



(19)

Евразийское
патентное
ведомство

(11) 027693

(13) В1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

- (45) Дата публикации и выдачи патента: 2017.08.31
(21) Номер заявки: 201170492
(22) Дата подачи: 2009.09.26

(51) Int. Cl. C12N 15/52 (2006.01)
C12N 5/10 (2006.01)
C12N 15/867 (2006.01)

(54) РЕКОМБИНАНТНЫЙ РЕПЛИКАЦИОННО-КОМПЕТЕНТНЫЙ РЕТРОВИРУС С ЦИТОЗИНДЕЗАМИНАЗНОЙ АКТИВНОСТЬЮ ДЛЯ ГЕННОЙ ТЕРАПИИ НАРУШЕНИЙ КЛЕТОЧНОЙ ПРОЛИФЕРАЦИИ

- (31) 61/100,666; 61/120,618; 61/186,823
(32) 2008.09.26; 2008.12.08; 2009.06.13
(33) US
(43) 2011.10.31
(86) PCT/US2009/058510
(87) WO 2010/045002 2010.04.22
(71)(73) Заявитель и патентовладелец:
ТОКАДЖЕН ИНК. (US)

- (72) Изобретатель:
Грубер Гарри Е., Джолли Дуглас, Перес
Омар, Логг Кристофер Р. (US)

- (74) Представитель:
Медведев В.Н. (RU)

(56) PHILIPPE ERBS et al. In Vivo Cancer Gene Therapy by Adenovirus-mediated Transfer of a Bifunctional Yeast Cytosine Deaminase/Uracil Phosphoribosyltransferase Fusion Gene. CANCER RESEARCH. 2000, vol. 60, no. 14, p. 3813-3822. See the whole document, especially, abstract, p. 3814-3818
GenBank Accession Number 1YSB_A, "Chain A, Yeast Cytosine Deaminase Triple Mutant", 24-SEP-2008, http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=Retrieve&db=protein&dopt=GenPept&RID=YENPGVH016&log%24=protalign&blast_rank=l&list_uids=67464318
GenBank Accession Number AAG33626, "Cytosine deaminase-uracil phosphoribosyltransferase fusion protein [synthetic construct]", 21-NOV-2000, http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=Retrieve&db=protein&dopt=GenPept&RID=YENPGVMH016&log%24=protalign&blast_rank=3&list_uids=11245466

B1

-
- (57) Изобретение относится к рекомбинантному репликационно-компетентному ретровирусу и его применению. Рекомбинантный ретровирус по настоящему изобретению способен трансформировать клетку-мишень человека полинуклеотидом, кодирующим полипептид с цитозиндезаминазной активностью и повышенной стабильностью в клетках человека, который катализирует превращение экзогенного 5-фторцитозина в 5-фторурацил, обеспечивая цитотокическое действие. Указанный рекомбинантный ретровирус применяется с 5-фторцитозином для получения лекарственного средства для лечения индивида с нарушением клеточной пролиферации.

027693

027693

B1

Перекрестная ссылка на родственные заявки

По заявке на данное изобретение испрашивается приоритет предварительной заявки США с серийным номером 61/100666, поданной 26 сентября 2008 г., предварительной заявки США с серийным номером 61/120618, поданной 8 декабря 2008 г., и предварительной заявки США с серийным номером 61/186823, поданной 13 июня 2009 г., описания которых включены в настоящий документ путем ссылки.

Предпосылки к созданию изобретения

Дрожжевая, или бактериальная, цитозинdezаминаза превращает нетоксичное антибиотическое про-лекарство 5-FC в цитотоксическое химиотерапевтическое средство 5-фторурацил (5-FU). У людей (и млекопитающих вообще) не известно наличие природного гена, кодирующего фермент со значительной цитозинdezаминазной активностью. Цитозинdezаминаза дрожжей или бактерий получила признание в лечении злокачественных заболеваний, путем использования доставки гена и вирусных векторов для доставки этого фермента, с последующей обработкой 5-FC, который затем превращался под действием этого фермента в цитотоксическое лекарственное средство (Miller et al., Can Res 62:773-780 2002; Kievit et al., Can Res 59:1417-1421, 1999).

Краткое изложение сущности изобретения

Цитозинdezаминазу или слитую конструкцию по настоящему изобретению доставляют с использованием системы для доставки генов (GDS). В другом аспекте полинуклеотид, кодирующий цитозинdezаминазу, доставляют с помощью GDS, которая представляет собой вирусный вектор или вектор, производный вируса. Вирусный вектор может быть реплицирующимся или нереплицирующимся и может представлять собой адено-вирусный вектор, коревой вектор, вектор вируса герпеса, ретровирусный вектор (в том числе лентивирусный вектор), рабдовирусный вектор, такой как вектор вируса визукулярного стоматита, реовирусный вектор, вектор вируса Сенеки Валли, поксовирусный вектор (в том числе векторы, полученные из вируса оспы животных или коровьей оспы), парвовирусный вектор (в том числе AAV вектор), альфа-вирусный вектор или другой вирусный вектор, известный специалисту в данной области. В одном варианте осуществления вирусный вектор может представлять собой репликативно-компетентный ретровирусный вектор, способный инфицировать реплицирующиеся клетки млекопитающих.

Репликативно-компетентный ретровирусный вектор может содержать ортогетровирус или, более характерно, гамма-ретровирусный вектор. В одном аспекте репликационно-компетентный ретровирусный вектор содержит внутренний сайт связывания рибосомы (IRES) 5' с полинуклеотидом, кодирующим цитозинdezаминазу по настоящему изобретению. В одном варианте осуществления полинуклеотид, кодирующий цитозинdezаминазу, находится в положении 3' по отношению к ENV полинуклеотиду ретровирусного вектора.

Настоящее изобретение относится к рекомбинантному репликационно-компетентному ретровирусу (RCR), содержащему: ретровирусный белок GAG; ретровирусный белок POL; оболочку ретровируса; ретровирусный полинуклеотид, содержащий последовательности длинных концевых повторов (LTR) на 3'-конце полинуклеотидной последовательности ретровируса, промоторную последовательность на 5'-конце ретровирусного полинуклеотида, указанный промотор подходит для экспрессии в клетках млекопитающих, домен нуклеиновой кислоты gag, домен нуклеиновой кислоты pol и домен нуклеиновой кислоты env; кассету, содержащую внутренний сайт связывания рибосомы (IRES) оперативно связанный с гетерологичным полинуклеотидом, где указанная кассета расположена от 5' к 3' LTR и 3' к домену нуклеиновой кислоты env, кодирующему оболочку ретровируса; и последовательности, действующие в цис-положении, необходимые для обратной транскрипции, упаковки и интеграции в клетку-мишень, где RCR сохраняет более высокую репливативную компетентность после 6 пассажей по сравнению с вектором pACE (SEQ ID NO: 21). В одном варианте осуществления полинуклеотидная последовательность ретровируса получена из вируса лейкоза мышей (MLV), вируса мышного лейкоза Молони (MoMLV), вируса лейкоза кошачьих или вируса лейкоза гиббонов (GALV). В другом варианте осуществления MLV представляет собой амфотропный MLV. Еще в одном варианте осуществления ретровирус представляет собой онкоретровирус. Еще в одном варианте осуществления клетка-мишень представляет собой клетку, имеющую нарушение клеточной пролиферации. Нарушение клеточной пролиферации может быть выбрано из группы, состоящей, но не ограничивающейся этим, из рака легких, рака толстой и прямой кишки, рака груди, рака предстательной железы, рака мочевыводящих путей, рака матки, рака головного мозга, рака тканей головы и шеи, рака поджелудочной железы, меланомы, рака желудка и рака яичников, ревматоидного артрита и других аутоиммунных заболеваний. В одном варианте осуществления промотор содержит CMV промотор, имеющий последовательность SEQ ID NO: 19, 20 или 22 от нуклеотида 1 до примерно нуклеотида 582 и может включать модификацию одного или нескольких оснований нуклеиновых кислот, и который способен направлять и инициировать транскрипцию. Еще в одном варианте осуществления промотор содержит последовательность SEQ ID NO: 19, 20 или 22 от нуклеотида 1 до примерно нуклеотида 582. В дополнительном варианте осуществления промотор содержит полинуклеотид с доменом CMV-R-U5. В одном варианте осуществления домен CMV-R-U5 содержит ранний промотор из цитомегаловируса человека, связанный с областью MLV R-U5. Еще в одном варианте осуществления полинуклеотид с доменом CMV-R-U5 содержит последовательность SEQ ID NO: 19, 20 или 22

примерно от нуклеотида 1 до примерно нуклеотида 1202 или последовательности, которые по меньшей мере на 95% идентичны последовательности SEQ ID NO: 19, 20 или 22, где указанный полинуклеотид активизирует транскрипцию молекулы нуклеиновой кислоты, оперативно связанной с ним. В другом варианте осуществления gag и pol полинуклеотида получены из онкоретровируса. Домен gag нуклеиновой кислоты может содержать последовательность примерно от нуклеотида номер 1203 до примерно нуклеотида 2819 последовательности SEQ ID NO: 19 или последовательность, по меньшей мере на 95%, 98, 99 или 99,8% идентичную ей. Домен pol может содержать последовательность примерно от нуклеотида номер 2820 до примерно нуклеотида 6358 последовательности SEQ ID NO: 19 или последовательность, по меньшей мере на 95, 98, 99 или 99,9% идентичную ей. В одном варианте осуществления домен env кодирует амфотерный белок env. Домен env может содержать последовательность примерно от нуклеотида номер 6359 до примерно нуклеотида 8323 последовательности SEQ ID NO: 19 или последовательность, по меньшей мере на 95, 98, 99 или 99,8% идентичную ей. Домен IRES может представлять собой любой IRES, однако в одном варианте осуществления IRES получают из вируса энцефаломиокардита. В дополнительном варианте осуществления IRES содержит последовательность примерно от нуклеотида номер 8327 до примерно нуклеотида 8876 последовательности SEQ ID NO: 19 или последовательность, по меньшей мере на 95, 98 или 99% идентичную ей. Еще в одном варианте осуществления гетерологичный полинуклеотид содержит полинуклеотид, имеющий последовательность SEQ ID NO: 3, 5, 11, 13, 15 или 17. В дополнительном варианте осуществления гетерологичная последовательность кодирует полипептид, содержащий последовательность SEQ ID NO: 4. Гетерологичная нуклеиновая кислота является оптимизированной для кодонов человека и кодирует полипептид последовательности SEQ ID NO: 4. В дополнительном варианте осуществления гетерологичная нуклеиновая кислота содержит последовательность SEQ ID NO: 19 примерно от нуклеотида номер 8877 примерно до 9353. В одном варианте осуществления 3' LTR получен из онкоретровируса. В дополнительном варианте осуществления 3' LTR содержит домен U3-R-U5. Еще в одном варианте осуществления 3' LTR содержит последовательность SEQ ID NO: 19 примерно от нуклеотида 9405 примерно до 9998 или последовательность, которая по меньшей мере на 95, 98 или 99,5% идентична ей.

Подробные описания одного или нескольких вариантов осуществления настоящего изобретения изложены в сопровождающих чертежах и в описании, приведенном ниже. Другие особенности, задачи и преимущества будут очевидны из описания и чертежей и формулы изобретения.

Краткое описание чертежей

На фиг. 1А-1Д представлено (а) схематическое изображение рекомбинантного ретровирусного вектора по настоящему изобретению; (б) плазмидная карта полинуклеотида по настоящему изобретению; (с и д) последовательность полинуклеотида по настоящему изобретению (SEQ ID NO: 19).

На фиг. 2А-2Д представлены схемы воспроизведения различных вариантов осуществления по настоящему изобретению, включая полипептиды с CD, OPRT и UPRT активностью.

На фиг. 3 показано, что наблюдаются более высокие уровни белка yCD2 по сравнению с белком yCD дикого типа в инфицированных клетках U-87.

На фиг. 4 показана стабильность вектора, содержащего полипептид CD по настоящему изобретению, и сравнение с другими векторами по настоящему изобретению.

На фиг. 5А-5В показано, что цитозиндезаминазная активность и вектор по настоящему изобретению обеспечивают сравнимую или лучшую экспрессию, и, следовательно, уничтожение инфицированных клеток крысы RG2 (5А) или U87 (5В) по сравнению с yCD активностью дикого типа (T5.0007), когда инфицированные клетки подвергаются воздействию повышенных уровней 5-FC.

На фиг. 6 показано, что специфическая активность в инфицированных клетках U87 T5.0002 (yCD2) выше, чем T5.0001 (частично оптимизированная yCD), которая выше, чем T5.0007 (wt yCD).

Однаковые обозначения на различных чертежах означают одинаковые элементы.

Подробное описание

Используемые в контексте настоящего изобретения и в прилагаемой формуле изобретения формы единственного числа включают обозначения множественного числа, если из контекста явным образом не следует иное. Таким образом, например, ссылка на "клетку" означает множество таких клеток, а ссылка на "средство" включает ссылку на одно или несколько средств, известных специалистам в данной области, и так далее.

Также использование "или" означает "и/или", если не указано иное. Аналогичным образом, "содержат", "содержит", "содержащий", "включают", "включает" и "включающий" являются взаимозаменяемыми и не предназначены для ограничения.

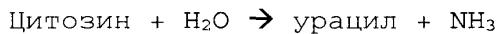
Далее должно быть понятно, что в тех случаях, когда в описаниях различных вариантов осуществления используется термин "включающий", специалисту в данной области было бы понятно, что в некоторых особых случаях вариант осуществления альтернативно может быть описан с использованием формулировки "по существу состоящий из" или "состоящий из".

Если не определено иное, все технические и научные термины, используемые в настоящем описании, имеют те же значения, обычно понимаемые специалистом в области, к которой относится настоящее изобретение. Хотя способы и материалы, сходные или эквивалентные описанным в настоящем до-

кументе, могут быть использованы при осуществлении описанных способов и композиций, в настоящем патенте описываются приводимые в качестве примера способы, устройства и материалы.

Публикации, рассмотренные выше и по всему тексту, представлены только для их раскрытия до даты подачи настоящей заявки. Ничто в настоящем описании не должно рассматриваться как признание того, что авторы изобретения не имеют права относить к более ранней дате такое описание в соответствии с предшествующим описанием.

Цитозиндезаминаза (ЕС 3.5.4.1) представляет собой фермент, который катализирует химическую реакцию:



Следовательно, два субстрата этого фермента представляют собой цитозин и H_2O , тогда как двумя продуктами этого фермента являются урацил и NH_3 . Этот фермент принадлежит семейству гидролаз, которые действуют на углерод-азотные связи, отличные от пептидных связей, в частности, в циклических амидах. Систематическим названием этого класса ферментов является цитозинаминогидролаза. Этот фермент также называют изоцитозиндезаминаза. Этот фермент участвует в метаболизме пириимида.

Более конкретно, цитозиндезаминаза представляет собой фермент, вовлеченный в метаболический путь пириимидинов, через который трансформируется экзогенный цитозин, посредством гидролитического дезаминирования, в урацил. Активности цитозиндезаминазы (CD-азы или CD) были продемонстрированы у прокариотов и низших эукариот, но они отсутствуют у млекопитающих (Koechlin et al., 1966, Biochem. Pharmacol. 15, 435-446; Polak et al., 1976, Chemotherapy, 22, 137-153). Ген FCY1 Saccharomyces cerevisiae (S.cerevisiae) и ген coda E.coli, которые кодируют соответственно CD-азу этих двух организмов, известны, и их последовательности опубликованы (EP 402108; Erbs et al., 1997, Curr. Genet. 31, 1-6; WO 93/01281). CD-аза также дезаминирует аналог цитозина, 5-фторцитозин (5-FC) в 5-фторурацил (5-FU), который является высокотоксичным соединением, в частности, когда он превращается в 5-фтор-UMP (5-FUMP). Клетки, лишенные CD-азной активности либо вследствие инактивирующей мутации гена, кодирующего этот фермент, либо вследствие природной недостаточности этого фермента (например, клетки млекопитающих), являются резистентными к 5-FC (Jund and Lacroute, 1970, J. Bacteriol. 102, 607-615; Kilstrup et al., 1989, J. Bacteriol., 171, 2124-2127). С другой стороны, было показано, что возможно передать 5-FC чувствительность клеткам млекопитающих, в которые была перенесена последовательность, кодирующая CD-азную активность (Huber et al., 1993, Cancer Res. 53, 4619-4626; Mullen et al., 1992, Proc. Natl. Acad. Sci. USA 89, 33-37; WO 93/01281). Соответственно, использование CD является предпочтительным применительно к генной терапии, в частности генной терапии злокачественных опухолей.

Однако чувствительность к 5-FC сильно варьирует в зависимости от клеточных линий. Низкая чувствительность наблюдается, например, в линиях опухолей человека PANC-1 (карцинома поджелудочной железы) и SK-BR-3 (аденокарцинома молочной железы), трансдуцированных ретровирусом, экспрессирующим ген coda E.coli (Harris et al., 1994, Gene Therapy 1, 170-175). Этот феномен объясняется отсутствием или слабой эндогенной конверсией 5-FU, образованного ферментативным действием CD-азы, в цитотоксический 5-FUMP. Эта стадия, которая обычно происходит в клетках млекопитающих под действием оротатфосфорибозилтрансферазы (OPRT-азы), может отсутствовать в некоторых опухолях и, следовательно, делать неэффективной генную терапию на основе CD-азы. У прокариот и низших эукариот урацил трансформируется в UMP через действие урацилфосфорибозилтрансферазы (UPRT-азной активности). Этот фермент также превращает 5-FU в 5-FUMP. Важно то, что бактериальная урацилфосфорибозилтрансфераза (UPRT) является функционально эквивалентной оротатфосфорибозилтрансферазе (OPRT) или уридин-5'-монофосфатсинтазе клеток млекопитающих. Эти ферменты опосредуют превращение 5-фторурацила (5-FU) (фторированного аналога урацила) в 5-фторуридин 5'-монофосфат (5-FUMP). 5-фторуридин 5' монофосфат впоследствии превращается в 5-FdUDP и FdUTP посредством de novo пириимидинового пути млекопитающих. Каждая 5-FdUTP является необратимым ингибитором тимидилатсинтазы (Thy-A) и в результате приводит к истощению dTTP. Принято считать, что это превращение является одним из необходимых путей для достижения цитотоксических эффектов 5-фторурацила и что бактериальная урацилфосфорибозилтрансфераза бактериального происхождения способна превращать 5-фторурацил в такой же активный метаболит, как оротатфосфорибозилтрансфераза млекопитающих. В отсутствие UPRT-азной активности или OPRT-азной активности 5-FU, возникающий в результате дезаминирования 5-FC под действием CD-азы, не трансформируется в цитотоксический 5-FUMP (Jund and Lacroute, 1970, J. Bacteriol. 102, 607-615).

"Мутация" означает любой процесс или механизм, который в результате приводит к получению мутантного белка, фермента, полинуклеотида, гена или клетки. Мутация включает в себя любую мутацию, в которой белок, фермент, полинуклеотид или генная последовательность изменены, любое определяемое изменение в клетке, возникающее в результате такой мутации. Обычно мутация происходит в полинуклеотидной или генной последовательности путем точковых мутаций, делеций или вставок единично-го или множественных нуклеотидных остатков. Мутация включает полинуклеотидные изменения, возникающие в области гена, кодирующего белок, а также изменения в областях вне последовательности,

кодирующей белок, такой как, но не только, регуляторные или промоторные последовательности. Мутация в гене может быть "молчашей", т.е. не отражающейся в изменении аминокислот при экспрессии, приводя к варианту гена с "консервативной последовательностью". В основном это происходит, когда одна аминокислота соответствует более чем одному кодону.

Неограничивающие примеры модифицированной аминокислоты включают гликозилированную аминокислоту, сульфатированную аминокислоту, пренилированную (например, фарнезилированную, геранилгеранилированную) аминокислоту, ацетилированную аминокислоту, ацилированную кислоту, пегилированную аминокислоту, биотинилированную аминокислоту, карбоксилированную аминокислоту, фосфорилированную аминокислоту и подобное. Ссылки, достаточные для инструктирования специалиста в отношении модификации аминокислот, в большом количестве встречаются повсюду в литературе. Примеры протоколов находятся у Walker (1998), Protein Protocols on CD-ROM (Humana Press, Towata, N.J.).

Для иллюстрации, последовательности нуклеиновой кислоты, кодирующие UPRT-азы *E.coli* (Anderson et al., 1992, Eur. J. Biochem 204, 51-56), *Lactococcus lactis* (Martinussen and Hammer, 1994, J. Bacteriol. 176, 6457-6463), *Mycobacterium bovis* (Kim et al., 1997, Biochem Mol. Biol. Int. 41, 1117-1124) и *Bacillus subtilis* (Martinussen et al., 1995, J. Bacteriol. 177, 271-274), могут быть использованы в контексте настоящего изобретения. Однако использование UPRT-азы дрожжей и, в частности, кодируемой геном FUR1 *S.cerevisiae*, последовательность которого описана у Kern et al. (1990, Gene 88, 149-157), включено в настоящее описание для ссылки. Чтобы обратить внимание, последовательности генов и последовательности соответствующих UPRT-аз могут быть найдены в литературе и специализированных базах данных (SWISSPROT, EMBL, Genbank Medline, etc.).

"Белок" или "полипептид" - термины, которые используются в настоящем описании взаимозаменяющими, содержат одну или несколько цепей химических структурных элементов, называемых аминокислотами, которые связаны вместе химическими связями, называемыми пептидными связями.

"Фермент" означает любое вещество, предпочтительно состоящее полностью или в значительной степени из белка, которое катализирует или активизирует более или менее специфично одну или несколько химических или биохимических реакций. Термин "фермент" также может относиться к катализитическому полинуклеотиду (например, РНК или ДНК). "Нативный" белок, фермент, полинуклеотид, ген или клетка, или белок, фермент, полинуклеотид, ген или клетка "дикого типа" означают белок, фермент, полинуклеотид, ген или клетку, которые встречаются в природе.

Термин "оперативно связанный" или "оперативно ассоциированный" относится к функциональной связи или ассоциации между регуляторной последовательностью и полинуклеотидом, регулируемым этой регуляторной последовательностью или между двумя различными полипептидами или полинуклеотидами, кодирующими такие полипептиды.

"Консервативная аминокислотная замена" или просто "консервативные варианты" конкретной последовательности относятся к замене одной аминокислоты или ряда аминокислот, по существу, на идентичные аминокислотные последовательности. Специалисту в данной области будет понятно, что индивидуальные замены, делеции или добавления, которые изменяют, добавляют или удаляют единичную аминокислоту или процентное содержание аминокислот в кодируемой последовательности, приводят к "консервативным вариантам", в которых изменения приводят к делеции аминокислоты, добавлению аминокислоты или замене аминокислоты химически сходной аминокислотой.

Таблицы консервативных замен, представляющие функционально сходные аминокислоты, хорошо известны из уровня техники. Например, одна группа консервативных замен включает аланин (A), серин (S) и треонин (T). Другая группа консервативных замен включает аспарагиновую кислоту (D) и глутаминовую кислоту (E). Другая группа консервативных замен включает аспарагин (N) и глутамин (Q). Еще одна группа консервативных замен включает аргинин (R) и лизин (K). Другая группа консервативных замен включает изолейцин (I), лейцин (L), метионин (M) и валин (V). Другая группа консервативных замен включает фенилаланин (F), тирозин (Y) и триптофан (W).

Таким образом, "консервативные аминокислотные замены" перечисленных полипептидных последовательностей (например, SEQ ID NO: 2, 4, 6, 8, 10, 11, 12, 13 и т.д.) включают замены процентного содержания, обычно менее 10%, аминокислот полипептидной последовательности, консервативно выбранной аминокислотой той же группы консервативных замен.

Соответственно, консервативно замещенный вариант полипептида по настоящему изобретению может содержать 100, 75, 50, 25 или 10 замен консервативно замещенных вариантом той же группы консервативных замен.

"Активность" фермента представляет собой критерий его способности катализировать реакцию, т.е. "функционировать", и может быть выражена как скорость, с которой образуется продукт реакции. Например, активность фермента может быть представлена как количество продукта, образованного за единицу времени или на единицу фермента (например, концентрация или масса) или в значениях аффинности или констант диссоциации. Используемая взаимозаменяющими в контексте настоящего изобретения "цитозиндезаминазная активность", "биологическая активность цитозиндезаминазы" или "функциональная активность цитозиндезаминазы" относится к активности, проявляемой белком или полипептидом

цитозиндезаминазы по настоящему изобретению, в отношении субстрата цитозиндезаминазы, как определено *in vivo* или *in vitro*, в соответствии со стандартными методиками. Анализы для измерения цитозиндезаминазной активности известны из уровня техники. Например, цитозиндезаминазную активность можно измерить путем определения скорости превращения 5-FC в 5-FU или цитозина в урацил. Определение 5-FC, 5-FU, цитозина и урацила может быть выполнено с помощью хроматографии и других способов, известных из уровня техники.

Специалисту в данной области будет понятно, что многие консервативные варианты конструкций нукleinовой кислоты, которые описаны, дают на выходе функционально идентичную конструкцию. Например, как рассмотрено выше, вследствие вырожденности генетического кода "молчание замены" (т.е. замены в последовательности нукleinовой кислоты, которая в результате не приводит к изменению в кодируемом полипептиде) означали свойство каждой последовательности нукleinовой кислоты, которая кодирует аминокислоту. Специалистам в данной области будет понятно, что вследствие вырожденности генетического кода может быть получено множество нуклеотидных последовательностей, кодирующих модифицированные цитозиндезаминазные полипептиды, некоторые из которых несут существенную идентичность с последовательностями нукleinовых кислот, явным образом описанных в настоящем изобретении. Например, кодоны AGA, AGG, CGA, CGC, CGG и CGU - все кодируют аминокислоту аргинин. Таким образом, в каждом положении в нукleinовых кислотах по настоящему изобретению, где аргинин определен кодоном, этот кодон может быть изменен в любой из соответствующих кодонов, описанных выше, без изменения кодируемого полипептида. Понятно, что U в последовательности РНК соответствует T в последовательности ДНК.

Аналогично, "консервативные аминокислотные замены", в одной или нескольких аминокислотах в аминокислотной последовательности, замещенные другими аминокислотами с очень похожими свойствами, также легко идентифицируются как в высокой степени сходные с описанной конструкцией. Такие консервативные варианты каждой описанной последовательности являются особенностью полипептидов, представленных в настоящем описании.

"Консервативные варианты" представляют собой белки или ферменты, в которых заданный аминокислотный остаток был заменен без изменения в целом конформации и функции белка или фермента, включая, но не ограничиваясь этим, замену аминокислоты аминокислотой, обладающей похожими свойствами, включая полярный или неполярный характер, размер, форму и заряд. Аминокислоты, отличные от аминокислот, указанных в качестве консервативных, могут отличаться в белке или ферменте так, чтобы сходство процентного состава белка или аминокислотной последовательности между любыми двумя белками аналогичной функции могли варьировать и могли составлять, например, по меньшей мере 30%, по меньшей мере 50%, по меньшей мере 70%, по меньшей мере 80% или по меньшей мере 90%, как определено в соответствии со схемой выравнивания. Как изложено в настоящем описании, "сходство последовательностей" означает степень, до которой нуклеотидные или белковые последовательности являются родственными. Степень сходства между двумя последовательностями может быть основана на проценте идентичности последовательностей и/или консервативности. "Идентичность последовательностей" в настоящем описании означает степень, до которой две нуклеотидные или аминокислотные последовательности являются инвариантными. "Выравнивание последовательностей" означает процесс выстраивания в одну линию двух или более последовательностей для достижения максимальных уровней идентичности (и, в случае аминокислотных последовательностей, консервативности) для целей определения степени сходства. Многочисленные способы выравнивания последовательностей и определения сходства/идентичности известны из уровня техники, например, такие как кластерный метод, в котором сходство основано на алгоритме MEGALIGN, а также BLASTN, BLASTP и FASTA (Lipman and Pearson, 1985; Pearson and Lipman, 1988). При использовании всех этих программ предпочтительными параметрами настройки являются те, которые в результате приводят к наибольшему сходству последовательностей. Например, "идентичность" или "процент идентичности" в отношении конкретной пары выравниваемых аминокислотных последовательностей может относиться к проценту идентичности аминокислотных последовательностей, полученному с помощью анализа ClustalW (версия W 1.8, доступная от Европейского института биоинформатики, Cambridge, UK), подсчета числа идентичных соответствий в выравнивании и деления этого числа идентичных соответствий на наибольшее из следующих показателей: (i) длина выравниваемых последовательностей и (ii) 96, и используя следующие параметры по умолчанию ClustalW для достижения медленных/точных попарных выравниваний-штраф за открытый разрыв: 10; штраф за удлинение разрыва: 0,10; матрица массы белка: Gonnet series; матрица массы ДНК: IUB; Toggle Slow/Fast попарные выравнивания = SLOW или FULL Alignment.

Две последовательности являются "оптимально выровненными", когда они выровнены по баллам сходства с использованием определенной матрицы аминокислотных замен (например, BLOSUM62), штрафа за разрыв и штрафа за удлинение разрыва, чтобы получить наиболее высокий балл, возможный для этой пары последовательностей. Матрицы аминокислотных замен и их использование в количественном определении сходства между двумя последовательностями хорошо известны из уровня техники и описаны, например, у Dayhoff et al. (1978), "A model of evolutionary change in proteins" в "Atlas of Protein Sequence and Structure", Vol. 5, Suppl. 3 (ed. M.O. Dayhoff), p. 345-352. Natl. Biomed. Res. Found., Washington-

ton, D.C. и Henikoff et al. (1992), Proc. Nat'l. Acad. Sci. USA 89: 10915-10919. Матрицу BLOSUM62 (фиг. 10) зачастую используют в качестве подстановочной матрицы по умолчанию для подсчета баллов в протоколах выравнивания последовательностей, таких как Gapped BLAST 2.0. Штраф за разрыв налагается на введение одного аминокислотного разрыва в одной из выровненных последовательностей, и штраф за удлинение разрыва налагается на каждое дополнительное пустое положение аминокислоты, включенное в уже открытый разрыв. Выравнивание определяется положениями аминокислот каждой последовательности, в которой выравнивание начинается и заканчивается, и необязательно вставкой разрыва или многочисленных разрывов в одной или обеих последовательностях для достижения наиболее возможного балла. Хотя оптимальное выравнивание и подсчет баллов может проводиться вручную, этот процесс облегчается использованием компьютерного алгоритма выравнивания, например, содержащий разрывы BLAST 2.0, описанный у Altschul et al. (1997) Nucl. Acids Res. 25: 3389-3402, и сделанный общедоступным на National Center for Biotechnology Information (NCBI) Website (www.ncbi.nlm.nih.gov). Оптимальные выравнивания, в том числе множественные выравнивания, могут быть получены с использованием, например, PSI-BLAST, доступного через веб-сайт NCB1 и описанный Altschul et al. (1997), Nucl. Acids Res. 25:3389-3402.

В отношении аминокислотной последовательности, которая является оптимально выровненной с эталонной последовательностью, аминокислотный остаток "соответствует" положению в эталонной последовательности, с которой этот остаток образует пару при выравнивании. "Положение" обозначается числом, которое последовательно определяет каждую аминокислоту в эталонной последовательности на основании ее положения относительно N-конца. Вследствие делеций, вставок, усечений, слияний и т.д., которые должны быть приняты во внимание при определении оптимального выравнивания, вообще, номер аминокислотного остатка в тестируемой последовательности, определяемый простым подсчетом от N-конца, не обязательно будет таким же, как номер его соответствующего положения в эталонной последовательности. Например, в том случае, когда имеет место делеция в выравниваемой тестируемой последовательности, там не будет аминокислоты, которая соответствует положению в эталонной последовательности в месте делеций. В том случае, когда имеет место вставка в выравниваемой эталонной последовательности, эта вставка не будет соответствовать какому-либо положению аминокислот в эталонной последовательности. В случае усечений или слияний, могут быть участки аминокислот либо в эталонной, либо выравниваемой последовательности, которые не соответствуют какой-либо аминокислоте в соответствующей последовательности.

Неконсервативными модификациями конкретного полипептида являются модификации, которые заменяют любую аминокислоту, не охарактеризованные как консервативная замена, например любая замена, которая пересекает связи шести групп, изложенных выше. Эти замены включают замены основных или кислых аминокислот нейтральными аминокислотами (например, Asp, Glu, Asn или Gln на Val, Ile, Leu или Met), ароматической аминокислоты - основными или кислыми аминокислотами (например, Phe, Тир или Трг на Asp, Asn, Glu или Gln) или любую другую замену, не заменяющую аминокислоту подобной аминокислотой. Основные боковые цепи включают лизин (K), аргинин (R), гистидин (H); кислые боковые цепи включают аспарагиновую кислоту (D), глутаминовую кислоту (E); незаряженные полярные боковые цепи включают глицин (G), аспарагин (N), глутамин (Q), серин (S), треонин (T), тирозин (Y), цистеин (C); неполярные боковые цепи включают аланин (A), валин (V), лейцин (L), изолейцин (I), пролин (P), фенилаланин (F), метионин (M), триптофан (W); бета-разветвленные боковые цепи включают треонин (T), валин (V), изолейцин (I); ароматические боковые цепи включают тирозин (Y), фенилаланин (F), триптофан (W), гистидин (H).

В соответствии с этим некоторые аминокислотные остатки в определенных положениях в полипептиде "исключены" из консервативных аминокислотных замен. Например, в настоящем описании представлен полипептид, содержащий последовательность SEQ ID NO: 4. В некоторых вариантах осуществления, полипептид может быть изменен консервативными аминокислотными заменами, как описано выше, однако некоторые остатки желательно оставлять незамещенными, например остатки в положениях 23, 108 и 140.

"Родительский" белок, фермент, полинуклеотид, ген или клетка представляют собой любой белок, фермент, полинуклеотид, ген или клетку, из которой любой другой белок, фермент, полинуклеотид, ген или клетка, происходят или получены с помощью любых способов, инструментов или методик, а сам родительский вариант может являться или может не являться природным или мутантным. Родительский полинуклеотид или ген кодирует родительский белок или фермент.

Полинуклеотид, полипептид или другой компонент является "выделенным", когда он частично или полностью отделен от компонентов, с которыми он обычно связан (других белков, нуклеиновых кислот, клеток, синтетических реагентов и т.д.). Нуклеиновая кислота или полипептид являются "рекомбинантными", когда они являются искусственными или сконструированными или получены из искусственного или сконструированного белка или нуклеиновой кислоты. Например, полинуклеотид, который встроен в вектор или любое другое гетерологичное положение, например в геном рекомбинантного организма, так, чтобы он не был связан с нуклеотидными последовательностями, которые обычно flankируют полинуклеотид, в том виде, как он находится в природе, представляет собой рекомбинантный полинуклеотид.

Белок, экспрессируемый *in vitro* или *in vivo* из рекомбинантного полинуклеотида является примером рекомбинантного полипептида. Аналогично этому, полинуклеотидная последовательность, которая не встречается в природе, например вариант природного гена, является рекомбинантной.

Важно отметить, что некоторые специфические, важные и достоверные применения СД полинуклеотидов не требуют, чтобы этот полинуклеотид кодировал полипептид со значительной СД активностью или даже активностью варианта СД. Например, СД полинуклеотиды, которые не кодируют активные ферменты, могут быть ценными источниками родительских полинуклеотидов для применения в методиках диверсификации для получения вариантов СД полинуклеотида или не-СД полинуклеотидов, с желаемыми функциональными свойствами (например, высокой k_{cat} или k_{cat}/K_m , низкой K_m , высокой стабильностью в отношении нагревания или других факторов окружающей среды, высокой скоростью транскрипции или трансляции, устойчивости к протеолитическому расщеплению и т.д.).

СД полинуклеотиды, в том числе нуклеотидные последовательности, которые кодируют СД полипептиды и их варианты, фрагменты полипептидов СД, родственные слитые белки или их функциональные эквиваленты, используют в рекомбинантных молекулах ДНК, которые направляют экспрессию полипептидов СД в соответствующих клетках-хозяевах, таких как бактериальные клетки. Вследствие присущей генетическому коду вырожденности другие последовательности нукleinовой кислоты, которые кодируют, по существу, такую же или функционально эквивалентную аминокислотную последовательность, также могут быть использованы для клонирования и экспрессии полинуклеотидов СД.

Используемый в настоящем описании термин "клетка-хозяин" включает клетку любого типа, которая подвержена трансформации конструкций нукleinовой кислоты. Термин "трансформация" означает введение чужеродного (например, постороннего или внеклеточного) гена, последовательности ДНК или РНК в клетку-хозяина так, чтобы клетка-хозяин экспрессировала введенный ген или последовательность для продукции желаемого вещества, обычно белка или фермента, кодируемого введенным геном или последовательностью. Введенный ген или последовательность может включать регуляторные или контрольные последовательности, например стартовую, стоп, промоторную, сигнальную, секрецию или другие последовательности, используемые генетическим аппаратом клетки. Клетка-хозяин, которая получает и экспрессирует введенную ДНК или РНК, была "трансформирована" и является "трансформантом" или "клоном". ДНК или РНК, введенная в клетку-хозяина, может быть получена из любого источника, в том числе клеток того же рода или вида, как клетка-хозяин или клеток другого рода или вида.

Как будет понятно специалисту в данной области, может быть предпочтительным модифицировать кодирующую последовательность для усиления ее экспрессии в конкретной клетке-хозяине. Генетический код дублируется 64 возможными кодонами, но большинство организмов предпочтительно используют подгруппу этих кодонов. Кодоны, которые используются наиболее часто в видах, называют оптимальными кодонами, а те, которые не используются очень часто, классифицируют как редкие или редко используемые кодоны (см., например, Zhang et al. (1991), Gene, 105:61-72). Кодоны могут быть заменены для отражения предпочтительной частоты использования кодонов клетки-хозяина, процесс - иногда называемый "оптимизацией кодона" или "контролем за отклонением разновидности кодонов".

Оптимизированные кодирующие последовательности, содержащие кодоны, предпочтительные у конкретной прокариотической или эукариотической клетки-хозяина (см., также Murray et al. (1989), Nucl. Acids Res. 17:477-508) могут быть получены, например, для увеличения скорости трансляции или для получения рекомбинантных РНК транскриптов, обладающих желаемыми свойствами, такими как более длинный период полужизни, по сравнению с транскриптами, продуцируемыми из неоптимизированной последовательности. Стоп-кодоны трансляции также могут быть модифицированы для отражения предпочтения клетки-хозяина. Например, предпочтительными стоп-кодонами для *S. cerevisiae* и млекопитающих являются UAA и UGA соответственно.

Предпочтительным стоп-кодоном для однодольных растений является UGA, тогда как насекомые и *E.coli* предпочитают использовать UAA в качестве стоп-кодона (Dalphin et al. (1996), Nucl. Acids Res. 24: 216-218).

Отклонения в частоте использования кодонов относятся к различиям среди организмов в частоте появления кодонов в последовательностях ДНК (генах), кодирующих белок. Кодон представляет собой серию трех нуклеотидов (триплетов), который кодирует определенный аминокислотный остаток в полипептидной цепи. Поскольку в ДНК существует четыре нуклеотида, аденин (A), гуанин (G), цитозин (C) и тимин (T), существует 64 возможных триплета, кодирующих 20 аминокислот, и три кодона терминации трансляции (нансенс). Вследствие этой избыточности, все кроме двух аминокислот кодируются более чем одним триплетом. Различные организмы зачастую демонстрируют конкретные предпочтения одного из нескольких кодонов, которые кодируют одну и ту же аминокислоту. То, как это предпочтение возникает, является очень обсуждаемой областью молекулярного развития.

В основном является общепризнанным то, что предпочтения кодонов отражают баланс между мутационными отклонениями и природным отбором оптимизации трансляции. Оптимальные кодоны в быстрорастущих микроорганизмах, наподобие *Escherichia coli* или *Saccharomyces cerevisiae* (пекарских дрожжах), отражают состав их соответствующего геномного пула тРНК. Считается, что оптимальные кодоны помогают достигать более высоких скоростей трансляции и высокой точности. В результате этих

факторов ожидается, что трансляционная селекция является более сильной в высокоэкспрессируемых генах, которая необходима в случае указанных выше микроорганизмов. В других организмах, которые не демонстрируют высокие скорости роста или которые представляют небольшие геномы, оптимизация частоты использования кодонов обычно отсутствует и предпочтения кодонов определяются характерными мутационными отклонениями, наблюдаемыми в этом конкретном геноме. Примерами этого являются *Homo sapiens* (человек) и *Helicobacter pylori*. Организмы, которые демонстрируют промежуточный уровень оптимизации использования кодонов, включают *Drosophila melanogaster* (плодовая мушка), *Caenorhabditis elegans* (нematодные черви) или *Arabidopsis thaliana* (резуха).

Термин "последовательности с оптимизированными кодонами" в основном относится к нуклеотидным последовательностям, которые были оптимизированы для конкретных видов клеток-хозяев путем замены любых кодонов, имеющих частоту использования примерно менее 20%. Нуклеотидные последовательности, которые были оптимизированы для экспрессии в заданных видах клеток-хозяев путем элиминации случайных последовательностей полиаденилирования, элиминации экзон/инtronных сигналов сплайсинга, элиминации транспозон-подобных повторов и/или оптимизации содержания GC в дополнение к оптимизации кодонов, называется в настоящем описании "последовательностями усиленной экспрессии".

