

(19)



**Евразийское
патентное
ведомство**

(21) **202392402** (13) **A1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки
2023.10.24

(51) Int. Cl. *C12N 15/86* (2006.01)
C07K 14/005 (2006.01)

(22) Дата подачи заявки
2022.02.25

(54) **РЕТРОВИРУСНЫЕ ВЕКТОРЫ**

(31) 2102832.9

(72) Изобретатель:

(32) 2021.02.26

Джилл Дебора, Хайд Стивен (GB)

(33) GB

(74) Представитель:

(86) PCT/GB2022/050524

Медведев В.Н. (RU)

(87) WO 2022/180411 2022.09.01

(71) Заявитель:

АйПи2АйПиО ИННОВЕЙШНЗ
ЛИМИТЕД (GB)

(57) Изобретение относится к ретровирусным векторам для переноса генов, в частности к лентивирусным векторам, псевдотипированным гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, и содержащим промотор и трансген; и к способам их получения. Настоящее изобретение также относится к применению указанных векторов в генотерапии, в частности для лечения заболеваний дыхательных путей, таких как кистозный фиброз (CF).

A1

202392402

202392402

A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-578822EA/042

РЕТРОВИРУСНЫЕ ВЕКТОРЫ

Настоящее изобретение относится к ретровирусным векторам для переноса генов, а в частности, к лентивирусным векторам, псевдотипированным гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, и содержащим промотор и трансген; и к способам их получения.

Предпосылки к созданию изобретения

Ретровирусы принадлежат к семейству РНК-вирусов (Retroviridae), которые кодируют фермент обратную транскриптазу. Лентивирусы относятся к роду семейства Retroviridae и характеризуются длительным инкубационным периодом. Ретровирусы, а в частности, лентивирусы, могут доставлять значительное количество вирусной РНК в ДНК клетки-хозяина и среди ретровирусов обладают уникальной способностью инфицировать неделящиеся клетки, поэтому они могут быть использованы в одном из наиболее эффективных методов, проводимых с применением вектора для доставки генов.

Псевдотипирование представляет собой процесс продукции вирусов или вирусных векторов в комбинации с чужеродными белками оболочки вируса. Таким образом, чужеродные белки оболочки вируса могут быть использованы для изменения тропизма хозяина или повышения/снижения стабильности вирусных частиц. Так, например, псевдотипирование позволяет уточнить свойства белков оболочки. Белком, часто используемым для псевдотипирования ретровирусных и лентивирусных векторов, является гликопротеин G вируса везикулярного стоматита (VSV), сокращенно VSV-G.

Лентивирусные векторы, особенно производные ВИЧ-1, представляют собой широко изученные и часто используемые векторы. Эволюция остова лентивирусных векторов и способность вирусов доставлять молекулы рекомбинантной ДНК (трансгены) в клетки-мишени позволяют использовать их во многих целях. Два возможных применения вирусных векторов включают восстановление функциональных генов в генотерапии и получение рекомбинантных белков *in vitro*.

При разработке ретровирусных/лентивирусных векторов, подходящих для применения в качестве векторов для доставки генов, одним из ключевых факторов является обеспечение максимально возможной безопасности вектора для пациентов. Вторым ключевым фактором является необходимость продукции достаточного количества вектора не только для лечения конкретного пациента, но и для обеспечения более широкого клинического доступа к терапии для всех пациентов, для которых она может быть эффективна. Эти два фактора могут конкурировать между собой, поскольку модификации, повышающие безопасность векторов, часто связаны со снижением выхода при получении векторов.

Одним из примеров клинических условий, в которых перенос генов в эпителий дыхательных путей будет давать позитивный эффект, является лечение кистозного фиброза (CF). CF является смертельным генетическим заболеванием, вызванным

мутациями в гене трансмембранного регулятора проводимости при CF (CFTR), который действует как хлоридный канал в эпителиальных клетках дыхательных путей. CF характеризуется рецидивирующими инфекциями грудной клетки, повышенной секрецией дыхательных путей и, в конечном итоге, дыхательной недостаточностью. В Великобритании в настоящее время средний возраст летального исхода от такого заболевания составляет ~25 лет. Для большинства генотипов не существует какого-либо лечения основного дефекта; при этом, современные методы лечения, а именно, симптоматического облегчения требуют нескольких часов самостоятельной терапии ежедневно. Генотерапия, в отличие от терапии низкомолекулярными препаратами, не зависит от класса мутаций CFTR и, таким образом, применима ко всем индивидуумам, страдающим CF. Однако, на сегодняшний день не существует вирусных векторов, одобренных для клинического применения при лечении CF, и то же самое относится и к другим заболеваниям, особенно, ко многим другим заболеваниям дыхательных путей.

Помимо проблем с безопасностью пациентов и выходом, существуют и другие проблемы, обычно связанные с переносом генов в эпителий дыхательных путей.

Эффективность переноса генов в эпителий дыхательных путей обычно является низкой, по меньшей мере частично, поскольку соответствующие рецепторы для многих вирусных векторов, по-видимому, преимущественно локализованы на базолатеральной поверхности эпителия дыхательных путей. Таким образом, до исследования, проведенного авторами настоящего изобретения, использование лентивирусных псевдотипов требовало нарушения целостности эпителия для трансдукции дыхательных путей, например, с помощью детергентов, таких как лизофосфатидилхолин или этиленгликоль-бис(2-аминоэтиловый эфир)-N, N,N',N'-тетрауксусная кислота, что было связано с повышенным риском сепсиса. Кроме того, обычные векторы для переноса генов с трудом проникают через слой слизистой оболочки дыхательных путей, что также снижает эффективность переноса генов. Возможность многократного введения обычных вирусных векторов, необходимого для обработки самообновляющегося эпителия в течение всей жизни, ограничена из-за адаптивных иммунных реакций пациентов, которые препятствуют успешному повторному введению.

Еще одним важным фактором является введение векторов для клинического применения. Следовательно, для эффективного лечения необходимо поддерживать стабильность вируса при использовании клинически ценных устройств (например, бронхоскопа и аэрозольного ингалятора).

Соответственно, существует потребность в векторе для генотерапии, который позволил бы решить одну или более проблем, описанных выше. В частности, целью настоящего изобретения является разработка способа получения псевдотипированного ретровирусного или лентивирусного (например, SIV) вектора и средства для осуществления указанного способа, где полученный вектор будет безопасным и адаптированным для повышения эффективности переноса генов через эпителий дыхательных путей и будет производиться в клинически значимом масштабе.

Сущность изобретения

Авторы настоящего изобретения ранее разработали лентивирусный вектор, который был псевдотипирован гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, и который содержит промотор и трансген. Обычно остов вектора состоит из вируса иммунодефицита обезьян (SIV), такого как SIV1 или SIV африканской зеленой мартышки (SIV-AGM). Предпочтительно, остов вирусного вектора согласно изобретению происходит от SIV-AGM. Белки HN и F функционируют соответственно, посредством присоединения к сиаловым кислотам и опосредования слияния клеток для проникновения вектора в клетки-мишени. Авторами настоящего изобретения было обнаружено, что этот специфически F/HN-псевдотипированный лентивирусный вектор может эффективно трансдуцировать эпителий дыхательных путей, что будет приводить к экспрессии трансгена, поддерживаемой в течение периодов времени, превышающих предполагаемую продолжительность жизни эпителиальных клеток дыхательных путей. Важно отметить, что авторы настоящего изобретения также обнаружили, что повторное введение не приводит к потере эффективности. Эти особенности делают векторы согласно изобретению привлекательными кандидатами для лечения заболеваний благодаря их использованию для экспрессии терапевтических белков: (i) в клетках дыхательных путей; (ii) их секрети в просвет дыхательных путей; и (iii) секрети в кровеносную систему.

Однако у этого лентивирусного вектора были потенциальные проблемы с точки зрения безопасности. В частности, между геномным вектором и вектором GagPol, используемым для его получения, наблюдалась значительная степень гомологии последовательностей. Эта гомология последовательностей создает теоретический риск того, что компетентный по репликации лентивирус (RCL) может образовываться либо во время продукции, либо при клиническом применении после введения пациенту. Это представляет угрозу безопасности пациента. Риск образования способных к репликации вирусных частиц является проблемой и для других ретровирусных/лентивирусных векторов.

Хотя было бы желательно уменьшить такой риск, однако сделать это непросто или, по меньшей мере не без выявления других неприемлемых недостатков. В частности, специалистами в данной области установлено, что модификации, направленные на снижение риска образования RCL, такие как оптимизация по кодонам получаемых генов gag-pol, обычно, негативно влияют на титр или выход вектора. Учитывая большие титры вектора, необходимые для лечения даже одного пациента, такое снижение выхода потенциально может сделать его производство коммерчески невыполнимым.

Авторы настоящего изобретения впервые продемонстрировали, что использование оптимизированных по кодонам генов gal-pol из SIV не оказывает негативного влияния на полученный титр вектора SIV, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса и может даже привести к увеличению титра вектора. Это неожиданно, учитывая то, что при обычных

производственных условиях (если ограничивающим фактором является плаزمид векторного генома, а не гены gag-pol), оптимизация по кодомам генов gag-pol обычно снижает выход вектора.

Таким образом, авторы настоящего изобретения первыми предложили разработать способ получения ретровирусного, в частности, лентивирусного вектора, такого как SIV, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса с пониженным риском образования RCL, не оказывая при этом, негативного влияния на титр вектора или даже повышая его. Таким образом, способы согласно изобретению будут обеспечивать получение более безопасных векторов с коммерчески желательными выходами.

Соответственно, настоящее изобретение относится к способу получения ретровирусного вектора, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми белками (F) респираторного парамиксовируса, где указанный вектор содержит промотор и трансген, причем, указанный способ включает использование оптимизированных по кодомам генов gag-pol. Предпочтительно, ретровирусный вектор представляет собой лентивирусный вектор, и необязательно, лентивирусный вектор выбирают из группы, состоящей из вектора на основе вируса иммунодефицита обезьян (SIV), вектора на основе вируса иммунодефицита человека (ВИЧ), вектора на основе вируса иммунодефицита кошек (FIV), вектора на основе вируса инфекционной анемии лошадей (EIAV) и вектора на основе вируса висны/maedi. Особенно предпочтительными являются способы получения вектора SIV.

Оптимизированные по кодомам гены gag-pol могут представлять собой гены gag-pol SIV. Оптимизированные по кодомам гены gag-pol могут содержать последовательность нуклеиновой кислоты или состоять из последовательности нуклеиновой кислоты, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 1. Оптимизированные по кодомам гены gag-pol могут содержать последовательность нуклеиновой кислоты или состоять из последовательности нуклеиновой кислоты SEQ ID NO: 1. Оптимизированные по кодомам гены gag-pol могут быть включены в плазмиду, которая содержит последовательность нуклеиновой кислоты или состоит из последовательности нуклеиновой кислоты, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 5. Оптимизированные по кодомам гены gag-pol могут содержаться в плазмиде, которая содержит последовательность нуклеиновой кислоты или состоит из последовательности нуклеиновой кислоты SEQ ID NO: 5.

Респираторный парамиксовирус может представлять собой вирус Сендай.

Титр ретровирусного вектора, полученного способом согласно изобретению, может быть: (a) эквивалентен титру ретровирусного вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодомам гены gag-pol; или (b) повышен по сравнению с титром ретровирусного вектора, полученного соответствующим способом, в котором не используются оптимизированные

по кодонам гены gal-pol. Необязательно, титр ретровирусного вектора может быть по меньшей мере в 1,5 раза, по меньшей мере в 2 раза или по меньшей мере в 2,5 раза выше, чем титр ретровирусного вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодонам гены gal-pol.

Промотор может быть выбран из группы, состоящей из промотора цитомегаловируса (CMV), промотора фактора элонгации 1a (EF1a) и гибридного промотора человеческого CMV-энхансера/EF1a (hCEF). Предпочтительно вектор содержит гибридный промотор человеческого CMV-энхансера/EF1a (hCEF).

Трансген может быть выбран из: (a) секретируемого терапевтического белка, необязательно, альфа-1-антитрипсина (A1AT), фактора VIII, поверхностно-активного белка В (SFTPВ), фактора VII, фактора IX, фактора X, фактора XI, фактора фон Виллебранда, гранулоцитарного-макрофагального колониестимулирующего фактора (GM-CSF) и моноклонального антитела против инфекционного агента; или (b) CFTR, ABCA3, DNAH5, DNAH11, DNAI1 и DNAI2. Предпочтительно, трансген кодирует: (i) CFTR; (ii) A1AT; или (iii) FVIII.

В особенно предпочтительных вариантах осуществления изобретения, этот способ позволяет продуцировать ретровирусный/лентивирусный (например, SIV) вектор, где: (a) промотор представляет собой промотор hCEF, а трансген кодирует CFTR; (b) промотор представляет собой промотор hCEF, а трансген кодирует A1AT; или (c) промотор представляет собой промотор hCEF или CMV, а трансген кодирует FVIII.

Способ согласно изобретению может включать следующие стадии или состоять из следующих стадий: (a) выращивания клеток в суспензии; (b) трансфекции клеток одной или более плазмидами; (c) добавления нуклеазы; (d) сбора лентивируса; (e) добавления трипсина; и (d) очистки. Одна или более плазмид могут содержать или состоять из: (a) векторной геномной плазмиды, предпочтительно выбранной из pGM830 и pGM326 или их вариантов, как определено в настоящей заявке; (b) плазмиды, содержащей gal-pol, предпочтительно pGM691 или ее варианта, как определено в настоящей заявке; (c) плазмиды Rev, предпочтительно pGM299 или ее варианта, определенного в настоящей заявке; (d) плазмиды слитого (F) белка, предпочтительно pGM301 или ее варианта, определенного в настоящей заявке; и (e) плазмиды гемагглютинирина-нейраминидазы (HN), предпочтительно pGM303 или ее варианта, определенного в настоящей заявке. Соотношение «плазида векторного генома:плазида, содержащая gal-pol: плазида Rev: плазида F: плазида HN» может составлять 20:9:6:6:6.

Стадии (a)-(f) этого способа могут быть проведены последовательно. Клетки могут представлять собой клетки НЕК293 (такие как клетки НЕК293F или НЕК293Т) или клетки 293Т/17. Добавление нуклеазы может осуществляться на стадии перед сбором клеток. Добавление трипсина может быть осуществлено на стадии после сбора. Стадия очистки может включать одну или более стадий хроматографии.

Плазида векторного генома может быть модифицирована для уменьшения числа ОРС ретровируса.

Настоящее изобретение также относится к нуклеиновой кислоте, содержащей оптимизированные по кодонам гены gag-pol, где указанная нуклеиновая кислота имеет последовательность, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 1. Предпочтительно, нуклеиновая кислота содержит последовательность нуклеиновой кислоты SEQ ID NO: 1 или состоит из нее.

Настоящее изобретение также относится к плазмиде, содержащей нуклеиновую кислоту согласно изобретению, где, необязательно: (a) плазида содержит последовательность нуклеиновой кислоты или состоит из последовательности нуклеиновой кислоты, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 5; или (b) плазида содержит последовательность нуклеиновой кислоты SEQ ID NO: 5 или состоит из нее. Необязательно, внутри плазмиды, нуклеиновая кислота функционально связана с промотором, регулирующим экспрессию белков Gag и Pol, предпочтительно, с промотором CAG.

Настоящее изобретение также относится к клетке-хозяину, содержащей нуклеиновую кислоту согласно изобретению и/или плазмиду согласно изобретению.

Настоящее изобретение также относится к ретровирусному вектору, псевдотипированному гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, где указанный вектор может быть получен способом согласно изобретению.

Настоящее изобретение также относится к способу лечения заболевания, включающему введение индивидууму, нуждающемуся в этом, ретровирусного вектора, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, где указанный вектор может быть получен способом согласно изобретению. Заболевание, подлежащее лечению, может представлять собой заболевание легких, а предпочтительно кистозный фиброз.

Краткое описание чертежей

На фигуре 1 показано выравнивание генов gag-pol дикого типа (не оптимизированных по кодонам) из pGM297 с типичными оптимизированными по кодонам генами gag-pol согласно изобретению из pGM691, что демонстрирует изменения в последовательности дикого типа.

На фигуре 2 А-Г показаны схематические чертежи иллюстративных плазмид, используемых для получения векторов согласно изобретению. На **Г** показана не оптимизированная по кодонам плазида gag-pol (pDNA2a, а в частности, pGM297), которая может быть оптимизирована по кодонам в соответствии с настоящим изобретением.

На фигуре 3 показаны схематические чертежи репрезентативной плазмиды pDNA1, используемой для получения векторов A1AT согласно изобретению.

На фигуре 4 А-Д показаны схематические чертежи репрезентативной плазмиды pDNA1, используемой для получения векторов FVIII согласно изобретению.

На фигуре 5: На **А** проиллюстрировано гомология между плазмидой pDNA1

pGM326 и плазмидой pGM297 pDNA2a, не оптимизированной по кодонам. На **В** показано сравнение неоптимизированной по кодонам плазмиды pDNA2a pGM297 и оптимизированной по кодонам плазмиды pDNA2a pGM691 согласно изобретению с различиями между двумя аннотированными последовательностями. На **С** показан график гомологии матрицы ДНК, иллюстрирующий гомологию между последовательностью ДНК, присутствующей в pGM297 (горизонтальная ось) и pGM691 (вертикальная ось). Сплошная диагональная линия представляет собой гомологию последовательностей, пунктирная линия выделяет области пониженной идентичности последовательностей; при этом, следует обратить внимание на пониженную идентичность последовательностей в областях оптимизации кодонов генов gag и pol в pGM691. Также следует обратить внимание на дополнительную последовательность, присутствующую в pGM297 (расположенную приблизительно в области от 6000 до 7000 оснований в соответствии с нумерацией, показанной на горизонтальной оси), где эта область представляет собой область RRE, присутствующую в pGM297, но отсутствующую в pGM691. На **Д** показано выравнивание последовательности ДНК ClustalW областей gag pol pGM297 (нижний ряд последовательности ДНК) и pGM691 (верхний ряд последовательности ДНК); при этом, гомология последовательностей указана заштрихованными областями в рамке; консенсусная последовательность ДНК показана под списками последовательностей pGM691 и pGM297. При этом, следует обратить внимание на полную гомологию ДНК между последовательностями pGM297 и pGM691 в (i) области сдвига gag-pol, в перекрывающейся части генов gag-pol и в (ii) последовательности полиаденилирования кроличьего бета-глобина (RBG pA). Следует также отметить, что pGM297 содержит последовательность RRE SIV, но при этом, она отсутствует в pGM691. На **Е** показана рестрикционная карта оптимизированных по кодонам генов gag-pol в плазмиде pGM693.

На фигуре 6: На **А** показано, что в условиях проведения эксперимента (DOE), использование оптимизированной по кодонам плазмиды pDNA2a pGM691 приводило к наблюдаемому увеличению титра вектора rSIV.F/HN hCEF-CFTR. На **В** показано, что увеличение титра вектора rSIV.F/HN hCEF-CFTR, достигнутое с использованием оптимизированной по кодонам плазмиды pDNA2a pGM691, наблюдается в двух различных сериях экспериментальных условий.

На фигуре 7 показано, что титр вектора rSIV.F/HN CMV-EGFP, полученный с использованием оптимизированной по кодонам плазмиды pDNA2a pGM691, выше, чем титр, полученный с использованием не оптимизированного по кодонам gagpol в pDNA2a-плазмиде pGM297. Это указывает на то, что благоприятные свойства оптимизированного по кодонам gagpol в F/HN-псевдотипированных векторах не ограничиваются rSIV.F/HN hCEF-CFTR, а являются общим свойством используемого оптимизированного по кодонам gagpol в F/HN-псевдотипированных векторах F/HN.

На фигуре 8 показана карта линейной плазмиды для неполной области Gag RRE cPPT hCEF векторной геномной плазмиды pGM326.

На фигуре 9 показана аннотированная схема векторной геномной плазмиды

pGM326 с идентифицированными OPC SIV. В частности, две крупные OPC, одна из 189 аминокислот (а.к.), а одна из 250 аминокислот, были идентифицированы выше промотора hCEF и трансгена soCFTR2.

На фигуре 10 показано, что плазида векторного генома pGM326 и модифицированная плазида векторного генома pGM830 в идентичных во всем остальном условиях (включая отсутствие GagPol) дают сопоставимые титры вектора как в клетках HEK293T (левая панель), так и в клетках A549 (правая панель).

На фигуре 11 показан титр вектора, полученного с использованием соGagPol и либо pGM326, либо pGM830, в других идентичных условиях, с наблюдаемой тенденцией к увеличению титра вектора, в том случае, если GagPol был объединен с pGM830.

Подробное описание изобретения

Определения

Если это не оговорено особо, то все технические и научные термины, используемые в настоящей заявке, имеют свое общепринятое значение, известное специалистам в области, к которой относится настоящее изобретение. В публикации Singleton, et al., DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 20 ED., John Wiley and Sons, New York (1994), and Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, NY (1991), специалистами в данной области предоставлен общий словарь многих терминов, используемых в настоящем изобретении. Значение и объем терминов должны быть понятны специалистам, однако в случае какой-либо скрытой двусмысленности определений, представленных в настоящей заявке, предпочтение следует отдать определениям, представленным в настоящей заявке, а не в каком-либо словаре или в других литературных источниках. Следует отметить, что настоящее изобретение не ограничивается конкретной методикой, протоколами, реагентами и т.п., описанными в настоящей заявке, и следовательно, они могут варьироваться.

Раскрытие настоящего изобретения не ограничивается проиллюстрированными способами и материалами, раскрытыми в настоящей заявке, и для осуществления или тестирования вариантов согласно изобретению могут быть применены любые способы и материалы, аналогичные или эквивалентные описанным в настоящей заявке. Используемая здесь терминология предназначена только для описания конкретных вариантов осуществления изобретения и не должна рассматриваться как ограничение объема настоящего изобретения, который определяется исключительно формулой изобретения.

Описание вариантов осуществления изобретения не рассматривается как исчерпывающее или ограничивающее раскрытие более конкретных форм. Хотя конкретные варианты осуществления изобретения и примеры такого раскрытия описаны здесь в иллюстративных целях, однако, в рамках объема такого раскрытия возможны и другие различные эквивалентные модификации, как будет понятно специалистам в данной области. Так, например, хотя стадии или функции способов представлены в

определенном порядке, однако, альтернативные варианты осуществления изобретения могут включать функции в другом порядке, или эти функции могут выполняться по существу одновременно. Принципы раскрытия, представленные в настоящей заявке, могут быть применены и к другим процедурам или способам, если это необходимо. Различные описанные здесь варианты осуществления изобретения могут быть объединены для создания дополнительных вариантов осуществления изобретения. Аспекты раскрытия изобретения могут быть изменены, если это необходимо для использования композиций, функций и концепций в приведенных выше документах и приложениях для создания дополнительных вариантов осуществления изобретения. Более того, из соображений биологически функциональной эквивалентности, некоторые изменения в структуре белка могут быть сделаны так, чтобы это не влияло на биологическое или химическое действие с точки зрения их вида или количества. Эти и другие изменения могут быть внесены в настоящее изобретение в свете его подробного описания. При этом, предполагается, что все такие модификации входят в объем прилагаемой формулы изобретения.

Если это не оговорено особо, то любые последовательности нуклеиновых кислот записываются слева направо в ориентации от 5' к 3', а аминокислотные последовательности записываются слева направо в ориентации от амино- до карбокси-конца, соответственно.

Представленные здесь заголовки не ограничиваются различными аспектами или вариантами осуществления изобретения.

Используемый здесь термин «способный к», если он используется вместе с глаголом, охватывает или означает действие, выраженное соответствующим глаголом. Так, например, словосочетание «способный к взаимодействию» также означает взаимодействие, «способный к расщеплению» также означает расщепление, «способный к «связыванию» также означает связывание, а «способность к специфическому нацеливанию» также означает специфическое нацеливание.

В настоящем описании могут встречаться и другие определения терминов. Перед более подробным описанием иллюстративных вариантов осуществления изобретения следует отметить, что такое раскрытие изобретения не ограничивается конкретно описанными здесь вариантами, и эти варианты могут варьироваться. Также следует отметить, что используемая здесь терминология применяется лишь в целях описания конкретных вариантов осуществления изобретения и не должна рассматриваться как ограничение объема изобретения, поскольку объем настоящего изобретения будет определяться только прилагаемой формулой изобретения.

Числовые диапазоны включают числа, определяющие этот диапазон. Если указан диапазон значений, то подразумевается, что каждое промежуточное значение приводится с точностью до десятой доли единицы нижнего предела между верхним и нижним пределами этого диапазона, если из контекста не следует иное. Каждый меньший диапазон между любым заявленным значением или промежуточным значением в

указанном диапазоне и любым другим заявленным или промежуточным значением в указанном диапазоне охватывается данным вариантом изобретения. Верхние и нижние пределы этих меньших диапазонов могут независимо включаться в этот диапазон или исключаться из него, и каждый диапазон, в котором любой из этих пределов, ни один из них или оба предела не включены в меньшие диапазоны, также охватывается настоящим вариантом с учетом любого конкретно исключенного предела в заявленном диапазоне. В тех случаях, когда указанный диапазон включает один или оба предела, то диапазоны, исключающие один или оба из этих включенных пределов, также включены в настоящее изобретение.

Используемые здесь формы единственного числа могут относиться к одному или более чем одному (например, по меньшей мере к одному) грамматическому объекту. Кроме того, если это не оговорено особо, то термины, употребляемые в единственном числе, могут относиться к терминам, употребляемым во множественном числе, а термины, употребляемым во множественном числе, могут относиться к терминам, употребляемым в единственном числе. В настоящей заявке, использование союза «или» означает «и/или», если это не оговорено особо. Кроме того, использование термина «включая», а также других форм, таких как «включает» и «включенный», не является ограничивающим.

«Приблизительно» обычно может означать приемлемую степень погрешности измеряемой величины с учетом характера или точности измерений. Репрезентативные степени погрешности находятся в пределах 20 процентов (%), а обычно, в пределах 10% и чаще всего в пределах 5% от заданного значения или диапазона значений. Предпочтительно, термин «приблизительно» следует понимать здесь как плюс или минус (\pm) 5%, предпочтительно $\pm 4\%$, $\pm 3\%$, $\pm 2\%$, $\pm 1\%$, $\pm 0,5\%$, $\pm 0,1\%$ от значения числа, с которым он используется.

Термин «состоящий из» относится к композициям, способам и их соответствующим компонентам, описанным в настоящей заявке, и этот термин исключает любой элемент, не упомянутый в описании изобретения.

Используемый здесь термин «состоящий в основном из» относится к тем элементам, которые необходимы для осуществления настоящего изобретения. Этот термин допускает наличие элементов, которые существенно не влияют на основную и новую или функциональную характеристику (характеристики) настоящего изобретения (то есть, неактивные или неиммуногенные ингредиенты).

Варианты осуществления изобретения, описанные здесь как «включающие» один или более признаков, также могут рассматриваться как раскрытие соответствующих вариантов осуществления изобретения, «состоящих из» и/или «по существу состоящих из» таких признаков.

Концентрации, количества, объемы, проценты и другие числовые значения могут быть представлены здесь в формате диапазона. Также следует отметить, что такой формат диапазона используется здесь просто для удобства и краткости изложения и должен интерпретироваться в широком смысле, то есть, он может включать не только числовые

значения, явно указанные в качестве пределов диапазона, но также и все отдельные числовые значения или поддиапазоны, включенные в этот диапазон так, как если бы каждое числовое значение и поддиапазон были конкретно включены.

Используемые здесь термины «вектор», «ретровирусный вектор» и «ретровирусный вектор F/HN» являются синонимами для обозначения ретровирусного вектора, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, если это не оговорено особо. Термины «лентивирусный вектор» и «лентивирусный вектор F/HN» используются здесь как синонимы для обозначения лентивирусного вектора, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, если это не оговорено особо. Во всем описании настоящего изобретения, термин «ретровирусные векторы» в равной степени и без оговорок относится к лентивирусным векторам согласно изобретению и к векторам SIV, которые были псевдотипированы гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса (эти векторы также обозначаются здесь как SIV F/HN или SIV-FHN).

Используемые здесь термины «титр» и «выход» используются как синонимы для обозначения количества лентивирусного (например, SIV) вектора, полученного способом согласно изобретению. Титр является основным признаком, характеризующим эффективность получения вектора, при этом, более высокие титры обычно указывают на то, что было продуцировано больше ретровирусных/лентивирусных (например, SIV) векторов (например, с использованием того же количества реагентов). Титр или выход может относиться к числу векторных геномов, которые интегрировались в геном клетки-мишени (интеграционный титр), что является показателем «активных» вирусных частиц, то есть, количества частиц, способных трансдуцировать клетку. Единицы трансдукции (TU/мл, также называемые TTU/мл) представляют собой биологический показатель количества клеток-хозяев, которые трансдуцируются при определенных условиях разведения культуры ткани/вируса, и являются мерой количества «активных» вирусных частиц. Общее количество (активных+неактивных) вирусных частиц также можно определить любым подходящим способом, например, путем оценки количества Gag в тестируемом растворе или количества копий вирусной РНК в тестируемом растворе. Затем можно сделать предположение, что лентивирусная частица содержит либо 2000 молекул Gag, либо 2 молекулы вирусной РНК. После оценки общего числа частиц и трансдуцирующего титра/TU рассчитывают соотношение частиц/инфекционности. Аминокислоты упоминаются здесь с указанием либо названия аминокислоты, либо трехбуквенной аббревиатуры, либо однобуквенной аббревиатуры.

Используемые здесь термины «белок» и «полипептид» являются синонимами и означают ряд аминокислотных остатков, соединенных друг с другом пептидными связями между альфа-амино- и карбоксильными группами соседних остатков. Термины «белок» и «полипептид» относятся к полимеру из аминокислот, включая модифицированные

аминокислоты (например, фосфорилированные, сахар-содержащие, гликозилированные и т.п.) и аналоги аминокислот, независимо от их размера или функции. Термины «белок» и «полипептид» часто используются здесь в отношении относительно крупных полипептидов, тогда как термин «пептид» часто используется в отношении небольших полипептидов, но при использовании в данной области, значения этих терминов частично совпадают. Термины «белок» и «полипептид» используются здесь как синонимы, если речь идет о продукте гена и его фрагментах. Таким образом, типичные полипептиды или белки включают генные продукты, встречающиеся в природе белки, гомологи, ортологи, паралоги, фрагменты и другие эквиваленты, варианты, фрагменты и аналоги вышеперечисленного.

Используемые здесь термины «полинуклеотиды», «нуклеиновая кислота» и «последовательность нуклеиновой кислоты» относятся к любой молекуле, а предпочтительно к полимерной молекуле, включающей звенья рибонуклеиновой кислоты, дезоксирибонуклеиновой кислоты или их аналога. Нуклеиновая кислота может быть одноцепочечной или двухцепочечной. Одноцепочечная нуклеиновая кислота может представлять собой одну цепь нуклеиновой кислоты денатурированной двухцепочечной ДНК. Альтернативно, она может быть одноцепочечной нуклеиновой кислотой, не происходящей от какой-либо двухцепочечной ДНК. В одном аспекте изобретения, нуклеиновая кислота может представлять собой ДНК. В другом аспекте изобретения, нуклеиновая кислота может представлять собой РНК. Подходящие молекулы нуклеиновой кислоты представляют собой ДНК, включая геномную ДНК или кДНК. Другими подходящими молекулами нуклеиновых кислот являются РНК, включая кРНК, кшРНК и антисмысловые олигонуклеотиды. Термины «трансен» и «ген» также используются как синонимы, и оба эти термина охватывают фрагменты или их варианты, кодирующие белок-мишень.

Трансены согласно изобретению включают последовательности нуклеиновых кислот, которые были удалены из их природного окружения, рекомбинантные или клонированные изоляты ДНК и химически синтезированные аналоги или аналоги, биологически синтезированные гетерологичными системами.

Незначительные изменения в аминокислотных последовательностях согласно изобретению рассматриваются как охватываемые настоящим изобретением при условии, что такие варианты в аминокислотной(ых) последовательности(ях) имеют последовательности, которые по меньшей мере на 60%, по меньшей мере на 70%, более предпочтительно по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95% и наиболее предпочтительно по меньшей мере на 97% или по меньшей мере на 99% идентичны аминокислотной последовательности согласно изобретению или ее фрагменту, как определено в настоящей заявке. Термин гомология используется здесь для обозначения идентичности. Таким образом, последовательность варианта или аналога аминокислотной последовательности может отличаться на основе замены (обычно консервативной замены), делеции или инсерции. Белки, содержащие

такие модификации, называются здесь вариантами.

Белки согласно изобретению могут включать варианты, в которых аминокислотные остатки одного вида заменены соответствующим остатком другого вида либо в консервативном, либо в неконсервативном положении. Варианты белковых молекул, раскрытые в настоящей заявке, могут быть получены и использованы в настоящем изобретении. В соответствии с примерами аналитической химии в случае применения методов многомерного анализа данных к отношениям структура/свойство-активность [см., например, Wold, et al. *Multivariate data analysis in chemistry. Chemometrics-Mathematics and Statistics in Chemistry* (Ed.: B. Kowalski); D. Reidel Publishing Company, Dordrecht, Holland, 1984 (ISBN 90-277-1846-6), количественные взаимосвязи между активностью и свойствами белков могут быть оценены с применением хорошо известных математических методов, таких как статистическая регрессия, распознавание паттернов и классификация [см. например, Norman et al. *Applied Regression Analysis*. Wiley-Interscience; 3rd edition (April 1998) ISBN: 0471170828; Kandel, Abraham et al. *Computer-Assisted Reasoning in Cluster Analysis*. Prentice Hall PTR, (May 11, 1995), ISBN: 0133418847; Krzanowski, Wojtek. *Principles of Multivariate Analysis: A User's Perspective* (Oxford Statistical Science Series, No 22 (Paper)). Oxford University Press; (December 2000), ISBN: 0198507089; Witten, Ian H. et al *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann; (October 11, 1999), ISBN:1558605525; Denison David G. T. (Editor) et al *Bayesian Methods for Nonlinear Classification and Regression* (Wiley Series in Probability and Statistics). John Wiley & Sons; (July 2002), ISBN: 0471490369; Ghose, Arup K. et al. *Combinatorial Library Design and Evaluation Principles, Software, Tools, and Applications in Drug Discovery*. ISBN: 0-8247-0487-8]. Свойства белков могут быть оценены исходя из эмпирических и теоретических моделей (например, анализа вероятных контактных остатков или определения физико-химических свойств) белковых последовательностей, функциональных и трехмерных структур, и эти свойства можно рассматривать по отдельности и в комбинации.

Аминокислоты упоминаются здесь по их названию, в виде трехбуквенной аббревиатуры или однобуквенной аббревиатуры. Используемый здесь термин «белок» включает белки, полипептиды и пептиды. Используемый здесь термин «аминокислотная последовательность» является синонимом термина «полипептид» и/или термина «белок». В некоторых случаях, термин «аминокислотная последовательность» является синонимом термина «пептид». Термины «белок» и «полипептид» используются здесь как синонимы. В настоящем описании и в формуле изобретения могут быть использованы общепринятые однобуквенные и трехбуквенные коды аминокислотных остатков. 3-буквенный код аминокислот определен специалистами Объединенной комиссии IUPAC/IUB по биохимической номенклатуре (JCBN). Также следует отметить, что полипептид может кодироваться более, чем одной нуклеотидной последовательностью из-за вырожденности генетического кода.

Аминокислотные остатки в неконсервативных положениях могут быть заменены

консервативными или неконсервативными остатками. В частности, рассматриваются консервативные замены аминокислот.

«Консервативная аминокислотная замена» представляет собой замену аминокислотного остатка другим аминокислотным остатком, имеющим аналогичную боковую цепь. Специалистами в данной области определены семейства аминокислотных остатков, имеющих сходные боковые цепи, включая основные боковые цепи (например, лизин, аргинин или гистидин), кислотные боковые цепи (например, аспарагиновая кислота или глутаминовая кислота), незаряженные полярные боковые цепи (например, глицин, аспарагин, глутамин, серин, треонин, тирозин или цистеин), неполярные боковые цепи (например, аланин, валин, лейцин, изолейцин, пролин, фенилаланин, метионин или триптофан), бета-разветвленные боковые цепи (например, треонин, валин, изолейцин) и ароматические боковые цепи (например, тирозин, фенилаланин, триптофан или гистидин). Таким образом, если аминокислота в полипептиде заменена другой аминокислотой из того же семейства боковых цепей, то такая аминокислотная замена считается консервативной. Включение консервативно модифицированных вариантов в белок согласно изобретению не исключает другие формы вариантов, например, полиморфные варианты, межвидовые гомологи и аллели.

«Неконсервативные аминокислотные замены» включают замены, в которых (i) остаток, имеющий электроположительную боковую цепь (например, Arg, His или Lys), замещен или заменен электроотрицательным остатком (например, Glu или Asp), (ii) гидрофильный остаток (например, Ser или Thr) замещен или заменен гидрофобным остатком (например, Ala, Leu, Ile, Phe или Val), (iii) цистеин или пролин замещен или заменен любым другим остатком, или (iv) остаток, имеющий объемную гидрофобную или ароматическую боковую цепь (например, Val, His, Ile или Trp), замещен или заменен остатком с меньшей боковой цепью (например, Ala или Ser) или без боковой цепи (например, Gly).

«Инсерции» или «делеции» обычно составляют приблизительно 1, 2 или 3 аминокислоты. Допустимые изменения могут быть определены экспериментально путем систематического введения инсерций или делеций аминокислот в белок с применением методов рекомбинантной ДНК и анализа на активность полученных рекомбинантных вариантов. Это может быть проведено с применением не более, чем рутинных экспериментов квалифицированным специалистом.

«Фрагмент» полипептида включает по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 70%, по меньшей мере 80%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 97% или более исходного полипептида.

Полинуклеотиды согласно изобретению могут быть получены любыми способами, известными специалистам в данной области. Так, например, большие количества полинуклеотидов могут быть получены путем репликации в подходящей клетке-хозяине. Фрагменты природной или синтетической ДНК, кодирующие желаемый фрагмент, будут включены в конструкции рекомбинантных нуклеиновых кислот, а обычно, в конструкции

ДНК, способные к встраиванию в прокариотическую или эукариотическую клетку и к репликации в этих клетках. Обычно, конструкции ДНК являются подходящими для автономной репликации в одноклеточном хозяине, таком как дрожжи или бактерии, но они также могут быть предназначены для введения и интеграции в геном культивируемых линий клеток насекомых, млекопитающих, растений или других эукариотических клеток.

Полинуклеотиды согласно изобретению также могут быть получены путем химического синтеза, например, фосфорамидитным методом или триэфирным методом, и могут быть получены на коммерчески доступных автоматических синтезаторах олигонуклеотидов. Двухцепочечный фрагмент может быть получен из одноцепочечного продукта химического синтеза, либо путем синтеза комплементарной цепи и одновременно отжига цепи в соответствующих условиях, либо путем добавления комплементарной цепи с использованием ДНК-полимеразы с соответствующей последовательностью праймера.

Применительно к последовательности нуклеиновой кислоты, термин «выделенный» в контексте настоящего изобретения означает, что полинуклеотидная последовательность была удалена из ее природного генетического окружения и, таким образом, не содержит других посторонних или нежелательных кодирующих последовательностей (но может включать встречающиеся в природе 5'- и 3'-нетранслируемые области, такие как промоторы и терминаторы), и присутствует в форме, подходящей для использования в генетически сконструированных системах получения белков. Такие выделенные молекулы представляют собой молекулы, которые были отделены от своего природного окружения.

Из-за вырожденности генетического кода, среди полинуклеотидов согласно изобретению возможны значительные модификации последовательностей. Вырожденные кодоны, охватывающие все возможные кодоны для данной аминокислоты, представлены ниже:

Аминокислота	Кодоны	Вырожденный кодон
Cys	TGC TGT	TGY
Ser	AGC AGT TCA TCC TCG TCT	WSN
Thr	ACA ACC ACG ACT	ACN
Pro	CCA CCC CCG CCT	CCN
Ala	GCA GCC GCG GCT	GCN
Gly	GGA GGC GGG GGT	GGN
Asn	AAC AAT	AAV
Asp	GAC GAT	GAY
Glu	GAA GAG	GAR
Gln	CAA CAG	CAR
His	CAC CAT	CAY
Arg	AGA AGG CGA CGC CGG CGT	MGN
Lys	AAA AAG	AAR
Met	ATG	ATG
Ile	ATA ATC ATT	ATH
Leu	CTA CTC CTG CTT TTA TTG	YTN
Val	GTA GTC GTG GTT	GTN

Phe	TTC TTT	TTY
Tyr	TAC TAT	TAY
Trp	TGG	TGG
Ter	TAA TAG TGA	TRR
Asn/Asp		RAY
Glu/Gln		SAR
Любая		NNN

Для специалиста в данной области очевидно, что существует некоторая гибкость при определении вырожденных кодонов, представляющих все возможные кодоны, кодирующие каждую аминокислоту. Так, например, некоторые полинуклеотиды, охватываемые вырожденной последовательностью, могут кодировать варианты аминокислотных последовательностей, но специалист в данной области может легко идентифицировать такие варианты последовательностей со ссылкой на аминокислотные последовательности согласно изобретению.

«Модифицированная» последовательность нуклеиновой кислоты имеет существенную гомологию или существенное сходство с эталонной последовательностью нуклеиновой кислоты (или ее фрагментом). Последовательность нуклеиновой кислоты или ее фрагмент является «по существу гомологичными» (или «по существу идентичными») эталонной последовательности, если при оптимальном выравнивании (с соответствующими инсерциями или делециями нуклеотидов) с другой нуклеиновой кислотой (или комплементарной ей цепью), идентичность нуклеотидных последовательностей составляет по меньшей мере приблизительно 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более нуклеотидных оснований. Способы определения гомологии последовательностей нуклеиновых кислот известны специалистам в данной области.

Альтернативно, «модифицированная» последовательность нуклеиновой кислоты по существу гомологична (или по существу идентична) эталонной последовательности (или ее фрагменту), если «вариант» и эталонная последовательность способны гибридизоваться в жестких (например, очень жестких) условиях. На гибридизацию последовательностей нуклеиновых кислот, помимо состава оснований, длины комплементарных цепей и количества ошибочных спариваний нуклеотидных оснований между гибридизующимися нуклеиновыми кислотами, будут влиять такие условия, как концентрация соли (например, NaCl), температура или органические растворители, как может быть легко оценено специалистами в данной области. Предпочтительно, используются жесткие температурные условия, которые обычно включают температуру выше 30°C, обычно выше 37°C, а предпочтительно выше 45°C. Жесткие солевые условия обычно составляют менее 1000 мМ, обычно менее 500 мМ, а предпочтительно менее 200 мМ. рН обычно составляет от 7,0 до 8,3. При этом, комбинация параметров гораздо важнее любого отдельного параметра.

Способы определения процента идентичности последовательностей нуклеиновой

кислоты известны специалистам в данной области. Так, например, при оценке идентичности последовательностей нуклеиновой кислоты, последовательность, имеющая определенное количество смежных нуклеотидов, может быть выровнена с последовательностью нуклеиновой кислоты (имеющей такое же количество смежных нуклеотидов) соответствующей части последовательности нуклеиновой кислоты согласно изобретению. Программы, известные специалистам в данной области для определения процента идентичности последовательностей нуклеиновой кислоты, включают программу BLAST для нуклеотидов (как описано ниже).

Специалисту в данной области очевидно, что различные виды демонстрируют «предпочтительный набор кодонов». Используемый в настоящей заявке термин «предпочтительный набор кодонов» относится к кодонам, которые наиболее часто используются в клетках определенного вида, при этом, предпочтение отдается одному или нескольким представителям возможных кодонов, кодирующих каждую аминокислоту. Так, например, аминокислота треонин (Thr) может кодироваться ACA, ACC, ACG или ACT, но в клетках-хозяевах млекопитающих, ACC является наиболее часто используемым кодоном; а у других видов, предпочтительными могут быть другие кодоны. Кодоны, предпочтительные для конкретных видов клеток-хозяев, могут быть введены в полинуклеотиды согласно изобретению различными способами, известными специалистам в данной области. Введение предпочтительных последовательностей кодонов в рекомбинантную ДНК может, например, усиливать продуцирование белка за счет повышения эффективности трансляции белка в клетках определенного типа или вида. Таким образом, в соответствии с настоящим изобретением, в дополнение к генам gag-pol, любая последовательность нуклеиновой кислоты может быть оптимизирована по кодонам для экспрессии в клетке-хозяине или в клетке-мишени. В частности, геном вектора (или соответствующая плаزمид), ген REV (или соответствующая плазмид), ген слитого белка (F) (или соответствующая плазмид) и/или ген гемагглютинаина-нейраминидазы (HN) (или соответствующая плазмид), или любая их комбинация могут быть оптимизированы по кодонам.

«Фрагмент» представляющего интерес полинуклеотида включает серию смежных нуклеотидов последовательности указанного полноразмерного полинуклеотида. Так, например, «фрагмент» представляющего интерес полинуклеотида может содержать (или состоять из них) по меньшей мере 30 смежных нуклеотидов последовательности указанного полинуклеотида (например, по меньшей мере 35, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 или 1000 смежных остатков нуклеиновой кислоты указанного полинуклеотида). Фрагмент может включать по меньшей мере одну антигенную детерминанту и/или может кодировать по меньшей мере один антигенный эпитоп соответствующего представляющего интерес полипептида. Обычно, фрагмент, определенный в настоящей заявке, сохраняет функцию полноразмерного полинуклеотида.

Используемые здесь термины «снижение», «сниженный», «уменьшение» или

«ингибирование» относятся к снижению на статистически значимую величину. Термины «снижать», «снижение», «уменьшать» или «ингибировать» обычно означают снижение по меньшей мере на 10% по сравнению с эталонным уровнем (например, в отсутствии лечения) и могут включать, например, снижение по меньшей мере приблизительно на 10%, по меньшей мере приблизительно на 20%, по меньшей мере приблизительно на 25%, по меньшей мере приблизительно на 30%, по меньшей мере приблизительно на 35%, по меньшей мере приблизительно на 40%, по меньшей мере приблизительно на 45%, по меньшей мере приблизительно на 50%, по меньшей мере приблизительно на 55%, по меньшей мере приблизительно на 60%, по меньшей мере приблизительно на 65%, по меньшей мере приблизительно на 70%, по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98%, по меньшей мере приблизительно на 99% или более. Используемый здесь термин «снижение» или «ингибирование» включает полное ингибирование или снижение по сравнению с эталонным уровнем. «Полное ингибирование» представляет собой 100% ингибирование (то есть, отмену) по сравнению с эталонным уровнем.

Все используемые здесь термины «увеличение», «увеличивать», «усиливать» или «активировать» относятся к увеличению на статически значимую величину. Термины «увеличенный», «повышать», «усиливать» или «активировать» могут означать увеличение по меньшей мере на 25%, по меньшей мере на 50% по сравнению с эталонным уровнем, например, увеличение по меньшей мере приблизительно на 50%, или по меньшей мере приблизительно на 75%, или по меньшей мере приблизительно на 80%, или по меньшей мере приблизительно на 90%, или по меньшей мере приблизительно на 100%, или по меньшей мере приблизительно на 150%, или по меньшей мере приблизительно на 200%, или по меньшей мере приблизительно на 250% или более по сравнению с эталонным уровнем, или по меньшей мере приблизительно в 1,5 раза, или по меньшей мере приблизительно в 2 раза, или по меньшей мере приблизительно в 2,5 раза, или по меньшей мере приблизительно в 3 раза, или по меньшей мере приблизительно в 4 раза или по меньшей мере приблизительно в 5 раз, или по меньшей мере приблизительно в 10 раз, или любое увеличение от 1,5 до 10 раз или более по сравнению с эталонным уровнем. С точки зрения выхода или титра, «увеличение» представляет собой наблюдаемое или статистически значимое увеличение такого уровня.

Термины «индивидуум», «субъект» и «пациент» используются здесь как синонимы и относятся к млекопитающему, для которого желательны диагностика, прогноз, мониторинг заболевания, лечение, терапия и/или оптимизация терапии. Млекопитающими могут быть (без ограничения) человек, примат, не являющийся человеком, мышь, крыса, собака, кошка, лошадь или корова. В предпочтительном варианте осуществления изобретения, индивидуумом, субъектом или пациентом является человек. «Индивидуум» может быть взрослым, несовершеннолетним или младенцем. «Индивидуум» может быть

мужчиной или женщиной.

«Индивидуумом, нуждающимся» в лечении определенного состояния, может быть человек, страдающий этим заболеванием, у которого было диагностировано это состояние или у которого имеется риск развития этого состояния.

Индивидуумом может быть индивидуум, который ранее был диагностирован или идентифицирован как страдающий состоянием или имеющий состояние, нуждающееся в лечении, или одно или несколько осложнений или симптомов, связанных с таким состоянием, и, необязательно, индивидуум, уже проходивший лечению состояния, определенного в настоящей заявке, или одного или нескольких осложнений или симптомов, связанных с указанным состоянием. Альтернативно, индивидуумом также может быть индивидуум, у которого ранее не было диагностировано состояние, определенное в настоящей заявке, или один или более симптомов или осложнений, связанных с указанным состоянием. Так, например, индивидуумом может быть индивидуум, у которого наблюдается один или более факторов риска развития состояния, или один или более симптомов или осложнений, связанных с указанным состоянием, или индивидуум, у которого отсутствуют такие факторы риска.

Используемый здесь термин «здоровый индивидуум» относится к индивидууму или к группе индивидуумов, например, индивидуумов, у которых не проявляются какие-либо симптомы заболевания, у которых не было диагностировано заболевание и/или у которых маловероятно развитие такого заболевания, например, кистозного фиброза (CF) или любого другого заболевания, описанного в настоящей заявке. Предпочтительно, указанный(ые) здоровый(е) индивидуум (индивидуумы) не принимает(ют) лекарства, воздействующие на CF, и у него (у них) не было диагностировано какое-либо другое заболевание. Один или более здоровых индивидуумов могут иметь тот же пол, возраст и/или индекс массы тела (ИМТ), как и обследуемый индивидуум. Применение стандартных статистических методов, используемых в медицине, позволяет определить как нормальные уровни экспрессии у здоровых людей, так и значительные отклонения от таких нормальных уровней.

В настоящей заявке, термины «контроль» и «контрольная группа» являются синонимами.

Используемый здесь термин «фармацевтически приемлемый» относится к лекарственному средству, одобренному Регуляторными органами Федерального правительства или Правительства штата или включенному в Фармакопею США, Европейскую фармакопею или другую общепризнанную фармакопею.

Обсуждаемые здесь публикации представлены исключительно для их раскрытия до даты подачи настоящей заявки. Ничто в настоящем изобретении не должно быть истолковано как допущение того, что такие публикации составляют предшествующий уровень техники для прилагаемой формулы изобретения.

Раскрытие, относящееся к различным способам согласно изобретению, предназначено для применения в равной степени к другим способам, терапевтическому

применению или к методам, к носителю данных или к устройству, к компьютерному программному продукту и т.п.

Ретровирусные и лентивирусные векторы

Настоящее изобретение относится к получению ретровирусной/лентивирусной (например, SIV) конструкции. Термин «ретровирус» относится к любому члену семейства РНК-вирусов Retroviridae, которые кодируют фермент обратную транскриптазу. Термин «лентивирус» относится к семейству ретровирусов. Примерами ретровирусов, подходящих для использования в настоящем изобретении, являются гаммаретровирусы, такие как вирус лейкоза мышей (MLV) и вирус лейкоза кошек (FLV). Примерами лентивирусов, подходящих для их использования в настоящем изобретении, являются вирус иммунодефицита обезьян (SIV), вирус иммунодефицита человека (HIV), вирус иммунодефицита кошек (FIV), вирус инфекционной анемии лошадей (EIAV) и вирус Висны/maedi. Предпочтительно, настоящее изобретение относится к лентивирусным векторам и к их получению. Особенно предпочтительным лентивирусным вектором является вектор SIV (включая все его штаммы и подтипы), такой как SIV-AGM (первоначально выделенный у африканских зеленых мартышек, *Cercopithecus aethiops*). Альтернативно, настоящее изобретение относится к векторам ВИЧ.

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению обычно представляют собой векторы, псевдотипированные гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса. Предпочтительно, респираторный парамиксовирус представляет собой вирус Сендай (вирус парагриппа мышей типа 1). Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению могут быть псевдотипированы белками другого вируса, при условии, что использование оптимизированных по кодонам генов gag-pol (например, SIV) не будет оказывать негативного влияния на титр полученного вектора, или даже будет приводить к увеличению титра вектора. Неограничивающими примерами других белков, которые могут быть использованы для псевдотипирования ретровирусных/лентивирусных (например, SIV) векторов согласно изобретению, являются гликопротеин G вируса везикулярного стоматита (G-VSV) и белок шипа коронавируса 2, вызывающего тяжелый острый респираторный синдром (SARS-CoV-2) или его модифицированные формы, такие как формы, описанные в заявках на патент Великобритании № 2118685.3 и 2105278.2, каждая из которых в полном объеме включена в настоящее описание посредством ссылки. Таким образом, настоящее изобретение может относиться к получению SIV, псевдотипированного G-VSV, или SIV, псевдотипированного белком шипа SARS-CoV-2, с использованием оптимизированных по кодонам генов gag-pol.

Ретровирусный/лентивирусный (например, SIV) вектор, полученный в соответствии с настоящим изобретением, может быть компетентным по интегразе (IC). Альтернативно, лентивирусный (например, SIV) вектор может быть дефицитным по интегразе (ID).

Ретровирусные/лентивирусные векторы, такие как векторы, полученные в соответствии с настоящим изобретением, могут интегрироваться в геном трансдуцированных клеток и обеспечивать длительную экспрессию, что делает их подходящими для трансдукции стволовых клеток/клеток-предшественников. В легких было идентифицировано несколько типов клеток с регенеративной способностью, ответственных за сохранение специфических клеточных линий в проводящих дыхательных путях и в альвеолах. К ним относятся базальные клетки и клетки протоков подслизистых желез в верхних дыхательных путях, булавовидные клетки и нейроэндокринные клетки в бронхиолярных дыхательных путях, бронхиоальвеолярные стволовые клетки в терминальных бронхиолах и пневмоциты типа II в альвеолах. Таким образом, не ограничиваясь какой-либо теорией, авторы лишь отмечают, что указанные ретровирусные/лентивирусные (например, SIV) векторы обеспечивают длительную экспрессию представляющего интерес трансгена после введения трансгена в одну или более долгоживущих эпителиальных клеток или клеток дыхательных путей определенных типов, таких как базальные клетки и клетки протоков подслизистых желез в верхних дыхательных путях, булавовидные клетки и нейроэндокринные клетки в бронхиолярных дыхательных путях, бронхиоальвеолярные стволовые клетки в терминальных бронхиолах и пневмоциты типа II в альвеолах.

Соответственно, ретровирусные/лентивирусные (например, SIV) векторы, полученные способами согласно изобретению, могут трансдуцировать одну или более клеток или клеточных линий с регенеративным потенциалом в легких (включая дыхательные пути и респираторный тракт) для достижения длительной экспрессии генов. Так, например, ретровирусные/лентивирусные (например, SIV) векторы могут трансдуцировать базальные клетки, такие как клетки верхних дыхательных путей/дыхательных путей. Базальные клетки играют центральную роль в процессах поддержания и восстановления эпителия после повреждения. Кроме того, базальные клетки широко распределены по всему респираторному эпителию человека с относительным распределением от 30% (в более крупных дыхательных путях) до 6% (в более мелких дыхательных путях).

Ретровирусные/лентивирусные (например, SIV) векторы, полученные в соответствии с настоящим изобретением, могут быть использованы для трансдукции выделенных и размноженных стволовых клеток/клеток-предшественников *ex vivo* перед введением пациенту. Предпочтительно, ретровирусные/лентивирусные (например, SIV) векторы, полученные в соответствии с настоящим изобретением, используют для трансдукции клеток в легкие (или в дыхательных путях/респираторном тракте) *in vivo*.

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению демонстрируют значительную устойчивость к сдвиговым усилиям с лишь умеренным снижением способности к трансдукции при прохождении через клинически значимые устройства для доставки, такие как бронхоскопы, аэрозольные баллончики и аэрозольные ингаляторы.

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению обеспечивают высокие уровни экспрессии трансгена, что приводит к повышению уровней (терапевтических уровней) экспрессии терапевтического белка. Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению обычно обеспечивают высокие уровни экспрессии трансгена при введении пациенту. Используемые здесь термины «высокая экспрессия» и «терапевтическая экспрессия» являются синонимами. Экспрессия может быть оценена любым подходящим методом (качественным или количественным, предпочтительно количественным), а концентрации приведены в любой подходящей единице измерения, например, в нг/мл или нМ.

Экспрессия представляющего интерес трансгена может быть приведена относительно экспрессии соответствующего эндогенного (дефектного) гена у пациента. Экспрессия может быть выражена как экспрессия мРНК или белка. Экспрессия трансгена согласно изобретению, такого как функциональный ген CFTR, может быть количественно определена относительно эндогенного гена, такого как эндогенные (дисфункциональные) гены CFTR, с точки зрения количества копий мРНК на клетку или любой другой подходящей единицы.

Уровни экспрессии трансгена и/или кодируемого терапевтического белка согласно изобретению могут быть оценены в легочной ткани, в жидкости эпителиальной выстилки и/или в сыворотке/плазме, в зависимости от ситуации. Таким образом, высокий и/или терапевтический уровень экспрессии может относиться к концентрации в легких, в жидкости эпителиальной выстилки и/или в сыворотке/плазме.

Трансген, включенный в вектор согласно изобретению, может быть модифицирован для облегчения экспрессии. Так, например, трансгенная последовательность может быть дефицитной по CpG (или она может не содержать CpG), и/или она может присутствовать в оптимизированной по кодонам форме для облегчения экспрессии гена. Таким образом, стандартные способы модификации последовательности трансгена известны специалистам в данной области.

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению демонстрируют эффективное поглощение клетками дыхательных путей, повышенную экспрессию трансгена и не теряют своей эффективности при повторном введении. Соответственно, ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению способны продуцировать длительную, воспроизводимую экспрессию высокого уровня в клетках дыхательных путей, не вызывая при этом чрезмерного иммунного ответа.

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению обеспечивают длительную экспрессию трансгена, что приводит к длительной экспрессии терапевтического белка. Как описано в настоящей заявке, словосочетания «долговременная экспрессия», «устойчивая экспрессия», «длительная экспрессия» и «стойкая экспрессия» являются синонимами. Долгосрочная экспрессия согласно изобретению означает экспрессию терапевтического гена и/или белка, предпочтительно

на терапевтических уровнях, в течение по меньшей мере 45 дней, по меньшей мере 60 дней, по меньшей мере 90 дней, по меньшей мере 120 дней, по меньшей мере 180 дней, по меньшей мере 250 дней, по меньшей мере 360 дней, по меньшей мере 450 дней, по меньшей мере 730 дней или более. Предпочтительно, длительная экспрессия означает экспрессию в течение по меньшей мере 90 дней, по меньшей мере 120 дней, по меньшей мере 180 дней, по меньшей мере 250 дней, по меньшей мере 360 дней, по меньшей мере 450 дней, по меньшей мере 720 дней или более, более предпочтительно по меньшей мере 360 дней, по меньшей мере 450 дней, по меньшей мере 720 дней или более. Эта долговременная экспрессия может быть достигнута путем введения повторных доз или однократной дозы.

Повторные дозы можно вводить два раза в день, ежедневно, два раза в неделю, еженедельно, ежемесячно, каждые два месяца, каждые три месяца, каждые четыре месяца, каждые шесть месяцев, ежегодно, каждые два года или реже. Введение доз можно продолжать до тех пор, пока это необходимо, например, в течение по меньшей мере шести месяцев, по меньшей мере одного года, двух лет, трех лет, четырех лет, пяти лет, десяти лет, пятнадцати лет, двадцати лет или более, или на протяжении всей жизни пациента, проходящего лечение.

Ретровирусный/лентивирусный (например, SIV) вектор содержит промотор, функционально связанный с трансгеном и обеспечивающий экспрессию трансгена. Обычно, промотор представляет собой гибридный промотор человеческого энхансера CMV/EF1a (hCEF). В этом промоторе hCEF может отсутствовать интрон, соответствующий нуклеотидам 570-709, и экзон, соответствующий нуклеотидам 728-733 промотора hCEF. Предпочтительный пример последовательности промотора hCEF согласно изобретению представлен SEQ ID NO: 10. Промотором может быть промотор CMV. Пример последовательности промотора CMV представлен SEQ ID NO: 11. Промотор может представлять собой промотор фактора элонгации 1a (EF1a) человека. Пример промотора EF1a представлен SEQ ID NO: 12. Другие промоторы для экспрессии трансгена известны специалистам в данной области, и возможность их применения для ретровирусных /лентивирусных (например, SIV) векторов согласно изобретению может быть определена с применением обычных методов, известных специалистам в данной области. Неограничивающие примеры других промоторов включают UbC и UC0E. Как описано в настоящей заявке, промотор может быть модифицирован для дополнительной регуляции экспрессии трансгена согласно изобретению.

Промотор, включенный в ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению, может быть специально выбран и/или модифицирован для дальнейшего уточнения регуляции экспрессии терапевтического гена. И в этом случае, специалистам в данной области известны подходящие промоторы и стандартные методы их модификации. В качестве неограничивающего примера может служить ряд подходящих (не содержащих CpG) промоторов, подходящих для использования в настоящем изобретении, и такие промоторы описаны в публикации Pringle et al. (J. Mol.

Med. Berl. 2012, 90(12): 1487-96), которая в полном объеме включена в настоящее описание посредством ссылки. Предпочтительно, ретровирусные/лентивирусные векторы (в частности, векторы SIV F/HN) согласно изобретению содержат промотор hCEF с низким содержанием динуклеотидов CpG или без них. В промоторе hCEF, все динуклеотиды CG могут быть заменены любым из AG, TG или GT. Таким образом, промотор hCEF может не содержать CpG. Предпочтительный пример последовательности промотора hCEF, не содержащей CpG, согласно изобретению представлен SEQ ID NO: 10. Отсутствие динуклеотидов CpG дополнительно улучшает свойства ретровирусных/лентивирусных (например, SIV) векторов согласно изобретению, а в частности, в тех случаях, когда нежелательно индуцировать иммунный ответ против экспрессированного антигена или воспалительный ответ против доставленной экспрессионной конструкции. Удаление динуклеотидов CpG снижает вероятность возникновения симптомов, подобных гриппу, и воспаления, которые могут возникать в результате введения конструкций, а особенно их введения в дыхательные пути.

Ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению может быть модифицирован для прекращения экспрессии гена. Стандартные методы такой модификации вектора известны специалистам в данной области. В качестве неограничивающего примера могут служить широко используемые Tet-чувствительные промоторы.

Предпочтительно, настоящее изобретение относится к ретровирусным/лентивирусным векторам F/HN, содержащим промотор и трансген, а в частности к векторам F/HN SIV. Псевдотипирование F/HN особенно эффективно для нацеливания на клетки эпителия дыхательных путей, и поэтому при терапевтическом применении, его обычно доставляют в клетки дыхательных путей, включая клетки эпителия дыхательных путей. Соответственно, ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению являются особенно подходящими для лечения заболеваний или расстройств дыхательных путей, респираторного тракта или легких. Обычно, ретровирусные/лентивирусные (например, SIV) векторы могут быть использованы для лечения генетического респираторного заболевания.

Ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению может содержать трансген, кодирующий полипептид или белок, который является терапевтическим и может быть применен для лечения таких заболеваний, а в частности, заболеваний или расстройств дыхательных путей, респираторного тракта или легких.

Соответственно, ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению может содержать трансген, кодирующий белок, выбранный из: (i) секретируемого терапевтического белка, необязательно альфа-1-антитрипсина (A1AT), фактора VIII, поверхностно-активного белка В (SFTPВ), фактора VII, фактора IX, фактора X, фактора XI, фактора фон Виллебранда, гранулоцитарного-макрофагального колониестимулирующего фактора (GM-CSF) и моноклонального антитела против инфекционного агента; или (ii) CFTR, ABCA3, DNAH5, DNAH11, DNAI1 и DNAI2. Другими

примерами трансгенов, которые могут быть включены в ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению, являются гены, связанные или ассоциированные с дефицитом других поверхностно-активных веществ.

