGAGGG	372
GGAAGGGGATGCAGTTGGGAGGG	373
CAAAGTCACTGTGTAGATGAAGG	374
GTGGCCAGGAGGAAGAAGAACGG	375
TCCAGCAGATGCAGAGGAAGGGG	376
AGGGAGGAGGGGTGGCTCGGGGG	377
AACCTATGGGGGGTGGGGGTGGG	378
GGGGAGGGGAGGAAGGTTATGGG	379
CACCCACCCCAACCCCATAGG	380
AAACCTATGGGGGGTGGGGGTGG	381
GGAGGAAGAAGAACGGGGATGGG	382
GAGGAAGAAGAACGGGGATGGGG	383
GGAGGGAGGAGGGGTGGCTCGGG	384
GGTTGGCCTGGGCTACACAGGGG	385
AGAGGAGGGAGTGGGGGATTGGG	386
GAGGTTGGCCTGGGCTACACAGG	387
AGGTTGGCCTGGGCTACACAGGG	388
TGGGGAGGGGAGGAAGGTTATGG	389
GAAAGTAGAGGCAGGAGGGTTGG	390
GAGGAGGGAGTGGGGGATTGGGG	391
GGGGAGGAAGGTTATGGGATGGG	392
AGAGTGCTTGCCTAGAGTGCAGG	393
AGGAGGGAGTGGGGGATTGGGGG	394
CAGAGGAGGGAGTGGGGGATTGG	395
AGGAGGAAGAAGAACGGGGATGG	396
TTTTTTCCCCTGTGTAGCCCAGG	397
GGTGCTTGGAAGTAGAGGCAGG	398
CTTGGAAGTAGAGGCAGGAGGG	399

	SEQ ID NO
CTGTAATCCCACAAATGAGGGGG	400
AGGACATGGGGAGGGGAGGAAGG	401
TCATCTACACAGTGACTTTGAGG	402
GAGGGAGTGGGGGATTGGGGGGG	403
GGGAGGGAGGAGGGGTGGCTCGG	404
AGGGAGTGGGGGATTGGGGGGGG	405
AACAACAAAAACAAAACCAGAGG	406
CTATGGGGGGTGGGGGTGGGTGG	407
GGAGGGAGTGGGGGATTGGGGGG	408
TGCCTGTAATCCCACAAATGAGG	409
GAGTGGGGGATTGGGGGGGGGGG	410
AGGGCAGGAGGCGGGGAGGGAGG	411
CAGGAGGCGGGGAGGGAGGAGGG	412
AGGAGGCGGGGAGGGAGGAGGGG	413
GGAGTGGGGGATTGGGGGGGGGG	414
GGGAGTGGGGGATTGGGGGGGGG	415
GCAGGAGGCGGGGAGGGAGGAGG	416
AGGCGGGGAGGGAGGAGGGGTGG	417
GCTTGGAAAGTAGAGGCAGGAGG	418
GAGAGAGAGAGAGAGAGTGTGAGG	419

[00120] Следует понимать, что описание, конкретные примеры и данные, указывающие примерные варианты реализации, даны в качестве иллюстрации и не предназначены для ограничения данного изобретения. Различные изменения и модификации в рамках данных изобретений, включая объединение вариантов реализации полностью и частично, станут очевидными для специалиста в данной области техники из обсуждения, раскрытия и данных, содержащихся в данном документе, и, таким образом, считаются частью данных изобретений.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Клетка млекопитающего, содержащая первый стабильный сайт интеграции, расположенный в геномной безопасной гавани, и второй стабильный сайт интеграции, который не расположен в геномной безопасной гавани, причем первый стабильный сайт интеграции содержит первый репортерный ген, кодирующий первый репортерный белок, и второй стабильный сайт интеграции содержит второй репортерный ген, кодирующий второй репортерный белок, при этом первый репортерный белок и второй репортерный белок различны.

2. Клетка млекопитающего по п. 1, отличающаяся тем, что первый и второй стабильные сайты интеграции содержат сайты узнавания рекомбиназы.

3. Клетка млекопитающего по п. 1, отличающаяся тем, что первый репортерный ген находится под контролем промотора SV40.

4. Клетка млекопитающего по п. 1, отличающаяся тем, что второй репортерный ген находится под контролем промотора SV40.

5. Клетка млекопитающего по п. 1, отличающаяся тем, что первый репортерный ген кодирует флуоресцентный белок.

6. Клетка млекопитающего по п. 1, отличающаяся тем, что второй репортерный ген кодирует флуоресцентный белок.

7. Клетка млекопитающего по п. 1, отличающаяся тем, что дополнительно содержит полинуклеотид, кодирующий репрессор под контролем промотора CMV.

8. Клетка млекопитающего по п. 1, отличающаяся тем, что представляет собой клетку человека.

9. Клетка млекопитающего по п. 8, отличающаяся тем, что клетка человека представляет собой амниотическую эпителиальную клетку человека.

10. Клетка млекопитающего по п. 8, отличающаяся тем, что клетка человека представляет собой клетку НЕК293.

11. Клетка млекопитающего по п. 1, отличающаяся тем, что представляет собой клетку СНО.

12. Клетка млекопитающего по п. 2, отличающаяся тем, что сайты узнавания для рекомбиназы представляют собой I_{ox}-сайты.

13. Клетка млекопитающего по п. 1, отличающаяся тем, что полинуклеотид, кодирующий представляющий интерес белок, вставлен в первый стабильный сайт интеграции или второй стабильный сайт интеграции.

14. Клетка млекопитающего по п. 1, отличающаяся тем, что второй стабильный сайт интеграции расположен во второй геномной безопасной гавани, которая отличается от первой геномной безопасной гавани.

15. Клетка млекопитающего по п. 1, отличающаяся тем, что второй стабильный сайт интеграции расположен в области, которая не является геномной безопасной гаванью.

16. Клетка млекопитающего, содержащая первый стабильный сайт интеграции,

расположенный в геномной безопасной гавани, и второй стабильный сайт интеграции, который не расположен в геномной безопасной гавани, причем первый стабильный сайт интеграции содержит первый полинуклеотид, кодирующий первый белок, и второй стабильный сайт интеграции содержит второй полинуклеотид, кодирующий второй белок.

17. Клетка млекопитающего по п. 16, отличающаяся тем, что первый белок представляет собой вирусный белок.

18. Клетка млекопитающего по п. 16, отличающаяся тем, что второй белок представляет собой вирусный белок.

19. Клетка млекопитающего по п. 16, отличающаяся тем, что первый белок представляет собой вирусный белок аденоассоциированного вируса.

20. Клетка млекопитающего по п. 17, отличающаяся тем, что второй белок представляет собой вирусный белок аденоассоциированного вируса.

21. Клетка млекопитающего по п. 17, отличающаяся тем, что первый белок представляет собой аденовирусный белок.

22. Клетка млекопитающего по п. 17, отличающаяся тем, что второй белок представляет собой аденовирусный белок.

23. Клетка млекопитающего по п. 16, отличающаяся тем, что содержит полинуклеотид, кодирующий аденоассоциированный вирусный белок, и полинуклеотид, кодирующий аденовирусный белок.

24. Клетка млекопитающего по п. 16, отличающаяся тем, что второй стабильный сайт интеграции расположен во второй геномной безопасной гавани, которая отличается от первой геномной безопасной гавани.

25. Клетка млекопитающего по п. 16, отличающаяся тем, что второй стабильный сайт интеграции расположен в области, которая не является геномной безопасной гаванью.

26. Клетка млекопитающего по п. 16, отличающаяся тем, что представляет собой клетку человека.

27. Клетка млекопитающего по п. 16, отличающаяся тем, что представляет собой клетку СНО.

28. Клетка млекопитающего, содержащая первый стабильный сайт интеграции, расположенный в геномной безопасной гавани, и второй стабильный сайт интеграции, который не расположен в геномной безопасной гавани, причем первый стабильный сайт интеграции содержит полинуклеотид, кодирующий первый репортерный ген, кодирующий первый репортерный белок и второй стабильный сайт интеграции содержат полинуклеотид, кодирующий Cas9, и полинуклеотид, кодирующий второй репортерный ген, кодирующий второй репортерный белок, при этом первый репортерный белок и второй репортерный белок различны.

29. Клетка млекопитающего по п. 28, отличающаяся тем, что второй стабильный сайт интеграции дополнительно содержит ген маркера селекции и внутренний сайт входа в рибосому (IRES).

30. Клетка млекопитающего по п. 28, отличающаяся тем, что первый и второй стабильные сайты интеграции содержат сайты узнавания для рекомбиназы.

31. Клетка млекопитающего по п. 28, отличающаяся тем, что представляет собой амниотическую эпителиальную клетку человека.

32. Клетка млекопитающего по п. 28, отличающаяся тем, что представляет собой клетку НЕК293.

33. Клетка млекопитающего по п. 28, отличающаяся тем, что представляет собой клетку ВНК.

34. Клетка млекопитающего по п. 28, отличающаяся тем, что представляет собой клетку СНО.

