



(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2021.02.26

(51) Int. Cl. **C12N 15/90** (2006.01)
C12N 15/85 (2006.01)

(21) Номер заявки
201790698

(22) Дата подачи заявки
2015.10.21

(54) НОВЫЕ САЙТЫ ИНТЕГРАЦИИ СНО И ИХ ПРИМЕНЕНИЯ

(31) **62/067,774**

(32) **2014.10.23**

(33) **US**

(43) **2018.01.31**

(86) **PCT/US2015/056653**

(87) **WO 2016/064999 2016.04.28**

(71)(73) Заявитель и патентовладелец:
**РЕГЕНЕРОН ФАРМАСЬЮТИКАЛЗ,
ИНК. (US)**

(72) Изобретатель:
**Шень Ин, Бураков Дарья, Чень Ган,
Фендл Джэймс П. (US)**

(74) Представитель:
**Карпенко О.Ю., Лыу Т.Н., Угрюмов
В.М., Дементьев В.Н., Глухарёва
А.О., Клюкин В.А., Строкова О.В.,
Христофоров А.А. (RU)**

(56) CHRISTINE LATTENMAYER ET AL.: "Identification of transgene integration loci of different highly expressing recombinant CHO cell lines by FISH", CYTOTECNOLOGY, vol. 51, no. 3, 15 November 2006 (2006-11-15), pages 171-182, XP019448503, KLUWER ACADEMIC PUBLISHERS, DO ISSN: 1573-0778, DOI:10.1007/S10616-006-9029-0 abstract page 180, right-hand column, page 178, right-hand column

XUN XU ET AL.: "The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line", NATURE BIOTECHNOLOGY, vol. 29, no. 8, 31 July 2011 (2011-07-31), pages 735-741, XP055223996, US

ISSN: 1087-0156, DOI: 10.1038/nbt.1932 the whole document

LEWIS NATHAN E. ET AL.: "Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. (Online Methods)", NATURE BIOTECHNOLOGY, vol. 31, no. 8, August 2013 (2013-08), pages 759-765+2PP, XP002753785, ISSN: 1546-1696 the whole document

WO-A1-2012138887

WO-A1-2008151219

HONG ZHOU ET AL.: "Generation of stable cell lines by site-specific integration of transgenes into engineered Chinese hamster ovary strains using an FLP-FRT system", JOURNAL OF BIOTECHNOLOGY, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL vol. 147, no. 2, 17 May 2010 (2010-05-17), pages 122-129, XP002711613, ISSN: 0168-1656, DOI:10.1016/J.JBIOTEC.2010.03.020 Retrieved from the Internet: URL: <http://www.sciencedirect.com/science/article/pii/S0168165610001653> abstract

SCOTT BAHR ET AL.: "Evaluating the effect of chromosomal context on zinc finger nuclease efficiency", BMC PROCEEDINGS, vol. 7, no. Suppl 6, P3, 4 December 2013 (2013-12-04), XP021170326, BIOMED CENTRAL LTD, LONDON UK ISSN: 1753-6561, DOI:10.1186/1753-6561-7-S6-P3 the whole document

MICHEL SADELAIN ET AL.: "Safe harbours for the integration of new DNA in the human genome", NATURE REVIEWS CANCER, vol. 12, no. 1, 1 January 2011 (2011-01-01), pages 51-58, XP055018235, ISSN: 1474-175X, DOI: 10.1038/nrc3179 abstract

(57) Предусматриваются усиливающие экспрессию нуклеотидные последовательности для систем экспрессии эукариот, которые обеспечивают усиленную и стабильную экспрессию рекомбинантных белков в эукариотических клетках. Предусматриваются сайты интеграции в геном, обеспечивающие усиленную экспрессию, и способы их применения для экспрессии гена, представляющего интерес, в эукариотической клетке. Предусматриваются локусы хромосом, последовательности и векторы для усиленной и стабильной экспрессии генов в эукариотических клетках.

Перекрестная ссылка на родственные заявки

Настоящая заявка испрашивает приоритет предварительной заявки на патент США № 62/067774, поданной 23 октября 2014 г., полное содержание которой включено в данный документ посредством ссылки.

Включение перечня последовательностей посредством ссылки

Перечень последовательностей в виде текстового файла ASCII размером 28 KB под названием 32353_T0045US01_SequenceListing.txt, созданного 20 октября 2015 г. и поданного в Ведомство США по патентам и товарным знакам через EFS-Web, включен в данный документ посредством ссылки.

Область техники

Настоящее изобретение обеспечивает стабильную интеграцию и/или экспрессию рекомбинантных белков в эукариотических клетках. В частности, настоящее изобретение предусматривает способы и композиции для повышенной экспрессии белков в эукариотических клетках, в частности в линиях клеток китайского хомячка (*Cricetulus griseus*), при помощи усиливающих экспрессию нуклеотидных последовательностей. Настоящее изобретение предусматривает полинуклеотиды и модифицированные клетки, которые облегчают обмен кассетами, опосредованный рекомбиназой (RMCE). Посредством способов по настоящему изобретению осуществляют интеграцию экзогенных нуклеиновых кислот в конкретные локусы хромосом в геноме клеток китайского хомячка с целью облегчения усиленной и стабильной экспрессии рекомбинантных белков модифицированными клетками.

Описание предшествующего уровня техники

Клеточные системы экспрессии направлены на обеспечение надежного и эффективного источника производства определенного белка для исследовательского и терапевтического применения. Экспрессия рекомбинантного белка в клетках млекопитающих является предпочтительным способом при производстве белков для терапевтического применения в связи, например, со способностью систем экспрессии млекопитающих соответствующим образом посттрансляционно модифицировать рекомбинантные белки.

Несколько клеточных систем доступны для экспрессии белков, при этом каждая содержит различные комбинации цис- и в некоторых случаях трансрегуляторных элементов с целью получения высоких уровней рекомбинантного белка при коротком времени инкубации. Несмотря на доступность многочисленных систем, проблема эффективного переноса гена и стабильности интегрированного гена для экспрессии рекомбинантного белка существует до сих пор. Многочисленные локальные генетические факторы будут определять не только то, когда целевой ген, представляющий интерес, будет экспрессироваться, но и то, сможет ли клетка функционально контролировать транскрипцию гена с точки зрения продуктивного результата или даже будет ли экспрессия поддерживаться в течение длительного времени. Хромосомные сайты интеграции, например сайты интеграции и локус-контролирующие области клеток яичника китайского хомячка (CHO) в определенных генах или рядом с определенными генами, были описаны в уровне техники (WO 2012/138887 A1; Li, Q. et al., 2002 Blood. 100:3077-3086). Таким образом, целевые регуляторные области, как правило, выявляют в области, кодирующей эндогенные белки. В то же время для длительной экспрессии целевого трансгена ключевым фактором является минимальное разрушение генов клетки для того, чтобы избежать изменения фенотипа линии клеток.

Разработка стабильных линий клеток с размещением дополнительных генов для экспрессии, например, дополнительных цепей антител, как в случае мультиспецифических антител, является особенно проблематичной. Могут иметь место широкие вариации уровней экспрессии интегрированных генов. Интеграция дополнительных генов может приводить к большей вариации экспрессии и нестабильности вследствие локального генетического окружения (т.е. эффектов положения). Соответственно в данной области техники существует потребность в улучшении систем экспрессии млекопитающих.

Краткое описание

В соответствии с одним аспектом настоящее изобретение обеспечивает клетку, содержащую экзогенную последовательность нуклеиновой кислоты, интегрированную в специфический сайт в локусе, где локус содержит нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 1 или SEQ ID NO: 4. В соответствии с некоторыми аспектами локус содержит нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 1. В соответствии с некоторыми аспектами локус содержит нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 4.

В соответствии с другим аспектом настоящее изобретение обеспечивает полинуклеотид, содержащий первую последовательность нуклеиновой кислоты, интегрированную в специфический сайт во второй последовательности нуклеиновой кислоты (например, локусе по настоящему изобретению). В соответствии с одним вариантом осуществления вторая последовательность нуклеиновой кислоты содержит нуклеотидную последовательность SEQ ID NO: 1. В соответствии с другим вариантом осуществления вторая последовательность нуклеиновой кислоты содержит нуклеотидную последовательность SEQ ID NO: 4.

В соответствии с одним вариантом осуществления вторая последовательность нуклеиновой кислоты представляет собой усиливающую экспрессию последовательность, выбранную из нуклеотидной по-

следовательности, по меньшей мере на 90% идентичной нуклеиновой кислоте SEQ ID NO: 1, или ее усиливающему экспрессию фрагменту. В соответствии с одним вариантом осуществления вторая последовательность нуклеиновой кислоты представляет собой усиливающую экспрессию последовательность, выбранную из нуклеотидной последовательности, по меньшей мере на 90% идентичной нуклеиновой кислоте SEQ ID NO: 4, или ее усиливающему экспрессию фрагменту. В соответствии с другим вариантом осуществления усиливающая экспрессию последовательность способна усиливать экспрессию белка, кодируемого экзогенной последовательностью нуклеиновой кислоты. В соответствии с другим вариантом осуществления усиливающая экспрессию последовательность способна усиливать экспрессию белка, кодируемого экзогенной последовательностью нуклеиновой кислоты, от по меньшей мере 1,5-кратного до по меньшей мере 3-кратного усиления экспрессии по сравнению с экспрессией, обычно наблюдаемой при случайной интеграции в геном.

В соответствии с другим вариантом осуществления экзогенная последовательность нуклеиновой кислоты интегрируется в специфический сайт в любом положении в SEQ ID NO: 1 или SEQ ID NO: 4.

В соответствии с некоторыми вариантами осуществления специфический сайт в некотором положении в SEQ ID NO: 1 или рядом с некоторым положением в SEQ ID NO: 1 выбирают из группы, состоящей из нуклеотидов, охватывающих положения под номерами

10-4000; 100-3900; 200-3800; 300-3700; 400-3600; 500-3500; 600-3400; 700-3300; 800-3200; 900-3100; 1000-3000; 1100-2900; 1200-2800; 1300-2700; 1200-2600; 1300-2500; 1400-2400; 1500-2300; 1600-2200; 1700-2100; 1800-2050; 1850-2050; 1900-2040; 1950-2025; 1990-2021; 2002-2021 и 2010-2015

в SEQ ID NO: 1. В соответствии с определенными вариантами осуществления специфический сайт в некотором положении в SEQ ID NO: 1 или рядом с некоторым положением в SEQ ID NO: 1 выбирают из группы, состоящей из нуклеотидов, охватывающих положения под номерами

1990-1991; 1991-1992; 1992-1993; 1993-1994; 1995-1996; 1996-1997; 1997-1998; 1999-2000; 2001-2002; 2002-2003; 2003-2004; 2004-2005; 2005-2006; 2006-2007; 2007-2008; 2008-2009; 2009-2010; 2010-2011; 2011-2012; 2012-2013; 2013-2014; 2014-2015; 2015-2016; 2016-2017; 2017-2018; 2018-2019; 2019-2020 и 2020-2021

в SEQ ID NO: 1.

В соответствии с другим вариантом осуществления специфический сайт в некотором положении в SEQ ID NO: 1 или рядом с некоторым положением в SEQ ID NO: 1 выбирают из группы, состоящей из нуклеотидов, охватывающих положения под номерами

10-500; 500-1000; 500-2100; 1000-1500; 1000-2100; 1500-2000; 1500-2500; 2000-2500; 2500-3000; 2500-3500; 3000-3500; 3000-4000 и 3500-4000

в SEQ ID NO: 1. В соответствии с определенными вариантами осуществления экзогенную последовательность нуклеиновой кислоты интегрируют в любой один или более специфических сайтов, описанных в данном документе, в их пределах или возле них.

В соответствии с другим вариантом осуществления экзогенная последовательность нуклеиновой кислоты содержит сайт узнавания, расположенный в усиливающей экспрессию последовательности, как описано выше, при условии, что усиливающая экспрессию последовательность содержит последовательность, которая по меньшей мере на приблизительно 90% идентична, по меньшей мере на приблизительно 91% идентична, по меньшей мере на приблизительно 92% идентична, по меньшей мере на приблизительно 93% идентична, по меньшей мере на приблизительно 94% идентична, по меньшей мере на приблизительно 95% идентична, по меньшей мере на приблизительно 96% идентична, по меньшей мере на приблизительно 97% идентична, по меньшей мере на приблизительно 98% идентична или по меньшей мере на приблизительно 99% идентична усиливающей экспрессию последовательности SEQ ID NO: 1 или SEQ ID NO: 4, ее усиливающему экспрессию фрагменту.

В соответствии с одним вариантом осуществления экзогенная последовательность нуклеиновой кислоты содержит сайт узнавания для рекомбиназы. В соответствии с некоторыми вариантами осуществления экзогенная последовательность нуклеиновой кислоты дополнительно содержит по меньшей мере один сайт узнавания для рекомбиназы, содержащий последовательность, независимо выбранную из сайта LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Lox5171, сайта Loxm2, сайта Lox71, сайта Lox66, сайта LoxFas и сайта f1t. В соответствии с одним вариантом осуществления сайт узнавания для рекомбиназы интегрирован в усиливающей экспрессию последовательности. В соответствии с другим вариантом осуществления сайт узнавания для рекомбиназы расположен непосредственно рядом в направлении 5' к концевому нуклеотиду 5'-конца генной кассеты или расположен непосредственно рядом в направлении 3' к концевому нуклеотиду 3'-конца генной кассеты. В соответствии с некоторыми вариантами осуществления по меньшей мере один сайт узнавания для рекомбиназы и генная кассета интегрированы в усиливающей экспрессию последовательности.

В соответствии с одним вариантом осуществления по меньшей мере два сайта узнавания для ре-

комбиназы присутствуют в усиливающей экспрессию последовательности. В соответствии с другим вариантом осуществления два сайта узнавания для рекомбиназы противоположной ориентации интегрированы в усиливающей экспрессию последовательности. В соответствии с другим вариантом осуществления три сайта узнавания для рекомбиназы интегрированы в усиливающей экспрессию последовательности.

В соответствии с одним аспектом предусматривается выделенная клетка яичника китайского хомячка (CHO), которая содержит разработанную усиливающую экспрессию последовательность SEQ ID NO: 1 или ее усиливающий экспрессию фрагмент. В соответствии с одним вариантом осуществления усиливающую экспрессию последовательность, содержащую нуклеотидную последовательность SEQ ID NO: 1 или SEQ ID NO: 4, или ее стабильный вариант разрабатывают с целью интеграции экзогенной последовательности нуклеиновой кислоты, как описано выше. В соответствии с другими вариантами осуществления настоящее изобретение обеспечивает выделенную клетку CHO, содержащую экзогенную последовательность нуклеиновой кислоты, вставленную в локус, содержащий усиливающую экспрессию последовательность SEQ ID NO: 1 или SEQ ID NO: 4, или ее стабильный вариант.