Таблица 1

Частота использования кодонов человека и предпочтение кодонов. Для каждого кодона в таблице показана частота использования каждого кодона (на тысячу) в кодирующих областях человека (первая колонка) и относительная частота каждого кодона среди синонимичных кодонов (вторая колонка)

Таблица частоты использования кодонов человека															
Gly	GGG	17,08	0,23	Arg	AGG	12,09	0,22	Trp	TGG	14,74	1,00	Arg	CGG	10,40	0,19
Gly	GGA	19,31	0,26	Arg	AGA	11,73	0,21	End	TGA	2,64	0,61	Arg	CGA	5,63	0,10
Gly	GGT	13,66	0,18	Ser	AGT	10,18	0,14	Cys	TGT	9,99	0,42	Arg	CGT	5,16	0,09
Gly	GGC	24,94	0,33	Ser	AGC	18,54	0,25	Cys	TGC	13,86	0,58	Arg	CGC	10,82	0,19
Glu	GAG	38,82	0,59	Lys	AAG	33,79	0,60	End	TAG	0,73	0,17	Gln	CAG	32,95	0,73
Glu	GAA	27,51	0,41	Lys	AAA	22,32	0,40	End	TAA	0,95	0,22	Gln	CAA	11,94	0,27
Asp	GAT	21,45	0,44	Asn	AAT	16,43	0,44	Tyr	TAT	11,80	0,42	His	CAT	9,56	0,41
Asp	GAC	27,06	0,56	Asn	AAC	21,30	0,56	Tyr	TAC	16,48	0,58	His	CAC	14,00	0,59
Val	GTG	28,60	0,48	Met	ATG	21,86	1,00	Leu	TTG	11,43	0,12	Leu	CTG	39,93	0,43
Val	GTA	6,09	0,10	Ile	ATA	6,05	0,14	Leu	TTA	5,55	0,06	Leu	CTA	6,42	0,07
Val	GTT	10,30	0,17	Ile	ATT	15,03	0,35	Phe	TTT	15,36	0,43	Leu	CTT	11,24	0,12
Val	GTC	15,01	0,25	Ile	ATC	22,47	0,52	Phe	TTC	20,72	0,57	Leu	CTC	19,14	0,20
Ala	GCG	7,27	0,10	Thr	ACG	6,80	0,12	Ser	TCG	4,38	0,06	Pro	CCG	7,02	0,11
Ala	GCA	15,50	0,22	Thr	ACA	15,04	0,27	Ser	TCA	10,96	0,15	Pro	CCA	17,11	0,27
Ala	GCT	20,23	0,28	Thr	ACT	13,24	0,23	Ser	TCT	13,51	0,18	Pro	CCT	18,03	0,29
Ala	GCC	28,43	0,40	Thr	ACC	21,52	0,38	Ser	TCC	17,37	0,23	Pro	CCC	20,51	0,33

Специалисту в данной области будет понятно, что изменения могут быть введены путем мутации в нуклеотидные последовательности любых полинуклеотидов, кодирующих полипептид, изложенный в любой из последовательностей: SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16 или 18 или имеющей нуклеотидную последовательность SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15 или 17, приводя, тем самым, к изменениям в аминокислотной последовательности кодируемого белков. В некоторых случаях изменение приведет к измененной функции этого полипептида. В других случаях это изменение не будет изменять функциональную способность кодируемого полипептида. В основном замены, которые не изменяют функцию полипептида, включают нуклеотидные замены, приводящие к аминокислотным заменам в "несущественных" аминокислотных остатках. "Несущественным" аминокислотным остатком является остаток, который может быть изменен из родительской последовательности без изменения биологической активности полученного в результате полипептида, например, катализирующего превращение 5-FC в 5-FU.

Также рассматриваются те ситуации, в которых желательно изменить активность родительского полинуклеотида, так, чтобы этот полипептид имел новую или повышенную активность в отношении конкретного субстрата или повышенную стабильность или сниженное разрушение. Понятно, что эти аминокислотные замены в основном не составляют "консервативные" замены. Вместо этого, эти замены составляют неконсервативные замены, введенные в последовательность для получения новой или улучшенной активности. Например, последовательность SEQ ID NO: 1 предоставляет родительскую последовательность нуклеиновой кислоты для цитозиндезаминазы *S.cervisiae* (SEQ ID NO: 2 представляет соответствующий полипептид). Последовательность SEQ ID NO: 3 представляет последовательность нуклеиновой кислоты мутантной последовательности, которая включает аминокислотные замены, которые придают повышенную стабильность этому полипептиду. Соответственно, молекула нуклеиновой кислоты, кодирующая аминокислотную последовательность SEQ ID NO: 2, представляет "родительскую" молекулу нуклеиновой кислоты, которая может быть мутирована для получения молекулы нуклеиновой

кислоты, которая кодирует модифицированный полипептид, который включает аминокислотные замены.

Также должно быть понятно, что модифицированный полипептид может составлять "родительский" полипептид из которого могут быть получены дополнительные замены. Соответственно, родительский полипептид и молекула нуклеиновой кислоты, которая кодирует родительский полипептид, включает модифицированные полипептиды и не только последовательности "дикого типа". Например, полинуклеотид последовательности SEQ ID NO: 5 представляет собой модифицированный полинуклеотид относительно последовательности SEQ ID NO: 1 (т.е. "родительский" полинуклеотид). Аналогично, полинуклеотид последовательности SEQ ID NO: 3 представляет собой модифицированный полинуклеотид относительно SEQ ID NO: 5. Соответственно, последовательность SEQ ID NO: 5 представляет собой родительскую последовательность SEQ ID NO: 3.

Мутационные способы создания разнообразия включают, например, сайт-направленный мутагенез (Ling et al. (1997), "Approaches to DNA mutagenesis: an overview", Anal Biochem. 254(2):157-178; Dale et al., (1996), "Oligonucleotide-directed random mutagenesis using the phosphorothioate method", Methods Mol. Biol. 57:369-374; Smith (1985), "In vitro mutagenesis", Ann. Rev. Genet. 19:423-462; Botstein & Shortle (1985), "Strategies and applications of in vitro mutagenesis", Science 229:1193-1201; Carter (1986), "Site-directed mutagenesis", Biochem. J. 237:1-7; и Kunkel (1987), "The efficiency of oligonucleotide directed mutagenesis" в Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D.M. J. eds., Springer Verlag, Berlin)); мутагенез с использованием матриц, содержащих урацил (Kunkel (1985), "Rapid and efficient site-specific mutagenesis without phenotypic selection", Proc. Natl. Acad. Sci. USA, 82:488-492; Kunkel et al. (1987), "Rapid and efficient site-specific mutagenesis without phenotypic selection", Methods в Enzymol. 154, 367-382; и Bass et al. (1988), "Mutant Trp repressors with new DNA-binding specificities", Science, 242:240-245); олигонуклеотид-направленный мутагенез (Methods in Enzymol. 100: 468-500 (1983); Methods in Enzymol. 154: 329-350 (1987); Zoller & Smith (1982), "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment", Nucleic Acids Res. 10:6487-6500; Zoller & Smith (1983), "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors" Methods in Enzymol. 100:468-500; и Zoller & Smith (1987), "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template", Methods в Enzymol. 154:329-350); фосфоротиоат-модифицированный ДНК мутагенез (Taylor et al. (1985), "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA", Nucl. Acids Res. 13:8749-8764; Taylor et al. (1985), "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA", Nucl. Acids Res. 13: 8765-8787; Nakamaye & Eckstein (1986), "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis", Nucl. Acids Res. 14:9679-9698; Sayers et al. (1988), "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis", Nucl. Acids Res. 16:791-802; и Sayers et al. (1988), "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" Nucl. Acids Res. 16:803-814); мутагенез с использованием двухспиральной ДНК, содержащей разрывы (Kramer et al. (1984), "The gapped duplex DNA approach to oligonucleotide-directed mutation construction", Nucl. Acids Res. 12:9441-9456; Kramer & Fritz (1987), Methods in Enzymol. "Oligonucleotide-directed construction of mutations via. gapped duplex DNA", 154:350-367; Kramer et al. (1988), "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations", Nucl. Acids Res. 16: 7207; и Fritz et al. (1988), "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro", Nucl. Acids Res. 16:6987-6999).

Дополнительные подходящие способы включают мочковое исправление ошибок спаривания (Kramer et al. (1984), "Point Mismatch Repair", Cell 38:879-887), мутагенез с использованием штаммов-хозяев с недостаточностью системы reparации (Carter et al. (1985), "Improved oligonucleotide site-directed mutagenesis using M13 vectors", Nucl. Acids Res. 13:4431-4443; и Carter (1987), "Improved oligonucleotide-directed mutagenesis using M13 vectors", Methods в Enzymol. 154:382-403), делеционный мутагенез (Eghitedarzadeh & Henikoff (1986), "Use of oligonucleotides to generate large deletions", Nucl. Acids Res. 14:5115), рестрикционную и рестрикционную очистку (Wells et al. (1986), "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin", Phil. Trans. R. Soc. Lond. A 317:415-423), мутагенез с помощью суммарного синтеза генов (Nambiar et al. (1984), "Total synthesis and cloning of a gene coding for the ribonuclease S protein", Science 223:1299-1301; Sakamar и Khorana (1988), "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)", Nucl. Acids Res. 14:6361-6372; Wells et al. (1985), "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites", Gene 34:315-323; и Grundstrom et al. (1985), "Oligonucleotide-directed mutagenesis by microscale "shot-gun", gene synthesis", Nucl. Acids Res. 13:3305-3316); reparацию двухспиральных разрывов (Mandecki (1986); Arnold (1993), "Protein engineering for unusual environments", Current Opinion в Biotechnology 4:450-455; и "Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis", Proc. Natl. Acad. Sci. USA, 83:7177-7181). Дополнительные подробности, касающиеся многих из указанных выше способов, можно найти в Methods in Enzymology Volume 154, которые также описывают полезные способы контро-

ля для решения проблем поиска неисправностей с использованием различных способов мутагенеза.

Дополнительные подробности, касающиеся различных способов создания разнообразия, можно найти в следующих патентах США, публикациях РСТ и публикациях ЕРО: патент США № 5605793 на имя Stemmer (25 февраля 1997 г.), "Methods for In vitro Recombination", патент США № 5811238 на имя Stemmer et al. (22 сентября 1998 г.), "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination", патент США № 5830721 на имя Stemmer et al. (3 ноября 1998 г.), "DNA Mutagenesis by Random Fragmentation and Reassembly", патент США № 5834252 на имя Stemmer, et al. (10 ноября 1998 г.), "End-Complementary Polymerase Reaction", патент США № 5837458 на имя Minshull, et al. (17 ноября 1998 г.), "Methods and Compositions for Cellular and Metabolic Engineering", WO 95/22625, Stemmer и Crameri, "Mutagenesis by Random Fragmentation and Reassembly", WO 96/33207, поданной Stemmer и Lipschutz "End Complementary Polymerase Chain Reaction", WO 97/20078, поданной Stemmer и Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination", WO 97/35966, поданной Minshull и Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering", WO 99/41402, поданной Punnonen et al. "Targeting of Genetic Vaccine Vectors", WO 99/41383, поданной Punnonen et al. "Antigen Library Immunization", WO 99/41369, поданной Punnonen et al. "Genetic Vaccine Vector Engineering", WO 99/41368, поданной Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines", EP 752008, поданной Stemmer и Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly", EP 0932670, поданной Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination", WO 99/23107, поданной Stemmer et al., "Modification of Virus Tropism and Host Range by Viral Genome Shuffling", WO 99/21979, поданной Apt et al., "Human Papillomavirus Vectors", WO 98/31837, поданной del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination", WO 98/27230, поданной Patten и Stemmer, "Methods and Compositions for Polypeptide Engineering", WO 98/13487, поданной Stemmer et al., "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection", WO 00/00632, "Methods for Generating Highly Diverse Libraries", WO 00/09679, "Methods for Obtaining in vitro Recombined Polynucleotide Sequence Banks and Resulting Sequences", WO 98/42832, поданной Arnold et al., "Recombination of Polynucleotide Sequences Using Random or Defined Primers", WO 99/29902, поданной Arnold et al., "Method for Creating Polynucleotide and Polypeptide Sequences", WO 98/41653, поданной Vind, "An in vitro Method for Construction of a DNA Library", WO 98/41622, поданной Borchert et al., "Method for Constructing a Library Using DNA Shuffling", WO 98/42727, поданной Pati and Zarling, "Sequence Alterations using Homologous Recombination", WO 00/18906, поданной Patten et al., "Shuffling of Codon-Altered Genes", WO 00/04190, поданной del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Recombination", WO 00/42561, поданной Crameri et al., "Oligonucleotide Mediated Nucleic Acid Recombination", WO 00/42559, поданной Selifonov и Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations", WO 00/42560, поданной Selifonov et al., "Methods for Making Character Strings, Polynucleotides & Polypeptides Having Desired Characteristics", WO 01/23401, поданной Welch et al., "Use of Codon-Varied Oligonucleotide Synthesis for Synthetic Shuffling" и WO 01/64864 "Single-Stranded Nucleic Acid Template-Mediated Recombination and Nucleic Acid Fragment Isolation", поданной Affholter.

Термины "вектор", "векторная конструкция" и "вектор экспрессии" означают носитель, с помощью которого последовательность ДНК или РНК (например, чужеродный ген) может быть встроена в клетку-хозяина так, чтобы трансформировать клетку-хозяина и активизировать экспрессию (например, транскрипцию и трансляцию) встроенной последовательности. Векторы обычно содержат ДНК трансмиссивного агента, в который чужеродную ДНК, кодирующую белок, встраивают с помощью технологии ферментов рестрикций. Обычным типом вектора является "плазмида", которая в основном представляет собой замкнутую молекулу двухспиральной ДНК, которая легко может принимать дополнительную (чужеродную) ДНК, и которую легко можно встраивать в подходящую клетку-хозяина. Большое число векторов, в том числе плазмид и грибных векторов, было описано для репликации и/или экспрессии в разнообразных эукариотических и прокариотических клетках-хозяевах. Неограничивающие примеры включают pKK плазмиды (Clonetech), pUC плазмиды, pET плазмиды (Novagen, Inc., Madison, Wis.), pRSET или pREP плазмиды (Invitrogen, San Diego, Calif.) или pMAL плазмиды (New England Biolabs, Beverly, Mass.) и многие соответствующие клетки-хозяева, используя способы, описанные или процитированные в настоящем описании или иным образом известные специалистам в релевантной области. Рекомбинантные клонирующие векторы зачастую будут включать одну или несколько систем репликации для клонирования или экспрессии, один или несколько маркеров для селекции в клетке-хозяине, например резистентность к антибиотикам, и одну или несколько экспрессионных кассет.

Термины "экспрессируют" и "экспрессия" означают предоставление возможности или осуществление реализации информации в гене или последовательности ДНК, например, продукцию белка путем активации клеточных функций, вовлеченных в транскрипцию и трансляцию соответствующего гена или последовательности ДНК. Последовательность ДНК экспрессируется в клетке или клеткой для образования "продукта экспрессии", такого как белок. Также можно сказать, что сам продукт экспрессии, например полученный в результате белок, "экспрессирован" клеткой. Полинуклеотид или полипептид экспрессируется рекомбинантно, например, когда он экспрессируется или продуцируется в чужеродной клетке-

хозяине под контролем чужеродного или нативного промотора или в нативной клетке-хозяине под контролем чужеродного промотора.

В одном варианте осуществления вирусный вектор может быть репликационно-компетентным ретровирусным вектором, способным инфицировать только реплицирующиеся клетки млекопитающих.

Ретровирусы были классифицированы различными путями, но номенклатура была стандартизована в последнее десятилетие (см. ICTVdB - The Universal Virus Database, v. 4 на World Wide Web www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/ и пособие "Retroviruses" Eds Coffin, Hughes and Varmus, Cold Spring Harbor Press 1997; описание которых включено в настоящий документ путем ссылки). Репликационно-компетентный ретровирусный вектор может включать ортогетровирусный или более характерно гамма-ретровирусный вектор. В одном аспекте репликационно-компетентный ретровирусный вектор содержит внутренний сайт связывания рибосомы (IRES) 5' с полинуклеотидом, кодирующим цитозинде-заминазу. В одном варианте осуществления полинуклеотид, кодирующий цитозинdezаминазу, находится 3' по отношению к ENV полинуклеотиду ретровирусного вектора.

Как рассмотрено ранее, общие сведения, в которых описываются методики молекулярной биологии, используемые в настоящем описании, в том числе использование векторов, промоторов и многие другие актуальные темы, включают Berger и Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, Volume 152 (Academic Press, Inc., San Diego, Calif.) ("Berger"); Sambrook et al., Molecular Cloning A Laboratory Manual, 2nd ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989 ("Sambrook") и Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, совместное предприятие Greene Publishing Associates, Inc. и John Wiley & Sons, Inc., (дополнено на протяжении 1999) ("Ausubel"). Примеры протоколов, достаточные, чтобы проинструктировать специалистов в данной области по *in vitro* методам амплификации, в том числе полимеразной цепной реакции (PCR), лигазной цепной реакции (LCR), Qβ-репликазной амплификации и другими методиками, опосредованных РНК полимеразой (например, NASBA), например, для продукции гомологичных нуклеиновых кислот по настоящему изобретению, можно найти у Berger, Sambrook и Ausubel, а также у Mullis et al. (1987) в патенте США № 4683202; Innis et al., eds. (1990), PCR Protocols: A Guide to Methods and Applications (Academic Press Inc. San Diego, Calif.) ("Innis"); Arnheim & Levinson (Oct. 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991), 3:81-94; Kwok et al. (1989), Proc. Natl. Acad. Sci. USA 86:1173; Guatelli et al. (1990), Proc. Nat'l. Acad. Sci. USA 87:1874; Lomell et al. (1989), J. Clin. Chem 35:1826; Landegren et al. (1988), Science, 241:1077-1080; Van Brunt (1990), Biotechnology, 8:291-294; Wu и Wallace (1989), Gene, 4:560; Barringer et al. (1990), Gene, 89:117 и Sooknanan and Malek (1995), Biotechnology, 13: 563-564. Усовершенствованные способы клонирования *in vitro* амплифицированных нуклеиновых кислот описаны у Wallace et al., в патенте США № 5426039. Усовершенствованные способы амплификации крупных нуклеиновых кислот с помощью ПЦР кратко изложены у Cheng et al. (1994), Nature, 369:684-685 и в ссылках, процитированных в этих публикациях, в которых получают ПЦР ампликоны вплоть до 40 тыс. пар оснований. Специалисту в данной области будет понятно, что, по существу, любая РНК может быть преобразована в двухспиральную ДНК, подходящую для рестрикционного расщепления, ПЦР удлинения и секвенирования с использованием обратной транскриптазы и полимеразы. См., например, Ausubel, Sambrook и Berger, все указано выше.

Также представлены сконструированные клетки-хозяева, которые трансдуцированы (трансформированы или трансфицированы) вектором, представленным в настоящем описании (например, клонирующим вектором или вектором экспрессии), а также получение полипептидов по настоящему изобретению с помощью рекомбинантных технологий. Вектором может быть, например, плазмида, вирусная частица, фаг и т.д. Сконструированные клетки-хозяева могут быть культивированы в обычной питательной среде, модифицированной соответствующим образом для активации промоторов, отбора трансформантов или амплификации кодирующего полинуклеотида. Условиями культивирования, такими как температура, pH и подобные, являются условия, использованные ранее с клетками-хозяевами, выбранными для экспрессии, и будут очевидны специалистам в данной области и в ссылках, процитированных в настоящем описании, в том числе, например, Sambrook, Ausubel и Berger, а также, например, Freshney (1994), Culture of Animal Cells: A Manual of Basic Technique, 3rd ed. (Wiley-Liss, New York) и процитированных там ссылках.

Векторы могут быть использованы для трансформации соответствующего хозяина, чтобы обеспечить экспрессию белка или полипептида клеткой-хозяином. Примеры соответствующих экспрессионных клеток-хозяев включают клетки бактерий, такие как *E.coli*, *B.subtilis*, *Streptomyces* и *Salmonella typhimurium*; клетки грибов, такие как *Saccharomyces cerevisiae*, *Pichia pastoris*, и *Neurospora crassa*; клетки насекомых, такие как *Drosophila* и *Spodoptera frugiperda*; клетки млекопитающих, такие как CHO, COS, BHK, HEK 293 или Bowes меланомы; или клетки растений или экспланта и т.д.

В бактериальных системах ряд векторов экспрессии может быть выбран в зависимости от использования, предназначенного для полипептида цитозинdezаминазы. Например, в тех случаях, когда большие количества полипептида СД или его фрагментов необходимы для коммерческой продукции или для индукции антител, могут быть желательными векторы, которые направляют высокий уровень экспрессии слитых белков, которые легко очистить. Такие векторы включают, но не только, многофункциональные

E.coli клонирующие векторы и векторы экспрессии, такие как BLUESCRIPT (Stratagene), в которых последовательность, кодирующая полипептид CD, может быть лигирована в вектор в рамке считывания с последовательностями для аминоконцевого Met и последующими 7 остатками бета-галактозидазы, так, чтобы продуцировался гибридный белок; pIN векторы (Van Heeke & Schuster (1989), J. Biol. Chem. 264:5503-5509); pET векторы (Novagen, Madison Wis.) и подобные.

Аналогично, у дрожжей *Saccharomyces cerevisiae* ряд векторов, содержащих конститтивные или индуцибельные промоторы, такие как альфа фактор, алкоголь-оксидаза и PGH, могут быть использованы для продукции полипептидов CD по настоящему изобретению. Обзоры см. у Ausubel (supra) и Grant et al. (1987), Methods in Enzymology, 153:516-544.

Настоящее изобретение также относится к репликационно-компетентным ретровирусным векторам, имеющим повышенную стабильность относительно предшествующих ретровирусных векторов. Такая повышенная стабильность во время инфицирования и репликации является важной для лечения нарушений клеточной пролиферации. Сочетание эффективности трансдукции, трансгенной стабильности и специфичности к мишени обеспечивается репликационно-компетентными ретровирусами. Композиции и способы обеспечивают стабильность вставки и поддерживают транскрипционную активность трансгена и трансляционную эффективность кодируемого полипептида.

Ретровирусы могут передаваться горизонтально и вертикально. Эффективная инфекционная трансмиссия ретровирусов требует экспрессии на клетке-мишени рецепторов, которые специфически распознают белки оболочки вируса, хотя вирусы могут использовать рецептор-независимые, неспецифические пути входления с низкой эффективностью. Кроме того, тип клетки-мишени должен быть способен поддерживать все стадии цикла репликации после связывания и пенетрации вируса. Вертикальная трансмиссия происходит, когда вирусный геном становится интегрированным в зародышевую линию хозяина. Затем провирус будет передаваться от поколения к поколению как будто клеточный ген. Следовательно, эндогенные провирусы укореняются, которые зачастую лежат латентно, но которые могут активироваться, когда хозяин подвергается воздействию соответствующих агентов.

Как указано выше, интегрированный посредник ДНК называется провирусом. В предшествующей генной терапии и системах доставки генов используются способы и ретровирусы, которые требуют транскрипции провируса и сборки в инфицирующий вирус, либо в присутствии соответствующего вируса-помощника, либо в клеточной линии, содержащий соответствующие последовательности, способствующие инкапсулированию без совпадающей продукции контаминирующего вируса-помощника. Как описано ниже, вирус-помощник не требуется для продукции рекомбинантного ретровируса по настоящему изобретению, поскольку последовательности для инкапсулирования обеспечиваются в геноме, таким образом предоставляя репликационно-компетентный ретровирусный вектор для доставки генов или генной терапии.

Ретровирусный геном, применимый в способах и композициях по настоящему изобретению, содержит провирусную ДНК, содержащую по меньшей мере три гена: gag, pol, и env, эти гены могут бытьflenkированы одним или двумя длинными концевыми повторами (LTR) или в провирусеflenkированы двумя длинными концевыми повторами (LTR) и последовательностями, содержащими действующие в цис-положении последовательности, такие как psi. Ген gag кодирует внутренние структурные белки (матрикс, капсид и нуклеокапсид); ген pol кодирует РНК-направленную ДНК полимеразу (обратную транскриптазу), протеазу и интегразу; и ген env кодирует гликопroteины вирусной оболочки. 5' и/или 3' LTR служат для активации транскрипции и полиаденилирования РНК вириона. LTR содержит все остальные последовательности, действующие в цис-положении, необходимые для репликации вируса. Лентивирусы имеют дополнительные гены, в том числе vif, vpr, tat, rev, vpu, nef, и upr (в HIV-1, HIV-2 и/или SIV).

Расположенные рядом с 5' LTR последовательности необходимы для обратной транскрипции генома (участок связывания праймера тРНК) и для эффективного инкапсулирования вирусной РНК в частицы (участок Psi). Если последовательности, необходимые для инкапсулирования (или упаковки ретровирусной РНК в инфекционный вирион), отсутствуют в вирусном геноме, результатом является цис-дефект, который предотвращает инкапсулирование геномной вирусной РНК. Этот тип модифицированного вектора представляет собой то, что обычно использовалось в предшествующих системах доставки гена (т.е. системы, лишенные элементов, которые необходимы для инкапсулирования вириона).

В одном варианте осуществления настоящее изобретение относится к рекомбинантному ретровирусу, способному инфицировать неделяющуюся клетку, делящуюся клетку или клетку, у которой имеется нарушение клеточной пролиферации.

Рекомбинантные репликационно-компетентные ретровирусы по настоящему изобретению содержат полинуклеотидную последовательность, кодирующую вирусный GAG, вирусный POL, вирусный ENV, гетерологичный полинуклеотид, с предшествующим внутренним сайтом связывания рибосомы (IRES), инкапсулированным в вирионе.

Гетерологичная последовательность нукleinовой кислоты оперативно связана с IRES. Используемый в настоящем описании термин "гетерологичная" последовательность нукleinовой кислоты или трансген относится к (i) к последовательности, которая обычно не существует в ретровирусах дикого

типа, (ii) к последовательности, которая происходит из чужеродных видов, или (iii) если из того же вида, она может быть, по существу, модифицирована от ее первоначальной формы. Альтернативно, неизмененная последовательность нуклеиновой кислоты, которая обычно не экспрессируется в клетке, является гетерологичной последовательностью нуклеиновой кислоты. В конкретном варианте осуществления гетерологичный полинуклеотид представляет собой полипептид по настоящему изобретению, обладающий цитозиндезаминазной активностью.

Выражение "неделящаяся" клетка относится к клетке, которая не проходит стадию митоза. Неделящиеся клетки могут быть блокированы в любой момент клеточного цикла (например, G₀/G₁, G_{1/S}, G_{2/M}) при условии, что эта клетка не является активно делящейся. Для инфицирования *ex vivo* делящаяся клетка может быть обработана для блокировки клеточного деления с использованием стандартных методик, используемых специалистами в данной области, в том числе излучением, обработкой афидоколином, истощением сыворотки и контактным ингибированием. Однако должно быть понятно, что инфицирование *ex vivo* зачастую проводят без блокировки клеток, поскольку многие клетки уже остановлены (например, стволовые клетки). Например, рекомбинантный лентивирусный вектор способен инфицировать любую неделяющуюся клетку, независимо от механизма, используемого для блокировки клеточного деления или момента в клеточном цикле, в который блокирована клетка. Примеры уже существующих неделяющихся клеток в организме включают нейрональные, мышечные клетки, клетки печени, кожи, сердца, легких и клетки костного мозга и их производные. Для делящихся клеток могут быть использованы онкоретровирусные векторы.

Под "делящейся" клеткой подразумевается клетка, которая подвергается активному митозу или мейозу. Такие делящиеся клетки включают стволовые клетки, клетки кожи (например, фибробласты и кератиноциты), гаметы и другие делящиеся клетки, известные из уровня техники. Представляющими особый интерес и охватываемые этим термином делящиеся клетки являются клетки, имеющие нарушения клеточной пролиферации, такие как неопластические клетки. Термин "нарушение клеточной пролиферации" относится к состоянию, характеризующемуся аномальным числом клеток. Это состояние может включать как гипертрофический (непрерывное умножение клеток, в результате приводящее к разрастанию клеточной популяции в пределах ткани), так и гипотрофический (отсутствие или недостаточность клеток в ткани) клеточный рост или чрезмерный приток или миграцию клеток в область организма. Клеточные популяции не являются обязательно трансформированными, образующими опухоль или злокачественными клетками, а могут включать нормальные клетки тоже. Нарушения клеточной пролиферации включают нарушения, связанные с разрастанием соединительных тканей, например, различные фиброзные состояния, в том числе, склеродермию, артрит и цирроз печени. Нарушения клеточной пролиферации включают неопластические нарушения, такие как карциномы тканей головы и шеи. Карциномы тканей головы и шеи включали бы, например, карциному ротовой полости, пищевода, горла, гортани, щитовидной железы, языка, губ, слюнных желез, носа, околоносовых пазух, носоглотки, верхнего свода носа и опухоли пазухи, обонятельную эстезионейроэпителиому, сквамозно-клеточный рак, злокачественную меланому, синоназальную недифференцированную карциному (SNUC), неоплазию головного мозга (в том числе глиобластомы) или крови. Также включены карциномы региональных лимфатических узлов, в том числе шейных лимфатических узлов, предгортанных лимфатических узлов, легочных юкстапищеводных лимфатических узлов и поднижечелюстных лимфатических узлов (Harrison's Principles of Internal Medicine (eds., Isselbacher, et al., McGraw-Hill, Inc., 13th Edition, p.1850-1853, 1994). Другие типы злокачественных опухолей включают, но этим не ограничиваются, рак легких, рак толстой и прямой кишки, рак груди, рак предстательной железы, рак мочевыводящих путей, рак матки, лимфому, рак ротовой полости, рак поджелудочной железы, лейкоз, меланому, рак желудка, рак кожи и рак яичников.

В одном варианте осуществления гетерологичный полинуклеотид в векторе содержит цитозиндезаминазу, которая была оптимизирована для экспрессии в клетках человека. В дополнительном варианте осуществления цитозиндезаминаза содержит последовательность, которая была оптимизирована в отношении кодонов человека, и содержит мутации, которые повышают стабильность цитозиндезаминазы (например, уменьшенное разрушение или повышенную термостабильность) по сравнению с цитозиндезаминазой дикого типа. Еще в одном варианте осуществления гетерологичный полинуклеотид кодирует слияную конструкцию, содержащую цитозиндезаминазу (либо кодон-оптимизированную для экспрессии в клетках человека, либо неоптимизированную, либо мутированную, либо немутированную), оперативно связанную с полинуклеотидом, кодирующим полипептид, обладающий UPRT или OPRT активностью. В другом варианте осуществления гетерологичный полинуклеотид содержит СД полинуклеотид или слияную конструкцию по настоящему изобретению (например, SEQ ID NO: 3, 5, 11, 13, 15 или 17).

В другом варианте осуществления репликационно-компетентный ретровирусный вектор может содержать гетерологичный полинуклеотид, кодирующий полипептид, содержащий цитозиндезаминазу (как описано выше), и дополнительно может содержать полинуклеотид, содержащий молекулу mi-RНК или si-RНК, связанную с промотором, специфичным к клеточному типу или к типу ткани.

Термин "регуляторная последовательность нукleinовой кислоты" относится обобщенно к промоторным последовательностям, сигналам полиаденилирования, последовательностям терминации транскрипции, регуляторным доменам, расположенным против хода транскрипции, точкам начала реплика-

ции, энхансерам и подобному, которые сообща обеспечивают репликацию, транскрипцию и трансляцию кодирующей последовательности в клетке-реципиенте. Не всегда нужно присутствие всех этих контрольных последовательностей, при условии, что выбранная кодирующая последовательность способна реплицироваться, транскрибироваться и транслироваться в соответствующей клетке-хозяине. Специалист в данной области без труда может идентифицировать регуляторную последовательность нуклеиновой кислоты из общедоступных баз данных и материалов. Более того, специалист в данной области может идентифицировать регуляторную последовательность, которая применима для предназначенного использования, например, *in vivo*, *ex vivo* или *in vitro*.

Термин "промоторная область" используется в настоящем описании в его первоначальном смысле, относящемся к нуклеотидной области, содержащей регуляторную последовательность ДНК, где регуляторная последовательность получена из гена, который способен связываться с РНК полимеразой и инициировать транскрипцию кодирующей последовательности, расположенной по ходу транскрипции (в 3'-направлении). Регуляторная последовательность может быть гомологичной или гетерологичной желаемой генной последовательности. Например, может быть использован широкий ряд промоторов, в том числе вирусный промотор или промотор млекопитающих, как описано выше.

Внутренние сайты связывания рибосомы ("IRES") относятся к сегменту нуклеиновой кислоты, которые способствуют связыванию или удерживанию рибосомы во время трансляции кодирующей последовательности обычно в положении 3' по отношению IRES. В некоторых вариантах осуществления IRES может содержать акцепторный сайт сплайсинга/донорный сайт, однако предпочтительные IRES лишены акцепторного сайта сплайсинга/донорного сайта. Обычно связывание рибосом с матричной РНК имеет место через кэп, расположенный на 5'-конце всех эукариотических мРНК. Однако существуют исключения из этого универсального правила. Отсутствие кэпа в некоторых вирусных мРНК предполагает существование альтернативных структур, позволяющих осуществлять связывание рибосом на внутреннем сайте этих РНК. На сегодняшний день ряд таких структур, обозначенных IRES на основании их функции, был идентифицирован в 5' некодирующей области некэпированных вирусных мРНК, например у пикоронавирусов, например вируса полиомиелита (Pelletier et al., 1988, Mol. Cell. Biol., 8, 1103-1112) и вируса EMCV (вирус энцефаломиокардита (Jang et al., J. Virol., 1988, 62, 2636-2643). Настоящее изобретение относится к использованию IRES применительно к репликативно-компетентному ретровирусному вектору.

Гетерологичная последовательность нуклеиновой кислоты обычно находится под контролем либо вирусных сигналов LTR промотор-энхансер, либо внутреннего промотора, и сигналов удерживания в ретровирусном LTR может все еще приводить к эффективной интеграции вектора в геном клетки-хозяина. В соответствии с этим рекомбинантные ретровирусные векторы по настоящему изобретению, желаемые последовательности, гены и/или генные фрагменты могут быть встроены в нескольких сайтах и под контролем различных регуляторных последовательностей. Например, сайтом для встраивания может быть вирусный проксимальный сайт энхансер/промотор (т.е. 5' LTR-контролируемый генний локус). Альтернативно, желаемые последовательности могут быть встроены в дистальный сайт (например, IRES последовательность 3' к гену env) или где присутствуют две или более гетерологичные последовательности, одна гетерологичная последовательность может быть под контролем первой регуляторной области, а вторая гетерологичная последовательность может быть под контролем второй регуляторной области. Другие дистальные сайты включают вирусные промоторные последовательности, где экспрессия желаемой последовательности или последовательностей происходит посредством сплайсинга цистрона проксимального промотора, может быть использован внутренний гетерологичный промотор, такой как SV40 или CMV или внутренний сайт связывания рибосомы (IRES).

В одном варианте осуществления ретровирусный геном по настоящему изобретению содержит IRES, содержащий клонирующий сайт для встраивания желаемой полинуклеотидной последовательности. В одном варианте осуществления IRES расположен 3' по отношению к гену env в ретровирусном векторе, но 5' по отношению к желаемой гетерологичной нуклеиновой кислоте. В соответствии с этим гетерологичная полинуклеотидная последовательность, кодирующая желаемый полипептид, может быть оперативно связана с IRES.

В другом варианте осуществления направляющая полинуклеотидная последовательность включена как часть рекомбинантного ретровирусного вектора по настоящему изобретению. Направляющая полинуклеотидная последовательность представляет собой направленный лиганд (например, пептидные гормоны, такие как герегулин, одноцепочечные антитела, рецептор или лиганд для рецептора), тканеспецифичный регуляторный элемент или регуляторный элемент, специфичный к клеточному типу (например, тканеспецифичный промотор или энхансер, или промотор или энхансер, специфичный к клеточному типу) или сочетание направленного лиганда и тканеспецифичного/специфичного к клеточному типу регуляторного элемента. Предпочтительно направленный лиганд оперативно связан с белком env ретровируса, создавая химерный ретровирусный белок env. Вирусные белки GAG, POL и ENV могут быть получены из любых подходящих ретровирусов (например, MLV или производных лентивирусов). В другом варианте осуществления вирусный белок ENV не является производным ретровируса (например, CMV или VSV).

Рекомбинантный ретровирус по настоящему изобретению, следовательно, является генетически модифицированным, таким образом, чтобы этот вирус был нацелен на конкретный тип клеток (например, гладкомышечные клетки, клетки печени, клетки почек, фибробласты, кератиноциты, мезенхимальные стволовые клетки, клетки костного мозга, хондроциты, эпителиальные клетки, интестинальные клетки, неопластические клетки, клетки глиомы, нейрональные клетки и другие, известные из уровня техники), так, чтобы полинуклеотидный геном доставлялся в неделяющуюся клетку-мишень, делящуюся клетку-мишень или клетку-мишень, имеющую нарушение клеточной пролиферации. Нацеливание может достигаться двумя путями. Первый путь направляет ретровирус к клетке-мишени путем связывания с клетками, имеющими молекулу на внешней поверхности клетки. В этом способе нацеливания ретровируса используется экспрессия направленного лиганда на оболочке ретровируса, чтобы облегчить нацеливание вируса на клетки или ткани, которые имеют рецептор или молекулу связывания, которая взаимодействует с направленным лигандом на поверхности ретровируса. После инфицирования клетки вирусом вирус вводит свою нуклеиновую кислоту в клетку и ретровирусный генетический материал может интегрироваться в геном клетки-хозяина. Во втором способе нацеливания используются регуляторные элементы, специфичные к клеткам или тканям для активации экспрессии и транскрипции вирусного генома в клетке-мишени, которая активно использует регуляторные элементы, как описано ниже полно. Перенесенный ретровирусный генетический материал затем транскрибируется и транслируется в белки в клетке-хозяине. Направляющий регуляторный элемент обычно связан с 5' и/или 3'LTR, создавая химерный LTR.

Путем встраивания гетерологической последовательности нуклеиновой кислоты, представляющей интерес, в вирусный вектор по настоящему изобретению, наряду с другим геном, который кодирует, например, лиганд для рецептора на специфической клетке-мишени, вектор теперь является специфичным к мишени. Вирусные векторы могут быть получены специальными к мишени путем присоединения, например, сахара, гликолипида или белка. Нацеливание может осуществляться путем использования антитела для нацеливания на вирусный вектор. Специалистам в данной области будет известно или они без труда могут определить специфические полинуклеотидные последовательности, которые могут быть встроены в вирусный геном или белки, которые могут быть присоединены к оболочке вируса, чтобы позволить осуществлять нацеленную специфическую доставку вирусного вектора, содержащего последовательность нуклеиновой кислоты, представляющей интерес.

Ретровирус и способы по настоящему изобретению предоставляют репликационно-компетентный ретровирус, для которого не требуется вирус-помощник, или дополнительная последовательность нуклеиновой кислоты, или белки для размножения и продукции вириона. Например, последовательности нуклеиновой кислоты ретровируса по настоящему изобретению кодируют, например, группаспецифический антиген и обратную транскриптазу (и интегразу, и протеазные ферменты, необходимые для созревания и обратной транскрипции) соответственно, как описано выше. Вирусные gag и pol могут быть получены из лентивируса, такого как HIV или онковирус, такой как MoMLV. Кроме того, полинуклеотидный геном ретровируса по настоящему изобретению включает последовательность, кодирующую белок оболочки вируса (ENV). Ген env может быть получен из любых ретровирусов. Белок env может представлять собой амфотропный оболочечный белок, которые позволяют осуществлять трансдукцию клеток человека или других видов или может представлять собой эктотрофный оболочечный белок, который способен трансдуцировать любые клетки мыши и крысы. Далее, может быть желательным нацелить рекомбинантный вирус путем связывания оболочечного белка с антителом или конкретным лигандом для нацеливания на рецептор конкретного клеточного типа. Как указано выше, ретровирусные векторы могут быть получены специальным к мишени путем встраивания, например, гликолипида или белка. Нацеливание зачастую осуществляют путем использования антитела для нацеливания ретровирусного вектора на антиген на конкретном клеточном типе (например, клеточный тип, обнаруженный в определенной ткани или типе злокачественных клеток). Специалистам в данной области будут известны или могут быть легко установлены, без проведения лишних экспериментов, специфические способы для достижения доставки ретровирусного вектора к специальной мишени. В одном варианте осуществления ген env доставляется не из ретровируса (например, CMV или VSV). Примеры ретровирусных env генов включают, но этим не ограничиваются, вирус мышиного лейкоза Молони (MoMuLV), вирус саркомы мышей Харви (HaMuSV), вирус опухоли молочной железы мышей (MuMTV), вирус лейкоза гибонов (GaLV), вирус иммунодефицита человека (HIV) и вирус саркомы Рауса (RSV). Также могут быть использованы другие env гены, например, вируса везикулярного стоматита (VSV) (белок G), оболочки цитомегаловируса (CMV) или гемагглютинин вируса гриппа (HA).

В одном варианте осуществления ретровирусный геном является производным онкоретровируса или гамма-ретровируса, более конкретно, онкоретровируса или гамма-ретровируса млекопитающих. Под "производным" подразумевается, что родительская полинуклеотидная последовательность представляет собой онковирус дикого типа, который был модифицирован путем встраивания или удаления природных последовательностей (например, вставки IRES, вставки гетерологического полинуклеотида, кодирующего, например, полипептид, обладающий цитозинdezаминаизной активностью по настоящему изобретению и подобного).

В другом варианте осуществления настоящего изобретение относится к ретровирусным векторам, которые направляются с использованием регуляторных последовательностей. Клеточно-специфичные или тканеспецифичные регуляторные последовательности (например, промоторы) могут быть использованы для направления экспрессии генных последовательностей в конкретных популяциях клеток. Подходящие промоторы млекопитающих и вирусные промоторы для настоящего изобретения описаны в других местах настоящего документа. В соответствии с этим в одном варианте осуществления настоящего изобретение относится к ретровирусу, имеющему тканеспецифичные промоторные элементы на 5'-конце ретровирусного генома. Предпочтительно тканеспецифичные регуляторные элементы/последовательности находятся в области U3 LTR ретровирусного генома, включая, например, клеточно- или тканеспецифичные промоторы и энхансеры для неопластических клеток (например, энхансеры и промоторы, специфичные для опухолевых клеток) и индуцибельные промоторы (например, тетракиклин).