35. Способ получения по меньшей мере одного представляющего интерес белка, включающий:

(а) предоставление клетки млекопитающего, содержащей первый стабильный сайт интеграции, расположенный в первой геномной безопасной гавани, и второй стабильный сайт интеграции, который не расположен в первой геномной безопасной гавани, причем первый стабильный сайт интеграции содержит первый репортерный ген, кодирующий первый репортерный белок и второй стабильный сайт интеграции содержит второй репортерный ген, кодирующий второй репортерный белок, при этом первый репортерный белок и второй репортерный белок различны, и при этом первый и второй стабильные сайты интеграции содержат сайты узнавания для рекомбиназы;

(б) введение полинуклеотида, кодирующего представляющий интерес белок, в стабильный сайт интеграции путем опосредованной рекомбиназой замены кассеты, и

(с) культивирование клетки млекопитающего в условиях, которые позволяют экспрессировать полинуклеотид, кодирующий представляющий интерес полинуклеотид.

36. Способ по п. 35, отличающийся тем, что клетка млекопитающего представляет собой амниотическую эпителиальную клетку человека.

37. Способ по п. 35, отличающийся тем, что клетка млекопитающего представляет собой клетку НЕК293.

38. Способ по п. 35, отличающийся тем, что клетка млекопитающего представляет собой клетку СНО.

39. Способ по п. 35, отличающийся тем, что первый стабильный сайт интеграции содержит первый полинуклеотид, кодирующий первый белок, и второй стабильный сайт интеграции содержит второй полинуклеотид, кодирующий второй белок.

40. Способ по п. 39, отличающийся тем, что первый белок представляет собой вирусный белок.

41. Способ по п. 39, отличающийся тем, что второй белок представляет собой вирусный белок.

42. Способ по п. 35, отличающийся тем, что клетка млекопитающего содержит полинуклеотид, кодирующий аденоассоциированный вирусный белок, и полинуклеотид, кодирующий аденовирусный белок.

43. Способ по п. 35, отличающийся тем, что второй стабильный сайт интеграции расположен во второй геномной безопасной гавани, которая отличается от первой геномной безопасной гавани.

44. Способ создания клетки млекопитающего с множеством стабильных сайтов интеграции, включающий:

(А) предоставление клетки млекопитающего, содержащей первую ДНК-кассету, содержащую в порядке от 5' до 3' полинуклеотид, кодирующий первый сайт lox, промотор, ген маркера селекции, кодирующий белок-маркер селекции, IRES, первый репортерный ген, кодирующий первый репортерный белок, промотор, функционально связанный с оператором, ген Cas9 и второй lox-сайт;

(В) интеграцию второй ДНК-кассеты, содержащей в порядке от 5' к 3' полинуклеотид, содержащий первое плечо гомологии геномной безопасной гавани, содержащее целевой сайт sgPНК CRISPR, третий сайт lox, второй репортерный ген, кодирующий второй репортерный белок, четвертый сайт lox и второе плечо гомологии геномной безопасной гавани, содержащее целевой сайт sgPНК CRISPR, при этом первый сайт lox, второй сайт lox, третий сайт lox и четвертый сайт lox различны, при этом первое и второе направляющие плечи могут содержать область с изменениями, и при этом второй репортерный белок отличается от первого репортерного белка;

(С) замену первой ДНК-кассеты на третью ДНК-кассету, при этом третья ДНК-кассета содержит в порядке от 5' к 3' полинуклеотид, кодирующий первый сайт lox, промотор, третий репортерный ген, кодирующий третий репортерный белок, и второй сайт lox, при этом третий репортерный белок отличается от второго репортерного белка, тем самым обеспечивая клетке млекопитающего множество стабильных сайтов интеграции.

45. Способ по п. 44, отличающийся тем, что изменения предотвращают воссоздание целевого сайта.

46. Способ по п. 44, отличающийся тем, что клетка млекопитающего представляет собой клетку НЕК293.

47. Способ по п. 44, отличающийся тем, что клетка млекопитающего представляет собой клетку СНО.

48. Способ по п. 44, отличающийся тем, что клетка на стадии (А) дополнительно содержит полинуклеотид, кодирующий репрессор под контролем промотора CMV.

49. Способ по п. 44, отличающийся тем, что клетка на стадии (В) дополнительно содержит полинуклеотид, кодирующий репрессор под контролем промотора CMV.

50. Способ по п. 44, отличающийся тем, что клетка на стадии (С) дополнительно содержит полинуклеотид, кодирующий репрессор под контролем промотора CMV.

51. Клетка млекопитающего, содержащая модифицированный геном, причем геном модифицирован путем вставки по меньшей мере трех ДНК-кассет в различные области генома, при этом модифицированный геном содержит

(1) первую последовательность дезоксирибонуклеиновой кислоты, которая на по

меньшей мере 90% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 1 и 2 до модификации;

(2) вторую последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 90% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и

(3) третью последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 90% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации,

при этом первая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки первой ДНК-кассеты, вторая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки второй ДНК-кассеты, и третья последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки третьей ДНК-кассеты.

52. Клетка млекопитающего по п. 51, отличающаяся тем, что

(1) первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 95% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 1 и 2 до модификации;

(2) вторая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 95% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и

(3) третья последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 95% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации.

53. Клетка млекопитающего по п. 51, отличающаяся тем, что

(1) первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 98% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 1 и 2 до модификации;

(2) вторая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 98% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и

(3) третья последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 98% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации.

54. Клетка млекопитающего по п. 51, отличающаяся тем, что

(1) первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 1 и 2 до модификации;

(2) вторая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и

(3) третья последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации.

55. Клетка млекопитающего по любому из пп. 51-54, отличающаяся тем, что

(a) первая ДНК-кассета содержит промотор и по меньшей мере один, выбранный из группы, состоящей из маркерного гена селекции и репортерного гена;

(b) вторая ДНК-кассета содержит промотор и по меньшей мере один, выбранный из группы, состоящей из маркерного гена селекции и репортерного гена; и

(c) третья ДНК-кассета содержит промотор и по меньшей мере один, выбранный из группы, состоящей из маркерного гена селекции и репортерного гена.

56. Клетка млекопитающего по п. 55, отличающаяся тем, что

(a) первая ДНК-кассета содержит промотор, маркерный ген селекции и репортерный ген;

(b) вторая ДНК-кассета содержит промотор, маркерный ген селекции и репортерный ген; и

(c) третья ДНК-кассета содержит промотор, маркерный ген селекции и репортерный ген.

57. Клетка млекопитающего по п. 55 или 56, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции.

58. Клетка млекопитающего по п. 57, отличающаяся тем, что представляющий интерес ген вставлен в стабильный сайт интеграции.

59. Клетка млекопитающего по п. 57, отличающаяся тем, что представляющий интерес ген кодирует представляющий интерес полипептид, выбранный из группы, состоящей из антител, цепей антител, рецепторов, Fc-содержащих белков, белков-ловушек, ферментов, факторов, репрессоров, активаторов, лигандов, репортерных белков, белков селекции, белковых гормонов, белковых токсинов, структурных белков, белков хранения, транспортных белков, нейротрансмиттеров и белков сокращения.

60. Клетка млекопитающего по любому из пп. 51-59, отличающаяся тем, что представляет собой клетку человека, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 90% идентична SEQ ID NO: 1.

61. Клетка млекопитающего по любому из пп. 51-59, отличающаяся тем, что представляет собой клетку CHO, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 90% идентична SEQ ID NO: 2.

62. Клетка млекопитающего по п. 60, отличающаяся тем, что представляет собой клетку человека, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 95% идентична SEQ ID NO: 1.

63. Клетка млекопитающего по п. 61, отличающаяся тем, что представляет собой клетку CHO, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 95% идентична SEQ ID NO: 2.

64. Клетка млекопитающего по п. 62, отличающаяся тем, что представляет собой клетку человека, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 98% идентична SEQ ID NO: 1.

65. Клетка млекопитающего по п. 63, отличающаяся тем, что представляет собой клетку СНО, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 98% идентична SEQ ID NO: 2.

66. Клетка млекопитающего по п. 64, отличающаяся тем, что представляет собой клетку человека, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 99% идентична SEQ ID NO: 1.

67. Клетка млекопитающего по п. 65, отличающаяся тем, что представляет собой клетку СНО, и первая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 99% идентична SEQ ID NO: 2.

68. Клетка млекопитающего по любому из пп. 61, 63, 65 или 67, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции, полученный с применением направляющей последовательности, выбранной из группы, состоящей из SEQ ID NO: с 13 по 419.

69. Клетка млекопитающего по любому из пп. 61, 63, 65 или 67, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции, полученный с применением направляющей последовательности, которая комплементарна целевым последовательностям в SEQ ID NO:2 в нуклеотидном положении в диапазонах, выбранных из группы, состоящей из: (a) от 1 до 2000; (b) от 2001 до 4000; (c) от 4001 до 6000; (d) от 6001 до 8000; (e) от 8001 до 10000; (f) от 10001 до 12000; (g) от 12001 до 14000; (h) от 14001 до 16000; (i) от 16001 до 18000; (j) от 18001 до 20000; (k) от 20001 до 22000; (l) от 22001 до 24000; (m) от 24001 до 26000; (n) от 26001 до 28000; (o) от 28001 до 30000; (p) от 30001 до 32000; (q) от 32001 до 34000; (r) от 34001 до 36000; (s) от 36001 до 38000; (t) от 38001 до 40000; (u) от 40001 до 42000; и (v) от 42001 до 44232.

70. Клетка млекопитающего, содержащая модифицированный геном, при этом модифицированный геном содержит последовательность дезоксирибонуклеиновой кислоты, содержащую AAVS1-подобную область, модифицированную вставкой по меньшей мере одной ДНК-кассеты, и

при этом направляющая последовательность выбрана из группы, состоящей из SEQ ID NO: с 13 по 419, которая комплементарна смысловой или антисмысловой цепи AAVS1-подобной области.

