

(19)



**Евразийское
патентное
ведомство**

(11) **044725**

(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2023.09.27

(21) Номер заявки
201892137

(22) Дата подачи заявки
2017.04.20

(51) Int. Cl. **C07K 16/00** (2006.01)
C07K 16/46 (2006.01)
C12N 15/90 (2006.01)

(54) КОМПОЗИЦИИ И СПОСОБЫ ПОЛУЧЕНИЯ АНТИТЕЛ, ОСНОВАННЫЕ НА ПРИМЕНЕНИИ ЛОКУСОВ, ОБЕСПЕЧИВАЮЩИХ ПОВЫШЕННУЮ ЭКСПРЕССИЮ

(31) 62/325,385

(32) 2016.04.20

(33) US

(43) 2019.03.29

(86) PCT/US2017/028552

(87) WO 2017/184831 2017.10.26

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

(72) Изобретатель:
**Бабб Роберт, Бураков Дарья, Чен Ганг,
Фандл Джеймс П. (US)**

(74) Представитель:
**Угрюмов В.М., Лыу Т.Н., Глухарёва
А.О., Гизатуллина Е.М., Строкова
О.В. (RU)**

(56) WO-A1-2008151219

WO-A1-2014121087

WO-A1-2013190032

CHRISTINE LATTENMAYER ET AL.
"Identification of transgene integration loci of different highly expressing recombinant CHO cell lines by FISH", CYTOTECNOLOGY, KLUWER ACADEMIC PUBLISHERS, DO, vol. 51, no. 3, 15 November 2006 (2006-11-15), pages 171-182, XP019448503, ISSN: 1573-0778, DOI:10.1007/510616-006-9029-0

LIN ZHANG ET AL. "Recombinase-mediated cassette exchange (RMCE) for monoclonal antibody expression in the commercially relevant CHO K1SV cell line", BIOTECHNOLOGY PROGRESS., vol. 31, no. 6, 13 October 2015 (2015-10-13), pages 1645-1656, XP055383248, US ISSN: 8756-7938, DOI: 10.1002/btpr.2175

SOEREN TURAN ET AL. "Recombinase-mediated cassette exchange (RMCE) - A rapidly-expanding toolbox for targeted genomic modifications", GENE., vol. 515, no. 1, 1 February 2013 (2013-02-01), pages 1-27, XP055383251, NL ISSN: 0378-1119, DOI:10.1016/j.gene.2012.11.016
WO-A1-2016064999

(57) Изобретение относится к сайт-специфической интеграции и экспрессии рекомбинантных белков в эукариотических клетках. В частности, настоящее изобретение предусматривает композиции и способы для обеспечения улучшенной экспрессии антител, в том числе биспецифических антител, в эукариотических клетках, в частности в линиях клеток китайского хомячка (*Cricetulus griseus*), посредством применения локуса, обеспечивающего повышенную экспрессию.

B1

044725

044725 B1

Перекрестная ссылка на родственную заявку

Настоящая заявка испрашивает приоритет по предварительной заявке на патент США № 62/325385, поданной 20 апреля 2016 г., полное содержание которой включено в данный документ посредством ссылки.

Область изобретения

Изобретение относится к сайт-специфической интеграции и экспрессии рекомбинантных белков в эукариотических клетках. В частности, настоящее изобретение относится к композициям и способам для обеспечения улучшенной экспрессии антигенсвязывающих белков, таких как биспецифические антитела, в эукариотических клетках, в частности в линиях клеток китайского хомячка (*Cricetulus griseus*), посредством применения локуса, обеспечивающего повышенную экспрессию.

Уровень техники

Клеточные системы экспрессии направлены на обеспечение надежного и эффективного источника для получения определенного белка для исследовательского или терапевтического применения. Экспрессия рекомбинантного белка в клетках млекопитающих представляет собой предпочтительный способ получения белков для терапевтического применения благодаря, например, способности систем экспрессии млекопитающих соответствующим образом посттрансляционно модифицировать рекомбинантные белки.

Несмотря на доступность различных систем экспрессии, проблема эффективного переноса гена и стабильности интегрированного гена для экспрессии рекомбинантного белка существует до сих пор. Один фактор для длительной экспрессии целевого трансгена представляет собой минимальное разрушение генов клетки для того, чтобы избежать изменений фенотипа линии клеток.

Разработка стабильных линий клеток с размещением нескольких генов для экспрессии, например, нескольких цепей антитела, как в случае полиспецифических 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). В системах экспрессии для получения полиспецифических антигенсвязывающих белков часто требуется экспрессия двух или больше разных цепей иммуноглобулина, предназначенных для образования пары в виде конкретного мультимерного формата, и при этом часто могут положительно влиять на получение гомодимера, а не требуемого гетеродимера или мультимерной комбинации. Соответственно, в данной области техники существует потребность в улучшенных системах экспрессии млекопитающих.

Краткое описание изобретения

В одном аспекте предусмотрена клетка, которая содержит экзогенную последовательность нуклеиновой кислоты, интегрированную в конкретный сайт в локусе, обеспечивающем повышенную экспрессию, где экзогенная последовательность нуклеиновой кислоты кодирует биспецифический антигенсвязывающий белок.

В некоторых вариантах осуществления экзогенная последовательность нуклеиновой кислоты включает первую экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый фрагмент легкой цепи (LCF), вторую экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый фрагмент тяжелой цепи (HCF), и третью экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую второй HCF (или указанный как HCF*, где второй HCF отличается от первого HCF), где первый и второй HCF и первый LCF образуют биспецифический антигенсвязывающий белок. В определенном варианте осуществления первый и второй HCF и первый LCF содержат по меньшей мере два переменных участка и два константных домена СН3 биспецифического антигенсвязывающего белка. В некоторых вариантах осуществления два переменных участка отличаются. В некоторых вариантах осуществления два участка СН3 отличаются. В некоторых вариантах осуществления каждая из экзогенных последовательностей нуклеиновой кислоты одновременно интегрируется в конкретный сайт в локусе, обеспечивающем повышенную экспрессию.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует аминокислоты из первого константного участка (например, кодирует один или более доменов СН1, СН2, шарнира или СН3), а нуклеотидная последовательность, кодирующая второй HCF, кодирует аминокислоты из второго константного участка. Аминокислоты из первого константного участка могут быть одинаковыми или отличающимися по сравнению с аминокислотами из второго константного участка. В конкретных вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует первый домен СН3, а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй домен СН3, где первый и второй домены СН3 могут быть одинаковыми или отличающимися. В некоторых вариантах осуществления первый и второй домены СН3 отличаются по меньшей мере одним аминокислотным положением; например, один из двух доменов СН3 представляет собой домен СН3 IgG человека, а другой представляет собой модифицированный домен СН3 IgG человека, и при этом два домена СН3 обладают различной характеристикой связывания с белком А. В других вариантах осуществления нуклеотидные последовательности, кодирующие первый и второй домены СН3,

отличаются друг от друга тем, что одна из нуклеотидных последовательностей была модифицирована по кодонам.

В других конкретных вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует первый переменный участок тяжелой цепи (VH), а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй участок VH, где первая и вторая тяжелые цепи могут иметь одинаковые или отличающиеся участки VH. В другом варианте осуществления первый и второй VH могут быть связаны с одними и теми же или различными константными участками.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый LCF, кодирует первый переменный участок легкой цепи (VL).

В некоторых вариантах осуществления экзогенная последовательность нуклеиновой кислоты содержит дополнительную экзогенную нуклеиновую кислоту, которая включает нуклеотидную последовательность, кодирующую второй LCF, такой как второй переменный участок легкой цепи (VL). В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая второй участок VL, также кодирует второй константный участок легкой цепи.

Относительные положения нескольких экзогенных нуклеиновых кислот в локусе могут меняться. В некоторых вариантах осуществления нуклеиновая кислота, кодирующая LCF, расположена выше или ниже относительно обеих нуклеиновых кислот, кодирующих HCF.

В некоторых вариантах осуществления каждая из последовательностей, кодирующих HCF или LCF, независимо связана с последовательностью, регулирующей транскрипцию. В конкретных вариантах осуществления первая экзогенная нуклеиновая кислота дополнительно содержит первый промотор, функционально связанный с нуклеотидной последовательностью, кодирующей первый LCF, вторая экзогенная нуклеиновая кислота дополнительно содержит второй промотор, функционально связанный с нуклеотидной последовательностью, кодирующей первый HCF, и третья экзогенная нуклеиновая кислота содержит третий промотор, функционально связанный с нуклеотидной последовательностью, кодирующей второй HCF, где первый, второй и третий промоторы являются одинаковыми или отличающимися, и/или указанные промоторы являются одинаковыми или отличающимися по сравнению с четвертым промотором, с которым функционально связана четвертая экзогенная нуклеиновая кислота. В некоторых вариантах осуществления первый, второй и третий промоторы являются одинаковыми.

В некоторых вариантах осуществления экзогенная последовательность нуклеиновой кислоты в сайте интеграции дополнительно содержит сайты распознавания для рекомбиназы, например, первый сайт распознавания для рекомбиназы (RRS), расположенный в направлении 5' относительно первой экзогенной нуклеиновой кислоты, и второй сайт распознавания для рекомбиназы (RRS), расположенный в направлении 3' относительно как второй, так и третьей экзогенных нуклеиновых кислот, где первый и второй RRS отличаются. В некоторых вариантах осуществления также включен третий RRS, и он расположен в направлении 3' относительно первой экзогенной нуклеиновой кислоты и в направлении 5' относительно одной из второй и третьей экзогенных нуклеиновых кислот или их обеих, где третий RRS отличается от первого и второго RRS.

В некоторых вариантах осуществления экзогенная последовательность нуклеиновой кислоты может включать четвертую экзогенную нуклеиновую кислоту, содержащую ген селективируемого маркера. В конкретных вариантах осуществления четвертая экзогенная нуклеиновая кислота расположена в направлении 3' относительно первой экзогенной нуклеиновой кислоты. В определенных вариантах осуществления четвертая экзогенная нуклеиновая кислота интегрирована в виде расщепленного гена. В других вариантах осуществления четвертая экзогенная нуклеиновая кислота, или селективируемый маркер, расположена в направлении 3' от третьего RRS, который находится в направлении 3' от четвертого промотора, который функционально связан с четвертой экзогенной нуклеиновой кислотой. В некоторых вариантах осуществления ген селективируемого маркера содержит третий RRS, который был вставлен, необязательно вставленный в интрон гена селективируемого маркера, где третий RRS отличается от первого и второго RRS.

В определенных вариантах осуществления порядок расположения экзогенных нуклеиновых кислот в локусе может быть следующим: в направлении от 5' к 3' первая экзогенная нуклеиновая кислота (кодирующая LCF), четвертая экзогенная нуклеиновая кислота (кодирующая селективируемый маркер), вторая экзогенная нуклеиновая кислота (кодирующая первый HCF) и третья экзогенная нуклеиновая кислота (кодирующая второй HCF); и в некоторых конкретных вариантах осуществления вторая экзогенная нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую модифицированный домен CH3 IgG человека, и третья экзогенная нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую нативный домен CH3 IgG человека.

В определенных вариантах осуществления порядок расположения экзогенных нуклеиновых кислот в локусе может быть следующим: в направлении от 5' к 3' первая экзогенная нуклеиновая кислота (кодирующая LCF), вторая экзогенная нуклеиновая кислота (кодирующая первый HCF), четвертая экзогенная нуклеиновая кислота (кодирующая селективируемый маркер) и третья экзогенная нуклеиновая кислота (кодирующая второй HCF), где вторая экзогенная нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую нативный домен CH3 IgG человека, и третья экзогенная нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую модифицированный домен CH3 IgG

человека.

В некоторых вариантах осуществления промоторы, связанные с последовательностями, кодирующими HCF или LCF, являются одинаковыми и отличаются от промотора, с которым функционально связан ген селектируемого маркера.

В некоторых вариантах осуществления биспецифический антигенсвязывающий белок специфически связывается с антигеном Т-клетки и антигеном опухолевой клетки. Предусмотрены также другие подходящие двойные специфичности в отношении антигенов.

В некоторых вариантах осуществления локус, обеспечивающий повышенную экспрессию, выбран из локуса, содержащего нуклеотидную последовательность, которая на по меньшей мере 90% идентична SEQ ID NO: 1, или локуса, содержащего нуклеотидную последовательность, которая на по меньшей мере 90% идентична SEQ ID NO: 2.

В различных вариантах осуществления клетка представляет собой клетку CHO.

В другом аспекте предусмотрены векторы, которые сконструированы для сайт-специфической интеграции нескольких экзогенных нуклеиновых кислот.

В некоторых вариантах осуществления настоящее изобретение предусматривает набор векторов, при этом набор включает первый вектор, содержащий в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF, и третий RRS; и второй вектор, содержащий в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый участок VH, второй RRS; где либо первая, либо вторая нуклеиновая кислота дополнительно содержит нуклеотидную последовательность, кодирующую второй HCF; и где первый и второй HCF и первый LCF образуют биспецифический антигенсвязывающий белок.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая второй HCF, включена в первую нуклеиновую кислоту, необязательно расположенную ниже нуклеотидной последовательности, кодирующей первый LCF. В других вариантах осуществления нуклеотидная последовательность, кодирующая второй HCF, включена во вторую нуклеиновую кислоту.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует первый химерный константный участок (например, кодирует один или более доменов CH1, шарнира, CH2 или CH3 или их фрагментов из любого изотипа), а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй химерный константный участок. Примеры химерного константного участка описаны в публикации международной заявки согласно PCT № WO 2014/121087 A1, опубликованной 7 августа 2014 г., включенной в данный документ посредством ссылки. Аминокислоты из первого константного участка могут быть одинаковыми или отличающимися по сравнению со вторым химерным константным участком. В конкретных вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует первый домен CH3, а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй домен CH3, где первый и второй домены CH3 могут быть одинаковыми или отличающимися. В некоторых вариантах осуществления первый и второй домены CH3 отличаются по меньшей мере одним аминокислотным положением; например, один из двух доменов CH3 представляет собой домен CH3 IgG человека, а другой представляет собой модифицированный домен CH3 IgG человека, и при этом два домена CH3 обладают различной характеристикой связывания с белком А. В других вариантах осуществления нуклеотидные последовательности, кодирующие первый и второй домены CH3, отличаются друг от друга тем, что одна из нуклеотидных последовательностей была модифицирована по кодонам.

В других конкретных вариантах осуществления нуклеотидная последовательность, кодирующая первый участок VH, кодирует первую тяжелую цепь, и нуклеотидная последовательность, кодирующая второй участок VH, кодирует вторую тяжелую цепь, где первая и вторая тяжелые цепи могут иметь одинаковые или отличающиеся константные участки.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый LCF, кодирует первый переменный участок легкой цепи (VL).

В некоторых вариантах осуществления каждая из последовательностей, кодирующих LCF или HCF, независимо связана с последовательностью, регулирующей транскрипцию, такой как промотор. В конкретных вариантах осуществления промотор, связанный с последовательностью, кодирующей первый HCF, и промотор, связанный с последовательностью, кодирующей второй HCF, являются одинаковыми. В конкретных вариантах осуществления все промоторы, связанные с LCF и HCF, являются одинаковыми.

В некоторых вариантах осуществления первая нуклеиновая кислота в первом векторе дополнительно содержит 5'-часть гена селектируемого маркера, расположенного в направлении 5' от третьего RRS в первом векторе; и вторая нуклеиновая кислота дополнительно содержит оставшуюся 3'-часть гена селектируемого маркера, расположенного в направлении 3' от третьего RRS во втором векторе - т. е. ген селектируемого маркера разделен по двум векторам. В других вариантах осуществления селектируемый маркер и промотор, с которым он функционально связан, разделены между двумя векторами, другими словами, промотор и ген селектируемого маркера расположены в разных векторах. В определенных ва-

риантах осуществления промотор, функционально связанный с геном маркера, расположен в первом векторе в направлении 5' относительно третьего RRS, и ген маркера расположен в направлении 3' от третьего RRS во втором векторе и расположен в направлении 5' от второго промотора, функционально связанного со второй нуклеиновой кислотой, и третьего промотора, функционально связанного с третьей нуклеиновой кислотой. В некоторых вариантах осуществления третий RRS в первом векторе находится в пределах 5'-части интрона гена селективируемого маркера; а третий RRS во втором векторе находится в пределах 3'-части интрона гена селективируемого маркера.

В конкретных вариантах осуществления первый вектор содержит в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту и третий RRS; и второй вектор содержит в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, где вторая нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую первый HCF, и нуклеотидную последовательность, кодирующую второй HCF, и второй RRS. В других конкретных вариантах осуществления первый вектор содержит в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, где первая нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую первый HCF, и нуклеотидную последовательность, кодирующую второй HCF, и третий RRS; и второй вектор содержит в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, где вторая нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую первый участок HCF, и второй RRS. В любом из данных конкретных вариантов осуществления первая нуклеиновая кислота может дополнительно содержать 5'-часть гена селективируемого маркера, расположенного в направлении 5' от третьего RRS в первом векторе; и вторая нуклеиновая кислота дополнительно содержит оставшуюся 3'-часть гена селективируемого маркера, расположенного в направлении 3' от третьего RRS во втором векторе; и где необязательно третий RRS в первом векторе находится в пределах 5'-части интрона гена селективируемого маркера, а третий RRS во втором векторе находится в пределах 3'-части интрона гена селективируемого маркера.

В некоторых вариантах осуществления набор векторов может включать дополнительный вектор или векторы; например, вектор, содержащий один или более RRS и нуклеотидную последовательность, кодирующую второй LCF, или вектор, кодирующий одну или более рекомбиназ, которые распознают RRS.

В других вариантах осуществления настоящее изобретение предусматривает вектор, сконструированный для обеспечения сайт-специфической интеграции нескольких экзогенных нуклеиновых кислот посредством гомологичной рекомбинации на основе гомологичных плеч. В некоторых вариантах осуществления вектор содержит экзогенную последовательность нуклеиновой кислоты, которая кодирует биспецифический антигенсвязывающий белок, фланкированную 5'-концевым гомологичным плечом и 3'-концевым гомологичным плечом для обеспечения интеграции в локус клетки, обеспечивающий повышенную экспрессию.

В дополнительном аспекте настоящее изобретение предусматривает систему, которая включает комбинацию клетки и одного или более векторов, и которую можно применять для получения клеток, характеризующихся интегрированными в локус, обеспечивающий повышенную экспрессию, экзогенными нуклеиновыми кислотами, которые вместе кодируют биспецифический антигенсвязывающий белок.

В определенных вариантах осуществления предусмотрена система, которая включает клетку и набор векторов, где клетка содержит интегрированный в локус ее генома, обеспечивающий повышенную экспрессию, набор RRS, которые отличаются друг от друга и отделены одной или более экзогенными нуклеиновыми кислотами, такими как селективируемые маркеры, для обмена посредством рекомбинации с генами, представляющими интерес, в наборе векторов; и где RRS в наборе векторов предусмотрены в том же порядке, что и RRS в клетке.

В некоторых вариантах осуществления предусмотрена система, которая включает клетку и набор векторов, где клетка содержит интегрированные в локус ее генома, обеспечивающий повышенную экспрессию, в направлении от 5' к 3': первый RRS, первую экзогенную нуклеиновую кислоту, второй RRS, вторую экзогенную нуклеиновую кислоту и третий RRS, где три RRS отличаются друг от друга; где набор векторов включает первый вектор, содержащий в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первую цепь иммуноглобулина или ее фрагмент, и второй RRS; второй вектор, содержащий второй RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую вторую цепь иммуноглобулина или ее фрагмент, и третий RRS; и где либо первая нуклеиновая кислота, либо вторая нуклеиновая кислота дополнительно содержит нуклеотидную последовательность, кодирующую третью цепь иммуноглобулина или ее фрагмент. При введении векторов в клетку первая и вторая нуклеиновые кислоты из векторов интегрируются в локус, обеспечивающий повышенную экспрессию, с помощью рекомбинации, опосредованной первым, вторым и третьим RRS.

В некоторых вариантах осуществления первая экзогенная нуклеиновая кислота в клетке содержит первый ген селективируемого маркера, и вторая экзогенная нуклеиновая кислота в клетке содержит второй ген селективируемого маркера, где первый и второй гены селективируемых маркеров отличаются. Селективируемые маркеры обмениваются интегрированными экзогенными нуклеиновыми кислотами в клетке.

В некоторых вариантах осуществления первый вектор содержит в направлении от 5' к 3' первый

RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF, и третий RRS; и второй вектор содержит в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, где вторая нуклеиновая кислота содержит как нуклеотидную последовательность, кодирующую первый HCF, так и нуклеотидную последовательность, кодирующую второй HCF, и второй RRS. В других вариантах осуществления первый вектор содержит в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF, и нуклеотидную последовательность, кодирующую второй HCF, и третий RRS; и второй вектор содержит в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF, и второй RRS.

В некоторых вариантах осуществления первый вектор содержит в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF, и третий RRS; и второй вектор содержит в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, где вторая нуклеиновая кислота содержит как нуклеотидную последовательность, кодирующую первый LCF, так и нуклеотидную последовательность, кодирующую второй HCF, и второй RRS. В других вариантах осуществления первый вектор содержит в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF, и нуклеотидную последовательность, кодирующую второй HCF, и третий RRS; и второй вектор содержит в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF, и второй RRS. В любом из данных вариантов осуществления первая нуклеиновая кислота в первом векторе может дополнительно содержать промотор, расположенный в направлении 5' от третьего RRS, и вторая нуклеиновая кислота во втором векторе дополнительно содержит ген селективируемого маркера, с которым будет функционально связан промотор, расположенный в направлении 3' от третьего RRS. В других вариантах осуществления первая нуклеиновая кислота в первом векторе может дополнительно содержать 5'-часть гена селективируемого маркера, расположенного в направлении 5' от третьего RRS, и вторая нуклеиновая кислота во втором векторе дополнительно содержит оставшуюся 3'-часть гена селективируемого маркера, расположенного в направлении 3' от третьего RRS; где необязательно третий RRS в первом векторе находится в пределах 5'-части интрона гена селективируемого маркера, а третий RRS во втором векторе находится в пределах 3'-части интрона гена селективируемого маркера.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая LCF, функционально связана с первым промотором, нуклеотидная последовательность, кодирующая первый HCF, функционально связана со вторым промотором, и нуклеотидная последовательность, кодирующая второй HCF, функционально связана с третьим промотором, где первый, второй и третий промоторы являются одинаковыми и отличаются от промотора, с которым ген селективируемого маркера, если он присутствует в одном из векторов, функционально связан.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует первый домен СНЗ, а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй домен СНЗ, где первый и второй домены СНЗ могут быть одинаковыми или отличающимися. В некоторых вариантах осуществления один из двух доменов СНЗ представляет собой нативный домен СНЗ IgG человека, а другой домен СНЗ представляет собой модифицированный домен СНЗ IgG человека. В конкретных вариантах осуществления нуклеотидная последовательность, кодирующая модифицированный домен СНЗ, находится в первом векторе (т. е. векторе, кодирующем первый LCF), необязательно ниже нуклеотидной последовательности, кодирующей первый LCF. В других конкретных вариантах осуществления нуклеотидная последовательность, кодирующая модифицированный домен СНЗ, находится во втором векторе и находится выше нуклеотидной последовательности, кодирующей немодифицированный домен СНЗ.

В другом аспекте настоящее изобретение предусматривает способы получения биспецифического антигенсвязывающего белка.

В некоторых вариантах осуществления способ включает обеспечение системы, описанной в данном документе, которая содержит клетку, имеющую RRS и набор векторов, содержащих несколько экзогенных нуклеиновых кислот, которые вместе кодируют биспецифический антигенсвязывающий белок и RRS, соответствующий RRS в клетке; введение векторов в клетку с помощью трансфекции; отбор трансфицированной клетки, где экзогенные нуклеиновые кислоты из векторов интегрировались в локус клетки, обеспечивающий повышенную экспрессию, с помощью рекомбинации, опосредованной RRS; обеспечение экспрессии полипептидов, кодируемых нуклеиновыми кислотами, в трансформированной клетке и получение биспецифического антигенсвязывающего белка из трансфицированной клетки.

В некоторых вариантах осуществления способ может включать обеспечение клетки, содержащей экзогенную последовательность нуклеиновой кислоты, которая кодирует биспецифический антигенсвязывающий белок, интегрированную в локус, обеспечивающий повышенную экспрессию, обеспечение экспрессии биспецифического антигенсвязывающего белка экзогенной последовательностью нуклеиновой кислоты и получение биспецифического антигенсвязывающего белка из клетки.

Краткое описание графических материалов

Фиг. 1. Иллюстративная стратегия биспецифического клонирования для интеграции в локус, обеспечивающий повышенную экспрессию. Вектор легкой цепи ("LC"), например, общей легкой цепи, и вектор двух тяжелых цепей ("HC") (при этом "*" указывает, что две HC отличаются, например, HC* содержит модификацию в домене СНЗ и/или является модифицированной по кодонам) получают с помощью клонирования переменных участков антитела, представляющего интерес, в соответствующие векторы. 3'-Сайт RRS в векторе LC и 5'-сайт RRS в векторе двух HC являются одинаковыми и включены в расщепленный интрон гена устойчивости к гигромицину, сконструированный с возможностью объединения и вырезания интрона для обеспечения экспрессии белка, кодируемого геном устойчивости к гигромицину, для эффективного отбора рекомбинантов. Стрелочками обозначены промоторы.