В соответствии с одним вариантом осуществления клетка CHO дополнительно содержит по меньшей мере одну последовательность узнавания для рекомбиназы в усиливающей экспрессию последовательности. В соответствии с другим вариантом осуществления по меньшей мере одну последовательность узнавания для рекомбиназы независимо выбирают из сайта LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Lox5171, сайта Loxm2, сайта Lox71, сайта Lox66, сайта LoxFas и сайта frt. В соответствии с другим вариантом осуществления сайт узнавания для рекомбиназы расположен непосредственно рядом в направлении 5' к концевому нуклеотиду 5'-конца генной кассеты или расположен непосредственно рядом в направлении 3' к концевому нуклеотиду 3'-конца генной кассеты. В соответствии с некоторыми вариантами осуществления по меньшей мере один сайт узнавания для рекомбиназы и генная кассета интегрированы в усиливающей экспрессию последовательности генома клетки CHO, описанной в данном документе.

В соответствии с другим вариантом осуществления по меньшей мере один сайт узнавания для рекомбиназы расположен, как описано выше, с оговоркой, что генная кассета содержит усиливающую экспрессию последовательность, по меньшей мере на 90% идентичную, по меньшей мере на приблизительно 91% идентичную, по меньшей мере на приблизительно 92% идентичную, по меньшей мере на приблизительно 93% идентичную, по меньшей мере на приблизительно 94% идентичную, по меньшей мере на приблизительно 95% идентичную, по меньшей мере на приблизительно 96% идентичную, по меньшей мере на приблизительно 97% идентичную, по меньшей мере на приблизительно 98% идентичную или по меньшей мере на приблизительно 99% идентичную 1001-2001 нуклеотиду SEQ ID NO: 1 (SEQ ID NO: 2) или ее усиливающему экспрессию фрагменту. В соответствии с другим вариантом осуществления по меньшей мере один сайт узнавания для рекомбиназы расположен, как описано выше, с оговоркой, что генная кассета содержит усиливающую экспрессию последовательность, по меньшей мере на 90% идентичную, по меньшей мере на приблизительно 91% идентичную, по меньшей мере на приблизительно 92% идентичную, по меньшей мере на приблизительно 93% идентичную, по меньшей мере на приблизительно 94% идентичную, по меньшей мере на приблизительно 95% идентичную, по меньшей мере на приблизительно 96% идентичную, по меньшей мере на приблизительно 97% идентичную, по меньшей мере на приблизительно 98% идентичную или по меньшей мере на приблизительно 99% идентичную 2022-3022 нуклеотиду SEQ ID NO: 1 (SEQ ID NO: 3) или ее усиливающему экспрессию фрагменту.

В соответствии с еще одним вариантом осуществления по меньшей мере один сайт узнавания для рекомбиназы вставляют в геном клетки CHO в пределах нуклеотидов

1990-1991, 1991-1992, 1992-1993, 1993-1994, 1995-1996, 1996-1997, 1997-1998, 1999-2000, 2001-2002, 2002-2003, 2003-2004, 2004-2005, 2005-2006, 2006-2007, 2007-2008, 2008-2009, 2009-2010, 2010-2011, 2011-2012, 2012-2013, 2013-2014, 2014-2015, 2015-2016, 2016-2017, 2017-2018, 2018-2019, 2019-2020, 2020-2021 или 2021-2022 SEQ ID NO:1.

В соответствии с другим вариантом осуществления экзогенную нуклеиновую кислоту вставляют в геном CHO в пределах нуклеотидов

1990-1991, 1991-1992, 1992-1993, 1993-1994, 1995-1996, 1996-1997, 1997-1998, 1999-2000, 2001-2002, 2002-2003, 2003-2004, 2004-2005, 2005-2006, 2006-2007, 2007-2008, 2008-2009, 2009-2010, 2010-2011, 2011-2012, 2012-2013, 2013-2014, 2014-2015, 2015-2016, 2016-2017, 2017-2018, 2018-2019, 2019-2020, 2020-2021 или 2021-2022 SEQ ID NO:1.

В соответствии с другим вариантом осуществления экзогенную нуклеиновую кислоту вставляют в геном CHO в пределах нуклеотидов 2001-2022 SEQ ID NO: 1. В соответствии с некоторыми вариантами осуществления экзогенную нуклеиновую кислоту вставляют в пределах нуклеотидов 2001-2002 или нуклеотидов 2021-2022 или SEQ ID NO: 1, а нуклеотиды 2002-2021 SEQ ID NO: 1 удаляют в результате

вставки. Аналогично, экзогенную нуклеиновую кислоту вставляют в геном CHO в пределах нуклеотидов 9302-9321 SEQ ID NO: 4. В соответствии с некоторыми вариантами осуществления экзогенную нуклеиновую кислоту вставляют в пределах нуклеотидов 9301-9302 или нуклеотидов 9321-9322 SEQ ID NO: 4, а нуклеотиды 9302-9321 SEQ ID NO: 4 удаляют в результате вставки.

В соответствии с некоторыми вариантами осуществления экзогенная последовательность нуклеиновой кислоты, интегрированная в специфический сайт в локусе, такая как нуклеотидная последовательность SEQ ID NO: 1 или SEQ ID NO: 4, содержит ген, представляющий интерес (GOI) (например, нуклеотидная последовательность, кодирующая представляющий интерес белок "POI"). В соответствии с определенными вариантами осуществления экзогенная последовательность нуклеиновой кислоты содержит один или более представляющих интерес генов. В соответствии с некоторыми вариантами осуществления один или более представляющих интерес генов выбирают из группы, состоящей из первого GOI, второго GOI и третьего GOI.

В соответствии с некоторыми вариантами осуществления экзогенная последовательность нуклеиновой кислоты, интегрированная в специфический сайт в локусе, такая как нуклеотидная последовательность SEQ ID NO: 1 или SEQ ID NO: 4, содержит GOI и по меньшей мере один сайт узнавания для рекомбиназы. В соответствии с одним вариантом осуществления первый GOI вставляют в усиливающую экспрессию последовательность SEQ ID NO: 1 или SEQ ID NO: 4, или усиливающую экспрессию последовательность, по меньшей мере на 90% идентичную нуклеотидам SEQ ID NO: 1 или SEQ ID NO: 4, или ее усиливающий экспрессию фрагмент, как описано выше, и первый GOI необязательно функционально связан с промотором, где связанный с промотором GOI (или GOI) фланкирован в направлении 5' первым сайтом узнавания для рекомбиназы и в направлении 3' вторым сайтом узнавания для рекомбиназы. В соответствии с другим вариантом осуществления второй GOI вставляют в направлении 3' от второго сайта узнавания для рекомбиназы, и при этом второй GOI фланкирован в направлении 3' третьим сайтом узнавания для рекомбиназы.

В соответствии с еще одним вариантом осуществления GOI функционально связан с промотором, способным управлять экспрессией GOI, где промотор содержит эукариотический промотор, который можно регулировать при помощи активатора или ингибитора. В соответствии с другими вариантами осуществления эукариотический промотор функционально связан с прокариотическим промотором, и эукариотическая клетка необязательно дополнительно содержит репрессорный белок прокариот.

В соответствии с другим вариантом осуществления один или более селективных маркеров включают между первым и вторым и/или вторым и третьим сайтами узнавания для рекомбиназы. В соответствии с некоторыми вариантами осуществления первый и/или второй представляющие интерес гены и/или один или более селективных маркеров функционально связаны с промотором, где промотор может быть один и тот же или иной. В соответствии с другим вариантом осуществления промотор содержит эукариотический промотор (такой как, например, промотор CMV или поздний промотор SV40), необязательно контролируемый прокариотическим оператором (таким как, например, tet-оператор). В соответствии с другими вариантами осуществления клетка дополнительно содержит ген, кодирующий прокариотический репрессор (такой как, например, репрессор tet).

В соответствии с другим вариантом осуществления клетка дополнительно содержит ген, способный экспрессировать рекомбиназу. В соответствии с некоторыми вариантами осуществления рекомбиназой является Cre-рекомбиназа.

В соответствии с одним аспектом предусматривается клетка-хозяин CHO, содержащая усиливающую экспрессию последовательность, выбранную из SEQ ID NO: 1 или SEQ ID NO: 4, или усиливающую экспрессию последовательность, по меньшей мере на 90% идентичную нуклеотидам SEQ ID NO: 1 или SEQ ID NO: 4, или ее усиливающему экспрессию фрагменту, содержащему первый сайт узнавания для рекомбиназы с последующим первым эукариотическим промотором, ген первого селективного маркера, второй эукариотический промотор, ген второго селективного маркера и второй сайт узнавания для рекомбиназы. В других вариантах осуществления клетка-хозяин CHO дополнительно предусматривает третий эукариотический промотор, ген третьего маркера и третий сайт узнавания для рекомбиназы. В соответствии с одним вариантом осуществления усиливающая экспрессию последовательность находится в SEQ ID NO: 1 или SEQ ID NO: 4, как описано выше.

В соответствии с одним вариантом осуществления первый, второй и третий сайты узнавания для рекомбиназы отличаются друг от друга. В соответствии с некоторыми вариантами осуществления сайты узнавания для рекомбиназы выбирают из сайта LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Lox5171, сайта Loxm2, сайта Lox71, сайта Lox66, сайта LoxFas и сайта fit.

В соответствии с одним вариантом осуществления геном первого селективного маркера является ген устойчивости к лекарственному средству. В соответствии с другим вариантом осуществления геном устойчивости к лекарственному средству является ген устойчивости к неомицину или ген устойчивости к гидромицину. В соответствии с другим вариантом осуществления гены второго и третьего селективных маркеров кодируют два различных флуоресцентных белка. В соответствии с одним вариантом осуществления два различных флуоресцентных белка выбирают из группы, состоящей из флуоресцентного красного белка, выделенного из *Discosoma* (DsRed), зеленого флуоресцентного белка (GFP), усиленного зе-

ленного флуоресцентного белка (eGFP), голубого флуоресцентного белка (CFP), усиленного голубого флуоресцентного белка (eCFP), желтого флуоресцентного белка (YFP), усиленного желтого флуоресцентного белка (eYFP) и флуоресцентного белка дальнего красного спектра (например, mKate, mKate2, mPlum, mRaspberry или E2-crimson).

В соответствии с одним вариантом осуществления первый, второй и третий промоторы являются одинаковыми. В соответствии с другим вариантом осуществления первый, второй и третий промоторы отличаются друг от друга. В соответствии с другим вариантом осуществления первый промотор отличается от второго и третьего промоторов, а второй и третий промоторы являются одинаковыми. В соответствии с другими вариантами осуществления первым промотором является поздний промотор SV40, а второй и третий промоторы представляют собой промотор CMV человека. В соответствии с другими вариантами осуществления первый и второй промоторы функционально связаны с прокариотическим оператором.

В соответствии с одним вариантом осуществления линия клетки-хозяина имеет экзогенно добавленный ген, кодирующий рекомбиназу, интегрированный в ее геном, функционально связанный с промотором. В соответствии с другим вариантом осуществления рекомбиназой является Cre-рекомбиназа. В соответствии с другим вариантом осуществления клетка-хозяин имеет ген, кодирующий регуляторный белок, интегрированный в ее геном, функционально связанный с промотором. В соответствии с другими вариантами осуществления регуляторным белком является репрессорный белок tet.

В соответствии с одним вариантом осуществления первый GOI и второй GOI кодируют легкую цепь или ее фрагмент антитела или тяжелую цепь или ее фрагмент антитела. В соответствии с другим вариантом осуществления первый GOI кодирует легкую цепь антитела, и второй GOI кодирует тяжелую цепь антитела.

В соответствии с определенными вариантами осуществления первый, второй и третий GOI кодируют полипептид, выбранный из группы, состоящей из первой легкой цепи или ее фрагмента, второй легкой цепи или ее фрагмента и тяжелой цепи или ее фрагмента. В соответствии с еще одним вариантом осуществления первый, второй и третий GOI кодируют полипептид, выбранный из группы, состоящей из легкой цепи или ее фрагмента, первой тяжелой цепи или ее фрагмента и второй тяжелой цепи или ее фрагмента.

В соответствии с одним вариантом осуществления предусматривается способ получения представляющего интерес белка, предусматривающий (а) введение в клетку-хозяина СНО гена, представляющего интерес (GOI), где GOI интегрируется в специфический локус, содержащий нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 1 или SEQ ID NO: 4; (b) культивирование клетки из (а) в условиях, которые обеспечивают экспрессию GOI; и (c) извлечение представляющего интерес белка. В соответствии с одним вариантом осуществления представляющий интерес белок выбирают из группы, состоящей из субъединицы иммуноглобулина или его фрагмента и рецептора или его лигандсвязывающего фрагмента. В соответствии с определенными вариантами осуществления представляющий интерес белок выбирают из группы, состоящей из легкой цепи антитела или ее фрагмента, участвующего в связывании с антигеном, и тяжелой цепи антитела или ее фрагмента, участвующего в связывании с антигеном.

В соответствии с некоторыми вариантами осуществления GOI вводят в клетку при помощи направленного вектора для обмена кассетами, опосредованного рекомбиназой (RMCE), и геном клетки-хозяина СНО содержит по меньшей мере одну экзогенную последовательность узнавания в специфическом локусе. В соответствии с другими вариантами осуществления геном клетки-хозяина СНО содержит по меньшей мере одну экзогенную последовательность узнавания и селективный маркер, необязательно функционально связанный с промотором, IRES и/или последовательность полиаденилирования (polyA) в специфическом локусе.

В соответствии с определенными вариантами осуществления геном клетки-хозяина СНО содержит один или более сайтов узнавания для рекомбиназы, как описано выше, и GOI вводят в специфический локус посредством действия рекомбиназы, которая распознает сайт узнавания для рекомбиназы.

В соответствии с другим вариантом осуществления GOI вводят в клетку при помощи направленного вектора для гомологичной рекомбинации, и при этом направленный вектор содержит 5'-область гомологии, гомологичную последовательности, присутствующей в специфическом локусе, GOI и 3'-область гомологии, гомологичную последовательности, присутствующей в специфическом локусе. В соответствии с другим вариантом осуществления направленный вектор дополнительно содержит два, три, четыре или пять или более представляющих интерес генов. В соответствии с другим вариантом осуществления один или более из представляющих интерес генов функционально связаны с промотором.

В соответствии с другим аспектом предусматривается направленный вектор, где направленный вектор содержит 5'-область гомологии, гомологичную последовательности, присутствующей в локусе, содержащем нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 1 или SEQ ID NO: 4, GOI и 3'-область гомологии, гомологичную последовательности, присутствующей в локусе, содержащем нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 1 или SEQ ID NO: 4. В соответствии с другим вариантом осуществления направленный век-

тор дополнительно содержит два, три, четыре или пять или более представляющих интерес генов.