Предпочтительно, трансген кодирует CFTR. Пример кДНК CFTR представлен SEQ ID NO: 13. Также включены ее варианты (как описано в настоящей заявке), а в частности, варианты, содержащие последовательность, которая по меньшей мере на 90% (например, по меньшей мере на 90, 92, 94, 95, 96, 97, 98, 99 или на 100%) идентична SEQ ID NO: 13.

Трансген может кодировать A1AT. Пример трансгена A1AT представлен SEQ ID NO: 14 или комплементарной последовательностью SEQ ID NO: 15. SEQ ID NO: 14 представляет собой оптимизированный по кодонам трансген A1AT, не содержащий CpG, и ранее разработанный авторами настоящего изобретения для усиления трансляции в клетках человека. Было показано, что такая оптимизация усиливает экспрессию генов до 15 раз. Варианты той же самой последовательности (как определено в настоящей заявке), которые обладают таким же техническим эффектом усиления трансляции по сравнению с немодифицированной последовательностью гена A1AT (дикого типа), также входят в объем настоящего изобретения. Примером полипептида, кодируемого указанным трансгеном A1AT, может служить полипептид SEQ ID NO: 16. Также включены ее варианты (как описано в настоящей заявке), а в частности, варианты, содержащие последовательность, которая по меньшей мере на 90% (например, по меньшей мере на 90, 92, 94, 95, 96, 97, 98, 99 или 100%) идентична SEQ ID NO: 14, 15 или 16.

Трансген может кодировать FVIII. Примеры трансгена FVIII представлены SEQ ID NO: 17 и 18 или соответствующими комплементарными последовательностями SEQ ID NO: 19 и 20. Примером полипептида, кодируемого трансгеном FVIII, может служить полипептид SEQ ID NO: 21 или 22. Также включены ее варианты (как описано в настоящей заявке), а в частности, варианты, содержащие последовательность, которая по меньшей мере на 90% (например, по меньшей мере на 90, 92, 94, 95, 96, 97, 98, 99 или 100%) идентична любой из SEQ ID NO: 17-22.

Трансген согласно изобретению может представлять собой любой один или более из DNAH5, DNAH11, DNAI1 и DNAI2 или другой известный родственный ген.

Если эпителий дыхательных путей предназначен для доставки ретровирусного/лентивирусного (например, SIV) вектора, то трансген может кодировать A1AT, SFTPВ или GM-CSF. Трансген может кодировать моноклональное антитело (mAb) против инфекционного агента. Трансген может кодировать антитело против TNF-альфа. Трансген может кодировать терапевтический белок, участвующий в развитии воспалительного, иммунного или метаболического состояния.

Ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению может быть доставлен в клетки дыхательных путей, для обеспечения продукции белков, секретируемых в систему кровообращения. В таких вариантах осуществления изобретения, трансген может кодировать фактор VII, фактор VIII, фактор IX, фактор X, фактор XI и/или фактор фон Виллебранда. Такой вектор может быть использован для

лечения заболеваний, а особенно, сердечно-сосудистых заболеваний и заболеваний крови, а предпочтительно, недостаточности свертывания крови, такой как гемофилия. И в этом случае, трансген может кодировать mAb против инфекционного агента или белка, участвующего в развитии воспалительного, иммунного или метаболического состояния, такого как болезнь накопления лизосом.

Ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению может не иметь интрона, расположенного между промотором и трансгеном. Точно так же может отсутствовать интрон между промотором и трансгеном в плазмиде векторного генома (pDNA1) (например, pGM326, как описано в настоящей заявке, и проиллюстрировано на фигуре 2A, с последовательностью SEQ ID NO: 3).

В некоторых предпочтительных вариантах осуществления изобретения, ретровирусный/лентивирусный (например, SIV) вектор содержит промотор hCEF и трансген CFTR, включая промотор и трансген, описанные в настоящей заявке. Необязательно, указанный ретровирусный/лентивирусный (например, SIV) вектор может не иметь интрона, расположенного между промотором и трансгеном. Такой ретровирусный/лентивирусный (например, SIV) вектор может быть получен описанным здесь способом с использованием геномной плазмиды, несущей трансген CFTR и промотор.

В некоторых предпочтительных вариантах осуществления изобретения, ретровирусный/лентивирусный (например, SIV) вектор содержит промотор hCEF и трансген A1AT, включая промотор и трансген, описанные в настоящей заявке. Необязательно, указанный ретровирусный/лентивирусный (например, SIV) вектор может не иметь интрона, расположенного между промотором и трансгеном. Такой ретровирусный/лентивирусный (например, SIV) вектор может быть получен описанным здесь способом с использованием геномной плазмиды, несущей трансген A1AT и промотор.

В некоторых предпочтительных вариантах осуществления изобретения, ретровирусный/лентивирусный (например, SIV) вектор содержит промотор hCEF или CMW и трансген FVIII, включая промотор и трансген, описанные в настоящей заявке. Необязательно, указанный ретровирусный/лентивирусный (например, SIV) вектор может не иметь интрона, расположенного между промотором и трансгеном. Такой ретровирусный/лентивирусный (например, SIV) вектор может быть получен описанным здесь способом с использованием геномной плазмиды, несущей трансген FVIII и промотор.

Ретровирусный/лентивирусный (например, SIV) вектор, описанный в настоящей заявке, содержит трансген. Трансген содержит последовательность нуклеиновой кислоты, кодирующую продукт гена, например белок, а в частности, терапевтический белок.

Так, например, в одном варианте осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая *CFTR*, *A1AT* или *FVIII*, содержит (или состоит из нее) последовательность нуклеиновой кислоты, которая по меньшей мере на 90% (например,

по меньшей мере на 90, 92, 94, 95, 96, 97, 98, 99 или 100%) идентична последовательности нуклеиновой кислоты CFTR, A1AT или FVIII, соответственно, примеры которых описаны в настоящей заявке. В другом варианте осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая *CFTR, A1AT или FVIII*, содержит (или состоит из нее) последовательность нуклеиновой кислоты, которая по меньшей мере на 90% (например, по меньшей мере на 90, 92, 94, 95, 96, 97, 98, 99 или на 100%) идентична последовательности нуклеиновой кислоты CFTR, A1AT или FVIII, соответственно, примеры которых описаны в настоящей заявке. В одном варианте осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая CFTR, представлена SEQ ID NO: 13, последовательность нуклеиновой кислоты, кодирующая A1AT, представлена SEQ ID NO: 14 или комплементарной последовательностью SEQ ID NO: 15 и/или последовательность нуклеиновой кислоты, кодирующая FVIII, представлена SEQ ID NO: 17 и 18, или соответствующими комплементарными последовательностями SEQ ID NO: 19 и 20, или их вариантами.

Аминокислотная последовательность трансгена CFTR, A1AT или FVIII может включать (или состоять из нее) аминокислотную последовательность, которая по меньшей мере на 90% (например, по меньшей мере на 90, 92, 94, 95, 96, 97, 98, 99 или на 100%) идентична функциональной полипептидной последовательности CFTR, A1AT или FVIII, соответственно.

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению могут содержать центральный полипуриновый тракт (сPPT) и/или посттранскрипционные регуляторные элементы вируса гепатита сурка (WPRE). Репрезентативная последовательность WPRE представлена SEQ ID NO: 23.

Методы получения

Как описано в настоящей заявке, авторы настоящего изобретения впервые продемонстрировали, что использование оптимизированных по кодонам генов *gal-pol SIV* не оказывают негативного влияния на полученный титр вектора SIV, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, и может даже приводить к увеличению титра вектора. Кроме того, авторы настоящего изобретения также показали, что использование оптимизированных по кодонам генов *gag-pol* можно также объединить с использованием плазмиды с модифицированным векторным геномом, как описано в настоящей заявке, при сохранении или даже повышении титра вектора.

Оптимизация по кодонам представляет собой метод максимизации экспрессии белка за счет повышения эффективности трансляции кодирующего гена. Эффективность трансляции повышается за счет модификации последовательности нуклеиновой кислоты. Оптимизация по кодонам является рутинной процедурой в данной области, и в рамках обычной практики, любой специалист может разработать оптимизированный по кодонам вариант данной последовательности нуклеиновой кислоты. Однако, не так просто предсказать влияние оптимизации кодонов на другие параметры. Так, например, как

описано в настоящей заявке, согласно общепринятому мнению, при нормальных технологических условиях (если лимитирующим фактором является векторная геномная плазида, а не гены gag-pol), оптимизация по кодомам генов gag-pol обычно снижает выход вектора.

Соответственно, настоящее изобретение относится к способу получения ретровирусного/лентивирусного (например, SIV) вектора, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса и содержащего промотор и трансген, где указанный способ включает использование оптимизированных по кодомам генов gag-pol. Предпочтительно, указанный вектор представляет собой лентивирусный вектор, а особенно предпочтительными являются векторы вируса иммунодефицита обезьян (SIV).

Обычно, оптимизированные по кодомам гены gag-pol, используемые в способах получения согласно изобретению, соответствуют продуцируемому ретровирусному/лентивирусному вектору. В неограничивающем примере, если лентивирусный вектор представляет собой вектор ВИЧ, то оптимизированные по кодомам гены gag-pol, используемые в способах получения согласно изобретению, представляют собой гены gag-pol ВИЧ. В неограничивающем примере, если лентивирусный вектор представляет собой вектор SIV, то оптимизированные по кодомам гены gag-pol, используемые в способах получения согласно изобретению, представляют собой гены gag-pol SIV.

Предпочтительно, оптимизированные по кодомам гены gag-pol, используемые в способах получения согласно изобретению, представляют собой гены gag-pol SIV. Примеры генов gag-pol SIV дикого типа, которые могут быть модифицированы для получения оптимизированных по кодомам генов gag-pol, приведены в SEQ ID NO: 2. Модификации генов gag-pol дикого типа SEQ ID NO: 2 для получения типичных оптимизированных по кодомам генов gag-pol согласно изобретению (SEQ ID NO: 1) показаны в виде сопоставления на Фигуре 1.

В дополнение к оптимизации по кодомам, оптимизированные по кодомам гены gag-pol, используемые в способах получения согласно изобретению, могут включать и другие модификации, такие как сдвиг трансляции (который позволяет смещать трансляцию из одной области в другую для обеспечения продукции обоих Gag и Pol). В оптимизированных по кодомам генах gag-pol согласно изобретению можно использовать любой подходящий вариант оптимизации по кодомам, при условии, что (i) гомология между плазмидой векторного генома и плазмидой GagPol будет снижена для минимизации риска продукции RCL и (ii) после оптимизации кодонов будет наблюдаться продуцирование достаточного количества GagPol без включения RRE (это дополнительно снижает гомологию и риск образования RCL).

Оптимизированные по кодомам гены gag-pol, используемые в способах получения согласно изобретению, могут быть полностью (на 100%) или частично оптимизированными по кодомам. Частичная оптимизация по кодомам включает

оптимизацию по кодонам, которая составляет по меньшей мере 70%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или более.

Предпочтительно, сами гены gag-pol являются полностью оптимизированными по кодонам, но могут содержать области, не включающие последовательности, оптимизированные по кодонам (например, между генами gag и pol). В качестве неограничивающего примера следует отметить, что для поддержания трансляционного сдвига рамок считывания между генами gag и pol, область вокруг последовательности сдвига трансляции не может быть оптимизирована по кодонам (например, в случае, если для этой функции важна точная последовательность сдвига трансляции). Примером не оптимизированной по кодонам последовательности сдвига трансляции в оптимизированных по кодонам генах gag-pol является SEQ ID NO: 1.

Предпочтительно, оптимизированные по кодонам гены gag-pol, используемые в способе согласно изобретению, содержат (или состоят из нее) последовательность нуклеиновой кислоты SEQ ID NO: 1 или ее вариант (как определено в настоящей заявке). В частности, оптимизированные по кодонам гены gag-pol, используемые в способе согласно изобретению, содержат (или состоят из нее) последовательность нуклеиновой кислоты, которая по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или более идентична последовательности SEQ ID NO: 1. Предпочтительно, оптимизированные по кодонам гены gag-pol, используемые в способе согласно изобретению, содержат (или состоят из нее) последовательность нуклеиновой кислоты, которая по меньшей мере на 90%, более предпочтительно по меньшей мере на 95%, а еще более предпочтительно по меньшей мере на 98% или более идентична последовательности SEQ ID NO: 1. Оптимизированные по кодонам гены gag-pol SEQ ID NO: 1 содержат последовательность сдвига трансляции, а поэтому не образуют единой стандартной открытой рамки считывания.

Способ согласно изобретению может быть масштабируемым способом, совместимым с GMP. Таким образом, способ согласно изобретению обычно позволяет получать очищенные ретровирусные/лентивирусные (например, SIV) векторы F/IN с высоким титром. Обычно, способ согласно изобретению дает титр ретровирусного/лентивирусного (например, SIV) вектора, который по меньшей мере эквивалентен титру ретровирусного/лентивирусного (например, SIV) вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодонам гены gag-pol. Используемый здесь термин «эквивалентный» может означать, что использование оптимизированных по кодонам генов gag-pol не приводит к значительному снижению титра ретровирусного/лентивирусного (например, SIV) вектора по сравнению с использованием соответствующих не оптимизированных по кодонам генов gag-pol. В неограничивающем примере, способ согласно изобретению дает титр ретровирусного/лентивирусного (например, SIV) вектора, который не более, чем в 2 раза,

не более, чем в 1,5 раза, не более, чем в 1,0 раза, не более, чем в 0,5 раза, не более, чем в 0,25 раза или менее ниже титра ретровирусного/лентивирусного (например, SIV) вектора по сравнению с титром соответствующих не оптимизированных по кодонам генов *gal-pol*. Термин «эквивалентный» может означать, что титр ретровирусного/лентивирусного (например, SIV) вектора, полученного способом с использованием оптимизированных по кодонам генов *gag-pol*, статистически не изменяется (например, $p < 0,05$, $p < 0,01$) по сравнению с титром ретровирусного/лентивирусного (например, SIV) вектора, полученного способом с использованием соответствующих не оптимизированных по кодонам генов *gal-pol*.

Предпочтительно, способ согласно изобретению дает титр ретровирусного/лентивирусного (например, SIV) вектора, который был выше по сравнению с титром ретровирусного/лентивирусного (например, SIV) вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодонам гены *gal-pol*. Титр ретровирусного/лентивирусного (например, SIV) вектора может быть по меньшей мере в 1,5 раза, по меньшей мере в 2 раза или по меньшей мере в 2,5 раза выше, чем титр ретровирусного/лентивирусного (например, SIV) вектора, полученного соответствующим способом, в котором не использовались оптимизированные по кодонам гены *gal-pol*.

Для получения ретровирусных/лентивирусных (например, SIV) векторов обычно используют одну или более плазмид, которые обеспечивают элементы, необходимые для получения вектора, такие как: геном для ретровирусного/лентивирусного вектора, *Gag-Pol*, *Rev*, *F* и *HN*. Множество элементов могут присутствовать на одной плазмиде. Предпочтительно, каждый элемент представлен на отдельной плазмиде, а поэтому имеется пять плазмид, по одной для каждого векторного генома, *Gag-Pol*, *Rev*, *F* и *HN*, соответственно.

Альтернативно, одна плазида может содержать элементы *Gag-Pol* и *Rev* и может называться упаковывающей плазмидой (*pDNA2*). Остальные элементы (геном, *F* и *HN*) могут быть обеспечены отдельными плазмидами (*pDNA1*, *pDNA3a*, *pDNA3b*, соответственно), а поэтому для получения ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению используются четыре плазмиды. В методах с четырьмя плазмидами, *pDNA1*, *pDNA3a* и *pDNA3b* могут быть такими, как это описано здесь в настоящей заявке применительно к методу с пятью плазмидами.

Предпочтительно, оптимизированные по кодонам гены *gag-pol*, используемые в способе согласно изобретению, содержатся в плазмиде, которая включает последовательность нуклеиновой кислоты или состоит из последовательности нуклеиновой кислоты SEQ ID NO: 5 (*pGM691*) или ее варианта (как определено в настоящей заявке). В частности, оптимизированные по кодонам гены *gag-pol*, используемые в способе согласно изобретению, содержатся в плазмиде, которая включает последовательность нуклеиновой кислоты или состоит из последовательности нуклеиновой кислоты, которая по меньшей мере на 80%, по меньшей мере на 85%, по

меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или более идентична последовательности SEQ ID NO: 5. Предпочтительно, оптимизированные по кодомам гены gag-pol, используемые в способе согласно изобретению, содержатся в плазмиде, которая включает последовательность нуклеиновой кислоты или состоит из последовательности нуклеиновой кислоты, которая по меньшей мере на 90%, более предпочтительно по меньшей мере на 95%, а еще более предпочтительно по меньшей мере на 98% или более идентична последовательности SEQ ID NO: 5. В плазмиде SEQ ID NO: 5 (или в ее вариантах): (i) оптимизированные по кодомам гены gag-pol SEQ ID NO: 1 содержат последовательность сдвига трансляции и, таким образом, не образуют одну обычную открытую рамку считывания; и (ii) оптимизированные по кодомам гены gag-pol SEQ ID NO: 1 функционально связаны с промотором CAG.

В предпочтительном способе согласно изобретению, в котором используются пять плазмид, плазида векторного генома кодирует весь генетический материал, который упаковывается в конечный ретровирусный/лентивирусный вектор, включая трансген. Обычно только часть генетического материала, обнаруженного в плазмиде векторного генома, присутствует в вирусе. Плазида векторного генома может быть обозначена здесь как «pDNA1» и обычно содержит трансген и промотор трансгена.

Остальные четыре плазмиды представляют собой полученные плазмиды, кодирующие белки Gag-Pol, Rev, F и HN. Эти плазмиды могут быть обозначены как «pDNA2a», «pDNA2b», «pDNA3a» и «pDNA3b», соответственно.

Модификации могут быть внесены в плазмиду генома вектора (pDNA1), в частности, для дальнейшего улучшения профиля безопасности вектора. Как показано в настоящей заявке, такие модификации могут включать модификацию или состоять из модификации последовательности pDNA1 для удаления вирусных, а особенно ретровирусных/лентивирусных (например, SIV) OРС из последовательности pDNA1. Таким образом, в способах согласно изобретению может быть использована модифицированная pDNA1, которая содержит уменьшенное количество нетрансгенных OРС. Указанная модифицированная pDNA1 может содержать модификации в любой области плазмидной последовательности. В частности, модифицированная pDNA1 может включать модификации для удаления: (i) 5'-3' OРС; (ii) OРС из ≥ 100 аминокислот; и/или (iii) OРС выше трансгена и/или промотора, функционально связанного с трансгеном. Хотя модифицированная pDNA1 может не содержать OРС, кроме трансгена, однако, это не играет важной роли. Скорее всего, модифицированная pDNA1 может по-прежнему содержать OРС, отличающиеся от трансгена, но может содержать меньшее количество нетрансгенных OРС по сравнению с немодифицированной pDNA1, из которой она была получена. В качестве неограничивающего примера может служить модифицированная pDNA1, которая может содержать по меньшей мере на 1, по меньшей мере на 2, по меньшей мере на 3, по меньшей мере на 4, по меньшей мере на 5 или т.п. меньше нетрансгенных OРС по сравнению с соответствующей немодифицированной pDNA1. В

качестве конкретного примера может служить рGM830 (которая происходит от рGM326), содержащая на 2 нетрансгенных ОРС меньше, чем рGM326. Модифицированная рDNA1 может содержать по меньшей мере 1, по меньшей мере 5, по меньшей мере 6, по меньшей мере 7, по меньшей мере 8, по меньшей мере 9, по меньшей мере 10, по меньшей мере 15, по меньшей мере 20 или более модификаций (например, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 или 20 модификаций) по сравнению с соответствующей немодифицированной рDNA1. В качестве неограничивающего примера может служить модифицированная рDNA1, которая может содержать от приблизительно 1 до приблизительно 20, например, от приблизительно 5 до приблизительно 15, или от приблизительно 5 до приблизительно 10 модификаций по сравнению с соответствующей немодифицированной рDNA1. В качестве конкретного примера может служить рGM830 (которая происходит от рGM326), содержащая 7 модификаций по сравнению с рGM326.

Как показано в примерах, использование рGM380 в качестве плазмиды рDNA1 потенциально может обеспечить повышенный титр SIV по сравнению с титром, полученным в способе получения, в котором плазмидой рDNA1 является рGM326 (фиг. 11), но в котором все другие плазмиды и параметры метода остаются постоянными. Другими словами, использование модифицированной рDNA1, такой как рGM830, не оказывает негативного влияния на повышенный титр, достигаемый с помощью оптимизированных по кодонам генов *gal-pol*, и даже потенциально может обеспечить дальнейшее повышение титра по сравнению с эффектом при использовании оптимизированных по кодонам генов *gal-pol*, таких как гены, полученные с использованием рGM691 в качестве рDNA2a. Термин «повышенный титр», определенный в настоящей заявке, в равной степени относится к способам согласно изобретению, в которых используются как оптимизированные по кодонам гены *gal-pol*, так и модифицированная рDNA1.

Обычно лентивирус представляет собой SIV, такой как SIV1, а предпочтительно SIV-AGM. Белки F и HN происходят от респираторного парамиксовируса, а предпочтительно от вируса Сендай.

В конкретном варианте осуществления изобретения, относящемся к CFTR, пять плазмид представлены на фигурах 2A-2F, и таким образом, рDNA1 представляет собой плазмиду рGM326 на фигуре 2A или плазмиду рGM830 на фигуре 2B, рDNA2a представляет собой плазмиду рGM691 на фигуре 2C, рDNA2b представляет собой плазмиду рGM299 на фигуре 2D, рDNA3a представляет собой плазмиду рGM301 на фигуре 2E, а рDNA3b представляет собой плазмиду рGM303 на фигуре 2F или варианты любой из этих плазмид (как описано в настоящей заявке). В этом варианте осуществления изобретения, конечный ретровирусный/лентивирусный вектор, содержащий CFTR, может обозначаться как vGM195 (см. Примеры). Плазмида рGM691 и вектор vGM195 являются предпочтительными вариантами осуществления изобретения.

Как показано в настоящей заявке, использование рGM691 в качестве плазмиды рDNA2a потенциально может обеспечивать повышенный титр SIV по сравнению с

титром, достигаемым в способе получения, в котором плазмидой pDNA2a является pGM297 (фигура 2G), но в котором все другие плазмиды и параметры метода поддерживаются постоянными.

Если способ согласно изобретению используют для получения A1AT, то пять плазмид могут быть охарактеризованы как показано на фигуре 3 (таким образом, плазида pDNA1 может представлять собой pGM407) и на всех фигурах 2C-F (как указано выше для конкретного варианта CFTR) или могут быть охарактеризованы варианты любой из этих плазмид (как описано в настоящей заявке).

Если способ согласно изобретению используют для получения FVIII, то пять плазмид могут быть охарактеризованы как указано на одной из фигур 4AD (таким образом, плазида pDNA1 может представлять собой pGM411, pGM412, pGM413 или pGM414) и на всех фигурах 2C-F или могут быть охарактеризованы варианты любой из этих плазмид (как описано в настоящей заявке).

Плазида, определенная на фигуре 2A, представлена SEQ ID NO: 3; плазида, определенная на фигуре 2B, представлена SEQ ID NO: 4; плазида, определенная на фигуре 2C, представлена SEQ ID NO: 5; плазида, определенная на фигуре 2D, представлена SEQ ID NO: 6; плазида, определенная на фигуре 2E, представлена SEQ ID NO: 7; плазида, определенная на фигуре 2F, представлена SEQ ID NO: 8; плазида, определенная на фигуре 2G, представлена SEQ ID NO: 9; плазида, определенная на фигуре 3, представлена SEQ ID NO: 24, а плазмиды F/HN-SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-со и/или F/HN-SIV-hCEF-HFVIII-N6-со, определенные на фигурах 4A-4D, представлены SEQ ID NO: 25-28, соответственно. Варианты (как определено в настоящей заявке) этих плазмид также входят в объем настоящего изобретения. В частности, в объем настоящего изобретения входят варианты, имеющие последовательность, которая по меньшей мере на 90% (например, по меньшей мере на 90, 92, 94, 95, 96, 97, 98, 99, 99,5 или 100%) идентична последовательности любой из SEQ ID NO: 3-9, 24 и 25-28.

В способе согласно изобретению, проводимом с использованием пяти плазмид, все пять плазмид вносят свой вклад в образование конечного ретровирусного/лентивирусного (например, SIV) вектора. Во время получения ретровирусного/лентивирусного (например, SIV) вектора, плазида векторного генома (pDNA1) будет включать энхансер/промотор, Psi, RRE, cPPT, mWPRE, SIN LTR, SV40 polyA (см. фигуру 2A или 2B), которые имеют важное значение для получения вирусов. С использованием pGM326 или pGM830 в качестве неограничивающих примеров pDNA1, было обнаружено, что энхансер/промотор CMV, polyA SV40, colE1 Ori и KanR участвуют в продуцировании ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению (например, vGM195 или vGM244), но они не были обнаружены в конечном ретровирусном/лентивирусном (например, SIV) векторе. RRE, cPPT (центральный полипуриновый тракт), hCEF, soCFTR2 (трансен) и mWPRE из pGM326 или pGM830 были обнаружены в конечном ретровирусном/лентивирусном (например, SIV) векторе.

SIN LTR (длинные концевые повторы, самоинактивирующиеся SIN/IN) и Psi (сигнал упаковки) могут быть обнаружены в конечном ретровирусном/лентивирусном (например, SIV) векторе.

Для других ретровирусных/лентивирусных (например, SIV) векторов согласно изобретению, соответствующие элементы из плазмид генома другого вектора (pDNA1) необходимы для получения вектора (но не обнаружены в конечном векторе) или присутствуют в конечном ретровирусном/лентивирусном (например, SIV) векторе.

Белки F и HN pDNA3a и pDNA3b (предпочтительно белки F и HN вируса Сендай) имеют важное значение для инфицирования клеток-мишеней конечным ретровирусным/лентивирусным (например, SIV) вектором, то есть, для проникновения в эпителиальные клетки пациента (обычно клетки легких или носа, как описано в настоящей заявке). Продукты плазмид pDNA2a и pDNA2b имеют важное значение для трансдукции вируса, то есть, для встраивания ретровирусной/лентивирусной (например, SIV) ДНК в геном хозяина. Промотор, регуляторные элементы (такие как WPRE) и трансген имеют важное значение для экспрессии трансгена в клетке(ах)-мишени(ях).

Способ согласно изобретению может включать, или состоять из них, следующие стадии: (a) выращивание клеток в суспензии; (b) трансфекцию клеток одной или более плазмидами; (c) добавление нуклеазы; (d) сбор лентивируса (например, SIV); (e) добавление трипсина; и (f) очистку лентивируса (например, SIV).

В этом способе может использоваться описанная здесь система на основе четырех или пяти плазмид. Таким образом, для проведения предпочтительного способа с пятью плазмидами, одна или более плазмид могут включать, или состоять из них, векторную геномную плазмиду pDNA1; плазмиду, содержащую gagpol, pDNA2a; плазмиду Rev, pDNA2b; плазмиду слитого белка (F), pDNA3a; и плазмиду гемагглютинирина-нейраминидазы (HN), pDNA3b. pDNA1 может быть выбрана из pGM326 и pGM830, а предпочтительно pGM830. pDNA2a может представлять собой pGM691. pDNA2b может представлять собой pGM299. pDNA3a может представлять собой pGM301. pDNA3b может представлять собой pGM303. При этом, может быть использована любая комбинация pDNA1, pDNA2a, pDNA2b, pDNA3a и pDNA3b. Предпочтительно, pDNA1 представляет собой pGM326 или pGM830 (особенно предпочтительной является pGM830); pDNA2a представляет собой pGM691; pDNA2b представляет собой pGM299; pDNA3a представляет собой pGM301; и pDNA3b представляет собой pGM303. Вектор SIV, полученный с использованием pGM830, pGM691, pGM299, pGM301 и pGM303, обозначается как vGM244. Вектор SIV, полученный с использованием pGM326, pGM691, pGM299, pGM301 и pGM303, обозначается как vGM195.

Любое подходящее соотношение плазмиды векторного генома: плазмиды gagpol : плазмиды Rev : плазмиды F : плазмиды HN может быть использовано для дальнейшей оптимизации (увеличения) титра продуцируемого ретровируса/лентивируса (например, SIV). В неограничивающем примере, соотношение плазмиды векторного генома : плазмиды co-gagpol : плазмиды Rev : плазмиды F : плазмиды HN может находиться в

диапазоне 10-40:4-20:3-12:3-12:3-12, а обычно 15-20:7-11:4-8:4-8:4-8, например, приблизительно 18-22:7-11:4-8:4-8:4-8,19-21:8-10:5-7:5-7:5-7. Предпочтительно, соотношение плазмиды векторного генома: плазмиды со-gagpol : плазмиды Rev: плазмиды F : плазмиды HN составляет приблизительно 20:9:6:6:6.

Стадии (a)-(f) этого способа обычно осуществляют последовательно, начиная со стадии (a) и продолжая до стадии (f). Этот способ может включать одну или более дополнительных стадий, таких как дополнительные стадии очистки, замену буфера, определение концентрации ретровирусного/лентивирусного (например, SIV) вектора после очистки и/или приготовление ретровирусного/лентивирусного (например, SIV) вектора после очистки (или определения концентрации). Каждая из этих стадий может содержать одну или более подстадий. Так, например, сбор может включать одну или более стадий или подстадий, и/или очистка может включать одну или более стадий или подстадий.

Клетки любого подходящего типа могут быть трансфицированы одной или более плазмидами (например, пятью плазмидами, описанными в настоящей заявке) с получением ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению. Обычно используют клетки млекопитающих, особенно линии клеток человека. Неограничивающими примерами клеток, подходящих для их использования в способах согласно изобретению, являются клетки HEK293 (такие как клетки HEK293F или HEK293T) и клетки 293T/17. Коммерчески доступные клеточные линии, подходящие для продукции вируса, также являются легко доступными (см. например, Gibco Viral Production Cells - Номер по каталогу A35347 от ThermoFisher Scientific).

Клетки могут быть выращены в среде, не содержащей компонентов животного происхождения, включая среду, не содержащую сыворотку. Клетки могут быть выращены в среде, содержащей человеческие компоненты. Клетки могут быть выращены в определенной среде, содержащей компоненты, полученные путем синтеза, или в среде, состоящей из этих компонентов.

В соответствии с настоящим изобретением можно использовать любые подходящие средства для трансфекции. Выбор подходящих средств для трансфекции является обычной практикой для специалиста в данной области. В неограничивающем примере, трансфекцию можно проводить с использованием PEIProTM, Липофектамина 2000TM или Липофектамина 3000TM.

В настоящем изобретении может быть использована любая подходящая нуклеаза. Выбор подходящей нуклеазы является обычной практикой для специалиста в данной области. Обычно, нуклеаза представляет собой эндонуклеазу. В неограничивающем примере, нуклеаза может представлять собой Бенсоназу® или Денаразу®. Добавление нуклеазы может быть проведено на стадии перед сбором или на стадии после сбора, или между стадиями сбора.

Активность трипсина может быть предпочтительно обеспечена рекомбинантным ферментом не животного происхождения, таким как фермент TrypLE SelectTM.

Добавление трипсина может быть проведено на стадии перед сбором или на стадии после сбора, или между стадиями сбора.

Для очистки ретровирусного/лентивирусного (например, SIV) вектора могут быть использованы любые подходящие средства очистки. Неограничивающие примеры подходящих стадий очистки включают глубокую/концевую фильтрацию, фильтрацию в тангенциальном потоке (TFF) и хроматографию. Стадия очистки обычно включает по меньшей мере стадию хроматографии. Неограничивающие примеры хроматографических стадий, которые могут быть проведены в соответствии с настоящим изобретением, включают смешанную эксклюзионную хроматографию (SEC) и/или анионообменную хроматографию. Элюирование может быть проведено с использованием или без использования солевого градиента, а предпочтительно без него.

Этот способ может быть применен для получения ретровирусных/лентивирусных (например, SIV) векторов согласно изобретению, таких как векторы, содержащие ген *CFTR*, *A1AT* и/или *FVIII*, как описано в настоящей заявке. Альтернативно, ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению содержит любой из вышеупомянутых генов или гены, кодирующие вышеупомянутые белки.

В способе согласно изобретению может быть использована любая комбинация одной или более специфических плазмидных конструкций, представленных на фигурах 2A-2F, фигуре 3 и/или на фигурах 4A-4D, для получения ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению. В частности, используют плазмидные конструкции, представленные на фигурах 2C-2F, предпочтительно в сочетании с плазмидой, представленной на фигуре 2B, фигуре 2A, фигуре 3 или фигуре 4A-4D, причем, особенно предпочтительной является плазида, изображенная на фигуре 2B.

Настоящее изобретение также относится к оптимизированным по кодонам генам gag-pol SIV. Эти оптимизированные по кодонам гены gag-pol SIV обычно являются подходящими для их использования в способах согласно изобретению. Оптимизированные по кодонам гены gag-pol согласно изобретению могут содержать, или состоять из них, последовательность нуклеиновой кислоты SEQ ID NO: 1 или ее вариант (как определено в настоящей заявке). В частности, оптимизированные по кодонам гены gag-pol согласно изобретению могут содержать, или состоять из нее, последовательность нуклеиновой кислоты, которая по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или более идентична последовательности SEQ ID NO: 1. Предпочтительно, оптимизированные по кодонам гены gag-pol согласно изобретению могут содержать, или состоять из нее, последовательность нуклеиновой кислоты, которая по меньшей мере на 90%, более предпочтительно по меньшей мере на 95%, еще более предпочтительно по меньшей мере на 98% или более идентична последовательности SEQ ID NO: 1. Соответственно, настоящее изобретение относится к нуклеиновой кислоте, содержащей оптимизированные по кодонам гены gag-pol, где указанная нуклеиновая кислота имеет последовательность,

которая по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или более идентична последовательности SEQ ID NO: 1, а предпочтительно по меньшей мере на 90%, более предпочтительно по меньшей мере на 95%, а еще более предпочтительно по меньшей мере на 98% или более идентична последовательности SEQ ID NO: 1. В своем особенно предпочтительном варианте, настоящее изобретение относится к нуклеиновой кислоте, которая содержит, или состоит из нее, последовательность нуклеиновой кислоты SEQ ID NO: 1. Оптимизированные по кодонам гены gag-pol (например, гены gag-pol SIV) согласно изобретению обычно функционально связаны с промотором для облегчения экспрессии белков gag-pol. При этом, может быть использован любой подходящий промотор, включая промоторы, описанные здесь как промоторы для трансгена. Предпочтительно, промотор представляет собой промотор CAG, который используется в репрезентативной плазмиде pGM691. Типичный промотор CAG представлен в SEQ ID NO: 29. Оптимизированные по кодонам гены gag-pol SEQ ID NO: 1 содержат последовательность сдвига трансляции и, таким образом, не образуют одну обычную открытую рамку считывания.

Настоящее изобретение также относится к плазмидам, содержащим оптимизированные по кодонам гены gag-pol SIV согласно изобретению, то есть, pDNA2a, включающую оптимизированные по кодонам гены gag-pol SIV согласно изобретению. Эти плазмиды обычно являются подходящими для их использования в способах согласно изобретению. Плазида (pDNA2a) согласно изобретению может содержать, или состоять из них, последовательность нуклеиновой кислоты SEQ ID NO: 5 (pGM691) или ее вариант (как определено в настоящей заявке). В частности, плазида (pDNA2a) согласно изобретению может содержать, или состоять из нее, последовательность нуклеиновой кислоты, которая по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или более идентична последовательности SEQ ID NO: 5. Предпочтительно, плазида (pDNA2a) согласно изобретению может содержать или состоять из нее, последовательность нуклеиновой кислоты, которая по меньшей мере на 90%, более предпочтительно, по меньшей мере на 95%, а еще более предпочтительно по меньшей мере на 98% или более идентична последовательности SEQ ID NO: 5. Соответственно, настоящее изобретение относится к плазмиде, содержащей оптимизированные по кодонам гены gag-pol SIV согласно изобретению (как определено в настоящей заявке), а в частности, к последовательности нуклеиновой кислоты, включающей, или состоящей из них, SEQ ID NO: 1 или ее вариант (как определено в настоящей заявке). Указанная плазида может содержать, или состоять из нее, последовательность нуклеиновой кислоты, которая по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98%, по меньшей мере на 99% или более идентична SEQ ID NO: 5, а предпочтительно по меньшей мере на 90%, более

предпочтительно по меньшей мере на 95%, еще более предпочтительно по меньшей мере на 98% или более идентична последовательности SEQ ID NO: 5. В своем особенно предпочтительном варианте, настоящее изобретение относится к плазмиде, которая содержит, или состоит из нее, последовательность нуклеиновой кислоты SEQ ID NO: 5. В плазмиде SEQ ID NO: 5 (или в ее вариантах): (i) оптимизированные по кодонам гены gag-pol SEQ ID NO: 1 содержат последовательность сдвига трансляции, а поэтому не образуют одну стандартную открытую рамку считывания; и (ii) оптимизированные по кодонам гены gag-pol SEQ ID NO: 1 функционально связаны с промотором CAG (например, как показано в настоящей заявке).

Оптимизированные по кодонам гены gag-pol (или нуклеиновые кислоты, содержащие или состоящие из них) и плазмиды, содержащие указанные гены или нуклеиновые кислоты, имеют преимущества при получении ретровирусных/лентивирусных (например, SIV) векторов с применением способов согласно изобретению, поскольку они позволяют получать ретровирусные/лентивирусные (например, SIV) векторы F/НN с высоким титром. Обычно, указанные оптимизированные по кодонам гены gag-pol (или нуклеиновые кислоты, содержащие или состоящие из них) и плазмиды, содержащие указанные гены или нуклеиновые кислоты, могут быть использованы для получения титра ретровирусного/лентивирусного (например, SIV) вектора, который по меньшей мере эквивалентен титру ретровирусного/лентивирусного (например, SIV) вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодонам гены gal-pol, как описано в настоящей заявке.

Предпочтительно, оптимизированные по кодонам гены gag-pol (или нуклеиновые кислоты, содержащие или состоящие из них) и плазмиды, содержащие указанные гены или нуклеиновые кислоты, позволяют получать титр ретровирусного/лентивирусного (например, SIV) вектора, который был повышен по сравнению с титром ретровирусного/лентивирусного (например, SIV) вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодонам гены gal-pol, как описано в настоящей заявке.

Настоящее изобретение также относится к клеткам-хозяевам, содержащим: (i) ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению, (ii) оптимизированные по кодонам гены gag-pol (или нуклеиновую кислоту, содержащую или состоящую из них) согласно изобретению; и/или (iii) плазмиду, содержащую указанные гены или нуклеиновую кислоту; или любую их комбинацию. Обычно, клеткой-хозяином является клетка млекопитающего, а в частности, клетка или клеточная линия человека. Неограничивающие примеры клеток-хозяев включают клетки HEK293 (такие как клетки HEK293F или HEK293T) и клетки 293T/17. Коммерчески доступные клеточные линии, подходящие для продукции вируса, также являются легко доступными (как описано в настоящей заявке).

Настоящее изобретение также относится к ретровирусному/лентивирусному

(например, SIV) вектору, который может быть получен способом согласно изобретению или с использованием оптимизированных по кодонам генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), плазмиды, содержащей указанные гены или нуклеиновые кислоты, или клетки-хозяина согласно изобретению.

Обычно ретровирусный/лентивирусный (например, SIV) вектор, который может быть получен способом согласно изобретению или с использованием оптимизированных по кодонам генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), плазмиды, содержащей указанные гены или нуклеиновые кислоты, или клетки-хозяина согласно изобретению, продуцируется с высоким титром. Титр может быть оценен в единицах трансдукции, как определено в настоящей заявке. Как описано в настоящей заявке, способы согласно изобретению обычно дают ретровирусный/лентивирусный (например, SIV) вектор с эквивалентными или более высокими титрами, чем титры, полученные соответствующими способами, в которых не используются оптимизированные по кодонам гены gag-pol. Соответственно, ретровирусный/лентивирусный (например, SIV) вектор, который может быть получен способом согласно изобретению или с использованием оптимизированных по кодонам генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), плазмиды, содержащей указанные гены или нуклеиновые кислоты, или клетки-хозяина согласно изобретению, может, но необязательно, иметь титр по меньшей мере приблизительно $2,5 \times 10^6$ ЕТ/мл, по меньшей мере приблизительно $3,0 \times 10^6$ ЕТ/мл, по меньшей мере приблизительно $3,1 \times 10^6$ ЕТ/мл, по меньшей мере приблизительно $3,2 \times 10^6$ ЕТ/мл, по меньшей мере приблизительно $3,3 \times 10^6$ ЕТ/мл, по меньшей мере приблизительно $3,4 \times 10^6$ ЕТ/мл, по меньшей мере около $3,5 \times 10^6$ ЕТ/мл, по меньшей мере около $3,6 \times 10^6$ ЕТ/мл, по меньшей мере около $3,7 \times 10^6$ ЕТ/мл, по меньшей мере приблизительно $3,8 \times 10^6$ ЕТ/мл, по меньшей мере приблизительно $3,9 \times 10^6$ ЕТ/мл, по меньшей мере приблизительно $4,0 \times 10^6$ ЕТ/мл или более. Предпочтительно, ретровирусный/лентивирусный (например, SIV) вектор получают с титром по меньшей мере приблизительно $3,0 \times 10^6$ ЕТ/мл или по меньшей мере приблизительно $3,5 \times 10^6$ ЕТ/мл.

Получение ретровирусных/лентивирусных (например, SIV) векторов с высоким титром может сообщать другие желательные свойства полученным векторным продуктам. Так, например, не ограничиваясь какой-либо конкретной теорией, авторы лишь отмечают, что получение высоких титров без необходимости интенсивного концентрирования с применением таких методов, как TFF, позволяет получить векторный продукт более высокого качества, чем ретровирусные/лентивирусные (например, SIV) векторы, полученные соответствующими методами без использования оптимизированных по кодонам генов gag-pol (и необязательно, плазмиды с модифицированным геномом вектора), поскольку векторы подвергаются меньшим сдвиговым усилиям, которые могут повредить вирусные частицы и их РНК-груз.

Настоящее изобретение также относится к способу увеличения титра ретровирусного/лентивирусного (например, SIV) вектора, включающему использование

оптимизированных по кодонам генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), плазмиды, содержащей указанные гены или нуклеиновые кислоты, или клетки-хозяина согласно изобретению. Указанный способ повышения титра ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению может повышать титр по меньшей мере в 1,5 раза, по меньшей мере в 2 раза или по меньшей мере в 2,5 раза или более по сравнению с соответствующим методом, в котором используются не оптимизированные по кодомам варианты генов gag-pol (или нуклеиновые кислоты, содержащие или состоящие из них), или плазмиды или клетки-хозяева, содержащие указанные неоптимизированные по кодомам гены или нуклеиновые кислоты. Альтернативно, способ увеличения титра ретровируса/лентивируса (например, SIV) согласно изобретению может повысить титр по меньшей мере приблизительно на 25%, по меньшей мере приблизительно на 50%, по меньшей мере приблизительно на 100%, по меньшей мере приблизительно на 150%, по меньшей мере приблизительно на 200% или более по сравнению с титром, полученным соответствующим способом, в котором используются не оптимизированные по кодомам варианты генов gag-pol (или нуклеиновые кислоты, содержащие или состоящие из них), или плазмиды или клетки-хозяева, содержащие указанные неоптимизированные по кодомам гены или нуклеиновые кислоты. Предпочтительно, способ увеличения титра ретровируса/лентивируса (например, SIV) согласно изобретению может повышать титр (a) по меньшей мере в 1,5 раза или по меньшей мере в 2 раза; и/или (b) по меньшей мере приблизительно на 25%, более предпочтительно по меньшей мере приблизительно на 50%, а еще более предпочтительно по меньшей мере приблизительно на 100%. Обычно, соответствующий способ идентичен способу согласно изобретению, за исключением использования оптимизированных по кодомам генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), плазмиды, содержащей указанные гены или нуклеиновые кислоты, или клетки-хозяина согласно изобретению. Все раскрытие здесь способы получения ретровирусного/лентивирусного (например, SIV) вектора применимы в равной степени и без оговорок к способам повышения титра ретровируса/лентивируса (например, SIV) согласно изобретению.

Настоящее изобретение также относится к применению оптимизированных по кодомам генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), плазмиды, содержащей указанные гены или нуклеиновые кислоты, или клетки-хозяина согласно изобретению для увеличения титра ретровирусного/лентивирусного (например, SIV) вектора. Указанное применение может повышать титр ретровирусного/лентивирусного (например, SIV) вектора по меньшей мере в 1,5 раза, по меньшей мере в 2 раза или по меньшей мере в 2,5 раза или более по сравнению с применением соответствующего неоптимизированного по кодомам варианта генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), или плазмид, или клеток-хозяев, содержащих указанные неоптимизированные по кодомам гены или нуклеиновые кислоты. В качестве альтернативы, указанное применение может повышать титр

ретровируса/лентивируса (например, SIV) по меньшей мере приблизительно на 25%, по меньшей мере приблизительно на 50%, по меньшей мере приблизительно на 100%, по меньшей мере приблизительно на 150%, по меньшей мере приблизительно на 200% или более по сравнению с применением соответствующего неоптимизированного по кодонам варианта генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), или плазмид, или клеток-хозяев, содержащих указанные не оптимизированные по кодонам гены или нуклеиновые кислоты. Предпочтительно, указанное применение увеличивает титр ретровируса/лентивируса (например, SIV) в (a) по меньшей мере в 1,5 раза или по меньшей мере в 2 раза; и/или (b) по меньшей мере приблизительно на 25%, более предпочтительно по меньшей мере приблизительно на 50%, а еще более предпочтительно по меньшей мере приблизительно на 100%. Обычно, соответствующее применение идентично способу согласно изобретению, за исключением использования оптимизированных по кодонам генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), плазмиды, содержащей указанные гены или нуклеиновые кислоты, или клетки-хозяина согласно изобретению. Все раскрытые в настоящей заявке способы получения ретровирусного/лентивирусного (например, SIV) вектора применимы в равной степени и без оговорок к использованию оптимизированных по кодонам генов gag-pol (или нуклеиновых кислот, содержащих или состоящих из них), плазмиды, содержащей указанные гены или нуклеиновые кислоты, или клетки-хозяина согласно изобретению для увеличения титра ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению. Использование оптимизированных по кодонам генов gag-pol в сочетании с модифицированной плазмидой векторного генома (с уменьшенными вирусными ОРС) может обеспечить дополнительное преимущество с точки зрения безопасности и/или титра вектора. Таким образом, повышенные выходы векторов, как описано в настоящей заявке, могут быть достигнуты с использованием только оптимизированных по кодонам генов gag-pol или в комбинации с модифицированной плазмидой генома вектора. Любые и все варианты осуществления изобретения, относящиеся к повышению титра вектора в способе с использованием оптимизированных по кодонам генов gag-pol, применимы в равной степени и без оговорок к способам, проводимым с использованием оптимизированных по кодонам генов gag-pol в комбинации с модифицированной плазмидой векторного генома согласно изобретению, и к векторам, полученным такими способами.

Терапевтические показания

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению обеспечивают более высокую и устойчивую экспрессию генов посредством эффективного переноса генов. F/HN-псевдотипированные ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению обладают такими свойствами, как: (i) трансдукция дыхательных путей без нарушения целостности эпителия; (ii) постоянная экспрессия генов; (iii) отсутствие хронической токсичности; и (iv) возможность эффективного повторного введения. Длительная/стойкая стабильная экспрессия гена, предпочтительно

на терапевтически эффективном уровне, может быть достигнута с использованием повторных доз вектора согласно изобретению. Альтернативно, однократная доза может быть использована для достижения желаемой длительной экспрессии.

Таким образом, ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению преимущественно можно использовать в генотерапии. Так, например, способность ретровирусных/лентивирусных (например, SIV) векторов согласно изобретению эффективно поглощать клетки дыхательных путей делает их в высокой степени подходящими для лечения заболеваний дыхательных путей. Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению также могут быть использованы в способах генотерапии для стимуляции секреции терапевтических белков. В дополнительном примере, настоящее изобретение относится к секреции терапевтических белков в просвет дыхательных путей или системы кровообращения. Таким образом, введение ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению и его поглощение клетками дыхательных путей позволяет использовать легкие (или нос, или дыхательные пути) в качестве «фабрики» для производства терапевтического белка, который затем секретируется и попадает в общий кровоток на терапевтических уровнях, где он может перемещаться в представляющие интерес клетки/ткани и вызывать терапевтический эффект. В отличие от внутриклеточных или мембранных белков, продуцирование таких секретлируемых белков не зависит от трансдукции клеток-мишеней при конкретном заболевании, что является значительным преимуществом и позволяет достичь высоких уровней экспрессии белка. Таким образом, другие заболевания, которые не являются заболеваниями дыхательных путей, такие как сердечно-сосудистые заболевания и заболевания крови, а в частности, недостаточность свертывания крови, также можно лечить с помощью ретровирусных/лентивирусных (например, SIV) векторов согласно изобретению.

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению позволяют эффективно лечить заболевание посредством введения трансгена для коррекции заболевания. Так, например, доставка функциональной копии гена CFTR *может быть* осуществлена для ослабления симптомов или профилактики заболевания легких у пациентов с CF, независимо от лежащей в их основе мутации. Соответственно, ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению могут быть использованы для лечения кистозного фиброза (CF), обычно, с помощью генотерапии с использованием трансгена CFTR, как описано в настоящей заявке.

В другом примере ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению могут быть использованы для устранения дефицита альфа-1-антитрипсина (A1AT), обычно, с помощью генотерапии с использованием трансгена A1AT, как описано в настоящей заявке. A1AT представляет собой секретлируемое антитело против протеазы, которое вырабатывается в основном в печени, а затем переносится в легкие, при этом, в самих легких также вырабатываются A1AT, но в меньших количествах. Основная функция A1AT заключается в связывании и

нейтрализации/ингибировании эластазы нейтрофилов. Генотерапия с помощью A1AT согласно изобретению применима к пациентам с дефицитом A1AT, а также к другим заболеваниям легких, таким как CF или хроническая обструктивная болезнь легких (ХОБЛ), и дает возможность решить некоторые проблемы, возникающие при традиционной фермент-заместительной терапии (при которой A1AT выделяют из крови человека и вводят внутривенно каждую неделю), что будет обеспечивать стабильную, длительную экспрессию в ткани-мишени (в эпителии легкого/носа), простоту введения и неограниченную доступность.

Трансдукция ретровирусным/лентивирусным (например, SIV) вектором согласно изобретению может приводить к секреции рекомбинантного белка в просвет легкого, а также в кровотоки. Одним из преимуществ такого процесса является то, что терапевтический белок достигает интерстициальной ткани. Таким образом, генотерапия с использованием A1AT также может быть полезной при лечении других заболеваний, не ограничиваемыми примерами которых являются диабет типа 1 и 2, острый инфаркт миокарда, ишемическая болезнь сердца, ревматоидный артрит, воспалительное заболевание кишечника, отторжение трансплантата, реакция «трансплантат против хозяина» (GvH), рассеянный склероз, заболевание печени, цирроз, васкулиты и инфекции, такие как бактериальные и/или вирусные инфекции.

A1AT обладает множеством других противовоспалительных и тканезащитных эффектов, например, в доклинических моделях диабета, реакции «трансплантат против хозяина» и воспалительных заболеваниях кишечника. Таким образом, продуцирование A1AT в легком и/или в носовой полости после трансдукции согласно изобретению может иметь более широкое применение, включая эти показания.

Другими примерами заболеваний, которые можно лечить с помощью генотерапии с использованием секретируемого белка согласно изобретению, являются сердечно-сосудистые заболевания и заболевания крови, а в частности, нарушения свертывания крови, такие как гемофилия (A, B или C), болезнь фон Виллебранда и дефицит фактора VII.

Другими примерами заболеваний или расстройств, подлежащих лечению, являются первичная цилиарная дискинезия (PCD), острое повреждение легких, дефицит поверхностно-активного белка В (SFTB), легочный альвеолярный протеиноз (PAP), хроническая обструктивная болезнь легких (ХОБЛ) и/или воспалительные, инфекционные, иммунные или метаболические состояния, такие как болезни запасания лизосом.

Соответственно, настоящее изобретение относится к способу лечения заболевания, включающему введение ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению индивидууму. Обычно, ретровирусный/лентивирусный (например, SIV) вектор получают с применением способа согласно изобретению. Любое описанное здесь заболевание может быть подвергнуто лечению в соответствии с настоящим изобретением. В частности, настоящее изобретение относится к способу лечения

заболевания легких с использованием ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению. Заболевание, подлежащее лечению, может быть хроническим заболеванием. Предпочтительным является способ лечения CF.

Настоящее изобретение также относится к описанному здесь ретровирусному/лентивирусному (например, SIV) вектору для применения в способе лечения заболевания. Обычно, ретровирусный/лентивирусный (например, SIV) вектор получают с применением способа согласно изобретению. Любое заболевание, описанное в настоящей заявке, можно лечить в соответствии с настоящим изобретением. В частности, настоящее изобретение относится к ретровирусному/лентивирусному (например, SIV) вектору согласно изобретению для применения в способе лечения заболевания легких. Заболевание, подлежащее лечению, может быть хроническим заболеванием. Предпочтительным является ретровирусный/лентивирусный (например, SIV) вектор для его применения при лечении CF.

Настоящее изобретение также относится к применению описанного здесь ретровирусного/лентивирусного (например, SIV) вектора в целях приготовления лекарственного средства для применения в способе лечения заболевания. Обычно, ретровирусный/лентивирусный (например, SIV) вектор получают с применением способа согласно изобретению. Любое описанное здесь заболевание можно лечить в соответствии с настоящим изобретением. В частности, настоящее изобретение относится к применению ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению в целях получения лекарственного средства для применения в способе лечения заболевания легких. Заболевание, подлежащее лечению, может быть хроническим заболеванием. Предпочтительным является применение ретровирусного/лентивирусного (например, SIV) вектора в целях приготовления лекарственного средства для применения в способе лечения CF.

Составы и введение

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению можно вводить в любой дозе, подходящей для достижения желаемого терапевтического эффекта. Соответствующие дозы могут быть определены врачом-клиницистом или другим практическим врачом с применением стандартных методов и в рамках обычного рабочего процесса. Неограничивающие примеры подходящих доз включают 1×10^8 единиц трансдукции (ЕТ), 1×10^9 ЕТ, 1×10^{10} ЕТ, 1×10^{11} ЕТ или более.

Настоящее изобретение также относится к композициям, содержащим ретровирусные/лентивирусные (например, SIV) векторы, описанные выше, и фармацевтически приемлемый носитель. Неограничивающие примеры фармацевтически приемлемых носителей включают воду, физиологический раствор и забуференный фосфатом физиологический раствор. Однако, в некоторых вариантах осуществления изобретения, композиция находится в лиофилизованной форме, и в этом случае, она может включать стабилизатор, такой как альбумин бычьей сыворотки (BSA). В некоторых вариантах осуществления изобретения может оказаться желательным получить

композицию с консервантом, таким как тиомерсал или азид натрия, для достижения более длительного хранения.

Ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению можно вводить любым подходящим способом. Может быть желательным введение композиций согласно изобретению (как описано выше) в дыхательные пути индивидуума. Эффективная доставка терапевтической/профилактической композиции или лекарственного средства к очагу инфекции в дыхательных путях может быть достигнута путем перорального или интраназального введения, например, в виде аэрозолей (например, назальных спреев) или с помощью катетеров. Обычно, ретровирусные/лентивирусные (например, SIV) векторы согласно изобретению являются стабильными в клинически релевантных аэрозольных ингаляторах, в ингаляторах (включая дозирующие ингаляторы), в катетерах и аэрозолях и т.п.

В некоторых вариантах осуществления изобретения, нос является предпочтительным участком для продукции терапевтического белка с использованием ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению по меньшей мере по одной из таких причин, как: (i) внеклеточные барьеры, такие как воспалительные клетки и мокрота, менее выражены в носу; (ii) простота введения вектора; (iii) потребность в меньших количествах вектора; и (iv) этические соображения. Таким образом, трансдукция клеток назального эпителия ретровирусным/лентивирусным (например, SIV) вектором согласно изобретению может давать эффективную (в высокой степени) и продолжительную экспрессию представляющего интерес терапевтического трансгена. Соответственно, предпочтительным может быть интраназальное введение ретровирусного/лентивирусного (например, SIV) вектора согласно изобретению.

Препараты для интраназального введения могут быть приготовлены в форме назальных капель или назального спрея. Интраназальный состав может содержать капли, имеющие приблизительный диаметр в диапазоне 100-5000 мкм, например, 500-4000 мкм, 1000-3000 мкм или 100-1000 мкм. Альтернативно, по объему, капли могут находиться в диапазоне приблизительно 0,001-100 мкл, например, 0,1-50 мкл или 1,0-25 мкл, или, например, 0,001-1 мкл.

Аэрозольный состав может иметь форму порошка, суспензии или раствора. Размер аэрозольных частиц зависит от способности аэрозоля к доставке. Более мелкие частицы, в отличие от более крупных частиц, могут перемещаться дальше по дыхательным путям к альвеолам. В одном варианте осуществления изобретения, аэрозольные частицы имеют определенное распределение по диаметру для облегчения доставки по всей длине бронхов, бронхиол и альвеол. В качестве альтернативы, распределение частиц по размерам может быть выбрано для нацеливания на конкретный участок дыхательных путей, например, на альвеолы. В случае аэрозольной доставки лекарственного средства, частицы могут иметь диаметр приблизительно в диапазоне 0,1-50 мкм, предпочтительно 1-25 мкм, а более предпочтительно 1-5 мкм.

Частицы аэрозоля могут быть предназначены для доставки с помощью

аэрозольного распылителя (например, через рот) или назального спрея. Аэрозольный состав может, но необязательно, содержать пропеллент и/или поверхностно-активное вещество.

Приготовление фармацевтических аэрозолей является обычной практикой для специалистов в данной области, см., например, Sciarra, J. в Remington's Pharmaceutical Sciences (*выше*). Агенты могут быть приготовлены в виде аэрозолей-растворов, аэрозолей-дисперсий или суспензий сухих порошков, эмульсий или полутвердых препаратов. Аэрозоль может быть доставлен с использованием любой системы пропеллента, известной специалистам в данной области. Аэрозоли можно наносить на верхние дыхательные пути, например, путем назальной ингаляции, или на нижние дыхательные пути, или на то и другое. Часть легкого, в которую доставляется лекарственное средство, может зависеть от конкретного заболевания. Композиции, содержащие вектор согласно изобретению, в частности, в случае использования интраназальной доставки, могут содержать увлажнитель. Это может способствовать уменьшению или предотвращению высыхания слизистой оболочки и предотвращению ее раздражения. Подходящие увлажнители включают, например, сорбит, минеральное масло, растительное масло и глицерин; успокаивающие средства; агенты, кондиционирующие слизистую оболочку; подсластители; и их комбинации. Композиции могут содержать поверхностно-активное вещество. Подходящие поверхностно-активные вещества включают неионные, анионные и катионные поверхностно-активные вещества. Примеры поверхностно-активных веществ, которые могут быть использованы, включают, например, полиоксиэтиленовые производные неполных сложных эфиров жирных кислот и ангидридов сорбита, такие как, например, Твин 80, стеарат полиоксила 40, стеарат полиоксиэтилена 50, фузиаты, соли желчных кислот и октоксипол.

В некоторых случаях, после первого введения может быть проведено последующее введение ретровирусного/лентивирусного (например, SIV) вектора. Введение может, например, осуществляться по меньшей мере через неделю, через две недели, через месяц, через два месяца, через шесть месяцев, через год или более после первого введения. В некоторых случаях, ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению можно вводить по меньшей мере один раз в неделю, один раз в две недели, один раз в месяц, каждые два месяца, каждые шесть месяцев, ежегодно или с более длительными интервалами. Предпочтительно, введение осуществляют каждые шесть месяцев, а более предпочтительно ежегодно. Ретровирусные/лентивирусные (например, SIV) векторы можно, например, вводить с интервалами, определяемыми уменьшением эффектов предыдущего введения.

Любые два или более ретровирусных/лентивирусных (например, SIV) векторов согласно изобретению могут быть введены отдельно, последовательно или одновременно. Таким образом, два ретровирусных/лентивирусных (например, SIV) вектора или более ретровирусных/лентивирусных (например, SIV) векторов, где по меньшей мере один ретровирусный/лентивирусный (например, SIV) вектор представляет собой

ретровирусный/лентивирусный (например, SIV) вектор согласно изобретению, могут быть введены по отдельности, одновременно или последовательно, а в частности, таким образом могут быть введены два или более ретровирусных/лентивирусных (например, SIV) векторов согласно изобретению. Два этих вектора могут быть введены в одной и той же или в разных композициях. В предпочтительном случае, два ретровирусных/лентивирусных (например, SIV) вектора могут быть доставлены в одной и той же композиции.

Гомология последовательностей

Для определения процента идентичности может быть применен любой из множества способов выравнивания последовательностей, включая, но не ограничиваясь ими, общие методы, локальные методы и гибридные методы, такие как, например, методы сегментного подхода. Протоколы для определения процента идентичности представляют собой рутинные процедуры, доступные специалисту в данной области. Общие методы позволяют выравнивать последовательности от начала до конца молекулы и определять наилучшее выравнивание путем суммирования оценок отдельных пар остатков и наложения штрафов за пробелы. Неограничивающие методы включают, например, CLUSTAL W, см., например, Julie D. Thompson et al., CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice, 22(22) *Nucleic Acids Research* 4673-4680 (1994); и методы итеративного уточнения, см., например, Osamu Gotoh, Significant Improvement in Accuracy of Multiple Protein. Sequence Alignments by Iterative Refinement as Assessed by Reference to Structural Alignments, 264(4) *J. Mol. Biol.* 823-838 (1996). Локальные методы позволяют выравнивать последовательности посредством идентификации одного или нескольких консервативных мотивов, общих для всех исходных последовательностей. Неограничивающие методы включают, например, Match-box, см., например, Eric Depiereux and Ernest Feytmans, Match-Box: A Fundamentally New Algorithm for the Simultaneous Alignment of Several Protein Sequences, 8(5) *CABIOS* 501-509 (1992); метод взятия образцов Гибса, см., например, C. E. Lawrence et al., Detecting Subtle Sequence Signals: A Gibbs Sampling Strategy for Multiple Alignment, 262(5131) *Science* 208-214 (1993); Align-M, см., например, Ivo Van Waile et al., Align-M - A New Algorithm for Multiple Alignment of Highly Divergent Sequences, 20(9) *Bioinformatics*:1428-1435 (2004).

Таким образом, процент идентичности последовательностей определяют обычными методами. См., например, Altschul et al., *Bull. Math. Bio.* 48: 603-16, 1986 and Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915-19, 1992. Вкратце, две аминокислотные последовательности выравнивают для оптимизации показателей выравнивания с использованием штрафа за пробел-пропуск, равного 10, штрафа за пробел-удлинение, равного 1, и матрицы оценки «blosum 62» Henikoff and Henikoff (там же), как показано ниже (аминокислоты обозначены стандартными однобуквенными кодами).

«Процент идентичности последовательностей» между двумя или более

последовательностями нуклеиновых кислот или аминокислот зависит от количества идентичных положений, общих для этих последовательностей. Таким образом, % идентичности можно вычислить как количество идентичных нуклеотидов/аминокислот, деленное на общее количество нуклеотидов/аминокислот, с умножением результата на 100. При вычислении % идентичности последовательностей также можно учитывать количество пробелов и длину каждого пробела, которые необходимо ввести для оптимизации выравнивания двух или более последовательностей. Сравнение последовательностей и определение процента идентичности между двумя или более последовательностями можно проводить с использованием конкретных математических алгоритмов, таких как BLAST, которые известны специалисту в данной области.

