

(19)



Евразийское
патентное
ведомство

(21) 201991747 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2020.06.04

(22) Дата подачи заявки
2018.02.27

(51) Int. Cl. *A61K 38/17* (2006.01)
A61K 9/127 (2006.01)
C07K 14/47 (2006.01)
C12N 15/10 (2006.01)

(54) НОВАЯ КОДОН-ОПТИМИЗИРОВАННАЯ мРНК CFTR

(31) 62/464,215

(32) 2017.02.27

(33) US

(86) PCT/US2018/020011

(87) WO 2018/157154 2018.08.30

(88) 2018.10.04

(71) Заявитель:
ТРАНСЛЕЙТ БИО, ИНК. (US)

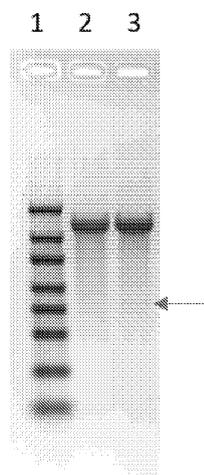
(72) Изобретатель:

Диас Ануша, Абисал Джонатан,
Беттанкур Брайан, Дероза Франк,
Хартлейн Майкл (US)

(74) Представитель:

Нилова М.И. (RU)

(57) В настоящем изобретении предложены, помимо прочего, улучшенные способы и фармацевтические композиции для лечения муковисцидоза на основе кодон-оптимизированной мРНК, кодирующей муковисцидозный регулятор трансмембранной проводимости (CFTR).



201991747 A1

201991747 A1

НОВАЯ КОДОН-ОПТИМИЗИРОВАННАЯ МРНК CFTR

РОДСТВЕННЫЕ ЗАЯВКИ

[0001] Настоящая заявка испрашивает приоритет по предварительной заявке на патент США № 62/464,215, поданной 27 февраля 2017 г., описания которой включены в настоящий документ путем ссылки.

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

[0002] В настоящем описании дана ссылка на перечень последовательностей (переданный в электронной форме в виде файла .txt с наименованием MRT-2001 WO_SL от 27 февраля 2018 г.). Файл .txt был создан в указанную дату и имеет размер 166293 байт. Полное содержание последовательности включено в настоящий документ путем ссылки.

ПРЕДПОСЫЛКИ СОЗДАНИЯ ИЗОБРЕТЕНИЯ

[0003] Муковисцидоз представляет собой аутосомное наследственное расстройство, вызванное мутацией гена CFTR, кодирующего хлоридный канал, который, как считается, участвует в регуляции множества других ионных каналов и транспортных систем в эпителиальных клетках. Утрата функции CFTR приводит к хроническому заболеванию легких, аномальной продукции слизи и значительному снижению ожидаемой продолжительности жизни. См. по существу публикацию Rowe et al., *New Engl. J. Med.* 352, 1992–2001 (2005).

[0004] В настоящее время лечения от муковисцидоза нет. В литературе отмечены многочисленные трудности, встретившиеся при попытке индуцировать экспрессию CFTR в легком. Например, вирусные векторы, содержащие CFTR-ДНК, активировали иммунные ответы, и после введения симптомы муковисцидоза сохранялись. Conese et al., *J. Cyst. Fibros.* 10 Suppl 2, S114–28 (2011); Rosenecker et al., *Curr. Opin. Mol. Ther.* 8, 439–45 (2006). Также отмечено, что невирусная доставка ДНК, включая CFTR-ДНК, вызывает иммунные ответы. Alton et al., *Lancet* 353, 947–54 (1999); Rosenecker et al., *J Gene Med.* 5, 49–60 (2003).

Более того, невирусные ДНК-векторы сталкиваются с дополнительной проблемой, заключающейся в том, что механизм комплекса ядерной поры обычно не импортирует ДНК в ядро, где происходит транскрипция. Pearson, Nature 460, 164–69 (2009).

ИЗЛОЖЕНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

[0005] В настоящем изобретении предложены, помимо прочего, фармацевтические композиции, содержащие матричную РНК (мРНК), кодирующую муковисцидозный регулятор трансмембранной проводимости (CFTR), а также способы их получения и использования. Эти фармацевтические композиции можно использовать для усовершенствованного лечения муковисцидоза.

[0006] В одном аспекте в настоящем изобретении предложены фармацевтические композиции для лечения муковисцидоза, содержащие мРНК, кодирующую муковисцидозный регулятор трансмембранной проводимости (CFTR), и причем мРНК, кодирующая белок CFTR, содержит полинуклеотидную последовательность, по меньшей мере на 85% идентичную SEQ ID NO: 1. В некоторых вариантах осуществления мРНК, кодирующая белок CFTR, содержит SEQ ID NO: 1. В некоторых вариантах осуществления мРНК дополнительно содержит последовательность 5'-нетранслируемой области (UTR) с SEQ ID NO: 4. В некоторых вариантах осуществления мРНК дополнительно содержит последовательность 3'-нетранслируемой области (UTR) с SEQ ID NO: 5 или SEQ ID NO: 6.

[0007] В некоторых вариантах осуществления мРНК, кодирующая белок CFTR, инкапсулирована внутри наночастицы. В некоторых вариантах осуществления наночастица представляет собой липосому. В некоторых вариантах осуществления липосома содержит один или более катионных липидов, один или более некатионных липидов, один или более липидов на основе холестерина и один или более ПЭГ-модифицированных липидов. В некоторых вариантах осуществления липосома содержит не более трех отдельных липидных компонентов. В некоторых вариантах осуществления один отдельный липидный компонент представляет собой катионный липид на основе стерола. В некоторых вариантах осуществления липосома имеет размер менее чем приблизительно 100 нм. В другом аспекте в настоящем изобретении предложены способы крупномасштабного производства мРНК, кодирующей муковисцидозный регулятор трансмембранной проводимости (CFTR). В некоторых

вариантах осуществления способ в соответствии с настоящим изобретением включает в себя синтез *in vitro* мРНК, кодирующей белок CFTR, с использованием РНК-полимеразы SP6, причем по меньшей мере 80% синтезированных молекул мРНК являются полноразмерными и причем в одной партии синтезируют по меньшей мере 100 мг мРНК.

[0008] В некоторых вариантах осуществления синтезированная *in vitro* мРНК, кодирующая CFTR, по существу не содержит вторичного полинуклеотидного фрагмента длиной приблизительно 1800 нуклеотидов. В некоторых вариантах осуществления синтез *in vitro* мРНК приводит к получению вторичного полинуклеотидного фрагмента, который составляет менее 10%, 5%, 4%, 3%, 2%, 1%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% от всей синтезированной мРНК.

[0009] В некоторых вариантах осуществления по меньшей мере 85%, 90%, 95%, 96%, 97%, 98% или 99% синтезированных молекул мРНК являются полноразмерными. В некоторых вариантах осуществления синтезированные молекулы мРНК являются по существу полноразмерными.

[0010] В некоторых вариантах осуществления в одной партии синтезируют по меньшей мере 200 мг, 300 мг, 400 мг, 500 мг, 600 мг, 700 мг, 800 мг, 900 мг, 1 г, 5 г, 10 г, 25 г, 50 г, 75 г, 100 г, 150 г, 200 г, 250 г, 500 г, 750 г, 1 кг, 5 кг, 10 кг, 50 кг, 100 кг, 1000 кг или более мРНК.

[0011] В некоторых вариантах осуществления белок CFTR содержит аминокислотную последовательность с SEQ ID NO: 3. В некоторых вариантах осуществления мРНК содержит полинуклеотидную последовательность, по меньшей мере на 85% идентичную SEQ ID NO: 1. В некоторых вариантах осуществления мРНК дополнительно содержит последовательность 5'-нетранслируемой области (UTR) с SEQ ID NO: 4. В некоторых вариантах осуществления мРНК дополнительно содержит последовательность 3'-нетранслируемой области (UTR) с SEQ ID NO: 5 или SEQ ID NO: 6.

[0012] В некоторых вариантах осуществления способ дополнительно включает в себя стадию добавления кэпа и/или хвоста к синтезированной мРНК CFTR.

[0013] Помимо прочего, в настоящем изобретении предложена мРНК, кодирующая муковисцидозный регулятор трансмембранной проводимости (CFTR),

синтезированный с использованием различных способов, описанных в настоящем документе, и содержащие ее фармацевтические композиции.

[0014] В еще одном аспекте в настоящем изобретении предложены способы доставки мРНК, кодирующей CFTR, которая описана в настоящем документе, для экспрессии белка *in vivo* и/или для лечения муковисцидоза. В некоторых вариантах осуществления в настоящем изобретении предложены способы лечения муковисцидоза, включающие в себя введение нуждающемуся в лечении субъекту композиции, содержащей мРНК, кодирующую муковисцидозный регулятор трансмембранной проводимости (CFTR), причем мРНК содержит полинуклеотидную последовательность, по меньшей мере на 85% (например, по меньшей мере на 90%, 92%, 94%, 95%, 96%, 97%, 98% или 99%) идентичную SEQ ID NO: 1.

[0015] В некоторых вариантах осуществления мРНК, кодирующая белок CFTR, содержит SEQ ID NO: 1. В некоторых вариантах осуществления мРНК дополнительно содержит последовательность 5'-нетранслируемой области (UTR) с SEQ ID NO: 4. В некоторых вариантах осуществления мРНК дополнительно содержит последовательность 3'-нетранслируемой области (UTR) с SEQ ID NO: 5 или SEQ ID NO: 6.

[0016] В некоторых вариантах осуществления мРНК, кодирующая белок CFTR, инкапсулирована внутри наночастицы. В некоторых вариантах осуществления наночастица представляет собой липосому. В некоторых вариантах осуществления липосома содержит один или более катионных липидов, один или более некаатионных липидов, один или более липидов на основе холестерина и один или более ПЭГ-модифицированных липидов. В некоторых вариантах осуществления липосома содержит не более трех отдельных липидных компонентов. В некоторых вариантах осуществления один отдельный липидный компонент представляет собой катионный липид на основе стерола. В некоторых вариантах осуществления катионный липид на основе стерола представляет собой сложный эфир имидазола и холестерина «ICE» (3S, 10R, 13R, 17R)-10, 13-диметил-17-((R)-6-метилгептан-2-ил)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-тетрадекагидро-1H-циклопента[а]фенантрен-3-ил 3-(1H-имидазол-4-ил)пропаноат. В некоторых вариантах осуществления липосома имеет размер менее чем приблизительно 100 нм.

[0017] В некоторых вариантах осуществления мРНК вводят субъекту посредством доставки через легкие. В некоторых вариантах осуществления доставка через легкие представляет собой распыление.

[0018] Другие характеристики, цели и преимущества настоящего изобретения станут понятны из следующего подробного описания, чертежей и формулы изобретения. Однако следует понимать, что несмотря на указание вариантов осуществления настоящего изобретения, подробное описание, чертежи и формула изобретения приведены только в качестве примера, но не ограничения. Специалисту в данной области станут очевидными разнообразные изменения и модификации в пределах объема изобретения.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

[0019] Рисунки приведены только в качестве примера, но не для ограничения.

[0020] На **Фиг. 1** показан пример геля, демонстрирующий, что синтез новой кодон-оптимизированной последовательности муковисцидозного регулятора трансмембранной проводимости (CFTR) с использованием промотора SP6 исключил вторичный полинуклеотидный фрагмент (полоса 2) по сравнению с предыдущей кодон-оптимизированной последовательностью CFTR (полоса 3). Стрелка указывает на вторичный полинуклеотидный фрагмент длиной приблизительно 1800 нуклеотидов.

ОПРЕДЕЛЕНИЯ

[0021] Чтобы настоящее изобретение было более понятным, ниже сначала приведены некоторые определения терминов. Дополнительные определения представленных ниже терминов и других терминов приведены в тексте описания. Публикации и другие справочные материалы, упоминаемые в настоящем документе для описания предпосылок создания изобретения и для предоставления дополнительной информации относительно его практического применения, включены в настоящий документ путем ссылки.

[0022] *Приблизительно* или *около*. В настоящем документе термин «приблизительно» или «около», применяемый к одной или более интересующим величинам, означает величину, аналогичную указанной эталонной величине. В

некоторых вариантах осуществления термин «приблизительно» или «около» обозначает диапазон величин, которые находятся в пределах 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1% или меньше в обе стороны (большую или меньшую) от заданной эталонной величины, кроме случаев, когда указано иное значение или иное значение очевидно из контекста (за исключением ситуаций, когда такое число превышает 100% от возможной величины).

[0023] В настоящем документе термин «партия» относится к количеству или объему мРНК, синтезированному за один раз, например, полученной по одному заказу на производство в течение одного и того же производственного цикла. Партия может означать количество мРНК, синтезированное в одной реакции, которая протекает с использованием одной аликвоты фермента и/или одной аликвоты ДНК-шаблона в ходе непрерывного синтеза при одном и том же наборе условий. В некоторых вариантах осуществления партия включает в себя мРНК, полученную в результате реакции, при которой по мере протекания реакции добавляют и/или пополняют не все реагенты и/или компоненты. Термин «не в одной партии» не означает мРНК, синтезированную в разное время и скомбинированную для получения требуемого количества.

[0024] *Доставка.* В настоящем документе термин «доставка» включает в себя как местную, так и системную доставку. Например, доставка мРНК охватывает ситуации, в которых мРНК доставляется в ткань-мишень, а кодируемый белок экспрессируется и остается внутри ткани-мишени (что также называется «местным распределением» или «местной доставкой»), и ситуации, в которых мРНК доставляется в ткань-мишень, а кодируемый белок экспрессируется и секретируется в систему циркуляции пациента (например, в сыворотку крови) и системно распределяется и поглощается другими тканями (что также называется «системным распределением» или «системной доставкой») В некоторых вариантах осуществления доставка представляет собой доставку через легкие, например, включающую в себя распыление.

[0025] *Инкапсулирование.* В настоящем документе термин «инкапсулирование» или его грамматический эквивалент обозначает процесс заключения молекулы мРНК внутрь наночастицы.

[0026] *Экспрессия.* В настоящем документе термин «экспрессия» нуклеотидной последовательности относится к трансляции мРНК в полипептид, сборке множества

полипептидов (например, тяжелой цепи или легкой цепи антитела) в интактный белок (например, антитело) и/или к посттрансляционной модификации полипептида или полностью собранного белка (например, антитела). В настоящей заявке термины «экспрессия» и «продукция», а также их грамматические эквиваленты используются взаимозаменяемо.

[0027] *Функциональность.* В настоящем документе «функциональная» биологическая молекула представляет собой биологическую молекулу в такой форме, в которой она демонстрирует свойство и/или активность, по которой ее характеризуют.

[0028] *Период полужизни.* В настоящем документе термин «период полужизни» означает время, необходимое для того, чтобы количественный показатель, такой как концентрация или активность нуклеиновой кислоты или белка, снизился до половины его значения, измеренного в начале временного периода.

[0029] *Улучшение, увеличение или уменьшение.* В настоящем документе термины «улучшать», «увеличивать» или «уменьшать» либо их грамматические эквиваленты указывают на значения относительно исходных измерений, таких как измерение у одного и того же индивида до начала лечения, описанного в настоящем документе, или измерение у контрольного субъекта (или множества контрольных субъектов) в отсутствие лечения, описанного в настоящем документе. «Контрольный субъект» представляет собой субъекта, страдающего такой же формой заболевания, что и субъект, получающий лечение, который приблизительно того же возраста, что и субъект, получающий лечение.

[0030] *Примеси.* В настоящем документе термин «примеси» относится к веществам, содержащимся в ограниченном количестве жидкости, газа или твердого вещества, которые отличаются по химическому составу от целевого материала или соединения. Примеси также называют загрязнителями.

[0031] *In vitro.* В настоящем документе термин «*in vitro*» относится к событиям, происходящим в искусственной среде, например, в пробирке или реакционном сосуде, в клеточной культуре и т. д., а не в многоклеточном организме.

[0032] *In vivo.* В настоящем документе термин «*in vivo*» относится к событиям, происходящим в многоклеточном организме, таком как человек или не относящееся к человеку животное. В контексте систем на основе клеток термин может

использоваться для обозначения событий, происходящих внутри живой клетки (в отличие, например, от систем *in vitro*).

[0033] *Выделенный*. В настоящем документе термин «выделенный» обозначает вещество и/или объект, который был (1) отделен от по меньшей мере некоторых компонентов, с которыми он был связан при изначальной продукции (в природе и/или в экспериментальных условиях), и/или (2) продуцирован, получен и/или произведен человеком. Выделенные вещества и/или объекты могут быть отделены от приблизительно 10%, приблизительно 20%, приблизительно 30%, приблизительно 40%, приблизительно 50%, приблизительно 60%, приблизительно 70%, приблизительно 80%, приблизительно 90%, приблизительно 91%, приблизительно 92%, приблизительно 93%, приблизительно 94%, приблизительно 95%, приблизительно 96%, приблизительно 97%, приблизительно 98%, приблизительно 99% или более чем приблизительно 99% других компонентов, с которыми они были изначально связаны. В некоторых вариантах осуществления выделенные агенты имеют степень чистоты, равную приблизительно 80%, приблизительно 85%, приблизительно 90%, приблизительно 91%, приблизительно 92%, приблизительно 93%, приблизительно 94%, приблизительно 95%, приблизительно 96%, приблизительно 97%, приблизительно 98%, приблизительно 99% или более чем приблизительно 99%. В настоящем документе вещество является «чистым», если оно по существу не содержит других компонентов. В настоящем документе расчет процентной чистоты выделенных веществ и/или объектов не должен включать в себя эксципиенты (например, буферный раствор, растворитель, воду и т. д.).

[0034] *Матричная РНК (мРНК)*. В настоящем документе термин «матричная РНК (мРНК)» означает полинуклеотид, который кодирует по меньшей мере один полипептид. В настоящем документе термин «мРНК» охватывает как модифицированную, так и немодифицированную РНК. мРНК может содержать одну или более кодирующих и не кодирующих областей. мРНК может быть очищена из природных источников, получена с использованием рекомбинантных экспрессионных систем и необязательно очищена, синтезирована химически и т. д. При необходимости, например, в случае химически синтезированных молекул мРНК может содержать аналоги нуклеозидов, такие как аналоги с химически модифицированными

основаниями или сахарами, модификациями основной цепи и т. д. Если не указано иное, последовательность мРНК приведена в направлении от 5'- к 3'-концу.

[0035] *Нуклеиновая кислота.* В настоящем документе термин «нуклеиновая кислота» в своем самом широком смысле обозначает любое соединение и/или вещество, которое находится в полинуклеотидной цепи или может быть встроено в нее. В некоторых вариантах осуществления нуклеиновая кислота представляет собой соединение и/или вещество, которое находится в полинуклеотидной цепи или может быть встроено в нее посредством фосфодиэфирной связи. В некоторых вариантах осуществления «нуклеиновая кислота» обозначает отдельные остатки нуклеиновой кислоты (например, нуклеотиды и/или нуклеозиды). В некоторых вариантах осуществления термин «нуклеиновая кислота» относится к полинуклеотидной цепи, содержащей отдельные остатки нуклеиновой кислоты. В некоторых вариантах осуществления «нуклеиновая кислота» охватывает РНК, а также одноцепочечные и/или двухцепочечные ДНК и/или кДНК. Более того, термины «нуклеиновая кислота», «ДНК», «РНК» и/или аналогичные термины включают в себя аналоги нуклеиновых кислот, т. е. аналоги, имеющие основную цепь, отличную от фосфодиэфирной основной цепи. Например, так называемые «пептидные нуклеиновые кислоты», известные в данной области, в каркасе которых присутствуют пептидные связи вместо фосфодиэфирных связей, считаются входящими в объем настоящего изобретения. Термин «нуклеотидная последовательность, кодирующая аминокислотную последовательность» включает в себя все нуклеотидные последовательности, которые являются вырожденными вариантами друг друга и/или кодируют одну и ту же аминокислотную последовательность. Нуклеотидные последовательности, которые кодируют белки и/или РНК, могут включать в себя интроны. Нуклеиновые кислоты могут быть очищены из природных источников, получены с использованием рекомбинантных экспрессионных систем и необязательно очищены, синтезированы химически и т. д. При необходимости, например, в случае химически синтезированных молекул, нуклеиновые кислоты могут содержать аналоги нуклеозидов, такие как аналоги с химически модифицированными основаниями или сахарами, модификациями основной цепи и т. д. Если не указано иное, последовательность нуклеиновой кислоты приведена в направлении от 5'- к 3'-концу. В некоторых вариантах осуществления нуклеиновая кислота представляет собой или содержит природные нуклеозиды (например, аденозин, тимидин, гуанозин, цитидин, уридин,

дезоксиаденозин, дезокситимидин, дезоксигуанозин и дезоксицитидин); аналоги нуклеозидов (например, 2-аминоаденозин, 2-тиотимидин, инозин, пирролопиримидин, 3-метиладенозин, 5-метилцитидин, С-5 пропинилцитидин, С-5 пропинилуридин, 2-аминоаденозин, С5-бромурин, С5-фторуридин, С5-йодуридин, С5-пропинилуридин, С5-пропинилцитидин, С5-метилцитидин, 2-аминоаденозин, 7-дезааденозин, 7-дезагуанозин, 8-оксоаденозин, 8-оксогуанозин, О(6)-метилгуанин и 2-тиоцитидин); химически модифицированные основания; биологически модифицированные основания (например, метилированные основания); интеркалированные основания; модифицированные сахара (например, 2'-фторрибоза, рибоза, 2'-дезоксирибоза, арабиноза и гексоза) и/или модифицированные фосфатные группы (например, фосфоротиоаты и 5'-N-фосфорамидитные связи). В некоторых вариантах осуществления настоящее изобретение конкретно относится к «немодифицированным нуклеиновым кислотам», что означает нуклеиновые кислоты (например, полинуклеотиды и остатки, включая нуклеотиды и/или нуклеозиды), которые не были химически модифицированы для облегчения или обеспечения доставки. В некоторых вариантах осуществления нуклеотиды Т и У используются в описаниях последовательностей взаимозаменяемо.

[0036] *Пациент.* В настоящем документе термин «пациент» или «субъект» относится к любому организму, которому может быть введена предложенная композиция, например, для экспериментальных, диагностических, профилактических, косметических и/или терапевтических целей. К типичным пациентам относятся животные (например, млекопитающие, такие как мыши, крысы, кролики, не относящиеся к человеку приматы и/или люди). В некоторых вариантах осуществления пациент представляет собой человека. Человек включает в себя пре- и постнатальные формы.

[0037] *Фармацевтически приемлемый.* В настоящей документе термин «фармацевтически приемлемый» относится к веществам, которые в пределах обоснованного медицинского мнения подходят для использования в контакте с тканями человека и животных, не вызывая излишнего токсического, раздражающего, аллергического ответа или другой проблемы или осложнения, соизмеримого с разумным соотношением пользы/риска.

[0038] *Субъект.* В настоящем документе термин «субъект» относится к человеку или любому не относящемуся к человеку животному (например, к мыши, крысе, кролику, собаке, кошке, крупному рогатому скоту, свинье, овце, лошади или примату). Человек включает в себя пре- и постнатальные формы. Во многих вариантах осуществления субъект представляет собой человека. Субъект может представлять собой пациента, который означает человека, обратившегося к поставщику медицинских услуг для диагностики или лечения заболевания. В настоящем документе термин «субъект» используется взаимозаменяемо с термином «индивид» или «пациент». Субъект может страдать от заболевания или расстройства или быть восприимчив к нему, но может демонстрировать или не демонстрировать симптомы заболевания или расстройства.

[0039] *По существу.* В настоящем документе термин «по существу» обозначает качественное состояние проявления в полной или практически полной мере или степени интересующей характеристики или свойства. Обычному специалисту в области биологии будет понятно, что биологические и химические явления редко или практически никогда не стремятся к завершенности, и/или не протекают до завершения, или не достигают, или не имеют абсолютного результата. Таким образом, термин «по существу» используется в настоящем документе для обозначения потенциального отсутствия завершенности, присущего многим биологическим и химическим явлениям.

[0040] *Лечение.* В настоящем документе термин «лечить», «лечение» или «лечебный» относится к любому способу, используемому для частичного или полного ослабления, облегчения, уменьшения, подавления, предотвращения, отсрочки появления, уменьшения степени тяжести и/или снижения частоты возникновения одного или более симптомов или признаков конкретного заболевания, расстройства и/или состояния. Лечение можно проводить субъекту, который не демонстрирует признаков заболевания и/или демонстрирует только ранние признаки заболевания, для снижения риска развития патологии, связанной с заболеванием.

ПОДРОБНОЕ ОПИСАНИЕ

[0041] В настоящем изобретении предложены, помимо прочего, улучшенные способы и фармацевтические композиции для лечения муковисцидоза на основе

кодон-оптимизированной матричной РНК (мРНК), кодирующей муковисцидозный регулятор трансмембранной проводимости (CFTR). В частности, эти кодон-оптимизированные мРНК можно эффективно синтезировать в крупном масштабе, например, с использованием РНК-полимеразы SP6. Определенная кодон-оптимизированная мРНК может быть особенно полезной для получения однородного, безопасного и эффективного клинического продукта.

В некоторых вариантах осуществления в настоящем изобретении предложены способы получения фармацевтической композиции, содержащей мРНК, причем мРНК представляет собой транскрибированную *in vitro* мРНК, кодирующую муковисцидозный регулятор трансмембранной проводимости (CFTR), причем транскрибированная *in vitro* мРНК синтезируется на основе ДНК-шаблона с использованием РНК-полимеразы SP6 и причем синтез транскрибированной *in vitro* мРНК не приводит к получению вторичного полинуклеотидного фрагмента длиной приблизительно 1800 нуклеотидов.

Муковисцидоз

[0042] Настоящее изобретение может использоваться для лечения субъекта, страдающего от муковисцидоза или подверженного ему. Муковисцидоз представляет собой генетическое расстройство, характеризующееся мутациями в гене муковисцидозного регулятора трансмембранной проводимости (CFTR). Белок CFTR функционирует как трансмембранный канал в клетках, продуцирующих слизь, пот, слюну, слезы и пищеварительные ферменты. Канал пропускает отрицательно заряженные частицы, называемые хлорид-ионами, в клетки и из них. Транспорт хлорид-ионов помогает контролировать перемещение воды в тканях, что необходимо для продукции жидкой и свободно текучей слизи. Слизь представляет собой скользкое вещество, которое смазывает и защищает выстилку дыхательных путей, пищеварительной системы, репродуктивной системы и других органов и тканей.

[0043] К симптомам муковисцидоза со стороны дыхательной системы относятся: стойкий кашель с отделением густой слизи (мокроты), свистящее дыхание, одышка, непереносимость нагрузок, рецидивирующие легочные инфекции и воспаленные носовые проходы или заложенный нос. К симптомам муковисцидоза со стороны пищеварительной системы относятся: зловонный, жирный стул,

недостаточный набор массы и роста, непроходимость кишечника, в особенности у новорожденных (меконияльный илеус), и тяжелый запор.

Кодон-оптимизированная мРНК, кодирующая CFTR

[0044] В некоторых вариантах осуществления в настоящем изобретении предложены способы и композиции для доставки кодон-оптимизированной мРНК, кодирующей CFTR, субъекту для лечения муковисцидоза. Подходящая кодон-оптимизированная мРНК CFTR кодирует любой полноразмерный белок CFTR, его фрагмент или часть, который может заместить активность встречающегося в природных условиях белка CFTR и/или снизить интенсивность, степень тяжести и/или частоту появления одного или более симптомов, связанных с муковисцидозом.

[0045] В некоторых вариантах осуществления подходящая последовательность кодон-оптимизированной мРНК представляет собой последовательность мРНК, кодирующую человеческий белок CFTR (hCFTR). Примеры кодирующей кодон-оптимизированной последовательности мРНК CFTR и соответствующей аминокислотной последовательности показаны в таблице 1.