В соответствии с другим аспектом предусматривается способ модификации генома клетки CHO с целью интеграции экзогенной последовательности нуклеиновой кислоты, предусматривающий стадию введения в клетку средства для переноса, которое включает вектор, где вектор содержит экзогенную последовательность нуклеиновой кислоты в локусе генома, содержащего нуклеотидную последовательность, по меньшей мере на 90% идентичную SEQ ID NO: 1 или SEQ ID NO: 4.

В соответствии с некоторыми вариантами осуществления вектор содержит 5'-область гомологии, гомологичную последовательности, присутствующей в локусе генома, содержащего нуклеотидную последовательность, по меньшей мере на 90% идентичную SEQ ID NO: 1 или SEQ ID NO: 4, экзогенную последовательность нуклеиновой кислоты, и 3'-область гомологии, гомологичную последовательности, присутствующей в локусе генома, содержащего нуклеотидную последовательность, по меньшей мере на 90% идентичную SEQ ID NO: 1 или SEQ ID NO: 4.

В соответствии с некоторыми вариантами осуществления экзогенная последовательность нуклеиновой кислоты в векторе содержит

одну или более последовательностей узнавания. В соответствии с другими вариантами осуществления экзогенная нуклеиновая кислота содержит один или более GOI, таких как селективный маркер, или нуклеиновую кислоту, кодирующую POI. В соответствии с еще одними вариантами осуществления экзогенная нуклеиновая кислота содержит один или более GOI и одну или более последовательностей узнавания.

В соответствии с одним вариантом осуществления средство для переноса содержит по меньшей мере один дополнительный вектор или мРНК. В соответствии с другим вариантом осуществления дополнительный вектор выбирают из группы, состоящей из аденовируса, лентивируса, ретровируса, аденоассоциированного вируса, интегрирующего фагового вектора, невирусного вектора, транспозона и/или транспозазы, субстрата интегразы и плазмиды. В соответствии с некоторыми вариантами осуществления дополнительный вектор содержит нуклеотидную последовательность, кодирующую сайт-специфическую нуклеазу, с целью интеграции экзогенной последовательности нуклеиновой кислоты.

В соответствии с определенными вариантами осуществления сайт-специфическая нуклеаза содержит цинк-пальцевую нуклеазу (ZFN), димер ZFN, эффекторную нуклеазу, подобную активаторам транскрипции (TALEN), белок слияния на основе домена TAL-эффектора или РНК-управляемую ДНК-эндонуклеазу.

В соответствии с другим аспектом предусматривается средство для переноса, предназначенное для модификации генома клетки CHO с целью интеграции экзогенной последовательности нуклеиновой кислоты, где средство для переноса включает вектор, где вектор содержит 5'-область гомологии, гомологичную последовательности, присутствующей в локусе генома, содержащего нуклеотидную последовательность, по меньшей мере на 90% идентичную SEQ ID NO: 1 или SEQ ID NO: 4, экзогенную последовательность нуклеиновой кислоты, и 3'-область гомологии, гомологичную последовательности, присутствующей в локусе генома, содержащего нуклеотидную последовательность, по меньшей мере на 90% идентичную SEQ ID NO: 1 или SEQ ID NO: 4.

В соответствии с некоторыми вариантами осуществления экзогенная последовательность нуклеиновой кислоты содержит одну или более последовательностей узнавания. В соответствии с другими вариантами осуществления экзогенная нуклеиновая кислота содержит один или более GOI, таких как селективный маркер, или нуклеиновую кислоту, кодирующую POI. В соответствии с еще одними вариантами осуществления экзогенная нуклеиновая кислота содержит один или более GOI и одну или более последовательностей узнавания.

В соответствии с еще одним аспектом предусматривается способ модификации генома клетки CHO с целью экспрессии терапевтического средства, содержащего средство для переноса, предназначенное для введения в геном экзогенной нуклеиновой кислоты, содержащей последовательность для экспрессии терапевтического средства, где средство для переноса содержит 5'-область гомологии, гомологичную последовательности, присутствующей в нуклеотидной последовательности SEQ ID NO: 1, нуклеиновую кислоту, кодирующую терапевтическое средство, и 3'-область гомологии, гомологичную последовательности, присутствующей в нуклеотидной последовательности SEQ ID NO: 1 или SEQ ID NO: 4.

В соответствии с еще одним аспектом настоящее изобретение обеспечивает модифицированную клетку-хозяина CHO и модифицированный геном CHO, где геном CHO модифицируют при помощи вставки экзогенной последовательности узнавания в локус генома, имеющий нуклеотидную последовательность, по меньшей мере на 90% идентичную SEQ ID NO: 1.

В соответствии с другим аспектом настоящее изобретение обеспечивает модифицированную эукариотическую клетку-хозяина, содержащую модифицированный эукариотический геном, где эукариотический геном модифицируют в сайте направленной интеграции в некодирующей области генома с целью вставки экзогенной нуклеиновой кислоты. В соответствии с некоторыми вариантами осуществления экзогенной нуклеиновой кислотой является последовательность узнавания. В соответствии с другими вариантами осуществления клеткой-хозяином является клетка-хозяин млекопитающего, такая как клетка CHO. В соответствии с другими вариантами осуществления сайт направленной интеграции содержит

усиливающую экспрессию последовательность, такую как SEQ ID NO: 1, при условии, что эта последовательность не кодирует никаких эндогенных белков. Настоящее изобретение также предусматривает способы получения такой модифицированной эукариотической клетки-хозяина.

В соответствии с любым из аспектов и вариантов осуществления, описанных выше, усиливающую экспрессию последовательность можно поместить в указанной ориентации, как в SEQ ID NO: 1, либо в обратном направлении от ориентации SEQ ID NO: 1.

Любые аспекты и варианты осуществления по настоящему изобретению можно применять совместно с любым другим аспектом или вариантом осуществления по настоящему изобретению, если иное не упомянуто или не видно из контекста.

Другие цели и преимущества будут очевидны из обзора следующего подробного описания.

Краткое описание графических материалов

Фиг. 1А и 1В. Фиг. 1А: схематическая диаграмма функциональной конструкции, использующей случайное введение молекулы нуклеиновой кислоты, экспрессирующей GOI (например, многоцепочечное антитело) и множественные копии селективного маркера, в геном клетки, например геном CHO, с целью выявления целевого локуса. Иллюстративная конструкция включает тяжелую цепь (HC); первую копию селективного маркера, такого как ген устойчивости к гидромицину (Hug); первую копию легкой цепи (LC); вторую копию селективного маркера (например, Hug), вторую копию легкой цепи (LC); третью копию селективного маркера (например, Hug). Фиг. 1В: иллюстративный донорный вектор с целью интеграции посредством гомологичной рекомбинации в нативный локус, определенный в виде SEQ ID NO: 1. 5'- и 3'-области гомологии получают из SEQ ID NO: 1.

На фиг. 2А-2С представлено, что локус SEQ ID NO: 1 (LOCUS 1), функционально связанный с геном, представляющим интерес (GOI), характеризуется усиленной экспрессией мРНК GOI по сравнению с тем же самым GOI, который функционально не связан с LOCUS 1, а вместо этого связан с контрольным локусом. Фиг. 2А: эквивалентное число копий генов, показанных для клеток, кодирующих ген антитела, представляющий интерес, т.е. одну тяжелую цепь (HC) и две легкие цепи (LC), функционально связанные с контрольным локусом и LOCUS 1. Фиг. 2В: уровни мРНК выше для GOI, экспрессируемого в LOCUS 1 по сравнению с мРНК контрольного локуса. Фиг. 2С: титр белка в 3 раза выше для клеток, экспрессирующих GOI в LOCUS 1 по сравнению с титром белка, образуемого клетками, экспрессирующими тот же самый GOI в контрольном локусе.

На фиг. 3А и 3В представлена иллюстративная кассета, содержащая флуоресцентный маркер и GOI, интегрированный в LOCUS 1 (например, mKate, фланкированный сайтами lox с целью обмена eYFP и GOI), по сравнению с такой же самой кассетой, интегрированной в контрольный локус (замененный другим флуоресцентным маркером, например, dsRed2, фланкированным сайтами lox), где для такой интеграции используется Cre-рекомбиназа и обмен кассетами, опосредованный рекомбиназой (RMCE). Такие кассеты использовали в экспериментах для измерения эффективности рекомбинации и транскрипции GOI.

На фиг. 4 показан более высокий уровень мРНК гена, представляющего интерес (GOI), как измерено в популяции клеток CHO, экспрессирующих GOI в LOCUS 1 (SEQ ID NO: 1), по сравнению с мРНК из популяции клеток CHO, экспрессирующих тот же самый GOI, в тех же самых регуляторных условиях, но интегрированный в контрольный локус, т.е. EESYR.

Подробное описание изобретения

Перед описанием способов настоящего изобретения необходимо понимать, что настоящее изобретение не ограничивается конкретными описанными способами и условиями экспериментов, поскольку такие способы и условия могут варьировать. Также следует понимать, что терминология, используемая в данном документе, предназначена только для целей описания конкретных вариантов осуществления и не предназначена для ограничения, поскольку объем настоящего изобретения будет ограничен только прилагаемой формулой изобретения.

Как используется в данном документе и прилагаемой формуле изобретения, формы единственного числа включают ссылки на форму множественного числа, не считая контекста, в котором четко продиктовано иное. Таким образом, например, ссылка на "способ" включает один или более способов и/или стадий, описанных в данном документе, и/или которые станут очевидными специалистам в данной области техники при прочтении настоящего раскрытия.

Если не определено иное или иное не упомянуто, все технические и научные термины, применяемые в данном документе, имеют такое же значение, как это обычно понимает специалист в данной области техники к которой принадлежит настоящее изобретение.

Хотя при практическом осуществлении или испытании настоящего изобретения можно применять любые способы и материалы, сходные с описанными в данном документе или эквивалентные им, в данном документе описаны только конкретные способы и материалы. Все публикации, упомянутые в данном документе, включены в данный документ посредством ссылки во всей своей полноте.

Определения.

Области ДНК являются функционально связанными, если они функционально связаны друг с другом. Например, промотор является функционально связанным с кодирующей последовательностью, если

промотор способен участвовать в транскрипции последовательности; сайт связывания рибосомы является функционально связанным с кодирующей последовательностью, если он расположен таким образом, чтобы обеспечить трансляцию. Как правило, функционально связанный может включать, но не требует, смежного расположения. В случае последовательностей, таких как секреторные лидерные последовательности, смежное расположение и правильность размещения в рамке считывания являются типичными признаками. Усиливающая экспрессию последовательность представляющего интерес локуса функционально связана с геном, представляющим интерес (GOI), где она функционально связана с GOI, например, где ее присутствие приводит к усиленной экспрессии и/или стабильной интеграции GOI.

Выражение "усиленный" при использовании для описания усиленной экспрессии включает усиление по меньшей мере от примерно 1,5-кратного по меньшей мере до примерно 3-кратного усиления экспрессии по сравнению с тем, что обычно наблюдается при случайной интеграции экзогенной последовательности в геном или при интеграции в другой локус, например, по сравнению с популяцией случайных интегрантов одной копии той же самой конструкции для экспрессии. Наблюдаемое усиление кратности экспрессии при использовании последовательностей по настоящему изобретению представлено по сравнению с уровнем экспрессии одного и того же гена, измеренной при фактически одинаковых условиях, в отсутствие последовательности по настоящему изобретению, например, по сравнению с интеграцией в другой локус в геном того же самого вида. Усиленная эффективность рекомбинации включает усиление способности локуса рекомбинировать (например, использовать сайты узнавания для рекомбиназы). Усиление относится к эффективности рекомбинации по сравнению со случайной рекомбинацией, например, без использования сайтов узнавания для рекомбиназы или подобного, которая обычно составляет 0,1%. Предпочтительная эффективность усиленной рекомбинации приблизительно в 10 раз выше случайной или приблизительно 1%. Если не указано иное, заявленное изобретение не ограничено конкретной эффективностью рекомбинации.

Если выражение "экзогенно добавляемый ген" или "экзогенно добавляемая нуклеиновая кислота" используется по отношению к представляющему интерес локусу, то это выражение относится к любой последовательности ДНК или гену, не присутствующему в представляющем интерес локусе, как в локусе, встречающемся в природе. Например, "экзогенно добавленный ген" в локусе CHO (например, в локусе, содержащем последовательность SEQ ID NO: 1), может быть геном хомячка, не встречающимся в определенном локусе CHO в природе (т.е. геном хомячка из другого локуса генома хомячка), геном от любого другого вида (например, геном человека), химерным геном (например, человека/мыши) или любым другим геном, который не существует в природе в представляющем интерес локусе CHO.

Процентная идентичность при описании представляющего интерес белка, такого как SEQ ID NO: 1 или SEQ ID NO: 4, или его фрагмента подразумевает включение гомологичных последовательностей, которые характеризуются упомянутой идентичностью в пределах областей со смежной гомологией, однако при расчете процентной идентичности не принимают во внимание присутствие гэпов, делеций или вставок, которые не имеют гомолога в сравниваемой последовательности.

Как используется в данном документе, определение "процентной идентичности" между, например, SEQ ID NO: 1 или ее фрагментом и видовым гомологом не включает сравнение последовательностей, где видовой гомолог не имеет гомологичной последовательности при выравнивании (т.е. SEQ ID NO: 1 или ее фрагмент имеет вставку в этой точке, а видовой гомолог имеет гэнг или делецию в зависимости от обстоятельств). Таким образом, "процентная идентичность" не включает штрафы за гэпы, делеции и вставки.

"Гомологичная последовательность" в контексте последовательностей нуклеиновых кислот относится к последовательности, которая фактически гомологична референтной последовательности нуклеиновой кислоты. В соответствии с некоторыми вариантами осуществления две последовательности считаются фактически гомологичными, если по меньшей мере 50, 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% или более их соответствующих нуклеотидов являются идентичными на протяжении рассматриваемого фрагмента из остатков. В соответствии с некоторыми вариантами осуществления рассматриваемый фрагмент представляет собой полную (т.е. целую) последовательность.

"Направленная вставка" относится к способам направленности на гены с целью вставки или интеграции гена или последовательности нуклеиновой кислоты в определенное положение в геноме, т.е. с целью направления ДНК в специфический сайт между двумя нуклеотидами в смежной полинуклеотидной цепи. Направленную вставку можно также выполнять для определенной генной кассеты, которая включает несколько генов, регуляторных элементов и/или последовательностей нуклеиновых кислот. "Вставка" и "интеграция" используются взаимозаменяемо. Подразумевается, что вставка гена или последовательности нуклеиновой кислоты (например, последовательности нуклеиновой кислоты, содержащей кассету экспрессии) может приводить к замене или делеции (или может быть выполнена для замены или делеции) одной или более нуклеиновых кислот в зависимости от используемой методики редактирования генома.

