

(19)



**Евразийское
патентное
ведомство**

(11) **046157**(13) **B1**(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(45) Дата публикации и выдачи патента

2024.02.12

(21) Номер заявки

202193162

(22) Дата подачи заявки

2020.05.29

(51) Int. Cl. *C12N 15/79* (2006.01)*A61P 25/08* (2006.01)*A61P 25/16* (2006.01)*A61P 25/28* (2006.01)**(54) КОМПОЗИЦИИ И СПОСОБЫ СЕЛЕКТИВНОЙ РЕГУЛЯЦИИ ЭКСПРЕССИИ ГЕНОВ**

(31) 62/854,238; 62/857,727; 63/008,569

(32) 2019.05.29; 2019.06.05; 2020.04.10

(33) US

(43) 2022.04.27

(86) PCT/US2020/035431

(87) WO 2020/243651 2020.12.03

(71)(73) Заявитель и патентовладелец:

**ИНКОУДИД ТЕРАПЬЮТИКС, ИНК.
(US)**

(72) Изобретатель:

**Тальятеля Стефани, Таненхаус Энн,
Рамамуртхи Картик, Янг Эндрю,
Оберкофлер Дэвид (US)**

(74) Представитель:

Медведев В.Н. (RU)

(56) WO-A2-2017180915

WO-A1-2018049079

JP-A-2018088888

YOUNG A. et al., "A GABA-Selective AAV Vector-Based Approach to Up-Regulate Endogenous Scn1a Expression Reverses Key Phenotypes in a Mouse Model of Dravet Syndrome", American Society of Gene & Cell Therapy 22nd Annual Meeting, Molecular Therapy. (2019, April), vol. 27, no. 4S1, page 420, abstract 915

WO-A1-2018187363

US-A1-2015071903

COLASANTE G. et al., "dCas9-Based Scn1a Gene Activation Restores Inhibitory Interneuron Excitability and Attenuates Seizures in Dravet Syndrome Mice", Molecular Therapy. (January 2020), vol. 28, no.1, pages 235-253. Entire document

YOUNG A. et al., "A GABA-selective AAV vector upregulate endogenous SCN1A expression and reverses multiple phenotypes in a mouse model of Dravet syndrome", American Epilepsy Association Annual Meeting Abstracts, abstract 3.1, published November 2019, (online journal) [retrieved from the internet on 26 August 2020], <URL:https://www.aesnet.org/meetings_events/annual_meeting_abstracts/view/2421999>Abstract

MILLER I. et al., "From gene replacement to gene regulation: developing a disease-modifying AAV gene therapy vector for SCN1A-positive (SCN1A+) pediatric epilepsy", American Epilepsy Association Annual Meeting Abstracts, abstract 1.091, published November 2019, (online journal) [retrieved from the internet on 26 August 2020], <URL: https://www.aesnet.org/meetings_events/annual_meeting_abstracts/view/2421087>Abstract

WO-A1-2019109051

WO-A1-2019224864

(57) В настоящем изобретении представлены сконструированные факторы транскрипции для селективной активации SCN1a и их применение для лечения заболеваний и нарушений, таких как синдром Драве. Также представлены сайты связывания микроРНК и их применение для селективной экспрессии в парвальбумин-содержащих нейронах.

B1**046157****046157 B1**

Ссылка

По заявке на настоящий патент испрашивается приоритет в соответствии с предварительной заявкой на патент США № 62/854238, поданной 29 мая 2019 г.; предварительной заявкой на патент США № 62/857727, поданной 5 июня 2019 г.; и предварительной заявкой на патент США № 63/008569, поданной 10 апреля 2020 г., каждая из которых полностью включена в настоящий документ посредством ссылки.

Перечень последовательностей

Настоящая заявка содержит Перечень последовательностей, который был представлен в электронном виде в формате ASCII и полностью включен в настоящий документ посредством ссылки. Указанная копия ASCII, созданная 28 мая 2020 г., называется 46482-724_601_SL.txt и имеет размер 418483 байта.

Предшествующий уровень техники настоящего изобретения

Широкий спектр заболеваний человека связан с аномальной экспрессией генов. В некоторых случаях генетическая мутация в гене вызывает нарушение его регуляции, подавление или отсутствие экспрессии вообще, что приводит к гипоплазии. В некоторых случаях генетическая мутация в гене вызывает его активацию, что приводит к сверхэкспрессии гена. При лечении генетических нарушений или заболеваний существует множество проблем. Один из подходов - генная терапия, которая включает в себя терапевтическую доставку нуклеиновой кислоты в клетки пациента. Тем не менее, различные проблемы, связанные с генной терапией, остаются нерешенными, такие как нежелательный иммунный ответ, вызванный генной терапией, нецелевые эффекты, ограничения клонирующей способности носителей генной терапии (например, вирусов), поддержание терапевтического эффекта в течение более длительного периода времени и т.д. Центральная нервная система (ЦНС) ставит множество уникальных задач для разработки терапии, направленной на устранение основного нарушения экспрессии гена и/или белка. Хотя существуют лекарственные средства, которые помогают управлять симптомами заболеваний/нарушений ЦНС, многие заболевания/нарушения ЦНС, например, синдром Драве, не имеют специальных способов лечения или устранения. Таким образом, существует потребность в новых композициях и способах, способных модулировать экспрессию любого эндогенного гена, чтобы помочь обратить вспять эффекты заболевания или нарушения, в частности, в терапии со сниженной иммуногенностью, уменьшенными нецелевыми эффектами, повышенной специфичностью для целевого гена и/или повышенной терапевтической эффективностью.

Сущность настоящего изобретения

Согласно одному аспекту в настоящей заявке представлена кассета экспрессии, содержащая последовательность, кодирующую не встречающийся в природе фактор транскрипции, который увеличивает экспрессию гена SCN1A в клетке, причем не встречающийся в природе фактор транскрипции содержит ДНК-связывающий домен (DBD), функционально связанный по меньшей мере с двумя активирующими транскрипцию доменами (TAD) следующим образом: TAD1-TAD2-DBD, DBD-TAD3-TAD4 или TAD1-TAD2-DBD-TAD3-TAD4. Согласно определенным вариантам осуществления TAD1, TAD2, TAD3 и TAD4 независимо выбраны из следующего: VP16, VP64, Viper, CITED2, CITED4, CREB3 или их функциональных фрагментов. Согласно определенным вариантам осуществления TAD1 и TAD2 представляют собой один и тот же TAD. Согласно определенным вариантам осуществления TAD1 и TAD2 представляют собой CITED2 или его функциональный фрагмент. Согласно определенным вариантам осуществления TAD3 и TAD4 представляют собой один и тот же TAD. Согласно определенным вариантам осуществления TAD3 и TAD4 представляют собой CITED2 или его функциональный фрагмент. Согласно определенным вариантам осуществления TAD3 и TAD4 представляют собой CITED4 или его функциональный фрагмент. Согласно определенным вариантам осуществления TAD1, TAD2, TAD3 и TAD4 представляют собой один и тот же TAD. Согласно определенным вариантам осуществления TAD1, TAD2, TAD3 и TAD4 представляют собой CITED2 или его функциональный фрагмент. Согласно определенным вариантам осуществления TAD1, TAD2, TAD3 и TAD4 представляют собой CITED4 или его функциональный фрагмент. Согласно определенным вариантам осуществления линкер отсутствует по меньшей мере между двумя доменами TAD.

Согласно определенным вариантам осуществления существует линкер по меньшей мере между двумя доменами TAD. Согласно определенным вариантам осуществления линкер содержит или состоит из GGSGGGSG (SEQ ID NO: 177) или GGSGGGSGGGSGGGSG (SEQ ID NO: 178).

Согласно определенным вариантам осуществления DBD связывается с областью генома, содержащей 18-27 нуклеотидов.

Согласно определенным вариантам осуществления DBD характеризуется идентичностью последовательности, составляющей по меньшей мере 80%, по отношению к его ближайшему человеческому аналогу. Согласно определенным вариантам осуществления DBD характеризуется идентичностью последовательности, составляющей по меньшей мере 90%, по отношению к его ближайшему человеческому аналогу. Согласно определенным вариантам осуществления каждый из DBD и по меньшей мере двух TAD характеризуется идентичностью последовательности, составляющей по меньшей мере 80%, по отношению к их ближайшим человеческим аналогам. Согласно определенным вариантам осуществления каждый из DBD и по меньшей мере двух TAD характеризуется идентичностью последовательности, состав-

ляющей по меньшей мере 90%, по отношению к их ближайшим человеческим аналогам.

Согласно определенным вариантам осуществления DBD содержит направляющую РНК и инактивированный нуклеазами белок Cas. Согласно определенным вариантам осуществления инактивированный нуклеазой белок Cas представляет собой инактивированный нуклеазой белок Cas9.

Согласно определенным вариантам осуществления DBD содержит домен типа "цинковые пальцы". Согласно определенным вариантам осуществления DBD содержит от шести до девяти доменов типа "цинковые пальцы". Согласно определенным вариантам осуществления DBD содержит шесть "цинковых пальцев". Согласно определенным вариантам осуществления DBD связывается с областью генома, содержащей 18 нуклеотидов. Согласно определенным вариантам осуществления DBD содержит девять "цинковых пальцев". Согласно определенным вариантам осуществления DBD связывается с областью генома, содержащей 27 нуклеотидов.

Согласно определенным вариантам осуществления DBD содержит последовательность, характеризующуюся идентичностью последовательности, составляющей по меньшей мере 95%, по отношению к любой из SEQ ID NO: 148-151. Согласно определенным вариантам осуществления DBD содержит последовательность, содержащую любую из SEQ ID NO: 148-151.

Согласно определенным вариантам осуществления DBD получен из EGR1 человека или EGR3 человека.

Согласно определенным вариантам осуществления DBD содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 77-98. Согласно определенным вариантам осуществления DBD содержит SEQ ID NO: 77-98.

Согласно определенным вариантам осуществления DBD содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к SEQ ID NO: 92. Согласно определенным вариантам осуществления DBD содержит SEQ ID NO: 92.

Согласно определенным вариантам осуществления не встречающийся в природе фактор транскрипции содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к SEQ ID NO: 130 или 131. Согласно определенным вариантам осуществления не встречающийся в природе фактор транскрипции содержит SEQ ID NO: 130 или 131.

Согласно определенным вариантам осуществления касета экспрессии содержит нуклеотидную последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 72 или 73. Согласно определенным вариантам осуществления касета экспрессии содержит нуклеотидную последовательность любой из SEQ ID NO: 72 или 73.

Согласно определенным вариантам осуществления касета экспрессии дополнительно содержит регуляторный элемент, который управляет экспрессией фактора транскрипции на более высоком уровне в PV-содержащих нейронах, чем в других типах клеток. Согласно определенным вариантам осуществления регуляторный элемент содержит любую из SEQ ID NO: 1-4. Согласно определенным вариантам осуществления регуляторный элемент содержит SEQ ID NO: 2 или 3.

Согласно определенным вариантам осуществления касета экспрессии дополнительно содержит селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления селективный в отношении PV сайт связывания микроРНК характеризуется идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 7, 14 или 15. Согласно определенным вариантам осуществления селективный в отношении PV сайт связывания микроРНК содержит любую из SEQ ID NO: 7, 14 или 15.

Согласно определенным вариантам осуществления касета экспрессии представляет собой часть вирусного вектора. Согласно определенным вариантам осуществления вирусный вектор представляет собой вирус AAV. Согласно определенным вариантам осуществления вирус AAV представляет собой вирус AAV9 или вирус scAAV9. Согласно определенным вариантам осуществления вирусный вектор представляет собой лентивирус.

Согласно другому аспекту в заявке представлена касета экспрессии, содержащая последовательность, кодирующую не встречающийся в природе фактор транскрипции, который увеличивает экспрессию гена SCN1A в клетке, причем не встречающийся в природе фактор транскрипции содержит ДНК-связывающий домен, функционально связанный с активирующим транскрипцию доменом, причем ДНК-связывающий домен представляет собой белок типа "цинковые пальцы", содержащий последовательность LEPGEKP - [YKSPCGKSFS X HQRTH TGEKP]_n - YKSPCGKSFS X HQRTH - TGKKTs (SEQ ID NO: 147), и причем тег HA отсутствует (SEQ ID NO: 303) между ДНК-связывающим доменом и активирующим транскрипцию доменом. Согласно определенным вариантам осуществления активирующий транскрипцию домен содержит последовательность VP16, VPR или VP64 или их функциональный фрагмент. Согласно определенным вариантам осуществления активирующий транскрипцию домен содержит VP64.

Согласно определенным вариантам осуществления ДНК-связывающий домен связывается с геномной областью, содержащей 18-27 нуклеотидов. Согласно определенным вариантам осуществления ДНК-связывающий домен представляет собой домен типа "цинковые пальцы", содержащий SEQ ID NO: 147, где n=6-9. Согласно определенным вариантам осуществления ДНК-связывающий домен представляет

собой домен типа "цинковые пальцы", содержащий SEQ ID NO: 147, где $n=6$. Согласно определенным вариантам осуществления ДНК-связывающий домен связывается с областью генома, содержащей 18 нуклеотидов. Согласно определенным вариантам осуществления ДНК-связывающий домен представляет собой домен типа "цинковые пальцы", содержащий SEQ ID NO: 147, где $n=9$. Согласно определенным вариантам осуществления ДНК-связывающий домен связывается с геномной областью, содержащей 27 нуклеотидов.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, характеризующуюся идентичностью последовательности, составляющей по меньшей мере 95%, по отношению к любой из SEQ ID NO: 148-151. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, содержащую любую из SEQ ID NO: 148-151.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 77-91. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит любую из SEQ ID NO: 77-91.

Согласно определенным вариантам осуществления касета экспрессии дополнительно содержит регуляторный элемент, который управляет экспрессией фактора транскрипции на более высоком уровне в PV-содержащих нейронах, чем в других типах клеток. Согласно определенным вариантам осуществления регуляторный элемент содержит любую из SEQ ID NO: 1-4. Согласно определенным вариантам осуществления регуляторный элемент содержит SEQ ID NO: 2 или 3.

Согласно определенным вариантам осуществления не встречающийся в природе фактор транскрипции содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к SEQ ID NO: 127. Согласно определенным вариантам осуществления не встречающийся в природе фактор транскрипции содержит SEQ ID NO: 127.

Согласно определенным вариантам осуществления касета экспрессии содержит нуклеотидную последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 93 или 71. Согласно определенным вариантам осуществления касета экспрессии содержит нуклеотидную последовательность любой из SEQ ID NO: 93 или 71.

Согласно определенным вариантам осуществления касета экспрессии дополнительно содержит селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления селективный в отношении PV сайт связывания микроРНК характеризуется идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 7, 14 или 15. Согласно определенным вариантам осуществления селективный в отношении PV сайт связывания микроРНК содержит любую из SEQ ID NO: 7, 14 или 15.

Согласно определенным вариантам осуществления касета экспрессии представляет собой часть вирусного вектора. Согласно определенным вариантам осуществления вирусный вектор представляет собой вирус AAV. Согласно определенным вариантам осуществления вирус AAV представляет собой вирус AAV9 или вирус scAAV9. Согласно определенным вариантам осуществления вирусный вектор представляет собой лентивирус. Согласно другому аспекту в заявке представлен полинуклеотид, содержащий селективный в отношении PV сайт связывания микроРНК, содержащий последовательность, характеризующуюся идентичностью последовательности, составляющей по меньшей мере 80%, по отношению к SEQ ID NO: 14 или 15, причем сайт связывания микроРНК снижает экспрессию трансгена в возбуждающих нейронах. Согласно определенным вариантам осуществления селективный в отношении PV сайт связывания микроРНК содержит SEQ ID NO: 14. Согласно определенным вариантам осуществления селективный в отношении PV сайт связывания микроРНК содержит SEQ ID NO: 15. Согласно другому аспекту в заявке представлена касета экспрессии, содержащая селективный в отношении PV сайт связывания микроРНК и промотор и/или энхансер. Согласно определенным вариантам осуществления промотор и/или энхансер представляет собой селективный в отношении PV регуляторный элемент, который управляет экспрессией трансгена на более высоком уровне в парвальбумин-содержащих нейронах (PV), чем в других типах клеток. Согласно определенным вариантам осуществления селективный в отношении PV регуляторный элемент функционально связан с трансгеном. Согласно другому аспекту в настоящей заявке представлена касета экспрессии, содержащая регуляторный элемент, функционально связанный с трансгеном, и по меньшей мере один сайт связывания микроРНК, причем регуляторный элемент управляет экспрессией трансгена на более высоком уровне в парвальбумин-содержащих (PV) нейронах, чем в других типах клеток, и причем сайт связывания микроРНК снижает экспрессию трансгена в возбуждающих нейронах. Согласно определенным вариантам осуществления касета экспрессии не содержит SEQ ID NO: 67. Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR128 (SEQ ID NO: 9). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR221 (SEQ ID NO: 11). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR222 (SEQ ID NO: 13). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR128 (SEQ ID NO: 9) и по меньшей мере один сайт связывания для MIR221 (SEQ ID NO:

11). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR128 (SEQ ID NO: 9), по меньшей мере один сайт связывания для MIR221 (SEQ ID NO: 11) и по меньшей мере один сайт связывания для MIR222 (SEQ ID NO: 13). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 7, 14 или 15. Согласно определенным вариантам осуществления сайт связывания микроРНК содержит SEQ ID NO: 7, 14 или 15.

Согласно определенным вариантам осуществления транскрипция кодирует полипептид, содержащий не встречающийся в природе фактор транскрипции, который увеличивает экспрессию гена SCN1A в клетке. Согласно определенным вариантам осуществления фактор транскрипции связывается с областью генома, содержащей 18-27 нуклеотидов. Согласно определенным вариантам осуществления фактор транскрипции содержит ДНК-связывающий домен. Согласно определенным вариантам осуществления фактор транскрипции содержит ДНК-связывающий домен и активирующий транскрипцию домен. Согласно определенным вариантам осуществления ДНК-связывающий домен характеризуется идентичностью последовательности, составляющей по меньшей мере 80%, по отношению к его ближайшему человеческому аналогу. Согласно определенным вариантам осуществления ДНК-связывающий домен характеризуется идентичностью последовательности, составляющей по меньшей мере 90%, по отношению к его ближайшему человеческому аналогу. Согласно определенным вариантам осуществления как ДНК-связывающий домен, так и активирующий транскрипцию домен характеризуются идентичностью последовательности, составляющей по меньшей мере 80%, по отношению к их ближайшим человеческим аналогам. Согласно определенным вариантам осуществления как ДНК-связывающий домен, так и активирующий транскрипцию домен характеризуется идентичностью последовательности, составляющей по меньшей мере 90%, по отношению к их ближайшим человеческим аналогам.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит направляющую РНК и инактивированный нуклеазой белок Cas. Согласно определенным вариантам осуществления инактивированный нуклеазой белок Cas представляет собой инактивированный нуклеазой белок Cas9.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит домен типа "цинковые пальцы". Согласно определенным вариантам осуществления ДНК-связывающий домен содержит от шести до девяти доменов типа "цинковые пальцы". Согласно определенным вариантам осуществления ДНК-связывающий домен содержит шесть "цинковых пальцев". Согласно определенным вариантам осуществления ДНК-связывающий домен связывается с областью генома, содержащей 18 нуклеотидов. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит девять "цинковых пальцев". Согласно определенным вариантам осуществления ДНК-связывающий домен связывается с геномной областью, содержащей 27 нуклеотидов.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, характеризующуюся идентичностью последовательности, составляющей по меньшей мере 95%, по отношению к любой из SEQ ID NO: 148-151. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, содержащую любую из SEQ ID NO: 148-151. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 92-98. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит любую из SEQ ID NO: 92-98.

Согласно определенным вариантам осуществления ДНК-связывающий домен представляет собой белок типа "цинковые пальцы", содержащий последовательность LEPGEKP - [YKSPCEGKSFS X HQRTH TGEKP]_n - YKSPCEGKSFS X HQRTH - TGKKT (SEQ ID NO: 147).

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 77-91. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит любую из SEQ ID NO: 77-91.

Согласно определенным вариантам осуществления ДНК-связывающий домен происходит от EGR1 человека или EGR3 человека.

Согласно определенным вариантам осуществления активирующий транскрипцию домен содержит последовательность VP16, VPR, VP64, CITED2, CITED4 или CREB3 или их функциональный фрагмент. Согласно определенным вариантам осуществления активирующий транскрипцию домен содержит последовательность CITED2, CITED4 или CREB3 человека или их функциональный фрагмент.

Согласно определенным вариантам осуществления регуляторный элемент содержит последовательность, содержащую любую из SEQ ID NO: 1-4. Согласно определенным вариантам осуществления регуляторный элемент содержит последовательность, содержащую SEQ ID NO: 2 или 3.

Согласно определенным вариантам осуществления не встречающийся в природе фактор транскрипции содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 105, 106 и 127-129. Согласно определенным вариантам

осуществления не встречающийся в природе фактор транскрипции содержит любую из SEQ ID NO: 105, 106 и 127-129.

Согласно определенным вариантам осуществления трансген содержит нуклеотидную последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 71, 74, 75, 76 или 184. Согласно определенным вариантам осуществления трансген содержит любую из SEQ ID NO: 71, 74, 75, 76 или 184.

Согласно определенным вариантам осуществления кассета экспрессии представляет собой часть вирусного вектора. Согласно определенным вариантам осуществления вирусный вектор представляет собой вирус AAV. Согласно определенным вариантам осуществления вирус AAV представляет собой вирус AAV9 или вирус scAAV9. Согласно определенным вариантам осуществления вирусный вектор представляет собой лентивирус.

Согласно другому аспекту в настоящей заявке представлен способ селективной экспрессии трансгена в парвальбумин-содержащих нейронах (PV) приматов, предусматривающий введение примату вирусного вектора, содержащего трансген и по меньшей мере один сайт связывания микроРНК, причем сайт связывания микроРНК снижает экспрессию трансгена в возбуждающих нейронах.

Согласно определенным вариантам осуществления вирусный вектор дополнительно содержит регуляторный элемент, функционально связанный с трансгеном, причем регуляторный элемент управляет экспрессией трансгена на более высоком уровне в парвальбумин-содержащих нейронах (PV), чем в других типах клеток. Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR128 (SEQ ID NO: 9). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR221 (SEQ ID NO: 11). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR222 (SEQ ID NO: 13). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR128 (SEQ ID NO: 9) и по меньшей мере один сайт связывания для MIR221 (SEQ ID NO: 11). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит по меньшей мере один сайт связывания для MIR128 (SEQ ID NO: 9), по меньшей мере один сайт связывания для MIR221 (SEQ ID NO: 11) и по меньшей мере один сайт связывания для MIR222 (SEQ ID NO: 13). Согласно определенным вариантам осуществления сайт связывания микроРНК содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 7, 14 или 15. Согласно определенным вариантам осуществления сайт связывания микроРНК содержит SEQ ID NO: 7, 14 или 15.

Согласно определенным вариантам осуществления трансген содержит последовательность, кодирующую не встречающийся в природе фактор транскрипции, который увеличивает экспрессию гена SCN1A в клетке.

Согласно определенным вариантам осуществления фактор транскрипции связывается с областью генома, содержащей 18-27 нуклеотидов.

Согласно определенным вариантам осуществления фактор транскрипции содержит ДНК-связывающий домен.

Согласно определенным вариантам осуществления фактор транскрипции содержит ДНК-связывающий домен и активирующий транскрипцию домен.

Согласно определенным вариантам осуществления ДНК-связывающий домен характеризуется идентичностью последовательности, составляющей по меньшей мере 80%, по отношению к его ближайшему человеческому аналогу. Согласно определенным вариантам осуществления ДНК-связывающий домен характеризуется идентичностью последовательности, составляющей по меньшей мере 90%, по отношению к его ближайшему человеческому аналогу. Согласно определенным вариантам осуществления как ДНК-связывающий домен, так и активирующий транскрипцию домен характеризуются идентичностью последовательности, составляющей по меньшей мере 80%, по отношению к их ближайшим человеческим аналогам. Согласно определенным вариантам осуществления как ДНК-связывающий домен, так и активирующий транскрипцию домен характеризуются идентичностью последовательности, составляющей по меньшей мере 90%, по отношению к их ближайшим человеческим аналогам.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит направляющую РНК и инактивированный нуклеазой белок Cas. Согласно определенным вариантам осуществления инактивированный нуклеазой белок Cas представляет собой инактивированный нуклеазой белок Cas9.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит домен типа "цинковые пальцы". Согласно определенным вариантам осуществления ДНК-связывающий домен содержит от шести до девяти доменов типа "цинковые пальцы". Согласно определенным вариантам осуществления ДНК-связывающий домен содержит шесть "цинковых пальцев". Согласно определенным вариантам осуществления ДНК-связывающий домен связывается с областью генома, содержащей 18 нуклеотидов. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит девять "цинковых пальцев". Согласно определенным вариантам осуществления ДНК-связывающий домен свя-

зывается с геномной областью, содержащей 27 нуклеотидов.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, характеризующуюся идентичностью последовательности, составляющей по меньшей мере 95%, по отношению к любой из SEQ ID NO: 148-151. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, имеющую любую из SEQ ID NO: 148-151.

Согласно определенным вариантам осуществления ДНК-связывающий домен происходит из EGR1 человека или EGR3 человека.

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 92-98. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит любую из SEQ ID NO: 92-98.

Согласно определенным вариантам осуществления ДНК-связывающий домен представляет собой белок типа "цинковые пальцы", содержащий последовательность LEPGEKP - [YKSPCGKSFS X HQRTH TGEKP]_n - YKSPCGKSFS X HQRTH - TGKKTS (SEQ ID NO: 147).

Согласно определенным вариантам осуществления ДНК-связывающий домен содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 77-91. Согласно определенным вариантам осуществления ДНК-связывающий домен содержит любую из SEQ ID NO: 77-91.

Согласно определенным вариантам осуществления активирующий транскрипцию домен содержит последовательность VP16, VPR, VP64, CITED2, CITED4 или CREB3 или их функциональный фрагмент. Согласно определенным вариантам осуществления активирующий транскрипцию домен содержит последовательность CITED2, CITED4 или CREB3 человека или их функциональный фрагмент.

Согласно определенным вариантам осуществления регуляторный элемент содержит последовательность, содержащую любую из SEQ ID NO: 1-4. Согласно определенным вариантам осуществления регуляторный элемент содержит последовательность, содержащую SEQ ID NO: 2 или 3.

Согласно определенным вариантам осуществления не встречающийся в природе фактор транскрипции содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 105, 106 и 127-129. Согласно определенным вариантам осуществления не встречающийся в природе фактор транскрипции содержит любую из SEQ ID NO: 105, 106 и 127-129.

Согласно определенным вариантам осуществления трансген содержит нуклеотидную последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к любой из SEQ ID NO: 71, 74, 75, 76 или 184. Согласно определенным вариантам осуществления трансген содержит любую из SEQ ID NO: 71, 74, 75, 76 или 184.

Согласно определенным вариантам осуществления вирусный вектор представляет собой вирус AAV. Согласно определенным вариантам осуществления вирус AAV представляет собой вирус AAV9 или вирус scAAV9. Согласно определенным вариантам осуществления вирусный вектор представляет собой лентивирус.

Согласно определенным вариантам осуществления притма представляет собой человека. Согласно определенным вариантам осуществления примат представляет собой отличного от человека примата. Согласно определенным вариантам осуществления отличный от человека примат представляет собой обезьяну старого мира, орангутана, гориллу, шимпанзе, мартышку, яванского макака, макака-резус или свиногохвостого макака.

Согласно другому аспекту в настоящей заявке представлена кассета экспрессии, содержащая последовательность, кодирующую не встречающийся в природе фактор транскрипции, который увеличивает экспрессию гена SCN1A в клетке, причем не встречающийся в природе фактор транскрипции содержит последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 90%, по отношению к SEQ ID NO: 128 или 129. Согласно определенным вариантам осуществления не встречающийся в природе фактор транскрипции содержит SEQ ID NO: 128 или 129.

Согласно другому аспекту в настоящей заявке представлен способ увеличения экспрессии SCN1A в клетке путем введения любой из кассет экспрессии, представленных в настоящем документе. Согласно определенным вариантам осуществления клетка представляет собой нервную клетку. Согласно определенным вариантам осуществления нервная клетка выбрана из группы, состоящей из униполярных, биполярных, мультиполярных или псевдоуниполярных нейронов. Согласно определенным вариантам осуществления клетка представляет собой ГАМКергический нейрон. Согласно определенным вариантам осуществления клетка представляет собой PV-содержащий нейрон. Согласно определенным вариантам осуществления клетка не представляет собой нервную клетку. Согласно определенным вариантам осуществления клетка представляет собой глиальную клетку. Согласно определенным вариантам осуществления клетка выбрана из группы, состоящей из астроцитов, олигодендроцитов, эпендимных клеток, шванновских клеток и сателлитных клеток. Согласно определенным вариантам осуществления клетка находится внутри субъекта. Согласно определенным вариантам осуществления субъект представляет собой млекопитающее. Согласно определенным вариантам осуществления субъект представляет

собой человека. Согласно определенным вариантам осуществления увеличение экспрессии SCN1A лечит заболевание, нарушение или симптом. Согласно определенным вариантам осуществления нарушение представляет собой нарушение центральной нервной системы. Согласно определенным вариантам осуществления нарушение представляет собой эпилепсию, связанную с гаплонедостаточностью SCN1A. Согласно определенным вариантам осуществления гаплонедостаточность является результатом того, что субъект является гетерозиготным по мутации потери функции гена SCN1A. Согласно определенным вариантам осуществления нарушение представляет собой эпилепсию, связанную со вставкой, делецией или заменой в гене SCN1A. Согласно определенным вариантам осуществления нарушение представляет собой эпилепсию, связанную с точечной мутацией в гене SCN1A. Согласно определенным вариантам осуществления нарушение представляет собой синдром Драве. Согласно определенным вариантам осуществления симптом нарушение центральной нервной системы представляет собой гиперактивность нейронов. Согласно определенным вариантам осуществления лечение нарушения центральной нервной системы предусматривает снижение гиперактивности нейронов. Согласно определенным вариантам осуществления симптомом нарушения центральной нервной системы являются припадки. Согласно определенным вариантам осуществления лечение нарушения центральной нервной системы предусматривает снижение частоты припадков. Согласно определенным вариантам осуществления лечение нарушения центральной нервной системы предусматривает уменьшение тяжести припадков.

Согласно другому аспекту в настоящей заявке представлен способ увеличения экспрессии SCN1A в ЦНС путем введения любой из кассет экспрессии, представленных в настоящем документе. Согласно определенным вариантам осуществления кассету экспрессии вводят посредством одностороннего интрацеребровентрикулярного (ICV) введения. Согласно определенным вариантам осуществления кассету экспрессии вводят посредством двустороннего интрацеребровентрикулярного (ICV) введения. Согласно определенным вариантам осуществления повышенная экспрессия SCN1A происходит в головном мозге. Согласно определенным вариантам осуществления повышенная экспрессия SCN1A происходит во фронтальной коре, теменной коре, височной коре, гиппокампе, продолговатом мозге и/или затылочной коре. Согласно определенным вариантам осуществления повышенная экспрессия SCN1A происходит в позвоночнике. Согласно определенным вариантам осуществления повышенная экспрессия SCN1A происходит в спинном мозге и/или межпозвоночном узле.

Включение по ссылке

Все публикации, патенты и заявки на патенты, упомянутые в настоящем описании, включены в настоящий документ посредством ссылки в той же степени, как если бы каждая отдельная публикация, патент или заявка на патент были специально и индивидуально указаны для включения посредством ссылки.

Краткое описание графических материалов

Новые особенности настоящего изобретения подробно изложены в прилагаемой формуле изобретения. Лучшее понимание особенностей и преимуществ настоящего изобретения будет получено при обращении к нижеследующему подробному описанию, в котором излагаются иллюстративные случаи, в которых используются принципы настоящего изобретения, и прилагаемым графическим материалам, на которых показано следующее.

На фиг. 1 показана активация эндогенного SCN1A с использованием сконструированных факторов транскрипции, которые связываются с различными областями на хромосоме 2 (со ссылкой на GRCh38.p12). Данные представлены в виде кратного изменения экспрессии SCN1A по сравнению с контрольным (EGFP-KASH) состоянием.

На фиг. 2А, фиг. 2В и фиг. 2С показана относительная экспрессия эндогенного SCN1A в клетках HEK293 с использованием SCN1A-специфических активаторов транскрипции (смотрите табл. 1). Данные представлены в виде кратного изменения относительно контрольных условий и показаны на шкале Log_{10} .

На фиг. 3А показана относительная экспрессия эндогенного SCN1A в нейронах ГАМК с использованием SCN1A-специфического активатора транскрипции (конструкция 30). Данные представлены в виде кратного изменения относительно контрольных условий (CBA-EGFP).

На фиг. 3В показана относительная экспрессия эндогенного SCN1A в нейронах ГАМК с использованием SCN1A-специфических активаторов транскрипции (конструкции 25 и 16). Данные представлены в виде кратного изменения относительно контрольных условий (CBA-EGFP) в Log_{10} .

На фиг. 4 показана относительная экспрессия эндогенного SCN1A и 40 ближайших соседних генов, управляемых специфическим фактором транскрипции SCN1A (конструкция 30). Данные представлены в виде кратного изменения относительно контрольных условий (CBA-EGFP-KASH) в Log_{10} .

На фиг. 5А и фиг. 5В показана экспрессия SCN1A-специфического активатора транскрипции *in vivo* по сравнению с контрольной кассетой экспрессии, которая экспрессировала eGFP. На фиг. 5А показана относительная экспрессию гена SCN1A у мышей, которым инъецировали либо контрольный eGFP, либо конструкцию 4, содержащую активатор транскрипции SCN1A. На фиг. 5В показано изменение экспрессии SCN1A с точки зрения среднего процента eGFP. Эти эксперименты показывают, что активация транскрипции с помощью конструкции 4 привела к усилению экспрессии SCN1A приблизительно на 20-30%.

На фиг. 6А, фиг. 6В, фиг. 6С, фиг. 6D, фиг. 6Е, фиг. 6F и фиг. 6G показано влияние на гипертерми-

ческие припадки на мышинной модели синдрома Драве с нокаутом $Scn1a^{tm1Kca}$ с использованием различных факторов транскрипции, специфических для SCN1A, по сравнению с контролем. Мышам P1 $Scn1a$ +/- (гетерозиготные; HET) вводили либо AAV9-EGFP, либо вектор AAV9, экспрессирующий фактор транскрипции, специфический для SCN1A (одна из конструкций 31-34, 42 и 43). Мышей P26-P28, подвергнутых инфузии, подвергали тесту на припадки, вызванные гипертермией, и регистрировали внутреннюю температуру, при которой они испытывали тонико-клонический припадок. На фиг. 6D показано прямое сравнение между конструкцией 32, которая содержит тег HA, расположенным между DBD и TAD, и конструкцией 34, которая не содержит тега HA. На фиг. 6E показано прямое сравнение между конструкцией 31, которая содержит сайт связывания микроРНК m1, расположенный между кодирующей областью и хвостом полиА, и конструкцией 32, которая не содержит сайт связывания микроРНК m1. На фиг. 6H проиллюстрировано влияние на гипертермические припадки в модели синдрома Драве у мышей с мутантным $Scn1a^{RX}$ использованием конструкции 31 (по сравнению с контролем в виде введенного PBS).

На фиг. 7A, фиг. 7B, фиг. 7C и фиг. 7D показана выживаемость в модели синдрома Драве у мышей с нокаутом $Scn1a^{tm1Kca}$ в различных условиях. На фиг. 7A показано сравнение мышей дикого типа (PBS WT) и $Scn1a$ +/- (PBS HET) в анализе выживаемости. Мышам P1 $Scn1a$ +/- (N=53) и $Scn1a$ +/- (N=54) вливали PBS. Мышей наблюдали в их домашней клетке ежедневно, и в случае какой-либо смертности регистрировали дату. Наблюдалась значительная разница в выживаемости между животными $Scn1a$ +/- и $Scn1a$ +/- (P<0,0001). На фиг. 7B-D показано влияние на выживаемость в мышинной модели синдрома Драве для мышей, обработанных различными факторами транскрипции, специфическими для SCN1A, по сравнению с контролем. Мышам P1 $Scn1a$ +/- вливали либо PBS, либо вектор AAV9, экспрессирующий фактор транскрипции, специфический для SCN1A (конструкции 31 или 33). Мышей наблюдали в их домашней клетке ежедневно, и в случае какой-либо смертности регистрировали дату. На фиг. 7D показано прямое сравнение между конструкцией 31, которая содержит сайт связывания микроРНК m1, расположенным между кодирующей областью и хвостом полиА, и конструкцией 33, которая не содержит сайт связывания микроРНК m1. На фиг. 7E показана выживаемость в модели синдрома Драве у мышей с мутантным $Scn1a^{RX}$ с использованием конструкции 31 (по сравнению с контролем, инъецированным PBS).

На фиг. 8 показана относительная экспрессия мРНК $Scn1A$ в различных тканях головного мозга после интрапаренхимальной доставки вектора AAV9, кодирующего SCN1A-специфический фактор транскрипции (конструкция 33), введенного двум яванским макакам в дозе $1,2 \times 10^{12}$ копий генома/животное, нормализованных по двум необработанным контрольным животным. Всех животных умерщвляли через 28 дней после инъекции и количественно определяли мРНК $Scn1A$ в образцах тканей с помощью Taqman PCR. Данные представлены как нормализованная экспрессия целевой мРНК в различных срезах ткани головного мозга. Аналогичные результаты были зарегистрированы с другим набором праймеров/зонда, полученных из гена $Scn1a$.

На фиг. 9A-F показан характер экспрессии EGFP в области зубчатой извилины гиппокампа мартышки после обработки векторами AAV9, содержащими трансген EGFP под контролем промотора EF1a, промотора RE 2 (SEQ ID NO: 2) или промотора RE 2 (SEQ ID NO: 2) с сайтом связывания микроРНК m1 (SEQ ID NO: 7), расположенным между кодирующей областью EGFP и сайтом polyA. Типичная область зубчатой извилины гиппокампа показана для каждого вектора обработки. В верхнем ряду показаны ядра клеток, окрашенные DAPI, а в нижнем ряду показаны GFP-положительные области, окрашенные антителом к GFP. На фиг. 9A (обработка EF1a) область hylus CA4 гиппокампа обведена, а стрелки указывают на слой тела гранулярных клеток зубчатых клеток (DG). Фиг. 9B и фиг. 9C сосредоточены в той же области. Область CA4, которая представляет собой смесь возбуждающих и тормозных интернейронов, выделена, поскольку это была единственная область со значительной экспрессией в условии RE 2 + m1. При управляемой EF1a и RE 2 экспрессии трансгена экспрессия GFP была более распространенной и включала в себя другие области гиппокампа. Считается, что слой клеток DG содержит в основном возбуждающие нейроны. Экспрессия GFP, управляемая EF1a и RE 2, видна в слое клеток DG (фиг. 9D и фиг. 9E), но не присутствует у животных, обработанных RE 2 + m1 (фиг. 9F) (белые стрелки).

На фиг. 10A-L показано, что характер экспрессии EGFP в области зубчатой извилины гиппокампа мартышки после обработки векторами AAV9, содержащими трансген EGFP под контролем промотора EF1a, промотора RE 2 (SEQ ID NO: 2) или промотора RE 2 (SEQ ID NO: 2) с сайтом связывания микроРНК m1 (SEQ ID NO: 7), расположенным между кодирующей областью EGFP и сайтом polyA, в первую очередь локализован на парвальбумин-положительных (PV) клетках у животных, получавших RE 2 и RE 2 + m1. Типичная область зубчатой извилины гиппокампа показана для каждого вектора обработки. В верхнем ряду показаны GFP-положительные области, а в следующем нижнем ряду показаны те же области, окрашенные ингибирующим маркером интернейронов для PV. Область в рамке на фиг. 10A-F показана при большем увеличении на фиг. 10G-L. Экспрессия GFP, управляемая RE 2 и RE 2 + m1, в основном совмещена с ингибирующим маркером PV-содержащих интернейронов (белые стрелки на фиг. 10H и 10K, 10I и 10L), тогда как в EG-EF1a экспрессия GFP не так легко локализуется в PV-положительных клетках (белые стрелки на фиг. 10G и 10J). Кроме того, GFP-положительные клетки характеризуются отчетливо выраженной морфологией интернейронов высокоразветвленных клеток с пи-

рамидным клеточным телом у животных, получавших RE 2 и RE 2 + m1 (стрелки на фиг. 10H и 10I), по сравнению с менее отчетливой морфологией клеточного тела у животных, получавших EF1a (стрелки на фиг. 10G).

На фиг. 11 показан VG/диплоидный геном в образцах тканей фронтальной коры (FC), ростральной теменной коры (Rostral PC), височной коры (TC), каудальной теменной коры (Caudal PC), гиппокампа (Hip), продолговатого мозга (Med) и затылочной коры (OC) животных, получавших AAV9-RE^{GABA-eTF^{SCN1A}}, вводимые в дозе 4,8E+13 или 8E+13 мкг/животное посредством одностороннего интрацеребровентрикулярного (ICV) введения (пример 10 и пример 11). Каждая точка данных представляет VG/диплоидный геном для образца ткани, а горизонтальные полосы представляют средний VG/диплоидный геном для всех образцов ткани для каждого животного.

На фиг. 12 показаны транскрипты/мкг РНК в образцах тканей фронтальной коры (FC), ростральной теменной коры (Rostral PC), височной коры (TC), каудальной теменной коры (Caudal PC), гиппокампа (Hip), продолговатого мозга (Med) и затылочной коры (OC) животных, получавших AAV9-RE^{GABA-eTF^{SCN1A}}, вводимые в дозе 4,8E+13 или 8E+13 геномов вектора/животное посредством одностороннего интрацеребровентрикулярного (ICV) введения (пример 10 и пример 11). Каждая точка данных представляет геномы вектора/диплоидный геном для образца ткани, а горизонтальные полосы представляют собой среднее отношение геномы вектора/диплоидный геном для всех образцов ткани для каждого животного. Средние транскрипты для ARFGAP2 составляли 1,85E+6/мкг РНК и обозначены пунктирной верхней граничной линией. Предел обнаружения обозначен пунктирной нижней граничной линией.

На фиг. 13 показано биораспределение вектора (геномы вектора/диплоидный геном) и экспрессия трансгена (транскрипты/мкг РНК) в образцах периферической ткани вне головного мозга. Показанные образцы периферической ткани представляют собой спинной мозг C2/L4 (SC C2/L4), межпозвоночный узел C2/L4 (DRG C2/L4), печень, селезенка, сердце, почки, легкие, поджелудочная железа и семенники/яичники. Среднее значение VCN (векторное биораспределение) и транскрипта (экспрессия трансгена) в головном мозге приматов показано пунктирной линией.

Подробное описание настоящего изобретения

В настоящем документе представлены сконструированные факторы транскрипции или eTF, которые не встречаются в природе и были разработаны для связывания с целевым сайтом генома и модуляции экспрессии представляющего интерес эндогенного гена. Такие eTF могут быть созданы для усиления или подавления экспрессии (экспрессии РНК и/или белка) представляющего интерес гена. В настоящем документе также представлены сайты связывания микроРНК, которые могут быть включены в вирусный вектор и обеспечивать селективную экспрессию трансгена в парвальбумин-содержащих нейронах (PV).

Согласно одному аспекту в настоящей заявке представлены eTF, которые способны повышать экспрессию гена альфа-субъединицы 1 натриевого потенциалзависимого канала (SCN1A) и повышать экспрессию его соответствующего белкового продукта Nav1.1, а также способы их применения для лечения заболеваний или нарушений, связанных с дефицитом Nav1.1, такие как, например, синдром Драве.

Согласно другому аспекту в настоящей заявке предусмотрены сайты связывания микроРНК, которые снижают экспрессию мРНК, содержащей сайт связывания микроРНК, в возбуждающих нейронах, тем самым приводя к селективной экспрессии гена в ГАМКергических нейронах или парвальбумин-содержащих нейронах (PV), и способы их применения для селективной экспрессии представляющего интерес гена в нейронах PV.

Определения.

Используемые в настоящем документе формы единственного числа предназначены для включения и форм множественного числа, если контекст явно не указывает на иное. Кроме того, в той степени, в которой термины "включая в себя", "включает в себя", "имеющий", "имеет", "с" или их варианты используются либо в подробном описании, либо в формуле изобретения, такие термины предназначены для включения аналогично термину "содержащий".

Термин "приблизительно" или "примерно" означает в пределах допустимого диапазона ошибок для конкретного значения, определенного специалистом в настоящей области техники, который будет частично зависеть от того, как значение измеряется или определяется, т.е. ограничения системы измерения. Например, "приблизительно" может означать в пределах одного или более чем одного стандартного отклонения в соответствии с практикой в настоящей области техники. В качестве альтернативы "приблизительно" может означать диапазон до 20%, до 15%, до 10%, до 5% или до 1% от заданного значения.

Термины "определение", "измерение", "оценка", "оценивание", "анализ" и их грамматические эквиваленты могут использоваться в настоящем документе взаимозаменяемо для обозначения любой формы измерения и включают в себя определение того, присутствует элемент или нет (например, обнаружение). Эти термины могут включать в себя как количественные, так и/или качественные определения. Оценка может быть относительной или абсолютной.

Термин "экспрессия" относится к процессу, с помощью которого последовательность нуклеиновой кислоты или полинуклеотид транскрибируется с матрицы ДНК (например, в мРНК или другой транскрипт РНК), и/или процессу, посредством которого транскрибированная мРНК впоследствии транслируется в пептиды, полипептиды или белки. Транскрипты и кодируемые полипептиды можно вместе назы-

вать "генным продуктом". Если полинуклеотид происходит из геномной ДНК, экспрессия может включать в себя сплайсинг мРНК в эукариотической клетке.

Используемые в настоящем документе термины "функционально связанный", "функциональная связь" или их грамматические эквиваленты относятся к сопоставлению генетических элементов, например, промотора, энхансера, последовательности полиаденилирования и т.д., причем элементы находятся во взаимосвязи, позволяющей им работать ожидаемым образом. Например, регуляторный элемент, который может включать в себя промоторные и/или энхансерные последовательности, функционально связан с кодирующей областью, если регуляторный элемент помогает инициировать транскрипцию кодирующей последовательности. Между регуляторным элементом и кодирующей областью могут быть промежуточные остатки, пока сохраняется эта функциональная связь. Используемый в настоящем документе термин "вектор" относится к макромолекуле или ассоциации макромолекул, которая содержит полинуклеотид или ассоциируется с ним и которая может использоваться для опосредования доставки полинуклеотида в клетку. Примеры векторов включают в себя плазмиды, вирусные векторы, липосомы и другие носители для доставки генов. Вектор, как правило, содержит генетические элементы, например, регуляторные элементы, функционально связанные с геном для облегчения экспрессии гена в мишени.

Используемые в настоящем документе термины "кассета экспрессии" и "кассета нуклеиновой кислоты" взаимозаменяемо относятся к комбинации последовательностей нуклеиновых кислот или элементов, которые экспрессируются вместе или функционально связаны для экспрессии. В некоторых случаях кассета экспрессии относится к комбинации регуляторных элементов и гена или генов, с которыми они функционально связаны для экспрессии.

Термин "AAV" представляет собой сокращение от аденоассоциированного вируса и может использоваться для обозначения самого вируса или его производного. Термин охватывает все серотипы, подтипы, а также встречающиеся в природе и рекомбинантные формы, если не требуется иное. Сокращение "гAAV" относится к рекомбинантному аденоассоциированному вирусу, также называемому рекомбинантным вектором AAV (или "вектором гAAV"). Термин "AAV" включает в себя AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, rh10 и их гибриды, AAV птиц, AAV крупного рогатого скота, AAV собак, AAV лошадей, AAV приматов, AAV не приматов и AAV овец. Геномные последовательности различных серотипов AAV, а также последовательности нагивных концевых повторов (TR), белков Rep и субъединиц капсида известны в настоящей области техники. Такие последовательности можно найти в литературе или в общедоступных базах данных, таких как GenBank. Используемый в настоящем документе термин "вектор гAAV" относится к вектору AAV, содержащему полинуклеотидную последовательность, не имеющую происхождения AAV (т.е. полинуклеотид, гетерологичный AAV), как правило, последовательность, представляющую интерес для генетической трансформации клетки. В общем, гетерологичный полинуклеотид фланкирован по меньшей мере одной, а обычно двумя последовательностями инвертированных концевых повторов AAV (ITR). Вектор гAAV может быть одноцепочечным (ssAAV) или самокомплементарным (scAAV). "Вирус AAV" или "вирусная частица AAV" относится к вирусной частице, состоящей по меньшей мере из одного капсидного белка AAV и инкапсидированного полинуклеотидного вектора гAAV. Если частица содержит гетерологичный полинуклеотид (т.е. полинуклеотид, отличный от генома AAV дикого типа, такой как трансген, который должен быть доставлен в клетку млекопитающего), ее обычно называют "векторной частицей гAAV" или просто "частицей гAAV". Таким образом, получение частицы гAAV обязательно включает в себя производство вектора гAAV, поскольку такой вектор содержится в частице гAAV.

Используемые в настоящем документе термины "лечить", "лечение", "терапия" и т.п. относятся к облегчению, задержке или замедлению прогрессирования, профилактике, ослаблению, уменьшению эффектов или симптомов, предотвращению возникновения, ингибированию или облегчению начало заболеваний или нарушений. Способы настоящего раскрытия можно использовать с любым млекопитающим. Примеры млекопитающих включают в себя, без ограничения, крыс, кошек, собак, лошадей, коров, овец, свиней и, более предпочтительно, людей. Терапевтический эффект включает в себя искоренение или улучшение основного подлежащего лечению заболевания. Кроме того, терапевтический эффект достигается за счет устранения или облегчения одного или нескольких физиологических симптомов, связанных с основным заболеванием, так что улучшение наблюдается у субъекта, несмотря на то, что субъект все еще может страдать от основного нарушения. В некоторых случаях с профилактической целью терапевтическое средство можно вводить субъекту с риском развития определенного заболевания или субъекту, сообщающему об одном или нескольких физиологических симптомах заболевания, даже если диагноз этого заболевания не был поставлен. Способы настоящего раскрытия можно использовать с любым млекопитающим. В некоторых случаях лечение может привести к уменьшению или прекращению симптомов (например, снижению частоты, продолжительности и/или тяжести припадков). Профилактический эффект включает в себя отсрочку или устранение появления заболевания или состояния, отсрочку или устранение появления симптомов заболевания или состояния, замедление, остановку или обращение вспять прогрессирования заболевания или состояния или любую их комбинацию.

Термин "эффективное количество" или "терапевтически эффективное количество" относится к тому количеству описанной в настоящем документе композиции, которого достаточно для воздействия на

предполагаемое применение, включая в себя, без ограничения, лечение заболевания, как определено ниже. Терапевтически эффективное количество может варьироваться в зависимости от предполагаемого лечебного применения (*in vivo*) или субъекта и подвергаемого лечению патологического состояния, например, массы и возраста субъекта, тяжести патологического состояния, способа введения и т.п., что может легко определить специалист в настоящей области техники. Этот термин также применяется к дозе, которая вызовет конкретный ответ в целевой клетке. Конкретная доза будет варьироваться в зависимости от конкретной выбранной композиции, схемы дозированного введения, которую необходимо соблюдать, вводится ли она в комбинации с другими соединениями, времени введения, ткани, в которую ее вводят, и физической системы доставки, в которой она вводится.

"Фрагмент" нуклеотидной или пептидной последовательности относится к последовательности, которая короче эталонной или "полноразмерной" последовательности.

"Вариант" молекулы относится к аллельным вариациям таких последовательностей, т.е. к последовательности, по существу сходной по структуре и биологической активности либо со всей молекулой, либо с ее фрагментом.

"Функциональный фрагмент" последовательности ДНК или белка относится к фрагменту, который сохраняет биологическую активность (функциональную или структурную), которая по существу аналогична биологической активности полноразмерной последовательности ДНК или белка. Биологическая активность последовательности ДНК может заключаться в ее способности влиять на экспрессию способом, который, как известно, приписывается полноразмерной последовательности.

Термины "субъект" и "индивидуум" используются в настоящем документе взаимозаменяемо для обозначения позвоночного животного, предпочтительно, млекопитающего, более предпочтительно, человека. Описанные в настоящем документе способы могут быть применимы в терапевтических целях для человека, в ветеринарии и/или в доклинических исследованиях на животных моделях заболевания или состояния.

Термин "*in vivo*" относится к событию, которое происходит в организме субъекта.

Термин "*in vitro*" относится к событию, которое происходит вне тела субъекта. Например, анализ *in vitro* включает в себя любой анализ, проводимый вне субъекта. Анализы *in vitro* включают в себя анализы на основе клеток, в которых используются живые или мертвые клетки. Анализы *in vitro* также включают в себя бесклеточный анализ, в котором не используются интактные клетки.

В общем, "идентичность последовательностей" или "гомология последовательностей", которые могут использоваться взаимозаменяемо, относятся к точному соответствию нуклеотид-нуклеотид или аминокислота-аминокислота двух полинуклеотидных или полипептидных последовательностей, соответственно. Как правило, способы определения идентичности последовательностей включают в себя сравнение двух нуклеотидных или аминокислотных последовательностей и определение их процентной идентичности. Сравнение последовательностей, например, с целью оценки идентичности, может быть выполнено с помощью любого подходящего алгоритма выравнивания, включая в себя, без ограничения, алгоритм Нидлмана-Вунша (см., например, выравниватель EMBOSS Needle, доступный на www.ebi.ac.uk/Tools/psa/emboss_needle/, обязательно с настройками по умолчанию), алгоритм BLAST (см., например, инструмент выравнивания BLAST, доступный на blast.ncbi.nlm.nih.gov/Blast.cgi, обязательно с настройками по умолчанию) и алгоритм Смита-Уотермана (см., например, выравниватель EMBOSS Water, доступный на www.ebi.ac.uk/Tools/psa/emboss_water/, обязательно с настройками по умолчанию). Оптимальное выравнивание можно оценить с использованием любых подходящих параметров выбранного алгоритма, включая в себя параметры по умолчанию. "Процент идентичности", также называемый "процентом гомологии", между двумя последовательностями может быть рассчитан как количество точных совпадений между двумя оптимально выровненными последовательностями, деленное на длину эталонной последовательности и умноженное на 100. Процент идентичности также может быть определен, например, путем сравнения информации о последовательностях с использованием усовершенствованной компьютерной программы BLAST, включая в себя версию 2.2.9, доступную в Национальных институтах здравоохранения. Программа BLAST основана на способе совмещения Karlin and Altschul, Proc. Natl. Acad. Sci. USA 87:2264-2268 (1990) и как описано в Altschul, et al., J. Mol. Biol. 215:403-410 (1990); Karlin and Altschul, Proc. Natl. Acad. Sci. USA 90:5873-5877 (1993) и Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997). Вкратце, программа BLAST определяет идентичность как количество идентичных выровненных символов (т.е. нуклеотидов или аминокислот), деленное на общее количество символов в более короткой из двух последовательностей. Программа может использоваться для определения процента идентичности по всей длине сравниваемых последовательностей. Параметры по умолчанию предоставляются для оптимизации поиска с помощью коротких последовательностей запросов, например, с помощью программы *blastp*. Программа также позволяет использовать фильтр SEG для маскировки сегментов последовательностей запросов, как это определено программой SEG Wootton and Federhen, Computers and Chemistry 17: 149-163 (1993). Высокая идентичность последовательностей, как правило, включает в себя диапазоны идентичности последовательностей от приблизительно 80% до 100% и целые числа между ними. Используемый в настоящем документе термин "сконструированный" применительно к белку относится к не встречающемуся в природе белку, включая в себя, без ограниче-

ния, белок, полученный из встречающегося в природе белка, или когда встречающийся в природе белок модифицирован или перепрограммирован, чтобы иметь определенное свойство.

Используемые в настоящем документе термины "синтетический" и "искусственный" используются взаимозаменяемо для обозначения белка или его домена, который характеризуется низкой идентичностью последовательности (например, идентичность последовательности составляет менее чем 50%) по отношению к встречающемуся в природе белку человека. Например, домены VPR и VP64 представляют собой синтетические домены трансаактивации.

Используемый в настоящем документе термин "сконструированный фактор транскрипции" или "eTF" относится к не встречающемуся в природе ДНК-связывающему белку или не встречающемуся в природе модулятору транскрипции, который был модифицирован или перепрограммирован для связывания со специфическим целевым сайтом связывания и/или включения модифицированного или замененного эффекторного домена транскрипции.

Используемый в настоящем документе термин "ДНК-связывающий домен" может использоваться для обозначения одного или нескольких ДНК-связывающих мотивов, таких как "цинковый палец" или основной мотив спираль-петля-спираль (bHLH), индивидуально или вместе как часть ДНК-связывающего белка.

Термины "активирующий транскрипцию домен", "домен активации транскрипции", "домен трансаактивации", "транс-активирующий домен" и "TAD" используются в настоящем документе взаимозаменяемо и относятся к домену белка, который вместе с ДНК-связывающим доменом может активировать транскрипцию с промотора, контактируя с транскрипционным аппаратом (например, с общими факторами транскрипции и/или РНК-полимеразой) либо напрямую, либо через другие белки, известные как ко-активаторы.

Термины "домен репрессора транскрипции" и "TRD" используются в настоящем документе взаимозаменяемо и относятся к домену белка, который в сочетании с ДНК-связывающим доменом может репрессировать транскрипцию с промотора, контактируя с транскрипционным механизмом (например, общие факторы транскрипции и/или РНК-полимераза) либо напрямую, либо через другие белки, известные как корепрессоры.

Термин "GRCh38.p12" относится к заплаточному релизу 12 Консорциума эталонных геномов Human Build 38 (GRCh38.p12), имеющему регистрационный номер сборки GenBank GCA_000001405.27 и датированному 21 декабря 2017 года.

Если не указано иное, все термины, использованные в настоящем документе, характеризуются тем же значением, что и для специалиста в настоящей области техники, и в практике настоящего изобретения будут использоваться традиционные способы молекулярной биологии, микробиологии и технологии рекомбинантных ДНК, которые находятся в пределах знания специалистов в настоящей области техники.

Сконструированные факторы транскрипции (eTF), которые активируют SCN1A

Согласно одному аспекту в настоящей заявке представлены eTF, которые способны повышать экспрессию гена альфа-субъединицы 1 натриевого канала с регулируемым напряжением (SCN1A) и увеличивать экспрессию его соответствующего белкового продукта Nav1.1. Ген SCN1A принадлежит к семейству генов, кодирующих субъединицы, используемые для сборки натриевых каналов. Эти каналы, которые транспортируют положительно заряженные ионы натрия в клетки, играют ключевую роль в способности клетки генерировать и передавать электрические сигналы. Ген SCN1A кодирует одну часть (альфа-субъединицу) натриевого канала, называемого Nav1.1. Эти каналы в основном находятся в головном мозге, где они контролируют поток ионов натрия в клетки. Каналы Nav1.1 участвуют в передаче сигналов от одной нервной клетки (или нейрона) к другой. Было обнаружено, что несколько мутаций в гене SCN1A вызывают генетическую эпилепсию с фебрильными припадками плюс (GEFS+), которая представляет собой спектр припадочных расстройств различной степени тяжести. Эти состояния включают в себя простые фебрильные (связанные с лихорадкой) припадки, которые начинаются в младенчестве и обычно прекращаются к 5 годам, и фебрильные припадки плюс (FS+). FS+ включает в себя фебрильные и другие типы припадков, в том числе не связанные с лихорадкой (афебрильные припадки), которые продолжаются во взрослом возрасте. Спектр GEFS+ также включает в себя другие состояния, такие как синдром Драве (также известный как тяжелая миоклоническая эпилепсия младенчества или SMEI), которые вызывают более серьезные припадки, которые длятся дольше и которые трудно контролировать. Эти повторяющиеся припадки (эпилепсия) могут со временем обостряться и часто сопровождаются снижением функции головного мозга. Многие другие мутации были связаны с семейной гемиплегической мигренью, формой мигренозной головной боли, которая передается в семье, и по меньшей мере одна мутация была связана с эффективностью некоторых противосудорожных лекарственных средств. Таким образом, представленный в настоящем документе eTF, который увеличивает экспрессию SCN1A, можно использовать для лечения множества заболеваний или нарушений, связанных с мутациями в канале Nav1.1.

Факторы транскрипции (TF) представляют собой белки, которые связываются с определенными последовательностями в геноме и контролируют экспрессию генов. Сконструированные факторы транскрипции или eTF, представленные в настоящем документе, которые активируют SCN1A, представляют собой не встречающиеся в природе белки, которые содержат ДНК-связывающий домен (DBD) и по

меньшей мере один домен, который является модулятором транскрипции, например, либо домен активации транскрипции (TAD), либо транскрипционный репрессорный домен (TRD). Согласно одному варианту осуществления eTF, который активирует SCN1A, может содержать DBD и TAD (например, TAD-DBD или DBD-TAD), причем DBD и TAD могут происходить из одного и того же белка или из разных белков. Согласно другому варианту осуществления eTF, который активирует SCN1A, может содержать DBD и два TAD, причем DBD и TAD происходят из одного и того же белка, DBD происходит из первого белка, и оба TAD происходят из второго белка, DBD и один TAD происходят из первого белка, а второй TAD происходит из второго белка, или DBD происходит из первого белка, один TAD происходит из второго белка, а второй TAD происходит из третьего белка (например, TAD1-DBD-TAD1, TAD1-DBD-TAD2, TAD1-TAD1-DBD, TAD1-TAD2-DBD, DBD-TAD1-TAD1 или DBD-TAD1-TAD2). Согласно другому варианту осуществления eTF, который активирует SCN1A, может содержать DBD и три TAD, причем DBD и TAD происходят из одного и того же белка, DBD происходит из первого белка, а TAD происходят из одного или нескольких разных белков, или причем DBD и все TAD происходят из разных белков, например, TAD_X-TAD_X-TAD_X-DBD, TAD_X-TAD_X-DBD-TAD_X, TAD_X-DBD-TAD_X-TAD_X или DBD-TAD_X-TAD_X-TAD_X, где каждый X выбран независимо и может быть таким же или отличным от одного или всех других TAD. Примеры включают в себя, например, TAD1-TAD1-DBD-TAD1, TAD1-TAD1-DBD-TAD2, TAD1-TAD2-DBD-TAD1, TAD1-TAD2-DBD-TAD2, TAD1-TAD2-DBD-TAD3, TAD1-DBD-TAD1-TAD1, TAD1-DBD-TAD2-TAD2, TAD1-DBD-TAD1-TAD2, TAD2-DBD-TAD1-TAD2, TAD1-DBD-TAD2-TAD3, TAD1-TAD1-TAD1-DBD, TAD1-TAD2-TAD2-DBD, TAD1-TAD2-TAD2-DBD, TAD1-TAD2-TAD3-DBD, DBD-TAD1-TAD1-TAD1, DBD-TAD1-TAD1-TAD2, DBD-TAD1-TAD2-TAD2 или DBD-TAD1-TAD2-TAD3 и т.д. Согласно другому варианту осуществления eTF, который активирует SCN1A, может содержать DBD и четыре TAD, причем DBD и TAD происходят из одного и того же белка, DBD происходит из первого белка, а TAD происходят из одного или нескольких разных белков, или причем DBD и все TAD происходят из разных белков, например, TAD_X-TAD_X-TAD_X-TAD_X-DBD, TAD_X-TAD_X-TAD_X-DBD-TAD_X, TAD_X-DBD-TAD_X-TAD_X-TAD_X или DBD-TAD_X-TAD_X-TAD_X-TAD_X, где каждый X выбирается независимо и может быть таким же или отличаться от одного или всех других TAD. Примеры включают в себя, например, TAD1-TAD1-DBD-TAD1-TAD1, TAD1-TAD1-DBD-TAD2-TAD2, TAD1-TAD2-DBD-TAD1-TAD2, TAD1-TAD2-DBD-TAD2-TAD1, TAD1-TAD2-DBD-TAD1-TAD3, TAD1-TAD3-DBD-TAD1-TAD2, TAD1-TAD2-DBD-TAD3-TAD4, TAD1-TAD1-TAD1-DBD-TAD2, TAD1-TAD2-TAD3-DBD-TAD4, TAD1-DBD-TAD1-TAD1-TAD2, TAD1-DBD-TAD2-TAD3-TAD4, TAD1-DBD-TAD1-TAD2-TAD3, TAD2-DBD-TAD1-TAD2-TAD3, TAD1-DBD-TAD2-TAD3-TAD4, TAD1-TAD1-TAD1-TAD1-DBD, TAD1-TAD2-TAD2-TAD3-DBD, TAD1-TAD2-TAD3-TAD4-DBD, DBD-TAD1-TAD1-TAD1-TAD1, DBD-TAD1-TAD1-TAD2-TAD2, DBD-TAD1-TAD2-TAD3-TAD4 или DBD-TAD1-TAD2-TAD3-TAD3 и т.д. Согласно одному варианту осуществления eTF, который активирует SCN1A, содержит DBD и два TAD, которые расположены на одном конце DBD (например, N-конце или C-конце), причем DBD происходит из первого белка, и оба TAD происходят из второго белка, или DBD происходит из первого белка, один TAD происходит из второго белка, а второй TAD происходит из третьего белка (например, TAD1-TAD1-DBD, TAD1-TAD2-DBD, DBD-TAD1-TAD1 или DBD-TAD1-TAD2). Согласно определенным вариантам осуществления DBD может представлять собой синтетическую конструкцию, содержащую домены из нескольких белков.

Согласно определенным вариантам осуществления DBD и TAD и/или два TAD могут быть напрямую конъюгированы, например, без промежуточной аминокислотной последовательности, DBD и TAD и/или два TAD могут быть конъюгированы с использованием пептидного линкера или их комбинаций. Согласно определенным вариантам осуществления DBD конъюгирован с TAD и/или один TAD конъюгирован со вторым TAD через линкер, содержащий 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 60, 70, 75, 80, 90 или 100 аминокислот или 1-5, 1-10, 1-20, 1-30, 1-40, 1-50, 1-75, 1-100, 5-10, 5-20, 5-30, 5-40, 5-50, 5-75, 5-100, 10-20, 10-30, 10-40, 10-50, 10-75, 10-100, 20-30, 20-40, 20-50, 20-75 или 20-100 аминокислот. В некоторых случаях DBD и TAD и/или два TAD конъюгированы через встречающиеся в природе промежуточные остатки, обнаруженные во встречающихся в природе белках, из которых происходят домены. Согласно другим вариантам осуществления DBD и TAD и/или два TAD конъюгированы через синтетическую или экзогенную линкерную последовательность. Подходящие линкеры могут быть гибкими, расщепляемыми, нерасщепляемыми, гидрофильными и/или гидрофобными. Согласно определенным вариантам осуществления DBD и TAD и/или два TAD могут быть слиты вместе через линкер, содержащий множество остатков глицина и/или серина. Примеры пептидных линкеров глицина/серина включают в себя [GS]_n, [GGGS]_n (SEQ ID NO: 179), [GGGGS]_n (SEQ ID NO: 180), [GGSG]_n (SEQ ID NO: 181), где n представляет собой целое число, равное или больше 1. Согласно определенным вариантам осуществления линкер, используемый для конъюгирования DBD и TAD и/или двух TAD, представляет собой GGSGGGSG (SEQ ID NO: 177). Согласно определенным вариантам осуществления линкер, используемый для конъюгирования DBD и TAD и/или двух TAD, представляет собой GGSGGGSGGGSGGGSG (SEQ ID NO: 178). Согласно определенным вариантам осуществления, когда DBD конъюгирован с двумя TAD, первый и второй TAD могут быть конъюгированы с DBD с одинаковыми или разными линкерами, или один TAD может быть конъюгирован с DBD с помощью линкера, а

другой TAD может быть конъюгирован напрямую с DBD (например, без промежуточной линкерной последовательности), или оба TAD могут быть непосредственно конъюгированы с DBD (например, без промежуточных линкерных последовательностей). Согласно определенным вариантам осуществления, когда DBD конъюгирован с двумя TAD на одном и том же конце (например, N-конце или C-конце), линкер, соединяющий два TAD, может быть таким же или отличаться от линкера, соединяющего TAD с DBD, или TAD могут быть конъюгированы друг с другом с помощью линкера, но TAD напрямую конъюгированы с DBD (например, без промежуточной линкерной последовательности), или TAD могут быть напрямую конъюгированы друг с другом (например, без промежуточных линкерных последовательностей), но TAD конъюгированы с DBD с помощью линкера. Согласно определенным вариантам осуществления представленные в настоящем документе eTF, которые активируют SCN1A, не содержат один или несколько тегов HA (например, SEQ ID NO: 171), расположенных между DBD и одним или несколькими TAD.

Представленные в настоящем документе eTF, которые активируют SCN1A, обладают свойствами, отличными от природных факторов транскрипции. Согласно определенным вариантам осуществления eTF, который активирует SCN1A, содержит DBD, полученный из встречающегося в природе белка, который был модифицирован таким образом, что DBD связывается с другим целевым сайтом по сравнению с встречающимся в природе белком, из которого он получен, и eTF, содержащий такой модифицированный DBD модулирует экспрессию другого гена (например, SCN1A) по сравнению с природным белком, из которого получен DBD (например, геном, отличным от SCN1A). Согласно другим вариантам осуществления eTF, представленный в настоящем документе, который активирует SCN1A, содержит TAD, полученный из встречающегося в природе белка, который был модифицирован таким образом, что eTF, содержащий такой модифицированный TAD, модулирует экспрессию другого гена (например, SCN1A) по сравнению с встречающимся в природе белком, из которого получен TAD (например, ген, отличный от SCN1A), и/или eTF, содержащий такой модифицированный TAD, по-разному модулирует экспрессию SCN1A (например, активирует или подавляет) по сравнению с встречающимся в природе белком, из которого получен TAD. Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A, содержит DBD, полученный из встречающегося в природе белка, и TAD, полученный из встречающегося в природе белка (одного и того же или разных белков), причем как DBD, так и TAD были модифицированы. Согласно таким вариантам осуществления DBD может связываться с другим целевым сайтом по сравнению с встречающимся в природе белком, из которого он был получен, eTF, содержащий такой модифицированный DBD и TAD, модулирует экспрессию другого гена (например, SCN1A) по сравнению с встречающимся в природе белком, из которых получены домены (например, ген(-ы), отличный от SCN1A), и/или eTF, содержащий такие модифицированные DBD и TAD, по-разному модулирует экспрессию SCN1A (например, активирует или подавляет) по сравнению с встречающимися в природе белками, из которых получены домены DBD и TAD.

ДНК-связывающие домены (DBD).

Приведенные в настоящем документе eTF, которые активируют SCN1A, могут содержать любой подходящей DBD, который связывается с представляющим интерес целевым сайтом (например, целевой сайт, который приводит к положительной регуляции SCN1A при связывании с помощью eTF, представленного в настоящем документе). Согласно определенным вариантам осуществления DBD может представлять собой синтетически разработанный DBD. Согласно другим вариантам осуществления DBD может быть получен из встречающегося в природе белка. Семейства DBD включают в себя основную спираль-петлю-спираль (bHLH) (например, c-Myc), основную лейциновую молнию (например, C/EBP), спираль-поворот-спираль (например, Oct-1) и "цинковые пальцы" (например, EGR1 или EGR3). Эти семейства демонстрируют широкий спектр специфичностей связывания ДНК и генных мишеней. Как предполагается в настоящем документе, любой из известных белков фактора транскрипции человека может служить в качестве белковой платформы для конструирования и/или перепрограммирования DBD для распознавания конкретного целевого сайта, что приводит к модуляции экспрессии эндогенного гена SCN1A. Согласно иллюстративным вариантам осуществления DBD, представленный в настоящем документе, содержит домен типа "цинковые пальцы", связывающий домен TALEN или комплекс gRNA/Cas.

Представленный в настоящем документе DBD может быть разработан для распознавания любого целевого сайта, который приводит к активации SCN1A. Согласно иллюстративным вариантам осуществления DBD разработан для распознавания местоположения в геноме и повышения экспрессии эндогенного гена SCN1A, связанного с eTF. Сайты связывания, способные модулировать экспрессию эндогенного гена SCN1A при связывании с помощью представленного в настоящем документе eTF, могут быть расположены в любом месте генома, что приводит к модуляции экспрессии гена SCN1A. Согласно различным вариантам осуществления сайт связывания может быть расположен на хромосоме, отличной от SCN1A, на той же хромосоме, что и SCN1A, выше против хода транскрипции от сайта начала транскрипции (TSS) гена SCN1A, ниже по ходу транскрипции от TSS гена SCN1A, проксимальнее от TSS гена SCN1A, дистальнее от TSS гена SCN1A, в кодирующей области гена SCN1A, в интроне гена SCN1A, ниже по ходу транскрипции от полиА-хвоста гена SCN1A, в промоторной последовательности, которая регулирует ген SCN1A, или в последовательности энхансера, которая регулирует ген SCN1A.

DBD может быть разработан для связывания с целевым сайтом связывания любой длины при условии, что он обеспечивает специфическое распознавание последовательности целевого сайта связывания посредством DBD, например, с минимальным или нулевым связыванием вне мишени. Согласно определенным вариантам осуществления целевой сайт связывания может модулировать экспрессию SCN1A при связывании с eTF на уровне, который является по меньшей мере 2-кратным, 5-кратным, 10-кратным, 20-кратным, 50-кратным, 75-кратным, 100-кратным, 150-кратным, 200-кратным, 250-кратным, 500-кратным или более по сравнению со всеми другими генами. Согласно определенным вариантам осуществления целевой сайт связывания может модулировать экспрессию SCN1A при связывании с eTF на уровне, который является по меньшей мере 2-кратным, 5-кратным, 10-кратным, 20-кратным, 50-кратным, 75-кратным, 100-кратным, 150-кратным, 200-кратным, 250-кратным, 500-кратным или более по сравнению с 40 ближайшими соседними генами (например, 40 генов, расположенных ближе всего к хромосоме, выше против хода транскрипции или ниже по ходу транскрипции от кодирующей последовательности SCN1A). Согласно определенным вариантам осуществления целевой сайт связывания может составлять по меньшей мере 5 п.н., 10 п.н., 15 п.н., 20 п.н., 25 п.н., 30 п.н., 35 п.н., 40 п.н., 45 п.н. или 50 п.н. или более. Конкретная длина сайта связывания определяется типом DBD в eTF. В общем, чем больше длина сайта связывания, тем выше специфичность связывания и модуляции экспрессии гена (например, более длинные сайты связывания характеризуются меньшим количеством нецелевых эффектов). Согласно определенным вариантам осуществления, eTF, содержащий DBD, распознающий более длинный целевой сайт связывания, имеет меньше нецелевых эффектов, связанных с неспецифическим связыванием (таких как, например, модуляция экспрессии нецелевого гена или отличного от SCN1A гена) относительно нецелевых эффектов, наблюдаемых с eTF, имеющим DBD, который связывается с более коротким целевым сайтом. В некоторых случаях снижение нецелевого связывания по меньшей мере в 1,2, 1,3, 1,4, 1,5, 2, 3, 4, 5, 6, 7, 8, 9 или 10 раз ниже по сравнению с сопоставимым eTF, имеющим DBD, который распознает более короткий целевой сайт связывания.

Согласно определенным вариантам осуществления представленный в настоящем документе DBD может быть модифицирован, чтобы иметь повышенную аффинность связывания, так что он связывается с целевым сайтом связывания в течение более длительного периода времени, так что TAD, конъюгированный с DBD, может рекрутировать больше факторов транскрипции и/или рекрутировать такой фактор транскрипции на более длительный период времени, чтобы оказывать большее влияние на уровень экспрессии эндогенного гена SCN1A. Согласно определенным вариантам осуществления DBD может быть модифицирован для увеличения его специфического связывания (или целевого связывания) с желаемым целевым сайтом и/или модифицирован для уменьшения его неспецифического или нецелевого связывания.

Согласно различным вариантам осуществления связывание между DBD или eTF и целевым сайтом связывания может быть определено с использованием различных способов. Согласно определенным вариантам осуществления специфическое связывание между DBD или eTF и целевым сайтом связывания может быть определено с использованием анализа определения изменения подвижности, анализа защиты от ДНКазы или любого другого способа *in vitro*, известного в настоящей области техники для анализа связывания белок-ДНК. Согласно другим вариантам осуществления специфическое связывание между eTF и целевым сайтом связывания можно определить с помощью функционального анализа, например, путем измерения экспрессии (РНК или белка) гена (например, SCN1A), когда целевой сайт связывания связан с eTF. Например, целевой сайт связывания может быть расположен выше против хода транскрипции от репортерного гена (такого как, например, eGFP) или гена SCN1A в векторе, содержащемся в клетке или интегрированном в геном клетки, причем клетка экспрессирует eTF. Альтернативно, экспрессирующий eTF вектор может быть введен в тип клеток, который в природе содержит ген SCN1A. Более высокие уровни экспрессии репортерного гена (или SCN1A) в присутствии eTF по сравнению с контролем (например, без eTF или eTF, который распознает другой целевой сайт) указывают на то, что DBD eTF связывается с целевым сайтом. Подходящие *in vitro* (например, неклеточные) системы транскрипции и трансляции также могут быть использованы аналогичным образом. Согласно определенным вариантам осуществления eTF, который связывается с целевым сайтом, может характеризоваться по меньшей мере 2-кратной, 3-кратной, 5-кратной, 10-кратной, 15-кратной, 20-кратной, 30-кратной, 50-кратной, 75-кратной, 100-кратной, 150-кратной или большей экспрессией репортерного гена или SCN1A по сравнению с контролем (например, без eTF или с eTF, который распознает другой целевой сайт).

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, размер которого составляет по меньшей мере 9 п.н., 12 п.н., 15 п.н., 18 п.н., 21 п.н., 24 п.н., 27 п.н., 30 п.н., 33 п.н. или 36 п.н.; более чем 9 п.н., 12 п.н., 15 п.н., 18 п.н., 21 п.н., 24 п.н., 27 п.н. или 30 п.н.; или 9-33 п.н., 9-30 п.н., 9-27 п.н., 9-24 п.н., 9-21 п.н., 9-18 п.н., 9-15 п.н., 9-12 п.н., 12-33 п.н., 12-30 п.н., 12-27 п.н., 12-24 п.н., 12-21 п.н., 12-18 п.н., 12-15 п.н., 15-33 п.н., 15-30 п.н., 15-27 п.н., 15-24 п.н., 15-21 п.н., 15-18 п.н., 18-33 п.н., 18-30 п.н., 18-27 п.н., 18-24 п.н., 18-21 п.н., 21-33 п.н., 21-30 п.н., 21-27 п.н., 21-24 п.н., 24-33 п.н., 24-30 п.н., 24-27 п.н., 27-33 п.н., 27-30 п.н. или 30-33 п.н. Согласно иллюстративным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который составляет

18-27 п.н., 18 п.н. или 27 п.н.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, расположенный на хромосоме 2. Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, расположенный на хромосоме 2 в пределах 110 т.п.н., 100 т.п.н., 90 т.п.н., 80 т.п.н., 70 т.п.н., 60 т.п.н., 50 т.п.н., 40 т.п.н., 30 т.п.н., 20 т.п.н., 10 т.п.н., 5 т.п.н., 4 т.п.н., 3 т.п.н., 2 т.п.н. или 1 т.п.н. выше против хода транскрипции или ниже по ходу транскрипции от TCC SCN1A. Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который расположен на хромосоме 2 в пределах 110 т.п.н. выше против хода транскрипции от TSS SCN1A. Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который расположен на хромосоме 2 в пределах 110 т.п.н. ниже по ходу транскрипции от TSS SCN1A. Согласно иллюстративным вариантам осуществления такие целевые сайты связывания составляют 18-27 п.н., 18 п.н. или 27 п.н.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который расположен на хромосоме 2 в пределах положений 166179652-165989571, в пределах положений 166128050-166127958, в пределах положений 166155414-166140590, в пределах положений 166179652-1661777272 или в пределах положений 1659990246-165989592 (все со ссылкой на GRCh38.p12). Согласно иллюстративным вариантам осуществления такие целевые сайты связывания составляют 18-27 п.н., 18 п.н. или 27 п.н.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который (i) составляет 18-27 п.н., 18 п.н. или 27 п.н., (ii) перекрывается с положением на хромосоме 2, выбранным из 166178880, 166177369, 166177362, 166177299, 166177299, 166155393, 166155264, 166149373, 166149176, 166149165, 166149118, 166148953, 166148565, 166142396, 166142391, 166142344, 166142239, 166141162, 166140928, 166140590, 165990076, 165989684, 165989571, 166155255, 166155099, 166148843, 166148361, 166142219, 166141090, 165990246, 165990193, 166149168, 166127991, 166128002, 166128037 или 166128025 (все со ссылкой на GRCh38.p12) и (iii) способен вызывать по меньшей мере 1,2-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, (i) связывается с целевым сайтом, содержащим любую из SEQ ID NO: 18, 25, 30, 31 или 35-66 или состоящим из нее, и (ii) способен вызывать по меньшей мере 1,2-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который (i) составляет 18-27 п.н., 18 п.н. или 27 п.н., (ii) перекрывается с положением на хромосоме 2, выбранным из 166155255, 166155099, 166148843, 166148361, 166142219, 166141090, 165990246, 165990193, 166149168, 166127991, 166128002, 166128037 или 166128025 (все со ссылкой на GRCh38.p12), и (iii) способен вызывать как минимум 2-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, (i) связывается с целевым сайтом, содержащим любую из SEQ ID NO: 18, 30, 31, 37, 38, 45, 47, 48, 49, 55, 61, 62 или 64 или состоящим из нее, и (ii) способен вызывать по меньшей мере 2-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который (i) составляет 18-27 п.н., 18 п.н. или 27 п.н., и (ii) перекрывается с положением на хромосоме 2, выбранным из 166149168, 166127991, 166128002, 166128037 или 166128025 (все со ссылкой на GRCh38.p12), и (iii) способен вызывать по меньшей мере 5-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, (i) связывается с целевым сайтом, содержащим любую из SEQ ID NO: 18, 30, 31, 37 или 38 или состоящим из нее, и (ii) способен вызывать по меньшей мере 5-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который (i) составляет 18-27 п.н., 18 п.н. или 27 п.н., (ii) перекрывается с положением на хромосоме 2, выбранным из 166128002, 166128037, или 166128025 (все со ссылкой на GRCh38.p12), и (iii) способен вызывать по меньшей мере 15-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, (i) связывается с целевым сайтом, содержащим любую из SEQ ID NO: 30, 37 или 38 или состоящим из нее, и (ii) способен вызывать по меньшей мере 15-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который (i) составляет 18-27 п.н., 18 п.н. или 27 п.н., (ii) перекрывается с положением на хромосоме 2, выбранным из 166128037 или 166128025 (все со ссылкой на GRCh38.p12), и (iii) способен вызывать по меньшей мере 20-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, (i) связывается с целевым сайтом, содержащим любую из SEQ ID NO: 30 или 38 или состоящим из нее, и (ii) способен вызывать по меньшей мере 20-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который (i) составляет 18-27 п.н., 18 п.н. или 27 п.н., (ii) перекрывается с положением на хромосоме 2 в положении 166128025 и (iii) способен вызывать по меньшей мере 25-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, (i) связывается с целевым сайтом, содержащим или состоящим из SEQ ID NO: 30, и (ii) способен вызывать по меньшей мере 25-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который (i) составляет 18-27 п.н., 18 п.н. или 27 п.н. и (ii) связывается с геномной областью, которая находится в пределах по меньшей мере 1 т.п.н., 750 п.н., 500 п.н., 400 п.н., 300 п.н., 200 п.н., 100 п.н. или 50 п.н. геномного местоположения, имеющего последовательность любой из SEQ ID NO: 18, 25, 30, 31 или 35-66. Согласно определенным вариантам осуществления целевой сайт связывания способен вызывать по меньшей мере 1,2-кратное, 2-кратное, 5-кратное, 15-кратное, 20-кратное или 25-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, который (i) составляет 18-27 п.н., 18 п.н. или 27 п.н. и (ii) связывается с областью генома, которая по меньшей мере частично перекрывается с геномным местоположением, имеющим последовательность любой из SEQ ID NO: 18, 25, 30, 31 или 35-66. Согласно определенным вариантам осуществления целевой сайт связывания способен вызывать по меньшей мере 1,2-кратное, 2-кратное, 5-кратное, 15-кратное, 20-кратное или 25-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, распознает целевой сайт связывания, имеющий любую из следующих последовательностей: SEQ ID NO: 18, 25, 30, 31 или 35-66. Согласно определенным вариантам осуществления целевой сайт связывания способен вызывать по меньшей мере 1,2-кратное, 2-кратное, 5-кратное, 15-кратное, 20-кратное или 25-кратное увеличение экспрессии SCN1A при связывании раскрытым в настоящем документе eTF.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, приводит по меньшей мере к 1,5-кратному, 2-кратному, 3-кратному, 4-кратному, 5-кратному, 6-кратному, 7-кратному, 8-кратному, 9-кратному, 10-кратному, 15-кратному, 20-кратному, 25-кратному, 50-кратному, 100-кратному или более или по меньшей мере 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более активации экспрессии SCN1A (РНК SCN1A и/или белок Nav1.1) в клетке или *in vivo* по сравнению с контролем (например, без eTF или с eTF, который не распознает целевой сайт). Согласно различным вариантам осуществления активация экспрессии SCN1A может быть обнаружена с помощью способов ПЦР, вестерн-блоттинга или иммуноанализов.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, связывается с целевым сайтом, который способен увеличивать экспрессию SCN1A по меньшей мере в 1,2 раза, 1,3 раза, 1,4 раза, 1,5 раз, 1,6 раз, 1,7 раз, 1,8 раз, 1,9 раз, 2 раза, 3 раза, 4 раза, 5 раз, 8 раз, 10 раз, 12 раз, 15 раз, 18 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз, 75 раз, 100 раз или больше или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более относительно контроля в анализе активации транскрипции. Иллюстративный анализ активации транскрипции SCN1A представлен в настоящем документе в примере 3. Вкратце, HEK293 трансфицируют плазмидой, несущей eTF или контрольную репортерную конструкцию eGFP. Через 48 часов после трансфекции клетки собирают, выделяют РНК и подвергают обратной транскрипции, а полученные образцы кДНК анализируют с помощью кПЦР (например, с использованием праймеров, имеющих SEQ ID NO: 185 и 186) для количественного определения содержания эндогенного транскрипта SCN1A. GAPDH можно использовать в качестве эталонного гена для определения относительных уровней экспрессии SCN1A.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, кото-

рый активирует SCN1A, характеризуется минимальными нецелевыми эффектами, например, нецелевыми эффектами, связанными с неспецифическим связыванием, такими как, например, модуляция экспрессии нецелевого гена или отличного от SCN1A гена. Согласно одному варианту осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, специфически активирует SCN1A по сравнению с контролем, по меньшей мере в 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз или 50 раз больше, чем экспрессия, произведенная eTF для одного или нескольких нецелевых генов по сравнению с контролем. Согласно иллюстративному варианту осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, специфически активирует транскрипцию гена SCN1A по сравнению с контролем, по меньшей мере, в 15 раз больше, чем транскрипция 40 ближайших соседних генов (например, 40 ближайших генов, расположенных к кодирующей последовательности SCN1A на хромосоме 2), производимая eTF относительно контроля, например, PLA2R1, ITGB6, RBMS1, TANK, PSMD14, TBR1, SLC4A10, DPP4, FAP, IFIH1, GCA, FIGN, GRB14, COBLL1, SLC38A11, SCN3A, SCN2A, CSRNP3, GALNT3, TTC21B, SCN9A, SCN7A, B3GALT1, STK39, CERS6, NOSTRIN, SPC25, ABCB11, DHRS9, BBS5, KLHL41, FASTKD1, PPIG, CCDC173, PHOSPHO2, KLHL23, SSB, METTL5, UBR3 и MYO3B (смотрите табл. 14). Согласно различным вариантам осуществления активация транскрипции гена SCN1A может быть обнаружена с помощью способов ПЦР.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, способен снижать частоту припадков в анализе гипертермических припадков (HTS) в модели синдрома Драве на мышцах Scn1a^{tm1^{ke}a}. Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF способен снижать частоту припадков при температуре 42,6°C в анализе HTS по меньшей мере в 1,2 раза, 1,3 раза, 1,4 раза, 1,5 раз, 1,6 раз, 1,7 раз, 1,8 раз, 1,9 раз, 2,0 раза или более или по меньшей мере на 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100% или более по сравнению с контролем (например, обработанный PBS или воздействие вектором AAV, содержащим последовательность, кодирующую eGFP). Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF способен снижать частоту припадков при температуре 42,6°C в анализе HTS, так что по меньшей мере у 60%, 62%, 65%, 70%, 75%, 76%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% мышей, участвовавших в анализе, не наблюдалось припадков при температуре 42,6°C. Иллюстративный анализ HTS описан в настоящем документе в примере 6. Вкратце, пометам детенышей, полученным от самцов мышей Scn1a +/-, скрещенных с самками мышей C57Bl/6J, может быть введен вектор AAV9, кодирующий eTF, который активирует SCN1A, как предусмотрено в настоящем документе, или контрольный вектор, кодирующий eGFP, посредством двустороннего ICV в P1. Мышам можно вводить дозу ~1,0E10-5,0E12 копий геномов/мышь. Анализ HTS проводят на гетерозиготных мышках P26-P28 SCN1A и SCN1A дикого типа на смешанном фоне 129Stac X C57Bl/6 путем повышения температуры тела мышей (в контролируемых условиях и с контролем температуры тела) на ~0,5°C каждые 2 минуты до начала первого тонико-клонического припадка, сопровождающегося потерей положения тела, или до достижения температуры тела 43°C. Считается, что у мыши не было припадка, если не было обнаружено припадка с потерей положения тела в течение всего периода эксперимента.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, способен увеличивать выживаемость мыши, гетерозиготной по SCN1A, например, линии мышей Scn1a^{tm1^{ke}a}. Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF способен увеличивать выживаемость гетерозиготных мышей SCN1A при P100 по меньшей мере в 1,2 раза, 1,3 раза, 1,4 раза, 1,5 раз, 1,6 раз, 1,7 раз, 1,8 раз, 1,9 раз, 2,0 раза или более, или по меньшей мере на 20%, 30%, 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 100% или более, как по сравнению с контролем (например, обработанным PBS или обработанным вектором AAV, содержащим последовательность, кодирующую eGFP). Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF способен увеличивать выживаемость гетерозиготных мышей SCN1A при P100, так что по меньшей мере 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% мышей, задействованных в анализе, все еще живы при P100. Иллюстративный анализ выживаемости описан в настоящем документе в примере 7. Вкратце, пометам детенышей, полученным от самцов мышей Scn1a +/-, скрещенных с самками мышей C57Bl/6J, может быть введен вектор AAV9 посредством двустороннего ICV в точке P1. Мышам можно вводить дозу ~1,0E10-5,0E12 копий геномов/мышь. Определяют количество мышей, доживших до P100.

Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A, может содержать DBD из белка типа "цинковые пальцы", полученного из белка типа "цинковые пальцы", или нуклеаза представляет собой инактивированный белок типа "цинковые пальцы". "Цинковый палец" представляет собой небольшой структурный мотив белка, который характеризуется координацией одного или нескольких ионов цинка (Zn²⁺) для стабилизации складки. Домены типа "цинковые пальцы" (Znf) представляют собой относительно небольшие белковые мотивы, которые содержат несколько пальцевидных выступов, которые создают tandemные контакты с целевым участком ДНК. Модульная природа мотива "цинкового пальца" позволяет связывать большое количество комбинаций последовательностей ДНК с высокой степенью аффинности и специфичности, и поэтому

идеально подходит для создания белка, который может быть нацелен на определенные последовательности ДНК и связывать их. Многие сконструированные наборы "цинковых пальцев" основаны на домене типа "цинковые пальцы" мышинового фактора транскрипции Zif268. Zif268 содержит три отдельных мотива "цинковые пальцы", которые вместе связывают последовательность из 9 п.н. с высокой аффинностью. Было идентифицировано большое количество белков типа "цинковые пальцы", и они были охарактеризованы на различные типы на основе структуры, которая дополнительно описана в настоящем документе. Любой такой белок типа "цинковые пальцы" применим в связи с описанными в настоящем документе DBD.

Доступны различные способы конструирования белков типа "цинковые пальцы". Например, описаны способы создания белков типа "цинковые пальцы" для связывания с представляющей интерес целевой последовательностью ДНК, смотрите, например, Liu Q, et al., Design of polydactyl zinc-finger proteins for unique addressing within complex genomes, Proc Natl Acad Sci USA. 94 (11): 5525-30 (1997); Wright DA et al., Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly, Nat Protoc. 2006; 1(3): 1637-52 и CA Gersbach and T Gaj, Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies, Am Chem Soc 47: 2309-2318 (2014). Кроме того, общедоступны различные веб-инструменты для создания белков типа "цинковые пальцы" для связывания с представляющей интерес целевой последовательностью ДНК, смотрите, например, программные инструменты для разработки нуклеаз типа "цинковые пальцы" и веб-сайт анализа сконструированных данных генома от OmicX, доступный в Интернете по адресу omictools.com/zfns-category; и веб-сайт дизайна Zinc Finger Tools от Scripps, доступный в Интернете по адресу scripps.edu/barbas/zfdesign/zfdesignhome.php. Кроме того, доступны различные коммерчески доступные услуги по разработке белков типа "цинковые пальцы" для связывания с представляющей интерес целевой последовательностью ДНК, смотрите, например, коммерчески доступные услуги или наборы, предлагаемые Creative Biolabs (в Интернете по адресу creative-biolabs.com/Design-and-Synthesis-of-Artificial-Zinc-Finger-Proteins.html), Modular Assembly Kit Zinc Finger Consortium, доступный от Addgene (в Интернете по адресу addgene.org/kits/zfc-modular-assembly/), или the CompoZr Custom ZFN Service от Sigma Aldrich (в Интернете по адресу sigmaaldrich.com/life-science/zinc-finger-nuclease-technology/custom-zfn.html).

Согласно определенным вариантам осуществления представленные в настоящем документе eTF, которые активируют SCN1A, содержат DBD, содержащий один или несколько "цинковых пальцев", или получены из DBD белка типа "цинковые пальцы". В некоторых случаях DBD содержит несколько "цинковых пальцев", причем каждый "цинковый палец" связан с другим "цинковым пальцем" или другим доменом либо на своем N-конце, либо на C-конце, или на обоих через аминокислотный линкер. В некоторых случаях представленный в настоящем документе DBD содержит один или несколько "цинковых пальцев" из одного или нескольких типов "цинковых пальцев", описанных в табл. 9. В некоторых случаях представленный в настоящем документе DBD содержит множество структур или мотивов "цинковых пальцев" или множество "цинковых пальцев", имеющих одну или несколько из SEQ ID NO: 152-167 или любую их комбинацию. Согласно определенным вариантам осуществления DBD содержит X-[ZF-X]_n и/или [X-ZF]_n-X, причем ZF представляет собой домен типа "цинковые пальцы", содержащий любой из мотивов, перечисленных в табл. 9 (например, любой из SEQ ID NO: 136-146), X представляет собой аминокислотный линкер, содержащий 1-50 аминокислот, и n представляет собой целое число от 1 до 15, например, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 или 15, причем каждый ZF может независимо иметь ту же последовательность или последовательность, отличную от других последовательностей ZF в DBD, и причем каждый линкер X может независимо иметь такую же последовательность или последовательность, отличную от других последовательностей X в DBD. Каждый "цинковый палец" может быть связан с другой последовательностью, "цинковым пальцем" или доменом на своем C-конце, N-конце или с обоими. В DBD каждый линкер X может быть идентичным по последовательности, длине и/или свойству (например, гибкости или заряду) или отличаться по последовательности, длине и/или свойству. В некоторых случаях два или более линкера могут быть идентичными, в то время как другие линкеры различны. Согласно иллюстративным вариантам осуществления линкер может быть получен или быть производным от последовательностей, соединяющих "цинковые пальцы", обнаруженных в одном или нескольких встречающихся в природе белках типа "цинковые пальцы", представленных в табл. 9. Согласно другим вариантам осуществления подходящие линкерные последовательности включают в себя, например, линкеры из 5 или более аминокислот в длину. См. также патент США № 6479626; 6903185 и 7153949 для иллюстративных линкерных последовательностей из 6 или более аминокислот в длину, каждая из которых полностью включена в настоящий документ. Предлагаемые в настоящем документе белки DBD могут включать в себя любую комбинацию подходящих линкеров между отдельными "цинковыми пальцами" белка. Описанные в настоящем документе белки DBD могут включать в себя любую комбинацию подходящих линкеров между отдельными "цинковыми пальцами" белка.

Согласно определенным вариантам осуществления представленные в настоящем документе eTF, которые активируют SCN1A, содержат DBD, содержащий один или несколько классических "цинковых пальцев". Классический "цинковый палец" C2H2 содержит два цистеина в одной цепи и два остатка гис-

тидина в другой цепи, координируемых ионом цинка. Классический домен типа "цинковый палец" содержит два β -листа и одну α -спираль, причем α -спираль взаимодействует с молекулой ДНК и формирует основу связывания DBD с целевым сайтом и может называться "спиралью распознавания". Согласно иллюстративным вариантам осуществления спираль распознавания "цинковых пальцев" содержит по меньшей мере одну аминокислотную замену в положении -1, 2, 3 или 6, тем самым изменяя специфичность связывания домена типа "цинковые пальцы". Согласно другим вариантам осуществления представленный в настоящем документе DBD содержит один или несколько неклассических "цинковых пальцев", например, C2-H2, C2-CN и C2-C2.

Согласно другому варианту осуществления представленный в настоящем документе eTF, который активирует SCN1A, содержит DBD, содержащий мотив типа "цинковые пальцы", имеющий следующую структуру: LEPGEKP - [YKPCGKSF X HQRTH TGEKP]_n - YKPCGKSF X HQRTH - TGKKT (SEQ ID NO: 147), где n представляет собой целое число от 1 до 15, например, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 или 15, и каждый X независимо представляет собой последовательность распознавания (например, спираль распознавания), способную связываться с 3 п.н. целевой последовательности. Согласно иллюстративным вариантам осуществления n равно 3, 6 или 9. Согласно особенно предпочтительному варианту осуществления n равно 6. Согласно различным вариантам осуществления каждый X может независимо иметь одинаковую аминокислотную последовательность или отличную аминокислотную последовательность по сравнению с другими последовательностями X в DBD. Согласно иллюстративному варианту осуществления каждый X представляет собой последовательность, содержащую 7 аминокислот, которая была разработана для взаимодействия с 3 п.н. представляющего интерес целевого сайта связывания с использованием Zinger Finger Design Tool от Scripps, расположенного в Интернете по адресу scripps.edu/barbas/zfdesign/zfdesignhome.php.

Поскольку каждый "цинковый палец" в DBD распознает 3 п.н., количество "цинковых пальцев", включенных в DBD, сообщает длину сайта связывания, распознаваемого DBD, например, DBD с 1 "цинковым пальцем" распознает целевой сайт связывания, имеющий 3 п.н., DBD с 2 "цинковыми пальцами" распознает целевой сайт связывания, имеющий 6 п.н., DBD с 3 "цинковыми пальцами" распознает целевой сайт связывания, имеющий 9 п.н., DBD с 4 "цинковыми пальцами" распознает целевой сайт связывания, имеющий 12 п.н., DBD с 5 "цинковыми пальцами" распознает целевой сайт связывания, имеющий 15 п.н., DBD с 6 "цинковыми пальцами" распознает целевой сайт связывания, имеющий 18 п.н., DBD с 9 "цинковыми пальцами" распознает целевой сайт связывания, имеющий 27 п.н., и т. д. В общем, DBD, которые распознают более длинные целевые сайты связывания, будут проявлять большую специфичность связывания (например, меньшее нецелевое связывание или неспецифическое связывание).

Согласно другим вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A, содержит DBD, который происходит из встречающегося в природе белка типа "цинковые пальцы" путем выполнения одной или нескольких аминокислотных замен в одной или нескольких спиральных распознавания доменов типа "цинковые пальцы" для изменения специфичности связывания DBD (например, изменение целевого сайта, распознаваемого DBD). Представленный в настоящем документе DBD может быть получен из любого встречающегося в природе белка типа "цинковые пальцы". Согласно различным вариантам осуществления такой DBD может быть получен из белка типа "цинковые пальцы" любого вида, например, мыши, крысы, человека и т.д. Согласно иллюстративному варианту осуществления представленный в настоящем документе DBD получен из белка типа "цинковые пальцы" человека. Согласно определенным вариантам осуществления представленный в настоящем документе DBD получен из встречающегося в природе белка, перечисленного в табл. 9. Согласно иллюстративному варианту осуществления представленный в настоящем документе белок DBD получен из белка типа "цинковые пальцы" EGR человека, например, EGR1, EGR2, EGR3 или EGR4.

Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A, содержит DBD, который происходит из встречающегося в природе белка, путем модификации DBD для увеличения количества доменов типа "цинковые пальцы" в белке DBD путем повторения одного или нескольких "цинковых пальцев" в пределах DBD встречающегося в природе белка. Согласно определенным вариантам осуществления такие модификации включают в себя дубликацию, утроение, учетверение или дальнейшее умножение "цинковых пальцев" в пределах DBD встречающегося в природе белка. В некоторых случаях умножается один "цинковый палец" из DBD человеческого белка, например, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 или более копий одного и того же мотива "цинкового пальца" повторяется в DBD eTF. В некоторых случаях набор "цинковых пальцев" из DBD встречающегося в природе белка умножается, например, набор из 3 "цинковых пальцев" из DBD встречающегося в природе белка дублируется, чтобы получить eTF, имеющий DBD с 6 "цинковыми пальцами", трехкратно повторяется для получения DBD eTF с 9 "цинковыми пальцами" или повторяется четыре раза для получения DBD eTF с 12 "цинковыми пальцами" и т.д. В некоторых случаях набор "цинковых пальцев" из DBD встречающегося в природе белка частично реплицируется для формирования DBD eTF, имеющего большее количество "цинковых пальцев", например, DBD eTF содержит четыре "цинковых пальца", причем "цинковые пальцы" представляют собой одну копию первого "цинкового пальца", одну копию второго "цинкового пальца" и две копии третьего "цинкового пальца" из встречающегося в природе белка,

всего четыре "цинковых пальца" в DBD eTF. Затем такие DBD дополнительно модифицируют, производя одну или несколько аминокислотных замен в одной или нескольких спиральных распознавания доменов типа "цинковые пальцы", чтобы изменить специфичность связывания DBD (например, изменяя целевой сайт, распознаваемый DBD). Согласно иллюстративным вариантам осуществления DBD получают из встречающегося в природе человеческого белка, такого как белок типа "цинковые пальцы" EGR человека, например, EGR1, EGR2, EGR3 или EGR4.

EGR1 и EGR3 человека характеризуются трехпальцевым DBD типа "цинковые пальцы" C2H2. Общие правила связывания для "цинковых пальцев" предусматривают, что все три пальца взаимодействуют со своей родственной последовательностью ДНК с аналогичной геометрией, используя одни и те же аминокислоты в альфа-спирали каждого "цинкового пальца" для определения специфичности или распознавания последовательности целевого сайта связывания. Такие правила связывания позволяют изменять DBD EGR1 или EGR3 для создания DBD, который распознает желаемый целевой сайт связывания. В некоторых случаях спираль распознавания ДНК из 7 аминокислот в мотиве "цинкового пальца" EGR1 или EGR3 модифицируется в соответствии с опубликованными правилами разработки "цинковых пальцев". Согласно определенным вариантам осуществления каждый "цинковый палец" в DBD с тремя пальцами EGR1 или EGR3 модифицируется, например, путем изменения последовательности одной или нескольких спиралей распознавания и/или путем увеличения количества "цинковых пальцев" в DBD. Согласно определенным вариантам осуществления EGR1 или EGR3 перепрограммированы для распознавания целевого сайта связывания по меньшей мере из 9, 12, 15, 18, 21, 24, 27, 30, 33, 36 или более пар оснований в желаемом целевом сайте. Согласно определенным вариантам осуществления такой DBD, полученный из EGR1 или EGR3, содержит по меньшей мере 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 или более "цинковых пальцев". Согласно иллюстративному варианту осуществления один или несколько "цинковых пальцев" в DBD содержат по меньшей мере одну аминокислотную замену в положении -1, 2, 3 или 6 спирали распознавания.

Согласно различным вариантам осуществления eTF, который активирует SCN1A, содержащий DBD, полученный из EGR1 или EGR3, характеризуется специфичностью связывания ДНК, которая отличается от специфичности связывания встречающихся в природе EGR1 или EGR3, например, DBD распознает целевой сайт связывания, имеющий последовательность, отличную от последовательности сайта связывания, распознаваемой модифицированными EGR1 или EGR3: (GCG(T/G)GGGCG) (SEQ ID NO: 182).

Согласно другим вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A, содержит DBD, который представляет собой комплекс gRNA/Cas. CRISPR (кластеризованные регулярные промежуточные короткие палиндромные повторы)/Cas9 представляет собой инструмент для редактирования генома, который позволяет направленно воздействовать на сайт-специфический геном. Система CRISPR/Cas типа II представляет собой прокариотическую систему адаптивного иммунного ответа, которая использует некодирующие РНК, чтобы направлять нуклеазу Cas9 для индукции сайт-специфического расщепления ДНК. Система CRISPR/Cas9 была использована для создания простого, программируемого с помощью РНК способа, обеспечивающего редактирование генома в клетках млекопитающих. Одна направляющая РНК (sgRNA) может быть создана для направления нуклеазы Cas9 в конкретное место генома, которое затем связывается комплексом gRNA/Cas9. gRNA может быть сконструирована для связывания с представляющим интерес целевым сайтом с использованием различных способов и инструментов. Например, способы создания gRNA для связывания с представляющей интерес целевой последовательностью ДНК описаны в Aach, et al. Flexible algorithm for identifying specific Cas9 targets in genomes. *BioRxiv*, Cold Spring Harbor Labs, doi: <http://dx.doi.org/10.1101/005074> (2014); Bae, et al. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 30(10): 1473-1475 (2014); Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotech* 34, 184-191 (2016); Gratz, et al. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics*. 196(4):961-971 (2014); Heigwer, et al. E-CRISP: fast CRISPR target site identification. *Nat Methods*. 11(2):122-123 (2014); Ma, et al. A guide RNA sequence design platform for the CRISPR/Cas9 system for model organism genomes. *Biomed Res Int*. doi:<http://doi.org/10.1155/2013/270805> (2013); Montague, et al. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res*. 42(W1):W401-W407 (2014); Liu, et al. CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. *Bioinformatics*. 31(22):3676-3678 (2015); Ran, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*. 520(7546):186-191 (2015); Wu, et al. Target specificity of the CRISPR-Cas9 system. *Quant Biol*. 2(2):59-70 (2015); Xiao, et al. CasOT: a genome-wide Cas9/gRNA off-target searching tool. *Bioinformatics*. 30(8):1180-1182 (2014); Zetsche, et al. Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas System. *Cell*. 163(3):759-771 (2015). Кроме того, общедоступны различные веб-инструменты для создания gRNA для связывания с представляющей интерес целевой последовательностью ДНК, смотрите, например, инструмент разработки CRISPR gRNA, доступный от AUTM в Интернете по адресу atum.bio/eCommerce/cas9/input?multipleContacts=false; инструмент разработки CRISPR/i gRNA, доступ-

ный от Broad Institute в Интернете по адресу portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design-crisprai; инструмент разработки E-CRISP, доступный в Немецком онкологическом исследовательском центре DKFZ, доступный в Интернете по адресу e-crisp.org/E-CRISP/; и инструмент Knockout Guide Design, доступный от Synthego в Интернете по адресу design.synthego.com/#/. Кроме того, доступны различные коммерчески доступные сервисы для создания gRNA для связывания с представляющей интерес целевой последовательностью ДНК, смотрите, например, коммерчески доступные услуги, предлагаемые IDT (в Интернете по адресу idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE), ThermoFisher (в Интернете по адресу thermofisher.com/order/custom-oligo/crispr) и GenScript (в Интернете по адресу gen-script.com/gRNA-design-tool.html).

Согласно иллюстративным вариантам осуществления DBD, который представляет собой комплекс gRNA/Cas, содержит дезактивированный нуклеазой белок Cas или dCas, такой как, например, dCas9, такой как дезактивированный нуклеазой *Staphylococcus aureus* (dSaCas9) или дезактивированный нуклеазой *Streptococcus pyogenes* Cas9 (dSpCas9). gRNA представлена в виде последовательности, содержащей нацеливающую область, которая нацеливает комплекс gRNA/Cas на желаемый целевой сайт, и каркасную область, которая облегчает взаимодействие с белком Cas. Любой подходящий каркас gRNA может использоваться в связи с представленными в настоящем документе gRNA. Согласно иллюстративному варианту осуществления gRNA представляет собой одиночную gRNA или sgRNA и содержит следующую каркасную последовательность: 5'-GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTGGCGAGA-3' (SEQ ID NO: 183). Нацеливающая область направляющей РНК присоединяется к 5'-концу каркасной последовательности с образованием полной sgRNA. Согласно определенным вариантам осуществления gRNA и белок dCas могут экспрессироваться из одной и той же кассеты экспрессии. Согласно определенным вариантам осуществления промотор U6 используется для экспрессии gRNA. Согласно другим вариантам осуществления gRNA может быть экспрессирована в клетке, которая была сконструирована для стабильной экспрессии белка dCas-TAD, например, либо путем стабильной интеграции dCas в геном, либо на плазмиде, которая стабильно поддерживается вне хромосом.

Согласно другим вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A, может содержать DBD из TALEN, полученный из TALEN, или который представляет собой инактивированную нуклеазой TALEN. Подобные активатору транскрипции эффекторным нуклеазы (TALEN) представляют собой рестрикционные ферменты, которые содержат DBD и домен нуклеазы, которые могут быть сконструированы для разрезания определенных последовательностей ДНК. TALEN создаются путем конъюгирования эффекторного ДНК-связывающего домена TAL с доменом расщепления ДНК (например, нуклеазой). Эффекторы, подобные активатору транскрипции (TALE), могут быть сконструированы так, чтобы связываться с желаемой целевой последовательностью ДНК, тем самым направляя домен нуклеазы в определенное место.

Эффекторы TAL представляют собой бактериальные белки из бактерий *Xanthomonas*. ДНК-связывающий домен содержит повторяющуюся высококонсервативную последовательность из 33-34 аминокислот с расходящимися 12-ой и 13-ой аминокислотами. Эти два положения, называемые Repeat Variable Di-residue (RVD), сильно изменчивы и показывают сильную корреляцию со специфическим распознаванием нуклеотидов. Эта прямая взаимосвязь между аминокислотной последовательностью и распознаванием ДНК позволяет конструировать DBD, которые специфически нацелены на желаемую последовательность, путем выбора комбинации повторяющихся сегментов, содержащих соответствующие RVD.

Существуют разные способы разработки TALE. Например, способы разработки TALE для связывания с представляющей интерес целевой последовательностью ДНК описаны в T. Cermak et al., *Nucleic Acids Research*. 39 (12):e82 (2011); F. Zhang F et al., *Nature Biotechnology*. 29 (2):149-53 (2011); R. Morbitzer et al., *Nucleic Acids Research*. 39 (13):5790-9 (2011); T. Li et al., *Nucleic Acids Research*. 39 (14):6315-25 (2011); R. Geissler et al., *PLOS One*. 6(5): e19509 (2011) и E. Weber et al., *PLOS One*. 6(5): e19722 (2011). Кроме того, общедоступны различные веб-инструменты для разработки TALE для связывания с представляющей интерес целевой последовательностью ДНК, смотрите, например, E-Talen, доступный в Интернете по адресу e-talen.org/E-TALEN/TAL и инструмент Effector Nucleotide Targeter 2.0, доступный в Интернете по адресу tall-nt.cac.cornell.edu/node/add/single-tale. Кроме того, доступны различные коммерчески доступные сервисы для создания TALE для связывания с представляющей интерес целевой последовательностью ДНК, смотрите, например, коммерчески доступные услуги, предлагаемые OmicX (в Интернете по адресу omictools.com/), Addgene (в Интернете по адресу addgene.org/talen/guide/ или ThermoFisher (в Интернете: thermofisher.com/us/en/home/life-science/genome-editing/genart-tals/tal-design-tool.html). Кроме того, общедоступная программа (DNAWorks) может использоваться для создания олигонуклеотидов, подходящих для сборки TALE, смотрите, например, D. Hoover D *Methods in Molecular Biology*. 852: 215-23 (2012).

Домены модуляции транскрипции.

Приведенные в настоящем документе eTF, которые активируют SCN1A, могут содержать любой подходящий домен, который способен рекрутировать один или несколько белковых факторов, которые

могут модулировать транскрипцию (например, РНК-полимеразу II, СВР/p300, CREB или KRAB) или уровень экспрессии гена от представляющего интерес гена, когда eTF связан с целевым сайтом через DBD (например, DBD с "цинковыми пальцами", gRNA/Cas DBD или TALE DBD). Согласно определенным вариантам осуществления такой домен рекрутирует белковые факторы, которые увеличивают уровень транскрипции или экспрессии представляющего интерес гена, и представляет собой домен активации транскрипции (TAD). Согласно другим вариантам осуществления такой домен рекрутирует белковые факторы, которые снижают уровень транскрипции или экспрессии представляющего интерес гена, и представляет собой домен репрессора транскрипции (TRD). Согласно определенным вариантам осуществления домен модуляции транскрипции (TAD или TRD) может представлять собой синтетически сконструированный домен. Согласно другим вариантам осуществления домен модуляции транскрипции (TAD или TRD) может происходить из встречающегося в природе белка, например, фактора транскрипции, коактиватора транскрипции, корепрессора транскрипции или белка-сайленсера. Согласно различным вариантам осуществления домен модуляции транскрипции (TAD или TRD) может происходить из белка любого вида, например, мыши, крысы, обезьяны, вируса или человека.

Согласно одному иллюстративному варианту осуществления TAD, подходящий для применения в представленных в настоящем документе eTF, которые активируют SNC1A, происходит из вирусного белка. Иллюстративные TAD, полученные из вирусных белков, включают в себя, например, TAD-домен VP64 (SEQ ID NO: 133), VPR (SEQ ID NO: 132), VP16, VP128, p65, p300 или любой его функциональный фрагмент или вариант, или последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к ней.

Согласно другому иллюстративному варианту осуществления TAD, подходящий для применения в представленных в настоящем документе eTF, которые активируют SCN1A, происходит из белка человека. Иллюстративные TAD, полученные из белков человека, включают в себя, например, TAD-домен взаимодействующего с СВР/p300 трансаактиватора 2 (CITED2) (SEQ ID NO: 134), взаимодействующего с СВР/p300 трансаактиватора 4 (CITED4) (SEQ ID NO: 135), EGR1 (SEQ ID NO: 176), CREB3 (SEQ ID NO: 224) или EGR3 (SEQ ID NO: 175), или любой их функциональный фрагмент или вариант, или последовательность, характеризующуюся идентичностью, составляющей по меньшей мере 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к ней.

Согласно определенным вариантам осуществления eTF, который активирует SCN1A, содержит DBD типа "цинковые пальцы", который конъюгирован с доменом активации транскрипции или TAD. Согласно различным вариантам осуществления DBD типа "цинковые пальцы" может быть конъюгирован с TAD из вирусного белка, такого как VP64 или VPR, или с TAD из человеческого белка, такого как CITED2, CITED4 или CREB3. Согласно определенным вариантам осуществления DBD типа "цинковые пальцы", полученный из человеческого белка, например, EGR1 или EGR3, конъюгирован с TAD, полученным из человеческого белка, например, CITED2, CITED4 или CREB3. Согласно определенным вариантам осуществления DBD типа "цинковые пальцы", полученный из человеческого белка, например, EGR1 или EGR3, конъюгирован с VP64 или VPR TAD. Согласно определенным вариантам осуществления синтетический DBD типа "цинковые пальцы" или DBD типа "цинковые пальцы", характеризующийся идентичностью последовательности, составляющей менее чем 75%, по отношению к человеческому белку, например, EGR1 или EGR3, конъюгирован с TAD, полученным из человеческого белка, например, CITED2, CITED4 или CREB3. Согласно определенным вариантам осуществления синтетический DBD типа "цинковые пальцы" или DBD типа "цинковые пальцы", характеризующийся идентичностью последовательности, составляющей менее чем 75%, по отношению к человеческому белку, например, EGR1 или EGR3, конъюгирован с VP64 или VPR TAD.

Согласно определенным вариантам осуществления белок dCas конъюгирован с TAD. Согласно различным вариантам осуществления dCas9 может быть конъюгирован с TAD из вирусного белка, такого как VP64 или VPR, или с TAD из человеческого белка, такого как CITED2, CITED4 или CREB3. Согласно иллюстративным вариантам осуществления dCas9 конъюгирован с VP64 или VPR TAD.

Согласно определенным вариантам осуществления белок TALE конъюгирован с TAD. Согласно различным вариантам осуществления TALE может быть конъюгирован с TAD из вирусного белка, такого как VP64 или VPR, или с TAD из человеческого белка, такого как CITED2, CITED4 или CREB3. Согласно иллюстративным вариантам осуществления TALE конъюгирован с VP64 или VPR TAD.

eTF, которые активируют SCN1A и высокоомологичны человеческим белкам.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, характеризуется высоким процентом идентичности по отношению к одному или нескольким белкам человека (как дополнительно описано ниже). Согласно определенным вариантам осуществления такие eTF характеризуются общей идентичностью последовательности, составляющей по меньшей мере 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к одному или нескольким белкам человека. Согласно определенным вариантам осуществления такие eTF проявляют пониженную иммуногенность по сравнению с eTF, имеющим более низкий общий процент идентичности последовательностей по отношению к одному или нескольким белкам человека. Согласно различным вариантам осуществления снижение иммуногенности можно из-

мерить с использованием анализа elispot, иммуноанализа или способа *in silico*. Согласно определенным вариантам осуществления такие eTF могут содержать DBD, полученный из EGR1 или EGR3 человека, и TAD, полученный из EGR1, EGR3, CITED2, CITED4 или CREB3 человека. Такие eTF практически не обладают иммуногенностью при введении субъекту или обладают пониженной иммуногенностью по сравнению с eTF, характеризующимися более низким процентом идентичности по отношению к последовательностям белков человека.

Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SNC1A, характеризуется идентичностью последовательности, составляющей по меньшей мере 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к одному или нескольким белкам человека. Когда представленный в настоящем документе eTF, который активирует SNC1A, содержит DBD и TAD, происходящие из одного и того же белка, процент идентичности с человеческим белком может быть определен путем расчета общего количества аминокислотных остатков в eTF, которые соответствуют человеческому белку, из которого он был получен (например, EGR1 или EGR3), разделенного на общее количество аминокислотных остатков в eTF. Когда eTF, обеспечивающий активную регуляцию SNC1A, содержит DBD из одного человеческого белка и TAD, полученный из другого человеческого белка, процент идентичности с человеком может быть определен путем отдельного расчета процента идентичности с человеческим для каждого домена и суммирования этих двух показателей вместе, например, (i) вычисление общего количества аминокислотных остатков в DBD, которые соответствуют человеческому белку, из которого он был получен (например, EGR1 или EGR3), деленного на общее количество аминокислотных остатков в eTF; (ii) вычисление общего количества аминокислотных остатков в TAD, которые соответствуют человеческому белку, из которого он был получен (например, CITED2, CITED4 или CREB3), деленного на общее количество аминокислотных остатков в eTF; и (iii) суммирование суммы (i) и (ii). Согласно такому варианту осуществления домены делятся следующим образом: первый домен проходит от N-конца eTF до начала кодирующей последовательности для второго домена, а второй домен проходит от начала кодирующей последовательности для второго домена через C-конец eTF (например, для eTF, имеющего конфигурацию NLS-DBD-линкер-NLS-TAD, первый домен будет NLS-DBD-линкер, а второй домен будет NLS-TAD). Когда представленный в настоящем документе eTF, который активирует SNC1A, содержит DBD из одного человеческого белка и двух TAD, полученных из одного или нескольких различных человеческих белков, процент идентичности с человеком может быть определен путем отдельного расчета процента идентичности с человеческим для каждого домена и суммирования всех троих вместе, например, (i) вычисление общего количества аминокислотных остатков в DBD, которые соответствуют белку человека, из которого он был получен (например, EGR1 или EGR3), деленного на общее количество аминокислотных остатков в eTF; (ii) вычисление общего количества аминокислотных остатков в первом TAD, которые соответствуют человеческому белку, из которого он был получен (например, CITED2, CITED4 или CREB3), деленного на общее количество аминокислотных остатков в eTF; (iii) вычисление общего количества аминокислотных остатков во втором TAD, которые соответствуют человеческому белку, из которого он был получен (например, CITED2, CITED4 или CREB3), деленного на общее количество аминокислотных остатков в eTF; и (iv) суммирование суммы (i), (ii) и (iii). Согласно такому варианту осуществления домены делятся следующим образом: первый домен проходит от N-конца eTF до начала кодирующей последовательности для второго домена, второй домен проходит от начала кодирующей последовательности для второго домена через начало кодирующей последовательности для третьего домена, а третий домен проходит от начала кодирующей последовательности для третьего домена через C-конец eTF (например, для eTF, имеющего конфигурацию NLS-TAD1-линкер-NLS-DBD-линкер-NLS-TAD2, первый домен будет NLS-TAD1-линкер, второй домен будет NLS-DBD-линкер, а третий домен будет NLS-TAD2). Процент идентичности по отношению к одному или нескольким человеческим белкам, как описано в этом разделе, может быть определен с использованием выходного процента идентичности, полученного с помощью стандартного инструмента BLAST для белков, доступного в NCBI (например, инструмента выравнивания набора blastp, используя алгоритм blastp (белок -> белок) с параметрами по умолчанию) доступен в Интернете на веб-сайте NCBI по адресу blast.ncbi.nlm.nih.gov/.

Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SNC1A, имеет преимущество в том, что вызывает слабый, минимальный или нулевой неблагоприятный иммунный ответ у человека из-за высокой степени идентичности последовательностей с природными белками человека. Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SNC1A, вызывает пониженную иммуногенность, например, снижение иммуногенности по меньшей мере в 0,5, 1, 1,5, 2, 2,5, 3, 3,5, 4, 4,5, 5, 10, 15, 20, 25, 30, 35, 40, 45 или 50 или более раз по сравнению с иммуногенностью, наблюдаемой с eTF, имеющим более низкий процент идентичности по отношению к одному или нескольким белкам человека, например, eTF, характеризующимся идентичностью последовательности, составляющей менее чем 50%, 55%, 65% или 70%, по отношению к одному или нескольким белкам человека. В некоторых случаях снижение иммуногенности можно измерить с помощью анализа elispot, иммуноанализа или способа *in*

silico. Генная терапия, обладающая низкой или минимальной иммуногенностью, имеет несколько преимуществ, включая в себя улучшенную переносимость пациентом, уменьшенную дозу, необходимую для достижения терапевтического эффекта, пролонгированные терапевтические эффекты после одного введения, возможность вводить несколько раз или в нескольких дозах, если необходимо, устойчивую терапевтическую эффективность в течение более длительного периода времени на одно введение, повышенную безопасность и/или повышенную эффективность генной терапии.

Согласно определенным вариантам осуществления представленные в настоящем документе eTF, которые активируют SCN1A и характеризуются высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, включают в себя DBD и TAD, полученные из одного или нескольких встречающихся в природе белков человека. Согласно определенным вариантам осуществления такие eTF могут содержать DBD, полученный из любого встречающегося в природе белка человека, содержащего DBD. Согласно иллюстративным вариантам осуществления, представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, включает в себя DBD, полученный из встречающегося в природе белка типа "цинковые пальцы", такого как, например, любая из конструкций 5-27, 36-41 или 44-53, перечисленные в табл. 1. Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, включает в себя DBD, полученный из белка EGR человека, такого как EGR1, EGR2, EGR3 или EGR4. Согласно иллюстративным вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, включает в себя DBD, полученный из EGR1 или EGR3 человека. Согласно различным вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, включает в себя DBD, полученный из белка типа "цинковые пальцы" человека, в котором минимальные аминокислотные изменения (например, 1, 2, 3, 4, 5, 6, 7 или 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 2-3, 2-4, 2-5, 2-6, 2-7, 3-4, 3-5, 36 или 3-7 аминокислотных замен) были сделаны в одном или нескольких доменах типа "цинковые пальцы" DBD, чтобы изменить специфичность связывания DBD для распознавания представляющего интерес целевого сайта связывания. Такие модификации последовательности предпочтительно выполняются в спиралах распознавания доменов типа "цинковые пальцы" DBD, в то время как остальная часть DBD или белка типа "цинковые пальцы" человека (включая в себя TAD) остается немодифицированной, чтобы сохранить как можно больше идентичности последовательности с природным человеческим белком по возможности.

Согласно определенным вариантам осуществления представленные в настоящем документе eTF, которые активируют SCN1A и характеризуются высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, содержат один или несколько доменов модуляции транскрипции (например, TAD), полученных из человеческого белка, конъюгированного с DBD, полученным из белка человека. Согласно различным вариантам осуществления домен модуляции транскрипции может происходить из любого встречающегося в природе человеческого белка, имеющего домен, способный рекрутировать один или несколько белковых факторов, которые могут модулировать транскрипцию (например, РНК-полимераза II, коактиваторный белок или ко-репрессорный белок) или уровень экспрессии гена из представляющего интерес гена, когда eTF связывается с целевым сайтом через DBD. Согласно иллюстративным вариантам осуществления TAD получен из белка EGR человека, такого как, например, EGR1, EGR2, EGR3 или EGR4 человека, или из указанного белка человека, такого как, например, белок CITED2 или CITED4 человека. Согласно иллюстративному варианту осуществления представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, включает в себя TAD из белка EGR1 или EGR3 человека. Согласно другому иллюстративному варианту осуществления представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, включает в себя TAD из белка CITED2 или CITED4 человека.

Согласно одному варианту осуществления представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, может содержать DBD человека (hDBD) и TAD человека (hTAD) (например, hTAD-hDBD или hDBD-hTAD), причем hDBD и hTAD могут происходить из одного и того же человеческого белка или из разных человеческих белков. Согласно другому варианту осуществления представленный в настоящем документе eTF, характеризующийся высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, может содержать hDBD и два hTAD, причем hDBD и hTAD происходят из одного и того же человеческого белка, hDBD происходит из первого человеческого белка и оба hTAD происходят из второго человеческого белка, hDBD и один hTAD происходят из первого человеческого белка, а второй hTAD происходит из

второго человеческого белка, или hDBD происходит из первого человеческого белка, один hTAD происходит из второго человеческого белка, а второй hTAD происходит из третьего человеческого белка (например, hTAD1-hDBD-hTAD1, hTAD1-hDBD-hTAD2, hTAD1-hTAD1-hDBD, hTAD1-hTAD2-hDBD, hDBD-hTAD1-hTAD1 или hDBD-hTAD1-hTAD2).

Согласно иллюстративным вариантам осуществления представленный в настоящем документе eTF, характеризующийся высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, содержит любую из следующих конфигураций: (i) hDBD и hTAD, оба получены из EGR1 человека; (ii) hDBD и hTAD, оба получены из EGR3 человека; (iii) hDBD, полученный из EGR1 человека, и hTAD, полученный из CITED2 (например, hEGR1 DBD-hCITED2 TAD или hCITED2 TAD-hEGR1 DBD); (iv) hDBD, полученный из EGR1 человека, и hTAD, полученный из CITED4 человека (например, hEGR1 DBD-hCITED4 TAD или hCITED4 TAD-hEGR1 DBD); (v) hDBD, полученный из EGR3 человека, и hTAD, полученный из CITED2 (например, hEGR3 DBD-hCITED2 TAD или hCITED2 TAD-hEGR3 DBD); (vi) hDBD, полученный из EGR3 человека, и hTAD, полученный из CITED4 человека (например, hEGR3 DBD-hCITED4 TAD или hCITED4 TAD-hEGR3 DBD); (vii) hDBD, полученный из EGR1 человека, и два hTAD, полученные из CITED2 (например, hCITED2 TAD-hEGR1 DBD-hCITED2 TAD, hCITED2 TAD-hCITED2 TAD-hEGR1 DBD или hEGR1 DBD-hCITED2 TAD-hCITED2 TAD); (viii) hDBD, полученный из EGR1 человека, и два hTAD, полученные из CITED4 человека (например, hCITED4 TAD-hEGR1 DBD-hCITED4 TAD, hCITED4 TAD-hCITED4 TAD-hEGR1 DBD или hEGR1 DBD-hCITED4 TAD-hCITED4 TAD); (ix) hDBD, полученный из EGR3 человека, и два hTAD, полученные из CITED2 человека (например, hCITED2 TAD-hEGR3 DBD-hCITED2 TAD, hCITED2 TAD-hCITED2 TAD-hEGR3 DBD или hEGR3 DBD-hCITED2 TAD-hCITED2 TAD); (x) hDBD, полученный из EGR3 человека, и два hTAD, полученные из CITED4 человека (например, hCITED4 TAD-hEGR3 DBD-hCITED4 TAD, hCITED4 TAD-hCITED4 TAD-hEGR3 DBD или hEGR3 DBD-hCITED4 TAD-hCITED4 TAD); (xi) hDBD, полученный из EGR1 человека, первый hTAD, полученный из CITED2 человека, второй hTAD, полученный из CITED4 человека (например, hCITED2 TAD-hEGR1 DBD-hCITED4 TAD, hCITED4 TAD-hEGR1 DBD-hCITED2 TAD, hCITED2 TAD-hCITED4 TAD-hEGR1 DBD, hCITED4 TAD-hCITED2 TAD-hEGR1 DBD, hEGR1 DBD-hCITED4 TAD-hCITED2 TAD или hEGR1 DBD-hCITED2 TAD-hCITED4 TAD); или (xii) hDBD, полученный из EGR3 человека, первый hTAD, полученный из CITED2 человека, второй hTAD, полученный из CITED4 человека (например, hCITED2 TAD-hEGR3 DBD-hCITED4 TAD, hCITED4 TAD-hEGR3 DBD-hCITED2 TAD, hCITED2 TAD-hCITED4 TAD-hEGR3 DBD, hCITED4 TAD-hCITED2 TAD-hEGR3 DBD, hEGR3 DBD-hCITED4 TAD-hCITED2 TAD или hEGR3 DBD-hCITED2 TAD-hCITED4 TAD).

Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, содержит любое из следующего: (i) последовательность, содержащая любую из SEQ ID NO: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 или 223; (ii) последовательность, содержащая любую из SEQ ID NO: 92-98; (iii) последовательность, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к любой из последовательностей (i) или (ii); или (iv) функциональный фрагмент или вариант любой из последовательностей (i), (ii) или (iii). Согласно иллюстративным вариантам осуществления такие eTF способны повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или больше по сравнению с контролем, или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или больше по сравнению с контролем. Согласно иллюстративным вариантам осуществления такие eTF характеризуются общей идентичностью последовательности, составляющей по меньшей мере 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% по отношению к одному или нескольким белкам человека. Согласно определенным вариантам осуществления такие eTF проявляют пониженную иммуногенность по сравнению с eTF, характеризующимся более низким общим процентом идентичности последовательностей по отношению к одному или нескольким белкам человека. Согласно различным вариантам осуществления снижение иммуногенности можно измерить с использованием анализа elispot, иммуноанализа или способа *in silico*.

Согласно определенным вариантам осуществления представленный в настоящем документе eTF, который активирует SCN1A и характеризуется высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, может дополнительно содержать одну или несколько аминокислотных последовательностей или доменов в дополнение к доменам DBD и TAD, например, сигнал ядерной локализации или линкер и т.д. Кроме того, полинуклеотид, кодирующий представленный в настоящем документе eTF, характеризующийся высоким процентом идентичности последовательности по отношению к одному или нескольким белкам человека, может дополнительно содержать одну или несколько последовательностей нуклеиновых кислот в дополнение к кодирующей последовательности для eTF, например, промотор, энхансер, полиА-хвост и т.д. Согласно таким вариантам осуществления одна или несколько дополнительных аминокислотных последовательностей и/или после-

довательностей нуклеиновых кислот предпочтительно представляют собой человеческие последовательности, происходящие из человеческих последовательностей, или характеризуются идентичностью последовательности, составляющей по меньшей мере 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% по отношению к белку человека.

Иллюстративные eTF SCN1A.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит ДНК-связывающий домен, содержащий один или несколько доменов типа "цинковые пальцы", содержащих спираль распознавания, содержащую любую из SEQ ID NO: 152-167. Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит ДНК-связывающий домен, содержащий по меньшей мере один, два, три, четыре, пять, шесть, семь, восемь, девять, десять, одиннадцать или двенадцать доменов типа "цинковые пальцы", причем каждый домен типа "цинковые пальцы" независимо содержит спираль распознавания, содержащую любую из SEQ ID NO: 152-167. Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит ДНК-связывающий домен, содержащий шесть доменов типа "цинковые пальцы", причем каждый домен типа "цинковые пальцы" независимо содержит спираль распознавания, содержащую любую из SEQ ID NO: 152-167. Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит ДНК-связывающий домен, содержащий девять доменов типа "цинковые пальцы", причем каждый домен типа "цинковые пальцы" независимо содержит спираль распознавания, содержащую любую из SEQ ID NO: 152-167. Согласно иллюстративным вариантам осуществления такие eTF содержат ДНК-связывающий домен, содержащий SEQ ID NO: 147, где каждый X независимо выбран из любой из SEQ ID NO: 152-167, а n равно 6 или 9.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит ДНК-связывающий домен, содержащий любое из следующего: (i) последовательность, содержащая RSDNLVR x REDNLHT x RSDLVLR x QSGNLTE x TSGHLVR x QNSTLTE (SEQ ID NO: 148), где x может представлять собой линкер из 1-50 аминокислот, (ii) последовательность, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99%, по отношению к SEQ ID NO: 148, или (ii) функциональный фрагмент из (i) или (ii). Согласно определенным вариантам осуществления такой eTF дополнительно содержит один или несколько TAD, выбранных из VP64, VPR, CITED2, CITED4 или CREB3. Согласно одному варианту осуществления такой eTF содержит домен TAD VPR, конъюгированный с С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит CITED2 TAD, конъюгированный с N-концом, С-концом или N-концом и С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит TAD CITED4, конъюгированный с N-концом, С-концом или N-концом и С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит два TAD CITED4, конъюгированных с N-концом или С-концом DBD. Согласно определенным вариантам осуществления такой eTF способен связываться с целевым сайтом, содержащим SEQ ID NO: 18, и повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или более раз по сравнению с контролем, или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более по сравнению с контролем.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит ДНК-связывающий домен, содержащий любое из следующего: (i) последовательность, содержащая RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT (SEQ ID NO: 149), где x может представлять собой линкер из 1-50 аминокислот, (ii) последовательность, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99%, по отношению к SEQ ID NO: 149, или (ii) функциональный фрагмент из (i) или (ii). Согласно определенным вариантам осуществления такой eTF дополнительно содержит один или несколько TAD, выбранных из VP64, VPR, CITED2, CITED4 или CREB3. Согласно одному варианту осуществления такой eTF содержит домен TAD VPR, конъюгированный с С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит CITED2 TAD, конъюгированный с N-концом, С-концом или N-концом и С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит TAD CITED4, конъюгированный с N-концом, С-концом или N-концом и С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит два TAD CITED4, конъюгированных с N-концом или С-концом DBD. Согласно определенным вариантам осуществления такой eTF способен связываться с целевым сайтом, содержащим SEQ ID NO: 30, и повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или более раз по сравнению с контролем, или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более по сравнению с контролем.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит ДНК-связывающий домен, содержащий любое из следующего: (i) последовательность, содержащая RRDELNV x RSDHLTN x RSDDLVR x RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT (SEQ ID NO: 151), (ii) последовательность, характеризую-

шаяся идентичностью последовательности, составляющей по меньшей мере 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99%, по отношению к SEQ ID NO: 151, или (ii) функциональный фрагмент (i) или (ii). Согласно определенным вариантам осуществления такой eTF дополнительно содержит один или несколько TAD, выбранных из VP64, VPR, CITED2, CITED4 или CREB3. Согласно одному варианту осуществления такой eTF содержит домен TAD VPR, конъюгированный с С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит CITED2 TAD, конъюгированный с N-концом, С-концом или N-концом и С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит TAD CITED4, конъюгированный с N-концом, С-концом или N-концом и С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит два TAD CITED4, конъюгированных с N-концом или С-концом DBD. Согласно определенным вариантам осуществления такой eTF способен связываться с целевым сайтом, содержащим SEQ ID NO: 32, и повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или более по сравнению с контролем, или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300% или 500% или более по сравнению с контролем.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит ДНК-связывающий домен, содержащий любое из следующего: (i) последовательность, содержащая DPGALVR x RSDNLVR x QSGDLRR x THLDLIR x TSGNLVR x RSDNLVR (SEQ ID NO: 150), (ii) последовательность, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 89%, 85%, 90%, 95%, 96%, 97%, 98% или 99%, по отношению к SEQ ID NO: 150, или (ii) функциональный фрагмент (i) или (ii). Согласно определенным вариантам осуществления такой eTF дополнительно содержит один или несколько TAD, выбранных из VP64, VPR, CITED2, CITED4 или CREB3. Согласно одному варианту осуществления такой eTF содержит домен TAD VPR, конъюгированный с С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит CITED2 TAD, конъюгированный с N-концом, С-концом или N-концом и С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит TAD CITED4, конъюгированный с N-концом, С-концом или N-концом и С-концом DBD. Согласно определенным вариантам осуществления такой eTF содержит два TAD CITED4, конъюгированных с N-концом или С-концом DBD. Согласно определенным вариантам осуществления такой eTF способен связываться с целевым сайтом, содержащим SEQ ID NO: 31, и повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или более по сравнению с контролем, или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более по сравнению с контролем.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит любое из следующего: (i) последовательность, содержащая любую из SEQ ID NO: 99-131, 205, 207, 209, 213, 217, 219, 221 или 223; (ii) последовательность, содержащая любую из SEQ ID NO: 77-98; (iii) последовательность, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к любой из последовательностей (i) или (ii); или (iv) функциональный фрагмент или вариант любой из последовательностей (i), (ii) или (iii). Согласно иллюстративным вариантам осуществления такие eTF способны повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или более по сравнению с контролем или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более по сравнению с контролем.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит любое из следующего: (i) последовательность, содержащая любую из SEQ ID NO: 99-102 или 125-127; (ii) последовательность, содержащая любую из SEQ ID NO: 77-91; (iii) последовательность, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к любой из последовательностей (i) или (ii); или (iv) функциональный фрагмент или вариант любой из последовательностей (i), (ii) или (iii). Согласно иллюстративным вариантам осуществления такие eTF способны повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или более по сравнению с контролем, или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более по сравнению с контролем.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит любое из следующего: (i) последовательность, содержащая любую из SEQ ID NO: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 или 223; (ii) последовательность, содержащая любую из SEQ ID NO: 92-98; (iii) последовательность, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к любой из последовательностей (i) или (ii); или (iv) функциональный фрагмент или вариант любой из последовательностей (i), (ii) или (iii). Согласно иллюстративным вариантам осуществления такие eTF способны повышать экспрессию SCN1A по меньшей

последовательностей (i) или (ii). Согласно иллюстративным вариантам осуществления такие eTF содержат SEQ ID NO: 92 и связываются с целевым сайтом, содержащим SEQ ID NO: 18. Согласно иллюстративным вариантам осуществления такие eTF способны повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или более по сравнению с контролем, или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более по сравнению с контролем. Согласно иллюстративным вариантам осуществления такие eTF характеризуются общей идентичностью последовательности, составляющей по меньшей мере 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к одному или нескольким белкам человека. Согласно определенным вариантам осуществления такие eTF проявляют пониженную иммуногенность по сравнению с eTF, характеризующимся более низким общим процентом идентичности последовательностей по отношению к одному или нескольким белкам человека. Согласно различным вариантам осуществления снижение иммуногенности можно измерить с использованием анализа elispot, иммуноанализа или способа *in silico*.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит любое из следующего: (i) последовательность, содержащая SEQ ID NO: 131; (ii) последовательность, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к SEQ ID NO: 131; или (iii) функциональный фрагмент или вариант любой из последовательностей (i) или (ii). Согласно иллюстративным вариантам осуществления такие eTF содержат SEQ ID NO: 92 и связываются с целевым сайтом, содержащим SEQ ID NO: 18. Согласно иллюстративным вариантам осуществления такие eTF способны повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или более по сравнению с контролем, или по меньшей мере на 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 100%, 125%, 150%, 200%, 250%, 300%, 400% или 500% или более по сравнению с контролем. Согласно иллюстративным вариантам осуществления такие eTF характеризуются общей идентичностью последовательности, составляющей по меньшей мере 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к одному или нескольким белкам человека. Согласно определенным вариантам осуществления такие eTF проявляют пониженную иммуногенность по сравнению с eTF, характеризующимся более низким общим процентом идентичности последовательностей по отношению к одному или нескольким белкам человека. Согласно различным вариантам осуществления снижение иммуногенности можно измерить с использованием анализа elispot, иммуноанализа или способа *in silico*.

Согласно определенным вариантам осуществления раскрытый в настоящем документе eTF, который активирует SCN1A, содержит DBD, содержащий комплекс gRNA/Cas, причем gRNA содержит нацеливающую последовательность, содержащую любую из SEQ ID NO: 35-66. Целевая последовательность gRNA прикреплена к 5'-концу последовательности каркаса, содержащей последовательность: 5'-GTTTTAGTACTCTGGAAACAGAAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA-3' (SEQ ID NO: 183). Согласно иллюстративным вариантам осуществления белок Cas представляет собой дезактивированный нуклеазой белок Cas9. Согласно определенным вариантам осуществления такой eTF дополнительно содержит один или несколько TAD, конъюгированных с белком Cas, причем TAD выбран из VP64, VPR, CITED2, CITED4 или CREB3. Согласно одному варианту осуществления такой eTF содержит домен VPR TAD, конъюгированный с C-концом белка Cas. Согласно определенным вариантам осуществления такой eTF содержит CITED2 TAD, конъюгированный с N-концом, C-концом или N-концом и C-концом белка Cas. Согласно определенным вариантам осуществления такой eTF содержит TAD CITED4, конъюгированный с N-концом, C-концом или N-концом и C-концом белка Cas. Согласно иллюстративным вариантам осуществления такие eTF способны повышать экспрессию SCN1A по меньшей мере в 2 раза, 5 раз, 10 раз, 15 раз, 20 раз, 25 раз, 30 раз, 40 раз, 50 раз или больше по сравнению с контролем.

Полинуклеотиды.

Согласно другому аспекту в настоящей заявке представлены полинуклеотиды, кодирующие любой из eTF, которые активируют SCN1A, раскрытые в настоящем документе. Согласно другому аспекту в настоящей заявке представлены полинуклеотиды, содержащие селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления в настоящей заявке представлены полинуклеотиды, содержащие селективный в отношении PV регуляторный элемент, функционально связанный с трансгеном, и селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления в настоящей заявке представлены полинуклеотиды, содержащие последовательность, кодирующую eTF, которая активирует SCN1A, как раскрыто в настоящем документе, и селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления в настоящей заявке представлен селективный в отношении PV регуляторный элемент, функционально связанный с трансгеном, кодирующим eTF, который активирует SCN1A, и селективный в отношении PV регуляторный элемент.

Полинуклеотиды, кодирующие eTF, которые активируют SCN1A.

Согласно определенным вариантам осуществления в настоящей заявке представлен полинуклеотид,

содержащий любое из следующего: (i) последовательность нуклеиновой кислоты, кодирующая eTF, который активирует SCN1A, содержащий любую из SEQ ID NO: 77-131, 205, 207, 209, 213, 217, 219, 221 или 223, или их вариант или функциональный фрагмент; (ii) нуклеиновая кислота, кодирующая функциональный фрагмент eTF, который активирует SCN1A, содержащий любую из SEQ ID NO: 77-131, 205, 207, 209, 213, 217, 219, 221 или 223; или (iii) нуклеиновая кислота, кодирующая eTF, который активирует SCN1A, характеризующийся идентичностью последовательности, составляющей по меньшей мере 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более, по отношению к eTF, который активирует SCN1A, содержащий любую из SEQ ID NO: 77-131, 205, 207, 209, 213, 217, 219, 221 или 223, или их вариант или функциональный фрагмент.

Согласно определенным вариантам осуществления в настоящей заявке представлен полинуклеотид, содержащий любое из следующего: (i) последовательность нуклеиновой кислоты, кодирующая DBD, содержащий любую из SEQ ID NO: 92-98, или ее вариант или функциональный фрагмент; (ii) нуклеиновая кислота, кодирующая функциональный фрагмент DBD, содержащий любую из SEQ ID NO: 92-98; или (iii) нуклеиновая кислота, кодирующая DBD, характеризующийся идентичностью последовательности, составляющей по меньшей мере 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более, по отношению к DBD, содержащему любую из SEQ ID NO: 92-98, или ее вариант или функциональный фрагмент, причем DBD способен связываться с целевым сайтом, связанным любой из SEQ ID NO: 92-98.

Согласно определенным вариантам осуществления в настоящей заявке представлен полинуклеотид, кодирующий eTF, который активирует эндогенный SCN1A, причем полинуклеотид содержит любое из следующего: (i) последовательность нуклеиновой кислоты, кодирующая eTF, содержащий любую из SEQ ID NO: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 или 223; (ii) нуклеиновая кислота, кодирующая функциональный фрагмент eTF, содержащий любую из SEQ ID NO: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 или 223; или (iii) нуклеиновая кислота, кодирующая eTF, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более, по отношению к eTF, содержащему любую из SEQ ID NO: 103-124, 128-131, 205, 207, 209, 213, 217, 219, 221 или 223, причем eTF способен активировать SCN1A.

Согласно определенным вариантам осуществления в настоящей заявке представлен полинуклеотид, кодирующий DBD, который связывается с целевым сайтом генома, способным повышать регуляцию эндогенного SCN1A при связывании с помощью раскрытого в настоящем документе eTF, причем полинуклеотид содержит любой из следующих элементов: (i) последовательность нуклеиновой кислоты, кодирующая DBD, содержащий любую из SEQ ID NO: 77-98; (ii) нуклеиновая кислота, кодирующая функциональный фрагмент DBD, содержащий любую из SEQ ID NO: 77-98; или (iii) нуклеиновая кислота, кодирующая eTF, характеризующийся идентичностью последовательности, составляющей по меньшей мере 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более, по отношению к DBD, содержащему любую из SEQ ID NO: 77-98, причем DBD способен связываться с целевым сайтом, связанным любой из SEQ ID NO: 77-98.

Согласно определенным вариантам осуществления в настоящей заявке представлен полинуклеотид, кодирующий DBD, который связывается с целевым сайтом генома, способным активировать эндогенный SCN1A при связывании раскрытым в настоящем документе eTF, причем полинуклеотид содержит любой из следующих элементов: (i) последовательность нуклеиновой кислоты, кодирующая DBD, содержащий любую из SEQ ID NO: 148-151; (ii) нуклеиновая кислота, кодирующая функциональный фрагмент DBD, содержащий любую из SEQ ID NO: 148-151; или (iii) нуклеиновая кислота, кодирующая eTF, характеризующийся идентичностью последовательности, составляющей по меньшей мере 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более, по отношению к DBD, содержащему любую из SEQ ID NO: 148-151, причем DBD способен связываться с целевым сайтом, связанным любой из SEQ ID NO: 92-98.

Согласно определенным вариантам осуществления в настоящей заявке представлен полинуклеотид, кодирующий eTF, способный регулировать эндогенный SCN1A, причем полинуклеотид содержит любое из следующего: (i) последовательность нуклеиновой кислоты, содержащая любую из SEQ ID NO: 70-76 или 184; (ii) нуклеиновая кислота, содержащая функциональный фрагмент любой из последовательностей (i); или (iii) нуклеиновая кислота, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более, по отношению к любой из последовательностей (i) или (ii), причем полинуклеотид кодирует eTF, который способен активировать SCN1A.

Согласно определенным вариантам осуществления в настоящей заявке представлен полинуклеотид, кодирующий eTF, способный регулировать эндогенный SCN1A, причем полинуклеотид содержит любое из следующего: (i) последовательность нуклеиновой кислоты, содержащая SEQ ID NO: 70; (ii) нуклеиновая кислота, содержащая функциональный фрагмент SEQ ID NO: 70; или (iii) нуклеиновая кислота, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более, по отноше-

шей мере 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более, по отношению к любой из последовательностей (i) или (ii). Согласно иллюстративным вариантам осуществления такие полинуклеотиды кодируют eTF, содержащий SEQ ID NO: 106, или его функциональный фрагмент или вариант, который способен активировать SCN1A.

Полинуклеотиды, содержащие сайты связывания микроРНК для селективной экспрессии в PV-содержащих нейронах.

Согласно другому аспекту в настоящей заявке представлены полинуклеотиды, содержащие сайты связывания микроРНК, которые приводят к селективной экспрессии представляющего интерес гена в парвальбумин-содержащих нейронах (PV). МикроРНК или миРНК представляют собой небольшие некодирующие РНК (~20 нуклеотидов), которые регулируют экспрессию генов посттранскрипционно путем гибридизации с комплементарными сайтами узнавания в молекуле мРНК и приводят к ингибированию экспрессии генов, способствуя деградации транскрипта мРНК или подавляя трансляцию белка, кодируемого мРНК. Предусмотренные в настоящем документе сайты связывания микроРНК ингибируют экспрессию представляющего интерес гена в возбуждающих нейронах, тем самым способствуя селективной экспрессии представляющего интерес гена в PV-содержащих нейронах (например, селективные к PV сайты связывания микроРНК). Согласно определенным вариантам осуществления возбуждающие нейроны представляют собой нейроны, которые экспрессируют один или несколько из STAC, Slc17a7, Car12, Syt17, ITPKA, Col6a1, CamKII, Sv2b, INHBA и/или DKK3. Согласно иллюстративному варианту осуществления возбуждающие нейроны представляют собой нейроны, экспрессирующие CamKII.

Согласно определенным вариантам осуществления в настоящей заявке представлены полинуклеотиды, содержащие один или несколько сайтов связывания микроРНК для одной или нескольких микроРНК, которые способствуют селективной экспрессии PV, например, способствуют деградации мРНК, содержащей сайт связывания микроРНК, в возбуждающих нейронах. Примеры микроРНК, которые способствуют селективной экспрессии PV, включают в себя, например, miR-128, miR-221 и miR-222. Согласно определенным вариантам осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более селективных в отношении PV сайтов связывания микроРНК. Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более сайтов связывания miR-128 (SEQ ID NO: 9). Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более сайтов связывания miR-221 (SEQ ID NO: 11). Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более сайтов связывания miR-222 (SEQ ID NO: 13). Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 1 сайт связывания miR-128, по меньшей мере один сайт связывания miR-221 и по меньшей мере один сайт связывания miR-222. Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере один сайт связывания miR-128 и по меньшей мере один сайт связывания miR-222. Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере один сайт связывания miR-221 и по меньшей мере один сайт связывания miR-222. Согласно иллюстративному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере один сайт связывания miR-128 (SEQ ID NO: 9) и по меньшей мере один сайт связывания miR-221 (SEQ ID NO: 11). Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 2 сайта связывания miR-128 (SEQ ID NO: 9) и по меньшей мере 2 сайта связывания miR-221 (SEQ ID NO: 11). Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 3 сайта связывания miR-128 (SEQ ID NO: 9) и по меньшей мере 3 сайта связывания miR-221 (SEQ ID NO: 11). Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 4 сайта связывания miR-128 (SEQ ID NO: 9) и по меньшей мере 4 сайта связывания miR-221 (SEQ ID NO: 11). Согласно одному варианту осуществления в настоящей заявке представлены полинуклеотиды, содержащие по меньшей мере 5 сайтов связывания miR-128 (SEQ ID NO: 9) и по меньшей мере 5 сайтов связывания miR-221 (SEQ ID NO: 11). Согласно таким вариантам осуществления сайты связывания могут быть расположены в любом порядке. Например, для конструкции, содержащей 2 сайта связывания miR-128 и 2 сайта связывания miR-221, сайты связывания могут быть расположены в любой из следующих конфигураций: miR-128 - miR-128 - miR-221 - miR221, miR-128 - miR-221 - miR-128 - miR-221, miR-128 - miR221 - miR221 - miR-128, miR-221 - miR128 - miR221 - miR128, miR-221 - miR128 - miR128 - miR221 или miR221 - miR221 - miR128 - miR128. Согласно иллюстративному варианту осуществления представленные в настоящем документе полинуклеотиды содержат последовательность, содержащую 4 сайта связывания miR-128 (SEQ ID NO: 9), за которыми следуют четыре сайта связывания miR-221 (SEQ ID NO: 11), например, miR-128 - miR-128 - miR128 - miR-128 - miR221 - miR221 - miR221 - miR221. Согласно другому иллюстративному варианту осуществления представленные в настоящем документе полинуклеотиды содержат последовательность, содержащую 1 последовательность miR-221 (SEQ ID NO: 11), 1 последовательность miR-222 (SEQ ID NO: 13) и 1 сайт связывания miR-128 (SEQ ID NO: 9), например, miR-

221 - miR222 - miR128. Согласно другому иллюстративному варианту осуществления представленные в настоящем документе полинуклеотиды содержат последовательность, содержащую 2 последовательности miR-221 (SEQ ID NO: 11), 2 последовательности miR-222 (SEQ ID NO: 13) и 2 сайта связывания miR-128 (SEQ ID NO: 9) расположены в следующем порядке: miR-221 - miR222 - miR128 - miR-221 - miR222 - miR128.

В полинуклеотидах, содержащих более одного сайта связывания микроРНК, сайты связывания могут быть непосредственно примыкающими друг к другу в полинуклеотидной последовательности (например, без линкера или промежуточной последовательности между сайтами связывания) или могут быть отделены друг от друга по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 или более нуклеотидами или 1-20, 1-15, 1-10, 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3 или 1-2 нуклеотидами. Согласно иллюстративным вариантам осуществления сайты связывания микроРНК разделены приблизительно 5 нуклеотидами или 5 нуклеотидами. Согласно иллюстративным вариантам осуществления последовательности, разделяющие сайты связывания микроРНК (а также соединения, образованные между сайтами связывания микроРНК, или соединения, образованные между сайтами связывания микроРНК и последовательностями, разделяющими сайты связывания микроРНК), не комплементарны никаким другим микроРНК или любым другим нейрональным микроРНК.

Согласно определенным вариантам осуществления представленные в настоящем документе полинуклеотиды содержат сайт связывания микроРНК, характеризующийся идентичностью, составляющей по меньшей мере 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100%, по отношению к SEQ ID NO: 7. Согласно иллюстративному варианту осуществления представленные в настоящем документе полинуклеотиды содержат сайт связывания микроРНК, содержащий SEQ ID NO: 7.

Согласно определенным вариантам осуществления представленные в настоящем документе полинуклеотиды содержат сайт связывания микроРНК, характеризующийся идентичностью, составляющей по меньшей мере 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100%, по отношению к SEQ ID NO: 14. Согласно иллюстративному варианту осуществления представленные в настоящем документе полинуклеотиды содержат сайт связывания микроРНК, содержащий SEQ ID NO: 14.

Согласно определенным вариантам осуществления представленные в настоящем документе полинуклеотиды содержат сайт связывания микроРНК, характеризующийся идентичностью, составляющей по меньшей мере 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100%, по отношению к SEQ ID NO: 15. Согласно иллюстративному варианту осуществления представленные в настоящем документе полинуклеотиды содержат сайт связывания микроРНК, содержащий SEQ ID NO: 15.

Согласно определенным вариантам осуществления указанные в настоящем документе сайты связывания микроРНК расположены в 3'-нетранслируемой области транскрипта мРНК, например, после кодона терминации трансляции (т.е. TAA, TGA или TAG) и перед полиА-хвостом. Сайт связывания микроРНК может быть расположен непосредственно рядом с кодоном терминации трансляции или может быть отделен от кодона терминации трансляции по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 или более нуклеотидами или 1-20, 1-15, 1-10, 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3 или 1-2 нуклеотидами, и/или может быть расположен рядом с полиА-хвостом или может быть отделен от полиА-хвоста по меньшей мере 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 или более нуклеотидами или 1-20, 1-15, 1-10, 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3 или 1-2 нуклеотидами.

Согласно определенным вариантам осуществления представленный в настоящем документе сайт связывания микроРНК приводит к селективной экспрессии гена в PV-содержащей клетке по сравнению с нецелевыми типами клеток. В некоторых случаях нецелевые типы клеток включают в себя, без ограничения, возбуждающие нейроны, не содержащие PV типы клеток ЦНС и ненейрональные типы клеток ЦНС. Согласно определенным вариантам осуществления селективные в отношении PV сайты связывания микроРНК приводят к селективной экспрессии генов в PV-содержащих нейронах по меньшей мере для одного, двух, трех, четырех, пяти или более не-PV типов клеток ЦНС. В некоторых случаях не-PV клетка ЦНС представляет собой возбуждающий нейрон, дофаминергический нейрон, астроцит, микроглию, мотонейрон, сосудистую клетку или не-ГАМКергический нейрон (например, клетку, которая не экспрессирует один или более GAD2, GAD1, NKX2.1, DLX1, DLX5, SST и VIP), не-PV нейрон (например, ГАМКергический нейрон, который не экспрессирует парвальбумин) или другие клетки ЦНС (например, типы клеток ЦНС, которые никогда не экспрессируют любой из PV, GAD2, GAD1, NKX2.1, DLX1, DLX5, SST и VIP). Согласно иллюстративному варианту осуществления представленный в настоящем документе селективный в отношении PV сайт связывания микроРНК приводит к повышенной селективности экспрессии генов в PV-содержащих нейронах по сравнению с возбуждающими нейронами (например, нейронами, которые экспрессируют один или несколько из STAC, Slc17a7, Car12, Syt17, IT-РКА, Col6a1, CamKII, Sv2b, INHBA и/или DKK3) за счет уменьшения экспрессии в возбуждающих нейронах. В некоторых случаях типы клеток различают по разным клеточным маркерам, морфологии, фенотипу, генотипу, функции и/или любым другим средствам классификации типов клеток.

Селективность экспрессии, управляемая селективным в отношении PV сайтом связывания микроРНК, может быть измерена несколькими способами. Согласно одному варианту осуществления селективность экспрессии гена в PV-клетке по сравнению с не содержащими PV клетками может быть изме-

рена путем сравнения количества PV-клеток, которые экспрессируют обнаруживаемый уровень транскрипта из гена, который содержит селективный в отношении PV сайт связывания микроРНК, с общим количеством клеток, которые экспрессируют ген (например, отношение PV-содержащих к общему количеству клеток (PV + не-PV клеток), экспрессирующих ген). Например, селективность в отношении PV-содержащих нейронов может быть определена с использованием анализа колокализации на основе иммуногистохимии с использованием кассеты экспрессии, содержащей ген, кодирующий флуоресцентный белок (например, eGFP), и селективный в отношении PV сайт связывания микроРНК для измерения экспрессии гена и антитела, которое идентифицирует PV-содержащие клетки (например, антитело к PV, которое специфически взаимодействует с PV-содержащими нейронами), связанное со второй флуоресцентной меткой (например, красным флуоресцентным белком). Селективность экспрессии в PV-содержащих клетках можно рассчитать путем деления количества клеток, которые экспрессируют как PV, так и eGFP (например, PV-содержащие клетки), на общее количество клеток, экспрессирующих eGFP (например, PV-содержащие клетки и не содержащие PV клетки), и умножения на 100, чтобы преобразовать в проценты. В другом примере селективность в отношении PV-содержащих нейронов может быть определена с использованием анализа колокализации на основе иммуногистохимии с использованием кассеты экспрессии, содержащей ген, кодирующий флуоресцентный белок (например, eGFP), и селективный в отношении PV сайт связывания микроРНК для измерения экспрессии гена, и первое антитело, которое идентифицирует PV-содержащие клетки (например, антитело к PV, которое специфически взаимодействует с PV-содержащими нейронами), связанное со второй флуоресцентной меткой (например, красным флуоресцентным белком) и вторым антителом, которое идентифицирует возбуждающие клетки (например, антитело к CamKII, которое специфически взаимодействует с возбуждающими нейронами). Селективность экспрессии в PV-содержащих клетках можно рассчитать путем деления количества клеток, экспрессирующих как PV, так и eGFP (например, PV-клетки), на количество клеток, экспрессирующих eGFP + PV и eGFP + CamKII (например, PV-содержащие клетки и возбуждающие клетки), и умножения на 100 для преобразования в процент. Чем выше процент PV-содержащих клеток, экспрессирующих трансген, тем более селективным является сайт связывания микроРНК в отношении PV-содержащих клеток. Согласно определенным вариантам осуществления представленный в настоящем документе селективный в отношении PV сайт связывания микроРНК может быть высокоселективным для экспрессии в PV-содержащих клетках. Например, представленный в настоящем документе селективный в отношении PV сайт связывания микроРНК может характеризоваться селективностью, составляющей приблизительно 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более чем приблизительно 99%, для PV-содержащих нейронов (например, PV-содержащие нейроны/все клетки \times 100 или PV-содержащие нейроны/PV + возбуждающие нейроны \times 100).

В некоторых случаях представленный в настоящем документе селективный в отношении PV сайт связывания микроРНК является коротким. В некоторых случаях размер селективного в отношении PV сайта связывания микроРНК совместим с клонирующей способностью вектора, например, вирусного вектора или gAAV, так что комбинированный размер трансгена, промотора (и необязательного энхансера) и сайта связывания микроРНК не превышает клонирующую способность вектора. В некоторых случаях селективный в отношении PV сайт связывания микроРНК составляет в длину до приблизительно 500 п.н., 400 п.н., 300 п.н., 250 п.н., 225 п.н., 215 п.н., 210 п.н., 200 п.н., 150 п.н., 140 п.н., 135 п.н., 130 п.н., 125 п.н., 120 п.н., 115 п.н., 110 п.н., 100 п.н., 90 п.н., 80 п.н., 75 п.н., 70 п.н., 65 п.н., 60 п.н. или 50 п.н.. В некоторых случаях селективный в отношении PV сайт связывания микроРНК составляет приблизительно 50-500 п.н., 50-400 п.н., 50-300 п.н., 50-250 п.н., 50-200 п.н., 50-100 п.н., 50-75 п.н., 50-70 п.н., 100-500 п.н., 100-400 п.н., 100-300 п.н., 100-250 п.н., 100-200 п.н., 100-150 п.н., 100-140 п.н., 100-135 п.н., 200-500 п.н., 200-400 п.н., 200-300 п.н. или 200-250 п.н..

Согласно иллюстративным вариантам осуществления представленный в настоящем документе полинуклеотид, который содержит один или несколько селективных в отношении PV сайтов связывания микроРНК не содержит SEQ ID NO: 67.

Кассеты экспрессии.

Согласно другому аспекту в настоящей заявке представлены кассеты экспрессии, содержащие представленный в настоящем документе полинуклеотид (например, полинуклеотид, содержащий последовательность, кодирующую eTF, которая активирует SCN1A и/или содержит селективный в отношении PV сайт связывания микроРНК) и один или несколько регуляторных элементов. Согласно определенным вариантам осуществления в настоящей заявке представлены кассеты экспрессии, содержащие представленный в настоящем документе полинуклеотид (например, полинуклеотид, содержащий последовательность, кодирующую eTF, которая активирует SCN1A и/или содержит селективный в отношении PV сайт связывания микроРНК) и селективный в отношении PV промотор.

Согласно определенным вариантам осуществления представленный в настоящем документе полинуклеотид (например, полинуклеотид, содержащий последовательность, кодирующую eTF, который активирует SCN1A и/или содержит селективный в отношении PV сайт связывания микроРНК) представляет собой часть кассеты экспрессии, содержащей один или несколько регуляторных элементов в добавок к

последовательности, кодирующей eTF. Согласно иллюстративным вариантам осуществления представленный в настоящем документе полинуклеотид (например, полинуклеотид, содержащий последовательность, кодирующую eTF, который активирует SCN1A и/или содержит селективный в отношении PV сайт связывания микроРНК), представляет собой часть кассеты экспрессии, содержащей промотор, расположенный выше против хода транскрипции от трансгенной последовательности, чтобы быть способным управлять экспрессией трансгена (например, eTF, который избирательно активирует SCN1A) в клетке.

Согласно определенным вариантам осуществления раскрытая в настоящем документе кассета экспрессии содержит представленный в настоящем документе полинуклеотид (например, полинуклеотид, содержащий последовательность, кодирующую eTF, которая активирует регуляцию SCN1A и/или содержит селективный в отношении PV сайт связывания микроРНК) и конститутивный промотор, расположенный выше против хода транскрипции от последовательности, кодирующей трансген, чтобы быть способным управлять экспрессией трансгена (например, eTF, который избирательно активирует SCN1A) в клетке. Примеры конститутивных промоторов включают в себя промотор GAD2, промотор синапсина человека, промотор CBA, промотор CMV, промотор minCMV, TATA-бокс, супер основной промотор или промотор EF1 α , или их комбинацию. Согласно определенным вариантам осуществления раскрытая в настоящем документе кассета экспрессии содержит представленный в настоящем документе полинуклеотид (например, полинуклеотид, содержащий последовательность, кодирующую eTF, которая активирует SCN1A и/или содержит селективный в отношении PV сайт связывания микроРНК) и короткий промотор, способный управлять экспрессией, трансгена (например, eTF, который избирательно активирует SCN1A) в клетке. Согласно определенным вариантам осуществления короткий промотор, подходящий для применения в соответствии с описанными в настоящем документе молекулами нуклеиновой кислоты, содержит менее чем 500 п.н., 450 п.н., 400 п.н., 350 п.н., 300 п.н., 250 п.н., 225 п.н., 200 п.н., 175 п.н., 150 п.н., 145 п.н., 140 п.н., 135 п.н., 130 п.н., 125 п.н., 120 п.н., 115 п.н., 110 п.н., 105 п.н., 100 п.н., 95 п.н., 90 п.н., 85 п.н., 80 п.н. или 75 п.н. или приблизительно 80-300 п.н., 80-275 п.н., 80-250 п.н., 80-200 п.н., 80-150 п.н., 80-125 п.н., 80-120 п.н., 80-115 п.н., 80-110 п.н., 80-105 п.н., 80-100 п.н., 85-300 п.н., 85-275 п.н., 85-250 п.н., 85-200 п.н., 85-150 п.н., 85-125 п.н., 85-120 п.н., 85-115 п.н., 85-110 п.н., 85-105 п.н., 85-100 п.н., 90-300 п.н., 90-275 п.н., 90-250 п.н., 90-200 п.н., 90-150 п.н., 90-125 п.н., 90-120 п.н., 90-115 п.н., 90-110 п.н., 90-105 п.н., 90-100 п.н., 95-300 п.н., 95-275 п.н., 95-250 п.н., 95-200 п.н., 95-150 п.н., 95-125 п.н., 95-120 п.н., 95-115 п.н., 95-110 п.н., 95-105 п.н., 95-100 п.н., 100-300 п.н., 100-275 п.н., 100-250 п.н., 100-200 п.н., 100-150 п.н., 100-125 п.н., 100-120 п.н., 100-115 п.н., 100-110 п.н. или 100-105 п.н. Согласно иллюстративным вариантам осуществления короткий промотор, подходящий для применения в соответствии с описанными в настоящем документе кассетами экспрессии, содержит приблизительно 100-120 п.н., приблизительно 117 п.н. или приблизительно 100 п.н.

Согласно определенным вариантам осуществления раскрытая в настоящем документе кассета экспрессии содержит короткий промотор, содержащий или состоящий из следующего: (i) SEQ ID NO: 1; (ii) его вариант или функциональный фрагмент или (iii) последовательность нуклеиновой кислоты, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к любому из (i) или (ii), функционально связанная с полинуклеотидом, кодирующим любой из eTF, который избирательно активирует раскрытый в настоящем документе SCN1A и необязательно содержащая раскрытый в настоящем документе сайт связывания микроРНК. Другие примеры короткой промоторной последовательности можно найти в публикации PCT №WO 2018/213786.

Согласно определенным вариантам осуществления раскрытая в настоящем документе кассета экспрессии содержит представленный в настоящем документе полинуклеотид (например, полинуклеотид, содержащий последовательность, кодирующую eTF, которая активирует SCN1A и/или содержит селективный в отношении PV сайт связывания микроРНК) и селективный промотор клеточного типа, расположенный выше против хода транскрипции от последовательности, кодирующей трансген (например, eTF, который избирательно активирует SCN1A), чтобы иметь возможность избирательно управлять экспрессией трансгена в представляющей интерес клетке. Согласно определенным вариантам осуществления селективный промотор клеточного типа может быть селективным (например, избирательно управлять экспрессией) в отношении любого представляющего интерес типа клеток, такого как, например, клетка сердца, клетка печени, мышечная клетка, костная клетка, нейрон или их субпопуляции. Согласно иллюстративному варианту осуществления раскрытая в настоящем документе кассета экспрессии содержит полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, и селективный в отношении PV регуляторный элемент (например, промотор, энхансер и/или промотор и энхансер), расположенный выше против хода транскрипции от последовательности, кодирующей eTF, так чтобы быть способным избирательно управлять экспрессией eTF в PV-клетке, и, необязательно, селективный в отношении PV сайт связывания микроРНК. Селективный в отношении PV регуляторный элемент относится к регуляторному элементу, который специфически модулирует экспрессию гена в PV-содержащем нейроне. Согласно определенным вариантам осуществления селективные в отношении PV регуляторные элементы усиливают экспрессию в PV-содержащем нейроне по сравнению с одним или несколькими другими типами клеток ЦНС. Согласно определенным вариантам осуществления селективный в отноше-

нии PV регуляторный элемент подавляет процессы транскрипции и/или трансляции в нецелевых типах клеток.

Согласно определенным вариантам осуществления представленный в настоящем документе селективный в отношении PV регуляторный элемент приводит к селективной экспрессии гена в PV-клетке по сравнению с нецелевыми типами клеток. В некоторых случаях нецелевые типы клеток включают в себя, без ограничения, возбуждающие нейроны, не-PV типы клеток ЦНС и ненейрональные типы клеток ЦНС. Согласно определенным вариантам осуществления селективные в отношении PV регуляторные элементы приводят к селективной экспрессии генов в PV-содержащих нейронах по меньшей мере одного, двух, трех, четырех, пяти или более не содержащих PV типов клеток ЦНС. В некоторых случаях не содержащая PV клетка ЦНС представляет собой возбуждающий нейрон, дофаминергический нейрон, астроцит, микроглию, мотонейрон, сосудистую клетку или не-ГАМКергический нейрон, (например, клетка, которая не экспрессирует один или несколько из GAD2, GAD1, NKX2.1, DLX1, DLX5, SST и VIP), не содержащий PV нейрон (например, ГАМКергический нейрон, который не экспрессирует парвальбумин) или другие клетки ЦНС (например, типы клеток ЦНС, которые никогда не экспрессируют любой из PV, GAD2, GAD1, NKX2.1, DLX1, DLX5, SST и VIP). В некоторых случаях представленный в настоящем документе селективный в отношении PV регуляторный элемент приводит к повышенной селективности экспрессии генов в PV-содержащих нейронах по сравнению с не содержащими PV ГАМКергическими клетками. В некоторых случаях типы клеток различают по разным клеточным маркерам, морфологии, фенотипу, генотипу, функции и/или любым другим средствам классификации типов клеток.

Селективность экспрессии, управляемую селективным в отношении PV регуляторным элементом, можно измерить несколькими способами. Согласно одному варианту осуществления селективность экспрессии гена в PV-содержащей клетке по сравнению с не содержащими PV клетками может быть измерена путем сравнения количества PV-содержащих клеток, которые экспрессируют обнаруживаемый уровень транскрипта от гена, который функционально связан с селективным в отношении PV регуляторным элементом, с общим количеством клеток, которые экспрессируют ген (например, отношение PV к общему количеству клеток (PV + не-PV клеток), экспрессирующих ген). Например, селективность в отношении PV-содержащих нейронов может быть определена с использованием анализа колокализации на основе иммуногистохимии с использованием гена, кодирующего флуоресцентный белок (например, eGFP), функционально связанный с селективным в отношении PV регуляторным элементом для измерения экспрессии гена и антитела, которое идентифицирует PV-содержащие клетки (например, антитело к PV, которое специфически взаимодействует с PV-содержащими нейронами), связанное со второй флуоресцентной меткой (например, красным флуоресцентным белком). Селективность экспрессии в PV-содержащих клетках можно рассчитать путем деления количества клеток, которые экспрессируют как PV, так и eGFP (например, PV-содержащие клетки), на общее количество клеток, экспрессирующих eGFP (например, PV-содержащие клетки и не содержащие PV клетки), и умножения на 100, чтобы преобразовать в проценты. Чем выше процент PV-содержащих клеток, экспрессирующих трансген, тем более селективным является регуляторный элемент для PV-содержащих клеток. Согласно определенным вариантам осуществления представленный в настоящем документе селективный в отношении PV регуляторный элемент может быть высокоселективным в отношении экспрессии в PV-содержащих клетках. Например, представленный в настоящем документе селективный в отношении PV регуляторный элемент может характеризоваться селективностью, составляющей приблизительно 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более чем приблизительно 99%, в отношении PV-содержащих нейронов (например, PV-содержащие нейроны/все клетки \times 100).

В некоторых случаях представленный в настоящем документе селективный в отношении PV регуляторный элемент является коротким. В некоторых случаях размер селективного в отношении PV регуляторного элемента совместим с клонирующей способностью вектора, например, вирусного вектора или gAAV, так что комбинированный размер трансгена и одного или нескольких селективных в отношении PV регуляторных элементов не превышает клонирующую способность вектора. В некоторых случаях селективный в отношении PV регуляторный элемент составляет в длину приблизительно до 2050 п.н., 2000 п.н., 1900 п.н., 1800 п.н., 1700 п.н., 1600 п.н., 1500 п.н., 1400 п.н., 1300 п.н., 1200 п.н., 1100 п.н., 1000 п.н., 900 п.н., 800 п.н., 700 п.н., 600 п.н., 500 п.н., 400 п.н., 300 п.н., 200 п.н. или 100 п.н. В некоторых случаях селективный в отношении PV регуляторный элемент составляет приблизительно 500-600 п.н., 500-700 п.н., 500-800 п.н., 500-900 п.н., 500-1000 п.н., 500-1500 п.н., 500-2000 п.н. или 500-2050 п.н.