Показатели выравнивания для определения идентичности последовательностей

A R N D C Q E G H I L K M F P S T W Y V

A 4

R -1 5

N -2 0 6

D -2 -2 1 6

C 0 -3 -3 -3 9

Q -1 1 0 0 -3 5

E -1 0 0 2 -4 2 5

G 0 -2 0 -1 -3 -2 -2 6

H -2 0 1 -1 -3 0 0 -2 8

I -1 -3 -3 -3 -1 -3 -3 -4 -3 4

L -1 -2 -3 -4 -1 -2 -3 -4 -3 2 4

K -1 2 0 -1 -3 1 1 -2 -1 -3 -2 5

M -1 -1 -2 -3 -1 0 -2 -3 -2 1 2 -1 5

F -2 -3 -3 -3 -2 -3 -3 -3 -1 0 0 -3 0 6

P -1 -2 -2 -1 -3 -1 -1 -2 -2 -3 -3 -1 -2 -4 7

S 1 -1 1 0 -1 0 0 0 -1 -2 -2 0 -1 -2 -1 4

T 0 -1 0 -1 -1 -1 -1 -2 -2 -1 -1 -1 -1 -2 -1 1 5

W -3 -3 -4 -4 -2 -2 -3 -2 -2 -3 -2 -3 -1 1 -4 -3 -2 11

Y -2 -2 -2 -3 -2 -1 -2 -3 2 -1 -1 -2 -1 3 -3 -2 -2 2 7

V 0 -3 -3 -3 -1 -2 -2 -3 -3 3 1 -2 1 -1 -2 -2 0 -3 -1 4

Затем вычисляют процент идентичности по формуле:

$$\frac{\text{Общее количество совпадений}}{\text{длина более длинной последовательности плюс}} \times 100$$

[длина более длинной последовательности плюс

количество пробелов, введенных в более длинную последовательность для выравнивания двух последовательностей]

По существу, гомологичные полипептиды характеризуются наличием одной или более аминокислотных замен, делеций или добавлений. Эти изменения предпочтительно носят незначительный характер, то есть, они представляют собой консервативные аминокислотные замены (как описано в настоящей заявке) и другие замены, которые существенно не влияют на укладку или активность полипептида; небольшие делеции, обычно от одной до приблизительно 30 аминокислот; и небольшие удлинения на аминокислотных или карбоксильных концах, такие как аминокислотный остаток метионина, небольшой линкерный пептид, состоящий приблизительно из 20-25 остатков, или аффинную метку.

В дополнение к 20 стандартным аминокислотам, аминокислотные остатки полипептидов согласно изобретению могут быть заменены нестандартными аминокислотами (такими как 4-гидроксипролин, 6-N-метиллизин, 2-аминоизомасляная кислота, изовалин и α -метилсерин). Аминокислотные остатки полипептида могут быть заменены ограниченным числом неконсервативных аминокислот; аминокислот, не кодируемых генетическим кодом; и неприродных аминокислот. Полипептиды согласно изобретению могут также содержать не встречающиеся в природе аминокислотные остатки.

Неприродными аминокислотами являются, но не ограничиваются ими, транс-3-метилпролин, 2,4-метанопронин, цис-4-гидроксипролин, транс-4-гидроксипролин, N-метилглицин, аллотреонин, метилтреонин, гидроксиэтилцистеин, гидроксиэтилгомоцистеин, нитроглутамин, гомоглутамин, пипеколиновая кислота, третлейцин, норвалин, 2-азафенилаланин, 3-азафенилаланин, 4-азафенилаланин и 4-фторфенилаланин. Специалистам в данной области известно несколько способов включения аминокислотных остатков неприродного происхождения в белки. Так, например, можно использовать систему *in vitro*, в которой нонсенс-мутации подавляются с использованием химически аминоацелированных супрессорных тРНК. Способы синтеза аминокислот и аминоацелирования тРНК известны специалистам в данной области. Транскрипцию и трансляцию плазмид, содержащих нонсенс-мутации, проводят в бесклеточной системе, включающей экстракт *E. coli* S30 и коммерчески доступные ферменты и другие реагенты. Белки очищают посредством хроматографии. См., например, Robertson et al., *J. Am. Chem. Soc.* 113:2722, 1991; Ellman et al., *Methods Enzymol.* 202:301, 1991; Chung et al., *Science* 259:806-9, 1993; и Chung et al., *Proc. Natl. Acad. Sci. USA* 90:10145-9, 1993). Во втором способе, трансляцию осуществляют в ооцитах *Xenopus* путем микроинъекции мутантной мРНК и химически аминоацелированных супрессорных тРНК (Turcatti et al., *J. Biol. Chem.* 271:19991-8, 1996). В третьем способе, клетки *E. coli* культивируют в отсутствие природной аминокислоты, подлежащей замене (например, фенилаланина), и в присутствии желаемой(ых) не встречающейся(ихся) в природе аминокислоты(аминокислот) (например, 2-

азафенилаланина, 3-азафенилаланина, 4-азафенилаланина или 4-фторфенилаланина). Не встречающуюся в природе аминокислоту включают в полипептид вместо ее природного аналога. См. Koide et al., *Biochem.* 33:7470-6, 1994. Встречающиеся в природе аминокислотные остатки могут быть превращены в не встречающиеся в природе виды посредством химической модификации *in vitro*. Химическую модификацию можно комбинировать с сайт-направленным мутагенезом для дальнейшего расширения диапазона замен (Wynn and Richards, *Protein Sci.* 2:395-403, 1993).

Аминокислотные остатки полипептидов согласно изобретению могут быть заменены ограниченным числом неконсервативных аминокислот; аминокислот, которые не кодируются генетическим кодом; не встречающихся в природе аминокислот и неприродных аминокислот.

Незаменимые аминокислоты в полипептидах согласно изобретению могут быть идентифицированы в соответствии с методами, известными специалистам в данной области, такими как сайт-направленный мутагенез или аланин-сканирующий мутагенез (Cunningham and Wells, *Science* 244: 1081-5, 1989). Сайты биологического взаимодействия также могут быть определены с помощью физического анализа структуры, определяемой такими методами, как ядерный магнитный резонанс, кристаллография, электронная дифракция или фотоаффинное мечение, в сочетании с мутацией аминокислот предполагаемого сайта контактирования. См., например, de Vos et al., *Science* 255:306-12, 1992; Smith et al., *J. Mol. Biol.* 224:899-904, 1992; Wlodaver et al., *FEBS Lett.* 309:59-64, 1992. Идентичность незаменимых аминокислот также может быть установлена на основании анализа гомологии с родственными компонентами (например, транслокационными или протеазными компонентами) полипептидов согласно изобретению.

Множественные аминокислотные замены могут быть сделаны и протестированы с применением известных методов мутагенеза и скрининга, таких как методы, описанные Reidhaar-Olson and Sauer (*Science* 241:53-7, 1988) или Bowie and Sauer (*Proc. Natl. Acad. Sci. USA* 86:2152-6, 1989). Вкратце, эти авторы раскрывают способы одновременной рандомизации двух или более положений в полипептиде, отбора функционального полипептида и последующего секвенирования мутагенизированных полипептидов для определения спектра допустимых замен в каждом положении. Другие способы, которые могут быть применены, включают фаговое представление (см., например, Lowman et al., *Biochem.* 30:10832-7, 1991; Ladner et al., патент США № 5223409; Huse, публикация ВОИС WO 92/06204) и мутагенез, направленный на нужную область (Derbyshire et al., *Gene* 46:145, 1986; Ner et al., *DNA* 7:127, 1988).

Множественные аминокислотные замены могут быть сделаны и протестированы с применением известных методов мутагенеза и скрининга, таких как методы, описанные Reidhaar-Olson and Sauer (*Science* 241:53-7, 1988) или Bowie and Sauer (*Proc. Natl. Acad. Sci. USA* 86:2152-6, 1989). Вкратце, эти авторы раскрывают способы одновременной рандомизации двух или более положений в полипептиде, отбора функционального

полипептида и последующего секвенирования мутагенизированных полипептидов для определения спектра допустимых замен в каждом положении. Другие способы, которые могут быть применены, включают фаговое представление (см., например, Lowman et al., Biochem. 30:10832-7, 1991; Ladner et al., патент США № 5223409; Huse, публикация ВОИС WO 92/06204) и мутагенез, направленный на нужную область (Derbyshire et al., Gene 46:145, 1986; Ner et al., DNA 7:127, 1988).

Примеры

Настоящее изобретение далее описано со ссылкой на нижеприведенные примеры. Эти примеры не ограничивают объем изобретения, и специалисту в данной области будет очевидно, что в объем настоящего изобретения могут входить подходящие эквиваленты. Таким образом, эти примеры можно рассматривать как составные части изобретения, а отдельные описанные в них аспекты можно рассматривать как раскрытые независимо или в любой комбинации.

Пример 1. Конструкция плазмиды pGM691.

Было проведено сравнение векторной геномной плазмиды (pDNA1) pGM326 с плазмидой GagPol (pDNA2a) pGM297. Как показано на фигуре 5A, существует значительная гомология между частичной последовательностью нуклеотидов gagpol в pGM326 и не оптимизированной по кодомам последовательностью gagpol в pGM297.

Модифицированная плазида pDNA2a была сконструирована для (i) уменьшения гомологии между частичной нуклеотидной последовательностью gagpol в pGM326 и не оптимизированной по кодомам последовательностью gagpol в pGM297; (ii) для оптимизации по кодомам генов gagpol в целях повышения экспрессии белка gagpol; (iii) для снижения теоретического риска образования компетентного по репликации лентивируса (RCL) во время продукции или клинического применения; и (iv) для устранения зависимости экспрессии gagpol от Rev. Сравнение pGM297 с модифицированной pDNA2a (pGM691) показано на фигурах 5B-5D с комментариями по изменениям.

pGM691 получали путем расщепления pGM297 рестриктирующими ферментами XhoI, EcoRV и BglII с получением фрагментов ДНК размером 4583 п.о., 3662 п.о. и 1641 п.о. Фрагмент 4583 п.о., содержащий плазмидный ориджин репликации и интрон промотора СВА, очищали и сохраняли. Плазида pGM693 была получена специалистами GeneArt/LifeTechnologies путем синтеза ДНК. pGM693 была сконструирована авторами настоящего изобретения так, чтобы она включала фрагмент ДНК от XhoI до BglII размером 4481 п.о., который включал оптимизированную по кодомам последовательность GagPol, в конечном счете обнаруженную в pGM691. pGM693 расщепляли ферментами XhoI и BglII с получением фрагментов ДНК размером 4481 п.о., 1236 п.о. и 1048 п.о. Фрагмент 4481 п.о., содержащий оптимизированную по кодомам последовательность GagPol, очищали и сохраняли (см. фигуру 5E). Два оставшихся фрагмента ДНК лигировали с помощью ДНК-лигазы, и полученную смесь лигированной ДНК переносили в клетки E. coli Stbl3, а затем клетки, содержащие плазмиды, способные к репликации,

отбирали по их резистентности к канамицину. Затем отбирали и размножали четко выделенные отдельные колонии резистентных к канамицину трансформированных клеток Stb13. Рестрикционный анализ ДНК полученных клонов выявил ряд клонов с ожидаемой структурой ДНК, при этом, один из них был зарезервирован и обозначен pGM691.

Пример 2. Получение вектора rSIV.F/HN hCEF-CFTR

Геном вектора pGM326, который включает трансген CFTR под транскрипционным контролем промотора hCEF, использовали в двух экспериментальных исследованиях (DoE) для оценки выходов продукта, обеспечиваемых с использованием либо pGM297 GagPol, либо pGM691 соGagPol.

В каждом исследовании DoE использовали широкий диапазон условий, включая низкие, умеренные и высокие концентрации каждого из используемых компонентов:

Функция	Код	Низкая	Умеренная	Высокая
Геном	pGM326	0,2	1,1	2
(со)GagPol	pGM297 или GM691	0,1	0,55	1
Rev	pGM299	0,1	0,55	1
F	pGM301	0,1	0,55	1
HN	pGM303	0,1	0,55	1
Реагент для трансфекции	Липофектамин 2000	4	7	10

Единицы измерения реагента для трансфекции представлены в мкл/мл, а для всех остальных реагентов - мкг/мл.

Был проведен 3-уровневый дробный факториальный анализ с дубликатами векторов, приготовленными для большинства условий, и с шестью дубликатами центральных точек. Всего был приготовлен 31 исходный вектор с использованием идентичных условий для pGM297 GagPol и pGM691 соGagPol.

Титр интегрирующей единицы трансдукции (ЕТ/мл), определенный путем вычисления соотношения последовательностей ДНК, специфичных для вектора и для генома, в трансдуцированных клетках с помощью количественной ПЦР после трансдукции клеток 293Т разведениями исходных векторов, представлен на графике на фигуре 6А (повторяющиеся штаммы векторов представлены точками, а линия указывает на все остальные идентичные условия).

После проведения экспериментов DOE, векторный геном pGM326, который включает трансген CFTR под транскрипционным контролем промотора hCEF, использовали для получения исходных векторов rSIV.F/HN с тремя повторностями с использованием либо pGM297 GagPol, либо pGM691 соGagPol, как уже было указано.

Для всех препаратов, Rev, F и HN были представлены плазмидами pGM299, pGM301 и pGM303, соответственно. Соотношение масс ДНК-генома вектора:GagPol:Rev:F:HN во всех случаях составляло 20:9:6:6:6. Для условий А и В, используемые уровни общей ДНК составляли 2,2 мкг/мл и 1,8 мкг/мл, соответственно.

Для условий А и В, общие уровни используемого Липофектамина 2000 составляли 7 мкл/мл и 8 мкл/мл, соответственно.

На графике представлен титр интегрирующей единицы трансдукции (ЕТ/мл), определенный путем вычисления соотношения последовательностей ДНК, специфичных для вектора и для генома, в трансдуцированных клетках с помощью количественной ПЦР после трансдукции клеток 293Т разведениями исходных векторов (отдельные штаммы векторов представлены точками, а линия указывает на медиану для группы).

Выходы векторов с соGagPol, представленные для рGM691, были приблизительно в 2,3 раза выше в условиях А и приблизительно в 1,5 раза выше в условиях В (фигура 6В). Таким образом, использование рGM691 в качестве рDNA2а заметно увеличивало вирусный титр SIV, независимо от других используемых условий культивирования. Это удивительно, поскольку существует множество независимых опубликованных исследований, в которых сообщается, что оптимизация по кодонам генов gagpol ассоциируется со снижением титра лентивирусов.

Пример 3. Получение rSIV.F/HN CMV-EGFP

Для того, чтобы выяснить, ограничивается ли способность оптимизированного по кодонам gagpol поддерживать или повышать титр вектора конкретной конструкцией rSIV.F/HN (rSIV.F/HN hCEF-CFTR), были проведены эксперименты с использованием плазмид для получения другого трансгена, функционально связанного с другим промотором.

Клетки НЕК293Т, Freestyle 293F (Life Technologies, Paisley, UK) и клетки 293Т/17 (CRL-11268; ATCC, Manassas, VA) хранили в модифицированной по способу Дульбекко минимальной среде Игла (Invitrogen, Carlsbad, CA), содержащей 10% фетальную бычью сыворотку и дополненной пенициллином (100 ед./мл) и стрептомицином (100 мкг/мл) или экспрессионной средой Freestyle™ 293 (Life Technologies).

SeV-F/HN-псевдотипированный вектор SIV был получен путем трансфекции клеток НЕК293Т или 293Т/17, культивированных в экспрессионной среде FreeStyle™ 293, смесью из пяти плазмид со следующими свойствами: рDNA1 (рGM311, которая включает трансген EGFP под транскрипционным контролем промотора CMV) кодирует мРНК лентивирусного вектора; рDNA2а (рGM691; фигура 2С) кодирует белки Gag и Pol SIV; рDNA2b (рGM299; фигура 2D) кодирует белки SIV Rev; рDNA3а (рGM301; фигура 2E) кодирует белок Fct4 вируса Сендай [Kobayashi et al., 2003 J. Virol. 77:2607]; и рDNA3b (рGM303; фигура 2F) кодирует SIVct+HN вируса Сендай [Kobayashi et al., 2003 J. Virol. 77:2607] в комплексе с PEIpro (Polyplus, Illkirch, France). Среду для культивирования клеток добавляли через 12-24 часа после трансфекции с использованием бутирата натрия. Бутират натрия стимулирует продуцирование векторов посредством ингибирования гистон-деацетилазы, что приводит к увеличению уровня экспрессии белковых компонентов слитого белка SIV и вируса Сендай, кодируемых пятью плазмидами. Среду для культивирования клеток добавляли через 44-52 часа и/или через 68-76 часов после трансфекции с использованием 5 единиц/мл бензоназной нуклеазы (Merck Millipore,

Nottingham, UK). Культуральный супернатант, содержащий вектор SIV, собирали через 68-76,5 часов после трансфекции и очищали путем фильтрации через мембрану размером 0,45 мкм. Вектор SIV обрабатывали путем расщепления с помощью TrypLE Select™. Затем вектор SIV дополнительно очищали и концентрировали путем анионообменной хроматографии и фильтрации в тангенциальном потоке.

Векторные штаммы rSIV.F/HN с тремя повторностями получали с использованием либо pGM297 GagPol, либо pGM691 соGagPol, как уже было указано. Массовое соотношение ДНК генома вектора:GagPol:Rev:F:HN во всех случаях составляло 20:9:6:6:6.

Титр функциональной единицы трансдукции (ФЕТ/мл), определенный путем детектирования EGFP-позитивных клеток с помощью проточной цитометрии после трансдукции клеток 293Т разведениями исходных векторов, был нанесен на график, показанный на фигуре 7 (отдельные векторные штаммы представлены точками, а линия указывает на медиану для группы). Что касается конструкций rSIV.F/HN hCEF-CFTR в Примере 2, то выход вектора rSIV.F/HN CMV-EGFP с соGagPol, представленный для pGM691, был приблизительно в 1,6 раза выше, чем при использовании не оптимизированного по кодонам gagpol pGM297. Это свидетельствует о том, что способность оптимизированного по кодонам gagpol поддерживать или повышать титр вектора не ограничивалась конкретной конструкцией rSIV.F/HN hCEF-CFTR, а скорее является функцией, обычно ассоциированной с соGagPol.

Пример 3. Уменьшение числа интактных OPC SIV в векторной геномной плазмиде

Дополнительные модификации одной или более сконструированных плазмид могут дополнительно повысить безопасность конечного векторного продукта, обеспечивая тем самым дополнительное клиническое преимущество.

Авторами настоящего изобретения были рассмотрены последовательности сконструированных плазмид и идентифицированы несколько представляющих интерес областей в векторной геномной плазмиде pGM326. В частности, частичная область Gag RRE cPPT hCEF pGM326 содержит:

- 77 старт-кодонов (ATG);
- 32 OPC длиной ≥ 10 аминокислот
- 2 крупных OPC в направлении от 5' до 3'
- 189 аминокислот из большинства 5'-ATG в геноме вектора (слитый белок Gag/RRE), кодирующего матрицу p17 и часть капсида p24;
- 250 аминокислот от внутреннего ATG до RRE (слитый белок RRE/cPPT/hCEF)

Эти конструкции проиллюстрированы на фигуре 8. Особое внимание привлекли 2 крупные OPC (показаны на фигуре 9).

Таким образом, акторами настоящего изобретения был разработан модифицированный вариант плазмиды pGM326 с комбинацией дополнительных модификаций, предназначенных для уменьшения количества интактных OPC SIV (а в частности, для удаления этих 2 крупных OPC) для повышения безопасности.

Модификации были внесены в 2 крупные ОРС, расположенные выше промотора hCEF и трансгена CFTR (soCFTR2). Внесенные изменения заключались в следующем:

- Было удалено 6 ATG (3хATG-ATTG, 1хATG-TTG, 2хATG-AAG)
- Был встроены 1 стоп-кодон (TCC-TAAA)
- Был изменен 1 рестрикционный сайт между частичным Gag и RRE (EcoRI GAATTC-GCCTGCAGG SbfI)

Полученная плазмида векторного генома представляет собой pGM830, как показано на Фигуре 2В, с последовательностью SEQ ID NO: 4.

Сравнение титра вектора с использованием векторных геномных плазмид pGM326 или pGM830 в соответствии с протоколом, проводимом в основном идентичных условиях показало, что использование pGM830 дает титр, сравнимый с титром pGM326, с использованием клеток НЕК293Т и А549 (см. Фигуру 10), что указывает на то, что улучшенный профиль безопасности может быть достигнут без негативного влияния на титр.

Пример 4. Комбинация соGagPol и плазмиды с модифицированным геномом вектора поддерживает или даже увеличивает титр вектора

Эксперименты, описанные в Примере 2, неожиданно продемонстрировали, что вместо ожидаемого снижения выхода, образование SIV.F/HN hCEF-CFTR с использованием соGagPol имело тенденцию к поддержанию или даже к повышению титра вектора. Эксперименты, описанные в Примере 3, продемонстрировали, что дальнейшее улучшение профиля безопасности вектора может быть достигнуто путем модификации геномной плазмиды вектора без негативного воздействия на титр вектора.

После этого были проведены дополнительные эксперименты, в которых использование соGagPol сочеталось с использованием векторной геномной плазмиды pGM830 для того, чтобы исследовать, можно ли комбинировать эти две модификации, ассоциированные с безопасностью, и поддерживать титр вектора.

Как показано на фигуре 11, авторы настоящего изобретения неожиданно обнаружили, что использование соGagPol не только может сочетаться с использованием плазмиды с модифицированным векторным геномом (pGM830), но и то, что эта комбинация дает наблюдаемую тенденцию к увеличению титра вектора.

Это говорит о том, что можно не только получить векторы с дополнительными улучшенными профилями безопасности путем комбинированного использования соGagPol с модифицированной плазмидой генома вектора, но также, что неожиданно, этого можно достичь при сохранении или даже увеличении титра rSIV.F/HN hCEF-трансгена.

Информация о последовательностях

Ключ к последовательностям

SEQ ID NO: 1: Оптимизированная по кодонам последовательность нуклеиновой кислоты SIV gal-pol

SEQ ID NO: 2: Последовательность нуклеиновой кислоты gag-pol SIV дикого типа

- SEQ ID NO: 3: Плазмида, определенная на Фигуре 2A (pDNA1 pGM326)
- SEQ ID NO: 4: Плазмида, определенная на Фигуре 2B (pDNA1 pGM830)
- SEQ ID NO: 5: Плазмида, определенная на Фигуре 2C (pDNA2a pGM691)
- SEQ ID NO: 6: Плазмида, определенная на Фигуре 2D (pDNA2b pGM299)
- SEQ ID NO: 7: Плазмида, определенная на Фигуре 2E (pDNA3a pGM301)
- SEQ ID NO: 8: Плазмида, определенная на Фигуре 2F (pDNA3b pGM303)
- SEQ ID NO: 9: Плазмида, определенная на Фигуре 2G (pDNA2a pGM297)
- SEQ ID NO: 10: Репрезентативный промотор hCEF
- SEQ ID NO: 11: Репрезентативный промотор CMV
- SEQ ID NO: 12: Репрезентативный промотор EF1a
- SEQ ID NO: 13: Репрезентативный трансген CFTR (soCFTR2)
- SEQ ID NO: 14: Репрезентативный трансген A1AT
- SEQ ID NO: 15: Цепь, комплементарная репрезентативному трансгену A1AT
- SEQ ID NO: 16: Репрезентативный полипептид A1A1
- SEQ ID NO: 17: Репрезентативный трансген FVIII (N6)
- SEQ ID NO: 18: Репрезентативный трансген FVIII (V3)
- SEQ ID NO: 19: Цепь, комплементарная репрезентативному трансгену FVIII (N6).
- SEQ ID NO: 20: Цепь, комплементарная репрезентативному трансгену FVIII (V3)
- SEQ ID NO: 21: Репрезентативный полипептид FVIII (N6)
- SEQ ID NO: 22: Репрезентативный полипептид FVIII (V3)
- SEQ ID NO: 23: Репрезентативный компонент WPRE (mWPRE)
- SEQ ID NO: 24: Плазмида F/HN-SIV-hCEF-soA1AT, определенная на фигуре 3 (pDNA1 pGM407).
- SEQ ID NO: 25: Плазмида F/HN-SIV-CMV-HFVIII-V3, определенная на фигуре 4A (pDNA1 pGM411)
- SEQ ID NO: 26: Плазмида F/HN-SIV-hCEF-HFVIII-V3, определенная на фигуре 4B (pDNA1 pGM413).
- SEQ ID NO: 27: Плазмида F/HN-SIV-CMV-HFVIII-N6-so, определенная на фигуре 4C (pDNA1 pGM412)
- SEQ ID NO: 28: Плазмида F/HN-SIV-hCEF-HFVIII-N6-so, определенная на фигуре 4D (pDNA1 pGM414).
- SEQ ID NO: 29: Репрезентативный промотор CAG

Последовательности

SEQ ID NO: 1: Оптимизированная по кодам последовательность нуклеиновой кислоты SIV gal-pol (из pGM691)

Длина: 4391; Тип молекулы: ДНК; Особенности локализации/Определители: источник, 1..4391; тип молекулы, другая ДНК; примечание: оптимизированная по кодам последовательность нуклеиновой кислоты SIV gal-pol (из pGM691); организм, синтетическая конструкция

ATGGGAGCTGCCACATCTGCCCTGAATAGACGGCAGCTGGACCAGTTCGAGA

AGATCAGACTGCGGCCCAACGGCAAGAAGAAGTACCAGATCAAGCACCTGATCTGG
GCCGGCAAAGAGATGGAAAGATTCGGCCTGCACGAGCGGCTGCTGGAAACCGAGG
AAGGCTGCAAGAGAATTATCGAGGTGCTGTACCCTCTGGAACCTACCGGCTCTGAG
GGCCTGAAGTCCCTGTTCAATCTCGTGTGCGTGCTGTACTGCCTGCACAAAGAACAG
AAAGTGAAGGACACCGAAGAGGGCCGTGGCCACAGTTAGACAGCACTGCCACCTGGT
GGAAAAAGAGAAGTCCGCCACAGAGACAAGCAGCGGCCAGAAGAAGAACGACAAG
GGAATTGCTGCCCTCCTGGCGGCAGCCAGAATTTTCTGCTCAGCAGCAGGGAAAC
GCCTGGGTGCACGTTCCACTGAGCCCTAGAACACTGAATGCCTGGGTCAAAGCCGT
GGAAGAGAAGAAGTTTGGCGCCGAGATCGTGCCCATGTTCCAGGCTCTGTCTGAGG
GCTGCACCCCTTACGACATCAACCAGATGCTGAACGTGCTGGGAGATCACCAGGGC
GCTCTGCAGATCGTGAAAGAGATCATCAACGAAGAGGCTGCCCAGTGGGACGTGAC
ACATCCATTGCCTGCTGGACCTCTGCCAGCCGGACAACACTGAGAGATCCTAGAGGCTC
TGATATCGCCGGCACCACCAGCTCTGTGCAAGAGCAGCTGGAATGGATCTACACCG
CCAATCCTAGAGTGGACGTGGGCGCCATCTACAGAAGATGGATCATCCTGGGCCTG
CAGAAATGCGTGAAGATGTACAACCCCGTGTCCGTGCTGGACATCAGACAGGGACC
CAAAGAGCCCTTCAAGGACTACGTGGACCGTTCTATAAGGCCATTAGAGCCGAGC
AGGCCAGCGGCGAAGTGAAGCAGTGGATGACAGAGAGCCTGCTGATCCAGAACGC
CAATCCAGACTGCAAAGTGATCCTGAAAGGCCTGGGCATGCACCCACACTGGAAG
AGATGCTGACAGCCTGTCAAGGCGTTGGCGGCCCTTCTTACAAAGCCAAAGTGATG
GCCGAGATGATGCAGACCATGCAGAACCAGAACATGGTGCAGCAAGGCGGCCCTAA
GAGACAGAGGCCTCCTCTGAGATGCTACAACACTGCGGCAAGTTCGGCCACATGCAGA
GACAGTGTCTGAGCCTAGGAAAACAAAATGTCTAAAGTGTGGAAAATTGGGACAC
CTAGCAAAGACTGCAGGGGACAGGTGAATTTTTTTAGGGTATGGACGGTGGATGGG
GGCAAACCGAGAAATTTTCCCGCCGCTACTCTTGGAGCGGAACCGAGTGCGCCTC
CTCCACCGAGCGGCACCACCCCATACGACCCAGCAAAGAAGCTCCTGCAGCAATAT
GCAGAGAAAGGGAAACAACACTGAGGGAGCAAAGAGGAATCCACCGGCAATGAATC
CGGATTGGACCGAGGGATATTCTTTGAACTCCCTCTTTGGAGAAGACCAATAAAGAC
CGTGTACATCGAGGGCGTGCCCATCAAGGCTCTGCTGGATACAGGCGCCGACGACA
CCATCATCAAAGAGAACGACCTGCAGCTGAGCGGCCCTTGGAGGCCTAAGATCATT
GGAGGAATCGGCGGAGGCCTGAACGTCAAAGAGTACAACGACCGGGAAGTGAAGA
TCGAGGACAAGATCCTGAGGGGCACAATCCTGCTGGGCGCCACACCTATCAACATC
ATCGGCAGAAATCTGCTGGCCCCTGCCGGCGCTAGACTGGTTATGGGACAGCTCTCT
GAGAAGATCCCCGTGACACCCGTGAAGCTGAAAGAAGGCGCTAGAGGACCTTGTGT
GCGACAGTGGCCTCTGAGCAAAGAGAAGATTGAGGCCCTGCAAGAAATCTGTAGCC
AGCTGGAACAAGAGGGCAAGATCAGCAGAGTTGGCGGCGAGAACGCCTACAATAC
CCCTATCTTCTGCATCAAGAAAAAGGACAAGAGCCAGTGGCGGATGCTGGTGGACT
TTAGAGAGCTGAACAAGGCTACCCAGGACTTCTTCGAGGTGCAGCTGGGAATTCCCTC
ATCCTGCCGGCCTGCGGAAGATGAGACAGATCACAGTGCTGGATGTGGGCGACGCC
TACTACAGCATCCCTCTGGACCCCAACTTCAGAAAGTACACCGCCTTCACAATCCCC
ACCGTGAACAATCAAGGCCCTGGCATCAGATACCAGTTCAACTGCCTGCCTCAAGG

CTGGAAGGGCAGCCCCACCATTTTTTCAGAATACCGCCGCCAGCATCCTGGAAGAAA
 TCAAGAGAAACCTGCCTGCTCTGACCATCGTGCAGTACATGGACGATCTGTGGGTCTG
 GAAGCCAAGAGAATGAGCACACCCACGACAAGCTGGTGGAACAGCTGAGAACAAA
 GCTGCAGGCCTGGGGCCTCGAAACCCCTGAGAAGAAGGTGCAGAAAGAACCTCCTT
 ACGAGTGGATGGGCTACAAGCTGTGGCCTCACAAGTGGGAGCTGAGCCGGATTCTAG
 CTCGAAGAGAAGGACGAGTGGACCGTGAACGACATCCAGAAACTCGTGGGCAAGCT
 GAATTGGGCAGCCCAGCTGTATCCCGGCCTGAGGACCAAGAACATCTGCAAGCTGA
 TCCGGGGAAAGAAGAACCTGCTGGAACCTGGTCACATGGACACCTGAGGCCGAGGCC
 GAATATGCCGAGAATGCCGAAATCCTGAAAACCGAGCAAGAGGGGACCTACTACAA
 GCCTGGCATTCCAATCAGAGCTGCCGTGCAGAAACTGGAAGGCGGCCAGTGGTCCT
 ACCAGTTTAAGCAAGAAGGCCAGGTCTGAAAGTGGGCAAGTACACCAAGCAGAA
 GAACACCCACACCAACGAGCTGAGGACACTGGCTGGCCTGGTCCAGAAAATCTGCA
 AAGAGGCCCTGGTCATTTGGGGCATCCTGCCTGTTCTGGAACCTGCCATTGAGCGGG
 AAGTGTGGGAACAGTGGTGGGCCGATTACTGGCAAGTGTCTTGGATCCCCGAGTGG
 GACTTCGTGTCTACCCCTCCTCTGCTGAAACTGTGGTACACCCTGACAAAAGAGCCC
 ATTCCTAAAGAGGACGTCTACTACGTTGACGGCGCCTGCAACCGGAACTCCAAAGA
 AGGCAAGGCCGGCTACATCAGCCAGTACGGCAAGCAGAGAGTGGAAACCCTGGAA
 AACACCACCAACCAGCAGGCCGAGCTGACCGCCATTAAGATGGCCCTGGAAGATAG
 CGGCCCAATGTGAACATCGTGACCGACTCTCAGTACGCCATGGGAATCCTGACAG
 CCCAGCCTACACAGAGCGATAGCCCTCTGGTTGAGCAGATCATTGCCCTGATGATTC
 AGAAGCAGCAAATCTACCTGCAGTGGGTGCCCGCTCACAAAGGCATCGGCGGAAAC
 GAAGAGATCGATAAGCTGGTGTCCAAGGGAATCAGACGGGTGCTGTTCTGGA
 GATTGAAGAGGCCCAAGAGGAACACGAGCGCTACCACAACAACCTGGAAGAATCTG
 GCCGACACCTACGGACTGCCCCAGATCGTGGCCAAAGAAATCGTGGCTATGTGCC
 CAAGTGTGAGATCAAGGGCGAACCTGTGCACGGCCAAGTGGATGCTTCTCCTGGCA
 CATGGCAGATGGACTGTACCCACCTGGAAGGCAAAGTGGTCATCGTGGCTGTGCAC
 GTGGCCTCCGGCTTTATTGAGGCCGAAGTATCCCCAGAGAGACAGGCAAAGAAAC
 CGCCAAGTTCCTGCTGAAGATCCTGTCCAGATGGCCCATCACACAGCTGCACACCGA
 CAACGGCCCTAACTTCACATCTCAAGAGGTGGCCGCCATCTGTTGGTGGGGAAAGA
 TTGAGCACACAACCGGCATTCCCTACAATCCACAGAGCCAGGGCAGCATCGAGTCC
 ATGAACAAGCAGCTCAAAGAGATTATCGGCAAGATCCGGGACGACTGCCAGTACAC
 AGAAACAGCCGTGCTGATGGCCTGTACATCCACAACCTCAAGCGGAAAGGCGGCA
 TCGGAGGACAGACATCTGCCGAGAGACTGATCAATATCATCACCCTCAGCTGGAA
 ATCCAGCACCTCCAGACCAAGATCCAGAAGATTCTGAACTTCCGGGTGTA
 CTACCGC
 GAGGGCAGAGATCCTGTTTGGAAAGGCCCAGCACAGCTGATCTGGAAGGCGAAGG
 TGCCGTGGTGTGCTGAAGGATGGCTCTGATCTGAAGGTGGTGGCCAGACGGAAGGCCA
 AGATTATCAAGGATTACGAGCCCAAACAGCGCGTGGGCAATGAAGGCGACGTTGAG
 GGCACAAGAGGCAGCGACAATTGA

SEQ ID NO: 2: Последовательность нуклеиновой кислоты gag-pol SIV дикого типа (из pGM297)

Длина: 4391; Тип молекулы: ДНК; Особенности локализации/Определители: источник, 1..4391; тип молекулы, неустановленная ДНК; организм, вирус иммунодефицита обезьян

ATGGGGGCGGCTACCTCAGCACTAAATAGGAGACAATTAGACCAATTTGAGA
AAATACGACTTCGCCCGAACGGAAAGAAAAAGTACCAAATTAACATTTAATATGG
GCAGGCAAGGAGATGGAGCGCTTCGGCCTCCATGAGAGGTTGTTGGAGACAGAGGA
GGGGTGTAAAAGAATCATAGAAGTCCTCTACCCCCTAGAACCAACAGGATCGGAGG
GCTTAAAAAGTCTGTTCAATCTTGTGTGCGTACTATATTGCTTGCACAAGGAACAGA
AAGTGAAAGACACAGAGGAAGCAGTAGCAACAGTAAGACAACACTGCCATCTAGT
GGAAAAAGAAAAAGTGCAACAGAGACATCTAGTGGACAAAAGAAAAATGACAAG
GGAATAGCAGCGCCACCTGGTGGCAGTCAGAATTTTCCAGCGCAACAACAAGGAAA
TGCCTGGGTACATGTACCCTTGTACCGCGCACCTTAAATGCGTGGGTAAAAGCAGT
AGAGGAGAAAAAATTTGGAGCAGAAATAGTACCCATGTTTCAAGCCCTATCAGAAG
GCTGCACACCCTATGACATTAATCAGATGCTTAATGTGCTAGGAGATCATCAAGGGG
CATTACAAATAGTGAAAGAGATCATTAATGAAGAAGCAGCCCAGTGGGATGTAACA
CACCCACTACCCGCAGGACCCCTACCAGCAGGACAGCTCAGGGACCCTCGCGGCTC
AGATATAGCAGGGACCACCAGCTCAGTACAAGAACAGTTAGAATGGATCTATACTG
CTAACCCCCGGGTAGATGTAGGTGCCATCTACCGGAGATGGATTATTCTAGGACTTC
AAAAGTGTGTCAAATGTACAACCCAGTATCAGTCCTAGACATTAGGCAGGGACCT
AAAGAGCCCTTCAAGGATTATGTGGACAGATTTTACAAGGCAATTAGAGCAGAACA
AGCCTCAGGGGAAGTGAAACAATGGATGACAGAATCATTACTATTCAAATGCTA
ATCCAGATTGTAAGGTCATCCTGAAGGGCCTAGGAATGCACCCACCCTTGAAGAA
ATGTTAACGGCTTGTACGGGGGTAGGAGGCCCAAGCTACAAAGCAAAAGTAATGGC
AGAAATGATGCAGACCATGCAAAATCAAAACATGGTGCAGCAGGGAGGTCCAAAA
AGACAAAGACCCCACTAAGATGTTATAATTGTGGAAAATTTGGCCATATGCAAAG
ACAATGTCCGGAACCAAGGAAAACAAAATGTCTAAAGTGTGGAAAATTTGGGACACC
TAGCAAAAGACTGCAGGGGACAGGTGAATTTTTTTAGGGTATGGACGGTGGATGGGG
GCAAAACCGAGAAATTTTCCCGCCGCTACTCTTGGAGCGGAACCGAGTGCGCCTCCT
CCACCGAGCGGCACCACCCCATACGACCCAGCAAAGAAGCTCCTGCAGCAATATGC
AGAGAAAGGGAAACAACACTGAGGGAGCAAAAGAGGAATCCACCGGCAATGAATCCG
GATTGGACCGAGGGATATTCTTTGAACTCCCTCTTTGGAGAAGACCAATAAAGACA
GTGTATATAGAAGGGGTCCCATTAAGGCACTGCTAGACACAGGGGCAGATGACAC
CCATAATTAAGAAAATGATTTACAATTATCAGGTCCATGGAGACCCAAAATTATA
GGGGGCATAGGAGGAGGCCTTAATGTAAAAGAATATAACGACAGGGAAGTAAAAA
TAGAAGATAAAAATTTTGAGAGGAACAATATTGTTAGGAGCAACTCCCATTAATATA
ATAGGTAGAAATTTGCTGGCCCCGGCAGGTGCCCGGTTAGTAATGGGACAATTATC
AGAAAAAATTCCTGTACACCTGTCAAATTGAAGGAAGGGGCTCGGGGACCCTGTG
TAAGACAATGGCCTCTCTCTAAAGAGAAGATTGAAGCTTTACAGGAAATATGTTCCC
AATTAGAGCAGGAAGGAAAAATCAGTAGAGTAGGAGGAGAAAATGCATACAATAC
CCCAATATTTTGCATAAAGAAGAAGGACAAATCCCAGTGGAGGATGCTAGTAGACT

TTAGAGAGTTAAATAAGGCAACCCAAGATTTCTTTGAAGTGCAATTAGGGATACCCC
ACCCAGCAGGATTAAGAAAGATGAGACAGATAACAGTTTTAGATGTAGGAGACGCC
TATTATTCCATACCATTGGATCCAAATTTTAGGAAATATACTGCTTTTACTATTCCCA
CAGTGAATAATCAGGGACCCGGGATTAGGTATCAATTCAACTGTCTCCCGCAAGGG
TGGAAGGATCTCCTACAATCTTCCAAAATACAGCAGCATCCATTTTGGAGGAGATA
AAAAGAACTTGCCAGCACTAACCATTGTACAATACATGGATGATTTATGGGTAGG
TTCTCAAGAAAATGAACACACCCATGACAAATTAGTAGAACAGTTAAGAACAAAAT
TACAAGCCTGGGGCTTAGAAACCCAGAAAAGAAGGTGCAAAAAGAACCACCTTAT
GAGTGGATGGGATACAACTTTGGCCTCACAAATGGGAACTAAGCAGAATACAACT
GGAGGAAAAAGATGAATGGACTGTCAATGACATCCAGAAGTTAGTTGGGAACTAA
ATTGGGCAGCACAATTGTATCCAGGTCTTAGGACCAAGAATATATGCAAGTTAATTA
GAGGAAAGAAAAATCTGTTAGAGCTAGTGACTTGGACACCTGAGGCAGAAGCTGAA
TATGCAGAAAATGCAGAGATTCTTAAAACAGAACAGGAAGGAACCTATTACAAACC
AGGAATACCTATTAGGGCAGCAGTACAGAAATTGGAAGGAGGACAGTGGAGTTACC
AATTCAAACAAGAAGGACAAGTCTTGAAAGTAGGAAAATACACCAAGCAAAAAGAA
CACCCATACAAATGAACTTCGCACATTAGCTGGTTTAGTGCAGAAGATTTGCAAAGA
AGCTCTAGTTATTTGGGGGATATTACCAGTTCTAGAACTCCCGATAGAAAGAGAGGT
ATGGGAACAATGGTGGGCGGATTACTGGCAGGTAAGCTGGATTCCCGAATGGGATT
TTGTCAGCACCCACCTTTGCTCAAATATGGTACACATTAACAAAAGAACCATAAC
CCAAGGAGGACGTTTACTATGTAGATGGAGCATGCAACAGAAATTCAAAGAAGGA
AAAGCAGGATACATCTCACAATACGGAAAACAGAGAGTAGAAACATTAGAAAACA
CTACCAATCAGCAAGCAGAATTAACAGCTATAAAAATGGCTTTGGAAGACAGTGGG
CCTAATGTGAACATAGTAACAGACTCTCAATATGCAATGGGAATTTTGACAGCACA
ACCCACACAAAGTGATTCACCATTAGTAGAGCAAATTATAGCCTTAATGATACAAA
AGCAACAAATATATTTGCAGTGGGTACCAGCACATAAAGGAATAGGAGGAAATGAG
GAGATAGATAAATTAGTGAGTAAAGGCATTAGAAGAGTTTTATTCTTAGAAAAAAT
AGAAGAAGCTCAAGAAGAGCATGAAAGATATCATAATAATTGGAAAAACCTAGCA
GATACATATGGGCTTCCACAAATAGTAGCAAAAAGAGATAGTGGCCATGTGTCCAAA
ATGTCAGATAAAGGGAGAACCAGTGCATGGACAAGTGGATGCCTCACCTGGAACAT
GGCAGATGGATTGTACTCATCTAGAAGGAAAAGTAGTCATAGTTGCGGTCCATGTA
GCCAGTGGATTCATAGAAGCAGAAGTCATACCTAGGGAAACAGGAAAAGAAACGG
CAAAGTTTCTATTA AAAATACTGAGTAGATGGCCTATAACACAGTTACACACAGAC
AATGGGCCTAACTTTACCTCCCAAGAAGTGGCAGCAATATGTTGGTGGGGAAAAAT
TGAACATACAACAGGTATACCATATAACCCCAATCTCAAGGATCAATAGAAAGCA
TGAACAAACAATTA AAAAGAGATAATTGGGAAAATAAGAGATGATTGCCAATATACA
GAGACAGCAGTACTGATGGCTTGCCATATTCACAATTTTAAAAGAAAGGGAGGAAT
AGGGGGACAGACTTCAGCAGAGAGACTAATTAATATAATAACAACACAATTAGAAA
TACAACATTTACAAACCAAAAATTCAAAAAATTTTAAATTTTAGAGTCTACTACAGAG
AAGGGAGAGACCCTGTGTGGAAAGGACCAGCACAATTAATCTGGAAAGGGGAAGG
AGCAGTGGTCCTCAAGGACGGAAGTGACCTAAAGGTTGTACCAAGAAGGAAAGCTA

AAATTATTAAGGATTATGAACCCAAACAAAGAGTGGGTAATGAGGGTGACGTGGAA
GGTACCAGGGGATCTGATAACTAA

SEQ ID NO: 3: Плазида, определенная на фигуре 2А (pDNA1 pGM326)

Длина: 10528; Тип молекулы: ДНК; Особенности локализации/Определители:
источник, 1..10528; тип молекулы, другая ДНК; примечание: pGM326; организм,
синтетическая конструкция

GGTACCTCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAA
TCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTA
TATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTGATTATTGACTAGTTATTA
TAGTAATCAATTACGGGGTCATTAAGTTCATAGCCCATATATGGAGTTCCGCGTTACA
TAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCATTGACG
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAA
TGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATG
CCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCC
CAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
GCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTT
GACTCACGGGGATTTCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTTTTGG
CACCAAATCAACGGGACTTTCCAAATGTCGTAACAACACTGCGATCGCCCGCCCCGT
TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGC
TGGCTTGTAACCTCAGTCTCTTACTAGGAGACCAGCTTGAGCCTGGGTGTTGCTGGT
TAGCCTAACCTGGTTGGCCACCAGGGGTAAGGACTCCTTGGCTTAGAAAGCTAATA
AACTTGCCTGCATTAGAGCTTATCTGAGTCAAGTGTCTCATTGACGCCTCACTCTCT
TGAACGGGAATCTTCCTTACTGGGTTCTCTCTCTGACCCAGGCGAGAGAACTCCAG
CAGTGGCGCCCGAACAGGGACTTGAGTGAGAGTGTAGGCACGTACAGCTGAGAAGG
CGTCGGACGCGAAGGAAGCGCGGGGTGCGACGCGACCAAGAAGGAGACTTGGTGA
GTAGGCTTCTCGAGTGCCGGGAAAAAGCTCGAGCCTAGTTAGAGGACTAGGAGAGG
CCGTAGCCGTAACCTACTCTGGGCAAGTAGGGCAGGCGGTGGGTACGCAATGGGGGC
GGCTACCTCAGCACTAAATAGGAGACAATTAGACCAATTTGAGAAAATACGACTTC
GCCCGAACGGAAAGAAAAAGTACCAAATTAACATTTAATATGGGCAGGCAAGGA
GATGGAGCGCTTCGGCCTCCATGAGAGGTTGTTGGAGACAGAGGAGGGGTGTA
GAATCATAGAAGTCTCTACCCCTAGAACCAACAGGATCGGAGGGCTTAAAAAGT
CTGTTCAATCTTGTGTGCGTGCTATATTGCTTGCACAAGGAACAGAAAGTGAAAGAC
ACAGAGGAAGCAGTAGCAACAGTAAGACAACACTGCCATCTAGTGGA
AAAAGAAA
AAAGTGCAACAGAGACATCTAGTGGACAAAAGAAAAATGACAAGGGAATAGCAGC
GCCACCTGGTGGCAGTCAGAAATTTCCAGCGCAACAACAAGGAAATGCCTGGGTAC
ATGTACCCTTGTCACCGCGCACCTTAAATGCGTGGGTAAAAGCAGTAGAGGAGAAA
AAATTTGGAGCAGAAATAGTACCCATGTTTCAAGCCCTATCGAATTCCCGTTTGTGC
TAGGGTTCTTAGGCTTCTTGGGGGCTGCTGGA
ACTGCAATGGGAGCAGCGGCGACA
GCCCTGACGGTCCAGTCTCAGCATTTGCTTGTGCTGGGATACTGCAGCAGCAGAAGAAT
CTGCTGGCGGCTGTGGAGGCTCAACAGCAGATGTTGAAGCTGACCATTTGGGGTGT

AAAAACCTCAATGCCCGCGTCACAGCCCTTGAGAAGTACCTAGAGGATCAGGCACG
ACTAAACTCCTGGGGGTGCGCATGGAAACAAGTATGTCATAACCACAGTGGAGTGGC
CCTGGACAAATCGGACTCCGGATTGGCAAATATGACTTGGTTGGAGTGGGAAAGA
CAAATAGCTGATTTGGAAAGCAACATTACGAGACAATTAGTGAAGGCTAGAGAACA
AGAGGAAAAGAATCTAGATGCCTATCAGAAGTAACTAGTTGGTCAGATTTCTGGTC
TTGGTTCGATTTCTCAAATGGCTTAACATTTTAAAAATGGGATTTTTAGTAATAGTA
GGAATAATAGGGTTAAGATTACTTTACACAGTATATGGATGTATAGTGAGGGTTAG
GCAGGGATATGTTCCCTCTATCTCCACAGATCCATATCCGCGGCAATTTTAAAAGAAA
GGGAGGAATAGGGGGACAGACTTCAGCAGAGAGACTAATTAATATAATAACAACA
CAATTAGAAATACAACATTTACAAACCAAAATTCAAAAAATTTTAAATTTTAGAGCC
GCGGAGATCTGTTACATAACTTATGGTAAATGGCCTGCCTGGCTGACTGCCCAATGA
CCCCTGCCCAATGATGTCAATAATGATGTATGTTCCCATGTAATGCCAATAGGGACT
TTCCATTGATGTCAATGGGTGGAGTATTTATGGTAACTGCCACTTGGCAGTACATC
AAGTGTATCATATGCCAAGTATGCCCCCTATTGATGTCAATGATGGTAAATGGCCTG
CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCT
ATGTATTAGTCATTGCTATTACCATGGGAATTCCTAGTGGAGAAGAGCATGCTTGA
GGGCTGAGTGCCCTCAGTGGGCAGAGAGCACATGGCCACAGTCCCTGAGAAGTT
GGGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAA
CTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACC
ATATATAAGTGCAGTAGTCTCTGTGAACATTCAAGCTTCTGCCTTCTCCCTCCTGTGA
GTTTGCTAGCCACCATGCAGAGAAGCCCTCTGGAGAAGGCCTCTGTGGTGAGCAAG
CTGTTCTTCAGCTGGACCAGGCCATCCTGAGGAAGGGCTACAGGCAGAGACTGGA
GCTGTCTGACATCTACCAGATCCCCTCTGTGGACTCTGCTGACAACCTGTCTGAGAA
GCTGGAGAGGGAGTGGGATAGAGAGCTGGCCAGCAAGAAGAACCCCAAGCTGATC
AATGCCCTGAGGAGATGCTTCTTCTGGAGATTCATGTTCTATGGCATCTTCTGTACC
TGGGGGAAGTGACCAAGGCTGTGCAGCCTCTGCTGCTGGGCAGAATCATTGCCAGC
TATGACCCTGACAACAAGGAGGAGAGGAGCATTGCCATCTACCTGGGCATTGGCCT
GTGCCTGCTGTTCAATTGTGAGGACCCTGCTGCTGCACCCTGCCATCTTTGGCCTGCAC
CACATTGGCATGCAGATGAGGATTGCCATGTTACAGCCTGATCTACAAGAAAACCCTG
AAGCTGTCCAGCAGAGTGCTGGACAAGATCAGCATTGGCCAGCTGGTGAGCCTGCT
GAGCAACAACCTGAACAAGTTTGATGAGGGCCTGGCCCTGGCCCCTTTGTGTGGAT
TGCCCCTCTGCAGGTGGCCCTGCTGATGGGCCTGATTTGGGAGCTGCTGCAGGCCTC
TGCCTTTTGTGGCCTGGGCTTCCCTGATTGTGCTGGCCCTGTTTCAGGCTGGCCTGGGC
AGGATGATGATGAAGTACAGGGACCAGAGGGCAGGCAAGATCAGTGAGAGGCTGG
TGATCACCTCTGAGATGATTGAGAACATCCAGTCTGTGAAGGCCTACTGTTGGGAGG
AAGCTATGGAGAAGATGATTGAAAACCTGAGGCAGACAGAGCTGAAGCTGACCAG
GAAGGCTGCCTATGTGAGATACTTCAACAGCTCTGCCTTCTTCTTCTCTGGCTTCTTT
GTGGTGTTCCTGTCTGTGCTGCCCTATGCCCTGATCAAGGGGATCATCCTGAGAAAAG
ATTTTCACCACCATCAGCTTCTGCATTGTGCTGAGGATGGCTGTGACCAGACAGTTC
CCCTGGGCTGTGCAGACCTGGTATGACAGCCTGGGGGCCATCAACAAGATCCAGGA

CTTCCTGCAGAAGCAGGAGTACAAGACCCTGGAGTACAACCTGACCACCACAGAAG
TGGTGATGGAGAATGTGACAGCCTTCTGGGAGGAGGGCTTTGGGGAGCTGTTT
AAGGCCAAGCAGAACAACAACAGAAAGACCAGCAATGGGGATGACTCCCTGT
TCTTCTCCAACCTTCTCCCTGCTGGGCACACCTGTGCTGAAGGACATCAACTTCAAGA
TTGAGAGGGGGCAGCTGCTGGCTGTGGCTGGATCTACAGGGGCTGGCAAGACCAGC
CTGCTGATGATGATCATGGGGGAGCTGGAGCCTTCTGAGGGCAAGATCAAGCACTC
TGGCAGGATCAGCTTTTGCAGCCAGTTCAGCTGGATCATGCCTGGCACCATCAAGGA
GAACATCATCTTTGGAGTGAGCTATGATGAGTACAGATACAGGAGTGTGATCAAGG
CCTGCCAGCTGGAGGAGGACATCAGCAAGTTTGTGCTGAGAAGGACAACATTGTGCTG
GGGGAGGGAGGCATTAACTGTCTGGGGGCCAGAGAGCCAGAATCAGCCTGGCCAG
GGCTGTGTACAAGGATGCTGACCTGTACCTGCTGGACTCCCCCTTTGGCTACCTGGA
TGTGCTGACAGAGAAGGAGATTTTTGAGAGCTGTGTGTGCAAGCTGATGGCCAACA
AGACCAGAATCCTGGTGACCAGCAAGATGGAGCACCTGAAGAAGGCTGACAAGATC
CTGATCCTGCATGAGGGCAGCAGCTACTTCTATGGGACCTTCTCTGAGCTGCAGAAC
CTGCAGCCTGACTTCAGCTCTAAGCTGATGGGCTGTGACAGCTTTGACCAGTTCTCT
GCTGAGAGGAGGAACAGCATCCTGACAGAGACCCTGCACAGATTCAGCCTGGAGGG
AGATGCCCTGTGAGCTGGACAGAGACCAAGAAGCAGAGCTTCAAGCAGACAGGG
GAGTTTGGGGAGAAGAGGAAGAACTCCATCCTGAACCCCATCAACAGCATCAGGAA
GTTCAGCATTGTGCAGAAAACCCCTGCAGATGAATGGCATTGAGGAAGATTCTG
ATGAGCCCCTGGAGAGGAGACTGAGCCTGGTGCCTGATTCTGAGCAGGGAGAGGCC
ATCCTGCCTAGGATCTCTGTGATCAGCACAGGCCCTAACTGCAGGCCAGAAGGAG
GCAGTCTGTGCTGAACCTGATGACCCACTCTGTGAACCAGGGCCAGAACATCCACA
GGAAAACCACAGCCTCCACCAGGAAAGTGAGCCTGGCCCCTCAGGCCAATCTGACA
GAGCTGGACATCTACAGCAGGAGGCTGTCTCAGGAGACAGGCCTGGAGATTTCTGA
GGAGATCAATGAGGAGGACCTGAAAGAGTGCTTCTTTGATGACATGGAGAGCATCC
CTGCTGTGACCACCTGGAACACCTACCTGAGATACATCACAGTGCACAAGAGCCTG
ATCTTTGTGCTGATCTGGTGCCTGGTGATCTTCTGGCTGAAGTGGCTGCCTCTCTGG
TGGTGCTGTGGCTGCTGGGAAACACCCCACTGCAGGACAAGGGCAACAGCACCCAC
AGCAGGAACAACAGCTATGCTGTGATCATCACCTCCACCTCCAGCTACTATGTGTTT
TACATCTATGTGGGAGTGGCTGATACCCTGCTGGCTATGGGCTTCTTTAGAGGCCTG
CCCCTGGTGCACACACTGATCACAGTGAGCAAGATCCTCCACCACAAGATGCTGCA
CTCTGTGCTGCAGGCTCCTATGAGCACCTGAATACCCTGAAGGCTGGGGGCATCCT
GAACAGATTCTCCAAGGATATTGCCATCCTGGATGACCTGCTGCCTCTCACCATCTT
TGACTTCATCCAGCTGCTGCTGATTGTGATTGGGGCCATTGCTGTGGTGGCAGTGCT
GCAGCCCTACATCTTTGTGGCCACAGTGCCTGTGATTGTGGCCTTCATCATGCTGAG
GGCCTACTTTCTGCAGACCTCCCAGCAGCTGAAGCAGCTGGAGTCTGAGGGCAGAA
GCCCCATCTTCAACCACCTGGTGACAAGCCTGAAGGGCCTGTGGACCCTGAGAGCCT
TTGGCAGGCAGCCCTACTTTGAGACCCTGTTCCACAAGGCCCTGAACCTGCACACAG
CCAACCTGGTTCTCTACCTGTCCACCCTGAGATGGTTCCAGATGAGAATTGAGATGA
TCTTTGTCATCTTCTTCATTGCTGTGACCTTCATCAGCATTCTGACCACAGGAGAGGG

AGAGGGCAGAGTGGGCATTATCCTGACCCTGGCCATGAACATCATGAGCACACTGC
AGTGGGCAGTGAACAGCAGCATTGATGTGGACAGCCTGATGAGGAGTGTGAGCAGA
GTGTTCAAGTTCATTGATATGCCACAGAGGGCAAGCCTACCAAGAGCACCAAGCC
CTACAAGAATGGCCAGCTGAGCAAAGTGATGATCATTGAGAACAGCCATGTGAAGA
AGGATGATATCTGGCCCAGTGGAGGCCAGATGACAGTGAAGGACCTGACAGCCAAG
TACACAGAGGGGGGCAATGCTATCCTGGAGAACATCTCCTTCAGCATCTCCCCTGGC
CAGAGAGTGGGACTGCTGGGAAGAACAGGCTCTGGCAAGTCTACCCTGCTGTCTGC
CTTCCTGAGGCTGCTGAACACAGAGGGAGAGATCCAGATTGATGGAGTGTCTGGG
ACAGCATCACTGCAGCAGTGGAGGAAGGCCTTTGGTGTGATCCCCCAGAAAGTG
TTCATCTTCAGTGGCACCTTCAGGAAGAACCTGGACCCCTATGAGCAGTGGTCTGAC
CAGGAGATTTGGAAAGTGGCTGATGAAGTGGGCCTGAGAAGTGTGATTGAGCAGTT
CCCTGGCAAGCTGGACTTTGTCCTGGTGGATGGGGGCTGTGTGCTGAGCCATGGCCA
CAAGCAGCTGATGTGCCTGGCCAGATCAGTGTGAGCAAGGCCAAGATCCTGCTGC
TGGATGAGCCTTCTGCCACCTGGATCCTGTGACCTACCAGATCATCAGGAGGACCC
TCAAGCAGGCCTTTGCTGACTGCACAGTCATCCTGTGTGAGCACAGGATTGAGGCCA
TGCTGGAGTGCCAGCAGTTCCTGGTATTGAGGAGAACAAAGTGAGGCAGTATGAC
AGCATCCAGAAGCTGCTGAATGAGAGGAGCCTGTTTCAGGCAGGCCATCAGCCCCTC
TGATAGAGTGAAGCTGTTCCCCACAGGAACAGCTCCAAGTGCAAGAGCAAGCCCC
AGATTGCTGCCCTGAAGGAGGAGACAGAGGAGGAAGTGCAGGACACCAGGCTGTG
AGGGCCCAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAA
CTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCT
ATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCT
TTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGC
TGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGAC
TTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACATCGCCGCCTGCCTTGCCCGC
TGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTGTCGGGGAA
ATCATCGTCCTTTCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACG
TCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGC
TGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTC
CCTTTGGGCCGCCTCCCCGCAAGCTTCGCACTTTTTAAAAGAAAAGGGAGGACTGGA
TGGGATTTATTACTCCGATAGGACGCTGGCTTGTAACTCAGTCTCTTACTAGGAGAC
CAGCTTGAGCCTGGGTGTTGCTGGTTAGCCTAACCTGGTTGGCCACCAGGGGTAAG
GACTCCTTGGCTTAGAAAGCTAATAAACTTGCCTGCATTAGAGCTCTTACGCGTCCC
GGGCTCGAGATCCGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGC
CCATCCCGCCCCTAACTCCGCCCAGTTCGCCCATTCTCCGCCCCATGGCTGACTAAT
TTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTA
GTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTTGCAAAAAGCTAACTTGTTTATTGCA
GCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATT
TTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCT
GTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGAGCGGT

ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAG
 GAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGC
 GTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACG
 CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
 CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC
 CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTC
 AGTTCGGTGTAGGTTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAG
 CCCGACCCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC
 GACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
 AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAA
 CAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA
 GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGC
 AGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG
 GGGTCTGACGCTCAGTGGAACGAAAACACTCACGTTAAGGGATTTTGGTCATGAGATT
 ATCAAAAAGGATCTTACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAATCAAT
 CTAAGTATATATGAGTAACTTGGTCTGACAGTTAGAAAAACTCATCGAGCATCA
 AATGAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCC
 GTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCC
 TGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCT
 CGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGT
 GAGAATGGCAACAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTA
 CGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCC
 TGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGA
 ATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAG
 GATATTCTTCTAATACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACC
 ATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATCC
 GTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGC
 CATGTTTCAGAAACA ACTCTGGCGCATCGGGCTTCCATAACAATCGATAGATTGTTCG
 CACCTGATTGCCCGACATTATCGCGAGCCATTTATACCCATATAAATCAGCATCCA
 TGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAA
 CACCCCTTGTAATTA CTGTTTATGTAAGCAGACAGTTTTTATTGTTTCATGATGATATATT
 TTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACAATTGGTCGACGGAT
 CC

SEQ ID NO: 4: Плазмида, определенная на фигуре 2B (pDNA1 pGM830)

Длина: 10536; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..10536; тип молекулы, другая ДНК; примечание: pGM830; организм,
 синтетическая конструкция

GGTACCTCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAA
 TCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTA
 TATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTGATTATTGACTAGTTATTA

TAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGTTACA
TAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCATTGACG
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAA
TGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATG
CCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCC
CAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
GCTATTACCATGGTGATGCGGTTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTT
GACTCACGGGGATTTCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTTTTGG
CACCAAATCAACGGGACTTTCCAAATGTCGTAACAACACTGCGATCGCCCGCCCCGT
TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGC
TGGCTTGTAACCTCAGTCTCTTACTAGGAGACCAGCTTGAGCCTGGGTGTTCGCTGGT
TAGCCTAACCTGGTTGGCCACCAGGGGTAAGGACTCCTTGGCTTAGAAAGCTAATA
AACTTGCCTGCATTAGAGCTTATCTGAGTCAAGTGTCTCATTGACGCCTCACTCTCT
TGAACGGGAATCTTCCTTACTGGGTTCTCTCTCTGACCCAGGCGAGAGAACTCCAG
CAGTGGCGCCCGAACAGGGACTTGAGTGAGAGTGTAGGCACGTACAGCTGAGAAGG
CGTCGGACGCGAAGGAAGCGCGGGGTGCGACGCGACCAAGAAGGAGACTTGGTGA
GTAGGCTTCTCGAGTGCCGGGAAAAAGCTCGAGCCTAGTTAGAGGACTAGGAGAGG
CCGTAGCCGTAACACTCTGGGCAAGTAGGGCAGGCGGTGGGTACGCAATTGGGGG
CGGCTACCTCAGCACTAAATAGGAGACAATTAGACCAATTTGAGAAAATACGACTT
CGCCCGAACGGAAAGAAAAAGTACCAAATTAACATTTAATATTGGGCAGGCAAGG
AGATTGGAGCGCTTCGGCCTCCATGAGAGGTTGTTGGAGACAGAGGAGGGGTGTAA
AAGAATCATAGAAGTCCTCTACCCCTAGAACCAACAGGATCGGAGGGCTTAAAAA
GTCTGTTCAATCTTGTGTGCGTGCTATATTGCTTGCACAAGGAACAGAAAGTGAAAG
ACACAGAGGAAGCAGTAGCAACAGTAAGACAACACTGCCATCTAGTGGAAAAAGA
AAAAAGTGCAACAGAGACATCTAGTGGACAAAAGAAAAATGACAAGGGAATAGCA
GCGCCACCTGGTGGCAGTCAGAATTTTCCAGCGCAACAACAAGGAAATTGCCTGGG
TACATGTACCCTTGTACCGCGCACCTTAAATGCGTGGGTAAAAGCAGTAGAGGAG
AAAAAATTTGGAGCAGAAATAGTACCCATGTTTCAAGCCCTATCGCCTGCAGGCCGT
TTGTGCTAGGGTTCTTAGGCTTCTTGGGGGCTGCTGGAAGTGCATTGGGAGCAGCGG
CGACAGCCCTGACGGTCCAGTCTCAGCATTGCTTGTGGGATACTGCAGCAGCAGA
AGAATCTGCTGGCGGCTGTGGAGGCTCAACAGCAGATGTTGAAGCTGACCATTTGG
GGTGTAAAAACCTCAATGCCCCGCGTCACAGCCCTTGAGAAGTACCTAGAGGATCA
GGCAGACTAACTCCTGGGGGTGCGCATGGAAACAAGTATGTCATAACCACAGTGG
AGTGGCCCTGGACAAATCGGACTCCGGATTGGCAAATAAGACTTGGTTGGAGTGG
GAAAGACAAATAGCTGATTTGGAAAGCAACATTACGAGACAATTAGTGAAGGCTAG
AGAACAAGAGGAAAAGAATCTAGATGCCTATCAGAAGTTAACTAGTTGGTCAGATT
TCTGGTCTTGGTTCGATTTCTCAAATGGCTTAACATTTTAAAAAAGGGATTTTLAGT
AATAGTAGGAATAATAGGGTTAAGATTACTTTACACAGTATATGGATGTATAGTGA
GGGTAGGCAGGGATATGTTCCCTCTATCTCCACAGATCCATATAAAGCGGCAATTTT
AAAAGAAAGGGAGGAATAGGGGGACAGACTTCAGCAGAGAGACTAATTAATATAA