"Сайт узнавания" или "последовательность узнавания" представляет собой определенную последовательность ДНК, распознаваемую нуклеазой или другим ферментом с целью связывания и контроля сайт-специфического расщепления остова ДНК. Эндонуклеазы расщепляют ДНК в молекуле ДНК. Сай-

ты узнавания также обозначаются в данной области техники в виде целевых сайтов узнавания.

"Сайт узнавания для рекомбиназы" представляет собой определенную последовательность ДНК, распознаваемую рекомбиназой, такой как Cre-рекомбиназа (Cre) или флиппаза (Flp). Сайт-специфические рекомбиназы могут осуществлять перестройки ДНК, в том числе делеции, инверсии и транслокации, если одну или более из их последовательностей узнавания помещают стратегически в геном организма. В одном примере Cre специфически опосредует явления рекомбинации в своем целевом сайте узнавания ДНК loxP, который состоит из двух инвертированных повторов по 13 п.о., разделенных спейсером из 8 п.о. Можно использовать несколько сайтов узнавания для рекомбиназы, например, для облегчения обусловленного рекомбиназой обмена ДНК. Также можно использовать варианты или мутанты сайтов узнавания для рекомбиназ, такие как сайты lox (Araki, N. et al., 2002, *Nucleic Acids Research*, 30:19, e103).

"Обмен кассетами, опосредованный рекомбиназой" относится к процессу точного замещения геномной целевой кассеты донорной кассетой. Молекулярные композиции, обычно предусматриваемые с целью осуществления этого процесса, включают 1) геномную целевую кассету, фланкированную как в направлении 5', так и в направлении 3' целевыми сайтами, специфическими по отношению к определенной рекомбиназе, 2) донорную кассету, фланкированную подходящими целевыми сайтами узнавания, и 3) сайт-специфическую рекомбиназу. Рекомбиназные белки хорошо известны из уровня техники (Turap, S. and Bode J., 2011, *FASEB J.*, 25, pp. 4088-4107) и обеспечивают точное расщепление ДНК в определенном целевом сайте узнавания (последовательности ДНК) без добавления или потери нуклеотидов. Распространенные комбинации рекомбиназа/сайт включают без ограничений Cre/lox и Flp/frt.

"Средство для переноса" представляет собой композицию, состоящую из любого полинуклеотида или совокупности полинуклеотидов, несущих экзогенную нуклеиновую кислоту для введения в клетку. Средство для переноса включает векторы, плазмиды и молекулы мРНК, которые доставляются в клетку при помощи хорошо известных способов трансфекции. В одном примере мРНК, вводимая в клетки, может быть временной и не интегрируется в геном, однако мРНК может нести экзогенную нуклеиновую кислоту, необходимую для осуществления процесса интеграции.

Общее описание.

Настоящее изобретение основано, по меньшей мере, частично на обнаружении уникальных последовательностей, т.е. локусов, в геноме, которые характеризуются более эффективной рекомбинацией, стабильностью при вставке и более высоким уровнем экспрессии, чем другие области или последовательности в геноме. Настоящее изобретение также основано, по меньшей мере, частично на том факте, что при выявлении таких усиливающих экспрессию последовательностей можно экзогенно добавлять подходящий ген или конструкцию в эти последовательности или возле этих последовательностей, и что экзогенно добавленный ген может преимущественно экспрессироваться или использоваться для дополнительных геномных модификаций. Такие последовательности, называемые усиливающими экспрессию последовательностями, считаются стабильными и не располагаются в кодирующей области генома. Эти усиливающие экспрессию и стабильные участки можно конструировать для дополнительных явлений клонирования или геномного редактирования. Таким образом, надежную систему экспрессии встраивают в геномный остов клетки.

Настоящее изобретение также основано на определенной направленности экзогенного гена в отношении сайта интеграции. Способы по настоящему изобретению обеспечивают эффективное "превращение" генома клетки в пригодную кассету клонирования, например, при помощи обмена кассетами, опосредованного рекомбиназой (RMCE). С этой целью в способах по настоящему изобретению используют сайты узнавания для рекомбиназы клеточного генома для размещения представляющих интерес генов с целью создания высокопродуктивных линий клеток для получения рекомбинантных белков.

Композиции по настоящему изобретению также можно включать в конструкции для экспрессии, например в векторы экспрессии для клонирования и разработки новых линий клеток. Векторы экспрессии, содержащие полинуклеотиды по настоящему изобретению, можно использовать для экспрессии белков временно или можно интегрировать в геном при помощи случайной или направленной рекомбинации, такой как гомологичная рекомбинация или рекомбинация, опосредованная рекомбиназами, которые распознают конкретные сайты рекомбинации (например, Cre-lox-опосредованная рекомбинация). Векторы экспрессии, содержащие полинуклеотиды по настоящему изобретению, можно также использовать для оценки эффективности других последовательностей ДНК, например регуляторных последовательностей, действующих в цис-положении.

Сайты интеграции обычно выявляют либо при случайной интеграции, либо при анализе явлений интеграции при помощи ретровирусов. Сайт интеграции СНО, подробно описанный в данном документе, был выявлен при помощи случайной интеграции ДНК, кодирующей мультицепочечное антитело, и, как было обнаружено, экспрессируемый белок характеризовался усиленной экспрессией.

Иллюстративное мультицепочечное антитело, содержащее одну тяжелую цепь (НС) и две копии легкой цепи (LC), случайно интегрировали в геном в виде кассеты экспрессии, содержащей чередующиеся гены устойчивости к гидромицину (см., например, три одинаковых гена *Hyg*, как обозначено на фиг. 1A). Один стабильный и высокоэкспрессирующий клон получали в результате интеграции кассеты экспрессии в локусы, выявленные в виде SEQ ID NO: 1.

В сравнении с интеграцией в другую область генома CHO (контрольный сайт интеграции) иллюстративное мультицепочечное антитело характеризуется более высокими уровнями экспрессии при интеграции в локус SEQ ID NO: 1. Интересным является то, что число копий гена сопоставимо для экспрессирующих антитела полинуклеотидов, интегрированных в SEQ ID NO: 1 по сравнению с контрольным сайтом интеграции, однако титры белка в 3 раза выше для экспрессирующих антитела полинуклеотидов, интегрированных в SEQ ID NO: 1.

Методы направленной рекомбинации использовали для превращения генома клетки CHO в конструкцию для клонирования, содержащую сайты узнавания для рекомбиназы (см., например, фиг. 3A-B).

Главным образом, после выявления сайта интеграции SEQ ID NO: 1 сайты узнавания для рекомбиназы (например, сайты lox) использовали в локусе для введения кассет экспрессии, которые содержат экспрессируемый GOI, такой как селективный маркер (см., например, фиг. 3A-B), вместе с любыми другими желательными элементами, такими как, например, промоторы, энхансеры, маркеры, операторы, сайты связывания рибосомы (например, сайты внутренней посадки рибосомы) и др.

Иллюстрация примера донорской конструкции, используемой для направленной интеграции сайтов lox в SEQ ID NO: 1, показана на фиг. 1B. Донорская конструкция содержит кассету экспрессии, контролируемую геном устойчивости к неомицину (neo) и сайтом внутренней посадки рибосомы (IRES), где кассета содержит флуоресцентный маркер (mKate) и фланкирована на 5'- и 3'-концах сайтами узнавания для рекомбиназы 5'- и 3'-областями гомологии (гомологичными SEQ ID NO: 1). Показана вставка в локус SEQ ID NO: 1, где вставка приводит к тому, что донорская конструкция neo/mKate замещает кассету экспрессии, содержащую маркер устойчивости к гидромицину, где кассета экспрессии в локусе SEQ ID NO: 1 фланкирована на 5'- и 3'-концах сайтами узнавания для рекомбиназы, соединенными с 5'- и 3'-областями гомологии (гомологичными SEQ ID NO: 1) (см. фиг. 1B).

Предусматриваются композиции и способы стабильной интеграции последовательности нуклеиновой кислоты в эукариотическую клетку, где последовательность нуклеиновой кислоты способна к усиленной экспрессии благодаря интеграции в SEQ ID NO: 1 или ее усиливающий экспрессию фрагмент. Предусматриваются клетки, которые содержат последовательность узнавания для рекомбиназы в SEQ ID NO: 1, подходящую для вставки GOI, с целью достижения экспрессии представляющего интерес белка GOI. Также предусматриваются композиции и способы направления сайтов интеграции вместе с конструкциями для экспрессии, например векторами экспрессии, и добавления экзогенной(экзогенных) нуклеиновой(нуклеиновых) кислоты(кислот) в представляющую интерес клетку CHO.

Физическая и функциональная характеристика сайта интеграции CHO.

Последовательность нуклеиновой кислоты SEQ ID NO: 1 (и более крупную последовательность нуклеиновой кислоты SEQ ID NO: 4) выявляли эмпирически при помощи последовательностей, расположенных выше и ниже от сайта интеграции конструкции нуклеиновой кислоты (содержащего кассету экспрессии) линии клеток, экспрессирующей белок при высоком уровне. Последовательности нуклеиновой кислоты по настоящему изобретению обеспечивают последовательности новыми функциональными возможностями, связанными с усиленной экспрессией и стабильностью нуклеиновой кислоты (например, экзогенной нуклеиновой кислоты, содержащей GOI), и без ограничения какой-либо теорией, могут функционировать аналогичным или отличным образом от таковых, ранее описанных для элементов, действующих в цис-положении, таких как промоторы, энхансеры, локус-контролирующие области, области прикрепления к ядерному скелету или области прикрепления к матриксу. По-видимому, SEQ ID NO: 1 не имеет открытых рамок считывания (ORF), что делает маловероятным то, что локус кодирует новые белки, действующие в транс-положении. Предполагаемый цинк-пальцевый белок был выявлен в геномном локусе в направлении 3' (ниже) от SEQ ID NO: 4.

Усиливающую экспрессию активность выявляли в отношении интеграции кассеты экспрессии, содержащей первый ген устойчивости к гидромицину (Hug), второй GOI, второй ген Hug, второй GOI, третий ген Hug и третий GOI, кодирующей последовательность в уникальном сайте некодирующей области ДНК генома CHO. Векторы экспрессии, содержащие, например, выделенную 5'-область размером 1 т.о. и выделенную 3'-область размером 1 т.о., выявленные в некодирующей области ДНК генома CHO по отношению к кассете экспрессии, экспрессирующей GOI, были способны обеспечивать трансфицированным клеткам CHO высокие уровни экспрессии рекомбинантных белков.

Настоящее изобретение охватывает векторы экспрессии, содержащие ориентированные в обратном порядке фрагменты SEQ ID NO: 1 или фрагменты SEQ ID NO: 4. Также можно разрабатывать другие комбинации фрагментов, описанных в данном документе. Примеры других комбинаций фрагментов, описанных в данном документе, которые также можно разработать, включают последовательности, которые содержат множественные копии усиливающих экспрессию последовательностей, раскрытых в данном документе, или последовательности, полученные при объединении раскрытых фрагментов SEQ ID NO: 1 или фрагментов SEQ ID NO: 4 с другими нуклеотидными последовательностями, с целью достижения оптимальных комбинаций регуляторных элементов. Такие комбинации могут быть смежно связаны или упорядочены с целью обеспечения оптимального разграничения фрагментов SEQ ID NO: 1 или SEQ ID NO: 4 (например, при помощи введения нуклеотидов спейсеров между фрагментами). Регуляторные элементы также могут быть упорядочены с целью обеспечения оптимального разграничения

фрагмента SEQ ID NO: 1 по отношению к регуляторным элементам.

SEQ ID NO: 1 и SEQ ID NO: 4, раскрытые в данном документе, выделяли из клеток СНО. Было обнаружено, что другие виды млекопитающих (такие как, например, человек или мышь) характеризовались ограниченной гомологией по отношению к выявленной усиливающей экспрессии области, однако гомологичные последовательности можно обнаружить в линиях клеток, полученных из других типов тканей *Cricetulus griseus* или других гомологичных видов, и можно выделить при помощи методик, которые хорошо известны из уровня техники. Например, можно выявить другие гомологичные последовательности при помощи межвидовой гибридизации или методик на основе ПЦР. Кроме того, можно внести изменения в нуклеотидной последовательности, изложенной в SEQ ID NO: 1, SEQ ID NO: 4 или ее фрагментах, при помощи методик сайт-направленного или случайного мутагенеза, которые хорошо известны из уровня техники. Варианты полученной последовательности затем можно исследовать на наличие усиливающей экспрессии активности, как описано в данном документе. ДНК, которые по меньшей мере на приблизительно 90% идентичны по составу нуклеиновых кислот SEQ ID NO: 1, SEQ ID NO: 4 или их фрагментам, имеющим усиливающую экспрессию активность, поддаются выделению при помощи стандартных экспериментов и, как предполагается, характеризуются усиливающей экспрессию активностью. В отношении фрагментов SEQ ID NO: 1 или SEQ ID NO: 4 процентная идентичность относится к той части референтной нативной последовательности, которая обнаружена в фрагменте SEQ ID NO: 1 или фрагменте SEQ ID NO: 4. Соответственно гомологи SEQ ID NO: 1, SEQ ID NO: 4 или их фрагментов и их варианты также охвачены вариантами осуществления по настоящему изобретению.

В соответствии с определенными вариантами осуществления фрагмент SEQ ID NO: 1 выбирают из группы, состоящей из нуклеотидов, охватывающих положения под номерами

10-4000; 100-3900; 200-3800; 300-3700; 400-3600; 500-3500;
600-3400; 700-3300; 800-3200; 900-3100; 1000-3000; 1100-2900; 1200-2800; 1300-2700;
1200-2600; 1300-2500; 1400-2400; 1500-2300; 1600-2200; 1700-2100; 1800-2050; 1850-
2050; 1900-2040; 1950-2025; 1990-2021; 2002-2021 и 2010-2015 SEQ ID NO:1.

В соответствии с другим вариантом осуществления фрагмент SEQ ID NO: 1 выбирают из группы, состоящей из нуклеотидов, охватывающих положения под номерами

10-
500; 500-1000; 500-2100; 1000-1500; 1000-2100; 1500-2000; 1500-2500; 2000-2500;
2500-3000; 2500-3500; 3000-3500; 3000-4000; и 3500-4000 SEQ ID NO:1.

В соответствии с определенными вариантами осуществления экзогенные последовательности нуклеиновой кислоты интегрируются в специфические сайты или возле специфических сайтов в фрагменте, описанном выше.

