



**ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ**

**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ**(21)(22) Заявка: **2008145487/10, 21.05.2007**(24) Дата начала отсчета срока действия патента:  
**21.05.2007**

Приоритет(ы):

(30) Конвенционный приоритет:

**23.05.2006 US 60/808,099****14.06.2006 US 60/813,856**(43) Дата публикации заявки: **27.06.2010** Бюл. № 18(45) Опубликовано: **27.01.2013** Бюл. № 3(56) Список документов, цитированных в отчете о  
поиске: **US 7045337 B, 16.05.2006. CHIN J. ET AL.,**  
**Proc. Natl. Acad. Sci. USA, 2002, v.99, no.17,**  
**pp.11020-11024. US 20030108885 A, 12.06.2003.**(85) Дата начала рассмотрения заявки РСТ на  
национальной фазе: **23.12.2008**(86) Заявка РСТ:  
**US 2007/012361 (21.05.2007)**(87) Публикация заявки РСТ:  
**WO 2007/139870 (06.12.2007)**

Адрес для переписки:

**190068, Санкт-Петербург, ул. Садовая, 51,**  
**офис 303, ООО "ПАТЕНТИКА", пат.пов.**  
**М.И.Ниловой, рег.№ 378**

(72) Автор(ы):

**ВАНГ Джиянгюн (US),****КСИ Джиянминг (US),****ШУЛЬЦ Питер Г. (US)**

(73) Патентообладатель(и):

**ЗЕ СКРИПС РЕСЕЧ ИНСТИТЮТ (US)****(54) ГЕНЕТИЧЕСКИ КОДИРУЕМЫЕ ФЛЮОРЕСЦЕНТНЫЕ КУМАРИНОВЫЕ  
АМИНОКИСЛОТЫ**

(57) Реферат:

Настоящее изобретение относится к области биохимии, в частности к способам трансляции белка, и может быть использовано в различных тест-системах, предназначенных для детектирования соединений полипептидной природы. Предложена ортогональная пара, включающая тРНК (О-тРНК) и аминоацил-тРНК синтетазу (О-РС), в которой О-тРНК является супрессорной тРНК, распознающей амбер-кодон, а О-РС представляет собой мутантную форму тирозил-тРНК синтетазы *Methanococcus janaschii*,

предпочтительно аминоацилирующую указанную тРНК неприродной аминокислотой L-(7-гидроксикумарин-4-ил)-этилглицином. Описана система трансляции, обеспечивающая введение в любое выбранное положение интересующего белка, синтезируемого в этой системе, неприродной аминокислоты L-(7-гидроксикумарин-4-ил)-этилглицина, которая включает ортогональную пару по изобретению, названную неприродную аминокислоту и нуклеиновую кислоту, которая кодирует представляющий интерес белок и содержит в выбранном положении селекторный кодон,

RU 2 4 7 3 6 9 0 C 2

RU 2 4 7 3 6 9 0 C 2



FEDERAL SERVICE  
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.

*C12N 15/54* (2006.01)*C12N 9/10* (2006.01)*C12N 15/74* (2006.01)*C12P 21/00* (2006.01)(12) **ABSTRACT OF INVENTION**(21)(22) Application: **2008145487/10, 21.05.2007**(24) Effective date for property rights:  
**21.05.2007**

Priority:

(30) Convention priority:

**23.05.2006 US 60/808,099****14.06.2006 US 60/813,856**(43) Application published: **27.06.2010 Bull. 18**(45) Date of publication: **27.01.2013 Bull. 3**(85) Commencement of national phase: **23.12.2008**

(86) PCT application:

**US 2007/012361 (21.05.2007)**

(87) PCT publication:

**WO 2007/139870 (06.12.2007)**

Mail address:

**190068, Sankt-Peterburg, ul. Sadovaja, 51, ofis  
303, OOO "PATENTIKA", pat.pov. M.I.Nilovoj,  
reg.№ 378**

(72) Inventor(s):

**VANG Dzhianjun (US),****KSI Dzhianming (US),****ShUL'Ts Piter G. (US)**

(73) Proprietor(s):

**ZE SKRIPS RESECh INSTIT'JuT (US)**(54) **GENETIC CODED FLUORESCENT COUMARINIC AMINO ACIDS**

(57) Abstract:

FIELD: medicine.

SUBSTANCE: what is presented is an orthogonal pair comprising tRNA (O-tRNA) and aminoacyl-tRNA synthetase (O-RS) wherein O-tRNA is suppressive tRNA recognizing an amber codon, while O-RS represents a mutant form of tyrosyl-tRNA synthetase *Methanococcus janaschii*, preferentially aminoacylating said tRNA by unnatural amino acid L-(7-hydroxycoumarin-4-yl)-ethyl glycine. What is described is a translation system providing the

introduction in any specified position of a protein of interest synthesised in this system, of unnatural amino acid L-(7-hydroxycoumarin-4-yl)-ethyl glycine which comprises the orthogonal pair under the invention, said unnatural amino acid and nucleic acid which codes the protein of interest and comprises a selector codon recognised by O-tRNA in the specified position.

EFFECT: method improvement.

20 cl, 9 dwg, 5 ex

Текст описания приведен в факсимильном виде.

## ПЕРЕКРЕСТНЫЕ ССЫЛКИ НА РОДСТВЕННЫЕ ЗАЯВКИ

[0001] Данная заявка испрашивает приоритет согласно:

Предварительной заявке на патент Соединенных Штатов Америки с  
порядковым № 60/808,099, поданной 23 мая 2006 г; и

Предварительной заявке на патент Соединенных Штатов Америки с  
порядковым № 60/813,856, поданной 14 июня 2006 г; раскрытие которых  
полностью включено в данную заявку посредством ссылки для любых целей.

## ЗАЯВЛЕНИЕ ОТНОСИТЕЛЬНО ПРАВ НА ИЗОБРЕТЕНИЕ, СДЕЛАННОЕ В РАМКАХ ФИНАНСИРУЕМЫХ ИЗ ФЕДЕРАЛЬНОГО БЮДЖЕТА НИОКР

[0002] Данное изобретение было сделано при государственной поддержке  
Министерства Энергетики, которое предоставило грант № ER46051. Государство  
может иметь определенные права на данное изобретение.

## ОБЛАСТЬ ТЕХНИКИ

[0003] Настоящее изобретение относится к области биохимии трансляции.  
Изобретение относится к соединениям и способам получения и применения  
ортогональных тРНК, ортогональных аминоксил-тРНК-синтетаз и их пар, которые  
встраивают не встречающиеся в природе аминокислоты в белки. Также  
изобретение относится к способам получения белков в клетках при помощи таких  
пар и белкам, полученным таким способом.

## УРОВЕНЬ ТЕХНИКИ

[0004] Флюоресценция стала одним из наиболее важных сигналов для  
детектирования в биотехнологии благодаря высокой чувствительности и  
безопасности в работе. Процессы типа резонансного переноса флюоресцентной  
энергии (FRET) или флюоресцентной поляризации делают возможным анализ  
событий построения, подвижности или конформационных изменений биомолекул

в режиме реального времени. Возможность селективно модифицировать белки с введением флюоресцентных зондов существенно облегчили исследования структуры и функции белков *in vitro* и *in vivo* (Hermanson (1996) in *Bioconjugate Techniques*, Academic Press: San Diego; and Tsien (1998) *Annu. Rev. Biochem.*, 67:509-544).

[0005] Современные флюоресцентные методы для изучения белков *in vivo* часто основаны на использовании конструкторов слияния с большими флюоресцентными белками, такими как зеленый флюоресцентный белок (GFP). Хотя полезность этого зонда в изучении экспрессии белков, локализации и взаимодействии биомолекул доказана, его большие размеры могут приводить к значительным нарушениям структуры. Присоединение GFP также ограничено С- или N-концом белка-мишени и относительно нечувствительно к окружению (Tsien (1998) *Annu. Rev. Biochem.*, 67:509-44). Чтобы получить соответствующий сигнал от GFP, необходимо много транскриптов, и требуется время задержки для его сворачивания и созревания флюорофора.

[0006] Химические методы также можно применять для селективной модификации белков разнообразными малыми, синтетическими флюорофорами, что минимизирует структурные нарушения (Hermanson, in *Bioconjugate Techniques*, Academic Press: San Diego (1996); Martin *et al.*, *Nat. Biotech.*, 23:1308-1314 (2005); Keppler *et al.*, *Nat. Biotech.*, 21:86-89 (2003); Lin *et al.*, *J. Am. Chem. Sci.*, 128:4542-3 (2006)). Однако, эти методы обычно ограничены остатками на поверхности изолированных белков, которые обладают особой реакционной способностью (например, модификация цистеина производными малеимида; см. Hermanson в *Bioconjugate Techniques* (1996) Academic Press: San Diego), проявляют плохую региоселективность, обладают цитотоксичностью или требуют введения мотивов белка, связывающих краситель.

[0007] Метод биосинтетического мечения с использованием химически неправильно ацилированных тРНК (Mendel *et al.*, *Annu. Rev. Biophys. Biomol. Struct.* (1995) 24:435-462; Hohsaka *et al.*, *FEBS Lett.* (2004) 560:173-177) продемонстрировали экспериментально, но он дает ограниченный выход белка и обычно его проводят только *in vitro*.

[0008] Введение генетически кодируемых флюоресцентных аминокислот в определенных сайтах в белках непосредственно в живых организмах позволил бы преодолеть многие ограничения флюоресцентного мечения белков (Wang and

Schultz, *Angew. Chem. Int. Ed.* (2005) 44:34-66; Wang *et al.*, *Annu. Rev. Biophys. Biomol. Struct.*, (2006) 35:225-249). Сайт-специфичное встраивание флюоресцентных аминокислот вызывало бы минимальные нарушения в белках-хозяевах и обеспечило бы возможность измерения резонансного переноса энергии флюоресценции (FRET) с гораздо более высокой точностью ((Truong and Ikura, *Curr. Opin. Struct. Bio.* 2001, 11:573-578). Кроме того, применение флюоресцентной аминокислоты позволит проводить зондирование окружения положения каждой аминокислоты, и точно определять остатки, которые опосредуют взаимодействие с другими компонентами клетки путем варьирования положения флюоресцентной аминокислоты в белке. Это также может быть очень полезно для исследования укладки белков (Lakowicz, *Principles of Fluorescence Spectroscopy* Ed. 2; Kluwer Academic/Plenum Publishers: New York, 1999), особенно в одномолекулярной системе (Lipman *et al.*, *Science* (2003) 301:1233-1235), поскольку одна молекула белка в норме содержит более одного остатка триптофана, а специфичное химическое мечение белков флюоресцентными зондами является проблематичным.

[0009] В данной области существует потребность в новых стратегиях введения (встраивания) флюоресцентных неприродных аминокислот в белки для изучения структуры и функции белков. Была разработана общая методология для сайт-специфичного встраивания разнообразных неприродных аминокислот в белки прокариотических и эукариотических организмов *in vivo*. В этих методах применяют ортогональные компоненты трансляции белков, которые распознают соответствующий селекторный кодон для встраивания желаемой неприродной аминокислоты в определенном положении во время трансляции полипептида *in vivo*. Эти методы основаны на использовании ортогональной тРНК (О-тРНК), которая распознает селекторный кодон, причем соответствующая специфичная ортогональная аминоацил-тРНК-синтетаза (О-РС) присоединяет к О-тРНК неприродную аминокислоту (т.е. «загружает» ее неприродной аминокислотой). Эти компоненты не дают перекрестного взаимодействия ни с одной из эндогенных тРНК, РС, аминокислот или кодонов в организме хозяина (т.е., они должны быть ортогональными). Применение таких пар тРНК-РС обеспечило возможность генетического кодирования большого количества различных по структуре неприродных аминокислот.

[0010] Практика применения ортогональных систем трансляции, которые пригодны для получения белков, включающих одну или более неприродных аминокислот в целом известна в данной области техники, так же как и обычные способы получения ортогональных систем трансляции. Например, см. номера Международной Публикации WO 2002/086075 «METHODS AND COMPOSITIONS FOR THE PRODUCTION OF ORTHOGONAL t-RNA-AMINOACYL-t-RNA SYNTHETASE PAIRS », WO 2002/085923 «IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS», WO 2004/094593 «EXPANDING THE EUCARIOTIC GENETIC CODE», WO 2005/019415, поданной 7 июля 2004 г; WO 2005/007870, поданной 7 июля 2004 г; WO 2005/007624, поданной 7 июля 2004 г и WO 2006/110182, поданной 27 октября 2005 г под названием « OTHOGONAL TRANSLATION COMPONENTS FOR THE IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS». Каждая из этих заявок полностью включена в настоящий документ путем ссылки. Более подробное обсуждение ортогональных систем трансляции, которые встраивают неприродные аминокислоты, и способов их получения и применения можно найти в Wang and Schultz, "Расширение генетического кода," *Chem. Commun. (Camb.)* 1:1-11 (2002); Wang and Schultz "Expanding genetic code", *Angewandte Chemie Int. Ed.*, 44(1):34-66 (2005); Xie and Schultz, " Expanding genetic code", *Methods* 36(3):227-238 (2005); Xie and Schultz, "Adding Amino Acids to Genetic Repertoire", *Curr. Opinion in Chemical Biology* 9(6):548-554 (2005); Wang et al., " Expanding genetic code", *Annu. Rev. Biophys. Biomol. Struct.*, 35:225-249 (2006); и Xie and Schultz, "A chemical toolkit for proteins – an expanded genetic code " *Nat. Rev. Mol. Cell Biol.*, 7(10):775-782 (2006), содержание которых полностью включено путем ссылки.

[0011] В данной области существует необходимость в разработке ортогональных компонентов трансляции, которые встраивают неприродные флюоресцентные аминокислоты в белки, причем возможно встраивание неприродных флюоресцентных аминокислот в определенных положениях. Изобретение, описываемое в настоящем документе, удовлетворяет этим и другим потребностям, что станет очевидно при рассмотрении приведенного ниже раскрытия.

## КРАТКОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[0012] Настоящее изобретение обеспечивает соединения и способы для встраивания неприродной кумариновой аминокислоты L-(7-гидроксикумарин-4-ил)-этилглицина в растущую цепь полипептида в ответ на селекторный кодон, например, терминирующий амбер-кодон, *in vivo* (например, в клетке-хозяине). Указанные композиции включают пары ортогональных тРНК (О-тРНК) и ортогональных аминоацил-тРНК-синтетаз (О-РС), которые не взаимодействуют с аппаратом трансляции клетки-хозяина. Т. е., эндогенная аминоацил-тРНК-синтетаза клетки-хозяина не нагружает (или по существу не нагружает) аминокислотой (природной или неприродной) О-РНК. Аналогично, О-РС согласно настоящему изобретению не нагружает ни одну эндогенную тРНК аминокислотой (природной или неприродной) до значимого или обнаруживаемого уровня. Эти новые композиции позволяют получать большие количества белков, в которые при трансляции встраивается L-(7-гидроксикумарин-4-ил)-этилглицин. Такие флюоресцентно меченые белки находят широкое применение.

[0013] Согласно некоторым аспектам настоящего изобретения предложены системы трансляции. Эти системы включают первую ортогональную аминоацил-тРНК-синтетазу (О-РС), первую ортогональную тРНК (О-тРНК) и первую неприродную аминокислоту, которой является L-(7-гидроксикумарин-4-ил)-этилглицин, причем первая О-РС предпочтительно аминоацилирует первую О-тРНК первой неприродной аминокислотой.

[0014] В системах трансляции можно применять компоненты, полученные из разных источников. В одном из вариантов первую О-РС получают из аминоацил-тРНК-синтетазы *Methanococcus jannaschii*, например, из тирозил-тРНК-синтетазы *Methanococcus jannaschii* дикого типа. О-РС, применяемая в системе, может содержать последовательность аминокислот, определенную в SEQ ID NO: 4, и консервативные варианты указанной последовательности. В некоторых вариантах О-тРНК представляет собой тРНК, являющуюся супрессором амбер-кодона. В некоторых вариантах О-тРНК включает EQ ID NO: 1 или кодируется указанной последовательностью.

[0015] Согласно некоторым аспектам система трансляции дополнительно включает нуклеиновую кислоту, кодирующую представляющий интерес белок, причем нуклеиновая кислота имеет по меньшей мере один селекторный кодон, распознаваемый О-тРНК.



[0016] Согласно некоторым аспектам система трансляции включает вторую ортогональную пару (т. е., вторую О-РС и вторую О-тРНК), которая использует вторую неприродную аминокислоту, благодаря чему такая система может  
5 выстраивать по меньшей мере две различных неприродных аминокислоты в разных выбранных сайтах в полипептиде. В такой двойной системе вторая О-РС предпочтительно аминоацилирует вторую О-тРНК второй неприродной  
10 аминокислотой, которая отлична от первой неприродной аминокислоты, а вторая О-тРНК распознает селекторный кодон, который отличен от селекторного кодона, распознаваемого первой О-тРНК.

[0017] Согласно некоторым вариантам реализации система трансляции  
15 находится в клетке-хозяине (и включает клетку хозяина). Применяемые клетки-хозяева особым образом не ограничены, при условии, что О-РС и О-тРНК сохраняют свою ортогональность внутри клетки-хозяина, которая их содержит. Клетка-хозяин может представлять собой клетку эубактерии, например, *E.coli*.  
20 Клетка-хозяин может содержать один или более полинуклеотидов, которые кодируют компоненты системы трансляции, включая О-РС или О-тРНК. В некоторых вариантах реализации полинуклеотид, кодирующий О-РС, содержит  
25 последовательность, определенную в SEQ ID NO: 5.

[0018] Настоящее изобретение также обеспечивает способы получения  
30 белков, содержащих одну или несколько неприродных аминокислот в избранном положении. Эти способы используют системы трансляции, описанные выше. В целом эти способы начинаются с этапа обеспечения системы трансляции, включающей: (i) первую неприродную аминокислоту, которая представляет собой  
35 кумариновую аминокислоту L-(7-гидроксикумарин-4-ил)-этилглицин; (ii) первую ортогональную аминоацил-тРНК-синтетазу (О-РС); (iii) первую ортогональную тРНК (О-тРНК), причем О-РС предпочтительно аминоацилирует эту О-тРНК неприродной аминокислотой; и (iv) нуклеиновую кислоту, кодирующую белок,  
40 причем данная нуклеиновая кислота содержит по меньшей мере один селекторный кодон, который распознает первая О-тРНК. В таком случае способ встраивает неприродную аминокислоту в выбранном положении в белок во время  
45 трансляции белка в ответ на селекторный кодон, что обеспечивает продуцирование белка, содержащего неприродную аминокислоту в выбранном положении.

[0019] Эти способы можно широко применять с использованием разнообразных реагентов и этапов. Согласно некоторым вариантам реализации предложен полинуклеотид, кодирующий О-РС. Согласно некоторым вариантам реализации предложена О-РС, полученная из аминоксил-тРНК-синтетазы *Methanococcus jannaschii*, например, тирозил-тРНК-синтетазы дикого типа *Methanococcus jannaschii*. Согласно некоторым вариантам реализации этап обеспечения включает обеспечение О-РС, содержащую последовательность аминокислот, определенную в SEQ ID NO: 4 и консервативные варианты указанной аминоксилоты.

[0020] Согласно некоторым вариантам реализации этих способов, этап обеспечения системы трансляции представляет собой осуществление мутации кармана связывания аминоксилоты в аминоксил-тРНК-синтетазе дикого типа методом сайт-направленного мутагенеза, и отбор полученной О-РС, которая предпочтительно аминоксилует О-тРНК неприродной аминоксилотой. Этап отбора может включать положительный отбор и отрицательный отбор О-РС из пула молекул аминоксил-тРНК-синтетазы, полученных в результате сайт-направленного мутагенеза. Согласно некоторым вариантам реализации, этап обеспечения обеспечивает полинуклеотид, кодирующий О-тРНК, например, О-тРНК, представляет собой тРНК, являющуюся супрессором амбер-кодона, или О-тРНК, которая содержит полинуклеотид, определенный в SEQ ID NO: 1, или кодируется таким полинуклеотидом. В данном способе этап обеспечения также может обеспечивать нуклеиновую кислоту, включающую селекторный амбер-кодон, который использует система трансляции.

[0021] Также, эти способы можно модифицировать таким образом, чтобы они обеспечивали встраивание более одной неприродной аминоксилоты. В таких способах совместно с первой системой трансляции используют вторую ортогональную систему трансляции, причем вторая система имеет другую специфичность в отношении аминоксилоты и селекторного кодона. Например, этап обеспечения может включать обеспечение второй О-РС и второй О-тРНК, причем вторая О-РС предпочтительно аминоксилует вторую О-тРНК второй неприродной аминоксилотой, которая отлична от первой неприродной аминоксилоты, и вторая О-тРНК распознает в нуклеиновой кислоте селекторный кодон, который отличен от селекторного кодона, распознаваемого первой О-тРНК.

[0022] Способы получения белка с неприродными аминокислотами также можно осуществить в клетке-хозяине. В таких случаях обеспечивают клетку хозяина, причем клетка-хозяин содержит неприродную аминокислоту, О-РС, О-тРНК и нуклеиновую кислоту с по меньшей мере одним селекторным кодоном, которая кодирует белок, и культивирование клетки-хозяина обеспечивает встраивание неприродной аминокислоты. Согласно некоторым вариантам реализации, этап обеспечения включает обеспечение эубактериальной клетки-хозяина (например, *E. coli*). Согласно некоторым вариантам реализации, этап обеспечения включает обеспечение клетки-хозяина, которая содержит полинуклеотид, кодирующий О-РС. Например, полинуклеотид, кодирующий О-РС, может содержать последовательность нуклеотидов, определенную в SEQ ID NO: 5. Согласно некоторым вариантам реализации, этап обеспечения системы трансляции осуществляют путем обеспечения экстракта клеток.

[0023] Согласно данному изобретению также предложены разнообразные композиции, включая нуклеиновые кислоты и белки. Природа композиции не ограничена особо, при условии, что композиция содержит указанную нуклеиновую кислоту или белок. Композиции согласно настоящему изобретению могут содержать любое количество дополнительных компонентов любой природы.

[0024] Например, согласно настоящему изобретению предложены композиции, включающие полипептиды О-РС, причем указанные полипептиды содержат последовательность аминокислот, определенную в SEQ ID NO: 4, или ее консервативный вариант, при этом полипептид, являющийся консервативным вариантом, аминоацилирует когнатную ортогональную тРНК (О-тРНК) неприродной аминокислотой с эффективностью, которая составляет по крайней мере 50% от эффективности, наблюдаемой в системе трансляции, включающей О-тРНК, неприродную аминокислоту и аминоацил-тРНК-синтазу, содержащую последовательность аминокислот, определенную в SEQ ID NO: 4. Также согласно данному изобретению предложены полинуклеотиды, которые кодируют какие-либо из полипептидов, перечисленных выше. Согласно некоторым вариантам реализации, эти полинуклеотиды могут включать последовательность нуклеотидов, определенную в SEQ ID NO: 5. Согласно некоторым вариантам реализации, полипептиды находятся в клетке.

[0025] Также согласно данному изобретению предложены композиции полинуклеотидов, содержащие последовательность нуклеотидов, определенную в

SEQ ID NO: 5. Согласно некоторым вариантам реализации настоящего изобретения предложены векторы, содержащие полинуклеотиды, например, векторы экспрессии. Согласно некоторым вариантам реализации настоящего изобретения предложены клетки, содержащие вектор, описанный выше.

## ОПРЕДЕЛЕНИЯ

[0026] Перед детальным описанием изобретения следует пояснить, что данное изобретение не ограничено конкретными биологическими системами, которые, безусловно, могут варьировать. Также следует понимать, что терминология, используемая в настоящем документе, предназначена исключительно для описания конкретных вариантов реализации, и не предполагается, что она будет ограничивать изобретение. В данном описании и прилагаемой формуле изобретения формы единственного числа, охватывают также определяемый объект во множественном числе, если контекст явно не указывает на другое. Так, например, термин «клетка» включает комбинации двух или более клеток; термин «полинуклеотид», фактически включает много копий этого полинуклеотида.

[0027] Если техническим и научным терминам, применяемым здесь, не даны особые определения в данном разделе и далее в оставшейся части документа, они имеют значение, принятое в одном из специальных разделов знаний, к которому имеет отношение данное изобретение.

[0028] *Ортогональный*: в настоящем документе термин «ортогональный» означает молекулу (например, ортогональная тРНК (О-тРНК) и/или ортогональная аминоксил-тРНК-синтетаза (О-РС)), которая функционирует с эндогенными компонентами клетки с меньшей эффективностью, чем соответствующая молекула, которая является эндогенной по отношению к клетке или системе трансляции, или которая не способна функционировать с эндогенными компонентами клетки. В отношении тРНК и аминоксил-тРНК синтетаз ортогональный означает неспособность или сниженную эффективность, например, менее 20% эффективности, менее 10% эффективности, менее 5% эффективности или менее 1% эффективности, с которой ортогональная тРНК функционирует вместе с эндогенной тРНК синтетазой, по сравнению с функционированием эндогенной тРНК вместе с эндогенной тРНК синтетазой, или по сравнению с функционированием ортогональной аминоксил-тРНК синтетазы с

эндогенной тРНК по сравнению с функционированием эндогенной тРНК синтетазы с эндогенной тРНК. Для ортогональной молекулы в клетке нет нормально функционирующей эндогенной комплементарной молекулы. Например, ортогональную тРНК в клетке любая эндогенная РС аминоацилирует клетки с пониженной или даже нулевой эффективностью, по сравнению с аминоацилированием эндогенной тРНК эндогенной РС. Другой пример: ортогональная РС аминоацилирует любую эндогенную тРНК клетки, представляющую интерес, с пониженной или даже нулевой эффективностью, по сравнению с аминоацилированием эндогенной тРНК эндогенной РС. В клетку можно ввести вторую ортогональную молекулу, которая будет действовать вместе с первой ортогональной молекулой. Например, ортогональная пара тРНК/РС содержит введенные комплементарные компоненты, которые эффективно функционируют в клетке (например, с 45%-ой эффективностью, 50%-ой эффективностью, 60%-ой эффективностью, 70%-ой эффективностью, 75%-ой эффективностью, 80%-ой эффективностью, 90%-ой эффективностью, 95%-ой эффективностью, 99 %-ой или большей эффективностью) по сравнению с контролем, например, соответствующей парой эндогенных тРНК/АРС, или активной ортогональной парой (например, парой ортогональная тРНК тирозила/РС).

[0029] *Ортогональная тирозил-тРНК*: в настоящей заявке ортогональная тирозил-тРНК (тирозил-О-тРНК) – это тРНК, которая является ортогональной по отношению к представляющей интерес системе трансляции, причем тРНК является (1) идентичной или по существу аналогичной существующей в природе тирозил-тРНК, (2) полученной из существующей в природе тирозил-тРНК путем естественного или искусственного мутагенеза, (3) полученной любым способом, который учитывает последовательность тирозил-тРНК дикого типа или последовательность мутантной тирозил-тРНК (1) или (2), (4) гомологичной тирозил-тРНК дикого типа или мутантной тирозил-тРНК; (5) гомологичной любому примеру тРНК, который обозначен как субстрат для тирозил-тРНК-синтетазы на *Фиг. 9*, или (6) консервативным вариантом любого примера тРНК, который обозначен как субстрат для тирозил-тРНК-синтетазы на *Фиг. 9*. Тирозил-тРНК может находиться в состоянии с присоединенной аминокислотой («нагруженном», аминоацилированном), или в состоянии без аминокислоты. Также очевидно, что когнатная синтетаза может присоединить к "тирозил-О-тРНК" (т.е.

аминоацилировать тирозил-О-тРНК) также аминокислоту, отличную от тирозина, соответственно, например, не встречающейся в природе аминокислотой. Действительно, очевидно, что тирозил-О-тРНК согласно настоящему изобретению

5 выгодно применять для встраивания по существу любой аминокислоты, природной или не встречающуюся в природе, в растущий полипептид в процессе трансляции в ответ на селекторный кодон.

10 [0030] *Ортогональная синтетаза аминокислоты тирозина*: применяемый в настоящем документе термин ортогональная синтетаза аминокислоты тирозила (тирозил-О-РС) – это фермент, который предпочтительно аминоацилирует тирозил-О-тРНК аминокислотой в интересующей нас системе трансляции.

15 Аминокислота, которую тирозил-О-РС загружает на тирозил-О-тРНК, может быть любой аминокислотой, природной, неприродной или искусственной, и здесь не имеет никаких ограничений. Синтетаза не обязательно является такой же, как

20 существующая в природе синтетаза аминокислоты тирозила, или ее гомологом, или такой же как синтетаза, разработанная как О-РС на *Фигуре 9*, или ее гомологом. Например, О-РС может быть консервативным вариантом тирозил-О-РС, определенной на *Фигуре 9*, и/или может быть по крайней мере идентичной О-

25 РС на *Фигуре 9* по своей последовательности на 50%, %, 60%, 70%, 80%, 90%, 95%, 98%, 99% или более.

30 [0031] *Когнатный*: термин "когатный" относится к компонентам, которые функционируют вместе, или имеют обладают некоторой степенью в отношении друг друга, например, ортогональная тРНК и ортогональная аминоацил-тРНК синтетаза. Компоненты могут быть также называться комплементарными.

35 [0032] *Предпочтительно аминоацилирует*: в настоящей заявке применительно к и ортогональным системам трансляции, О-РС "предпочтительно аминоацилирует" когатную О-тРНК, в том случае, если О-РС присоединяет к О-

40 тРНК аминокислоту с более высокой эффективностью, чем к любой эндогенной тРНК в системе экспрессии. Таким образом, если О-тРНК и любая заданная эндогенная тРНК присутствуют в системе трансляции приблизительно в равном молярном отношении, О-РС будет «нагружать» О-тРНК с большей частотой, чем

45 эндогенную тРНК. Предпочтительно, относительное соотношение О-тРНК, которую «нагружает» О-РС, к эндогенной тРНК, которую нагружает О-РС, является высоким, предпочтительно приводящее к тому, что О-РС нагружает

50 исключительно или почти исключительно О-тРНК, если О-тРНК и эндогенная тРНК

присутствуют в системе трансляции в равных молярных концентрациях. Относительное соотношение между О-тРНК и эндогенной тРНК, которые «нагружает» О-РС, если О-тРНК и О-РС присутствуют в равных молярных концентрациях, более 1:1, предпочтительно, по меньшей мере, около 2:1, более предпочтительно 5:1, еще более предпочтительно 10:1, еще более предпочтительно 20:1, еще более предпочтительно 50:1, еще более предпочтительно 75:1, еще более предпочтительно 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1 или выше.