TAACAACACAATTAGAAATACAACATTTACAAAACCAAATTCAAAAAATTTTAAAT
TTTAGAGCCGCGGAGATCTGTTACATAACTTATGGTAAATGGCCTGCCTGGCTGACT
GCCCAATGACCCCTGCCCAATGATGTCAATAATGATGTATGTTCCCATGTAATGCCA
ATAGGGACTTTCCATTGATGTCAATGGGTGGAGTATTTATGGTAACTGCCCACTTGG
CAGTACATCAAGTGTATCATATGCCAAGTATGCCCCCTATTGATGTCAATGATGGTA
AATGGCCTGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGC
AGTACATCTATGTATTAGTCATTGCTATTACCATGGGAATTCCTAGTGGAGAAGAG
CATGCTTGAGGGCTGAGTGCCCCCTCAGTGGGCAGAGAGCACATGGCCACAGTCCC
TGAGAAGTTGGGGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCT
TGGGTAACTGGGAAAGTGTGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGG
GGAGAACCATATATAAGTGCAGTAGTCTCTGTGAACATTCAAGCTTCTGCCTTCTCC
CTCCTGTGAGTTTGCTAGCCACCATGCAGAGAAGCCCTCTGGAGAAGGCCTCTGTGG
TGAGCAAGCTGTTCTTCAGCTGGACCAGGCCATCCTGAGGAAGGGCTACAGGCAG
AGACTGGAGCTGTCTGACATCTACCAGATCCCCTCTGTGGACTCTGCTGACAACCTG
TCTGAGAAGCTGGAGAGGGAGTGGGATAGAGAGCTGGCCAGCAAGAAGAACCCCA
AGCTGATCAATGCCCTGAGGAGATGCTTCTTCTGGAGATTCATGTTCTATGGCATCT
TCCTGTACCTGGGGGAAGTGACCAAGGCTGTGCAGCCTCTGCTGCTGGGCAGAATC
ATTGCCAGCTATGACCCTGACAACAAGGAGGAGAGGAGCATTGCCATCTACCTGGG
CATTGGCCTGTGCCTGCTGTTCAATTGTGAGGACCCTGCTGCTGCACCCTGCCATCTTT
GGCCTGCACCACATTGGCATGCAGATGAGGATTGCCATGTTTACGCCTGATCTACAAG
AAAACCCTGAAGCTGTCCAGCAGAGTGTGGACAAGATCAGCATTGGCCAGCTGGT
GAGCCTGCTGAGCAACAACCTGAACAAGTTTGATGAGGGCCTGGCCCTGGCCCACT
TTGTGTGGATTGCCCTCTGCAGGTGGCCCTGCTGATGGGCCTGATTTGGGAGCTGC
TGCAGGCCTCTGCCTTTTGTGGCCTGGGCTTCTGATTGTGCTGGCCCTGTTTCAGGC
TGGCCTGGGCAGGATGATGATGAAGTACAGGGACCAGAGGGCAGGCAAGATCAGT
GAGAGGCTGGTGATCACCTCTGAGATGATTGAGAACATCCAGTCTGTGAAGGCCTA
CTGTTGGGAGGAAGCTATGGAGAAGATGATTGAAAACCTGAGGCAGACAGAGCTGA
AGCTGACCAGGAAGGCTGCCTATGTGAGATACTTCAACAGCTCTGCCTTCTTCTTCT
CTGGCTTCTTTGTGGTGTTCCTGTCTGTGCTGCCCTATGCCCTGATCAAGGGGATCAT
CCTGAGAAAGATTTTCACCACCATCAGCTTCTGCATTGTGCTGAGGATGGCTGTGAC
CAGACAGTTCCCCTGGGCTGTGCAGACCTGGTATGACAGCCTGGGGGCCATCAACA
AGATCCAGGACTTCTGCAGAAGCAGGAGTACAAGACCCTGGAGTACAACCTGACC
ACCACAGAAGTGGTGATGGAGAATGTGACAGCCTTCTGGGAGGAGGGCTTTGGGGA
GCTGTTTGAGAAGGCCAAGCAGAACAACAACAACAGAAAGACCAGCAATGGGGAT
GACTCCCTGTTCTTCTCCAACCTTCTCCCTGCTGGGCACACCTGTGCTGAAGGACATCA
ACTTCAAGATTGAGAGGGGGCAGCTGCTGGCTGTGGCTGGATCTACAGGGGCTGGC
AAGACCAGCCTGCTGATGATGATCATGGGGGAGCTGGAGCCTTCTGAGGGCAAGAT
CAAGCACTCTGGCAGGATCAGCTTTTGCAGCCAGTTCAGCTGGATCATGCCTGGCAC
CATCAAGGAGAACATCATCTTTGGAGTGAGCTATGATGAGTACAGATAACAGGAGTG
TGATCAAGGCCTGCCAGCTGGAGGAGGACATCAGCAAGTTTGCTGAGAAGGACAAC

ATTGTGCTGGGGGAGGGAGGCATTACACTGTCTGGGGGCCAGAGAGCCAGAATCAG
CCTGGCCAGGGCTGTGTACAAGGATGCTGACCTGTACCTGCTGGACTCCCCCTTTGG
CTACCTGGATGTGCTGACAGAGAAGGAGATTTTTGAGAGCTGTGTGTGCAAGCTGAT
GGCCAACAAGACCAGAATCCTGGTGACCAGCAAGATGGAGCACCTGAAGAAGGCT
GACAAGATCCTGATCCTGCATGAGGGCAGCAGCTACTTCTATGGGACCTTCTCTGAG
CTGCAGAACCTGCAGCCTGACTTCAGCTCTAAGCTGATGGGCTGTGACAGCTTTGAC
CAGTTCTCTGCTGAGAGGAGGAACAGCATCCTGACAGAGACCCTGCACAGATTGAG
CCTGGAGGGAGATGCCCCCTGTGAGCTGGACAGAGACCAAGAAGCAGAGCTTCAAGC
AGACAGGGGAGTTTGGGGAGAAGAGGAAGAAGACTCCATCCTGAACCCCATCAACAGC
ATCAGGAAGTTCAGCATTGTGCAGAAAACCCCCCTGCAGATGAATGGCATTGAGGA
AGATTCTGATGAGCCCCTGGAGAGGAGACTGAGCCTGGTGCCTGATTCTGAGCAGG
GAGAGGCCATCCTGCCTAGGATCTCTGTGATCAGCACAGGCCCTACACTGCAGGCC
AGAAGGAGGCAGTCTGTGCTGAACCTGATGACCCACTCTGTGAACCAGGGCCAGAA
CATCCACAGGAAAACCACAGCCTCCACCAGGAAAGTGAGCCTGGCCCCCTCAGGCCA
ATCTGACAGAGCTGGACATCTACAGCAGGAGGCTGTCTCAGGAGACAGGCCTGGAG
ATTTCTGAGGAGATCAATGAGGAGGACCTGAAAGAGTGCTTCTTTGATGACATGGA
GAGCATCCCTGCTGTGACCACCTGGAACACCTACCTGAGATACATCACAGTGCACA
AGAGCCTGATCTTTGTGCTGATCTGGTGCCTGGTGTCTTCTGGCTGAAGTGGCTG
CCTCTCTGGTGGTGTGCTGTGGCTGCTGGGAAACACCCCACTGCAGGACAAGGGCAAC
AGCACCCACAGCAGGAACAACAGCTATGCTGTGATCATCACCTCCACCTCCAGCTAC
TATGTGTTCTACATCTATGTGGGAGTGGCTGATACCCTGCTGGCTATGGGCTTCTTTA
GAGGCCTGCCCCCTGGTGCACACACTGATCACAGTGAGCAAGATCCTCCACCACAAG
ATGCTGCACTCTGTGCTGCAGGCTCCTATGAGCACCCCTGAATACCCTGAAGGCTGGG
GGCATCCTGAACAGATTCTCCAAGGATATTGCCATCCTGGATGACCTGCTGCCTCTC
ACCATCTTTGACTTCATCCAGCTGCTGCTGATTGTGATTGGGGCCATTGCTGTGGTGG
CAGTGCTGCAGCCCTACATCTTTGTGGCCACAGTGCCTGTGATTGTGGCCTTCATCAT
GCTGAGGGCCTACTTTCTGCAGACCTCCCAGCAGCTGAAGCAGCTGGAGTCTGAGG
GCAGAAGCCCCATCTTCACCCACCTGGTGACAAGCCTGAAGGGCCTGTGGACCCTG
AGAGCCTTTGGCAGGCAGCCCTACTTTGAGACCCTGTTCCACAAGGCCCTGAACCTG
CACACAGCCAACCTGGTTCCTCTACCTGTCCACCCTGAGATGGTTCAGATGAGAATT
GAGATGATCTTTGTCATCTTCTTCATTGCTGTGACCTTCATCAGCATTCTGACCACAG
GAGAGGGAGAGGGCAGAGTGGGCATTATCCTGACCCTGGCCATGAACATCATGAGC
ACACTGCAGTGGGCAGTGAACAGCAGCATTGATGTGGACAGCCTGATGAGGAGTGT
GAGCAGAGTGTTCAAGTTCATTGATATGCCACAGAGGGCAAGCCTACCAAGAGCA
CCAAGCCCTACAAGAATGGCCAGCTGAGCAAAGTGATGATCATTGAGAACAGCCAT
GTGAAGAAGGATGATATCTGGCCCAGTGGAGGCCAGATGACAGTGAAGGACCTGAC
AGCCAAGTACACAGAGGGGGGCAATGCTATCCTGGAGAACATCTCCTTCAGCATCT
CCCCTGGCCAGAGAGTGGGACTGCTGGGAAGAACAGGCTCTGGCAAGTCTACCCTG
CTGTCTGCCTTCCTGAGGCTGCTGAACACAGAGGGAGAGATCCAGATTGATGGAGT
GTCCTGGGACAGCATCACACTGCAGCAGTGGAGGAAGGCCTTTGGTGTGATCCCCC

AGAAAGTGTTTCATCTTCAGTGGCACCTTCAGGAAGAACCTGGACCCCTATGAGCAGT
GGTCTGACCAGGAGATTTGGAAAGTGGCTGATGAAGTGGGCCTGAGAAGTGTGATT
GAGCAGTTCCCTGGCAAGCTGGACTTTGTCCTGGTGGATGGGGGCTGTGTGCTGAGC
CATGGCCACAAGCAGCTGATGTGCCTGGCCAGATCAGTGCTGAGCAAGGCCAAGAT
CCTGCTGCTGGATGAGCCTTCTGCCCACCTGGATCCTGTGACCTACCAGATCATCAG
GAGGACCCTCAAGCAGGCCTTTGCTGACTGCACAGTCATCCTGTGTGAGCACAGGAT
TGAGGCCATGCTGGAGTGCCAGCAGTTCCTGGTGATTGAGGAGAACAAAGTGAGGC
AGTATGACAGCATCCAGAAGCTGCTGAATGAGAGGAGCCTGTTTCAGGCAGGCCATC
AGCCCCTCTGATAGAGTGAAGCTGTTCCCCACAGGAACAGCTCCAAGTGCAAGAG
CAAGCCCCAGATTGCTGCCCTGAAGGAGGAGACAGAGGAGGAAGTGCAGGACACC
AGGCTGTGAGGGCCCAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGT
ATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGT
ATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTG
CTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACT
GTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTT
TCCGGGACTTTCGCTTTCCTCCCTATTGCCACGGCGGAATCATCGCCGCCTGCC
TTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGTTGGTGTGTG
CGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCG
CGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGC
GGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTC
GGATCTCCCTTTGGGCCGCCTCCCCGCAAGCTTCGCACTTTTTAAAAGAAAAGGGAG
GACTGGATGGGATTTATTACTCCGATAGGACGCTGGCTTGTAACCTCAGTCTCTTACT
AGGAGACCAGCTTGAGCCTGGGTGTTGCTGGTTAGCCTAACCTGGTTGGCCACCAG
GGGTAAGGACTCCTTGGCTTAGAAAGCTAATAAACTTGCCTGCATTAGAGCTCTTAC
GCGTCCCGGGCTCGAGATCCGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTA
ACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCGCCCATTTCTCCGCCCATGGCT
GACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCA
GAAGTAGTGAGGAGGCTTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTAACTTGTTT
ATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAA
AGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTAT
CATGTCTGTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGGC
AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATA
ACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAA
GGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAA
TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
CCTGTCCGCTTTTCTCCCTTCGGGAAGCGTGGCGCTTTTCTCATAGCTCACGCTGTAGG
TATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
GTTTCAGCCCGACCCTGCGCCTTATCCGGTAACCTATCGTCTTGAGTCCAACCCGGTA
AGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG

GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAG
 AAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGT
 TGGTAGCTCTTGATCCGGCAAACAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTG
 CAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTT
 CTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTAAAGGGATTTTGGTCATG
 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAA
 TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTAGAAAAACTCATCGAGC
 ATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAA
 AGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAG
 ATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTC
 CCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATC
 CGGTGAGAATGGCAACAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCC
 ATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTG
 CGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAA
 TCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAA
 TCAGGATATTCTTCTAATACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGT
 AACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAA
 TTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCT
 TTGCCATGTTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCATACAATCGATAGATT
 GTCGCACCTGATTGCCCGACATTATCGCGAGCCATTTATACCCATATAAATCAGCA
 TCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTC
 ATAACACCCCTTGTATTA CTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGATGATA
 TATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACAATTGGTCGAC
 GGATCC

SEQ ID NO: 5: Плазмида, определенная на фигуре 2B (pDNA2a pGM691)

Длина: 9064; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..9064; тип молекулы, другая ДНК; примечание: pGM691; организм,
 синтетическая конструкция

ATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
 AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGA
 CCGCCCAACGACCCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCAC
 TTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGAC
 GGTAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACT
 TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAGGTGAGCCCCACG
 TTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTAT
 TTTTAAATATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGGGCGCGCGCCAG
 GCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGGCGG
 CAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGGCGGGG
 CGGCGGCCCTATAAAAAGCGAAGCGCGCGGGCGGGGAGTCGCTGCGCGCTGCCT

TCGCCCCGTGCCCCGCTCCGCCGCCGCCTCGCGCCGCCGCCCGCCCCGGCTCTGACTGAC
CGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTA
GCGCTTGGTTTAATGACGGCTTGTCTTTCTGTGGCTGCGTGAAAGCCTTGAGGG
GCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
TGTGCGTGGGGAGCGCCGCGTGC GGCTCCGCGCTGCCCGCGGGCTGTGAGCGCTGC
GGGCGCGGCGCGGGGCTTTGTGCGCTCCGCAGTGTGCGCGAGGGGAGCGCGGCCGG
GGGCGGTGCCCGCGGTGCGGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGG
GGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGTCGGTCGGGCTGCAACC
CCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCT
CCGTACGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGG
GGGTGCCGGGCGGGGCGGGGCCCGCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCG
CGGCGGCCCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGCAGCCATTGCC
TTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGG
AGCCGAAATCTGGGAGGCGCCGCCGCACCCCCCTTAGCGGGCGCGGGGCGAAGCGG
TGCGGCGCCGCGCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCG
CCGTCCCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
GGGGGGACGGGGCAGGGCGGGGTTCCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAG
CCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCT
GGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTGCTCGAGCCACCATGGGAGCT
GCCACATCTGCCCTGAATAGACGGCAGCTGGACCAGTTCGAGAAGATCAGACTGCG
GCCAACGGCAAGAAGAAGTACCAGATCAAGCACCTGATCTGGGCCGGCAAAGAG
ATGGAAAGATTCCGGCCTGCACGAGCGGCTGCTGGAAACCGAGGAAGGCTGCAAGA
GAATTATCGAGGTGCTGTACCCTCTGGAACCTACCGGCTCTGAGGGCCTGAAGTCCC
TGTTCAATCTCGTGTGCGTGCTGTACTGCCTGCACAAAGAACAGAAAGTGAAGGAC
ACCGAAGAGGCCGTGGCCACAGTTAGACAGCACTGCCACCTGGTGGAAAAAGAGA
AGTCCGCCACAGAGACAAGCAGCGGCCAGAAGAAGAACGACAAGGGAATTGCTGC
CCCTCCTGGCGGCAGCCAGAATTTCTGCTCAGCAGCAGGGAAACGCCTGGGTGC
ACGTTCCACTGAGCCCTAGAACACTGAATGCCTGGGTCAAAGCCGTGGAAGAGAAG
AAGTTTGGCGCCGAGATCGTGCCCATGTTCCAGGCTCTGTCTGAGGGCTGCACCCCT
TACGACATCAACCAGATGCTGAACGTGCTGGGAGATCACCAGGGCGCTCTGCAGAT
CGTGAAAGAGATCATCAACGAAGAGGCTGCCAGTGGGACGTGACACATCCATTGC
CTGCTGGACCTCTGCCAGCCGGACAACCTGAGAGATCCTAGAGGCTCTGATATCGCCG
GCACCACCAGCTCTGTGCAAGAGCAGCTGGAATGGATCTACACCGCCAATCCTAGA
GTGGACGTGGGCGCCATCTACAGAAGATGGATCATCCTGGGCCTGCAGAAATGCGT
GAAGATGTACAACCCCGTGTCCGTGCTGGACATCAGACAGGGACCCAAAGAGCCCT
TCAAGGACTACGTGGACCGGTTCTATAAGGCCATTAGAGCCGAGCAGGCCAGCGGC
GAAGTGAAGCAGTGGATGACAGAGAGCCTGCTGATCCAGAACGCCAATCCAGACTG
CAAAGTGATCCTGAAAGGCCTGGGCATGCACCCCACACTGGAAGAGATGCTGACAG
CCTGTCAAGGCGTTGGCGGCCCTTCTTACAAAGCCAAAGTGATGGCCGAGATGATG
CAGACCATGCAGAACCAGAACATGGTGCAGCAAGGCGGCCCTAAGAGACAGAGGC

CTCCTCTGAGATGCTACAACCTGCGGCAAGTTCGGCCACATGCAGAGACAGTGTCTCTG
AGCCTAGGAAAACAAAATGTCTAAAGTGTGGAAAATTGGGACACCTAGCAAAAAGAC
TGCAGGGGACAGGTGAATTTTTTAGGGTATGGACGGTGGATGGGGGCAAAACCGAG
AAATTTTCCCGCCGCTACTCTTGGAGCGGAACCGAGTGCGCCTCCTCCACCGAGCGG
CACCACCCCATACGACCCAGCAAAGAAGCTCCTGCAGCAATATGCAGAGAAAGGGA
AACAACTGAGGGAGCAAAAGAGGAATCCACCGGCAATGAATCCGGATTGGACCGA
GGGATATTCTTTGAACTCCCTCTTTGGAGAAGACCAATAAAGACCGTGTACATCGAG
GGCGTGCCCATCAAGGCTCTGCTGGATACAGGCGCCGACGACACCATCATCAAAGA
GAACGACCTGCAGCTGAGCGGCCCTTGGAGGCCTAAGATCATTGGAGGAATCGGCG
GAGGCCTGAACGTCAAAGAGTACAACGACCGGGAAGTGAAGATCGAGGACAAGAT
CCTGAGGGGCACAATCCTGCTGGGCGCCACACCTATCAACATCATCGGCAGAAATC
TGCTGGCCCCTGCCGGCGCTAGACTGGTTATGGGACAGCTCTCTGAGAAGATCCCCG
TGACACCCGTGAAGCTGAAAGAAGGCGCTAGAGGACCTTGTGTGCGACAGTGGCCT
CTGAGCAAAGAGAAGATTGAGGCCCTGCAAGAAATCTGTAGCCAGCTGGAACAAGA
GGGCAAGATCAGCAGAGTTGGCGGCGAGAACGCCTACAATACCCCTATCTTCTGCA
TCAAGAAAAAGGACAAGAGCCAGTGGCGGATGCTGGTGGACTTTAGAGAGCTGAAC
AAGGCTACCCAGGACTTCTTCGAGGTGCAGCTGGGAATTCCTCATCCTGCCGGCCTG
CGGAAGATGAGACAGATCACAGTGCTGGATGTGGGCGACGCCTACTACAGCATCCC
TCTGGACCCCAACTTCAGAAAGTACACCGCCTTCACAATCCCCACCGTGAACAATCA
AGGCCCTGGCATCAGATAACAGTTCAACTGCCTGCCTCAAGGCTGGAAGGGCAGCC
CCACCATTTTTTCAGAATACCGCCGCCAGCATCCTGGAAGAAATCAAGAGAAACCTG
CCTGCTCTGACCATCGTGCAGTACATGGACGATCTGTGGGTCGGAAGCCAAGAGAA
TGAGCACACCCACGACAAGCTGGTGGAAACAGCTGAGAACAAAGCTGCAGGCCTGGG
GCCTCGAAACCCCTGAGAAGAAGGTGCAGAAAGAACCTCCTTACGAGTGGATGGGC
TACAAGCTGTGGCCTCACAAGTGGGAGCTGAGCCGGATTCAGCTCGAAGAGAAGGA
CGAGTGGACCGTGAACGACATCCAGAACTCGTGGGCAAGCTGAATTGGGCAGCCC
AGCTGTATCCCGCCTGAGGACCAAGAACATCTGCAAGCTGATCCGGGGAAAGAAG
AACCTGCTGGAACCTGGTCACATGGACACCTGAGGCCGAGGCCGAATATGCCGAGAA
TGCCGAAATCCTGAAAACCGAGCAAGAGGGGACCTACTACAAGCCTGGCATTCCAA
TCAGAGCTGCCGTGCAGAACTGGAAGGCGGCCAGTGGTCCTACCAGTTTAAGCAA
GAAGGCCAGGTCCTGAAAGTGGGCAAGTACACCAAGCAGAAGAACACCCACACCA
ACGAGCTGAGGACACTGGCTGGCCTGGTCCAGAAAATCTGCAAAGAGGCCCTGGTC
ATTTGGGGCATCCTGCCTGTTCTGGAACCTGCCATTGAGCGGGAAGTGTGGGAACAG
TGGTGGGCCGATTACTGGCAAGTGTCTTGGATCCCCGAGTGGGACTTCGTGTCTACC
CCTCCTCTGCTGAAACTGTGGTACACCCTGACAAAAGAGCCATTCTAAAGAGGGAC
GTCTACTACGTTGACGGCGCCTGCAACCGGAACTCCAAAGAAGGCAAGGCCGGCTA
CATCAGCCAGTACGGCAAGCAGAGAGTGGAAACCCTGGAAAACACCACCAACCAG
CAGGCCGAGCTGACCGCCATTAAGATGGCCCTGGAAGATAGCGGCCCCAATGTGAA
CATCGTGACCGACTCTCAGTACGCCATGGGAATCCTGACAGCCCAGCCTACACAGA
GCGATAGCCCTCTGGTTGAGCAGATCATTGCCCTGATGATTCAGAAGCAGCAAATCT

ACCTGCAGTGGGTGCCCGCTCACAAAGGCATCGGCGGAAACGAAGAGATCGATAAG
CTGGTGTCCAAGGGAATCAGACGGGTGCTGTTCCCTGGAAAAGATTGAAGAGGCCCA
AGAGGAACACGAGCGCTACCACAACAACCTGGAAGAATCTGGCCGACACCTACGGAC
TGCCCCAGATCGTGGCCAAAGAAATCGTGGCTATGTGCCCAAGTGTGAGATCAAG
GGCGAACCTGTGCACGGCCAAGTGGATGCTTCTCCTGGCACATGGCAGATGGACTG
TACCCACCTGGAAGGCAAAGTGGTCATCGTGGCTGTGCACGTGGCCTCCGGCTTTAT
TGAGGCCGAAGTGATCCCCAGAGAGACAGGCAAAGAAACCGCCAAGTTCCTGCTGA
AGATCCTGTCCAGATGGCCCATCACACAGCTGCACACCGACAACGGCCCTAACTTCA
CATCTCAAGAGGTGGCCGCCATCTGTTGGTGGGGAAAGATTGAGCACACAACCGGC
ATTCCCTACAATCCACAGAGCCAGGGCAGCATCGAGTCCATGAACAAGCAGCTCAA
AGAGATTATCGGCAAGATCCGGGACGACTGCCAGTACACAGAAACAGCCGTGCTGA
TGGCCTGTACATCCACAACCTTCAAGCGGAAAGGCGGCATCGGAGGACAGACATCT
GCCGAGAGACTGATCAATATCATCACTCAGCTGGAAATCCAGCACCTCCAGAC
CAAGATCCAGAAGATTCTGAACTTCCGGGTGTACTACCGCGAGGGCAGAGATCCTG
TTTGAAAGGCCAGCACAGCTGATCTGGAAAGGCGAAGGTGCCGTGGTGTGCTGAAG
GATGGCTCTGATCTGAAGGTGGTGGCCAGACGGAAGGCCAAGATTATCAAGGATTA
CGAGCCCAAACAGCGCGTGGGCAATGAAGGCGACGTTGAGGGCACAAGAGGCAGC
GACAATTGAAATCACTCCTCAGGTGCAGGCTGCCTATCAGAAGGTGGTGGCTGGTG
TGGCCAATGCCCTGGCTCACAAATACCACTGAGATCTTTTTCCCTCTGCCAAAATT
ATGGGGACATCATGAAGCCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAAATTTA
TTTTATTGCAATAGTGTGTTGGAATTTTTGTGTCTCTCACTCGGAAGGACATATGG
GAGGGCAAATCATTTAAAACATCAGAATGAGTATTTGGTTTAGAGTTTGGCAACATA
TGCCCATATGCTGGCTGCCATGAACAAAGGTTGGCTATAAAGAGGTCATCAGTATAT
GAAACAGCCCCCTGCTGTCCATTCTTATTCCATAGAAAAGCCTTGACTTGAGGTTA
GATTTTTTTTATATTTTGTGTTTGTGTTATTTTTTTCTTTAACATCCCTAAAATTTTCTT
ACATGTTTTACTAGCCAGATTTTTCTCCTCTCCTGACTACTCCAGTCATAGCTGTC
CCTCTTCTCTTATGGAGATCCCTCGACCTGCAGCCCAAGCTTGGCGTAATCATGGTC
ATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAAATCCACACAACATACGAGC
CGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAA
TTGCGTTGCGCTCACTGCCCCTTTCCAGTCGGGAAACCTGTGCGTGCCAGCGGATCC
GCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCT
AACTCCGCCAGTTCGGCCATTCTCCGCCCATGGCTGACTAATTTTTTTTATTTAT
GCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTT
TTTTGGAGGCCTAGGCTTTTGCAAAAAGCTAACTTGTTTATTGCAGCTTATAATGGTT
ACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATT
CTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTCCGCTTCCTCG
CTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCGAGCGGTATCAGCTCACTCA
AAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG
AGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT
TCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGG

TGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
 CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCT
 TCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAG
 GTCGTTGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
 GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCA
 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC
 AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTAT
 CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG
 CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCG
 CAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
 GTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCT
 TCACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATG
 AGTAAACTTGGTCTGACAGTTAGAAAAACTCATCGAGCATCAAATGAAACTGCAAT
 TTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAA
 GGAGAAAACTCACCGAGGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTCTGC
 GATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAG
 GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAACA
 GCTTATGCATTTCTTTCCAGACTTGTTC AACAGGCCAGCCATTACGCTCGTCATCAAA
 ATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAA
 TACGCGATCGCTGTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCA
 GGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATATTCTTCTAATA
 CCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAG
 TACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCGTCAGCCAGTTTAGTC
 TGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACA
 ACTCTGGCGCATCGGGCTTCCCATACAATCGATAGATTGTCGCACCTGATTGCCCGA
 CATTATCGCGAGCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATC
 GCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTAC
 TGTTTTATGTAAGCAGACAGTTTTTATTGTTTCATGATGATATATTTTTATCTTGTGCAAT
 GTAACATCAGAGATTTTGAGACACAACAATTGGTTCGAC

SEQ ID NO: 6: Плазмида, определенная на фигуре 2D (pDNA2b pGM299)

Длина: 3384; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..3384; тип молекулы, другая ДНК; примечание: pGM299; организм,
 синтетическая конструкция

TCAATATTGGCCATTAGCCATATTATTCATTTGGTTATATAGCATAAATCAATA
 TTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGG
 CTCATGTCCAATATGACCGCCATGTTGGCATTGATTATTGACTAGTTATTAATAGTA
 ATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGGTTACATAACT
 TACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAAT
 AATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGT
 GGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAG

TCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCGCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTGGTGGCACCAAATCAACGGGACTTTCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCA GTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGAAGTTGGTCGTGAGGCACTGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAACTGGGCTTGTCGAGACAGAGAAGACTCTTGCCTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACTATAGGCTAGCCTCGAGAATTCGATTATGCCCC TAGGACCAGAAGAAAGAAGATTGCTTCGCTTGATTTGGCTCCTTTACAGCACCAATCCATATCCACCAAGTGGGGAAGGGACGGCCAGACAACGCCGACGAGCCAGGAGAAGGTGGAGACAACAGCAGGATCAAATTAGAGTCTTGGTAGAAAGACTCCAAGAGCAGGTGTATGCAGTTGACCGCCTGGCTGACGAGGCTCAACACTTGGCTATAACAACAGTTGCTGACCCTCCTCATTAGCTTAGAATCACTAGTGAATTCACGCGTGGTACCTCTAGAGTCGACCCGGGCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACACTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAA AACCTCTACAAATGTGGTAAAATCGATAAGGATCCGTCGACCAATTGTTGTGTCTCAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATAACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGCTCTAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTA CTCTGATGATGCATGGTTACTCACC ACTGCGATCCCCGGAAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTGATGCGCTGGCAGTGTTCCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCTGCTCGCTCAGGCGCAATCACGAATGAATAACGGTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTGTTGCCATTCTCACCGGATTCAGTCGTCACCTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATAACCAGGATCTTGCCATCCTATGGAACCTCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTTTTAATTTAA AAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAAATCCCTTAACGTGA

GTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGA
 TCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGC
 GGTGGTTTGTGGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTT
 CAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCA
 CTTCAAGAACTCTGTAGCACCCGCCTACATACTCGCTCTGCTAATCCTGTTACCAGT
 GGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTT
 ACCGGATAAGGCGCAGCGGTCTGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCT
 TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGC
 GCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
 GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGT
 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG
 GGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTT
 GCTGGCCTTTTGCTCACATGGCTCGACAGATCT

SEQ ID NO: 7: Плазмида, определенная на фигуре 2Е (pDNA3a pGM301)

Длина: 6264; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..6264; тип молекулы, другая ДНК; примечание: pGM301; организм,
 синтетическая конструкция

ATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
 AGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGA
 CCGCCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCAC
 TTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGAC
 GGTAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACT
 TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAGGTGAGCCCCACG
 TTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTAT
 TTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGGGGGCGCGCGCCAG
 GCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGGCGG
 CAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGGCGG
 CGGCGGCCCTATAAAAAGCGAAGCGCGCGGGCGGGGAGTCGCTGCGCGCTGCCT
 TCGCCCCGTGCCCCGCTCCGCCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGAC
 CGCGTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTA
 GCGCTTGGTTAATGACGGCTTGTTCCTTTCTGTGGCTGCGTGAAAGCCTTGAGGG
 GCTCCGGGAGGGCCCTTTGTGCGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGGAGCGCCGCGTGC GGCTCCGCGCTGCCCGGCGGCTGTGAGCGCTGC
 GGGCGCGGCGCGGGGCTTTGTGCGCTCCGCAGTGTGCGCGAGGGGAGCGCGGGCCGG
 GGGCGGTGCCCCGCGGTGCGGGGGGGGCTGCGAGGGGAACAAGGCTGCGTGCGG
 GGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGTCGGTCCGGGCTGCAACC
 CCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCT
 CCGTACGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGG
 GGGTGCCGGGCGGGGCGGGGCCCGCCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCG

CGGCGGCCCGGAGCGCCGGCGGCTGTTCGAGGCGCGGCGAGCCGCAGCCATTGCC
TTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGG
AGCCGAAATCTGGGAGGCGCCGCCGCACCCCCTCTAGCGGGCGCGGGGGCGAAGCGG
TGCGGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCG
CCGTCCCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
GGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAG
CCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTTCTACAGCTCCTGGGCAACGTGCT
GGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGATTGCCATGGCAACATATA
TCCAGAGAGTACAGTGCATCTCAACATCACTACTGGTTGTTCTCACCACATTGGTCT
CGTGTGAGATTCCCAGGGATAGGCTCTCTAACATAGGGGTCATAGTCGATGAAGGG
AAATCACTGAAGATAGCTGGATCCCACGAATCGAGGTACATAGTACTGAGTCTAGT
TCCGGGGGTAGACTTTGAGAATGGGTGCGGAACAGCCCAGGTTATCCAGTACAAGA
GCCTACTGAACAGGCTGTTAATCCATTGAGGGATGCCTTAGATCTTCAGGAGGCTC
TGATAACTGTCACCAATGATACGACACAAAATGCCGGTGCTCCCCAGTCGAGATTCT
TCGGTGCTGTGATTGGTACTATCGCACTTGGAGTGGCGACATCAGCACAAATCACCG
CAGGGATTGCACTAGCCGAAGCGAGGGAGGCCAAAAGAGACATAGCGCTCATCAA
AGAATCGATGACAAAACACACAAGTCTATAGAAGTCTGCAAAAACGCTGTGGGGG
AACAAATTCTTGCTCTAAAGACACTCCAGGATTTTCGTGAATGATGAGATCAAACCCG
CAATAAGCGAATTAGGCTGTGAGACTGCTGCCTTAAGACTGGGTATAAAATTGACA
CAGCATTACTCCGAGCTGTTAACTGCGTTCGGCTCGAATTTCCGAACCATCGGAGAG
AAGAGCCTCACGCTGCAGGCGCTGTCTTCACTTTACTCTGCTAACATTACTGAGATT
ATGACCACAATCAGGACAGGGCAGTCTAACATCTATGATGTCATTTATACAGAACA
GATCAAAGGAACGGTGATAGATGTGGATCTAGAGAGATACATGGTCACCCTGTCTG
TGAAGATCCCTATTCTTTCTGAAGTCCCAGGTGTGCTCATAACAAGGCATCATCTA
TTTCTTACAACATAGACGGGGAGGAATGGTATGTGACTGTCCCCAGCCATATACTCA
GTCGTGCTTCTTTCTTAGGGGGTGCAGACATAACCGATTGTGTTGAGTCCAGATTGA
CCTATATATGCCCCAGGGATCCCGCACAACTGATACCTGACAGCCAGCAAAAGTGT
ATCCTGGGGGACACAACAAGGTGTCCTGTCACAAAAGTTGTGGACAGCCTTATCCCC
AAGTTTGTCTTTGTGAATGGGGGCGTTGTTGCTAACTGCATAGCATCCACATGTACC
TGCGGGACAGGCCGAAGACCAATCAGTCAGGATCGCTCTAAAGGTGTAGTATTCT
AACCCATGACAACCTGTGGTCTTATAGGTGTCAATGGGGTAGAATTGTATGCTAACCG
GAGAGGGCACGATGCCACTTGGGGGGTCCAGAACTTGACAGTCGGTCCTGCAATTG
CTATCAGACCCGTTGATATTTCTCTCAACCTTGCTGATGCTACGAATTTCTTGCAAGA
CTCTAAGGCTGAGCTTGAGAAAGCACGGAAAATCCTCTCGGAGGTAGGTAGATGGT
ACAACTCAAGAGAGACTGTGATTACGATCATAGTAGTTATGGTCGTAATATTGGTGG
TCATTATAGTGATCATCATCGTGCTTTATAGACTCAGAAGGTGAAATCACTAGTGAA
TTCCTCCTCAGGTGCAGGCTGCCTATCAGAAGGTGGTGGCTGGTGTGGCCAATGCC
CTGGCTCACAAATACCACTGAGATCTTTTTCCCTCTGCCAAAATTATGGGGACATC
ATGAAGCCCCTTGAGCATCTGACTTCTGGCTAATAAAGGAAATTTATTTTCATTGCA
ATAGTGTGTTGGAATTTTTTGTGTCTCTCACTCGGAAGGACATATGGGAGGGCAAAT

CATTTAAAACATCAGAATGAGTATTTGGTTTAGAGTTTGGCAACATATGCCCATATG
CTGGCTGCCATGAACAAAGGTTGGCTATAAAGAGGTCATCAGTATATGAAACAGCC
CCCTGCTGTCCATTCTTATTCCATAGAAAAGCCTTGACTTGAGGTTAGATTTTTTTT
ATATTTTGTGGTGTATTATTTTTCTTTAACATCCCTAAAATTTCTTACATGTTTT
ACTAGCCAGATTTTTCTCTCTCTGACTACTCCAGTCATAGCTGTCCCTCTTCTC
TTATGGAGATCCCTCGACCTGCAGCCCAAGCTTGGCGTAATCATGGTCATAGCTGTT
TCCTGTGTGAAATTGTTATCCGCTCACAATTCACACAACATACGAGCCGGAAGCAT
AAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCG
CTCACTGCCCCTTCCAGTCGGGAAACCTGTCGTGCCAGCGGATCCGCATCTCAAT
TAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCC
AGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCG
AGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCC
TAGGCTTTTGCAAAAAGCTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG
CAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGT
TTGTCCAAACTCATCAATGTATCTTATCATGTCTGTCCGCTTCTCGCTCACTGACTC
GCTGCGCTCGGTCGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAAT
ACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGC
CAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT
CCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC
CGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC
CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGT
GGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCC
AAGCTGGGCTGTGTGCACGAACCCCCGTTTCAGCCCGACCCTGCGCCTTATCCGGT
AACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC
ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT
GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA
CCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA
ACTCACGTAAAGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCC
TTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGT
CTGACAGTTAGAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAG
GATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCAC
CGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTC
CAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGA
AATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAACAGCTTATGCATTTCTT
TCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAAATCACTCGCATCAA
CCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGT
TAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAG
CGCATCAACAATATTTTACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTT

TTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAAATG
 CTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATC
 TGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATC
 GGGCTTCCCATAACAATCGATAGATTGTTCGCACCTGATTGCCCCGACATTATCGCGAGC
 CCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCA
 AGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCA
 GACAGTTTTATTGTTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGAT
 TTTGAGACACAACAATTGGTCGAC

SEQ ID NO: 8: Плазмида, определенная на фигуре 2F (pDNA3b pGM303)

Длина: 6522; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1.6522; тип молекулы, другая ДНК; примечание: pGM303; организм,
 синтетическая конструкция

ATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
 AGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGA
 CCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCAC
 TTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGAC
 GGTAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACT
 TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAGGTGAGCCCCACG
 TTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTAT
 TTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGGGGGCGCGCGCCAG
 GCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGGCGG
 CAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGGCGGGCGG
 CGGCGGCCCTATAAAAAGCGAAGCGCGCGGGCGGGCGGGAGTCGCTGCGCGCTGCCT
 TCGCCCCGTGCCCGCTCCGCGCCGCCCTCGCGCCGCCCGCCCCGGCTCTGACTGAC
 CGCGTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTA
 GCGCTTGGTTTAAATGACGGCTTGTCTTTCTTCTGTGGCTGCGTGAAAGCCTTGAGGG
 GCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGAGCGCCGCGTGC GGCTCCGCGCTGCCCGGCGGCTGTGAGCGCTGC
 GGGCGCGGCGCGGGGCTTTGTGCGCTCCGCAGTGTGCGCGAGGGGAGCGCGGCCGG
 GGGCGGTGCCCGCGGTGCGGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGG
 GGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGTCGGTCGGGCTGCAACC
 CCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCT
 CCGTACGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGG
 GGGTGCCGGGCGGGGCGGGGCCCGCCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCG
 CGGCGGCCCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGCAGCCATTGCC
 TTTTATGGTAATCGTGCGAGAGGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGG
 AGCCGAAATCTGGGAGGCGCCGCGCACCCCCCTTAGCGGGCGCGGGGCGAAGCGG
 TGCGGCGCCGCGCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCG
 CCGTCCCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGGGCAGGGCG

GGGTTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCAT
GCCTTCTTCTTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATC
ATTTTGGCAAAGAATTCCTCGAGCATGTGGTCTGAGTTAAAAATCAGGAGCAACGA
CGGAGGTGAAGGACCAGAGGACGCCAACGACCCCCGGGGAAAGGGGGTGCAACAC
ATCCATATCCAGCCATCTCTACCTGTTTATGGACAGAGGGTTAGGGATGGTGATAGG
GGCAAACGTGACTCGTACTGGTCTACTTCTCCTAGTGGTAGCACCACAAAACCAGCA
TCAGGTTGGGAGAGGTCAAGTAAAGCCGACACATGGTTGCTGATTCTCTCATTACC
CAGTGGGCTTTGTCAATTGCCACAGTGATCATCTGTATCATAATTTCTGCTAGACAA
GGGTATAGTATGAAAGAGTACTCAATGACTGTAGAGGCATTGAACATGAGCAGCAG
GGAGGTGAAAGAGTCACTTACCAGTCTAATAAGGCAAGAGGTTATAGCAAGGGCTG
TCAACATTCAGAGCTCTGTGCAAACCGGAATCCCAGTCTTGTTGAACAAAAACAGC
AGGGATGTCATCCAGATGATTGATAAGTCGTGCAGCAGACAAGAGCTCACTCAGCA
CTGTGAGAGTACGATCGCAGTCCACCATGCCGATGGAATTGCCCCACTTGAGCCACA
TAGTTTCTGGAGATGCCCTGTGCGAGAACCGTATCTTAGCTCAGATCCTGAAATCTC
ATTGCTGCCTGGTCCGAGCTTGTTATCTGGTTCTACAACGATCTCTGGATGTGTTAGG
CTCCCTTCACTCTCAATTGGCGAGGCAATCTATGCCTATTCATCAAATCTCATTACAC
AAGGTTGTGCTGACATAGGGAAATCATATCAGGTCCTGCAGCTAGGGTACATATCA
CTCAATTCAGATATGTTCCCTGATCTTAACCCCGTAGTGTCCCACTTATGACATCA
ACGACAATCGGAAATCATGCTCTGTGGTGGCAACCGGGACTAGGGGTATCAGCTTT
GCTCCATGCCGACTGTAGACGAAAGAACCGACTACTCTAGTGATGGTATTGAGGAT
CTGGTCCTTGATGTCCTGGATCTCAAAGGGGAGAATAAGTCTCACCGGTATCGCAAC
AGCGAGGTAGATCTTGATCACCCGTTCTCTGCACTATACCCCAGTGTAGGCAACGGC
ATTGCAACAGAAGGCTCATTGATATTTCTTGGGTATGGTGGACTAACCACCCTCTG
CAGGGTGATACAAAATGTAGGACCCAAGGATGCCAACAGGTGTCGCAAGACACATG
CAATGAGGCTCTGAAAATTACATGGCTAGGAGGGAAACAGGTGGTCAGCGTGATCA
TCCAGGTCAATGACTATCTCTCAGAGAGGCCAAAGATAAGAGTCACAACCATTCCA
ATCACTCAAACTATCTCGGGGCGGAAGGTAGATTATTAATAATTGGGTGATCGGGT
GTACATCTATAACAAGATCATCAGGCTGGCACTCTCAACTGCAGATAGGAGTACTTGA
TGTCAGCCACCCTTTGACTATCAACTGGACACCTCATGAAGCCTTGTCTAGACCAGG
AAATAAAGAGTGCAATTGGTACAATAAGTGTCCGAAGGAATGCATATCAGGCGTAT
ACACTGATGCTTATCCATTGTCCCCTGATGCAGCTAACGTCGCTACCGTCACGCTAT
ATGCCAATACATCGCGTGTCAACCCAACAATCATGTATTCTAACACTACTAACATTA
TAAATATGTTAAGGATAAAGGATGTTCAATTAGAGGCTGCATATAACCACGACATCGT
GTATCACGCATTTTGGTAAAGGCTACTGCTTTCACATCATCGAGATCAATCAGAAGA
GCCTGAATACCTTACAGCCGATGCTCTTTAAGACTAGCATCCCTAAATTATGCAAGG
CCGAGTCTTAAGCGGCCGCGCATGCGAATTCCTCCTCAGGTGCAGGCTGCCTATCA
GAAGGTGGTGGCTGGTGTGGCCAATGCCCTGGCTCACAAATACCACTGAGATCTTTT
TCCCTCTGCCAAAATTATGGGGACATCATGAAGCCCCTTGAGCATCTGACTTCTGG
CTAATAAAGGAAATTTATTTTCATTGCAATAGTGTGTTGGAATTTTTTGTGTCTCTCA
CTCGGAAGGACATATGGGAGGGCAAATCATTTAAAACATCAGAATGAGTATTTGGT

TTAGAGTTTGGCAACATATGCCCATATGCTGGCTGCCATGAACAAAGGTTGGCTATA
AAGAGGTCATCAGTATATGAAACAGCCCCCTGCTGTCTATTCCTTATTCCATAGAAA
AGCCTTGACTTGAGGTTAGATTTTTTTTTATATTTTGTTTTGTGTTATTTTTTTCTTTAA
CATCCCTAAAATTTTCCTTACATGTTTTACTAGCCAGATTTTTTCCTCCTCTCCTGACTA
CTCCCAGTCATAGCTGTCCCTCTTCTCTTATGGAGATCCCTCGACCTGCAGCCCAAGC
TTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTC
CACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTG
AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCTTTCCAGTCGGGAAACCTG
TCGTGCCAGCGGATCCGCATCTCAATTAGTCAGCAACCATAGTCCC GCCCTAACTC
CGCCCATCCCGCCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACT
AATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAA
GTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTAACTTGTTTATT
GCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGC
ATTTTTTTCACTGCATTCTAGTTGTGGTTTTGTCCAACTCATCAATGTATCTTATCAT
GTCTGTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCCGGCTGCGGCGAG
CGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC
GCAGGAAAGAACATGTGAGCAAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGG
CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATC
GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
CCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA
TCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT
TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAAG
ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGT
ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA
AGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT
GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGC
AAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTC
TACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGA
GATTATCAAAAAGGATCTTACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAAT
CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTAGAAAAACTCATCGAGC
ATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAA
AGCCGTTTCTGTAATGAAGGAGAAAACCTCACCGAGGCAGTTCCATAGGATGGCAAG
ATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTC
CCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATC
CGGTGAGAATGGCAACAGCTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCC
ATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTG
CGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAA
TCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAA
TCAGGATATTCTTCTAATACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGT

AACCATGCATCATCAGGAGTACGGATAAAAATGCTTGATGGTCGGAAGAGGCATAAA
 TTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTTGGCAACGCTACCT
 TTGCCATGTTTTAGAAACAACCTCTGGCGCATCGGGCTTCCCATACAATCGATAGATT
 GTCGCACCTGATTGCCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCA
 TCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTC
 ATAACACCCCTTGTATTAAGTGTATGTAAGCAGACAGTTTTATTGTTTCATGATGATA
 TATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACAATTGGTCGAC

SEQ ID NO: 9: Плазмида, определенная на фигуре 2G (pDNA2a pGM297)

Длина: 9886; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..9886; тип молекулы, другая ДНК; примечание: pGM297; организм,
 синтетическая конструкция

ATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
 AGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGA
 CCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCAC
 TTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGAC
 GGTAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACT
 TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAGGTGAGCCCCACG
 TTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTAT
 TTTTAAATATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGGGCGCGCGCCAG
 GCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGGCGG
 CAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGGCGG
 CGGCGGCCCTATAAAAAGCGAAGCGCGCGGGCGGGGCGGGAGTCGCTGCGCGCTGCCT
 TCGCCCCGTGCCCCGCTCCGCCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGAC
 CGCGTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTA
 GCGCTTGGTTTAATGACGGCTTGTTCCTTTCTGTGGCTGCGTGAAAGCCTTGAGGG
 GCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
 TGTGCGTGGGAGCGCCGCGTGC GGCTCCGCGCTGCCCGGCGGCTGTGAGCGCTGC
 GGGCGCGGCGCGGGGCTTTGTGCGCTCCGCAGTGTGCGCGAGGGGAGCGCGGCCGG
 GGGCGGTGCCCCGCGGTGCGGGGGGGGCTGCGAGGGGAACAAGGCTGCGTGCGG
 GGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGTCGGTCCGGGCTGCAACC
 CCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCT
 CCGTACGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGG
 GGGTGCCGGGCGGGGCGGGGCCGCCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCG
 CGGCGGCCCCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGCAGCCATTGCC
 TTTTATGGTAATCGTGCAGAGGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGG
 AGCCGAAATCTGGGAGGCGCCGCCACCCCCCTTAGCGGGCGCGGGGGCGAAGCGG
 TGCGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCG
 CCGTCCCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
 GGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAG

CCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTTCTACAGCTCCTGGGCAACGTGCT
GGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTGCTCGAGACTAGTGACTTGGT
GAGTAGGCTTCGAGCCTAGTTAGAGGACTAGGAGAGGCCGTAGCCGTAACACTCT
GGGCAAGTAGGGCAGGCGGTGGGTACGCAATGGGGGCGGCTACCTCAGCACTAAAT
AGGAGACAATTAGACCAATTTGAGAAAATACGACTTCGCCCCGAACGGAAAGAAAA
AGTACCAAATTAACATTTAATATGGGCAGGCAAGGAGATGGAGCGCTTCGGCCTC
CATGAGAGGTTGTTGGAGACAGAGGAGGGGTGTAAAAGAATCATAGAAGTCCTCTA
CCCCCTAGAACCAACAGGATCGGAGGGCTTAAAAAGTCTGTTCAATCTTGTGTGCGT
ACTATATTGCTTGCACAAGGAACAGAAAGTGAAAGACACAGAGGAAGCAGTAGCA
ACAGTAAGACAACACTGCCATCTAGTGGAAAAAGAAAAAAGTGCAACAGAGACAT
CTAGTGGACAAAAGAAAAATGACAAGGGAATAGCAGCGCCACCTGGTGGCAGTCA
GAATTTTCCAGCGCAACAACAAGGAAATGCCTGGGTACATGTACCCTTGTACCCGCG
CACCTTAAATGCGTGGGTAAAAGCAGTAGAGGAGAAAAAATTTGGAGCAGAAATA
GTACCCATGTTTCAAGCCCTATCAGAAGGCTGCACACCCTATGACATTAATCAGATG
CTTAATGTGCTAGGAGATCATCAAGGGGCATTACAAATAGTGAAAGAGATCATTA
TGAAGAAGCAGCCCAGTGGGATGTAACACACCCACTACCCGCAGGACCCCTACCAG
CAGGACAGCTCAGGGACCCTCGCGGCTCAGATATAGCAGGGACCACCAGCTCAGTA
CAAGAACAGTTAGAATGGATCTATACTGCTAACCCCCGGGTAGATGTAGGTGCCAT
CTACCGGAGATGGATTATTCTAGGACTTCAAAGTGTGTCAAATGTACAACCCAGT
ATCAGTCCTAGACATTAGGCAGGGACCTAAAGAGCCCTTCAAGGATTATGTGGACA
GATTTTACAAGGCAATTAGAGCAGAACAAGCCTCAGGGGAAGTGAAACAATGGATG
ACAGAATCATTACTATTCAAATGCTAATCCAGATTGTAAGGTCATCCTGAAGGGC
CTAGGAATGCACCCACCCTTGAAGAAATGTTAACGGCTTGTACAGGGGGTAGGAGG
CCCAAGCTACAAAGCAAAAGTAATGGCAGAAATGATGCAGACCATGCAAAATCAA
AACATGGTGCAGCAGGGAGGTCCAAAAAGACAAAGACCCCCACTAAGATGTTATAA
TTGTGGAAAATTTGGCCATATGCAAAGACAATGTCCGGAACCAAGGAAAACAAAAT
GTCTAAAGTGTGGAAAATTTGGGACACCTAGCAAAAGACTGCAGGGGACAGGTGAAT
TTTTTAGGGTATGGACGGTGGATGGGGGCAAAACCGAGAAATTTCCCGCCGCTACT
CTTGGAGCGGAACCGAGTGCGCCTCCTCCACCGAGCGGCACCACCCCATACGACCC
AGCAAAGAAGCTCCTGCAGCAATATGCAGAGAAAGGGAAACAACACTGAGGGAGCAA
AAGAGGAATCCACCGGCAATGAATCCGGATTGGACCGAGGGATATTCTTTGAACTC
CCTCTTTGGAGAAGACCAATAAAGACAGTGTATATAGAAGGGGTCCCCATTAAGGC
ACTGCTAGACACAGGGGCAGATGACACCATAATTAAGAAAATGATTTACAATTAT
CAGGTCCATGGAGACCCAAAATTATAGGGGGCATAGGAGGAGGCCTTAATGTAAAA
GAATATAACGACAGGGAAGTAAAAATAGAAGATAAAATTTTGAGAGGAACAATATT
GTTAGGAGCAACTCCCATTAATATAATAGGTAGAAATTTGCTGGCCCCGGCAGGTGC
CCGGTTAGTAATGGGACAATTATCAGAAAAAATTCCTGTACACCTGTCAAATTGAA
GGAAGGGGCTCGGGGACCCTGTGTAAGACAATGGCCTCTCTCTAAAGAGAAGATTG
AAGCTTTACAGGAAATATGTTCCCAATTAGAGCAGGAAGGAAAAATCAGTAGAGTA
GGAGGAGAAAATGCATACAATACCCCAATATTTTGCATAAAGAAGAAGGACAAATC

CCAGTGGAGGATGCTAGTAGACTTTAGAGAGTTAAATAAGGCAACCCAAGATTTCT
TTGAAGTGCAATTAGGGATACCCACCCAGCAGGATTAAGAAAGATGAGACAGATA
ACAGTTTTAGATGTAGGAGACGCCTATTATTCCATACCATTGGATCCAAATTTTAGG
AAATATACTGCTTTTACTATTCCCACAGTGAATAATCAGGGACCCGGGATTAGGTAT
CAATTCAACTGTCTCCCGCAAGGGTGGAAAGGATCTCCTACAATCTTCCAAAATACA
GCAGCATCCATTTTGGAGGAGATAAAAAGAACTTGCCAGCACTAACCATTGTACA
ATACATGGATGATTTATGGGTAGGTTCTCAAGAAAATGAACACACCCATGACAAAT
TAGTAGAACAGTTAAGAACAAAATTACAAGCCTGGGGCTTAGAAAACCCAGAAAAG
AAGGTGCAAAAAGAACCACCTTATGAGTGGATGGGATACAACTTTGGCCTCACAA
ATGGGAACTAAGCAGAATACAACCTGGAGGAAAAAGATGAATGGACTGTCAATGAC
ATCCAGAAGTTAGTTGGGAACTAAATTGGGCAGCACAATTGTATCCAGGTCTTAG
GACCAAGAATATATGCAAGTTAATTAGAGGAAAGAAAAATCTGTTAGAGCTAGTGA
CTTGGACACCTGAGGCAGAAGCTGAATATGCAGAAAATGCAGAGATTCTTAAAACA
GAACAGGAAGGAACCTATTACAAACCAGGAATACCTATTAGGGCAGCAGTACAGAA
ATTGGAAGGAGGACAGTGGAGTTACCAATTCAAACAAGAAGGACAAGTCTTGAAAG
TAGGAAAATACACCAAGCAAAAAGAACCCATACAAATGAACTTCGCACATTAGCT
GGTTTAGTGCAGAAGATTTGCAAAGAAGCTCTAGTTATTTGGGGGATATTACCAGTT
CTAGAACTCCCGATAGAAAGAGAGGTATGGGAACAATGGTGGGCGGATTACTGGCA
GGTAAGCTGGATTCCCGAATGGGATTTTGTGTCAGCACCCACCTTTGCTCAAACCTATG
GTACACATTAACAAAAGAACCATACCCAAGGAGGACGTTTACTATGTAGATGGAG
CATGCAACAGAAATTCAAAGAAGGAAAAGCAGGATACATCTCACAATACGGAAA
ACAGAGAGTAGAAACATTAGAAAACACTACCAATCAGCAAGCAGAATTAACAGCTA
TAAAATGGCTTTGGAAGACAGTGGGCCTAATGTGAACATAGTAACAGACTCTCAA
TATGCAATGGGAATTTTGACAGCACAACCCACACAAAGTGATTCACCATTAGTAGA
GCAAATTATAGCCTTAATGATACAAAAGCAACAAATATATTTGCAGTGGGTACCAG
CACATAAAGGAATAGGAGGAAATGAGGAGATAGATAAATTAGTGAGTAAAGGCAT
TAGAAGAGTTTTATTCTTAGAAAAAATAGAAGAAGCTCAAGAAGAGCATGAAAGAT
ATCATAATAATTGGAAAAACCTAGCAGATACATATGGGCTTCCACAAATAGTAGCA
AAAGAGATAGTGGCCATGTGTCCAAAATGTCAGATAAAGGGAGAACCAGTGCATGG
ACAAGTGGATGCCTCACCTGGAACATGGCAGATGGATTGTACTIONCTAGAAAGGAA
AAGTAGTCATAGTTGCGGTCCATGTAGCCAGTGGATTATAGAAGCAGAAGTCATA
CCTAGGGAAACAGGAAAAGAAACGGCAAAGTTTCTATTA AAAAATACTGAGTAGATG
GCCTATAACACAGTTACACACAGACAATGGGCCTAACTTTACCTCCCAAGAAGTGG
CAGCAATATGTTGGTGGGGAAAAATTGAACATACAACAGGTATACCATATAACCCC
CAATCTCAAGGATCAATAGAAAGCATGAACAAACAATTA AAAAGAGATAAATTGGGAA
AATAAGAGATGATTGCCAATATACAGAGACAGCAGTACTGATGGCTTGCCATATTC
ACAATTTTAAAAGAAAGGGAGGAATAGGGGGACAGACTTCAGCAGAGAGACTAAT
TAATATAATAACAACACAATTAGAAATACAACATTTACAAACCAAATTCAAAAAA
TTTTAAATTTTAGAGTCTACTACAGAGAAGGGAGAGACCCTGTGTGGAAAGGACCA
GCACAATTAATCTGGAAAGGGGAAGGAGCAGTGGTCCTCAAGGACGGAAGTGACCT

AAAGGTTGTACCAAGAAGGAAAGCTAAAATTATTAAGGATTATGAACCCAAACAAA
GAGTGGGTAATGAGGGTGACGTGGAAGGTACCAGGGGATCTGATAACTAAATGGCA
GGGAATAGTCAGATATTGGATGAGACAAAGAAATTTGAAATGGAECTATTATATGC
ATCAGCTGGCGGCCGCGAATTCAGTAGTGATTCCCGTTTGTGCTAGGGTTCTTAGGC
TTCTTGGGGGCTGCTGGAAGTCAATGGGAGCAGCGGCGACAGCCCTGACGGTCCA
GTCTCAGCATTGCTTGCTGGGATACTGCAGCAGCAGAAGAATCTGCTGGCGGCTGT
GGAGGCTCAACAGCAGATGTTGAAGCTGACCATTTGGGGTGTTAAAAACCTCAATG
CCCGCGTCACAGCCCTTGAGAAGTACCTAGAGGATCAGGCACGACTAAACTCCTGG
GGGTGCGCATGGAAACAAGTATGTCATAACCACAGTGGAGTGGCCCTGGACAAATCG
GACTCCGGATTGGCAAAATATGACTTGGTTGGAGTGGGAAAGACAAATAGCTGATT
TGGAAGCAACATTACGAGACAATTAGTGAAGGCTAGAGAACAAGAGGAAAAGAA
TCTAGATGCCTATCAGAAGTTAACTAGTTGGTCAGATTTCTGGTCTTGGTTCGATTT
TCAAATGGCTTAACATTTTAAAAATGGGATTTTATAGTAATAGTAGGAATAATAGGG
TTAAGATTACTTTACACAGTATATGGATGTATAGTGAGGGTTAGGCAGGGATATGTT
CCTCTATCTCCACAGATCCATATCCAATCGAATCCCGCGGCCGCAATTCCTCCTC
AGGTGCAGGCTGCCTATCAGAAGGTGGTGGCTGGTGTGGCCAATGCCCTGGCTCAC
AAATACCACTGAGATCTTTTTCCCTCTGCCAAAATTATGGGGACATCATGAAGCCC
CTTGAGCATCTGACTTCTGGCTAATAAAGGAAATTTATTTTCATTGCAATAGTGTGTT
GGAATTTTTTGTGTCTCTCACTCGGAAGGACATATGGGAGGGCAAATCATTTAAAC
ATCAGAATGAGTATTTGGTTTAGAGTTTGGCAACATATGCCCATATGCTGGCTGCCA
TGAACAAAGGTTGGCTATAAAGAGGTCATCAGTATATGAAACAGCCCCCTGCTGTC
CATTCCCTTATTCCATAGAAAAGCCTTGACTTGAGGTTAGATTTTTTTTATATTTTGT
TGTGTTATTTTTTTCTTTAACATCCCTAAAATTTTCTTACATGTTTTACTAGCCAGAT
TTTTCTCCTCTCCTGACTACTCCCAGTCATAGCTGTCCCTCTTCTCTTATGGAGATCC
CTCGACCTGCAGCCCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAA
TTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAG
CCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCG
CTTTCCAGTCGGGAAACCTGTCGTGCCAGCGGATCCGCATCTCAATTAGTCAGCAAC
CATAGTCCCGCCCCTAACTCCGCCATCCCGCCCCTAACTCCGCCCAGTTCGCCCA
TTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCG
GCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGC
AAAAAGCTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA
CAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACT
CATCAATGTATCTTATCATGTCTGTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGG
TCGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC
CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGA
CGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT
AAAGATAACCAGGCGTTTTCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCC
TGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA

TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTG
 TGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
 TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACA
 GGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT
 AACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT
 ACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
 CGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
 AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACACTCACGTT
 AAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATT
 AAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT
 AGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTTCATATCAGGATTATCAA
 TACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGT
 TCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAA
 TACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCAT
 GAGTGACGACTGAATCCGGTGAGAATGGCAACAGCTTATGCATTTCTTTCCAGACTT
 GTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGT
 TATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGAC
 AATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACA
 ATATTTTACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTTCCGGGGA
 TCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTC
 GGAAGAGGCATAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCA
 TTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCA
 TACAATCGATAGATTGTGCGCACCTGATTGCCCGACATTATCGCGAGCCATTTATAC
 CCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCC
 CGTTGAATATGGCTCATAACACCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTTA
 TTGTTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACAC
 AACAAATTGGTCGAC