В соответствии с другими вариантами осуществления последовательность экзогенной нуклеиновой кислоты располагают в SEQ ID NO: 1 или ее фрагментах, как описано выше, или в последовательности, которая по меньшей мере на приблизительно 90% идентична, по меньшей мере на приблизительно 91% идентична, по меньшей мере на приблизительно 92% идентична, по меньшей мере на приблизительно 93% идентична, по меньшей мере на приблизительно 94% идентична, по меньшей мере на приблизительно 95% идентична, по меньшей мере на приблизительно 96% идентична, по меньшей мере на приблизительно 97% идентична, по меньшей мере на приблизительно 98% идентична или по меньшей мере на приблизительно 99% идентична усиливающей экспрессии последовательности SEQ ID NO: 1 или ее усиливающим экспрессию фрагменте.

Можно получить клеточные популяции, экспрессирующие повышенные уровни представляющего интерес белка с применением способов, предусматриваемых в данном документе. Абсолютный уровень экспрессии будет меняться с конкретным белком в зависимости от того, как эффективно белок процессируется в клетке. Клеточные популяции, полученные при помощи экзогенной(экзогенных) последовательности(последовательностей), интегрированной(интегрированных) в усиливающие экспрессию последовательности по настоящему изобретению, стабильны во времени и для большинства целей их можно обрабатывать, как стабильные линии клеток. Стадии рекомбинации также можно задерживать до последующего времени в процессе получения линий клеток по настоящему изобретению.

Усиливающий экспрессию локус СНО и его фрагменты.

Настоящее изобретение охватывает усиливающий экспрессию фрагмент нуклеотидной последовательности, которая по меньшей мере на приблизительно 90% идентична, по меньшей мере приблизительно 91% идентична, по меньшей мере приблизительно 92% идентична, по меньшей мере приблизительно 93% идентична, по меньшей мере приблизительно 94% идентична, по меньшей мере приблизительно 95% идентична, по меньшей мере приблизительно 96% идентична, по меньшей мере приблизительно 97% идентична, по меньшей мере приблизительно 98% идентична или по меньшей мере приблизительно 99% идентична нуклеотидной последовательности SEQ ID NO: 1 или SEQ ID NO: 4. Настоящее изобретение включает векторы, содержащие фрагмент, в том числе для временной или стабильной трансфекции, охватывающие положения под номерами

10-4000; 100-3900; 200-3800;
 300-3700; 400-3600; 500-3500; 600-3400; 700-3300; 800-3200; 900-3100; 1000-3000;
 1100-2900; 1200-2800; 1300-2700; 1200-2600; 1300-2500; 1400-2400; 1500-2300; 1600-
 2200; 1700-2100; 1800-2050; 1850-2050; 1900-2040; 1950-2025; 1990-2021; 2002-2021 и
 2010-2015 SEQ ID NO:1.

Настоящее изобретение также включает эукариотическую клетку, содержащую такой фрагмент, где фрагмент является экзогенным по отношению к клетке и интегрируется в геном клетки, и клетки, содержащие такой фрагмент, имеющий по меньшей мере один сайт узнавания для рекомбиназы, который находится во фрагменте, непосредственно в направлении 5' или непосредственно в направлении 3' к фрагменту.

В соответствии с одним вариантом осуществления усиливающий экспрессию фрагмент SEQ ID NO: 1 расположен в положении в SEQ ID NO: 1, охватывающем положения под номерами

10-500; 500-1000; 500-2100; 1000-1500;
 1000-2100; 1500-2000; 1500-2500; 2000-2500; 2500-3000; 2500-3500; 3000-3500; 3000-
 4000 или 3500-4000 SEQ ID NO:1.

Если поддерживается стабильная интеграция и/или повышенная транскрипция интегрированных полинуклеотидов, то точное расположение сайта вставки в локус (т.е. интеграции) по отношению к иллюстративным сайтам не является обязательным. Более того, сайт интеграции может находиться в любом положении, которое находится в или рядом с SEQ ID NO: 1 или фрагментом SEQ ID NO: 1 или SEQ ID NO: 4 или фрагментом SEQ ID NO: 4, как описано в данном документе. Поддерживает ли определенное расположение в хромосоме в представляющем интерес локусе или рядом с представляющим интерес локусом стабильную интеграцию и эффективную транскрипцию интегрированного экзогенного гена можно определить согласно стандартным процедурам, хорошо известным из уровня техники, или способам, приведенным в качестве примера, в данном документе.

Сайты интеграции, рассматриваемые в данном документе, расположены в локусе, содержащем нуклеотидную последовательность SEQ ID NO: 1 или SEQ ID NO: 4, или в непосредственной близости от представляющего интерес локуса, например, на менее приблизительно 1 т.о., 500 пар оснований (п.о.), 250 п.о., 100 п.о., 50 п.о. 25 п.о., 10 п.о. или менее приблизительно 5 п.о. выше (5') или ниже (3') по отношению к расположению SEQ ID NO: 1 в хромосомной ДНК. В соответствии с еще одними вариантами осуществления используемый сайт интеграции расположен на приблизительно 1000, 2500, 5000 или более пар оснований выше (5') или ниже (3') по отношению к расположению SEQ ID NO: 1 или SEQ ID NO: 4 в хромосомной ДНК.

В данной области техники подразумевается, что крупные области генома, такие как области прикрепления к ядерному скелету/матриксу, используются для эффективной репликации и транскрипции хромосомной ДНК. Область прикрепления к ядерному скелету/матриксу (S/MAR), также называемая областью прикрепления к ядерному скелету (SAR), или матрикс-ассоциированной областью или областью прикрепления к матриксу (MAR), представляет собой область ДНК генома эукариот, где прикрепляется ядерный матрикс. Без ограничения какой-либо теорией, S/MAR обычно располагаются в некодирующих областях, отделяют определенную область транскрипции (например, хроматиновый домен) от соседних областей, а также предоставляют площадки для производства и/или связывания факторов, которые способствуют транскрипции, такие как сайты узнавания ДНКаз или полимераз. Были описаны некоторые S/MAR, приблизительно 14-20 т.о. в длину (Klar, et al., 2005, Gene 364:79-89). В связи с этим предполагается, что интеграция генов в LOCUS 1 (в или возле SEQ ID NO: 1 или SEQ ID NO: 4) способствует усиленной экспрессии.

Специалистам в данной области будет понятно, что с целью высокой транскрипционной активности в исследуемом локусе можно оптимизировать несколько элементов, приводя к высокой экспрессии вставленного гена, кодирующего представляющий интерес белок. Рассматриваемые элементы включают сильный промотор для контроля транскрипции, соответствующую транскрипционную систему и ДНК, имеющую открытую и доступную конфигурацию. Вставку в исследуемый локус можно оптимизировать в пределах компетенции специалиста в данной области при направлении сайта интеграции, выбранного в SEQ ID NO: 1 или SEQ ID NO: 4.

В соответствии с одним вариантом осуществления для усиления экспрессии GOI используют усиливающую экспрессию последовательность SEQ ID NO: 1. На фиг. 2A представлены результаты оперативного связывания GOI с SEQ ID NO: 1 (LOCUS 1) по сравнению с тем же самым GOI, интегрированным в другой локус в геноме клеток CHO (контрольный локус). Число копий гена, измеренное для каждой линии клеток, одинаково, однако эксперименты показывают, что уровень мРНК и титр белка клеток, экспрессирующих GOI, в 3 раза выше для GOI, функционально связанного с LOCUS 1.

В соответствии с различными вариантами осуществления экспрессию GOI можно усилить при расположении GOI в SEQ ID NO: 1 или SEQ ID NO: 4. В соответствии с различными вариантами осуществления усиление экспрессии составляет по меньшей мере от приблизительно 1,5-кратного до прибли-

тельно 3-кратного или более.

Генетическая модификация целевого локуса.

Способы генетического конструирования генома клетки в определенном расположении (т.е. целевом локусе) можно осуществлять несколькими путями. Методики генетического редактирования использовали с целью стабильной интеграции последовательности нуклеиновой кислоты в эукариотическую клетку, где последовательность нуклеиновой кислоты представляет собой экзогенную последовательность, обычно не встречающуюся в таких клетках. Клональная экспансия необходима для обеспечения того, что потомство клетки будет иметь общие генотипические и фенотипические характеристики с работанной линией клеток. В некоторых примерах нативные клетки модифицируют при помощи методики гомологичной рекомбинации с целью интеграции экзогенной последовательности нуклеиновой кислоты в SEQ ID NO: 1 или SEQ ID NO: 4. В другом примере предусматриваются клетки, которые содержат по меньшей мере одну последовательность узнавания для рекомбиназы в SEQ ID NO: 1 или SEQ ID NO: 4, подходящую для интеграции экзогенной последовательности нуклеиновой кислоты или гена, представляющего интерес.

В некоторых примерах предусматриваются клетки, которые содержат первую последовательность узнавания для рекомбиназы и вторую последовательность узнавания для рекомбиназы, где каждую из первой и второй последовательностей узнавания для рекомбиназы выбирают из группы, состоящей из LoxP, Lox511, Lox5171, Lox2272, Lox2372, Loxm2, Lox-FAS, Lox71, Lox66 и их мутантных форм. В этом случае, если желателен обмен кассетами, опосредованный рекомбиназой (RMCE), то сайт-специфической рекомбиназой является Cre-рекомбиназа или ее производное. В другом примере каждую из первой и второй последовательностей узнавания для рекомбиназы выбирают из группы, содержащей FRT, F3, F5, мутантную форму -10 FRT, мутантную форму +10 FRT и их мутантные формы, и согласно этому сценарию, если желателен RCME, то сайт-специфической рекомбиназой является рекомбиназа Flp или ее производное. В еще одном примере каждую из указанных первой и второй последовательностей узнавания для рекомбиназы выбирают из группы, содержащей attB, attP и их мутантные формы, и в этом случае, если желателен RMCE, то сайт-специфической рекомбиназой является интеграза *phi* C31 или ее производное.

В соответствии с одним аспектом способы и композиции для стабильной интеграции последовательности нуклеиновой кислоты в SEQ ID NO: 1 или SEQ ID NO: 4 или ее усиливающего экспрессию фрагмента реализовывают при помощи гомологичной рекомбинации. Молекулу нуклеиновой кислоты, т.е. представляющий интерес ген или полинуклеотид, можно вставить в целевой локус (т.е. SEQ ID NO: 1) при помощи гомологичной рекомбинации или при помощи способов с использованием сайт-специфической нуклеазы, в результате которых происходит специфическое направление последовательностей в сайтах интеграции. При гомологичной рекомбинации гомологичные полинуклеотидные молекулы (т.е. гомологичные области) выстраивают и обменивают область из их последовательностей. Трансген можно вводить во время такого обмена, когда трансген фланкирован гомологичными последовательностями генома. В одном примере сайт узнавания для рекомбиназы можно вводить в геном клетки-хозяина в сайтах интеграции.

Гомологичную рекомбинацию в эукариотических клетках можно облегчать при введении разрыва в сайте интеграции хромосомной ДНК. На модельных системах продемонстрировано, что частота гомологичных рекомбинаций во время направленного воздействия на ген увеличивается, если двухцепочечный разрыв введен в хромосомную последовательность-мишень. Это может быть выполнено при направлении определенных нуклеаз в специфические сайты интеграции. ДНК-связывающие белки, которые распознают последовательности ДНК в целевом локусе, известны из уровня техники. Векторы, направленные на ген, также используют для облегчения гомологичной рекомбинации. При отсутствии вектора, направленного на ген, для направленного восстановления гомологии, клетки зачастую закрывают двухцепочечный разрыв путем нехомологичного соединения концов (NHEJ), что может привести к делеции или вставке нескольких нуклеотидов в сайт расщепления. В случае если возникают вставки или делеции (InDels), фактически небольшое число нуклеотидов либо вставляют, либо удаляют случайным образом в сайт разрыва и эти InDels могут сдвигать или разрывать любую открытую рамку считывания (ORF) гена в целевом локусе. Следует понимать, что локус, выявляемый в виде SEQ ID NO: 1 (или SEQ ID NO: 4), не является областью, кодирующей ген. Поэтому нарушения транскрипции эндогенного гена при вставке и/или делеции в этом локусе не предполагается.

Направленное восстановление гомологии (или направленная рекомбинация гомологии) (HDR) является особо пригодным для вставки или интеграции генов в исследуемый локус. Донорская конструкция содержит гомологичные области, полученные из SEQ ID NO: 1 или SEQ ID NO: 4, как описано в данном документе.

Конструкция вектора, направленного на ген, и селекция нуклеазы находятся в пределах квалификации специалиста в данной области техники, к которой принадлежит настоящее изобретение.

В некоторых примерах цинк-пальцевые нуклеазы (ZFN), которые характеризуются модулярной структурой и содержат отдельные цинк-пальцевые домены, распознают конкретную 3-нуклеотидную последовательность в последовательности-мишени (например, в сайте направленной интеграции). В не-

которых вариантах осуществления можно использовать ZFN в комбинации с отдельными доменами цинковых пальцев, нацеленных на множество последовательностей-мишеней.

Эффекторные нуклеазы (TALEN), подобные активатору транскрипции (TAL), также можно использовать для сайт-специфического изменения генома. ДНК-связывающий домен белка TAL-эффектора, как правило, используют в комбинации с неспецифическим доменом расщепления нуклеазы рестрикции, такой как FokI. В некоторых вариантах осуществления белок слияния, содержащий ДНК-связывающий домен белка TAL-эффектора и домен расщепления нуклеазы рестрикции, используют для распознавания и расщепления ДНК в локусе последовательности-мишени по настоящему изобретению (Boch J. et al., 2009 Science 326:1509-1512).

РНК-управляемые эндонуклеазы (RGEN) представляют собой программируемые инструменты для геномной инженерии, которые разработаны благодаря бактериальному адаптивному иммунному механизму. В данной системе короткие палиндромные повторы, регулярно расположенные группами, (CRISPR)/CRISPR-ассоциированный (Cas) иммунный ответ, представляют собой белок Cas9, который формирует эндонуклеазу, специфическую к определенной последовательности при образовании комплекса с двумя РНК, одна из которых направлена на выбор мишени. RGEN включают компоненты (Cas9 и tracrRNA) и направленные CRISPR РНК (crRNA). Как эффективность направленного расщепления ДНК, так и расположение сайтов расщепления меняется на основании положения мотива, примыкающего к протоспейсеру (PAM), дополнительно необходимого для целевого узнавания (Chen, H. et al., J. Biol. Chem., опубликованная онлайн 14 марта 2014 г., как Manuscript M113.539726).