71. Клетка млекопитающего по п. 70, дополнительно содержащая вторую последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 90% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и

третью последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 90% идентична по меньшей мере одной, выбранной из группы, состоящей

из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации,

при этом первая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки первой ДНК-кассеты, вторая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки второй ДНК-кассеты, и третья последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки третьей ДНК-кассеты.

72. Клетка млекопитающего по п. 71, отличающаяся тем, что вторая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 90-99%, на 95-99% или на 98-99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из последовательности, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и третья последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 90-99%, на 95-99% или на 98-99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации.

73. Клетка млекопитающего по любому из пп. 70-72, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции, полученный с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 50-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

74. Клетка млекопитающего по п. 73, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции, полученный с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 75%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

75. Клетка млекопитающего по п. 74, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт

интеграции, полученный с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 85%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

76. Клетка млекопитающего по п. 75, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции, полученный с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 90%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

77. Клетка млекопитающего по п. 76, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции, полученный с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 95%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

78. Клетка млекопитающего по п. 77, отличающаяся тем, что первая последовательность дезоксирибонуклеиновой кислоты содержит стабильный сайт интеграции, полученный с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 98%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до

8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

79. Клетка млекопитающего, содержащая модифицированный геном, причем модифицированный геном содержит стабильный сайт интеграции в AAVS1-подобной области, при этом стабильный сайт интеграции получен с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 50%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

80. Клетка млекопитающего по п. 79, отличающаяся тем, что модифицированный геном содержит стабильный сайт интеграции в AAVS1-подобной области, при этом стабильный сайт интеграции получен с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 75%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

81. Клетка млекопитающего по п. 80, отличающаяся тем, что модифицированный геном содержит стабильный сайт интеграции в AAVS1-подобной области, при этом стабильный сайт интеграции получен с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 85%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до

28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

82. Клетка млекопитающего по п. 81, отличающаяся тем, что модифицированный геном содержит стабильный сайт интеграции в AAVS1-подобной области, при этом стабильный сайт интеграции получен с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 90%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

83. Клетка млекопитающего по п. 82, отличающаяся тем, что модифицированный геном содержит стабильный сайт интеграции в AAVS1-подобной области, при этом стабильный сайт интеграции получен с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 95%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

84. Клетка млекопитающего по п. 83, отличающаяся тем, что модифицированный геном содержит стабильный сайт интеграции в AAVS1-подобной области, при этом стабильный сайт интеграции получен с применением направляющей последовательности, которая комплементарна по меньшей мере одной целевой последовательности, на по меньшей мере 98%-99% идентичной с SEQ ID NO:2 в положениях нуклеотидов:

(a) от 1 до 2000; или (b) от 2001 до 4000; или (c) от 4001 до 6000; или (d) от 6001 до 8000; или (e) от 8001 до 10000; или (f) от 10001 до 12000; или (g) от 12001 до 14000; или (h) от 14001 до 16000; или (i) от 16001 до 18000; или (j) от 18001 до 20000; или (k) от 20001 до 22000; или (l) от 22001 до 24000; или (m) от 24001 до 26000; или (n) от 26001 до 28000; или (o) от 28001 до 30000; или (p) от 30001 до 32000; или (q) от 32001 до 34000; или (r) от 34001 до 36000; или (s) от 36001 до 38000; или (t) от 38001 до 40000; или (u) от 40001 до 42000; или (v) от 42001 до 44232.

85. Клетка млекопитающего по любому из пп. 78-84, дополнительно содержащая

вторую последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 90% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и

третью последовательность дезоксирибонуклеиновой кислоты, которая на по меньшей мере 90% идентична по меньшей мере одной, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации,

при этом первая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки первой ДНК-кассеты, вторая последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки второй ДНК-кассеты, и третья последовательность дезоксирибонуклеиновой кислоты модифицируется путем вставки третьей ДНК-кассеты.

86. Клетка млекопитающего по п. 85, отличающаяся тем, что вторая последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 90-99%, на 95-99% или на 98-99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из одной, выбранной из группы, состоящей из SEQ ID NO: от 5 до 10 до модификации; и третья последовательность дезоксирибонуклеиновой кислоты на по меньшей мере 90-99%, на 95-99% или на 98-99% идентична по меньшей мере одной последовательности, выбранной из группы, состоящей из SEQ ID NO: 11 и SEQ ID NO: 12 до модификации.

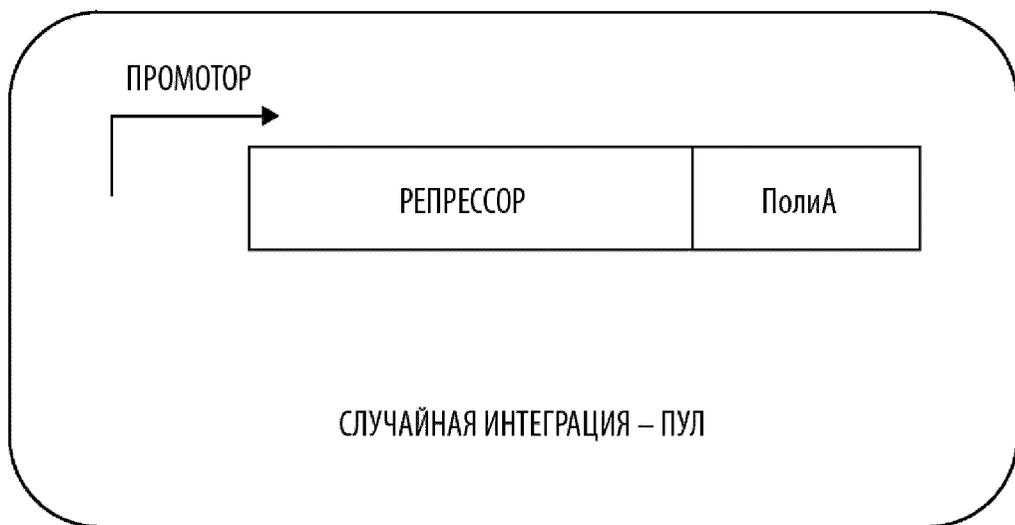
87. Способ получения белка, включающий стадии:

- (1) культивирования клеток млекопитающих по пп. 51-86; и
- (2) сбор белка.

88. Клетка по любому из вышеприведенных способов.

89. Способ, применяющий любую из вышеприведенных клеток.