Таблица 1. Пример кодон-оптимизированного человеческого CFTR

SEQ ID NO: 1	<p>AUGCAACGCUCUCCUCUUGAAAAGGCCUCGGUGGUGUCCAAGCUC UUCUUCUCGUGGACUAGACCCAUCCUGAGAAAGGGGUACAGACAG CGCUUGGAGCUGUCCGAUAUCAAAAUCCCUUCCGUGGACUCC GCGGACAACCUGUCCGAGAAGCUCGAGAGAGAAUGGGACAGAGAA CUCGCCUCAAGAAGAACCCGAAGCUGAUUAAUGCGCUUAGGCGG UGCUIUUUCUGGCGGUUCAUGUUCUACGGCAUCUCCUCUACCUG GGAGAGGUCACCAAGGCCGUGCAGCCCCUGUUGCUGGGACGGAUU AUUGCCUCCUACGACCCCGACAACAAGGAAGAAAGAAGCAUCGCU AUCUACUUGGGCAUCGGUCUGUGCCUGCUUUUCAUCGUCCGGACC CUCUUGUUGCAUCCUGCUAUUUUCGGCCUGCAUCACAUUGGCAUG CAGAUGAGAAUUGCCAUGUUUCCCUGAUCUACAAGAAAACUCUG AAGCUCUCGAGCCGCGUGCUUGACAAGAUUCCAUCGGCCAGCUC GUGUCCCUGCUCUCCAACAUCUGAACAAGUUCGACGAGGGCCUC GCCUGGCCACUUCGUGUGGAUCGCCCCUCUGCAAGUGGCGCUU CUGAUGGGCCUGAUCUGGGAGCUGCUGCAAGCCUCGGCAUUCUGU GGGCUUGGAUCCUGAUCGUGCUGGCACUGUCCAGGCCGGACUG</p>
-------------------------	---

GGGCGGAUGAUGAUGAAGUACAGGGACCAGAGAGCCGGAAAGAUAU
UCCGAACGGCUGGUGAUCACUUCGGAAAUGAUCGAAAACAUC CAG
UCAGUGAAGGCCUACUGCUGGGAAGAGGCCAUGGAAAAGAUGAUA
GAAAACCUCGGCAAACCGAGCUGAAGCUGACCCGCAAGGCCGCU
UACGUGCGCUAUUUCAACUCGUCGCGCUUUCUUCUUCUCGGGUUC
UUCGUGGUGUUCUCUCUCCGUGCUCUCCCUACGCCUGAUUAAGGGA
AUCAUCCUCAGGAAGAUCUUCACCACCAUUCUUCUGUAUCGUG
CUCCGCAUGGCCGUGACCCGGCAGUUCUCAUGGGCCGUGCAGACU
UGGUACGACUCCUGGGAGCCAUUAACAAGAUC CAGGACUUCUUC
CAAAGCAGGAGUACAAGACCCUCGAGUACAACCGACUACUACC
GAGGUCGUGAUGGAAAACGUCACCGCCUUUUGGGAGGAGGGAUUU
GGCGAACUGUUCGAGAAGGCCAAGCAGAACAACAACAACCGCAAG
ACCUCGAACGGUGACGACUCCUCUUCUUCUUCUCAAACUUCAGCCUG
CUCGGGACGCCCGUGCUGAAGGACAUUAACUUC AAGAUCGAAAGA
GGACAGCUCUCCUGGCCGUGGCCGGAUCGACCGGAGCCGGAAAGACU
UCCUCGUGAUGGUGAUC AUGGGAGAGCUUGAACCUAGCGAGGGA
AAGAUC AAGCACUCCGGCCGAUCAGCUUCUGUAGCCAGUUUUC
UGGAUCAUGCCCGAAACCAUUAAGGAAAACAUCAUCUUCGGCGUG
UCCUACGAUGAAUACCGCUACCGGUC CGUGAUC AAAGCCUGCCAG
CUGGAAGAGGAUAUUUCAAGUUCGCGGAGAAAGAUACAUCGUG
CUGGGCGAAGGGGGUAUUACCUUGUCGGGGGGCCAGCGGGCUAGA
AUCUCGCUGGCCAGAGCCGUGUAUAAGGACGCCGACCUGUAUCUC
CUGGACUCCCCUUCGGAUACCUGGACGUCCUGACCGAAAAGGAG
AUCUUCGAAUCGUGCGUGUGCAAGCUGAUGGCUAACAAGACUCGC
AUCCUCGUGACCUC AAAUUGGAGCACCUGAAGAAGGCAGACAAG
AUUCUGAUUCUGCAUGAGGGGUCUCCUACUUUUACGGCACCUUC
UCGGAGUUGCAGAACUUGCAGCCGACUUCUCAUCGAAGCUGAUG
GGUUGCGACAGCUUCGACCAGUUCUCCGCCGAAAGAAGGAACUCG
AUCCUGACGGAACCUUGCACCGCUUCUCUUGGAAGGGCGACGCC
CCUGUGUCAUGGACCGAGACUAAGAAGCAGAGCUUCAAGCAGACC
GGGGAAUUCGGCGAAAAGAGGAAGAACAGCAUCUUGAACCCCAU
AACUCCAUCCGCAAGUUCUCAAUCGUGCAAAAGACGCCACUGCAG
AUGAACGGCAUUGAGGAGGACUCCGACGAACCCCUUGAGAGGGCGC
CUGUCCUGGUGCCGGACAGCGAGCAGGGAGAAGCCAUC CUGCCU
CGGAUUUCGUGAUCUCCACUGGUCGACGCUCCAAGCCC GGCGG
CGGCAGUCCGUGCUGAACCUGAUGACCCACAGCGUGAACCAGGGC
CAAAACAUCACCGCAAGACUACCGCAUCCACCCGGAAAGUGUCC
CUGGCACCUCAAGCGAAUCUUAACCGAGCUCGACAUCUACUCCCGG
AGACUGUCGCAGGAAACCGGGCUCGAAAUUUCCGAAGAAAUCAAC
GAGGAGGAUCUGAAAGAGUGCUUCUUCGACGAUAUGGAGUCGAUA
CCCGCCGUGACGACUUGGAACACUUAUCUGCGGUACAUCACUGUG
CACAAGUCAUUGAUCUUCGUGCUGAUUUGGUGCCUGGUGAUUUUC
CUGGCCGAGGUCGCGGCCUCACUGGUGGUGCUCUGGCUGUUGGGA

	AACACGCCUCUGCAAGACAAGGGAAACUCCACGCACUCGAGAAAC AACAGCUAUGCCGUGAUUAUCACUCCACCUCUUAUACGUG UUCUACAUCUACGUCGGAGUGGGCGGAUACCCUGCUCGCGAUGGGU UUCUUCAGAGGACUGCCGCUGGUCCACACCUUGAUCACCGUCAGC AAGAUUCUUCACCACAAGAUGUUGCAUAGCGUGCUGCAGGCCCCC AUGUCCACCCUCAACACUCUGAAGGCCGGAGGCAUUCUGAACAGA UUCUCCAAGGACAUCGCUAUCUGGACGAUCUCCUGCCGCUUACC AUCUUUGACUUCAUCCAGCUGCUGCUGAUCGUGAUUGGAGCAAUC GCAGUGGGUGGGCUGCUGCAGCCUACAUUUUCGUGGCCACUGUG CCGGUCAUUGUGGGCGUUCAUCAUGCUGCGGGCCUACUCCUCCAA ACCAGCCAGCAGCUGAAGCAACUGGAAUCCGAGGGACGAUCCCCC AUCUUCACUCACCUUGUGACGUCGUUGAAGGGACUGUGGACCCUC CGGGCUUUCGGACGGCAGCCCUACUUCGAAACCCUCUCCACAAG GCCUGAACCUCACACCGCCA AUUGGUUCCUGUACCUGUCCACCC UGCGGUGGUUCCAGAUGCACAUCGAGAUGAUUUUCGUCAUCUUCU UCAUCGCGGUCACAUUCAUUCAGCAUCCUGACUACCGGAGAGGGAG AGGGACGGGUCGGAUAAUCCUGACCCUCGCAUGAACAUUAUGA GCACCCUGCAGUGGGCAGUGAACAGCUCGAUCGACGUGGACAGCC UGAUGC GAAGCGUCAGCCGCGUGUUCAAGUUCAUCGACAUGCCUA CUGAGGGAAAACCCACUAAGUCCACUAAGCCCUACAAAAAUGGCC AGCUGAGCAAGGUCAUGAUCAUCGAAAACUCCACGUGAAGAAGG ACGAUAAUUGGCCCUCCGGAGGUCAA AUGACCGUGAAGGACCUGA CCGCAAAGUACACCGAGGGAGGAAACGCCAUUCUGAAAACAUCA GCUUCUCCAUUUCGCCGGGACAGCGGGUCGGCCUUCUCGGGCGGA CCGGUUCGGGAAGUCAACUCUGCUGUCGGCUUUCUCCGGCUGC UGAAUACCGAGGGGGAAAUCCAAAUUGACGGCGUGUCUUGGGAU CCAUAUCUCUGCAGCAGUGGCGGAAGGCCUUCGGCGUGAUCCCCC AGAAGGUGUUCAUCUUCUCGGGUACCUUCCGGAAGAACCUGGAUC CUUACGAGCAGUGGAGCGACCAAGAAAUCUGGAAGGUCGCCGACG AGGUCGGCCUGCGCUCGUGAUUGAACAAUUCUGGAAAGCUGG ACUUCGUGCUCGUCGACGGGGGAUGUGUCCUGUCGCACGGACUA AGCAGCUCAUGUGCCUCGCACGGUCCGUGCUCUCAAGGCCAAGA UUCUGCUGCUGGACGAACCUUCGGCCACCUGGAUCCGGUCACCU ACCAGAUCAUCAGGAGGACCUGAAGCAGGCCUUGCCGAUUGCA CCGUGAUUCUCUGCGAGCACCGCAUCGAGGCCAUGCUGGAGUGCC AGCAGUUCUGGUCAUCGAGGAGAACAAAGGUCCGCCAAUACGACU CCAUUCAAAAGCUCCUCAACGAGCGGUCGUGUCAGACAAGCUA UUUCACCGUCCGAUAGAGUGAAGCUCUUCGCCGAUCGGAACAGCU CAAAGUGCAAUCGAAGCCGCAGAUCGCAGCCUUGAAGGAAGAGA CUGAGGAAGAGGUGCAGGACACCCGGCUUUA
SEQ ID NO: 2	AUGCAGCGGUCCCCGUCGAAAAGGCCAGUGUCGUGUCCAAACUC UUCUUCUCAUGGACUCGGCCUAUCCUUAAGAAAGGGGUAUCGGCAG AGGCUUGAGUUGUCUGACAUCUACCAGAUCCCCUCGGUAGAUUCG

GCGGAUAACCUCUCGGAGAAGCUCGAACGGGAAUGGGACCGCGAA
CUCGCGUCUAAGAAAAACCCGAAGCUCAUCAAACGCACUGAGAAGG
UGCUCUCUCUGGGCGGUUCAUGUUCUACGGUAUCUUCUUGUAUCUC
GGGGAGGUCACAAAAGCAGUCCAACCCUGUUGUUGGGUCGCAUU
AUCGCCUCGUACGACCCCGAUAACAAAGAAGAACGGAGCAUCGCG
AUCUACCUCGGGAUCGGACUGUGUUUGCUUUUCAUCGUCAGAACA
CUUUUGUUGCAUCCAGCAAUCUUCGGCCUCCAUCACAUCGGUAUG
CAGAUGCAGAAUCGCUAUGUUUAGCUUGAUCUACAAAAGACACUG
AAACUCUCGUCGCGGGUGUUGGAUAAGAUUCCAUCGGUCAGUUG
GUGUCCCUGCUUAGUAAUAACCUCAACAAAUUCGAUGAGGGACUG
GCGCUGGCACAUUUCGUGUGGAUUGCCCCGUUGCAAGUCGCCCUU
UUGAUGGGCCUUAUUUGGGAGCUGUUGCAGGCAUCUGCCUUUUGU
GGCCUGGGAAUUCUGAUUGUGUUGGCAUUGUUUCAGGCUGGGCUU
GGGCGGAUGAUGAUGAAGUAUCGCGACCAGAGAGCGGGUAAAUC
UCGGAAAGACUCGUCAUCACUUCGGAAAUGAUCGAAAACAUC CAG
UCGGUCAAAAGCCUAUUGCUGGGAAGAAGCUAUGGAGAAGAUGAUU
GAAAACCUCGCGCAACUGAGCUGAAACUGACCCGCAAGGCGGGCG
UAUGUCCGGUAUUUCAAUUCGUCAGCGUUCUUCUUUCCGGGUUC
UUCGUUGUCUUUCUCUCGGUUUUGCCUUAUGCCUUGAUUAAGGGG
AUUAUCCUCGCAAGAUUUUCACCACGAUUUCGUUCUGCAUUGUA
UUGCGCAUGGCAGUGACACGGCAAUUUCGUGGGCCGUGCAGACA
UGGUAUGACUCGCUUGGAGCGAUCACAAAUAUCAAGACUUCUUG
CAAAAGCAAGAGUACAAGACCCUGGAGUACAUCUACUACUACG
GAGGUAGUAAUGGAGAAUGUGACGGCUUUUUGGGAAGAGGGUUU
UGGAGAACUGUUUGAGAAAGCAAAGCAGAAUAACAACAACCGCAA
GACCUCAAAUGGGGACGAUUCUUCGUUUUUCUCGAACUUCUCCU
GCUCGGAACACCCGUGUUGAAGGACAUCAAUUUCAAGAUUGAGAG
GGGACAGCUUCUCGCGGUAGCGGGAAGCACUGGUGCGGGAAAAC
UAGCCUCUUGAUGGUGAUUAUGGGGGAGCUUGAGCCAGCGAGGG
GAAGAUUAAACACUCCGGGCGUAUCUCAUUCUGUAGCCAGUUUUC
AUGGAUCAUGCCC GGAACCAUUAAGAGAACAUCAUUUUCGGAGU
AUCCUAUGAUGAGUACCGAUACAGAUCGGUCAUUAAGGCGUGCCA
GUUGGAAGAGGACAUUUCUAAGUUCGCCGAGAAGGAUAACAUCGU
CUUGGGAGAAGGGGGUAUUACAUUGUCGGGAGGGCAGCGAGCGCG
GAUCAGCCUCGCGAGAGCGGUUAUACAAGAUGCAGAUUUGUAUCU
GCUUGAUUCACCGUUUGGAUACCUCGACGUAUUGACAGAAAAGA
AAUCUUCGAGUCGUGCGUGUGUAAACUUAUGGCUAAUAAGACGAG
AAUCCUGGUGACAUCAAAAAUGGAACACCUUAAGAAGGCGGACAA
GAUCCUGAUCCUCCACGAAGGAUCGUCCUACUUUACGGCACUUU
CUCAGAGUUGCAAAACUUGCAGCCGGACUUCUCAAGCAAACUCAU
GGGGUGUGACUCAUUCGACCAGUUCAGCGCGGAACGGCGGAACUC
GAUCUUGACGGAAACGCUGCACCGAUUCUCGCUUGAGGGUGAUGC
CCCGGUAUCGUGGACCGAGACAAAGAAGCAGUCGUUUAAGCAGAC

AGGAGAAUUUGGUGAGAAAAGAAAGAACAGUAUCUUGAAUCCUA
UUAACUCAAUUCGCAAGUUCUCAAUUCGUCCAGAAAACUCCACUGC
AGAUGAAUGGAAUUGAAGAGGAUUCGGACGAACCCUGGAGCGCA
GGCUUAGCCUCGUGCCGGAUUCAGAGCAAGGGGAGGCCAUUCUUC
CCCGGAUUUCGGUGAUUUCAACCGGACCUACACUUCAGGCGAGGC
GAAGGCAAUCCGUGCUCAACCUCAUGACGCAUUCGGUAAAACCAGG
GGCAAAACAUUCACCGCAAAACGACGGCCUCAACGAGAAAAGUGU
CACUUGCACCCCAGGCGAAUUUGACUGAACUCGACAUCUACAGCC
GUAGGCUUUCGCAAGAAACCGGACUUGAGAUCAGCGAAGAAAUCA
AUGAAGAAGAUUUGAAAGAGUGUUUCUUUGAUGACAUGGAAUCA
AUCCCAGCGGUGACAACGUGGAACACAUAUCUUGCGUUACAUCACG
GUGCACAAGUCCUUGAUUUUCGUCCUCAUCUGGUGUCUCGUGAUC
UUUCUCGCUGAGGUCGCAGCGUCACUUGUGGUCCUCUGGCUGCUU
GGUAAUACGCCCUUGCAAGACAAAGGCAAUUCUACACACUCAAGA
AACAAUCCUAUGCCGUGAUUAUCACUUCUACAAGCUCGUAAUAC
GUGUUUUAUCAUCUACGUAGGAGUGGCCGACACUCUGCUCGCGAUG
GGUUUCUUCGAGGACUCCACUCGUUCACACGCUUAUCACUGUC
UCCAAGAUUCUCCACCAUAAGAUGCUCUACUAGCGUACUGCAGGCU
CCAUGUCCACCUUGAAUACGCUCAAGGCGGGAGGUAAUUUGAAU
CGCUUCUCAAAAAGAUUUUGCAAUUUUGGAUGACCUUCUGCCCCUG
ACGAUCUUCGACUUCAUCCAGUUGUUGCUGAUCGUGAUUGGGGCU
AUUGCAGUAGUCGCUGUCCUCCAGCCUACAUUUUUGUCGCGACC
GUUCCGGUGAUCGUGGCGUUUAUCAUGCUCGCGGGCCUAAUUCUUG
CAGACGUCACAGCAGCUUAAGCAACUGGAGUCUGAAGGGAGGUCG
CCUAUCUUUACGCAUCUUGUGACCAGUUUGAAGGGAUUGUGGACG
UUGCGCGCCUUUGGCAGGCAGCCCUACUUUGAAACACUGUCCAC
AAAGCGCUGAAUCUCCAUACGGCAAUUGGUUUUUGUAUUUGAGU
ACCCUCCGAUGGUUUCAGAUGCAGCAUUGAGAUGAUUUUUUGUGAUC
UUCUUUAUCGCGGUGACUUUAUCUCCAUCUUGACCACGGGAGAG
GGCGAGGGACGGGUCGGUAUUAUCCUGACACUCGCCAUGAACAUU
AUGAGCACUUUGCAGUGGGCAGUGAACAGCUCGAUUGAUGUGGAU
AGCCUGAUGAGGUCCGUUUCGAGGGUCUUUAAGUUCAUCGACAUG
CCGACGGAGGGAAAGCCACAAAAGUACGAAACCCUAUAAGAAU
GGCAAUUGAGUAAGGUAAUGAUCUUCGAGAACAGUCACGUGAAG
AAGGAUGACAUCUGGCCUAGCGGGGGUCAGAUGACCGUGAAGGAC
CUGACGGCAAAAUACACCGAGGGAGGGAACGCAAUCCUUGAAAAC
AUCUCGUUCAGCAUUAGCCCCGGUCAGCGUGUGGGGUUGCUCGGG
AGGACCGGGUCAGGAAAUCGACGUUGCUCUGCGGCCUUCUUGAGA
CUUCUGAAUACAGAGGGUGAGAUCAGAUUCGACGGCGUUUCGUGG
GAUAGCAUCACCUUGCAGCAGUGGCGGAAAGCGUUUGGAGUAAUC
CCCCAAAAGGUCUUUAUCUUUAGCGGAACCUUCCGAAAGAAUCUC
GAUCCUUAUGAACAGUGGUCAGAUC AAGAGAUUUGGAAAGUCGCG
GACGAGGUUGGCCUUCGGAGUGUAAUCGAGCAGUUUCCGGGAAAA

	<p>CUCGACUUUGUCCUUGUAGAUGGGGGAUGCGUCCUGUCGCAUGGG CACAAGCAGCUCAUGUGCCUGGCGCGAUCCGUCCUCUCUAAAGCG AAAAUUCUUCUCUUGGAUGAACCUUCGGCCCAUCUGGACCCGGUA ACGU AUCAGAUCAUCAGAAGGACACUUAAGCAGGCGUUUGCCGAC UGCACGGUGAUUCUCUGUGAGCAUCGUAUCGAGGCCAUGCUCGAA UGCCAGCAAUUUCUUGUCAUCGAAGAGAAUAAGGUCCGCCAGUAC GACUCCAUCAGAAAGCUGCUUAAUGAGAGAUCAUUGUUCGGCAG GCGAUUUCACCAUCCGAUAGGGUGAAACUUUUUCCACACAGAAU UCGUCGAAGUGCAAGUCCAAACCGCAGAUCGCGGCCUUGAAAGAA GAGACUGAAGAAGAAGUUCAAGACACGCGUCUUUAA</p>
<p>Последова- тельность челове- ческого белка CFTR</p>	<p>MQRSPLEKASVVSKLFFSWTRPILRKGYRQRLELSDIYQIPSVDSADNLS EKLEREWRELASKKNPKLINALRRCFFWRFMFYGFILYLGEVTKAVQ PLLLGRIIASYDPDNKEERSIAIYLGIGLCLLFIVRLLLHPAIFGLHHIGM QMRIAMFSLIYKKTLLSSRVLDKISIGQLVSLLSNNLNKFDLGLALAHF VWIAPLQVALLMGLIWELLQASAFGLGFLIVLALFQAGLGRMMMKY RDQRAGKISERLVITSEMIENIQSVKAYCWEAMEKMIENLRQTELKLT RKAAYVRYFNSSAFFSFFVFLSVLPYALIKGIILRKIFTTISFCIVLRM AVTRQFPWAVQTWYDSLGAINKIQDFLQKQEYKTLEYNLTTTEVME NVTAFWEEGFELFEKAKQNNNNRKT SNGDDSLFFSNFSLLGTPVLKDI NFKIERGQLLAVAGSTGAGKTSLLMVIMGELEPSEGKIKHSGRISFCSQF SWIMPGTIKENIIFGVSYDEYRYSVIKACQLEEDISKFAEKDNIVLGEGG ITLSGGQRARISLARAVYKDADLYLLDSPFGYLDVLTEKEIFESCVCCKL MANKTRILVTSKMEHLKKADKILILHEGSSYFYGTSELQNLQPDFSSKL MGCDSFDQFSAERRNSILTETLHRFSLEGDAPVSWTETKQSFQKQTFGEF GEKRKNSILNPINSIRKFSIVQKTPLQMNGIEEDSDEPLERRLSLVPDSEQ GAILPRISVISTGPTLQARRRQSVLNLMTHSV NQGQNIHRKTTASTRKV SLAPQANLTELDIYSRRLSQETGLEISEEINEEDLKECFDDMESIPAVTT WNTYLRYITVHKSLIFVLIWCLVIFLAEVAASLVVLWLLGNTPLQDKGN STHSRNNSYAVIITSTSSYYVFYIYVGVADTLLAMGFFRGLPLVHTLITV SKILHHKMLHSV LQAPMSTLNTLKAGGILNRFSKDIAILDDLPLTIFDFI QLLLIVIGAI VVAVLQPYIFVATVPVIVAFIMLRAYFLQTSQQLKQLESE GRSPIFTHL VTSKGLWTLRAFGRPYFETLFHKALNLHTANWFLYLST LRWFQMRIEMIFVIFIAVTFISILTTGEGEGRVGIILTLAMNIMSTLQWA VNSSIDVDSL MRSVSRVFKFIDMPTEGKPTKSTKPYKNGQLSKVMIIENS HVKKDDIWPSGGQMTVKDLTAKYTEGGNAILENISFSISPGQRVGLLGR TGSGKSTLLSAFLRLLNTEGEIQIDGVS WDSITLQQWRKAFGVIPQKVFI FSGTFRKNLDPYEQWSDQEIWKVADEVGLRSVIEQFPGKLDLFDVLDGG CVLSHGHKQLMCLARSVLSKAKILLDEPSAHLDPV TYQIIRRTLKQAF ADCTVILCEHRIEAMLECCQFLVIEENKVRQYDSIQKLLNERSLFRQAIS PSDRVKLFPHRNSSKCKSKPQIAALKEETEEVQDTRL (SEQ ID NO: 3)</p>

[0046] Дополнительные примеры кодон-оптимизированных последовательностей мРНК описаны в разделе «Примеры» ниже, например, SEQ ID NO: 7 и SEQ ID NO: 8, обе из которых включают в себя нетранслируемые области 5' и 3', кадрирующие кодон-оптимизированную кодирующую hCFTR мРНК, и SEQ ID NO: 27–SEQ ID NO: 40.

[0047] В некоторых вариантах осуществления подходящая последовательность мРНК может представлять собой последовательность мРНК, кодирующую гомолог или аналог человеческого белка CFTR (hCFTR). Например, гомолог или аналог белка hCFTR может представлять собой модифицированный белок hCFTR, содержащий одну или более аминокислотных замен, делеций и/или вставок по сравнению с белком hCFTR дикого типа или встречающегося в природных условиях, сохраняя при этом существенную активность белка hCFTR. В некоторых вариантах осуществления мРНК, подходящая для настоящего изобретения, кодирует аминокислотную последовательность, которая по меньшей мере на 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более гомологична SEQ ID NO: 3. В некоторых вариантах осуществления мРНК, подходящая для настоящего изобретения, кодирует белок, по существу идентичный белку hCFTR. В некоторых вариантах осуществления мРНК, подходящая для настоящего изобретения, кодирует аминокислотную последовательность, которая по меньшей мере на 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более идентична SEQ ID NO: 3. В некоторых вариантах осуществления мРНК, подходящая для настоящего изобретения, кодирует фрагмент или участок белка hCFTR. В некоторых вариантах осуществления мРНК, подходящая для настоящего изобретения, кодирует фрагмент или участок белка hCFTR, причем фрагмент или участок белка продолжает сохранять активность CFTR, аналогичную активности белка дикого типа. В некоторых вариантах осуществления мРНК, подходящая для настоящего изобретения, имеет нуклеотидную последовательность, которая по меньшей мере на 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более идентична SEQ ID NO: 1, SEQ ID NO: 7 или SEQ ID NO: 8.

[0048] В некоторых вариантах осуществления мРНК, подходящая для настоящего изобретения, имеет нуклеотидную последовательность, которая по

меньшей мере на 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более идентична любой одной из SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39 или SEQ ID NO: 40.

[0049] В некоторых вариантах осуществления подходящая мРНК кодирует слитый белок, содержащий полноразмерный белок hCFTR, его фрагмент или участок, слитый с другим белком (например, слияние на N- или C-конце). В некоторых вариантах осуществления белок, слитый с мРНК, кодирующей полноразмерный белок hCFTR, его фрагмент или участок, кодирует сигнальную или нацеливающую на клетки последовательность.

Синтез мРНК

[0050] мРНК в соответствии с настоящим изобретением могут быть синтезированы в соответствии с любым из разнообразных известных способов. Например, мРНК в соответствии с настоящим изобретением могут быть синтезированы посредством транскрипции *in vitro* (IVT). Вкратце, IVT, как правило, выполняется с помощью содержащего промотор линейного или кольцевого ДНК-шаблона, пула рибонуклеотидтрифосфатов, буферной системы, которая может содержать ДТТ и ионы магния, и соответствующей РНК-полимеразы (например, РНК-полимеразы Т3, Т7 или SP6), ДНКазы I, пирофосфатазы и/или ингибитора РНКазы. Точные условия могут меняться в зависимости от конкретного применения.

[0051] В некоторых вариантах осуществления для получения мРНК в соответствии с изобретением ДНК-шаблон транскрибируют *in vitro*. Подходящий ДНК-шаблон, как правило, содержит промотор, например, промотор Т3, Т7 или SP6, для транскрипции *in vitro*, за которым следует требуемая нуклеотидная последовательность для требуемой мРНК и сигнал терминации.

Синтез мРНК с использованием РНК-полимеразы SP6

[0052] В некоторых вариантах осуществления мРНК CFTR получают с использованием РНК-полимеразы SP6. РНК-полимераза SP6 представляет собой ДНК-

зависимую РНК-полимеразу с высокой специфичностью к промоторным последовательностям SP6. Полимераза SP6 катализирует синтез РНК 5' → 3' *in vitro* либо на одноцепочечной ДНК, либо на двухцепочечной ДНК после своего промотора; она встраивает нативные рибонуклеотиды, и/или модифицированные рибонуклеотиды, и/или меченые рибонуклеотиды в полимеризованный транскрипт. К примерам таких меченых рибонуклеотидов относятся нуклеотиды, меченные биотином, флуоресцеином, дигоксигенином, аминоаллилом и изотопами.

[0053] Первоначально последовательность РНК-полимеразы бактериофага SP6 (GenBank: Y00105.1) была описана как имеющая следующую аминокислотную последовательность:

MQDLHAIQLQLEEEMFNGGIRRFADQQRQIAAGSESDTAWNRRLLSELIAPMAEGI
QAYKEEYEGKKGRAPRALAFLQCVENEVAAYITMKVVM DMLNTDQAIAMSVA
ERIEDQVRFSKLEGHAAKYFEKVKKSLKASRTKSYRHAHNVAVVAEKSVAEKDADF
DRWEAWPKETQLQIGTTLLEILEGSVIFYNGEPVFMRAMRTYGGKTIYYLQTSSEVGG
WISAFKEHVAQLSPAYAPCVIPPRPWRTPFNNGGFHTEKVASRIRLVKGNREHVRKLT
QKQMPKVYKAINALQNTQWQINKDVLAVIEEVIRLDLGYGVPSFKPLIDKENKPANP
VPVEFQHILRGRELKEMLSPEWQQFINWKGEARLYTAETKRGSKSAAVVRMVGQ
ARKYSAFESIYFVYAMDSRSRVYVQSSTLSPQSN DLGKALLRFTEGRPVNGVEALK
WFCINGANLWGW DKKTFDVRVSNVLDEEFQDMCRDIAADPLTFTQWAKADAPYEF
LAWCFEYAQYLDLVDEGRADEFRTHLPVHQDGSCSGIQHYSAMLRDEVGAKAVNL
KPSDAPQDIYGAVAQVVIKKNALYMDADDATFTTSGSVT LSGTELRAMASAWDSIGI
TRSLTKKPVMTLPYGSTRLTCRESVIDYIVDLEEKEAQKAVAEGRTANKVHPFEDDR
QDYLTPGAAYNYMTALIWPSISEVVKAPIVAMKMIRQLARFAAKRNEGLMYTLPTG
FILEQKIMATEMLRVRTCLMGDIKMSLQVETDIVDEAAMMGAAAPNFVHGH DASHL
ILTVCELVDKGVTSIAVIHDSFGTHADNTLTLRV ALKGQMVAMYIDGNALQKLLEEH
EVRWMVDTGIEVPEQGEFDLNEIMDSEYVFA (SEQ ID NO: 9).

[0054] РНК-полимераза SP6, подходящая для настоящего изобретения, может представлять собой любой фермент, имеющий по существу ту же полимеразную активность, что и РНК-полимераза бактериофага SP6. Таким образом, в некоторых вариантах осуществления РНК-полимеразы SP6, подходящая для настоящего изобретения, может быть модифицирована на основе SEQ ID NO: 9. Например, подходящая РНК-полимераза SP6 может содержать одну или более аминокислотных

замен, делеций или добавлений. В некоторых вариантах осуществления подходящая РНК-полимераза SP6 имеет аминокислотную последовательность, которая приблизительно на 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, 85%, 84%, 83%, 82%, 81%, 80%, 75%, 70%, 65% или 60% идентична или гомологична SEQ ID NO: 9. В некоторых вариантах осуществления подходящая РНК-полимераза SP6 может представлять собой усеченный белок (с N-конца, C-конца или внутри последовательности), но с сохранением полимеразной активности. В некоторых вариантах осуществления подходящая РНК-полимераза SP6 представляет собой слитый белок.

[0055] РНК-полимераза SP6, подходящая для изобретения, может представлять собой доступный в продаже продукт, например, производства компании Aldevron, Ambion, New England Biolabs (NEB), Promega и Roche. SP6 можно заказать и/или специально разработать с использованием коммерческого источника или некоммерческого источника в соответствии с аминокислотной последовательностью SEQ ID NO: 9 или вариантом SEQ ID NO: 9, как описано в настоящем документе. SP6 может представлять собой полимеразу со стандартной точностью или может представлять собой полимеразу с высокой точностью / высокой эффективностью / высокой производительностью, которая была модифицирована для стимулирования активности РНК-полимеразы, например, с мутациями в гене РНК-полимеразы SP6 или с посттрансляционными модификациями самой РНК-полимеразы SP6. Примеры таких модифицированных SP6 включают в себя SP6 RNA Polymerase-Plus™ от Ambion, HiScribe SP6 от NEB и системы RiboMAX™ и Riboprobe® от Promega.

[0056] В некоторых вариантах осуществления подходящая РНК-полимераза SP6 представляет собой слитый белок. Например, РНК-полимераза SP6 может включать в себя одну или более меток для облегчения выделения, очистки или растворимости фермента. Подходящая метка может быть расположена на N-конце, C-конце и/или внутри последовательности. К не имеющим ограничительного характера примерам подходящей метки относятся кальмодулин-связывающий белок (CBP); 8 кДа антиген *Fasciola hepatica* (Fh8); маркерный пептид FLAG; глутатион-S-трансфераза (GST); гистиридиновая метка (например, гексагистиридиновая метка (His6)); связывающий мальтозу белок (MBP); N-утилизирующее вещество (NusA); гибридная метка с малым модификатором, связанным с убиквитином (SUMO); стрептавидин-связывающий

пептид (STREP); тандемная аффинная очистка (TAP) и тиоредоксин (TrxA). В настоящем изобретении можно использовать и другие метки. Эти и другие слитые метки были описаны, например, в публикации Costa et al. *Frontiers in Microbiology* 5 (2014): 63 и в публикации PCT/US16/57044, содержание которых полностью включено в настоящий документ путем ссылки. В некоторых вариантах осуществления гистидиновая метка расположена на N-конце SP6.

Промотор SP6

[0057] В настоящем изобретении можно использовать любой промотор, который может быть распознан РНК-полимеразой SP6. Как правило, промотор SP6 содержит последовательность 5'-ATTTAGGTGACACTATAG-3' (SEQ ID NO: 10). Были обнаружены и/или созданы варианты промотора SP6 для оптимизации распознавания и/или связывания SP6 с ее промотором. Не имеющие ограничительного характера варианты включают в себя, без ограничений: 5'-ATTTAGGGGACACTATAGAAGAG-3'; 5'-ATTTAGGGGACACTATAGAAGG-3'; 5'-ATTTAGGGGACACTATAGAAGGG-3'; 5'-ATTTAGGTGACACTATAGAA-3'; 5'-ATTTAGGTGACACTATAGAAGA-3'; 5'-ATTTAGGTGACACTATAGAAGAG-3'; 5'-ATTTAGGTGACACTATAGAAGG-3'; 5'-ATTTAGGTGACACTATAGAAGGG-3'; 5'-ATTTAGGTGACACTATAGAAGNG-3' и 5'-CATACGATTTAGGTGACACTATAG-3' (SEQ ID NO: 11 – SEQ ID NO: 20).

[0058] Кроме того, подходящий промотор SP6 для настоящего изобретения может быть приблизительно на 95%, 90%, 85%, 80%, 75% или 70% идентичным или гомологичным любому одному из SEQ ID NO: 10 – SEQ ID NO: 20. Более того, промотор SP6, используемый в настоящем изобретении, может включать в себя один или более дополнительных нуклеотидов 5' и/или 3' к любой из промоторных последовательностей, описанных в настоящем документе.

ДНК-шаблон

[0059] Как правило, ДНК-шаблон CFTR является либо полностью двухцепочечным, либо преимущественно одноцепочечным с двухцепочечной промоторной последовательностью SP6.

[0060] В качестве шаблонов для транскрипции *in vitro* с использованием SP6 можно использовать линейризованную плазмидную ДНК (линейризованную посредством одного или более рестрикционных ферментов), линейризованные

геномные ДНК-фрагменты (посредством рестрикционного фермента и/или физических средств), ПЦР-продукты и/или синтетические ДНК-олигонуклеотиды при условии, что они содержат двухцепочечный промотор SP6, расположенный перед транскрибируемой последовательностью ДНК (и в правильной ориентации).

[0061] В некоторых вариантах осуществления линейаризованный ДНК-шаблон имеет тупой конец.

[0062] В некоторых вариантах осуществления транскрибируемая последовательность ДНК может быть оптимизирована для обеспечения более эффективной транскрипции и/или трансляции. Например, последовательность ДНК может быть оптимизирована в отношении цис-регуляторных элементов (например, ТАТА-боксы, сигналы терминации и сайты связывания с белками), сайтов искусственной рекомбинации, *chi*-сайтов, содержания динуклеотидов CpG, негативных по CpG островков, содержания GC, сайтов проскальзывания полимеразы и/или других элементов, относящихся к транскрипции; последовательность ДНК может быть оптимизирована в отношении скрытых сайтов сплайсинга, вторичной структуры мРНК, устойчивой свободной энергии мРНК, повторяющихся последовательностей, мотива нестабильности РНК и/или других элементов, относящихся к процессингу и стабильности мРНК; последовательность ДНК может быть оптимизирована в отношении отклонений в использовании кодонов, возможности адаптации кодонов, внутренних *chi*-сайтов, сайтов связывания с рибосомами (например, IRES), незрелых сайтов polyA, последовательностей Шайна — Дальгарно (SD) и/или других элементов, связанных с трансляцией; и/или последовательность ДНК может быть оптимизирована в отношении кодонного контекста, взаимодействий кодон — антикодон, сайтов трансляционных пауз и/или других элементов, связанных со сворачиванием белков. В настоящем изобретении можно использовать известные в данной области способы оптимизации, например, GeneOptimizer от ThermoFisher и OptimumGene™, описанные в публикации US 20110081708, содержание которой полностью включено в настоящий документ путем ссылки.

