



(51) МПК
C12N 15/11 (2006.01)
C12P 21/00 (2006.01)
C12N 15/63 (2006.01)

**ФЕДЕРАЛЬНАЯ СЛУЖБА
 ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: **2008135114/10**, **13.03.2007**

(24) Дата начала отсчета срока действия патента:
13.03.2007

Приоритет(ы):

(30) Конвенционный приоритет:
16.03.2006 US 60/783,272
28.11.2006 US 60/861,456

(43) Дата публикации заявки: **27.04.2010** Бюл. № 12

(45) Опубликовано: **20.10.2012** Бюл. № 29

(56) Список документов, цитированных в отчете о поиске: **US 20050287639**, **29.12.2005. WANG L. et al., Chem. Biol., 8(9), 94720, 32001. KOHRER C. et al., Nucl. Acids Res., 32, no.21, 6200-6211, 2004. XIE et al., Curr. Opin. Chem. Biol., 9(6), 548-554, 2005. RU 2157239 C2, 10.10.2000.**

(85) Дата начала рассмотрения заявки РСТ на национальной фазе: **16.10.2008**

(86) Заявка РСТ:
US 2007/006382 (13.03.2007)

(87) Публикация заявки РСТ:
WO 2007/109035 (27.09.2007)

Адрес для переписки:

190000, Санкт-Петербург, ул. Малая Морская, 15, офис 5, BOX 1125, ООО "ПАТЕНТИКА", пат.пов. М.И.Ниловой, рег.№ 378

(72) Автор(ы):

**ВАНГ Джиангюн (US),
 ШУЛЬЦ Питер Г. (US)**

(73) Патентообладатель(и):

ЗЕ СКРИПС РЕСЕЧ ИНСТИТЬЮТ (US)

(54) ГЕНЕТИЧЕСКИ ЗАПРОГРАММИРОВАННАЯ ЭКСПРЕССИЯ БЕЛКОВ, СОДЕРЖАЩИХ НЕ ВСТРЕЧАЮЩУЮСЯ В ПРИРОДЕ АМИНОКИСЛОТУ ФЕНИЛСЕЛЕНОЦИСТЕИН

(57) Реферат:

Изобретение относится к области молекулярной биологии и биотехнологии и может быть использовано при получении рекомбинантных белков различного назначения, включающих не встречающуюся в природе аминокислоту фенилселеноцистеин (ФСЦ). Предлагается композиция, обеспечивающая возможность трансляции

ФСЦ-содержащего белка, которая состоит из ФСЦ, ортогональной тРНК (О-тРНК) и ортогональной аминоацил-тРНК синтетазы (О-РС), избирательно ацилирующей используемую тРНК фенилселеноцистеином, а также способ получения такого белка, где ФСЦ находится в заранее выбранном положении, в системе трансляции, представленной эубактериальной клеткой. Раскрыты возможности последующих

полезных модификаций ФСЦ, находящегося в
составе рекомбинантного продукта. 8 н. и 29

з.п., 1 табл., 2 ил., 1 пр.

R U 2 4 6 4 3 1 5 C 2

R U 2 4 6 4 3 1 5 C 2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.
C12N 15/11 (2006.01)
C12P 21/00 (2006.01)
C12N 15/63 (2006.01)

(12) **ABSTRACT OF INVENTION**

(21)(22) Application: **2008135114/10, 13.03.2007**

(24) Effective date for property rights:
13.03.2007

Priority:

(30) Convention priority:
16.03.2006 US 60/783,272
28.11.2006 US 60/861,456

(43) Application published: **27.04.2010 Bull. 12**

(45) Date of publication: **20.10.2012 Bull. 29**

(85) Commencement of national phase: **16.10.2008**

(86) PCT application:
US 2007/006382 (13.03.2007)

(87) PCT publication:
WO 2007/109035 (27.09.2007)

Mail address:

**190000, Sankt-Peterburg, ul. Malaja Morskaja, 15,
ofis 5, VOKh 1125, OOO "PATENTIKA", pat.pov.
M.I.Nilovoj, reg.№ 378**

(72) Inventor(s):

**VANG Dzhiangjun (US),
ShUL'Ts Piter G. (US)**

(73) Proprietor(s):

ZE SKRIPS RESECh INSTIT'JuT (US)

(54) **GENETICALLY PROGRAMMED EXPRESSION OF PROTEINS CONTAINING UNNATURAL AMINO ACID PHENYL-SELENOCYSTEINE**

(57) Abstract:

FIELD: medicine.

SUBSTANCE: invention may be used in preparing recombinant proteins for different applications containing unnatural amino acid phenyl-selenocysteine (PSC). What is offered is a composition enabling translation of PSC-containing protein, consisting of PSC, orthogonal tRNA (O-tRNA) and orthogonal aminoacyl-tRNA synthetase

(O-RS) selectively acylating the used tRNA by phenyl-selenocysteine, as well as a method for preparing such a protein wherein PSC is found in the pre-set position, in the translation system presented by an eubacterial cell.

EFFECT: there are disclosed possibilities of the future effective PSC modifications found in the composition of the recombinant product.

37 cl, 1 tbl, 2 dwg, 1 ex

ПЕРЕКРЕСТНЫЕ ССЫЛКИ НА РОДСТВЕННЫЕ ЗАЯВКИ

[0001] Настоящая заявка испрашивает приоритет согласно предварительной заявке на патент США №60/783,272, поданной 16 марта 2006; и предварительной заявке на патент США 60/861,456, поданной 28 ноября 2006, содержание которых полностью включено в настоящую заявку посредством ссылки для любых целей.

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

[0002] Настоящее изобретение сделано при государственной поддержке Национального Института Здоровья по гранту номер GM62159. Государство может иметь определенные права на изобретение.

ОБЛАСТЬ ТЕХНИКИ

[0003] Изобретение относится к области биохимии трансляции. Изобретение относится к композициям, а также способам получения и применения, ортогональной тРНК, ортогональной аминокил-тРНК синтетазы, а также их пар, способных к встраиванию в белки не встречающихся в природе аминокислот. Изобретение также относится к способам синтеза белков в клетках с использованием таких пар и белков, полученных этими способами.

УРОВЕНЬ ТЕХНИКИ

[0004] Исследование структуры и функции белка исторически основано на свойствах и химических реакциях, которые доступны при использовании реакционно-способных групп встречающихся в природе аминокислот. К сожалению, каждый известный организм, от бактерии до человека, кодирует одни и те же двадцать стандартных аминокислот. Эти 20 аминокислот содержат удивительно ограниченное число функциональных групп: азотистые основания, карбоксильные кислоты и амиды, спирты и тиоловые группы. Этот ограниченный выбор реакционно-способных групп сводит исследование структуры и функций белка к исследованиям химических свойств аминокислот, встречающихся в природе. Например, ограниченное число встречающихся в природе реакционных групп сузило возможность осуществления точно направленных модификаций белка, до исключения других аминокислот в белке.

[0005] Реакции химически селективного лигирования, в которые вовлечены белки, чрезвычайно важны для разнообразных целей, включая, без ограничения: изучение белок-белковых взаимодействий, передачу сигналов внутри клетки, создание новой белковой терапии. Самые селективные реакции модификации белка, применяемые в настоящее время в данной области техники, включают: образование ковалентных связей между нуклеофильными и электрофильным партнерами реакции, которые нацеливают нуклеофильные остатки природного происхождения в боковых цепях аминокислот белка, например, реакция α -гало кетонов с боковыми цепями гистидина или цистеина. Селективность в таких случаях определяется числом и доступностью нуклеофильных остатков в белке. К сожалению, встречающиеся в природе белки часто содержат неудачно расположенные (например, недоступные) сайты реакций или многочисленные мишени реакций (например, остатки лизина, гистидина и цистеина), что приводит к низкой селективности реакций модификации, что, в свою очередь, затрудняет осуществление точно направленной модификации белка, с помощью нуклеофильных/электрофильных реагентов. Кроме того, участки модификации обычно ограничены природными нуклеофильными боковыми цепями лизина, гистидина или цистеина. Модификация на других участках затруднена или невозможна.

[0006] В данной области техники существует потребность в разработке новых

стратегий встраивания не встречающихся в природе аминокислот в белки с целью изменения и изучения структуры и функции белка, в которых не встречающиеся в природе аминокислоты обладают новыми реакционными группами или другими особенностями, например, биологическими свойствами, не найденными у аминокислот, встречающихся в природе. В технике существует значительная потребность создания новых стратегий реакций модификации белка, которые обеспечивали бы высокоселективную модификацию белков и, кроме того, позволили бы модифицировать белки в физиологических условиях. Существует необходимость в разработке новых способов осуществления модификации белка, которые обеспечивали бы высокоспецифичную модификацию белка, например модификацию, при которой ни одна из природных аминокислот не подвергается перекрестной или побочной реакции. Новые химические составы для высокоспецифичной модификации белка находят множество разнообразных применений в исследованиях структуры и функции белка, а также в производстве терапевтических белков.

ЛИПИДИЗАЦИЯ БЕЛКОВ

[0007] Липидизация белков является ключевой посттрансляционной модификацией, которая вовлечена в локализацию белка, нормальный внутриклеточный транспорт белка и белок-белковые взаимодействия. Липидизация белков часто является необходимой для нормальной биологической активности. Эта особенность также важна для создания некоторых терапевтических белков. Липидизация также является важной для изучения белок-белковых взаимодействий и передачи сигналов в клетке. К сожалению, получить липидизированные белки, используя природную нелипидизированную форму белка, путем хемоселективного сшивания *in vitro* чрезвычайно трудно, и, как правило, липидизация ограничивается модификацией небольшого числа находящихся на поверхности белка остатков цистеина.

[0008] Для биологической активности многих белков клетки необходима ассоциация с мембраной клетки, которая зависит от посттрансляционной модификации цистеина остатками липидов, таких как фарнезил, миристоил и пальмитоил (Chernomordik and Kozlov (2003), *Annual Review of Biochemistry* 72:175-207). Например, многие рецепторы, сопряженные с G-белками, пальмитоилизированы, белки Ras и фарнезилированы и пальмитоилизированы (Chernomordik and Kozlov (2003), *Annual Review of Biochemistry* 72:175-207). В то время как фарнезилирование белка представляет собой устойчивую и необратимую модификацию, пальмитоилирование обратимо, что обеспечивает динамическое регулирование функции белка, и специфичное нацеливание на клеточные мембраны (Rocks et al. (2005), *Science* 307(5716):1746-1752). Кроме того, γ -карбоксиглутамовая кислота является необходимой модификацией, важной для кальцийзависимой мембранной адгезии в каскаде коагуляции (Davie et al. (1991), *Biochemistry* 30(43):10363-10370).

ОРТОГОНАЛЬНЫЕ СИСТЕМЫ ТРАНСЛЯЦИИ

[0009] Одна стратегия преодоления ограничений ограниченного генетического кода состоит в том, чтобы расширить генетический код и добавить аминокислоты, которые имеют новые реакционные свойства, связанные с биологической активностью. Была разработана общая методика для *in vivo* сайт-специфичного встраивания разнообразных не встречающихся в природе аминокислот в белки в организмах как эукариот, так и прокариот. Эта методика основана на ортогональных компонентах трансляции белка, которые распознают подходящий «селекторный» кодон, встраивая желаемую не встречающуюся в природе аминокислоту в определенное положение в процессе трансляции полипептида *in vivo*. В этих способах используют ортогональную

тРНК (О-тРНК), которая узнает селекторный кодон, а соответствующая специфичная ортогональная аминоацил-тРНК синтетаза (О-РС) присоединяет к О-тРНК не встречающуюся в природе аминокислоту. Эти компоненты не реагируют перекрестно ни с одной из эндогенных тРНК, РС, аминокислот или кодонов в организме-хозяине (то есть, они, очевидно, ортогональны). Применение таких ортогональных пар тРНК-РС позволило осуществить генетическое кодирование большого количества не встречающихся в природе аминокислот различной структуры.

[0010] Способы практического применения ортогональных систем трансляции, которые подходят для получения белков, содержащих одну или более не встречающихся в природе аминокислот, широко известны в данной области, так же как и общие способы получения ортогональных систем трансляции. См., например международные публикации: WO 2002/086075 "METHODS AND COMPOSITION FOR THE PRODUCTION OF ORTHOGONAL tRNA-AMINOACYL-tRNA SYNTHETASE PAIRS;" WO 2002/085923 "IN VIVO INCORPORATION OF UNNATURAL AMINO ACIDS;" WO 2004/094593 "EXPANDING THE EUKARYOTIC GENETIC CODE;" WO 2005/019415, поданную 7 июля 2004; WO 2005/007870, поданную 7 июля; WO 2005/007624, поданную 7 июля и WO 2006/110182, поданную 27 октября 2005, озаглавленную "ORTHOGONAL TRANSLATION COMPONENTS FOR THE VIVO INCORPORATION OF UNNATURAL AMINO ACIDS." Каждая из этих заявок полностью включена в настоящую заявку посредством ссылки. Дополнительное обсуждение ортогональных систем трансляции, которые включают не встречающиеся в природе аминокислоты, и способов их получения и применения, можно найти в Wang and Schultz, "Expanding the Genetic Code," Chem. Commun. (Camb.) 1:1-11 (2002); 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); и Xie and Schultz, "A Chemical Toolkit for Proteins - an Expanded Genetic Code," Nat. Rev. Mol. Cell Biol., 7(10):775-782 (2006).

КРАТКОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[0012] Настоящее изобретение обеспечивает композиции и способы встраивания не встречающейся в природе аминокислоты фенилселеноцистеина в растущую цепь полипептида в ответ на селекторный кодон, например терминирующий амбер-кодон, *in vivo* (например, в клетке). Указанные композиции включают пары ортогональных-тРНК (О-тРНК) и ортогональных аминоацил-тРНК синтетаз (О-РС), которые не взаимодействуют с механизмом трансляции клетки-хозяина. Т.е., эндогенная клеточная аминоацил-тРНК синтетазой хозяина не присоединяет (или по существу не присоединяет) аминокислоту (природной или не встречающейся в природе) к О-тРНК. Точно также О-РС согласно настоящему изобретению не присоединяют к эндогенным тРНК аминокислоты (природные или не встречающиеся в природе) до значительного или, в некоторых случаях, обнаруживаемого уровня. Новые композиции делают возможным получение в большом количестве белка, содержащего встроенные в процессе трансляции не встречающиеся в природе аминокислоты. Химические свойства не встречающейся в природе аминокислоты фенилселеноцистеина также делают возможными направленные модификации того остатка, который необходимо модифицировать для получения желаемой конъюгации, например, но не обязательно, конъюгации липида. Такие специфически модифицированные полипептиды могут найти применение в терапии и в биомедицинских исследованиях.

[0013] В некоторых аспектах, изобретение обеспечивает системы трансляции. Указанные системы включают первую ортогональную аминоацил-тРНК синтетазу (О-РС), первую ортогональную тРНК (О-тРНК), и первую не встречающуюся в природе аминокислоту, которая является фенилселеноцистеином, причем первая О-РС предпочтительно аминоацилирует первую О-тРНК первой не встречающейся в природе аминокислотой фенилселеноцистеином. В некоторых аспектах, О-РС предпочтительно аминоацилирует О-тРНК указанным фенилселеноцистеином с эффективностью, составляющей не менее 50% эффективности, наблюдаемой для системы трансляции, включающей ту же О-тРНК, фенилселеноцистеин и аминоацил-тРНК синтетазу, содержащую последовательность аминокислот SEQ ID NO:4, 6 или 8.

[0014] В системах трансляции можно применять компоненты, полученные из разнообразных источников. В одном варианте, О-РС, используемая в системе, может содержать последовательность аминокислот SEQ ID NOS:4, 6 или 8, и консервативные варианты этой последовательности. В некоторых вариантах реализации, О-тРНК - это янтарная супрессорная тРНК. В некоторых вариантах реализации, О-тРНК содержит или закодирована последовательностью SEQ ID NO:1.

[0015] В некоторых аспектах, система трансляции, кроме того, содержит нуклеиновую кислоту, кодирующую белок, представляющий интерес, при этом нуклеиновая кислота содержит, по меньшей мере, один селекторный кодон, распознаваемый О-тРНК.

[0016] В некоторых аспектах, система трансляции содержит вторую ортогональную пару (то есть, вторую О-РС и вторую О-тРНК), которая использует вторую не встречающуюся в природе аминокислоту, так что теперь система в состоянии встраивать, по меньшей мере, две различные не встречающиеся в природе аминокислоты на различных выбранных участках в полипептиде. В этой двойной системе вторая О-РС предпочтительно аминоацилирует вторую О-тРНК второй не встречающейся в природе аминокислотой, которая отличается от первой не встречающейся в природе аминокислоты, и вторая О-тРНК узнает селекторный кодон, который отличается от селекторного кодона, узнаваемого первой О-тРНК.

[0017] В некоторых вариантах реализации, система трансляции находится в клетке-хозяине (и включает клетку-хозяина). Клетка-хозяин используется без определенных ограничений, при условии, что О-РС и О-тРНК сохраняют свою ортогональность в организме клетки-хозяина. Клетка-хозяин может быть клеткой эубактерии, такой как E.coli. Клетка-хозяин может включать один или более полинуклеотидов, которые кодируют компоненты системы трансляции, включая О-РС или О-тРНК. В некоторых вариантах реализации, полинуклеотид, кодирующий О-РС, содержит последовательность нуклеотидов SEQ ID NO:5, 7 или 9.

[0018] Изобретение также обеспечивает способы получения белков, содержащих одну или более не встречающихся в природе аминокислот в выбранных положениях. В этих способах применяют системы трансляции, описанные выше. В целом, эти способы начинаются с этапа обеспечения системы трансляции, включающей: (i) первую не встречающуюся в природе аминокислоту, которая является фенилселеноцистеином; (ii) первую ортогональную аминоацил-тРНК синтетазу (О-АРС); (iii) первую ортогональную тРНК (О-тРНК), при этом О-РС предпочтительно аминоацилирует О-тРНК не встречающейся в природе аминокислотой; и (iv) нуклеиновую кислоту, кодирующую белок, причем нуклеиновая кислота содержит, по меньшей мере, один селекторный кодон (возможно амбер-кодон), который распознает первая О-тРНК. Затем способ обеспечивает встраивание не встречающейся в природе аминокислоты в

выбранное положение в белке в процессе трансляции белка в ответ на селекторный кодон, в результате чего происходит синтез белка, содержащего не встречающуюся в природе аминокислоту в выбранном положении. В некоторых аспектах описанных способов, О-РС предпочтительно аминоацилирует О-тРНК фенилселеноцистеином с эффективностью, которая составляет, по меньшей мере, 50% эффективности, наблюдаемой в системе трансляции, включающей ту же самую О-тРНК фенилселеноцистеин, и аминоацил-тРНК синтетазу, содержащую последовательность аминокислот SEQ ID NO:4, 6 или 8. В некоторых аспектах, система трансляции содержит нуклеиновую кислоту, которая кодирует О-РС.

[0019] Эти способы могут быть широко применяться с использованием разнообразных реактивов. В некоторых вариантах реализации обеспечивается полинуклеотид, кодирующий О-АРС. В некоторых вариантах реализации О-АРС содержит последовательность аминокислот SEQ ID NO:4, 6 или 8 или их консервативные варианты. Система трансляции, используемая в способах согласно настоящему изобретению, может содержать нуклеиновую кислоту, которая кодирует О-АРС, например, нуклеиновую кислоту SEQ ID NO:5, 7 или 9.

[0020] В некоторых вариантах реализации описанных способов этап обеспечения системы трансляции включает осуществление мутации кармана связывания аминокислоты аминоацил-тРНК синтетазы дикого типа путем сайт-направленного мутанеза и отбор полученных О-РС, которые предпочтительно аминоацилируют О-тРНК не встречающейся в природе аминокислотой. Этап отбора может включать положительный и отрицательный отбор О-РС из набора полученных после сайт-направленного мутагенеза молекул аминоацил-тРНК синтетазы. В некоторых вариантах реализации, на этапе обеспечения получают полинуклеотид, кодирующий О-тРНК, например, О-тРНК, которая представляет собой тРНК, являющуюся супрессором амбер-мутации, или О-тРНК, которая содержит или закодирована полинуклеотидом SEQ ID NO:1. В описанных способах на этапе обеспечения, возможно, также получают нуклеиновую кислоту, содержащую селекторный амбер-кодон, который используется системой трансляции.

[0021] Указанные способы могут также быть модифицированы таким образом, чтобы встраивать в белок больше одной не встречающейся в природе аминокислоты. В таких способах помимо первой системы трансляции применяют вторую ортогональную систему трансляции, причем вторая система встраивает другую аминокислоту и обладает специфичностью к другому селекторному кодону. Например, этап обеспечения может включать обеспечение второй пары О-РС и второй О-тРНК, в которой вторая О-РС предпочтительно аминоацилирует вторую О-тРНК второй не встречающейся в природе аминокислотой, которая отличается от первой не встречающейся в природе аминокислоты, и где вторая О-тРНК распознает селекторный кодон в нуклеиновой кислоте, который отличается от селекторного кодона, распознаваемого первой О-тРНК.

[0022] Способы получения белка, содержащего не встречающуюся в природе аминокислоту, можно также реализовать в клетке-хозяине. В таких случаях, обеспечивают клетку-хозяин, причем указанная клетка-хозяин содержит не встречающуюся в природе аминокислоту, О-РС, О-тРНК и нуклеиновую кислоту, содержащую, по меньшей мере, один селекторный кодон, и кодирующую белок, при этом культивирование клетки-хозяина приводит к встраиванию не встречающейся в природе аминокислоты. В некоторых вариантах реализации, этап обеспечения включает обеспечение клетки-хозяина эубактерии (например, E.coli). В некоторых

вариантах реализации этап обеспечения включает обеспечение клетки-хозяина, которая содержит полинуклеотид, кодирующий О-РС. Например, полинуклеотид, кодирующий О-РС, может содержать последовательность нуклеотидов SEQ ID NO:5, 7 или 9.