В некоторых обстоятельствах может быть желательным регулировать экспрессию. Например, различные вирусные промоторы с различной силой активности могут быть использованы в зависимости от желаемого уровня экспрессии. В клетках млекопитающих ранний промотор CMV зачастую используют для обеспечения сильной транскрипционной активации.

Модифицированные варианты промотора CMV, которые являются менее сильными, также были использованы, когда требовались сниженные уровни экспрессии трансгена. В тех случаях, когда желательна экспрессия трансгена в гематopoэтических клетках, могут быть использованы ретровирусные промоторы, такие как LTR из MLV или MMTV. Другие вирусные промоторы, которые могут быть использованы, включают SV40, RSV LTR, HIV-1 и HIV-2 LTR, адено-вирусные промоторы, например, из области E1A, E2A или MLP, AAV LTR, вируса мозаики цветной капусты, HSV-TK, и вируса саркомы птиц.

Аналогично, тканеспецифичные или селективные промоторы могут быть использованы для осуществления транскрипции в специфических тканях или клетках, чтобы снизить возможную токсичность или нежелательные эффекты на ткани, не являющиеся мишениями. Например, промоторы, такие как PSA, пробазин, простатическая кислая фосфатаза или простат-специфический глангулярный калликреин (hK2), могут быть использованы для направления генной экспрессии в предстательной железе. Другие промоторы/регуляторные домены, которые могут быть использованы, изложены в табл. 1.

При некоторых показаниях может быть желательным активировать транскрипцию в конкретное время после введения вектора для генной терапии. Это может быть сделано с такими промоторами, которые регулируются гормонами или цитокинами. Например, в терапевтических применениях, при которых показанием является ткань половых желез, в которой продуцируются или распространяются специфические стероидные гормоны, предпочтительным может быть использование андроген- или эстроген-регулируемых промоторов. Такие промоторы, которые являются гормонально регулируемыми, включают MMTV, MT-1, экдизон и RuBisco. Могут быть использованы другие гормон-регулируемые промоторы, например, чувствительные к гормонам щитовидной железы, гипофиза и надпочечников. Промоторы, чувствительные к цитокинам и белкам воспаления, которые могли быть использованы, включают кининоген K и T (Kageyama et al., 1987), c-fos, TNF-альфа, С-реактивный белок (Arcone et al., 1988), гаптоглобин (Oliviero et al., 1987), сывороточный амилоид A2, C/EBP альфа, IL-1, IL-6 (Poli и Cortese, 1989), комплемент C3 (Wilson et al., 1990), IL-8, альфа-1 кислый гликопротеин (Prowse и Baumann, 1988), альфа-1 антитрипсин, липопротеинлипазу (Zechner et al., 1988), ангиотензиноген (Ron et al., 1990), фибриноген, c-jun (индивидуируемый сложными эфирами форбала, TNF-альфа, УФ-излучением, ретиноевой кислотой и перекисью водорода), коллагеназу (индивидуируемую сложными эфирами форбала и ретиноевой кислотой), металлотионеин (индивидуируемый тяжелым металлом и глюкокортикоидом), стромелизин (индивидуируемый сложным эфиром форбала, интерлейкином-1 и EGF), альфа-2 макроглобулин и альфа-1 антихимотрипсин. Опухоль-специфичные промоторы, такие как остеокальцин, фактор, чувствительный к гипоксии (HRE), MAGE-4, CEA, альфа-фетопротеин, GRP78/BiP и тирозиназа, также могут быть использованы для регуляции генной экспрессии в опухолевых клетках.

Кроме того, этот перечень промоторов не следует рассматривать как исчерпывающий или ограничивающий, специалистам в данной области будут известны другие промоторы, которые могут быть использованы в сочетании с промоторами и способами, описанными в настоящем документе.

Таблица 2

Тканеспецифичные промоторы

Ткань	Промотор
Поджелудочная железа	Инсулинэластинамилаза pdr-1 pdx-1 глюкокиназа
Печень	Альбумин РЕРСК HBV энхансер α фетопротеин аполипопротеин С α -1 антитрипсин вителлогенин, NF-AB транстиреин
Скелетные мышцы	Н цепь миозина мышечная креатинкиназа дистрофин калпаин p94 скелетный альфа-актин быстрый тропонин 1
Кожа	Кератин K6 Кератин K1
Легкие	CFTR цитокератин человека 18 (K18) белки сурфактанта легких А, В и С СС-10 Р1
Гладкая мускулатура	sm22 α SM-альфа-актин
Эндотелий	Эндотелин-1 Е-селектин фактор фон Виллебранда TIE, KDR/flk-1 тирозиназа меланоцитов
Жировая ткань	липопротеинлипаза (Zechner <i>et al.</i> , 1988) адипсин (Spiegelman <i>et al.</i> , 1989) ацетил-СоА карбоксилаза (Pape и Kim, 1989) глицерофосфатдегидрогеназа (Dani <i>et al.</i> , 1989) адipoцит Р2 (Hunt <i>et al.</i> , 1986)
Молочная железа	кислый белок молочной сыворотки (WAP) (Andres <i>et al.</i> . PNAS 84:1299-1303 1987
Кровь	β -глобин

"Тканеспецифичные регуляторные элементы" представляют собой регуляторные элементы (например, промоторы), которые способны направлять транскрипцию гена в одной ткани, оставаясь в основном "молчачими" в тканях других типов. Однако будет понятно, что тканеспецифичные промоторы могут иметь определяемое количество "фоновой" или "базовой" активности в тех тканях, в которых они являются молчачими. Степень, в которой промотор селективно активируется в ткани-мишени, может быть выражена как индекс селективности (активность в ткани-мишени/активность в контрольной ткани). В этом отношении тканеспецифичный промотор, используемый при осуществлении настоящего изобретения, обычно имеет индекс селективности примерно больше 5. Предпочтительно индекс селективности составляет примерно больше 15.

Далее будет понятно, что некоторые промоторы, хотя не ограниченные в активности в отношении одного типа ткани, могут, тем не менее, демонстрировать селективность в том, что они могут быть активными в одной группе тканей и менее активными или молчачими в другой группе. Такие промоторы также называют "тканеспецифичными", и их использование предусматривается настоящим изобретением. В соответствии с этим тканеспецифичные регуляторные элементы, используемые в настоящем изобретении, применимы для регуляции гетерологичных белков, таких как полипептиды, обладающие цито-зинdezаминаизной активностью по настоящему изобретению, а также применимые в качестве направляющей полинуклеотидной последовательности в ретровирусных векторах.

Ретровирусные векторы и полинуклеотиды СД и полипептиды по настоящему изобретению могут быть использованы для лечения широкого ряда заболеваний и нарушений, включающих ряд клеточно-пролиферативных заболеваний и нарушений (см., например, патенты США № 4405712 и 4650764; Friedmann, 1989, Science, 244:1275-1281; Mulligan, 1993, Science, 260:926-932, R. Crystal, 1995, Science 270:404-410, каждый из которых включен в настоящее описание посредством ссылки в полном объеме, см. также публикацию The Development of Human Gene Therapy, Theodore Friedmann, Ed., Cold Spring

Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1999. ISBN 0-87969-528-5, которая включена в настоящее описание путем ссылки в полном объеме).

Таким образом, настоящее изобретение относится к рекомбинантному ретровирусу, способному инфицировать неделяющиеся клетки, делящиеся клетки или неопластические клетки, где рекомбинантный ретровирус содержит вирусный GAG; вирусный POL; вирусный ENV; гетерологичную нуклеиновую кислоту (например, содержащему полипептид по настоящему изобретению, обладающий цитозинdezаминазной активностью), оперативно связанную с IRES; и действующие в цис-положении последовательности нуклеиновой кислоты, необходимые для упаковки, обратной транскрипции и интеграции. Рекомбинантный ретровирус может представлять собой лентивирус, такой как HIV, или может представлять собой онковирус. Как описано выше для способа получения рекомбинантного ретровируса, рекомбинантный ретровирус по настоящему изобретению дополнительно может содержать по меньшей мере один из белков VPR, VIF, NEF, VPX, TAT, REV и VPU. Не желая связываться конкретной теорией, считается, что один или несколько из этих генных/белковых продуктов являются важными для повышения титра вируса продуцированного рекомбинантного ретровируса (например, NEF) или могут быть необходимыми для инфицирования и упаковки вириона.

Например, промотор может содержать CMV промотор, имеющий последовательность SEQ ID NO: 19, 20 или 22 от нуклеотида 1 до примерно нуклеотида 582, и может включать модификацию одного или нескольких (например, 2-5, 5-10, 10-20, 20-30, 30-50, 50-100 или более) оснований нуклеиновой кислоты при условии, что модифицированный промотор способен направлять и инициировать транскрипцию. В одном варианте осуществления промотор или регуляторная область содержит полинуклеотидный домен CMV-R-U5. Домен CMV-R-U5 содержит ранний промотор цитомегаловируса человека для области MLV R-U5. В одном варианте осуществления полинуклеотидный домен CMV-R-U5 содержит последовательность SEQ ID NO: 19, 20 или 22 примерно от нуклеотида 1 до примерно нуклеотида 1202 или последовательность, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 19, 20 или 22, где указанный полинуклеотид активизирует транскрипцию молекулы нукleinовой кислоты, оперативно связанной с ним. Домен gag этого полинуклеотида может быть получен из любого числа ретровирусов, но обычно его получают из онкоретровируса и, более конкретно, из онкоретровируса млекопитающих. В одном варианте осуществления домен gag содержит последовательность примерно от нуклеотида номер 1203 до примерно нуклеотида 2819 или последовательность, по меньшей мере на 95, 98, 99 или 99,8% (округленно до ближайшего 10-го) идентичную ей. Домен pol этого полинуклеотида может быть получен из любого числа ретровирусов, но обычно его получают из онкоретровируса и, более конкретно, из онкоретровируса млекопитающих. В одном варианте осуществления домен pol содержит последовательность примерно от нуклеотида номер 2820 до примерно нуклеотида 6358 или последовательность, по меньшей мере на 95, 98, 99 или 99,9% (округленно до ближайшего 10-го) идентичную ей. Домен env этого полинуклеотида может быть получен из любого числа ретровирусов, но обычно его получают из онкоретровируса и, более конкретно, из онкоретровируса млекопитающих. В некоторых вариантах осуществления кодирующий домен env содержит амфотропный домен env. В одном варианте осуществления домен env содержит последовательность примерно от нуклеотида номер 6359 до примерно нуклеотида 8323 или последовательность, по меньшей мере на 95, 98, 99 или 99,8% (округленно до ближайшего 10-го) идентичную ей. Домен IRES домен этого полинуклеотида может быть получен из любого числа внутренних сайтов связывания рибосомы. В одном варианте осуществления IRES получен из вируса энцефаломиокардита. В одном варианте осуществления домен IRES содержит последовательность примерно от нуклеотида номер 8327 до примерно нуклеотида 8876 или последовательность, по меньшей мере на 95, 98 или 99% (округленно до ближайшего 10-го) идентичную ей, при условии, что этот домен позволяет осуществлять связывание рибосомы. Гетерологичный домен может содержать цитозинdezаминазу по настоящему изобретению. В одном варианте осуществления полинуклеотид CD содержит последовательность кодонов, оптимизированных для экспрессии в клетках человека. Еще в одном варианте осуществления полинуклеотид CD кодирует мутантный полипептид, обладающий цитозинdezаминазной активностью, которому мутации придают повышенную термостабильность, повышающей температуру плавления (T_m) на 10°C, позволяя задерживать кинетическую активность в широком температурном интервале и повышать накопленные уровни белка. В одном варианте осуществления цитозинdezаминаза содержит последовательность SEQ ID NO: 19 примерно от нуклеотида номер 8877 до примерно 9353. За гетерологичным доменом может следовать домен, обогащенный пурином. 3'LTR может быть получен из любого числа ретровирусов, обычно онкоретровируса и предпочтительно онкоретровируса млекопитающих. В одном варианте осуществления 3'LTR содержит домен U3-R-U5. Еще в одном варианте осуществления LTR содержит последовательность SEQ ID NO: 19 примерно от нуклеотида 9405 примерно до 9998 или последовательность, которая по меньшей мере на 95, 98 или 99,5% (округленно до ближайшего 10-го) идентичную ей.

Настоящее изобретение также относится к рекомбинантному ретровирусному вектору, содержащему от 5' к 3' CMV-R-U5, слияние раннего промотора цитомегаловируса человека с областью MLV R-U5; PBS, сайт связывания праймера с обратной транскриптазой; 5' сайт сплайсинга; сигнал упаковки Ψ ; gag, ORF для MLV группоспецифичного антигена; pol, ORF для полипротеина MLV полимеразы; 3' сайт

сплайсинга; 4070A env, ORF для белка оболочки MLV штамма 4070A; IRES, внутренний сайт связывания рибосомы цитомегаловируса; модифицированную цитозинdezаминазу (термостабильную с оптимизированными кодонами); PPT, полипуриновый тракт; и U3-R-U5, MLV длинный концевой повтор. Эта структура дополнительно представлена на фиг. 1.

Настоящее изобретение также относится к ретровирусному вектору, содержащему последовательность SEQ ID NO: 19, 20 или 22.

Для примера, а не для ограничения, ретровирусный вектор, используемый при лечении нарушения клеточной пролиферации, будет включать амфотропный белок ENV, GAG и POL белки, промоторную последовательность в области U3 ретровирусного генома и любую последовательность, действующую в цис-положении, необходимую для репликации, упаковки и интеграции ретровирусного генома в клетку-мишень.

Следующие примеры предназначены для иллюстрации, а не для ограничения этого изобретения. Хотя такие примеры являются характерными из тех, которые могли быть использованы, другие процедуры, известные специалистам в данной области, могут быть использованы альтернативно.

Примеры

Пример 1. Конструирование модифицированных генов CD и встраивание в плазмидные векторы.

Была проведена генетическая оптимизация гена дикого типа цитозинdezаминазы дрожжей, чтобы он включал (1) три позиционные мутации, которые меняют три аминокислоты (A23L, I140L и V108I) для повышения термостабильности белка цитозинdezаминазы дрожжей; и (2) дополнительные модификации генной последовательности для повышения частоты использования кодонов человека для повышения эффективности трансляции белка в клетках человека без дополнительных изменений аминокислотной последовательности.

Конструкция последовательности для CD, включала CD оптимизированную, CD-UPRT (+/- линкер) и CD-OPRT-азу (+/-линкер). Конечная последовательность, кодирующая цитозинdezаминазу, может содержать на 5'-конце сайт PSI1 (полноразмерный) и на 3'-конце сайт Not1 плюс поли А хвост для методики на основе кассеты PSI1/Not1.

Следующая последовательность, содержащая цитозинdezаминазу дрожжей, была использована для клонирования, оптимизации и мутации (нуклеиновые кислоты в рамке содержат сайты рестрикции, используемые в последующих способах клонирования):

```

AACACGA[TTATAA]ATGGTGACAGGGGAATGGCAAGCAAGTGGATCAGAAGGGTATGGACAT
TGCCTATGAGGAGGCGCCCTAGGTTACAAAGAGGGTGGTGTCCATTGGCGGATGTCTTAT
CAATAACAAAGACGGAAGTGTCTCGGTCGTGGTCACAACATGAGATTCAAAAGGGATCCGC
CACACTACATGGTGAGATCTCCACTTGGAAAAGTGTGGGAGATTAGAGGGCAAAGTGTACAA
AGATACCACTTGTATACGACGCTGTCTCCATGCGACATGTGTACAGGTGCCATCATCATGTA
TGGTATTCCACGCTGTGTTGCGTGAGAACGTTAATTCAAAAGTAAGGGCGAGAAATATT
ACAAACTAGAGGTACCGAGGTTGTTGACGATGAGAGGTGTAAAAAGATCATGAAACA
ATTTATCGATGAAAGACCTCAGGATTGGTTGAAGATATTGGTGAGTAG[GCGGCCGC]GCCATA
GATAAAATAAAAGATTATTTAGTCTCCAGAAAAAGGGGG (SEQ ID NO: 31)

```

В табл. кратко изложены гены и полученные в результате плазмидные векторы, которые были получены, и их названия.

Таблица 3

Векторные конструкции и названия

Tocagen Код	Наименование	Первоначальное название	5'LTR пром	оболочка	Вектор	IRES	Трансген	3'LTR
T5.0000	pACE-yCD	pACE-CD (Tai et al. 2005)	CMV	Ampho (4070A)	pACE	EMCV	CD дрожжей дикого типа	MLV U3
T5.0001	pAC3-yCD1	CDopt последовательность	CMV	Ampho (4070A)	pAC3	EMCV	Модифицированная CD	MLV U3
T5.0002	pAC3-yCD2	CDopt+3pt	CMV	Ampho (4070A)	pAC3	EMCV	Модифицированная CD	MLV U3
T5.0003	pAC3-yCD2-U	CDopt+3pt-UPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-UPRT	MLV U3
T5.0004	pAC3-yCD2-O	CDopt+3pt-OPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-OPRT	MLV U3
T5.0005	pAC3-yCD2-LO	CDopt+3pt-LINK-OPRT	CMV	Ampho (4070A)	pAC3	EMCV	CD2-L-OPRT	MLV U3
T5.0006	pAC3-eGFP	pAC3-emd, pAC3GFP	CMV	Ampho (4070A)	pAC3	EMCV	Изумрудный GFP	MLV U3
T5.0007	pAC3-yCD	pAC3-yCD	CMV	Ampho (4070A)	pAC3	EMCV	CD дрожжей дикого типа	MLV U3

Репликационно-компетентный ретровирусный вектор, описанный Kasahara et al. pACE-CD (патент США № 6899871, описание которого включено в настоящий документ), был использован в качестве основы для дополнительных модификаций. Вектор (pAC3-yCD) был модифицирован для экспрессии модифицированного гена цитозинdezаминазы дрожжей, как описано в настоящей заявке, и был использован в этих конструкциях. См. 1А ниже для диаграммы векторной конструкции для исходного трансфицированного репликационно-компетентного ретровируса. CMV представляет собой ранний промотор CMV человека, U3, R и U5 представляют собой соответствующие области вирусного длинного концевого повтора (LTR). Gag, pol и env представляют собой области, кодирующие вирусные белки. 1B и 1D демонстрируют плазмидную структуру и последовательность по настоящему изобретению. После синтеза генов в контракторе (Bio Basic Inc., Markham, Ontario, Canada) их встраивали в сайт Psi1-Not1 основы вектора pAC3 (фиг. 1). Основу плазмиды обычно получают путем разрезания плазмиды pAC3-eGFP с использованием Psi1 и Not1 и очистки крупного (около 11 тыс. пар оснований) фрагмента, кодирующего плазмиду и основу ретровируса).

А. Гуманизированный кодон-оптимизированный ген CD (CDopt, aka CD1, T5.0001).

Сравнение цитозинdezаминазы с кодонами, оптимизированными для экспрессии в клетках человека по Conrad et al. и PCT WO 99/60008 указывает на 91 суммарный кодон, оптимизированный у обоих, 36 кодонов идентичных, 47 кодонов имели изменения в третьей паре оснований (все кодируют одну и ту же аминокислоту) и 9 кодонов были различными (однако они кодировали одну и ту же аминокислоту). Из 9 кодонов, которые были различными:

AGC (Ser) на TCC (Ser)

CGT (Arg) на AGG (Arg)

CCA (Pro) на CCT (Pro)

все имели эквивалентное содержание GC и кодируют одну и ту же аминокислоту.

Нативная последовательность гена дрожжей была отдельно кодон-оптимизирована для получения следующего гена CD (CD1) и была названа T5.0001 при встраивании в плазмидный вектор pAC3, который кодирует репликационно-компетентный ретровирус (RCR) с IRES.

```
TTATAATGGTGACCCGGCGCATGGCTCCAAGTGGATCAAAAGGGATGGATATCGCTTAC
GAGGAGGCCCTGGCTACAAGGAGGGCGGTGCCTATCGGCGGCTGTCTGATCAACAAC
AAGGACGGCAGTGTGGCTGGCAGGGCCACAACATGAGGTTCCAGAAGGGCTCCGCCACCCTG
CACGGCGAGATCTCCACCCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAGGACACC
ACCCCTGTACACCAACCCCTGTCCCCCTGTGACATGTGTACCGGCCTATCATCATGTACGGCATT
CCTAGGTGTGGTGGCGAGAACGTGAACCTCAAGTCAAGGGCGAGAAGTACCTGCAAACC
AGGGGCCACGAGGTGGTGGTTGTGACGATGAGAGGTGTAAGAAGATCATGAAGCAGTTCATC
GACGAGAGGCCTCAGGACTGGTCAGGATATCGGCGAGTGATAAGCGGGCCAGATAAAAATA
```

AAAGATTTATTTAGTCTCCAGAAAAGGGGGG. (SEQ ID NO:32)

Б. Термостабилизированный ген CD.

Были проведены дополнительные модификации для повышения стабильности цитозинdezаминазы. Генетические модификации гена дикого типа цитозинdezаминазы дрожжей были проведены таким образом, чтобы этот ген содержал три позиционные мутации, которые меняют три аминокислоты (A23L, I140L и V108I) для повышения термостабильности белка цитозинdezаминазы дрожжей.

Следующие пары праймеров использовали в получении гена для полипептида цитозиндезаминазы по настоящему изобретению:

смысловой: 5'-tcgaggatatcgccgagtgaaacccgttatttttggc-3'

(SEQ ID NO:25)

антисмысловой: 5' -

gccaaaaagaataacgggtttcaactcgccgatatcctcga-3' (SEQ ID NO:26)

смысловой:

5' tcggcgagtatccggcggcgccctccggcggcgccctccggcggcgcc

tccggcgccgcggccaacccgttatt-3' (SEQ ID NO:27)

антисмысловой: 5' -

aataaacgggttggcgccgcgcggaggcgccgcccggaggcgccgcgc

cggaggcgccgcggatcaactcgccga-3' (SEQ ID NO:28)

Для повышения стабильности нативной CD дрожжей три аминокислотные замены были встроены в этот белок. Эти замены были отдельными или в сочетании с оптимизацией кодонов для экспрессии в клетках человека.

Три аминокислотные замены представляли собой: A23L, V108I, I140L. Последовательность, кодирующая эти замены, показана ниже.

```

ATGGTGACAGGGGAATGGCAAGCAAGTGGGATCAGAAGGTATGGACATTGCCTATG
AGGAGGCGTTATTAGGTTACAAAGAGGGTGGTGCCTATTGGCGGATGTCTTATCAATAACA
AAGACGGAAGTGTCTCGGTGTTGTCACAAACATGAGATTCAAAAGGGATCCGCCACACTAC
ATGGTGAGATCTCCACTTGAAAAGTGTGGGAGATTAGAGGGCAAAGTGTACAAAGATAACCA
CTTGATACGACGCTGTCTCCATGCGACATGTGTACAGGTGCCATCATGTATGGTATTG
CACGCTGTGTCATCGGTGAGAACGTTAATTCAAAAGTAAGGGCGAGAAATATTACAAACTA
GAGGTACGGAGGGTGTGTTGACGATGAGAGGTGTAAGGTTAATGAAACAATTATCG
ATGAAAGACCTCAGGATTGGTTGAAGATATTGGTGAGTAGGCGGCCGCGCCATAGATAAAAT
AAAAGATTATTCAGTCTCCAGAAAAAGGGGG (SEQ ID NO:33)

```

Кодируемый полипептид содержит следующую последовательность (замещенные аминокислоты выделены жирным шрифтом и подчеркнуты):

```

1 MVTGGMASKWDQKGMDIAYEEALLGYKEGGVPIGGCLINNKDGSVLGRGHNMRFQKGSAT
61 LHGEISTLENCGRLEGKVYKDTTLSPCDMCTGAIIMYGIPRCVIGENVNFKSKGEK
121 YLQTRGHEVVVVDERCKKLMKQFIDERPQDWFDEDIE-
```

Создание конечной конструкции, которая включает 3 аминокислотные замены A23L/V108I/I140L, используя предпочтительные кодоны, и использует предпочтительную частоту использования кодонов человека для полной последовательности (этот ген называется CDopt+3pt [aka CD2] и T5.0002 при встраивании в плазмидный вектор pAC3, который кодирует RCR с IRES) (SEQ ID NO: 34):

```

1 ATGGTGACCGGGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATATCGCTTACGAG
61 GAGGCCCTGCTGGGCTACAAGGAGGGCGCGTGCCTATGGCGGCTGTCTGATCAACAAAC
121
AAGGACGGCAGTGTGCTGGCAGGGCCACAAACATGAGGTTCCAGAAGGGCTCCGCCACC
181
CTGCACGGCGAGATCTCCACCCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAAG
241
GACACCACCCGTACACCACCCGTCCCCTGTGACATGTGTACCGGCGCTATCATCATG
301
TACGGCATCCCTAGGTGTGATCGGGAGAACGTGAACTTCAAGTCCAAGGGCGAGAAG
361
TACCTGCAAACCAGGGCCACGAGGTGGTGGTGTGACGATGAGAGGTGTAAGAAGCTG
421
ATGAAGCAGTTCATCGACGAGAGGCCTCAGGACTGGTCAGGATATCGGCGAGTGATAA

```

Подчеркнутые кодоны обозначают предпочтительные кодоны для аминокислотных замен.

Конструкция CD оптимизированной последовательности (предпочтительность кодонов человека + 3 аминокислотные замены)

```
AACACGATTATAAATGGTGACCGGCATGGCCTCCAAGTGGATCAAAAGGGCATGGATAT
CGCTTACGAGGAGGCCCTGCTGGCTACAAGGAGGGCGCGTGCCTATCGGCGCTGTCTGAT
CAACAAACAAGGACGGCAGTGTGCTGGCAGGGGCCAACACATGAGGTTCCAGAAGGGCTCCGC
CACCTGCACGGCGAGATCTCACCCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAA
GGACACCACCCCTGTACACCACCCCTGTCCCCTTGTGACATGTGTACCGGCCTATCATCATGTA
CGGCATCCCTAGGTGTGATCGCGAGAACGTGAACCTCAAGTCCAAGGGCGAGAAGTACST
GCAAACCAGGGGCCACGAGGTGGTGGTTGACGATGAGAGGTGTAAGAAGCTGATGAAGCA
GTTCATCGACGAGAGGCCCTCAGGACTGGTCGAGGATATCGGCGAGTAA[GC GG C C C]GCCATA
```

GATAAAATAAAAGATTTATTTAGTCTCCAGAAAAAGGGGGG (SEQ ID NO:35)

С. Конструкция слитого гена CD-UPRT (CDopt+3pt-UPRT, [aka CDopt-UPRT и CD2-UPRT], T5.0003 в pAC3 плазмидном RCR векторе).

Также была разработана слитая конструкция, содержащая полипептид CD, как описано выше, связанный с полипептидом UPRT для получения CD оптимизированной-UPRT последовательности с использованием схемы I, как изложено на фиг. 2А. Следующие праймеры были использованы для делекции стоп-старт между CD и UPRT.

Последовательности праймеров

Название праймера	Последовательность праймера (5'-3') (SEQ ID NO:)			
del118-123	5'-tcgaggatatcgccgagtgaaacccgttattcttttggc-3' (36)			
del118-123 - антисмысловой	5'-gccaaaaagaataacgggtttcactcgccgatatcctcga-3' (37)			
del118-123	40	79,06°C	-44,37 ккал/моль	21,1%
del118-123 - антисмысловой	40	79,06°C	-47,95 ккал/моль	20,3%
Название праймера	Дуплекс праймер-матрица			
del118-123	5'-tcgaggatatcgccgagtga-----aaccgttattcttttggc-3' ccaaagctcctatacgccgctcaactatctacttggcaataagaaaaaccgaag			
del118-123 - антисмысловой	ggttcgaggatatcgccgagtgatagatgaacccgttattcttttggcttc 3'-agctcctatacgccgctcaact-----ttggcaataagaaaaaccg			

Полученный в результате слитый полинуклеотид содержит 1296 пар оснований и последовательность, изложенную непосредственно ниже:

```

AACACGATTATAAATGGTGACCGCGGCATGGCCTCCAAGTGGGATCAAAAGGGCATGGATAT
CGCTTACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGCGTGCCTATCGGCGCTGTCTGAT
CAACAACAAGGACGGCAGTGTCTGGGCAGGGGCCACAACATGAGGTTCCAGAAGGGCTCCGC
CACCTGACGGCGAGATCTCCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAA
GGACACCACCCGTACACCACCCGTCCCCGTGACATGTGTACCGCGCTATCATCATGTA
CGGCATCCCTAGGTGTGATCGCGAGAACGTGAACCTCAAGTCCAAGGGCGAGAAGTACCT
GCAAACCAGGGGCCACGAGGTGGTGGTTGACGATGAGAGGTGTAAGAAGCTGATGAAGCA
GTTCATCGACGAGAGGCCCTCAGGACTGGTCGAGGATATCGCGAGAACCGTTATTCTTTT
GGCTTCTCCATTCTGTACCTTACATATCTTATATATTATCAAACAAAGGGCTTCGTTAG
CAAACCTAGAAATCTGCAAAAATGTCTTCGGAACCATTAAAGAACGTCTACTTGCTACCTCA
AACAAACCAATTGCTGGGTTGTACACCATCATCAGAAATAAGAATACAACTAGACCTGATT
CATTTCTACTCCGATAGAACATCATCAGATTGTTGGTTGAAGAAGGTTGAACCATCTACCTGT
GCAAAGCAAATTGTGGAAACTGACACCAACGAAAATTCTGAAGGTGTCTCATGGTAA
AATCTGTGGTGTTCATTGTCAGAGCTGGTGAATCGATGGAGCAAGGATTAAGAGACTGTTG
TAGGTCTGTGCGTATCGGTAAAATTAAATTCAAAGGGACGAGGAGACTGCTTACCAAAGTT
ATTCTACGAAAATTACCAAGGAGTATATCTGAAAGGTATGTCTTCCTATTAGACCCAATGCT
GGCCACCGGTGGTAGTGCTATCATGGCTACAGAAGTCTGATTAAGAGAGGTGTTAACCCAGA
GAGAATTCTACTCTAAACCTAATCTGTAGTAAGGAAGGGATTGAAAAATACCATGCCGCCTT
CCCAGAGGTAGAACATTGTTACTGGTGCCTCGACAGAGGTCTAGATGAAAACAAGTATCTAGT
TCCAGGGTTGGGTGACTTGGTGACAGATACTACTGTGTTAAGCGGCCGCCATAGATAAA
ATAAAAGATTTATTAGTCTCCAGAAAAGGGGGG (SEQ ID NO:38)

```

D. Конструирование слитого гена CD-линкер UPRT (CDopt+3pt-LINK-UPRT [aka CDopt-LINKER-UPRT и CD2-L-UPRT]).

Слитая конструкция также была получена путем клонирования линкерного (Ser-Gly-Gly-Gly)₄ домена между и в рамке считывания с CD полипептидом и UPRT полипептидом для получения CD оптимизированной-линкер-UPRT последовательности с использованием схемы II, как представлено в 2В. Для встраивания линкера были использованы следующие праймеры.

Название праймера	Последовательность праймера (5' - 3') (SEQ ID NO:)			
	5' -			
ins_60nt_after_477	5' - tcggcgagtgatccggcggcggcgccctccggcgccctccggcg gcggcgccctccggcggcggcgccaacccgttatt-3' (39)			
ins_60nt_after_477- антисмысловой	5' - aataaacgggttggcgccgcccggagggcgccgcggaggcgcc gccgcccggaggcgccgcccggatcactcgccga-3' (40)			
Название праймера	Длина (нуклеотидов)	Tm	Энергия дуплекса при 68 °C	Затрата энергии ошибочно спаренных оснований
ins_60nt_after_477	82	79,77°C	-30,19 ккал/моль	83,3%
ins_60nt_after_477- антисмысловой	82	79,77°C	-32,31 ккал/моль	82,2%

Полученная в результате конструкция имеет размер: 1356 пар оснований и последовательность, представленную непосредственно ниже:

AACACGATTATAAATGGTGACCGCGGCATGGCCTCAAGTGGATCAAAGGGCATGGATAT
CGCTTACGAGGAGGCCCTGCTGGCTACAAGGAGGGCGGTGCCTATCGGCGCTGTCGAT
CAACAACAAGGACGGCAGTGTGCTGGCAGGGCACAAACATGAGGTTCCAGAAGGGCTCCGC
CACCCCTGCACGGCAGATCTCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAA
GGACACCACCCGTACACCACCCGTCCCCCTGTGACATGTGTACCGCGCTATCATCATGTA
CGGCATCCCTAGGTGTGATCGCGAGAACGTGAACCTCAAGTCCAAGGGCGAGAAGTACCT
GCAAACCAGGGCCACGAGGTGGTGGTTGACGATGAGAGGTGTAAGAAGCTGATGAAGCA
GTTCATCGACGAGAGGCCCTCAGGACTGGTCGAGGATATCGGCCAGTCCGGCGGCCCTC
CGGGCGCGCGCCTCCGGCGGCCCTCCGGCGGCCAACCCGTTATTCTTTGGC
TTCTCCATTCTGTACCTTACATATCTTATATTATCCAAACAAAGGGCTTTGTTAGCAA
ACCTAGAAATCTGAAAAATGTCTCGGAACCATTAAAGAACGTCTACTTGCTACCTCAAAC
AAACCAATTGCTGGTTGTACACCATCATCAGAAAATAAGAATACAACTAGACCTGATTCAT
TTTCTACTCCGATAGAATCATCAGATTGTTGGTTGAAGAAGGTTGAACCACACCTGTGCA
AAAGCAAATTGTGAAACTGACACCAACGAAAACCTCGAAGGTGTCTCATTGATGGTAAAT
CTGTGGTGTTCATTGTCAGAGCTGGTGAATCGATGGAGCAAGGATTAAGAGACTGTTGTAG
GTCTGTGCGTATCGTAAATTAAATTCAAAGGGACGAGGAGACTGCTTACCAAAGTTATT
CTACGAAAATTACAGAGGATATCTGAAAGGTATGTCTTCTTATTAGACCCAATGCTGCC
CACCGGTGGTAGTGTATCATGGCTACAGAAGTCTGATTAAGAGAGGTGTTAGCCAGAGAG
AATTACTTCTTAAACCTAATCTGAGTAAGGAAGGGATTGAAAAATACCATGCCGCTTCCC
AGAGGTCAAGAATTGTTACTGGTGCCTCGACAGAGGTCTAGATGAAAACAAGTATCTAGTTCC
AGGGTTGGGTGACTTGGTACAGATACTACTGTGTTAA

GCGGCCGCCATAGATAAAATAAAGATTTATTTAGTCTCCAGAAAAAGGGGGG (SEQ

E. Конструирование слитого гена CD-OPRT (CDopt+3pt-OPRT [aka CDopt-OPRT и CD2-OPRT], T5.0004 при встраивании в рАС3 плазмидный RCR вектор).

Также была получена слитая конструкция, содержащая CD полипептид, как описано выше, соединенный с полипептидом OPRT для получения CD оптимизированной-OPRT-азы (CD гуманизированный + Зрт мутантный + OPRT-аза функциональный домен человека) с использованием схемы III, как показано на фиг. 2С.

Полученная в результате конструкция имеет размер 1269 пар нуклеотидов и последовательность, представленную непосредственно ниже:

AACACGA**TATATAA**ATGGTGACCGCGGCATGGCCTCCAAGTGGGATCAAAAGGCATGGATAT
CGCTTACGAGGAGGCCCTGCTGGGCTACAAGGAGGGCGCGTGCCTATCGGCGGTGCTGTGAT
CAACAACAAGGACGGCAGTGTGCTGGCAGGGCCACAAACATGAGGTTCCAGAAGGGCTCCGC
CACCCCTGCACGGCGAGATCTCACCCTGGAGAACACTGTGGCAGGCTGGAGGGCAAGGTGTACAA
GGACACCACCCCTGTACACCACCCGTCCCCCTGTGACATGTGTACCGGCCTATCATCATGTA
CGGCATCCCTAGGTGTGATCGCGAGAACGTGAACCTCAAGTCCAAGGGCAGAACGTACCT
GCAAACCAGGGCCACGAGGTGGTGGTTGTTGACGATGAGAGGTGTAAGAAGCTGATGAAGCA
GTTCATCGACGAGAGGCCTCAGGACTGGTTGAGGATATCGGCGAGG**GCGGTCGCTCGT**gcagc
tttggggccattggtgacgggtctgtacgacgtgcaggcttcaagttgggacttcgtgct
gaagagcgggctttcccccacatctacatcgatctgcggggcatcgatctcgaccgcgtct
tctgagtcaaggttgcaagatatttattccaaactgcacaaatgcaggcatcagtttgcac
cgtgtgtggagtgccctatacagcttgcattggctacagttatctgttcaaccaaataat
tccaaatgcttattagaaggaaagaaacaaaggattatggactaagcgtctgtagaaggaac
tattaatccaggagaacactgttaatcattgaagatgttgcaccagtggatctagtgttt
ggaaactgttggagttctcagaaggaggcttgaaggtcaactgtatgccatagtgttgg
cagagagcaggaggcaaggacaagttgcaggcgcacgggatccgcctccactcagttgtac
attgtccaaaatgtggagattctcgagcagcagaaaaagttgatgtgagacagttggag
agtgaagaggttattcaggagaatgtttgtggcagcgaatcataatggttctcccttc
tataaaggaagcacccaaagaactcaGCTTCGGTGCACGTGCAGAGCTGCCAGGATCCACCC
AGTTGCATCGAAGTAA**GCGGCCG**GCCATAGATAAAAAGATTATTAGTCTCCAGAA

AAAGGGGGG (SEQ ID NO: 42)

F. Конструирование слитого гена CD-линкер-OPRT (CDopt+3pt-LINKER-OPRT, [aka CDopt-LINKER-OPRT и CD2-L-OPRT], T5.0005 в pAC3 плазмидном RCR векторе).

Слитая конструкция также была получена путем клонирования линкерного (Ser-Gly-Gly-Gly-Gly)₄ домена между и в рамках считывания с CD полипептидом и OPRT полипептидом для получения CD оптимизированной-линкер-OPRT последовательности с использованием схемы IV, как показано на фиг. 2D.

Полученная в результате конструкция имеет размер 1329 пар нуклеотидов и последовательность, представленную непосредственно ниже:

AACACGA**TTATAA**ATGGTGACCGCGGCATGGCCTCAAGTGGGATCAAAAGGCATGGATAT
CGCTTACGAGGAGGCCCTGCTGGCTACAAGGAGGGCGCGTGCCTATCGGCGGTGCTGAT
CAACAACAAGGACGGCAGTGTGCTGGCAGGGCCACAACATGAGGTTCCAAGGGCTCCGC
CACCCCTGCACGGCGAGATCTCACCCTGGAGAACTGTGGCAGGCTGGAGGGCAAGGTGTACAA
GGACACCACCCCTGTACACCACCCGTCCCCCTGTGACATGTGTACCGCGCTATCATCATGTA
CGGCATCCTAGGTGTGTACGGCGAGAACGTGAACCTCAAGTCCAAGGGCGAGAACGTACCT
GCAAACCAGGGGCCACGAGGTGGTGGTTGACGATGAGAGGTGTAAGAACGCTGATGAAAGCA
GTTCATCGACGAGAGGCCTCAGGACTGGTTCGAGGATATCGGCAGTCGGCGGGCGCCCTC
CGGCGGCGGCCTCCGGCGGCCCTCCGGCGGCCGCGCCGGTCGCTCGTGCagctt
qqqqccattqqtqacqqqtctqtaqcacqtcqaqqcttcaaqttqqqqacttcgtqctqaa

gagcgggcttcctccccatcacatcgatctgcggggcatcgtgtcgaccgcgtttct
 gagtcaggttgcagatatttattccaaactgcccataatgcaggcatcagtttgacaccgt
 gtgtggagtgcctatacagcttgcattggctacagttatctgttcaaccaatcaaattcc
 aatgcttatagaagaaaacaaggattatgaaactaagcgtctgttagaaggaactat
 taatccaggagaaacctgttaatcattgaagatgttgcaccagtggatctagtgtttgga
 aactgtttaggttcttcagaaggaggcttgaaggtaactgtatgcctactcagtgatgttgcacag
 agagcagggaggcaaggacaagttgcaggcgacggatccgcctccactcagtgatgttgcacatt
 gtccaaaatgctggagattctcgagcagcagaaaaagttgtatgcgtgagacagttggagagtt
 gaagaggttattcaggagaatgtcttgcggcagcgaatcataatggttctccctttctat
 aaaggaagcacccaaagaactcaGCTTCGGTGCACGTGCAGAGCTGCCAGGATCCACCCAGT
 TGCATCGAAGTAAGCGGCCGC GCCATAGATAAAAGATTATTTAGTCTCCAGAAAAAA
 GGGGG. (SEQ ID NO: 43)

Пример 2. Получение инфицирующего вектора.

