



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(51) МПК
C12N 9/02 (2006.01)
C12N 15/53 (2006.01)
C12N 15/63 (2006.01)
C12N 15/85 (2006.01)
C12Q 1/66 (2006.01)
C12Q 1/68 (2006.01)

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

C12N 9/0069 (2006.01); C12N 15/52 (2006.01); C12N 15/63 (2006.01); C12N 15/85 (2006.01); C12Q 1/66 (2006.01); C12Q 1/68 (2006.01); C12Y 113/12 (2006.01)

(21)(22) Заявка: 2017102986, 30.01.2017

(24) Дата начала отсчета срока действия патента:
30.01.2017

Дата регистрации:
13.12.2018

Приоритет(ы):

(22) Дата подачи заявки: 30.01.2017

(43) Дата публикации заявки: 30.07.2018 Бюл. № 22

(45) Опубликовано: 13.12.2018 Бюл. № 35

Адрес для переписки:

143026, Москва, территория инновационного центра "Сколково", 4, офис 125.4, ООО "Центр интеллектуальной собственности "Сколково"

(72) Автор(ы):

Ямпольский Илья Викторович (RU)

(73) Патентообладатель(и):

Общество с ограниченной ответственностью "Планта" (RU)

(56) Список документов, цитированных в отчете о поиске: GenBank: GAT53638.1, 05.04.2016. WO 2012109470 A2, 16.08.2012 A2. RU 2601128 C2, 27.10.2016. PAKULA et al., "Genetic analysis of protein stability and function", Annu Rev Genet., 1989; 23.

(54) НОВЫЕ ЛЮЦИФЕРАЗЫ И СПОСОБЫ ИХ ИСПОЛЬЗОВАНИЯ

(57) Реферат:

Группа изобретений относится к биохимии. Предложена нуклеиновая кислота, кодирующая люциферазу или ее функциональный фрагмент и выбранная из группы, состоящей из нуклеиновой кислоты, кодирующей белок, характеризующийся аминокислотной последовательностью, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34, нуклеиновой кислоты, кодирующей белок, имеющий последовательность, которая по крайней мере на 65% идентична аминокислотной последовательности, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34, и нуклеиновой кислоты, кодирующей

белок, в состав которого входит консенсусная последовательность SEQ ID NO: 35. Предложены также люцифераза или ее функциональный фрагмент, кодируемые указанной нуклеиновой кислотой. Также предложены экспрессионная кассета, клетка-хозяин, трансгенный организм, содержащие указанные молекулы нуклеиновых кислот. Кроме того, предложены набор для мечения организмов и способ мечения клеток и клеточных органелл. Группа изобретений обеспечивает расширение технических средств в области применения билюминесцентных систем. 8 н. и 18 з.п. ф-лы, 10 ил., 1 табл., 14 пр.



(51) Int. Cl.
C12N 9/02 (2006.01)
C12N 15/53 (2006.01)
C12N 15/63 (2006.01)
C12N 15/85 (2006.01)
C12Q 1/66 (2006.01)
C12Q 1/68 (2006.01)

(12) ABSTRACT OF INVENTION

(52) CPC

C12N 9/0069 (2006.01); *C12N 15/52* (2006.01); *C12N 15/63* (2006.01); *C12N 15/85* (2006.01); *C12Q 1/66* (2006.01); *C12Q 1/68* (2006.01); *C12Y 113/12* (2006.01)

(21)(22) Application: **2017102986, 30.01.2017**(24) Effective date for property rights:
30.01.2017Registration date:
13.12.2018

Priority:

(22) Date of filing: **30.01.2017**(43) Application published: **30.07.2018** Bull. № 22(45) Date of publication: **13.12.2018** Bull. № 35

Mail address:

**143026, Moskva, territoriya innovatsionnogo
tsentra "Skolkovo", 4, ofis 125.4, OOO "Tsentr
intellektualnoj sobstvennosti "Skolkovo"**

(72) Inventor(s):

Yampolskij Ilya Viktorovich (RU)

(73) Proprietor(s):

**Obshchestvo s ogranichennoj otvetstvennostyu
"Planta" (RU)**

(54) NEW LUCIFERASES AND METHODS OF THEIR USE

(57) Abstract:

FIELD: biochemistry.

SUBSTANCE: group of inventions relates to biochemistry. Nucleic acid encoding a luciferase or its functional fragment is proposed and selected from the group consisting of a nucleic acid encoding a protein, characterized by the amino acid sequence represented in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, or 34, a nucleic acid encoding a protein having a sequence, which is at least 65 % identical to the amino acid sequence shown in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 or 34, and a nucleic acid that encodes a protein, which includes

the consensus sequence SEQ ID NO: 35. Luciferase or its functional fragment encoded by the specified nucleic acid is also proposed. Expression cassette, a host cell, and a transgenic organism containing these nucleic acid molecules is also proposed. In addition, a kit for labeling organisms and a method for labeling cells and cell organelles is proposed.

EFFECT: group of inventions provides for the expansion of technical means in the application of bioluminescent systems.

26 cl, 10 dwg, 1 tbl, 14 ex

Область изобретения

[001] Это изобретение относится в основном к области биологии и химии. В частности, изобретение относится к люциферазам.

Уровень техники

5 [002] Биолюминесценцией называется способность биологических организмов или биомолекул к производству и эмиссии света. Способность к биолюминесценции определяется наличием специфического белка люциферазы или фотопротеина. Люциферазы - это ферменты, которые катализируют окисление низкомолекулярных соединений - люциферинов, превращая их в оксилюциферины. Окисление сопровождается
10 выделением света и высвобождением оксилюциферина.

[003] Фотопротеины также катализируют окисление люциферинов, однако, в этом случае люциферин выступает в роли простетической группы и образует стабильный комплекс с фотопротеином. Количество света, производимого фотопротеином, примерно пропорционально его концентрации, в то время как для люциферазы оно
15 зависит от концентраций как фермента, так и люциферина. Активация биолюминесцентной реакции, катализируемой фотопротеином, во многих случаях происходит в ответ на появление в среде ионов металлов. Например, фотопротеин акворин катализирует окисление люциферина (целентеразина) в ответ на высвобождение ионов кальция, что приводит к появлению коротких вспышек света.

20 [004] Люциферазы находят применение как репортерные гены во многих приложениях биомедицины и биотехнологии. В частности, они используются в методах диагностики, способах выявления в среде микроорганизмов, токсических агентов, при определении концентрации различных веществ, при выявлении активации сигнальных каскадов и т.д. [Scott et al., *Annu Rev Anal Chem*, 2011, 4: 297-319; Badr and Tannous, *Trends Biotechnol.*
25 2011, 29: 624-33; Andreu et al., *FEMS Microbiol Rev.* 2011, 35: 360-94]. Многие способы применения люцифераз описаны в обзорах [Kaskova et al., *Chem Soc Rev.*, 2016, 45: 6048-6077; Scott et al., *Annu Rev Anal Chem*, 2011, 4: 297-319; Widder and Falls, *IEEE Journal of Selected Topics in Quantum Electronics*, 2014, 20: 232-241].

[005] В настоящее время известны несколько типов биолюминесцентных систем,
30 показано, что у различных организмов они возникали в эволюции независимо более сорока раз [Herring, *Journal of Bioluminescence and Chemiluminescence*, 1987, 1: 147-63; Haddock et al., *Annual Review of Marine Science*, 2010; 2: 443-93].

[006] Описана люцифераза североамериканского светлячка *Photinus pyralis*, катализирующая окисление D-люциферина [de Wet et al., *Proc. Natl. Acad. Sci. USA*, 1985,
35 82: 7870-3; de Wet et al., *Proc. Natl. Acad. Sci. USA*, 1987, 7: 725-37]. Окисление D-люциферина сопровождается выделением желто-зеленого света с максимумом эмиссии при 560 нм. Тот же D-люциферин окисляется другими люциферазами насекомых: в настоящее время клонировано более 30 ферментов из различных видов насекомых из семейств Phengodidae, Elateridae and Lampyridae, излучающих свет с максимумами эмиссии
40 в диапазоне от 536 до 630 нм. Также описаны мутантные формы люцифераз насекомых и изготовлены синтетические аналоги D-люциферина, позволяющие получать пары люциферин-люцифераза с различными свойствами [Thorne et al., *Chem Biol.*, 2010, 17: 646-57. Несмотря на значительное разнообразие аналогов, сам D-люциферин остается наиболее предпочтительным субстратом биолюминесценции *in vivo* вследствие высокого квантового выхода реакции (0.88±0.25 [Seliger and McElroy, *Arch Biochem Biophys*, 1960,
45 88: 136-141]). Значительным затруднением при использовании данной системы является относительно большой размер люциферазы *Photinus pyralis* (61 кДа), что делает ее непригодной для создания ряда химерных конструкций (например, при изучении вирусов)

вследствие низкой стабильности их увеличенного генома [Tran et al., J Virol, 2013, 87: 13321-13329; Tran et al., Viruses, 2015, 7: 5319-5327], а также необходимость получения D-люциферина в чистой энантиомерной форме, поскольку его изомер L-люциферин является сильным конкурентным ингибитором реакции [Lembert, Biochem J, 1996, 317: 273-277]. Тот факт, что люцифераза *Photinus pyralis* не является секретлируемой, вносит дополнительные ограничения для количественной оценки сигнала билюминесценции *in vivo*.

[007] Описана также большая группа люцифераз и фотопротеинов, катализирующих окисление целентеразина. Целентеразин-зависимые билюминесцентные системы описаны, например, у *Renilla*, *Gaussia* и *Metridia longa* [O. Shimomura, Bioluminescence: Chemical Principles and Methods, World Scientific Publishing Co. Pte. Ltd, Singapore, 2006, 470 p.] и имеют широкое практическое применение. Также получены мутантные формы целентеразин-зависимых люцифераз и фотопротеинов и синтетические аналоги целентеразина [Kaskova et al., Chem Soc Rev., 2016, 45: 6048-6077]. Несмотря на наличие большого числа преимуществ целентеразиновой системы (секретлируемость, малые размеры и большое разнообразие доступных люцифераз), основные ограничения для ее применения связаны в первую очередь с нахождением максимума билюминесценции в зоне голубого света, вследствие чего большая его часть *in vivo* поглощается исследуемыми тканями. Кроме того, сам субстрат билюминесценции может испускать свет в процессе неферментативного окисления кислородом воздуха (этот процесс усиливается в присутствии супероксид-анионов и ионов пероксинитрита в тканях), внося шум в измеряемый билюминесцентный сигнал.

[008] Другой пример билюминесцентной системы описан у морских бактерий. Эта система значительно отличается от других билюминесцентных систем. Бактериальный люциферин (миристиновый альдегид), окисляется в процессе реакции, но не является эмиттером билюминесценции [O. Shimomura, Bioluminescence: Chemical Principles and Methods, World Scientific Publishing Co. Pte. Ltd, Singapore, 2006? 470 p.]. Помимо люциферина в качестве ключевых компонентов люминесцентной реакции выступают НАДН (никотинамидадениндинуклеотид) и ФМН-Н₂ (флавиномононуклеотид), окисленное производное которого и выступает в качестве истинного источника света. Билюминесцентные системы морских бактерий имеют ограничение по использованию, так как пригодны только для прокариотических систем экспрессии.

[009] Также известны билюминесцентные системы остракод рода *Cypridina*, которые характеризуются химически высокоактивным люциферинном и высокостабильной люциферазой. [Shimomura et al., Science, 1969, 164: 1299-300]. Одним из основных недостатков данной билюминесцентной системы является исключительно низкая стабильность самого люциферина *Cypridina* на воздухе, особенно в присутствии примесей. Пик билюминесценции люциферина лежит в диапазоне 448-463 нм (в зависимости от ионной силы раствора), что делает данную систему неподходящей для применения в неизменной форме в глубоких тканях *in vivo*.

[010] Также известны билюминесцентные системы динофлагеллят и эвфаузиид. В настоящее время клонированы гены, кодирующие три люциферазы из этой группы [O. Shimomura, Bioluminescence: Chemical Principles and Methods, World Scientific Publishing Co. Pte. Ltd, Singapore, 2006]. Существенным недостатком данной системы является ее неполная изученность: полные люциферазные последовательности еще не были установлены, и спектр возможных применений данной системы не определен. Сведения о механизме билюминесценции динофлагеллят и эвфаузиид пока что отрывочны.

[011] Несмотря на большое количество билюминесцентных систем, используемых

на сегодняшний день, сохраняется потребность в расширении линейки люциферин-люциферазных пар, обладающих новыми свойствами. В частности, интерес представляют АТФ- и NAD(P)H-независимые люциферазы, способные к окислению проникающих в клетки водорастворимых люциферинов.

5 [012] В этом плане большой интерес представляют люциферазы из грибов. Биолюминесценция грибов хорошо известна и упоминается еще в трактатах Аристотеля. Однако, биолюминесцентные системы грибов описаны весьма слабо. В 1959 году Airth and McElroy показали, что в состав биолюминесцентной системы грибов входят по
10 крайней мере теплочувствительный компонент - люцифераза и тепло-нечувствительный компонент - люциферин, и НАД(Ф)Н [Airth and McElroy, Journal of Bacteriology, 1959, 77: 249-50]. В 2015 году Purtov и коллеги определили люциферин грибов: им оказалась способная проникать через мембраны клеток молекула - 3-гидроксигиспидин [Purtov et al., Angewandte Chemie, 2015, 54: 8124-28]. Однако ни одной люциферазы грибов не было идентифицировано.

15 Сущность изобретения

[013] Настоящее изобретение обеспечивает выделенные молекулы нуклеиновых кислот, кодирующих новые люциферазы и их функциональные мутанты. Указанные люциферазы окисляют 3-гидроксигиспидин, что приводит к появлению света. Указанные люциферазы не зависят от АТФ и НАД(Ф)Н. В преимущественных воплощениях
20 указанные нуклеиновые кислоты выделены из грибов или получены с помощью генно-инженерных методов.

[014] В некоторых воплощениях нуклеиновая кислота настоящего изобретения кодирует люциферазу, выбранную из группы SEQ ID NO: 02, 04, 06, 08, 10, 12, 14, 16 или 18. Примеры нуклеотидных последовательностей показаны в SEQ ID NO: 1, 3, 5, 7, 9,
25 11, 13, 15, 17.

[015] В некоторых воплощениях нуклеиновая кислота настоящего изобретения кодирует люциферазу, содержащую характерную консенсусную аминокислотную последовательность, показанную в SEQ ID NO: 35.

[016] В некоторых воплощениях нуклеиновая кислота настоящего изобретения
30 кодирует функциональный фрагмент люциферазы, который короче с С- и/или N-конца природной люциферазы.

[017] В некоторых воплощениях нуклеиновая кислота настоящего изобретения кодирует люциферазу, аминокислотная последовательность которой по существу сходна люциферазе, выбранной из группы SEQ ID NO: 02, 04, 06, 08, 10, 12, 14, 16 или 18. В
35 некоторых воплощениях нуклеиновая кислота настоящего изобретения кодирует функциональный фрагмент люциферазы, аминокислотная последовательность которого по существу сходна функциональному фрагменту, выбранному из группы в SEQ ID NO: 20, 22, 24, 26, 28, 30, 32 или 34.

[018] Молекулы нуклеиновых кислот, которые отличаются от представленных
40 нуклеотидных последовательностей вследствие вырожденности генетического кода или гибридизуются с ними, также входят в рамки настоящего изобретения.

[019] В других воплощениях также обеспечиваются векторы, включающие нуклеиновую кислоту настоящего изобретения. Кроме того, настоящее изобретение обеспечивает кассеты экспрессии, включающие нуклеиновую кислоту настоящего
45 изобретения и регуляторные элементы, необходимые для экспрессии нуклеиновой кислоты в выбранной клетке-хозяине. Кроме того, обеспечиваются клетки, стабильные клеточные линии, трансгенные организмы (например, растения, животные, грибы, микроорганизмы), включающие нуклеиновые кислоты, векторы или экспрессионные

кассеты настоящего изобретения.

[020] В других воплощениях обеспечиваются функциональные люциферазы настоящего изобретения, которые кодируются нуклеиновыми кислотами указанными выше.

5 [021] Кроме того, обеспечивается набор, содержащий нуклеиновые кислоты или векторы или экспрессионные кассеты, содержащие нуклеиновые кислоты настоящего изобретения.

[022] Кроме того, обеспечиваются антитела, специфически связывающие белки настоящего изобретения или их фрагменты.

10 [023] Кроме того, обеспечиваются способы мечения клеток, клеточных структур и биомолекул с помощью нуклеиновых кислот и белков настоящего изобретения.

[024] Технический результат состоит в расширении технических средств в области применения билюминесцентных систем и достигается за счет идентификации аминокислотных и нуклеотидных последовательностей новой группы ферментов,
15 обладающих способностью катализировать окисление 3-гидроксигиспидина, сопровождающееся испусканием света.

Краткое описание фигур

[025] Фигура 1 показывает спектры билюминесценции люциферазы из *Neonothopanus nambi* при экспрессии в естественной (1) и гетерологической (2) системах.

20 [026] Фигуры 2 и 3 иллюстрируют характер изменения интенсивности билюминесценции клетками линии HeLa Kyoto (Фиг. 2) и HEK293T (Фиг. 3), экспрессирующими люциферазы *Neonothopanus nambi* (1), *Armillaria mellea* (2), *Mycena citricolor* (3), *Armillaria ostoyae* (4), *Mycena chlorophos* (5), *Armillaria gallica* (6), *Panellus stipticus* (7), *Omphalotus olearius* (8), а также нетрансфицированными контрольными
25 клетками (9) со временем.

[027] Фигура 4 показывает фотографию вестерн-блота с антителами к гистидиновой метке, конъюгированными с пероксидазой хрена. Дорожка 1: пост-ядерный супернатант клеток, экспрессирующих люциферазу *Neonothopanus nambi*, не несущую гистидиновой метки. Дорожка 2: пост-ядерный супернатант клеток, экспрессирующих люциферазу
30 *Neonothopanus nambi*, имеющую гистидиновую метку. Дорожка 3: супернатант, полученный центрифугированием при 140000 g. Дорожка 4: осадок, полученный центрифугированием при 140000 g.

[028] Фигура 5 иллюстрирует множественное выравнивание аминокислотных последовательностей люцифераз настоящего изобретения. Трансмембранный N-концевой
35 домен подчеркнут. Сверху показана консенсусная последовательность.

[029] Фигура 6 показывает фотографию градиентного полиакриламидного геля (10-25%), окрашенного Coomassie Blue с нанесением лизата клеток *E. coli* до индукции IPTG (дорожка 1); лизата клеток *E. coli* после ночной экспрессии при комнатной температуре (дорожка 2), супернатанта после лизиса клеток и центрифугирования при 35000 g
40 (дорожка 3); экстракта из телец включения *E. Coli* (дорожка 4).

[030] Фигура 7 иллюстрирует зависимость интенсивности билюминесценции рекомбинантой люциферазы из *Neonothopanus nambi* от pH.

[031] Фигура 8 показывает наложение микрофотографий, снятых в проходящем свете с микрофотографиями, снятыми в красном канале пропускания (сверху, mKate2)
45 и в зеленом канале пропускания (снизу, люцифераза).

[032] Фигура 9 иллюстрирует мечение клеток опухоли внутри живой мыши с помощью люциферазы *Neonothopanus nambi* (вживлена в левую часть спины мыши) и люциферазы *Photinus pyralis* (вживлена в правую часть). Показано наложение фотографий мыши и

зарегистрированного светового сигнала от вживленных опухолей.

[033] Фигура 10 иллюстрирует мечение нервной системы эмбриона *Xenopus laevis* (стадии 16-17 нейруляции) с помощью люциферазы *Neonothopanus nambi*. Показана биоллюминесценция от люциферазы.

5 Подробное описание изобретения

[034] Настоящее изобретение направлено на молекулы нуклеиновых кислот, которые кодируют новые люциферазы, их функциональные фрагменты, гомологи и мутанты, а также на белки, кодируемые указанными нуклеиновыми кислотами. Молекулы нуклеиновых кислот, представляющие интерес, выделены из грибов или получены
10 методами генной инженерии. Также обеспечиваются клетки-хозяева, стабильные клеточные линии и трансгенные организмы, содержащие указанные молекулы нуклеиновых кислот. Кроме того, обеспечиваются антитела, специфичные к белкам настоящего изобретения.

[035] Указанные белковые и нуклеотидные композиции применяются во многих
15 приложениях и методах, в частности, в приложениях мечения организмов, клеток, клеточных органелл или белков. Также указанные белковые и нуклеотидные композиции применяются в методах выявления белок-белковых взаимодействий, при тестировании активности промоторов в различных условиях. Наконец, обеспечиваются наборы для их использования в таких методах и приложениях.

20 Определения

[036] Различные термины, относящиеся к объектам настоящего изобретения, используются выше и также в описании и в формуле изобретения.

[037] Как здесь используется, термин "люцифераза" означает белок, который обладает способностью к окислению люциферина, где реакция окисления сопровождается
25 выделением света (люминесценцией) и происходит освобождение окисленного люциферина.

[038] Термин "гуманизированный" относится к изменению нуклеотидной последовательности белка, сделанной для оптимизации генетического кода кодонов для экспрессии в клетках млекопитающих (Yang et al., 1996, *Nucleic Acids Research* 24:
30 4592-4593).

[039] Как здесь используется, термин "выделенный" означает молекулу или клетку, которые находятся в среде, отличной от среды, в которой молекула или клетка находятся в естественных условиях.

[040] Как здесь используется, термин "мутант" или "производное" относятся к белку,
35 раскрытому в настоящем изобретении, в котором одна или более аминокислот добавлены и/или замещены и/или удалены (делетированы) и/или вставлены (инsertированы) в N-конец и/или C-конец, и/или в пределах нативных аминокислотных последовательностей белков настоящего изобретения. Как здесь используется, термин "мутант" относится к молекуле нуклеиновой кислоты, которая кодирует мутантный
40 белок. Кроме того, термин "мутант" здесь относится к любому варианту, который короче или длиннее белка или нуклеиновой кислоты.

[041] Термин "гомология" используется для описания взаимосвязи последовательностей нуклеотидов или аминокислот с другими последовательностями нуклеотидов или аминокислот, которая определена степенью идентичности и/или
45 сходства между указанными сравниваемыми последовательностями.

[042] Как здесь используется, аминокислотная или нуклеотидная последовательности "по существу сходны" или "по существу такие же" как референсная последовательность, если аминокислотная или нуклеотидная последовательности имеют по крайней мере

40% идентичности с указанной последовательностью внутри выбранного для сравнения региона. Таким образом, по существу сходные последовательности включают те, которые имеют, например, по крайней мере, 45% идентичности, или по крайней мере, 50% идентичности, или по крайней мере, 55% идентичности, или по крайней мере, 60% идентичности, или по крайней мере, 62% идентичности, или по крайней мере 65% идентичности, или по крайней мере 70% идентичности, или по крайней мере, 75% идентичности, например, по крайней мере, 80% идентичности, по крайней мере, 85% идентичности, по крайней мере, 90% идентичности (например, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 98% или 99% идентичности). Две последовательности, которые идентичны одна другой, так же по существу сходны. Для целей настоящего изобретения длина сравниваемых последовательностей люцифераз составляет по крайней мере 100 аминокислот, предпочтительно, по крайней мере, 200 аминокислот, например, 200-230 аминокислот или полную длину аминокислотной последовательности белка или функционального фрагмента. Для нуклеиновых кислот длина сравниваемых последовательностей в основном составляет, по крайней мере, 300 нуклеотидов; предпочтительно, по крайней мере, 600 нуклеотидов.

[043] Идентичность последовательностей определяется на основании референсной последовательности. Алгоритмы для анализа последовательности известны в данной области, такие как BLAST, описанный в Altschul et al., J. Mol. Biol., 215, pp. 403-10 (1990). Для целей настоящего изобретения для определения уровня идентичности и сходства между нуклеотидными последовательностями и аминокислотными последовательностями может быть использовано сравнение нуклеотидных и аминокислотных последовательностей, производимое с помощью пакета программ Blast, предоставляемого National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/blast>) с использованием содержащего разрывы выравнивания со стандартными параметрами.

[044] Как здесь используется, термин "функциональный" означает, что нуклеотидная или аминокислотная последовательность может функционировать для указанного испытания или задачи. Термин "функциональный", используемый для описания люцифераз, означает, что белок обладает способностью производить сопровождающуюся люминесценцией реакцию окисления люциферина.

[045] Как здесь используется, "биохимические свойства" относятся к белковому фолдингу (сворачиванию) и скорости созревания, времени полужизни, скорости катализа, pH и температурной стабильности, и другим подобным свойствам.

[046] Как здесь используется, "спектральные свойства" относятся к спектрам, квантовому выходу и интенсивности люминесценции и другим подобным свойствам.

[001] Термин "оперативно связанный" или ему подобный при описании химерных белков относиться к полипептидным последовательностям, которые находятся в физической и функциональной связи одна с другой. В наиболее предпочтительных воплощениях, функции полипептидных компонентов химерной молекулы не изменены по сравнению с функциональными свойствами выделенных полипептидных компонентов. Например, люцифераза настоящего изобретения может быть сшита с представляющим интерес партнером слияния. В этом случае химерный белок сохраняет свойства люциферазы, а представляющий интерес полипептид сохраняет его оригинальную биологическую активность. В некоторых воплощениях настоящего изобретения, активность как люциферазы, так и партнера слияния могут быть снижены по сравнению с активностями изолированных белков. Такие химерные белки также находят применение в рамках настоящего изобретения. Термин "оперативно связанный"

или ему подобный при описании нуклеиновых кислот означает, что нуклеиновые кислоты ковалентно связаны таким образом, что в местах их соединения отсутствуют 'сбои' рамки считывания и стоп-кодоны. Как очевидно для любого специалиста в данной области техники, нуклеотидные последовательности, кодирующие химерный белок, включающий "оперативно связанные" компоненты (белки, полипептиды, линкерные последовательности, белковые домены и т.д.), состоят из фрагментов, кодирующих указанные компоненты, где эти фрагменты ковалентно связаны таким образом, что в ходе транскрипции и трансляции нуклеотидной последовательности продуцируется полноразмерный химерный белок.

[047] Как здесь используется, термин "специфически гибридизуется" относится к ассоциации между двумя одноцепочечными молекулами нуклеиновых кислот или в достаточной степени комплементарными последовательностями, что разрешает такую гибридизацию в predetermined условиях, обычно используемых в данной области (иногда используется термин "по существу комплементарный").

[048] Ссылка на нуклеотидную последовательность, "кодирующую" полипептид, означает, что с нуклеотидной последовательности в ходе транскрипции мРНК и трансляции продуцируется этот полипептид. При этом может быть указана как кодирующая цепь, идентичная мРНК и обычно используемая в списке последовательностей, так и комплементарная цепь, которая используется как матрица при транскрипции. Как очевидно для любого специалиста в данной области техники, термин также включает любые вырожденные нуклеотидные последовательности, кодирующие одинаковую аминокислотную последовательность. Нуклеотидные последовательности, кодирующие полипептид, включают последовательности, содержащие интроны.

Молекулы нуклеиновых кислот

[049] Настоящее изобретение обеспечивает молекулы нуклеиновых кислот кодирующие люциферазы и их функциональные фрагменты (например кодирующие укороченные или удлиненные варианты люциферазы). В предпочтительных воплощениях нуклеиновая кислота настоящего изобретения кодирует люциферазу, аминокислотная последовательность которой выбрана из групп SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18. Молекулы нуклеиновых кислот, кодирующие гомологи и мутанты указанных люцифераз, по существу сходные с вышеперечисленными люциферазами, также находятся в рамках настоящего изобретения. Каждый из этих специфических типов молекул нуклеиновой кислоты, представляющей интерес, раскрывается более детально ниже в экспериментальной части.

[050] Как здесь используется, молекула нуклеиновой кислоты - это молекула ДНК, такая как геномная ДНК или кДНК, или молекула РНК, такая как молекула мРНК. В некоторых воплощениях, молекула нуклеиновой кислоты настоящего изобретения - это молекула ДНК (или кДНК), содержащая открытую рамку считывания, которая кодирует люциферазу настоящего изобретения и способна в подходящих условиях (например, при физиологических внутриклеточных условиях) быть использована для экспрессии люциферазы в гетерологической системе экспрессии.

[051] В некоторых воплощениях молекула нуклеиновой кислоты настоящего изобретения, кодирующая люциферазу, выделена из гриба. Особый интерес представляют нуклеиновые кислоты, кодирующие люциферазы из *Neonothopanus nambi* (SEQ ID NO: 2), *Armillaria gallica* (SEQ ID NO: 4), *Armillaria mellea* (SEQ ID NO: 6), *Armillaria ostoyae* (SEQ ID NO: 8), *Mycena chlorophos* (SEQ ID NO: 10), *Mycena citricolor* (SEQ ID NO: 12), *Omphalotus olearius* (SEQ ID NO: 14), и *Panellus stipticus* (SEQ ID NO: 16 и 18).

Нуклеотидные последовательности таких нуклеиновых кислот показаны в SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17.

[052] В некоторых воплощениях молекула нуклеиновой кислоты настоящего изобретения, кодирующая люциферазу, получена методами генной инженерии. Способы получения нуклеиновых кислот хорошо известны из области техники. Например, доступность информации о последовательности аминокислот (например, SEQ ID NO: 2-18) или информации о нуклеотидной последовательности (например, SEQ ID NO: 3, 5, 7, 9, 11, 13, 15 или 17) дает возможность изготовить выделенные молекулы нуклеиновых кислот настоящего изобретения с помощью олигонуклеотидного синтеза. В случае информации о последовательности аминокислот, может быть синтезировано несколько нуклеиновых кислот отличающихся друг от друга вследствие вырожденности генетического кода. Методы выбора вариантов кодонов для требуемого хозяина хорошо известны в данной области.

[053] Синтетические олигонуклеотиды могут быть приготовлены с помощью фосфорамидитного метода, и полученные конструкторы могут быть очищены с помощью методов хорошо известных в данной области, таких как высокоэффективная жидкостная хроматография (ВЭЖХ) или других методов как описано, например, в Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., (1989) Cold Spring Harbor Press, Cold Spring Harbor, NY, и по инструкции, описанной в, например, United States Dept. of HHS, National Institute of Health (NIH) Guidelines for Recombinant DNA Research. Длинные двухцепочечные молекулы ДНК настоящего изобретения могут быть синтезированы следующим образом: могут быть синтезированы несколько меньших фрагментов с необходимой комплементарностью, которые содержат подходящие концы, способные к когезии с соседним фрагментом. Соседние фрагменты могут быть сшиты с помощью ДНК-лигазы, методов, основанных на рекомбинации, или метода, основанного на ПЦР.

[054] Молекулы нуклеиновой кислоты, кодирующие люциферазы настоящего изобретения, могут быть также клонированы из биологических источников из гриба, относящегося к типу Basidiomycota, преимущественно к классу Basidiomycetes, в частности, к отряду Agaricales.

[055] Настоящее изобретение также охватывает нуклеиновые кислоты, которые гомологичны, по существу сходны, идентичны, или получены из нуклеиновых кислот, кодирующих белки настоящего изобретения.

[056] Нуклеиновые кислоты изобретения находятся в среде, отличной от среды, в которой они находятся в естественных условиях, например, они выделены, представлены в увеличенном количестве, находятся или экспрессированы в системах *in vitro* или в клетках или организмах, отличных от тех, в которой они находятся в естественных условиях.

[057] Изменения или различия в нуклеотидной последовательности между высоко сходными нуклеотидными последовательностями могут представлять нуклеотидные замены в последовательности, которые возникают в процессе нормальной репликации или дубликации. Другие замены могут быть специально рассчитаны и вставлены в последовательность для определенных целей, таких, как изменение кодонов определенных аминокислот или нуклеотидной последовательности регуляторного региона. Такие специальные замены могут быть произведены *in vitro* с помощью различных технологий мутагенеза или получены в организмах-хозяевах, находящихся в специфических селекционных условиях, которые индуцируют или отбирают эти изменения. Такие специально полученные варианты последовательности могут быть названы "мутантами" или "производными" исходной последовательности.

[058] Нуклеиновая кислота, кодирующая люциферазу, по существу сходную с люциферазой, аминокислотная последовательность которой показана в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, также обеспечивается настоящим изобретением. Нуклеиновая кислота, кодирующая такой полипептид или его фрагмент, может быть выделена любым из многих известных методов. Фрагмент кДНК настоящего изобретения может использоваться как гибридизационная проба против библиотеки кДНК из целевого организма в условиях высокой строгости. Проба может быть большим фрагментом или более коротким вырожденным праймером. Нуклеиновые кислоты, имеющие сходство последовательности, могут быть детектированы гибридизацией в условиях высокой строгости, например, при 50°C или выше (например, 60°C или 65°C), 50% формамид, 0.1 × SSC (15 mM хлорид натрия / 1.5 mM натрия цитрат), 0.1% SDS. Нуклеиновые кислоты, имеющие области, по существу идентичные к референсной последовательности, например, аллельные варианты, генетически-измененные варианты нуклеиновых кислот и т.д., связываются с референсной последовательностью в гибридных условиях высокой строгости. Используя пробы, в частности меченые пробы комплементарные к референсной последовательности ДНК, можно выделить подобные нуклеотидные последовательности.

[059] Нуклеиновая кислота, кодирующая такой полипептид или его фрагмент, может быть также идентифицирована в ходе геномного или транскриптомного секвенирования. В частности, по существу сходная люцифераза может быть идентифицирована среди последовательностей гипотетических белков, предсказанных на основе данных, полученных при секвенировании организмов, например, при секвенировании грибов различных видов, преимущественно относящихся к типу Basidiomycota, например, к классу Basidiomycetes, в частности, к отряду Agaricales.

[060] Нуклеиновая кислота, кодирующая люциферазу, содержащую специфическую консенсусную аминокислотную последовательность, по существу сходную с последовательностью, показанной в SEQ ID NO: 35, также обеспечивается настоящим изобретением. Термин «консенсусная последовательность» означает здесь усредненную аминокислотную последовательность, регулярно встречающуюся (с небольшими вариациями по отдельным аминокислотам) в разных люциферазах настоящего изобретения.