5 [0003] В некоторых вариантах описанных способов процедуры дополнительно включают модификацию фенилселеноцистеина после его встраивания в полипептид. Например, в определенных условиях можно осуществить реакцию фенилселеноцистеина, которая приведет его к преобразованию в дегидроаланин, расположенный в выбранном положении. Эта реакция может представлять собой окислительное отщепление. Указанную реакцию можно осуществить путем приведения фенилселеноцистеина в контакт с перекисью водорода.

10 [0024] Изобретение также обеспечивает способы получения липидизированных белков, в которых липид присоединяют (конъюгируют) в определенном выбранном положении. В этих способах применяют системы трансляции, описанные выше. Как правило, указанные способы начинают с этапа обеспечения системы трансляции, включающей: (i) не встречающуюся в природе аминокислоту фенилселеноцистеин; (ii) ортогональную аминоацил-тРНК синтетазу (О-РС); (iii) ортогональную тРНК (О-тРНК), причем О-РС предпочтительно аминоацилирует О-тРНК фенилселеноцистеином; и, (iv) кодирующую белок нуклеиновую кислоту, которая содержит, по меньшей мере, один селекторный кодон (например амбер-кодон), распознаваемый О-тРНК. Следовательно, согласно данному способу, фенилселеноцистеин встраивают в выбранное положение в белке в процессе трансляции белка в ответ на селекторный кодон. Фенилселеноцистеин далее подвергают химическому преобразованию, в результате которого он превращается в дегидроаланин, находящийся в выбранном положении, который, в свою очередь, сам реагирует с липидом с образованием липидизированного фрагмента аминокислоты. В итоге такой способ обеспечивает образование белка, содержащего липид в выбранном положении в белке.

25 [0025] В некоторых аспектах указанных способов, реакцию фенилселеноцистеина осуществляют путем окислительного отщепления или приведения в контакт с перекисью водорода. В некоторых аспектах, конъюгируемый липид, реагирующий с дегидроаланином, может быть тиопальмитиновой кислотой, фарнезилмеркаптаном или 1-гексадекантиолом. Соответственно, полученные липидизированные аминокислоты представляют собой пальмитоилцистеин, фарнезилцистеин и S-гексадецилцистеин. Модификацию дегидроаланина обычно осуществляют путем реакции присоединения Майкла.

35 [0026] Изобретение также обеспечивает способы получения белка, содержащего остаток дегидроаланина в выбранном положении. В таких способах применяют системы трансляции, описанные выше. Как правило, указанные способы начинают с шага обеспечения системы трансляции, включающей: (i) первую не встречающуюся в природе аминокислоту, которая является фенилселеноцистеином; (ii) первую ортогональную аминоацил-тРНК синтетазу (О-РС); (iii) первую ортогональную тРНК (О-тРНК), при этом О-РС предпочтительно аминоацилирует О-тРНК не встречающейся в природе аминокислотой; и (iv) нуклеиновую кислоту, кодирующую белок, причем указанная нуклеиновая кислота содержит, по меньшей мере, один селекторный кодон (например амбер-кодон), распознаваемый указанной О-тРНК. Соответственно, в данном способе встраивают фенилселеноцистеин в выбранное положение в белке в процессе трансляции белка в ответ на селекторный кодон. Затем

осуществляют реакцию фенилселеноцистеина с образованием дегидроаланина, находящегося в выбранном положении. В некоторых аспектах описанных способов, реакцию фенилселеноцистеина осуществляют путем окислительного отщепления или приведения в контакт с перекисью водорода.

5 [0027] Изобретение также обеспечивает разнообразные композиции, включающие нуклеиновые кислоты и белки. Природа композиций не ограничена конкретно, при условии, что композиция содержит указанную нуклеиновую кислоту или белок. Композиции согласно настоящему изобретению могут включать любое число
10 дополнительных компонентов любой природы.

[0028] Например, изобретение обеспечивает композиции, включающие полипептиды О-РС, причем указанные полипептиды включают последовательности аминокислот SEQ ID NO:4, 6 или 8, или их консервативный вариант. В некоторых аспектах, консервативный вариант полипептида аминоацилирует когнатную
15 ортогональную тРНК (О-тРНК) не встречающейся в природе аминокислотой с эффективностью, которая составляет, по меньшей мере, 50% эффективности, наблюдаемой в системе трансляции, включающей О-тРНК, не встречающуюся в природе аминокислоту и аминоацил-тРНК синтетазу, содержащую
20 последовательность аминокислот SEQ ID NO:4, 6, 8 или 10. Изобретение также обеспечивает полинуклеотиды, которые кодируют любой из указанных полипептидов. В некоторых вариантах реализации указанные полинуклеотиды могут содержать последовательность нуклеотидов SEQ ID NO:5, 7 или 9. В некоторых вариантах реализации, полипептиды находятся в клетке.

25 [0029] Изобретение также обеспечивает композиции полинуклеотидов, содержащих последовательность нуклеотидов SEQ ID NO:5, 7 или 9. В некоторых вариантах реализации, изобретение обеспечивает векторы, содержащие полинуклеотиды, например, векторы экспрессии. В некоторых вариантах реализации, изобретение
30 обеспечивает клетки, содержащие вектор, описанный выше.

ТЕРМИНОЛОГИЯ

[0030] Перед подробным описанием согласно настоящему изобретению следует пояснить, что данное изобретение не ограничено определенными биологическими системами, которые могут, конечно, варьировать. Следует также пояснить, что
35 терминология, используемая здесь, предназначена исключительно для описания конкретных вариантов реализации и не является ограничивающей. В настоящем описании, а также в прилагающемся формуле формы единственного числа включают множественные объекты, если контекст не указывает ясно на противоположное.
40 Таким образом, например, "клетка" включает комбинации двух или более клеток; "полинуклеотид" включает, в качестве практического значения, множество копий этого полинуклеотида.

[0031] Если не определено здесь и ниже в остальной части описания, все употребляемые в настоящему описании технические и научные термины имеют
45 значение, обычно понимаемое средними специалистами в области техники, к которой относится настоящее изобретение.

[0032] Ортогональный: в настоящем описании термин "ортогональный" относится к молекуле (например, ортогональная тРНК (О-тРНК) и/или ортогональная
50 аминоацил-тРНК синтетаза (О-РС)), которая действует с эндогенными компонентами клетки с меньшей эффективностью, чем соответствующая молекула, которая является эндогенной в отношении к клетке или системе трансляции, или эта молекула (ортогональная) не может функционировать вместе с эндогенными компонентами

клетки. В контексте тРНК и аминоксил-тРНК синтетазы, термин «ортогональный» относится к неспособности или пониженной эффективности, например, менее чем 20%-ой эффективности, менее 10%-ой эффективности, менее 5%-ой эффективности, или менее 1%-ой 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. Ортогональная молекула не имеет нормально функционирующей эндогенной комплементарной молекулы в клетке. Например, ортогональную тРНК в клетке любая эндогенная РС аминоксилует клетки с пониженной или даже нулевой эффективностью, по сравнению с аминоксилацией эндогенной тРНК эндогенной РС. В другом примере, ортогональная РС аминоксилует любую эндогенную тРНК клетки, представляющую интерес, с пониженной или даже нулевой эффективностью, по сравнению с аминоксилацией 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 контролем, например, соответствующей парой эндогенных тРНК/АРС, или активной ортогональной парой (например, парой ортогональная тРНК тирозила/РС).

[0004] Ортогональная тирозил-тРНК. В настоящей заявке ортогональная тирозил-тРНК - это тирозил-тРНК (тирозил-О-тРНК), которая является ортогональной по отношению к системе трансляции, представляющий интерес, причем тРНК: (1) идентична или по существу подобна тирозил-тРНК природного происхождения, (2) получена из тирозил-тРНК естественного происхождения путем естественного или искусственного мутагенеза, (3) получена любым способом, в котором применяют последовательность тирозил-тРНК дикого типа или мутанта с последовательностью (1) или (2), (4) гомологична дикому типу или мутанту тирозил-тРНК; (5) гомологична любому примеру тРНК, обозначенному как субстрат для тирозил-тРНК синтетазы на фиг.2, или (6) консервативный вариант любого примера тРНК, который обозначен как субстрат для тирозил-тРНК синтетазы на фиг.2. Тирозил-тРНК может находиться в состоянии с присоединенной аминокислотой или в состоянии без аминокислоты. Также очевидно, что когнатная синтетаза может присоединить к "тирозил-О-тРНК" (т.е. аминоксилуют тирозил-О-тРНК) также аминокислоту, отличную от тирозина, соответственно, например, не встречающейся в природе аминокислотой. Действительно, очевидно, что тирозил-О-тРНК согласно настоящему изобретению выгодно применять для встраивания по существу любой аминокислоты, природной или не встречающейся в природе, в растущий полипептид в процессе трансляции в ответ на селективный кодон.

[0034] Ортогональная тирозил синтетаза: В настоящем описании ортогональная тирозил синтетаза (тирозил-О-РС) представляет собой фермент, который предпочтительно аминоксилует аминокислотой тирозил-О-тРНК в системе трансляции, представляющий интерес. Аминокислота, которую тирозил-О-РС присоединяет к тирозил-О-тРНК, может в данном случае без ограничения

представлять собой любую аминокислоту, природную, не встречающуюся в природе или искусственную. Синтетаза может быть идентична либо гомологична природной тирозил синтетазе, либо идентична или гомологична синтетазе, обозначенной как О-РС на фиг.2. Например, О-РС может быть консервативным вариантом тирозил-О-РС, представленной на фиг.2, и/или может быть, по меньшей мере, на 50%, 60%, 70%, 80%, 90%, 95%, 98%, 99% или более идентична последовательности О-РС, представленной на фиг.2.

[0035] Когнатный: термин "когатный" относится к компонентам, которые функционируют вместе или имеют некоторый аспект специфики друг к другу, например, ортогональная тРНК и ортогональная аминоацил-тРНК синтетаза. Компоненты могут быть также называться комлементарными.

[0036] Предпочтительно аминоацилирует: в настоящей заявке применительно к ортогональным системам трансляции, О-РС "предпочтительно аминоацилирует" когатную О-тРНК, в том случае, если О-РС присоединяет к О-тРНК аминокислоту более эффективно, чем к любой эндогенной тРНК в системе экспрессии. Таким образом, когда О-тРНК и любая данная эндогенная тРНК присутствуют в системе трансляции в приблизительно равном молярном соотношении, О-РС присоединяет аминокислоты к О-тРНК (нагружает) чаще, чем к эндогенной тРНК.

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

[0037] О-РС "предпочтительно аминоацилирует О-тРНК не встречающейся в природе аминокислотой", в том случае, если (а) О-РС предпочтительно аминоацилирует О-тРНК по сравнению с эндогенными тРНК, и (b) причем это аминоацилирование является специфичным в отношении не встречающейся в природе аминокислоты, по сравнению с присоединением ортогональной тРНК синтетазой к О-тРНК любой природной аминокислоты. Таким образом, если не встречающаяся в природе и природная аминокислоты присутствуют в равных молярных количествах в системе трансляции, включающей О-РС и О-тРНК, О-РС будет присоединять к О-тРНК чаще не встречающуюся в природе аминокислоту, чем природную аминокислоту. Предпочтительно, относительное отношение О-тРНК, «нагруженной» не встречающейся в природе аминокислотой, к О-тРНК, «нагруженной» природной аминокислотой, высоко. Более предпочтительно, О-РС присоединяет к О-тРНК исключительно, или почти исключительно, не встречающую в природе аминокислоту. Относительное соотношение между присоединением к О-тРНК не встречающейся в природе аминокислоты и присоединением к О-тРНК природной аминокислоты, если природная и не встречающаяся в природе аминокислоты присутствуют в системе трансляции в равных молярных концентрациях, составляет 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.

[0005] Селекторный кодон: термин "селекторный кодон" относится к кодонам, распознаваемым О-тРНК в процессе трансляции и не распознаваемым эндогенной тРНК. Петля антикодона 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.

[0039] Супрессорная тРНК: супрессорная тРНК - это тРНК, которая изменяет чтение матричной РНК (мРНК) в данной системе трансляции, обычно, путем обеспечения встраивания аминокислоты в ответ на стоп-кодон (то есть, прочитывания "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.

[0040] 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). Супрессорную (подавляющую) 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.

[0041] Настоящее изобретение обеспечивает различные способы количественного анализа супрессорной активности. Процент супрессии селекторного кодона (например, амбер-кодона) определенными О-тРНК и О-РС, представляющих интерес, равен процентной доли активности данного экспрессируемого тестового маркера (например, LacZ), который включает селекторный кодон, в нуклеиновой кислоте, кодирующей экспрессируемый маркер, в системе трансляции, представляющей интерес, где система трансляции, представляющая интерес, содержит О-РС и О-тРНК, по сравнению с конструкцией положительного контроля, в которой положительный контроль не содержит О-тРНК, О-РС и селекторного кодона. Таким образом, например, если конструкция активного маркера положительного контроля, которая не содержит селекторного кодона, имеет наблюдаемую активность X в данной системе трансляции, 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, представляющей интерес.

[0042] 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 (derivatизированной) 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 соответствующий смысловой, а не селекторный кодон в желаемом положении.

[0043] Система трансляции: термин "система трансляции" относится к компонентам, которые встраивают аминокислоту в растущую цепь полипептида (белка).

Компоненты системы трансляции могут включать, например, рибосомы, тРНК, синтетазы, мРНК и т.п. О-тРНК и/или О-РС согласно настоящему изобретению могут быть добавлены к или быть частью *in vivo* или *in vitro* систем трансляции, например, в неэукариотической клетке, например, в клетке бактерии (такой, как *E. coli*), или в эукариотической клетке, например, клетке дрожжей, клетке млекопитающего, клетке растения, клетке водорослей, клетке гриба, клетке насекомого, и/или в другой подобной клетке.

[0044] Не встречающаяся в природе аминокислота: В настоящей заявке термин "не встречающаяся в природе аминокислота" относится к любой аминокислоте, модифицированной аминокислоте, и/или аналогу аминокислоты, которые не являются одной из 20 стандартных аминокислот природного происхождения. Например, в настоящем изобретении находит применение не встречающаяся в природе аминокислота фенилселеноцистеин (см. фиг.1, структура 1).

[0045] 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 может можно осуществлять соответствующими методами скрининга, например, описанными в настоящей заявке.

5 Мутагенез полипептида обычно влечет за собой манипуляцию над полинуклеотидом, который кодирует данный полипептид.

[0046] 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.

[0047] Маркер отрицательного отбора или скрининга: В настоящей 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).

[0048] Репортер: В настоящей заявке термин "репортер" относится к компоненту, который можно применять для идентификации и/или отбора целевых компонентов системы, представляющей интерес. Например, репортер может включать белок, например фермент, который обеспечивает устойчивость или чувствительность к антибиотикам (например, β -лактамаза, хлорамфеникол ацетилтрансфераза (ХАТ) и т.п.), флуоресцентный маркер скрининга (например, зеленый флуоресцентный белок, например, (GFP), YFP, EGFP, RFP, и т.д.), люминесцентный маркер (например, белок люциферазы светлячка), маркер скрининга, основанного на аффинности, или маркерные гены положительного или отрицательного отбора, такие как *lacZ*, β -gal/*lacZ* (β -галактозидаза), *ADH* (алкоголь дегидрогеназа), *his3*, *ura3*, *leu2*, *lys2*, или тому подобное.

[0049] Эукариот: В настоящей заявке термин "эукариот" (эукариотический) относится к организмам, принадлежащим царству Eucarya (Эукариоты). Эукариоты, в целом, отличаются от прокариот своей, обычно многоклеточной, организацией (но не исключительно многоклеточной, например, дрожжи), наличием ограниченного мембраной ядра и других ограниченных мембранами органелл, линейным генетическим материалом (то есть, линейными хромосомами), отсутствием оперонов, наличием интронов, кэппированием и полиаденилированием мРНК, и другими биохимическими особенностями, такими как другая структура рибосом.

40 Эукариотические организмы включают, например, животных (например, млекопитающих, насекомых, рептилий, птиц, и т.д.), инфузорий, растения (например, однодольные, двудольные, водоросли, и т.д.), грибы, дрожжи, жгутиконосцев, микроспоридий, протистов, и т.д.

[0050] Прокариот: в настоящей заявке термин "прокариот" (прокариотический) относится к организмам, принадлежащим к царству Monera (также называемому Prokarya - Прокариоты). Прокариотические организмы в целом отличаются от эукариот одноклеточной организацией, бесполом размножением путем почкования или деления, отсутствием ограниченного мембраной ядра или других ограниченных мембраной органелл, кольцевыми хромосомами, наличием оперонов, отсутствием интронов, кэппирования и полиаденилирования мРНК, и другими биохимическими особенностями, такими как другая структура рибосом. Прокариоты

включают подцарства Eubacteria и Archaea (иногда называемое "Archaeobacteria").
Cyanobacteria (синие-зеленые водоросли) и микоплазмы иногда выделяют в отдельное царство Monera.

5 [0051] Бактерия: В настоящей заявке термин "бактерия" и "эубактерия" относится к прокариотическим организмам, которые не являются архебактериями (Archaea). Аналогично, термин Archaea относится к прокариотам, которые не являются эубактериями (eubacteria). Eubacteria и Archaea можно различить по множеству морфологических и биохимических признаков. Например, чтобы отнести организм
10 к Eubacteria или Archaea используют различия в последовательностях рибосомных РНК, структуре РНК-полимеразы, наличие или отсутствие интронов, чувствительность к антибиотикам, наличие или отсутствие пептидогликанов клеточной стенки и других компонентов клеточной стенки, по разветвленной структуре мембранных липидов или структуре без ветвей, и наличию/отсутствию
15 гистонов и гистоноподобных белков.

[0052] Примеры Eubacteria включают *Escherichia coli*, *Thermus thermophilus*, *Bacillus subtilis* и *Bacillus stearothermophilus*. Примеры Archaea включают *Methanococcus jannaschii* (Mj), *Methanosarcina mazei* (Mm), *Methanobacterium thermoautotrophicum* (Mt),
20 *Methanococcus maripaludis*, *Methanopyrus kandleri*, *Halobacterium*, типа *Haloferax volcanii* и *Halobacterium* вида NRC-1, *Archaeoglobus fulgidus* (Af), *Pyrococcus furiosus* (Pf), *Pyrococcus horikoshii* (Ph), *Pyrobaculum aerophilum*, *Pyrococcus abyssi*, *Sulfolobus solfataricus* (Ss), *Sulfolobus tokodaii*, *Aeuropyrum pernix* (Ap), *Thermoplasma acidophilum* и *Thermoplasma volcanium*.

25 [0053] Консервативный вариант: В настоящей заявке термин "консервативный вариант", применительно к компоненту трансляции, обозначает компонент трансляции, например, консервативный вариант О-тРНК или консервативный вариант О-РС, который функционирует аналогично основному компоненту, аналогом
30 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 этой О-РС будет аминоацилировать когнатную О-тРНК не встречающейся в природе аминокислотой фенилселеноцистеином. В этом примере, О-РС и консервативный вариант О-РС не
35 обладают идентичными последовательностями аминокислот. 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,
40 работает с) соответствующей когнатной О-тРНК или О-РС. В вариантах способах реализации, консервативный вариант О-РС содержит одну или более консервативную замену аминокислоты по сравнению с О-РС, из которого данный вариант получен. В некоторых вариантах реализации, консервативный вариант О-РС содержит одну или более консервативную замену - аминокислоты по сравнению с О-РС, из которого он
45 был получен, и кроме того, сохраняет биологическую активность О-РС; например, консервативный вариант О-РС, который сохраняет, по меньшей мере, 10% биологической активности материнской молекулы О-РС, из которой он был получен, или, по меньшей мере, 20%, по меньшей мере, 30%, или, по меньшей мере, 40%. В
50 некоторых предпочтительных вариантах реализации, консервативный вариант О-РС сохраняет, по меньшей мере, 50% биологической активности материнской молекулы О-РС, из которой он был получен. Консервативные замены аминокислот консервативного варианта О-РС могут иметь место в любом домене О-РС, включая

карман связывания аминокислоты.

[0055] Агент отбора или скрининга: В настоящей заявке термин "агент отбора или скрининга" относится к агенту, 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 может быть изменяться, например, концентрация, интенсивность и т.д.

[0056] В ответ на: В настоящей заявке термин "в ответ на" относится к процессу, при котором О-тРНК согласно настоящему изобретению распознает селекторный кодон и осуществляет встраивание не встречающейся в природе аминокислоты, которая присоединена к тРНК, в растущую цепь полипептида.

[0057] Кодировать: В настоящей заявке термин "кодировать" (кодирует) здесь относится к любому процессу, благодаря которому информация, содержащаяся в полимерной макромолекуле или строке последовательности, обеспечивает управление получением второй молекулы или строки последовательности, которая отличается от первой молекулы или строки последовательности. В настоящей заявке данный термин употребляется в широком смысле и может иметь разнообразные приложения. В некоторых аспектах, термин "кодирует" 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.

[0058] В другом аспекте, термин "кодировать" относится к любому процессу, посредством которого информация, 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, посредством процесса транскрипции с участием фермента ДНК-зависимой РНК полимеразы). Кроме того, молекула РНК может кодировать полипептид, как в процессе трансляции. Если данный термин используется для описания процесса трансляции, термин "кодирует" также охватывает триплетный кодон, который кодирует аминокислоту. В некоторых случаях, молекула РНК может кодировать молекулу ДНК, например, как в процессе обратной транскрипции с участием РНК-зависимой ДНК-полимеразы. В другом аспекте, молекула ДНК может кодировать полипептид, причем предполагается, что употребляемый в этом случае термин "кодировать" включает как процесс транскрипции, так и процесс трансляции.