[0033] О-РС «предпочтительно аминоацилирует О-тРНК неприродной аминокислотой» в том случае, если (а) О-РС преимущественно аминоацилирует О-тРНК по сравнению с эндогенной тРНК, и (b) когда такое аминоацилирование специфично для неприродной аминокислоты по сравнению с аминоацилированием О-тРНК с участием О-РС любой природной аминокислотой. То есть, если неприродные и природные аминокислоты присутствуют в эквимольном количестве в системе трансляции, содержащей О-РС и О-тРНК, О-РС будет нагружать О-тРНК неприродной аминокислотой с большей частотой, чем природной аминокислотой. Желательно, чтобы относительное отношение О-тРНК, нагруженной неприродной аминокислотой, к О-тРНК, нагруженной природной аминокислотой, было высоким. Более предпочтительно, чтобы О-РС нагружала О-тРНК исключительно или почти исключительно неприродной аминокислотой. Относительное соотношение между нагрузкой О-тРНК неприродной аминокислотой и нагрузкой О-тРНК природной аминокислотой, в случае, когда и природная и неприродная аминокислота присутствуют в эквимольных концентрациях, выше, чем 1:1, желательно по меньшей мере около 2:1, более предпочтительно 5:1, еще более предпочтительно 10:1, еще более предпочтительно 20:1, еще более предпочтительно 50:1, еще более предпочтительно 75:1, еще более предпочтительно 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1 или выше.

[0034] *Селекторный кодон*: Термин «селекторный кодон» обозначает кодоны, распознаваемые О-тРНК в процессе трансляции и не распознаваемые эндогенной тРНК. Антикодонная петля О-тРНК распознает селекторный кодон на мРНК и встраивает соответствующую этому кодону аминокислоту, например, не встречающуюся в природе аминокислоту, в данном сайте полипептида. Селекторные кодоны могут включать, например, нонсенс кодоны, такие как, стоп-

кодоны, например, амбер-кодон (янтарный), охра-кодон (охровый), и опал-кодон (опаловый); кодоны содержащие, четыре или более оснований; редкие кодоны; кодоны, полученные из природных или не встречающихся в природе пар оснований и/или тому подобные.

[0035] *Супрессорная тРНК*: тРНК-супрессор – это тРНК, которая изменяет считывание матричной РНК (мРНК) в данной системе трансляции, обычно, путем обеспечения встраивания аминокислоты в ответ на стоп-кодон (то есть, «проскакивания» “read-through” стоп-кодона) в процессе трансляции полипептида. В некоторых аспектах, селекторный кодон согласно настоящему изобретению – это стоп-кодон, например, кодон терминации (например, амбер-, охра- или опал-кодон), кодон, состоящий из четырех оснований, редкий кодон, и т.д..

[0036] *Супрессорная активность*: в настоящей заявке термин “супрессорная активность” относится, в целом, к способности тРНК (например, супрессорной тРНК), обеспечивать «проскок» трансляции (read-through) через кодон (например, селекторный кодон, который представляет собой амбер-кодон или кодон, состоящий из четырех или более оснований), который в ином случае привел бы к терминации трансляции или к ошибке трансляции (например, к сдвигу рамки). Супрессорная (подавляющая) активность супрессорной тРНК можно выразить как процент от активности «проскакивания» трансляции, наблюдаемой в сравнении со второй супрессорной тРНК, или по сравнению с контрольной системой, например, контрольной системой без О-РС.

[0037] Согласно настоящему изобретению предложены различные способы количественного анализа супрессорной активности. Процент супрессии, представляющего интерес селекторного кодона (например, амбер-кодона) конкретной О-тРНК и О-РС, обозначает процентную долю активности данного экспрессируемого тестового маркера (например, LacZ), который включает селекторный кодон, в нуклеиновой кислоте, кодирующей экспрессируемый тестовый маркер, в интересующей нас системе трансляции, где интересующая нас система трансляции включает О-РС и О-тРНК, относительно конструкту положительного контроля, где положительный не имеет О-тРНК, О-РС и селекторного кодона. Таким образом, например, если конструкция активного маркера положительного контроля, которая не содержит селекторного кодона, имеет в данной системе трансляции наблюдаемую активность X выражаемую в единицах, соответствующих используемой тест-системе для анализа маркера, то



процент супрессии тестовой конструкции, включающей селекторный кодон - это процентная доля X, которую конструкция тестового маркера демонстрирует в условиях, по существу идентичных тем, в которых происходила экспрессия маркера положительного контроля, за исключением того, что конструкцию тестового маркера осуществляют в системе трансляции, которая содержит дополнительно О-тРНК и О-РС. Как правило, система трансляции, экспрессирующая тестовый маркер также содержит аминокислоту, которая распознается О-РС и О-тРНК. Полученное значение процентной супрессии можно уточнить путем сравнения тестового маркера и конструкций "фонового" или "отрицательного" контроля, которые содержат тот же селекторный кодон, что и тестовый маркер, но находятся в системе, которая не содержит О-тРНК, О-РС и/или подходящей аминокислоты, распознаваемой О-тРНК и/или О-РС. Такой отрицательный контроль можно применять для нормализации полученных значений процентной супрессии, с целью учета влияния фоновых сигналов маркера в системе трансляции, представляющий интерес.

[0038] Эффективность супрессии может быть определена любым из множества способов измерения, известных в данной области. Например, можно применять способ анализа с использованием репортера β-галактозидазы, например, измененную (дериватизированной) плазмиду lacZ (причем конструкция содержит селекторный кодон в последовательности нуклеиновой кислоты lacZ) вводят в клетки подходящего организма (например, организма, в котором можно применять ортогональные компоненты) вместе с плазмидой, включающей О-тРНК согласно настоящему изобретению. Также можно ввести когнатную синтетазу (в форме полипептида или в форме экспрессируемого полинуклеотида, кодирующего когнатную синтетазу). Клетки выращивают на среде до желаемой плотности, например, до OD600 равной приблизительно 0.5, и выполняют на β-галактозидазу, например, с использованием набора BetaFluor™ (Novagen). Процентную супрессию можно рассчитать как процентную долю активности образца относительно контроля, например, наблюдаемого значения для измененной конструкции lacZ, которая содержит в желаемом положении соответствующий смысловой, а не селекторный кодон.

[0039] Система трансляции: термин «система трансляции» обозначает компоненты, которые встраивают аминокислоту в растущую цепь полипептида (белка). Компоненты системы трансляции могут включать, например, рибосомы,

тРНК, синтетазы, мРНК и подобные компоненты. О-тРНК и/или О-РС согласно настоящему изобретению, можно добавлять, или они могут быть частью системы трансляции *in vitro* или *in vivo*, например, в клетке не эукариот, например, бактерии (такой как *E.coli*), или в клетке эукариот, например, клетке дрожжей, клетке млекопитающих, клетке растений, клетке водорослей, клетке грибов, клетке насекомых и/или подобных.

[0040] *Неприродная аминокислота*: в настоящей заявке термин «неприродная аминокислота» обозначает любую аминокислоту, модифицированную аминокислоту и/или аналог аминокислоты, которые не являются одной из 20 стандартных существующих в природе аминокислот или селеноцистеином, или пирролизинном. Например, неприродная аминокислота L-(7-гидроксикумарин-4-ил)-этилглицин (также обозначаемая 7-гидроксикумарин-этиленглицин; см. *Фигуру 1*, структуру 1) находит применение в данном изобретении.

[0041] *Полученный из*: в настоящей заявке термин «полученный из» относится к компоненту, который выделен или создан с применением указанных молекул или организмов, или информации из указанной молекулы или организма. Например, полипептид, который получен из второго полипептида, может содержать последовательность аминокислот, которая является идентичной или по существу подобной последовательности аминокислот второго полипептида. В случае полипептидов, производные полипептидов могут быть получены, например, путем естественного мутагенеза, искусственного направленного мутагенеза или искусственного случайного мутагенеза. Для получения полипептидов можно применять намеренно направленный мутагенез или намеренно случайный мутагенез, или комбинацию указанных типов мутагенеза. Мутагенез полипептида, приводящий к образованию другого полипептида, полученного из первого, может представлять собой случайное событие (*например*, вызванное неточностью полимеразы), а идентификация полученного полипептида можно осуществлять соответствующими методами скрининга, *например*, описанными в настоящей заявке. После мутагенеза полипептида обычно следует манипуляция над полинуклеотидом, который кодирует данный полипептид.

[00422] *Маркер положительного отбора или скрининга*: в настоящей заявке термин «маркер положительного отбора или скрининга» относится к

маркерам, которые, в случае присутствия, например, экспрессируются, активируются или тому подобное, что обеспечивает отличие клеток, имеющих характерный признак, например, клеток с маркером положительного отбора, от клеток, не имеющих данного признака.

[00433] *Маркер отрицательного отбора или скрининга:* в настоящей заявке термин “маркер отрицательного отбора или скрининга” относится к маркеру который, в случае присутствия, например, экспрессируется, активируется или тому подобное, что позволяет отличить клетки, которые не имеют особенности или признака, на которые осуществляют отбор (например, в сравнении с клеткой, которая обладает такой особенностью или признаком).

[0044] *Репортер:* в настоящей заявке термин «репортер» обозначает компонент, который можно применять для идентификации и/или отбора целевых компонентов представляющей интерес системы. Например, репортер может включать белок, например, фермент, который придает устойчивость или чувствительность к антибиотику (например,  $\beta$ -лактамаза, хлорамфеникол-ацетилтрансфераза (ХАТ), и т.п.), флюоресцентный маркер скрининга (например, зеленый флюоресцентный белок (например, GFP), YFP, EGFP, RFP и др.), люменсцентный маркер (например, люциферазный белок светлячков), маркер скрининга, основанного на аффинности, или маркерные гены положительного или отрицательного отбора, такие как lacZ,  $\beta$ -gal/lacZ ( $\beta$ -галактозидазы), АДГ (алкогольдегидрогеназа), his3, ura3, leu2, lys2, или т.п.

[0045] *Эукариоты:* в настоящей заявке термин «эукариоты» обозначает организмы, принадлежащие к царству Eucarya (Эукариоты - ядерные). Эукариоты обычно отличаются от прокариот своей характерной многоклеточной организацией (но не исключительно многоклеточной, например, водоросли), наличием ограниченного мембраной ядра и других ограниченных мембранами органелл, линейным генетическим материалом (т.е. линейными хромосомами), отсутствием оперонов, наличием интронов, экспонирования и поли-А-мРНК, и другими биохимическими особенностями, такими как другая структура рибосом. Эукариотические организмы включают, например, животных (например, млекопитающих, насекомых, рептилий, птиц и др.), реснитчатых, растения (например, однодольные, двудольные, водоросли и др.), грибы, дрожжи, жгутиковых, микроспоридины, простейших и др.

[0046] *Прокариоты*: в настоящей заявке термин «прокариоты» обозначает организмы, принадлежащие к царству доядерных (также называемых прокариотами). Прокариотические организмы обычно отличаются от эукариот 5 одноклеточной организацией, бесполом размножением путем почкования или деления, отсутствием ограниченного мембраной ядра и других ограниченных мембраной органелл, кольцевой хромосомой, наличием оперонов, отсутствием 10 интронов, смыслового экспонирования и поли-А-мРНК, и другими биохимическими характеристиками, такими как отличительная структура рибосом. Прокариоты включают подцарство эубактерий и архебактерий. По отдельной классификации царства доядерных иногда выделяют цианобактерии (сине-зеленые водоросли) и 15 микоплазму.

[0047] *Бактерии*: в настоящей заявке термин «бактерии» обозначает прокариотические организмы, отличающиеся от *архебактерий*. Аналогично, 20 архебактерии обозначают прокариот, которые не являются эубактериями. Эубактерии и архебактерии можно различить по нескольким морфологическим и биохимическим критериям. Например, чтобы отнести организм к Eubacteria или Archaea используют различия в последовательности рибосомальной РНК, 25 структуре РНК-полимеразы, наличие или отсутствие интронов, чувствительность к антибиотиками, наличие или отсутствие пептидогликанов клеточной стенки и других компонентов клеточной стенки, разветвленная и неразветвленная 30 структура липидов мембраны и наличие/отсутствие гистонов и гистоноподобных белков.

[0048] Примеры эубактерий включают *Escherichia coli*, *Thermus thermophilus*, 35 *Bacillus subtilis* и *Bacillus stearothermophilus*. Примеры архебактерий включают *Methanococcus jannaschii* (Mj), *Methanosarcina mazei* (Mm), *Methanobacterium thermoautotrophicum* (Mt), *Methanococcus maripaludis*, *Methanopyrus kandleri*, 40 *Halobacterium* такие как *Haloferax volcanii* и *Halobacterium* виды NRC-1, *Archaeoglobus fulgidus* (Af), *Pyrococcus furiosus* (Pf), *Pyrococcus horikoshii* (Ph), *Pyrobaculum aerophilum*, *Pyrococcus abyssi*, *Sulfolobus solfataricus* (Ss), *Sulfolobus tokodaii*, *Aeuryopyrum pernix* (Ap), *Thermoplasma acidophilum* и *Thermoplasma* 45 *volcanium*.

[0049] *Консервативный вариант*: в настоящей заявке термин «консервативный вариант» применительно к компоненту трансляции, обозначает 50 компонент трансляции, например, консервативный вариант О-тРНК или

консервативный вариант О-РС, который функционирует аналогично основному компоненту, аналогом которого является данный консервативный вариант, например, О-тРНК или О-РС, имеющие изменения в последовательности по сравнению с эталонной О-тРНК или О-РС. Например, О-РС, или консервативный вариант той О-РС, будет аминоацилировать когнатную О-тРНК неприродной аминокислотой, например, неприродной аминокислотой L-(7-гидроксикумарин-4-ил)-этилглицином). В данном примере последовательности аминокислот О-РС и консервативного варианта О-РС не являются идентичными. Консервативный вариант может содержать, например, одно изменение, два изменения, три изменения, четыре изменения, или пять или больше изменений в последовательности, при условии, что данный консервативный вариант сохраняет комплементарность (например, функционирует с) родственной соответствующей О-тРНК или О-РС.

[0050] Согласно некоторым вариантам реализации консервативный вариант О-РС содержит одну или несколько консервативных замен аминокислот по сравнению с О-РС, из которой он получен. Согласно некоторым вариантам реализации консервативный вариант О-РС содержит одну или несколько консервативных замен аминокислот по сравнению с О-РС, из которой он получен, и кроме того, сохраняет биологическую активность О-РС; например, консервативный вариант О-РС, который сохраняет по меньшей мере 10% биологической активности родительской молекулы О-РС, из которой он был получен, или, в качестве альтернативы, по меньшей мере 20%, по меньшей мере 30% или по меньшей мере 40%. Согласно некоторым предпочтительным вариантам реализации консервативный вариант сохраняет по меньшей мере 50% биологической активности родительской молекулы О-РС, из которой он был получен. Консервативные замены аминокислот в консервативном варианте О-РС могут происходить в любом домене О-РС, включая карман связывания аминокислоты.

[0051] *Агент отбора или скрининга*: в настоящей заявке термин «агент отбора или скрининга» обозначает агент, который, в случае его присутствия, обеспечивает отбор/скрининг определенных компонентов из популяции. Например, агент отбора или скрининга может быть, без ограничения, например, питательным веществом, антибиотиком, длиной волны света, антителом,

экспрессируемым полинуклеотидом и т.д.. Агент отбора можно варьировать, например, концентрация, интенсивность, и т.д.

[0052] *В ответ на:* в настоящей заявке термин «в ответ на» обозначает любой процесс, в котором О-тРНК согласно настоящему изобретению распознает селекторный кодон и опосредует встраивание неприродной аминокислоты, которая соединена с тРНК, в растущую цепь полипептида.

[0053] *Кодировать:* в настоящей заявке термин «кодировать» обозначает любой процесс, в результате которого информация, содержащаяся в макромолекуле или последовательности полимера, обеспечивает управление образованием второй молекулы или нити последовательности, которая отличается от первой молекулы или нити последовательности. В настоящем описании данный термин применяют в широком смысле, и он может иметь разные применения. Согласно некоторым аспектам, термин «кодировать» описывает процесс полуконсервативной репликации ДНК, при которой одна нить двуцепочечной молекулы ДНК выступает в качестве матрицы, кодирующей вновь синтезируемую комплементарную сестринскую цепь ДНК, для ДНК ДНК-полимеразы.

[0054] Согласно другому аспекту, термин «кодировать» обозначает любой процесс, посредством которого информацию, содержащаяся в одной молекуле, обеспечивает направление синтеза второй молекулы, которая имеет отличную от первой молекулы химическую природу. Например, молекула ДНК может кодировать молекулу РНК (например, через процесс транскрипции с участием фермента ДНК-зависимой РНК-полимеразы). Также, молекула РНК может кодировать полипептид, как в процессе трансляции. Если термин «кодировать» применяют для описания процесса трансляции, этот термин также распространяется на триплетный кодон, который кодирует аминокислоту. Согласно некоторым аспектам, молекула РНК может кодировать молекулу ДНК, например, при процессе обратной транскрипции, включающий РНК-зависимую ДНК-полимеразу. Согласно другому аспекту, молекула ДНК может кодировать полипептид, причем предполагается, что термин «кодировать» в этом случае охватывает и процесс транскрипции, и процесс трансляции.

## КРАТКОЕ ОПИСАНИЕ ФИГУР

[0055] *Фигура 1.* Представлена схема синтеза L-(7-гидроксикумарин-4-ил)-этилглицина (структура 1).

5 [0056] *Фигура 2.* Представлены спектры поглощения и испускания L-(7-гидроксикумарин-4-ил)-этилглицина. Оба спектра были получены при pH 7,4, 100 mM натрий-фосфатного буфера. При этом pH L-(7-гидроксикумарин-4-ил)-  
10 этилглицин присутствует как в форме фенола, так и в форме фенолята (приблизительно в соотношении 2:1). Форма фенолята имеет коэффициент экстинкции 17000 при 360 нМ и квантовый выход 0,63.

15 [0057] *Фигура 3.* Представлена структура миоглобина кашалота (код базы данных о белках (PDB) 105M). Указаны R-группы остатков Ser4 и His37.

[0058] *Фигура 4.* Представлены результаты анализа кругового дихроизма индуцированного мочевиной разворачивания форм миоглобина – дикого типа и  
20 мутантной. Анализируемые мутантные формы представляли собой либо Ser4→7-гидроксикумарин-этилглицин, либо His37→7-гидроксикумарин-этилглицин мутантные холомиоглобины. Разворачивание отслеживали по плотности  
25 электрического тока и по интенсивности флюоресценции 7-гидроксикумарин-этилглицина. И те, и другие эксперименты проводили в 100 mM натрий-фосфатном буфере, pH 7,4, с различными концентрациями мочевины, как  
указано. В присутствии 5 М мочевины оба мутантных миоглобина  
30 демонстрировали повышение флюоресцентного сигнала по сравнению с 0 М мочевины. «Долю свернутого белка» вычисляли путем нормирования флюоресцентного сигнала или молярной эллиптичности при 5 М мочевины на  
35 полностью несвернутое состояние.

[0059] *Фигура 5.* Представлены фотографии SDS-PAGE (электрофорезах в  
волиактриламидном геле с додецилсульфатом натрия) анализа экспрессии  
40 мутантного миоглобина TAG-4 (показан черной стрелкой). На центральной панели показано окрашенное кумасси разделение белков в SDS-PAGE, аккумулялированных в присутствии и в отсутствие 1 mM L-(7-гидроксикумарин-4-ил)-  
этилглицина. На правой панели показано флюоресцентное изображение  
45 миоглобина дикого типа и мутантного миоглобина TAG4 после разделения путем SDS-PAGE. На левой панели показаны окрашенные маркеры молекулярной массы.

[0060] *Фигура 6.* Представлены спектры масс-спектрометрии с ионизацией методом электроспрея (ESI-MS) мутантного миоглобина TAG4. На врезке показан восстановленный спектр. Предполагаемая масса: 18512 Да; экспериментальная: 18511 Да.

[0061] *Фигура 7.* Представлен спектр поглощения 60 мкМ кумариновой аминокислоты L-(7-гидроксикумарин-4-ил)-этилглицина в 100 мМ натрий-фосфатного буфера.

[0062] *Фигура 8.* Представлена структурная модель активного центра мутантного MjTyrRS (CouRS-D8). Модель получили с использованием структуры MjTyrRS дикого типа (код базы данных белков 1J1U). Субстрат тирозин заменяли на L-(7-гидроксикумарин-4-ил)-этилглицин, были показаны следующие мутации: Tyr32Glu, Leu65His, Ala67Gly, His70Gly, Phe108Tyr, Gln109His, Asp158Gly и Leu162Gly. Структуру оптимизировали методом минимизации энергии при помощи программного продукта InsightII (Accelrys Software, Inc., Сан-Диего, Калифорния, США).

[0063] *Фигура 9.* Представлены различные последовательности нуклеотидов и аминокислот.

## ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[0064] В настоящем описании представлены ортогональные пары тРНК/аминоацил-тРНК-синтетаза, которые позволяют осуществлять *in vivo* селективное встраивание флюоресцентной аминокислоты L-(7-гидроксикумарин-4-ил)-этилглицина (см. *Фигуру 1*, структуру 1) в белки в *E.coli* в ответ на селекторный кодон, например, терминирующий амбер-кодон TAG. Согласно данному изобретению предложены новые полипептиды ортогональных аминоацил-тРНК-синтетаз (O-PC), которые специфичным образом нагружают родственную ортогональную тРНК (O-тРНК) неприродной аминокислотой L-(7-гидроксикумарин-4-ил)-этилглицином.

[0065] Молекулу 7-гидроксикумарина первоначально изучали в связи с его высоким квантовым выходом флюоресценции, относительно большим сдвигом Стокса (см. *Фигуру 2*), малым размером, чувствительностью к pH (см. *Фигуру 7*) и полярностью в качестве растворителя (Zinsli, *J. Photochem* (1974) 3:55-69).

[0066] Согласно некоторым аспектам, чтобы продемонстрировать (но не ограничить) настоящее изобретение, раскрытие в настоящем документе



демонстрирует, что фрагмент неприродной аминокислоты можно встроить в модельный белок (см. *примеры 4 и 5*). Не предусматривается, что встраивание неприродной аминокислоты будет ограничено каким-либо конкретным белком. Из настоящего раскрытия станет ясно, что включение неприродной флюоресцентной аминокислоты в конкретный белок, представляющий интерес, полезно для широкого спектра задач.

[0067] В настоящем раскрытии описана разработка новых ортогональных пар тРНК/аминоацил-тРНК-синтетаза, которые функционируют в эубактерии, обеспечивая сайт-специфичное встраивание неприродной флюоресцентной аминокислоты (например, неприродной аминокислоты L-(7-гидроксикумарин-4-ил)-этилглицина, представленной на *Фигуре 1*, структура 1) в ответ на селекторный кодон. Говоря кратко, мы идентифицировали новые мутанты тирозил-тРНК-синтетазы *Methanococcus jannaschii*, которые селективно нагружают супрессорную тРНК неприродной аминокислотой в клетке-хозяине *E.coli*.

[0068] Выделенные пары тРНК-синтетаз можно использовать для сайт-специфичного встраивания неприродной флюоресцентной аминокислоты в белок. Включение неприродной аминокислоты в белок можно программировать таким образом, чтобы оно происходило в любом желаемом положении, путем конструирования полинуклеотида, кодирующего белок, представляющий интерес, таким образом, чтобы он содержал селекторный кодон, который дает сигнал к встраиванию неприродной аминокислоты.

[0069] Описанное в настоящей заявке изобретение обеспечивает ортогональные пары для генетического кодирования и встраивания неприродной аминокислоты L-(7-гидроксикумарин-4-ил)-этилглицина в белки в клетке эубактерии, например, *E.coli*, причем ортогональные компоненты не дают перекрестной реакции с эндогенными компонентами аппарата трансляции клетки-хозяина *E.coli*, но распознают желаемую неприродную аминокислоту и встраивают ее в белки в ответ на селекторный кодон (например, несмысловой амбер-кодон, TAG). Ортогональные компоненты, предложенные согласно изобретению, включают ортогональные аминоацил-тРНК-синтетазы, полученные из тирозил-тРНК-синтетазы *Methanococcus jannaschii*, и мутантную тирозил-тРНК<sub>CUA</sub>, являющуюся супрессором амбер-кодона, которые функционируют как ортогональная пара в эубактериальной клетке-хозяине.

[0070] Согласно данному изобретению предложены соединения и способы идентификации и получения дополнительных ортогональных пар тРНК-аминоацил-тРНК-синтетазы, например, пар ОтРНК/О-РС, которые можно использовать для включения неприродной аминокислоты L-(7-гидроксикумарин-4-ил)-этилглицина в белки. Пара тРНК/О-РС согласно настоящему изобретению обладает способностью опосредовать встраивание 7-гидроксикумарин-этилглицина в белок, который кодируется полинуклеотидом, причем полинуклеотид содержит селекторный кодон, распознаваемый О-тРНК. Антикодонная петля О-тРНК распознает селекторный кодон на тРНК-синтетазе и встраивает неприродную аминокислоту в этот сайт в полипептиде. Обычно ортогональная аминоацил-тРНК-синтетаза согласно настоящему изобретению, предпочтительно аминоацилирует (или нагружает) свою О-тРНК только конкретной неприродной аминокислотой.

#### ТЕХНОЛОГИЯ ОРТОГОНАЛЬНЫХ тРНК/АМИНОАЦИЛ-тРНК-СИНТЕАЗ

[0071] Понимание новых композиций и способов согласно настоящему изобретению требует понимания действий, связанных с парой ортогональной тРНК и ортогональной аминоацил-тРНК синтетазы. Чтобы добавлять в генетический код дополнительные не встречающиеся в природе аминокислоты, необходимо, чтобы новые ортогональные пары, включающие аминоацил-тРНК синтетазу и подходящую тРНК, могли эффективно функционировать в механизме трансляции хозяина, на в то же время быть ортогональными по отношению к системе трансляции хозяина, что означает, что они функционируют независимо от синтетаз и тРНК эндогенных систем трансляции. Желательные признаки ортогональной пары включают тРНК, которая декодирует или распознает только определенный кодон, *например*, селекторный кодон, который не декодируют никакие эндогенные тРНК, и аминоацил-тРНК синтетазу, которая предпочтительно аминоацилирует (или "нагружает") когнатную тРНК только одной определенной не встречающейся в природе аминокислотой. О-тРНК - также обычно не подвергается аминоацилированию (или плохо аминоацилируется, то есть, нагружается) эндогенными синтетазами. Например, в системе хозяина *E. coli*, ортогональная пара будет включать аминоацил-тРНК синтетазу, которая не обладает способностью перекрестной реакции ни с одной из эндогенных тРНК (*например*, *E. coli* существует 40 эндогенных тРНК), и ортогональную тРНК,

которую не аминокацилирует ни одна из эндогенных синтетаз (которых в *E. coli*, например, двадцать одна).

[0072] Общие принципы ортогональных систем трансляции, которые являются подходящими для создания белков и включают одну или более не встречающихся в природе аминокислот, известны в данной области техники, как и общие способы создания ортогональных систем трансляции. Например, см. публикацию международной патентной заявки номер WO 2002/086075, озаглавленную "METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL tRNA-AMINOACYL-tRNA SYNTHETASE PAIRS;" WO 2002/085923, озаглавленную "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS;" WO 2004/094593, озаглавленную "EXPANDING THE EUKARYOTIC GENETIC CODE;" WO 2005/019415, поданную 7 июля, 2004; WO 2005/007870, поданную 7 июля, 2004; WO 2005/007624, поданную 7 июля, 2004; и WO 2006/110182, поданную 27 октября, 2005, озаглавленную "ORTHOGONAL TRANSLATION COMPONENTS FOR THE IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS." Каждая из этих публикаций полностью включена в настоящее описание посредством ссылки ссылкой. Обсуждение ортогональных систем трансляции, которые встраивают не встречающиеся в природе аминокислоты, и способов их создания и использования можно найти также в Wang and Schultz "Expanding the Genetic Code," *Angewandte Chemie Int. Ed.*, 44(1):34-66 (2005), Xie and Schultz, "An Expanding Genetic Code," *Methods* 36(3):227-238 (2005); Xie and Schultz, "Adding Amino Acids to the Genetic Repertoire," *Curr. Opinion in Chemical Biology* 9(6):548-554 (2005); Wang et al., "Expanding the Genetic Code," *Annu. Rev. Biophys. Biomol. Struct.*, 35:225-249 (2006); and Ryu and Schultz, "Efficient Incorporation of Unnatural Amino Acids into Proteins in *Escherichia coli*," *Nat. Methods* (4):263-265 (2006); содержание каждой из них полностью включено в настоящую заявку посредством ссылки.

#### •Ортогональные системы трансляции

[0073] Ортогональные системы трансляции обычно содержат клетки (которые могут быть клетками прокариот, такими как *E. coli*), которые включают ортогональную тРНК (О-тРНК), ортогональную аминокацил-тРНК-синтетазу (О-РС), и неприродную аминокислоту, причем О-РС аминокацилирует О-тРНК неприродной аминокислотой. Ортогональная пара согласно настоящему изобретению может

включать О-тРНК, например, супрессорную тРНК, тРНК со сдвигом рамки считывания или подобную и когнатную О-РС. Ортогональные системы согласно настоящему изобретению обычно могут содержать пары О-тРНК/О-РС либо в клетке-хозяине, либо без клетки-хозяина. Помимо многокомпонентной системы согласно изобретению также предложены новые индивидуальные компоненты, например, новые полипептиды (например, как определено в SEQ ID NO: 4) и полинуклеотиды (например, как определено в SEQ ID NO: 5) ортогональной аминоксил-тРНК-синтетазы

[0074] Как правило, если ортогональная пара распознает селекторный кодон и присоединяет аминокислоту в ответ на селекторный кодон, говорят, что данная ортогональная пара "супрессирует" (подавляет) селекторный кодон. Таким образом, селекторный кодон, который не распознают (например, клетки) эндогенные механизмы системы трансляции, обычно не вызывает присоединения аминокислоты, что приводит к блокированию синтеза полипептида, который был бы в ином случае транслирован с нуклеиновой кислоты. В ортогональной системе, состоящей из пары компонентов, О-РС аминокислирует О-тРНК определенной не встречающейся в природе аминокислотой. Нагруженная О-тРНК распознает селекторный кодон и супрессирует (подавляет) остановку трансляции, вызываемую селекторным кодоном.

[0075] Согласно некоторым аспектам, О-тРНК согласно настоящему изобретению распознает селекторный кодон и обладает по меньшей мере, например, 45%, 50%, 60%, 75%, 80% или 90%, или большей эффективностью супрессии в присутствии когнатной синтетазы в ответ на селекторный кодон по сравнению с эффективностью супрессии О-тРНК, включающей последовательность полинуклеотида, которая представлена далее в настоящем документе в перечне последовательностей, или кодируемой такой последовательностью.

[0076] Согласно некоторым вариантам реализации, эффективность супрессии О-РС и О-тРНК вместе приблизительно, например, в 5 раз, 10 раз, 15 раз, 20 раз или в 25 раз или более выше, чем эффективность супрессии О-тРНК без О-РС. Согласно некоторым аспектам, эффективность супрессии О-РС и О-тРНК вместе составляет по меньшей мере, например, 35%, 40%, 45%, 50%, 60%, 75%, 80% или 90%, или более от эффективности супрессии пары ортогональных

синтетаз, которые указаны в настоящем документа в перечне последовательностей.