Согласно определенным вариантам осуществления представленный в настоящем документе селективный в отношении PV регуляторный элемент содержит или состоит из любого из следующего: (i) SEQ ID NO: 2-4; (ii) вариант, функциональный фрагмент или их комбинация или (iii) последовательность нуклеиновой кислоты, характеризующаяся идентичностью последовательности, составляющей по меньшей мере 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% по отношению к любому из (i) или (ii). В некоторых случаях регуляторный элемент включает в себя любую из SEQ ID NO: 2-4. Другие примеры селективных в отношении PV регуляторных элементов можно найти в публикации PCT № WO 2018/187363.

Согласно иллюстративным вариантам осуществления в настоящей заявке представлены кассеты

цию (i) или (iii) последовательность нуклеиновой кислоты, характеризующуюся идентичностью последовательности, составляющей по меньшей мере 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к любым из (i) или (ii), и (3) селективный в отношении PV сайт связывания микроРНК, содержащий (i) SEQ ID NO: 15; (ii) вариант, функциональный фрагмент или комбинацию (i) или (iii) последовательность нуклеиновой кислоты, характеризующуюся идентичностью последовательности, составляющей по меньшей мере 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99%, по отношению к любому из (i) или (ii).

Согласно определенным вариантам осуществления представленная в настоящем документе кассета экспрессии, содержащая селективный в отношении PV регуляторный элемент и селективный в отношении PV сайт связывания микроРНК, характеризующийся размером менее чем 5 т.п.н., 4,9 т.п.н., 4,8 т.п.н., 4,7 т.п.н., 4,6 т.п.н., 4,5 т.п.н., 4,4 т.п.н., 4,3 т.п.н., 4,2 т.п.н., 4,1 т.п.н., 4,0 т.п.н., 3,9 т.п.н., 3,8 т.п.н., 3,7 т.п.н., 3,6 т.п.н., 3,5 т.п.н., 3,4 т.п.н., 3,3 т.п.н., 3,2 т.п.н., 3,1 т.п.н., 3,0 т.п.н., 2,9 т.п.н., 2,8 т.п.н., 2,7 т.п.н., 2,6 т.п.н., 2,5 т.п.н., 2,4 т.п.н., 2,3 т.п.н., 2,2 т.п.н., 2,1 т.п.н., 2,0 т.п.н., 1,9 т.п.н., 1,8 т.п.н., 1,7 т.п.н., 1,6 т.п.н. или 1,5 т.п.н. или менее, или приблизительно 1,5-5 т.п.н., 1,5-4,7 т.п.н., 1,5-4,5 т.п.н., 1,5-4,0 т.п.н., 1,5-3,5 т.п.н., 1,5-3,0 т.п.н., 1,5-2,5 т.п.н., 1,5-2,0 т.п.н.

Согласно определенным вариантам осуществления представленная в настоящем документе кассета экспрессии может содержать еще один дополнительный регуляторный элемент в дополнение к промотору, такой как, например, последовательности, связанные с инициацией или окончанием транскрипции, энхансерные последовательности и эффективные сигналы процессинга РНК. Иллюстративные регуляторные элементы включают в себя, например, интрон, энхансер, UTR, элемент стабильности, последовательность WPRE, консенсусную последовательность Kozak, посттрансляционный элемент ответа, сайт связывания микроРНК или последовательность полиаденилирования (полиА) или их комбинацию. Регуляторные элементы могут функционировать, чтобы модулировать экспрессию генов на фазе транскрипции, посттранскрипционной фазе или на фазе трансляции экспрессии генов. На уровне РНК регуляция может происходить на уровне трансляции (например, элементы стабильности, которые стабилизируют мРНК для трансляции), расщепления РНК, сплайсинга РНК и/или терминации транскрипции. Согласно различным вариантам осуществления регуляторные элементы могут рекрутировать факторы транскрипции в кодирующую область, которые повышают избирательность экспрессии генов в представляющем интерес типе клеток, увеличивают скорость, с которой производятся транскрипты РНК, повышают стабильность производимой РНК и/или увеличивают скорость образования синтеза белка из транскриптов РНК. Согласно иллюстративному варианту осуществления представленная в настоящем документе кассета экспрессии содержит по меньшей мере один селективный в отношении PV сайт связывания микроРНК, как предусмотрено в настоящем документе.

Согласно определенным вариантам осуществления описанные в настоящем документе кассеты экспрессии дополнительно содержат последовательность полиА. Подходящие последовательности полиА включают в себя, например, искусственный полиА длиной приблизительно 75 п.н. (РА75) (см., например, WO 2018/126116), полиА бычьего гормона роста, ранний сигнал полиА SV40, поздний сигнал полиА SV40, полиА бета-глобина кролика, полиА тимидинкиназы HSV, полиА гена протамина, полиА аденовируса 5 E1b, полиА гормона роста или полиА PBGD. Согласно иллюстративным вариантам осуществления последовательность полиА, подходящая для применения в представленных в настоящем документе кассетах экспрессии, представляет собой полиА hGH (SEQ ID NO: 17) или синтетический полиА (SEQ ID NO: 16). Как правило, последовательность полиА расположена ниже по ходу транскрипции от полинуклеотида, кодирующего eTF, в описанных в настоящем документе кассетах экспрессии.

Согласно определенным вариантам осуществления представленные в настоящем документе кассеты экспрессии дополнительно содержат одну или несколько последовательностей нуклеиновых кислот, кодирующих один или несколько сигналов ядерной локализации (NLS). Может быть использован любой пептид NLS, который облегчает импорт белка, к которому прикреплен, в ядро клетки. Примеры NLS включают в себя, например, NLS большого Т-антигена SV40, NLS нуклеоплазмина, NLS EGL-13, NLS с-Мус и NLS TUS-белка. Смотрите, например, C. Dingwall et al., *J. Cell Biol.* 107: 841-9 (1988); J.P. Makkerh et al., *Curr Biol.* 6: 1025-7 (1996) и M. Ray et al., *Bioconj. Chem.* 26: 1004-7 (2015). NLS может быть расположен в любом месте последовательности белка eTF, но согласно предпочтительным вариантам осуществления он конъюгирован с N-концом eTF или доменом eTF. Согласно иллюстративным вариантам осуществления представленные в настоящем документе кассеты нуклеиновых кислот кодируют eTF с NLS, слитым с N-концом eTF. Согласно другим вариантам осуществления представленные в настоящем документе кассеты нуклеиновых кислот кодируют eTF с первым NLS, слитым с N-концом eTF, и вторым NLS, расположенным между DBD и доменом TAD eTF.

Векторы экспрессии.

Согласно определенным вариантам осуществления описанные в настоящем документе кассеты экспрессии могут быть включены в вектор экспрессии. Векторы экспрессии можно использовать для доставки кассеты экспрессии в целевую клетку посредством трансфекции или трансдукции. Вектор может представлять собой интегрирующий или неинтегрирующий вектор, что относится к способности вектора

интегрировать кассету экспрессии или трансген в геном клетки-хозяина. Примеры векторов экспрессии включают в себя, без ограничения, (а) невирусные векторы, такие как векторы нуклеиновых кислот, включая в себя линейные олигонуклеотиды и кольцевые плазмиды; искусственные хромосомы, такие как искусственные хромосомы человека (HAC), искусственные хромосомы дрожжей (YAC) и бактериальные искусственные хромосомы (BAC или PAC)); эписомальные векторы; транспозоны (например, PiggyBac); и (b) вирусные векторы, такие как ретровирусные векторы, лентивирусные векторы, аденовирусные векторы и аденоассоциированные вирусные векторы.

Векторы экспрессии могут представлять собой линейные олигонуклеотиды или кольцевые плазмиды и могут быть доставлены в клетку с помощью различных способов трансфекции, включая в себя физические и химические способы. Физические способы, как правило, относятся к способам доставки, использующим физическую силу для противодействия барьеру клеточной мембраны для облегчения внутриклеточной доставки генетического материала. Примеры физических способов включают в себя использование иглы, баллистической ДНК, электропорацию, сонопорацию, фотопорацию, магнитофекцию и гидропорацию. Химические способы, как правило, относятся к способам, при которых химические носители доставляют молекулу нуклеиновой кислоты в клетку, и могут включать в себя неорганические частицы, векторы на основе липидов, векторы на основе полимеров и векторы на основе пептидов.

Согласно определенным вариантам осуществления вектор экспрессии вводят в целевую клетку с использованием катионного липида (например, катионной липосомы). Для доставки генов были исследованы различные типы липидов, такие как, например, липидная наноэмульсия (например, которая представляет собой дисперсию одной несмешивающейся жидкости в другой, стабилизированной эмульгирующим средством) или твердые липидные наночастицы.

Согласно определенным вариантам осуществления вектор экспрессии вводят в целевую клетку с использованием носителя для доставки на основе пептидов. Носители на основе пептидов могут иметь преимущества, заключающиеся в защите доставляемого генетического материала, нацеливании на специфические клеточные рецепторы, разрушении эндосомных мембран и доставке генетического материала в ядро. Согласно определенным вариантам осуществления вектор экспрессии вводят в целевую клетку с использованием носителя для доставки на основе полимера. Носители на основе полимеров могут содержать природные белки, пептиды и/или полисахариды или синтетические полимеры. Согласно одному варианту осуществления средство доставки на основе полимера содержит полиэтиленимин (PEI). PEI может конденсировать ДНК в положительно заряженные частицы, которые связываются с остатками анионной поверхности клетки и попадают в клетку посредством эндоцитоза. Согласно другим вариантам осуществления носитель для доставки на основе полимера может содержать поли-L-лизин (PLL), поли(DL-молочную кислоту) (PLA), поли(DL-лактид-ко-гликозид) (PLGA), полиорнитин, полиаргинин, гистоны, протамины, дендримеры, хитозаны, синтетические аминопроизводные декстрана и/или катионные акриловые полимеры. Согласно определенным вариантам осуществления носители для доставки на основе полимеров могут включать в себя смесь полимеров, такую как, например, PEG и PLL.

Согласно определенным вариантам осуществления вектор экспрессии может представлять собой вирусный вектор, подходящий для генной терапии. Предпочтительные характеристики векторов вирусной генной терапии или векторов доставки генов могут включать в себя способность воспроизводимо и стабильно размножаться и очищаться до высоких титров; для опосредования направленной доставки (например, для доставки трансгена конкретно в представляющую интерес ткань или орган без широкого распространения вектора в другом месте); и опосредовать доставку гена и экспрессию трансгена, не вызывая вредных побочных эффектов.

Несколько типов вирусов, например, непатогенный парвовирус, называемый аденоассоциированным вирусом, были сконструированы для целей генной терапии, используя путь вирусной инфекции, но избегая последующей экспрессии вирусных генов, которая может привести к репликации и токсичности. Такие вирусные векторы можно получить путем удаления всех или некоторых кодирующих областей из вирусного генома, но оставив нетронутыми те последовательности (например, концевые повторяющиеся последовательности), которые могут быть необходимы для таких функций, как упаковка векторного генома в вирусный капсид или интеграция векторной нуклеиновой кислоты (например, ДНК) в хроматин хозяина.

Согласно различным вариантам осуществления подходящие вирусные векторы включают в себя ретровирусы (например, вирусы А-типа, В-типа, С-типа и D-типа), аденовирус, парвовирус (например, аденоассоциированные вирусы или AAV), коронавирусы, вирусы с отрицательно-полярной нитью РНК, такие как ортомиксовирус (например, вирус гриппа), рабдовирус (например, вирус бешенства и везикулярного стоматита), парамиксовирус (например, корь и Сендай), вирусы с положительно-полярной нитью РНК, такие как пикорнавирус и альфавирус, и двухцепочечные ДНК-вирусы, включая в себя аденовирус, вирус герпеса (например, вирус простого герпеса типов 1 и 2, вирус Эпштейна-Барра, цитомегаловирус) и поксвирус (например, коровьей оспы, оспы птиц и оспы канареек). Примеры ретровирусов включают в себя вирус лейкемии-саркомы птиц, человеческий Т-лимфотрофный вирус типа 1 (HTLV-1), вирус лейкемии крупного рогатого скота (BLV), лентивирус и спумавирус. Другие вирусы включают в себя, например, вирус Норуолк, тогавирус, флавивирус, реовирусы, паповавирус, гепаднавирус и вирус

гепатита. Вирусные векторы можно разделить на две группы в зависимости от их способности интегрироваться в геном хозяина - интегрирующиеся и неинтегрирующиеся. Онкоретровирусы и лентивирусы могут интегрироваться в хроматин клетки-хозяина, в то время как аденовирусы, аденоассоциированные вирусы и вирусы герпеса преимущественно сохраняются в ядре клетки в виде внехромосомных эписом.

Согласно определенным вариантам осуществления подходящий вирусный вектор представляет собой ретровирусный вектор. Ретровирусы относятся к вирусам семейства Retroviridae. Примеры ретровирусов включают в себя онкоретровирусы, такие как вирус лейкемии мышей (MLV), и лентивирусы, такие как вирус иммунодефицита человека 1 (HIV-1). Ретровирусные геномы представляют собой одноцепочечные (оц) РНК и содержат различные гены, которые могут быть в цис- или транс-форме. Например, ретровирусный геном может содержать цис-действующие последовательности, такие как два длинных концевых повтора (LTR), с элементами для экспрессии генов, обратной транскрипции и интеграции в хромосомы хозяина. Другие компоненты включают в себя сигнал упаковки (ψ или Ψ) для специфической упаковки РНК во вновь образованные вирионы и полипуриновый тракт (PPT), место инициации синтеза положительно-полярной нити ДНК во время обратной транскрипции. Кроме того, ретровирусный геном может содержать гены gag, pol и env. Ген gag кодирует структурные белки, ген pol кодирует ферменты, которые сопровождают оцРНК и осуществляют обратную транскрипцию вирусной РНК в ДНК, а ген env кодирует вирусную оболочку. Как правило, gag, pol и env предоставляются в транс для вирусной репликации и упаковки.

Согласно определенным вариантам осуществления представленный в настоящем документе ретровирусный вектор может представлять собой лентивирусный вектор. Распознаются по меньшей мере пять серогрупп или серотипов лентивирусов. Вирусы разных серотипов могут по-разному инфицировать определенные типы клеток и/или хозяев. Лентивирусы, например, включают в себя ретровирусы приматов и ретровирусы неprimатов. Ретровирусы приматов включают в себя HIV и вирус иммунодефицита обезьян (SIV). Ретровирусы неprimатов включают в себя вирус иммунодефицита кошек (FIV), вирус иммунодефицита крупного рогатого скота (BIV), вирус артрита-энцефалита коз (CAEV), вирус инфекционной анемии лошадей (EIAV) и виснавирус. Лентивирусы или лентивекторы могут быть способны трансдуцировать покоящиеся клетки. Как и в случае онкоретровирусных векторов, конструкция лентивекторов может быть основана на разделении цис- и транс-действующих последовательностей.

Согласно определенным вариантам осуществления в настоящей заявке представлены векторы экспрессии, которые были разработаны для доставки с помощью оптимизированного терапевтического ретровирусного вектора. Ретровирусный вектор может представлять собой лентивирус, содержащий левый (5') LTR; последовательности, которые способствуют упаковке и/или ядерному импорту вируса; промотор; необязательно, один или несколько дополнительных регуляторных элементов (таких как, например, энхансер или последовательность полиА); необязательно лентивирусный элемент обратного ответа (RRE); конструкция, содержащая селективный регуляторный элемент PV, функционально связанный с последовательностью, кодирующей eTF; необязательно инсулятор и правый (3') ретровирусный LTR.

Согласно иллюстративным вариантам осуществления представленный в настоящем документе вирусный вектор представляет собой аденоассоциированный вирус (AAV). AAV представляет собой небольшой вирус животных без оболочки, дефектной репликации, который поражает людей и некоторые другие виды приматов. Известно, что AAV вызывает заболевание человека и вызывает умеренный иммунный ответ. Векторы AAV могут также инфицировать как делящиеся, так и покоящиеся клетки, не встраиваясь в геном клетки-хозяина.

Геном AAV состоит из линейной одноцепочечной ДНК длиной ~4,7 т.п.н. Геном состоит из двух открытых рамок считывания (ORF), фланкированных последовательностью инвертированного концевого повтора (ITR) длиной приблизительно 145 п.н. ITR состоит из нуклеотидной последовательности на 5'-конце (5' ITR) и нуклеотидной последовательности, расположенной на 3'-конце (3' ITR), которые содержат палиндромные последовательности. ITR функционируют в цис-системе путем сворачивания с образованием Т-образных шпильчатых структур за счет комплементарного спаривания оснований, которые действуют как праймеры во время инициации репликации ДНК для синтеза второй цепи. Две открытые рамки считывания кодируют гены гер и сар, которые участвуют в репликации и упаковке вириона. Согласно иллюстративному варианту осуществления представленный в настоящем документе вектор AAV не содержит генов гер или сар. Такие гены могут быть представлены в транс для производства вирионов, как описано ниже.

Согласно определенным вариантам осуществления вектор AAV может включать в себя "лишнюю" нуклеиновую кислоту. Согласно определенным вариантам осуществления "лишняя" нуклеиновая кислота может кодировать зеленый флуоресцентный белок или ген устойчивости к антибиотикам, таким как канамицин или ампициллин. Согласно определенным вариантам осуществления "лишняя" нуклеиновая кислота может быть расположена вне последовательностей ITR (например, по сравнению с последовательностью трансгена eTF и регуляторными последовательностями, которые расположены между 5' и 3' ITR-последовательностями).

Существуют различные серотипы AAV, включая в себя AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12 и AAV13. Эти серотипы различаются по своему тропизму

или по типам клеток, которые они заражают. AAV могут содержать геном и капсиды из множества серотипов (например, псевдотипов). Например, AAV может содержать геном серотипа 2 (например, ITR), упакованный в капсид серотипа 5 или серотипа 9. Псевдотипы могут повышать эффективность трансдукции, а также изменять тропизм.

В некоторых случаях предпочтителен серотип AAV, который может пересекать гематоэнцефалический барьер или инфицировать клетки ЦНС. В некоторых случаях AAV9 или его вариант используется для доставки кассеты экспрессии согласно настоящему раскрытию, содержащей селективный регуляторный элемент PV, функционально связанный с трансгеном, кодирующим eTF, который избирательно активирует SCN1A. В некоторых случаях AAV9 или его вариант используется для доставки кассеты экспрессии согласно настоящему раскрытию, содержащей селективный в отношении PV сайт связывания микроРНК. В некоторых случаях AAV9 или его вариант используется для доставки кассеты экспрессии согласно настоящему раскрытию, содержащей селективный в отношении PV регуляторный элемент, функционально связанный с трансгеном, кодирующим eTF, который селективно активирует SCN1A, и селективный в отношении PV сайт связывания микроРНК.

Согласно иллюстративным вариантам осуществления в настоящей заявке представлены векторы экспрессии, которые были разработаны для доставки с помощью AAV. AAV может быть любого серотипа, например, AAV1, AAV2, AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV-DJ, или химерным, гибридным или вариантом AAV. AAV также может представлять собой самокомплементарный AAV (scAAV). Согласно определенным вариантам осуществления вектор экспрессии, разработанный для доставки с помощью AAV, содержит 5' ITR и 3' ITR. Согласно определенным вариантам осуществления вектор экспрессии, разработанный для доставки с помощью AAV, содержит 5' ITR, промотор, трансген, кодирующий eTF, и 3' ITR. Согласно определенным вариантам осуществления вектор экспрессии, разработанный для доставки с помощью AAV, содержит 5' ITR, энхансер, промотор, трансген, кодирующий eTF, последовательность полиА и 3' ITR.

Клетки-хозяева.

Согласно другому аспекту настоящее изобретение относится к клетке-хозяину, содержащей кассету экспрессии или вектор экспрессии, как раскрыто в настоящем документе. Клетки-хозяева могут представлять собой бактериальную клетку, дрожжевую клетку, клетку насекомого или клетку млекопитающего. Согласно иллюстративному варианту осуществления под клеткой-хозяином понимается любая линия клеток, которая восприимчива к заражению представляющим интерес вирусом и поддается культивированию *in vitro*.

Согласно определенным вариантам осуществления предложенная в настоящем документе клетка-хозяин может использоваться для целей генной терапии *ex vivo*. Согласно таким вариантам осуществления клетки трансфицируют молекулой нуклеиновой кислоты или вектором экспрессии, содержащим селективный в отношении PV сайт связывания микроРНК и/или последовательность, кодирующую eTF, которая избирательно активирует SCN1A, как описано в настоящем документе, и затем трансплантируют пациенту или субъекту. Трансплантированные клетки могут иметь аутологичное, аллогенное или гетерологичное происхождение. Для клинического применения выделение клеток, как правило, проводят в условиях надлежащей производственной практики (GMP). Перед трансплантацией, как правило, проверяется качество клеток и отсутствие микробов или других загрязнителей, и может проводиться предварительное кондиционирование, такое как облучение и/или иммуносупрессивное лечение. Кроме того, клетки-хозяева можно трансплантировать вместе с факторами роста для стимуляции пролиферации и/или дифференцировки клеток.

Согласно определенным вариантам осуществления клетка-хозяин может использоваться для генной терапии *ex vivo*. Предпочтительно указанные клетки представляют собой эукариотические клетки, такие как клетки млекопитающих, они включают в себя, без ограничения, людей, отличных от людей приматов, таких как обезьяны; шимпанзе; мартышки и орангутаны, домашних животных, включая в себя собак и кошек, а также домашний скот, такой как лошади, крупный рогатый скот, свиньи, овцы и козы, или другие виды млекопитающих, включая в себя, без ограничения, мышей, крыс, морских свинок, кроликов, хомяков и т.п. Специалист в настоящей области техники выберет более подходящие клетки в соответствии с пациентом или объектом трансплантации.

Согласно определенным вариантам осуществления представленная в настоящем документе клетка-хозяин может представлять собой клетку со свойствами самообновления и плюрипотентности, например, стволовые клетки или индуцированные плюрипотентные стволовые клетки. Стволовые клетки предпочтительно представляют собой мезенхимальные стволовые клетки. Мезенхимальные стволовые клетки (MSC) способны дифференцироваться по меньшей мере в один из остеобластов, хондроцитов, адипоцитов или миоцитов и могут быть выделены из любого типа ткани. Как правило, MSC выделяют из костного мозга, жировой ткани, пуповины или периферической крови. Способы их получения хорошо известны специалисту в настоящей области техники. Индуцированные плюрипотентные стволовые клетки (также известные как iPS-клетки или iPSC) представляют собой тип плюрипотентных стволовых клеток, которые могут быть получены непосредственно из взрослых клеток. Yamanaka с соавт. индуцировали iPS-клетки путем переноса генов Oct3/4, Sox2, Klf4 и c-Myc в фибробласты мыши и человека и вынуждая

клетки экспрессировать гены (WO 2007/069666). Thomson с соавт. впоследствии производили iPS-клетки человека с использованием Nanog и Lin28 вместо Klf4 и c-Myc (WO 2008/118820).

Согласно иллюстративному варианту осуществления представленная в настоящем документе клетка-хозяин представляет собой упаковывающую клетку. Указанные клетки могут представлять собой прикрепленные или суспензионные клетки. Упаковывающая клетка и вспомогательный вектор или вирус, или конструкция(и) ДНК обеспечивают вместе в транс все недостающие функции, которые требуются для полной репликации и упаковки вирусного вектора.

Предпочтительно указанные упаковывающие клетки представляют собой эукариотические клетки, такие как клетки млекопитающих, включая в себя клетки обезьян, человека, собак и грызунов. Примерами клеток человека являются клетки PER.C6 (WO01/38362), MRC-5 (ATCC CCL-171), WI-38 (ATCC CCL-75), клетки HEK-293 (ATCC CRL-1573), клетки HeLa (ATCC CCL2) и фетальные клетки легких режус (ATCC CL-160). Примерами клеток отличных от людей приматов являются клетки Vero (ATCC CCL81), клетки COS-1 (ATCC CRL-1650) или клетки COS-7 (ATCC CRL-1651). Примерами клеток собаки являются клетки MDCK (ATCC CCL-34). Примерами клеток грызунов являются клетки хомяка, такие как BHK21-F, клетки НКСС или клетки CHO.

В качестве альтернативы источникам млекопитающих клеточные линии для применения в настоящем изобретении могут быть получены из источников птиц, таких как курица, утка, гусь, перепел или фазан. Примеры птичьих клеточных линий включают в себя птичьи эмбриональные стволовые клетки (WO01/85938 и WO03/076601), иммортализованные клетки сетчатки утки (WO2005/042728) и клетки, полученные из птичьих эмбриональных стволовых клеток, включая в себя клетки курицы (WO2006/108846) или клетки утки, такие как клеточная линия EB66 (WO2008/129058 и WO2008/142124).

Согласно другому варианту осуществления указанные клетки-хозяева представляют собой клетки насекомых, такие как клетки SF9 (ATCC CRL-1711), клетки Sf21 (IPLB-Sf21), клетки MG1 (BTI-TN-MG1) или клетки High Five™ (BTI-TN-5B1-4).

Согласно определенным вариантам осуществления представленные в настоящем документе клетки-хозяева, содержащие рекомбинантный вектор/геном AAV по настоящему изобретению (например, содержащие селективный в отношении PV сайт связывания микроРНК и/или последовательность, кодирующую eTF, который избирательно активирует SCN1A), могут дополнительно содержать одну или несколько конструкций нуклеиновых кислот, такие как, например, (i) конструкция нуклеиновой кислоты (например, плаزمид-помощник AAV), которая кодирует гены гер и сар, но не несет последовательностей ITR; и/или (ii) конструкция нуклеиновой кислоты (например, плазмид), обеспечивающая аденовирусные функции, необходимые для репликации AAV. Согласно иллюстративному варианту осуществления представленная в настоящем документе клетка-хозяин содержит: i) вектор экспрессии, содержащий селективный в отношении PV сайт связывания микроРНК и/или последовательность, кодирующую eTF, который избирательно активирует SCN1A, как предусмотрено в настоящем документе (т.е. рекомбинантный геном AAV); ii) конструкцию нуклеиновой кислоты, кодирующую гены гер и сар AAV, которая не несет последовательности ITR; и iii) конструкцию нуклеиновой кислоты, содержащую аденовирусные гены-помощники (как описано ниже).

Согласно определенным вариантам осуществления гены-помощники гер, сар и аденовируса могут быть объединены в одной плазмиде (Blouin V et al. *J Gene Med.* 2004; 6(suppl): S223-S228; Grimm D. et al. *Hum. Gene Ther.* 2003; 7: 839-850). Таким образом, согласно другому иллюстративному варианту осуществления представленная в настоящем документе клетка-хозяин содержит: i) вектор экспрессии, содержащий селективный в отношении PV сайт связывания микроРНК и/или последовательность, кодирующую eTF, которая избирательно активирует SCN1A, как раскрыто в настоящем документе (т.е. рекомбинантный геном AAV); и ii) плазмиду, кодирующую гены гер и сар AAV, которая не несет последовательности ITR и дополнительно содержит гены-помощники аденовирусов.

Согласно другому варианту осуществления представленная в настоящем документе клетка-хозяин содержит: а) вектор экспрессии, содержащий селективный в отношении PV сайт связывания микроРНК и/или последовательность, кодирующую eTF, которая избирательно активирует SCN1A, как описано в настоящем документе (т.е. геном рекомбинантного AAV); б) плазмиду, кодирующую гены гер и сар AAV, которая не несет последовательности ITR; и с) плазмиду, содержащую аденовирусные РНК генов-помощников E2a, E4 и VA; причем котрансфекция выполняется в клетках, предпочтительно в клетках млекопитающих, которые конститутивно экспрессируют и характеризуются транскомплементацией к аденовирусному гену E1, как клетки HEK-293 (ATCC CRL-1573).

Согласно определенным вариантам осуществления клетка-хозяин, подходящая для крупномасштабного производства векторов AAV, представляет собой клетки насекомых, которые можно инфицировать комбинацией рекомбинантных бакуловирусов (Urabe et al. *Hum. Gene Ther.* 2002; 13: 1935-1943). Например, клетки SF9 могут быть совместно инфицированы тремя бакуловирусными векторами, соответственно экспрессирующими гер AAV, сар AAV и упаковываемый вектор AAV. Рекомбинантные бакуловирусные векторы будут обеспечивать функции вирусного гена-помощника, необходимые для репликации и/или упаковки вируса.

Дополнительное руководство по конструированию и производству вирионов для генной терапии

согласно настоящему изобретению можно найти в: *Viral Vectors for Gene Therapy, Methods and Protocols*. Series: *Methods in Molecular Biology*, Vol. 737. Merten and Al-Rubeai (Eds.); 2011 Humana Press (Springer); *Gene Therapy*. M. Giacca. 2010 Springer-Verlag; Heilbronn R. and Weger S. *Viral Vectors for Gene Transfer: Current Status of Gene Therapeutics*. In: *Drug Delivery, Handbook of Experimental Pharmacology* 197; M. Schafer-Korting (Ed.). 2010 Springer-Verlag; pp. 143-170; *Adeno-Associated Virus: Methods and Protocols*. R. O. Snyder and P. Moullier (Eds). 2011 Humana Press (Springer); Bunning H. et al. Recent developments in adeno-associated virus technology. *J. Gene Med.* 2008; 10:717-733 и *Adenovirus: Methods and Protocols*. M. Chillon and A. Bosch (Eds.); Third. Edition. 2014 Humana Press (Springer).

Вирионы и способы получения вирионов.

Согласно определенным вариантам осуществления в настоящей заявке представлены вирусные частицы, содержащие вирусный вектор, содержащий селективный в отношении PV сайт связывания микроРНК и/или последовательность, кодирующую eTF, который избирательно активирует SCN1A, как раскрыто в настоящем документе. Термины "вирусная частица" и "вирион" используются в настоящем документе взаимозаменяемо и относятся к инфекционной и, как правило, дефектной по репликации вирусной частице, содержащей вирусный геном (например, вектор вирусной экспрессии), упакованный внутри капсида, и, в зависимости от случая, например, для ретровирусов - липидную оболочку, окружающую капсид. "Капсид" относится к структуре, в которую упакован вирусный геном. Капсид состоит из нескольких олигомерных структурных субъединиц, состоящих из белков. Например, у AAV есть икосаэдрический капсид, образованный взаимодействием трех капсидных белков: VP1, VP2 и VP3. Согласно одному варианту осуществления представленный в настоящем документе вирион представляет собой рекомбинантный вирион AAV или вирион gAAV, полученный путем упаковки вектора AAV, содержащего селективный в отношении PV регуляторный элемент и селективный в отношении PV сайт связывания микроРНК. Согласно другому варианту осуществления представленный в настоящем документе вирион представляет собой рекомбинантный вирион AAV или вирион gAAV, полученный путем упаковки вектора AAV, содержащего селективный в отношении PV регуляторный элемент, функционально связанный с последовательностью, кодирующей eTF, которая избирательно активирует SCN1A, как описано в настоящем документе, в белковой оболочке. Согласно другому варианту осуществления представленный в настоящем документе вирион представляет собой рекомбинантный вирион AAV или вирион gAAV, полученный путем упаковки вектора AAV, содержащего селективный в отношении PV регуляторный элемент, функционально связанный с последовательностью, кодирующей eTF, которая избирательно активирует SCN1A, как описано в настоящем документе, и селективный в отношении PV сайт связывания микроРНК в белковой оболочке.

Согласно определенным вариантам осуществления представленный в настоящем документе рекомбинантный вирион AAV может быть получен путем инкапсидации генома AAV, полученного из определенного серотипа AAV, в вирусную частицу, образованную природными белками Cap, соответствующими AAV того же конкретного серотипа. Согласно другим вариантам осуществления представленная в настоящем документе вирусная частица AAV содержит вирусный вектор, содержащий ITR данного серотипа AAV, упакованный в белки из другого серотипа. Смотрите, например, Bunning H et al. *J Gene Med* 2008; 10: 717-733. Например, вирусный вектор, имеющий ITR из данного серотипа AAV, может быть упакован в: а) вирусную частицу, состоящую из капсидных белков, полученных из того же или другого серотипа AAV (например, ITR AAV2 и капсидных белков AAV9; ITR AAV2 и капсидных белков AAV8 и т.п.); б) мозаичную вирусную частицу, состоящую из смеси капсидных белков из различных серотипов или мутантов AAV (например, ITR AAV2 с капсидными белками AAV1 и AAV9); в) химерную вирусную частицу, состоящую из капсидных белков, которые были усечены перестановкой доменов между различными серотипами или вариантами AAV (например, ITR AAV2 с белками капсида AAV8 с доменами AAV9); или д) нацеленную вирусную частицу, сконструированную для отображения селективных связывающих доменов, обеспечивающих строгое взаимодействие со специфическими рецепторами целевой клетки (например, ITR AAV5 с капсидными белками AAV9, генетически усеченными путем вставки пептидного лиганда; или капсидные белки AAV9, не модифицированные генетически путем связывания пептидного лиганда на поверхности капсида).

Квалифицированному специалисту будет понятно, что представленный в настоящем документе вирион AAV может содержать капсидные белки любого серотипа AAV. Согласно одному варианту осуществления вирусная частица содержит капсидные белки из серотипа AAV, выбранного из группы, состоящей из AAV1, AAV2, AAV5, AAV8 и AAV9, которые более подходят для доставки в ЦНС (M. Nosque-miller et al., *Hum Gene Ther* 27(7): 478-496 (2016)). Согласно конкретному варианту осуществления вирусная частица содержит кассету экспрессии по настоящему изобретению, в которой последовательности 5'ITR и 3'ITR кассеты экспрессии относятся к серотипу AAV2, а белки капсида относятся к серотипу AAV9.

В настоящей области техники известны многочисленные способы получения вирионов gAAV, включая в себя трансфекцию, получение стабильных клеточных линий и системы производства инфекционных гибридных вирусов, которые включают в себя гибриды аденовирус-AAV, гибриды герпесвирус-AAV (Conway, J E et al., (1997) *J. Virology* 71(11):8780-8789) и гибриды бакуловирин-AAV. Все куль-

туры производства гAAV для производства вирусных частиц гAAV требуют: 1) подходящие клетки-хозяева, включая в себя, например, клеточные линии человеческого происхождения, такие как клетки HeLa, A549 или 293, или клеточные линии, полученные из насекомых, такие как SF-9, в случае систем производства бакуловирусов; 2) подходящей функции вируса-помощника, обеспечиваемой аденовирусом дикого типа или мутантным аденовирусом (таким как чувствительный к температуре аденовирус), вирусом герпеса, бакуловирусом или плазмидной конструкцией, обеспечивающей вспомогательные функции; 3) гены гер и сар AAV и генные продукты; 4) трансген (например, содержащий один или несколько из: селективный в отношении PV сайт связывания микроРНК, последовательность, кодирующую eTF, которая избирательно активирует SCN1A, как описано в настоящем документе, и/или селективный в отношении PV промотор), фланкированный последовательностями ITR AAV; и 5) подходящие носители и компоненты среды для поддержки производства гAAV.

Согласно различным вариантам осуществления описанные в настоящем документе клетки-хозяева содержат следующие три компонента: (1) ген гер и ген сар, (2) гены, обеспечивающие вспомогательные функции, и (3) трансген (например, содержащий один или несколько из: селективного в отношении PV сайта связывания микроРНК, последовательности, кодирующей eTF, который избирательно активирует SCN1A, как описано в настоящем документе, и/или селективного в отношении PV промотора), фланкированных ITR. Ген гер AAV, ген сар AAV и гены, обеспечивающие вспомогательные функции, могут быть введены в клетку путем включения указанных генов в вектор, такой как, например, плаزمид, и введения указанного вектора в клетку-хозяина. Гены гер, сар и гены, обеспечивающие вспомогательные функции, могут быть включены в одну и ту же плазмиду или в разные плазмиды. Согласно предпочтительному варианту осуществления гены гер и сар AAV включены в одну плазмиду, а гены, обеспечивающие вспомогательные функции, включены в другую плазмиду. Различные плазмиды для создания клетки-хозяина для производства вирионов (например, содержащие гены гер и сар AAV, вспомогательных функций или трансген) могут быть введены в клетку с использованием любого подходящего способа, хорошо известного в настоящей области техники. Примеры способов трансфекции включают в себя, без ограничения, соосаждение с фосфатом кальция, DEAE-декстраном, полибренном, электропорацию, микроинъекцию, опосредованное липосомами слияние, липофекцию, ретровирусную инфекцию и библистическую трансфекцию. Согласно определенным вариантам осуществления плазмиды, обеспечивающие гены гер и сар, вспомогательные функции и трансген, фланкированный ITR, могут быть введены в клетку одновременно. Согласно другому варианту осуществления плазмиды, обеспечивающие гены гер и сар и вспомогательные функции, могут быть введены в клетку до или после введения плазмиды, содержащей трансген. Согласно иллюстративному варианту осуществления клетки трансфицируют одновременно тремя плазмидами (например, способом тройной трансфекции): (1) плазмидой, содержащей трансген (например, содержащей один или несколько из: селективного в отношении PV сайта связывания микроРНК, последовательности, кодирующей eTF, который избирательно активирует SCN1A, как описано в настоящем документе, и/или селективного в отношении PV промотора), фланкированного ITR, (2) плазмидой, содержащей гены гер и сар AAV, и (3) плазмидой, содержащей гены, обеспечивающие вспомогательные функции. Иллюстративные клетки-хозяева могут представлять собой клетки 293, A549 или HeLa.

Согласно другим вариантам осуществления один или несколько из (1) генов гер и сар AAV, (2) генов, обеспечивающих вспомогательные функции, и (3) трансгена (например, содержащего один или несколько из: селективного в отношении PV сайта связывания микроРНК, последовательности, кодирующей eTF, который избирательно активирует SCN1A, как описано в настоящем документе, и/или селективного в отношении PV промотора), фланкированного ITR, может переноситься упаковывающей клеткой либо эпизомально, либо интегрироваться в геном упаковывающей клетки. Согласно одному варианту осуществления клетки-хозяева могут представлять собой упаковывающие клетки, в которых гены гер и сар AAV и гены, обеспечивающие вспомогательные функции, стабильно поддерживаются в клетке-хозяине, и клетка-хозяин временно трансфицируется плазмидой, содержащей трансген (например, содержащей один или несколько из: селективного в отношении PV сайта связывания микроРНК, последовательности, кодирующей eTF, который избирательно активирует SCN1A, как описано в настоящем документе, и/или селективного в отношении PV промотора), фланкированный ITR. Согласно другому варианту осуществления клетки-хозяева представляют собой упаковывающие клетки, в которых гены гер и сар AAV стабильно поддерживаются в клетке-хозяине, и клетка-хозяин временно трансфицируется плазмидой, содержащей трансген (например, содержащей один или несколько из следующего: селективный в отношении PV сайт связывания микроРНК, последовательность, кодирующая eTF, который избирательно активирует SCN1A, как описано в настоящем документе, и/или селективный в отношении PV промотор), фланкированный ITR и плазмидой, содержащей вспомогательные функции. Согласно другому варианту осуществления клетки-хозяева могут представлять собой упаковывающие клетки, в которых вспомогательные функции стабильно поддерживаются в клетке-хозяине, и клетка-хозяин временно трансфицируется плазмидой, содержащей трансген (например, содержащий один или несколько из следующего: селективный в отношении PV сайт связывания микроРНК, последовательность, кодирующая eTF, который избирательно активирует SCN1A, как описано в настоящем документе, и/или селективный

в отношении PV промотор), фланкированный ITR, и плазмидой, содержащей гены гер и сар. Согласно другому варианту осуществления клетки-хозяева могут представлять собой клеточные линии-продуценты, которые стабильно трансфицируются генами гер и сар, генами, обеспечивающими вспомогательные функции, и последовательностью трансгена (например, содержащего одно или несколько из следующего: селективный в отношении PV сайт связывания микроРНК, последовательность, кодирующая eTF, который избирательно активирует SCN1A, как описано в настоящем документе, и/или селективный в отношении PV промотор), фланкированного ITR. Иллюстративные упаковывающие клетки и клетки-продуценты могут происходить из клеток 293, A549 или HeLa.

Согласно другому варианту осуществления линия клеток-продуцентов представляет собой линию клеток насекомых (как правило, клетки Sf9), которые инфицированы векторами экспрессии бакуловируса, которые обеспечивают белки Rep и Cap. Для этой системы не требуются гены-помощники аденовируса (Ayuso E, et al., *Curr. Gene Ther.* 2010, 10:423-436).

Используемый в настоящем документе термин "белок сар" относится к полипептиду, обладающему по меньшей мере одной функциональной активностью нативного белка сар AAV (например, VP1, VP2, VP3). Примеры функциональной активности белков сар включают в себя способность индуцировать образование капсида, способствовать накоплению одноцепочечной ДНК, облегчать упаковку ДНК AAV в капсиды (т.е. инкапсидацию), связываться с клеточными рецепторами и облегчать проникновение вириона в клетки-хозяева. В принципе, в контексте настоящего изобретения можно использовать любой белок Сар.

Сообщалось, что белки сар влияют на тропизм хозяина, специфичность клеток, тканей или органов, использование рецепторов, эффективность заражения и иммуногенность вирусов AAV. Соответственно, сар AAV для применения в gAAV может быть выбран с учетом, например, вида субъекта (например, человека или не человека), иммунологического состояния субъекта, его пригодности для длительного или краткосрочного лечения или конкретного терапевтического применения (например, лечение конкретного заболевания или нарушения или доставка в определенные клетки, ткани или органы). Согласно определенным вариантам осуществления белок сар происходит из группы AAV, состоящей из серотипов AAV1, AAV2, AAV5, AAV8 и AAV9. Согласно иллюстративному варианту осуществления белок сар происходит из AAV9.

Согласно некоторым вариантам осуществления сар AAV для применения в способе по настоящему изобретению может быть получен путем мутагенеза (т.е. путем вставок, делеций или замен) одного из вышеупомянутых сар AAV или кодирующей его нуклеиновой кислоты. Согласно определенным вариантам осуществления сар AAV по меньшей мере на 70%, 75%, 80%, 85%, 90%, 95%, 98% или 99% или более аналогичен одному или нескольким из вышеупомянутых сар AAV.

Согласно некоторым вариантам осуществления сар AAV является химерным и содержит домены из двух, трех, четырех или более из вышеупомянутых сар AAV. Согласно определенным вариантам осуществления сар AAV представляет собой мозаику мономеров VP1, VP2 и VP3, происходящих из двух или трех различных AAV или рекомбинантного AAV. Согласно определенным вариантам осуществления композиция gAAV содержит более одного из вышеупомянутых сар.

Согласно некоторым вариантам осуществления сар AAV для применения в вирионе gAAV сконструирован таким образом, чтобы он содержал гетерологичную последовательность или другую модификацию. Например, последовательность пептида или белка, обеспечивающая избирательное нацеливание или уклонение от иммунитета, может быть преобразована в белок сар. В качестве альтернативы или в дополнение, сар может быть химически модифицирован так, чтобы поверхность gAAV была полиэтиленгликолированной (т.е. пегилированной), что может способствовать уклонению от иммунитета. Белок сар также может быть мутагенизирован (например, для удаления его естественного связывания с рецептором или для маскировки иммуногенного эпитопа).

Используемый в настоящем документе термин "белок гер" относится к полипептиду, обладающему по меньшей мере одной функциональной активностью нативного белка гер AAV (например, гер 40, 52, 68, 78). Примеры функциональной активности белка гер включают в себя любую активность, связанную с физиологической функцией белка, включая в себя облегчение репликации ДНК посредством распознавания, связывания и отсечения ориджина AAV репликации ДНК, а также активности ДНК-геликазы. Дополнительные функции включают в себя модуляцию транскрипции с промоторов AAV (или других гетерологичных) и сайт-специфическую интеграцию ДНК AAV в хромосому хозяина. Согласно конкретному варианту осуществления гены гер AAV могут происходить из серотипов AAV1, AAV2, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 или AAVrh10; более предпочтительно из серотипа AAV, выбранного из группы, состоящей из AAV1, AAV2, AAV5, AAV8 и AAV9.

Согласно определенным вариантам осуществления белок гер AAV для применения в способе по настоящему изобретению может быть получен путем мутагенеза (т.е. путем вставок, делеций или замен) одного из вышеупомянутых гер AAV или его кодирующей нуклеиновой кислоты. Согласно определенным вариантам осуществления гер AAV по меньшей мере на 70%, 75%, 80%, 85%, 90%, 95%, 98% или 99% или более аналогичен одному или нескольким из вышеупомянутых гер AAV.

Использованные в настоящем документе выражения "вспомогательные функции" или "гены-

помощники" относятся к вирусным белкам, репликация которых зависит от AAV. Вспомогательные функции включают в себя те белки, которые необходимы для репликации AAV, включая в себя, без ограничения, те белки, которые участвуют в активации транскрипции гена AAV, стадией специфическом сплайсинге мРНК AAV, репликации ДНК AAV, синтезе продуктов экспрессии сар и сборке капсида AAV. Вспомогательные функции на основе вирусов могут быть получены из любого из известных вспомогательных вирусов, таких как аденовирус, вирус герпеса (кроме вируса простого герпеса типа 1) и вирус осповакцины. Вспомогательные функции включают в себя, без ограничения, аденовирус E1, E2a, VA и E4 или герпесвирус UL5, ULB, UL52 и UL29 и герпесвирусную полимеразу. Согласно предпочтительному варианту осуществления белки, от которых зависит репликация AAV, происходят от аденовируса.

Согласно определенным вариантам осуществления вирусный белок, репликация которого зависит от AAV, для применения в способе по настоящему изобретению может быть получен путем мутагенеза (т.е. путем вставок, делеций или замен) одного из вышеупомянутых вирусных белков или кодирующей его нуклеиновой кислоты. Согласно определенным вариантам осуществления вирусный белок по меньшей мере на 70%, 75%, 80%, 85%, 90%, 95%, 98% или 99% или более аналогичен одному или нескольким из вышеупомянутых вирусных белков.

Способы анализа функций белков сар, белков гер и вирусных белков, от которых зависит репликация AAV, хорошо известны в настоящей области техники.

Клетки-хозяева для экспрессии представляющего интерес трансгена (например, содержащие одно или несколько из следующего: селективный в отношении PV сайт связывания микроРНК, последовательность, кодирующая eTF, который избирательно активирует SCN1A, как описано в настоящем документе, и/или селективный в отношении PV промотор) могут быть выращены в условиях, подходящих для сборки вирионов AAV. Согласно определенным вариантам осуществления клетки-хозяева выращивают в течение подходящего периода времени, чтобы способствовать сборке вирионов AAV и высвобождению вирионов в среду. Как правило, клетки можно выращивать в течение приблизительно 24 часов, приблизительно 36 часов, приблизительно 48 часов, приблизительно 72 часов, приблизительно 4 дней, приблизительно 5 дней, приблизительно 6 дней, приблизительно 7 дней, приблизительно 8 дней, приблизительно 9 дней или до приблизительно 10 дней. Приблизительно через 10 дней (или раньше, в зависимости от условий культивирования и конкретной используемой клетки-хозяина) уровень продукции, как правило, значительно снижается. Как правило, время культивирования измеряется с момента образования вируса. Например, в случае AAV производство вируса, как правило, начинается после обеспечения функции вируса-помощника в соответствующей клетке-хозяине, как описано в настоящем документе. Как правило, клетки собирают от приблизительно 48 до приблизительно 100, предпочтительно от приблизительно 48 до приблизительно 96, предпочтительно от приблизительно 72 до приблизительно 96, предпочтительно от приблизительно 68 до приблизительно 72 часов после заражения вирусом-помощником (или после начала производства вируса).

Культуры-продуценты гAAV можно выращивать в различных условиях (в широком диапазоне температур, в течение различных периодов времени и т.п.), подходящих для конкретной используемой клетки-хозяина. Культуры производства гAAV включают в себя культуры, зависящие от прикрепления, которые можно культивировать в подходящих емкостях, зависящих от прикрепления, таких как, например, вращающиеся флаконы, фильтры из полого волокна, микроносители и биореакторы с уплотненным слоем или с псевдооживленным слоем. Культуры производства вектора гAAV могут также включать в себя адаптированные к суспензии клетки-хозяева, такие как клетки HeLa, 293 и SF-9, которые можно культивировать различными способами, включая в себя, например, вращающиеся колбы, биореакторы с мешалкой и одноразовые системы, такие как система одноразовых реакторов фирмы Wave.

Подходящие среды, известные в настоящей области техники, могут использоваться для производства вирионов гAAV. Эти среды включают в себя, без ограничения, среды, производимые Nuclone Laboratories и JRH, включая в себя модифицированную среду Игла (MEM), модифицированную Дульбекко среду Игла (DMEM), каждая из которых полностью включена в настоящий документ посредством ссылки. Согласно определенным вариантам осуществления культуральная среда для производства гAAV может быть дополнена сывороткой или полученными из сыворотки рекомбинантными белками на уровне 0,5-20% (в объемном отношении или массово-объемном отношении). Альтернативно, векторы гAAV могут быть получены в бессывороточных условиях, которые также могут называться средами без продуктов животного происхождения.

После культивирования клеток-хозяев, чтобы обеспечить производство вирионов AAV, полученные вирионы могут быть собраны и очищены. Согласно определенным вариантам осуществления вирионы AAV могут быть получены из (1) клеток-хозяев производственной культуры путем лизиса клеток-хозяев и/или (2) культуральной среды указанных клеток через период времени после трансфекции, предпочтительно 72 часа. Вирионы гAAV могут быть собраны из отработанной среды из производственной культуры при условии, что клетки культивируются в условиях, которые вызывают высвобождение вирионов гAAV в среду из интактных клеток (см., например, патент США № 6566118). Подходящие способы лизирования клеток также известны в настоящей области техники и включают в себя, например, несколько циклов замораживания/оттаивания, обработку ультразвуком, микрофлюидизацию и обработку химиче-

скими веществами, такими как детергенты и/или протеазы.

После сбора вирионы гAAV могут быть очищены. Используемый в настоящем документе термин "очищенный" включает в себя препарат вирионов гAAV, лишенный по меньшей мере некоторых других компонентов, которые также могут присутствовать там, где вирионы гAAV встречаются в природе или из которых изначально получены. Таким образом, например, очищенные вирионы гAAV могут быть получены с использованием способа выделения для обогащения их из исходной смеси, такой как лизат культуры или супернатант производственной культуры. Обогащение можно измерить множеством способов, например, по доле устойчивых к ДНКазам частиц (DRP) или копий генома (gc), присутствующих в растворе, или по инфекционности, или его можно измерить по отношению к второму, потенциально мешающему веществу, присутствующему в исходной смеси, такому как загрязняющие вещества, включая в себя загрязняющие вещества производственной культуры или технологические загрязняющие вещества, включая в себя вирус-помощник, компоненты среды и т.п.

Согласно определенным вариантам осуществления выход культуры производства гAAV можно осветлить для удаления остатков клеток-хозяев. Согласно некоторым вариантам осуществления выход культуры производства можно осветлить с использованием множества стандартных технологий, таких как центрифугирование или фильтрация через фильтр с размером пор 0,2 мкм или более (например, фильтр из ацетата целлюлозы или серия глубинных фильтров).

Согласно определенным вариантам осуществления выход культуры для производства гAAV дополнительно обрабатывают бензоназой™ для расщепления любой высокомолекулярной ДНК, присутствующей в производственной культуре. Согласно некоторым вариантам осуществления расщепление бензоназой™ проводят в стандартных условиях, например, при конечной концентрации 1-2,5 единиц/мл бензоназы™ при температуре от температуры окружающей среды до 37°C в течение периода от 30 минут до нескольких часов.

Согласно определенным вариантам осуществления вирионы гAAV могут быть выделены или очищены с использованием одной или нескольких из следующих стадий очистки: равновесное центрифугирование; проточная анионообменная фильтрация; тангенциальная проточная фильтрация (TFF) для концентрирования частиц гAAV; захват гAAV с помощью апатитовой хроматографии; тепловая инактивация вируса-помощника; захват гAAV с помощью хроматографии гидрофобного взаимодействия; замена буфера с помощью эксклюзионной хроматографии (SEC); наночистка и захват гAAV с помощью анионообменной хроматографии, катионообменной хроматографии или аффинной хроматографии. Эти стадии можно использовать по отдельности, в различных комбинациях или в разном порядке. Способы очистки частиц гAAV можно найти, например, в Xiao et al., (1998) *Journal of Virology* 72:2224-2232; патентах США № 6989264 и 8137948 и WO 2010/148143.

Согласно определенным вариантам осуществления очищенные вирионы AAV можно подвергнуть диализу против PBS, отфильтровать и хранить при температуре -80°C. Титры вирусных геномов можно определить с помощью количественной ПЦР с использованием линейаризованной плазмидной ДНК в качестве стандартной кривой (см., например, Lock M, et al., *Hum. Gene Ther.* 2010; 21:1273-1285).

Фармацевтические композиции.

Согласно определенным вариантам осуществления в настоящей заявке представлены композиции, содержащие селективный в отношении PV сайт связывания микроРНК и/или последовательность, кодирующую eTF, который избирательно активирует SCN1A, и фармацевтически приемлемый носитель. Согласно другим вариантам осуществления в настоящей заявке представлены вирионы, содержащие селективный в отношении PV сайт связывания микроРНК и/или последовательность, кодирующую eTF, который избирательно активирует SCN1A, и фармацевтически приемлемый носитель. Согласно иллюстративным вариантам осуществления такие композиции подходят для применения в генной терапии. Фармацевтические композиции предпочтительно стерильны и стабильны в условиях производства и хранения. Стерильные растворы могут быть получены, например, путем фильтрации через стерильные фильтрующие мембраны.

Приемлемые носители и вспомогательные вещества в фармацевтических композициях предпочтительно нетоксичны для реципиентов в используемых дозировках и концентрациях. Приемлемые носители и вспомогательные вещества могут включать в себя такие буферы, как фосфатный, цитратный, HEPES и TAE, такие антиоксиданты, как аскорбиновая кислота и метионин, такие консерванты, как гексаметоний хлорид, октадецилдиметилбензиламмонийхлорид, резорцин и хлорид бензалкония, такие белки, как сывороточный альбумин человека, желатин, декстран и иммуноглобулины, такие гидрофильные полимеры, как поливинилпирролидон, такие аминокислоты, как глицин, глутамин, гистидин и лизин, и такие углеводы, как глюкоза, манноза, сахароза и сорбитол. Фармацевтические композиции по настоящему раскрытию можно вводить парентерально в форме инъекционной композиции. Фармацевтические композиции для инъекций могут быть составлены с использованием стерильного раствора или любой фармацевтически приемлемой жидкости в качестве носителя. Фармацевтически приемлемые носители включают в себя, без ограничения, стерильную воду и физиологический раствор.

Фармацевтические композиции по настоящему раскрытию могут быть приготовлены в микрокап-

сулах, таких как гидроксиметилцеллюлозные или желатиновые микрокапсулы и микрокапсулы из полиметилметакрилата. Фармацевтические композиции по настоящему раскрытию также могут быть приготовлены в других системах доставки лекарственных средств, таких как липосомы, микросферы альбумина, микроэмульсии, наночастицы и микрокапсулы. Фармацевтическая композиция для генной терапии может быть в приемлемом разбавителе или может содержать матрицу с медленным высвобождением, в которую встроен носитель для доставки гена.

Представленные в настоящем документе фармацевтические композиции могут быть составлены для парентерального введения, подкожного введения, внутривенного введения, внутримышечного введения, внутриартериального введения, интрапаренхимального введения, интратекального введения, введения в мостомозжечковую цистерну, интрацеребровентрикулярного введения или внутрибрюшинного введения. Фармацевтическая композиция также может быть приготовлена или вводиться путем назального введения, посредством спрея, перорального, аэрозольного, ректального или вагинального введения. Согласно одному варианту осуществления представленная в настоящем документе фармацевтическая композиция вводится в ЦНС или спинномозговую жидкость (CSF), т.е. путем интрапаренхимальной инъекции, интратекальной инъекции, инъекции в мостомозжечковую цистерну или интрацеребровентрикулярной инъекции. Тканевая мишень может быть специфической, например, ЦНС, или может представлять собой комбинацию нескольких тканей, например, мышечных тканей и тканей ЦНС. Иллюстративная ткань или другие мишени могут включать в себя печень, скелетные мышцы, сердечную мышцу, жировые отложения, почки, легкие, эндотелий сосудов, эпителиальные клетки, кроветворные клетки, ЦНС и/или CSF. Согласно предпочтительному варианту осуществления представленная в настоящем документе фармацевтическая композиция, содержащая селективный в отношении PV сайт связывания микроРНК и/или eTF, который избирательно активирует SCN1A, вводится инъекцией в ЦНС или спинномозговую жидкость, т.е. путем интрапаренхимальной инъекции, интратекальной инъекции, инъекции в мостомозжечковую цистерну или интрацеребровентрикулярной инъекции. Один или несколько из этих способов можно использовать для введения фармацевтической композиции по настоящему раскрытию.

Согласно определенным вариантам осуществления представленная в настоящем документе фармацевтическая композиция содержит "эффективное количество" или "терапевтически эффективное количество". В контексте настоящего описания такие количества относятся к количеству, эффективному в дозировках и в течение периодов времени, необходимых для достижения желаемого терапевтического результата, такого как повышение уровня экспрессии SCN1A и/или уменьшение частоты и/или продолжительности припадков.

Дозировка фармацевтических композиций по настоящему раскрытию зависит от факторов, включающих в себя путь введения, заболевание, которое необходимо лечить, и физические характеристики (например, возраст, массу, общее состояние здоровья) субъекта. Дозировка может быть скорректирована для обеспечения оптимального терапевтического ответа. Как правило, дозировка может представлять собой количество, которое эффективно лечит заболевание, не вызывая значительной токсичности. Согласно одному варианту осуществления представленный в настоящем документе вектор AAV можно вводить пациенту для лечения дефицита SCN1A (включая в себя, например, синдром Драве) в количестве или дозе в диапазоне от 5×10^{11} до 1×10^{14} гк/кг (количество геномных копий на килограмм массы тела пациента (гк/кг)). Согласно более конкретному варианту осуществления вектор AAV вводят в количестве, которое находится в диапазоне от приблизительно 5×10^{11} гк/кг до приблизительно 3×10^{13} гк/кг или от приблизительно 1×10^{12} до приблизительно 1×10^{14} гк/кг, или от приблизительно 1×10^{12} до приблизительно 1×10^{13} гк/кг, или приблизительно 5×10^{11} гк/кг, 1×10^{12} гк/кг, $1,5 \times 10^{12}$ гк/кг, $2,0 \times 10^{12}$ гк/кг, $2,5 \times 10^{12}$ гк/кг, 3×10^{12} гк/кг, $3,5 \times 10^{12}$ гк/кг, 4×10^{12} гк/кг, $4,5 \times 10^{12}$ гк/кг, 5×10^{12} гк/кг, $5,5 \times 10^{12}$ гк/кг, 6×10^{12} гк/кг, $6,5 \times 10^{12}$ гк/кг, 7×10^{12} гк/кг, $7,5 \times 10^{12}$ гк/кг, 8×10^{12} гк/кг, $8,5 \times 10^{12}$ гк/кг, 9×10^{12} гк/кг или $9,5 \times 10^{12}$ гк/кг. гк/кг можно определить, например, с помощью кПЦР или цифровой капельной ПЦР (ddPCR) (см., например, M. Lock et al, Hum Gene Ther Methods. 2014 Apr;25(2): 115-25). Согласно другому варианту осуществления представленный в настоящем документе вектор AAV можно вводить пациенту для лечения дефицита SCN1A (включая в себя, например, синдром Драве) в количестве или дозе в диапазоне от 1×10^9 до 1×10^{11} iu/кг (инфекционные единицы вектора (iu)/масса тела субъекта или пациента (кг)). Согласно определенным вариантам осуществления фармацевтическая композиция может быть приготовлена в виде стандартной дозы по мере необходимости. Такие единичные дозированные единицы могут содержать от приблизительно 1×10^9 гс до приблизительно 1×10^{15} гс.

Фармацевтические композиции по настоящему раскрытию можно вводить нуждающемуся в этом субъекту, например, один или несколько раз (например, 1-10 раз или более) ежедневно, еженедельно, ежемесячно, два раза в год, ежегодно или по медицинским показаниям. Согласно иллюстративному варианту осуществления достаточно однократного введения. Согласно одному варианту осуществления фармацевтическая композиция, содержащая кассету экспрессии, кодирующую селективный в отношении PV сайт связывания микроРНК и/или eTF, который избирательно активирует SCN1A, подходит для применения у людей и вводится путем интрапаренхимальной инъекции, интратекальной инъекции, инъекции в мостомозжечковую цистерну или интрацеребровентрикулярной инъекции. Согласно одному вари-

анту осуществления фармацевтическая композиция доставляется через периферическую вену путем болюсной инъекции. Согласно другим вариантам осуществления фармацевтическая композиция доставляется через периферическую вену путем инфузии в течение приблизительно 10 минут (± 5 минут), в течение приблизительно 20 минут (± 5 минут), в течение приблизительно 30 минут (± 5 минут), в течение приблизительно 60 минут (± 5 минут) или в течение более чем 90 минут (± 10 минут).

Согласно другому аспекту в настоящей заявке дополнительно представлен набор, содержащий молекулу нуклеиновой кислоты, вектор, клетку-хозяина, вирион или фармацевтическую композицию, как описано в настоящем документе, в одном или нескольких контейнерах. Набор может содержать инструкции или упаковочные материалы, которые описывают, как вводить пациенту молекулу нуклеиновой кислоты, вектор, клетку-хозяина или вирион, содержащиеся в наборе. Контейнеры набора могут быть из любого подходящего материала, например, стекла, пластика, металла и т.д., и быть любого подходящего размера, формы или конфигурации. Согласно определенным вариантам осуществления наборы могут содержать одну или несколько ампул или шприцев, которые содержат молекулу нуклеиновой кислоты, вектор, клетку-хозяина, вирион или фармацевтическую композицию в подходящей жидкости или форме раствора.

Способы лечения.

Согласно одному аспекту в настоящей заявке представлены способы применения eTF, которые избирательно активируют SCN1A, как раскрыто в настоящем документе. Согласно определенным вариантам осуществления в настоящей заявке представлены способы введения кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, как раскрыто в настоящем документе, для усиления экспрессии SCN1A в клетке. Согласно различным вариантам осуществления eTF, который избирательно активирует SCN1A, как раскрыто в настоящем документе, можно использовать для модуляции экспрессии SCN1A в клетке *in vitro*, *in vivo* или *ex vivo*.

Согласно определенным вариантам осуществления в настоящей заявке представлены способы лечения заболевания или нарушения, связанного с SCN1A, путем введения субъекту кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, как раскрыто в настоящем документе, при необходимости. Согласно определенным вариантам осуществления нарушение представляет собой нарушение центральной нервной системы. Согласно иллюстративным вариантам осуществления заболевание или нарушение связано с гаплонедостаточностью SCN1A. Согласно определенным вариантам осуществления нарушение представляет собой эпилепсию, связанную с гаплонедостаточностью SCN1A. Согласно определенным вариантам осуществления гаплонедостаточность является результатом того, что субъект является гетерозиготным по мутации потери функции гена SCN1A. Согласно определенным вариантам осуществления нарушение представляет собой эпилепсию, связанную со вставкой, делецией или заменой в гене SCN1A. Согласно определенным вариантам осуществления нарушение представляет собой эпилепсию, связанную с точечной мутацией в гене SCN1A. Согласно определенным вариантам осуществления способ лечения заболевания или нарушения предусматривает введение кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, как раскрыто в настоящем документе, так что недостаточная экспрессия SCN1A корректируется, доводится до пределов уровня здорового человека или доводится до нормального диапазона, определенного стандартом медицинской помощи. Согласно определенным вариантам осуществления раскрытые в настоящем документе способы применяются для лечения заболевания или нарушения, связанного с эндогенным SCN1A, содержащим одну или несколько мутаций, которые приводят к аномальной экспрессии SCN1A.

Согласно определенным вариантам осуществления в настоящей заявке представлены способы облегчения симптома, связанного с заболеванием или нарушением, путем введения кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, как раскрыто в настоящем документе, нуждающемуся в этом субъекту.

Согласно иллюстративному варианту осуществления в настоящей заявке представлены способы лечения заболевания, нарушения или симптома, связанного с мутацией в SCN1A (например, точечной мутацией, заменой, делецией, инверсией и т.д.), дефицитом в Nav1.1 и/или снижением активности Nav1.1 путем введения нуждающемуся в этом субъекту кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно усиливает экспрессию гена SCN1A или его белкового продукта Nav1.1. Управляемые напряжением ионные натриевые каналы важны для генерации и распространения потенциалов действия в поперечнополосатых мышцах и нервных тканях. Управляемые напряжением ионные натриевые каналы представляют собой гетеромерные комплексы, состоящие из большой центральной порообразующей гликозилированной альфа-субъединицы и двух меньших вспомогательных бета-субъединиц. Большая альфа-субъединица субъединица Nav1.1, кодируемая геном SCN1A, актуальна для различных заболеваний или нарушений, таких как синдром Драве. Nav1.1 экспрессируется в нейронах и может быть собрана с различными бета-

субъединицами, включая в себя Nav β 1, экспрессируемую геном SCN1B.

Согласно определенным вариантам осуществления в настоящей заявке представлены способы лечения заболеваний, связанных с мутацией в SCN1A (например, делецией, вставкой, инверсией, точечной мутацией (например, нонсенс-мутацией, миссенс-мутацией) и т.д.) или сниженной активностью Nav1.1 с использованием eTF, который избирательно повышает экспрессию эндогенного гена SCN1A. Заболевания и нарушения, связанные с мутациями SCN1A, включают в себя, без ограничения: синдром Драве, синдром Охтаха, эпилепсию, раннюю детскую эпилептическую энцефалопатию 6 (EIEE6), семейные фебрильные припадки 3A (FEB3A), трудноизлечимую детскую эпилепсию с генерализованными тонико-клоническими припадками (ICEGTC), мигрень, семейную гемиплегию 3 (FHM3), синдром Панайотопулоса, семейную фибрилляцию предсердий 13 (ATFB13), генерализованную эпилепсию с фебрильными судорогами плюс типа 1 (gefs + тип 1), синдром Бругада, неспецифический дефект сердечной проводимости, генерализованную эпилепсию плюс, доброкачественный ранний младенческий эпилептический синдром, раннюю детскую эпилептическую энцефалопатию (EIEE11), доброкачественную семейную детскую эпилепсию, нейродегенерацию, таупатии и болезнь Альцгеймера. В некоторых случаях неврологическим заболеванием является синдром Драве. Мутации или аномалии в SCN1A также были связаны с судорожными расстройствами, эпилепсией, аутизмом, семейной гемиплегической мигренью 3 типа (FHM3), генетической эпилепсией с фебрильными припадками плюс (GEFS+) и эффективностью некоторых противосудорожных лекарственных средств. Например, мутация ICS5N+5G>A в SCN1A связана с максимально безопасным количеством (дозой) противосудорожных лекарственных средств фенитоина и карбамазепина.

Согласно определенным вариантам осуществления в настоящей заявке представлен способ лечения субъекта с синдромом Драве или с риском его развития путем введения кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A. Синдром Драве характеризуется продолжительными фебрильными и нефебрильными припадками в течение первого года жизни ребенка. Это заболевание прогрессирует до других типов припадков, таких как миоклонические и парциальные припадки, задержка психомоторного развития и атаксия. Оно характеризуется когнитивными нарушениями, поведенческими нарушениями и двигательными нарушениями. Поведенческие нарушения часто включают в себя гиперактивность и импульсивность, а в более редких случаях - аутистическое поведение. Синдром Драве также связан с нарушениями сна, включая в себя сонливость и бессонницу. У многих пациентов синдром Драве вызван генетическими мутациями, которые приводят к выработке нефункциональных белков. При лечении заболеваний, связанных с генетическими причинами, существует множество проблем. Таким образом, большинство существующих способов лечения было направлено на профилактическое лечение припадков и других симптомов.

У 70-90% пациентов синдром Драве вызывается нонсенс-мутациями в гене SCN1A, что приводит к преждевременному стоп-кодону и, следовательно, к нефункциональному белку. Как правило, миссенс-мутация в сегменте S5 или S6 поры натриевого канала приводит к потере функции канала и развитию синдрома Драве. Гетерозиготное наследование мутации SCN1A, например, нонсенс-мутация, миссенс-мутация, делеция, вставка, инверсия и т.д. - это все, что необходимо для развития дефектного натриевого канала; у пациентов с синдромом Драве останется одна нормальная копия гена. Таким образом, заболевание характеризуется как гаплонедостаточность, и, таким образом, увеличение экспрессии функционирующей копии SCN1A может восстанавливать нормальные уровни продукции Nav1.1.

Симптомы, связанные с синдромом Драве, включают в себя судороги, дефекты памяти, задержку развития, низкий мышечный тонус и/или когнитивные проблемы. Лечение кассетой экспрессии, вектором экспрессии или вириальной частицей, описанными в настоящем документе, может привести к улучшению одного или нескольких симптомов, таких как уменьшение количества, продолжительности и/или интенсивности припадков. Введение генной терапии, как описано в настоящем документе, субъекту с риском развития синдрома Драве может предотвратить развитие или замедлить прогрессирование одного или нескольких симптомов Драве.

Согласно определенным вариантам осуществления лечение кассетой экспрессии, вектором экспрессии или вириальной частицей, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, как описано в настоящем документе, снижает продолжительность и/или частоту припадков, например, припадков, связанных с синдромом Драве, по меньшей мере на 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% 95% или более по сравнению с не подвергнутым лечению контролем или по сравнению с уровнем до лечения.

У некоторых пациентов с болезнью Альцгеймера выработка амилоида β (A β) с участием многих пептидов и протеаз может влиять на возбудимость нейронов, вызывая припадки и подавляя регуляцию натриевого канала Nav1.1 в нейронах PV. Согласно другому варианту осуществления в настоящей заявке представлены способы лечения субъекта, страдающего болезнью Альцгеймера, путем введения описан-

ной в настоящем документе кассеты экспрессии, вектора экспрессии или вирусной частицы, которая содержит полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A. Симптомы, связанные с болезнью Альцгеймера, включают в себя кратковременную потерю памяти, когнитивные трудности, припадки и трудности с речью, исполнительными функциями, восприятием (агнозия) и выполнением движений (апраксия). Лечение кассетой экспрессии, вектором экспрессии или вирусной частицей, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, может привести к улучшению одного или нескольких симптомов болезни Альцгеймера, таких как снижение прогрессирования потери памяти или предотвращение одного или нескольких симптомов. В некоторых случаях лечение может привести к коррекции активности головного мозга с высокой мощностью гамма-излучения. Лечение может привести к снижению частоты припадков и/или тяжести припадков или снижению активности высокой мощности гамма-излучения по меньшей мере на 10%, 20%, 30%, 40%, 50%, 60%, 70% или более, как по сравнению с отсутствием лечения. В некоторых случаях лечение может привести к улучшению когнитивных функций. Обучение и/или память могут быть улучшены по меньшей мере на 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% или более чем на 100% по сравнению с отсутствием лечения или перед лечением полинуклеотидом, кодирующим eTF, который избирательно активирует SCN1A, как раскрыто в настоящем документе.

В некоторых случаях лечение кассетой экспрессии, вектором экспрессии или вирусной частицей, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, снижает активность высокой мощности гамма-излучения (например, активность высокой мощности гамма-излучения, связанную с болезнью Альцгеймера) по меньшей мере на 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% или 95% по сравнению с не подвергнутым лечению контролем или по сравнению с уровнем до лечения.

Паркинсонизм относится к совокупности признаков и симптомов, обнаруживаемых при болезни Паркинсона (PD), включая в себя медлительность (брадикинезию), тугоподвижность (ригидность), тремор и дисбаланс (постуральную нестабильность). В некоторых случаях введение кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует SCN1A, как описано в настоящем документе, субъекту с риском развития или страдающим болезнью Паркинсона может предотвратить развитие одного или нескольких симптомов или замедлить прогрессирование болезни Паркинсона по меньшей мере на 10%, по меньшей мере на 20%, по меньшей мере на 30%, по меньшей мере на 40%, по меньшей мере на 50%, по меньшей мере на 60%, по меньшей мере на 70%, по меньшей мере на 80% или по меньшей мере на 90% по сравнению с отсутствием лечения.

Согласно определенным вариантам осуществления в настоящей заявке представлены способы, которые можно применять для лечения субъекта, который находится в группе риска развития заболевания. Может быть известно, что субъект предрасположен к заболеванию, например, неврологическому заболеванию или заболеванию, связанному с эпилепсией, припадками и/или энцефалопатией. Субъект может быть предрасположен к заболеванию из-за генетического события или из-за известных факторов риска. Например, субъект может нести мутацию в SCN1A, которая связана с эпилепсией (такой как, например, синдром Драве). Любая мутация в гене SCN1A, которая снижает его активность (за счет снижения уровня экспрессии, нарушения функции белка или комбинации обоих), может предрасполагать субъекта к заболеванию, включая в себя любое одно или несколько из вставок, делеций, инверсий, транслокаций или замен (например, точечные мутации, включая в себя нонсенс-мутации и/или миссенс-мутации) в гене SCN1A. В некоторых случаях субъект может быть предрасположен к такому заболеванию, как болезнь Альцгеймера, из-за возраста субъекта. В некоторых случаях у субъекта может быть недостаточное количество белка SCN1A, и лечение заболевания, связанного с SCN1A, предусматривает введение кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует эндогенный SCN1A, как описано в настоящем документе.

Согласно определенным вариантам осуществления лечение с использованием кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей полинуклеотид, кодирующий eTF, который избирательно активирует эндогенный SCN1A, представленный в настоящем документе, может привести к уменьшению или прекращению симптомов, связанных с синдромом Драве или другим заболеванием, связанным с SCN1A, или нарушением, например, эпилепсией, связанной с гаплонедостаточностью SCN1A. Например, лечение может улучшить обучение, память, когнитивные и/или двигательные функции; уменьшить частоту и/или продолжительность припадков; и/или снизить температурную чувствительность (или увеличить температурный порог для запуска припадка).

Согласно другому аспекту в настоящей заявке представлены способы селективной экспрессии трансгена в PV-содержащих нейронах путем введения кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей по меньшей мере один селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления в настоящей заявке представлены способы селективной экспрессии трансгена в PV-содержащих нейронах приматов путем введения кассеты экс-

прессии, вектора экспрессии или вирусной частицы, содержащей трансген и по меньшей мере один селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления в настоящей заявке представлены способы селективной экспрессии трансгена в PV-содержащих нейронах путем введения кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей селективный в отношении PV регуляторный элемент, функционально связанный с трансгеном, и по меньшей мере один селективный в отношении PV сайт связывания микроРНК. Согласно иллюстративным вариантам осуществления трансген содержит последовательность, кодирующую любой из eTF, который избирательно активирует SCN1A, как описано в настоящем документе.

Согласно определенным вариантам осуществления в настоящей заявке представлен способ генной терапии, предусматривающий введение субъекту кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей трансген и по меньшей мере один селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления в заявке представлены способы генной терапии, предусматривающие введение субъекту кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей селективный в отношении PV регуляторный элемент, функционально связанный с трансгеном, и по меньшей мере один селективный в отношении PV сайт связывания микроРНК. Согласно иллюстративным вариантам осуществления трансген содержит последовательность, кодирующую любой из eTF, который избирательно активирует SCN1A, как описано в настоящем документе.

Согласно определенным вариантам осуществления в настоящей заявке представлен способ лечения заболевания или нарушения, предусматривающий введение кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей трансген и по меньшей мере один селективный в отношении PV сайт связывания микроРНК. Согласно определенным вариантам осуществления в заявке представлены способы лечения заболевания или нарушения, предусматривающие введение субъекту кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей селективный в отношении PV регуляторный элемент, функционально связанный с трансгеном, и по меньшей мере один селективный в отношении PV сайт связывания микроРНК. Согласно иллюстративным вариантам осуществления трансген содержит последовательность, кодирующую любой из eTF, который избирательно активирует SCN1A, как описано в настоящем документе. Согласно определенным вариантам осуществления кассета экспрессии, вектор экспрессии или вирусная частица, содержащая трансген и селективный в отношении PV сайт связывания микроРНК, и, необязательно, селективный в отношении PV регуляторный элемент, могут быть использованы для лечения заболевания или нарушения, в которое вовлечены PV-содержащие нейроны. Согласно определенным вариантам осуществления кассета экспрессии, вектор экспрессии или вирусная частица, содержащая трансген и селективный в отношении PV сайт связывания микроРНК и, возможно, селективный в отношении PV регуляторный элемент, используются для лечения состояния нейронов. Нейрональные заболевания или нарушения, подходящие для лечения, включают в себя, без ограничения, синдром Драве, болезнь Альцгеймера, болезнь Паркинсона, болезнь Хантингтона, боковой амиотрофический склероз (ALS), спинальную мышечную атрофию (SMA), эпилепсию, нейродегенеративные нарушения, двигательные нарушения, нарушения настроения, болезни двигательных нейронов, прогрессирующую мышечную атрофию (PMA), прогрессирующий бульбарный паралич, псевдобульбарный паралич, первичный боковой склероз, неврологические последствия СПИДа, нарушения развития, рассеянный склероз, нейрогенетические нарушения, инсульт, травму спинного мозга, черепно-мозговую травму, таупатию, гиповозбудимость нейронов и/или припадки. Согласно некоторым вариантам осуществления вирусный вектор, вирусная частица или фармацевтическая композиция, содержащая трансген и селективный в отношении PV сайт связывания микроРНК и необязательно селективный в отношении PV регуляторный элемент, используются для лечения психического расстройства (например, шизофрении, обсессивно-компульсивного расстройства, наркомании, депрессии, тревоги, психоза); расстройства аутистического спектра (например, синдрома ломкой X-хромосомы, синдрома Ретта); эпилепсии (например, синдрома Драве, хронической травматической энцефалопатии, генерализованной эпилепсии с фебрильными припадками плюс (GEFS+), эпилептической энцефалопатии, височной эпилепсии, фокальной эпилепсии, туберозного склероза, эпилепсии, связанной с гаплонедостаточностью SCN1A); и/или нейродегенерации (например, болезни Альцгеймера, болезни Паркинсона). Заболевания, связанные с дисфункциональными нейронами PV, например, вызванные мутациями потери функции в SCN1A или Nav1.1, включают в себя: синдром Драве, синдром Охтаха, эпилепсию, раннюю детскую эпилептическую энцефалопатию 6 (EIEE6), семейные фебрильные припадки 3A (FEB3A), трудноизлечимую детскую эпилепсию с генерализованными тонико-клоническими припадками (ICEGTC), мигрень, семейную гемиплегию 3 (FHM3), синдром Панайотопулоса, семейную фибрилляцию предсердий 13 (ATFB13), генерализованную эпилепсию с фебрильными припадками плюс типа 1 (gefs + тип 1), неспецифический кардиальный синдром, неспецифический синдром Бругада дефект проводимости, генерализованную эпилепсию с фебрильными припадками плюс, доброкачественный ранний младенческий эпилептический синдром, раннюю детскую эпилептическую энцефалопатию 11 (EIEE11), доброкачественную семейную детскую эпилепсию, нейродегенерацию, таупатию и болезнь Альцгеймера.

Согласно определенным вариантам осуществления лечения с использованием кассеты экспрессии, вектора экспрессии или вирусной частицы, содержащей трансген и селективный в отношении PV сайт

связывания микроРНК, и, возможно, селективный в отношении PV регуляторный элемент, описанный в настоящем документе, приводит к улучшению симптомов, связанных с заболеванием или нарушением нейронов. Например, за пациентом с болезнью Паркинсона можно наблюдать симптоматически на предмет улучшения двигательных функций, что указывает на положительный ответ на лечение. Введение терапии с использованием описанного в настоящем документе способа субъекту с риском развития нейронального нарушения может предотвратить развитие или замедлить прогрессирование одного или нескольких симптомов.

Согласно определенным вариантам осуществления касета экспрессии, вектор экспрессии или вирусная частица, содержащая трансген и селективный в отношении PV сайт связывания микроРНК, и, необязательно, селективный в отношении PV регуляторный элемент, представленные в настоящем документе, могут быть использованы для лечения субъекта, у которого было диагностировано заболевание нейронов, например, эпилепсия, связанная с гаплонедостаточностью SCN1A, такая как, например, синдром Драве. Согласно различным вариантам осуществления любое из нейрональных заболеваний или нарушений, описанных в настоящем документе, вызвано известным генетическим событием (например, любой из мутаций SCN1A, известных в настоящей области техники) или имеет неизвестную причину.

Согласно определенным вариантам осуществления касета экспрессии, вектор экспрессии или вирусная частица, содержащая трансген и селективный в отношении PV сайт связывания микроРНК, и, возможно, селективный в отношении PV регуляторный элемент, представленные в настоящем документе, могут быть использованы для лечения субъекта, который находится в группе риска развития заболевания или нарушения. Согласно определенным вариантам осуществления может быть известно, что субъект предрасположен к заболеванию, например, нейрональному заболеванию (например, эпилепсии, связанной с гаплонедостаточностью SCN1A, такой как, например, синдром Драве). Согласно определенным вариантам осуществления субъект может быть предрасположен к заболеванию из-за генетического события или из-за известных факторов риска. Например, субъект может нести мутацию в SCN1A, которая связана с эпилепсией или синдромом Драве, например, вставку, делецию, инверсию, транслокацию или замену (например, точечную мутацию, включая в себя нонсенс-мутацию и/или миссенс-мутацию).

Согласно определенным вариантам осуществления касета экспрессии, вектор экспрессии или вирусная частица, содержащая трансген и селективный в отношении PV сайт связывания микроРНК, и, возможно, селективный регуляторный элемент PV, представленные в настоящем документе, могут быть использованы для уменьшения одного или нескольких симптомов, связанных с заболеванием или нарушением. Например, симптомы, связанные с синдромом Драве, включают в себя судороги, дефекты памяти, задержку развития, плохой мышечный тонус и/или когнитивные проблемы. Лечение вирусным вектором, вирусной частицей или фармацевтической композицией, содержащей трансген и селективный в отношении PV сайт связывания микроРНК и, необязательно, селективный в отношении PV регуляторный элемент, представленные в настоящем документе, может привести к улучшению одного или нескольких симптомов, таких как уменьшение количества, продолжительности и/или интенсивности припадков.

Согласно определенным вариантам осуществления описанные в настоящем документе способы используются для увеличения экспрессии трансгена в нейроне PV, генной терапии или лечения заболевания или нарушения у приматов. Согласно определенным вариантам осуществления примат представляет собой человека. Согласно определенным вариантам осуществления примат представляет собой отличного от человека примата. Согласно определенным вариантам осуществления отличный от человека примат представляет собой обезьяну старого мира, орангутана, гориллу, шимпанзе, яванского макака, макака-резус или свинохвостого макака.

Термины "субъект" и "индивидуум" используются в настоящем документе взаимозаменяемо для обозначения позвоночного животного, предпочтительно млекопитающего, более предпочтительно человека. Описанные в настоящем документе способы могут быть применимы в терапевтических целях для человека, в ветеринарии и/или в доклинических исследованиях на животных моделях заболевания или состояния. Согласно различным вариантам осуществления субъект, которого можно лечить в соответствии с описанными в настоящем документе способами, представляет собой млекопитающее, такое как, например, мышь, крыса, хомяк, морская свинка, песчанка, корова, овца, свинья, коза, осел, лошадь, собака, кошка, лама, обезьяна (например, обезьяна старого мира, мартышка или макак, такой как макак-резус, свинохвостый макак или яванский макак (например, макак-крабед)), обезьяна (например, орангутан, горилла или шимпанзе) или человек. Согласно иллюстративному варианту осуществления субъект представляет собой человека.

В следующих таблицах представлены раскрытые в настоящем документе последовательности.

Таблица 1. Раскрытые в настоящем документе иллюстративные сконструированные факторы транскрипции. Последовательности регуляторных элементов (RE) раскрыты ниже в табл. 2 и 8. Для RE, когда указан m1, это означает, что сайт связывания микроРНК m1 (SEQ ID NO: 7, табл. 8) включен между кодирующей областью и полиА-хвостом. Последовательности ДНК-связывающих доменов (DBD) раскрыты ниже в табл. 3. Для DBD сконструированный "цинковый палец" (eZF) указывает на то, что конструкция имеет формулу, изложенную в SEQ ID NO: 147 (табл. 10); EGR1 указывает на то, что DBD происхо-

дит от человеческого EGR1 дикого типа (SEQ ID NO: 176; табл. 12); и EGR3 указывает на то, что DBD происходит от человеческого EGR3 дикого типа (SEQ ID NO: 175, табл. 12). Последовательности для целевых сайтов (например, последовательности, связанные с DBD) представлены в табл. 4 ниже. Последовательности для доменов активации транскрипции (TAD) раскрыты ниже в табл. 5. Для TAD (c) указывает на то, что TAD расположен на с-конце DBD, (n) указывает на то, что TAD расположен на n-конце DBD, (n/c) означает, что существует TAD, расположенный как на n-конце, так и на с-конце DBD, а 2x CITED4 (n) указывает на то, что есть 2 копии TAD CITED4, расположенные на n-конце DBD. Последовательности сконструированных факторов транскрипции полной длины (DBD + TAD) представлены в табл. 6 ниже.