КРАТКОЕ ОПИСАНИЕ РИСУНКОВ

[0059] На фиг.1 представлены химические структуры не встречающейся в природе аминокислоты фенилселеноцистеина (структура 1), а также структуры, которые могут быть получены из фенилселеноцистеина. Эти структуры включают дегидроаланин (2), пальмитоилцистеин (3), фарнезилцистеин (4), S-гексадецилцистеин (5) и гамма-карбоксиглютамовую кислоту (6).

[0060] На фиг.2 представлены различные нуклеотиды и последовательности аминокислот, которые находят применение в настоящем изобретении.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[0061] Изобретение обеспечивает решения проблем, связанных с ограничениями, свойственными применению системы трансляции, ограниченной 20 аминокислотами природного происхождения. Указанные решения включают композиции и способы, связанные с запрограммированным сайт-специфичным встраиванием не

встречающейся в природе аминокислоты фенилселеноцистеина (фиг.1, структура 1) в белки в процессе биосинтеза с применением ортогональных систем трансляции.

Можно запрограммировать встраивание фенилселеноцистеина в белок таким образом, что оно будет происходить в любом желаемом положении, путем конструирования полинуклеотида, который кодирует белок, представляющий интерес, и который должен содержать селекторный кодон, являющийся сигналом встраивания не встречающейся в природе аминокислоты в растущую цепь полипептида.

[0062] Изобретение обеспечивает новые композиции, включающие новые аминоксил-тРНК синтетазы (О-РС), которые способны присоединять («нагружать») к подходящей когнатной супрессорной О-тРНК (например, О-тРНК SEQ ID NO:1) фенилселеноцистеин. Указанные О-РС - это новые мутанты тирозил-тРНК синтетазы *Methanococcus jannaschii*, которые селективно присоединяют к О-тРНК не встречающуюся в природе аминокислоту фенилселеноцистеин в бактериальных клетках-хозяевах. В наиболее предпочтительном варианте ортогональные компоненты не реагируют перекрестно с эндогенными компонентами механизмов трансляции клетки-хозяина (например, клетки *E.coli*).

[0063] О-РС согласно настоящему изобретению может включать О-РС последовательностей SEQ ID NOS:4, 6 или 8. Изобретение также обеспечивает полинуклеотиды, которые кодируют полипептиды указанных О-РС. Данное описание также раскрывает способы разработки новых пар О-тРНК/О-РС, которые функционируют в эубактериях и включают не встречающуюся в природе аминокислоту фенилселеноцистеин в определенный сайт (сайт-специфически) в ответ на селекторные кодоны.

[0064] Изобретение также обеспечивает способы очень эффективного и сайт-специфического генетически запрограммированного встраивания фенилселеноцистеина (фиг.1, структура 1) в белки (предпочтительно *in vivo*) в ответ на селекторный кодон (например, янтарный («амбер») нонсенс-кодон, TAG). Указанные новые способы и композиции можно применять, например, в бактериальной системе-хозяине.

[0065] В некоторых случаях, не встречающуюся в природе аминокислоту фенилселеноцистеин можно, кроме того, специфично и региоселективно модифицировать после ее встраивания в полипептид, как описано в настоящем описании. Благодаря химическим реакционным свойствам фенилселено-группы, белки, в которые включена данная не встречающаяся в природе аминокислота, можно модифицировать с чрезвычайно высокой селективностью. В некоторых случаях, реакционная группа не встречающейся в природе аминокислоты фенилселеноцистеина предпочтительна из-за того, что является полностью чуждой для *in vivo* систем, и, благодаря этому обеспечивает повышение селективности реакции. В некоторых аспектах, возможно проведение реакций модификации в относительно мягких условиях реакции, которые обеспечивают как *in vitro*, так и *in vivo* реакции конъюгации, в которых участвуют белки, и сохраняют жизнеспособность клетки-хозяина и/или биологическую активность белка.

[0066] В некоторых аспектах, модифицируют фрагмент включенного фенилселеноцистеина, и полученный модифицированный продукт затем в свою очередь модифицируют, например, путем осуществления реакции конъюгации.

[0067] Природа объекта, который в конечном счете конъюгируют в выбранном положении в белке (соответствующем селекторному кодону в открытой рамке считывания, кодирующей белок), не ограничена отдельно, и такой объект представляет собой любой желательный объект, например липиды, красители,

флюорофоры, сшивающие агенты, производные сахаров, полимеры (например, производные полиэтиленгликоля), фотоактивируемые сшивающие агенты, цитостатические соединения, аффинные метки, производные биотина, смолы, гранулы, вторые (или более) белки или полипептиды, полинуклеотиды (например, ДНК, РНК и т.д.), хелаторы металлов, кофакторы, жирные кислоты, углеводы и т.п.

[0068] В некоторых аспектах, чтобы продемонстрировать (но не ограничивать) существующее изобретение, в настоящей заявке описан фрагмент не встречающейся в природе аминокислоты фенилселеноцистеина, включенный в модельный белок, например, миоглобин, гормон роста человека и GFP. Это не означает, что встраивание не встречающейся в природе аминокислоты фенилселеноцистеина ограничено каким-либо определенным белком. Очевидно, идеи настоящего изобретения позволяют осуществить включение не встречающейся в природе аминокислоты фенилселеноцистеина в любой желаемый белок, представляющий интерес. Создание белков, содержащих одну или более не встречающуюся в природе аминокислоту фенилселеноцистеин (или одну или в комбинации с другими отличными неестественными аминокислотами), выгодно для множества целей, например, для применения в терапевтических белках и в исследовательских целях.

МОДИФИКАЦИЯ ФЕНИЛСЕЛЕНОЦИСТЕИНА

[0069] В некоторых аспектах, настоящее изобретение обеспечивает способы модификации остатка аминокислоты фенилселеноцистеина после его включения (встраивания) в полипептид. Одна такая модификация, например, окислительное расщепление, может обеспечивать преобразование остатка аминокислоты фенилселеноцистеина в α -, β -ненасыщенную аминокислоту дегидроаланин (см. фиг.1, структура 2). Не встречающаяся в природе аминокислота дегидроаланин также является реакционно-способной и может быть впоследствии модифицирована.

[0070] Не подразумевается, что настоящее изобретение ограничено каким-либо конкретным механизмом (например, окислительного отщепления) или определенными условиями реакции (например, контакт с перекисью водорода) для преобразования фенилселеноцистеина в дегидроаланин. Обычному специалисту в этой области техники будут очевидны разнообразные подходящие альтернативные механизмы и условия реакции, которые находят в настоящем изобретении равное применение для преобразования остатка фенилселеноцистеина в дегидроаланин.

МОДИФИКАЦИЯ ДЕГИДРОАЛАНИНА

[0071] В некоторых аспектах, изобретение обеспечивает способы дальнейшей модификации остатка не встречающейся в природе аминокислоты дегидроаланина в полипептиде. Остаток дегидроаланина является реакционно-способным и может быть мишенью в высокоспецифичных реакциях модификации. Указанные способы особенно полезны в получении конъюгатов липидов с получением липидизированных белков.

[0072] Как описано в настоящей заявке, реакции присоединения Майкла с не встречающейся в природе аминокислотой дегидроаланином приводят к получению белков с запрограммированными сайт-специфичными посттрансляционными модификациями. Например, реакция дегидроаланина с тиолипидом может дать липидизированный белок. Например, реакция с тиопальмитиновой кислотой дает пальмитоилцистеин (см. фиг.1, структура 3), реакция с фарнезилмеркаптаном дает фарнезилцистеин (см. фиг.1, структура 4), и реакция с малонатом дает γ -карбоксиглютамовую кислоту (см. фиг.1, структура 6). Кроме того, реакция с 1-гексадекантиолом дает S-гексадецилцистеин (см. фиг.1, структура 5).

[0073] Хотя в настоящем описании приведены конкретные примеры, это не

означает, что изобретение ограничено каким-либо конкретным механизмом (например, путем присоединения Майкла) или определенными реактивами (тиолипиды) для модификации (конъюгации) дегидроаланина. Обычному специалисту в этой области техники очевидно существование разнообразных подходящих
 5 альтернативных механизмов, условий реакции и реактивов, которые находят равное применение в настоящем изобретении для модификации остатка дегидроаланина. Действительно, модификация остатка дегидроаланина не ограничена конъюгацией липида, и данный механизм конъюгации можно применять для конъюгирования с
 10 полипептидом любого желательного компонента в положении дегидроаланина.

МЕТОДИКА ПРИМЕНЕНИЯ ОРТОГОНАЛЬНОЙ ТРНК/АМИНОАЦИЛ-ТРНК СИНТЕТАЗЫ

[0074] Понимание новых композиций и способов согласно настоящему изобретению требует понимания действий, связанных с парой ортогональной тРНК и ортогональной аминоксил-тРНК синтетазы. Чтобы добавлять в генетический код
 15 дополнительные не встречающиеся в природе аминокислоты, необходимо, чтобы новые ортогональные пары, включающие аминоксил-тРНК синтетазу и подходящую тРНК, могли эффективно функционировать в механизме трансляции хозяина, но в то же время быть «ортогональными по отношению к системе трансляции хозяина, что
 20 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 аминоксил-тРНК
 25 синтетазу, которая предпочтительно аминоксилует (или "нагружает") когнатную тРНК только одной определенной не встречающейся в природе аминокислотой. О-тРНК также обычно не подвергается аминоксилации (или плохо аминоксилуется, то есть, нагружается) эндогенными синтетазами. Например, в
 30 системе хозяина *E.coli* ортогональная пара будет включать аминоксил-тРНК синтетазу, которая не обладает способностью перекрестной реакции ни с одной из эндогенных тРНК (например, *E.coli* существует 401 эндогенных тРНК), и ортогональную тРНК, которую не аминоксилует ни одна из эндогенных синтетаз (которых в *E.coli*, например, двадцать одна).

[0075] 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
 40 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,
 45 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
 50 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

[0076] Ортогональные системы трансляции обычно включают клетки (которые могут быть клетками прокариот, такими как клетки *E.coli*), которые содержат ортогональную тРНК (О-тРНК), ортогональную аминоацил тРНК синтетазу (О-РС), и не встречающуюся в природе аминокислоту, причем О-РС аминоацилирует О-тРНК не встречающейся в природе аминокислотой. Ортогональная пара согласно настоящему изобретению может включать О-тРНК, например, супрессорную тРНК, тРНК, сдвига рамки, или тому подобное, и когнатную О-РС. Ортогональные системы согласно настоящему изобретению обычно могут включать пары О-тРНК/О-РС, либо в клетке-хозяине, либо без клетки-хозяина. В дополнение к многокомпонентным системам, изобретение также обеспечивает отдельные новые компоненты, например, полипептиды новой ортогональной аминоацил-тРНК синтетазы (например, SEQ ID NO:4, 6 или 8), и полинуклеотиды, которые кодируют указанные полипептиды (например, SEQ ID NO:5, 7 или 9).

[0077] 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 аминоацилирует О-тРНК определенной не встречающейся в природе аминокислотой. Нагруженная О-тРНК распознает селекторный кодон и супрессирует (подавляет) остановку трансляции, вызванную селекторным кодоном.

[0078] В некоторых аспектах, О-тРНК согласно настоящему изобретению распознает селекторный кодон и обладает, по меньшей 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 подавления О-тРНК, включающей или кодируемой последовательностью, представленной в списке последовательностей.

[0079] В некоторых вариантах реализации, эффективность супрессии парой О-РС и О-тРНК примерно в 5 раз, 10 раз, 15 раз, 20 раз, или 25 раз или более превосходит эффективность супрессии О-тРНК без О-РС. В некоторых аспектах, эффективность супрессии парой О-РС и О-тРНК составляет, по меньшей мере, около, например, 35%, 40%, 45%, 50%, 60%, 75%, 80%, или 90% или более эффективности супрессии парой ортогональной синтетазы, представленной в списке последовательностей.

[0080] Клетка-хозяин использует пару О-тРНК/ О-РС для встраивания не встречающейся в природе аминокислоты в растущую цепь полипептида, например, с участием нуклеиновой кислоты, которая содержит полинуклеотид, который кодирует полипептид, представляющий интерес, причем указанный полинуклеотид содержит селекторный кодон, распознаваемый О-тРНК. В определенных предпочтительных аспектах, клетка может включать одну или более дополнительную пару О-тРНК/ О-

РС, при этом к дополнительной О-тРНК дополнительная О-РС присоединяет другую не встречающуюся в природе аминокислоту. Например, одна из О-тРНК может распознавать кодоны, содержащие четыре основания, а другая О-тРНК может распознавать стоп-кодон. В качестве альтернативы, одна кодирующая нуклеиновая кислота может содержать множество различных стоп-кодонов или множество различных кодонов, состоящих из четырех оснований.

[0081] Как было отмечено, в некоторых вариантах реализации, в клетке или другой системе трансляции присутствуют многочисленные пары О-тРНК/О-РС, что обеспечивает встраивание более одной не встречающейся в природе аминокислоты в полипептид. Например, клетка может дополнительно содержать дополнительную другую пару О-тРНК/О-РС и вторую не встречающуюся в природе аминокислоту, причем указанная дополнительная О-тРНК распознает второй селекторный кодон, и указанная дополнительная О-РС предпочтительно аминоацилирует О-тРНК второй не встречающейся в природе аминокислотой. Например, клетка, которая содержит пару О-тРНК/О-РС (где О-тРНК узнает, например, янтарный селекторный кодон), может дополнительно включать вторую ортогональную пару, при этом вторая О-тРНК распознает другой селекторный кодон, например, опал-кодон, кодон из четырех оснований, или тому подобное. Желательно, чтобы различные ортогональные пары были получены из различных источников, что может способствовать распознаванию различных селекторных кодонов.

[0082] В определенных вариантах реализации, системы включают клетку, такую как клетка *E.coli*, которая содержит ортогональную тРНК (О-тРНК), ортогональную аминоацил-тРНК синтетазу (О-РС), не встречающуюся в природе 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.

[0083] 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 импортирование 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 объединять.

Ортогональная тРНК (О-тРНК)

[0084] Ортогональная тРНК (О-тРНК) согласно настоящему изобретению желательно опосредует встраивание не встречающейся в природе аминокислоты в белок, кодируемый полинуклеотидом, который содержит селекторный кодон,

распознаваемый данной О-тРНК, например, *in vivo* или *in vitro*. В определенных вариантах, О-тРНК согласно настоящему изобретению обладает, по меньшей мере, около, например, 45%, 50%, 60%, 75%, 80%, или 90% или больше эффективностью супрессии в присутствии когнатой синтетазы в ответ на селекторный кодон, по сравнению с О-тРНК, содержащей или кодируемой последовательностью полинуклеотида, содержащейся в списке последовательностей.

[0085] 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 (derivatизированной) 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™ для анализа β-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.

[0086] Примеры О-тРНКs согласно настоящему изобретению приведены в списке последовательностей, прилагающемся к настоящей заявке, см, например, фиг.2 и SEQ ID NO:1. Настоящее раскрытие также обеспечивает идеи для создания дополнительных эквивалентных молекул О-тРНК. В молекуле РНК, такой как мРНК О-РС, или молекуле О-тРНК, тимин (Т) заменяют на урацил (U) в данной последовательности (или, наоборот, в кодирующей ДНК), или в последовательности, комплементарной данной. Также можно осуществлять дополнительные модификации оснований с получением по существу эквивалентных молекул.

[0087] Изобретение охватывает также консервативные варианты О-тРНК, соответствующие отдельным О-тРНК, описанным в настоящей заявке. Например, консервативные варианты О-тРНК включают молекулы, которые функционируют аналогично отдельным О-тРНК, например, О-тРНК с последовательностями, представленными в списке последовательностей, и которые сохраняют L-подобную структуру тРНК на основе соответствующей самокомплементарности, но не содержат последовательности, идентичной тем, которые, например, приведены в списке последовательностей или на фиг.2, и желательны, отличные от молекул тРНК дикого типа.

[0088] Композиция, содержащая О-тРНК, может дополнительно включать ортогональную аминоксил-тРНК синтетазу (О-АРСазу), причем О-РС предпочтительно аминоксилует данную О-тРНК не встречающейся в природе аминокислотой. В определенных вариантах, композиция, включающая О-тРНК, может дополнительно включать систему трансляции (например, *in vitro* или *in vivo*). Нуклеиновая кислота, которая содержит полинуклеотид, который кодирует представляющий интерес полипептид, причем указанный полинуклеотид содержит селекторный кодон, распознаваемый О-тРНК, либо комбинация из одного или большего количества перечисленных компонентов может присутствовать в клетке.

[0089] Способы создания ортогональной тРНК (О-тРНК) также являются признаком настоящего изобретения. О-тРНК, полученная таким способом, также является признаком настоящего 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-специфичного 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 способов, например, О-тРНК SEQ ID NO: 1.

[0090] 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, вариабельной 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 родства к желаемой АРСазе.

[0091] Способы могут включать анализ сходства (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.

[0092] Как правило, О-тРНК получают подвергая, например, отрицательному отбору популяцию клеток первого вида, где клетки включают одну из множества потенциальных О-тРНК. Отрицательный отбор отсеивает клетки, которые содержат член библиотеки потенциальных О-тРНК, который аминокислирует аминоксил-тРНК синтетаза (АРСаза), которая является эндогенной по отношению к данной клетке. Это обеспечивает множество тРНК, которые являются ортогональными для клетки первого вида.

[0093] В определенных вариантах реализации, при отрицательном отборе в полинуклеотид вводят селекторный кодон (ы), который кодирует маркер отрицательного отбора, например, фермент, который обеспечивает устойчивость к антибиотикам, например, β -лактамазу, фермент, который дает детектируемый продукт, например, β -галактозидазу, хлорамфеникол ацетилтрансферазу (ХАТ), например, токсичный продукт, такой как барназа, в несущественное положение (например, по-прежнему дающее функциональную барназу), и т.д. Скрининг/отбор может проводиться путем выращивания популяции клеток в присутствии агента

отбора (например, антибиотика, такого как ампициллин). В одном варианте реализации, концентрация агента отбора может быть различной.

5 [0094] Например, чтобы измерить активность супрессорных тРНК, применяют систему отбора, которая основана на *in vivo* супрессии селекторного кодона, например, нонсенс-кодона (например, стоп-кодона) или мутаций со сдвигом рамки, введенных в полинуклеотид, который кодирует маркер отрицательного отбора, например, ген β -лактамазы (*bla*). Например, конструируют варианты полинуклеотида, например, варианты *bla*, содержащие селекторный кодон в определенном положении 10 (например, A184). Клетки, например, бактерии, трансформируют этими полинуклеотидами. В случае ортогональной тРНК, которую не могут эффективно аминоацилировать эндогенные синтетазы *E. coli*, устойчивость к антибиотикам, например, устойчивость к ампициллину, должна быть приблизительно равна или меньше устойчивости бактерий, не трансформированных плазмидой. Если тРНК не 15 ортогональна, или если в системе происходит коэкспрессия гетерологичной синтетазы, способной аминоацилировать тРНК, наблюдают более высокий уровень устойчивости к антибиотикам, например, ампициллину. Отбирают клетки, например, бактерии, которые не способны расти на среде Лурия-Бертани (LB), содержащей антибиотика в концентрациях, приблизительно равных таковой для клеток, не трансформированных 20 плазмидами.

[0095] В случае токсичного продукта (например, рибонуклеазы или барназа), если одну из множества потенциальных тРНК аминоацилируют эндогенные синтетазы 25 хозяина, например, *Escherichia coli* (то есть, она не является ортогональной для синтетазы 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.

30 [0096] В одном варианте реализации, пул тРНК, которые являются ортогональными для желаемого организма, подвергают положительному отбору, при котором селекторный кодон вводят в положительный маркер отбора, например, кодирующий ген устойчивости к лекарственному средству, такой как ген β -лактамазы. Положительный отбор выполняют для клеток, содержащих полинуклеотид, 35 кодирующий или содержащий член пула тРНК, которые являются ортогональными для клетки, полинуклеотид, кодирующий положительный маркер отбора, и полинуклеотид, кодирующий когнатную АРСазу. В определенных вариантах реализации, вторая популяция клеток содержит клетки, которые не были отсеяны в ходе отрицательного отбора. Полинуклеотиды экспрессируются в клетках, а клетки 40 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 45 не могут быть эффективно распознаны синтетазой, представляющей интерес. Клетка, содержащая нефункционирующие тРНК или тРНК, которые синтетазы, представляющая интерес, не может эффективно распознавать, является чувствительной к антибиотику. Поэтому тРНК: (i) не являются субстратами для 50 эндогенных синтетаз хозяина, например, *Escherichia coli*, (ii) их может аминоацилировать представляющая интерес синтетазы; и (iii) участвуют в трансляции, переживают оба отбора.

[0097] Соответственно, один и тот же маркер может быть либо положительным, либо отрицательным маркером, в зависимости от условий, в которых проводят скрининг. Таким образом, маркер является положительным маркером в случае скрининга проводится на, но отрицательным маркером в случае скрининга проводится для отсева (против).

[0098] 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 желательна 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.

[0099] 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 флуоресцирует 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 USA. 90:10444-8.

[0100] 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 AMINOACYL-tRNA SYNTHETASE PAIRS"; WO 2004/094593, озаглавленной "EXPANDING THE EUKARYOTIC 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 tРНК recognition of a tРНК synthetase by a single amino acid change, PNAS 100(10):5676-5681.

Ортогональная синтетаза аминоксил-тРНК (О-РС)

[0101] О-РС согласно настоящему изобретению избирательно аминоксилует О-тРНК не встречающейся в природе аминоксилотой, *in vitro* или *in vivo*. О-РС согласно настоящему изобретению можно обеспечить в системе трансляции, например, в клетке,

в виде полипептида, который содержит О-РС, и/или полинуклеотида, который кодирует О-РС, или ее часть. Например, пример О-РС содержит последовательность аминокислот, которая обозначена SEQ ID NO:4, 6 или 8, или ее консервативный вариант. В другом примере, О-РС, или ее часть, кодирует последовательность полинуклеотида, которая кодирует последовательность аминокислот из списка последовательностей или примеров, приведенных в настоящей заявке, или комплементарная им последовательность полинуклеотида. См., например, полинуклеотид SEQ ID NO:5, 7 или 9.