[0077] Клетка-хозяин использует пару О-тРНК/О-РС для встраивания неприродной аминокислоты в растущую цепь полипептида, например, с участием нуклеиновой кислоты, которая содержит полинуклеотид, который кодирует представляющий интерес полипептид, причем полинуклеотид содержит селекторный кодон, распознаваемый О-тРНК. Согласно определенным предпочтительным аспектам, клетка может включать одну или более дополнительных пар О-тРНК/О-РС, причем дополнительная О-РС нагружает дополнительную О-тРНК другой неприродной аминокислотой. Например, одна из О-тРНК может распознавать кодон из четырех оснований, а другая О-тРНК может распознавать стоп-кодон. В качестве альтернативы, одна кодирующая нуклеиновая кислоты может содержать множество различных стоп кодонов или множество различных кодонов, состоящих из четырех оснований.

[0078] Как отмечалось, согласно некоторым вариантам реализации, в клетке или другой системе трансляции существует много пар О-тРНК/О-РС, которые позволяют осуществлять встраивание более одной неприродной аминокислоты в полипептид. Например, клетка может дополнительно включать дополнительную другую пару О-тРНК/О-РС и вторую неприродную аминокислоту, причем эта дополнительная О-тРНК распознает второй селекторный кодон и эта дополнительная О-РС предпочтительно аминоацилирует О-тРНК второй неприродной аминокислотой. Например, клетка, которая включает пару О-тРНК/О-РС (где О-тРНК распознает, например, селекторный амбер- кодон), может также содержать вторую ортогональную пару, в которой вторая О-тРНК распознает другой селекторный кодон, например, опал-кодон, кодон из четырех оснований, или подобный. Желательно, чтобы разные ортогональные пары были получены из разных источников, что может облегчить распознавание разных селекторных кодонов.

[0079] В некоторых вариантах реализации, системы включают клетку, такую как клетка *E. coli*, которая содержит ортогональную тРНК (О-тРНК), ортогональную аминоацил-тРНК синтетазу (О-РС), неприродную аминокислоту и нуклеиновую кислоту, которая содержит полинуклеотид, который кодирует полипептид представляющий интерес, причем указанный полинуклеотид содержит селекторный кодон, распознаваемый О-тРНК. Система трансляции может также

быть бесклеточной системой, например, любой из разнообразных коммерчески доступных систем *"in vitro"* транскрипции/трансляции в комбинации с парой О-тРНК/О-РС и не встречающейся в природе аминокислотой, согласно настоящему описанию.

[0080] О-тРНК и/или О-РС могут быть природного происхождения или могут быть получены, например, путем мутации тРНК и/или РС природного происхождения, например, методом создания библиотек тРНК и/или библиотек РС, из любого из разнообразных организмов и/или при использовании любого из разнообразных доступных методов мутации. Например, один способ создания ортогональной пары тРНК/аминоацил-тРНК синтетазы включает в себя импортное гетерологичной (для клетки-хозяина) пары тРНК/синтетазы из, например, любого источника, отличного от данной клетки-хозяина, или множества источников, в клетку-хозяина. Признаки кандидата в гетерологичные синтетазы включают, например, следующие: она не аминоацилирует ни одну тРНК клетки-хозяина, и признаки кандидата в гетерологичные тРНК включают, например, то что ее не аминоацилирует никакая синтетаза клетки-хозяина. Кроме того, гетерологичная тРНК ортогональна всем синтезам клетки-хозяина. Второй способ создания ортогональной пары включает в себя создание библиотек мутантов для скрининга и/или отбора О-тРНК или О-РС. Эти стратегии также можно сочетать.

#### •Ортогональная тРНК (О-тРНК)

[0081] Ортогональная тРНК (О-тРНК), предлагаемая изобретением, желательно должна опосредовать включение неприродной аминокислоты в белок, кодируемый полинуклеотидом, который содержит селекторный кодон, распознаваемый О-тРНК, например, *in vivo* или *in vitro*. Согласно конкретным вариантам реализации, О-тРНК согласно настоящему изобретению обладает по меньшей мере, например, 45%, 50%, 60%, 75%, 80% или 90%, или более эффективности супрессии в присутствии когнатной синтетазы в ответ на селекторный кодон по сравнению с эффективностью супрессии О-тРНК, содержащей последовательность полинуклеотида, которая приведена далее в настоящем документе в перечне последовательностей О-тРНК, или кодируемой такой последовательностью.

[0082] Эффективность супрессии может быть определена любым из множества способов измерения, известных в данной области. Например, можно применять способ анализа с использованием репортера  $\beta$ -галактозидазы, например, измененную (derivatизированную) плазмиду lacZ (причем конструкция содержит селекторный кодон, в последовательности нуклеиновой кислоты lacZ) вводят в клетки подходящего организма (например, организма, в котором можно применять ортогональные компоненты) вместе с плазмидой, включающей О-тРНК согласно настоящему изобретению. Также можно ввести когнатную синтетазу (в форме полипептида или в форме экспрессируемого полинуклеотида, кодирующего когнатную синтетазу). Клетки выращивают на среде до желаемой плотности, например, до OD<sub>600</sub> равной приблизительно 0.5, и выполняют на  $\beta$ -галактозидазу, например, с использованием набора BetaFluor™ для анализа  $\beta$ -галактозидазы (Novagen). Процентную супрессию можно рассчитать как процентную долю активности образца относительно контроля, например, наблюдаемого значения для измененной конструкции lacZ, которая содержит в желаемом положении соответствующий смысловой, а не селекторный кодон.

[0083] Примеры О-тРНК согласно настоящему изобретению приведены в перечне последовательностей в настоящем документе, например, см. *Фиг. 9* и SEQ ID NO: 1. Настоящее описание также содержит руководство по разработке дополнительных эквивалентных молекул О-тРНК. В молекуле тРНК, такой как мРНК О-РС или молекула О-тРНК, тимин (Т) замещен урацилом (U) в данной последовательности (или, наоборот, в кодирующей ДНК), или в последовательности, комплементарной данной. Также можно осуществлять дополнительные модификации оснований с получением по существу функционально эквивалентных молекул.

[0084] Изобретение охватывает также консервативные варианты О-тРНК, соответствующие конкретным О-тРНК, представленным в настоящем описании. Например, консервативные варианты О-тРНК включают молекулы, которые функционируют аналогично конкретным О-тРНК, например, как в перечне последовательностей в настоящем документе, и которые сохраняют L-подобную структуру тРНК за счет соответствующей самокомплементарности, но не содержат последовательности, идентичной той, которая например, приведена в перечне последовательностей или на *Фигуре 9*, и желательно, отличаются от молекул тРНК дикого типа.

[0085] Композиция, содержащая О-тРНК, может также включать ортогональную аминоксил-тРНК-синтетазу (О-РСазу), причем О-РС предпочтительно аминоксилует О-тРНК неприродной аминокислотой. Согласно некоторым вариантам реализации, соединение, включающее О-тРНК, может также включать систему трансляции (например, *in vivo* или *in vitro*). Также, в клетке может присутствовать нуклеиновая кислота, которая содержит полинуклеотид, кодирующий представляющий интерес полипептид, причем полинуклеотид содержит селекторный кодон, распознаваемый О-тРНК, или в клетке может присутствовать комбинация одного или более из перечисленных компонентов.

[0086] Способы получения ортогональной тРНК (О-тРНК) также являются аспектом данного изобретения. О-тРНК, полученная таким способом, также является аспектом изобретения. В определенных вариантах реализации настоящего изобретения, О-тРНК может быть получена путем создания библиотеки мутантов. Библиотеку мутантных тРНК можно создать с использованием различных методов мутагенеза, известных в данной области. Например, мутантную тРНК можно получить путем сайт-специфического мутагенеза, методом случайных точечных мутаций, гомологичной рекомбинации, перетасовкой (shuffling) ДНК или другими методами рекурсивного мутагенеза, при помощи химерных конструкций или с применением любой комбинации указанных методов, например, О-тРНК, имеющей последовательность SEQ ID NO: 1.

[0087] Можно вводить дополнительные мутации в определенном положении(ях), например, в неконсервативном положении(ях), или в консервативном положении, в рандомизированном положении(ях), или в комбинации указанных положений, в желательной петле или области тРНК, например, в петле антикодона, акцепторном стебле, D-ветви или петле, варибельной петле, ТРС-ветви или петле, другой области молекулы тРНК, или их комбинации. Как правило, мутации в тРНК включают мутацию петли антикодона каждого члена библиотеки мутантной тРНК, для обеспечения распознавания селекторного кодона. Способ может дополнительно включать присоединение дополнительных последовательностей к О-тРНК. Как правило, О-тРНК обладает улучшенной ортогональностью для желаемого организма по сравнению с исходным материалом, например, множеству последовательностей тРНК, при сохранении сродства к желаемой АРСазе.



[0088] Способы могут включать анализ подобия (и/или предполагаемой гомологии) последовательностей тРНК и/или аминоксил-тРНК синтетаз для выявления потенциальных кандидатов на роль О-тРНК, О-РС и/или пар О-тРНК/О-РС которые, скорее всего, будут ортогональны определенному организму. Для анализа можно применять компьютерные программы, известные в данной области и описанные в настоящей заявке, например, BLAST, а можно использовать другие программы, которые появятся в будущем. В одном примере, для отбора потенциальных ортогональных компонентов трансляции для использования в *E. coli*, выбирают такую синтетазу и/или тРНК, которые не демонстрируют близкого подобия последовательности с эубактериальными организмами.

[0089] Обычно О-тРНК получают путем, например, отрицательного отбора популяции клеток первого вида, в которой клетки содержат одну из множества потенциальных О-тРНК. Отрицательный отбор отсеивает клетки, которые содержат элемент библиотеки потенциальных О-тРНК, который аминокислирует аминоксил-тРНК синтетазу (АРСаза), которая является эндогенной по отношению к данной клетке. Это обеспечивает множество тРНК, которые являются ортогональными для клетки первого вида.

[0090] В некоторых вариантах реализации при отрицательном отборе селекторный(е) кодон(ы) вводят в полинуклеотид, который кодирует маркер отрицательного отбора, например, фермент, который обеспечивает резистентность к антибиотику, например,  $\beta$ -лактамаза – фермент, дающий определяемый продукт, например,  $\beta$ -галактозидаза, хлорамфениколацетилтрансфераза (CAT), например, какой-либо токсический продукт, такой как барназа, в бессмысловом положении (например, все же продуцирующий функциональную барназу), и др. Скрининг/отбор можно осуществлять путем выращивания популяции клеток в присутствии агента отбора (например, антибиотика, такого как ампициллин). Согласно одному варианту реализации, варьируют концентрацию агента отбора.

[0091] Например, для измерения активности супрессорной тРНК применяют систему селекции, которая основана на супрессии селекторного кодона *in vivo*, например, нонсенс-кодона (например, стоп-кодона) или мутаций сдвига рамки считывания, введенных в полинуклеотид, который кодирует маркер отрицательного отбора, например, в ген  $\beta$ -лактамазы (*bla*). Например,

конструируют варианты полинуклеотида, например, варианты *bla*, с селекторным кодоном в определенном положении (например, A184). Этим полинуклеотидом трансформируют клетки, например, бактерии. В случае ортогональной тРНК, которую не могут эффективно нагружать эндогенные синтетазы *E.coli*, устойчивость к антибиотикам, например, устойчивость к ампициллину, должна быть примерно равна или ниже устойчивости бактерии, не трансформированной плазмидой. Если тРНК не является ортогональной, или в этой системе параллельно происходит экспрессия гетерологичной синтетазы, способной нагружать тРНК, будет наблюдаться более высокий уровень устойчивости к антибиотику, например, к ампициллину. Отбирают клетки, например, бактерии, которые не способны расти на среде Лурия-Бертани (LB), содержащей антибиотики в концентрациях приблизительно равных таковой для клеток, не трансформированных плазмидами.

[0092] В случае токсического продукта (например, рибонуклеазы или барназы), когда элемент множества потенциальных тРНК аминоацилируют эндогенные синтетазы хозяина, например, *E.coli* (т.е., она не ортогональна синтезам хозяина, например, *Escherichia coli*), имеет место супрессия селекторного кодона, и образование токсичного продукта полинуклеотида приводит к смерти клетки. Клетки, содержащие ортогональные тРНК или не функционирующие тРНК, выживают.

[0093] Согласно одному варианту реализации, пул тРНК, которые являются ортогональными по отношению к желаемому организму, затем подвергают положительному отбору, при котором селекторный кодон вводят в маркер положительного отбора, например, кодируемый геном устойчивости к лекарственному препарату, таким как ген  $\beta$ -лактамазы. Выполняют положительный отбор клеток, содержащих полинуклеотид, который кодирует или содержит элемент пула тРНК, которые ортогональны по отношению к клетке и полинуклеотид, кодирующий когнатную РС. В определенных вариантах реализации, вторая популяция клеток содержит клетки, которые не были отсеяны в ходе отрицательного отбора. Полинуклеотиды экспрессируются в клетке, и клетку выращивают в присутствии агента отбора, например, ампициллина. Затем, тРНК отбирают по их способности к аминоацилированию коэкспрессируемой когнатной синтетазой и встраиванию аминокислоту в ответ на данный селекторный кодон. Как правило, эти клетки демонстрируют эффективность

супрессии, повышенную по сравнению с клетками, содержащими нефункционирующие тРНК, или тРНК, которые не могут быть эффективно распознаны синтетазой, представляющей интерес. Клетка, содержащая нефункционирующие тРНК или тРНК, которые синтетаза, представляющая интерес, не может эффективно распознавать, является чувствительной к антибиотикам. Поэтому, тРНК: (i) не являются субстратами для эндогенных синтетаз хозяина, например, *Escherichia coli*, (ii) их может аминокислотировать представляющая интерес синтетаза; и (iii) они участвуют в трансляции, переживают оба отбора.

[0094] Соответственно, один и тот же маркер может быть либо положительным, либо отрицательным маркером, в зависимости от контекста, в котором проводят скрининг. То есть, маркер является положительным маркером, если на него проводят скрининг, но он же является отрицательным маркером, если проводят скрининг против него.

[0095] Строгость отбора, например, положительного отбора, отрицательного отбора или и положительного, и отрицательного отбора, в указанных способах, может включать изменение строгости отбора. Например, по причине того, что барназа является чрезвычайно токсичным белком, строгость отрицательного отбора можно контролировать путем введения различного количества селекторных кодонов в ген барназы и/или с помощью применения индуцибельного промотора. В другом примере, изменяют концентрацию агента отбора или скрининга (например, концентрацию ампициллина). В некоторых аспектах согласно настоящему изобретению, строгость изменяют, поскольку желательная низкая активность в ходе первых циклов отбора. Таким образом, менее строгие критерии отбора применяют в ранних циклах отбора, а более строгие критерии применяют в более поздних циклах отбора. В определенных вариантах реализации, отрицательный отбор, положительный отбор или и отрицательный, и положительный отборы могут повторять многократно. Можно применять различные множественные маркеры отрицательного отбора, маркеры положительного отбора или и маркеры и положительного, и отрицательного отборов. В определенных вариантах реализации, маркеры положительного и отрицательного отборов могут быть одинаковыми.

[0096] В настоящем изобретении для получения ортогональных компонентов трансляции, например, О-тРНК, О-РС и пары О-тРНК/О-РС, которая

присоединяет не встречающуюся в природе аминокислоту в ответ на селекторный кодон можно применять и другие методы отбора/скрининга. Например, указанные маркер отрицательного отбора, маркер положительного отбора или маркеры и положительного и отрицательного отборов могут включать маркер, который флюоресцирует или катализирует люминесцентную реакцию в присутствии подходящего реагента. В другом варианте реализации, продукт маркерного гена обнаруживают путем флуоресцентно-активированной сортировки клеток (FACS) или по люминесценции. Маркер может включать основанный на аффинности маркер скрининга. См. также, Франсиско, Francisco, J. A., et al., (1993) Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. Proc Natl Acad Sci U S A. 90:10444-8.

[0097] Дополнительные способы получения рекомбинантной ортогональной тРНК можно найти, например, в публикациях международных заявок на патент: WO 2002/086075, озаглавленной "METHODS AND COMPOSITIONS FOR THE PRODUCTION OF ORTHOGONAL tRNA АМИНОАЦИЛ-tRNA SYNTHETASE PAIRS;" WO 2004/094593, озаглавленной "EXPANDING THE ЭУКАРИОТИЧЕСКИЙ GENETIC CODE;" и WO 2005/019415, поданной 7 июля 2004 года, смотри также Forster et al., (2003) Programming peptidomimetic synthetases by translating genetic codes designed de novo PNAS 100(11):6353-6357; и Feng et al., (2003), Expanding тРНК recognition of a тРНК synthetase by a single amino acid change, PNAS 100(10): 5676-5681.

#### •Ортогональная аминокацил-тРНК-синтетаза (О-РС)

[0098] О-РС согласно настоящему изобретению предпочтительно аминокацилирует О-тРНК неприродной аминокислотой, *in vitro* или *in vivo*. О-РС согласно настоящему изобретению может быть представлена в системе трансляции, например, в клетке, в виде полипептида, который содержит О-РС, и/или в виде полинуклеотида, который кодирует О-РС или ее часть. Например, образец О-РС содержит последовательность аминокислот, как определено в четвертом пункте в SEQ ID NO: 4, или ее консервативным вариантом. В другом примере О-РС или ее часть кодируется последовательностью полинуклеотида, который кодирует последовательность, содержащую аминокислоты, определенную в перечне последовательностей или примеров в настоящем

документе, или комплиментарную ей последовательность полинуклеотида. Например, см. полинуклеотид SEQ ID NO: 5.

[0099] Способы идентификации ортогональной аминоксил-тРНК-синтетазы (О-РС), например, О-РС для применения с какой-либо О-тРНК, также являются аспектом настоящего изобретения. Например, способ включает проведение отбора, например, положительного отбора, популяции клеток первого вида, клетки которого индивидуально включают: 1) элемент множества аминоксил-тРНК-синтетаз (РС), (например, множество РС может включать мутантные РС, РС, полученные из вида, отличного от первого вида); 2) ортогональную тРНК (О-тРНК) (например, из одного или нескольких видов); и 3) полинуклеотид, который кодирует маркер (например, положительного) отбора и содержит, по крайней мере, один селекторный кодон. Отбирают или выбирают путем скрининга те клетки, которые демонстрируют усиление эффективности супрессии по сравнению с клетками, не имеющими или имеющими сниженное количество данного элемента множества РС. Эффективность супрессии можно измерить при помощи методик, известных в технике и описанных выше. Клетки, имеющие повышенную эффективность супрессии, содержат активную РС, которая аминокислирует О-тРНК. Уровень аминокислирования (*in vitro* или *in vivo*) активной РС первого набора тРНК из первого вида сравнивают с уровнем аминокислирования (*in vitro* или *in vivo*) активной РС второго набора тРНК из второго вида. Уровень аминокислирования можно определять по детектируемому веществу (например, меченой неприродной аминокислоте). Обычно отбирают активную РС, которая аминокислирует второй набор тРНК с большей эффективностью, чем первый набор тРНК, что обеспечивает эффективную (оптимизированную) ортогональную аминоксил-тРНК-синтазу для применения с О-тРНК. О-РС, идентифицированная таким способом, также является аспектом настоящего изобретения.

[0100] Для определения аминокислирования можно использовать любой из множества тестов. Эти анализы можно проводить *in vitro* или *in vivo*. Например, способы анализа аминокислирования *in vitro* описаны в, например, Hoben and Soll (1985) *Methods Enzymol.* 113:55-59. Аминокислирование также можно определять с использованием репортера вместе с ортогональными компонентами трансляции и детектированием репортера в клетке, экспрессирующий полинуклеотид, содержащий, по меньшей мере, один селекторный кодон. См.

также, WO 2002/085923, озаглавленную "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS;" и WO 2004/094593, озаглавленную "EXPANDING THE ЭУКАРИОТИЧЕСКИЙ GENETIC CODE."

5 [0101] С идентифицированной О-РС также можно в дальнейшем производить манипуляции, позволяющие изменить субстратную специфичность синтетазы, таким образом, чтобы О-РС присоединяла к О-тРНК только желаемую  
10 желаемая неприродная аминокислота, а не любая из обычных 20 аминокислот. Способы получения ортогональной аминоацил-тРНК-синтетазы с субстратной специфичностью к неприродной аминокислоте включают осуществление мутаций синтетазы, например, в активном центре синтетазы, в сайте механизма редактирования синтетазы, в других сайтах, путем комбинирования различных доменов синтетазы, или тому подобное, и применение процесса отбора. Используемая стратегия основана на сочетании положительного отбора, после  
15 которого следует отрицательный отбор. При положительном отборе супрессия селекторного кодона, введенного в заменимое положение(я) положительного маркера, позволяет клеткам выживать при давлении положительного отбора. Таким образом, в присутствии и природной, и неприродной аминокислоты, выжившие клетки кодируют активные синтетазы, нагружающие ортогональную супрессорную тРНК либо природной, либо неприродной аминокислотой. При отрицательном отборе супрессия селекторного кодона, вводимого в заменимое  
20 положение(я) отрицательного маркера, удаляет синтетазы со специфичностью к природной аминокислоте. Выжившие во время положительного и отрицательного отбора клетки кодируют синтетазы, которые аминоацилируют (нагружают) ортогональную супрессорную тРНК только неприродной аминокислотой. Эти синтетазы затем можно подвергать дальнейшему мутагенезу, например, перестановкам в ДНК (shuffling) или другим рекурсивным способам мутагенеза.

40 [0102] Библиотеку мутантных О-РС можно получить при помощи различных методов мутагенеза, известных в данной области. Например, мутантную РС можно путем осуществления сайт-специфичных мутаций, случайных точечных мутаций, гомологичной рекомбинации, перестановки в ДНК или других рекурсивных способов мутагенеза, химерных конструкций или сочетания этих  
45 методов. Например, библиотеку мутантных РС можно сделать из двух или более разных, например, более менее объемных, менее многообразных «под-библиотек». Химерные библиотеки РС также включены в изобретение. Следует  
50

отметить, что библиотеки тРНК-синтетаз из различных организмов (например, микроорганизмов, таких как эубактерии или ахебактерии), такие как библиотеки, содержащие природное разнообразие (см. например, Патент США № 6 238 884 Short *et al.*; Патент США № 5 756 316 Schallenberger *et al.*; Патент США № 5 783 431 Petersen *et al.*; Патент США № 5 824 485 to Thompson *et al.*; Патент США № 5 958 672 Short *et al.*), могут по выбору конструироваться и отсортировываться на ортогональные пары.

[0103] После того как синтетазы подвергли положительному и отрицательному отбору/скринингу, эти синтетазы затем можно дополнительно подвергнуть мутагенезу. Например, можно выделить нуклеиновую кислоту, которая кодирует О-РС; из нуклеиновой кислоты можно получить набор полинуклеотидов, который кодирует мутированные О-РС (например, путем случайного мутагенеза, сайт-специфичного мутагенеза, рекомбинации или любого их сочетания); и эти отдельные этапы или комбинацию этих этапов можно повторять до тех пор, пока не будет получена мутированная О-РС, которая будет предпочтительно аминоацилировать О-тРНК неприродной аминокислотой. Согласно некоторым аспектам данного изобретения эти этапы осуществляют много раз, например, по меньшей мере два раза.

[0104] В способах согласно изобретению можно также применять дополнительные уровни строгости отбора/скрининга для получения О-тРНК, О-РС, или их пары. Строгость отбора или скрининга можно менять на одном или двух этапах способа получения О-РС. Изменение может включать, например, изменение количества применяемого агента отбора/скрининга и т.д. Также быть выполнены дополнительные циклы положительного и/или отрицательного отбора. Отбор или скрининг может также включать одно или более изменений в доступности аминокислоты, изменение эффективности трансляции, изменение точности трансляции и т.д. Как правило, указанные одно или более изменений основаны на мутации одного или большего количества генов организма, в котором для получения белка применяют ортогональную пару тРНК-синтетаза-тРНК

[0105] Дополнительные общие детали получения О-РС и изменения субстратной специфичности синтетазы можно найти в публикации международной патентной заявки № WO 2002/086075, озаглавленной "METHODS AND COMPOSITIONS FOR THE PRODUCTION OF ORTHOGONAL tRNA AMINOACYL-tRNA SYNTHETASE PAIRS;" и WO 2004/094593, озаглавленной

"EXPANDING THE EUKARYOTIC GENETIC CODE." Смотри также, Wang and Schultz "Expanding the Genetic Code," *Angewandte Chemie Int. Ed.*, 44(1):34-66 (2005), содержание которых полностью включено в настоящую заявку посредством ссылки.

#### ОРГАНИЗМЫ-ИСТОЧНИКИ И ОРГАНИЗМЫ ХОЗЯИНА

[0106] Ортогональные компоненты трансляции (О-тРНК и О-РС) согласно настоящему изобретению можно получить из любого организма (или комбинации организмов) и применять в системе трансляции хозяина любого другого вида, при условии, что компоненты О-тРНК/О-РС и система хозяина функционируют ортогональным образом. О-тРНК и О-РС ортогональной пары не обязательно должны быть получены из одного и того же организма. В некоторых аспектах, ортогональные компоненты получают из генов *Archaea* (то есть, археобактерии) для использования в системе хозяина, принадлежащего к эубактериям.

[0107] Например, ортогональная О-тРНК может быть получена из организмов археобактерий, например, таких как *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium* таких как *Haloferax volcanii* и *Halobacterium* species NRC-1, *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, *Pyrococcus horikoshii*, *Aeuropyrum pernix*, *Methanococcus maripaludis*, *Methanopyrus kandleri*, *Methanosarcina mazei* (Mm), *Pyrobaculum aerophilum*, *Pyrococcus abyssi*, *Sulfolobus solfataricus* (Ss), *Sulfolobus tokodaii*, *Thermoplasma acidophilum*, *Thermoplasma volcanium* или подобных, или эубактерий, таких как *Escherichia coli*, *Thermus thermophilus*, *Bacillus subtilis*, *Bacillus stearothermophilus* или подобных, а ортогональная О-РС может быть получена из организма или сочетания организмов, например, археобактерий, таких как *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium* such as *Haloferax volcanii* и *Halobacterium* вида NRC-1, *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, *Pyrococcus horikoshii*, *Aeuropyrum pernix*, *Methanococcus maripaludis*, *Methanopyrus kandleri*, *Methanosarcina mazei*, *Pyrobaculum aerophilum*, *Pyrococcus abyssi*, *Sulfolobus solfataricus*, *Sulfolobus tokodaii*, *Thermoplasma acidophilum*, *Thermoplasma volcanium* или подобных, или эубактерий, таких как *Escherichia coli*, *Thermus thermophilus*, *Bacillus subtilis*, *Bacillus stearothermophilus* или подобных. Согласно одному варианту реализации эукариотические источники, например, растения, водоросли, простейшие, грибы, дрожжи, животные (например, млекопитающие, насекомые,



членистоногие, и др.) или подобные, также можно применять в качестве источников О-тРНК и О-РС.

[0108] Отдельные компоненты пары О-тРНК/О-РС могут быть получены из одного и того же организма или из разных организмов. Согласно одному варианту реализации, пара О-тРНК/О-РС происходит из одного и того же организма. В другом случае О-тРНК и О-РС из пары О-тРНК/О-РС происходят из разных организмов.

[0109] О-тРНК, О-РС или пару О-тРНК/О-РС можно отбирать или отсортировывать *in vivo* или *in vitro* и/или применять в клетке, для получения полипептида с неприродной аминокислотой. Применение клеток эубактерий не ограничивается: например, *Escherichia coli*, *Thermus thermophilus*, *Bacillus subtilis*, *Bacillus stearothermophilus*, или подобные. Композиции эубактериальных клеток, содержащих компоненты трансляции согласно настоящему изобретению также являются аспектом изобретения.

[0110] См. также, публикацию международной патентной заявки номер WO 2004/094593, озаглавленную "EXPANDING THE THE EUKARYOTIC GENETIC CODE," содержащую информацию по методам скрининга О-тРНК и/или О-РС в одном виде для применения в другом виде.

[0111] Хотя в ортогональных системах трансляции (например, содержащие О-РС, О-тРНК неприродную аминокислоту) можно использовать культивируемые клетки-хозяина для получения белков, содержащих неприродные аминокислоты, что ортогональная система трансляции согласно настоящему изобретению не обязательно требует интактной, жизнеспособной клетки-хозяина. Например, в ортогональной системе трансляции можно использовать бесклеточную систему в присутствии экстракта клеток. Действительно, применение бесклеточных *in vitro* систем транскрипции/трансляции для наработки белка также является общепринятой методикой. Адаптация таких *in vitro* систем для синтеза белков, содержащих не встречающиеся в природе аминокислоты, с применением ортогональных компонентов систем трансляции, описанных в настоящей заявке, находится в рамках данного изобретения.

## СЕЛЕКТОРНЫЕ КОДОНЫ

[0112] Селекторные кодоны согласно настоящему изобретению расширяют объем кодонов генетического кода механизма биосинтеза белков. Например, селекторные кодоны включает, например, уникальные кодоны, содержащие три основания, нонсенс-кодоны, такие как терминирующие кодоны, например, амбер-кодон (UAG), или опал-кодон (UGA), не встречающиеся в природе кодоны, кодоны содержащие, по меньшей мере, четыре основания, редкие кодоны и т.п. В желательный ген можно встроить ряд селекторных кодонов, например, один или более, два или более, более трех, и т.д. Применение различных селекторных кодонов обеспечивает возможность применения многочисленных ортогональных пар тРНК/синтетаза, что делает возможным одновременное сайт-специфичное встраивание множества не встречающихся в природе аминокислот, например, включение, по меньшей мере, одной не встречающейся в природе аминокислоты, с использованием этих различных селекторных кодоны.

[0113] Согласно одному варианту реализации способы включают применение селекторного кодона, который является стоп-кодоном (терминирующим кодоном) для встраивания неприродной аминокислоты в клетку *in vivo*. Например, получают О-тРНК, которая распознает стоп-кодон и которую О-РС аминоацилирует неприродной аминокислотой. Такую О-тРНК не распознают существующие в природе аминоацил-тРНК синтетазы хозяина. Для введения стоп-кодона в представляющий интерес сайт в полинуклеотиде, кодирующем представляющий интерес полипептид, можно применять традиционный сайт-направленный мутагенез. Например, см. Sayers *et al.* (1988), "5',3' Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis," *Nucleic Acids Res*, 791-802. При объединении О-РС, О-тРНК и нуклеиновой кислоты, которая кодирует представляющий интерес полипептид, например, *in vivo*, в ответ на стоп-кодон происходит встраивание неприродной аминокислоты, в результате чего получают полипептид, содержащий неприродную аминокислоту в определенном положении. Согласно одному варианту реализации данного изобретения стоп-кодон, используемый в качестве селекторного кодона, представляет собой амбер-кодон, UAG, и/или опал-кодон, UGA. В одном примере генетический код, в котором используют и UAG, и UGA, в качестве селекторного кодона, может кодировать 22 аминокислоты, при этом сохраняют несмысловой

охра-кодон, UAA, который является наиболее распространенным сигналом терминации.