[0063] В некоторых вариантах осуществления ДНК-шаблон включает в себя 5'-и/или 3'-нетранслируемую область. В некоторых вариантах осуществления 5'-нетранслируемая область включает в себя один или более элементов, которые влияют на стабильность или трансляцию мРНК, например, чувствительный к железу элемент.

В некоторых вариантах осуществления 5'-нетранслируемая область может иметь длину от приблизительно 50 до 500 нуклеотидов.

[0064] В некоторых вариантах осуществления 3'-нетранслируемая область включает в себя один или более из сигнала полиаденилирования, сайта связывания белков, который влияет на стабильность местоположения мРНК в клетке, или одного или более сайтов связывания для миРНК. В некоторых вариантах осуществления 3'-нетранслируемая область может иметь длину в диапазоне от 50 до 500 нуклеотидов или более.

[0065] Примеры последовательностей 3' и/или 5' UTR могут быть получены на основе молекул мРНК, которые являются стабильными (например, глобин, актин, GAPDH, тубулин, гистон или ферменты цикла лимонной кислоты), для повышения стабильности смысловой молекулы мРНК. Например, последовательность 5' UTR может включать в себя частичную последовательность немедленного раннего гена 1 (IE1) ЦМВ или ее фрагмент для повышения устойчивости к нуклеазам и/или улучшения периода полужизни полинуклеотида. Также предусмотрено включение последовательности, кодирующей гормон роста человека (hGH) или его фрагмент, на 3'-конце или в нетранслируемой области полинуклеотида (например, мРНК) для дополнительной стабилизации полинуклеотида. Как правило, эти модификации повышают стабильность и/или улучшают фармакокинетические свойства (например, время полужизни) полинуклеотида по сравнению с их немодифицированными аналогами и включают в себя, например, модификации, внесенные для повышения устойчивости таких полинуклеотидов к расщеплению нуклеазами *in vivo*.

Крупномасштабный синтез мРНК

[0066] Настоящее изобретение относится к крупномасштабному производству кодон-оптимизированной мРНК CFTR. В некоторых вариантах осуществления способ в соответствии с изобретением позволяет синтезировать мРНК в количестве по меньшей мере 100 мг, 150 мг, 200 мг, 300 мг, 400 мг, 500 мг, 600 мг, 700 мг, 800 мг, 900 мг, 1 г, 5 г, 10 г, 25 г, 50 г, 75 г, 100 г, 250 г, 500 г, 750 г, 1 кг, 5 кг, 10 кг, 50 кг, 100 кг, 1000 кг или более за одну партию. В настоящем документе термин «партия» относится к количеству или объему мРНК, синтезированному за один раз, например, полученному по одному заказу на производство. Партия может означать количество мРНК, синтезированное в одной реакции, которая протекает с использованием одной

аликвоты фермента и/или одной аликвоты ДНК-шаблона в ходе непрерывного синтеза при одном и том же наборе условий. мРНК, синтезированная в одной партии, не будет содержать мРНК, синтезированную в другое время и скомбинированную для получения требуемого количества. Как правило, реакционная смесь включает в себя РНК-полимеразу SP6, линейный ДНК-шаблон и РНК-полимеразный реакционный буфер (который может включать в себя рибонуклеотиды или может требовать добавления рибонуклеотидов).

[0067] В соответствии с настоящим изобретением, как правило, используют 1–100 мг полимеразы SP6 на грамм (г) получаемой мРНК. В некоторых вариантах осуществления на грамм получаемой мРНК используют приблизительно 1–90 мг, 1–80 мг, 1–60 мг, 1–50 мг, 1–40 мг, 10–100 мг, 10–80 мг, 10–60 мг, 10–50 мг полимеразы SP6. В некоторых вариантах осуществления для получения приблизительно 1 грамма мРНК используют приблизительно 5–20 мг полимеразы SP6. В некоторых вариантах осуществления для получения приблизительно 0,5–2 граммов мРНК используют приблизительно 100 граммов полимеразы SP6. В некоторых вариантах осуществления для получения приблизительно 5–20 граммов мРНК используют приблизительно 1 килограмм полимеразы SP6. В некоторых вариантах осуществления для получения приблизительно 1 грамма мРНК используют приблизительно 5 мг полимеразы SP6. В некоторых вариантах осуществления для получения приблизительно 100 граммов мРНК используют приблизительно 500 мг полимеразы SP6. В некоторых вариантах осуществления для получения приблизительно 1 килограмма мРНК используют приблизительно 5 граммов полимеразы SP6. В некоторых вариантах осуществления на грамм получаемой мРНК используют приблизительно 10 мг, 20 мг, 30 мг, 40 мг, 50 мг, 60 мг, 70 мг, 80 мг, 90 мг или 100 мг плазмидной ДНК. В некоторых вариантах осуществления для получения приблизительно 1 грамма мРНК используют приблизительно 10–30 мг плазмидной ДНК. В некоторых вариантах осуществления для получения приблизительно 100 граммов мРНК используют приблизительно 1–3 грамма плазмидной ДНК. В некоторых вариантах осуществления для получения приблизительно 1 килограмма мРНК используют приблизительно 10–30 граммов плазмидной ДНК. В некоторых вариантах осуществления для получения приблизительно 1 грамма мРНК используют приблизительно 10 мг плазмидной ДНК. В некоторых вариантах осуществления для получения приблизительно 100 граммов мРНК используют приблизительно 1 грамм плазмидной ДНК. В некоторых вариантах

осуществления для получения приблизительно 1 килограмма мРНК используют приблизительно 10 граммов плазмидной ДНК.

[0068] В некоторых вариантах осуществления концентрация РНК-полимеразы SP6 в реакционной смеси может составлять от приблизительно 1 до 100 нМ, от 1 до 90 нМ, от 1 до 80 нМ, от 1 до 70 нМ, от 1 до 60 нМ, от 1 до 50 нМ, от 1 до 40 нМ, от 1 до 30 нМ, от 1 до 20 нМ или от приблизительно 1 до 10 нМ. В некоторых вариантах осуществления концентрация РНК-полимеразы SP6 составляет от приблизительно 10 до 50 нМ, от 20 до 50 нМ или от 30 до 50 нМ. Можно использовать концентрацию от 100 до 10 000 единиц/мл РНК-полимеразы SP6, например, от 100 до 9000 единиц/мл, от 100 до 8000 единиц/мл, от 100 до 7000 единиц/мл, от 100 до 6000 единиц/мл, от 100 до 5000 единиц/мл, от 100 до 1000 единиц/мл, от 200 до 2000 единиц/мл, от 500 до 1000 единиц/мл, от 500 до 2000 единиц/мл, от 500 до 3000 единиц/мл, от 500 до 4000 единиц/мл, от 500 до 5000 единиц/мл, от 500 до 6000 единиц/мл, от 1000 до 7500 единиц/мл и от 2500 до 5000 единиц/мл.

[0069] Концентрация каждого рибонуклеотида (например, АТФ, УТФ, ГТФ и ЦТФ) в реакционной смеси составляет от приблизительно 0,1 мМ до приблизительно 10 мМ, например, от приблизительно 1 мМ до приблизительно 10 мМ, от приблизительно 2 мМ до приблизительно 10 мМ, от приблизительно 3 мМ до приблизительно 10 мМ, от приблизительно 1 мМ до приблизительно 8 мМ, от приблизительно 1 мМ до приблизительно 6 мМ, от приблизительно 3 мМ до приблизительно 10 мМ, от приблизительно 3 мМ до приблизительно 8 мМ, от приблизительно 3 мМ до приблизительно 6 мМ, от приблизительно 4 мМ до приблизительно 5 мМ. В некоторых вариантах осуществления в реакционной смеси каждый рибонуклеотид содержится в концентрации приблизительно 5 мМ. В некоторых вариантах осуществления общая концентрация рибонуклеозидтрифосфатов (рНТФ) (например, АТФ, ГТФ, ЦТФ и УТФ в совокупности), используемых в реакционной смеси, находится в диапазоне от 1 мМ до 40 мМ. В некоторых вариантах осуществления общая концентрация рНТФ (например, АТФ, ГТФ, ЦТФ и УТФ в совокупности), используемых в реакционной смеси, находится в диапазоне от 1 мМ до 30 мМ, или от 1 мМ до 28 мМ, или от 1 мМ до 25 мМ, или от 1 мМ до 20 мМ. В некоторых вариантах осуществления общая концентрация рНТФ составляет менее 30 мМ. В некоторых вариантах осуществления общая концентрация рНТФ составляет

менее 25 мМ. В некоторых вариантах осуществления общая концентрация рНТФ составляет менее 20 мМ. В некоторых вариантах осуществления общая концентрация рНТФ составляет менее 15 мМ. В некоторых вариантах осуществления общая концентрация рНТФ составляет менее 10 мМ.

[0070] РНК-полимеразный реакционный буфер, как правило, содержит солевой/буферный агент, например, Tris, HEPES, сульфат аммония, бикарбонат натрия, цитрат натрия, ацетат натрия, фосфат калия, фосфат натрия и хлорид магния.

[0071] Значение рН реакционной смеси может находиться в диапазоне от приблизительно 6 до 8,5, от 6,5 до 8,0, от 7,0 до 7,5, а в некоторых вариантах осуществления значение рН составляет 7,5.

[0072] Для образования реакционной смеси комбинируют линейный или линеаризованный ДНК-шаблон (например, описанный выше, в количестве/концентрации, достаточном для получения требуемого количества РНК), РНК-полимеразный реакционный буфер и РНК-полимеразу SP6. Реакционную смесь инкубируют при температуре от приблизительно 37 °С до приблизительно 42 °С в течение периода от тридцати минут до шести часов, например, от приблизительно шестидесяти до приблизительно девяноста минут.

[0073] В некоторых вариантах осуществления приблизительно 5 мМ НТФ, приблизительно 0,05 мг/мл полимеразы SP6 и приблизительно 0,1 мг/мл ДНК-шаблона в подходящем РНК-полимеразном реакционном буфере (конечная рН реакционной смеси приблизительно 7,5) инкубируют при температуре от приблизительно 37 °С до приблизительно 42 °С в течение периода от шестидесяти до девяноста минут.

[0074] В некоторых вариантах осуществления реакционная смесь содержит линеаризованный двухцепочечный ДНК-шаблон со специфичным к полимеразе SP6 промотором, РНК-полимеразу SP6, ингибитор РНКаз, пирофосфатазу, 29 мМ НТФ, 10 мМ ДТТ и реакционный буфер (10х буфер содержит 800 мМ HEPES, 20 мМ спермидина, 250 мМ MgCl₂, рН 7,7) и достаточное количество (QS) для доведения до требуемого объема реакционной смеси не содержащей РНКаз воды; далее эту реакционную смесь инкубируют при 37 °С в течение 60 минут. Затем полимеразную реакцию гасят путем добавления ДНКазы I и буфера ДНКазы I (10х буфер содержит 100 мМ Tris-HCl, 5 мМ MgCl₂ и 25 мМ CaCl₂, рН 7,6) для облегчения расщепления

двухцепочечного ДНК-шаблона при подготовке к очистке. Было показано, что этого варианта осуществления достаточно для получения 100 граммов мРНК.

[0075] В некоторых вариантах осуществления реакционная смесь содержит НТФ в концентрации в диапазоне 1–10 мМ, ДНК-шаблон в концентрации в диапазоне 0,01–0,5 мг/мл и РНК-полимеразу SP6 в концентрации в диапазоне 0,01–0,1 мг/мл, например, реакционная смесь содержит НТФ в концентрации 5 мМ, ДНК-шаблон в концентрации 0,1 мг/мл и РНК-полимеразу SP6 в концентрации 0,05 мг/мл.

Нуклеотиды

[0076] Для получения мРНК в соответствии с настоящим изобретением можно использовать различные встречающиеся в природных условиях или модифицированные нуклеозиды. В некоторых вариантах осуществления мРНК представляет собой или содержит природные нуклеозиды (например, аденозин, гуанозин, цитидин, уридин); аналоги нуклеозидов (например, 2-аминоаденозин, 2-тиотимидин, инозин, пирролопиримидин, 3-метиладенозин, 5-метилцитидин, C-5 пропинилцитидин, C-5 пропинилуридин, 2-аминоаденозин, C5-бромурин, C5-фторуридин, C5-йодуридин, C5-пропинилуридин, C5-пропинилцитидин, C5-метилцитидин, 2-аминоаденозин, 7-дезааденозин, 7-дезагуанозин, 8-оксоаденозин, 8-оксогуанозин, O(6)-метилгуанин, псевдоуридин (например, N-1-метилпсевдоуридин), 2-тиоуридин и 2-тиоцитидин); химически модифицированные основания; биологически модифицированные основания (например, метилированные основания); интеркалированные основания; модифицированные сахара (например, 2'-фторрибоза, рибоза, 2'-дезоксирибоза, арабиноза и гексоза); и/или модифицированные фосфатные группы (например, фосфотиоаты и 5'-N-фосфорамидитные связи).

В некоторых вариантах осуществления мРНК содержит один или более нестандартных нуклеотидных остатков. К нестандартным нуклеотидным остаткам могут относиться, например, 5-метилцитидин («5mC»), псевдоуридин («ψU») и/или 2-тиоуридин («2sU»). См., например, патент США № 8,278,036 или WO2011012316, где приведено описание таких остатков и их встраивания в мРНК. мРНК может представлять собой РНК, которая определяется как РНК, в которой 25% остатков U представляют собой 2-тиоуридин, а 25% остатков C представляют собой 5-метилцитидин. Способы использования РНК описаны в патентной публикации США US20120195936 и международной публикации WO2011012316, обе из которых

полностью включены в настоящий документ путем ссылки. Наличие нестандартных нуклеотидных остатков может сделать мРНК более стабильной и/или менее иммуногенной, чем контрольная мРНК с той же последовательностью, но содержащая только стандартные остатки. В дополнительных вариантах осуществления мРНК может содержать один или более нестандартных нуклеотидных остатков, выбранных из изоцитозина, псевдоизоцитозина, 5-бром урацила, 5-пропинилурацила, 6-аминопурина, 2-аминопурина, инозина, диаминопурина и 2-хлор-6-аминопурицидозина, а также комбинации этих модификаций с другими модификациями нуклеотидных оснований. Некоторые варианты осуществления могут дополнительно включать в себя дополнительные модификации в фуранозном кольце или нуклеотидном основании. Дополнительные модификации могут включать в себя, например, модификации или замены сахара (например, одну или более из 2'-О-алкильной модификации и «закрытой» нуклеиновой кислоты (LNA)). В некоторых вариантах осуществления РНК может образовывать комплекс или гибридизоваться с дополнительными полинуклеотидами и/или пептидными полинуклеотидами (PNA). В некоторых вариантах осуществления, в которых модификация сахара представляет собой алкильную модификацию 2'-О-алкильную модификацию, такая модификация может включать в себя, без ограничений, модификацию 2'-дезоксигруппой, модификацию 2'-О-метильной группой, модификацию 2'-О-метоксиэтильной группой и модификацию 2'-дезоксигруппой. В некоторых вариантах осуществления любая из этих модификаций может присутствовать в 0–100% нуклеотидов, например, более чем в 0%, 1%, 10%, 25%, 50%, 75%, 85%, 90%, 95% или 100% составляющих нуклеотидов по отдельности или в комбинации. *Обработка после синтеза*

[0077] Как правило, после синтеза можно добавлять кэп на 5'-конце и/или хвост на 3'-конце. Наличие кэпа важно для обеспечения устойчивости к нуклеазам, присутствующим в большинстве эукариотических клеток. Наличие «хвоста» служит для защиты мРНК от расщепления экзонуклеазами.

[0078] 5'-кэп обычно добавляют следующим образом: сначала РНК-терминальная фосфатаза удаляет одну из концевых фосфатных групп из 5'-нуклеотида, оставляя два концевых фосфата; затем к концевым фосфатам добавляют гуанозинтрифосфат (ГТФ) посредством гуанилилтрансферазы, получая 5'5'-трифосфатную связь; и 7-азот гуанина метилируют метилтрансферазой. К примерам

структур кэпов относятся, без ограничений, m⁷G(5')ppp (5'(A,G(5')ppp(5')A и G(5')ppp(5')G. Дополнительные структуры кэпов описаны в опубликованной заявке на патент США № US 2016/0032356 и предварительной заявке на патент США № 62/464,327, поданной 27 февраля 2017 г., которые включены в настоящий документ путем ссылки.

[0079] Как правило, структура хвоста включает в себя поли(A) и/или поли(C). Хвост поли-A или поли-C на 3'-конце мРНК, как правило, содержит по меньшей мере 50 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 150 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 200 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 250 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 300 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 350 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 400 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 450 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 500 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 550 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 600 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 650 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 700 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 750 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 800 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 850 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 900 аденозиновых или цитозиновых нуклеотидов, по меньшей мере 950 аденозиновых или цитозиновых нуклеотидов или по меньшей мере 1 тыс. аденозиновых или цитозиновых нуклеотидов соответственно. В некоторых вариантах осуществления хвост поли(A) или поли(C) может содержать приблизительно 10–800 аденозиновых или цитозиновых нуклеотидов (например, приблизительно 10–200 аденозиновых или цитозиновых нуклеотидов, приблизительно 10–300 аденозиновых или цитозиновых нуклеотидов, приблизительно 10–400 аденозиновых или цитозиновых нуклеотидов, приблизительно 10–500 аденозиновых или цитозиновых нуклеотидов, приблизительно 10–550 аденозиновых или цитозиновых нуклеотидов, приблизительно 10–600 аденозиновых или цитозиновых нуклеотидов, приблизительно 50–600 аденозиновых или цитозиновых нуклеотидов, приблизительно 100–600 аденозиновых или цитозиновых нуклеотидов, приблизительно 150–600 аденозиновых или цитозиновых нуклеотидов, приблизительно 200–

600 аденозиновых или цитозиновых нуклеотидов, приблизительно 250–
600 аденозиновых или цитозиновых нуклеотидов, приблизительно 300–
600 аденозиновых или цитозиновых нуклеотидов, приблизительно 350–
600 аденозиновых или цитозиновых нуклеотидов, приблизительно 400–
600 аденозиновых или цитозиновых нуклеотидов, приблизительно 450–
600 аденозиновых или цитозиновых нуклеотидов, приблизительно 500–
600 аденозиновых или цитозиновых нуклеотидов, приблизительно 10–
150 аденозиновых или цитозиновых нуклеотидов, приблизительно 10–
100 аденозиновых или цитозиновых нуклеотидов, приблизительно 20–70 аденозиновых
или цитозиновых нуклеотидов или приблизительно 20–60 аденозиновых или
цитозиновых нуклеотидов) соответственно. В некоторых вариантах осуществления
структура хвоста включает в себя комбинацию хвостов поли (А) и поли (С) с
различными длинами, описанными в настоящем документе. В некоторых вариантах
осуществления структура хвоста включает в себя по меньшей мере 50%, 55%, 65%,
70%, 75%, 80%, 85%, 90%, 92%, 94%, 95%, 96%, 97%, 98% или 99% аденозиновых
нуклеотидов. В некоторых вариантах осуществления структура хвоста включает в себя
по меньшей мере 50%, 55%, 65%, 70%, 75%, 80%, 85%, 90%, 92%, 94%, 95%, 96%,
97%, 98% или 99% цитозиновых нуклеотидов.

[0080] Как описано в настоящем документе, добавление 5'-кэпа и/или 3'-хвоста
облегчает обнаружение незавершенных транскриптов, образованных во время синтеза
in vitro, поскольку без добавления кэпа и/или хвоста размер этих преждевременно
прерванных транскриптов мРНК может быть слишком мал для обнаружения. Таким
образом, в некоторых вариантах осуществления 5'-кэп и/или 3'-хвост добавляют к
синтезированной мРНК до проверки мРНК на чистоту (например, на уровень
незавершенных транскриптов, присутствующих в мРНК). В некоторых вариантах
осуществления 5'-кэп и/или 3'-хвост добавляют к синтезированной мРНК перед
очисткой мРНК, описанной в настоящем документе. В других вариантах
осуществления 5'-кэп и/или 3'-хвост добавляют к синтезированной мРНК после
очистки мРНК, описанной в настоящем документе.

[0081] мРНК, синтезированную в соответствии с настоящим изобретением,
можно использовать без дополнительной очистки. В частности, мРНК,
синтезированную в соответствии с настоящим изобретением, можно использовать без

стадии удаления коротких полимеров. В некоторых вариантах осуществления мРНК, синтезированную в соответствии с настоящим изобретением, можно дополнительно очищать. Для очистки мРНК, синтезированной в соответствии с настоящим изобретением, можно использовать различные способы. Например, очистку мРНК можно проводить с использованием способов центрифугирования, фильтрации и/или хроматографии. В некоторых вариантах осуществления синтезированную мРНК очищают посредством осаждения этанолом, или фильтрования, или хроматографии, или очистки в геле, или любыми другими подходящими способами. В некоторых вариантах осуществления мРНК очищают методом ВЭЖХ. В некоторых вариантах осуществления мРНК экстрагируют в стандартный раствор фенол : хлороформ : изоамиловый спирт, хорошо известный специалисту в данной области. В некоторых вариантах осуществления мРНК очищают с помощью тангенциальной поточной фильтрации. Подходящие способы очистки включают в себя описанные в публикациях US 2016/0040154, US 2015/0376220, заявке на договор о патентной кооперации PCT/US18/19954, озаглавленной METHODS FOR PURIFICATION OF MESSENGER RNA, поданной 27 февраля 2018 г., и заявке на договор о патентной кооперации PCT/US18/19978, озаглавленной METHODS FOR PURIFICATION OF MESSENGER RNA, поданной 27 февраля 2018 г., все из которых включены в настоящий документ путем ссылки и могут использоваться при практической реализации настоящего изобретения.

[0082] В некоторых вариантах осуществления мРНК очищают перед добавлением кэпа и хвоста. В некоторых вариантах осуществления мРНК очищают после добавления кэпа и хвоста. В некоторых вариантах осуществления мРНК очищают как до, так и после добавления кэпа и хвоста.

[0083] В некоторых вариантах осуществления мРНК очищают до, после или как до, так и после добавления кэпа и хвоста путем центрифугирования.

[0084] В некоторых вариантах осуществления мРНК очищают до, после или как до, так и после добавления кэпа и хвоста путем фильтрации.

[0085] В некоторых вариантах осуществления мРНК очищают до, после или как до, так и после добавления кэпа и хвоста путем тангенциальной проточной фильтрации (TFF).

[0086] В некоторых вариантах осуществления мРНК очищают до, после или как до, так и после добавления кэпа и хвоста путем хроматографии.

Характеризация мРНК

[0087] Полноразмерные или незавершенные транскрипты мРНК можно обнаруживать и количественно оценивать с помощью любых способов, доступных в данной области. В некоторых вариантах осуществления синтезированные молекулы мРНК обнаруживают с использованием блоттинга, капиллярного электрофореза, хроматографии, флуоресценции, гель-электрофореза, ВЭЖХ, окрашивания серебром, спектроскопии, ультрафиолета (УФ) или СВЭЖХ либо их комбинации. В настоящее изобретение включены другие способы обнаружения, известные в данной области. В некоторых вариантах осуществления синтезированные молекулы мРНК обнаруживают с использованием спектроскопии поглощения УФ-света с разделением капиллярным электрофорезом. В некоторых вариантах осуществления мРНК сначала денатурируют красителем глиоксалем, а затем проводят электрофорез в геле («глиоксальный гель-электрофорез»). В некоторых вариантах осуществления синтезированную мРНК характеризуют до добавления кэпа или хвоста. В некоторых вариантах осуществления синтезированную мРНК характеризуют после добавления кэпа и хвоста.

[0088] В некоторых вариантах осуществления мРНК, полученная способом, описанным в настоящем документе, содержит менее 10%, менее 9%, менее 8%, менее 7%, менее 6%, менее 5%, менее 4%, менее 3%, менее 2%, менее 1%, менее 0,5%, менее 0,1% примесей, отличных от полноразмерной мРНК. К примесям относятся IVT-загрязнители, например, белки, ферменты, свободные нуклеотиды и/или короткие полимеры.

[0089] В некоторых вариантах осуществления мРНК, полученная в соответствии с изобретением, по существу не содержит коротких полимеров или незавершенных транскриптов. В частности, мРНК, полученная в соответствии с изобретением, содержит необнаружимый уровень коротких полимеров или незавершенных транскриптов при определении капиллярным электрофорезом или глиоксальным гель-электрофорезом. В настоящем документе термин «короткие полимеры» или «незавершенные транскрипты» относится к любым транскриптам, которые меньше полноразмерных. В некоторых вариантах осуществления «короткие полимеры» или

«незавершенные транскрипты» имеют длину менее 100 нуклеотидов, менее 90, менее 80, менее 70, менее 60, менее 50, менее 40, менее 30, менее 20 или менее 10 нуклеотидов. В некоторых вариантах осуществления короткие полимеры обнаруживают или количественно определяют после добавления 5'-кэпа и/или 3'-полиА-хвоста.

Раствор мРНК

[0090] В некоторых вариантах осуществления мРНК может быть предоставлена в растворе, который смешивают с липидным раствором таким образом, чтобы мРНК можно было инкапсулировать в липидные наночастицы. Подходящим раствором мРНК может быть любой водный раствор, содержащий мРНК для инкапсулирования, с различными концентрациями. Например, подходящий раствор мРНК может содержать мРНК в концентрации, равной или составляющей более чем приблизительно 0,01 мг/мл, 0,05 мг/мл, 0,06 мг/мл, 0,07 мг/мл, 0,08 мг/мл, 0,09 мг/мл, 0,1 мг/мл, 0,15 мг/мл, 0,2 мг/мл, 0,3 мг/мл, 0,4 мг/мл, 0,5 мг/мл, 0,6 мг/мл, 0,7 мг/мл, 0,8 мг/мл, 0,9 мг/мл или 1,0 мг/мл. В некоторых вариантах осуществления подходящий раствор мРНК может содержать мРНК в концентрации в диапазоне от приблизительно 0,01–1,0 мг/мл, 0,01–0,9 мг/мл, 0,01–0,8 мг/мл, 0,01–0,7 мг/мл, 0,01–0,6 мг/мл, 0,01–0,5 мг/мл, 0,01–0,4 мг/мл, 0,01–0,3 мг/мл, 0,01–0,2 мг/мл, 0,01–0,1 мг/мл, 0,05–1,0 мг/мл, 0,05–0,9 мг/мл, 0,05–0,8 мг/мл, 0,05–0,7 мг/мл, 0,05–0,6 мг/мл, 0,05–0,5 мг/мл, 0,05–0,4 мг/мл, 0,05–0,3 мг/мл, 0,05–0,2 мг/мл, 0,05–0,1 мг/мл, 0,1–1,0 мг/мл, 0,2–0,9 мг/мл, 0,3–0,8 мг/мл, 0,4–0,7 мг/мл или 0,5–0,6 мг/мл. В некоторых вариантах осуществления подходящий раствор мРНК может содержать мРНК в концентрации до приблизительно 5,0 мг/мл, 4,0 мг/мл, 3,0 мг/мл, 2,0 мг/мл, 1,0 мг/мл, 0,09 мг/мл, 0,08 мг/мл, 0,07 мг/мл, 0,06 мг/мл или 0,05 мг/мл.

[0091] Как правило, подходящий раствор мРНК может также содержать буферный агент и/или соль. Как правило, буферные агенты могут включать в себя HEPES, сульфат аммония, бикарбонат натрия, цитрат натрия, ацетат натрия, фосфат калия и фосфат натрия. В некоторых вариантах осуществления подходящая концентрация буферного агента может находиться в диапазоне от приблизительно 0,1 мМ до 100 мМ, от 0,5 мМ до 90 мМ, от 1,0 мМ до 80 мМ, от 2 мМ до 70 мМ, от 3 мМ до 60 мМ, от 4 мМ до 50 мМ, от 5 мМ до 40 мМ или от 6 мМ до 30 мМ, от 7 мМ до 20 мМ, от 8 мМ до 15 мМ или от 9 мМ до 12 мМ. В некоторых вариантах

осуществления подходящая концентрация буферного агента равна или составляет более чем приблизительно 0,1 мМ, 0,5 мМ, 1 мМ, 2 мМ, 4 мМ, 6 мМ, 8 мМ, 10 мМ, 15 мМ, 20 мМ, 25 мМ, 30 мМ, 35 мМ, 40 мМ, 45 мМ или 50 мМ.

[0092] Примеры солей могут включать в себя хлорид натрия, хлорид магния и хлорид калия. В некоторых вариантах осуществления подходящая концентрация солей в растворе мРНК может находиться в диапазоне от приблизительно 1 мМ до 500 мМ, от 5 мМ до 400 мМ, от 10 мМ до 350 мМ, от 15 мМ до 300 мМ, от 20 мМ до 250 мМ, от 30 мМ до 200 мМ, от 40 мМ до 190 мМ, от 50 мМ до 180 мМ, от 50 мМ до 170 мМ, от 50 мМ до 160 мМ, от 50 мМ до 150 мМ или от 50 мМ до 100 мМ. Концентрация соли в подходящем растворе мРНК равна или составляет более чем приблизительно 1 мМ, 5 мМ, 10 мМ, 20 мМ, 30 мМ, 40 мМ, 50 мМ, 60 мМ, 70 мМ, 80 мМ, 90 мМ или 100 мМ.

[0093] В некоторых вариантах осуществления подходящий раствор мРНК может иметь рН в диапазоне от приблизительно 3,5–6,5, 3,5–6,0, 3,5–5,5, 3,5–5,0, 3,5–4,5, 4,0–5,5, 4,0–5,0, 4,0–4,9, 4,0–4,8, 4,0–4,7, 4,0–4,6 или 4,0–4,5. В некоторых вариантах осуществления подходящий раствор мРНК может иметь значение рН, равное или составляющее не более чем приблизительно 3,5, 4,0, 4,1, 4,2, 4,3, 4,4, 4,5, 4,6, 4,7, 4,8, 4,9, 5,0, 5,2, 5,4, 5,6, 5,8, 6,0, 6,1, 6,3 и 6,5.

[0094] Для получения раствора мРНК, подходящего для настоящего изобретения, можно использовать различные способы. В некоторых вариантах осуществления мРНК можно непосредственно растворять в буферном растворе, описанном в настоящем документе. В некоторых вариантах осуществления раствор мРНК может быть получен посредством смешивания маточного раствора мРНК с буферным раствором перед смешиванием с липидным раствором для инкапсулирования. В некоторых вариантах осуществления раствор мРНК может быть получен путем смешивания маточного раствора мРНК с буферным раствором непосредственно перед смешиванием с липидным раствором для инкапсулирования. В некоторых вариантах осуществления подходящий маточный раствор мРНК может содержать мРНК в воде в концентрации, равной или составляющей более чем приблизительно 0,2 мг/мл, 0,4 мг/мл, 0,5 мг/мл, 0,6 мг/мл, 0,8 мг/мл, 1,0 мг/мл, 1,2 мг/мл, 1,4 мг/мл, 1,5 мг/мл, 1,6 мг/мл, 2,0 мг/мл, 2,5 мг/мл, 3,0 мг/мл, 3,5 мг/мл, 4,0 мг/мл, 4,5 мг/мл или 5,0 мг/мл.

[0095] В некоторых вариантах осуществления маточный раствор мРНК смешивают с буферным раствором с использованием насоса. К примерам насосов относятся, без ограничений, шестеренчатые насосы, перистальтические насосы и центробежные насосы.

[0096] Как правило, буферный раствор перемешивают с большей скоростью, чем маточный раствор мРНК. Например, буферный раствор можно перемешивать со скоростью, которая по меньшей мере в 1х, 2х, 3х, 4х, 5х, 6х, 7х, 8х, 9х, 10х, 15х или 20х больше скорости перемешивания маточного раствора мРНК. В некоторых вариантах осуществления буферный раствор перемешивают при скорости потока в диапазоне приблизительно 100–6000 мл/мин (например, приблизительно 100–300 мл/мин, 300–600 мл/мин, 600–1200 мл/мин, 1200–2400 мл/мин, 2400–3600 мл/мин, 3600–4800 мл/мин, 4800–6000 мл/мин или 60–420 мл/мин). В некоторых вариантах осуществления буферный раствор перемешивают со скоростью потока более чем приблизительно 60 мл/мин, 100 мл/мин, 140 мл/мин, 180 мл/мин, 220 мл/мин, 260 мл/мин, 300 мл/мин, 340 мл/мин, 380 мл/мин, 420 мл/мин, 480 мл/мин, 540 мл/мин, 600 мл/мин, 1200 мл/мин, 2400 мл/мин, 3600 мл/мин, 4800 мл/мин или 6000 мл/мин.