Вектор может быть получен целым рядом способов, но первым этапом является введение плазмидной ДНК вектора в клетки, для возможности продукции инфекционных частиц, которые затем могут быть собраны из клеточного супернатанта. После получения инфекционных частиц другие способы получения могут быть выполнены специалистами в данной области. Векторные частицы получали путем транзиторной трансфекции клеток 293T (Pear et al. Proc. Natl. Acad. Sci. USA. 90:8392-8396 1993). Клетки 293T размораживали и помещали в культуру, затем дважды пересевали во флаконы T-75, содержащие 15 мл среды DMEM, которая была получена путем перемешивания среды DMEM с высоким содержанием глюкозы (Hyclone# 30081, 500 мл) с FBS (Hyclone# SH30070, 50 мл), L-глутамином (Cellgro # 25-005-CI, 5 мл), NEAA (Hyclone #SH30238, 5 мл), и пенициллин-стреп (Cellgro # 30-002-CI, 5 мл). Эти флаконы инкубировали при 37°C и 5% CO₂. После 3-го пассажа клетки высевали в 6 флаконов T-25, содержащих каждый 5 мл среды, при плотности клеток 1,8×10⁶ клеток/T-25 (или 7,2×10⁴ клеток/см²). Через один день после высеваания T-25 клетки трансфицировали плазмидой T5.0002, которая экспрессировала вирусный вектор, с использованием набора для трансфекции фосфатом кальция (Calcium Phosphate Transfection Kit) от компании Promega (Cat# E1200). Через 18 ч после трансфекции среду в одном комплекте флаконов (3 флакона в каждом комплекте) заменяли свежей средой, содержащей 10 mM NaB. Среду во 2-м комплекте флаконов не заменяли, который служил в качестве контроля (ноль NaB). Через 8 ч после обработки NaB среду во всех флаконах заменяли свежей средой, не содержащей NaB. В обоих комплектах флаконов допускали продолжение экспрессии до следующего дня (продолжительность 22 ч). Супернатанты из обоих комплектов флаконов собирали и исследовали титры с помощью qPCR, выраженные в единицах трансдукции (TU)/мл (смотри пример 3).

Результаты титрования показаны в следующей табл. 4.

Таблица 4

Условие	Первый титр	Второй титр (после хранения при -80°C в течение 68 дней)
Без NaB	1,5 ($\pm 0,05$) × 10 ⁶ TU/мл	1,2 ($\pm 0,2$) × 10 ⁶ TU/мл
10 mM NaB	1,4 ($\pm 0,3$) × 10 ⁶ TU/мл	7,0 ($\pm 0,14$) × 10 ⁵ TU/мл

Последующие препараты вектора были получены таким же образом, без бутират анатрия. Другие векторные плазмиды были использованы таким же образом для получения векторных препаратов с титрами от 10⁵ до 10⁷ TU/мл. Такой материал дополнительно может быть очищен и концентрирован, при желании, как описано ниже, а также см. US 5792643; T. Rodriguez et al. J. Gene Med. 9:233 2007; заявка США 61218063. В некоторых вариантах осуществления настоящего изобретения дозу вычисляли по граммам массы головного мозга. В таких вариантах осуществления доза репликационно-компетентного ретровирусного вектора по настоящему изобретению, применимая в способах для лечения, может находиться в интервале от 10³ до 10⁷ TU на 1 г массы головного мозга.

Пример 3. Количественный ПЦР анализ титрования.

Функциональную концентрацию вектора, или титр, определяют с использованием количественного способа на основе ПЦР (qPCR). В этом способе вектор титруют путем инфицирования трансдуцируемой линии клеток-хозяев (например, клетки карциномы предстательной железы человека PC-3, ATCC Cat # CRL-1435) со стандартным объемом вектора и измерения результирующего количество провируса, находящегося в клетках-хозяевах после трансдукции. Клетки и вектор инкубируют в стандартных усло-

виях культивирования (37°C , 5% CO_2) в течение 24 ч, чтобы дать возможность завершить инфицирование до добавления антиретровирусного AZT для остановки репликации вектора.

Далее, клетки собирают с культуральной чашки и геномную ДНК (гДНК) очищают с использованием набора для очистки геномной ДНК Invitrogen Purelink и элюируют из колонки для очистки с использованием стерильной воды без содержания РНазы/ДНазы. Соотношение поглощения A260/A280 изменяют на спектрофотометре для определения концентрации и относительной чистоты образца. Концентрации гДНК нормализуют с использованием дополнительного количества воды без содержания РНКазы-/ДНКазы до наиболее низкой концентрации любого заданного набора препаратов гДНК, так, чтобы входящая ДНК для qPCR оставалась неизменной для всех анализируемых образцов.

Чистоту геномной ДНК дополнительно определяют с помощью электрофореза аликвоты каждого образца на окрашенном этидия бромидом 0,8% агарозном геле. Если образец проходит диапазон поглощения при A260/A280 1,8-2,0 и демонстрирует одну полосу гДНК, тогда этот образец готов для проведения qPCR анализа провирусного числа копий вектора. Используя праймеры, которые задают область LTR провируса (обратно транскрибированная векторная ДНК и векторная ДНК, которая интегрируется в гДНК хозяина), qPCR проводят для оценки суммарного числа событий трансдукции, которые происходят, когда известный объем вектора использовали для трансдукции известного числа клеток. Количество событий трансдукции на реакцию вычисляют по стандартной кривой, в которой используется плазмида, несущая мишень известного числа копий, которая серийно разбавлена от $1\text{E}7$ до $1\text{E}1$ копий и измерена в идентичных условиях qPCR, как образцы. Зная, сколько геномных эквивалентов было использовано для каждой реакции qPCR (из концентрации, определенной ранее) и сколько событий трансдукции, которые произошли на реакцию, авторы изобретения определили суммарное число событий трансдукции, которые произошли, исходя из суммарного числа клеток, которые присутствовали во время трансдукции. Это значение является титром вектора после разбавления в среде, содержащей клетки во время исходной трансдукции. Для вычисления правильного значения титра, разбавление корректируется путем умножения на коэффициент, равный сумме объема культуры и объема титра, деленные на объем титра. Эти эксперименты проводят в реплицирующихся культурах и анализируют с помощью qPCR, используя измерения в трех повторах для каждого условия для определения среднего титра и связанного с ним стандартного отклонения и коэффициента вариантности.

Пример 4. Уровни экспрессии, измеренные с помощью вестерн-блоттинга.

На фиг. 3 показано, что наблюдаются более высокие уровни кодонов, оптимизированных для человека с тремя мутациями для более высокой стабильности по сравнению с белком yCD дикого типа в вестерн-блот анализе клеток U-87, инфицированных вирусом, кодирующим либо дикого типа (ACE-yCD), либо полностью оптимизированные (AC3-yCD2) гены цитозинdezаминазы.

Пример 5. Генетическая стабильность вирусных векторов.

Известно, что после обратной транскрипции и первого события интеграции в обработанные клетки, ДНК провируса и любое последующее потомство ретровируса имеет обычную LTR структуру из MLV на любом конце. Было показано, что эта конфигурация является стабильной после многократных циклов инфицирования (см. фиг. 4).

Приблизительно 10^6 интактных клеток U-87 первоначально было инфицировано вирусным вектором при MOI 0,01, и их выращивали до полного инфицирования для завершения одного цикла инфекции. Затем супернатант снова возвращали на инфицированные клетки, а затем цикл повторяли. Геномную стабильность последовательности yCD2 оценивали с помощью ПЦР амплификации интегрированного провируса из инфицированных клеток, используя MLV специфические праймеры, фланкирующие сайт встраивания трансгена. Для каждой группы инфекций, амплификацию векторной плазмида (pAC3-yCD2 и Kasahara et al. вектор pACE-CD) также поводили, чтобы проследить расположение по размеру полноразмерного ампликона на геле. Появление любых полос, меньших полноразмерного ампликона, будет показателем нестабильности вектора. Такие эксперименты демонстрировали, что вектор по настоящему изобретению (T5.0002 - содержащий модифицированный вектор, и CDopt+3pt (CD2) гетерологичный полинуклеотид сохраняли стабильность на протяжении большего количества пассажей, чем pACE-CD или T5.007, каждый из которых несет (полинуклеотид) дрожжей дикого типа.

Пример 6. Эксперименты по цитолизу клеток.

В экспериментах на клеточных культурах *in vitro* показано, что цитозинdezаминаза в клетках, экспрессирующих белок yCD2, по меньшей мере, является такой же активной, как и цитозинdezаминаза из клеток, экспрессирующих белок yCD дикого типа, путем проведения титрований 5-FC с использованием клеток крыс RG2 (фиг. 5A) или клеток U-87 (фиг. 5B), инфицированных либо вирусом, полученным из pAC3-yCD2/T5.0002) [AC3-yCD2(V)] из pAC3-yCD/T5.0007 [AC3-yCD(V)], либо другими векторами. Вкратце, для клеток U-87 через 5 дней после инфицирования при множественности заражения 0,1 (т.е. 100% инфицированных) либо вектором AC3-yCD (CD дикого типа), либо вектором AC3-yCD2 (термо-стабильным и кодон-оптимизированным) подвергали воздействию повышенных количеств 5-FC или 0,1 mM 5-FU в качестве положительного контроля в течение 8 дней. На 8 день обработки 5-FC жизнеспособность клеточных культур оценивали с помощью анализа MTS (Promega CellTiter 96 AQUEOUS One Solution Proliferation Assay). Данные показывают сравнимый цитолиз между двумя ретровирусными век-

торами при повышенных дозах обработки с использованием 5-FC. Культуры RG2 обрабатывали подобным образом, и также показано, если вообще что-либо, слабый сдвиг кривой цитолиза в направлении более низкой концентрации 5-FC для вируса из анализа экспрессии T5.0002CD. Клетки U87 трансдудировали при множественности заражения (MOI) 0,1, культивировали в течение 5 дней, чтобы дать возможность для распространения вируса и собирали клетки с 5 дня после трансдукции. Затем клетки собирали центрифугированием при 800×g в течение 5 мин. Супернатант аспирировали из клеточного осадка и промывали 5 мл фосфатно-буферного солевого раствора (PBS) и снова центрифугировали при 800×g в течение 5 мин. Полученный в результате клеточный осадок переносили в 1,5 мл PBS, ресуспендировали пропусканием через наконечник пипетки, и помещали в морозильную камеру при -20°C. Клетки лизировали методом замораживания/оттаивания. Ранее ресуспендированные клетки размораживали при комнатной температуре, пропускали через наконечник пипетки, перемешивали со смесью ингибиторов протеаз и снова замораживали при -20°C. Перед проведением ферментативного исследования, образцы снова размораживали при комнатной температуре и пропускали через носик пипетки. Затем суспензию центрифугировали при 14000 об/мин в настольной центрифуге в течение 5 мин. Супернатант отбрасывали из осадка и помещали в новую пробирку эпендорфа и помещали на лед. Ферментативную активность yCD оценивали, путем использования анализа HPLC.

Анализ HPLC проводили на Shimadzu устройстве LC20AT, соединенном последовательно с фотометрическим детектором и автоинжектором. Твердая фаза представляла собой Hypersil BDS C18 HPLC колонку с 5 мкм размером сфер и 4,0×250 мм размерами колонки. Подвижная фаза представляла собой 50 mM фосфат аммония, pH 2,1, содержащий 0,01% трет-бутиламмония перхлорат и 5% метanol; эта система была уравновешена при 22°C. Все реактивы были ACS степени чистоты, а растворители были степени чистоты "для ВЭЖХ". Получали реакционную смесь, состоящую из 800 мкл с конечной концентрацией 0,125 мг/мл 5FC (1 мМ) в PBS и помещали в 1,5 мл флаконы для автоматического дозатора. Затем инициировали реакцию путем добавления 200 мкл каждого клеточного лизата. Реакционную смесь/флаконы для автоматического дозатора помещали в автодозатор и вводили 5 мкл реакционной смеси. Моменты времени были взяты периодически, путем извлечения 5 мкл аликвоты из каждого реакционного флакона и анализа на ВЭЖХ колонке. Скорости превращения 5FC в 5FU вычисляли путем сравнения площадей пика с известными количествами из ранее полученной стандартной кривой 5FU.

Скорость превращения 5FC в 5FU была получена путем нанесения на график количества 5FU (в нмолях) в зависимости от соответствующего интервала времени. Была получена концентрация белка для образца клеток и специфическую активность образцов клеточного лизата вычисляли путем деления скорости превращения (нмоль/мин) на количество белка, используемого в исследовании, в мг. На фиг. 6 показана специфическая активность векторов после 5 дней трансдукции при MOI 0,1. Данные демонстрируют, что pACE-yCD (T5.0000) < pAC3-yCD1(T5.0001) < pAC3-CD2 (T5.0002) в значении специфической активности цитозинdezаминазы в клетках культуры ткани.

Пример 7. Опухоли, обработанные полностью модифицированным геном CD (yCD2), элиминируются более эффективно, чем опухоли, обработанные немодифицированным геном CD дрожжей, и не рецидивируют (PR-01-08-001).

Для определения, какая векторная конструкция приводит к получению наиболее эффективного вектора в модели подкожного ксенотрансплантата мыши/человека. Оценивали три различные конструкции: T5.0001 (частично модифицированный (ген) CD); T5.0007 (немодифицированный ген CD дрожжей); T5.0002 (полностью модифицированный ген yCD2). Рост опухоли, выживание и регрессию опухоли после введения пролекарства 5-FC оценивали в модели подкожной глиомы человека (U87) у мышей с иммунодефицитом. Оценивали две различные концентрации 5-FC для определения ответа на дозу, наблюдаемого между 5-FC и конструкциями опухоль-вектор.

Всего было исследовано 12 групп, состоящих из 9-11 самок мышей на группу. Все мыши подвергались имплантации на спине с правого бока в 0 день либо: 98% смеси линии неинфицированных опухолевых клеток U-87 и 2% смеси линии клеток U-87, которые были инфицированы одной из трех векторных конструкций TOCA 511, T5.0002, T5.0001, T5.0007; или неинфицированной контрольной линии клеток U-87 (100%). Мыши инокулировали 2×10^6 клеток/мышь. Мыши в группах 1-3 не были заражены клетками U87, в группах 4-6 получали смесь клеток U87, трансдудированных T5.0002 U87, группы 7-9 получали смесь клеток U87, трансдудированных T5.0001, и группы 10-12 получали смесь клеток U87, трансдудированных T5.0007. Опухоли выращивали в течение 6 дней, пока размер опухоли не достигал приблизительно 100 mm³. Каждая группа мышей, которым вводили вектор, была рандомизирована для введения одной из двух доз 5-FC (200 или 500 мг/кг/день), вводимых в виде однократной внутрибрюшинной инъекции начиная с 6 дня без введения 5-FC. Введение 5-FC продолжали ежедневно в течение 28 последующих дней. Выживание мышей после 29 дня оценивали по размеру опухоли, а затем скарифицировали в различные дни, извлекая опухоли для анализа, при их наличии. На 29 день мыши из групп 4-12 снова были рандомизированы и мышей подразделяли на 2 подгруппы (либо продолжали введение 5-FC для наблюдения за регрессией опухоли, либо прекращали воздействие 5-FC для наблюдения за возобновлением опухолевого роста).

Результаты: во всех опухолях, обработанных T5.0002, T5.0000 и T5.0007, демонстрировалась регрессия опухоли до неопределенной при дозе 200 и 500 мг/кг на 29 день. У всех животных, получивших воздействие T5.0002, опухоли заново не образовывались, когда прекращалось воздействие 5FC, до 39 дней. В отличие от этого, некоторые опухоли, обработанные T5.0001 и T5.0007, рецидивировали к 39 дню. Эксперимент заканчивали на 39 день.

Заключение: T5.0002 плюс 5-FC представляет собой более эффективную противоопухолевую терапию, чем T5.0001 или T5.0007 плюс 5-FC.

Был описан ряд вариантов осуществления настоящего изобретения. Тем не менее, будет понятно, что различные модификации могут быть сделаны, не отклоняясь от сущности и объема настоящего изобретения. В соответствии с этим другие варианты осуществления входят в объем следующей формулы изобретения.

Список последовательностей

<110>	Tocagen Inc. Gruber, Harry E. Jolly, Douglas Perez, Omar	
<120>	ВЕКТОРЫ ДЛЯ ГЕННОЙ ТЕРАПИИ И ЦИТОЗИНДЕЗАМИНАЗЫ	
<130>	00014-007W01	
<140>	Еще не известно	
<141>	2009-09-26	
<150>	US 61/100,666	
<151>	2008-09-26	
<150>	US 61/120,618	
<151>	2008-12-08	
<150>	US 61/186,823	
<151>	2009-06-13	
<160>	22	
<170>	PatentIn version 3.5	
<210>	1	
<211>	477	
<212>	ДНК	
<213>	Saccharomyces cerevisiae	
<220>		
<221>	CDS	
<222>	(1)..(477)	
<400>	1	
	atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac	48
	Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
	1 5 10 15	
	att gcc tat gag gag gcg gcc tta ggt tac aaa gag ggt ggt gtt cct	96
	Ile Ala Tyr Glu Ala Ala Leu Gly Tyr Lys Glu Gly Val Pro	
	20 25 30	
	att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt	144
	Ile Gly Cys Leu Ile Asn Lys Asp Gly Ser Val Leu Gly Arg	
	35 40 45	
	ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag	192
	Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
	50 55 60	
	atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa	240
	Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
	65 70 75 80	
	gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt	288
	Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
	85 90 95	
	gcc atc atc atg tat ggt att cca cgc tgt gtc ggt gag aac gtt	336
	Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val	

027693

100	105	110	
aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu 115 120 125			384
gtt gtt gtt gtt gac gat gag agg tgt aaa aag atc atg aaa caa ttt Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe 130 135 140			432
atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu 145 150 155			477
<210> 2			
<211> 158			
<212> Белок			
<213> <i>Saccharomyces cerevisiae</i>			
<400> 2			
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp 1 5 10 15			
Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro 20 25 30			
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg 35 40 45			
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu 50 55 60			
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys 65 70 75 80			
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly 85 90 95			
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val 100 105 110			
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu 115 120 125			
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe 130 135 140			
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu 145 150 155			
<210> 3			
<211> 477			

027693

<212> ДНК
<213> Искусственная последовательность

<220>
<223> Сконструированная цитозинdezаминаза

<220>
<221> CDS
<222> (1)..(477)

<400> 3
atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac 48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

att gcc tat gag gag gcg tta tta ggt tac aaa gag ggt ggt gtt cct 96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt 144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag 192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa 240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt 288
Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

gcc atc atc atg tat ggt att cca cgc tgt gtc atc ggt gag aac gtt 336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag 384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

gtt gtt gtt gac gat gag agg tgt aaa aag tta atg aaa caa ttt 432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag 477
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 4
<211> 158
<212> Белок
<213> Искусственная последовательность

<220>
<223> Синтетическая конструкция

<400> 4
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp

027693

1	5	10	15
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro			
20	25	30	
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg			
35	40	45	
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu			
50	55	60	
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys			
65	70	75	80
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly			
85	90	95	
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val			
100	105	110	
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu			
115	120	125	
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe			
130	135	140	
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu			
145	150	155	
<210>	5		
<211>	480		
<212>	ДНК		
<213>	Искусственная последовательность		
<220>			
<223>	Цитозинdezаминаза с оптимизированными кодонами человека		
<220>			
<221>	CDS		
<222>	(1)...(480)		
<400>	5		
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat		48	
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp			
1	5	10	15
atc gct tac gag gag gcc gca ctg ggc tac aag gag ggc ggc gtg cct		96	
Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro			
20	25	30	
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg		144	
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg			
35	40	45	

027693

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu 50 55 60	192
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys 65 70 75 80	240
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly 85 90 95	288
gct atc atc atg tac ggc atc cct agg tgt gtg gtc ggc gag aac gtg Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val 100 105 110	336
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu 115 120 125	384
gtg gtg gtt gtt gac gat gag agg tgt aag aag atc atg aag cag ttc Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe 130 135 140	432
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tga taa Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu 145 150 155	480
 <210> 6 <211> 158 <212> Белок <213> Искусственная последовательность	
 <220> <223> Синтетическая конструкция	
 <400> 6	
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp 1 5 10 15	
Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro 20 25 30	
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg 35 40 45	
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu 50 55 60	
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys 65 70 75 80	
Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly 85 90 95	

027693

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
145 150 155

<210> 7
<211> 756
<212> ДНК
<213> *Saccharomyces cerevisiae*

<220>
<221> CDS
<222> (1)..(756)

<400> 7
atg aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat 48
Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr
1 5 10 15

ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat 96
Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn
20 25 30

ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct 144
Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro
35 40 45

caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat 192
Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn
50 55 60

aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga ttg 240
Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu
65 70 75 80

ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg 288
Leu Val Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val
85 90 95

gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa 336
Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys
100 105 110

atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga 384
Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly
115 120 125

tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa 432
Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln
130 135 140

agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta cca 480

027693

Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro			
145	150	155	160
gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc			528
Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala			
165	170	175	
acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga ggt			576
Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly			
180	185	190	
gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag gaa			624
Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu			
195	200	205	
ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act			672
Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr			
210	215	220	
ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg			720
Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly			
225	230	235	240
ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa			756
Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val			
245	250		
<210> 8			
<211> 251			
<212> Белок			
<213> <i>Saccharomyces cerevisiae</i>			
<400> 8			
Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr			
1	5	10	15
Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn			
20	25	30	
Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro			
35	40	45	
Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn			
50	55	60	
Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu			
65	70	75	80
Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val			
85	90	95	
Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys			
100	105	110	
Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly			

027693

115	120	125
Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln		
130	135	140
Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro		
145	150	155
Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala		
165	170	175
Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly		
180	185	190
Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu		
195	200	205
Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr		
210	215	220
Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly		
225	230	235
Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val		
245	250	
<210> 9		
<211> 1443		
<212> ДНК		
<213> homo sapiens		
<220>		
<221> CDS		
<222> (1)..(1443)		
<400> 9		
atg gct gtt gct cgt gct gct ctt ggt cct ctt gtt act ggt ctt tat		48
Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr		
1	5	10
		15
gat gtt caa gct ttt aaa ttt ggt gat ttt gtt ctt aaa tct ggt ctt		96
Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu		
20	25	30
tct tct cct att tat att gat ctt cgt ggt att gtt tct cgt cct cgt		144
Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg		
35	40	45
ctt ctt tct caa gtt gct gat att ctt ttt caa act gct caa aat gct		192
Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala		
50	55	60
ggt att tct ttt gat act gtt tgt ggt gtt cct tat act gct ctt cct		240
Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro		

027693

65	70	75	80	
ctt gct act gtt att tgt tct act aat caa att cct atg ctt att cgt Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg 85 90 95				288
cgt aaa gaa act aaa gat tat ggt act aaa cgt ctt gtt gaa ggt act Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr 100 105 110				336
att aat cct ggt gaa act tgt ctt att att gaa gat gtt gtt act tct Ile Asn Pro Gly Glu Thr Cys Leu Ile Glu Asp Val Val Thr Ser 115 120 125				384
ggg tct tct gtt ctt gaa act gtt gaa gtt ctt caa aaa gaa ggt ctt Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu 130 135 140				432
aaa gtt act gat gct att gtt ctt ctt gat cgt gaa caa gaa ggt ggt aaa Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys 145 150 155 160				480
gat aaa ctt caa gct cat ggt att cgt ctt cat tct gtt tgt act ctt Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu 165 170 175				528
tct aaa atg ctt gaa att ctt gaa caa caa aaa gtt gat gct gaa Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu 180 185 190				576
act gtt ggt cgt gtt aaa cgt ttt att caa gaa aat gtt ttt gtt gct Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala 195 200 205				624
gct aat cat aat ggt tct ctt tct att aaa gaa gct cct aaa gaa Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu 210 215 220				672
ctt tct ttt ggt gct cgt gct gaa ctt cct cgt att cat cct gtt gct Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala 225 230 235 240				720
tct aaa ctt ctt cgt ctt atg caa aaa aaa gaa act aat ctt tgt ctt Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu 245 250 255				768
tct gct gat gtt tct ctt gct cgt gaa ctt ctt caa ctt gct gat gct Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala 260 265 270				816
ctt ggt cct tct att tgt atg ctt aaa act cat gtt gat att ctt aat Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn 275 280 285				864
gat ttt act ctt gat gtt atg aaa gaa ctt att act ctt gct aaa tgt Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys 290 295 300				912
cat gaa ttt ctt att ttt gaa gat cgt aaa ttt gct gat att ggt aat His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn 305 310 315 320				960
act gtt aaa aaa caa tat gaa ggt ggt att ttt aaa att gct tct tgg Thr Val Lys Lys Gln Tyr Glu Gly Ile Phe Lys Ile Ala Ser Trp				1008

027693

325	330	335	
gct gat ctt gtt aat gct cat gtt gtt cct ggt tct ggt gtt aaa Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys 340 345 350			1056
ggt ctt caa gaa gtt ggt ctt cct cat cgt ggt tgt ctt ctt att Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile 355 360 365			1104
gct gaa atg tct tct act ggt tct ctt gct act ggt gat tat act cgt Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg 370 375 380			1152
gct gct gtt cgt atg gct gaa gaa cat tct gaa ttt gtt gtt ggt ttt Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe 385 390 395 400			1200
att tct ggt tct cgt gtt tct atg aaa cct gaa ttt ctt cat ctt act Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr 405 410 415			1248
cct ggt gtt caa ctt gaa gct ggt ggt gat aat ctt ggt caa caa tat Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr 420 425 430			1296
aat tct cct caa gaa gtt att ggt aaa cgt ggt tct gat att att att Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile 435 440 445			1344
gtt ggt cgt ggt att att tct gct gat cgt ctt gaa gct gct gaa Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu 450 455 460			1392
atg tat cgt aaa gct gct tgg gaa gct tat ctt tct cgt ctt ggt gtt Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val 465 470 475 480			1440
taa			1443
<210> 10 <211> 480 <212> Белок <213> homo sapiens			
<400> 10			
Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr 1 5 10 15			
Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu 20 25 30			
Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg 35 40 45			
Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala 50 55 60			

027693

Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro
65 70 75 80

Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg
85 90 95

Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr
100 105 110

Ile Asn Pro Gly Glu Thr Cys Leu Ile Glu Asp Val Val Thr Ser
115 120 125

Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu
130 135 140

Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Lys
145 150 155 160

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

Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu
180 185 190

Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala
195 200 205

Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu
210 215 220

Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala
225 230 235 240

Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu
245 250 255

Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala
260 265 270

Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn
275 280 285

Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys
290 295 300

His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn
305 310 315 320

027693

Thr Val Lys Lys Gln Tyr Glu Gly Gly Ile Phe Lys Ile Ala Ser Trp
325 330 335

Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys
340 345 350

Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile
355 360 365

Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg
370 375 380

Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe
385 390 395 400

Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr
405 410 415

Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr
420 425 430

Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile
435 440 445

Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu
450 455 460

Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val
465 470 475 480

<210> 11
<211> 1227
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Слитая конструкция CDopt-UPRT

<220>
<221> CDS
<222> (1)..(1227)

<400> 11
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg 144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg

027693

35	40	45	
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu 50 55 60			192
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys 65 70 75 80			240
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly 85 90 95			288
gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val 100 105 110			336
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu 115 120 125			384
gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe 130 135 140			432
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag aac ccg Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro 145 150 155 160			480
tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat ctt ata tat Leu Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr 165 170 175			528
tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat ctg caa aaa Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys 180 185 190			576
atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct caa aca aac Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn 195 200 205			624
caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat aca act aga Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg 210 215 220			672
cct gat ttc att ttc tac tcc gat aga atc atc aga ttg ttg gtt gaa Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu 225 230 235 240			720
gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg gaa act gac Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp 245 250 255			768
acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa atc tgt ggt Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly 260 265 270			816
gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga tta aga gac Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp 275 280 285			864
tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa agg gac gag Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu			912

027693

290	295	300	
gag act gct tta cca aag tta ttc tac gaa aaa tta cca gag gat ata Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile 305 310 315 320			960
tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc acc ggt ggt Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly 325 330 335			1008
agt gct atc atg gct aca gaa gtc ttg att aag aga ggt gtt aag cca Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro 340 345 350			1056
gag aga att tac ttc tta aac cta atc tgt agt aag gaa ggg att gaa Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu 355 360 365			1104
aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act ggt gcc ctc Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu 370 375 380			1152
gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg ttg ggt gac Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp 385 390 395 400			1200
ttt ggt gac aga tac tac tgt gtt taa Phe Gly Asp Arg Tyr Tyr Cys Val 405			1227
<210> 12			
<211> 408			
<212> Белок			
<213> Искусственная последовательность			
<220>			
<223> Синтетическая конструкция			
<400> 12			
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp 1 5 10 15			
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro 20 25 30			
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg 35 40 45			
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu 50 55 60			
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys 65 70 75 80			
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly 85 90 95			

027693

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro
145 150 155 160

Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr
165 170 175

Tyr Pro Asn Lys Gly Ser Phe Val Scr Lys Pro Arg Asn Leu Gln Lys
180 185 190

Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn
195 200 205

Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg
210 215 220

Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu
225 230 235 240

Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp
245 250 255

Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly
260 265 270

Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp
275 280 285

Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu
290 295 300

Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile
305 310 315 320

Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly
325 330 335

Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro
340 345 350

027693

Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu
355 360 365

Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu
370 375 380

Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
385 390 395 400

Phe Gly Asp Arg Tyr Tyr Cys Val
405

<210> 13
<211> 1287
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Слитая конструкция - CDopt - линкер - UPRT

<220>
<221> CDS
<222> (1)..(1287)

<400> 13
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
Ile Ala Tyr Glu Ala Leu Leu Gly Tyr Lys Glu Gly Val Pro
20 25 30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg 144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag 192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag 240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc 288
Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg 336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag 384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

gtg gtg gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc 432

027693

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe 130 135 140	
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly 145 150 155 160	480
ggc ggc gcc tcc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc Gly Gly Ala Ser Gly Gly Ala Ser Gly Gly Ala Ser Gly Gly 165 170 175	528
ggc gcc aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr 180 185 190	576
tat ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg 195 200 205	624
aat ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta Asn Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu 210 215 220	672
cct caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag Pro Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys 225 230 235 240	720
aat aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg 245 250 255	768
ttg ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile 260 265 270	816
gtg gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly 275 280 285	864
aaa atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln 290 295 300	912
gga tta aga gac tgt tgt agg tct gtc cgt atc ggt aaa att tta att Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile 305 310 315 320	960
caa agg gag gag act gct tta cca aag tta ttc tac gaa aaa tta Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu 325 330 335	1008
cca gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu 340 345 350	1056
gcc acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg 355 360 365	1104
ggt gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys 370 375 380	1152
gaa ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt	1200

027693

Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val
385 390 395 400

act ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca 1248
Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro
405 410 415

ggg ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa 1287
Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
420 425

<210> 14
<211> 428
<212> Белок
<213> Искусственная последовательность

<220>
<223> Синтетическая конструкция

<400> 14

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
145 150 155 160

Gly Gly Ala Ser Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
165 170 175

027693

Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr
180 185 190

Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg
195 200 205

Asn Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu
210 215 220

Pro Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys
225 230 235 240

Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg
245 250 255

Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile
260 265 270

Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly
275 280 285

Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln
290 295 300

Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile
305 310 315 320

Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu
325 330 335

Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu
340 345 350

Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg
355 360 365

Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys
370 375 380

Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val
385 390 395 400

Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro
405 410 415

Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
420 425

027693

<210> 15
<211> 1200
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Слитая конструкция - CDopt3 - OPRT

<220>
<221> CDS
<222> (1)..(1200)

<400> 15
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat 48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg 144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag 192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag 240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc 288
Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg 336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag 384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

gtg gtg gtt gtc gac gat gag agg tgt aag aag ctg atg aag cag ttc 432
Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag ggc gtc 480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Val
145 150 155 160

gct cgt gca gct ttg ggg cca ttg gtg acg ggt ctg tac gac gtg cag 528
Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr Asp Val Gln
165 170 175

gct ttc aag ttt ggg gac ttc gtg ctg aag agc ggg ctt tcc tcc ccc 576
Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu Ser Ser Pro
180 185 190

027693

atc tac atc gat ctg cg ^g ggc atc gtg tct cga ccg cgt ctt ctg agt Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg Leu Leu Ser 195 200 205	624
cag gtt gca gat att tta ttc caa act gcc caa aat gca ggc atc agt Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala Gly Ile Ser 210 215 220	672
ttt gac acc gtg tgt gga gtg cct tat aca gct ttg cca ttg gct aca Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro Leu Ala Thr 225 230 235 240	720
gtt atc tgt tca acc aat caa att cca atg ctt att aga agg aaa gaa Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu 245 250 255	768
aca aag gat tat gga act aag cgt ctt gta gaa gga act att aat cca Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro 260 265 270	816
gga gaa acc tgt tta atc att gaa gat gtt gtc acc agt gga tct agt Gly Glu Thr Cys Leu Ile Glu Asp Val Val Thr Ser Gly Ser Ser 275 280 285	864
gtt ttg gaa act gtt gag gtt ctt cag aag gag ggc ttg aag gtc act Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr 290 295 300	912
gat gcc ata gtg ctg ttg gac aga gag cag gga ggc aag gac aag ttg Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Lys Asp Lys Leu 305 310 315 320	960
cag gcg cac ggg atc cgc ctc cac tca gtg tgt aca ttg tcc aaa atg Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met 325 330 335	1008
ctg gag att ctc gag cag cag aaa aaa gtt gat gct gag aca gtt ggg Leu Glu Ile Leu Glu Gln Gln Lys Val Asp Ala Glu Thr Val Gly 340 345 350	1056
aga gtg aag agg ttt att cag gag aat gtc ttt gtg gca gcg aat cat Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His 355 360 365	1104
aat ggt tct ccc ctt tct ata aag gaa gca ccc aaa gaa ctc agc ttc Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu Leu Ser Phe 370 375 380	1152
ggt gca cgt gca gag ctg ccc agg atc cac cca gtt gca tcg aag taa Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys 385 390 395	1200
 <210> 16 <211> 399 <212> Белок <213> Искусственная последовательность	
 <220> <223> Синтетическая конструкция	
 <400> 16	
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	

027693

1	5	10	15
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro			
20	25	30	
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg			
35	40	45	
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu			
50	55	60	
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys			
65	70	75	80
Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly			
85	90	95	
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val			
100	105	110	
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu			
115	120	125	
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe			
130	135	140	
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Val			
145	150	155	160
Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr Asp Val Gln			
165	170	175	
Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu Ser Ser Pro			
180	185	190	
Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg Leu Leu Ser			
195	200	205	
Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala Gly Ile Ser			
210	215	220	
Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro Leu Ala Thr			
225	230	235	240
Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu			
245	250	255	

027693

Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro
260 265 270

Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser Gly Ser Ser
275 280 285

Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr
290 295 300

Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys Asp Lys Leu
305 310 315 320

Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met
325 330 335

Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu Thr Val Gly
340 345 350

Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His
355 360 365

Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu Leu Ser Phe
370 375 380

Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys
385 390 395

<210> 17
<211> 1260
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Слитая конструкция - CDopt3 - линкер - OPRT

<220>
<221> CDS
<222> (1)..(1260)

<400> 17 48
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct 96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg 144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag 192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu

027693

50	55	60	
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys 65 70 75 80			240
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc Asp Thr Thr Leu Tyr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly 85 90 95			288
gct atc atc atg tac ggc atc cct agg tgt gtc atc ggc gag aac gtg Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val 100 105 110			336
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu 115 120 125			384
gtg gtg gtt gtc gac gat gag agg tgt aag aag ctg atg aag cag ttc Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe 130 135 140			432
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly 145 150 155 160			480
ggc ggc gcc tcc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc Gly Gly Ala Ser Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly 165 170 175			528
ggc gcc gcg gtc gct cgt gca gct ttg ggg cca ttg gtc acg ggt ctg Gly Ala Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu 180 185 190			576
tac gac gtg cag gct ttc aag ttt ggg gac ttc gtg ctg aag agc agg Tyr Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly 195 200 205			624
ctt tcc tcc ccc atc tac atc gat ctg cgg ggc atc gtg tct cga ccg Leu Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro 210 215 220			672
cgt ctt ctg agt cag gtt gca gat att tta ttc caa act gcc caa aat Arg Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn 225 230 235 240			720
gca ggc atc agt ttt gac acc gtc tgt gga gtc cct tat aca gct ttg Ala Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu 245 250 255			768
cca ttg gct aca gtt atc tgt tca acc aat caa att cca atg ctt att Pro Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile 260 265 270			816
aga agg aaa gaa aca aag gat tat gga act aag cgt ctt gta gaa gga Arg Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly 275 280 285			864
act att aat cca gga gaa acc tgt tta atc att gaa gat gtt gtc acc Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr 290 295 300			912
agt gga tct agt gtt ttg gaa act gtt gag gtc ctt cag aag gag ggc Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly			960

027693

305	310	315	320	
ttg aag gtc act gat gcc ata gtg ctg ttg gac aga gag cag gga ggc Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly 325 330 335				1008
aag gac aag ttg cag gcg cac ggg atc cgc ctc cac tca gtg tgt aca Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr 340 345 350				1056
ttg tcc aaa atg ctg gag att ctc gag cag cag aaa aaa gtt gat gct Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala 355 360 365				1104
gag aca gtt ggg aga gta aag agg ttt att cag gag aat gtc ttt gtg Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val 370 375 380				1152
gca gcg aat cat aat ggt tct ccc ctt tct ata aag gaa gca ccc aaa Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys 385 390 395 400				1200
gaa ctc agc ttc ggt gca cgt gca gag ctg ccc agg atc cac cca gtt Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val 405 410 415				1248
gca tcg aag taa Ala Ser Lys				1260

<210> 18
<211> 419
<212> Белок
<213> Искусственная последовательность

<220>
<223> Синтетическая конструкция

<400> 18

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

027693

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Leu Met Lys Gln Phe
130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
145 150 155 160

Gly Gly Ala Ser Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
165 170 175

Gly Ala Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu
180 185 190

Tyr Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly
195 200 205

Leu Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro
210 215 220

Arg Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn
225 230 235 240

Ala Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu
245 250 255

Pro Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile
260 265 270

Arg Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly
275 280 285

Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr
290 295 300

Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly
305 310 315 320

Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly
325 330 335

Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr
340 345 350

027693

Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala
355 360 365

Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val
 370 375 380

Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys
 385 390 395 400

Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val
405 410 415

Ala Ser Lys

```

<210> 19
<211> 11892
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Вектор RCR - pAC3-yCD2

<400> 19
tagttattaa tagtaatcaa ttacgggtc attagttcat agccatata tggagttccg      60
cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc cccgcccatt    120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggacttcc attgacgtca    180
atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc   240
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta  300
catgacctta tgggacttgc ctacttggca gtacatctac gtattagtca tcgctattac  360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg  420
atttccaagt ctccacccca ttgacgtcaa tggagtttg tttggcacc aaaatcaacg  480
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatggcgc gtaggcgtgt  540
acggtgtggag gtctatataa gcagagctgg ttagtgaac cggcgccagt cctccgattg  600
actgagtcgc cgggtaccc gtgtatccaa taaacctct tgcaaggta tccgacttgt  660
gttctcgctg ttccctggga gggtctccctc tgagtgattg actaccgc acgggggggtc  720
tttcatttgg gggctcgcc gggatcgga gaccctgcc cagggaccac cgacccacca  780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac  840
tgattttatg cgccctcgctc ggtactagtt agctaactag ctctgtatct ggccggacccg 900
tggtggaaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt 960
cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcggtt ggactctttg 1020

```

gtgcacccccc	ctttagaggag	ggatatgtgg	ttctggtagg	agacgagaac	ctaaaacagt	1080
tcccgccctcc	gtctgaattt	ttgcttcgg	tttgggaccg	aagccgcgcc	gcgcgtcttg	1140
tctgctgcag	catcgttctg	tgttgtctct	gtctgactgt	gtttctgtat	ttgtctgaga	1200
atatggcca	gactgttacc	actcccttaa	gtttgacctt	aggtcactgg	aaagatgtcg	1260
agcggatcgc	tcacaaccag	tcggtagatg	tcaagaagag	acgttgggtt	accttctgct	1320
ctgcagaatg	gccaacccctt	aacgtcggat	ggccgcgaga	cggcaccttt	aaccgagacc	1380
tcatcaccca	ggttaagatc	aaggtctttt	cacctggccc	gcatggacac	ccagaccagg	1440
tccctacat	cgtgacctgg	gaaggcctgg	cttttacccc	ccctccctgg	gtcaagccct	1500
ttgtacaccc	taagcctccg	cctcccttcc	ctccatccgc	cccgtctctc	cccccttgaac	1560
ctcctcggttc	gaccggcct	cgatccccc	tttatccagc	cctcaactcct	tctctaggcg	1620
ccaaacctaa	acctaagtt	ctttctgaca	gtggggggcc	gctcatcgac	ctacttacag	1680
aagacccccc	gccttatagg	gaccaagac	caccccttc	cgacaggac	ggaaatggtg	1740
gagaagcgcac	ccctgcggga	gaggcaccgg	accctcccc	aatggcatct	cgcctacgtg	1800
ggagacggga	gccccctgtg	gccgactcca	ctacctcgca	ggcattcccc	ctccgcgcag	1860
gagggaaacgg	acagcttcaa	tactggccgt	tccctcttc	tgacctttac	aactggaaaa	1920
ataataaccc	ttctttttct	gaagatccag	gtaaactgac	agctctgate	gagtctgttc	1980
tcatcaccca	tcagccacc	tgggacgact	gtcagcagct	gttggggact	ctgctgaccg	2040
gagaagaaaa	acaacgggtg	ctcttagagg	ctagaaaggc	ggtgcggggc	gatgatggc	2100
gccccactca	actgccaat	gaagtcgatg	ccgctttcc	cctcgagcgc	ccagactggg	2160
attacaccac	ccaggcaggt	aggaaccacc	tagtccacta	tcgcccagttg	ctcctagcgg	2220
gtctccaaaa	cgcggcaga	agccccacca	atttggccaa	gtaaaagga	ataacacaag	2280
ggcccaatga	gtctccctcg	gccttcctag	agagacttaa	ggaagcctat	cgcaggtaca	2340
ctccttatga	ccctgaggac	ccagggcaag	aaactaatgt	gtctatgtct	ttcatttggc	2400
agtctgcccc	agacattggg	agaaaatgg	agaggtaga	agatttaaaa	aacaagacgc	2460
ttggagattt	ggttagagag	gcagaaaaga	tcttaataa	acgagaaacc	ccggaagaaa	2520
gagggaaacg	tatcaggaga	gaaacagagg	aaaaagaaga	acgcccgtagg	acagaggatg	2580
agcagaaaaga	gaaagaaaaga	gatcgtagga	gacatagaga	gatgagcaag	ctattggca	2640
ctgtcgtag	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccaactcg	2700
atcgcgacca	gtgtgcctac	tgcaaagaaa	aggggactg	ggctaaagat	tgtcccaaga	2760
aaccacgagg	acctcgggga	ccaagacccc	agacctccct	cctgacccta	gatgacttagg	2820
gaggtcaggg	tcaggagccc	ccccctgaac	ccaggataac	cctcaaagtc	ggggggcaac	2880
ccgtcacctt	cctggtagat	actggggccc	aacactccgt	gctgacccaa	aatcctggac	2940