[0114] Встраивание неприродной аминокислоты *in vivo* можно осуществлять без значительной перестройки клетки-хозяина. Например, в неэукариотических клетках, таких как *Escherichia coli*, благодаря тому, что эффективность экспрессии кодона UAG зависит от конкурирования между О-тРНК, например, амбер-супрессорной тРНК и релизинг-фактором 1 (РФ1) (который связывает кодон UAG и инициирует выделение растущего пептида из рибосомы), эффективность супрессии можно модулировать, например, либо путем повышения уровня экспрессии О-тРНК, например, супрессорной тРНК, либо используя штамм, не имеющий РФ1. В эукариотических клетках, поскольку эффективность экспрессии кодона UAG зависит от конкуренции между О-тРНК, например, амбер-супрессорной тРНК и эукариотическим релизинг-фактором эукариот (например, РФ) (который связывает кодон UAG и инициирует выделение растущего пептида из рибосомы), эффективность супрессии можно модулировать, например, путем повышения уровня экспрессии О-тРНК, например, супрессорной тРНК. Кроме того, также могут присутствовать дополнительные соединения, например, восстановители, такие как дитиотриетол (ДТТ).

[0115] Неприродные аминокислоты также можно кодировать редкими кодонами. Например, было показано, что в случае если концентрация аргинина в реакции синтеза белка *in vitro* снижена, редкий кодон для аргинина AGG обеспечивает эффективное встраивание Ala синтетической тРНК аланином. См. например, Ma *et al.* (1993), *Biochemistry* 32:7939. В данном случае синтетическая тРНК конкурирует с существующей в природе тРНК<sup>Arg</sup>, которая существует как минорная разновидность в *Escherichia coli*. Кроме того, некоторые организмы не используют все триплетные кодоны. Свободный кодон AGA у *Micrococcus luteus* удалось использовать для встраивания аминокислот в экстракт транскрипции/трансляции *in vitro*. См., например, Kowal and Oliver (1997), *Nucl. Acid. Res.*, 25:4685. Согласно настоящему изобретению, могут быть получены компоненты для применения указанных редких кодонов *in vivo*.

[0116] Селекторные кодоны также могут включать расширенные кодоны (т.е. кодоны с увеличенным числом оснований), например, кодоны из четырех и более оснований, такие как кодоны из четырех, пяти, шести и более оснований. Примеры кодонов из четырех оснований включают, например, AGGA, CUAG,

UAGA, CCCU и т.п.. Примеры кодонов из пяти оснований включают, например, AGGAC, CCCCUC, CUAGA, CUACU, UAGGC и т. п. Способы согласно настоящему изобретению включают применение расширенных кодонов для супрессии сдвига рамки считывания. Кодоны из четырех или более оснований могут обеспечивать встраивание одной или более неприродных аминокислот в один и тот же белок. Согласно другим вариантам реализации антикодонные петли могут декодировать, например, кодон из по меньшей мере четырех оснований, кодон из по меньшей мере пяти оснований или кодон из по меньшей мере шести оснований или более. Поскольку существует 256 возможных кодонов из четырех оснований, в одной и той же клетке при помощи кодонов из четырех или более оснований можно закодировать множество неприродных аминокислот. Также см. Anderson *et al.*, (2002) "Exploring the Limits of Codon and Anticodon Size," *Chemistry and Biology*, 9:237-244; and, Magliery, (2001) "Expanding the Genetic Code: Selection of Efficient Suppressors of Four-base Codons and Identification of "Shifty" Four-base Codons with a Library Approach in *Escherichia coli*," *J. Mol. Biol.*, 307:755-769.

[0117] Например, кодоны из четырех оснований использовали для встраивания неприродных аминокислот в белки, с использованием методов биосинтеза *in vitro*. См., например, Ma *et al.*, (1993) *Biochemistry*, 32:7939; and Hohsaka *et al.*, (1999) *J. Am. Chem. Soc.*, 121:34. CGGG и ACGU использовали для одновременного встраивания 2-нафтилаланина и NBD-производного лизина в стрептавидин *in vitro* при помощи двух химически ацилированных тРНК-супрессоров сдвига рамки считывания. См., например, Hohsaka *et al.*, (1999) *J. Am. Chem. Soc.*, 121:12194. Moore *u др.* исследовали *in vivo* способность производных тРНК<sup>Leu</sup> с антикодонами NCUA, супрессировать кодоны UAGN (N может быть U, A, G, или C), и обнаружили, что тРНК<sup>Leu</sup> с антикодоном UCUA сможет декодировать квадруплетный кодон UAGA с эффективностью 13 - 26 %, при слабом декодировании в положении рамки 0 или-1. См. Moore *et al.*, (2000) *J. Mol. Biol.*, 298:195. В одном варианте реализации, в настоящем изобретении можно применять расширенные кодоны, основанные на редких кодонах или нонсенс-кодонах, что может уменьшить бессмысленные (миссенс) проскоки и супрессию сдвига рамки считывания на других нежелательных участках. Четырехосновные кодоны применяли в качестве селекторных кодонов в разнообразных ортогональных системах. См., например, WO 2005/019415; WO 2005/007870 и WO 2005/07624. Смотри также: Wang and Schultz "Expanding the Genetic Code,"

*Angewandte Chemie Int. Ed.*, 44(1):34-66 (2005), содержание которой полностью включено в настоящую заявку посредством ссылки. Хотя в приведенных ниже примерах в качестве селекторного кодона применяют амбер-кодон, кодоны с четырьмя или с большим числом оснований также можно применять. Это можно осуществить путем изменения приведенных примеров с включением О-тРНК с четырьмя основаниями и синтетазы, измененной таким образом, чтобы она включала мутации, аналогичные описанным ранее, для различных не встречающихся в природе аминокислот О-РС.

[0118] В данной системе селекторный кодон также может включать один из природных кодонов из трех оснований, в случае если эндогенная система не использует (или редко использует), естественный кодон. Примеры включают систему, в которой отсутствует тРНК, которая распознает естественные трехосновные кодоны, и/или систему, в котором трехосновный кодон представляет собой редкий кодон.

[0119] Селекторные кодоны могут включать неприродные пары оснований. Такие неприродные пары оснований дополнительно расширяют существующий генетический алфавит. Одна дополнительная пара оснований увеличивает число триплетных кодонов с 64 до 125. Свойства пар третьих оснований включают стабильное и селективное спаривание оснований, эффективное ферментативное встраивание в ДНК с высокой точностью с помощью полимеразы и эффективное непрерывное удлинение праймера после синтеза новой не встречающейся в природе пары оснований. Описания не встречающихся в природе пар х оснований, которые можно адаптировать к способам и композициям, содержатся например, в Hiraio, *et al.*, (2002) *An unnatural base pair for incorporating amino acid analogues into protein*, *Nature Biotechnology*, 20:177-182. См. также Wu, Y., *et al.*, (2002) *J. Am. Chem. Soc.* 124:14626-14630. Другие публикации по этому вопросу перечислены далее.

[0120] Не встречающийся в природе нуклеозид для применения *in vivo* обладает способностью проникать через мембрану и находится в фосфорилированной форме с образованием соответствующего трифосфата. Кроме того, увеличенная генетическая информация устойчива и не подвержена разрушению ферментами клетки. Ранее Беннер и другие использовали в своих работах преимущество структуры водородных связей, которая отличается от структуры в канонических парах Уотсона-Крика, самый примечательный пример

это пара iso-C:iso-G. См., например, Switzer *et al.*, (1989) *J. Am. Chem. Soc.*, 111:8322; и Piccirilli *et al.*, (1990) *Nature*, 343:33; Kool, (2000) *Curr. Opin. Chem. Biol.*, 4:602. Эти основания обычно в некоторой степени не спариваются с природными основаниями и не способны к ферментативной репликации. Kool и сотрудники показали, что гидрофобные уплотняющие взаимодействия между основаниями могут замещать водородные связи, что обеспечивает регуляцию образования пар азотистых оснований. См., Kool, (2000) *Curr. Opin. Chem. Biol.*, 4:602; and Guckian and Kool, (1998) *Angew. Chem. Int. Ed. Engl.*, 36, 2825. Чтобы получить не встречающуюся в природе пару азотистых оснований, удовлетворяющую всем указанным требованиям, Schultz, Romesberg и сотрудники систематически синтезировали и изучали ряд не встречающихся в природе гидрофобных оснований. Оказалось, что пара PICS:PICS является более устойчивой, чем природные пары азотистых оснований, и способная к эффективному включению в ДНК фрагментом Кленова ДНК полимеразы 1 (KF) *Escherichia coli*. См., например, McMinn *et al.*, (1999) *J. Am. Chem. Soc.*, 121:11586; and Ogawa *et al.*, (2000) *J. Am. Chem. Soc.*, 122:3274. KF способная синтезировать пару оснований 3MN:3MN с эффективностью и селективностью, достаточными для биологической функции. См., например, Ogawa *et al.*, (2000) *J. Am. Chem. Soc.*, 122:8803. Однако, оба основания действуют как терминаторы дальнейшей репликации цепи. Недавно была получена мутантная ДНК полимераза, которую можно применять для репликации пары оснований PICS. Кроме того, возможна репликация пары оснований 7Al. См., например, Tae *et al.*, (2001) *J. Am. Chem. Soc.*, 123:7439. Была также получена новая пара оснований, содержащая атом металла, Dipic:Py, которая образует устойчивую пару при связывании Cu(II). См. Meggers *et al.*, (2000) *J. Am. Chem. Soc.*, 122:10714. Поскольку кодоны с большим числом оснований и не встречающиеся в природе кодоны по своей природе ортогональны природным кодонам, в способах согласно настоящему изобретению можно использовать преимущество этого свойства для получения для них ортогональных тРНК.

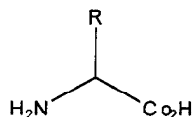
[0121] Также для встраивания неприродной аминокислоты в желаемый полипептид можно использовать систему параллельной трансляции. В системе параллельной трансляции в ген вводят большую последовательность, но не транслируют в белок. Эта последовательность содержит структуру, которая

служит ключом для того, чтобы заставить рибосому перепрыгнуть через последовательность и продолжить трансляцию ниже вставки.

## НЕПРИРОДНЫЕ АМИНОКИСЛОТЫ

[0122] В настоящем документе понятие неприродная аминокислота относится к любой аминокислоте, модифицированной аминокислоте или аналогу аминокислоты, отличным от селеноцистеина и/или пирролизина и следующих двадцати генетически кодируемых альфа-аминокислот: аланин, аргинин, аспарагин, аспарагиновая кислота, цистеин, глутамин, глутаминовая кислота, глицин, гистидин, изолейцин, лейцин, лизин, метионин, фенилаланин, пролин, серин, треонин, триптофан, тирозин, валин. Общую структура альфа-аминокислот иллюстрирует Формула I:

I



[0123] Неприродная аминокислота – это обычно любая структура, имеющая Формулу I, где R-группа – это любой заместитель, отличный от заместителей, встречающихся в двадцати природных аминокислотах. Структура двадцати природных аминокислот описана, например, в *Biochemistry* by L. Stryer, 3<sup>rd</sup> ed. 1988, Freeman and Company, New York. Заметим, что неприродные аминокислоты согласно настоящему изобретению могут представлять собой соединения, существующие в природе, которые отличаются от двадцати альфа-аминокислот, перечисленных выше.

[0124] Поскольку неприродные аминокислоты согласно настоящему изобретению обычно отличаются от природных аминокислот по боковой цепи, неприродные аминокислоты образуют амидные связи с другими аминокислотами, например, с природными или неприродными, таким же образом, что и в существующем в природе белке. Однако неприродные аминокислоты имеют боковые группы, которые отличаются от боковых групп природных аминокислот.

[0125] Особого интереса здесь заслуживает неприродная кумариновая аминокислота L-(7-гидроxicумарин-4-ил)-этилглицин (также обозначаемая 7-гидроxicумарин-этилглицин; см. Фиг.1, структуру 1). Помимо неприродной аминокислоты 7-гидроxicумарина-этилглицина одновременно в представляющий

интерес полипептид можно встраивать другие неприродные аминокислоты, например, путем применения соответствующей второй пары О-РС/О-тРНК в сочетании с ортогональной парой, предложенной согласно настоящему изобретению. Известно много таких дополнительных неприродных аминокислот и подходящих ортогональных пар. См. настоящее описание и ссылки, цитируемые в настоящем документе. Например, см. Wang and Schultz "Expanding the Genetic Code," *Angewandte Chemie Int. Ed.*, 44(1):34-66 (2005); Xie and Schultz, "An Expanding Genetic Code," *Methods* 36(3):227-238 (2005); Xie and Schultz, "Adding Amino Acids to the Genetic Repertoire," *Curr. Opinion in Chemical Biology* 9(6):548-554 (2005); and Wang *et al.*, "Expanding the Genetic Code," *Annu. Rev. Biophys. Biomol. Struct.*, 35:225-249 (2006); содержание которых включено в настоящую заявку путем ссылки.

[0126] Хотя неприродная кумариновая аминокислота L-(7-гидроксикумарин-4-ил)-этилглицин (показанная на *Fig.1*, структура 1) представляет основной интерес в Примерах, описанных в настоящей заявке, не предполагается, что изобретение будет строго ограничено этой структурой. Действительно, можно легко получать разнообразные легко получаемые структурно близкие аналоги, которые будут сохранять принципиальное свойство флюоресценции, и также будут специфичным образом распознаваться аминоксил-тРНК-синтетазами согласно настоящему изобретению (например, О-РС последовательности, определенной в SEQ ID NO: 4). Предполагается, что эти родственные флюоресцентные аналоги находятся входят в объем настоящего.

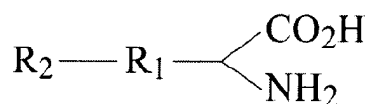
[0127] В других неприродных аминокислотах, например, R в формуле I, возможно, включает алкил-, арил-, ацил-, гидразин-, гало-, гидразид, алкенил, эфир, борат, боронат, фосфо, фосфоно, фосфин, энон, имин, эфир, гидроксилмин, амин, и тому подобные или сочетание указанных фрагментов. Другие представляющие интерес неприродные аминокислоты включают, но не ограничиваются, аминокислоты, включающие фотоактивируемые кросс-линкеры, спин-меченые аминокислоты, флуоресцентные аминокислоты, связывающие металлы аминокислоты, содержащие металл аминокислоты, радиоактивные аминокислоты, аминокислоты с новыми функциональными группами, аминокислоты, которые ковалентно или нековалентно взаимодействуют с другими молекулами, активируемые светом (photocaged) и/или фотоизомеризуемые аминокислоты, биотин или аналог биотина, содержащий аминокислоту, кето-



содержащие аминокислоты, гликозилированные аминокислоты, аминокислоты к боковой цепи которых присоединен фрагмент сахара, аминокислоты, включающие полиэтиленгликоль или полиэфир, аминокислоты с замещающим тяжелым атомом, химически расщепляемые или фоторасщепляемые аминокислоты, аминокислоты с боковой цепью, удлиненной по сравнению с природными аминокислотами (например, полиэфирами или длинными углеводородными цепями, например, более приблизительно 5, более приблизительно 10 атомов углерода, и т.д.), аминокислоты, содержащие связанный с атомом углерода сахар, аминокислоты, содержащие аминокислоты, содержащие один или более токсичных фрагментов.

[0128] Согласно другому аспекту настоящее изобретение обеспечивает аминокислоты, имеющие общую структуру, проиллюстрированную формулой IV ниже:

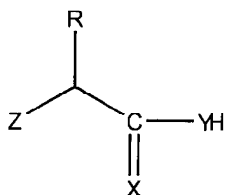
IV



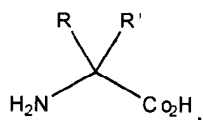
[0129] Неприродная аминокислота, имеющая такое строение обычно представляет собой любую структуру, в которой  $R_1$  – это заместитель, встречающийся в одной из двадцати природных аминокислот (например, тирозин или фенилаланин), а  $R_2$  – это заместитель. Таким образом, этот тип неприродной аминокислоты можно рассматривать как производное природной аминокислоты.

[0130] Помимо неприродных аминокислот, которые содержат кумариновую боковую цепь, как показано на *Фиг.1*, структура 1, неприродные аминокислоты могут, возможно, содержать модифицированные структуры скелета, например, как проиллюстрировано структурами в Формуле II и III:

II



## III



где Z обычно содержит OH, NH<sub>2</sub>, SH, NH-R' или S-R'; X и Y, которые могут  
 быть одинаковыми или разными, обычно содержат S или O, а R и R', которые  
 могут быть одинаковыми или разными, обычно выбраны из того же перечня  
 заместителей для группы R, который приведен выше для неприродных  
 аминокислот, имеющих Формулу I, а также водорода. Например, неприродные  
 аминокислоты согласно настоящему изобретению могут содержать заместители в  
 amino- или карбоксильной группе, как показано в Формулах II и III. Неприродные  
 аминокислоты этого типа включают, без ограничения, α-гидрокси кислоты, α-  
 тио кислоты, α-аминотиокарбоксилаты, например, с боковыми цепями,  
 соответствующими обычным двадцати природным аминокислотам или  
 неприродным боковым цепям. Кроме того, заместители на α-углероде могут  
 включать L, D, или α-α-замещенные аминокислоты, такие как D-глутамат, D-  
 аланин, D-метил-О-тирозин, аминомасляную кислоту и т. п.. Другие структурные  
 варианты включают циклические аминокислоты, такие как аналоги пролина, а  
 также 3, 4, 6, 7, 8 и 9-членные циклические аналоги пролина, β и γ-аминокислоты,  
 такие как замещенные β-аланин и γ-аминомасляная кислота.

[0131] Согласно некоторым аспектам, в настоящем изобретении применяют  
 неприродные аминокислоты в L-конфигурации. Однако, не предполагается что  
 изобретение ограничивается применением неприродных аминокислот в L-  
 конфигурации. Предполагается, что D-энантиомеры этих неприродных  
 аминокислот также найдут применение в настоящем изобретении.

[0132] Неприродные аминокислоты, которые можно применять согласно  
 настоящему изобретению, не ограничены строго кумариновой неприродной  
 аминокислотой L-(7-гидроксикумарин-4-ил)-этилглицин. Любой специалист в  
 данной области поймет, что можно легко получить широкое разнообразие не  
 встречающихся в природе аналогов естественно встречающихся аминокислот. В  
 качестве пример, но без ограничения, легко получают не встречающиеся в  
 природе производные тирозина. Аналоги тирозина включают, например, пара-  
 замещенные тирозины орто-замещенные тирозины и мета-замещенные тирозины,

причем замещенный тирозин содержит алкинильную группу, ацетильную группу, бензоил, аминогруппу, гидразин, гидроксиамин, тиоловую группу, карбоксигруппу, изопропиловую группу, метиловую группа, C6 - C20 углеводородные линейные или разветвленные цепи, насыщенный или ненасыщенный углеводород, группу О-метила, полиэфирную группу, нитро группу или подобные группы. Кроме того, предусмотрены арильные кольца со множеством заместителей. Аналогии

глутамин согласно настоящему изобретению включают, без ограничения, α-гидрокси производные, γ-замещенные производные, циклические производные, и амид-замещенные производные глутамин. Примеры аналогов фенилаланина включают, без ограничения: пара-замещенные фенилаланины, орто-замещенные фенилаланины и мета-замещенные фенилаланины, в которых заместитель содержит алкинильную группу, гидрокси-группу, метокси-группу, метиловую группу, группу аллила, альдегид, нитро, тиоловую группу, кето-группу или подобную группу. Конкретные примеры не встречающихся в природе аминокислот включают, без ограничения: фенилселеноцистеин, сульфотирозин, *p*-этилтиокарбонил -L-фенилаланин, *p*-(3-оксобутаноил) -L-фенилаланин, 1,5-дансил-аланин, 7-амино-кумариновую аминокислоту, 7-гидрокси-кумариновую аминокислота, нитробензил-серин, О-(2-нитробензил)-L-тирозин, *p*-карбоксиметил-L-фенилаланин, *p*-циано-L-фенилаланин, *m*-циано-L-фенилаланин, бифенилаланин, 3-амино-L-тирозин, бипиридил аланин, *p*-(2-амино-1-гидроксиэтил)-L-фенилаланин, *p*-изопропилтиокарбонил -L-фенилаланин, 3-нитро-L-тирозин и *p*-нитро-L-фенилаланин. Кроме того, *p*-пропаргиллоксифенилаланин, 3,4-дигидрокси-L-фенилаланин (DHP), 3, 4, 6-тригидрокси-L-фенилаланин, 3,4,5-тригидрокси-L-фенилаланин, 4-нитро-фенилаланин, *p*-ацетил-L-фенилаланин, О-метил-L-тирозин, L-3-(2-нафтил) аланин, 3-метил-фенилаланина, О-4-аллил-L-тирозин, 4-пропил-L-тирозин, 3-нитро-тирозин, 3-тиол-тирозин, три-О-ацетил-GlcNAcβ-серин, L-дигидроксифенилаланин, фторированный фенилаланин, изопропил-L-фенилаланин, *p*-азидо-L-фенилаланин, *p*-ацил-L-фенилаланин, *p*-бензоил-L-фенилаланин, L-фосфосерин, фосфоносерин, фосфотирозин, *p*-йодо-фенилаланин, *p*-бромфенилаланин, *p*-амино-L-фенилаланин, и изопропил-L-фенилаланин, и т.п.. Разнообразные структуры не встречающихся в природе аминокислот описаны в цитируемых в настоящем описании работах. См. также, опубликованные международные заявки на патент WO 2004/094593,

озаглавленную "EXPANDING THE EUKARYOTIC GENETIC CODE;" и WO 2006/110182, озаглавленную "ORTHOGONAL TRANSLATION COMPONENTS FOR THE *IN VIVO* INCORPORATION OF UNNATURAL AMINO ACIDS," поданную 27 октября, 2005.

•Химический синтез неприродных аминокислот

[0133] Многие из не встречающихся в природе аминокислот, перечисленных выше коммерчески доступны, например в Sigma (США) или Aldrich (Milwaukee, Висконсин, США). Те аминокислоты, которые нельзя приобрести, можно синтезировать, как описано в различных публикациях или с использованием стандартных методов, известных специалистам в данной области. Методы органического синтеза, см., например, в *Organic Chemistry* Fessenden and Fessenden, (1982, Second Edition, Willard Grant Press, Boston Mass.); *Advanced Organic Chemistry* by March (Third Edition, 1985, Wiley and Sons, New York); and *Advanced Organic Chemistry* by Carey and Sundberg (Third Edition, Parts A and B, 1990, Plenum Press, New York). Дополнительные публикации, в которых описан синтез не встречающихся в природе аминокислот включают, например: WO 2002/085923 озаглавленную "In vivo incorporation of Unnatural Amino Acids;" Matsoukas *et al.*, (1995) *J. Med. Chem.*, 38, 4660-4669; King and Kidd (1949) *A New Synthesis of Glutamine and of  $\gamma$ -Dipeptides of Glutamic Acid from Phthylated Intermediates*. *J. Chem. Soc.*, 3315-3319; Friedman and Chatterji (1959) *Synthesis of Derivatives of Glutamine as Model Substrates for Anti-Tumor Agents*. *J. Am. Chem. Soc.* 81, 3750-3752; Craig *et al.* (1988) *Absolute Configuration of the Enantiomers of 7-Chloro-4 [[4-(diethylamino)-1-methylbutyl]amino]quinoline (Chloroquine)*. *J. Org. Chem.* 53, 1167-1170; Azoulay *et al.* (1991) *Glutamine analogues as Potential Antimalarials*, *Eur. J. Med. Chem.* 26, 201-5; Koskinen, and Rapoport (1989) *Synthesis of 4-Substituted Prolines as Conformationally Constrained Amino Acid Analogues*. *J. Org. Chem.* 54, 1859-1866; Christie and Rapoport (1985) *Synthesis of Optically Pure Pípecolates from L-Asparagine. Application to the Total Synthesis of (+)-Apovincamine through Amino Acid Decarbonylation and Iminium Ion Cyclization*. *J. Org. Chem.* 1989:1859-1866; Barton *et al.*, (1987) *Synthesis of Novel  $\alpha$ -Amino-Acids and Derivatives Using Radical Chemistry: Synthesis of L- and D- $\alpha$ -Amino-Adipic Acids, L- $\alpha$ -aminopimelic Acid and Appropriate Unsaturated Derivatives*. *Tetrahedron Lett.* 43:4297-4308; and, Subasinghe *et al.*, (1992) *Quisqualic acid analogues: synthesis of beta-*

*heterocyclic 2-aminopropanoic acid derivatives and their activity at a novel quisqualate-sensitized site. J. Med. Chem. 35:4602-7. См. также, публикацию международной патентной заявки WO 2004/058946, озаглавленной "PROTEIN ARRAYS," поданной 22, 2003.*

•Поглощение неприродных аминокислот клетками

[0134] Поглощение неприродной аминокислоты клеткой - это одна из проблем, которую обычно рассматривают при разработке и отборе не встречающихся в природе аминокислот, например, для встраивания в белок. Например, высокая плотность заряда  $\alpha$ -аминокислот наводит на мысль, что эти соединения вряд ли будут проникать в клетку. Природные аминокислоты транспортируются в клетку рядом систем транспорта основе белка, часто демонстрирующих различные степени специфичности к аминокислотам. Можно провести быстрый скрининг для определения того, какие не встречающиеся в природе аминокислоты, поглощают клетки (и имеет ли вообще место поглощение). См., например, определение токсичности в, например, международной публикации WO 2004/058946, озаглавленной "PROTEIN ARRAYS," поданной 22 декабря 2003 года; и Liu and Schultz (1999) *Progress toward the evolution of an organism with an expanded genetic code. PNAS* 96:4780-4785. Хотя поглощение легко исследовать в различных тест-системах, альтернатива созданию не встречающихся в природе аминокислот, которые поддаются транспорту путей поглощения клеткой, - это обеспечение путей биосинтеза для создания аминокислот *in vivo*.

•Биосинтез неприродных аминокислот

[0135] В клетке изначально существует много путей биосинтеза для продукции аминокислот и других соединений. Если путь биосинтеза для конкретной неприродной аминокислоты не существует в природе, например, в клетке, настоящее изобретение обеспечивает такие способы. Например, можно создать пути биосинтеза не встречающихся в природе аминокислот в клетке-хозяине, путем добавления новых ферментов или изменения существующих путей в клетке-хозяине. Дополнительные новые ферменты могут быть естественно встречающимися ферментами или искусственно созданными ферментами. Например, биосинтез *p*-аминофенилаланин (как показано в примере

в WO 2002/085923) основан на добавлении комбинации известных ферментов других организмов. Гены ферментов могут быть введены в клетку путем трансформации клетки плазмидой, содержащей гены. Экспрессия генов в клетке обеспечивает путь ферментативного синтеза желаемого соединения. Примеры типов ферментов, которые можно добавлять, приведены в изложенных ниже примерах. Последовательности дополнительных ферментов можно найти, например, в Genbank. Таким же образом в клетку можно добавить искусственно измененные ферменты. В таком способе осуществляют манипуляции с механизмами и ресурсами клетки с получением не встречающихся в природе аминокислот.

[0136] Многие доступные способы изменения ферментов и других компонентов путей биосинтеза можно применять в данном изобретении для получения не встречающихся в природе аминокислот (или, действительно, изменять синтетазы для получения новых субстратных специфичностей или других видов активности, представляющих интерес). Например, можно применять перетасовку (shuffling) ДНК для получения новых ферментов и/или путей, включающих такие ферменты, для получения не встречающихся в природе аминокислот (или продукции новых синтетаз), *in vitro* или *in vivo*. См., например, Stemmer (1994), *Rapid evolution of a protein in vitro by DNA shuffling*, *Nature* 370(4):389-391; и, Stemmer, (1994), *DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution*, *Proc. Natl. Acad. Sci. USA.*, 91:10747-10751. В схожем подходе применяют перетасовку семейств родственных (например, гомологичных) генов, для быстрого получения ферментов с желательными характеристиками. Пример такого способа "перетасовки генов семейства" (family gene shuffling) можно найти в Cramer *et al.* (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" *Nature*, 391(6664): 288-291. Новые ферменты (или компоненты биосинтетических путей или синтетазы) могут также быть получены с использованием метода рекомбинации ДНК, известного как "постепенное усеменение для создания гибридных ферментов" ("ITCHY"-incremental truncation for the creation of hybrid enzymes), например, как описано в Ostermeier *et al.* (1999) "A combinatorial approach to hybrid enzymes independent of DNA homology" *Nature Biotech* 17:1205. Этот подход можно также применять для получения библиотеки фермента или других вариантов пути, которые могут служить субстратами для одного или более

способов рекомбинации *in vitro* или *in vivo*. См., также, Ostermeier *et al.* (1999) "Combinatorial Protein Engineering by Incremental Truncation," *Proc. Natl. Acad. Sci. USA*, 96: 3562-67, and Ostermeier *et al.* (1999), "Incremental Truncation as a Strategy in the Engineering of Novel Biocatalysts," *Biological and Medicinal Chemistry*, 7: 2139-44. В другом подходе применяют экспоненциальный мутагенез ансамблей (exponential ensemble mutagenesis) для получения библиотеки фермента или других вариантов пути, которые, например, отбирают по способности катализировать реакцию биосинтеза, подходящую для синтеза не встречающейся в природе аминокислоты (или новой синтетазы). В этом подходе, маленькие группы остатков в последовательности, представляющей интерес, одновременно рандомизируют, чтобы определить в каждом измененном положении аминокислоты, которые приводят к образованию функциональных белков. Примеры таких процедур, которые можно адаптировать для наработки новых ферментов для наработки неприродных аминокислот (или новых синтетаз) согласно настоящему изобретению можно найти в Delegrave and Youvan (1993) *Biotechnology Research* 11:1548-1552. В еще одном подходе можно применять случайный или полуслучайный (semi-random) мутагенез, с использованием олигонуклеотидов, в различной степени близких исходной последовательности (doped от dope - примесь) или вырожденных олигонуклеотидов для разработки фермента и/или компонента пути, например, путем применения общих методов мутагенеза например, Arkin and Youvan (1992) "Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis" *Biotechnology* 10:297-300; или Reidhaar-Olson *et al.* (1991) "Random mutagenesis of protein sequences using oligonucleotide cassettes" *Methods Enzymol.* 208:564-86. Еще один подход, часто называемый "нестохастическим" мутагенезом, в котором задействуют реассоциацию полинуклеотида и сайт-насыщающий (site-saturation) мутагенез, в который можно применять для получения ферментов и/или компонентjd пути, которые затем можно подвергнуть скринингу на способность выполнять функцию одной или более синтетаз или функцию биосинтетического пути (например, синтеза не встречающихся в природе аминокислот *in vivo*). См., например, Short "NON-STOCHASTIC GENERATION OF GENETIC VACCINES AND ENZYMES" WO 00/46344.