Фиг. 2. Иллюстративная стратегия биспецифического клонирования для интеграции в локус, обеспечивающий повышенную экспрессию. Применение универсальной легкой цепи (см., например, гуманизованную универсальную легкую цепь (ULC) мыши VelocImmune®, как описано в WO 2013022782), имеющей 5'-RRS (RRS1), позволяет эффективно конструировать новые биспецифические антитела путем вставки одной тяжелой цепи (HC*), фланкированной третьим RRS (RRS3), в ранее полученные плазмиды, содержащие кассеты экспрессии для универсальной легкой цепи. Вторую тяжелую цепь (HC) клонируют во вторую плазмиду с помощью сайтов RRS2 и RRS3.

Фиг. 3. Иллюстративная стратегия биспецифического клонирования для интеграции в локус, обеспечивающий повышенную экспрессию. Три различные цепи антител (AbC1, AbC2 и AbC3) биспецифического антитела сперва клонируют в отдельные векторы. Каждый из векторов с AbC1 и AbC3 имеет сайты RRS, фланкирующие кассету экспрессии антитела. Кассету экспрессии AbC2 вырезают из плазмиды AbC2, а затем субклонировать в плазмиду для экспрессии AbC3 с получением плазмиды, которая содержит в направлении от 5' к 3': сайт RRS3, кассету экспрессии AbC2, кассету экспрессии AbC3 и сайт RRS2. Данную плазмиду вместе с плазмидой AbC1 и рекомбиназой вводят в клетки-хозяева, которые содержат RRS1 и RRS2 в локусе, обеспечивающем повышенную экспрессию. После опосредованного рекомбиназой обмена кассетами выделяют линии клеток, экспрессирующие биспецифическое антитело.

Фиг. 4. Иллюстративная стратегия биспецифического клонирования для интеграции в локус, обеспечивающий повышенную экспрессию. Три различные цепи антител (AbC1, AbC2 и AbC3) биспецифического антитела сперва клонируют в отдельные векторы. Каждый из векторов с AbC1 и AbC3 имеет сайты RRS, фланкирующие кассету экспрессии антитела. Кассету экспрессии AbC2 вырезают из плазмиды AbC2, а затем субклонировать в плазмиду для экспрессии AbC1 с получением плазмиды, которая содержит в направлении от 5' к 3': сайт RRS1, кассету экспрессии AbC1, кассету экспрессии AbC2 и сайт RRS3. Данную плазмиду вместе с плазмидой AbC3 и рекомбиназой вводят в клетки-хозяева, которые содержат RRS1 и RRS2 в локусе, обеспечивающем повышенную экспрессию. После опосредованного рекомбиназой обмена кассетами выделяют линии клеток, экспрессирующие биспецифическое антитело.

Фиг. 5. Экспрессия биспецифического антитела из кассет экспрессии, интегрированных в один сайт генома (EESYR®). Линии клеток CHO, RSX4189-1, RSX4187-1, RSX4191-1, RSX4188-1, получали с помощью опосредованного рекомбиназой обмена кассетами в локусе EESYR®. Порядок расположения кассет экспрессии для трех разных цепей антитела (AbC1, AbC2 и AbC3) биспецифического Ab в локусе EESYR® изображен слева. Титры каждого биспецифического антитела в отработанной среде после 4-дневного культивирования во встряхиваемых колбах определяли с помощью HPLC, и они представлены на гистограмме справа.

Подробное описание изобретения

Определения

Используемый в данном документе термин "антитело" предусматривает молекулы иммуноглобулинов, содержащие четыре полипептидные цепи, две тяжелые (H) цепи и две легкие (L) цепи, соединенные между собой дисульфидными связями. Каждая тяжелая цепь содержит переменный участок тяжелой цепи (сокращенно обозначаемый в данном документе как HCVR или VH) и константный участок тяжелой цепи. Константный участок тяжелой цепи содержит три домена - CH1, CH2 и CH3. Каждая легкая цепь содержит переменный участок легкой цепи (сокращенно обозначаемый в данном документе как LCVR или VL) и константный участок легкой цепи. Константный участок легкой цепи содержит один домен CL. VH- и VL-участки могут быть дополнительно подразделены на участки гипервариабельности, называемые определяющими комплементарность участками (CDR), чередующиеся с более консервативными участками, называемыми каркасными участками (FR). Каждый VH и VL состоит из трех CDR и четырех FR, расположенных от аминоконца к карбоксиконцу в следующем порядке: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4 (CDR тяжелой цепи могут сокращенно называться как HCDR1, HCDR2 и HCDR3; CDR легкой цепи могут сокращенно называться как LCDR1, LCDR2 и LCDR3).

Фраза "антигенсвязывающий белок" предусматривает белок, который имеет по меньшей мере один CDR и способен к селективному распознаванию антигена, т. е. способен связывать антиген с K_D , которая находится по меньшей мере в микромолярном диапазоне. Для терапевтических антигенсвязывающих белков (например, терапевтических антител) часто требуется K_D , которая находится в наномолярном или

пикомолярном диапазоне. Обычно антигенсвязывающий белок содержит два или больше CDR, например, 2, 3, 4, 5 или 6 CDR. Примеры антигенсвязывающих белков включают антитела, антигенсвязывающие фрагменты антител, такие как полипептиды, содержащие переменные участки тяжелых цепей и легких цепей антитела (например, Fab-фрагмент, F(ab')₂-фрагмент), и белки, содержащие переменные участки тяжелых цепей и легких цепей антитела и содержащие дополнительные аминокислоты из константных участков тяжелых и/или легких цепей (как, например, одного или более константных доменов, т. е. одного или более доменов CL, CH1, CH2 и CH3).

Фраза "биспецифический антигенсвязывающий белок" предусматривает антигенсвязывающие белки, способные к селективному связыванию с двумя или более эпитопами на двух разных молекулах (например, антигенах) либо на одной и той же молекуле (например, на одном и том же антигене) или обладающие отличающимися специфичностями в отношении них. Антигенсвязывающая часть или фрагмент антигенсвязывающей части (Fab) такого белка обуславливают специфичность в отношении конкретного антигена и обычно состоят из переменного участка тяжелой цепи и переменного участка легкой цепи иммуноглобулина. В некоторых случаях переменный участок тяжелой цепи и переменный участок легкой цепи могут не представлять собой родственную пару, другими словами, характеризоваться разными специфичностями связывания.

Примером биспецифического антигенсвязывающего белка является "биспецифическое антитело", которое предусматривает антитело, способное селективно связывать два или более эпитопов. Как правило, биспецифические антитела содержат две различные тяжелые цепи, при этом каждая тяжелая цепь специфически связывается с отличающимся эпитопом на двух разных молекулах (например, антигенах) либо на одной и той же молекуле (например, на одном и том же антигене). Если биспецифический антигенсвязывающий белок способен селективно связывать два разных эпитопа (первый эпитоп и второй эпитоп), то аффинность переменного участка первой тяжелой цепи в отношении первого эпитопа, как правило, будет на по меньшей мере один, два, или три, или четыре порядка ниже аффинности переменного участка первой тяжелой цепи в отношении второго эпитопа и наоборот. Биспецифические антигенсвязывающие белки, такие как биспецифические антитела, могут содержать переменные участки тяжелых цепей, которые распознают разные эпитопы одного и того же антигена. Типичное биспецифическое антитело имеет две тяжелые цепи, каждая из которых имеет три CDR тяжелой цепи, за которыми следуют (от N-конца к C-концу) домен CH1, шарнир, домен CH2 и домен CH3, и легкую цепь иммуноглобулина, которая либо не придает специфичность в отношении связывания антигена, но может объединяться с каждой тяжелой цепью, либо которая может объединяться с каждой тяжелой цепью и может связываться с одним или более эпитопами, связываемыми антигенсвязывающими участками тяжелой цепи, либо которая может объединяться с каждой тяжелой цепью и обеспечивать связывание одной или обеих тяжелых цепей с одним или обоими эпитопами. В одном варианте осуществления Fc-домен содержит по меньшей мере CH2 и CH3. Fc-домен может содержать шарнир, домен CH2 и домен CH3.

Один вариант осуществления биспецифического формата включает первую тяжелую цепь (HC), вторую тяжелую цепь, которая имеет модифицированный CH3 (HC*), и общую легкую цепь (LC) (две копии одной и той же легкой цепи). Другой вариант осуществления включает первую тяжелую цепь (HC), общую LC и слитый полипептид HC-ScFv (где вторая HC является слитой с N-концом ScFv). Другой вариант осуществления включает первую HC, родственную LC, слитый полипептид HC-ScFv (где второй HC является слитым с N-концом ScFv). Другой вариант осуществления включает первую тяжелую цепь (HC), LC и Fc-домен. Другой вариант осуществления включает первую HC, LC, слитый полипептид ScFv-Fc (где Fc является слитым с C-концом ScFv). Другой вариант осуществления включает первую HC, общую LC и слитый полипептид Fc-ScFv (где Fc является слитым с N-концом ScFv). Другой вариант осуществления включает первую HC, LC и ScFv-HC (где второй HC является слитым с C-концом ScFv).

В определенных вариантах осуществления одна тяжелая цепь (HC) может представлять собой нативную последовательность или последовательность "дикого типа", а вторая тяжелая цепь может являться модифицированной в Fc-домене. В других вариантах осуществления одна тяжелая цепь (HC) может представлять собой нативную последовательность или последовательность "дикого типа", а вторая тяжелая цепь может являться модифицированной по кодонам.

Термин "клетка" предусматривает любую клетку, которая является подходящей для экспрессии рекомбинантной последовательности нуклеиновой кислоты и имеет локус, который обеспечивает стабильную интеграцию и повышенную экспрессию экзогенной нуклеиновой кислоты. Клетки предусматривают клетки млекопитающих, такие как клетки животного, отличного от человека, клетки человека или продукты слияния клеток, такие как, например, гибридомы или квадромы. В некоторых вариантах осуществления клетка представляет собой клетку человека, обезьяны, человекообразной обезьяны, хомячка, крысы или мыши. В некоторых вариантах осуществления клетка представляет собой клетку млекопитающего, выбранную из следующих клеток: CHO K1, DXB-11 CHO, Veggie-CHO, COS (например, COS-7), клетки сетчатки, Vero, CV1, клетки почки (например, HEK293, 293 EBNA, MSR 293, MDCK, HaK, BHK), HeLa, HepG2, WI38, MRC 5, Colo205, HB 8065, HL-60 (например, BHK21), Jurkat, Daudi, A431 (эпидермальные), CV-1, U937, 3T3, L-клетки, клетки C127, SP2/0, NS-0, MMT 060562, клет-

ки Сертоли, клетки BRL 3A, клетки HT1080, клетки миеломы, опухолевые клетки и линии клеток, полученные из вышеупомянутых клеток. В некоторых вариантах осуществления клетка содержит один или более вирусных генов, например, клетка сетчатки, которая экспрессирует вирусный ген (например, клетка PER.C6™).

"Плотность клеток" относится к количеству клеток на объем образца, например, в виде общего количества клеток (жизнеспособных и мертвых) на мл. Количество клеток можно подсчитывать вручную или с помощью автоматизированной системы, такой как проточный питомер. Автоматические счетчики клеток были адаптированы для подсчета количества жизнеспособных или мертвых клеток или как жизнеспособных, так и мертвых клеток с применением, например, стандартной методики с поглощением трипанового синего. Фраза "плотность жизнеспособных клеток" или "концентрация жизнеспособных клеток" относится к количеству жизнеспособных клеток на объем образца (также называемому "числом жизнеспособных клеток"). Любое количество хорошо известных проводимых вручную или автоматизированных методик можно применять для определения плотности клеток. Параметры биомассы культуры можно измерять в реальном времени, где электропроводность или оптическая плотность коррелирует с количеством клеток на объем. Конечная плотность клеток в культуре клеток, такой как производственная культура, варьирует в зависимости от исходной линии клеток, например, в диапазоне от приблизительно 1,0 до 10×10^6 клеток/мл. В некоторых вариантах осуществления достигают конечной плотности клеток, составляющей от 1,0 до 10×10^6 клеток/мл, перед сбором белка, представляющего интерес, из производственной культуры клеток. В других вариантах осуществления достигают конечной плотности клеток, составляющей по меньшей мере $5,0 \times 10^6$ клеток/мл, по меньшей мере 6×10^6 клеток/мл, по меньшей мере 7×10^6 клеток/мл, по меньшей мере 8×10^6 клеток/мл, по меньшей мере 9×10^6 клеток/мл или по меньшей мере 10×10^6 клеток/мл.

Термин "модифицированная по кодонам" подразумевает, что нуклеотидная последовательность, кодирующая белок, была модифицирована по одному или более нуклеотидам, т. е. по одному или более кодонам без изменений аминокислот, кодируемых кодоном, с получением модифицированной по кодонам версии нуклеотидной последовательности. Модификация по кодонам нуклеотидной последовательности может обеспечить подходящее основание для того, чтобы различать нуклеотидную последовательность от ее модифицированной по кодонам версии в анализе, основанном на применении нуклеиновой кислоты (например, среди прочего, в анализе на основе гибридизации, ПЦР). В некоторых случаях кодоны нуклеотидной последовательности модифицировали для обеспечения улучшенной или оптимизированной экспрессии кодируемого белка в клетке-хозяине посредством применения хорошо известных из уровня техники методик оптимизации кодонов (Gustafsson, C, et al., 2004, Trends in Biotechnology, 22:346-353; Chung, B.K.-S., et al., 2013, Journal of Biotechnology, 167:326-333; Gustafsson, C, et al., 2012, Protein Expr Purif, 83(1): 37-46). Инструменты в виде программного обеспечения для конструирования последовательностей с применением таких методик также являются хорошо известными из уровня техники, включая без ограничения, среди прочего, Codon optimizer (Fuglsang A. 2003, Protein Expr Purif, 31:247-249), Gene Designer (Villalobos A, et al., 2006, BMC Bioinforma, 7:285) и OPTIMIZER (Puigbo P, et al., 2007, Nucleic Acids Research, 35:W126-W131).

Фраза "определяющий комплементарность участок" или термин "CDR" предусматривают аминокислотную последовательность, кодируемую последовательностью нуклеиновой кислоты генов иммуноглобулина организма, которая в норме (т. е. у животного дикого типа) находится между двумя каркасными участками в переменном участке легкой или тяжелой цепи молекулы иммуноглобулина (например, антитела или рецептора Т-клетки). CDR может кодироваться, например, последовательностью зародышевого типа либо перегруппированной или неперегруппированной последовательностью, и, например, непримированной или зрелой В-клеткой или Т-клеткой. В некоторых случаях (например, в случае CDR3) CDR могут кодироваться двумя или более последовательностями (например, последовательностями зародышевого типа), которые не являются смежными (например, в случае неперегруппированной последовательности нуклеиновой кислоты), но они являются смежными в последовательности нуклеиновой кислоты в В-клетке, например, как результат сплайсинга или соединения последовательностей (например, V-D-J-рекомбинация с образованием CDR3 тяжелой цепи).

Термин "локус, обеспечивающий повышенную экспрессию" относится к локусу в геноме клетки, который содержит последовательность или последовательности и характеризуется более высоким уровнем экспрессии в сравнении с другими участками или последовательностями в геноме, если подходящий ген или конструкция являются экзогенно добавленными (т. е. интегрированными) в или вблизи последовательности или последовательностей или "функционально связанными" с последовательностью или последовательностями.

Термин "повышенный" при использовании для описания повышенной экспрессии предусматривает повышение экспрессии от по меньшей мере в приблизительно 1,5 раза до повышения по меньшей мере в приблизительно 3 раза по сравнению с тем, что обычно наблюдают при случайной интеграции экзогенной последовательности в геном или при интеграции в другой локус, например, по сравнению с популяцией случайных интегрантов отдельной копии одной и той же конструкции для экспрессии. Наблюдает-

мое кратное повышение экспрессии при применении последовательностей по настоящему изобретению сравнивают с уровнем экспрессии одного и того же гена, измеряемым при практически таких же условиях в отсутствие последовательности по настоящему изобретению, например, в сравнении с интеграцией в другой локус в геноме одного и того же вида. Повышенная эффективность рекомбинации предусматривает повышение способности локуса к рекомбинации (например, с применением сайтов распознавания для рекомбиназы ("RRS")). Повышение относится к эффективности рекомбинации по сравнению со случайной рекомбинацией, например, без применения сайтов распознавания для рекомбиназы или подобно-го, которая обычно составляет 0,1%. Предпочтительная эффективность повышенной рекомбинации приблизительно в 10 раз превышает случайную или составляет приблизительно 1%. Если не указано иное, заявленное изобретение не ограничено конкретной эффективностью рекомбинации. Локусы, обеспечивающие повышенную экспрессию, обычно поддерживают высокий уровень выработки белка, представляющего интерес, в клетке-хозяине. Следовательно, повышенная экспрессия предусматривает высокий уровень выработки белка, представляющего интерес (повышенный титр в граммах белка), в клетке, а не достижение высоких титров просто посредством большого количества копий клеток в культуре. Удельная продуктивность Q_p (пг/клетка/день, т. е. pcd) считается мерой устойчивой продуктивности. Желательными являются рекомбинантные клетки-хозяева, характеризующиеся Q_p , составляющей более 5 pcd, или более 10 pcd, или более 15 pcd, или более 20 pcd, или более 25 pcd, или даже более 30 pcd. Клетки-хозяева с геном, представляющим интерес, вставленным в локус, обеспечивающий повышенную экспрессию, или в "горячую точку", характеризуются высокой удельной продуктивностью.

Если фраза "экзогенно добавленный ген", "экзогенно добавленная нуклеиновая кислота" или просто "экзогенная нуклеиновая кислота" применяется в отношении локуса, представляющего интерес, то данная фраза относится к любой последовательности ДНК или гену, не присутствующим в локусе, представляющем интерес, как в локусе, встречающемся в природе. Например, "экзогенная нуклеиновая кислота" в локусе СНО (например, в локусе, содержащем последовательность под SEQ ID NO: 1 или SEQ ID NO: 2), может представлять собой ген хомячка, не встречающийся в конкретном локусе СНО в природе (т. е. ген хомячка из другого локуса генома хомячка), ген от любого другого вида (например, ген человека), химерный ген (например, человека/мыши) или любой другой ген, который не встречается в природе как содержащийся в локусе СНО, представляющем интерес.

Фраза "тяжелая цепь" или "тяжелая цепь иммуноглобулина" предусматривает последовательность константного участка тяжелой цепи иммуноглобулина из любого организма и, если не указано иное, включает переменный домен тяжелой цепи. Переменные домены тяжелой цепи содержат три CDR тяжелой цепи и четыре FR-участка, если не указано иное. Типичная тяжелая цепь после переменного домена имеет (от N-конца к С-концу) домен CH1, шарнир, домен CH2 и домен CH3. Термин "фрагмент из тяжелой цепи" или "фрагмент тяжелой цепи" (также называемый в данном документе "HCF") предусматривает пептид, состоящий из по меньшей мере 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 или больше аминокислот тяжелой цепи, и может предусматривать один или более CDR, один или более CDR, комбинированных с одним или более FR, один или более CH1, шарниров, CH2 или CH3, переменный участок, константный участок, фрагменты константного участка (например, CH1, CH2, CH3) или их комбинации. Примеры HCF включают VH и полные Fc-участки или их части. Фраза "нуклеотидная последовательность, кодирующая HCF" предусматривает нуклеотидные последовательности, кодирующие полипептид, состоящий из HCF, и нуклеотидные последовательности, кодирующие полипептид, содержащий HCF, например, полипептиды, которые могут содержать дополнительные аминокислоты кроме указанного HCF. Например, нуклеотидная последовательность, кодирующая HCF, включает нуклеотидные последовательности, кодирующие полипептиды, состоящие из VH, состоящие из VH, связанного с CH3, состоящие, среди прочего, из полной тяжелой цепи.

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

Фраза "легкая цепь" предусматривает последовательность константного участка легкой цепи иммуноглобулина из любого организма и, если не указано иное, включает каппа и лямбда легкие цепи человека. Переменные домены легкой цепи (VL) обычно содержат три CDR легкой цепи и четыре каркасных (FR) участка, если не указано иное. Как правило, полноразмерная легкая цепь содержит от аминоконца к карбоксильному концу: домен VL, который содержит FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, и константный домен легкой цепи. Легкие цепи, которые можно применять согласно настоящему изобретению, включают, например, такие, которые не связываются селективно либо с первым, либо со вторым эпитопом, селективно связываемым биспецифическим антителом. Подходящие легкие цепи также включают такие, которые могут связываться с одним или обоими эпитопами, которые связываются антиген-связывающими участками антитела, или способствовать их связыванию. Термин "фрагмент из легкой

цепи" или "фрагмент легкой цепи" (или "LCF") предусматривает пептид, состоящий из по меньшей мере 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 или больше аминокислот легкой цепи, и может включать один или более CDR, один или более CDR, комбинированных с одним или более FR, варибельный участок, константный участок, фрагменты константного участка или их комбинации. Примеры LCF включают VL и полные константные участки легкой цепи ("CL") или их части. Фраза "нуклеотидная последовательность, кодирующая LCF" предусматривает нуклеотидные последовательности, кодирующие полипептид, состоящий из LCF, и нуклеотидные последовательности, кодирующие полипептид, содержащий LCF, например, полипептиды, которые могут содержать дополнительные аминокислоты кроме указанного LCF. Например, нуклеотидная последовательность, кодирующая LCF, включает нуклеотидные последовательности, кодирующие полипептиды, состоящие из VL, или состоящие, среди прочего, из полной легкой цепи.

Фраза "функционально связанный" относится к связыванию нуклеиновых кислот или белков таким образом, что связанные молекулы функционируют надлежащим образом. Участки ДНК являются функционально связанными, если они функционально сопряжены 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, что обеспечивает трансляцию. Как правило, "функционально связанный" может включать, но не требует, смежное расположение. В случае последовательностей, таких как лидерные последовательности, отвечающие за секрецию, смежное расположение и правильность размещения в рамке считывания представляют собой типичные признаки. Последовательность, обеспечивающая повышенную экспрессию, в локусе, представляющем интерес, функционально связана с геном, представляющим интерес (GOI), при этом она функционально сопряжена с GOI, например, где ее присутствие приводит к повышенной экспрессии GOI.

"Процентная идентичность" при описании локуса, представляющего интерес, такого как SEQ ID NO: 1 или SEQ ID NO: 2, или его фрагмента подразумевает включение гомологичных последовательностей, которые характеризуются упомянутой идентичностью в пределах участков непрерывной гомологии, но 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 будет включать сравнение последовательностей, где видовой гомолог не имеет гомологичной последовательности для сравнения при выравнивании (т. е. SEQ ID NO: 1 или ее фрагмент имеет вставку в этой точке, или видовой гомолог имеет гэп или делецию, в зависимости от обстоятельств). Таким образом, "процентная идентичность" не включает штрафы за гэпы, делеций и вставки.

"Сайт распознавания" или "последовательность распознавания" представляет собой специфическую последовательность ДНК, распознаваемую нуклеазой или другим ферментом с целью связывания и направления сайт-специфического расщепления остова ДНК. Эндонуклеазы расщепляют ДНК в пределах молекулы ДНК. Сайты распознавания также называются в данной области техники целевыми сайтами распознавания.

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

"Обмен кассетами, опосредованный рекомбиназой", или "RMCE", относится к способу точного замещения геномной целевой кассеты донорной кассетой. Молекулярные композиции, обычно предусмотренные для осуществления данного способа, включают: 1) геномную целевую кассету, фланкированную как с 5'-конца, так и с 3'-конца целевыми сайтами распознавания, специфическими для конкретной рекомбиназы, 2) донорную кассету, фланкированную соответствующими целевыми сайтами распознавания, и 3) сайт-специфическую рекомбиназу. Белки-рекомбиназы хорошо известны из уровня техники (Turan, S. and Bode J., 2011, *FASEB J.*, 25, pp. 4088-4107) и обеспечивают точное расщепление ДНК в конкретном целевом сайте распознавания (последовательности ДНК) без добавления или потери нуклеотидов. Общепринятые комбинации рекомбиназа/сайт включают без ограничения Cre/lox и Flp/rt. В коммерчески доступных наборах также предусмотрены векторы, содержащие сайт R4-attP, и вектор, кодирующий интегразу phiC31 для RMCE. (см. также, например, опубликованную заявку на патент США с № US20130004946.)

"Сайт-специфическая интеграция" или "нацеленная вставка" относятся к способам нацеливания на

гены, применяемым для непосредственной вставки или интеграции гена или последовательности нуклеиновой кислоты в конкретное положение в геноме, т. е. для направления ДНК к конкретному сайту между двумя нуклеотидами в непрерывной полинуклеотидной цепи. Сайт-специфическую интеграцию или нацеленную вставку можно также осуществлять в отношении конкретной нуклеиновой кислоты, которая включает несколько единиц или кассет экспрессии, таких как несколько генов, каждый из которых имеет собственные регуляторные элементы (такие как промоторы, энхансеры и/или последовательности терминации транскрипции). "Вставку" и "интеграцию" применяют 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 применяемой методики редактирования генома.