Стратегии выявления последовательностей, уникальных для определенного целевого локуса SEQ ID NO: 1, известны из уровня техники, однако при выравнивании многих из этих последовательностей по отношению к геному CHO обнаруживаются потенциальные нецелевые сайты с совпадением в 16-17 пар оснований. Одна иллюстративная ведущая РНК размером 20 п.о., кодируемая последовательностью, изложенной в SEQ ID NO: 5 (соответствующая 1990-2001 нуклеотиду из SEQ ID NO: 1) является пригодной для РНК-управляемого CRISPR/Cas редактирования генов SEQ ID NO: 1 или SEQ ID NO: 4. Плазмиду, содержащую промотор, который контролирует экспрессию малой ведущей РНК и tracrRNA (например, SEQ ID NO: 6), а также несущую подходящий фермент Cas9 под контролем промотора, можно ко-трансфицировать донорным вектором (несущим ген, представляющий интерес, фланкированный 5'- и 3'-областями гомологии) с целью направленной интеграции этим способом. Различные модификации и варианты молекул РНК, помимо описанных выше в данном документе, очевидны специалистам в данной области техники и находятся в пределах объема настоящего изобретения.

В соответствии с некоторыми вариантами осуществления средство для переноса, предназначенное для введения в геном экзогенной нуклеиновой кислоты, содержащей последовательность, кодирующую ген, представляющий интерес, или последовательность узнавания, или генную кассету, в зависимости от ситуации, содержит вектор, несущий экзогенную нуклеиновую кислоту, и один или более дополнительных векторов или мРНК. В соответствии с одним вариантом осуществления один или более дополнительных векторов или мРНК содержат нуклеотидную последовательность, кодирующую сайт-специфическую нуклеазу, в том числе без ограничений цинк-пальцевую нуклеазу (ZFN), димер ZFN, эффекторную нуклеазу, подобную активаторам транскрипции (TALEN), белок слияния на основе домена TAL-эффектора или РНК-управляемую ДНК-эндонуклеазу. В соответствии с определенными вариантами осуществления один или более векторов или мРНК содержат первый вектор, содержащий ведущую РНК, tracrRNA и нуклеотидную последовательность, кодирующую фермент Cas, и второй вектор, содержащий донорную (экзогенную) нуклеотидную последовательность. Такая донорная последовательность содержит нуклеотидную последовательность, кодирующую ген, представляющий интерес, или последовательность узнавания, или генную кассету, содержащую любой из этих экзогенных элементов, предусмотренных для направленной вставки. Если используется мРНК, то мРНК можно трансфицировать в клетку при помощи обычных способов трансфекции, известных специалистам в данной области техники, и она может кодировать фермент, например транспозазу или эндонуклеазу. Несмотря на то, что мРНК, вводимая в клетки, может быть временной и не интегрируется в геном, мРНК может нести экзогенную нуклеиновую кислоту, необходимую или эффективную для осуществления интеграции. В некоторых примерах мРНК выбирают с целью устранения любого риска длительных побочных эффектов добавочного полинуклеотида, если требуется лишь кратковременная экспрессия для достижения желательной интеграции GOI.

Специалистам в данной области техники доступны и другие способы гомологичной рекомбинации, такие как BuD-получение нуклеаз (BuDN) с определенной ДНК-связывающей специфичностью (Stella, S. et al., Acta Cryst. 2014, D70, 2042-2052). Выбор способов точной модификации генома основан на доступных инструментах, совместимых с уникальными целевыми последовательностями в SEQ ID NO: 1, поэтому избегают нарушения фенотипа клеток.

Конструкции для направленности на гены.

Полинуклеотидная последовательность, подлежащей интеграции в геном хозяина, может быть любая пригодная в промышленном отношении последовательность ДНК, такая как последовательность узнавания, для создания клеточных систем экспрессии. Полинуклеотидная последовательность, подле-

жащая интеграции в геном хозяина, может кодировать любой терапевтически пригодный или пригодный в промышленном отношении белок или белки, как описано в данном документе. Выявление целевой последовательности в целевом локусе с целью интеграции экзогенной последовательности нуклеиновой кислоты зависит от ряда факторов. В зависимости от используемого способа гомологической рекомбинации в пределах компетенции специалиста в данной области находится выбор последовательностей, гомологичных SEQ ID NO: 1 или SEQ ID NO: 4. Векторы с сайт-специфическими нуклеазами при использовании требуют дополнительных компонентов (композиций последовательностей), которые распознают специфический сайт, предусмотренный для разрыва ДНК.

В связи с этим конструкция для направленности на гены включает такие нуклеотидные последовательности, которые облегчают направленную интеграцию такой экзогенной последовательности нуклеиновой кислоты в представляющий интерес локус. В соответствии с некоторыми вариантами осуществления конструкция содержит первую гомологичную область и вторую гомологичную область. В соответствии с другими вариантами осуществления конструкция (например, генная кассета) содержит гомологичные области, полученные из SEQ ID NO: 1 или SEQ ID NO: 4. В соответствии с некоторыми вариантами осуществления области гомологии содержат нуклеотидную последовательность, гомологичную нуклеотидной последовательности, присутствующей в SEQ ID NO: 1 или SEQ ID NO: 4. В соответствии с конкретными вариантами осуществления конструкция содержит 5'-область гомологии, имеющую нуклеотидную последовательность SEQ ID NO: 2 (соответствующую 1001-2001 нуклеотиду SEQ ID NO: 1), и 3'-область гомологии, имеющую нуклеотидную последовательность SEQ ID NO: 3 (соответствующую 2022-2001 нуклеотиду SEQ ID NO: 1). Гомологичные области, например первая гомологичная область (также называемая 5'-областью гомологии) и вторая гомологичная область (также называемая 3'-областью гомологии) являются гомологичными целевой последовательности в локусе. Гомологичные области в направлении от 5' до 3' могут распространяться на область или целевую последовательность в локусе, которая содержит по меньшей мере 1 т.о., или по меньшей мере приблизительно 2 т.о., или по меньшей мере приблизительно 3 т.о., или по меньшей мере приблизительно 4 т.о., или по меньшей мере 5 т.о., или по меньшей мере приблизительно 10 т.о. В соответствии с другими вариантами осуществления суммарное число нуклеотидов целевой последовательности, выбранной в качестве первой и второй гомологичных областей, содержит по меньшей мере 1 т.о., или по меньшей мере приблизительно 2 т.о., или по меньшей мере приблизительно 3 т.о., или по меньшей мере приблизительно 4 т.о., или по меньшей мере 5 т.о., или по меньшей мере приблизительно 10 т.о. В некоторых примерах расстояние между 5'-областью гомологии и 3'-областью гомологии (гомологичными целевой последовательности) составляет по меньшей мере 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900 п.о. или по меньшей мере 1 т.о., или по меньшей мере приблизительно 2 т.о., или по меньшей мере приблизительно 3 т.о., или по меньшей мере приблизительно 4 т.о., или по меньшей мере 5 т.о., или по меньшей мере приблизительно 10 т.о. В примерах, где SEQ ID NO: 2 и SEQ ID NO: 3 выбирают в качестве 5'- и 3'-областей гомологии, расстояние между двумя областями гомологии может составлять 20 нуклеотидов (соответствующих нуклеотидам 2002-2021 SEQ ID NO: 1); и такие области гомологии могут опосредовать интеграцию экзогенной последовательности нуклеиновой кислоты в локус, содержащем SEQ ID NO: 1, например, в нуклеотидах 1990-2021 или 2002-2021 SEQ ID NO: 1, и одновременную делецию нуклеотидов 2002-2021 SEQ ID NO: 1.

В соответствии с другими вариантами осуществления конструкция содержит первую гомологичную область и вторую гомологичную область, где первая и вторая гомологичные области совместно содержат целевую последовательность, которая замещает эндогенную последовательность в локусе. В соответствии с еще одними вариантами осуществления первая и вторая гомологичные области содержат целевую последовательность, которая интегрируется или вставляется в эндогенную последовательность в локусе.

Модифицированные линии клеток создавали при помощи интеграции одного или более сайтов узнавания для рекомбиназы в положение в SEQ ID NO: 1. Эти модифицированные линии клеток могут также включать дополнительные экзогенные гены для негативного или позитивного отбора экспрессируемого гена, представляющего интерес.

Настоящее изобретение обеспечивает способы модификации генома клетки CHO, предусматривающие введение одного или более средств для переноса в клетку, где одно или более средств для переноса содержат экзогенную нуклеиновую кислоту, содержащую последовательность для интеграции, 5'-область гомологии, гомологичную последовательности, присутствующей в нуклеотидной последовательности SEQ ID NO: 1, и 3'-область гомологии, гомологичную последовательности, присутствующей в нуклеотидной последовательности SEQ ID NO: 1. В соответствии с некоторыми вариантам осуществления способы дополнительно предусматривают одно или более средств для переноса, содержащих нуклеазу и композиции для сайт-специфического расщепления ДНК в сайте интеграции.

Модифицированные линии клеток можно использовать в качестве удобных и стабильных систем экспрессии для обмена кассетами, опосредованного рекомбиназой (RMCE). Последовательность нуклеиновой кислоты, кодирующей представляющий интерес белок, можно подходящим образом интегрировать в модифицированную клетку, содержащую SEQ ID NO: 1 или ее усиливающий экспрессию фрагмент, имеющий по меньшей мере один сайт узнавания для рекомбиназы, например, при помощи способа

RMCE.

Рекомбинантные векторы экспрессии могут содержать синтетические или полученные из cDNA фрагменты ДНК, кодирующие белок, функционально связанные с подходящим транскрипционным и/или трансляционным регуляторным элементом, полученным из генов млекопитающих, вирусов или насекомых. Такие регуляторные элементы включают транскрипционные промоторы, энхансеры, последовательности, кодирующие подходящие сайты связывания рибосом мРНК и последовательности, которые контролируют терминацию транскрипции и трансляции, описанную подробно ниже. Векторы экспрессии млекопитающих могут также содержать нетранслируемые элементы, такие как точка начала репликации, другие 5'- или 3'-фланкирующие нетранслируемые последовательности, и 5' или 3' нетранслируемые последовательности, такие как донор сплайсированного фрагмента и акцепторный сайт. Также можно включить селективный маркерный ген для облегчения узнавания трансфектантов.

Флуоресцентные маркеры представляют собой селективные маркерные гены для распознавания генных кассет, которые были или не были успешно вставлены и/или замещены, в зависимости от ситуации. Примеры флуоресцентных маркеров хорошо известны из уровня техники, включая без ограничений флуоресцентный красный белок, выделенный из *Discosoma* (DsRed), зеленый флуоресцентный белок (GFP), усиленный зеленый флуоресцентный белок (eGFP), голубой флуоресцентный белок (CFP), усиленный голубой флуоресцентный белок (eCFP), желтый флуоресцентный белок (YFP), усиленный желтый флуоресцентный белок (eYFP) и флуоресцентный белок дальнего красного спектра (например, mKate, mKate2, mPlum, mRaspberry или E2-crimson). См. также, например, Nagai, T., et al., 2002 *Nature Biotechnology* 20:87-90; Heim, R. et al., 23 February 1995 *Nature* 373:663-664 и Strack, R.L. et al., 2009 *Biochemistry* 48:8279-81.

Из вирусных источников можно обеспечить транскрипционные и трансляционные контролирующие последовательности в векторах экспрессии, пригодных для трансфицирования клеток позвоночных. Например, обычно применяемые промоторы и энхансеры получают из вирусов, таких как полиома, аденовирус 2, вирус обезьян 40 (SV40) и цитомегаловирус человека (CMV). Вирусные геномные промоторы, контролирующие и/или сигнальные последовательности можно использовать для контроля экспрессии, при условии, что такие контролирующие последовательности являются совместимыми с выбранной клеткой-хозяином. В зависимости от типа клетки, в которой рекомбинантный белок должен экспрессироваться, можно также применять невирусные клеточные промоторы (например, β -глобин и EF-1 α промоторы).

Последовательности ДНК, полученные из вирусного генома SV40, например, точки начала репликации SV40, ранний и поздний промотор, энхансер, сплайсер и сайты полиаденилирования, можно применять для обеспечения других генетических элементов, пригодных для экспрессии гетерологичной последовательности ДНК. Ранние и поздние промоторы, как правило, пригодны в связи с тем, что их обоих легко получают из вируса SV40 в качестве фрагмента, который также содержит вирусную точку начала репликации SV40 (Fiers et al., *Nature* 273:113, 1978). Можно также применять более мелкие или более крупные фрагменты SV40. Как правило, последовательность, включающая примерно 250 п.о., простирается от сайта Hind III до сайта BglII, расположенного в точке начала репликации SV40.

Бицистронные векторы экспрессии, используемые для экспрессии нескольких транскриптов, были описаны ранее (Kim S.K. and Wold B.J., *Cell* 42:129, 1985) и их можно использовать в комбинации с усиливающей экспрессию последовательностью по настоящему изобретению, например SEQ ID NO: 1, или ее фрагментом. Другие типы векторов экспрессии будут также пригодны, например, описанные в патенте США № 4634665 (Axel et al.) и патенте США № 4656134 (Ringold et al.).

Представляющие интерес белки.

Можно использовать любой представляющий интерес белок, пригодный для экспрессии в эукариотических клетках. Например, представляющий интерес белок включает без ограничений антитело или его антигенсвязывающий фрагмент, химерное антитело или его антигенсвязывающий фрагмент, ScFv или его фрагмент, белок слияния, содержащий Fc-фрагмент, или его фрагмент, фактор роста или его фрагмент, цитокин или его фрагмент или внеклеточный домен рецептора клеточной поверхности или его фрагмент. Представляющие интерес белки могут быть простыми полипептидами, состоящими из одной субъединицы, или сложными белками из нескольких субъединиц, содержащими две или более субъединицы.

Клетки-хозяева и трансфекция.

Клетками-хозяевами, используемыми в способах по настоящему изобретению, являются клетки-хозяева млекопитающих, в том числе, например, клетки яичника китайского хомячка (CHO) и клетки мышцы. В соответствии с предпочтительным вариантом осуществления настоящее изобретение обеспечивает фрагмент последовательности нуклеиновой кислоты SEQ ID NO: 1, которая кодирует усиливающую экспрессию последовательность в клетке CHO. Сайт интеграции может встречаться в SEQ ID NO: 1 или любом фрагменте SEQ ID NO: 1. Сайт интеграции, например, может представлять собой сайт узнавания для рекомбиназы, расположенный в SEQ ID NO: 1 или любом фрагменте SEQ ID NO: 1. Одним примером подходящего сайта интеграции является сайт LoxP. Другим примером подходящего сайта интеграции являются два сайта узнавания для рекомбиназы, например, выбранные из группы, состоящей из сай-

та LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Loxm2, сайта Lox71, сайта Lox66 и сайта Lox5171. В соответствии с другими вариантами осуществления сайт интеграции расположен в некотором положении в последовательности или рядом с некоторым положением в последовательности, выбранной из группы, состоящей из нуклеотидов, охватывающих положения под номерами

10-4000; 100-3900; 200-3800; 300-3700; 400-3600; 500-3500; 600-3400; 700-3300; 800-3200; 900-3100; 1000-3000; 1100-2900; 1200-2800; 1300-2700; 1200-2600; 1300-2500; 1400-2400; 1500-2300; 1600-2200; 1700-2100; 1800-2050; 1850-2050, 1900-2040; 1950-2025, 1990-2021, 2002-2021 и 2010-2015 в SEQ ID NO:1.