SEQ ID NO: 10: Репрезентативный промотор hCEF

Длина: 574; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..574; тип молекулы, другая ДНК; примечание: промотор hCEF; организм,
 синтетическая конструкция

1 AGATCTGTTA CATAACTTAT GGTA AATGGC CTGCCTGGCT GACTGCCCAA
 TGACCCCTGC
 61 CCAATGATGT CAATAATGAT GTATGTTCCC ATGTAATGCC AATAGGGACT
 TTCCATTGAT
 121 GTCAATGGGT GGAGTATTTA TGGTAACTGC CCACTTGGCA GTACATCAAG
 TGTATCATAT
 181 GCCAAGTATG CCCCTATTG ATGTCAATGA TGGTAAATGG CCTGCCTGGC
 ATTATGCCCA
 241 GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC TATGTATTAG

TCATTGCTAT

301 TACCATGGGA ATTCACTAGT GGAGAAGAGC ATGCTTGAGG GCTGAGTGCC
CCTCAGTGGG

361 CAGAGAGCAC ATGGCCCACA GTCCCTGAGA AGTTGGGGGG AGGGGTGGGC
AATTGAACTG

421 GTGCCTAGAG AAGGTGGGGC TTGGGTAAAC TGGGAAAGTG ATGTGGTGTA
CTGGCTCCAC

481 CTTTTTCCCC AGGGTGGGGG AGAACCATAT ATAAGTGCAG TAGTCTCTGT
GAACATTCAA

541 GCTTCTGCCT TCTCCCTCCT GTGAGTTTGC TAGC

SEQ ID NO: 11: Репрезентативный промотор CMV

Длина: 873; Тип молекулы: ДНК; Особенности локализации/Определители:
источник, 1..873; тип молекулы, неустановленная ДНК; организм, человеческий
цитомегаловирус

CCGCGGAGATCTCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGC
ATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTAC
ATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGGCATTGATTATTGACTAGTT
ATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCG
TTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCAT
TGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGAC
GTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC
ATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT
ATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAG
TCATCGCTATTACCATGGTGATGCGGTTTTTGGCAGTACACCAATGGGCGTGGATAGC
GGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT
TTTGGCACCAAAATCAACGGGACTTTCCAAATGTCGTAATAACCCCGCCCCGTTGA
CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA
GTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCT
AACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGC
AGAAGTTGGTCGTGAGGCACTGGGCAGGCTAGC

SEQ ID NO: 12 Репрезентативный промотор EF1a

Длина: 395; Тип молекулы: ДНК; Особенности локализации/Определители:
источник, 1..395; тип молекулы, неустановленная ДНК; организм, Homo sapiens

AGATCCATATCCGCGGCAATTTTAAAAGAAAGGGAGGAATAGGGGGACAGA
CTTCAGCAGAGAGACTAATTAATATAATAACAACAATTAGAAATACAACATTTA
CAAACCAAAATTCAAAAAATTTTAAATTTTAGAGCCGCGGAGATCCCGTGAGGCTC
CGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCGAGAAGTTGGGGG
GAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGA
AAGTGATGTCGTGTAAGTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATAT
AAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACA

GGCTAGC

SEQ ID NO: 13: Репрезентативный трансген CFTR (soCFTR2)

Длина: 4459; Тип молекулы: ДНК; Особенности локализации/Определители: источник, 1..4459; тип молекулы, другая ДНК; примечание: soCFTR2; организм, синтетическая конструкция

1 GCTAGCCACC ATGCAGAGAA GCCCTCTGGA GAAGGCCTCT GTGGTGAGCA
AGCTGTTCTT

61 CAGCTGGACC AGGCCCATCC TGAGGAAGGG CTACAGGCAG AGACTGGAGC
TGTCTGACAT

121 CTACCAGATC CCCTCTGTGG ACTCTGCTGA CAACCTGTCT GAGAAGCTGG
AGAGGGAGTG

181 GGATAGAGAG CTGGCCAGCA AGAAGAACCC CAAGCTGATC AATGCCCTGA
GGAGATGCTT

241 CTTCTGGAGA TTCATGTTCT ATGGCATCTT CCTGTACCTG GGGGAAGTGA
CCAAGGCTGT

301 GCAGCCTCTG CTGCTGGGCA GAATCATTGC CAGCTATGAC CCTGACAACA
AGGAGGAGAG

361 GAGCATTGCC ATCTACCTGG GCATTGGCCT GTGCCTGCTG TTCATTGTGA
GGACCCTGCT

421 GCTGCACCCT GCCATCTTTG GCCTGCACCA CATTGGCATG CAGATGAGGA
TTGCCATGTT

481 CAGCCTGATC TACAAGAAAA CCCTGAAGCT GTCCAGCAGA GTGCTGGACA
AGATCAGCAT

541 TGGCCAGCTG GTGAGCCTGC TGAGCAACAA CCTGAACAAG TTTGATGAGG
GCCTGGCCCT

601 GGCCCACTTT GTGTGGATTG CCCCTCTGCA GGTGGCCCTG CTGATGGGCC
TGATTTGGGA

661 GCTGCTGCAG GCCTCTGCCT TTTGTGGCCT GGGCTTCCTG ATTGTGCTGG
CCCTGTTTCA

721 GGCTGGCCTG GGCAGGATGA TGATGAAGTA CAGGGACCAG AGGGCAGGCA
AGATCAGTGA

781 GAGGCTGGTG ATCACCTCTG AGATGATTGA GAACATCCAG TCTGTGAAGG
CCTACTGTTG

841 GGAGGAAGCT ATGGAGAAGA TGATTGAAAA CCTGAGGCAG ACAGAGCTGA
AGCTGACCAG

901 GAAGGCTGCC TATGTGAGAT ACTTCAACAG CTCTGCCTTC TTCTTCTCTG
GCTTCTTTGT

961 GGTGTTCCCTG TCTGTGCTGC CCTATGCCCT GATCAAGGGG ATCATCCTGA
GAAAGATTTT

1021 CACCACCATC AGCTTCTGCA TTGTGCTGAG GATGGCTGTG ACCAGACAGT

TCCCCTGGGC

1081 TGTGCAGACC TGGTATGACA GCCTGGGGGC CATCAACAAG ATCCAGGACT
TCCTGCAGAA

1141 GCAGGAGTAC AAGACCCTGG AGTACAACCT GACCACCACA GAAGTGGTGA
TGGAGAATGT

1201 GACAGCCTTC TGGGAGGAGG GCTTTGGGGA GCTGTTTGAG AAGGCCAAGC
AGAACAACAA

1261 CAACAGAAAG ACCAGCAATG GGGATGACTC CCTGTTCTTC TCCAACCTCT
CCCTGCTGGG

1321 CACACCTGTG CTGAAGGACA TCAACTTCAA GATTGAGAGG GGGCAGCTGC
TGGCTGTGGC

1381 TGGATCTACA GGGGCTGGCA AGACCAGCCT GCTGATGATG ATCATGGGGG
AGCTGGAGCC

1441 TTCTGAGGGC AAGATCAAGC ACTCTGGCAG GATCAGCTTT TGCAGCCAGT
TCAGCTGGAT

1501 CATGCCTGGC ACCATCAAGG AGAACATCAT CTTTGGAGTG AGCTATGATG
AGTACAGATA

1561 CAGGAGTGTG ATCAAGGCCT GCCAGCTGGA GGAGGACATC AGCAAGTTTG
CTGAGAAGGA

1621 CAACATTGTG CTGGGGGAGG GAGGCATTAC ACTGTCTGGG GGCCAGAGAG
CCAGAATCAG

1681 CCTGGCCAGG GCTGTGTACA AGGATGCTGA CCTGTACCTG CTGGACTCCC
CCTTTGGCTA

1741 CCTGGATGTG CTGACAGAGA AGGAGATTTT TGAGAGCTGT GTGTGCAAGC
TGATGGCCAA

1801 CAAGACCAGA ATCCTGGTGA CCAGCAAGAT GGAGCACCTG AAGAAGGCTG
ACAAGATCCT

1861 GATCCTGCAT GAGGGCAGCA GCTACTTCTA TGGGACCTTC TCTGAGCTGC
AGAACCTGCA

1921 GCCTGACTTC AGCTCTAAGC TGATGGGCTG TGACAGCTTT GACCAGTTCT
CTGCTGAGAG

1981 GAGGAACAGC ATCCTGACAG AGACCCTGCA CAGATTCAGC CTGGAGGGAG
ATGCCCCTGT

2041 GAGCTGGACA GAGACCAAGA AGCAGAGCTT CAAGCAGACA GGGGAGTTTG
GGGAGAAGAG

2101 GAAGAACTCC ATCCTGAACC CCATCAACAG CATCAGGAAG TTCAGCATTG
TGCAGAAAAC

2161 CCCCTGCAG ATGAATGGCA TTGAGGAAGA TTCTGATGAG CCCCTGGAGA
GGAGACTGAG

2221 CCTGGTGCCT GATTCTGAGC AGGGAGAGGC CATCCTGCCT AGGATCTCTG

TGATCAGCAC

2281 AGGCCCTACA CTGCAGGCCA GAAGGAGGCA GTCTGTGCTG AACCTGATGA
CCCACTCTGT

2341 GAACCAGGGC CAGAACATCC ACAGGAAAAC CACAGCCTCC ACCAGGAAAG
TGAGCCTGGC

2401 CCCTCAGGCC AATCTGACAG AGCTGGACAT CTACAGCAGG AGGCTGTCTC
AGGAGACAGG

2461 CCTGGAGATT TCTGAGGAGA TCAATGAGGA GGACCTGAAA GAGTGCTTCT
TTGATGACAT

2521 GGAGAGCATC CCTGCTGTGA CCACCTGGAA CACCTACCTG AGATACATCA
CAGTGCACAA

2581 GAGCCTGATC TTTGTGCTGA TCTGGTGCCT GGTGATCTTC CTGGCTGAAG
TGGCTGCCTC

2641 TCTGGTGGTG CTGTGGCTGC TGGGAAACAC CCCACTGCAG GACAAGGGCA
ACAGCACCCA

2701 CAGCAGGAAC AACAGCTATG CTGTGATCAT CACCTCCACC TCCAGCTACT
ATGTGTTCTA

2761 CATCTATGTG GGAGTGGCTG ATACCCTGCT GGCTATGGGC TTCTTTAGAG
GCCTGCCCCT

2821 GGTGCACACA CTGATCACAG TGAGCAAGAT CCTCCACCAC AAGATGCTGC
ACTCTGTGCT

2881 GCAGGCTCCT ATGAGCACCC TGAATACCCT GAAGGCTGGG GGCATCCTGA
ACAGATTCTC

2941 CAAGGATATT GCCATCCTGG ATGACCTGCT GCCTCTCACC ATCTTTGACT
TCATCCAGCT

3001 GCTGCTGATT GTGATTGGGG CCATTGCTGT GGTGGCAGTG CTGCAGCCCT
ACATCTTTGT

3061 GGCCACAGTG CCTGTGATTG TGGCCTTCAT CATGCTGAGG GCCTACTTTC
TGCAGACCTC

3121 CCAGCAGCTG AAGCAGCTGG AGTCTGAGGG CAGAAGCCCC ATCTTCACCC
ACCTGGTGAC

3181 AAGCCTGAAG GGCCTGTGGA CCCTGAGAGC CTTTGGCAGG CAGCCCTACT
TTGAGACCCT

3241 GTTCCACAAG GCCCTGAACC TGCACACAGC CAACTGGTTC CTCTACCTGT
CCACCCTGAG

3301 ATGGTTCCAG ATGAGAATTG AGATGATCTT TGTCATCTTC TTCATTGCTG
TGACCTTCAT

3361 CAGCATTCTG ACCACAGGAG AGGGAGAGGG CAGAGTGGGC ATTATCCTGA
CCCTGGCCAT

3421 GAACATCATG AGCACACTGC AGTGGGCAGT GAACAGCAGC ATTGATGTGG

ACAGCCTGAT

3481 GAGGAGTGTG AGCAGAGTGT TCAAGTTCAT TGATATGCCC ACAGAGGGCA
AGCCTACCAA

3541 GAGCACCAAG CCCTACAAGA ATGGCCAGCT GAGCAAAGTG ATGATCATTG
AGAACAGCCA

3601 TGTGAAGAAG GATGATATCT GGCCCAGTGG AGGCCAGATG ACAGTGAAGG
ACCTGACAGC

3661 CAAGTACACA GAGGGGGGCA ATGCTATCCT GGAGAACATC TCCTTCAGCA
TCTCCCCTGG

3721 CCAGAGAGTG GGA CTGCTGG GAAGAACAGG CTCTGGCAAG TCTACCCTGC
TGTCTGCCTT

3781 CCTGAGGCTG CTGAACACAG AGGGAGAGAT CCAGATTGAT GGAGTGTCTT
GGGACAGCAT

3841 CACACTGCAG CAGTGGAGGA AGGCCTTTGG TGTGATCCCC CAGAAAGTGT
TCATCTTCAG

3901 TGGCACCTTC AGGAAGAACC TGGACCCCTA TGAGCAGTGG TCTGACCAGG
AGATTTGGAA

3961 AGTGGCTGAT GAAGTGGGCC TGAGAAGTGT GATTGAGCAG TTCCTGGCA
AGCTGGACTT

4021 TGTCTGGTG GATGGGGGCT GTGTGCTGAG CCATGGCCAC AAGCAGCTGA
TGTGCCTGGC

4081 CAGATCAGTG CTGAGCAAGG CCAAGATCCT GCTGCTGGAT GAGCCTTCTG
CCCACCTGGA

4141 TCCTGTGACC TACCAGATCA TCAGGAGGAC CCTCAAGCAG GCCTTTGCTG
ACTGCACAGT

4201 CATCCTGTGT GAGCACAGGA TTGAGGCCAT GCTGGAGTGC CAGCAGTTCC
TGGTGATTGA

4261 GGAGAACAAA GTGAGGCAGT ATGACAGCAT CCAGAAGCTG CTGAATGAGA
GGAGCCTGTT

4321 CAGGCAGGCC ATCAGCCCCT CTGATAGAGT GAAGCTGTTC CCCACAGGA
ACAGCTCCAA

4381 GTGCAAGAGC AAGCCCCAGA TTGCTGCCCT GAAGGAGGAG ACAGAGGAGG
AAGTGCAGGA

4441 CACCAGGCTG TGAGGGCCC

SEQ ID NO: 14: Репрезентативный трансген А1АТ

Длина: 1257; Тип молекулы: ДНК; Особенности локализации/Определители:
источник, 1..1257; тип молекулы, другая ДНК; примечание: sohAAT; организм,
синтетическая конструкция

ATGCCAGCTCTGTGTCCTGGGGCATTCTGCTGCTGGCTGGCCTGTGCTGTCT
GGTGCCTGTGTCCCTGGCTGAGGACCCTCAGGGGGATGCTGCCAGAAAACAGACA

CCTCCCACCATGACCAGGACCACCCACCTTCAACAAGATCACCCCCAACCTGGCAG
 AGTTTGCCTTCAGCCTGTACAGACAGCTGGCCCACCAGAGCAACAGCACCAACATCT
 TTTTCAGCCCTGTGTCCATTGCCACAGCCTTTGCCATGCTGAGCCTGGGCACCAAGG
 CTGACACCCATGATGAGATCCTGGAAGGCCTGAACTTCAACCTGACAGAGATCCCT
 GAGGCCCAGATCCATGAGGGCTTCCAGGAACTGCTGAGAACCCTGAACCAGCCAGA
 CAGCCAGCTGCAGCTGACAACAGGCAATGGGCTGTTCTGTCTGAGGGCCTGAAGC
 TGGTGGACAAGTTTCTGGAAGATGTGAAGAAGCTGTACCACTCTGAGGCCTTACA
 GTGAACTTTGGGGACACAGAAGAGGCCAAGAAACAGATCAATGACTATGTGGAAA
 AGGGCACCCAGGGCAAGATTGTGGACCTTGTGAAAGAGCTGGACAGGGACACTGTG
 TTTGCCCTTGTGAACTACATCTTCTTCAAGGGCAAGTGGGAGAGGCCCTTTGAAGTG
 AAGGACACTGAGGAAGAGGACTTCCATGTGGACCAAGTGACCACAGTGAAGGTGCC
 AATGATGAAGAGACTGGGGATGTTCAATATCCAGCACTGCAAGAACTGAGCAGCT
 GGGTGCTGCTGATGAAGTACCTGGGCAATGCTACAGCCATATTCTTTCTGCCTGATG
 AGGGCAAGCTGCAGCACCTGGAAAATGAGCTGACCCATGACATCATCACAAATTT
 CTGAAAATGAGGACAGAAGATCTGCCAGCCTGCATCTGCCCAAGCTGAGCATCAC
 AGGCACATATGACCTGAAGTCTGTGCTGGGACAGCTGGGAATCACCAAGGTGTTCA
 GCAATGGGGCAGACCTGAGTGGAGTGACAGAGGAAGCCCCTCTGAAGCTGTCCAAG
 GCTGTGCACAAGGCAGTGCTGACCATTGATGAGAAGGGCACAGAGGCTGCTGGGGC
 CATGTTTCTGGAAGCCATCCCCATGTCCATCCCCCAGAAGTGAAGTTCAACAAGCC
 CTTTGTGTTCTGATGATTGAGCAGAACACCAAGAGCCCCCTGTTTCATGGGCAAGGT
 TGTGAACCCACCCAGAAATGA

SEQ ID NO: 15: Цепь, комплементарная репрезентативному трансгену A1AT

Длина: 1257; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..1257; тип молекулы, другая ДНК; примечание: комплементарная цепь
 sohAAT; организм, синтетическая конструкция

TACGGGTCGAGACACAGGACCCCGTAAGACGACGACCGACCGGACACGACA
 GACCACGGACACAGGGACCGACTCCTGGGAGTCCCCCTACGACGGGTCTTTTGTCTG
 TGGAGGGTGGTACTGGTCCTGGTGGGGTGGAAAGTTGTTCTAGTGGGGGTGGACCGT
 CTCAAACGGAAGTCGGACATGTCTGTCGACCGGGTGGTCTCGTTGTCGTGGTTGTAG
 AAAAAGTCGGGACACAGGTAACGGTGTGCGAAACGGTACGACTCGGACCCGTGGTT
 CCGACTGTGGGTACTACTCTAGGACCTTCCGGACTTGAAGTTGGACTGTCTCTAGGG
 ACTCCGGGTCTAGGTACTCCCGAAGGTCCTTGACGACTCTTGGGACTTGGTCGGTCT
 GTCGGTCGACGTCGACTGTTGTCCGTTACCCGACAAGGACAGACTCCCGGACTTCGA
 CCACCTGTTCAAAGACCTTCTACACTTCTTCGACATGGTGAGACTCCGGAAGTGTCA
 CTTGAAACCCCTGTGTCTTCTCCGGTCTTTGTCTAGTTACTGATACACCTTTTCCCGT
 GGGTCCCGTTCTAACACCTGGAACACTTTCTCGACCTGTCCCTGTGACACAAACGGG
 AACACTTGATGTAGAAGAAGTTCCCGTTCACCCTCTCCGGGAAACTTCACTTCCTGT
 GACTCCTTCTCCTGAAGGTACACCTGGTTCCTGACTGGTGTCACTTCCACGGTACTACTT
 CTCTGACCCCTACAAGTTATAGGTCGTGACGTTCTTTGACTCGTCGACCCACGACGA
 CTACTTCATGGACCCGTTACGATGTCGGTATAAGAAAGACGGACTACTCCCGTTCGA

CGTCGTGGACCTTTTACTCGACTGGGTACTGTAGTAGTGGTTTAAAGACCTTTTACTC
 CTGTCTTCTAGACGGTCGGACGTAGACGGGTTCGACTCGTAGTGTCCGTGTATACTG
 GACTTCAGACACGACCCTGTTCGACCCTTAGTGGTTCCACAAGTCGTTACCCCGTCTG
 GACTCACCTCACTGTCTCCTTCGGGGAGACTTCGACAGGTTCCGACACGTGTTCCGT
 CACGACTGGTAACTACTCTTCCCCTGTCTCCGACGACCCCGGTACAAAGACCTTCGG
 TAGGGGTACAGGTAGGGGGGTCTTCACTTCAAGTTGTTTCGGGAAACACAAGGACTA
 CТААСТСГТСТТГТГГТТСТСГГГГГГАСААГТАСССГТТССААСАСТТГГГГТГГГТ
 СТТТАСТ

SEQ ID NO: 16: Репрезентативный полипептид A1AT

Длина: 419; Тип молекулы: АА; Особенности локализации/Определители:
 источник, 1..419; тип молекулы, белок; организм, Homo sapiens

AEDPQGDAQAQKTDTSHHQDHPDFAEDPQGDAQAQKTDTSHHQDHPDFAEDPQ
 NLAFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPE
 AQIHEGFQELLRTLNPDSQLQLTTGNGLFLSEGLKLVDFLEDVKKLYHSEAFTVNFG
 DTEEAQKQINDYVEKGTQGKIVDLVKELDRDVFALVNYIFFKKGKWERPFVVKDTEED
 FHVDQVTTVKVPMKRLGMFNIQHCKKLSWVLLMKYLGNATAIFFLPDEGKLGQHLE
 NELTHDIITKFLNEDRRSASLHLPKLSITGTYDLKSVLGLGITKVFNSGADLSGVTEEA
 PLKLSKAVHKAFLTIDEKGTAAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFM
 GKVVNPTQK

SEQ ID NO: 17: Репрезентативный трансген FVIII (N6)

Длина: 5013; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..5013; тип молекулы, другая ДНК; примечание: оптимизированный по
 кодонам трансген FVIII (N6); организм, синтетическая конструкция

ATGCAGATTGAGCTGAGCACCTGCTTCTCCTGTGCCTGCTGAGGTTCTGCTT
 CTCTGCCACCAGGAGATACTACCTGGGGGCTGTGGAGCTGAGCTGGGACTACATGC
 AGTCTGACCTGGGGGAGCTGCCTGTGGATGCCAGGTCCCCCCCAGAGTGCCCAAG
 AGCTTCCCCTTCAACACCTCTGTGGTGTACAAGAAGACCCTGTTTGTGGAGTTCACT
 GACCACCTGTTCAACATTGCCAAGCCCAGGCCCCCCTGGATGGGCCTGCTGGGCCCC
 ACCATCCAGGCTGAGGTGTATGACACTGTGGTGTATCACCTGAAGAACATGGCCAG
 CCACCCTGTGAGCCTGCATGCTGTGGGGGTGAGCTACTGGAAGGCCTCTGAGGGGG
 CTGAGTATGATGACCAGACCAGCCAGAGGGAGAAGGAGGATGACAAGGTGTTCCCT
 GGGGGCAGCCACACCTATGTGTGGCAGGTGCTGAAGGAGAATGGCCCCATGGCCTC
 TGACCCCTGTGCCTGACCTACAGCTACCTGAGCCATGTGGACCTGGTGAAGGACCT
 GAACTCTGGCCTGATTGGGGCCCTGCTGGTGTGCAGGGAGGGCAGCCTGGCCAAGG
 AGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTTGTCTGTGTTTGTATGAGGGCA
 AGAGCTGGCACTCTGAAACCAAGAACAGCCTGATGCAGGACAGGGATGCTGCCTCT
 GCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCC
 TGGCCTGATTGGCTGCCACAGGAAGTCTGTGTACTGGCATGTGATTGGCATGGGCAC
 CACCCCTGAGGTGCACAGCATCTTCTGGAGGGCCACACCTTCTGGTCAGGAACCA
 CAGGCAGGCCAGCCTGGAGATCAGCCCCATCACCTTCTGACTGCCAGACCCTGCT

GATGGACCTGGGCCAGTTCCTGCTGTTCTGCCACATCAGCAGCCACCAGCATGATGG
CATGGAGGCCTATGTGAAGGTGGACAGCTGCCCTGAGGAGCCCCAGCTGAGGATGA
AGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGAT
GTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCC
AAGAAGCACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGA
CTATGCCCCCTGGTGCTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTGA
ACAATGGCCCCCAGAGGATTGGCAGGAAGTACAAGAAGGTCAGGTTCATGGCCTAC
ACTGATGAAACCTTCAAGACCAGGGAGGCCATCCAGCATGAGTCTGGCATCCTGGG
CCCCCTGCTGTATGGGGAGGTGGGGGACACCCTGCTGATCATCTTCAAGAACCAGG
CCAGCAGGCCCTACAACATCTACCCCCATGGCATCACTGATGTGAGGCCCTGTACA
GCAGGAGGCTGCCAAGGGGGTGAAGCACCTGAAGGACTTCCCCATCCTGCCTGGG
GAGATCTTCAAGTACAAGTGGACTGTGACTGTGGAGGATGGCCCCACCAAGTCTGA
CCCCAGGTGCCTGACCAGATACTACAGCAGCTTTGTGAACATGGAGAGGGACCTGG
CCTCTGGCCTGATTGGCCCCCTGCTGATCTGCTACAAGGAGTCTGTGGACCAGAGGG
GCAACCAGATCATGTCTGACAAGAGGAATGTGATCCTGTTCTCTGTGTTTGATGAGA
ACAGGAGCTGGTACCTGACTGAGAACATCCAGAGGTTCCCTGCCAACCCCTGCTGGG
GTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACAGCATCAATGG
CTATGTGTTTGACAGCCTGCAGCTGTCTGTGTGCCTGCATGAGGTGGCCTACTGGTA
CATCCTGAGCATTGGGGCCCAGACTGACTTCCTGTCTGTGTTCTTCTCTGGCTACACC
TTCAAGCACAAGATGGTGTATGAGGACACCCTGACCCTGTTCCCCTTCTCTGGGGAG
ACTGTGTTTCATGAGCATGGAGAACCCTGGCCTGTGGATTCTGGGCTGCCACAACCTCT
GACTTCAGGAACAGGGGCATGACTGCCCTGCTGAAAGTCTCCAGCTGTGACAAGAA
CACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGC
CTACCTGCTGAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGAACAGCA
GGCACCCCAGCACCAGGCAGAAGCAGTTCAATGCCACCACCATCCCTGAGAATGAC
ATAGAGAAGACAGACCCATGGTTTGCCCACCGGACCCCCATGCCCAAGATCCAGAA
TGTGAGCAGCTCTGACCTGCTGATGCTGCTGAGGCAGAGCCCCACCCCCCATGGCCT
GAGCCTGTCTGACCTGCAGGAGGCCAAGTATGAAACCTTCTCTGATGACCCCAGCCC
TGGGGCCATTGACAGCAACAACAGCCTGTCTGAGATGACCCACTTCAGGCCCCAGC
TGCACCACTCTGGGGACATGGTGTTCACCCCTGAGTCTGGCCTGCAGCTGAGGCTGA
ATGAGAAGCTGGGCACCACTGCTGCCACTGAGCTGAAGAAGCTGGACTTCAAAGTC
TCCAGCACCAGCAACAACCTGATCAGCACCATCCCCTCTGACAACCTGGCTGCTGGC
ACTGACAACACCAGCAGCCTGGGCCCCCCCCAGCATGCCTGTGCACTATGACAGCCA
GCTGGACACCACCCTGTTTGGCAAGAAGAGCAGCCCCCTGACTGAGTCTGGGGGCC
CCCTGAGCCTGTCTGAGGAGAACAATGACAGCAAGCTGCTGGAGTCTGGCCTGATG
AACAGCCAGGAGAGCAGCTGGGGCAAGAATGTGAGCAGCAGGGAGATCACCAGGA
CCACCCTGCAGTCTGACCAGGAGGAGATTGACTATGATGACACCATCTCTGTGGAG
ATGAAGAAGGAGGACTTTGACATCTACGACGAGGACGAGAACCAGAGCCCCAGGA
GCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGAGGCTGTGGGAC
TATGGCATGAGCAGCAGCCCCCATGTGCTGAGGAACAGGGCCCAGTCTGGCTCTGT

GCCCCAGTTCAAGAAGGTGGTGTTCAGGAGTTCCTGATGGCAGCTTCACCCAGCC
 CCTGTACAGAGGGGAGCTGAATGAGCACCTGGGCCTGCTGGGCCCTACATCAGGG
 CTGAGGTGGAGGACAACATCATGGTGACCTTCAGGAACCAGGCCAGCAGGCCCTAC
 AGCTTCTACAGCAGCCTGATCAGCTATGAGGAGGACCAGAGGCAGGGGGCTGAGCC
 CAGGAAGAACTTTGTGAAGCCCAATGAAACCAAGACCTACTTCTGGAAGGTGCAGC
 ACCACATGGCCCCCACCAGGATGAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTG
 ATGTGGACCTGGAGAAGGATGTGCACTCTGGCCTGATTGGCCCCCTGCTGGTGTGCC
 ACACCAACACCCTGAACCCTGCCATGGCAGGCAGGTGACTGTGCAGGAGTTTGCC
 CTGTTCTTCACCATCTTTGATGAAACCAAGAGCTGGTACTTCACTGAGAACATGGAG
 AGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAGGACCCACCTTCAAGGAGAA
 CTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGCCTGGTGAT
 GGCCCAGGACCAGAGGATCAGGTGGTACCTGCTGAGCATGGGCAGCAATGAGAACA
 TCCACAGCATCCACTTCTCTGGCCATGTGTTCACTGTGAGGAAGAAGGAGGAGTACA
 AGATGGCCCTGTACAACCTGTACCCTGGGGTGTGTTGAGACTGTGGAGATGCTGCCCA
 GCAAGGCTGGCATCTGGAGGGTGGAGTGCCTGATTGGGGAGCACCTGCATGCTGGC
 ATGAGCACCTGTTCTTGGTGTACAGCAACAAGTGCCAGACCCCCCTGGGCATGGCC
 TCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGCCAGTATGGCCAGTGGGCC
 CCCAAGCTGGCCAGGCTGCACTACTCTGGCAGCATCAATGCCTGGAGCACCAAGGA
 GCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAA
 GACCCAGGGGGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCA
 TGTACAGCCTGGATGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACTGGCACC
 CTGATGGTGTCTTTGGCAATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAAC
 CCCCCATCATTGCCAGATACATCAGGCTGCACCCACCCACT
 ACAGCATCAGGAGCACCCCTGAGGATGGAGCTGATGGGCTGTGACCTGAACAGCTGC
 AGCATGCCCTGGGCATGGAGAGCAAGGCCATCTCTGATGCCCAGATCACTGCCAG
 CAGCTACTTCACCAACATGTTTGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCT
 GCAGGGCAGGAGCAATGCCTGGAGGCCCCAGGTCAACAACCCCAAGGAGTGGCTGC
 AGGTGGACTTCCAGAAGACCATGAAGGTGACTGGGGTGACCACCCAGGGGGTGAAG
 AGCCTGCTGACCAGCATGTATGTGAAGGAGTTCCTGATCAGCAGCAGCCAGGATGG
 CCACCAGTGGACCCTGTTCTTCCAGAATGGCAAGGTGAAGGTGTTCCAGGGCAACC
 AGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCTGCTGACCAGATACC
 TGAGGATTCACCCCAGAGCTGGGTGCACCAGATTGCCCTGAGGATGGAGGTGCTG
 GGCTGTGAGGCCCAGGACCTGTAAGTGA

SEQ ID NO: 18: Репрезентативный трансген FVIII (V3)

Длина: 4425; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1..4425; тип молекулы, другая ДНК; примечание: оптимизированный по
 кононам трансген FVIII (V3); организм, синтетическая конструкция

ATGCAGATTGAGCTGAGCACCTGCTTCTTCTGCTGCTGAGGTTCTGCTT
 CTCTGCCACCAGGAGATACTACCTGGGGGCTGTGGAGCTGAGCTGGGACTACATGC
 AGTCTGACCTGGGGGAGCTGCCTGTGGATGCCAGGTTCCCCCCCAGAGTGCCCAAG

AGCTTCCCCTTCAACACCTCTGTGGTGTACAAGAAGACCCTGTTTGTGGAGTTCACT
GACCACCTGTTCAACATTGCCAAGCCCAGGCCCCCCTGGATGGGCCTGCTGGGCCCC
ACCATCCAGGCTGAGGTGTATGACACTGTGGTGTATCACCTGAAGAACATGGCCAG
CCACCCTGTGAGCCTGCATGCTGTGGGGGTGAGCTACTGGAAGGCCTCTGAGGGGG
CTGAGTATGATGACCAGACCAGCCAGAGGGAGAAGGAGGATGACAAGGTGTTCCCT
GGGGGCAGCCACACCTATGTGTGGCAGGTGCTGAAGGAGAATGGCCCCATGGCCTC
TGACCCCCTGTGCCTGACCTACAGCTACCTGAGCCATGTGGACCTGGTGAAGGACCT
GAACTCTGGCCTGATTGGGGCCCTGCTGGTGTGCAGGGAGGGCAGCCTGGCCAAGG
AGAAGACCCAGACCCTGCACAAGTTCATCCTGCTGTTTGTCTGTGTTTGTATGAGGGCA
AGAGCTGGCACTCTGAAACCAAGAACAGCCTGATGCAGGACAGGGATGCTGCCTCT
GCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCC
TGGCCTGATTGGCTGCCACAGGAAGTCTGTGTACTGGCATGTGATTGGCATGGGCAC
CACCCCTGAGGTGCACAGCATCTTCCTGGAGGGCCACACCTTCCTGGTCAGGAACCA
CAGGCAGGCCAGCCTGGAGATCAGCCCCATCACCTTCCTGACTGCCAGACCCTGCT
GATGGACCTGGGCCAGTTCCTGCTGTTCTGCCACATCAGCAGCCACCAGCATGATGG
CATGGAGGCCTATGTGAAGGTGGACAGCTGCCCTGAGGAGCCCCAGCTGAGGATGA
AGAACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGAT
GTGGTGTAGGTTTGTATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCC
AAGAAGCACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGGACTGGGA
CTATGCCCCCTGGTGTCTGGCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTGA
ACAATGGCCCCCAGAGGATTGGCAGGAAGTACAAGAAGGTCAGGTTCATGGCCTAC
ACTGATGAAACCTTCAAGACCAGGGAGGCCATCCAGCATGAGTCTGGCATCCTGGG
CCCCCTGCTGTATGGGGAGGTGGGGGACACCCTGCTGATCATCTTCAAGAACCAGG
CCAGCAGGCCCTACAACATCTACCCCCATGGCATCACTGATGTGAGGCCCTGTACA
GCAGGAGGCTGCCAAGGGGGTGAAGCACCTGAAGGACTTCCCCATCCTGCCTGGG
GAGATCTTCAAGTACAAGTGGACTGTGACTGTGGAGGATGGCCCCACCAAGTCTGA
CCCCAGGTGCCTGACCAGATACTACAGCAGCTTTGTGAACATGGAGAGGGACCTGG
CCTCTGGCCTGATTGGCCCCCTGCTGATCTGCTACAAGGAGTCTGTGGACCAGAGGG
GCAACCAGATCATGTCTGACAAGAGGAATGTGATCCTGTTCTCTGTGTTTGTATGAGA
ACAGGAGCTGGTACCTGACTGAGAACATCCAGAGGTTCCCTGCCAACCCTGCTGGG
GTGCAGCTGGAGGACCCTGAGTTCAGGCCAGCAACATCATGCACAGCATCAATGG
CTATGTGTTTGTACAGCCTGCAGCTGTCTGTGTGCCTGCATGAGGTGGCCTACTGGTA
CATCCTGAGCATTGGGGCCCAGACTGACTTCCTGTCTGTGTTCTTCTCTGGCTACACC
TTCAAGCACAAGATGGTGTATGAGGACACCCTGACCCTGTTCCCCTTCTCTGGGGAG
ACTGTGTTTATGAGCATGGAGAACCCTGGCCTGTGGATTCTGGGCTGCCACAACCTCT
GACTTCAGGAACAGGGGCATGACTGCCCTGCTGAAAGTCTCCAGCTGTGACAAGAA
CACTGGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCTACCTGCTGAGCA
AGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGAATGCCACTAATGTGTCTAAC
AACAGCAACACCAGCAATGACAGCAATGTGTCTCCCCAGTGCTGAAGAGGCACCA
GAGGGAGATCACCAGGACCACCCTGCAGTCTGACCAGGAGGAGATTGACTATGATG

ACACCATCTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTACGACGAGGACGAG
AACCAGAGCCCCAGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGT
GGAGAGGCTGTGGGACTATGGCATGAGCAGCAGCCCCATGTGCTGAGGAACAGGG
CCCAGTCTGGCTCTGTGCCCCAGTTCAAGAAGGTGGTGTTCAGGAGTTCAGTATG
GCAGCTTACCCAGCCCCTGTACAGAGGGGAGCTGAATGAGCACCTGGGCCTGCTG
GGCCCCTACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTTCAGGAACCA
GGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTATGAGGAGGACCAGA
GGCAGGGGGCTGAGCCCAGGAAGAACTTTGTGAAGCCCAATGAAACCAAGACCTAC
TTCTGGAAGGTGCAGCACCATGGCCCCACCAAGGATGAGTTTGACTGCAAGGC
CTGGGCCTACTTCTCTGATGTGGACCTGGAGAAGGATGTGCACTCTGGCCTGATTGG
CCCCCTGCTGGTGTGCCACACCAACACCCTGAACCCTGCCATGGCAGGCAGGTGAC
TGTGCAGGAGTTTGCCCTGTTCTTCACCATCTTTGATGAAACCAAGAGCTGGTACTT
CACTGAGAACATGGAGAGGAACTGCAGGGCCCCCTGCAACATCCAGATGGAGGACC
CCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCC
TGCCTGGCCTGGTGTGGCCAGGACCAGAGGATCAGGTGGTACCTGCTGAGCATG
GGCAGCAATGAGAACATCCACAGCATCCACTTCTCTGGCCATGTGTTCACTGTGAGG
AAGAAGGAGGAGTACAAGATGGCCCTGTACAACCTGTACCCTGGGGTGTTTGAGAC
TGTGGAGATGCTGCCAGCAAGGCTGGCATCTGGAGGGTGGAGTGCCTGATTGGGG
AGCACCTGCATGCTGGCATGAGCACCTGTTCTGTTGTACAGCAACAAGTGCCAG
ACCCCCCTGGGCATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACTCTGGCAGCATCAA
TGCCTGGAGCACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCA
TGATCATCCATGGCATCAAGACCCAGGGGGCCAGGCAGAAGTTCAGCAGCCTGTAC
ATCAGCCAGTTCATCATCATGTACAGCCTGGATGGCAAGAAGTGGCAGACCTACAG
GGGCAACAGCACTGGCACCCCTGATGGTGTCTTTGGCAATGTGGACAGCTCTGGCAT
CAAGCACAACATCTTCAACCCCCCATCATTGCCAGATACATCAGGCTGCACCCAC
CCACTACAGCATCAGGAGCACCCCTGAGGATGGAGCTGATGGGCTGTGACCTGAACA
GCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGATGCCAGATCACT
GCCAGCAGCTACTTCACCAACATGTTTGCCACCTGGAGCCCCAGCAAGGCCAGGCT
GCACCTGCAGGGCAGGAGCAATGCCTGGAGGGCCCCAGGTCAACAACCCCAAGGAGT
GGCTGCAGGTGGACTTCCAGAAGACCATGAAGGTGACTGGGGTGACCACCCAGGGG
GTGAAGAGCCTGCTGACCAGCATGTATGTGAAGGAGTTCCTGATCAGCAGCAGCCA
GGATGGCCACCAGTGGACCCCTGTTCTTCCAGAATGGCAAGGTGAAGGTGTTCCAGG
GCAACCAGGACAGCTTCAACCCCTGTGGTGAACAGCCTGGACCCCCCCTGCTGACCA
GATACCTGAGGATTCACCCCCAGAGCTGGGTGCACCAGATTGCCCTGAGGATGGAG
GTGCTGGGCTGTGAGGCCCAGGACCTGTAAGTGA

SEQ ID NO: 19: Цепь, комплементарная репрезентативному трансгену FVIII (N6)

Длина: 5013; Тип молекулы: ДНК; Особенности локализации/Определители: источник, 1..5013; тип молекулы, другая ДНК; примечание: комплементарная цепь

оптимизированного по кодонам трансгена FVIII (N6); организм, синтетическая конструкция

TACGTCTAACTCGACTCGTGGACGAAGAAGGACACGGACGACTCCAAGACG
AAGAGACGGTGGTCCTCTATGATGGACCCCCGACACCTCGACTCGACCCTGATGTAC
GTCAGACTGGACCCCCTCGACGGACACCTACGGTCCAAGGGGGGGTCTCACGGGTT
CTCGAAGGGGAAGTTGTGGAGACACCACATGTTCTTCTGGGACAAACACCTCAAGT
GACTGGTGGACAAGTTGTAACGGTTCGGGTCCGGGGGGACCTACCCGGACGACCCG
GGGTGGTAGGTCCGACTCCACATACTGTGACACCACTAGTGGGACTTCTTGTACCGG
TCGGTGGGACACTCGGACGTACGACACCCCCACTCGATGACCTTCCGGAGACTCCCC
CGACTCATACTACTGGTCTGGTCGGTCTCCCTCTTCCTCCTACTGTTCCACAAGGGAC
CCCCGTGGTGTGGATAACACCCGTCCACGACTTCCTCTTACCGGGGTACCGGAGAC
TGGGGGACACGGACTGGATGTTCGATGGACTCGGTACACCTGGACCACTTCCTGGAC
TTGAGACCGGACTAACCCCGGGACGACCACACGTCCCTCCCGTTCGGACCGGTTCCCTC
TTCTGGGTCTGGGACGTGTTCAAGTAGGACGACAAACGACACAAACTACTCCCGTTC
TCGACCGTGAGACTTTGGTTCTTGTTCGGACTACGTCCTGTCCCTACGACGGAGACGG
TCCCGGACCGGGTTCTACGTGTGACACTTACCGATACACTTGTCCCTCGGACGGACCG
GACTAACCGACGGTGTCCCTCAGACACATGACCGTACACTAACCGTACCCGTGGTGG
GGACTCCACGTGTCGTAGAAGGACCTCCCGGTGTGGAAGGACCAGTCCTTGGTGTCC
GTCCGGTTCGGACCTCTAGTCGGGGTAGTGGAAGGACTGACGGGTCTGGGACGACTA
CCTGGACCCGGTCAAGGACGACAAGACGGTGTAGTCGTCGGTGGTTCGTACTIONCGT
ACCTCCGATACACTTCCACCTGTCGACGGGACTCCTCGGGGTTCGACTCCTACTTCT
TGTTACTCCTCCGACTCCTGATACTACTACTGGACTGACTGAGACTCTACCTACACC
ACTCCAAACTACTACTGTTGTTCGGGGTTCGAAGTAGGTCTAGTCCAGACACCGGTTCT
TCGTGGGGTTCTGGACCCACGTGATGTAACGACGACTCCTCCTCCTGACCCTGATAC
GGGGGGACCACGACCGGGGACTACTGTCCTCGATGTTCTCGGTTCATGGACTTGTTAC
CGGGGGTCTCCTAACCGTCCCTCATGTTCTTCCAGTCCAAGTACCGGATGTGACTAC
TTTGAAGTTCTGGTCCCTCCGGTAGGTTCGTACTIONCGTACAGACCGTAGGACCCGGGGGACG
ACATACCCCTCCACCCCCTGTGGGACGACTAGTAGAAGTTCTTGGTCCGGTTCGTCGG
GGATGTTGTAGATGGGGGTACCGTAGTGACTACACTCCGGGGACATGTCGTCCTCCG
ACGGGTTCCCCACTTCGTGGACTTCTGAAGGGGTAGGACGGACCCCTCTAGAAGT
TCATGTTACCTGACACTGACACCTCCTACCGGGGTGGTTCAGACTGGGGTCCACGG
ACTGGTCTATGATGTCGTCGAAACACTTGTACCTCTCCCTGGACCGGAGACCGGACT
AACCGGGGGACGACTAGACGATGTTCCCTCAGACACCTGGTCTCCCCGTTGGTCTAGT
ACAGACTGTTCTCCTTACACTAGGACAAGAGACACAAACTACTCTTGTCCCTCGACCA
TGGACTGACTCTTGTAGGTCTCCAAGGACGGGTTGGGACGACCCACGTCGACCTCC
TGGGACTCAAGGTCCGGTTCGTTGTAGTACGTGTCGTACTIONCGGATACACAAACTGT
CGGACGTCGACAGACACACGGACGTACTCCACCGGATGACCATGTAGGACTCGTAA
CCCCGGGTCTGACTGAAGGACAGACACAAGAAGAGACCGATGTGGAAGTTCGTGTT
CTACCACATACTCCTGTGGGACTGGGACAAGGGGAAGAGACCCCTCTGACACAAGT
ACTCGTACCTCTTGGGACCGGACACCTAAGACCCGACGGTGTGAGACTGAAGTCCT

TGTCCCCGTA CTGACGGGACGACTTTCAGAGGTCGACACTGTTCTTGTGACCCCTGA
TGATACTCCTGTCGATACTCCTGTAGAGACG

GATGGACGACTCGTTCTTGTACGGTAACTCGGGTCCTCGAAGTCGGTCTTGT
CGTCCGTGGGGTCGTGGTCCGTCTTCGTCAAGTTACGGTGGTGGTAGGGACTCTTAC
TGTATCTCTTCTGTCTGGGTACCAAACGGGTGGCCTGGGGGTACGGGTCTAGGTCT
TACACTCGTCGAGACTGGACGACTACGACGACTCCGTCTCGGGGTGGGGGGTACCG
GACTCGGACAGACTGGACGTCCTCCGGTTCATACTTTGGAAGAGACTACTGGGGTCCG
GGACCCCGTA AACTGTCGTTGTTGTCGGACAGACTCTACTGGGTGAAGTCCGGGGTC
GACGTGGTGAGACCCCTGTACCACAAGTGGGGACTCAGACCCGGACGTCGACTCCGA
CTTACTCTTCGACCCGTGGTGACGACGGTGACTCGACTTCTTCGACCTGAAGTTTCA
GAGGTCGTGGTCGTTGTTGGACTAGTCGTGGTAGGGGAGACTGTTGGACCGACGAC
CGTGACTGTTGTGGTCGTCGGACCCGGGGGGTTCGTACGGACACGTGATACTGTCG
GTCGACCTGTGGTGGGACAAACCGTTCTTCTCGTCGGGGGACTGACTCAGACCCCG
GGGGACTCGGACAGACTCCTCTTGTACTGTCGTTTCGACGACCTCAGACCCGGACTAC
TTGTCGGTCCTCTCGTCGACCCCGTTCTTACACTCGTCGTCCTCTAGTGGTCCTGGT
GGGACGTCAGACTGGTCCTCCTCTAACTGATACTACTGTGGTAGAGACACCTCTACT
TCTTCTCCTGAAACTGTAGATGCTGCTCCTGCTCTTGGTCTCGGGGTCTCGAAGGT
CTTCTTCTGGTCCGTGATGAAGTAACGACGACACCTCTCCGACACCCTGATACCGTA
CTCGTCGTCGGGGGTACACGACTCCTTGTCCCGGGTTCAGACCGAGACACGGGGTCA
AGTTCTTCCACCACAAGGTCCTCAAGTGACTACCGTCGAAGTGGGTTCGGGGACATGT
CTCCCCTCGACTTACTCGTGGACCCGGACGACCCGGGGATGTAGTCCCGACTCCACC
TCCTGTTGTAGTACCACTGGAAGTCCTTGGTCCGGTCGTCCGGGATGTGGAAGATGT
CGTCGGACTAGTCGATACTCCTCCTGGTCTCCGTCCCCCGACTCGGGTCTTCTTCAA
AACTTCGGGTACTTTGGTTCTGGATGAAGACCTTCCACGTCGTGGTGTACCGGGG
GTGGTTCCTACTCAA AACTGACGTTCCGGACCCGGATGAAGAGACTACACCTGGACCT
CTTCTACACGTCGAGACCGGACTAACC GGGGGACGACCACACGGTGTGGTTGTGGG
ACTTGGGACGGGTACCGTCCGTCCACTGACACGTCCTCAAACGGGACAAGAAGTGG
TAGAACTACTTTGGTTCTCGACCATGAAGTGACTCTTGTACCTCTCCTTGACGTCCC
GGGGGACGTTGTAGGTCTACCTCCTGGGGTGGAAAGTTCCTCTTGATGTCCAAGGTAC
GGTAGTTACCGATGTAGTACCTGTGGGACGGACCCGGACCACTACCGGGTCTGGTCT
CCTAGTCCACCATGGACGACTCGTACCCGTCGTTACTCTTGTAGGTGTCGTAGGTGA
AGAGACCGGTACACAAGTGACACTCCTTCTTCTCCTCATGTTCTACCGGGACATGT
TGGACATGGGACCCCAAAACTCTGACACCTCTACGACGGGTCGTTCCGACCGTAG
ACCTCCCACCTCACGACTAACCCTCGTGGACGTACGACCGTACTCGTGGGACAAG
GACCACATGTCGTTGTTACGGTCTGGGGGGACCCGTACCGGAGACCGGTGTAGTCC
CTGAAGGTCTAGTGACGGAGACCGGTCATAACCGGTCACCCGGGGGTTTCGACCGGTC
CGACGTGATGAGACCGTCGTAGTTACGGACCTCGTGGTTCCTCGGGAAGTCGACCTA
GTTCCACCTGGACGACCGGGGGTACTAGTAGGTACCGTAGTTCTGGGTCCCCCGGTC
CGTCTTCAAGTCGTCGGACATGTAGTCGGTCAAGTAGTAGTACATGTCGGACCTACC
GTTCTTACCGTCTGGATGTCCCCGTTGTCGTGACCGTGGGACTACCACAAGAAACC

GTTACACCTGTCGAGACCGTAGTTCGTGTTGTAGAAAGTTGGGGGGGTAGTAACGGTC
TATGTAGTCCGACGTGGGGTGGGTGA

TGTCGTAGTCCCTCGTGGGACTCCTACCTCGACTACCCGACACTGGACTTGTCG
ACGTCGTACGGGGACCCGTACCTCTCGTTCCGGTAGAGACTACGGGTCTAGTGACGG
TCGTTCGATGAAGTGGTTGTACAAACGGTGGACCTCGGGGTCGTTCCGGTCCGACGTG
GACGTCCCGTCCTCGTTACGGACCTCCGGGGTCCAGTTGTTGGGGTTCCTCACCGAC
GTCCACCTGAAGGTCTTCTGGTACTTCCACTGACCCCACTGGTGGGTCCCCCACTTCT
CGGACGACTGGTCGTACATACTTCCTCAAGGACTAGTCGTCGTCGGTCCCTACCGG
TGGTCACCTGGGACAAGAAGGTCTTACCGTTCCACTTCCACAAGGTCCCGTTGGTCC
TGTCGAAGTGGGGACACCACTTGTCGGACCTGGGGGGGGACGACTGGTCTATGGAC
TCCTAAGTGGGGTCTCGACCCACGTGGTCTAACGGGACTCCTACCTCCACGACCCG
ACACTCCGGGTCCCTGGACATGACT

SEQ ID NO: 20: Цепь, комплементарная репрезентативному трансгену FVIII (V3)

Длина: 4425; Тип молекулы: ДНК; Особенности локализации/Определители: источник, 1..4425; тип молекулы, другая ДНК; примечание: комплементарная цепь оптимизированного по кодонам трансгена FVIII (V3); организм, синтетическая конструкция

TACGTCTAACTCGACTCGTGGACGAAGAAGGACACGGACGACTCCAAGACG
AAGAGACGGTGGTCCCTCTATGATGGACCCCGACACCTCGACTCGACCCTGATGTAC
GTCAGACTGGACCCCTCGACGGACACCTACGGTCCAAGGGGGGTCTCACGGGTT
CTCGAAGGGGAAGTTGTGGAGACACCACATGTTCTTCTGGGACAAACACCTCAAGT
GACTGGTGGACAAGTTGTAACGGTTCGGGTCCGGGGGGACCTACCCGGACGACCCG
GGGTGGTAGGTCCGACTCCACATACTGTGACACCACTAGTGGGACTTCTTGTACCGG
TCGGTGGGACACTCGGACGTACGACACCCCACTCGATGACCTTCCGGAGACTCCCC
CGACTCATACTACTGGTCTGGTCCGTCTCCCTCTCCTCCTACTGTTCCACAAGGGAC
CCCCGTCCGTGTGGATAACACACCGTCCACGACTTCCTCTTACCGGGGTACCGGAGAC
TGGGGGACACGGACTGGATGTCGATGGACTCGGTACACCTGGACCACTTCCTGGAC
TTGAGACCGGACTAACCCCGGGACGACCACACGTCCCTCCCGTCCGACCGGTTCCCTC
TTCTGGGTCTGGGACGTGTTCAAGTAGGACGACAAACGACACAACTACTCCCGTTC
TCGACCGTGAGACTTTGGTTCTTGTCGGACTACGTCCTGTCCCTACGACGGAGACGG
TCCCGGACCGGGTTCTACGTGTGACACTTACCGATACTTGTCCCTCGGACGGACCG
GACTAACCGACGGTGTCCCTCAGACACATGACCGTACACTAACCGTACCCGTGGTGG
GGACTCCACGTGTCGTAGAAGGACCTCCCGGTGTGGAAGGACCAGTCCTTGGTGTCC
GTCCGGTCCGACCTCTAGTCGGGGTAGTGGAAGGACTGACGGGTCTGGGACGACTA
CCTGGACCCCGTCAAGGACGACAAGACGGTGTAGTCGTCGGTGGTTCGTACTACCGT
ACCTCCGGATACTTCCACCTGTCGACGGGACTCCTCGGGGTGACTCCTACTTCT
TGTTACTCCTCCGACTCCTGATACTACTACTGGACTGACTGAGACTCTACCTACACC
ACTCCAACTACTACTGTTGTCCGGGTTCGAAGTAGGTCTAGTCCAGACACCGGTTCT
TCGTGGGGTTCTGGACCCACGTGATGTAACGACGACTCCTCCTCCTGACCCTGATAC

GGGGGGACCACGACCGGGGACTACTGTCCTCGATGTTCTCGGTCATGGACTTGTTAC
CGGGGGTCTCCTAACCGTCCTTCATGTTCTTCCAGTCCAAGTACCGGATGTGACTAC
TTTGAAGTTCTGGTCCCTCCGGTAGGTCGTA CT CAGACCGTAGGACCCGGGGGACG
ACATACCCCTCCACCCCCTGTGGGACGACTAGTAGAAGTTCTTGGTCCGGTCGTCCG
GGATGTTGTAGATGGGGGTACCGTAGTGACTACACTCCGGGGACATGTCGTCTCCTCCG
ACGGGTTCCCCCACTTCGTGGACTTCTGAAGGGGTAGGACGGACCCCTCTAGAAGT
TCATGTTACCTGACACTGACACCTCCTACCGGGGTGGTTCAGACTGGGGTCCACGG
ACTGGTCTATGATGTCGTGAAACACTTGTACCTCTCCCTGGACCGGAGACCGGACT
AACCGGGGGACGACTAGACGATGTTCTCAGACACCTGGTCTCCCCGTTGGTCTAGT
ACAGACTGTTCTCCTTACACTAGGACAAGAGACACAACTACTCTTGTCTCGACCA
TGGACTGACTCTTGTAGGTCTCCAAGGACGGGTTGGGACGACCCACGTCGACCTCC
TGGGACTCAAGGTCCGGTCGTTGTAGTACGTGTCGTAGTTACCGATACACAACTGT
CGGACGTCGACAGACACACGGACGTACTCCACCGGATGACCATGTAGGACTCGTAA
CCCCGGGTCTGACTGAAGGACAGACACAAGAAGAGACCGATGTGGAAGTTCGTGTT
CTACCACATACTCCTGTGGGACTGGGACAAGGGGAAGAGACCCCTCTGACACAAGT
ACTCGTACCTCTTGGGACCGGACACCTAAGACCCGACGGTGTGAGACTGAAGTCCT
TGTCCCCGTACTGACGGGACGACTTTCAGAGGTGACACTGTTCTTGTGACCCCTGA
TGATACTCCTGTCGATACTCCTGTAGAGACGGATGGACGACTCGTTCTTGTACGGT
AACTCGGGTCTCGAAGTCGGTCTTACGGTGATTACACAGATTGTTGTCGTTGTGGT
CGTTACTGTCGTTACACAGAGGGGGTCACGACTTCTCCGTGGTCTCCCTCTAGTGGT
CCTGGTGGGACGTCAGACTGGTCTCCTCTAACTGATACTACTGTGGTAGAGACACC
TCTACTTCTTCTCCTGAAACTGTAGATGCTGCTCCTGCTCTTGGTCTCGGGGTCTC
GAAGGTCTTCTTCTGGTCCGTGATGAAGTAACGACGACACCTCTCCGACACCCTGAT
ACCGTACTCGTCGTCGGGGGTACACGACTCCTTGTCCCGGGTCAGACCGAGACACG
GGGTCAAGTCTTCCACCACAAGGTCCTCAAGTGACTACCGTCGAAGTGGGTCCGG
GACATGTCTCCCTCGACTTACTCGTGGACCCGGACGACCCGGGGATGTAGTCCCGA
CTCCACCTCCTGTTGTAGTACCCTGGAAGTCCTTGGTCCGGTCGTCCGGGATGTGCG
AAGATGTCGTCGGACTAGTCGATACTCCTCCTGGTCTCCGTCCCCCGACTCGGGTCC
TTCTTGAACACTTCGGGTACTTTGGTCTGGATGAAGACCTTCCACGTCGTGGTGT
ACCGGGGGTGGTTCCTACTCAAACCTGACGTTCCGGACCCGGATGAAGAGACTACAC
CTGGACCTCTTCTACACGTGAGACCGGACTAACCGGGGGACGACCACACGGTGTG
GTTGTGGGACTTGGGACGGGTACCGTCCGTCCACTGACACGTCCTCAAACGGGACA
AGAAGTGGTAGAACTACTTTGGTCTCGACCATGAAGTGACTCTTGTACCTCTCCT
TGACGTCCCGGGGGACGTTGTAGGTCTACCTCCTGGGGTGGAAGTTCCTCTTGATGT
CCAAGGTACGGTAGTTACCGATGTAGTACCTGTGGGACGGACCGGCCACTACCGGG
TCCTGGTCTCCTAGTCCACCATGGACGACTCGTACCCGTGCTTACTCTTGTAGGTGTC
GTAGGTGAAGAGACCGGTACACAAGTGACACTCCTTCTTCTCCTCATGTTCTACCG
GGACATGTTGGACATGGGACCCACAACTCTGACACCTCTACGACGGGTGCTTCCG
ACCGTAGACCTCCACCTCACGGACTAACCCCTCGTGGACGTACGACCGTACTCGTG
GGACAAGGACCACATGTCGTTGTTACGGTCTGGGGGGACCCGTACCGGAGACCGG

TGTAGTCCCTGAAGGTCTAGTGACGGAGACCGGTCATACCGGTCACCCGGGGGTTTC
 GACCGGTCCGACGTGATGAGACCGTCGTAGTTACGGACCTCGTGGTTCCTCGGGAA
 GTCGACCTAGTTCCACCTGGACGACCGGGGGTACTAGTAGGTACCGTAGTTCTGGGT
 CCCCCGGTCCGTCTTCAAGTCGTCGGACATGTAGTCGGTCAAGTAGTAGTACATGTC
 GGACCTACCGTTCTTACCGTCTGGATGTCCCCGTTGTCGTGACCGTGGGACTACCA
 CAAGAAACCGTTACACCTGTCGAGACCGTAGTTCGTGTTGTAGAAGTTGGGGGGGT
 AGTAACGGTCTATGTAGTCCGACGTGGGGTGGGTGATGTCGTAGTCCTCGTGGGACT
 CCTACCTCGACTACCCGACACTGGACTTGTGACGTCGTACGGGGACCCGTACCTCT
 CGTTCGGTAGAGACTACGGGTCTAGTGACGGTTCGTGATGAAGTGGTTGTACAAA
 CGGTGGACCTCGGGGTCGTTCCGGTCCGACGTGGACGTCCCGTCCTCGTTACGGACC
 TCCGGGGTCCAGTTGTTGGGGTTCCTACCGACGTCCACCTGAAGGTCTTCTGGTAC
 TTCCACTGACCCCACTGGTGGGTCCCCACTTCTCGGACGACTGGTTCGTACATACAC
 TTCCTCAAGGACTAGTCGTCGTCGGTTCCTACCGGTGGTCACCTGGGACAAGAAGGTC
 TTACCGTTCCACTTCCACAAGGTCCCGTTGGTTCCTGTCGAAGTGGGGACACCACTTG
 TCGGACCTGGGGGGGGACGACTGGTCTATGGACTCCTAAGTGGGGGTCTCGACCCA
 CGTGGTCTAACGGGACTCCTACCTCCACGACCCGACACTCCGGGTCTGGACATGAC
 T

SEQ ID NO: 21: Репрезентативный полипептид FVIII (N6)

Длина: 1670; Тип молекулы: AA; Особенности локализации/Определители:
 источник, 1..1670; тип молекулы, белок; организм, Homo sapiens

MQIELSTCFLLRFLCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPK
 SFPFNTSVVYKKTFLVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPV
 SLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLC
 LTYSYLSHVDLVKDLNSGLIGALLVCREGLAKEKTQTLHKFILLFAVFDEGKSWHSET
 KNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPPEVHSI
 FLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSC
 PEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDNNSPSFIQIRSVAKKHPKTWVHYIA
 AEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHE
 SGILGPLLYGEVGD TLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGE
 IFKYKWTVTVEDGPTKSDPRCLTRYYSFVNMERDLASGLIGPLLYCYKESVDQRGNQI
 MSDKRNVLFSVFDENRSWYL TENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQ
 LSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKVMYEDTLTLFPFSGETVFM MENP
 GLWILGCHNSDFRNRGMTALLKVSSCDKNTGDY YEDSYEDISAYLLSKNNAIEPRFSQ
 NSRHPSTRQKQFNATTIPENDIEKTDPWFAHRT PMPKIQNVSSSDLMLLRQSPTPHGLS
 LSDLQEAKEYETFSDDPSPGAIDSNNSLSEMTHFRPQLHHS GDMVFTPESEGLQLRLNEKLG
 TTAATELKKLDFKVSSTSNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLD TTLFGK
 KSSPLTESGGPLSLSEENNSKLLSGLMNSQESSWGKNVSSREITRTTLQSDQEEIDYDD
 TISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLNRNRAQSG
 SVPQFKKVVVFQEFTDGSFTQPL YRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSF
 YSSLISYEEDQRQGAEPKRFVKNPNETKTYFWKVQH HMAPTKDEFDCKAWAYFSDVD

LEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIARYIRLHPHTHSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

SEQ ID NO: 22: Репрезентативный полипептид FVIII (V3)

Длина: 1474; Тип молекулы: AA; Особенности локализации/Определители: источник, 1..1474; тип молекулы, белок; организм, Homo sapiens

MQIELSTCFLLCCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPEFNTSVVYKKTFLVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTPVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDNNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLYGEVGDITLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYYSFVNMERDLASGLIGPLLICYKESVDQQRGNQIMSDKRNVLFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDLSQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSEMPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYEDSYEDISAYLLSKNNAIEPRFSQ NATNVSNNNSNTSNDNSVSPVVKRHQREITRTLQSDQEEIDYDDTISVEMKKEDFDIYDENQSPRSFQKTRHYFIAAVERLWDYGMSSSPHVLNRNAQSGSVPQFKKVVFEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGA EPRKNFVKPNETKTYFWKVQHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFS SLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIARYIRLHPHTHSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

SEQ ID NO: 23: Репрезентативный компонент WPRE (mWPRE)

Длина: 600; Тип молекулы: ДНК; Особенности локализации/Определители:

источник, 1..600; тип молекулы, неустановленная ДНК; организм, вирус гепатита сурка