[061] Мутантные или производные нуклеиновые кислоты могут быть получены на матричной нуклеиновой кислоте, выбранной из вышеописанных нуклеиновых кислот, путем модификации, делеции или добавления одного или более нуклеотидов в матричной последовательности или их комбинации, для получения варианта матричной нуклеиновой кислоты. Модификации, добавления или делеции могут быть выполнены любым способом, известным в данной области (см. например, Gustin et al., *Biotechniques* (1993) 14: 22; Barany, *Gene* (1985) 37: 111-123; и Colicelli et al., *Mol. Gen. Genet.* (1985) 199: 537-539, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, (1989), CSH Press, pp. 15.3-15.108), включая подверженный ошибкам ПЦР (error-prone PCR), shuffling, олигонуклеотид-направленный мутагенез, ПЦР со сборкой, парный ПЦР мутагенез, мутагенез in vivo, кассетный мутагенез, рекурсивный множественный мутагенез, экспоненциальный множественный мутагенез, сайт-специфический мутагенез, случайный мутагенез, генное реассемблирование (gene reassembly), генный сайт-насыщающий мутагенез (GSSM), искусственную перестройку с лигированием (SLR) или их комбинации. Модификации, добавления или делеции могут быть также выполнены методом, включающим рекомбинацию, рекурсивную рекомбинацию последовательностей, фосфотиоат-модифицированный мутагенез ДНК, мутагенез на урацил-содержащей

матрице, мутагенез с двойным пропуском, точечный восстановительный по
рассогласованию мутагенез, мутагенез штамма, дефицитного по восстановлению,
химический мутагенез, радиоактивный мутагенез, делетационный мутагенез,
рестрикционно-избирательный мутагенез, рестрикционный мутагенез с очисткой, синтез
искусственных генов, множественный мутагенез, создание химерных множественных
нуклеиновых кислот и их комбинации. В некоторых воплощениях люциферазы,
кодируемые мутантными или производными нуклеиновыми кислотами, имеют те же
самые спектральные или биохимические свойства как люцифераза дикого типа. В других
воплощениях, мутантные или производные нуклеиновые кислоты кодируют люциферазы
с измененными свойствами.

[062] Также обеспечиваются вырожденные варианты нуклеиновых кислот, которые
кодируют белки настоящего изобретения. Вырожденные варианты нуклеиновых кислот
включают замены кодонов нуклеиновой кислоты на другие кодоны, кодирующие те
же самые аминокислоты. В частности, вырожденные варианты нуклеиновых кислот
создаются, чтобы увеличить экспрессию в клетке-хозяине. В этом воплощении, кодоны
нуклеиновой кислоты, которые не являются предпочтительными или являются менее
предпочтительными в генах клетки-хозяина, заменены кодонами, которые обильно
представлены в кодирующих последовательностях генов в клетке-хозяине, где указанные
замененные кодоны кодируют ту же самую аминокислоту. Примеры вырожденных
вариантов, представляющих интерес, описаны более подробно в экспериментальной
части ниже.

[063] Нуклеиновые кислоты, кодирующие укороченные и удлиненные варианты
указанных люцифераз, также входят в рамки настоящего изобретения. Как здесь
используется, эти варианты белков содержат аминокислотные последовательности с
измененными С-, N-, или обоими концами полипептидной цепи.

[064] В укороченных вариантах один или более (обычно до 39, чаще 37 или менее)
аминокислотных остатков могут быть удалены из последовательности или заменены
на любые другие аминокислотные остатки. В частности, может быть полностью или
частично удалена последовательность, кодирующая трансмембранный домен и
предшествующую ему аминокислотную последовательность с N-конца люциферазы.
Трансмембранный домен может быть идентифицирован с помощью способов, известных
из уровня техники: например, с использованием алгоритмов, описанных в [Krogh et al.,
Journal of Molecular Biology 2001, 305(3): 567-580] и [Sonnhammer et al., Proceedings of the
Sixth International Conference on Intelligent Systems for Molecular Biology, pages 175-182,
Menlo Park, CA, 1998. AAAI Press]. Для анализа может быть использовано основанное
на указанном алгоритме программное обеспечение, приведенное в [[http://www.cbs.dtu.dk/
services/ТМНММ/](http://www.cbs.dtu.dk/services/ТМНММ/)]. В удлиненных вариантах, С- или N-конец белка может содержать
дополнительные аминокислотные остатки. Примеры аминокислотных
последовательностей функциональных фрагментов и кодирующие их нуклеотидные
последовательности показаны в SEQ ID NO: 19-34. Для экспрессии функциональных
фрагментов кодирующие их нуклеиновые кислоты оперативно сшиваются с нуклеиновыми
кислотами, содержащими, по крайней мере, регуляторные последовательности и сайт
начала транскрипции. Также эти нуклеиновые кислоты могут содержать
последовательности, кодирующие гистидиновую метку (6 His tag), сигнальный пептид
или функциональные белковые домены.

[065] Вышеуказанные модификации не изменяют по существу спектральные свойства
люцифераз, но могут изменять внутриклеточную локализацию, облегчать белковый
фолдинг в клетке-хозяине, снижать способность к агрегации или модулировать другие

биохимические свойства белков, например, период полураспада. В некоторых воплощениях, эти модификации не изменяют биохимические свойства белка. Все виды модификаций и мутаций, указанные выше, как правило осуществляются на уровне нуклеиновой кислоты.

5 [066] Молекулы нуклеиновой кислоты изобретения могут кодировать весь или часть заявленного белка. Двух- и одноцепочечные фрагменты могут быть получены из последовательности ДНК путем химического синтеза олигонуклеотидов в соответствии со стандартными способами, энзиматической рестрикцией, амплификацией ПЦР и т.д. В основном, фрагменты ДНК будут иметь размер по меньшей мере приблизительно 15 нуклеотидов, обычно по меньшей мере приблизительно 18 нуклеотидов или 10 приблизительно 25 нуклеотидов и могут быть по меньшей мере приблизительно 50 нуклеотидов. В некоторых воплощениях, заявленные молекулы нуклеиновой кислоты могут иметь размер приблизительно 100, приблизительно 200, приблизительно 300, приблизительно 400, приблизительно 500, приблизительно 600 или более. Заявленные 15 нуклеиновые кислоты могут кодировать фрагменты заявленных белков или полные белки; например, заявленные нуклеиновые кислоты могут кодировать полипептиды приблизительно из 25 аминокислот, приблизительно из 50, приблизительно из 75, приблизительно из 100, приблизительно из 125, приблизительно из 150, приблизительно из 200 аминокислот вплоть до полной длины белка.

20 [067] Заявленные нуклеиновые кислоты могут быть выделены и получены по существу в очищенной форме. По существу очищенная форма означает, что нуклеиновые кислоты являются по меньшей мере приблизительно на 50% чистыми, обычно по меньшей мере приблизительно на 90% чистыми и обычно являются "рекомбинантными", то есть фланкированными одним или более нуклеотидами, с которыми последовательность 25 обычно не связана в хромосоме, встречающейся в природе в ее естественном организме-хозяине.

[068] Также обеспечиваются нуклеиновые кислоты, которые кодируют слитые белки, включающие белок настоящего изобретения, или их фрагменты, которые ниже рассматриваются более детально.

30 [069] Также обеспечиваются вектор и другие конструкции нуклеиновой кислоты, содержащие заявленные нуклеиновые кислоты. Подходящие векторы включают вирусные и невирусные векторы, плазмиды, космиды, фаги и т.д., предпочтительно плазмиды, и используются для клонирования, амплификации, экспрессии, переноса и т.д., последовательности нуклеиновой кислоты настоящего изобретения в подходящего 35 хозяина. Выбор подходящего вектора является понятным для квалифицированного специалиста в данной области. Полноразмерная нуклеиновая кислота или ее часть обычно вставляются в вектор посредством прикрепления ДНК-лигазой к расщепленному ферментами рестрикции сайту в векторе. Альтернативно, желательная нуклеотидная последовательность может быть вставлена гомологичной рекомбинацией *in vivo*, 40 обычно, присоединением гомологичных участков к вектору на флангах желательной нуклеотидной последовательности. Гомологичные участки добавляются лигированием олигонуклеотидов или полимеразной цепной реакцией, с использованием праймеров, включающих, например, как гомологичные участки, так и часть желательной нуклеотидной последовательности. Вектор, как правило, имеет ориджин репликации, 45 обеспечивающий его размножение в клетках-хозяевах в результате его введения в клетку как внехромосомного элемента. Вектор также может содержать регуляторные элементы, обеспечивающие экспрессию нуклеиновой кислоты в клетке-хозяине и получение рекомбинантной функциональной люциферазы. В экспрессионном векторе,

указанная нуклеиновая кислота является функционально связанной с регуляторной последовательностью, которая может включить промоторы, энхансеры, терминаторы, операторы, репрессоры, сайленсеры, инсуляторы и индукторы.

5 [070] Также обеспечиваются кассеты экспрессии или системы, использованные *inter alia* для получения заявленных люцифераз или химерных белков на их основе или для репликации заявленных молекул нуклеиновой кислоты. Кассета экспрессии может существовать как внехромосомный элемент или может быть включена в геном клетки в результате введения указанной кассеты экспрессии в клетку. Для экспрессии белковый продукт, кодируемый нуклеиновой кислотой изобретения, экспрессируется в любой 10 удобной системе экспрессии, включая, например, бактериальные системы, дрожжи, растения, насекомых, земноводных или клетки млекопитающих. В кассете экспрессии целевая нуклеиновая кислота оперативно соединяется с регуляторными последовательностями, которые могут включать промоторы, энхансеры, терминирующие последовательности, операторы, репрессоры и индукторы. Способы 15 получения кассет или систем экспрессии, способных экспрессировать желательный продукт, известны специалистам в данной области.

[071] Клеточные линии, которые устойчиво экспрессируют белки настоящего изобретения, могут быть выбраны способами, известными в данной области (например, ко-трансфекция с селективируемым маркером, таким как *dhfr*, *gpt*, генами устойчивости 20 к неомицину или гигромицину, что делает возможным выявление и выделение трансфицированных клеток, которые содержат ген, включенный в геном).

[072] Вышеописанные системы экспрессии могут использоваться в прокариотических или эукариотических хозяевах. Для получения белка могут использоваться клетки-хозяева, такие как *E. coli*, *B. subtilis*, *S. cerevisiae*, клетки насекомого, или клетки высшего 25 организма, не являющиеся эмбриональными клетками человека, такие как дрожжи, растения (например, *Arabidopsis thaliana*, *Nicotiana benthamiana*, *Physcomitrella patens*), позвоночные, например, COS 7 клетки, НЕК 293, СНО, ооциты *Xenopus* и т.д.

[073] Если используется любая вышеупомянутая клетка-хозяин или другие подходящие для репликации и/или экспрессии нуклеиновых кислот изобретения клетки-хозяева или организмы, то полученная реплицированная нуклеиновая кислота, 30 экспрессированный белок или полипептид находятся в рамках притязания изобретения как продукт клетки-хозяина или организма. Продукт может быть выделен подходящим способом, известным в данной области.

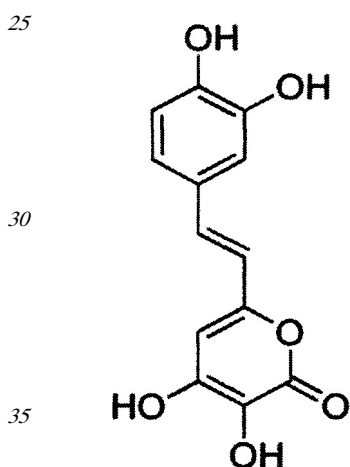
[074] Также обеспечиваются короткие фрагменты ДНК заявленных нуклеиновых 35 кислот, которые применяются как праймеры для ПЦР, амплификации *rolling circle*, гибридизационные скрининговые пробы и т.д. Длинные фрагменты ДНК применяются для получения кодируемых полипептидов, как ранее описано. Однако, для использования в геометрических реакциях амплификации, таких как ПЦР, используется пара коротких фрагментов ДНК, то есть праймеров. Точная последовательность 40 праймера не является критической для изобретения, но для большинства применений праймеры будут гибридизоваться с заявленной последовательностью в условиях строгости, как известно в данной области. Предпочтительно выбрать пару праймеров, которые дадут продукт амплификации по меньшей мере приблизительно из 50 нуклеотидов, предпочтительно по меньшей мере приблизительно из 100 нуклеотидов 45 и могут простираться на полную последовательность нуклеиновой кислоты. Алгоритмы отбора последовательностей праймеров обычно известны и доступны в коммерческих пакетах программ. Праймеры для амплификации гибридизуются с комплементарными цепочками ДНК и будут затравлять встречные реакции амплификации.

[075] Молекулы нуклеиновых кислот настоящего изобретения также могут применяться для определения экспрессии гена в биологическом образце. Способ, в котором исследуются клетки на наличие специфических нуклеотидных последовательностей, таких как геномная ДНК или РНК, хорошо отработан в данной области. Кратко, выделяют ДНК или мРНК из образца клетки. мРНК может быть амплифицирована с помощью ОТ-ПЦР, с использованием обратной транскриптазы для формирования комплементарной цепочки ДНК, с последующей амплификацией с помощью полимеразной цепной реакцией с использованием праймеров, специфических для заявленных последовательностей ДНК. Альтернативно, образец мРНК отделяют с помощью гель-электрофореза, переносят на подходящий носитель, например, нитроцеллюлозу, нейлон и т.д., и затем тестируют фрагментом заявленной ДНК в качестве пробы. Могут также использоваться другие способы, такие как анализы сшивания олигонуклеотидов, гибридизация *in situ* и гибридизация ДНК-пробами, иммобилизованными на твердый чип. Обнаружение мРНК, гибридизующейся с заявленной последовательностью указывает на экспрессию гена в образце.

Белки

[076] Также обеспечиваются в соответствии с заявленным изобретением люциферазы, их гомологи, производные и мутанты, включая полноразмерные белки и их части или фрагменты.

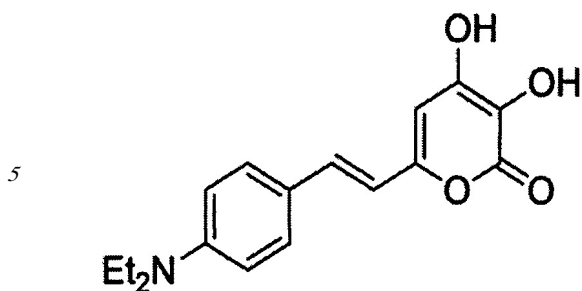
[077] Заявленные белки являются люциферазами, способными катализировать окисление люциферина в присутствии кислорода. Реакция окисления не зависит от присутствия в среде АТФ, NAD(P)H и других метаболитов. Заявленные белки отличаются от известных ранее люцифераз, так как окисляют 3-гидроксигиспидин, имеющий структуру:



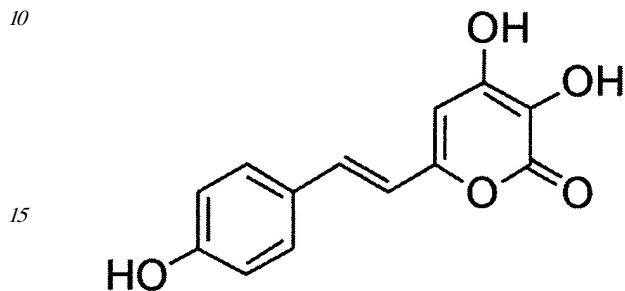
[078] Заявленные люциферазы могут катализировать окисление других химических соединений. Выявление таких соединений производится путем объединения выделенной люциферазы и указанного химического соединения в подходящих условиях и выявлении испускаемого в ходе реакции окисления света. Примерами соединений, способных выступать в качестве люциферинных для заявленных люцифераз, являются, например

[079] (E)-6-(4-(диэтиламино)стирил)-3,4-дигидрокси-2H-пиран-2-он, имеющий структуру:

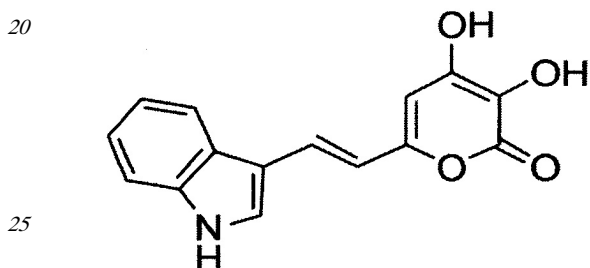
45



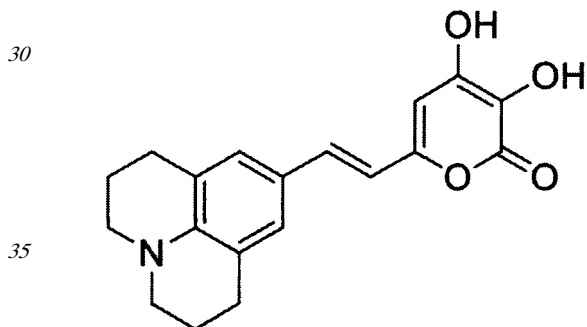
[080] (E)-3,4-дигидрокси-6-(4-гидроксистирил)-2H-пиран-2-он, имеющий структуру:



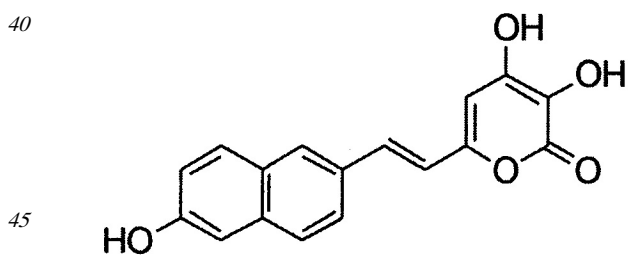
[081] (E)-6-(2-(1H-индол-3-ил)винил)-3,4-дигидрокси-2H-пиран-2-он, имеющий структуру:



[082] (E)-6-(2-(1,2,3,5,6,7-гексагидропиридо[3,2,1-ij]хинолин-9-ил)винил)-3,4-дигидрокси-2H-пиран-2-он, имеющий структуру:



[083] и (E)-3,4-дигидрокси-6-(2-(6-гидроксинафталин-2-ил)винил)-2H-пиран-2-он, имеющий структуру:



[084] Реакция окисления люциферинов люциферазой настоящего изобретения сопровождается выделением детектируемого света.

[085] В некоторых вариантах осуществления настоящего изобретения выделяемый

в процессе реакции свет может быть выявлен обычными методами (например, при визуальном осмотре, осмотре с помощью приборов ночного видения, спектрофотометрией, спектрофлуориметрией, с использованием фотографической регистрации изображения, с использованием специализированного оборудования для детекции люминесценции и флуоресценции, такого, как, например, IVIS Spectrum In Vivo Imaging System (Perkin Elmer) и т.п.). Регистрируемый свет может испускаться в диапазоне интенсивностей от одного фотона до легко заметного глазу света, например, с интенсивностью 1 кд, и яркого света с интенсивностью, например, 100 кд, и более.

[086] Испускаемый при окислении 3-гидроксигиспидина свет обладает находится в диапазоне от 400 до 700 нм, чаще в диапазоне от 450 до 650 нм, с максимумом эмиссии при 520-590 нм.

[087] Заявленные белки сохраняют активность при температурах ниже 50°C, чаще при температурах до 45°C, то есть они сохраняют активность при температурах 30-42°C и могут быть использованы в системах гетерологической экспрессии *in vitro* и *in vivo*.

[088] Заявленные белки обладают рН стабильностью в диапазоне от 4 до 10, чаще в диапазоне от 6.5 до 9.5. Оптимум рН стабильности заявленных белков лежит в диапазоне между 7.0 и 8.0, например, между 7.3-7.5.

[089] Белки, представляющие интерес, включают природные люциферазы из *Neonothopanus nambi* (SEQ ID NO: 2), *Armillaria gallica* (SEQ ID NO: 4), *Armillaria mellea* (SEQ ID NO: 6), *Armillaria ostoyae* (SEQ ID NO: 8), *Mycena chlorophos* (SEQ ID NO: 10), *Mycena citricolor* (SEQ ID NO: 12), *Omphalotus olearius* (SEQ ID NO: 14), и *Panellus stipticus* (SEQ ID NO: 16 и 18), их рекомбинантные и укороченные варианты, например, содержащие аминокислотные последовательности, показанные в SEQ ID NO: 20, 22, 24, 26, 28, 30, 32, 34 и детально описанные в экспериментальной секции, ниже.

[090] Также обеспечиваются люциферазы, по существу сходные вышеуказанным люциферазам и их функциональным фрагментам. Во многих вариантах осуществления настоящего изобретения представляющие интерес аминокислотные последовательности характеризуются значительной идентичностью по последовательности, которая составляет, например, не менее 40% идентичности, например, не менее 45% идентичности, или не менее 50% идентичности, или не менее 55% идентичности, или не менее 60% идентичности, или не менее 65% идентичности, или не менее 70% идентичности, или не менее 75% идентичности, например, не менее 80% идентичности, не менее 85% идентичности, не менее 90% идентичности (например, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 98% или 99% идентичности). В особенности это относится к последовательности аминокислот, которые обеспечивают функциональные участки белка, то есть к последовательности белка, расположенной после последовательности трансмембранного домена, входящего в состав природных люцифераз настоящего изобретения.

[091] Также обеспечиваются люциферазы, которые содержат характерную консенсусную последовательность, присущую природным люциферазам настоящего изобретения и показанную в SEQ ID NO: 35. Консенсусная последовательность создается с помощью множественного сравнения люцифераз настоящего изобретения путем выявления аминокислот, наиболее часто встречающихся в данном положении во всей совокупности родственных последовательностей.

[092] Белки, представляющие собой мутанты указанных выше белков, также рассматриваются в настоящем изобретении. Мутанты могут сохранять биологические свойства белков, из которого они были получены, или могут обладать биологическим

свойствами, которые отличаются от свойств белков дикого типа. Термин «биологическое свойство» белков согласно настоящему изобретению относится, без ограничения, к способности окислять различные люциферины; биохимическим свойствам, таким, как стабильность *in vivo* и/или *in vitro* (например, период полувыведения); скорость созревания; тенденция к агрегации или олигомеризации, а также другие подобные свойства. Мутации включают изменения одной или нескольких аминокислот, делеции или вставки одной или нескольких аминокислот; N-концевые усечения или расширения, C-концевые усечения или расширения и т.п.

[093] Мутанты могут быть получены с использованием стандартных методик молекулярной биологии, таких как подробно описанные выше в разделе «Молекулы нуклеиновой кислоты».

[094] Белки согласно настоящему изобретению находятся в изолированной форме, то есть это означает, что данный белок по существу свободен от присутствия других белков или других природных биологических молекул, таких как олигосахариды, нуклеиновые кислоты и их фрагменты и т.п., где термин «по существу свободен» в данном случае означает, что менее чем 70%, обычно менее чем 60% и чаще менее чем 50% указанной композиции, содержащей выделенный белок, представляет собой другую природную биологическую молекулу. В некоторых вариантах указанные белки присутствуют по существу в очищенной форме, где термин «по существу очищенная форма» обозначает чистоту, равную по меньшей мере 95%, обычно равную по меньшей мере 97% и чаще равную по меньшей мере 99%.

[095] В предпочтительном варианте целевые белки получают методом синтеза, например, путем экспрессии рекомбинантной нуклеиновой кислоты, кодирующей последовательность, которая кодирует представляющий интерес белок в подходящей клетке-хозяине, как было указано выше. Может использоваться любая удобная процедура очистки белка, где подходящие методики очистки белка описаны в соответствующем руководстве *Guide to Protein Purification*, (Deuthser ed.) (Academic Press, 1990). Например, может быть получен лизат из исходного источника и далее подвергнут очистке с использованием методики ВЭЖХ, вытеснительной хроматографии, геле-электрофореза, аффинной хроматографии и т.п.

[096] В настоящем изобретении также описываются белки слияния, включающие белок согласно настоящему изобретению или его функциональные фрагменты, слитые, например, с последовательностью субклеточной локализации (например, с сигналом ядерной локализации, сигналом локализации в пероксисомах, митохондриях, аппарате Гольджи и т.п.), с сигнальным пептидом или любым белком или полипептидом, представляющим интерес. Белки слияния могут включать, например, люциферазу настоящего изобретения и второй полипептид («партнер слияния»), оперативно слитые в рамке считывания с N-концом и/или C-концом люциферазы. Партнеры слияния включают, без ограничения, полипептиды, которые могут связываться с антителами, специфичными к партнеру слияния (например, эпитопными метками), с антителами или их связывающими фрагментами, с полипептидами, которые обеспечивают каталитическую функцию или индуцируют клеточный ответ, лигандами или рецепторами или их миметиками и т.п.

[097] В настоящем изобретении также рассматриваются антитела, которые специфически связываются с люциферазами согласно настоящему изобретению. Подходящие антитела могут быть получены с использованием методик, известных в данной области. Например, поликлональные антитела могут быть получены по методике, описанной в работе (Harlow and Lane *Antibodies: A Laboratory Manual*, (1988)

Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). Моноклональные антитела могут быть получены по методике, описанной в работе (Goding Monoclonal Antibodies: Principles and Practice: Production and Application of Monoclonal Antibodies in Cell Biology, Biochemistry and Immunology; 3rd edition, (1996) Academic Press). Химерные антитела, включающие гуманизированные антитела, а также одноцепочечные антитела и фрагменты антител, такие как Fv, F(ab')₂ и FAb, также представляют интерес.

Трансгенные организмы

[098] Нуклеиновые кислоты настоящего изобретения могут использоваться для создания трансгенных организмов или для сайт-специфических генных модификаций клеточных линий. Трансгенные клетки согласно настоящему изобретению включают одну или несколько нуклеиновых кислот, рассматриваемых в настоящем изобретении, которые присутствуют в качестве трансгена. Для целей настоящего изобретения может использоваться любая подходящая клетка-хозяин, включающая прокариотические (например, *Escherichia coli*, *Streptomyces* sp., *Bacillus subtilis*, *Lactobacillus acidophilus* и т.п.) или эукариотические клетки-хозяева, не являющиеся эмбриональными клетками человека. Трансгенные организмы согласно настоящему изобретению могут представлять прокариотические или эукариотические организмы, включающие бактерии, цианобактерии, грибы, растения и животные, в которые вводится одна или большее число клеток организма, содержащего гетерологичную нуклеиновую кислоту согласно настоящему изобретению, путем встраивания ее за счет манипуляции человеком, например, в рамках трансгенных методик, известных в данной области.

[099] Выделенная нуклеиновая кислота согласно настоящему изобретению может быть встроена в организм-хозяин с использованием методик, известных в данной области, например, путем инфекции, трансфекции, трансформации, доставки с помощью генной пушки или трансконъюгации. Методики переноса молекул нуклеиновой кислоты (например, ДНК) в такие организмы хорошо известны и описаны в стандартных руководствах таких, как Sambrook et al. (*Molecular Cloning: A Laboratory Manual*, 3rd Ed., (2001) Cold Spring Harbor Press, Cold Spring Harbor, NY).

[0100] В одном варианте осуществления настоящего изобретения трансгенный организм может представлять собой прокариотический организм. Способы трансформации прокариотических хозяйских клеток хорошо известны в данной области (см., например, Sambrook et al. (*Molecular Cloning: A Laboratory Manual*, 2nd Ed., (1989) Cold Spring Harbor Laboratory Press и Ausubel et al., *Current Protocols in Molecular Biology* (1995) John Wiley & Sons, Inc).

[0101] В другом варианте осуществления настоящего изобретения указанный трансгенный организм может представлять собой гриб, например, дрожжи. Дрожжи широко используются в качестве носителя для гетерологичной генной экспрессии (см., например, Goodey et al., *Yeast biotechnology*, D R Berry et al., eds, (1987) Allen and Unwin, London, pp. 401-429, и Kong et al., *Molecular and Cell Biology of Yeasts*, E.F. Walton and G.T. Yarronton, eds, Blackie, Glasgow (1989) pp. 107-133). Доступно несколько типов дрожжевых векторов, включающих интегрирующиеся векторы, которые требуют рекомбинации с геномом-хозяином для своего поддержания, а также автономно реплицирующиеся плазмидные векторы.

[0102] Другим организмом-хозяином является организм животного. Трансгенные животные могут быть получены с использованием трансгенных методик, известных в данной области и описанных в стандартных руководствах (таких как: Pinkert, *Transgenic Animal Technology: A Laboratory Handbook*, 2nd edition (2003) San Diego: Academic Press; Gersenstein and Vinterstein, *Manipulating the Mouse Embryo: A Laboratory Manual*, 3rd ed,

(2002) Nagy A. (Ed), Cold Spring Harbor Laboratory; Blau et al., Laboratory Animal Medicine, 2nd Ed., (2002) Fox J.G., Anderson L.C., Loew F.M., Quimby F.W. (Eds), American Medical Association, American Psychological Association; Gene Targeting: A Practical Approach by Alexandra L. Joyner (Ed.) Oxford University Press; 2nd edition (2000)). Например, трансгенные животные могут быть получены посредством гомологичной рекомбинации, в рамках которой меняется эндогенный локус. Альтернативно, конструкцию нуклеиновой кислоты интегрируют в случайном режиме в геном. Векторы для стабильной интеграции включают плазмиды, ретровирусы и другие вирусы животных, YAC и т.п.

[0103] Нуклеиновая кислота может быть введена в клетку, непосредственно или опосредованно, за счет введения в предшественник клетки, с помощью осторожной генетической манипуляции, такой как микроинъекция, или с помощью инфекции рекомбинантным вирусом или с использованием рекомбинантного вирусного вектора и т.п. Термин «генетическая манипуляция» не включает классический кроссбридинг или оплодотворение *in vitro*, но, скорее, обозначает введение рекомбинантной молекулы нуклеиновой кислоты. Указанная молекула нуклеиновой кислоты может быть интегрирована в хромосому или может представлять собой внехромосомно реплицирующуюся ДНК.

[0104] Конструкции ДНК для гомологичной рекомбинации включают по меньшей мере часть нуклеиновой кислоты согласно настоящему изобретению, где указанный ген содержит желательную одну или несколько генетических модификаций и включает участки гомологии к целевому локусу. В конструкции ДНК для проведения случайной интеграции нет необходимости включать участки гомологии для облегчения рекомбинации. Могут быть также включены маркеры позитивной и негативной селекции. Способы получения клеток, содержащих целевые генные модификации, посредством гомологической рекомбинации, известны в данной области. Различные методики трансфекции клеток млекопитающих описаны, например, в работе Keown et al., Meth. Enzymol. (1990) 185: 527-537).

[0105] В случае эмбриональных стволовых клеток (ES), может быть использована клеточная линия ES или эмбриональные клетки могут быть получены в свежем виде от организма-хозяина, такого как мышь, крыса, морская свинка и т.п. Такие клетки выращивают на соответствующем питательном слое для фибробластов или выращивают в присутствии фактора ингибирования лейкозных клеток (LIF). Трансформированные ES или эмбриональные клетки могут быть использованы для создания трансгенных животных с использованием соответствующей методики, известной в данной области.

[0106] Трансгенные животные могут представлять собой любых животных, отличных от человека, включая млекопитающее, отличное от человека (например, мышь, крыса), птицу или земноводное и т.п., и используются в функциональных исследованиях, при скрининге лекарственных препаратов и т.п.

[0107] Могут быть также получены трансгенные растения. Способы получения трансгенных растительных клеток и растений описаны в патентах №№5767367, 5750870, 5739409, 5689049, 5689045, 5674731, 5656466, 5633155, 5629470, 5595896, 5576198, 5538879 и 5484956, описание которых включено в настоящее изобретение в качестве ссылок. Способы получения трансгенных растений обобщены также в следующих обзорах: Plant Biochemistry and Molecular Biology (eds. Lea and Leegood, John Wiley & Sons (1993) pp. 275-295 и в Plant Biotechnology and Transgenic Plants (eds. Oksman-Caldentey and Barz) (2002) 719 p.

[0108] Для получения трансгенного организма-хозяина могут использоваться, например, эмбрионные эксплантаты, содержащие соматические клетки. После сбора

клеток или тканей интересующую экзогенную ДНК вводят в растительные клетки, и для такого введения доступно множество различных методик. При наличии выделенных протопластов возникает возможность введения с использованием ДНК-опосредованных протоколов генного переноса, включающих инкубирование протопластов с оголенной ДНК, такой как плаزمид, включающей экзогенную кодирующую последовательность, представляющую интерес, в присутствии поливалентных катионов (например, ПЭГ или PLO); или по методу электропорации протопластов в присутствии оголенной ДНК, включающей интересующую экзогенную последовательность. Далее отбирают протопласты, которые успешно захватили экзогенную ДНК, растягивают их до образования калюса и в итоге получают трансгенные растения посредством контакта с соответствующими количествами и взятыми в соответствующих соотношениях стимулирующими факторами, такими как ауксины и цитокинины.

[0109] Могут использоваться другие подходящие методы получения растений, такие как подход, основанный на использовании «генной пушки» или трансформация, опосредованная использованием *Agrobacterium*, известные специалистам в данной области.

Способы использования

[0110] Полипептиды и нуклеиновые кислоты настоящего изобретения находят применение в различных приложениях. Например, они используются в качестве реагентов для диагностики, контроля качества, экологического тестирования и других подобных исследований в биотехнологии и медицине. Кроме того, они находят применение в бытовых и ориентированных на развлечения приложениях, например, при получении биолюминесцентных трансгенных растений и животных, которые могут быть использованы в качестве источников света.