Под "стабильной интеграцией" подразумевается, что экзогенная нуклеиновая кислота, интегрированная в геном клетки-хозяина, остается интегрированной в течение длительного периода времени в культуре клеток, например, по меньшей мере 7 дней, по меньшей мере 10 дней, по меньшей мере 15 дней, по меньшей мере 20 дней, по меньшей мере 25 дней, по меньшей мере 30 дней, по меньшей мере 35 дней, по меньшей мере 40 дней, по меньшей мере 45 дней, по меньшей мере 50 дней, по меньшей мере 55 дней, по меньшей мере 60 дней или дольше. Понятно, что получение биспецифических антигенсвязывающих белков с целью получения и очищения в крупных масштабах представляет собой сложную задачу. Стабильность и клональные свойства являются необходимыми для воспроизводимости любой биомолекулы, особенно предназначенной для терапевтического применения. Стабильные клоны, экспрессирующие биспецифические антитела, полученные посредством способов согласно настоящему изобретению, обеспечивают надежный и воспроизводимый способ получения терапевтических биомолекул.

Общее описание

В настоящем изобретении предусмотрены композиции и способы для обеспечения улучшенной экспрессии нескольких полипептидов в клетке-хозяине, в частности в линиях клеток китайского хомячка (*Cricetus griseus*), посредством применения локуса, обеспечивающего повышенную экспрессию, в клетке-хозяине.

Более конкретно, в настоящем изобретении предусмотрены композиции и способы, предназначенные для интеграции нескольких экзогенных нуклеиновых кислот, которые вместе кодируют биспецифический антигенсвязывающий белок, в конкретный сайт в локусе, обеспечивающем повышенную экспрессию, в клетке-хозяине, такой как клетка CHO. В частности, в настоящем изобретении предусмотрены клетки, содержащие несколько экзогенных нуклеиновых кислот, интегрированных в конкретный сайт в локусе, обеспечивающем повышенную экспрессию, где несколько экзогенных нуклеиновых кислот вместе кодируют биспецифический антигенсвязывающий белок. В настоящем изобретении дополнительно предусмотрены векторы на основе нуклеиновой кислоты, сконструированные для сайт-специфической интеграции нескольких экзогенных нуклеиновых кислот в локус, обеспечивающий повышенную экспрессию. В настоящем изобретении дополнительно предусмотрены системы, которые включают клетку-хозяина, содержащую два или более сайтов распознавания для рекомбиназы (RRS), и набор векторов, содержащих соответствующие RRS и несколько экзогенных нуклеиновых кислот, для сайт-специфической интеграции нескольких экзогенных нуклеиновых кислот из векторов в локус, обеспечивающий повышенную экспрессию. Кроме того, в настоящем изобретении предусмотрены способы получения биспецифического антигенсвязывающего белка с применением клеток, векторов и систем, раскрытых в данном документе.

Клетки, содержащие несколько экзогенных нуклеиновых кислот, интегрированных в конкретный сайт в локусе, обеспечивающем повышенную экспрессию

В одном аспекте настоящее изобретение предусматривает клетку, содержащую экзогенную последовательность нуклеиновой кислоты, интегрированную в конкретный сайт в локусе, обеспечивающем повышенную экспрессию, где экзогенная последовательность нуклеиновой кислоты кодирует биспецифический антигенсвязывающий белок.

Клетки, предусмотренные в данном документе, способны продуцировать биспецифический антигенсвязывающий белок (например, биспецифическое антитело) с высокими титрами и/или высокой удельной продуктивностью (пг/клетка/день). В некоторых вариантах осуществления клетка продуцирует биспецифический антигенсвязывающий белок с титром, составляющим по меньшей мере 5 мг/л, 10 мг/л, 15 мг/л, 20 мг/л, 25 мг/л, 30 мг/л, 35 мг/л, 40 мг/л, 45 мг/л, 50 мг/л или больше. В некоторых вариантах осуществления клетка продуцирует биспецифический антигенсвязывающий белок при соотношении титра биспецифического антигенсвязывающего белка и общего титра антигенсвязывающих белков, составляющем по меньшей мере 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50% или больше. В некоторых вариантах осуществления клетка, которая продуцирует биспецифический антигенсвязывающий белок, характеризуется удельной продуктивностью, составляющей по меньшей мере 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 пикограмм/клетка/день (pgd) или больше, которую определяют на основании общего количества антигенсвязывающих белков (в пг), продуцируемых клеткой в день.

Клетки-хозяева, содержащие экзогенную последовательность нуклеиновой кислоты, кодирующую

биспецифический антигенсвязывающий белок, интегрированную в конкретный сайт в локусе, обеспечивающем повышенную экспрессию, характеризуются высокой плотностью клеток в производственной культуре, например, $1-10 \times 10^6$ клеток/мл. В других вариантах осуществления клетка-хозяин, в которой кодируется биспецифический антигенсвязывающий белок, характеризуется конечной плотностью клеток, составляющей по меньшей мере 5×10^6 клеток/мл, 6×10^6 клеток/мл, 7×10^6 клеток/мл, 8×10^6 клеток/мл, 9×10^6 клеток/мл или 10×10^6 клеток/мл.

В некоторых вариантах осуществления биспецифический антигенсвязывающий белок содержит два HC-фрагмента ("HCF"), характеризующихся различными специфичностями в отношении антигена, и два LCF. В случаях, если применяются два участка VL, они могут быть одинаковыми или отличающимися. В конкретных вариантах осуществления два участка VL являются одинаковыми, как, например, общая легкая цепь.

В некоторых вариантах осуществления каждый из двух HCF содержит аминокислоты из константного участка тяжелой цепи, такого как CH1, CH2 или CH3. В конкретных вариантах осуществления каждый из двух HCF содержит домен CH3. В конкретных вариантах осуществления каждый из двух HCF содержит константный участок, т. е. полный константный участок.

В некоторых вариантах осуществления каждый из двух HCF содержит VH, и при этом два VH могут быть одинаковыми или отличающимися.

В некоторых вариантах осуществления биспецифический антигенсвязывающий белок содержит две тяжелые цепи (т. е. две полные тяжелые цепи).

В некоторых вариантах осуществления каждый из двух LCF содержит VL. В конкретных вариантах осуществления каждый LCF состоит из участка VL, который функционально связан с аминокислотной последовательностью, которая содержит аминокислоты из константного участка легкой цепи. В конкретных вариантах осуществления каждый участок VL функционально связан с участком CL, т. е. биспецифический антигенсвязывающий белок содержит легкую цепь (т. е. полную легкую цепь).

В некоторых вариантах осуществления экзогенная последовательность нуклеиновой кислоты, интегрированная в локус, обеспечивающий повышенную экспрессию, включает первую экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF, вторую экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF, и третью экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую второй HCF.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый LCF, может кодировать последовательность переменного участка легкой цепи (VL). В конкретных вариантах осуществления нуклеотидная последовательность, кодирующая первый участок VL, кодирует первую легкую цепь.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует аминокислоты из константного участка первой тяжелой цепи, например, одного или более из CH1, шарнира, CH2 или CH3), а нуклеотидная последовательность, кодирующая второй HCF, кодирует аминокислоты из константного участка второй тяжелой цепи. Аминокислоты из константного участка первой тяжелой цепи могут быть одинаковыми или отличающимися по сравнению с аминокислотами из константного участка второй тяжелой цепи. Например, нуклеотидная последовательность, кодирующая первый HCF, кодирует первый домен CH3, а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй домен CH3, где первый и второй домены CH3 могут быть одинаковыми или отличаться в одном или более аминокислотных положений, как описано в данном документе ниже для биспецифических антигенсвязывающих белков.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует первый VH, а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй VH.

В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая первый HCF, кодирует первую тяжелую цепь, а нуклеотидная последовательность, кодирующая второй HCF, кодирует вторую тяжелую цепь. Первая и вторая тяжелые цепи могут иметь одинаковые константные участки или отличаться по одной или более аминокислотам. Различные примеры биспецифических антигенсвязывающих белков, которые имеют отличающиеся константные домены тяжелых цепей (например, отличающиеся домены CH3), дополнительно описаны в данном документе ниже. Независимо от кодируемых аминокислотных последовательностей, нуклеотидные последовательности, кодирующие аминокислоты из двух константных участков тяжелых цепей, могут отличаться тем, что одна из двух кодирующих нуклеотидных последовательностей может быть модифицированной по кодонам, что обеспечивает подходящее основание для того, чтобы различать две нуклеотидные последовательности в анализе выявления, основанном на применении нуклеиновых кислот.

В некоторых вариантах осуществления каждая нуклеотидная последовательность, кодирующая HCF или LCF, независимо и функционально связана с последовательностью, регулирующей транскрипцию, которая предусматривает промотор. Под "независимо" подразумевается, что каждая кодирующая

последовательность функционально связана с отдельной последовательностью, регулирующей транскрипцию, такой как промотор, таким образом, что транскрипция кодирующих последовательностей находится под регуляцией и контролем отдельных элементов. В некоторых вариантах осуществления промоторы, управляющие транскрипцией двух полипептидов, содержащих HCF, являются одинаковыми. В некоторых вариантах осуществления все промоторы, управляющие транскрипцией двух полипептидов, содержащих HCF, а также промотор, управляющий транскрипцией полипептида, содержащего LCF, являются одинаковыми, например, представляют собой промотор CMV. В некоторых вариантах осуществления каждая нуклеотидная последовательность, кодирующая HCF или LCF, независимо и функционально связана с индуцируемым или репрессируемым промотором. Индуцируемые или репрессируемые промоторы обеспечивают возможность того, чтобы продуцирование происходило, например, только в фазе продуцирования (культура с подпиткой), а не в ходе фазы роста (культура, выращиваемая в системе посевных ферментеров). Точного контроля продуцирования (экспрессии) продукта каждого гена можно достигать посредством разных промоторов.

В одном таком примере клетки сначала конструируют так, чтобы они экспрессировали тетрациклинзависимый репрессорный белок (TetR), и каждую нуклеотидную последовательность, кодирующую HCF и LCF, помещают под транскрипционный контроль промотора, активность которого регулирует TetR. Два расположенных друг за другом оператора TetR (TetO) помещают непосредственно ниже промотора CMV. В некоторых вариантах осуществления каждая нуклеотидная последовательность, кодирующая HCF и/или LCF, независимо и функционально связана с промотором, расположенным выше по меньшей мере одного оператора TetR (TetO) или оператора Arc (ArcO). В других вариантах осуществления каждая нуклеотидная последовательность, кодирующая HCF и/или LCF, независимо и функционально связана с гибридными промоторами CMV/TetO или CMV/ArcO. Дополнительные подходящие промоторы описаны в данном документе ниже.

Относительные положения нескольких экзогенных нуклеиновых кислот в локусе могут меняться. Не вдаваясь в какую-либо теорию, считается, что важно достичь сбалансированных (т. е., сравнимых) уровней экспрессии двух полипептидов, содержащих HCF. В некоторых вариантах осуществления нуклеиновая кислота, кодирующая LCF, расположена выше относительно обеих нуклеиновых кислот, кодирующих HCF. В случаях, если три промотора для управления экспрессией полипептида, содержащего LCF, и двух полипептидов, содержащих HCF, являются одинаковыми, то подходящий порядок расположения может предусматривать: в направлении от 5' к 3' нуклеотидную последовательность, кодирующую LCF, нуклеотидную последовательность, кодирующую первый HCF, дополнительный отличающийся промотор, функционально связанный с нуклеотидной последовательностью (такой как ген селективируемого маркера), и нуклеотидную последовательность, кодирующую второй HCF. Другие варианты подходящего порядка расположения предусматривают: в направлении от 5' к 3' нуклеотидную последовательность, кодирующую LCF, дополнительный отличающийся промотор, функционально связанный с нуклеотидной последовательностью (такой как ген селективируемого маркера), нуклеотидную последовательность, кодирующую первый HCF, и нуклеотидную последовательность, кодирующую второй HCF. Если нуклеотидные последовательности, кодирующие HCF, кодируют последовательность константного участка, то либо нуклеотидная последовательность, расположенная выше, может кодировать модифицированную версию последовательности константного участка (например, модифицированный CH3), либо нуклеотидная последовательность, расположенная ниже, может кодировать модифицированную версию последовательности константного участка, при этом одна из них кодирует немодифицированную версию последовательности константного участка.

В некоторых вариантах осуществления клетка дополнительно содержит один или более RRS, также интегрированных в локус. В некоторых вариантах осуществления клетка содержит первый и второй RRS, которые отличаются друг от друга и фланкируют экзогенную последовательность нуклеиновой кислоты, где экзогенная последовательность нуклеиновой кислоты, в свою очередь, содержит нуклеиновую кислоту, кодирующую первый LCF, нуклеиновую кислоту, кодирующую первый HCF, и нуклеиновую кислоту, кодирующую второй HCF. В конкретных вариантах осуществления нуклеиновая кислота, кодирующая LCF, расположена выше относительно обеих нуклеиновых кислот, кодирующих HCF, и клетка содержит третий RRS, расположенный в направлении 3' относительно нуклеиновой кислоты, кодирующей первый LCF, и в направлении 5' относительно одной или обеих экзогенных нуклеиновых кислот, кодирующих HCF, где третий RRS отличается от первого и второго RRS. Третий RRS можно сконструировать для включения в интрон гена, который можно поместить между любыми двумя из последовательностей, кодирующих HCF или LCF.

Биспецифические антигенсвязывающие белки

Биспецифические антигенсвязывающие белки, такие как биспецифические антитела, подходящие для клонирования и продуцирования в клетках с помощью векторов и систем, описанных в настоящем изобретении, не ограничиваются любым конкретным форматом биспецифических антигенсвязывающих белков.

В различных вариантах осуществления биспецифический антигенсвязывающий белок содержит два полипептида, каждый из которых содержит антигенсвязывающий фрагмент (например, участок VH) и

домен СНЗ, где антигенсвязывающий фрагмент в двух полипептидах характеризуется разными специфичностями в отношении антигена, и где два домена СНЗ являются гетеродимерными по отношению друг к другу, поскольку один из доменов СНЗ был модифицирован в по меньшей мере одном аминокислотном положении для получения различных характеристик связывания с белком А у двух полипептидов. См., например, биспецифические антитела, описанные в патенте США 8586713. Таким образом, схему выделения, основанную на различном связывании с белком А, можно использовать для легкого отделения гетеродимерных биспецифических антигенсвязывающих белков от гомодимеров.

В некоторых вариантах осуществления биспецифический антигенсвязывающий белок включает две тяжелые цепи, которые характеризуются разными специфичностями в отношении антигена и отличаются доменом СНЗ в по меньшей мере одном аминокислотном положении, что обеспечивает различные характеристики связывания с белком А у двух тяжелых цепей.

В некоторых вариантах осуществления два полипептида содержат домены СНЗ IgG человека, где один из двух полипептидов содержит домен СНЗ IgG человека, выбранного из IgG1, IgG2 и IgG4, и другой из двух полипептидов содержит модифицированный домен СНЗ IgG человека, выбранного из IgG1, IgG2 и IgG4, где модификация снижает или устраняет связывание модифицированного участка СНЗ с белком А. В конкретных вариантах осуществления один из двух полипептидов содержит домен СНЗ IgG1 человека, и другой из двух полипептидов содержит модифицированный домен СНЗ IgG1 человека, где модификация выбрана из группы, состоящей из (i) 95R и (ii) 95R и 96F согласно системе нумерации экзонов IMGT. В других конкретных вариантах осуществления модифицированный домен СНЗ содержит от одной до пяти дополнительных модификаций, выбранных из группы, состоящей из 16E, 18M, 44S, 52N, 57M и 82I согласно системе нумерации экзонов IMGT.

В других различных вариантах осуществления два полипептида содержат домены СНЗ IgG мыши, где один из двух полипептидов содержит домен СНЗ немодифицированного IgG мыши, и другой из двух полипептидов содержит модифицированный домен СНЗ IgG мыши, где модификация снижает или устраняет связывание модифицированного участка СНЗ с белком А. В различных вариантах осуществления участок СНЗ IgG мыши модифицируют так, чтобы он содержал конкретные аминокислоты в конкретных положениях (нумерация согласно EU), выбранных из группы, состоящей из 252T, 254T и 256T; 252T, 254T, 256T и 258K; 247P, 252T, 254T, 256T и 258K; 435R и 436F; 252T, 254T, 256T, 435R и 436F; 252T, 254T, 256T, 258K, 435R и 436F; 24tP, 252T, 254T, 256T, 258K, 435R и 436F; а также 435R. В конкретном варианте осуществления сделана конкретная группа модификаций, выбранная из групп, состоящих из: M252T, S254T, S256T; M252T, S254T, S256T, I258K; I247P, M252T, S254T, S256T, I258K; H435R, H436F; M252T, S254T, S256T, H435R, H436F; M252T, S254T, S256T, I258K, H435R, H436F; I247P, M252T, S254T, S256T, I258K, H435R, H436F; а также H435R.

В различных вариантах осуществления биспецифический антигенсвязывающий белок представляет собой гибрид моноклонального антитела или антигенсвязывающего белка мыши и крысы, например, гибрид IgG2a мыши и IgG2b крысы. Согласно данным вариантам осуществления биспецифическое антитело состоит из гетеродимера на основе двух антител, содержащего по одной паре тяжелой/легкой цепей каждого, которые соединены посредством их Fc-частей. Требуемый гетеродимер можно легко очистить из смеси двух исходных антител, являющихся гомодимерами, и биспецифического гетеродимера, поскольку свойства связывания биспецифического антитела с белком А отличаются от таковых у исходных антител: IgG2b крысы не связывается с белком А, тогда как IgG2a мыши связывается. Следовательно, гетеродимер мыши-крысы связывается с белком А, но элюируется при более высоком значении pH, чем гомодимер IgG2a мыши, и это обеспечивает возможность селективной очистки биспецифического гетеродимера.

В других различных вариантах осуществления биспецифический антигенсвязывающий белок представлен в формате, который в данной области техники называется "выступы-во-впадины" (см., например, патент США № 7183076). В данных вариантах осуществления Fc-части двух антител сконструированы с обеспечением одного выпяченного "выступа" и другой комплементарной "впадины". В случае продуцирования в одной и той же клетке считается, что тяжелые цепи предпочтительно образуют гетеродимеры, а не гомодимеры, посредством связывания сконструированных "выступов" со сконструированными "впадинами".

В другом варианте осуществления первая тяжелая цепь и вторая тяжелая цепь содержат одну или более аминокислотных модификаций в домене СНЗ для обеспечения возможности взаимодействия между двумя тяжелыми цепями. Аминокислотные остатки в области контакта СНЗ-СНЗ можно заменять заряженной аминокислотой с обеспечением электростатически неблагоприятного состояния для образования гомодимера. (см., например, публикацию согласно PCT № WO 2009089004 и публикацию европейского патента № EP 1870459).

В других вариантах осуществления первая тяжелая цепь содержит домен СНЗ изотипа IgA, а вторая тяжелая цепь содержит домен СНЗ IgG (или наоборот) для содействия предпочтительному образованию гетеродимеров. (см., например, публикацию согласно PCT № WO 2007110205).

В других вариантах осуществления посредством способов конструирования цепей иммуноглобулина можно обеспечивать различные форматы для содействия образованию гетеродимеров, как, например,

обмен Fab-плечами (публикация согласно PCT №WO 2008119353; публикация согласно PCT № WO 2011131746), взаимодействие с суперспирализованным доменом (публикация согласно PCT № WO 2011034605) или образование пептидов с лейциновыми застежками (Kostelny, et al., J. Immunol. 1992, 148(5): 1547-1553).

Вариабельные участки тяжелой цепи иммуноглобулина, которые можно применять для получения биспецифических антигенсвязывающих белков, можно получать с применением любого способа, известного из уровня техники.

Например, первая тяжелая цепь содержит вариабельный участок, кодируемый нуклеиновой кислотой, которая получена из генома зрелой В-клетки первого животного, которого иммунизировали первым антигеном, при этом первая тяжелая цепь специфически распознает первый антиген; и вторая тяжелая цепь содержит вариабельный участок, кодируемый нуклеиновой кислотой, которая получена из генома зрелой В-клетки второго животного, которого иммунизировали вторым антигеном, при этом вторая тяжелая цепь специфически распознает второй антиген. Последовательности вариабельного участка тяжелой цепи иммуноглобулина также можно получить с помощью любого другого способа, известного из уровня техники, например, с помощью фагового дисплея. В других примерах нуклеиновые кислоты, кодирующие вариабельные участки тяжелой цепи, включают таковые от антител, которые были описаны или иным образом доступны из уровня техники. В некоторых вариантах осуществления одна из двух последовательностей, кодирующих тяжелую цепь, была модифицирована по кодонам с целью обеспечения подходящего основания для того, чтобы различать две кодирующие последовательности в анализах, основанных на применении нуклеиновой кислоты.

Биспецифические антитела, содержащие две тяжелые цепи, которые распознают два разных эпитопа (или два разных антигена), легче выделять, если они могут образовывать пары с одной и той же легкой цепью (т. е. легкими цепями, содержащими идентичные вариабельные и константные домены). Из уровня техники известен ряд способов получения легких цепей, которые могут образовывать пары с двумя тяжелыми цепями с различной специфичностью, при этом не нарушая или существенно не нарушая селективность и/или аффинность вариабельного домена тяжелой цепи в отношении его целевого антигена, как описано, например, в патенте США 8586713 и источниках, раскрытых в нем.

Биспецифические антигенсвязывающие белки могут характеризоваться двойными специфичностями в отношении различных антигенов и связанными с этим полезными применениями.

В некоторых примерах можно получить биспецифические антигенсвязывающие белки, демонстрирующие специфичность в отношении связывания опухолевого антигена и Т-клеточного антигена, которые нацеливаются на антиген на клетке, например, CD20, а также нацеливаются на антиген на Т-клетке, например, на рецептор Т-клетки, такой как CD3. Таким образом, биспецифический антигенсвязывающий белок нацеливается как на представляющую интерес клетку у пациента (например, В-клетку у пациента с лимфомой, посредством связывания с CD20), так и на Т-клетку у пациента. В различных вариантах осуществления биспецифический антигенсвязывающий белок сконструирован таким образом, чтобы активировать Т-клетку при связывании с Т-клеточным рецептором, например, при связывании с CD3, сопрягая, таким образом, активацию Т-клетки с конкретной выбранной опухолевой клеткой.

В контексте биспецифических антигенсвязывающих белков, где один фрагмент связывается с CD3, а другой фрагмент связывается с целевым антигеном, целевой антиген может представлять собой опухоль-ассоциированный антиген. Неограничивающие примеры специфических опухоль-ассоциированных антигенов включают, например, AFP, ALK, белки BAGE, BIRC5 (сурвивин), BIRC7, β -катенин, bcr-ab1, BRCA1, BCMA, BORIS, CA9, карбоангидразу IX, каспазу-8, CALR, CCR5, CD19, CD20(MS4A1), CD22, CD30, CD40, CDK4, CEA, CLEC-12, CTLA4, циклин-B1, CYP1B1, EGFR, EGFRvIII, ErbB2/Her2, ErbB3, ErbB4, ETV6-AML, EpCAM, EphA2, Fra-1, FOLR1, белки GAGE (например, GAGE-1, -2), GD2, GD3, GloboH, глипикан-3, GM3, gp100, Her2, HLA/B-raf, HLA/k-ras, HLA/MAGE-A3, hTERT, LMP2, белки MAGE (например, MAGE-1, -2, -3, -4, -6 и -12), MART-1, мезотелин, ML-IAP, Muc1, Muc2, Muc3, Muc4, Muc5, Muc16 (CA-125), MUM1, NA17, NY-BR1, NY-BR62, NY-BR85, NY-ESO1, OX40, p15, p53, PAP, PAX3, PAX5, PCTA-1, PLAC1, PRLR, PRAME, PSMA (FOLH1), белки RAGE, Ras, RGS5, Rho, SART-1, SART-3, Steap-1, Steap-2, TAG-72, TGF- β , TMPRSS2, антиген Thompson-nouvelle (Tn), TRP-1, TRP-2, тирозиназу и уроплакин-3.

В некоторых вариантах осуществления биспецифический антигенсвязывающий белок выбран из группы, состоящей из биспецифического антитела к CD3 и CD20 (как описано в публ. заяв. на патент США №№ US 2014/0088295 A1 и US 20150266966A1, включенных в данный документ посредством ссылки), биспецифического антитела к CD3 и муцину-16 (например, биспецифического антитела к CD3 и Muc16) и биспецифического антитела к CD3 и простатспецифическому мембранному антигену (например, биспецифического антитела к CD3 и PSMA). В других вариантах осуществления биспецифический антигенсвязывающий белок содержит один фрагмент, который связывает CD3. Иллюстративные фрагменты антитела к CD3 описаны в публ. заяв. на патент США №№ US 2014/0088295 A1 и US 20150266966 A1 и в международной публикации № WO 2017/053856, опубликованной 30 марта 2017 г., все из которых включены в данный документ посредством ссылки). В еще одних вариантах осуществления биспе-

цифический антигенсвязывающий белок содержит один фрагмент, который связывается с CD3, и один фрагмент, который связывается с BCMA, CD19, CD20, CD28, CLEC-12, Ней, белком HLA, белком MAGE, Muc16, PSMA или Steap-2.