027693

ccctaagtga taagtctgcc tgggtccaag gggctactgg agggaaagcgg tatcgctgga	3000
ccacggatcg caaagtacat ctagctaccg gtaaggcac ccactcttc ctccatgtac	3060
cagactgtcc ctatcctctg ttaggaagag atttgctgac taaaactaaaa gcccaaatcc	3120
actttgaggg atcaggagcc caggttatgg gaccaatggg gcagccccctg caagtgtga	3180
ccctaaatat agaagatgag catcggtac atgagaccc aaaagagcca gatgtttctc	3240
tagggccac atggctgtct gatttccctc aggccctggc gggaaaccggg ggcacatggac	3300
tggcagttcg ccaagctcct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420
tgttggacca gggaaatactg gtacctgccc agtccccctg gaacacgccc ctgctacc	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggtgaaaga catccacccc accgtgccc acccttacaa cctcttgagc gggctccac	3600
cgtcccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc	3660
accccacccag tcagcctc ttcgccttg agtggagaga tccagagatg ggaatctcag	3720
gacaatttgc ctggaccaga ctccccacagg gtttcaaaaa cagttccacc ctgtttgatg	3780
aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttgc atcctgtac	3840
agtacgtgta tgacttactg ctggccgcca cttctgagct agactgcca caaggtactc	3900
ggggccctgtt acaaacccta gggaaacctcg ggtatccggc ctcggccaag aaagcccaa	3960
tttgcagaa acaggtaag tatctgggtt atcttctaaa agagggtcag agatggctga	4020
ctgaggccag aaaagagact gtgatgggc agcctactcc gaagacccct cgacaactaa	4080
gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggtt gcagaaatgg	4140
cagccccctt gtaccctc accaaaacgg ggactctgtt taattgggc ccagaccaac	4200
aaaaggccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgcca	4260
atttgactaa gcccttgaa ctcttgc acgagaagca ggctacgccc aaagggtgtcc	4320
taacgcacaaa actgggaccc tggcgctggc cgggtggccta cctgtccaaa aagctagacc	4380
cagtagcagc tgggtggccc cttgcctac ggatggtagc agccattgcc gtactgacaa	4440
aggatgcagg caagctaacc atgggacagc cactagtcat tctggccccc catgcagtag	4500
aggcactagt caaacaaccc cccgaccgct ggcttccaa cgcccgatg actcactatc	4560
aggccttgcgt tttggacacg gaccgggtcc agttcgacc ggtggtagcc ctgaacccgg	4620
ctacgctgtc cccactgcct gaggaaggc tgcaacacaa ctgccttgat atcctggccg	4680
aagccccacgg aacccgaccc gacctaacgg accagccgct cccagacgccc gaccacaccc	4740
ggtacacgga tggaaagcgt ctcttacaag agggacagcg taaggcggga gctgcgggtga	4800
ccaccgagac cgaggtaatc tgggctaaag ccctgcccagc cgggacatcc gctcagcggg	4860

027693

ctgaactgat agcactcacc caggccctaa agatggcaga aggtaagaag ctaaatgttt	4920
atactgatag ccgttatgct tttgctactg cccatatcca tggagaaata tacagaaggc	4980
gtgggttgc cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac	5040
taaaagccct ctittctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg	5100
gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcgccccga aaggcagcca	5160
tcacagagac tccagacacc tctaccctcc tcataaaaaa ttcatcaccc tacacctcag	5220
aacattttca ttacacagt actgatataa aggacctaac caagttgggg gccattttatg	5280
ataaaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt	5340
ttgaattatt agactttctt catcagctga ctacacccag ctittcaaaaa atgaaggctc	5400
tccttagagag aageccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata	5460
tcactgagac ctgcaaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg	5520
gaacttagggt ccgcgggcat cggccggca ctcattggga gatcgatttc accgagataa	5580
agcccgatt gtatggctat aaatatcttc tagttttat agataccctt tctggctgga	5640
tagaaggcctt cccaaaccaag aaagaaaaccg ccaaggctgt aaccaagaag ctactagagg	5700
agatcttccc caggttcggc atgcctcagg tattggaaac tgacaatggg cctgcctcg	5760
tctccaaggt gagtcagaca gtggccgatc tttttttttt tgattggaaa ttacattgt	5820
catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt	5880
taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc	5940
tgtaccgagc ccgcaacacg cggggcccc atggcctcac cccatatgag atcttatatg	6000
gggcacccccc gcccctgtta aacttccctg acctgtacat gacaagagtt actaacagcc	6060
cctctctcca agctcaactt caggctctt acttagtcca gcacgaagtc tggagaccc	6120
tggccggcgc ctaccaagaa caactggacc gaccgggtt acctcaccct taccgagtgc	6180
gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcg tggaaaggac	6240
cttacacagt cctgctgacc accccacccg ccctcaaagt agacggcatc gcagcttgg	6300
tacacgccc ccacgtgaag gctggccgacc cgggggggtgg accatcctct agactgacat	6360
ggcgcgttca acgtctcaa aacccttca agataagatt aaccctgttga agcccttaat	6420
agtcatgggat gtcctgttag gagtagggat ggcagagac ccccatcagg tctttatgt	6480
aacctggaga gtcaccaacc ttagtactgg gcttaccgacc aatgccacct ccctcctgg	6540
aactgtacaa gatgccttcc caaaattata tttgatcta tttgtatctt tcggagagga	6600
gtgggaccc tcaagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag	6660
acagcggacc cggactttt acttttacgt gtggccctggg cataccgtaa agtcgggggtt	6720
tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccc gacaggctt	6780

027693

ctggaagccc acatcatcgt gggaccta at ctccttaag cgcgtaaca cccctggg a	6840
cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaa ag tatccaattc	6900
cttccaagg gctactcgag gggcagatg caaccctcta gtcctagaat tcactgtgc	6960
aggaaaaaag gctaactggg acgggccaa atcggtggg ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt	7080
ccccataggg cccaacccag tattaccga ccaaagactc cttccctcac caatagagat	7140
tgtaccggct ccacagccac cttagccccct caataccagt tacccccc tt ccactaccag	7200
tacaccctca acctccctca caagtccaa tggtccacag ccaccccgag gaactggaga	7260
tagactacta gctctagtc aaggagccta tcaggcgctt aacctcacca atcccgaaca	7320
gacccaagaa tggtggctgt gcttagtg tc gggacctcct tattacgaag gagtagcgg	7380
cgtgggcaact tataccaatc attccaccgc tccggccaac tgcgtggca cttcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggctatgc atggggcag tacctaaaac	7500
tcaccaggcc ttatgtaca ccacccaa ag cgcggctca ggatcctact accttgcagc	7560
acccgcgg acaatgtggg cttgcagcac tggattgact ccctgcttgc ccaccacgg	7620
gctcaatcta accacagatt atttgttatt agttaactc tggccagag taatttacca	7680
ctcccccgat tatatgtatg gtcagcttga acagcgtacc aaataaaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggaggattg cagctggaa	7800
agggacgggg accactgcct taattaaaac ccagcagttt gagcagctt atgccgctat	7860
ccagacagac ctcaacgaag tcgaaaatgc aattaccaac cttagaaaatg cactgacctc	7920
gttgtctgaa gtgtcctac agaaccgcag aggcttagat ttgttattcc taaaggaggg	7980
aggctctgc gcagccctaa aagaagaatg ttgttttat gcagaccaca cggggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca ccttaatctc	8160
caccatcatg ggacctctaa tagtactctt actgatctt ctctttggac cttgcattct	8220
caatcgattt gtccaaattt ttaaagacag gatctcagtg gtccaggctc tggtttgac	8280
tcaagcaatat caccagctaa aaccataga gtacgagcc tgaacgcgtt actggccgaa	8340
gccgcttgaa ataaggccgg tgtgcgtttg tctatatgtt atttccacc atattgccgt	8400
cttttggcaa tgtgaggggcc cgaaacctg gccctgtctt cttgacgagc attcctaggg	8460
gtctttcccc tctcgccaaa ggaatgcaag gtctgttgaat gtcgtgaag gaagcagttc	8520
ctcttggaaac ttcttgaaga caaacaaacgt ctgtacgac ctttgcagg cagcggaaacc	8580
ccccacaccttgg cgacaggtgc ctctcgccccc aaaagccacg tgtataagat acacctgcaaa	8640
aggcggcaca accccagtgac cacgttgc gttggatagt tgtggaaaga gtcaaattggc	8700

027693

tctcctcaag cgtattcaac aaggggctga aggatgccca gaaggtaccc cattgtatgg	8760
gatctgatct ggggcctcg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaaac	8820
gtcttaggccc cccgaaccac ggggacgtgg tttcccttg aaaaacacga ttataaatgg	8880
tgaccggcgg catggcctcc aagtgggatc aaaaggcat ggatatcgct tacgaggagg	8940
ccctgctgg ctacaaggag ggcggcgtgc ctatccgcg ctgtctgatc aacaacaagg	9000
acggcagtgt gctggcagg ggcacaaca tgaggttcca gaaggctcc gccaccctgc	9060
acggcgagat ctccaccctg gagaactgtg gcaggctgga gggcaaggtg tacaaggaca	9120
ccaccctgta caccaccctg tcccctgtg acatgtgtac cggcgctatc atcatgtacg	9180
gcatccctag gtgtgtgatc ggcgagaacg tgaactcaa gtccaaggc gagaagtacc	9240
tgcaaaccag gggccacgag gtggggttt ttgacgatga gaggtgtaa aagctgtga	9300
acgagttcat cgacgagagg cctcaggact ggttcgagga tatcggcgg taagcggccg	9360
cagataaaaat aaaagatttt atttagtctc cagaaaaagg gggaaatgaa agaccccacc	9420
tgtaggttt gcaagcttagc ttaagtaacg ccattttgca aggcatggaa aaatacataa	9480
ctgagaatag agaagtcag atcaaggtca ggaacagatg gaacagctga atatggcca	9540
aacaggatata ctgtggtaag cagttctgc cccggctcg ggcacaagac agatggaaaca	9600
gctgaatatg ggccaaacag gatactgtg gtaagcagtt cctgccccgg ctcagggcca	9660
agaacagatg gtccccagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg	9720
tttccaggt gccccaaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcag	9780
ttcgcttc gcttctgttc gcgcgttct gctccccgag ctcaataaaa gagcccacaa	9840
cccccactc gggcgccag tcctccgatt gactgagtcg cccgggtacc cgtgtatcca	9900
ataaaccctc ttgcagttgc atccgacttg tggtctcgct gttccttggg agggtctcct	9960
ctgagtgatt gactaccgt cagcgggggt cttcattac atgtgagcaa aaggccagca	10020
aaaggccagg aaccgtaaaa aggccgcgtt gctggcgaaa ttccataggc tccgcccccc	10080
tgacgagcat cacaaaaatc gacgctcaag tcagagggtt cgaaacccga caggactata	10140
aagataccag gcgtttcccc ctgaaagctc ctcgtgcgc tctcctgttc cgaccctgcc	10200
gcttaccgga tacctgtccg cctttctccc ttccggaaagc gtggcgcttt ctcaatgctc	10260
acgctgttagt tatctcagtt cggtgttaggt cgttcgctcc aagctgggt gtgtgcacga	10320
accccccgtt cagccgcacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc	10380
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgg	10440
gtatgttaggc ggtgctacag agttctgaa gtggggcct aactacggct acactagaag	10500
gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag	10560
ctcttgatcc ggcaaaacaaa ccaccgctgg tagcggtggg tttttgttt gcaaggcagca	10620

027693

gattacgcgc agaaaaaaag gatctaaga agatccttg atctttcta cgggtctga 10680
cgctcagtgg aacgaaaact cacgttaagg gatttggtc atgagattat caaaaaggat 10740
cttcacccatg atccctttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga 10800
gtaaacttgg tctgacagtt accaatgctt aatcagttag gcacctatct cagcgatctg 10860
tctatttcgt tcataccatag ttgcctgact ccccgctcgtagataacta cgatacggg 10920
gggcttacca tctggccccca gtgctgcaat gataccgcga gaccacgct caccggctcc 10980
agatttatca gcaataaacc agccagccgg aaggccgag cgccagaatg gtcctgcaac 11040
tttatccgcc tccatccagt ctattaattt tgccgggaa gctagagtaa gtagttcgcc 11100
agttaatagt ttgcgcaacg ttgtgccat tgctgcaggc atcgttgtgt cacgctcgcc 11160
gtttggatg gcttcattca gctccggttc ccaacgatca aggccgatgta catgatcccc 11220
catgttgtgc aaaaaagccgg ttagctcctt cggccctccg atcgttgtca gaagtaagtt 11280
ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 11340
atccgtaa tgctttctg tgactggatgta gtactcaacc aagtattctt gagaatagtg 11400
tatgcggcga ccgagttgct cttggccggc gtaaacacgg gataataccg cgccacatag 11460
cagaacttta aaagtgctca tcattggaaa acgttctcg gggcgaaaac tctcaaggat 11520
cttaccgctg ttgagatcca gttcgatgta acccactcgat gcacccaaact gatttcagc 11580
atcttttact ttaccaggcg tttctgggtg agcaaaaaca ggaaggcaaa atgcccggaaa 11640
aaagggaata agggcgacac ggaaatgttg aatactcata ctctccctt ttcaatattt 11700
ttgaagcatt tatecagggtt attgttcat gagcggatac atatttgaat gtattttggaa 11760
aaataaaacaa ataggggttc cgcgcacatt tccccggaaa gtgcacccctg acgtctaaga 11820
aaccattatt atcatgacat taacctataa aaataggcgat atcacgaggc ccttcgtct 11880
tcaagaattc at 11892

<210> 20
<211> 11892
<212> ДНК
<213> Искусственная последовательность

<220>
<223> вектор RCR - pAC3-yCD

<400> 20
tagttattaa tagtaatcaa ttacgggtc attagttcat agcccatata tggagttccg 60
cgttacataa cttacggtaa atggccggcc tggctgaccg cccaaacgacc cccgcccatt 120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggacttcc attgacgtca 180
atgggtggag tatttacggtaa aactgcccc cttggcagta catcaagtgt atcatatgcc 240
aagtacgccc cctattgacgg tcaatgacgg taaatggccc gcctggcatt atgcccagta 300

027693

catgaccta tgggacttc ctactggca gtacatctac gtattagtca tcgctattac	360
catggtgatg cggtttggc agtacatcaa tggcggtgga tagcggttg actcacgggg	420
atttccaagt ctccaccca ttgacgtcaa tggagttt tttggcacc aaaatcaacg	480
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatggcg gtaggcgtgt	540
acggtggag gtctatataa gcagagctgg ttagtgaac cggcgccagt cctccgattg	600
actgagtcgc ccgggtaccc gtgtatccaa taaaccctct tcgacttgca tccgacttgt	660
ggtctcgctg ttccctggga gggctccctc tgagtgattt actaccgc acgggggtc	720
tttcatttgg gggctcgcc gggatcgga gaccctgcc cagggaccac cgaccacca	780
ccggggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
tgatttatg cgcctgcgtc ggtactagtt agctaactag ctctgtatct ggccggacccg	900
tggtggact gacgagttcg gaacacccgg ccccaaccct gggagacgtc ccagggactt	960
cggggccgt ttttggcccg cgacctgagt caaaaatcc cgatcgcccc ggactcttg	1020
gtgcacccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt	1080
tcccgcctcc gtctgaattt ttgcttcgg tttggaccc aagccgcgc gcgcgtcttgc	1140
tctgctgcag catcgttctg tttgtctct gtctgactgt gttctgtat ttgtctgaga	1200
atatggcca gactgttacc actcccttaa gtttgaccc aggtcaactgg aaagatgtcg	1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt accttctgct	1320
ctgcagaatg gccaaccttt aacgtcgat ggcgcgaga cggcacctt aaccgagacc	1380
tcatcaccca ggttaagatc aaggctttt cacctggccc gcatggacac ccagaccagg	1440
tcccctacat cgtgacctgg gaaggcttgg ctttgaccc ccctccctgg gtcaagccct	1500
ttgtacaccc taaggctccg cctcccttcc ctccatccgc cccgtctctc ccccttgaac	1560
ctcctcggttc gaccccgctt cgatccccc tttatccgc cctcactcct tctctaggcg	1620
ccaaacctaa acctcaagtt ctttgacca gtggggggcc gctcatcgac ctacttacag	1680
aagacccccc gccttatagg gaccaagac cacccccc cgcacaggac gggaaatggtg	1740
gagaagcgac ccctgcggga gaggcacccg acccctcccc aatggcatct cgcctacgtg	1800
ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag	1860
gaggaaacgg acagcttcaa tactggccgt tctctcttc tgaccttac aactggaaaa	1920
ataataaccc ttcttttctt gaagatccag gtaaaactgac agctctgatc gagtctgttc	1980
tcatcaccca tcagccacc tgggacact gtcagcagct gttggggact ctgctgaccg	2040
gagaagaaaa acaacgggtg ctcttagagg ctagaaagcc ggtgcggggc gatgtgggc	2100
gccccactca actgccaat gaagtcgtat ccgcctttcc cctcgagcgc ccagactggg	2160
attacaccac ccaggcaggt aggaaccacc tagtccacta tcgcccgttg ctccctagcgg	2220

027693

gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag	2280
ggcccaatga gtctccctcg gccttcctag agagacttaa ggaaggcctat cgcatgtaca	2340
ctccttatga ccctgaggac ccaggcagaag aaactaatgt gtctatgtct ttcatttggc	2400
agtctgcccc agacatggg agaaaatgg tagaggtttaga agatttaaa aacaagacgc	2460
ttggagatgg ggttagagag gcagaaaaga tctttaataa acgagaaacc ccgaaagaaa	2520
gagaggaacg tattcaggaga gaaacagagg aaaaagaaga acgcccgtagg acagaggatg	2580
agcagaaaaga gaaagaaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca	2640
ctgtcgtagg tggacgaaaa caggatagac agggaggaga acgaaggagg tcccaactcg	2700
atcgcgacca gtgtgcctac tgcaagaaa agggcactg ggctaaagat tgcccaaga	2760
aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgactagg	2820
gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc gggggcaac	2880
ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac	2940
ccctaagtga taagtctgcc tgggtccaag gggctactgg agaaaagcgg tattcgctgga	3000
ccacggatcg caaagtacat cttagtaccg gtaaggtcac ccactcttc ctccatgtac	3060
cagactgtcc ctatcctctg ttaggaagag atttgctgac taaaactaaaa gccaaatcc	3120
actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga	3180
ccctaaatat agaagatgag catcggtac atgagaccc aaaagagcca gatgttctc	3240
tagggtccac atggctgtct gatttcctc aggctgggc gggaaaccggg ggcattggac	3300
tggcagttcg ccaagctcct ctgatcatac ctctqaaagc aacctctacc cccgtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420
tgttggacca gggataactg gtaccctgcc agtccccctg gaacacgccc ctgctaccgg	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggttggaga catccacccc accgtgcccc acccttacaa cctcttgagc gggctccac	3600
cgtccccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc	3660
accccacccag tcagcccttc ttcccttttgg agtggagaga tccagagatg ggaatctcag	3720
gacaatttgc ctggaccaga ctcccacagg gtttcaaaaa cagttccacc ctgtttgtatg	3780
aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atccctgctac	3840
agtacgttgg tgacttactg ctggccgcca cttctgagct agactgccaa caaggactc	3900
ggccctgtt acaaacccta gggAACCTCG ggtatcgggc ctccggccaag aaagccaaaa	3960
tttggcagaa acaggtcaag tatctgggtt atcttctaaa agagggttag agatggctga	4020
ctgaggccag aaaagagact gtgtatggggc agccctactcc gaagaccctt cgacaactaa	4080
gggagttcct agggacggca ggcttctgtc gcctctggat ccctgggtt gcagaaatgg	4140

027693

cagccccctt	gtaccctctc	acccaaacgg	ggactctgtt	taattggggc	ccagaccaac	4200
aaaaggccta	tcaagaaaatc	aagcaagctc	ttcttaactgc	cccagccctg	gggttgccag	4260
atttgactaa	gcccttgaa	ctctttgtcg	acgagaagca	gggtacgccc	aaaggtgtcc	4320
taacgcaaaa	actgggacct	tggcgtcggc	cgggtggccta	cctgtccaaa	aagctagacc	4380
cagtagcagc	tgggtggccc	ccttgctac	ggatggtagc	agccattgcc	gtactgacaa	4440
aggatgcagg	caagctaaccc	atggacagc	cactagtcat	tctggccccc	catgcagtag	4500
aggcactagt	caaacaaccc	cccgaccgct	ggcttccaa	cgcggatg	actcactatc	4560
aggcctgct	tttggacacg	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaacccgg	4620
ctacgctgct	cccaactgcct	gaggaagggc	tgcaacacaa	ctgccttgat	atcctggccg	4680
aagcccacgg	aacccgaccc	gacctaacgg	accagccgct	cccagacgccc	gaccacaccc	4740
ggtacacgga	tggaagcagt	ctcttacaag	aggacacgac	taaggcggga	gctgcgggtga	4800
ccacccgagac	cgaggtataatc	tgggctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
ctgaactgat	agcacttcacc	caggccctaa	agatggcaga	aggtagaag	ctaaatgttt	4920
atactgatag	ccgttatgct	tttgcgtactg	cccatatatcc	tggagaataa	tacagaaggc	4980
gtgggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
taaaaggcct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggctgacca	agcggcccgaa	aaggcagcca	5160
tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcaccc	tacacctcag	5220
aacatttca	ttacacagt	actgatataa	aggacctaac	caagttgggg	gccattttatc	5280
ataaaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggctc	5400
tccttagagag	aagccacagt	ccctactaca	tgctgaaccg	ggatcgaaca	ctcaaaaata	5460
tcactgagac	ctgcaaagct	tgtgcacaag	tcaacgccc	caagtctgcc	gttaaacagg	5520
gaacttagggt	ccgcggcat	cggccggca	ctcattggga	gatcgatttc	accgagataa	5580
agcccgatt	gtatggctat	aaatatcttc	tagttttat	agataccctt	tctggctgga	5640
tagaaggcctt	cccaaccaag	aaagaaaaccg	ccaaaggctgt	aaccaagaag	ctactagagg	5700
agatcttccc	cagggtcggc	atgcctcagg	tattggaaac	tgacaatggg	cctgccttcg	5760
tctccaaggt	gagtcagaca	gtggccgatc	tgttggggat	tgattggaaa	ttacattgtg	5820
catacagacc	ccaaagctca	ggccaggtag	aaagaatgaa	tagaaccatc	aggagactt	5880
taactaaatt	aacgcttgca	actggctcta	gagactgggt	gctctactc	cccttagccc	5940
tgtaccgagc	ccgcaacacg	ccggcccccc	atggcctcac	cccatatgag	atcttataatg	6000
gggcacccccc	gccccttgta	aactccctg	accctgacat	gacaagagtt	actaacagcc	6060

027693

cctctctcca agtcactta caggctctc acttagtcca gcacgaagtc tggagacctc	6120
tggcgccagc ctaccaagaa caactggacc gaccgggtt acctcaccct taccgagtcg	6180
gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac	6240
cttacacagt cctgctgacc acccccacccg ccctcaaagt agacggcatc gcagcttgg	6300
tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat	6360
ggcgcgttca acgctctcaa aacccctca agataagatt aacccgtgga agcccttaat	6420
agtcatggga gtcctgttag gagtagggat ggcagagagc cccatcagg tcttaatgt	6480
aacctggaga gtcaccaacc ttagtactgg gcgtaccgac aatgccacct ccctcctgg	6540
aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga	6600
gtgggacccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag	6660
acagcggacc cggactttt acttttacgt gtgccttggg cataccgtaa agtcggggtg	6720
tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccc gacaggctt	6780
ctggaagccc acatcatcgt gggacctaatt ctccttaag cgccgttaaca ccccttggg	6840
cacgggatgc tctaaagttt cctgtggccc ctgtacgac ctctccaaag tatccattc	6900
cttccaaggg gctactcgag gggcagatg caaccctcta gtcctagaat tcactgatgc	6960
aggaaaaaaag gctaactggg acggggccaa atcgtggggc ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggc cttatgtgg gaccccgagt	7080
ccccataggg cccaacccag tattaccga ccaaagactc cttccctcac caatagagat	7140
tgtaccggct ccacaggccac ctggccctt caataccagt tacccctt ccactaccag	7200
tacaccctca acctccctta caagtccaaag tgtcccacag ccaccccaag gaactggaga	7260
tagactacta gctcttagtca aaggagccta tcaggcgctt aacccatcca atcccgacaa	7320
gacccaagaa tggggctgt gcttagtgc gggacccctt tattacgaag gagtagcggt	7380
cgtgggact tataccaatc attccaccgc tccggccaac tgtacggca cttcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcttatgc atggggcag tacctaaaac	7500
tcaccaggcc ttatgtaaaca ccacccaaag cgccggctca ggatcctact accttgcagc	7560
acccggccga acaatgtggg cttgcagcac tggattgact ccctgctgt ccaccacgg	7620
gctcaatcta accacagatt attgtgtatt agtgaactc tggccagag taatttacca	7680
ctccccctat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattt cagctggaaat	7800
agggacgggg accactgcct taattaaaac ccagcgttt gagcagcttc atgccgttat	7860
ccagacacac ctcaacgaag tcgaaaagtc aattaccaac ctgaaaagt cactgaccc	7920
gttgtctgaa gtagtcctac agaaccgcag aggcttagat ttgctattcc taaaggagg	7980

027693

aggtctctgc gcagccctaa aagaagaatg ttgttttat gcagaccaca cggggctagt 8040
gagagacagc atggccaaat taagagaaag gcttaatcg agacaaaaac tatttgagac 8100
aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca ccttaatctc 8160
caccatcatg ggacacctaa tagtacttta actgatctt ctcttggac ctgcattct 8220
caatcgattt gtccaaattt taaaagacag gatctcagtg gtccaggctc tggtttgc 8280
tcagcaatat caccagctaa aacccataga gtacgagcca tgaacgcgtt actggccgaa 8340
gccgcttgga ataaggccgg tgtgcgtttg tctatatgtt atttccacc atattgccgt 8400
cttttgcaa tgtgagggcc cgaaaacctg gccctgttctt cttgacgagc attcctaggg 8460
gtctttcccc tctcgccaaa ggaatgcaag gtctgttgaat tgtcgtgaag gaagcagttc 8520
ctctggaaagc ttcttgaaga caaacaacgt ctgtagcgac ccttgcagg cagcggaaacc 8580
ccccacacctgg cgacagggtgc ctctcgccccc aaaagccacg tggataagat acacctgcaa 8640
aggcggcaca accccagtgc cacgttgtga gttggatagt tggaaaga gtcaaattggc 8700
tctcctcaag cgtattcaac aagggtgtga aggtgccca gaaggatccc cattgtatgg 8760
gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg taaaaaaac 8820
gtcttaggccc cccgaaccac ggggacgtgg tttcccttg aaaaacacga ttataaatgg 8880
tgacaggggg aatggcaagc aagtgggatc agaagggtat ggacattgcc tatgaggagg 8940
cgcccttagg ttacaaagag ggtgggttgc ctattggcgg atgtcttatac aataacaaag 9000
acggaagtgt tctcggtcgt ggtcacaaca tggatattca aaaggatcc gccacactac 9060
atggtgagat ctccactttg gaaaactgtg ggagattaga gggcaaagtg tacaagata 9120
ccactttgtt tacgacgctg tctccatgctg acatgtgtac aggtgccatc atcatgtatg 9180
gtattccacg ctgtgtgtc ggtgagaacg ttaatttcaa aagtaaggc gagaatattt 9240
tacaaactag agtcacgag gttttgttg ttgacgatga gaggtgtaaa aagatcatga 9300
aacaatttat cgtgaaaga cctcaggatt ggttgaaga tattggtgag taggcggccg 9360
cagataaaat aaaagatttt attagtctc cagaaaaagg gggaaatgaa agacccacc 9420
tgttaggttt gcaagctagc ttaagtaacg ccattttgca aggcatggaa aaatacataa 9480
ctgagaataag agaagttcag atcaaggatca ggaacagatg gaacagctga atatggccca 9540
aacaggatat ctgtggtaag cagttcctgc cccggctcag ggccaagaac agatggaaca 9600
gctgaatatg ggccaaacag gatatctgt gtaaggcgtt cctgccccgg ctcaggccca 9660
agaacagatg gtcccaagat gcggtccagc cctcagcagt ttctagagaa ccatcagatg 9720
tttccaggg gccccaaagga cctgaaatga ccctgtgcct tatttgaact aaccaatcg 9780
ttcgcttc gcttctgttc ggcgcgttct gctcccccgg ctcaataaaa gagcccacaa 9840
cccctcactc ggggcggccag tcctccgatt gactgagtcg cccgggtacc cgtgtatcca 9900

027693

ataaacccctc ttgcagttgc atccgacttgc tggctcgct gttccttggg agggtctccct	9960
ctgagtgatt gactaccgt cagccccgt ctccattac atgtgagcaa aaggccagca	10020
aaaggccagg aaccgtaaaa aggcccggtt gctggcggtt ttccataggc tccggccccc	10080
tgacgagcat cacaaaaatc gacgctcaag tcagagggtgg cgaaacccga caggactata	10140
aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgcc	10200
gcttaccgga tacctgtccg ccttctccc ttccggaaagc gtggcgctt ctcaatgctc	10260
acgctgttagg tatctcagtt cggtgttaggt cgttcgctcc aagctggct gtgtgcacga	10320
accccccgtt cagcccgacc gctgcgcctt atccgtaac tatcgtctt agtccaaccc	10380
ggtaagacac gacttatcgc cactggcagc aacctgtgtt aacaggatta gcagagcgag	10440
gtatgttaggc ggtgctacag agtttcttga gtggtggct aactacggct acactagaag	10500
gacagtattt ggtatctgcg ctctgtgaa gccagttacc ttccggaaaaa gagttggtag	10560
ctcttgatcc ggcaaaacaaa ccaccgctgg tagcggtggt tttttgttt gcaaggcagca	10620
gattacgcgc agaaaaaaag gatctaaga agatccttgc atctttctta cggggctgtga	10680
cgcctcagttt aacgaaaact cacgttaagg gatttggtc atgagattat caaaaaggat	10740
cttcacccat tagttttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatgaa	10800
gtaaacttgg tctgacagtt accaatgtt aatcagttag gcacctatct cagcgatctg	10860
tctatttctgt tcatccatag ttgcctgact ccccgctgtg tagataacta cgatacggga	10920
gggcttacca tctggcccca gtgctgaat gataccgcga gaccacgcgtt caccggctcc	10980
agatttatca gcaataaaacc agccagccgg aaggcccgag cgccagaatgt gtcctgcaac	11040
tttatccgc tccatccagt ctattaattt tgccggaa gctagagtaa gtagttcgcc	11100
agttaatagt ttgcgcaacg ttgttgcctt tgctgcaggg atcgtgggtt cacgctcgct	11160
gtttggtagt gcttcattca gctccggttc ccaacgatca aggccgatgtt catgatcccc	11220
catgttgtgc aaaaaagccgg tttagtccctt cggccctccg atcgttgtca gaagtaatgtt	11280
ggccgcagtg ttatcactca tggttatggc agcaactgcatt aattctctta ctgtcatgcc	11340
atccgttaaga tgctttctgt tgactggta gtactcaacc aagtatttctt gagaatagtg	11400
tatgcggcga ccgagttgtt cttggccggc gtcaacacgg gataataccg cgccacatacg	11460
cagaacttta aaagtgcata tcattggaaa acgttcttcg gggcgaaaaac tctcaaggat	11520
cttaccgcgtt tttagatcca gttcgatgtt acccaactcggtt gcacccaaact gatcttcagg	11580
atcttttact ttaccacgcg tttctgggtt agcaaaaaaca ggaaggcaaa atgcccggaaa	11640
aaagggaaata agggcgacac ggaaatgttg aataactcata ctctccctt ttcaatattt	11700
ttgaagcatt tattcagggtt attgtctcat gagccgatgtt atatttgaat gtatttttagaa	11760
aaataaaacaa ataggggttc cgccacatt tcccccggaaaaa gtgccacactg acgtctaaga	11820

027693

aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc ccttcgtct	11880
tcaagaattc at	11892
<210> 21	
<211> 12007	
<212> ДНК	
<213> Искусственная последовательность	
<220>	
<223> Вектор RCR - pACE-CD	
<400> 21	
tagttattaa tagtaatcaa ttacgggtc attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt	120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggacttcc attgacgtca	180
atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc	240
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	300
catgaccta tgggacttcc ctacttggca gtacatctac gtattagtca tcgctattac	360
catggtgatg cggtttggc agtacatcaa tggcgtgga tagcggttt actcacgggg	420
atttccaagt ctccacccca ttgacgtcaa tggagtttgc tttggcacc aaaatcaacg	480
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatggcg gtaggcgtgt	540
acggtgaggag gtctatataa gcagagctgg ttttgtaac cggcgccagt cctccgatttgc	600
actgagtcgc cgggtaccc gtgtatccaa taaaccctct tgcaatgtca tcggacttgt	660
ggtctcgctg ttccctggga gggctccctc tgagtgatttgc actacccgtc agcgggggtc	720
tttcatttgg gggctcgcc gggatcgaaa gaccctgc cagggaccac cgacccacca	780
ccggggagta agctggccag caacttatct gtgtctgtcc gattgtcttag tgcgttatgac	840
tgatttatg cgcctcgctc ggtactagtt agctaacttag ctctgtatct ggcggacccg	900
tggttggaaact gacgagttcg gaacacccgg ccgcaaccctt gggagacgtc ccagggactt	960
cggggccgt ttttggcc cgacctgagt ccaaaaatcc cgatcgaaaa ggactcttttgc	1020
gtgcacccca ctttagaggag gatatgtgg ttctggtagg agacgagaac ctaaaacagt	1080
tcccgccctcc gtctgaattt ttgtttcggtt tttgggaccg aagccgcgc ggcgtcttg	1140
tctgtcgacg catcggtctcg tggttctct gtctgactgt gtttctgtat ttgtctgaga	1200
atatggccca gactgttacc actccctaa gtttgacctt aggtcaactgg aaagatgtcg	1260
agcggatcgcc tcacaaccag tcggtagatg tcaagaagag acgtgggtt accttctgt	1320
ctgcagaatg gccaacccaa aacgtcgatgg ccgcgcgaga cggcacccctt aaccgagacc	1380
tcatcacccaa ggttaagatc aaggtctttt cacctggccc gcatggacac ccagaccagg	1440
tccctacat cgtgacctgg gaagccttgg cttttgaccc ccctccctgg gtcaagccct	1500

027693

ttgtacaccc taagcctccg ctccttcc tcctatccgc cccgtctctc ccccttgaac	1560
ctcctcggttc gaccccgct cgatcctccc ttatccagc ctcactctt tctctaggcg	1620
ccaaacctaa acctaagtt ctcttgaca gtggggggcc gctcatcgac ctacttacag	1680
aagacccccc gccttatagg gacccaagac cacccttcc cgacaggac ggaatggtg	1740
gagaagcgac ccctgcggga gaggcacgg accccctccc aatggcatct cgctacgtg	1800
ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag	1860
gaggaaacgg acagctcaa tactggcgct tctcttcc tgaccttac aactggaaaa	1920
ataataaccc ttcttttct gaagatccag gtaaactgac agtctgatc gagtctgttc	1980
tcatcaccca tcagccacc tgggacgact gtcaagact gttggggact ctgctgaccg	2040
gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgtgggc	2100
gccccactca actgccaat gaagtcgatg ccgtttcc ctcgagcgc ccagactggg	2160
attacaccac ccaggcaggt aggaaccacc tagtccacta tcgcccagtgc ctcctagcgg	2220
gtctccaaaa cgcggcaga agccccacca atttggccaa ggtaaaagga ataacacaag	2280
ggcccaatga gtctccctcg gccttcctag agagacttaa ggaagcctat cgcaaggata	2340
ctccttatga ccctgaggac ccaggcaag aaactaatgt gtctatgtct ttcatggc	2400
agtctgcccc agacattggg agaaagttag agaggttaga agattaaaa aacaagacgc	2460
ttggagatt ggttagagag gcagaaaaga tcttaataa acgagaaacc ccggaagaaaa	2520
gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgcccgtagg acagaggatg	2580
agcagaaaaga gaaagaaaga gatcgttaga gacatagaga gatgagcaag ctattggca	2640
ctgtcgtag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg	2700
atcgcgacca gtgtgcctac tgcaaaagaaa aggggcactg ggctaaagat tgtcccaaga	2760
aaccacgagg acctcgggga ccaagacccc agacctccct cctgacccta gatgacttagg	2820
gaggtcgagg tcaggagccc cccctgaaac ccaggataac cctcaaagtc gggggcaac	2880
ccgtcacctt cctggtagat actggggccc aacactccgt gctgacccaa aatcctggac	2940
ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaagcgg tatkctgg	3000
ccacggatcg caaagtacat ctatcgatcg gtaaggtcac ccactcttcc tcctatgtac	3060
cagactgtcc ctatctctg ttaggaagag atttgcgtac taaaactaaaa gccaaatcc	3120
actttgaggg atcaggagcc caggttatgg gaccaatggg gcagccctg caagtgttga	3180
ccctaaatat agaagatgag catcggtac atgagaccc aaaaagagcca gatgttctc	3240
tagggtcac atggctgtct gatttcctc aggctgggc gaaacccggg ggcacggac	3300
tggcagttcg ccaagctct ctgatcatac ctctgaaagc aacctctacc cccgtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420

027693

tgttggacca gggataactg gtaccctgcc agtccccctg gaacacgccc ctgctacccg	3480
ttaagaaacc agggactaat gattataggc ctgtccagga tctgagagaa gtcaacaagc	3540
gggttggaga catccacccc accgtgccca acccttacaa cctcttgagc gggctccac	3600
cgtccccacca gtggtacact gtgcttgatt taaaggatgc cttttctgc ctgagactcc	3660
accccaccag tcagcctctc ttccgccttg agtggagaga tccagagatg ggaatcttag	3720
gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtccacc ctgtttgatg	3780
aggcactgca cagagaccta gcagacttcc ggatccagca cccagacttg atcctgctac	3840
agtacgtgga tgacttactg ctggccgcca ctctcgatc agactgcca caaggtactc	3900
gggcctgtt acaaacccta gggAACCTCG ggtatcgggc ctccggcaag aaagcccaa	3960
tttgcagaa acaggtcaag tatctgggtt atcttctaaa agagggtcag agatggctga	4020
ctgaggccag aaaagagact gtgatggggc agcctactcc gaagaccctt cgacaactaa	4080
gggagttcct agggacggca ggctctgtc gcctctggat ccctgggtt gcagaaatgg	4140
cagcccccgtt gtaccctctc accaaaacgg ggactctgtt taattggggc ccagaccaac	4200
aaaaggcccta tcaagaaatc aagcaagctc ttctaactgc cccagccctg gggttgccc	4260
atttgactaa gcccttgaa ctctttgtcg acgagaagca gggctacgcc aaaggtgtcc	4320
taacgcaaaa actgggacact tggcgctggc cgggtggcta cctgtccaaa aagcttagacc	4380
cagtagcagc tgggtggccc cttgtctac ggatggtagc agccattgcc gtactgacaa	4440
aggatgcagg caagctaacc atggacagc cactagtcat tctggccccc catgcagtag	4500
aggcactagt caaacaaaccc cccgaccgct ggcttccaa cggccggatg actcactatc	4560
aggccttgct tttggacacg gaccgggtcc agttcgacc ggtggtagcc ctgaaccgg	4620
ctacgctgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg	4680
aagcccacgg aacccgaccc gacctaacgg accagccgct cccagacgcc gaccacacct	4740
ggtacacggg tggaaacagt ctcttacaag agggacagcg taaggcggga gctgcgggtga	4800
ccaccggagac cgaggtataat tgggttaaag ccctgccagc cgggacatcc gctcagcggg	4860
ctgaactgat agcactcacc caggccctaa agatggcaga aggttggaaat tacagaaggc	4920
atactgatag ccgttatgct tttgtactg cccatatcca tggagaaata tacagaaggc	4980
gtgggttgct cacatcagaa ggcaaaagaga tcaaaaataa agacgagatc ttggccctac	5040
taaaaggccct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg	5100
gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggccgaa aaggcagccca	5160
tcacagagac tccagacacc tctaccctcc tcatagaaaa ttcatcaccc tacaccttag	5220
aacatttca ttacacagtg actgtatataa aggacctaac caagttgggg gccattttatg	5280
ataaaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt	5340