[0111] Например, нуклеиновые кислоты композиции могут быть применены для детекции различных внешних сигналов в среде. Например, для выявления сигналов в кишечнике млекопитающих. Для реализации применения в геном микроорганизма-хозяина вводят кассеты экспрессии, кодирующие сигнальный каскад, в котором триггером выступает анализируемый сигнал внешней среды, а репортером - экспрессия гена люциферазы. Экспрессия гена люциферазы может вызываться как непосредственно прохождением сигнального каскада, так и через индукцию генетических изменений в геноме клетки, с помощью методик, описанных, например, в [Kotula et al., Proc. Natl. Acad. Sci. USA, 2014, 111: 4838-4843]. Так, например, для создания штамма *E. coli*, регистрирующего присутствие во внешней среде тетрациклина, в геном *E. coli* вводят кодирующую последовательность люциферазы под контролем промотора гена *tetA*, а также также последовательность, кодирующую тетрациклин-чувствительный репрессор TetR из транспозона Tn10, под контролем конститутивного промотора. Геномное редактирование *E. coli* обеспечивается за счет методик, хорошо известных в области, например, описанных в [Sabri et al., Microbial Cell Factories, 2013, 12: 60]. Детекция присутствия в среде тетрациклина достигается сравнением интенсивности биолюминесценции при добавлении раствора 3-гидроксигиспидина к генетически модифицированным клеткам *E. coli*, прошедшим инкубацию в анализируемой среде, и к контрольным клеткам, не прошедшим такую инкубацию. Аналогичным образом могут быть созданы другие штаммы микроорганизмов, чувствительные к другим сигналам внешней среды, путем замены соответствующих элементов сигнального каскада на элементы, специфически чувствительные к желаемым сигналам. Независимость эмиссии света заявленными люциферазами от внутриклеточной доступности АТФ, NAD(P)H и других метаболитов обеспечивают стабильность

репортерного сигнала в различных физиологических состояниях клетки.

[0112] Также нуклеиновые кислоты настоящего изобретения находят применение в способах определения наличия в воде токсических веществ, например, производных гексахлорциклогексана и других. В микроорганизм-хозяин (например, в *E. coli*) вводят кассету экспрессии, обеспечивающую конститутивную продукцию люциферазы, оперативно слитой с сигналом быстрой белковой деградации. Для определения токсичности лиофилизированных бактерий инкубируют в течение 90 минут с анализируемой жидкостью в кювете для определения биолюминесценции, и в то же время другую аликвоту бактерий инкубируют с контрольной жидкостью. Критерием токсического действия является изменение интенсивности биолюминесценции в исследуемой пробе по сравнению с таковой для образца с раствором, не содержащим токсических веществ. По уменьшению интенсивности люминесценции судят о размере токсического эффекта.

[0113] Также молекулы настоящего изобретения могут быть использованы для определения концентрации 3-гидроксигиспидина в анализируемом растворе. Для этого, реагент, содержащий очищенную люциферазу в известной концентрации помещают в кювету и регистрируют фоновый сигнал биолюминесценции с помощью люминометра. Затем проводят контрольные калибровочные измерения: в кювету добавляют раствор 3-гидроксигиспидина с известной концентрацией, не прерывая регистрации интенсивности биолюминесценции. Разность между величиной сигнала интенсивности биолюминесценции в присутствии и в отсутствие 3-гидроксигиспидина является величиной, пропорциональной концентрации 3-гидроксигиспидина. Операцию повторяют для растворов 3-гидроксигиспидина в том диапазоне концентраций, в котором предполагается проводить измерения. На основании полученных данных строят калибровочный график зависимости интенсивности биолюминесценции от концентрации 3-гидроксигиспидина. Концентрацию 3-гидроксигиспидина в неизвестном образце определяют по интенсивности биолюминесценции, возникающей при добавлении раствора 3-гидроксигиспидина неизвестной концентрации к реагенту, используя калибровочный график, описанный выше. Заявленные белки и нуклеиновые кислоты - единственные существующие на сегодняшний день реагенты, способные специфически определять концентрацию 3-гидроксигиспидина в сложных смесях с другими соединениями.

[0114] Также нуклеиновые кислоты композиции могут быть использованы для создания трансгенных растений или животных, излучающих свет. Например, для создания трансгенного мха *Physcomitrella patens* необходимо интегрировать в его геном кодирующую последовательность люциферазы, оптимизированной для экспрессии в клетках организма-хозяина, под контролем конститутивного промотора, например, промотора гена *Akt1* риса, или индуцируемого промотора, например, термочувствительного промотора гена *Gmhsp17.3B* сои. Интеграция в геном может быть произведена с помощью любого из методов, известных в области, например, с помощью гомологичной рекомбинации, индуцируемой возникновением двуцепочечного разрыва в геномной ДНК мха с помощью нуклеазы *Cas9* *Streptococcus pyogenes* и гидовой РНК, комплементарной геномному локусу. В этом случае, интегрируемая генетическая конструкция должна быть фланкирована участками, гомологичными участкам генома в области двуцепочечного разрыва, длиной около 250 нуклеотидов. Для доставки ДНК в клетки мха может также применяться любая из известных методик, например, трансформация протопластов мха оголенной ДНК с помощью полиэтиленгликоля. После проведения трансформации, протопласты мха должны подращиваться в

питательной среде для регенерации клеточной стенки, а затем высеваться на твердую среду для регенерации гаметофитов. Генетически модифицированные растения могут служить источниками света в случае добавления 3-гидроксигиспидина или родственных молекул в среду или в почву, или автономно биоллюминесцировать, если биосинтез 3-гидроксигиспидина происходит в клетках хозяина. Заявленные люциферазы излучают преимущественно зеленый свет и, таким образом, оптимально подходят для эмиссии света сквозь ткани фотосинтезирующих растений в связи со сниженным поглощением таких тканей в зеленой области видимого спектра.

[0115] Нуклеиновые кислоты композиции также могут быть применены для визуализации клеточных белков, органелл, отдельных клеток или тканей. Например, для визуализации миграции раковых клеток в организме нуклеиновые последовательности люцифераз вводят в составе кассеты экспрессии или экспрессионного вектора в раковые клетки. Доставка может осуществляться любым из способов, известных в области. Например, нуклеиновые кислоты, кодирующие люциферазу, могут быть интегрированы в геном раковых клеток перед имплантацией в организм хозяина. В других вариантах, нуклеиновые кислоты водятся во все клетки организма, но быть помещены под контроль промоторов, активных только в раковых клетках. Так как 3-гидроксигиспидин способен проникать через клеточные мембраны, заявленные люциферазы можно визуализировать *in vivo* в живых организмах без фиксации и пермеабиллизации тканей. Для визуализации развития раковых опухолей и метастазирования в исследуемый организм должен быть введен раствор люциферина, а ткани визуализированы с помощью подходящего для детекции биоллюминесценции оборудования.

Наборы

[0116] В настоящем изобретении также обеспечиваются наборы для осуществления одного или нескольких направлений применения нуклеиновых кислот и белков настоящего изобретения.

[0117] Указанные наборы в типичном случае включают белок согласно настоящему изобретению или нуклеиновую кислоту, кодирующую данный белок, предпочтительно с элементами для обеспечения экспрессии целевого белка в клетке-хозяине, например, вектор или кассету экспрессии, содержащую нуклеиновую кислоту, кодирующую целевой белок. Указанные наборы также включают люциферин для проведения люминесцентной реакции с люциферазой настоящего изобретения.

[0118] В некоторых воплощениях люциферин выбран из группы, включающей 3-гидроксигиспидин, (E)-6-(4-(диэтиламино)стирил)-3,4-дигидрокси-2H-пиран-2-он, (E)-3,4-дигидрокси-6-(4-гидроксистирил)-2H-пиран-2-он, (E)-6-(2-(1H-индол-3-ил)винил)-3,4-дигидрокси-2H-пиран-2-он, (E)-6-(2-(1,2,3,5,6,7-гексагидропиридо[3,2,1-ij]хинолин-9-ил)винил)-3,4-дигидрокси-2H-пиран-2-он и (E)-3,4-дигидрокси-6-(2-(6-гидроксинафталин-2-ил)винил)-2H-пиран-2-он.

[0119] Указанные компоненты набора обычно присутствуют в подходящей для хранения среде, такой как буферный раствор, обычно в подходящем контейнере. Указанные компоненты могут присутствовать в наборе в лиофилизированной форме.

[0120] Также в наборе могут содержаться антитела, специфичные для предлагаемого белка. В некоторых вариантах указанный набор включает множество различных векторов, каждый из которых кодирует целевой белок, где указанные векторы разработаны для целей экспрессии в разных средах и/или в разных условиях, таких как, например, конститутивная экспрессия, где указанный вектор включает сильный промотор для экспрессии в клетках млекопитающих, или беспромоторный вектор с

сайтом множественного клонирования для обычного встраивания промотора и пролонгированной экспрессии и т.п.

[0121] Кроме указанных выше компонентов, целевые наборы также включают инструкции для осуществления предлагаемых методов. Инструкции могут присутствовать в целевых наборах в разных формах, где в каждом наборе может присутствовать одна или несколько таких форм.

[0122] Приведенные ниже примеры даны для целей иллюстрации и не должны быть интерпретированы в плане ограничения.

Примеры (экспериментальная часть)

10 Пример 1. Выделение кодирующей последовательности люциферазы из *Neonothopanus nambi*

[0123] Суммарная РНК из мицелия *Neonothopanus nambi* была выделена по методу, описанному в [Chomczynski and Sacchi, Anal. Biochem., 1987, 162, 156-159]. кДНК была амплифицирована с помощью SMART PCR cDNA Synthesis Kit (Clontech, США) согласно протоколу производителя. ПЦР-продукт был клонирован в вектор pGAPZ (Invitrogen, США) и трансформирован в компетентные клетки *E. coli* штамма XL1 Blue. Бактерии выращивали на чашках Петри в присутствии антибиотика зеоцина. Через 16 ч колонии были смыты с чашек, интенсивно перемешаны, и из них была выделена плазмидная ДНК с помощью набора для выделения плазмидной ДНК (Евроген, Россия). Выделенная плазмидная ДНК была лиnearизована по сайту рестрикции *AvrII* и использована для трансформации клеток *Pichia pastoris* GS115. Электропорация была проведена по методу с использованием ацетата лития и дитиотреитола, описанному в [Wu and Letchworth, Biotechniques, 2004, 36: 152-4]. Электропорированные клетки были рассеяны на чашки Петри со средой RDB medium, содержащей 1 М сорбитола, 2% (w/v) глюкозы, 1.34% (w/v) дрожжевую основу азотного агара (YNB), 0.005% (w/v) смеси аминокислот, 0.00004% (w/v) биотина и 2% (w/v) агара. Разнообразие итоговой библиотеки кДНК *Neonothopanus nambi* в дрожжах составило порядка одного миллиона клонов.

Полученные колонии опрыскивали раствором 3-гидроксигиспидина, детектируя присутствие в клетках люциферазы по появлению света. Испускаемый колониями свет детектировали с помощью IVIS Spectrum CT (PerkinElmer, США). Колонии, в которых детектировалось свечение в ответ на добавление 3-гидроксигиспидина отбирали и использовали для ПЦР в качестве матрицы со стандартными плазмидными праймерами. Продукты ПЦР секвенировали по методу Сенгера. Полученная последовательность нуклеиновой кислоты показана в SEQ ID NO: 01. Кодируемая ею аминокислотная последовательность показана в SEQ ID NO: 02.

[0124] Для анализа биолюминесценции экспрессированной люциферазы, колонии *Pichia pastoris*, в которых наблюдалось свечение, изолировали и растили в колбах объемом 750 мл, заполненных 250 мл среды YPD, в течение 72 часов при температуре 30°C и перемешивании со скоростью 200 об/мин. Затем клетки, экспрессирующие люциферазу, осаждали центрифугированием при 5000 g в течение 15 минут при 4°C. Полученный осадок ресуспендировали в фосфатном буфере с ионной силой 0.1 М и pH 7.4, содержащем 0.1% DDM, в течение двух часов при 4°C. Суспензию центрифугировали при 21000 g в течение 30 min при 4°C. Реакцию биолюминесценции активировали добавляя 750 мкл полученного супернатанта к 250 мкл раствора 3-гидроксигиспидина (25 мкМ) в 1% DDM. Биолюминесценцию детектировали с помощью спектрофлюориметра Varian Cary Eclipse Spectrofluorimeter. Полученный спектр в целом совпадал со спектром биолюминесценции мицелия *Neonothopanus nambi* (Фиг. 1).

Пример 2. Выделение кодирующих последовательностей люцифераз из различных

видов грибов

[0125] Из грибов *Armillaria gallica*, *Armillaria mellea*, *Armillaria ostoyae*, *Mycena chlorophos*, *Omphalotus olearius* и *Panellus stipticus* была выделена геномная ДНК и проведено полногеномное секвенирование по технологии Illumina HiSeq (Illumina, США) согласно рекомендациям производителя. Результаты секвенирования были использованы для предсказания аминокислотных последовательностей гипотетических белков и использованы для поиска гомологов люциферазы из *Noeonothopanus nambi*, выявленной как описано в Примере 1. Поиск гомологов осуществлялся с помощью алгоритма для анализа последовательности, описанного в Altschul et al., J. Mol. Biol., 215, pp. 403-10 (1990) и программного обеспечения, предоставляемого National Center for Biotechnology Information. Был также проведен поиск аминокислотных последовательностей в данных геномного секвенирования грибов в базе данных NCBI Genbank. При поиске использовали стандартные параметры поиска blastp. В результате были идентифицированы последовательности гипотетических белков - гомологов люциферазы из *Noeonothopanus nambi* в *Armillaria gallica*, *Armillaria mellea*, *Armillaria ostoyae*, *Mycena chlorophos*, *Omphalotus olearius*, *Panellus stipticus* и *Mycena citricolor*. Все выявленные люциферазы по существу сходны друг с другом. Степень идентичности аминокислотных последовательностей показан в Таблице 1.

20

25

30

35

40

45

Таблица 1. Степень идентичности аминокислотных последовательностей
люцифераз

	SEQ ID NO:2	SEQ ID NO:4	SEQ ID NO:6	SEQ ID NO:8	SEQ ID NO:10	SEQ ID NO:12	SEQ ID NO:14	SEQ ID NO:16	SEQ ID NO:18	
5										
10	SEQ ID NO:2	1.000	0.745	0.760	0.745	0.658	0.629	0.789	0.655	0.658
	SEQ ID NO:4		1.000	0.953	0.964	0.687	0.658	0.684	0.665	0.662
15	SEQ ID NO:6			1.000	0.945	0.684	0.651	0.687	0.669	0.665
	SEQ ID NO:8				1.000	0.687	0.651	0.684	0.669	0.665
20	SEQ ID NO:10					1.000	0.669	0.662	0.731	0.735
	SEQ ID NO:12						1.000	0.633	0.684	0.680
25	SEQ ID NO:14							1.000	0.662	0.665
	SEQ ID NO:16								1.000	0.996
30	SEQ ID NO:18									1.000

[0126] На основании множественного выравнивания на консервативные участки
35 нуклеотидных последовательностей были сконструированы вырожденные праймеры.
Структуры праймеров показаны в SEQ ID No: 36-43. Из плодовых тел и мицелия грибов
Armillaria gallica, *Armillaria mellea*, *Armillaria ostoyae*, *Mycena chlorophos*, *Mycena citricolor*,
Omphalotus olearius, and *Panellus stipticus* была выделена суммарная РНК и приготовлена
кДНК как описано в Примере 1. Полученная кДНК была использована для ПЦР с
40 вышеуказанными праймерами. ПЦР проводили в реакционной смеси объемом 50 мкл,
содержащей 1 мкл 20-кратно амплифицированной «ДНК, смесь полимераз Encyclo
(Евроген), однократный буфер, предоставленный производителем, 200 мкМ dNTPs и
0.5 мкМ одного из праймеров 1711-1714 и 0.5 мкМ одного из праймеров 1715-1718. 30
циклов ПЦР были проведены в амплификаторе PTC-200 MJ Research Thermal Cycler в
45 режиме расчета температуры блока (каждый цикл проводили в условиях 95°C - 10 сек;
55°C - 10 сек; 72°C - 1 мин). ПЦР-продукт был клонирован в вектор pTAdv (Clontech),
плазмидную ДНК выделяли и секвенировали по методу Сенгера с помощью
универсальных праймеров M13. Во всех случаях в кДНК исследованных грибов были

обнаружены последовательности люцифераз.

[0127] При амплификации «ДНК *Panellus stipticus* были обнаружены два варианта последовательности люциферазы, отличающиеся единичной аминокислотной заменой валина на изолейцин в положении 146 (SEQ ID No: 15-18).

5 Пример 3. Экспрессия люцифераз в клетках млекопитающих

[0128] Кодированные последовательности люцифераз, полученные как описано в примерах 1 и 2, оптимизировали для экспрессии в клетках млекопитающих (гуманизировали). Гуманизированные нуклеиновые кислоты получали методом олигонуклеотидного синтеза по стандартным методикам. Нуклеотидный состав гуманизированных нуклеиновых кислот показан в SEQ ID NOs: 44-51, аминокислотный состав соответствующих белков идентичен белкам дикого типа, показанным в SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16. Полученные нуклеиновые кислоты клонировали в вектор pmKate2-keratin (Евроген, Россия), используя сайты рестрикции NheI и NotI вместо последовательности, кодирующей белок слияния mKate2-keratin. Плазмидную ДНК
10 очищали и трансфицировали в клетки НЕК293NT и HeLa с помощью трансфекционного реагента FuGENE HD (Promega, США) по протоколу производителя. Спустя 24 часа после трансфекции в среду добавляли 3-гидроксигиспидин в концентрации 660 мкг/мл, а свечение клеток детектировали с помощью IVIS Spectrum CT (PerkinElmer). Все проанализированные образцы испускали свет с интенсивностью на порядок и более
15 превышающей сигнал, исходящий от нетрансфицированных контрольных клеток (Фиг. 2).

[0129] Нуклеотидную последовательность SEQ ID No: 51 в составе вектора подвергли сайт-направленному мутагенезу, внося замену валина на изолейцин в положении 146. Вектор, содержащий полученную последовательность трансфицировали в клетки
25 НЕК293NT. Спустя 24 часа после трансфекции, в среду добавляли 3-гидроксигиспидин в концентрации 660 мкг/мл, а свечение клеток детектировали с помощью IVIS Spectrum CT (PerkinElmer). Было показано, что этот вариант также способен к билюминесцентной реакции окисления 3-гидроксигиспидина.

Пример 4. Экспрессия люциферазы *Neonothopanus nambi* в клетках *Pichia pastoris*
30 [0130] ДНК, кодирующая люциферазу *Neonothopanus nambi*, была получена как описано в Примере 1. Ген люциферазы был амплифицирован с помощью ген-специфичных концевых праймеров и клонирован в экспрессионный вектор GAP-pPicZA (Invitrogen) и использованием сайтов рестрикции эндонуклеаз BstBI и Sall, в одну рамку считывания с последовательностью, кодирующей С-концевую гистидиновую метку. Полученная
35 генетическая конструкция была трансформирована в штамм *Pichia pink*, характеризующийся со сниженной активностью протеаз, с помощью электропорации. Электропорация осуществлялась по методу с использованием ацетата лития и дитиотреитола, описанному в [Wu and Letchworth, Biotechniques, 2004, 36: 152-4]. Электропорированные клетки рассеивали на чашки Петри, содержащие среду YPD (2%
40 (w/v) пептона, 1% (w/v) дрожжевого экстракта, 2% (w/v) глюкозы, 2% (w/v) агара) и антибиотик зеоцин в концентрации 100 мкг/мл.

[0131] Клоны *Pichia pink*, производящие люциферазу, были идентифицированы опрыскиванием колоний раствором 3-гидроксигиспидина и визуализацией билюминесценции (испускаемого света) с помощью IVIS Spectrum CT (PerkinElmer).
45 Были отобраны клоны, для которых интенсивность свечения была наибольшей. Клоны изолировали и растили в колбах объемом 750 мл, заполненных 250 мл среды YPD, в течение 72 часов при температуре 30°C и перемешивании со скоростью 200 об/мин. Клетки, экспрессирующие люциферазу, были осаждены центрифугированием при 5000

g в течение 15 минут при 4°C. Осадок затем был ресуспендирован в 100 мл лизирующего буфера (0.1 М фосфата натрия, 0.1 М KCl, 4 мМ EDTA, 2 мМ TCEP, 1 мМ PMSF, pH 7.4) и разрушены пропусканием через гомогенизатор высокого давления (APV2000, SPX) 20 раз при давлении в 600 бар. Лизат центрифугировали при 8000 g в течение 15 мин при температуре 4°C. Полученный пост-ядерный супернатант центрифугировали при 140000 g в течение 2 ч при температуре 4°C. Полученный осадок (микросомальная фракция) ресуспендировали в 10 мл воды. Образцы разрешали с помощью денатурирующего электрофореза по Лэммли в полиакриламидном геле (10-25%), окрашенного Coomassie Blue. и использовали для вестерн-блота с конъюгатом антител на гистидиновую метку и пероксидазы хрена, с хемилюминесцентной детекцией сигнала. (Фиг. 4). Вестерн-блот продемонстрировал специфическое окрашивание люциферазы, перемещающейся в геле в области белков с массой 28 кДа, что примерно соответствует ожидаемой молекулярной массе люциферазы *Neonothopanus nambi*. Также Вестерн-блот показал соосаждение рекомбинантной люциферазы с осадком, полученным центрифугированием при 140000 g.

Пример 5. Получение функциональных фрагментов люцифераз

[0132] Открытые рамки считывания, кодирующие укороченные фрагменты люциферазы из *Neonothopanus nambi*, получали с помощью олигонуклеотидного синтеза. Нуклеиновые кислоты, кодирующие фрагменты люциферазы с отрезанными N-концевыми 6, 9, 12, 15, 21, 25, 31, 33, 35, 37 и 40 аминокислотными остатками, оперативно сшивали с нуклеиновой кислотой, кодирующей сайт инициации транскрипции и гистидиновую метку и клонировали в вектор pET-23 с помощью эндонуклеаз рестрикции BamHI and HindIII. Вектор использовали для трансформации клеток *Escherichia coli* штамма BL21-CodonPlus (производного высокопроизводительного штамма Stratagene BL21-Gold). Клетки высевали на чашки Петри со средой (1% NaCl, 1% триптон, 0.5% дрожжевой экстракт, 1.5% агар, ампициллин 100 мкг/мл, хлорамфеникол 100 мкг/мл и вода), и инкубировали в течение ночи при 37°C. Колонии *Escherichia coli* затем опрыскивали раствором люциферина и визуализировали в IVIS Spectrum CT (PerkinElmer, США) для определения функциональности экспрессируемых фрагментов люцифераз. Было обнаружено, что фрагменты люцифераз с отрезанными N-концевыми 6, 9, 12, 15, 21, 25, 31, 33, 35, 37 аминокислотными остатками испускают свет при опрыскивании раствором люциферина.

[0133] Нуклеотидная и аминокислотная последовательности N-конца, с которым оперативно сшивали фрагменты люциферазы показаны в SEQ ID No: 52 и 53, соответственно.

[0134] Анализ аминокислотной последовательности люциферазы из *Neonothopanus nambi* с помощью программного обеспечения, приведенного в [<http://www.cbs.dtu.dk/services/ТМНММ/>] показал, что в состав первых 39 аминокислот входит трансмембранный домен (фиг. 5). На основании полученных данных был сделан вывод, что удаление последовательности, содержащей трансмембранный домен, не влияет на способность люцифераз грибов катализировать реакцию окисления 3-гидроксиаиспидина, сопровождающуюся испусканием света.

[0135] С помощью программного обеспечения, приведенного на сайте [<http://www.cbs.dtu.dk/services/ТМНММ/>] был проведен анализ люцифераз из *Armillaria gallica*, *Armillaria mellea*, *Armillaria ostoyae*, *Mycena chlorophos*, *Mycena citricolor*, *Omphalotus olearius* и *Panellus stipticus*, клонированных как описано в Примере 2. Во всех случаях в составе аминокислотной последовательности были выявлены N-концевые фрагменты, содержащие трансмембранные домены (фиг. 5). Нуклеотидные и аминокислотные

последовательности фрагментов люцифераз, полученные после удаления N-концевых последовательностей, содержащих трансмембранные домены, показаны в SEQ ID NO: 19-34.

[0136] Множественное выравнивание люцифераз показано на фиг. 5. Видно, что на С-конце белки содержат неконсервативный фрагмент длиной 8-11 аминокислот, который также может быть удален или заменен на другой С-конец без потери функциональности люцифераз. Также видно, что люциферазы содержат высокомолекулярный центральный участок, содержащий консенсусную последовательность, показанную в SEQ ID No: 35.

Пример 6. Использование люцифераз с другими люциферинами

[0137] Плазмиды, содержащие кодирующие последовательности грибных люцифераз для экспрессии в клетках млекопитающих, были получены как описано в примере 3 и использованы для трансфекции клеток НЕК293NT. Спустя 24 часа после трансфекции клетки смывали с чашек с помощью раствора Версена с добавлением 0.025% трипсина, меняли среду на фосфатно-солевой буфер с pH 8.0 центрифугированием, ресуспендировали клетки, лизировали ультразвуком, а в среду добавляли 3-гидроксигиспидин или один из его аналогов ((E)-6-(4-(диэтиламино)стирил)-3,4-дигидрокси-2Н-пиран-2-он, (E)-3,4-дигидрокси-6-(4-гидроксистирил)-2Н-пиран-2-он, (E)-6-(2-(1Н-индол-3-ил)винил)-3,4-дигидрокси-2Н-пиран-2-он, (E)-6-(2-(1,2,3,5,6,7-гексагидропиридо[3,2,1-ij]хинолин-9-ил)винил)-3,4-дигидрокси-2Н-пиран-2-он или (E)-3,4-дигидрокси-6-(2-(6-гидроксинафталин-2-ил)винил)-2Н-пиран-2-он) в концентрации 660 мкг/мл.

[0138] Спектры биолюминесценции детектировали с помощью спектрофлуориметра Varian Cary Eclipse. Все проанализированные образцы испускали свет со всеми проверенными люциферинами В зависимости от используемого люциферина наблюдалось смещение оптимума люминесценции: так при окислении (E)-3,4-дигидрокси-6-(2-(6-гидроксинафталин-2-ил)винил)-2Н-пиран-2-она наблюдался сдвиг в длинноволновую область со значительным испусканием фотонов с длиной волны более 580 нм, а при окислении (E)-6-(2-(1Н-индол-3-ил)винил)-3,4-дигидрокси-2Н-пиран-2-она - в коротковолновую область.

Пример 7. Получение рекомбинантных люцифераз

[0139] На 5'-конец нуклеиновых кислот, кодирующих функциональные фрагменты люцифераз (SEQ ID NOs: 19, 21, 23, 25, 27, 29, 31, 33), была оперативно присоединена последовательность SEQ ID No: 52 и полученные конструкции были клонированы в вектор рЕТ-23 с помощью эндонуклеаз рестрикции BamHI and HindIII. Вектор использовали для трансформации клеток Escherichia coli штамма BL21-CodonPlus (производного высокопроизводительного штамма Stratagene BL21-Gold). Клетки высевали на чашки Петри со средой LB, содержащей 1.5% агар, ампициллин 100 мкг/мл и хлорамфеникол 100 мкг/мл, и инкубировали в течение ночи при 37°C. Колонии Escherichia coli затем переносили в 4 мл жидкой среды LB с добавлением ампициллина и хлорамфеникола, инкубировали в течение ночи при покачивании при 37°C. 1 мл ночной культуры переносили в 100 мл среду Overnight Express Autoinduction medium (Novagen), в которую были предварительно добавлены ампициллин и хлорамфеникол. Культуру растили при 37°C в течение 2.5 часов до достижения ей оптической плотности 0.6 ОЕ при 600 нм, а затем растили на комнатной температуре в течение 16 часов. Затем клетки осаждали центрифугированием при 4500 об/мин в течение 20 минут в центрифуге Eppendorf 5810R, ресуспендировали в 35 мл буфера (50 мМ Tris HCl pH 8.0, 150 мМ NaCl) и разрушали ультразвуком. Клеточный лизат центрифугировали при 7500 об/мин в течение 15 минут, сохраняли осадок, в то время как супернатант дополнительно

подвергали центрифугированию при 35000 об/мин в течение одного часа для разделения микросомальной и растворимой фракций. Осадок растворяли в 10 мл буфера, содержащего мочевины (8 M urea, 50 mM Tris, pH 8.0) в течение ночи при комнатной температуре.

5 [0140] Наличие ожидаемого рекомбинантного продукта подтверждали с помощью электрофореза. Пример подобного анализа для фрагмента люциферазы *Neonothopanus pambí* показан на фиг. 6. Видно наличие полосы в области 28 kDa, примерно соответствующей ожидаемой молекулярной массе люциферазы *Neonothopanus pambí*.

Аликвоты выделенных рекомбинантных белков наносили на полиакриламидный
10 гель-элетрофорз, а также использовали для проверки функциональности и стабильности рекомбинантных белков. Выделенные рекомбинантные белки испускали свет при соединении с 3-гидроксигиспидином с максимумами эмиссии в области 520-535 нм, причем интенсивность испускаемого света была наибольшей у функционального
фрагмента люциферазы *Neonothopanus pambí*, и наименьшей - у функционального
15 фрагмента *Omphalotus olearius*. Рекомбинантные белки были активны в буферных растворах с pH в дапазоне 7-9 и демонстрировали максимальную интенсивность билюминесцентной реакции при pH 7.3-8. Пример графика зависимости интенсивности билюминесцентного свечения от pH раствора показан на фиг. 7.

[0141] Для анализа температурной стабильности люциферазы инкубировали при
20 различных температурах в течение 10 мин при pH 7.4. По окончании инкубации к белкам добавляли 3-гидроксигиспидин и анализировали интенсивность билюминесценции как описано выше. Люциферазы сохраняли более 10% максимальной активности после инкубации при температурах ниже 50°C, более 30% активности при температурах ниже 40°C, более 70% при температурах ниже 38°C и 100% активности при температурах
25 ниже 34°C включительно.

Пример 8. Использование люциферазы *Neonothopanus pambí* для мечения клеток

[0142] Вектор, содержащий люциферазу *Neonothopanus pambí* под контролем
цитомегаловирусного промотора, полученный как описано в Примере 3,
котрансфицировали с вектором pTurboFP635-N (Евроген, Россия), кодирующим красный
30 флуоресцентный белок, в клетки линии НЕК293NT. Трансфекцию проводили трансфекционным реагентом FuGENE HD (Promega) по протоколу, рекомендованному производителем. Спустя 24 часа после трансфекции в среду добавляли 3-гидроксигиспидин до конечной концентрации 660 мкг/мл, а свечение клеток анализировали с помощью микроскопа Leica DM6000 с объективом 20x. Клетки
35 визуализировали в проходящем свете, в канале для детекции зеленой флуоресценции и в канале для детекции красной флуоресценции (Фигура 6). Экспрессия люциферазы *Neonothopanus pambí* в клетках человека приводила к появлению отчетливого светового сигнала в зеленой области спектра. Никаких проявлений токсичности экспрессии люциферазы для клеток обнаружено не было.

40 Пример 9. Мечение белков с помощью люциферазы *Neonothopanus pambí*

[0143] Гуманизованную нуклеиновую кислоту, кодирующую люциферазу *Neonothopanus pambí*, полученную как описано в Примере 3, оперативно сшивали (клонировали в одну рамку считывания) с нуклеиновыми кислотами, кодирующими
цитоплазматический бета-актин и фибрилларин человека. Полученные конструкции
45 клонировали в вектор pmKate2-keratin, как описано в Примере 3. Спустя 24 часа после трансфекции, в среду культивирования был добавлен 3-гидроксигиспидин до конечной концентрации 660 мкг/мл. Регистрируемая при этом билюминесценция соответствовала паттернам внутриклеточной локализации, характерным для соответствующих клеточных

белков.

Пример 10. Мечение органелл с помощью люциферазы *Neonothopanus nambi*

[0144] Вариант последовательности люциферазы *Neonothopanus nambi*,

оптимизированный для экспрессии в клетках человека, полученный как описано в

5 Примере 3, оперативно сшивали считывания со следующими сигналами внутриклеточной локализации: сигналом локализации в митохондриях (MTS) из 7-ой субъединицы человеческой цитохром-оксидазы; сигналом, кодируемым N-концевыми 81 аминокислотами человеческой бета-1,4-галактозил-трансферазы [Watzel and Berger, *Nucleic Acids. Res.*, 1990, 18: 7174]; сигналом пероксисомальной локализации [Gould et al. *J. Biol. Chem.*, 1989, 108: 1657-1664; Gould et al. *EMBO J.*, 1990, 9: 85-90; Monosov et al., *J. Histo. Cytochem.*, 1996, 44: 581-589]; тремя копиями сигнала внутриядерной локализации (NLS) Т-антигена вируса SV40 [Kalderon et al., *Cell*, 1984, 39: 499-509; Lanford et al., *Cell*, 1986, 46: 575-582]. Трансфекция клеток HeLa Kyoto плазмидами, экспрессирующими химерные конструкции с люциферазой *Neonothopanus nambi* приводила к эффективному
15 перемещению химерных белков в соответствующие органеллы внутри клеток. Билюминесценция была детектирована через 24 часа после трансфекции при добавлению в среду 3-гидроксигипсидина в конечной концентрации 660 мкг/мл.

Пример 11. Мечение клеток внутри целого организма

[0145] Вектор, содержащий кодирующую последовательность люциферазы

20 *Neonothopanus nambi* под контролем цитомегаловирусного промотора, был получен как описано в Примере 3. Дополнительно была синтезирована гуманизированная нуклеотидная последовательность, кодирующая люциферазу светлячка *Photinus pyralis*, и клонирована в такой же вектор.

[0146] Полученные конструкции использовали для трансфекции клеток линии CT26

25 (клетки карциномы *Mus musculus*). Клетки, экспрессирующие люциферазу *Neonothopanus nambi*, были инъецированы подкожно в левую часть спины мыши, в то время, как клетки, экспрессирующие люциферазу *Photinus pyralis*, - инъецированы таким же образом в правую половину спины мыши. Спустя 10 минут после инъекции, смесь грибного люциферина (0.5 мг) и люциферина светлячка *Photinus pyralis* (0.5 мг) была введена
30 мышам внутривентрально. Билюминесценция мыши была затем визуализирована с помощью IVIS Spectrum CT (PerkinElmer). Опухоль, экспрессирующая люциферазу *Neonothopanus nambi*, демонстрировала интенсивность испускаемого света в 2040000 у.е., а опухоль, экспрессирующая люциферазу *Photinus pyralis* - 21000000 у.е. (Фиг. 9).