Таблица 1

Конструкция	RE	DBD	Целевой сайт	TAD (положение)	SEQ ID NO (DBD)	SEQ ID NO (DBD + TAD)
1	RE 1	eZF	Z13	VPR (c)	89	99
2	RE 1	eZF	Z1	VPR (c)	77	100
3	RE 1	eZF	Z13	VP64 (c)	89	101
4	RE 1	eZF	Z1	VP64 (c)	77	102
5	RE 1	EGR1	Z13	CITED4 (n/c)	93	103
6	RE 1	EGR1	Z13	CITED4 (n)	93	104
7	RE 1	EGR1	Z1	CITED4 (n/c)	92	105
8	RE 1	EGR1	Z1	CITED4 (n)	92	106
9	RE 1	EGR1	Z1	CITED4 (c)	92	107
10	RE 1	EGR1	Z1	CITED2 (c)	92	108
11	RE 1	EGR1	Z1	CITED2 (n)	92	109
12	RE 1	EGR3	Z1	CITED4 (n/c)	96	110
13	RE 1	EGR3	Z1	CITED4 (c)	96	111
14	RE 1	EGR3	Z1	CITED2 (c)	96	112
15	RE 1	EGR3	Z1	CITED2 (n)	96	113
16	CBA	EGR3	Z15	N/A	98	114
17	RE 1	EGR1	Z13	N/A	93	115
18	RE 1	EGR1	Z15	N/A	94	116
19	RE 1	EGR1	Z13	N/A	93	117
20	RE 1	EGR1	Z13	N/A	93	115
21	RE 1	EGR1	Z1	N/A	92	118
22	CBA	EGR3	Z13	N/A	97	119
23	RE 1	EGR1	Z17	N/A	95	120
24	RE 1	EGR1	Z13	N/A	93	121
25	CBA	EGR3	Z1	N/A	96	122
26	RE 1	EGR1	Z1	N/A	92	123
27	RE 1	EGR1	Z1	N/A	92	124
28	RE 1	eZF	Z8	VP64 (c)	84	125
29	RE 1	eZF	Z14	VP64 (c)	90	126
30	CBA	eZF	Z13	VPR (c)	89	99
31	RE 2 (m1)	eZF	Z1	VP64 (c)	77	102
32	RE 2	eZF	Z1	VP64 (c)	77	102
33	RE 2	eZF	Z1	VPR (c)	77	100
34	RE 2	eZF	Z1	VP64 (c)	77	127
35	RE 2 (m1)	eZF	Z1	VP64 (c)	77	127
36	RE 2	EGR1	Z1	CITED4 (n)	92	128
37	RE 2 (m1)	EGR1	Z1	CITED4 (n)	92	106
38	RE 2 (m1)	EGR1	Z1	CITED4 (n)	92	129
39	RE 2 (m1)	EGR1	Z1	CITED4 (n/c)	92	105
40	RE 2 (m1)	EGR1	Z1	2x CITED4 (n)	92	130

41	RE 2 (m1)	EGR1	Z1	2x CITED4 (n)	92	131
42	RE 2 (m2)	eZF	Z1	VP64 (c)	77	102
43	RE 2 (m3)	eZF	Z1	VP64 (c)	77	102
44	RE 2 (m2)	EGR1	Z1	CITED4 (n)	92	106
45	RE 2 (m3)	EGR1	Z1	CITED4 (n)	92	106
46	RE 1	EGR1	Z1	CITED4 (n)	92	205
47	RE 1	EGR1	Z1	2x CITED4 (n)	92	207
48	RE 1	EGR1	Z1	2x CITED4 (n)	92	209
49	CBA	eZF	Z1	CREB3 (n)	77	213
50	CBA	EGR1	Z1	CREB3 (n)	92 (без C-конц. Lys)	217
51	CBA	EGR1	Z13	CREB3 (n)	93 (без C-конц. Lys)	219
52	CBA	eZF	Z1	CREB3 (n); без домена ТМ на (c)	77	221
53	CBA	EGR1	Z1	CREB3 (n); без домена ТМ на (c)	92 (без C-конц. Lys)	223

Таблица 2

Последовательности нуклеиновых кислот для различных регуляторных элементов (RE), раскрытых в настоящем документе

RE	Последовательность	SEQ ID NO
RE 1	GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAG CGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCATT GACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGG GACTTTC AAAATGTCGTAACA ACTCCGCCCCATTGACG CAAATGGGCGGTAGGCGGTACGGTGGGAGGTCTATATA AGCAGAGCTGGTACCGTGTGTATGCTCAGGGGCTGGGAA AGGAGGGGAGGGAGCTCCGGCTCAG	1
RE 2	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatgtaacataat gaagtctgttgacataaagaatcatgatattaatgccatggaatgaaagggcgatcaacac tatggttgaaaaggggaaattgtagagcacagatgtgtcgtgtggcagtgctgtctctag caatactcagagaagagagagaacaatgaaattctgattggccccagtgtagccagatga ggttcagctgccaactttctttcacatctatgaaagtcatttaagcaactaacttttttttt tttttttgagacagagtctgtctgtgccaggacagagtgcagtgtgactcaatctcgct cactgcagcctccactcctaggctcaaacggtcctcctgcatcagcctccaagtagctgga atacaggagtgcccaccatgccagctaattttgtatttttaatagatacgggggttcacat atcaccaggctggtctcgaactcctggcctcaagtgatccacgtcctcgccctcccaagt gctgggattataggcgtcagccactatgcccaaccgaccaacctttttaaataatatttaa aaatggatttcacatatatactagtatttacatttatccacacaaaacggacgggctccgctg aaccagtgaggccccagacgtgcgcataaataaccctgctgtgctgcaccacctggggaga ggggaggaccaggtaaatggagcgcgcatagcaaaagggacgccccgtcctttctc tgcgggtggcactggtagctgtggccaggtgtgtactttgatgggcccaaggctggagc	2

	<p>tcaaggaagcgtcgcagggtcacagatctggggaaacccggggaaaagcactgaggcaa aaccgccgctcgtcctacaatatatgggagggggaggttgagtacgttctggattactcata agaccttttttttctccggcgcaaaaccgtgagctggattataatgccctataaagctcc agaggcgtcagggcacctgcagaggagccccgccgctccgactagctgccccgcga gcaacggcctcgtgattccccgccgatccggtccccgctccccactctgccccgctacc ccggagccgtgcagccgctcctccaatctctcttctcctggcgtcgcgtgcgagaggga actagcgagaacgaggaagcagctggaggtgacccgggcaattacgctgtcagggcc gagccgagcggatcgtggcgctgtgcagaggaaagcgaggagtccccgctcgtctgc gcagagccgaggtggtaagctagcagaccactggactcccagcgccaaccgtggctttt cagccagctcctctcccgcgcttcaaccaacccatcccagcgccggccaccaac ctcccgaatgagtctcctgccccagcagccgaaggcgtactaggaacgtaacctgtta ctttccaggggcgtagtcgacccgctcccagttgctgtcgcgactgcgcgcggggt agagtcaagggtgactgtggttctctgccaagtccgaggagaacgtaaatatgggc cttttccccctcactctgtctaccaaaagtcctatgccccggagcagttagccttttttcc aggaattagccagacacaacaacgggaaccagacaccgaaccagacatccccccccgt gcgccctccccgctcgtgccttctcctcttctctcagagccggatctcaaggggag cctcgtgcccccgctgctcagtcctcctcggtgcaaggacccggagtcctccccgcac agctctcgtctctttgagcctgtttctgcgcccggaccagtcaggactctggacagtagag gccccgggacaccgagctg</p>	
<p>RE 3</p>	<p>TCAACAGGGGGACACTTGGGAAAGAAGGATGGGGACAG AGCCGAGAGGACTGTTACACATTAGAGAAACATCAGTG ACTGTGCCAGCTTTGGGGTAGACTGCACAAAAGCCCTGA GGCAGCACAGGCAGGATCCAGTCTGCTGGTCCCAGGAA GCTAACCGTCTCAGACAGAGCACAAAGCACCGAGACAT GTGCCACAAGGCTTGTGTAGAGAGGTCAGAGGACAGCG TACAGGTCCCAGAGATCAAACCTCAACCTCACCAGGCTTG GCAGCAAGCCTTACCAACCCACCCCCACCCACCACC CTGCACGCGCCCCCTCTCCCTCCCCATGGTCTCCCATGGC TATCTCACTTGGCCCTAAAATGTTAAGGATGACACTGG CTGCTGAGTGGAATGAGACAGCAGAAGTCAACAGTAG ATTTTAGGAAAGCCAGAGAAAAAGGCTTGTGCTGTTTTT AGAAAGCCAAGGGACAAGCTAAGATAGGGCCCAAGTAA TGCTAGTATTTACATTTATCCACACAAAACGGACGGGCC TCCGCTGAACCAGTGAGGCCCCAGACGTGCGCATAAATA ACCCCTGCGTGCTGCACCACCTGGGGAGAGGGGGAGGA CCACGGTAAATGGAGCGAGCGCATAGCAAAGGGACGC GGGGTCTTTTCTCTGCCGGTGGCACTGGGTAGCTGTGG CCAGGTGTGGTACTTTGATGGGGCCAGGGCTGGAGCTC AAGGAAGCGTCGCAGGGTCACAGATCTGGGGGAACCC GGGAAAAGCACTGAGGCAAACCGCCGCTCGTCTCCT ACAAATATATGGGAGGGGGAGGTTGAGTACGTTCTGGATT ACTCATAAGACCTTTTTTTTTTCTTCCGGGCGCAAACC GTGAGCTGGATTTATAATCGCCCTATAAAGCTCCAGAGG CGGTCAGGCACCTGCAGAGGAGCCCCGCCGCTCCGCCGA CTAGCTGCCCCCGCGAGCAACGGCCTCGTGATTTCCCCG CCGATCCGGTCCCCGCCTCCCCACTCTGCCCCCGCCTACC CCGGAGCCGTGCAGCCGCTCTCCGAATCTCTCTTCTC CTGGCGCTCGCGTGCGAGAGGGAAC TAGCGAGAACGAG GAAGCAGCTGGAGGTGACGCCGGCAGATTACGCCTGT</p>	<p>3</p>

	<p>CAGGGCCGAGCCGAGCGGATCGCTGGGCGCTGTGCAGAG GAAAGGCGGGAGTGCCCGCTCGCTGTCGAGAGCCG AGTGGGTAAGCTAGCGACCACCTGGACTTCCCAGCGCC CAACCGTGGCTTTTCAGCCAGGTCCTCTCCTCCCGCGCT TCTCAACCAACCCCATCCCAGCGCCGGCCACCCAACCTC CCGAAATGAGTGCTTCTGCCCCAGCAGCCGAAGGCGCT ACTAGGAACGGTAACCTGTTACTTTTCCAGGGGCCGTAG TCGACCCGCTGCCCGAGTTGCTGTGCGACTGCGCGCGCG GGGCTAGAGTGCAAGGTGACTGTGGTTCTTCTCTGGCCA AGTCCGAGGGAGAACGTAAGATATGGGCCTTTTTCCCC CTCTACCTTGTCTCACCAAAGTCCCTAGTCCCCGGAGC AGTTAGCCTCTTTCTTCCAGGGAATTAGCCAGACACAA CAACGGGAACCAGACACCGAACAGACATGCCCGCCCC GTGCGCCCTCCCCGCTCGCTGCCTTTCCTCCCTTTGTCT CTCCAGAGCCGATCTTCAAGGGGAGCCTCCGTGCCCCC GGCTGCTCAGTCCCTCCGGTGTGCAGGACCCCGGAAGTC CTCCCCGCACAGCTCTCGCTTCTCTTTGCAGCCTGTTTCT GCGCCGGACCAGTCGAGGACTCTGGACAGTAGAGGCC CCGGACGACCGAGCTG</p>	
<p>RE 4</p>	<p>GCCCTTAGGCCACCTGACCAGGTCCCCTCAGTCCCCC CTTCCACACTCCCACACTCAGCCCCCTCCCCCCCCC GACCCCTGCAGGATTATCCTGTCTGTGTTCTGACTCAGC CTGGGAGCCACCTGGGCAGCAGGGGCCAAGGGTGTCT AGAAGGGACCTGGAGTCCACGCTGGGCCAAGCCTGCCCT TTCTCCCTCTGTCTCCGTCCTGCTTGCAGTTCTGCTGA ATGTGGTTATTTCTCTGGCTCCTTTTACAGAGAATGCTGC TGCTAATTTTATGTGGAGCTCTGAGGCAGTGTAATTGGA AGCCAGACACCCTGTCAGCAGTGGGCTCCCGTCTGAGC TGCCATGCTTCTGCTCTCCTCCCGTCCCGGCTCCTCATT TCATGCAGCCACCTGTCCCAGGGAGAGAGGAGTCACCCA GGCCCTCAGTCCGCCCTTAAATAAGAAAGCCTCCGTT GCTCGGCACACATAACCAAGCAGCCGCTGGTGCAATCT</p>	<p>4</p>
<p>СВА (энхансер СМV + промотор бета- актина кур)</p>	<p>CGTTACATAAATTACGGTAAATGGCCCCGCTGGCTGACC GCCAACGACCCCCGCCATTGACGTCAATAATGACGTA TGTTCCTTAGTAACGCCAATAGGGACTTTCCATTGACG TCAATGGGTGGAGTATTTACGGTAAACTGCCACTTGGC AGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTAT TGACGTCAATGACGGTAAATGGCCCCGCTGGCATTATGC CCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTA CATCTACGTATTAGTCATCGCTATTACCATGggtcgaggtgagc cccacgttctgcttactctccccatctccccccctccccaccccaattttgtattattatttttt aattattttgtcagcagcagtgggggcggggggggggggggcgcgcccagcggggcggg gggggcgggggggggggggggggggggggggggggggggggggggggggggggggggggg gggggcgggggggggggggggggggggggggggggggggggggggggggggggggggggg cggcgcgctccgaaagtctctttatggcgagggcgggggggggggggggggggggggggggg gaagcgcgcgggggggggggagtcgctgctgctgcttgccttcgccccgtccccgctccgcgcg cctcgcgcccggccccggctctgactgaccgctgactcccacaggtgagcggggcggga cggcccttctcctcgggctgtaattagccttggttaatgacggctcgtttctttctgtggctg cgtgaaagccttaaggctccgggagggccctttgtcggggggggagcggctcggggggg tgcgtgcgtgtgtgtgcgtggggagcggcggcgggggggggggggggggggggggggggg agcgtcggggcggcgggggggggggggggggggggggggggggggggggggggggggggg</p>	<p>5</p>

	<p>ccggggcgggtccccgcgggtgcgggggggctgcgaggggaacaaggctgcgtgcgg ggtgtgtcgtgggggggtgagcaggggggtgtgggcgcggcggctcgggctgaaccccc cctgcacccccctccccagttgtgtagcacggccccgcttcgggtgcggggctccgtgcg gggcgtggcgcggggctcgcctgcccggcgggggtgcccggcaggtgggggtcccgg gcggggcggggccgctcggccgggggagggctcgggggaggggcggcggcggccccg gagcgcggcggctgtcagggcgcggcgagccgaccattgcctttatgtaatcgtgcg agagggcgcagggactcctttgtccaaatctggcggagccaaaatctgggagggcggc cgcaccccccttagcgggcggggcgaagcgggtgcggcggcggcaggaaggaaatggg cggggagggcctcgtgctcggcggcggcggcggcggcggcggcggcggcggcggcggc gccgagggggacggctcctcggggggacggggcagggcgggggttcggcttctggc gtgtgaccggcggcttagagcctctgtaaccatgttcatgcttcttctttctacagctcct gggcaacgtgctggtgtgtgctgtctcatcatfttggcaagaatt</p>	
EF1 альфа	<p>GAGTAATTCATACAAAAGGACTCGCCCCTGCCTTGGGGA ATCCCAGGGACCGTCGTTAAACTCCCACTAACGTAGAAC CCAGAGATCGCTGCGTTCGCGCCCCCTCACCCGCCCGCT CTCGTACACTGAGGTGGAGAAGAGCATGCGTGAGGCT CCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAG TCCCCGAGAAGTTGGGGGGAGGGTTCGGCAATTGAACC GGTGCCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAG TGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGG GGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTT CTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGT GCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTT ATGGCCCTTGCGTGCCTTGAATTACTTCCACGCCCTGGC TGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTGGAA GTGGGTGGGAGAGTTCGAGGCCTTGCCTTAAGGAGCCC CTTCCGCTCGTGCTTGAGTTGAGGCCTGGCTTGGGCGCT GGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCT GTCTCGTGTTCGATAAGTCTCTAGCCATTTAAAATTT TTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGATAGT CTTGTAAATGCGGGCCAAGATCTGCACACTGGTATTTTCG GTTTTTGGGGCCCGGGCGGCGACGGGGCCCGTGCCTCC CAGCGCACATGTTCCGGCAGGGCGGGCCTGCGAGCGCG GCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCC GGCTGCTCTGGTGCCTGGCCTCGCGCCGCCGTGTATCG CCCC GCCCTGGGCGGCAAGGCTGGCCCGGTTCGGCACCA TTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTG CAGGGAGCTCAAAAATGGAGGACGCGGCGCTCGGGAGAG CGGGCGGGTGAGTACCCACACAAAGGAAAAGGGCCTT TCCGTCTCAGCCGTCGCTTCATGTGACTCCACGGAGTA CCGGGCGCCGTCAGGCACCTCGATTAGTTCTCGAGCTT TTGGAGTACGTGCTCTTAGGTTGGGGGGAGGGGTTTA TGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGA AGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGA ATTTGCCCTTTTGTAGTTTGGATCTTGGTTCAATTCAG CCTCAGACAGTGGTTCAAAGTTTTTTCTTCCATTTACAG TGTCGTGA</p>	6

Таблица 3. Аминокислотные последовательности для иллюстративных ДНК-связывающих доменов (DBD), представленных в настоящем документе. Для DBD сконструированный "цинковый палец" (eZF) указывает на то, что конструкция имеет формулу, изложенную в SEQ ID NO: 147 (табл. 10); EGR1 указывает на то, что DBD происходит от человеческого EGR1 дикого типа (SEQ ID NO: 176; табл. 12); и EGR3 указывает на то, что DBD происходит от человеческого EGR3 дикого типа (SEQ ID NO: 175, табл. 12). Целевые сайты представляют собой последовательности, связанные с DBD, и представлены в табл. 4 ниже.

Таблица 3

DBD/целевой сайт	Последовательность	SEQ ID NO
eZF/z1	LEPGEKPYKCPECGKSFSSRDNLVRHQRTHTEKPYKCPECGKSFSSREDNLHQRTHTEKPYKCPECGKSFSSRDNLVRHQRTHTEKPYKCPECGKSFSSQSGNLTEHQRTHTEKPYKCPECGKSFSSGHLVRHQRTHTEKPYKCPECGKSFSSQNSTLTHQRTHTGKKT	77
eZF/z2	LEPGEKPYKCPECGKSFSTKNSLTHQRTHTEKPYKCPECGKSFSSRADNLTEHQRTHTEKPYKCPECGKSFSSQLAHLRAHQRTHTEKPYKCPECGKSFSTKNSLTHQRTHTEKPYKCPECGKSFSSQAGHLASHQRTHTEKPYKCPECGKSFSTHLDLIRHQRTHTGKKT	78
eZF/z3	LEPGEKPYKCPECGKSFSSQAGHLASHQRTHTEKPYKCPECGKSFSSREDNLHQRTHTEKPYKCPECGKSFSSGHLVRHQRTHTEKPYKCPECGKSFSSQAGHLASHQRTHTEKPYKCPECGKSFSSQAGHLASHQRTHTGKKT	79
eZF/z4	LEPGEKPYKCPECGKSFSTTGNLTVHQRTHTEKPYKCPECGKSFSSGELVRHQRTHTEKPYKCPECGKSFSSREDNLHQRTHTEKPYKCPECGKSFSSGHLVRHQRTHTEKPYKCPECGKSFSSQAGHLASHQRTHTEKPYKCPECGKSFSSQAGHLASHQRTHTGKKT	80
eZF/z5	LEPGEKPYKCPECGKSFSSRRTCRASHQRTHTEKPYKCPECGKSFSSSTGALTEHQRTHTEKPYKCPECGKSFSSRDNLVRHQRTHTEKPYKCPECGKSFSSRDALTEHQRTHTEKPYKCPECGKSFSSGDLRRHQRTHTEKPYKCPECGKSFSSSHLTHQRTHTGKKT	81
eZF/z6	LEPGEKPYKCPECGKSFSSRDKNLKNHQRTHTEKPYKCPECGKSFSSDPGALVRHQRTHTEKPYKCPECGKSFSSREDNLHQRTHTEKPYKCPECGKSFSSDPGALVRHQRTHTEKPYKCPECGKSFSSGELVRHQRTHTEKPYKCPECGKSFSSRDKNLKNHQRTHTGKKT	82
eZF/z7	LEPGEKPYKCPECGKSFSSKALTEHQRTHTEKPYKCPECGKSFSSPADLTHQRTHTEKPYKCPECGKSFSSRDNLVRHQRTHTEKPYKCPECGKSFSSREDNLHQRTHTEKPYKCPECGKSFSSRDNLVRHQRTHTEKPYKCPECGKSFSSQSGNLTEHQRTHTEKPYKCPECGKSFSSGHLVRHQRTHTEKPYKCPECGKSFSSQNSTLTHQRTHTGKKT	83

eZF/z8	LEPGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPEC GKSFSDNLRHQRTHTGEKPYKCPECGKSFSDNLHT HQRTHTGEKPYKCPECGKSFSDLRHQRTHTGEKPYK CPECGKSFSDQSNLTEHQRTHTGEKPYKCPECGKSFSDSGH LVHQRTHTGKKT	84
eZF/z9	LEPGEKPYKCPECGKSFSSKALTEHQRTHTGEKPYKCPEC GKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSDNLR HQRTHTGEKPYKCPECGKSFSDNLHTHQRTHTGEKPYK CPECGKSFSDLRHQRTHTGEKPYKCPECGKSFSDQSNL LTEHQRTHTGKKT	85
eZF/z10	LEPGEKPYKCPECGKSFSDCRDLARHQRTHTGEKPYKCE CGKSFSDNDALTEHQRTHTGEKPYKCPECGKSFSDNALT EHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPY KCPECGKSFSDPGNLVRHQRTHTGEKPYKCPECGKSFSDR AHLERHQRTHTGEKPYKCPECGKSFSDSSSLVRHQRTHTG EKPYKCPECGKSFSDRTLTNQRTHTGKKT	86
eZF/z11	LEPGEKPYKCPECGKSFSDNDALTEHQRTHTGEKPYKCE CGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSDPGNLV RHQRTHTGEKPYKCPECGKSFSDRAHLERHQRTHTGEKPY KCPECGKSFSDSSSLVRHQRTHTGEKPYKCPECGKSFSDRT LTNQRTHTGKKT	87
eZF/z12	LEPGEKPYKCPECGKSFSDNDALTEHQRTHTGEKPYKCE CGKSFSDPGHLVRHQRTHTGEKPYKCPECGKSFSDSGELV RHQRTHTGEKPYKCPECGKSFSDHLDLRHQRTHTGEKPY KCPECGKSFSSKALTEHQRTHTGEKPYKCPECGKSFSDQL AHLRAHQRTHTGEKPYKCPECGKSFSDHLNQRTHTG KKT	88
eZF/z13	LEPGEKPYKCPECGKSFSDNLRHQRTHTGEKPYKCE CGKSFSDRTLTNQRTHTGEKPYKCPECGKSFSDNLHT HQRTHTGEKPYKCPECGKSFSDSHLTEHQRTHTGEKPY KCPECGKSFSDSSSLVRHQRTHTGEKPYKCPECGKSFSDR NLHTHQRTHTGKKT	89
eZF/z14	LEPGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCE CGKSFSDNLRHQRTHTGEKPYKCPECGKSFSDQSGDLR RHQRTHTGEKPYKCPECGKSFSDHLDLRHQRTHTGEKPY KCPECGKSFSDSGNLVRHQRTHTGEKPYKCPECGKSFSDR NLVRHQRTHTGKKT	90
eZF/z15	LEPGEKPYKCPECGKSFSDDELNVHQRTHTGEKPYKCE CGKSFSDHLDLRHQRTHTGEKPYKCPECGKSFSDDLV RHQRTHTGEKPYKCPECGKSFSDNLRHQRTHTGEKPY KCPECGKSFSDRTLTNQRTHTGEKPYKCPECGKSFSDR NLHTHQRTHTGEKPYKCPECGKSFSDSHLTEHQRTHTGE KPYKCPECGKSFSDSSSLVRHQRTHTGEKPYKCPECGKSF REDNLHTHQRTHTGKKT	91
EGR1/z1	RPYACPVESCDRRFSRSDNLRHIRHTGQKPFQCRICMRN FSREDNLHTHQRTHTGEKPFACDICGRKFARSDLRHTKI HLRQKDRPYACPVESCDRRFSQSGNLTEHIRHTGQKPFQ CRICMRNFSTSGHLVRHRTHTGEKPFACDICGRKFAQNSTL TEHTKIHLRQKDK	92

EGR1/z13	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRN FSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKI HLRQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQC RICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNL HTHTKIHLRQKDK	93
EGR1/z15	RPYACPVESCDRRFSRRDELNVHIRIHTGQKPFQCRICMRN FSRSDHLTNHIRTHTGEKPFACDICGRKFARSDDLVRHTKI HLRQKDRPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQC RICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNL HTHTKIHLRQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQ KPFQCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFA REDNLHTHTKIHLRQKD	94
EGr1/z17	RPYACPVESCDRRFSDPGALVRHIRIHTGQKPFQCRICMRN FSRSDNLVRHIRTHTGEKPFACDICGRKFAQSGDLRRHTKI HLRQKDRPYACPVESCDRRFSTHLDLIRHIRIHTGQKPFQC RICMRNFSTSGNLVRHIRTHTGEKPFACDICGRKFARSDNL VRHTKIHLRQKDRPYACPVESCDRRFSQSGHLTEHIRIHTG QKPFQCRICMRNFSERSHLREHIRTHTGEKPFACDICGRKF AQAGHLASHTKIHLRQKD	95
EGR3/z1	RPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRS FSREDNLHTHIRTHTGEKPFACEFCGRKFARSDELVRHAKI HLKQKEHACPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRI CMRSFSTSGHLVRHIRTHTGEKPFACEFCGRKFAQNSTLTE HAKIHLKQKEK	96
EGR3/z13	RPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRS FSHRTTLTNHIRTHTGEKPFACEFCGRKFAREDNLHTHAKI HLKQKEHACPAEGCDRRFSTSHSLTEHLRIHTGHKPFQCRI CMRSFSQSSSLVRHIRTHTGEKPFACEFCGRKFAREDNLHT HAKIHLKQKEK	97
EGR3/z15	RPHACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICMRS FSRSDHLTNHIRTHTGEKPFACEFCGRKFARSDDLVRHAKI HLKQKEHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCR ICMRSFSHRTTLTNHIRTHTGEKPFACEFCGRKFAREDNLH THAKIHLKQKEHACPAEGCDRRFSTSHSLTEHLRIHTGHK PFQCRICMRSFSQSSSLVRHIRTHTGEKPFACEFCGRKFA REDNLHTHAKIHLKQKEK	98

Таблица 4

Последовательности целевых сайтов и хромосомная локализация иллюстративных целевых сайтов, свя-
занных с раскрытыми в настоящем документе ДНК-связывающими доменами

Начальн. положени е хром. 2	Последовательность целевого сайта	Целевой сайт	SEQ ID NO для последовательнос ти целевого сайта
166149168	CTAGGTCAAGTGTAGGAG	z1	18
166149158	ACTTGACCTAGACAGCCT	z2	19
166073978	TGAATAACTCATTAGTGA	z3	20

166073933	AAAGTACATTAGGCTAAT	z4	21
166149199	CCAGCACTGGTGCTTCGT	z5	22
166149176	AAGGCTGTCTAGGTCAAG	z6	23
166149168	CTAGGTCAAGTGTAGGAGACACAC	z7	24
166149165	GGTCAAGTGTAGGAGACA	z8	25
166149162	CAAGTGTAGGAGACACAC	z9	26
166149160	AGTGTAGGAGACACACTGCTGGCC	z10	27
166149160	AGTGTAGGAGACACACTG	z11	28
166149155	AGGAGACACACTGCTGGCCTG	z12	29
166128025	TAGGTACCATAGAGTGAG	z13	30
166127991	GAGGATACTGCAGAGGTC	z14	31
166127999	TAGGTACCATAGAGTGAGGCGAGGATG	z15	32
166127991	ATAGAGTGAGGCGAGGATGAAGCCGAG	z16	33
166127974	TGAAGCCGAGAGGATACTGCAGAGGTC	z17	34

Таблица 5

Аминокислотные последовательности для раскрытых в настоящем документе иллюстративных доменов активации транскрипции (TAD)

TAD	Последовательность	SEQ ID NO
VPR	DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMGLNSRSGSPKKRKRKVGSQLPDTDDRHRIEEKRRKRTYETFKSIMKKSPFSGPTDPRPPRRRIAVPSRSSASVPKPAPQPY PFTSSLSTINYDEFPTMVFPSPGQISQASALAPAPPQVLPQAPAPAPAMVSALAQAPAPVPVLPAGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDDELGALLGNSTDPVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAAPLGLNGLLSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPEASHLLEDPEETSQAVKALREMATVPIQKEEAICGQMDLSHPPPRGHLDELTTLESMTEDLNLDSPITPELNEILDTFLNDECLLHAMHISTGLSIFDTSLF	132
VP64	DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMGL	133
CITED 2	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPPHHHQQQPQHAFNALMGEHIHYGAGNMNATSGIRHAMGPGTVNNGHPPSALAPAARFNNSQFMGPPVASQGGSLPASMQLQKLNQYFNHHPYPHNHYPDLHPAAGHQMNGTNQHFRDCNPKHSGGSSTPGGSGGSSTPGGSGSSSGGAGSSNSGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMMSLVIEMGLDRIKELPELWLGQNEFDFMTDFVCKQQPSRVSC	134
CITED 4	ADHMLAEGYRLVQRPPSAAAAGHGHALRTLPPYAGPGLDSGLRPRGAPLGGPPPRQPGALAYGAFGPPSSFQPPFVPPPAAGIAHLQPVATPYPGRAAAPNPAGGPPGPQAPSAAPPAAHALGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGAPPAGSVSC	135
CREB3	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPOHMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLTKTEEQILKRV	224

Таблица 6

Аминокислотные последовательности для раскрытых в настоящем документе иллюстративных сконструированных факторов транскрипции (DBD + TAD)

Конструкция	Последовательность	SEQ ID NO
1	MAPKKKRVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTEKPYKCPECGKSFHRTTLTNH QRTHTGEKPYKCPECGKSFREDNLHTHQRTHTEG KPYKCPECGKSFSTSHSLTEHQRTHTEKPYKCPEC GKSFSQSSSLVRHQRTHTEKPYKCPECGKSFRED NLHTHQRTHTGKKTSKRPAATKAGQAKKKKGSY PYDVPDYALEEASGSGRADALDDFDLMLGSDALD DFDLMLGSDALDDFDLMLGSDALDDFDLMLIN SRSSGSPKKKRVGSQYLPDTDDRHRIEEKRRKTYE TFKSIMKKSFGPTDPRPPRRIVPSRSSASVPKPA PQYPFTSSLSTINYDEFPTMVFPQISQASALAPAP PQVLPQAPAPAPAMVSALAQAPAPVPLAPGPPQ AVAPPAPKPTQAGEGTLSEALLQLQFDDDEDLALLG NSTDPAVFTDLASVDNSEFQQLLNQIPVAPHTTEP MLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPGL LSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPK PEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPL PASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPE ASHLEDPDEETSQAVKALREMADEVIPQKEEAIC GQMDLSHPPRGLDELTTTLESMTEDLNLDSPPLTP ELNEILDFTLNDECLLHAMHISTGLSIFDTSLF	99
2	MAPKKKRVGIHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTEKPYKCPECGKSFREDNLHT HQRTHTEKPYKCPECGKSFSDDELVRHQRTHTEG KPYKCPECGKSFQSGNLTEHQRTHTEKPYKCPEC GKSFSSTSGHLVRHQRTHTEKPYKCPECGKSFQNS TLTEHQRTHTGKKTSKRPAATKAGQAKKKKGSYP YDVPDYALEEASGSGRADALDDFDLMLGSDALD DFDLMLGSDALDDFDLMLGSDALDDFDLMLIN SRSSGSPKKKRVGSQYLPDTDDRHRIEEKRRKTYE TFKSIMKKSFGPTDPRPPRRIVPSRSSASVPKPA PQYPFTSSLSTINYDEFPTMVFPQISQASALAPAP PQVLPQAPAPAPAMVSALAQAPAPVPLAPGPPQ AVAPPAPKPTQAGEGTLSEALLQLQFDDDEDLALLG	100

	NSTDPVFTDLASVDNSEFQQLLNQGIPVAPHTTEP MLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGL LSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPK PEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPL PASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPE ASHLLEDPEETSQAVKALREMADTVIPQKEEAIC GQMDLSHPPPRGHLDELTTTLESMTEDLNLDSP LTP ELNEILD TFLNDECLLHAMHISTGLSIFDTSLF	
3	MAPKKKRKVG IHG VPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSF SHRTTLTNH QRTHTGEKPYKCPECGKSF SREDNLH THQRTHTGE KPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCPEC GKSFSQSSSLVRHQRTHTGEKPYKCPECGKSF SRED NLH THQRTHTGKKT SKRPAATKKAGQAKKKKGSY PYDVPDYALEDALDDFDL DMLGSDALDDFDL DML GSDALDDFDL DMLGSDALDDFDL DML	101
4	MAPKKKRKVG IHG VPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTGEKPYKCPECGKSF SREDNLHT HQRTHTGEKPYKCPECGKSF SRSDELVRHQRTHTGE KPYKCPECGKSF S QSGNLTEHQRTHTGEKPYKCPEC GKSFSSTSGHLVRHQRTHTGEKPYKCPECGKSF SQNS TLTEHQRTHTGKKT SKRPAATKKAGQAKKKKGSYP YDVPDYALEDALDDFDL DMLGSDALDDFDL DMLG SDALDDFDL DMLGSDALDDFDL DML	102
5	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSA AAPP PPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFS DLGSAPPAG SVSCGGSGGGSGQS QLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFARE DNLHTHTKIH LRQKDRPYACPVESCDRRFSTSHSLT EHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGE KPFACDICGRKFAREDNLHTHTKIH LRQKDKLEMA DHLMLAEGYRLVQRPPSAAA AHGPHALRTLPPYAG PGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQ PFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGP PGPQPAPSA AAPP PPAHALGGMDAELIDEEALT SLE LELGLHRVRELPELFLGQSEFDCFS DLGSAPPAGSVS C	103
6	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSA AAPP PPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFS DLGSAPPAG SVSCGGSGGGSGQS QLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFARE DNLHTHTKIH LRQKDRPYACPVESCDRRFSTSHSLT	104

	EHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGE KPFACDICGRKFAREDNLHTHTKIHLRQKDK	
7	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPLDGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQFPFVPPPAAGIAHLQPVATPYPGRAAAPNP GGPPGPQAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCGGSGGGSGSOLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCR CMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDKLEMAD HMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGP GLDGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFP FPAVPPPAAGIAHLQPVATPYPGRAAAPNPAGGPP GPQAPSAAAPPPPAHALGGMDAELIDEEALTSLEL ELGLHRVRELPELFLGQSEFDCFDLGSAPPAGSVSC	105
8	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPLDGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQFPFVPPPAAGIAHLQPVATPYPGRAAAPNP GGPPGPQAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCGGSGGGSGSOLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCR CMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDK	106
9	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCD RRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDNL HTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPF QCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKF AQNSTLTEHTKIHLRQKDKLEMADHMLAEGYRLV QRPPSAAAAHGPHALRTLPPYAGPLDGLRPRGAP LGPPPPRQPGALAYGAFGPPSSFPFPAVPPPAAGIA HLQPVATPYPGRAAAPNPAGGPPGPQAPSAAAPP PPAHALGGMDAELIDEEALTSLELELGLHRVRELPE LFLGQSEFDCFDLGSAPPAGSVSC	107
10	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCD RRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDNL HTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPF QCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKF AQNSTLTEHTKIHLRQKDKLEMSGLEMADHMMAM NHGRFPDGTNGLHHHPAHRMGMGQFPPSPHHHQQQ QPQHAFNALMGEHIIHYGAGNMNATSGIRHAMGPG TVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLP ASMQLQKLNNQYFNHHPYPHNHYPDLHPAAGHQ MNGTNQHFRDCNPKHSGGSSTPGGSGGSSTPGGSG	108

	SSSGGAGSSNSGGGSGSGNMPASVAHVPAAMLPP NVIDTDFIDEEVLMSLVIEMGLDRIKELPELWLGQN EFDFTDFVCKQQPSRVSC	
11	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRM GMGQFPPSPHHHQQQPQHAFNALMGEHIIHYGAGN MNATSGVRHAMGPGTVNGGHPPSALAPAARFNNS QFMGPPVASQGGSLPASMQLQKLNQYFNHHPYPH NHYPDLHPAAGHQMNGTNQHFRDCNPKHSGSS TPGGSGSSSTPGGSGSSSGGAGSSNSGGGSGSGNM PASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGL DRIKELPELWLGQNEFDFTDFVCKQQPSRVSCQSQ LIKPSRMRKYPNRPSTPPHERPYACPVESCDRRFSR SDNLVRHIRHTGQKPFQCRICMRNFSREDNLHTHIR THTGEKPFACDICGRKFARSDDELVRHTKIHLRQKDR PYACPVESCDRRFSQSGNLTEHIRHTGQKPFQCRIC MRNFSSTGHLVRHIRTHTGEKPFACDICGRKFAQNS TLTEHTKIHLRQKDK	109
12	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQFPFPAVPPPAAGIAHLQPVATPYPGRAAAPNPAP GGPPGPQAPSA AAPPAAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGSGRPHACPAEGCDRRFSRSDNLVRH LRIHTGHKPFQCRICMRFSREDNLHTHIRHTHTGEK FACEFCGRKFARSDDELVRHAKIHLKQKEHACPAEG CDRRFSQSGNLTEHLRIHTGHKPFQCRICMRFSSTG HLVRHIRHTHTGEKPFACEFCGRKFAQNSTLTEHAKI HLKQKEKLEMADHMLAEGYRLVQRPPSAAA AHG PHALRTLPPYAGPGLDGLRPRGAPLGPPPPRQPGA LAYGAFGPPSSFQFPFPAVPPPAAGIAHLQPVATPYPG RAAAPNPAPGGPPGPQAPSA AAPPAAHALGGMD AELIDEEALTSLELELGLHRVRELPELFLGQSEFDC SDLGSAPPAGSVSC	110
13	MRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQ RICMRFSREDNLHTHIRHTHTGEKPFACEFCGRKFAR SDELVRHAKIHLKQKEHACPAEGCDRRFSQSGNLTE HLRIHTGHKPFQCRICMRFSSTGHLVRHIRHTHTGEK PFACEFCGRKFAQNSTLTEHAKIHLKQKEKLEMAD HMLAEGYRLVQRPPSAAA AHGPHALRTLPPYAGP GLDGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQ FPVPPPAAGIAHLQPVATPYPGRAAAPNPAPGGPP GPQAPSA AAPPAAHALGGMDAELIDEEALTSLEL ELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC	111
14	MRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQ RICMRFSREDNLHTHIRHTHTGEKPFACEFCGRKFAR SDELVRHAKIHLKQKEHACPAEGCDRRFSQSGNLTE HLRIHTGHKPFQCRICMRFSSTGHLVRHIRHTHTGEK PFACEFCGRKFAQNSTLTEHAKIHLKQKEKKAKEG GAPSASSAPPVSLAPVVTTCALEMSGLEMADHMM MNHGRFPDGTNGLHHHPAHRMGMGQFPPSPHHHQ QQPQHAFNALMGEHIIHYGAGNMNATSGIRHAMGP	112

	GTVNNGHPPSALAPAARFNNSQFMGPPVASQGGSL PASMQLQKLNNQYFNHHPYPHNHYMPDLHPAAGH QMNGTNQHFRDCNPKHSGGSSTPGGSGGSSTPGGS GSSSGGGAGSSNSGGGSGSGNMPASVAHVPAAML PNVIDTDFIDEVLMMLVEMGLDRIKELPELWLGQN EFDFTDFVCKQQPSRVSC	
15	MSGLEMADHMMAMNHGRFPDGTNGLHHPAHRM GMGQFPSPHHHQQQPQHAFNALMGEHIIYGAGN MNATSGVRHAMGPGTVNNGHPPSALAPAARFNNS QFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPH NHYMPDLHPAAGHQMNGTNQHFRDCNPKHSGGS TPGGSGGSSTPGGSGSSGGGAGSSNSGGGSGSGNM PASVAHVPAAMLPPNVIDTDFIDEVLMMLVEMGL DRIKELPELWLGQNEFDFTDFVCKQQPSRVSCRPH ACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICM RSFSREDNLHHTHRTHTGEKPFACEFCGRKFARSDEL VRHAKIHLKQKEHACPAEGCDRRFSQSGNLTEHLRI HTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEKPF ACEFCGRKFAQNSTLTEHAKIHLKQKEKKAKEGGAP SASSAPPVSLAPVVTCA	113
16	MTGKLAEKLPVTMSSLNQLPDNLYPEEIPALNLF SGSSDSVVHYNQMATENVMDIGLTNEKPNPELSYS GSFQPAPGNKTVTYLGKFAFDSPSNWCQDNIISLMS AGILGVPPASGALSTQTSTASMVQPPQGDVEAMYP ALPPYSNCGDLYSEPVSFHDPQGNPGLAYSPQDYQS AKPALDSNLFPMIPDYNLYHHPNDMGSIPEHKPFQ MDPIRVNPPITPLETIKAFKDKQIHPGFGSLPQPPLT LKPIRPRKYPNRPSTPLHERPHACPAEGCDRRFSRR DELNVHLRIHTGHKPFQCRICMRSFSRSDHLTNHIRT HTGEKPFACEFCGRKFARSDDLVRHAKIHLKQKEH ACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICM RSFSHRTTLNHIRTHTGEKPFACEFCGRKFAREDNL HTHAKIHLKQKEHACPAEGCDRRFSTSHSLTEHLRI HTGHKPFQCRICMRSFSQSSSLVRHIRTHTGEKPFAC EFCGRKFAREDNLHTHAKIHLKQKEKKAKEGGAPS ASSAPPVSLAPVVTCA	114
17	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSTTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVGLVSMNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAKGGFQVPMIPDYLF PQQQDGLGLGTPDQKPFQGLSRTQQPSLTPLSTIKA FATQSGSODLKALNTSYQSOLIKPSRMRKYPNRPST TPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQK FQCRICMRNFSHRTTLNHIRTHTGEKPFACDICGRK FAREDNLHTHTKIHLRQKDRPYACPVESCDRRFSTS HSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTH TGEKPFACDICGRKFAREDNLHTHTKIHLRQKDKKA DKSVVASSATSSLSSYSPVATSYSPVTTSSYSPATT	115

	SYPSPVPTSFSSPGSSTYSPVHSGFPSPSVATTYSSV PPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTI EIC	
18	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSLVSMNPPASSSSAPSPAASSA SASQPPLSCAVPSNDSSPIYSAAPTFFTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQDGLGLGTPDQKPFQGLSRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRRDELNVHIRIHTGQKP FQCRICMRNFSRDHLTNHIRTHTGEKPFACDICGRK FARSDDLVRHTKIHRLRQKDRPYACPVESCDRRFSRS DNLVRHIRIHTGQKPFQCRICMRNFSTRTTLTNHIRT HTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDRP YACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICM RNFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNL HTHTKIHLRQKDKKADKS VVASSATSSLSSYPSPVA TSYSPVTTSYSPATTSYSPVPTSFSSPGSSTYSPV HSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFS ASTGLSDMTATFSPRTIEIC	116
19	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSLVSMNPPASSSSAPSPAASSA SASQPPLSCAVPSNDSSPIYSAAPTFFTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQDGLGLGTPDQKPFQGLSRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRDNLVRHIRIHTGQKP FQCRICMRNFSTRTTLTNHIRTHTGEKPFACDICGRK FAREDNLHTHIRTHTGEKPFACDICGRKFSTSHSLTE HIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEK PFACDICGRKFAREDNLHTHTKIHLRQKDKKADKS VVASSATSSLSSYPSPVATSYPSPVTTSYSPATTSY SPVPTSFSSPGSSTYSPVHSGFPSPSVATTYSSVPPA FPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	117
20	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSLVSMNPPASSSSAPSPAASSA SASQPPLSCAVPSNDSSPIYSAAPTFFTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQDGLGLGTPDQKPFQGLSRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRDNLVRHIRIHTGQKP FQCRICMRNFSTRTTLTNHIRTHTGEKPFACDICGRK	115

	FAREDNLHTHTKIHRLRQKDRPYACPVESCDRRFSTS HSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTH TGEKPFACDICGRKFAREDNLHTHTKIHRLRQKDKKA DKSVVASSATSSLSSYPSPVATSYSPVTTSYSPATT SYSPVPTSFSSPGSSSTYSPVHSGFSPSPVATTYSSV PPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTI EIC	
21	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSTTRLPITYTGRFSLEPAPNS GNTLWPEPLFSLVGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQDGLGLGTPDQKPFQGLESRTOQPSTPLSTIKA FATQSGSODLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQK FQCRICMRNFSREDNLHTHIRTHTGEKPFACDICGR KFARSDDELVRHTKIHRLRQKDRPYACPVESCDRRFSQ SGNLTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHIR THTGEKPFACDICGRKFAQNSTLTHEHTKIHRLRQKDK KADKSVVASSATSSLSSYPSPVATSYSPVTTSYSP ATTSYSPVPTSFSSPGSSSTYSPVHSGFSPSPVATTY SSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFS PRTIEIC	118
22	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPALNLF SGSSDSVVHYNQMATENVMDIGLTNEKPNPELSYS GSFQPAPGNKTVTYLGKFAFDSPSNWCQDNIISLMS AGILGVPPASGALSTQTSTASMVQPPQGDVEAMYP ALPPYNSCGDLYSEPVSFHDPQGNPGLAYSPQDYQS AKPALDSNLFPMIPDYNLYHHPNDMGSIPEHKPFQ MDPIRVNPPITPLETIKAFKDKQIHPGFGLPQPPLT LKPIRPRKYPNRPSKTPLHERPHACPAEGCDRRFSR DNLVRHLRIHTGHKPFQCRICMRSFSHRITTLNHIRT HTGEKPFACEFCGRKFAREDNLHTHAKIHLKQKEH ACPAEGCDRRFSTSHSLTEHLRIHTGHKPFQCRICMR SFSQSSSLVRHIRTHTGEKPFACEFCGRKFAREDNLH THAKIHLKQKEKKAEEKGGAPSASSAPPVSLAPVVT CA	119
23	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSTTRLPITYTGRFSLEPAPNS GNTLWPEPLFSLVGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQDGLGLGTPDQKPFQGLESRTOQPSTPLSTIKA FATQSGSODLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSDPGALVRHIRIHTGQK FQCRICMRNFSRSDNLVRHIRTHTGEKPFACDICGR	120

	KFAQSGDLRRHTKIHRLRQKDRPYACPVESCDRRFST HLDLIRHIRIHTGQKPFQCRICMRNFSTSGNLVRHIR THTGEKPFACDICGRKFARSDNLVRHTKIHRLRQKDR PYACPVESCDRRFSQSGHLTEHIRIHTGQKPFQCRIC MRNFSERSHLREHIRTHTGEKPFACDICGRKFAQAG HLASHTKIHLRQKDKKADKSVVASSATSSLSSYPSP VATSYPSPVTTSPSPATTSPSPVPTSFSSPGSSSTYP SPVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNS FSASTGLSDMTATFSPRTIEIC	
24	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSTTRLPPIYTGFRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAKGGFQVPMIPDYLF PQQQGDGLGTPDQKPFQGLSRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLTRHIRIHTGQK FQCRICMRNFHSHSTLTNHIRTHTGEKPFACDICGRK FARSDNRKTHIRHTHTGEKPFACDICGRKFSTSHSLTE HIRIHTGQKPFQCRICMRNFSQSSSLTRHIRHTHTGEK FACDICGRKFARSDNRKTHTKIHRLRQKDKKADKSV VASSATSSLSSYPSPVATSYPSPVTTSPSPATTSPSP PVPTSFSSPGSSSTYPSPVHSGFPSPSVATTYSSVPPAF PAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	121
25	MTGKLAEKLPVTMSLLNQLPDNLYPEEIPALNLF SGSSDSVVHYNQMATENVMIDIGLTNEKPNPELSYS GSFQPAPGNKTVTYLGKFAFDSPSNWCQDNIILMS AGILGVPPASGALSTQTSTASMVQPPQGDVEAMYP ALPPYSNCGDLYSEPVSFHDPQGNPGLAYSPQDYQS AKPALDSNLFPMIPDYNLYHHPNDMGSIPEHKPFQ MDPIRVNPPITPLETIKAFKDKQIHPGFGSLPQPPLT LKPIRPRKYPNRPSKTPLHERPHACPAEGCDRRFSRS DNLVRHLRIHTGHKPFQCRICMRSFSREDNLHHTHIRT HTGEKPFACEFCGRKFARSDDELVRHAKIHLKQKEH ACPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRICM RSFSTSGHLVRHIRHTHTGEKPFACEFCGRKFAQNSTL TEHAKIHLKQKEKKAKEGGAPSASSAPPVSLAPVVT TCA	122
26	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSTTRLPPIYTGFRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQSPPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAKGGFQVPMIPDYLF PQQQGDGLGTPDQKPFQGLSRTQQPSLTPLSTIKA FATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLTRHIRIHTGQK FQCRICMRNFSRSDNLTHIRHTHTGEKPFACDICGRK	123

	FARSDERKRHIRTHTGEKPFACDICGRKFSQSGNLTE HIRIHTGQKPFQCRICMRNFSTSGHLTRHIRTHTGEK PFACDICGRKFAQSSTRKEHTKIHLRQKDKKADKSV VASSATSSLSSYPSPVATSYPSPVTTSYSPATTSYSP PVPTSFSSPGSSTYSPVHSGFPPSPVATTYSSVPPAF PAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	
27	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKL EEMMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGG GGGGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDI SLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNS GNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSA SASQPPLSCAVPSNDSSPIYSAAPTFTPNTDIFPEPQ SQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLF PQQQDGLGLGTPDQKPFQGLESRTOQPSTLPLSTIKA FATQSGSQDLKALNTSYQSOLIKPSRMRKYPNRPSK TPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQK FQCRICMRNFREDNLHHTHIRTHTGEKPFACDICGR KFARSDDELVRHIRTHTGEKPFACDICGRKFSQSGNL EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTHEHTKIHLRQKDKKADKS VVASSATSSLSSYPSPVATSYPSPVTTSYSPATTSYSP SPVPTSFSSPGSSTYSPVHSGFPPSPVATTYSSVPPA FPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	124
28	MAPKKKRKVGIIHGVPAAALEPGEKPYKCPECGKSFSS PADLTRHQRTHHTGEKPYKCPECGKSFSDNLVRH QRTHHTGEKPYKCPECGKSFSDNLHHTHQRTHHTGE KPYKCPECGKSFSDNLHHTHQRTHHTGEKPYKCPE GKSFSSQSGNLTEHQRTHHTGEKPYKCPECGKSFSTSG HLVRHQRTHHTGKTKSRPAATKAGQAKKKKGSY PYDVDPDYALEDALDDFDLMLGSDALDDFDLDM GSDALDDFDLMLGSDALDDFDLDM	125
29	MAPKKKRKVGIIHGVPAAALEPGEKPYKCPECGKSFSS DPGALVRHQRTHHTGEKPYKCPECGKSFSDNLVR HQRTHHTGEKPYKCPECGKSFSSQSGDLRRHQRTHHTG EKPYKCPECGKSFSTHLDLIRHQRTHHTGEKPYKCPE CGKSFSTSGNLVRHQRTHHTGEKPYKCPECGKSFSD DNLVRHQRTHHTGKTKSRPAATKAGQAKKKKGSY PYDVDPDYALEDALDDFDLMLGSDALDDFDLDM LGSDALDDFDLMLGSDALDDFDLDM	126
30	MAPKKKRKVGIIHGVPAAALEPGEKPYKCPECGKSFSS RSDNLVRHQRTHHTGEKPYKCPECGKSFSDNLVRH QRTHHTGEKPYKCPECGKSFSDNLHHTHQRTHHTGE KPYKCPECGKSFSTSHSLTEHQRTHHTGEKPYKCPE GKSFSSQSSSLVRHQRTHHTGEKPYKCPECGKSFSD NLHHTHQRTHHTGKTKSRPAATKAGQAKKKKGSY PYDVDPDYALEEASGSGRADALDDFDLMLGSDALD DFDLMLGSDALDDFDLMLGSDALDDFDLMLIN SRSSGSPKKRKGVSQYLPDTPDRHRIEERKRRTYE TFKSIMKKSPPSGPTDPRPPRRIVPSRSSASVPKPA PQPYPFTSSLSTINYDEFPTMVFPQISQASALAPAP PQVLPQAPAPAPAMVSALAQAPAPVPLAPGPPQ	99

	AVAPPAPKPTQAGEGTLSEALLQLQFDDDEDLGALLG NSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEP MLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGL LSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPK PEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPL PASLAPTPTGVPHEPVGSLTPAPVPQPLDPAPAVTPE ASHLLEDPEETSQAVKALREMADTVIPQKEEAIC GQMDLSHPPPRGHLDELTTTLESMTEDLNLDSP LTP ELNEILD TFLNDECLLHAMHISTGLSIFDTSLF	
31	MAPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTH TGEKPYKCPECGKSFSREDNLHT HQRTH TGEKPYKCPECGKSFSRSDELVRHQRTH TGE KPYKCPECGKSFSQSGNLTEHQRTH TGEKPYKCPEC GKSFSSTSGHLVRHQRTH TGEKPYKCPECGKSFSQNS TLTEHQRTH TGGKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEDALDDFDL DMLGSDALDDFDL DMLG SDALDDFDL DMLGSDALDDFDL DML	102
32	MAPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTH TGEKPYKCPECGKSFSREDNLHT HQRTH TGEKPYKCPECGKSFSRSDELVRHQRTH TGE KPYKCPECGKSFSQSGNLTEHQRTH TGEKPYKCPEC GKSFSSTSGHLVRHQRTH TGEKPYKCPECGKSFSQNS TLTEHQRTH TGGKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEDALDDFDL DMLGSDALDDFDL DMLG SDALDDFDL DMLGSDALDDFDL DML	102
33	MAPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTH TGEKPYKCPECGKSFSREDNLHT HQRTH TGEKPYKCPECGKSFSRSDELVRHQRTH TGE KPYKCPECGKSFSQSGNLTEHQRTH TGEKPYKCPEC GKSFSSTSGHLVRHQRTH TGEKPYKCPECGKSFSQNS TLTEHQRTH TGGKTSKRPAATKKAGQAKKKKGSYP YDVPDYALEEASGSGRADALDDFDL DMLGSDALD DFDL DMLGSDALDDFDL DMLGSDALDDFDL DMLIN SRSSGSPKKR KVG SQYLPDTDDRHRIEEKRKR TYE TFKSIMKKS PFGPTDPRPPRRIAVPSRSSASVPKPA PQYPFTSSLSTINYDEFPTMVFP SGQISQASALAPAP PQVLPQAPAPAPAMVSALAQAPAPVPVLAPGPPQ AVAPPAPKPTQAGEGTLSEALLQLQFDDDEDLGALLG NSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEP MLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGL LSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLPK PEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPL PASLAPTPTGVPHEPVGSLTPAPVPQPLDPAPAVTPE ASHLLEDPEETSQAVKALREMADTVIPQKEEAIC GQMDLSHPPPRGHLDELTTTLESMTEDLNLDSP LTP ELNEILD TFLNDECLLHAMHISTGLSIFDTSLF	100
34	MAPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTH TGEKPYKCPECGKSFSREDNLHT HQRTH TGEKPYKCPECGKSFSRSDELVRHQRTH TGE KPYKCPECGKSFSQSGNLTEHQRTH TGEKPYKCPEC GKSFSSTSGHLVRHQRTH TGEKPYKCPECGKSFSQNS	127

	TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGS ALDDFDLMLGSDALDDFDLMLGSDALDDFDL MLGSDALDDFDLML	
35	MAPKKRKGVIHGVPAALEPGEKPYKCPECCKSFS RSDNLVRHQRTHTGEKPYKCPECCKSFSREDNLHT HQRTHTGEKPYKCPECCKSFSRSDNLVRHQRTHTGE KPYKCPECCKSFSQSGNLTEHQRTHTGEKPYKCPEC CKSFSSTSGHLVRHQRTHTGEKPYKCPECCKSFSQNS TLTEHQRTHTGKKTSKRPAATKKAGQAKKKKGS ALDDFDLMLGSDALDDFDLMLGSDALDDFDL MLGSDALDDFDLML	127
36	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQFPFVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSRED NLHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIH LRQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQK PFQCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGR KFAQNSTLTEHTKIHLRQKDK	128
37	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQFPFVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRIC MRNFSREDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLTE HIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDK	106
38	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQFPFVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCGGSGGGSGGGSGGSGQSQLIKPSRMRKYPN RPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHT GQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFACDI CGRKFARSDELVRHTKIHLRQKDRPYACPVESCDRR FSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGHLVR HIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHLRQK DK	129
39	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQFPFVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE	105

	<p>RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSREDNLHTHIRHTHTGEKPFACDICGRKFARS DELVRHTKIHRLRQKDRPYACPVESCDRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRHTHTGE KPFACDICGRKFAQNSTLTEHTKIHRLRQKDKLEMAD HLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGP GLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFP FPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLEL ELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC</p>	
40	<p>MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCADHMLAEGYRLVQRPPSAAAAHGPHALRTL PPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFG PPSFPFPAVPPPAAGIAHLQPVATPYPGRAAAPPN APGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEA LTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPP AGSVSCQSOLIKPSRMRKYPNRPSKTPPHERPYACP VESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNF REDNLHTHIRHTHTGEKPFACDICGRKFARSDELVRH TKIHRLRQKDRPYACPVESCDRRFSQSGNLT GQKPFQCRICMRNFSTSGHLVRHIRHTHTGEKPFACDI CGRKFQNSTLTEHTKIHRLRQKDK</p>	130
41	<p>MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGADHMLAEGYRLVQRPPSAAAAH GPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPG ALAYGAFGPPSSFPFPAVPPPAAGIAHLQPVATPY GRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGGM DAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSCGGSGGGSGQSOLIKPSRMRKY PNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRI HTGQKPFQCRICMRNFREDNLHTHIRHTHTGEKPFAC DICGRKFARSDELVRHTKIHRLRQKDRPYACPVESC DRRFSQSGNLT EHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRHTHTGEKPFACDICGRKFAQNSTLTEHTKIHL RQKDK</p>	131
42	<p>MAPKKRKRKVGIVPAALEPGEKPYKCPECGKSFS RSDNLVRHQRTHTHTGEKPYKCPECGKSFSREDNLHT HQRTHTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGE KPYKCPECGKSFSQSGNLT TEHQRTHTHTGEKPYKCPECGKSFSQSGNLT TEHQRTHTHTGKTSKRPAATKAGQAKKKKGSYP YDVPDYALEDALDDFDLMLGSDALDDFDLMLG SDALDDFDLMLGSDALDDFDLML</p>	102

43	MAPKKKRVGIHGVPAALEPGEKPYKCPECCKSFS RSDNLVRHQRTHTGEKPYKCPECCKSFSREDNLHT HQRTHTGEKPYKCPECCKSFSRSDNLVRHQRTHTGE KPYKCPECCKSFSQSGNLTEHQRTHTGEKPYKCPEC GKSFSTSGHLVRHQRTHTGEKPYKCPECCKSFSQNS TLTEHQRTHTGKKTSKRPAATKAGQAKKKKGSYP YDVPDYALEDALDDFDLMLGSDALDDFDLMLG SDALDDFDLMLGSDALDDFDLML	102
44	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSTPPHE RPYACPVESCDRRFSRSDNLVRHIRHTGQKPFQCRI CMRNFSDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDK	106
45	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCGGSGGGSGQSQLIKPSRMRKYPNRPSTPPHE RPYACPVESCDRRFSRSDNLVRHIRHTGQKPFQCRI CMRNFSDNLHTHIRTHTGEKPFACDICGRKFARS DELVRHTKIHLRQKDRPYACPVESCDRRFSQSGNLT EHIRHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDK	106
46	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCGGSGGGSGGGSGGGSGQSQLIKPSRMRKYPN RPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRHT GQKPFQCRICMRNFSDNLHTHIRTHTGEKPFACDI CGRKFARSDDELVRHTKIHLRQKDRPYACPVESCDRR FSQSGNLTEHIRHTGQKPFQCRICMRNFSTSGHLVR HIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHLRQ DK	205
47	MAADHMLAEGYRLVQRPPSAAA AHGPHALRTLPP YAGPGLDGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAP GGPPGPQPAPSAAAPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFDLGSAPPAG SVSCADHMLAEGYRLVQRPPSAAA AHGPHALRTL PPYAGPGLDGLRPRGAPLGPPPPRQPGALAYGAFG PPSSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPN APGGPPGPQPAPSAAAPPPAHALGGMDAELIDEEA	207

	LTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPP AGSVSCQSQLIKPSRMRKYPNRPSKTPPHERPYACP VESCRRFRSDNLVRHIRIHTGQKPFQCRICMRNFS REDNLHTHIRHTHTGEKPFACDICGRKFARSEDELVRH TKIHLRQKDRPYACPVESCRRFRFSQSGNLTEHIRIHT GQKPFQCRICMRNFSTSGHLVRHIRHTHTGEKPFACDI CGRKFAQNSTLTEHTKIHLRQKDK	
48	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPP YAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVA TPYPGRAAAPNP GGPPGPQPAPSAAAPPPAHALGGMDAELIDEEALT SLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSCGGSGGGSGADHMLAEGYRLVQRPPSAAAAH GPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPG ALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVA TPYP GRAAAPNPAGPPGPQPAPSAAAPPPAHALGGM DAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSCGGSGGGSGSQLIKPSRMRKY PNRPSKTPPHERPYACPVESCRRFRSDNLVRHIRI HTGQKPFQCRICMRNFSREDNLHTHIRHTHTGEKPFAC DICGRKFARSEDELVRHTKIHLRQKDRPYACPVESC RRFRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRHTHTGEKPFACDICGRKFAQNSTLTEHTKIHL RQKDK	209
49	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPQ HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRLEPGEKPYKCPECGKSFSRSDNLVR HQRHTHTGEKPYKCPECGKSFSREDNLHTHQRHTHTG EKPYKCPECGKSFSRSEDELVRHQRHTHTGEKPYKCPE CGKSFSQSGNLTEHQRHTHTGEKPYKCPECGKSFS GHLVRHQRHTHTGEKPYKCPECGKSFSQNSTLTEHQ RTHTGKKTSVYVGGLESRLKYTAQNMELQNKVQ LLEEQLNSLLDQLRKLQAMVIEISNKTSSSSTCILVLL VSFCLLLVPAMYSSDTRGSLPAEHGVL SRQLRALPS EDPYQLELPALQSEVPKSTHQWLDGSDCVLQAPG NTSCLLHYMPQAPSAEPPELWPFDFSEPLCRGPIL PLQANLTRKGGWLPTGSPSVILQDRYSG	213
50	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPQ HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRRPYACPVESCRRFRSDNLVRHIRI HTGQKPFQCRICMRNFSREDNLHTHIRHTHTGEKPFAC DICGRKFARSEDELVRHTKIHLRQKDRPYACPVESC RRFRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRHTHTGEKPFACDICGRKFAQNSTLTEHTKIHL RQKDVYVGGLESRLKYTAQNMELQNKVQLLEE QLNSLLDQLRKLQAMVIEISNKTSSSSTCILVLLVSFCL LLVPAMYSSDTRGSLPAEHGVL SRQLRALPSEDYPYQ	217

	LLELPALQSEVPKSTHQWLDGSDCVLQAPGNTSCL LHYMPQAPSAEPPLEWPFDFLSEPLCRGPILPLQAN LTRKGGWLPTGSPSVILQDRYSG	
51	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPO HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRPPYACPVESCDRRFRSDNLVRHIRI HTGQKPFQCRICMRNFSHRTTTLTNHIRTHTGKPF CDICGRKFAREDNLHTHTKIHLRQKDRPYACPVESC DRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSL VRHIRTHTGKPFACDICGRKFAREDNLHTHTKIHL RQKDVYVGGLESRLVLYTAQNMELQNKVQLLEEQ NLSLLDQLRKLQAMVIEISNKTSSSSTCILVLLVFSCL LLVPAMYSSDTRGSLPAEHGVLRSRQLRALPSEDPYQ LLELPALQSEVPKSTHQWLDGSDCVLQAPGNTSCL LHYMPQAPSAEPPLEWPFDFLSEPLCRGPILPLQAN LTRKGGWLPTGSPSVILQDRYSG	219
52	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPO HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRLEPGKPYKCPECGKSFSRSDNLVR HQRTHHTGKPYKCPECGKSFSREDNLHHTHQRTHHTG EKPYKCPECGKSFSRDELVRHQRTHHTGKPYKCPE CGKSFSQSGNLTEHQRTHHTGKPYKCPECGKSFS GHLVRHQRTHHTGKPYKCPECGKSFSQNSTLTEHQ RTHHTGKTSVYVGGLESRLVLYTAQNMELQNKVQ LLEEQLNLSLLDQLRKLQAMVIEIS	221
53	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPL DWALPLSEVPSDWEVDDLLCSLLSPPASLNILSSNP CLVHHDHTYSLPRETVSMDLESESCRKEGTQMTPO HMEELAEQEIARLVLTDEEKSLLEKEGLILPETLPLT KTEEQILKRVRPPYACPVESCDRRFRSDNLVRHIRI HTGQKPFQCRICMRNFSREDNLHTHIRTHTGKPF CDICGRKFARSDELVRHTKIHLRQKDRPYACPVESC DRRFSSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRTHTGKPFACDICGRKFAQNSTLTEHTKIHL RQKDVYVGGLESRLVLYTAQNMELQNKVQLLEEQ NLSLLDQLRKLQAMVIEIS	223

Таблица 7

Последовательности нуклеиновой кислоты, кодирующие раскрытые в настоящем документе иллюстративные сконструированные факторы транскрипции

Конструкция	Последовательность	SEQ ID NO
31 (RE + кодирующая)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatatttgaagtctgttgacataaagaatcatgatattaatgccatggaaatg aaaggcgatcaaacactatggttgaagggggaaattgtagagcacagatgt	67
	gttcgtgtggcagtgctgctcttagcaatactcagagaagagagagaacaatg aaatctgattggcccagtgtagcccagatgaggtcagctgccaactttctctt tcacatctatgaaagtcatttaagcacaactaactttttttttttttttttttgagaca gactcttgcctgtgcccaggacagagtgtagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctccaagtagc tggatfacaggagtggccaccatgccagctaattttgtattttaatagatacg ggggttcaccatatacaccaggctggtctcgaactcctggcctcaagtatcca cctgcctcggcctccaaagtgtggattatagcgctcagcactatgcccaac ccgaccaactttttaataaataatttaaaaaattggtattccatataatactagta ttacattatccacacaaaacggacgggctccgctgaaccagtgaggcccca gacgtgcgataaataaccctgctgctcaccacctggggaaggggggag gaccaggtaaatggagcgcgcataagcggagcgcgggtccttttct ctgccgggtggcactggtagctgtgccaggtgtgtaactttgatggggcccag ggctggagctcaaggaaagcgtcgcagggtcacagatctgggggaacccccg ggaaaagcactgaggcaaaccccgctcctcctcaatataatggagggg gagggtgagtctgattactcataagaccttttttcttccggcgcaaa accgtgagctggattataatcgcctataaagctccagaggcggcagccact gcagaggagcccccgctcgcgcgactagctgccccgcgagcaacggcct cgtgatttcccccgatccggctccccgcctccccactctgccccgcctacccc ggagccgtgagcgcctcctccgaatctctctctctctcgcgctcgcgctgga gagggaactagcagaacgaggaagcagctggaggtgacgccgggagcagatt acgctgtcaggggccgagccgagcggatcctggcgctgtgcagaggaaa ggagggagtggcggctcgtgtcagagcccaggtgggtaagctagcagac cacctggacttcccagcgcaccaaccgtggctttcagccaggctcctctcctccg cggcttctcaaccaacccateccagcgcggccaccaacctcccgaatga gtgttctcctcccagcagccgaaggcctactaggaacgtaacctgttacttt tccaggggctgtagtcgacccgctcggagtgctgtgcgactgcgcgcgcg gggctagagtgaaggtgactgtggttctctcggccaagtccgagggaagc gtaaagatatgggccttttccccctctacctgtctcaccaaagtccctagtc ccgagcagtagcctttttccagggaattagccagacacaacaacgggaac cagacaccgaaccagacatgccggcccgtgcccctcccgtcgtcgtccttt cctcctctgtctcctcagagccggtcttcaaggggagcctcctgccccgg ctgctcagtcctccggtgtgagggaccgggaggtcctcccgcacagctcct gttctcttgcagcctggttctcgcggaccagctcaggactctggacagtga ggccccgggacgaccagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA CTGAGCATCAACGAATCATAACGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA	

	<p>ATGCGGTA AAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACTGGGAAAAAAA ACGAGTAAAAGGCCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGACGCGC TGGACGATTTTCGATCTCGACATGCTGGGTTCTGA TGCCCTCGATGACTTTGACCTGGATATGTTGGGA AGCGACGCATTGGATGACTTTGATCTGGACATGC TCGGCTCCGATGCTCTGGACGATTTTCGATCTCGA TATGTTATAAACTAGTaaagagaccggtcactgtgacagtaaaa gagaccggtcactgtgagaatgaaagagaccggtcactgtgatcgaaaaga gaccggtcactgtgagcggcctgaaaccagcagacaatgtagctcagtaga aaccagcagacaatgtagctgaatggaaccagcagacaatgtagctcgg agaaaccagcagacaatgtagctAAGCTTGGGTGGCATCCC TGTGACCCCTCCCCAGTGCCCTCCTGGCCCTGG AAGTTGCCACTCCAGTGCCACCAGCCTTGTCT AATAAAATTAAGTTGCATCATTTTGTCTGACTAG GTGTCTTCTATAATATTATGGGGTGGAGGGGG TGGTATGGAGCAAGGGGCAAGTTGGGAAGACAA CCTGTAGGGCCTGCGGGGTCTATTGGGAACCAA GCTGGAGTGCAGTGGCACAATCTTGGCTCACTGC AATCTCCGCCTCCTGGGTCAAGCGATTCTCCTG CCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCAT GCATGACCAGGCTCAGCTAATTTTTGTTTTTTGG TAGAGACGGGGTTTTACCATATTGGCCAGGCTGG TCTCCAACCTAATCTCAGGTGATCTACCCACC TTGGCCTCCCAAATTGCTGGGATTACAGGCGTGA ACCACTGCTCCCTTCCCTGTCTT</p>	
32 (RE + кодирующая)	<p>ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatafttgaagtctgtgacataaagaatcatgatattaatgccatggaaatg aaagggcgatcaacactatggttgaaggggaaattgtagacacagatg gttcgtgtgacagtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattgccccagtgtagccagatgaggtcagctccaactttctctt tcacatctatgaaagtcatttaagcacaactactttttttttttttttttttttgagaca gagtctgtctgtgcccagacagagtgagtagtactcaatctcggctcact gcagcctccacctcctaggctcaaacggtcctcctgcatcagcctccaagtagc tgaattacaggagtgcccacatgcccagctaattttgtatttttaatagatacg ggggttcaccatataccagcgtggtctcgaactcctggcctcaagtatcca cctgcctcggcctccaaagtgtgggattataggcgtcagccactatgcccac ccgaccaacttttttaaaataaataatlaaaaaattggtatttcacataatactagta ttacattatccacacaaaacggacggcctccgctgaaccagtgaggcccca gacgtgcgataaataaccctgctgctgcaccacctggggagagggggag gaccacggtaatggagcgcgcatagcaaaaaggacgcggggtcctttct ctccgggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaacccgg ggaaaagcactgaggcaaaccccgctcgtctcctcaataatattggaggggg gaggttgagtactgtgattactataagaccttttttttctccgggcgcaaa accgtgagctggattataatcgccctataaagctccagagggcgtcaggcacct gcagaggagccccccgctcggcctcagtagctgcccccgagcaacggcct cgtgattccccccgatccggctccccgctccccactctcccccgctacccc</p>	68

<p> ggagccgtgcagccgcctctccgaatctctctctctctctggcgtcgcgtgcga gagggaaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcctggcgctgtgcagaggaaa ggcgggagtgcccggctcgtgtcgagagccgaggtgggtaagctagcgac cactggactcccagcgcceaacctggcttttcagccaggtcctctctccc cgcttctcaaccaacccatcccagcgcggccaccaacctcccgaatga gtgcttctgccccagcagccgaaggcgtactaggaacgtaacctgttacttt tccaggggctgtagtcgacccgctccccgagttgctgtgcgactgcgcgcg gggtagagtgcaaggtgactgtggttctctggccaagtccgagggagaac gtaaagatatggccttttcccctctcacctgtctcaccaaagtcctagtccc cggagcagtagccttttttccaggaattagccagacacaacaacgggaac cagacaccgaaccagacatcccgcctcctcccgcctcgtcgtccttt cctcctctgtctctccagagccgatctcaagggagcctccgtgcccccg ctgctcagtcctccggtgtgcaggaccgccgaagtctcccgcacagctctc gcttcttttcagcctgttctgcgccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTAGTTCG AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCGGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTACAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGACGCGC TGGACGATTCGATCTCGACATGCTGGGTTCTGA TGCCCTCGATGACTTTGACCTGGATATGTTGGGA AGCGACGCATTGGATGACTTTGATCTGGACATGC TCGGCTCCGATGCTCTGGACGATTTTCGATCTCGA TATGTTATAAACTAGTAAAGCTGGGTGGCATCCC TGTGACCCCTCCCCAGTGCTCTCCTGGCCCTGG AAGTTGCCACTCCAGTGCCACCAGCCTTGTCT AATAAAATTAAGTTGCATATTTTGTCTGACTAG GTGTCCTTCTATAATATTATGGGGTGGAGGGGG TGGTATGGAGCAAGGGGCAAGTTGGGAAGACAA CCTGTAGGGCCTGCGGGGTCTATTGGGAACCAA GCTGGAGTGCAGTGGCACAATCTTGGCTCACTGC AATCTCCGCTCCTGGGTTCAAGCGATTCTCCTG CCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCAT </p>	
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

	GCATGACCAGGCTCAGCTAATTTTTGTTTTTTTGG TAGAGACGGGGTTTCACCATATTGGCCAGGCTGG TCTCCAACCTCCTAATCTCAGGTGATCTACCCACC TTGGCCTCCCAAATTGCTGGGATTACAGGCGTGA ACCACTGCTCCCTTCCCTGTCCTT	
33 (RE + кодирующая)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatatttgaagtctgttgacataaagaatcatgatattaatgccatggaaatg aaagggcgatcaacactatggttgaaaaggggaaattgtagagcacagatgt gttcgtgtgagcagtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggcccagtgtagcccagatgaggtcagctccaactttctctt tcacatcttatgaaagtcatttaagcacaactaactttttttttttttttttttttgagaca gagcttctctgttgccaggacagagtgtagtagtactcaatctcggctcact gcagcctccacctctaggctcaaacggctcctctgcatcagcctccaagtagc tggattacagaggatgcccaccatgcccagtaattttgtatttttaataagatac ggggttcaccatataccaggctggtctcgaactcctggcctcaagtatcca cctgcctcggcctccaaagtgtgggattataggcgtcagccactatgccaaac ccgaccaacctttttaaataaataatftaaaaaattggtatttcacatataata ttacatttatccacaaaacggagcggcctcgtgaaccagttagggccca gacgtgcgataaataaccctgctgctgcaccacctggggagagggggag gaccacggtaatggagcgcgcatagcaaaaggacgcggggtcctttct ctccgggtggcactggtagctgtggcaggtgtgtagtttgatggggcccag ggctggagctcaaggagcgtcgcaggggtcacagatctgggggaaccccgg ggaaaagcactgaggcaaacccgctcgtctcctcaatatagggagggg gaggtgagtagctctgtagtactataagaccttttttttcttccggcgcaaa accgtgagctggattataatcgcctataaagctccagaggcgtcaggcacct cgagaggagccccccgctcgcgcgactagctgccccgcgagcaacggcct cgtgatttccccccgatccggtccccgctccccactctgccccgctacc ggagccgtgcagccgctctccgaatctctctctctctcctgctcgcgtcgcga gagggaaactagcagaaacgaggaagcagctggaggtgacgccgggagatt acgctgtcaggggccagccgagcggatcgtggcgtctgtgcagagaaa ggcgggagtgcccggctcgtctgcagagccgaggtgggtaagtagcagac cactggactcccagcggcaaccgtggctttcagccaggtcctctctccc cggtctcaaccaaccateccagcggccggccaccaacctcccgaatga gtgcttctccccagcagccgaaggcgtactaggaacggtaacctgttacttt tcaggggcctgtagtcgaccgctccccgagttgctgtcgcactgcgcgcg gggctagagtgcaaggtgactgtggttctctctgccaagtccgagggagaac gtaaagatatgggccttttccccctctcacctgtctcacaagtccttagtccc cggagcagtagcctctttttccaggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccggccgctgcgccctccccgctcgtccttt cctccctctgtctctccagagccggtcttcaaggggagcctccgtgccccgg ctgctcagtcctccggtgtgcaggaccgggaagtcctccccgcacagctctc gcttctcttgcagcctgttctgcgcccggaccagtcgaggactctggacagtaga ggccccgggagaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCTTGAATGTGGGAAATCATTAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA	69

<p> TACAAGTGTCCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGAGGCC AGCGGTTCCGGACGGGCTGACGCATTGGACGAT TTTGATCTGGATATGCTGGGAAGTGACGCCCTCG ATGATTTTGACCTTGACATGCTTGGTTCCGATGC CCTTGATGACTTTGACCTCGACATGCTCGGCAGT GACGCCCTTGATGATTTTCGACCTGGACATGCTGA TTAACTCTAGAAGTTCCGGATCTCCGAAAAAGAA ACGCAAAGTTGGTAGCCAGTACCTGCCCGACAC CGACGACCGGCACCGGATCGAGGAAAAGCGGAA GCGGACCTACGAGACATTCAAGAGCATCATGAA GAAGTCCCCCTTCAGCGGCCCCACCGACCCTAGA CCTCCACCTAGAAGAATCGCCGTGCCCAGCAGAT CCAGCGCCAGCGTGCCAAAACCTGCCCCCCAGC CTTACCCCTTACCAGCAGCCTGAGCACCATCAA CTACGACGAGTTCCCTACCATGGTGTTCCTCCAGC GGCCAGATCTCTCAGGCCTCTGCTCTGGCTCCAG CCCCTCCTCAGGTGCTGCCTCAGGCTCCTGCTCC TGCACCAGCTCCAGCCATGGTGTCTGCACTGGCT CAGGCACCAGCACCCGTGCCTGTGCTGGCTCCTG GACCTCCACAGGCTGTGGCTCCACCAGCCCCTAA ACCTACACAGGCCGCGAGGGCACACTGTCTGA AGCTCTGCTGCAGCTGCAGTTCGACGACGAGGAT CTGGGAGCCCTGCTGGGAAACAGCACCGATCCT GCCGTGTTACCGACCTGGCCAGCGTGGACAAC AGCGAGTTCCAGCAGCTGCTGAACCAGGGCATC CCTGTGGCCCCCTCACACCACCGAGCCCATGCTGA TGGAATACCCCGAGGCCATCACCCGGCTCGTGAC AGGCGCTCAGAGGCCTCCTGATCCAGCTCCTGCC CCTCTGGGAGCACAGGCCTGCCTAATGGACTGC TGTCTGGCGACGAGGACTTCAGCTCTATCGCCGA TATGGATTTCTCAGCCTTGCTGGGCTCTGGCAGC GGCAGCCGGGATTCCAGGGAAGGGATGTTTTTG CCGAAGCCTGAGGCCGGCTCCGCTATTAGTGACG TGTTTGAGGGCCGCGAGGTGTGCCAGCCAAAAC GAATCCGGCCATTTATCCTCCAGGAAGTCCATG GGCCAACCGCCACTCCCCGCCAGCCTCGACCA ACACCAACCGGTCCAGTACATGAGCCAGTCGGG TCACTGACCCCGGCACCAGTCCCTCAGCCACTGG </p>

	<p>ATCCAGCGCCCGCAGTACTCCCAGGCCAGTC ACCTGTTGGAGGATCCCGATGAAGAGACGAGCC AGGCTGTCAAAGCCCTTCGGGAGATGGCCGATA CTGTGATTCCCCAGAAGGAAGAGGCTGCAATCT GTGGCCAAATGGACCTTCCCATCCGCCCCCAAG GGGCCATCTGGATGAGCTGACAACCACACTTGA GTCCATGACCGAGGATCTGAACCTGGACTCACCC CTGACCCCGGAATTGAACGAGATTCTGGATACCT TCCTGAACGACGAGTGCCTCTTGCATGCCATGCA TATCAGCACAGGACTGTCCATCTTCGACACATCT CTGTTTTAACTAGTaaataaagatctttatttcattagatctgtgtg tggtttttgtgtg</p>	
<p>34 (RE + кодирующая)</p>	<p>ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacataatttgaagtctgttgacataaagaatcatgatattaatgccatggaaatg aaaggcgatcaacactatggttgaagggggaaattgtagagcacagatgt gttcgtgtggcagtgctgtctctagcaatactcagagaagagagagaacaatg aaattctgattggcccagtgtagccagatgaggctcagctgccaactttctt tcacatctatgaaagtatttaagcacaactaactttttttttttttttttttttgagaca gagtcttctctgttgcccagagacagagtgtagtagtcaatctcggctcact gcagcctccacctctaggtcaaacggctcctctgcatcagcctccaagtagc tgggaattacaggagtgcccaccatgccagctaattttgtatttttaatagatacg ggggttcaccatataccagcctggtctcgaactcctggcctcaagtatcca cctgctcggcctccaaagtctgggattataggcgtcagccactatgcccaac ccgaccaactttttaataataatftaaaaaattggtatttcacatataactagta ttacatttatecacacaaaacggagcggcctccgctgaaccagtgaggcccca gacgtgcccataaataaccctgctgctgaccacctggggagaggggggag gaccacggtaaatggagcagcgcatagcaaaaggagcgcgggtcttttct ctgcccgtggcactgggtagctgtgcccaggtgtgtaactttgatggggcccag ggctggagctcaaggaagcgtcgcaggggtcacagatctgggggaacccggg ggaaaagcactgaggcaaacccgctcgtctcctcaatataatggagggg gaggtgagtagtcttgattactataagaccttttttcttccgggagcaaa accgtgagctggattataatcgcctataaagctccagaggcggcagggacct gcagaggagcccccgctcgcgactagctgccccgcgagcaacggcct cgtgatttccccgcatccggtccccgcctccccactctgccccgcctacccc ggagccgtgcagccgctctccgaatctctctctctctgctcgcgtcgcgga gagggaaactagcgagaacgaggaagcagctggaggtgacgccgggagatt acgctgtcagggccgagccgagcggatcctgggcctgtgcagaggaaa ggccggagtgcccggctcgtgtcgcagagccaggtgggtaagctagcgac cacctggacttcccagcgcacaaccgtggctttcagccaggtcctctcctccc cggcttctcaaccaaccateccagcgcggccaccaactcccgaatga gtgcttctccccagcagccgaaggcgtactaggaacgtaacctgtacttt tccagggccgtagtcgaccgctccccaggtgctgtgactgcccgcg gggctagagtgcaaggtgactgtggttctctctgccaagtccaggggagaac gtaaagataatggccttttccccctctcacctgtctaccaaaagtcctagccc cggagcagtagcctttttccaggaattagccagacacaacaacgggaac cagacaccgaaccagacatgccgccccgtgcccctccccgctcgtgctttt cctccctctgtctccagagccgattctcaaggggacctccgtgccccgg ctgctcagctcctcgggtgagcagaccgggaagtctccccgcacagctctc gcttcttttgagcctgtttctgcccggaccagtcgaggactctggacagtaga ggccccggagcagcagctgGAATTCGCCACCATGGCCCC</p>	<p>70</p>

	AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTCACAATCAGGGAATCTGA CTGAGCATCAACGAACTCATACCGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAAGTGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACACTGGGAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCGACGC GCTGGACGATTTTCGATCTCGACATGCTGGGTCT GATGCCCTCGATGACTTTGACCTGGATATGTTGG GAAGCGACGCATTGGATGACTTTGATCTGGACAT GCTCGGCTCCGATGCTCTGGACGATTTTCGATCTC GATATGTTATAAAAGCTTGGGTGGCATCCCTGTG ACCCCTCCCAGTGCCTCTCCTGGCCCTGGAAGT TGCCACTCCAGTGCCACCAGCCTTGTCCTAATA AAATTAAGTTGCATCATTTTGTCTGACTAGGTGT CCTTCTATAATATTATGGGGTGGAGGGGGTGGT ATGGAGCAAGGGGCAAGTTGGGAAGACAACCTG TAGGGCCTGCGGGTCTATTGGGAACCAAGCTG GAGTGCAGTGGCACAATCTTGGCTCACTGCAATC TCCGCTCCTGGGTTCAAGCGATTCTCCTGCCTC AGCCTCCCAGTTGTTGGGATTCCAGGCATGCAT GACCAGGCTCAGCTAATTTTTGTTTTTTTGGTAG AGACGGGGTTTTACCATATTGGCCAGGCTGGTCT CCAACTCCTAATCTCAGGTGATCTACCCACCTTG GCCTCCCAAATTGCTGGGATTACAGGCGTGAACC ACTGCTCCCTTCCCTGTCCTT	
26 (RE + кодирующая)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatatttgaagtctgttgacataaagaatcatgatattaatgccatggaaatg aaagggcgatcaactatggttgaaggggaaattgtaggcacagatgt gttcgtgagcagtgctgtcttagcaatactagagaagagagagaacaatg aaattctgattggcccagtgtagcccagatgaggttcagctgccaactttctct tcacatctatgaaagtcatttaagcacaactactttttttttttttttttttttgagaca gagtcttctctgttggccaggacagagtgtagtgactcaatctcggctcact gcagcctccactcctaggtcaaacggctcctcctgcatcagcctccaagtagc tggaaattacaggagtgcccaccatgccagctaattttgtatttttaatagatacg ggggttcccatatcaccaggctggtctcgaactcctgcctcaagtgatcca cctgcctcggcctccaaagtgtgggattatagcgctcagccactatgcccaac ccgaccaactttttaaataaatatttaaaaattggtatttcacatatatactagta	71

	<p>tttacattatccacacaaaacggacggcctccgctgaaccagtgaggcccca gacgtgcgcataaataacccctgctgctgcaccacctggggagagggggag gaccacggtaaatggagcgcgcatagcaaaaggacgcggggtcctttct ctgccggtggcactgggtagctgtggccaggtgtggtactttgatgggcccag ggctggagctcaaggaagcgtcgcagggtcacagatctggggaacccggg ggaaaagcactgaggcaaacccgctcgtctcctacaatatgggagggg gaggtgagtagcttgattactataagacctttttttcctccggcgcaaa accgtgagctggattataatcgcctataaagctccagaggcgtcaggcacct gcagaggagccccgctcctccgactagctccccgcgagcaacggcct cgtgattcccccgatccggtccccgctccccactctccccgcctacccc ggagccgtgcagccgctctccgaatctctctctctctggcgtcgcgtgcga gagggaaactagcgagaacgaggaagcagctggaggtgacgccggcagatt acgctgtcaggggccgagccgagcggatcgtgggcgtgtgcagagggaaa ggcgggagtgccggctcgtctgcagagccgaggtgggtaagctagcgac cacctggactcccagcggcaaccgtggctttcagccaggtcctctctccc cggcttcaaccaacccatcccagcggcggccaccaacctcccgaatga gtgcttctgccccagcagccgaaggcgtactaggaacgtaacctgtacttt tcaggggctgtagtcgaccgctgcccagttgctgtgcgactgcgcgcg gggctagagtgaaggtgactgtggttctctgccaagtccgaggggaaac gtaaagatatggccttttcccctctcaccttgtctaccaaagtcctagtccc cggagcagttagccttttctccaggaattagccagacacacacgggaac cagacaccgaaccagacatgcccggcctgcccctcccgcctcgtccttt cctcctctgtctcagagccggatctcaaggggagcctccgtgccccgg ctgctcagtcctccggtgtgcaggaccgggaagtccccgcacagctctc gcttctttgcagcctgtttctgcccggaccagtcgaggactctggacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTACAATCAGGGAATCTGA CTGAGCATCAACGAACCTCATAACGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACTGGGAAAAAAA ACGAGTAAAAGGCCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCGACGC GCTGGACGATTTTCGATCTCGACATGCTGGGTTCT GATGCCCTCGATGACTTTGACCTGGATATGTTGG GAAGCGACGCATTGGATGACTTTGATCTGGACAT GCTCGGCTCCGATGCTCTGGACGATTTTCGATCTC GATATGTTATAAaaagagaccggttactgtgacagtaaaagagac</p>	
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