027693

ttgaattatt agacttctt catcagctga ctcacctcg cttctcaaaa atgaaggctc	5400
tccttagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata	5460
tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg	5520
gaacttagggc cccgggcat cggccggca ctcattggga gatcgatttc accgagataa	5580
agccccgatt gtatggctat aaatatcttc tagttttat agatacctt tctggctgga	5640
tagaaggc tt cccaaccaag aaagaacccg ccaaggtcgt aaccaagaag ctactagagg	5700
agatcttccc caggttcggc atgectcagg tattggAAC tgacaatggg cctgccttcg	5760
tctccaaggt gagtcagaca gtggccgatc tggtgggat tgattggaaa ttacattgtg	5820
catacagacc ccaaagctca ggcaggtag aaagaatgaa tagaaccatc aaggagactt	5880
taactaaatt aacgcttgca actggctcta gagactgggt gctctactc cccttagccc	5940
tgtaccgagc cgcacacacg cggggcccccc atggcctcac cccatatgag atcttatatg	6000
gggcacccccc gccccttta aactccctg accctgacat gacaagagtt actaacagcc	6060
cctctctcca agtcactta caggctctct acttagtcca gcacgaagtc tggagacctc	6120
tggcggcagc ctaccaagaa caactggacc gaccgggtt acctcaccct taccgagtcg	6180
gcgcacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac	6240
cttacacagt cctgctgacc acccccacccg ccctcaaagt agacggcatc gcagcttgg	6300
tacacgccc ccacgtgaag gctggcggacc ccgggggtgg accatcctct agactgacat	6360
ggcgcgttca acgctctcaa aacccctca agataagatt aaccctgga agcccttaat	6420
agtcatggg gtcctgttag gagtagggat ggcagagac ccccatcagg tcttaatgt	6480
aacctggaga gtcaccaacc ttagtactgg gcgtaccgac aatgccaccc ctccctggg	6540
aactgtacaa gatgccttcc caaaattata tttgtatcta tggatctgg tcggagagga	6600
gtgggaccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag	6660
acagcggacc cggactttt actttacgt gtggccctggg cataccgtaa agtcggggtg	6720
tggggacca ggagagggtt actgtggtaa atgggggtgt gaaaccaccc gacaggctta	6780
ctggaaagccc acatcatcgt gggaccta at ccccttaag cgcggtaaca cccctggg	6840
cacgggatgc tctaaagttt cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaaagg gctactcgag gggcagatg caaccctcta gtcctagaat tcaactgatgc	6960
aggaaaaaaag gctaactggg acggggccaa atcgtggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggac cttaatgtgg gaccccgagt	7080
ccccataggg cccaaacccag tattaccctga ccaaagactc ctttcctcac caatagagat	7140
tgtaccggct ccacagccac ctggccctt caataccagt taccggccctt ccactaccag	7200
tacaccctca acctccctca caagtccaa tgcccacag ccacccctcag gaactggaga	7260

027693

tagactacta gctctagtca aaggagccta tcagggcgctt aacctcacca atcccgacaa	7320
gacccaagaa tgttggctgt gcttagtgta cggacccctt tattacgaag gagtagcggt	7380
cgtgggcact tataccaaatc attccaccgc tccggccaac tgtacggca cttcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcttatgc atggggcgag tacctaaaac	7500
tcaccagggc ttatgttaaca ccacccaaag cgccggctca ggatcctact accttgcagc	7560
acccggccga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacggt	7620
gctcaatcta accacagatt attgtgtatt agttgaactc tggcccagag taatttacca	7680
ctcccccgat tatatgtatg gtcagcttga acagcgtacc aaatataaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattt cagctggat	7800
agggacgggg accactgcct taataaaaac ccacgagttt gacgagcttc atgcccstat	7860
ccagacagac ctcaacgaag tcgaaaagtca aattaccaac ctagaaaagt cactgaccc	7920
gttgtctgaa gtagtcctac agaaccgcag aggccatagtttgcattcc taaaggaggg	7980
aggctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tattttagac	8100
aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca ccttaatctc	8160
caccatcatg ggacctctaa tagtactctt actgatctt ctctttggac cttgcattct	8220
caatcgatta gtccaatttgc ttaaagacag gatatacgatg gtccaggctc tagtttgcac	8280
tcaacaatat caccagctga agcctataga gtacgagcca tgacgtacgt tactggccg	8340
agccgcttgg aataaggccg gtgtgcgtt gtctatatgt tattttccac catattgcgg	8400
tcttttggca atgtgaggcgc ccggaaacctt ggcctgtct tcttgacgag cattcctagg	8460
ggtctttccc ctctcgccaa aggaatgcaa ggtctgttga atgtcgtgaa ggaagcagtt	8520
cctctggaaag cttcttgaag acaaacaacg tctgtacgtc ccctttgcag gcagcggAAC	8580
cccccacctg gcgacaggtg cctctgcggc caaaagccac gtgtataaga tacacctgca	8640
aaggcggcac aaccccaatgtt ccacgttgc agttggatag ttgtggaaag agtcaaattgg	8700
ctctcctcaa gcgtattcaa caagggctg aaggatgccc agaaggatacc ccattgtatg	8760
ggatctgtatc tggggcctcg gtgcacatgc ttatcatgtt tttatgtcgag gttaaaaaaaa	8820
cgtcttaggcc ccccgaacca cggggacgtt gttttccctt gaaaaacacg ataataaccat	8880
ggtgacaggg ggaatggcaa gcaagtggta tcagaagggtt atggacatttgc cctatgagga	8940
ggggggctta gtttacaaag aggggtgggtt tcttattggc ggtatgtctta tcaataacaa	9000
agacggaaatgtt gttctcggtc gtggtcacaa catgagattt caaaagggtt ccggccacact	9060
acatggtgatc atctccactt tggaaaactg tggagatattt gaggcggaaag tgtacaaaga	9120
taccacttttgc tatacgacgc tgcgtccatg cgacatgtgtt acaggtgccatc tcatcatgtt	9180

027693

tggtattcca cgctgtgttgcggtgagaa cgttaatttc aaaagtaagg gcgagaaata	9240
tttacaaact agaggtcacg aggttggtgt tggacat gagagggtgtaaaagatcat	9300
gaaacaattt atcgatgaaa gacccagga ttggttgaa gatattggtagtggcggc	9360
cgcgccatag ataaaataaa agattttttt tagtctccag aaaaaggggg gaatgaaaga	9420
ccccacctgt aggtttggca agctagctt agtaacgcca ttttgcaggcatggaaaa	9480
tacataactg agaatagaga agttcagatc aaggtcagga acagatggaa cagctgaata	9540
tgggccaac aggatatctg tggttaaggat ttcctgcccc ggctcaggc caagaacaga	9600
tggAACAGCT gaatatggc caaacaggat atctgtggta agcagttcct gccccggctc	9660
aggggccaaga acagatggc cccagatgcg gtccagccct cagcagttc tagagaacca	9720
tcagatgttt ccagggtgcc ccaaggaccc gaaatgaccc tggccttgc ttAAACTAAC	9780
caatcagttc gcttctcgct tctgttcgcg cgcttctgc ccccgagctc aataaaagag	9840
cccacaaccc ctcaactcggg ggcgcagtcc tccgattgac tgagtgcggg gggtaaccgt	9900
gtatccaata aacccttgc cagttgcac tgcattgtgg tctcgctgtt cttggggagg	9960
gtctccctcg agtgatgtac taccgtcag cgggggttt tcatttgggg gctcgccgg	10020
gatcgggaga cccctgccc gggaccaccg acccaccacc gggaggtaag ctggctgcct	10080
cgcgcgttcc ggtgtatgcg gtgaaaacct ctgacatgtg agcaaaggc cagcaaagg	10140
ccaggaaccg taaaaaggcc gcgttgcgg cgttttcca taggctccgc cccctgacg	10200
agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaaagat	10260
accaggcggtt tccccctggaa agctccctcg tgcgtctcc tggccgacc ctggcgctta	10320
ccggataacct gtccgcctt ctccttcgg gaagcgtggc gctttctcaa tgctcacgc	10380
gtaggtatct cagttcggtg tagtcgttc gctccaagct gggctgtgtg cacgaacccc	10440
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa	10500
gacacgactt atcgccactg gcagcggcc ctggtaacag gattacgaga gcgaggatg	10560
taggcgggtgc tacagagttc ttgaagtggt ggctctaacta cggctacact agaaggacag	10620
tatTTGGTAT ctgcgtctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt	10680
gatccggcaa acaaaccacc gctggtagcg gtgggggggg tggcaag cagcagatta	10740
cgcgcagaaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgc	10800
agtggAACGA aaactcacgt taaggattt tggcatgat attatcaaaa aggttctca	10860
cctagatct ttAAATTAA aaatgaagtt taaatcaat ctAAAGTATA tatgagtaaa	10920
ttcggctctga cagttaccaa tgcttaatca gtggggcacc tatctcagcg atctgtctat	10980
ttcgttcatc catagttgcc tgactccccg tcgtgttagat aactacgata cgggaggc	11040
taccatctgg ccccagtgc gcaatgatac cgcgagaccc acgttcaccc gctccagatt	11100

tatcagcaat aaaccagcca	gccggaaggg	ccgagcgcag	aagtggtcct	gcaacttat	11160
ccgcctccat ccagtctatt	aatttgc	ggaaagctag	agtaagtagt	tcgcccgtta	11220
atagtttgcg	caacgttgtt	gccattgctg	caggcatcgt	ggtgtcacgc	11280
gtatggcttc	attcagctcc	ggttccaaac	gatcaaggcg	atttacatga	11340
tgtgaaaaaa	agcggttagc	tccttcggtc	ctccgatcgt	tgtcagaagt	11400
cagtgttatac	actcatggtt	atggcagcac	tgcatatc	tcttactgtc	11460
taagatgc	ttctgtgact	ggtgagact	caaccaagtc	attctgagaa	11520
ggcgaccgag	ttgctcttgc	ccggcgtcaa	cacggataa	taccgcgcca	11580
ctttaaaagt	gctcatcatt	ggaaaacgtt	cttcggggcg	aaaactctca	11640
cgctgttgag	atccagttcg	atgtaaacca	ctcgtgcacc	caactgatct	11700
ttactttcac	cagcgttict	gggtgagcaa	aaacaggaag	gcaaaatgcc	11760
gaataaggc	gacacgaaa	tgttgaatac	tcatactctt	ccttttcaa	11820
gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	11880
aacaatagg	ggttccgcgc	acattcccc	gaaaagtgcc	acctgacgtc	11940
ttatttatcat	gacattaacc	tataaaaata	ggcgtatcac	gaggccctt	12000
aattcat					12007

<210> 22

<211> 11893

<212> ДНК

<213> Искусственная последовательность

<220>

<223> вектор RCR - pAC3-yCD2

<400> 22

tagttattaa	tagtaatcaa	ttacgggtc	attagttcat	agccatata	tggagttccg	60
cgttacataa	cttacgtaa	atggccgc	tggctgaccg	cccaacgacc	ccgcgcatt	120
gacgtcaata	atgacgtat	ttcccatagt	aacccaata	gggactttcc	attgacgtca	180
atgggtggag	tatCACG	aaactgccc	cttggcagta	catcaagtgt	atcatatgcc	240
aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	atgcccagta	300
catgaccta	tgggactt	ctactggca	gtacatctac	gtttagtca	tcgcttattac	360
catggtgatg	cgttttggc	agtacatcaa	tggcgtgga	tagcggttt	actcacgggg	420
atttccaagt	ctccacccca	ttgacgtcaa	tggagttt	tttggcacc	aaaatcaacg	480
ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatggcg	gtaggcgtgt	540
acggtgggag	gtctatataa	gcagagctgg	tttagtgaac	cggcgccagt	cctccgattg	600
actgagtcgc	ccgggtaccc	gtgtatccaa	taaacccctt	tgcagttca	tccgacttgt	660

027693

ggtctcgctg ttccttgga gggctccctc tgagtgattg actaccgc tc	720
tttcatttgg gggctcgccc gggatcgaaa gacccctgcc cagggaccac cgacccacca	780
ccggggaggt a gctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
t gat tttatg cgcctgcgtc ggtacttagtt agctaactag ctctgtatct ggccggacccg	900
tggtggaaact gacgagttcg gaacacccgg ccgcaaccct gggagacgtc ccagggactt	960
cggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgttt ggactctttg	1020
gtgcacccccc ctttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt	1080
tcccgccccc gtctgaattt ttgcttccgg tttgggaccg aagccgcgc gcgcgtcttg	1140
tctgctgcag catcgttctg tttttctct gtctgactgt gtttctgtat ttgtctgaaa	1200
atatggccca gactgttacc actccctaa gtttgacctt aggtcactgg aaagatgtcg	1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgtggggtt accttctgtc	1320
ctgcagaatg gccaaccctt aacgtcggat ggccgcgaga cggcacctt aaccgagacc	1380
tcatcacca ggttaagatc aaggctttt cacctggccc gcatggacac ccagaccagg	1440
tccctacat cgtgacctgg gaaggcttgg cttttgaccc ccctccctgg gtcaaggccct	1500
ttgtacaccc taaggctccg cctcccttcc ctccatccgc cccgtctctc ccccttgaac	1560
ctccctcggtc gaccccgctt cgatcctccc ttatccagc cctcactcct tctctaggcg	1620
ccaaacctaa acctcaagtt ct tcttgaca gtggggggcc gctcatcgac ctacttacag	1680
aagacccccc gccttatagg gaccaagac cacccttcc cgacaggac g gaaatggtg	1740
gagaagcgac ccctgcggga gaggcaccgg acccctcccc aatggcatct cccttacgt	1800
ggagacggga gccccctgtg gccgactcca ctacctcgca ggcattcccc ctccgcgcag	1860
gaggaaacgg acagctcaa tactggccgt tctcccttcc tgacctttac aactggaaaa	1920
ataataaccc ttcttttct gaagatccag gtaaaactgac agctctgatc gagtctgtcc	1980
tcatcacca tcagccacc tgggacgact gtcagcagct gttgggact ctgctgaccg	2040
gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgcggggc gatgatggc	2100
gccccactca actgcccattt gaagtcgatg ccgccttcc cctcgagcgc ccagactgg	2160
attacaccac ccaggcaggta aggaaccacc tagtccacta tgcggcgttg ctccctcg	2220
gtctccaaaa cgcggcaga agccccca atttggccaa gttaaaagga ataacacaag	2280
ggcccaatga gtctccctcg gccttccatg agagacttaa ggaagcctat cgcaggtaca	2340
ctcccttatga ccctggggac ccaggcaga aactaatgt gtctatgtct ttcatggc	2400
agtctgcccc agacattggg agaaagtttag agaggttaga agatttaaaa aacaagacgc	2460
ttggagatt gtttagagag gcagaaaaga tcttaataa acgagaaacc ccgaaagaaa	2520
gagagaaacg tatcaggaga gaaacagagg aaaaagaaga acgcgcgttgg acagaggatg	2580

agcagaaaaga	gaaagaaaaga	gatcgtagga	gacatagaga	gatgagcaag	ctattggcca	2640
ctgtcgtag	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccaactcg	2700
atcgcgacca	gtgtgcctac	tgcaaagaaa	aggggactg	ggctaaagat	tgtcccaaga	2760
aaccacgagg	acctcgaaaa	ccaagacccc	agacctccct	cctgacccta	gatgactagg	2820
gaggtcaggg	tcaggagccc	ccccctgaac	ccaggataac	cctcaaagtc	ggggggcaac	2880
ccgtcacctt	cctggtagat	actggggccc	aacactccgt	gctgacccaa	aatcctggac	2940
ccctaagtga	taagtctgcc	tgggtccaag	gggctactgg	aggaaagcgg	tatcgctgga	3000
ccacggatcg	caaagtacat	ctagttaccg	gtaaggtcac	ccacttttc	ctccatgtac	3060
cagactgtcc	ctatccctctg	ttaggaagag	atttgctgac	taaactaaaa	gccccaaatcc	3120
actttgaggg	atcaggagcc	caggttatgg	gaccaatggg	gcagccctg	caagtgttg	3180
ccctaaatat	agaagatgag	tatcgctac	atgagaccc	aaaagagcca	gatgtttctc	3240
tagggtccac	atggctgtct	gatttcctc	aggcctggc	ggaaaccggg	ggcatgggac	3300
tggcagttcg	ccaagctcct	ctgatcatac	ctctgaaagc	aacctctacc	cccgtgtcca	3360
taaaacaata	ccccatgtca	caagaagcca	gactggggat	caagccccac	atacagagac	3420
tgttggacca	ggaaataactg	gtaccctgcc	agtccccctg	gaacacgccc	ctgctaccgg	3480
ttaagaaacc	agggactaat	gattataggc	ctgtccagga	tctgagagaa	gtcaacaagg	3540
gggttggaaaga	catccaccccc	accgtgccc	acccttacaa	cctttgagc	gggctccac	3600
cgtcccacca	gtggcacact	gtgcttgatt	taaaggatgc	cttttctgc	ctgagactcc	3660
accccaccaag	tcagccctctc	ttcgcccttg	agtggagaga	tccagagatg	ggaatctca	3720
gacaattgac	ctggaccaga	ctcccacagg	gtttcaaaaa	cagtcccacc	ctgtttgatg	3780
aggcactgca	cagagaccta	gcagacttcc	ggatccagca	cccagacttg	atccctgctac	3840
agtacgtgga	tgacttactg	ctggccgcca	cttctgagct	agactgcca	caaggtactc	3900
gggcctgtt	acaaaccta	gggaacctcg	ggtatcgggc	ctcgccaaag	aaagccaaa	3960
tttgccagaa	acaggtcaag	tatctgggg	atcttctaaa	agagggtcag	agatggctga	4020
ctgaggccag	aaaagagact	gtgtggggc	agctctactcc	gaagacccct	cgacaactaa	4080
gggagttcct	agggacggca	ggcttctgtc	gectctggat	ccctgggtt	gcagaaatgg	4140
cagccccctt	gtaccctctc	acccaaacgg	ggactctgtt	taattggggc	ccagaccaac	4200
aaaaggecta	tcaagaaatc	aagcaagctc	ttctaactgc	cccagccctg	gggttgcag	4260
atttgactaa	gcccttgaa	ctctttgtcg	acgagaagca	gggctacgca	aaaggtgtcc	4320
taacgcaaaa	actgggacct	tggcgtcgcc	cggtggccta	cctgtccaaa	aagctagacc	4380
cagtagcaga	tgggtggccc	ccttgccctac	ggatggtagc	agccattgcc	gtactgacaa	4440
aggatgcagg	caagctaacc	atggacagc	cactagtcat	tctggccccc	catgcagtag	4500

027693

aggcactagt caaaacaaccc cccgaccgct ggctttccaa cgccccgatg actcactatc	4560
aggccttgct tttggacacg gaccgggtcc agttcgacc ggtggtagcc ctgaacccgg	4620
ctacgctgct cccactgcct gaggaagggc tgcaacacaa ctgccttgat atcctggccg	4680
aagcccacgg aaccggaccc gacctaacgg accagccgct cccagacgcc gaccacacct	4740
ggtacacgga tggaaagcagt ctcttacaag agggacageg taaggcggga gctgcgggtga	4800
ccaccggagac cgaggtaatc tgggctaaag ccctgcccagc cgggacatcc gctcagcggg	4860
ctgaactgat agcactcacc caggccctaa agatggcaga aggtaaaagccttata tacagaaggc	4920
atactgatag ccgttatgct tttgtactg cccatatacc tggagaaaata tacagaaggc	4980
gtgggttgct cacatcagaa ggcaaagaga tcaaaaataa agacgagatc ttggccctac	5040
taaaaggcct ctttctgccc aaaagactta gcataatcca ttgtccagga catcaaaagg	5100
gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggcccgaa aaggcagcc	5160
tcacagagac tccagacacc tctaccctcc tcataaaaaa ttcatcaccc tacacctcag	5220
aacattttca ttacacagt actgatataa aggacctaac caagttgggg gccattttatg	5280
ataaaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgcctgac cagtttactt	5340
ttgaattatt agactttctt catcagctga ctcacccctg cttctcaaaa atgaaggctc	5400
tccttagagag aagccacagt ccctactaca tgctgaaccg ggatcgaaca ctcaaaaata	5460
tcactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaaacagg	5520
gaacttaggggt ccgcgggcat cggccggca ctcattggga gatcgatttc accgagataa	5580
agcccgatt gtatggctat aaatatcttc tagttttat agataccctt tctggctgga	5640
tagaagcctt cccaaaccaag aaagaaaaccg ccaaggctgt aaccaagaag ctactagagg	5700
agatcttccc caggttcggc atgcctcagg tattggAAC tgacaatggg cctgccttcg	5760
tctccaaagggt gагtсагакаа gtggccgatc tggggggat tgattggaaa ttacattgt	5820
catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt	5880
taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc	5940
tgtaccgagc ccgcaacacg ccggggcccc atggcctcac cccatatgag atcttatatg	6000
gggcacccccc gccccttgta aacttccctg accctgacat gacaagagtt actaacagcc	6060
cctctctcca agctcactta caggctctct acttagtcca gcacgaagtc tggagacctc	6120
tggccggcagc ctaccaagaa caactggacc gaccgggtgt acctcaccct taccgagtcg	6180
gcgacacagt gtgggtccgc cgacaccaga ctaagaacctt agaacctcgc tggaaaggac	6240
cttacacagt cctgctgacc acccccacccg ccctcaaagt agacggcatc gcagcttgaa	6300
tacacgcccgc ccacgtgaag gctgccgacc ccgggggtgg accatcctct agactgacat	6360
ggcgcggttca acgctctcaa aaccctcata agataagatt aaccctgtggaa agcccttaat	6420

027693

agtcatggga gtcctgttag gagtagggat ggcagagagc ccccatcagg tctttaatgt	6480
aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg	6540
aactgtacaa gatgccttcc caaaattata ttttgatcta tgtgatctgg tcggagagga	6600
gtggggccct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcaggag	6660
acagcggacc cggaactttg acttttacgt gtgccttggg cataccgtaa agtcggggtg	6720
tgggggacca ggagagggtc actgtggtaa atgggggtgt gaaaccacgg gacaggctta	6780
ctggaagccc acatcatcgt gggaccta at ctccttaag cgcgtaaca cccctggg	6840
cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatccaattc	6900
cttccaagg gctactcgag gggcagatg caaccctcta gtcctagaat tcactgatgc	6960
aggaaaaaag gctaactggg acgggccc aa atcgtgggg ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctccctgac ccggcaggc cttatgtgg gaccccgagt	7080
ccccataggg cccaacccag tattaccga ccaaagactc ccttctcac caatagagat	7140
tgtaccggct ccacagccac ctatccccctt caataccagt tacccctt ccactaccag	7200
tacaccctca acctccctca caagtccaa tgccccacag ccaccccccag gaactggaga	7260
tagactacta gctctagtc aaggagccta tcaggcgctt aacccatcca atcccgacaa	7320
gacccaaagaa tttggctgt gcttagtgc gggacccctt tattacgaag gagtagcggt	7380
cgtggcact tataccaatc attccaccgc tccggcaac tgtacggcca cttcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggctatgc atggggcag tacctaaaac	7500
tcaccaggcc ttatgtaaaca ccacccaaag cgccggctca ggatctact accttgcagc	7560
acccgcccga acaatgtggg cttgcagcac tggattgact ccctgcttgt ccaccacgg	7620
gctcaatcta accacagatt attgtgtatt agtgaactc tggcccagag taatttacca	7680
ctccccctat tatatgtatg gtcagttga acagcgtacc aaatataaaa gagagccagt	7740
atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattt cagctggat	7800
agggacgggg accactgcct taataaaac ccagcagttt gagcagcttc atgccgttat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctagaaaagt cactgaccc	7920
gttgtctgaa gtgtctac agaaccgcag aggcttagat ttgttattcc taaaggaggg	7980
aggtctctgc gcagccctaa aagaagaatg ttgttttat gcagaccaca cggggctagt	8040
gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
aggccaagga tggttcgaag ggctgtttaa tagatcccc tggtttacca cttaatctc	8160
caccatcatg ggacctctaa tagtactctt actgatctta ctcttggac cttgcattct	8220
caatcgattt gtcctttaa ttaaagacag gatctcgtt gtcaggctc tggttttgac	8280
tcagcaatat caccagctaa aaccataga gtacgagcca tgaacgcgtt actggccgaa	8340

027693

gcccgttggaa ataaggccgg tgtgcgtttg tctatatgtt atttccacc atattgccgt	8400
cttttggcaa tgtgagggcc cgaaacctg gcccgttgc cttgacgagc attccttaggg	8460
gtctttcccc tctcgccaaa ggaatgcaag gtctgttga tgctgtgaag gaagcagtcc	8520
ctctggaaac ttcttgaaaga caaacaacgt ctgtacgac ccttgcagg cagcggaaacc	8580
ccccacactgg cgacagggtgc ctctcgccccc aaaagccacg tgtataagat acacctgcaa	8640
aggcggcaca accccagtgac cacgttgtga gttggatagt tgtggaaaga gtcaaattggc	8700
tctcctcaag cgtattcaac aaggggctga aggtaccca gaaggtaccc cattgtatgg	8760
gatctgatct ggggcctcgg tgcacatgtt ttacatgtgt ttagtcgagg ttaaaaaaac	8820
gtcttaggccc cccgaaccac ggggacgtgg ttttccttgg aaaaacacga ttataaatgg	8880
tgaccggccg catggcctcc aagtgggatc aaaagggcat ggatatcgct tacgaggagg	8940
ccctgctggg ctacaaggag ggcggcgtgc ctatcgccgg ctgtctgatc aacaacaagg	9000
acggcagtgt gctgggcagg ggccacaaca tgaggttcca gaagggctcc gccaccctgc	9060
acggcggat ctccaccctg gagaactgtg gcaggctggg gggcaagggtg tacaaggaca	9120
ccaccctgta caccaccctg tccccctgtg acatgtgtac cggcgctate atcatgtacg	9180
gtatccctag gtgtgtgatc ggcgagaacg tgaacttcaa gtccaaaggc gagaagtacc	9240
tgcaaaccag gggccacgag gtgggggttgg ttgacgatga gaggtgttaag aagctgtatga	9300
agcagttcat cgacgagagg cctcaggact gggtcgagga tatcggcgag taagcggccg	9360
cagataaaat aaaagatttt atttagtctc cagaaaaagg gggaaatgaa agaccccacc	9420
tgttaggttg gcaagcttgc ttaagtaacg ccatttttgc aaggcatggaa aaatacataa	9480
ctgagaatag agaagttcag atcaaggatca ggaacagatg gaacagctga atatggcca	9540
aacaggatat ctgtggtaag cagttctgc cccggctcag ggccaaagaaac agatggaaaca	9600
gctgaatatg gcccaaacag gatatctgtg gtaaggcgtt cctgccccgg ctcaggccca	9660
agaacagatg gtcccccagat gcggtccagc cctcagcgtt ttctagagaa ccatcagatg	9720
tttccaggt gccccaaagga cctgaaatgaa ccctgtgcct tatttgaact aaccaatcag	9780
ttcgcttctc gcttctgttc gcgcgcttc gctcccccgg ctcataaaaa gagcccacaa	9840
ccccctcactc gggcgccag tcctccgatt gactgagtcg cccgggtacc cgttatcca	9900
ataaacccttc ttgcagttgc atccgacttg tggctcgct gttccttggg agggtctcct	9960
ctgagtgatt gactaccgt cagcgggggtt cttcattac atgtgagcaa aaggccagca	10020
aaaggccagg aaccgtaaaa aggccgcgtt gctggcggtt ttccataggc tccgcccccc	10080
tgacgagcat cacaaaaatc gacgtcaag tcagaggtgg cggaaaccga caggactata	10140
aagataccag gcgtttcccc ctggaaagctc cctcgtgcgc tctcctgttc cgaccctgcc	10200
gcttaccggaa tacctgtccg cctttctccc ttccggaaagc gtggcgctt ctcatagtctc	10260

027693

acgctgttagg tatctcagtt cggtgttaggt cgttcgctcc aagctggct gtgtgcacga 10320
accccccggtt cagcccgacc gctgcgcctt atccggtaac tatcgcttg agtccaaccc 10380
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcag 10440
gtatgttaggc ggtgctacag agttctgaa gtggtggct aactacggct acaactagaag 10500
gacagtattt ggtatctcgctt ctctgctgaa gccagttacc ttcgaaaaa gagttggtag 10560
ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt tttttgttt gcaagcaga 10620
gattacgcgc agaaaaaaaaag gatctaaga agatccttg atctttctta cggggctgt 10680
cgctcagtgaa aacgaaaact cacgttaagg gatttggtc atgagattat caaaaaggat 10740
cttcaccttag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatataatg 10800
gtaaaacttgg tctgacagtt accaatgctt aatcagttagt gcacctatct cagcgatctg 10860
tctatttcgt tcatccatag ttgcctgact cccgcgtgt tagataacta cgatacggga 10920
ggcttacca tctggcccca gtgctgaat gataccgcga gaccacgct caccggctcc 10980
agatttatca gcaataaacc agccagccgg aaggccgag cgccagaatgt gtcctgcaac 11040
tttatccgcc tccatccagt ctattaattt tgccggaa gctagagtaa gtagttcgcc 11100
agtttaatgt ttgcgcacg ttgttgcatt tgctgcaggc atcggttgtt cacgctcgct 11160
gtttggatgt gtttcattca gctccggttc ccaacgatca aggccgatgtt catgatcccc 11220
catgttgcggc aaaaaagcgg tttagtcctt cggccctccg atcggtgtca gaagtaatgt 11280
ggccgcagggtt ttagtcactca tggttatggc agcactgcat aattctctta ctgtcatgcc 11340
atccgttaaga tgctttctg tgactggta gtaactcaacc aagtcttctt gagaatagt 11400
tatgcggcga ccgagttgtt cttgcggcgtt gtcaacacgg gataataccg cgccacatag 11460
cagaacttta aaagtgcata tcattggaaa acgttctcg gggcggaaac tctcaaggat 11520
cttaccgcgtt tttagatcca gttcgatgtt accactcgat gcacccaaact gatcttcaggc 11580
atctttact ttccaccagcg ttctgggtt agcaaaaaaca ggaaggcaaa atgccgcaaa 11640
aaagggaata agggcgacac ggaaatgtt aatactcata ctcttccttt ttcaatattt 11700
ttgaaggatt tatcagggtt attgtctcat gagcggatatac atatttgaat gtattttagaa 11760
aaataaacaa ataggggttc cgccacatt tcccgaaaaa gtgccacctg acgtctaaga 11820
aaccatttattt atcatgacat taacctataa aaataggcgt atcacgaggc ccttcgtct 11880
tcaagaattt cat 11893

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Рекомбинантный репликационно-компетентный ретровирус (RCR), способный трансформировать клетку-мишень человека полинуклеотидом, кодирующим полипептид с цитозиндезаминазной активностью и повышенной стабильностью в клетках человека, который катализирует превращение экзогенного 5-фторцитозина в 5-фторурацил, обеспечивая цитотоксическое действие у индивида с нарушением клеточной пролиферации, где ретровирус содержит

ретровирусный белок GAG;

ретровирусный белок POL;

оболочку ретровируса;

ретровирусный полинуклеотид, содержащий подходящий для экспрессии в клетке млекопитающего промотор на 5'-конце ретровирусного полинуклеотида, домены gag, pol и env ретровирусного полинуклеотида, последовательность длинных концевых повторов (LTR) на 3'-конце ретровирусного полинуклеотида;

кассету, содержащую внутренний сайт связывания рибосомы (IRES), функционально связанный с полинуклеотидом, кодирующим полипептид с цитозиндезаминазной активностью и содержащим нуклеотиды от приблизительно 8877 до приблизительно 9353 последовательности SEQ ID NO: 19, оптимизированной по кодонам для экспрессии в клетках человека, где указанная кассета расположена между доменом env и последовательностью LTR; и

последовательности, действующие в цис-положении, необходимые для обратной транскрипции, упаковки и интеграции RCR в клетку-мишень, где последовательности, действующие в цис-положении, расположены рядом с последовательностью LTR.

2. Ретровирус по п.1, где ретровирусный полинуклеотид получен из вируса лейкоза мышей (MLV), вируса мышного лейкоза Молони (MoMLV), вируса лейкоза кошек или вируса лейкоза гиббонов (GALV).

3. Ретровирус по п.2, где MLV представляет собой амфотропный MLV.

4. Ретровирус по п.1, где ретровирус представляет собой гамма-ретровирус.

5. Ретровирус по п.1, где клетка-мишень представляет собой клетку с нарушением клеточной пролиферации.

6. Ретровирус по п.1, где клетка-мишень представляет собой неопластическую клетку.

7. Ретровирус по п.5, где нарушение клеточной пролиферации выбрано из рака легких, рака толстой и прямой кишки, рака груди, рака предстательной железы, рака мочевыводящих путей, рака матки, рака головного мозга, рака тканей головы и шеи, рака поджелудочной железы, меланомы, рака желудка и рака яичников.

8. Ретровирус по п.1, где промотор представляет собой промотор, регулируемый гормоном, или промотор, регулируемый цитокином.

9. Ретровирус по п.1, где промотор представляет собой тканеспецифичный промотор.

10. Ретровирус по п.9, где тканеспецифичный промотор содержит по меньшей мере один элемент андрогенного ответа (ARE).

11. Ретровирус по п.10, где элемент андрогенного ответа получен из промотора пробазина.

12. Ретровирус по п.9, где тканеспецифичный промотор включает промотор пробазина.

13. Ретровирус по п.1, где промотор представляет собой промотор CMV, содержащий нуклеотиды от приблизительно 1 до приблизительно 582 последовательности SEQ ID NO: 19, 20 или 22, причем промотор может включать модификацию одного или нескольких нуклеотидов без изменения функциональной активности.

14. Ретровирус по п.1, где промотор содержит нуклеотиды от приблизительно 1 до приблизительно 582 последовательности SEQ ID NO: 19 или 20.

15. Ретровирус по п.1, где промотор содержит полинуклеотидный домен CMV-R-U5.

16. Ретровирус по п.15, где домен CMV-R-U5 содержит ранний промотор цитомегаловируса человека, функционально связанный с областью R-U5 MLV.

17. Ретровирус по п.16, где домен CMV-R-U5 содержит нуклеотиды от приблизительно 1 до приблизительно 1202 последовательности SEQ ID NO: 19, 20 или 22 или последовательности, которая по меньшей мере на 95% идентична последовательности SEQ ID NO: 19, 20 или 22, где указанный полинуклеотидный домен активизирует транскрипцию молекулы нуклеиновой кислоты, функционально связанной с ним.

18. Ретровирус по п.1, где домен gag полинуклеотида получен из гамма-ретровируса.

19. Ретровирус по п.18, где домен gag содержит нуклеотиды от приблизительно 1203 до приблизительно 2819 последовательности SEQ ID NO: 19 или последовательности, которая по меньшей мере на 95, 98, 99 или 99,8% идентична ей.

20. Ретровирус по п.1, где домен pol полинуклеотида получен из гамма-ретровируса.

21. Ретровирус по п.20, где домен pol содержит нуклеотиды от приблизительно 2820 до приблизи-

тельно 6358 последовательности SEQ ID NO: 19 или последовательности, которая по меньшей мере на 95, 98, 99 или 99,9% идентична ей.

22. Ретровирус по п.1, где домен env содержит нуклеотиды от приблизительно 6359 до приблизительно 8323 последовательности SEQ ID NO: 19 или последовательности, которая по меньшей мере на 95, 98, 99 или 99,8% идентична ей.

23. Ретровирус по п.1, где IRES получен из вируса энцефаломиокардита.

24. Ретровирус по п.23, где IRES содержит нуклеотиды от приблизительно 8327 до приблизительно 8876 последовательности SEQ ID NO: 19 или последовательности, которая по меньшей мере на 95, 98 или 99% идентична ей.

25. Ретровирус по п.1, где 3'-концевая последовательность LTR получена из гамма-ретровируса.

26. Ретровирус по п.25, где 3'-концевая последовательность LTR содержит домен U3-R-U5.

27. Ретровирус по п.26, где 3'-концевая последовательность LTR содержит нуклеотиды от приблизительно 9405 до приблизительно 9998 последовательности SEQ ID NO: 19 или последовательности, которая по меньшей мере на 95, 98 или 99,5% идентична ей.

28. Ретровирус по п.1, где ретровирусный полинуклеотид содержит последовательность SEQ ID NO: 19, 20 или 22.

29. Ретровирус по п.1, где ретровирусный полинуклеотид находится в проретровирусной форме и содержит в направлении от 5'-конца к 3'-концу

домен CMV-R-U5, содержащий ранний промотор цитомегаловируса человека, функционально связанный с областью R-U5 MLV;

последовательности, действующие в цис-положении, содержащие участок PBS для связывания праймера для обратной транскриптазы;

5'-концевой сайт сплайсинга и

Ψ сигнал упаковки;

домены gag и pol MLV, 3'-концевой сайт сплайсинга и домен env MLV штамма 4070A;

кассету, содержащую IRES из вируса энцефаломиокардита, полинуклеотид, кодирующий полипептид с цитозинdezаминаизной активностью, содержащий нуклеотиды от приблизительно 8877 до приблизительно 9353 последовательности SEQ ID NO: 19, и полипуриновый тракт;

3'-концевую последовательность LTR, содержащую домен U3-R-U5 MLV.

30. Ретровирус по п.29, где ретровирусный полинуклеотид содержит последовательность SEQ ID NO: 19 или 22.

31. Ретровирус по п.29, где ретровирусный полинуклеотид получен из вируса лейкоза мышей (MLV), вируса мышного лейкоза Молони (MoMLV), вируса лейкоза кошек или вируса лейкоза гиббонов (GALV).

32. Ретровирус по п.31, где MoMLV представляет собой амфотропный MoMLV.

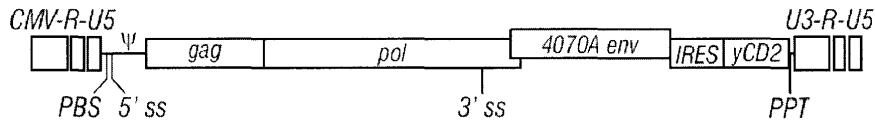
33. Применение ретровируса по п.29 для получения лекарственного средства, катализирующего превращение экзогенного 5-фторцитозина в 5-фторурацил в клетках индивида с нарушением клеточной пролиферации.

34. Применение по п.33, где ретровирусный полинуклеотид содержит последовательность SEQ ID NO: 19, 20 или 22.

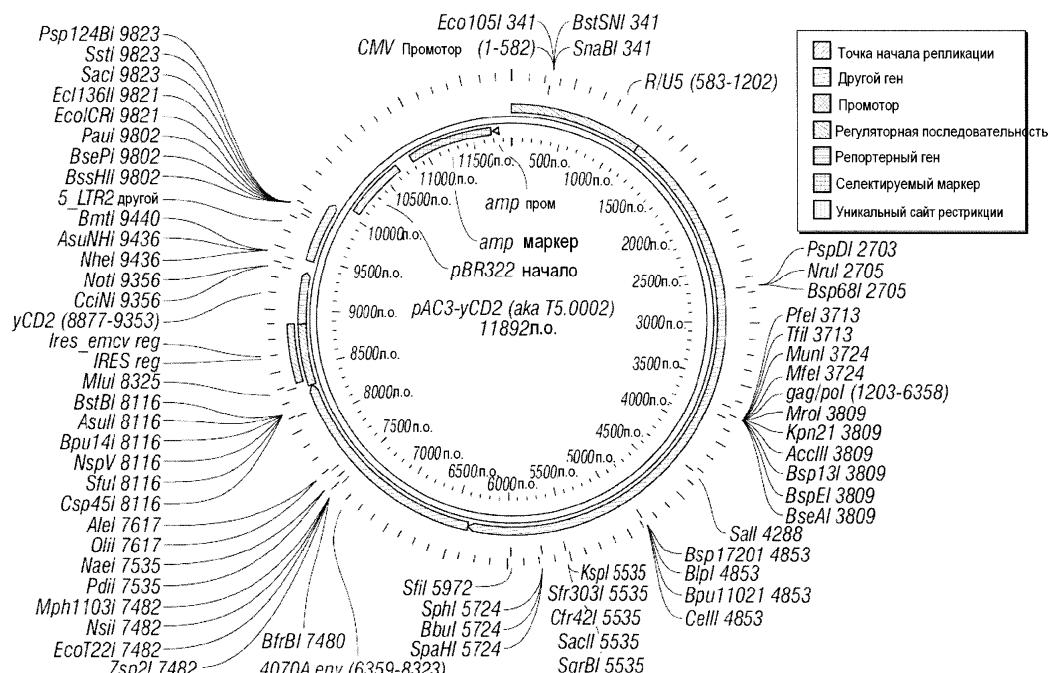
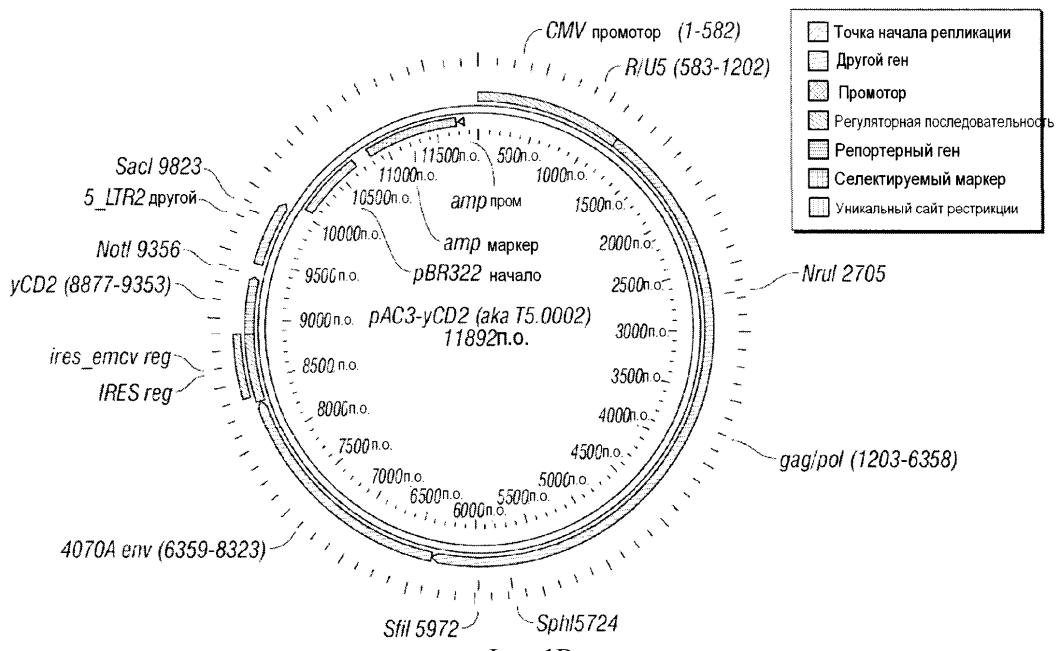
35. Применение по п.33, где нарушением клеточной пролиферации является мультиформная глиобластома.