[0102] Способы идентификации ортогональной аминоксил-тРНК синтетазы (О-РС), например, О-РС, для применения с О-тРНК, являются признаком настоящего изобретения. Например, способ включает отбор, например, положительный отбор, среди популяции клеток первого вида, при котором клетки индивидуально включают: 1) член множества аминоксил-тРНК синтетаз (АРС), (например, множество АРС может включать мутантные АРС, АРС, полученные из любого вида, за исключением первых видов, или и мутантные АРС, и АРС, полученные из любого вида, кроме первых видов); 2) ортогональную тРНК (О-тРНК) (например, из одного или более видов); и 3) полинуклеотид, который кодирует маркер отбора (например, положительного) и содержит, по меньшей мере, один селекторный кодон. Клетки, которые выбирают в результате отбора или скрининга, представляют собой клетки, которые демонстрируют повышенную эффективность супрессии в сравнении с клетками, не содержащими данный член множеств АРС или содержащими его в меньшем количестве. Эффективность супрессии может быть измерена способами, известными в данной области техники или описанными в настоящей заявке. Клетки, обладающие повышенной эффективностью супрессии, включают активную РС, которая аминокислирует О-тРНК. Уровень аминокислирования (in vitro или in vivo) активной АРСазой первого набора тРНК из первого вида сравнивают с уровнем аминокислирования (in vitro или in vivo) активной АРСазой второго набора тРНК из второго вида. Уровень аминокислирования можно определить с помощью детектируемого вещества (например, меченой не встречающейся в природе аминокислоты). Обычно отбирают активную АРС, которая более эффективно аминокислирует второй набор тРНК, чем первый набор тРНК, что обеспечивает эффективную (оптимизированную) ортогональную аминоксил-тРНК синтетазу для применения с О-тРНК. О-РС, идентифицированная способом согласно настоящему изобретению, также является признаком настоящего изобретения.

[0103] Для определения аминокислирования можно применять любой из множества подходов. 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, экспрессирующей полинуклеотид, содержащей, 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 EUKARYOTIC GENETIC CODE."

[0104] Идентифицированную О-РС можно дополнительно подвергать манипуляциям, позволяющим изменить субстратную специфичность синтетазы, таким образом, чтобы О-РС присоединяла к О-тРНК только желаемую не встречающуюся в природе аминокислоту, но не присоединяла ни одну из 20 стандартных аминокислот. Способы получения ортогональной аминоксил-тРНК синтетазы с субстратной

специфичностью не встречающейся в природе аминокислоты включают осуществление мутации синтетазы, например, в активном центре синтетазы, в сайте механизма редактирования синтетазы, на других участках, комбинацию мутаций в различных сайтах синтетаз, или подобные способы, а также применение отбора.

5 Применяемая стратегия основана на комбинации положительного отбора с последующим отрицательным отбором. В положительном отборе супрессия селекторного кодона, введенного в несущественном положении (ях) маркера положительного отбора, позволяет клеткам выживать под давлением
10 положительного отбора. Таким образом, в присутствии природных, а также в присутствии не встречающихся в природе аминокислот, выжившие клетки кодируют активные синтетазы, аминоацилирующие ортогональной супрессорной тРНК природной или не встречающейся в природе аминокислотой. При отрицательном
15 отборе, супрессия селекторного кодона, введенного в несущественном положении (ях) отрицательного маркера, исключает синтетазы, специфичные к природным аминокислотам. Клетки, выжившие при отрицательном и положительном отборе, кодируют синтетазы, которые аминоацилируют (нагружают) ортогональную супрессорную тРНК исключительно не встречающейся в природе аминокислотой.
20 Затем эти синтетазы можно подвергнуть дальнейшим мутациям, например, перетасовыванию ДНК или другим рекурсивным методам мутагенеза.

[0105] Библиотеку мутантных О-РС можно получить при помощи различных методов мутагенеза, известных в данной области. Например, мутантную АРС можно
25 получить путем осуществления сайт-специфичных мутаций, случайных точечных мутаций, гомологичной рекомбинации, перетасовки ДНК или другими рекурсивными методами мутагенеза, с применением химерных конструкции или при помощи любой комбинации указанных способов. Например, библиотеку мутантной АРС можно
30 получить из двух или более других, например, меньших, менее разнообразных "суббиблиотек". Химерные библиотеки АРС также включены в объем изобретения. Нужно отметить, что для ортогональных пар могут быть сконструированы и скринированы библиотеки тРНК-синтетаз различных организмов (например, микроорганизмов, таких как зубактерии или архебактерии), такие как библиотеки,
35 которые включают естественное разнообразие (см., например, патент США №6,238,884 Short et al; U.S. Patent No. 5,756,316 to Schallenger et al; патент США №5,783,431, авторы Petersen et al; U.S. Patent No. 5,824,485 Thompson et al; патент США №5,958,672 Short et al).

[0106] После того, как синтетазы подвергли положительному и отрицательному
40 отбору/скринингу, их можно дополнительно подвергнуть дальнейшему мутагенезу. Например, можно выделить нуклеиновую кислоту, которая кодирует О-РС; из нуклеиновой кислоты можно получить ряд полинуклеотидов, которые кодируют мутантную О-РС (например, путем неспецифического мутагенеза, сайт-специфического
45 мутагенеза, рекомбинации или любой комбинацией указанных методов), эти отдельные шаги или комбинацию этих шагов можно повторять до получения мутантной О-РС, избирательно аминоацилирующей О-тРНК не встречающейся в природе аминокислотой. В некоторых аспектах согласно настоящему изобретению, указанные шаги выполняют многократно, например, по меньшей мере, два раза.

50 [0107] 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 могут 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.

[0108] Дополнительные общие подробности получения О-РС и изменения специфичности субстрата 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.

ИСХОДНЫЕ ОРГАНИЗМЫ И ОРГАНИЗМЫ-ХОЗЯЕВА

[0109] 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.

[0110] Например, ортогональная О-тРНК может быть получена из организма Archaea, например, archaeobacterium, такого как *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, такого как *Haloferax volcanii* и *Halobacterium species NRC-1*, *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, *Pyrococcus horikoshii*, *Aeuryopyrum pernix*, *Methanococcus maripaludis*, *Methanopyrus kandleri*, *Methanosarcina mazei* (Мм), *Pyrobaculum aerophilum*, *Pyrococcus abyssi*, *Sulfolobus solfataricus* (Ss), *Sulfolobus tokodaii*, *Thermoplasma acidophilum*, *Thermoplasma volcanium*, или тому подобных, или эубактерий, таких как *Escherichia coli*, *Thermus thermophilus*, *Bacillus subtilis*, *Bacillus stearothermophilus*, или тому подобных, в то время как ортогональная О-РС может быть получена из организма или комбинации организмов, например, archaeobacterium, таких как *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Halobacterium*, таких как *Haloferax volcanii* и *Halobacterium species NRC-1*, *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, *Pyrococcus horikoshii*, *Aeuryopyrum pernix*, *Methanococcus maripaludis*, *Methanopyrus kandleri*, *Methanosarcina mazei*, *Pyrobaculum aerophilum*, *Pyrococcus abyssi*, *Sulfolobus solfataricus*, *Sulfolobus tokodaii*, *Thermoplasma acidophilum*, *Thermoplasma volcanium*, или подобного, или eubacterium, типа *Escherichia coli*, *Thermus thermophilus*, *Бацилла subtilise* *Бацилла stearothermophilus*, или тому подобных. В одном варианте реализации, источники-эукариоты, например, растения, водоросли, протисты, грибы, дрожжи, животные (например, млекопитающие, насекомые, членистоногие, и т.д.), или тому подобные, можно также применять в качестве источников О-тРНК и О-РС.

[0111] Отдельные компоненты пары О-тРНК/О-РС могут быть получены из одного и того же организма или различных организмов. В одном варианте реализации, пару О-тРНК/О-РС получают из одного и того же организма. В альтернативном варианте,

О-тРНК и О-РС пары О-тРНК/О-РС получают из различных организмов.

[0112] О-тРНК, О-РС или пара О-тРНК/О-РС могут быть подвергнуты отбору или скринингу *in vivo* или *in vitro* и/или применяться в клетке, например, зубактериальной

5 клетке, с получением полипептида с не встречающейся в природе аминокислотой. Применение клеток зубактерий не ограничивается, например, *Escherichia coli*, *Thermus thermophilus*, *Bacillus subtilis*, *Bacillus stearothermophilus*, или подобными клетками. Композиции клеток зубактерий, содержащих компоненты трансляции согласно

10 настоящему изобретению, также являются признаком настоящего изобретения. [0113] 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.

15 [0114] Хотя ортогональные системы трансляции (например, включающие О-РС, О-тРНК и не встречающуюся в природе аминокислоту) можно применять в культивируемых клетках-хозяевах для синтеза белков, содержащих не встречающиеся в природе аминокислоты, это не означает, что ортогональная система трансляции согласно настоящему изобретению требует интактной, жизнеспособной клетки-хозяина. Например, ортогональную систему трансляции можно применять с

20 бесклеточной системой в присутствии экстракта клеток. Действительно, применение бесклеточной *in vitro* системы транскрипции/трансляции для синтеза белков является хорошо разработанной методикой. Адаптация этих систем *in vitro* для синтеза белков, содержащих не встречающиеся в природе аминокислоты, с применением

25 ортогональных компонентов систем трансляции, описанных в настоящей заявке, находится в рамках данного изобретения.

СЕЛЕКТОРНЫЕ КОДОНЫ

[0115] Селекторные кодоны согласно настоящему изобретению расширяют объем

30 кодонов генетического кода механизма биосинтеза белков. Например, селекторные кодоны включает, например, уникальные кодоны, содержащие три основания, нонсенс-кодона, 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

35 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

40 встречающихся в природе аминокислот, например, включение, по меньшей мере, одной не встречающейся в природе аминокислоты, с использованием этих различных селекторных кодонов.

[0116] В одном варианте реализации, способы включают применение селекторного кодона, который является терминирующим кодоном, для встраивания не

45 встречающейся в природе аминокислоты *in vivo* в клетке. Например, получают О-тРНК, которая распознает терминирующий кодон и которую О-РС аминоацилирует не встречающейся в природе аминокислотой. Эту О-тРНК не распознают природные аминоацил-тРНК синтетазы хозяина. Для введения терминирующего кодона в представляющем интерес сайте в полинуклеотид, кодирующий представляющий

50 интерес полипептид, можно применять обычный сайт-направленный мутагенез. См., например, Sayers, J.R., et al. (1988), 5',3' Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis. *Nucleic Acids Res.* 791-802. При объединении,

например, *in vivo*, O-PC, O-тРНК и нуклеиновой кислоты, которая кодирует полипептид, представляющий интерес, происходит встраивание, не встречающаяся в природе аминокислота встраивается в ответ на терминирующий кодон с образованием полипептида, содержащего не встречающуюся в природе аминокислоту в указанном
 5 положении. В одном варианте реализации настоящего изобретения, стоп-кодон, применяемый в качестве селекторного кодона, представляет собой амбер-кодон, UAG, и/или опал-кодон, UGA. В одном примере, генетический код, в котором и UAG, и UGA применяют в качестве селекторных кодонов, может кодировать 22 аминокислоты, при
 10 этом сохраняется охра-нонсенс-кодон UAA, который является самым распространенным сигналом терминации.

[0117] Встраивание не встречающихся в природе аминокислот *in vivo* можно осуществить без значительного вмешательства в клетку-хозяина. Например в неэукариотических клетках, таких как *Escherichia coli*, по причине того, что
 15 эффективность супрессии кодона UAG зависит от конкуренции между O-тРНК, например, тРНК, являющейся супрессором амбер-кодона, и фактора высвобождения 1 (RF1) (который связывается с кодоном UAG и инициирует высвобождение растущего пептида из рибосомы), эффективность супрессии может
 20 модулировать, например, увеличение уровня экспрессии O-тРНК, например, супрессорной тРНК, или использование дефицитного по RF1 штамма. Поскольку в эукариотических клетках (клетках эукариот) эффективность супрессии кодона UAG зависит от конкуренции между O-тРНК, например, тРНК, являющейся супрессором амбер-кодона, и фактора высвобождения эукариот (например, eRF), который
 25 связывается с терминирующим кодоном и инициирует высвобождение растущего пептида из рибосомы, эффективность супрессии можно модулировать, например, увеличением уровня экспрессии O-тРНК, например, супрессорной тРНК. Кроме того, могут присутствовать дополнительные соединения, например восстановители, такие как дитиотретиол (DTT).
 30

[0118] Не встречающиеся в природе аминокислоты могут кодироваться редкими кодонами. Например, оказалось, что если концентрация аргинина в реакции синтеза белка *in vitro* уменьшена, редкий кодон аргинина, AGG, обеспечивает встраивание аланина синтетической тРНК, ацилированной аланином. См., например, Ma et al.,
 35 *Biochemistry*, 32:7939 (1993). В этом случае, синтетическая тРНК конкурирует с природной тРНК^{Arg}, которая существует как минорная разновидность в *Escherichia coli*. Кроме того, некоторые организмы не используют все триплетные кодоны. В *Micrococcus luteus* кодон AGA, не имеющий «присвоенной» аминокислоты,
 40 применяли для встраивания аминокислот в *in vitro* экстракте транскрипции/трансляции. См., например, Kowal and Oliver, *Nucl. Acid. Res.*, 25:4685 (1997). Согласно настоящему изобретению, могут быть получены компоненты для применения указанных редких кодонов *in vivo*.

[0119] Селекторные кодоны могут также включать кодоны с увеличенным числом оснований, например, кодоны, содержащие четыре или более оснований, например, четыре, пять, шесть или более оснований. Примеры кодонов, состоящих из четырех оснований, включают, например, AGGA, CUAG, UAGA, CCCU, и т.п. Примеры кодонов, состоящих из пяти оснований, включают, например, AGGAC, CCCCU, CCCUC, CUAGA, CUACU, UAGGC и т.п. Способы согласно настоящему изобретению
 45 включают применения кодонов с большим числом оснований, для супрессии сдвига рамки считывания. С помощью кодонов, состоящих из четырех или более оснований, можно встраивать в один белок, например, одну или более не встречающихся в
 50

природе аминокислот. В других вариантах реализации, антикодоновые петли могут расшифровать, 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, содержащих четыре
 5 основания, применение кодонов, содержащих четыре или более оснований, позволяет закодировать в одной клетке множество не встречающихся в природе аминокислот. См. также, Anderson et al., (2002) Exploring the Limits of Codon and Anticodon Size, Chemistry and Biology, 9:237-244; и, Magliery, (2001) Expanding the Genetic Code: Selection of
 10 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.

[0120] Например, кодоны из четырех оснований применяли для встраивания не встречающихся в природе аминокислот в белки, с применением методов
 биосинтеза *in vitro*. См., например, Ma et al., (1993) Biochemistry, 32:7939; and Hohsaka
 15 et al., (1999) J. Am. Chem. Soc., 121:34. CGGG и AGGU применяли для одновременного встраивания 2-нафтилаланина и NBD-производного лизина в стрептавидин *in vitro* с помощью двух химически ацилированных тРНК, супрессирующих сдвиг рамки. См., например, Hohsaka et al., (1999) J Am. Chem. Soc., 121:12194. Moore и др. исследовали *in*
 20 *vivo* способность производных тРНК^{Leu} с антикодонами NCUA супрессировать кодоны UAGN (N может быть U, A, G, или C) и обнаружили, что тРНК^{Leu} с антикодоном UCUA может декодировать квадруплетный кодон UAGA с эффективностью 13-26%, при слабом декодировании в положении рамки 0 или -1. См.
 25 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
 30 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
 35 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
 40 аминокислот О-РС.

[0121] В данной системе, селекторный кодон может также включать один из естественных трехосновных кодонов, в случае если эндогенная система не использует (или редко использует) естественный кодон. Примеры включают систему, в которой
 45 отсутствует тРНК, которая распознает естественные трехосновные кодоны, и/или систему, в котором трехосновный кодон представляет собой редкий кодон.

[0122] Селекторные кодоны могут включать не встречающиеся в природе пары азотистых оснований. Эти не встречающиеся в природе пары азотистых оснований дополнительно расширяют существующий генетический алфавит. Одна
 50 дополнительная пара оснований увеличивает количество триплетных кодонов с 64 до 125. Свойства третьих пар оснований включают устойчивое и селективное спаривание оснований, эффективное ферментативное встраивание в ДНК полимеразой с высокой точностью, и эффективное непрерывное удлинение праймера

после синтеза новой не встречающейся в природе пары оснований. Описания не встречающихся в природе пар x оснований, которые можно адаптировать к способам и композициям, содержатся например, в Hirao, 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 релевантные публикации указаны ниже.

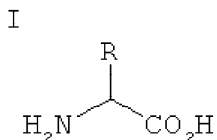
[0123] Не встречающийся в природе нуклеозид для применения *in vivo* обладает способностью проникать через мембрану и находится в фосфорилированной форме с образованием соответствующего трифосфата. Кроме того, увеличенная генетическая информация устойчива и не подвержена разрушению ферментами клетки. 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 I (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 способна 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 полимеразы, которую можно применять для репликации пары оснований PICS. Кроме того, возможна репликация пары оснований 7AI. 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.

[0124] Также для встраивания не встречающейся в природе аминокислоты в желательный полипептид можно применять обходную (bypassing) систему трансляции. В обходной системе трансляции в ген встраивают большую последовательность, но ее трансляция в белок не происходит. Последовательность содержит структуру, которая служит сигналом, который индуцирует сдвиг рибосомы по последовательности и возобновление трансляции далее по ходу трансляции от вставки.

НЕ ВСТРЕЧАЮЩИЕСЯ В ПРИРОДЕ АМИНОКИСЛОТЫ

[0125] В настоящем описании термин «не встречающаяся в природе аминокислота»

относится к любой аминокислоте, модифицированной аминокислоте или аналогу аминокислоты, отличным от селеноцистеина и/или пирролизина и следующих двадцати генетически кодируемых альфа-аминокислот: аланина, аргинина, аспарагина, аспарагиновой кислоты, цистеина, глутамина, глутаминовой кислоты, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина, валина. Общую структура альфа-аминокислот иллюстрирует Формула I:



[0126] Не встречающаяся в природе аминокислота - это, обычно, любая структура, которая соответствует Формуле I, где группа R - это любой заместитель, отличный от заместителей, присутствующих в двадцати природных аминокислотах. См. например *Biochemistry* by L. Stryer, 3rd ed. 1988, Freeman and Company, New York, где приведены структуры двадцати естественных аминокислот. Обратите внимание, что не встречающиеся в природе аминокислоты согласно настоящему изобретению могут быть естественно встречающимися соединениями, исключая двадцать альфа-аминокислот, упомянутых выше.

[0127] Поскольку не встречающиеся в природе аминокислоты согласно настоящему изобретению обычно отличаются от природных аминокислот боковой цепью, не встречающиеся в природе аминокислоты образуют амидные связи с другими аминокислотами, например, природными или не встречающимися в природе, так же как и во встречающихся в природе белках. Однако не встречающиеся в природе аминокислоты имеют группы боковой цепи, которые отличают их от естественных аминокислот.

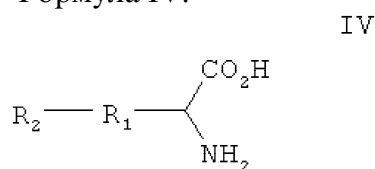
[0128] Особенный интерес для настоящего изобретения представляет не встречающаяся в природе аминокислота фенилселеноцистеин (см. фиг.1, структура 1). В представляющий интерес полипептид одновременно с не встречающейся в природе аминокислотой фенилселеноцистеином можно встраивать другие не встречающиеся в природе аминокислоты, например, с использованием подходящей второй пары O-PC/O-гРНК вместе с ортогональной парой, обеспеченной в соответствии с настоящим изобретением. Известно много таких дополнительных не встречающихся в природе аминокислот и подходящих ортогональных пар. См. настоящее описание и цитируемые в нем источники. 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); и Wang et al., "Expanding the Genetic Code," *Annu. Rev. Biophys. Biomol. Struct.*, 35:225-249 (2006); содержание каждого из источников полностью включено в настоящее описание посредством ссылки.

[0129] Хотя не встречающаяся в природе аминокислота фенилселеноцистеин, показанная на фиг.1, структура 1, представляет главный интерес в описанных здесь примерах, это не означает, что настоящее изобретение строго ограничено этой структурой. Действительно, можно легко получить разнообразные легкомодифицируемые, структурно близкие аналоги, которые сохраняют принципиальные особенности фенилселеноцистеина, показанного на фиг.1, структура 1, и также специфично распознаваемые аминокислот-тРНК синтетазами

согласно настоящему изобретению (например, O-PC SEQ ID NO:4, 6 и 8). Это означает, что данные родственные аналоги-аминокислоты входят в объем настоящего изобретения.

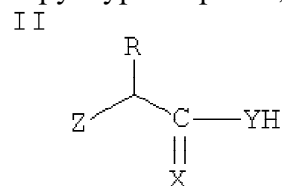
[0130] В других не встречающихся в природе аминокислотах, например, R в
 5 Формуле I возможно включает алкил-, арил-, ацил-, гидразин, циано-, гало-, гидразин, алкенил, эфир, борат, боронат, фосфо, фосфоно, фосфин, енон, имин, сложный эфир, гидроксилламин, амин, и т.п., или любые из перечисленных групп. Другие не
 10 встречающиеся в природе аминокислоты, представляющие интерес, включают, без ограничения, 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
 15 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
 20 полиэфир, аминокислоты с замещающим тяжелым атомом, химически расщепляемые или фоторасщепляемые аминокислоты, аминокислоты с боковой цепью, удлиненной по сравнению с природными аминокислотами (например, полиэфирными или длинными углеводородными цепями, например, более приблизительно 5, более приблизительно 10 атомов углерода, и т.д.), аминокислоты, содержащие связанный с
 25 атомом углерода сахар, аминокислоты, содержащие аминокислоты, содержащие один или более токсичных фрагментов.