В контексте биспецифических антигенсвязывающих белков, где один фрагмент связывается с рецептором Т-клетки, как, например, связывается с CD3, а другой фрагмент связывается с целевым антигеном, целевой антиген может представлять собой антиген, ассоциированный с инфекционным заболеванием. Неограничивающие примеры антигенов, ассоциированных с инфекционным заболеванием, включают, например, антиген, который экспрессируется на поверхности вирусной частицы или предпочтительно экспрессируется на клетке, инфицированной вирусом, где вирус выбран из группы, состоящей из HIV, вируса гепатита (А, В или С), вируса герпеса (например, HSV-1, HSV-2, CMV, HAV-6, VZV, вируса Эпштейна-Барра), аденовируса, вируса гриппа, флавивируса, эховируса, риновируса, вируса Коксаки, коронавируса, респираторно-синцитиального вируса, вируса эпидемического паротита, ротавируса, вируса кори, вируса краснухи, парвовируса, вируса корьей оспы, HTLV, вируса денге, папилломавируса, вируса контактиозного моллюска, вируса полиомиелита, вируса бешенства, вируса JC и вируса арбовирусного энцефалита. В качестве альтернативы целевой антиген может представлять собой антиген, который экспрессируется на поверхности бактерии или предпочтительно экспрессируется на клетке, инфицированной бактерией, где бактерия выбрана из группы, состоящей из хламидий, риккетсий, микобактерий, стафилококков, стрептококков, пневмококков, менингококков, гонококков, клебсиелл, протеев, серратий, псевдомонад, легионелл, дифтерийной палочки, сальмонелл, бацилл, бактерий, вызывающих холеру, столбняк, ботулизм, сибирскую язву, чуму, лептоспир и бактерий, вызывающих болезнь Лайма. В определенных вариантах осуществления целевой антиген представляет собой антиген, который экспрессируется на поверхности гриба или предпочтительно экспрессируется на клетке, инфицированной грибом, где гриб выбран из группы, состоящей из *Candida* (*albicans*, *krusei*, *glabrata*, *tropicalis* и т. д.), *Cryptococcus neoformans*, *Aspergillus* (*fumigatus*, *niger* и т. д.), представителей семейства Мукоровых (*mucor*, *absidia*, *rhizopus* и т. д.), *Sporothrix schenckii*, *Blastomyces dermatitidis*, *Paracoccidioides brasiliensis*, *Coccidioides immitis* и *Histoplasma capsulatum*. В определенных вариантах осуществления целевой антиген представляет собой антиген, экспрессируемый на поверхности паразита или предпочтительно экспрессируемый на клетке, инфицированной паразитом, где паразит выбран из группы, состоящей из *Entamoeba histolytica*, *Balantidium coli*, *Naegleria fowleri*, *Acanthamoeba* sp., *Giardia lamblia*, *Cryptosporidium* sp., *Pneumocystis carinii*, *Plasmodium vivax*, *Babesia microti*, *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania donovani*, *Toxoplasma gondii*, *Nippostrongylus brasiliensis*, *Taenia crassiceps* и *Brugia malayi*. Неограничивающие примеры специфических патоген-ассоциированных антигенов включают, например, gp120 ВИЧ, CD4 ВИЧ, гликопротеид L вируса гепатита В, гликопротеид М вируса гепатита В, гликопротеид S вируса гепатита В, E1 вируса гепатита С, E2 вируса гепатита С, гепатоцит-специфический белок, gB вируса простого герпеса, gB цитомегаловируса и белок оболочки HTLV.

Также можно получить биспецифические связывающие белки, которые содержат два связывающих фрагмента, каждый из которых направлен на определенного партнера по связыванию (т. е. каждый направлен на разную мишень) на поверхности одной и той же клетки. Данная структура особенно подходит для нацеливания на конкретные клетки или типы клеток, в случае которых обе мишени экспрессируются на поверхности одной и той же клетки. Однако мишени могут находиться по отдельности на разных клетках, при этом связывающие фрагменты для данных связывающих белков выбирают таким образом, чтобы каждый связывающий фрагмент связывал свою мишень с относительно низкой аффинностью (например, при K_D в низком микромолярном или высоком наномолярном диапазоне, например, при K_D , составляющей более ста наномоль, например, 500, 600, 700, 800 наномоль). Таким образом, продолжительному связыванию мишени способствуют только ситуации, когда две мишени находятся поблизости на одной и той же клетке.

Можно получить биспецифические связывающие белки, которые содержат два связывающих фрагмента, которые связывают одну и ту же мишень, при этом каждый связывает разный эпитоп на одной и той же мишени. Данная структура особенно подходит для повышения до максимума вероятности успешного блокирования мишени связывающим белком. Одна и та же биспецифическая связывающая молекула может нацеливаться на несколько внеклеточных петель, например, трансмембранного канала или рецептора клеточной поверхности.

Можно получить биспецифические связывающие белки, которые содержат два связывающих фрагмента, которые образуют кластер с отрицательными регуляторами передачи сигнала иммунными клетками и активируют их, что приводит к супрессии иммунного ответа. Репрессии в цис-положении можно достичь, если мишени находятся на одной и той же клетке; репрессии в транс-положении можно достичь, если мишени находятся на разных клетках. Репрессии в цис-положении, например, можно достичь с помощью биспецифического связывающего белка, содержащего связывающий фрагмент антитела IgGRIIb и связывающий фрагмент антитела к FelD1, в результате чего кластер с IgGRIIb образуется только в присутствии FelD1 с целью подавления иммунного ответа на FelD1. Репрессии в транс-положении, например, можно достичь с помощью биспецифического связывающего белка, содержащего связывающий фрагмент антитела к BTLA, и связывающийся фрагмент, который специфически связывает

тканеспецифический антиген, представляющий интерес, в результате чего образование кластера с ингибирующей молекулой BTLA происходит только в выбранной целевой ткани, что потенциально обеспечивает лечение аутоиммунных заболеваний.

Можно получить биспецифические связывающие белки, которые активируют многокомпонентные рецепторы. При такой структуре два связывающих фрагмента, направленных на два компонента рецептора, связываются, обеспечивают поперечную сшивку рецептора и активируют передачу сигнала от рецептора. Этого можно добиться, например, с применением биспецифического связывающего белка со связывающим фрагментом, который связывает IFNAR1, и связывающим фрагментом, который связывает IFNAR2, при этом связывание обеспечивает поперечную сшивку рецептора. Такой биспецифический связывающий белок может обеспечивать альтернативу лечению интерфероном.

Можно получить биспецифические связывающие белки, связывающие фрагменты которых перемещаются через полупроницаемый барьер, например, гематоэнцефалический барьер. При такой структуре один связывающий фрагмент связывает мишень, которая может проходить через конкретный селективный барьер; при этом другой связывающий фрагмент нацеливается на молекулу с терапевтической активностью, где целевая молекула с терапевтической активностью обычно не может пересечь барьер. Данный тип биспецифического связывающего белка пригоден для доставки терапевтических средств в ткани, которые терапевтическое средство иным образом не достигает. Некоторые примеры включают нацеливание на рецептор rIGR для перемещения терапевтического средства в кишечник или легкое или нацеливание на рецептор трансферрина для перемещения терапевтического средства через гематоэнцефалический барьер.

Можно получить биспецифические связывающие белки, связывающие фрагменты которых перемещаются в конкретные клетки или типы клеток. При такой структуре один связывающий фрагмент нацеливается на белок клеточной поверхности (например, рецептор), который легко интернализуется в клетку. Другой связывающий фрагмент нацеливается на внутриклеточный белок, при этом связывание внутриклеточного белка приводит к терапевтическому эффекту.

Биспецифические связывающие белки, которые связывают рецептор на поверхности фагоцитарной иммунной клетки и молекулу на поверхности инфекционного патогена (например, дрожжей или бактерии), сближают инфекционного патогена с фагоцитарной иммунной клеткой для обеспечения фагоцитоза патогена. Примером такой структуры может являться биспецифическое антитело, которое нацеливается на молекулу CD64 или CD89, а также патоген.

Биспецифические связывающие белки, которые содержат переменный участок антитела в качестве одного связывающего фрагмента и фрагмент, отличный от Ig, в качестве второго связывающего фрагмента. Переменный участок антитела обеспечивает нацеливание, тогда как фрагмент, отличный от Ig, представляет собой эффектор или токсин, связанный с Fc. Таким образом, лиганд (например, эффектор или токсин) доставляется к мишени, связанной с переменным участком антитела.

Биспецифические связывающие белки, которые содержат два фрагмента, каждый из которых связан с участком Ig (например, последовательностью Ig, содержащей участок CH2 и CH3) таким образом, что любые два фрагмента белка могут сближаться по отношению друг к другу в пределах Fc. Примеры данной структуры включают "ловушки", например, гомо- или гетеродимерные молекулы-ловушки.

Локусы, обеспечивающие повышенную экспрессию

Локусы, обеспечивающие повышенную экспрессию, которые подходят для применения в настоящем изобретении, включают, например, локус, который содержит нуклеотидную последовательность, характеризующуюся существенной гомологией с SEQ ID NO: 1, описанной в патенте США № 8389239 (также называемый в данном документе "локусом EESYR®"), локус, который содержит нуклеотидную последовательность, характеризующуюся существенной гомологией с SEQ ID NO: 2 или SEQ ID NO: 3, описанными в заявке на патент США с регистрационным № 14/919300 (также называемый в данном документе "локусом YARS"), а также другие локусы и последовательности, обеспечивающие повышенную экспрессию, которые документально подтверждены из уровня техники (например, US 20150167020 A1 и патент США № 6800457).

В некоторых вариантах осуществления локус, обеспечивающий повышенную экспрессию, который применяется в настоящем изобретении, выбран из локуса, который содержит нуклеотидную последовательность, характеризующуюся существенной гомологией с SEQ ID NO: 1, или локуса, который содержит нуклеотидную последовательность, характеризующуюся существенной гомологией с SEQ ID NO: 2 или SEQ ID NO: 3. Данные локусы содержат последовательности, которые не только обеспечивают повышенную экспрессию генов, интегрированных посредством функциональной связи с данными последовательностями (т. е. в пределах данных последовательностей или в непосредственной близости к данным последовательностям), но также характеризуются более высокой эффективностью рекомбинации и улучшенной интеграционной стабильностью по сравнению с другими последовательностями в геноме.

SEQ ID NO: 1, SEQ ID NO: 2 и SEQ ID NO: 3 были идентифицированы из клеток CHO. Было обнаружено, что другие виды млекопитающих (такие как, например, люди или мыши) характеризуются ограниченной гомологией в отношении идентифицированного участка, обеспечивающего повышенную экспрессию, однако гомологичные последовательности можно обнаружить в линиях клеток, полученных из

других типов тканей *Cricetulus griseus* или других гомологичных видов, и их можно выделить посредством методик, которые хорошо известны из уровня техники. Например, можно выявить другие гомологичные последовательности при помощи межвидовой гибридизации или методик на основе ПЦР. Кроме того, в нуклеотидную последовательность, изложенную под SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3, можно вносить изменения посредством методик сайт-направленного или случайного мутагенеза, которые хорошо известны из уровня техники. Варианты полученной последовательности затем можно тестировать в отношении активности повышения экспрессии. ДНК, которые по меньшей мере приблизительно на 90% идентичны по составу нуклеиновой кислоты SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3, характеризующимися активностью, обеспечивающей повышение экспрессии, можно выделять с помощью стандартных экспериментов, и, как предполагается, характеризуются активностью, обеспечивающей повышение экспрессии.

Сайт интеграции, сайт или положение нуклеотида для вставки одной или более экзогенных нуклеиновых кислот могут находиться в любом положении, которое находится в пределах любой из последовательностей, обеспечивающих повышенную экспрессию (таких как SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3), или рядом с ней. То, обеспечивает ли конкретное местоположение в хромосоме в пределах представляющего интерес локуса или рядом с ним стабильную интеграцию и эффективную транскрипцию интегрированного экзогенного гена, можно определить согласно стандартным процедурам, хорошо известным из уровня техники, например, описанным в патенте США 8389239 и заявке на патент США с регистрационным № 14919300.

Сайты интеграции, рассматриваемые в данном документе, расположены в пределах последовательностей, обеспечивающих повышенную экспрессию, или в непосредственной близости к данным последовательностям, 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 местоположению последовательности, обеспечивающей повышенную экспрессию, в хромосомной ДНК. В еще одних вариантах осуществления используемый сайт интеграции расположен на приблизительно 1000, 2500, 5000 или более пар оснований выше (5') или ниже (3') по отношению к местоположению последовательности, обеспечивающей повышенную экспрессию, в хромосомной ДНК.

В данной области техники подразумевается, что крупные участки генома, такие как участки прикрепления к ядерному скелету/матриксу, используются для эффективной репликации и транскрипции хромосомной ДНК. Участок прикрепления к ядерному скелету/матриксу (S/MAR), также называемый участком прикрепления к ядерному скелету (SAR), или матрикс-ассоциированным участком или участком прикрепления к матриксу (MAR), представляет собой участок ДНК генома эукариот, где прикрепляется ядерный матрикс. Без ограничения какой-либо теорией S/MAR обычно картируются в некодирующих участках, отделяют указанный участок транскрипции (например, домен хроматина) от соседних участков, а также предоставляют площадки для транскрипционной машины и/или связывания факторов, которые способствуют транскрипции, такие как сайты распознавания ДНКазы или полимеразы. Были описаны некоторые S/MAR, имеющие длину приблизительно 14-20 т.о. (Klar, et al., 2005, Gene 364:79-89). В связи с этим предполагается, что интеграция генов в локус, обеспечивающий повышенную экспрессию (например, в пределах SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3 или вблизи них), придает повышенную экспрессию. В некоторых вариантах осуществления клетки-хозяева, содержащие экзогенную последовательность нуклеиновой кислоты, кодирующую биспецифический антигенсвязывающий белок, интегрированную в конкретный сайт в пределах локуса, обеспечивающего повышенную экспрессию, характеризуются высокой удельной продуктивностью. В других вариантах осуществления клетка-хозяин, в которой кодируется биспецифический антигенсвязывающий белок, характеризуется удельной продуктивностью, составляющей по меньшей мере 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25 или 30 пикограмм/клетка/день (pcd).

В некоторых вариантах осуществления сайт интеграции находится в локусе, который содержит нуклеотидную последовательность под SEQ ID NO: 1. В конкретных вариантах осуществления сайт интеграции находится в пределах нуклеотидной последовательности под SEQ ID NO: 1 или в непосредственной близости к ней. В конкретных вариантах осуществления сайт интеграции находится в определенном положении в пределах SEQ ID NO: 1, выбранном из положений, охватывающих нуклеотиды под номером 10-13515; 20-12020; 1020-11020; 2020-10020; 3020-9020; 4020-8020; 5020-7020; 6020-6920; 6120-6820; 6220-6720; 6320-6620; 6420-6520; 6460-6500; 6470-6490 и 6475-6485. В других вариантах осуществления сайт интеграции находится в последовательности, которая выбрана из группы, состоящей из нуклеотидов 5000-7400, 5000-6500, 6400-7400 из SEQ ID NO: 1 и нуклеотидов 6400-6500 из SEQ ID NO: 1. В конкретном варианте осуществления сайт интеграции находится перед, после или в пределах триплета "act", соответствующего нуклеотидам 6471-6473 из SEQ ID NO: 1.

В некоторых вариантах осуществления сайт интеграции находится в локусе, который содержит нуклеотидную последовательность под SEQ ID NO: 2 или SEQ ID NO: 3. В конкретных вариантах осуществления сайт интеграции находится в пределах нуклеотидной последовательности под SEQ ID NO: 2 или в непосредственной близости к ней. В конкретных вариантах осуществления сайт интеграции находится в пределах нуклеотидной последовательности под SEQ ID NO: 3 или в непосредственной близости

к ней. В некоторых вариантах осуществления сайт интеграции находится 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 из SEQ ID NO: 3. В конкретных вариантах осуществление интеграция происходит на месте или в пределах нуклеотидов 2001-2022 из SEQ ID NO: 3. В некоторых вариантах осуществления экзогенную нуклеиновую кислоту вставляют на место или в пределах нуклеотидов 2001-2002 или нуклеотидов 2021-2022 из SEQ ID NO: 3, при этом нуклеотиды 2002-2021 из SEQ ID NO: 3 удаляются в результате вставки.

Сайт-специфическая интеграция в локус, обеспечивающий повышенную экспрессию

Интеграции нескольких экзогенных нуклеиновых кислот в локус, обеспечивающий повышенную экспрессию, сайт-специфическим образом, т. е. в один конкретный сайт в пределах локуса, обеспечивающего повышенную экспрессию, раскрытого в данном документе, можно достичь несколькими способами, в том числе, например, посредством гомологичной рекомбинации и опосредованного рекомбиназой обмена кассетами, как описано в уровне техники (см., например, патент США № 8389239 и источники, раскрытые в нем).

В некоторых вариантах осуществления предусмотрены клетки, которые содержат по меньшей мере две, т. е. две или больше, разные последовательности распознавания для рекомбиназы (RRS) в пределах локуса, обеспечивающего повышенную экспрессию, которые подходят для интеграции последовательности нуклеиновой кислоты, содержащей несколько экзогенных нуклеиновых кислот или генов, представляющих интерес. Такие клетки можно получить путем введения экзогенной последовательности нуклеиновой кислоты, содержащей два или более RRS, в требуемый локус посредством различных способов, в том числе гомологичной рекомбинации, описанной в данном документе ниже и в уровне техники, например, патенте США № 8389239 и источниках, раскрытых в нем.

В конкретных вариантах осуществления предусмотрены клетки, которые содержат более двух разных последовательностей распознавания для рекомбиназы (RRS) в пределах локуса, обеспечивающего повышенную экспрессию, которые подходят для интеграции нескольких экзогенных нуклеиновых кислот. В конкретных вариантах осуществления предусмотрены клетки, которые содержат три разные последовательности распознавания для рекомбиназы (RRS) в пределах локуса, обеспечивающего повышенную экспрессию, которые могут опосредовать интеграцию двух отдельных экзогенных нуклеиновых кислот, например, где 5'-RRS и расположенный в середине RRS в геноме соответствуют 5'-RRS и 3'-RRS, фланкирующим первую экзогенную нуклеиновую кислоту, подлежащую интеграции, и расположенный в середине RRS и 3'-RRS в геноме соответствуют 5'-RRS и 3'-RRS, фланкирующим вторую экзогенную нуклеиновую кислоту, подлежащую интеграции.

Подходящие RRS могут быть выбраны из группы, включающей LoxP, Lox511, Lox5171, Lox2272, Lox2372, Loxm2, Lox-FAS, Lox71, Lox66 и их мутантные формы, при этом для обеспечения опосредованного рекомбиназой обмена кассетами (RMCE) применяют сайт-специфическую рекомбиназу, которая представляет собой рекомбиназу Cre или ее производное. В других примерах подходящий RRS может быть выбран из группы, включающей FRT, F3, F5, мутантную форму FRT с делецией 10 аминокислот, мутантную форму FRT с добавлением 10 аминокислот и их мутантные формы, и в данном случае для обеспечения RMCE применяют сайт-специфическую рекомбиназу, которая представляет собой рекомбиназу Flp или ее производное. В еще одном примере RRS могут быть выбраны из группы, включающей attB, attP и их мутантные формы, и в данном случае для обеспечения RMCE применяют сайт-специфическую рекомбиназу, которая представляет собой интегразу ϕ C31.

В других вариантах осуществления нативные клетки модифицируют посредством методики гомологичной рекомбинации с интеграцией последовательности нуклеиновой кислоты, содержащей несколько экзогенных нуклеиновых кислот, в конкретный сайт в пределах локуса, обеспечивающего повышенную экспрессию.

При гомологичной рекомбинации гомологичные молекулы полинуклеотидов (т. е. гомологичные плечи) выстраиваются и обмениваются частью их последовательностей. Трансген можно вводить во время такого обмена, когда трансген фланкирован гомологичными последовательностями генома. В одном примере сайт распознавания для рекомбиназы можно вводить в геном клетки-хозяина по сайтам интеграции посредством гомологичной рекомбинации. В других примерах последовательность нуклеиновой кислоты, содержащая несколько экзогенных нуклеиновых кислот, например, несколько нуклеиновых кислот, которые вместе кодируют биспецифический антигенсвязывающий белок, где нуклеотидная последовательность фланкирована последовательностями, гомологичными последовательностям в целевом локусе ("гомологичные плечи"), вставляется в геном хозяина.

Гомологичной рекомбинации в эукариотических клетках можно содействовать путем введения разрыва в хромосомную ДНК по сайту интеграции. Это можно осуществлять путем нацеливания определенных нуклеаз на конкретный сайт 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.

Конструирование вектора, нацеливающегося на ген, и выбор нуклеазы для обеспечения гомологичной рекомбинации находятся в пределах компетенции специалиста в данной области техники, к которой принадлежит настоящее изобретение. В некоторых примерах нуклеазы с цинковыми пальцами (ZFN), которые характеризуются модулярной структурой и содержат отдельные домены с цинковыми пальцами, распознают конкретную 3-нуклеотидную последовательность в последовательности-мишени (например, в сайте направленной интеграции). В некоторых вариантах осуществления можно использовать ZFN в комбинации с отдельными доменами с цинковыми пальцами, нацеленными на множество последовательностей-мишеней. Нуклеазы на основе подобного активатору транскрипции (TAL) эффектора (TALEN) также можно использовать для сайт-специфического редактирования 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 на основе механизма бактериального адаптивного иммунитета. 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 расщепления ДНК-мишени, так и местоположение сайтов расщепления меняются в зависимости от положения мотива, смежного с протоспейсером (PAM), дополнительно необходимого для распознавания мишени (Chen, H. et al., Biol. Chem., опубликованная онлайн 14 марта 2014 г. в виде оригинала M1 13.539726). Последовательности, уникальные для конкретного нацеливаемого локуса (такие как SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3), можно идентифицировать посредством выравнивания многих из данных последовательностей по отношению к геному CHO, за счет чего можно выявить потенциальные нецелевые сайты с совпадением в 16-17 пар оснований.

В некоторых вариантах осуществления нацеливающий вектор, несущий нуклеиновую кислоту, представляющую интерес (например, нуклеиновую кислоту, содержащую один или более RRS, необязательно фланкирующих один или более генов селективируемого маркера, или нуклеиновую кислоту, содержащую несколько экзогенных нуклеиновых кислот, которые вместе кодируют биспецифический антигенсвязывающий белок), фланкированную 5'- и 3'-концевыми гомологичными плечами, вводят в клетку с одним или более дополнительными векторами или мРНК. В одном варианте осуществления один или более дополнительных векторов или мРНК содержат нуклеотидную последовательность, кодирующую сайт-специфическую нуклеазу, в том числе без ограничений нуклеазу с цинковыми пальцами (ZFN), ZFN в виде димера, нуклеазу на основе подобного активатору транскрипции эффектора (TALEN), слитый белок на основе домена TAL-эффектора или РНК-управляемую ДНК-эндонуклеазу. В определенных вариантах осуществления один или более векторов или мРНК предусматривают первый вектор, содержащий направляющую 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 распознавания, или генную кассету, содержащую любой из этих экзогенных элементов, предусмотренных для нацеленной вставки. Если применяют мРНК, то мРНК можно трансфицировать в клетку посредством общепринятых способов трансфекции, известных специалисту 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 от вспомогательного полинуклеотида, если требуется лишь кратковременная экспрессия для достижения требуемой интеграции нуклеиновой кислоты.

Векторы для сайт-специфической интеграции

В данном документе предусмотрены векторы на основе нуклеиновой кислоты для введения экзогенных нуклеиновых кислот в локус, обеспечивающий повышенную экспрессию, посредством сайт-специфической интеграции. Подходящие векторы включают векторы, сконструированные так, чтобы содержать экзогенную нуклеиновую последовательность, фланкированную RRS, для интеграции посредством RMCE, и векторы, сконструированные так, чтобы содержать экзогенную нуклеиновую последовательность, представляющую интерес, фланкированную гомологичными плечами, для интеграции посредством гомологичной рекомбинации.

В различных вариантах осуществления предусмотрены векторы для обеспечения сайт-специфической интеграции посредством RMCE. В некоторых вариантах осуществления векторы сконструированы так, чтобы обеспечивать одновременную интеграцию нескольких нуклеиновых кислот в це-

левой локус. В отличие от последовательной интеграции, одновременная интеграция обеспечивает эффективность и быстрое выделение требуемых клонов, которые продуцируют биспецифический антигенсвязывающий белок.

В некоторых вариантах осуществления предусмотрен набор векторов, и он включает два или более векторов, при этом каждый вектор содержит по меньшей мере два RRS, фланкирующих одну или более нуклеиновых кислот, где нуклеиновые кислоты в векторах из набора вместе кодируют биспецифический антигенсвязывающий белок.

В одном варианте осуществления набор векторов включает первый вектор, содержащий в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF, и третий RRS; второй вектор, содержащий в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF, второй RRS; где либо первая, либо вторая нуклеиновая кислота дополнительно содержит нуклеотидную последовательность, кодирующую второй HCF; и где первый и второй HCF и первый LCF кодируют участки (например, переменные участки) биспецифического антигенсвязывающего белка. В некоторых вариантах осуществления нуклеотидная последовательность, кодирующая второй HCF, включена в первую нуклеиновую кислоту в первом векторе (т. е. первый LCF и второй HCF в одном векторе), необязательно расположенную, например, ниже нуклеотидной последовательности, кодирующей первый LCF; и в других вариантах осуществления нуклеотидная последовательность, кодирующая второй HCF, включена во вторую нуклеиновую кислоту во втором векторе (первый HCF и второй HCF в одном векторе).