1 GGGCCCAATC AACCTCTGGA TTACAAAATT TGTGAAAGAT TGA CTGGTAT
TCTTAACTAT
61 GTTGCTCCTT TTACGCTATG TGGATACGCT GCTTTAATGC CTTTGTATCA
TGCTATTGCT
121 TCCCGTATGG CTTTCATTTT CTCCTCCTTG TATAAATCCT GGTGCTGTC
TCTTTATGAG
181 GAGTTGTGGC CCGTTGTCAG GCAACGTGGC GTGGTGTGCA CTGTGTTTGC
TGACGCAACC
241 CCCACTGGTT GGGGCATTGC CACCACCTGT CAGCTCCTTT CCGGGACTTT
CGCTTTCCCC
301 CTCCCTATTG CCACGGCGGA ACTCATCGCC GCCTGCCTTG CCCGCTGCTG
GACAGGGGCT
361 CGGCTGTTGG GCACTGACAA TTCCGTGGTG TTGTCGGGGA AATCATCGTC
CTTCCTTGG
421 CTGCTCGCCT GTGTTGCCAC CTGGATTCTG CGCGGGACGT CCTTCTGCTA
CGTCCCTTCG
481 GCCCTCAATC CAGCGGACCT TCCTCCCGC GGCCTGCTGC CGGCTCTGCG
GCCTCTTCCG
541 CGTCTTCGCC TTCGCCCTCA GACGAGTCGG ATCTCCCTTT GGGCCGCCTC
CCC GCAAGCT

SEQ ID NO: 24: Плазмида F/HN-SIV-hCEF-soA1AT, определенная на фигуре 3 (pDNA1 pGM407)

Длина: 7349; Тип молекулы: ДНК; Особенности локализации/Определители: источник, 1..7349; тип молекулы, другая ДНК; примечание: pGM407; организм, синтетическая конструкция

1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
301 TCCCATAGTA ACGCCAATAG GGA CTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
361 AACTGCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT

421 CAATGACGGT AAATGGCCCG CCTGGCATTG TGCCCAGTAC ATGACCTTAC
GGGACTTTCC

481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA

541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT

601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA

661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGC GTGTACG
GTGGGAGGTC

721 TATATAAGCA GAGCTCGCTG GCTTGTA ACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC

781 TGGGTGTTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA

841 GAAAGCTAAT AAACCTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCTT
CATTGACGCC

901 TCACTCTCTT GAACGGGAAT CTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA

961 CTCCAGCAGT GGC GCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA

1021 GGC GTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA

1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC

1141 CGTAACTACT CTTGGGCAAG TAGGGCAGGC GGTGGGTACG CAATGGGGGC
GGCTACCTCA

1201 GCACTAAATA GGAGACAATT AGACCAATTT GAGAAAATAC GACTTCGCC
GAACGGAAAG

1261 AAAAAGTACC AAATTAACA TTTAATATGG GCAGGCAAGG AGATGGAGCG
CTTCGGCCTC

1321 CATGAGAGGT TGTTGGAGAC AGAGGAGGGG TGTA AAAAGAA TCATAGAAGT
CCTCTACCCC

1381 CTAGAACCAA CAGGATCGGA GGGCTTAAAA AGTCTGTTCA ATCTTGTGTG
CGTGCTATAT

1441 TGCTTGCACA AGGAACAGAA AGTGAAAGAC ACAGAGGAAG CAGTAGCAAC
AGTAAGACAA

1501 CACTGCCATC TAGTGGAAAA AGAAAAAAGT GCAACAGAGA CATCTAGTGG
ACAAAAGAAA

1561 AATGACAAGG GAATAGCAGC GCCACCTGGT GGCAGTCAGA ATTTTCCAGC
GCAACAACAA

1621 GGAAATGCCT GGGTACATGT ACCCTTGTC A CCGCGCACCT TAAATGCGTG
GGTAAAAGCA

1681 GTAGAGGAGA AAAAATTTGG AGCAGAAATA GTACCCATTT TTTTGTTTCA
AGCCCTATCG

1741 AATTCCC GTT TGTGCTAGGG TTCTTAGGCT TCTTGGGGGC TGCTGGAACT
GCAATGGGAG

1801 CAGCGGCGAC AGCCCTGACG GTCCAGTCTC AGCATTGCT TGCTGGGATA
CTGCAGCAGC

1861 AGAAGAATCT GCTGGCGGCT GTGGAGGCTC AACAGCAGAT GTTGAAGCTG
ACCATTTGGG

1921 GTGTAAAAA CCTCAATGCC CGCGTCACAG CCCTTGAGAA GTACCTAGAG
GATCAGGCAC

1981 GACTAAACTC CTGGGGGTGC GCATGGAAAC AAGTATGTCA TACCACAGTG
GAGTGGCCCT

2041 GGACAAATCG GACTCCGGAT TGGCAAATA TGA CTGGTT GGAGTGGGAA
AGACAAATAG

2101 CTGATTTGGA AAGCAACATT ACGAGACAAT TAGTGAAGGC TAGAGAACAA
GAGGAAAAGA

2161 ATCTAGATGC CTATCAGAAG TTA ACTAGTT GGTCAGATTT CTGGTCTTGG
TTCGATTTCT

2221 CAAAATGGCT TAACATTTTA AAAATGGGAT TTTTAGTAAT AGTAGGAATA
ATAGGGTTAA

2281 GATTACTTTA CACAGTATAT GGATGTATAG TGAGGGTTAG GCAGGGATAT
GTTCTCTAT

2341 CTCCACAGAT CCATATCCGC GGCAATTTTA AAAGAAAGGG AGGAATAGGG
GGACAGACTT

2401 CAGCAGAGAG ACTAATTAAT ATAATAACAA CACAATTAGA AATACAACAT
TTACAAACCA

2461 AAATTCAAAA AATTTTAAAT TTTAGAGCCG CGGAGATCTG TTACATAACT
TATGGTAAAT

2521 GGCCTGCCTG GCTGACTGCC CAATGACCCC TGCCCAATGA TGTCATAAAT
GATGTATGTT

2581 CCCATGTAAT GCCAATAGGG ACTTTCCATT GATGTCAATG GGTGGAGTAT
TTATGGTAAC

2641 TGCCCACTTG GCAGTACATC AAGTGTATCA TATGCCAAGT ATGCCCCCTA
TTGATGTCAA

2701 TGATGGTAAA TGGCCTGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
ACTTTCCTAC

2761 TTGGCAGTAC ATCTATGTAT TAGTCATTGC TATTACCATG GGAATTCCT
AGTGGAGAAG

2821 AGCATGCTTG AGGGCTGAGT GCCCCTCAGT GGGCAGAGAG CACATGGCCC
ACAGTCCCTG

2881 AGAAGTTGGG GGGAGGGGTG GGCAATTGAA CTGGTGCCTA GAGAAGGTGG
GGCTTGGGTA

2941 AACTGGGAAA GTGATGTGGT GTACTGGCTC CACCTTTTTC CCCAGGGTGG
GGGAGAACCA

3001 TATATAAGTG CAGTAGTCTC TGTGAACATT CAAGCTTCTG CCTTCTCCCT
CCTGTGAGTT

3061 TGCTAGCCAC CATGCCAGC TCTGTGCCT GGGGCATTCT GCTGCTGGCT
GGCCTGTGCT

3121 GTCTGGTGCC TGTGTCCCTG GCTGAGGACC CTCAGGGGGA TGCTGCCAG
AAAACAGACA

3181 CCTCCACCA TGACCAGGAC CACCCACCT TCAACAAGAT CACCCCAAC
CTGGCAGAGT

3241 TTGCCTTCAG CCTGTACAGA CAGCTGGCCC ACCAGAGCAA CAGCACCAAC
ATCTTTTCA

3301 GCCCTGTGTC CATTGCCACA GCCTTTGCCA TGCTGAGCCT GGGCACCAAG
GCTGACACCC

3361 ATGATGAGAT CCTGGAAGGC CTGAACTCA ACCTGACAGA GATCCCTGAG
GCCAGATCC

3421 ATGAGGGCTT CCAGGAACTG CTGAGAACCC TGAACCAGCC AGACAGCCAG
CTGCAGCTGA

3481 CAACAGGCAA TGGGCTGTTC CTGTCTGAGG GCCTGAAGCT GGTGGACAAG
TTTCTGGAAG

3541 ATGTGAAGAA GCTGTACCAC TCTGAGGCCT TCACAGTGAA CTTTGGGGAC
ACAGAAGAGG

3601 CCAAGAAACA GATCAATGAC TATGTGGAAA AGGGCACCCA GGGCAAGATT
GTGGACCTTG

3661 TGAAAGAGCT GGACAGGGAC ACTGTGTTTG CCCTTGTGAA CTACATCTTC
TTCAAGGGCA

3721 AGTGGGAGAG GCCCTTTGAA GTGAAGGACA CTGAGGAAGA GGA CTTCAT
GTGGACCAAG

3781 TGACCACAGT GAAGGTGCCA ATGATGAAGA GACTGGGGAT GTTCAATATC
CAGCACTGCA

3841 AGAACTGAG CAGCTGGGTG CTGCTGATGA AGTACCTGGG CAATGCTACA
GCCATATTCT

3901 TTCTGCCTGA TGAGGGCAAG CTGCAGCACC TGGAAAATGA GCTGACCCAT
GACATCATCA

3961 CCAAATTTCT GGAAAATGAG GACAGAAGAT CTGCCAGCCT GCATCTGCCC
AAGCTGAGCA

4021 TCACAGGCAC ATATGACCTG AAGTCTGTGC TGGGACAGCT GGAATCACC
AAGGTGTTCA
4081 GCAATGGGGC AGACCTGAGT GGAGTGACAG AGGAAGCCCC TCTGAAGCTG
TCCAAGGCTG
4141 TGCACAAGGC AGTGCTGACC ATTGATGAGA AGGGCACAGA GGCTGCTGGG
GCCATGTTTC
4201 TGGAAGCCAT CCCCATGTCC ATCCCCCAG AAGTGAAGTT CAACAAGCCC
TTTGTGTTCC
4261 TGATGATTGA GCAGAACACC AAGAGCCCC TGTTTCATGGG CAAGGTTGTG
AACCCACCC
4321 AGAAATGAGG GCCCAATCAA CCTCTGGATT ACAAATTTG TGAAAGATTG
ACTGGTATTC
4381 TTAATATGT TGCTCCTTTT ACGCTATGTG GATACGCTGC TTTAATGCCT
TTGTATCATG
4441 CTATTGCTTC CCGTATGGCT TTCATTTTCT CCTCCTTGTA TAAATCCTGG
TTGCTGTCTC
4501 TTTATGAGGA GTTGTGGCCC GTTGTCAGGC AACGTGGCGT GGTGTGCACT
GTGTTTGCTG
4561 ACGCAACCCC CACTGGTTGG GGCATTGCCA CCACCTGTCA GCTCCTTCC
GGGACTTTCG
4621 CTTTCCCCCT CCCTATTGCC ACGGCGGAAC TCATCGCCGC CTGCCTTGCC
CGCTGCTGGA
4681 CAGGGGCTCG GCTGTTGGGC ACTGACAATT CCGTGGTGTT GTCGGGGAAA
TCATCGTCCT
4741 TTCCTTGGCT GCTCGCCTGT GTTGCCACCT GGATTCTGCG CGGGACGTCC
TTCTGCTACG
4801 TCCCTTCGGC CCTCAATCCA GCGGACCTTC CTTCGCGG CCTGCTGCCG
GCTCTGCGGC
4861 CTCTCCGCG TCTTCGCCTT CGCCTCAGA CGAGTCGGAT CTCCCTTTGG
GCCGCCTCCC
4921 CGCAAGCTTC GCACTTTTTA AAAGAAAAGG GAGGACTGGA TGGGATTTAT
TACTCCGATA
4981 GGACGCTGGC TTGTAATCA GTCTCTTACT AGGAGACCAG CTTGAGCCTG
GGTGTTCGCT
5041 GGTTAGCCTA ACCTGGTTGG CCACCAGGGG TAAGGACTCC TTGGCTTAGA
AAGCTAATAA
5101 ACTTGCTGC ATTAGAGCTC TTACGCGTCC CGGGCTCGAG ATCCGCATCT
CAATTAGTCA
5161 GCAACCATAG TCCCGCCCCT AACTCCGCC ATCCCGCCCC TAACTCCGCC
CAGTTCCGCC

5221 CATTCTCCGC CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA
GGCCGCCTCG

5281 GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG
CTTTTGCAAA

5341 AAGCTAACTT GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC
ATCACAAATT

5401 TCACAAATAA AGCATTTTTT TCACTGCATT CTAGTTGTGG TTTGTCCAAA
CTCATCAATG

5461 TATCTTATCA TGTCTGTCCG CTCCTCGCT CACTGACTCG CTGCGCTCGG
TCGTTCCGGT

5521 GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG
AATCAGGGGA

5581 TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC
GTAAAAAGGC

5641 CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCTGAC GAGCATCACA
AAAATCGACG

5701 CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT
TTCCCCCTGG

5761 AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATAACC
TGTCCGCCTT

5821 TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC
TCAGTTCGGT

5881 GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCG TTCAGC
CCGACCGCTG

5941 CGCCTTATCC GGTA ACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT
TATCGCCACT

6001 GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG
CTACAGAGTT

6061 CTTGAAGTGG TGGCCTAACT ACGGCTACAC TAGAAGAACA GTATTTGGTA
TCTGCGCTCT

6121 GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA
AACAACCAC

6181 CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA
AAAAAGGATC

6241 TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG
AAA ACTCACG

6301 TTAAGGGATT TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC
TTTTAAATTA

6361 AAAATGAAGT TTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG
ACAGTTAGAA

6421 AAACATCATCG AGCATCAAAT GAAACTGCAA TTTATTCATA TCAGGATTAT
 CAATACCATA
 6481 TTTTTGAAAA AGCCGTTTCT GTAATGAAGG AGAAAACTCA CCGAGGCAGT
 TCCATAGGAT
 6541 GGCAAGATCC TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC
 AACCTATTAA
 6601 TTTCCCCTCG TCAAAAATAA GGTATCAAG TGAGAAATCA CCATGAGTGA
 CGACTGAATC
 6661 CGGTGAGAAT GGCAACAGCT TATGCATTTT TTTCCAGACT TGTTC AACAG
 GCCAGCCATT
 6721 ACGCTCGTCA TCAAAATCAC TCGCATCAAC CAAACCGTTA TTCATTCGTG
 ATTGCGCCTG
 6781 AGCGAGACGA AATACGCGAT CGCTGTAAA AGGACAATTA CAAACAGGAA
 TCGAATGCAA
 6841 CCGGCGCAGG AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG
 GATATTCTTC
 6901 TAATACCTGG AATGCTGTTT TTCCGGGGAT CGCAGTGGTG AGTAACCATG
 CATCATCAGG
 6961 AGTACGGATA AAATGCTTGA TGGTCGGAAG AGGCATAAAT TCCGTCAGCC
 AGTTTAGTCT
 7021 GACCATCTCA TCTGTAACAT CATTGGCAAC GCTACCTTTG CCATGTTTCA
 GAAACAATC
 7081 TGGCGCATCG GGCTTCCCAT ACAATCGATA GATTGTCGCA CCTGATTGCC
 CGACATTATC
 7141 GCGAGCCCAT TTATACCCAT ATAAATCAGC ATCCATGTTG GAATTTAATC
 GCGGCCTAGA
 7201 GCAAGACGTT TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT
 TTATGTAAGC
 7261 AGACAGTTTT ATTGTTTCATG ATGATATATT TTTATCTTGT GCAATGTAAC
 ATCAGAGATT
 7321 TTGAGACACA ACAATTGGTC GACGGATCC

SEQ ID NO: 25: Плазида F/HN-SIV-CMV-HFVIII-V3, определенная на фигуре 4A (pDNA1 pGM411)

Длина: 10812; Тип молекулы: ДНК; Особенности локализации/Определители:
 источник, 1.. 10812; тип молекулы, другая ДНК; примечание: pGM411; организм,
 синтетическая конструкция

1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
 AAATCAATAT
 61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
 ATATTGGCTC

121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT

181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA

241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCATTG ACGTCAATAA
TGACGTATGT

301 TCCCATAGTA ACGCCAATAG GGACTIONTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA

361 AACTGCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT

421 CAATGACGGT AAATGGCCCC CCTGGCATTG TGCCCAGTAC ATGACCTTAC
GGGACTTTC

481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA

541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT

601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA

661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGC GTGTACG
GTGGGAGGTC

721 TATATAAGCA GAGCTCGCTG GCTTGTA ACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC

781 TGGGTGTTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA

841 GAAAGCTAAT AAACCTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCCT
CATTGACGCC

901 TCACTCTCTT GAACGGGAAT CTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA

961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA

1021 GGC GTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA

1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC

1141 CGTAACTACT CTGGGCAAGT AGGGCAGGCG GTGGGTACGC AATGGGGGCG
GCTACCTCAG

1201 CACTAAATAG GAGACAATTA GACCAATTTG AGAAAATACG ACTTCGCCCCG
AACGGAAAGA

1261 AAAAGTACCA AATTAAACAT TTAATATGGG CAGGCAAGGA GATGGAGCGC
TTCGGCCTCC

1321 ATGAGAGGTT GTTGGAGACA GAGGAGGGGT GTAAAAGAAT CATAGAAGTC
CTCTACCCCC

1381 TAGAACCAAC AGGATCGGAG GGCTTAAAAA GTCTGTTCAA TCTTGTGTGC
GTGCTATATT

1441 GCTTGCACAA GGAACAGAAA GTGAAAGACA CAGAGGAAGC AGTAGCAACA
GTAAGACAAC

1501 ACTGCCATCT AGTGGAAAAA GAAAAAAGTG CAACAGAGAC ATCTAGTGGA
CAAAAGAAAA

1561 ATGACAAGGG AATAGCAGCG CCACCTGGTG GCAGTCAGAA TTTTCCAGCG
CAACAACAAG

1621 GAAATGCCTG GGTACATGTA CCCTTGTCAC CGCGCACCTT AAATGCGTGG
GTAAAAGCAG

1681 TAGAGGAGAA AAAATTTGGA GCAGAAATAG TACCCATGTT TCAAGCCCTA
TCGAATTCCC

1741 GTTTGTGCTA GGGTTCTTAG GCTTCTTGGG GGCTGCTGGA ACTGCAATGG
GAGCAGCGGC

1801 GACAGCCCTG ACGGTCCAGT CTCAGCATTT GCTTGCTGGG ATACTGCAGC
AGCAGAAGAA

1861 TCTGCTGGCG GCTGTGGAGG CTCAACAGCA GATGTTGAAG CTGACCATTT
GGGGTGTTAA

1921 AAACCTCAAT GCCCGCGTCA CAGCCCTTGA GAAGTACCTA GAGGATCAGG
CACGACTAAA

1981 CTCCTGGGGG TGCGCATGGA AACAAGTATG TCATACCACA GTGGAGTGGC
CCTGGACAAA

2041 TCGGACTCCG GATTGGCAAA ATATGACTTG GTTGGAGTGG GAAAGACAAA
TAGCTGATTT

2101 GGAAAGCAAC ATTACGAGAC AATTAGTGAA GGCTAGAGAA CAAGAGGAAA
AGAATCTAGA

2161 TGCCTATCAG AAGTTAACTA GTTGGTCAGA TTTCTGGTCT TGGTTCGATT
TCTCAAATG

2221 GCTTAACATT TTAAAAATGG GATTTTTAGT AATAGTAGGA ATAATAGGGT
TAAGATTACT

2281 TTACACAGTA TATGGATGTA TAGTGAGGGT TAGGCAGGGA TATGTTCCCTC
TATCTCCACA

2341 GATCCATATC CGCGGCAATT TTAAAAGAAA GGGAGGAATA GGGGGACAGA
CTTCAGCAGA

2401 GAGACTAATT AATATAATAA CAACACAATT AGAAATACAA CATTTACAAA
CCAAAATTCA

2461 AAAAATTTTA AATTTTAGAG CCGCGGAGAT CTCAATATTG GCCATTAGCC
ATATTATTCA

2521 TTGGTTATAT AGCATAAATC AATATTGGCT ATTGGCCATT GCATACGTTG
TATCTATATC
2581 ATAATATGTA CATTTATATT GGCTCATGTC CAATATGACC GCCATGTTGG
CATTGATTAT
2641 TGACTIONAGTTA TTAATAGTAA TCAATTACGG GGTCATTAGT TCATAGCCCA
TATATGGAGT
2701 TCCGCGTTAC ATAACCTACG GTAAATGGCC CGCCTGGCTG ACCGCCCAAC
GACCCCCGCC
2761 CATTGACGTC AATAATGACG TATGTTCCCA TAGTAACGCC AATAGGGACT
TTCCATTGAC
2821 GTCAATGGGT GGAGTATTTA CGGTAAACTG CCCACTTGGC AGTACATCAA
GTGTATCATA
2881 TGCCAAGTCC GCCCCCTATT GACGTCAATG ACGGTAAATG GCCCGCCTGG
CATTATGCCC
2941 AGTACATGAC CTTACGGGAC TTTCTACTT GGCAGTACAT CTACGTATTA
GTCATCGCTA
3001 TTACCATGGT GATGCGGTTT TGGCAGTACA CCAATGGGCG TGGATAGCGG
TTTGACTIONAC
3061 GGGGATTTCC AAGTCTCCAC CCCATTGACG TCAATGGGAG TTTGTTTTGG
CACCAAAATC
3121 AACGGGACTT TCCAAAATGT CGTAATAACC CCGCCCCGTT GACGCAAATG
GGCGGTAGGC
3181 GTGTACGGTG GGAGGTCTAT ATAAGCAGAG CTCGTTTAGT GAACCGTCAG
ATCACTIONAGAA
3241 GCTTTATTGC GGTAGTTTAT CACAGTTAAA TTGCTAACGC AGTCAGTGCT
TCTGACACAA
3301 CAGTCTCGAA CTTAAGCTGC AGAAGTTGGT CGTGAGGCAC TGGGCAGGCT
AGCCACCAAT
3361 GCAGATTGAG CTGAGCACCT GCTTCTTCCT GTGCCTGCTG AGGTTCTGCT
TCTCTGCCAC
3421 CAGGAGATAC TACCTGGGGG CTGTGGAGCT GAGCTGGGAC TACATGCAGT
CTGACCTGGG
3481 GGAGCTGCCT GTGGATGCCA GGTTCCCCC CAGAGTGCCC AAGAGCTTCC
CCTTCAACAC
3541 CTCTGTGGTG TACAAGAAGA CCCTGTTTGT GGAGTTCACT GACCACCTGT
TCAACATTGC
3601 CAAGCCCAGG CCCCCCTGGA TGGGCCTGCT GGGCCCCACC ATCCAGGCTG
AGGTGTATGA
3661 CACTGTGGTG ATCACCTGA AGAACATGGC CAGCCACCCT GTGAGCCTGC
ATGCTGTGGG

3721 GGTGAGCTAC TGGAAGGCCT CTGAGGGGGC TGAGTATGAT GACCAGACCA
GCCAGAGGGA

3781 GAAGGAGGAT GACAAGGTGT TCCCTGGGGG CAGCCACACC TATGTGTGGC
AGGTGCTGAA

3841 GGAGAATGGC CCCATGGCCT CTGACCCCCT GTGCCTGACC TACAGCTACC
TGAGCCATGT

3901 GGACCTGGTG AAGGACCTGA ACTCTGGCCT GATTGGGGCC CTGCTGGTGT
GCAGGGAGGG

3961 CAGCCTGGCC AAGGAGAAGA CCCAGACCCT GCACAAGTTC ATCCTGCTGT
TTGCTGTGTT

4021 TGATGAGGGC AAGAGCTGGC ACTCTGAAAC CAAGAACAGC CTGATGCAGG
ACAGGGATGC

4081 TGCCTCTGCC AGGGCCTGGC CCAAGATGCA CACTGTGAAT GGCTATGTGA
ACAGGAGCCT

4141 GCCTGGCCTG ATTGGCTGCC ACAGGAAGTC TGTGTACTGG CATGTGATTG
GCATGGGCAC

4201 CACCCCTGAG GTGCACAGCA TCTTCCTGGA GGGCCACACC TTCCTGGTCA
GGAACCACAG

4261 GCAGGCCAGC CTGGAGATCA GCCCCATCAC CTCCTGACT GCCCAGACCC
TGCTGATGGA

4321 CCTGGGCCAG TTCCTGCTGT TCTGCCACAT CAGCAGCCAC CAGCATGATG
GCATGGAGGC

4381 CTATGTGAAG GTGGACAGCT GCCCTGAGGA GCCCCAGCTG AGGATGAAGA
ACAATGAGGA

4441 GGCTGAGGAC TATGATGATG ACCTGACTGA CTCTGAGATG GATGTGGTGA
GGTTTGATGA

4501 TGACAACAGC CCCAGCTTCA TCCAGATCAG GTCTGTGGCC AAGAAGCACC
CCAAGACCTG

4561 GGTGCACTAC ATTGCTGCTG AGGAGGAGGA CTGGGACTAT GCCCCCCTGG
TGCTGGCCCC

4621 TGATGACAGG AGCTACAAGA GCCAGTACCT GAACAATGGC CCCCAGAGGA
TTGGCAGGAA

4681 GTACAAGAAG GTCAGGTTCA TGGCCTACAC TGATGAAACC TTCAAGACCA
GGGAGGCCAT

4741 CCAGCATGAG TCTGGCATCC TGGGCCCCCT GCTGTATGGG GAGGTGGGGG
ACACCCTGCT

4801 GATCATCTTC AAGAACCAGG CCAGCAGGCC CTACAACATC TACCCCATG
GCATCACTGA

4861 TGTGAGGCCC CTGTACAGCA GGAGGCTGCC CAAGGGGGTG AAGCACCTGA
AGGACTTCCC

4921 CATCCTGCCT GGGGAGATCT TCAAGTACAA GTGGACTGTG ACTGTGGAGG
ATGGCCCCAC

4981 CAAGTCTGAC CCCAGGTGCC TGACCAGATA CTACAGCAGC TTTGTGAACA
TGGAGAGGGA

5041 CCTGGCCTCT GGCCTGATTG GCCCCCTGCT GATCTGCTAC AAGGAGTCTG
TGGACCAGAG

5101 GGGCAACCAG ATCATGTCTG ACAAGAGGAA TGTGATCCTG TTCTCTGTGT
TTGATGAGAA

5161 CAGGAGCTGG TACCTGACTG AGAACATCCA GAGGTTCCCTG CCCAACCCCTG
CTGGGGTGCA

5221 GCTGGAGGAC CCTGAGTTCC AGGCCAGCAA CATCATGCAC AGCATCAATG
GCTATGTGTT

5281 TGACAGCCTG CAGCTGTCTG TGTGCCTGCA TGAGGTGGCC TACTGGTACA
TCCTGAGCAT

5341 TGGGGCCCAG ACTGACTTCC TGTCTGTGTT CTTCTCTGGC TACACCTTCA
AGCACAAGAT

5401 GGTGTATGAG GACACCCTGA CCCTGTTCCC CTTCTCTGGG GAGACTGTGT
TCATGAGCAT

5461 GGAGAACCCT GGCCTGTGGA TTCTGGGCTG CCACAACCTCT GACTTCAGGA
ACAGGGGCAT

5521 GACTGCCCTG CTGAAAGTCT CCAGCTGTGA CAAGAACACT GGGGACTACT
ATGAGGACAG

5581 CTATGAGGAC ATCTCTGCCT ACCTGCTGAG CAAGAACAAT GCCATTGAGC
CCAGGAGCTT

5641 CAGCCAGAAT GCCACTAATG TGTCTAACAA CAGCAACACC AGCAATGACA
GCAATGTGTC

5701 TCCCCCAGTG CTGAAGAGGC ACCAGAGGGA GATCACCAGG ACCACCCTGC
AGTCTGACCA

5761 GGAGGAGATT GACTATGATG ACACCATCTC TGTGGAGATG AAGAAGGAGG
ACTTTGACAT

5821 CTACGACGAG GACGAGAACC AGAGCCCCAG GAGCTTCCAG AAGAAGACCA
GGCACTACTT

5881 CATTGCTGCT GTGGAGAGGC TGTGGGACTA TGGCATGAGC AGCAGCCCCC
ATGTGCTGAG

5941 GAACAGGGCC CAGTCTGGCT CTGTGCCCCA GTTCAAGAAG GTGGTGTTC
AGGAGTTCAC

6001 TGATGGCAGC TTCACCCAGC CCCTGTACAG AGGGGAGCTG AATGAGCACC
TGGGCCTGCT

6061 GGGCCCCTAC ATCAGGGCTG AGGTGGAGGA CAACATCATG GTGACCTTCA
GGAACCAGGC

6121 CAGCAGGCCC TACAGCTTCT ACAGCAGCCT GATCAGCTAT GAGGAGGACC
AGAGGCAGGG

6181 GGCTGAGCCC AGGAAGAACT TTGTGAAGCC CAATGAAACC AAGACCTACT
TCTGGAAGGT

6241 GCAGCACCAC ATGGCCCCCA CCAAGGATGA GTTTGACTGC AAGGCCTGGG
CCTACTTCTC

6301 TGATGTGGAC CTGGAGAAGG ATGTGCACTC TGGCCTGATT GGCCCCCTGC
TGGTGTGCCA

6361 CACCAACACC CTGAACCCTG CCCATGGCAG GCAGGTGACT GTGCAGGAGT
TTGCCCTGTT

6421 CTTACCATC TTTGATGAAA CCAAGAGCTG GTACTTCACT GAGAACATGG
AGAGGAACTG

6481 CAGGGCCCCC TGCAACATCC AGATGGAGGA CCCACCTTC AAGGAGAACT
ACAGGTTCCA

6541 TGCCATCAAT GGCTACATCA TGGACACCCT GCCTGGCCTG GTGATGGCCC
AGGACCAGAG

6601 GATCAGGTGG TACCTGCTGA GCATGGGCAG CAATGAGAAC ATCCACAGCA
TCCACTTCTC

6661 TGGCCATGTG TTCACTGTGA GGAAGAAGGA GGAGTACAAG ATGGCCCTGT
ACAACCTGTA

6721 CCCTGGGGTG TTTGAGACTG TGGAGATGCT GCCCAGCAAG GCTGGCATCT
GGAGGGTGGA

6781 GTGCCTGATT GGGGAGCACC TGCATGCTGG CATGAGCACC CTGTTCTCTGG
TGTACAGCAA

6841 CAAGTGCCAG ACCCCCCTGG GCATGGCCTC TGGCCACATC AGGGACTTCC
AGATCACTGC

6901 CTCTGGCCAG TATGGCCAGT GGGCCCCCAA GCTGGCCAGG CTGCACTACT
CTGGCAGCAT

6961 CAATGCCTGG AGCACCAAGG AGCCCTTCAG CTGGATCAAG GTGGACCTGC
TGGCCCCCAT

7021 GATCATCCAT GGCATCAAGA CCCAGGGGGC CAGGCAGAAG TTCAGCAGCC
TGTACATCAG

7081 CCAGTTCATC ATCATGTACA GCCTGGATGG CAAGAAGTGG CAGACCTACA
GGGGCAACAG

7141 CACTGGCACC CTGATGGTGT TCTTTGGCAA TGTGGACAGC TCTGGCATCA
AGCACAACAT

7201 CTTCAACCCC CCCATCATTG CCAGATACAT CAGGCTGCAC CCCACCCACT
ACAGCATCAG

7261 GAGCACCTG AGGATGGAGC TGATGGGCTG TGACCTGAAC AGCTGCAGCA
TGCCCCCTGGG

7321 CATGGAGAGC AAGGCCATCT CTGATGCCCA GATCACTGCC AGCAGCTACT
TCACCAACAT

7381 GTTTGCCACC TGGAGCCCCA GCAAGGCCAG GCTGCACCTG CAGGGCAGGA
GCAATGCCTG

7441 GAGGCCCCAG GTCAACAACC CCAAGGAGTG GCTGCAGGTG GACTTCCAGA
AGACCATGAA

7501 GGTGACTGGG GTGACCACCC AGGGGGTGAA GAGCCTGCTG ACCAGCATGT
ATGTGAAGGA

7561 GTTCCTGATC AGCAGCAGCC AGGATGGCCA CCAGTGGACC CTGTTCTTCC
AGAATGGCAA

7621 GGTGAAGGTG TTCCAGGGCA ACCAGGACAG CTTACCCCT GTGGTGAACA
GCCTGGACCC

7681 CCCCTGCTG ACCAGATACC TGAGGATTCA CCCCAGAGC TGGGTGCACC
AGATTGCCCT

7741 GAGGATGGAG GTGCTGGGCT GTGAGGCCCA GGACCTGTAC TGAGCGGCCG
CGGGCCCAAT

7801 CAACCTCTGG ATTACAAAAT TTGTGAAAGA TTGACTGGTA TTCTTAACTA
TGTTGCTCCT

7861 TTTACGCTAT GTGGATACGC TGCTTTAATG CCTTTGTATC ATGCTATTGC
TTCCCGTATG

7921 GCTTTCATTT TCTCCTCCTT GTATAAATCC TGGTTGCTGT CTCTTTATGA
GGAGTTGTGG

7981 CCCGTTGTCA GGCAACGTGG CGTGGTGTGC ACTGTGTTTG CTGACGCAAC
CCCCACTGGT

8041 TGGGGCATTG CCACCACCTG TCAGCTCCTT TCCGGGACTT TCGCTTTCCC
CCTCCCTATT

8101 GCCACGGCGG AACTCATCGC CGCCTGCCTT GCCCGCTGCT GGACAGGGGC
TCGGCTGTTG

8161 GGC ACTGACA ATTCCGTGGT GTTGTGCGGG AAATCATCGT CCTTTCCTTG
GCTGCTCGCC

8221 TGTGTTGCCA CCTGGATTCT GCGCGGGACG TCCTTCTGCT ACGTCCCTTC
GGCCCTCAAT

8281 CCAGCGGACC TTCCTTCCC GCGCCTGCTG CCGGCTCTGC GGCCTCTTCC
GCGTCTTCGC

8341 CTTCGCCCTC AGACGAGTCG GATCTCCCTT TGGGCCGCCT CCCC GCAAGC
TTCGCACTTT

8401 TTAAAAGAAA AGGGAGGACT GGATGGGATT TATTACTCCG ATAGGACGCT
GGCTTGTAAC

8461 TCAGTCTCTT ACTAGGAGAC CAGCTTGAGC CTGGGTGTTC GCTGGTTAGC
CTAACCTGGT

8521 TGGCCACCAG GGGTAAGGAC TCCTTGGCTT AGAAAGCTAA TAAACTTGCC
TGCATTAGAG

8581 CTCTTACGCG TCCCGGGCTC GAGATCCGCA TCTCAATTAG TCAGCAACCA
TAGTCCCGCC

8641 CCTAACTCCG CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATTCTC
CGCCCCATGG

8701 CTGACTAATT TTTTTTATTT ATGCAGAGGC CGAGGCCGCC TCGGCCTCTG
AGCTATTCCA

8761 GAAGTAGTGA GGAGGCTTTT TTGGAGGCCT AGGCTTTTGC AAAAAGCTAA
CTTGTTTATT

8821 GCAGCTTATA ATGGTTACAA ATAAAGCAAT AGCATCACAA ATTCACAAA
TAAAGCATTT

8881 TTTTCACTGC ATTCTAGTTG TGGTTTGTCC AAACATCA ATGTATCTTA
TCATGTCTGT

8941 CCGCTTCTC GCTCACTGAC TCGCTGCGCT CGGTCGTTCC GCTGCGGCGA
GCGGTATCAG

9001 CTCACTCAA GCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA
GGAAAGAACA

9061 TGTGAGCAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG
CTGGCGTTTT

9121 TCCATAGGCT CCGCCCCCT GACGAGCATC AAAAAATCG ACGCTCAAGT
CAGAGGTGGC

9181 GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC
CTCGTGCGCT

9241 CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
TCGGGAAGCG

9301 TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC
GTTCGCTCCA

9361 AGCTGGGCTG TGTGCACGAA CCCCCGTTT AGCCCGACCG CTGCGCCTTA
TCCGGTAACT

9421 ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA
GCCACTGGTA

9481 ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG
TGGTGGCCTA

9541 ACTACGGCTA CACTAGAAGA ACAGTATTTG GTATCTGCGC TCTGCTGAAG
CCAGTTACCT

9601 TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT
AGCGGTGGTT

9661 TTTTGTGTTG CAAGCAGCAG ATTACGCGCA GAAAAAAGG ATCTCAAGAA
GATCCTTTGA

9721 TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAAC TC ACGTTAAGGG
ATTTTGGTCA
9781 TGAGATTATC AAAAAGGATC TTCACCTAGA TCCTTTTAAA TAAAAAATGA
AGTTTTAAAT
9841 CAATCTAAAG TATATATGAG TAAACTTGGT CTGACAGTTA GAAAAACTCA
TCGAGCATCA
9901 AATGAAACTG CAATTTATTC ATATCAGGAT TATCAATACC ATATTTTTGA
AAAAGCCGTT
9961 TCTGTAATGA AGGAGAAAAC TCACCGAGGC AGTTCCATAG GATGGCAAGA
TCCTGGTATC
10021 GGTCTGCGAT TCCGACTCGT CCAACATCAA TACAACCTAT TAATTTCCCC
TCGTCAAAAA
10081 TAAGGTTATC AAGTGAGAAA TCACCATGAG TGACGACTGA ATCCGGTGAG
AATGGCAACA
10141 GCTTATGCAT TTCTTTCCAG ACTTGTTCAA CAGGCCAGCC ATTACGCTCG
TCATCAAAAT
10201 CACTCGCATC AACCAAACCG TTATTCATTC GTGATTGCGC CTGAGCGAGA
CGAAATACGC
10261 GATCGCTGTT AAAAGGACAA TTACAAACAG GAATCGAATG CAACCGGCGC
AGGAACACTG
10321 CCAGCGCATC AACAATATTT TCACCTGAAT CAGGATATTC TTCTAATACC
TGGAATGCTG
10381 TTTTCCGGG GATCGCAGTG GTGAGTAACC ATGCATCATC AGGAGTACGG
ATAAAATGCT
10441 TGATGGTCGG AAGAGGCATA AATTCCGTCA GCCAGTTTAG TCTGACCATC
TCATCTGTAA
10501 CATCATTGGC AACGCTACCT TTGCCATGTT TCAGAAACAA CTCTGGCGCA
TCGGGCTTCC
10561 CATACAATCG ATAGATTGTC GCACCTGATT GCCCGACATT ATCGCGAGCC
CATTTATACC
10621 CATATAAATC AGCATCCATG TTGGAATTTA ATCGCGGCCT AGAGCAAGAC
GTTTCCCGTT
10681 GAATATGGCT CATAACACCC CTTGTATTAC TGTTTATGTA AGCAGACAGT
TTTATTGTTC
10741 ATGATGATAT ATTTTTATCT TGTGCAATGT AACATCAGAG ATTTTGAGAC
ACAACAATTG
10801 GTCGACGGAT CC

**SEQ ID NO: 26: Плазмида F/HN-SIV-hCEF-HFVIII-V3, определенная на
фигуре 4B (pDNA1 pGM413)**

Длина: 10519; Тип молекулы: ДНК; Особенности локализации/Определители:

источник, 1..10519; тип молекулы, другая ДНК; примечание: pGM413; организм, синтетическая конструкция

1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT
61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
TGACGTATGT
301 TCCCATAGTA ACGCCAATAG GGA CTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
361 AACTGCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
421 CAATGACGGT AAATGGCCCCG CCTGGCATTG TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGC GTGTACG
GTGGGAGGTC
721 TATATAAGCA GAGCTCGCTG GCTTGTA ACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
781 TGGGTGTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
841 GAAAGCTAAT AA ACTTGCCCT GCATTAGAGC TTATCTGAGT CAAGTGTCTT
CATTGACGCC
901 TCACTCTCTT GAACGGGAAT CTTCCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA
1021 GGC GTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGA CTAGGAG
AGGCCGTAGC

1141 CGTAACTACT CTGGGCAAGT AGGGCAGGCG GTGGGTACGC AATGGGGGCG
GCTACCTCAG

1201 CACTAAATAG GAGACAATTA GACCAATTTG AGAAAATACG ACTTCGCCCCG
AACGGAAAGA

1261 AAAAGTACCA AATTAAACAT TTAATATGGG CAGGCAAGGA GATGGAGCGC
TTCGGCCTCC

1321 ATGAGAGGTT GTTGGAGACA GAGGAGGGGT GTAAAAGAAT CATAGAAGTC
CTCTACCCCC

1381 TAGAACCAAC AGGATCGGAG GGCTTAAAAA GTCTGTTCAA TCTTGTGTGC
GTGCTATATT

1441 GCTTGCACAA GGAACAGAAA GTGAAAGACA CAGAGGAAGC AGTAGCAACA
GTAAGACAAC

1501 ACTGCCATCT AGTGGAAAAA GAAAAAAGTG CAACAGAGAC ATCTAGTGGA
CAAAAGAAAA

1561 ATGACAAGGG AATAGCAGCG CCACCTGGTG GCAGTCAGAA TTTTCCAGCG
CAACAACAAG

1621 GAAATGCCTG GGTACATGTA CCCTTGTCAC CGCGCACCTT AAATGCGTGG
GTAAAAGCAG

1681 TAGAGGAGAA AAAATTTGGA GCAGAAATAG TACCCATGTT TCAAGCCCTA
TCGAATTCCC

1741 GTTTGTGCTA GGGTTCTTAG GCTTCTTGGG GGCTGCTGGA ACTGCAATGG
GAGCAGCGGC

1801 GACAGCCCTG ACGGTCCAGT CTCAGCATTT GCTTGCTGGG ATACTGCAGC
AGCAGAAGAA

1861 TCTGCTGGCG GCTGTGGAGG CTCAACAGCA GATGTTGAAG CTGACCATTT
GGGGTGTTAA

1921 AAACCTCAAT GCCCGCGTCA CAGCCCTTGA GAAGTACCTA GAGGATCAGG
CACGACTAAA

1981 CTCCTGGGGG TGCGCATGGA AACAAGTATG TCATACCACA GTGGAGTGGC
CCTGGACAAA

2041 TCGGACTCCG GATTGGCAAA ATATGACTTG GTTGGAGTGG GAAAGACAAA
TAGCTGATTT

2101 GGAAAGCAAC ATTACGAGAC AATTAGTGAA GGCTAGAGAA CAAGAGGAAA
AGAATCTAGA

2161 TGCCTATCAG AAGTTAACTA GTTGGTCAGA TTTCTGGTCT TGGTTCGATT
TCTCAAATG

2221 GCTTAACATT TAAAAAATGG GATTTTTAGT AATAGTAGGA ATAATAGGGT
TAAGATTACT

2281 TTACACAGTA TATGGATGTA TAGTGAGGGT TAGGCAGGGA TATGTTCCCTC
TATCTCCACA

2341 GATCCATATC CGCGGCAATT TAAAAAGAAA GGGAGGAATA GGGGGACAGA
CTTCAGCAGA

2401 GAGACTAATT AATATAATAA CAACACAATT AGAAATACAA CATTTACAAA
CCAAAATTCA

2461 AAAAATTTTA AATTTTAGAG CCGCGGAGAT CTGTTACATA ACTTATGGTA
AATGGCCTGC

2521 CTGGCTGACT GCCCAATGAC CCCTGCCCAA TGATGTCAAT AATGATGTAT
GTTCCCATGT

2581 AATGCCAATA GGGACTTTCC ATTGATGTCA ATGGGTGGAG TATTTATGGT
AACTGCCCAC

2641 TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTATGCCCC CTATTGATGT
CAATGATGGT

2701 AAATGGCCTG CCTGGCATTG TGCCCAGTAC ATGACCTTAT GGGACTTTCC
TACTTGGCAG

2761 TACATCTATG TATTAGTCAT TGCTATTACC ATGGGAATTC ACTAGTGGAG
AAGAGCATGC

2821 TTGAGGGCTG AGTGCCCCTC AGTGGGCAGA GAGCACATGG CCCACAGTCC
CTGAGAAGTT

2881 GGGGGGAGGG GTGGGCAATT GAACTGGTGC CTAGAGAAGG TGGGGCTTGG
GTAAACTGGG

2941 AAAGTGATGT GGTGTACTGG CTCCACCTTT TTCCCCAGGG TGGGGGAGAA
CCATATATAA

3001 GTGCAGTAGT CTCTGTGAAC ATTCAAGCTT CTGCCTTCTC CCTCCTGTGA
GTTTGCTAGC

3061 CACCAATGCA GATTGAGCTG AGCACCTGCT TCTTCCTGTG CCTGCTGAGG
TTCTGCTTCT

3121 CTGCCACCAG GAGATACTAC CTGGGGGCTG TGGAGCTGAG CTGGGACTAC
ATGCAGTCTG

3181 ACCTGGGGGA GCTGCCTGTG GATGCCAGGT TCCCCCCCAG AGTGCCCAAG
AGCTTCCCCT

3241 TCAACACCTC TGTGGTGTAC AAGAAGACCC TGTTTGTGGA GTTCACTGAC
CACCTGTTCA

3301 ACATTGCCAA GCCCAGGCC CCCTGGATGG GCCTGCTGGG CCCCACCATC
CAGGCTGAGG

3361 TGTATGACAC TGTGGTGATC ACCCTGAAGA ACATGGCCAG CCACCCTGTG
AGCCTGCATG

3421 CTGTGGGGGT GAGCTACTGG AAGGCCTCTG AGGGGGCTGA GTATGATGAC
CAGACCAGCC

3481 AGAGGGAGAA GGAGGATGAC AAGGTGTTCC CTGGGGGCAG CCACACCTAT
GTGTGGCAGG

3541 TGCTGAAGGA GAATGGCCCC ATGGCCTCTG ACCCCCTGTG CCTGACCTAC
AGCTACCTGA

3601 GCCATGTGGA CCTGGTGAAG GACCTGAACT CTGGCCTGAT TGGGGCCCTG
CTGGTGTGCA

3661 GGGAGGGCAG CCTGGCCAAG GAGAAGACCC AGACCCTGCA CAAGTTCATC
CTGCTGTTTG

3721 CTGTGTTTGA TGAGGGCAAG AGCTGGCACT CTGAAACCAA GAACAGCCTG
ATGCAGGACA

3781 GGGATGCTGC CTCTGCCAGG GCCTGGCCCA AGATGCACAC TGTGAATGGC
TATGTGAACA

3841 GGAGCCTGCC TGGCCTGATT GGCTGCCACA GGAAGTCTGT GTACTGGCAT
GTGATTGGCA

3901 TGGGCACCAC CCCTGAGGTG CACAGCATCT TCCTGGAGGG CCACACCTTC
CTGGTCAGGA

3961 ACCACAGGCA GGCCAGCCTG GAGATCAGCC CCATCACCTT CCTGACTGCC
CAGACCCTGC

4021 TGATGGACCT GGGCCAGTTC CTGCTGTTCT GCCACATCAG CAGCCACCAG
CATGATGGCA

4081 TGGAGGCCTA TGTGAAGGTG GACAGCTGCC CTGAGGAGCC CCAGCTGAGG
ATGAAGAACA

4141 ATGAGGAGGC TGAGGACTAT GATGATGACC TGACTGACTC TGAGATGGAT
GTGGTGAGGT

4201 TTGATGATGA CAACAGCCCC AGCTTCATCC AGATCAGGTC TGTGGCCAAG
AAGCACCCCA

4261 AGACCTGGGT GCACTACATT GCTGCTGAGG AGGAGGACTG GGACTATGCC
CCCCTGGTGC

4321 TGGCCCCTGA TGACAGGAGC TACAAGAGCC AGTACCTGAA CAATGGCCCC
CAGAGGATTG

4381 GCAGGAAGTA CAAGAAGGTC AGGTTCATGG CCTACACTGA TGAAACCTTC
AAGACCAGGG

4441 AGGCCATCCA GCATGAGTCT GGCATCCTGG GCCCCCTGCT GTATGGGGAG
GTGGGGGACA

4501 CCCTGCTGAT CATCTTCAAG AACCAGGCCA GCAGGCCCTA CAACATCTAC
CCCCATGGCA

4561 TCACTGATGT GAGGCCCTG TACAGCAGGA GGCTGCCCAA GGGGGTGAAG
CACCTGAAGG

4621 ACTTCCCCAT CCTGCCTGGG GAGATCTTCA AGTACAAGTG GACTGTGACT
GTGGAGGATG

4681 GCCCCACCAA GTCTGACCCC AGGTGCCTGA CCAGATACTA CAGCAGCTTT
GTGAACATGG

4741 AGAGGGACCT GGCCTCTGGC CTGATTGGCC CCCTGCTGAT CTGCTACAAG
GAGTCTGTGG

4801 ACCAGAGGGG CAACCAGATC ATGTCTGACA AGAGGAATGT GATCCTGTTC
TCTGTGTTTG

4861 ATGAGAACAG GAGCTGGTAC CTGACTGAGA ACATCCAGAG GTTCCTGCCC
AACCTGCTG

4921 GGGTGCAGCT GGAGGACCCT GAGTTCAGG CCAGCAACAT CATGCACAGC
ATCAATGGCT

4981 ATGTGTTTGA CAGCCTGCAG CTGTCTGTGT GCCTGCATGA GGTGGCCTAC
TGGTACATCC

5041 TGAGCATTGG GGCCAGACT GACTTCCTGT CTGTGTTCTT CTCTGGCTAC
ACCTTCAAGC

5101 ACAAGATGGT GTATGAGGAC ACCCTGACCC TGTTCCTT CTCTGGGGAG
ACTGTGTTCA

5161 TGAGCATGGA GAACCCTGGC CTGTGGATTC TGGGCTGCCA CAACTCTGAC
TTCAGGAACA

5221 GGGGCATGAC TGCCCTGCTG AAAGTCTCCA GCTGTGACAA GAACACTGGG
GACTACTATG

5281 AGGACAGCTA TGAGGACATC TCTGCCTACC TGCTGAGCAA GAACAATGCC
ATTGAGCCCA

5341 GGAGCTTCAG CCAGAATGCC ACTAATGTGT CTAACAACAG CAACACCAGC
AATGACAGCA

5401 ATGTGTCTCC CCCAGTGCTG AAGAGGCACC AGAGGGAGAT CACCAGGACC
ACCCTGCAGT

5461 CTGACCAGGA GGAGATTGAC TATGATGACA CCATCTCTGT GGAGATGAAG
AAGGAGGACT

5521 TTGACATCTA CGACGAGGAC GAGAACCAGA GCCCCAGGAG CTTCCAGAAG
AAGACCAGGC

5581 ACTACTTCAT TGCTGCTGTG GAGAGGCTGT GGGACTATGG CATGAGCAGC
AGCCCCATG

5641 TGCTGAGGAA CAGGGCCCAG TCTGGCTCTG TGCCCCAGTT CAAGAAGGTG
GTGTTCCAGG

5701 AGTTCATGA TGGCAGCTTC ACCCAGCCCC TGTACAGAGG GGAGCTGAAT
GAGCACCTGG

5761 GCCTGCTGGG CCCCTACATC AGGGCTGAGG TGGAGGACAA CATCATGGTG
ACCTTCAGGA

5821 ACCAGGCCAG CAGGCCCTAC AGCTTCTACA GCAGCCTGAT CAGCTATGAG
GAGGACCAGA

5881 GGCAGGGGGC TGAGCCCAGG AAGAACTTTG TGAAGCCCAA TGAAACCAAG
ACCTACTTCT

5941 GGAAGGTGCA GCACCACATG GCCCCCACCA AGGATGAGTT TGA CTGCAAG
GCCTGGGCCT

6001 ACTTCTCTGA TGTGGACCTG GAGAAGGATG TGCACTCTGG CCTGATTGGC
CCCCTGCTGG

6061 TGTGCCACAC CAACACCCTG AACCTGCCC ATGGCAGGCA GGTGACTGTG
CAGGAGTTTG

6121 CCCTGTTCTT CACCATCTTT GATGAAACCA AGAGCTGGTA CTTCACTGAG
AACATGGAGA

6181 GGAAGTGCAG GGCCCCCTGC AACATCCAGA TGGAGGACCC CACCTTCAAG
GAGA ACTACA

6241 GGTTCCATGC CATCAATGGC TACATCATGG ACACCCTGCC TGGCCTGGTG
ATGGCCCAGG

6301 ACCAGAGGAT CAGGTGGTAC CTGCTGAGCA TGGGCAGCAA TGAGA ACATC
CACAGCATCC

6361 ACTTCTCTGG CCATGTGTTC ACTGTGAGGA AGAAGGAGGA GTACAAGATG
GCCCTGTACA

6421 ACCTGTACCC TGGGGTGTTT GAGACTGTGG AGATGCTGCC CAGCAAGGCT
GGCATCTGGA

6481 GGGTGGAGTG CCTGATTGGG GAGCACCTGC ATGCTGGCAT GAGCACCTG
TTCCTGGTGT

6541 ACAGCAACAA GTGCCAGACC CCCCTGGGCA TGGCCTCTGG CCACATCAGG
GACTTCCAGA

6601 TCACTGCCTC TGGCCAGTAT GGCCAGTGGG CCCCCAAGCT GGCCAGGCTG
CACTACTCTG

6661 GCAGCATCAA TGCCTGGAGC ACCAAGGAGC CCTTCAGCTG GATCAAGGTG
GACCTGCTGG

6721 CCCCATGAT CATCCATGGC ATCAAGACCC AGGGGGCCAG GCAGAAGTTC
AGCAGCCTGT

6781 ACATCAGCCA GTTCATCATC ATGTACAGCC TGGATGGCAA GAAGTGGCAG
ACCTACAGGG

6841 GCAACAGCAC TGGCACCTG ATGGTGTTCT TTGGCAATGT GGACAGCTCT
GGCATCAAGC

6901 ACAACATCTT CAACCCCCC ATCATTGCCA GATACATCAG GCTGCACCCC
ACCCACTACA

6961 GCATCAGGAG CACCCTGAGG ATGGAGCTGA TGGGCTGTGA CCTGAACAGC
TGCAGCATGC

7021 CCCTGGGCAT GGAGAGCAAG GCCATCTCTG ATGCCCAGAT CACTGCCAGC
AGCTACTTCA

7081 CCAACATGTT TGCCACCTGG AGCCCCAGCA AGGCCAGGCT GCACCTGCAG
GGCAGGAGCA

7141 ATGCCTGGAG GCCCCAGGTC AACAAACCCCA AGGAGTGGCT GCAGGTGGAC
TTCCAGAAGA

7201 CCATGAAGGT GACTGGGGTG ACCACCCAGG GGGTGAAGAG CCTGCTGACC
AGCATGTATG

7261 TGAAGGAGTT CCTGATCAGC AGCAGCCAGG ATGGCCACCA GTGGACCCTG
TTCTTCCAGA

7321 ATGGCAAGGT GAAGGTGTTT CAGGGCAACC AGGACAGCTT CACCCCTGTG
GTGAACAGCC

7381 TGGACCCCCC CCTGCTGACC AGATACCTGA GGATTCACCC CCAGAGCTGG
GTGCACCAGA

7441 TTGCCCTGAG GATGGAGGTG CTGGGCTGTG AGGCCCAGGA CCTGTACTGA
GCGGCCGCGG

7501 GCCCAATCAA CCTCTGGATT ACAAATTTG TGAAAGATTG ACTGGTATTC
TTAACTATGT

7561 TGCTCCTTTT ACGCTATGTG GATACGCTGC TTTAATGCCT TTGTATCATG
CTATTGCTTC

7621 CCGTATGGCT TTCATTTTCT CTCCTTGTA TAAATCCTGG TTGCTGTCTC
TTTATGAGGA

7681 GTTGTGGCCC GTTGTCAGGC AACGTGGCGT GGTGTGCACT GTGTTTGCTG
ACGCAACCCC

7741 CACTGGTTGG GGCATTGCCA CCACCTGTCA GCTCCTTTCC GGGACTTTCG
CTTTCCCCCT

7801 CCCTATTGCC ACGGCGGAAC TCATCGCCGC CTGCCTTGCC CGCTGCTGGA
CAGGGGCTCG

7861 GCTGTTGGGC ACTGACAATT CCGTGGTGTG GTCGGGGAAA TCATCGTCCT
TTCCTTGGCT

7921 GCTCGCCTGT GTTGCCACCT GGATTCTGCG CGGGACGTCC TTCTGCTACG
TCCCTTCGGC

7981 CCTCAATCCA GCGGACCTTC CTTCCCGCGG CCTGCTGCCG GCTCTGCGGC
CTCTTCCGCG

8041 TCTTCGCCTT CGCCCTCAGA CGAGTCGGAT CTCCCTTTGG GCCGCCTCCC
CGCAAGCTTC

8101 GCACTTTTTA AAAGAAAAGG GAGGACTGGA TGGGATTTAT TACTCCGATA
GGACGCTGGC

8161 TTGTA ACTCA GTCTCTTACT AGGAGACCAG CTTGAGCCTG GGTGTTTCGCT
GGTTAGCCTA

8221 ACCTGGTTGG CCACCAGGGG TAAGGACTCC TTGGCTTAGA AAGCTAATAA
ACTTGCCTGC

8281 ATTAGAGCTC TTACGCGTCC CGGGCTCGAG ATCCGCATCT CAATTAGTCA
GCAACCATAG

8341 TCCCGCCCCT AACTCCGCCC ATCCCGCCCC TAACTCCGCC CAGTTCCGCC
CATTCTCCGC

8401 CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCTCG
GCCTCTGAGC

8461 TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG CTTTTGCAA
AAGCTAACTT

8521 GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT
TCACAAATAA

8581 AGCATTTTTT TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG
TATCTTATCA

8641 TGTCTGTCCG CTCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT
GCGGCGAGCG

8701 GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA
TAACGCAGGA

8761 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC
CGCGTTGCTG

8821 GCGTTTTTCC ATAGGCTCCG CCCCCTGAC GAGCATCACA AAAATCGACG
CTCAAGTCAG

8881 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCTGG
AAGCTCCCTC

8941 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT
TCTCCCTTCG

9001 GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT
GTAGGTCGTT

9061 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAGC CCGACCGCTG
CGCCTTATCC

9121 GGTA ACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT
GGCAGCAGCC

9181 ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT
CTTGAAGTGG

9241 TGGCCTAACT ACGGCTACAC TAGAAGAACA GTATTTGGTA TCTGCGCTCT
GCTGAAGCCA

9301 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC
CGCTGGTAGC

9361 GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC
TCAAGAAGAT

9421 CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCAG
TTAAGGGATT

9481 TTGGTCATGA GATTATCAA AAGGATCTTC ACCTAGATCC TTTTAAATTA
AAAATGAAGT

9541 TTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTAGAA
AAACTCATCG
9601 AGCATCAAAT GAAACTGCAA TTTATTCATA TCAGGATTAT CAATACCATA
TTTTTGAAAA
9661 AGCCGTTTCT GTAATGAAGG AGAAAACCTCA CCGAGGCAGT TCCATAGGAT
GGCAAGATCC
9721 TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC AACCTATTA
TTCCCCCTCG
9781 TCAAAAATAA GGTTATCAAG TGAGAAATCA CCATGAGTGA CGACTGAATC
CGGTGAGAAT
9841 GGCAACAGCT TATGCATTTT TTTCCAGACT TGTTC AACAG GCCAGCCATT
ACGCTCGTCA
9901 TCAAAATCAC TCGCATCAAC CAAACCGTTA TTCATTCGTG ATTGCGCCTG
AGCGAGACGA
9961 AATACGCGAT CGCTGTAAA AGGACAATTA CAAACAGGAA TCGAATGCAA
CCGGCGCAGG
10021 AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG GATATTCTTC
TAATACCTGG
10081 AATGCTGTTT TTCCGGGGAT CGCAGTGGTG AGTAACCATG CATCATCAGG
AGTACGGATA
10141 AAATGCTTGA TGGTCGGAAG AGGCATAAAT TCCGTCAGCC AGTTTAGTCT
GACCATCTCA
10201 TCTGTAACAT CATTGGCAAC GCTACCTTTG CCATGTTTCA GAAACAACCTC
TGGCGCATCG
10261 GGCTTCCCAT ACAATCGATA GATTGTCGCA CCTGATTGCC CGACATTATC
GCGAGCCCAT
10321 TTATACCCAT ATAAATCAGC ATCCATGTTG GAATTTAATC GCGGCCTAGA
GCAAGACGTT
10381 TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT TTATGTAAGC
AGACAGTTTT
10441 ATTGTTTCATG ATGATATATT TTTATCTTGT GCAATGTAAC ATCAGAGATT
TTGAGACACA
10501 ACAATTGGTC GACGGATCC

**SEQ ID NO: 27: Плазмида F/HN-SIV-CMV-HFVIII-N6-co, определенная на
фигуре 4C (pDNA1 pGM412)**

Длина: 11400; Тип молекулы: ДНК; Особенности локализации/Определители:
источник, 1..11400; тип молекулы, другая ДНК; примечание: pGM412; организм,
синтетическая конструкция

1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT

61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC
121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCATTG ACGTCAATAA
TGACGTATGT
301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
361 AACTGCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT
421 CAATGACGGT AAATGGCCCC CCTGGCATTG TGCCCAGTAC ATGACCTTAC
GGGACTTTCC
481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
GGTTTTGGCA
541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT
601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA
661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGC GTGTACG
GTGGGAGGTC
721 TATATAAGCA GAGCTCGCTG GCTTGTA ACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC
781 TGGGTGTTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA
841 GAAAGCTAAT AAACCTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCTT
CATTGACGCC
901 TCACTCTCTT GAACGGGAAT CTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA
961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA
1021 GCGTTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA
1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTAGGAG
AGGCCGTAGC
1141 CGTAACTACT CTGGGCAAGT AGGGCAGGCG GTGGGTACGC AATGGGGGCG
GCTACCTCAG
1201 CACTAAATAG GAGACAATTA GACCAATTTG AGAAAATACG ACTTCGCCCCG
AACGGAAAGA