[0131] В другом аспекте, изобретение обеспечивает не встречающиеся в природе аминокислоты с общей структурой, которую иллюстрирует показанная ниже
 30 Формула IV:

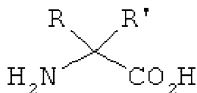


[0132] Не встречающаяся в природе аминокислота, имеющая такое строение, - это, обычно, любая конструкция, в которой R1 представляет собой заместитель, присутствующий в одной из двадцати естественных аминокислот (например, тирозин или фенилаланин), и R2 - заместитель. Таким образом, этот тип не встречающейся в
 40 природе аминокислоты можно рассматривать как производное природной аминокислоты.

[0133] В дополнение к не встречающимся в природе аминокислотам, которые содержат структуру фенилселеноцистеина, показанную на фиг.1, структура 1, не
 45 встречающиеся в природе аминокислоты могут также содержать измененные структуры каркаса, например, как показано структурами Формулы:



III



5 где Z обычно содержит OH, NH₂, SH, NH-R', или S-R'; X и Y, которые могут быть одинаковыми или разными, обычно включают S или O, и R и R', которые могут быть одинаковыми или разными, обычно выбраны из списка компонентов для группы R, описанной выше для не встречающихся в природе аминокислот, имеющих Формулу I, или могут быть водородом. Например, не встречающиеся в природе аминокислоты согласно настоящему изобретению могут содержать заместители в amino- или карбоксильных группах, как показано в Формулах II и III. Не встречающиеся в природе аминокислоты этого типа включают, без ограничения, α-оксикислоты, α-тиоациды α-амиотиокарбоксилаты, например, с боковыми цепями, соответствующими общим двадцати природным аминокислотам или не встречающимся в природе боковым цепям. Кроме того, замены у α-атома углерода могут включать: L, D, или α-α-двузамещенные аминокислоты, такие как D-глутамат, D-аланин, D-метил-О-тирозин, аминокислотная кислота, и т.п. Другие структурные варианты включают циклические аминокислоты, такие как аналоги пролина и 3-, 4-, 6-, 7-, 8-, и 9-членные циклические аналоги пролина, β- и γ-замещенные аминокислоты, такие как β-аланин и γ-аминомасляная кислота.

[0134] В некоторых аспектах, в настоящем изобретении применяют не встречающиеся в природе аминокислоты в L-конфигурации. Однако это не означает, что изобретение ограничено применением L-конфигурации не встречающихся в природе аминокислот. Предполагается, что D-энантиомеры этих не встречающихся в природе аминокислот можно применять в настоящем изобретении.

[0135] Не встречающиеся в природе аминокислоты, находящие применение в настоящем изобретении, не ограничены строго не встречающейся в природе аминокислотой фенилселеноцистеином, показанной на фиг.1, структура 1. Любой специалист в данной области поймет, что можно легко получить широкое разнообразие не встречающихся в природе аналогов естественно встречающихся аминокислот. В качестве примера, 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, гидроксамин, тиоловую группу, карбоксигруппу, изопропиловую группу, метиловую группа, C6-C20 углеводородные линейные или разветвленные цепи, насыщенный или ненасыщенный углеводород, группу O-метила, полиэфирную группу, нитро группу или подобные группы. Кроме того, предусмотрены арильные кольца со множеством заместителей. Аналоги глутамина согласно настоящему изобретению включают, без ограничения, α-гидрокси производные, γ-замещенные производные, циклические производные и амид-замещенные производные глутамина. Приведенные в качестве примера аналоги фенилаланина включают, без ограничения: пара-замещенные фенилаланины, орто-замещенные фенилаланины и мета-замещенные фенилаланины, в которых заместитель содержит алкинильную группу, гидроксигруппу, 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 аминокислоту, нитробензил-серин, О-(2-нитробензил)-L-тирозин, р-карбоксиметил-L-фенилаланин, р-циано-L-фенилаланин, m-циано-L-фенилаланин, бифенилаланин, 3-амино-L-тирозин, бипиридил аланин, р-(2-амино-1-гидроксиэтил)-L-фенилаланин, р-изопропилтиокарбонил -L-фенилаланин, 3-нитро-L-тирозин и р-нитро-L-фенилаланин. Кроме того, р-пропаргиллоксифенилаланин, 3,4-дигидрокси-L-фенилаланин (DHP), 3,4,6-тригидрокси-L-фенилаланин, 3,4,5-тригидрокси-L-фенилаланин, 4-нитро-фенилаланин, р-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-GlcNAc β -серин, L-дигидроксифенилаланин, фторированный фенилаланин, изопропил-L-фенилаланин, р-азидо-L-фенилаланин, р-ацил-L-фенилаланин, р-бензоил-L-фенилаланин, L-фосфосерин, фосфоносерин, фосфонотирозин, р-йодо-фенилаланин, р-бромфенилаланин, р-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 не встречающихся в природе аминокислот

[0136] 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 γ -Dipeptides of Glutamic Acid from Phthylated Intermediates. J. Chem. Soc., 3315-3319; Friedman and Chatterrji (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 Pipecolates from L-Asparagine. Application to the Total Synthesis of (+)-Apovincamine through Amino Acid Decarboxylation and Iminium Ion Cyclization. J. Org. Chem. 1989:1859-1866; Barton et al., (1987) Synthesis of Novel α -Amino-Acids and Derivatives Using Radical Chemistry: Synthesis of L- and D- α -Amino-Adipic Acids, L- α -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 не встречающихся в природе аминокислот клетками

[0137] Поглощение не встречающейся в природе 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 α -аминокислот наводит на мысль, что эти соединения вряд ли будут проникать в клетку. Природные аминокислоты транспортируются в клетку рядом систем транспорта на 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 не встречающихся в природе аминокислот

[0138] В клетке изначально существует много биосинтетических путей продукции аминокислот и других соединений. Хотя может не существовать путей биосинтеза определенных не встречающихся в природе аминокислот, 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 ρ -аминофенилаланина (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.

[0139] Действительно, любые разнообразные способы можно применять для получения новых ферментов для применения в путях биосинтеза, для изменения существующих возбуждающих путей, для продукции не встречающихся в природе аминокислот, *in vitro* или *in vivo*. 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 (suffling) 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

5 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

10 или синтетазы) могут также быть получены с использованием метода рекомбинации ДНК, известного как "постепенное уечение 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

15 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

20 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

25 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, аминокислоту, которая дает функциональный белок. Примеры таких процедур, которые можно адаптировать для

30 данного изобретения, для получения новых ферментов синтеза не встречающихся в природе аминокислот (или новых синтетазы) 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 похожих на исходную

35 последовательность (doped от dope - ghbvctm) 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;

40 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, который можно применять для получения ферментов и/или компонент пути, которые затем можно

45 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.

50 [0140] Альтернатива таким способам мутагенеза включает рекомбинацию полных геномов организмов и отбор полученного потомства для функции конкретных путей (часто называемый "перетасовкой целого генома" (whole genome shuffling)). Этот подход можно применять в данном изобретении, например, путем геномной

рекомбинации и отбора организма (например, E.coli или другой клетки) по способности продуцировать не встречающуюся в природе аминокислоту (или ее промежуточное соединение). Например, в разработке способов развития существующих и/или новых путей в клетках для продуцирования не встречающихся в природе аминокислот можно применять способы, изложенные в следующих публикациях: *in vivo*: 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* 415(6872):644-646.

[0141] 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 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* 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) "Rapid Evolution of Novel Traits in Microorganisms" *Applied and Environmental Microbiology*, 67:3645, и многие другие.

[0142] Независимо от применяемого способа, обычно, не встречающаяся в природе аминокислота, полученная с помощью искусственного пути биосинтеза согласно настоящему изобретению, продуцируется в концентрации, достаточной для эффективного биосинтеза белка, например, в количестве, в которой она продуцируется в клетке в природе, но не в такой степени, чтобы значительно повлиять на концентрацию других аминокислот клетки или исчерпать ресурсы клетки. Обычные концентрации, получаемые *in vivo* этим способом, составляют от приблизительно 10 мм до приблизительно 0.05 мм. После конструирования клетки для синтеза ферментов, желательных для определенного пути, и синтеза не встречающейся в природе аминокислоты можно применить отбор *in vivo* для дальнейшей оптимизации продукции не встречающейся в природе аминокислоты и на синтез белка рибосом и рост клетки.

Ортогональные компоненты для встраивания не встречающихся в природе аминокислот

[0143] Изобретение обеспечивает композиции и способы получения ортогональных компонентов для встраивания не встречающейся в природе аминокислоты фенилселеноцистеина (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 можно применять для встраивания не встречающейся в природе аминокислоты в растущую цепь полипептида.

[0144] Композиция согласно настоящему изобретению содержит ортогональную аминоацил-тРНК синтетазу (О-РС), причем указанная О-РС избирательно аминоацилирует О-тРНК фенилселеноцистеином. В определенных вариантах реализации, О-РС содержит последовательность аминокислоты, содержащую SEQ ID NO:4, 6 или 8, и ее консервативные варианты. В определенных вариантах реализации согласно настоящему изобретению, О-РС избирательно аминоацилирует О-тРНК

более предпочтительно, чем любую эндогенную тРНК, определенную не встречающейся в природе аминокислотой, причем указанная О-РС имеет тенденцию взаимодействовать с О-тРНК, и где соотношение О-тРНК, нагруженной не 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.

[0145] Композиция, которая содержит О-РС, может дополнительно включать ортогональную тРНК (О-тРНК), где О-тРНК узнает селекторный кодон. Как правило, О-тРНК согласно настоящему изобретению обладает, по меньшей 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 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 отсутствие О-РС, например, в 5 раз, 10 раз, 15 раз, 20 раз, 25 раз или более. В некоторых аспектах, эффективность супрессии О-РС и О-тРНК вместе составляет, по меньшей мере, 45% эффективности супрессии ортогональной пары синтетазы тирозил-тРНК, полученной из *Methanococcus jannaschii*.

[0146] 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.

[0147] Клетка (например, эубактериальная клетка или дрожжевая клетка), содержащая систему трансляции, также предусмотрена настоящим изобретением, причем система трансляции включает ортогональную тРНК (О-тРНК); ортогональную аминоацил-тРНК синтетазу (О-РС); и не встречающуюся в природе аминокислоту фенилселеноцистеин. Как правило, О-РС аминоацилирует 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 аминоацилирует 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, 6, 8 или 10, либо консервативный вариант одной из указанных последовательностей.

[0148] Клетка согласно настоящему изобретению может дополнительно включать другую дополнительную пару О-тРНК/О-РС и вторую не встречающуюся в природе аминокислоту, при этом, например, О-тРНК распознает второй селекторный кодон и О-РС предпочтительно аминоацилирует соответствующую О-тРНК второй не встречающейся в природе аминокислотой, при этом вторая аминокислота отличается от первой не встречающейся в природе аминокислоты. Клетка согласно настоящему изобретению может включать нуклеиновую кислоту, которая содержит

полинуклеотид, который кодирует представляющий интерес полипептид, причем указанный полинуклеотид содержит селекторный кодон, распознаваемый указанной О-тРНК.

5 [0149] В определенных вариантах реализации, клетка согласно настоящему изобретению - это эубактериальная клетка (такая, как E.coli), которая содержит ортогональную тРНК (О-тРНК), ортогональную синтетазу аминоксил-тРНК (О-РС), не встречающуюся в природе аминокислоту, и нуклеиновую кислоту, которая
10 содержит полинуклеотид, кодирующий представляющий интерес полипептид, при этом указанный полинуклеотид содержит селекторный кодон, распознаваемый указанной О-тРНК. В определенных вариантах реализации настоящего изобретения, О-РС предпочтительно аминоксилует О-тРНК не встречающейся в природе аминокислотой с эффективностью, превышающей эффективность, с которой О-РС аминоксилует любую эндогенную тРНК.

15 [0150] В определенных вариантах реализации настоящего изобретения, О-тРНК согласно настоящему изобретению содержит или кодируется последовательностью полинуклеотида, которая представлена в списке последовательностей (например, SEQ ID NO:1) или приведенных в настоящем описании примерах, или
20 последовательностью, комплементарной такому полинуклеотиду. В определенных вариантах реализации настоящего изобретения, О-РС содержит последовательность аминокислот, которая представлена в списке последовательностей, или ее консервативный вариант. В одном варианте реализации, О-РС или ее часть закодированы последовательностью полинуклеотидов, кодирующей
25 последовательность аминокислот, которая представлена в списке последовательностей или приведенных здесь примерах, или последовательностью, комплементарной такому полинуклеотиду.

[0151] О-тРНК и/или О-РС согласно настоящему изобретению могут быть
30 получены из разнообразных организмов (например, эукариотических и/или неэукариотических организмов).

[0152] Полинуклеотиды также являются признаком настоящего изобретения. Полинуклеотид согласно настоящему изобретению (например, SEQ ID NO:5, 7 или 9) содержит искусственный (например, созданный человеком и не встречающийся в
35 природе) полинуклеотид, содержащий кодирующую полипептид последовательность нуклеотидов, которая приведена в списке последовательностей, прилагающемся к настоящей заявке, и/или комплементарную такой последовательности полинуклеотида. Полинуклеотид согласно настоящему изобретению может также
40 включать нуклеиновую кислоту, которая гибридизирует с полинуклеотидом, описанным выше, при очень жестких условиях, по существу, по всей длине нуклеиновой кислоты. Полинуклеотид согласно настоящему изобретению также содержит 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,
45 на 98% или более идентичен полинуклеотиду, содержащемуся в природной тРНК, или соответствующей кодирующей нуклеиновой кислоте (но полинуклеотид согласно настоящему изобретению отличается от встречающейся в природе тРНК или соответствующей кодирующей нуклеиновой кислоты), при этом тРНК распознает
50 селекторный кодон, например кодон, состоящий 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 указанных выше полинуклеотидов, и/или полинуклеотид,

содержащий консервативный вариант любого из указанных выше полинуклеотидов, также относятся к полинуклеотидам согласно настоящему изобретению.

[0153] Векторы, содержащие полинуклеотид согласно настоящему изобретению, также представляют собой признаки настоящего изобретения. Например, вектор
5 согласно настоящему изобретению может включать плазмиду, космиду, бактериофаг, вирус, вектор экспрессии, и/или тому подобное. Клетка, содержащая вектор согласно настоящему изобретению, также является признаком настоящего изобретения.

[0154] Способы получения компонентов пары О-тРНК/О-РС также являются
10 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
15 антикодоновой 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
20 потенциальных О-тРНК, который аминокатирует эндогенная аминокатил-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.

[0155] В определенных вариантах реализации, способы дополнительно включают
25 положительный отбор второй популяции клеток первого вида, где клетки включают член пула тРНК, ортогональный по отношению к клетке первого вида, когнатную аминокатил-тРНК синтетазу, и маркер положительного отбора. С помощью
30 положительного отбора клетки отбирают или скринируют с получением клеток, которые содержат член пула тРНК, аминокатируемых когнатной аминокатил-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
35 отрицательным отбором.

[0156] Также предусмотрены способы идентификации ортогональной аминокатил-тРНК синтетазы, которая нагружает О-тРНК не встречающейся в природе
40 аминокислотой. Например, способы включают отбор популяции клеток первого вида, каждая клетка которой содержит: 1) член множества аминокатил-тРНК синтаз (РСаз), (например, множество РСаз может включать мутантную РСазу, РСазу, полученную из вида, отличного от первого вида, или и мутантную АРСазу, и АРСазу, полученную из любого вида, отличного от первого вида); 2) ортогональную тРНК (О-тРНК) (например, из одного или более видов); и 3) полинуклеотид, который кодирует
45 маркер положительного отбора и содержит, по меньшей мере, один селекторный кодон.

[0157] Клетки (например, клетку-хозяина) подвергают отбору или скринингу на
50 клетки, которые демонстрируют повышение эффективности супрессии по сравнению с клетками, у которых или снижено количество данного члена множества АРС, или он отсутствует. Эти отобранные/скринированные клетки содержат активную АРС, которая аминокатирует О-тРНК. Ортогональная аминокатил-тРНК синтетаза, идентифицированная таким способом, также представляет собой признак настоящего

изобретения.

[0158] Способы получения белка в клетке (например, в эубактериальной клетке, такой как клетка *E.coli* или подобной клетке, или в дрожжевой клетке), содержащей не встречающуюся в природе аминокислоту в выбранном положении, также являются признаком настоящего изобретения. Например, способ включает выращивание на подходящей питательной среде клеток, которые содержат нуклеиновую кислоту, которая содержит, по меньшей мере, один селекторный кодон и кодирует белок, обеспечение не встречающейся в природе аминокислоты и встраивание не встречающейся в природе аминокислоты в указанное положение в белке в процессе трансляции нуклеиновой кислоты, содержащей, по меньшей мере, один селекторный кодон, 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 избирательно аминоацилирует О-тРНК не встречающейся в природе аминокислотой. Белок, полученный таким способом, также является признаком настоящего изобретения.

[0159] Изобретение также обеспечивает композиции, которые включают белки, содержащие фенилселеноцистеин. В определенных вариантах реализации, белок содержит последовательность аминокислот, которая является, по меньшей мере, на 75% идентичной последовательности известного белка, например, гормона роста человека, терапевтического белка, диагностического белка, промышленного фермента, или их части. Композиция может включать фармацевтически приемлемый носитель.

НУКЛЕИНОВАЯ КИСЛОТА, И ПОСЛЕДОВАТЕЛЬНОСТИ ПОЛИПЕПТИДА И ВАРИАНТЫ

[0160] 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, включающие указанные последовательности полипептида или полинуклеотида. Примеры указанных последовательностей, например, последовательности аминокислот О-тРНК и О-РС и последовательности нуклеотида, описаны в настоящей заявке (см. фиг.2). 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 настоящем изобретении, например, полинуклеотиды и полипептиды, кодирующие консервативные варианты О-РС, описанной в заявке.

[0161] Конструкция и анализ видов ортогональных синтаз (О-РС), которые обладают способностью аминоацилировать когнатные О-тРНК фенилселеноцистеином, описаны в Примере 1. В этом примере описана конструкция и анализ видов О-РС, которые способны встраивать не встречающуюся в природе аминокислоту фенилселеноцистеин.

[0162] Изобретение обеспечивает полипептиды (О-РС) и полинуклеотиды, например, О-тРНК, полинуклеотиды, которые кодируют О-РС или ее части, олигонуклеотиды, которые применяют для выделения клонов аминоацил-тРНК синтазы, 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, 7 или 9, и полинуклеотид, который является
5 комплементарным к такому полинуклеотиду или кодирует последовательность такого полинуклеотида. Полинуклеотиды согласно настоящему изобретению дополнительно включают любой полинуклеотид, который кодирует последовательность аминокислот O-PC, содержащую SEQ ID NO:4, 6 или 8. Аналогично, искусственная нуклеиновая
10 кислота, которая гибридизируется с полинуклеотидом, указанным выше, при очень жестких условиях по существу по всей длине нуклеиновой кислоты (и отлична от встречающегося в природе полинуклеотида) представляет собой полинуклеотид согласно настоящему изобретению. В одном варианте реализации, композиция
15 содержит полипептид согласно настоящему изобретению и носитель (например, буфер, воду, фармацевтически приемлемый носитель", и т.д.). Изобретение также обеспечивает антитело или антисыворотку, специфично связывающие полипептид согласно настоящему изобретению. Искусственный полинуклеотид - это полинуклеотид, который создан человеком и не встречается в природе.

[0163] Полинуклеотиды согласно настоящему изобретению также включают искусственный полинуклеотид, то есть, например, по меньшей мере, 75%, по меньшей мере, 80%, по меньшей мере, 90%, по меньшей мере, 95%, по меньшей мере, 98% или
20 более идентичный полинуклеотиду встречающейся в природе тРНК (но отличающийся от полинуклеотида встречающейся в природе тРНК). Полинуклеотид также содержит
25 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 (но не на 100% идентичен) полинуклеотиду встречающейся в природе тРНК.

[0164] В определенных вариантах реализации, вектор (например, плазида, космида, бактериофаг, вирус, и т.д.) содержит полинуклеотид согласно настоящему изобретению. В одном варианте реализации, вектор представляет собой вектор
30 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
35 содержит полинуклеотид согласно настоящему изобретению.

[0165] Специалисту будет ясно, что настоящее изобретение включает множество вариантов описанных последовательностей. Например, в изобретение включены
40 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
45 здесь, как определено, например, стандартными способами сравнения последовательностей, также включены в изобретение.

Консервативные варианты

[0166] Вследствие вырожденности генетического кода, "молчащие замены" (то есть, замены в последовательности нуклеиновой кислоты, которые не приводят к
50 изменению кодируемого полипептида) являются косвенным признаком 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 подобными описанной конструкции. Такие консервативные варианты каждой описанной последовательности представляют собой признак настоящего изобретения.

[0167] Термин "консервативные варианты" определенной последовательности нуклеиновой кислоты относится к таким нуклеиновым кислотам, которые кодируют идентичные или по существу идентичные последовательности аминокислот, или, в случае, если нуклеиновая кислота не кодирует последовательность аминокислот, к по существу идентичным последовательностям. Специалисту будет ясно, что отдельные замены, делеции или вставки, которые приводят к замене, добавлению или удалению одной природной аминокислоты или небольшого процента аминокислот (обычно меньше 5%, чаще меньше 4%, 2% или 1%) в кодируемой последовательности, представляют собой "консервативно модифицированные варианты", в которых указанные изменения приводят к 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 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 последовательности, является консервативным изменением основной нуклеиновой кислоты.