[0097] В некоторых вариантах осуществления маточный раствор мРНК перемешивают при скорости потока в диапазоне приблизительно 10–600 мл/мин (например, приблизительно 5–50 мл/мин, приблизительно 10–30 мл/мин, приблизительно 30–60 мл/мин, приблизительно 60–120 мл/мин, приблизительно 120–240 мл/мин, приблизительно 240–360 мл/мин, приблизительно 360–480 мл/мин или приблизительно 480–600 мл/мин). В некоторых вариантах осуществления маточный раствор мРНК перемешивают при скорости потока, равной или составляющей более чем приблизительно 5 мл/мин, 10 мл/мин, 15 мл/мин, 20 мл/мин, 25 мл/мин, 30 мл/мин, 35 мл/мин, 40 мл/мин, 45 мл/мин, 50 мл/мин, 60 мл/мин, 80 мл/мин, 100 мл/мин, 200 мл/мин, 300 мл/мин, 400 мл/мин, 500 мл/мин или 600 мл/мин.

Доставляющие носители

[0098] В соответствии с настоящим изобретением мРНК, кодирующую белок CFTR (например, полноразмерный белок, фрагмент или часть белка CFTR), описанную в настоящем документе, можно доставлять в виде свободной мРНК (неупакованной) или с помощью доставляющих носителей. В настоящем документе термины

«доставляющий носитель», «транспортный носитель», «наночастица» или их грамматические эквиваленты используются взаимозаменяемо.

[0099] Доставляющие носители можно составлять в комбинации с одной или более дополнительными нуклеиновыми кислотами, носителями, нацеливающими лигандами или стабилизирующими реагентами либо использовать в фармакологических композициях, где их смешивают с подходящими эксципиентами. Методики составления и введения лекарственных средств можно найти в публикации Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa., последнее издание. Конкретный доставляющий носитель выбирают на основании его способности облегчать трансфекцию нуклеиновой кислоты в клетку-мишень.

[0100] В некоторых вариантах осуществления доставляющий носитель, содержащий мРНК CFTR, вводят путем доставки через легкие, например, путем распыления. В этих вариантах осуществления доставляющий носитель может находиться в аэрозольной композиции, которую можно вдыхать. В некоторых вариантах осуществления мРНК экспрессируется в ткани, в которую введен доставляющий носитель, например, в носовой полости, трахее, бронхах, бронхиолах и/или других клетках или тканях, связанных с легочной системой. Дополнительное описание доставки через легкие и распыления приведено в родственной международной заявке PCT/US17/61100, поданной 10 ноября 2017 г. автором настоящей заявки и озаглавленной NOVEL ICE-BASED LIPID NANOPARTICLE FORMULATION FOR DELIVERY OF MRNA, и в предварительной заявке на патент США USSN 62/507,061, каждая из которых полностью включена в настоящий документ путем ссылки.

[0101] В некоторых вариантах осуществления мРНК, кодирующие белок CFTR, могут доставляться посредством одного доставляющего носителя. В некоторых вариантах осуществления мРНК, кодирующие белок CFTR, могут доставляться с помощью одного или более доставляющих носителей, каждый из которых имеет разный состав. В соответствии с различными вариантами осуществления к подходящим доставляющим носителям относятся, без ограничений, носители на основе полимера, такого как полиэтиленимин (PEI), липидные наночастицы и липосомы, нанолипосомы, церамид-содержащие нанолипосомы, протеолипосомы, натуральные и синтетические экзосомы, натуральные, синтетические и

полусинтетические ламеллярные тела, твердые наночастицы, твердые наночастицы фосфосиликата кальция, твердые наночастицы фосфата кальция, твердые наночастицы диоксида кремния, нанокристаллические твердые частицы, полупроводниковые наночастицы, поли(D-аргинин), золь-гели, нанодендримеры, доставляющие системы на основе крахмала, мицеллы, эмульсии, ниосомы, мультидоменные блок-полимеры (виниловые полимеры, полимеры полипропилакриловой кислоты, динамические поликонъюгаты), сухие порошковые составы, плазмиды, вирусы, кальцийфосфатнуклеотиды, аптамеры, пептиды и другие векторные метки. Также предусмотрено использование бионанокансул и других сборок вирусных капсидных белков в качестве подходящего транспортного носителя (Hum. Gene Ther. 2008 September; 19(9):887–95).

[0102] Доставляющий носитель, содержащий мРНК CFTR, можно вводить и дозировать в соответствии с существующей медицинской практикой с учетом клинического состояния субъекта, места и способа введения (например, местного и системного, включая пероральное, легочное и посредством инъекции), схемы введения, возраста субъекта, пола, массы тела и других факторов, значимых для медицинских специалистов в данной области. «Эффективное количество» для целей настоящего изобретения может быть определено с учетом таких значимых соображений, которые известны обычным специалистам в экспериментальных клинических исследованиях, фармакологии, клинике и медицине. В некоторых вариантах осуществления введенное количество является эффективным для достижения по меньшей мере некоторой стабилизации, улучшения или устранения симптомов и других показателей, выбранных специалистами в данной области в качестве подходящих мер прогрессирования, регрессирования или улучшения течения заболевания. Например, подходящим количеством и схемой введения является такое количество, которое вызывает по меньшей мере временную продукцию белка.

[0103] В некоторых вариантах осуществления доставляющие носители состоят таким образом, что они подходят для длительного высвобождения содержащейся в них мРНК. Такие композиции с длительным высвобождением могут быть удобно вводить субъекту с продолжительными интервалами дозирования.

Липосомные доставляющие носители

[0104] В некоторых вариантах осуществления подходящим доставляющим носителем является липосомный доставляющий носитель, например, липидная наночастица. В настоящем документе липосомные доставляющие носители, например, наночастицы липидов, как правило, характеризуются микроскопическими везикулами, имеющими внутреннее водное пространство, изолированное от внешней среды мембраной из одного или более двойных слоев. Двуслойные мембраны липосом, как правило, образованы амфифильными молекулами, такими как липиды синтетического или природного происхождения, которые содержат пространственно разделенные гидрофильные и гидрофобные домены (Lasic, Trends Biotechnol., 16: 307–321, 1998). Двухслойные мембраны липосом также могут быть образованы амфифильными полимерами и поверхностно-активными веществами (например, полимеросомы, ниосомы и т. д.). В контексте настоящего изобретения липосомный доставляющий носитель, как правило, служит для транспортировки требуемой мРНК в клетку или ткань-мишень. В некоторых вариантах осуществления доставляющий носитель в виде наночастиц представляет собой липосому. В некоторых вариантах осуществления липосома содержит один или более катионных липидов, один или более некаатионных липидов, один или более липидов на основе холестерина и один или более ПЭГ-модифицированных липидов. В некоторых вариантах осуществления липосома содержит не более трех отдельных липидных компонентов. В некоторых вариантах осуществления один отдельный липидный компонент представляет собой катионный липид на основе стерола.

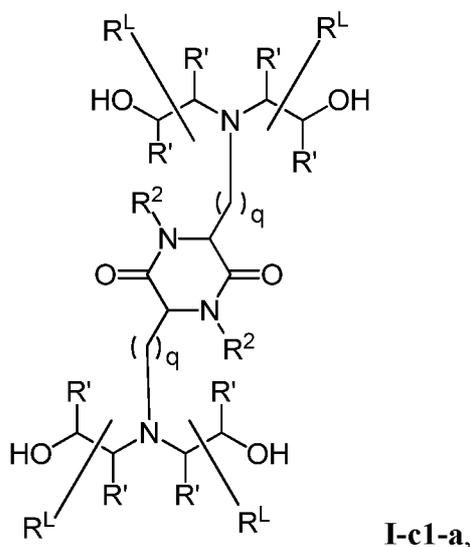
Катионные липиды

[0105] В некоторых вариантах осуществления липосома могут содержать один или более катионных липидов. В настоящем документе фраза «катионный липид» относится к любому из множества липидных фрагментов, которые имеют чистый положительный заряд при выбранном рН, таком как физиологический рН. В литературе описаны несколько катионных липидов, многие из которых доступны в продаже. К примерам подходящих катионных липидов для использования в композициях и способах по изобретению относятся описанные в международных патентных публикациях WO 2010/053572 (например, С1 2-200, описанный в пункте [00225]) и WO 2012/170930, обе из которых включены в настоящий документ

путем ссылки. В некоторых вариантах осуществления в композициях и способах по изобретению используются липидные наночастицы, содержащие ионизируемый катионный липид, описанный в предварительной заявке на патент США № 61/617,468, поданной 29 марта 2012 г. (включена в настоящий документ путем ссылки), такие как, например, (15Z, 18Z)-N,N-диметил-6-(9Z, 12Z)-октадека-9, 12-диен-1-ил)тетракоза-15,18-диен-1-амин (HGT5000), (15Z, 18Z)-N,N-диметил-6-((9Z, 12Z)-октадека-9, 12-диен-1-ил)тетракоза-4,15,18-триен-1-амин (HGT5001) и (15Z,18Z)-N,N-диметил-6-((9Z, 12Z)-октадека-9, 12-диен-1-ил)тетракоза-5, 15, 18-триен-1-амин (HGT5002).

[0106] В некоторых вариантах осуществления предложенные липосомы включают в себя катионный липид, описанный в публикации WO 2013/063468 и в предварительной заявке на патент США, озаглавленной Lipid Formulations for Delivery of Messenger RNA, поданной одновременно с настоящей заявкой и в ту же дату, обе из которых включены в настоящий документ путем ссылки.

[0107] В некоторых вариантах осуществления катионный липид содержит соединение формулы **I-c1-a**:



или его фармацевтически приемлемую соль, где:

каждый R^2 независимо представляет собой водород или C_{1-3} алкил;

каждый q независимо равен 2–6;

каждый R' независимо представляет собой водород или C_{1-3} алкил

и каждый R^L независимо представляет собой C_{8-12} алкил.

[0108] В некоторых вариантах осуществления каждый R^2 независимо представляет собой водород, метил или этил. В некоторых вариантах осуществления каждый R^2 независимо представляет собой водород или метил. В некоторых вариантах осуществления каждый R^2 представляет собой водород.

[0109] В некоторых вариантах осуществления каждый q независимо равен 3–6. В некоторых вариантах осуществления каждый q независимо равен 3–5. В некоторых вариантах осуществления каждый q равен 4.

[0110] В некоторых вариантах осуществления каждый R^1 независимо представляет собой водород, метил или этил. В некоторых вариантах осуществления каждый R^1 независимо представляет собой водород или метил. В некоторых вариантах осуществления каждый R^1 независимо представляет собой водород.

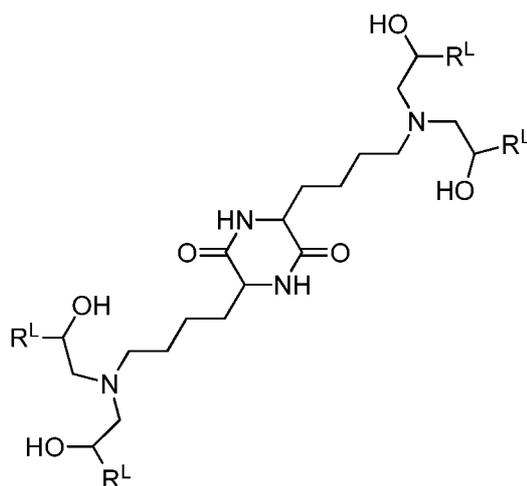
[0111] В некоторых вариантах осуществления каждый R^L независимо представляет собой C_{8-12} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой n - C_{8-12} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой C_{9-11} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой n - C_{9-11} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой C_{10} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой n - C_{10} алкил.

[0112] В некоторых вариантах осуществления каждый R^2 независимо представляет собой водород или метил; каждый q независимо равен 3–5; каждый R^1 независимо представляет собой водород или метил и каждый R^L независимо представляет собой C_{8-12} алкил.

[0113] В некоторых вариантах осуществления каждый R^2 представляет собой водород; каждый q независимо равен 3–5 и каждый R^L независимо представляет собой C_{8-12} алкил.

[0114] В некоторых вариантах осуществления каждый R^2 представляет собой водород; каждый q равен 4; каждый R^1 представляет собой водород; а каждый R^L независимо представляет собой C_{8-12} алкил.

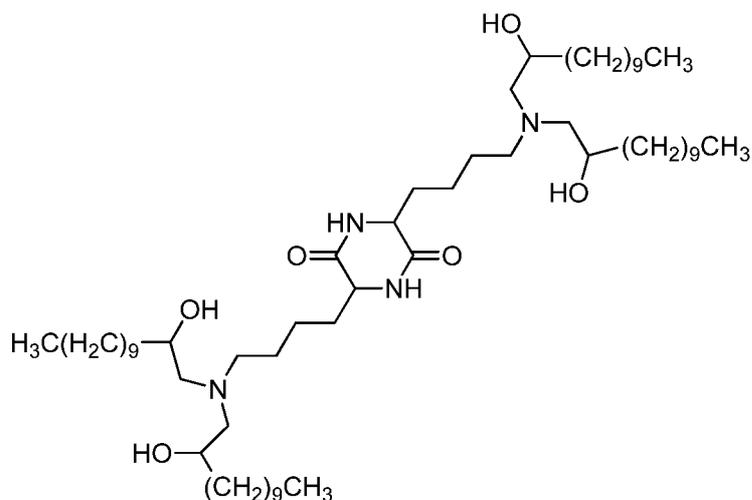
[0115] В некоторых вариантах осуществления катионный липид содержит соединение формулы **I-g**:



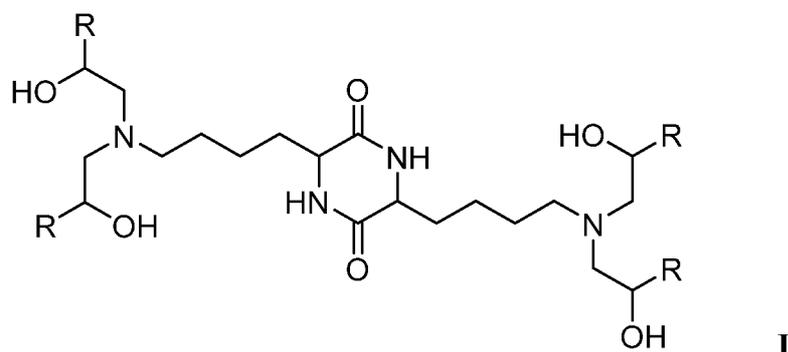
I-g,

или его фармацевтически приемлемую соль, где каждый R^L независимо представляет собой C_{8-12} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой *n*- C_{8-12} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой C_{9-11} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой *n*- C_{9-11} алкил. В некоторых вариантах осуществления каждый R^L независимо представляет собой C_{10} алкил. В некоторых вариантах осуществления каждый R^L представляет собой *n*- C_{10} алкил.

[0116] В конкретных вариантах осуществления предложенные липосомы включают в себя катионный липид сКК-Е12 или (3,6-бис(4-(бис(2-гидроксидодецил)амино)бутил)пиперазин-2,5-дион). Структура сКК-Е12 показана ниже:



[0117] Дополнительные примеры катионных липидов включают в себя соединения формулы I:



и их фармацевтически приемлемые соли,

где

R представляет собой («OF-00»),

R представляет собой («OF-01»),

R представляет собой («OF-02») или

R представляет собой («OF-03»)

(см., например, Fenton, Owen S., et al. Bioinspired Alkenyl Amino Alcohol Ionizable Lipid Materials for Highly Potent In Vivo mRNA Delivery. *Advanced materials* (2016)).

[0118] В некоторых вариантах осуществления один или более катионных липидов могут представлять собой N-[1-(2,3-диолеилокси)пропил]-N,N,N-триметиламмония хлорид, или «DOTMA» (Feigner et al. (Proc. Nat'l Acad. Sci. 84, 7413 (1987); патент США № 4,897,355). DOTMA может быть составлен отдельно или может быть скомбинирован с нейтральным липидом, диолеилфосфатидилэтаноламином, или «DOPE», либо другими катионными или некаатионными липидами для получения липосомного транспортного носителя или липидной наночастицы, и такие липосомы могут использоваться для улучшения доставки нуклеиновых кислот в клетки-мишени. К другим подходящим катионным липидам относятся, например, 5-карбоксиспермилглициндиоктадециламид, или «DOGS», 2,3-диолеилокси-N-[2(сперминкарбоксамидо)этил]-N,N-диметил-1-пропанаминий, или «DOSPA» (Behr et al. Proc. Nat'l Acad. Sci. 86, 6982 (1989); патент США № 5,171,678; патент США

№ 5,334,761), 1,2-диолеоил-3-диметиламмонийпропан, или «DODAP», 1,2-диолеоил-3-триметиламмонийпропан, или «DOTAP».

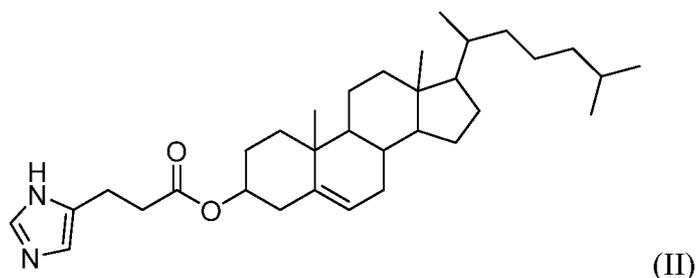
[0119] К дополнительным примерам катионных липидов также относятся 1,2-дистеарилокси-N,N-диметил-3-аминопропан, или «DSDMA», 1,2-диолеилокси-N,N-диметил-3-аминопропан, или «DODMA», 1,2-дилинолеилокси-N,N-диметил-3-аминопропан, или «DLinDMA», 1,2-дилиноленилокси-N,N-диметил-3-аминопропан, или «DLenDMA», N-диолеил-N,N-диметиламмония хлорид, или «DODAC», N,N-дистеарил-N,N-диметиламмония бромид, или «DDAB», N-(1,2-димиристилоксипроп-3-ил)-N,N-диметил-N-гидроксиэтиниламмония бромид, или «DMRIE», 3-диметиламино-2-(холест-5-ен-3-бетаоксибутан-4-окси)-1-(цис,цис-9,12-октадекадиеноксипропан, или «CLinDMA», 2-[5'-(холест-5-ен-3-бетаокси)-3'-оксапентокси]-3-диметил-1-(цис,цис-9',1-2'-октадекадиеноксипропан, или «CpLinDMA», N,N-диметил-3,4-диолеилоксибензиламин, или «DMOVA», 1,2-N,N'-диолеилкарбамил-3-диметиламинопропан, или «DOcarbDAP», 2,3-дилинолеилокси-N,N-диметилпропиламин, или «DLinDAP», 1,2-N,N'-дилинолеилкарбамил-3-диметиламинопропан, или «DLincarbDAP», 1,2-дилинолеоилкарбамил-3-диметиламинопропан, или «DLinCDAP», 2,2-дилинолеил-4-диметиламиноэтил-[1,3]-диоксолан, или «DLin- -DMA», 2,2-дилинолеил-4-диметиламиноэтил-[1,3]-диоксолан, или «DLin-K-ХТС2-DMA», и 2-(2,2-ди((9Z,12Z)-октадека-9,12-диен-1-ил)-1,3-диоксолан-4-ил)-N,N-диметилэтанамин (DLin-KC2-DMA)) (см. WO 2010/042877; Semple et al., Nature Biotech. 28: 172–176 (2010)) или их смеси. (Heyes, J., et al., J Controlled Release 107: 276–287 (2005); Morrissey, DV., et al., Nat. Biotechnol. 23(8): 1003–1007 (2005); публикация PCT WO2005/121348A1). В некоторых вариантах осуществления один или более катионных липидов содержат по меньшей мере одну из имидазольной, диалкиламино или гуанидиновой групп.

[0120] В некоторых вариантах осуществления один или более катионных липидов могут быть выбраны из ХТС (2,2-дилинолеил-4-диметиламиноэтил-[1,3]-диоксолан), МСЗ (((6Z,9Z,28Z,31Z)-гептатриаконта-6,9,28,31-тетраен-19-ил 4-(диметиламино)бутаноат), ALNY-100 ((3aR,5s,6aS)-N,N-диметил-2,2-ди((9Z,12Z)-октадека-9,12-диенил)тетрагидро-3aH-циклопента[d] [1,3]диоксол-5-амин)), NC98-5 (4,7,13-трис(3-оксо-3-(ундециламино)пропил)-N1,N16-диундецил-4,7,10,13-тетраазагексадекан-1,16-диамид), DODAP (1,2-диолеил-3-диметиламмонийпропан),

HGT4003 (публикация WO 2012/170889, содержимое которой полностью включено в настоящий документ путем ссылки), ICE (публикация WO 2011/068810, содержимое которой полностью включено в настоящий документ путем ссылки), HGT5000 (предварительная заявка на патент США № 61/617,468, содержимое которой полностью включено в настоящий документ путем ссылки) или HGT5001 (цис- или транс-) (предварительная заявка на патент № 61/617,468), аминспиртовые липидоиды, такие как описанные в публикации WO2010/053572, DOTAP (1,2-диолеил-3-триметиламмонийпропан), DOTMA (1,2-ди-О-октадецил-3-триметиламмонийпропан), DLinDMA (Heyes, J.; Palmer, L.; Bremner, K.; MacLachlan, I. Cationic lipid saturation influences intracellular delivery of encapsulated nucleic acids J. Contr. Rel. 2005, 107, 276–287), DLin-KC2-DMA (Semple, S.C. et al. “Rational Design of Cationic Lipids for siRNA Delivery” Nature Biotech. 2010, 28, 172–176), C12-200 (Love, K.T. et al. Lipid-like materials for low-dose in vivo gene silencing PNAS 2010, 107, 1864–1869).

Стероловые катионные липиды

[0121] В некоторых вариантах осуществления катионные липиды на основе стерола представляют собой диалкиламино-, имидазол- и гуанидин-содержащие катионные липиды на основе стерола. Например, некоторые варианты осуществления относятся к композиции, содержащей один или более катионных липидов на основе стерола, содержащих имидазол, например сложный эфир имидазола и холестерина «ICE» (3S, 10R, 13R, 17R)-10, 13-диметил-17-((R)-6-метилгептан-2-ил)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-тетрадекагидро-1H-циклопента[а]фенантрен-3-ил 3-(1H-имидазол-4-ил)пропаноат, представленный структурой (II) ниже. В некоторых вариантах осуществления липидная наночастица для доставки РНК (например, мРНК), кодирующей функциональный белок, может содержать один или более катионных липидов на основе имидазола, например, сложный эфир имидазола и холестерина «ICE» (3S, 10R, 13R, 17R)-10, 13-диметил-17-((R)-6-метилгептан-2-ил)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-тетрадекагидро-1H-циклопента[а]фенантрен-3-ил 3-(1H-имидазол-4-ил)пропаноат, представленный структурой (II).



[0122] В некоторых вариантах осуществления процентная доля катионного липида в липосоме может составлять более чем 10%, более чем 20%, более чем 30%, более чем 40%, более чем 50%, более чем 60% или более чем 70%. В некоторых вариантах осуществления катионный (-ые) липид (-ы) составляет (-ют) приблизительно 30–50% (например, приблизительно 30–45%, приблизительно 30–40%, приблизительно 35–50%, приблизительно 35–45% или приблизительно 35–40%) липосомы по массе. В некоторых вариантах осуществления катионный липид (например, липид ICE) составляет приблизительно 30%, приблизительно 35%, приблизительно 40%, приблизительно 45% или приблизительно 50% липосомы в молярном соотношении.

Некатионные/вспомогательные липиды

[0123] В некоторых вариантах осуществления предложенные липосомы содержат один или более некатионных («вспомогательных») липидов. В настоящем документе фраза «некатионный липид» относится к любому нейтральному, цвиттерионному или анионному липиду. В настоящем документе фраза «анионный липид» относится к любому из множества липидных фрагментов, несущих чистый отрицательный заряд при выбранном рН, таком как физиологический рН. К некатионным липидам относятся, без ограничений, дистеароилфосфатидилхолин (DSPC), диолеоилфосфатидилхолин (DOPC), дипальмитоилфосфатидилхолин (DPPC), диолеоилфосфатидилглицерин (DOPG), дипальмитоилфосфатидилглицерин (DPPG), диолеоилфосфатидилэтаноламин (DOPE), пальмитоилолеоилфосфатидилхолин (POPC), пальмитоилолеоилфосфатидилэтаноламин (POPE), диолеоилфосфатидилэтаноламин 4-(N-малеимидометил)-циклогексан-1-карбоксилат (DOPE-мал), дипальмитоилфосфатидилэтаноламин (DPPE), димиристоилфосфоэтаноламин (DMPE), дистеароилфосфатидилэтаноламин (DSPE), фосфатидилсерин, сфинголипиды, цереброзиды, ганглиозиды, 16-О-монометил PE, 16-

О-диметил PE, 18-1-транс PE, 1-стеароил-2-олеоилфосфатидилэтаноламин (SOPE) или их смеси.

[0124] В некоторых вариантах осуществления такие некаатионные липиды могут использоваться по отдельности, но предпочтительно используются в комбинации с другими липидами, например, катионными липидами. В некоторых вариантах осуществления некаатионный липид может составлять в молярном соотношении от приблизительно 5% до приблизительно 90% или от приблизительно 10% до приблизительно 70% от общего количества липида, присутствующего в липосоме. В некоторых вариантах осуществления некаатионный липид представляет собой нейтральный липид, т. е. липид, который не несет чистого заряда в условиях, в которых композицию составляют и/или вводят. В некоторых вариантах осуществления процентная доля некаатионного липида в липосоме может составлять более чем 5%, более чем 10%, более чем 20%, более чем 30% или более чем 40%.

Липиды на основе холестерина

[0125] В некоторых вариантах осуществления предложенные липосомы содержат один или более липидов на основе холестерина. Например, к подходящим катионным липидам на основе холестерина относятся, например, DC Choi (N,N-диметил-N-этилкарбоксамидохолестерин), 1,4-бис(3-N-олеиламинопропил)пиперазин (Gao, et al. Biochem. Biophys. Res. Comm. 179, 280 (1991); Wolf et al. BioTechniques 23, 139 (1997); патент США № 5,744,335) или ICE. В некоторых вариантах осуществления липид на основе холестерина может составлять в молярном соотношении от приблизительно 2% до приблизительно 30% или от приблизительно 5% до приблизительно 20% от общего количества липида, присутствующего в липосоме. В некоторых вариантах осуществления процентная доля липида на основе холестерина в липидной наночастице может составлять более чем 5%, более чем 10%, более чем 20%, более чем 30% или более чем 40%.

ПЭГ-модифицированные липиды

[0126] В настоящем изобретении также предусмотрено использование модифицированных полиэтиленгликолем (ПЭГ) фосфолипидов и дериватизированных липидов, таких как производные церамиды (ПЭГ-CER), включая N-октаноилсфингозин-1-[сукцинил(метоксиполиэтиленгликоль)-2000] (C8 PEG-2000

церамид), либо отдельно, либо в комбинации с другими липидными составами, образующими транспортный носитель (например, липидную наночастицу). К предусмотренным ПЭГ-модифицированным липидам относятся, без ограничений, полиэтиленгликолевая цепь длиной до S кДа, ковалентно связанная с липидом с алкильной (-ыми) цепью (-ями) длиной C₆–C₂₀. Добавление таких компонентов может предотвращать агрегацию комплексов и также может обеспечивать средство увеличения срока циркуляции и улучшения доставки композиции из липида и нуклеиновой кислоты в ткани-мишени (Klibanov et al. (1990) FEBS Letters, 268 (1): 235–237) или же их можно выбирать для быстрого обмена составом *in vivo* (см. патент США № 5,885,613). Особенно полезными обмениваемыми липидами являются ПЭГ-церамиды, имеющие более короткие ацильные цепи (например, C14 или C18). ПЭГ-модифицированный фосфолипид и производные липиды по настоящему изобретению могут содержать в молярном соотношении от приблизительно 0% до приблизительно 20%, от приблизительно 0,5% до приблизительно 20%, от приблизительно 1% до приблизительно 15%, от приблизительно 4% до приблизительно 10% или приблизительно 2% от общего количества липидов, присутствующих в липосомном транспортном носителе.

[0127] В соответствии с различными вариантами осуществления выбор катионных липидов, неcatiонных липидов и/или ПЭГ-модифицированных липидов, которые образуют липидную наночастицу, а также относительное молярное соотношение таких липидов друг к другу зависят от характеристик выбранного (-ых) липида (-ов), природы предполагаемых клеток-мишеней, характеристик доставляемой МСНА. К дополнительным факторам относятся, например, насыщенность алкильной цепи, а также размер, заряд, рН, рКа, эффективность слияния с клетками и токсичность выбранного (-ых) липида (-ов). Таким образом, молярные соотношения можно корректировать соответствующим образом.

Полимеры

[0128] В некоторых вариантах осуществления подходящий доставляющий носитель составляют с использованием полимера в качестве носителя, отдельно или в комбинации с другими носителями, включая различные липиды, описанные в настоящем документе. Таким образом, в настоящем документе в некоторых вариантах осуществления к липосомным доставляющим носителям также относятся

наночастицы, содержащие полимеры. К подходящим полимерам могут относиться, например, полиакрилаты, полиалкилцианоакрилаты, полилактид, сополимеры полилактид-полигликолид, поликапролактоны, декстран, альбумин, желатин, альгинат, коллаген, хитозан, циклодекстрины, протамин, пегилированный протамин, PLL, пегилированный PLL и полиэтиленимин (PEI). Если присутствует PEI, он может представлять собой разветвленный PEI с молекулярной массой в диапазоне от 10 до 40 кДа, например, разветвленный PEI с молекулярной массой 25 кДа (Sigma № 408727).

[0129] Подходящая для настоящего изобретения липосома может включать в себя один или более из любых из катионных липидов, некатионных липидов, холестеринных липидов, липидов, ПЭГ-модифицированных липидов и/или полимеров, описанных в настоящем документе, в различных соотношениях. В качестве не имеющих ограничительного характера примеров подходящий липосомный состав может включать в себя комбинацию, выбранную из сКК-E12, DOPE, холестерина и DMG-PEG2K; C12-200, DOPE, холестерина и DMG-PEG2K; HGT4003, DOPE, холестерина и DMG-PEG2K; ICE, DOPE, холестерина и DMG-PEG2K или ICE, DOPE и DMG-PEG2K.

[0130] В различных вариантах осуществления катионные липиды (например, сКК-E12, C12-200, ICE и/или HGT4003) составляют приблизительно 30–60% (например, приблизительно 30–55%, приблизительно 30–50%, приблизительно 30–45%, приблизительно 30–40%, приблизительно 35–50%, приблизительно 35–45% или приблизительно 35–40%) липосомы в молярном соотношении. В некоторых вариантах осуществления процентная доля катионных липидов (например, сКК-E12, C12-200, ICE и/или HGT4003) равна или составляет более чем приблизительно 30%, приблизительно 35%, приблизительно 40%, приблизительно 45%, приблизительно 50%, приблизительно 55%, приблизительно 60% липосомы в молярном соотношении.

[0131] В некоторых вариантах осуществления соотношение катионного (-ых) липида (-ов) и некатионного (-ых) липида (-ов), липида (-ов) на основе холестерина, ПЭГ-модифицированного (-ых) липида (-ов) может находиться в диапазоне приблизительно 30–60 : 25–35 : 20–30 : 1–15 соответственно. В некоторых вариантах осуществления соотношение катионного (-ых) липида (-ов) и некатионного (-ых) липида (-ов), липида (-ов) на основе холестерина, ПЭГ-модифицированного (-ых)

липидов (-ов) составляет приблизительно 40 : 30 : 20 : 10 соответственно. В некоторых вариантах осуществления соотношение катионного (-ых) липидов (-ов) и некатионного (-ых) липидов (-ов), липидов (-ов) на основе холестерина, ПЭГ-модифицированного (-ых) липидов (-ов) составляет приблизительно 40 : 30 : 25 : 5 соответственно. В некоторых вариантах осуществления соотношение катионного (-ых) липидов (-ов) и некатионного (-ых) липидов (-ов), липидов (-ов) на основе холестерина, ПЭГ-модифицированного (-ых) липидов (-ов) составляет приблизительно 40 : 32 : 25 : 3 соответственно. В некоторых вариантах осуществления соотношение катионного (-ых) липидов (-ов) и некатионного (-ых) липидов (-ов), липидов (-ов) на основе холестерина, ПЭГ-модифицированного (-ых) липидов (-ов) составляет приблизительно 50 : 25 : 20 : 5.