[0147] мРНК люциферазы *Neonothopanus nambi* была получена путем транскрипции

35 *in vitro* линеаризованного по сайту Acc65I вектора pCS2+, содержащего ген люциферазы, с помощью полимеразы SP6 из набора SP6 mMessage mMachine Kits (Ambion, США). мРНК далее очищали с помощью CleanRNA Standard Kit (Evrogen) и инъецировали в эмбрионы *Xenopus laevis* на двухклеточной стадии в оба бластомера, по 500 пг мРНК на бластомер. Для визуализации, на стадии ранней гаструлы (стадия 10.5) в эмбрионы
40 инъецировали раствор люциферина 660 мкг/мл в бластоцель. Свечение эмбриона после окрашивания родамином детектировали во время нейруляции (стадии 16-17) при помощи микроскопа Leica DM6000, в каналах микроскопа для детекции зеленой и красной флуоресценции, с использованием объектива 5x (Фигура 10). Билюминесцентный сигнал был выявлен в нервной ткани эмбриона.

45 Пример 12. Приготовление поликлональных антител

[0148] Кодирующая последовательность люциферазы *Neonothopanus nambi* с

удаленным трансмембранным доменом, показанная в SEQ ID No: 19 была синтетически получена в виде двуцепочечной ДНК и клонирована в экспрессионный вектор pQE-30

(Qiagen, Германия) таким образом, что полученный рекомбинантный белок содержал на N-конце гистидиновую метку. После экспрессии в *E. coli*, рекомбинантный белок был очищен с помощью металл-аффинной смолы TALON (Clontech) в денатурирующих условиях. Препарат очищенного белка, эмульсифицированный в адьюванте Фрейнда, был использован для четырех иммунизаций кроликов с месячными интервалами. Кровь кроликов отбирали на десятый или одиннадцатый дни после иммунизаций. Активность полученной поликлональной антисыворотки была продемонстрирована на рекомбинантном белке с помощью методов ELISA и Вестерн-иммуоблоттинга.

Пример 13. Получение трансгенных растений

[0149] Кодирующая последовательность люциферазы *Neonothopanus nambi* была оптимизирована для экспрессии как в клетках мха *Physcomitrella patens* (SEQ ID No: 54). Затем *in silico* была создана транскрипционная единица, содержащую промотор гена *akt1* риса, 5'-нетранслируемую область цитомегаловируса человека, кодирующую последовательность люциферазы, стоп-кодон, и последовательность терминатора из гена *osc* агробактерии *Agrobacterium tumefaciens*. Полученная последовательность была создана синтетически и клонирована с помощью методики Gibson assembly [Gibson et al., Nat Methods, 2009, 6: 343-5] в экспрессионный вектор pLand#1 между фрагментами ДНК, совпадающими с локусом геномной ДНК мха *Physcomitrella patens* между последовательностями высокоэкспрессируемых генов мха Pp3c16_6440V3.1 и Pp3c16_6460V3.1. Вектор pLand#1 также содержал последовательность гидовой РНК (sgRNA) для нуклеазы Cas9, комплементарную участку того же локуса ДНК. Препарат плазмидной ДНК был котрансформирован вместе с экспрессионным вектором, содержащим последовательность нуклеазы Cas9, в протопласты *Physcomitrella patens* согласно протоколу полиэтиленгликольной трансформации, описанному в [Cove et al., Cold Spring Harb Protoc, 2009, 2]. Протопласты затем инкубировали в среде BCD в течение двух дней в темноте при покачивании с интенсивностью 50 об/мин для регенерации клеточной стенки. Протопласты затем переносили на чашки Петри, содержащие агар и среду BCD и выращивали при 16-часовом освещении в течение недели.

Трансформированные колонии мха скринировали с внешних геномных праймеров с помощью ПЦР для определения успешности интеграции генетической конструкции в геном, переносили на свежие чашки Петри и выращивали в тех же условиях освещения в течение 30 дней.

[0150] Полученные гаметофиты мха вымачивали в среде BCD, содержащей 3-гидроксигиспидин в концентрации 660 мкг/мл, и анализировали с помощью IVIS Spectrum In Vivo Imaging System (Perkin Elmer). Все проанализированные трансгенные растения демонстрировали биолюминесценцию с интенсивностью, как минимум на два порядка превышающую интенсивность сигнала контрольных растений дикого типа, инкубированных в том же растворе с 3-гидроксигиспидином.

Пример 14. Получение трансгенных животных

[0151] Трансгенные рыбы *Danio rerio*, содержащие ген люциферазы *Neonothopanus nambi*, были созданы по методу, описанному в [Hisano et al., Sci Rep., 2015, 5: 8841]. Для создания трансгенных животных были синтетически созданы фрагменты ДНК, содержащие последовательности гидовой РНК и мРНК нуклеазы Cas9 под контролем промотора полимераз бактериофага T7. Полученные фрагменты использовались для транскрипции *in vitro* с помощью реагентов из набора MAXIscript T7 kit (Life Technologies, USA), а синтезированная РНК очищалась с помощью набора для выделения РНК (Евроген, Россия). Также была синтетически получена последовательность донорного вектора, содержащая ген люциферазы *Neonothopanus nambi*, фланкированный 50-

нуклеотидными последовательностями из гена *krttl1c19e* *Danio rerio*. Донорный вектор, мРНК нуклеазы Cas9 и гидовая РНК были растворены в инъекционном буфере (40 мМ HEPES (pH 7.4), 240 мМ KCl с добавлением 0.5% фенольного красного) и инъецированы в зародыши 1-2-клеточные зародыши *Danio rerio* в объеме около 1-2 нл. Из 70 зародышей, около 50 зародышей пережило инъекцию и демонстрировало нормальное развитие на четвертый день после оплодотворения.

[0152] Для регистрации билюминесцентного сигнала, раствор 3-гидроксигиспидина инъецировали внутривенно в личинки *Danio rerio* согласно методике, описанной в [Cosentino et al., J Vis Exp. 2010; (42): 2079]. Билюминесценцию регистрировали с помощью IVIS Spectrum In Vivo Imaging System (Perkin Elmer). После регистрации из личинок выделяли геномную ДНК для подтверждения интеграции гена люциферазы в геном. Все личинки с корректной интеграцией гена люциферазы *Neonothopanus nambi* в геном демонстрировали билюминесценцию с интенсивностью, как минимум на порядок превышающую интенсивность сигнала, исходящего от рыб дикого типа после инъекции раствора 3-гидроксигиспидина.

[0153] Все публикации, патентные заявки, процитированные в данном описании, включены в качестве ссылки в той мере, в какой каждая отдельная публикация или отдельная патентная заявка специфически и индивидуально соответствует включению в качестве ссылки. Цитирование каждой публикации приводится для целей иллюстрации текста и лучшего понимания настоящего изобретения, и не должно рассматриваться как допущение, что каждая такая публикация соответствует достигнутому уровню в данной области.

>SEQ ID NO:1

ATGCGCATTAACATTAGCCTCTCGTCTCTCTTCGAACGTCTCTCCAACTTAGCAG
 TCGCAGCATAGCGATTACATGTGGAGTTGTTCTCGCCTCCGCAATCGCCTTTCCCAT
 CATCCGCAGAGACTACCAGACTTTCCTAGAAAGTGGGACCCTCGTACGCTCCGCAGA
 ACTTTAGAGGATACATCATCGTCTGTGTCTCTCGCTATTCCGCCAAGAGCAGAAAG
 GGCTCGCCATCTATGATCGTCTTCCCGAGAAACGCAGGTGGTTGGCCGACCTTCCCT
 TTCGTGAAGGAACCAGACCCAGCATTACCAGCCATATCATTACAGCGACAGCGCACT
 CAACTGGTTCGATCAGGAGTTTGCCACCAGGGAGCTCATAGACAAGGTCATCCCTCG
 CGTGCAAGCACGACACACCGACAAAACGTTCCCTCAGCACATCAAAGTTCGAGTTTC
 ATGCGAAGGCCATATTTCTCTTGCCTTCTATCCCAATCAACGACCCTCTGAATATCC
 CTAGCCACGACACTGTCCGCCGAACGAAGCGCGAGATTGCACATATGCATGATTAT
 CATGATTGCACACTTCATCTTGCTCTCGCTGCGCAGGATGGAAAGGAGGTGCTGAA
 GAAAGGTTGGGGACAACGACATCCTTTGGCTGGTCCCTGGAGTTCCTGGTCCACCAA
 CGGAATGGACTTTTCTTTATGCGCCTCGCAACGAAGAAGAGGCTCGAGTAGTGGAG
 ATGATCGTTGAGGCTTCCATAGGGTATATGACGAACGATCCTGCAGGAAAGATTGT
 AGAAAACGCCAAG

>SEQ ID NO:2

MRINISLSSLFERLSKLSRSIAITCGVVLASAIAPFIIRRDYQTFLEVGPSYAPQNFRGY
 IIVCVLSLFRQEQKGLAIYDRLPEKRRWLADLPFREGTRPSITSHIIQRQRTQLVDQEFAT
 RELIDKVIPRVQARHTDKTFLSTSKFEFHAKAIFLLPSIPINDPLNIPSHDVTVRRTKREIAH
 MHDYHDCTHLALAAQDGKEVLKKGWQQRHPLAGPGVPGPPEWTFLYAPRNEEEAR
 VVEMIVEASIGYMTNDPAGKIVENAK

>SEQ ID NO:3

ATGTCCTTCATCGACAGCATGAACTTGACCTCGTCGGACACCTCTTTGGCATCA
 GGAATCGCGGCTTAGCCGCCGCTTGTTGTCTCTAGCAGTCGCCTCTACTATCGCCT
 TCCCTTACATTCGTAGGGACTACCAGACATTTTTATCTGGCGGTCCCTCTTACGCTCC

CCAGAATATCAGAGGATATTTTCATCGTCTGCGTTCTGGCCTTGTTCCGTCAGGAGCA
 AAAGGGCCTTGCGATATATGATCGCCTTCCCGAGAAGCGCAGGTGGCTGCCTGACT
 TGCCTCCTCGCAATGGCCCGCGGCCGATCACGACCAGCCATATAATCCAAAGACAG
 CGCAACCAGGGCGCCGGACCCCAAGTTCGCCCTCGAGGAACTCAAGGCCACGGTTAT
 5 TCCACGGGTGCAGGCTCGCCATACTGACCTCACCCATCTCAGCCTATCCAAATTCGA
 GTTCCATGCTGAAGCAATTTTCTGCTCCCCTCTGTACCCATCGATGATCCAAAAA
 TGTCCAAGTCACGACACGGTGCAGGACGAAAAGGGAGATCGCGCATATGCACG
 ACTACCATGACTTACGCTGCATCTTGCACTGGCCGCCAAGACGGGAAGGAAGTC
 GTGTGAAGGGATGGGGGCAGCGACACCCCTAGCAGGCCCTGGCGTTCCTGGTCC
 10 ACCTACGGAGTGGACATTTATTTATGCGCCACGTAACGAAGAGGAACTGGCAGTGG
 TGGAAATGATTATCGAGGCATCAATAGGCTATATGACCAATGACCCTGCTGGAGTA
 GTTATCGCA

>SEQ ID NO:4

MSFIDSMKLDLIVGHLFGIRNRGLAAACCALAVASTIAFPYIRRDYQTFLSGGPSYAPQ
 15 NIRGYFIVCVLALFRQEQKGLAIYDRLPEKRRWLPDLPPRNGPRPITTSIIQRQRNQAPD
 PKFALEELKATVIPRVQARHTDLTHLSLSKFEFHAEAIPLLPSVPIDDPKNVPSHDTVRR
 KREIAHMHDYHDFTLHLALAAQDGKEVVSKEGWQRHPLAGPGVPGPPTTEWTFIYAPR
 NEEELAVVEMIIASIGYMTNDPAGVIA

>SEQ ID NO:5

ATGTCCTTCTTCGACAGCGTGAAACTTGACCTCGTTCGGACGCCTCTTTGGCATCA
 GGAATCGCGGCTTAGCTGTTACTTGTGTGCTGTGGCAGTCGCCTCTATCATCGCGT
 TCCCTTACATTTCGTAGGGACTACCAGACATTTTATCTGGGGGTCCCTCCTACGCTCC
 CCAGAACATCAGAGGATACCTCATTGTCTGCGTCCTGGCCTTGTTCCGTCAGGAGCA
 AAAAGGCCTTGCGATATACGACCGCCTTCCCGAGAAGCGCAGGTGGCTACCTGACT
 25 TGCCTCCTCGCGATGGCCACGGCCATCACGACCAGCCATATAATCCAAAGACAG
 CGCAACCAGGGCGCCGGACCTCAAGTTCGCCCTCGAGGAACTCAAGGCCACGGTCAT
 TCCACGGGTGCAGGCTCGCCACACTGACCTCACCCATCTCAGCCTATCCAAGTTCGA
 GTTCCATGCTGAAGCAATCTTCTGCTCCCCTCTGTACCCATCGATGATCCAAAGAA
 TGTGCCAAGTCACGACACGGTGCAGGACGAAAGAGGGAAATTGCGCATATGCACG
 30 ACTACCATGACTACACGCTGCATCTTGCGTTGGCCGCCAAGACGGGAAGGAAGTC
 GTATCAAAGGGATGGGGGCAGCGACACCCGCTGGCAGGCCCTGGCGTTCCTGGTCC
 ACCGACGGAGTGGACGTTTATTTATGCGCCACGTAACGAAGAGGAGCTGGCAGTGG
 TGGAAATGATTATCGAGGCATCGATAGGCTATATGACCAATGACCCTGCAGGAAA
 ACTATCGCATAG

35 >SEQ ID NO:6

MSFFDSVKLDLIVGRLFGIRNRGLAVTCCAVAVASIIAFPYIRRDYQTFLSGGPSYAPQ
 NIRGYLIVCVLALFRQEQKGLAIYDRLPEKRRWLPDLPPRDGPRPITTSIIQRQRNQAPD
 LKFALEELKATVIPRVQARHTDLTHLSLSKFEFHAEAIPLLPSVPIDDPKNVPSHDTVRR
 KREIAHMHDYHDYTLHLALAAQDGKEVVSKEGWQRHPLAGPGVPGPPTTEWTFIYAPR
 40 NEEELAVVEMIIASIGYMTNDPAGKTIA

>SEQ ID NO:7

ATGTCCTTCATCGACAGCATGAAACTTGACTTCGTCGGACACCTCTTTGGCATCA
 GGAATCGCGGCTTAGCCACCGCTTGTGTGCTGTGGCAGTCGCTTCTGCCATCGCCT
 TCCCTTACATTTCGTAGGGACTACCAGACATTTATCTGGCGGTCCCTCTTACGCTCC
 45 CCAGAACATCAAAGGATATCTCATCGTCTGCGTCCTGGCCTTGTTCCGTCAGGAGCA
 AAAGGGCCTTGCGATATATGACCGCCTTCCCGAGAAGCGCAGGTGGCTACCTGACT
 TGCCTCCTCGCAATGGCCCGCGGCCATCACGACCAGCCATATAATCCAAAGACAG
 CGCAACCAGGGCGCCAGACTCCAAGTTCGCCCTCGAGGAACTCAAGGCTACGGTCAT

TCCACGGGTGCAGGCTCGCCACACTGACCTCACCCATCTCAGCCTATCCAAGTTCGA
 GTTCCATGCTGAAGCAATCTTCCTGCTCCCCCTCTGTACCCATCGATGATCCAAAAA
 TGTCCAAGTCATGACACGGTGCAGGACGAAGAGGGAGATCGCGCATATGCACG
 ACTACCATGACTTTACGTTGCATCTTGCCTGGCCGCCAAGACGGGAAGGAAGTC
 5 GTGGCGAAGGGATGGGGGCAGCGACACCCGCTGGCAGGCCCTGGCGTTCCTGGTCC
 ACCTACGGAGTGGACGTTTATTTATGCGCCACGTAACGAAGAGGAACTGGCAGTGG
 TGAAATGATTATCGAGGCATCAATAGGCTATATGACCAATGACCCTGCTGGAACA
 GTTATCGTATAG

>SEQ ID NO:8

10 MSFIDSMKLDVFGHLFGIRNRGLATACCAV AVASAIAPYIRRDYQTFLSGGPSYAPQ
 NIKGYLIVCVLALFRQEQKGLAIYDRLPEKRRWLPDLPPRNGPRPITTSIIQRQRNQAPD
 SKFALEELKATVIPRVQARHTDLTHLSLSKFEFHAEAIFFLLPSVPIDDPKNVPSHDTVRR
 KREIAHMHDYHDFTLHLALAAQDGKEVVAKGWGQRHPLAGPGVPGPPTWTFIYAPR
 NEEELAVVEMIIIEASIGYMTNDPAGTVIV

15 >SEQ ID NO:9

ATGGTCCAACCTCACCCGAACCTCCGGATTCATCGCCGCTGCGGCCATTGTTGCTG
 CCATCGCCTTCCCCTTCATTCGTCGAGACTACCAGACGTTCCCTTCGTGGTGGGCCGT
 CCTATGCCCCACAGAACATCCGCGGCTATATCATCGTCCTGGTTCTGTCCCTCTTCCG
 CGGCGAGGAGAAGGGTCTTGCAATCTACGAGCCCCTTCTTGAGAAGCGCACATGGC
 20 TGCCGAGCTTCCGCGGCGCGCGGGAGACCGGCCCAAGACGACGAGCCACATCATC
 CAACGGCAGCTCGACCAGTACCCCGACCCGGACTTTGTCTCAAAGCCCTGAAAGC
 GACGGTCATCCCGCGTGTCCAAGCCCGGCACACAGACAAGACTCACCTCGCGCTGT
 CCAAGTTCGAGTTCATGCTGAGGCCATCTTCGTGCGCCCGGAAATCGCCATCGACG
 ACCCGAAGCATATCCCCAGCCACGACACGGTGCGACGGACGAAGCGCGAGATTGCG
 25 CACATGCACGACTATCACGACTGCACGCTGCATTTGGCGCTAGCGGCGCAGGACGC
 GAAGCAGGTGCTGCAGAAGGGCTGGGGCCAGCGCCATCCGCTGGCAGGGCCTGGG
 ATGCCCCGGGCCGCCACGGAGTGGACGTTCTTGTATGCCCCGAGGACCGAGGAGGA
 AGTGAAGGTTGTGGAGACCATTGTGCGAGGCCTCTATCGCGTACATGACGAACGCGG
 AGAAGCCGGTCGAGCTGGTGCGAG

30 >SEQ ID NO:10

MVQLTRTSGFIAAAAIVAAIAFPFIRRDYQTFRLRGGPSYAPQNIRGYIIVLVLSLFRGEE
 KGLAIYEPLPEKRTWLPELPRRAGDRPKTTSHIIQRQLDQYDPDFVLKALKATVIPRVQ
 ARHTDKTHLALSLSKFEFHAEAIFFVRPEIAIDDPKHIPSHDTVRRRKREIAHMHDYHDCTLH
 LALAAQDAKQVLQKGWGQRHPLAGPGMPGPPTWTFIYAPRTEEEVKVETIVEASIA
 35 YMTNAEKPVELVQ

>SEQ ID NO:11

ATGGCTTATCAGCTCACTTGGATTCAGACTCTCGTGCTGGGTGCCCTTGTGGCAA
 TGGCAGTAGCGTTCCTTCATCAAGAAAGACTACGAGACGTTCCCTGAAGGGCGGC
 CCCTCCTATGCGCCCCAAAACGTTTCGCGGATACATCATCGTGCTCGTGCTCGCGCTC
 40 TTCCGCCAAGAGCAGCTCGGGCTGGAGATCTACGACCGCATGCCCGAGAAACGTGCG
 CTGGCTCGCGAATCTCCCTCAGCGCGAGGGCCCCCGCCCCAAGACCACAAGTCACA
 TCATCCAGCGGCAGCTCAGCCAGCACACGGACCCCGCATTCGGCGCCGCGTACCTC
 AAAGACACCGTCATTCCGCGCGTCCAGGCGCGGCACGCAGCCAACACGCACATCGC
 GCGCTCGACGTTTCGAGTTCACGCCCGCCGCGATCTTCCTGAACGCGGACGTGCCGCT
 45 GCCCGAGGGCCTGCCCGCAAGCGAGACGGTGCGGCGGACCAAGGGCGAGATCGCG
 CACATGCACGACTACCACGACTTCACGCTGCACCTCGCGCTCGCAGCAGCGGATGG
 GAAGGAGGTGGTTCGCAAGGGCTGGGGGCAGCGCCATCCGCTGGCGGGACCCGGC
 GTGCCGGGTCCGCCGAACGAGTGGACCTTTGTGTATGCGCCGAGGAATGAAGAGGA

GATGGGCGTGGTCGAGCAGATCGTAGAGGCGGCGATTGGGTACATGTCTGAACGTGC
CTGCGCTGGAA

>SEQ ID NO:12

MAYQLTWIQTIVLGLVAMAVAFPFIKDYETFLKGGPSYAPQNVRGYIIVLVLALF
5 RQEQLGLEIYDRMPEKRRWLANLPQREGPRPKTTSHIIQRQLSQHTDPAFGAAYLKDTV
IPRVQARHAANTHARSTFEFHAAAIFLNADVPLPEGLPASETVRRTKGEIAHMHDYHDF
TLHLALAAADGKEVVGKGWQQRHPLAGPGVPGPPNEWTFVYAPRNEEEMGVVEQIVE
AAIGYMSNPALE

>SEQ ID NO:13

10 ATGCTCCCAGCTTTCATCTACAAACCAAGGCTAGTGATCACTTGTGTATTCGTTCT
GGCCTCCGCACTCGCATTTCCTTCATACGCAAAGATTACCAGACTTTCCTGGAGGT
GGGACCCTCGTACGCCCCGAGAACCTCCAAGGATACATCATCGTCTGTGTACTCTC
TCTGTTCCGGCAAGAACAGAAAGACGTAGCGATTTATGATCGCCTTCCTGAGAAAA
GGAGGTGGTTAGGAGACCTCCCGTTTCGCGAGGGGCCAAGACCGAGTATCACTAGC
15 CATATCATCCAGCGACAGCGCACCCAATTGGCTGACGCCGAGTTCGCTACCAAAGA
GCTGATAGGCAAATCATCCCTCGCGTCCAAGCCCGACACACCAACACAACATTCC
TCAGCACATCTAAATTCGAATTCACGCCAGGCCATCTTCCTTTTGCCCTCTATCCC
AATCAACGACCCTCAAACATTCCAAGCCACGATACCGTTCGTGCGACGAAACGCG
AGATCGCGCATATGCATGATTATCACGACTGTACGTTGCATCTCGCACTTGCTGCTC
20 AAGATGGGAAGGAGGTTTTAGAGAAAGGATGGGGTCAGCGACATCCTCTTGCTGGA
CCTGGTGTTCCTGGCCCGCCGACGGAGTGGACGTTTTCTTTATGCACCGCGCAGCGAA
GAGGAGGTTCCGGGTTGTGGAGATGATTGTTGAGGCATCAGTTGTGTATATGACGAA
TGATCCTGCGGATAAAATCGTAGAAGCTACTGTGCAGGGTACTGAAGAATAG

>SEQ ID NO:14

25 MLPAFIYKPRIVITCVFVLASALAFPFIKDYQTFLEVGPSYAPQNLQGYIIVCVLSLF
RQEQLDVAIYDRLPEKRRWLGDLPFREGPRPSITSHIIQRQRTQLADAEFATKELIGKIIPR
VQARHTNTTFLSTSKFEFHAQAIFLLPSIPINDPQNIPSHDTVRRTKREIAHMHDYHDCTL
HLALAAQDGKEVLEKGWQQRHPLAGPGVPGPPTEWTFLYAPRSEEEVVRVEMIVEASV
VYMTNDPADKIVEATVQGTEE

30 >SEQ ID NO:15

ATGAACATCAACCTGAAAGCTCTGATCGGAGTCTGTGCCGTGCTCATCACCGCTG
CAGTGTCCCCTTCGTTTCGTAAAGACTATCACACCTTTCCTTGAAGGTGGACCATCCT
ACGCGCCGCGAATTTGCAAGGCTATATCATCGTGTGGTGCTCTCACTCTTTCGAG
GGGAGGAGACGGGATTGAAATATACGACCGCTTGCCCGAAAAACGCCGCTGGCTC
35 GAGGAGCTGCCTGTTTCGCGAAGGCCCGCGCCCAAAGACAACCAGCCACATCATTCA
GAGACAGTTGAATCAGCACGTTGACCCGGACTTCGGAATGA ACTCTTTGAAAGGCT
CCGTCATCCGGCGCCTTCAATCCCGCCACCAGGACATAACTCAACTCGCACTCTCGA
AATTCGAATCCACGCCGAGGCCATATTTCTGCGCCCCGATGTCGCGATCAACGATC
CCAAACACGTCCCAGCCACGACACGGTGCGCCGCACAAAGCGCGAGATAGCTCAC
40 ATGCACGACTACCATGATTACACGTGTCATTTGGCGCTCGCAGCGCAGGATGGGAA
GCAAGTGATTGCAAAGGGTGGGGCCAGAGACATCCGCTCGCCGACCGGGCATGC
CGGGGCCCGCCGACGGAGTGGACATTTTGTATGCGCCGAGGAATGAGGCGGAGGTT
CAAGTGTTGGAGACGATTATCGAAGCGTCAATCGGGTACATGTCTGAACGCACCAGC
CTTGGGTGGGAGCGAG

45 >SEQ ID NO:16

MNINLKALIGVCAVLITAAVFPFVRKDYHTFLEGGPSYAPQNLQGYIIVLVLVSLFRGE
ETGLEIYDRLPEKRRWLEELPVREGPRPKTTSHIIQRQLNQHVDPDFGMNSLKGSVIRRL
QSRHQDITQLALSKFEFHAEAIFLRPDVAINDPKHVPSHDTVRRTKREIAHMHDYHDYT

CHLALAAQDGKQVIAKGWGQRHPLAGPGMPGPPEWTFlyAPRNEAEVQVLETHIEAS
IGYMSNAPALGGSE

>SEQ ID NO:17

ATGAACATCAACCTGAAAGCTCTGATCGGAGTCTGTGCCGTGCTCATCACCGCTG
5 CAGTGTCCCCTTCGTTTCGTAAGACTATCACACCTTTCTTGAAGGTGGACCATCCT
ACGCGCCGCAGAATTTGCAAGGCTATATCATCGTGTGGTGCTCTCACTCTTTTCGAG
GGGAGGAGACGGGATTGAAATATACGACCGCTTGCCCGAAAAACGCCGCTGGCTC
GAGGAGCTGCCTGTTTCGCGAAGGCCCGCGCCCAAAGACAACCAGCCACATCATTCA
GAGACAGTTGAATCAGCACGTTGACCCGGACTTCGGAATGAACTCTTTGAAAGGCT
10 CCGTCATCCGGCGCCTTCAATCCCGCCACCAGGACATAACTCAACTCGCACTCTCGA
AATTCGAATCCACGCCGAGGCCATATTTCTGCGCCCCGATATCGCGATCAACGATC
CCAAACACGTCCCGAGCCACGACACGGTGCGCCGCACAAAGCGCGAGATAGCTCAC
ATGCACGACTACCATGATTACACGTGTCATTTGGCGCTCGCAGCGCAGGATGGGAA
GCAAGTGATTGCAAAGGGTGGGGCCAGAGACATCCGCTCGCCGACCGGGCATGC
15 CGGGGCCCGCCGACGGAGTGGACATTTTTGTATGCGCCGAGGAATGAGGCGGAGGTT
CAAGTGTTGGAGACGATTATCGAAGCGTCAATCGGGTACATGTTCGAACGCACCAGC
CTTGGGTGGGAGCGAG

>SEQ ID NO:18

MNINLKALIGVCAVLITAAVFPFVRKDYHTFLEGGPSYAPQNLQGYIIVLVLSLFRGE
20 ETGLEIYDRLPEKRRWLEELPVREGPRPKTTSHIIQRQLNQHVDPDFGMNSLKGSVIRRL
QSRHQDITQLALSKFEFHAEAIFLRPDIAINDPKHVPSHDTVRRTKREIAHMHDYHDYTC
HLALAAQDGKQVIAKGWGQRHPLAGPGMPGPPEWTFlyAPRNEAEVQVLETHIEASIG
YMSNAPALGGSE

>SEQ ID NO: 19

ATCATCCGCAGAGACTACCAGACTTTCCTAGAAAGTGGGACCCTCGTACGCTCCGC
25 AGAACTTTAGAGGATACATCATCGTCTGTGTCCTCTCGCTATTCCGCCAAGAGCAGA
AAGGGCTCGCCATCTATGATCGTCTTCCCGAGAAACGCAGGTGGTTGGCCGACCTTC
CCTTTCGTGAAGGAACCAGACCCAGCATTACCAGCCATATCATTACGCGACAGCGC
ACTCAACTGGTCGATCAGGAGTTTGCCACCAGGGAGCTCATAGACAAGGTCATCCC
30 TCGCGTGCAAGCACGACACACCGACAAAACGTTCCCTCAGCACATCAAAGTTCGAGT
TTCATGCGAAGGCCATATTTCTCTTGCCTTCTATCCCAATCAACGACCCTCTGAATAT
CCCTAGCCACGACACTGTCCGCCGAACGAAGCGCGAGATTGCACATATGCATGATT
ATCATGATTGCACACTTCATCTTGCTCTCGCTGCGCAGGATGGAAAGGAGGTGCTGA
AGAAAGGTTGGGGACAACGACATCCTTTGGCTGGTCCTGGAGTTCCTGGTCCACCA
35 ACGGAATGGACTTTTCTTTATGCGCCTCGCAACGAAGAAGAGGCTCGAGTAGTGGA
GATGATCGTTGAGGCTTCCATAGGGTATATGACGAACGATCCTGCAGGAAAGATTG
TAGAAAACGCCAAG

>SEQ ID NO:20

IIRRDYQTFLEVGPSYAPQNFGRGYIIVCVLSLFRQEKGLAIYDRLPEKRRWLADLPFR
40 EGTRPSITSHIIQRQRTQLVDQEFATRELIDKVIPRVQARHTDKTFLSTSKFEFHAKAIFLL
PSIPINDPLNIPSHDTVRRTKREIAHMHDYHDCTLHLALAAQDGKEVLKKGWQRHPLA
PGVPGPPEWTFlyAPRNEEEARVEMIVEASIGYMTNDPAGKIVENAK

>SEQ ID NO:21

TACATTCGTAGGGACTACCAGACATTTTTATCTGGCGGTCCTCTTACGCTCCCCA
45 GAATATCAGAGGATATTTATCATCGTCTGCGTTCTGGCCTTGTTCCGTCAGGAGCAAAA
GGGCCTTGCGATATATGATCGCCTTCCCGAGAAGCGCAGGTGGCTGCCTGACTTGCC
TCCTCGCAATGGCCC GCGGCCGATCACGACCAGCCATATAATCCAAAGACAGCGCA
ACCAGGCGCCGACCCCAAGTTCGCCCTCGAGGAACTCAAGGCCACGGTTATTCCA

CGGGTGCAGGCTCGCCATACTGACCTCACCCATCTCAGCCTATCCAAATTCGAGTTC
 CATGCTGAAGCAATTTTCTGCTCCCCTCTGTACCCATCGATGATCCAAAAATGTT
 CCAAGTCACGACACGGTGCGCAGGACGAAAAGGGAGATCGCGCATATGCACGACT
 ACCATGACTTCACGCTGCATCTTGCACTGGCCGCCCAAGACGGGAAGGAAGTCGTG
 5 TCGAAGGGATGGGGGCAGCGACACCCCCTAGCAGGCCCTGGCGTTCCTGGTCCACC
 TACGGAGTGGACATTTATTTATGCGCCACGTAACGAAGAGGAACTGGCAGTGGTGG
 AAATGATTATCGAGGCATCAATAGGCTATATGACCAATGACCCTGCTGGAGTAGTT
 ATCGCA

>SEQ ID NO:22

10 YIRRDYQTFLSGGPSYAPQNIRGYFIVCVLALFRQEQKGLAIYDRLPEKRRWLPDLPP
 RNGPRPITTSIIQRQRNQAPDPKFALEELKATVIPRVQARHTDLTHLSLSKFEFHAEAI
 LLPSVPIDDPKNVPSHDTVRRTKREIAHMHDYHDFTLHLALAAQDGKEVVSKGWGQRH
 PLAGPGVPGPTEWTFIYAPRNEEELAVVEMIIIEASIGYMTNDPAGVVIA

>SEQ ID NO:23

15 TACATTCGTAGGGACTACCAGACATTTTTATCTGGGGGTCCCTCCTACGCTCCCCA
 GAACATCAGAGGATACCTCATTGTCTGCGTCCTGGCCTTGTTCCGTCAGGAGCAAAA
 AGGCCTTGCGATATACGACCGCCTTCCCGAGAAGCGCAGGTGGCTACCTGACTTGC
 CTCCTCGCGATGGCCCACGGCCCATCACGACCAGCCATATAATCCAAGACAGCGC
 AACCAGGCGCCGGACCTCAAGTTCGCCCTCGAGGAACTCAAGGCCACGGTCATTCC
 20 ACGGGTGCAGGCTCGCCACACTGACCTCACCCATCTCAGCCTATCCAAGTTCGAGTT
 CCATGCTGAAGCAATCTTCCTGCTCCCCTCTGTACCCATCGATGATCCAAAGAATGT
 GCCAAGTCACGACACGGTGCGCAGGACGAAGAGGGAAATTGCGCATATGCACGAC
 TACCATGACTACACGCTGCATCTTGCGTTGGCCGCCCAAGACGGGAAGGAAGTCGT
 ATCAAAGGGATGGGGGCAGCGACACCCGCTGGCAGGCCCTGGCGTTCCTGGTCCAC
 25 CGACGGAGTGGACGTTTATTTATGCGCCACGTAACGAAGAGGAGCTGGCAGTGGTG
 GAAATGATTATCGAGGCATCGATAGGCTATATGACCAATGACCCTGCAGGAAAAAC
 TATCGCA