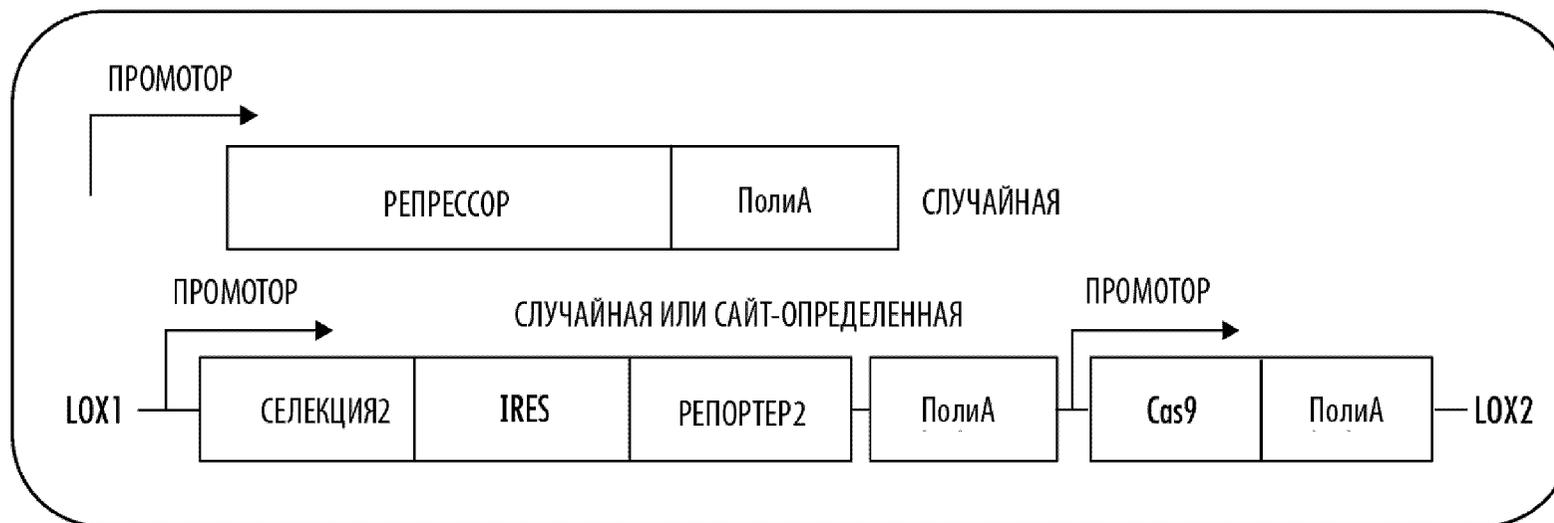
По доверенности



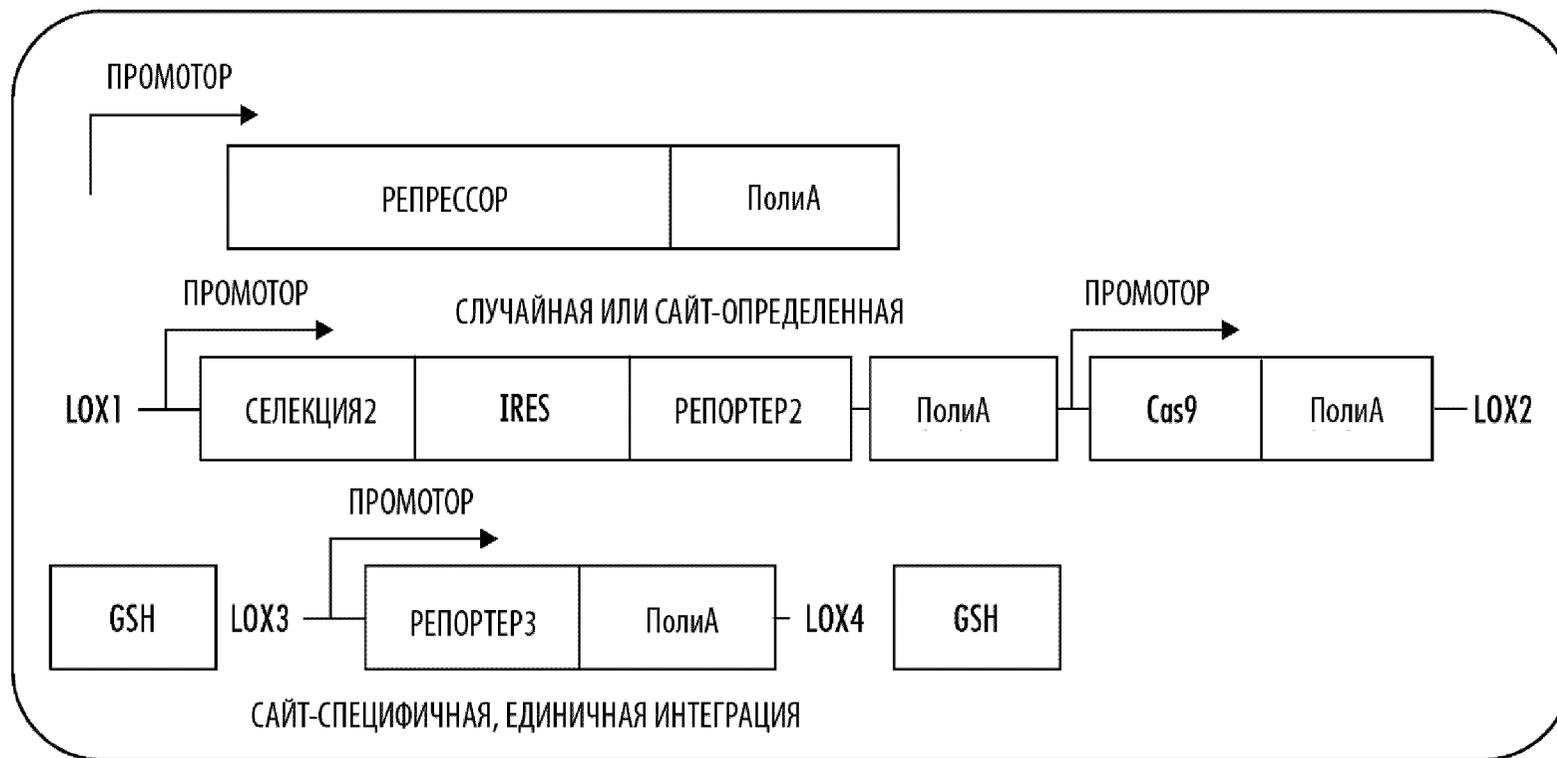
Фиг. 1



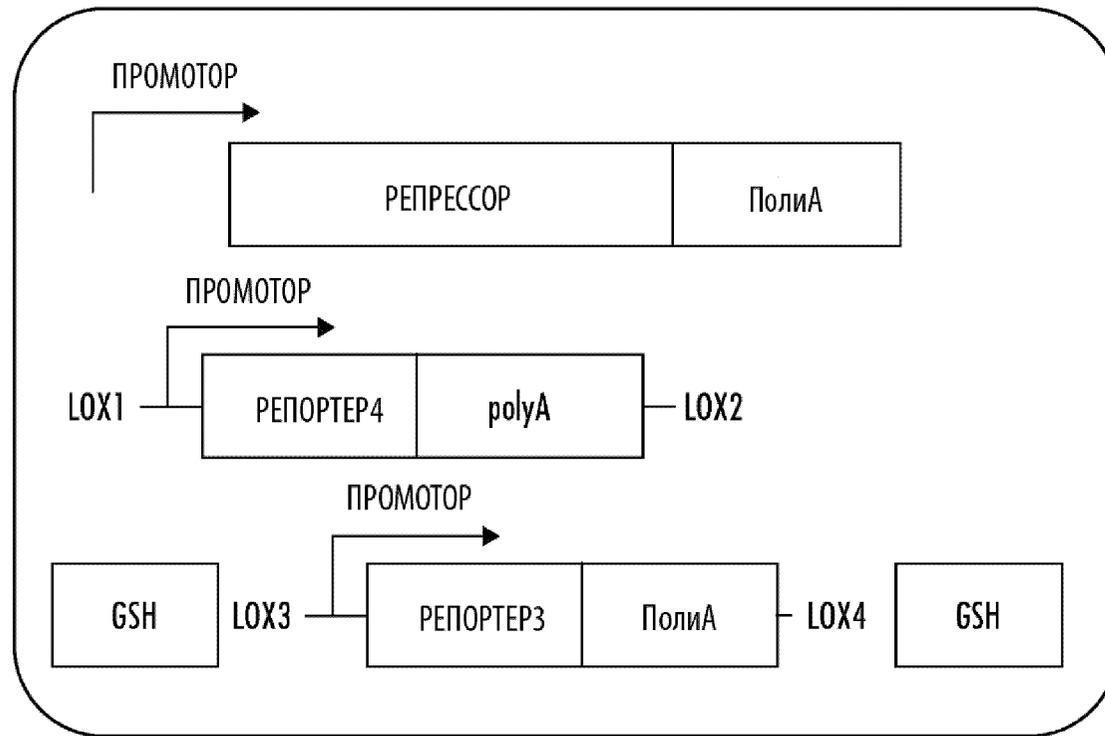
Фиг. 2



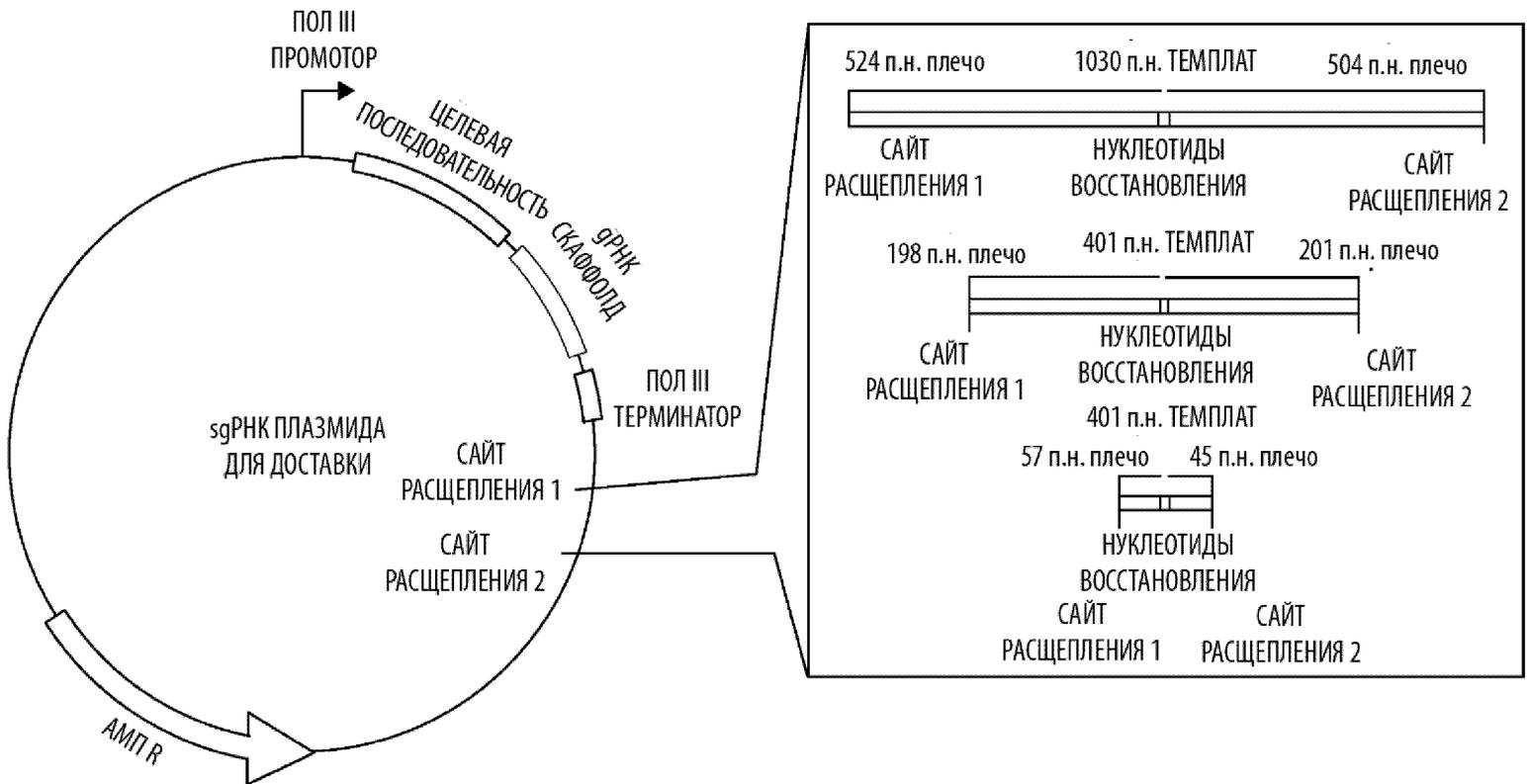
Фиг. 3



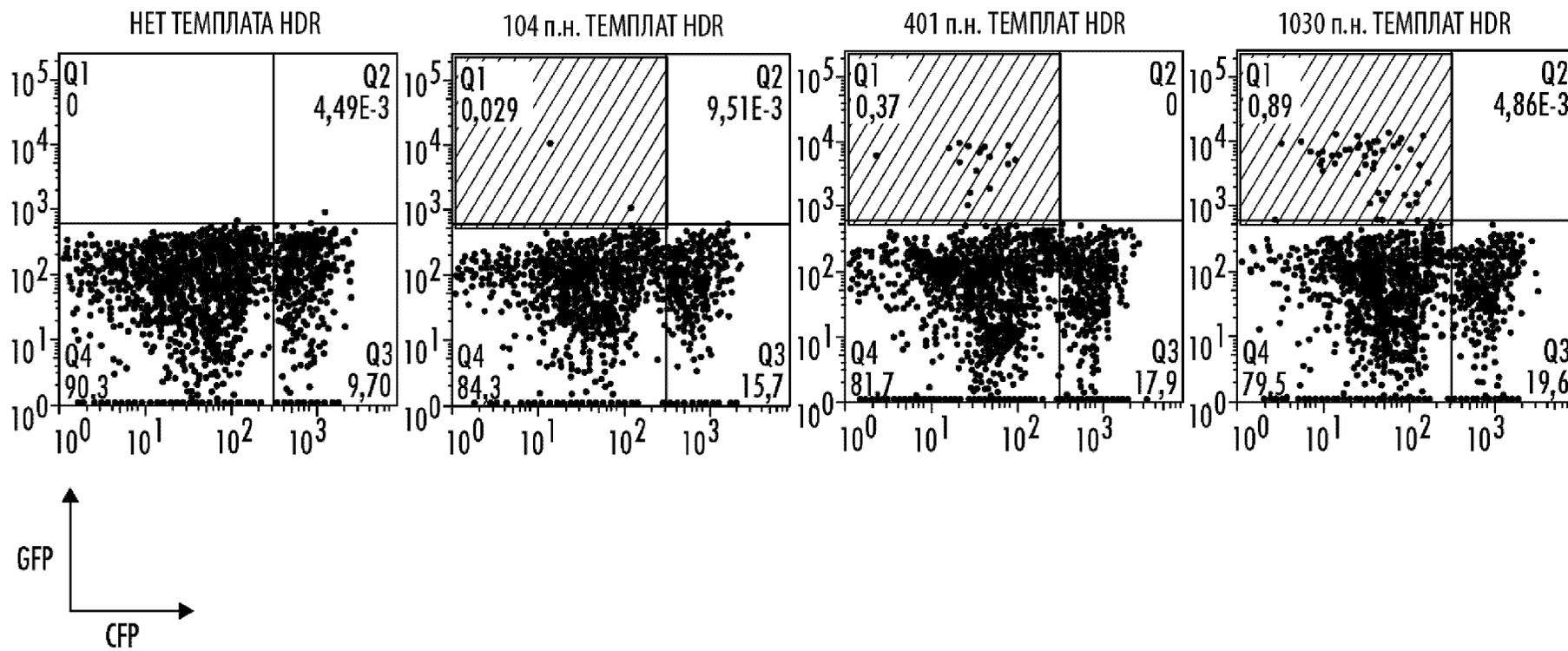