Нуклеотидные последовательности, кодирующие HCF, могут кодировать аминокислоты, например, аминокислоты или домены(домены) из константного участка, или кодировать целый константный участок. В конкретных вариантах осуществления нуклеотидные последовательности, кодирующие HCF или LCF, могут кодировать один или более константных доменов, таких как CL, CH1, шарнир, CH2, CH3 или их комбинации. В определенных вариантах осуществления нуклеотидная последовательность, кодирующая HCF, может кодировать домен CH3. Например, нуклеотидная последовательность, кодирующая первый HCF, может кодировать первый домен CH3, а нуклеотидная последовательность, кодирующая второй HCF, может кодировать второй домен CH3. Первый и второй домены CH3 могут быть одинаковыми или отличаться по меньшей мере одной аминокислотой. Отличия в доменах CH3 или в константных участках могут принимать любой из форматов биспецифических антигенсвязывающих белков, описанных в данном документе, например, отличия, которые приводят к различным характеристикам связывания с белком А, электростатическому взаимодействию или к формату "выступ-и-впадина". Независимо от любых отличий в аминокислотной последовательности две нуклеотидные последовательности, кодирующие HCF, также могут отличаться тем, что одна из двух нуклеотидных последовательностей была модифицирована по кодам.

В некоторых вариантах осуществления каждая нуклеотидная последовательность, кодирующая HCF, независимо и функционально связана с последовательностью, регулирующей транскрипцию, предусматривающей, например, промотор. В некоторых вариантах осуществления промоторы, управляющие транскрипцией двух полипептидов, содержащих HCF, являются одинаковыми. В некоторых вариантах осуществления все промоторы, управляющие транскрипцией двух полипептидов, содержащих HCF, а также промотор, управляющий транскрипцией полипептида, содержащего LCF, являются одинаковыми (например, промотор CMV). В некоторых вариантах осуществления каждая нуклеотидная последовательность, кодирующая HCF или LCF, независимо и функционально связана с индуцируемым или репрессируемым промотором. Индуцируемые или репрессируемые промоторы обеспечивают возможность того, чтобы продуцирование происходило, например, только в фазе продуцирования (культура с подпиткой), а не в ходе фазы роста (культура, выращиваемая в системе посевных ферментеров). Индуцируемые или репрессируемые промоторы также обеспечивают возможность дифференциальной экспрессии одного или более генов, представляющих интерес. В некоторых вариантах осуществления каждая нуклеотидная последовательность, кодирующая HCF и/или LCF, независимо и функционально связана с промотором, расположенным выше по меньшей мере одного оператора TetR (TetO) или оператора Arc (ArcO). В еще одних вариантах осуществления каждая нуклеотидная последовательность, кодирующая HCF и/или LCF, независимо и функционально связана с гибридным промотором CMV/TetO или CMV/ArcO. Примеры гибридных промоторов (также называемых регуляторными слитыми белками) можно найти в международной публикации № WO 03101189 A1, опубликованной 11 декабря 2003 года (включенной в данный документ посредством ссылки).

В некоторых вариантах осуществления первая нуклеиновая кислота в первом векторе дополнительно содержит 5'-часть гена селективируемого маркера, расположенного в направлении 5' от третьего RRS в первом векторе; и вторая нуклеиновая кислота дополнительно содержит оставшуюся 3'-часть гена селективируемого маркера, расположенного в направлении 3' от третьего RRS во втором векторе. В данных вариантах осуществления первый, второй и третий RRS опосредуют сайт-специфическую интеграцию первой и второй нуклеиновых кислот, что приводит к соединению 5'-части и 3'-части гена селективируемого маркера должным образом и одновременно обеспечению интегрированных клонов для удобного отбора.

В определенных вариантах осуществления третий RRS в первом векторе сконструирован так, чтобы находиться в пределах 5'-части интрона гена селективируемого маркера; а третий RRS во втором векторе сконструирован так, чтобы находиться в пределах 3'-части интрона гена селективируемого маркера. В еще одних вариантах осуществления третий RRS в первом векторе сконструирован так, чтобы находится между промотором и геном селективируемого маркера, с которым он функционально связан (но отделен от него в другом векторе); при этом третий RRS в первом векторе сконструирован так, чтобы находится в направлении 3' от промотора; а третий RRS во втором векторе сконструирован так, чтобы находится в направлении 5' от гена селективируемого маркера.

Набор векторов, описанный выше, может включать более двух векторов. Например, кроме двух векторов, описанных выше, набор может включать третий вектор, содержащий по меньшей мере два RRS, фланкирующих нуклеотидную последовательность, кодирующую второй LCF. Набор также может включать вектор, кодирующий одну или более рекомбиназ, которые распознают RRS.

В других вариантах осуществления предусмотрены векторы для обеспечения сайт-специфической интеграции посредством гомологичной рекомбинации. В некоторых примерах полинуклеотидная последовательность, подлежащая интеграции в геном хозяина, может представлять собой ДНК-последовательность, такую как RRS или несколько RRS, фланкирующих один или более генов селективируемого маркера, для получения клеток, имеющих один или более RRS, интегрированных в требуемый локус для последующей интеграции нуклеиновых кислот, кодирующих биспецифический антигенсвязывающий белок. В других примерах полинуклеотидная последовательность, подлежащая интеграции в геном хозяина, содержит несколько нуклеиновых кислот, которые вместе кодируют биспецифический антигенсвязывающий белок. Например, полинуклеотидная последовательность содержит нуклеиновые кислоты, кодирующие две отличающиеся тяжелые цепи и общую легкую цепь биспецифического антитела. В некоторых вариантах осуществления каждая из нескольких нуклеиновых кислот, которые вместе кодируют биспецифический антигенсвязывающий белок, независимо (т. е. по отдельности) функционально связана с регуляторными последовательностями (такими как промотор, энхансер, последовательность терминации транскрипции или их комбинация) - то есть регуляторные последовательности (такие как промоторы) являются отдельными для каждой из нескольких нуклеиновых кислот, которые при этом могут быть одинаковыми или отличающимися (т. е. содержащими одинаковые или отличающиеся нуклеотидные последовательности). В случаях, если нуклеиновая кислота среди нескольких нуклеиновых кислот содержит несколько кодирующих последовательностей, то каждая кодирующая последовательность или каждая нуклеотидная последовательность, кодирующая N-концевую часть полипептида, независимо и функционально связана со своими собственными регуляторными последовательностями (такими как промотор).

Выбор последовательностей, гомологичных последовательностям в пределах локуса, обеспечивающего повышенную экспрессию, и включение выбранных последовательностей в качестве гомологичных плеч в нацеливающийся вектор находятся в пределах компетенции специалиста в данной области техники. В некоторых вариантах осуществления вектор или конструкция содержат первое гомологичное плечо и второе гомологичное плечо, где первое и второе гомологичные плечи совместно содержат целевую последовательность, которая замещает эндогенную последовательность в пределах локуса. В других вариантах осуществления первое и второе гомологичные плечи содержат целевую последовательность, которая интегрируется или вставляется в пределах эндогенной последовательности в локусе. В некоторых вариантах осуществления гомологичные плечи содержат нуклеотидную последовательность, гомологичную нуклеотидной последовательности, присутствующей в SEQ ID NO: 1, SEQ ID NO: 2 или SEQ ID NO: 3. В конкретных вариантах осуществления вектор содержит 5'-концевое гомологичное плечо, содержащее нуклеотидную последовательность, соответствующую нуклеотидам 1001-2001 из SEQ ID NO: 3, и 3'-концевое гомологичное плечо, содержащее нуклеотид, соответствующий нуклеотидам 2022-3022 из SEQ ID NO: 3. Гомологичные плечи, например, первое гомологичное плечо (также называемое 5'-концевым гомологичным плечом) и второе гомологичное плечо (также называемое 3'-концевым гомологичным плечом), являются гомологичными целевой последовательности в локусе. Гомологичные плечи в направлении от 5'-конца к 3'-концу могут охватывать участок или целевую последовательность в локусе, которые содержат 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 других вариантах осуществления общее число нуклеотидов в целевой последовательности, выбранной для первого и второго гомологичных плеч, составляет 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'-концевым гомологичным плечом и 3'-концевым гомологичным плечом (гомологичными целевой последовательности) составляет по меньшей мере 5 п.о., 10 п.о., 20 п.о., 30 п.о., 40 п.о., 50 п.о., 60 п.о., 70 п.о., 80 п.о., 90 п.о., 100 п.о., 200 п.о., 300 п.о., 400 п.о., 500 п.о., 600 п.о., 700 п.о., 800 п.о., 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 случа-

ях, где нуклеотиды 1001-2001 и 2022-3022 из SEQ ID NO: 3 выбраны в качестве 5'- и 3'-гомологичных плеч, расстояние между двумя гомологичными плечами может составлять 20 нуклеотидов (соответствующих нуклеотидам 2002-2021 из SEQ ID NO: 3); и такие гомологичные плечи могут опосредовать интеграцию экзогенной последовательности нуклеиновой кислоты в локус, содержащий SEQ ID NO: 3, например, в пределах нуклеотидов 1990-2021 или 2002-2021 из SEQ ID NO: 3, и одновременно делению нуклеотидов 2002-2021 из SEQ ID NO: 3.

Раскрытые в данном документе векторы для введения экзогенных нуклеиновых кислот с целью сайт-специфической интеграции в локус, обеспечивающий повышенную экспрессию, могут включать дополнительные гены и последовательности для управления экспрессией экзогенных нуклеиновых кислот, представляющих интерес и кодирующих полипептиды, и для обеспечения селекции и идентификации клеток, в которые экзогенные нуклеиновые кислоты, представляющие интерес, успешно интегрировались. Такие дополнительные последовательности включают, например, последовательности, регулирующие транскрипцию и трансляцию, гены селективируемого маркера и т. п., которые также описаны в данном документе ниже.

Регуляторные последовательности

Раскрытые в данном документе векторы для введения экзогенных нуклеиновых кислот в локус, обеспечивающий повышенную экспрессию, сайт-специфическим образом и клетки, полученные в результате сайт-специфической интеграции, могут включать регуляторные последовательности для управления экспрессией экзогенных нуклеиновых кислот, представляющих интерес, и кодируемых полипептидов. Регуляторные последовательности включают транскрипционные промоторы, энхансеры, последовательности, кодирующие подходящие сайты связывания рибосом на мРНК, и последовательности, которые контролируют терминацию транскрипции и трансляции. Из вирусных источников можно получить последовательности, контролирующие транскрипцию и трансляцию. Например, широко применяемые промоторы и энхансеры, полученные из вирусов, таких как вирус полиомы, аденовирус 2, вирус обезьяны 40 (SV40), цитомегаловирус (CMV) мыши или человека, представляют собой промотор ранних генов CMV (CMV-IE) или промотор основного IE CMV (CMV-MIE), а также RSV, промотор поздних генов SV40, SL3-3, MMTV, промотор гена убиквитина (Ubi), промотор гена убиквитина C (Ubc) и промоторы LTR ВИЧ. Промоторы, контролирующие и/или сигнальные последовательности из генома вирусов можно использовать для управления экспрессией при условии, что такие контролирующие последовательности являются совместимыми с выбранной клеткой-хозяином. В зависимости от типа клетки, в которой белки, представляющие интерес, должны экспрессироваться, можно также применять отличные от вирусных клеточные промоторы (например, промоторы гена β -глобина и EF-1 α). Последовательности ДНК, полученные из генома вируса SV40, например, точку начала репликации, промотор ранних и поздних генов, энхансер, сайты сплайсинга и полиаденилирования SV40 можно применять для обеспечения других генетических элементов, пригодных для экспрессии экзогенной последовательности ДНК. Промоторы ранних и поздних генов особенно пригодны в связи с тем, что их обоих легко получают из вируса SV40 в виде фрагмента, который также содержит точку начала репликации вируса SV40 (Fiers et al., *Nature* 273:113, 1978). Можно также применять более мелкие или более крупные фрагменты SV40. Обычно включают последовательность из примерно 250 п.о., которая простирается от сайта Hind III до сайта BglI, расположенного в точке начала репликации SV40. Можно применять индуцируемые (например, индуцируемые химическим соединением, кофактором, регуляторным белком), и они особенно пригодны для обеспечения возможности того, чтобы продуцирование антигенсвязывающих белков происходило только в фазе продуцирования (культура с подпиткой), а не в ходе фазы роста (культура, выращиваемая в системе посевных ферментеров). Примеры индуцируемых или репрессируемых промоторов включают промоторы гена алкогольдегидрогеназы I, тетрациклин-чувствительные промоторные системы, промоторы генов глюкокортикоидного рецептора, промотор гена эстрогенового рецептора, промоторы гена рецептора экизона, промоторы на основе гена металлотионеина и промоторы на основе гена T7-полимеразы. Последовательности, подходящие для экспрессии нескольких транскриптов посредством бицистронного вектора, были описаны ранее (Kim S. K. and Wold B. J., *Cell* 42:129, 1985), и их можно применять в настоящем изобретении. Примеры подходящих стратегий мультицистронной экспрессии белков включают применение пептида 2A (Szymczak et al., *Expert Opin Biol Ther* 5: 627-638 (2005)) и применение сайта внутренней посадки рибосомы ("IRES"), обе из которых хорошо известны из уровня техники. Другие типы векторов экспрессии будут также пригодны, например, описанные в патенте США № 4634665 (Axel et al.) и патенте США № 4656134 (Ringold et al.).

Селективируемые маркеры

Раскрытые в данном документе векторы для введения экзогенных нуклеиновых кислот в локус, обеспечивающий повышенную экспрессию, сайт-специфическим образом и клетки, полученные в результате сайт-специфической интеграции, могут включать один или более генов селективируемых маркеров.

В некоторых вариантах осуществления ген селективируемого маркера придает устойчивость к лекарственному средству, как, например, таковые, описанные в табл. 1 Kaufman, R. J. (1988) *Meth. Enzymology* 185:537, и предусматривают устойчивость к DHFR-MTX, P-гликопротеину и различным липофильным

цитотоксическим средствам с множественной устойчивостью к лекарственным средствам (например, адриамицину, колхицину, винкристину) и аденозиндеаминазу (ADA)-Xyl-A или аденозину и 2'-дезоксикоформицину. Другие доминантные селективируемые маркеры включают выделенные из микроорганизмов гены устойчивости к антибиотикам, например, устойчивости к неомизину, канамицину или гиромизину. Существует несколько подходящих систем для отбора клеток млекопитающих (Sambrook, выше, с. 16.9-16.15). Также были описаны протоколы котрансфекции, в которых используют два доминантных селективируемых маркера (Okayama and Berg, *Mol. Cell Biol* 5:1136, 1985).

В других вариантах осуществления ген селективируемого маркера кодирует полипептид, который обеспечивает или способен генерировать выявляемый сигнал для распознавания генных кассет, которые были успешно вставлены и/или заменены или не были успешно вставлены и/или заменены, в зависимости от конкретного случая. Подходящие примеры включают среди прочего флуоресцентный маркер или белок, фермент, который катализирует химическую реакцию, которая генерирует выявляемый сигнал. 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.

Системы для получения биспецифических антигенсвязывающих белков

В дополнительном аспекте настоящее изобретение предусматривает системы, которые включают комбинацию клетки и одного или более векторов, и которые можно применять для получения клеток, характеризующихся интегрированными в локус, обеспечивающий повышенную экспрессию, экзогенными нуклеиновыми кислотами, которые вместе кодируют биспецифический антигенсвязывающий белок. Системы могут быть представлены, например, в виде набора.

В некоторых вариантах осуществления система разработана так, чтобы обеспечивать эффективное конструирование вектора и одновременную интеграцию нескольких экзогенных нуклеиновых кислот посредством RMCE в конкретный сайт в локусе, обеспечивающем повышенную экспрессию. Одновременная интеграция обеспечивает быстрое выделение требуемых клонов, и применение одного локуса, обеспечивающего повышенную экспрессию, также является важным для создания стабильной линии клеток.

В некоторых вариантах осуществления предусмотрена система, которая включает любой из наборов векторов, описанных выше, сконструированных для интеграции нескольких экзогенных нуклеиновых кислот посредством RMCE, и клетку, содержащую RRS, которые интегрированы в конкретный сайт в локусе, обеспечивающем повышенную экспрессию, и которые соответствуют RRS в наборе векторов. Например, система включает клетку и набор векторов, где клетка содержит интегрированные в локус ее генома, обеспечивающий повышенную экспрессию, в направлении от 5' к 3' первый RRS, первую экзогенную нуклеиновую кислоту, второй RRS, вторую экзогенную нуклеиновую кислоту и третий RRS, где три RRS отличаются друг от друга; где набор векторов включает первый вектор, содержащий в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF (например, первый VL), и второй RRS; второй вектор, содержащий второй RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF (например, первый VH), и третий RRS; и где либо первая нуклеиновая кислота, либо вторая нуклеиновая кислота дополнительно содержит нуклеотидную последовательность, кодирующую второй HCF (например, второй VH). При введении векторов в клетку первая и вторая нуклеиновые кислоты из векторов интегрируются в локус, обеспечивающий повышенную экспрессию, с помощью рекомбинации, опосредованной первым, вторым и третьим RRS. Для облегчения скрининга трансфектантов, содержащих нуклеиновые кислоты, должным образом интегрированные из векторов в локус, первая экзогенная нуклеиновая кислота в клетке из системы может содержать первый ген селективируемого маркера, и вторая экзогенная нуклеиновая кислота в клетке может содержать второй ген селективируемого маркера, где первый и второй гены селективируемого маркера отличаются друг от друга, а также отличаются от любого гена селективируемого маркера, предусмотренного в векторах; и в конкретных вариантах осуществления первый и второй гены селективируемого маркера кодируют флуоресцентные белки (которые могут обеспечивать отрицательный отбор), и первая и вторая нуклеиновые кислоты в векторах предусматривают дополнительный ген селективируемого маркера в расщепленном формате для обеспечения положительного отбора. Отрицательный отбор в отдельности может обеспечивать быстрое выделение клонов, хотя эффективность выделения клонов с заданной рекомбинацией может быть ограничена (приблизительно 1%). Отрицательный отбор в сочетании с положительным отбором (новая флуоресценция или устойчивость к лекарственному средству или антибиотику) может значительно улучшить эффективность выделения положительных клонов (до приблизительно 80%).

Данные системы могут включать дополнительные компоненты, реактивы или информацию, напри-

мер, протоколы для введения вектора(векторов) из системы в клетку из системы посредством трансфекции. 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 полиэтиленимин. Отличные от химических 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 применением 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). 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. 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). 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 липид-нуклеиновая кислота (или липосом), которые облегчают захват нуклеиновой кислоты в клетки при нанесении на культивируемые клетки.

Способы получения биспецифических антигенсвязывающих белков

В настоящем изобретении также предусмотрены способы получения биспецифических антигенсвязывающих белков. Путем применения способов по настоящему изобретению биспецифический антигенсвязывающий белок (например, биспецифическое антитело) может продуцироваться с высокими титрами и/или высокой удельной продуктивностью (пг/клетка/день). В некоторых вариантах осуществления биспецифический антигенсвязывающий белок продуцируется с титром, составляющим по меньшей мере 5 мг/л, 10 мг/л, 15 мг/л, 20 мг/л, 25 мг/л, 30 мг/л, 35 мг/л, 40 мг/л, 45 мг/л, 50 мг/л или больше. В некоторых вариантах осуществления биспецифический антигенсвязывающий белок продуцируется при соотношении титра биспецифического антигенсвязывающего белка и общего титра антигенсвязывающих белков, составляющем по меньшей мере 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50% или больше. В некоторых вариантах осуществления биспецифический антигенсвязывающий белок продуцируется при удельной продуктивности, составляющей по меньшей мере 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 пикограмм/клетка/день или больше, которую определяют на основании общего количества антигенсвязывающих белков (в пг), продуцируемых клеткой в день.

В одном варианте осуществления способа применяют систему, раскрытую в данном документе, и в клетку из системы вводят векторы из системы посредством трансфекции. Трансфицированные клетки, где экзогенные нуклеиновые кислоты были надлежащим образом интегрированы посредством RMCE в целевой локус клетки, обеспечивающий повышенную экспрессию, можно подвергать скринингу и идентифицировать. Два полипептида, содержащие HCF, и по меньшей мере один полипептид, содержащий LCF, могут экспрессироваться интегрированными нуклеиновыми кислотами, и биспецифический антигенсвязывающий белок, содержащий все три полипептида, можно получить из идентифицированных трансфицированных клеток и очистить с применением известных способов.

В некоторых вариантах осуществления способ включает (i) обеспечение системы, которая включает клетку и набор векторов, где клетка содержит интегрированные в локус ее генома, обеспечивающий повышенную экспрессию, в направлении от 5' к 3' первый RRS, первую экзогенную нуклеиновую кислоту, второй RRS, вторую экзогенную нуклеиновую кислоту и третий RRS, где три RRS отличаются друг от друга; где набор векторов включает первый вектор, содержащий в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF (например, первый VL), и второй RRS; второй вектор, содержащий второй RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF (например, первый VH), и третий RRS; и где либо первая нуклеиновая кислота, либо вторая нуклеиновая кислота дополнительно содержит нуклеотидную последовательность, кодирующую второй HCF (например, второй VH); (ii) одновременное введение векторов в клетку и (iii) осуществление скрининга в отношении трансформированных клеток, в которых первая и вторая нуклеиновые кислоты из векторов одновременно интегрировались в локус, обеспечивающий повышенную экспрессию, с помощью рекомбинации, опосредованной первым, вторым и третьим RRS.

В конкретном варианте осуществления способа для облегчения скрининга трансформантов, содержащих нуклеиновые кислоты, должным образом интегрированные из векторов в локус, первая экзогенная нуклеиновая кислота в клетке из системы может содержать первый ген селективируемого маркера, и вторая экзогенная нуклеиновая кислота в клетке может содержать второй ген селективируемого маркера,

где первый и второй гены селективируемого маркера отличаются друг от друга; и первая и вторая нуклеиновые кислоты в векторах вместе кодируют дополнительный селективируемый маркер в расщепленном формате, где полная последовательность, кодирующая дополнительный ген селективируемого маркера, обеспечивается после одновременной интеграции. Скрининг в отношении трансформантов можно проводить для отбора по первому и второму селективируемым маркерам (отрицательный отбор) и по дополнительному селективируемому маркеру (положительный отбор).

В другом варианте осуществления в способе просто применяют клетку, содержащую экзогенную последовательность нуклеиновой кислоты, интегрированную в конкретный сайт в локусе клетки, обеспечивающую повышенную экспрессию, где экзогенная последовательность нуклеиновой кислоты кодирует биспецифический антигенсвязывающий белок и биспецифический антигенсвязывающий белок экспрессируется в клетке. Клонированная кассета экспрессии находится в непрерывном виде в конкретном сайте интеграции.

Настоящее описание дополнительно проиллюстрировано следующими примерами, которые не следует понимать как ограничивающие каким-либо образом. Содержание всех приведенных ссылок (включая ссылки на литературу, выданные патенты и опубликованные заявки на патенты, приводимые в настоящей заявке) явным образом включено в данный документ посредством ссылки.

Примеры

Пример 1. Клонирование плазмид, обеспечивающих экспрессию биспецифического антитела

Компоненты биспецифических антител в виде тяжелых и легких цепей можно клонировать из клеток гибридомы, В-клеток, плазматических клеток или библиотек генов рекомбинантных антител с применением способов, известных из уровня техники. Например, антитела можно клонировать из клеток гибридомы или В-клеток с помощью пяти штрих RACE-ПНР или ПЦР с применением праймеров, направленных к лидерным пептидам, последовательностям каркасного участка 1, последовательностям каркасного участка 4 или последовательностям константных участков. В качестве альтернативы гены антител или мРНК в клетках, экспрессирующих антитело, можно секвенировать с помощью секвенирования нового поколения и затем идентифицировать посредством способов биоинформатики. Также возможно секвенировать белки антител и клонировать соответствующие гены антител с помощью технологии синтеза ДНК. Библиотеки рекомбинантных антител, такие как дрожжевые и фаговые библиотеки, также являются источниками генов антител.

Каждая из экспрессирующих линий клеток CHO, RSX4189-1, RSX4187-1, RSX4191-1 и RSX4188-1, продуцирует биспецифическое антитело, состоящее из трех различных полипептидов: AbC1, AbC2 и AbC3 (фиг. 5). Для получения плазмид для конструирования RSX4189-1 плазмиду AbC1 подвергали линейаризации путем расщепления с помощью Mfe I, который расположен в направлении 3' от гена AbC1. Кассету экспрессии AbC2, вырезанную из плазмиды AbC2 с помощью расщепления Mfe I, лигировали с сайтом Mfe I линейаризированной плазмиды AbC1. Продукты лигирования трансформировали в *E. coli* DH10B. После трансформации и выращивания в среде LB, содержащей ампициллин, отдельные колонии *E. coli* анализировали на наличие требуемой плазмиды, содержащей гены AbC1 и AbC2. Последовательности тахи-ргер ДНК для плазмиды AbC3 и плазмиды с двойной экспрессией AbC1-AbC2 подтверждали с помощью секвенирования по Сэнгеру. Данные две плазмиды вместе с плазмидой экспрессии Cre pRG858 трансфицировали в клетки-хозяева EESYR®, содержащие сайты RRS1 и RRS3 в локусе EESYR®, с применением липофектамина. Трансфицированные клетки отбирали с помощью антибиотиков в течение 12 дней, и рекомбинантные клетки затем объединяли в качестве RSX4189-1.