[0137] Альтернативу таким методом мутаций представляет собой рекомбинация целых геномов организмов и отбор полученного в результате

потомства для конкретных функций путей (часто именуемые «перетасовкой  
целого генома»). Такой подход можно применять к настоящему изобретению,  
например, путем рекомбинации генома и обора организмов (например, *E.coli* или  
5 другой клетки) на способность продуцировать неприродную аминокислоту (или ее  
промежуточный продукт). Например, способы, изученные в приведенных ниже  
публикациях, можно применять для разработки путей для развития существующих  
и/или новых путей в клетках с целью наработки неприродных аминокислот *in vivo*:  
10 Patnaik *et al.* (2002) "Genome shuffling of lactobacillus for improved acid tolerance"  
*Nature Biotechnology*, 20(7): 707-712; и Zhang *et al.* (2002) "Genome shuffling leads to  
rapid phenotypic improvement in bacteria" *Nature*, February 7, 415(6872): 644-646.

15 [0138] Также доступны другие способы конструирования организмов и путей  
метаболизма, например, для продуцирования желаемых соединений, которые  
также можно применять для продуцирования не встречающихся в природе  
аминокислот. Примеры публикаций, в которых изложены подходящие подходы  
20 конструирования путей включают: Nakamura and White (2003) "Metabolic  
engineering for the microbial production of 1,3 propanediol" *Curr. Opin. Biotechnol.*  
14(5):454-9; Berry *et al.* (2002) "Application of Metabolic Engineering to improve both  
the production and use of Biotech Indigo" *J. Industrial Microbiology and Biotechnology*  
25 28:127-133; Banta *et al.* (2002) "Optimizing an artificial metabolic pathway: Engineering  
the cofactor specificity of Corynebacterium 2,5-diketo-D-gluconic acid reductase for use  
in vitamin C biosynthesis" *Biochemistry*, 41(20), 6226-36; Selivonova *et al.* (2001)  
30 "Rapid Evolution of Novel Traits in Microorganisms" *Applied and Environmental  
Microbiology*, 67:3645, и многие другие.

35 [0139] Независимо от того, какой способ применяют, обычно неприродную  
аминокислоту, продуцируемую при помощи искусственных путей биосинтеза  
согласно настоящему изобретению, получают в концентрациях, достаточных для  
эффективного биосинтеза белка, например, в естественном количестве в клетке,  
40 но не в такой степени, которая могла бы значительно повлиять на концентрацию  
других внутриклеточных аминокислот или истощала бы ресурсы клетки. Обычно  
концентрации, нарабатываемые *in vivo* таким способом, составляют около 10 мМ  
– 0,05 мМ. Когда клетку подвергли модификации с целью наработки ферментов,  
45 нужных для конкретного пути, и синтезировали неприродную аминокислоту, можно  
применить отбор *in vivo* для дальнейшей оптимизации продукции не



встречающейся в природе аминокислоты и на синтез белка рибосом и рост клетки.

5           •Ортогональные компоненты для встраивания неприродных аминокислот

10           [0140] Изобретение обеспечивает композиции и способы получения ортогональных компонентов для встраивания неприродной аминокислоты L-(7-гидроксикумарин-4-ил)этилглицина (см. *фиг. 1*, структура 1) в растущую цепь полипептида в ответ на селекторный кодон, *например*, терминирующий амбер-кодон, нонсенс-кодон, кодон, содержащий четыре или более оснований, и т.д., *например, in vivo*. Например, изобретение обеспечивает ортогональные тРНК (О-тРНК), ортогональные аминоацил-тРНК синтетазы (О-РС) и их пары. Эти пары могут можно применять для встраивания не встречающейся в природе аминокислоты в растущую цепь полипептида.

20           [0141] Композиция согласно настоящему изобретению содержит ортогональную аминоацил-тРНК синтазу (О-РС), причем указанная О-РС избирательно аминоацилирует О-тРНК 7-гидроксикумарин-этилглицином. В 25 определенных вариантах реализации, О-РС содержит последовательность аминокислот, содержащую SEQ ID NO: 4 и ее консервативные варианты. В определенных вариантах реализации настоящего изобретения согласно настоящему изобретению, О-РС предпочтительно аминоацилирует О-тРНК более 30 предпочтительно, чем любую эндогенную тРНК, определенной не встречающейся в природе аминокислотой, причем указанная О-РС имеет тенденцию взаимодействовать с О-тРНК, и где соотношение О-тРНК, нагруженной не 35 встречающейся в природе аминокислотой и эндогенной тРНК, нагруженной той же не встречающейся в природе аминокислоте, более 1:1, и более предпочтительно: О-РС нагружает О-тРНК исключительно или почти исключительно.

40           [0142] Композиция, которая содержит О-РС, может дополнительно включать ортогональную тРНК (О-тРНК), которая О-тРНК распознает селекторный кодон. Как правило, О-тРНК согласно настоящему изобретению обладает, по меньшей 45 мере, около, *например*, 45 %, 50 %, 60 %, 75 %, 80 % или 90 %, или большей эффективностью супрессии в присутствии когнатной синтетазы в ответ на селекторный кодон по сравнению с эффективностью супрессии О-тРНК, содержащей последовательность полинуклеотида, которая представлена в 50 списках последовательностей (*например*, SEQ ID NO: 1) и примерах приведенных

здесь, или кодируемой такой последовательностью. В одном варианте реализации, эффективность супрессии О-РС и О-тРНК вместе больше, чем эффективность супрессии О-тРНК в отсутствии О-РС, например, в 5 раз, 10 раз, 15 раз, 20 раз, 25 раз или более. В некоторых вариантах эффективность супрессии О-РС и О-тРНК вместе составляет, по меньшей мере, 45 % эффективности супрессии ортогональной пары синтетазы тирозил-тРНК, полученной из *Methanococcus jannaschii*.

[0143] Композиция, которая содержит О-тРНК, может включать клетку (например, клетку эубактерии, такую как клетка *E. coli* и т.п., или клетку эукариот, такую как клетка дрожжей), и/или системы трансляции.

[0144] Также настоящим изобретением предусмотрена клетка (например, эубактериальная клетка или дрожжевая клетка), содержащая систему трансляции, также предусмотрена настоящим изобретением, причем система трансляции включает ортогональную тРНК (О-тРНК); ортогональную аминоктил-тРНК синтетазу (О-РС); и неприродную аминокислоту 7-гидроксикумарин-этилглицин. Как правило, О-РС аминоктилирует О-тРНК предпочтительнее, чем любую эндогенную тРНК не встречающейся в природе аминокислотой, причем О-РС имеет тенденцию реагировать с О-тРНК, при этом соотношение О-тРНК, нагруженной не встречающейся в природе аминокислотой, к эндогенной тРНК, нагруженной не встречающейся в природе аминокислотой, более 1:1, и более предпочтительно: О-РС нагружает исключительно или почти исключительно О-тРНК. О-тРНК распознает первый селекторный кодон, и О-РС избирательно аминоктилирует О-тРНК не встречающейся в природе аминокислотой. В одном варианте реализации, О-тРНК содержит или кодирована последовательностью полинуклеотида, которая обозначена SEQ ID NO: 1, или комплементарной ей последовательностью полинуклеотида. В одном варианте реализации, О-РС содержит последовательность аминокислот, которая обозначена SEQ ID NO:4, и консервативные варианты указанной последовательности.

[0145] Клетка согласно настоящему изобретению, возможно, дополнительно содержит другую пару О-тРНК/О-РС и вторую неприродную аминокислоту, причем эта О-тРНК распознает второй селекторный кодон и эта О-РС предпочтительно аминоктилирует соответствующую О-тРНК второй неприродной аминокислотой, где вторая аминокислота отлична от первой аминокислоты. Клетка согласно

настоящему изобретению, возможно, включает нуклеиновую кислоту, которая содержит полинуклеотид, который кодирует представляющий интерес полипептид, причем указанный полинуклеотид содержит селекторный кодон, распознаваемый О-тРНК.

[0146] В некоторых вариантах реализации клетка согласно настоящему изобретению является эубактериальной клеткой (такой как *E.coli*), которая содержит ортогональную т-РНК (О-тРНК), ортогональную аминокил-тРНК-синтетазу, неприродную аминокислоту и нуклеиновую кислоту, которая содержит полинуклеотид, кодирующий белок, представляющий интерес, причем указанный полинуклеотид содержит селекторный кодон, распознаваемый О-тРНК. В некоторых вариантах реализации данного изобретения О-РС предпочтительно аминокилирует О-тРНК неприродной аминокислотой с эффективностью, которая выше эффективности, с которой О-РС аминокилирует эндогенную тРНК.

[0147] В некоторых вариантах реализации данного изобретения О-тРНК согласно настоящему изобретению содержит или кодируется последовательностью полинуклеотида, определенной в перечне последовательностей (например, SEQ ID NO: 1) или в приведенных в настоящем описании примерах, или последовательностью полинуклеотида, комплементарной указанной. В некоторых вариантах реализации данного изобретения О-РС содержит последовательность аминокислот, определенную в перечне последовательностей, или консервативный вариант указанной последовательности. В одном варианте реализации, О-РС или ее часть закодированы последовательностью полинуклеотидов, кодирующей последовательность аминокислот, которая представлена в списке последовательностей или приведенных здесь примерах, или последовательностью, комплементарной такому полинуклеотиду.

[0148] О-тРНК и/или О-РС согласно настоящему изобретению могут быть получены из разнообразных организмов (например, эукариотических и/или неэукариотических организмов).

[0149] Полинуклеотиды также являются признаком данного изобретения. Полинуклеотиды согласно настоящему изобретению (например, SEQ ID NO: 5) включают искусственный полинуклеотид (например, созданный человеком и не существующий в природе), содержащий последовательность нуклеотидов, кодирующую полипептид, определенный в приведенном здесь перечне

последовательностей, и/или комплиментарную указанной последовательности полинуклеотида. Полинуклеотид согласно настоящему изобретению также может включать нуклеиновую кислоту, которая гибридизуется с полинуклеотидом, описанным выше, в очень жестких условиях, по существу по всей длине нуклеиновой кислоты. Полинуклеотиды согласно настоящему изобретению также включают полинуклеотид, который, например, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98% или более идентичен полинуклеотиду существующей в природе тРНК или соответствующей кодирующей нуклеиновой кислоты (но полинуклеотид согласно настоящему изобретению отличается от существующей в природе тРНК или соответствующей кодирующей нуклеиновой кислоты), при этом тРНК распознает селекторный кодон, например, кодон из четырех оснований. Искусственные полинуклеотиды, которые, например, по меньшей мере на 80%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98% или более идентичны любому из перечисленных выше полинуклеотидов и/или полинуклеотид, содержащий консервативный вариант любого из перечисленных выше полинуклеотидов, также относятся к полинуклеотидам согласно настоящему изобретению.

[0150] Векторы, содержащие полинуклеотиды согласно настоящему изобретению также являются признаками изобретения. Например, вектор согласно настоящему изобретению может включать плазмиду, фаг, вирус, вектор экспрессии и/или подобное. Клетка, содержащая вектор согласно настоящему изобретению, также является признаком изобретения.

[0151] Способы получения компонентов пары О-тРНК/О-РС также являются признаками настоящего изобретения. Компоненты, полученные такими способами также являются признаками настоящего изобретения. Например, способы создания, по меньшей мере, одной тРНК, которая является ортогональной по отношению к клетке (О-тРНК), включают получение библиотеки мутантной тРНК; мутагенез антикодонной петли каждого члена библиотеки мутантной тРНК, с обеспечением возможности распознавания селекторного кодона, в результате чего получают библиотеку потенциальных О-тРНК, и отрицательный отбор первой популяции клеток первого вида, в котором клетки содержат член библиотеки потенциальных О-тРНК. Отрицательный отбор исключает клетки, которые включают член библиотеки потенциальных О-тРНК, который аминокатирует

эндогенная аминоксил-тРНК синтетаза (РТС) клетки. Это обеспечивает пул тРНК, которые являются ортогональными по отношению к клетке первого вида, таким образом, получают, по меньшей мере, одну О-тРНК. Также предусмотрена О-тРНК, полученная способами согласно настоящему изобретению.

[0152] В некоторых вариантах реализации способы также включают проведение положительного отбора второй популяции клеток первого вида, в которой клетки содержат элемент пула тРНК, которые ортогональны по отношению к клетке первого вида, родственную тРНК-синтетазу и маркер положительного отбора. Путем положительного отбора отбирают или отсортировывают клетки с получением клеток, которые содержат элемент пула тРНК, аминокисляруемых когнатной аминоксил-тРНК синтетазой, и которые демонстрируют желаемый ответ в присутствии маркера положительного отбора, в результате получают О-тРНК. В определенных вариантах реализации, вторая популяция клеток содержит клетки, которые не были исключены отрицательным отбором.

[0153] Также предложены способы идентификации ортогональной аминоксил-тРНК-синтетазы, которая нагружает О-тРНК неприродной аминокислотой. Например, способы включают проведение отбора популяции клеток первого вида, где клетки содержат: 1) одну из множества аминоксил-тРНК-синтетаз (РС) (например, множество РС может включать мутантные РС, РС, полученные из видов, отличных от первого вида или мутантные РС и РС, полученные из вида, отличного от первого вида); 2) ортогональную тРНК (О-тРНК) (например, из одного или нескольких видов); и 3) полинуклеотид, который кодирует маркер положительного отбора и содержит по меньшей мере один селекторный кодон.

[0154] Клетки (например, клетку-хозяина) подвергают отбору или скринингу на клетки, которые демонстрируют повышение эффективности супрессии по сравнению с клетками, у которых или снижено количество данного члена множества APC или он отсутствует. Эти отобранные/скринированные клетки содержат активную APC, которая аминокислярует О-тРНК. Ортогональная аминоксил-тРНК синтетаза, идентифицированная таким способом, также представляет собой признак настоящего изобретения.

[0155] Способы наработки белков в клетке (например, в зубактериальной клетке, такой как клетка *E.coli*, или подобных, или в клетке дрожжей), имеющих в

избранном положении неприродную аминокислоту, также является признаком изобретения. Например, способ включает выращивание клетки в соответствующей среде, причем клетка содержит нуклеиновую кислоту, содержащую по меньшей мере один селекторный кодон, и кодирующую белок, обеспечение неприродной аминокислоты и встраивание неприродной аминокислоты в конкретном положении в белке в ходе трансляции нуклеиновой кислоты по меньшей мере с одним селекторным кодоном, в результате чего получают белок. Клетка дополнительно содержит: ортогональную тРНК (О-тРНК), которая функционирует в клетке и распознает данный селекторный кодон; и ортогональную аминоацил-тРНК-синтетазу (О-РС), которая предпочтительно аминоацилирует О-тРНК неприродной аминокислотой. Белок, полученный таким способом, также является признаком настоящего изобретения.

[0156] Согласно настоящему изобретению также предложены композиции, которые включают белки, причем белки содержат 7-гидроксикумарин-этилглицин. В некоторых вариантах реализации белок содержит последовательность аминокислот, которая по меньшей мере на 75% идентична последовательности известного белка, например, терапевтического белка, диагностического белка, промышленного фермента или части такого белка. Композиция может включать фармацевтически приемлемый носитель.

#### ■Нуклеиновая кислота и последовательность полипептидов и варианты

[0157] Как описано в настоящей заявке, изобретение обеспечивает последовательности полинуклеотида, кодирующие, например, О-тРНК и О-РС, и последовательности аминокислот полипептида, например, О-РС, и, например, композиции, системы и способы с использованием указанных последовательностей полипептида или полинуклеотида. Примеры указанных последовательностей, например, последовательности аминокислот и нуклеотидов в О-тРНК и О-РС раскрыты в настоящем документе (см. *Фиг.9*, например, SEQ ID NO: 4 и 5). Однако, специалисту будет ясно, что настоящее изобретение не ограничено последовательностями, конкретно указанными в настоящем описании, например, в примерах и списке последовательностей. Специалисту в данной области техники очевидно, что настоящее изобретение обеспечивает также многие родственные последовательности с функциями, описанными в настоящей,

например, полинуклеотиды и полипептиды, кодирующие консервативные варианты О-РС, описанной в заявке.

[0158] Конструкция и анализ вида ортогональной синетазы (О-РС), которая способна аминокислотами О-тРНК не природной аминокислотой 7-гидроксикумарин-этилглицином, описаны в Примерах 3-5. В этих примерах описаны конструирование и анализ видов О-РС, которые способны встраивать не природную аминокислоту 7-гидроксикумарин-этилглицин.

[0159] Изобретение обеспечивает полипептиды (О-РС) и полинуклеотиды, например, О-тРНК, полинуклеотиды, которые кодируют О-РС или ее части, олигонуклеотиды, которые применяют для выделения клонов аминокислот-тРНК синетазы, и т.д. Полинуклеотиды согласно настоящему изобретению включают полинуклеотиды, содержащие один или более селекторный кодон и кодирующие представляющие интерес белки или полипептиды согласно настоящему изобретению. Кроме того, полинуклеотиды согласно настоящему изобретению включают, например, полинуклеотид, включающий последовательность нуклеотидов, которая приведена в SEQ ID NO: 5, и полинуклеотид, который является комплементарным к такому полинуклеотиду или кодирует последовательность такого полинуклеотида. Полинуклеотид согласно настоящему изобретению также включает любой полинуклеотид, который кодирует последовательность аминокислот О-РС, определенную в SEQ ID NO: 4. Аналогично, искусственная нуклеиновая кислота, которая гибридизуется с полинуклеотидом, указанным выше, при очень жестких условиях по существу по всей длине нуклеиновой кислоты и отлична от существующего в природе полинуклеотида) является полинуклеотидом согласно настоящему изобретению. Согласно одному варианту реализации композиция включает полипептид согласно настоящему изобретению и наполнитель (например, буфер, воду, принятый в фармацевтике наполнитель и др.). Также согласно изобретению предложено антитело или антисыворотку, которые специфичным образом вступают в иммунные реакции с полипептидом согласно настоящему изобретению. Искусственный полинуклеотид – это полинуклеотид, который создан человеком и не встречается в природе.

[0160] Полинуклеотид согласно настоящему изобретению также включает искусственный полинуклеотид, который, например, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 90%, по меньшей мере на 95%, по

меньшей мере на 98% или более идентичен полинуклеотиду существующей в природе тРНК (но отличен от существующей в природе тРНК). Полинуклеотид также включает искусственный полинуклеотид, который, например, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98% или более идентичен полинуклеотиду существующей в природе тРНК (но не на 100% идентичный).

[0161] В определенных вариантах реализации, вектор (например, плаزمид, космида, бактериофаг, вирус, и т.д.) содержит полинуклеотид согласно настоящему изобретению. В одном варианте реализации, вектор представляет собой вектор экспрессии. В другом варианте реализации, вектор экспрессии содержит промотор, функционально связанный с одним или более полинуклеотидом согласно настоящему изобретению. В другом варианте реализации, клетка содержит вектор, который содержит полинуклеотид согласно настоящему изобретению.

[0162] Для специалиста будет очевидно, что настоящее изобретение включает множество вариантов описанных последовательностей. Например, в изобретение включены консервативные варианты описанных последовательностей, которые дают функционально идентичную последовательность. Варианты последовательностей полинуклеотида нуклеиновой кислоты, которые гибридизируются, по меньшей мере, с одной описанной последовательностью, считают включенными в настоящее изобретение. Уникальные субпоследовательности последовательностей, описанных здесь, как определено, например, стандартными способами сравнения последовательностей, также включены в изобретение.

#### •Консервативные варианты

[0163] Вследствие вырожденности генетического кода, "молчащие замены" (то есть, замены в последовательности нуклеиновой кислоты, которые не приводят к изменению кодируемого полипептида) являются косвенным признаком, присущим каждой последовательности нуклеиновой кислоты, которая кодирует последовательность аминокислот. Аналогично, "консервативные замены аминокислот", при которых происходит замена одной или ограниченного количества аминокислот в последовательности аминокислот на аминокислоты с очень похожими свойствами, также легко идентифицируют как являющиеся в



высокой степени аналогичные описанной конструкции. Такие консервативные варианты каждой описанной последовательности представляют собой признак настоящего изобретения.

5 [0164] "Консервативные варианты" определенной последовательности нуклеиновой кислоты - это нуклеиновые кислоты, которые кодируют идентичные или по существу идентичные последовательности аминокислот, или, в случае, 10 если нуклеиновая кислота не кодирует последовательность аминокислот, по существу идентичные последовательности. Специалисту будет ясно, что отдельные замены, делеции или вставки, которые приводят к замене, добавлению, или одной удалению природной аминокислоты или небольшого 15 процента аминокислот (обычно меньше 5 %, чаще меньше 4 %, 2 % или 1 %) в кодируемой последовательности, представляют собой "консервативно модифицированные варианты", в котором указанные изменения приводят к 20 делеции аминокислоты, добавлению аминокислоты, или замене аминокислоты химически подобной аминокислотой. Таким образом, "консервативные варианты" представленной в списке последовательности полипептида согласно настоящему изобретению включают замены небольшого процента, обычно менее 5 %, чаще 25 менее 2 % или 1 % аминокислот последовательности полипептида, на аминокислоту из той же группы консервативных замен. Наконец, добавление последовательностей, которые не изменяют кодируемую активность молекулы нуклеиновой кислоты, такое как добавление нефункционирующей 30 последовательности, является консервативным изменением основной нуклеиновой кислоты.

35 [0165] В технике хорошо известны таблицы консервативных замен, в которых представлены функционально тождественные аминокислоты, где один остаток аминокислоты замещен другим остатком аминокислоты, имеющей сходные химические свойства (например, ароматическая боковая цепь или 40 положительно заряженные боковые цепи), и это не приводит к существенному изменению функциональных свойств молекулы полипептида. Далее приведены примеры групп, которые содержат природные аминокислоты с подобными 45 химическими свойствами; замены в пределах группы являются «консервативными заменами».

*Консервативные замены аминокислот*

Неполярные и/или алифатические боковые цепи	Полярные, незаряжен ные боковые цепи	Ароматические боковые цепи	Положитель но заряженные боковые цепи	Отрицатель но заряженные боковые цепи
Глицин Аланин Валин Лейцин Изолейцин Пролин	Серин Треонин Цистеин Метионин Аспарагин Глутамин	Фенилаланин Тирозин Триптофан	Лизин Аргинин Гистидин	Аспартат Глутамат

•Гибридизация нуклеиновых кислот

[0166] Сравнительную гибридизацию можно применять для идентификации нуклеиновых кислот согласно настоящему изобретению, включая консервативные варианты нуклеиновых кислот согласно настоящему изобретению, и этот сравнительный способ гибридизации является предпочтительным способом различения нуклеиновых кислот согласно настоящему изобретению. Кроме того, целевые нуклеиновые кислоты, которые гибридизируются с нуклеиновой кислотой, представленной в SEQ ID NO: 5, 7, 9 или 11, в условиях высокой, ультравысокой и сверхвысокой (ультра-ультравысокой) жесткости являются признаком настоящего изобретения. Примеры таких нуклеиновых кислот включают нуклеиновые кислоты с одной или несколькими молчащими или консервативными заменами нуклеиновой кислоты по сравнению с данной последовательностью нуклеиновой кислоты.

[0167] Говорят, что тестовая нуклеиновая кислота специфично гибридизируется с нуклеиновой кислотой-зондом, если она гибридизируется с зондом по меньшей мере на 50 % также хорошо, как с полностью совпадающей комплементарной мишенью (target), то есть, с отношением сигнала к шуму, составляющим по меньшей мере, половину от уровня гибридизации зонда с мишенью в условиях, при которых идеально совпадающий комплементарный зонд связывается с идеально совпадающей комплементарной мишенью с отношением сигнала к шуму, которое составляет, по меньшей мере, около 5х-10х уровня,

который наблюдается при гибридизации любой из несовпадающих нуклеиновых кислот-мишеней.

[0168] Говорят, что нуклеиновые кислоты "гибридизируются", если они ассоциируют, обычно, в растворе. Гибридизация нуклеиновых кислот происходит за счет разнообразных хорошо описанных физико-химических сил, таких как водородные связи, вытеснение растворителя, межплоскостное взаимодействие оснований и т.п.. Подробное руководство по гибридизации нуклеиновых кислот можно найти в Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2, "Overview of Principles of Hybridization and the Strategy of Nucleic Acid Probe Assays," (Elsevier, New York), а также в *Current Protocols In Molecular Biology*, Ausubel *et al.*, eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2004); и Hames and Higgins (1995), *Gene Probes 1* и *Gene Probes 2*, оба IRL Press at Oxford University Press, Oxford, England в этих публикациях приведены детали, касающиеся синтеза, мечения, обнаружения и количественного анализа ДНК и РНК, включая олигонуклеотиды.

[0169] Пример жестких условий гибридизации комплементарных нуклеиновых кислот, которые содержат более 100 комплементарных остатков на фильтре в Саузерн или Нозерн блоте: 50%-ый формалин с 1 мг гепарина при 42°C, при гибридизации, выполняемой в течение ночи. Пример условий жесткой промывки: промывка 0.2x SSC, при 65°C в течение 15 минут (см. Sambrook *et al.* для описания буфера SSC; Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001). Часто промывки высокой жесткости, предшествует промывание низкой жесткости для исключения фонового сигнала зонда. Пример промывания низкой жесткости: 2x SSC при 40°C в течение 15 минут. Как правило, отношение сигнала к шуму составляет 5x (или выше) наблюдаемого для неродственного зонда, который при определенном способе анализа гибридизации указывает на обнаружение специфичной гибридизации.

[0170] "Жесткие условия промывки гибридизации" в контексте экспериментов гибридизации нуклеиновой кислоты, таких как Саузерн- и Нозерн-гибридизация, зависят от последовательности, и различаются при различных параметрах окружающей среды. Подробное руководство по гибридизации нуклеиновых кислот можно найти в Tijssen (1993) *Laboratory Techniques in*

*Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2, "Overview of Principles of Hybridization and the Strategy of Nucleic Acid Probe Assays," (Elsevier, New York); и в Hames and Higgins (1995), *Gene Probes 1* and *Gene Probes 2*, обе IRL Press at Oxford University Press, Oxford, England.

Жесткие условия гибридизации и промывки можно легко установить для любой тестовой нуклеиновой кислоты эмпирическим путем. Например, в определении жестких условий гибридизации и промывки, жесткость условий гибридизации и промывки постепенно увеличивают (например, путем увеличения температуры, уменьшения концентрации солей, увеличения концентрации детергента и/или увеличения концентрации органических растворителей, таких как формалин при гибридизации или отмывке), пока не будет достигнут выбранный набор критериев. Например, при очень жестких условиях гибридизации и промывки, условия гибридизации и промывки постепенно увеличивают, пока зонд не начнет связываться с идеально совпадающей комплементарной мишенью с соотношением сигнала и шума, составляющим, по меньшей мере, 5x уровня, наблюдаемого в случае гибридизации зонда и несовпадающей мишени.

[0171] "Очень жесткие" условия выбирают таким образом, чтобы они были эквивалентны точке температурного плавления ( $T_m$ ) конкретного зонда  $T_m$  - это температура (при определенной ионной силе и pH), при которой 50 % тестовой последовательности гибридизируются с идеально совпадающим зондом. Для целей настоящего изобретения, обычно, "очень жесткие" условия гибридизации и промывки выбирают таким образом, чтобы они были приблизительно на 5° C ниже  $T_m$  конкретной последовательности при определенной ионной силе и pH.

[0172] "Ультравысокая жесткость" условий гибридизации и промывки – это условия, при которых жесткость условий гибридизации и промывки увеличивают до тех пор, пока соотношение сигнала шуму для связывания зонда с идеально совпадающей комплементарной нуклеиновой кислотой-мишенью не будет составлять, по меньшей мере, 10x уровня, который наблюдается при гибридизации любой из несовпадающих нуклеиновых кислот-мишеней. Говорят, что нуклеиновая кислота-мишень, которая гибридизируется с зондом при таких условиях, с соотношением сигнала к шуму, составляющим, по меньшей мере, 1/2 уровня для идеально совпадающей комплементарной нуклеиновой кислоты-мишени, связывается с зондом в условиях ультравысокой жесткости.

[0173] Аналогично, еще более высокие уровни жесткости можно определить путем постепенного увеличения жесткости условий гибридизации и/или промывки подходящего гибридизационного теста. Например, в котором, жесткость условий гибридизации и промывки увеличивают, пока соотношение сигнала к шуму для связывания зонда с идеально совпадающей комплементарной нуклеиновой кислотой-мишенью не составит, по меньшей мере, 10х, 20Х, 50Х, 100Х, 500Х или более от уровня, который наблюдают при гибридизации с любой из несовпадающих нуклеиновых кислот-мишеней. Говорят, что нуклеиновая кислота-мишень, которая гибридизирует с зондом при таких условиях, с соотношением сигнала к шуму равным, по меньшей мере,  $\frac{1}{2}$  от уровня для идеально совпадающей комплементарной нуклеиновой кислоты-мишени, связывается с зондом в сверхжестких (ультра-ультра жестких) условиях.

[0174] Нуклеиновые кислоты, которые не гибридизируются друг с другом в жестких условиях, тем не менее являются по существу идентичными, если полипептиды, которые они кодируют, по существу идентичны. Это имеет место, например, если копия нуклеиновой кислоты создана с использованием максимальной вырожденности кодонов, разрешенной генетическим кодом.

#### •Уникальные субпоследовательности

[0175] В некоторых аспектах, настоящее изобретение обеспечивает нуклеиновую кислоту, которая содержит уникальную субпоследовательность в нуклеиновой кислоте, выбранной из последовательностей О-тРНК и О-РС, описанных в настоящей заявке. Уникальная субпоследовательность уникальна, по сравнению с нуклеиновой кислотой, соответствующей любой известной О-тРНК или последовательностью нуклеиновой кислоты О-РС. Выравнивание может быть выполнено с использованием, например, программы BLAST с параметрами по умолчанию. Любую уникальную субпоследовательность можно применять, например, в качестве для идентификации нуклеиновых кислот согласно настоящему изобретению или родственных нуклеиновых кислот.

[0176] Аналогично, настоящее изобретение включает полипептид, который содержит уникальную подпоследовательность в полипептиде, выбранном из последовательностей О-РС, описанных в настоящей заявке. В настоящей заявке, уникальная подпоследовательность уникальна по сравнению с полипептидом, соответствующим любой известной последовательности полипептида.

[0177] Изобретение также включает нуклеиновые кислоты-мишени, которые гибридизируются в жестких условиях с уникальным кодирующим олигонуклеотидом, который кодирует уникальную субпоследовательность в полипептиде, выбранном из последовательностей О-РС, причем указанная уникальная субпоследовательность уникальна по сравнению с полипептидом, соответствующим любому из полипептидов контроля (например, материнских последовательностей, из которых были получены синтетазы согласно настоящему изобретению, например, путем мутации). Уникальные последовательности определяют, как отмечено выше.

•Сравнение, идентичность и гомология последовательностей

[0178] Термины "идентичный" или "процент идентичности" применительно к двум или более нуклеиновым кислотам или последовательностям полипептида, относится к двум или более последовательностям или субпоследовательностям, которые являются одинаковыми или имеют указанный процент остатков аминокислот или нуклеотидов, которые совпадают, при сравнении и выравнивании с максимальным совпадением, которые определены при помощи одного из алгоритмов сравнения последовательностей, описанных ниже (или других алгоритмов, доступных специалистам) или путем визуального осмотра.