Фиг. 4



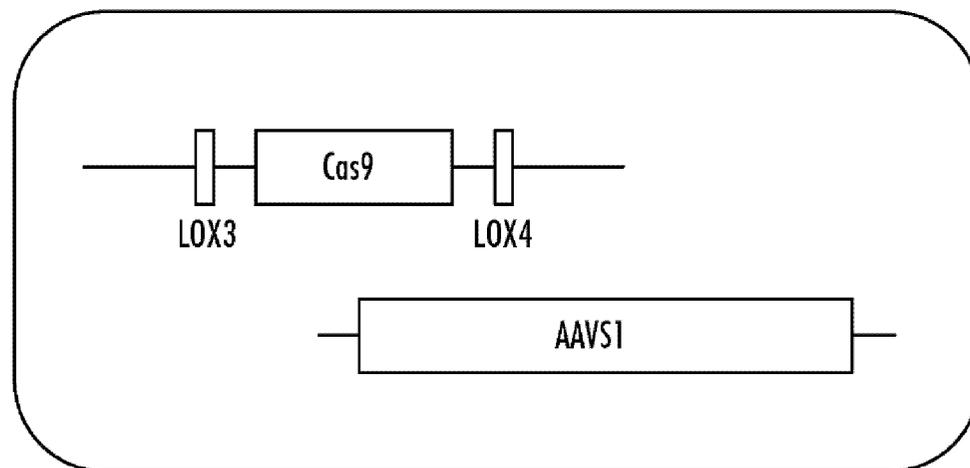
Фиг. 5



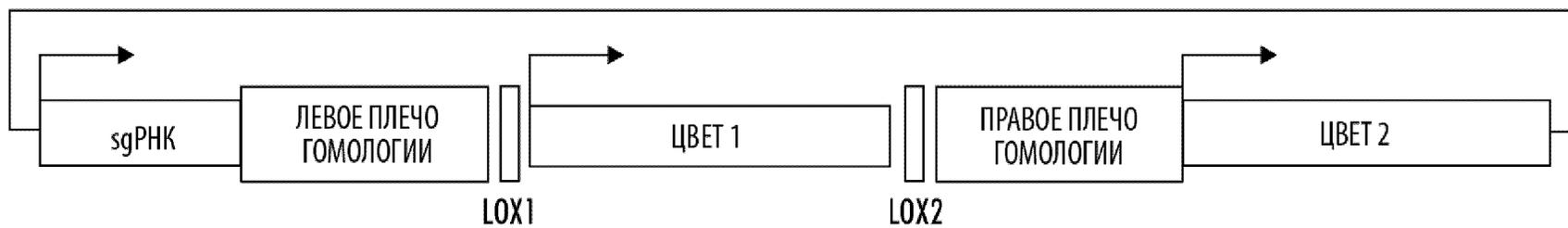
Фиг. 6



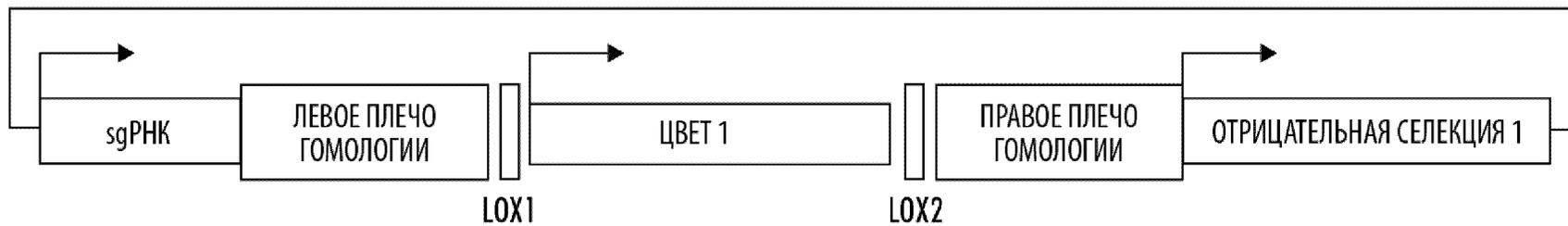
Фиг. 7



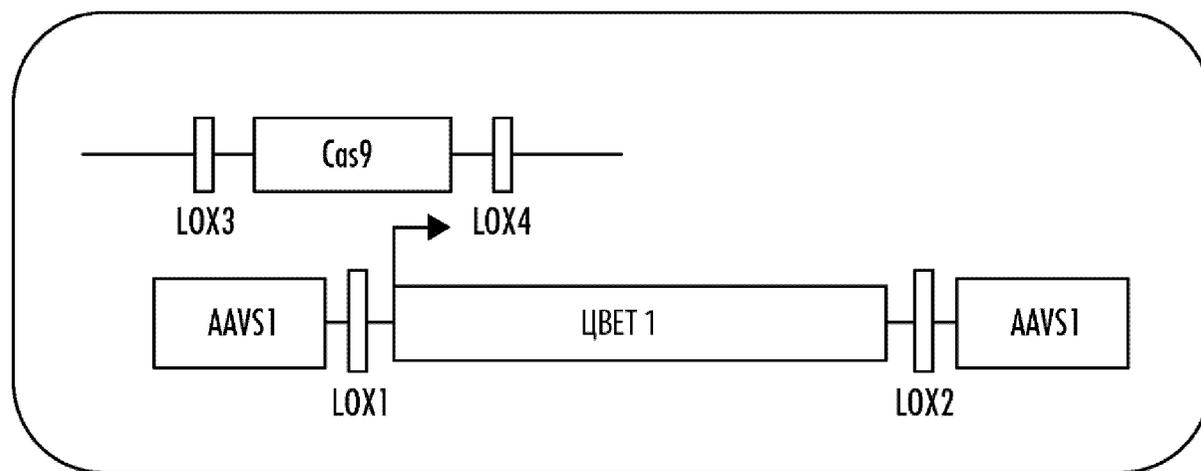
Фиг. 8