	<p>cggttactgtgagaatgaaagagaccggttactgtgatcggaaaagagaccg gttactgtgagcggcctgaaaccagcagacaatgtagctcagtagaaacc agcagacaatgtagctaatggaaccagcagacaatgtagcttcggagaaa cccagcagacaatgtagctAAGCTTGGGTGGCATCCCTGTG ACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGT TGCCACTCCAGTGCCACCAGCCTTGTCTAATA AAATTAAGTTGCATCATTTTGTCTGACTAGGTGT CCTTCTATAATATTATGGGGTGGAGGGGGTGGT ATGGAGCAAGGGCAAGTTGGAAGACAACCTG TAGGGCTGCGGGTCTATTGGAACCAAGCTG GAGTGCAGTGGCACAATCTGGCTCACTGCAATC TCCGCTCCTGGGTTCAAGCGATTCTCCTGCCTC AGCCTCCCGAGTTGTTGGGATTCCAGGCATGCAT GACCAGGCTCAGCTAATTTTGTTTTTTGGTAG AGACGGGGTTTACCATATTGGCCAGGCTGGTCT CCAACCTAATCTCAGGTGATCTACCCACCTTG GCCTCCCAAATTGCTGGGATTACAGGCGTGAACC ACTGCTCCCTCCCTGTCTT</p>	
40 (кодирующая)	<p>ATGGCCGAGATCACCTGATGCTGGCTGAAGGCT ACAGACTGGTGCAGCGCCTCCATCTGCCGCTGC CGCCACGGCCCCACGCCCTGAGAACACTGCC CCCTACGCCGGCCCTGGTCTTGATAGCGGACTCA GACCTAGAGGGCGCCCTCTGGGCCCTCCACCTCC AAGACAGCCTGGAGCCCTGGCCTACGGCGCCTTC GGCCCTCCTTCTAGCTTCCAGCCCTCCCCGCCGT GCCTCCTCCAGCGCTGGCATCGCCACCTGCAG CCTGTGGCCACCCCTTACCCGGAAGAGCCGCCG CCCCTCAAACGCCCTGGCGGACCTCCTGGCCC CCAGCCTGCTCAAGCGCCGCTGCCCTCCACCT CCTGCTCATGCCCTGGGCGGCATGGACGCCGAGC TGATCGACGAGGAAGCCCTGACCAGCCTGGAAC TGGAAGTGGCCTGCACAGAGTGCAGGAACTGC CTGAGCTGTTCTGGGACAGAGCGAGTTCGACTG CTTCAGCGACCTGGGCAGCGCCCTCCTGCCGGC TCTGTGTCTGCgccgaccactgatgetcggagggetaccgct ggtgcagaggccgccgtccgccgccgccatggccctcatgctccgg actctgccccgtacgcccggccggcctggacagtgggctgaggccggg ggggctccgctggggccgccgcccccggcaaccggggccctggcgtac ggggcctcgggcccggcctcctccagcccttccggccgtgctccgccg gccgcccggcatcgcgacactgcagcctgtggcgacgccgtacccggccgc gcCgccgcccccaacgctcgggaggccccggcccgcagccggc cccaagcggcagccccgccgccgccgcgacgccctggggcgcagga cgccgaactcatcagcagggaggcgtgacgtgctggagctggagctggg ctgcaccgctgcgagctgcccagctgtcctggccagagcagttcga ctgcttctcgactggggtccgcgccgccggcctccgtgagctgcccagtc ccagctcataaaccagccgatgcaagtacccaaccggccagcaag acgccccccacgaacgcccttacgctgcccagtgagtcctgtgatgccgc ttctccCGCAGCGACAACCTGGTGAGAcacatccgatccac acaggccagaagcccttcagtgccgatctgcatgAGAAactcagcCG AGAGGATAACTTGCACACTcacatccgaccacacagggc</p>	72

	aaaagcccttcgctgcacatctgtggaagaaagtttccCGGAGCGA TGAACCTGTCCGAcataccaagatccacttgcggcagaaggaccgc ccttacgcttcccagtgaggctctgtgatcgccgttctccCAATCAGG GAATCTGACTGAGcacaatccgcatccacagggcagaagcccttc cagtccgcatctgcatGAAaacttcagcACAAGTGGACATCT GGTACGCcacaatccgacccacagggcgaagcccttcgctgcgac atctgtggaagaaagtttccCAGAATAGTACCCTGACCGAA cataccaagatccacttgcggcagaaggacaag	
41 (кодирующая)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCT ACAGACTGGTGCAGCGGCCCTCCATCTGCCGCTGC CGCCACGGCCCCACGCCCTGAGAACACTGCC CCTACGCCGGCCCTGGTCTTGATAGCGGACTCA GACCTAGAGGGCGCCCTCTGGGCCCTCCACCTCC AAGACAGCCTGGAGCCCTGGCCTACGGCGCCTTC GGCCCTCCTTCTAGCTTCCAGCCCTTCCCCGCCGT GCCTCCTCAGCTGCTGGCATCGCCACCTGCAG CCTGTGGCCACCCCTTACCCCGAAGAGCCGCCG CCCCTCCAAACGCCCTGGCGGACCTCCTGGCCC CCAGCCTGCTCCAAGCGCCGCTGCCCTCCACCT CCTGCTCATGCCCTGGGCGGCATGGACGCCGAGC TGATCGACGAGGAAGCCCTGACCAGCCTGGAAC TGAACTGGGCTGCACAGAGTGCGGGAAGTGC CTGAGCTGTTCTGGGACAGAGCGAGTTCGACTG CTTCAGCGACCTGGGCAGCGCCCTCCTGCCGGC TCTGTCTCCTGCGGCGCAGCGGCGGCGGAAGC GGCgcccaccactgatgctcgcggaggctaccgctggtgcagaggcc gccgtccgccgcccccacatggccctatgctcctccgactctgccgccgt acgcccggcccggcctggacagtgggctgaggccgggggctccgctgg ggccgcccccgcccaaccggggcctggcgtacggggccttcgggc cgccgtcctctccagcccttccggcgtgctccgccggccgggcatcg cgacactgcagcctgtggcgacgccgtacccggccgcccggccgccc ccaacgctccgggaggccccgggcccgcagccggcccgaagcggcga gccccgccgcccgcgacgcccctgggggcatggacgcccgaactc gacgaggaggcgtgacgtcgtgagctggagctgggctgcaccgctgc gcgagctgcccagctgtcctggccagagcgagttcactgcttctggact ggggtccgcccggccggcctcgtgagctgcggtggttctggtggtggt ctggtcagctccagctcaaaaccagccgcatgcgcaagtacccaaccgg cccgcaagagccccccacgaagccccttacgcttcccagtgagctctgt gatcgccgttctccCGCAGCGACAACCTGGTGAGAcacatc cgcacacacagggcagaagccctccagtgcccatctgcatGAAaact tcagcCGAGAGGATAACTTGCACACTcacaatccgacccac acagggcaagcccttcgctgcacatctgtggaagaaagtttccCGG AGCGATGAACTTGTCCGAcataccaagatccacttgcggcaga aggaccgcccctacgcttcccagtgaggctctgtgatgccgcttctccCAA TCAGGGAATCTGACTGAGcacaatccgcatccacagggccag aagccctccagtgccgcatctgcatGAAaacttcagcACAAGTGG ACATCTGGTACGCcacaatccgacccacacagggcaagcccttc gcctgcacatctgtggaagaaagtttccCAGAATAGTACCCTG ACCGAAcataccaagatccacttgcggcagaaggacaag	73

<p>42 (RE + кодирующая)</p>	<p>ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacatatttgaagtctgttgacataaagaatcatgatattaatgccatggaaatg aaaggcgcgatcaacactatggttgaagggggaaattgtagcacagatg gtcgtgtgacagtggtgctcttagcaatactagagaagagagagaacaatg aaattctgattggccccagtgtagcccagatgaggttcagctgccaaactttctt tcacatctatgaaagtcaatlaagcacaactactttttttttttttttttttttttgagaca gagcttctctgtgtgccaggacagagtgtagtgactcaatctcggctcact gcagcctccacctctaggtcaaacggctcctctgcatcagcctcccaagtagc tggaattacaggagtggtgcccaccatgcccagctaattttgtatttttaagatacg gggggttcaccatataccaggctggctcgaactcctggcctcaagtatcca cctgctcggcctccaaagtgtgggattataggcgtcagccactatgcccaac ccgaccaacttttttaataaataatttaaaaaattggtatttcacatatatactagta ttacatttatccacaaaaacggacgggctcctcgtgaaccagtgaggcccca gacgtgcgataaataaccctgctgctgcaccacctggggagagggggag gaccacgtaaatggagcgcagcatagcaaaaggacgcggggtccttttct ctgccgtggcactgggtagctgtggccaggtgtgtactttgatggggccag ggctggagctcaaggaagcgtcgcagggtcacagatctgggggaaccccg ggaaaagcactgaggcaaacggcgtcgtctcctacaatatatgggagggg gaggtgagtagctgtggtactcataagacctttttttctcctccgggcgcaaa accgtgagctggattataatcgccataaagctccagaggcgggtcaggcacct gcagaggagccccgctcctcgcgactagctcccccgagacaacggcct cgtgattccccccgatccggtccccgctccccactctcccccgctacccc ggagccgtcagcgcctctccgaatctctctctctctctggcgtcgcgtgcga gagggaaactagcgagaacgaggaagcagctggaggtgacgccgggcagatt acgcctgtcagggccgagccgagcggatcgtgggcgtgtgcagaggaaa ggcgggagtgcccggctcgtctgcagagccgaggtgggtaagctagcgac cacctggactccagcgccaaccgtggcttttcagccaggtcctctctctccg cggttctcaaccaacccatcccagcgcggccaccaactcccgaatga gtgcttctgccccagcagccgaaggcgtactaggaacggtaacctgttacttt tccaggggcccgtatgcacccgtcggcaggttgctgtgcgactgcccgcgcg gggctagagtgcaaggtgactgtgttctctctgccaagtccgaggagaac gtaaagatatggccttttccccctctcacctgtctcaccaaagtcctagtccc cggagcagttagcctttttccagggaattagccagacacaacaacgggaac cagacaccgaccagacatgcccggcctgcccctccccgctcgtgctcttt cctccctctgtctctccagagccgatcttcaaggggagcctccgtgccccgg ctgctcagtcctcgggtgtgcaggaccccgaagtctccccgcacagctctc gcttctcttgacagcctgttctgcccggaccagtcgaggactctgacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCCC AAAGAAGAAGCGGAAGGTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGCAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCGAATGTGGTAAAAGTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTACAATCAGGGAATCTGA CTGAGCATCAACGAACATACCGGGGAAAAAC</p>	<p>74</p>
---------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------

	<p>CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACTGGGAAAAAAA ACGAGTAAAAGGCCGGCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGACGCGC TGGACGATTTGATCTCGACATGCTGGGTTCTGA TGCCCTCGATGACTTTGACCTGGATATGTTGGGA AGCGACGCATTGGATGACTTTGATCTGGACATGC TCGGCTCCGATGCTCTGGACGATTTGATCTCGA TATGTTATAAACTAGTGAAACCCAGCAGACAAT GTAGCTAGACCCAGTAGCCAGATGTAGCTAAAG AGACCGGTTCACTGTGAAAGCTTGGGTGGCATCC CTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTG GAAGTTGCCACTCCAGTGCCACCAGCCTTGTC TAATAAAATTAAGTTGCATCATTTTGTCTGACTA GGTGTCTTCTATAATATTATGGGGTGGAGGGGG GTGGTATGGAGCAAGGGGCAAGTTGGGAAGACA ACCTGTAGGGCCTGCGGGGTCTATTGGGAACCA AGCTGGAGTGCAGTGGCACAATCTTGGCTCACTG CAATCTCCGCCTCCTGGGTCAAGCGATTCTCCT GCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCA TGCATGACCAGGCTCAGCTAATTTTTGTTTTTTG GTAGAGACGGGGTTTACCATATTGGCCAGGCTG GTCTCCAACCTCTAATCTCAGGTGATCTACCCAC CTTGGCCTCCCAATTGCTGGGATTACAGGCGTG AACCACTGCTCCCTTCCCTGTCTT</p>	
<p>43 (RE + кодирующая)</p>	<p>ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggt aacataattgaagtctgttgacataaagaatcatgatattaatgccatggaaatg aaagggcgatcaacactatggttgaagggggaattgtagagcacagatgt gttcgtgtggcagtgtgctgtcttagcaatactcagagaagagagagaacaatg aaattctgattggcccagtgtagcccagatgaggttcagctccaactttctctt tcacatctatgaaagtcatfaagcacaactaactttttttttttttttttttttgagaca gagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcact gcagcctccacctcctaggctcaaacggctcctcctgcatcagcctccaagtagc tggattacaggagtgggccaccatgccagctaatgtttttatattatagatac ggggttcaccatataccagcgtggtctcgaactcctggcctcaagtatcca cctgcctcggcctccaaaagtgtgggattataggcgtcagccactatgccaac ccgaccaacttttttaataataatfttaaaaaattggtatttcacatataactagta tttacattatccacaaaaacggacggcctccgctgaaccagtgaggcccca gacgtcgcataaataaccctcgtgctgcaccacctggggagagggggag gaccacggtaatggagcgcgcgatagcaaaaggacgcggggtcctttct ctgccgtggcactgggtagctgtggccaggtgtgtactttgatggggcccag ggctggagctcaaggagcgtcgcagggtcacagatctgggggaacccgg ggaaaagcactgaggcaaaccccgctcgtctcctacaatatatgggagggg gagggtgagtacttctgattactataagacctttttttcctccggcgcaaa accgtgagctgattataatcgcctataaagctccagaggcggcagccact gcagaggagccccgcccctccgagtagctgccccgcgagcaacggcct</p>	<p>75</p>

	<p> cgtgatttccccgccgatccggtccccgcctccccactctgccccgcctaccc ggagccgtgcagccgcctctccgaatctctctctctctgcccgcctgcgctgcga gagggaactagcgagaacgaggaagcagctggaggtgacgcccgggcagatt acgcctgtcagggccgagccgagcggatcgtgggcgctgtgcagagggaaa ggcgggagtgccccggctcgtctgcagagccgaggtgggtaagctagcgcac cacctggactcccagcggccaaccgtggctttcagccaggtcctctctctccc cggcttctcaaccaaccatcccagcggccggccaccaacctcccgaatga gtgcttcgccccagcagccgaaggcgctactaggaacggtaacctgtacttt tccaggggcccgtagtcgaccgctcccaggttgcgtgcgactgcgcgcgcg gggctagagtcaaggtgactgtggttctctctgccaagtccgaggggagaac gtaaagatatggcctttttcccctctcaccttgtctcaccaaagtcctagtccc cggagcagttagcctcttttccagggaaatagccagacacaacaacgggaac cagacaccgaaccagacatgcccggccgtgcccctcccgcctcgtgccttt cctccctctgtctctccagagccggatctcaaggggagcctccgtgccccgg ctgctcagtcctccggtgtgcaggaccgggaagtctcccgcacagctctc gcttctttgcagcctgtttctgcccggaccagtcgaggacttgacagtaga ggccccgggacgaccgagctgGAATTTCGCCACCATGGCCCC AAAGAAGAAAGCGGAAGGTTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTA CAAATGTCCTGAATGTGGGAAATCATTAGTCGC AGCGACAACCTGGTGAGACATCAACGCACCCAT ACAGGAGAAAAACCTTATAAATGTCCAGAATGT GGAAAGTCCTTCTCACGAGAGGATAACTTGAC ACTCATCAACGAACACATACTGGTGAAAAACCA TACAAGTGTCCGAATGTGGTAAAAGTTTTAGCC GGAGCGATGAACTTGTCCGACACCAACGAACCC ATACAGGCGAGAAGCCTTACAAATGTCCCGAGT GTGGCAAGAGCTTCTACAATCAGGGAATCTGA CTGAGCATCAACGAACACTACACGGGGAAAAAC CTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTTC CACAAGTGGACATCTGGTACGCCACCAGAGGAC ACATACAGGGGAGAAGCCCTACAAATGCCCCGA ATGCGGTAAAAGTTTCTCTCAGAATAGTACCCTG ACCGAACACCAGCGAACACACTGGGAAAAAAA ACGAGTAAAAGGCCGGCCACGAAAAAGGCC GGCCAGGCAAAAAAGAAAAAGGGATCCTACCCA TACGACGTACCAGATTACGCTCTCGAGGACGCGC TGGACGATTTTCGATCTCGACATGCTGGGTTCTGA TGCCCTCGATGACTTTGACCTGGATATGTTGGGA AGCGACGCATTGGATGACTTTGATCTGGACATGC TCGGCTCCGATGCTCTGGACGATTTTCGATCTCGA TATGTTATAAACTAGTGAAACCCAGCAGACAAT GTAGCTAGACCCAGTAGCCAGATGTAGCTAAAG AGACCGGTTCACTGTGAGAAACCCAGCAGACAA TGTAGCTAGACCCAGTAGCCAGATGTAGCTAAA GAGACCGGTTCACTGTGAAAGCTTGGGTGGC CCTGTGACCCCTCCCAGTGCCTCTCCTGGCCCT GGAAGTTGCCACTCCAGTGCCACCCAGCCTTGT CTAATAAAATTAAGTTGCATCATTTTGTCTGACT AGGTGTCCTTCTATAATATTATGGGGTGGAGGGG </p>	
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

	<p>GGTGGTATGGAGCAAGGGGCAAGTTGGGAAGAC AACCTGTAGGGCCTGCGGGTCTATTGGGAACC AAGCTGGAGTGCAGTGGCACAATCTTGGCTCACT GCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCC TGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGC ATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTT GGTAGAGACGGGGTTTACCATATTGGCCAGGCT GGTCTCCAACCTCCTAATCTCAGGTGATCTACCCA CCTTGGCCTCCCAAATTGCTGGGATTACAGGCGT GAACCACTGCTCCCTTCCCTGTCCTT</p>	
<p>44 (RE + кодирующая)</p>	<p>ggaggaagccatcaactaaactacaatgactgtaagatacaaatgggaatggt aacatatttgaagtctgttgacataaagaatcatgatattaatgccatggaatg aaagggcgaatcaacactatggttgaagggggaaattgtagagcacagatgt gttcggtggcagtgctgctctagcaatactagagaagagagagaacaatg aaattctgattggccccagtgtagcccagatgaggtcagctgccaaactttctt tcacatctatgaaagtcaattaaagcacaactacttttttttttttttttggaca gagcttgctctgtggccagacagagtgagtagtgactcaatctcggtcact gcagcctccactcctagggctcaaacggctcctcctcatcagctcccaaagtgc tgaattacaggagtgcccaccatgccagctaattttgtattttaaagatacag ggggttcacatatacaccaggctggctcgaactcctggcctcaagtatcca cctgcctggcctccaaagtgtgggattataggcgcagccactatgcccac ccgaccaactttttaaaataaataattaaaaaattggtatttcacataatactagta ttacattatcacacaaaacggacgggctccgctgaaccagtgagggccca gacgtgcgataaataaccctgctgctgacacactggggagagggggag gaccacggtaaatggagcagcgcatagcaaaagggagcgggtccttttct ctgccgtggcactgggtagctggtgcccaggtgtgtactttgatggggccag ggctggagctcaaggaagcgtcgcaggggtcacagatctgggggaacccgg ggaaaagcactgaggcaaacggcgtctgctcctacaatataatgggagggg gaggtgagtagcttgattactataagaccttttttccctcggcgcaaa accgtgagctggattataatgccctataaagctccagaggggtcagccact gcagaggagcccccgctcgcgactagctgccccgcgagcaacggcct cgtgatttccccccgatcgggtccccgcctccccactctgcccgctacccc ggagccgtgcagccgctcctccaatctctcttctcctggcgctgcgtgcga gagggaaactagcagaacgaggaagcagctggaggtgacccggcgagatt acgctgtcagggccagccagcggatcgtgggcctgtgcagaggaaa ggcgggagtgcccggctcgtgtcgcagagccaggtggttaagctagcgcac cacctggactcccagcgcacaaccgtggcttttagccaggtcctctcctccc cggtctcaaccaacccatcccagcggccacccaactccgaaatga gtcttctgccccagcagccgaaggcgctactaggaacgtaactgttacttt tccagggccgtagtcgaccgctgcccaggtgctgctgcgactcgcgcg gggctagtgcaaggtgactgtgttctctggccaagtccgaggggagaa gtaagatgatggccttttcccctctcactgtctcaccaaagtccctagtccc cggagcagtagcctttttccaggaattagccagacacaacaacgggaac cagacaccgaaccagacatgcccgccccgctgcctccccgctgcctttt cctccctctgctctccagagccgatctcaaggggacctcctgccccgg ctgctcagtcctccggtgtaggaccccggaaagtccccgcacagctctc gcttcttgcagctgttctgcggggaccagtcgaggactctggacagtaga ggccccggacgaccgagctgGAATTCGCCACCATGGCCgc cgaccactgatgctcggagggctaccgctggtgcagagggcccgctcc gccggccgccccatggcctcatgcctccgactctgccccgtacgggg</p>	<p>76</p>

	<p>cccgggctggacagtgggctgaggccgsgggggctccgctgggcccgc gcccggcccgaacccgggcccctggcgtaggggcttcggcgccgctc ctcctccagcccttccggccgtgcctccgccggcgccggcgcacatcgcgacac gcagcctgtggcgacgcccgtaccggccgagccggccgcccccaagc tccgggagggccccggcccgcagccggcccaagcgccgcagccccgc cgccggccgagccgacccctgggcccagcggcccccaagcgccgcagccccgc aggcgtgacgctcgtggagctggagctggggctgaccggcgtgcgagct gccgagctgtcctggccagagcagttcgactgcttctggactggggtc cgcgcccggctccgtgagctcggggtgtgtgtgtgtgtgtgtgtgtgt gtcccagctcataaaccagccgtgcaagtacccaaccggcccagca agaccccccaacgccttacgcttcccagtgagctcgtgatcgcc gtttctccCGCAGCGACAACCTGGTGAGAcacatccgcatcc acacaggccagaagccctccagtgccgcatctgcatGAAaactcagcC GAGAGGATAACTTGCACACTcacatccgacccacacagc gaaaagcccttcgctgacatctgtggaagaagtccCGGAGCG ATGAACTTGTCAGAcataccaagtcacttgcggcagaagc gcccttacgcttcccagtgagctcgtgatcgccgcttcccAATCAG GGAATCTGACTGAGAcacatccgcatccacagccagaagcct tccagtgccgcatctgcatGAAaactcagcACAAGTGGACATC TGGTACGCCacatccgacccacacagccgaaaagcccttcgctgcg acatctgtggaagaagtccCAGAATAGTACCCTGACCG AAcataccaagtcacttgcggcagaaggacaagtaaCTCGAGGA AACCCAGCAGACAATGTAGCTAGACCCAGTAGC CAGATGTAGCTAAAGAGACCGGTTCACTGTGAA AGCTTGGGTGGCATCCCTGTGACCCCTCCCCAGT GCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTG CCCACCAGCCTTGTCTAATAAAAATTAAGTTGCA TCATTTTGTCTGACTAGGTGTCTTCTATAAATATT ATGGGGTGGAGGGGGTGGTATGGAGCAAGGGG CAAGTTGGGAAGACAACCTGTAGGGCCTGCGGG GTCTATTGGGAACCAAGCTGGAGTGCAGTGGCA CAATCTTGGCTCACTGCAATCTCCGCCTCCTGGG TTCAAGCGATTCTCCTGCCTCAGCCTCCCCGATT GTTGGGATTCCAGGCATGCATGACCAGGCTCAGC TAATTTTTGTTTTTTTGGTAGAGACGGGGTTTTCAC CATATTGGCCAGGCTGGTCTCCAACCTTAATCT CAGGTGATCTACCCACCTTGGCCTCCCCAATTGC TGGGATTACAGGCGTGAACCACTGCTCCCTTCCC TGTCTT</p>	
<p>45 (RE + кодирующая)</p>	<p>ggaggaagccatcaactaaactacaatgactgtaagatacaaatgggaatggt aacatatttgaagtctgtgacataaagaatcatgatattaatgccatggaaatg aaaggcagatcaactatggttgaagaagggaagttagagcacagatgt gttgctgtgcagtgctgtcttagcaatactagagaagagagagaacaatg aaattctgattggcccagtgtagccccagatgaggttcagctgccaaacttctct tcacatcttagaaagtcatttaagcacaactacttttttttttttttttttttgagaca gagcttctgtgtgccaggacagagtgagtagtgactcaatctcggctcact gcagcctccacctctaggctcaaacggtcctcctgcatcagcctccaagtagc tggaaftacaggagtggcccaccatgccagctaattttgtattttaatagatacg gggtttacatatacaccagctggtctcgaactcctggcctcaagtatcca cctgcctcggcctccaaagtctgggattataggcgtcagccactatgcccac</p>	<p>184</p>

	<p>ccgaccaacctttttaaataaatatftaaaaaattggtatttcacatatatactagta ttacatttatccacacaaaacggacgggctccgctgaaccagtgaggcccca gacgtgcgcataaataaccctgctgctgaccacctggggagagggggag gaccacggtaaatggagcgcgcatagcaaaaggagcgcggggtccttttct ctgccgtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaaagcgtcgcagggtcacagatctgggggaacccgg ggaaaagcactgaggcaaacccgctcgtctcctaataatagggaagggg gaggttgagtacgttctggattactataagaccttttttctccggcgcaaa accgtgagctgattataatcgcctataaagctccagagcgggtcaggcacct gcagaggagccccggctcggcgaactagctgccccggcagcaacggcct cgtgatttccccggatccggtccccgctccccactctgccccgctacccc ggagccgtgcagccgctctccgaatctctctctctgctgctcgcgtgcga gagggaaactagcgagaacgaggaagcagctggaggtgacggggcagatt acgctgtcagggccgagccgagcggatcgtggcgctgtgcagaggaaa ggcgggagtgcccggctcgtctgcagagccgaggtgggtaagctagcgac cactggactcccagcggcaaccgtggctttcagccaggtcctctctccc cggcttcaaccaaccatcccagcggccggcaccacactcccgaatga gtgcttctgccccagcagccgaaggcgtactaggaacgtaacctgtacttt tcaggggccgtagtgcaccgctcccaggtgctgctgactgcgcgcg gggctagagtcaagggtactgtggttctctgccaagtccgagggagaac gtaaagatatggccttttcccctctcacctgtctaccaaagtcctagtccc cggagcagttagccttttcttccagggaatagccagacacaacaagggaac cagacaccgaaccagacatgcccggccgtgcccctcccgcctgctgcttt cctcccttctgctctccagagccgatctcaaggggagcctccgtgccccgg ctgctcagtcctccggtgtgcaggaccggaaagtctcccgcacagctctc gcttctttgcagcctgttctgcccggaccagtcgaggacttgacagtaga ggccccgggacgaccgagctgGAATTCGCCACCATGGCCgc cgaccacctgatgctcggcagggtaccgctggtgcagaggccgcccctc gcccggccccatggccctcatgctccggactctgcccgtacgcccggg cccggcctggacagtgggctgaggccgcccggggctccgctggggccgccc gcccggcccaaccggggcctggcgtacggggccttcgggcccggcgtc ctctccagcccttccggcctgctcccggcccggcggcatcgcgacct gcagcctgtggcgacccgtaccccggcggcggccgcccccaacgc tccgggagggccccggcccagcccggcccaagcggccgagccccgc cgccggccgacgcccggggcgtgagccgaaactcatcgacgagg aggcgtgacgtcgtggagctggagctggggtgcaccgctgcccgagct gcccagctgttctggccagagcagttcgaactgttctcgactgggggtc cgcggccccggcctccgtgagctgcggtggttctggtggtgttctggtca gtcccagctcaaaaccagccgatgcaagtacccaaccggcccagca agacggcccccaacgccccttacgcttggccagtgaggctctgtgacgcc gcttctccCGCAGCGACAACCTGGTGAGAcacatccgcatcc acacaggccagaaacccttccagtcgcatctgcatGAAaacttcagcC GAGAGGATAACTTGACACTTcacatccgcaccacacagggc gaaaagcccttcgctgacatctgtggaagaagtgtgcccCGGAGCG ATGAACTTGTCCGAcataccaagatcacttggcggcagaaggacc gcccttacgcttcccagtgagctctgtatcggccttctccCAATCAG GGAATCTGACTGAGcacatccgcatccacagggcagaagccct tccagtcggcatctgcatGAAaacttcagcACAAGTGGACATC TGGTACGCcacatccgacccacacagggcaaaagcccttcgctgag acatctgtggaagaagtgtgcccCAGAATAGTACCCTGACCG</p>	
--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

	<p>AAcataccaagatccacttgcggcagaaggacaagtaaCTCGAGGA AACCCAGCAGACAATGTAGCTAGACCCAGTAGC CAGATGTAGCTAAAGAGACCGGTTCACTGTGAG AAACCCAGCAGACAATGTAGCTAGACCCAGTAG CCAGATGTAGCTAAAGAGACCGGTTCACTGTGA AAGCTTGGGTGGCATCCCTGTGACCCCTCCCAG TGCTCTCCTGGCCCTGGAAGTTGCCACTCCAGT GCCCACCAGCCTTGTCTTAATAAAATTAAGTTGC ATCATTTTTGTCTGACTAGGTGTCTTCTATAATAT TATGGGGTGGAGGGGGGTTGGTATGGAGCAAGGG GCAAGTTGGGAAGACAACCTGTAGGGCCTGCGG GGTCTATTGGGAACCAAGCTGGAGTGCAGTGGC ACAATCTTGGCTCACTGCAATCTCCGCCTCCTGG GTTCAAGCGATTCTCCTGCCCTCAGCCTCCCGAGT TGTTGGGATTCCAGGCATGCATGACCAGGCTCAG СТААТТТТТГТТТТТТТГГТАGAGACGGGGTТTCA CCАТАТTGGCCAGGCTGGTCTCCAАCTCCTААTС TCAGGTGATCTACCCACCTTGGCCTCCCAAATG CTGGGATTACAGGCGTGAACCACTGCTCCCTTCC CTGTCTT</p>	
<p>8 (кодирующая)</p>	<p>ATGGCCgcccaccacctgatgctcgcggagggtaccgctggtgcaga ggccgcctccgcccggcccggcccatggccctcatcgctccggactctgccg ccgtacgcggcccgccctggacagtgggctgaggccgccccgggctccg ctggggccgcccggcccgaaccggggccctggcgtagcggggccttc ggcccgctctcctccagcccttccggccgtgcctccggcgccgcccg catcgcgacctgcagcctgtggcgaccggtacccccggcgccgcccgcg ccccccaacctccgggagggccccggccgagccggcccccaagcgc gcagccccgcgcccggcgcacgcccggggcgcatggacgccgaactc atcgacgaggagcgctgacctgctggagctggagctgggctgcaccg tgcgcgagtccccgagctgctggcggccagagcgagttcgactgcttctcg acttgggtccgcccggccggtccgctgagctgggtggtgctggtg gttctgctagctccagctcaaaaccagccgcatgcgcaagtacccaacc ggcccagcaagacccccccagcaacgccccttacgcttcccagtgaggctc tgtgatcgccccttcccCGCAGCGACAACCTGGTGAGAcac atccgcatccacacagccagaagcccttccagtggcgcctctgatgAGAA acttcagcCGAGAGGATAACTTGCACACTccatccgcacc cacacaggcgaagccctcgcctgacatctgtggaagaaagtggccCG GAGCGATGAACTTGTCCGAcataccaagatccacttgcggca gaaggaccgcccttacgcttcccagtgaggctctgtatgccgcttctccCA ATCAGGGAATCTGACTGAGcacatcccatccacacagggca gaagccctccagtggcgatctgatgAGAAacttcagcCAAGTG GACATCTGGTACGCcacatccgaccacacaggcgaagccct tcgctgcgacatctgtggaagaaagtggcCAGAATAGTACCCT GACCGAAcataccaagatccacttgcggcagaaggacaag</p>	<p>203</p>
<p>46 (кодирующая)</p>	<p>ATGGCCgcccaccacctgatgctcgcggagggtaccgctggtgcaga ggccgcctccgcccggcccggcccatggccctcatcgctccggactctgccg ccgtacgcggcccgccctggacagtgggctgaggccgccccgggctccg ctggggccgcccggcccgaaccggggccctggcgtagcggggccttc ggcccgctctcctccagcccttccggccgtgcctccggcgccgggg catcgcgacctgcagcctgtggcgaccggtacccccggcgccgcccgcg</p>	<p>204</p>

	<p>cccccaacgctccgggaggccccgggcccgcagccggccccaaagcgc gcagccccgcccgcggccgcacgccctggcgcatggacgccgaactc atcgacgaggaggcgtgactcgctggagctggagctgggctgcaccgcg tgcgagctgcccagctgttctgggccaagcagcttgcactgcttctcg actggggtccgcccggccggctccgtgagctgcggtggttctggtggtg gttctggtGGTGGCAGCGGGGAGGTTCTGGTcagtcca gctcataaaaccagcccatgcaagtacccaaccggcccagcaagacg ccccccacgaacgcccttacgcttggccagtggagtcctgtgatcgccgtct ccCGCAGCGACAACCTGGTGAGAcacatccgatccacaca ggccagaagcccttcagtgccgcatctgcatgAGAaactcagcCGAG AGGATAACTTGCACACTcacatccgacccacacagggcaaaa gcccttcgctgcgacatctgtgaaagaaagtttccCGGAGCGATGA ACTTGTCCGAcataccaagatccacttgcggcagaaggaccgccc cgttggccagtggaagctgtgatcgccgttctccCAATCAGGGAA TCTGACTGAGcacatccgatccacacagggcagaagcccttcagtg ccgcatctgcatgAGAaactcagcACAAGTGGACATCTGGT ACGCcacatccgacccacacagggcaaaagcccttcgctgcgacatctg tgaagaaagtttggcCAGAATAGTACCCTGACCGAAcatac caagatccacttgcggcagaaggacaag</p>	
47 (кодирующая)	<p>ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCT ACAGACTGGTGCAGCGGCCTCCATCTGCCGCTGC CGCCACGGCCCCACGCCCTGAGAACACTGCC CCCTACGCCGGCCCTGGTCTTGATAGCGGACTCA GACCTAGAGGCGCCCTCTGGGCCCTCCACCTCC AAGACAGCCTGGAGCCCTGGCCTACGGCGCCTTC GGCCCTCCTTCTAGCTTCCAGCCCTTCCCCGCCGT GCCTCCTCCAGC₆GCTGGCATCGCCACCTGCAG CCTGTGGCCACCCCTTACCCCGAAGAGCCGCCG CCCCTCAAACGCCCTGGCGGACCTCCTGGCCC CCAGCCTGCTCCAAGCGCCGCTGCCCTCCACCT CCTGCTCATGCCCTGGGCGGCATGGACGCCGAGC TGATCGACGAGGAAGCCCTGACCAGCCTGGAAC TGGAACTGGGCCTGCACAGAGTGCGGGAACTGC CTGAGCTGTTCTGGGACAGAGCGAGTTCGACTG CTTCAGCGACCTGGGCAGCGCCCTCCTGCCGGC TCTGTGTCCTGCgccgaccactgatgctcggagggtaccgct ggtgcagaggccgctccgccgcccgcgccatgcccctcatgctcgg actctgccgccgtacgcccggccgctggacagtgaggctgagccgagg ggggctccgctgggcccgcgcccggcccaaccggggccctggcgtac ggggccttgggcccgtcctcctccagcccttccggcgtgctccgccc gccgaggcatcgcgacactgacgctgtggcgacgcccgtacccggccc gcCgccgcgcccccaacgctcgggaggccccggcccagccggc cccaagcggcagccccgcccggccgcccgcacgcccctgggcccagtgga cggcgaactcatcgacgaggagcgtgactgctggagctggagctgggg ctgaccgctgcccagctgcccagctgttctggccagagcagttcga ctgcttctggacttgggtccgcccggcggcctcctgagctgcccagtc ccagctcateaaaccagcccatgcaagtacccaaccggcccagcaag acgccccccacgaacgcccttacgcttggccagtggagtcctgtgatcgccg ttctccCGCAGCGACAACCTGGTGAGAcacatccgatccac acaggccagaagcccttcagtgccgcatctgcatgAGAaactcagcCG</p>	206

	AGAGGATAACTTGCACACTcacatccgcaccacacagcg aaaagcccttcgctgcacatctgtggaagaaagttgccCGGAGCGA TGAACCTGTCCGAcataccaagatccacttgcggcagaaggaccgc ccttacgcttgcagtgaggctctgatcgccgttctccCAATCAGG GAATCTGACTGAGcacatccgcatccacagggcagaagcccttc cagtgccgcatctgcatGAAaacttcagcACAAGTGGACATCT GGTACGCcacatccgcaccacacagcgaaaagcccttcgctgcgac atctgtggaagaaagttgccCAGAATAGTACCCTGACCGAA cataccaagatccacttgcggcagaaggacaag	
48 (кодирующая)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCT ACAGACTGGTGCAGCGGCCTCCATCTGCCGCTGC CGCCACGGCCCCACGCCCTGAGAACACTGCC CCCTACGCCGGCCCTGGTCTTGATAGCGGACTCA GACCTAGAGGGCGCCCTCTGGGCCCTCCACCTCC AAGACAGCCTGGAGCCCTGGCCTACGGCGCCTTC GGCCCTCCTTCTAGCTTCCAGCCCTTCCCCGCCGT GCCTCCTCCAGCTGCTGGCATCGCCACCTGCAG CCTGTGGCCACCCCTTACCCCGGAAGAGCCGCCG CCCCTCAAACGCCCTGGCGGACCTCCTGGCCC CCAGCCTGCTCCAAGCGCCGCTGCCCTCCACCT CCTGCTCATGCCCTGGGCGGCATGGACGCCGAGC TGATCGACGAGGAAGCCCTGACCAGCCTGGAAC TGGAACCTGGGCTGCACAGAGTGCGGGAACTGC CTGAGCTGTTCTGGGACAGAGCGAGTTCGACTG CTTACGCGACCTGGGCAGCGCCCTCCTGCCGGC TCTGTGCTCCTGCGGCGCAGCGCGGCGGAAGC GGCgccgaccactgatgctcgcgagggtaccgcctggtgcagaggcc gectccgcccgcgcccctggcctcatgctccggactctgcgect acgcccgggggctggacagtgaggctgagccgcccgggctccgctgg ggccgcccggcccgaaccggggcctggctgacggggcttcgggc cgccgtcctcctccagcccttcggcgtgctccgcccggcgccgcatg cgacactgcagcctgtggcgacgcccgtaccggccgcccggccgccc ccaacgctccggaggccccggcccgcagcccggcccgaagcgccga gccccgcccggcccgcacgcccggggcagtggaagcgaactcagc gacgaggaggcgtgacgtcgtggagctggagctgggctgacaccgctg gcgagctgcccagctgttctggccagagcgagttcgactgcttctggact ggggtccgcccggcccggcctcgtgagctcgggtggttctggtggtt ctggtcagtcagctcatcaaacccagccgcatgcaagtagcccaaccgg cccagcaagagccccccacgaagccccttacgcttcccagtgagctctgt gatgcccgttctccCGCAGCGACAACCTGGTGAGAcacatc cgcatccacagggcagaagccctccagtgccgcatctgcatGAAaact tcagcCGAGAGGATAACTTGCACACTcacatccgcaccac acagggaaaagcccttcgctgcgacatctgtggaagaaagttgccCGG AGCGATGAACTTGTCCGAcataccaagatccacttgcggcaga aggaccgccccttacgcttcccagtgaggctctgatgcccgttctccCAA TCAGGGAATCTGACTGAGcacatccgcatccacacagggccag aagcccttccagtgccgcatctgcatGAAaacttcagcACAAGTGG ACATCTGGTACGCcacatccgcaccacacagggcgaagcccttc gctgcgacatctgtggaagaaagttgccCAGAATAGTACCCTG ACCGAAcataccaagatccacttgcggcagaaggacaag	208

49 (кодирующая)	<p>atggagctggaattggatgctggtgaccaagacctgctggccttctgctagag gaaagtggagatttggggacggcaccgatgagccgtgagggccccactgg actgggctgctgcccgttctgaggtGccgagcgactgggaagtagatgatttgc tgtgctccctgctgagtccccagcgtcgttgaacattctcagctcctccaacccc tgccttgtccacatgaccacactactcctcccacgggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggaggggcttattctgctgagacacttctctactaa gacagaggaacaattctgaaacgtgtgctggCTCGAACCAGGTGA AAAACCTTACAAATGTCTGAATGTGGGAAATC ATTCAGTCGCAGCGACAACCTGGTGAGACATCA ACGCACCCATACAGGAGAAAAACCTTATAAATG TCCAGAATGTGGAAAGTCCTTCTCACGAGAGGAT AACTTGCACACTCATCAACGAACACATACTGGTG AAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCA ACGAACCCATACAGGCGAGAAGCCTTACAAATG TCCCGAGTGTGGCAAGAGCTTCTACAATCAGGG AATCTGACTGAGCATCAACGAACTCATAACGGG GAAAAACCTTACAAGTGTCCAGAGTGTGGGAAG AGCTTTTCCACAAGTGGACATCTGGTACGCCACC AGAGGACACATACAGGGGAGAAGCCCTACAAAT GCCCCGAATGCGGTAAAAGTTTCTCTCAGAATAG TACCCTGACCGAACACCAGCGAACACACACTGG GAAAAAACGAGTgtgtaCgttgggggttagagagcCgggtctt gaaatacacagcccagaatattggagcttcagaacaaagtacagcttctgagga acagaattgtcccttctagatcaactgaggaaactccaggccatggtgattgaga tCtaaacaaccagcagcagcagcacctgcatcttggctctGctagtctctt ctgcctcctctgtactgctatgtactctctgacacaagggggagcctgccag ctgagcatggaggtgtgtcccgcagcttctgcccctcccagtgaggaccctta ccagctggagctgcctgccctgagtcagtcagaagtgccaaagacagcacacacc agtgttggacggctcagactgtgtactccaggccctggcaacacttctgct gctgcattacatgctcaggctcccagtcagagcctcccctggagtgcccCt ccctgaccttctcagagcctctctgccgaggtcccacctcccctgaggca aatctcacaaggaaggaggatggcttctactggtagcccctctgctatttga ggacagatactcaggc</p>	212
50 (кодирующая)	<p>atggagctggaattggatgctggtgaccaagacctgctggccttctgctagag gaaagtggagatttggggacggcaccgatgagccgtgagggccccactgg actgggctgctgcccgttctgaggtGccgagcgactgggaagtagatgatttgc tgtgctccctgctgagtccccagcgtcgttgaacattctcagctcctccaacccc tgccttgtccacatgaccacactactcctcccacgggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggaggggcttattctgctgagacacttctctactaa gacagaggaacaattctgaaacgtgtgctggcgccttacgcttcccagtgga gtcctgtgatcgccgttctccCGCAGCGACAACCTGGTGAG AcacatccgatccacacaggccagaagccttccagtgccgatctgcatGA GAaactcagcCGAGAGGATAACTTGCACACTcacatccg caccacacagggcaaaagccttgcctgagacatctgtggaagaaagttgc cCGGAGCGATGAACTTGTCCGAcataccaagatccacttgcg</p>	216

	<p>gcagaaggaccgcccttacgcttgcccagtgaggctctgtgatgccgcttctcc CAATCAGGGAATCTGACTGAGcacatccgcatccacacagg ccagaagcccttccagtcccgcatctcatgAGAaacttcagcACAAGT GGACATCTGGTACGCcacatccgcacccacacaggcgaaaagcc cttgcctgcgacatctgtggaagaagtgtgccAGAATAGTACCC TGACCGAAAcataccaagatccacttggcgcagaaggacgtgtaCgtt ggggttagagagcCgggtcttgaatacacagccagaatatggagctcag aacaagtacagcttctggaggaacagaattgtcccttctagatcaactgagga aactccaggccatggtgattgagatCtaaacaaaaccagcagcagcagcacc tgcacttggctctGctagtctcttctgcctcctctgtactgtctactctc tgacacaagggggagcctgccagctgagcatggagtgtgtcccgcagcttc gtgccctcccagtgaggaccctaccagctggagctgcctgcctgcagtcag aagtccgaaagacagcacaccagtggttggacggctcagactgttactcc aggccctggcaacacttctgcctgctgcattacatgcctcaggctcccagtcg agagcctcccctggagtggccCttccctgaccttctcagagcctctctgccga ggtcccatctcccctgcaggcaaatctcacaaggaaggaggatggcttct actggtagcccctctgtcatttgcaggacagatactcaggc</p>	
<p>51 (кодирующая)</p>	<p>atggagctggaattggatgctggtgaccaagacctgctggccttctctagag gaaagtggagatttggggacggcaccgatgaggccgtgaggccccactgg actggcgctgcccgttctgaggtGccgagcgactgggaagtagatgattgc tgtctccctgctgagtcccccagctcgttgaacattctcagctcctcaacccc tgccttgcaccatgaccacactactcctcccacgggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggaccagatgactccacagcata tggaggagctggcagagcaggagattgtaggctagtactgacagatgaggag aagagtctattggagaaggagggttattctgcctgagacacttctcactaa gacagaggaacaaattctgaaactgtgcccgcgaccttacgcttcccagtgga gtcctgtgatccgcttctccCGCTCAGACAACCTCGTTCTGA cacatccgcatccacagccagaagccctccagtgccgcatctgcatgAG AaacttcagcCACCGGACTACACTACGAACcacatccgca cccacacaggcgaaaagcccttgcctgagacatctggaagaagtgtgccc AGAGAAGACAATCTCCATACTcataccaagatccacttgcg gcagaaggaccgcccttacgcttgcccagtgaggctctgtgatgccgcttctcc ACCAGCCATTCTCTACTGAAcacatccgcatccacacagg ccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCAGTCT AGCTCACTGGTGAGGcacatccgcacccacacaggcgaaaagc ccttgcctgagacatctgtggaagaagtgtgccAGGGAGGATAAC CTGCATACGcataccaagatccacttgcggcagaaggacgtgtaCgtt gggggttagagagcCgggtcttgaatacacagccagaatatggagcttca gaacaagtacagcttctggaggaacagaattgtcccttctagatcaactgagg aaactccaggccatggtgattgagatCtaaacaaaaccagcagcagcagcac ctgcatcttggctctGctagtctccttgcctcctctgtacctgctatgactcct ctgacacaagggggagcctgccagctgagcatggagtgtgtcccgcagctt cgtgccctcccagtgaggaccctaccagctggagctgcctgccctgcagtc gaagtccgaaagacagcacaccagtggttggacggctcagactgtgtact ccaggcccctggcaacacttctgcctgctgcattacatgcctcaggctcccag gcagagcctcccctggagtggccCttccctgaccttctcagagcctctctgcc gaggtcccatctcccctgcaggcaaatctcacaaggaaggaggatggcttct ctactggtagcccctctgtcatttgcaggacagatactcaggc</p>	<p>218</p>
<p>52 (кодирующая)</p>	<p>atggagctggaattggatgctggtgaccaagacctgctggccttctctagag gaaagtggagatttggggacggcaccgatgaggccgtgaggccccactgg</p>	<p>220</p>

	<p>actggcgctgccgctttctgaggtGccgagcgactgggaagtagatgattgc tgtgctccctgctgagtgccccagcgtcgttgaacattctcagctcctccaacccc tgccttgccaccatgaccacacctactcctccccagggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggagggttattctgcctgagacacttctctactaa gacagaggaacaaattctgaaacgtgtgctgCGAACCAGGTGA AAAACCTTACAAATGTCCTGAATGTGGGAAATC ATTCAGTCGCAGCGACAACCTGGTGAGACATCA ACGCACCCATACAGGAGAAAAACCTTATAAATG TCCAGAATGTGGAAAGTCCTTCTCACGAGAGGAT AACTTGCACACTCATCAACGAACACATACTGGTG AAAAACCATAACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCA ACGAACCATAACAGGCGAGAAGCCTTACAAATG TCCCGAGTGTGGCAAGAGCTTCTCACAATCAGGG AATCTGACTGAGCATCAACGAACCTACACCGGG GAAAAACCTTACAAGTGTCCAGAGTGTGGGAAG AGCTTTTCCACAAGTGGACATCTGGTACGCCACC AGAGGACACATACAGGGGAGAAGCCCTACAAAT GCCCGAATGCGGTAAAAGTTTCTCTCAGAATAG TACCCTGACCGAACACCAGCGAACACACACTGG GAAAAAACGAGTgtgaCgttgggggttagagagcCgggtctt gaaatacacagcccagaatatggagcttcagaacaaagtacagcttctggagga acagaattgtccctctagatcaactgaggaaactccaggccatggtgattgaga tatca</p>	
<p>53 (кодирующая)</p>	<p>atggagctggaattggatgctggtgaccaagacctgctggccttctgctagag gaaagtggagatttggggacggcaccgatgagccgtgagggcccactgg actggcgctgccgctttctgaggtGccgagcgactgggaagtagatgattgc tgtgctccctgctgagtgccccagcgtcgttgaacattctcagctcctccaacccc tgccttgccaccatgaccacacctactcctccccagggaaactgtctctatgga tctagagagtgagagctgtagaaaagaggggacccagatgactccacagcata tggaggagctggcagagcaggagattgctaggctagtactgacagatgaggag aagagtctattggagaaggagggttattctgcctgagacacttctctactaa gacagaggaacaaattctgaaacgtgtgctgCGCAGCGACAACCTGGTGAG AcacatccgatccacacagggcagaagccctccagtccgcacatctcatgA GAaacttcagcCGAGAGGATAACTTGCACACTcacatccg caccacacagggcgaagcccttcgctgcacatctgtggaagaaagtggc cCGGAGCGATGAACTTGTCCGAcataccaagatccacttgcg gcagaaggaccgcttaccgcttcccagtgaggctctgtatcgcgcttctcc CAATCAGGGAATCTGACTGAGcacatccgatccacacagg ccagaagccctccagtccgcacatctcatgAGAaacttcagcACAAGT GGACATCTGGTACGCacatccgaccacacagggcgaagacc cttcgcctgcgacatctgtggaagaaagttgccCAGAATAGTACCC TGACCGAAcataccaagatccacttgcggcagaaggacgtgtaCgtt ggggttagagagcCgggtcttgaatacacagcccagaatatggagcttcag aacaaagtacagcttctggaggaacagaattgtccctctagatcaactgagga aactccaggccatggtgattgagatCtca</p>	<p>222</p>

Таблица 8
Последовательности нуклеиновых кислот для иллюстративных
сайтов связывания микроРНК и микроРНК

Описание	Последовательность	SEQ ID NO
Сайт связывания M1	aaagagaccgggttctactgtgacagtaaaagagaccgggttctactgtgagaatgaaagagaccgggttctactgtgatcggaagagaccgggttctactgtgagcggcctgaaaccagcagacaatgtagctcagtagaaccagcagacaatgtagctgaatggaaccagcagacaatgtagcttcggagaaccagcagacaatgtagct	7
Последовательность miR128	UCACAGUGAACCGGUCUCUUU	8
Сайт связывания miR128	AAAGAGACCGGTTCACTGTGA	9
Последовательность miR221	AGCUACAUUGUCUGCUGGGUUUC	10
Сайт связывания miR221	GAAACCCAGCAGACAATGTAGCT	11
Последовательность miR222	AGCUACAUCUGGCUACUGGGUCU	12
Сайт связывания miR222	AGACCCAGTAGCCAGATGTAGCT	13
Сайт связывания M2	GAAACCCAGCAGACAATGTAGCTAGACCCAGTAGCCAGATGTAGCTAAAGAGACCGGTTCACTGTGA	14
Сайт связывания M3	GAAACCCAGCAGACAATGTAGCTAGACCCAGTAGCCAGATGTAGCTAAAGAGACCGGTTCACTGTGAGAAACCCAGCAGACAATGTAGCTAGACCCAGTAGCCAGATGTAGCTAAAGAGACCGGTTCACTGTGA	15

Таблица 9

Различные типы структур с "цинковыми пальцами" и иллюстративные белки типа "цинковые пальцы" для производства eTF

Название типа ZF	SEQ ID NO	Структура ZF (причем каждый x может независимо быть любым остатком)	Иллюстративные белки, которые могут служить в качестве белковой платформы для eTF или ДНК-связывающего домена eTF, раскрытых в настоящем документе
«Цинковые пальцы» C2H2-типа (ZNF)	136	C-x-C-x-H-x-H	KLF4, KLF5, EGR3, ZFP637, SLUG, ZNF750, ZNF281, ZFP.89, GLIS1, GLIS3
Белки типа «пальцы Ring» (RNF)	137	C-x-C-x-C-x-H-xxx-C-x-C-x-C-x-C	MDM2, BRCA1, ZNF179
Белки типа «пальцы PHD» (PHF)	138	C-x-C-x-C-x-C-xxx-H-x-C-x-C-x-C	KDM2A, PHF1, ING1

Содержащие домен LIM	139	C-x-C-x-H-x-C-x-C-x-C-x-C-x-(C,H,D)	ZNF185, LIMK1, PXN
Ядерные гормональные рецепторы (NR)	140	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	VDR, ESR1, NR4A1
«Цинковые пальцы» СССН-типа (ZC3H)	141	C-x-C-x-C-x-H	RC3H1, HELZ, MBNL1, ZFP36, ZFP36L1
«Цинковые пальцы» FYVE-типа (ZFYVE)	140	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	EEA1, HGS, PIKFYVE
«Цинковые пальцы» ССНС-типа (ZCCHC)	142	C-x-C-x-H-x-C	CNPI.H., SF1, LIN28A
«Цинковые пальцы» ДННС-типа (ZDHHC)	143	C-x-C-x-H-x-C-xxx-C-x-C-x-H-x-C	ZDHHC2, ZDHHC8, ZDHHC9
«Цинковые пальцы» MYND-типа(ZMYND)	144	C-x-C-x-C-x-C-xxx-C-x-C-x-H-x-C	PDCD2, RUNX1T1, SMYD2, SMYD1
«Цинковые пальцы» RANBP2-типа(ZRANB)	145	C-x-C-x-C-x-C	YAF2, SHARPIN, EWSR1
«Цинковые пальцы» ZZ-типа (ZZZ)	145	C-x-C-x-C-x-C	HERC2, NBR1, CREBП.Н.
«Цинковые пальцы» С2НС-типа (ZC2HC)	142	C-x-C-x-H-x-C	IKBKKG, L3MBTL1, ZNF746
Содержащ. домен «цинковые пальцы» GATA (GATAD)	145	C-x-C-x-C-x-C	GATA4, GATA6, MTA1
ZF класс гомеобоксы и псевдогены	136	C-x-C-x-H-x-H	ADNP, ZEB1, ZHX1
Содержащ. домен THAP (THAP)	141	C-x-C-x-C-x-H	THAP1, THAP4, THAP11

«Цинковые пальцы» CXXC-типа (CXXC)	140	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	CXXC1, CXXC5, MBD1, DNMT1
«Цинковые пальцы» SWIM-типа (ZSWIM)	141	C-x-C-x-C-x-H	MAP3K1, ZSWIM5, ZSWIM6
«Цинковые пальцы» AN1-типа (ZFAND)	146	C-x-C-x-C-x-C-xxx-C-x-H-x-H-x-C	ZFAND3, ZFAND6, IGHMPLH.2
«Цинковые пальцы» 3CxxC-типа (Z3CXXC)	142	C-x-C-x-H-x-C	ZAR1, RTP1, RTP4
«Цинковые пальцы» CW-типа (ZCW)	145	C-x-C-x-C-x-C	MORC1, ZCWPW1, KDM1B
«Цинковые пальцы» GRF-типа (ZGRF)	145	C-x-C-x-C-x-C	TTF2, NEIL3, TOP3A
«Цинковые пальцы» MIZ-типа (ZMIZ)	142	C-x-C-x-H-x-C	PIAS1, PIAS3, PIAS4
«Цинковые пальцы» BED-типа (ZBED)	136	C-x-C-x-H-x-H	ZBED1, ZBED4, ZBED6
«Цинковые пальцы» HIT-типа (ZNHIT)	144	C-x-C-x-C-x-C-xxx-C-x-C-x-H-x-C	ZNHIT3, DDX59, INO80B
«Цинковые пальцы» MYM-типа (ZMYM)	145	C-x-C-x-C-x-C	ZMYM2, ZMYM3, ZMYM4
«Цинковые пальцы» matrin-типа (ZMAT)	136	C-x-C-x-H-x-H	ZNF638, ZMAT1, ZMAT3, ZMAT5
«Цинковые пальцы» C2H2C-типа	136	C-x-C-x-H-x-H	MYT1, MYT1L, ST18
«Цинковые пальцы» DBF-типа (ZDBF)	136	C-x-C-x-H-x-H	DBF4, DBF4B, ZDBF2
«Цинковые пальцы» PARP-типа	142	C-x-C-x-H-x-C	LIG3, PARP1

Таблица 10

Аминокислотные последовательности иллюстративных ДНК-связывающих доменов типа "цинковые пальцы"

DBD/целевой сайт	Последовательность	SEQ ID NO
εZF	LEPGEKP – [YKCPGKGSFS X HQRTH TGEKP] _n - YKCPGKGSFS X HQRTH – TGKKTS, где n - целое число от 1 до 15 и каждый X - это последовательность распознавания, способная связываться с 3 п.н. целевой последовательности	147
Целевой сайт Z1	RSDNLVR x REDNLHT x RSDELVR x QSGNLTE x TSGHLVR x QNSTLTE, где каждый x представляет собой линкер, содержащий 1-50 аминокислотных остатков	148
Целевой сайт Z13	RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT, где каждый x представляет собой линкер, содержащий 1-50 аминокислотных остатков	149
Целевой сайт Z14	DPGALVR x RSDNLVR x QSGDLRR x THLDLIR x TSGNLVR x RSDNLVR, где каждый x представляет собой линкер, содержащий 1-50 аминокислотных остатков	150
Целевой сайт Z15	RRDELNV x RSDHLTN x RSDDLVR x RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT, где каждый x представляет собой линкер, содержащий 1-50 аминокислотных остатков	151

Таблица 11

Аминокислотные последовательности для раскрытых в настоящем документе иллюстративных последовательностей распознавания типа "цинковые пальцы"

Последовательность	SEQ ID NO
RSDNLVR	152
REDNLHT	153
RSDELVR	154
QSGNLTE	155
TSGHLVR	156
QNSTLTE	157
DPGALVR	158
HRTTLTN	159
QSGDLRR	160
TSHSLTE	161
THLDLIR	162
QSSSLVR	163
TSGNLVR	164
RRDELNV	165
RSDDLVR	166
RSDHLTN	167

Таблица 12

Другие раскрытые в настоящем документе нуклеотидные и аминокислотные последовательности

Описание	Последовательность	SEQ ID NO
Домен EGR1 NLS	LIKPSRMRKYPNRP SK	168
SV40 NLS	PKKKRKV	169
Нуклеоплазмиды NLS	KRPAATKKAGQAKKKK	170
Тэг HA	YPYDVPDYA	171
spA (синтетич. полиА)	AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGG TTTTTGTGTGCGGACCGCACGTG	16
hGH (полиА гормона роста человека)	GGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCT GGCCCTGGAAGTTGCCACTCCAGTGCCACCAGCCTT GTCCTAATAAAATTAAGTTGCATCATTTTGTCTGACTA GGTGTCTTCTATAATATTATGGGGTGGAGGGGGGTG GTATGGAGCAAGGGGCAAGTTGGGAAGACAACCTGT AGGGCCTGCGGGTCTATTGGGAACCAAGCTGGAGTG CAGTGGCACAATCTTGGCTCACTGCAATCTCCGCCTCC TGGGTTCAAGCGATTCTCTGCCTCAGCCTCCCGAGTT GTTGGGATCCAGGCATGCATGACCAGGCTCAGCTAA TTTTTGTTTTTTGGTAGAGACGGGGTTTACCATATT GGCCAGGCTGGTCTCCAACCTCTAATCTCAGGTGATCT ACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGC GTGAACCACTGCTCCCTTCCCTGTCCTT	17
Белок SCN1A	MEQTVLVPPGPDSFNFFTRRESLAAIERRIAEEKAKNPKPD KKDDDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLED LDPYYINKKTFIVLNKGKAI FRFSATSALYILTPFNPLRKI AIKILVHSLFSMLIMCTILTNCVFM TMSNPPDWTKNVEY TFTGIYTFESLIKIIARGFCLEDFTLRDPWNWLDFTVITF AYVTEFVDLGNVSALRTRFVLRALKTISVIPGLKTIVGAL IQSVKCLSDVMILTVFCLSVFALIGLQLFMGNLRNKCIQ WPPTNASLEEHSIEKNITVNYNGTLINETVFEFDWKS YIQ DSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGRN PNYGYTSFDTFSWAFLSLFRLMTQDFWENLYQLTLRAA GKTYMIFVFLVIFLGSFYLINLILAVVAMAYEEQNQATLE EAEQKEAEFQQMIEQLKKQEEAAQQAATATASEHSREP SAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGG EEKDEDEFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSSP HQSLLSIRGSLFSPRRNSR TSLFSFRGRAKDVGSENFAD DEHSTFEDNESRRDSL FVPRRHGERRNSNLSQTSRSSRM LAVFPANGKMHSTVDCNGVVSLVGGPSVPTSPVGQLLP EVIIDKPATDDNGTTTETEMRKRSSSFHVSMDFLEDPSQ RQRAMSIASILTNTVEELESRQKCPWCWYKFSNIFLIWD CSPYWLKVKHVVNLVVMDFVDLAITICIVLNTLFMAM EHYPMTDHFNNVLT VGNLVFTGIFTAEMFLKIIAMDPYY YFQEGWNIFDGFIVT LSLVELGLANVEGLSVLRSFLLRV FKLAKSWPTLNMLIKIIGNSVGALGNLTLVLAIIVFIFAVV GMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVF RVLCGEWIETMWDCMEVAGQAMCLTVFMMVMVIGNL VVLNLFALLLSSFSADNLAATDDD NEMNNLQIAVDRM HKGVA YVKRKIYEFIQSFIRKQKILDEIKPLDDLNNKKD SCMSNHTAEIGKDL DYLKDVNGTTSIGIGTGSSVEKYIIDE SDYMSFINNPSLTVTVPIAVGESDFENLNTEDFSSESDLEE	172

	<p>SKEKLNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACF TEGCVQRFKCCQINVEEGRGKQWWNLRRTCFRIVEHNW FETFIVFMILLSGALAFEDIYIDQRKTIKTMLEYADKVFT YIFILEMLLKWVAYGYQTYFTNAWCWLDLIVDVSLVS LTANALGYSELGAIKSLRTLRLRPLRALS RFEGMRVVV NALLGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFYHCI NTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKNVF DNVGFGYLSLLQVATFKGWMDIMYAAVDSRNVELQPK YEESLYMYLYFVIFIFGSFFTLNLFIVIIDNFNQKKKFG GQDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFG MVFDFVTRQVFDISIMILICLNMVTMMVETDDQSEYVTT ILSRINLVFIVLFTGECVLKLSLRHYYFTIGWNIFDFVVVI LSIVGMFLAELIEKYFVSPTLFRVIRLARIGRILRLIKGAK GIRTLFALMMSLPALFNIGLLLFLVMFIYAIFGMSNFAY VKREVGIDDMFNFETFGNSMICLFQITTSAGWDGLLAPIL NSKPPDCDPNKVNPSSVKGDCGNPSVGIFFFVSYIIISFL VVVNMYIAVILENFVATEESAEPLEDDFEMFYEVWEK FDPDATQFMFEKLSQFAAALEPPLNLPQPNKQLIAMD LPMVSGDRIHCLDILFAFTKRVLGEGEMDALRIQMEER FMASNPSKVSYPITTTLKRKQEEVSAVIIQRAYRRHLLK RTVKQASFTYNKNKIKGGANLLIKEDMIIDRINENSITEK TDLTMSTAACPPSYDRVTKPIVEKHEQEGKDEKAKGK</p>	
Белок dCAS	<p>KRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANV ENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTD HSELGINPYEARVKGSLSQKLSSEEFSAALLHLAKRRGV HNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLER LKKDGEVRSINRFKTSYVKEAKQLKVQKAYHQLDQ SFIDTYIDLETRRTYEGPGEGSPFGWKDIKEWYEMLM GHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDEN EKLEYEYKQIENVFQKQKPTLKQIAKEILVNEEDIK YRVSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIA KILTIYQSSEDIQEELTNLSELTQEEIEQISNLKGYTGTH NLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQ KEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIE LAREKNSKDAQKMINEMQKRNROTNERIEEIRTTGKEN AKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFYEV DHIIPRSVSFDNSFNKVLVKQEEASKKGNRTPFQYLS DSKISYETFKKHILNLAKGGRISKTKKEYLLEERDINRF SVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKV KSINGGFTSFLRRKWKFKERNKGYKHAEDALIINANAD FIFKEWKLDKAKKVMENQMFEKQAESMPEIETEY KEIFITPHQIKHIKDFKDYKYSHRVDKKNRELINDTLYS TRKDDKGNTLIVNNLNGLYDKDNDKLLKLINKSPEKLL MYHHDPQTYQKLLIMEQYGDEKNPLYKYEETGNYL TKYSKKNPVIKIKYYGNKLNALHDITDDYPNSRNK VVKLSLKPYPFDVYLDNGVYKFVTVKNLDVIKENYYE VNSKCYEEAKLKKISNQAEFIAFYNNDLIKINGELYRV IGVNNDLLNRIEVMIDITYREYLENMNDKRPPRIIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQIIKKG</p>	173

<p>Конструкция dCAS9-VP64</p>	<p>MAPKKRRKVGIHGVPAAKRNYILGLAIGITSVGYGIIDYE TRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRIIRH RIQRVKLLFDYNLLTDHSELGINPYEARVKGLSQKLSE EEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISR SKALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEA KQLLKVQKAYHQLDQSFIDTYIDLLETRRTTYEGPGE PFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLY NALNDLNNLVITRDENEKLEYEYEFQIENVFKQKKPT LKQIAKEILVNEEDIKGYRVSTGKPEFTNLKVYHDIKDI TARKEIENAEELLDQIAKILTIYQSSEDIQEELTNLSEL TQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQ IAIFNRLKLVPKKVDLSQKKEIPTLVDDFILSPVVKRSF IQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMINEM QKRNRQTNERIEEIRTTGKENAKYLIEKIKLHDMQEG KCLYSLEAIPLEDLLNPFNYEVDHIIIPRSVSFDNSF NKKVLVKQEEASKKGNRTPFQYLSSSDSKISYETFF KKHILNLAAGKGRISKTKKEYLLEERDINRFSVQKDF INRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSIN GGFTSFLRRKWKFKKERNKGYKHAEDALIINANAD FIFKEWKKLDAKAKVMENQMFEEKQAESMPEIETE QEYKEIFITPHQIKHIKDFKDYKYSHRVDKKNREL INDTLYSTRKDDKGNLIVNNLNGLYDKDNDKLLK LINKSPEKLLMYHHPQTYQKLLIMEQYGDENPLYK YEEETGNLYTKYSKDNQGPVIKKIKYYGNKLN AHLDITDDYPNSRNKVVKLSLKPYPFDVYLDNGV YKFTVKNLDVIKKENYEVNSKCYEEAKKLLKISN QAEFIASFYNNDLIKINGELYRVIGVNNDDLNR IEVNMIDITYREYLENMNDKRPPRIIKTIASKTQ SIKKYSTDILGNLYEVKSKKHPQIIKKGKRPAA TKKAGQAKKKKGSYPYDVPDYALEDALDDFDL DMLGSDALDDFDLMLGSDALDDFDLMLGSDAL DDFDLMLGSDALDDFDLMLGSDALDDFDLMLG</p>	<p>174</p>
<p>Белок WT EGR3 (человека)</p>	<p>MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLF SGSSDSVVHYNQMATENVMIDIGLTNEKPNPELS YSYSGSFQAPGNKTVTYLGKFAFDSPSNWCQDNI SLMSAGILGVPPASGALSTQTSTASMVQPPQGD VEAMYPALPPYSNCGDLYSEPVSFHDPQGNP GLAYSPQDYQSAKPAALDSNLFPMIPDYNLY HHPNDMGSIPEHKPFQGMPIRVNPPPITPLETI KAFKDKQIHPGFGSLPQPPLTLKPIRPRKYPNR PSKTPLHERPHACPAEGCDRRFSRDELTRHLRI HTGHKPFQCRICMRSFSRSDHLTTHIRHTHTG EKPFAFEFCGRKFARSDERKRHAKIHLKQKEK KAKEGGAPSASSAPPVSLAPVVTCA</p>	<p>175</p>
<p>Белок WT EGR1 (человека)</p>	<p>maaakaemqlmsplqisdpgsfphsptmdnypkle emmlsngapqflgaagapegsgnsstsssgggg ggsgnsstssstfnpqadtgeqpyehhtaesfpdis lnnekvlvetsypsqttrppitytgrfslpapnsgnt lwpeplfslvsglvsmtnppasssapspaassasas qspplscavpsndsspiyaaptfptntdifpepq sqaifgsagtalqyppaypaakggfvmpidylfp qqqgdglgtpdqkpfqglesrtqqpsltplstik afatqsgsqdlkalntsyqsqlikpsrmrkypnr psktppherpyacpvescdrfhrsdeltrhiritg qkpfqcricmrfsrsdhlthirhtgekpfacdic grkfarsderkrhtkihrqdkkadksvassatss lssypsvatsypsvttsypspattsyvspvptfs spgsstypsvhsgfpspsvattysvppafpaqv ssfssavtnsfstglsdmtatfsprtieic</p>	<p>176</p>

Линкер	GGSGGGSG	177
Линкер	GGSGGGSGGGSGGGSG	178
Линкер	[GGGS]n	179
Линкер	[GGGS]n	180
Линкер	[GGSG]n	181
Сайт распознавания для WT EGR1 или EGR3	GCG(T/G)GGGCG	182
Каркас sgRNA	GTTTTAGTACTCTGGAACAGAATCTACTAAAACAAG GCAAAATGCCGTGTTTATCTCGTCAACTTGTGCGGAG A	183
Кодирующая последовательность CREB3 человека	atggagctggaattggatgctggtgaccaagacctgctggccttctgctagagaaagt ggagatttggggacggcaccgatgaggccgtgaggcccccactggactgggcgctg ccgtttctgaggtGccgagcgaactggaagtagatgattgctgtgctccctgctgagtc ccccagctcgtgaacattctcagctcctcaaccctgctgtccaccatgaccacac ctactcctcccacgggaaactgtctctatggatctagagagtgagagctgtagaaaaga ggggaccagatgactccacagcatatggaggagctggcagagcaggagattgctagg ctagtactgacagataggagaagagtctattggagaaggaggccttattctgctgag acacttctctactaagacagaggaacaattctgaacgtgtgctggaggagattcga aataaaagatctgctcaagagagccgaggaagaaaggtgtaCgttgggggttaga gagcGgggtctgaaatacacagcccagaatggagcttcagaacaaagtacagctct ggaggaacagaatttgccttctagatcaactgaggaactccaggccatggtgattgag atCtaaacaaccagcagcagcagcactgcatcttggctcctGctagtctcctctgc ctcctcctgtacctgctatgtactcctctgacacaagggggagcctgacagctgagcatg gagtgtgtcccgccagctctgctccctcccagtgaggaccctaccagctggagctgc ctgccctgagtcagaagtgcgaaagacagcacaccagtggtggagcggctcaga ctgtgtactccagcccctggcaactctgctgctgctcattacatgctcaggtccca gtgcagagcctcccctggagtggccCttccctgacctctctcagagcctctctgccgag gtcccacctcccctgaggcaaatctcacaaggaaggaggatggcttctactggtgta gcccctctgctatttgcaggacagatactcaggc	210
Последовательность AA CREB3 человека	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWA LPLSEVPSDWEVDDLCSLLSPPASLNILSSSNPCLVHHD HTYSLPRETVSMDLESESCRKEGTQMPQHMEELAEQEI ARLVLTDEEKSLLEKEGLIPETLPLTKTEEQILKRVRRI RNKRSAQESRRKKKVVYVGGLESRVLYKYTAQNMELQNK VQLLEEQNLSLLDQLRKLQAMVIEISNKTSSSSTCILVLL VSFCLLLVPAMYSSDTRGSLPAEHGVLRSQRLRALPSEDP YQLELPALQSEVPKDSHQWLDGSDCVLQAPGNTSCLL HYMPQAPSAEPPLEWFPDLFSEPLCRGPILPLQANLTRK GGWLPTGSPSVILQDRYSG	211
Человеческий С-концевой домен CREB3 с трансмембранной областью	VYVGGLESRVLYKYTAQNMELQNKVQLLEEQNLSLLDQL RKLQAMVIEISNKTSSSSTCILVLLVSFCLLLVPAMYSSD TRGSLPAEHGVLRSQRLRALPSEDPYQLELPALQSEVPKD STHQWLDGSDCVLQAPGNTSCLLHYMPQAPSAEPPLEW FPDLFSEPLCRGPILPLQANLTRKGGWLPTGSPSVILQD RYSG	225
Человеческий С-концевой	VYVGGLESRVLYKYTAQNMELQNKVQLLEEQNLSLLDQL RKLQAMVIEIS	226

домен CREB3 без трансмембранной области		
Кодирующая последовательность CREB3-TRE	atggagctggaattggatgctggtgaccaagacctgctggccttctgctagaggaaagtggagattggggacggcacccgatgaggccgtgagggccctactggactgggcgctg ccgctttctgaggtGccgagcgactgggaagtagatgattgctgtgctccctgctgagtc ccccagcgtcgttgaacattctcagctcctccaaccctgcttgcaccatgaccacac ctactcctcccacgggaaactgtctctatggatctagagagtgagagctgtagaaaaga ggggaccagatgactccacagcatatggaggagctggcagagcaggagattgctagg ctagtactgacagatgaggagaagagtctattggagaaggaggccttattctgctgag acacttctctactaagacagaggaacaattctgaacgtgtgctggCTTGAGCC CGGAGAGAAGCCGTACAAGTGCCCTGAGTGCGGCAA GTCTTTTAGCAGAAGAGACGAACCTAATGTCCACCAG CGAACGCATACTGGTGAAGGCCCTATAAATGTCCTG AATGTGGGAAATCATCTCCAGCCGAGAACCTGTAG GGCTCACCAGCGAACACACACCCGGCGAAAAACCATA CAAATGTCCAGAATGCGGGAAATCCTTTTCTCAGTCAT CCAACTTGGTGAGACATCAACGCACGCACACTGGAGA AAAGCCTTACAAATGCCCGGAATGTGGAAAGTCTTTT TCCCAATTGGCCCATTTGCGAGCCCATCAGAGGACTC ACACGGGCGAGAAACCTTACAAATGCCCGGAATGCGG GAAATCTTTTCAACGAGTGGAACCTCGTAAGACAC CAAAGAACGCATACAGGCGAAAAGCCATATAAGTGTC CTGAGTGTGGTAAATCATTCTCACACAGGACCACCT GACAAATCACCAGCGCACGCACACCCGGCAAGAAGAC AAGCgtgtaCgttgggggttagagagcCgggtcttgaaatacacagcccagaat atggagcttcagaacaagtagccttctggggaacagaattgtcccttctagatcaactg aggaactccaggccatggtgattgagatCtaaacaaaaccagcagcagcacct gcatcttggctctGctagtctccttctgctcctcctctgtactgctatgtactcctctgacac aagggggagcctgccagctgagcatggagtgtgtcccagcttctgcccctccca gtgaggacccttaccagctggagctgacctgacctgagtcagaagtcccgaagacagac acacaccagtggttgacggctcagactgtgtactccaggcccctggcaacttctgct ctgctgcattacatgcctcaggctcccagtcagagcctcccctggagtggccCttcct gaccttctcagagcctctctgcccaggtcccctcctcccctgaggcaaatctcaaa ggaaggaggaggtgcttctactgtagcccctctgtcattttgaggacagatactcagg c	214
CREB3-TRE	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWA LPLSEVPSDWEVDDLCSLLSPPASLNILSSSNPCLVHHD HTYSLPRETVSMDLESESCRKEGTQMT PQHMEELAEQEI ARLVLTDEEKSLLEKEGLILPETLPLTKTEEQILKRVRLP GEKPYKCPEC GKSFSSRRDELNVHQRTHTGEKPYKCPEC GKSFSSRRTCRAHQRTHTGEKPYKCPEC GKSFSSQSSNLV RHQRTHTGEKPYKCPEC GKSFSQLAHLRAHQRTHTGEK PYKCPEC GKSFSTSGNLVRHQRTHTGEKPYKCPEC GKSF SHRTTLTNHQRTHTGKKT SVYVGGLESRVLKYTAQNME LQNKVQLLEE QNLSLLDQLRKLQAMVIEISNK TSSSSTCI LVLLVSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALP SEDPYQLELPALQSEVPK DSTHQWLDGSDCVLQAPGNTS	215
	CLLHYMPQAPSAEPPLEWPFDFLSEPLCRGPILPLQANL TRKGGWLP TGSPSVILQDRYSG	

Примеры

Следующие ниже примеры включены для дальнейшего описания некоторых аспектов настоящего раскрытия и не должны использоваться для ограничения объема настоящего изобретения.

Пример 1.

Идентификация целевых областей, способных активировать SCN1A, с использованием SCN1A-специфических активаторов транскрипции.

Чтобы идентифицировать области генома, способные повышать экспрессию эндогенного SCN1A, разрабатывали различные сконструированные факторы транскрипции (либо нуклеазы типа "цинковые пальцы", либо конструкции gRNA/daCas9), которые нацелены на различные области генома, как указано в табл. 4 и 13 выше. Для конструкций gRNA/daCas9 эта gRNA имела ту же последовательность, что и целевая область, потому что gRNA разрабатывали для нацеливания на комплементарную геномную цепь. Белок dCas9 представлял собой конструкцию dCAS9-VP64 (SEQ ID NO: 174). Клетки HEK293 культиви-

ровали стандартными способами и трансфицировали (FugeneHD, Promega) 3 мкг плазмиды, несущей сконструированный фактор транскрипции или контрольную конструкцию, на лунку 6-луночного планшета. Клетки трансфицировали плазмидами, экспрессирующими конструкции, указанные в табл. 13 ниже. Через 48 часов после трансфекции клетки собирали, выделяли РНК (набор Qiagen RNeasy Mini) и обрабатывали ДНКазой. РНК (3 мкг) подвергали обратной транскрипции с использованием праймеров OligoDT (Superscript IV, Invitrogen). Образцы кДНК анализировали с помощью кПЦР с использованием Phusion Polymerase (New England Biolabs) и SYBR Green I: (30 секунд при температуре 98°C, 40-кратно [10 секунд при температуре 98°C, 15 секунд при температуре 66°C, 15 секунд при температуре 72°C]). Праймеры против SCN1A (5'-TGTCTCGGCATTGAGAACATTTC-3' (SEQ ID NO: 185); 5'-ATTGGTGGGAGGCCATTGTAT-3' (SEQ ID NO: 186)) использовали для количественной оценки уровня эндогенного транскрипта SCN1A, а относительные уровни экспрессии SCN1A определяли способом дельта-дельта Ct с GAPDH в качестве эталонного гена (5'-ACCACAGTCCATGCCATCAC'-3' (SEQ ID NO: 187); 5'-TCCACCACCTGTTGCTGTA-3' (SEQ ID NO: 188)). Данные представлены в виде кратных изменений относительно контрольного условия.

Результаты показаны на фиг. 1 и в табл. 13 ниже в виде кратного изменения транскрипции SCN1A относительно контрольных условий (например, репортерная конструкция EGFP-KASH). В табл. 13 приведены значения для конструкций, которые привели по меньшей мере к 1,5-кратному увеличению транскрипции по сравнению с контрольными условиями.

Таблица 13. Влияние различных целевых сайтов генома и соответствующих eTF на транскрипцию. CON указывает на конструкцию типа "цинковые пальцы", которая использовалась в эксперименте (смотрите табл. 1). Для конструкций gRNA целевой сайт и последовательность gRNA являются одинаковыми, поскольку gRNA разрабатывали для нацеливания на комплементарную цепь ДНК.

Таблица 13

CO N	Начальн. положение в хромосоме 2	Целевая последовательность	Целевой сайт с фиг. 1	SEQ ID NO для целевого сайта	Среднее	T-значение
4	166149168	ctaggtcaagttaggag	z1	18	5,12090848	0,0096628
28	166149165	GGTCAAGTGTAGGAGACA	z8	25	1,52068773	0,62403349
3	166128025	taggtaccatagagttag	z13	30	25,4730028	0,14942042
29	166127991	gaggatactgcagaggctc	z14	31	8,6766618	0,16432794
	166149176	aaggctgtctaggtcaagtgt	g9	35	1,36378425	0,18753821
	166149118	tgttctccagattaacactt	g10	36	1,63040825	0,46710683
	166128002	gatgaagccgagaggatactg	g11	37	18,7579211	0,13148732
	166128037	gctgattgtattaggtacca	g12	38	22,4892633	0,09291316
	166177299	AGAAAGCTGATACAGATACA A	g15	39	1,7542842	0,34519408
	166178880	ggtacggccaagatttcttg	g17	40	1,36801947	0,48762102
	166177299	AGAAAGCTGATACAGATACA A	g19	41	1,45636874	0,44464045
	166177369	ACACAATGAGCCACCTACAAG	g20	42	1,31187425	0,42605224
	166177362	GTGGCTCATTGTGTGTGTGCC	g22	43	1,25217773	0,26572657
	166155264	CATATCCCTGCAGGTTTCAGAA	g24	44	1,75688991	0,28984533
	166155099	agagagagagagagagaga	g25	45	2,05701745	0,42409102
	166155393	TTCTCAGTTTTGAAATTA AAA	g26	46	1,64498972	0,21705582
	166155255	TGGATTCTCTTCTGAACCTGC	g27	47	2,27026665	0,43195546
	166148361	TGCTGAGGCAGGACACAGTGT	g29	48	2,4290169	0,30364553
	166148843	ATCATCTGTAACCATCAAGGA	g30	49	2,58328006	0,0748197
	166148565	TCCTGCCTACTTAGTTTCAAG	g31	50	1,97097781	0,25980859
	166148953	ATTACAGTTCTGTGTCAGCATGC	g32	51	1,34500323	0,32186367
	166149373	TGGTCTCATCTTTTTGTGGG	g33	52	1,71471378	0,32104302
	166142239	CGATATTTTCATGGATTCTT	g34	53	1,7735976	0,21954265

	166142391	CTGACACTTACTTTGTCTAAA	g35	54	1,95513108	0,02069095
	166142219	AAAACCTGGAACCGCATTCCTCA	g36	55	2,08698135	0,0454403
	166142396	ACAAAGTAAGTGTCAAGTGTGG	g37	56	1,30739959	0,72725347
	166142344	ATAATAGTTGTGTCTTTATAA	g38	57	1,55783618	0,29846459
	166141162	TGTACAAGCAGGGCTGCAAA G	g39	58	1,4663605	0,02946062
	166140590	GTTAACAAATACACTAAACAC	g40	59	1,37399196	0,33638238
	166140928	ttcaacaagctcccaagaagt	g42	60	1,46929899	0,24465271
	166141090	ATGTTCAAGGTGCAGAAGGA A	g43	61	2,04547409	0,09880194
	165990246	TGTTTGCTCAAACGTGCACCA	g44	62	2,13402102	0,25583999
	165989684	AAATAAGACATGAAAACAAG A	g45	63	1,27016182	0,32368695
	165990193	AAATATGTACCACAAGAAATG	g46	64	2,29522738	0,41829497
	165990076	TATCTGGTTTCTCTCACTGCT	g47	65	1,44542116	0,0947106
	165989571	ATTGCAAAGCATAATTTGGAT	g48	66	1,42246971	0,18117243

Пример 2.

Активация эндогенного SCN1A в клетках HEK293 с использованием специфических для SCN1A факторов транскрипции.

Клетки HEK293 культивировали стандартными способами и высевали в 6-луночные планшеты. Клетки в каждой лунке трансфицировали (FugeneHD, Promega) 3 мкг плазмиды, несущей либо единственную сконструированную конструкцию фактора транскрипции, CREB3 WT человека (SEQ ID NO: 211), либо контрольную конструкцию EGFP. Исследуемые сконструированные конструкции фактора транскрипции включали в себя: конструкции 1-27 и 46-53 (табл. 1) и плазмиду, экспрессирующую CREB3-TRE (SEQ ID NO: 215; CREB3 с ДНК-связывающим доменом bZIP, замененным синтетическим доменом ZF, нацеленным на промотор TET) (каждый исследовали отдельно). Через 48 часов после трансфекции клетки собирали и выделяли РНК (набор Qiagen RNeasy Mini) и обрабатывали ДНКазой. РНК (3 мкг) подвергали обратной транскрипции с использованием праймеров OligoDT (Superscript IV, Invitrogen). Образцы кДНК анализировали с помощью кПЦР с использованием Phusion Polymerase (New England Biolabs) и SYBR Green I: (30 секунд при температуре 98°C, 40-кратно [10 секунд при температуре 98°C, 15 секунд при температуре 66°C, 15 секунд при температуре 72°C]). Праймеры против SCN1A (5'-TGTCTCGGCATGAGAATTC-3' (SEQ ID NO: 185); 5'-ATTGGTGGGAGGCCATGTAT-3' (SEQ ID NO: 186)) использовали для количественной оценки уровней эндогенного транскрипта SCN1A, а относительные уровни экспрессии SCN1A определяли способом дельта-дельта Ct с GAPDH в качестве эталонного гена (5'-ACCACAGTCCATGCCATCAC'3' (SEQ ID NO: 187); 5'-TCCACCACCCTGTTGCTGTA-3' (SEQ ID NO: 188)). Данные представлены в виде кратных изменений относительно контрольного условия (смотрите фиг. 2А, фиг. 2В и фиг. 2С). Контрольная конструкция состояла из EGFP, экспрессированного под контролем RE 1 (SEQ ID NO: 1). Доставка сконструированных факторов транскрипции индуцировала различную степень активации эндогенного транскрипта SCN1A по отношению к состоянию EGFP.

Пример 3.

Активация эндогенного SCN1A в ГАМК-нейронах с использованием специфических для SCN1A факторов транскрипции.

ГАМК-нейроны iCell (Cellular Dynamics) помещали в 6-луночный планшет (~1Е6 клеток/лунку) и поддерживали в соответствии с протоколом, рекомендованным производителем. Через 72 часа после посева рекомбинантный AAV (серотип AAV-DJ), экспрессирующий EGFP, или конструкцию активатора 30 на фиг. 3А, или конструкцию 25, или конструкцию 16 на фиг. 3В под контролем универсального промотора (промотор СВА) добавляли в культуральную среду приблизительно в количестве 2Е11 копий генома на лунку. Через одну неделю (фиг. 3А) или две недели (фиг. 3В) после инфицирования РНК выделяли из культивированных клеток (набор Qiagen RNeasy Mini) и обрабатывали ДЦКазой. Восстановленную РНК подвергали обратной транскрипции с использованием праймеров OligoDT (Superscript IV, Invitrogen). Образцы кДНК анализировали с помощью кПЦР с использованием Phusion Polymerase (New England Biolabs) и SYBR Green I: (30 секунд при температуре 98°C, 40-кратно [10 секунд при температуре 98°C, 15 секунд при температуре 66°C, 15 секунд при температуре 72°C]). Праймеры против SCN1A (5'-TGTCTCGGCATGAGAATTC-3' (SEQ ID NO: 185); 5'-ATTGGTGGGAGGCCATGTAT-3' (SEQ ID NO: 186)) использовали для количественной оценки уровней эндогенного транскрипта SCN1A, а относительные уровни экспрессии SCN1A определяли способом дельта-дельта Ct с GAPDH в качестве контрольного гена (5'-ACCACAGTCCATGCCATCAC'3' (SEQ ID NO: 187); 5'-TCCACCACCCTGTTGCTGTA-3' (SEQ ID NO: 188)). Данные представлены в виде кратных изменений

относительно контрольного условия (смотрите фиг. 3А и 3В). Управляемая AAV экспрессия сконструированных факторов транскрипции вызывала значительную активацию эндогенного транскрипта SCN1A в культивируемых полученных из iPS ГАМК-нейронах.