Соотношение отдельных липидных компонентов

[0132] В вариантах осуществления, в которых липидная наночастица содержит три и не более трех отдельных компонентов липидов, соотношение общего содержания липидов (т. е. соотношения липидный компонент (1) : липидный компонент (2) : липидный компонент (3)) можно представить как $x : y : z$, где

$$(y + z) = 100 - x.$$

[0133] В некоторых вариантах осуществления каждый из «x», «y» и «z» представляет собой молярные процентные доли этих трех отдельных компонентов липидов, а соотношение представляет собой молярное соотношение.

[0134] В некоторых вариантах осуществления каждый из «x», «y» и «z» представляет собой массовые процентные доли этих трех отдельных компонентов липидов, а соотношение представляет собой массовое соотношение.

[0135] В некоторых вариантах осуществления липидный компонент (1), представленный переменной «x», представляет собой катионный липид на основе стерола.

[0136] В некоторых вариантах осуществления липидный компонент (2), представленный переменной «y», представляет собой вспомогательный липид.

[0137] В некоторых вариантах осуществления липидный компонент (3), представленный переменной «z», представляет собой ПЭГ-липид.

[0138] В некоторых вариантах осуществления переменная «х», представляющая молярную процентную долю липидного компонента (1) (например, катионного липида на основе стерола), составляет по меньшей мере приблизительно 10%, приблизительно 20%, приблизительно 30%, приблизительно 40%, приблизительно 50%, приблизительно 55%, приблизительно 60%, приблизительно 65%, приблизительно 70%, приблизительно 75%, приблизительно 80%, приблизительно 85%, приблизительно 90% или приблизительно 95%.

[0139] В некоторых вариантах осуществления переменная «х», представляющая молярную процентную долю липидного компонента (1) (например, катионного липида на основе стерола), составляет не более чем приблизительно 95%, приблизительно 90%, приблизительно 85%, приблизительно 80%, приблизительно 75%, приблизительно 70%, приблизительно 65%, приблизительно 60%, приблизительно 55%, приблизительно 50%, приблизительно 40%, приблизительно 30%, приблизительно 20% или приблизительно 10%. В вариантах осуществления переменная «х» составляет не более чем приблизительно 65%, приблизительно 60%, приблизительно 55%, приблизительно 50%, приблизительно 40%.

[0140] В некоторых вариантах осуществления переменная «х», представляющая молярную процентную долю липидного компонента (1) (например, катионного липида на основе стерола), составляет: по меньшей мере приблизительно 50%, но не менее чем приблизительно 95%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 90%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 85%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 80%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 75%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 70%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 65% или по меньшей мере приблизительно 50%, но не менее чем приблизительно 60%. В вариантах осуществления переменная «х» составляет по меньшей мере приблизительно 50%, но менее чем приблизительно 70%, по меньшей мере приблизительно 50%, но менее чем приблизительно 65% или по меньшей мере приблизительно 50%, но менее чем приблизительно 60%.

[0141] В некоторых вариантах осуществления переменная «х», представляющая массовую процентную долю липидного компонента (1) (например, катионного липида

на основе стерола), составляет по меньшей мере приблизительно 10%, приблизительно 20%, приблизительно 30%, приблизительно 40%, приблизительно 50%, приблизительно 55%, приблизительно 60%, приблизительно 65%, приблизительно 70%, приблизительно 75%, приблизительно 80%, приблизительно 85%, приблизительно 90% или приблизительно 95%.

[0142] В некоторых вариантах осуществления переменная «x», представляющая массовую процентную долю липидного компонента (1) (например, катионного липида на основе стерола), составляет не более чем приблизительно 95%, приблизительно 90%, приблизительно 85%, приблизительно 80%, приблизительно 75%, приблизительно 70%, приблизительно 65%, приблизительно 60%, приблизительно 55%, приблизительно 50%, приблизительно 40%, приблизительно 30%, приблизительно 20% или приблизительно 10%. В вариантах осуществления переменная «x» составляет не более чем приблизительно 65%, приблизительно 60%, приблизительно 55%, приблизительно 50%, приблизительно 40%.

[0143] В некоторых вариантах осуществления переменная «x», представляющая массовую процентную долю липидного компонента (1) (например, катионного липида на основе стерола), составляет: по меньшей мере приблизительно 50%, но не менее чем приблизительно 95%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 90%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 85%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 80%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 75%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 70%; по меньшей мере приблизительно 50%, но не менее чем приблизительно 65% или по меньшей мере приблизительно 50%, но не менее чем приблизительно 60%. В вариантах осуществления переменная «x» составляет по меньшей мере приблизительно 50%, но менее чем приблизительно 70%, по меньшей мере приблизительно 50%, но менее чем приблизительно 65% или по меньшей мере приблизительно 50%, но менее чем приблизительно 60%.

[0144] В некоторых вариантах осуществления переменная «z», представляющая молярную процентную долю липидного компонента (3) (например, ПЭГ-липид), составляет не более чем приблизительно 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20% или 25%. В вариантах осуществления переменная «z», представляющая

молярную процентную долю липидного компонента (3) (например, ПЭГ-липида), составляет приблизительно 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%. В вариантах осуществления переменная «z», представляющая молярную процентную долю липидного компонента (3) (например, ПЭГ-липида), составляет от приблизительно 1% до приблизительно 10%, от приблизительно 2% до приблизительно 10%, от приблизительно 3% до приблизительно 10%, от приблизительно 4% до приблизительно 10%, от приблизительно 1% до приблизительно 7,5%, от приблизительно 2,5% до приблизительно 10%, от приблизительно 2,5% до приблизительно 7,5%, от приблизительно 2,5% до приблизительно 5%, от приблизительно 5% до приблизительно 7,5% или от приблизительно 5% до приблизительно 10%.

[0145] В некоторых вариантах осуществления переменная «z», представляющая массовую процентную долю липидного компонента (3) (например, ПЭГ-липида), составляет не более чем приблизительно 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20% или 25%. В вариантах осуществления переменная «z», представляющая массовую процентную долю липидного компонента (3) (например, ПЭГ-липида), составляет приблизительно 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%. В вариантах осуществления переменная «z», представляющая массовую процентную долю липидного компонента (3) (например, ПЭГ-липида), составляет от приблизительно 1% до приблизительно 10%, от приблизительно 2% до приблизительно 10%, от приблизительно 3% до приблизительно 10%, от приблизительно 4% до приблизительно 10%, от приблизительно 1% до приблизительно 7,5%, от приблизительно 2,5% до приблизительно 10%, от приблизительно 2,5% до приблизительно 7,5%, от приблизительно 2,5% до приблизительно 5%, от приблизительно 5% до приблизительно 7,5% или от приблизительно 5% до приблизительно 10%.

[0146] Для композиций, содержащих три и только три отдельных липидных компонента, переменные «x», «y» и «z» могут находиться в любой комбинации при условии, что сумма всех этих трех переменных составляет 100% от общего содержания липидов.

Образование липосом, инкапсулирующих мРНК

[0147] Липосомные транспортные носители для использования в композициях по изобретению можно получить различными методиками, известными в настоящее время в данной области. Липосомы для использования в представленных композициях

могут быть получены с помощью различных методик, известных в настоящее время в данной области. Например, мультиламеллярные везикулы (MLV) можно получить в соответствии с традиционными методиками, такими как путем осаждения выбранного липида на внутреннюю стенку подходящего контейнера или сосуда в результате растворения липида в подходящем растворителе и последующего выпаривания растворителя с образованием тонкой пленки на внутренней стороне сосуда или путем распылительной сушки. Затем в сосуд можно добавить водную фазу встряхивающим движением, что приведет к образованию MLV. После этого можно образовать униламеллярные везикулы (ULV) путем гомогенизации, обработки ультразвуком или экструзии мультиламеллярных везикул. Кроме того, униламеллярные везикулы могут быть получены методиками удаления детергента.

[0148] В некоторых вариантах осуществления предложенные композиции содержат липосому, причем мРНК связана как с поверхностью липосомы, так и инкапсулирована внутри этой же липосомы. Например, при получении композиций по настоящему изобретению катионные липосомы могут связываться с мРНК посредством электростатических взаимодействий. Например, при получении композиций по настоящему изобретению катионные липосомы могут связываться с мРНК посредством электростатических взаимодействий.

[0149] В некоторых вариантах осуществления композиции и способы по изобретению включают в себя мРНК, инкапсулированную в липосоме. В некоторых вариантах осуществления в одну липосому могут быть инкапсулированы один или более типов мРНК. В некоторых вариантах осуществления один или более типов мРНК могут быть инкапсулированы в разные липосомы. В некоторых вариантах осуществления мРНК инкапсулируют в одну или более липосом, которые отличаются по своему липидному составу, молярному соотношению липидных компонентов, размеру, заряду (дзета-потенциалу), нацеливающим лигандам и/или их комбинациям. В некоторых вариантах осуществления одна или более липосом могут иметь разный состав катионных липидов на основе стерола, нейтрального липида, ПЭГ-модифицированного липида и/или их комбинаций. В некоторых вариантах осуществления одна или более липосом могут иметь другое молярное соотношение катионного липида на основе холестерина, нейтрального липида и ПЭГ-модифицированного липида, используемого для создания липосомы.

[0150] Процесс встраивания требуемой мРНК в липосому часто называют «загрузкой». Примеры способов описаны в публикации Lasic, et al., FEBS Lett., 312: 255–258, 1992, которая включена в настоящий документ путем ссылки. Встроенные в липосому нуклеиновые кислоты могут полностью или частично размещаться во внутреннем пространстве липосомы, в двухслойной мембране липосомы или могут быть связаны с внешней поверхностью мембраны липосомы. Встраивание нуклеиновой кислоты в липосому также в настоящем документе называется «инкапсуляцией», и в этом случае нуклеиновая кислота полностью находится во внутреннем пространстве липосомы. Цель встраивания мРНК в транспортный носитель, такой как липосома, часто заключается в защите нуклеиновой кислоты от окружающей среды, которая может содержать ферменты или химические вещества, которые расщепляют нуклеиновые кислоты, и/или системы или рецепторы, вызывающие быстрое выведение нуклеиновых кислот. Соответственно, в некоторых вариантах осуществления подходящий доставляющий носитель выполнен с возможностью повышения стабильности содержащейся в ней мРНК и/или облегчения доставки мРНК в клетку или ткань-мишень.

[0151] Подходящие липосомы в соответствии с настоящим изобретением могут быть выполнены с различными размерами. В некоторых вариантах осуществления предложенные липосомы могут быть выполнены меньшими, чем ранее известные липосомы, инкапсулирующие мРНК. В некоторых вариантах осуществления уменьшенный размер липосом ассоциируется с более эффективной доставкой мРНК. При выборе подходящего размера для липосомы может быть принято во внимание местоположение клетки или ткани-мишени и в некоторой степени область применения, для которой изготавливается липосома.

[0152] В некоторых вариантах осуществления подходящий размер липосомы выбирают так, чтобы упростить системное распределение антитела, кодируемого мРНК. В некоторых вариантах осуществления может быть желательно ограничить трансфекцию мРНК определенными клетками или тканями. Например, для нацеливания на гепатоциты размеры липосомы можно подобрать таким образом, чтобы ее размеры были меньше, чем поры эндотелиального слоя, выстилающего печеночные синусоиды в печени; в таких случаях липосома сможет легко проникать в такие эндотелиальные поры и достигать гепатоциты-мишени.

[0153] Альтернативно или дополнительно размеры липосомы можно подобрать таким образом, чтобы она имела достаточный диаметр для ограничения или явного предотвращения попадания в определенные клетки или ткани.

[0154] Для определения размеров популяции липосом можно использовать разнообразные известные в данной области альтернативные способы. Один такой способ формирования размеров описан в патенте США № 4,737,323, включенном в настоящий документ путем ссылки. Ультразвуковая обработка суспензии липосом либо с помощью ванны, либо с помощью зонда позволяет получить прогрессивное уменьшение размера вплоть до небольших ULV менее чем приблизительно 0,05 мкм в диаметре. Гомогенизация представляет собой другой способ, основанный на подаче сдвигового усилия для фрагментации крупных липосом на более мелкие. При типичной процедуре гомогенизации MLV рециркулируют по стандартному гомогенизатору эмульсий до тех пор, пока не будут обнаруживаться липосомы выбранного размера, как правило, в диапазоне от приблизительно 0,1 до 0,5 мкм. Размер липосом можно определить методом квазиэластического рассеяния света (QEELS), как описано в публикации Bloomfield, Ann. Rev. Biophys. Bioeng., 10:421–150 (1981), включенной в настоящий документ путем ссылки. Средний диаметр липосомы можно уменьшить путем обработки сформированных липосом ультразвуком. Периодические циклы обработки ультразвуком можно чередовать с оценкой методом QEELS, регулируя эффективный синтез липосом.

ПРИМЕРЫ

[0155] Несмотря на то что некоторые соединения, композиции и способы по настоящему изобретению были описаны специфично в соответствии с некоторыми вариантами осуществления, представленные ниже примеры служат только для иллюстрации соединений по настоящему изобретению и не предназначены для того, чтобы ограничивать их.

Пример 1. Синтез и сравнение конструктов мРНК hCFTR

[0156] Кодон-оптимизированную матричную РНК человеческого муковисцидозного регулятора трансмембранной проводимости (CFTR) синтезировали путем транскрипции *in vitro* из плазмидного ДНК-шаблона, кодирующего ген, после

чего была добавлена структура 5'-кэпа (Кэп 1) (Fechter, P.; Brownlee, G.G. "Recognition of mRNA cap structures by viral and cellular proteins" J. Gen. Virology 2005, 86, 1239–1249) и хвост 3' поли(А) длиной приблизительно 250 нуклеотидов по данным геле-электрофореза. Нетранслируемые области (UTR) 5' и 3', присутствующие в каждом мРНК-продукте, представлены как X и Y соответственно и соответствуют приведенному определению (см. ниже).

Пример кодон-оптимизированных мРНК человеческого муковисцидозного регулятора трансмембранной проводимости (CFTR)

Длина конструктора:

X – SEQ ID NO: 1 – Y

Последовательности 5' и 3' UTR:

X (последовательность 5' UTR) =

GGACAGAUCGCCUGGAGACGCCAUCCACGCUGUUUUGACCUCCAUAAGAAGACA
CCGGGACCGAUCCAGCCUCCGCGGCCGGGAACGGUGCAUUGGAACGCGGAUUC
CCCGUGCCAAGAGUGACUCACCGUCCUUGACACG (**SEQ ID NO: 4**)

Y (последовательность 3' UTR) =

CGGGUGGCAUCCCUGUGACCCCUCUCCAGUGCCUCUCCUGGCCUUGGAAGUUG
CCACUCCAGUGCCCACCAGCCUUGUCCUAAUAAAUAAGUUGCAUCAAGCU
(**SEQ ID NO: 5**)

ИЛИ

GGGUGGCAUCCCUGUGACCCCUCUCCAGUGCCUCUCCUGGCCUUGGAAGUUGC
CACUCCAGUGCCCACCAGCCUUGUCCUAAUAAAUAAGUUGCAUCAAGCU
(**SEQ ID NO: 6**)

[0157] Пример кодон-оптимизированной последовательности мРНК человеческого CFTR включает в себя SEQ ID NO: 1, как описано в разделе подробного описания.

[0158] Пример полноразмерной кодон-оптимизированной последовательности мРНК человеческого CFTR показан ниже:

GGACAGAUCGCCUGGAGACGCCAUCCACGCUGUUUUGACCUCCAUAAGAAGACA
CCGGGACCGAUCCAGCCUCCGCGGCCGGGAACGGUGCAUUGGAACGCGGAUUC
CCCGUGCCAAGAGUGACUCACCGUCCUUGACACGAUGCAACGCUCUCCUCUUG
AAAAGGCCUCGGUGGUGUCCAAGCUCUUCUUCUCGUGGACUAGACCCAUCCUG
AGAAAGGGGUACAGACAGCGCUUGGAGCUGUCCGAUAUCUAUCAAUCCCUUC
CGUGGACUCCGCGGACAACCUGUCCGAGAAGCUCGAGAGAGAAUGGGACAGAG
AACUCGCCUCAAGAAGAACCCGAAGCUGAUUAAUGCGCUUAGGCGGUGCUUU
UUCUGGCGGUUCAUGUUCUACGGCAUCUUCUCCUCUACCUGGGAGAGGUCACCAA
GGCCGUGCAGCCCCUGUUGCUGGGACGGAUUAUUGCCUCCUACGACCCCGACA
ACAAGGAAGAAAGAAGCAUCGCUAUCUACUUGGGCAUCGGUCUGUGCCUGCU
UUUCAUCGUCCGGACCCUCUUGUUGCAUCCUGCUAUUUUCGGCCUGCAUCACA
UUGGCAUGCAGAUGAGAAUUGCCAUGUUUCCUGAUCUACAAGAAAACUCU
GAAGCUCUCGAGCCGCGUGCUUGACAAGAUUCCAUCGGCCAGCUCGUGUCCC
UGCUCUCCAACAUCUGAACAAGUUCGACGAGGGCCUCGCCUUGGCCACUUC
GUGUGGAUCGCCCCUCUGCAAGUGGCGCUUCUGAUGGGCCUGAUCUGGGAGCU
GCUGCAAGCCUCGGCAUUCUGUGGGCUUGGAUUCUGAUCGUGCUGGCACUGU
UCCAGGCCGGACUGGGGCGGAUGAUGAUGAAGUACAGGGACCAGAGAGCCGG
AAAGAUUUCGAACGGCUGGUGAUCACUUCGGAAAUGAUCGAAAACAUCAG
UCAGUGAAGGCCUACUGCUGGGAAGAGGCCAUGGAAAAGAUGAUUGAAAACC
UCCGGCAAACCGAGCUGAAGCUGACCCGCAAGGCCGCUUACGUGCGCUAUUUC
AACUCGUCCGCUUUCUUCUUCUCCGGGUUCUUCGUGGUGUUUCUCUCCGUGCU
CCCCUACGCCUGAUUAAGGGAAUCAUCCUCAGGAAGAUCUUCACCACCAUUU
CCUUCUGUAUCGUGCUCGCAUGGCCGUGACCCGGCAGUUCCCAUGGGCCGUG
CAGACUUGGUACGACUCCCUGGGAGCCAUAACAAGAUCAGGACUCCUUCA
AAAGCAGGAGUACAAGACCCUCGAGUACAACCUGACUACUACCGAGGUCGUGA
UGGAAAACGUCACCGCCUUUUGGGAGGAGGGAUUUGGCGAACUGUUCGAGAA
GGCCAAGCAGAACAACAACAACCGCAAGACCUCGAACGGUGACGACUCCCUCU
UCUUUCAAACUUCAGCCUGCUCGGGACGCCCGUGCUGAAGGACAUUAACUUC
AAGAUCGAAAGAGGACAGCUCCUGGCGGUGGCCGGAUCGACCGGAGCCGGAAA
GACUUCCUGCUGAUGGUGAUCAUGGGAGAGCUUGAACCUAGCGAGGGAAAG
AUCAAGCACUCCGGCCGCAUCAGCUUCUGUAGCCAGUUUUCUGGAUCAUGCC
CGGAACCAUUAAGGAAAACAUCAUCUUCGGCGUGUCCUACGAUGAAUACCGCU
ACCGGUCCGUGAUCAAAGCCUGCCAGCUGGAAGAGGAUAUUUCAAGUUCGCG

GAGAAAGAUAAACAUCGUGCUGGGCGAAGGGGGUAAUACCUUGUCGGGGGGCC
AGCGGGCUAGAAUCUCGCUGGCCAGAGCCGUGUAUAAGGACGCCGACCUGUAU
CUCCUGGACUCCCCUUCGGAUACCUGGACGUCCUGACCGAAAAGGAGAUCU
CGAAUCGUGCGUGUGCAAGCUGAUGGCUAACAAGACUCGCAUCCUCGUGACCU
CCAAAUGGAGCACCUGAAGAAGGCAGACAAGAUUCUGAUUCUGCAUGAGGG
GUCCUCCUACUUUUACGGCACCUUCUCGGAGUUGCAGAACUUGCAGCCCGACU
UCUCAUCGAAGCUGAUGGGUUGCGACAGCUUCGACCAGUUCUCCGCCGAAAGA
AGGAACUCGAUCCUGACGGAAACCUUGCACCGCUUCUCUUUGGAAGGCGACGC
CCCUGUGUCAUGGACCGAGACUAAGAAGCAGAGCUUCAAGCAGACCGGGGAAU
UCGGCGAAAAGAGGAAGAACAGCAUCUUGAACCCCAUUAACUCCAUCCGCAAG
UUCUCAUUCGUGCAAAGACGCCACUGCAGAUGAACGGCAUUGAGGAGGACUC
CGACGAACCCCUUGAGAGGGCGCCUGUCCUGGUGCCGGACAGCGAGCAGGGAG
AAGCCAUCCUGCCUCGGAUUUCCGUGAUCUCCACUGGUCCGACGCUCCAAGCC
CGGCGGGCGGCAGUCCGUGCUGAACCUGAUGACCCACAGCGUGAACCAGGGCCA
AAACAUUCACCGCAAGACUACCGCAUCCACCCGGAAAGUGUCCUGGCACCUC
AAGCGAAUCUUAACCGAGCUCGACAUCUACUCCCGGAGACUGUCGCAGGAAACC
GGGCUCGAAAUUUCCGAAGAAAUCAACGAGGAGGAUCUGAAAGAGUGCUUCU
UCGACGAUAUGGAGUCGAUACCCGCCGUGACGACUUGGAACACUUAUCUGCGG
UACAUCACUGUGCACAAGUCAUUGAUCUUCGUGCUGAUUUGGUGCCUGGUGA
UUUUCUGGCCGAGGUCGCGGCCUCACUGGUGGUGCUCUGGCUGUUGGGAAAC
ACGCCUCUGCAAGACAAGGGAAACUCCACGCACUCGAGAAACAACAGCUAUGC
CGUGAUUAUCACUUCACCUCUCUUAUUACGUGUUCUACAUCUACGUCGGAG
UGGCGGAUACCCUGCUCGCGAUGGGUUUCUUCAGAGGACUGCCGCUGGUCCAC
ACCUUGAUCACCGUCAGCAAGAUUCUUCACCACAAGAUGUUGCAUAGCGUGCU
GCAGGCCCCCAUGUCCACCCUCAACACUCUGAAGGCCGGAGGCAUUCUGAACA
GAUUCUCCAAGGACAUCGCUAUCUGGACGAUCUCCUGCCGCUUACCAUCUUU
GACUUCAUCCAGCUGCUGCUGAUCGUGAUUGGAGCAAUCGCAGUGGUGGGCGG
UGCUGCAGCCUUAUAUUUCGUGGCCACUGUGCCGGUCAUUGUGGGCGUUCAUC
AUGCUGCGGGCCUACUUCUCCAACCAGCCAGCAGCUGAAGCAACUGGAAUC
CGAGGGACGAUCCCCAUCUUCACUCACCUUGUGACGUCGUUGAAGGGACUGU
GGACCCUCCGGGCUUUCGGACGGCAGCCUACUUCGAAACCCUCUCCACAAG
GCCUGAACCUCACACCGCCA AUUGGUUCCUGUACCUGUCCACCCUGCGGUG
GUUCCAGAUGC GCAUCGAGAUGAUUUUCGUCAUCUUCUUAUCGCGGUCACAU

UCAUCAGCAUCCUGACUACCGGAGAGGGAGAGGGACGGGUCGGAUAAUCCU
GACCCUCGCCAUGAACAUUAUGAGCACCCUGCAGUGGGCAGUGAACAGCUCGA
UCGACGUGGACAGCCUGAUGCGAAGCGUCAGCCGCGUGUUCAAGUUCAUCGAC
AUGCCUACUGAGGGAAAACCCACUAAGUCCACUAAGCCCUACAAAAAUGGCCA
GCUGAGCAAGGUCAUGAUCUACGAAAACUCCACGUGAAGAAGGACGAUAAU
UGGCCCUCCGGAGGUCAAAUGACCGUGAAGGACCUGACCGCAAAGUACACCGA
GGGAGGAAACGCCAUUCUCGAAAACAUCAGCUUCUCCAUUUCGCCGGGACAGC
GGGUCGGCCUUCUCGGGCGGACCGGUUCCGGGAAGUCAACUCUGCUGUCGGCU
UUCUCCGGCUGCUGAAUACCGAGGGGGAAAUCCAAAUUGACGGCGUGUCUUG
GGAUUCCAUUACUCUGCAGCAGUGGCGGAAGGCCUUCGGCGUGAUCCCCAGA
AGGUGUUCAUCUUCUCGGGUACCUUCCGGAAGAACCUGGAUCCUUACGAGCAG
UGGAGCGACCAAGAAAUCUGGAAGGUCGCCGACGAGGUCGGCCUGCGCUCCGU
GAUUGAACAAUUUCCUGGAAAGCUGGACUUCGUGCUCGUCGACGGGGGAUGU
GUCCUGUCGCACGGACAUAAGCAGCUCAUGUGCCUCGCACGGUCCGUGCUCUC
CAAGGCCAAGAUUCUGCUGCUGGACGAACCUUCGGCCCACCUGGAUCCGGUCA
CCUACCAGAUCAUCAGGAGGACCCUGAAGCAGGCCUUUGCCGAUUGCACCGUG
AUUCUCUGCGAGCACCGCAUCGAGGCCAUGCUGGAGUGCCAGCAGUUCUGGU
CAUCGAGGAGAACAAGGUCCGCCAAUACGACUCCAUUCAAAGCUCCUCAACG
AGCGGUCGCUGUUCAGACAAGCUAUUUCACCGUCCGAUAGAGUGAAGCUCUUC
CCGCAUCGGAACAGCUCAAAGUGCAAUUCGAAGCCGCAGAUCGCAGCCUUGAA
GGAAGAGACUGAGGAAGAGGUGCAGGACACCCGGCUUUAACGGGUGGCAUCC
CUGUGACCCCUCCCAGUGCCUCUCCUGGCCUUGGAAGUUGCCACUCCAGUGC
CCACCAGCCUUGUCCUAAUAAAAUUAAGUUGCAUCAAGCU (SEQ ID NO: 7)

[0159] В другом примере полноразмерная кодон-оптимизированная последовательность мРНК человеческого CFTR показана ниже:

GGACAGAUCGCCUGGAGACGCCAUCCACGCUGUUUUGACCUCCAUAGAAGACA
CCGGGACCGAUCCAGCCUCCGCGGCCGGGAACGGUGCAUUGGAACGCGGAUUC
CCCGUGCCAAGAGUGACUCACCGUCCUUGACACGAUGCAACGCUCUCCUCUUG
AAAAGGCCUCGGUGGUGUCCAAGCUCUUCUUCUCGUGGACUAGACCCAUCCUG
AGAAAGGGGUACAGACAGCGCUUGGAGCUGUCCGAUAUCAUCAAUCCUUC
CGUGGACUCCGCGGACAACCUUGUCCGAGAAGCUCGAGAGAGAAUGGGACAGAG
AACUCGCCUCAAGAAGAACCCGAAGCUGAUUAAUGCGCUUAGGCGGUGCUUU

UUCUGGCGGUUCAUGUUCUACGGCAUCUUCCUCUACCUGGGAGAGGUCACCAA
GGCCGUGCAGCCCCUGUUGCUGGGACGGAUUAUUGCCUCCUACGACCCCGACA
ACAAGGAAGAAAGAAGCAUCGCUAUCUACUUGGGCAUCGGUCUGUGCCUGCU
UUUCAUCGUCCGGACCCUCUUGUUGCAUCCUGCUAUUUUCGGCCUGCAUCACA
UUGGCAUGCAGAUGAGAAUUGCCAUGUUUCCUGAUCUACAAGAAAACUCU
GAAGCUCUCGAGCCGCGUGCUUGACAAGAUUCCAUCGGCCAGCUCGUGUCCC
UGCUCUCCAACAUCUGAACAAGUUCGACGAGGGCCUCGCCUGGCCACUUC
GUGUGGAUCGCCCCUCUGCAAGUGGCGCUUCUGAUGGGCCUGAUCUGGGAGCU
GCUGCAAGCCUCGGCAUUCUGUGGGCUUGGAUUCUGAUCGUGCUGGCACUGU
UCCAGGCCGGACUGGGGCGGAUGAUGAUGAAGUACAGGGACCAGAGAGCCGG
AAAGAUUUCGAAACGGCUGGUGAUCACUUCGGAAAUGAUCGAAAACAUCAG
UCAGUGAAGGCCUACUGCUGGGAAGAGGCCAUGGAAAAGAUGAUUGAAAACC
UCCGGCAAACCGAGCUGAAGCUGACCCGCAAGGCCGCUUACGUGCGCUAUUUC
AACUCGUCCGCUUUCUUCUUCUCCGGGUUCUUCGUGGUGUUUCUCUCCGUGCU
CCCCUACGCCUGAUUAAGGGAAUCAUCCUCAGGAAGAUCUUCACCACCAUUU
CCUUCUGUAUCGUGCUCGCAUGGCCGUGACCCGGCAGUUCCCAUGGGCCGUG
CAGACUUGGUACGACUCCCUGGGAGCCAUAACAAGAUCAGGACUCCUUC
AAAGCAGGAGUACAAGACCCUCGAGUACAACCUGACUACUACCGAGGUCGUGA
UGGAAAACGUCACCGCCUUUUGGGAGGAGGGAUUUGGCGAACUGUUCGAGAA
GGCCAAGCAGAACAACAACAACCGCAAGACCUCGAACGGUGACGACUCCCUCU
UCUUUCAAACUUCAGCCUGCUCGGGACGCCCGUGCUGAAGGACAUAACUUC
AAGAUCGAAAGAGGACAGCUCCUGGCGGUGGCGGAUCGACCGGAGCCGGAAA
GACUUCCUGCUGAUGGUGAUCAUGGGAGAGCUUGAACCUAGCGAGGGAAAG
AUCAAGCACUCCGGCCGCAUCAGCUUCUGUAGCCAGUUUUCUGGAUCAUGCC
CGGAACCAUUAAGGAAAACAUCAUCUUCGGCGUGUCCUACGAUGAAUACCGCU
ACCGGUCCGUGAUCAAAGCCUGCCAGCUGGAAGAGGAUAUUUCAAGUUCGCG
GAGAAAGAUACAUCGUGCUGGGCGAAGGGGGUAUUACCUUGUCGGGGGGCC
AGCGGGCUAGAAUCUCGCUGGCCAGAGCCGUGUAUAAGGACGCCGACCUGUAU
CUCCUGGACUCCCCUUCGGAUACCUGGACGUCCUGACCGAAAAGGAGAUCU
CGAAUCGUGCGUGUGCAAGCUGAUGGCUAACAAGACUCGCAUCCUCGUGACCU
CCAAAUGGAGCACCUGAAGAAGGCAGACAAGAUUCUGAUUCUGCAUGAGGG
GUCCUCCUACUUUACGGCACCUUCUCGGAGUUGCAGAACUUGCAGCCCGACU
UCUCAUCGAAGCUGAUGGGUUGCGACAGCUUCGACCAGUUCUCCGCCGAAAGA