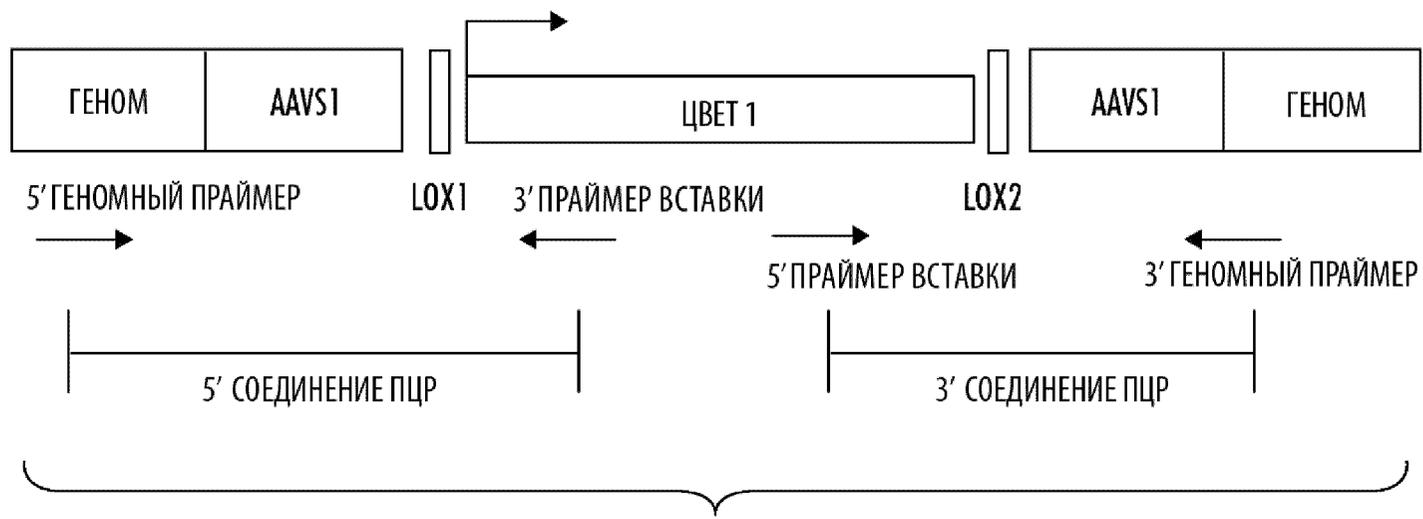
Фиг. 9А



Фиг. 9В

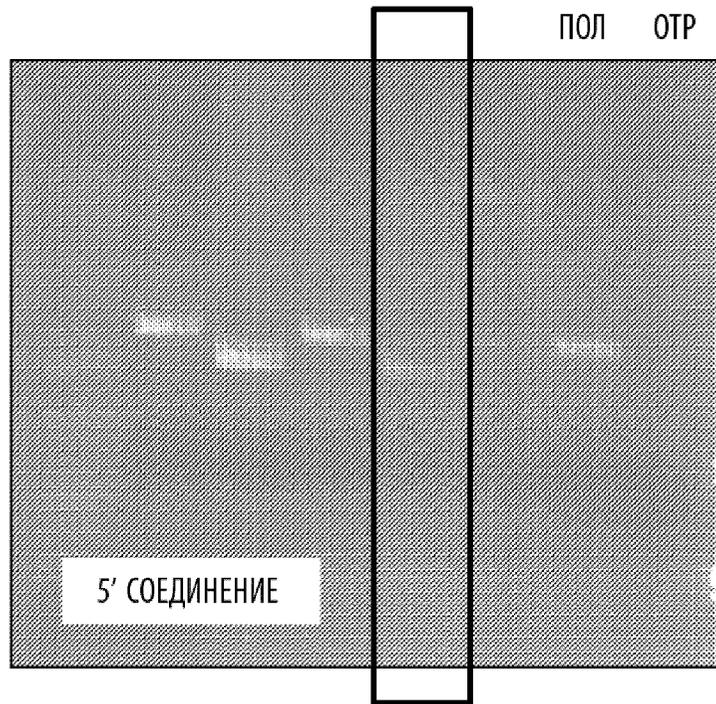


Фиг. 10

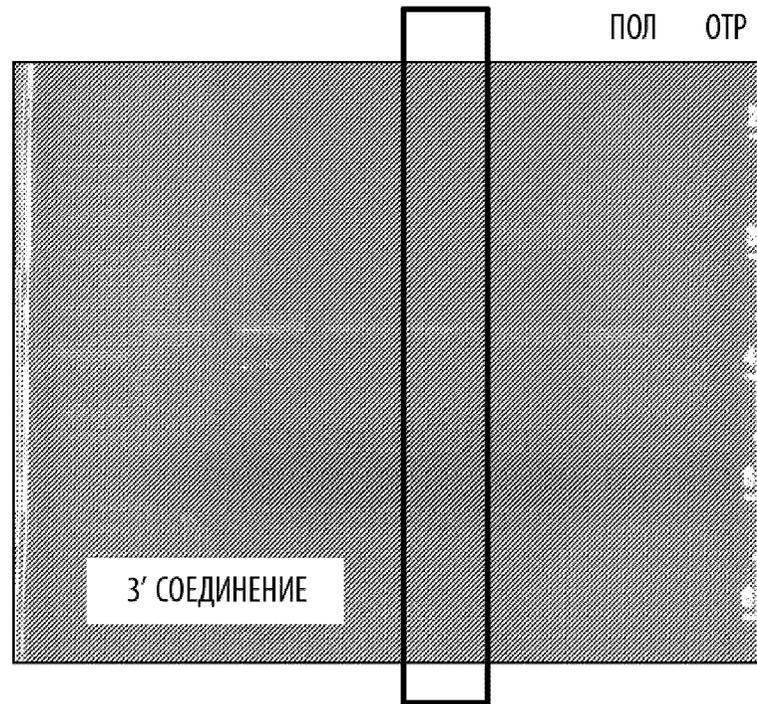


Фиг. 11

СТАБИЛЬНЫЕ Cas9
НАПРАВЛЕННЫЕ КЛЕТКИ

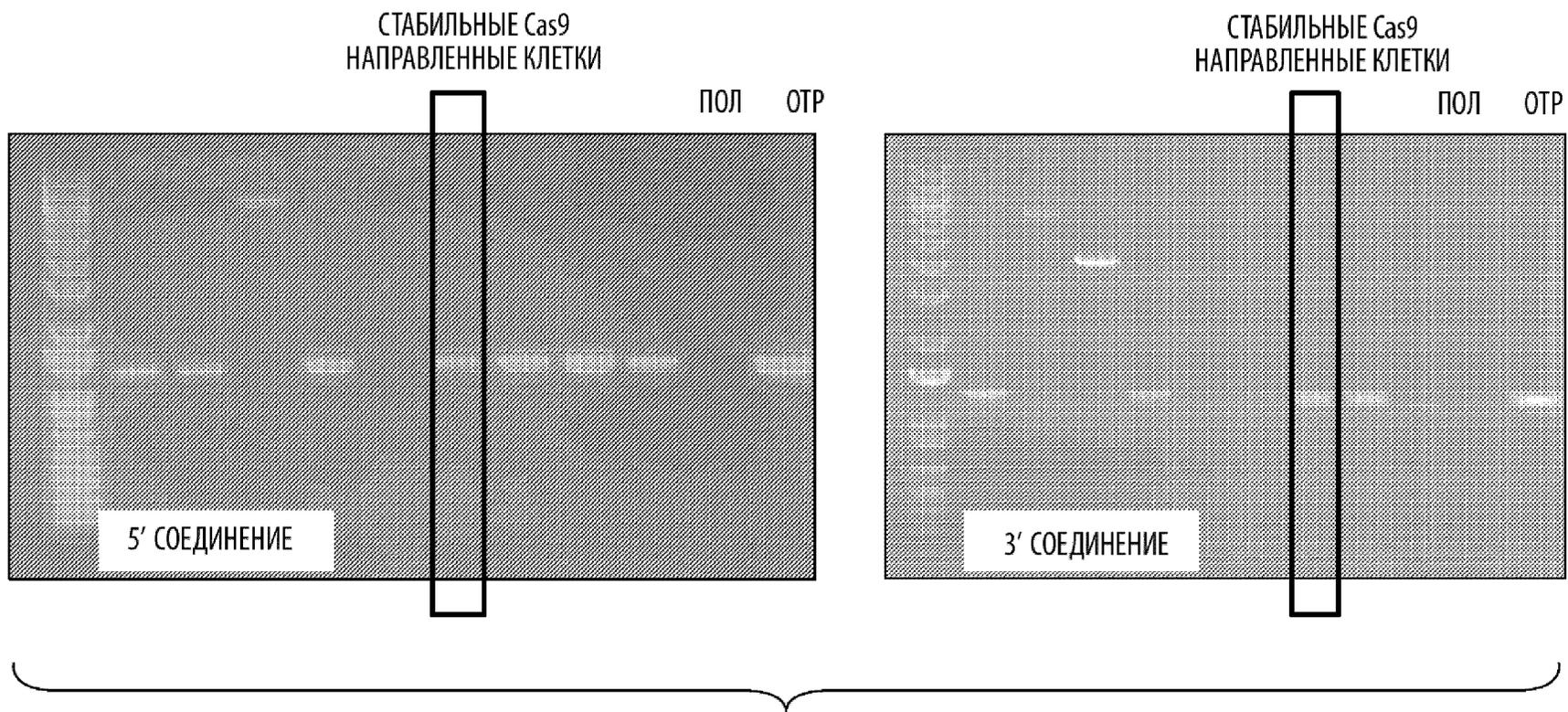