>SEQ ID NO:24

30 >YIRRDYQTFLSGGPSYAPQNIRGYLIVCVLALFRQEQKGLAIYDRLPEKRRWLPDLP
 PRDGP RPITTSIIQRQRNQAPDLKFALEELKATVIPRVQARHTDLTHLSLSKFEFHAEAI
 FLLPSVPIDDPKNVPSHDTVRRTKREIAHMHDYHDYTLHLALAAQDGKEVVSKGWGQR
 HPLAGPGVPGPTEWTFIYAPRNEEELAVVEMIIIEASIGYMTNDPAGKTIA

>SEQ ID NO:25

ATGTCCTTCATCGACAGCATGAACTTGACTTCGTCGGACACCTCTTTGGCATCA
 35 GGAATCGCGGCTTAGCCACCGCTTGTTGTGCTGTGGCAGTCGCTTCTGCCATCGCCT
 TCCCTTACATTCGTAGGGACTACCAGACATTCTTATCTGGCGGTCCCTCTTACGCTCC
 CCAGAACATCAAAGGATATCTCATCGTCTGCGTCCTGGCCTTGTTCCGTCAGGAGCA
 AAAGGGCCTTGCGATATATGACCGCCTTCCCGAGAAGCGCAGGTGGCTACCTGACT
 TGCCTCCTCGCAATGGCCCGCGGCCCATCACGACCAGCCATATAATCCAAGACAG
 40 CGCAACCAGGCGCCAGACTCCAAGTTCGCCCTCGAGGAACTCAAGGCTACGGTCAT
 TCCACGGGTGCAGGCTCGCCACACTGACCTCACCCATCTCAGCCTATCCAAGTTCGA
 GTTCCAAGTCATGACACGGTGCGCAGGACGAAGAGGGAGATCGCGCATATGCACG
 ACTACCATGACTTTACGTTGCATCTTGCACTGGCCGCCCAAGACGGGAAGGAAGTC
 45 GTGGCGAAGGGATGGGGGCAGCGACACCCGCTGGCAGGCCCTGGCGTTCCTGGTCC
 ACCTACGGAGTGGACGTTTATTTATGCGCCACGTAACGAAGAGGAACTGGCAGTGG
 TGAAATGATTATCGAGGCATCAATAGGCTATATGACCAATGACCCTGCTGGAACA
 GTTATCGTA

>SEQ ID NO:26

MSFIDSMKLDVFGHLFGIRNRGLATACCAVAVASAIAPYIRRDYQTFLSGGPSYAPQ
 NIKGYLIVCVLALFRQEQKGLAIYDRLPEKRRWLPDLPPRNGPRPITTSIIIQRQRNQAPD
 SKFALEELKATVIPRVQARHTDLTHLSLSKFEFHAEAIFFLLPSVPIDDPKNVPSHDTVRR
 5 KREIAHMHDYHDFTLHLALAAQDGKEVVAKGWGQRHPLAGPGVPGPPTEWTFIYAPR
 NEEELAVVEMIIIEASIGYMTNDPAGTVIV

>SEQ ID NO:27

TTCATTTCGTCGAGACTACCAGACGTTCTTCGTGGTGGGCGTCCTATGCCCCAC
 AGAACATCCGCGGCTATATCATCGTCCTGGTTCTGTCCCTCTTCCGCGGCGAGGAGA
 10 AGGGTCTTGCAATCTACGAGCCCCCTTCTGAGAAGCGCACATGGCTGCCGGAGCTTC
 CGCGGCGCGCGGGAGACCGGCCCAAGACGACGAGCCACATCATCCAACGGCAGCT
 CGACCAGTACCCCGACCCGGACTTTGTCTTCAAAGCCCTGAAAGCGACGGTCATCC
 CGCGTGTCCAAGCCCGGCACACAGACAAGACTCACCTCGCGCTGTCCAAGTTCGAG
 TTCCATGCTGAGGCCATCTTCGTGCGCCCGGAAATCGCCATCGACGACCCGAAGCA
 15 TATCCCCAGCCACGACACGGTTCGACGGACGAAGCGCGAGATTGCGCACATGCACG
 ACTATCACGACTGCACGCTGCATTTGGCGCTAGCGGGCGCAGGACGCGAAGCAGGTG
 CTGCAGAAGGGCTGGGGCCAGCGCCATCCGCTGGCAGGGCCTGGGATGCCCGGGCC
 GCCCACGGAGTGGACGTTCTTGTATGCCCGAGGACCGAGGAGGAAGTGAAGGTTG
 TGGAGACCATTGTCGAGGCCTCTATCGCGTACATGACGAACGCGGAGAAGCCGGTC
 20 GAGCTGGTGCAG

>SEQ ID NO:28

FIRRDYQTFLRGGPSYAPQNIRGYIIVLVLSLFRGEEKGLAIYEPLPEKRTWLPELPRRA
 GDRPKTTSHIIQRQLDQYPDPDFVLKALKATVIPRVQARHTDKTHLALSFEFHAEAIFFV
 RPEIAIDDPKHIPSHDTVRRRKREIAHMHDYHDCTLHLALAAQDAKQVLQKGWGQRHP
 25 LAGPGMPGPPEWTFYAPRTEEEVKVETIVEASIAIYMTNAEKPVVIVQ

>SEQ ID NO:29

TTCATCAAGAAAGACTACGAGACGTTCTGAAGGGCGGCCCTCCTATGCGCCCC
 AAAACGTTTCGCGGATACATCATCGTGCTCGTGCTCGCGCTCTTCCGCCAAGAGCAGC
 TCGGGCTGGAGATCTACGACCCGCATGCCCGAGAAACGTCGCTGGCTCGCGAATCTC
 30 CCTCAGCGCGAGGGCCCCCGCCCCAAGACCACAAGTCACATCATCCAGCGGCAGCT
 CAGCCAGCACACGGACCCCGCATTCGGCGCCGCGTACCTCAAAGACACCGTCATTC
 CGCGCGTCCAGGCGCGGCACGCAGCCAACACGCACATCGCGCGCTCGACGTTTCGAG
 TTCCACGCCCGCGGATCTTCTGAACGCGGACGTGCCGCTGCCCGAGGGCCTGCC
 GCAAGCGAGACGGTTCGCGCGGACCAAGGGCGAGATCGCGCACATGCACGACTACC
 35 ACGACTTCACGCTGCACCTCGCGCTCGCAGCAGCGGATGGGAAGGAGGTGGTCGGC
 AAGGGCTGGGGGCAGCGCCATCCGCTGGCGGGACCCGGCGTGCCGGGTCCGCCGAA
 CGAGTGGACCTTTGTGTATGCGCCGAGGAATGAAGAGGAGATGGGCGTGGTTCGAGC
 AGATCGTAGAGGCGGCGATTGGGTACATGTCGAACGTGCCTGCGCTGGAA

>SEQ ID NO:30

FIKKDYETFLKGGPSYAPQNVIRGYIIVLVLALFRQEQQLGLEIYDRMPEKRRWLANLPQ
 REGPRPKTTSHIIQRQLSQHTDPAFGAAYLKDTVIPRVQARHAANTHIAIRSTFEFHAAAI
 FLNADVPLPEGLPASETVRRTKGEIAHMHDYHDFTLHLALAAADGKEVVGKGWGQRH
 PLAGPGVPGPPNEWTFVYAPRNEEEMGVVEQIVEAAIGYMSNPALE

>SEQ ID NO:31

TTCATACGCAAAGATTACCAGACTTTCCTGGAGGTGGGACCCTCGTACGCCCCGC
 AGAACCTCCAAGGATACATCATCGTCTGTGTACTCTCTCTGTTCCGGCAAGAACAGA
 AAGACGTAGCGATTTATGATCGCCTTCTGAGAAAAGGAGGTGGTTAGGAGACCTC
 45 CCGTTTCGCGAGGGGCCAAGACCGAGTATCACTAGCCATATCATCCAGCGACAGCG

CACCCAATTGGCTGACGCCGAGTTCGCTACCAAAGAGCTGATAGGCAAATCATCC
CTCGCGTCCAAGCCCCGACACACCAACACAACATTCCTCAGCACATCTAAATTCGAA
TTCCACGCCAGGCCATCTTCTTTTGGCCCTCTATCCCAATCAACGACCCTCAAAC
ATTCCAAGCCACGATACCGTTCGTGCGACGAAACGCGAGATCGCGCATATGCATGA
5 TTATCACGACTGTACGTTGCATCTCGCACTTGCTGCTCAAGATGGGAAGGAGGTTTT
AGAGAAAGGATGGGGTCAGCGACATCCTCTTGCTGGACCTGGTGTTCCTGGCCCCG
CGACGGAGTGGACGTTTCTTTATGCACCGCGCAGCGAAGAGGAGGTTTCGGGTTGTG
GAGATGATTGTTGAGGCATCAGTTGTGTATATGACGAATGATCCTGCGGATAAAAT
CGTAGAAGCTACTGTGCAGGGTACTGAAGAA
10 >SEQ ID NO:32
FIRKDYQTFLEVGPSYAPQNLQGYIIVCVLSLFRQEQKDVAIYDRLPEKRRWLGDLPF
REGPRPSITSHIIQRQRTLADAEFATKELIGKIIPRVQARHTNTTFLSTSKFEFHAQAIPLL
PSIPINDPQNIPSHDTVRRTKREIAHMHDYHDCTLHLALAAQDGKEVLEKGGWQRHPLA
GPGVPGPTEWTFLYAPRSEEEVRVEMIVEASVVYMTNDPADKIVEATVQGTEE
15 >SEQ ID NO:33
TTCGTTTCGTAAAGACTATCACACCTTCTTGAAGGTGGACCATCCTACGCGCCGC
AGAATTTGCAAGGCTATATCATCGTGTTGGTGCTCTCACTCTTTCGAGGGGAGGAGA
CGGGATTGAAATATACGACCGCTTGCCCGAAAAACGCCGCTGGCTCGAGGAGCTG
CCTGTTTCGCGAAGGCCCGCGCCCAAAGACAACCAGCCACATCATTACAGAGACAGTT
20 GAATCAGCACGTTGACCCGGACTTCGGAATGAACTCTTTGAAAGGCTCCGTCATCCG
GCGCCTTCAATCCCGCCACCAGGACATAACTCAACTCGCACTCTCGAAATTCGAATT
CCACGCCGAGGCCATATTTCTGCGCCCCGATGTCGCGATCAACGATCCCAAACACG
TCCCGAGCCACGACACGGTGCGCCGACAAAGCGCGAGATAGCTCACATGCACGAC
TACCATGATTACACGTGTCATTTGGCGCTCGCAGCGCAGGATGGGAAGCAAGTGAT
25 TGAAAAGGGTGGGGCCAGAGACATCCGCTCGCCGGACCGGGCATGCCGGGGCCG
CCGACGGAGTGGACATTTTTGTATGCGCCGAGGAATGAGGCGGAGGTTCAAGTGTT
GGAGACGATTATCGAAGCGTCAATCGGGTACATGTGGAACGCACCAGCCTTGGGTG
GGAGCGAG
>SEQ ID NO:34
30 FVRKDYHTFLEGGPSYAPQNLQGYIIVLVLVSLFRGEETGLEIYDRLPEKRRWLEELPV
REGPRPKTTSHIIQRQLNQHVDPDFGMNSLKGSVIRRLQSRHQDITQLALSKFEFHAEAIPL
LRPDVAINDPKHVP SHDTVRRTKREIAHMHDYHDYTCHLALAAQDGKQVIKGGWQR
HPLAGPGMPGPTEWTFLYAPRNEAEVQVLETIIEASIGYMSNAPALGGSE
>SEQ ID NO: 35
35 DYXTFLXXGPSYAPQNXXGYXIVXVXLXFRXEXXXIYXXXPEKRXWLXXLPXRXG
XRPXTSHIIQRQXXQXXDXFXXXLXXXIXRXQXRHXXXTXXXSXFEFHAXAIFX
XXXXXXXXPXXXPXXXTVRRTKXEIAHMHDYHDXTXHLALAAADXKVXXKGGWQRH
PLAGPGXPGPPXEWTFXYAPRXEXXVXEXIXEAXXXYMXN
>SEQ ID NO:36
40 ATCGCNTTYCCNTWCATHCCNARRGAYTAC SARACNTT
>SEQ ID NO:37
ATCGCNTTYCCNTWCATHCCNARRGAYTAC SARAC
>SEQ ID NO:38
ATCGCNTTYCCNTWCATHCCNARRGAYTACSA
45 >SEQ ID NO:39
ATCGCNTTYCCNTWCATHCCNARRGAYTA
>SEQ ID NO:40
NGCAGGGNCNTTSGTCATATASCCNATNGANGCYTCNAYNAT

>SEQ ID NO:41

NGCAGGGNCNTTSGTCATATASCCNATNGANGCYTC

>SEQ ID NO:42

NGCAGGGNCNTTSGTCATATASCCNATNGANGC

5 >SEQ ID NO:43

NGCAGGGNCNTTSGTCATATASCCNATNGA

>SEQ ID NO:44

ATGCGCATTAAACATCTCCCTTTCATCTCTTTTTCGAGCGATTGAGCAAACCTGAGTTC
CAGGAGTATTGCAATCACTTGTGGGGTTGTCCCTCGCGAGCGCCATCGCATTCCCCAT
10 CATCCGGAGAGATTATCAGACGTTTCTTGAGGTGGGCCCTAGCTATGCACCACAGA
ACTTCCGAGGATATATCATCGTGTGTGTACTGTCAGTGTAGGCAAGAACAAAAGG
GATTGGCTATCTATGATAGGTTGCCTGAGAAACGGCGGTGGCTCGCTGATCTCCCAT
TTAGAGAGGGGACACGACCGAGCATCACTTCACACATCATACAAAGACAGCGAACG
CAGCTCGTTGACCAAGAGTTCGCAACTAGGGAACCTGATTGATAAGGTGATACCCAG
15 AGTACAGGCGCGACACACCGATAAACTTTTCTTTCCACCTCTAAATTCGAGTTCCA
TGCCAAAGCTATTTTCTTGTGCTTCCATACCGATTAATGATCCTCTGAATATTCCA
TCCCACGACACAGTTCGACGGACGAAACGCGAAATTGCGCACATGCACGACTATCA
CGATTGCACTTTGCACCTGGCACTGGCTGCTCAAGACGGAAAAGAAGTTCTGAAAA
AGGGTTGGGGGCAAAGACATCCGCTGGCGGGACCCGGTGTACCTGGGCCCGCCTACG
20 GAATGGACATTTTGTACGCACCGAGGAACGAAGAGGAGGCCAGGGTTCGTTGAGAT
GATTGTTGAGGCTAGTATTGGGTACATGACGAATGATCCGGCTGGTAAAATTGTTGA
AAATGCAAAG

>SEQ ID NO:45

ATGAGCTTTATAGACTCAATGAAACTTGATCTTGTAGGACACCTCTTCGGCATAAC
25 GCAACAGAGGACTCGCAGCAGCCTGTTGCGCCCTTGACAGTAGCTTCCACGATAGCT
TTCCCTTATATCCGCCGGGATTACCAAACGTTCCCTTAGCGGTGGCCCTTCTTACGCG
CCACAGAACATACGAGGCTATTTTATCGTGTGTGTGCTCGCTCTGTTTCGCCAAGAA
CAGAAAGGGCTGGCGATTTATGACCGACTTCCTGAGAAACGAAGGTGGCTTCCTGA
CCTTCCCTCCCCGGAATGGCCCTCGCCCGATTACGACCTCTCATATTATACAAAGACA
30 GAGGAACCAAGCGCCAGATCCGAAATTTGCCCTGGAGGAGCTCAAGGCCACTGTGA
TTCCACGAGTCCAAGCCCGCCATACGGATCTTACACATCTCAGTCTTAGCAAATTTG
AGTTTCATGCCGAGGCAATTTTCTTTTGCCGTCCGTTCCGATAGACGACCCCAAAA
ATGTACCCTCACACGATACGGTCAGGCGCACAAAGAGGGAGATCGCCCATATGCAT
GACTATCACGACTTTACCCTGCATCTCGCTCTGGCGGCACAAGACGGCAAGGAGGT
35 AGTCAGCAAAGGCTGGGGCCAAAGGCATCCGCTCGCGGGCCAGGGGTTCGGGGGC
CACCTACAGAGTGGACTTTCATCTATGCGCCGCGAAACGAGGAAGAACTCGCAGTA
GTTGAGATGATCATAGAGGCATCAATTGGCTACATGACGAACGATCCTGCCGGAGT
CGTTATAGCC

>SEQ ID NO: 46

40 ATGAGTTTTTTTACTCTGTCAAACCTTGATCTGGTGGGACGGCTCTTCGGCATCCG
AAACAGGGGACTTGCGGTACTTGCTGCGCGGTAGCGGTCCCTCTATTATCGCCTT
CCCATACATTCGCCGCGACTATCAGACCTTCCCTGTCAGGCGGACCTTCTTATGCACC
GCAGAAATAAGGGGCTATCTTATTGTTTGTGTACTGGCTCTCTTTTCGGCAAGAACAA
AAAAGGTCTGGCTATCTACGATAGACTTCCCTGAGAAACGACGCTGGCTGCCAGATC
45 TTCCCCCCCCGCGACGGCCCAAGGCCATAACAACAAGTCATATTATACAGAGACAA
CGGAACCAGGCCCCAGACCTCAAGTTTGCCCTCGAGGAGCTGAAAGCGACTGTTAT
ACCTAGAGTACAGGCGAGACATACAGACTTGACCCATTTGAGCCTGAGCAAGTTTG
AATTCCATGCCGAAGCAATCTTTCTTCTGCCTTCCGTACCCATTGACGATCCAAAAA

ACGTCCCCTCACATGATACTGTAAGACGCACGAAGCGAGAGATCGCACACATGCAC
 GACTACCATGACTACACTCTGCACCTTGCCTTGGCAGCTCAGGATGGCAAGGAAGT
 TGTATCCAAGGGGTGGGGCCAAAGACACCCCTTGGCTGGCCCGGGAGTGCCTGGAC
 CCCCAGCCGAATGGACGTTTATTTACGCACCAAGGAACGAGGAAGAATTGGCCGTC
 5 GTTGAAATGATTATTGAAGCGTCCATAGGATATATGACCAACGATCCGGCCGGAAA
 GACAATCGCG

>SEQ ID NO: 47

ATGTCTTTTATTGACAGTATGAAGCTTGATTTCTGATAGGTCACCTTGTTTGGTATACG
 AAACCGAGGTCTCGCTACTGCGTGCTGCGCGGTGGCAGTAGCGAGTGCTATAGCAT
 10 TCCCGTATATCCGCCGAGATTACCAGACATTTCTTAGCGGTGGACCTAGCTATGCAC
 CCCAGAATATTAAGGGTATTTGATCGTTTGTGTGCTTGCCTGTTTCGCCAGGAGC
 AGAAGGGTCTTGCTATTTACGACCGACTGCCTGAGAAAAGGCGATGGTTGCCTGAC
 CTTCCACCCAGAAACGGTCCCAGGCCCATCACGACCTCACACATAATCCAACGGCA
 GAGGAACCAGGCTCCAGATTCAAAGTTTGCCTTGAGGAACTCAAGGCTACTGTAA
 15 TACCTAGAGTCCAGGCCAGACATACGGATCTCACCCATCTCAGCCTCAGTAAGTTTG
 AGTTTCATGCGGAAGCTATTTTCTGCTCCCGAGCGTACCCATTGATGACCCCAAGA
 ATGTCCCCTCACATGATACGGTACGACGCACGAAGCGGGAGATAGCTCACATGCAC
 GACTACCATGATTTCACTCTGCATCTCGCCTTGGCAGCGCAGGATGGCAAGGAGGT
 AGTAGCGAAAGGATGGGGGCAGAGACACCCGCTGGCAGGCCCTGGGGTTCCCGGC
 20 CCTCCGACCGAATGGACATTTATCTATGCGCCCCGCAACGAAGAGGAGCTCGCAGT
 TGTTGAGATGATCATTGAAGCATCCATTGGGTACATGACAAATGACCCAGCCGGTA
 CAGTCATTGTT

>SEQ ID NO: 48

ATGGTACAGCTCACGCGAACCTCAGGGTTTATTGCAGCGGCTGCGATTGTTGCCG
 25 CCATCGCGTTCCTCGTTTATTTCGACGGGACTATCAGACTTTTCTGCGAGGGGGGCCCT
 CATATGCCCCACAGAATATTCGCGGCTATATCATTGTCCTGGTCTTGTTCCCTCTTCAG
 AGGAGAAGAAAAGGGTCTGGCTATCTATGAGCCGCTTCCAGAGAAACGGACGTGGT
 TGCCAGAGTTGCCTAGACGGGCGGGCGATAGACCCAAAACCTACCAGCCATATAATA
 CAACGACAACCTCGACCAATATCCCGACCCGGACTTCGTACTTAAAGCATTGAAAGC
 30 AACTGTAATACCGAGAGTTCAGGCAAGGCACACTGATAAAACCCATCTTGCACTTT
 CCAAGTTTGAATTCCACGCCGAGGCAATCTTCGTCCGACCAGAGATAGCCATAGAC
 GATCCGAAACATATCCCAAGTCATGACACAGTCAGACGAACGAAACGGGAGATAG
 CCCATATGCATGATTATCATGACTGCACCCTGCACCTTGCCTTGCGGCACAGGACG
 CAAAACAGGTCTTGCAAAAAGGATGGGGACAGAGACACCCCTTGCCTGGACCTGGC
 35 ATGCCGGGTCCGCCTACTGAATGGACATTCCTCTACGCCCCGAGAAGTGGAGGGA
 GGTGAAAGTGGTTGAAACCATAGTGGAGGCGAGTATAGCTTACATGACCAATGCAG
 AAAAACAGTAGAACTCGTACAA

>SEQ ID NO: 49

ATGGCTTATCAGTTGACATGGATTCAGACTCTTGTATTGGGTGCTCTCGTGGCTAT
 40 GGCTGTTGCATTTCCATTCATCAAAAAGATTATGAGACCTTCTCAAGGGCGGCC
 TTCATACGCCCCCAGAATGTCCGCGGTTATATTATCGTACTCGTCTCGCGCTGTTT
 AGACAAGAACAAATTGGGGCTTGAGATATATGACCGCATGCCCGAAAACGCAGGTG
 GCTCGCTAACCTCCCTCAACGGGAAGGTCCGCGACCGAAAACCACTAGCCATATCA
 TTCAGCGCCAGCTCTCTCAGCATACCGACCCCGCCTTTGGAGCGGCTTATCTGAAAG
 45 ACACCGTAATCCCGCGAGTGCAGGCGCGGCATGCAGCGAACACTCACATAGCGCGA
 AGTACGTTTCGAATTTACGCAGCGGCTATCTTTCTCAACGCAGACGTGCCGTTGCC
 GAAGGACTGCCAGCTAGCGAAACAGTGAGACGAACCAAAGGAGAAATCGCGCACA
 TGCACGATTACCACGATTTTACGCTGCACCTGGCACTTGCGGCAGCGGACGGGAAA

GAGGTGGTCGGCAAAGGATGGGGACAGAGACATCCACTCGCTGGCCCTGGGGTGCC
TGGCCCTCCGAATGAGTGGACATTCGTGTACGCTCCCAGGAACGAGGAAGAGATGG
GCGTAGTAGAGCAGATTGTAGAGGCTGCAATCGGTTACATGTCAAATGTCCCTGCTC
TCGAG

5 >SEQ ID NO: 50

ATGCTCCCGGCTTTCATATATAAACCGAGATTGGTCATTACGTGCGTTTTTGTCT
CGTAGTGCATTGGCTTTTCCCTTCATAAGAAAGGACTATCAGACTTTCCTCGAAGT
CGGGCCGTCTTATGCTCCTCAAACCTTCAAGGGTATATCATAGTCTGCGTATTGTC
ACTCTTCAGACAAGAGCAAAAAGATGTAGCCATATACGATAGACTCCCCGAGAAAA
10 GAAGGTGGCTTGGTGACCTGCCGTTACAGAGAAGGGCCGCGCCCGTCTATTACCTCCC
ACATTATCCAACGCCAAAGAACGCAACTTGCTGATGCAGAGTTCGCAACAAAGGAG
TTGATTGGGAAGATCATTCCGAGGGTGCAGGCAAGGCATACAAATACCACTTTCTT
GTCAACGAGTAAATTTGAGTTTCACGCCAGGCAATCTTTCTGCTTCCCTTCAATTCCC
ATAAACGATCCACAAAATATCCCAAGTCACGATACGGTACGGCGCACGAAGCGCGA
15 GATCGCCACATGCATGATTATCATGATTGCACGCTTCACTTGGCTCTCGCCGCGCA
GGACGGGAAAGAGGTGCTTGAGAAGGGGTGGGGTACGCGGCATCCACTCGCCGGG
CCAGGTGTTCCCGGTCCCCCAACCGAATGGACATTTCTCTATGCGCCCCGCTCTGAA
GAAGAGGTACGGGTCGTGGAAATGATAGTCGAGGCTAGCGTGGTTTATATGACGAA
TGATCCTGCGGACAAGATTGTAGAGGCCACCGTCCAGGGCACCGAAGAG

20 >SEQ ID NO: 51

ATGAATATAAATCTGAAAGCACTTATTGGTGTCTGTGCGGTACTCATCACCGCTG
CTGTCTTTCCTTTTGTACGGAAAGACTATCACACCTTTTGGAGGGCGGCCCTTCCTA
TGCTCCACAGAACTTGCAGGGCTACATAATAGTCCTCGTCTTGTCACTCTTCAGGGG
AGAGGAGACTGGTTTGGAATTTATGATCGCTTGCCAGAAAAGCGGCGATGGCTGG
25 AAGAACTGCCAGTGCGGGAGGGACCGCGGCCTAAGACTACGAGTCACATCATCCAA
CGCCAGCTCAACCAGCACGTTGACCCCGACTTCGGCATGAATAGCTTGAAAGGCTC
TGTGATACGGAGGCTTCAATCCCGCCACCAAGATATAACGCAGCTTGCTCTTTCCAA
GTTTGAGTTTACGCCGAAGCAATTTTTCTCCGGCCTGACGTGCGCAATCAATGACCC
CAAACACGTGCCATCTCATGACACTGTCCGACGCACAAAGAGGGAGATCGCTCACA
30 TGCACGACTACCACGATTACACTTGCCATTTGGCCCTTGCCGCCCAAGATGGCAAAC
AAGTGATAGCGAAGGGCTGGGGTCAAAGACATCCGCTTGCTGGGCCCCGGTATGCC
GGACCTCCTACCGAGTGGACCTTCTTTATGCCCCACGCAACGAAGCAGAGGTGCA
AGTGCTGGAAACTATTATTGAGGCAAGTATAGGATATATGAGTAACGCACCTGCC
TGGGAGGTTCCGAA

35 >SEQ ID NO: 52

ATGGCTAGCATGACTGGTGGACAGCAAATGGGTGCGGGATCCCACCACCACCAC
CACCAC

>SEQ ID NO: 53

MASMTGGQQMGRGSHHHHHH

40 >SEQ ID NO: 54

ATGAGGATAAATATCTCTTTGTCTAGTCTCTTTGAGAGACTGAGCAAATTGTCAT
CCCGATCCATCGCAATTACCTGCGGCGTAGTCTTGGCAAGTGCAATAGCATTCCCAA
TTATCCGAAGAGACTATCAAACGTTCCCTTGAAGTCCGAGCTATGCCCCACAGA
ACTTCCGAGGCTATATCATCGTTTGTGTCTTGTCACTTTTTAGGCAAGAGCAGAAAG
45 GGTTGGCAATCTATGACAGGCTTCCAGAAAAGAGGCGTTGGCTTGCCGACTTGCCG
TTTCGTGAAGGGACGAGGCCATCCATAACCTCCCACATTATACAGCGTCAGCGTACT
CAGCTGGTAGATCAAGAGTTTGCAACCAGAGAACTTATCGATAAGGTGATCCCACG
TGTGCAAGCAAGACACACAGACAAAACCTTTTCTCAGTACGTCAAATTTGAGTTTCA

TGCAAAGGCCATCTTCCTTCTCCCCTCTATCCCTATTAATGATCCTCTTAAACATTCCC
 TCACACGATACGGTAAGAAGAACCAAGCGTGAGATCGCTCACATGCATGATTACCA
 TGATTGCACTCTGCACTTGGCTCTTGCTGCTCAGGATGGTAAGGAAGTTTTGAAGAA
 GGGGTGGGGCCAGCGTCACCCACTGGCCGGACCAGGAGTCCCAGGCCCTCCTACTG
 5 AGTGGACCTTCSTTTACGCACCAAGGAACGAAGAGGAGGCCAGAGTTGTTCGAGATG
 ATAGTCGAAGCTAGTATTGGCTACATGACTAACGATCCTGCTGGTAAGATTGTTCGAA
 AATGCCAAG

(57) Формула изобретения

10 1. Нуклеиновая кислота, кодирующая люциферазу или ее функциональный фрагмент, и выбранная из группы, состоящей из:

(а) нуклеиновой кислоты, кодирующей белок, характеризующийся аминокислотной последовательностью, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34;

15 (б) нуклеиновой кислоты, кодирующей белок, имеющий последовательность, которая по крайней мере на 65% идентична аминокислотной последовательности, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34;

(в) нуклеиновой кислоты, кодирующей белок, в состав которого входит консенсусная последовательность SEQ ID NO: 35.

20 2. Нуклеиновая кислота по п.1, которая кодирует белок, характеризующийся аминокислотной последовательностью, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34.

3. Нуклеиновая кислота по п.1, которая кодирует белок, имеющий последовательность, которая по крайней мере на 62%, 65%, 70%, 75%, 80%, 85%, 90% идентична аминокислотной последовательности, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34.

4. Нуклеиновая кислота по п.1, которая кодирует белок, имеющий последовательность, которая на 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 98% или 99% идентична аминокислотной последовательности, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34.

5. Нуклеиновая кислота по п.1, которая кодирует белок, в состав которого входит консенсусная последовательность SEQ ID NO: 35.

6. Нуклеиновая кислота по п.1, кодирующая люциферазу, которая обладает способностью к окислению люциферина, представляющего собой 3-гидроксигиспидин или его аналог.

7. Нуклеиновая кислота по п.6, где аналог 3-гидроксигиспидина представляет собой (Е)-6-(4-(диэтиламино)стирил)-3,4-дигидрокси-2Н-пиран-2-он, (Е)-3,4-дигидрокси-6-(4-гидроксистирил)-2Н-пиран-2-он, (Е)-6-(2-(1Н-индол-3-ил)винил)-3,4-дигидрокси-2Н-пиран-2-он, (Е)-6-(2-(1,2,3,5,6,7-гексагидропиридо[3,2,1-ij]хинолин-9-ил)винил)-3,4-дигидрокси-2Н-пиран-2-он или (Е)-3,4-дигидрокси-6-(2-(6-гидроксиафталин-2-ил)винил)-2Н-пиран-2-он.

8. Экспрессионная кассета, содержащая молекулу нуклеиновой кислоты по любому из пп.1-7 под контролем регуляторных элементов, необходимых для экспрессии нуклеиновой кислоты в клетке-хозяине, которая будучи интегрированной в геном клетки или при введении в клетку в виде внехромосомного элемента способна обеспечить экспрессию люциферазы, кодируемой указанной нуклеиновой кислотой.

9. Клетка-хозяин, продуцирующая люциферазу, кодируемую нуклеиновой кислотой по любому из пп.1-7, содержащая экспрессионную кассету по п.8 в виде

внехромосомного элемента или элемента, интегрированного в геном этой клетки.

10. Люцифераза или ее функциональный фрагмент, выбранная из группы, состоящей из:

(а) белка, характеризующегося аминокислотной последовательностью, соответствующей аминокислотной последовательности, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34;

(б) белка, имеющего аминокислотную последовательность, которая по меньшей мере на 65% идентична аминокислотной последовательности, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34;

(в) белка, имеющего аминокислотную последовательность в состав которой входит консенсусная последовательность SEQ ID NO: 35.

11. Люцифераза по п.10, характеризующаяся аминокислотной последовательностью, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34.

12. Люцифераза по п.10, характеризующаяся аминокислотной последовательностью, которая по крайней мере на 62%, 65%, 70%, 75%, 80%, 85%, 90% идентична аминокислотной последовательности, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34.

13. Люцифераза по п.10, характеризующаяся аминокислотной последовательностью, которая на 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 98% или 99% идентична аминокислотной последовательности, представленной в SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32 или 34.

14. Люцифераза по п.10, в состав которой входит консенсусная последовательность SEQ ID NO: 35.

15. Люцифераза по п.10, которая обладает способностью к окислению люциферина, представляющего собой 3-гидроксигиспидин или его аналог.

16. Люцифераза по п.15, где аналог 3-гидроксигиспидина представляет собой (E)-6-(4-(диэтиламино)стирил)-3,4-дигидрокси-2H-пиран-2-он, (E)-3,4-дигидрокси-6-(4-гидроксистирил)-2H-пиран-2-он, (E)-6-(2-(1H-индол-3-ил)винил)-3,4-дигидрокси-2H-пиран-2-он, (E)-6-(2-(1,2,3,5,6,7-гексагидропиридо[3,2,1-ij]хинолин-9-ил)винил)-3,4-дигидрокси-2H-пиран-2-он или (E)-3,4-дигидрокси-6-(2-(6-гидроксиафталин-2-ил)винил)-2H-пиран-2-он.

17. Набор для мечения организмов, клеток, клеточных органелл или белков, содержащий (1) нуклеиновую кислоту по любому из пп.1-7, или экспрессионную кассету по п. 8 либо содержащий люциферазу по любому из пп.10-16, и (2) окисляемый указанной люциферазой люциферин.

18. Набор по п.17, в котором люциферин представляет собой 3-гидроксигиспидин или его аналог.