Для получения плазмид для конструирования RSX4187-1 кассету экспрессии AbC3, фланкированную сайтами Mlu I и Nhe I, клонировали по сайтам Mlu I и Spe I в плазмиду AbC2 в направлении 3' от гена AbC2. Комбинированную плазмиду AbC2-AbC3, плазмиду AbC1 и плазмиду Cre pRG858 совместно трансфицировали в клетки-хозяева EESYR® с применением липофектамина. Клетки, которые подверглись RMCE, объединяли в качестве RSX4187-1.

Для получения плазмид для конструирования RSX4191-1 кассету экспрессии AbC3 клонировали по сайту Mfe I в плазмиду AbC1 в направлении 3' от гена AbC1. Комбинированную плазмиду AbC1-AbC3, плазмиду AbC2 и плазмиду Cre pRG858 совместно трансфицировали в клетки-хозяева EESYR® с применением липофектамина. Клетки, которые подверглись RMCE, объединяли в качестве RSX4191-1.

Для получения плазмид для конструирования RSX4188-1 кассету экспрессии AbC2, фланкированную сайтами Mlu I и Nhe I, клонировали по сайтам Mlu I и Spe I в плазмиду AbC3 в направлении 3' от гена AbC3. Комбинированную плазмиду AbC3-AbC2, плазмиду AbC1 и плазмиду Cre pRG858 совместно трансфицировали в клетки-хозяева EESYR® с применением липофектамина. Клетки, которые подверглись RMCE, объединяли в качестве RSX4188-1.

Пример 2. Экспрессия биспецифического антитела из локуса EESYR®

Линии клеток, экспрессирующие биспецифическое антитело, RSX4189-1, RSX4187-1, RSX4191-1 и RSX4188-1, культивировали в виде суспензии в бессывороточной среде. Для количественного определения их уровней экспрессии биспецифического антитела количества клеток в культурах подсчитывали на проточном питомере Guava и начинали культивирование свежих культур, содержащих 2 миллиона кле-

ток на мл среды, во встряхиваемых колбах. Через 4 дня после центрифугирования с удалением клеток собирали отработанную среду. Титры биспецифического антитела определяли с применением анализа методом HPLC на белке А, который являлся специфичным к биспецифическому антителу. Титры белка, представляющего собой биспецифическое антитело, которое экспрессировалось в RSX4189-1, RSX4187-1, RSX4191-1 и RSX4188-1, составляли 37,8 мг/л, 40,5 мг/л, 48,3 мг/л и 21,8 мг/л соответственно. Общий титр всех белков, представляющих собой антитела (в том числе белка, представляющего собой биспецифическое антитело, и белков, представляющих собой моноспецифические антитела), которые экспрессировались в данных линиях клеток, и соотношение титра белка, представляющего собой биспецифическое антитело, и общего титра белков, представляющих собой антитела, показаны в таблице ниже.

| AbPID/RSX | Титр биспецифических антител (мг/л) | Общий титр (мг/л) | % биспецифических антител |
|------------------|--|----------------------------------|--|
| RSX4187-1 | 40,5 | 211,7 | 19,10% |
| RSX4188-1 | 21,8 | 218,3 | 10,00% |
| RSX4189-1 | 37,8 | 199,5 | 19,00% |
| RSX419-1 | 48,3 | 172,9 | 27,90% |

Перечень последовательностей

SEQ ID NO: 1

13515 оснований ДНК

Cricetulus griseus

tctagaaca aaaccaaaaa tattaagtca ggcttggctt caggtgctgg ggtggagtgc 60

tgacaaaaat acacaaatc ctggctttct aaggctttt cggggattca ggtattgggt 120

gatgtagaa taaaaatctg aaacataggt gatgtatctg ccatactgca tgggtgtgta 180

tgtgtgtgta tgtgtgtctg tgtgtgtgcc cagacagaaa taccatgaag gaaaaaaca 240

cttcaaagac aggagagaag agtgacctgg gaaggactcc ccaatgagat gagaactgag 300

cacatgccag aggaggtgag gactgaacca ttcaacacaa gtggtgaata gtctgcaga 360

cacagagagg gccagaagca ctcagaactc cagggggctca ggagtggtc tctggaggct 420

tctgcccttg gaggttctg aggaggagge ttccatattg aaaatgtagt tagtggccgt 480

ttccattagt acagtgacta gagagagctg aggaccact ggactgagge ctagatgctc 540

agtcagatgg ccatgaaagc ctagacaagc acttccgggt ggaaggaaa cagcaggtgt 600

gaggggtcag gggcaagta gtgggagagg tctccagat gaagtagcag gaacggagac 660

gcactggatg gccccacttg tcaaccagca aaagcttga tcttgttcta agaggccagg 720

gacatgaca ggtgatctc ggttttaaa aggctttgt ttacctaatc acttctatta 780

gtcagatact ttgtaacaca aatgagtact tggcctgtat ttagaaaact tctgggatcc 840

tgaaaaaaca caatgacatt ctggctgcaa cacctggaga ctcccagcca ggcctggac 900

ccgggtccat teatgcaat actcagggac agattcttca ctaggactg atgagctgtc 960

ttgatgcaa atgtggcctc ttcatftac tacaagtcac catgagtcag gagtgetgt 1020

ttgcacagtg tgactaagtg atggagtgtt gactgcagcc attcccggcc ccagcttgtg 1080

agagagatcc ttttaattg aaagtaaget caaagtacc acgaagccac acatgtataa 1140

actgttgaa taatctgtc acatacaca accatgtgaa taatctgtgt acatgtataa 1200

actgttgaa taatctgtgt gcagccttc ctacctact acctccagt gatcaggttt 1260

ggactgctg tgtgtactg gacctgaat gteccaccg ctgtcccctg tctttacga 1320

ttctgacatt ttaataaat tcageggctt ccctctgct ctgtgcctag ctatacctg 1380

gtactctgca ttttggttc tgtgacatt ctctgtgact ctgctacatt ctacagatgac 1440
 atgtgacaca gaaggtgttc cctctggaga catgtgatgt ccctgtcatt agtggaatca 1500
 gatgccccca aactgttgc cagtgttgg gaaagtgaca cgtgaaggag gatcaggaaa 1560
 agaggggtgg aaatcaagat gtgtctgagt atctcatgtc cctgagtggc ccaggctgct 1620
 gacttactc cccaagtga gggaggccat ggtgagtaca cacacctcac acatactata 1680
 tccaacacac acacacacac acacacacac acgcacgcac gcacgcacgc acgcacacat 1740
 gcacacacac gaactacatt tcacaaacca catacgcata ttacacccca aacgtateac 1800
 ctatacata cacacataca caccctcca cacatcacac acataccaca cccacacaca 1860
 gcacacacat acataggcac acattcacac accacacata tacatttgg tatgcataca 1920
 tgcatacaca cacaggcaca cagacaccac acacatgcat tgtgtacgca cacatgcata 1980
 cacacacata ggcacacatt gagcacacac atacatttgt gtacgcacac tacatagaca 2040
 tatatgcatt tgtatatgca cacatgcatg cacacataca taggcacaca tagagcacac 2100
 acatacattt gtgtatgcac acatgcacac accaatcaca tgggaagact caggttcttc 2160
 actaagggtc acatgaactt agcagttcct ggttatctcg tgaaactgg aagattgctg 2220
 tggagaagag gaagcgttgg cttgagccct ggcagcaatt aaccccgccc agaagaagta 2280
 ggtttaaaaa tgagagggtc tcaatgtgga acccgcaggg cggcagttca gagaagagac 2340
 ctaccaagc caactgagag caaaggcaga gggatgaacc tgggatgtag tttgaacctc 2400
 tgtaccagct gggctcatg ctattttgtt atatctttat taaatattct ttagtttta 2460
 tgtgcgtgaa taccttgctt gcataaatgt atgggcactg tatgtgttct tggcgccgtt 2520
 ggaggccagg agagggcatg gatcctccgg agctggcggt tgagacagtt gtgaccaca 2580
 gtgtggggtc tgggaactgg gtcttagtgt tccgcaagtg cagctggggc tctaacctc 2640
 tgagccatcc ctccagctc aagaaactta tttcttagg acatggggga agggatccag 2700
 ggcttaggc ttgttgttc agcaaaact ctttctgtgt atttgaatt ttattttatt 2760
 ttactttttt gggatagaat cacattctgc agctcaggct gggcctgaac tcatcaaat 2820
 cctcctgtct cagtctacca ggtgataaga ttactgatgt gagcctggct ttgacaagca 2880
 ctttagagtc cccagccctt ctggacactt gttccaagta taatatatat atatatatat 2940

atatataat atatataat atatattgtg tgtgtgtgtt tgtgtgtgta tgagacactt 3000
 gctctaaggg tatcatatat atccttgatt tgcttttaatt ttattttta attaaaaatg 3060
 attagctaca tgcacctgt atgcgtctgt atcatctata tatecttct tctttctctc 3120
 tctttctctc ttctttctt caceccaag catctatttt caaatccttg tgccgaggag 3180
 atgccaagag tctcgttggg ggagatggg agggggcgat acaggggaag agcaggagga 3240
 aagggggaca gactgggtg ggtctttgga gagctcagga gaatagcagc gatcttccct 3300
 gtccctggg tcaacctta cagccaacac cattttgtgg cctggcagaa gagttgcaa 3360
 gctggtcgca ggtctgccac acaaccccaa tctggccca agaaaaggca cctgtgtgtg 3420
 actctggggg taaaggcgt gctgtgtct ctccagctgg acttgaaact cccgtttaat 3480
 aaagagtct gcaaaataat acccgagag tcacagtgc aggttcccgt gcttctctga 3540
 agcggcaggc acgggttccc taggaaatgg ggccttgctt gccaaagctcc cacggcttg 3600
 cctgcaaacg gcctgaatga tctggcactc tgcgttgcca ctgggatgaa atggaaaaaa 3660
 gaaaaagaag aagtgtctct ggaagcgggc gcgctcacac aaaccgcaa cgattgtgta 3720
 aacactctcc atgagaatc tggagtgcgg ttgccctta ctggggagct gaagacagct 3780
 agtgggggcg gggggaggac ctgtctagca tcttccacg gtgctcgtg gctgtgtgtc 3840
 atgccgggaa ccgaaacgc gaaactaaagt caagtctgc ttgggtgaa ctgacaatca 3900
 acgaaatcac ttgattgtt ttctctttt tactggaatt ctggatttg atagatgggg 3960
 gaggatcaga gggggagggg aggggcgggg agacggaggg aggaggggag gaggggagga 4020
 ggggaggagg ggaggagggg aagggatgga ggaaaataact aactttcta attcaacatg 4080
 acaaagattc ggagaaagt caccgctagt gaccggagg aggaatgcc tattggcat 4140
 tatattcct gtcgtctaat ggaatcaac tcttggctc agcaccaagg attctgagcc 4200
 tatctattc aagacagtaa ctacagccca cacggaagag gctatacaac tgaagaata 4260
 aaatttcaac ttatttcat ttctgtgact gcatgttca atgtagagag ccacctgtg 4320
 ctaggggctg atgtgctggg cagtagagtt ctgagcccgt taactggaac aaccagaac 4380
 tcccaccaca gtagagctt gctgagagag ggaggccctt ggtgagattt ctttgtgtat 4440
 ttatttagag acagggtctc atactgtagt ccaagctagc ctccagctca cagaaattct 4500

cctgttccgg ttccaaagt actggagtta tgagtgtg ttaattgaac gctaagaatt 4560
 tgctgattga agaaaacctc aagtgggttt ggctaataccc cacgaccca gaggctgagg 4620
 caggaggaat gagagaattc aaggtttgcc agagccacag ggtgagctca atgtggagac 4680
 tgtgaggggtg agctcaatgt ggagactgtg aggggtgagct caatgtggag actgtgaggg 4740
 tgagctcaat gtggagactg tgagggtgag ctcaatgtgg agactgtgag ggtgagctca 4800
 atgtggagac ctgtatcaag ataataatag tagtagtaac aatgcaggcg aggggtgtggt 4860
 tgagtggtag agcagttagt tgattgaca tgcttgaggt ctcccgtcc atctgtggcc 4920
 ctgcaacagg aaggaggga ggaagggggg gaacgagaga gaggaaagag agacagaagc 4980
 taagataggg aatgagagag gaaggaagaa acgggaagaa atcagactc cttctgagt 5040
 tccccaacg cctagtgaca tctgtgcac accctaaggt ggcctttgtg tggcactggc 5100
 ttgggtggtc gggaaaggca tttcagctt gttgcagaac tgccacagta gcatgctggg 5160
 tccgtgaaag ttctgccc ttaacaagaa gtcttacta cttgtgacct caccagtga 5220
 aatttctta attgtctct ggtgttctgg gttttgact tttgttcta aggatacatt 5280
 cctgggtgat gtcatgaagt ccccaaagac acagtggggc tgtgttgat tgggaaagat 5340
 gatttatctg ggggttcaaa aggaaaagaa gggaaacagg cacttgggaa aatgtctcc 5400
 cgcccaccg aattttggct tggcaaccgt ggtggaggag caagaaacac gtggacgttt 5460
 gaggaggcat ggggtcctag gaggacagga agcagaagga gagagctggg ctgacagcct 5520
 gcaggcattg cacagttca gaaggagatt acagcatgac tgagtttta gggatccaac 5580
 agggacctgg gtagagattc tgtgggtct gaggcaactt gacctagcc agatggtatt 5640
 tgaataacct gctcttagag ggaaaacaga catagcaaac agagccacgt ttagtgatga 5700
 aactctact ttgcctgagt catgtgcggc catgcccagg ggtcaggctg aactcaact 5760
 caaaaacaag tgagaaattg aagacaatcc gtggtggcag ctactggaag ggccaccaca 5820
 tcccagaaa gactggagct gctaaaaagc catttggat aggcacagtt atcttgaatg 5880
 catggagcag agattacgga aaaatcgaga atgtaatga ggcaacattc gagttgagtc 5940
 atcagtggt gaaaccag acgttccat ccctaaaag gaacatctt ctctcagta 6000
 aatggaaat aaaaattggg gcttgaattt ggcaaatgat tcagaactct gtgtaggtat 6060

ttccacacgc acagtggata atttcatgt tggagttat ttgtgctaaa aggcagaaaa 6120
 gggtaaaaag cacatcttaa gagttatgag gttctacgaa taaaaataat gttacttaca 6180
 gctattcett aattagtacc cccctccacc tgggtaatt tctgagata gtcagtgggg 6240
 aaaagatctc tcctctctt cttctcccc ctcctctct ctcctctct cctcctctc 6300
 ctcctctctc tcctcctc ccccttctt tcttctttg ctcctctctc tctgctctc 6360
 tctccttctc ttctcattt attctaagta gctttaaca gcacaccaat tacctgtgta 6420
 taacgggaaa acacaggctc aagcagctta gagaagattg atctgtgttc actagcgtgc 6480
 aattcagagg tgggtgaaga taaaaggcaa acatttgagg ccatttctt attggcacg 6540
 gcacttagga agtgaacat gcctaacta ctggtttgta ccaccttcc ctataatgga 6600
 ctgtttggga agctcctggg caaccgattc tggcatctca ttggtcagag gcctgttaaa 6660
 tggtagctt atttgcaaag aaggctgtaa cttgtagctt taaaagcctc tcctcaagaa 6720
 agaagggaga aaggatatgg ctagacatat ctaatagact taaccactgt gaaaagcctt 6780
 agtatgaate agatagaacc tatttttaac tcagttttga aaaaaataat ctttatattt 6840
 atttgtgtg gtgtgtgtg gtgtgtgtg gtgtgtgtg gtgtgtgtg gaaccacatg 6900
 tagcaggtgc tggaggaggc cagaagaggg caccagatct cctggaactg acaccacaca 6960
 tggttatgag ctgcctgatg tgggtgctgg gaactgaact ctcgtgttct gcaagagcag 7020
 caactgttct cftaactgat gagccatctc tcagccccc ccataattt taattgttca 7080
 ttttagtaaa ttttattcat aatcaattat cacagtataa aacaatgatt ttatatatat 7140
 catatacata tcaaggatga cagtgagggg gatatgtgtg tgtgtgtgtg tgtgtgtgtg 7200
 tgtgtgtgtg tgtgttattt gtgtgtgtgc ttttaagaa ggtgcatag tcaactgcatt 7260
 tctctgaagg atttcaaagg aatgagacat gtctgtctgc caggaaccct atcttctct 7320
 ttgggaatct gacccaaatg aggtattctg aggaactgaa tgaagagctc aagtagcagt 7380
 gtcttaaacc caaatgtgct gtctagagaa agtcaacgctc atcagtgagc tgaggagaga 7440
 ttactgagc ggaagacaag cgctctttga tftaagtggc tgaacagctc acggctgtgg 7500
 agtggagcct gtgctcaggt ctgaggcagt ctttgctagc cagctgtgat gagcagtgaa 7560
 gaaaggggtgg agatggaggc aggggtgggag cagggctatg gttcagacta ggtatcgtga 7620

gcacaccagc tggttgactt gtggtctgtg ggtcagcgt tgtaaaccgc ctcagggtca 7680
ggcagtcaca ttgcttgaag ctgaatgggt gaggcaacac agagagtgc aagaaggcaa 7740
agtaccacct ctccccgac ccaggtcact tctgggttat agctgagact ccggacagca 7800
tgcaaccagc tggtagagc tcagggaaa acttgatgtc tgcattgtgc tatgaaatgt 7860
gattcggtag atctggagaa aatttataat gctggctcag tcaagcactg aacaaaggta 7920
ccttggcttt gggagctaca tgacattgac tttaggcag acttttttt ttctgcccgc 7980
caattcccag ataaccaata tggaggctca atattaatta taaatgctcg gctgatagct 8040
caggettgtt actagctaac tctccaact taaatgaacc cattctatt atctacattc 8100
tgccacgtga cttaccttg tacttctgt ttctctct tctctgactc tgccctctg 8160
cttcccagag tcttagtct ggttctctg cctaacctta tctgcccag ctgctgacca 8220
agcattfata attaataa agtctcccag tgagactctc atccaggag gacttgggtg 8280
ctccccctc ctattgcca tctgtgtctt cctcttctt cgttcccc tctcttctt 8340
gctctctc ctccacct ctttcatag tattgatgc aagggtgtc tagaatggag 8400
gagtgccc atggcatgcaa agaaccagt taggatgctc tctgagggt tgtaacata 8460
agcagtgag acaattcaag ccacagagt aagacggaag gatgactgt gctctagagc 8520
aactctggg gcagaatcac agggtaggt tctgactga gggcgaagag gccacgagga 8580
agggagtgag ttgtctgag ctagaagcta cggcccact ctggttagca gacctgcca 8640
caagcatgct ttgtaatca tctgggatct gatttctc taaatctatg tcaactctt 8700
aagaaaatgt gaattctcac attaaattt agatatactt ctttgggtg ggggggtgta 8760
aaaaatctc aagaatatg atttctggg gccggagaga tggctcagag gtaagagaa 8820
ctggtgctc tttagacat tctgagtca attcccagca accacatgt ggctcacaac 8880
catctgtaat gcgacctgt gccatctct gacatgcatg gatacatgca ggcagaaagc 8940
tgtatacata gtaaattgat aaatctttt ttaaaaagag tatggattct gccgggtgtt 9000
ggtggcgcac gccttaac ccagcactct ggaggcagag gcaggtgat ctctgtgagt 9060
tcgagaccag cctggtctat aagagctagt tccaggacag cctccaaagc cacagagaaa 9120
ccctgtctcg aaaaacaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaga gtatggattc 9180

taagaaagcc gtaacagctg gagctgtgta cggagttcag cgtggtacta gaagaacaga 9240
 cattcatgat gaaacacccc aggattttta cttagtatct agtttccatt gttgttttga 9300
 gaccggctct tatgctctcc aggctggcct caaactgctg atcttcccgc ctctacctct 9360
 caagtctctgg gactacttgg ctcataaac agtttttgc gggctccctg aagttatggt 9420
 tgtacaaacc gtgggggtca atatactcac ttgggcagag agagaaggte tgaatcccag 9480
 acaatgactg catctcagga cagttgggaa gaggacaatg gcagaaggac ttagaaaaga 9540
 tagactggag ggtggaaaag cagcaggaac agagaacaa aacaggaagc ttgctatcca 9600
 gggccactct ggagtctgt ggcaagatgg aagcgggcta ggggaataca tttgtgctac 9660
 tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgat caatgcctat caatgttga 9720
 ggggaaatat gtataccaca ttgattctgg gagcaattct cagtatctgg cctagagaaa 9780
 ggaatggccc ctgcagaata gacagagtga atgggtccct ttatcattg ctaaagtga 9840
 ggagaaataa acatccttcc atagagttc aggtaaatga accccacagt tcatctgtgc 9900
 cgtggtggag gcctggccaa cagttaaaaa gattagacac ggacaaagtc tgaaggaaac 9960
 acctcgaata ggaagaggag agccacctca ttctgtaact ttctcaagg ggaagatgtt 10020
 ccaagagtgg gaataaatgg tcaaaggggg gatttttaat taggaaaacg atttctgta 10080
 tcacttgtga aactggaggt tgatttggg cataggacaa tagattttag gctttgcaa 10140
 aagctgtttc aaagcagaga aatggaatag agacaattat gtagcgagga gggagggtgg 10200
 ggcgaagatg gagacagaga agtggaaagct gactttaggg aagaggaaca tagaccacag 10260
 gggcggggcg gggggcaggg gcggggggcg gggctcaaag gaggcagtgg gaacgttget 10320
 agtgttcgca gcgtaagcgt gaatgtgcaa gcgtctttgt ggtgtgtgac caggagtagc 10380
 tgggctggct tgtgtgctgc ttgtaatccc agtctttgag gttccacac tgtccacag 10440
 tgggtgtgat tttcctcgg agagcatgag ggctctgctt tccccacatc ctccccagc 10500
 ttcgttgga tttgtttcca agatgttagt gggtagaca aagcctctct gttgatttgc 10560
 ctftaacagg tgacaaaaaa agctcaacca ggagacattt ttgccttctt ggaaggtaat 10620
 gctcccatgt agagcaatgg gacctctc taaggtgagg ctactcttgc agtttgcacc 10680
 cagctcttct gatgcaggaa ggaagttggt gggcaagcaa gactgtttgc ttcttgcgat 10740

ggacacattc tgcacacaaa ggctcaggag gggagaaggc tgtttgatgt ttagcactca 10800
 ggaaggcccc tgatgcatct gtgattagct gtctccatct gtggagcaga cacggactaa 10860
 ctaaaaacca gtgttttaaa attgtcaagc cttaagggtg aggaaattga ctattgtgc 10920
 tgggccatac gtagagcaag tgctctgcat tgggccaacc cccggctctg gtttctagcc 10980
 accagaatgg ctagaacta actcacaate ctcccattcc aggtctcagg tgetagaatg 11040
 aaccactata ccagcctgcc tgctgccta cctgccttcc taaatttaa atcatgggga 11100
 gtaggggaga atacacttat cttagttagg gtttctattg ctgtgaagag acaccatgag 11160
 catggcaact ctataaagg aaaacattta gtgggtggc agtttcagag gttttagtac 11220
 attgtcatca tggctgggaa catgatggca tgcagacaga catggtgctg gagaaaggga 11280
 tgagagtctt acatcttga ggcaacagga cctcagctga gacactggct ggtaccctga 11340
 gcataggaaa cctcacagcc caccctcaca gtgacatatt tcttcaaca aagccatacc 11400
 tctaatagt gccactcct atgagatgac agggccaatt acattcaaac tgctataaca 11460
 cttaaagta tttattttt attattgtaa attatgatg tagctgggtg gtggcagccg 11520
 aggtgcacgc cttaatccc agcacttggg aggagagggc agatggatct ctgtgagttc 11580
 aagaccagcc tggctataa gagctagtgt caaggaagga tatacaaaga acagttctag 11640
 gatagccttc aaagccacag agaagtgtg tcttgaaaac caaaaattgt gctgggacct 11700
 gtctctgctt tgggtgcttc cactcccc agagctggac tcttggtcaa cactgaatca 11760
 gctgcaaaat aaactcctgg attcctctct tgaacagga gcccgaagtc aggcgccac 11820
 ttgtctctc gcaggattgc catagacttt ttctgtgtgc ccaccattcc agactgaagt 11880
 agagatggca gtggcagaga ctgggaagggc tgcaacgaaa acaggaagtt attgcacct 11940
 gggaaatagtc tggaaatgaa gcttcaaac ttgcttcatg ttcagttgta cacagactca 12000
 ctcccaggtt gactcacacg tgtaaattt cctgactatg tetgactgc tttatctga 12060
 tgcttcttc caaaaatgcc aagtgtacaa ggtgagggaa tcacccttgg attcagagcc 12120
 cagggtctc ctcttaacc tggactgtc tttctcggc agcctctgac accctcccc 12180
 ccatttctc tatcagaagg tctgagcaga gttggggcac gctcatgtcc tgatactc 12240
 cttgtcttc tgaagatcta actctgacc cagaaagatg gctaaggtgg tgaagtgtt 12300

gacatgaaga cttggtctta agaactggag caggggaaaa aagtcggatg tggcagcatg 12360
taccgaaat ccagaactg gggaggtaga gacggatgag tgcccggggc tagctggctg 12420
ctcagccagc ctagctgaat tgccaaatc caactcctat tgaaaaacct ttaccaaca 12480
aacaacaaa caaataataa caacaacaac aacaacaac taccatac aagtgggcg 12540
gctcttggct cttgaggaat gactcacca aacccaaagc ttgccacagc tgttctctgg 12600
cctaaatggg gtgggggtgg ggcagagaca gagacagaga gagacatgac ttctgggct 12660
gggctgtgtg ctctaggcca ccaggaactt tctgtcttg ctctctgtct ggcacagcca 12720
gagcaccagc acccagcagg tgacacacacc tcctccgtg cttctgagc aaacacaggt 12780
gccttggct gtctattgaa cggagtaag ttctgcaga tgtatgcatg gaaacaacat 12840
tgtctgggt ttattctac tgtgtgata aaaaccgggg aactccagga agcagctgag 12900
gcagaggcaa atgcaaggaa tgctgcctcc tagctgctc cccatggctt gccgggctg 12960
ctttctgaa gccctctct cccattggc atgctgaca tgaacagcgt ttgaaatgct 13020
ctcaaatgct acttcaaag aaggtctc tgatctgct aactaaatca gaccatggtt 13080
caccgtgcat tatcttctg ctgtctgtct gtctgtctgt ctgtctatct gtctatc 13140
tatcaatcat ctatctatct atctctatt tatctaccta teattcaatc atctatctc 13200
taactagtta tcatttatt atttgttac ttacttttt tatttgagac agtattctc 13260
tgagtacag ccttggctgt cctggaacce attctgtaac caggctgtcc taaactcac 13320
agagatcaa ctgcctctgc ctctctggtg ctgggggtaa agacgtgcac caccaacgcc 13380
ccgctctatc atctattat gtaactatta ttcagtcatt atctatctc taactatcca 13440
tcactgtct atccatc 13500
atctataatc aattg 13515