СТАБИЛЬНЫЕ Cas9
НАПРАВЛЕННЫЕ КЛЕТКИ

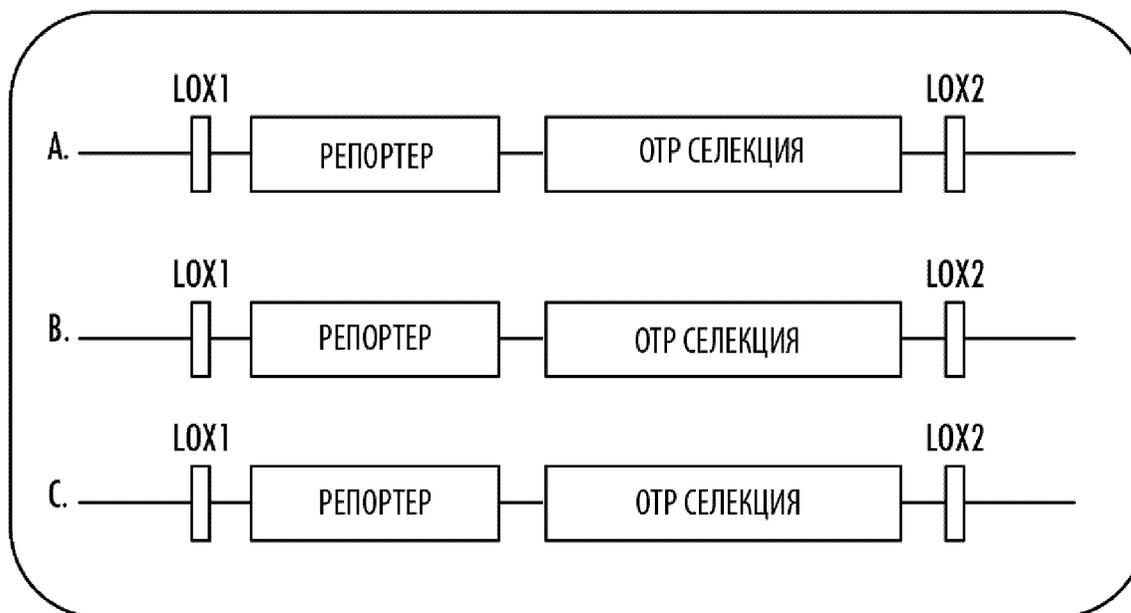


12/18

Фиг. 12

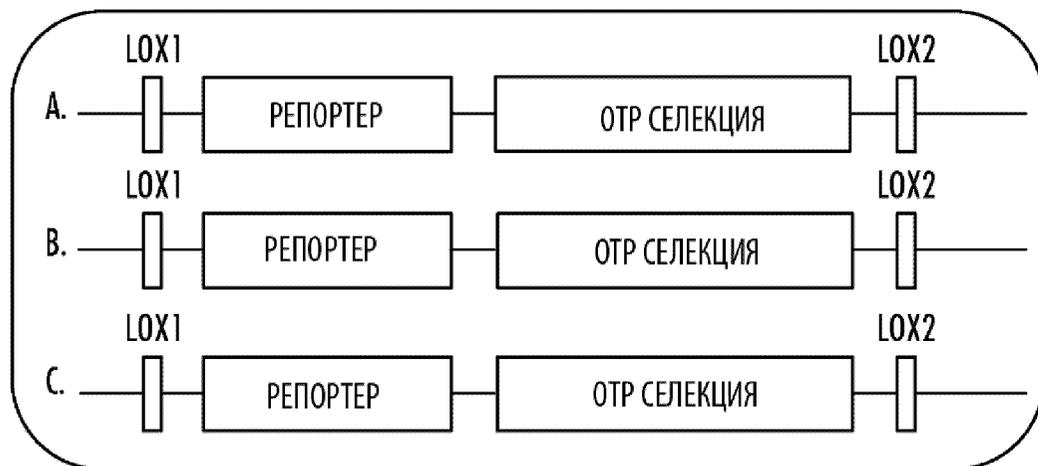


Фиг. 13

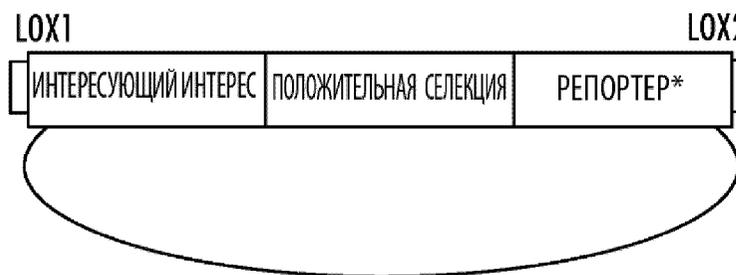
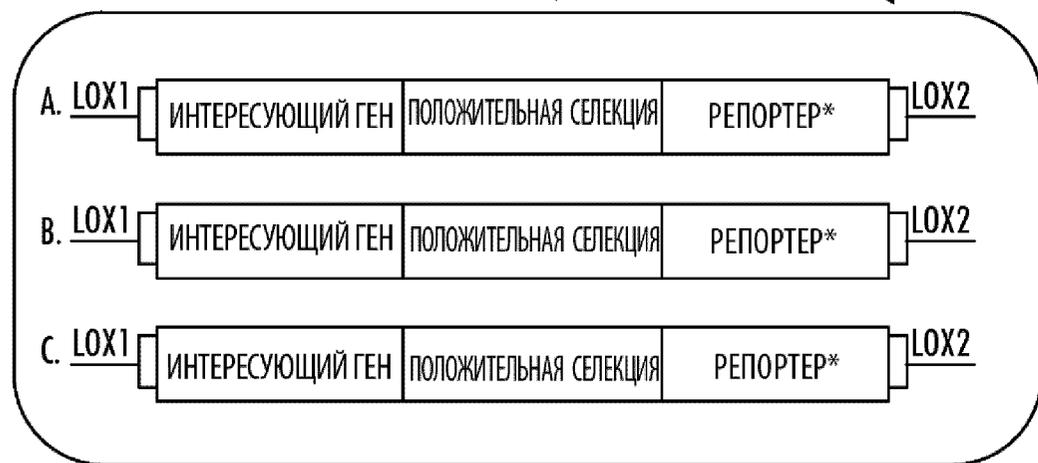