Пример 4.

Специфическая активация эндогенного SCN1A в ГАМК-нейронах с использованием специфического для SCN1A фактора транскрипции.

ГАМК-нейроны iCell (Cellular Dynamics) помещали в 6-луночный планшет (~1Е6 клеток/лунку) и поддерживали в соответствии с протоколом, рекомендованным производителем. Через 72 часа после посева в культуральную среду добавляли рекомбинантный AAV (серотип AAV-DJ), экспрессирующий EGFP или активатор (конструкция 30), который содержит DBD типа "цинковые пальцы", слитый с VPR TAD, управляемый промотором CBA под контролем промотора CBA приблизительно в количестве 2Е11 копий генома на лунку.

Через неделю после инфицирования РНК выделяли из культивированных клеток (набор Qiagen RNeasy Mini) и обрабатывали ДНКазой. Библиотеки RNAseq получали из восстановленной РНК с использованием набора библиотек TruSeq Stranded mRNA (Illumina) и секвенировали на Illumina NextSeq (секвенирование парных концов 2×75 циклов). Считанные данные секвенирования сопоставляли с геномом человека (RNASTAR), и анализ дифференциальной экспрессии выполняли с помощью DESeq2. Данные представлены в виде кратного изменения относительно контрольных (AAVDJ-CBA-EGFP) образцов (смотрите фиг. 4). Результаты показаны в табл. 14, а на фиг. 4 показана относительная экспрессия эндогенного SCN1A и транскриптов 40 ближайших соседних генов, представленная в виде кратных изменений относительно контрольного условия. Конструкция 30, как описано в табл. 1, была способна специфически увеличивать экспрессию гена SCN1A или белка Nav1.1 по сравнению с другими исследованными генами. Это указывает на то, что целевой сайт, распознаваемый активатором транскрипции конструкции 30, специфичен для гена SCN1A, что приводит к увеличению экспрессии гена SCN1A в нейронах ГАМК.

Таблица 14

Влияние на транскрипцию эндогенного SCN1A и 40 ближайших соседних генов в ГАМК-нейронах, обработанных специфическим для SCN1A фактором транскрипции (конструкция 30)

Название гена	Начало хром. 2	Конец хром. 2	Нить хромос.	Кратность изменения против контроля
PLA2R1	160788518	160919121	-	0,16367458
ITGB6	160956176	161128399	-	0,20679884
RBMS1	161128661	161350305	-	1,63514667
TANK	161993418	162092732	+	0,90946407
PSMD14	162164548	162268228	+	0,92699237
TBR1	162272604	162282381	+	0,53199642
SLC4A10	162280842	162841792	+	1,89407328
DPP4	162848750	162931052	-	2,82345284
FAP	163027193	163101661	-	2,26977379
IFIH1	163123588	163175213	-	1,46146481
GCA	163175349	163228105	+	2,58702426
FIGN	164449905	164592522	-	0,46785861
GRB14	165349321	165478358	-	0,5631965
COBLL1	165510133	165700189	-	0,43199257
SLC38A11	165752695	165812035	-	4,06730119
SCN3A	165944031	166060577	-	1,0807866
SCN2A	166095911	166248818	+	1,24475196
CSRNP3	166326156	166545917	+	0,82971233
GALNT3	166604100	166651192	-	0,33804418

TTC21B	166713984	166810353	-	1,58661143
SCN1A	166845669	166984523	-	62,9552975
SCN9A	167051694	167232503	-	1,71659087
SCN7A	167260082	167350757	-	0,29331967
B3GALT1	168675181	168730551	+	0,64436013
STK39	168810529	169104651	-	1,19821739
CERS6	169312371	169631644	+	0,86828378
NOSTRIN	169643048	169722024	+	1,82142718
SPC25	169690641	169769881	-	0,86880697
ABCB11	169779447	169887832	-	3,1441368
DHRS9	169921298	169952677	+	1,10381777
BBS5	170335687	170382432	+	0,65476347
KLHL41	170366211	170382772	+	0,87373377
FASTKD1	170386258	170430385	-	1,02786927
PPIG	170440849	170497916	+	1,09866236
CCDC173	170501934	170550943	-	0,67290779
PHOSPHO2	170550974	170558218	+	0,91339152
KLHL23	170550997	170633499	+	0,73926347
SSB	170648442	170668574	+	1,00631994
METTL5	170666590	170681441	-	1,21271497
UBR3	170683967	170940641	+	1,21350908
MYO3B	171034654	171511681	+	0,52839217

Пример 5.

Экспрессия SCN1A из кассеты экспрессии *in vivo*.

Чтобы проверить экспрессию активаторов транскрипции SCN1A *in vivo*, создавали рекомбинантные векторы AAV9 компанией VectorBiolabs (Malvern, PA). Самцам мышей C57Bl/6 (N=5 на группу, возраст 7-8 недель) проводили двустороннюю инфузию 1,5 мкл очищенного вектора AAV в дорсальный гиппокамп (AP -2,0 мм, латерально $\pm 1,5$, DV -1,4 мм от твердой мозговой оболочки) и вентральный гиппокамп (AP -3,1 мм, латерально $\pm 2,8$, DV -3,8 мм от твердой мозговой оболочки), всего 4 места инъекции. AAV доставляли со скоростью 0,3 мкл/мин с периодом отдыха 4 мин после каждой инъекции. Через четыре недели после воздействия мышей умерщвляли и иссекали ткань гиппокампа. Для каждой группы ткань левого и правого гиппокампа собирали вместе для гомогенизации у большинства животных (N=4), за исключением одного животного, у которого собирали и гомогенизировали только левый гиппокамп. РНК выделяли из гомогената (набор Qiagen RNeasy Mini) и обрабатывали ДНКазой. РНК (3 мкг) подвергали обратной транскрипции с использованием праймеров OligoDT (Superscript IV, Invitrogen). Образцы кДНК анализировали с помощью кПЦР на экспрессию мышинного SCN1A с использованием Phusion Polymerase (New England Biolabs) и SYBR Green I: 30 секунд при температуре 98°C, 40-кратно [10 секунд при температуре 98°C, 15 секунд при температуре 64°C, 15 секунд при температуре 72°C]. Праймеры против SCN1A мыши (5'-CAAAAAGCCACAAAAGCCT-3' (SEQ ID NO: 189); 5'-TTAGCTCCGCAAGAAACATC-3' (SEQ ID NO: 190)) использовали для количественной оценки уровней эндогенного транскрипта SCN1A, а относительные уровни экспрессии SCN1A *in vivo* определяли способом дельта-дельта Ct с GAPDH в качестве эталонного гена (5'-ACCACAGTCCATGCCATCAC'3' (SEQ ID NO: 187); 5'-TCCACCACCCTGTTGCTGTA-3' (SEQ ID NO: 188)).

На фиг. 5А и фиг. 5В показаны средние результаты для пяти животных, каждому из которых вводили конструкцию AAV9. Контрольная конструкция eGFP содержала репортерный трансген eGFP. Конструкция 4 (смотрите табл. 1) содержала активатор транскрипции, который распознавал целевую последовательность, содержащую SEQ ID NO: 18, как описано в табл. 1 выше. На фиг. 5А показана относительная экспрессия SCN1A *in vivo*. На фиг. 5В показано изменение экспрессии SCN1A *in vivo* в процентах от средней экспрессии eGFP. Эти результаты показали, что активатор транскрипции SCN1A экспрессионной кассеты А приводил к усилению экспрессии SCN1A приблизительно на 20-30% *in vivo*.

Такие кассеты экспрессии могут быть адаптированы для применения на людях для лечения синдрома Драве, эпилепсии, судорог, болезни Альцгеймера, болезни Паркинсона и/или любых других заболеваний или состояний, связанных с дефицитом и/или нарушением активности SCN1A.

Пример 6.

Анализ гипертермических припадков (HTS) на мышах с синдромом Драве.

А. Гетерозиготная модель мыши с нокаутом Scn1a.

Лечение синдрома Драве и/или его симптомов с использованием кассет экспрессии исследовали на линии мышей Scn1a^{tm1Kca}. Эта линия мышей представляет собой признанную мышиную модель для синдрома Драве. Линии мыши Scn1a^{tm1Kca} не требуют рекомбиназы CRE. Мышь Scn1a^{tm1Kca} (доступная из лаборатории the Jackson Laboratory; описанная в Hawkins et al., Scientific Reports, vol. 7: 15327 (2017)) содержит делецию первого кодирующего экзона SCN1A. Мыши, гомозиготные по аллелю нокаута SCN1A, характеризуются тремором, атаксией, припадками и умирают к 16-му дню после рождения. У гетерозиготных мышей на фоне C57BL/6 развиваются спонтанные припадки, и большая часть их умирает в течение нескольких недель. Такая линия мышей может быть использована для изучения безопасности и эффективности лечения эпилепсии и синдрома Драве. См. Miller et al., Genes Brain Behav. 2014 Feb;13(2):163-72 для дополнительной информации.

Чтобы исследовать эффективность активаторов транскрипции в линии мышей Scn1a^{tm1Kca}, в пометы детенышей, полученных от самцов Scn1a +/-, скрещенных с самками C57BL/6J, вводили вектор AAV посредством двусторонней ICV в точке P1. Мышам вводили дозы конструкций 31-34 (табл. 1). Мышей оставляли без воздействий вместе с их матерью до отлучения в P18, а затем снова не трогали до P26-P28, когда начинали анализ гипертермического припадков (HTS). Отдельные пометы мышей, которым вводили дозу в P1, отсеивали на P18 и ежедневно наблюдали за смертностью. Индукцию гипертермического припадков проводили у гетерозиготных (HET) мышей и мышей WT Scn1a на смешанном фоне 129Stac X C57BL/6 в P26-P28. Перед исследованием мышам вставляли смазанный ректальный датчик температуры (Ret-4) и подсоединяли к модулю контроля температуры (TCAT 2DF, Physitemp), который был последовательно соединен с нагревательной лампой (HL-1). Затем мышей помещали в большой стеклянный стакан и на короткое время давали уравновеситься с окружающей средой. После этого температура тела повышалась на ~0,5°C каждые 2 минуты до начала первого тонико-клонического припадков, сопровождающегося потерей положения стоя или до достижения температуры 43°C. Если у мыши случался припадок с потерей положения тела, эксперимент заканчивали и регистрировали внутреннюю температуру тела мыши. Если в течение всего эксперимента не было обнаружено приступов с потерей положения тела, эта мышь считалась без припадков, и анализ завершали. Образцы тканей получали от мышей в точке P1, а генотипирование мышей выполняли в ходе эксперимента с использованием ПЦР в реальном времени. После завершения анализа генотипирование было не слепым, и статус мышей как HET или WT коррелировал с полученными данными. Данные наносили на кривую выживаемости Каплана-Мейера, а значимость определяли с помощью теста Мантеля-Кокса. Результаты показаны в табл. 15 и 16 и на фиг. 6А-Е.

Таблица 15

Обобщение условий, использованных в примере 6

Конструкция	Дозировка (копии генов/мышь)
Конструкция 31 (фиг. 6А и 6Е))	6,0E+10
Конструкция 32 (фиг. 6В, 6D, 6Е)	3,1E+11
Конструкция 33 (фиг. 6С)	5,8E+10
Конструкция 34 (фиг. 6D)	4,9E+13

Таблица 16

Обобщение результатов анализа на гипертермические припадки

Конструкция eTF (FIG.)	Кол-во контрольных животных (обработ. PBS)	Кол-во обработанных животных	% без припадков при темп. 42,6°C	P-значение
Репортер EGFP	16	N/A	44%	
Конструкция 31 (фиг. 6А и 6Е))	16	18	95	P<0,0001
Конструкция 32 (фиг. 6В, 6D, 6Е)	16	21	76	P<0,05
Конструкция 33 (фиг. 6С)	16	14	93	P<0,001
Конструкция 34 (фиг. 6D)	16	12	83	P<0,05

Дополнительные эксперименты проводили на мышах Scn1a^{tm1Kca}, как описано выше, чтобы проверить влияние конструкций 42 и 43 на припадки в анализе HTS. В этих экспериментах конструкцию 42 вводили в дозе 9×10^{10} копий генома/мышь посредством двусторонней ICV в P1, а конструкцию 43 вводили в дозе 6×10^{10} копий генома/мышь посредством двусторонней ICV в P1 или P5. Результаты показаны на фиг. 6F (конструкция 42) и фиг. 6G (конструкция 43). Обе конструкции показали значительное снижение припадков по сравнению с контрольными EGFP (P<0,0001 как для конструкции 42, так и для конструкции 43).

В. Гетерозиготная мутантная модель мыши $Scn1a^{RX}$.

Лечение синдрома Драве и/или его симптомов с использованием кассеты экспрессии по настоящему раскрытию исследовали на линии мышей $Scn1a^{RX}$. Эта линия мышей представляет собой признанную модель мышей для синдрома Драве. Линии мыши $Scn1a^{RX}$ не требуют рекомбиназы CRE. Мышь $Scn1a^{RX}$ (доступная из лаборатории Jackson Laboratory; описанная в Ogiwara et al., J. Neuroscience, vol. 27: 5903-5914 (2007)), содержит нонсенс-мутацию с потерей функции в экзоне 21 гена $SCN1A$ (CgG на TgA; R1407X). У гетерозиготных мышей на фоне C57BL/6 развиваются спонтанные припадки, и большой процент мышей умирает в течение нескольких недель.

Для проверки эффективности активаторов транскрипции в линии мышей $Scn1a^{RX}$ в пометы детенышей, полученных от скрещивания посредством ЭКО сперматозоидов самцов $Scn1a^{RX/+}$ с яйцеклеткой C57BL/6J с эмбрионами, имплантированными самкам CD-1, вводили вектор AAV (конструкция 31) в дозе $5,1 \times 10^{10}$ копий генома (gc)/мышь или контроль PBS посредством двусторонней ICV в P1. Мышей оставляли без воздействий вместе с их матерью до отлучения в P18, а затем снова оставляли без воздействий до P26-P28, когда начинали анализ HTS. Отдельные пометы мышей, которым вводили дозу в P1, осматривали в P18 и ежедневно наблюдали за смертностью. Индукцию гипертермического припадка проводили у гетерозиготных (HET) мышей P26-P28 и мышей WT $Scn1a$ на фоне C57BL/6. Перед исследованием мышам вставляли смазанный ректальный датчик температуры (Ret-4) и подсоединяли к модулю контроля температуры (TCAT 2DF, Physitemp), который был последовательно соединен с нагревательной лампой (HL-1). Затем мышей помещали в большой стеклянный стакан и на короткое время давали уравновеситься с окружающей средой. После этого температура тела повышалась на $\sim 0,5^\circ\text{C}$ каждые 2 минуты до начала первого тонико-клонического припадка, сопровождающегося потерей положения тела, или до достижения температуры 43°C . Если у мыши случался припадок с потерей положения тела, эксперимент заканчивали и регистрировали внутреннюю температуру тела мыши. Если в течение всего эксперимента не было обнаружено припадков с потерей положения тела, эта мышь считалась без припадков, и анализ завершался. Образцы тканей получали от мышей в точке P1, и генотипирование мышей выполняли в ходе эксперимента с использованием ПЦР в реальном времени. После завершения анализа генотипирование было не слепым, и статус мышей как HET или WT коррелировал с полученными данными. Ни у одной из исследованных мышей WT $Scn1a$ не было припадка. Данные для мышей HET, обработанных конструкцией 31 ($n=13$) и обработанных PBS контролем ($n=14$), наносили на кривую выживаемости Каплана-Мейера, и значимость определяли с помощью критерия Кокса-Мантеля. Как показано на фиг. 6H, мыши HET, обработанные конструкцией 31, показывают значительное снижение индукции гипертермического припадка по сравнению с мышами HET, получавшими контрольный PBS ($P < 0,01$).

Пример 7.

Анализ выживаемости на мышинной модели синдрома Драве.

А. Гетерозиготная модель мыши с нокаутом $Scn1a$.

Для проверки эффективности активаторов транскрипции в линии мышей $Scn1a^{tm1Kea}$ в пометы детенышей, полученных от самцов $Scn1a +/-$, скрещенных с самками C57BL/6J, вводили вектор AAV посредством двусторонней ICV в точке P1. Мышей не оставляли без воздействия с их матерью до отлучения. Наблюдение за состоянием здоровья мышей $Scn1a +/-$ проводили ежедневно после отлучения в P18. Мышам, которые были найдены мертвыми в своей домашней клетке по любой причине, была записана дата. Данные наносили на кривую выживаемости Каплана-Мейера, а значимость определяли с помощью критерия Кокса-Мантеля.

Результаты показаны в табл. 17 и на фиг. 7A-D.

Таблица 17

Обобщение условий и результатов анализа выживаемости

SEQ ID	Дозировка (gc/мышь)	Кол-во контрольных животных (обработ. PBS)	Кол-во обработанных животных	% выживших в P100 (*в P83)	P-значение
PBS	N/A	53	N/A	49%	
Конструкция 33 (фиг. 7C и 7D)	$5,8E+10$	53	29	76%	$P < 0,05$
Конструкция 31 (фиг. 7B и 7D)	$6,0E+10$	53	34	97%	$P < 0,0001$

В. Гетерозиготная модель мутантной мыши $Scn1a^{RX}$.

Чтобы проверить эффективность активаторов транскрипции в линии мышей $Scn1a^{RX}$, в пометы детенышей, полученных от самцов и скрещивания ЭКО сперматозоидов $Scn1a^{RX/+}$ с яйцеклетками C57BL/6J, в эмбрионы, имплантированных самками CD-1, вводили вектор AAV (конструкция 31) при $5,1 \times 10^{10}$ копий генома (gc)/мышь или контроль PBS посредством двусторонней ICV в P1. Мышей оставляли без воздействия с их матерью до отлучения. Наблюдение за состоянием здоровья мышей $Scn1a$

RX/+ проводили ежедневно после отлучения в P18. Мышам, которые были найдены мертвыми в своей домашней клетке по любой причине, была записана дата. Данные для обработанных конструкцией 31 (n=27) и обработанных контролем (n=18) наносили на кривую выживаемости Каплана-Мейера, а значимость определяли с помощью критерия Кокса-Мантеля.

Как показано на фиг. 7E, мыши *Scn1a^{RX/+}*, получившие конструкцию 31, характеризовались повышенной выживаемостью по сравнению с мышами *Scn1a^{RX/+}*, получавшими контрольный PBS (P<0,0001).

Пример 8.

Уровни транскрипции *SCN1A* у отличных от человека приматов после лечения AAV, кодирующим специфический фактор транскрипции *SCN1A*.

В исследовании использовали самцов яванского макака (*macaca fascicularis*) в возрасте от 2 до 3 лет. Животных перед включением в исследование предварительно проверяли на наличие перекрестно-реактивных антител к AAV9 с помощью клеточного анализа нейтрализующих антител. AAV9, экспрессирующий фактор транскрипции, специфический для *SCN1A* (конструкция 33), или контроль разводили в PBS и вводили интрапаренхимально в концентрации 1,2E12 gc/животное. Для инъекций определяли три различных стереотаксических координаты в каждом полушарии, шесть участков инъекции на животное. В каждую точку вводили объем 10 мкл. Инъекции в правое полушарие были симметричны левым. Два необработанных животных использовали в качестве контроля.

Для оценки экспрессии мРНК *Scn1A* проводили обратную транскрипцию с последующим способом кПЦР. При вскрытии трупа через 28 дней после введения дозы срезы тканей из различных областей головного мозга (лобная кора, теменная кора, височная кора, затылочная кора, гиппокамп, продолговатый мозг, мозжечок; по 200 мг каждого) от контрольных и обработанных животных собирали в RNeasy Lysis Buffer, а затем замораживали. Вкратце, иссекали 30 мг ткани, экстрагировали РНК (с помощью мини-набора для ткани Qiagen RNeasy Lipid, номер по каталогу 1023539), преобразовывали в кДНК путем обратной транскрипции (с использованием набора для обратной транскрипции кДНК высокой емкости Applied Biosystems, номер по каталогу 4368814) и проводили кПЦР с использованием набора праймеров/зондов для *Scn1A* и гена домашнего хозяйства *GAPDH* (Applied Biosystems, номер по каталогу Rh02621745-g1 FAM).

Наборы праймеров/зондов для *SCN1A* приведены ниже.

Таблица 18

Последовательности праймеров, использованные в примере 8

Ген	SEQ ID NO	Последовательность (5'-3')	Примечание
Scn1A	191	CCATGGAAGTGGCTCGATTTAC	F-праймер
	192	ATTGGTGGGAGGCCACTGTAT	R-праймер
	193	AGGCCTGAAAACCATTGTGGGAGCCCT	Зонд (FAM)

Экспрессию гена *Scn1A* в каждом исследуемом образце определяли относительным количественным анализом (RQ) с использованием способа сравнительного Ct (Δ Ct). Этот способ измеряет разность Ct (Δ Ct) между целевым геном и геном домашнего хозяйства, а затем сравнивает значения Δ Ct образцов для лечения с контрольными образцами.

Δ Ct = среднее Ct целевого гена - среднее Ct гена домашнего хозяйства.

$\Delta\Delta$ Ct = Δ Ct образца лечения - Δ Ct контрольного образца.

Относительная экспрессия (образец лечения) = $2^{-\Delta\Delta$ Ct}

Данные представлены в виде нормализованной экспрессии целевой мРНК в различных тканях головного мозга (см. фиг. 8). Как показано на фиг. 8, участки головного мозга, проксимальные к участкам интрапаренхиматозных инъекций, показали самые высокие уровни экспрессии транскрипта *SCN1A*.

Пример 9.

Селективная экспрессия трансгена в PV-содержащих нейронах у отличных от человека приматов после лечения AAV, имеющего селективный в отношении PV промотор и сайт связывания микроРНК.

В исследовании использовали шесть мартышек (*Callithrix jacchus*), которых предварительно проверяли на перекрестно-реактивные антитела к AAV9 перед включением в исследование. Двум обезьянам вводили AAV9, содержащий трансген EGFP под контролем промотора EF1alpha, двум обезьянам вводили AAV9, содержащий трансген EGFP, под контролем RE 2 (SEQ ID NO: 2), и двум обезьянам вводили AAV9, содержащий трансген EGFP под контролем RE 2 (SEQ ID NO: 2), а также содержащий сайт связывания микроРНК (SEQ ID NO: 7), расположенный между кодирующей областью EGFP и полиА-хвостом. Векторы AAV9 разводили в PBS и животным вводили три внутримозговые инъекции (2 мкл каждая) в гиппокамп/энторинальную кору каждого полушария, всего 6 участков инъекции на животное. Два животных получали вектор AAV9, содержащий EF1alpha-EGFP, каждое получило общую дозу 5,8E+11 копий генома/животное, два животных получали вектор AAV9, содержащий RE 2-EGFP, каждое получило общую дозу 3,0E+11 копий генома/животное, и два животных получали вектор AAV9, содержащий RE 2 + m1-EGFP, каждое получило общую дозу 2,3E+11 копий генома/животное.

Иммуногистохимию использовали для оценки селективной экспрессии парвальбумина (PV). При вскрытии трупа, через 28 дней после введения дозы, собирали срезы тканей из различных областей головного мозга. Плавающие срезы головного мозга мартышек (35 мкм) фиксировали в 4% параформальдегиде, блокировали буфером (PBS, 3% BSA, 3% ослиная сыворотка, 0,3% Triton-X 100, 0,2% Tween-20), а затем окрашивали анти-GFP (Abcam ab290), а затем вторичным антителом, конъюгированным с Alexa-488 (Thermo A21206). За этим последовали антитело к PV (Swant) и вторичное антитело, конъюгированное с Alexa-647 (Thermo A31571) и 4',6-диамидино-2-фенилиндолом (DAPI). Срезы монтировали и визуализировали с помощью PerkinElmer Vectra3.

Результаты показаны на фиг. 9A-F и 10A-L.

Пример 10.

Биораспространение eTF^{SCN1A}.

Целью этого исследования было сравнить биораспределение eTF^{SCN1A} в центральной нервной системе (ЦНС) молодых обезьян яванского макака при введении в дозе 4,8E+13 посредством односторонней интрацеребровентрикулярной (ICV) инъекции. Каждому животному инъецировали AAV9, содержащий кассету экспрессии, кодирующую eTF^{SCN1A}, под контролем селективного регуляторного элемента ГАМК (GABA) (RE^{GABA}-eTF^{SCN1A}). Частицы AAV9 составляли в PBS + 0,001% плороновой кислоты и вводили в дозе 4,8E+13 или 8E+13 vg/животное. Каждому животному вводили 2 мл приготовленных вирусных частиц. Дизайн исследования представлен в табл. 19.

Яванских макаков в возрасте двадцать четыре месяца группировали, как указано в табл. 19. Перед началом исследования образцы крови животных исследовали на уровне титра нейтрализующих антител к AAV9 с использованием анализа титра NAb, описанного выше. Для исследования выбирали животных с низкими или отрицательными результатами на антитела. Образцы вводили посредством инъекции ICV с использованием стандартных хирургических процедур. Размороженный дозируемый материал ненадолго хранили на влажном льду и нагревали до комнатной температуры непосредственно перед дозированием. Животных анестезировали, готовили к операции и помещали в стереотаксическую рамку, совместимую с MPT (Корф). Для определения координат цели выполняли базовую MPT. Делали разрез и просверливали одно отверстие в черепе над целевым участком. Шприц BD объемом 3 мл, прикрепленный к набору удлинителей с микроканальным отверстием длиной 36 дюймов, подготавливали с образцом и помещали в инфузионный насос. Линию экстензии примировали. Твердую мозговую оболочку открывали, и дозирующую иглу продвигали на глубину от 13,0 до 18,1 мм от мягкой мозговой оболочки. Инъекцию контрастного вещества и рентгеноскопию использовали для подтверждения введения спинномозговой иглы в правый боковой желудочек. В иглу Quinke BD размером 3,0 дюйма 22 г для спинномозговой хирургии добавляли контраст, чтобы определить ее расположение, перед тем, как прикрепить заправленную линию экстензии и шприц. Настройки насоса составляли 0,1 мл/мин в течение 19-20 минут. Буфер вводили вручную после введения дозы, чтобы очистить линию экстензии. Иглу оставляли на месте в течение 1-2 минут после завершения инфузии, а затем иглу извлекали. Носитель и исследуемое вещество вводили один раз в день 1, и субъектов поддерживали в течение 27- или 29-дневного периода восстановления.

Таблица 19
Дизайн исследования биораспределения

Группа	Пол	ID	Доза (VG/животное)
Группа 1 (контроль-буфер)	М	21001	
	Ж	11501	
Группа 2 (RE ^{GABA} -eTF ^{SCN1A})	М	2001	4,8E+13
	Ж	2501	4,8E+13
	М	3001	4,8E+13
	М	3002	8E+13

После введения дозы за животными регулярно наблюдали на протяжении всего исследования, и периодически отбирали образцы крови. Введение eTF^{SCN1A} не было связано с неожиданной смертностью, клиническими данными или макроскопическими наблюдениями. Животные, получившие AAV9-RE^{GABA}-eTF^{SCN1A}, выжили до запланированного вскрытия трупа на 28 ± 2 дня. Никаких клинических или поведенческих признаков, повышения температуры тела или снижения массы тела во время ежедневных или еженедельных медицинских осмотров не наблюдалось. Временное повышение печеночных трансаминаз (АЛТ и АСТ) у животных, получавших AAV9-RE^{GABA}-eTF^{SCN1A}, наблюдалось, но полностью исчезало к концу исследования без иммуномодуляции, и не было отмечено сопутствующего повышения сывороточного билирубина или щелочной фосфатазы. Ни один из других измеренных результатов клинической химии не был выдающимся. В гистопатологических исследованиях печени не сообщалось о микроскопических наблюдениях. Лейкоциты CSF были повышены при окончательном сборе по сравнению со значениями до воздействия, но сравнимы между контрольными животными и животными, получившими AAV9-RE^{GABA}-eTF^{SCN1A}. Плеоцитоза, связанного с AAV9-RE^{GABA}-eTF^{SCN1A}, не наблюдалось. Макрообследование и подробное микрогистопатологическое исследование ненейрональных тканей у всех животных не было примечательным. Ткани включали в себя основные периферические органы (например,

сердце, легкие, селезенку, печень и гонады). Макрообследования и подробная микрогистопатология нейрональных тканей не показали каких-либо заметных результатов. Ткани включали в себя головной, спинной мозг и связанные с ним межпозвоночные узлы (из шейного, грудного и поясничного отделов). Исследования проводились тремя независимыми патологами, в том числе одним в специализированном центре невропатологии.

ICV введение AAV9 не предотвращало постдозовый иммунный ответ в сыворотке, поскольку через четыре недели после введения дозы наблюдались нейтрализующие антитела к капсиду AAV9. Однако уровни нейтрализующих антител к AAV9 в CSF оставались неизменными и сравнимыми с уровнями до введения дозы (табл. 20).

Таблица 20
Титр сывороточных NAb к AAV9

Номер субъекта	Титр сывороточных NAb к AAV9			Титр NAb CSF к AAV9	
	До скрининга	При инъекциях.	Через 4 недели после инъекций.	При инъекциях.	Через 4 недели после инъекций.
21001	1:5	< 1:5	< 1:5	< 1:5	< 1:5
11501	< 1:5	< 1:5	< 1:5	< 1:5	< 1:5
2001	< 1:5	< 1:5	1:405	< 1:5	1:5
2501	< 1:5	< 1:5	1:135	< 1:5	1:5
3001	< 1:5	< 1:5	1:1215	< 1:5	< 1:5
3002	< 1:5	< 1:5	1:135	< 1:5	< 1:5

Образцы собирали через 27-29 дней после введения дозы из основных органов (желудочки сердца, долей печени, долей сердца легких, почек, селезенки, поджелудочной железы и шейных лимфатических узлов) во время плановой некроскопии. Прицельную биопсию собирали с помощью восьмимиллиметровых биопсийных щипцов и обрабатывали, как описано ниже.

Пример 11.

Биораспределение eTF^{SCN1A} в головном мозге.

ddPCR использовали для измерения биораспределения eTF^{SCN1A} в головном мозге. Образцы из различных областей ткани головного мозга яванского макака (FC: лобная кора; PC: теменная кора; TC: височная кора; Hip: гиппокамп; Med: продолговатый мозг; OC: затылочная кора) измеряли на количество копий вектора для оценки биораспределения eTF^{SCN1A} под контролем селективного регуляторного элемента ГАМК (RE^{GABA}-eTF^{SCN1A}) при введении AAV9 посредством односторонней ICV. Тканевую ДНК выделяли с помощью наборов DNeasy Blood & Tissues (Qiagen). Количество ДНК определяли и нормализовали с помощью УФ-спектрофотометра. 20 нанограмм тканевой ДНК добавляли к реакционной смеси объемом 20 микролитров вместе с ddPCR Super Mix for Probes (без dUTP) (Bio-Rad) и праймерами Taq-Man и зондами, направленными против областей последовательности eTF^{SCN1A}. Создавали капли и амплифицировали шаблоны с использованием автоматического генератора капель и термоциклера (Bio-Rad). После этапа ПЦР планшеты загружали и считывали с помощью QX2000 Droplet Reader для определения числа копий вектора в тканях. Ген обезьяньего альбумина (MfAlb) служил внутренним контролем для нормализации содержания геномной ДНК и был амплифицирован в той же реакции. Праймеры и зонды для eTF^{SCN1A} и MfAlb представлены в табл. 21.

Таблица 21
Праймеры и зонды для eTF^{SCN1A} и MfAlb

Название праймеров/зонда		Последовательность (5'-3')
eTF ^{SCN1A}	Прямой праймер eTF ^{SCN1A}	GAATGTGGGAAATCATTCAGTCGC (SEQ ID NO: 194)
	Обратный праймер eTF ^{SCN1A}	GCAAGTTATCCTCTCGTGAGAAGG (SEQ ID NO: 195)
	Зонд eTF ^{SCN1A}	GCGACAACCTGGTGAGACATCAACGCACC (SEQ ID NO: 196)
MfAlbumin	Прямой праймер MfAlb	GCTGTATCTCTTGTGGGCTGT (SEQ ID NO: 197)
	Обратный праймер MfAlb	AAATCATGGGAGCTGCCGGTT (SEQ ID NO: 198)
	Зонд MfAlb	CCACACAAATCTCTCCCTGGCATTG (SEQ ID NO: 199)

eTF^{SCN1A} широко распределялся по всему головному мозгу при дозировке 4,8E+13 вирусных геномов на животное со средним значением 1,3-3,5 вирусных геномов/диплоидный геном (фиг. 11). Кроме того, при сравнении переноса гена по всему головному мозгу RE^{GABA}-eTF^{SCN1A}, после дозированного вве-

деня 4,8E+13 вирусных геномов на животное, с переносом гена через головной мозг eGFP, введенного посредством ICV в различных дозах, увеличение вирусного генома/диплоидного генома наблюдалось с увеличением дозы. Это указывает на то, что перенос гена в головном мозге происходит дозозависимым образом при введении в AAV9 через ICV.

Пример 12.

Транскрипция eTF^{SCN1A} в головном мозге.

Транскрипцию eTF^{SCN1A} под контролем селективного регуляторного элемента ГАМК, RE^{GABA} ($RE^{GABA}-eTF^{SCN1A}$), оценивали путем измерения мРНК eTF^{SCN1A} с использованием анализа экспрессии генов на основе ddPCR. Тканевую РНК выделяли с помощью наборов RNeasy Plus Mini (Qiagen) или RNeasy Lipid Tissue Mini (Qiagen) для тканей головного мозга. Количество РНК определяли и нормализовали с помощью УФ-спектрофотометра, а качество РНК (RIN) проверяли с помощью Bioanalyzer RNA Chip. Один микрограмм тканевой РНК использовали для обработки ДНКазой и синтеза кДНК с помощью набора для синтеза кДНК SuperScript VILO с наборами ферментов ezDNase™ (Thermo Fisher). 50 микрограммов РНК преобразовывали в кДНК. кДНК добавляли в реакционную смесь объемом 20 микролитров вместе с ddPCR Super Mix for Probes (без dUTP) (Bio-Rad) и праймерами и зондами TaqMan, направленными против областей последовательности eTF^{SCN1A} (табл. 22). Создавали капли и амплифицировали шаблоны с использованием автоматического генератора капель и термоциклера (Bio-Rad). После ПЦР-амплификации планшет загружали и считывали с помощью QX2000 Droplet Reader, чтобы получить уровни экспрессии генов в тканях. Ген обезьяны ARFGAP2 (MfARFGAP2) (Thermo Fisher Scientific) служил эндогенным контролем для нормализации уровней экспрессии гена и был амплифицирован в той же реакции. Среднее количество транскриптов для ARFGAP2 составляло 1,85E+6 мкг РНК (фиг. 12, верхняя граница). Предел обнаружения обозначен нижней границей.

мРНК eTF^{SCN1A} наблюдалась по всему головному мозгу у всех животных, что указывает на то, что ГАМК-селективный промотор, RE^{GABA} , был транскрипционно активен в ткани головного мозга для всех макак, получивших AAV9- $RE^{GABA}-eTF^{SCN1A}$ (фиг. 12). FC: лобная кора; PC: теменная кора; TC: височная кора; Hip: гиппокамп; Med: продолговатый мозг; OC: затылочная кора.

Таблица 22

Праймеры и зонды TaqMan, направленные против областей последовательности eTF^{SCN1A}

Название праймеров/зонда	Описание	Последовательность (5'-3')
eTF^{SCN1A}	Прямой eTF^{SCN1A} праймер	GAATGTGGGAAATCATTTCAGTCGC (SEQ ID NO: 200)
	Обратный eTF^{SCN1A} праймер	GCAAGTTATCCTCTCGTGAGAAGG (SEQ ID NO: 201)
	Зонд eTF^{SCN1A}	GCGACAACCTGGTGAGACATCAACGCACC (SEQ ID NO: 202)
MfARFGAP2	Прямой, обратный праймеры, зонд	Thermo Fisher (номер в каталоге: 4448491)

Пример 13.

Биораспределение и транскрипция eTF^{SCN1A} в периферических тканях.

Число копий вектора дополнительно измеряли в различных органах для оценки трансдукции $RE^{GABA}-eTF^{SCN1A}$ в тканях по всему организму при введении AAV9 посредством односторонней ICV. Уровни транскрипции eTF^{SCN1A} также измеряли с помощью ddPCR для оценки транскрипционной активности eTF^{SCN1A} под контролем ГАМК-селективного регуляторного элемента RE^{GABA} в тканях по всему организму при введении в AAV9 посредством односторонней ICV. Оба способа выполняли, как в общем описано выше. Трансдукция $RE^{GABA}-eTF^{SCN1A}$ и транскрипция eTF^{SCN1A} в спинном мозге (SC) и межпозвоночном узле (DRG) были сопоставимы с уровнями, наблюдаемыми в головном мозге. За исключением печени, трансдукция $RE^{GABA}-eTF^{SCN1A}$ была ниже в периферических тканях за пределами головного мозга (фиг. 13). Трансдукция $RE^{GABA}-eTF^{SCN1A}$ в печени была выше, чем в головном мозге. Транскрипция eTF^{SCN1A} не была обнаружена в периферических тканях, включая в себя сердце, легкие и гонады. Однако уровни транскрипта eTF^{SCN1A} в печени были сопоставимы с уровнями eTF^{SCN1A} , измеренными в головном мозге. Кроме того, транскрипция eTF^{SCN1A} в печени является чрезвычайно низкой, если ее нормализовать по количеству присутствующих копий вектора (приблизительно в 1000 раз ниже по сравнению с транскрипцией eTF^{SCN1A} в головном мозге). В целом это продемонстрировало, что транскрипция eTF^{SCN1A} под контролем ГАМК-селективного регуляторного элемента RE^{GABA} ограничена ЦНС.

Перечень последовательностей

<110> ИНКОУДИД ТЕРАПЬЮТИКС, ИНК.

<120> КОМПОЗИЦИИ И СПОСОБЫ СЕЛЕКТИВНОЙ РЕГУЛЯЦИИ ЭКСПРЕССИИ ГЕНОВ

<130> 46482-724.601

<140>

<141>

<150> 63/008,569

<151> 2020-04-10

<150> 62/857,727

<151> 2019-06-05

<150> 62/854,238

<151> 2019-05-29

<160> 226

<170> PatentIn version 3.5

<210> 1

<211> 259

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 1

gtgatgсgggt tttggсagta catcaatggg сgtggatagc ggtttgactc acggggattt 60

ссаagtctcc acccattga сgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac 120

tttсcaaaat gtcgтаасаа ctccgccccа ttgacгсааа tgggсggtag гсgtgtacгг 180

tgggaggtct atataagсag agctggтacc gtgtgtatgc tcaggggctg ggaaaggagг 240

ggagggagct ссggctсag 259

<210> 2

<211> 2051

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 2

ggaggaagcc atсаастааа ctасаатgac tgтааgатаc аааattggga atggтаасат 60

atттtгаagt tctgttgaca тааагаатса tgататтаат gcccatggaa atgaaagggc 120

gatсаасact atggттtgaa aagggggaaa ttgтаgаgca сagатgtgtt сgtgtggсag 180

tgtgctgtct ctagсаатac tcagагаага gagагаасаа tgaaattctg attggccccа 240

gtgtgagccc agatgagggt cagctgocaa ctttctcttt cacatcttat gaaagtcatt 300
 taagcacaac taactttttt tttttttttt tttttttgag acagagtctt gctctgttgc 360
 ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcttaggct 420
 caaacgggtcc tcttgcatac gcctcccaag tagctggaat tacaggagtg gccaccatg 480
 cccagctaata ttttgtattt ttaatagata cgggggtttc accatatac ccaggctggt 540
 ctggaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600
 ggcgtcagcc actatgccca acccgaccaa ccttttttaa aataaatatt taaaaaattg 660
 gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720
 aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag 780
 aggggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840
 tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900
 gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960
 gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020
 tactcataag accttttttt tttccttccg ggcgcaaac cgtgagctgg atttataatc 1080
 gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140
 tagctgcccc cgcgagcaac ggctcgtga tttccccgcc gatccggtcc ccgctcccc 1200
 actctgcccc cgctacccc ggagcgtgc agccgctct cgaatctct ctcttctct 1260
 ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320
 cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380
 gggagtgccc ggctcgtctg cgagagccg aggtgggtaa gctagcgacc acctggactt 1440
 cccagcggcc aaccgtggct tttcagccag gtctctctct cccgcggctt ctcaaccaac 1500
 cccatcccag cgccggccac ccaacctccc gaaatgagtg cttctgccc cagcagccga 1560
 aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620
 gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680
 gccaagtccg agggagaacg taaagatatg ggcttttttc cccctctcac cttgtctcac 1740
 caaagtccct agtccccgga gcagttagcc ttttcttttc caggaatta gccagacaca 1800
 acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgtg 1860
 ctttctctcc ctcttctctc tccagagccg gatcttcaag gggagcctcc gtgcccccg 1920
 ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgttct 1980
 ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040
 cgaccgagct g 2051

<210> 3
 <211> 1878
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 3

tcaacagggg gacacttggg aaagaaggat ggggacagag ccgagaggac tgttacacat	60
tagagaaaca tcagtgactg tgccagcttt ggggtagact gcacaaaagc cctgaggcag	120
cacaggcagg atccagtctg ctggtcccag gaagctaacc gtctcagaca gagcacaag	180
caccgagaca tgtgccaca ggcttgtgta gagaggctcag aggacagcgt acaggccca	240
gagatcaaac tcaacctcac caggcttggc agcaagcctt taccaacca cccccacc	300
accaccctg cacgcgccc tctcccctcc ccatggctc ccatggctat ctcaactggc	360
cctaaaatgt ttaaggatga cactggctgc tgagtggaaa tgagacagca gaagtcaaca	420
gtagatttta ggaaagccag agaaaaaggc ttgtgctggt tttagaaagc caagggacaa	480
gctaagatag ggcccagta atgctagtat ttacatttat ccacacaaa cggacgggcc	540
tccgctgaac cagtgaggcc ccagacgtgc gcataaataa cccctgcgtg ctgcaccacc	600
tggggagagg gggaggacca cggtaaattg agcgagcgca tagcaaaagg gacgcggggt	660
ccttttctct gccggtggca ctgggtagct gtggccagggt gtggtaacttt gatggggccc	720
agggctggag ctcaaggaag cgtcgcagggt tcacagatct gggggaacc cggggaaaag	780
cactgaggca aaaccgccc tcgtctccta caatatatgg gagggggagg ttgagtacgt	840
tctggattac tcataagacc tttttttttt ccttccgggc gcaaaaccgt gagctggatt	900
tataatcgcc ctataaagct ccagaggcgg tcaggcacct gcagaggagc cccgccgctc	960
cgccgactag ctgccccgc gagcaacggc ctcgtgattt ccccgccgat ccggtccccg	1020
cctccccact ctgccccgc ctaccgccga gccgtgcagc cgcctctccg aatctctctc	1080
ttctcctggc gctcgcgtgc gagagggaac tagcgagaac gaggaagcag ctggagggtga	1140
cgccgggcag attacgctg tcagggccga gccgagcggga tcgctgggcg ctgtgcagag	1200
gaaagggcgg agtgcccggc tcgctgtcgc agagccgagg tgggtaagct agcgaccacc	1260
tggacttccc agcgcacca cgtggctttt cagccaggct ctctcctccc gcggcttctc	1320
aaccaacccc atcccagcgc cggccacca acctcccga atgagtgctt cctgccccag	1380
cagccgaagg cgctactagg aacggtaacc tgttactttt ccaggggccg tagtcgacc	1440
gctgccccgag ttgctgtgcg actgcgcgcg cggggctaga gtgcaagggt actgtggttc	1500
ttctctggcc aagtccgagg gagaacgtaa agatatgggc ctttttcccc ctctcacctt	1560
gtctcaccaa agtccctagt ccccgagca gttagcctct ttctttccag ggaattagcc	1620

agacacaaca acgggaacca gacaccgaac cagacatgcc cgccccgtgc gccctccccg 1680
ctcgctgcct ttcctccctc ttgtctctcc agagccggat cttcaagggg agcctccgtg 1740
ccccggctg ctcagtcctt ccggtgtgca ggacccccga agtcctcccc gcacagctct 1800
cgcttctctt tgcagcctgt ttctgogccg gaccagtcga ggactctgga cagtagaggg 1860
cccgggacga ccgagctg 1878

<210> 4
<211> 509
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический
полинуклеотид

<400> 4
gccctctagg ccacctgacc aggtccctc agtccccccc ttcccacact cccacactca 60
gccccctcc ccccccccg acccctgcag gattatcctg tctgtgttcc tgactcagcc 120
tgggagccac ctgggcagca ggggccaaagg gtgtcctaga agggacctgg agtccacgct 180
gggccaagcc tgccctttct ccctctgtct tccgtccctg cttgcggttc tgctgaatgt 240
ggttatttct ctggctcctt ttacagagaa tgctgctgct aattttatgt ggagctctga 300
ggcagtgtaa ttggaagcca gacaccctgt cagcagtggg ctcccgtcct gagctgccat 360
gcttctctgct ctctcccgt cccggctcct catttcatgc agccacctgt cccagggaga 420
gaggagtcac ccaggcccct cagtccgccc cttaaataag aaagcctccg ttgctcggca 480
cacataccaa gcagccgctg gtgcaatct 509

<210> 5
<211> 1644
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический
полинуклеотид

<400> 5
cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc cccgccatt 60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120
atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc 180
aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta 240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300
catgggtcga ggtgagcccc acgttctgct tcactctccc catctcccc cctccccac 360

ccccaatttt gtatttattt attttttaat tattttgtgc agcgatgggg gcgggggggg 420
 ggggggcgcg cgccaggcgg ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga 480
 ggtgcggcgg cagccaatca gagcggcggc ctccgaaagt ttccttttat ggcgagggcg 540
 cggcggcggc ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgcct 600
 tcgccccgtg ccccgctccg cgccgcctcg cgccgcccgc cccggctctg actgaccgcg 660
 ttactcccac aggtgagcgg gcgggacggc ccttctcctc cgggctgtaa ttagcgcttg 720
 gtttaatgac ggctcgtttc ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag 780
 ggccctttgt gcggggggga gcggctcggg ggggtgcgtgc gtgtgtgtgt gcgtggggag 840
 cgccgcgtgc ggcccgcgct gcccgggcggc tgtgagcgtc gcgggcggcg cgcggggcct 900
 tgtgcgctcc gcgtgtgcgc gaggggagcg cggccggggg cggtgccccg cggtgcgggg 960
 gggctgcgag gggaaacaaag gctgcgtgcg ggggtgtgtgc gtgggggggt gagcaggggg 1020
 tgtgggcgcg gcggtcgggc tgtaaccccc ccctgcaccc ccctccccga gttgctgagc 1080
 acggccccggc ttcgggtgcg gggctcogtg cggggcgtgg cgcggggctc gccgtgccgg 1140
 gcgggggggtg gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg 1200
 gctcggggga ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cggcgagccg 1260
 cagccattgc cttttatggt aatcgtgcga gagggcgcag ggacttcctt tgtcccaaat 1320
 ctggcggagc cgaaatctgg gagggcggc cgacccccct ctacggggcg cgggcgaagc 1380
 ggtgcggcgc cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc 1440
 gtcccccttct ccatctccag cctcggggct gccgcagggg gacggctgcc ttcggggggg 1500
 acggggcagg gcggggttcg gcttctggcg tgtgaccggc ggctctagag cctctgctaa 1560
 ccatgttcat gccttcttct ttttctaca gctcctgggc aacgtgctgg ttgttgtgct 1620
 gtctcatcat tttggcaaag aatt 1644

<210> 6

<211> 1335

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 6

gagtaattca taaaaagga ctgccccctg ccttggggaa tcccagggac cgtcgttaaa 60
 ctcccactaa cgtagaaccc agagatcgct gcgttcccgc cccctcaccg gcccgctctc 120
 gtcatcactg aggtggagaa gagcatgogt gaggctccgg tgcccgtcag tgggcagagc 180
 gcacatcgcc cacagtcccc gagaagttgg ggggaggggt cggcaattga accggtgcct 240

agagaaggtg gcgcggggta aactgggaaa gtgatgtcgt gtactggctc cgcctttttc 300
 ccgaggggtg gggagaaccg tatataagtg cagtagtcgc cgtgaacggt ctttttcgca 360
 acgggtttgc cgccagaaca caggtaagtg ccgtgtgtgg ttcccgcggg cctggcctct 420
 ttacgggtta tggcccttgc gtgccttgaa ttacttccac gccctggct gcagtacgtg 480
 attcttgatc ccgagcttcg ggttggaagt ggggtgggaga gttcgaggcc ttgcgcttaa 540
 ggagcccctt cgcctcgtgc ttgagttgag gcctggcttg ggcgctgggg ccgccgcgtg 600
 cgaatctggt ggcaccttcg cgcctgtctc gctgcttttcg ataagtctct agccatttaa 660
 aatttttgat gacctgctgc gacgcttttt ttctggcaag atagtcttgt aaatgcggggc 720
 caagatctgc aactgggtat ttcggttttt ggggccgcgg gcggcgacgg ggcccgtgcg 780
 tcccagcgca catgttcggc gagggcggggc ctgcgagcgc ggccaccgag aatcggacgg 840
 gggtagtctc aagctggccg gcctgctctg gtgcctggcc tcgcgcccgc gtgtatcgcc 900
 ccgccctggg cggcaaggct ggcccggctg gcaccagttg cgtgagcggg aagatggccg 960
 cttcccggcc ctgctgcagg gagctcaaaa tggaggacgc ggcgctcggg agagcggggc 1020
 ggtgagtcac ccacacaaag gaaaagggcc tttccgtcct cagccgtcgc ttcattgtgac 1080
 tccacggagt accggggcgc gtccaggcac ctcgattagt tctcgagctt ttggagtacg 1140
 tcgtcttttag gttgggggga ggggttttat gcgatggagt ttccccacac tgagtgggtg 1200
 gagactgaag ttagggcagc ttggcacttg atgtaattct ccttgggaatt tgcccttttt 1260
 gagtttgat cttggttcat tctcaagcct cagacagtgg ttcaaagttt ttttcttcca 1320
 tttcaggtgt cgtga 1335

<210> 7

<211> 214

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 7

aaagagaccg gttcactgtg acagtaaaag agaccggttc actgtgagaa tgaaagagac 60
 cggttcactg tgatcggaaa agagaccggt tcaactgtgag cggccttgaa acccagcaga 120
 caatgtagct cagtagaaaac ccagcagaca atgtagctga atggaaacc agcagacaat 180
 gtagcttcgg agaaaccag cagacaatgt agct 214

<210> 8

<211> 21

<212> ПНК

<213> Homo sapiens

<400> 8 усасагуаа ссггусисии и	21
<210> 9 <211> 21 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический олигонуклеотид	
<400> 9 ааагагассг гттсактгтг а	21
<210> 10 <211> 23 <212> РНК <213> Homo sapiens	
<400> 10 агсиасаиуг усигсигггу ииц	23
<210> 11 <211> 23 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический олигонуклеотид	
<400> 11 гааасссагс агасаатгта гсг	23
<210> 12 <211> 23 <212> РНК <213> Homo sapiens	
<400> 12 агсиасауци ггсиасиггг уци	23
<210> 13 <211> 23 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический олигонуклеотид	
<400> 13 агаассгагта гссгагатгта гсг	23
<210> 14 <211> 67	

<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	Описание искусственной последовательности: Синтетический олигонуклеотид	
<400>	14	
	gaaacccagc agacaatgta gctagaccsa gtagccagat gtagctaaag agaccggttc	60
	actgtga	67
<210>	15	
<211>	134	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	Описание искусственной последовательности: Синтетический полинуклеотид	
<400>	15	
	gaaacccagc agacaatgta gctagaccsa gtagccagat gtagctaaag agaccggttc	60
	actgtgagaa acccagcaga caatgtagct agaccagta gccagatgta gctaaagaga	120
	ccggttcact gtga	134
<210>	16	
<211>	62	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	Описание искусственной последовательности: Синтетический олигонуклеотид	
<400>	16	
	aataaaagat ctttatcttc attagatctg tgtggtgggt ttttgtgtgc ggaccgcacg	60
	tg	62
<210>	17	
<211>	477	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	Описание искусственной последовательности: Синтетический полинуклеотид	
<400>	17	
	gggtggcatc cctgtgaccs ctccccagtg cctctcctgg ccctggaagt tgccactcca	60
	gtgcccacca gccttgtcct aataaaatta agttgcatca ttttgtctga ctaggtgtcc	120
	ttctataata ttatggggtg gaggggggtg gtatggagca aggggcaagt tgggaagaca	180
	acctgtaggg cctgcggggt ctattgggaa ccaagctgga gtgcagtggc acaatcttgg	240

046157

ctcactgcaa tctccgcctc ctgggttcaa gcgattctcc tgcctcagcc tcccagattg	300
ttgggattcc aggcattgat gaccaggctc agctaatttt tgtttttttg gtagagacgg	360
ggtttcacca tattggccag gctgggtctcc aactcctaata ctcaggatgat ctaccacact	420
tggcctccca aattgctggg attacaggcg tgaaccactg ctcccttccc tgtcctt	477
<210> 18	
<211> 18	
<212> ДНК	
<213> Homo sapiens	
<400> 18	
ctaggtcaag tgtaggag	18
<210> 19	
<211> 18	
<212> ДНК	
<213> Homo sapiens	
<400> 19	
acttgacctg gacagcct	18
<210> 20	
<211> 18	
<212> ДНК	
<213> Homo sapiens	
<400> 20	
tgaataactc attagtga	18
<210> 21	
<211> 18	
<212> ДНК	
<213> Homo sapiens	
<400> 21	
aaagtacatt aggctaata	18
<210> 22	
<211> 18	
<212> ДНК	
<213> Homo sapiens	
<400> 22	
ccagcactgg tgcttcgt	18
<210> 23	
<211> 18	
<212> ДНК	
<213> Homo sapiens	
<400> 23	
aaggctgtct aggtcaag	18

<210> 24
 <211> 24
 <212> ДНК
 <213> Homo sapiens

 <400> 24
 ctaggtcaag tgtaggagac acac 24

<210> 25
 <211> 18
 <212> ДНК
 <213> Homo sapiens

 <400> 25
 ggtcaagtgt aggagaca 18

<210> 26
 <211> 18
 <212> ДНК
 <213> Homo sapiens

 <400> 26
 caagtgtagg agacacac 18

<210> 27
 <211> 24
 <212> ДНК
 <213> Homo sapiens

 <400> 27
 agtgtaggag acacactgct ggcc 24

<210> 28
 <211> 18
 <212> ДНК
 <213> Homo sapiens

 <400> 28
 agtgtaggag acacactg 18

<210> 29
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 29
 aggagacaca ctgctggcct g 21

<210> 30
 <211> 18
 <212> ДНК
 <213> Homo sapiens

 <400> 30
 taggtaccat agagtgag 18

<210> 31
 <211> 18
 <212> ДНК
 <213> Homo sapiens

 <400> 31
 gaggatactg cagaggtc 18

 <210> 32
 <211> 27
 <212> ДНК
 <213> Homo sapiens

 <400> 32
 taggtaccat agagtgagggc gaggatg 27

 <210> 33
 <211> 27
 <212> ДНК
 <213> Homo sapiens

 <400> 33
 atagagtgag gcgaggatga agccgag 27

 <210> 34
 <211> 27
 <212> ДНК
 <213> Homo sapiens

 <400> 34
 tgaagccgag aggatactgc agaggtc 27

 <210> 35
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 35
 aaggctgtct aggtcaagtg t 21

 <210> 36
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 36
 tgttcctcca gattaacact t 21

 <210> 37
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 37
 gatgaagccg agaggatact g 21

<210> 38
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 38
 gctgatttgt attaggtacc a 21

<210> 39
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 39
 agaaagctga tacagataca a 21

<210> 40
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 40
 ggtacgggca aagatttctt g 21

<210> 41
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 41
 agaaagctga tacagataca a 21

<210> 42
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 42
 acacaatgag ccacctaca g 21

<210> 43
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 43
 gtggctcatt gtgtgtgtgc c 21

<210> 44
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 44
 catatccctg caggttcaga a 21

<210> 45
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 45
 agagagagag agagagagag a 21

<210> 46
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 46
 ttctcagttt tgaaattaa a 21

<210> 47
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 47
 tggattctct tctgaacctg c 21

<210> 48
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 48
 tgctgaggca ggacacagtg t 21

<210> 49
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 49
 atcatctgta accatcaagg a 21

<210> 50
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 50
 tcctgcctac ttagtttcaa g 21

<210> 51
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 51
 attacagttc tgtcagcatg c 21

<210> 52
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 52
 tgggtctcatt ctttttgtgg g 21

<210> 53
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 53
 cgatattttc atggattcct t 21

<210> 54
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 54
 ctgacactta ctttgtctaa a 21

<210> 55
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 55
 aaaactggaa ccgcattccc a 21

<210> 56
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 56
 acaaagtaag tgtcagtg g 21

<210> 57
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 57
 ataatagttg tgtctttata a 21

<210> 58
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 58
 tgtacaagca gggctgcaaa g 21

<210> 59
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 59
 gttaacaaat aactaaaca c 21

<210> 60
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 60
 ttcaacaagc tccaagaag t 21

<210> 61
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 61
 atggtcaagg tgcagaagga a 21

<210> 62
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 62
 tgtttgctca aacgtgcacc a 21

<210> 63
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 63
 aaataagaca tgaaaacaag a 21

<210> 64
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 64
 aaatatgtac cacaagaat g 21

<210> 65
 <211> 21
 <212> ДНК
 <213> Homo sapiens

 <400> 65
 tatctggttt ctctcactgc t 21

<210> 66
 <211> 21
 <212> ДНК
 <213> Homo sapiens

<400> 66
 attgcaaagc ataatttgga t 21

<210> 67
 <211> 3585
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический
 полинуклеотид

<400> 67
 ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60
 attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120
 gatcaacact atggtttgaa aagggggaaa ttgtagagca cagatgtggt cgtgtggcag 180
 tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240
 gtgtgagccc agatgagggt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300
 taagcacaac taactttttt tttttttttt tttttttgag acagagtctt gctctgttgc 360
 ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctccataggt 420
 caaacgggtcc tcttgcatac gcctcccaag tagctggaat tacaggagtg gccaccatg 480
 cccagctaат ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540
 ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600
 ggcgtcagcc actatgcccа acccgaccaa ccttttttaa aataaatatt taaaaaattg 660
 gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720
 aaccagtgag gcccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag 780
 agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840
 tctgccggtg gcaactgggt gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900
 gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960
 gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020
 tactcataag accttttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080
 gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140
 tagctgcccc cgcgagcaac ggctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200
 actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260
 ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380
gggagtgcc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440
cccagcgcgc aaccgtggct tttcagccag gtccctctct cccgcggctt ctcaaccaac 1500
cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560
aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620
gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680
gccaagtccg agggagaacg taaagatatg ggocctttttc cccctctcac cttgtctcac 1740
caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800
acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgcctcc ccgctcgctg 1860
cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccg 1920
ctgctcagtc cctccgggtg gcaggacccc ggaagtctc cccgcacagc tctcgcttct 1980
ctttgcagcc tgtttctgcy ccggaccagt cgaggactct ggacagtaga ggccccggga 2040
cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100
gagtcccagc agccctcgaa ccagggtgaaa aaccttaca atgtcctgaa tgtgggaaat 2160
cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220
ataaatgtcc agaatgtgga aagtcttct cagcagagga taacttgac actcatcaac 2280
gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340
gcgatgaact tgtccgacac caacgaacct atacaggcga gaagccttac aaatgtcccg 2400
agtgtggcaa gagcttctca caatcagggg atctgactga gcatcaacga actcataccg 2460
gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520
tacgccacca gaggacacat acaggggaga agccctaca atgccccgaa tgcggtaaaa 2580
gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640
gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700
catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760
tgggttctga tgccctcgat gactttgacc tggatatggt gggaagcgac gcattggatg 2820
actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880
aaactagtaa agagaccggt tcaactgtgac agtaaaagag accggttcac tgtgagaatg 2940
aaagagaccg gttcaactgtg atcggaaaag agaccggttc actgtgagcg gccttgaaac 3000
ccagcagaca atgtagctca gtagaaaacc agcagacaat gtagctgaat ggaaaccag 3060
cagacaatgt agcttcggag aaaccagca gacaatgtag ctaagcttgg gtggcatccc 3120
tgtgaccct ccccagtgcc tctcctggcc ctggaagttg cactccagt gccaccagc 3180
cttgcctaa taaaattaag ttgcatcatt ttgtctgact aggtgtcctt ctataatatt 3240

atgggggtgga ggggggtggt atggagcaag gggcaagttg ggaagacaac ctgtagggcc 3300
 tgcgggggtct attgggaacc aagctggagt gcagtggcac aatcttggct cactgcaatc 3360
 tccgcctcct gggttcaagc gattctcctg cctcagcctc ccgagttggt gggattccag 3420
 gcatgcatga ccaggctcag ctaatTTTTTg tTTTTTggt agagacgggg tttcaccata 3480
 ttggccaggc tggctcctaa ctctaactc caggtgatct acccaccttg gcctccaaa 3540
 ttgctgggat tacaggcgtg aaccactgct cccttccttg tcctt 3585

<210> 68

<211> 3371

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический
 полинуклеотид

<400> 68

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60
 attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120
 gatcaacact atggtttgaa aagggggaaa ttgtagagca cagatgtggt cgtgtggcag 180
 tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaatctctg attggcccca 240
 gtgtgagccc agatgagggt cagctgcca ctttctcttt cacatcttat gaaagtcatt 300
 taagcacaac taactTTTTT tTTTTTTTT tTTTTTgag acagagtctt gctctgttgc 360
 ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctctaaggct 420
 caaacgggtc tcctgcatca gcctcccaag tagctggaat tacaggagtg gccaccatg 480
 cccagctaat ttttgtatTT ttaatagata cgggggtttc accatatcac ccaggctggt 540
 ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600
 ggcgtcagcc actatgcca acccgaccaa ctttttttaa aataaatatt taaaaaattg 660
 gtatttcaca tatatactag tattttacatt tatccacaca aaacggacgg gcctccgctg 720
 aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag 780
 agggggagga ccacggtaaa tggagcagc gcatagcaaa agggacgcgg ggtccttttc 840
 tctgccggtg gcaactgggt gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900
 gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960
 gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020
 tactcataag acctTTTTTT tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080
 gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccc ctccgcccac 1140
 tagctgcccc cgcgagcaac ggctcgtga tttccccgcc gatccgggtc ccgcctcccc 1200

actctgcccc	cgcctacccc	ggagccgtgc	agccgcctct	ccgaatctct	ctcttctcct	1260
ggcgctcgcg	tgcgagaggg	aactagcgag	aacgaggaag	cagctggagg	tgacgccggg	1320
cagattacgc	ctgtcagggc	cgagccgagc	ggatcgctgg	gcgctgtgca	gaggaaaggc	1380
gggagtgccc	ggctcgctgt	cgcagagccg	aggtgggtaa	gctagcgacc	acctggactt	1440
cccagcgcgc	aaccgtggct	tttcagccag	gtcctctcct	cccgcggctt	ctcaaccaac	1500
cccatcccag	cgccggccac	ccaacctccc	gaaatgagtg	cttcctgccc	cagcagccga	1560
aggcgctact	aggaacggta	acctgttact	tttccagggg	ccgtagtcga	cccgtgccc	1620
gagttgctgt	gcgactgcgc	gcgcggggct	agagtgcaag	gtgactgtgg	ttcttctctg	1680
gccaagtccg	agggagaacg	taaagatatg	ggcctttttc	ccctctcac	cttgtctcac	1740
caaagtccct	agtccccgga	gcagttagcc	tctttctttc	cagggaaatta	gccagacaca	1800
acaacgggaa	ccagacaccg	aaccagacat	gcccgcctcc	tgcgccctcc	ccgctcgctg	1860
cctttcctcc	ctcttgtctc	tccagagccg	gatcttcaag	gggagcctcc	gtgcccccg	1920
ctgctcagtc	cctccggtgt	gcaggacccc	ggaagtcctc	cccgcacagc	tctcgcttct	1980
ctttgcagcc	tgtttctgcg	ccggaccagt	cgaggactct	ggacagtaga	ggccccggga	2040
cgaccgagct	ggaattcgcc	accatggccc	caaagaagaa	gcggaaggtc	ggtatccacg	2100
gagtcccagc	agccctcgaa	ccaggtgaaa	aaccttacia	atgtcctgaa	tgtgggaaat	2160
cattcagtcg	cagcgacaac	ctggtgagac	atcaacgcac	ccatacagga	gaaaaacctt	2220
ataaatgtcc	agaatgtgga	aagtccttct	cacgagagga	taacttgcac	actcatcaac	2280
gaacacatac	tggtgaaaaa	ccatacaagt	gtcccgaatg	tggtaaaagt	tttagccgga	2340
gcgatgaact	tgtccgacac	caacgaacct	atacaggcga	gaagccttac	aaatgtcccg	2400
agtgtggcaa	gagcttctca	caatcagggg	atctgactga	gcatcaacga	actcataccg	2460
gggaaaaacc	ttacaagtgt	ccagagtgtg	ggaagagctt	ttccacaagt	ggacatctgg	2520
tacgccacca	gaggacacat	acaggggaga	agccctacia	atgccccgaa	tgcggtaaaa	2580
gtttctctca	gaatagtacc	ctgaccgaac	accagcgaac	acacactggg	aaaaaaacga	2640
gtaaaaggcc	ggcggccacg	aaaaaggccg	gccaggcaaa	aaagaaaaag	ggatcctacc	2700
catacgacgt	accagattac	gctctcgagg	acgcgctgga	cgatttcgat	ctcgacatgc	2760
tgggttctga	tgccctcgat	gactttgacc	tggatatggt	gggaagcgac	gcattggatg	2820
actttgatct	ggacatgctc	ggctccgatg	ctctggacga	tttcgatctc	gatatgttat	2880
aaactagtaa	gcttgggtgg	catccctgtg	acccctcccc	agtgcctctc	ctggccctgg	2940
aagttgccac	tccagtgccc	accagccttg	tctaataaaa	attaagttgc	atcattttgt	3000
ctgactaggt	gtccttctat	aatattatgg	ggtggagggg	ggtggtatgg	agcaaggggc	3060

046157

aagttgggaa gacaacctgt agggcctgcg gggctctattg ggaaccaagc tggagtgcag 3120
 tggcacaatc ttggctcact gcaatctccg cctcctgggt tcaagcgatt ctctgcctc 3180
 agcctcccga gttgttggga ttccaggcat gcatgaccag gctcagctaa tttttgtttt 3240
 tttggtagag acgggggttc accatattgg ccaggctggc ctccaactcc taatctcagg 3300
 tgatctaccc accttggcct cccaaattgc tgggattaca ggcgtgaacc actgctccct 3360
 tccctgtcct t 3371

<210> 69

<211> 4380

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический
 полинуклеотид

<400> 69

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60
 attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120
 gatcaacact atggtttgaa aagggggaaa ttgtagagca cagatgtggt cgtgtggcag 180
 tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240
 gtgtgagccc agatgagggt cagctgcca ctttctcttt cacatcttat gaaagtcatt 300
 taagcacaac taactttttt tttttttttt tttttttgag acagagtctt gctctgttgc 360
 ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcttaggct 420
 caaacgggtc tctgcatca gcctccaag tagctggaat tacaggagtg gccaccatg 480
 cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggc 540
 ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600
 ggcgtcagcc actatgcca acccgaccaa ctttttttaa aataaatatt taaaaattg 660
 gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720
 aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag 780
 agggggagga ccacggtaaa tggagcagc gcatagcaaa agggacgagg ggtccttttc 840
 tctgccggtg gcaactggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900
 gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960
 gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020
 tactcataag accttttttt tttccttccg ggcgcaaac cgtgagctgg atttataatc 1080
 gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccc ctccgcccac 1140
 tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccgggtc ccgcctcccc 1200

actctgcccc	cgctaccccc	ggagccgtgc	agccgcctct	ccgaatctct	ctcttctcct	1260
ggcgctcgcg	tgcgagaggg	aactagcgag	aacgaggaag	cagctggagg	tgacgccggg	1320
cagattacgc	ctgtcagggc	cgagccgagc	ggatcgctgg	gcgctgtgca	gaggaaaaggc	1380
gggagtgccc	ggctcgctgt	cgcagagccg	aggtgggtaa	gctagcgacc	acctggactt	1440
cccagcgcgc	aaccgtggct	tttcagccag	gtcctctcct	cccgcggctt	ctcaaccaac	1500
cccatcccag	cgccggccac	ccaacctccc	gaaatgagtg	cttcctgccc	cagcagccga	1560
aggcgctact	aggaacggta	acctgttact	tttccagggg	ccgtagtcga	cccgtgccc	1620
gagttgctgt	gcgactgcgc	gcgcggggct	agagtgcgaag	gtgactgtgg	ttcttctctg	1680
gccaagtccg	agggagaacg	taaagatatg	ggcctttttc	cccctctcac	cttgtctcac	1740
caaagtccct	agtccccgga	gcagttagcc	tctttctttc	caggaatta	gccagacaca	1800
acaacgggaa	ccagacaccg	aaccagacat	gcccgcctcc	tgcgccctcc	ccgctcgctg	1860
cctttcctcc	ctcttgtctc	tccagagccg	gatcttcaag	gggagcctcc	gtgcccccg	1920
ctgctcagtc	cctccggtgt	gcaggacccc	ggaagtcctc	cccgcacagc	tctcgcttct	1980
ctttgcagcc	tgtttctgcg	ccggaccagt	cgaggactct	ggacagtaga	ggccccggga	2040
cgaccgagct	ggaattcgcc	accatggccc	caaagaagaa	gcggaaggtc	ggtatccacg	2100
gagtcccagc	agccctcgaa	ccaggtgaaa	aaccttacia	atgtcctgaa	tgtgggaaat	2160
cattcagtcg	cagcgacaac	ctggtgagac	atcaacgcac	ccatacagga	gaaaaacctt	2220
ataaatgtcc	agaatgtgga	aagtccttct	cacgagagga	taacttgcac	actcatcaac	2280
gaacacatac	tggtgaaaaa	ccatacaagt	gtcccgaatg	tggtaaaagt	tttagccgga	2340
gcgatgaact	tgtccgacac	caacgaacct	atacaggcga	gaagccttac	aaatgtcccg	2400
agtgtggcaa	gagcttctca	caatcagggg	atctgactga	gcatcaacga	actcataccg	2460
gggaaaaacc	ttacaagtgt	ccagagtgtg	ggaagagctt	ttccacaagt	ggacatctgg	2520
tacgccacca	gaggacacat	acaggggaga	agccctacia	atgccccgaa	tgcggtaaaa	2580
gtttctctca	gaatagtacc	ctgaccgaac	accagcgaac	acacactggg	aaaaaaacga	2640
gtaaaaggcc	ggcggccacg	aaaaaggccg	gccaggcaaa	aaagaaaaag	ggatcctacc	2700
catacgacgt	accagattac	gctctogagg	aggccagcgg	ttccggacgg	gctgacgcat	2760
tggacgattt	tgatctggat	atgctgggaa	gtgacgccct	cgatgatttt	gacctgaca	2820
tgcttggttc	ggatgccctt	gatgactttg	acctogacat	gctcggcagt	gacgcccttg	2880
atgatttcga	cctggacatg	ctgattaact	ctagaagttc	cggatctccg	aaaaagaaac	2940
gcaaagttag	tagccagtac	ctgcccgaca	ccgacgaccg	gcaccggatc	gaggaaaagc	3000
ggaagcggac	ctacgagaca	ttcaagagca	tcatgaagaa	gtcccccttc	agcggcccca	3060
ccgaccctag	acctccacct	agaagaatcg	ccgtgccccag	cagatccagc	gccagcgtgc	3120

caaaacctgc cccccagcct tacccttca ccagcagcct gagcaccatc aactacgacg 3180
 agttccctac catggtgttc cccagcggcc agatctctca ggcctctgct ctggctccag 3240
 cccctcctca ggtgctgct caggctcctg ctctgcacc agctccagcc atggtgtctg 3300
 cactggctca ggcaccagca cccgtgctg tgctggctcc tggacctca caggctgtgg 3360
 ctccaccagc ccctaaacct acacaggccg gcgagggcac actgtctgaa gctctgctgc 3420
 agctgcagtt cgacgacgag gatctgggag ccctgctggg aaacagcacc gatcctgccg 3480
 tgttcaccga cctggccagc gtggacaaca gcgagttcca gcagctgctg aaccagggca 3540
 tcctgtggc ccctcacacc accgagccca tgctgatgga ataccccgag gccatcacc 3600
 ggctcgtgac aggcgctcag aggcctcctg atccagctcc tgcccctctg ggagcaccag 3660
 gcctgcctaa tggactgctg tctggcgacg aggacttcag ctctatcgcc gatatggatt 3720
 tctcagcctt gctgggctct ggcagcggca gccgggattc caggggaaggg atgtttttgc 3780
 cgaagcctga ggcgggctcc gctattagtg acgtgtttga gggccgcgag gtgtgccagc 3840
 caaaacgaat ccggccattt catcctccag gaagtccatg ggccaaccgc cactccccg 3900
 ccagcctcgc accaacacca accggtccag tacatgagcc agtcgggtca ctgacccccg 3960
 caccagtccc tcagccactg gatccagcgc ccgcagtgac tcccgaggcc agtcacctgt 4020
 tggaggatcc cgatgaagag acgagccagc ctgtcaaagc ccttcgggag atggccgata 4080
 ctgtgattcc ccagaaggaa gaggctgcaa tctgtggcca aatggacctt tccatccgc 4140
 ccccaagggg ccatctggat gagctgacaa ccacacttga gtccatgacc gaggatctga 4200
 acctggactc acccctgacc ccggaattga acgagattct ggataccttc ctgaacgacg 4260
 agtgccctt gcatgccatg catatcagca caggactgtc catcttcgac acatctctgt 4320
 tttaaactag taataaaaga tctttatctt cattagatct gtgtgttggt tttttgtgtg 4380

<210> 70

<211> 3332

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 70

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60
 attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120
 gatcaacact atggtttgaa aagggggaaa ttgtagagca cagatgtggt cgtgtggcag 180
 tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240
 gtgtgagccc agatgagggt cagctgcca ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taactttttt tttttttttt tttttttgag acagagtctt gctctgttgc	360
ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcttaggct	420
caaacggtcc tcttgcatac gcctcccaag tagctggaat tacaggagtg gccaccatg	480
cccagctaata ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt	540
ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata	600
ggcgtcagcc actatgccca acccgaccaa ccttttttaa aataaatatt taaaaaattg	660
gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg	720
aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag	780
agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc	840
tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg	900
gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag	960
gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat	1020
tactcataag accttttttt tttccttccg ggcgcaaac cgtgagctgg atttataatc	1080
gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac	1140
tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccgggtcc ccgcctcccc	1200
actctgcccc cgctacccc ggagcctgac agccgcctct ccgaatctct ctcttctcct	1260
ggcgtcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg	1320
cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc	1380
gggagtgcc ggctcgtgt cgagagccg aggtgggtaa gctagcgacc acctggactt	1440
cccagcgcgc aaccgtggct tttcagccag gtccctcctc cccgcggctt ctcaaccaac	1500
cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga	1560
aggcgtact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc	1620
gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg	1680
gccaagtccg agggagaacg taaagatatg ggcctttttc cccctctcac cttgtctcac	1740
caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca	1800
acaacgggaa ccagacaccg aaccagacat gccgccccg tgcgcctcc ccgctcgtg	1860
cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccg	1920
ctgctcagtc cctccggtgt gcaggacccc ggaagtctc cccgcacagc tctcgcttct	1980
ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga	2040
cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg	2100
gagtcccagc agccctcgaa ccagggtgaaa aaccttaca atgtcctgaa tgtgggaaat	2160

cattcagtcg cagcgcacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220
 ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgac actcatcaac 2280
 gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340
 gcgatgaact tgtccgacac caacgaacc atacaggcga gaagccttac aaatgtcccg 2400
 agtgtggcaa gagcttctca caatcagggg atctgactga gcatcaacga actcataccg 2460
 gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520
 tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580
 gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640
 gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatccgacg 2700
 cgctggacga tttcgatctc gacatgctgg gttctgatgc cctcgatgac tttgacctgg 2760
 atatgttggg aagcgcgcga ttggatgact ttgatctgga catgctcggc tccgatgctc 2820
 tggacgattt cgatctcgat atgttataaa agcttgggtg gcatccctgt gaccctccc 2880
 cagtgcctct cctggccctg gaagttgcca ctccagtgcc caccagcctt gtcctaataa 2940
 aattaagttg catcattttg tctgactagg tgtccttcta taatattatg ggggtggaggg 3000
 ggggtggtatg gagcaagggg caagttggga agacaacctg tagggcctgc ggggtctatt 3060
 gggaaaccaag ctggagtgca gtggcacaat cttggctcac tgcaatctcc gcctcctggg 3120
 ttcaagcgat tctcctgcct cagcctcccg agttggtggg attccaggca tgcattgacca 3180
 ggctcagcta atttttgttt ttttggtaga gacggggttt caccatattg gccaggctgg 3240
 tctccaactc ctaatctcag gtgatctacc caccttggcc tcccaaattg ctgggattac 3300
 aggcgtgaac cactgctccc ttccctgtcc tt 3332

<210> 71

<211> 3546

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 71

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60
 attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120
 gatcaacact atggtttgaa aagggggaaa ttgtagagca cagatgtggt cgtgtggcag 180
 tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240
 gtgtgagccc agatgagggt cagctgocaa ctttctcttt cacatcttat gaaagtcatt 300
 taagcacaac taactttttt tttttttttt tttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctctaggct 420
caaacgggtcc tcttgcatac gcctcccaag tagctggaat tacaggagtg gccaccatg 480
cccagctaata ttttgtatct ttaatagata cgggggtttc accatatcac ccaggctggt 540
ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600
ggcgtcagcc actatgcccc acccgaccaa ccttttttaa aataaatatt taaaaaattg 660
gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720
aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag 780
agggggagga ccacggtaaa tggagcagac gcatagcaaa agggacgcgg ggtccttttc 840
tctgccggtg gcaactggta gctgtggcca ggtgtggtac tttgatgggg ccaggggctg 900
gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960
gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020
tactcataag accttttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080
gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccc ctccgcccac 1140
tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200
actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260
ggcgtcgcg tgcgagaggg aactagcag aacgaggaag cagctggag tgcgcccggg 1320
cagattacgc ctgtcagggc cgagccagac ggatcgctgg gcgctgtgca gagaaaaggc 1380
gggagtgcc ggctcgtgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440
cccagcggcc aaccgtggct tttcagccag gtccctctct cccgcggctt ctcaaccaac 1500
cccatcccag cgcgcccac ccaacctccc gaaatgagtg cttcctgccc cagcagcccga 1560
aggcgtact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620
gagttgctgt gcgactgccc gcgccccgct agagtcaag gtgactgtgg ttcttctctg 1680
gccaagtccg agggagaacg taaagatatg ggcctttttc cccctctcac cttgtctcac 1740
caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800
acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgcccctc ccgctcgtg 1860
cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgccccgg 1920
ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980
ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040
cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100
gagtcccagc agccctcgaa ccagggtgaaa aaccttaca atgtcctgaa tgtgggaaat 2160
cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220
ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340
 gcgatgaact tgtccgacac caacgaacco atacaggcga gaagccttac aaatgtcccg 2400
 agtgtggcaa gagcttctca caatcagggga atctgactga gcatcaacga actcataccg 2460
 gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520
 tacgccacca gaggacacat acaggggaga agccctaca atgccccgaa tgcggtaaaa 2580
 gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640
 gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatccgacg 2700
 cgctggacga tttcgatctc gacatgctgg gttctgatgc cctcgatgac tttgacctgg 2760
 atatgttggg aagcgacgca ttggatgact ttgatctgga catgctcggc tccgatgctc 2820
 tggacgattt cgatctcgat atgttataaa aagagaccgg ttcactgtga cagtaaaaga 2880
 gaccggttca ctgtgagaat gaaagagacc ggttactgt gatcggaaaa gagaccggtt 2940
 cactgtgagc ggccttgaaa cccagcagac aatgtagctc agtagaaacc cagcagacaa 3000
 tgtagctgaa tggaaaccca gcagacaatg tagcttcgga gaaaccagc agacaatgta 3060
 gctaagcttg ggtggcatcc ctgtgacccc tcccagtgct ctctcctggc cctggaagtt 3120
 gccactccag tgcccaccag ccttgtccta ataaaattaa gttgcatcat tttgtctgac 3180
 taggtgtcct tctataatat tatgggggtg aggggggtgg tatggagcaa ggggcaagtt 3240
 gggaaagaaa cctgtagggc ctgccccgtc tattgggaac caagctggag tgcagtggca 3300
 caatcttggc tcaactgcaat ctccgcctcc tgggttcaag cgattctcct gcctcagcct 3360
 cccgagttgt tgggattcca ggcatgcatg accaggctca gctaattttt gtttttttgg 3420
 tagagacggg gtttcacat attggccagg ctggtctcca actcctaata tcaggtgatc 3480
 taccacctt ggctcccaa attgctggga ttacaggcgt gaaccactgc tcccttcct 3540
 gtcctt 3546

<210> 72

<211> 1707

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 72

atggccgag atcacctgat gctggctgaa ggctacagac tggcgcagcg gcctccatct 60
 gccgctgccg cccacggccc ccacgcctg agaaactgc cccctacgc cggccctgg 120
 cttgatagcg gactcagacc tagaggcgcc cctctgggccc ctccacctcc aagacagcct 180
 ggagccctgg cctacggcgc cttcggccct cctctagct tccagccctt ccccgccgtg 240

cctcctccag ccgctggcat cgcccacctg cagcctgtgg ccaccctta ccccggaaga 300
 gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360
 gctgcccctc cacctcctgc tcatgccttg ggcggcatgg acgccgagct gatcgacgag 420
 gaagccctga ccagcctgga actggaactg ggctgcaca gagtgcggga actgcctgag 480
 ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540
 ggctctgtgt cctgcgccga ccacctgatg ctcgccgagg gctaccgcct ggtgcagagg 600
 ccgccgtccg ccgccgcgc ccattggcct catgcgctcc ggactctgcc gccgtacgcg 660
 ggccccggcc tggacagtgg gctgaggccg cggggggctc cgctggggcc gccgccgccc 720
 cgccaaccgg gggccctggc gtacggggcc ttcgggcccgc cgtcctcctt ccagcccttt 780
 ccggccgtgc ctccgccggc cgcgggcac cgcacactgc agcctgtggc gacgccgtac 840
 cccggccgcg ccgccgcgc ccccaacgct ccgggaggcc ccccgggccc gcagccggcc 900
 ccaagcgccg cagccccgcc gccgcccgcg cacgccttg gcgcatgga cgccgaactc 960
 atcgacgagg aggcgctgac gtcgctggag ctggagctgg ggctgcaccg cgtgcgcgag 1020
 ctgcccgagc tgttcctggg ccagagcgag ttcgactgct tctcggactt ggggtccgcg 1080
 ccgcccgccg gctccgtgag ctgccagtcc cagctcatca aaccagccg catgcgcaag 1140
 taccccaacc ggcccagcaa gacgcccccc cacgaacgcc cttacgcttg cccagtggag 1200
 tcctgtgac gccgcttctc ccgcagcgac aacctggtga gacacatccg catccacaca 1260
 ggccagaagc ccttcagtg ccgcatctgc atgagaaact tcagccgaga ggataacttg 1320
 cacactcaca tccgcaccca cacaggcgaa aagcccttcg cctgcgacat ctgtggaaga 1380
 aagtttgccc ggagcgatga acttgctcga cataccaaga tccacttgcg gcagaaggac 1440
 cgcccttacg cttgccagtg ggagtctgt gatcgccgct tctcccaatc agggaatctg 1500
 actgagcaca tccgcatcca cacaggccag aagcccttcc agtgccgcat ctgcatgaga 1560
 aacttcagca caagtggaca tctggtagc cacatccgca cccacacagg cgaaaagccc 1620
 ttcgctgcg acatctgtgg aagaaagttt gcccagaata gtaccctgac cgaacatacc 1680
 aagatccact tgcggcagaa ggacaag 1707

<210> 73

<211> 1755

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 73

atggccgag atcacctgat gctggctgaa ggctacagac tgggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgcacctg agaacactgc ccccctacgc cggccctggt 120
cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180
ggagccctgg cctacggcgc ctteggccct ccttctagct tccagccctt ccccgccgtg 240
cctcctccag ctgctggcat cgcccacctg cagcctgtgg ccaccctta ccccggaaga 300
gccgcccgcc ctccaaacgc ccctggcgga cctcctggcc ccagcctgc tccaagcgcc 360
gctgcccctc cacctcctgc tcatgccttg ggcggcatgg acgcccagct gatcgacgag 420
gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480
ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540
ggctctgtgt cctgcgggcg cagcggcgcc ggaagcggcg ccgaccacct gatgctcgcc 600
gagggctacc gcctggtgca gaggcgcgcg tccgcgcgcg ccgcccattg ccctcatgcg 660
ctccggactc tgccgcgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg 720
gctccgctgg ggccgcgcgc gcccgcgcaa cccggggccc tggcgtagcg ggccttcggg 780
ccgccgtcct ccttcagcc ctttcggcc gtgcctccgc cggccgcggg catcgcgcac 840
ctgcagcctg tggcgacgcc gtaccccgcc cgcgcggccg cccccccaa cgctccggga 900
ggcccccccg gcccgagcc ggccccaaag gccgcagccc cgcgcgcgcc cgcgcacgcc 960
ctggggcgga tggacgcga actcatcgac gaggaggcgc tgacgtcgct ggagctggag 1020
ctggggctgc accgcgtgcg cgagctgccc gagctgttcc tgggcccagag cgagttcgac 1080
tgcttctcgg acttggggtc cgcgcgcgcc gccggctccg tgagctgcgg tggttctggt 1140
ggtggttctg gtcagtccca gctcatcaaa cccagccgca tgcgcaagta cccaaccgg 1200
cccagcaaga cccccccca cgaacgcct tacgcttgcc cagtggagtc ctgtgatcgc 1260
cgcttctccc gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc 1320
ttccagtgcc gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc 1380
cgcaccaca caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg 1440
agcgatgaac ttgtccgaca taccaagatc cacttgccggc agaaggaccg cccttacgct 1500
tgcccagtgg agtctgtgga tcgcccgttc tcccaatcag ggaatctgac tgagcacatc 1560
cgcatccaca caggccagaa gcccttcag tgcgcacatc gcatgagaaa cttcagcaca 1620
agtggacatc tggtagcca catccgcacc cacacaggcg aaaagccctt cgcctgcgac 1680
atctgtggaa gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg 1740
cggcagaagg acaag 1755