В соответствии с определенными вариантами осуществления сайт интеграции в некотором положении в SEQ ID NO: 1 или рядом с некоторым положением в SEQ ID NO: 1 выбирают из группы, состоящей из нуклеотидов, охватывающих положения под номерами

1990-1991; 1991-1992; 1992-1993; 1993-1994; 1995-1996; 1996-1997; 1997-1998; 1999-2000; 2001-2002; 2002-2003; 2003-2004; 2004-2005; 2005-2006; 2006-2007; 2007-2008; 2008-2009; 2009-2010; 2010-2011; 2011-2012; 2012-2013; 2013-2014; 2014-2015; 2015-2016, 2016-2017; 2017-2018, 2018-2019, 2019-2020 и 2020-2021 в SEQ ID NO:1.

Настоящее изобретение включает клетку-хозяина млекопитающего, трансфицированную вектором экспрессии или мРНК по настоящему изобретению. Если используют какую-либо клетку млекопитающих, то в соответствии с определенным вариантом осуществления клеткой-хозяином является клетка CHO.

Трансфицированные клетки-хозяева включают клетки, которые были трансфицированы векторами экспрессии или молекулами мРНК, которые содержат последовательность, кодирующую белок или полипептид. Экспрессируемые белки могут быть секретированы в среде культивирования в зависимости от выбранной последовательности нуклеиновой кислоты, а могут оставаться в клетке или откладываться в клеточной мембране. Для экспрессии рекомбинантных белков можно применять различные системы культуры клеток млекопитающих. Другие линии клеток, разработанные для схем специфического отбора и амплификации, также будут пригодны согласно способам и композициям, предусматриваемым в данном документе, при условии, что был выявлен целевой локус, по меньшей мере на 80% гомологичный SEQ ID NO: 1. Предусматриваемой в вариантах осуществления линией клеток является линия клеток CHO, обозначенная K1. Для достижения массового производства рекомбинантных белков в подходящем случае линию клетки-хозяина можно предварительно адаптировать к среде биореактора.

Несколько протоколов трансфекции известно из уровня техники и описано в Kaufman (1988) Meth. Enzymology 185:537. Выбранный протокол трансфекции будет зависеть от типа клетки-хозяина и природы GOI и его можно выбрать на основании стандартного эксперимента. Основными требованиями любого такого протокола являются, во первых, введение ДНК, кодирующей представляющий интерес белок в подходящую клетку-хозяина, и затем определение и выделение клеток-хозяев, в которых включена гетерологичная ДНК, относительно стабильным экспрессируемым образом. Молекулы мРНК, кодирующие белки, пригодные для интеграции в геном клетки-хозяина или другой функции, могут быть временными и, таким образом, ограниченными во времени.

Протоколы трансфекции, а также протоколы введения полипептидных или полинуклеотидных последовательностей в клетки, могут варьировать. Неограничивающие способы трансфекции включают способы химической трансфекции, включающие применение липосом; наночастиц; фосфата кальция (Graham et al. (1973). Virology 52 (2): 456-67, Bacchetti et al. (1977) Proc Natl Acad Sci USA 74 (4): 1590-4, и Kriegler, M. (1991). Transfer and Expression: A Laboratory Manual. New York: W.H. Freeman and Company, pp. 96, 97); дендримеров или катионных полимеров, таких как DEAE-декстран или полиэтиленмин. Нехимические способы включают электропорацию, сонопорацию и оптическую трансфекцию. Трансфекция на основе частиц включает применение генной пушки, магнитную трансфекцию (Bertram, J. (2006) Current Pharmaceutical Biotechnology 7, 277-28). Для трансфекции также можно применять способы на основе вирусов. Доставка мРНК включает способы с использованием TransMessenger™ и TransIT® (Bire et al., BMC Biotechnology 2013, 13:75).

Одним широко применяемым способом введения гетерологичной ДНК в клетку является осаждение фосфатом кальция, например, как описано в Wigler et al. (Proc. Natl. Acad. Sci. USA 77:3567, 1980). ДНК, вводимая в клетку-хозяина этим способом, часто подвергается перестройке, делая эту процедуру пригодной для котрансфекции независимых генов.

Полиэтилен-индуцированное слияние бактериальных протопластов с клетками млекопитающих (Schaffner et al. (1980) Proc. Natl. Acad. Sci. USA 77:2163) представляет собой другой пригодный способ введения гетерологичной ДНК. Протокол слияния протопластов часто приводит к образованию нескольких копий плазмидной ДНК, интегрированной в геном клетки-хозяина млекопитающего, и данная мето-

дика требует, чтобы маркер отбора и амплификации находился в одной и той же плазмиде в виде GOI.

Для введения ДНК непосредственно в цитоплазму клетки-хозяина можно также использовать электропорацию, например, как описано в Potter et al. (Proc. Natl. Acad. Sci. USA 81:7161, 1988) или в Shigekawa et al. (BioTechniques 6:742, 1988). В отличие от слияния протопласта, электропорация не требует, чтобы маркер отбора и GOI находились в одной и той же плазмиде.

Были описаны другие реактивы, пригодные для введения гетерологичной ДНК в клетку млекопитающего, такие как реактив Lipofectin™ и реактив Lipofectamine™ (Gibco BRL, Гейтерсберг, Мэриленд). Оба этих коммерчески доступных реактива применяются для образования комплексов из липида и нуклеиновой кислоты (или липосом), которые облегчают захват нуклеиновой кислоты в клетки при нанесении на культивируемые клетки.

В соответствии с одним вариантом осуществления введение одного или более полинуклеотидов в клетки опосредуется электропорацией, внутрицитоплазматической инъекцией, вирусной инфекцией, аденовирусом, лентивирусом, ретровирусом, трансфекцией, опосредованной липидами трансфекцией или опосредуется Nucleofection™.

Способ амплификации GOI также желателен для экспрессии рекомбинантного белка и обычно он включает применение маркера отбора (описано в Kaufman, выше). Устойчивость к токсическим лекарственным средствам является характеристикой, наиболее часто используемой в качестве маркера отбора, и может являться результатом либо доминантного признака (например, может использоваться независимо от типа клетки-хозяина), либо рецессивного признака (например, пригодна в определенных типах клеток-хозяев, которые характеризуются недостаточностью той активности, в отношении которой проводят отбор). Несколько амплифицируемых маркеров пригодны для применения в векторах экспрессии по настоящему изобретению (например, как описано в Sambrook, Molecular Biology: A Laboratory Manual, Cold Spring Harbor Laboratory, NY, 1989; pgs 16.9-16.14).

Пригодные селективные маркеры для амплификации генов в устойчивых к лекарственным средствам клетках представлены в таблице Kaufman, R.J., выше и предусматривают устойчивость к DHFR-MTX, Р-гликопротеин и различные липофильные цитотоксические средства с множественной устойчивостью к лекарственным средствам (MDR) (например, адриамицин, колхицин, винкристин) и аденозин-деаминазу (ADA)-Xyl-A или аденозин и 2'-дезоксикоформин.

Другие доминантные селективные маркеры включают выделенные из микроорганизмов гены устойчивости к антибиотикам, например устойчивости к неомицину, канамицину или гидромицину. Однако не было показано, что эти маркеры отбора являются амплифицируемыми (Kaufman, R.J., выше). Несколько подходящих систем отбора существуют для клеток млекопитающих (Sambrook, выше, pgs 16.9-16.15). Также были описаны протоколы котрансфекции, использующие два доминантных селективных маркера (Okayama and Berg, Mol. Cell Biol 5:1136, 1985).

Пригодные регуляторные элементы, описанные ранее или известные из уровня техники, можно также включать в конструкции нуклеиновых кислот, применяемые для трансфекции клеток млекопитающих. Выбранный протокол трансфекции и элементы, выбранные для применения в данном документе, будут зависеть от используемого типа клетки-хозяина. Специалисты в данной области техники знают много различных протоколов и клеток-хозяев и могут выбрать подходящую систему для экспрессии требуемого белка на основании требований к применяемой системе культуры клеток.

Другие признаки по настоящему изобретению будут становиться очевидными в следующих описаниях иллюстративных вариантов осуществления, которые даны в иллюстративных целях по настоящему изобретению и не предназначены для его ограничения.

Примеры

Следующие примеры изложены с тем, чтобы обеспечить специалистов в данной области техники тем, как осуществлять и применять способы, а также получать и применять композиции, описанные в данном документе, и при этом они не предполагают ограничение объема настоящего изобретения. Были приложены усилия для обеспечения точности по отношению к применяемым числам (например, количеству, температуре и т.п.), однако, необходимо учитывать некоторые экспериментальные ошибки и отклонения. Если не указано иное, то части являются частями по массе, молекулярная масса является средней молекулярной массой, температура представлена в градусах Цельсия, а давление является атмосферным или близким к атмосферному.

Пример 1. Выявление представляющего интерес локуса и характеристика сайтов интеграции.

Клетки K1 CHO трансфицировали двумя плазмидами, содержащими гены последовательностей антител и селективные гены устойчивости к антибиотикам в качестве селективных маркеров. Отбор стабильных трансфектантов выполняли при помощи экспансии клеток в присутствии антибиотиков. Отдельные клоны клеток, экспрессирующих высокие уровни антител, выделяли при помощи технологии сортировки FASTR® (см. патент США № 8673589 B2). Выявляли несколько клонов, характеризующихся самыми высокими уровнями экспрессии антител.

Геномную ДНК из этих клонов фрагментировали при помощи технологии Covaris Adaptive Focused Acoustics (AFA)™ (Fisher, S. et al., 2011, Genome Biology 12:R1). Создавали библиотеки ДНК (Agilent

SureSelectXT #G9612A) и инкубировали с разработанными в индивидуальном порядке биотинилированными РНК-приманками (Agilent SureSelectXT #5190-4811), разработанными в отношении целых последовательностей плазмид, которые вводили в клетки CHO. Фрагменты геномной ДНК, которые содержали плазмидные последовательности, обогащали магнитными гранулами со стрептавидином и подвергали секвенированию при помощи Illumina MiSeq с целью выявления сайтов интеграции плазмид. Последовательности слияния, которые содержали как последовательность плазмид, так и последовательность генома CHO, анализировали и выравнивали по отношению к геному CHO. Один сайт интеграции подтверждали при помощи Саузерн-блоттинга и ПЦР с последующим секвенированием. Сайт интеграции, имеющий нуклеотидную последовательность SEQ ID NO: 1, выявляли в качестве горячей точки экспрессии (см. также ID локуса GenBank № AFTDO1150902.1, nt35529:39558). Сайты интеграции анализировали с целью определения их пригодности для дальнейшего получения линий клеток. Было желательно, чтобы сайты интеграции располагались в некодирующей области, чтобы не нарушать нормального геномного аппарата клетки, например трансляции белков, или не изменять фенотип клетки.

Исходя из выравнивания с использованием поиска Blat (Kent W.J., BLAT - the BLAST-like alignment tool. *Genome Res.* 2002 Apr; 12(4):656-64), SEQ ID NO: 1 характеризовалась очень низкой гомологией с геномными последовательностями мыши и человека. Blast-анализ последовательностей SEQ ID NO: 1 по отношению к CHO-1[ATCC]_refseq_transcript (www.chogenome.org) показал, что последовательность выявленного локуса не содержала никаких кодирующих областей каких-либо известных генов. Более крупную последовательность под SED ID NO: 4, которая охватывала SEQ ID NO: 1, также выявляли в качестве локуса, пригодного для направленной интеграции.

Определяли, что последовательности сайтов интеграции находились в некодирующих областях геномов CHO и мыши, и их дополнительно использовали в экспериментах, описанных ниже.

Пример 2. Эффективность экзогенной ДНК, включенной в сайты интеграции клеток-хозяев.

Направленную вставку экзогенных генов в специфический локус генома CHO в виде SEQ ID NO: 1 выполняли при помощи TALE-нуклеазы (TALEN). На конструкцию, содержащую последовательности тяжелой и легкой цепей антитела, случайным образом интегрированные в геном клетки, как в примере 1, направляли TALEN. TALEN направляли на положения в трех идентичных генах Нуг конструкции для экспрессии антитела (см. фиг. 1А). Целевой сайт расщепления TALEN для последовательностей Нуг выбирали на основе ZiFit.partners.org (ZiFit Targeter, версия 4.2). TALEN разрабатывали на основе известных способов (Boch J. et al., 2009 *Science* 326:1509-1512).

Донорный вектор mKate (см. фиг. 1В) и TALEN-кодирующий вектор трансфицировали в клетки-хозяева CHO при помощи стандартного протокола с использованием липофектина (LIPOFECTAMINE, Life Technologies, Гейтерсберг, Мэриленд). Клетки культивировали и стабильные клоны с желательными признаками выделяли и сортировали при помощи FACS. Одну интеграцию в желаемый локус подтверждали при помощи Саузерн-блоттинга и ПЦР.

Пример 3. Направленная рекомбинация разработанных клеток в представляющем интерес локусе при помощи RMCE.

Линию клеток CHO, экспрессирующую высокие уровни флуоресцентного гена, например mKate, где ген фланкирован сайтами lox в представляющем интерес локусе, отбирали с целью выделения. Вторую линию клеток CHO, экспрессирующую второй флуоресцентный ген, dsRed, где ген фланкирован сайтами lox, располагали в контрольном локусе, т.е. EESYR (патент США № 8389239B2, опубликованный 5 марта 2013 г.).

Трансфицированные клетки CHO адаптировали для роста в суспензии в бессывороточной среде. Затем клетки трансфицировали в планшете размером 10 см донорным вектором экспрессии и плазмидой, кодирующей Cre-рекомбиназу. Донорный вектор экспрессии содержал ген, представляющий интерес, кодирующий белок слияния с Fc-фрагментом, фланкированный сайтами Lox (см. фиг. 3А или 3В). Клетки культивировали в среде для культивирования с 400 мкг/мл гидромицина в течение двух недель после трансфекции, и клетки, экспрессирующие eYFP, но не mKate (или dsRed в случае локуса интеграции EESYR), выделяли при помощи проточной цитометрии. Клетки, экспрессирующие eYFP, экспандировали в суспензионных культурах в бессывороточной среде и уровни мРНК определяли при помощи qRT-PCR с использованием стандартных процедур для каждой популяции клеток, кодирующих белок слияния с Fc-фрагментом (см. фиг. 4).