19. Набор по п.18, в котором аналог 3-гидроксигиспидина представляет собой (E)-6-(4-(диэтиламино)стирил)-3,4-дигидрокси-2H-пиран-2-он, (E)-3,4-дигидрокси-6-(4-гидроксистирил)-2H-пиран-2-он, (E)-6-(2-(1H-индол-3-ил)винил)-3,4-дигидрокси-2H-пиран-2-он, (E)-6-(2-(1,2,3,5,6,7-гексагидропиридо[3,2,1-ij]хинолин-9-ил)винил)-3,4-дигидрокси-2H-пиран-2-он или (E)-3,4-дигидрокси-6-(2-(6-гидроксиафталин-2-ил)винил)-2H-пиран-2-он.

20. Трансгенный организм, обладающий способностью к биоллюминесценции, содержащий нуклеиновую кислоту по любому из пп.1-7, где указанный трансгенный организм не является человеком.

21. Трансгенный организм по п.20, где указанный трансгенный организм способен к биоллюминесценции в присутствии 3-гидроксигиспидина или его аналога.

22. Трансгенный организм по п.21, где аналог 3-гидроксигиспицина представляет собой (Е)-6-(4-(диэтиламино)стирил)-3,4-дигидрокси-2Н-пиран-2-он, (Е)-3,4-дигидрокси-6-(4-гидроксистирил)-2Н-пиран-2-он, (Е)-6-(2-(1Н-индол-3-ил)винил)-3,4-дигидрокси-2Н-пиран-2-он, (Е)-6-(2-(1,2,3,5,6,7-гексагидропиридо[3,2,1-ij]хинолин-9-ил)винил)-3,4-дигидрокси-2Н-пиран-2-он или (Е)-3,4-дигидрокси-6-(2-(6-гидроксиафталин-2-ил)винил)-2Н-пиран-2-он.

23. Экспрессионная кассета, содержащая молекулу нуклеиновой кислоты по любому из пп.1-7, оперативно слитую с сигналом внутриклеточной локализации, под контролем регуляторных элементов, необходимых для экспрессии нуклеиновой кислоты в клетке-хозяине, которая будучи интегрированной в геном клетки или при введении в клетку в виде внехромосомного элемента способна обеспечить экспрессию люциферазы, кодируемой указанной нуклеиновой кислотой, связанной с сигналом внутриклеточной локализации.

24. Способ мечения клеток и клеточных органелл, включающий введение в клетку экспрессионной кассеты по п.8 или 24.

25. Способ по п.24, дополнительно включающий введение в клетку 3-гидроксигиспицина или его аналога.

26. Способ по п.25, где аналог 3-гидроксигиспицина представляет собой (Е)-6-(4-(диэтиламино)стирил)-3,4-дигидрокси-2Н-пиран-2-он, (Е)-3,4-дигидрокси-6-(4-гидроксистирил)-2Н-пиран-2-он, (Е)-6-(2-(1Н-индол-3-ил)винил)-3,4-дигидрокси-2Н-пиран-2-он, (Е)-6-(2-(1,2,3,5,6,7-гексагидропиридо[3,2,1-ij]хинолин-9-ил)винил)-3,4-дигидрокси-2Н-пиран-2-он или (Е)-3,4-дигидрокси-6-(2-(6-гидроксиафталин-2-ил)винил)-2Н-пиран-2-он.

25

30

35

40

45

SEQUENCE LISTING

<110> Ямпольский, Илья
 <120> Новые люциферазы и способы их использования
 <130> Yamp-luc
 <160> 54
 <170> PatentIn version 3.5
 <210> 1
 <211> 801
 <212> DNA
 <213> Neonothopanus nambi
 <400> 1
 atgcgcatcattacattagcctctcgtctctcttcgaacgctctccaaacttagcagtcgc 60
 agcatagcga ttacatgtgg agttgttctc gctccgcaa tcgctttcc catcatccgc 120
 agagactacc agactttcct agaagtggga ccctcgtacg ctccgcagaa cttagagga 180
 tacatcatcg tctgtgtcct ctgctattc cgccaagagc agaaaggct cgccatctat 240
 gatcgtcttc ccgagaaacg caggtgggtg gccgacctc ctttctgta aggaaccaga 300
 cccagcatta ccagccatat cattcagcga cagcgcactc aactggtcga tcaggagttt 360
 gccaccaggg agctcataga caaggtcatc cctcgcgtgc aagcacgaca caccgacaaa 420
 acgttcctca gcacatcaaa gttcgagttt catgcaagg ccatatttct cttgccttct 480
 atcccaatca acgacctct gaatatccct agccacgaca ctgtccgccg aacgaagcgc 540
 gagattgcac atatgcatga ttatcatgat tgcacacttc atcttgctct cgctgcgag 600
 gatggaaagg aggtgctgaa gaaagggttg ggacaacgac atcctttggc tggctctgga 660
 gttcctggtc caccaacgga atggactttt ctttatgctc ctgcaacga agaagaggct 720
 cgagtagtgg agatgatcgt tgaggcttcc atagggtata tgacgaacga tctgagga 780
 aagattgtag aaaacgcaa g 801
 <210> 2
 <211> 267
 <212> PRT
 <213> Neonothopanus nambi
 <400> 2
 Met Arg Ile Asn Ile Ser Leu Ser Ser Leu Phe Glu Arg Leu Ser Lys
 1 5 10 15
 Leu Ser Ser Arg Ser Ile Ala Ile Thr Cys Gly Val Val Leu Ala Ser
 20 25 30

Ala Ile Ala Phe Pro Ile Ile Arg Arg Asp Tyr Gln Thr Phe Leu Glu
 35 40 45

Val Gly Pro Ser Tyr Ala Pro Gln Asn Phe Arg Gly Tyr Ile Ile Val
 50 55 60

Cys Val Leu Ser Leu Phe Arg Gln Glu Gln Lys Gly Leu Ala Ile Tyr
 65 70 75 80

Asp Arg Leu Pro Glu Lys Arg Arg Trp Leu Ala Asp Leu Pro Phe Arg
 85 90 95

Glu Gly Thr Arg Pro Ser Ile Thr Ser His Ile Ile Gln Arg Gln Arg
 100 105 110

Thr Gln Leu Val Asp Gln Glu Phe Ala Thr Arg Glu Leu Ile Asp Lys
 115 120 125

Val Ile Pro Arg Val Gln Ala Arg His Thr Asp Lys Thr Phe Leu Ser
 130 135 140

Thr Ser Lys Phe Glu Phe His Ala Lys Ala Ile Phe Leu Leu Pro Ser
 145 150 155 160

Ile Pro Ile Asn Asp Pro Leu Asn Ile Pro Ser His Asp Thr Val Arg
 165 170 175

Arg Thr Lys Arg Glu Ile Ala His Met His Asp Tyr His Asp Cys Thr
 180 185 190

Leu His Leu Ala Leu Ala Ala Gln Asp Gly Lys Glu Val Leu Lys Lys
 195 200 205

Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Val Pro Gly Pro
 210 215 220

Pro Thr Glu Trp Thr Phe Leu Tyr Ala Pro Arg Asn Glu Glu Glu Ala
 225 230 235 240

Arg Val Val Glu Met Ile Val Glu Ala Ser Ile Gly Tyr Met Thr Asn
 245 250 255

Asp Pro Ala Gly Lys Ile Val Glu Asn Ala Lys
 260 265

<210> 3
 <211> 798
 <212> DNA
 <213> Armillaria gallica

<400> 3
 atgtccttca tcgacagcat gaaacttgac ctcgtcggac acctctttgg catcaggaat 60
 cgcggcttag cgcgcccttg ttgtgctcta gcagtcgcct ctactatcgc cttcccttac 120
 attcgtaggg actaccagac atttttatct ggcggtcctt cttacgctcc ccagaatata 180
 agaggatatt tcacgtctct cgtttctggc ttgttccgctc aggagcaaaa ggccttgcg 240
 atatatgata gccttccga gaagcgcagg tggctgcctg acttgcctcc tcgcaatggc 300
 ccgcgccga tcacgaccag ccatataatc caaagacagc gcaaccaggc gccggacccc 360
 aagttcgccc tcgaggaact caaggccacg gttattccac ggtgacaggc tcgccatact 420
 gacctacccc atctcagcct atccaaattc gagttccatg ctgaagcaat tttcctgctc 480
 ccctctgtac ccacgatga tccaaaaaat gttccaagtc acgacacggt gcgcaggacg 540
 aaaagggaga tcgcgcatat gcacgactac catgacttca cgctgcatct tgcaactggc 600
 gcccaagacg ggaaggaagt cgtgtcgaag ggatgggggc agcgacaccc cctagcaggc 660
 cctggcgctc ctggtccacc tacggagtgg acatttattt atgcgccacg taacgaagag 720
 gaactggcag tggtggaat gattatcgag gcatcaatag gctatatgac caatgaccct 780
 gctggagtag ttatcgca 798

<210> 4
 <211> 266
 <212> PRT
 <213> Armillaria gallica

<400> 4

Met Ser Phe Ile Asp Ser Met Lys Leu Asp Leu Val Gly His Leu Phe
 1 5 10 15

Gly Ile Arg Asn Arg Gly Leu Ala Ala Ala Cys Cys Ala Leu Ala Val
 20 25 30

Ala Ser Thr Ile Ala Phe Pro Tyr Ile Arg Arg Asp Tyr Gln Thr Phe
 35 40 45

Leu Ser Gly Gly Pro Ser Tyr Ala Pro Gln Asn Ile Arg Gly Tyr Phe
 50 55 60

Ile Val Cys Val Leu Ala Leu Phe Arg Gln Glu Gln Lys Gly Leu Ala
 65 70 75 80

Ile Tyr Asp Arg Leu Pro Glu Lys Arg Arg Trp Leu Pro Asp Leu Pro
 85 90 95

Pro Arg Asn Gly Pro Arg Pro Ile Thr Thr Ser His Ile Ile Gln Arg
 100 105 110

Gln Arg Asn Gln Ala Pro Asp Pro Lys Phe Ala Leu Glu Glu Leu Lys
 115 120 125

Ala Thr Val Ile Pro Arg Val Gln Ala Arg His Thr Asp Leu Thr His
 130 135 140

Leu Ser Leu Ser Lys Phe Glu Phe His Ala Glu Ala Ile Phe Leu Leu
 145 150 155 160

Pro Ser Val Pro Ile Asp Asp Pro Lys Asn Val Pro Ser His Asp Thr
 165 170 175

Val Arg Arg Thr Lys Arg Glu Ile Ala His Met His Asp Tyr His Asp
 180 185 190

Phe Thr Leu His Leu Ala Leu Ala Ala Gln Asp Gly Lys Glu Val Val
 195 200 205

Ser Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Val Pro
 210 215 220

Gly Pro Pro Thr Glu Trp Thr Phe Ile Tyr Ala Pro Arg Asn Glu Glu
 225 230 235 240

Glu Leu Ala Val Val Glu Met Ile Ile Glu Ala Ser Ile Gly Tyr Met
 245 250 255

Thr Asn Asp Pro Ala Gly Val Val Ile Ala
 260 265

<210> 5
 <211> 801
 <212> DNA
 <213> Armillaria mellea

<400> 5
 atgtccttct tcgacagcgt gaaacttgac ctgctcggac gcctctttgg catcaggaat 60
 cgcggcttag ctgttacttg ttgtgctgtg gcagtcgcct ctatcatcgc gttcccttac 120
 attcgtaggg actaccagac atttttatct gggggtcct cctacgctcc ccagaacatc 180

agaggataacc tcattgtctg cgtcctggcc ttgttccgtc aggagcaaaa aggccttgcg 240
 atatacgacc gccttcccga gaagcgagg tggctacctg acttgcctcc tcgcatggc 300
 ccacggccca tcacgaccag ccatataatc caaagacagc gcaaccaggc gccggacctc 360
 aagttcgccc tcgaggaact caaggccacg gtcattccac ggggtgcaggc tcgccacct 420
 gacctcacc atctcagcct atccaagttc gagttccatg ctgaagcaat cttcctgctc 480
 ccctctgtac ccctcagatga tccaagaat gtgccaagtc acgacacggt gcgcaggacg 540
 aagagggaaa ttgcgcatat gcacgactac catgactaca cgctgcatct tgcgttgcc 600
 gcccaagacg ggaaggaagt cgtatcaaag ggatgggggc agcgacaccc gctggcaggc 660
 cctggcgctc ctgggtccacc gacggagtgg acgtttatct atgcgccacg taacgaagag 720
 gagctggcag tggtggaat gattatcgag gcatcgatag gctatatgac caatgacct 780
 gcaggaaaaa ctatcgata g 801

<210> 6
 <211> 266
 <212> PRT
 <213> Armillaria mellea

<400> 6

Met Ser Phe Phe Asp Ser Val Lys Leu Asp Leu Val Gly Arg Leu Phe
 1 5 10 15

Gly Ile Arg Asn Arg Gly Leu Ala Val Thr Cys Cys Ala Val Ala Val
 20 25 30

Ala Ser Ile Ile Ala Phe Pro Tyr Ile Arg Arg Asp Tyr Gln Thr Phe
 35 40 45

Leu Ser Gly Gly Pro Ser Tyr Ala Pro Gln Asn Ile Arg Gly Tyr Leu
 50 55 60

Ile Val Cys Val Leu Ala Leu Phe Arg Gln Glu Gln Lys Gly Leu Ala
 65 70 75 80

Ile Tyr Asp Arg Leu Pro Glu Lys Arg Arg Trp Leu Pro Asp Leu Pro
 85 90 95

Pro Arg Asp Gly Pro Arg Pro Ile Thr Thr Ser His Ile Ile Gln Arg
 100 105 110

Gln Arg Asn Gln Ala Pro Asp Leu Lys Phe Ala Leu Glu Glu Leu Lys
 115 120 125

Ala Thr Val Ile Pro Arg Val Gln Ala Arg His Thr Asp Leu Thr His
 130 135 140

Leu Ser Leu Ser Lys Phe Glu Phe His Ala Glu Ala Ile Phe Leu Leu
 145 150 155 160

Pro Ser Val Pro Ile Asp Asp Pro Lys Asn Val Pro Ser His Asp Thr
 165 170 175

Val Arg Arg Thr Lys Arg Glu Ile Ala His Met His Asp Tyr His Asp
 180 185 190

Tyr Thr Leu His Leu Ala Leu Ala Ala Gln Asp Gly Lys Glu Val Val
 195 200 205

Ser Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Val Pro
 210 215 220

Gly Pro Pro Thr Glu Trp Thr Phe Ile Tyr Ala Pro Arg Asn Glu Glu
 225 230 235 240

Glu Leu Ala Val Val Glu Met Ile Ile Glu Ala Ser Ile Gly Tyr Met
 245 250 255

Thr Asn Asp Pro Ala Gly Lys Thr Ile Ala
 260 265

<210> 7
 <211> 801
 <212> DNA
 <213> Armillaria ostoyae

<400> 7
 atgtccttca tcgacagcat gaaacttgac ttcgtcggac acctctttgg catcaggaat 60
 cgcggttag ccaccgcttg ttgtgctgtg gcagtcgctt ctgccatcgc cttcccttac 120
 attcgtaggg actaccagac attcttatct ggcggtccct cttacgctcc ccagaacatc 180
 aaaggatata tcatcgtctg cgtcctggcc ttgttccgtc aggagcaaaa ggccttgcg 240
 atatatgacc gccttcccga gaagcgcagg tggctacctg acttgcctcc tcgcaatggc 300
 ccgcggccca tcacgaccag ccatataatc caaagacagc gcaaccaggc gccagactcc 360
 aagttcgccc tcgaggaact caaggctacg gtcattccac ggggtgcaggc tcgccacact 420
 gacctacccc atctcagcct atccaagttc gagttccatg ctgaagcaat cttcctgctc 480
 ccctctgtac ccatcgaatg tccaaaaaat gttccaagtc atgacacggt gcgcaggacg 540

aagagggaga tcgcatat gcacgactac catgacttta cgttgcatct tgactggcc 600
 gcccaagacg ggaaggaagt cgtggcgaag ggatgggggc agcgacaccc gctggcaggc 660
 cctggcgcttc ctggtccacc tacggagtgg acgtttatatt atgcgccacg taacgaagag 720
 gaactggcag tggtggaat gattatcgag gcatcaatag gctatatgac caatgacct 780
 gctggaacag ttatcgata g 801

<210> 8
 <211> 266
 <212> PRT
 <213> Armillaria ostoyae

<400> 8

Met Ser Phe Ile Asp Ser Met Lys Leu Asp Phe Val Gly His Leu Phe
 1 5 10 15

Gly Ile Arg Asn Arg Gly Leu Ala Thr Ala Cys Cys Ala Val Ala Val
 20 25 30

Ala Ser Ala Ile Ala Phe Pro Tyr Ile Arg Arg Asp Tyr Gln Thr Phe
 35 40 45

Leu Ser Gly Gly Pro Ser Tyr Ala Pro Gln Asn Ile Lys Gly Tyr Leu
 50 55 60

Ile Val Cys Val Leu Ala Leu Phe Arg Gln Glu Gln Lys Gly Leu Ala
 65 70 75 80

Ile Tyr Asp Arg Leu Pro Glu Lys Arg Arg Trp Leu Pro Asp Leu Pro
 85 90 95

Pro Arg Asn Gly Pro Arg Pro Ile Thr Thr Ser His Ile Ile Gln Arg
 100 105 110

Gln Arg Asn Gln Ala Pro Asp Ser Lys Phe Ala Leu Glu Glu Leu Lys
 115 120 125

Ala Thr Val Ile Pro Arg Val Gln Ala Arg His Thr Asp Leu Thr His
 130 135 140

Leu Ser Leu Ser Lys Phe Glu Phe His Ala Glu Ala Ile Phe Leu Leu
 145 150 155 160

Pro Ser Val Pro Ile Asp Asp Pro Lys Asn Val Pro Ser His Asp Thr
 165 170 175

7

Val Arg Arg Thr Lys Arg Glu Ile Ala His Met His Asp Tyr His Asp
 180 185 190

Phe Thr Leu His Leu Ala Leu Ala Ala Gln Asp Gly Lys Glu Val Val
 195 200 205

Ala Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Val Pro
 210 215 220

Gly Pro Pro Thr Glu Trp Thr Phe Ile Tyr Ala Pro Arg Asn Glu Glu
 225 230 235 240

Glu Leu Ala Val Val Glu Met Ile Ile Glu Ala Ser Ile Gly Tyr Met
 245 250 255

Thr Asn Asp Pro Ala Gly Thr Val Ile Val
 260 265

<210> 9
 <211> 753
 <212> DNA
 <213> Mycena chlorophos

<400> 9
 atggtccaac tcacccgaac ctccggattc atcgccgctg cggccattgt tgctgccatc 60
 gccttcccgt tcattcgtcg agactaccag acgttccttc gtgggtgggcc gtcctatgcc 120
 ccacagaaca tccgcggcta tatcatcgtc ctggttctgt ccctcttccg cggcgaggag 180
 aagggtcttg caatctacga gcccttctcct gagaagcgca catggctgcc ggagcttccg 240
 cggcgcgcgg gagaccggcc caagacgacg agccacatca tccaacggca gctcgaccag 300
 taccocgacc cggactttgt cctcaaagcc ctgaaagcga cggatcatccc gcgtgtccaa 360
 gcccggcaca cagacaagac tcacctcgcg ctgtccaagt tcgagttcca tgctgaggcc 420
 atcttcgtgc gcccgaaat cgccatcgac gaccogaagc atatcccag ccacgacacg 480
 gtgcgacgga cgaagcgcga gattgcgcac atgcacgact atcacgactg cagcgtgcat 540
 ttggcgctag cggcgcagga cgcgaagcag gtgctgcaga agggctgggg ccagcgcgat 600
 ccgctggcag ggcctgggat gcccgggccg cccacggagt ggacgttctt gtatgccccg 660
 aggaccgagg aggaagtgaa gtttgtggag accattgtcg aggcctctat cgcgtacatg 720
 acgaacgcgg agaagccggt cgagctggtg cag 753

<210> 10
 <211> 251
 <212> PRT

<213> Mycena chlorophos

<400> 10

Met Val Gln Leu Thr Arg Thr Ser Gly Phe Ile Ala Ala Ala Ala Ile
1 5 10 15

Val Ala Ala Ile Ala Phe Pro Phe Ile Arg Arg Asp Tyr Gln Thr Phe
20 25 30

Leu Arg Gly Gly Pro Ser Tyr Ala Pro Gln Asn Ile Arg Gly Tyr Ile
35 40 45

Ile Val Leu Val Leu Ser Leu Phe Arg Gly Glu Glu Lys Gly Leu Ala
50 55 60

Ile Tyr Glu Pro Leu Pro Glu Lys Arg Thr Trp Leu Pro Glu Leu Pro
65 70 75 80

Arg Arg Ala Gly Asp Arg Pro Lys Thr Thr Ser His Ile Ile Gln Arg
85 90 95

Gln Leu Asp Gln Tyr Pro Asp Pro Asp Phe Val Leu Lys Ala Leu Lys
100 105 110

Ala Thr Val Ile Pro Arg Val Gln Ala Arg His Thr Asp Lys Thr His
115 120 125

Leu Ala Leu Ser Lys Phe Glu Phe His Ala Glu Ala Ile Phe Val Arg
130 135 140

Pro Glu Ile Ala Ile Asp Asp Pro Lys His Ile Pro Ser His Asp Thr
145 150 155 160

Val Arg Arg Thr Lys Arg Glu Ile Ala His Met His Asp Tyr His Asp
165 170 175

Cys Thr Leu His Leu Ala Leu Ala Ala Gln Asp Ala Lys Gln Val Leu
180 185 190

Gln Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Met Pro
195 200 205

Gly Pro Pro Thr Glu Trp Thr Phe Leu Tyr Ala Pro Arg Thr Glu Glu
210 215 220

Glu Val Lys Val Val Glu Thr Ile Val Glu Ala Ser Ile Ala Tyr Met

225	230	235	240									
Thr	Asn	Ala	Glu	Lys	Pro	Val	Glu	Leu	Val	Gln		
				245						250		
<p><210> 11 <211> 738 <212> DNA <213> Mycena citricolor</p> <p><400> 11</p>												
atggc	ttatc	agctc	acttg	gattc	agact	ctcgt	gctgg	gtgcc	cttgt	ggcaat	ggca	60
gtagc	gttcc	ccttc	atcaa	gaaag	actac	gagac	gttcc	tgaag	ggcgg	cccct	cctat	120
gcgccc	aaaa	acgtt	cgcg	ataca	tcatc	gtgct	cgtc	tcg	gctctt	cgcca	aagag	180
cagctc	gggc	tggag	atcta	cgacc	gcatg	cccg	gaaaac	gtc	gctggct	cgcaat	ctc	240
cctcag	cgcg	agggc	ccccg	cccc	aagacc	acaag	tcaca	tcat	ccagcg	gcag	ctcagc	300
cagc	acacg	g	accc	gcatt	cg	g	ccg	cg	tac	ctca	aag	360
acacc	gtcat	tcc	g	cg	cg	cg	cg	cg	cg	cg	cg	420
cgat	cttcc	tga	acg	cg	ga	cgt	g	ccc	gag	ggc	cg	480
cg	gg	gacca	agg	g	gagat	cg	g	cac	gact	acc	g	540
g	ctc	g	cag	cg	gat	gg	ga	agg	gtc	g	gca	600
g	ctc	g	cag	cg	gat	gg	ga	agg	gtc	g	gca	660
ct	gg	gg	gg	gac	cc	gg	cg	gt	g	cc	g	720
aat	ga	a	g	agg	agat	gg	cg	t	gag	cag	at	780
aat	gc	c	ctg	cg	ctg	gaa						

<210> 12	<211> 246	<212> PRT	<213> Mycena citricolor												
<400> 12															
Met	Ala	Tyr	Gln	Leu	Thr	Trp	Ile	Gln	Thr	Leu	Val	Leu	Gly	Ala	Leu
1				5					10					15	
Val	Ala	Met	Ala	Val	Ala	Phe	Pro	Phe	Ile	Lys	Lys	Asp	Tyr	Glu	Thr
			20					25					30		
Phe	Leu	Lys	Gly	Gly	Pro	Ser	Tyr	Ala	Pro	Gln	Asn	Val	Arg	Gly	Tyr
		35					40					45			
Ile	Ile	Val	Leu	Val	Leu	Ala	Leu	Phe	Arg	Gln	Glu	Gln	Leu	Gly	Leu
	50					55					60				

Glu Ile Tyr Asp Arg Met Pro Glu Lys Arg Arg Trp Leu Ala Asn Leu
 65 70 75 80

Pro Gln Arg Glu Gly Pro Arg Pro Lys Thr Thr Ser His Ile Ile Gln
 85 90 95

Arg Gln Leu Ser Gln His Thr Asp Pro Ala Phe Gly Ala Ala Tyr Leu
 100 105 110

Lys Asp Thr Val Ile Pro Arg Val Gln Ala Arg His Ala Ala Asn Thr
 115 120 125

His Ile Ala Arg Ser Thr Phe Glu Phe His Ala Ala Ala Ile Phe Leu
 130 135 140

Asn Ala Asp Val Pro Leu Pro Glu Gly Leu Pro Ala Ser Glu Thr Val
 145 150 155 160

Arg Arg Thr Lys Gly Glu Ile Ala His Met His Asp Tyr His Asp Phe
 165 170 175

Thr Leu His Leu Ala Leu Ala Ala Asp Gly Lys Glu Val Val Gly
 180 185 190

Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Val Pro Gly
 195 200 205

Pro Pro Asn Glu Trp Thr Phe Val Tyr Ala Pro Arg Asn Glu Glu Glu
 210 215 220

Met Gly Val Val Glu Gln Ile Val Glu Ala Ala Ile Gly Tyr Met Ser
 225 230 235 240

Asn Val Pro Ala Leu Glu
 245

<210> 13
 <211> 786
 <212> DNA
 <213> *Omphalotus olearius*

<400> 13
 atgctcccag ctttcatcta caaaccaagg ctagtgatca cttgtgtatt cgttctggcc 60
 tccgcactcg catttccctt catacgcaaa gattaccaga ctttcctgga ggtgggaccc 120
 tcgtacgcc cgcagaacct ccaagatac atcatcgtct gtgtactctc tctgttccgg 180

caagaacaga aagacgtagc gatttatgat cgccttcctg agaaaaggag gtggttagga 240
gacctccgt ttcgcgaggg gccaaagaccg agtatcacta gccatatcat ccagcgacag 300
cgcaccaat tggctgacgc cgagttcgct accaaagagc tgataggcaa aatcatccct 360
cgcgtccaag cccgacacac caacacaaca ttcctcagca catctaaatt cgaattccac 420
gccagggcca tcttcctttt gccctctatc ccaatcaacg accctcaaaa cattccaagc 480
cacgataccg ttcgtcgcac gaaacgagc atcgcgcata tgcattgatta tcacgactgt 540
acgttgcatc tcgcacttgc tgctcaagat ggaagaggag ttttagagaa aggatggggg 600
cagcgacatc ctcttgctgg acctggtgtt cctggccgc cgacggagtg gacgtttctt 660
tatgcaccgc gcagcgaaga ggaggttcgg gttgtggaga tgattgttga ggcattcagt 720
gtgtatatga cgaatgatcc tgcggataaa atcgtagaag ctactgtgca ggttactgaa 780
gaatag 786

<210> 14
<211> 261
<212> PRT
<213> *Omphalotus olearius*
<400> 14

Met Leu Pro Ala Phe Ile Tyr Lys Pro Arg Leu Val Ile Thr Cys Val
1 5 10 15
Phe Val Leu Ala Ser Ala Leu Ala Phe Pro Phe Ile Arg Lys Asp Tyr
20 25 30
Gln Thr Phe Leu Glu Val Gly Pro Ser Tyr Ala Pro Gln Asn Leu Gln
35 40 45
Gly Tyr Ile Ile Val Cys Val Leu Ser Leu Phe Arg Gln Glu Gln Lys
50 55 60
Asp Val Ala Ile Tyr Asp Arg Leu Pro Glu Lys Arg Arg Trp Leu Gly
65 70 75 80
Asp Leu Pro Phe Arg Glu Gly Pro Arg Pro Ser Ile Thr Ser His Ile
85 90 95
Ile Gln Arg Gln Arg Thr Gln Leu Ala Asp Ala Glu Phe Ala Thr Lys
100 105 110
Glu Leu Ile Gly Lys Ile Ile Pro Arg Val Gln Ala Arg His Thr Asn
115 120 125

Thr Thr Phe Leu Ser Thr Ser Lys Phe Glu Phe His Ala Gln Ala Ile
 130 135 140

Phe Leu Leu Pro Ser Ile Pro Ile Asn Asp Pro Gln Asn Ile Pro Ser
 145 150 155 160

His Asp Thr Val Arg Arg Thr Lys Arg Glu Ile Ala His Met His Asp
 165 170 175

Tyr His Asp Cys Thr Leu His Leu Ala Leu Ala Ala Gln Asp Gly Lys
 180 185 190

Glu Val Leu Glu Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro
 195 200 205

Gly Val Pro Gly Pro Pro Thr Glu Trp Thr Phe Leu Tyr Ala Pro Arg
 210 215 220

Ser Glu Glu Glu Val Arg Val Val Glu Met Ile Val Glu Ala Ser Val
 225 230 235 240

Val Tyr Met Thr Asn Asp Pro Ala Asp Lys Ile Val Glu Ala Thr Val
 245 250 255

Gln Gly Thr Glu Glu
 260

<210> 15
 <211> 747
 <212> DNA
 <213> *Panellus stipticus*

<400> 15
 atgaacatca acctgaaagc tctgatcgga gtctgtgccg tgctcatcac cgctgcagtg 60
 ttccccttcg ttcgtaaaga ctatcacacc tttcttgaag gtggaccatc ctacgcgccg 120
 cagaatttgc aagctatat catcgtgttg gtgctctcac tctttcgagg ggaggagacg 180
 ggattggaaa tatacgaccg cttgcccga aaacgccgct ggctcgagga gctgcctggt 240
 cggaaggcc cgcgccc aaa gacaaccagc cacatcattc agagacagtt gaatcagcac 300
 gttgaccggt acttcggaat gaactctttg aaaggtccg tcatccggcg cettcaatcc 360
 cgccaccagg acataactca actcgcactc tcgaaattcg aattccacgc cgaggccata 420
 tttctgcgcc ccgatgtcgc gatcaacgat cccaacacg tcccagacca cgacacggtg 480
 cgccgcaaaa agcgcgagat agctcacatg cactgactacc atgattacac gtgtcatttg 540

gcgctcgcag cgcaggatgg gaagcaagtg attgcaaaag ggtggggcca gagacatccg 600
 ctcgccggac cgggcatgcc ggggccgccc acggagtgga catttttgta tgcgccgagg 660
 aatgaggcgg aggttcaagt gttggagacg attatcgaag cgtcaatcgg gtacatgtcg 720
 aacgcaccag ccttgggtgg gagcgag 747

<210> 16
 <211> 249
 <212> PRT
 <213> Panellus stipticus

<400> 16

Met Asn Ile Asn Leu Lys Ala Leu Ile Gly Val Cys Ala Val Leu Ile
 1 5 10 15
 Thr Ala Ala Val Phe Pro Phe Val Arg Lys Asp Tyr His Thr Phe Leu
 20 25 30
 Glu Gly Gly Pro Ser Tyr Ala Pro Gln Asn Leu Gln Gly Tyr Ile Ile
 35 40 45
 Val Leu Val Leu Ser Leu Phe Arg Gly Glu Glu Thr Gly Leu Glu Ile
 50 55 60
 Tyr Asp Arg Leu Pro Glu Lys Arg Arg Trp Leu Glu Glu Leu Pro Val
 65 70 75 80
 Arg Glu Gly Pro Arg Pro Lys Thr Thr Ser His Ile Ile Gln Arg Gln
 85 90 95
 Leu Asn Gln His Val Asp Pro Asp Phe Gly Met Asn Ser Leu Lys Gly
 100 105 110
 Ser Val Ile Arg Arg Leu Gln Ser Arg His Gln Asp Ile Thr Gln Leu
 115 120 125
 Ala Leu Ser Lys Phe Glu Phe His Ala Glu Ala Ile Phe Leu Arg Pro
 130 135 140
 Asp Val Ala Ile Asn Asp Pro Lys His Val Pro Ser His Asp Thr Val
 145 150 155 160
 Arg Arg Thr Lys Arg Glu Ile Ala His Met His Asp Tyr His Asp Tyr
 165 170 175

Thr Cys His Leu Ala Leu Ala Ala Gln Asp Gly Lys Gln Val Ile Ala
 180 185 190

Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Met Pro Gly
 195 200 205

Pro Pro Thr Glu Trp Thr Phe Leu Tyr Ala Pro Arg Asn Glu Ala Glu
 210 215 220

Val Gln Val Leu Glu Thr Ile Ile Glu Ala Ser Ile Gly Tyr Met Ser
 225 230 235 240

Asn Ala Pro Ala Leu Gly Gly Ser Glu
 245

<210> 17
 <211> 747
 <212> DNA
 <213> Panellus stipticus
 <400> 17
 atgaacatca acctgaaagc tctgatcgga gtctgtgccg tgctcatcac cgctgcagtg 60
 ttccccttcg ttcgtaaaga ctatcacacc tttcttgaag gtggaccatc ctacgcgccg 120
 cagaatttgc aagctatat catcgtgttg gtgctctcac tctttcgagg ggaggagacg 180
 ggattgaaa tatacgaccg cttgcccga aaacgccgt ggctcgagga gctgcctgtt 240
 cggaaggcc cgcgccaaa gacaaccagc cacatcattc agagacagtt gaatcagcac 300
 gttgaccgg acttcggaat gaactctttg aaagctccg tcatccggcg cttcaatcc 360
 cgccaccagg acataactca actgcactc tcgaaattcg aattccacgc cgaggccata 420
 tttctgcgcc cggatatacgc gatcaacgat cccaacacg tcccgagcca cgacacggtg 480
 cgccgcacaa agcgcgagat agctcacatg cagcactacc atgattacac gtgtcatttg 540
 ggcctcgcag cgcaggatgg gaagcaagtg attgcaaaag ggtggggcca gagacatccg 600
 ctcgccggac cgggcatgcc ggggccgccg acggagtgga catttttgta tgcgccgagg 660
 aatgaggcgg aggttcaagt gttggagacg attatcgaag cgtcaatcgg gtacatgtcg 720
 aacgcaccag ccttgggtgg gagcgag 747