AGGAACUCGAUCCUGACGGAAACCUUGCACCGCUUCUCUUUGGAAGGCGACGC
CCCUGUGUCAUGGACCGAGACUAAGAAGCAGAGCUUCAAGCAGACCGGGGAAU
UCGGCGAAAAGAGGAAGAACAGCAUCUUGAACCCCAUUAACUCCAUCGCAAG
UUCUCAAUUCGUGCAAAGACGCCACUGCAGAUGAACGGCAUUGAGGAGGACUC
CGACGAACCCCUUGAGAGGGCGCCUGUCCCUGGUGCCGGACAGCGAGCAGGGAG
AAGCCAUCCUGCCUCGGAUUUCCGUGAUCUCCACUGGUCCGACGCUCCAAGCC
CGGCGGCGGCAGUCCGUGCUGAACCUGAUGACCCACAGCGUGAACCAGGGCCA
AAACAUUCACCGCAAGACUACCGCAUCCACCCGGAAAGUGUCCUGGCACCUC
AAGCGAAUCUUACCGAGCUCGACAUCUACUCCC GGAGACUGUCGCAGGAAACC
GGGCUCGAAAUUUCCGAAGAAAUCAACGAGGAGGAUCUGAAAGAGUGCUUCU
UCGACGAUAUGGAGUCGAUACCCGCCGUGACGACUUGGAACACUUAUCUGCGG
UACAUCACUGUGCACAAGUCAUUGAUCUUCGUGCUGAUUUGGUGCCUGGUGA
UUUUCUGGCCGAGGUCGCGGCCUCACUGGUGGUGCUCUGGCUGUUGGGAAAC
ACGCCUCUGCAAGACAAGGGAAACUCCACGCACUCGAGAAACAACAGCUAUGC
CGUGAUUAUCACUUCACCUCUCUUAUACGUGUUCUACAUCUACGUCGGAG
UGGCGGAUACCCUGCUCGCGAUGGGUUUCUUCAGAGGACUGCCGCUGGUCCAC
ACCUUGAUCACCGUCAGCAAGAUUCUUCACCACAAGAUGUUGCAUAGCGUGCU
GCAGGCCCCCAUGUCCACCCUCAACACUCUGAAGGCCGGAGGCAUUCUGAACA
GAUUCUCCAAGGACAUCGCUAUCUCCUGGACGAUCUCCUGCCGCUUACCAUCUU
GACUUCAUCCAGCUGCUGCUGAUCGUGAUUGGAGCAAUCGCAGUGGUGGCGG
UGCUGCAGCCUUAUAUUUCGUGGCCACUGUGCCGGUCAUUGUGGCGUUCAUC
AUGCUGCGGGCCUACUUCUCCAAACCAGCCAGCAGCUGAAGCAACUGGAAUC
CGAGGGACGAUCCCCAUCUUCACUCACCUUGUGACGUCGUUGAAGGGACUGU
GGACCCUCCGGGCUUUCGGACGGCAGCCUACUUCGAAACCCUCUCCACAAG
GCCUGAACCUCACACCGCCA AUUGGUUCCUGUACCUGUCCACCCUGCGGUG
GUUCCAGAUGC GCAUCGAGAUGAUUUUCGUCAUCUUCUUAUCGCGGUCACAU
UCAUCAGCAUCCUGACUACCGGAGAGGGAGAGGGACGGGUCGGAAUAAUCCU
GACCCUCGCCAUGAACAUUAUGAGCACCCUGCAGUGGGCAGUGAACAGCUCGA
UCGACGUGGACAGCCUGAUGCGAAGCGUCAGCCGCGUGUUAAGUUAUCGAC
AUGCCUACUGAGGGAAAACCCACUAAGUCCACUAAGCCCUACAAAAAUGGCCA
GCUGAGCAAGGUCAUGAUCAUUCGAAAACUCCACGUGAAGAAGGACGAUAUU
UGGCCCUCCGGAGGUCAAAUGACCGUGAAGGACCUGACCGCAAAGUACACCGA
GGGAGGAAACGCCAUUCUCGAAAACAUCAGCUUCUCCA UUCGCCGGGACAGC

GGGUCGGCCUUCUCGGGCGGACCGGUUCCGGGAAGUCAACUCUGCUGUCGGCU
UCCUCCGGCUGCUGAAUACCGAGGGGGAAAUCCAAAUUGACGGCGUGUCUUG
GGAUUCCAUUACUCUGCAGCAGUGGCGGAAGGCCUUCGGCGUGAUCCCCAGA
AGGUGUUCAUCUUCUCGGGUACCUUCCGGAAGAACCUGGAUCCUUACGAGCAG
UGGAGCGACCAAGAAAUCUGGAAGGUCGCCGACGAGGUCGGCCUGCGCUCCGU
GAUUGAACAAUUUCCUGGAAAGCUGGACUUCGUGCUCGUCGACGGGGGAUGU
GUCCUGUCGCACGGACAUAAGCAGCUCAUGUGCCUCGCACGGUCCGUGCUCUC
CAAGGCCAAGAUUCUGCUGCUGGACGAACCUUCGGCCCACCUGGAUCCGGUCA
CCUACCAGAUCAUCAGGAGGACCCUGAAGCAGGCCUUUGCCGAUUGCACCGUG
AUUCUCUGCAGCACCACAUCGAGGCCAUGCUGGAGUGCCAGCAGUUCUGGU
CAUCGAGGAGAACAAGGUCCGCCAAUACGACUCCAUUCAAAGCUCUCAACG
AGCGGUCGCUGUUCAGACAAGCUAUUUCACCGUCCGAUAGAGUGAAGCUCUUC
CCGCAUCGGAACAGCUCAAAGUGCAAUUCGAAGCCGCAGAUCGCAGCCUUGAA
GGAAGAGACUGAGGAAGAGGUGCAGGACACCCGGCUUUAAGGGUGGCAUCCC
UGUGACCCCUCCCCAGUGCCUCUCCUGGCCUUGGAAGUUGCCACUCCAGUGCC
CACCAGCCUUGUCCAAUAAAUAAGUUGCAUCAAAAGCU (SEQ ID NO: 8)

Сравнение конструктов мРНК hCFTR

[0160] Предыдущая последовательность hCFTR (SEQ ID NO: 2) была кодон-оптимизирована с использованием промотора T7. После изменения промотора, используемого для синтеза мРНК hCFTR на SP6, была синтезирована «более чистая» мРНК в отношении преждевременно прерванных последовательностей, но получили вторые фрагменты размером приблизительно 1800 нуклеотидов («длинные полимеры») в малых количествах. Их визуализировали с помощью гель-электрофореза, как показано на **Фиг. 1**. На **Фиг. 1** полоса 1 содержит РНК-лэддер, полоса 2 содержит мРНК с SEQ ID NO: 1, а полоса 3 содержит мРНК с SEQ ID NO: 2. Как показано стрелкой, в полосе 3 присутствует вторичный полинуклеотидный фрагмент, имеющий длину приблизительно 1800 нуклеотидов. Были сконструированы несколько новых последовательностей (относительно SEQ ID NO: 2) с сайтовыми мутациями для удаления предполагаемых скрытых промоторов, но это не приводило к исчезновению вторичных фрагментов ~ 1800 нуклеотидов. Выполняли полную кодоновую реоптимизацию с созданием SEQ ID NO: 1, которая успешно привела к получению

мРНК-продукта без дополнительной продукции вторых фрагментов

~ 1800 нуклеотидов (полоса 1).

[0161] Таким образом, SEQ ID NO: 1 особенно пригодна для использования в однородной, безопасной и эффективной фармацевтической композиции.

Пример 2. Дополнительные примеры кодон-оптимизированных последовательностей CFTR

[0162] Представленные ниже дополнительные примеры кодон-оптимизированных последовательностей используются для синтеза мРНК CFTR для безопасного и эффективного клинического использования:

ATGCAGAGGAGCCCACTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCAGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAAAAACCCCAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTTACAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCCAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGAAGATATCAG
AGCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCT
ACTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAG
CTCAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCTT
CTTCAGTGGCTTCTTTGTTGTCTTCTGTCTGTTCTGCCATATGCACTGATAAAAG
GCATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATG
GCCGTCACAAGACAGTTCCTGCGGCTGTGCAGACCTGGTACGATTCCTTGGGGG
CCATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAAT

ACAACCTCACCACCACTGAAGTGGTCATGGAAAATGTGACAGCCTTTTGGGAGG
AGGGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAG
ACGAGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCACTGCTCGGGACCC
CTGTGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGC
AGGCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTG
GAGCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAG
TTTTCTGGATCATGCCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGTCT
ATGATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACA
TCTCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCT
TTCTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCCGGGCAGTCTACAAGGATGC
AGACCTCTACTTGTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAA
GAAATTTTTGAAAGCTGTGTGTGCAAACCTGATGGCAAACAAGACCAGGATTCTTG
TCACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATG
AAGGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGA
CTTCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGA
AGAAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCCC
CAGTTTCTTGGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTG
GTGAAAAGAGGAAAAATTCAATTCTCAATCCAATTAACAGTATTCGCAAGTTCA
GCATTGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACG
AGCCGCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCA
TCCTGCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCG
GCAGAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCAC
AGGAAGACTACAGCTTCTACCCGAAAGTGTCTCTGGCCCCTCAGGCCAATCTGA
CCGAGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTGGAAATAT
CTGAAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGA
GCATCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTACTGTCCACAA
GAGCCTCATATTTGTCCTCATCTGGTGCCTGGTTATTTTCCCTCGCTGAGGTGGCGG
CCAGTCTTGTTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAA
TAGTACTCACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGC
TACTACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTT
CTCCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCAC
CATAAAATGCTTCATTCTGTCTGCAGGCACCCATGAGCACTTTGAACACATTGA
AGGCTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCT

CCTCCCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCA
TAGCAGTGGTTGCTGTCCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATT
GTTGCCTTTATTATGCTCAGAGCTTACTTCCTGCAAACCTTCTCAACAGCTCAAACA
GCTAGAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAG
GGACTGTGGACTCTGAGAGCATTCTGGGCGACAGCCTTACTTTGAGACACTGTTCC
ACAAGGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCG
CTGGTTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCT
TCATTTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGAATCATCCTCA
CGCTGGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGA
TGTGGATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCTA
CTGAGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCA
AGGTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCAGCG
GGGGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATG
CCATTTTGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCT
GGGTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTG
AATACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCTG
CAGCAGTGGAGAAAAGCATTGTTGGGTCATTCCACAGAAAGTTTTTCATCTTCTCTG
GCACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCT
GGAAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTTCTGGCA
AGCTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACA
GCTGATGTGCCTCGCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATG
AGCCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAA
GCAAGCTTTTGCCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATG
CTGGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGAC
AGCATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCAT
CTGACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCC
CCAGATCGCTGCCCTCAAGGAGGAAACTGAGGAAGAGGTGCAGGATACCCGCCT
GTGA (SEQ ID NO: 21)

ATGCAGAGGAGCCCACTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCTGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAACCCTAAGCTTATCAATGCTCT

GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTTACAAAAAGACCCTG
AAACTTTCCTCAAGAGTGTTAGATAAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGCCAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCCTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATTCAGA
GCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCTA
CTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAGCT
CAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCTTCT
TCAGTGGCTTCTTTGTTGTCTTCCTGTCTGTTCTGCCATATGCACTGATAAAAGGC
ATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATGGC
CGTCACAAGACAGTTCCTTGGGCTGTGCAGACCTGGTACGATTCCTTGGGGGCC
ATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAATAC
AACCTCACCACCCTGAAGTGGTCATGGAAAATGTGACAGCCTTTTTGGGAGGAG
GGTTTTGGAGAATTGTTTCGAGAAGGCCAAAGCAGAATAACAACAACAGGAAGACG
AGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCACTGCTCGGGACCCCTG
TGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGCAG
GCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTGGA
GCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAGTT
TTCCTGGATCATGCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGTCTAT
GATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACATC
TCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCTTT
CTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGCAG
ACCTCTACTTGTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAAGA
AATTTTTGAAAGCTGTGTGTGCAAACTGATGGCAAACAAGACCAGGATTCTTGTC
ACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATGAA
GGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGACT
TCTCCTCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGAAG
AAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCCCCA

GTTTCTTGGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTGGTG
AAAAGAGGAAAAATTCAATTCTCAATCCAATTAACAGTATTCGCAAGTTCAGCAT
TGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACGAGCC
GCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCATCCT
GCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCGGCA
GAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCACAGG
AAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCTCAGGCCAATCTGACCG
AGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTGGAAATATCTG
AAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGAGCA
TCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTACTGTCCACAAGAG
CCTCATATTTGTCCCTCATCTGGTGCCTGGTTATTTTCCTCGCTGAGGTGGCGGCCA
GTCTTGTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAATAG
TACTCACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGCTAC
TACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTTCTT
CCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCACCAT
AAAATGCTTCATTCTGTCTGCAGGCACCCATGAGCACTTTGAACACATTGAAGG
CTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCTCCTC
CCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCATAGC
AGTGGTTGCTGTCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATTGTTG
CCTTTATTATGCTCAGAGCTTACTTCTGCAAACCTTCTCAACAGCTCAAACAGCTA
GAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAGGGAC
TGTGGACTCTGAGAGCATTTCGGGCGACAGCCTTACTTTGAGACACTGTTCCACAA
GGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCGCTGG
TTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCTTCAT
TTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGGAATCATCCTCACGCT
GGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGATGTG
GATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCTACTG
AGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCAAG
GTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGCGGG
GGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATGCC
ATTTTGGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCTGG
GTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCCTCGGCTCCTGAA
TACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCCCTGCA

GCAGTGGAGAAAAGCATTGTTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTCTGGC
ACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCTGG
AAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTCCTGGCAAG
CTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACAGC
TGATGTGCCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATGAG
CCCAGTGCTCACCTTGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAAGC
AAGCTTTTGCCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATGCT
GGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGACAG
CATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCATCT
GACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCCCC
AGATCGCTGCCCTCAAGGAGGAACTGAGGAAGAGGTGCAGGATACCCGCCTGT
GA (SEQ ID NO: 22)

ATGCAGAGGAGCCCCTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCAGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAACCCTAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTTACAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCCAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTAAGTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATATCAG
AGCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCT
ACTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAG
CTCAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCT
CTTCAGTGGCTTCTTTGTTGTCTTCTGTCTGTTCTGCCATATGCACTGATAAAAG
GCATTATTTTACGAAAGATCTTACCACCATCAGTTTTTGCATCGTTCTCAGGATG
GCCGTCACAAGACAGTTCCTTGGGCTGTGCAGACCTGGTACGATTCTTGGGGG

CCATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAAT
ACAACCTCACCACCCTGAAGTGGTCATGGAAAATGTGACAGCCTTTTGGGAGG
AGGGTTTTGGAGAATTGTTGAGAAGGCAAAGCAGAATAACAACAACAGGAAG
ACGAGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCACTGCTCGGGACCC
CTGTGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGC
AGGCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTG
GAGCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAG
TTTTCTGGATCATGCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGTCT
ATGATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACA
TCTCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCT
TTCTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGC
AGACCTCTACTTGTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAA
GAAATTTTTGAAAGCTGTGTGTGCAAACCTGATGGCAAACAAGACCAGGATTCTTG
TCACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATG
AAGGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGA
CTTCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGA
AGAAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCCC
CAGTTTCTTGGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTG
GTGAAAAGAGGAAAAATTCAATTCTCAATCCAATTAACAGTATTCGCAAGTTCA
GCATTGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACG
AGCCGCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCA
TCCTGCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCG
GCAGAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCAC
AGGAAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCCTCAGGCCAATCTGA
CCGAGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTTGAAATAT
CTGAAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGA
GCATCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTACTGTCCACAA
GAGCCTCATATTTGTCCTCATCTGGTGCCTGGTTATTTTCTCGCTGAGGTGGCGG
CCAGTCTTGTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAA
TAGTACACACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGC
TACTACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTT
CTTCCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCAC
CATAAAATGCTTCATTCTGTCTGCAGGCACCCATGAGCACTTTGAACACATTGA

AGGCTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCT
CCTCCCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCA
TAGCAGTGGTTGCTGTCCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATT
GTTGCCTTTATTATGCTCAGAGCTTACTTCCTGCAAACCTTCTCAACAGCTCAAACA
GCTAGAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAG
GGACTGTGGACTCTGAGAGCATTGCGGGCGACAGCCTTACTTTGAGACACTGTTCC
ACAAGGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCG
CTGGTTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCT
TCATTTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGGAATCATCCTCA
CGCTGGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGA
TGTGGATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCTA
CTGAGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCA
AGGTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGCG
GGGGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATG
CCATTTTGGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCT
GGGTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTG
AATACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCTG
CAGCAGTGGAGAAAAGCATTGTTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTCTG
GCACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCT
GGAAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTCCTGGCA
AGCTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACA
GCTGATGTGCCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATG
AGCCCAGTGCTCACCTTGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAA
GCAAGCTTTTGGCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATG
CTGGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGAC
AGCATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCAT
CTGACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCC
CCAGATCGCTGCCCTCAAGGAGGAACTGAGGAAGAGGTGCAGGATACCCGCCT
GTGA (SEQ ID NO: 23)

ATGCAGAGGAGCCCACTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCAGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG

GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAAAAACCCCAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTTACAAAAAGACCCTG
AACTTTCTCAAGAGTGTTAGATAAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATATCAG
AGCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCT
ACTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAG
CTCAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCT
CTTCAGTGGCTTCTTTGTTGTCTTCCTGTCTGTTCTGCCATATGCACTGATAAAAG
GCATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATG
GCCGTCACAAGACAGTTCCCCTGGGCTGTGCAGACCTGGTACGATTCCTTGGGGG
CCATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAAT
ACAACCTCACCACCACTGAAGTGGTCATGGAAAATGTGACAGCCTTTTGGGAGG
AGGGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAG
ACGAGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCACTGCTCGGGACCC
CTGTGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGC
AGGCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTG
GAGCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAG
TTTTCTGGATCATGCCCCGCACCATTAAGGAAAACATCATATTTGGAGTGTCT
ATGATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACA
TCTCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCT
TTCTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGC
AGACCTCTACTTGTTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAA
GAAATTTTTGAAAGCTGTGTGTGCAAACTGATGGCAAACAAGACCAGGATTCTTG
TCACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATG
AAGGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGA
CTTCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGA

AGAAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCCC
CAGTTTCTTGGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTG
GTGAAAAGAGGAAAAATTCAATTCTCAATCCAATTAACAGTATTCGCAAGTTCA
GCATTGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACG
AGCCGCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCA
TCCTGCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCG
GCAGAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCAC
AGGAAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCCTCAGGCCAATCTGA
CCGAGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTGGAAATAT
CTGAAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGA
GCATCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTACTGTCCACAA
GAGCCTCATATTTGTCCTCATCTGGTGCCTGGTTATTTTCCTCGCTGAGGTGGCGG
CCAGTCTTGTTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAA
TAGTACTCACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGC
TACTACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTT
CTCCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCAC
CATAAAATGCTTCATTCTGTCTGCAGGCACCCATGAGCACTTTGAACACATTGA
AGGCTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCT
CCTCCCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCA
TAGCAGTGGTTGCTGTCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATT
GTTGCCTTTATTATGCTCAGAGCTTACTTCCTGCAAACCTTCTCAACAGCTCAAACA
GCTAGAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAG
GGACTGTGGACTCTGAGAGCATTTCGGGCGACAGCCTTACTTTGAGACACTGTTCC
ACAAGGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCG
CTGGTTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCT
TCATTTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGGGAATCATCCTCA
CGCTGGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGA
TGTGGATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCA
ACTGAGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGC
AAGGTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGC
GGGGGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAAT
GCCATTTTGGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGC
TGGGTTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCT

GAATACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCTT
GCAGCAGTGGAGAAAAGCATTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTCT
GGCACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATC
TGGAAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTCCTGGC
AAGCTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAAC
AGCTGATGTGCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGAT
GAGCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAA
AGCAAGCTTTTGCCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAAT
GCTGGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGA
CAGCATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCA
TCTGACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGC
CCCAGATCGCTGCCCTCAAGGAGGAAACTGAGGAAGAGGTGCAGGATACCCGCC
TGTGA (SEQ ID NO: 24)

ATGCAGAGGAGCCCCTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCAGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAACCCTAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTTACAAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCACAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATATCAG
AGCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCT
ACTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAG
CTCAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCTT
CTTCAGTGGCTTCTTTGTTGTCTTCTGCTGTTCTGCCATATGCACTGATAAAAG
GCATTATTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATG

GCCGTCACAAGACAGTTCCCCTGGGCTGTGCAGACCTGGTACGATTCCCTTGGGGG
CCATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAAT
ACAACCTCACCACCACTGAAGTGGTCATGGAAAATGTGACAGCCTTTTGGGAGG
AGGGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAG
ACGAGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCTACTGCTCGGGACCC
CTGTGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGC
AGGCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTG
GAGCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAG
TTTTCTGGATCATGCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGTCT
ATGATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACA
TCTCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCT
TTCTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGC
AGACCTCTACTTGTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAA
GAAATTTTTGAAAGCTGTGTGTGCAAACACTGATGGCAAACAAGACCAGGATTCTTG
TCACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATG
AAGGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGA
CTTCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGA
AGAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCCC
CAGTTTCTTGGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTG
GTGAAAAGAGGAAAAATTCAATTCTCAATCCTATTAACAGTATTCGCAAGTTCAG
CATTGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACGA
GCCGCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCAT
CCTGCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCGG
CAGAGTGTTTTAAATCTCATGACCATTAGTGAACCAGGGCCAAAATATCCACA
GGAAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCTCAGGCCAATCTGAC
CGAGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTTGAAATATC
TGAAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGAG
CATCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTAATGTCCACAAG
AGCCTCATATTTGTCTCATCTGGTGCCTGGTTATTTTCTCGCTGAGGTGGCGGC
CAGTCTTGTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAAT
AGTACTCACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGCT
ACTACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTT
CTTCCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCAC

CATAAAATGCTTCATTCTGTCCTGCAGGCACCCATGAGCACTTTGAACACATTGA
AGGCTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCT
CCTCCCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCA
TAGCAGTGGTTGCTGTCCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATT
GTTGCCTTTATTATGCTCAGAGCTTACTTCCTGCAAACCTTCTCAACAGCTCAAACA
GCTAGAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAG
GGACTGTGGACTCTGAGAGCATTGCGGGCGACAGCCTTACTTTGAGACACTGTTCC
ACAAGGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCG
CTGGTTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCT
TCATTTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGGAATCATCCTCA
CGCTGGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGA
TGTGGATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCTA
CTGAGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCA
AGGTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGCG
GGGGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATG
CCATTTTGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCT
GGGTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTG
AATACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCCCTG
CAGCAGTGGAGAAAAGCATTGTTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTG
GCACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCT
GGAAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTTCTGGCA
AGCTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACA
GCTGATGTGCCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATG
AGCCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAA
GCAAGCTTTTGCCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATG
CTGGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGAC
AGCATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCAT
CTGACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCC
CCAGATCGCTGCCCTCAAGGAGGAACTGAGGAAGAGGTGCAGGATACCCGCCT
GTGA (SEQ ID NO: 25)

ATGCAGAGGAGCCCCTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCAGATA

TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAAAAACCCCAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTACAAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCCAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATTCAGA
GCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCTA
CTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAGCT
CAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCTTCT
TCAGTGGCTTCTTTGTTGTCTTCTGTCTGTTCTGCCATATGCACTGATAAAAGGC
ATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATGGC
CGTCACAAGACAGTTCCTTGGGCTGTGCAGACCTGGTACGATTCTTGGGGGCC
ATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAAACTTTAGAATAC
AACCTCACCACCCTGAAGTGGTCATGGAAAATGTGACAGCCTTTTTGGGAGGAG
GGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAGACG
AGCAATGGGGACGACTCTCTTCTTTCAGCAACTTTTCACTGCTCGGGACCCCTG
TGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGCAG
GCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTGGA
GCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAGTT
TTCCTGGATCATGCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGCCTAT
GATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACATC
TCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCTTT
CTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGCAG
ACCTCTACTTGTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAAGA
AATTTTTGAAAGCTGTGTGTGCAAACCTGATGGCAAACAAGACCAGGATTCTTGTC
ACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATGAA
GGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGACT

TCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGAAG
AAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCCCA
GTTTCTTGGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTGGTG
AAAAGAGGAAAAATTCAATTCTCAATCCAATTAACAGTATTCGCAAGTTCAGCAT
TGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACGAGCC
GCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCATCCT
GCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCGGCA
GAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCACAGG
AAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCTCAGGCCAATCTGACCG
AGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTGGAAATATCTG
AAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGAGCA
TCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTACTGTCCACAAGAG
CCTCATATTTGTCTCATCTGGTGCCTGGTTATTTTCTCGCTGAGGTGGCGGCCA
GTCTTGTTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAATAG
TACTCACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGCTAC
TACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTTCTT
CCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCACCAT
AAAATGCTTCATTCTGTCTGCAGGCACCCATGAGCACTTTGAACACATTGAAGG
CTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCTCCTC
CCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCATAGC
AGTGGTTGCTGTCCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATTGTTG
CCTTTATTATGCTCAGAGCTTACTTCTGCAAACCTTCTCAACAGCTCAAACAGCTA
GAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAGGGAC
TGTGGACTCTGAGAGCATTTCGGGCGACAGCCTTACTTTGAGACACTGTTCCACAA
GGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCGCTGG
TTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCTTCAT
TTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGGAATCATCCTCACGCT
GGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGATGTG
GATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCAACTG
AGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCAAG
GTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGCGGG
GGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATGCC
ATTTTGGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCTGG

GTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTGAA
TACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCCCTGCA
GCAGTGGAGAAAAGCATTTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTCTGGC
ACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCTGG
AAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTTCTGGCAAG
CTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACAGC
TGATGTGCCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATGAG
CCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAAGC
AAGCTTTTGCCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATGCT
GGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGACAG
CATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCATCT
GACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCCCC
AGATCGCTGCCCTCAAGGAGGAACTGAGGAAGAGGTGCAGGATACCCGCCTGT
GA (SEQ ID NO: 26)

ATGCAGAGGAGCCCACTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCTGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAACCCEAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGCCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTTACAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGCCAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATATCAG
AGCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCT
ACTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAG
CTCAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCT
CTTCAGTGGCTTCTTTGTTGTCTTCCTGTCTGTTCTGCCATATGCACTGATAAAAG

GCATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATG
GCCGTCACAAGACAGTTCCCCTGGGCTGTGCAGACCTGGTACGATTCCTTGGGGG
CCATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAAT
ACAACCTCACCACCACTGAAGTGGTCATGGAAAATGTGACAGCCTTTTGGGAGG
AGGGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAG
ACGAGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCACTGCTCGGGACCC
CTGTGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGC
AGGCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTG
GAGCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAG
TTTTCTGGATCATGCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGTCT
ATGATGAGTACCGCTACCGGTCCGTCATCAAAGCCTGTCAGTTGGAGGAGGACA
TCTCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCT
TTCTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGC
AGACCTCTACTTGTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAA
GAAATTTTTGAAAGCTGTGTGTGCAAACTGATGGCAAACAAGACCAGGATTCTTG
TCACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATG
AAGGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGA
CTTCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGA
AGAAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCC
CAGTTTCTTGGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTG
GTGAAAAGAGGAAAAATTCAATTCTCAATCCAATTAACAGTATTCGCAAGTTCA
GCATTGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACG
AGCCGCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCA
TCCTGCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCG
GCAGAGTGTTTTAAATCTCATGACCATTCAAGTGAACCAGGGCCAAAATATCCAC
AGGAAGACTACAGCTTCTACCCGGAAGTGTCTCTGGCCCTCAGGCCAATCTGA
CCGAGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTTGAAATAT
CTGAAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGA
GCATCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTAAGTCCACAA
GAGCCTCATATTTGTCCTCATCTGGTGCCTGGTTATTTTCTCGCTGAGGTGGCGG
CCAGTCTTGTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAA
TAGTACTCACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGC
TACTACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTT

CTTCCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCAC
CATAAAATGCTTCATTCTGTCTGCAGGCACCCATGAGCACTTTGAACACATTGA
AGGCTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCT
CCTCCCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCA
TAGCAGTGGTTGCTGTCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATT
GTTGCCTTTATTATGCTCAGAGCTTACTTCCTGCAAACCTTCTCAACAGCTCAAACA
GCTAGAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAG
GGACTGTGGACTCTGAGAGCATTCGGGCGACAGCCTTACTTTGAGACACTGTTCC
ACAAGGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCG
CTGGTTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCT
TCATTTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGGAATCATCCTCA
CGCTGGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGA
TGTGGATTCTCTAATGAGGAGTGTCTCCCGGGTGTAAATTCATTGATATGCCTA
CTGAGGGGAAACCCACCAAGTCAACAAAGCCTTATAAGAATGGACAGCTGAGCA
AGGTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCAGCG
GGGGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATG
CCATTTTGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCT
GGGTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTG
AATACAGAAGGCGAAATCCAAATTGACGGGGTGTGAGCTGGGACAGCATCACCTG
CAGCAGTGGAGAAAAGCATTGTTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTG
GCACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCT
GGAAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTCCTGGCA
AGCTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACA
GCTGATGTGCCTCGCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATG
AGCCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAA
GCAAGCTTTTGCCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATG
CTGGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGAC
AGCATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCAT
CTGACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCC
CCAGATCGCTGCCCTCAAGGAGGAACTGAGGAAGAGGTGCAGGATACCCGCCT
GTGA (SEQ ID NO: 27)

ATGCAGAGGAGCCACTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCAGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAACCCCAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTTACAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCACAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATATCAG
AGCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCT
ACTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAG
CTCAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCTT
CTTCAGTGGCTTCTTTGTTGTCTTCCTGTCTGTTCTGCCATATGCACTGATAAAAG
GCATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATG
GCCGTCACAAGACAGTTCCCCTGGGCTGTGCAGACCTGGTACGATTCCTTGGGGG
CCATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAAT
ACAACCTCACCACCCTGAAGTGGTCATGGAAAATGTGACAGCCTTTTGGGAGG
AGGGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAG
ACGAGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCACTGCTCGGGACCC
CTGTGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGC
AGGCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTG
GAGCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAG
TTTTCTGGATCATGCCCCGCACCATTAAGGAAAACATCATATTTGGAGTGTCT
ATGATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACA
TCTCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCT
TTCTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGC
AGACCTCTACTTGTTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAA
GAAATTTTTGAAAGCTGTGTGTGCAAACTGATGGCAAACAAGACCAGGATTCTTG

TCACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATG
AAGGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGA
CTTCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGA
AGAAACTCTATACTCACAGAGACCTCCACCGCTTCTCCCTTGAGGGAGATGCCC
CAGTTTCTTGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTG
GTGAAAAGAGGAAAAATTCAATTCTCAATCCTATTAACAGTATTCGCAAGTTCAG
CATTGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACGA
GCCGCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCAT
CCTGCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCGG
CAGAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCACA
GGAAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCTCAGGCCAATCTGAC
CGAGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTTGAATATC
TGAAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGAG
CATCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATACTGTCCACAAG
AGCCTCATATTTGTCTCATCTGGTGCCTGGTTATTTTCCTCGCTGAGGTGGCGGC
CAGTCTTGTTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAAT
AGTACACACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGCT
ACTACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTT
CTTCCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCAC
CATAAAATGCTTCATTCTGTCCTGCAGGCACCCATGAGCACTTTGAACACATTGA
AGGCTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCT
CCTCCCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCA
TAGCAGTGGTTGCTGTCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATT
GTTGCCTTTATTATGCTCAGAGCTTACTTCCTGCAAACCTTCTCAACAGCTCAAACA
GCTAGAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAG
GGACTGTGGACTCTGAGAGCATTCGGGCGACAGCCTTACTTTGAGACACTGTTCC
ACAAGGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCG
CTGGTTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCT
TCATTTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGAATCATCCTCA
CGCTGGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCAGTATAGA
TGTGGATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCTA
CTGAGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCA
AGGTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGCG

GGGGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATG
CCATTTTGGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCT
GGGTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTG
AATACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCTG
CAGCAGTGGAGAAAAGCATTTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTCTG
GCACTTTCAGAAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCT
GGAAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTTCTGGCA
AGCTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACA
GCTGATGTGCCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATG
AGCCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAA
GCAAGCTTTTGGCCACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATG
CTGGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGAC
AGCATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCAT
CTGACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCC
CCAGATCGCTGCCCTCAAGGAGGAAACTGAGGAAGAGGTGCAGGATACCCGCCT
GTGA (SEQ ID NO: 28)

ATGCAGAGGAGCCCACTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCAGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAACCCEAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTTACAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATATCAG
AGCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCT
ACTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAG

CTCAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCTT
CTTCAGTGGCTTCTTTGTTGTCTTCCTGTCTGTTCTGCCATATGCACTGATAAAAG
GCATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATG
GCCGTCACAAGACAGTTCCTCCCTGGGCTGTGCAGACCTGGTACGATTCCTTGGGGG
CCATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAAT
ACAACCTCACCACCACTGAAGTGGTCATGGAAAATGTGACAGCCTTTTGGGAGG
AGGGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAG
ACGAGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCACTGCTCGGGACCC
CTGTGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGC
AGGCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTG
GAGCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAG
TTTTCTGGATCATGCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGTCT
ATGATGAGTACCGCTACCGGTCCGTCATCAAAGCCTGTCAGTTGGAGGAGGACA
TCTCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCT
TTCTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGC
AGACCTCTACTTGTTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAA
GAAATTTTTGAAAGCTGTGTGTGCAAACCTGATGGCAAACAAGACCAGGATTCTTG
TCACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATG
AAGGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGA
CTTCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGA
AGAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCC
CAGTTTCTTGGACAGAAACCAAGAAGCAGTCTTTAAGCAGACTGGCGAGTTTG
GTGAAAAGAGGAAAAATTCAATTCTCAATCCAATTAACAGTATTTCGCAAGTTCA
GCATTGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACG
AGCCGCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCA
TCCTGCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCG
GCAGAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCAC
AGGAAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCCTCAGGCCAATCTGA
CCGAGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTGGAAATAT
CTGAAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGA
GCATCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTAAGTGTCCACAA
GAGCCTCATATTTGTCCTCATCTGGTGCCTGGTTATTTTCTCGCTGAGGTGGCGG
CCAGTCTTGTTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAA

TAGTACTCACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGC
TACTACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTT
CTTCCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCAC
CATAAAATGCTTCATTCTGTCTGCAGGCACCCATGAGCACTTTGAACACATTGA
AGGCTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCT
CCTCCCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCA
TAGCAGTGGTTGCTGTCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATT
GTTGCCTTTATTATGCTCAGAGCTTACTTCCTGCAAACCTTCTCAACAGCTCAAACA
GCTAGAGTCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAG
GGACTGTGGACTCTGAGAGCATTCCGGGCGACAGCCTTACTTTGAGACACTGTTCC
ACAAGGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCG
CTGGTTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCT
TCATTTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGGGAATCATCCTCA
CGCTGGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGA
TGTGGATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCTA
CTGAGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCA
AGGTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGCG
GGGGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATG
CCATTTTGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCT
GGGTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTG
AATACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCTG
CAGCAGTGGAGAAAAGCATTGTTGGGGTCATTCCACAGAAAGTTTTATCTTCTCTG
GCACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCT
GGAAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTTCTGGCA
AGCTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACA
GCTGATGTGCCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATG
AGCCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAA
GCAAGCTTTTGCCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATG
CTGGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGAC
AGCATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCAT
CTGACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCC
CCAGATCGCTGCCCTCAAGGAGGAACTGAGGAAGAGGTGCAGGATAACCCGCCT
GTGA (SEQ ID NO: 29)

ATGCAGAGGAGCCACTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCTGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAACCCCAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTACAAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCACAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATTCAGA
GCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCTA
CTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAGCT
CAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCTTCT
TCAGTGGCTTCTTTGTTGTCTTCCTGTCTGTTCTGCCATATGCACTGATAAAAGGC
ATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATGGC
CGTCACAAGACAGTTCCCCTGGGCTGTGCAGACCTGGTACGATTCCTTGGGGGCC
ATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAATAC
AACCTCACCACCCTGAAGTGGTCATGGAAAATGTGACAGCCTTTTTGGGAGGAG
GGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAGACG
AGCAATGGGGACGACTCTCTTCTTTCAGCAACTTTTCACTGCTCGGGACCCCTG
TGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGCAG
GCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTGGA
GCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAGTT
TTCCTGGATCATGCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGTCTAT
GATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACATC
TCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCTTT
CTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGCAG
ACCTCTACTTGTTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAAGA
AATTTTTGAAAGCTGTGTGTGCAAACACTGATGGCAAACAAGACCAGGATTCTTGTC

ACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATGAA
GGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGACT
TCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGAAG
AAACTCTATACTCACAGAGACCCTCCACCGCTTCTCCCTTGAGGGAGATGCCCA
GTTTCTTGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTGGTG
AAAAGAGGAAAAATTCAATTCTCAATCCTATTAACAGTATTCGCAAGTTCAGCAT
TGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACGAGCC
GCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCATCCT
GCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCGGCA
GAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCACAGG
AAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCTCAGGCCAATCTGACCG
AGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTGGAAATATCTG
AAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGAGCA
TCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTACTGTCCACAAGAG
CCTCATATTTGTCTCATCTGGTGCCTGGTTATTTTCCTCGCTGAGGTGGCGGCCA
GTCTTGTTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAATAG
TACACACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGCTAC
TACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTTCTT
CCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCACCAT
AAAATGCTTCATTCTGTCTGCAGGCACCCATGAGCACTTTGAACACATTGAAGG
CTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCTCCTC
CCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCATAGC
AGTGGTTGCTGTCCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATTGTTG
CCTTTATTATGCTCAGAGCTTACTTCTGCAAACCTTCTCAACAGCTCAAACAGCTA
GAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAGGGAC
TGTGGACTCTGAGAGCATTTCGGGCGACAGCCTTACTTTGAGACACTGTTCCACAA
GGCCCTGAACTTGCACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCGCTGG
TTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCTTCAT
TTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGGAATCATCCTCACGCT
GGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCCAGTATAGATGTG
GATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCTACTG
AGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCAAG
GTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGCGGG

GGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATGCC
ATTTTGGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCTGG
GTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTGAA
TACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCCCTGCA
GCAGTGGAGAAAAGCATTTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTCTGGC
ACTTTCAGAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCTGG
AAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTTCTGGCAAG
CTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACAGC
TGATGTGCCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATGAG
CCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAAGC
AAGCTTTTGCCGACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATGCT
GGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGACAG
CATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCATCT
GACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCCCC
AGATCGCTGCCCTCAAGGAGGAAACTGAGGAAGAGGTGCAGGATACCCGCCTGT
GA (SEQ ID NO: 30)

Пример 3. Дополнительные примеры кодон-оптимизированных последовательностей CFTR

[0163] Представленные ниже дополнительные примеры кодон-оптимизированных последовательностей используются для генерации человеческой мРНК CFTR для безопасного и эффективного клинического использования:

ATGCAGAGAAGCCCCCTGGAGAAGGCCTCTGTGGTGAGCAAGCTGTTCTTCAGCT
GGACCAGACCCATCCTGAGAAAGGGCTACAGACAGAGACTGGAGCTGTCTGACA
TCTACCAGATCCCCTCTGTGGACTCTGCCGACAACCTGTCTGAGAAGCTGGAGAG
AGAGTGGGACAGAGAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAATGCCCT
GAGAAGATGCTTCTTCTGGAGATTCATGTTCTATGGCATCTTCCTGTACCTGGGA
GAGGTGACCAAGGCCGTGCAGCCCCTGCTGCTGGGCAGGATCATTGCCAGCTAT
GACCCTGACAACAAGGAGGAGAGAAGCATTGCCATCTACCTGGGCATTGGCCTG
TGCTGCTGTTTATTGTGAGAACCCTGCTGCTGCACCCTGCCATCTTTGGCCTGCA
CCACATTGGCATGCAGATGAGAATTGCCATGTTTCAGCCTGATCTACAAGAAGACC

CTGAAGCTGAGCAGCAGAGTGCTGGACAAGATCAGCATTGGCCAGCTGGTGAGC
CTGCTGAGCAACAACCTGAACAAGTTTGATGAGGGCCTGGCCCTGGCCCACTTTG
TGTGGATTGCCCCCTGCAGGTGGCCCTGCTGATGGGCCTGATCTGGGAGCTGCT
GCAGGCCTCTGCCTTCTGTGGCCTGGGCTTCCTGATTGTGCTGGCCCTGTTCCAGG
CCGGCCTGGGCAGAATGATGATGAAGTACAGAGACCAGAGAGCCGGCAAGATCT
CTGAGAGACTGGTGATCACCTCTGAGATGATTGAGAACATCCAGTCTGTGAAGG
CCTACTGCTGGGAGGAGGCCATGGAGAAGATGATTGAGAACCTGAGACAGACAG
AGCTGAAGCTGACCAGGAAGGCCCTATGTGAGATACTTCAACAGCTCTGCCTT
CTTCTTCTCTGGCTTCTTTGTGGTGTTCTGTCTGTGCTGCCCTATGCCCTGATCAA
GGGCATCATCCTGAGGAAGATCTTACCACCATCAGCTTCTGCATTGTGCTGAGG
ATGGCCGTGACCAGGCAGTTCCCCTGGGCCGTGCAGACCTGGTATGACAGCCTG
GGGGCCATCAACAAGATCCAGGACTTCCTGCAGAAGCAGGAGTACAAGACCCTG
GAGTACAACCTGACCACCACAGAGGTGGTGATGGAGAATGTGACAGCCTTCTGG
GAGGAGGGCTTTGGAGAGCTGTTTGAGAAGGCCAAGCAGAACAACAACAACAG
AAAGACCAGCAATGGAGATGACAGCCTGTTCTTCAGCAACTTCAGCCTGCTGGG
CACCCCTGTGCTGAAGGACATCAACTTCAAGATTGAGAGGGGCCAGCTGCTGGC
CGTGGCCGGCAGCACAGGAGCCGGCAAGACCAGCCTGCTGATGGTGATCATGGG
AGAGCTGGAGCCCTCTGAGGGCAAGATCAAGCACTCTGGCAGAATCAGCTTCTG
CAGCCAGTTCAGCTGGATCATGCCTGGCACCATCAAGGAGAACATCATCTTTGGG
GTGAGCTATGATGAGTACAGGTACAGATCTGTGATCAAGGCCTGCCAGCTGGAG
GAGGACATCTCCAAGTTTGCCGAGAAGGACAACATTGTGCTGGGGGAGGGAGGC
ATCACCTGTCTGGGGGCCAGAGAGCCAGAATCAGCCTGGCCAGAGCCGTGTAC
AAGGATGCCGACCTGTACCTGCTGGACAGCCCCTTTGGCTACCTGGATGTGCTGA
CAGAGAAGGAGATCTTTGAGAGCTGTGTGTGCAAGCTGATGGCCAACAAGACCA
GGATCCTGGTGACCAGCAAGATGGAGCACCTGAAGAAGGCCGACAAGATCCTGA
TCCTGCATGAGGGCAGCAGCTACTTCTATGGCACCTTCTCTGAGCTGCAGAACCT
GCAGCCTGACTTCAGCAGCAAGCTGATGGGCTGTGACAGCTTTGACCAGTTCTCT
GCTGAGAGAAGAAACAGCATCCTGACAGAGACCCTGCACAGGTTTCAGCCTGGAG
GGGGATGCCCCTGTGAGCTGGACAGAGACCAAGAAGCAGAGCTTCAAGCAGACA
GGAGAGTTTGGGGAGAAGAGGAAGAAGCAGCATCCTGAACCCCATCAACAGCATC
AGGAAGTTCAGCATTGTGCAGAAGACCCCCCTGCAGATGAATGGCATTGAGGAG
GACTCTGATGAGCCCCTGGAGAGAAGACTGAGCCTGGTGCCAGACTCTGAGCAG
GGAGAGGCCATCCTGCCCAGGATCTCTGTGATCAGCACAGGCCCCACCCTGCAG

GCCAGAAGAAGACAGTCTGTGCTGAACCTGATGACCCACTCTGTGAACCAGGGC
CAGAATATCCACAGAAAGACCACAGCCAGCACCAGAAAGGTGAGCCTGGCCCC
CAGGCCAACCTGACAGAGCTGGACATCTACAGCAGAAGGCTGAGCCAGGAGAC
AGGCCTGGAGATCTCTGAGGAGATCAATGAGGAGGACCTGAAGGAGTGCTTCTT
TGATGACATGGAGAGCATCCCTGCCGTGACCACCTGGAACACCTACCTGAGATA
CATCACAGTGCACAAGAGCCTGATCTTTGTGCTGATCTGGTGCCTGGTGATCTTC
CTGGCCGAGGTGGCCGCCAGCCTGGTGGTGCTGTGGCTGCTGGGCAACACCCCC
CTGCAGGACAAGGGCAACAGCACCCACAGCAGAAACAACAGCTATGCTGTGATC
ATCACCAGCACCAGCAGCTACTATGTGTTCTACATCTATGTGGGAGTGGCTGACA
CCCTGCTGGCCATGGGCTTCTTCAGAGGCCTGCCCCTGGTGCACACCCTGATCAC
AGTGAGCAAGATCCTGCACCACAAGATGCTGCACTCTGTGCTGCAGGCCCCCAT
GAGCACCTGAACACCCTGAAGGCTGGAGGCATCCTGAACAGATTCAGCAAGGA
CATTGCCATCCTGGATGACCTGCTGCCCCTGACCATCTTTGACTTCATCCAGCTGC
TGCTGATTGTGATTGGAGCCATTGCCGTGGTGGCCGTGCTGCAGCCCTACATCTT
TGTGGCCACAGTGCCTGTGATTGTGGCCTTCATCATGCTGAGGGCCTACTTCCTG
CAGACCAGCCAGCAGCTGAAGCAGCTGGAGTCTGAGGGCAGAAGCCCCATCTTC
ACCCACCTGGTGACCAGCCTGAAGGGCCTGTGGACCCTGAGGGCCTTTGGCAGA
CAGCCCTACTTTGAGACCCTGTTCCACAAGGCCCTGAACCTGCACACAGCCAACT
GGTTCCTGTACCTGAGCACCTGAGATGGTTCAGATGAGGATTGAGATGATCTT
TGTGATCTTCTTCATTGCCGTGACCTTCATCAGCATCCTGACCACAGGGGAGGGC
GAGGGCAGAGTGGGCATCATCCTGACCCTGGCCATGAACATCATGAGCACCTG
CAGTGGGCCGTGAACAGCAGCATTGATGTGGACAGCCTGATGAGATCTGTGAGC
AGAGTGTTCAAGTTCATTGACATGCCACAGAGGGCAAGCCCACCAAGAGCACC
AAGCCCTACAAGAATGGCCAGCTGAGCAAGGTGATGATCATTGAGAACAGCCAT
GTGAAGAAGGATGACATCTGGCCCTCTGGAGGCCAGATGACAGTGAAGGACCTG
ACAGCCAAGTACACAGAGGGGGGCAATGCCATCCTGGAGAACATCAGCTTCAGC
ATCAGCCCTGGCCAGAGGGTGGGCCTGCTGGGCAGAACAGGCTCTGGCAAGAGC
ACCCTGCTGTCTGCCTTCTGAGGCTGCTGAACACAGAGGGAGAGATCCAGATTG
ATGGGGTGAGCTGGGACAGCATCACCTGCAGCAGTGGAGGAAGGCCTTTGGGG
TGATCCCCAGAAGGTGTTTCATCTTCTCTGGCACCTTCAGGAAGAACCTGGACCC
CTATGAGCAGTGGTCTGACCAGGAGATCTGGAAGGTGGCCGATGAGGTGGGCCT
GAGATCTGTGATTGAGCAGTTCCTGGCAAGCTGGACTTTGTGCTGGTGGATGGA
GGCTGTGTGCTGAGCCATGGCCACAAGCAGCTGATGTGCCTGGCCAGATCTGTGC

TGAGCAAGGCCAAGATCCTGCTGCTGGATGAGCCCTCTGCCACCTGGACCCTGT
GACCTACCAGATCATCAGAAGAACCCTGAAGCAGGCCTTTGCCGACTGCACAGT
GATCCTGTGTGAGCACAGAATTGAGGCCATGCTGGAGTGCCAGCAGTTCCTGGTG
ATTGAGGAGAACAAGGTGAGGCAGTATGACAGCATCCAGAAGCTGCTGAATGAG
AGAAGCCTGTTTCAGACAGGCCATCAGCCCCTCTGACAGAGTGAAGCTGTTCCCC
ACAGGAACAGCAGCAAGTGCAAGAGCAAGCCCCAGATTGCCGCCCTGAAGGAG
GAGACAGAGGAGGAGGTGCAGGACACCAGACTGTGA (SEQ ID NO: 31)

ATGCAGAGGAGCCCCCTGGAGAAGGCCAGCGTGGTGAGCAAGCTGTTCTTCAGC
TGGACCAGGCCATCCTGAGGAAGGGCTACAGGCAGAGGCTGGAGCTGAGCGAC
ATCTACCAGATCCCCAGCGTGGACAGCGCCGACAACCTGAGCGAGAAGCTGGAG
AGGGAGTGGGACAGGGAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAACGC
CCTGAGGAGGTGCTTCTTCTGGAGGTTTCATGTTCTACGGCATCTTCCTGTACCTGG
GCGAGGTGACCAAGGCCGTGCAGCCCCTGCTGCTGGGCAGGATCATCGCCAGCT
ACGACCCCGACAACAAGGAGGAGAGGAGCATCGCCATCTACCTGGGCATCGGCC
TGTGCCTGCTGTTTCATCGTGAGGACCCTGCTGCTGCACCCCGCCATCTTCGGCCT
GCACCACATCGGCATGCAGATGAGGATCGCCATGTTTCAGCCTGATCTACAAGAA
GACCCTGAAGCTGAGCAGCAGGGTGCTGGACAAGATCAGCATCGGCCAGCTGGT
GAGCCTGCTGAGCAACAACCTGAACAAGTTCGACGAGGGCCTGGCCCTGGCCCA
CTTCGTGTGGATCGCCCCCTGCAGGTGGCCCTGCTGATGGGCCTGATCTGGGAG
CTGCTGCAGGCCAGCGCCTTCTGCGGCCTGGGCTTCCTGATCGTGCTGGCCCTGT
TCCAGGCCGGCCTGGGCAGGATGATGATGAAGTACAGGGACCAGAGGGCCGGCA
AGATCAGCGAGAGGCTGGTGATCACCAGCGAGATGATCGAGAACATCCAGAGCG
TGAAGGCCTACTGCTGGGAGGAGGCCATGGAGAAGATGATCGAGAACCTGAGGC
AGACCGAGCTGAAGCTGACCAGGAAGGCCGCCTACGTGAGGTACTTCAACAGCA
GCGCCTTCTTCTTCAGCGGCTTCTTCGTGGTGTTTCCTGAGCGTGCTGCCCTACGCC
CTGATCAAGGGCATCATCCTGAGGAAGATCTTCACCACCATCAGCTTCTGCATCG
TGCTGAGGATGGCCGTGACCAGGCAGTTCCTGGGCCGTGCAGACCTGGTACG
ACAGCCTGGGCGCCATCAACAAGATCCAGGACTTCCTGCAGAAGCAGGAGTACA
AGACCCTGGAGTACAACCTGACCACCACCGAGGTGGTGATGGAGAACGTGACCG
CCTTCTGGGAGGAGGGCTTCGGCGAGCTGTTTCGAGAAGGCCAAGCAGAACAACA
ACAACAGGAAGACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTTCAGCC
TGCTGGGCACCCCGTGCTGAAGGACATCAACTTCAAGATCGAGAGGGGCCAGC

TGCTGGCCGTGGCCGGCAGCACCGGCGCCGGCAAGACCAGCCTGCTGATGGTGA
TCATGGGCGAGCTGGAGCCCAGCGAGGGCAAGATCAAGCACAGCGGCAGGATC
AGCTTCTGCAGCCAGTTCAGCTGGATCATGCCCGGCACCATCAAGGAGAACATC
ATCTTCGGCGTGAGCTACGACGAGTACAGGTACAGGAGCGTGATCAAGGCCTGC
CAGCTGGAGGAGGACATCAGCAAGTTCGCCGAGAAGGACAACATCGTGCTGGGC
GAGGGCGGCATCACCTGAGCGGGCGGCCAGAGGGCCAGGATCAGCCTGGCCAG
GGCCGTGTACAAGGACGCCGACCTGTACCTGCTGGACAGCCCCTTCGGCTACCTG
GACGTGCTGACCGAGAAGGAGATCTTCGAGAGCTGCGTGTGCAAGCTGATGGCC
AACAAAGACCAGGATCCTGGTGACCAGCAAGATGGAGCACCTGAAGAAGGCCGA
CAAGATCCTGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCTTCAGCGA
GCTGCAGAACCTGCAGCCCGACTTCAGCAGCAAGCTGATGGGCTGCGACAGCTT
CGACCAGTTCAGCGCCGAGAGGAGGAACAGCATCCTGACCGAGACCCTGCACAG
GTTCAGCCTGGAGGGCGACGCCCCCGTGAGCTGGACCGAGACCAAGAAGCAGAG
CTTCAAGCAGACCGGCGAGTTCGGCGAGAAGAGGAAGAACAGCATCCTGAACCC
CATCAACAGCATCAGGAAGTTCAGCATCGTGCAGAAGACCCCCCTGCAGATGAA
CGGCATCGAGGAGGACAGCGACGAGCCCCTGGAGAGGAGGCTGAGCCTGGTGC
CCGACAGCGAGCAGGGCGAGGCCATCCTGCCAGGATCAGCGTGATCAGCACCG
GCCCCACCCTGCAGGCCAGGAGGAGGCAGAGCGTGCTGAACCTGATGACCCACA
GCGTGAACCAGGGCCAGAACATCCACAGGAAGACCACCGCCAGCACCAAGGAA
GTGAGCCTGGCCCCCAGGCCAACCTGACCGAGCTGGACATCTACAGCAGGAGG
CTGAGCCAGGAGACCGGCCTGGAGATCAGCGAGGAGATCAACGAGGAGGACCT
GAAGGAGTGCTTCTTCGACGACATGGAGAGCATCCCCGCCGTGACCACCTGGAA
CACCTACCTGAGGTACATCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGG
TGCTGGTGTGATCTTCTGGCCGAGGTGGCCGCCAGCCTGGTGGTGTGCTGTGGCTGC
TGGGCAACACCCCCCTGCAGGACAAGGGCAACAGCACCCACAGCAGGAACAAC
AGCTACGCCGTGATCATCACAGCACCAGCAGCTACTACGTGTTCTACATCTACG
TGGGCGTGGCCGACACCCTGCTGGCCATGGGCTTCTTCAGGGGCCTGCCCTGGT
GCACACCCTGATCACCGTGAGCAAGATCCTGCACCACAAGATGCTGCACAGCGT
GCTGCAGGCCCCCATGAGCACCCCTGAACACCCTGAAGGCCGGCGGCATCCTGAA
CAGGTTCAAGCAAGGACATCGCCATCCTGGACGACCTGCTGCCCTGACCATCTTC
GACTTCATCCAGCTGCTGCTGATCGTGATCGGCGCCATCGCCGTGGTGGCCGTGC
TGCAGCCCTACATCTTCGTGGCCACCGTGCCCGTGATCGTGGCCTTCATCATGCT
GAGGGCCTACTTCTGCAGACCAGCCAGCAGCTGAAGCAGCTGGAGAGCGAGGG

CAGGAGCCCCATCTTCACCCACCTGGTGACCAGCCTGAAGGGCCTGTGGACCCTG
AGGGCCTTCGGCAGGCAGCCCTACTTCGAGACCCTGTTCCACAAGGCCCTGAACC
TGCACACCGCCA ACTGGTTCCTGTACCTGAGCACCTGAGGTGGTTCAGATGAG
GATCGAGATGATCTTCGTGATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTGA
CCACCGGCGAGGGCGAGGGCAGGGTGGGCATCATCCTGACCCTGGCCATGAACA
TCATGAGCACCTGCAGTGGGCCGTGAACAGCAGCATCGACGTGGACAGCCTGA
TGAGGAGCGTGAGCAGGGTGTCAAGTTCATCGACATGCCACCGAGGGCAAGC
CCACCAAGAGCACCAAGCCCTACAAGAACGGCCAGCTGAGCAAGGTGATGATCA
TCGAGAACAGCCACGTGAAGAAGGACGACATCTGGCCCAGCGGGCCAGATG
ACCGTGAAGGACCTGACCGCCAAGTACACCGAGGGCGGCAACGCCATCCTGGAG
AACATCAGCTTCAGCATCAGCCCCGGCCAGAGGGTGGGCCTGCTGGGCAGGACC
GGCAGCGGCAAGAGCACCTGCTGAGCGCCTTCCTGAGGCTGCTGAACACCGAG
GGCGAGATCCAGATCGACGGCGTGAGCTGGGACAGCATCACCTGCAGCAGTGG
AGGAAGGCCTTCGGCGTGATCCCCAGAAGGTGTTTCATCTTCAGCGGCACCTTCA
GGAAGAACCTGGACCCCTACGAGCAGTGGAGCGACCAGGAGATCTGGAAGGTG
GCCGACGAGGTGGGCCTGAGGAGCGTGATCGAGCAGTTCCCCGGCAAGCTGGAC
TTCGTGCTGGTGGACGGCGGCTGCGTGCTGAGCCACGGCCACAAGCAGCTGATG
TGCTGGCCAGGAGCGTGCTGAGCAAGGCCAAGATCCTGCTGCTGGACGAGCCC
AGCGCCACCTGGACCCCGTGACCTACCAGATCATCAGGAGGACCCTGAAGCAG
GCCTTCGCCGACTGCACCGTGATCCTGTGCGAGCACAGGATCGAGGCCATGCTG
GAGTGCCAGCAGTTCCTGGTGATCGAGGAGAACAAGGTGAGGCAGTACGACAGC
ATCCAGAAGCTGCTGAACGAGAGGAGCCTGTTTCAGGCAGGCCATCAGCCCCAGC
GACAGGGTGAAGCTGTTCCCCACAGGAACAGCAGCAAGTGCAAGAGCAAGCCC
CAGATCGCCGCCCTGAAGGAGGAGACCGAGGAGGAGGTGCAGGACACCAGGCT
GTGA (SEQ ID NO: 32)

ATGCAGAGATCCCCTCTGGAGAAGGCCTCAGTGGTGTCCAAGCTTTTCTTCTCCT
GGACCAGGCCCATTTTAAGAAAGGGCTACAGGCAGAGACTTGAGCTGTCTGACA
TCTATCAGATCCCTTCTGTGGATTCTGCTGACAATCTTAGTGAAAAATTGGAAAG
GGAGTGGGACAGAGAGCTGGCAAGTAAAAAGAACCCCAAGCTGATTAATGCCCT
GAGGCGCTGCTTTTTTTGGAGATTCATGTTCTATGGCATATTCCTCTACCTTGGAG
AAGTAACCAAGCTGTACAGCCTCTCCTCCTTGGCAGAATCATTGCCTCCTATGA
TCCTGATAACAAGGAGGAGAGAAGCATAGCCATCTACCTGGGCATTGGGCTGTG

CCTCTTGTTTATTGTGAGGACCCTTCTCTTGCACCCTGCCATCTTTGGCCTTCATC
ACATTGGCATGCAAATGAGAATAGCAATGTTTAGTCTTATTTACAAAAAACATT
AAAACCTCTCTCCAGGGTGTGGACAAGATCAGTATTGGACAACCTGGTCAGCCTG
CTGAGCAACAACCTGAACAAGTTTGATGAAGGACTGGCCCTGGCCCACTTTGTCT
GGATTGCCCCCTTCAGGTGGCTCTTTTGATGGGCCTGATCTGGGAACTCCTGCA
GGCCTCTGCCTTCTGTGGGTTAGGCTTCCTGATAGTGCTAGCTCTCTTTCAGGCAG
GGTTGGGTAGAATGATGATGAAGTACAGAGACCAGAGGGCTGGGAAGATATCTG
AGAGGCTGGTCATTACTTCTGAAATGATAGAAAACATCCAGTCTGTAAAGCTTA
CTGCTGGGAGGAGGCTATGGAAAAGATGATTGAGAACTTGAGGCCAACAGAGCT
CAAGCTGACTAGGAAGGCAGCCTATGTCAGGTATTTCAACAGCAGTGCTTTCTTC
TTCTCAGGCTTTTTTCGTGGTCTTCTTGAGTGTCTGCCCTATGCCCTCATCAAGGG
GATAATTTTGAGAAAGATTTTCACCACTATTTCTTTTGCATTGTCCTGAGGATGG
CTGTCACCAGGCAATTCCCCTGGGCTGTGCAGACATGGTATGACTCTCTGGGGGC
CATCAACAAAATCCAAGATTTCTGCAGAAGCAGGAGTACAAGACCCTGGAATA
CAACCTCACCACCACAGAAGTTGTGATGGAGAATGTGACTGCATTCTGGGAGGA
AGGATTTGGGGAGCTGTTTGAGAAAGCAAACAAAACAATAATAACAGGAAAA
CCAGCAATGGAGATGACTCCCTGTTCTTTTCCAACCTTCTCTTTGTTGGGCACCCCT
GTCCTGAAAGATATAAACTTTAAAATTGAAAGAGGGCAGCTGTTGGCAGTTGCT
GGCTCCACAGGAGCTGGAAAACTTCACTACTGATGGTGATCATGGGGGAGTTA
GAACCCTCTGAAGGGAAAATAAAACATTCTGGGAGGATTAGTTTCTGCAGCCAG
TTCAGCTGGATCATGCCTGGGACCATTAAGAAAATATTATATTTGGAGTGAGCT
ATGATGAATATAGATATAGGAGTGTCATCAAAGCCTGTCAGTTGGAGGAAGACA
TCAGCAAATTTGCAGAGAAAGACAACATTGTTCTGGGTGAAGGTGGCATCACCC
TGTCAGGAGGGCAAAGGGCCAGGATCAGCTTGCCAGAGCAGTCTATAAAGATG
CTGATCTGTACCTCCTGGATAGCCCTTTTGGCTATCTGGATGTTTTGACAGAGAA
GGAAATTTTTGAGTCCTGTGTCTGCAAGTTAATGGCAAATAAAACAAGGATACTT
GTGACCTCAAAAATGGAACACCTGAAGAAGGCTGACAAAATTCTGATCCTGCAT
GAGGGCAGCAGCTACTTTTATGGAACATTTTCTGAACTGCAGAATTTGCAACCAG
ACTTTTCATCAAAGCTCATGGGATGTGACAGTTTTGATCAGTTTTCTGCAGAAAG
GAGAACTCCATTTTGACTGAGACCCTGCACAGGTTTCAGTCTGGAGGGGGATGC
CCCAGTGAGTTGGACTGAGACAAAGAAACAGAGCTTCAAGCAGACTGGAGAGTT
TGGAGAAAAGAGGAAAACTCAATTCTCAATCCCATCAATAGCATCAGGAAGTT
CAGCATAGTTCAGAAGACTCCTTTGCAGATGAATGGGATTGAAGAGGACTCAGA

TGAGCCCCTGGAAAGGAGACTCTCCTTGGTGCCAGATTCAGAGCAGGGGGAAGC
CATACTGCCAAGGATCTCTGTGATTTCTACAGGGCCCACCCTCCAAGCAAGAAGG
AGACAGTCAGTTTTAAACCTGATGACCCACTCTGTCAACCAGGGACAGAACATTC
ATAGAAAGACAACAGCATCTACAAGAAAAGTTTCACTGGCCCCTCAAGCCAATT
TAACTGAACTAGATATCTACAGCAGGAGGCTCAGCCAAGAAACAGGCCTGGAGA
TCTCAGAAGAAATAAATGAGGAGGATTTGAAGGAATGCTTCTTTGATGATATGG
AGAGCATCCCAGCTGTCACAACCTGGAACACCTACCTGAGATACATCACAGTGC
ACAAATCCCTCATCTTTGTACTIONATATGGTGCCCTTGTCATCTTCTTAGCTGAGGTG
GCTGCTTCCCTGGTGGTGCTGTGGCTGCTGGGAAACACACCCCTCCAGGATAAAG
GGA ACTCTACTCACAGCAGGAACAACAGTTATGCTGTGATCATCACCAGTACCTC
CTCCTACTATGTGTTCTACATTTATGTTGGAGTTGCAGACACATTGCTTGCCATGG
GTTTTTTTAGAGGACTCCCCCTGGTGCATACTCTCATCACTGTTTCCAAAATCCTT
CACCACAAGATGCTGCACAGTGTACTACAGGCTCCCATGAGCACCCCTCAACACTC
TTAAAGCAGGAGGAATCTTGAACAGATTTAGCAAGGACATTGCAATTCTTGATG
ACCTGCTTCCACTGACCATCTTTGACTTCATCCAGCTTCTGCTCATTGTAATTGGT
GCCATTGCTGTGGTAGCAGTGCTCCAGCCATATATTTTTGTGGCCACTGTGCCTGT
TATTGTGGCCTTCATTATGTTGAGAGCCTACTTCCTGCAGACCTCTCAGCAGCTCA
AGCAACTTGAAAGTGAGGGCAGGAGCCCCATATTTACACACTTGGTCACTTCCCT
CAAAGGCCTCTGGACACTCAGAGCTTTTGGAAGACAACCTTATTTTGAACTCTC
TTCCACAAGGCTCTGAATCTCCACACAGCCAACTGGTTTCTGTATCTTTCAACACT
GCGCTGGTTCCAGATGAGGATTGAGATGATCTTTGTTATCTTCTTCATAGCTGTTA
CCTTCATCTCTATTCTGACA ACTGGTGAGGGGGAAGGGAGAGTAGGCATCATCCT
CACACTAGCCATGAACATAATGTCTACCTTACAATGGGCCGTGAACAGCTCCATA
GATGTGGACAGCCTCATGAGAAGTGTGTCAAGAGTTTTCAAATTCATTGACATGC
CCACAGAAGGCAAACCAACCAAGAGCACAAAACCCTACAAGAATGGCCAGCTG
AGTAAGGTCATGATCATTGAAAATTCTCATGTGAAGAAGGATGATATTTGGCCCA
GTGGGGGCCAGATGACAGTCAAGGACCTCACTGCCAAATACACAGAGGGTGGAA
ATGCTATCCTAGAGAACATCTCCTTCTCCATCTCCCCAGGCCAAAGAGTTGGCTT
GCTGGGCAGGACTGGCAGTGGCAAGTCCACCTTGCTCTCAGCATTCTCAGGCTT
TTAAATACAGAGGGAGAGATTCAAATTGATGGGGTGTCTTGGGATAGTATAACA
CTTCAACAGTGGAGGAAAGCCTTTGGTGTGATTCCTCAGAAAGTGTTTATCTTCT
CTGGCACTTTCAGAAAAAATCTGGACCCCTATGAACAGTGGAGTGACCAGGAAA
TCTGGAAGGTGGCAGATGAAGTGGGCCTAAGATCAGTCATAGAGCAGTTTCCTG