<210> 74

<211> 3438

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 74

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat	60
atdddgaagt tctgdtgaca taaagaatca tgatattaat gcccatggaa atgaaagggc	120
gatcaacact atggdttgaa aagggggaaa ttgtagagca cagatgtgtt cgtgtggcag	180
tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca	240
gtgtgagccc agatgaggtt cagctgcca ctttctcttt cacatcttat gaaagtcatt	300
taagcacaac taactttttt tttttttttt tttttttgag acagagtctt gctctgttgc	360
ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcttaggct	420
caaacgggcc tctgcatca gcctcccaag tagctggaat tacaggagtg gccaccatg	480
cccagctaat ttttgtattd ttaatagata cgggggtttc accatatcac ccaggctggt	540
ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata	600
ggcgtcagcc actatgccca acccgaccaa ctttttttaa aataaatatt taaaaattg	660
gtattdcaca tatatactag tattttacatt tatccacaca aaacggacgg gcctccgctg	720
aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag	780
agggggagga ccacggtaaa tggagcagc gcatagcaaa agggacgcgg ggtccttttc	840
tctgccggtg gcaactgggtg gctgtggcca ggtgtggtac tttgatgggg cccagggctg	900
gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag	960
gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat	1020
tactcataag accttttttt tttccttccg ggcgcaaaac cgtgagctgg attdataatc	1080
gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgcc ctccgccgac	1140
tagctgcccc cgcgagcaac ggctcgtga tttccccgcc gatccgggtc ccgcctcccc	1200
actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct	1260
ggcgtcgcg tgcgagaggg aactagcag aacgaggaag cagctggagg tgacgccggg	1320
cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc	1380
gggagtgcc ggctcgtgt cgcagagcc aggtgggtaa gctagcgacc acctggactt	1440
cccagcgcgc aaccgtggct tttcagccag gtctctctct cccgcggctt ctcaaccaac	1500
cccatcccag cgcgggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga	1560
agggcgtact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc	1620
gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg	1680

gccaagtccg agggagaacg taaagatatg ggcccttttc ccctctcac cttgtctcac 1740
caaagtcctt agtccccgga gcagttagcc tctttctttc caggaatta gccagacaca 1800
acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgcccctcc ccgctcgctg 1860
cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccg 1920
ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980
ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040
cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100
gagtcccagc agccctcgaa ccaggtgaaa aaccttaca atgtcctgaa tgtgggaaat 2160
cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220
ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280
gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340
gcgatgaact tgtccgacac caacgaacct atacaggcga gaagccttac aaatgtcccg 2400
agtgtggcaa gagcttctca caatcagggg atctgactga gcatcaacga actcataccg 2460
gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520
tacgccacca gaggacacat acaggggaga agccctaca atgccccgaa tgcggtaaaa 2580
gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640
gtaaaaggcc gggggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggtatcctacc 2700
catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760
tgggttctga tgccctcgat gactttgacc tggatatggt ggggaagcgac gcattggatg 2820
actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatggtat 2880
aaactagtga aaccagcag acaatgtagc tagaccagc agccagatgt agctaaagag 2940
accggttcac tgtgaaagct tgggtggcat cctgtgacc cctccccagt gcctctcctg 3000
gccctggaag ttgccactcc agtgcccacc agccttgtcc taataaaatt aagttgcatc 3060
atthttgtctg actaggtgtc cttctataat attatggggg ggaggggggt ggtatggagc 3120
aaggggcaag ttgggaagac aacctgtagg gcctgcgggg tctattggga accaagctgg 3180
agtgcagtg cacaatcttg gctcactgca atctccgct cctgggttca agcgattctc 3240
ctgcctcagc ctcccagatt gttgggattc caggcatgca tgaccaggct cagctaattt 3300
ttgttttttt ggtagagacg gggtttcacc atattggcca ggctgggtctc caactcctaa 3360
tctcaggtga tctaccacc ttggcctccc aaattgctgg gattacaggc gtgaaccact 3420
gctcccttcc ctgtcctt 3438

<210> 75

<211> 3505

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 75

```

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat      60
atdddgaagt tctggtgaca taaagaatca tgatattaat gcccatggaa atgaaagggc      120
gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtggt cgtgtggcag      180
tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca      240
gtgtgagccc agatgaggtt cagctgocaa ctttctcttt cacatcttat gaaagtcatt      300
taagcacaac taactttttt tttttttttt tttttttgag acagagtctt gctctgttgc      360
ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctccataggct      420
caaacgggcc tccctgcatca gcctcccaag tagctggaat tacaggagtg gccaccatg      480
cccagctaat ttttgtattd ttaatagata cgggggtttc accatatcac ccaggctggt      540
ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata      600
ggcgtcagcc actatgccca acccgaccaa ccttttttaa aataaatatt taaaaaattg      660
gtattdcaca tatatactag tattttacatt tatccacaca aaacggacgg gcctccgctg      720
aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag      780
agggggagga ccacggtaaa tggagcgcgc gcatagcaaa agggacgcgg ggtccttttc      840
tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg      900
gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag      960
gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat      1020
tactcataag accttttttd tttccttccg ggcgcaaac cgtgagctgg attdataatc      1080
gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac      1140
tagctgcccc cgcgagcaac ggctcgtgta tttccccgcc gatccgggtcc ccgcctcccc      1200
actctgcccc cgctacccc ggagcctgtc agccgcctct ccgaatctct ctcttctcct      1260
ggcgtcgcg tgcgagaggg aactagcgcg aacgaggaag cagctggagg tgacgccggg      1320
cagattacgc ctgtcagggc cgagccgcgc ggatcgctgg gcgctgtgca gaggaaaggc      1380
gggagtgccc ggctcgtctg cgcagagccg aggtgggtaa gctagcgcacc acctggactt      1440
cccagcgcgc aaccgtggct tttcagccag gtccctctct cccgcggctt ctcaaccaac      1500
cccatcccag cgcgcggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga      1560
agggcgtact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc      1620
gagttgctgt gcgactgcgc gcgcggggct agagtgcagg gtgactgtgg ttcttctctg      1680

```

gccaagtccg agggagaacg taaagatatg ggcctttttc cccctctcac cttgtctcac 1740
caaagtccct agtccccgga gcagttagcc tctttctttc caggggaatta gccagacaca 1800
acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgcctcc cgcctcgcctg 1860
cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccg 1920
ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgccttct 1980
ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040
cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100
gagtcccagc agccctcgaa ccaggtgaaa aaccttaca atgtcctgaa tgtgggaaat 2160
cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220
ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280
gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340
gcgatgaact tgtccgacac caacgaacct atacaggcga gaagccttac aaatgtcccg 2400
agtgtggcaa gagcttctca caatcagggg atctgactga gcatcaacga actcataccg 2460
gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520
tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580
gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640
gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700
catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760
tgggttctga tgcctcgcgt gactttgacc tggatatggt ggggaagcgac gcattggatg 2820
actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880
aaactagtga aaccagcag acaatgtagc tagaccagc agccagatgt agctaaagag 2940
accggttcac tgtgagaaac ccagcagaca atgtagctag acccagtagc cagatgtagc 3000
taaagagacc ggttactgtg gaaagcttgg gtggcatccc tgtgaccct cccagtgcc 3060
tctcctggcc ctggaagtgt cactccagt gccaccagc cttgtcctaa taaaattaag 3120
ttgcatcatt ttgtctgact aggtgtcctt ctataatatt atgggggtgga ggggggtggt 3180
atggagcaag gggcaagtgt ggaagacaac ctgtagggcc tgcggggtct attgggaacc 3240
aagctggagt gcagtggcac aatcttggct cactgcaatc tccgcctcct gggttcaagc 3300
gattctcctg cctcagcctc ccgagttggt gggattccag gcatgcatga ccaggctcag 3360
ctaatttttg ttttttgggt agagacgggg tttcaccata ttggccaggc tggctccaa 3420
ctcctaattc caggtgatct acccaocttg gcctccaaa ttgctgggat tacaggcgtg 3480
aaccactgct ccctccctg tcctt 3505

<210> 76
 <211> 3804
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 76
 ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60
 attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120
 gatcaacact atggtttgaa aagggggaaa ttgtagagca cagatgtggt cgtgtggcag 180
 tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaattctg attggcccca 240
 gtgtgagccc agatgagggt cagctgcca ctttctcttt cacatcttat gaaagtcatt 300
 taagcacaac taactttttt tttttttttt ttttttgag acagagtctt gctctgttgc 360
 ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctccataggct 420
 caaacgggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gccaccatg 480
 cccagctaат ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540
 ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600
 ggcgtcagcc actatgccc acccgaccaa ctttttttaa aataaatatt taaaaaattg 660
 gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720
 aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag 780
 agggggagga ccacggtaaa tggagcgcgc gcatagcaaa agggacgcgg ggtccttttc 840
 tctgccggtg gcaactggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900
 gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960
 gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020
 tactcataag accttttttt tttccttccg ggcgcaaac cgtgagctgg atttataatc 1080
 gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgcc ctccgccgac 1140
 tagctgcccc cgcgagcaac ggctcgtga tttccccgcc gatccgggtcc ccgcctcccc 1200
 actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260
 ggcgctcgcg tgcgagaggg aactagcgcg aacgaggaag cagctggagg tgacgccggg 1320
 cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380
 gggagtgccc ggctcgtctg cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440
 cccagcggcc aaccgtggct tttcagccag gtctctctct cccgcggctt ctcaaccaac 1500
 cccatcccag cgcggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620
 gagttgctgt gcgactgcgc gcgcggggct agagtgcgaag gtgactgtgg ttcttctctg 1680
 gccaaagtccg agggagaacg taaagatatg ggcctttttc cccctctcac cttgtctcac 1740
 caaagtcctt agtccccgga gcagttagcc tctttctttc cagggaaatta gccagacaca 1800
 acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgcctcc cgcctcgctg 1860
 cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccg 1920
 ctgctcagtc cctccggtgt gcaggacccc ggaagtccct cccgcacagc tctcgcttct 1980
 ctttgacagc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040
 cgaccgagct ggaattcgcc accatggccg ccgaccacct gatgctcgcc gagggctacc 2100
 gcctggtgca gaggccgccc tccgcgccc cgcgccatgg ccctcatgcg ctccggactc 2160
 tgccgcccga cgcgggccc ggcctggaca gtgggctgag gccgcggggg gctccgctgg 2220
 ggccgcccgc gccccgcaa cccggggccc tggcgtaagg ggccttcggg ccgcccctct 2280
 ccttccagcc ctttccggcc gtgcctccgc cggccgccc catcgcgac ctgcagcctg 2340
 tggcgacgcc gtaccccggc cgcgcggccc cgcccccaa cgctccggga ggccccccg 2400
 gcccgcagcc ggccccaaag gccgcagccc cgcgcgccc cgcgcacgccc ctgggcccga 2460
 tggacgcccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag ctggggctgc 2520
 accgctgctg cgagctgccc gagctgttcc tgggcccagag cgagttcgac tgcttctcgg 2580
 acttggggtc cgcgcgccc gccggctccg tgagctgccc tggttctggt ggtggttctg 2640
 gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg cccagcaaga 2700
 cgccccccca cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc cgcttctccc 2760
 gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc ttccagtgcc 2820
 gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc cgcaccaca 2880
 caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg agcgatgaac 2940
 ttgtccgaca taccaagatc cacttgccgc agaaggaccg cccttacgct tgcccagtgg 3000
 agtcctgtga tcgccccttc tcccattcag ggaatctgac tgagcacatc cgcattccaca 3060
 caggccagaa gcccttcag tgcgcgatct gcatgagaaa cttcagcaca agtggacatc 3120
 tggtagcca catccgcacc cacacaggcg aaaagccctt cgcctgcgac atctgtggaa 3180
 gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg cggcagaagg 3240
 acaagtaact cgaggaaacc cagcagacaa tgtagctaga cccagtagcc agatgtagct 3300
 aaagagaccg gttcactgtg aaagcttggg tggcatccct gtgaccctc cccagtgcct 3360
 ctctggccc tggaagttgc cactccagtg cccaccagcc ttgtcctaat aaaattaagt 3420
 tgcatcattt tgtctgacta ggtgtccttc tataatatta tggggtggag gggggtggta 3480

046157

tggagcaagg ggcaagttgg gaagacaacc tgtagggcct gcgggggtcta ttgggaacca 3540
agctggagtg cagtggcaca atcttggctc actgcaatct cgcctcctg ggttcaagcg 3600
attctcctgc ctacagctcc cgagttggtg ggattccagg catgcatgac caggetcagc 3660
taatttttgt ttttttggtg gagacggggg ttcaccatat tggccaggct ggtctccaac 3720
tcctaattctc aggtgatcta cccaccttgg cctcccaaat tgctgggatt acaggcgtga 3780
accactgctc ccttcctgt cctt 3804

<210> 77

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 77

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
35 40 45

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

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

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn
145 150 155 160

Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
 165 170 175

<210> 78

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 78

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

Phe Ser Thr Lys Asn Ser Leu Thr Glu His Gln Arg Thr His Thr Gly
 20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ala
 35 40 45

Asp Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Leu Ala His Leu Arg
 65 70 75 80

Ala His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 85 90 95

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

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 115 120 125

Phe Ser Gln Ala Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly
 130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr His
 145 150 155 160

Leu Asp Leu Ile Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
 165 170 175

<210> 79

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 79

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

Phe Ser Gln Ala Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly
 20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
 35 40 45

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly Asn Leu Thr
 65 70 75 80

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 85 90 95

Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln Arg
 100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 115 120 125

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

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ala
 145 150 155 160

Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
 165 170 175

<210> 80

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 80

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

046157

Phe Ser Thr Thr Gly Asn Leu Thr Val His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser
35 40 45

Gly Glu Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

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

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg
145 150 155 160

Ala Asn Leu Arg Ala His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 81

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 81

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Ser Arg Arg Thr Cys Arg Ala His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Thr
35 40 45

046157

Gly Ala Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

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

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Gly Asp Leu Arg Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser
145 150 155 160

His Ser Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 82

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 82

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Lys Asp Asn Leu Lys Asn His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro
35 40 45

Gly Ala Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu

046157

85 90 95

Cys Gly Lys Ser Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Thr Ser Gly Glu Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Lys
145 150 155 160

Asp Asn Leu Lys Asn His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 83
<211> 232
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический
полипептид

<400> 83
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Ser Lys Lys Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Pro
35 40 45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Asn Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr His Thr Gly
 130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser
 145 150 155 160

Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly His Leu Val
 180 185 190

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 195 200 205

Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu His Gln Arg
 210 215 220

Thr His Thr Gly Lys Lys Thr Ser
 225 230

<210> 84

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 84

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly
 20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
 35 40 45

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
 65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 85 90 95

Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg
 100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 115 120 125

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly
 130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser
 145 150 155 160

Gly His Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
 165 170 175

<210> 85

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 85

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

Phe Ser Ser Lys Lys Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
 20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Pro
 35 40 45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Asn Leu Val
 65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 85 90 95

Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg
 100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 115 120 125

046157

Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr His Thr Gly
 130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser
 145 150 155 160

Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
 165 170 175

<210> 86

<211> 232

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 86

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

Phe Ser Asp Cys Arg Asp Leu Ala Arg His Gln Arg Thr His Thr Gly
 20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Asn
 35 40 45

Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Asn Asp Ala Leu Thr
 65 70 75 80

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 85 90 95

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

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 115 120 125

Phe Ser Asp Pro Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly
 130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg
 145 150 155 160

Ala His Leu Glu Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr

046157

165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Ser Ser Leu Val
180 185 190

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
195 200 205

Cys Gly Lys Ser Phe Ser His Arg Thr Thr Leu Thr Asn His Gln Arg
210 215 220

Thr His Thr Gly Lys Lys Thr Ser
225 230

<210> 87
<211> 176
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический
полипептид

<400> 87
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Asn Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Pro
35 40 45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro Gly Asn Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Gln Arg Ala His Leu Glu Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg
145 150 155 160

Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 88

<211> 204

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 88

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Asn Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro
35 40 45

Gly His Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly Glu Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

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

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Leu
145 150 155 160

Ala His Leu Arg Ala His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

046157

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp His Leu Thr
 180 185 190

Asn His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
 195 200

<210> 89

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 89

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly
 20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg
 35 40 45

Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
 65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 85 90 95

Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln Arg
 100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly
 130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
 145 150 155 160

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
 165 170 175

046157

<210> 90

<211> 176

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 90

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg Thr His Thr Gly
 20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
 35 40 45

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asp Leu Arg
 65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 85 90 95

Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln Arg
 100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 115 120 125

Phe Ser Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly
 130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
 145 150 155 160

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
 165 170 175

<210> 91

<211> 260

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

046157

<400> 91

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Arg Asp Glu Leu Asn Val His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
35 40 45

Asp His Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Asp Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser His Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
145 150 155 160

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr
180 185 190

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
195 200 205

Cys Gly Lys Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg
210 215 220

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
225 230 235 240

Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly
245 250 255

Lys Lys Thr Ser
260

<210> 92

<211> 177

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 92

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg
1 5 10 15

Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro
20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu
35 40 45

His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
50 55 60

Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr
65 70 75 80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu
85 90 95

Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile
100 105 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
115 120 125

Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr
130 135 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln
145 150 155 160

Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys Asp
165 170 175

Lys

<210> 93

<211> 177

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 93

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg
 1 5 10 15

Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro
 20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu
 35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
 50 55 60

Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr
 65 70 75 80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu
 85 90 95

Ser Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Ile
 100 105 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
 115 120 125

Asn Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr
 130 135 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
 145 150 155 160

Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp
 165 170 175

Lys

<210> 94

<211> 264

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 94

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg
 1 5 10 15

Arg Asp Glu Leu Asn Val His Ile Arg Ile His Thr Gly Gln Lys Pro
 20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu
 35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
 50 55 60

Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Thr
 65 70 75 80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu
 85 90 95

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
 100 105 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
 115 120 125

Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr
 130 135 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
 145 150 155 160

Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp
 165 170 175

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Thr
 180 185 190

Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
 195 200 205

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser Leu
 210 215 220

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp

046157

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
 180 185 190

Ser Gly His Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
 195 200 205

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Glu Arg Ser His Leu
 210 215 220

Arg Glu His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
 225 230 235 240

Ile Cys Gly Arg Lys Phe Ala Gln Ala Gly His Leu Ala Ser His Thr
 245 250 255

Lys Ile His Leu Arg Gln Lys Asp
 260

<210> 96

<211> 175

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 96

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
 1 5 10 15

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro
 20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn Leu
 35 40 45

His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
 50 55 60

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Ala
 65 70 75 80

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
 85 90 95

Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg Ile
 100 105 110

046157

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
 115 120 125

Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu
 130 135 140

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn Ser
 145 150 155 160

Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
 165 170 175

<210> 97

<211> 175

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 97

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
 1 5 10 15

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro
 20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser His Arg Thr Thr Leu
 35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
 50 55 60

Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ala
 65 70 75 80

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
 85 90 95

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

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
 115 120 125

Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu
 130 135 140

046157

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
 145 150 155 160

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
 165 170 175

<210> 98

<211> 261

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 98

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
 1 5 10 15

Arg Asp Glu Leu Asn Val His Leu Arg Ile His Thr Gly His Lys Pro
 20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Ser Asp His Leu
 35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
 50 55 60

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Ala
 65 70 75 80

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
 85 90 95

Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile
 100 105 110

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
 115 120 125

Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly Glu
 130 135 140

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
 145 150 155 160

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu His Ala
 165 170 175

Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu

046157

180 185 190

Thr Glu His Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg
195 200 205

Ile Cys Met Arg Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile
210 215 220

Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg
225 230 235 240

Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ala Lys Ile His Leu
245 250 255

Lys Gln Lys Glu Lys
260

<210> 99
<211> 753
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический полипептид

<400> 99
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
50 55 60

Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu
85 90 95

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110

Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln
115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 130 135 140

Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr
 145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
 165 170 175

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Lys Lys Thr
 180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
 195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Glu Ala
 210 215 220

Ser Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met
 225 230 235 240

Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
 245 250 255

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
 260 265 270

Asp Asp Phe Asp Leu Asp Met Leu Ile Asn Ser Arg Ser Ser Gly Ser
 275 280 285

Pro Lys Lys Lys Arg Lys Val Gly Ser Gln Tyr Leu Pro Asp Thr Asp
 290 295 300

Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe
 305 310 315 320

Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg
 325 330 335

Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg Ser Ser Ala Ser Val
 340 345 350

Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr
 355 360 365

Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile

046157

370 375 380

Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln
385 390 395 400

Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln
405 410 415

Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val
420 425 430

Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser
435 440 445

Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu
450 455 460

Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr Asp Leu Ala Ser Val
465 470 475 480

Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala
485 490 495

Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr
500 505 510

Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro
515 520 525

Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp
530 535 540

Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Gly Ser Gly
545 550 555 560

Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro Glu
565 570 575

Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys Gln
580 585 590

Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala Asn
595 600 605

Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val His
610 615 620

046157

Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro Leu Asp
625 630 635 640

Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro
645 650 655

Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp
660 665 670

Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met Asp
675 680 685

Leu Ser His Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr Thr Thr
690 695 700

Leu Glu Ser Met Thr Glu Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro
705 710 715 720

Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu
725 730 735

His Ala Met His Ile Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu
740 745 750

Phe

<210> 100

<211> 753

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 100

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro

046157

65					70											75				80	
Tyr	Lys	Cys	Pro	Glu	Cys	Gly	Lys	Ser	Phe	Ser	Arg	Ser	Asp	Glu	Leu						
				85					90				95								
Val	Arg	His	Gln	Arg	Thr	His	Thr	Gly	Glu	Lys	Pro	Tyr	Lys	Cys	Pro						
			100					105					110								
Glu	Cys	Gly	Lys	Ser	Phe	Ser	Gln	Ser	Gly	Asn	Leu	Thr	Glu	His	Gln						
		115					120					125									
Arg	Thr	His	Thr	Gly	Glu	Lys	Pro	Tyr	Lys	Cys	Pro	Glu	Cys	Gly	Lys						
	130					135					140										
Ser	Phe	Ser	Thr	Ser	Gly	His	Leu	Val	Arg	His	Gln	Arg	Thr	His	Thr						
145					150					155					160						
Gly	Glu	Lys	Pro	Tyr	Lys	Cys	Pro	Glu	Cys	Gly	Lys	Ser	Phe	Ser	Gln						
				165					170					175							
Asn	Ser	Thr	Leu	Thr	Glu	His	Gln	Arg	Thr	His	Thr	Gly	Lys	Lys	Thr						
			180					185					190								
Ser	Lys	Arg	Pro	Ala	Ala	Thr	Lys	Lys	Ala	Gly	Gln	Ala	Lys	Lys	Lys						
		195					200					205									
Lys	Gly	Ser	Tyr	Pro	Tyr	Asp	Val	Pro	Asp	Tyr	Ala	Leu	Glu	Glu	Ala						
	210					215					220										
Ser	Gly	Ser	Gly	Arg	Ala	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met						
225					230					235					240						
Leu	Gly	Ser	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	Ser						
				245					250					255							
Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	Ser	Asp	Ala	Leu						
			260					265					270								
Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Ile	Asn	Ser	Arg	Ser	Ser	Gly	Ser						
		275					280					285									
Pro	Lys	Lys	Lys	Arg	Lys	Val	Gly	Ser	Gln	Tyr	Leu	Pro	Asp	Thr	Asp						
	290					295					300										
Asp	Arg	His	Arg	Ile	Glu	Glu	Lys	Arg	Lys	Arg	Thr	Tyr	Glu	Thr	Phe						
305					310					315					320						

046157

Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg
 325 330 335

Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg Ser Ser Ala Ser Val
 340 345 350

Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr
 355 360 365

Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile
 370 375 380

Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln
 385 390 395 400

Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln
 405 410 415

Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val
 420 425 430

Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser
 435 440 445

Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu
 450 455 460

Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr Asp Leu Ala Ser Val
 465 470 475 480

Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala
 485 490 495

Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr
 500 505 510

Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro
 515 520 525

Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp
 530 535 540

Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Gly Ser Gly
 545 550 555 560

Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro Glu
 565 570 575

046157

Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys Gln
580 585 590

Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala Asn
595 600 605

Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val His
610 615 620

Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro Leu Asp
625 630 635 640

Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro
645 650 655

Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp
660 665 670

Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met Asp
675 680 685

Leu Ser His Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr Thr Thr
690 695 700

Leu Glu Ser Met Thr Glu Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro
705 710 715 720

Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu
725 730 735

His Ala Met His Ile Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu
740 745 750

Phe

<210> 101

<211> 272

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 101

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

046157

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
 35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
 50 55 60

Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro
 65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu
 85 90 95

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
 100 105 110

Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln
 115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 130 135 140

Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr
 145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
 165 170 175

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Lys Lys Thr
 180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
 195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
 210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
 225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
 245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
 260 265 270

<210> 102

<211> 272

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 102

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu
85 90 95Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln
115 120 125Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140Ser Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln
165 170 175Asn Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205

046157

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
260 265 270

<210> 103

<211> 580

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 103

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu

046157

Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala
 405 410 415

Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala
 420 425 430

Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly
 435 440 445

Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly
 450 455 460

Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala
 465 470 475 480

Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala
 485 490 495

Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala
 500 505 510

Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met
 515 520 525

Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu
 530 535 540

Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln
 545 550 555 560 565

Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly
 565 570 575

Ser Val Ser Cys
 580

<210> 104

<211> 394

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 104

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
 1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr

046157

20 25 30
 Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
 35 40 45
 Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
 50 55 60
 Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
 65 70 75 80
 Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
 85 90 95
 Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
 100 105 110
 Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
 115 120 125
 Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
 130 135 140
 Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
 145 150 155 160
 Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
 165 170 175
 Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
 180 185 190
 Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
 195 200 205
 Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
 210 215 220
 Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
 225 230 235 240
 Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
 245 250 255
 Arg Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His
 260 265 270

046157

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
275 280 285

Arg Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys
290 295 300

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
305 310 315 320

Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
325 330 335

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser
340 345 350

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
355 360 365

Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His
370 375 380

Thr Lys Ile His Leu Arg Gln Lys Asp Lys
385 390

<210> 105

<211> 580

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 105

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro

046157

				85					90					95			
Tyr	Pro	Gly	Arg	Ala	Ala	Ala	Ala	Pro	Pro	Asn	Ala	Pro	Gly	Gly	Pro	Pro	
			100						105					110			
Gly	Pro	Gln	Pro	Ala	Pro	Ser	Ala	Ala	Ala	Pro	Pro	Pro	Pro	Ala	His		
		115							120				125				
Ala	Leu	Gly	Gly	Met	Asp	Ala	Glu	Leu	Ile	Asp	Glu	Glu	Ala	Leu	Thr		
	130						135				140						
Ser	Leu	Glu	Leu	Glu	Leu	Gly	Leu	His	Arg	Val	Arg	Glu	Leu	Pro	Glu		
145					150					155					160		
Leu	Phe	Leu	Gly	Gln	Ser	Glu	Phe	Asp	Cys	Phe	Ser	Asp	Leu	Gly	Ser		
				165					170					175			
Ala	Pro	Pro	Ala	Gly	Ser	Val	Ser	Cys	Gly	Gly	Ser	Gly	Gly	Gly	Ser		
			180					185						190			
Gly	Gln	Ser	Gln	Leu	Ile	Lys	Pro	Ser	Arg	Met	Arg	Lys	Tyr	Pro	Asn		
		195					200					205					
Arg	Pro	Ser	Lys	Thr	Pro	Pro	His	Glu	Arg	Pro	Tyr	Ala	Cys	Pro	Val		
	210						215				220						
Glu	Ser	Cys	Asp	Arg	Arg	Phe	Ser	Arg	Ser	Asp	Asn	Leu	Val	Arg	His		
225					230					235					240		
Ile	Arg	Ile	His	Thr	Gly	Gln	Lys	Pro	Phe	Gln	Cys	Arg	Ile	Cys	Met		
				245					250					255			
Arg	Asn	Phe	Ser	Arg	Glu	Asp	Asn	Leu	His	Thr	His	Ile	Arg	Thr	His		
			260					265					270				
Thr	Gly	Glu	Lys	Pro	Phe	Ala	Cys	Asp	Ile	Cys	Gly	Arg	Lys	Phe	Ala		
		275					280					285					
Arg	Ser	Asp	Glu	Leu	Val	Arg	His	Thr	Lys	Ile	His	Leu	Arg	Gln	Lys		
	290					295					300						
Asp	Arg	Pro	Tyr	Ala	Cys	Pro	Val	Glu	Ser	Cys	Asp	Arg	Arg	Phe	Ser		
305					310					315					320		
Gln	Ser	Gly	Asn	Leu	Thr	Glu	His	Ile	Arg	Ile	His	Thr	Gly	Gln	Lys		
			325						330					335			

046157

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
 340 345 350

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
 355 360 365

Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
 370 375 380

Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ala Asp His
 385 390 395 400

Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala
 405 410 415

Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala
 420 425 430

Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly
 435 440 445

Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly
 450 455 460

Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala
 465 470 475 480

Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala
 485 490 495

Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala
 500 505 510

Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met
 515 520 525

Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu
 530 535 540

Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln
 545 550 555 560

Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly
 565 570 575

Ser Val Ser Cys
 580

<210> 106

<211> 394

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 106

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
195 200 205

046157

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
 210 215 220

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
 225 230 235 240

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
 245 250 255

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His
 260 265 270

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
 275 280 285

Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys
 290 295 300

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
 305 310 315 320

Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
 325 330 335

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
 340 345 350

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
 355 360 365

Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
 370 375 380

Thr Lys Ile His Leu Arg Gln Lys Asp Lys
 385 390

<210> 107

<211> 388

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 107

Met Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
 1 5 10 15

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val

046157

20 25 30
 Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
 35 40 45
 Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
 50 55 60
 Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His
 65 70 75 80
 Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
 85 90 95
 Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys
 100 105 110
 Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
 115 120 125
 Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
 130 135 140
 Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
 145 150 155 160
 Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
 165 170 175
 Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
 180 185 190
 Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ala Asp His
 195 200 205
 Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala
 210 215 220
 Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala
 225 230 235 240
 Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly
 245 250 255
 Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly
 260 265 270

046157

Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala
 275 280 285

Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala
 290 295 300

Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala
 305 310 315 320

Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met
 325 330 335

Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu
 340 345 350

Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln
 355 360 365

Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly
 370 375 380

Ser Val Ser Cys
 385

<210> 108

<211> 479

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 108

Met Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
 1 5 10 15

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
 20 25 30

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
 35 40 45

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
 50 55 60

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His
 65 70 75 80

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

046157

85 90 95
 Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys
 100 105 110
 Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
 115 120 125
 Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
 130 135 140
 Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
 145 150 155 160
 Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
 165 170 175
 Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
 180 185 190
 Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ser Gly Leu
 195 200 205
 Glu Met Ala Asp His Met Met Ala Met Asn His Gly Arg Phe Pro Asp
 210 215 220
 Gly Thr Asn Gly Leu His His His Pro Ala His Arg Met Gly Met Gly
 225 230 235 240
 Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln Pro Gln His Ala
 245 250 255
 Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly Ala Gly Asn Met
 260 265 270
 Asn Ala Thr Ser Gly Ile Arg His Ala Met Gly Pro Gly Thr Val Asn
 275 280 285
 Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala Arg Phe Asn Asn
 290 295 300
 Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly Gly Ser Leu Pro
 305 310 315 320
 Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr Phe Asn His His
 325 330 335

046157

Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His Pro Ala Ala Gly
340 345 350

His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp Cys Asn Pro Lys
355 360 365

His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly Gly Ser Ser Thr
370 375 380

Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala Gly Ser Ser Asn
385 390 395 400

Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala Ser Val Ala His
405 410 415

Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp Thr Asp Phe Ile
420 425 430

Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met Gly Leu Asp Arg
435 440 445

Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn Glu Phe Asp Phe
450 455 460

Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg Val Ser Cys
465 470 475

<210> 109

<211> 476

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 109

Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly
1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg
20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln
35 40 45

Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly
50 55 60

Ala Gly Asn Met Asn Ala Thr Ser Gly Val Arg His Ala Met Gly Pro

046157

Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile
 325 330 335

Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg
 340 345 350

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
 355 360 365

Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg
 370 375 380

Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg
 385 390 395 400

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
 405 410 415

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
 420 425 430

Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
 435 440 445

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr
 450 455 460

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys
 465 470 475

<210> 110

<211> 554

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 110

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
 1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
 20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
 35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala

046157

50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser
195 200 205

Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys
210 215 220

Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn
225 230 235 240

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
245 250 255

Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
260 265 270

Ala Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly
275 280 285

Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg
290 295 300

046157

Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser
305 310 315 320

Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly
325 330 335

Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn
340 345 350

Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
355 360 365

Leu Glu Met Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val
370 375 380

Gln Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg
385 390 395 400

Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro
405 410 415

Arg Gly Ala Pro Leu Gly Pro Pro Pro Arg Gln Pro Gly Ala Leu
420 425 430

Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala
435 440 445

Val Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr
450 455 460

Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro
465 470 475 480

Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala
485 490 495

His Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu
500 505 510

Thr Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro
515 520 525

Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly
530 535 540

Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
545 550

<210> 111

<211> 362

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 111

Met Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser
1 5 10 15Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys
20 25 30Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn
35 40 45Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
50 55 60Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
65 70 75 80Ala Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly
85 90 95Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg
100 105 110Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser
115 120 125Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly
130 135 140Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn
145 150 155 160Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
165 170 175Leu Glu Met Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val
180 185 190Gln Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg
195 200 205

Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro
 210 215 220

Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu
 225 230 235 240

Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala
 245 250 255

Val Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr
 260 265 270

Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro
 275 280 285

Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala
 290 295 300

His Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu
 305 310 315 320

Thr Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro
 325 330 335

Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly
 340 345 350

Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
 355 360

<210> 112

<211> 479

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 112

Met Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser
 1 5 10 15

Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys
 20 25 30

Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn
 35 40 45

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys

046157

Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly Gly Ser Leu Pro
305 310 315 320

Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr Phe Asn His His
325 330 335

Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His Pro Ala Ala Gly
340 345 350

His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp Cys Asn Pro Lys
355 360 365

His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly Gly Ser Ser Thr
370 375 380

Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala Gly Ser Ser Asn
385 390 395 400

Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala Ser Val Ala His
405 410 415

Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp Thr Asp Phe Ile
420 425 430

Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met Gly Leu Asp Arg
435 440 445

Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn Glu Phe Asp Phe
450 455 460

Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg Val Ser Cys
465 470 475

<210> 113

<211> 476

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 113

Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly
1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg
20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln

046157

35 40 45
 Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly
 50 55 60
 Ala Gly Asn Met Asn Ala Thr Ser Gly Val Arg His Ala Met Gly Pro
 65 70 75 80
 Gly Thr Val Asn Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala
 85 90 95
 Arg Phe Asn Asn Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly
 100 105 110
 Gly Ser Leu Pro Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr
 115 120 125
 Phe Asn His His Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His
 130 135 140
 Pro Ala Ala Gly His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp
 145 150 155 160
 Cys Asn Pro Lys His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly
 165 170 175
 Gly Ser Ser Thr Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala
 180 185 190
 Gly Ser Ser Asn Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala
 195 200 205
 Ser Val Ala His Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp
 210 215 220
 Thr Asp Phe Ile Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met
 225 230 235 240
 Gly Leu Asp Arg Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn
 245 250 255
 Glu Phe Asp Phe Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg
 260 265 270
 Val Ser Cys Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg
 275 280 285

046157

Phe Ser Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly
 290 295 300

His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu
 305 310 315 320

Asp Asn Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
 325 330 335

Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val
 340 345 350

Arg His Ala Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala
 355 360 365

Glu Gly Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
 370 375 380

Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met
 385 390 395 400

Arg Ser Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
 405 410 415

Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala
 420 425 430

Gln Asn Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys
 435 440 445

Glu Lys Lys Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro
 450 455 460

Pro Val Ser Leu Ala Pro Val Val Thr Thr Cys Ala
 465 470 475

<210> 114

<211> 559

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический
 полипептид

<400> 114

Met Thr Gly Lys Leu Ala Glu Lys Leu Pro Val Thr Met Ser Ser Leu
 1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala

046157

20 25 30
 Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln
 35 40 45
 Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro
 50 55 60
 Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn
 65 70 75 80
 Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
 85 90 95
 Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
 100 105 110
 Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
 115 120 125
 Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu
 130 135 140
 Pro Pro Tyr Ser Asn Cys Gly Asp Leu Tyr Ser Glu Pro Val Ser Phe
 145 150 155 160
 His Asp Pro Gln Gly Asn Pro Gly Leu Ala Tyr Ser Pro Gln Asp Tyr
 165 170 175
 Gln Ser Ala Lys Pro Ala Leu Asp Ser Asn Leu Phe Pro Met Ile Pro
 180 185 190
 Asp Tyr Asn Leu Tyr His His Pro Asn Asp Met Gly Ser Ile Pro Glu
 195 200 205
 His Lys Pro Phe Gln Gly Met Asp Pro Ile Arg Val Asn Pro Pro Pro
 210 215 220
 Ile Thr Pro Leu Glu Thr Ile Lys Ala Phe Lys Asp Lys Gln Ile His
 225 230 235 240
 Pro Gly Phe Gly Ser Leu Pro Gln Pro Pro Leu Thr Leu Lys Pro Ile
 245 250 255
 Arg Pro Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Leu His Glu
 260 265 270

046157

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
 275 280 285

Arg Asp Glu Leu Asn Val His Leu Arg Ile His Thr Gly His Lys Pro
 290 295 300

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Ser Asp His Leu
 305 310 315 320

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
 325 330 335

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Ala
 340 345 350

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
 355 360 365

Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile
 370 375 380

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
 385 390 395 400

Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly Glu
 405 410 415

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
 420 425 430

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu His Ala
 435 440 445

Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu
 450 455 460

Thr Glu His Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg
 465 470 475 480

Ile Cys Met Arg Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile
 485 490 495

Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg
 500 505 510

Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ala Lys Ile His Leu
 515 520 525

046157

Lys Gln Lys Glu Lys Lys Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser
530 535 540

Ser Ala Pro Pro Val Ser Leu Ala Pro Val Val Thr Thr Cys Ala
545 550 555

<210> 115

<211> 631

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 115

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

046157

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
 180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
 195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
 210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
 225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
 245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
 260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
 275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
 290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
 305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
 325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
 340 345 350

Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
 355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu Thr
 370 375 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
 385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr Lys
 405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser
 420 425 430

046157

Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg
435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn
450 455 460

Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly
465 470 475 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu
485 490 495

Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys
500 505 510

Lys Ala Asp Lys Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser
515 520 525

Ser Tyr Pro Ser Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr
530 535 540

Ser Tyr Pro Ser Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr
545 550 555 560

Ser Phe Ser Ser Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser
565 570 575

Gly Phe Pro Ser Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro
580 585 590

Ala Phe Pro Ala Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn
595 600 605

Ser Phe Ser Ala Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser
610 615 620

Pro Arg Thr Ile Glu Ile Cys
625 630

<210> 116

<211> 719

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 116

046157

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
 1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
 20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
 35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
 50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
 65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
 85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
 100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
 115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
 130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
 145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
 165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
 180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
 195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
 210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
 225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
 245 250 255

046157

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
 260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
 275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
 290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
 305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
 325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Arg
 340 345 350

Asp Glu Leu Asn Val His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
 355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu Thr
 370 375 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
 385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Thr Lys
 405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser
 420 425 430

Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile Arg
 435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn
 450 455 460

Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly
 465 470 475 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu
 485 490 495

Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Arg
 500 505 510

046157

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Thr Ser
515 520 525

His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
530 535 540

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser Leu Val
545 550 555 560

Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
565 570 575

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr Lys
580 585 590

Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala
595 600 605

Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr
610 615 620

Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr
625 630 635 640

Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser
645 650 655

Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala
660 665 670

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser
675 680 685

Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu
690 695 700

Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys
705 710 715

<210> 117

<211> 627

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

046157

<400> 117

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
 1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
 20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
 35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
 50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
 65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
 85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
 100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
 115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
 130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
 145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
 165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
 180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
 195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
 210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
 225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
 245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
 260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
 275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
 290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
 305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
 325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
 340 345 350

Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
 355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu Thr
 370 375 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
 385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ile Arg
 405 410 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
 420 425 430

Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly
 435 440 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser
 450 455 460

Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
 465 470 475 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His
 485 490 495

Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys

046157

500 505 510
 Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
 515 520 525
 Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
 530 535 540
 Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
 545 550 555 560
 Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
 565 570 575
 Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
 580 585 590
 Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
 595 600 605
 Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
 610 615 620

 Glu Ile Cys
 625

 <210> 118
 <211> 631
 <212> Белок
 <213> Искусственная последовательность

 <220>
 <223> Описание искусственной последовательности: Синтетический
 полипептид

 <400> 118
 Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
 1 5 10 15

 Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
 20 25 30

 Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
 35 40 45

 Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
 50 55 60

 Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
 65 70 75 80

046157

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg

046157

				325						330							335
Pro	Tyr	Ala	Cys	Pro	Val	Glu	Ser	Cys	Asp	Arg	Arg	Phe	Ser	Arg	Ser		
			340					345					350				
Asp	Asn	Leu	Val	Arg	His	Ile	Arg	Ile	His	Thr	Gly	Gln	Lys	Pro	Phe		
		355					360					365					
Gln	Cys	Arg	Ile	Cys	Met	Arg	Asn	Phe	Ser	Arg	Glu	Asp	Asn	Leu	His		
	370					375					380						
Thr	His	Ile	Arg	Thr	His	Thr	Gly	Glu	Lys	Pro	Phe	Ala	Cys	Asp	Ile		
385					390					395					400		
Cys	Gly	Arg	Lys	Phe	Ala	Arg	Ser	Asp	Glu	Leu	Val	Arg	His	Thr	Lys		
				405					410					415			
Ile	His	Leu	Arg	Gln	Lys	Asp	Arg	Pro	Tyr	Ala	Cys	Pro	Val	Glu	Ser		
			420					425					430				
Cys	Asp	Arg	Arg	Phe	Ser	Gln	Ser	Gly	Asn	Leu	Thr	Glu	His	Ile	Arg		
		435					440					445					
Ile	His	Thr	Gly	Gln	Lys	Pro	Phe	Gln	Cys	Arg	Ile	Cys	Met	Arg	Asn		
	450					455					460						
Phe	Ser	Thr	Ser	Gly	His	Leu	Val	Arg	His	Ile	Arg	Thr	His	Thr	Gly		
465					470					475					480		
Glu	Lys	Pro	Phe	Ala	Cys	Asp	Ile	Cys	Gly	Arg	Lys	Phe	Ala	Gln	Asn		
				485					490					495			
Ser	Thr	Leu	Thr	Glu	His	Thr	Lys	Ile	His	Leu	Arg	Gln	Lys	Asp	Lys		
			500					505					510				
Lys	Ala	Asp	Lys	Ser	Val	Val	Ala	Ser	Ser	Ala	Thr	Ser	Ser	Leu	Ser		
		515					520						525				
Ser	Tyr	Pro	Ser	Pro	Val	Ala	Thr	Ser	Tyr	Pro	Ser	Pro	Val	Thr	Thr		
	530					535					540						
Ser	Tyr	Pro	Ser	Pro	Ala	Thr	Thr	Ser	Tyr	Pro	Ser	Pro	Val	Pro	Thr		
545					550					555					560		
Ser	Phe	Ser	Ser	Pro	Gly	Ser	Ser	Thr	Tyr	Pro	Ser	Pro	Val	His	Ser		
				565					570					575			

046157

Gly Phe Pro Ser Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro
580 585 590

Ala Phe Pro Ala Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn
595 600 605

Ser Phe Ser Ala Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser
610 615 620

Pro Arg Thr Ile Glu Ile Cys
625 630

<210> 119

<211> 473

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 119

Met Thr Gly Lys Leu Ala Glu Lys Leu Pro Val Thr Met Ser Ser Leu
1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala
20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln
35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro
50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn
65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
85 90 95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
100 105 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
115 120 125

Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu
130 135 140

Pro Pro Tyr Ser Asn Cys Gly Asp Leu Tyr Ser Glu Pro Val Ser Phe

046157

Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu
405 410 415

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
420 425 430

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu Lys Lys
435 440 445

Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro Pro Val Ser
450 455 460

Leu Ala Pro Val Val Thr Thr Cys Ala
465 470

<210> 120

<211> 719

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 120

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn

046157

130
 Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
 145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
 165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
 180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
 195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
 210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
 225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
 245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
 260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
 275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
 290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
 305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
 325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Asp Pro
 340 345 350

Gly Ala Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
 355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn Leu Val
 370 375 380

046157

Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
 385 390 395 400

Cys Gly Arg Lys Phe Ala Gln Ser Gly Asp Leu Arg Arg His Thr Lys
 405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser
 420 425 430

Cys Asp Arg Arg Phe Ser Thr His Leu Asp Leu Ile Arg His Ile Arg
 435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn
 450 455 460

Phe Ser Thr Ser Gly Asn Leu Val Arg His Ile Arg Thr His Thr Gly
 465 470 475 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser
 485 490 495

Asp Asn Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp Arg
 500 505 510

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser
 515 520 525

Gly His Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
 530 535 540

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Glu Arg Ser His Leu Arg
 545 550 555 560

Glu His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
 565 570 575

Cys Gly Arg Lys Phe Ala Gln Ala Gly His Leu Ala Ser His Thr Lys
 580 585 590

Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala
 595 600 605

Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr
 610 615 620

Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr
 625 630 635 640

046157

Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser
645 650 655

Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala
660 665 670

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser
675 680 685

Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu
690 695 700

Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys
705 710 715

<210> 121

<211> 627

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 121

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

046157

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
 130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
 145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
 165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
 180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
 195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
 210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
 225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
 245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
 260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
 275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
 290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
 305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
 325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
 340 345 350

Asp Asn Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
 355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Ser Thr Thr Leu Thr
 370 375 380

046157

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
 385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Asn Arg Lys Thr His Ile Arg
 405 410 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
 420 425 430

Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly
 435 440 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser
 450 455 460

Ser Ser Leu Thr Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
 465 470 475 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Asn Arg Lys
 485 490 495

Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
 500 505 510

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
 515 520 525

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
 530 535 540

Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
 545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
 565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
 580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
 595 600 605

Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
 610 615 620

Glu Ile Cys
 625

<210> 122

<211> 473

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 122

Met Thr Gly Lys Leu Ala Glu Lys Leu Pro Val Thr Met Ser Ser Leu
 1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala
 20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln
 35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro
 50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn
 65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
 85 90 95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
 100 105 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
 115 120 125

Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu
 130 135 140

Pro Pro Tyr Ser Asn Cys Gly Asp Leu Tyr Ser Glu Pro Val Ser Phe
 145 150 155 160

His Asp Pro Gln Gly Asn Pro Gly Leu Ala Tyr Ser Pro Gln Asp Tyr
 165 170 175

Gln Ser Ala Lys Pro Ala Leu Asp Ser Asn Leu Phe Pro Met Ile Pro
 180 185 190

Asp Tyr Asn Leu Tyr His His Pro Asn Asp Met Gly Ser Ile Pro Glu
 195 200 205

046157

His Lys Pro Phe Gln Gly Met Asp Pro Ile Arg Val Asn Pro Pro Pro
 210 215 220

Ile Thr Pro Leu Glu Thr Ile Lys Ala Phe Lys Asp Lys Gln Ile His
 225 230 235 240

Pro Gly Phe Gly Ser Leu Pro Gln Pro Pro Leu Thr Leu Lys Pro Ile
 245 250 255

Arg Pro Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Leu His Glu
 260 265 270

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
 275 280 285

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro
 290 295 300

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn Leu
 305 310 315 320

His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
 325 330 335

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Ala
 340 345 350

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
 355 360 365

Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg Ile
 370 375 380

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
 385 390 395 400

Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu
 405 410 415

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn Ser
 420 425 430

Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys Lys
 435 440 445

Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro Pro Val Ser
 450 455 460

Leu Ala Pro Val Val Thr Thr Cys Ala
465 470

<210> 123

<211> 627

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 123

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

046157

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
340 345 350

Asp Asn Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn Leu Thr
370 375 380

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Ile Arg
405 410 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
420 425 430

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
435 440 445

046157

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
450 455 460

Gly His Leu Thr Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
465 470 475 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Ser Ser Thr Arg Lys
485 490 495

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
500 505 510

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
515 520 525

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
530 535 540

Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
595 600 605

Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620

Glu Ile Cys
625

<210> 124

<211> 627

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 124

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

046157

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
 20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
 35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
 50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
 65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
 85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
 100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
 115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
 130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
 145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
 165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
 180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
 195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
 210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
 225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
 245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
 260 265 270

046157

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
 275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
 290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
 305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
 325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
 340 345 350

Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
 355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His
 370 375 380

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
 385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Ile Arg
 405 410 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
 420 425 430

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
 435 440 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
 450 455 460

Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
 465 470 475 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr
 485 490 495

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
 500 505 510

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser

046157

515 520 525

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
530 535 540

Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
595 600 605

Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620

Glu Ile Cys
625

<210> 125
<211> 272
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический полипептид

<400> 125
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60

Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu
85 90 95

046157

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
 100 105 110

Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln
 115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 130 135 140

Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr
 145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr
 165 170 175

Ser Gly His Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr
 180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
 195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
 210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
 225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
 245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
 260 265 270

<210> 126

<211> 272

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 126

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
 1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
 20 25 30

046157

Ser Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60

Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asp Leu
85 90 95

Arg Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110

Glu Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln
115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140

Ser Phe Ser Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
165 170 175

Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
260 265 270

<210> 127

<211> 261

046157

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 127

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu
85 90 95Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln
115 120 125Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140Ser Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln
165 170 175Asn Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205Lys Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
210 215 220

046157

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
225 230 235 240

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
245 250 255

Asp Leu Asp Met Leu
260

<210> 128

<211> 386

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 128

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

046157

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gln Ser Gln Leu Ile Lys Pro
180 185 190

Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His
195 200 205

Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
210 215 220

Arg Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys
225 230 235 240

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn
245 250 255

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
260 265 270

Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
275 280 285

Thr Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val
290 295 300

Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
305 310 315 320

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
325 330 335

Arg Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
340 345 350

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
355 360 365

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys
370 375 380

Asp Lys
385

<210> 129

<211> 402

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 129

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gln Ser Gln Leu Ile Lys Pro
195 200 205

Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His
210 215 220

046157

Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
225 230 235 240

Arg Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys
245 250 255

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn
260 265 270

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
275 280 285

Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
290 295 300

Thr Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val
305 310 315 320

Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
325 330 335

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
340 345 350

Arg Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
355 360 365

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
370 375 380

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys
385 390 395 400

Asp Lys

<210> 130

<211> 569

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 130

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr

046157

20 25 30
 Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
 35 40 45
 Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
 50 55 60
 Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
 65 70 75 80
 Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
 85 90 95
 Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
 100 105 110
 Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
 115 120 125
 Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
 130 135 140
 Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
 145 150 155 160
 Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
 165 170 175
 Ala Pro Pro Ala Gly Ser Val Ser Cys Ala Asp His Leu Met Leu Ala
 180 185 190
 Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala Ala Ala Ala His
 195 200 205
 Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu
 210 215 220
 Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro
 225 230 235 240
 Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser
 245 250 255
 Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala Gly Ile Ala His
 260 265 270

046157

Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro
 275 280 285

Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala
 290 295 300

Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met Asp Ala Glu Leu
 305 310 315 320

Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu Leu Gly Leu His
 325 330 335

Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp
 340 345 350

Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
 355 360 365

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg
 370 375 380

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
 385 390 395 400

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
 405 410 415

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
 420 425 430

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr
 435 440 445

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
 450 455 460

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
 465 470 475 480

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
 485 490 495

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
 500 505 510

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu
 515 520 525

046157

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
530 535 540

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
545 550 555 560

Lys Ile His Leu Arg Gln Lys Asp Lys
565

<210> 131

<211> 585

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 131

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

046157

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg
195 200 205

Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu
210 215 220

Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly
225 230 235 240

Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr
245 250 255

Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro
260 265 270

Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr
275 280 285

Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly
290 295 300

Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala
305 310 315 320

Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser
325 330 335

Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu
340 345 350

Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala
355 360 365

Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser Gly
370 375 380

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg
385 390 395 400

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
405 410 415

046157

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
 420 425 430

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
 435 440 445

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr
 450 455 460

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
 465 470 475 480

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
 485 490 495

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
 500 505 510

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
 515 520 525

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu
 530 535 540

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
 545 550 555 560

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
 565 570 575

Lys Ile His Leu Arg Gln Lys Asp Lys
 580 585

<210> 132

<211> 523

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 132

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
 1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
 20 25 30

046157

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp
 35 40 45

Met Leu Ile Asn Ser Arg Ser Ser Gly Ser Pro Lys Lys Lys Arg Lys
 50 55 60

Val Gly Ser Gln Tyr Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu
 65 70 75 80

Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys
 85 90 95

Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile
 100 105 110

Ala Val Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln
 115 120 125

Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe
 130 135 140

Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu
 145 150 155 160

Ala Pro Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro
 165 170 175

Ala Pro Ala Met Val Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro
 180 185 190

Val Leu Ala Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys
 195 200 205

Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu
 210 215 220

Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp
 225 230 235 240

Pro Ala Val Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu Phe Gln
 245 250 255

Gln Leu Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr Glu Pro
 260 265 270

Met Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val Thr Gly Ala
 275 280 285

046157

Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu
 290 295 300

Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp
 305 310 315 320

Met Asp Phe Ser Ala Leu Leu Gly Ser Gly Ser Gly Ser Arg Asp Ser
 325 330 335

Arg Glu Gly Met Phe Leu Pro Lys Pro Glu Ala Gly Ser Ala Ile Ser
 340 345 350

Asp Val Phe Glu Gly Arg Glu Val Cys Gln Pro Lys Arg Ile Arg Pro
 355 360 365

Phe His Pro Pro Gly Ser Pro Trp Ala Asn Arg Pro Leu Pro Ala Ser
 370 375 380

Leu Ala Pro Thr Pro Thr Gly Pro Val His Glu Pro Val Gly Ser Leu
 385 390 395 400

Thr Pro Ala Pro Val Pro Gln Pro Leu Asp Pro Ala Pro Ala Val Thr
 405 410 415

Pro Glu Ala Ser His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln
 420 425 430

Ala Val Lys Ala Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys
 435 440 445

Glu Glu Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro
 450 455 460

Arg Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu
 465 470 475 480

Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile Leu
 485 490 495

Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser
 500 505 510

Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe
 515 520

<210> 133

<211> 50

046157

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 133

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
 1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
 20 25 30

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp
 35 40 45

Met Leu
 50

<210> 134

<211> 275

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 134

Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly
 1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg
 20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln
 35 40 45

Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly
 50 55 60

Ala Gly Asn Met Asn Ala Thr Ser Gly Ile Arg His Ala Met Gly Pro
 65 70 75 80

Gly Thr Val Asn Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala
 85 90 95

Arg Phe Asn Asn Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly
 100 105 110

Gly Ser Leu Pro Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr

046157

115 120 125

Phe Asn His His Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His
 130 135 140

Pro Ala Ala Gly His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp
 145 150 155 160

Cys Asn Pro Lys His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly
 165 170 175

Gly Ser Ser Thr Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala
 180 185 190

Gly Ser Ser Asn Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala
 195 200 205

Ser Val Ala His Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp
 210 215 220

Thr Asp Phe Ile Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met
 225 230 235 240

Gly Leu Asp Arg Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn
 245 250 255

Glu Phe Asp Phe Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg
 260 265 270

Val Ser Cys
 275

<210> 135

<211> 183

<212> Бeлoк

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 135

Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro
 1 5 10 15

Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro
 20 25 30

Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala
 35 40 45

046157

Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly
50 55 60

Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro
65 70 75 80

Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro
85 90 95

Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro
100 105 110

Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu
115 120 125

Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu
130 135 140

Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe
145 150 155 160

Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro
165 170 175

Pro Ala Gly Ser Val Ser Cys
180

<210> 136

<211> 7

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<220>

<221> MOD_RES

<222> (2)..(2)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (4)..(4)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (6)..(6)

<223> Любая аминокислота

<400> 136
 Cys Хаа Cys Хаа His Хаа His
 1 5

<210> 137
 <211> 17
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический пептид

<220>
 <221> MOD_RES
 <222> (2)..(2)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (4)..(4)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (6)..(6)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (8)..(10)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (12)..(12)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (14)..(14)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (16)..(16)
 <223> Любая аминокислота

<400> 137
 Cys Хаа Cys Хаа Cys Хаа His Хаа Хаа Хаа Cys Хаа Cys Хаа Cys Хаа
 1 5 10 15

Cys

<210> 138
 <211> 17
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический пептид

<220>
 <221> MOD_RES
 <222> (2)..(2)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (4)..(4)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (6)..(6)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (8)..(10)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (12)..(12)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (14)..(14)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (16)..(16)
 <223> Любая аминокислота

<400> 138
 Cys Хаа Cys Хаа Cys Хаа Cys Хаа Хаа Хаа His Хаа Cys Хаа Cys Хаа
 1 5 10 15

Cys

<210> 139
 <211> 15
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический пептид

<220>
 <221> MOD_RES
 <222> (2)..(2)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (4)..(4)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (6)..(6)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (8)..(8)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (10)..(10)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (12)..(12)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (14)..(14)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (15)..(15)

<223> Cys, His, or Asp

<400> 139

Cys Хаа Cys Хаа His Хаа Cys Хаа Cys Хаа Cys Хаа Cys Хаа Хаа

1 5 10 15

<210> 140

<211> 17

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<220>

<221> MOD_RES

<222> (2)..(2)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (4)..(4)

<223> Любая аминокислота

<220>

<221> MOD_RES
 <222> (6)..(6)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (8)..(10)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (12)..(12)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (14)..(14)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (16)..(16)
 <223> Любая аминокислота

<400> 140
 Cys Хаа Cys Хаа Cys Хаа Cys Хаа Хаа Хаа Cys Хаа Cys Хаа Cys Хаа
 1 5 10 15

Cys

<210> 141
 <211> 7
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический пептид

<220>
 <221> MOD_RES
 <222> (2)..(2)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (4)..(4)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (6)..(6)
 <223> Любая аминокислота

<400> 141
 Cys Хаа Cys Хаа Cys Хаа His
 1 5

<210> 142
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<220>
<221> MOD_RES
<222> (2)..(2)
<223> Любая аминокислота

<220>
<221> MOD_RES
<222> (4)..(4)
<223> Любая аминокислота

<220>
<221> MOD_RES
<222> (6)..(6)
<223> Любая аминокислота

<400> 142
Cys Хаа Cys Хаа His Хаа Cys
1 5

<210> 143
<211> 17
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<220>
<221> MOD_RES
<222> (2)..(2)
<223> Любая аминокислота

<220>
<221> MOD_RES
<222> (4)..(4)
<223> Любая аминокислота

<220>
<221> MOD_RES
<222> (6)..(6)
<223> Любая аминокислота

<220>
<221> MOD_RES
<222> (8)..(10)
<223> Любая аминокислота

<220>
<221> MOD_RES
<222> (12)..(12)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (14)..(14)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (16)..(16)

<223> Любая аминокислота

<400> 143

Cys Хаа Cys Хаа His Хаа Cys Хаа Хаа Хаа Cys Хаа Cys Хаа His Хаа
 1 5 10 15

Cys

<210> 144

<211> 17

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<220>

<221> MOD_RES

<222> (2)..(2)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (4)..(4)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (6)..(6)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (8)..(10)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (12)..(12)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (14)..(14)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (16)..(16)

<223> Любая аминокислота

<400> 144

Cys Хаа Cys Хаа Cys Хаа Cys Хаа Хаа Хаа Cys Хаа Cys Хаа His Хаа
1 5 10 15

Cys

<210> 145

<211> 7

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<220>

<221> MOD_RES

<222> (2)..(2)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (4)..(4)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (6)..(6)

<223> Любая аминокислота

<400> 145

Cys Хаа Cys Хаа Cys Хаа Cys
1 5

<210> 146

<211> 17

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<220>

<221> MOD_RES

<222> (2)..(2)

<223> Любая аминокислота

<220>

<221> MOD_RES

<222> (4)..(4)

<223> Любая аминокислота

<220>

<221> MOD_RES
 <222> (6)..(6)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (8)..(10)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (12)..(12)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (14)..(14)
 <223> Любая аминокислота

<220>
 <221> MOD_RES
 <222> (16)..(16)
 <223> Любая аминокислота

<400> 146
 Cys Хаа Cys Хаа Cys Хаа Cys Хаа Хаа Хаа Cys Хаа His Хаа His Хаа
 1 5 10 15

Cys

<210> 147
 <211> 344
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический полипептид

<220>
 <221> MISC_FEATURE
 <222> (8)..(322)
 <223> Эта область может содержать 1-15 повторяющихся звеньев "Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro"

<220>
 <221> NON_CONS
 <222> (18)..(19)

<220>
 <221> NON_CONS
 <222> (39)..(40)

<220>
 <221> NON_CONS
 <222> (60)..(61)

<220>
 <221> NON_CONS
 <222> (81)..(82)

<220>
 <221> NON_CONS
 <222> (102)..(103)

<220>
 <221> NON_CONS
 <222> (123)..(124)

<220>
 <221> NON_CONS
 <222> (144)..(145)

<220>
 <221> NON_CONS
 <222> (165)..(166)

<220>
 <221> NON_CONS
 <222> (186)..(187)

<220>
 <221> NON_CONS
 <222> (207)..(208)

<220>
 <221> NON_CONS
 <222> (228)..(229)

<220>
 <221> NON_CONS
 <222> (249)..(250)

<220>
 <221> NON_CONS
 <222> (270)..(271)

<220>
 <221> NON_CONS
 <222> (291)..(292)

<220>
 <221> NON_CONS
 <222> (312)..(313)

<220>
 <221> NON_CONS
 <222> (333)..(334)

<220>
 <223> Смотри первоначальное описание заявки для подробного описания
 замен и предпочтительных вариантов осуществления

<400> 147
 Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
 1 5 10 15

Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
 20 25 30

046157

Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys
 35 40 45
 Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr
 50 55 60
 His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
 65 70 75 80
 Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
 85 90 95
 Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro
 100 105 110
 Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His
 115 120 125
 Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser
 130 135 140
 His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
 145 150 155 160
 Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
 165 170 175
 Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr
 180 185 190
 Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
 195 200 205
 Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly
 210 215 220
 Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
 225 230 235 240
 Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly
 245 250 255
 Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln
 260 265 270
 Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys

046157

275 280 285

Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys
290 295 300

Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu
305 310 315 320

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg
 325 330 335

Thr His Thr Gly Lys Lys Thr Ser
 340

- <210> 148
- <211> 292
- <212> Белок
- <213> Искусственная последовательность

- <220>
- <223> Описание искусственной последовательности: Синтетический полипептид

- <220>
- <221> MOD_RES
- <222> (8)..(57)
- <223> Любая аминокислота

- <220>
- <221> MISC_FEATURE
- <222> (8)..(57)
- <223> Эта область может содержать 1-50 остатков

- <220>
- <221> MOD_RES
- <222> (65)..(114)
- <223> Любая аминокислота

- <220>
- <221> MISC_FEATURE
- <222> (65)..(114)
- <223> Эта область может содержать 1-50 остатков

- <220>
- <221> MOD_RES
- <222> (122)..(171)
- <223> Любая аминокислота

- <220>
- <221> MISC_FEATURE
- <222> (122)..(171)
- <223> Эта область может содержать 1-50 остатков

- <220>
- <221> MOD_RES
- <222> (179)..(228)
- <223> Любая аминокислота

<220>
 <221> MISC_FEATURE
 <222> (179)..(228)
 <223> Эта область может содержать 1-50 остатков

<220>
 <221> MOD_RES
 <222> (236)..(285)
 <223> Любая аминокислота

<220>
 <221> MISC_FEATURE
 <222> (236)..(285)
 <223> Эта область может содержать 1-50 остатков

<400> 148

Arg Ser Asp Asn Leu Val Arg Xaa
 1 5 10 15

Xaa
 20 25 30

Xaa
 35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp Asn Leu His Thr
 50 55 60

Xaa
 65 70 75 80

Xaa
 85 90 95

Xaa
 100 105 110

Xaa Xaa Arg Ser Asp Glu Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 115 120 125

Xaa
 130 135 140

Xaa
 145 150 155 160

Xaa Gln Ser Gly Asn Leu
 165 170 175

Thr Glu Xaa
 180 185 190

Xaa
 195 200 205

Xaa
 210 215 220

Xaa Xaa Xaa Xaa Thr Ser Gly His Leu Val Arg Xaa Xaa Xaa Xaa Xaa
 225 230 235 240

Xaa
 245 250 255

Xaa
 260 265 270

Xaa Gln Asn Ser
 275 280 285

Thr Leu Thr Glu
 290

<210> 149
 <211> 292
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический полипептид

<220>
 <221> MOD_RES
 <222> (8)..(57)
 <223> Любая аминокислота

<220>
 <221> MISC_FEATURE
 <222> (8)..(57)
 <223> Эта область может содержать 1-50 остатков

<220>
 <221> MOD_RES
 <222> (65)..(114)
 <223> Любая аминокислота

<220>
 <221> MISC_FEATURE
 <222> (65)..(114)
 <223> Эта область может содержать 1-50 остатков

<220>
 <221> MOD_RES
 <222> (122)..(171)

<223> Любая аминокислота

<220>

<221> MISC_FEATURE

<222> (122)..(171)

<223> Эта область может содержать 1-50 остатков

<220>

<221> MOD_RES

<222> (179)..(228)

<223> Любая аминокислота

<220>

<221> MISC_FEATURE

<222> (179)..(228)

<223> Эта область может содержать 1-50 остатков

<220>

<221> MOD_RES

<222> (236)..(285)

<223> Любая аминокислота

<220>

<221> MISC_FEATURE

<222> (236)..(285)

<223> Эта область может содержать 1-50 остатков

<400> 149

Arg Ser Asp Asn Leu Val Arg Xaa
1 5 10 15

Xaa
20 25 30

Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Arg Thr Thr Leu Thr Asn
50 55 60

Xaa
65 70 75 80

Xaa
85 90 95

Xaa
100 105 110

Xaa Xaa Arg Glu Asp Asn Leu His Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa
130 135 140

046157

Xaa
 145 150 155 160

Xaa Thr Ser His Ser Leu
 165 170 175

Thr Glu Xaa
 180 185 190

Xaa
 195 200 205

Xaa
 210 215 220

Xaa Xaa Xaa Xaa Gln Ser Ser Ser Leu Val Arg Xaa Xaa Xaa Xaa Xaa
 225 230 235 240

Xaa
 245 250 255

Xaa
 260 265 270

Xaa Arg Glu Asp
 275 280 285

Asn Leu His Thr
 290

<210> 150

<211> 292

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<220>

<221> MOD_RES

<222> (8)..(57)

<223> Любая аминокислота

<220>

<221> MISC_FEATURE

<222> (8)..(57)

<223> Эта область может содержать 1-50 остатков

<220>

<221> MOD_RES

<222> (65)..(114)

<223> Любая аминокислота

<220>

<221> MISC_FEATURE

<222> (65)..(114)

<223> Эта область может содержать 1-50 остатков

<220>

<221> MOD_RES

<222> (122)..(171)

<223> Любая аминокислота

<220>

<221> MISC_FEATURE

<222> (122)..(171)

<223> Эта область может содержать 1-50 остатков

<220>

<221> MOD_RES

<222> (179)..(228)

<223> Любая аминокислота

<220>

<221> MISC_FEATURE

<222> (179)..(228)

<223> Эта область может содержать 1-50 остатков

<220>

<221> MOD_RES

<222> (236)..(285)

<223> Любая аминокислота

<220>

<221> MISC_FEATURE

<222> (236)..(285)

<223> Эта область может содержать 1-50 остатков

<400> 150

Asp Pro Gly Ala Leu Val Arg Xaa
1 5 10 15

Xaa
20 25 30

Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp Asn Leu Val Arg
50 55 60

Xaa
65 70 75 80

Xaa
85 90 95

046157

Xaa
 100 105 110

Xaa Xaa Gln Ser Gly Asp Leu Arg Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 115 120 125

Xaa
 130 135 140

Xaa
 145 150 155 160

Xaa Thr His Leu Asp Leu
 165 170 175

Ile Arg Xaa
 180 185 190

Xaa
 195 200 205

Xaa
 210 215 220

Xaa Xaa Xaa Xaa Thr Ser Gly Asn Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa
 225 230 235 240

Xaa
 245 250 255

Xaa
 260 265 270

Xaa Arg Ser Asp
 275 280 285

Asn Leu Val Arg
 290

<210> 151
 <211> 463
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический полипептид

<220>

<221> MOD_RES
<222> (8)..(57)
<223> Любая аминокислота

<220>
<221> MISC_FEATURE
<222> (8)..(57)
<223> Эта область может содержать 1-50 остатков

<220>
<221> MOD_RES
<222> (65)..(114)
<223> Любая аминокислота

<220>
<221> MISC_FEATURE
<222> (65)..(114)
<223> Эта область может содержать 1-50 остатков

<220>
<221> MOD_RES
<222> (122)..(171)
<223> Любая аминокислота

<220>
<221> MISC_FEATURE
<222> (122)..(171)
<223> Эта область может содержать 1-50 остатков

<220>
<221> MOD_RES
<222> (179)..(228)
<223> Любая аминокислота

<220>
<221> MISC_FEATURE
<222> (179)..(228)
<223> Эта область может содержать 1-50 остатков

<220>
<221> MOD_RES
<222> (236)..(285)
<223> Любая аминокислота

<220>
<221> MISC_FEATURE
<222> (236)..(285)
<223> Эта область может содержать 1-50 остатков

<220>
<221> MOD_RES
<222> (293)..(342)
<223> Любая аминокислота

<220>
<221> MISC_FEATURE
<222> (293)..(342)
<223> Эта область может содержать 1-50 остатков

<220>
<221> MOD_RES
<222> (350)..(399)
<223> Любая аминокислота

<220>
 <221> MISC_FEATURE
 <222> (350)..(399)
 <223> Эта область может содержать 1-50 остатков

<220>
 <221> MOD_RES
 <222> (407)..(456)
 <223> Любая аминокислота

<220>
 <221> MISC_FEATURE
 <222> (407)..(456)
 <223> Эта область может содержать 1-50 остатков

<400> 151
 Arg Arg Asp Glu Leu Asn Val Xaa
 1 5 10 15
 Xaa
 20 25 30
 Xaa
 35 40 45
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp His Leu Thr Asn
 50 55 60
 Xaa
 65 70 75 80
 Xaa
 85 90 95
 Xaa
 100 105 110
 Xaa Xaa Arg Ser Asp Asp Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 115 120 125
 Xaa
 130 135 140
 Xaa
 145 150 155 160
 Xaa Arg Ser Asp Asn Leu
 165 170 175
 Val Arg Xaa
 180 185 190

046157

Xaa
 195 200 205

Xaa
 210 215 220

Xaa Xaa Xaa Xaa His Arg Thr Thr Leu Thr Asn Xaa Xaa Xaa Xaa Xaa
 225 230 235 240

Xaa
 245 250 255

Xaa
 260 265 270

Xaa Arg Glu Asp
 275 280 285

Asn Leu His Thr Xaa
 290 295 300

Xaa
 305 310 315 320

Xaa
 325 330 335

Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser His Ser Leu Thr Glu Xaa Xaa Xaa
 340 345 350

Xaa
 355 360 365

Xaa
 370 375 380

Xaa Gln
 385 390 395 400

Ser Ser Ser Leu Val Arg Xaa
 405 410 415

Xaa
 420 425 430

Xaa Xaa

435

440

445

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp Asn Leu His Thr
 450 455 460

<210> 152

<211> 7

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 152

Arg Ser Asp Asn Leu Val Arg

1 5

<210> 153

<211> 7

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 153

Arg Glu Asp Asn Leu His Thr

1 5

<210> 154

<211> 7

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 154

Arg Ser Asp Glu Leu Val Arg

1 5

<210> 155

<211> 7

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 155

Gln Ser Gly Asn Leu Thr Glu

1 5

<210> 156
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 156
Thr Ser Gly His Leu Val Arg
1 5

<210> 157
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 157
Gln Asn Ser Thr Leu Thr Glu
1 5

<210> 158
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 158
Asp Pro Gly Ala Leu Val Arg
1 5

<210> 159
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 159
His Arg Thr Thr Leu Thr Asn
1 5

<210> 160
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический

пептид

<400> 160
Gln Ser Gly Asp Leu Arg Arg
1 5

<210> 161
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 161
Thr Ser His Ser Leu Thr Glu
1 5

<210> 162
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 162
Thr His Leu Asp Leu Ile Arg
1 5

<210> 163
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 163
Gln Ser Ser Ser Leu Val Arg
1 5

<210> 164
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 164
Thr Ser Gly Asn Leu Val Arg
1 5

<210> 165
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 165
Arg Arg Asp Glu Leu Asn Val
1 5

<210> 166
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 166
Arg Ser Asp Asp Leu Val Arg
1 5

<210> 167
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 167
Arg Ser Asp His Leu Thr Asn
1 5

<210> 168
<211> 16
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<400> 168
Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys
1 5 10 15

<210> 169
<211> 7
<212> Белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический

пептид

<400> 169

Pro Lys Lys Lys Arg Lys Val
1 5

<210> 170

<211> 16

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 170

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 171

<211> 9

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 171

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5

<210> 172

<211> 2009

<212> Белок

<213> Homo sapiens

<400> 172

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

Lys Ala Ile Phe Arg Phe Ser Ala Thr Ser Ala Leu Tyr Ile Leu Thr
 100 105 110
 Pro Phe Asn Pro Leu Arg Lys Ile Ala Ile Lys Ile Leu Val His Ser
 115 120 125
 Leu Phe Ser Met Leu Ile Met Cys Thr Ile Leu Thr Asn Cys Val Phe
 130 135 140
 Met Thr Met Ser Asn Pro Pro Asp Trp Thr Lys Asn Val Glu Tyr Thr
 145 150 155 160
 Phe Thr Gly Ile Tyr Thr Phe Glu Ser Leu Ile Lys Ile Ile Ala Arg
 165 170 175
 Gly Phe Cys Leu Glu Asp Phe Thr Phe Leu Arg Asp Pro Trp Asn Trp
 180 185 190
 Leu Asp Phe Thr Val Ile Thr Phe Ala Tyr Val Thr Glu Phe Val Asp
 195 200 205
 Leu Gly Asn Val Ser Ala Leu Arg Thr Phe Arg Val Leu Arg Ala Leu
 210 215 220
 Lys Thr Ile Ser Val Ile Pro Gly Leu Lys Thr Ile Val Gly Ala Leu
 225 230 235 240
 Ile Gln Ser Val Lys Lys Leu Ser Asp Val Met Ile Leu Thr Val Phe
 245 250 255
 Cys Leu Ser Val Phe Ala Leu Ile Gly Leu Gln Leu Phe Met Gly Asn
 260 265 270
 Leu Arg Asn Lys Cys Ile Gln Trp Pro Pro Thr Asn Ala Ser Leu Glu
 275 280 285
 Glu His Ser Ile Glu Lys Asn Ile Thr Val Asn Tyr Asn Gly Thr Leu
 290 295 300
 Ile Asn Glu Thr Val Phe Glu Phe Asp Trp Lys Ser Tyr Ile Gln Asp
 305 310 315 320
 Ser Arg Tyr His Tyr Phe Leu Glu Gly Phe Leu Asp Ala Leu Leu Cys
 325 330 335
 Gly Asn Ser Ser Asp Ala Gly Gln Cys Pro Glu Gly Tyr Met Cys Val

046157

340 345 350
 Lys Ala Gly Arg Asn Pro Asn Tyr Gly Tyr Thr Ser Phe Asp Thr Phe
 355 360 365
 Ser Trp Ala Phe Leu Ser Leu Phe Arg Leu Met Thr Gln Asp Phe Trp
 370 375 380
 Glu Asn Leu Tyr Gln Leu Thr Leu Arg Ala Ala Gly Lys Thr Tyr Met
 385 390 395 400
 Ile Phe Phe Val Leu Val Ile Phe Leu Gly Ser Phe Tyr Leu Ile Asn
 405 410 415
 Leu Ile Leu Ala Val Val Ala Met Ala Tyr Glu Glu Gln Asn Gln Ala
 420 425 430
 Thr Leu Glu Glu Ala Glu Gln Lys Glu Ala Glu Phe Gln Gln Met Ile
 435 440 445
 Glu Gln Leu Lys Lys Gln Gln Glu Ala Ala Gln Gln Ala Ala Thr Ala
 450 455 460
 Thr Ala Ser Glu His Ser Arg Glu Pro Ser Ala Ala Gly Arg Leu Ser
 465 470 475 480
 Asp Ser Ser Ser Glu Ala Ser Lys Leu Ser Ser Lys Ser Ala Lys Glu
 485 490 495
 Arg Arg Asn Arg Arg Lys Lys Arg Lys Gln Lys Glu Gln Ser Gly Gly
 500 505 510
 Glu Glu Lys Asp Glu Asp Glu Phe Gln Lys Ser Glu Ser Glu Asp Ser
 515 520 525
 Ile Arg Arg Lys Gly Phe Arg Phe Ser Ile Glu Gly Asn Arg Leu Thr
 530 535 540
 Tyr Glu Lys Arg Tyr Ser Ser Pro His Gln Ser Leu Leu Ser Ile Arg
 545 550 555 560
 Gly Ser Leu Phe Ser Pro Arg Arg Asn Ser Arg Thr Ser Leu Phe Ser
 565 570 575
 Phe Arg Gly Arg Ala Lys Asp Val Gly Ser Glu Asn Asp Phe Ala Asp
 580 585 590

046157

Asp Glu His Ser Thr Phe Glu Asp Asn Glu Ser Arg Arg Asp Ser Leu
 595 600 605

Phe Val Pro Arg Arg His Gly Glu Arg Arg Asn Ser Asn Leu Ser Gln
 610 615 620

Thr Ser Arg Ser Ser Arg Met Leu Ala Val Phe Pro Ala Asn Gly Lys
 625 630 635 640

Met His Ser Thr Val Asp Cys Asn Gly Val Val Ser Leu Val Gly Gly
 645 650 655

Pro Ser Val Pro Thr Ser Pro Val Gly Gln Leu Leu Pro Glu Val Ile
 660 665 670

Ile Asp Lys Pro Ala Thr Asp Asp Asn Gly Thr Thr Thr Glu Thr Glu
 675 680 685

Met Arg Lys Arg Arg Ser Ser Ser Phe His Val Ser Met Asp Phe Leu
 690 695 700

Glu Asp Pro Ser Gln Arg Gln Arg Ala Met Ser Ile Ala Ser Ile Leu
 705 710 715 720

Thr Asn Thr Val Glu Glu Leu Glu Glu Ser Arg Gln Lys Cys Pro Pro
 725 730 735

Cys Trp Tyr Lys Phe Ser Asn Ile Phe Leu Ile Trp Asp Cys Ser Pro
 740 745 750

Tyr Trp Leu Lys Val Lys His Val Val Asn Leu Val Val Met Asp Pro
 755 760 765

Phe Val Asp Leu Ala Ile Thr Ile Cys Ile Val Leu Asn Thr Leu Phe
 770 775 780

Met Ala Met Glu His Tyr Pro Met Thr Asp His Phe Asn Asn Val Leu
 785 790 795 800

Thr Val Gly Asn Leu Val Phe Thr Gly Ile Phe Thr Ala Glu Met Phe
 805 810 815

Leu Lys Ile Ile Ala Met Asp Pro Tyr Tyr Tyr Phe Gln Glu Gly Trp
 820 825 830

Asn Ile Phe Asp Gly Phe Ile Val Thr Leu Ser Leu Val Glu Leu Gly
 835 840 845

046157

Leu Ala Asn Val Glu Gly Leu Ser Val Leu Arg Ser Phe Arg Leu Leu
850 855 860

Arg Val Phe Lys Leu Ala Lys Ser Trp Pro Thr Leu Asn Met Leu Ile
865 870 875 880

Lys Ile Ile Gly Asn Ser Val Gly Ala Leu Gly Asn Leu Thr Leu Val
885 890 895

Leu Ala Ile Ile Val Phe Ile Phe Ala Val Val Gly Met Gln Leu Phe
900 905 910

Gly Lys Ser Tyr Lys Asp Cys Val Cys Lys Ile Ala Ser Asp Cys Gln
915 920 925

Leu Pro Arg Trp His Met Asn Asp Phe Phe His Ser Phe Leu Ile Val
930 935 940

Phe Arg Val Leu Cys Gly Glu Trp Ile Glu Thr Met Trp Asp Cys Met
945 950 955 960

Glu Val Ala Gly Gln Ala Met Cys Leu Thr Val Phe Met Met Val Met
965 970 975

Val Ile Gly Asn Leu Val Val Leu Asn Leu Phe Leu Ala Leu Leu Leu
980 985 990

Ser Ser Phe Ser Ala Asp Asn Leu Ala Ala Thr Asp Asp Asp Asn Glu
995 1000 1005

Met Asn Asn Leu Gln Ile Ala Val Asp Arg Met His Lys Gly Val
1010 1015 1020

Ala Tyr Val Lys Arg Lys Ile Tyr Glu Phe Ile Gln Gln Ser Phe
1025 1030 1035

Ile Arg Lys Gln Lys Ile Leu Asp Glu Ile Lys Pro Leu Asp Asp
1040 1045 1050

Leu Asn Asn Lys Lys Asp Ser Cys Met Ser Asn His Thr Ala Glu
1055 1060 1065

Ile Gly Lys Asp Leu Asp Tyr Leu Lys Asp Val Asn Gly Thr Thr
1070 1075 1080

Ser Gly Ile Gly Thr Gly Ser Ser Val Glu Lys Tyr Ile Ile Asp
1085 1090 1095

Glu Ser Asp Tyr Met Ser Phe Ile Asn Asn Pro Ser Leu Thr Val
 1100 1105 1110

Thr Val Pro Ile Ala Val Gly Glu Ser Asp Phe Glu Asn Leu Asn
 1115 1120 1125

Thr Glu Asp Phe Ser Ser Glu Ser Asp Leu Glu Glu Ser Lys Glu
 1130 1135 1140

Lys Leu Asn Glu Ser Ser Ser Ser Ser Glu Gly Ser Thr Val Asp
 1145 1150 1155

Ile Gly Ala Pro Val Glu Glu Gln Pro Val Val Glu Pro Glu Glu
 1160 1165 1170

Thr Leu Glu Pro Glu Ala Cys Phe Thr Glu Gly Cys Val Gln Arg
 1175 1180 1185

Phe Lys Cys Cys Gln Ile Asn Val Glu Glu Gly Arg Gly Lys Gln
 1190 1195 1200

Trp Trp Asn Leu Arg Arg Thr Cys Phe Arg Ile Val Glu His Asn
 1205 1210 1215

Trp Phe Glu Thr Phe Ile Val Phe Met Ile Leu Leu Ser Ser Gly
 1220 1225 1230

Ala Leu Ala Phe Glu Asp Ile Tyr Ile Asp Gln Arg Lys Thr Ile
 1235 1240 1245

Lys Thr Met Leu Glu Tyr Ala Asp Lys Val Phe Thr Tyr Ile Phe
 1250 1255 1260

Ile Leu Glu Met Leu Leu Lys Trp Val Ala Tyr Gly Tyr Gln Thr
 1265 1270 1275

Tyr Phe Thr Asn Ala Trp Cys Trp Leu Asp Phe Leu Ile Val Asp
 1280 1285 1290

Val Ser Leu Val Ser Leu Thr Ala Asn Ala Leu Gly Tyr Ser Glu
 1295 1300 1305

Leu Gly Ala Ile Lys Ser Leu Arg Thr Leu Arg Ala Leu Arg Pro
 1310 1315 1320

Leu Arg Ala Leu Ser Arg Phe Glu Gly Met Arg Val Val Val Asn

046157

1325						1330						1335			
Ala	Leu	Leu	Gly	Ala	Ile	Pro	Ser	Ile	Met	Asn	Val	Leu	Leu	Val	
1340						1345					1350				
Cys	Leu	Ile	Phe	Trp	Leu	Ile	Phe	Ser	Ile	Met	Gly	Val	Asn	Leu	
1355						1360					1365				
Phe	Ala	Gly	Lys	Phe	Tyr	His	Cys	Ile	Asn	Thr	Thr	Thr	Gly	Asp	
1370						1375					1380				
Arg	Phe	Asp	Ile	Glu	Asp	Val	Asn	Asn	His	Thr	Asp	Cys	Leu	Lys	
1385						1390					1395				
Leu	Ile	Glu	Arg	Asn	Glu	Thr	Ala	Arg	Trp	Lys	Asn	Val	Lys	Val	
1400						1405					1410				
Asn	Phe	Asp	Asn	Val	Gly	Phe	Gly	Tyr	Leu	Ser	Leu	Leu	Gln	Val	
1415						1420					1425				
Ala	Thr	Phe	Lys	Gly	Trp	Met	Asp	Ile	Met	Tyr	Ala	Ala	Val	Asp	
1430						1435					1440				
Ser	Arg	Asn	Val	Glu	Leu	Gln	Pro	Lys	Tyr	Glu	Glu	Ser	Leu	Tyr	
1445						1450					1455				
Met	Tyr	Leu	Tyr	Phe	Val	Ile	Phe	Ile	Ile	Phe	Gly	Ser	Phe	Phe	
1460						1465					1470				
Thr	Leu	Asn	Leu	Phe	Ile	Gly	Val	Ile	Ile	Asp	Asn	Phe	Asn	Gln	
1475						1480					1485				
Gln	Lys	Lys	Lys	Phe	Gly	Gly	Gln	Asp	Ile	Phe	Met	Thr	Glu	Glu	
1490						1495					1500				
Gln	Lys	Lys	Tyr	Tyr	Asn	Ala	Met	Lys	Lys	Leu	Gly	Ser	Lys	Lys	
1505						1510					1515				
Pro	Gln	Lys	Pro	Ile	Pro	Arg	Pro	Gly	Asn	Lys	Phe	Gln	Gly	Met	
1520						1525					1530				
Val	Phe	Asp	Phe	Val	Thr	Arg	Gln	Val	Phe	Asp	Ile	Ser	Ile	Met	
1535						1540					1545				
Ile	Leu	Ile	Cys	Leu	Asn	Met	Val	Thr	Met	Met	Val	Glu	Thr	Asp	
1550						1555					1560				