[0168] Таблицы консервативных замен, обеспечивающие функционально подобные аминокислоты, хорошо известны в данной области, где одним остатком аминокислоты заменяют другой остаток аминокислоты, имеющий подобные химические свойства (например, ароматические боковые цепи или положительно заряженные боковые цепи), что не вызывает существенных изменений функциональных свойств молекулы полипептида. Следующие приведенные в качестве примера группы содержат природные аминокислоты с подобными (близкими) химическими свойствами, в которых замены в пределах группы представляют собой "консервативные замены".

Консервативные замены аминокислот

Неполярные и/или алифатические боковые цепи	Полярные, ненаполненные боковые цепи	Ароматические боковые цепи	Положительно заряженные боковые цепи	Отрицательно заряженные боковые цепи
Глицин	Серин			
Аланин	Треонин	Фенилаланин	Лизин	Аспарагиновая кислота
Валин	Цистеин	Тирозин	Аргинин	
Лейцин	Метионин	Триптофан	Гистидин	Глютаминовая кислота
Изолейцин	Аспарагин			
Пролин	Глутамин			

Гибридизация нуклеиновых кислот

[0169] 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

различия нуклеиновых кислот согласно настоящему изобретению. Кроме того, целевые нуклеиновые кислоты, которые гибридизируют с нуклеиновой кислотой, представленной в SEQ ID NO:5, 7, 9 или 11, при 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.

[0170] 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.

[0171] Говорят, что нуклеиновые кислоты "гибридизируются", если они соединяются, 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.

[0172] 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 0.2х 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 промывке высокой жесткости предшествует промывание низкой жесткости для исключения фонового сигнала зонда. Пример промывания низкой жесткости: 2х SSC при 40°C в течение 15 минут. Как правило, отношение сигнала к шуму составляет 5х (или выше) наблюдаемого для неродственного зонда, который при определенном способе анализа гибридизации указывает на обнаружение специфичной гибридизации.

[0173] "Жесткие условия промывки при 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.

5 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
10 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 связывается с идеально совпадающей комплементарной мишенью с соотношением сигнала и шума, составляющим, по меньшей мере, 5x уровня, наблюдаемого в случае
15 гибридизации зонда и несовпадающей мишени.

[0174] "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
20 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.

[0175] "Ультравысокая жесткость" условий гибридизации и промывки - это такая, при которой жесткость условий гибридизации и промывки увеличивают до тех пор, пока соотношение сигнала шуму для связывания зонда с идеально совпадающей
25 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
30 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.

[0176] 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
35 гибридизации и промывки увеличивают, пока соотношение сигнала к шуму для связывания зонда с идеально совпадающей комплементарной нуклеиновой кислотой-мишенью не составит, по меньшей мере, 10x, 20X, 50X, 100X, 500X или более от
40 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,
45 по меньшей мере, 1/2 от уровня для идеально совпадающей комплементарной нуклеиновой кислоты-мишени, связывается с зондом при условиях сверхвысокой (ультра-ультра высокой) жесткости.

[0177] Нуклеиновые кислоты, которые не гибридизируются друг с другом в жестких
50 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

[0178] В некоторых аспектах, настоящее изобретение обеспечивает нуклеиновую кислоту, которая содержит уникальную подпоследовательность 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, для идентификации нуклеиновых кислот согласно настоящему изобретению или родственных нуклеиновых кислот.

[0179] 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.

[0180] Изобретение также включает нуклеиновые кислоты-мишени, которые гибридизируются в жестких условиях с уникальным кодирующим олигонуклеотидом, который кодирует уникальную подпоследовательность 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 синтазы согласно настоящему изобретению, например, путем мутации).

Уникальные последовательности определяют, как отмечено выше.

Сравнение последовательностей, идентичность и гомология

[0181] Термины "идентичная" или "процент идентичности" применительно к двум или более нуклеиновым кислотам или последовательностям полипептида, относятся 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.

[0182] 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.

[0183] Белки и/или последовательности белков являются "гомологичными", если они получены природным или искусственным путем из общего материнского белка или последовательности белка. Аналогично, нуклеиновые кислоты и/или последовательности нуклеиновой кислоты являются гомологичными, если они
5 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
10 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,
15 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, но обычно для установления
20 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, 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.

[0184] В сравнении последовательностей и определении гомологии обычно одна последовательность выступает в качестве контрольной последовательности, с которой сравнивают тестовые последовательности. Используя алгоритм сравнения последовательностей, тестовую и контрольную последовательности вводят в
30 компьютер, при необходимости определяют координаты субпоследовательности и определяют параметры алгоритма программы последовательности. После этого алгоритм сравнения последовательности на основании заданных параметров программы вычисляет процент идентичности последовательностей для тестовой последовательности (ей) относительно контрольной последовательности.

[0185] 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
40 подобия 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), или визуальным осмотром (смотри в общем Current Protocols In Molecular Biology, Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and
45 John Wiley & Sons, Inc., (supplemented through 2004).

[0186] Один пример алгоритма, подходящего для определения процента идентичности последовательностей и подобия последовательностей - это алгоритм BLAST, который описан в Altschul et al., J. Mol. Biol., 215:403-410 (1990). Программное обеспечение для реализации BLAST-анализа находится в свободном
50 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 (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 порог количества баллов для соседнего слова (Altschul et al., (1990) *J. Mol. Biol.*, 215:403-410). Указанные исходные совпадения соседних слов выполняют роль затравки для начала поисков для выявления более длинных HSP, содержащих их. Совпадения слов продлевают в обоих направлениях вдоль каждой последовательности до тех пор, пока общее количество баллов выравнивания продолжает увеличиваться. Суммарные баллы рассчитывают, используя, 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 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).

[0187] В дополнение к вычислению процента идентичности последовательностей, алгоритм BLAST также выполняет статистический анализ подобия между двумя последовательностями (см., например, Karlin and Altschul, *Proc Nat'l. 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

[0188] 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 темы, связанные с, например, получением генов, которые включают селекционные кодоны для синтеза содержащих не встречающиеся в природе аминокислоты белков, ортогональных тРНК, ортогональных синтетаз, и их пар.

[0189] В настоящем изобретении применяют различные типы мутагенеза, например

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

[0190] Клетки-хозяева подвергают генетической модификации (например, трансформируют, трансдуцируют или трансфецируют) полинуклеотидами согласно 25 настоящему изобретению или конструктами, которые включают полинуклеотид согласно настоящему изобретению, например, вектор согласно настоящему изобретению, который может быть, например, вектором клонирования или вектором экспрессии. Например, области, кодирующие ортогональную тРНК, ортогональную тРНК синтетазу, и белок, который необходимо изменить, функционально связаны с 30 элементами контроля экспрессии гена, которые являются функциональными в желаемой 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 35 мере одну независимую последовательность терминатора, последовательности, обеспечивающие репликацию кассеты в эукариотах, или прокариотах, или и в эукариотах, и в прокариотах (например, шаттл-векторы) и маркеры отбора и для прокариотических, и для эукариотических систем. Векторы подходят для репликации 40 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 45 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, 50 вируса, «голового» 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, высокоскоростное 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.

5 [0191] Высокоэффективная и гибкая система, содержащая единственную плазмиду, была разработана для сайт-специфичного встраивания не встречающихся в природе аминокислот в белки в ответ на терминирующий амбер-кодон (UAG) в *E.coli*. В новой системе пара супрессорной тРНК^{tyr} (CUA) *M. jannaschii* и тирозил-тРНК синтетазы закодирована в единственной плазмиде, которая совместима с большинством
10 векторов экспрессии *E.coli*. Для оптимизации вторичной структуры и процессинга тРНК был создан моноцистронный оперон тРНК под контролем промотора *proK* и терминатора. Введение мутантной формы *glnS* промотора в ген синтетазы обеспечило значительное увеличение как эффективности, так и точности супрессии. Увеличение
15 эффективности супрессии было также достигнуто с помощью увеличения числа копий гена тРНК и специфической мутации (D286R) (Kobayashi et al., "Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion," *Nat. Struct. Biol.*, 10(6):425-432 [2003]). Универсальность оптимизированной системы была также продемонстрирована очень эффективным и точным встраиванием нескольких
20 различных не встречающихся в природе аминокислот, уникальные полезные свойства которых для изучения функции и структуры белка были предварительно доказаны.

[0192] Каталог бактерий и бактериофагов, пригодных для клонирования, представлен, например, ATCC, e.g., *The ATCC Catalogue of Bacteria and Bacteriophage* (1996) Gherna et al. (eds), опубликованный Американской коллекцией типовых культур.
25 Другие основные процедуры для секвенирования, клонирования 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 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
30 существу, любую нуклеиновую кислоту (и фактически любую меченную нуклеиновую кислоту, стандартную или нестандартную) можно изготовить на заказ или
35 стандартным заказать из любого из разнообразных коммерческих источников, таких как Midland Certified Reagent Company (Midland, TX), The Great American Gene Company (Ramona, CA), ExpressGen Inc. (Chicago, IL), Operon Technologies Inc. (Alameda, CA) и многие другие.

40 [0193] Модифицированные клетки-хозяева можно культивировать на 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, например, по
45 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 Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds) (1995) *Plant Cell, Tissue and Organ Culture; Fundamental
50 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.

БЕЛКИ И ПОЛИПЕПТИДЫ, ПРЕДСТАВЛЯЮЩИЕ ИНТЕРЕС

[0194] 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 кодирующий 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.

[0195] В определенных вариантах реализации, О-РС аминоацилирует предпочтительно О-тРНК, 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.

[0196] Изобретение также обеспечивает композиции, которые содержат белки, которые содержат не встречающуюся в природе аминокислоту. В определенных вариантах реализации, белок содержит последовательность аминокислот, которая является, по меньшей мере, на 75% идентичной последовательности терапевтического белка, диагностического белка, промышленного фермента, или их фрагмента.

[0197] 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 не встречающейся в природе аминокислоты в белок. Этот процесс описан в публикациях международных заявок на патент номер 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," 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 не встречающейся в природе аминокислоты, такой как фенилселеноцистеин, 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 (ax).

[0198] Клетка согласно 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 микрограмм, по меньшей мере 1 миллиграмм, по меньшей мере 10 миллиграмм или более белка, который содержит не встречающуюся в природе

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 белка на литр, или по меньшей мере 10 миллиграмм белка на литр или большей, в, например, клеточном лизате, буфере, фармацевтическом буфере, или другой жидкой суспензии (например, в объеме, равном, например, любому значению от приблизительно 1 мл до приблизительно 100 л). Получение больших количеств (например, больше количеств, которые обычно возможно получить другими способами, например, *in vitro* трансляцией) белка в клетке, содержащего, по меньшей мере, одну не встречающуюся в природе аминокислоту, является признаком настоящего изобретения.

[0199] Встраивание не встречающейся в природе аминокислоты можно осуществлять для, например, изменения для приспособления структуры и/или функции белка, например, изменение размера, кислотности, нуклеофильности, водородных связей, гидрофобности, доступности участков являющихся целями протеаз, нацеливания на какой-либо фрагмент (например, для белкового микрочипа), встраивания меток или реакционно-способных групп, и т.д. Белки, которые содержат не встречающуюся в природе аминокислоту, могут иметь усовершенствованные 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 не встречающуюся 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.

[0200] В некоторых аспектах настоящего изобретения, композиция содержит 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 не встречающихся в природе аминокислот. Не встречающиеся в природе аминокислоты могут 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 не встречающихся в природе аминокислоты. Для данного белка, содержащего более одной не

встречающейся в природе аминокислоты, не встречающиеся в природе аминокислоты могут быть одинаковыми или различными (например, белок может содержать два или более различных типов не встречающихся в природе аминокислот, или может включать две одинаковые не встречающиеся в природе аминокислоты). Для данного белка с более чем двумя не встречающимися в природе аминокислотами указанные не встречающиеся в природе аминокислоты могут быть одинаковыми, различными или представлять собой комбинацию множества не встречающейся в природе аминокислоты одного вида (т.е. несколько одинаковых аминокислот) с по меньшей мере с одной другой не встречающейся в природе аминокислотой.

[0201] По существу любой белок (или часть белка), который содержит не встречающуюся в природе аминокислоту (и любую соответствующую, кодирующую нуклеиновую кислоту, например, которая содержит один или более селекторных кодонов), можно получить с использованием композиций и способов, изложенных в настоящей заявке. Не было предпринято никаких попыток идентифицировать сотни тысяч известных белков, любой из которых может быть изменен таким образом, чтобы он включал одну или более не встречающихся в природе аминокислот, например, с помощью применения любых доступных методов мутагенеза для включения одного или более подходящих селекторных кодонов в соответствующую систему трансляции. Общие базы последовательностей известных белков включают: GenBank EMBL, DDBJ и NCBI. Другие базы можно легко найти путем поиска в сети Интернет.

[0202] Как правило, белки, например, по меньшей мере на 60%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 90%, по меньшей мере на 95%, или по меньшей мере на 99% или более идентичны любому доступному белку (например, терапевтическому белку, диагностическому белку, промышленному ферменту, или фрагменту перечисленных белков и т.п.), и они включают одну или более не встречающихся в природе аминокислот. Примеры терапевтических, диагностических и других белков, которые можно модифицировать таким образом, чтобы они включали одну или более не встречающихся в природе аминокислот, можно найти в публикациях международных заявок на патент: 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 публикациях. Примеры терапевтических, диагностических и других белков, которые могут быть модифицированы таким образом, чтобы они содержали одну или более не встречающихся в природе аминокислот, включают, без ограничения, следующие белки, например, гирудин, гормон роста человека, RAS, альфа-1 антитрипсин, ангиостатин, антигемолитический фактор, антитела (детали, касающиеся антител, приведены ниже), аполипопротеин, апопротеин, предсердный натрийуретический фактор, предсердный натрийуретический полипептид, предсердные пептиды, 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, CC хемокины (например, моноцитарный хемотаксический белок -1, моноцитарный хемотаксический белок -2, моноцитарный хемотаксический белок -3, воспалительный белок моноцитов-1 альфа, воспалительный белок моноцитов-1 бета, RANTES, I309, R83915, R91733, HCC1, T58847, D31065, T64262), лиганд CD40, C-kit лиганд, коллаген, колонстимулирующий фактор (CSF), фактор комплемента 5a, ингибитор комплемента, рецептор комплемента 1, цитокины, (например, пептид 78, активирующий нейтрофилы эпителия, GROa/MGSA, GROb, GROg, MIP-1a, MIP-1d,

МСП-1), эпидермальный фактор роста (EGF), эритропоэтин ("ЕРО"), расслаивающие токсины А и В, фактор IX, фактор VII, фактор VIII, фактор X, фактор роста фибробластов (FGF), фибриноген, фибронектин, G-CSF, GM-CSF, глюкоцереброзидазу, гонадотропин, факторы роста, хэджджог (hedgehog)-белки (например, Sonic, Indian, Desert), гемоглобин, фактор роста гепатоцитов (HGF), гирудин, сывороточный альбумин человека, инсулин, инсулиноподобный фактор роста (IGF), интерфероны (например, IFN- α , IFN- β , IFN- γ), 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 т.д.), фактор роста кератиноцитов (KGF), лактоферрин, фактор ингибирования лейкоза, люцифераза, неуртурин, фактор ингибирования нейтрофилов (NIF), онкостатин М., остеогенный белок, паратгормон, PD-ECSF (колониестимулирующий фактор тромбоцитов-эритроцитов), PDGF (тромбоцитарный фактор ростапт), пептидные гормоны (например, соматотропин человека), плейотропин, белок А, белок G, дирогенные экзотоксины А, В и С, 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, то есть, стафилококковые энтеротоксины (SEA, SEB, SEC1, SEC2, SEC3, SED, SEE), супероксид дисмутаза (СОД), токсин синдрома токсического шока (TSST-1), тимозин альфа 1, тканевой активатор плазминогена, бета фактора некроза опухоли (бета TFN), рецептор фактора некроза опухоли (TNFR), фактор некроза опухоли альфа (TNF альфа), фактор роста эндотелиаля сосудов (VEGEF), урокиназа и многие другие.

[0203] Один класс белков, которые можно получить при помощи описанных в настоящей заявке композиций и способов *in vivo* встраивания не встречающихся в природе аминокислот, включает модуляторы транскрипции или их фрагменты. Модуляторы транскрипции, приводимые в качестве примера, включают гены и белки модуляторов транскрипции, которые модулируют рост, дифференцировку, регуляцию клетки и т.п. Модуляторы транскрипции, найденные у прокариот, вирусов и эукариот, включая грибы, растения, дрожжи, насекомых, и животных, включая млекопитающих, обеспечивают широкий диапазон мишеней терапии. Очевидно, что активаторы экспрессии и транскрипции регулируют транскрипцию посредством многих механизмов, например, путем связывания с рецепторами, инициации каскадов передачи сигнала, регуляции экспрессия факторов транскрипции, связывающихся с промоторами и энхансерами, связывания с белками, которые связываются с промоторами и энхансерами, раскручивания ДНК, участия в сплайсинге пре-мРНК, полиаденилирования РНК, и участия в разрушении РНК.

[0204] Один класс белков согласно настоящему изобретению (например, белки с одной или более не встречающимися в природе аминокислотами) включает биологически активные белки, такие как цитокины, молекулы воспаления, факторы роста, их рецепторы, и продукты онкогенов, например, интерлейкины (например, IL-1, IL-2, IL-8, и т.д.), интерфероны, FGF, IGF-I, IGF-, FGF, PDGF, TNF, TGF-a, TGF-b, EGF, KGF, SCF/c-Kit, CD40L/CD40, VLA-4/VCAM-1, ICAM-1/LFA-1, и гиалулин/CD44; молекулы трансдукции сигнала и соответствующие продукты онкогена, например, Mos, Ras, Raf, и MET; и активаторы и супрессоры трансляции, например, p53, TAT, Fos, Muc, Jun, Myb, Rel, и рецепторы стероидных гормонов, такие как эстроген, прогестерон, тестостерон, альдостерон, лиганд рецептора LDL и кортикостерон.

[0205] Ферменты (например, промышленные ферменты) или фрагменты ферментов, содержащие по меньшей мере одну не встречающуюся в природе аминокислоту, также

предусмотрены настоящим изобретением. Примеры ферментов включают, например, амидазы, рацемазы аминокислот, ацилазы, дегалогеназы, диоксигеназы, диарилпропан пероксидазы, эпимеразы, гидролазы эпоксида, эстеразы, изомеразы, киназы, глюкоза изомеразы, гликозидазы, гликозил трансферазы, галопероксидазы, монооксигеназы (например, р450s), липазы, лигниновые пероксидазы, нитрил гидратазы, нитрилазы, протеазы, фосфатазы, субтилизины, трансминазы, и нуклеазы, но не ограничены перечисленными ферментами.

[0206] Многие из этих белков коммерчески доступны (См., например, Sigma BioSciences), и соответствующие последовательности белков и генов, и обычно, множество вариантов белков и генов являются известными (см., например, Genbank). Любой из указанных белков можно модифицировать путем вставки одной или более не встречающихся в природе аминокислот согласно настоящему 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 LD₅₀ 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.

[0207] Другие различные белки также можно модифицировать таким образом, чтобы они содержали одну или более не встречающихся в природе аминокислот, с применением композиций и способов согласно настоящему изобретению. Например, изобретение может включать замену одной или более природной аминокислоты в одном более белках вакцин на не встречающуюся в природе аминокислоту, например, в белках из инфекционных грибов, например, *Aspergillus*, *Candida species*; бактерий, особенно *E.coli*, которая служит моделью патогенных бактерий, а также важных с медицинской точки зрения бактерий, таких как *Staphylococci* (например, *aureus*), или *Streptococci* (например, *pneumoniae*); протозоа, таких как спорозоа (например, *Plasmodia*), ризопода (например, *Entamoeba*) и флагеллят (*Trypanosoma*, *Leishmania*, *Trichomonas*, *Giardia*, и т.д.); вирусов, таких как вирусы (+)РНК (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.

[0208] Белки, имеющие отношение к сельскому хозяйству, такие как белки устойчивости к насекомым (например, белки Cry), 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.

5 [0209] В определенных 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
10 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.

15 [0210] Гены, кодирующие белки или полипептиды, представляющие интерес, можно подвергнуть мутагенезу, с применением методов, известных любому специалисту в данной области техники и описанных в разделе "Мутагенез и другие методы молекулярной биологии" настоящей заявки, с включением, например, одного или
20 более селекторных кодонов для встраивания не встречающейся в природе аминокислоты. Например, нуклеиновую кислоту, кодирующую белок, представляющий интерес, подвергают мутагенезу с включением одного или более селекторных кодонов для обеспечения вставки одной или более не встречающихся в природе аминокислот. Настоящее изобретение включает любой такой вариант,
25 например, мутант, вариант любого белка, например, содержащий по меньшей мере одну не встречающуюся в природе аминокислоту. Аналогично, настоящее изобретение включает соответствующие нуклеиновые кислоты, то есть, любую нуклеиновую кислоту с одним или более селекторными кодонами, которые кодируют
30 одну или более не встречающихся в природе аминокислот.

[0211] Чтобы получить белок, который содержит не встречающуюся в природе аминокислоту, можно применять клетки-хозяева и организмы, подготовленные к *in vivo* включению не встречающейся в природе аминокислоты с помощью ортогональных пар тРНК/РС. Клетки-хозяева подвергают генетической
35 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
40 отдельном векторе, либо два компонента могут быть в одном векторе, а третий компонент - во втором векторе. Вектор может иметь, например, форму плазмиды, бактерии, вируса, «голового» полинуклеотида или конъюгированного полинуклеотида.