<210> 18
 <211> 249
 <212> PRT
 <213> Panellus stipticus
 <400> 18
 Met Asn Ile Asn Leu Lys Ala Leu Ile Gly Val Cys Ala Val Leu Ile

1 5 10 15

Thr Ala Ala Val Phe Pro Phe Val Arg Lys Asp Tyr His Thr Phe Leu
 20 25 30

Glu Gly Gly Pro Ser Tyr Ala Pro Gln Asn Leu Gln Gly Tyr Ile Ile
 35 40 45

Val Leu Val Leu Ser Leu Phe Arg Gly Glu Glu Thr Gly Leu Glu Ile
 50 55 60

Tyr Asp Arg Leu Pro Glu Lys Arg Arg Trp Leu Glu Glu Leu Pro Val
 65 70 75 80

Arg Glu Gly Pro Arg Pro Lys Thr Thr Ser His Ile Ile Gln Arg Gln
 85 90 95

Leu Asn Gln His Val Asp Pro Asp Phe Gly Met Asn Ser Leu Lys Gly
 100 105 110

Ser Val Ile Arg Arg Leu Gln Ser Arg His Gln Asp Ile Thr Gln Leu
 115 120 125

Ala Leu Ser Lys Phe Glu Phe His Ala Glu Ala Ile Phe Leu Arg Pro
 130 135 140

Asp Ile Ala Ile Asn Asp Pro Lys His Val Pro Ser His Asp Thr Val
 145 150 155 160

Arg Arg Thr Lys Arg Glu Ile Ala His Met His Asp Tyr His Asp Tyr
 165 170 175

Thr Cys His Leu Ala Leu Ala Ala Gln Asp Gly Lys Gln Val Ile Ala
 180 185 190

Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Met Pro Gly
 195 200 205

Pro Pro Thr Glu Trp Thr Phe Leu Tyr Ala Pro Arg Asn Glu Ala Glu
 210 215 220

Val Gln Val Leu Glu Thr Ile Ile Glu Ala Ser Ile Gly Tyr Met Ser
 225 230 235 240

Asn Ala Pro Ala Leu Gly Gly Ser Glu
 245

<210> 19
 <211> 690
 <212> DNA
 <213> artificial sequence

<220>
 <223> Фрагмент нуклеотидной последовательности, кодирующий функциональный фрагмент люциферазы из *Neonothopanus nambi*

<400> 19
 atcatccgca gagactacca gactttccta gaagtgggac cctcgtacgc tccgcagaac 60
 tttagaggat acatcatcgt ctgtgtcctc tcgctattcc gccaaagca gaaagggctc 120
 gccatctatg atcgtcttcc cgagaaacgc aggtggttgg ccgaccttcc ctttcgtgaa 180
 ggaaccagac ccagcattac cagccatata attcagcgac agcgcactca actggtcgat 240
 caggagttag ccaccagga gtcctatagac aaggatcacc ctccgctgca agcagcagac 300
 accgacaaaa cgttcctcag cacatcaaag ttcgagttag atgccaagga catatttctc 360
 ttgccttcta tcccaatcaa cgacctctg aatatcccta gccacgacac tgcgccgca 420
 acgaagcgcg agattgcaca tatgcatgat tatcatgatt gcacacttca tcttgctctc 480
 gctgcgcagg atgaaaagga ggtgctgaag aaaggttggg gacaacgaca tcctttggct 540
 ggtcctggag ttctgtgtcc accaacggaa tggacttttc tttatgcgcc tcgcaacgaa 600
 gaagaggctc gagtagtga gatgatcgtt gaggcttcca taggtatat gacgaacgat 660
 cctgcaggaa agattgtaga aaacgccaag 690

<210> 20
 <211> 230
 <212> PRT
 <213> artificial sequence

<220>
 <223> аминокислотная последовательность функционального фрагмента люциферазы из *Neonothopanus nambi*

<400> 20
 Ile Ile Arg Arg Asp Tyr Gln Thr Phe Leu Glu Val Gly Pro Ser Tyr
 1 5 10 15
 Ala Pro Gln Asn Phe Arg Gly Tyr Ile Ile Val Cys Val Leu Ser Leu
 20 25 30
 Phe Arg Gln Glu Gln Lys Gly Leu Ala Ile Tyr Asp Arg Leu Pro Glu
 35 40 45
 Lys Arg Arg Trp Leu Ala Asp Leu Pro Phe Arg Glu Gly Thr Arg Pro

	50		55		60															
Ser	Ile	Thr	Ser	His	Ile	Ile	Gln	Arg	Gln	Arg	Thr	Gln	Leu	Val	Asp					
65					70					75					80					
Gln	Glu	Phe	Ala	Thr	Arg	Glu	Leu	Ile	Asp	Lys	Val	Ile	Pro	Arg	Val					
				85					90						95					
Gln	Ala	Arg	His	Thr	Asp	Lys	Thr	Phe	Leu	Ser	Thr	Ser	Lys	Phe	Glu					
			100					105						110						
Phe	His	Ala	Lys	Ala	Ile	Phe	Leu	Leu	Pro	Ser	Ile	Pro	Ile	Asn	Asp					
		115					120							125						
Pro	Leu	Asn	Ile	Pro	Ser	His	Asp	Thr	Val	Arg	Arg	Thr	Lys	Arg	Glu					
	130					135						140								
Ile	Ala	His	Met	His	Asp	Tyr	His	Asp	Cys	Thr	Leu	His	Leu	Ala	Leu					
145					150					155					160					
Ala	Ala	Gln	Asp	Gly	Lys	Glu	Val	Leu	Lys	Lys	Gly	Trp	Gly	Gln	Arg					
				165					170						175					
His	Pro	Leu	Ala	Gly	Pro	Gly	Val	Pro	Gly	Pro	Pro	Thr	Glu	Trp	Thr					
			180					185						190						
Phe	Leu	Tyr	Ala	Pro	Arg	Asn	Glu	Glu	Glu	Ala	Arg	Val	Val	Glu	Met					
		195					200						205							
Ile	Val	Glu	Ala	Ser	Ile	Gly	Tyr	Met	Thr	Asn	Asp	Pro	Ala	Gly	Lys					
	210					215					220									
Ile	Val	Glu	Asn	Ala	Lys															
225					230															
<210>	21																			
<211>	681																			
<212>	DNA																			
<213>	artificial																			
<220>																				
<223>	фрагмент нуклеотидной последовательности, кодирующий функциональный фрагмент люциферазы из Armillaria gallica																			
<400>	21																			
tacattcgta	gggactacca	gacattttta	tctggcggtc	cctcttacgc	tccccagaat															60
atcagaggat	atctcatcgt	ctgcgttctg	gccttgttcc	gtcaggagca	aaagggcctt															120


```

gcgatatatg atcgccttcc cgagaagcgc aggtggctgc ctgacttgcc tcctcgcaat      180
ggcccgcggc cgatcacgac cagccatata atccaagac agcgcaacca ggcgccggac      240
cccaagtctg ccctcgagga actcaaggcc acggttattc cacgggtgca ggctcgccat      300
actgacctca cccatctcag cctatccaaa ttcgagtacc atgctgaagc aattttcctg      360
ctcccctctg taccatcga tgatccaaaa aatgttccaa gtcacgacac ggtgcgcagg      420
acgaaaaggg agatcgcgca tatgcacgac taccatgact tcacgctgca tcttgactg      480
gccgccaag acggaagga agtcgtgtcg aagggatggg ggcagcgaca cccctagca      540
ggccctggcg ttctgtgtcc acctacggag tggacattta tttatgcgcc acgtaacgaa      600
gaggaactgg cagtgggtga aatgattatc gaggcatcaa taggctatat gaccaatgac      660
cctgctggag tagttatcgc a                                             681

```

```

<210> 22
<211> 227
<212> PRT
<213> artificial

```

```

<220>
<223> Аминокислотная последовательность функционального фрагмента
люциферазы из Armillaria gallica

```

```

<400> 22

```

```

Tyr Ile Arg Arg Asp Tyr Gln Thr Phe Leu Ser Gly Gly Pro Ser Tyr
1           5           10           15

```

```

Ala Pro Gln Asn Ile Arg Gly Tyr Phe Ile Val Cys Val Leu Ala Leu
                20           25           30

```

```

Phe Arg Gln Glu Gln Lys Gly Leu Ala Ile Tyr Asp Arg Leu Pro Glu
          35           40           45

```

```

Lys Arg Arg Trp Leu Pro Asp Leu Pro Pro Arg Asn Gly Pro Arg Pro
50           55           60

```

```

Ile Thr Thr Ser His Ile Ile Gln Arg Gln Arg Asn Gln Ala Pro Asp
65           70           75           80

```

```

Pro Lys Phe Ala Leu Glu Glu Leu Lys Ala Thr Val Ile Pro Arg Val
          85           90           95

```

```

Gln Ala Arg His Thr Asp Leu Thr His Leu Ser Leu Ser Lys Phe Glu
          100           105           110

```

```

Phe His Ala Glu Ala Ile Phe Leu Leu Pro Ser Val Pro Ile Asp Asp

```

115 120 125

Pro Lys Asn Val Pro Ser His Asp Thr Val Arg Arg Thr Lys Arg Glu
 130 135 140

Ile Ala His Met His Asp Tyr His Asp Phe Thr Leu His Leu Ala Leu
 145 150 155 160

Ala Ala Gln Asp Gly Lys Glu Val Val Ser Lys Gly Trp Gly Gln Arg
 165 170 175

His Pro Leu Ala Gly Pro Gly Val Pro Gly Pro Pro Thr Glu Trp Thr
 180 185 190

Phe Ile Tyr Ala Pro Arg Asn Glu Glu Glu Leu Ala Val Val Glu Met
 195 200 205

Ile Ile Glu Ala Ser Ile Gly Tyr Met Thr Asn Asp Pro Ala Gly Val
 210 215 220

Val Ile Ala
 225

<210> 23
 <211> 681
 <212> DNA
 <213> artificial

<220>
 <223> Фрагмент нуклеотидной последовательности, кодирующий функциональный фрагмент люциферазы из *Armillaria mellea*

<400> 23
 tacattcgta gggactacca gacattttta tctggggggtc cctcctacgc tccccagaac 60
 atcagaggat acctcattgt ctgcgctcctg gccttggtcc gtcaggagca aaaaggcctt 120
 gcgatatacg accgccttcc cgagaagcgc aggtggctac ctgacttgcc tcctcgcgat 180
 ggcccacggc ccatcacgac cagccatata atccaaagac agcgcaacca ggcgccggac 240
 ctcaagttcg ccctcgagga actcaaggcc acggtcattc cacgggtgca ggctcgccac 300
 actgacctca cccatctcag cctatccaag ttcgagttcc atgctgaagc aatcttctg 360
 ctcccctctg taccatcga tgatccaaag aatgtgcaa gtcacgacac ggtgcgcagg 420
 acgaagaggg aaattgca tatgacgac taccatgact acacgctgca tcttgctgtg 480
 gccgccaag acgggaagga agtcgatca aaggatggg ggcagcgaca cccgctggca 540
 ggccctggcg ttctggtcc accgacggag tggacgttta tttatcgcc acgtaacgaa 600

gaggagctgg cagtgggtgga aatgattatc gaggcacoga taggctatat gaccaatgac 660
 cctgcaggaa aaactatcgc a 681

<210> 24
 <211> 227
 <212> PRT
 <213> artificial

<220>
 <223> Аминокислотная последовательность функционального фрагмента
 люциферазы из *Armillaria mellea*

<400> 24

Tyr Ile Arg Arg Asp Tyr Gln Thr Phe Leu Ser Gly Gly Pro Ser Tyr
 1 5 10 15

Ala Pro Gln Asn Ile Arg Gly Tyr Leu Ile Val Cys Val Leu Ala Leu
 20 25 30

Phe Arg Gln Glu Gln Lys Gly Leu Ala Ile Tyr Asp Arg Leu Pro Glu
 35 40 45

Lys Arg Arg Trp Leu Pro Asp Leu Pro Pro Arg Asp Gly Pro Arg Pro
 50 55 60

Ile Thr Thr Ser His Ile Ile Gln Arg Gln Arg Asn Gln Ala Pro Asp
 65 70 75 80

Leu Lys Phe Ala Leu Glu Glu Leu Lys Ala Thr Val Ile Pro Arg Val
 85 90 95

Gln Ala Arg His Thr Asp Leu Thr His Leu Ser Leu Ser Lys Phe Glu
 100 105 110

Phe His Ala Glu Ala Ile Phe Leu Leu Pro Ser Val Pro Ile Asp Asp
 115 120 125

Pro Lys Asn Val Pro Ser His Asp Thr Val Arg Arg Thr Lys Arg Glu
 130 135 140

Ile Ala His Met His Asp Tyr His Asp Tyr Thr Leu His Leu Ala Leu
 145 150 155 160

Ala Ala Gln Asp Gly Lys Glu Val Val Ser Lys Gly Trp Gly Gln Arg
 165 170 175

His Pro Leu Ala Gly Pro Gly Val Pro Gly Pro Pro Thr Glu Trp Thr

180 185 190

Phe Ile Tyr Ala Pro Arg Asn Glu Glu Leu Ala Val Val Glu Met
 195 200 205

Ile Ile Glu Ala Ser Ile Gly Tyr Met Thr Asn Asp Pro Ala Gly Lys
 210 215 220

Thr Ile Ala
 225

<210> 25
 <211> 798
 <212> DNA
 <213> artificial

<220>
 <223> Фрагмент нуклеотидной последовательности, кодирующий функциональный фрагмент люциферазы из *Armillaria ostoyae*

<400> 25
 atgtccttca tcgacagcat gaaacttgac ttcgtcggac acctcttgg catcaggaat 60
 cgcggttag ccaccgctg ttgtgctgtg gcagtcgctt ctgccatcgc ctcccttac 120
 attcgtaggg actaccagac attcttatct ggcggtccct cttacgctcc ccagaacatc 180
 aaaggatata tcatcgtctg cgtcctggcc ttgttccgtc aggagcaaaa ggccttgcg 240
 atatatgacc gccttcccga gaagcgcagg tggctacctg acttgcctcc tcgcaatggc 300
 ccgcgggcca tcacgaccag ccatataatc caaagacagc gcaaccaggc gccagactcc 360
 aagttcgccc tcgaggaact caaggctacg gtcattccac gggtgcaagg tcgccacact 420
 gacctcacc atctcagcct atccaagttc gagttccatg ctgaagcaat cttcctgctc 480
 ccctctgtac ccatcgatga tccaaaaaat gttccaagtc atgacacggt gcgcaggacg 540
 aagagggaga tcgcgcatat gcacgactac catgacttta cgttgcatct tgcaactggc 600
 gcccaagacg ggaaggaagt cgtggcgaag ggatgggggc agcgacacc gctggcaggc 660
 cctggcgctt ctggctcacc tacggagtgg acgtttatatt atgcgccacg taacgaagag 720
 gaactggcag tggtggaat gattatcgag gcatcaatag gctatatgac caatgaccct 780
 gctggaacag ttatcgta 798

<210> 26
 <211> 266
 <212> PRT
 <213> artificial

<220>
 <223> Аминокислотная последовательность функционального фрагмента

люциферазы из *Armillaria ostoyae*
 <400> 26

Met Ser Phe Ile Asp Ser Met Lys Leu Asp Phe Val Gly His Leu Phe
 1 5 10 15

Gly Ile Arg Asn Arg Gly Leu Ala Thr Ala Cys Cys Ala Val Ala Val
 20 25 30

Ala Ser Ala Ile Ala Phe Pro Tyr Ile Arg Arg Asp Tyr Gln Thr Phe
 35 40 45

Leu Ser Gly Gly Pro Ser Tyr Ala Pro Gln Asn Ile Lys Gly Tyr Leu
 50 55 60

Ile Val Cys Val Leu Ala Leu Phe Arg Gln Glu Gln Lys Gly Leu Ala
 65 70 75 80

Ile Tyr Asp Arg Leu Pro Glu Lys Arg Arg Trp Leu Pro Asp Leu Pro
 85 90 95

Pro Arg Asn Gly Pro Arg Pro Ile Thr Thr Ser His Ile Ile Gln Arg
 100 105 110

Gln Arg Asn Gln Ala Pro Asp Ser Lys Phe Ala Leu Glu Glu Leu Lys
 115 120 125

Ala Thr Val Ile Pro Arg Val Gln Ala Arg His Thr Asp Leu Thr His
 130 135 140

Leu Ser Leu Ser Lys Phe Glu Phe His Ala Glu Ala Ile Phe Leu Leu
 145 150 155 160

Pro Ser Val Pro Ile Asp Asp Pro Lys Asn Val Pro Ser His Asp Thr
 165 170 175

Val Arg Arg Thr Lys Arg Glu Ile Ala His Met His Asp Tyr His Asp
 180 185 190

Phe Thr Leu His Leu Ala Leu Ala Ala Gln Asp Gly Lys Glu Val Val
 195 200 205

Ala Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Val Pro
 210 215 220

Gly Pro Pro Thr Glu Trp Thr Phe Ile Tyr Ala Pro Arg Asn Glu Glu

Ala Pro Gln Asn Ile Arg Gly Tyr Ile Ile Val Leu Val Leu Ser Leu
 20 25 30

Phe Arg Gly Glu Glu Lys Gly Leu Ala Ile Tyr Glu Pro Leu Pro Glu
 35 40 45

Lys Arg Thr Trp Leu Pro Glu Leu Pro Arg Arg Ala Gly Asp Arg Pro
 50 55 60

Lys Thr Thr Ser His Ile Ile Gln Arg Gln Leu Asp Gln Tyr Pro Asp
 65 70 75 80

Pro Asp Phe Val Leu Lys Ala Leu Lys Ala Thr Val Ile Pro Arg Val
 85 90 95

Gln Ala Arg His Thr Asp Lys Thr His Leu Ala Leu Ser Lys Phe Glu
 100 105 110

Phe His Ala Glu Ala Ile Phe Val Arg Pro Glu Ile Ala Ile Asp Asp
 115 120 125

Pro Lys His Ile Pro Ser His Asp Thr Val Arg Arg Thr Lys Arg Glu
 130 135 140

Ile Ala His Met His Asp Tyr His Asp Cys Thr Leu His Leu Ala Leu
 145 150 155 160

Ala Ala Gln Asp Ala Lys Gln Val Leu Gln Lys Gly Trp Gly Gln Arg
 165 170 175

His Pro Leu Ala Gly Pro Gly Met Pro Gly Pro Pro Thr Glu Trp Thr
 180 185 190

Phe Leu Tyr Ala Pro Arg Thr Glu Glu Glu Val Lys Val Val Glu Thr
 195 200 205

Ile Val Glu Ala Ser Ile Ala Tyr Met Thr Asn Ala Glu Lys Pro Val
 210 215 220

Glu Leu Val Gln
 225

<210> 29
 <211> 666
 <212> DNA
 <213> artificial

<220>
 <223> Фрагмент нуклеотидной последовательности, кодирующий функциональный фрагмент люциферазы из *Musca citricolor*

<400> 29
 ttcacccaaga aagactacga gacgttcctg aagggcgcc cctcctatgc gccccaaaac 60
 gttcgcggat acatcatcgt gctcgtgctc gcgctcttcc gccaaagagca gctcgggctg 120
 gagatctacg accgcatgcc cgagaaacgt cgctggctcg cgaatctccc tcagcgcgag 180
 ggcccccgcc ccaagaccac aagtcacatc atccagcggc agctcagcca gcacacggac 240
 cccgcattcg gcgcccgcga cctcaaagac accgtcattc gcgcgctcca ggcgcgccac 300
 gcagccaaca cgcacatcgc gcgctcgacg ttcgagttcc acgcccgcgc gatcttctctg 360
 aacgcggacg tgccgctgcc cgagggcctg cccgcaagcg agacggtgcg gcggaccaag 420
 ggcgagatcg cgcacatgca cgactaccac gacttcacgc tgcacctcgc gctcgcagca 480
 gcggatggga aggaggtggt cggcaagggc tgggggcagc gccatccgct ggcgggaccc 540
 ggcgtgccgg gtccgccgaa cgagtggacc tttgtgtatg cgccgaggaa tgaagaggag 600
 atgggcgtgg tcgagcagat cgtagaggcg gcgattgggt acatgtcgaa cgtgcctgcg 660
 ctggaa 666

<210> 30
 <211> 222
 <212> PRT
 <213> artificial

<220>
 <223> Аминокислотная последовательность функционального фрагмента люциферазы из *Musca citricolor*

<400> 30

Phe Ile Lys Lys Asp Tyr Glu Thr Phe Leu Lys Gly Gly Pro Ser Tyr
 1 5 10 15

Ala Pro Gln Asn Val Arg Gly Tyr Ile Ile Val Leu Val Leu Ala Leu
 20 25 30

Phe Arg Gln Glu Gln Leu Gly Leu Glu Ile Tyr Asp Arg Met Pro Glu
 35 40 45

Lys Arg Arg Trp Leu Ala Asn Leu Pro Gln Arg Glu Gly Pro Arg Pro
 50 55 60

Lys Thr Thr Ser His Ile Ile Gln Arg Gln Leu Ser Gln His Thr Asp
 65 70 75 80

Pro Ala Phe Gly Ala Ala Tyr Leu Lys Asp Thr Val Ile Pro Arg Val
85 90 95

Gln Ala Arg His Ala Ala Asn Thr His Ile Ala Arg Ser Thr Phe Glu
100 105 110

Phe His Ala Ala Ala Ile Phe Leu Asn Ala Asp Val Pro Leu Pro Glu
115 120 125

Gly Leu Pro Ala Ser Glu Thr Val Arg Arg Thr Lys Gly Glu Ile Ala
130 135 140

His Met His Asp Tyr His Asp Phe Thr Leu His Leu Ala Leu Ala Ala
145 150 155 160

Ala Asp Gly Lys Glu Val Val Gly Lys Gly Trp Gly Gln Arg His Pro
165 170 175

Leu Ala Gly Pro Gly Val Pro Gly Pro Pro Asn Glu Trp Thr Phe Val
180 185 190

Tyr Ala Pro Arg Asn Glu Glu Glu Met Gly Val Val Glu Gln Ile Val
195 200 205

Glu Ala Ala Ile Gly Tyr Met Ser Asn Val Pro Ala Leu Glu
210 215 220

<210> 31
<211> 705
<212> DNA
<213> artificial

<220>
<223> Фрагмент нуклеотидной последовательности, кодирующий функциональный фрагмент люциферазы из *Omphalotus olearius*

<400> 31
ttcatatcgca aagattacca gactttcctg gaggtgggac cctcgtacgc cccgcagaac 60
ctccaaggat acatcatcgt ctgtgtactc tctctgttcc ggcaagaaca gaaagacgta 120
gcgatttatg atcgccttcc tgagaaaagg aggtggttag gagacctccc gtttcgagag 180
gggccaagac cgagtatcac tagccatata atccagcgcac agcgcacca attggctgac 240
gccgagttcg ctaccaaaga gctgataggc aaaatcatcc ctgcggtcca agcccgacac 300
accaacacaa cattcctcag cacatctaaa ttcgaattcc acgcccaggc catcttcctt 360
ttgccctcta tccaatcaa cgaccctcaa aacattccaa gccacgatac cgttcgtcgc 420
acgaaacgcg agatcgcgca tatgcatgat tatcacgact gtacggttga tctcgcactt 480

gctgctcaag atggaagga ggttttagag aaaggatggg gtcagcgaca tcctcttgct 540
 ggacctgggtg ttcttgccc gccgacggag tggacgtttc tttatgcacc gcgcagcgaa 600
 gaggaggttc gggttgtgga gatgattggt gaggcacag ttgtgtatat gacgaatgat 660
 cctgctgata aaatcgtaga agctactgtg cagggctactg aagaa 705

<210> 32
 <211> 235
 <212> PRT
 <213> artificial

<220>
 <223> Аминокислотная последовательность функционального фрагмента
 люциферазы из *Opthalotus olearius*

<400> 32

Phe Ile Arg Lys Asp Tyr Gln Thr Phe Leu Glu Val Gly Pro Ser Tyr
 1 5 10 15

Ala Pro Gln Asn Leu Gln Gly Tyr Ile Ile Val Cys Val Leu Ser Leu
 20 25 30

Phe Arg Gln Glu Gln Lys Asp Val Ala Ile Tyr Asp Arg Leu Pro Glu
 35 40 45

Lys Arg Arg Trp Leu Gly Asp Leu Pro Phe Arg Glu Gly Pro Arg Pro
 50 55 60

Ser Ile Thr Ser His Ile Ile Gln Arg Gln Arg Thr Gln Leu Ala Asp
 65 70 75 80

Ala Glu Phe Ala Thr Lys Glu Leu Ile Gly Lys Ile Ile Pro Arg Val
 85 90 95

Gln Ala Arg His Thr Asn Thr Thr Phe Leu Ser Thr Ser Lys Phe Glu
 100 105 110

Phe His Ala Gln Ala Ile Phe Leu Leu Pro Ser Ile Pro Ile Asn Asp
 115 120 125

Pro Gln Asn Ile Pro Ser His Asp Thr Val Arg Arg Thr Lys Arg Glu
 130 135 140

Ile Ala His Met His Asp Tyr His Asp Cys Thr Leu His Leu Ala Leu
 145 150 155 160

Ala Ala Gln Asp Gly Lys Glu Val Leu Glu Lys Gly Trp Gly Gln Arg
 165 170 175

His Pro Leu Ala Gly Pro Gly Val Pro Gly Pro Pro Thr Glu Trp Thr
 180 185 190

Phe Leu Tyr Ala Pro Arg Ser Glu Glu Glu Val Arg Val Val Glu Met
 195 200 205

Ile Val Glu Ala Ser Val Val Tyr Met Thr Asn Asp Pro Ala Asp Lys
 210 215 220

Ile Val Glu Ala Thr Val Gln Gly Thr Glu Glu
 225 230 235

<210> 33
 <211> 681
 <212> DNA
 <213> artificial

<220>
 <223> Фрагмент нуклеотидной последовательности, кодирующий
 функциональный фрагмент люциферазы из *Panellus stipticus*

<400> 33
 ttcgttcgta aagactatca cacctttctt gaaggtggac catcctacgc gccgcagaat 60
 ttgcaaggct atatcatcgt gttggtgctc tcaactcttc gaggggagga gacgggattg 120
 gaaatatacg accgcttgcc cgaaaaacgc cgctggctcg aggagctgcc tgttcgcgaa 180
 ggcccgcgcc caaagacaac cagccacatc attcagagac agttgaatca gcacgttgac 240
 ccggacttcg gaatgaactc tttgaaaggc tccgctcatcc ggcgccttca atcccgccac 300
 caggacataa ctcaactcgc actctcgaaa ttcgaattcc acgccgaggc catatctctg 360
 cgccccgatg tcgcatcaaa cgatcccaaa cacgtcccga gccacgacac ggtgcgcccgc 420
 acaaagcgcg agatagctca catgcacgac taccatgatt acacgtgtca tttggcgctc 480
 gcagcgcagg atgggaagca agtgattgca aaaggggtggg gccagagaca tccgctcgcc 540
 ggaccgggca tgccggggcc gccgacggag tggacatctt tgtatgcgcc gaggaatgag 600
 gcggagggtc aagtgttga gacgattatc gaagcgtcaa tcgggtacat gtcgaacgca 660
 ccagccttgg gtgggagcga g 681

<210> 34
 <211> 227
 <212> PRT
 <213> artificial
 <220>

<223> Аминокислотная последовательность функционального фрагмента люциферазы из *Panellus stipticus*

<400> 34

Phe Val Arg Lys Asp Tyr His Thr Phe Leu Glu Gly Gly Pro Ser Tyr
1 5 10 15

Ala Pro Gln Asn Leu Gln Gly Tyr Ile Ile Val Leu Val Leu Ser Leu
20 25 30

Phe Arg Gly Glu Glu Thr Gly Leu Glu Ile Tyr Asp Arg Leu Pro Glu
35 40 45

Lys Arg Arg Trp Leu Glu Glu Leu Pro Val Arg Glu Gly Pro Arg Pro
50 55 60

Lys Thr Thr Ser His Ile Ile Gln Arg Gln Leu Asn Gln His Val Asp
65 70 75 80

Pro Asp Phe Gly Met Asn Ser Leu Lys Gly Ser Val Ile Arg Arg Leu
85 90 95

Gln Ser Arg His Gln Asp Ile Thr Gln Leu Ala Leu Ser Lys Phe Glu
100 105 110

Phe His Ala Glu Ala Ile Phe Leu Arg Pro Asp Val Ala Ile Asn Asp
115 120 125

Pro Lys His Val Pro Ser His Asp Thr Val Arg Arg Thr Lys Arg Glu
130 135 140

Ile Ala His Met His Asp Tyr His Asp Tyr Thr Cys His Leu Ala Leu
145 150 155 160

Ala Ala Gln Asp Gly Lys Gln Val Ile Ala Lys Gly Trp Gly Gln Arg
165 170 175

His Pro Leu Ala Gly Pro Gly Met Pro Gly Pro Pro Thr Glu Trp Thr
180 185 190

Phe Leu Tyr Ala Pro Arg Asn Glu Ala Glu Val Gln Val Leu Glu Thr
195 200 205

Ile Ile Glu Ala Ser Ile Gly Tyr Met Ser Asn Ala Pro Ala Leu Gly
210 215 220

30

Gly Ser Glu
225

<210> 35
<211> 211
<212> PRT
<213> artificial

<220>
<223> Консенсусная последовательность

<220>
<221> MISC_FEATURE
<222> (3)..(214)
<223> X - любая аминокислота

<400> 35

Asp Tyr Xaa Thr Phe Leu Xaa Xaa Gly Pro Ser Tyr Ala Pro Gln Asn
1 5 10 15

Xaa Xaa Gly Tyr Xaa Ile Val Xaa Val Leu Xaa Leu Phe Arg Xaa Glu
20 25 30

Xaa Xaa Xaa Xaa Ile Tyr Xaa Xaa Xaa Pro Glu Lys Arg Xaa Trp Leu
35 40 45

Xaa Xaa Leu Pro Xaa Arg Xaa Gly Xaa Arg Pro Xaa Xaa Thr Ser His
50 55 60

Ile Ile Gln Arg Gln Xaa Xaa Gln Xaa Xaa Asp Xaa Xaa Phe Xaa Xaa
65 70 75 80

Xaa Xaa Leu Xaa Xaa Xaa Xaa Ile Xaa Arg Xaa Gln Xaa Arg His Xaa
85 90 95

Xaa Xaa Thr Xaa Xaa Xaa Ser Xaa Phe Glu Phe His Ala Xaa Ala Ile
100 105 110

Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Pro Xaa Xaa
115 120 125

Xaa Thr Val Arg Arg Thr Lys Xaa Glu Ile Ala His Met His Asp Tyr
130 135 140

His Asp Xaa Thr Xaa His Leu Ala Leu Ala Ala Xaa Asp Xaa Lys Val
145 150 155 160

Xaa Xaa Lys Gly Trp Gly Gln Arg His Pro Leu Ala Gly Pro Gly Xaa

31

165

170

175

Pro Gly Pro Pro Xaa Glu Trp Thr Phe Xaa Tyr Ala Pro Arg Xaa Glu
 180 185 190

Xaa Glu Xaa Xaa Val Xaa Glu Xaa Ile Xaa Glu Ala Xaa Xaa Xaa Tyr
 195 200 205

Met Xaa Asn
 210

<210> 36
 <211> 38
 <212> DNA
 <213> artificial

<220>
 <223> Праймер 1711

<220>
 <221> misc_feature
 <222> (6)..(36)
 <223> R - A или G, Y - C или T, S- G или C, W - A или T, H - A или или
 T, N - любое основание

<400> 36
 atcgcnttys cntwcahcc narrgaytac saracntt 38

<210> 37
 <211> 35
 <212> DNA
 <213> artificial

<220>
 <223> Праймер 1712

<220>
 <221> misc_feature
 <222> (6)..(3)
 <223> R - A или G, Y - C или T, S- G или C, W - A или T, H - A или или
 T, N - любое основание

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n is a, c, g, t or u

<220>
 <221> misc_feature

<222> (21)..(21)
 <223> n is a, c, g, t or u

 <400> 37
 atcgcnttgc cntwcathcc narrgaytac sarac 35

 <210> 38
 <211> 32
 <212> DNA
 <213> artificial

 <220>
 <223> Праймер 1713

 <220>
 <221> misc_feature
 <222> (6)..(31)
 <223> R - А или G, Y - С или Т, S- G или C, W - А или Т, Н - А или или
 Т, N - любое основание

 <400> 38
 atcgcnttgc cntwcathcc narrgaytac sa 32

 <210> 39
 <211> 29
 <212> DNA
 <213> artificial

 <220>
 <223> Праймер 1714

 <220>
 <221> misc_feature
 <222> (6)..(27)
 <223> R - А или G, Y - С или Т, S- G или C, W - А или Т, Н - А или или
 Т, N - любое основание

 <400> 39
 atcgcnttgc cntwcathcc narrgayta 29

 <210> 40
 <211> 42
 <212> DNA
 <213> artificial

 <220>
 <223> Праймер 1715

 <220>
 <221> misc_feature
 <222> (1)..(40)
 <223> R - А или G, Y - С или Т, S- G или C, W - А или Т, Н - А или или
 Т, N - любое основание

 <400> 40

ngcagggncn ttsgtcatat ascncatnga ngcytcnayn at 42

<210> 41
 <211> 36
 <212> DNA
 <213> artificial

<220>
 <223> Праймер 1716

<220>
 <221> misc_feature
 <222> (1)..(34)
 <223> R - А или G, Y - С или Т, S- G или С, W - А или Т, Н - А или или
 Т, N - любое основание

<400> 41
 ngcagggncn ttsgtcatat ascncatnga ngcytc 36

<210> 42
 <211> 32
 <212> DNA
 <213> artificial

<220>
 <223> Праймер 1717

<220>
 <221> misc_feature
 <222> (7)..(30)
 <223> R - А или G, Y - С или Т, S- G или С, W - А или Т, Н - А или или
 Т, N - любое основание