046157

Asp Gln Ser Glu Tyr Val Thr Thr Ile Leu Ser Arg Ile Asn Leu
 1565 1570 1575

Val Phe Ile Val Leu Phe Thr Gly Glu Cys Val Leu Lys Leu Ile
 1580 1585 1590

Ser Leu Arg His Tyr Tyr Phe Thr Ile Gly Trp Asn Ile Phe Asp
 1595 1600 1605

Phe Val Val Val Ile Leu Ser Ile Val Gly Met Phe Leu Ala Glu
 1610 1615 1620

Leu Ile Glu Lys Tyr Phe Val Ser Pro Thr Leu Phe Arg Val Ile
 1625 1630 1635

Arg Leu Ala Arg Ile Gly Arg Ile Leu Arg Leu Ile Lys Gly Ala
 1640 1645 1650

Lys Gly Ile Arg Thr Leu Leu Phe Ala Leu Met Met Ser Leu Pro
 1655 1660 1665

Ala Leu Phe Asn Ile Gly Leu Leu Leu Phe Leu Val Met Phe Ile
 1670 1675 1680

Tyr Ala Ile Phe Gly Met Ser Asn Phe Ala Tyr Val Lys Arg Glu
 1685 1690 1695

Val Gly Ile Asp Asp Met Phe Asn Phe Glu Thr Phe Gly Asn Ser
 1700 1705 1710

Met Ile Cys Leu Phe Gln Ile Thr Thr Ser Ala Gly Trp Asp Gly
 1715 1720 1725

Leu Leu Ala Pro Ile Leu Asn Ser Lys Pro Pro Asp Cys Asp Pro
 1730 1735 1740

Asn Lys Val Asn Pro Gly Ser Ser Val Lys Gly Asp Cys Gly Asn
 1745 1750 1755

Pro Ser Val Gly Ile Phe Phe Phe Val Ser Tyr Ile Ile Ile Ser
 1760 1765 1770

Phe Leu Val Val Val Asn Met Tyr Ile Ala Val Ile Leu Glu Asn
 1775 1780 1785

Phe Ser Val Ala Thr Glu Glu Ser Ala Glu Pro Leu Ser Glu Asp
 1790 1795 1800

046157

Asp Phe Glu Met Phe Tyr Glu Val Trp Glu Lys Phe Asp Pro Asp
 1805 1810 1815

Ala Thr Gln Phe Met Glu Phe Glu Lys Leu Ser Gln Phe Ala Ala
 1820 1825 1830

Ala Leu Glu Pro Pro Leu Asn Leu Pro Gln Pro Asn Lys Leu Gln
 1835 1840 1845

Leu Ile Ala Met Asp Leu Pro Met Val Ser Gly Asp Arg Ile His
 1850 1855 1860

Cys Leu Asp Ile Leu Phe Ala Phe Thr Lys Arg Val Leu Gly Glu
 1865 1870 1875

Ser Gly Glu Met Asp Ala Leu Arg Ile Gln Met Glu Glu Arg Phe
 1880 1885 1890

Met Ala Ser Asn Pro Ser Lys Val Ser Tyr Gln Pro Ile Thr Thr
 1895 1900 1905

Thr Leu Lys Arg Lys Gln Glu Glu Val Ser Ala Val Ile Ile Gln
 1910 1915 1920

Arg Ala Tyr Arg Arg His Leu Leu Lys Arg Thr Val Lys Gln Ala
 1925 1930 1935

Ser Phe Thr Tyr Asn Lys Asn Lys Ile Lys Gly Gly Ala Asn Leu
 1940 1945 1950

Leu Ile Lys Glu Asp Met Ile Ile Asp Arg Ile Asn Glu Asn Ser
 1955 1960 1965

Ile Thr Glu Lys Thr Asp Leu Thr Met Ser Thr Ala Ala Cys Pro
 1970 1975 1980

Pro Ser Tyr Asp Arg Val Thr Lys Pro Ile Val Glu Lys His Glu
 1985 1990 1995

Gln Glu Gly Lys Asp Glu Lys Ala Lys Gly Lys
 2000 2005

<210> 173

<211> 1052

<212> Белок

<213> Homo sapiens

<400> 173

046157

Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val Gly
1 5 10 15

Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly Val
20 25 30

Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg Ser
35 40 45

Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile Gln
50 55 60

Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His Ser
65 70 75 80

Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu Ser
85 90 95

Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu Ala
100 105 110

Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr Gly
115 120 125

Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala Leu
130 135 140

Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys Asp
145 150 155 160

Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr Val
165 170 175

Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln Leu
180 185 190

Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg Arg
195 200 205

Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys Asp
210 215 220

Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe Pro
225 230 235 240

Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr Asn
245 250 255

046157

Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn Glu
260 265 270

Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe Lys
275 280 285

Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu Val
290 295 300

Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys Pro
305 310 315 320

Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr Ala
325 330 335

Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala Lys
340 345 350

Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu Thr
355 360 365

Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser Asn
370 375 380

Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile Asn
385 390 395 400

Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala Ile
405 410 415

Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln Gln
420 425 430

Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro Val
435 440 445

Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile Ile
450 455 460

Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg Glu
465 470 475 480

Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys Arg
485 490 495

Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr Gly
500 505 510

046157

Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp Met
 515 520 525

Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu Asp
 530 535 540

Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro Arg
 545 550 555 560

Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys Gln
 565 570 575

Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu Ser
 580 585 590

Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile Leu
 595 600 605

Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu Tyr
 610 615 620

Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp Phe
 625 630 635 640

Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu Met
 645 650 655

Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys Val
 660 665 670

Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp Lys
 675 680 685

Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp Ala
 690 695 700

Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys Leu
 705 710 715 720

Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys Gln
 725 730 735

Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu Ile
 740 745 750

Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp Tyr

046157

	755						760					765				
Lys	Tyr	Ser	His	Arg	Val	Asp	Lys	Lys	Pro	Asn	Arg	Glu	Leu	Ile	Asn	
	770					775					780					
Asp	Thr	Leu	Tyr	Ser	Thr	Arg	Lys	Asp	Asp	Lys	Gly	Asn	Thr	Leu	Ile	
785					790					795					800	
Val	Asn	Asn	Leu	Asn	Gly	Leu	Tyr	Asp	Lys	Asp	Asn	Asp	Lys	Leu	Lys	
				805					810					815		
Lys	Leu	Ile	Asn	Lys	Ser	Pro	Glu	Lys	Leu	Leu	Met	Tyr	His	His	Asp	
			820					825					830			
Pro	Gln	Thr	Tyr	Gln	Lys	Leu	Lys	Leu	Ile	Met	Glu	Gln	Tyr	Gly	Asp	
		835					840					845				
Glu	Lys	Asn	Pro	Leu	Tyr	Lys	Tyr	Tyr	Glu	Glu	Thr	Gly	Asn	Tyr	Leu	
	850					855					860					
Thr	Lys	Tyr	Ser	Lys	Lys	Asp	Asn	Gly	Pro	Val	Ile	Lys	Lys	Ile	Lys	
865					870					875					880	
Tyr	Tyr	Gly	Asn	Lys	Leu	Asn	Ala	His	Leu	Asp	Ile	Thr	Asp	Asp	Tyr	
				885					890					895		
Pro	Asn	Ser	Arg	Asn	Lys	Val	Val	Lys	Leu	Ser	Leu	Lys	Pro	Tyr	Arg	
			900					905						910		
Phe	Asp	Val	Tyr	Leu	Asp	Asn	Gly	Val	Tyr	Lys	Phe	Val	Thr	Val	Lys	
		915					920					925				
Asn	Leu	Asp	Val	Ile	Lys	Lys	Glu	Asn	Tyr	Tyr	Glu	Val	Asn	Ser	Lys	
	930					935					940					
Cys	Tyr	Glu	Glu	Ala	Lys	Lys	Leu	Lys	Lys	Ile	Ser	Asn	Gln	Ala	Glu	
945					950					955					960	
Phe	Ile	Ala	Ser	Phe	Tyr	Asn	Asn	Asp	Leu	Ile	Lys	Ile	Asn	Gly	Glu	
				965					970					975		
Leu	Tyr	Arg	Val	Ile	Gly	Val	Asn	Asn	Asp	Leu	Leu	Asn	Arg	Ile	Glu	
			980						985					990		
Val	Asn	Met	Ile	Asp	Ile	Thr	Tyr	Arg	Glu	Tyr	Leu	Glu	Asn	Met	Asn	
		995					1000						1005			

046157

Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr
 1010 1015 1020

Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr
 1025 1030 1035

Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
 1040 1045 1050

<210> 174

<211> 1148

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 174

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
 1 5 10 15

Ala Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
 20 25 30

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
 35 40 45

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
 50 55 60

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
 65 70 75 80

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
 85 90 95

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
 100 105 110

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
 115 120 125

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
 130 135 140

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
 145 150 155 160

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys

046157

				165						170					175	
Asp	Gly	Glu	Val	Arg	Gly	Ser	Ile	Asn	Arg	Phe	Lys	Thr	Ser	Asp	Tyr	
			180					185					190			
Val	Lys	Glu	Ala	Lys	Gln	Leu	Leu	Lys	Val	Gln	Lys	Ala	Tyr	His	Gln	
		195					200					205				
Leu	Asp	Gln	Ser	Phe	Ile	Asp	Thr	Tyr	Ile	Asp	Leu	Leu	Glu	Thr	Arg	
	210					215					220					
Arg	Thr	Tyr	Tyr	Glu	Gly	Pro	Gly	Glu	Gly	Ser	Pro	Phe	Gly	Trp	Lys	
225					230					235					240	
Asp	Ile	Lys	Glu	Trp	Tyr	Glu	Met	Leu	Met	Gly	His	Cys	Thr	Tyr	Phe	
				245					250					255		
Pro	Glu	Glu	Leu	Arg	Ser	Val	Lys	Tyr	Ala	Tyr	Asn	Ala	Asp	Leu	Tyr	
			260					265					270			
Asn	Ala	Leu	Asn	Asp	Leu	Asn	Asn	Leu	Val	Ile	Thr	Arg	Asp	Glu	Asn	
		275					280					285				
Glu	Lys	Leu	Glu	Tyr	Tyr	Glu	Lys	Phe	Gln	Ile	Ile	Glu	Asn	Val	Phe	
	290					295					300					
Lys	Gln	Lys	Lys	Lys	Pro	Thr	Leu	Lys	Gln	Ile	Ala	Lys	Glu	Ile	Leu	
305					310					315					320	
Val	Asn	Glu	Glu	Asp	Ile	Lys	Gly	Tyr	Arg	Val	Thr	Ser	Thr	Gly	Lys	
				325					330					335		
Pro	Glu	Phe	Thr	Asn	Leu	Lys	Val	Tyr	His	Asp	Ile	Lys	Asp	Ile	Thr	
			340					345					350			
Ala	Arg	Lys	Glu	Ile	Ile	Glu	Asn	Ala	Glu	Leu	Leu	Asp	Gln	Ile	Ala	
		355					360					365				
Lys	Ile	Leu	Thr	Ile	Tyr	Gln	Ser	Ser	Glu	Asp	Ile	Gln	Glu	Glu	Leu	
	370					375					380					
Thr	Asn	Leu	Asn	Ser	Glu	Leu	Thr	Gln	Glu	Glu	Ile	Glu	Gln	Ile	Ser	
385					390					395					400	
Asn	Leu	Lys	Gly	Tyr	Thr	Gly	Thr	His	Asn	Leu	Ser	Leu	Lys	Ala	Ile	
				405					410					415		

046157

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
 420 425 430

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
 435 440 445

Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
 450 455 460

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
 465 470 475 480

Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
 485 490 495

Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
 500 505 510

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
 515 520 525

Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
 530 535 540

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
 545 550 555 560

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
 565 570 575

Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
 580 585 590

Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
 595 600 605

Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
 610 615 620

Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
 625 630 635 640

Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
 645 650 655

Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
 660 665 670

046157

Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
675 680 685

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
690 695 700

Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
705 710 715 720

Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
725 730 735

Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
740 745 750

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
755 760 765

Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
770 775 780

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
785 790 795 800

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
805 810 815

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
820 825 830

Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
835 840 845

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
850 855 860

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
865 870 875 880

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
885 890 895

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
900 905 910

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
915 920 925

046157

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
 930 935 940

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
 945 950 955 960

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
 965 970 975

Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
 980 985 990

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
 995 1000 1005

Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn
 1010 1015 1020

Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser
 1025 1030 1035

Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn
 1040 1045 1050

Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys
 1055 1060 1065

Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys
 1070 1075 1080

Lys Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu
 1085 1090 1095

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala
 1100 1105 1110

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp
 1115 1120 1125

Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
 1130 1135 1140

Asp Leu Asp Met Leu
 1145

<210> 175

046157

<211> 387

<212> Белок

<213> Homo sapiens

<400> 175

Met Thr Gly Lys Leu Ala Glu Lys Leu Pro Val Thr Met Ser Ser Leu
 1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala
 20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln
 35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro
 50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn
 65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
 85 90 95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
 100 105 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
 115 120 125

Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu
 130 135 140

Pro Pro Tyr Ser Asn Cys Gly Asp Leu Tyr Ser Glu Pro Val Ser Phe
 145 150 155 160

His Asp Pro Gln Gly Asn Pro Gly Leu Ala Tyr Ser Pro Gln Asp Tyr
 165 170 175

Gln Ser Ala Lys Pro Ala Leu Asp Ser Asn Leu Phe Pro Met Ile Pro
 180 185 190

Asp Tyr Asn Leu Tyr His His Pro Asn Asp Met Gly Ser Ile Pro Glu
 195 200 205

His Lys Pro Phe Gln Gly Met Asp Pro Ile Arg Val Asn Pro Pro Pro
 210 215 220

Ile Thr Pro Leu Glu Thr Ile Lys Ala Phe Lys Asp Lys Gln Ile His
 225 230 235 240

Pro Gly Phe Gly Ser Leu Pro Gln Pro Pro Leu Thr Leu Lys Pro Ile
 245 250 255

Arg Pro Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Leu His Glu
 260 265 270

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
 275 280 285

Ser Asp Glu Leu Thr Arg His Leu Arg Ile His Thr Gly His Lys Pro
 290 295 300

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Ser Asp His Leu
 305 310 315 320

Thr Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
 325 330 335

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Ala
 340 345 350

Lys Ile His Leu Lys Gln Lys Glu Lys Lys Ala Glu Lys Gly Gly Ala
 355 360 365

Pro Ser Ala Ser Ser Ala Pro Pro Val Ser Leu Ala Pro Val Val Thr
 370 375 380

Thr Cys Ala
 385

<210> 176

<211> 543

<212> Белок

<213> Homo sapiens

<400> 176

Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
 1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
 20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
 35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
 50 55 60

046157

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
 65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
 85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
 100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
 115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
 130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
 145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
 165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
 180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
 195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
 210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
 225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
 245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
 260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
 275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
 290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
 305 310 315 320

046157

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
340 345 350

Asp Glu Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu Thr
370 375 380

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Thr Lys
405 410 415

Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala
420 425 430

Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr
435 440 445

Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr
450 455 460

Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser
465 470 475 480

Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala
485 490 495

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser
500 505 510

Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu
515 520 525

Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys
530 535 540

<210> 177

<211> 8

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 177

Gly Gly Ser Gly Gly Gly Ser Gly
1 5

<210> 178

<211> 16

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 178

Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly
1 5 10 15

<210> 179

<211> 4

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<220>

<223> Смотрите первоначальное описание заявки для подробного описания замен и предпочтительных вариантов осуществления

<400> 179

Gly Gly Gly Ser
1

<210> 180

<211> 5

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<220>

<223> Смотрите первоначальное описание заявки для подробного описания замен и предпочтительных вариантов осуществления

<400> 180

Gly Gly Gly Gly Ser
1 5

<210> 181

<211> 4

<212> Белок

<213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический пептид

<220>
 <223> Смотри первоначальное описание заявки для подробного описания замен и предпочтительных вариантов осуществления

<400> 181
 Gly Gly Ser Gly
 1

<210> 182
 <211> 9
 <212> ДНК
 <213> Homo sapiens

<400> 182
 gсgkgggсg 9

<210> 183
 <211> 76
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический олигонуклеотид

<400> 183
 gtttttagtac tctggaааса gaatctacta ааасаaggса ааатgссgtg tttatctcgt 60
 сааcttgттg gсgаgа 76

<210> 184
 <211> 3871
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 184
 ggaggaagcc atсааctааа ctасаатgас tgтааgаtас аааатtgggа atggтаасат 60
 атттtgааgt tctgттgаса таааgаатса tgататтаат gсссатggаа atgааagggc 120
 gatсаасасt atggттtgаа аagggggааа ttgтаgаgса саgатgtgtt сgtgtggсag 180
 tgtgctgtct ctаgсаатас tcаgаgааgа gagаgаасаа tgааатtctg атtgгсссса 240
 gtgtgаgссс агатgаggтт саgctgссаа ctttctcttt cacatcttat gаааgtcatt 300
 тааgсасаас таactttttt tttttttttt tttttttgag асаgаgtctt gctctgttgс 360
 ссaggасаgа gtgсagtagt gactсаатct сggctcactg саgсctccас ctсctaggct 420
 сааасggтсс tcctgсатса gcctccсаag таgctggaат тасаggаgtg gсссассатg 480

cccagctaatt ttttgtatatt ttaatagata cgggggtttc accatatcac ccaggctggt 540
ctcgaactcc tggcctcaag tgatccaact gcctcggcct cccaaagtgc tgggattata 600
ggcgtcagcc actatgcccc acccgaccaa ccttttttaa aataaatatt taaaaaattg 660
gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720
aaccagtgag gccccagacg tgcgcataaa taaccctgc gtgctgcacc acctggggag 780
agggggagga ccacggtaaa tggagcgcgc gcatagcaaa agggacgcgg ggtccttttc 840
tctgcccgtg gcaactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900
gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960
gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020
tactcataag accttttttt tttccttccg ggcgcaaac cgtgagctgg atttataatc 1080
gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140
tagctgcccc cgcgagcaac ggctcgtga tttccccgcc gatccggtcc ccgctcccc 1200
actctgcccc cgctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260
ggcgtcgcg tgcgagaggg aactagcgcg aacgaggaag cagctggagg tgacgccggg 1320
cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380
gggagtgccc ggctcgtgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440
cccagcggcc aaccgtggct tttcagccag gtctctctct cccgcggctt ctcaaccaac 1500
cccatcccag cgccggccac ccaacctccc gaaatgagtg ctctctgccc cagcagccga 1560
aggcgtact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620
gagttgctgt gcgactgcgc gcgcggggct agagtgcagg gtgactgtgg ttcttctctg 1680
gccaagtccg agggagaacg taaagatatg ggcttttttc cccctctcac cttgtctcac 1740
caaagtccct agtccccgga gcagttagc tctttctttc caggaatta gccagacaca 1800
acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgcctcc ccgctcgtg 1860
cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccg 1920
ctgctcagtc cctccggtgt gcaggacccc ggaagtctc cccgcacagc tctcgttct 1980
ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040
cgaccgagct ggaattcgcc accatggccg ccgaccacct gatgctcgcc gagggctacc 2100
gcctggtgca gaggccgccg tccgcgcgcg ccgcccattg cctcatgcg ctccggactc 2160
tgccgccgta cgcgggcccc ggctggaca gtgggctgag gccgcggggg gctccgctgg 2220
ggccgccgcc gccccgcaa cccggggccc tggcgtacgg ggcttcggg ccgcccctct 2280
ccttcagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac ctgcagcctg 2340

tggcgacgcc gtaccccggc cgcgcggccg cccccccaa cgctccggga ggccccccgg 2400
 gcccgcagcc ggcccgaagc gccgcagccc cgcgcgcgcc cgcgcacgcc ctggggcgga 2460
 tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag ctggggctgc 2520
 accgcgtgcg cgagctgccc gagctgttcc tggggcagag cgagttcgac tgcttctcgg 2580
 acttggggtc cgcgcgcgcc gccggctccg tgagctgcgg tggttctggt ggtggttctg 2640
 gtcagtccca gctcatcaaa cccagccgca tgcgcaagta cccaaccgg cccagcaaga 2700
 cccccccca cgaacgcct tacgcttgcc cagtggagtc ctgtgatcgc cgcttctccc 2760
 gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc ttccagtgcc 2820
 gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc cgcaccaca 2880
 caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg agcgatgaac 2940
 ttgtccgaca taccaagatc cacttgccgc agaaggaccg cccttacgct tgcccagtgg 3000
 agtctgtga tcgccccttc tcccaatcag ggaatctgac tgagcacatc cgcattcaca 3060
 caggccagaa gcccttcag tgccgcatct gcatgagaaa cttcagcaca agtggacatc 3120
 tggtagcca catccgcacc cacacaggcg aaaagccctt cgctgcgac atctgtggaa 3180
 gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg cggcagaagg 3240
 acaagtaact cgaggaaacc cagcagacaa tgtagctaga cccagtagcc agatgtagct 3300
 aaagagaccg gttcactgtg agaaaccag cagacaatgt agctagacc agtagccaga 3360
 tgtagctaaa gagaccggtt cactgtgaaa gcttgggtgg catccctgtg acccctccc 3420
 agtgcctctc ctggccctgg aagttgccac tccagtgcc accagccttg tcctaataaa 3480
 attaagttgc atcattttgt ctgactaggt gtccttctat aatattatgg ggtggagggg 3540
 ggtggtatgg agcaaggggc aagttgggaa gacaacctgt agggcctgcg ggtctattg 3600
 ggaaccaagc tggagtgcag tggcacaatc ttggctcact gcaatctccg cctcctgggt 3660
 tcaagcgatt ctctgcctc agcctcccga gttgttggga ttccaggcat gcatgaccag 3720
 gctcagctaa tttttgtttt tttggtagag acggggtttc accatattgg ccaggctggt 3780
 ctccaactcc taatctcagg tgatctaccc accttggcct cccaaattgc tgggattaca 3840
 ggcgtgaacc actgctccct tcctgtcct t 3871

<210> 185

<211> 22

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический праймер

<400> 185

tgtctcggca ttgagaacat tc	22
<210> 186	
<211> 21	
<212> ДНК	
<213> Искусственная последовательность	
<220>	
<223> Описание искусственной последовательности: Синтетический праймер	
<400> 186	
attggtggga ggscattgta t	21
<210> 187	
<211> 20	
<212> ДНК	
<213> Искусственная последовательность	
<220>	
<223> Описание искусственной последовательности: Синтетический праймер	
<400> 187	
ассасаgtcc atgscatcac	20
<210> 188	
<211> 20	
<212> ДНК	
<213> Искусственная последовательность	
<220>	
<223> Описание искусственной последовательности: Синтетический праймер	
<400> 188	
tccaccacc tgttgctgta	20
<210> 189	
<211> 20	
<212> ДНК	
<213> Искусственная последовательность	
<220>	
<223> Описание искусственной последовательности: Синтетический праймер	
<400> 189	
сааааагсс асаааагсст	20
<210> 190	
<211> 20	
<212> ДНК	
<213> Искусственная последовательность	
<220>	
<223> Описание искусственной последовательности: Синтетический праймер	

<400> 190 ttagctccgc aagaacatc	20
<210> 191 <211> 23 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический праймер	
<400> 191 ccatggaact ggctcgattt cac	23
<210> 192 <211> 21 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический праймер	
<400> 192 attggtggga ggcactgta t	21
<210> 193 <211> 27 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический probe	
<400> 193 aggcctgaaa accattgtgg gagscct	27
<210> 194 <211> 24 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический праймер	
<400> 194 gaatgtggga aatcattcag tcgc	24
<210> 195 <211> 24 <212> ДНК <213> Искусственная последовательность	
<220>	

<223> Описание искусственной последовательности: Синтетический праймер	
<400> 195 gcaagttatc ctctcgtgag aagg	24
<210> 196 <211> 29 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический probe	
<400> 196 gsgasaacst ggtgagacat caacgcacc	29
<210> 197 <211> 22 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический праймер	
<400> 197 gctgttatct cttgtgggct gt	22
<210> 198 <211> 22 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический праймер	
<400> 198 aaactcatgg gagctgcccgg tt	22
<210> 199 <211> 25 <212> ДНК <213> Искусственная последовательность	
<220> <223> Описание искусственной последовательности: Синтетический probe	
<400> 199 ссасасааат ctctccctgg cattg	25
<210> 200 <211> 24 <212> ДНК <213> Искусственная последовательность	

<220>
 <223> Описание искусственной последовательности: Синтетический праймер

<400> 200
 gaatgtggga aatcattcag tcgc 24

<210> 201
 <211> 24
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический праймер

<400> 201
 gsaagttatc ctctcgtgag aagg 24

<210> 202
 <211> 29
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический probe

<400> 202
 gcgacaacct ggtgagacat caacgcacc 29

<210> 203
 <211> 1182
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 203
 atggccgscg accacctgat gctcgccgag ggctaccgcc tgggtgcagag gccgccgtcc 60
 gccgccgscg cccatggccc tcatgcgctc cggactctgc cgccttacgc gggcccgggc 120
 ctggacagtg ggctgaggcc gcggggggct ccgctggggc cgcgcgcgcc ccgccaacct 180
 ggggcccctg cgtacggggc cttcggggcc cgcctcctct tccagccctt tccggccgtg 240
 cctccgcccg ccgcccggcat cgcgcacctg cagcctgtgg cgcgcgcgta ccccggccgc 300
 gcggcccgcg cccccaacgc tccgggaggc cccccgggcc cgcagccggc cccaagcgcc 360
 gcagccccgc cgcgcccgc gcacgccttg ggcggcatgg acgcccgaact catcgacgag 420
 gaggcgctga cgtcgctgga gctggagctg gggctgcacc gcgtgcgcga gctgcccgag 480
 ctgttcctgg gccagagcga gttcgactgc ttctcggact tgggggtccgc gccgcccgcc 540

ggctccgtga gctgCGgtgg ttctGGtggt ggttctGGtc agtcccagct catcaaacc 600
 agccgcatgc gcaagtacc caaccggccc agcaagacgc cccccacga acgcccctac 660
 gcttgcccag tggagtctg tgatcgccgc ttctcccga gcgacaacct ggtgagacac 720
 atccgcatcc acacaggcca gaagcccttc cagtgcgca tctgcatgag aaacttcagc 780
 cgagaggata acttgcacac tcacatccgc acccacacag gcgaaaagcc cttcgctgc 840
 gacatctgtg gaagaaagt tgcCCggagc gatgaactg tccgacatac caagatccac 900
 ttgCGgcaga aggaccgccc ttacgcttgc ccagtggagt cctgtgatcg ccgcttctcc 960
 caatcaggga atctgactga gcacatccgc atccacacag gccagaagcc cttccagtgc 1020
 cgcatctgca tgagaaactt cagcacaagt ggacatctgg tacgccacat ccgcaccac 1080
 acaggcgaaa agcccttcgc ctgcgacatc tgtggaagaa agtttgccca gaatagtacc 1140
 ctgaccgaac ataccaagat ccacttgCGg cagaaggaca ag 1182

<210> 204

<211> 1206

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 204

atggccgCCg accacctgat gctCGccgag ggctaccgCC tGGtgCagag gccCCgtCC 60
 gccCGccgCC cccatggccc tcatgcgctc cggactctgc cGCCgtacgc gggccCGggc 120
 ctggacagtG ggctgaggcc gcgggggggct ccgctggggc cGCCgCCgCC ccGCCaaccC 180
 ggggCCctgG cgtacggggc cttCGgggCG ccgtcctcct tccagccctt tccggccgtG 240
 cctccgCCgG ccgCGggcat cgcgcacctg cagcctgtgG cGacgCCgta cccCGgCCgc 300
 gcggccgCGc ccccaacgc tccgggagGc cccCGgggCC cGcagCCgGc cccaagCGcC 360
 gcagccccgc cGCCgCCcgc gcacgCCctg ggCGgcatgG acGCCgaact catcGacgag 420
 gaggcgctga cgtcGctgga gctggagctG gggctgcacc gcgtGcGcga gctgcccgag 480
 ctgttctctgG gccagagcga gttcGactgc ttctCGgact tggggTccgc gccCCcCGcC 540
 ggctccgtga gctgCGgtgg ttctGGtggt ggttctGGtg gTggcagCGg gggaggttct 600
 ggtcagtccc agtcatcaa acccagcCGc atgcGcaagt accccaaccG gccCagcaag 660
 acgccccccc acgaacgccc ttacgcttgc ccagtggagt cctgtgatcg ccgcttctcc 720
 cgcagcgaca acctggtgag acacatccgc atccacacag gccagaagcc cttccagtgc 780
 cgcatctgca tgagaaactt cagccgagag gataacttgc aactcacat ccgcaccac 840
 acaggcgaaa agcccttcgc ctgcgacatc tgtggaagaa agtttgccCG gagcGatgaa 900

046157

cttgtccgac ataccaagat ccacttgccg cagaaggacc gcccttacgc ttgcccagtg 960
gagtcctgtg atcgccgctt ctcccaatca gggaatctga ctgagcacat ccgcatccac 1020
acaggccaga agcccttcca gtgccgcata tgcatagaaa acttcagcac aagtggacat 1080
ctggtagccc acatccgcac ccacacaggc gaaaagccct tcgcctgcga catctgtgga 1140
agaaagtttg cccagaatag taccctgacc gaacatacca agatccactt gcggcagaag 1200
gacaag 1206

<210> 205

<211> 402

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 205

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

046157

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
 165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
 180 185 190

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gln Ser Gln Leu Ile Lys Pro
 195 200 205

Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His
 210 215 220

Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
 225 230 235 240

Arg Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys
 245 250 255

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn
 260 265 270

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
 275 280 285

Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
 290 295 300

Thr Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val
 305 310 315 320

Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
 325 330 335

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
 340 345 350

Arg Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
 355 360 365

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
 370 375 380

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys
 385 390 395 400

Asp Lys

<210> 206

<211> 1707

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 206

```

atggccgсag atсacctgat gctggctgaa ggctacagac tgggtgcagcg gcctccatct      60
gccgctgсcg cccacggccc ccacgcccctg agaaсactgc ccccctacgc cggccctggt      120
cttgatagcg gactcagacc tagagggcgcc cctctgggсc ctccacctcc aagacagcct      180
ggagccctgg cctacggcgсc cttcgгccct ccttctagct tccagccctt ccccgccgtg      240
cctcctccag ccgctggcat cgcccacctg cagcctgtgg ccaccцctta ccccggaaga      300
gccgсcgccc ctccaaacgc ccctggcgga cctcctggсc cccagcctgc tccaagcgсc      360
gctgccccctc cacctcctgc tcatgcцctg ggcggcatgg acgсcgagct gatcgacgag      420
gaagccctga ccagcctgga actggaactg ggсctgcaca gagtgcggga actgcctgag      480
ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgсc ccctcctgcc      540
ggctctgtgt cctgсgссga ccacctgatg ctcgсcgagg gctaccgсct ggtgcagagg      600
ccgсcgtcсg ccgсcgссgc ccatggccct catgсgctcc ggactctgcc gccgtacсgс      660
ggccccgggсc tggacagtgg gctgaggсcg cggggggctc cgctggggсc gccсcgссc      720
cgccaaccсg gggccctggc gtacggggсc ttcgggссgc cgtcctcctt ccagcccttt      780
ccggсcgtgc ctccgссggc cgcgggсcатc gcgcacctgc agcctgtggc gacгссgtac      840
cccggсcgсg ccgсcgссc ccccaacgct ccgggaggсc ccccgggccc gcagccggсc      900
ccaagcgссg cagccccгсc gccgccccгсg cagгсcctgg gcggcatgga cgссgaactc      960
atcgacgagg aggcgctgac gtcgctggag ctggagctgg ggctgcaccg cgtgсgсgag      1020
ctgсссgagc tgttcctggg ccagagcgag ttcgactgct tctcgгactt ggggtccгсg      1080
ccgсссгссg gctcсgtgag ctgссagtcc cagctcatca aaccсagссg catgсгсaaag      1140
taccсcaacc ggсссagcaa gacгссcccc caggaacгсc cttacгсcttg cссagtggag      1200
tcctgtgatc gccгсcttctc ccгсcagсгac aacctggтga gacacatссg catccacaca      1260
ggссagaagc cttcсagtг ccгсatctgc atgagaaact tcagссgaga ggataactтg      1320
сacactcaca tccгсaccсa cacaggсgaa aagcccttcg cctgсгacat ctgtggaaga      1380
aagtttgссc ggagcgatga acttgтccga cataccaaga tccactтгсg gcagaaggac      1440
cgссcttacg cttgсссagt ggagtсctgt gatгсссгct tctcccaatc agggaatctg      1500
actgagсaca tccгсatсca cacaggссag aagcccttcc agтгссгсat ctгсatgaga      1560

```

046157

aacttcagca caagtggaca tctggtacgc cacatccgca cccacacagg cgaaaagccc 1620
 ttcgcctgcg acatctgtgg aagaaagttt gccagaata gtaccctgac cgaacatacc 1680
 aagatccact tgcggcagaa ggacaag 1707

<210> 207

<211> 569

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 207

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
 1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
 20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
 35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
 50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
 65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
 85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
 100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
 115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
 130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
 145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
 165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Ala Asp His Leu Met Leu Ala
 180 185 190

Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala Ala Ala Ala His
 195 200 205

Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu
 210 215 220

Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro
 225 230 235 240

Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser
 245 250 255

Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala Gly Ile Ala His
 260 265 270

Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro
 275 280 285

Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala
 290 295 300

Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met Asp Ala Glu Leu
 305 310 315 320

Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu Leu Gly Leu His
 325 330 335

Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp
 340 345 350

Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
 355 360 365

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg
 370 375 380

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
 385 390 395 400

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
 405 410 415

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
 420 425 430

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr

ggctctgtgt cctgcgggcg cagcggcggc ggaagcggcg ccgaccacct gatgctcgcc 600
 gagggctacc gcctggtgca gaggcgcccg tccgcccggc cggcccatgg ccctcatgcg 660
 ctccggactc tgccgcccga cgcggggccc ggccctggaca gtgggctgag gccgcggggg 720
 gctccgctgg ggccgcccgc gccccgcaa cccggggccc tggcgtagcg ggccttcggg 780
 ccgcccgtcct ccttcagcc ctttcgggc gtgcctccgc cggccgcccg catcgcgcac 840
 ctgcagcctg tggcgacgcc gtaccccggc cgcgcccggc cggcccccaa cgctccggga 900
 ggccccccgg gcccgagcc ggccccaagc gccgcagccc cggcccccgc cgcgcacgcc 960
 ctggggggca tggagcccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag 1020
 ctgggggctgc accgctgcg cgagctgcc gagctgttcc tgggcccagag cgagttcgac 1080
 tgcttctcgg acttggggtc cgcgcccgc gccggctccg tgagctgccc tggttctggt 1140
 ggtggttctg gtcagtccca gtcatacaaa cccagcccga tgcgcaagta cccaaccgg 1200
 cccagcaaga cggccccca cgaacgcct tacgcttgcc cagtggagtc ctgtgatcgc 1260
 cgcttctccc gcagcgacaa cctggtgaga cacatccga tccacacagg ccagaagccc 1320
 ttccagtgcc gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc 1380
 cgcaccaca caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg 1440
 agcgatgaac ttgtccgaca taccaagatc cacttgccgc agaaggaccg cccttacgct 1500
 tgcccagtg agtctctgga tgcgcccgtc tcccaatcag ggaatctgac tgagcacatc 1560
 cgcattccaca caggccagaa gcccttcag tgcgcatct gcatgagaaa cttcagcaca 1620
 agtggacatc tggtagccca catccgcacc cacacaggcg aaaagccctt cgcctgcgac 1680
 atctgtggaa gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg 1740
 cggcagaagg acaag 1755

<210> 209

<211> 585

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 209

Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
 1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
 20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
 35 40 45

046157

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg
195 200 205

Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu
210 215 220

Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly
225 230 235 240

Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr
245 250 255

Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro
260 265 270

Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr
275 280 285

Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly

046157

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
545 550 555 560

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
565 570 575

Lys Ile His Leu Arg Gln Lys Asp Lys
580 585

<210> 210

<211> 1113

<212> ДНК

<213> Homo sapiens

<400> 210

```

atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt      60
ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg      120
ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc      180
ccagcgtcgt tgaacattct cagctcctcc aaccctgcc ttgtccacca tgaccacacc      240
tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag      300
gggaccacaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta      360
gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca      420
cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgccggaggaa gattcgaat      480
aaaagatctg ctcaagagag ccgcaggaaa aagaaggtgt acgttggggg ttagagagc      540
cgggtcttga aatacacagc ccagaatatg gagcttcaga acaaagtaca gcttctggag      600
gaacagaatt tgtcccttct agatcaactg aggaaactcc aggccatggg gattgagatc      660
tcaaacaaaa ccagcagcag cagcacctgc atcttgggtcc tgctagtctc cttctgcctc      720
ctccttgtag ctgctatgta ctctctgac acaaggggga gcctgccagc tgagcatgga      780
gtgttgtccc gccagcttcg tgccctcccc agtgaggacc cttaccagct ggagctgcct      840
gccctgcagt cagaagtgcc gaaagacagc acacaccagt ggttggacgg ctcagactgt      900
gtactccagg ccctggcaa cacttctcgc ctgctgcatt acatgcctca ggctcccagt      960
gcagagcctc ccctggagtg gcccttcctc gacctcttct cagagcctct ctgccgaggt     1020
cccacctccc ccctgcaggc aaatctcaca aggaagggag gatggcttcc tactggtagc     1080
ccctctgtca ttttgagga cagatactca ggc                                     1113

```

<210> 211

<211> 371

<212> Белок

<213> Homo sapiens

<400> 211

046157

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
 1 5 10 15
 Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
 20 25 30
 Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
 35 40 45
 Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
 50 55 60
 Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
 65 70 75 80
 Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
 85 90 95
 Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
 100 105 110
 Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
 115 120 125
 Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
 130 135 140
 Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Lys Ile Arg Asn
 145 150 155 160
 Lys Arg Ser Ala Gln Glu Ser Arg Arg Lys Lys Lys Val Tyr Val Gly
 165 170 175
 Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu
 180 185 190
 Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp
 195 200 205
 Gln Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr
 210 215 220
 Ser Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu
 225 230 235 240
 Leu Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro
 245 250 255

046157

Ala Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu
260 265 270

Asp Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys
275 280 285

Asp Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala
290 295 300

Pro Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser
305 310 315 320

Ala Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro
325 330 335

Leu Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys
340 345 350

Gly Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg
355 360 365

Tyr Ser Gly
370

<210> 212

<211> 1590

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 212

atggagctgg aattggatgc tggtagacca gacctgctgg ccttcctgct agaggaaagt	60
ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg	120
ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc	180
ccagcgtcgt tgaacattct cagctcctcc aaccctgcc ttgtccacca tgaccacacc	240
tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag	300
gggaccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta	360
gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca	420
cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggctcga accaggtgaa	480
aaaccttaca aatgtcctga atgtgggaaa tcattcagtc gcagcgacaa cctggtgaga	540
catcaacgca cccatacagg agaaaaacct tataaatgtc cagaatgtgg aaagtccttc	600

tcacgagagg ataacttgca cactcatcaa cgaacacata ctggtgaaaa accatacaag 660
 tgtcccgaat gtggtaaaag ttttagccgg agcgatgaac ttgtccgaca ccaacgaacc 720
 catacaggcg agaagcctta caaatgtccc gagtgtggca agagcttctc acaatcaggg 780
 aatctgactg agcatcaacg aactcatacc ggggaaaaac cttacaagtg tccagagtgt 840
 gggaagagct tttccacaag tggacatctg gtacgccacc agaggacaca tacaggggag 900
 aagccctaca aatgccccga atgcggtaaa agtttctctc agaatagtac cctgaccgaa 960
 caccagcgaa cacacactgg gaaaaaaaaag agtgtgtacg ttggggggtt agagagccgg 1020
 gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080
 cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140
 aaaaaacca gcagcagcag cacctgcac ttggtcctgc tagtctcctt ctgcctcctc 1200
 cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260
 ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320
 ctgcagtcag aagtgccgaa agacagcaca caccagtggg tggacggctc agactgtgta 1380
 ctccaggccc ctggcaacac ttctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440
 gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500
 atcctcccc tgcaggcaaa tctcacaagg aaggaggat ggcttcctac tggtagcccc 1560
 tctgtcattt tgcaggacag atactcaggc 1590

<210> 213

<211> 530

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 213

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
 1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
 20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
 35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
 50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
 65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
 85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
 100 105 110

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

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
 130 135 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Leu Glu Pro Gly Glu
 145 150 155 160

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp
 165 170 175

Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
 180 185 190

Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr
 195 200 205

His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
 210 215 220

Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr
 225 230 235 240

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
 245 250 255

Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu
 260 265 270

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly
 275 280 285

His Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
 290 295 300

Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu
 305 310 315 320

His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly

<400> 214
 atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60
 ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120
 ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180
 ccagcgtcgt tgaacattct cagctcctcc aaccctgcc ttgtccacca tgaccacacc 240
 tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300
 gggaccacaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360
 gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420
 cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcttga gcccgagag 480
 aagccgtaca agtgcctga gtgcggcaag tcttttagca gaagagacga acttaatgtc 540
 caccagcгаа cgcatactgg tgaaaagccc tataaatgtc ctgaatgtgg gaaatcattc 600
 tccagccgca gaacctgtag ggctcaccag cgaacacaca ccggcgaaaa accatacaaa 660
 tgtccagaat gcgggaaatc cttttctcag tcatccaact tggtagagaca tcaacgcacg 720
 cacactggag aaaagcctta caaatgcccg gaatgtggaa agtctttttc ccaattggcc 780
 catttgcgag cccatcagag gactcacacg ggcgagaaac cttacaaatg cccggaatgc 840
 gggaaatctt tttcaacgag tggcaacctc gtaagacacc aaagaacgca tacaggcgaa 900
 aagccatata agtgtcctga gtgtggtaaa tcattctcac acaggaccac cctgacaaat 960
 caccagcгca cgcacaccgg caagaagaca agcgtgtacg ttggggggtt agagagccgg 1020
 gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080
 cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggatgat tgagatctca 1140
 aacaaaacca gcagcagcag cacctgcate ttggctctgc tagtctcctt ctgcctcctc 1200
 cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtг 1260
 ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320
 ctgcagtcag aagtgccгаа agacagcaca caccagtggт tgгacггctc agactgtgta 1380
 ctccaggccc ctggcaaac ttctgctctg ctgcattaca tgctcaggc tcccagtгca 1440
 gagcctcccc tggagtggcc cttcctgac ctcttctcag agcctctctг ccgaggtccc 1500
 atcctcccc tgсaggcaaа tctcacaagg aaggгaggat ggcttcctac тггtagcccc 1560
 tctgtcattt тгсaggacag atactcagгc 1590

<210> 215

<211> 530

<212> Белок

<213> Искусственная последовательность

<220>

046157

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 215

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
 1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
 20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
 35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
 50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
 65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
 85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
 100 105 110

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

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
 130 135 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Leu Glu Pro Gly Glu
 145 150 155 160

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Arg Asp
 165 170 175

Glu Leu Asn Val His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
 180 185 190

Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Arg Arg Thr Cys Arg Ala
 195 200 205

His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
 210 215 220

Gly Lys Ser Phe Ser Gln Ser Ser Asn Leu Val Arg His Gln Arg Thr
 225 230 235 240

046157

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
 245 250 255

Ser Gln Leu Ala His Leu Arg Ala His Gln Arg Thr His Thr Gly Glu
 260 265 270

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly
 275 280 285

Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
 290 295 300

Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg Thr Thr Leu Thr Asn
 305 310 315 320

His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly
 325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
 340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
 355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser
 370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu
 385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala
 405 410 415

Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp
 420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp
 435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro
 450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala
 465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu
 485 490 495

Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly
 500 505 510

Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr
 515 520 525

Ser Gly
 530

<210> 216

<211> 1590

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический
 полинуклеотид

<400> 216

atggagctgg aattggatgc tggtagacca gacctgctgg ccttcctgct agaggaaagt	60
ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg	120
ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc	180
ccagcgtcgt tgaacattct cagctcctcc aaccctgcc ttgtccacca tgaccacacc	240
tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag	300
gggaccacaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta	360
gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca	420
cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc	480
ccagtggagt cctgtgatcg ccgcttctcc cgcagcgaca acctggtgag acacatccgc	540
atccacacag gccagaagcc cttccagtgc cgcactgca tgagaaactt cagccgagag	600
gataacttgc aactcacat ccgcaccac acaggcgaaa agcccttcgc ctgcgacatc	660
tgtggaagaa agtttgcccg gagcgatgaa cttgtccgac ataccaagat ccacttgccg	720
cagaaggacc gcccttacgc ttgcccagtg gagtccctgtg atcgccgctt ctcccaatca	780
gggaatctga ctgagcacat ccgcatccac acaggccaga agcccttcca gtgccgcatc	840
tgcatgagaa acttcagcac aagtggacat ctggtacgcc acatccgcac ccacacaggc	900
gaaaagccct tcgcctgcga catctgtgga agaaagtttg ccagaatag taccctgacc	960
gaacatacca agatccactt gcggcagaag gacgtgtacg ttggggggtt agagagccgg	1020
gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa	1080
cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca	1140
aacaaaacca gcagcagcag cacctgcac ttggtcctgc tagtctcctt ctgcctcctc	1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260
 ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320
 ctgcagtcag aagtgccgaa agacagcaca caccagtggg tggacggctc agactgtgta 1380
 ctccaggccc ctggcaacac ttcttgctg ctgcattaca tgcctcaggc tcccagtgca 1440
 gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500
 atcctcccc tgcaggcaaa tctcacaagg aaggaggat ggcttcctac tggtagcccc 1560
 tctgtcattt tgcaggacag atactcaggc 1590

<210> 217

<211> 530

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 217

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
 1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
 20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
 35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
 50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
 65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
 85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
 100 105 110

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

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
 130 135 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Pro Tyr Ala Cys

046157

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala
 405 410 415

Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp
 420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp
 435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro
 450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala
 465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu
 485 490 495

Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly
 500 505 510

Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr
 515 520 525

Ser Gly
 530

<210> 218

<211> 1590

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 218

atggagctgg aattggatgc tggtgacca gacctgctgg ccttcctgct agaggaaagt 60
 ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120
 ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180
 ccagcgtcgt tgaacattct cagctcctcc aaccctgcc ttgtccacca tgaccacacc 240
 tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300
 gggaccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360
 gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420
 cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc 480
 ccagtggagt cctgtgatcg ccgcttctcc cgctcagaca acctcgttcg acacatccgc 540

atccacacag gccagaagcc cttccagtgc cgcactctgca tgagaaactt cagccaccgg 600
 actaactca cgaaccacat ccgcaccac acaggcgaaa agcccttcgc ctgcgacatc 660
 tgtggaagaa agtttgccag agaagacaat ctccatactc ataccaagat ccacttgccg 720
 cagaaggacc gcccttacgc ttgcccagtg gaggctctgtg atcgccgctt ctccaccagc 780
 cattctctca ctgaacacat ccgcattcac acaggccaga agcccttcca gtgccgcatc 840
 tgcattgagaa acttcagcca gtctagctca ctgggtgaggc acatccgcac ccacacaggc 900
 gaaaagccct tcgcttcgca catctgtgga agaaagtttg ccaggaggga taacctgcat 960
 acgcatacca agatccactt ggggcagaag gacgtgtacg ttggggggtt agagagccgg 1020
 gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080
 cagaatttgt cccttctaga tcaactgagg aactccagg ccatgggtgat tgagatctca 1140
 aacaaaacca gcagcagcag cacctgcatc ttggctctgc tagtctcctt ctgcctcctc 1200
 cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260
 ttgtcccgcc agcttcgtgc cctcccagc gaggaccctt accagctgga gctgcctgcc 1320
 ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380
 ctccaggccc ctggcaacac ttcttgctg ctgcattaca tgctcaggc tcccagtgca 1440
 gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500
 atcctcccc tgcaggcaaa tctcacaagg aaggaggat ggcttctac tggtagcccc 1560
 tctgtcattt tgcaggacag atactcaggc 1590

<210> 219

<211> 530

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 219

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
 1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
 20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
 35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
 50 55 60

046157

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
 65 70 75 80
 Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
 85 90 95
 Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
 100 105 110
 Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
 115 120 125
 Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
 130 135 140
 Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Pro Tyr Ala Cys
 145 150 155 160 165
 Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val
 165 170 175
 Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile
 180 185 190
 Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg
 195 200 205
 Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
 210 215 220
 Phe Ala Arg Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg
 225 230 235 240
 Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg
 245 250 255
 Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly
 260 265 270
 Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser
 275 280 285
 Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
 290 295 300
 Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His
 305 310 315 320

046157

Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Val Tyr Val Gly Gly
 325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
 340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
 355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser
 370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu
 385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala
 405 410 415

Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp
 420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp
 435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro
 450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala
 465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu
 485 490 495

Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly
 500 505 510

Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr
 515 520 525

Ser Gly
 530

<210> 220

<211> 1140

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 220

```

atggagctgg aattggatgc tggtgacca gacctgctgg ccttcctgct agaggaaagt      60
ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg      120
ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc      180
ccagcgtcgt tgaacattct cagctcctcc aaccctgcc ttgtccacca tgaccacacc      240
tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag      300
gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta      360
gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca      420
cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggctcga accaggtgaa      480
aaaccttaca aatgtcctga atgtgggaaa tcattcagtc gcagcgacaa cctggtgaga      540
catcaacgca ccatacagg agaaaaacct tataaatgtc cagaatgtgg aaagtccttc      600
tcacgagagg ataacttgca cactcatcaa cgaacacata ctggtgaaaa accatacaag      660
tgtcccgaat gtggtaaaag ttttagccgg agcgatgaac ttgtccgaca ccaacgaacc      720
catacaggcg agaagcctta caaatgtccc gagtgtggca agagcttctc acaatcaggg      780
aatctgactg agcatcaacg aactcatacc ggggaaaaaac cttacaagtg tccagagtgt      840
gggaagagct tttccacaag tggacatctg gtacgccacc agaggacaca tacaggggag      900
aagccctaca aatgccccga atgcggtaaa agtttctctc agaatagtac cctgaccgaa      960
caccagcgaa cacacactgg gaaaaaaacg agtgtgtacg ttggggggtt agagagccgg     1020
gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa     1080
cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatatca     1140

```

<210> 221

<211> 380

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 221

```

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1           5           10           15
Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
                20           25           30
Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp

```

046157

35 40 45
 Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
 50 55 60
 Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
 65 70 75 80
 Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
 85 90 95
 Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
 100 105 110
 Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
 115 120 125
 Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
 130 135 140
 Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Leu Glu Pro Gly Glu
 145 150 155 160
 Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp
 165 170 175
 Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
 180 185 190
 Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr
 195 200 205
 His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
 210 215 220
 Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr
 225 230 235 240
 His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
 245 250 255
 Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu
 260 265 270
 Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly
 275 280 285

046157

His Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
290 295 300

Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu
305 310 315 320

His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser
370 375 380

<210> 222

<211> 1140

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полинуклеотид

<400> 222

```

atggagctgg aattggatgc tggtagacca gacctgctgg ccttcctgct agaggaaagt      60
ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg      120
ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc      180
ccagcgtcgt tgaacattct cagctcctcc aaccctgcc ttgtccacca tgaccacacc      240
tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag      300
gggaccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta      360
gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca      420
cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc      480
ccagtggagt cctgtgatcg ccgcttctcc cgcagcgaca acctggtgag acacatccgc      540
atccacacag gccagaagcc cttccagtgc cgcactctgca tgagaaactt cagccgagag      600
gataacttgc aactcacat ccgcacccac acaggcgaaa agcccttcgc ctgcgacatc      660
tgtggaagaa agtttgcccg gagcgatgaa cttgtccgac ataccaagat ccaattgctg      720
cagaaggacc gcccttacgc ttgcccagtg gagtctgtg atcgccgctt ctcccaatca      780
gggaatctga ctgagcacat ccgcatccac acaggccaga agcccttcca gtgccgcatc      840
tgcattgagaa acttcagcac aagtggacat ctggtagccc acatccgcac ccacacaggg      900

```

gaaaagccct tcgcctgcga catctgtgga agaaagtttg cccagaatag taccctgacc 960
 gaacatacca agatccactt gcggcagaag gacgtgtacg ttggggggtt agagagccgg 1020
 gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080
 cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

<210> 223

<211> 380

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 223

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
 1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
 20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
 35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
 50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
 65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
 85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
 100 105 110

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

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
 130 135 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Pro Tyr Ala Cys
 145 150 155 160

Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val
 165 170 175

046157

Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile
 180 185 190

Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg
 195 200 205

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
 210 215 220

Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg
 225 230 235 240

Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg
 245 250 255

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
 260 265 270

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
 275 280 285

Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
 290 295 300

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr
 305 310 315 320

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Val Tyr Val Gly Gly
 325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
 340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
 355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser
 370 375 380

<210> 224

<211> 155

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 224

046157

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
100 105 110

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

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
130 135 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg
145 150 155

<210> 225

<211> 199

<212> Белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический полипептид

<400> 225

Val Tyr Val Gly Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln
1 5 10 15

Asn Met Glu Leu Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu
20 25 30

Ser Leu Leu Asp Gln Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile
35 40 45

Ser Asn Lys Thr Ser Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Полинуклеотид, содержащий последовательность нуклеиновой кислоты, кодирующую не встречающийся в природе фактор транскрипции, причем не встречающийся в природе фактор транскрипции содержит аминокислотную последовательность, имеющую по меньшей мере 97% идентичность последовательности SEQ ID NO: 127.

2. Полинуклеотид по п.1, в котором не встречающийся в природе фактор транскрипции содержит аминокислотную последовательность SEQ ID NO: 127.

3. Полинуклеотид по п.1 или 2, где полинуклеотид дополнительно содержит регуляторный элемент, который управляет экспрессией фактора транскрипции на более высоком уровне в парвальбумин (PV)-содержащих нейронах, чем в других типах клеток, и где регуляторный элемент содержит SEQ ID NO: 1, 2, 3 или 4.

4. Полинуклеотид по п.3, в котором регуляторный элемент содержит SEQ ID NO: 2.

5. Полинуклеотид по любому из пп.1-4, дополнительно содержащий микроРНК связывающий сайт.

6. Полинуклеотид по п.5, в котором микроРНК связывающий сайт содержит SEQ ID NO: 9, 11 или 13.

7. Полинуклеотид по п.5, в котором микроРНК связывающий сайт содержит SEQ ID NO: 7, 14 или 15.

8. Полинуклеотид по п.5, где полинуклеотид содержит последовательность нуклеиновой кислоты, характеризующуюся идентичностью, составляющей по меньшей мере 95% по отношению к SEQ ID NO: 71.

9. Полинуклеотид по п.5, где полинуклеотид содержит SEQ ID NO: 71.

10. Полинуклеотид по любому из пп.1-9, где не встречающийся в природе фактор транскрипции увеличивает экспрессию гена SCN1A в клетке.

11. Экспрессионный вектор, содержащий полинуклеотид по любому из пп.1-10.

12. Экспрессионный вектор по п.11, где экспрессионный вектор представляет собой вирусный вектор.

13. Экспрессионный вектор по п.11, где вирусный вектор представляет собой аденоассоциированный вирусный (AAV) вектор.

14. Экспрессионный вектор по п.13, где AAV вектор имеет серотип AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, rh10, AAV птиц, AAV крупного рогатого скота, AAV собак, AAV лошадей, AAV приматов, AAV не приматов, AAV овец, scAAV, scAAV1, scAAV2, scAAV5, scAAV8 или scAAV9, или их гибриды.

15. Экспрессионный вектор по п.14, где полинуклеотид дополнительно содержит последовательность 5' AAV инвертированного концевой повтора (ITR) и последовательность 3' AAV ITR.

16. Экспрессионный вектор по п.15, где каждая из последовательностей 5' AAV ITR и 3' AAV ITR независимо представляет собой AAV1, AAV2, AAV5, AAV8 или AAV9.

17. Фармацевтическая композиция, содержащая полинуклеотид по любому из пп.1-10 или экспрессионный вектор по любому из пп.11-16.

18. Клетка, содержащая полинуклеотид по любому из пп.1-10 или экспрессионный вектор по любому из пп.11-16.

19. Применение полинуклеотида по любому из пп.1-10 в производстве лекарственного средства для лечения эпилепсии у нуждающегося в этом субъекта.

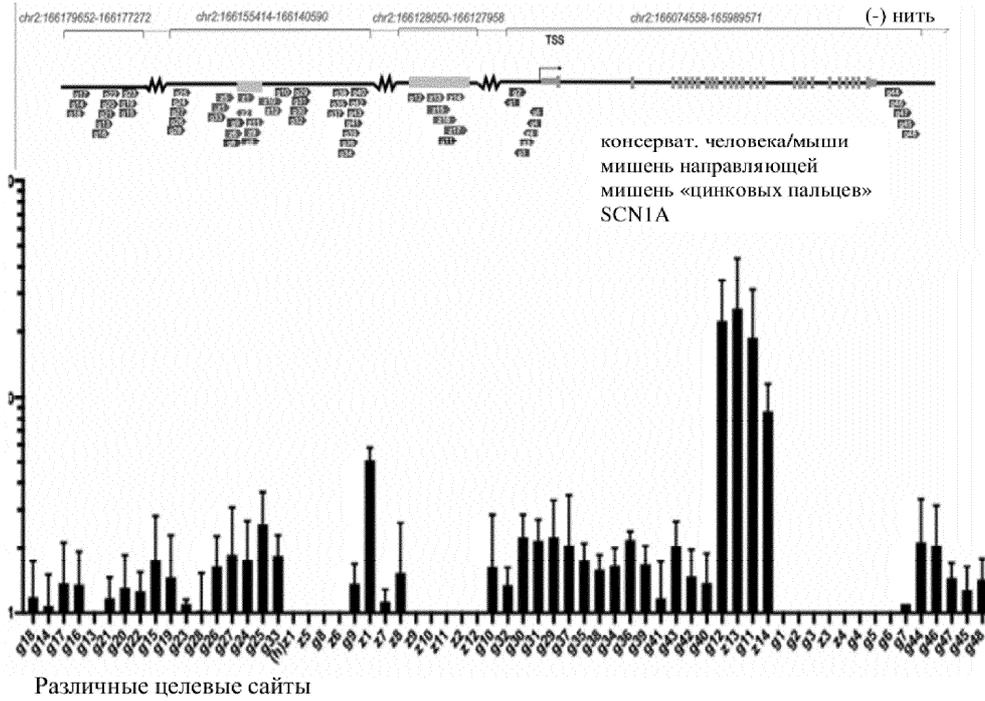
20. Применение полинуклеотида по любому из пп.1-10 в производстве лекарственного средства для лечения синдрома Драве у нуждающегося в этом субъекта.

21. Применение полинуклеотида по любому из пп.1-10 в производстве лекарственного средства для снижения частоты или продолжительности припадков.

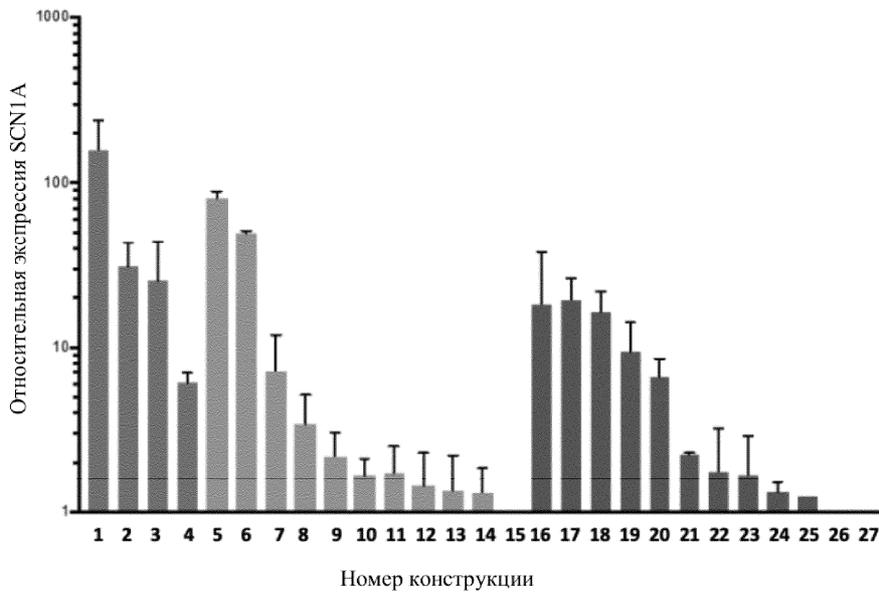
22. Применение экспрессионного вектора по любому из пп.11-16 в производстве лекарственного средства для лечения эпилепсии у нуждающегося в этом субъекта.

23. Применение экспрессионного вектора по любому из пп.11-16 в производстве лекарственного средства для лечения синдрома Драве у нуждающегося в этом субъекта.

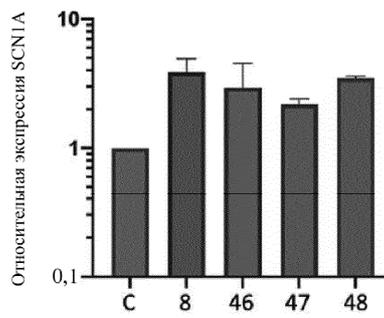
24. Применение экспрессионного вектора по любому из пп.11-16 в производстве лекарственного средства для снижения частоты или продолжительности припадков.



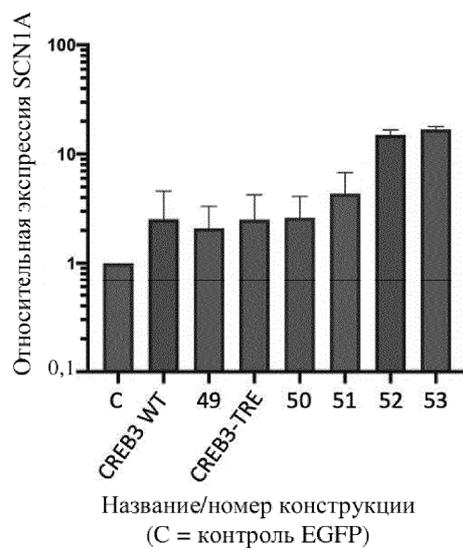
Фиг. 1



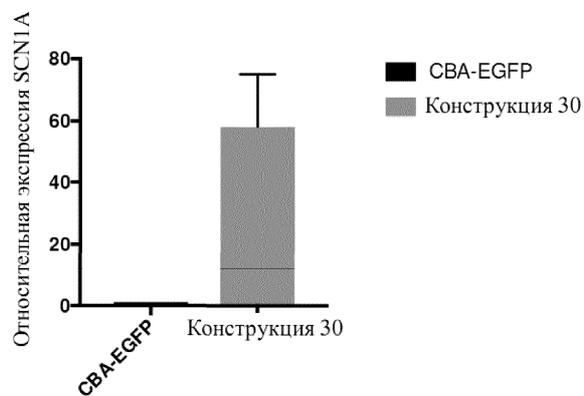
Фиг. 2А



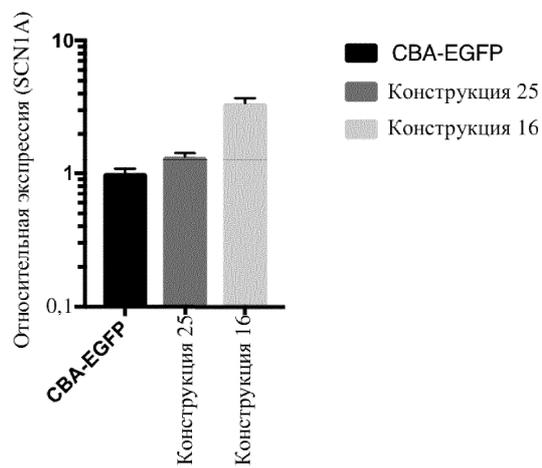
Фиг. 2В



Фиг. 2С

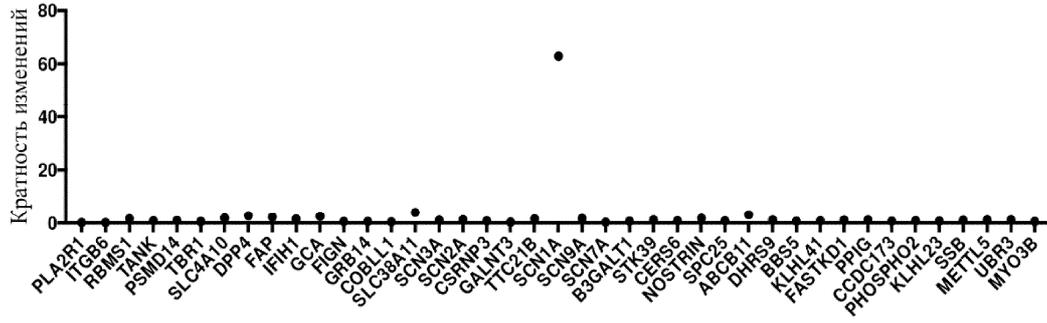


Фиг. 3А

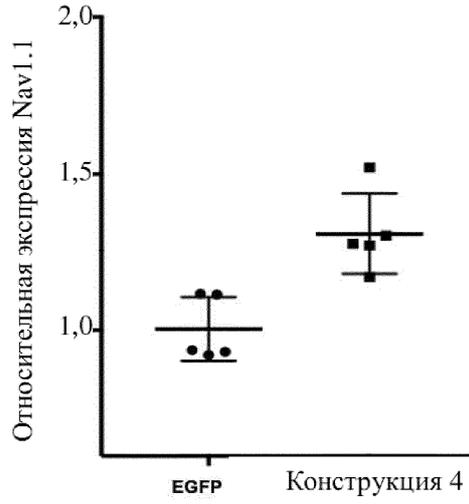


Фиг. 3В

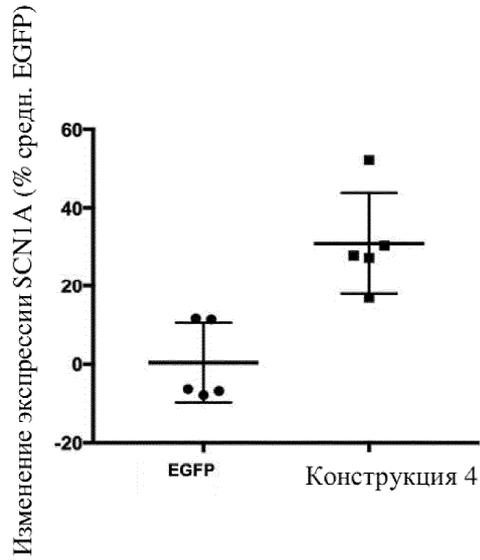
046157



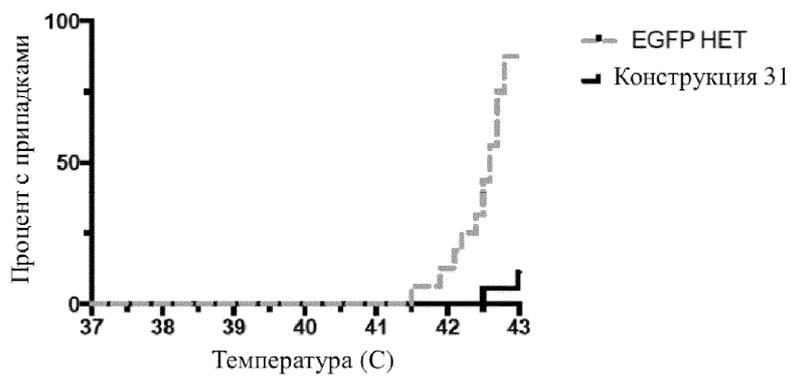
Фиг. 4



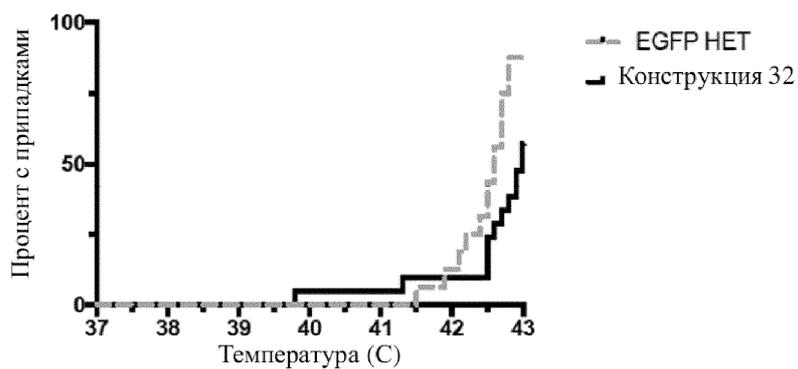
Фиг. 5A



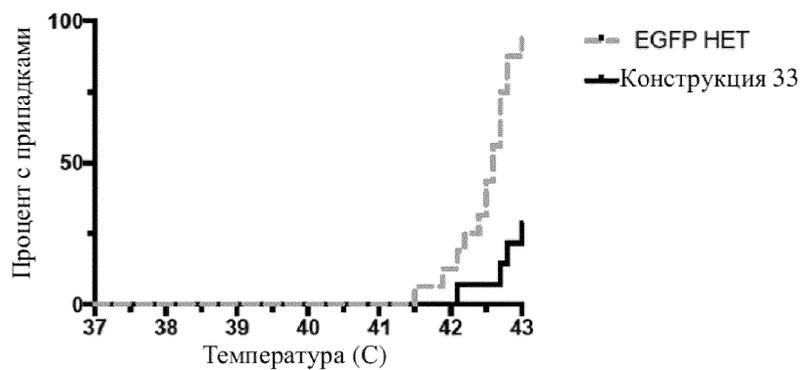
Фиг. 5B



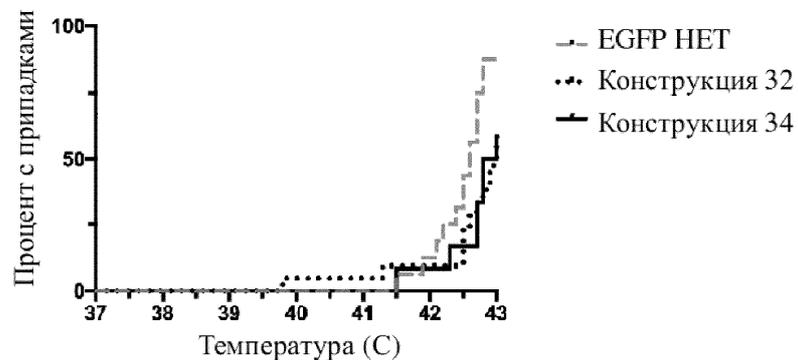
Фиг. 6А



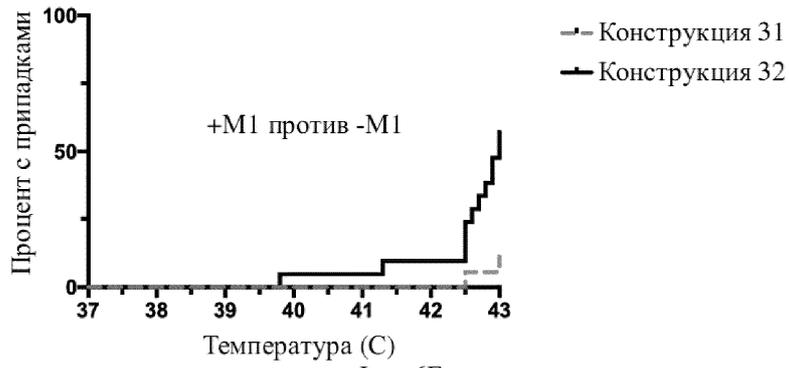
Фиг. 6В



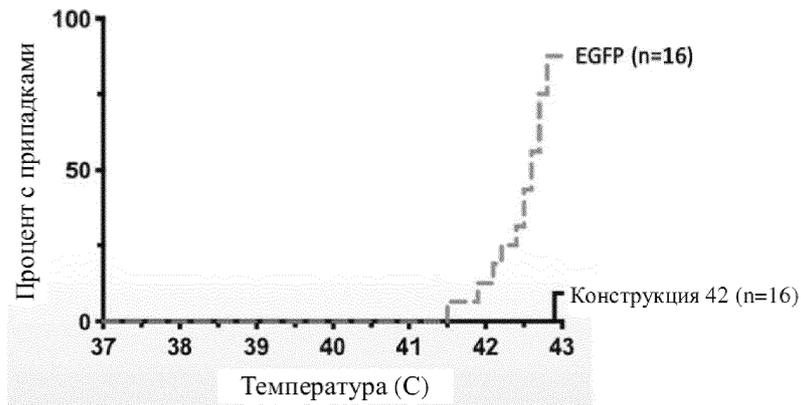
Фиг. 6С



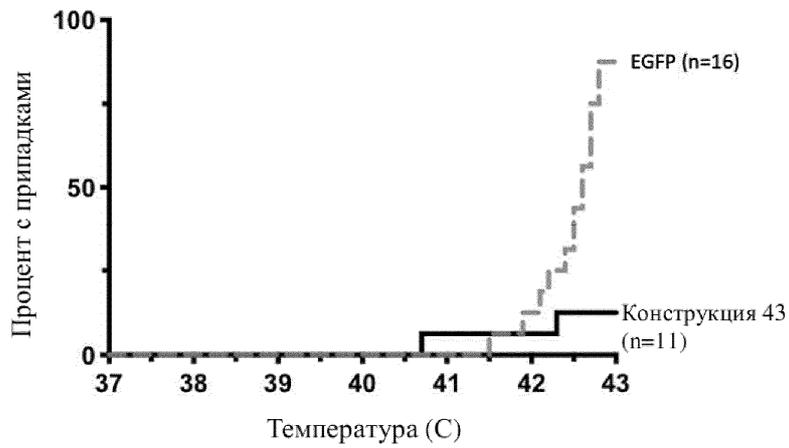
Фиг. 6D



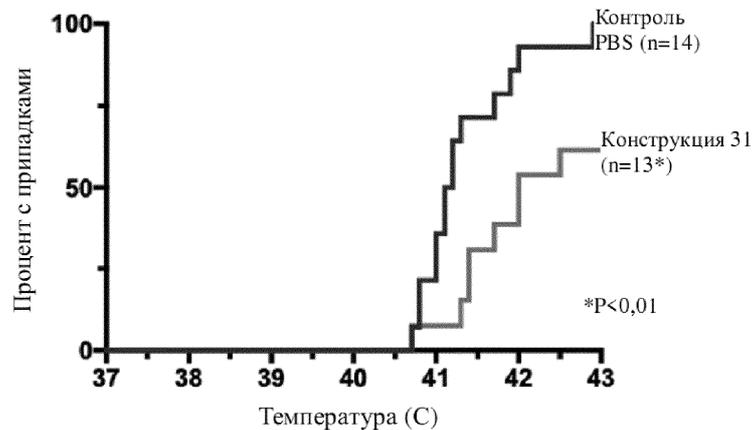
Фиг. 6E



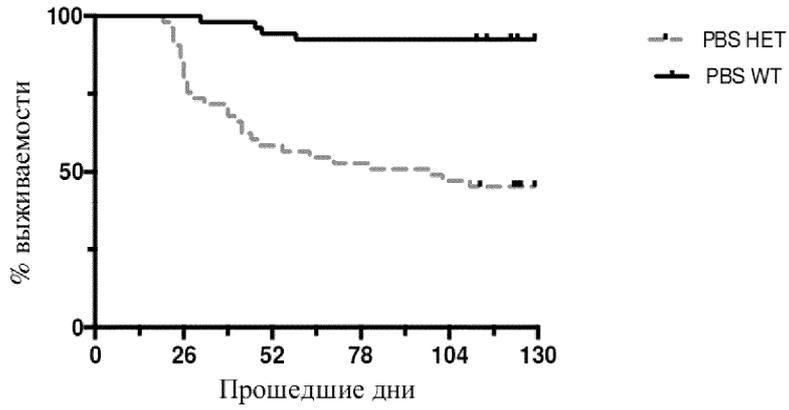
Фиг. 6F



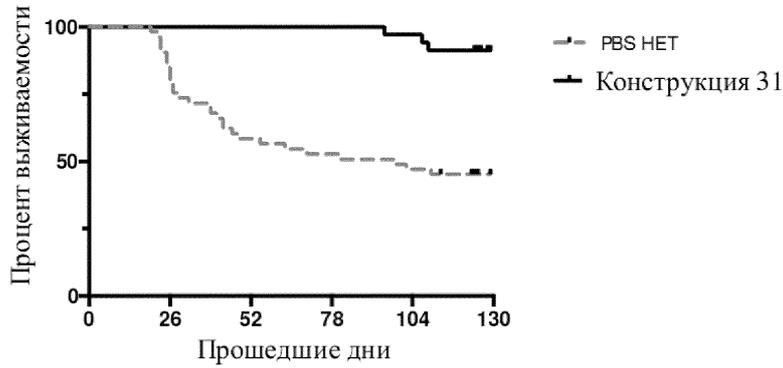
Фиг. 6G



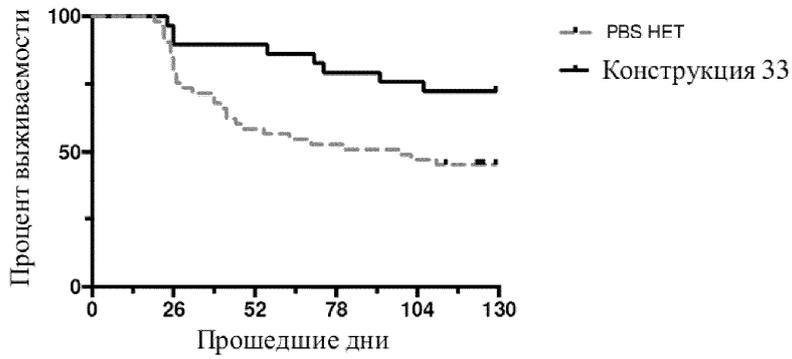
Фиг. 6H



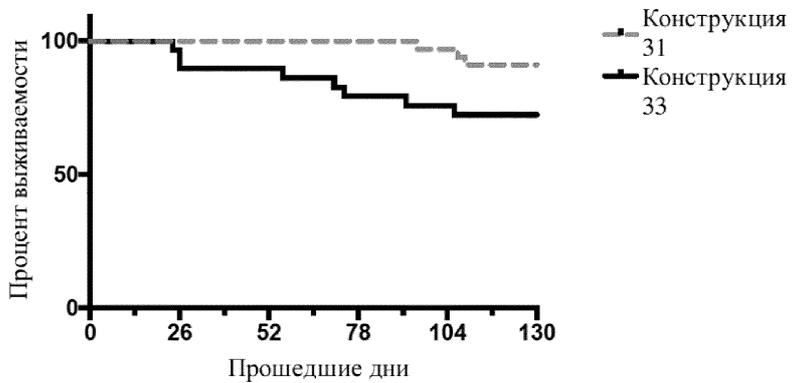
Фиг. 7А



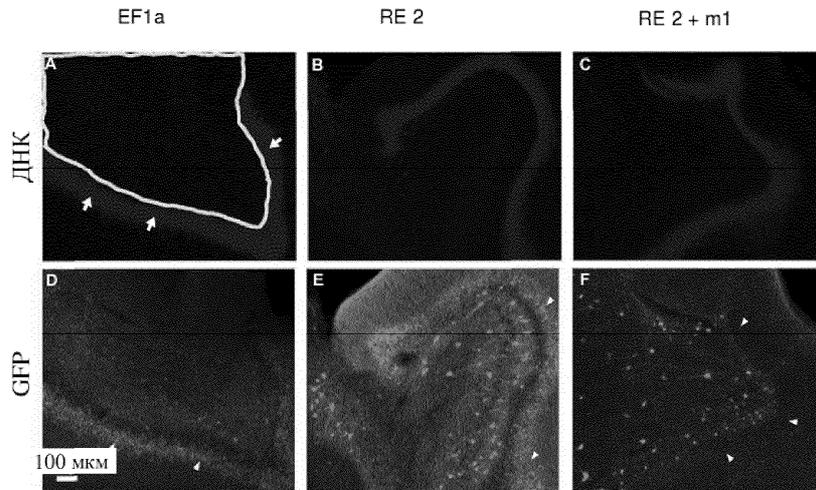
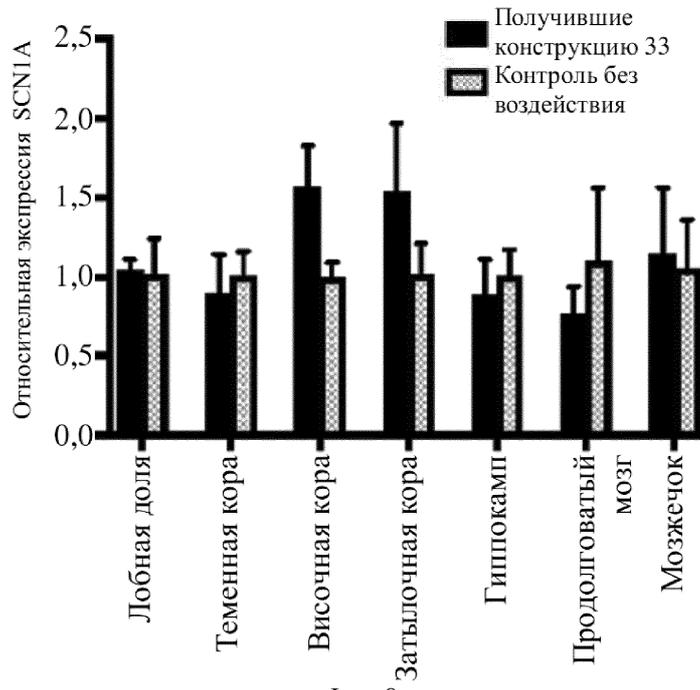
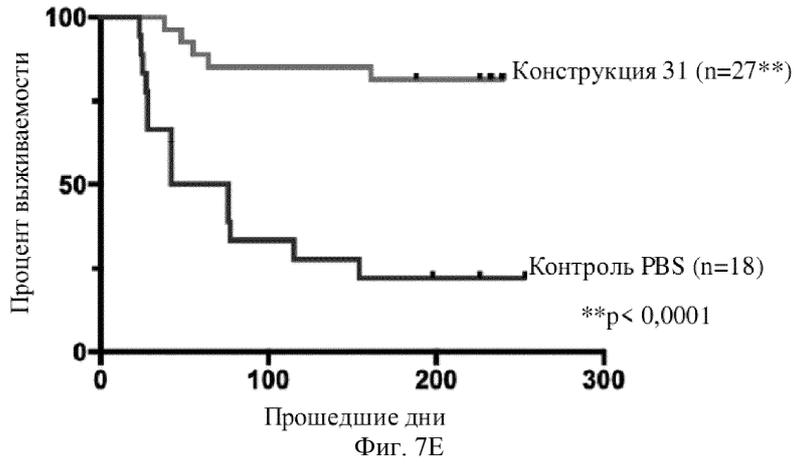
Фиг. 7В

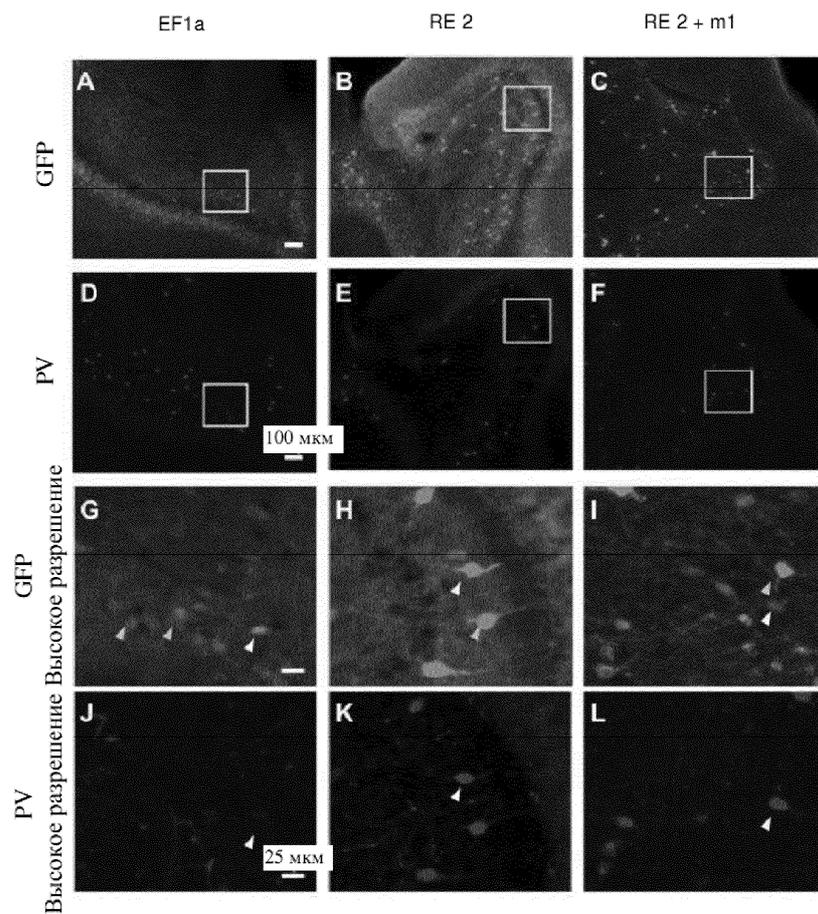


Фиг. 7С

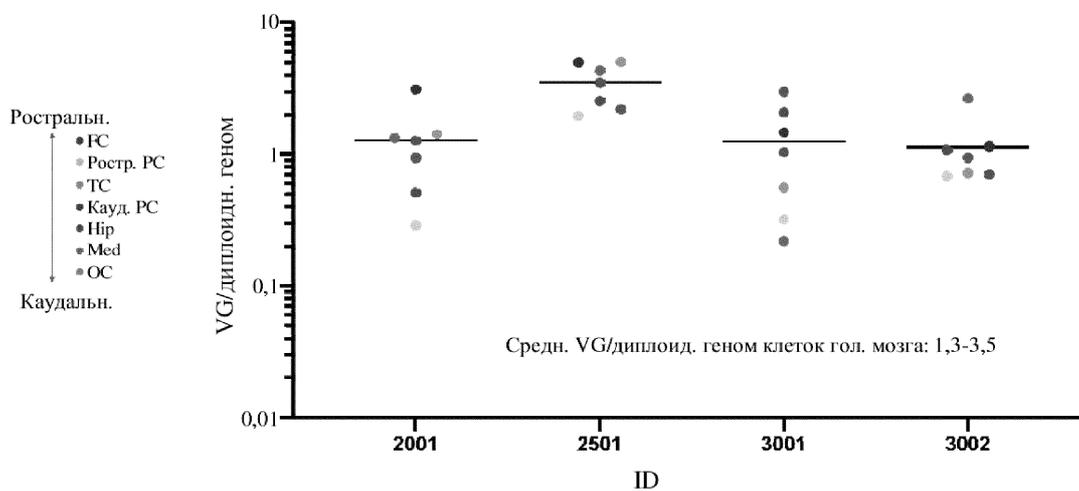


Фиг. 7D

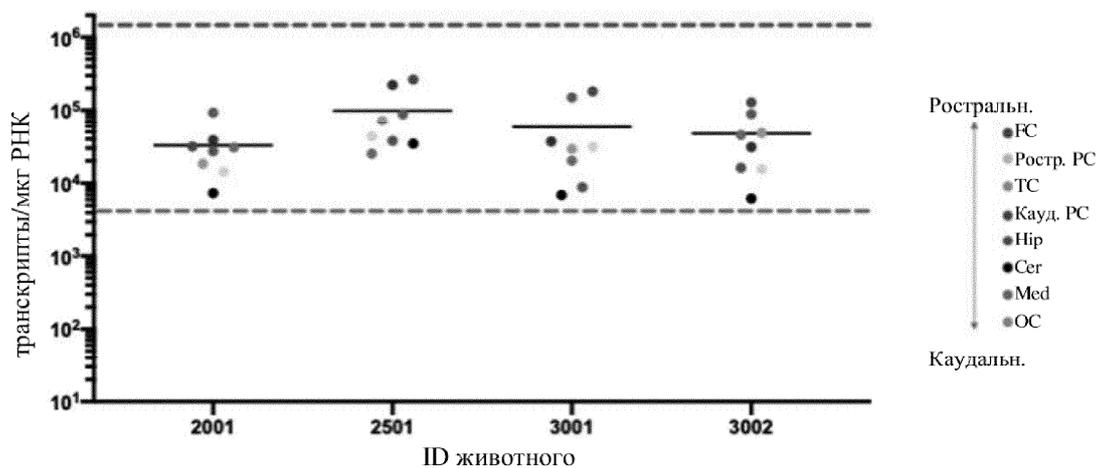




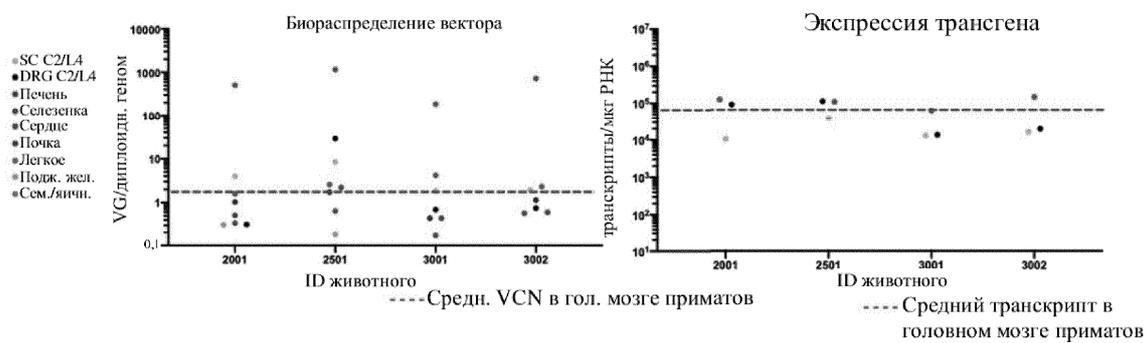
Фиг. 10



Фиг. 11



Фиг. 12



Фиг. 13