Эффективность рекомбинационного обмена (процент популяции из выживших клеток, экспрессирующих из маркера донорной кассеты, т.е. eYFP, в результате обмена с красным маркером, т.е. mKate или dsRed) сравнивали между популяциями клеток (таблица). Высокую эффективность рекомбинационного обмена наблюдали в каждом локусе.

Эффективность рекомбинации

	Красный маркер	Эффективность обмена (%) (красный маркер+/eYFP-)	Случайная интеграция (%) (красный маркер+/eYFP+)
LOCUS 1 (SEQ ID NO:1)	mKate	72	27
Контрольный локус (EESYR)	dsRed	92	7

Транскрипцию наблюдали при более высоком уровне (в 1,5 раза выше) в клеточной популяции, имеющей сконструированный LOCUS 1 по сравнению с контрольным локусом (фиг. 4).

Объем настоящего изобретения не ограничен конкретными вариантами осуществления, описанными в данном документе. Действительно, различные модификации настоящего изобретения, кроме описанных в данном документе, будут очевидны специалистам в данной области техники из вышеизложенного описания и сопровождающих фигур. Такие модификации входят в объем прилагаемой формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Выделенная клетка для стабильной интеграции и/или экспрессии экзогенных нуклеиновых кислот, содержащая экзогенную последовательность нуклеиновой кислоты, интегрированную в локус генома клетки, где локус содержит нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 1 или SEQ ID NO: 4, где клетка представляет собой клетку CHO.

2. Клетка по п.1, где экзогенная последовательность нуклеиновой кислоты содержит одну или более последовательностей узнавания для рекомбиназы.

3. Клетка по п.2, где экзогенная последовательность нуклеиновой кислоты содержит по меньшей мере две последовательности узнавания для рекомбиназы и селективный маркер, расположенный между двумя последовательностями узнавания для рекомбиназы.

4. Клетка по п.2, где одна или более последовательностей узнавания для рекомбиназы выбраны из группы, состоящей из сайта LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Lox5171, сайта Loxm2, сайта Lox71, сайта Lox66, сайта LoxFas и сайта ftr.

5. Клетка по п.1, где экзогенная последовательность нуклеиновой кислоты содержит первый экзогенный ген, представляющий интерес (GOI), и первый экзогенный промотор, где первый экзогенный GOI функционально связан с первым экзогенным промотором.

6. Клетка по п.5, где экзогенная последовательность нуклеиновой кислоты дополнительно содержит второй экзогенный GOI и второй экзогенный промотор, где второй экзогенный GOI расположен в направлении 3' от первого GOI и функционально связан со вторым экзогенным промотором.

7. Клетка по п.6, где экзогенная последовательность нуклеиновой кислоты дополнительно содержит первую последовательность узнавания для рекомбиназы в направлении 5' от первого экзогенного GOI и вторую последовательность узнавания для рекомбиназы в направлении 3' от второго экзогенного GOI.

8. Клетка по п.7, где первая и вторая последовательности узнавания для рекомбиназы являются различными и выбраны из группы, состоящей из сайта LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Lox5171, сайта Loxm2, сайта Lox71, сайта Lox66, сайта LoxFas и сайта ftr.

9. Клетка по п.7, где первый экзогенный GOI кодирует легкую цепь антитела или его антигенсвязывающего фрагмента, и второй экзогенный GOI кодирует тяжелую цепь антитела или его антигенсвязывающего фрагмента.

10. Клетка по п.6, где экзогенная последовательность нуклеиновой кислоты дополнительно содержит третий экзогенный GOI и третий экзогенный промотор, где третий экзогенный GOI функционально связан с третьим экзогенным промотором.

11. Клетка по п.10, где третий экзогенный GOI и функционально связанный третий экзогенный промотор расположены в направлении 3' от второго экзогенного GOI.

12. Клетка по п.10, где первый, второй и третий GOI кодируют полипептид, выбранный из группы, состоящей из первой легкой цепи антитела или его антигенсвязывающего фрагмента, второй легкой цепи антитела или его антигенсвязывающего фрагмента и тяжелой цепи антитела или его антигенсвязывающего фрагмента.

13. Клетка по п.10, где первый, второй и третий GOI кодируют полипептид, выбранный из группы, состоящей из легкой цепи антитела или его антигенсвязывающего фрагмента, первой тяжелой цепи антитела или его антигенсвязывающего фрагмента и второй тяжелой цепи антитела или его антигенсвязывающего фрагмента.

14. Способ для интеграции экзогенной нуклеиновой кислоты в геном клетки CHO, включающий введение в клетку CHO экзогенной нуклеиновой кислоты, где экзогенная нуклеиновая кислота интегрируется в локус генома, содержащий нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 1 или SEQ ID NO: 4.

15. Способ по п.14, где экзогенная нуклеиновая кислота содержит одну или несколько последовательностей узнавания для рекомбиназы.

16. Способ по п.15, где экзогенная нуклеиновая кислота содержит по меньшей мере две последовательности узнавания для рекомбиназы и селективный маркер, расположенный между двумя последовательностями узнавания для рекомбиназы.

17. Способ по п.15, где одна или несколько последовательностей узнавания для рекомбиназы выбраны из группы, состоящей из сайта LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Lox5171, сайта Loxm2, сайта Lox71, сайта Lox66, сайта LoxFas и сайта frr.

18. Способ по п.14, где экзогенная нуклеиновая кислота содержит первый экзогенный ген, представляющий интерес (GOI), и первый экзогенный промотор, где первый экзогенный GOI функционально связан с первым экзогенным промотором.

19. Способ по п.18, где экзогенная нуклеиновая кислота дополнительно содержит второй экзогенный GOI и второй экзогенный промотор, где второй экзогенный GOI функционально связан со вторым экзогенным промотором и расположен в направлении 3' от первого экзогенного GOI.

20. Способ по п.19, где первый экзогенный GOI кодирует легкую цепь антитела или его антигенсвязывающего фрагмента, и второй экзогенный GOI кодирует тяжелую цепь антитела или его антигенсвязывающего фрагмента.

21. Способ по п.14, где экзогенную нуклеиновую кислоту вводят с помощью вектора нуклеиновой кислоты, содержащего:

- a) 5'-область гомологии, гомологичную последовательности, присутствующей в нуклеотидной последовательности SEQ ID NO: 1 или SEQ ID NO: 4,
- b) экзогенную нуклеиновую кислоту и
- c) 3'-область гомологии, гомологичную последовательности, присутствующей в нуклеотидной последовательности SEQ ID NO: 1 или SEQ ID NO: 4.

22. Способ по п.21, где экзогенную нуклеиновую кислоту вводят с помощью, по меньшей мере, дополнительного вектора нуклеиновой кислоты или молекулы мРНК.

23. Способ по п.22, где дополнительный вектор выбран из группы, состоящей из аденовируса, лентивируса, ретровируса, аденоассоциированного вируса, интегрирующего фагового вектора, невирусного вектора, транспозона и/или транспозазы, субстрата интегразы и плазмиды, и/или дополнительный вектор содержит нуклеиновую кислоту, кодирующую сайт-специфическую нуклеазу, сайт-специфическая нуклеаза необязательно выбрана из группы, состоящей из цинк-пальцевой нуклеазы (ZFN), димера ZFN, эффекторной нуклеазы, подобной активаторам транскрипции (TALEN), белка слияния на основе домена TAL-эффектора и РНК-управляемой ДНК-эндонуклеазы.

24. Способ по п.14, где перед введением экзогенной нуклеиновой кислоты клетка CHO содержит одну или более последовательностей узнавания для рекомбиназы, интегрированных в указанный локус, и где экзогенная нуклеиновая кислота интегрируется в одну или более последовательностей узнавания для рекомбиназы в указанном локусе.

25. Способ по п.24, где перед введением экзогенной нуклеиновой кислоты клетка CHO содержит по меньшей мере две последовательности узнавания для рекомбиназы, интегрированные в указанный локус, и селективный маркер, расположенный между двумя последовательностями узнавания для рекомбиназы.

26. Способ по п.25, где по меньшей мере две последовательности узнавания для рекомбиназы являются различными и выбраны из группы, состоящей из сайта LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Lox5171, сайта Loxm 2, сайта Lox71, сайта Lox66, сайта LoxFas и сайта frr.

27. Способ по п.25, где экзогенная нуклеиновая кислота содержит первый экзогенный ген, представляющий интерес (GOI), функционально связанный с первым экзогенным промотором, первую последовательность узнавания для рекомбиназы, расположенную в направлении 5' от первого экзогенного промотора, и вторую последовательность узнавания для рекомбиназы, расположенную в направлении 3' от первого экзогенного GOI, и где первая и вторая последовательности узнавания для рекомбиназы являются идентичными соответственно по меньшей мере двум последовательностям узнавания для рекомбиназы, присутствующим в локусе генома перед введением экзогенной нуклеиновой кислоты.

28. Способ по п.27, где первая и вторая последовательности узнавания для рекомбиназы являются различными и выбраны из группы, состоящей из сайта LoxP, сайта Lox511, сайта Lox2272, сайта Lox2372, сайта Lox5171, сайта Loxm2, сайта Lox71, сайта Lox66, сайта LoxFas и сайта frr.

29. Способ по п.27, где экзогенная нуклеиновая кислота дополнительно содержит второй экзогенный GOI, функционально связанный со вторым экзогенным промотором, где второй экзогенный GOI расположен в направлении 3' от первого экзогенного GOI и в направлении 5' от второй последовательности узнавания для рекомбиназы.

30. Способ по п.29, где первый экзогенный GOI кодирует легкую цепь антитела или его антигенсвязывающего фрагмента, и второй экзогенный GOI кодирует тяжелую цепь антитела или его антигенсвязывающего фрагмента.

31. Способ для экспрессии гена, представляющего интерес (GOI), включающий:

- (a) получение клетки CHO, содержащей экзогенную последовательность нуклеиновой кислоты, ин-

тегрированную в локус генома клетки CHO, где экзогенная последовательность нуклеиновой кислоты содержит первый экзогенный GOI, функционально связанный с первым экзогенным промотором, и где локус содержит нуклеотидную последовательность, которая по меньшей мере на 90% идентична SEQ ID NO: 1 или SEQ ID NO: 4, и

(b) культивирование клетки CHO из (a) в условиях, которые обеспечивают экспрессию первого экзогенного GOI.

32. Способ по п.31, где первый экзогенный GOI кодирует первый представляющий интерес белок (POI), и где способ дополнительно предусматривает:

а) извлечение первого POI.

33. Способ по п.31, где экзогенная последовательность нуклеиновой кислоты дополнительно содержит второй экзогенный GOI, функционально связанный со вторым экзогенным промотором, и где условия обеспечивают экспрессию первого экзогенного GOI и второго экзогенного GOI.

34. Способ по п.31, где первый экзогенный GOI кодирует легкую цепь антитела, и второй экзогенный GOI кодирует тяжелую цепь антитела.

35. Композиция для модификации генома клетки CHO, содержащая вектор, который содержит:

а) 5'-область гомологии, гомологичную последовательности, присутствующей в нуклеотидной последовательности SEQ ID NO: 1 или SEQ ID NO: 4,

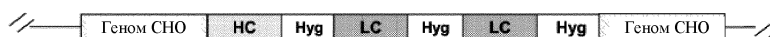
б) экзогенную последовательность нуклеиновой кислоты и

с) 3'-область гомологии, гомологичную последовательности, присутствующей в нуклеотидной последовательности SEQ ID NO: 1 или SEQ ID NO: 4.

36. Композиция по п.35, где последовательность экзогенной нуклеиновой кислоты содержит одну или более последовательностей узнавания для рекомбиназы и/или содержит первый экзогенный ген, представляющий интерес (GOI).

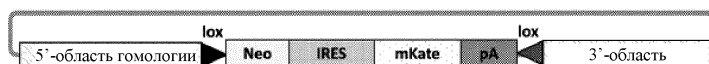
37. Композиция по п.35 или 36, где композиция содержит, по меньшей мере, дополнительный вектор или молекулу мРНК, при этом необязательно дополнительный вектор содержит нуклеиновую кислоту, кодирующую сайт-специфическую нуклеазу для интеграции последовательности узнавания, и где необязательно сайт-специфическая нуклеаза выбрана из группы, состоящей из цинк-пальцевой нуклеазы (ZFN), димера ZFN, эффекторной нуклеазы, подобной активаторам транскрипции (TALEN), белка слияния на основе домена TAL-эффектора или PНК-управляемой ДНК-эндонуклеазы.

Интеграция в локусе SEQ ID NO: 1



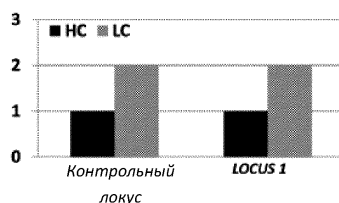
Фиг. 1А

Донорный вектор с флуоресцентным маркером



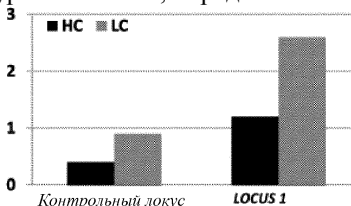
Фиг. 1В

Число копий гена, определяемое при помощи qPCR



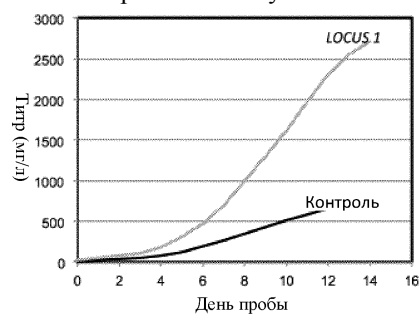
Фиг. 2А

Относительный уровень мРНК, определяемый при помощи qPCR

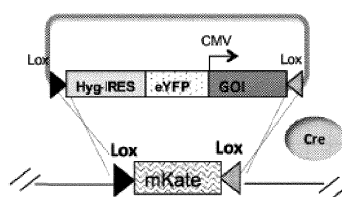


Фиг. 2В

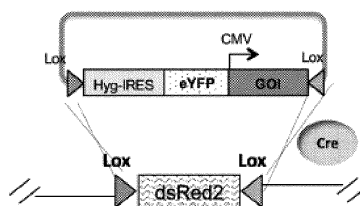
Титры биореактора в отношении продуцирующих
линий с контрольным локусом и LOCUS 1



Фиг. 2С

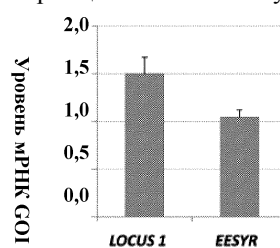
**LOCUS 1**

Фиг. 3А

**Контрольный локус**

Фиг. 3В

Сконструированный LOCUS 1 поддерживает RMCE
и транскрипцию на высоком уровне



Фиг. 4



Евразийская патентная организация, ЕАПВ

Россия, 109012, Москва, Малый Черкасский пер., 2