1261 AAAAGTACCA AATTAACAT TTAATATGGG CAGGCAAGGA GATGGAGCGC
TTCGGCCTCC

1321 ATGAGAGGTT GTTGGAGACA GAGGAGGGGT GTAAAAGAAT CATAGAAGTC
CTCTACCCCC

1381 TAGAACCAAC AGGATCGGAG GGCTTAAAAA GTCTGTTCAA TCTTGTGTGC
GTGCTATATT

1441 GCTTGCACAA GGAACAGAAA GTGAAAGACA CAGAGGAAGC AGTAGCAACA
GTAAGACAAC

1501 ACTGCCATCT AGTGGAAAAA GAAAAAAGTG CAACAGAGAC ATCTAGTGGA
CAAAAGAAAA

1561 ATGACAAGGG AATAGCAGCG CCACCTGGTG GCAGTCAGAA TTTTCCAGCG
CAACAACAAG

1621 GAAATGCCTG GGTACATGTA CCCTTGTCAC CGCGCACCTT AAATGCGTGG
GTAAAAGCAG

1681 TAGAGGAGAA AAAATTTGGA GCAGAAATAG TACCCATGTT TCAAGCCCTA
TCGAATTCCC

1741 GTTTGTGCTA GGGTTCTTAG GCTTCTTGGG GGCTGCTGGA ACTGCAATGG
GAGCAGCGGC

1801 GACAGCCCTG ACGGTCCAGT CTCAGCATTT GCTTGCTGGG ATRACTGCAGC
AGCAGAAGAA

1861 TCTGCTGGCG GCTGTGGAGG CTCAACAGCA GATGTTGAAG CTGACCATTT
GGGGTGTTAA

1921 AAACCTCAAT GCCCGCGTCA CAGCCCTTGA GAAGTACCTA GAGGATCAGG
CACGACTAAA

1981 CTCCTGGGGG TGCGCATGGA AACAAGTATG TCATACCACA GTGGAGTGGC
CCTGGACAAA

2041 TCGGACTCCG GATTGGCAAA ATATGACTTG GTTGGAGTGG GAAAGACAAA
TAGCTGATTT

2101 GGAAAGCAAC ATTACGAGAC AATTAGTGAA GGCTAGAGAA CAAGAGGAAA
AGAATCTAGA

2161 TGCCTATCAG AAGTTAACTA GTTGGTCAGA TTTCTGGTCT TGGTTCGATT
TCTCAAATG

2221 GCTTAACATT TAAAAATGG GATTTTTAGT AATAGTAGGA ATAATAGGGT
TAAGATTA

2281 TTACACAGTA TATGGATGTA TAGTGAGGGT TAGGCAGGGA TATGTTCCCTC
TATCTCCACA

2341 GATCCATATC CGCGGCAATT TAAAAGAAA GGGAGGAATA GGGGGACAGA
CTTCAGCAGA

2401 GAGACTAATT AATATAATAA CAACACAATT AGAAATACAA CATTTACAAA
CCAAAATTCA

2461 AAAAATTTTA AATTTTAGAG CCGCGGAGAT CTCAATATTG GCCATTAGCC
ATATTATTCA

2521 TTGGTTATAT AGCATAAATC AATATTGGCT ATTGGCCATT GCATACGTTG
TATCTATATC

2581 ATAATATGTA CATTTATATT GGCTCATGTC CAATATGACC GCCATGTTGG
CATTGATTAT

2641 TGACTIONGTTA TTAATAGTAA TCAATTACGG GGTCATTAGT TCATAGCCCA
TATATGGAGT

2701 TCCGCGTTAC ATAACTTACG GTAAATGGCC CGCCTGGCTG ACCGCCAAC
GACCCCGCC

2761 CATTGACGTC AATAATGACG TATGTTCCCA TAGTAACGCC AATAGGGACT
TTCCATTGAC

2821 GTCAATGGGT GGAGTATTTA CGGTAAACTG CCCACTTGGC AGTACATCAA
GTGTATCATA

2881 TGCCAAGTCC GCCCCCTATT GACGTCAATG ACGGTAAATG GCCCGCCTGG
CATTATGCCC

2941 AGTACATGAC CTTACGGGAC TTTCTACTT GGCAGTACAT CTACGTATTA
GTCATCGCTA

3001 TTACCATGGT GATGCGGTTT TGGCAGTACA CCAATGGGCG TGGATAGCGG
TTTGACTIONCAC

3061 GGGGATTTCC AAGTCTCCAC CCCATTGACG TCAATGGGAG TTTGTTTTGG
CACCAAATC

3121 AACGGGACTT TCCAAAATGT CGTAATAACC CCGCCCCGTT GACGCAAATG
GGCGGTAGGC

3181 GTGTACGGTG GGAGGTCTAT ATAAGCAGAG CTCGTTTAGT GAACCGTCAG
ATCACTIONGAA

3241 GCTTTATTGC GGTAGTTTAT CACAGTTAAA TTGCTAACGC AGTCAGTGCT
TCTGACACAA

3301 CAGTCTCGAA CTTAAGCTGC AGAAGTTGGT CGTGAGGCAC TGGGCAGGCT
AGCCACCAAT

3361 GCAGATTGAG CTGAGCACCT GCTTCTTCCT GTGCCTGCTG AGGTTCTGCT
TCTCTGCCAC

3421 CAGGAGATAC TACCTGGGGG CTGTGGAGCT GAGCTGGGAC TACATGCAGT
CTGACCTGGG

3481 GGAGCTGCCT GTGGATGCCA GGTTCCCCC CAGAGTGCCC AAGAGCTTCC
CCTTCAACAC

3541 CTCTGTGGTG TACAAGAAGA CCCTGTTTGT GGAGTTCACTION GACCACCTGT
TCAACATTGC

3601 CAAGCCCAGG CCCCCCTGGA TGGGCCTGCT GGGCCCCACC ATCCAGGCTG
AGGTGTATGA

3661 CACTGTGGTG ATCACCTGA AGAACATGGC CAGCCACCCT GTGAGCCTGC
ATGCTGTGGG

3721 GGTGAGCTAC TGGAAGGCCT CTGAGGGGGC TGAGTATGAT GACCAGACCA
GCCAGAGGGA

3781 GAAGGAGGAT GACAAGGTGT TCCCTGGGGG CAGCCACACC TATGTGTGGC
AGGTGCTGAA

3841 GGAGAATGGC CCCATGGCCT CTGACCCCT GTGCCTGACC TACAGCTACC
TGAGCCATGT

3901 GGACCTGGTG AAGGACCTGA ACTCTGGCCT GATTGGGGCC CTGCTGGTGT
GCAGGGAGGG

3961 CAGCCTGGCC AAGGAGAAGA CCCAGACCCT GCACAAGTTC ATCCTGCTGT
TTGCTGTGTT

4021 TGATGAGGGC AAGAGCTGGC ACTCTGAAAC CAAGAACAGC CTGATGCAGG
ACAGGGATGC

4081 TGCCTCTGCC AGGGCCTGGC CCAAGATGCA CACTGTGAAT GGCTATGTGA
ACAGGAGCCT

4141 GCCTGGCCTG ATTGGCTGCC ACAGGAAGTC TGTGTACTGG CATGTGATTG
GCATGGGCAC

4201 CACCCCTGAG GTGCACAGCA TCTTCCTGGA GGGCCACACC TTCCTGGTCA
GGAACCACAG

4261 GCAGGCCAGC CTGGAGATCA GCCCATCAC CTCCTGACT GCCCAGACCC
TGCTGATGGA

4321 CCTGGGCCAG TTCCTGCTGT TCTGCCACAT CAGCAGCCAC CAGCATGATG
GCATGGAGGC

4381 CTATGTGAAG GTGGACAGCT GCCCTGAGGA GCCCAGCTG AGGATGAAGA
ACAATGAGGA

4441 GGCTGAGGAC TATGATGATG ACCTGACTGA CTCTGAGATG GATGTGGTGA
GGTTTGATGA

4501 TGACAACAGC CCCAGCTTCA TCCAGATCAG GTCTGTGGCC AAGAAGCACC
CCAAGACCTG

4561 GGTGCACTAC ATTGCTGCTG AGGAGGAGGA CTGGGACTAT GCCCCCCTGG
TGCTGGCCCC

4621 TGATGACAGG AGCTACAAGA GCCAGTACCT GAACAATGGC CCCCAGAGGA
TTGGCAGGAA

4681 GTACAAGAAG GTCAGGTTCA TGGCCTACAC TGATGAAACC TTCAAGACCA
GGGAGGCCAT

4741 CCAGCATGAG TCTGGCATCC TGGGCCCCCT GCTGTATGGG GAGGTGGGGG
ACACCCTGCT

4801 GATCATCTTC AAGAACCAGG CCAGCAGGCC CTACAACATC TACCCCATG
GCATCACTGA

4861 TGTGAGGCC CTGTACAGCA GGAGGCTGCC CAAGGGGGTG AAGCACCTGA
AGGACTTCCC

4921 CATCCTGCCT GGGGAGATCT TCAAGTACAA GTGGACTGTG ACTGTGGAGG
ATGGCCCCAC

4981 CAAGTCTGAC CCCAGGTGCC TGACCAGATA CTACAGCAGC TTTGTGAACA
TGGAGAGGGA

5041 CCTGGCCTCT GGCCTGATTG GCCCCCTGCT GATCTGCTAC AAGGAGTCTG
TGGACCAGAG

5101 GGGCAACCAG ATCATGTCTG ACAAGAGGAA TGTGATCCTG TTCTCTGTGT
TTGATGAGAA

5161 CAGGAGCTGG TACCTGACTG AGAACATCCA GAGGTTCTG CCCAACCTG
CTGGGGTGCA

5221 GCTGGAGGAC CCTGAGTTCC AGGCCAGCAA CATCATGCAC AGCATCAATG
GCTATGTGTT

5281 TGACAGCCTG CAGCTGTCTG TGTGCCTGCA TGAGGTGGCC TACTGGTACA
TCCTGAGCAT

5341 TGGGGCCCAG ACTGACTTCC TGTCTGTGTT CTTCTCTGGC TACACCTTCA
AGCACAAGAT

5401 GGTGTATGAG GACACCCTGA CCCTGTTCCC CTTCTCTGGG GAGACTGTGT
TCATGAGCAT

5461 GGAGAACCCT GGCCTGTGGA TTCTGGGCTG CCACAACCTCT GACTTCAGGA
ACAGGGGCAT

5521 GACTGCCCTG CTGAAAGTCT CCAGCTGTGA CAAGAACACT GGGGACTACT
ATGAGGACAG

5581 CTATGAGGAC ATCTCTGCCT ACCTGCTGAG CAAGAACAAT GCCATTGAGC
CCAGGAGCTT

5641 CAGCCAGAAC AGCAGGCACC CCAGCACCAG GCAGAAGCAG TTCAATGCCA
CCACCATCCC

5701 TGAGAATGAC ATAGAGAAGA CAGACCCATG GTTTGCCCAC CGGACCCCCA
TGCCCAAGAT

5761 CCAGAATGTG AGCAGCTCTG ACCTGCTGAT GCTGCTGAGG CAGAGCCCCA
CCCCCATGG

5821 CCTGAGCCTG TCTGACCTGC AGGAGGCCAA GTATGAAACC TTCTCTGATG
ACCCAGCCC

5881 TGGGGCCATT GACAGCAACA ACAGCCTGTC TGAGATGACC CACTTCAGGC
CCCAGCTGCA

5941 CCACTCTGGG GACATGGTGT TCACCCCTGA GTCTGGCCTG CAGCTGAGGC
TGAATGAGAA

6001 GCTGGGCACC ACTGCTGCCA CTGAGCTGAA GAAGCTGGAC TTCAAAGTCT
CCAGCACCAG

6061 CAACAACCTG ATCAGCACCA TCCCCTCTGA CAACCTGGCT GCTGGCACTG
ACAACACCAG

6121 CAGCCTGGGC CCCCCAGCA TGCCTGTGCA CTATGACAGC CAGCTGGACA
CCACCCTGTT

6181 TGGCAAGAAG AGCAGCCCCC TGA CTGAGCCTGT
CTGAGGAGAA

6241 CAATGACAGC AAGCTGCTGG AGTCTGGCCT GATGAACAGC CAGGAGAGCA
GCTGGGGCAA

6301 GAATGTGAGC AGCAGGGAGA TCACCAGGAC CACCCTGCAG TCTGACCAGG
AGGAGATTGA

6361 CTATGATGAC ACCATCTCTG TGGAGATGAA GAAGGAGGAC TTTGACATCT
ACGACGAGGA

6421 CGAGAACCAG AGCCCCAGGA GCTTCCAGAA GAAGACCAGG CACTACTTCA
TTGCTGCTGT

6481 GGAGAGGCTG TGGGACTATG GCATGAGCAG CAGCCCCCAT GTGCTGAGGA
ACAGGGCCCA

6541 GTCTGGCTCT GTGCCCCAGT TCAAGAAGGT GGTGTTCCAG GAGTTCCTG
ATGGCAGCTT

6601 CACCAGCCC CTGTACAGAG GGGAGCTGAA TGAGCACCTG GGCCTGCTGG
GCCCCTACAT

6661 CAGGGCTGAG GTGGAGGACA ACATCATGGT GACCTTCAGG AACCAGGCCA
GCAGGCCCTA

6721 CAGCTTCTAC AGCAGCCTGA TCAGCTATGA GGAGGACCAG AGGCAGGGGG
CTGAGCCCAG

6781 GAAGA ACTTT GTGAAGCCCA ATGAAACCAA GACCTACTTC TGGAAGGTGC
AGCACCACAT

6841 GGCCCCCACC AAGGATGAGT TTGACTGCAA GGCCTGGGCC TACTTCTCTG
ATGTGGACCT

6901 GGAGAAGGAT GTGCACTCTG GCCTGATTGG CCCCCTGCTG GTGTGCCACA
CCAACACCCT

6961 GAACCCTGCC CATGGCAGGC AGGTGACTGT GCAGGAGTTT GCCCTGTTCT
TCACCATCTT

7021 TGATGAAACC AAGAGCTGGT ACTTCACTGA GAACATGGAG AGGAACTGCA
GGGCCCCCTG

7081 CAACATCCAG ATGGAGGACC CCACCTTCAA GGAGA ACTAC AGGTTCCATG
CCATCAATGG

7141 CTACATCATG GACACCCTGC CTGGCCTGGT GATGGCCCAG GACCAGAGGA
TCAGGTGGTA

7201 CCTGCTGAGC ATGGGCAGCA ATGAGA ACAT CCACAGCATC CACTTCTCTG
GCCATGTGTT

7261 CACTGTGAGG AAGAAGGAGG AGTACAAGAT GGCCCTGTAC AACCTGTACC
CTGGGGTGTT

7321 TGAGACTGTG GAGATGCTGC CCAGCAAGGC TGGCATCTGG AGGGTGGAGT
GCCTGATTGG

7381 GGAGCACCTG CATGCTGGCA TGAGCACCTT GTTCCTGGTG TACAGCAACA
AGTGCCAGAC

7441 CCCCTGGGC ATGGCCTCTG GCCACATCAG GGACTTCCAG ATCACTGCCT
CTGGCCAGTA

7501 TGGCCAGTGG GCCCCCAAGC TGGCCAGGCT GCACTACTCT GGCAGCATCA
ATGCCTGGAG

7561 CACCAAGGAG CCCTTCAGCT GGATCAAGGT GGACCTGCTG GCCCCCATGA
TCATCCATGG

7621 CATCAAGACC CAGGGGGCCA GGCAGAAGTT CAGCAGCCTG TACATCAGCC
AGTTCATCAT

7681 CATGTACAGC CTGGATGGCA AGAAGTGGCA GACCTACAGG GGCAACAGCA
CTGGCACCTT

7741 GATGGTGTTC TTTGGCAATG TGGACAGCTC TGGCATCAAG CACAACATCT
TCAACCCCC

7801 CATCATTGCC AGATACATCA GGCTGCACCC CACCCACTAC AGCATCAGGA
GCACCCTGAG

7861 GATGGAGCTG ATGGGCTGTG ACCTGAACAG CTGCAGCATG CCCCTGGGCA
TGGAGAGCAA

7921 GGCCATCTCT GATGCCCAGA TCACTGCCAG CAGCTACTTC ACCAACATGT
TTGCCACCTG

7981 GAGCCCCAGC AAGGCCAGGC TGCACCTGCA GGGCAGGAGC AATGCCTGGA
GGCCCCAGGT

8041 CAACAACCCC AAGGAGTGGC TGCAGGTGGA CTTCAGAAAG ACCATGAAGG
TGAAGGGGT

8101 GACCACCCAG GGGGTGAAGA GCCTGCTGAC CAGCATGTAT GTGAAGGAGT
TCCTGATCAG

8161 CAGCAGCCAG GATGGCCACC AGTGGACCCT GTTCTTCCAG AATGGCAAGG
TGAAGGTGTT

8221 CCAGGGCAAC CAGGACAGCT TCACCCCTGT GGTGAACAGC CTGGACCCCC
CCCTGCTGAC

8281 CAGATACCTG AGGATTCACC CCCAGAGCTG GGTGCACCAG ATTGCCCTGA
GGATGGAGGT

8341 GCTGGGCTGT GAGGCCCAGG ACCTGTACTG AGCGGCCGCG GGCCCAATCA
ACCTCTGGAT

8401 TACAAAATTT GTGAAAGATT GACTGGTATT CTTAACTATG TTGCTCCTTT
TACGCTATGT

8461 GGATACGCTG CTTAATGCC TTTGTATCAT GCTATTGCTT CCCGTATGGC
TTTCATTTTC

8521 TCCTCCTTGT ATAAATCCTG GTTGCTGTCT CTTTATGAGG AGTTGTGGCC
CGTTGTCAGG

8581 CAACGTGGCG TGGTGTGCAC TGTGTTTGCT GACGCAACCC CCACTGGTTG
GGGCATTGCC

8641 ACCACCTGTC AGCTCCTTTC CGGGACTTTC GCTTTCCCCC TCCCTATTGC
CACGGCGGAA

8701 CTCATCGCCG CCTGCCTTGC CCGCTGCTGG ACAGGGGCTC GGCTGTTGGG
CACTGACAAT

8761 TCCGTGGTGT TGTCGGGGAA ATCATCGTCC TTTCCTTGGC TGCTCGCCTG
TGTTGCCACC

8821 TGGATTCTGC GCGGGACGTC CTTCTGCTAC GTCCCTTCGG CCCTCAATCC
AGCGGACCTT

8881 CCTTCCCGCG GCCTGCTGCC GGCTCTGCGG CCTCTTCCGC GTCTTCGCCT
TCGCCCTCAG

8941 ACGAGTCGGA TCTCCCTTTG GGCCGCCTCC CCGCAAGCTT CGCACTTTTT
AAAAGAAAAG

9001 GGAGGACTGG ATGGGATTTA TTA CTCCGAT AGGACGCTGG CTTGTA ACTC
AGTCTCTTAC

9061 TAGGAGACCA GCTTGAGCCT GGGTGTTCGC TGGTTAGCCT AACCTGGTTG
GCCACCAGGG

9121 GTAAGGACTC CTTGGCTTAG AAAGCTAATA AACTTGCCTG CATTAGAGCT
CTTACGCGTC

9181 CCGGGCTCGA GATCCGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC
TAACTCCGCC

9241 CATCCCGCCC CTA ACTCCGC CCAGTTCCGC CCATTCTCCG CCCCATGGCT
GACTAATTTT

9301 TTTTATTTAT GCAGAGGCCG AGGCCGCCTC GGCCTCTGAG CTATTCCAGA
AGTAGTGAGG

9361 AGGCTTTTTT GGAGGCCTAG GCTTTTGCAA AAAGCTAACT TGTTTATTGC
AGCTTATAAT

9421 GGTTACAAAT AAAGCAATAG CATCACAAAT TTCACAAATA AAGCATTTTT
TTC ACTGCAT

9481 TCTAGTTGTG GTTTGTCCAA ACTCATCAAT GTATCTTATC ATGTCTGTCC
GCTTCCTCGC

9541 TACTGACTC GCTGCGCTCG GTCGTTCCGC TGC GGCGAGC GGTATCAGCT
CACTCAAAGG

9601 CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGA ACATG
TGAGCAAAAAG

9661 GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GCGGTTTTTC
CATAGGCTCC

9721 GCCCCCTGA CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA
AACCCGACAG

9781 GACTATAAAG ATACCAGGCG TTTCCCCTG GAAGCTCCCT CGTGCGCTCT
CCTGTTCCGA

9841 CCCTGCCGCT TACCGGATAC CTGTCCGCCT TTCTCCCTTC GGAAGCGTG
GCGCTTTCTC

9901 ATAGCTCACG CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT TCGCTCCAAG
CTGGGCTGTG

9961 TGCACGAACC CCCCGTTCAG CCCGACCGCT GCGCCTTATC CGGTA ACTAT
CGTCTTGAGT

10021 CCAACCCGGT AAGACACGAC TTATCGCCAC TGGCAGCAGC CACTGGTAAC
AGGATTAGCA

10081 GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC
TACGGCTACA

10141 CTAGAAGAAC AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC
GGAAAAAGAG

10201 TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG CCGTGGTTTT
TTTGTGTTGCA

10261 AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA TCCTTTGATC
TTTTCTACGG

10321 GGTCTGACGC TCAGTGGAAC GAAAACCTCAC GTTAAGGGAT TTTGGTCATG
AGATTATCAA

10381 AAAGGATCTT CACCTAGATC CTTTAAATT AAAAAATGAAG TTTTAAATCA
ATCTAAAGTA

10441 TATATGAGTA AACTTGGTCT GACAGTTAGA AAAACTCATC GAGCATCAAA
TGAAACTGCA

10501 ATTTATTCAT ATCAGGATTA TCAATACCAT ATTTTTGAAA AAGCCGTTTC
TGTAATGAAG

10561 GAGAAAACCTC ACCGAGGCAG TTCCATAGGA TGGCAAGATC CTGGTATCGG
TCTGCGATTC

10621 CGACTCGTCC AACATCAATA CAACCTATTA ATTTCCCCTC GTCAAAAATA
AGGTTATCAA

10681 GTGAGAAATC ACCATGAGTG ACGACTGAAT CCGGTGAGAA TGGCAACAGC
TTATGCATTT

10741 CTTTCCAGAC TTGTTCAACA GGCCAGCCAT TACGCTCGTC ATCAAAATCA
CTCGCATCAA

10801 CCAAACCGTT ATTCATTCGT GATTGCGCCT GAGCGAGACG AAATACGCGA
TCGCTGTAA

10861 AAGGACAATT ACAAACAGGA ATCGAATGCA ACCGGCGCAG GAACACTGCC
AGCGCATCAA

10921 CAATATTTTC ACCTGAATCA GGATATTCTT CTAATACCTG GAATGCTGTT
TTTCCGGGGA

10981 TCGCAGTGGT GAGTAACCAT GCATCATCAG GAGTACGGAT AAAATGCTTG
ATGGTCGGAA

11041 GAGGCATAAA TTCCGTCAGC CAGTTTAGTC TGACCATCTC ATCTGTAACA
TCATTGGCAA

11101 CGCTACCTTT GCCATGTTTC AGAAACAACCT CTGGCGCATC GGGCTTCCCA
TACAATCGAT

11161 AGATTGTCGC ACCTGATTGC CCGACATTAT CGCGAGCCCA TTTATACCCA
TATAAATCAG

11221 CATCCATGTT GGAATTTAAT CGCGGCCTAG AGCAAGACGT TTCCCGTTGA
ATATGGCTCA

11281 TAACACCCCT TGTATTACTG TTTATGTAAG CAGACAGTTT TATTGTTTAT
GATGATATAT

11341 TTTTATCTTG TGCAATGTAA CATCAGAGAT TTTGAGACAC AACAATTGGT
CGACGGATCC

**SEQ ID NO: 28: Плазмида F/HN-SIV-hCEF-HFVIII-N6-co, определенная на
фигуре 4D (pDNA1 pGM414)**

Длина: 11108; Тип молекулы: ДНК; Особенности локализации/Определители:
источник, 1..11108; тип молекулы, другая ДНК; примечание: pGM414; организм,
синтетическая конструкция

1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT
AAATCAATAT

61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT
ATATTGGCTC

121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT

181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA

241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCATTG ACGTCAATAA
TGACGTATGT

301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA

361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT

421 CAATGACGGT AAATGGCCCCG CCTGGCATTG TGCCCAGTAC ATGACCTTAC
GGGACTTTCC

481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC

GGTTTTGGCA

541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
TCCACCCCAT

601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
AATGTCGTAA

661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
GTGGGAGGTC

721 TATATAAGCA GAGCTCGCTG GCTTGTA ACT CAGTCTCTTA CTAGGAGACC
AGCTTGAGCC

781 TGGGTGTTTCG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
CCTTGGCTTA

841 GAAAGCTAAT AAAGTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCTT
CATTGACGCC

901 TCACTCTCTT GAACGGGAAT CTCCTTACT GGGTTCTCTC TCTGACCCAG
GCGAGAGAAA

961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
CAGCTGAGAA

1021 GGCGTCTGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
TTGGTGAGTA

1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGAAGTAGGAG
AGGCCGTAGC

1141 CGTAACTACT CTTGGGCAAG TAGGGCAGGC GGTGGGTACG CAATGGGGGC
GGCTACCTCA

1201 GCACTAAATA GGAGACAATT AGACCAATTT GAGAAAATAC GACTTCGCCC
GAACGGAAAG

1261 AAAAAGTACC AAATTAACA TTTAATATGG GCAGGCAAGG AGATGGAGCG
CTTCGGCCTC

1321 CATGAGAGGT TGTTGGAGAC AGAGGAGGGG TGTAAGAGAA TCATAGAAGT
CCTCTACCCC

1381 CTAGAACCAA CAGGATCGGA GGGCTTAAAA AGTCTGTTCA ATCTTGTGTG
CGTGCTATAT

1441 TGCTTGCACA AGGAACAGAA AGTGAAAGAC ACAGAGGAAG CAGTAGCAAC
AGTAAGACAA

1501 CACTGCCATC TAGTGGAAAA AGAAAAAAGT GCAACAGAGA CATCTAGTGG
ACAAAAGAAA

1561 AATGACAAGG GAATAGCAGC GCCACCTGGT GGCAGTCAGA ATTTTCCAGC
GCAACAACAA

1621 GGAAATGCCT GGGTACATGT ACCCTTGTCA CCGCGCACCT TAAATGCGTG
GGTAAAAGCA

1681 GTAGAGGAGA AAAAATTTGG AGCAGAAATA GTACCCATGT TTCAAGCCCT

ATCGAATTCC

1741 CGTTTGTGCT AGGGTTCTTA GGCTTCTTGG GGGCTGCTGG AACTGCAATG
GGAGCAGCGG

1801 CGACAGCCCT GACGGTCCAG TCTCAGCATT TGCTTGCTGG GATACTGCAG
CAGCAGAAGA

1861 ATCTGCTGGC GGCTGTGGAG GCTCAACAGC AGATGTTGAA GCTGACCATT
TGGGGTGTTA

1921 AAAACCTCAA TGCCCGCGTC ACAGCCCTTG AGAAGTACCT AGAGGATCAG
GCACGACTAA

1981 ACTCCTGGGG GTGCGCATGG AAACAAGTAT GTCATACCAC AGTGGAGTGG
CCCTGGACAA

2041 ATCGGACTCC GGATTGGCAA AATATGACTT GGTTGGAGTG GGAAAGACAA
ATAGCTGATT

2101 TGGAAAGCAA CATTACGAGA CAATTAGTGA AGGCTAGAGA ACAAGAGGAA
AAGAATCTAG

2161 ATGCCTATCA GAAGTTAACT AGTTGGTCAG ATTTCTGGTC TTGGTTCGAT
TTCTCAAAT

2221 GGCTTAACAT TTTAAAAATG GGATTTTTAG TAATAGTAGG AATAATAGGG
TTAAGATTAC

2281 TTTACACAGT ATATGGATGT ATAGTGAGGG TTAGGCAGGG ATATGTTCTT
CTATCTCCAC

2341 AGATCCATAT CCGCGGCAAT TTTAAAAGAA AGGGAGGAAT AGGGGGACAG
ACTTCAGCAG

2401 AGAGACTAAT TAATATAATA ACAACACAAT TAGAAATACA ACATTTACAA
ACCAAATTC

2461 AAAAAATTTT AAATTTTAGA GCCGCGGAGA TCTGTTACAT AACTTATGGT
AAATGGCCTG

2521 CCTGGCTGAC TGCCCAATGA CCCCTGCCA ATGATGTCAA TAATGATGTA
TGTTCCCATG

2581 TAATGCCAAT AGGGACTTTC CATTGATGTC AATGGGTGGA GTATTTATGG
TAACTGCCA

2641 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTATGCC CCTATTGATG
TCAATGATGG

2701 TAAATGGCCT GCCTGGCATT ATGCCAGTA CATGACCTTA TGGGACTTTC
CTACTTGGCA

2761 GTACATCTAT GTATTAGTCA TTGCTATTAC CATGGGAATT CACTAGTGGA
GAAGAGCATG

2821 CTTGAGGGCT GAGTGCCCCT CAGTGGGCAG AGAGCACATG GCCCACAGTC
CCTGAGAAGT

2881 TGGGGGGAGG GGTGGGCAAT TGAAGTGGTG CCTAGAGAAG GTGGGGCTTG

GGTAAACTGG

2941 GAAAGTGATG TGGTGTACTG GCTCCACCTT TTTCCCCAGG GTGGGGGAGA
ACCATATATA

3001 AGTGCAGTAG TCTCTGTGAA CATTCAAGCT TCTGCCTTCT CCCTCCTGTG
AGTTTGCTAG

3061 CCACCAATGC AGATTGAGCT GAGCACCTGC TTCTTCCTGT GCCTGCTGAG
GTTCTGCTTC

3121 TCTGCCACCA GGAGATACTA CCTGGGGGCT GTGGAGCTGA GCTGGGACTA
CATGCAGTCT

3181 GACCTGGGGG AGCTGCCTGT GGATGCCAGG TTCCCCCCA GAGTGCCCAA
GAGCTTCCCC

3241 TTCAACACCT CTGTGGTGTA CAAGAAGACC CTGTTTGTGG AGTTCAGTGA
CCACCTGTTC

3301 AACATTGCCA AGCCCAGGCC CCCCTGGATG GGCCTGCTGG GCCCCACCAT
CCAGGCTGAG

3361 GTGTATGACA CTGTGGTGAT CACCCTGAAG AACATGGCCA GCCACCCTGT
GAGCCTGCAT

3421 GCTGTGGGGG TGAGCTACTG GAAGGCCTCT GAGGGGGCTG AGTATGATGA
CCAGACCAGC

3481 CAGAGGGAGA AGGAGGATGA CAAGGTGTTC CCTGGGGGCA GCCACACCTA
TGTGTGGCAG

3541 GTGCTGAAGG AGAATGGCCC CATGGCCTCT GACCCCCTGT GCCTGACCTA
CAGCTACCTG

3601 AGCCATGTGG ACCTGGTGAA GGACCTGAAC TCTGGCCTGA TTGGGGCCCT
GCTGGTGTGC

3661 AGGGAGGGCA GCCTGGCCAA GGAGAAGACC CAGACCCTGC ACAAGTTCAT
CCTGCTGTTT

3721 GCTGTGTTTG ATGAGGGCAA GAGCTGGCAC TCTGAAACCA AGAACAGCCT
GATGCAGGAC

3781 AGGGATGCTG CCTCTGCCAG GGCCTGGCCC AAGATGCACA CTGTGAATGG
CTATGTGAAC

3841 AGGAGCCTGC CTGGCCTGAT TGGCTGCCAC AGGAAGTCTG TGTAAGTGGCA
TGTGATTGGC

3901 ATGGGCACCA CCCCTGAGGT GCACAGCATC TTCCTGGAGG GCCACACCTT
CCTGGTCAGG

3961 AACCACAGGC AGGCCAGCCT GGAGATCAGC CCCATCACCT TCCTGACTGC
CCAGACCCTG

4021 CTGATGGACC TGGGCCAGTT CCTGCTGTTC TGCCACATCA GCAGCCACCA
GCATGATGGC

4081 ATGGAGGCCT ATGTGAAGGT GGACAGCTGC CCTGAGGAGC CCCAGCTGAG

GATGAAGAAC

4141 AATGAGGAGG CTGAGGACTA TGATGATGAC CTGACTGACT CTGAGATGGA
TGTGGTGAGG

4201 TTTGATGATG ACAACAGCCC CAGCTTCATC CAGATCAGGT CTGTGGCCAA
GAAGCACCCC

4261 AAGACCTGGG TGCACTACAT TGCTGCTGAG GAGGAGGACT GGGACTATGC
CCCCCTGGTG

4321 CTGGCCCCTG ATGACAGGAG CTACAAGAGC CAGTACCTGA ACAATGGCCC
CCAGAGGATT

4381 GGCAGGAAGT ACAAGAAGGT CAGGTTCATG GCCTACACTG ATGAAACCTT
CAAGACCAGG

4441 GAGGCCATCC AGCATGAGTC TGGCATCCTG GGCCCCCTGC TGTATGGGGA
GGTGGGGGAC

4501 ACCCTGCTGA TCATCTTCAA GAACCAGGCC AGCAGGCCCT ACAACATCTA
CCCCCATGGC

4561 ATCACTGATG TGAGGCCCT GTACAGCAGG AGGCTGCCA AGGGGGTGAA
GCACCTGAAG

4621 GACTTCCCCA TCCTGCCTGG GGAGATCTT CAGTACAAGT GGACTGTGAC
TGTGGAGGAT

4681 GGCCCCACCA AGTCTGACCC CAGGTGCCTG ACCAGATACT ACAGCAGCTT
TGTGAACATG

4741 GAGAGGGACC TGGCCTCTGG CCTGATTGGC CCCCTGCTGA TCTGCTACAA
GGAGTCTGTG

4801 GACCAGAGGG GCAACCAGAT CATGTCTGAC AAGAGGAATG TGATCCTGTT
CTCTGTGTTT

4861 GATGAGAACA GGAGCTGGTA CCTGACTGAG AACATCCAGA GGTCCTGCC
CAACCCTGCT

4921 GGGGTGCAGC TGGAGGACCC TGAGTTCCAG GCCAGCAACA TCATGCACAG
CATCAATGGC

4981 TATGTGTTTG ACAGCCTGCA GCTGTCTGTG TGCCTGCATG AGGTGGCCTA
CTGGTACATC

5041 CTGAGCATTG GGGCCCAGAC TGAATCCTG TCTGTGTTCT TCTCTGGCTA
CACCTTCAAG

5101 CACAAGATGG TGTATGAGGA CACCCTGACC CTGTTCCCCT TCTCTGGGGA
GACTGTGTTC

5161 ATGAGCATGG AGAACCTGG CCTGTGGATT CTGGGCTGCC ACAACTCTGA
CTTCAGGAAC

5221 AGGGGCATGA CTGCCCTGCT GAAAGTCTCC AGCTGTGACA AGAACACTGG
GGACTACTAT

5281 GAGGACAGCT ATGAGGACAT CTCTGCCTAC CTGCTGAGCA AGAACAATGC

CATTGAGCCC

5341 AGGAGCTTCA GCCAGAACAG CAGGCACCCC AGCACCAGGC AGAAGCAGTT
CAATGCCACC

5401 ACCATCCCTG AGAATGACAT AGAGAAGACA GACCCATGGT TTGCCACCG
GACCCCATG

5461 CCCAAGATCC AGAATGTGAG CAGCTCTGAC CTGCTGATGC TGCTGAGGCA
GAGCCCCACC

5521 CCCCATGGCC TGAGCCTGTC TGACCTGCAG GAGGCCAAGT ATGAAACCTT
CTCTGATGAC

5581 CCCAGCCCTG GGGCCATTGA CAGCAACAAC AGCCTGTCTG AGATGACCCA
CTTCAGGCC

5641 CAGCTGCACC ACTCTGGGGA CATGGTGTTT ACCCCTGAGT CTGGCCTGCA
GCTGAGGCTG

5701 AATGAGAAGC TGGGCACCAC TGCTGCCACT GAGCTGAAGA AGCTGGACTT
CAAAGTCTCC

5761 AGCACCAGCA ACAACCTGAT CAGCACCATC CCCTCTGACA ACCTGGCTGC
TGGCACTGAC

5821 AACACCAGCA GCCTGGGCCC CCCAGCATG CCTGTGCACT ATGACAGCCA
GCTGGACACC

5881 ACCCTGTTTG GCAAGAAGAG CAGCCCCCTG ACTGAGTCTG GGGGCCCCCT
GAGCCTGTCT

5941 GAGGAGAACA ATGACAGCAA GCTGCTGGAG TCTGGCCTGA TGAACAGCCA
GGAGAGCAGC

6001 TGGGGCAAGA ATGTGAGCAG CAGGGAGATC ACCAGGACCA CCCTGCAGTC
TGACCAGGAG

6061 GAGATTGACT ATGATGACAC CATCTCTGTG GAGATGAAGA AGGAGGACTT
TGACATCTAC

6121 GACGAGGACG AGAACCAGAG CCCAGGAGC TTCCAGAAGA AGACCAGGCA
CTACTTCATT

6181 GCTGCTGTGG AGAGGCTGTG GGACTATGGC ATGAGCAGCA GCCCCATGT
GCTGAGGAAC

6241 AGGGCCCAGT CTGGCTCTGT GCCCCAGTTC AAGAAGGTGG TGTTCAGGA
GTTCCTGAT

6301 GGCAGCTTCA CCCAGCCCCT GTACAGAGGG GAGCTGAATG AGCACCTGGG
CCTGCTGGGC

6361 CCCTACATCA GGGCTGAGGT GGAGGACAAC ATCATGGTGA CCTTCAGGAA
CCAGGCCAGC

6421 AGGCCCTACA GCTTCTACAG CAGCCTGATC AGCTATGAGG AGGACCAGAG
GCAGGGGGCT

6481 GAGCCCAGGA AGAACTTTGT GAAGCCCAAT GAAACCAAGA CCTACTTCTG

GAAGGTGCAG

6541 CACCACATGG CCCCCACCAA GGATGAGTTT GACTGCAAGG CCTGGGCCTA
CTTCTCTGAT

6601 GTGGACCTGG AGAAGGATGT GCACTCTGGC CTGATTGGCC CCCTGCTGGT
GTGCCACACC

6661 AACACCCTGA ACCCTGCCCA TGGCAGGCAG GTGACTGTGC AGGAGTTTGC
CCTGTTCTTC

6721 ACCATCTTTG ATGAAACCAA GAGCTGGTAC TTCACTGAGA ACATGGAGAG
GAACTGCAGG

6781 GCCCCCTGCA ACATCCAGAT GGAGGACCCC ACCTTCAAGG AGAACTACAG
GTTCCATGCC

6841 ATCAATGGCT ACATCATGGA CACCCTGCCT GGCCTGGTGA TGGCCCAGGA
CCAGAGGATC

6901 AGGTGGTACC TGCTGAGCAT GGCAGCAAT GAGAACATCC ACAGCATCCA
CTTCTCTGGC

6961 CATGTGTTCA CTGTGAGGAA GAAGGAGGAG TACAAGATGG CCCTGTACAA
CCTGTACCCT

7021 GGGGTGTTTG AACTGTGGA GATGCTGCCC AGCAAGGCTG GCATCTGGAG
GGTGGAGTGC

7081 CTGATTGGGG AGCACCTGCA TGCTGGCATG AGCACCTGT TCCTGGTGTA
CAGCAACAAG

7141 TGCCAGACCC CCCTGGGCAT GGCCTCTGGC CACATCAGGG ACTTCCAGAT
CACTGCCTCT

7201 GGCCAGTATG GCCAGTGGGC CCCCAAGCTG GCCAGGCTGC ACTACTCTGG
CAGCATCAAT

7261 GCCTGGAGCA CCAAGGAGCC CTCAGCTGG ATCAAGGTGG ACCTGCTGGC
CCCCATGATC

7321 ATCCATGGCA TCAAGACCCA GGGGGCCAGG CAGAAGTTCA GCAGCCTGTA
CATCAGCCAG

7381 TTCATCATCA TGTACAGCCT GGATGGCAAG AAGTGGCAGA CCTACAGGGG
CAACAGCACT

7441 GGCACCCTGA TGGTGTCTT TGGCAATGTG GACAGCTCTG GCATCAAGCA
CAACATCTTC

7501 AACCCCCCA TCATTGCCAG ATACATCAGG CTGCACCCCA CCCACTACAG
CATCAGGAGC

7561 ACCCTGAGGA TGGAGCTGAT GGGCTGTGAC CTGAACAGCT GCAGCATGCC
CCTGGGCATG

7621 GAGAGCAAGG CCATCTCTGA TGCCAGATC ACTGCCAGCA GCTACTTCAC
CAACATGTTT

7681 GCCACCTGGA GCCCCAGCAA GGCCAGGCTG CACCTGCAGG GCAGGAGCAA

TGCCTGGAGG

7741 CCCAGGTCA ACAACCCCAA GGAGTGGCTG CAGGTGGACT TCCAGAAGAC
CATGAAGGTG

7801 ACTGGGGTGA CCACCCAGGG GGTGAAGAGC CTGCTGACCA GCATGTATGT
GAAGGAGTTC

7861 CTGATCAGCA GCAGCCAGGA TGGCCACCAG TGGACCCTGT TCTTCCAGAA
TGGCAAGGTG

7921 AAGGTGTTCC AGGGCAACCA GGACAGCTTC ACCCCTGTGG TGAACAGCCT
GGACCCCCC

7981 CTGCTGACCA GATACCTGAG GATTCACCCC CAGAGCTGGG TGCACCAGAT
TGCCCTGAGG

8041 ATGGAGGTGC TGGGCTGTGA GGCCAGGAC CTGTACTGAG CGGCCGCGGG
CCCAATCAAC

8101 CTCTGGATTA CAAAATTTGT GAAAGATTGA CTGGTATTCT TAACTATGTT
GCTCCTTTTA

8161 CGCTATGTGG ATACGCTGCT TTAATGCCTT TGTATCATGC TATTGCTTCC
CGTATGGCTT

8221 TCATTTTCTC CTCCTTGTAT AAATCCTGGT TGCTGTCTCT TTATGAGGAG
TTGTGGCCCG

8281 TTGTCAGGCA ACGTGGCGTG GTGTGCACTG TGTTTGCTGA CGCAACCCCC
ACTGGTTGGG

8341 GCATTGCCAC CACCTGTCAG CTCCTTCCG GGAATTCGC TTTCCCCCTC
CCTATTGCCA

8401 CGGCGGAACT CATCGCCGCC TGCCTTGCCC GCTGCTGGAC AGGGGCTCGG
CTGTTGGGCA

8461 CTGACAATTC CGTGGTGTG TCGGGGAAAT CATCGTCCTT TCCTTGGCTG
CTCGCCTGTG

8521 TTGCCACCTG GATTCTGCGC GGGACGTCCT TCTGCTACGT CCCTTCGGCC
CTCAATCCAG

8581 CGGACCTTCC TTCCCGCGGC CTGCTGCCGG CTCTGCGGCC TCTTCCGCGT
CTTCGCCTTC

8641 GCCCTCAGAC GAGTCGGATC TCCCTTTGGG CCGCCTCCCC GCAAGCTTCG
CACTTTTAA

8701 AAGAAAAGGG AGGACTGGAT GGGATTTATT ACTCCGATAG GACGCTGGCT
TGTAATCAG

8761 TCTCTACTA GGAGACCAGC TTGAGCCTGG GTGTTGCTG GTTAGCCTAA
CCTGGTTGGC

8821 CACCAGGGGT AAGGACTCCT TGGCTTAGAA AGCTAATAAA CTTGCCTGCA
TTAGAGCTCT

8881 TACGCGTCCC GGGCTCGAGA TCCGCATCTC AATTAGTCAG CAACCATAGT

CCCGCCCCTA

8941 ACTCCGCCCA TCCCGCCCCT AACTCCGCC AGTTCGCC ATTCTCGCC
CCATGGCTGA

9001 CTAATTTTTT TTATTTATGC AGAGGCCGAG GCCGCCTCGG CCTCTGAGCT
ATTCCAGAAG

9061 TAGTGAGGAG GCTTTTTTGG AGGCCTAGGC TTTTGCAAAA AGCTAACTTG
TTTATTGCAG

9121 CTTATAATGG TTACAAATAA AGCAATAGCA TCACAAATTT CACAAATAAA
GCATTTTTTT

9181 CACTGCATTC TAGTTGTGGT TTGTCCAAAC TCATCAATGT ATCTTATCAT
GTCTGTCCGC

9241 TTCCTCGCTC ACTGACTCGC TGCCTCGGT CGTTCGGCTG CGGCGAGCGG
TATCAGCTCA

9301 CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA
AGAACATGTG

9361 AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG
CGTTTTTCCA

9421 TAGGCTCCGC CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA
GGTGGCGAAA

9481 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCTGGA AGCTCCCTCG
TGCCTCTCC

9541 TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT CTCCCTTCGG
GAAGCGTGGC

9601 GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC
GCTCCAAGCT

9661 GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG
GTAACATCG

9721 TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA
CTGGTAACAG

9781 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT
GGCCTAACTA

9841 CGGCTACACT AGAAGAACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG
TTACCTTCGG

9901 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG
GTGGTTTTTT

9961 TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC
CTTTGATCTT

10021 TTCTACGGGG TCTGACGCTC AGTGGAACGA AACTCACGT TAAGGGATTT
TGGTCATGAG

10081 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTAAATTA AAATGAAGTT

TTAAATCAAT

10141 CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTAGAAA AACTCATCGA
GCATCAAATG

10201 AAAGTCAAT TTATTCATAT CAGGATTATC AATACCATAT TTTTGAAAA
GCCGTTTCTG

10261 TAATGAAGGA GAAAACAC CGAGGCAGTT CCATAGGATG GCAAGATCCT
GGTATCGGTC

10321 TGCGATTCCG ACTCGTCCAA CATCAATACA ACCTATTAAT TTCCCCTCGT
CAAAAATAAG

10381 GTTATCAAGT GAGAAATCAC CATGAGTGAC GACTGAATCC GGTGAGAATG
GCAACAGCTT

10441 ATGCATTTCT TTCCAGACTT GTTCAACAGG CCAGCCATTA CGCTCGTCAT
CAAAATCACT

10501 CGCATCAACC AAACCGTTAT TCATTCGTGA TTGCGCCTGA GCGAGACGAA
ATACGCGATC

10561 GCTGTAAAA GGACAATTAC AAACAGGAAT CGAATGCAAC CGGCGCAGGA
ACACTGCCAG

10621 CGCATCAACA ATATTTTCAC CTGAATCAGG ATATTCTTCT AATACCTGGA
ATGCTGTTTT

10681 TCCGGGGATC GCAGTGGTGA GTAACCATGC ATCATCAGGA GTACGGATAA
AATGCTTGAT

10741 GGTCGGAAGA GGCATAAATT CCGTCAGCCA GTTTAGTCTG ACCATCTCAT
CTGTAACATC

10801 ATTGGCAACG CTACCTTTGC CATGTTTCAG AAACAACCTCT GGCGCATCGG
GCTTCCATA

10861 CAATCGATAG ATTGTCGCAC CTGATTGCCC GACATTATCG CGAGCCCATT
TATACCCATA

10921 TAAATCAGCA TCCATGTTGG AATTTAATCG CGGCCTAGAG CAAGACGTTT
CCCGTTGAAT

10981 ATGGCTCATA ACACCCCTTG TATTAATGTT TATGTAAGCA GACAGTTTTA
TTGTTTATGA

11041 TGATATATTT TTATCTTGTG CAATGTAACA TCAGAGATTT TGAGACACAA
CAATTGGTTCG

11101 ACGGATCC

SEQ ID NO: 29: Репрезентативный промотор CAG

Длина: 1738; Тип молекулы: ДНК; Особенности локализации/Определители:
источник, 1..1738; тип молекулы, другая ДНК; примечание: промотор CAG; организм,
синтетическая конструкция

ATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
AGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGA

CCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCCAC
TTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGAC
GGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACT
TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAGGTGAGCCCCACG
TTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCCAATTTTGTATTTATTTAT
TTTTTAATTATTTTGTGACGCGATGGGGGCGGGGGGGGGGGGGGGGGGGCGCGCGCCAG
GCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGGCGG
CAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGGCGGGCGG
CGGCGGCCCTATAAAAAGCGAAGCGCGCGGGCGGGGAGTCGCTGCGCGCTGCCT
TCGCCCCGTGCCCGCTCCGCCCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGAC
CGCGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTA
GCGCTTGGTTTAAATGACGGCTTGTTCCTTTCTGTGGCTGCGTGAAAGCCTTGAGGG
GCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTG
TGTGCGTGGGGAGCGCCGCGTGC GGCTCCGCGCTGCCCGGCGGCTGTGAGCGCTGC
GGGCGCGGCGCGGGGCTTTGTGCGCTCCGCAGTGTGCGCGAGGGGAGCGCGGCCGG
GGGCGGTGCCCGCGGTGCGGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGG
GGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGTCGGTCGGGCTGCAACC
CCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCT
CCGTACGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGG
GGGTGCCGGGCGGGGCGGGGCCCGCCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCG
CGGCGGCCCCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGCAGCCATTGCC
TTTTATGGTAATCGTGCGAGAGGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGG
AGCCGAAATCTGGGAGGCGCCGCGCACCCCCCTTAGCGGGCGCGGGGCGAAGCGG
TGCGGCGCCGCGCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCG
CCGTCCCCTTCTCCCTCTCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
GGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAG
CCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCTCCTGGGCAACGTGCT
GGTTATTGTGCTGTCTCATATTTTGGCAAAGAATTGCTCGAGCCACC.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ получения ретровирусного вектора, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, и содержащего промотор и трансген, где указанный способ включает использование оптимизированных по кодонам генов gag-pol.
2. Способ по п.1, где ретровирусный вектор представляет собой лентивирусный вектор.
3. Способ по п.2, где лентивирусный вектор выбран из группы, состоящей из вектора вируса иммунодефицита обезьян (SIV), вектора вируса иммунодефицита человека (ВИЧ), вектора вируса иммунодефицита кошек (FIV), вектора вируса инфекционной анемии лошадей (EIAV) и вектора вируса Висны/maedi.
4. Способ по п.2 или 3, где лентивирусный вектор представляет собой вектор SIV.
5. Способ по любому из пп. 1-4, где оптимизированные по кодонам гены gag-pol представляют собой гены gag-pol SIV.
6. Способ по любому из пп. 1-5, где оптимизированные по кодонам гены gag-pol содержат, или состоят из нее, последовательность нуклеиновой кислоты, которая по меньшей мере на 80% идентична SEQ ID NO: 1.
7. Способ по п.6, где оптимизированные по кодонам гены gag-pol содержат, или состоят из нее, последовательность нуклеиновой кислоты SEQ ID NO: 1.
8. Способ по любому из пп. 1-7, где оптимизированные по кодонам гены gag-pol содержатся в плазмиде, которая включает, или состоит из нее, последовательность нуклеиновой кислоты, которая по меньшей мере на 80% идентична SEQ ID NO: 5.
9. Способ по п.8, где оптимизированные по кодонам гены gag-pol содержатся в плазмиде, которая включает, или состоит из нее, последовательность нуклеиновой кислоты SEQ ID NO: 5.
10. Способ по любому из пп. 1-9, где респираторный парамиксовирус представляет собой вирус Сендай.
11. Способ по любому из пп. 1-10, где титр полученного ретровирусного вектора: эквивалентен титру ретровирусного вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодонам гены gal-pol; или повышен по сравнению с титром ретровирусного вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодонам гены gal-pol.
12. Способ по п.11, где титр ретровирусного вектора по меньшей мере в 2 раза или по меньшей мере в 2,5 раза превышает титр ретровирусного вектора, полученного соответствующим способом, в котором не используются оптимизированные по кодонам гены gal-pol.
13. Способ по любому из пп. 1-12, где промотор выбран из группы, состоящей из промотора цитомегаловируса (CMV), промотора фактора элонгации 1a (EF1a) и гибридного промотора человеческого CMV-энхансера/EF 1a (hCEF).

14. Способ по любому из пп. 1-13, где вектор содержит гибридный промотор человеческого CMV-энхансера/EF1a (hCEF).

15. Способ по любому из пп. 1-14, где трансген выбран из:

секретируемого терапевтического белка, необязательно альфа-1-антитрипсина (A1AT), фактора VIII, поверхностно-активного белка В (SFTPВ), фактора VII, фактора IX, фактора X, фактора XI, фактора фон Виллебранда, гранулоцитарного-макрофагального колониестимулирующего фактора (GM-CSF) и моноклонального антитела против инфекционного агента; или

CFTR, ABCA3, DNAH5, DNAH11, DNAI1 и DNAI2.

16. Способ по любому из пп. 1-15, где трансген кодирует:

CFTR;

A1AT; или

FVIII.

17. Способ по любому из пп. 1-16, где:

промотор представляет собой промотор hCEF, а трансген кодирует CFTR;

промотор представляет собой промотор hCEF, а трансген кодирует A1AT; или

промотор представляет собой промотор hCEF или CMV, а трансген кодирует

FVIII.

18. Способ по любому из пп. 1-17, где указанный способ включает следующие стадии:

выращивание клеток в суспензии;

трансфекцию клеток одной или более плазмидами;

добавление нуклеазы;

сбор лентивируса;

добавление трипсина; и

очистку.

19. Способ по п.18, где одна или более плазмид содержат или состоят из:

векторной геномной плазмиды, предпочтительно выбранной из pGM830 и pGM326;

плазмиды со-galpol, предпочтительно pGM691;

плазмиды Rev, предпочтительно pGM299;

плазмиды слитого (F) белка, предпочтительно pGM301; и

плазмиды гемагглютинаина-нейраминидазы (HN), предпочтительно pGM303.

20. Способ по п.19, где отношение «плазида векторного генома : плазида со-galpol : плазида Rev : плазида F: плазида HN» составляет 20:9:6:6:6.

21. Способ по любому из пп. 18-20, где стадии (a)-(f) осуществляют последовательно.

22. Способ по любому из пп. 18-21, где клетки представляют собой клетки НЕК293Т или 293Т/17.

23. Способ по любому из пп. 18-22, где добавление нуклеазы осуществляют на

стадии перед сбором.

24. Способ по любому из пп. 18-23, где добавление трипсина осуществляют после стадии сбора.

25. Способ по любому из пп. 18-24, где стадия очистки включает стадию хроматографии.

26. Способ по любому из пп. 19-24, где плазида векторного генома модифицирована для уменьшения количества ОРС ретровируса.

27. Нуклеиновая кислота, содержащая оптимизированные по кодонам гены gag-pol, где указанная нуклеиновая кислота по меньшей мере на 80% идентична последовательности SEQ ID NO: 1.

28. Нуклеиновая кислота по п.27, которая содержит, или состоит из нее, последовательность нуклеиновой кислоты SEQ ID NO: 1.

29. Плазида, содержащая нуклеиновую кислоту по п.27 или 28, где, необязательно:

плазида содержит, или состоит из нее, последовательность нуклеиновой кислоты, которая по меньшей мере на 80% идентична последовательности SEQ ID NO: 5; или

плазида содержит, или состоит из нее, последовательность нуклеиновой кислоты SEQ ID NO: 5.

30. Клетка-хозяин, содержащая нуклеиновую кислоту по п.27 или 28 и/или плазмиду по п.29.

31. Ретровирусный вектор, псевдотипированный гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, где указанный вектор может быть получен способом по любому из пп. 1-26.

32. Способ лечения заболевания, включающий введение индивидууму, нуждающемуся в этом, ретровирусного вектора, псевдотипированного гемагглютинином-нейраминидазой (HN) и слитыми (F) белками респираторного парамиксовируса, где указанный вектор может быть получен способом по любому из пп. 1-26.

33. Способ лечения по п.32, где заболевание, подлежащее лечению, представляет собой заболевание легких, предпочтительно, кистозный фиброз.

По доверенности

ФИГ. 1

Программа выравнивания множества последовательностей ClustalW

2 выровненных последовательности

Время обработки: 1,9 секунд

Вставленные пробелы = 0

Оценка = 199998

Консервативные идентичности = 3375

Режим попарного выравнивания: Медленный

Параметры попарного выравнивания:

Штраф за пробел-пропуск = 15,0 Штраф за пробел-удлинение = 6,7

Параметры множественного выравнивания:

Штраф за пробел-пропуск = 15,0

Штраф за пробел-удлинение = 6,7

Дивергентное снижение = 30%

Транзиции: Взвешенные

1. GagPol pGM297 соGagPol pGM691

Длина выровненной последовательности = 4391, Сегменты с пробелами = 0, Общие несоответствия (включая пробелы) = 1016,

Идентичности = 3375 (76,9%)

GagPol pGM297	1	ATGGGGGGCGGCTACCTCAGCACTAAATAGGAGACAATTAGACCAATTTGAGAAAAATACGA	60
соGagPol pGM691	1	ATGGGAGCTGCCACATCTGCCCTGAATAGACGGCAGCTGGACCAGTTCGAGAAGATCAGA	60
		***** ** ** ** **	
GagPol pGM297	61	CTTCGCCCGAACGGAAAGAAAAAGTACCAAATTAACATTTAATATGGGCAGGCCAAGGAG	120
соGagPol pGM691	61	CTGCGGCCCAACGGCAAGAAGAAGTACCAGATCAAGCACCTGATCTGGGCCGGCAAGAG	120
		** ** ** ***** ***** ***** ** ** ** * ** ***** ***** **	
GagPol pGM297	121	ATGGAGCGCTTCGGCCTCCATGAGAGGTTGTTGGAGACAGAGGAGGGGTGTAAGAAGATC	180
соGagPol pGM691	121	ATGGAAGATTCGGCCTGCACGAGCGGCTGCTGGAACCGAGGAAGGCTGCAAGAGAATT	180
		***** * ***** ** ** ** **	
GagPol pGM297	181	ATAGAAGTCTCTACCCCTAGAACCAACAGGATCGGAGGGCTTAAAAAGTCTGTCAAT	240
соGagPol pGM691	181	ATCGAGGTGCTGTACCCCTGGAACCTACCGGCTCTGAGGGCCTGAAGTCCCTGTCAAT	240
		** ** ** ** ***** ** ***** ** ** ** ***** * ** *****	
GagPol pGM297	241	CTTGTGTGCGTACTATATGCTTGCACAAGGAACAGAAAGTGAAGACACAGAGGAAGCA	300
соGagPol pGM691	241	CTCGTGTGCGTGTACTGCTGCACAAGGAACAGAAAGTGAAGACACCGAAGAGGCC	300
		** ***** ** ** ** ***** ***** ***** ***** ** ** **	
GagPol pGM297	301	GTAGCAACAGTAAGACAACACTGCCATCTAGTGGAAAAAGAAAAAGTGCACAGAGACA	360
соGagPol pGM691	301	GTGGCCACAGTTAGACAGCACTGCCACCTGGTGGAAAAAGAGAAGTCCGCCACAGAGACA	360
		** ** ***** ***** ***** ** ***** ** ** *****	
GagPol pGM297	361	TCTAGTGGACAAAAGAAAAATGACAAGGGGAATAGCAGCGCCACCTGGTGGCAGTCAGAAT	420
соGagPol pGM691	361	AGCAGCGGCCAGAAGAAGACGACAGGGGAATGCTGCCCTCCTGGCGGCAGCCAGAAT	420
		* ** ** ***** ** ***** ***** ** ** ***** ***** *****	
GagPol pGM297	421	TTTCCAGCGCAACAACAAGGAAATGCCTGGGTACATGTACCCCTTGTACC CGCACCTTA	480
соGagPol pGM691	421	TTTCCTGCTCAGCAGCAGGAAACCGCTGGGTGCACGTTCCACTGAGCCCTAGAACA CTG	480
		***** ** ** ** ***** ***** ***** ** ** ** **	
GagPol pGM297	481	AATGCGTGGGTAAAAGCAGTAGAGGAGAAAAAATTTGGAGCAGAAATAGTACCCATGTTT	540
соGagPol pGM691	481	AATGCCTGGGTCAAAGCCGTGGAAGAGAAGAAGTTTGGCGCCGAGATCGTGCCCATGTTT	540
		***** ***** ***** ** ** ***** ** ***** ** ** ** *****	
GagPol pGM297	541	CAAGCCCTATCAGAAGGCTGCACACCCTATGACATTAATCAGATGCTTAATGTGCTAGGA	600
соGagPol pGM691	541	CAGGCTCTGTCTGAGGGCTGCACCCTTACGACATCAACCAGATGCTGAACGTGCTGGGA	600
		** ** ** ** ***** ** ** ***** ** ***** ** ***** **	
GagPol pGM297	601	GATCATCAAGGGGCATTACAAATAGTGAAGAGATCATTAATGAAGAAGCAGCCAGTGG	660
соGagPol pGM691	601	GATCACCAGGGCGCTCTGCAGATCGTGAAGAGATCATCAACGAAGAGGCTGCCAGTGG	660
		***** ** ** ** * ** ** ***** ***** ** ***** ** *****	
GagPol pGM297	661	GATGTAACACACCCCACTACCCCGAGGACCCTACCAGCAGGACAGCTCAGGGACCCTCGC	720
соGagPol pGM691	661	GACGTGACACATCCATGCTGCTGGACCTCTGCCAGCCGGACA ACTGAGAGATCCTAGA	720
		** ** ***** ** * ** ** ***** ** ***** ***** ** ** ** **	
GagPol pGM297	721	GGCTCAGATATAGCAGGGACCACCAGCTCAGTACAAGAACAGTTAGAATGGATCTATACT	780
соGagPol pGM691	721	GGCTCTGATATCGCCGGCACCACCAGCTCTGTGCAAGAGCAGCTGGAATGGATCTACACC	780
		***** ***** ** ** ***** ***** ** ***** ** * ***** **	
GagPol pGM297	781	GCTAACCCCGGGTAGATGTAGGTGCCATCTACCGGAGATGGATTAATCTAGGACTTCAA	840
соGagPol pGM691	781	GCCAACTTAGAGTGGACGTGGCGCCATCTACAGAAGATGGATCATCTCTGGCCTCGCAG	840

*** ** * ** ** ** * ** ** ** **

GagPol pGM297 841 AAGTGTGTCAAATGTACAACCCAGTATCAGTCCCTAGACATTAGGCAGGGACCTAAAGAG 900
coGagPol pGM691 841 AAATGCGTGAAGATGTACAACCCCGTGTCCGTGCTGGACATCAGACAGGGACCCAAAGAG 900
** * ** * ** ** ** **

GagPol pGM297 901 CCCTTCAAGGATTATGTGGACAGATTTTACAAGGCAATTAGAGCAGAACAAGCCCTCAGGG 960
coGagPol pGM691 901 CCCTTCAAGGACTACGTGGACCGGTTCTATAAGGCCATTAGAGCCGAGCAGGCCAGCGGC 960
***** ** ** * ** ** ** * ** **

GagPol pGM297 961 GAAGTGAACAATGGATGACAGAATCATTACTCATTCAAAATGCTAATCCAGATTGTAAG 1020
coGagPol pGM691 961 GAAGTGAAGCAGTGGATGACAGAGAGCCTGCTGATCCAGAACGCCAATCCAGACTGCAA 1020
***** ** ** * ** ** * ** ** * ** **

GagPol pGM297 1021 GTCATCCTGAAGGGCCTAGGAATGCACCCACCCTTGAAGAAATGTTAACGGCTTGTGAC 1080
coGagPol pGM691 1021 GTGATCCTGAAAGGCCTGGGCATGCACCCACACTGGAAGAGATGCTGACAGCCTGTCAA 1080
** ***** ** * ** ** * ** ** * ** **

GagPol pGM297 1081 GGGGTAGGAGGCCCAAGCTACAAGCAAAGTAAATGGCAGAAATGATGCAGACCATGCAA 1140
coGagPol pGM691 1081 GGCGTTGGCGGCCCTTCTTACAAGCCAAAGTGAATGGCCGAGATGATGCAGACCATGCA 1140
** * ** * ** ** * ** ** * ** ** * ** **

GagPol pGM297 1141 AATCAAACATGGTGCAGCAGGGAGGTCCAAAAGACAAAGACCCCCACTAAGATGTTAT 1200
coGagPol pGM691 1141 AACCAGAACATGGTGCAGCAAGCGGCCCTAAGAGACAGAGGCCTCCTCTGAGATGCTAC 1200
** * ** * ** ** * ** ** * ** ** * ** **

GagPol pGM297 1201 AATTGTGGAAAATTTGGCCATATGCAAAGACAATGTCCGGAACCAAGGAAAAACAAATGT 1260
coGagPol pGM691 1201 AACTGCGGCAAGTTCGGCCACATGCAGAGACAGTGTCTCTGAGCCTAGGAAAAACAAATGT 1260
** * ** * ** ** * ** ** * ** ** * ** **

GagPol pGM297 1261 CTAAAGTGTGGAAAATTTGGGACACCTAGCAAAGACTGCAGGGGACAGGTGAATTTTTTA 1320
coGagPol pGM691 1261 CTAAAGTGTGGAAAATTTGGGACACCTAGCAAAGACTGCAGGGGACAGGTGAATTTTTTA 1320
***** ** ** * ** ** * ** **

GagPol pGM297 1321 GGGTATGGACGGTGGATGGGGGCAAACCGAGAAATTTCCCGCCGCTACTCTTGGAGCG 1380
coGagPol pGM691 1321 GGGTATGGACGGTGGATGGGGGCAAACCGAGAAATTTCCCGCCGCTACTCTTGGAGCG 1380
***** ** ** * ** ** * ** **

GagPol pGM297 1381 GAACCGAGTGCGCCCTCCTCCACCGAGCGGCACACCCCATACGACCCAGCAAAGAAGCTC 1440
coGagPol pGM691 1381 GAACCGAGTGCGCCCTCCTCCACCGAGCGGCACACCCCATACGACCCAGCAAAGAAGCTC 1440
***** ** ** * ** ** * ** **

GagPol pGM297 1441 CTGCAGCAATATGCAGAGAAAGGGAAACAACCTGAGGGAGCAAAGAGGAATCCACCGGCA 1500
coGagPol pGM691 1441 CTGCAGCAATATGCAGAGAAAGGGAAACAACCTGAGGGAGCAAAGAGGAATCCACCGGCA 1500
***** ** ** * ** ** * ** **

GagPol pGM297 1501 ATGAATCCGGATTGGACCGAGGGATATTCTTTGAACCTCCCTCTTTGGAGAAGACCAATA 1560
coGagPol pGM691 1501 ATGAATCCGGATTGGACCGAGGGATATTCTTTGAACCTCCCTCTTTGGAGAAGACCAATA 1560
***** ** ** * ** ** * ** **

GagPol pGM297 1561 AGACAGTGTATATAGAAGGGGTCCCCATTAAAGGCACTGCTAGACACAGGGGCAGATGACA 1620
coGagPol pGM691 1561 AGACCGTGTACATCGAGGGCGTGCCATCAAGGCTCTGCTGGATACAGGCCCGCAGGACA 1620
**** * ** * ** * ** * ** * ** * ** * ** * ** * **