Фиг. 14

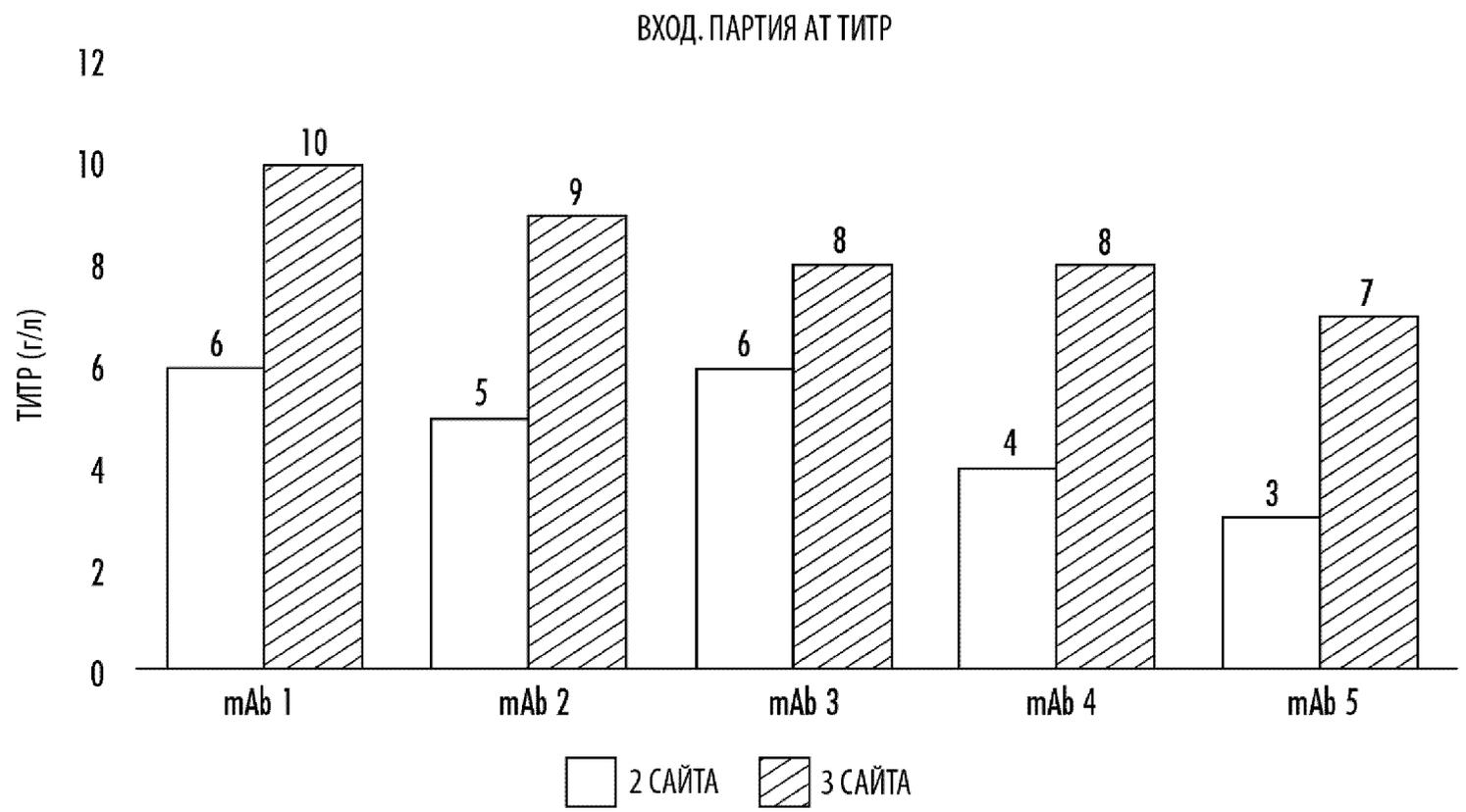
КЛЕТКА-ХОЗЯИН



ПРОИЗВОДСТВО БЕЛКА ИНТЕРЕСА КЛЕТКИ

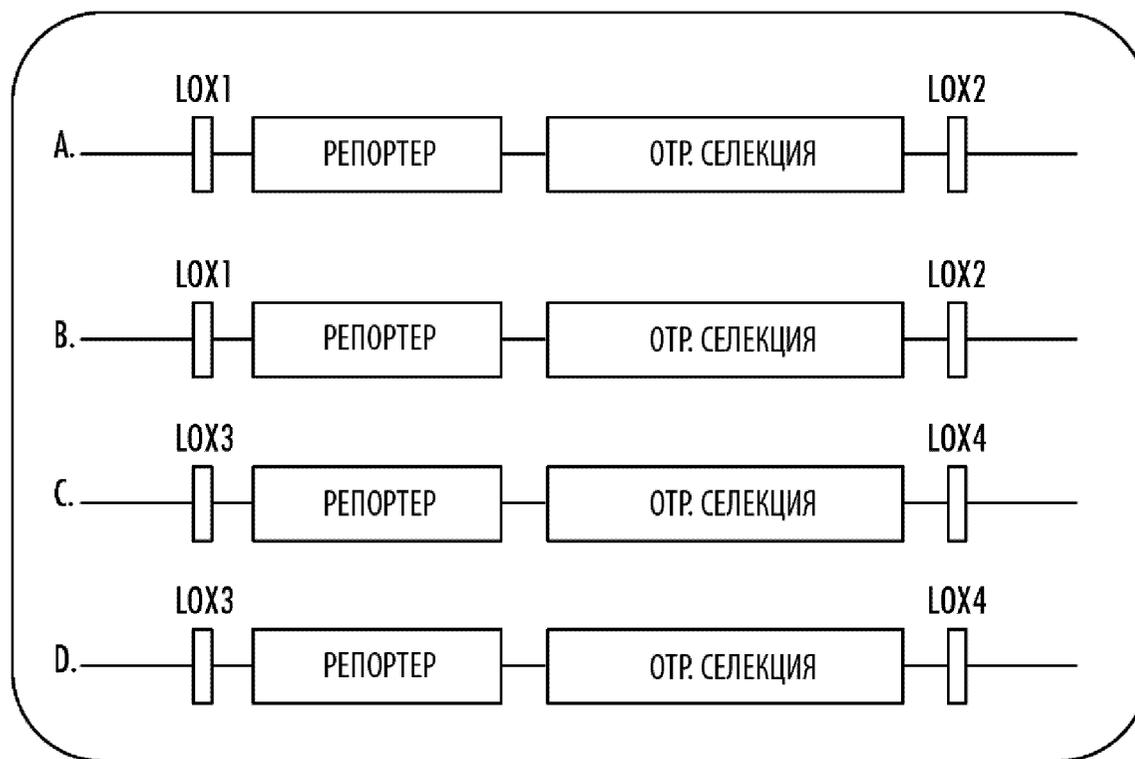


Фиг. 15



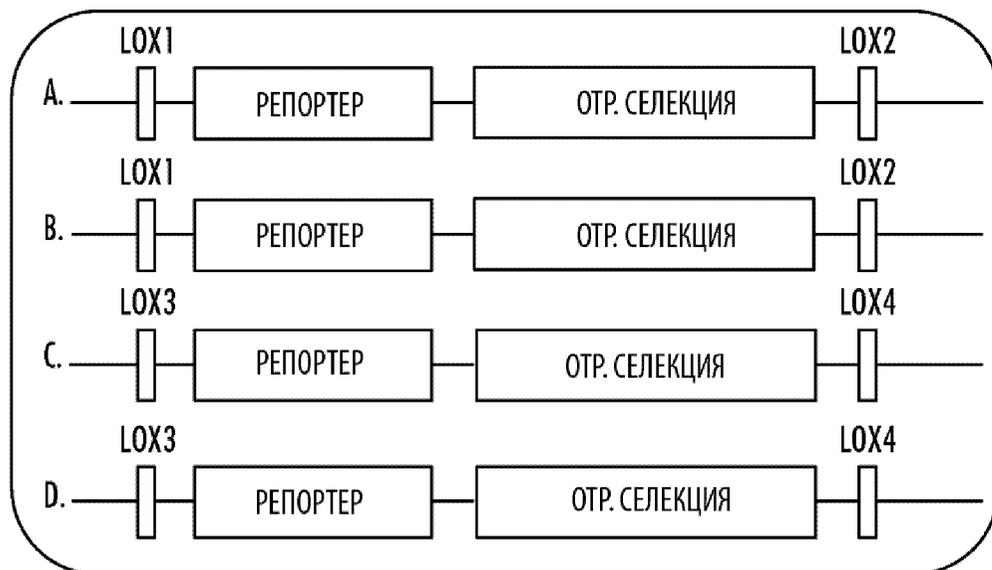
16/18

Фиг. 16

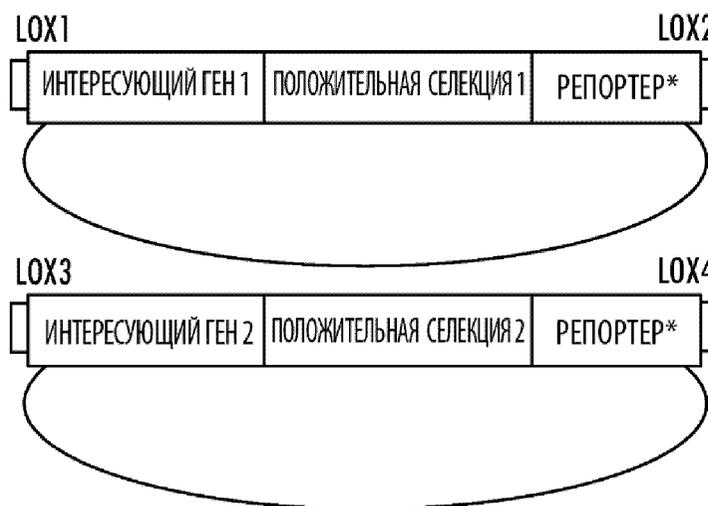
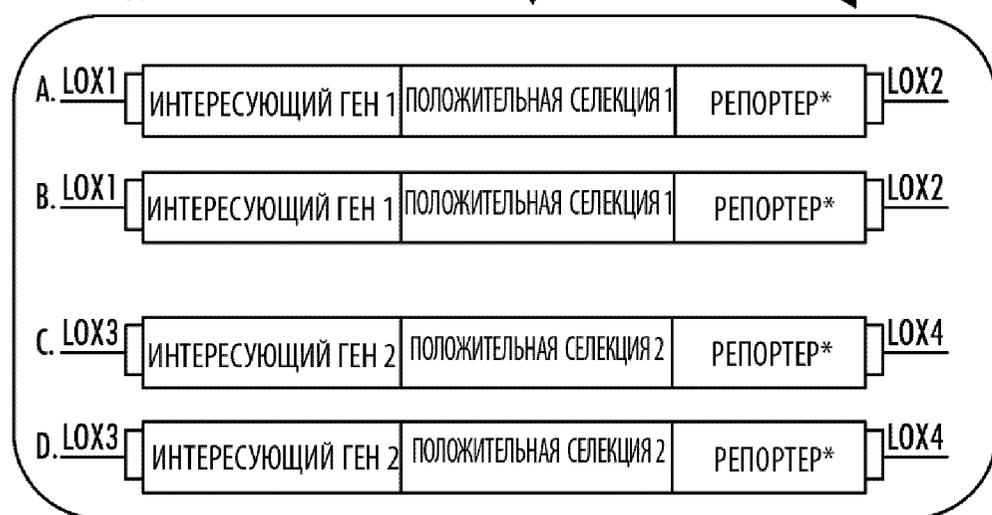


Фиг. 17

КЛЕТКА-ХОЗЯИН



ПРОИЗВОДСТВО БЕЛКА ИНТЕРЕСА КЛЕТКИ



Фиг. 18