SEQ ID NO: 2

14931 оснований

ДНК

Stricetulus griseus

другой признак (2176).. (2239)

n представляет собой а, с, g, t или нуклеотид отсутствует

catgtacact tatgcaagta tgatatggcc caacacagta ttttacacca attttatct 60
 ataaaatata catgtacatc aaaatatatt attaataata acatcattat tctttcttc 120
 caagtaataa acacatacac tgaaatttg gttctgtgg ataatttaa tgaacagga 180
 aatgcaaatt tatcttagca tgttacttc actttcttg catagataac cagtaatcac 240
 attgatggat catgtagtga aatgtatct taggtatcta aggaatttg gcttcgttt 300
 gtgcttggtg aactgaatt ctattcctaa caacagtgtg taaggattct gtctgattc 360
 tttaccagt atttgtccat ttgcatttc tttatttc atggctgctg ttctagaag 420
 tgggaagtag tgtgtcaagt ctgttaaca tgttccctg atgatcagtg tctaacacc 480
 tctctgagta catgttgcc aatgtcttt ctagaccat ctattctgc ttgacttate 540
 ctggtacatg cctgccaaga aattctcct catccttct gtctctcac tgatttact 600
 gatgtgtgga tttcacattg atcatatgga aatagaagat acaatttct ttattcacag 660
 tttggaagac tttcaatctc atagatcacc attatcttt gctactgttc cctatgctat 720
 ggtgaaattt ccatttgaat aattgctta acaattaaca agaaagaatc tattttact 780
 tgcaataact tccattcag aacatttact aactgttac tatatcaaa aactagttt 840
 atatatcatg tgagaaatga ctaattcata atttggccat gacattttt tcagaaacag 900
 aaaaagtgac caatacatac acaatgctat aaatattaag acttcagcaa attaaatatt 960
 tattcatgat atcacataaa attcattat tatgtttat taaatgtgt ttttaaaca 1020
 gtggtatcac taaatattaa gttagatgtg tttatgtgct taatgaatt atattttaga 1080
 atgttataag ttgtatatag tcaaatatgt aataaattt atttttagg tcttctcat 1140
 taaggattt taattttggg tccctttcc agagtactc tagctcatga tgagtgaca 1200
 taaaactaa acagtacaaa atgtacattg cattcagtat tgcacttgat cttgcactg 1260
 aagtttgagt cagttcatac atttagtact tgggaagtac attaagctaa ctttcattgc 1320
 tctggcaaaa tgctcgataa gataagagtc tattgtggaa agccatggca gcaggaaagt 1380
 aagactgctg atgatgtta atccatagtc aagacgcaga aggagatgaa tgctggtatc 1440
 caacatttt tgctgttcat tttctctaga accctagtc ataaagatgt atgactgca 1500

ttcaaaatgc gtccccctca gttgttcaac tttctgtaa atacccttc aggcattgtct 1560
 agaagattgt ttgcgaaata cttctcaate cattcaagtt gatagtgacg attaatcaact 1620
 gcagaataaa agcctgtaac ttggctcacg tgccaaggaa tatgcacact cctgacacat 1680
 caataagtaa atcaaagtgt agcttttgcc tftaacattg ccagacttat gtaatgttct 1740
 gcacgttctt cctccatcac ttttattct aatgggtgtt ccttgacatt gaatcacgct 1800
 gtggaagctg cttagaatta acattgaaat ctactgatat atttatgatg cagcaattta 1860
 gattfactat ttacttaga atttttata attgagagaa tataatattt tcacagttat 1920
 ctatctgctg taaatagagg atttataaaa aaatctctat aactttttt tacaacacac 1980
 agtaaaatta agttaaatt taataaagtc actatgttga tttcaaagtg tgctacgccc 2040
 acgggtgctca cgcaggtgta gcagaagatg ccactaagggt gggctaaggc cgatgggttg 2100
 gggctcgcgc tcctggaga tgagccccag gcggttccct ggcaatcagc tgcgatcatg 2160
 atgccccgatg agccannnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2220
 nnnnnnnnnn nnnnnnnnnc tgggtgactt tatggaaaga atttgataga tttcatgatg 2280
 tagaagaatt ttattaggct tattttacag gagactaaga ccttgggacc taaagatata 2340
 tgggtcctga gaatcaggaa atgggtagag acgtggttga tggatgaga cagattttag 2400
 agaactctta gatcatgggc aatgaccgca atctgatgct tagaatagat catctataaa 2460
 caattatgct gttcttttc tttctgtgt atgatctgat gatgtagecc ccttgccaag 2520
 ttccctgatc ccccttgcca agttccctga ttgtaacagt atataagcat tgcttgagag 2580
 catattcaac tacattgagt gtgtctgtct gtcatttctt cgccgattec tgatttctc 2640
 ttgagccttt tccttgctc tcctcgggc ggtggtctcc acgagaggcg gtccgtggca 2700
 aaagtgtata aatgttctaa aacatttgaa ctctaaaaca tgcaaatga aaaattaaaa 2760
 taaataaaca tgaaaattaa aatatattag ctgctaaaag ttaaacaata ctatataata 2820
 tttgttatt agaattcaaa atcacattag ttggatttaa tttgaacatt gcattcttc 2880
 aataataatt tcaataaaaa aagttcccc atgatagtag aaaataataa catatgtatc 2940
 tatctattta tftaactaca catatatagc atttgttca actaaaataa atgaatgagc 3000
 aaagcaccta agtaattgggt gtctattata ttatgaagc caatagttc aaataaatta 3060

tcatgcataa ggaggtattg caaatgttaa acctttttg aaacagatat tcccagttac 3120
 agaaattata atttctaate ttectataa gtagaatgat gataattaat ataggccatt 3180
 tgtaaataat gttcagatta aaatattctc tatttcacta gagaagaatg atattaaatg 3240
 tattatattt tatttccat tttgtttgca ccactattct atatecctca gcagtttaaa 3300
 tttgtttcac catatgtgtg tgtgtttgta tcttaaatat ggcactaaaa ttagaataat 3360
 ttaatataaa tcttaggag aaaagatatt gaattatttt atgttgatag gaaaatatct 3420
 ttaattgtc caagaatact tttcttcta ttttaggact gatcagacc aggactaata 3480
 ttttatatgt actaatteta tgtacaaaa tatgttatta tctcatgaat tctgtctcaa 3540
 tattgagga ataaaaatag tccatcatga actttaaata taaaataatg attaattaat 3600
 ttttattcat atttgtttg tatgaatggt tatacatcac atgtgtgcct ggtgactgtg 3660
 aatgtcagga gaaggtatga aagccactgg aattggaata agagataata tttgagatgt 3720
 tatgtgggtg ctgagaatta gacgcaagcc atctcaaga atagccagca tactatacca 3780
 ctgagtaate cattcatccc tcaataatta tctttgtaga cagtaaataat atttctaaac 3840
 tataaatgac cagaaaaatt aatgtattat taatgaagac attcatctca tgtgacacac 3900
 ttcacctgtc taaatcagta acactctctc cactaattaa gattttctaa gtgcatgaca 3960
 cttactattt ctaaagctgt ccaatggggg ccagtcccca gtcagcacc agtgagataa 4020
 tccatgaatg catttatate ttaggaaaaa ttcttateta tgtagtattt agaacatttt 4080
 catgtgaggg gataaacaag gaagcacaga tgctttctga tagaaacttt ctcttaatt 4140
 catctagaaa aaaaaaacct ctgagaaaa tctctcttgc tctctccca atgctctatt 4200
 cagcatctc tcctactta attctagatc tttttcteta tgcctcttg ctgctgcct 4260
 gctggctctg ctctatgct cccatgtea cttttcttg ctatctcacc gttacctct 4320
 ctgcctcact ctctgcctc ttctctgctt ctcatatggc caggctctgg acaattatag 4380
 ttatatgta cattctcata acacatgata tgcacatag tttctctcag gctagggata 4440
 tcacaatgac tggccaatga gcaagtggcc ttgcatgtag ctctaagttg gtgatggtc 4500
 ccagacagta agtagccatt tggttgaaat ttgaggttgg gtagtacatg aagactgaat 4560
 tttcttcaa ctctggcctt gaaatagtaa aacaacacct atgaaaatga cgacctgtat 4620

ttgtctttag aggcaaccac atattgtctg cagggcctgc ttgaatttg ctctgaagtt 4680
 agcttgtttg tgtaaaagga agaatcctat atcagcctga gaaatgtaa atacctagc 4740
 attcaagtc atcaaaatta tatggagagt ataaatcatc cttctgacta ttcatagtca 4800
 tatttgtgtc caccaagtat aaaacacact accaaagggc tgtggaaaa atcgccataa 4860
 ctgttcttat tagggaggca tagcagtggc acctgaggaa gttacagcaa caaccagtca 4920
 tccagtcaat aacccatgg cttgcccact tggaggtacc caataatgtt tggctttgcc 4980
 gagtaggact ccaacaaatt cagagggta attttaaat gctggttgc actgctgaac 5040
 agteccattg ccctctgcat aattccacaa tggaaagctt ttacactga ttgccaatca 5100
 ttaacagcc tactcagcat aacaggtat gatattatc tgcatttgt tacattacta 5160
 gatgaattcc tatttctcc tacaatagtg gaactgaaaa aagatacaca atcactactac 5220
 cccttacta atcttatgac ttatatcatt tcaatttca gaccataatg caaactattg 5280
 accaaaacat gtgaagatga aaaatagaaa tgtagaataa tattacatat aaaaagaaaa 5340
 ggcggactta tttgtttta tttcttagca tgcatagcaa tacatgattt gaggtttata 5400
 taataaaggg acaataaate ttcaagaaac ttaccctac tgaattaaa tattaagaa 5460
 ggtcacacat ttactcaat atattagact actgggcaaa tagacatgaa aagtagagtt 5520
 aatattgagg taggccttct gtgaaatgc taaggaaatt atgtttcata cagtgtgtaa 5580
 ccaagtggga atcatatcag aaagcagtc aaagcttata ttacaagtaa cagatgcttg 5640
 gttatatgac ctcccagagc ttgactgtct atacacaaaa agtgggtgta ataaaactgt 5700
 aattgggct atgtttttt aaatggctc accaacatga aaggaaggga atgagcatgt 5760
 catggatgct tagagattat gctccagca agaagaattg agctttggct cttattacag 5820
 aaacatgaca aggtgtgagt ttatttatt agaaattata taatattta agctggggac 5880
 taaaaattht attgaaacaa acaggcaagg gataggeatg tactagaagc aaaaatagga 5940
 tgtcaatgct gtaatgttat ttttggacc aaaatagtat ttctataga aatgacaatg 6000
 atcttaggtt attattctc ataaagatga caagtcaca agatccta gttcattaa 6060
 atcgttttag tcatttaata gagtgtgtg atagattaca caaggaag cacttacgat 6120
 gagaaataat gatatccaca attatttct taattcttag aaacatteta ttgttatatc 6180

tcaatctcag aagccactta ttgctttatt attgaaacat atgaaattgt aagttatata 6240
 ttgtctatgg tgacattca aagaacatgt gacgtacagt gtagcacaga taaagaacat 6300
 aactgcagct gaatcagtaa ctaaacttac atacattaaa tctgccatgt tggcaacagt 6360
 gtgtgcaacta ccaaaggatg tactaatgct cacgacactc ccctatgtca ccctttgttc 6420
 atcattacat cataggtcta tttgtttgc tttgaaatc tagaccaagt cttttgtgtc 6480
 ttccaagca cagagctcat taattacct catagacttg ttaaacttct tctggttcat 6540
 caattgaata gaaatactca ctactaatta tgtgagaccc tgccagtacc atagcacatg 6600
 gataatfff acataaaaca tgcatacaag taagattatt cagactgaac atgaatfita 6660
 gagaaatcag gaaggagat atgggagtg ttggagtgag actagagaaa tgtaattaa 6720
 ctataatctc aatacaaga tctactaagc aaaaaacatg aaacattgtc attcaagtga 6780
 aacatcagtc ttcaaattgg aaagatattt ttactaggaa aatgtctggt agatggttat 6840
 tatctagaaa acacaaaaat tagaaaacgg taaactfita taaaaagaat aatacaatga 6900
 gactacatga aaagttctta actaatgaaa caaatatctt gaaactfff tcttaaaagt 6960
 ttaatcaaa taaccatcat ggaaattcaa attaaaacta ttacatatt acccctgaaa 7020
 taataactaa taccaataa aaataatata aacaaaaat ggcaatgcat gccatcatgg 7080
 attgggaga gagaatgttc attgcagttc tgaatggata ctggtgccac cacggtgaaa 7140
 atctctgtat agtccttcc aaaagctgaa aatagacata tcacaagacc tgccacacat 7200
 tttcaagca aataccaaa ggactctacc tgactgcaga gacacttct cataaaatat 7260
 tattgttgat ctattcataa tatctggaaa atagaacag ccaagatgcc catcaactga 7320
 ttaatagatg ataaaattat tgtacatttc agtgaatat tattcagttt ttaagaaaa 7380
 tgaaattatg taataagcat gtaaatggat atatcttgaa acaaccattc ccattatat 7440
 tacctaaaca ttgaaagtcc aaaatcatat gatctfita gtggatctac taatctttg 7500
 ctatatgtat tttattgaac tacccatgga tgtgagataa ttgtaacaa cagcacatgg 7560
 gagagcatgg gatcattcaa ggaagattag agagaatgca tttttagga gaaatggag 7620
 gagcaataga aaggattaaa tgaggftact gatgaaagtg atggftagag aaggcaatat 7680
 gaggaggat aactagcact tagggcctft tgaaaaagac atagagaaaa tactattgta 7740

gaaacttct ataattggtg tatagttata tacaccaaag agctcagatg gagttaccct 7800
ataatggaaa tattaactac tttttatcac tgtgataaaa catcctgaac agagcaacat 7860
agattgggaa gcatttactt tggcttacag ttctaacggg ataaaaattc atgatgaaag 7920
aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga agcaggtgag agctcacatc 7980
ttgaagtgtg agaatgtagc agagagaaca aactgcaaat gaccagaaaa tgcttttggg 8040
tcagagccca taccctctg actgacttct ccagaaattc tgaacaaata aaactcccca 8100
aacagagcca taactgaagg tccagtgtct gagactacta ggggtattc ttattcaaac 8160
cactacaatg ggggtgggggg agcaatcctc caagtaggca ctacacacag acaataaaaa 8220
actctagtaa ctggaatgga ttgacttatt tgaattactt gccagtgag ctacatagag 8280
cacaattatt gtatttaaat tacccttat gatcttaca aactgacag taagtcata 8340
ttgctaaaga aaccacatat ttgaatcagg gaacatggtg atatctagtt gttcttcaac 8400
tggaacttc atgctttctg cccagcattc atgttgctgg aaagagcaat gtacactacc 8460
agtgtagaaa taaatcadc aatcttaca agatgtgat cctataagtt acaataaaaa 8520
ttagcctgat aagatatccc caccagaaga atattccat aaatgctatg ggagcaacaa 8580
gctattttct aaattagctt taatcctatt ctacaagaga gaatccatat ctagaatagt 8640
tatagggatc aagaacccat ggcttgattg gtcataggcc caatgggaga tctaatatt 8700
attgttctac aaaatgaaaa taactcctaa tgactgttg ctgcagtaat aagttagtat 8760
gttctcaac tctacaaga gaagtttgt ctacaataa atggcaatta aagcagcccc 8820
acaagattta tateatccg atctctcat ggctatgca tctagaagct aggaaacaaa 8880
gaggacccta agagagacat acatggtccc cctggagaag gggaaggggg caagacctc 8940
aaagctaatt gggagcatgg gggaggggag agggagtag aagaaagaga aggggataaa 9000
aggagggaga ggaggacaag agagagaagg aagatctagt caagagaaga tagaggagag 9060
caagaaaaga gataccatag tagagggagc cttgtatgtt taaatagaaa actggcacta 9120
gggaattgc caagatcca caaggtcca ctaataatct aagcaatagt cgagaggcta 9180
ccttaaagc ctttctctga taatgagatt gatgactacc ttatatacca tctagagcc 9240
tcatccagt agctgatgga agcagaagca gacatctaca gctaaacact gagctagttg 9300

cagacagga ggagtgatga gcaaagtcaa gaccaggctg gagaaacaca cagaaacagc 9360
 agacctgaaa aaaatgttc acatggacce cagactgata gctgggagtc cagcatagga 9420
 cttttctaga aacctgaat gaggatatca gtttgagggt ctggttaatc tatggggaca 9480
 ctggtagtgg atcaatattt atccctagtt catgactgga atttgggtac ccattccaca 9540
 tggaggaatt ctctgtcagc ctagacacat gggggagggt ctaggtcctg ctccaaataa 9600
 tgtgttagac tttgaagaac tcccttgaga agactcacce tccctgggga gcagaaaggg 9660
 gatgggatga gggttggtga gggacaggag aggaggggag ggtgagggaa ctgggattga 9720
 caagtaaatg atgcttgttt ctaatttaa tgaataaagg aaaagtaaaa gaagaaaaga 9780
 aaacaggcca aaagattata aaagacagag gtggtgggtg actataaaga aacactatta 9840
 tctaaataaa aatatgtcag aagcacacat gaacttatag tgtttatgaa agtatgtata 9900
 ataactacat aatctcaagc caagaaaaaa atatcatctt tcagtgatga aggtgatttt 9960
 atttctcca gaattaaagc caaagaccta atgaaagtaa ttatcttcaa aaggttgaaa 10020
 atacatactt tgcaatacac agatctgcct agaaatctca tgttcacaat acacatgatg 10080
 ctcaattgaa ttccattcaa tgttacagtt tagataaaca gttttagat aaactcaca 10140
 tgtatcattt cttttattt ttgaccaa cagcttctca tctgttattc agaataattc 10200
 ctcgatggca ggatatccat cccaattggg ggaaggggag aatttgaaga aaacctagac 10260
 cacatacata ttgccattg ggaacaaaag tctaaaatga tgttgttcac atcttctca 10320
 ctagtctct ccccgtecca aagaacctg gtatatgtgc ctcatcttac agagagagga 10380
 aagcaggaac tgagcatecc ttactgcca tectcaacce aaaatttgc tcatgtctca 10440
 gctctgccct tctcatatga cagttacaag tcaaggcttc caaagtcct ctgtcatgtt 10500
 tgggtgcaat agtttataca gatgacttca tgtctcata tctaatgtct tatatagatt 10560
 aatattaac aatgttattt cttaaccac attttaaatt aatttaaaaa tccattaatt 10620
 gtgtctataa aatgcagaca gaggctgag acacaatata agcctgatga tctgaatttg 10680
 aaactcacac ccaccacatg gagaatcaac ttcaaaaat tttctatta ctccacact 10740
 tacaccattg tacaacaca ataataatga acaaaatgaa atgaaataaa aaattaagtc 10800
 tctgtagta atgctactgt gcagcaaaag taaaaatggc agcttaagct tgctttatgg 10860

ttacacttta ccacttcca ttaattataa ggacttcaat catggcagaa ctatgctgtt 10920
 attgtctcag tgtaacctaa ccaggtgttc cagatgttct taatgtggac acctaaacta 10980
 tttgatattt gggtaagat ctttcctct ttcagaagaa acctcaggac agaggaatc 11040
 ttgtcttita atttgagtc tgtagacttt tccatttca aatatacatg aaacaagtga 11100
 tgaagaaaat taatcaaaag gtgggaattg caatgatatt aggttcaata ttaagcttca 11160
 atattatcat ggaatcgctt gttatacact gagtgtttgg caataaggga ttttagaag 11220
 aaggagtttt tattctcaac aggttcctta agtttagctc aaataaatct aagcaatcca 11280
 ctctagaatt aatagtttc ctaagggcac agctatgaat agagetcaat ttacatataa 11340
 aatttgttc accatttatg tcattccagt ttcattagt acaaggaaaa tacaaaatat 11400
 tttagatgca atatcaagt aatagttcat ctctttttt aatatatc acctaaatca 11460
 ccattttctc agaaaaatct ggctgaagt tctgtctgga acttcaacat gaaaaatatg 11520
 cacagcttgc tattataaat cctagttgat ttttaagatt catgtctggg gtctgactca 11580
 gaggggccag aggctagaca aatattttt gaatcttcat tgtgaagatt ttaatgatt 11640
 atttaatat aaataacaaa gatgatggat aatgtaactt tgtacagttc atagacgctg 11700
 aactactttg tgcttaaaat gttagttccc tatcataaat gataggtgat aagtgtatgt 11760
 ttaactttt cctctgagc tatattcatg tactagagaa ttattttaa catgaaaaga 11820
 ctgtgtttat agtctcagct cctgagaact ggtccaacct taggcaggtg aatgccagga 11880
 gcaacgtttt tctctacag aggatgtttt gctgccaage aacctggtg tgtggaaatg 11940
 ttctttttt aatcaagttt aaagggtctt catcatgctg ttgctccaca tatttcagg 12000
 tttagagcttg gtccttgag tattatcttt taccagaaaa tcatagtat tcttcaata 12060
 actaacaact aaacttttcg ataaaaaga attggaattt caattttaa gcctgagtaa 12120
 aattctgtg aatcaggata tttatttta agtcttatct tttaaaaagt tattttattt 12180
 tttaaaaaat tataatatac ttcataattt tctctcttc acttttttt acaaacactt 12240
 ctatagatca ccatgtgttt tttttttac atttatggcc tctttctgtt cattgttatt 12300
 acatacaaat agtcttgctt atagaagaac accacaattt gttacctgat acaaattat 12360
 caaccttaa aacctacaaa ctattgatat tactgaaaag actatactta tagatgtaaa 12420

gatatatgtg tgtgcacata tatagataca catatatgta ggatttttaa ttttagatt 12480
tagacatcaa aattattat atgactgaga aactagacac tataaatgag cattcagat 12540
tcaacaccgt gatttagat attgcacaa tgacagaaaa tttcttata gaaaattta 12600
agtttgtga ttgctctgtg cacttagtga agtctcacag aaaaagaatc atagtattt 12660
tagttataa taaaaagtac atataattaa aatggttggc aaaaacaac attgagcat 12720
tttctatt tactatcaag tagtatcatt ttgaaataat aattgacta gttcaaaaa 12780
tgaaaacaaa atttaacta aatgcctaact ctgacctgat aacatttta tgaatgaaat 12840
tattcaatag tgttataat taggggceca aaactttcc taaaataaaa ctttaattt 12900
tttccattt ttatttaaat tagaaacaaa attgtttac atgtaaatca gagttcctc 12960
acctccct tctcctgtc cctcactaac acctactg tccatacca tttctgctc 13020
ccaggagggg tgaggcctc catggggaaa ctcagagtc tgtctatcct ttcgtagag 13080
gcctaggecc taccattt gtctaggcta aggtcacia agtttactcc tatgctagt 13140
ataagtactg atctactaca agagacacca tagatttctt aggttctc actgacacc 13200
atgttcatgg ggtctggaac aatcatatgc tagtttcta ggtatcagtc tggggaccat 13260
gagctcccc ttgtcaggt caactgttc tgtgggttc accacctgg tctgactgc 13320
ttgctcacc actcctcct ttctgtaact gggttccagt acaattcctg gtttagctg 13380
gggtgtctac ttctacttc atcagcttct gggatggagc ctctaggata gcatacaatt 13440
agtcacatc tcaatcag ggaaggcat taaagtagc ctctcattg ttgcttggat 13500
tgtagttgg tgcatttt gtagatctt ggacattcc ctagtccag atatctttt 13560
aaacctaaa gactacctt attatggtat ctctttctt gctctctt attctccag 13620
aaaaatctt cctgctcct tatatttcc tctcctcc tctctccc ttctcattt 13680
cctagatcca tctccttc cccatgctc ccaagagaga tttgctcag gagatctgt 13740
tcttaacce tttcttggg gatctgtctc tcttagggtt gctctgtt cctagcttct 13800
ctggaagtgt ggattgtaag ctggtaatca tttgctccat gctctaaatc catatatgag 13860
tgatgttgt cttttgtga ctgggttacc tctcctaaa tggttcttc catatgctg 13920
tggattcaa tagcaciaac aacatacagt atcttggggc aactaacc aacaagtga 13980