GagPol pGM297 1621 CCATAATTAAGAAAATGATTTACAATTATCAGGTCCATGGAGACCCAAAATATAGGGG 1680
coGagPol pGM691 1621 CCATCATCAAAGAGAACGACCTGCAGCTGAGCGGCCCTTGGAGGCCAATGATCATTGGAG 1680
**** * ** * ** * ** * ** * ** * ** * ** * ** * **

GagPol pGM297 1681 GCATAGGAGGAGGCCCTTAATGTAAGAATATAACGACAGGGAAGTAAAAATAGAAGATA 1740
coGagPol pGM691 1681 GAATCGGCGGAGGCCTGAACGTCAAAGAGTACAACGACCGGGAAGTGAAGATCGAGGACA 1740
* ** * ** * ** ** * ** * ** * ** * ** * ** * ** * **

GagPol pGM297 1741 AAATTTTGAGAGGAACAATATTGTTAGGAGCAACTCCATTAAATATAATAGGTAGAAATT 1800
coGagPol pGM691 1741 AGATCCTGAGGGGCACAATCCTGCTGGGGCCACACCTATCAACATCATCGGCAGAAATC 1800
* ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * **

GagPol pGM297 1801 TGCTGGCCCCGCGCAGGTGCCCGGTTAGTAATGGGACAATTATCAGAAAAAATTCCTGTCA 1860
coGagPol pGM691 1801 TGCTGGCCCCGCGCGCGCTAGACTGGTTATGGGACAGCTCTCTGAGAAGATCCCGGTGA 1860
***** ** ** * ** ** * ** * ** * ** * ** * ** * ** * **

GagPol pGM297 1861 CACCTGTCAAATTGAAGGAAGGGGCTCGGGGACCCTGTGTAAGACAATGGCCTCTCTCTA 1920
coGagPol pGM691 1861 CACCCGTGAAGCTGAAAGAAGGCGCTAGAGGACCTTGTGTGCGACAGTGGCCTCTGAGCA 1920
**** * ** * ** * ** * ** * ** * ** * ** * ** * ** * **

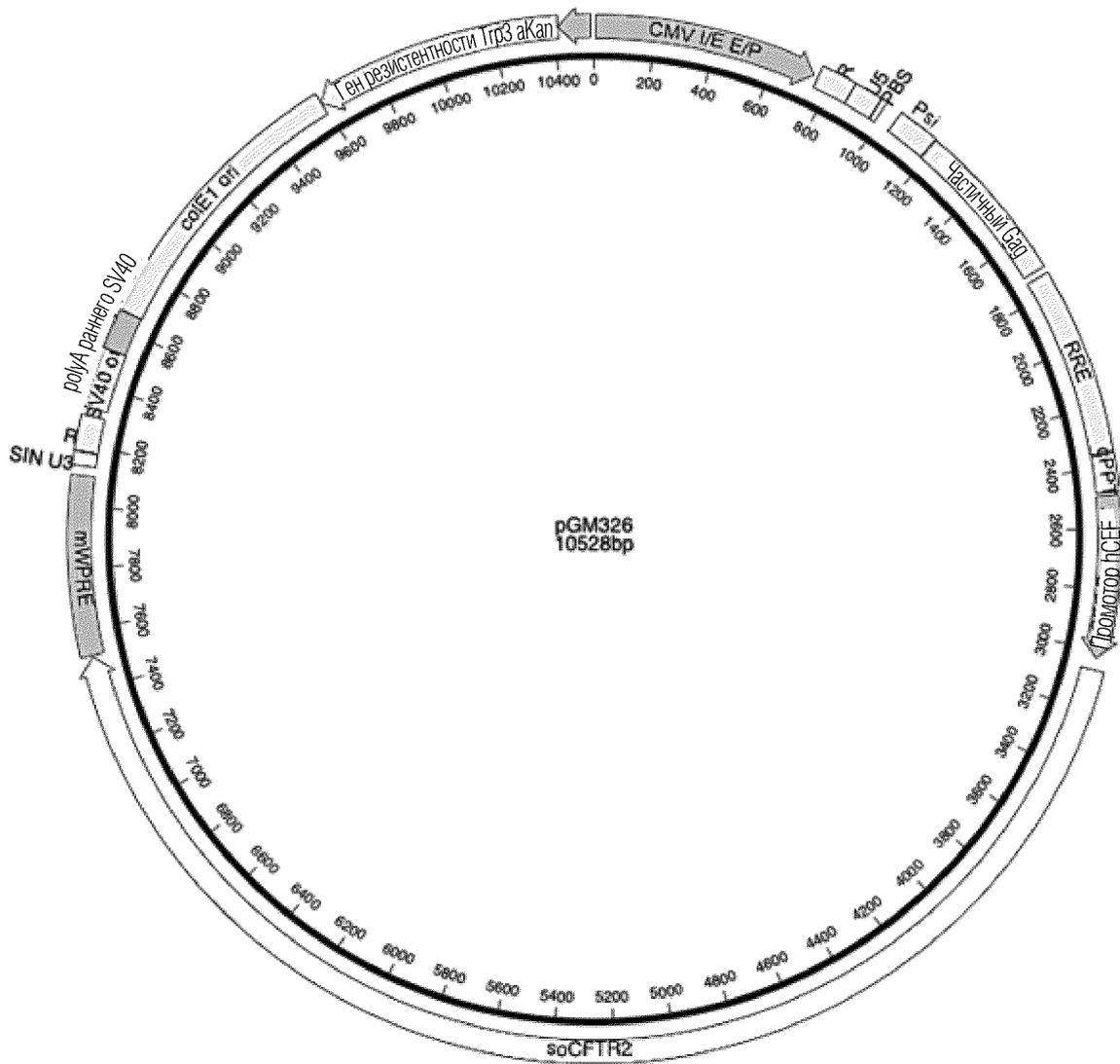
GagPol pGM297 1921 AAGAGAAGATTGAAGCTTTACAGGAAATATGTTCCCAATTAGAGCAGGAAGGAAAAATCA 1980
coGagPol pGM691 1921 AAGAGAAGATTGAGGCCCTGCAAGAAATCTGTAGCCAGCTGGAACAAAGAGGGCAAGATCA 1980
***** ** * ** * ** * ** * ** * ** * ** * ** * ** * **

coGagPol pGM691 4261 TGGTGCTGAAGGATGGCTCTGATCTGAAGGTGGTGCCACAGACGGAAGGCCAAGATTATCA 4320
***** ** ***** ** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

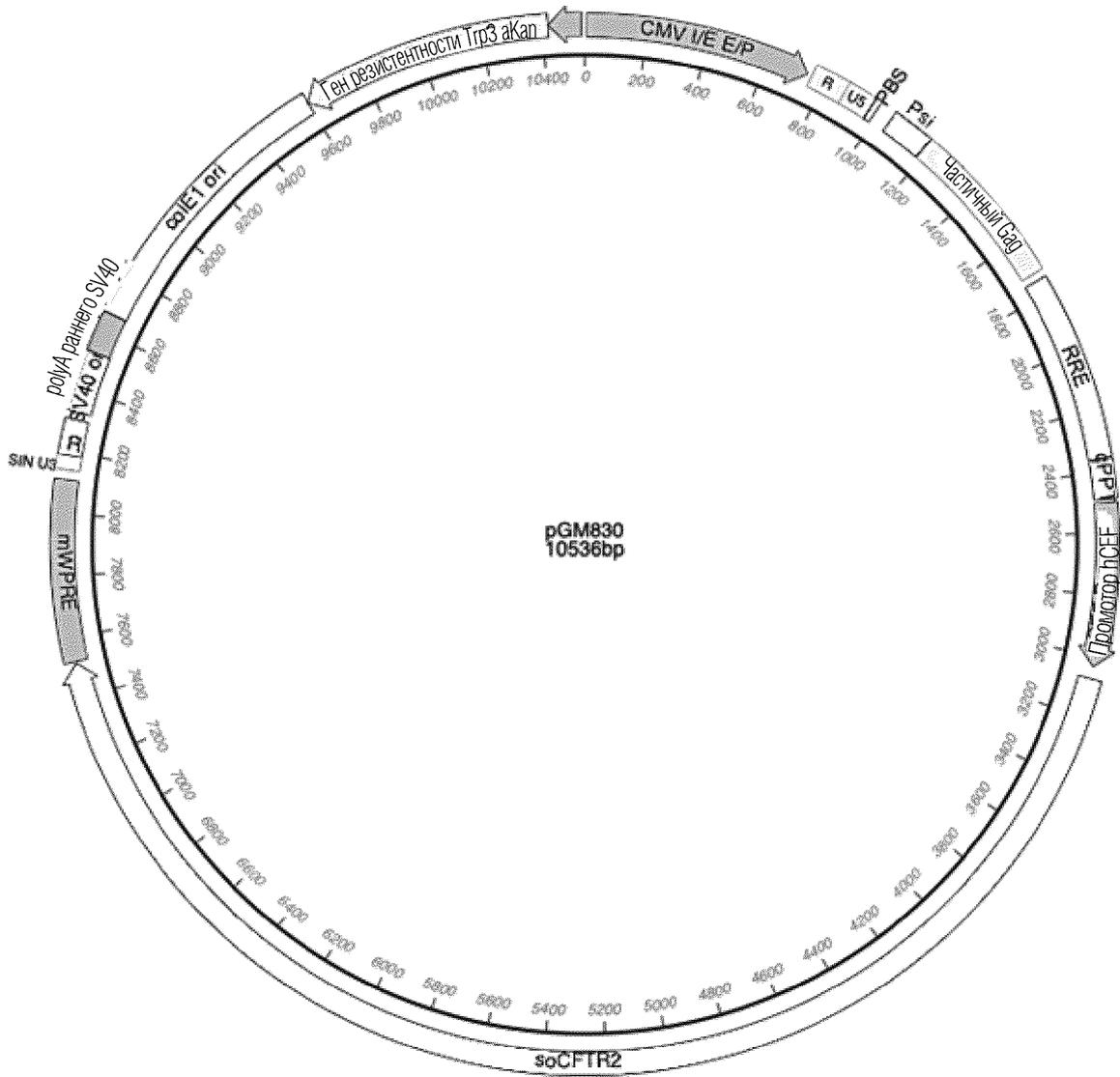
GagPol pGM297 4321 AGGATTATGAACCCAAACAAAGAGTGGGTAATGAGGGTGACGTGGAAGGTACCAGGGGAT 4380
coGagPol pGM691 4321 AGGATTACGAGCCCAACAGCGCGTGGGCAATGAAGGCGACGTTGAGGGCACAAGAGGCA 4380
***** ** ***** * ** * ** * ** * ** * ** * ** * ** * ** * ** * ** *

GagPol pGM297 4381 CTGATAACTAA 4391
coGagPol pGM691 4381 GCGACAATTGA 4391
* * * * *

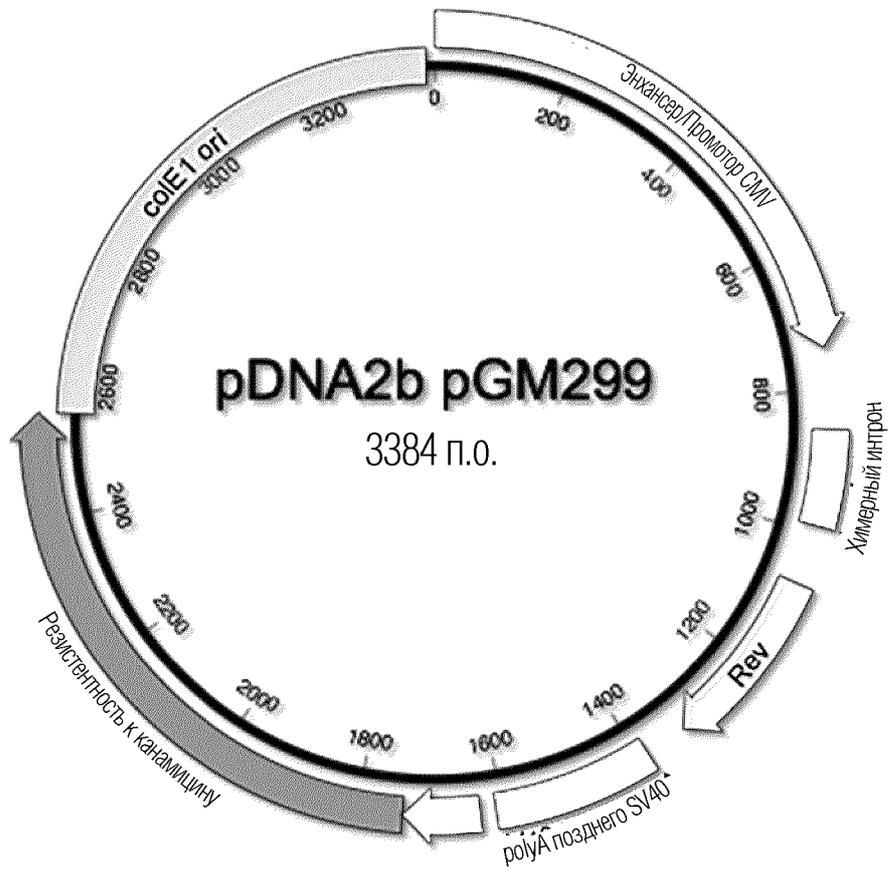
ФИГ. 2А



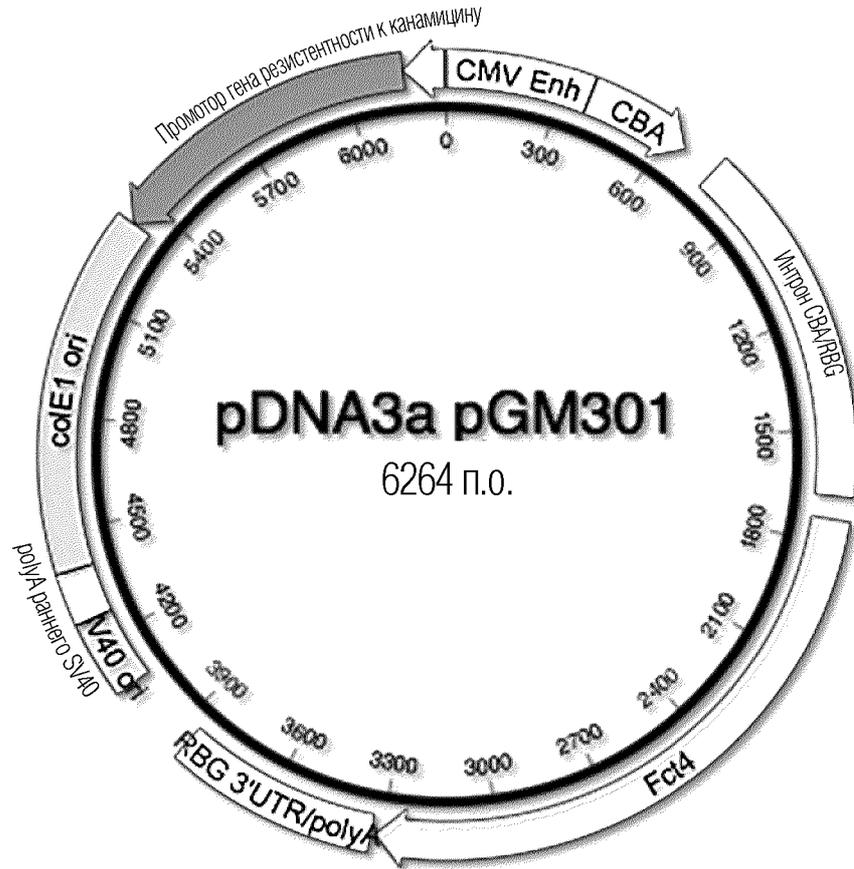
ФИГ. 2В



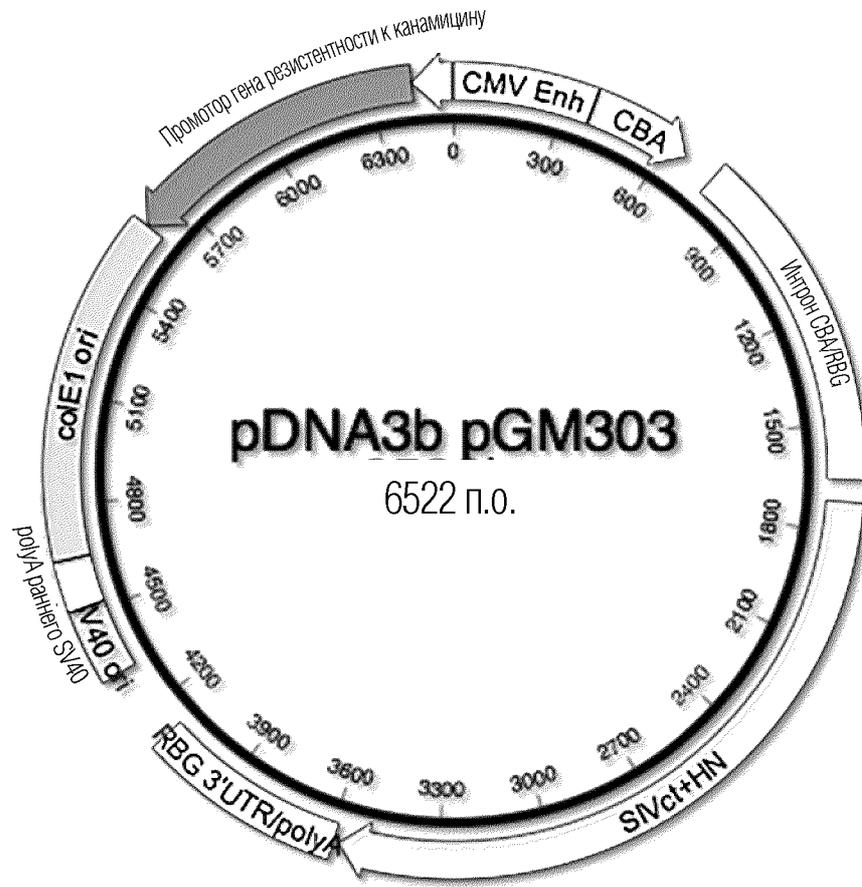
ФИГ. 2D



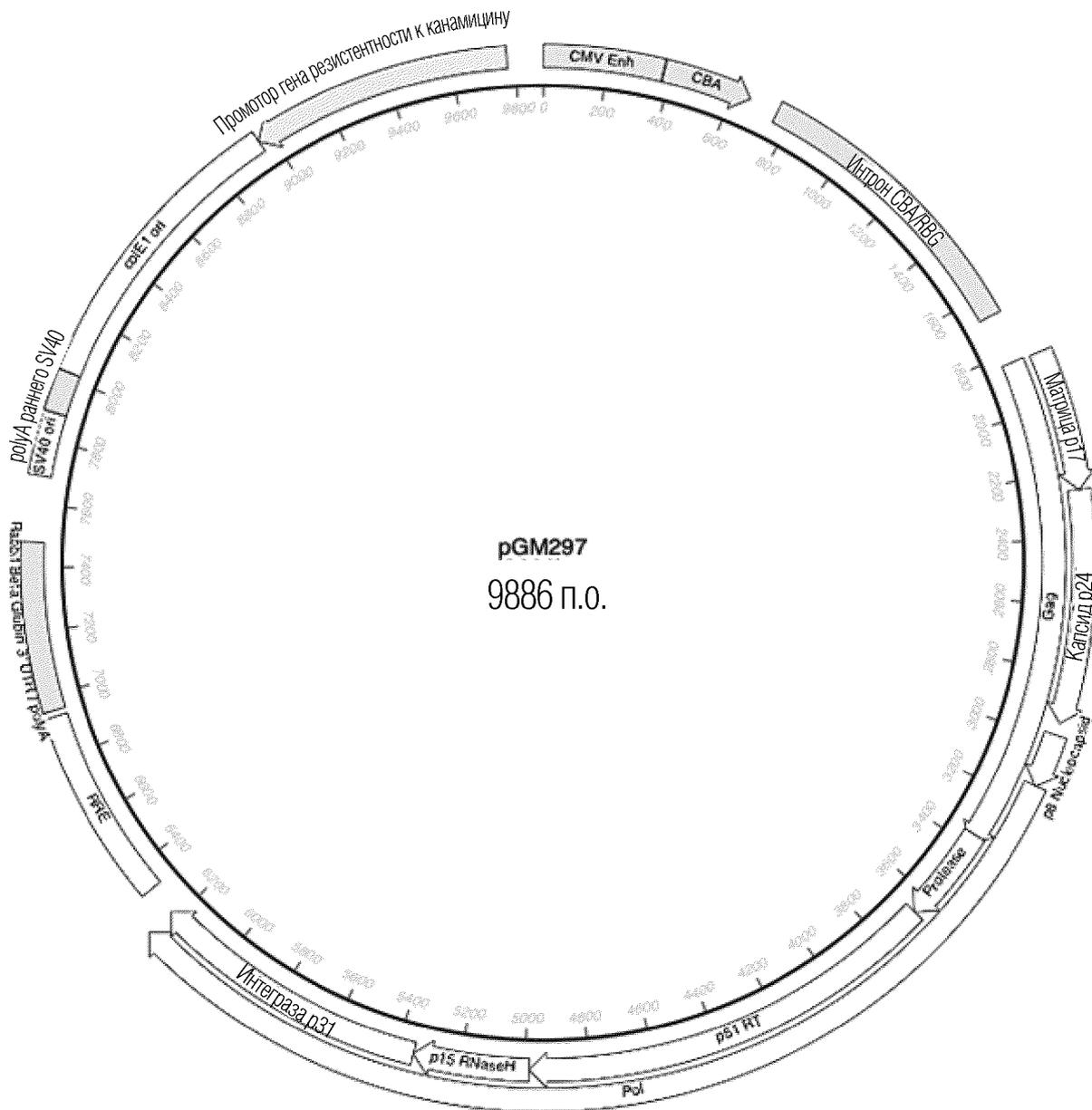
ФИГ. 2Е



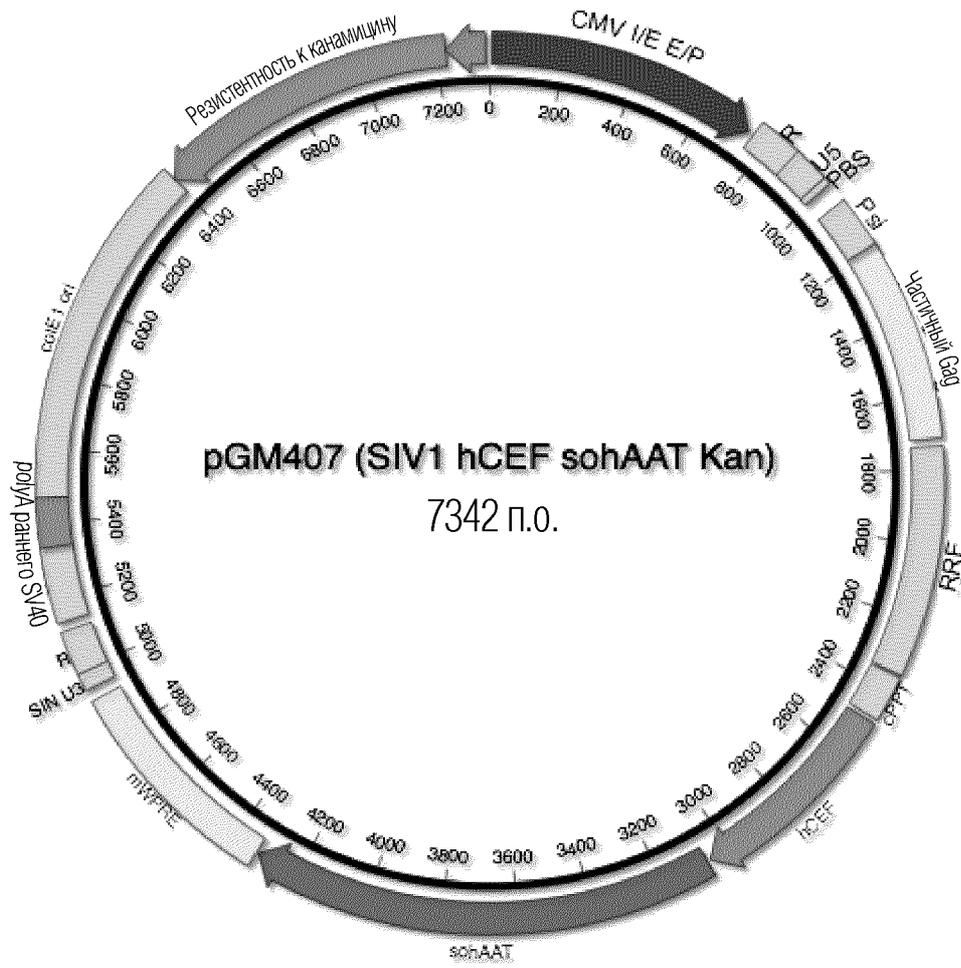
ФИГ. 2F



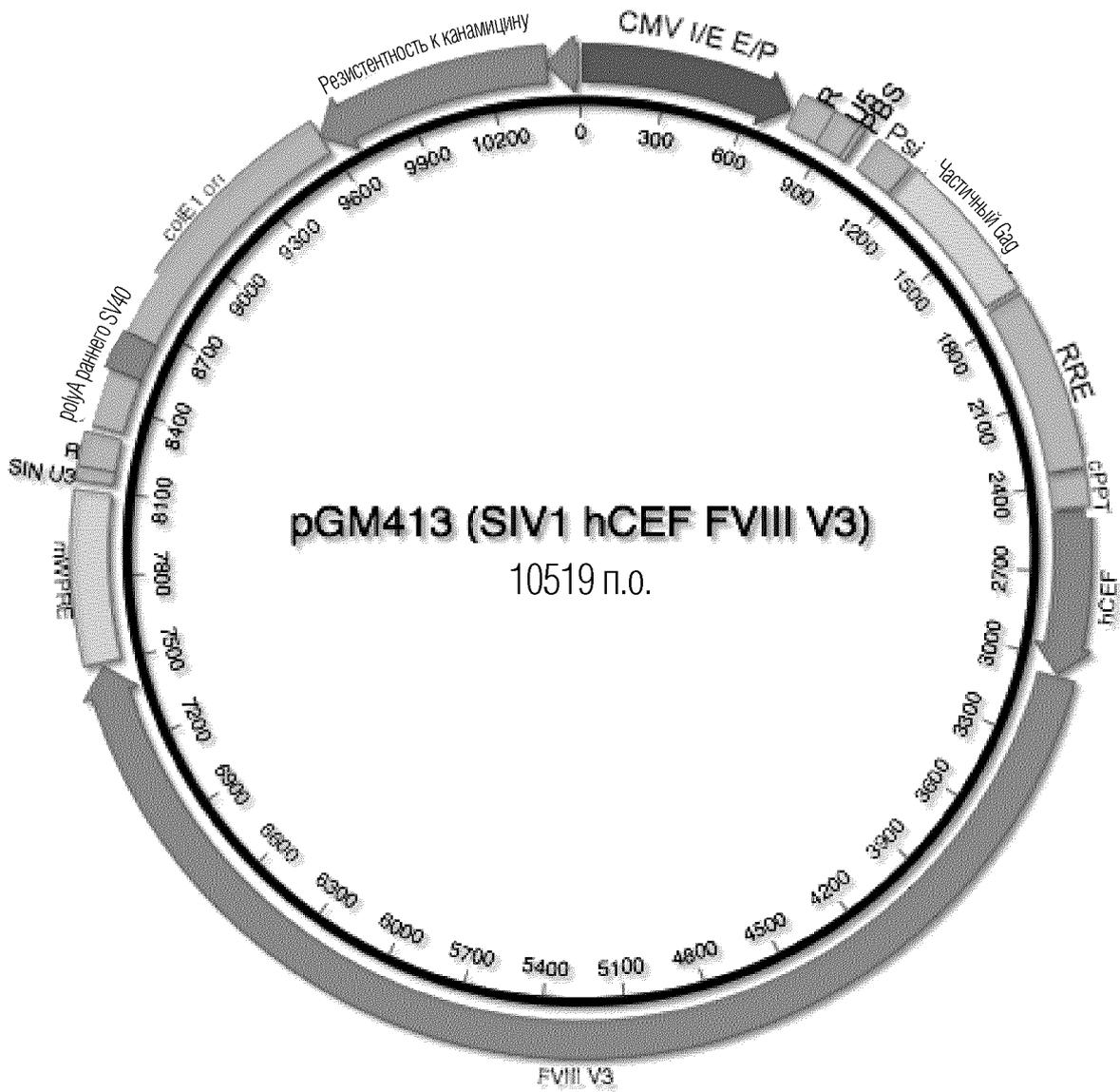
ФИГ. 2G



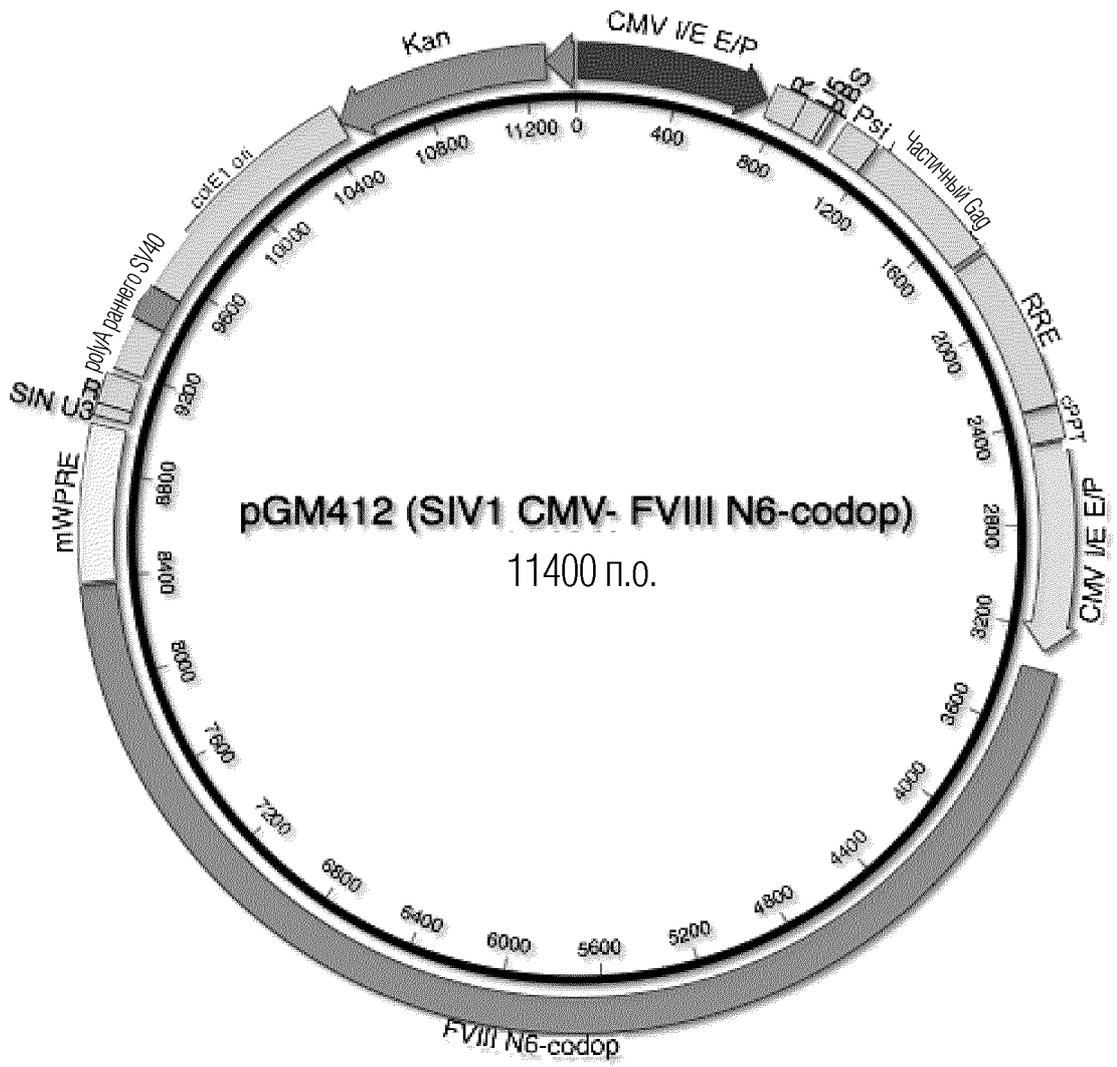
ФИГ. 3



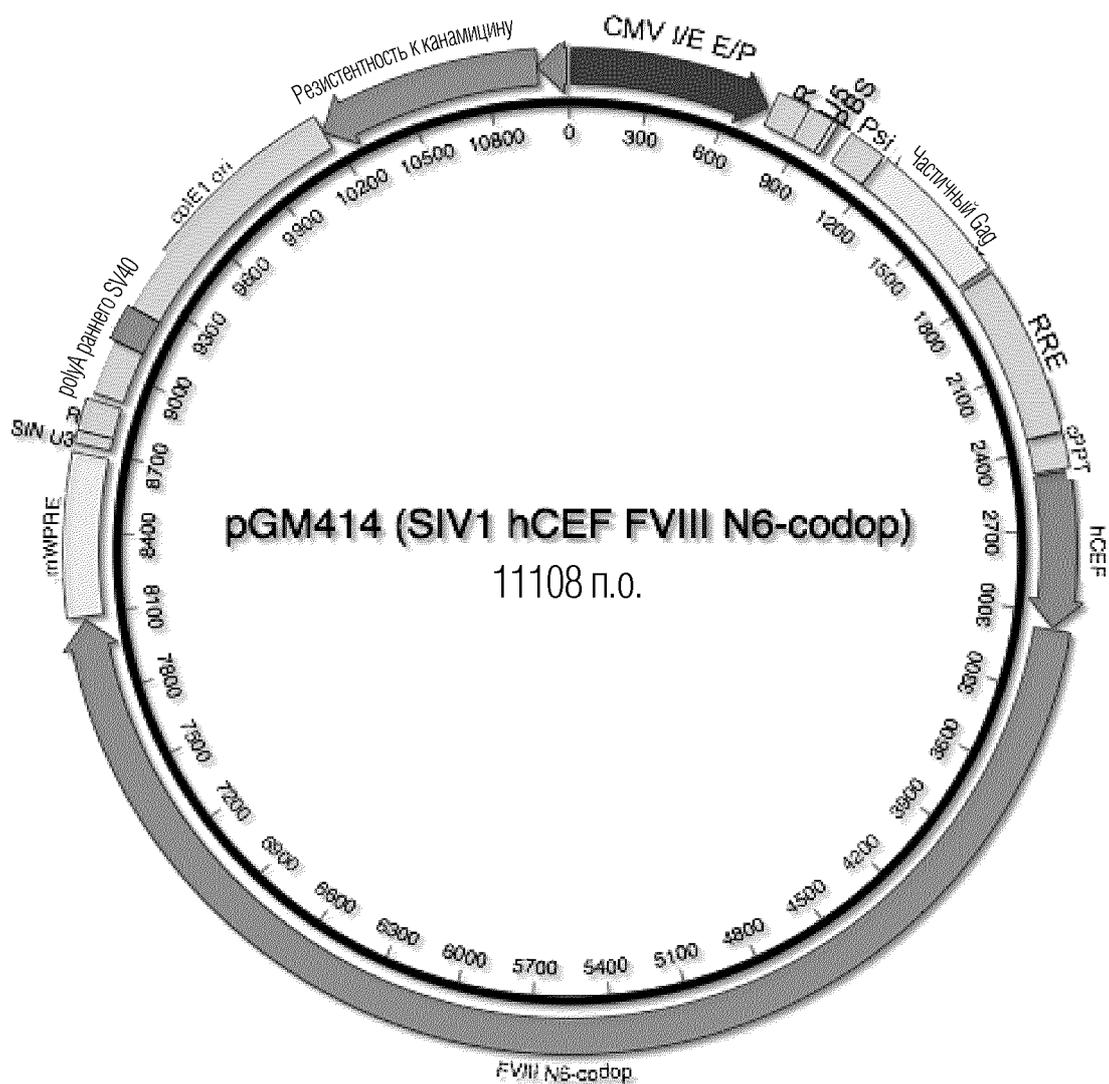
ФИГ. 4В



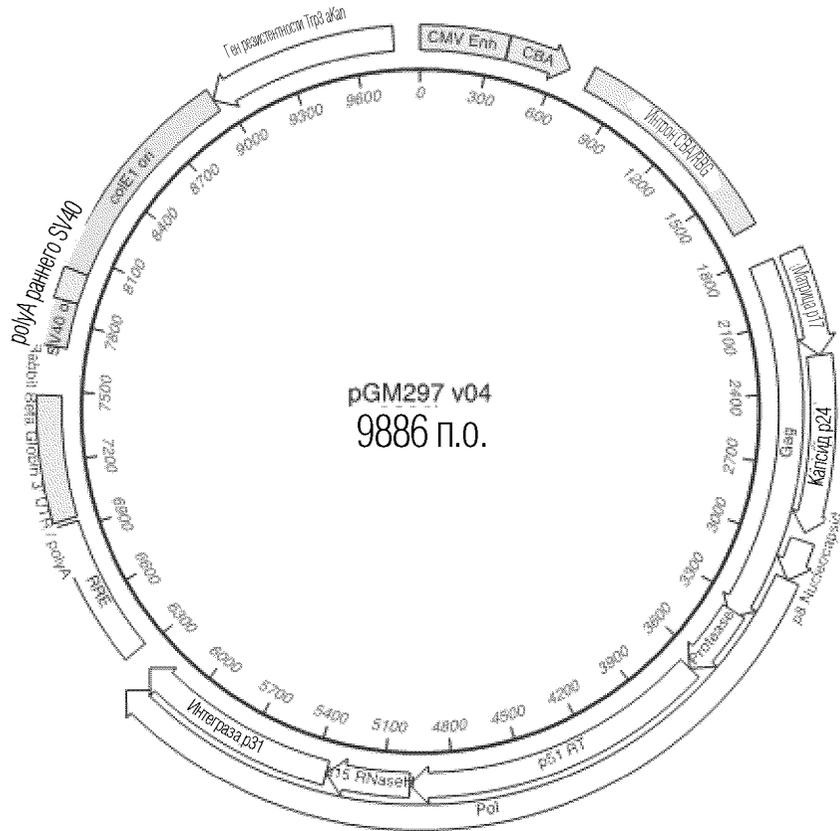
ФИГ. 4С



ФИГ. 4D



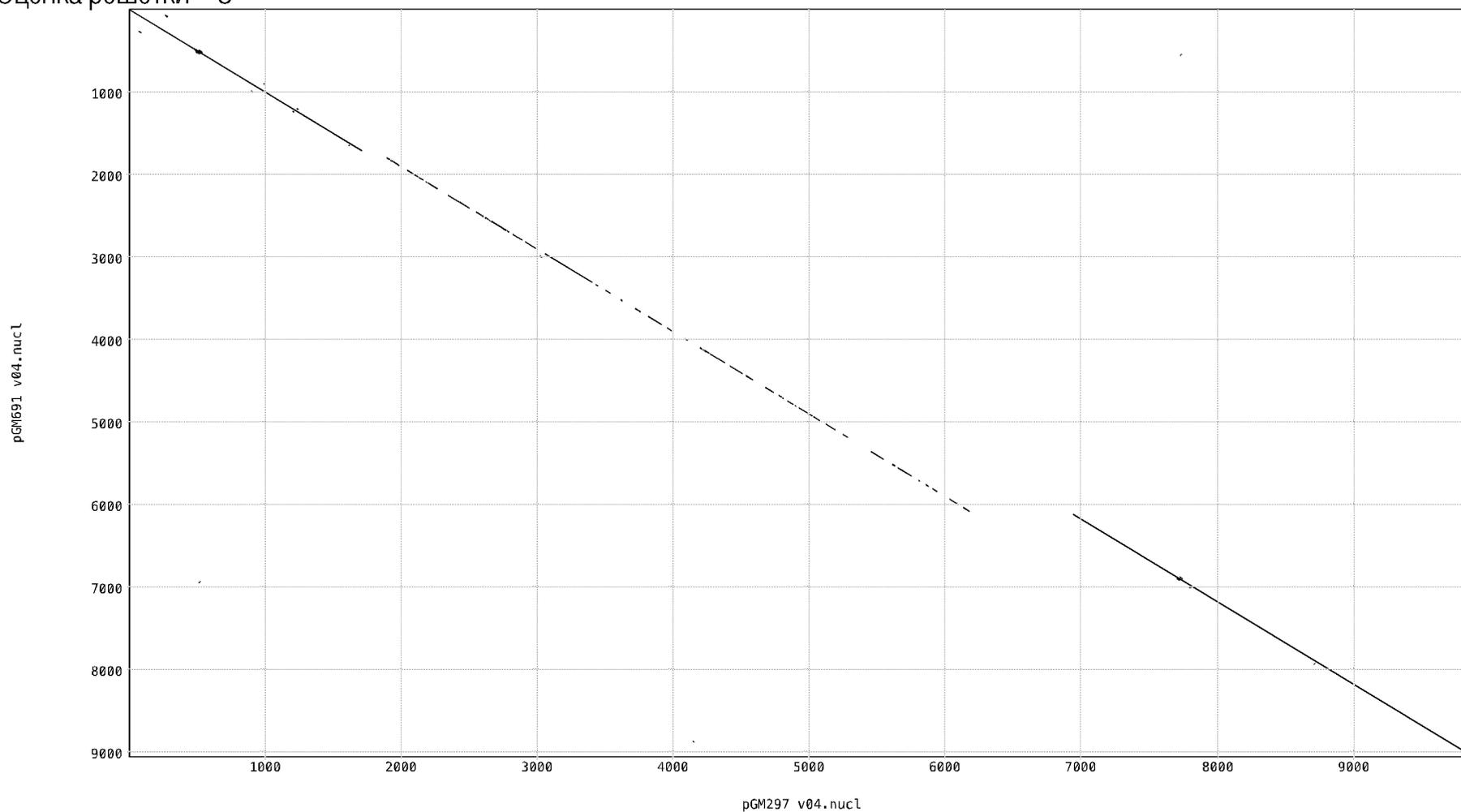
ФИГ. 5В



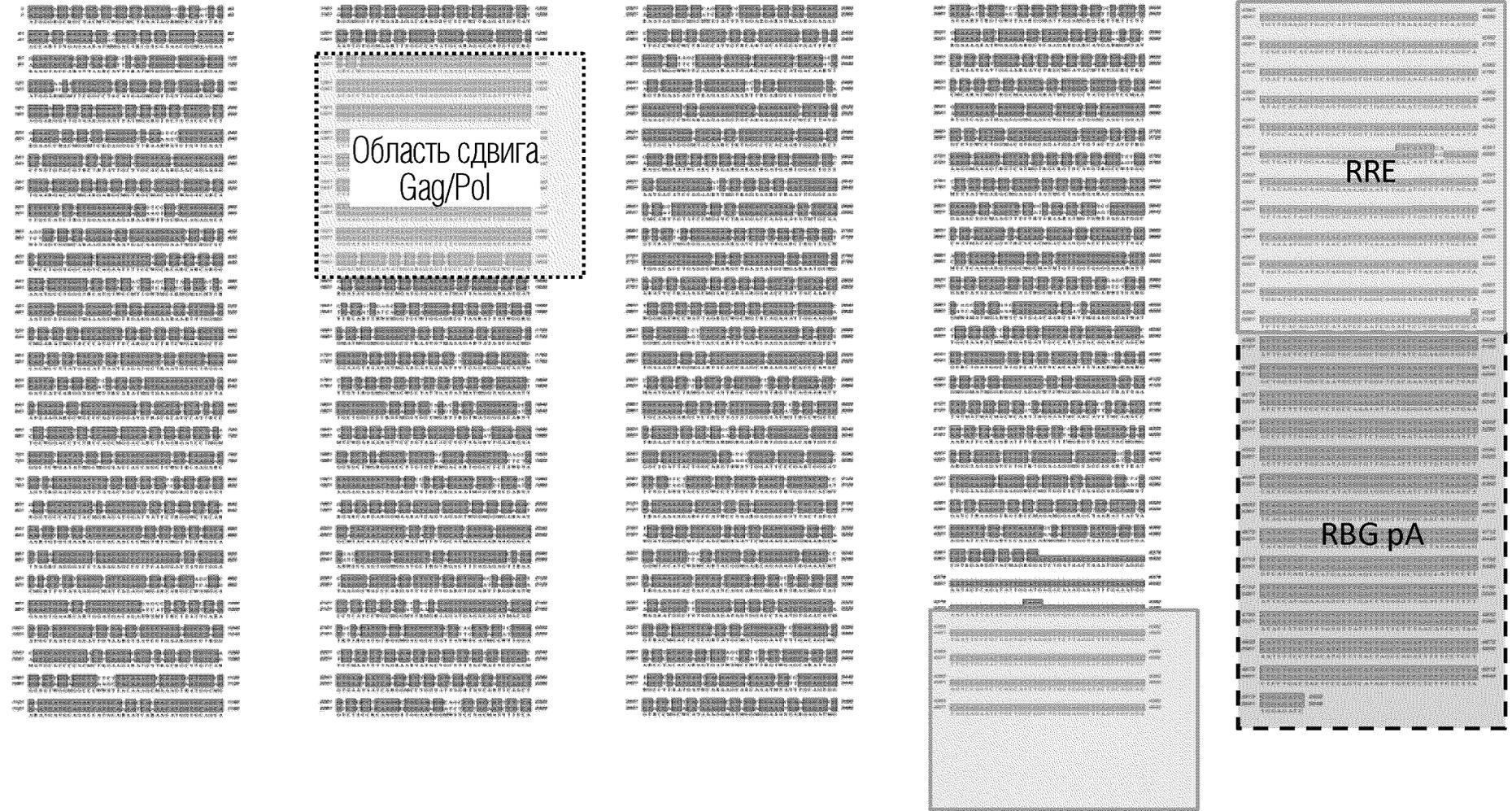
ФИГ. 5С

Размер окна = 30
Мин. % баллов = 60
Оценка решетки = 8

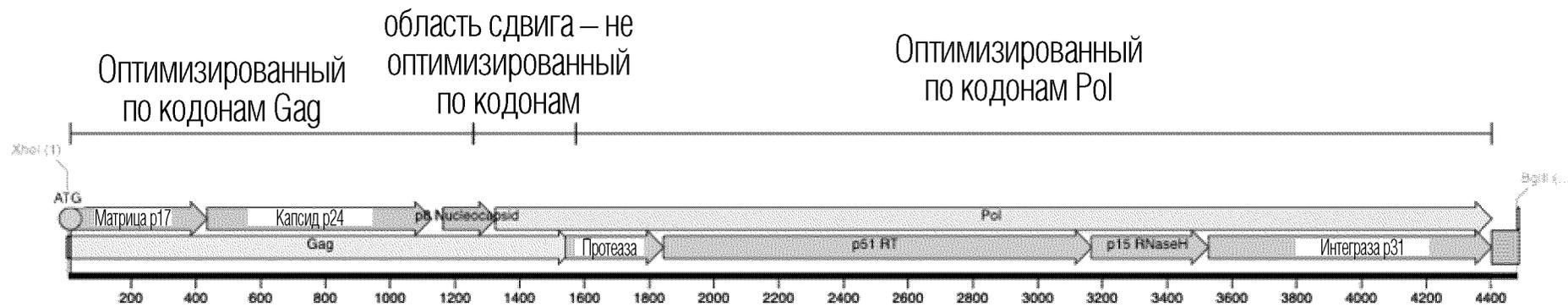
Цепь = Обе Оценочная матрица: матрица для базы даннх ДНК.nnat
Переход = 1



ФИГ. 5D

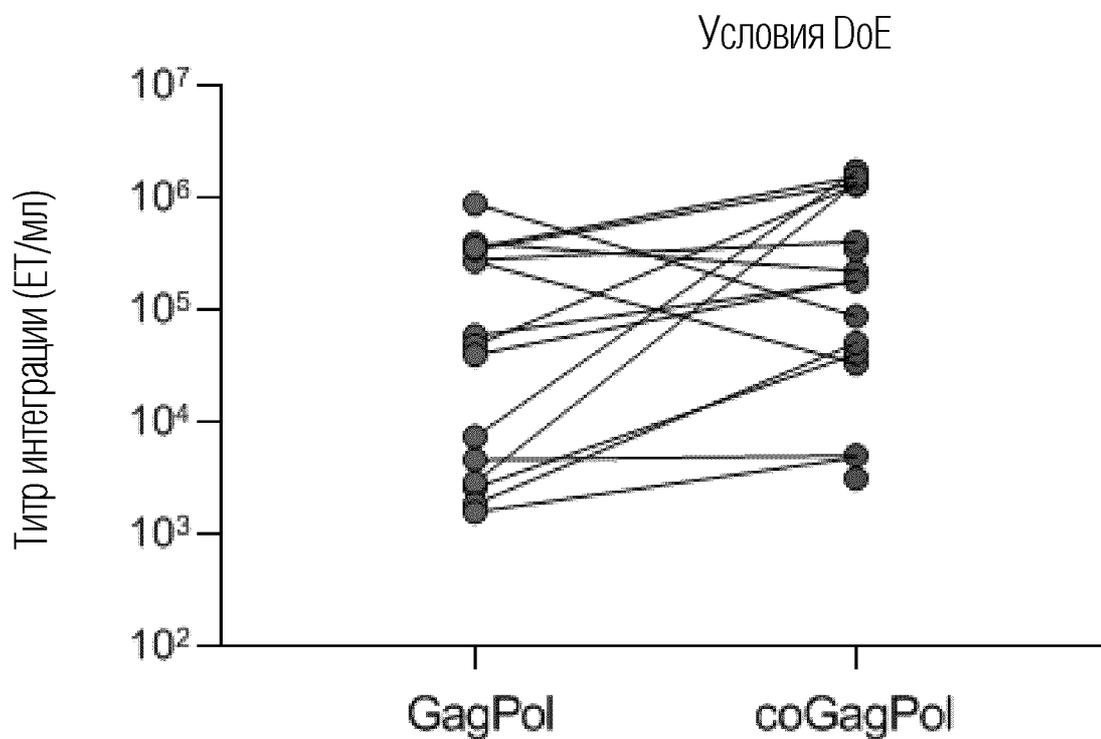


ФИГ. 5Е

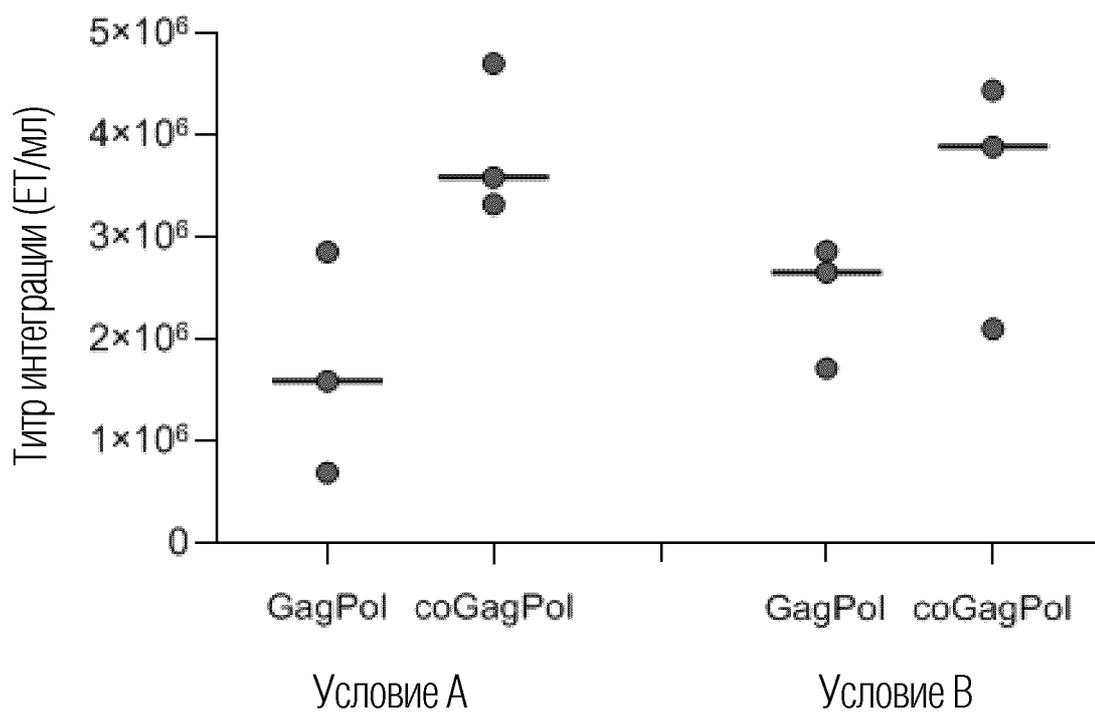


ФИГ. 6

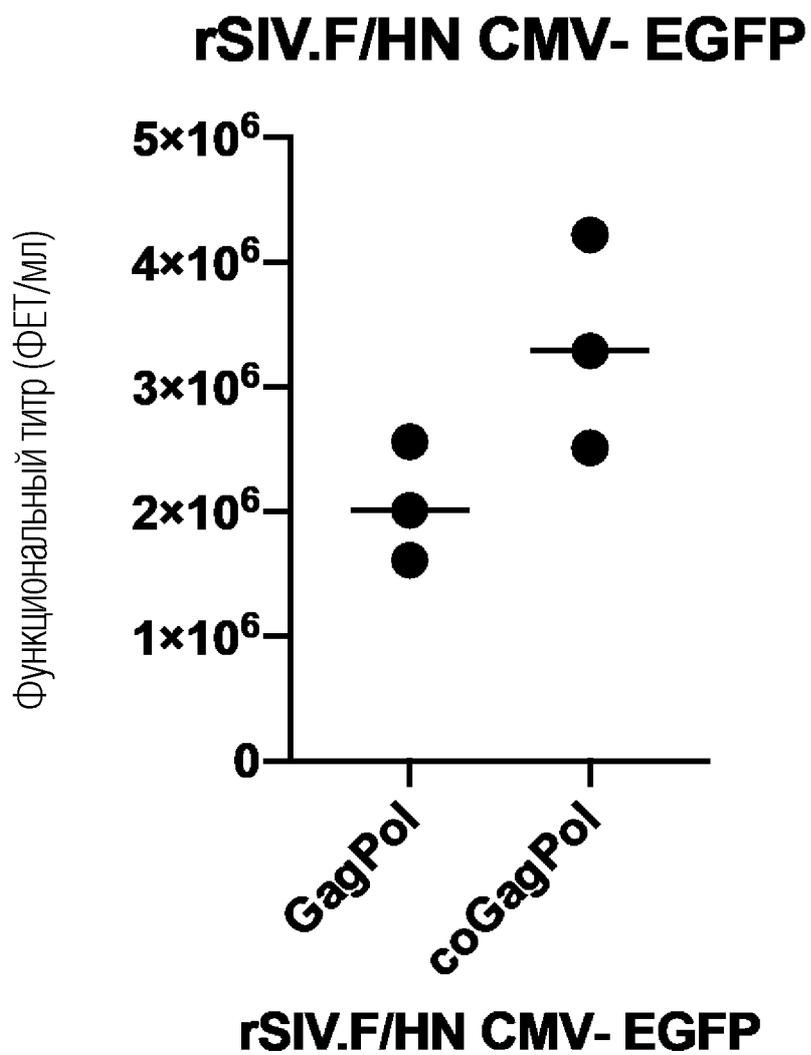
А



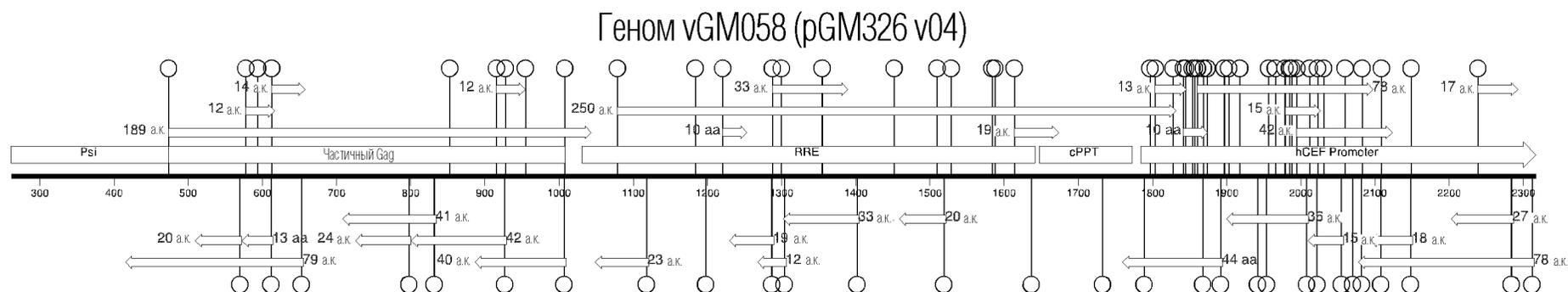
В



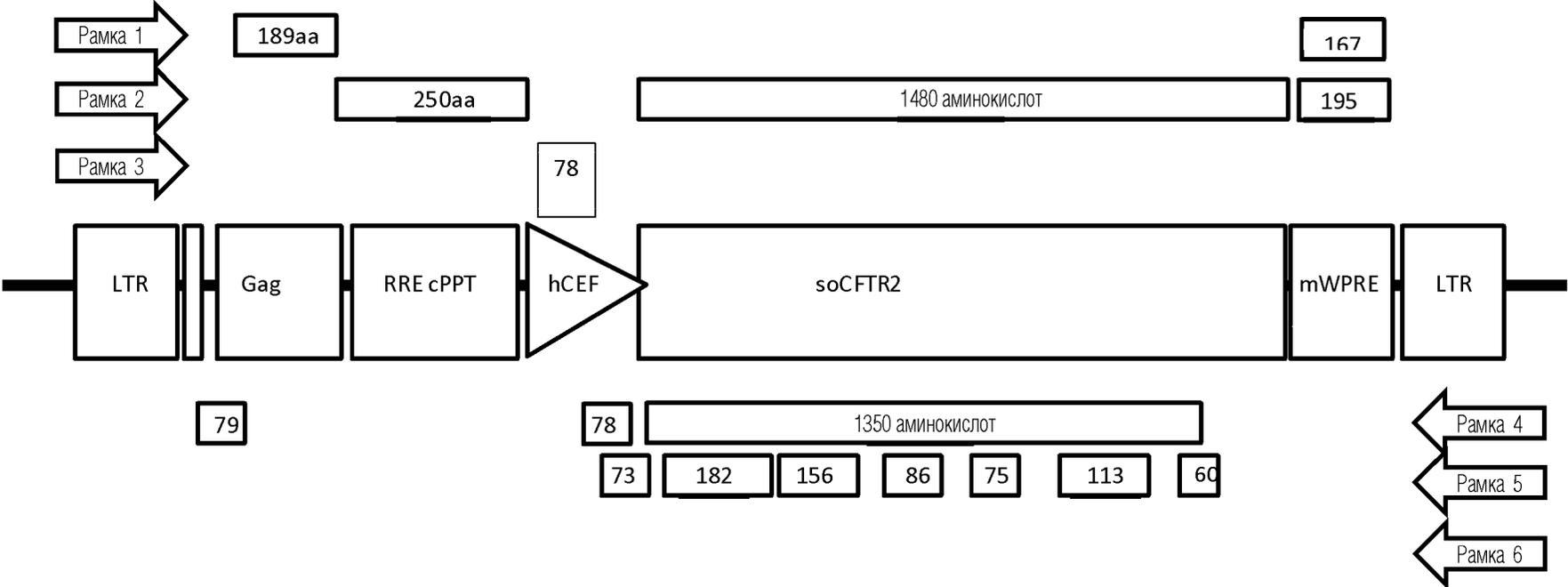
ФИГ. 7



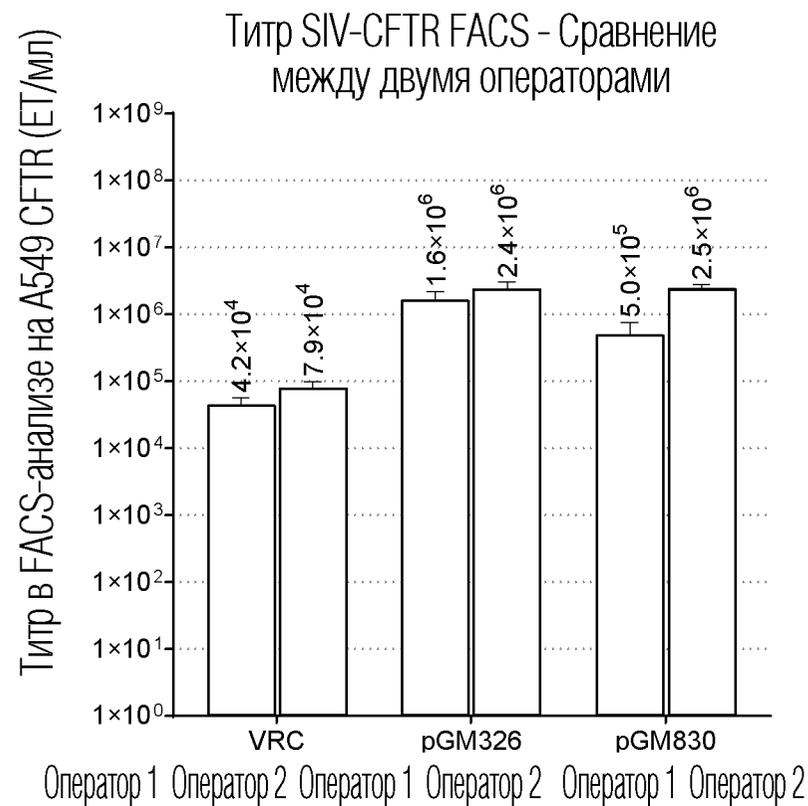
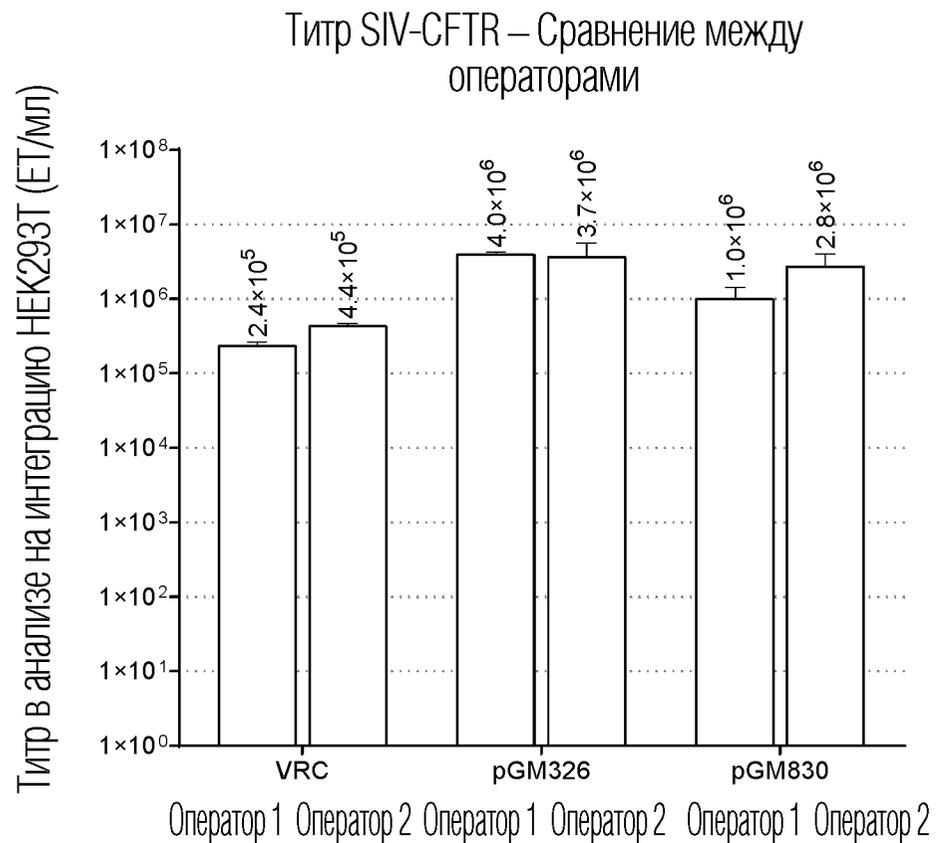
ФИГ. 8



ФИГ. 9



ФИГ. 10



ФИГ. 11