36. Применение по п.33, где нарушение клеточной пролиферации выбрано из рака легких, рака толстой и прямой кишки, рака груди, рака предстательной железы, рака мочевыводящих путей, рака матки, рака головного мозга, рака тканей головы и шеи, рака поджелудочной железы, меланомы, рака желудка и рака яичников.

37. Применение ретровируса по п.1 для лечения нарушения клеточной пролиферации у индивида посредством каталитического превращения в его клетках экзогенного 5-фторцитозина в 5-фторурацил.



Фиг. 1А



[CACTTATTAAAGTAATCAAATTACGGGTCAATTAGTCATAACCCCAATAATGGAGTCGGCGTACATAACTTACCGTAATA
 |GGCCCGGCGCTGGCTACCGCCCAACGACCCCCGCCMTGACGTCATAAAAGACGTATGTTCCCATAGTAAACGCCAATAGGGA
 |CTTTCCATTGACGTCAATGGGTGGACTATTACGGTAAACTGCCACTTGGCAGTACACAIACTGTCATATGCCAAGTGAC
 |GCCCGCTATTGACGTCAATGGGTGGACTATTACGGTAAACTGCCACTTGGCAGTACACAIACTGTCATATGCCAAGTGAC
 |CACTACATCTACGTATTAGTCATCGCTATTACCATGGTGTACGGTGTTCGGCAGTACACAIACTGTCATATGCCAAGTGAC
 |ACTCACGGGATTCCAACTTCACCCCAATGACGTCATAAGGAGTTCTTGGCACCACAAATCAACGGGACTTTCCAAA
 |ATGTCGTAACAACTCCGCCCATGACGTCATAAGGAGTTCTTGGCACCACAAATCAACGGGACTTTCCAAA
 |GTGAAACCGGCCAGTCCTCCGATGACTGACTGCCCGTACCCCTCATGCCAAACCCCTTGCACCTGGCACTCCAC
 |CTGGGATCGGAGACCCCTGCCACGGGACCCACCCACGGGAGCTTACGGTGGCAGCAACTTATCTGTGTGTGTC
 |GATTGTCCTAGTGTCTATGACTGTATTATGCGCTTGGCTGGTGTACTAGTGTACTAGCTCTGTAATCTGGGGACCCGTG
 |GTGGAACTGACGAGTTGGAACACCCGGCGCAACACCTGGAGACGTCGGAGCTGGGGCGTGTGTC
 |CTGACTCCAAAATCCCGATCGTTGGACTCTTGGCACCCCTTAGAGGAGGAAATGTGGTTCTGGTAGGAGACGA
 |GAACCTAAAACACTCCCGCCTCCGTCGATTTTGCTTGGGACCGAAGCCGCGCCGCGTGTGTGCTG
 |AGCACTGGTCTGTTGCTCTGCTGACTGTGTCTGTTGCTGAGATAAGGGCCAGACTGTACCACTCCCTAA
 |GTTGACCTTGGTACTGGAAAGATGTCGAGGGATGCTACACCAACTCGGTAGATGTCAGAAAGAGACGTTGGGTTAC
 |CTTCCTGCTTCGAGAAATGGCCAACCTTAACTGCGGATGGCGGAGACGGCACCTTAAACCGAGACCTCATCACCCAGTT
 |AAGATCAAGGCTTTACCTGGCCCGATGGACACCCAGACCCAGGTCCTCATCTGTCACCTGGGAGGCTTGGGTTTG
 |ACCCCGCTCCCTGGCTCAAGCCCTTGACACCCCTAACGCTCCGCTCCCTGCTCCCTGCTCCCTG
 |ACCTCTCTGCTGACCCCGCTGATCTCCCTTATCCAGCCCTCACTCTCTGCTAGGGCCCAAACCTAAACCTCAAGT
 |CTTCTGACACTGGGGGGCCGCTACCGACCTACTACAGAAGACCCCGCTTATAGGGACCCAAAGACACCCCGCTCG
 |ACAGGGACGGAATGGGGAGAGCAGCCCTGGGGAGAGGACCCGACCCCTCCCAATGGCATCTCGCTACGTTGGGAG
 |ACGGGAGCCCCCTGTCGCGACTCACTACCTCGCAGGCACTTCCGCGAGGAGGAAACGGGACGCTCAAACTGG
 |CGGTTCTCTTTGACCTTACAACGGAAAATAACCCCTTTTGTGAGATCCAGGTTAAACTGACAGCTCTGA
 |TGGAGTCTGTCATCACCCCATCAGCCACCTGGGACACTGTCACCGAGCTGGGACTCTGCTGACCGGAGAAGAAA
 |ACAACGGGTGCTCTAGAGGCTAGAAAGGGCTGGGGGGCGATGATGGGGCCCTCACTAACTGCCCAATGAACTCCGATGCC
 |GCTTTCCCTCGAGCCGAGACTGGGATTACACCAACCCAGGGTACGGAAACCCCTGCTGCTACGAGGTTAAAGGATTTAAACAGAC
 |TAGCGGGCTCCCAAACGGGGCAGAAGCCCCACCAATTGGCAAGGTTAAAGGATAACACAGGGCCCAATGAGCTCTCC
 |CTCGGCGCTCCCTAGAGAGACTAACGGACCTATGGCAGGTTACCTCTTATGACCTGAGGACCCAGGGCAAGAAACTAAAT
 |GIGTGTATGTTCTTCTTGGAGTCTGCCCCAGACATTGGGAGAAAGTTAGAGAGGTTAGAAGGTTAAAGGATTTAAACAGAC
 |TTGGAGATTGGGTTAGAGAGGCAAGAAAAGATTTAAATAAACGAGAAACCCCGGAAGAAGAGACGGAAACGTTACGGAGAGA
 |AACAGAGGAAAAGAGAACCCCTACGACAGGGAGTACGGAGAAACAGAGAAAGAGATCTGGAGGAGACATAGAGAGATG
 |AGCAAGCTTATGGCCACTGTGCTTACTGGACGAAACAGGTTAGACAGGGAGGAGAACGAAAGGAGGTCCCAACTCGATGCC
 |ACCAAGCTGCTACTGCAAGAAAAGGGGCACTGGCTAAAGATGTCCTAAAGAAGACCCACGGACCTGGGGACCCAGAAC
 |CMV промотор
 R
 U
 G = 5'ss

Фиг. 1D-1

CCAGACCTCCCTCCTGACCCCTAGATGACTAGGGAGGTCAAGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCCTCAAAGTC
 GGGGGCAACCCGTCACTTCTGGTAGATACTGGGCCAACACTCGCTGCTGACCCAAATCTGGACCCCTAAGTGATA
 AGTCTGCCTGGGTCCAAGGGCTACTGGAGGAAGGGTATCGCTGGACCAAGGTGCTAAAGTACATCTAGCTACCGTAA
 GGTCAACCCACTCTTCTCCATGTACAGACTGTCCTATCCCTGTAGGAAGAGATTGCTGACTAACTAAAAGCCAA
 ATCCACTTGAGGGATCAGGAGCCAGGTATGGCACCAATGGGCAGCCCCCTGAACCTGTCACCCCTAAATATAAGACATC
 AGCATGGCTACATGAGACCTCAAAGAGCCAGATGTTCTCTAGGGTCCACATGGCTGCTGATTTCTCAGGCCTGGC
 GGAAACCGGGGGCATGGACTGGCACTTCCTCTGTATCATACCTCTGAAACCAACCTTAACCCCGTGTCCATA
 AAAACAATACCCATGTACAAGAAGCCAGACTGGGATCAAGCCCCACATAAGAGACTGTTGGACCAGGGAAATCTGGTAC
 CCTGCCACTCCCCCTGGAACACGCCCTGOTACCGCTTAAGAACAGGGACTAATGATATAGGCCTGTCAGGAATCTGAG
 AGAAGTCAACAAGGGGTGGAAGACATCCACCCACCGTCCCCAACCTTACAACCTTGTGAGACTCCACCCACCACTGAG
 CAGTGGTACACTGTCTTGATTTAAGGATGCCCTTCTGAGACTCCACCCACCACTGAGCTOTCTGCTTTC
 AGTGGAGAGATCCAGAGATGGAATCTCAGGACAATTGACCTGGACAGACTCCACAGGGTTCAAAAACAGTCCCACCC
 GTTTGATGAGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCACCCAGACTTGATCTGCTACAGTACGGATGAC
 TTACTGCTGCCGCCACTCTGAGCTAGACTGCCAACAGGTACTCGGGCTGTACAAACCCCTAGGGACACTCGGGTATC
 GGGCTCGGCCAGAAAAGAGACTGTGATGGGGCAGCCTACTCCGAGAACCCCTCGACAACTAAGGGAGTTCTAGGGACGGCA
 GGCTCTGTGCCCTGGCTCCCTGGTTTCCAGAATGGCAGCCCCCTGTACCCCTCTCACAAAACGGGACTCTGTTA
 ATGGGGGCCAGACCAAAAGGCTATCAAGAAATCAAGCAAGCTTCTAATGCCCCAGCCCTGGGTIGCCAGATT
 GACTAAGCCCTTGAACCTTGTGACGAGAAGCAGGGCTACGCCAAAGGTGCTAACGCAAAACTGGGACCTTGGCGT
 CGGGGGTGGCTTACCTGTCAAAAAGCTAGACCCACTAGCAGCTGGTGGCCCTTGCCTACGGATGGTAGCAGCCATTG
 CGTACTGACAAGGATGAGGCAAGCTAACCATGGACACCCACTAGCTCTGCCCCCATGCACTAGAGGCACTAGT
 CAAACAACCCCCCGACCGCTGGCTTCCAACGCCGATGACTCACTATCAGGCTTGTCTGGACACGGACCGGGTCCAG
 TTGGGACCGGTGGTAGCCCTGAACCCGCTACGCTGCTCCCACTGCGTGGAGGGCTGCAACACAACCTGCTGATATCC
 TGGCGAACGCCACCGAACCCGACCTAACGGACCAGCCGCTCCAGACGCCGACACCTGGTACACGGATGGAG
 CAGCTCTTACAAGAGGGACAGGTAGGGGGAGCTGGGTGACCAACCGAGACCCGGAGGTAAATCTGGCTAAAGCCCTGCGA
 GCCGGGACATCCGCTAGGGGCTGAACTGATACCACTGCCATATCCATGGAGAAATATACAGAACGGCTGGGTGCTCACATCAGAAGG
 CAAAGAGATCAAAATAAGAGGAGACTTGGCCACTAAAGCCCTTCTGCCCCAAAGACTTAGCATAATCCATTG
 CCAGGACATCAAAAGGGACACAGCGCCGAGGCTAGAGGCAACCGGATGGCTGACCAAGCGCCGAAAGGCAGCCATCACAG
 AGACTCCAGACACCTTACCCCTCATAGAAAATTCATCACCCCTACACCTCAGAACATTTTCAATTACACAGTGACTGATAT

[P01]

Фиг. 1D-2

AAAGGACCTAACCAAGTTGGGCCATTATGATAAAACAAAGAAGTATTGGGTACCAAGGAAACCTGTATGCCCTGAC
 CAGTTTACTTTGAATTATTAGCTTCTCATCAGCTGACTCACCTCAGCTCTCAAATGAAGGCTCTCTAGAGAGA
 GCCACAGTCCTACTACATGCTAACCGGGATCGAACACTCAAATATCACTGAGACCTGCAAAGCTTGTCACAAGTCAA
 CGCCAGTAAGCTGCCCTTAAACAGGGAACTAGGGTCCGGGCATCGGCCCGGACTCATGGGAGATCGATTTCAACCGAG
 ATAAAGCAGGATGTTATGSCATAAATATCTAGTTTATAGATACTTCTGGGATAGAAGCCTTCCAACCA
 AGAAGAAACGCCAGGTGTAACCAAGAACCTAGAGGAGATTTCAGCTGGCATGCCATCTCAGGTATTGGGAA
 TGACAAATGGGCTGCCCTTCGTCAGGGTACAGGGAGATCTGGGATAGTGGGAATTACATTTGCA
 TACAGACCCAAACCTAGGCCAGGTAGAAAGAATGAATAGAACCATCAAGGAGATTAACTAAATTAAACGCTGCAACTG
 GCTTACAGAGCTGGGCTCTACTCCCTTAGGCCCTACCGAGCCGCAACACGCCGGGCCAATGCCCTACCCCTA
 TGAGATTTATGGGCACCCCGCCCCCTGTAACCTTCTGACATGACAAGAGTTACTAACAGCCCCCTCTCTC
 CAACCTACCTACAGGCTCTACTTAGTCAGCAGCAACTCTGGGACCTCTGGGGAGCTACCAAGAACACTGGAC
 GACCGGGCTACCTACCCCTAACGGAGTOGGGAGACAGCTGGGCTCCGGACACAGACTAAAGAACCTAGAACCTCGCTG
 GAAAGGACCTTACACAGTCTGACCCCTAACGGGCTAACAGTACGGGATCGGACCTGCTGGGATACACGCCGCCAC
 GTGAAGGCTGCCACGGGGGGGGGGGACATCTAGACTGACCTGGGCTTCAAGCTTCAAAACCCCTCAAGATAA
 GATTAAACCTGGGAGACCCCTTAATACTCTAGGGACTCTGGAGACTAGGGATGCCAGAGGCCCATCAGCTCTTAA
 GTAACTGGGAGCTACCAACCTGATGACTGGGCTACCGCCTGGGACTCCCTCTGGGAATGTAACAGATGCTTCC
 CAAAATTATATTGATCTATGATCTGGTGGAGAGGGTGGGACCTTCAGACAGGAAACCGTATGTCGGTATGGCTG
 CAAGTACCCCGCAGGGAGACAGGGACCCGGACTTTGACTTTACGTGTGCCCCTGGGATACCGTAAAGTGGGTTGGG
 GGACAGGAGAGGGTACTGTGTAATGGGTTGAAACCACCCGACAGGTTACTGGAGGCCACATCATGTGGGAC
 TAATTCCTTAACCGGGTAACACCCCTGGGACACGGATCTTAAAGGTTGGCCCTGCTACGACTCTCCAA
 AGTATCCAATTCTTCAAGGGGACTCGAGGGGGAGATGCAACCCCTAGCTCAAGTACGTGAGGAAAAAG
 GCTAACTGGGACGGGCCAAATGGGAGCTGAGACTGTACCGGACAGGACATGTTACCGTGTCTGGGACCC
 GGCAGGGCTTAACTGGGACCCGAGCTTACGGGACCCAAACGGCTACTGGGACCC
 AGAGATTCTACGGCTCACACCCACTAGCCCCCTAATACCGTAAACCCCTTCAACTACAGTACACCTCAACCTC
 CCTACAAGTCCAAGTCTCCACAGGCCACCCCAAGGAACTGGGAGATAGACTACTAGCTCTAGTCAAGGACCTATCAGGCC
 TTAACTCTACCAATCCCGAACAGACCAAGAATGTTGGCTGTGCTTACTGTGCGGACCTCTTATACGAGGAGTAGCGGT
 CGTGGGCACTTACCAATCATCCACCGCTGGGCAACTGTACCGGCACTTCCACATAAGCTTACCCCTATCTGAAGTG
 ACAGGACAGGGCTATGATGGGGCACTACCTAAACACTACCGGCTTATGAAACCCACCCAAACGCCGGCTCAGGAT
 CCTACTGCTGACAGACCCGGGGGGGAAATGCTGGGCTTCTGGGACCTCTGGGACCCACCTGGGACCCACGGGTCTCAA
 TCTAACACAGATTATGTTGTTAGTGTGACTCTGGCCCTGGGAGCTTACCTGGGGGAGGATTATATGTTATGGTCAGCTT
 GAACAGGCTACCAATAAAAGAGGGCACTATCTGGGCTTCTACTAGGGGATTAACCTGGGGGGATTTG
 CAGCTGGAAATAGGGACGGGGACCACTGGGCTTAAACCCAGCAGTTGACGCTCTGGGCTATCCAGACAGACCT
 CAACGAAGTCAAAACTCAATTACCAACCTAGAAAAGTCACTGACCTCTGCTGAAGTAGCTTACAGAACUGCAGAGGC
 CTAGATTGCTTACCTAAAGGAGGGAGGTCTGGCAGGCCCTAAAGAAGAATGTTTTATGGAGACCACACGGGG
 TAGTGAGGAGCAGCATGGGAAATTAAAGGAGAAGGCTTAACTGAGGACAAAAACTATTGAGACAGGGCAAGGATGTTG
 AGGGCTTAAATAGCTCCCTGGTTACCCATTAACTCCACCATCTAGGGCTTAAATAGTACTTCTACTGATCTT
 CTCTGGGACCTGCTTACCTGATGTCCTGACCTTAAAGGAGGATCTACTGGTCCAGGCTCTGGTTTGACTC
 NGCAATATCAGCTAAACACCCATAGACTGAGCTGCTGCTACTGGCCGAGGCTTGGAAATAAGCCGGTGTG
 CGTGGCTTATGTTTATTCACATATTGGCTTCTGGGAAATGIGAGGGGGGGAAACCTGGCCTGCTTCTGTGAC
 GAGCATTCCIAAGGGCTTCTCCCTGGGAAAGGAATGCAAGGCTGTTGAATGIGCTGAAAGGAAGGCTTCTGTGGA
 GCTCTTGGAAGACAACACCCATGCTGAGGCTTCTGGGACCCCTGGGAGACCCGGGACGGTCTCTGGG
 AAAAGGCCACCTGATAAGTACACCTGCAAGGGGACACCCCGAGTGGCCACCTGTTGAGTTGGGATAGTGTGGAAGAG
 CAAATGGCTCTGAAAGCTATCAAAAGGGCTGAAGGATGCCAGAAGGTTCCCTGTTGATCTGATCTGGG
 CCTGGGTGACATGCTTACATGTTGTTAGTGAGGCTTAAAGGAGGCTTACGGGCCCCGAAACCCAGGGGAGCTGGTTTCC
 TTGAAAAACGATTATAAAATGGTGAAGGGGCGGGATGGGCTTCAAGGAGTCAAAGGGCATGGATATGCTTACAGGA
 GGGCTCTGGGCTACAGGGGGGGGGCTGGGCTTATGGGGGCTGCTGATCAACAAACAGGGGAGGCTGCTGGGAG
 GGGCACACATGGGTTCCAGAGGGCTCCACCTGCAACGGAGGCTGCACTGAGGAGACTCTGGCAGGTTGGG
 GCAAGGCTACAAGGACACCCACCTGACACCCCTGGGCTTGTGACATGTTACCGGGCTATCATGATGAGG
 CCTCTGGGACCTGATGGGGAGAACGTAACCTCAAGTCAAGGGGAGAAGTACCTGAAACCAGGGGGAGGGTGG
 GTGTTGACCATGAGGGTGTAGAAGGCTGATGAGGCTGACCTGAGGCTTACGGACTGGTCCAGGATATCGGG

[AG] = 3' ss

4070A env

M11

EMCV IRES

Ps11

yCD2

Фиг. 1D-3

AGTAA~~CGGGCGG~~AGATAAAAATAAAGATTATTTAGTCACCAAAAAAGGGGGAAATGAAAGACCCACCTG~~TG~~AGGT~~TG~~
 GCAGCTAGCTTAACTAACGCCATTTCGAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGITCAGATCAAGGTAGG
 AACAGATGGAACAGCTGAATATGGCCAACAGGATATCTGTGTAAGCAGTCTGCCCCGGCTCAGGGCAAGAACAGAT
 GGAACAGCTGAATATGGCCAACAGGATATCTGTGTAAGCAGTCTGCCCCGGCTCAGGGCAAGAACAGATGGTCCC
 AGATGCGGTCAAGCCCTCAGCAGTCTAGAGAACCATCAGATGTTTCAAGGTGCCCAAGGACCTGAAATGACCCGTG
 CTTATTGACCAACAAATCAGITCGCTTCTCGCTTCTGTTGCGCCGCTTCTGCTCCCGAGCTCAATAAAAGAGCCACA
 CCCCTCACTCGGGGCGGAGTCTCGGATITGACTGAGTGCGCCGGTACCCCTGTAATCCAAATAACCCCTCTGCACTG
 CGGACTTGTGCTGCTGTTCTTGGAGGTCTCTCTGAGTCACTACCCCTCACGGGCGCTTCTCATTCATTCATCT
 GAGCAAAGGGCAGCAAAGGGCAGGAACCTAAAGGGCGCGGTTGCGGTTTCCATAGGCTCCGCCCCCTGACGA
 GCTACAAAAATCAGCTCAACTCAGGGTGGCAAACCCGACAGGACTATAAAGATAACAGGGTTTCCCTGGAAAGC
 TCCCTCGTGGCTCTGCTGTTCCGACCCCTGCGCTTACCGGATACCTCTCCGCTTCTCCCTGGSAAGCGTGGCGCTT
 CTCAATGCTCACCGCTAGGTATCTCAGITCGGTAGGTCTGCTCAGGCTGACGAAACCCCCCTGCA
 GCGGACCGCTGCGCTTACCGTAACATCTGCTTGTAGGTTGCTACAGAGTTCTGAAGTGGCTTAACACGGCTACACTA
 ACTGTTAACAGGATTAGCAGAGCGAGGTATGTTAGGGGTGCTACAGAGTTCTGAAGTGGCTTAACACGGCTACACTA
 GAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCTCTGATCCGCAAACA
 AACCCACCGCTGGTAGCGGTTGGTTTTTTGTTGCAAGCAGCAGATTACCCGCAAGAAAAGGATCTAAGAAGATCTTIG
 ATCTTCTACCGGGCTGACGCTCAGTGGAAACAAACTCACGTTAAGGATTTGGTCATGAGATTATCAAAAGGATCT
 TCACCTAGATTTAAATTAAAGGTTAAATCACTAAAGTATATGTTAGGTTAACCCTGGCTCTGACAGTTACCA
 ATGCTTAATCAGTGAGGACCTATCTCAGGATCTGTTCTCATCCATAGTGGCTGACTCCCCCTCGTGTAGATA
 ACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCCGAGACCCACGCTCACCGGCTCCAGATTAT
 CAGCAATAAACGCCAGCGGAAGGGCAGGGCAGAAGTGGCTCTGCAACTTTACCGCTCCATCCAGTCTATTAAATTG
 TTGGCGGAGCTAGTAAAGTAGTTTGGCTTAAATAGTTTGGCAACGTTGTTGCTTCTGCACTGCTGGTGTCA
 CGCTGCTGTTGGTATGGCTTCACTCAGCTCCGGTCCGATCGTGTAGAAGTGGCCGACTGTTTACTCATGGTTATGGCAGCACT
 GCATAATTCTTACTGTCATGCCATCCGTAAGATGTTTCTGACTGGTCACTACTICAACCAAGTCATTCTGAGAATAG
 TGATGCGGCGACCGAGTTGCTTGGCCGCTCAACAGGGATAATACCCGCCACATACCGAAACTTTAAAGTCCTCA
 TCATTGGAAAACCTTCTCGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTGTAAACCCACTGG
 ACCCAACTGATTTCAAGCATTTTACTTCACCCAGCTTCTGAGGAAACAGGAAGGAAATGCGGAAAAG
 GGAATAAGGGGACACGGAAATGTTGAATACCTCTCTCTTCAATATTATGAGCATTATCAGGGTTATTGTC
 TCATGAGGGATACATATTGAGATGTTAGAAAATAACAAATAAGGGGTTCCGGCACATTCCCCGAAAAGTGCACC
 TGACGTCAAGAAACCATTATTATCATGACATTAAACCTATAAAATAAGGGTATCACGAGGCCCTTGTCTCAAGAATT
 AT

Фиг. 1D-4

CMV Промотор (1-582) >>>

1 TAGTTATTAAAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATAATGGAGTTCCG 60
 ATCAAATAATTATCATTAGTAATGCCCATGCCCCAGTAATCAAGTATCGGGTATACCTCAAGGC

 61 CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCAACGACCCCCGCCATT 120
 GCAATGTATTGAATGCCATTACCGGGCGGACCGACTGGCGGGTTGCTGGGGGGTAA

 121 GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCATTGACGTCA 180
 CTGCAGTTATTACTGCATACAAGGGTATCATTGCGGTTATCCCTGAAAGGTAACGTGAGT

 181 ATGGGTGGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCC 240
 TACCCACCTCATAAATGCCATTGACGGGTGAACCGTCATGTAGTTCACATAGTATAACGG

 241 AAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTA 300
 TTCATGCGGGGATAACTGCAGTTACTGCCATTACCGGGCGGACCGTAATACGGGTCA

EcoI05I

SnaBI

BstSNI

301 CATGACCTTATGGGACTTCCACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC 360
 GTACTGGAATACCCTGAAAGGATGAACCGTCATGTAGATGCATAATCAGTAGCGATAATG

 361 CATGGTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTGACTCACGGGG 420
 GTACCACTACGCCAAAACCGTCATGTAGTTACCCGCACTATGCCAAACTGAGTGGCCC

 421 ATTTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTGTTGGCACCAAAATCAACG 480
 TAAAGGTTAGGGTGGGTAACTGCAGTTACCCCTCAAACAAAACCGTGGTTAGTTGC

 481 GGACTTCCAAAATGCGTAACAACCTCCGCCCCATTGACGCAAATGGCGGTAGGCGTGT 540
 CCTGAAAGGTTTACAGCATTGAGGCGGGTAACTGCAGTTACCCGCCATCCGCACA

R область (583-650) >>>

541 ACGGTGGGAGGTCTATAAGCAGAGCTGGTTAGTGAACCGCGCCAGTCCTCCGATTG 600
 TGCCACCCCTCCAGATATATCGTCGACCAAATCACTGGCCCGGGTCAGGAGGCTAAC

Фиг. 1D-4

U5 область (651-1202) >>>

601 ACTGAGTCGCCGGTACCCGTATCCAATAAACCCCTCTGGCAGTTGCATCCGACTTGT 660
 TGACTCAGCGGGCCATGGGCACATAGGTTATTGGGAGAACGTCACAGTAGGCTGAACA

 661 GGTCTCGCTGTTCTGGGAGGGTCTCCTCTGAGTGATTGACTACCCGTCAAGGGGGTC 720
 CCAGAGCGACAAGGAACCCCTCCAGAGGGAGACTCACTAACTGATGGCAGTCGCCCG

 721 TTTCATTTGGGGGCTCGTCCGGGATCGGGAGACCCCTGCCAGGGACCACCGACCCA 780
 AAAGTAAACCCCGAGCAGGCCCTAGCCCTCTGGGACGGGCTGGTGGTGGGTGGT

5'SS (788)

781 CCGGGAGGTAAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGAC 840
 GGCCCTCCATTGACCGGTGTTGAATAGACACAGACAGGCTAACAGATCACAGATACTG

 841 TGATTTATGCGCCTGCGTCGGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCG 900

Фиг. 1D-5

ACTAAAATACGCGGACGCAGCCATGATCAATGATTGATCGAGACATAGACCGCCTGGC
 901 TGGTGGAACTGACGAGTTCGGAACACCCGGCGCAACCCTGGGAGACGTCCCAGGGACTT 960
 ACCACCTTGACTGCTCAAGCCTTGTGGCCGGCGTTGGGACCCCTCTGCAGGGTCCCTGAA
 961 CGGGGGCGCTTTGTGGCCGACCTGAGTCAAATCCCGATCGTTTGGACTCTTG 1020
 GCCCCCGGCAAAACACCGGGCTGGACTCAGGTTTAGGGCTAGCAAAACCTGAGAAC
 1021 GTGCACCCCCCTTAGAGGAGGGATATGTGGTCTGGTAGGAGACGAGAACCTAAACAGT 1080
 CACGTGGGGGAATCTCCTCCCTATACACCAAGACCATCCTCTGCTTGGATTTGTCA
 1081 TCCCGCCTCCGTCTGAATTTCGTTGGGACCGAAGCCGCGCGCGCGTCTTG 1140
 AGGGCGGAGGCAACTAAAAACGAAAGCCAACCCCTGGCTTCGGCGCGCGCAGAAC
 1141 TCTGCTGCAGCATCGTCTGTGTTCTGTACTGTGACTGTGTTCTGTATTGCTGAGA 1200
 AGACGACGTCTGAGCAAGACACAAGAGACAGACTGACACAAAGACATAACAGACTCT
 gag(1203, 2819) >>>
 1201 ATATGGGCCAGACTGTTACCACTCCCTTAAGTTGACCTTAGGTACTGAAAGATGTCG 1260
 TATAACCGGTCTGACAATGGTAGGGAATTCAAACGTGAAATCCAGTGACCTTCTACAGC
 1261 AGCGGATCGCTCACAAACCACTGGTAGATGTCAAGAAGAGACGTTGGTTACCTTCGCT 1320
 TCGCTAGCGAGTGTGGTCAGCCATCTACAGTTCTCTGCAACCCATGGAAGACGA
 1321 CTGCAGAATGCCAACCTTAAAGTCGGATGGCCGCGAGACGGCACCTTAACCGAGACC 1380
 GACGTCTACCGGTTGAAATTGCAAGCCTACCGCGCTCTGGCTGAAATTGGCTCTGG
 1381 TCATCACCCAGGTTAAGATCAAGGTCTTTCACCTGGCCGCATGGACACCCAGACCGG 1440
 AGTAGTGGGTCATTCTAGTTCCAGAAAAGTGGACGGCGTACCTGTGGTCTGGTCC
 1441 TCCCTACATCGTGCACCTGGGAAGCCTGGCTTTGACCCCCCTCCCTGGGTCAGCCCT 1500
 AGGGGATGTACCACTGGACCCCTCGGAACCGAAAAGTGGGGGAGGGACCCAGTCGGGA
 1501 TTGTACACCTAAACCTCGCCCTCCTCTCCATCCGCCCCGCTCTCCCTTGAAC 1560
 AACATGTGGGATTGGAGGAGGAGGAGAAGGAGGTAGGGGGGAGAGAGGGGAACCTG
 1561 CTCCTCGTTGACCCCGCCTCGATCCTCCCTTATCCAGCCCTCACTCCTCTAGGCG 1620
 GAGGAGCAAGCTGGGCGGAGCTAGGAGGAAATAGGTGGAGTGAGGAAGAGATCCGC
 1621 CCAAACCTAAACCTCAAGTCTTCTGACAGTGGGGGCCGCTCATGACCTACTTACAG 1680
 GGTTGGATTGGAGTTCAAGAAAGACTGTCACCCCCCGCGAGTAGCTGGATGAATGTC
 1681 AAGACCCCCCGCCTTATAGGGACCCAAGACCACCCCTCCGACAGGGACGGAAATGGTG 1740
 TTCTGGGGGGCGGAATATCCCTGGTTCTGGTGGGGAGGCTGTCCCTGCCTTACAC

Фиг. 1D-6

1741 GAGAAGCGACCCCTGCAGGGAGAGGCACCGGACCCCTCCCCAATGGCATCTGCCTACGTG 1800
 CTCTTCGCTGGGGACGCCCTCTCGTGGCTGGGGAGGGTTACCGTAGAGCGGATGCAC
 1801 GGAGACGGGAGGCCCTGTGGCGACTCCACTACCTCGCAGGCATTCCCGCTCGCGCAG 1860
 CCTCTGCCCTCGGGGACACCGGCTGAGGTGATGGAGCGTCCGTAAGGGGAGGCGCGTC
 1861 GAGGAAACGGACAGCTTCAATACTGGCGTTCTCCTCTTGACCTTACAATGGAAAAA 1920
 CCTCTTGCTGTCGAAGTTATGACCGGAAGAGGAGACTGGAAATGTTGACCTTT
 1921 ATAATAACCCCTTTCTGAAGATCCAGTAAACTGACAGCTCTGATCGAGTCTGTTTC 1980
 TATTATTGGGAAGAAAAGACTCTAGGTCCATTGACTGTCGAGACTAGCTCAGACAAAG
 1981 TCATCACCCATCAGCCCACCTGGGACGACTGTCAGCAGCTGTTGGGACTCTGCTGACCG 2040
 AGTAGTGGTAGTCGGGTGACAGTCGTCGACAACCCCTGAGACGACTGGC
 3121 ACTTTGAGGGATCAGGAGCCAGGTTATGGGACCAATGGGGCAGCCCTGCAAGTGGTGA 3180
 TGAAACTCCCTAGTCCTCGGGCCAATACCTGGTTACCCGTCGGGACGTTCACAACT
 3181 CCCTAAATATAAGAGATGAGCATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTCTC 3240
 GGGATTATATCTTCACTCGTAGCCGATGTACTCTGGAGTTTCTCGGTCTACAAAGAG
 3241 TAGGGTCCACATGGCTGTCTGATTTCCCTCAGGCCTGGGCGGAAACCGGGGATGGGAC 3300
 ATCCCAGGTGACCGACAGACTAAAGGAGTCCGGACCCGCTTGGCCCCCGTACCCCTG
3' SS (3314)
 3301 TGGCAGTTGCCAAGCTCCTCTGATCATACCTCTGAAAGCAACCTTACCCCGTGTCCA 3360
 ACCGTCAGCGGTTGAGGAGACTAGTATGGAGACTTCGTTGGAGATGGGGCACAGGT
 3361 TAAAACAATACCCATGTACAAGAAGCCAGACTGGGATCAAGCCCCACATACAGAGAC 3420
 ATTTGTTATGGGTACAGTGTCTCGGTCTGACCCCTAGTCGGGTGTATGTCCTG
 3421 TGTGGACCAAGGAATACTGGTACCCCTGCCAGTCCCCCTGGAACACGCCCTGCTACCCG 3480
 ACAACCTGGTCCCTATGACCATGGACGGTCAGGGGACCTTGTGCGGGGACGATGGC
 3481 TTAAGAAACCAAGGACTAATGATTAGGCCTGTCCAGGATCTGAGAGAAGTCACAAAGC 3540
 AATTCTTGGTCCCTGATTACTAATATCCGGACAGGTCTAGACTCTTCAGTTGTC
 3541 GGGTGAAGACATCCACCCACCGTGCCCAACCCTACAAACCTTTGAGCGGGCTCCAC 3600
 CCCACCTTCTGAGGTGGGTGGCACGGGTGGAAACTCGCCGAGGGT
 3601 CGTCCCACCAAGTGGTACACTGTGCTTGATTAAAGGATGCCTTTCTGCCTGAGACTCC 3660
 GCAGGGTGGTCACCATGTGACACGAACAAATTCTACGGAAAAGACGGACTCTGAGG
 PfeI
 |
 TfI
 |
 3661 ACCCCCACCAAGTCAGCCTCTTCGCTTGGAGAGAGATCCAGAGATGGGAATCTCAG 3720
 TGGGGTGGTCAGTCGGAGAGAAGCGGAAACTCACCTCTAGGTCTCACCTTAGAGTC

Фиг. 1D-7

MfeI
|
MunI
|
3721 GACAATTGACCTGGACCAGACTCCCACAGGGTTCAAAAACAGTCCCACCCGTGATG 3780
CTGTTAACTGGACCTGGTCTGAGGTGTCCTGAGGGTCAAGTTGTCAGGGTGGGACAACTAC

MrnI
|
BseAI
|
Bsp13I
|
BspEI
|
Kpn2I
|
AccIII
|
3781 AGGCAC TGACAGAGACCTAGCAGACTTCCGGATCCACAGCACCCAGACTTGATCCTGCTAC 3840
TCCGTGACGTGTCCTGGATCGTCTGAAGGCCTAGGTGTCGGTCTGAACTAGGACGATG

3841 ACTACGTGGATGACTTACTGCTGGCCCACTTCTGAGCTAGACTGCCAACAAAGGTACTC 3900
TCATGCACCTACTGAATGACGACCGCCGGTAAGACTCGATCTGACGGTTGTTCCATGAG

3121 ACTTTGAGGGATCAGGAGCCCCAGGTTATGGGACCAATGGGAGCAGCCCTGCAAGTGTGA 3180
TGAAACTCCCTAGTCCTCGGGTCCAATACCCCTGGTACCCCGTCGGGACGTTACAACACT

3181 CCCTAAATATAGAAGATGAGCATCGGCTACATGAGACCTCAAAAGAGCCAGATGTTCTC 3240
GGGATT TATATCTCTACTCGTAGCCGATGTACTCTGGAGTTTCTCGGTCTACAAAGAG

3241 TAGGGTCCACATGGCTGTCTGATTTCTCGAGGCTGGCGGAAACGGGGGCATGGAC 3300
ATCCCAGGTGACCGACAGACTAAAGGAGTCCGGACCCGCTTGCCCCCGTACCCCTG

3' SS (3314)

3301 TGGCAGTTGCCAAGCTCCTCTGATCATACTCTGAAAGCAACCTCTACCCCGTGTCCA 3360
ACCGTCAAGCGGTTCGAGGAGACTAGTATGGAGACTTCTGGAGATGGGGCACAGGT

3361 TAAAACAATACCCATGTACAAGAAGCCAGACTGGGATCAAGCCCCACATACAGAGAC 3420
ATTTGTTATGGGTACAGTGTCTCGGTCTGACCCCTAGTTCTGGGTGTATGTCCTG

3421 TGTGGACCAGGAATACTGGTACCCCTGCCAGTCCCCCTGGAACACGGCCCTGCTACCCG 3480
ACAACCTGGTCCCTATGACCATGGGACGGTCAGGGGACCTGTGCGGGGACGATGGC

3481 TTAAGAAACCAAGGACTAATGATTATAGGCCTGTCCAGGATCTGAGAGAAAGTCAAAGC 3540
AATTCTTGGTCCCTGATTACTAATATCCGGACAGGTCTAGACTCTCTCAGTTGTCG

Фиг. 1D-8

3541 GGGTGGAAGACATCCACCCACCGTGCCAACCTTACAACCTTGAGCGGCTCCAC 3600
 CCCACCTCTGTAGGTGGGTTGGCACGGTTGGGAATGTTGGAGAACTCGCCCAGGGTG
 3601 CGTCCCACCAGTGGTACACTGTGCTTGATTAAAGGATGCCTTTCTGCCTGAGACTCC 3660
 GCAGGGTGGTCACCATGTGACACGAACTAACTACGGAAAAGACGGACTCTGAGG
 |
 PfeI
 |
 TfiI
 |
 3661 ACCCCACCAGTCAGCCTCTTCGCCTTGAGTGGAGAGATCCAGAGATGGGAATCTCAG 3720
 TGGGGTGGTCAGTCGGAGAGAAGCGGAAACTCACCTCTAGGTCTACCCCTAGAGTC
 |
 MfeI
 |
 MunI
 |
 3721 GACAATTGACCTGGACCAGACTCCCACAGGGTTCAAAAACAGTCCCACCTGTTGATG 3780
 CTGTTAACTGGACCTGGTCTGAGGGTGTCCCAAAGTTTGTAGGGTGGGACAAACTAC
 |
 MroI
 |
 BseAI
 |
 Bsp13I
 |
 BspEI
 |
 Kpn2I
 |
 AccIII
 |
 3781 AGGCACTGCACAGAGACCTAGCAGACTTCCGGATCCAGCACCCAGACTTGATCCTGCTAC 3840
 TCCGTGACGTGTCCTGGATCGTCTGAAGGCTAGTCGTGGTCTGAACTAGGACGATG
 3841 AGTACGTGGATGACTTACTGCTGGCCGCACTTCTGAGCTAGACTGCCAACAGGTACTC 3900
 TCATGCACCTACTGAATGACGACCGGCGGTGAAGACTCGATCTGACGGTTGTTCCATGAG

Фиг. 1D-9

3901 GGGCCCTGTTACAAACCTAGGAACCTCGGGTATCGGGCTCGGCCAAGAAAGCCAAA 3960
 CCGGGACAATGTTGGATCCCTGGAGCCCATAGCCCAGGAGCCGGTTCTTCGGTT
 3961 TTTGCCAGAACAGGTCAAGTATCTGGGTATCTCTAAAAGAGGGTCAGAGATGGCTGA 4020
 AACGGTCTTGTCCAGTCATAGACCCATAGAAGATTCTCCAGTCTACCGACT
 4021 CTGAGGCCAGAACAGAGACTGTGATGGGCAGCCTACTCCGAAGACCCCTCGACAACATAA 4080
 GACTCCGGTCTTCTCTGACACTACCCGTCGGATGAGGCTCTGGGAGCTGTGATT
 4081 GGGAGTTCCCTAGGGACGGCAGGCTCTGTCGCCCTCGGATCCCTGGGTTGAGAAATGG 4140
 CCCTCAAGGATCCCTGCCGCCAGACAGGGAGACCTAGGGACCCAAACGTCTTACC
 4141 CAGCCCCCTTGACCCCTCACAAAACGGGACTCTGTTAATTGGGGCCCAGACCAAC 4200
 GTCGGGGAAACATGGAGAGTGGTTTGGCCCTGAGACAAATTAAACCCGGGCTGGTTG
 4201 AAAAGGCCTATCAAGAAATCAAGCAAGCTCTCTAACCTGCCCCAGCCCTGGGTTGCCAG 4260
 TTTCCGGATAGTTCTTAGTCGAGAAGATTGACGGGTCGGACCCAAACGGTC