Определение полипептидов по иммунореактивности

45 [0212] Поскольку полипептиды согласно настоящему изобретению обеспечивают разнообразные новые последовательности полипептидов (например, полипептиды, содержащие не встречающиеся в природе аминокислоты в случае белков, синтезируемых в системах трансляции, описанных в настоящей заявке, или, например, в случае новых синтетаз, новых последовательностей стандартных аминокислот),
50 полипептиды также обеспечивают новые характерные структуры, которые можно распознавать, например, в иммунологических пробах. Получение антисыворотки, которая специфично связывает полипептиды согласно настоящему изобретению, а также 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 дисплей, 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.

[0213] Чтобы получать антисыворотку для применения в иммунологическом анализе, получают и очищают один или более иммуногенных полипептидов, описанных в настоящей заявке. Например, рекомбинантный белок можно получить 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

[0214] 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 не встречающейся в природе 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 не встречающейся в природе 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 (ax). Белки с не встречающейся в природе аминокислотой можно применять в разнообразных целях. Например, не встречающийся в природе 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 не встречающейся в природе 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.

5 НАБОРЫ

[0215] 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
10 содержит по меньшей мере один контейнер, содержащий последовательность полинуклеотида, кодирующую O-тРНК, и/или O-тРНК, и/или последовательность полинуклеотида, кодирующую O-РС, и/или O-РС. В одном варианте реализации, набор дополнительно содержит не встречающуюся в природе аминокислоту
15 фенилселеноцистеин. В другом варианте, набор дополнительно содержит инструкции по получению белка и/или клетки-хозяина.

ПРИМЕРЫ

[0216] Приведенные ниже примеры приведены с целью иллюстрации, но не ограничения настоящего изобретения. Для специалиста будут очевидны
20 разнообразные второстепенные параметры, можно изменять, не выходя за пределы объема настоящего изобретения. Очевидно, что примеры и варианты реализации описаны в настоящей заявке исключительно в целях иллюстрации, и что различные их модификации или изменения, соответственно, очевидные для специалистов в данной области техники, находятся в пределах идеи и объема данной заявки и объема
25 формулы согласно настоящему изобретению.

ПРИМЕР

Генетический отбор мутантных синтетаз, специфичных к фенилселеноцистеину

РЕАКЦИОННАЯ ХИМИЯ ФЕНИЛСЕЛЕНОЦИСТЕИНА

[0217] Селен - это необходимый элемент как в органическом синтезе, так и в
30 биологии. Было показано, что фенилселено-группа может быть окислена в мягких условиях, и что самопроизвольное отщепление сульфонильной группы от селеноксида приводит к формированию олефина. Эта стратегия широко применяют в общем синтезе.

[0218] Встраивание не встречающейся в природе аминокислоты, содержащей селен,
35 в белки обеспечило бы привлекательную мишень для высокоселективной посттрансляционной модификации белков, например, получения селективно липидизированных белков. Одна из не встречающихся в природе селеноаминокислот - это фенилселеноцистеин (регистрационный номер CAS 71128-82-0), называемый также
40 фенилселенидом или Se-фенил-L-селеноцистеином. Структура этого соединения показана на фиг.1, структура 1. Эту не встречающуюся в природе аминокислоту можно получить, например в Sigma®-Aldrich®, Inc. номер по каталогу 50827.

[0219] Продуктами окислительного расщепления остатка аминокислоты фенилселеноцистеина являются α -, β -ненасыщенная аминокислота дегидроаланин (см.
45 фиг.1, структура 2). Дегидроаланин обнаружен во множестве природных пептидов (Chatterjee et al. (2005), Chemical Reviews 105(2):633-683). Дегидропептиды, содержащие незащищенные остатки цистеина, подвержены внутримолекулярному стереоселективному сопряженному присоединению, в результате которого образуется
50 лантионин (Chatterjee et al. (2005), Chemical Reviews 105(2):633-683). Лантионины - это структуры, которые найдены в лантибиотиках, представляющих собой класс антибиотиков, подвергнувшихся посттрансляционной модификации (Chatterjee et al. (2005), Chemical Reviews 105(2):633-683).

[0220] В результате реакции присоединения Майкла с участием не встречающейся в природе аминокислоты дегидроаланина (см. фиг.1, структура 2) могут быть получены белки с посттрансляционными модификациями, например, реакция с тио-липидом обеспечивает получение липидизированного белка. Более конкретно, реакция с тиопальмитиновой кислотой дает пальмитоилцистеин (см. фиг.1, структура 3), реакция с фарнезилмеркаптаном дает фарнезилцистеин (см. фиг.1, структура 4) и реакция с малонатом дает γ -карбоксиглутаминовую кислоту (см. фиг.1 структура 6). Кроме того, реакция с 1-гексадеканэтиолом дает S-гексодецилцистеин (см. фиг.1, структура 5). Хотя S-гексодецилцистеин не является нативной посттрансляционной модификацией, наличие этого остатка в белке может придать ему желаемые свойства; например, может обеспечить связывание сывороточного альбумина человека, и более высокую стабильность белка *in vivo*.

[0221] Липидизация полипептидов путем направленной модификации остатков фенилселеноцистеина может придать полипептиду желательные свойства, такие как локализация в мембране, улучшенная стабильность *in vivo* (то есть, улучшенный период полураспада), и растворимость в липидах. Кроме того, липидизированные формы некоторых белков являются биологически активными.

ПОЛУЧЕНИЕ БЕЛКОВ, ВКЛЮЧАЮЩИХ ФЕНИЛСЕЛЕНОЦИСТЕИН

[0222] Способы, которые позволяют систематическое добавление не встречающихся в природе аминокислот в генетический код клеток *E.coli*, дрожжей и млекопитающих, были предварительно описаны. Такие способы основаны на получении пары нонсенс супрессорной тРНК и аминоксил-тРНК синтетазы (РС), которая имеет свойство ортогональности, определенное как способность селективно встраивать данную аминокислоту в ответ на уникальный кодон без перекрестной реакции с эндогенными тРНК, аминоксил-тРНК синтетазами или аминокислотами организма.

[0223] Настоящее изобретение обеспечивает композиции и способы получения полипептидов *in vivo*, содержащих не встречающуюся в природе аминокислоту фенилселеноцистеин путем создания ортогональных реагентов, которые обеспечивают генетически запрограммированное встраивание этой не встречающейся в природе аминокислоты непосредственно в растущую цепь полипептида в ходе трансляции. Кроме того, применение таких ортогональных систем обеспечивает получение в больших количествах полипептидов, содержащих не встречающуюся в природе аминокислоту (то есть, намного большего количества полипептидов, чем было бы возможно получить с использованием других методов синтеза). Полипептиды, содержащие фенилселеноцистеин, можно применять в реакциях направленной модификации, например, при синтезе искусственно липидизированных полипептидов, согласно идеям, изложенным в настоящем описании, и других реакциях. Такие модифицированные белки находят разнообразные применения, включая, без ограничения, улучшение терапевтических средств.

[0224] Для получения ортогональной пары тРНК/аминоацил-РС, которая единственным образом встраивает фенилселеноцистеин (фиг.1., структура 1), применяли библиотеку мутантов по активному центру тирозил-тРНК синтетазы (MjTyrRS) *Methanococcus jannaschii*, которая специфично присоединяет сконструированный супрессор нонсенс-мутации *M. jannaschii* (MjmRNA^{Tyr}CUA), не распознаваемый синтетазой *E.coli*. Библиотеку синтетаз рВК-lib5, описанную в Wang et al. (2006), Annual Review of Biophysics and Biomolecular Structure 35:225-249, применяли для скрининга и подвергли ряду положительных и отрицательных отборов. Выживание при положительном отборе зависело от подавления амбер-мутации в гене

хлорамфеникол ацетилтрансферазы (САТ) в присутствии фенилселеноцистеина; выживание при отрицательном отборе зависело от неадекватного подавления амбер-мутации в гене, кодирующем токсичный белок барназу, в отсутствие фенилселеноцистеина. Клоны выживают после циклов положительного и отрицательного отбора, только если они обязательно встраивают фенилселеноцистеин в ответ на янтарный кодон.

[0225] После описанных отборов идентифицировали клоны-кандидаты, что позволило клеткам, несущим ген САТ, содержащий мутацию селекторного амбер-кодона на доступном сайте, оставаться живыми в присутствии хлорамфеникола только в присутствии фенилселеноцистеина. В отсутствие фенилселеноцистеина те же самые клетки не росли в присутствии хлорамфеникола, что согласуется с эффективным поглощением фенилселеноцистеина при низком фоновом выживании или отсутствии фонового выживания в результате встраивания эндогенных аминокислот.

[0226] Секвенирование клонов мутантной синтетазы кандидата выявило три различных изолята синтетазы, каждый из которых способен к функционированию в ортогональной системе трансляции.

Клон1 (PhSeRS-SD): Y32W, L65E, H70G, D158Q, L162S

Клон2 (PhSeRS-K4): Y32W, L65H, H70G, F108N, Q109S, D158S, L162E

5 Клон3 (PhSeRS-K5): Y32W, L65H, A67G, H70G, F108K, Q109S, D158E, L162E

[0227] Указанные клоны, дополнительно описаны в приведенной ниже Таблице.

10 **Последовательности аминокислот тирозил-тРНК синтетазы *Methanococcus jannaschii***

Дикий тип и мутанты

	32	65	67	70	108	109	158	162	SEQ ID NO:	
15										
20	дикий тип	Tyr	Leu	Ala	His	Phe	Gln	Asp	Leu	2
25	Клон Мутанта 1(SD)	Trp	Glu	Ala	Gly		Gln	Gln	Ser	4
30	клон мутанта 2(K4)	Trp	His	Ala	Gly			Ser	Glu	6
35	клон мутанта 3(K5)	Trp	His	Gly	Gly	Lys	Ser	Glu	Glu	8

40 [0228] Из трех изолированных мутантных синтетаз, клон 2 синтетазы (PhSeRS-K4) имел самые высокие активность и специфичность. Полные последовательности нуклеотидов и аминокислот каждого из этих клонов и соответствующих видов дикого типа приведены на фиг.2.

ПРОВЕРКА ОРТОГОНАЛЬНОЙ СИНТЕАЗЫ

45 [0229] С применением мутантных синтетаз, описанных в настоящей заявке, экспрессировали модельный белок миоглобин, содержащий селекторный амбер-кодон в положении 4 (TAG4), в количестве 1.5-2 мг/л. Фоновая экспрессия в отсутствие фенилселеноцистеина отсутствовала. Затем этот модельный белок, содержащий фенилселеноцистеин, подвергли посттрансляционной модификации с получением γ -
50 карбоксиглютаминной кислоты (фиг.1 структура б), что было подтверждено методом масс-спектрофотометрии высокого разрешения.

[0230] Два дополнительных модельных белка, содержащих фенилселеноцистеин, также экспрессировали в количестве 3-5 мг/л, с использованием мутантных синтетаз и

ортогональной системы получения *in vivo*, описанной в данной заявке. Был получен гормон роста человека (hGH), содержащий фенилселеноцистеин. После получения формы с фенилселеноцистеином белок подвергли модификации с получением hGH, содержащего S-гексадецилцистеин. Связывающие свойства сывороточного альбумина человека и период полураспада *in vivo* модифицированного hGH исследуют в настоящее время на мышах.

[0231] Ортогональные реагенты, описанные в настоящей заявке, также применяли для получения зеленого флуоресцентного белка (GFP), содержащего фенилселеноцистеин. Эту форму белка также впоследствии подвергли модификации с получением, в качестве альтернативы, формы GFP, содержащей S-гексадецилцистеин, фарнезилцистеин и пальмитоилцистеин. В настоящее время идет тестирование свойств нацеливания в мембраны этих форм липидизированных GFP.

[0232] Таким образом, настоящее изобретение обеспечивает композиции и способы генетически запрограммированной продукции белков, содержащих не встречающуюся в природе аминокислоту фенилселеноцистеин (см. фиг.1, структура 1). Последующее окислительное отщепление путем реакции с перекисью водорода приводит к образованию дегидроаланина (см. фиг.1, структура 2) в желательном положении. Этот фрагмент дегидроаланина затем может быть мишенью различных реакций вторичной модификации.

[0233] Хотя изложенное выше изобретение в целях ясности описано с упоминанием некоторых деталей, для специалиста в данной области техники после прочтения этого описания станет очевидно, что в форму и детали можно внести различные изменения, не выходя за пределы истинного объема согласно настоящему изобретению. Все публикации, патенты, заявки на патент и/или другие документы, процитированные в этой заявке, включены в нее посредством ссылки в полном объеме для любых целей, так же как если бы для каждой отдельной публикации, патента, заявки на патент и/или другого документа было отдельно указано, что они включены в настоящую заявку для любых целей.

Формула изобретения

1. Композиция для трансляции белка, содержащего фенилселеноцистеин, включающая:

(а) фенилселеноцистеин;

(б) ортогональную аминоксил-тРНК синтетазу (О-РС), где последовательность указанной О-РС является по меньшей мере на 90% идентичной SEQ ID NO:4, 6 или 8 и содержит Tgr в положении, соответствующем положению 32 последовательности SEQ ID NO:4, 6 или 8; Glu или His в положении, соответствующем положению 65 последовательности SEQ ID NO:4, 6 или 8; Gly в положении, соответствующем положению 70 последовательности SEQ ID NO:4, 6 или 8; Gln, Ser или Glu в положении, соответствующем положению 158 последовательности SEQ ID NO:4, 6 или 8; Ser или Glu в положении, соответствующем положению 162 последовательности SEQ ID NO:4, 6 или 8; и

(с) ортогональную тРНК (О-тРНК), где указанная О-тРНК имеет последовательность полинуклеотида, которая по меньшей мере на 90% идентична SEQ ID NO:1;

причем указанная О-РС избирательно аминоксилуирует указанную О-тРНК указанным фенилселеноцистеином.

2. Композиция по п.11, в которой указанная О-РС избирательно аминоксилуирует

указанную О-тРНК указанным фенилселеноцистеином с эффективностью, которая составляет по меньшей мере 50% эффективности, которую наблюдают у системы трансляции, включающей указанную О-тРНК, указанный фенилселеноцистеин и аминоксил-тРНК синтетазу, содержащую последовательность аминокислот SEQ ID NO:4, 6 или 8.

3. Композиция по п.1, в которой указанная О-РС содержит последовательность аминокислот, выбранную из группы, состоящей из последовательностей аминокислот, обозначенных SEQ ID NO:4, 6, 8, и их консервативных вариантов.

4. Композиция по п.1, в которой указанная О-тРНК представляет собой тРНК, являющуюся супрессором амбер-кодона.

5. Композиция по п.1, в которой указанная О-тРНК содержит последовательность полинуклеотида, обозначенную SEQ ID NO:1, или кодируется указанной последовательностью.

6. Композиция по п.1, отличающаяся тем, что содержится в клетке-хозяине.

7. Композиция по п.6, отличающаяся тем, что указанная клетка-хозяин представляет собой клетку зубактерии.

8. Композиция по п.6, отличающаяся тем, что указанная клетка-хозяин представляет собой клетку E.coli.

9. Композиция по п.6, отличающаяся тем, что указанная клетка-хозяин содержит полинуклеотид, кодирующий указанную О-РС.

10. Композиция по п.9, отличающаяся тем, что указанный полинуклеотид содержит последовательность нуклеотидов, обозначенную SEQ ID NO:5, 7 или 9.

11. Композиция по п.6, отличающаяся тем, что указанная клетка-хозяин содержит полинуклеотид, кодирующий указанную О-тРНК.

12. Способ получения в системе трансляции белка, содержащего фенилселеноцистеин в выбранном положении, который включает:

(a) обеспечение системы трансляции, включающей (i) композицию по любому из пп.1-11,

(ii) нуклеиновую кислоту, находящуюся в контакте с (i), кодирующую указанный белок, причем указанная нуклеиновая кислота содержит по меньшей мере один селекторный кодон, который распознает используемая О-тРНК, и

(iii) клеточные факторы, необходимые для процесса трансляции;

(b) создание и поддержание условий, обеспечивающих трансляцию указанного белка.

13. Способ по п.12, согласно которому указанная О-РС избирательно аминокислирует указанную О-тРНК указанным фенилселеноцистеином с эффективностью, которая составляет по меньшей мере 50% эффективности, наблюдаемой для системы трансляции, содержащей указанную О-тРНК, указанный фенилселеноцистеин, и аминоксил-тРНК синтетазу, содержащую последовательность аминокислот SEQ ID NO:4, 6 или 8.

14. Способ по п.12, согласно которому указанная система трансляции содержит полинуклеотид, кодирующий указанную О-РС.

15. Способ по п.12, согласно которому указанная или в котором указанное обеспечение системы трансляции включает обеспечение О-РС, содержащей последовательность аминокислот, выбранную из группы, состоящей из последовательностей аминокислот, обозначенных SEQ ID NO:4, 6, 8, и их консервативных вариантов.

16. Способ по п.12, согласно которому указанное обеспечение ортогональной

аминоацил-тРНК синтетазы включает осуществление мутации кармана связывания аминокислоты аминоацил-тРНК синтетазы дикого типа способом сайт-направленного мутагенеза и отбор полученной О-РС, которая избирательно аминоацилирует указанную О-тРНК фенилселеноцистеином.

5 17. Способ по п.16, согласно которому указанная стадия отбора включает положительный отбор и отрицательный отбор указанной О-РС из пула, включающего множество молекул мутантной аминоацил-тРНК синтетазы, полученной в результате указанного сайт-направленного мутагенеза.

10 18. Способ по п.12, согласно которому указанное обеспечение системы трансляции включает обеспечение полинуклеотида, кодирующего указанную О-тРНК.

19. Способ по п.12, согласно которому указанное обеспечение системы трансляции включает обеспечение О-тРНК, которая представляет собой тРНК, являющуюся супрессором амбер-кодона.

15 20. Способ по п.12, согласно которому указанное обеспечение системы трансляции включает обеспечение нуклеиновой кислоты, которая кодирует или содержит указанную О-тРНК, причем указанная нуклеиновая кислота содержит последовательность полинуклеотида, обозначенную SEQ ID NO:1.

20 21. Способ по п.12, согласно которому указанное обеспечение системы трансляции включает обеспечение нуклеиновой кислоты, которая кодирует указанную О-РС, причем указанная нуклеиновая кислота содержит последовательность полинуклеотида, обозначенную SEQ ID NO: 5, 7 или 9.

25 22. Способ по п.12, согласно которому указанное обеспечение системы трансляции включает обеспечение нуклеиновой кислоты, содержащей селекторный амбер-кодон.

23. Способ по п.12, согласно которому указанное обеспечение системы трансляции включает обеспечение клетки-хозяина, где указанная клетка-хозяин содержит фенилселеноцистеин, указанную О-РС, указанную О-тРНК и указанную нуклеиновую кислоту; указанная стадия (b) включает культивирование указанной клетки-хозяина, причем указанная клетка-хозяин представляет собой клетку-хозяина эубактерии.

24. Способ по п.23, согласно которому клетка-хозяин представляет собой клетку-хозяина E.coli.

35 25. Способ получения белка, содержащего липидный фрагмент в выбранном положении, причем указанный способ включает:

(a) обеспечение системы трансляции, включающей

(i) композицию по любому из пп.1-11,

40 (ii) нуклеиновую кислоту, находящуюся в контакте с (i), кодирующую указанный белок, где указанная нуклеиновая кислота содержит селекторный кодон в указанном выбранном положении, распознаваемом используемой О-тРНК, и

(iii) клеточные факторы, необходимые для процесса трансляции;

(b) создание и поддержание условий, обеспечивающих трансляцию указанного белка;

45 (c) выделение указанного белка;

(d) осуществление реакции окислительного расщепления, в результате чего происходит преобразование указанного фенилселеноцистеина в выбранном положении в указанном белке в дегидроаланин;

50 (e) осуществление реакции полученного на стадии (d) продукта с липидным фрагментом с образованием белка, включающего липидированный фрагмент аминокислоты в выбранном положении в белке.

26. Способ по п.25, согласно которому указанная реакция указанного

фенилселеноцистеина в выбранном положении в указанном белке представляет собой окислительное отщепление или в котором указанная реакция указанного фенилселеноцистеина в выбранном положении в указанном белке включает

5 приведение указанного фенилселеноцистеина в контакт с перекисью водорода.
27. Способ по п.25, согласно которому указанная реакция указанного дегидроаланина в выбранном положении в указанном белке представляет собой реакцию присоединения Майкла.

28. Способ по п.25, согласно которому указанный липидный фрагмент выбран из 10 группы, состоящей из тиопальмитиновой кислоты, фарнезилмеркаптана и 1-гексадекантиола, или в котором указанный липидированный фрагмент аминокислоты выбран из группы, состоящей из пальмитоилцистеина, фарнезилцистеина и S-гексадецилцистеина.

29. Способ получения в системе трансляции белка, содержащего не встречающуюся в природе аминокислоту дегидроаланин в выбранном положении, причем указанный 15 способ включает:

(a) обеспечение системы трансляции, включающей

(i) композицию по любому из пп.1-11,

20 (ii) нуклеиновую кислоту, находящуюся в контакте с (i), кодирующую указанный белок, где указанная нуклеиновая кислота содержит селекторный кодон в указанном выбранном положении, распознаваемом используемой О-тРНК, и (iii) клеточные факторы, необходимые для процесса трансляции;

25 (b) создание и поддержание условий, обеспечивающих трансляцию указанного белка;

(c) выделение указанного белка;

30 (d) осуществление реакции окислительного расщепления, в результате чего происходит преобразование указанного фенилселеноцистеина в выбранном положении в указанном белке в дегидроаланин.

30. Способ по п.29, в котором указанная реакция указанного фенилселеноцистеина в выбранном положении в указанном белке представляет собой окислительное отщепление.

31. Способ по п.29, согласно которому указанная реакция указанного 35 фенилселеноцистеина в выбранном положении в указанном белке включает приведение указанного фенилселеноцистеина в контакт с перекисью водорода.