[0179] Фраза "по существу идентичный" применительно к двум нуклеиновым кислотам или полипептидам (например, ДНК кодирующую О-тРНК или О-РС, или последовательность аминокислот О-РС) относится к двум или более последовательностям или субпоследовательностям, которые имеют, по меньшей мере, приблизительно 60 %, приблизительно 80 %, приблизительно 90-95 %, приблизительно 98 %, приблизительно 99 % или более идентичных нуклеотидов или остатков аминокислот, при сравнении и выравнивании с максимальным совпадением, согласно измерению с использованием алгоритма сравнения последовательностей или путем визуального осмотра. Такие "по существу идентичные" последовательности, как обычно предполагают, являются "гомологичными", независимо от фактического происхождения. Предпочтительно, "существенная идентичность" существует в области последовательностей, которая составляет, по меньшей мере, приблизительно 50 остатков в длину, более предпочтительно по области, по меньшей мере приблизительно 100 остатков, и наиболее предпочтительно, последовательности по существу в

области, составляющей, по меньшей мере, приблизительно 150 остатков, или по всей длине двух последовательностей, которые сравнивают.

[0180] Белки и/или последовательности белков являются «гомологичными», если они получены, естественным путем или искусственно, из общего материнского белка или последовательности белка. Аналогично, нуклеиновые кислоты и/или последовательности нуклеиновой кислоты являются гомологичными, если они получены, природным или искусственным путем, из общей материнской нуклеиновой кислоты или последовательности нуклеиновой кислоты. Например, любую встречающуюся в природе нуклеиновую кислоту можно изменить любым доступным методом мутагенеза, таким образом, чтобы она включала один или более селекторных кодонов. При экспрессии такая подвергшаяся мутагенезу нуклеиновая кислота, кодирует полипептид, включающий одну или более не встречающихся в природе аминокислот. Процесс мутации может, конечно, изменить дополнительно один или более стандартных кодонов, что приведет к изменению одной или более стандартных аминокислот в конечном мутантном белке. Гомологию, как правило, определяют по подобию последовательностей между двумя или более нуклеиновыми кислотами или белками (или их последовательностями). Точный процент подобия последовательностей, который можно применять для определения гомологии, зависит от рассматриваемых нуклеиновой кислоты и белка, которые, но обычно для установления гомологии необходимо как минимум 25%-ое подобие. Более высокие уровни подобия последовательностей, например, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, или 99 % или более, могут можно применять для установления гомологии. Способы определения процента подобия последовательностей (например, BLASTP и BLASTN с использованием параметров по умолчанию) описаны в настоящей заявке и являются общедоступными.

[0181] При сравнении и определении гомологии последовательностей обычно одна последовательность выступает в роли эталонной последовательности, с которой сравнивают тестируемую последовательность. При применении алгоритма сравнения последовательности тестируемую и эталонную последовательность вводят в компьютер, задают координаты последовательностей, если необходимо, и задают параметры последовательной программы алгоритма. Затем алгоритм сравнения последовательностей

вычисляет процент идентичности последовательностей тестируемой последовательности(ей) относительно эталонной последовательности, основываясь на заданных параметрах программы.

5 [0182] Оптимальное выравнивание последовательностей для сравнения можно осуществить, например, при помощи алгоритма локальной гомологии Смита и Ватермана, *Adv. Appl. Math.* 2:482 (1981), алгоритма выравнивания гомологии Needleman и Wunsch, *J. Mol. Biol.* 48:443 (1970), метода поиска подобия  
10 Pearson and Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), и компьютерной реализации указанных алгоритмов (GAP, BESTFIT, FASTA, и TFASTA в Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI),  
15 или визуальным осмотром (смотри вообще *Current Protocols In Molecular Biology*, Ausubel *et al.*, eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2004).

20 [0183] Одним из примеров алгоритмов, которые подходят для определения процента идентичности последовательностей и сходства последовательностей, является алгоритм BLAST, который описан у Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990). Программное обеспечение для реализации BLAST-анализа находится в  
25 свободном доступе на сайте Национального центра биотехнологической информации (см. вебсайт NCBI). Данный алгоритм включает сначала идентификацию пар последовательностей с высоким количеством баллов (HSP) путем идентификации коротких слов длиной W в исследуемой  
30 последовательности, которые либо совпадают, либо удовлетворяют некоторому положительно оцениваемому пороговому количеству баллов T при выравнивании со словом такой же длины в последовательности базы данных. T обозначает  
35 порог количества баллов для соседнего слова (Altschul *et al.*, (1990) *J. Mol. Biol.*, 215:403-410). Указанные исходные совпадения соседних слов выполняют роль заправки для начала поисков для выявления более длинных HSP, содержащих их.  
40 Совпадения слов продлевают в обоих направлениях вдоль каждой последовательности до тех пор, пока общее количество баллов выравнивания продолжает увеличиваться. Суммарные баллы рассчитывают, используя,  
45 например, последовательности полинуклеотидов, параметры M (вознаграждение за совпадающую пару остатков) и N (штраф за несовпадающие остатки, всегда <0). Для последовательностей аминокислот для расчета суммарного балла применяют матрицу баллов. Расширение для совпадения слов в каждом  
50



направлении прекращают, когда: 1) совокупное количество баллов выравнивания уменьшается на количество  $X$  от своего максимального достигнутого значения; 2) совокупное количество баллов снижается до нуля или ниже вследствие накопления одного или нескольких отрицательно оцениваемых выравниваний остатков; или 3) достигнут конец одной из последовательностей. Параметры алгоритма Blast  $W$ ,  $T$  и  $X$  определяют чувствительность и скорость выравнивания. В программе Blast по умолчанию применяют следующие значения параметров: длина слова ( $W$ ) = 11, матрица подсчета баллов BLOSUM62 (см. Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

[0184] В дополнение к вычислению процента идентичности последовательностей, алгоритм BLAST также позволяет осуществить статистический анализ подобия между двумя последовательностями (см., например, Karlin и Altschul, *Proc. Natl. Acad. Sci. USA* 90:5873-5787 (1993)). Одна мера подобия, которую обеспечивает алгоритм BLAST, - это наименьшая сумма вероятности ( $P(N)$ ), которая показывает вероятность, с которой совпадение двух последовательностей нуклеотидов или аминокислот происходит случайно. Например, нуклеиновую кислоту считают подобной контрольной последовательности, если наименьшая сумма вероятности, при сравнении тестовой нуклеиновой кислоты и контрольной нуклеиновой кислоты, составляет менее приблизительно 0.1, более предпочтительно менее приблизительно 0.01, и наиболее предпочтительно менее приблизительно 0.001.

#### •Мутагенез и другие методы молекулярной биологии

[0185] Можно осуществлять различные манипуляции с полинуклеотидами и полипептидами согласно настоящему изобретению, а также применяемыми в настоящем изобретении, с использованием методов молекулярной биологии. Общие тексты, в которых описаны методы молекулярной биологии, включают: Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology*, volume 152 (1987), Academic Press, Inc., San Diego, CA; Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001 and *Current Protocols In Molecular Biology*, Ausubel *et al.*, eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2004)). В этих публикациях описаны мутагенез, использование векторов, промоторов и многие другие

вопросы, связанные с, например, получением генов, которые включают селекторные кодоны для синтеза содержащих не встречающиеся в природе аминокислоты белков, ортогональных тРНК, ортогональных синтетаз, и их пар.

5 [0186] В изобретении применяют различные типы мутагенеза, например, для того, для мутирования молекулы тРНК, для получения библиотек тРНК, чтобы наработать библиотеки синтетаз, чтобы включать селекторные кодоны, которые  
10 кодируют неприродные аминокислоты в представляющем интерес белке или полипептиде. Они включают, без ограничения: сайт-направленный, случайный точечный мутагенез, гомологичную рекомбинацию, перетасовка ДНК или другие рекурсивные методы мутагенеза, конструирование химер, мутагенез с  
15 использованием урацил-содержащих матриц, олигонуклеотид-направленный мутагенез, мутагенез фосфоротиоат-модифицированной ДНК, мутагенез с использованием содержащей разрывы двухцепочечной ДНК и тому подобные методы, или любую их комбинацию. Дополнительные подходящие способы  
20 включают: точечную репарацию несовпадений, мутагенез с использованием штаммов хозяина с нарушенной репарацией, рестрикцию-отбор и рестрикцию-очистку, делеционный мутагенез, мутагенез путем полного синтеза гена, репарацией двунитевого разрыва, и т.п.. Мутагенез, например, с применением  
25 химерных конструкторов, также включен в данное изобретение. В одном варианте реализации, мутагенез можно осуществлять, следуя известной информации о природной молекуле, либо модифицированной (измененной) или подвергшейся  
30 мутации природной молекуле молекуле, например, последовательности, сравнениям последовательности, физическим свойствам, кристаллической структуре или тому подобное.

35 [0187] Клетки-хозяева подвергают генетической модификации (например, трансформируют, трансдуцируют или трансфецируют) полинуклеотидами согласно настоящему изобретению или конструктами, которые включают  
40 полинуклеотид согласно настоящему изобретению, например, вектор согласно настоящему изобретению, который может быть, например, вектором клонирования или вектором экспрессии. Например, области, кодирующие  
45 ортогональную тРНК, ортогональную тРНК синтазу, и белок, который необходимо изменить (дериватизировать), функционально связаны с элементами контроля экспрессии гена, которые являются функциональными в желаемой  
50 клетке-хозяине. Типичные векторы содержат терминаторы

транскрипции и трансляции, последовательности инициации транскрипции и трансляции, и промоторы, подходящие для регулирования экспрессии конкретной нуклеиновой кислоты-мишени. Векторы могут включать стандартные кассеты экспрессии, содержащие по меньшей мере одну независимую последовательность терминатора, последовательности, обеспечивающие репликацию кассеты в эукариотах, или прокариотах, или и в эукариотах, и в прокариотах (например, «челночные» векторы) и маркеры отбора и для прокариотических и для эукариотических систем. Векторы подходят для репликации и/или интеграции в прокариотическую, эукариотическую, или, предпочтительно, и в эукариотическую и в прокариотическую клетку. См. Gilman and Smith, *Gene* 8:81 (1979); Roberts *et al.*, *Nature*, 328:731 (1987); Schneider *et al.*, *Protein Expr. Purif.* 6435:10 (1995); *Current Protocols In Molecular Biology*, Ausubel *et al.*, Eds., *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2004); Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001; and Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology*, volume 152 (1987), Academic Press, Inc., San Diego, CA.. Вектор может быть, например, в форме плазмиды, бактерии, вируса, «голого» полинуклеотида, или конъюгированного полинуклеотида. Векторы вводят в клетки и/или микроорганизмы стандартными способами, включая электропорацию (Fromm *et al.*, *Proc. Natl. Acad. Sci. USA* 82, 5824 (1985), инфекцию вирусными векторами, высоко скоростное баллистическое проникновение маленьких частиц, содержащих нуклеиновую кислоту, или внутри матрикса маленьких гранул или частиц, или на их поверхности (Klein *et al.*, *Nature* 327, 70-73 (1987)), и/или тому подобное.

[0188] Для сайт-направленного включения неприродных аминокислот в белки в ответ на амбер-стоп-кодон (UAG) у *E. Coli* была разработана высокоэффективная и гибкая система, содержащая единственную плазмиду. В новой системе пара супрессора тРНК<sub>тир</sub> (CUA) *M. Jannaschii* и тирозил-тРНК-синтетазы кодируются в одиночной плазмиде, которая совместима с большинством векторов экспрессии *E. coli*. Для оптимальной вторичной структуры и процессинга тРНК сконструировали моноцистронный тРНК-оперон, находящийся под контролем проК-промотора и терминатора. Введение мутированной формы *glnS*-промотора для синтетазы привел к значительному

увеличению и эффективности супрессии, и точности. Увеличение эффективности супрессии также было получено за счет множества копий гена тРНК, а также за счет специфической мутации (D286R) в синтетазе (Kobayashi *et al.* (2003),  
5 "Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion," *Nat. Struct. Biol.*, 10(6):425-432). Универсальность оптимизированной системы была также продемонстрирована очень эффективным  
10 и точным встраиванием нескольких различных не встречающихся в природе аминокислот, уникальные полезные свойства которых для изучения функции и структуры белка были предварительно доказаны.

[0189] Каталог бактерий и бактериофагов, пригодных для клонирования  
15 представлен, например, ATCC, *e.g.*, *The ATCC Catalogue of Bacteria and Bacteriophage* (1996) Gherna *et al.* (eds) опубликованный Американской коллекцией типовых культур.. Другие основные процедуры для секвенирования, клонирования  
20 и других методов молекулярной биологии и лежащая их теоретические обоснования можно также найти в Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 2001; *Current Protocols In Molecular Biology*, Ausubel *et al.*, eds., *Current*  
25 *Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2004); and in Watson *et al.* (1992) *Recombinant DNA Second Edition* Scientific American Books, NY. Кроме того, по существу, любую  
30 нуклеиновую кислоту (и фактически любую меченную нуклеиновую кислоту, стандартную или нестандартную) можно изготовить на заказ или стандартным заказать из любого из разнообразных коммерческих источников, таких как Midland  
35 Certified Reagent Company (Midland, TX), The Great American Gene Company (Ramona, CA), ExpressGen Inc. (Chicago, IL), Operon Technologies Inc. (Alameda, CA) и многие другие.

[0190] Модифицированные клетки-хозяева можно культивировать на  
40 обычных питательных средах, измененных для соответствия таким целям как, например, скрининг, активация промоторов или отбор трансформантов. Эти клетки можно культивировать в трансгенных организмах. Другие полезные ссылки,  
45 например, по изоляции клеток и культивированию (*например*, для последующей изоляции нуклеиновой кислоты) включают: Freshney (1994) *Culture of Animal Cells, a Manual of Basic Technique*, third edition, Wiley- Liss, New York and the references cited therein; Payne *et al.* (1992) *Plant Cell and Tissue Culture in Liquid Systems* John  
50

Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds) (1995) *Plant Cell, Tissue and Organ Culture*; Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) and Atlas and Parks (eds) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL.

#### ПРЕДСТАВЛЯЮЩИЕ ИНТЕРЕС БЕЛКИ И ПОЛИПЕПТИДЫ

[0191] Способы синтеза в клетке белка, содержащего не встречающуюся в природе аминокислоту в указанном положении, также являются признаком настоящего изобретения. Например, способ включает культивирование на подходящей питательной среде клетки, которая содержит кодирующую белок нуклеиновую кислоту, содержащую по меньшей мере один селекторный кодон и обеспечивающую не встречающуюся в природе аминокислоту; причем клетка также содержит: ортогональную тРНК (О-тРНК), которая функционирует в клетке и распознает селекторный кодон, и ортогональную аминоацил-тРНК синтетазу (О-РС), которая избирательно аминоацилирует О-тРНК не встречающейся в природе аминокислотой. Белок, полученный таким способом также является признаком настоящего изобретения.

[0192] В некоторых вариантах реализации О-РС аминоацилирует О-тРНК предпочтительнее, чем эндогенную тРНК в системе экспрессии. Относительное соотношение О-тРНК и эндогенной тРНК, нагружаемых О-РС, если О-тРНК и О-РС присутствуют при равных молярных концентрациях, более 1:1, предпочтительно, по меньшей мере, около 2:1, более предпочтительно 5:1, еще более предпочтительно 10:1, еще более предпочтительно 20:1, еще более предпочтительно 50:1, еще более предпочтительно 75:1, еще более предпочтительно 95:1, 98:1, 99:1, 100:1, 500:1, 1,000:1, 5,000:1 или выше.

[0193] Также согласно настоящему изобретению предложены соединения, которые включают белки, где белки содержат неприродную аминокислоту. Согласно определенным вариантам реализации белок содержит последовательность аминокислот, которая по меньшей мере на 75% идентична последовательности терапевтического белка, диагностического белка, промышленного фермента или его часть.

[0194] Композиции согласно настоящему изобретению и композиции, полученные способами согласно настоящему изобретению могут находиться в клетке. Пары О-тРНК/О-РС или отдельные компоненты согласно настоящему изобретению можно, в этом случае, применять в системе механизма трансляции хозяина, что обеспечит встраивание неприродной в природе аминокислоты в белок. Этот процесс описан в публикациях международных заявок на патент номер WO 2004/094593, поданной 16 апреля 2004 "EXPANDING THE EUKARYOTIC GENETIC CODE", и WO 2002/085923 "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS", которые включены в настоящее описание посредством ссылки. Например, если пару О-тРНК/О-РС вводят в организм, например, клетку *Escherichia coli*, это приводит к *in vivo* встраиванию неприродной аминокислоты, такой как 7-гидроксикумарин-этилглицина в белок в ответ на селекторный кодон. Не встречающаяся в природе аминокислота, которую добавляют в систему, может быть синтетической аминокислотой, такой как производное фенилаланина или тирозина, которое может быть экзогенно добавлено в среду для культивирования. Композиции согласно настоящему изобретению могут находиться в *in vitro* системе трансляции, или в *in vivo* системе (системах).

[0195] Клетка согласно настоящему изобретению обеспечивает возможность синтезировать белки, которые включают не встречающиеся в природе аминокислоты в больших, пригодных для применения количествах. В некоторых аспектах, композиция может содержать, например, по меньшей мере 10 микрограмм, по меньшей мере 50 микрограмм, по меньшей мере 75 микрограмм, по меньшей мере 100 микрограмм, по меньшей мере 200 микрограмм, по меньшей мере 250 микрограмм, по меньшей мере 500 микрограммов, по меньшей мере 1 миллиграмм, по меньшей мере 10 миллиграмм или более белка, который содержит неприродную аминокислоту, или количество, которого можно достичь с помощью *in vivo* способов синтеза белков (подробности относительно получения и очистки рекомбинантных белков приведены в настоящей заявке). В другом аспекте, белок может присутствовать в композиции в концентрации, например, по меньшей мере 10 микрограмм белка на литр, по меньшей мере 50 микрограмм белка на литр, по меньшей мере 75 микрограмм белка на литр, по меньшей мере 100 микрограмм белка на литр, по меньшей мере 200 микрограмм белка на литр, по меньшей мере 250 микрограмм белка на литр, по меньшей мере 500 микрограмм белка на литр, по меньшей мере 1 миллиграмм

приблизительно 1 нл до приблизительно 100 л). Получение больших количеств (например, больше количеств, которые обычно возможно получить другими способами, например, путем *in vitro* трансляции) белка в клетке, содержащего, по меньшей мере, одну не встречающуюся в природе аминокислоту является признаком настоящего изобретения.

[0196] Встраивание неприродной аминокислоты можно производить, например, для моделирования изменения в структуре и/или функции белка, например, для изменения размера, кислотности, нуклеофильности, водородных связей, гидрофобности, доступности сайтов-мишеней для протеаз, нацеливания на какой-либо фрагмент (например, для белкового микрочипа), встраивания меток или реакционноспособных групп, и т.д. Белки, которые содержат не встречающуюся в природе аминокислоту, могут иметь улучшенные или даже полностью новые каталитические или физические свойства. Например, следующие свойства можно изменять путем включения не встречающейся в природе аминокислоты в белок: токсичность, биораспределение, структурные свойства, спектральные свойства, химические и/или фотохимические свойства, каталитическая способность, период полураспада (например, период полураспада в сыворотке), способность к реакциям с другими молекулами, например, ковалентному или нековалентному, и т.п.. Композиции, включающие белки, которые включают, по меньшей мере, одну неприродную в природе аминокислоту, можно применять, *например*, в новых методах лечения, диагностики, в качестве каталитических ферментов, промышленных ферментов, связывающих белков (например, антител), и *например*, в исследовании структуры и функции белка. См., например, Dougherty, (2000) *Unnatural Amino Acids as Probes of Protein Structure and Function*, *Current Opinion in Chemical Biology*, 4:645-652.

[0197] Согласно некоторым аспектам изобретения композиция включает по меньшей мере один белок, содержащий по меньшей мере одну, *например*, по меньшей мере две, по меньшей мере три, по меньшей мере четыре, по меньшей мере пять, по меньшей мере шесть, по меньшей мере семь, по меньшей мере восемь, по меньшей мере девять, или по меньшей мере десять или больше неприродных аминокислот. Неприродные аминокислоты могут быть одинаковыми или различными, например, может быть 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10, или более различных участков в белке, которые включают 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10, или более различных не встречающихся в природе аминокислот. В другом аспекте,

композиция содержит белок, в котором по меньшей мере одна, но меньше, чем все, отдельные аминокислоты присутствующие в белке, представляют собой неприродные аминокислоты. Для данного белка, содержащего более одной не встречающейся в природе аминокислоты, неприродные в природе аминокислоты могут быть одинаковыми или различными (например, белок может содержать два или более различных типов неприродных аминокислот, или может включать две одинаковые неприродные аминокислоты). Для данного белка с более чем двумя неприродными аминокислотами, указанные неприродные аминокислоты могут быть одинаковыми, различными или представлять собой комбинацию множества неприродных в природе аминокислот одного вида (т.е. несколько одинаковых аминокислот) с по меньшей мере одной другой неприродной аминокислотой.

[0198] По существу любой белок (или его часть), который включает неприродную аминокислоту (и любую соответствующую кодирующую нуклеиновую кислоту, например, которая включает один или более селекторных кодонов) можно получить при помощи композиций и способов согласно настоящему изобретению. Не предпринималось попыток идентифицировать сотни тысяч известных белков, любой из которых можно модифицировать путем включения в него одной или более неприродной аминокислоты, например, путем моделирования доступных способов мутации, например, путем применения любых доступных методов мутагенеза для включения одного или более подходящих селекторных кодонов в соответствующую систему трансляции. К известным базам последовательностей известных белков относятся GenBank EMBL, DDBJ и NCBI. Другие базы можно найти без труда путем поиска в интернете.

[0199] Обычно белки, например, по меньшей мере на 60%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 99% или более идентичны любому доступному белку (например, терапевтическому белку, диагностическому белку, промышленному ферменту или их части, и тому подобное), и они содержат одну или несколько неприродных аминокислот. Примеры терапевтических, диагностических и других белков, которые можно модифицировать таким образом, чтобы они содержали одну или несколько неприродных аминокислот, можно найти в следующих Международных Публикациях: WO 2004/094593, поданной 16 апреля 2004, озаглавленной "EXPANDING THE EUKARYOTIC GENETIC CODE" и WO



2002/085923, озаглавленной "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS." и других источниках. Примеры терапевтических, диагностических и других белков, которые можно модифицировать таким образом, чтобы они содержали одну или несколько неприродных аминокислот, включают, без ограничения, например, альфа-1-антитрипсин, ангиостатин, антигемолитический фактор, антитела (более подробная информация об антителах представлена ниже), аполипопротеин, апопротеин, натрийуретический фактор предсердий, натрийуретический полипептид предсердий, пептиды предсердий, хемокины C-X-C (например, T39765, NAP-2, ENA-78, Gro-a, Gro-b, Gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG), кальцитонин, хемокины CC (например, хемоаттрактаный белок-1 моноцитов, хемоаттрактаный белок-2 моноцитов, хемоаттрактаный белок-3 моноцитов, воспалительный белок-1 альфа моноцитов, воспалительный белок-1 бета моноцитов, RANTES, I309, R83915, R91733, HCC1, T58847, D31065, T64262), лиганд CD40, лиганд C-kit, коллаген, колоний-стимулирующий фактор (CSF), комплемент-активирующий белок-78, GRO $\alpha$ /MGSA, GRO $\beta$ , GRO $\gamma$ , MIP-1 $\alpha$ , MIP-1 $\delta$ , MCP-1), эпидермальный фактор роста (EGF), эритропоэтин ("EPO"), эксфолиационные токсины A и B, фактор IX, фактор VII, фактор VIII, фактор X, фактор роста фибробластов (FGF), фибронектин, фибриноген, G-CSF, GM-CSF, глюкоцереброзидаза, гонадотропин, факторы роста, белок Хедгехофа (например, Sonic, Indian, Desert) гемоглобин, фактор роста гепатоцитов (HGF), гирудин, сывороточный альбумин человека, инсулин, инсулиноподобный фактор роста (IGF), интерфероны (например, IFN- $\alpha$ , IFN- $\beta$ , IFN- $\gamma$ ), интерлейкины (например, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12 и др.), фактор роста кератиноцитов (NIF), онкостатин M, остеогенный белок, парат-гормон, PD-ECSF, PDGF, пептидные гормоны (например, гормон роста человека), плейотропин, Белок A, Белок G, пирогенные экзотоксины A, B и C, релаксин, ренин, SCF, растворимый рецептор комплемента I, растворимый I-CAM 1, растворимые рецепторы интерлейкинов (IL-1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15), растворимый рецептор ФНО, соматомедин, соматостатин, соматотропин, стрептокиназа, суперантигены, т.е., энтеротоксины стафилококка (SEA, SEB, SEC1, SEC2, SEC3, SED, SEE), супероксиддисмутаза (SOD), токсин токсического шок-синдрома (TSST-1), тимозин-альфа-1, активатор тканевого плазминогена, фактор некроза опухоли бета (TNF beta), рецептор фактора некроза опухолей

(TNFR), фактор некроза опухоли альфа (TNF alpha), фактор роста эндотелия сосудов (VEGEF), урокиназа и многие другие.

[0200] Один класс белков, которые можно создать при помощи композиций и способов для включения неприродных аминокислот *in vivo*, описанных в настоящем документе, включает модуляторы транскрипции или их части. Пример модуляторов транскрипции включает гены и белки модуляторов транскрипции, которые модулируют рост, дифференцировку, регуляцию клеток. И т.п. Модуляторы транскрипции найдены в прокариотах, вирусах и эукариотах, включая грибы, растения, дрожжи, насекомых и животных, включая млекопитающих, что означает широкий спектр терапевтических мишеней. Очевидно, что активаторы экспрессии и транскрипции регулируют транскрипцию посредством многих механизмов, например, путем связывания с рецепторами, стимулирования каскадов передачи сигнала, регулирования экспрессии факторов транскрипции, связывания с промоторами и энхансерами, связывания с белками, которые связывают промоторы и энхансеры, раскручивания ДНК, спайсирования пре-м-РНК, полиаденилирования РНК и разрушения РНК.

[0201] Один из классов белков согласно настоящему изобретению (например, белки с одной или несколькими неприродными аминокислотами) включает биологически активные белки, такие как цитокины, молекулы воспаления, факторы роста, их рецепторы и продукты онкогенов, например, интерлейкины (например, IL-1, IL-2 и IL-8 и др.), интерфероны FGF, IGF-I, IGF-II, FGF, PDGF, TNF, TGF- $\alpha$ , TGF- $\beta$ , EGF, KGF, SCF/c-Kit, CD40L/CD40, VLA-4/VCAM-1, ICAM-1/LFA-1, гиалуриин/CD44; молекулы сигнальной трансдукции и соответствующие продукты онкогенов, например, Mos, Ras, Raf, и Met; и активаторы и супрессоры транскрипции, например, p53, Tat, Fos, Myc, Jun, Myb, Rel и рецепторы стероидных гормонов, такие как рецепторы эстрогена, прогестерона, тестостерона, альдостерона, лиганды LDL-рецепторов и кортикостерона.

[0202] Ферменты (например, промышленные ферменты) или их части по меньшей мере с одной неприродной аминокислотой также предложены согласно данному изобретению. Примеры ферментов включают, но не ограничиваются, например, амидазы, рацемазы аминокислот, ацилазы, деалогеназы, диоксигеназы, диарилпропанпероксидазы, эпимеразы, эпоксид-гидролазы, эстеразы, изомеразы, киназы, изомеразы глюкозы, гликозидазы, гликозил-

трансферазы, галопероксидазы, монооксигеназы (например, р450), липазы, лигнин-пероксидазы, нитрил-гидратазы, нитрилазы, протеазы, фосфатазы, субтилины, трансаминазы и нуклеазы.

[0203] Многие из этих белков имеются в продаже (см. например, Каталог и прайс-лист Sigma BioSciences 2002), а соответствующие последовательности белков и обычно, множество вариантов белков и генов являются известными, (см., например, Genbank). Любые из них можно модифицировать путем всраивания одной или нескольких неприродных аминокислот согласно данному изобретению, например, с изменением одного или более терапевтических, диагностических или ферментативных свойств, представляющих интерес. Примеры терапевтически значимых свойств включают: период полураспада в сыворотке, полупериод хранения, стабильность, иммуногенность, терапевтическую активность, детектируемость (например, путем включения репортерных групп (например, меток или сайтов связывания метки) в не встречающиеся в природе аминокислоты), уменьшение LD<sub>50</sub> или других побочных эффектов, способность поступать в организм через желудочный тракт (например, например возможность орального введения), или тому подобное. Примеры свойств диагностических белков включают: полупериод хранения, стабильность, диагностическую активность, детектируемость и тому подобное. Примеры подходящих свойств ферментов включают полупериод хранения, стабильность, ферментативную активность, способность продукции, или тому подобное.

[0204] Также можно модифицировать другие разнообразные белки, включая одну или несколько неприродных аминокислот при помощи соединений и способов, предложенных изобретением. Например, изобретение может включать замену одной или нескольких природных аминокислот в одном или нескольких белках-вакцинах на неприродную аминокислоту, например, в белках из инфицированных грибов, например, вида *Aspergillus*, *Candida*; бактерий, в частности, *E. coli*, которая служит моделью патогенной бактерии, а также важные для медицины бактерии, такие как *Staphylococci* (например, *aureus*) или *Streptococci* (например, *pneumoniae*); простейших, таких как споровики (например, *Plasmodia*), корненожки (например, *Entamoeba*) и жгутиковые (*Trypanosoma*, *Leishmania*, *Trichomonas*, *Giardia*, др.); вирусы. Такие как (+)РНК-вирусы (примеры включают поксвирусы например, вирус осповакцины пикорновирусы, например полиомиелита; тогавирусы, например, краснухи; флавивирсы, например, вирус

гепатита С (HCV); и коронавирусы), вирусы (-)РНК (например, рабдовирусы, например, вирус везикулярного стоматита (VSV); парамиксовирусы, например, респираторно-синцитиальный вирус (RSV); ортомиксовирусы, например, грипп; буньявирусы; и аренавирусы), вирусы с двухцепочечной ДНК (реовирусы, например), вирусы с РНК, превращающейся в ДНК, то есть, ретровирусы, например, ВИЧ (HIV) и вирус Т-клеточного лейкоза человека (HTLV), и определенные вирусы с ДНК ,превращающейся в РНК, такие как вирус гепатита В.

[0205] Белки, важные для сельского хозяйства, такие как белки устойчивости к насекомым (например, Cry-белок), ферменты синтеза крахмала и липидов, токсины растений и насекомых, белки устойчивости к токсинам, белки детоксикации микотоксина, ферменты роста растений (например, Рибулоза 1,5-бифосфат карбоксилаза/оксигеназа, "РУБИСКО"), липооксигеназа (LOX), и фосфоенолпируват карбоксилаза (PEP) также являются подходящими мишенями для модификации с применением не встречающейся в природе аминокислоты.