SalI

|

4261 ATTTGACTAAGCCCTTGAACTCTTGTGCGACGAGAACGCAGGGCTACGCCAAAGGTGTCC 4320
 TAAACTGATTGGAAACTTGAGAACAGCTGCTCTCGTCCGATGCGGTTCCACAGG
 4321 TAACGCAAAACGGACCTTGGCGTGGCGGTGGCTACCTGTCCAAGCTAGACC 4380
 ATTGCGTTTTGACCTGGAAACCGCAGCCGGCACCGGATGGACAGGTTTCGATCTGG
 4381 CAGTAGCAGCTGGTGGCCCTTGCCCTACGGATGGTAGCAGCCATTGCCGTACTGACAA 4440
 GTCATCGTCGACCCACCGGGGAAACGGATGCCCTACCATCGTGGTAACGGCATGACTGTT
 4441 AGGATGCAGGAAGCTAACCATGGGACAGCCACTAGTCATTGGCCCCCATGCAGTAG 4500
 TCCTACGTCCGTTGATTGGTACCCCTGTCGGTGTACGTAAGACCGGGGGTACGTAC
 4501 AGGCACTAGTCAAACAACCCCCCGACCGCTGGCTTCCAACGCCGGATGACTCACTATC 4560
 TCCGTGATCAGTTGTTGGGGCTGGCACCGAAAGGTTGGGGCTACTGAGTGATAG
 4561 AGGCCTTGCTTTGGACACGGACCGGGTCCAGTTGGACCCGGTGGTAGCCCTGAACCCGG 4620
 TCCGGAACGAAAACCTGTGCCCTGGCCAGGTCAAGCCTGGCACCATCGGACTTGGGCC
 4621 CTACGCTGCTCCACTGCCTGAGGAAGGGCTGCAACACAACGCCTGATATCCTGGCCG 4680
 GATGCGACGAGGGTGACGGACTCCTCCGACGTTGTTGACGGAACATAGGACCGGC
 4681 AAGCCCACGAAACCGACCCGACCTAACCGGACCGAGCCCTCCAGACGCCGACCACACCT 4740
 TTCCGGTGCCTGGCTGGCTGGATTGCCCTGGTCGGAGGGTCTGGCTGGTGGTGG
 4741 GGTACACGGATGGAAGCAGTCTTACAAGAGGGACAGCGTAAGGCGGGAGCTGCCGTGA 4800
 CCATGTGCCTACCTCGTCAGAGAATGTTCTCCCTGTCGCATTCCGCCCTGACGCCACT

Фиг. 1D-10

BspI
 |
 CeuII
 |
 Bpu1102I
 |
 Bsp1720I
 |

4801	CCACCGAGACCGAGGTAACTGGGCTAAAGCCCTGCCAGCCGGACATCCGCTCAGCGGG 4860 GGTGGCTCTGGCTCATTAGACCGATTGGGACGGTCGGCCCTGTAGGCAGTCGCC
4861	CTGAACTGATAGCACTCACCCAGGCCCTAAAGATGGCAGAAGGTAAAGACTAAATGTTT 4920 GACTTGACTATCGTAGTGAGTGGTCCGGGATTCTACCGCTTCCATTCTCGATTTACAAA
4921	ATACTGATAGCCGTTATGCTTTGCTACTGCCCATATCCATGGAGAAATATAAGAAGGC 4980 TATGACTATCGCAATACGAAACGATGACGGGTATAGGTACCTCTTATATGCTTCG
4981	GTGGGTTGCTCACATCAGAAGGCAAAGAGATCAAAATAAAAGACGAGATCTGGCCCTAC 5040 CACCCAAACGAGTGTAGTCTCCGTTCTAGTTTATTCGCTCTAGAACCGGGATG
5041	TAAAAGCCCTTTCTGCCAAAAGACTTACGATAATCCATTGTCAGGACATCAAAGG 5100 ATTTTCGGGAGAAAGACGGGTTTCTGAATCGTATTAGGTAAACAGGTCTGTAGTTTCC
5101	GACACAGCGCCGAGGCTAGAGGAACCGGATGGCTGACCAAGCGGCCGAAGGCAGCCA 5160 CTGTGTCGCGCTCCGATCTCCGTTGGCTACCGACTGGTCCGGCTTCCGCGT
5161	TCACAGAGACTCCAGACACCTCTACCCCTCTCATAGAAAATTCTACACCTACACCTCAG 5220 AGTGTCTCTGAGGCTGTGGAGATGGGAGGATCTTTAAGTAGTGGGATGTGGAGTC
5221	AACATTTTCATTACACAGTGACTGATATAAGGACCTAACCAAGTTGGGGCCATTATG 5280 TTGTAAAAGTAATGTGTCACTGACTATATTCCGTTGGATGGTCAACCCCCGGTAAATAC
5281	ATAAAACAAAAGAAGTATTGGGCTACCAAGGAAACCTGTGATGCCCTGACCAGTTACTT 5340 TATTTGTTCTCATAACCCAGATGGTCCTTTGGACACTACGGACTGGTCAAATGAA
5341	TTGAATTATTAGACTTCTTCATCAGCTGACTCACCTCAGCTCTCAAAATGAAGGCTC 5400 AACTTAATAATCTGAAAGAAGTAGTCGACTGAGTGGAGTCGAAGAGTTTACTCCGAG
5401	TCCTAGAGAGAACGCCACAGTCCTACTACATGCTGAACCGGGATCGAACACTCAAATA 5460 AGGATCTCTTCGGTGTAGGGATGATGACTGAGTGGCCCTAGCTTGAGTTTAT

Фиг. 1D-11

5461 TCACTGAGACCTGCAAAGCTGTGCACAAGTCACGCCAGCAAGTCTGCCGTTAACAGG 5520
 AGTGACTCTGGACGTTCGAACACGTGTTAGTTGCGGTCGTTCAGACGGCAATTGTCC

SacII
 |
 SgrBI
 |
 Cfr42I
 |
 Sfr303I
 |
 KspI
 |

5521 GAACTAGGGTCCGGCATCGGCCGGCACTCATTGGGAGATCGATTCACCGAGATAA 5580
 CTTGATCCCAGGCGCCGTAGCCGGCCGTAGTAACCTCTAGCTAAAGTGGCTCTATT

5581 AGCCCCGGATTGTATGGCTATAATATCTCTAGTTTATAGATACTTTCTGGCTGGA 5640
 TCGGGCCTAACATACCGATATTATAGAAGATCAAAATATCTATGAAAAGACCGACCT

5641 TAGAACCTTCCCACCAAGAAAGAACCGCCAAGGTCGTAACCAAGAAGCTACTAGAGG 5700
 ATCTCGGAAGGGTGGTTCTTCTTGGCGGGTCCAGCATTGGTCTTCGATGATCTCC

PaeI
 |
 BbvI
 |
 SpeHI
 |
 SphI
 |

5701 AGATCTTCCCAGGTTGGCATGCCCTCAGGTATTGGGAACGTGACAATGGGCCTGCCCTCG 5760
 TCTAGAAGGGGTCCAAGCCGTACGGAGTCCATAACCCTTGACTGTTACCCGGACGGAAGC

6901 CTTCCAAGGGGCTACTCGAGGGGGCAGATGCAACCCCTAGTCCTAGAAATTCACTGATGC 6960
 GAAGGTTCCCCGATGAGCTCCCGTCTACGTTGGGAGATCAGGATCTTAAAGTGAACGAC

6961 AGGAAAAAAGGCTAACTGGGACGGGCCAAATGTGGGACTGAGACTGTACCGGACAGG 7020
 TCCCTTTTCCGATTGACCCCTGCCGGGTTAGCACCCCTGACTCTGACATGGCCTGTC

7021 AACAGATCCTATTACCATGTTCTCCCTGACCCGGCAGGTCTTAATGTGGGACCCGAGT 7080
 TTGTCTAGGATAATGGTACAAGAGGGACTGGCGTCCAGGAATTACACCCGGGCTCA

7081 CCCCATAGGGCCCAACCCAGTATTACCGACCAAAGACTCCCTCCCTACCAATAGAGAT 7140
 GGGGTATCCC GGTTGGGTCTAAATGGGTGGTTCTGAGGGAAAGGAGTGGTTATCTCTA

7141 TGTACCGGCTCCACAGCCACCTAGCCCCCTCAATACCAAGTTACCCCCCTCCACTACCAAG 7200
 ACATGCCCGAGGTGTCGGTGGATCGGGGAGTTATGGTCAATGGGGGAAGGTGATGGTC

7201 TACACCCCTAACCTCCCTACAAGTCCAAGTGTCCCACAGCCACCCCCCAGGAACGGAGA 7260
 ATGTGGGAGTTGGAGGGGATGTTCAAGGTTACAGGGTGTGGTGGGGTCTTGACCTCT

7261 TAGACTACTAGCTCTAGTCAAAGGAGCCTATCAGGCCGTTAACCTACCAATCCGACAA 7320
 ATCTGATGATCGAGATCAAGTTCTCGGATAGTCCGGAATTGGAGTGGTTAGGGCTGTT

Фиг. 1D-12

027693

7321 GACCCAAGAATGTTGGCTGTGCTTAGTGTCTGGACCTCCTTATTACGAAGGAGTAGCGGT 7380
CTGGGTTCTTACAACCGACAGAACATCACAGCCTGGAGGAATAATGCTTCCTCATCGCCA

7381 CGTGGGCACTTATACCAATCATCCACCGCTCCGGCAACTGTACGGCCACTTCCAACA 7440
GCACCCGTGAATATGGTTAGTAAGGTGGCGAGGCCGGTGACATGCCGGTGAAGGGTTGT

EcoT22I
|
NsiI
|
Mph1103I
|
Zsp2I
|
BfrBI
| |

7441 TAAGCTTACCCCTATCTGAAGTGACAGGACAGGGCTATGCATGGGGCAGTACCTAAAAC 7500
ATTCAATGGATAGACTTCACTGTCTGCCCCATACGTACCCCCGTATGGATTTG

NaeI
|
PdiI
|
MrONI
| |
NgoMIV
| |

7501 TCACCAGGCCTTATGTAACACCACCAAAGCGCCGGCTCAGGATCCTACTACCTTGCAGC 7560
AGTGGTCCGGAATACATTGTGGTGGGTTCGCGGCCGAGTCCTAGGATGATGAAACGTCG

AleI
|
OliI
|

7561 ACCCGCCGGAACATGTGGGCTTGCAGCACTGGATTGACTCCCTGCTTGTCCACCGT 7620
TGGGCGGCCTTGTACACCGAACGTCGTGACCTAACTGAGGGACGAAACAGGTGGTCCA

7621 GCTCAATCTAACCAACAGATTATTGTGTATTAGTGAACCTCTGGCCCAGAGTAATTACCA 7680
CGAGTTAGATTGGTGTCTAATAACACATAATCAACTTGAGACCGGGTCTCATTAATGGT

Фиг. 1D-13

027693

6901 CTTCCAAGGGGCTACTCGAGGGGGCAGATGCAACCCTAGTCCTAGAATTCACTGATGC 6960
GAAGGTTCCCCGATGAGCTCCCCGTCTACGTTGGGAGATCAGGATCTTAAGTGACTACG
6961 AGGAAAAAAGGCTAACTGGGACGGGCCAAATCGTGGGACTGAGACTGTACCGGACAGG 7020
TCCTTTTTCGATTGACCTGCCGGTTAGCACCCCTGACTCTGACATGGCTGTCC
7021 AACAGATCCTATTACCATGTTCTCCCTGACCCGGCAGGTCTTAATGTGGGACCCCAGT 7080
TTGTCTAGGATAATGGTACAAGAGGGACTGGCCGTCCAGGAATTACACCCCTGGGGCTCA
7081 CCCCATAGGGCCAACCCAGTATTACCGACCAAAGACTCCCTCCTACCAATAGAGAT 7140
GGGGTATCCCCGGTTGGGTCTATAATGGGCTGGTTCTGAGGGAAAGGAGTGGTTATCTCTA
7141 TGTACCGGCTCACAGCCACCTAGCCCCCTCAATACCAAGTTACCCCCCTTCCACTACCAG 7200
ACATGGCCGAGGTGTCGGTGGATCGGGGAGTTATGGTCAATGGGGGAAGGTGATGGTC
7201 TACACCCCTCAACCTCCCCTACAAGTCCAAGTGTCCCACAGCCACCCCCAGGAACGTGGAGA 7260
ATGTGGGAGTTGGAGGGGATGTTCAAGGTCACAGGGTGTGGTGGGGTCTTGACCTCT
7261 TAGACTACTAGCTCTAGTCAAAGGAGCCTATCAGGCCTTAACCTCACCAATCCCACAA 7320
ATCTGATGATCGAGATCAGTTCCCTCGGATAGTCGGCGAATTGGAGTGGTTAGGGCTGTT
7321 GACCCAAGAATGTTGGCTGTGCTTAGTGTGGGACCTCTTATTACGAAGGGTAGCGGT 7380
CTGGGTTCTTACAACCGACACGAATCACAGCCCTGGAGGAATAATGCTCCTCATGCCA
7381 CGTGGGCACTTATACCAATCATTCCACCGCTCGGCCAACTGTACGGCCACTCCAAACA 7440
GCACCCGTGAATATGGTTAGTAAGGTGGGAGGCCGGTTGACATGCCGGTGAAGGGTTGT
7441 TAAGCTTACCCATATCTGAAGTGACAGGACAGGGCTATGCATGGGGCAGTACCTAAAC 7500
ATTGGAATGGGATAGACTTCAGTGTCTGTCGGGATACGTACCCCGTCATGGATTTG
7501 TCACCAGGCCTATGTAACACCACCCAAAGCGCCGGCTCAGGATCCTACTACCTGCAGC 7560
AGTGGTCCGAATACATTGTTGGGTTCGCGGCCGAGTCCTAGGATGATGGAACGTCG

EcoT22I
|
NsiI
|
Mph1103I
|
Zsp2I
|
BfrBI
| |
7441 TAAGCTTACCCATATCTGAAGTGACAGGACAGGGCTATGCATGGGGCAGTACCTAAAC 7500
ATTGGAATGGGATAGACTTCAGTGTCTGTCGGGATACGTACCCCGTCATGGATTTG
7501 TCACCAGGCCTATGTAACACCACCCAAAGCGCCGGCTCAGGATCCTACTACCTGCAGC 7560
AGTGGTCCGAATACATTGTTGGGTTCGCGGCCGAGTCCTAGGATGATGGAACGTCG

Фиг. 1D-14

AleI
|
OliI
|

7561 ACCCGCCGGAACAATGGGGCTTGCAGCACTGGATTGACTCCCTGCTTGTCCACCAACGGT 7620
 TGGGCGGCCCTTGTACACCCGAACTGCTGACCTAAGTGGGACGAAACAGGTGGTGCCA

 7621 GCTCAATCTAACCAACAGATTATTGTGTATTAGTTGAACCTCTGGCCCAAGAGTAATTACCA 7680
 CGAGTTAGATTGGTGTCTAAACACATAATCAACTGAGACCGGGTCTATTAAATGGT
 7681 CTCCCCGATTATATGTATGGTCAGCTGAACAGCGTACCAAATATAAAAGAGAGCCAGT 7740
 GAGGGGGCTAATATACATACCAAGTCGAACATTGTCGATGGTTATATTTCTCGGTCA

 7741 ATCATTGACCCCTGGCCCTTCACTAGGAGGATTAACCATGGGAGGGATTGCAGCTGAAT 7800
 TAGTAACCTGGGACCGGAAAGATGATCCTCTAAATTGGTACCCCTCCCTAACGTCGACCTA

 7801 AGGGACGGGGACCACTGCCTTAATTAAAACCCAGCAGTTGAGCAGCTCATGCCGTAT 7860
 TCCCTGCCCTGGTGACGGAAATTATTTGGTCTGCAAACCTCGTCAAGTACGGCGATA

 7861 CCAGACAGACCTCAACGAAGTCGAAAAGTCATTACCAACCTAGAAAAGTCAGTACGACCTC 7920
 GGTCCTGCTGGAGTTGCTCAGTTTCAGTTAATGGTGGATCTTCAGTCAGTGGAG

 7921 GTTGTCTGAAGTAGTCCTACAGAACCGCAGAGGCCTAGATTGCTATTCTAAAGGAGGG 7980
 CAACAGACTTCATCAGGATGTCCTGGCGCTCCGGATCTAAACGATAAGGATTCTCCC

 7981 AGGTCTCTGCCGAGCCCTAAAAAGAAGAATGTTTATGCAGACACACGGGGCTAGT 8040
 TCCAGAGACGCGTCGGATTCTTACAACAAAACGTCTGGTGTGCCCGATCA

 8041 GAGAGACAGCATGCCAAATTAAAGAGAAGGCTTAATCAGAGACAAAAACTATTGAGAC 8100
 CTCTCTGTCGTACCGGTTAATTCTCTTCGAATTAGTCCTGTTTGATAAAACTCTG

NspV
|
BstBI
|
Bsp119I
|
AsuII
|
Csp45I
|
SfuI
|
Bpu14I
|
BspT104I
|

8101 AGGCCAAGGATGGTTCGAAGGGCTGTTAATAGATCCCCCTGGTTACCACTTAATCTC 8160
 TCCGGTTCTACCAAGCTTCCCGACAAATTATCTAGGGGGACCAATGGTGAATTAGAG

Фиг. 1D-15

8161 CACCATCATGGGACCTCTAATAGTACTCTTACTGATCTTACTCTTGACCTGCATTCT 8220
 GTGGTAGTACCCCTGGAGATTATCATGAGAAATGACTAGAATGAGAAACCTGGAACGTAAGA
 8221 CAATCGATTGGTCAATTGTTAAAGACAGGATCTCAGTGGTCCAGGCTCTGGTTTGAC 8280
 GTTAGCTAACCAACAGGTTAAACAATTCTGTCTAGAGTCACCAGGTCGAGACCAAACTG
 IRES reg(8327, 8876)>>>
 |
 MluI(8325)
 |
 |
 8281 TCAGCAATATCACCAAGCTAAAACCCATAGAGTACCGAGCCATGAACGCCGTTACTGGCCCAA 8340
 AGTCGTTATAGGGTCGATTTGGGTATCTCATGCTCGGTACTTGCACATGACCGGCTT
 ires_emcv reg(8378, 8876)>>>
 |
 8341 GCGCCTTGGAAATAAGGCCGGTGCCTTGTCTATATGTATTTCCACCATATTGCCGT 8400
 CGGCGAACCTTATTCCGCCAACACGCAAACAGATATAACAATAAAAGGTGGTATAACGCA
 8401 CTTTGGCAATGTGAGGGCCCGAAACCTGGCCCTGTCTCTGACGAGCATTCCCTAGGG 8460
 GAAAACCGTTACACTCCC GGCTTTGGACCGGGACAGAAGAACTGCTCGTAAGGATCCC
 8461 GTCTTCCCCTCTGCCAACAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTC 8520
 CAGAAAGGGGAGCGGGTTCCCTACGTTCCAGACAACCTACGCACTTCCTCGTCAAG
 8521 CTCTGGAAGCTCTTGAAGACAAACACCTCTGTAGCGACCCCTTGCAGGCAGCGGAACC 8580
 GAGACCTTCGAAGAACCTCTGTTGTCAGACATCGCTGGAAACGTCGCGCTTGG
 8581 CCCCACCTGGGACAGGTGCCTCTGCCAAAAGCCACGTGTATAAGATAACCTGCAA 8640
 GGGGTGGACCGCTGTCCACGGAGACGCCGGTTTCGGTGCACATATTCTATGTGGACGTT
 8641 AGGC GGCAACCCCCAGTGCCACGTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGC 8700
 TCCGCCGTGTTGGGTCAAGGTCAACACTCAACACCTATCAACACCTTCTCAGTTACCG
 8701 TCTCCTCAACCGTATTCAACAAGGGGCTGAAGGATGCCAGAAGGTACCCATTGTATGG 8760
 AGAGGAGTTCCATAAGTTCTCCCGACTTCCACGGGTCTTCCATGGGTAAACATACC
 8761 GATCTGATCTGGGGCCTCGGTGCACATGCTTACATGTGTTAGTCGAGGTTAAAAAAAC 8820
 CTAGACTAGACCCCCGGAGCCACGTGTACCAAATGTACAAATCAGCTCCAATTTCCTG

Фиг. 1D-16

yCD2 (8877, 9353) >>>

|

PsiI (8874)

|

|

8821 |GCTCTAGGCCCCCGAACCACGGGACGTGTTTCCCTTGAAACACGATTATAATGC 8830
CAGATCCGGGGGCTTGGTCCCCCTGCACCAAAAGGAAACTTTCTGCTAATATTACG|

8881 |TCACCGGGCGCATGGCCTCCAAGTGGCATCAAAGGGCATGGATATGCTTACGGAGGAGG 8940
ACTGGCCGCCGTACCGGACGTTACCCCTAGTTCCCGTACCTATAGCGAATGCTCTCC|

8941 |CCCTGCTGGCCTACAGGAGGGCGGCTTCCCTATCGGGCTCTCTGATCAACAAAGG 9000
GGGACCAACCCGATGTTCCCTCCGCCGACGGATAGCCCGACAGACTACTTGTTCTCC|

9001 |ACGGCAGTGTCTGGCCAGGGCCACACATGAGGTTCCAGAAGGGCTCCGCCACCTGC 9060
TCCCGTACAACGACCCGTCCCCGGTGTCTACTCCAAGGTCTTCCCGAGGGCGTGGGACG|

9061 |ACGGCGAGATCTCCACCCCTGGAGAACTGTGGCAGGCTGGAGGCAAGGTGTAACAGGACA 9120
TGCCGCTCTAGGGTGGACCTCTTGACACCGTCCGACCTCCCGTCCACATGTTCTGT|

9121 |CCACCCCTCTACACCAACCTGCTCCCTTGACATGTGTAACGGGGCTATCATCATGTAACG 9180
CGTGGGACATGTGGTGGACAGGGGAACTGTACACATGGCCCGATAGTACTACATGC|

9181 |GCATCCCTAGGTGTGATGCCGAGAACGTGAACCTCAACTCCAAGGGGAGAACATACC 9240
CGTAGGGATCCACACACTACCCGCTTGCACCTTGAAAGTCAGGTCCCGCTTCAATGG|

9241 |TCCAAACCAGGGGCCACGAGGTCTGGTTGACCATGAGGGTGTAAAGAAGCTGATGA 9300
ACGTTTGGTCCCCGTGCTCCACCAACAACTGCTACTCTCACATCTTGACTACT|

NciI (9356)

|

CciNI

|

9301 |AGCAGTTCATCGACGAGAGGCCCTAGGACTGGTCCAGGATATGGCGAGTAAGCGGGCG 9360
TCGTCAAGTAGCTGCTCTCCCGAGTCTGACCAAGCTCTATAGCCGCTCATTCGCCGCC|

U3 область (9405, 9854) >>>

|

9361 |CAGATAAAATAAAAGATTTATTTAGTCTCCAGAAAAGGGGGGAATGAAAGACCCCACC 9420
GCTATTTATTTCTAAATAATCAGAGGTCTTTCCCCCTTACTTCTGGGTGG

Фиг. 1D-17

Фиг. 1D-18

pBR322 начало (10045,10664)<<<

10021 AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTCCATAGGCTCCGCCCCC 10080
TTCCGGTCCTGGCATTTCCGGCGAACGACCGAAAAAGGTATCCGAGGCAGGGGGGGG

10081 TGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCAAACCCGACAGGACTATA 10140
ACTGCTCGTAGTGTAGCTCGAGTTCAAGTCCTCCACCGCTTGGGCTGTCCTGATAT

10141 AAGATACCAGGGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCC 10200
TTCTATGGTCCGAAAGGGGACCTTCGAGGGAGCACCGAGAGGACAAGGCTGGACGG

10201 GCCTTACCGGATACCTGTCGGCCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 10260
CGAATGGCCTATGGACAGGGAAAGAGGGAAAGCCCTCGCACCGCAAAGAGTTACGAG

10261 ACGCTGTAGGTATCTCAGTCGGTGTAGGTCGTTCGCTCCAAGCTGGCTGTGACGA 10320
TGCACATCCATAGAGTCAGGCCACATCCAGCAAGCGAGGTTGACCCGACACACGTGCT

10321 ACCCCCCGTTCAGCCGACCGCTGCGCCTATCCGTAACATCGTCTGAGTCCAACCC 10380
TGGGGGGCAAGTCGGCTGGCGACGCCATTGATAGCAGAACTCAGGTTGGG

10381 GGTAAGACGACTTATGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 10440
CCATTCTGTGCTGAATAGCGGTGACCGTCGCGTGACCATTCGCTTAATCGTCTCGCTC

10441 GTATGTTAGGCGGTGCTACAGAGTTCTGAAAGTGGTGGCCTAACTACGGCTACACTAGAAAG 10500
CATACATCCGCCACGATGTCTCAAGAACTCACCACGGATTGATGCCGATGTGATCTTC

10501 GACAGTATTGGTATCGCCTCTGCTGAAGGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 10560
CTGTCATAAACCATAGACGCCAGACGACTCGGTCAATGGAAGCCCTTCTCAACCATC

10561 CTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTGTTGCAAGCAGCA 10620
GAGAACTAGGCCGTTGGTGGCGACCATGCCACCAAAAAACAAACGTTCGTCG

10621 GATTACGCCAGAAAAAGATCTCAAGAAGATCCTTGATCTTCTACGGGCTGA 10680
CTAATGCCGCTTTCTAGAGTTCTAGGAAACTAGAAAAGATGCCCGACT

10681 CGCTCAGTGGAACGAAAACCTACCGTTAAGGGATTTGGTCATGAGATTATCAAAAGGAT 10740
GCGAGTCACCTGCTTTGAGTGCATTCCCTAAACAGTACTCTAATAGTTTCTA

10741 CTTCACCTAGATCCTTAAATTAAAAATGAAGTTAAATCAATCTAAAGTATATATGA 10800
GAAGTGGATCTAGGAAATTAAATTCTAAACTTAAAGTTAGTTAGATTCATATATACT

Фиг. 1D-19

027693

amp маркер (10819,11679) <<<

10801 GTAAACTTGGCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTG 10860
CATTTGAACCAGACTGTCAATGGTTACGAATTAGTCACTCCGGATAGAGTCGCTAGAC

10861 TCTATTTCGTCATCCATAGTTGCCCTGACTCCCCCTCGTGTAGATAACTACGATAACGGGA 10920
AGATAAAGCAAGTAGGTATCACCGACTGAGGGCAGCACATCTATTGATGCTATGCCCT

10921 GGGCTTACCATCTGGCCCCAGTGTGCAATGATAACCGCGAGACCCACGCTCACCGCTCC 10980
CCCGAATGGTAGACCGGGGTACGACGTTACTATGGCGCTCGGGTGCAGTGGCCGAGG

10981 AGATTATCAGCAATAAACCCAGCCAGCCGAAGGGCCGAGCAGCAGAAGTGGCTCTGCAAC 11040
TCTAAATAGTCGTTATTGGTCGGTCGGCCTCCGGCTCGCTTCACCAGGACGTTG

11041 TTATCCGCCTCCATCCAGTCTATTATTGTCGGGGAAAGCTAGAGTAAGTAGTCGCC 11100
AAATAGCGCGAGGTAGGTAGATAATTAACAACGGCCCTCGATCTCATTCAACAGCGG

11101 AGTTAATAGTTGCGAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTACGCTCGTC 11160
TCAATTATCAAACCGCTGCAACAAACGTAACGACGTCGTTAGCACCACAGTGCAGCAG

11161 GTTGGTATGGCTCATTCAAGCTCCGGTCCCAACGATCAAGGGAGTTACATGATCCCC 11220
CAAACCATAACCGAAGTAAGTCGAGGCCAAGGGTTGCTAGTCCGCTCAATGTAAGGGG

11221 CATGTTGTGCAAAAAGCGGTTAGCTCCTCGGTCTCCGATCGTTGTCAGAAGTAAGTT 11280
GTACAACACGTTTCGCAATCGAGGAAGCCAGGAGGCTAGCAACAGTCTTCATTCAA

11281 GGCGCTCAGTGTATCACTCATGGTTATGGCAGCACTGCATAATTCTTACTGTATGCC 11340
CCGGCGTCACAATAGTGAGTACCAATACCGTCGTCAGTATTAAGAGAAATGACAGTACGG

11341 ATCCGTAAGATGCTTTCTGTGACTGGTAGTACTCAACCAAGTCATTCTGAGAATAGTG 11400
TAGGCATTCTACGAAAAGACACTGACCACCATGAGTTGGTCAAGTAAAGACTCTTATCAC

11401 TATGCGGCACCGAGTTGCTCTGCCCGCGTCAACACGGGATAATACCGGCCACATAG 11460
ATACGCCGCTGGCTCAACGAGAACGGCCGAGTTGTCGCCATTATGGCGCGGTATC

11461 CAGAACTTAAAGTGTCACTATTGAAAACGTTCTCGGGGAAACTCTAAGGAT 11520
GTCTTGAATTTACGAGTAGTAACCTTTGCAAGAACGGCCGTTTGAGACTTCCTA

11521 CTTACCGCTGTGAGATCCAGTTGATGTAACCCACTCGTGCACCCACTGATCTCAGC 11580
GAATGGGACAACCTCTAGGTCAGCTACATTGGTGAGCACGTGGTTGACTAGAAAGTCG

11581 ATCTTTACTTCAACCAGCGTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGAAA 11640
TAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTGTCCCTCGTTACGGGTTT

11641 AAAGGAAATAAGGGCGACACGGAAATGTTGAATCTACTCTTCCTTTCAATATTA 11700
TTTCCTTATTCCCGCTGTGCCCTTACAACTTATGAGTATGAGAAGGAAAAGTTATAAT

Фиг. 1D-20

amp промотор (11721,11749) <<<

11701 TTGAAGCATTATCAGGGTTATTGTCATGAGCGGATAACATATTGAAATGTATTAGAA 11760
AACTTCGTAATAGTCCCAATAACAGAGTACTCGCCTATGTATAAACTTACATAATCTT

11761 AAATAAACAAATAGGGTTCCGCGCACATTCCCCGAAAAGTGCCACCTGACGTCAAGA 11820
TTTATTGTTTATCCCAAGGCGCGTGTAAAGGGCTTTCACGGTGGACTGCAGATTCT

11821 AACCAATTATCATGACATTACCTATAAAATAGGCATACGGAGGCCCTTCGTCT 11880
TTGGTAATAATAGTACTGTAATTGGATATTTCATCGCATAGTGCCTGGAAAGCAGA

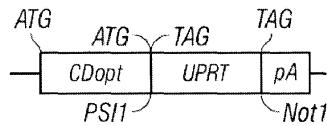
11881 TCAAGAATTCA 11892
AGTTCTTAAGTA

Фиг. 1D-21

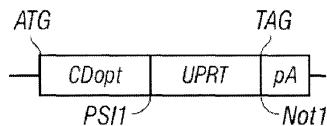
1. Отдельный синтез CDopt и UPRT генов с сайтами *PSI1/Not1*



2. Расщепление *PSI1* и проведение лигирования слепых концов



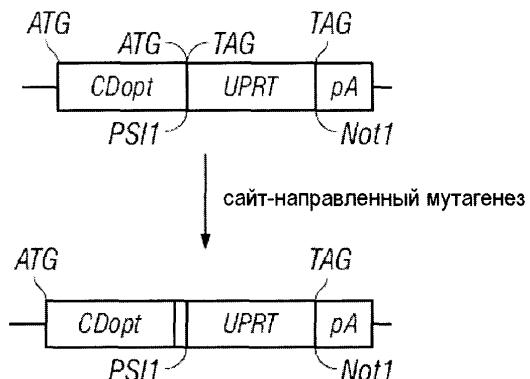
3. Делекция СТОП кодона гена CDopt и СТАРТ кодона UPRT путем сайт-направленного мутагенеза



4. Размер гена CDopt – UPRT: 1,3 тыс. осн.

Фиг. 2А

1. Синтез CDopt–ЛИНКЕР–UPRT: встраивание линкера сайтом-направленным мутагенезом



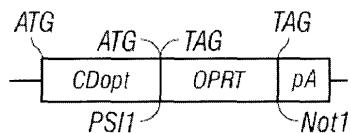
2. Размер гена CDopt–ЛИНКЕР–UPRT: 1,36 тыс. осн.

Фиг. 2В

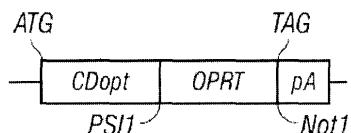
1. Отдельный синтез гена CDopt и домена OPRT с сайтами PstI/Not1



2. Расщепление PstI и проведение лигирования слепых концов



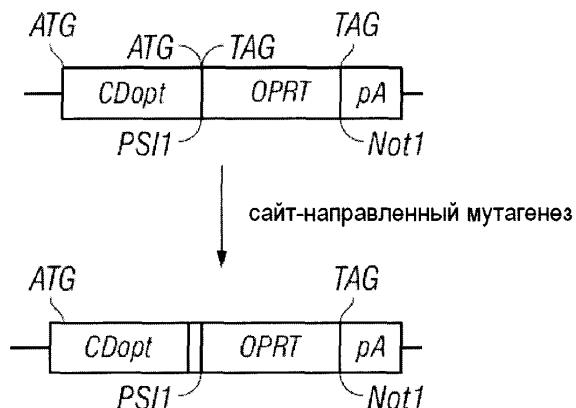
3. Делекция СТОП кодона гена CDopt и СТАРТ кодона OPRT путем сайт-направленного мутагенеза



4. Размер гена CDopt – OPRT: 1,3 тыс. осн.

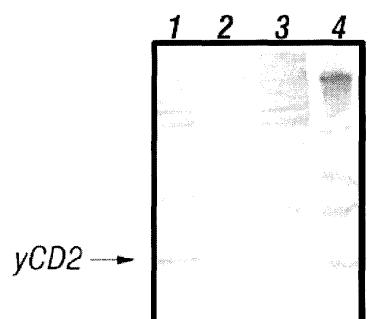
Фиг. 2С

1. Синтез CDopt-ЛИНКЕР-OPRT: встраивание линкера сайт-направленным мутагенезом



2. Размер гена CD-OPRT: 1,33 тыс. осн.

Фиг. 2D



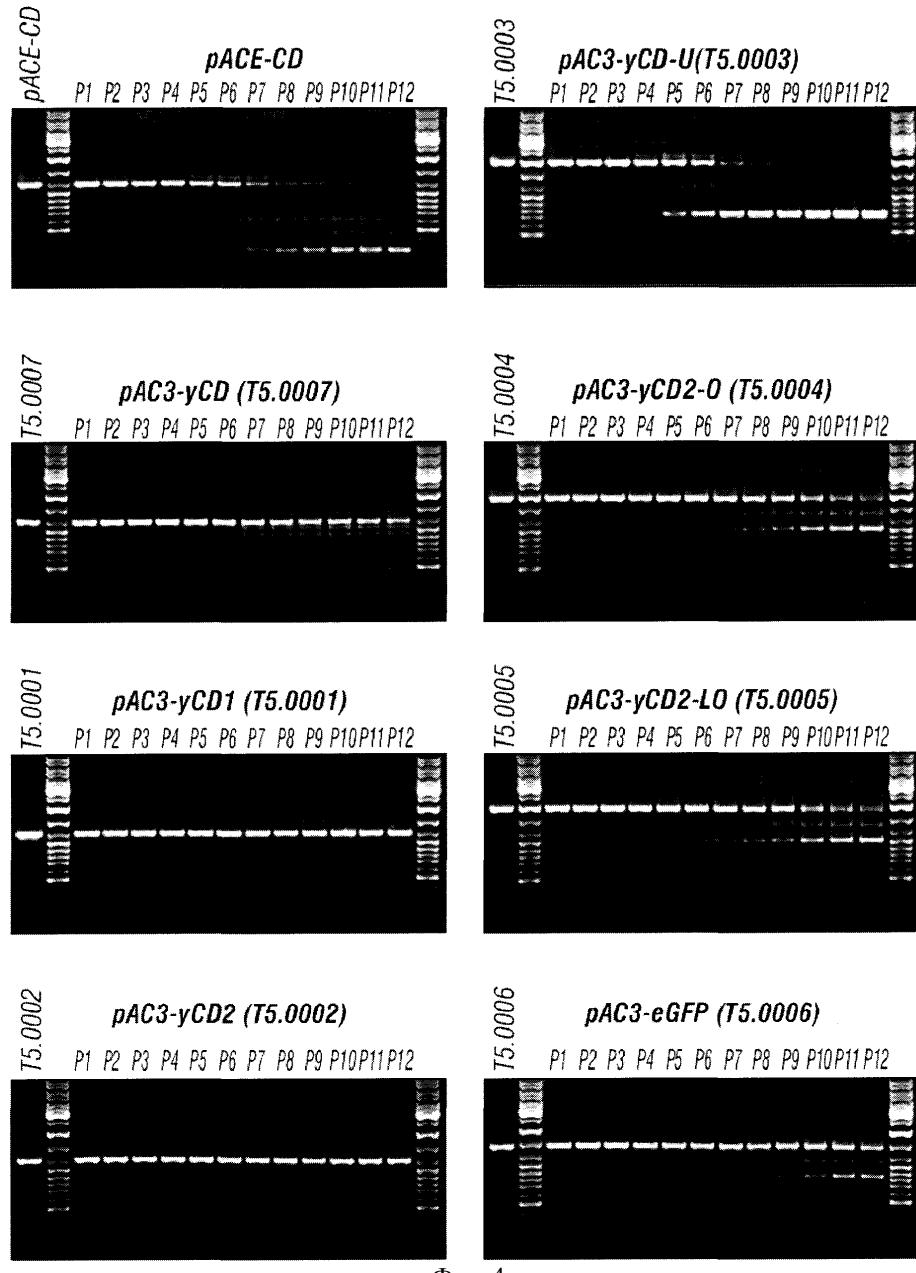
1 - U87 + AC3-yCD2 (V)

2 - U87 + ACE-yCD (V)

3 - U87 (неинфицированный)

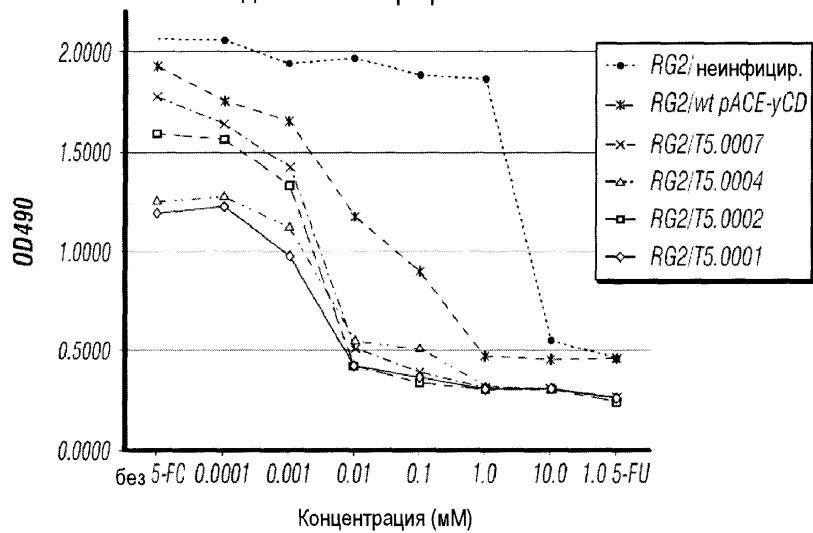
4 - стандартная молекулярная масса

Фиг. 3

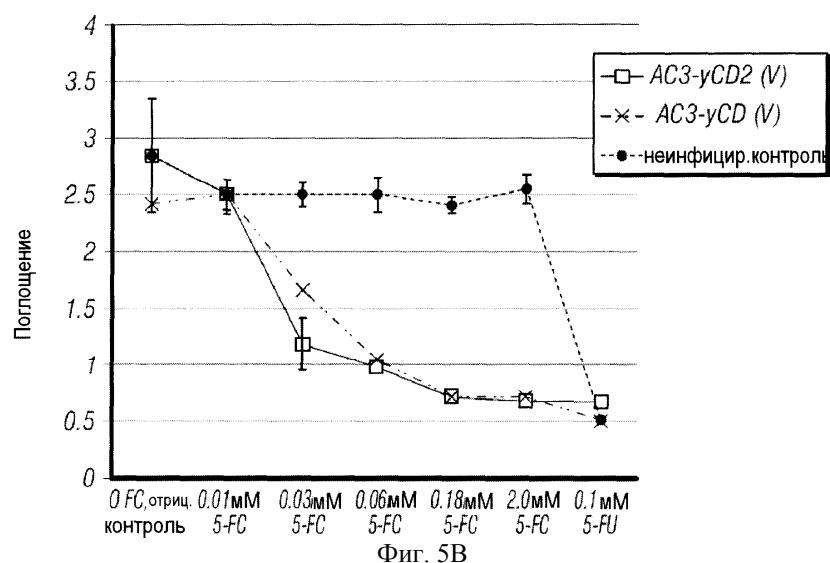


Фиг. 4

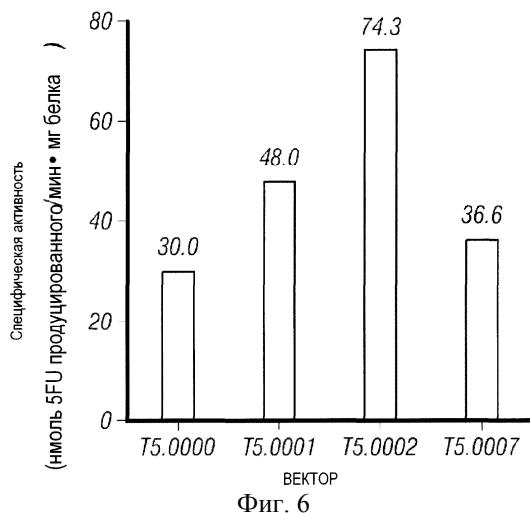
RG2/день 8 Вектор сравнения



Фиг. 5А



Фиг. 5В



Фиг. 6