32. Полипептид О-РС, характеризующийся последовательностью аминокислот, 40 выбранной из группы, состоящей из последовательностей аминокислот, которые по меньшей мере на 90% идентичны SEQ ID NO:4, 6 или 8 и содержат Trp в положении, соответствующем положению 32 последовательности SEQ ID NO:4, 6 или 8; Glu или His в положении, соответствующем положению 65 последовательности SEQ ID NO:4, 6 или 8; Gly в положении, соответствующем положению 70 последовательности SEQ ID NO:4, 6 или 8; Gin, Ser или Glu в положении, соответствующем положению 158 45 последовательности SEQ ID NO:4, 6 или 8; Ser или Glu в положении, соответствующем положению 162 последовательности SEQ ID NO:4, 6 или 8; или консервативных вариантов указанных последовательностей.

33. Полипептид по п.32, отличающийся тем, что указанный полипептид находится в 50 клетке.

34. Полинуклеотид, кодирующий полипептид О-РС, характеризующийся нуклеотидной последовательностью, которая определяет аминокислотную последовательность полипептида по п.32.

35. Полинуклеотид по п.34, отличающийся тем, что указанный полинуклеотид имеет последовательность нуклеотидов SEQ ID NO:5, 7 или 9.

36. Вектор экспрессии, содержащий полинуклеотид по п.34.

5 37. Клетка-хозяин, содержащая полинуклеотид по п.34, которая способна продуцировать полипептид O-PC, избирательно ацилирующий O-тРНК фенилселеноцистеином.

10

15

20

25

30

35

40

45

50

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

- <110> ЗЕ СКРИПС РЕСЕЧ ИНСТИТЬЮТ
 ВАНГ, Джиангюн
 ШУЛЬЦ, Питер Г.
- <120> ГЕНЕТИЧЕСКИ ЗАПРОГРАММИРОВАННАЯ ЭКСПРЕССИЯ БЕЛКОВ СОДЕРЖАЩИХ НЕ
 ВСТРЕЧАЮЩИХСЯ В ПРИРОДЕ АМИНОКИСЛОТУ ФЕНИЛСЕЛЕНОЦИСТЕИН
- <130> 54-001820US
- <140> US 11/724,039
- <141> 2007-03-13
- <160> 9
- <170> патентная версия 3.4
- <210> 1
- <211> 77
- <212> RNA
- <213> Artificial
- <220>
- <223> nonsense suppressor tyrosyl-tRNA CUA derived from Methanococcus
 jannaschii
- <400> 1
 ccggcgguag uucagcaggg cagaacggcg gacucuaaaу ccgcauggcg cugguucaaa 60
 uccggcccgc cggacca 77
- <210> 2
- <211> 306
- <212> PRT
- <213> Methanococcus jannaschii
- <400> 2
- 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
- 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
- Leu Leu Ala Asp Leu His Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp
 65 70 75 80
- 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 aaagagaaac acatctgaaa ttatcagcga ggaagagtta	60
agagaggttt taaaaaaga tgaaaaatct gttacatag gttttgaacc aagtggtaaa	120
atacatttag ggcattatct ccaataaaa aagatgattg atttacaana tgctggattt	180
gatataatta tattgttggc tgatttacac gcctatttaa accagaaagg agagttggat	240
gagattagaa aataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca	300
aaatatgttt atggaagtga attccagctt gataaggatt atacactgaa tgtctataga	360

ttggctttaa aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420
gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcagggtta tgatattcat 480
tatttaggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgtttagca 540
agggagcttt taccaaaaaa ggttgtttgt attcacaacc ctgtcttaac gggtttggat 600
ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660
gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720
ataatggaga tagctaaata cttccttgaa tatkctttaa ccataaaaag gccagaaaaa 780
tttgggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
gaattgcatc caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
ccaattagaa agagatta 918

<210> 4
<211> 306
<212> PRT
<213> Artificial

<220>
<223> Phenylselenocysteine aminoacyl tRNA synthetase

<400> 4

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 Trp
20 25 30

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

Glu Leu Ala Asp Leu Gly Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp
65 70 75 80

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 Gln Ile His
145 150 155 160

Tyr Ser 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> 5
- <211> 921
- <212> DNA
- <213> Artificial

- <220>
- <223> Phenylselenocysteine aminoacyl tRNA synthetase

<400> 5
 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
 agagagggtt taaaaaaga tgaaaaatct gcttggatag gttttgaacc aagtggtaaa 120
 atacatttag ggcattatct ccaaataaaa aagatgattg atttcaaaa tgctggattt 180
 gatataatta tagagttggc tgatttaggg gcctatttaa accagaaagg agagttggat 240
 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
 aaatatgttt atggaagtga attccagctt gataaggatt atacactgaa tgtctataga 360
 ttggctttaa aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420
 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcaggttaa tcaaattcat 480
 tatagtggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catggttagca 540

aggagcttt taccaaaaaa ggttgttgt attcacaacc ctgtcttaac gggtttggat 600
 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660
 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttggtga aggaaatcca 720
 ataatggaga tagctaaata cttccttgaa taccctttaa ccataaaaag gccagaaaaa 780
 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
 gaattgcatc caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
 ccaattagaa agagattata a 921

<210> 6
 <211> 306
 <212> PRT
 <213> Artificial

<220>
 <223> Phenylselenocysteine aminoacyl tRNA synthetase
 <400> 6

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 Trp
 20 25 30
 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
 His Leu Ala Asp Leu Gly Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp
 65 70 75 80
 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 Asn Ser 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 Ser Ile His
 145 150 155 160
 Tyr Glu 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> 7
- <211> 921
- <212> DNA
- <213> Artificial

<220>
 <223> Phenylselenocysteine aminoacyl tRNA synthetase

<400> 7
 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
 agagaggttt taaaaaaga tgaaaaatct gcttggatag gttttgaacc aagtggtaaa 120
 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaaa tgctggattt 180
 gatataatta tacatttggc tgatttaggc gcctatttaa accagaaagg agagttggat 240
 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
 aaatatgttt atggaagtga aaattctctt gataaggatt atacactgaa tgtctataga 360
 ttggctttaa aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420
 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcagggtta ttctattcat 480
 tatgagggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540
 agggagcttt taccaaaaaa gttgtttgt attcacaacc ctgtcttaac gggtttgat 600
 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660
 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720

ataatggaga tagctaaata cttccttgaa tatkctttaa ccataaaaag gccagaaaa 780
 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
 gaattgcatc caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
 ccaattagaa agagattata a 921

<210> 8
 <211> 306
 <212> PRT
 <213> Artificial

<220>
 <223> Phenylselenocysteine aminoacyl tRNA synthetase
 <400> 8

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 Trp
 20 25 30

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

His Leu Gly Asp Leu Gly Ala Tyr Leu Asn Gln Lys Gly Glu Leu Asp
 65 70 75 80

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 Lys Ser 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 Glu Ile His
 145 150 155 160

Tyr Glu 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

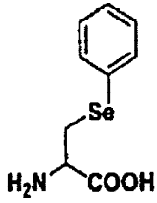
Asn Ala Val Ala Glu Glu Leu Ile Lys Ile Leu Glu Pro Ile Arg Lys
 290 295 300

Arg Leu
 305

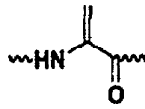
<210> 9
 <211> 921
 <212> DNA
 <213> Artificial

<220>
 <223> Phenylselenocysteine aminoacyl tRNA synthetase

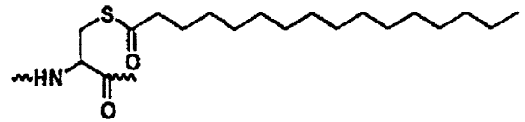
<400> 9
 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
 agagagggtt taaaaaaga tgaaaaatct gcttggatag gttttgaacc aagtggtaaa 120
 atacatttag ggcattatct ccaaataaaa aagatgattg atttcaaaa tgctggattt 180
 gatataatta tacatttggg tgatttaggc gcctatttaa accagaaagg agagttggat 240
 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
 aaatatgttt atggaagtga aaagtctctt gataaggatt atacactgaa tgtctataga 360
 ttggctttaa aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420
 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcagggtta tgagattcat 480
 tatgagggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgtttagca 540
 agggagcttt taccaaaaaa ggttgtttgt attcacaacc ctgtcttaac gggtttggat 600
 ggagaaggaa agatgagttc ttcaaaaagg aattttatag ctgttgatga ctctccagaa 660
 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720
 ataatggaga tagctaaata cttccttgaa tctcctttaa ccataaaaag gccagaaaaa 780
 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
 gaattgcatc caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
 ccaattagaa agagattata a 921



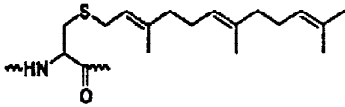
phenylselenocysteine
(1)



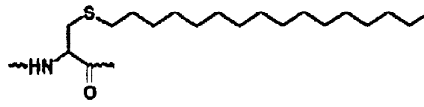
dehydroalanine
(2)



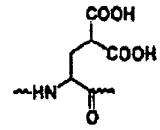
palmitoylcysteine
(3)



farnesylcysteine
(4)



S-hexadecylcysteine
(5)



γ-carboxyglutamic acid
(6)

ФИГ. 1

Nucleotide and Amino Acid Sequences

SEQ ID NO:	Description	SEQUENCE
1	<i>Methanococcus jannaschii</i> - suppressor tyrosyl-tRNA ^{CUA} aka Mj tRNA-Tyr(CUA) or mutRNA ^{Tyr} _{CUA}	CCGGCGGUAGUUCAGCAGGGCAGAACGGCGGACUCUAAAUCCG CAUGGGCCUGGUUCAAAAUCGGGCCCGCGGACCA
2	Wild-type <i>Methanococcus jannaschii</i> tyrosyl-tRNA synthetase (MjTyrRS) amino acid sequence	MDEFEMIKRNTSEI I SEELREVLKDKDEKSAYIGFEPSSGKIHL GHYLQIKKIMIDLQNAQFDI I ILLADLHAYLNQKGELEIRKIG DYNKKVFEAMGLKAKYVYVYSEFQLDKDYTLNVYRLALKTTLKR ARRSMELIAREDENPKVAEVIYPIMQVNDIHYLGVDAVGGME QRKIHMLARELLPKKVVCIHNPVLTGLDGEKMSSSKGNFIAV DDSPEEIRAKIKKAYCPAGVVEGNPIMEIAKYFLEYPLTIKRP EKPFGGDLTVNSYEELES LFKNKELHPMDLKNVAEELIKILEP IRKRL
3	Wild-type <i>Methanococcus jannaschii</i> tyrosyl-tRNA synthetase (MjTyrRS) nucleotide sequence	ATGGACGAATTGAAATGATAAAGAGAAACACATCTGAAATTA TCAGCGAGGAGAGTTAAGAGAGGTTTAAAAAAGATGAAAA ATCTGCTTACATAGGTTTGAACCAAGTGGTAAAAATACATTA GGCATTATCTCCAAATAAAAAAGATGATTGATTTACAAAATG CTGGATTTGATAAATATATATTGTTGGCTGATTTACACGCCA TTTAAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATAGGA GATTATAACAAAAAAGTTTGAAGCAATGGGGTTAAAGGCAA AATATGTTTATGGAAGTGAATTCAGCTTGATAAGGATTATAC ACTGAATGCTATAGATTGGCTTTAAAACTACCTTAAAAAGA GCAAGAAGGAGTATGGAACCTATAGCAAGAGAGGATGAAAATC CAAAGGTTGCTGAAGTTATCTATCCAAATAATGCAGGTTAATGA TATTCATTATTTAGGCCTTGATGTTGCAGTTGGAGGGATGGAG CAGAGAAAAATACACATGTTAGCAAGGGAGCTTTTACAAAAA AGGTTGTTTGTATTCAACAACCTGTC TTAACGGGTTTGGATGG AGAAGGAAAGATGAGTTCTTCAAAGGGAATTTTATAGCTGTT GATGACTCTCCAGAGAGATTAGGGCTAAGATAAAGAAAGCAT ACTGCCAGCTGGAGTTGTTGAAGGAAATCCAATAATGCAGAT AGCTAAATACTTCTTGAATATCCTTTAACCATAAAAAGGCCA GAAAAATTTGGTGGAGATTTGACAGTTAATAGCTATGAGGAGT TAGAGAGTTTATTTAAAAATAAGGAATGCATCCAATGGATTT AAAAATGCTGTAGCTGAAGAACTTATAAGATTTTAGAGCCA ATTAGAAAGAGATTA
4	CLONE 1 (SD) Phenylselenocysteine aminocyl-tRNA synthetase clone SD (PhSeRS-SD) amino acid sequence (derived from wild-type <i>Methanococcus jannaschii</i> tyrosyl tRNA-synthetase)	MDEFEMIKRNTSEI I SEELREVLKDKDEKSAWIGFEPSSGKIHL GHYLQIKKIMIDLQNAQFDI I IELADL GAYLNQKGELEIRKI GDYNNKVF EAMGLKAKYVYVYSEFQLDKDYTLNVYRLALKTTLK RARRSMELIAREDENPKVAEVIYPIMQVNDIHYSGVDVAVGG MEQRKIHMLARELLPKKVVCIHNPVLTGLDGEKMSSSKGNFI AVDDSP E EIRAKIKKAYCPAGVVEGNPIMEIAKYFLEYPLTIK RPEKPFGGDLTVNSYEELES LFKNKELHPMDLKNVAEELIKIL EPIRKRL

Фиг. 2

SEQ ID NO:	Description	SEQUENCE
5	<p style="text-align: center;">CLONE 1 (SD) Phenylselenocysteine aminoacyl-tRNA synthetase clone SD (PhSeRS-SD) nucleotide sequence</p>	<p>ATGGACGAATTTGAAAATGATAAAGAGAAACACATCTGAAATTA TCAGCGAGGAAGAGTTAAGAGAGGTTTTAAAAAAGATGAAAA ATCTGCTTGGATAGGTTTTGAACCAAGTGGTAAAATACATTTA GGGCATTATCTCCAAATAAAAAAGATGATTTGATTTACAAAATG CTGGATTTGATATAAATATAGAGTTGGCTGATTTAGGGGCCTA TTTAAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATAGGA GATTATAACAAAAAAGTTTTTGAAGCAATGGGGTTAAAGGCCAA AATATGTTTTATGGAAGTGAATCCAGCTTGATAAGGATATAC ACTGAATGTCTATAGATTGGCTTTAAAAACTACCTTAAAAAGA GCAAGAAGGAGTATGGAACCTATAGCAAGAGAGGATGAAAATC CAAAGGTTGCTGAAGTTATCTATCCAATAATGCAGGTTAATCA AATTCATTATAGTGGCGTTGATGTTGCAGTTGGAGGGATGGAG CAGAGAAAAATACACATGTTAGCAAGGGAGCTTTTACCAAAAA AGGTTGTTTTGTATTCACAACCCTGTCTTAACGGGTTTGGATGG AGAAGGAAAGATGAGTTCTTCAAAGGGAATTTTATAGCTGTT GATGACTCTCCAGAAGAGATTAGGGCTAAGATAAAGAAAGCAT ACTGCCAGCTGGAGTTGTTGAAGGAAATCCAATAATGGAGAT AGCTAAATACTTCCTTGAATATCCTTAAACCATAAAAAGGCCA GAAAAATTTGGTGGAGATTTGACAGTTAATAGCTATGAGGAGT TAGAGAGTTTTATTTAAAAATAAGGAATTCATCCAATGGATTT AAAAATGCTGTAGCTGAAGAACTTATAAAGATTTTAGAGCCA ATTAGAAAAGAGATTATAA</p>
6	<p style="text-align: center;">CLONE 2 (K4) Phenylselenocysteine aminoacyl-tRNA synthetase clone K4 (PhSeRS-K4) amino acid sequence (derived from wild-type <i>Methanococcus jannaschii</i> tyrosyl tRNA-synthetase)</p>	<p>MDEFEMIKRNTSEIISEBELREVLKKEKSAWIGFEPSPGKIHL GHYLOIKKMIDLQNAFGDIIHBLADLGAYLNQKGLDEIRKI GDYNNKVFEMGLKAKYVYVGSNSLDKDYTLNVYRLALKKTL KRARRSMELIAREDENPKVAEVIYPIQVNSIHYEGVDVAVG GMEQRKIHMLARELLPKKVVCIHNPVLTGLDGEKMSSSKGNF IAVDDSPREIRAKIKKAYCPAGVVEGNPIMEIAKYFLEYPLTI KRPEKFGDLTVNSYEELESFLPKNELHPMDLKNVAEBLIKI LEPIRKRL</p>
7	<p style="text-align: center;">CLONE 2 (K4) Phenylselenocysteine aminoacyl-tRNA synthetase clone K4 (PhSeRS-K4) nucleotide sequence</p>	<p>ATGGACGAATTTGAAAATGATAAAGAGAAACACATCTGAAATTA TCAGCGAGGAAGAGTTAAGAGAGGTTTTAAAAAAGATGAAAA ATCTGCTTGGATAGGTTTTGAACCAAGTGGTAAAATACATTTA GGGCATTATCTCCAAATAAAAAAGATGATTTGATTTACAAAATG CTGGATTTGATATAAATATACATTTGGCTGATTTAGGCGCCTA TTTAAACCAGAAAGGAGAGTTGGATGAGATTAGAAAAATAGGA GATTATAACAAAAAAGTTTTTGAAGCAATGGGGTTAAAGGCCAA AATATGTTTTATGGAAGTGAATAATCTCTTGATAAGGATATAC ACTGAATGTCTATAGATTGGCTTTAAAAACTACCTTAAAAAGA GCAAGAAGCAGTATGGAACCTATAGCAAGAGAGGATGAAAATC CAAAGGTTGCTGAAGTTATCTATCCAATAATGCAGGTTAATTC TATTCATTATGAGGGCGTTGATGTTGCAGTTGGAGGGATGGAG CAGAGAAAAATACACATGTTAGCAAGGGAGCTTTTACCAAAAA AGGTTGTTTTGTATTCACAACCCTGTCTPAACGGGTTTGGATGG AGAAGGAAAGATGAGTTCTTCAAAGGGAATTTTATAGCTGTT GATGACTCTCCAGAAGAGATTAGGGCTAAGATAAAGAAAGCAT ACTGCCAGCTGGAGTTGTTGAAGGAAATCCAATAATGGAGAT AGCTAAATACTTCCTTGAATATCCTTAAACCATAAAAAGGCCA GAAAAATTTGGTGGAGATTTGACAGTTAATAGCTATGAGGAGT TAGAGAGTTTTATTTAAAAATAAGGAATTCATCCAATGGATTT AAAAATGCTGTAGCTGAAGAACTTATAAAGATTTTAGAGCCA ATTAGAAAAGAGATTATAA</p>

ФИГ. 2 (продолжение)

SEQ ID NO:	Description	SEQUENCE
8	<p>CLONE 3 (K5) Phenylselenocysteine aminoacyl-tRNA synthetase clone K5 (PhSeRS-K5) amino acid sequence (derived from wild-type <i>Methanococcus jannaschii</i> tyrosyl tRNA-synthetase)</p>	<p>MDEPEMIKRNTSEI I SEBELREVLKKDEKSAWIGFEPGSKIHL GHYLOIKKMIDLQNA GFDI I IHLGDDL GAYLNQKGE LDEIRKI GDYNNKVF EAMGLKAKYVYGSEKSLDKDYTLNVYRLALKTTL KRARRSMELIARE DENPKVAEVIYPIMQVNEIHYEGVDVAVG GMEQRKI HMLARELLPKKVVCIHNPVLTGLDGE GKMS SSKGNF IAVDDSP EIRAKIKKAYCPAGVVEGNPIMETAKYFLEYPLTI KRPEKFGGDLTVNSYEELES LFKNKLHPMDLKNVABELIKI LEPIRKRL</p>
9	<p>CLONE 3 (K5) Phenylselenocysteine aminoacyl-tRNA synthetase clone K5 (PhSeRS-K5) nucleotide sequence</p>	<p>ATGGACGAATTTGAAATGATAAAGAGAAACACATCTGAAATTA TCAGCGAGGAAGAGTTAAGAGAGGTTTTAAAAAAGATGAAAA ATCTGCTTGGATAGGTTTTGAACCAAGTGGTAAAAATACATTTA GGCATTATCTCCAAATAAAAAAGATGATGATTTACAAAATG CTGGATTTGATATAAATTATACATTTGGGTGATTTAGGCGCCTA TTTAAACCAGAAAGGAGAGTTGGATGACATTAGAAAAATAGGA GATTATAACAAAAAAGCTTTTTGAAGCAATGGGGTTAAAGGCAA AATATGTTTATGGAAGTGA AAAAGTCTCTTGATAAGGATTATAC ACTGAAATGTCTATAGATGGCTTTAAAAACTACCTTAAAAAGA GCAAGAAGGAGTATGGA ACTTATAGCAAGAGAGGATGAAAAATC CAAAGGTTGCTGAAOTTATCTATCCAAATAATGCAGGTTAATGA GATTCATTATGAGGCCGTTGATGTTGCAGTTGGAGGGATGGAG CAGAGAAAAATACACATGTTAGCAAGGGAGCTTTTACAAAAA AGGTTGTTTGTATTCA CAACCCTGTCTTACGGGTTTGGATGG AGAAGGAAAAGATGAGTCTTCAA AAGGGAATTTTATAGCTGTT GATGACTCTCCAGAAGAGATTAGGGCTAAGATAAAGAAAGCAT ACTGCCAGCTGGAGTTGTTGAAGGAAATCCAATAATGGAGAT AGCTAAATACTTCTTGAATATCCTTAAACCATAAAAAGGCCA GAAAAATTTGGTGGAGATTTGACAGTTAATAGCTATGAGGAGT TAGAGAGTTTATTTAAAAATAAGGAATTGCATCCAATGGATTT AAAAATGCTGTAGCTGAAGAACTTATAAAGATTTTAGAGCCA ATTAGAAAGAGATTATAA</p>

ФИГ. 2 (продолжение)