aagaccagta tagcaagaac tttgagttta aagaaagaaa ttaaagaaga taccagaaaa 14040
tggaaagatc tcccatgctc tttgataggc agaatcaaca tagtaaaaat ggcaatcttg 14100
ccaaaatcca tctacagact caatgcaatc cccattaaat accagcacac ttcttcacag 14160
acctgaaaga ataatactta actttatag gagaaacaaa agaccagga taggccaac 14220
aacctgtac aatgaaggca ctccagagg catcccac cctgacttca agctctatta 14280
tagagtaata atcctgaaaa cagcttgga atggcacaaa aatagacagg tagaccaatg 14340
gaattgagtt gaaaaccctg atattaacc acatatctat gaacacctga ctttgacaaa 14400
gaagctaagg ttatacaatg taagaaagaa agcatcttca acaaatctg ctggcataac 14460
tggatgctgg catgtagaag actgcagata gatccatgctc taatgcatg cacaaaactt 14520
aagtccaat ggatcaaaaa cctcaacata aatccagcca cactgaacct catagaagag 14580
aaagtgggaa gtatccttga ataaattggt acaggagacc acatcttgaa cttaacacca 14640
gtagcacaga caatcagatc aataatcaat aatgggacc tctgaaact gagaagcttc 14700
tgtaaggcaa tggataagtc aacaggacaa aatggcagcc cacggaatgg gaaaagatat 14760
tcaccaatcc tatatctgac agagggtgctc tctctattg caagaacac aataagctag 14820
ttttaaaac accaattaat cggattataa agttgggtag agaactaat aaagaattgt 14880
taacagagca atctaactg gcagaaagac acataagaaa gtgctacca t 14931
SEQ ID NO: 3
4001 оснований
ДНК
Cricetulus griseus
ccaagatgcc catcaactga ttaatagatg ataaaattat tgtacattc agtgtaatat 60
tattcagttt ttaagaaaaa tgaaattatg taataagcat gtaaattgat atatcttgaa 120
acaaccattc cccattatat tacctaaaca ttgaaagtcc aaaatcatat gatctttta 180
gtggatctac taatctttg ctatatgtat ttattgaac tacccatgga tgtgagataa 240
ttgtaacaa cagcacatgg gagagcatgg gatcattcaa ggaagattag agagaatgca 300
tttttagga gataatggag gagcaataga aaggattaaa tgaggttact gatgaaagtg 360

atggttagag aaggcaatat gaggaggat aactagcact tagggccttt tgaaaaagac 420
 atagagaaaa tactattgta gaaacttct ataattggtg tatagtata tacaccaaag 480
 agctcagatg gagttaccct ataatggaaa tattaactac tttttatcac tgtgataaaa 540
 catcctgaac agagcaacat agattgggaa gcatttactt tggcttacag ttctaacggg 600
 ataaaaattc atgatgaaag aatgaatatg tcagcaaaca gcagtagcaa tggcctgaga 660
 agcaggtgag agctcacatc ttgaagtgtg agaattgtgc agagagaaca aactgcaaat 720
 gaccagaaaa tgcttttggg tcagagccca taccctctg actgacttct ccagaaattc 780
 tgaacaaata aaactcccca aacagagcca taactgaagg tccagtgtct gagactacta 840
 ggggtatttc ttattcaaac cactacaatg ggggtggggg agcaatcctc caagtaggca 900
 ctacacacag acaataaaaa actctagtaa ctggaatgga ttgacttatt tgaattactt 960
 gccagtggag ctacatagag cacaattatt gtatttaaat tacccttat gatcttaca 1020
 aacttgacag taagatcata ttgctaaaga aaccacatat ttgaatcagg gaacatggtg 1080
 atatctagtt gttctcaac tggaaacttc atgctttctg cccagcattc atgttgctgg 1140
 aaagagcaat gtacactacc agtgtagaaa ttaatcctc aatcttatca agatgtggat 1200
 cctataagtt acaataaaaa ttgcctgat aagatatccc caccagaaga atattcacat 1260
 aaatgctatg ggagcaaca gctattttct aaattagctt taatcctatt ctacaagaga 1320
 gaatccatat ctagaatagt tatagggatc aagaacccat ggcttgattg gtcataggcc 1380
 caatgggaga tectaattatt attgttctac aaaatgaaaa taactcctaa tgacttgtg 1440
 ctgcagtaat aagttagtat gttgctcaac tctacaaga gaagttttgt cttacaataa 1500
 atggcaatta aagcagcccc acaagattta taccataccg atctctcat ggcctatgca 1560
 tctagaagct aggaaacaaa gaggacccta agagagacat acatggctcc cctggagaag 1620
 gggaaagggg caagacctcc aaagctaatt gggagcatgg gggaggggag agggagttag 1680
 aagaaagaga aggggataaa aggagggaga ggaggacaag agagagaagg aagatctagt 1740
 caagagaaga tagaggagag caagaaaaga gataccatag tagagggagc cttgtatgtt 1800
 taaatagaaa actggcacta gggaattgtc caaagatcca caaggtcca ctaataatct 1860
 aagcaatagt cgagaggcta cctfaaaagc ctttctctga taatgagatt gatgactacc 1920

ttatatacca tctagagacc ttcattccagt agctgatgga agcagaagca gacatctaca 1980
 gctaaacact gagctagttg cagacaggga ggagtgatga gcaaagtcaa gaccaggctg 2040
 gagaaacaca cagaaacagc agacctgaaa aaaatgftgc acatggacce cagactgata 2100
 gctgggagtc cagcatagga cttttctaga aacctgaat gaggatatca gtttgaggt 2160
 ctggttaatc tatggggaca ctggtagtgg atcaatattt atccctagt catgactgga 2220
 attgggtac ccattccaca tggaggaatt ctctgtcagc ctagacacat gggggaggtt 2280
 ctaggctctg ctccaaataa tgtgttagac ttgaagaac tcccttgaga agactcacc 2340
 tcctgggga gcagaaagg gatgggatga gggttggtga gggacaggag aggaggggag 2400
 ggtgaggaa ctgggattga caagtaaatg atgcttgtt ctaatttaa tgaataaagg 2460
 aaaagtaaaa gaagaaaaga aaacaggcca aaagattata aaagacagag gtggtgggtg 2520
 actataaaga aacactatta tctaaataaa aacatgtcag aagcacacat gaacttatag 2580
 tgtttatgaa agtatgtata ataactacat aatctcaagc caagaaaaa atatcatctt 2640
 tcagtgatga aggtgattt atttctcca gaattaaagc caaagacctg atgaaagtaa 2700
 ttatctcaa aaggtgaaa atacatactt tgcaatacac agatctgcct agaaatctca 2760
 tgttcacaat acacatgatg ctcaattgaa ttccattcaa tgttacagtt tagataaaca 2820
 gttgtgat aaactcaca tgtatcattt cttttattt ttgaccaa cagcttctca 2880
 tctgttattc agaataatc ctgatggca ggatatccat cccaattggg ggaaggggag 2940
 aattgaaga aaacctagac cacatacata ttgcccattg ggaacaaaag tctaaatga 3000
 tgtgttcac atttctcta ctagtctct cccctgcca aagaacctg gtatatgtgc 3060
 ctcatctac agagagagga aagcaggaac tgagcatccc ttacttgcca tctcaacce 3120
 aaaattgca tcattgctca gctctgccct tctcatatga cagttacaag tcaaggcttc 3180
 caaagtcct ctgcatgtt tgggtgcaat agtttataca gatgacttca tgtctcata 3240
 tctaatgtct tatatagatt aatattaaac aatgttattt cttaaccac attttaaatt 3300
 aatttaaaa tccattaatt gtgtctataa aatgcagaca gactgctgag acacaatata 3360
 agcctgatga tctgaattg aaactcacac ccaccacatg gagaatcaac ttcaaaaat 3420
 ttcttatta ctccacact tacaccattg tacaacaca ataataatga acaaaatgaa 3480

atgaaataaa aaattaagtc tctgtaggta atgctactgt gcagcaaaag taaaaatggc 3540
agcttaagct tgctttatgg ttacacttta ccatctcca ttaattataa ggacttcaat 3600
catggcagaa ctatgctgtt attgtctcag tgtaacctaa ccagggtgtc cagatgttct 3660
taatgtggac acctaaacta tttgatattt gggtaagat cttccctct ttcagaagaa 3720
acctcaggac agaggggaatc ttgtcttta atttgagtc tgtagacttt ttccatttca 3780
aatatacatg aaacaagtga tgaagaaaat taatcaaaag gtgggaattg caatgatatt 3840
aggttcaata ttaagcttca atattatcat ggaatcgcct gttatacact gagtgtttgg 3900
caataaggga tttttagaag aaggagtttt tattctcaac aggttcctta agttagctc 3960
aaataaatct aagcaatcca ctctagaatt aaatagttc c 4001

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Клетка для экспрессии биспецифического антитела или его антигенсвязывающего фрагмента, содержащая экзогенную последовательность нуклеиновой кислоты, интегрированную в конкретный сайт в локусе, обеспечивающем повышенную экспрессию,

где локус, обеспечивающий повышенную экспрессию, содержит нуклеотидную последовательность, которая на по меньшей мере 90% идентична SEQ ID NO: 2 или SEQ ID NO: 3,

где экзогенная последовательность нуклеиновой кислоты кодирует указанное биспецифическое антитело или его антигенсвязывающий фрагмент и содержит первую экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый фрагмент легкой цепи (LCF), вторую экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый фрагмент тяжелой цепи (HCF), и третью экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую второй HCF, и

где первый HCF, второй HCF и первый LCF являются фрагментами указанного биспецифического антитела.

2. Клетка по п.1, где нуклеотидная последовательность, кодирующая первый HCF, кодирует первый домен CH2 и первый домен CH3, а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй домен CH2 и второй домен CH3.

3. Клетка по п.2, где первый и второй домены CH3 отличаются по меньшей мере одним аминокислотным положением.

4. Клетка по п.2, где первый и второй домены CH3 представляют собой домены CH3 IgG человека, и при этом один из двух доменов CH3 модифицирован в по меньшей мере одном аминокислотном положении, в результате чего характеристики связывания с белком А отличаются от таковых у немодифицированного домена CH3.

5. Клетка по п.2, где нуклеотидные последовательности, кодирующие первый и второй домены CH3, отличаются друг от друга тем, что одна из нуклеотидных последовательностей модифицирована по кодонам.

6. Клетка по п.1, где экзогенная последовательность нуклеиновой кислоты дополнительно содержит дополнительную экзогенную нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую второй LCF.

7. Клетка по п.1, где первая экзогенная нуклеиновая кислота дополнительно содержит первый промотор, функционально связанный с нуклеотидной последовательностью, кодирующей первый LCF, вторая экзогенная нуклеиновая кислота дополнительно содержит второй промотор, функционально связанный с нуклеотидной последовательностью, кодирующей первый HCF, и третья экзогенная нуклеиновая кислота содержит третий промотор, функционально связанный с нуклеотидной последовательностью, кодирующей второй HCF, где первый, второй и третий промоторы являются одинаковыми или отличающимися.

8. Клетка по п.1, где первая экзогенная нуклеиновая кислота расположена выше относительно второй и третьей экзогенных нуклеиновых кислот.

9. Клетка по п.8, дополнительно содержащая первый сайт распознавания для рекомбиназы (RRS),

расположенный в направлении 5' относительно первой экзогенной нуклеиновой кислоты, и второй сайт распознавания для рекомбиназы (RRS), расположенный в направлении 3' относительно как второй, так и третьей экзогенных нуклеиновых кислот, где первый и второй RRS отличаются.

10. Клетка по п.8, дополнительно содержащая третий RRS, расположенный в направлении 3' относительно первой экзогенной нуклеиновой кислоты и в направлении 5' относительно одной из второй и третьей экзогенных нуклеиновых кислот или их обеих, где третий RRS отличается от первого и второго RRS.

11. Клетка по п.9, дополнительно содержащая четвертую экзогенную нуклеиновую кислоту, содержащую ген селективируемого маркера.

12. Клетка по п.11, где четвертая экзогенная нуклеиновая кислота дополнительно содержит третий RRS и расположена в направлении 3' относительно первой экзогенной нуклеиновой кислоты.

13. Клетка по п.1, где вторая экзогенная нуклеиновая кислота расположена выше относительно первой и третьей экзогенных нуклеиновых кислот.

14. Клетка по п.12, где порядок расположения экзогенных нуклеиновых кислот в локусе является следующим: в направлении от 5' к 3' первая экзогенная нуклеиновая кислота, четвертая экзогенная нуклеиновая кислота, вторая экзогенная нуклеиновая кислота и третья экзогенная нуклеиновая кислота, где вторая экзогенная нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую модифицированный домен CH3 IgG человека, и третья экзогенная нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую нативный домен CH3 IgG человека.

15. Клетка по п.12, где порядок расположения экзогенных нуклеиновых кислот в локусе является следующим: в направлении от 5' к 3' первая экзогенная нуклеиновая кислота, вторая экзогенная нуклеиновая кислота, четвертая экзогенная нуклеиновая кислота и третья экзогенная нуклеиновая кислота, где вторая экзогенная нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую нативный домен CH3 IgG человека, и третья экзогенная нуклеиновая кислота содержит нуклеотидную последовательность, кодирующую модифицированный домен CH3 IgG человека.

16. Клетка по п.14 или 15, где первый, второй и третий промоторы являются одинаковыми и отличаются от промотора, с которым функционально связан ген селективируемого маркера.

17. Клетка по п.1, где биспецифический антигенсвязывающий белок специфически связывается с антигеном Т-клетки и антигеном опухолевой клетки.

18. Клетка по п.1, где локус, обеспечивающий повышенную экспрессию, содержит нуклеотидную последовательность SEQ ID NO: 2 или SEQ ID NO: 3.

19. Клетка по любому из предыдущих пунктов, где клетка представляет собой клетку CHO.

20. Система для получения биспецифического антитела или его антигенсвязывающего фрагмента, содержащая клетку и набор векторов,

где клетка содержит интегрированные в локус ее генома, обеспечивающий повышенную экспрессию, в направлении от 5' к 3' первый RRS, первую экзогенную нуклеиновую кислоту, третий RRS, вторую экзогенную нуклеиновую кислоту и второй RRS, где три RRS отличаются друг от друга;

где локус, обеспечивающий повышенную экспрессию, содержит нуклеотидную последовательность, по меньшей мере на 90% идентичную SEQ ID NO: 2 или SEQ ID NO: 3,

где набор векторов содержит первый вектор, содержащий в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый LCF, и третий RRS;

второй вектор, содержащий третий RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF, и второй RRS; и

где либо первая нуклеиновая кислота, либо вторая нуклеиновая кислота дополнительно содержит нуклеотидную последовательность, кодирующую второй HCF;

где первый HCF, второй HCF и первый LCF являются фрагментами указанного биспецифического антитела; и

где при введении векторов в клетку первая и вторая нуклеиновые кислоты из векторов интегрируются в локус, обеспечивающий повышенную экспрессию, с помощью рекомбинации, опосредованной первым, вторым и третьим RRS.

21. Система по п.20, где первая экзогенная нуклеиновая кислота содержит первый ген селективируемого маркера, а вторая экзогенная нуклеиновая кислота содержит второй ген селективируемого маркера, где первый и второй гены селективируемого маркера отличаются.

22. Система по п.20, где

первый вектор содержит в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, содержащую первый LCF, и третий RRS; и

второй вектор содержит в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, где вторая нуклеиновая кислота содержит как нуклеотидную последовательность, кодирующую первый HCF, так и нуклеотидную последовательность, кодирующую второй HCF, и второй RRS.

23. Система по п.20, где

первый вектор содержит в направлении от 5' к 3' первый RRS, первую нуклеиновую кислоту, со-

державшую нуклеотидную последовательность, кодирующую первый LCF, и нуклеотидную последовательность, кодирующую второй HCF, и третий RRS; и

второй вектор содержит в направлении от 5' к 3' третий RRS, вторую нуклеиновую кислоту, содержащую нуклеотидную последовательность, кодирующую первый HCF, и второй RRS.

24. Система по п.20, где нуклеотидная последовательность, кодирующая первый HCF, кодирует первый домен СНЗ, а нуклеотидная последовательность, кодирующая второй HCF, кодирует второй домен СНЗ.

25. Система по п.24, где один из первого и второго доменов СНЗ представляет собой домен СНЗ IgG человека, при этом другой домен представляет собой модифицированный домен СНЗ IgG человека, содержащий модификацию в по меньшей мере одном аминокислотном положении.

26. Система по п.25, где нуклеотидная последовательность, кодирующая модифицированный домен СНЗ, находится в первом векторе.

27. Система по п.25, где нуклеотидная последовательность, кодирующая модифицированный домен СНЗ, находится во втором векторе и находится выше нуклеотидной последовательности, кодирующей немодифицированный домен СНЗ.

28. Способ для отбора клетки, включающий:

(i) обеспечение системы по любому из пп.20-27;

(ii) одновременное введение векторов в клетку посредством трансфекции и

(iii) отбор трансфицированной клетки, где первая и вторая нуклеиновые кислоты из векторов интегрировались в локус клетки, обеспечивающий повышенную экспрессию, с помощью рекомбинации, опосредованной первым, вторым и третьим RRS.

29. Способ получения биспецифического антитела или его антигенсвязывающего фрагмента, включающий:

отбор трансфицированной клетки согласно способу по п.28,

обеспечение экспрессии первого LCF, первого HCF и второго HCF в отобранной трансфицированной клетке и

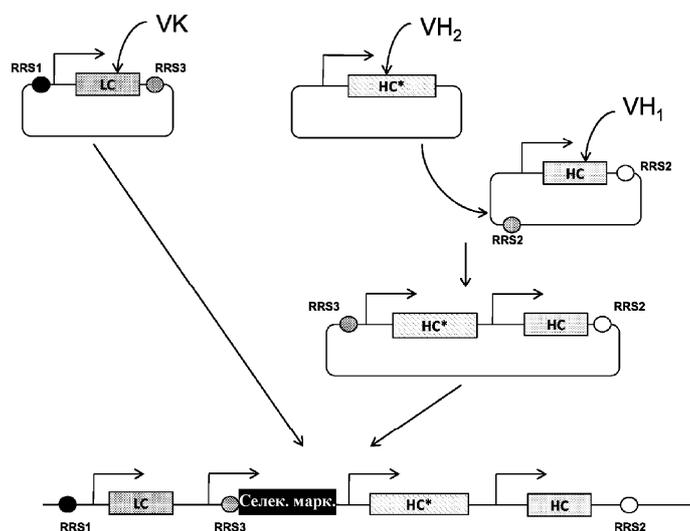
получение биспецифического антитела или его антигенсвязывающего фрагмента, содержащего первый LCF, первый HCF и второй HCF, из отобранной трансфицированной клетки.

30. Способ получения биспецифического антитела или его антигенсвязывающего фрагмента, включающий:

(i) обеспечение клетки по любому из пп.1-19;

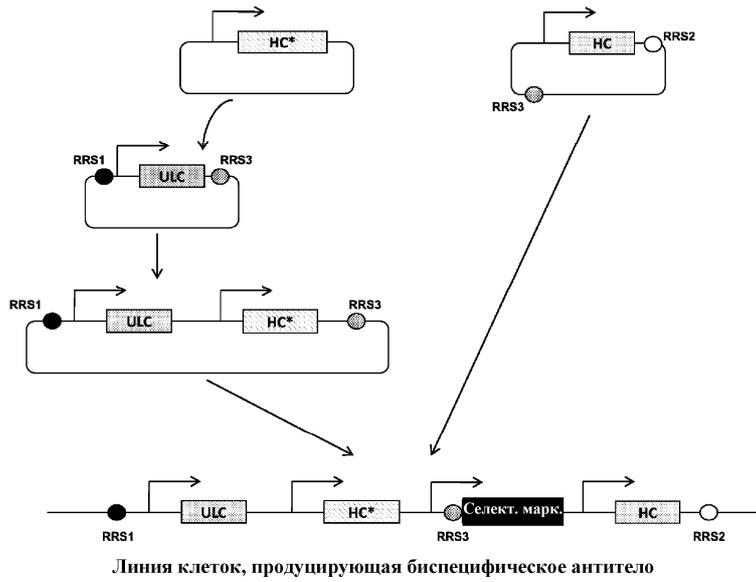
(ii) обеспечение экспрессии биспецифического антитела или его антигенсвязывающего фрагмента экзогенной последовательностью нуклеиновой кислоты и

(iii) получение биспецифического антитела или его антигенсвязывающего фрагмента из клетки.

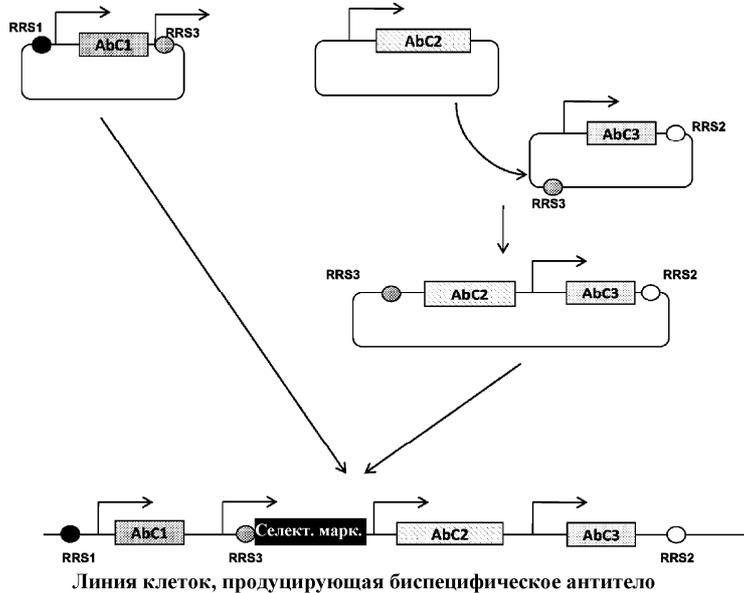


Линия клеток, продуцирующая биспецифическое антитело

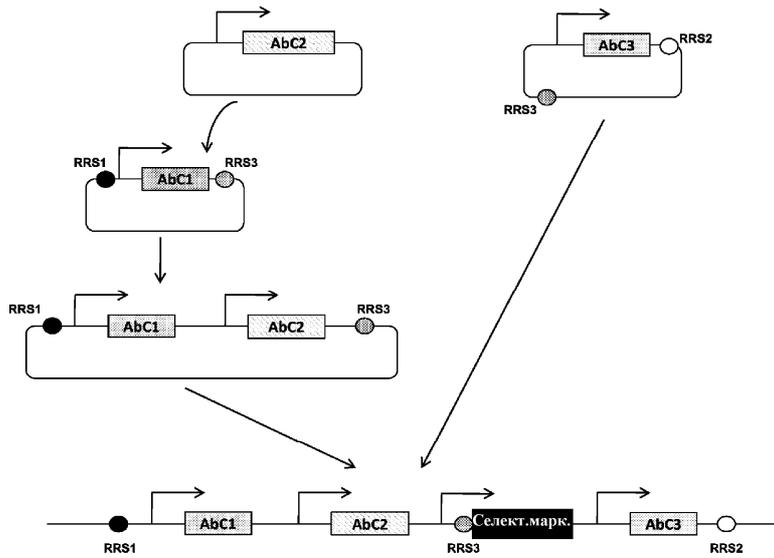
Фиг. 1



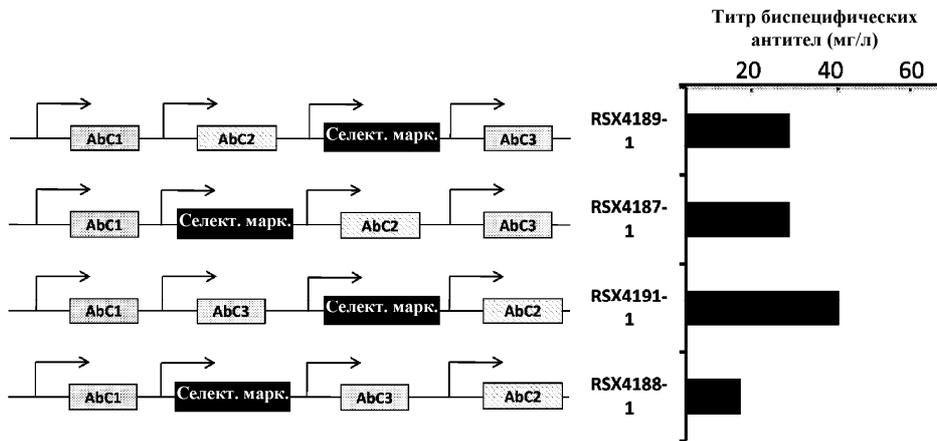
Фиг. 2



Фиг. 3



Линия клеток, продуцирующая биспецифическое антитело
Фиг. 4



Фиг. 5