<400> 42
 gcagggncnt tsgtcatata scncatngan gc 32

<210> 43
 <211> 30
 <212> DNA
 <213> artificial

<220>
 <223> Праймер 1718

<220>
 <221> misc_feature
 <222> (1)..(28)
 <223> R - А или G, Y - С или Т, S- G или С, W - А или Т, Н - А или или
 Т, N - любое основание

<400> 43
 ngcagggncn ttsgtcatat ascncatnga 30

<210> 44

<211> 801
 <212> DNA
 <213> artificial

 <220>
 <223> Гуманизированная нуклеотидная последовательность для люциферазы
 из *Neonothopanus nambi*

 <400> 44
 atgcgatta acatctccct ttcattcttt ttcgagcgat tgagcaaac gagttccagg 60
 agtattgcaa tcacttggtg ggttgctctc gcgagcgcca tcgcattccc catcatccgg 120
 agagattatc agacgtttct tgagggtggc cctagctatg caccacagaa cttccgagga 180
 tataatcatc tgtgtgtact gtcactgttt aggcaagaac aaaagggatt ggctatctat 240
 gataggttgc ctgagaaacg gcggtggctc gctgatctcc catttagaga ggggacacga 300
 ccgagcatca cttcacacat catacaaaga cagcgaacgc agctcgttga ccaagagttc 360
 gcaactaggg aactgattga taagggtgata cccagagtac aggcgcgaca caccgataaa 420
 acttttcttt ccacctctaa attcgagttc catgccaaag ctattttctt gttgccttcc 480
 ataccgatta atgatcctct gaattattcca tcccacgaca cagttcgacg gacgaaacgc 540
 gaaattgcbc acatgcacga ctatcacgat tgcactttgc acctggcact ggctgctcaa 600
 gacgaaaaag aagtctgaa aaagggttg gggcaaacgc atccgctggc gggacccggg 660
 gtacctgggc cgcctacgga atggacattt ttgtacgcac cgaggaacga agaggaggcc 720
 agggctcgtt agatgattgt tgaggctagt attgggtaca tgacgaatga tccggctggt 780
 aaaattggtg aaaatgcaaa g 801

 <210> 45
 <211> 798
 <212> DNA
 <213> artificial

 <220>
 <223> Гуманизированная нуклеотидная последовательность для люциферазы
 из *Armillaria gallica*

 <400> 45
 atgagcttta tagactcaat gaaacttgat cttgtaggac acctcttcgg catacgsaac 60
 agaggactcg cagcagcctg ttgcgccctt gcagtagctt ccacgatagc tttcccttat 120
 atccgccggg attaccaaac gttccttagc ggtggccctt cttacgcgcc acagaacata 180
 cgaggctatt ttatcgtgtg tgtgctcgct ctgtttcgcc aagaacagaa agggctggcg 240
 atttatgacc gacttcctga gaaacgaagg tggcttctg accttcctcc ccggaatggc 300
 cctcgcccga ttacgacctc tcatattata caaagacaga ggaaccaagc gccagatccg 360
 aaatttgccc tggaggagct caaggccact gtgattccac gagtccaagc ccgccatacg 420

gatcttacac atctcagtct tagcaaattt gagtttcatg ccgaggcaat ttttcttttg 480
 ccgtccgttc cgatagacga cccccaaaat gtaccctcac acgatacggc caggcgacaca 540
 aagagggaga tcgcccatac gcatgactat cacgacttta ccctgcatct cgctctggcg 600
 gcacaagacg gcaaggaggt agtcagcaaa ggctggggcc aaaggcatcc gctcgcgggc 660
 ccaggggttc cggggccacc tacagagtgg actttcatct atgcgccgag aaacgaggaa 720
 gaactcgcag tagttgagat gatcatagag gcatcaattg gctacatgac gaacgatcct 780
 gccggagtcg ttatagcc 798

<210> 46
 <211> 798
 <212> DNA
 <213> artificial

<220>
 <223> Гуманизированная нуклеотидная последовательность для люциферазы
 из *Armillaria mellea*

<400> 46
 atgagttttt ttgactctgt caaacttgat ctggtgggac ggctcttcgg catccgaaac 60
 aggggacttg cggttacttg ctgcgcggta gcggtcgcct ctattatcgc cttcccatac 120
 attcgccgag actatcagac cttcctgtca ggcggacctt cttatgcacc gcagaatata 180
 aggggctatc ttattgtttg tgtactggct ctctttcggc aagaacaaaa aggtctggct 240
 atctacgata gacttcctga gaaacgacgc tggctgccag atcttcccc cgcgcacggc 300
 ccaaggccca taacaacaag tcatattata cagagacaac ggaaccaggc cccagacctc 360
 aagtttgccc tcgaggagct gaaagcgact gttataccta gactacaggc gagacataca 420
 gacttgacc atttgagcct gagcaagttt gaattccatg ccgaagcaat ctttcttctg 480
 ccttccgtac ccattgacga tccaaaaaac gtcccctcac atgatactgt aagacgcacg 540
 aagcgagaga tcgcacacat gcacgactac catgactaca ctctgcacct tgccttgga 600
 gctcaggatg gcaaggaagt tgtatccaag ggggtggggc aaagacacc cttggctggc 660
 ccgggagtg ctggaccccc gaccgaatgg acgtttattt acgcaccaag gaacgaggaa 720
 gaattggccg tcgttgaaat gattattgaa gcgtccatag gatatatgac caacgatccg 780
 gccggaaaaga caatcgcg 798

<210> 47
 <211> 798
 <212> DNA
 <213> artificial

<220>

<223> Гуманизированная нуклеотидная последовательность для люциферазы из *Armillaria ostoyae*

<400> 47
atgtccttta ttgacagtat gaagccttgat ttcgtaggtc acttgtttgg tatacgaaac 60
cgaggtctcg ctactgcgtg ctgcgcggtg gcagtagcga gtgctatagc attcccgtat 120
atccgccgag attaccagac atttcttagc ggtggaccta gctatgcacc ccagaatatt 180
aaaggttatt tgatcgtttg tgtgcttgca ctgtttcgcc aggagcagaa gggctctgct 240
atttacgacc gactgcctga gaaaaggcga tggttgcctg accttccacc cagaaacggc 300
cccagggcca tcacgacctc acacataatc caacggcaga ggaaccaggc tccagattca 360
aagtttgccg ttgaggaact caaggctact gtaataccta gagtccaggc cagacatacg 420
gatctcacc atctcagcct cagtaagttt gaggttcatg cggaagctat tttcctgctc 480
ccgagcgtac ccattgatga cccaagaat gtcacctcac atgatacggc acgacgcacg 540
aagcgggaga tagctcacat gcacgactac catgatttca ctctgcatct cgccttgcca 600
gcgcaggatg gcaaggaggt agtagcgaag gtagggggc agagacacc cgtggcaggc 660
cctggggttc ccggccctcc gaccgaatgg acatttatct atgcccccg caacgaagag 720
gagctcgcag ttgttgagat gatcattgaa gcatccattg ggtacatgac aaatgaccca 780
gccggtacag tcattggt 798

<210> 48
<211> 753
<212> DNA
<213> artificial

<220>
<223> Гуманизированная нуклеотидная последовательность для люциферазы из *Muscena chlorophos*

<400> 48
atggtacagc tcacgcgaac ctcagggttt attgcagcgg ctgcgattgt tgccgccatc 60
gcgttcccgt ttattcgacg ggactatcag acttttctgc gagggggccc ctcatatgcc 120
ccacagaata ttcgcggtta tatcattgtc ctggctctgt ccctcttcag aggagaagaa 180
aaggttcttg ctatctatga gccgcttcca gagaaacgga cgtggttgcc agagttgcct 240
agacgggagg gcgatagacc caaaactacc agccatataa tacaacgaca actcgaccaa 300
tatcccagcc cggacttctg acttaaagca ttgaaagcaa ctgtaatacc gagagttcag 360
gcaaggcaca ctgataaaac ccattctgca ctttccaagt ttgaattcca cgccgaggca 420
atcttcgtcc gaccagagat agccatagac gatccgaaac atatccaag tcatgacaca 480
gtcagacgaa cgaaacggga gatagcccat atgcatgatt atcatgactg caccctgcac 540

cttgcgcttg cggcacagga cgcaaaacag gtcttgcaaa aaggatgggg acagagacac 600
ccccttgccg gacctggcat gccgggtccg cctactgaat ggacattcct ctacgccccg 660
agaactgagg aggaggtgaa agtggttgaa accatagtgg aggcgagtat agcttacatg 720
accaatgcag aaaaaccagt agaactcgta caa 753

<210> 49
<211> 738
<212> DNA
<213> artificial

<220>
<223> Гуманизированная нуклеотидная последовательность для люциферазы из *Muscena citricolor*

<400> 49
atggcttatac agttgacatg gattcagact ctgtatttgg gtgctctcgt ggctatggct 60
gttgcatttc cattcatcaa aaaagattat gagaccttcc tcaagggcgg cccttcatac 120
gccccccaga atgtccgcgg ttatattatc gtactcgtcc tcgctgtgtt cagacaagaa 180
caattggggc ttgagatata tgaccgcatg cccgaaaaac gcaggtggct cgctaacctc 240
cctcaacggg aagggtccgcg accgaaaacc actagccata tcattcagcg ccagctctct 300
cagcataacc accccgcctt tggagcggct tatctgaaag acaccgtaat cccgcgagtg 360
caggcgcggc atgcagcga cactcacata gcgcgaagta cgttcgaatt tcacgcagcg 420
gctatctttc tcaacgcaga cgtgccgttg cccgaaggac tgccagctag cgaaacagtg 480
agacgaacca aaggagaaat cgcgcacatg cacgattacc acgattttac gctgcacctg 540
gcacttgccg cagcggacgg gaaagaggtg gtcggcaaa gatggggaca gagacatcca 600
ctcgtggcc ctggggtgcc tggccctccg aatgagtgga cattcgtgta cgctcccagg 660
aacgaggaag agatgggcgt agtagagcag attgtagagg ctgcaatcgg ttacatgtca 720
aatgtccctg ctctcgag 738

<210> 50
<211> 783
<212> DNA
<213> artificial

<220>
<223> Гуманизированная нуклеотидная последовательность для люциферазы из *Omphalotus olearius*

<400> 50
atgctcccgg ctttcatata taaaccgaga ttggctatta cgtgcgtttt tgttctcgct 60
agtgcattgg cttttccctt cataagaaag gactatcaga ctttctcga agtcgggccg 120
tcttatgctc ctcaaacct tcaagggtat atcatagtct gcgtattgtc actcttcaga 180

caagagcaaa aagatgtagc catatacgat agactccccg agaaaagaag gtggcttggt 240
gacctgccgt tcagagaagg gccgcgcccg tctattacct cccacattat ccaacgcaa 300
agaacgcaac ttgctgatgc agagttcgca acaaggagt tgattgggaa gatcattccg 360
agggtgcagg caaggcatac aaataccact ttcttgcaaa cgagtaaatt tgagtttcac 420
gcccaggcaa tctttctgct tccttcaatt cccataaacg atccacaaaa tatccaagt 480
cacgatacgg tacggcgcac gaagcgcgag atcgcccaca tgcatgatta tcatgattgc 540
acgcttcaact tggctctcgc cgcgcaggac gggaaagagg tgcttgagaa ggggtggggt 600
cagcggcatc cactcgcccg gccagggtgt cccgggtccc caaccgaatg gacatttctc 660
tatgcgcccc gctctgaaga agaggtacgg gtcgtgaaa tgatagtcga ggctagcgtg 720
gtttatatga cgaatgatcc tgcggacaag attgtagagg ccaccgtcca gggcaccgaa 780
gag 783

<210> 51
<211> 747
<212> DNA
<213> artificial

<220>
<223> Гуманизированная нуклеотидная последовательность для люциферазы
из *Panellus stipticus*

<400> 51
atgaatataa atctgaaagc acttattggt gtctgtgagg tactcatcac cgctgctgtc 60
tttcttttg tacggaaga cttacacacc tttttggagg gcggcccttc ctatgctcca 120
cagaacttgc agggctacat aatagtcttc gtcttgcac tcttcagggg agaggagact 180
ggtttgaaa tttatgatcg cttgccagaa aagcggcgat ggctggaaga actgccagtg 240
cgggagggac cgcggcctaa gactacgagt cacatcatcc aacgccagct caaccagcac 300
gttgaccccg acttcggcat gaatagcttg aaaggctctg tgatacggag gcttcaatcc 360
cgccaccaag atataacgca gcttgctctt tccaagtttg agtttcacgc cgaagcaatt 420
tttctccggc ctgacgtcgc aatcaatgac cccaaacacg tgccatctca tgacactgtc 480
cgacgcacaa agaggagat cgctcacatg cacgactacc acgattacac ttgccatttg 540
gcccttgccg cccaagatgg caaacaagtg atagcgaagg gctgggggtca aagacatccg 600
cttgctgggc ccggtatgcc tggacctcct accgagtgga ccttccttta tgccccacgc 660
aacgaagcag aggtgcaagt gctggaaact attattgagg caagtatagg atatatgagt 720
aacgcacctg ccctgggagg ttccgaa 747

<210> 52
 <211> 60
 <212> DNA
 <213> artificial

<220>
 <223> Нуклеотидная последовательность, кодирующая сайт инициации трансляции и гистидиновую метку

<400> 52
 atggctagca tgactggtgg acagcaaatg ggtcgcggat cccaccacca ccaccaccac 60

<210> 53
 <211> 20
 <212> PRT
 <213> artificial

<220>
 <223> N-концевая аминокислотная последовательность, содержащая сайт инициации трансляции и гистидиновую метку

<400> 53
 Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser His His
 1 5 10 15

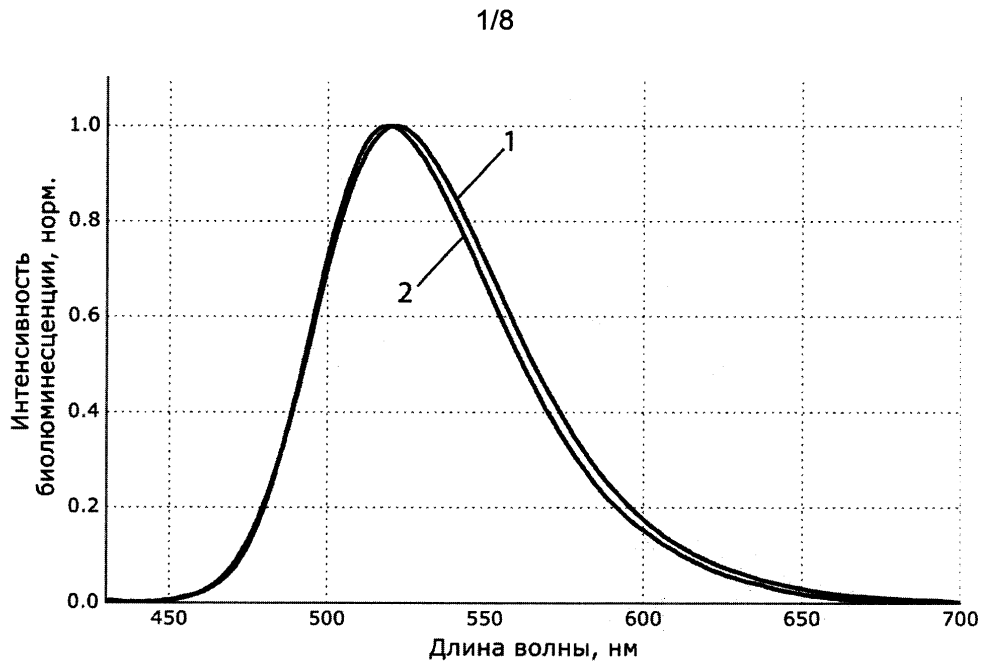
His His His His
 20

<210> 54
 <211> 801
 <212> DNA
 <213> artificial

<220>
 <223> Нуклеотидная последовательность для люциферазы из *Neonothopanus nambi*, оптимизированная для экспрессии в *Physcomitrella patens*

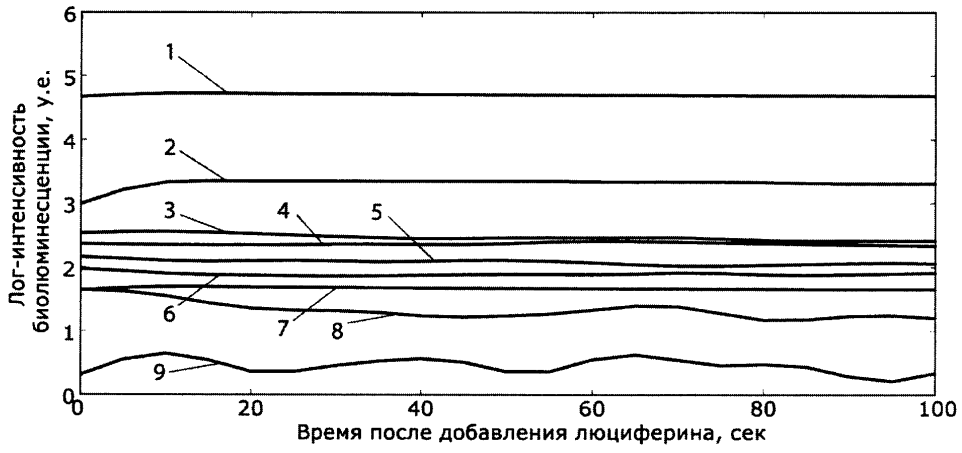
<400> 54
 atgaggataa atatctcttt gtctagtctc tttgagagac tgagcaaat gtcacccga 60
 tccatcgcaa ttacctgagg cgtagtcttg gcaagtgcaa tagcattccc aattatccga 120
 agagactatc aaacgttctt tgaagtcggt ccgagctatg cccacagaa cttccgaggc 180
 tataatcatcg tttgtgtctt gtcacttttt aggcaagagc agaaagggtt ggcaatctat 240
 gacaggcttc cagaaaagag gcggttgctt gccgacttgc cgtttcgtga agggacgagg 300
 ccattccataa cctcccat tatacagcgt cagcgtactc agctggtaga tcaagagttt 360
 gcaaccagag aacttatcga taaggtgatc ccacgtgtgc aagcaagaca cacagacaaa 420
 acttttctca gtacgtcaaa atttgagttt catgcaaagg ccatcttctt tctcccctct 480
 atccctatta atgatcctct taacattccc tcacacgata cggtaagaag aaccaagcgt 540
 gagatcgctc acatgcatga ttaccatgat tgactctgc acttggtctt tgctgctcag 600

gatggtaagg aagttttgaa gaaggggtgg ggccagcgtc acccactggc cggaccagga	660
gtcccaggcc ctctactga gtggaccttc ctttacgcac caaggaacga agaggaggcc	720
agagttgtcg agatgatagt cgaagctagt attggctaca tgactaacga tcctgctggt	780
aagattgtcg aaaatgccaa g	801

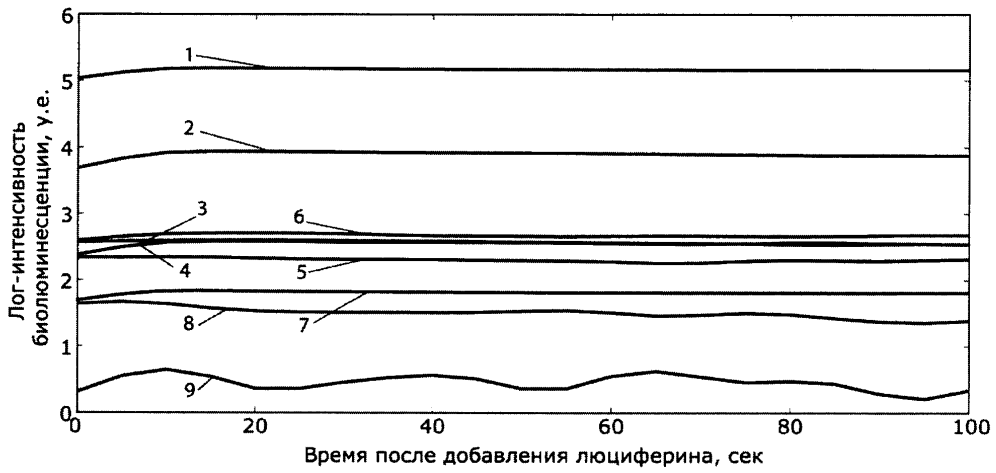


Фиг. 1

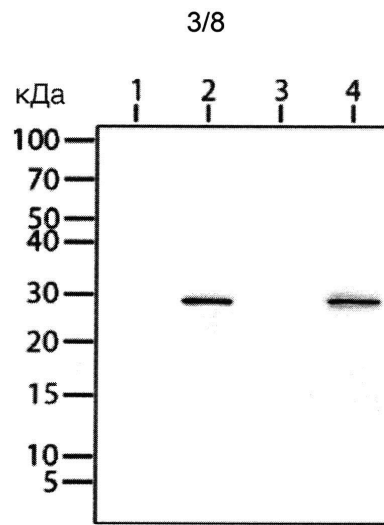
2/8



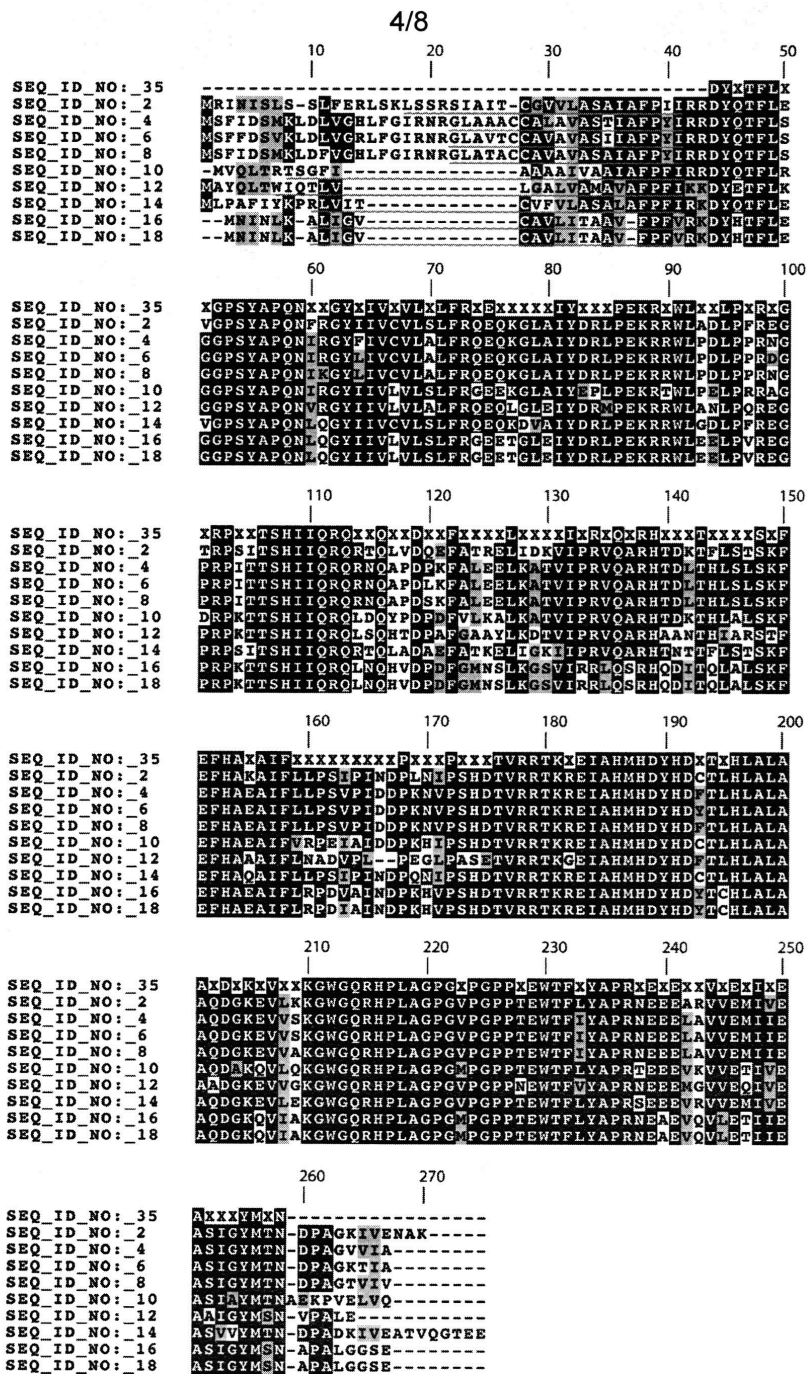
Фиг. 2



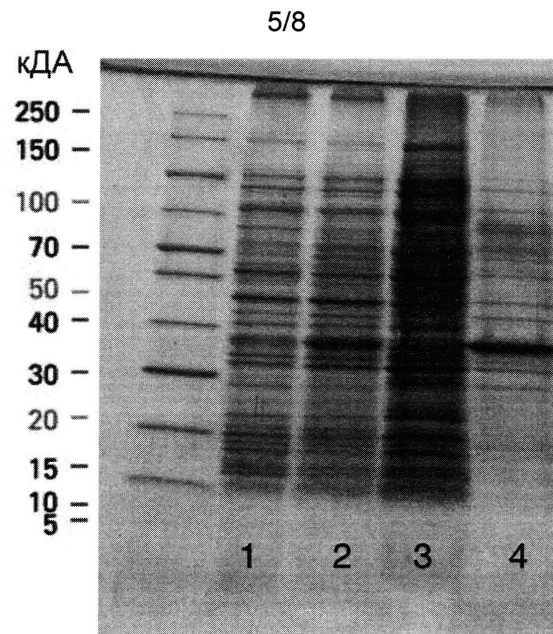
Фиг. 3



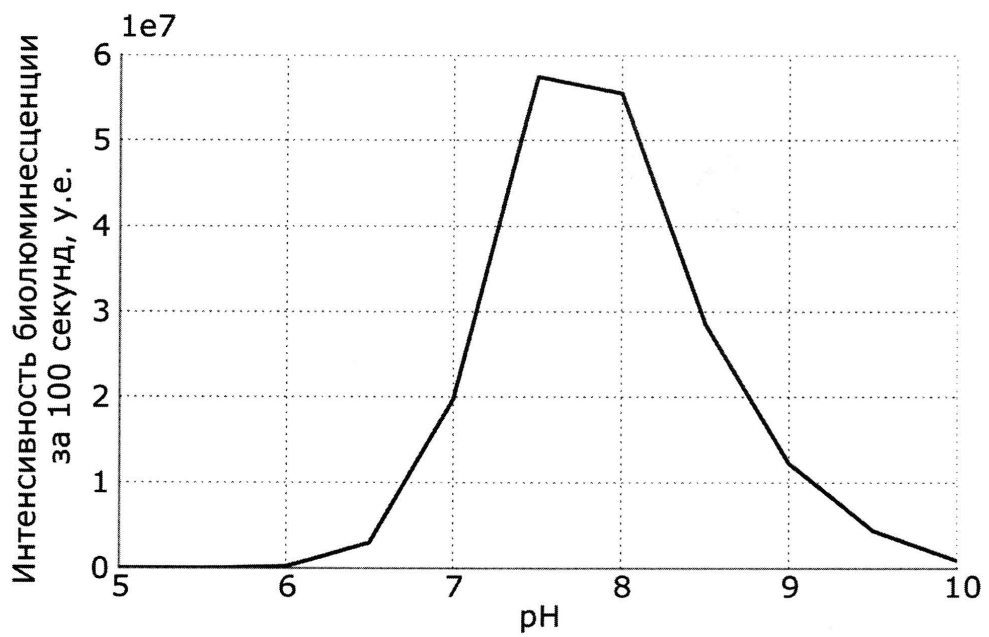
Фиг. 4



Фиг. 5

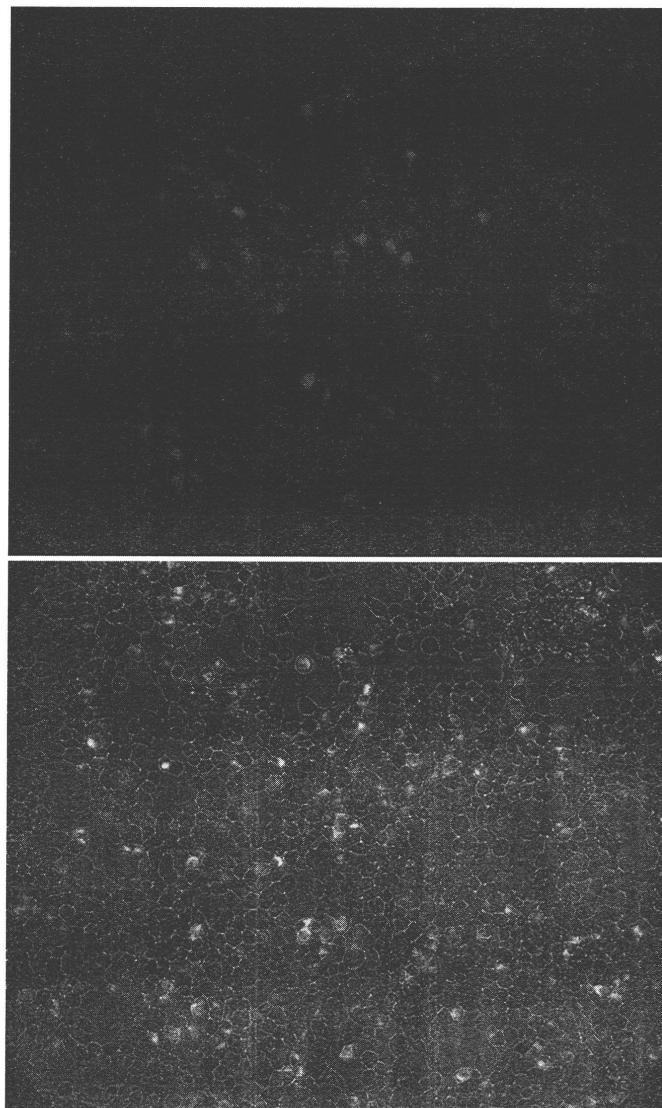


Фиг. 6



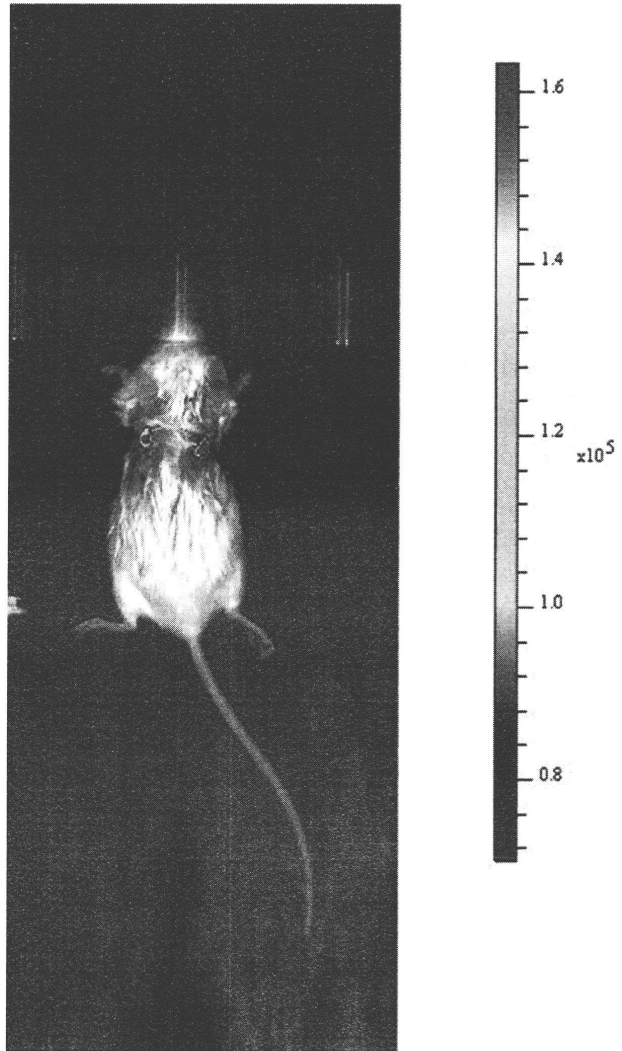
Фиг. 7

6/8



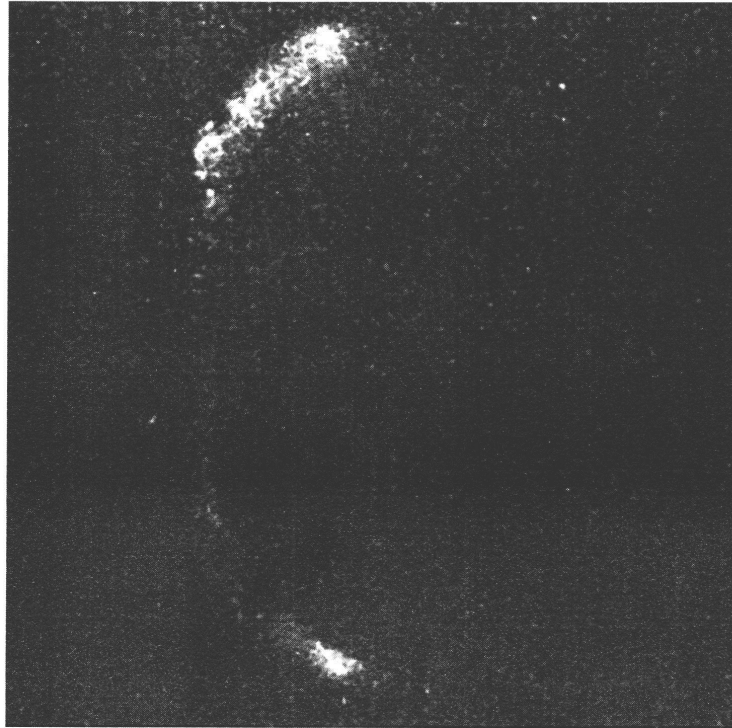
Фиг. 8

7/8



Фиг. 9

8/8



Фиг. 10