[0206] В некоторых вариантах реализации белок или полипептид, представляющий интерес, (или его фрагмент) в способах и/или композициях согласно настоящему изобретению кодирует нуклеиновая кислота. Как правило, нуклеиновая кислота содержит по меньшей мере один селекторный кодон, по меньшей мере два селекторного кодона, по меньшей мере три селекторного кодона, по меньшей мере четыре селекторного кодона, по меньшей мере пять селекторных кодонов, по меньшей мере шесть селекторных кодонов, по меньшей мере семь селекторных кодонов, по меньшей мере восемь селекторных кодонов, по меньшей мере девять селекторных кодонов, десять или больше селекторных кодонов.

[0207] Гены, кодирующие белки или полипептиды, представляющие интерес, можно подвергнуть мутагенезу при помощи методов, хорошо известных в одной области техники и описанных здесь в разделе «Мутагенез и другие методы молекулярной биологии», с включением, например, одного или более селекторных кодонов для встраивания не встречающейся в природе аминокислоты. Например, нуклеиновую кислоту для представляющего интерес белка подвергают мутагенезу с целью включения одного или нескольких селекторных кодонов, обеспечивая встраивания одной или нескольких неприродных аминокислот. Данное изобретение включает любой такой вариант, например, мутантные варианты любого белка, например, включающие по

меньшей мере одну неприродную аминокислоту. Аналогично, данное изобретение также включает соответствующие нуклеиновые кислоты, т.е., любую нуклеиновую кислоту с одним или несколькими селекторными кодонами, которые кодируют одну или несколько неприродных аминокислот.

[0208] Чтобы получить белок, включающий неприродную аминокислоту, можно использовать клетку-хозяина и организмы, которые адаптированы для включения неприродной аминокислоты *in vivo* посредством пар ортогональных тРНК/РС. Клетки-хозяева подвергают генетической модификации (например, трансформации, трансдукции или трансфекции) одним или более векторами, которые экспрессируют ортогональную тРНК, ортогональную тРНК синтетазу, и вектор, который кодирует белок, подлежащий модификации. Все эти компоненты могут находиться в одном векторе, или каждый может быть в отдельном векторе, либо два компонента могут быть в одном векторе, а третий компонент - во втором векторе. Вектор может иметь, например, форму плазмиды, бактерии, вируса, «голого» полинуклеотида или конъюгированного полинуклеотида.

#### •Определение полипептидов по иммунореактивности

[0209] Поскольку полипептиды согласно настоящему изобретению обеспечивают разнообразные новые последовательности полипептидов (например, полипептиды, содержащие неприродные аминокислоты в случае белков, синтезированных в системе трансляции, представленной здесь, или, например, в случае новых синтетаз, новые последовательности стандартных аминокислот), полипептиды также обеспечивают новые структурные характеристики, которые можно распознавать, например, при помощи иммунологического анализа. Получение антисыворотки, которая специфично связывает полипептиды согласно настоящему изобретению, а так же полипептиды, которые связываются такой антисывороткой, являются признаком настоящего изобретения. Термин "антитело" в настоящей заявке включает, без ограничения, полипептид, по существу кодируемый геном иммуноглобулина или генами иммуноглобулина, или фрагментов таких генов, которые специфично связывают и распознают анализируемое вещество (антиген). Примеры включают антитела: поликлональные, моноклональные, химерные, одноцепочечные, и т.п. Фрагменты иммуноглобулинов, включая Fab-домены и фрагменты, полученные с помощью библиотеки экспрессии, включая бактериофаговый диспле, также

включены в термин "антитело" в том значении в котором он употребляется в настоящей заявке. См., например, Paul, *Fundamental Immunology*, 4th Ed., 1999, Raven Press, New York, по структуре антител и терминологии..

[0210] Для того, чтобы получить антисыворотки для применения в иммунологическом анализе, нарабатывают один или более иммуногенных полипептидов и очищают их, как описано выше. Например, рекомбинантный белок можно нарабатывать в рекомбинантной клетке. Инбредная линия мышей (применяемая в этом анализе, потому что результаты более воспроизводимы благодаря фактической генетической идентичности мышей) иммунизируют с применением иммуногенного белка(ов) в комбинации со стандартным адъювантом, таким как адъювант Фрейнда, по стандартному протоколу иммунизации мыши (см., например, Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York, содержащую стандартное описание получения антител, форматов иммунологического анализа и условий, которые можно применять для определения специфической иммунореактивности. Дополнительные подробности относительно белков, антител, антисыворотки, и т.д. можно найти в публикациях международных заявок на патент: WO 2004/094593, озаглавленной "EXPANDING THE EUKARYOTIC GENETIC CODE;" WO 2002/085923, озаглавленной "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS;" WO 2004/035605, озаглавленной "GLYCOPROTEIN SYNTHESIS;" и WO 2004/058946, озаглавленной "PROTEIN ARRAYS."

#### ПРИМЕНЕНИЕ О-тРНК И О-РС И ПАР О-тРНК/О-РС

[0211] Композиции согласно настоящему изобретению и композиции, полученные способами согласно настоящему изобретению могут находиться в клетке. Пары О-тРНК/О-РС или отдельные компоненты согласно настоящему изобретению можно, кроме того, задействовать в механизмах трансляции системы хозяина, что приводит к встраиванию неприродной аминокислоты в белок. Этот процесс описан в международной публикации номер: WO 2002/085923 Schultz, *et al.*, озаглавленной "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS," которая включена в настоящую заявку посредством ссылки. Например, когда пару О-тРНК/О-РС вводят в хозяина, например, *Escherichia coli* или дрожжи, указанная пара вызывает *in vivo* встраивание неприродной аминокислоты, которую можно экзогенно добавлять в среду для выращивания, в белок,

когда пару О-тРНК/О-РС вводят в хозяина, например, *Escherichia coli* или дрожжи, указанная пара вызывает *in vivo* встраивание неприродной аминокислоты, которую можно экзогенно добавлять в среду для выращивания, в белок, например, тестовый белок миоглобин или терапевтический белок, в ответ на селекторный кодон, например, нонсенс амбер-кодон. Композиция согласно настоящему изобретению может находиться в *in vitro* системе трансляции, или в клеточной *in vivo* системе (ах). Белки с неприродной аминокислотой можно применять в разнообразных целях. Например, неприродный фрагмент, встроенный в белок может, служить мишенью для широкого диапазона модификаций, например, перекрестных связей с другими белками, с небольшими молекулами, такими как метки или красители, и/или биомолекулами. С такими модификациями, встраивание неприродной аминокислоты может привести к улучшению терапевтических белков, и его можно применять для улучшения каталитической функции ферментов. В некоторых аспектах, встраивание и последующая модификация не встречающейся в природе аминокислоты в белке могут облегчать исследование структуры белка, взаимодействия с другими белками, и т.п..

## НАБОРЫ

[0212] Наборы также представляют собой признак настоящего изобретения. Например, предусмотрен набор для получения в клетке белка, который содержит по меньшей мере одну не встречающуюся в природе аминокислоту, причем набор содержит по меньшей мере один контейнер, содержащий последовательность полинуклеотида, кодирующую О-тРНК, и/или О-тРНК, и/или последовательность полинуклеотида, кодирующую О-РС, и/или О-РС. В одном варианте реализации, набор, дополнительно содержит неприродную аминокислоту 7-гидроксикумарин-этилглицин. В другом варианте, набор дополнительно содержит инструкции по получению белка и/или клетки-хозяина.

## ПРИМЕРЫ

[0213] Следующие ниже примеры приведены исключительно с целью иллюстрации, но не ограничения описанного в настоящей заявке изобретение. Одно из направлений будет распознавать разнообразные некритические параметры, которые можно изменять, не выходя за пределы описанного в заявке

области техники, и должны быть включены в общие тенденции и сферу применения данного приложения и в рамки прилагаемой формулы изобретения.

## ПРИМЕР 1

### Общие методологии

[0214] Все химические вещества были получены из коммерческих источников, их применяли без дальнейшей очистки. Спектры  $^1\text{H}$  ЯМР регистрировали на многоядерном спектрофотометре Bruker AMX-400 (Bruker BioSpin GmbH, Rheinstetten/Karlsruhe, Германия) с химическими сдвигами, указываемыми относительно тетраметилсилана. Масс-спектры белков были получены в Центре Scripps при помощи масс-спектрометрии (La Jolla, CA).

### ПЛАЗМИДЫ И ЛИНИИ КЛЕТОК

[0215] Плазмида pBK-lib5 кодирует библиотеку мутантов тирозил-тРНК-синтетазы *M. Jannaschii* (MjTyrRS), в которых остатки Tyr32, Leu65, Phe108, Gln109, Asp158 и Leu162 рандомизированы; кроме того, His70 в результате мутации заменен на Gly, а Ala67 в результате мутации заменен на Ala, либо на Gly. Плазмида pREP(2)/YC кодирует *MjtRNA<sub>CUA</sub><sup>Tyr</sup>*, ген хлорамфеникол-ацетилтрансферазы (CAT) с кодоном TAG в остатке 112, ген GFP, находящийся под контролем промотора T7, мутантную T7-РНК-полимеразу с кодоном TAG на остатках 1 и 107, и маркер Tet<sup>r</sup>. Плазмида pLWJ17B3 кодирует *MjtRNA<sub>CUA</sub><sup>Tyr</sup>* под контролем промотора *lpp* и терминатора *rrnC*, ген барназы (с тремя амбер-кодонами в остатках 2, 44 и 65) под контролем промотором *ara* и маркер Amp<sup>r</sup>. Плазмида pBAD/JYAMB-4TAG-Myo кодирует ген миоглобина с промотором арабинозы и терминатором *rrnB*, *MjtRNA<sub>CUA</sub><sup>tyr</sup>* с промотором *lpp* и терминатора *rrnC*, и маркер резистентности к тетрациклину. Штамм *E. coli* GeneHog®-Fis, F<sup>-</sup> *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*) 80/*lacZ* $\Delta$ M15  $\Delta$ *lacX74* *recA1* *endA1* *araD139*  $\Delta$ (*ara-leu*)7697 *galU* *galK* *lrpsL*(Str<sup>R</sup>) *nupG*, *fis::Tn7* (DHFR) использовали для исследования того, что экспрессируемые белки содержат неприродную аминокислоту L-(7-гидроксикумарин-4-ил)-этилглицин.



## ПРИМЕР 2

Синтез неприродной аминокислоты L-(7-гидроксикумарин-4-ил)-этилглицин

[0216] Неприродную аминокислоту L-(7-гидроксикумарин-4-ил)-этилглицин (которую также называют 7-гидроксикумарин-этилглицин; см. ФИГ.1, структура 1) синтезировали следующим образом: сначала превращали  $\alpha$ -бензиловый эфир N- $\alpha$ -Cbz-L-глутаминовой кислоты в  $\beta$ -кето-эфир с боковой цепью, который затем использовали в реакции с резорцинолом в метанесульфоновой кислоте (реакция фон Пекманна; Brun *et al.*, *Angew. Chem., Int. Ed.*, (2004) 43:3432-3436), в результате чего получали L-(7-гидроксикумарин-4-ил)-этилглицин. Эти этапы более подробно описаны ниже.

## ЭТАП 1 – Синтез Этил-магния-малоната

[0217] К перемешиваемому водному раствору 1,5 М хлорида магния и 3 М этил-калия малоната добавляли пять объемов изопропанола. После фильтрации этил-магниий-малонат высушивали под вакуумом.

ЭТАП 2 – Синтез 7-этил-эфира (2S)-2-бензилоксикарбониламино-5-оксо-гептанедиевой кислоты

[0218] Z-Glu-Obzl ( $\alpha$ -бензиловый эфир N- $\alpha$ -Cbz-L-глутаминовой кислоты) (5г, 13,46 ммоль) растворяли в 50 мл сухого ТГФ при комнатной температуре. Добавляли N,N'-карбонил-диимидазол медленно и затем перемешивали в течение 2 ч при комнатной температуре. Добавляли этил-магниий малонат (0,55 эк) добавляли, и полученную смесь перемешивали при комнатной температуре в течение ночи. Продукт экстрагировали эфиром и промывали 10% NaHCO<sub>3</sub>, водой и соляным раствором. Остаток, содержащий 3, очищали путем флэш-хроматографии на силикагеле (этилацетат: гексаны 50:50) и концентрировали на вращательном испарителе с целью получить белое твердое вещество с 40% выходом. <sup>1</sup>H-ЯМР (400 МГц, CDCl<sub>3</sub>):  $\delta$  1,26 (t, J = 7,2, 3H), 1,9-2,01 (m, 1H), 2,1-2,28 (m, 1H), 2,5-2,7 (m, 2H), 3,37 (s, 2H), 4,18 (q, J = 7,2, 2H), 4,41 (s, 1H), 5,11 (s, 2H), 5,18 (s, 2H), 5,44 (d, J = 8, 1H), 7,25-7,41 (m, 10H). LC-MS (ESI) вычисленные для C<sub>24</sub>H<sub>2</sub>NO<sub>7</sub> (M+ H<sup>+</sup>) 442,18, obsd. 442,0.

ЭТАП 3 – Синтез L-(7-гидроксикумарин-4-ил)-этилглицина (ФИГ.1, структура 1)

[0219] 1-бензил-эфир 7-этил-эфира (2S)-2-бензилоксикарбониламино-5-оксо-гептанедиевой кислоты из предыдущего этапа (2,35 г, 5 ммоль) медленно

добавляли к резорцинолу (3 г, 25 ммоль) в метансульфониевой кислоте (20 мл) и перемешивали в течение двух часов при комнатной температуре. Затем к данной смеси добавляли пять объемов эфира и охлаждали до  $-30^{\circ}\text{C}$ . Преципитат промывали холодным эфиром, растворяли в воде, фильтровали и лиофилизировали. L-(7-гидроксикумарин-4-ил)-этилглицин очищали путем обращенно-фазовой ЖХВД с получением коричневого твердого вещества с 50% выходом (колонок YMC AA12S052503WT, скорость потока 12 мл/мин, от 10% до 90%  $\text{CH}_3\text{CN}$ , 0,1% ТФА (вес./об.) в воде, на протяжении 12 мин).  $^1\text{H}$ -ЯМР (ДМСО):  $\delta$  2,06 (s, 2H), 2,70-2,90 (m, 2H), 3,82-4,0 (m, 1H), 6,06 (s, 1H), 6,75 (dd,  $J = 8,8, 2,4$ , 1H), 6,82 (d,  $J = 2,4$ , 1H), 7,62 (d,  $J = 8,8$ , 1H), 8,17 (s, 3H). LC-MS (ESI) расч. для  $\text{C}_{13}\text{H}_{14}\text{NO}_5$  ( $\text{M} + \text{H}^+$ ) 264,08 Да, эксп. 264,0 Да.

### ПРИМЕР 3

Генетический отбор мутантной синтетазы, специфичной к 7-гидроксикумарин-этилглицину

[0220] Для селективного встраивания 7-гидроксикумарин-этилглицина в определенные сайты в белках *E.coli* применяли пару мутантной амбер-супрессорной тирозил--тРНК из *Methanococcus jannaschii* ( $\text{Mj}t\text{RNA}^{\text{Tyr}}_{\text{CUA}}$ )/тирозил-тРНК-синтетазы ( $\text{MjTyrRS}$ ), которая уникально включает 7-гидроксикумарин-этилглицин в ответ на селекторный кодон TAG.

### КОНСТРУИРОВАНИЕ БИБЛИОТЕКИ

[0221] Чтобы разместить крупную боковую цепь кумарина неприродной аминокислоты 7-гидроксикумарин-этилглицин, сделали библиотеку pBK-lib5  $\text{MjTyrRS}$ , в которой His70 в тирозил-тРНК-синтетазе дикого типа *Methanococcus jannaschii* (ФИГ.9, SEQ ID NO: 2 и 3) в результате мутагенеза заменен на Gly, а Ala67 фиксировали либо как Ala, либо как Gly, что позволило увеличить размер активного центра. Были рандомизированы шесть остатков: Tyr-32, Leu-65, Phe-108, Gln-109, Asp-158 и Leu-162, находящиеся вблизи связанного тирозина (Kobayashi *et al.*, *Nat. Struct. Biol.*, (2003) 10:425-432; Zhang *et al.*, *Protein Sci.*, (2005) 14:1340-1349). Библиотека pBK-lib5 включала  $2 \times 10^9$  независимых клонов Tyr-PC, и ее конструировали при помощи методов ПЦР.

## ПОЛОЖИТЕЛЬНЫЙ/ОТРИЦАТЕЛЬНЫЙ ОТБОР

[0222] Затем эту библиотеку подвергали циклам как положительного, так и отрицательного отбора. При положительном отборе выживание клетки зависело от супрессии амбер-кодона, введенного в разрешающем сайте в ген хлорамфеникол-ацетилтрансферазы (CAT), при этом клетки, трансформированные одновременно рBK-lib и *MjtRNA*<sup>Tyr<sub>CUA</sub></sup>, выращивали в присутствии 1мМ неприродной аминокислоты (НПАК) и хлорамфеникола (Xie and Schultz, *Methods* (2005) 36:227-238). Прошедшие положительный отбор клоны затем подвергали трансформации в клетках, содержащих *MjtRNA*<sup>Tyr<sub>CUA</sub></sup> и ген, кодирующий токсичный белок барназу с тремя амбер-мутациями, введенными в разрешающие сайты. Эти клетки выращивали в отсутствие НПАК, что обеспечивало устранение любых клонов, которые утилизируют эндогенные аминокислоты (отрицательный отбор). Процедуры положительного и отрицательного отбора описаны более подробно ниже.

[0223] *Положительный отбор* – в качестве штамма-хозяина для положительного отбора применяли *E. coli* DH10B, содержащие плазмиду рREP(2)/YC. Клетки трансформировали библиотекой рBK-lib5, вводили в SOC в течение 1 часа, промывали два раза минимальной средой с глицерином, и лейцином (GMMML), затем высевали на чашку с GMMML-агаром, дополненным канамицином, хлорамфениколом, тетрациклином и 7-гидроксикумарин-этилглицином в концентрациях 50 мкг/мл, 60 мкг/мл, 15 мкг/мл и 1 мМ, соответственно. Чашки инкубировали при 37<sup>0</sup>С в течение 60 часов, выжившие клетки соскабливали, эксрагировали и очищали плазмидные ДНК путем гель-электрофореза.

[0224] *Отрицательный отбор* – затем ДНК рBK-lib5 трансформировали в электрокомпонентные клетки, содержащие плазмиду отрицательного отбора рLWJ17B3, вводили в течение 1 часа в SOC, и затем высевали на чашки на LB-агар, содержащий 0,2% арабинозы, 50 мкг/мл ампициллина и 50 мкг/мл канамицина. Затем чашки инкубировали при 37<sup>0</sup>С в течение 8-12 часов, и из выживших клонов экстрагировали ДНК рBK-lib5, как описано выше.

[0225] Затем библиотеку подвергали следующему циклу положительного отбора, затем – отрицательному отбору и последнему циклу положительного отбора (с хлорамфениколом при 70 мкг/ил). На этой стадии 96 отдельных клонов отобрали и суспендировали в 50 мкл GMMML в 96-луночном планшете и

переносили колонии мутодом реплик на два набора планшетов(чашек) с GMML. Один набор планшето́в с GMML-агаром дополнили тетрациклином (15 мкг/мл), канамицином (50 мкг/мл) и хлорамфениколом в концентрациях 60, 80, 100 и 120 мг/мл с 1 мМ 7-гидроксикумарин-этилглицина. Другой набор планшетов был идентичен первому, но не содержал 7-гидроксикумарин-этилглицина, а применяемые концентрации хлорамфеникола были 0, 20, 40 и 60 мкг/мл. Через 60 часов инкубации при 37<sup>0</sup>С, было выявлено, что один клон выжил при 100 мкг/мл хлорамфеникола в присутствии 1 мМ 7-гидроксикумарин-этилглицина, но только при 20 мкг/мл хлорамфеникола в отсутствие 7-гидроксикумарин-этилглицина. Таким образом, после трех циклов положительного отбора и двух циклов отрицательного отбора эта библиотека сдала один клон, который рос при 100 мкг/мл хлорамфеникола в присутствии 1 мМ 7-гидроксикумарин-этилглицина, но только при 20 мкг/мл хлорамфеникола в отсутствие 7-гидроксикумарин-этилглицина.

[0226] Этот клон, названный CouRS-D8, секвенировали и обнаружили, что он содержит следующие восемь мутаций Tyr32Glu, Leu65His, Ala67Gly, His70Gly, Phe108Tyr, Gln109His, Asp158Gly и Leu162Gly (см. ФИГ. 9 и SEQ ID NOs: 4 и 5). Последовательность мутантного полипептида синтета́зы, представленная в SEQ ID NO: 4, на ФИГ. 9 положения мутаций показаны жирными буквами. Белок мутантной синтета́зы также представлен в таблице ниже.

	Положение аминокислоты (кодон)								EQ ID NO:
	32	65	67	0	108	109	158	162	
<i>Mj</i> тирозил- тРНК- синтета́за									
Дикий тип	Tyr (TAC)	Leu (TTG)	Ala (GCT)	is (CAC)	Phe (TTC)	Gln (CAG)	Asp (GAT)	Leu (TTA)	3)
CouRS-D8 мутант	Glu (GAG)	His (CAT)	Gly (GGT)	ly (GGC)	Tyr (TAT)	His (CAT)	Gly (GGG)	Gly (GGT)	5)

[0227] Четыре остатка в результате мутации были заменены на глицин, что, вероятно, обеспечило образование достаточного пространства для включения остатка 7-гидроксикумарин-этилглицина. Tyr32 и Asp158, которые образуют водородные связи для связывания тирозина в ферменте дикого типа

также мутировали, на что указывает и потеря активности в отношении тирозина (Kobayashi *et al.*, *Nat. Struct. Biol.*, (2003) 10:425-432; Zhang *et al.*, *Protein Sci.*, (2005) 14:1340-1349).

[0050] На ФИГ.8 показаны предсказанная структурная модель активного центра мутантной *MjTyrRS* (CouRS-D8). Эту модель получили с использованием структуры *MjTyrRS* дикого типа (код базы данных 1J1U). Субстрат тирозин заменили на L-(7-гидроксикумарин-4-ил)-этилглицин; показаны следующие мутации: Tyr32Glu, Leu65His, Ala67Gly, His70Gly, Phe108Tyr, Gln109His, Asp158Gly и Leu162Gly. Затем структуру оптимизировали методом минимизации энергии при помощи программного обеспечения InsightII® software (Accelrys Software, Inc.®, San, Diego, CA, USA).

#### ПРИМЕР 4

Экспрессия и характеристика модельного белка мутантного миоглобина, содержащего 7-гидроксикумарин-этилглицин

[0229] Для определения эффективности и точности встраивания 7-гидроксикумарин-этилглицина в белки, заменяли селекторный амбер-кодон либо в положении 4 (Ser-4), либо в положении 37 (His-37) в миоглобине кашалота (см. ФИГ.3), содержащем С-концевую гистидиновую метку (His<sub>6</sub>). Экспрессию белка осуществляли в присутствии выбранной синтетазы CouRS-D8 и *MjtRNA*<sup>Tyr</sup><sub>CUA</sub> у *E. coli*, выращенной либо на минимальной среде, или в среде, дополненной 1 мМ 7-гидроксикумарин-этилглицина. В качестве отрицательного контроля осуществляли также экспрессию белков в отсутствие 7-гидроксикумарин-этилглицина. Получение и анализ белка описаны более подробно далее.

#### ПОЛУЧЕНИЕ МУТАНТНОГО БЕЛКА

[0230] Плазмиды pBAD/JYAMB-4TAG-Myo или pBAD/JYAMB-37TAG-Myo подвергали независимой совместной трансформации вектором экспрессии pBK CouRS-D8 в клетки GeneHog®-Fis *E. coli*. Клетки размножали в среде 2YT (5 мл), дополненной канамицином (50 мкг/мл) и тетрациклином (15 мкг/мл). Начальную культуру (1 мл) использовали для инокуляции 100 мл жидкой GMMML или 2YT, дополненной соответствующими антибиотиками и 1 мМ 7-гидроксикумарин-этилглицином. Затем клетки выращивали при 37°C до ОП<sub>600</sub> 0,5, и индуцировали экспрессию белков путем добавления 0,2% арабинозы. По прошествии еще 4-12

часов клетки собирали путем центрифугирования. Мутанты миоглобина TAG4 и TAG37 очищали путем афинной хроматографии на Ni-NTA в нормальных условиях. Затем образцы обессоливали. Выход мутантного миоглобина TAG4 и TAG37 в любой среде составлял 2 мг/л. Для сравнения, выход миоглобина дикого типа В близких условиях составляет 5 мг/л.

#### МАСС-СПЕКТРОМЕТРИЧЕСКИЙ АНАЛИЗ БЕЛКОВ

[0231] Очищенный мутантный миоглобин TAG4 подвергали масс-спектрометрическому анализу в модификации ESI (ионизация методом электроспрея). ESI-масс-спектрометрический анализ мутантного миоглобина дал наблюдаемую в среднем массу  $M_{avg}$  18511 Да, что хорошо согласуется с расчетной массой  $M_{avg}$  18512 Да (см. ФИГ.6). Поскольку молекулы, соответствующие включению природной аминокислоты в масс-спектрах ESI не были обнаружены при пороге всерткии 5%, из соотношения сигнал-шум из масс-спектра белка получили минимальную чистоту включения для 7-гидроксикумарин-этилглицина, равную 95%. Порог детектирования задан таким образом, чтобы позволял детектировать только такие молекулы, количество которых составляет 5% основного пика. Поскольку наблюдали только основной пик, чистота мутантного белка составляет > 95%.

#### АНАЛИЗ БЕЛКОВ ПУТЕМ РАЗДЕЛЕНИЯ МЕТОДОМ ЭЛЕКТРОФОРЕЗА В ПААГ-ДСН

[0232] Анализ очищенного белка методом электрофореза в ПААГ-СДС показал, что полноразмерный белок экспрессировался только в присутствии 7-гидроксикумарин-этилглицина (см. ФИГ.5). Кроме того, мутантный миоглобин TAG4, включающий генетически кодируемый 7-гидроксикумарин-этилглицин, не демонстрирует флюоресценции. Для визуализации флюоресценции применяли флюоресцентный сканер STORM<sup>®</sup> (Molecular Dynamics/GE<sup>®</sup> Healthcare Life Sciences, Piscataway, NJ).

## ПРИМЕР 5

Применение генетически кодируемого 7-гидроксикумарин-этилглицина в качестве зонда для определения конфигурации белков

5 [0233] Чтобы продемонстрировать одно из потенциальных применений 7-гидроксикумарин-этилглицина, эту аминокислоту применяли в качестве зонда зависимой от мочевины денатурации холомиоглобина. Поскольку флюоресценция кумарина чувствительна к полярности растворителя, интенсивность его флюоресценции должна коррелировать с местным раскручиванием белка вблизи 7-гидроксикумарин-этилглицина. Структура миоглобина состоит из восьми спиралей (A-H), соединенных короткими петлями и повернутыми. (Delli Castelli *et al.*, *Protein Sci.*, (2002) 11:2273-2276). Ser4 в спирали A и His37 в спирали C (См. ФИГ.3; оба остатка в значительной степени контактируют с растворителем и незначительно взаимодействуют с другими близлежащими остатками), причем какждый независимо путем мутации заменен на селекторный кодон TAG, что обеспечивает к генетически запрограммированное встраивание *in vivo* 7-гидроксикумарин-этилглицина в присутствии специфичной ортогональной пары О-ТРНК/О-РС.

25 [0234] Как показано на ФИГ.4, индуцируемое мочевиной раскручивание кривых дикого типа, Ser4 → 7-гидроксикумарин-этилглицин и His37 → 7-гидроксикумарин-этилглицин хемоглобина, которое, как показал мониторинг методом кругового дихроизма, практически идентично, что указывает на то, что включение 7-гидроксикумарин-этилглицина и в петлю A, и в петлю C не вызывает значительного нарушения стабильности белка. При концентрации мочевины 2 М интенсивность флюоресценции 7-гидроксикумарин-этилглицина в положении 4 в холомиоглобине увеличивается на 30% (и остается на этом уровне при содержании мочевины от 2 М до 5 М), что указывает на нарушения в данном участке белка (Zhang *et al.*, *Protein Sci.*, (2005) 14:1340-1349). Мутантный миоглобин с 7-гидроксикумарин-этилглицином в остатке 37, напротив, демонстрирует небольшое изменение интенсивности флюоресценции при 2 М мочевины, но претерпевает аналогичное повышение флюоресценции при 3М мочевины. В соответствии с этими результатами, проведенное ранее исследование методом ЯМР показало, что при концентрации мочевины выше 2,2 М, спирали A и B сильно нарушены, как показало исчезновение короткого и среднего диапазона ОЯЭ в этом участке, тогда

как спирали C, D и F разворачиваются позже, при концентрации мочевины выше 3,0 М (Delli Castelli *et al.*, *Protein Sci.*, (2002) 11:2273-2276). Таким образом, очевидно, что аминокислота 7-гидроксикумарин-этилглицин является сайт-специфическим зондом конформационных изменений белка, в отличие от кругового дихроизма, который указывает на глобальные изменения, усредненные по всей структуре.

[0235] Чувствительность 7-гидроксикумарин-этилглицина к pH также тестировали, как показано на ФИГ.7. Спектр поглощения 60 мкМ кумариновой аминокислоты 7-гидроксикумарин-этилглицина измеряли в 100 мМ натрий-фосфатном буфере. Поглощение на 360 нм увеличивается при повышении pH. По данным этих измерений рассчитали pKa, равное 7,8.

[0236] Чувствительность 7-гидроксикумарин-этилглицина к pH (См. ФИГ.7) и к полярности растворителя (Zinsli, *J. Photochem* (1974) 3:55-69) должна сделать его полезным зондом для многих биологических исследований, как *in vitro*, так и *in vivo* (Brun *et al.*, *Angew. Chem., Int. Ed.*, (2004) 43:3432-3436). Например, аминокислоту 7-гидроксикумарин-этилглицин можно применять для отслеживания молекулярных взаимодействий или конформационных изменений в белках, или топологии мембраносвязанных белков. Кроме того, поскольку 7-гидроксикумарин-этилглицин имеет pKa 7,8, которое можно искусственно изменять заменой кумаринового кольца (Sun *et al.*, *Bioorg. Med. Chem. Lett.*, (1998) 8:3107-3110), он должен стать полезным зондом pH в органеллах, и pH-зависимых процессов в клетке. Кроме того, в возбужденном состоянии 7-гидроксикумарин-этилглицин является сильной фотокислотой (Cohen and Huppert, *Phys. Chem. A* (2001) 105:7157-7164) и может облегчать исследование процессов переноса протонов в белках.

[0051] Хотя в целях прояснения и понимания приведенное здесь изобретение было описано с указанием некоторых деталей для специалистов в данной области техники при чтении данного раскрытия будет очевидно, что можно осуществить различные изменения формы и деталей, не выходя за рамки данного изобретения. Все публикации, патенты, заявки на патент и/или документы, цитируемые в данной заявке, полностью включены в описание посредством ссылки для любых целей в такой же степени, как если бы каждая отдельная публикация, патент, заявка на патент, и/или документ были указаны по отдельности.



## ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

5	<110>	ЗЕ СКРИПС РЕСЕЧ ИНСТИТУТ ВАНГ, Джаангюн КСИ, Джаанминг ШУЛЬЦ, Питер Г	
	<120>	ГЕНЕТИЧЕСКИ КОДИРУЕМЫЕ ФЛЮОРЕСЦЕНТНЫЕ КУМАРИНОВЫЕ АМИНОКИСЛОТЫ	
	<130>	54-002020US	
	<140>	US 11/805,247	
10	<141>	2007-05-21	
	<160>	5	
	<170>	патент в версии 3.4	
15	<210>	1	
	<211>	77	
	<212>	RNA	
	<213>	Искусственная последовательность	
20	<220>		
	<223>	мутантная супрессорная тирозил-ТРНК CUA	
25	<400>	1 ccggcgguag uucagcaggg сагаасggcg gacucuaaaи ccgcauggcg cugguисааа uiccggcccgс cggacca	60 77
	<210>	2	
	<211>	306	
30	<212>	PRT	
	<213>	Methanococcus jannaschii	
	<400>	2	
35		Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser 1 5 10 15	
		Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Tyr 20 25 30	
40		Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln 35 40 45	
		Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile 50 55 60	
45		Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp 65 70 75 80	
50			

Glu Ile Arg Lys Ile Gly Asp Tyr Asn Lys Lys Val Phe Glu Ala Met  
 85 90 95  
 Gly Leu Lys Ala Lys Tyr Val Tyr Gly Ser Glu Phe Gln Leu Asp Lys  
 100 105 110  
 Asp Tyr Thr Leu Asn Val Tyr Arg Leu Ala Leu Lys Thr Thr Leu Lys  
 115 120 125  
 Arg Ala Arg Arg Ser Met Glu Leu Ile Ala Arg Glu Asp Glu Asn Pro  
 130 135 140  
 Lys Val Ala Glu Val Ile Tyr Pro Ile Met Gln Val Asn Asp Ile His  
 145 150 155 160  
 Tyr Leu Gly Val Asp Val Ala Val Gly Gly Met Glu Gln Arg Lys Ile  
 165 170 175  
 His Met Leu Ala Arg Glu Leu Leu Pro Lys Lys Val Val Cys Ile His  
 180 185 190  
 Asn Pro Val Leu Thr Gly Leu Asp Gly Glu Gly Lys Met Ser Ser Ser  
 195 200 205  
 Lys Gly Asn Phe Ile Ala Val Asp Asp Ser Pro Glu Glu Ile Arg Ala  
 210 215 220  
 Lys Ile Lys Lys Ala Tyr Cys Pro Ala Gly Val Val Glu Gly Asn Pro  
 225 230 235 240  
 Ile Met Glu Ile Ala Lys Tyr Phe Leu Glu Tyr Pro Leu Thr Ile Lys  
 245 250 255  
 Arg Pro Glu Lys Phe Gly Gly Asp Leu Thr Val Asn Ser Tyr Glu Glu  
 260 265 270  
 Leu Glu Ser Leu Phe Lys Asn Lys Glu Leu His Pro Met Asp Leu Lys  
 275 280 285  
 Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys  
 290 295 300  
 Arg Leu  
 305  
 <210> 3  
 <211> 918  
 <212> DNA  
 <213> Methanococcus jannaschii

<400> 3

atggacgaat ttgaaatgat aaagagaaaac acatctgaaa ttatcagcga ggaagagtta 60  
 agagagggtt taaaaaaaga tgaaaaatct gcttacatag gttttgaacc aagtggtaaa 120  
 5 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaaa tgctggattt 180  
 gatataatta tattgttggc tgatttacac gcctatttaa accagaaagg agagttggat 240  
 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300  
 10 aaatatgttt atggaagtga attccagctt gataaggatt atacactgaa tgtctataga 360  
 ttggctttta aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420  
 gatgaaaatc caaagggttc tgaagttatc tatccaataa tgcagggtta tgatattcat 480  
 tatttaggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540  
 15 agggagcttt taccaaaaaa gggtgtttgt attcacaacc ctgtcttaac gggtttggat 600  
 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660  
 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720  
 20 ataattggaga tagctaaata cttccttgaa tatcctttta ccataaaaag gccagaaaaa 780  
 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840  
 gaattgcata caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900  
 ccaattagaa agagatta 918  
 25

<210> 4

<211> 306

<212> PRT

30

<213> Искусственная последовательность

<220>

<223> L-(7-гидроксикумарин-4-ил)-этилглицин аминоксил-ТРНК синтетазы

35

<400> 4

Met Asp Glu Phe Glu Met Ile Lys Arg Asn Thr Ser Glu Ile Ile Ser  
 1 5 10 15

40

Glu Glu Glu Leu Arg Glu Val Leu Lys Lys Asp Glu Lys Ser Ala Glu  
 20 25 30

Ile Gly Phe Glu Pro Ser Gly Lys Ile His Leu Gly His Tyr Leu Gln  
 35 40 45

45

Ile Lys Lys Met Ile Asp Leu Gln Asn Ala Gly Phe Asp Ile Ile Ile  
 50 55 60

His Leu Gly Asp Leu Gly Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp

50

	65		70		75		80									
	Glu	Ile	Arg	Lys	Ile	Gly	Asp	Tyr	Asn	Lys	Lys	Val	Phe	Glu	Ala	Met
				85					90						95	
5	Gly	Leu	Lys	Ala	Lys	Tyr	Val	Tyr	Gly	Ser	Glu	Tyr	His	Leu	Asp	Lys
			100						105					110		
	Asp	Tyr	Thr	Leu	Asn	Val	Tyr	Arg	Leu	Ala	Leu	Lys	Thr	Thr	Leu	Lys
10			115					120					125			
	Arg	Ala	Arg	Arg	Ser	Met	Glu	Leu	Ile	Ala	Arg	Glu	Asp	Glu	Asn	Pro
		130					135					140				
	Lys	Val	Ala	Glu	Val	Ile	Tyr	Pro	Ile	Met	Gln	Val	Asn	Gly	Ile	His
15	145					150					155				160	
	Tyr	Gly	Gly	Val	Asp	Val	Ala	Val	Gly	Gly	Met	Glu	Gln	Arg	Lys	Ile
					165					170					175	
	His	Met	Leu	Ala	Arg	Glu	Leu	Leu	Pro	Lys	Lys	Val	Val	Cys	Ile	His
20				180					185					190		
	Asn	Pro	Val	Leu	Thr	Gly	Leu	Asp	Gly	Glu	Gly	Lys	Met	Ser	Ser	Ser
			195					200					205			
25	Lys	Gly	Asn	Phe	Ile	Ala	Val	Asp	Asp	Ser	Pro	Glu	Glu	Ile	Arg	Ala
		210					215					220				
	Lys	Ile	Lys	Lys	Ala	Tyr	Cys	Pro	Ala	Gly	Val	Val	Glu	Gly	Asn	Pro
30	225					230					235				240	
	Ile	Met	Glu	Ile	Ala	Lys	Tyr	Phe	Leu	Glu	Tyr	Pro	Leu	Thr	Ile	Lys
				245						250					255	
	Arg	Pro	Glu	Lys	Phe	Gly	Gly	Asp	Leu	Thr	Val	Asn	Ser	Tyr	Glu	Glu
35				260					265					270		
	Leu	Glu	Ser	Leu	Phe	Lys	Asn	Lys	Glu	Leu	His	Pro	Met	Asp	Leu	Lys
			275					280					285			
40	Asn	Ala	Val	Ala	Glu	Glu	Leu	Ile	Lys	Ile	Leu	Glu	Pro	Ile	Arg	Lys
		290					295					300				
	Arg	Leu														
45	305															
	<210>	5														
	<211>	921														
50	<212>	DNA														

<213> Искусственная последовательность

<220>

<223> L-(7-гидроксикумарин-4-ил)-этилглицин аминоксил-тРНК синтетазы

<400> 5

atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60  
agagagggttt taaaaaaga tgaaaaatct gctgagatag gttttgaacc aagtggtaaa 120  
atacatttag ggcattatct ccaaataaaa aagatgattg atttacaata tgctggattt 180  
gatataatta tacatttggg tgatttaggc gcctatttaa accagaaagg agagttggat 240  
gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300  
aaatatgttt atggaagtga atatcatctt gataaggatt atacactgaa tgtctataga 360  
ttggctttta aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420  
gatgaaaatc caaagggttg tgaagttatc tatccaataa tgcagggtta tgggattcat 480  
tatggtggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540  
agggagcttt taccaaaaaa gggtgtttgt attcacaacc ctgtcttaac gggtttggat 600  
ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660  
gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720  
ataatggaga tagctaaata cttccttgaa tatcctttaa ccataaaaag gccagaaaaa 780  
tttgggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840  
gaattgcac caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900  
ccaattagaa agagattata a 921

### Формула изобретения

1. Система трансляции, включающая

(а) первую неприродную аминокислоту, которая представляет собой неприродную аминокислоту L-(7-гидроксикумарин-4-ил)-этилглицин;

(b) первую ортогональную аминоксил-тРНК синтетазу (О-РС), содержащую последовательность аминокислот, указанную в SEQ ID NO:2, или консервативный вариант указанной последовательности, при этом SEQ ID NO:2 или консервативный вариант указанной последовательности дополнительно содержит Glu в положении, соответствующем Tyr32; His в положении, соответствующем Leu65; Gly в положении, соответствующем His70; Tyr в положении, соответствующем Phe108; His в положении, соответствующем Gln109; Gly в положении, соответствующем Asp158; и Gly в положении, соответствующем Leu162; и

(с) первую ортогональную тРНК (О-тРНК), кодируемую последовательностью SEQ ID NO:1,

причем указанная первая О-РС предпочтительно аминокислирует указанную первую О-тРНК указанной первой неприродной аминокислотой.

2. Система трансляции по п.1, отличающаяся тем, что указанная первая О-РС получена из аминоксил-тРНК-синтетазы *Methanococcus jannaschii*.

3. Система трансляции по п.1, отличающаяся тем, что указанная первая О-тРНК представляет собой супрессорную тРНК, распознающую амбер-кодон.

4. Система трансляции по п.1, которая дополнительно включает нуклеиновую

кислоту, кодирующую представляющий интерес белок, причем указанная нуклеиновая кислота содержит, по меньшей мере, один селекторный кодон, который распознает указанная первая О-тРНК.

5 Система трансляции по п.4, дополнительно содержащая вторую О-РС и вторую О-тРНК, причем указанная вторая О-РС предпочтительно аминоацилирует вторую указанную О-тРНК второй неприродной аминокислотой, которая отлична от указанной первой неприродной аминокислоты, и указанная вторая О-тРНК  
10 распознает селекторный кодон, отличный от селекторного кодона, распознаваемого указанной первой О-тРНК.

6. Система трансляции по п.1, отличающаяся тем, что указанная первая О-РС и указанная первая О-тРНК указанной системы экспрессируются в клетке-хозяине, при этом указанная клетка-хозяин представляет собой клетку эубактерии.

7. Система трансляции по п.6, отличающаяся тем, что указанная клетка-хозяин  
15 представляет собой клетку *E.coli*.

8. Система трансляции по п.6, отличающаяся тем, что указанная клетка-хозяин содержит полинуклеотид, кодирующий указанную первую О-РС, и полинуклеотид, кодирующий указанную первую О-тРНК.

20 9. Способ получения в системе трансляции белка, содержащего неприродную аминокислоту в выбранном положении, который включает:

(a) обеспечение системы трансляции, содержащей

(i) первую неприродную аминокислоту, которая представляет собой неприродную аминокислоту L-(7-гидроксикумарин-4-ил)-этилглицин,

25 (ii) первую ортогональную аминоацил-тРНК синтетазу (О-РС), содержащую последовательность аминокислот, соответствующую SEQ ID NO:2, или консервативный вариант указанной последовательности, при этом SEQ ID NO:2 или консервативный вариант указанной последовательности дополнительно  
30 содержит Glu в положении, соответствующем Tyr32; His в положении, соответствующем Leu65; Gly в положении, соответствующем His70; Tyr в положении, соответствующем Phe108; His в положении, соответствующем Gln109; Gly в положении, соответствующем Asp158; и Gly в положении, соответствующем Leu162;

35 (iii) первую ортогональную тРНК (О-тРНК), кодируемую последовательностью SEQ ID NO:1, и

(b) обеспечение нуклеиновой кислоты, кодирующей указанный белок, причем указанная нуклеиновая кислота содержит, по меньшей мере, один селекторный кодон, распознаваемый указанной первой О-тРНК;

40 (c) обеспечение трансляции указанного белка;

(d) выделение указанного белка.

10. Способ по п.9, отличающийся тем, что указанная О-РС получена из аминоацил-тРНК синтетазы *Methanococcus jannaschii*, или тирозил-тРНК синтетазы *Methanococcus jannaschii* дикого типа.

45 11. Способ по п.9, отличающийся тем, что указанное обеспечение системы трансляции включает обеспечение полинуклеотида, кодирующего указанную О-тРНК.

12. Способ по п.9, в котором указанный белок дополнительно содержит вторую неприродную аминокислоту, которая отлична от указанной первой неприродной аминокислоты, и указанная система трансляции дополнительно содержит вторую О-РС и вторую О-тРНК, причем указанная вторая О-РС предпочтительно  
50 аминоацилирует указанную вторую О-тРНК второй неприродной аминокислотой,

которая отлична от указанной первой неприродной аминокислоты, и указанная вторая О-тРНК распознает селекторный кодон в нуклеиновой кислоте, который отличен от селекторного кодона, распознаваемого первой О-тРНК.

5 13. Способ по п.9, отличающийся тем, что указанная первая О-РС и указанная первая О-тРНК указанной системы трансляции экспрессируются в клетке-хозяине, содержащей указанную нуклеиновую кислоту, кодирующую указанный белок, и при этом указанная клетка-хозяин представляет собой клетку эубактерии.

10 14. Способ по п.13, отличающийся тем, что указанная клетка-хозяин представляет собой клетку E.coli.

15 15. Способ по п.13, отличающийся тем, что указанная клетка-хозяин содержит полинуклеотид, кодирующий указанную первую О-РС, и полинуклеотид, кодирующий указанную первую О-тРНК.

16. Аминоацил-тРНК синтетаза, содержащая последовательность аминокислот, соответствующую SEQ ID NO:4, или консервативный вариант указанной последовательности.

17. Полинуклеотид, кодирующий ортогональную аминоацил-тРНК синтетазу (О-РС) по п.16.

20 18. Полинуклеотид по п.17, отличающийся тем, что указанный полинуклеотид содержит последовательность нуклеотидов, соответствующую SEQ ID NO:5.

19. Экспрессионный вектор, содержащий полинуклеотид по п.17.

25 20. Клетка-хозяин для получения белка, содержащего неприродную аминокислоту, содержащая вектор по п.19.

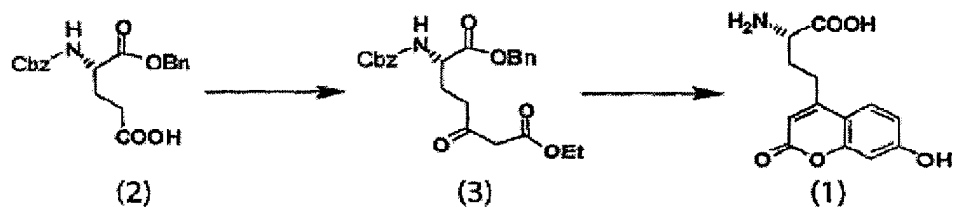
30

35

40

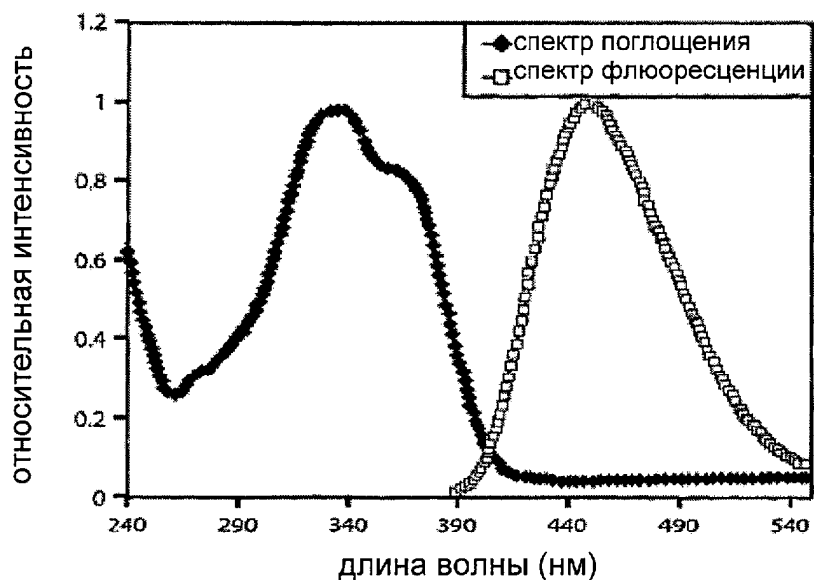
45

50

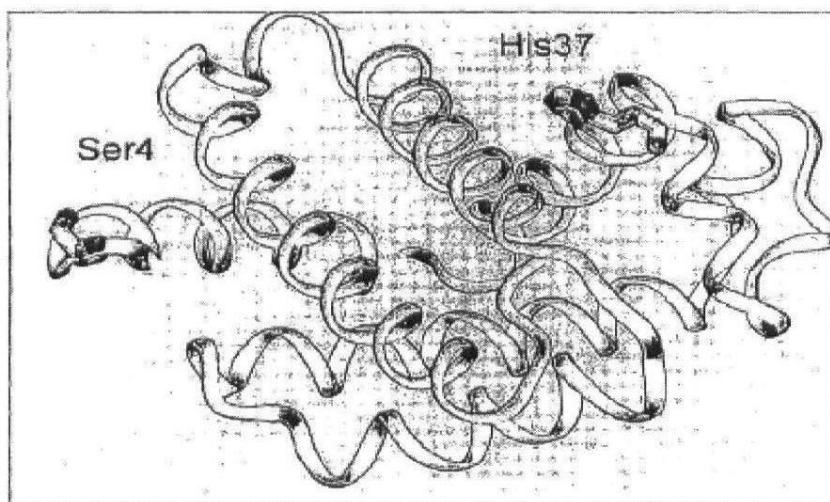


- 1) L-(7-гидроксикумарин-4-ил)-этилглицин
- 2) N-α-Cbz-L-глутаминовой кислоты α-бензиловый спирт
- 3) 7-этиловый эфир 1-бензилового эфира (2S)-2-бензилоксикарбониламино-5-оксо-гептандиойной кислоты

ФИГ.1

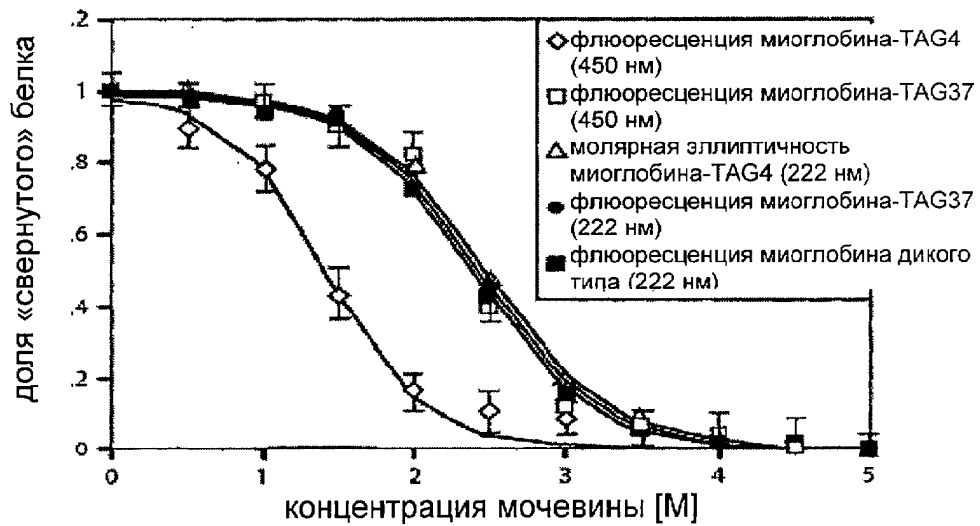


ФИГ.2

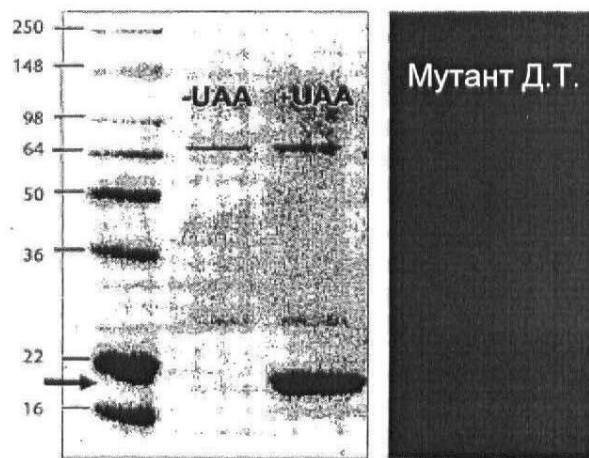


ФИГ.3

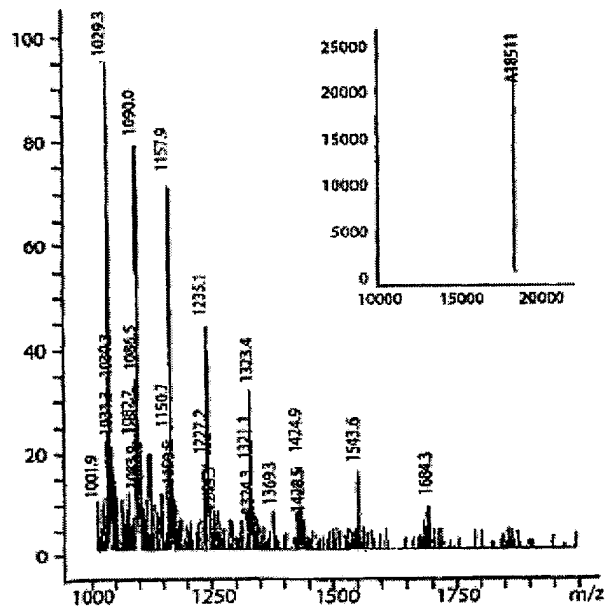




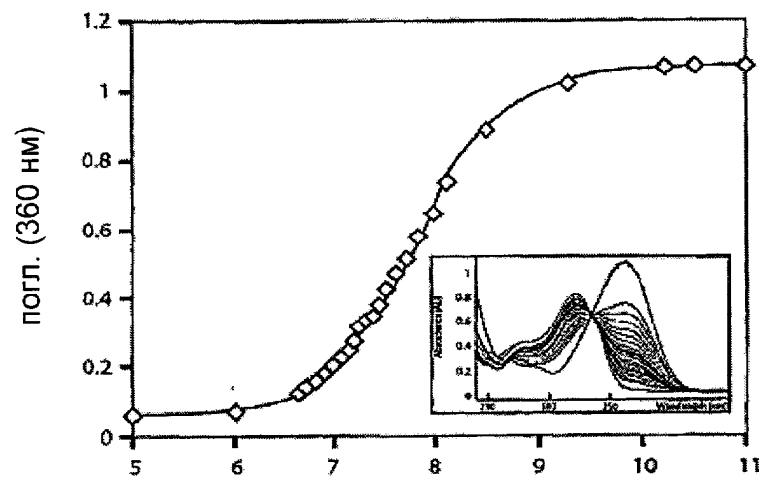
ФИГ.4



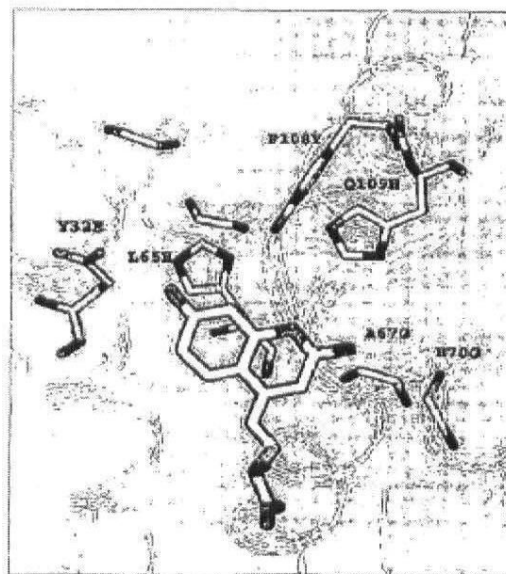
ФИГ.5



ФИГ.6



ФИГ.7



ФИГ.8

## Последовательность нуклеотидов и аминокислот

SEQ ID NO:	описание	последовательность
1	1. Супрессорная тирозил-тРНК <sup>сua</sup> <i>Methanococcus jannaschii</i> <i>MjtRNA-Tyr(CUA)</i> или <i>mutRNA<sup>Tyr</sup><sub>CUA</sub></i>	CCGGCGGUAGUUCAGCAGGGCAGAACGGCGGACUCUAAAUCCG CAUGGCGCUGGUUCAAUCCGGCCCGCCGGACCA
2	2. Последовательность аминокислот тирозил-тРНК синтетазы (MjTyrRS) <i>Methanococcus jannaschii</i> дикого типа	MDEFEMIKRNTSEIISEEELREVLKKDEKSAYIGFEPGSKIHL GHYLGQIKKMDLQONAGFDIIILLADLHAYLNQKGELDEIRKIG DYNKKVFEAMGLKAKYVYGSEFQLDKDYTLNVYRLALKTTLKR ARRSMELIAREDENPKVAEVIYPIMQVNDIHYLGVDVAVGGME QRKIHMLARELLPKKVVCIHNPVLTGLDGEKMSSSKGNFIAV DDSPSEIRAKIKKAYCPAGVVEGNPIMEIAKYFLEYPLTIKRP EKFGGDLTVNSYEELESLEFKNKELHPMDLKNVAEELIKILEP IRKRL
3	2. Последовательность нуклеотидов тирозил-тРНК синтетазы (MjTyrRS) <i>Methanococcus jannaschii</i> дикого типа	ATGGACGAATTTGAAATGATAAAGAGAAACACATCTGAAATTATCAGC GAGGAAGAGTTAAGAGAGGTTTAAAAAAGATGAAAAATCTGCTTAC ATAGGTTTTGAACCAAGTGGTAAATACATTTAGGGCATTATCTCCAA ATAAAAAAGATGATTTGATTTACAAAATGCTGGATTTGATATAATTATA TTGTTGGCTGATTTACACGCCCTATTTAAACCAGAAAGGAGAGTTGGAT GAGATTAGAAAAATAGGAGATTATAACAAAAAAGTTTTTGAAGCAATG GGTTAAAGGCAAAATATGTTTATGGAAGTGAATTCACGCTTGATAAG GATTATACACTGAATGCTATAGATTGSCCTTAAAAACTACCTTAAAA AGAGCAAGAAGGAGTATGGAACCTATAGCAAGAGAGGATGAAAAATCCA AAGGTTGCTGAAGTTATCTATCCAATAATGCAGGTTAATGATATTTCAT TATTTAGGCGTTGATGTTGCACTTGGAGGGATGGAGCAGAGAAAAATA CACATGTTAGCAAGGGAGCTTTTACCAAAAAAGGTTGTTTCTATTTCAC AACCTGTCTTAAACGGGTTTGGATGGAGAAGGAAAGATGAGTTCTTCA AAAGGGAATTTTATAGCTGTTGATGACTCTCCAGAAGAGATTAGGGCT AAGATAAAGAAAGCATACTGCCAGCTGGAGTTGTTGAACGAATCCA ATAATGGAGATAGCTAATACTTCCTTGAATATCCTTTAACCATAAAA AGGCCAGAAAAATTTGGTGGAGATTTGACAGTTAATAGCTATGAGGAG TTAGAGAGTTTATTTAAAAATAAGGAATTGCATCCAATGGATTAAAA AATGCTGTAGCTGAAGAACTTATAAAGATTTTAGAGCCAATTAGAAAG AGATTAA
4	2. Последовательность аминокислот L-(7-гидроксикумарин-4-ил)-этилглицин аминоцил-тРНК-синтетазы  (получается из тирозил-тРНК-синтетазы <i>Methanococcus jannaschii</i> дикого типа)	MDEFEMIKRNTSEIISEEELREVLKKDEKSAEIGFEPGSKIHL GHYLGQIKKMDLQONAGFDIIILLADLHAYLNQKGELDEIRKI GDYNKKVFEAMGLKAKYVYGSEYHLDDKDYTLNVYRLALKTTL KRARRSMELIAREDENPKVAEVIYPIMQVNGIHYGGVDVAVG GMEQRKIHMLARELLPKKVVCIHNPVLTGLDGEKMSSSKGNF IAVDDSPSEIRAKIKKAYCPAGVVEGNPIMEIAKYFLEYPLTI KRPEKFGGDLTVNSYEELESLEFKNKELHPMDLKNVAEELIKI LEPIRKRL

ФИГ.9

SEQ ID NO:	описание	последовательность
5	5. Последовательность нуклеотидов L-(7-гидроксикумарин-4-ил)-этилглицин аминоцил-тРНК-синтазы	ATGGACGAATTGTGAATGATAAAGAGAAAACACATCTGAAATTATCAGC GAGGAAGAGTTAAGAGAGGTTTAAAAAAGATGAAAAATCTGCTGAG ATAGGTTTGTGAACCAAGTGGTAAAAATACATTTAGGGCATTTATCTCCAA ATAAAAAAGATGATTGATTACAAAAATGCTGGATTTCATATTAATTATA CATTTGGGTGATTTAGGCGCCTATTTAAACCAGAAAGGAGAGTTGGAT GAGATTAGAAAAATAGGAGATTATAACAAAAAGTTTGTGAAGCAATG GGGTTAAAGGCCAAAAATATGTTTATGGAGTGAATATCATCTTGATAAG GATTATACACTGAATGTCTATAGATTGGCTTTAAAACTACCTTAAAA AGAGCAAGAAGGAGTATGGAACCTTATAGCAAGAGAGGATGAAAAATCCA AAGCTTCCTGAAGTTATCTATCCAATAATGCAGGTTAATGGGATTTCAT TATGGTGGCGTTGATGTTGCAGTTGGAGGGATGGAGCAGAGAAAAATA CACATCTTAGCAAGGGAGCTTTTACCAAAAAAGGTTGTTGTATTTCAC AACCCCTGCTTAAACGGGTTTGGATGGAGAAGGAAAGATGAGTTCTTCA AAAGGGAATTTTATAGCTGTTGATGACTCTCCAGAAGAGATTAGGGCT AAGATAAAGAAAGCATACTGCCAGCTGGAGTTGTTGAAGGAAATCCA ATAATGGAGATAGCTAAATACCTTCTTGAATATCCTTTAACCATAAAA AGGCCAGAAAAATTTGGTGGAGATTTGACAGTTAATAGCTATGAGGAG TTAGAGAGTTTATTTAAAAATAAGGAATTGCATCCAATGGATTAAAAA AATGCTGTAGCTGAAGAACTTATAAAGATTTTAGAGCCAATTAGAAAG AGATTATAA

**ФИГ.9**  
(продолж)