GAAAGTTGGATTTTGTGCTTGTAGATGGAGGCTGTGTGCTGTCCCATGGCCATAA
ACAGCTAATGTGCCTGGCTAGGTCAGTGCTGAGCAAGGCCAAGATCCTGCTGTTA
GATGAGCCTTCAGCCCATCTGGACCCTGTGACATAACCAGATTATCAGAAGAATC
TGAAGCAGGCCTTTGCTGACTGCACTGTCATCCTGTGTGAGCACAGAATTGAGGC
CATGCTGGAGTGCCAGCAGTTCCTTGTATAGAAGAGAATAAGGTTAGGCAGTAT
GACAGCATTTCAGAACTGCTAAATGAAAGATCTCTCTTCAGGCAAGCTATTTTAC
CATCTGATAGAGTGAACTTTTTCCCCACAGAAATTCCTCTAAATGTAAATCTAA
GCCCCAGATAGCTGCCTTGAAAGAGGAGACTGAAGAAGAAGTCCAGGACACCAG
ACTGTGA (SEQ ID NO: 33)

ATGCAGAGATCCCCGCTGGAGAAGGCATCTGTGGTGTCAAACTGTTCTTTAGCT
GGACAAGGCCATCCTTAGGAAAGGGTACAGACAGAGGTTGGAGCTGTCAGACA
TATATCAGATCCCTTCAGTGGACTCTGCAGACAACCTCTCTGAAAAGCTGGAGAG
GGAATGGGACAGGGAAGTGGCCAGCAAAAAAAAAACCCTAACTGATTAATGCCCT
GAGGAGGTGCTTCTTTTGGAGATTCATGTTCTATGGGATCTTCTTTACCTGGGGG
AGGTGACTAAAGCTGTTTCAGCCTCTTCTTCTGGGGAGGATTATTGCCTCCTATGA
CCCAGACAACAAAGAAGAAAGAAGCATAGCCATTTACTTAGGCATAGGCCTCTG
CTTGCTCTTCATAGTTAGAACCCTCCTACTCCACCCAGCCATCTTTGGTCTCCACC
ACATAGGTATGCAGATGAGAATAGCAATGTTCTCCTTGATCTACAAGAAGACCCT
CAAGCTGTCCAGCAGGGTGTGGACAAGATCTCCATAGGCCAGTTAGTCAGTCTA
CTGTCCAATAACTTAAATAAGTTTGATGAGGGACTGGCACTGGCACATTTTGTGT
GGATTGCCCCCTCCAAGTGGCCCTTCTTATGGGCCTTATCTGGGAGCTGTTGCA
GGCCTCTGCTTTCTGTGGCCTGGGTTTCCTCATAGTCCTAGCCTTATTCCAGGCTG
GACTGGGCAGAATGATGATGAAGTATAGGGACCAAAGAGCAGGGAAGATTTCTG
AAAGGCTGGTTATAACTTCTGAGATGATTGAGAACATTCAGTCAGTGAAAGCTTA
CTGCTGGGAAGAAGCTATGGAAAAATGATTGAAAATCTCAGACAGACTGAATT
AAAGTTGACCAGGAAAGCTGCTTATGTCAGATACTTCAACTCCTCAGCCTTCTTT
TTTTCTGGCTTCTTTGTTGTATTCCTTTCAGTCCTCCCCTATGCCCTGATTAAGGGC
ATTATCTTGAGGAAAATTTTACAACCATCTCCTTTTGTATTGTCTCAGGATGGC
TGTTACAAGGCAATTTCTTGGGCTGTGCAAACTTGGTATGATAGCCTTGGAGCA
ATCAACAAGATCCAGGATTTCTGCAAAAAGCAGGAGTACAAGACATTGGAATAC
AACCTTACCACCACTGAGGTGGTGTGATGGAAAATGTGACTGCCTTCTGGGAGGAG
GGGTTTGGAGAGCTGTTTGGAGAAAGCCAAACAGAACAACAACAATAGAAAGACC

TCTAATGGTGATGATTCCCTGTTCTTTTCTAACTTTAGTCTTCTGGGGACCCCAGT
TCTGAAAGATATTAACCTTTAAAATTGAAAGGGGACAGTTGCTGGCTGTGGCTGGG
TCCACTGGGGCTGGGAAGACAAGCCTGCTCATGGTGATCATGGGAGAGCTGGAA
CCCAGTGAAGGAAAGATCAAACACTCAGGCAGGATCTCCTTCTGCAGCCAGTTCT
CATGGATTATGCCAGGCACTATTAAGAAAATATCATCTTTGGTGTAAGCTATGA
TGAGTACAGGTATAGATCTGTAATTAAGCCTGCCAGCTGGAGGAAGACATCTCT
AAGTTTGCTGAGAAGGATAACATTGTGTTGGGGGAAGGGGGCATCACCCCTTCTG
GTGGGCAGAGGGCTAGGATCTCCCTTGCTAGGGCAGTATAACAAGGATGCTGACT
TGTACCTCTTGGATAGTCCTTTTGGCTACCTAGATGTGCTGACAGAGAAAGAAAT
ATTTGAAAGCTGTGTGTGTAAGCTCATGGCTAACAAGACCAGGATCCTGGTCACC
AGTAAAATGGAACACCTCAAAAAGCAGACAAGATCCTTATTCTCCATGAGGGC
TCCTCCTACTTCTATGGGACCTTCAGTGAGCTGCAGAATCTGCAGCCAGACTTCT
CCTCAAACTTATGGGCTGTGACTCCTTTGACCAATTCTCTGCAGAAAGAAGGAA
TAGCATACTGACAGAAACACTGCATAGATTCTCCCTGGAAGGAGATGCCCCAGT
GAGTTGGACAGAAACCAAAAAGCAGAGCTTCAAGCAGACTGGTGAGTTTGGTGA
AAAGAGGAAGAATTCTATCCTGAACCCCATCAATAGCATCAGGAAATTTAGCAT
AGTCCAAAAGACCCCCCTCCAGATGAATGGAATAGAGGAGGATAGTGATGAGCC
TCTTGAGAGAAGGCTGTCCCTGGTTCCAGACAGTGAACAGGGTGAAGCCATTCTT
CCGAGGATCAGTGTCATCTCCACTGGGCCACATTGCAGGCCAGAAGAAGACAG
TCTGTTCTGAATTTGATGACACATTCTGTGAATCAAGGCCAGAATATCCATAGAA
AAACCACTGCCAGCACCAGAAAAGTTTCTCTAGCCCCCAGGCTAACCTGACTG
AGTTAGACATCTACAGCAGAAGGCTGAGCCAAGAGACTGGCTTGGAAATATCTG
AGGAGATCAATGAGGAGGACCTCAAGGAGTGCTTCTTTGATGACATGGAGTCAA
TCCCTGCAGTCACTACATGGAACACTTACCTAAGGTACATCACAGTTCATAAGAG
CCTCATCTTTGTCCTCATATGGTGTCTGGTCATCTTTTTAGCAGAAGTGGCTGCCA
GCCTAGTTGTGCTGTGGTTACTGGGCAATACACCTCTTCAGGACAAAGGCAATAG
CACACACAGCAGAAACAACCTCCTATGCAGTGATCATCACCTCTACAAGCTCTTAC
TATGTATTCTATATATATGTGGGAGTGGCAGATACTCTCCTGGCCATGGGATTCTT
CAGGGGATTACCTCTAGTTCACACATTGATCACAGTGTCAAAAATTCTCCACCAC
AAGATGTTACACAGTGTCTGCAAGCCCCAATGTCTACTCTGAACACACTTAAGG
CAGGTGGAATTTTGAATAGGTTTAGCAAGGACATAGCTATCCTGGATGATCTCCT
CCCTCTGACCATCTTTGACTTCATCCAGTTACTGCTCATTGTAATTGGAGCCATTG
CAGTGGTAGCAGTCCTACAGCCTTACATTTTTGTGGCTACTGTTCCCTGTTATTGTG

GCCTTCATTATGCTAAGAGCTTACTTCCTGCAAACAAGCCAACAGTTGAAACAGC
TAGAAAGTGAGGGAAGGTCCCCATCTTCACCCACCTGGTGACATCACTCAAGG
GGCTATGGACTCTTAGGGCTTTTGGGAGACAGCCGTACTTTGAGACCTTATTCCA
TAAGGCCCTTAACCTCCATACAGCAAACCTGGTTCTTATACCTGAGTACTCTGAGG
TGGTTTCAAATGAGGATTGAAATGATTTTTGTGATCTTCTTCATTGCTGTGACCT
CATCTCAATCTTGACCACAGGAGAGGGGGAGGGCAGGGTGGGCATCATACTGAC
CTTGGCCATGAACATTATGTCAACCCTGCAGTGGGCTGTCAATAGCTCCATTGAT
GTGGACAGTCTGATGAGGAGTGTCTCCAGGGTCTTCAAGTTTATTGACATGCCAA
CTGAGGGCAAACCCACCAAAAAGCACTAAGCCATATAAAAATGGCCAACTGTCCA
AAGTGATGATCATTGAAAATTCACATGTAAAGAAGGATGATATCTGGCCCTCTGG
AGGACAGATGACAGTGAAAGACCTGACTGCCAAGTACACAGAGGGTGGTAATGC
CATTCTTGAGAACATTAGTTTCAGTATTTCCCCGGGGCAAAGGGTGGGCCTCCTT
GGCAGAACAGGCTCTGGCAAGAGTACCCTGCTGTCAGCCTTTTTAAGACTGTTGA
ACACTGAGGGAGAAATTCAGATTGATGGTGTCTCCTGGGATAGCATCACCCCTCA
GCAGTGGAGAAAAGCTTTTGGAGTGATCCCGCAAAGGTTTTTCATCTTTTCAGGC
ACCTTCCGGAAGAACCTGGACCCCTATGAGCAGTGGTCTGACCAGGAAATATGG
AAGGTAGCTGATGAAGTTGGGCTTAGGTCAGTCATAGAGCAGTTCCCAGGCAA
CTGGACTTTGTCTTGGTGGATGGTGGATGTGTACTGAGTCATGGGCACAAACAGC
TGATGTGCCTAGCCAGGTCTGTGCTCAGCAAGGCAAAGATATTGCTGCTTGATGA
ACCCAGTGCCCATCTGGACCCAGTCACATATCAGATCATCAGAAGAACATTGAA
GCAGGCCTTTGCTGATTGCACAGTTATCCTCTGTGAGCACAGGATTGAGGCCATG
CTGGAGTGCCAGCAGTTTCTGGTGATTGAGGAGAATAAAGTAAGGCAGTATGAC
TCCATCCAGAAGCTGCTCAATGAAAGAAGCCTCTTTAGACAAGCTATCTCCCCCT
CAGACAGGGTCAAATTGTTCCCTCACAGAAACAGCAGCAAGTGCAAGAGCAAGC
CCCAAATTGCAGCCTTGAAAGAGGAGACAGAGGAAGAGGTGCAGGACACCAGA
CTCTGA (SEQ ID NO: 34)

ATGCAGAGAAGCCCCCTGGAGAAGGCCAGCGTGGTGAGCAAGCTGTTCTTCAGC
TGGACCAGACCCATCCTGAGAAAGGGCTACAGACAGAGACTGGAGCTGAGCGAC
ATCTACCAGATCCCAGCGTGGACAGCGCCGACAACCTGAGCGAGAAGCTGGAG
AGAGAGTGGGACAGAGAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAACGC
CCTGAGAAGATGCTTCTTCTGGAGATTCATGTTCTACGGCATCTTCTGTACCTGG
GCGAGGTGACCAAGGCCGTGCAGCCCCTGCTGCTGGGCAGAATCATCGCCAGCT

ACGACCCCGACAACAAGGAGGAGAGAAGCATCGCCATCTACCTGGGCATCGGCC
TGTGCCTGCTGTTTCATCGTGAGAACCCTGCTGCTGCACCCCGCCATCTTCGGCCT
GCACCACATCGGCATGCAGATGAGAATCGCCATGTTTCAGCCTGATCTACAAGAA
GACCCTGAAGCTGAGCAGCAGAGTGCTGGACAAGATCAGCATCGGCCAGCTGGT
GAGCCTGCTGAGCAACAACCTGAACAAGTTCGACGAGGGCCTGGCCCTGGCCCA
CTTCGTGTGGATCGCCCCCTGCAGGTGGCCCTGCTGATGGGCCTGATCTGGGAG
CTGCTGCAGGCCAGCGCCTTCTGCGGCCTGGGCTTCCTGATCGTGCTGGCCCTGT
TCCAGGCCGGCCTGGGCAGAATGATGATGAAGTACAGAGACCAGAGAGCCGGCA
AGATCAGCGAGAGACTGGTGATCACCAGCGAGATGATCGAGAACATCCAGAGCG
TGAAGGCCTACTGCTGGGAGGAGGCCATGGAGAAGATGATCGAGAACCTGAGAC
AGACCGAGCTGAAGCTGACCAGAAAGGCCGCCTACGTGAGATACTTCAACAGCA
GCGCCTTCTTCTTCAGCGGCTTCTTCGTGGTGTTTCCTGAGCGTGCTGCCCTACGCC
CTGATCAAGGGCATCATCCTGAGAAAGATCTTCACCACCATCAGCTTCTGCATCG
TGCTGAGAATGGCCGTGACCAGACAGTTCCTGAGGCGGTGCAGACCTGGTACG
ACAGCCTGGGCGCCATCAACAAGATCCAGGACTTCCTGCAGAAGCAGGAGTACA
AGACCCTGGAGTACAACCTGACCACCACCGAGGTGGTGATGGAGAACGTGACCG
CCTTCTGGGAGGAGGGCTTCGGCGAGCTGTTTCGAGAAGGCCAAGCAGAACAACA
ACAACAGAAAGACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTTCAGCC
TGCTGGGCACCCCGTGCTGAAGGACATCAACTTCAAGATCGAGAGAGGCCAGC
TGCTGGCCGTGGCCGGCAGCACCGGCGCCGGCAAGACCAGCCTGCTGATGGTGA
TCATGGGCGAGCTGGAGCCCAGCGAGGGCAAGATCAAGCACAGCGGCAGAATC
AGCTTCTGCAGCCAGTTCAGCTGGATCATGCCCGGCACCATCAAGGAGAACATC
ATCTTCGGCGTGAGCTACGACGAGTACAGATACAGAAGCGTGATCAAGGCCTGC
CAGCTGGAGGAGGACATCAGCAAGTTCGCCGAGAAGGACAACATCGTGCTGGGC
GAGGGCGGCATCACCTGAGCGGCGGCCAGAGAGCCAGAATCAGCCTGGCCAG
AGCCGTGTACAAGGACGCCGACCTGTACCTGCTGGACAGCCCTTCGGCTACCTG
GACGTGCTGACCGAGAAGGAGATCTTCGAGAGCTGCGTGTGCAAGCTGATGGCC
ACAAGACCAGAATCCTGGTGACCAGCAAGATGGAGCACCTGAAGAAGGCCGA
CAAGATCCTGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCTTCAGCGA
GCTGCAGAACCTGCAGCCCGACTTCAGCAGCAAGCTGATGGGCTGCGACAGCTT
CGACCAGTTCAGCGCCGAGAGAAGAAACAGCATCCTGACCGAGACCCTGCACAG
ATTCAGCCTGGAGGGCGACGCCCCGTGAGCTGGACCGAGACCAAGAAGCAGAG
CTTCAAGCAGACCGGCGAGTTCGGCGAGAAGAGAAAGAACAGCATCCTGAACCC

CATCAACAGCATCAGAAAGTTCAGCATCGTGCAGAAGACCCCCCTGCAGATGAA
CGGCATCGAGGAGGACAGCGACGAGCCCCTGGAGAGAAGACTGAGCCTGGTGC
CCGACAGCGAGCAGGGCGAGGCCATCCTGCCAGAATCAGCGTGATCAGCACCG
GCCCCACCCTGCAGGCCAGAAGAAGACAGAGCGTGCTGAACCTGATGACCCACA
GCGTGAACCAGGGCCAGAACATCCACAGAAAGACCACCGCCAGCACCAGAAAG
GTGAGCCTGGCCCCCAGGCCAACCTGACCGAGCTGGACATCTACAGCAGAAGA
CTGAGCCAGGAGACCGGCCTGGAGATCAGCGAGGAGATCAACGAGGAGGACCT
GAAGGAGTGCTTCTTCGACGACATGGAGAGCATCCCCGCCGTGACCACCTGGAA
CACCTACCTGAGATACATCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGG
TGCTGGTGATCTTCTGGCCGAGGTGGCCGCCAGCCTGGTGGTGCTGTGGCTGC
TGGGCAACACCCCCCTGCAGGACAAGGGCAACAGCACCCACAGCAGAAACAAC
AGCTACGCCGTGATCATCACAGCACCAGCAGCTACTACGTGTTCTACATCTACG
TGGGCGTGGCCGACACCCTGCTGGCCATGGGCTTCTTCAGAGGCCTGCCCTGGT
GCACACCCTGATCACCGTGAGCAAGATCCTGCACCACAAGATGCTGCACAGCGT
GCTGCAGGCCCCCATGAGCACCCCTGAACACCCTGAAGGCCGGCGGCATCCTGAA
CAGATTCAGCAAGGACATCGCCATCCTGGACGACCTGCTGCCCTGACCATCTTC
GACTTCATCCAGCTGCTGCTGATCGTGATCGGCGCCATCGCCGTGGTGGCCGTGC
TGCAGCCCTACATCTTCGTGGCCACCGTGCCCGTGATCGTGGCCTTCATCATGCT
GAGAGCCTACTTCCTGCAGACCAGCCAGCAGCTGAAGCAGCTGGAGAGCGAGGG
CAGAAGCCCCATCTTCACCCACCTGGTGACCAGCCTGAAGGGCCTGTGGACCCTG
AGAGCCTTCGGCAGACAGCCCTACTTCGAGACCCTGTTCCACAAGGCCCTGAACC
TGCACACCGCCAACCTGGTTCCTGTACCTGAGCACCCCTGAGATGGTTCAGATGAG
AATCGAGATGATCTTCGTGATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTGA
CCACCGGCGAGGGCGAGGGCAGAGTGGGCATCATCCTGACCCTGGCCATGAACA
TCATGAGCACCCCTGCAGTGGGCCGTGAACAGCAGCATCGACGTGGACAGCCTGA
TGAGAAGCGTGAGCAGAGTGTTCAAGTTCATCGACATGCCACCGAGGGCAAGC
CCACCAAGAGCACCAAGCCCTACAAGAACGGCCAGCTGAGCAAGGTGATGATCA
TCGAGAACAGCCACGTGAAGAAGGACGACATCTGGCCCAGCGGCGGCCAGATG
ACCGTGAAGGACCTGACCGCCAAGTACACCGAGGGCGGCAACGCCATCCTGGAG
AACATCAGCTTCAGCATCAGCCCCGCCAGAGAGTGGGCCTGCTGGGCAGAACC
GGCAGCGGCAAGAGCACCCCTGCTGAGCGCCTTCTGAGACTGCTGAACACCGAG
GGCGAGATCCAGATCGACGGCGTGAGCTGGGACAGCATCACCCCTGCAGCAGTGG
AGAAAGGCCTTCGGCGTGATCCCCAGAAGGTGTTTCATCTTCAGCGGCACCTTCA

GAAAGAACCTGGACCCCTACGAGCAGTGGAGCGACCAGGAGATCTGGAAGGTG
GCCGACGAGGTGGGCCTGAGAAGCGTGATCGAGCAGTTCCCCGGCAAGCTGGAC
TTCGTGCTGGTGGACGGCGGCTGCGTGCTGAGCCACGGCCACAAGCAGCTGATG
TGCTGGCCAGAAGCGTGCTGAGCAAGGCCAAGATCCTGCTGCTGGACGAGCCC
AGCGCCACCTGGACCCCGTGACCTACCAGATCATCAGAAGAACCCTGAAGCAG
GCCTTCGCCGACTGCACCGTGATCCTGTGCGAGCACAGAATCGAGGCCATGCTG
GAGTGCCAGCAGTTCCTGGTGATCGAGGAGAACAAGGTGAGACAGTACGACAGC
ATCCAGAAGCTGCTGAACGAGAGAAGCCTGTTTCAGACAGGCCATCAGCCCCAGC
GACAGAGTGAAGCTGTTCCCCACAGAAACAGCAGCAAGTGCAAGAGCAAGCCC
CAGATCGCCGCCCTGAAGGAGGAGACCGAGGAGGAGGTGCAGGACACCAGACT
GTGA (SEQ ID NO: 35)

ATGCAGCGCAGCCCCCTGGAGAAGGCCAGCGTGGTGAGCAAGCTGTTCTTCAGC
TGGACCCGCCCATCCTGCGCAAGGGCTACCGCCAGCGCCTGGAGCTGAGCGAC
ATCTACCAGATCCCAGCGTGGACAGCGCCGACAACCTGAGCGAGAAGCTGGAG
CGCGAGTGGGACCGCGAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAACGCC
CTGCGCCGCTGCTTCTTCTGGCGCTTCATGTTCTACGGCATCTTCCTGTACCTGGG
CGAGGTGACCAAGGCCGTGCAGCCCCTGCTGCTGGGCCGCATCATCGCCAGCTA
CGACCCCGACAACAAGGAGGAGCGCAGCATCGCCATCTACCTGGGCATCGGCCT
GTGCCTGCTGTTTCATCGTGCGCACCCCTGCTGCTGCACCCCGCCATCTTCGGCCTGC
ACCACATCGGCATGCAGATGCGCATCGCCATGTTTCAGCCTGATCTACAAGAAGA
CCCTGAAGCTGAGCAGCCGCGTGCTGGACAAGATCAGCATCGGCCAGCTGGTGA
GCCTGCTGAGCAACAACCTGAACAAGTTCGACGAGGGCCTGGCCCTGGCCCACT
TCGTGTGGATCGCCCCCTGCAGGTGGCCCTGCTGATGGGCCTGATCTGGGAGCT
GCTGCAGGCCAGCGCCTTCTGCGGCCTGGGCTTCTGATCGTGCTGGCCCTGTTC
CAGGCCGGCCTGGGCCGCATGATGATGAAGTACCGCGACCAGCGCGCCGGCAAG
ATCAGCGAGCGCCTGGTGATCACCAGCGAGATGATCGAGAACATCCAGAGCGTG
AAGGCCTACTGCTGGGAGGAGGCCATGGAGAAGATGATCGAGAACCTGCGCCAG
ACCGAGCTGAAGCTGACCCGCAAGGCCGCCTACGTGCGCTACTTCAACAGCAGC
GCCTTCTTCTTCAGCGGCTTCTTCGTGGTGTTCTGAGCGTGCTGCCCTACGCCCT
GATCAAGGGCATCATCCTGCGCAAGATCTTACCACCATCAGCTTCTGCATCGTG
CTGCGCATGGCCGTGACCCGCCAGTCCCCTGGGCCGTGCAGACCTGGTACGACA
GCCTGGGCGCCATCAACAAGATCCAGGACTTCTGCAGAAGCAGGAGTACAAGA

CCCTGGAGTACAACCTGACCACCACCGAGGTGGTGGATGGAGAACGTGACCGCCT
TCTGGGAGGAGGGCTTCGGCGAGCTGTTTCGAGAAGGCCAAGCAGAACAACAACA
ACCGCAAGACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTTCAGCCTGC
TGGGCACCCCCGTGCTGAAGGACATCAACTTCAAGATCGAGCGCGGCCAGCTGC
TGGCCGTGGCCGGCAGCACCGGCGCCGGCAAGACCAGCCTGCTGATGGTGGATCA
TGGGCGAGCTGGAGCCCAGCGAGGGCAAGATCAAGCACAGCGGCCGCATCAGCT
TCTGCAGCCAGTTCAGCTGGATCATGCCCGGCACCATCAAGGAGAACATCATCTT
CGGCGTGAGCTACGACGAGTACCGCTACCGCAGCGTGATCAAGGCCTGCCAGCT
GGAGGAGGACATCAGCAAGTTCGCCGAGAAGGACAACATCGTGCTGGGCGAGG
GCGGCATCACCTGAGCGGCGGCCAGCGCGCCCGCATCAGCCTGGCCCCGCGCCG
TGTACAAGGACGCCGACCTGTACCTGCTGGACAGCCCCCTTCGGCTACCTGGACGT
GCTGACCGAGAAGGAGATCTTCGAGAGCTGCGTGTGCAAGCTGATGGCCAACAA
GACCCGCATCCTGGTGACCAGCAAGATGGAGCACCTGAAGAAGGCCGACAAGAT
CCTGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCTTCAGCGAGCTGCAG
AACCTGCAGCCCGACTTCAGCAGCAAGCTGATGGGCTGCGACAGCTTCGACCAG
TTCAGCGCCGAGCGCCGCAACAGCATCCTGACCGAGACCCTGCACCGCTTCAGC
CTGGAGGGCGACGCCCCGTGAGCTGGACCGAGACCAAGAAGCAGAGCTTCAAG
CAGACCGGCGAGTTCGGCGAGAAGCGCAAGAACAGCATCCTGAACCCCATCAAC
AGCATCCGCAAGTTCAGCATCGTGCAGAAGACCCCCCTGCAGATGAACGGCATC
GAGGAGGACAGCGACGAGCCCCTGGAGCGCCGCCTGAGCCTGGTGCCCGACAGC
GAGCAGGGCGAGGCCATCCTGCCCCGCATCAGCGTGATCAGCACCGGCCCCACC
CTGCAGGCCCGCCGCCGCCAGAGCGTGCTGAACCTGATGACCCACAGCGTGAAC
CAGGGCCAGAACATCCACCGCAAGACCACCGCCAGCACCCGCAAGGTGAGCCTG
GCCCCCAGGCCAACCTGACCGAGCTGGACATCTACAGCCGCCGCCTGAGCCAG
GAGACCGGCCTGGAGATCAGCGAGGAGATCAACGAGGAGGACCTGAAGGAGTG
CTTCTTCGACGACATGGAGAGCATCCCCGCCGTGACCACCTGGAACACCTACCTG
CGCTACATCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGGTGCCTGGTGA
TCTTCTGGCCGAGGTGGCCGCCAGCCTGGTGGTGGCTGTGGCTGCTGGGCAACAC
CCCCCTGCAGGACAAGGGCAACAGCACCCACAGCCGCAACAACAGCTACGCCGT
GATCATCACAGCACAGCAGCTACTACGTGTTCTACATCTACGTGGGCGTGGCC
GACACCCTGCTGGCCATGGGCTTCTTCCGCGGCCCTGCCCTGGTGCACACCCTGA
TCACCGTGAGCAAGATCCTGCACCACAAGATGCTGCACAGCGTGCTGCAGGCC
CCATGAGCACCTGAACACCCTGAAGGCCGGCGGCATCCTGAACCGCTTCAGCA

AGGACATCGCCATCCTGGACGACCTGCTGCCCTGACCATCTTCGACTTCATCCA
GCTGCTGCTGATCGTGATCGGCGCCATCGCCGTGGTGGCCGTGCTGCAGCCCTAC
ATCTTCGTGGCCACCGTGCCCGTGATCGTGGCCTTCATCATGCTGCGCGCCTACTT
CCTGCAGACCAGCCAGCAGCTGAAGCAGCTGGAGAGCGAGGGCCGCAGCCCCAT
CTTCACCCACCTGGTGACCAGCCTGAAGGGCCTGTGGACCCTGCGCGCCTTCGGC
CGCCAGCCCTACTTCGAGACCCTGTTCCACAAGGCCCTGAACCTGCACACCGCCA
ACTGGTTCCTGTACCTGAGCACCTGCGCTGGTTCAGATGCGCATCGAGATGAT
CTTCGTGATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTGACCACCGGCGAG
GGCGAGGGCCGCGTGGGCATCATCCTGACCCTGGCCATGAACATCATGAGCACC
CTGCAGTGGGCCGTGAACAGCAGCATCGACGTGGACAGCCTGATGCGCAGCGTG
AGCCGCGTGTTCAAGTTCATCGACATGCCACCGAGGGCAAGCCCACCAAGAGC
ACCAAGCCCTACAAGAACGGCCAGCTGAGCAAGGTGATGATCATCGAGAACAGC
CACGTGAAGAAGGACGACATCTGGCCAGCGGGCCAGATGACCGTGAAGGA
CCTGACCGCCAAGTACACCGAGGGCGGCAACGCCATCCTGGAGAACATCAGCTT
CAGCATCAGCCCCGGCCAGCGCGTGGGCTGCTGGGCGCACCGGCAGCGGCAA
GAGCACCTGCTGAGCGCCTTCCTGCGCCTGCTGAACACCGAGGGCGAGATCCA
GATCGACGGCGTGAGCTGGGACAGCATCACCTGCAGCAGTGGCGCAAGGCCTT
CGGCGTGATCCCCCAGAAGGTGTTTCATCTTCAGCGGCACCTTCGCAAGAACCTG
GACCCCTACGAGCAGTGGAGCGACCAGGAGATCTGGAAGGTGGCCGACGAGGTG
GGCCTGCGCAGCGTGATCGAGCAGTTCGCCGCAAGCTGGACTTCGTGCTGGTGG
ACGGCGGCTGCGTGCTGAGCCACGGCCACAAGCAGCTGATGTGCCTGGCCCCGA
GCGTGCTGAGCAAGGCCAAGATCCTGCTGCTGGACGAGCCCAGCGCCCACCTGG
ACCCCGTGACCTACCAGATCATCCGCCGCACCCTGAAGCAGGCCTTCGCCGACTG
CACCGTGATCCTGTGCGAGCACCGCATCGAGGCCATGCTGGAGTGCCAGCAGTT
CCTGGTGATCGAGGAGAACAAGGTGCGCCAGTACGACAGCATCCAGAAGCTGCT
GAACGAGCGCAGCCTGTTCCGCCAGGCCATCAGCCCCAGCGACCGCGTGAAGCT
GTTCCCCCACCGCAACAGCAGCAAGTGCAAGAGCAAGCCCCAGATCGCCGCCCT
GAAGGAGGAGACCGAGGAGGAGGTGCAGGACACCCGCCTGTAA (SEQ ID NO:

36)

ATGCAGAGAAGCCCCCTGGAGAAGGCCAGCGTGGTGGAGCAAGCTGTTCTTCAGC
TGGACCAGACCCATCCTGAGAAAGGGCTACAGACAGAGACTGGAGCTGAGCGAC
ATCTACCAGATCCCCAGCGTGGACAGCGCCGACAACCTGAGCGAGAAGCTGGAG

AGAGAGTGGGACAGAGAGCTGGCCAGCAAGAAGAACCCCAAGCTGATCAACGC
CCTGAGAAGATGCTTCTTCTGGAGATTCATGTTCTACGGCATCTTCCTGTACCTGG
GCGAGGTGACCAAGGCCGTGCAGCCCCTGCTGCTGGGCAGAATCATCGCCAGCT
ACGACCCCGACAACAAGGAGGAGAGAAGCATCGCCATCTACCTGGGCATCGGCC
TGTGCCTGCTGTTTCATCGTGAGAACCCTGCTGCTGCACCCCGCCATCTTCGGCCT
GCACCACATCGGCATGCAGATGAGAATCGCCATGTTTCAGCCTGATCTACAAGAA
GACCCTGAAGCTGAGCAGCAGAGTGCTGGACAAGATCAGCATCGGCCAGCTGGT
GAGCCTGCTGAGCAACAACCTGAACAAGTTCGACGAGGGCCTGGCCCTGGCCCA
CTTCGTGTGGATCGCCCCCTGCAGGTGGCCCTGCTGATGGGCCTGATCTGGGAG
CTGCTGCAGGCCAGCGCCTTCTGCGGCCTGGGCTTCCTGATCGTGCTGGCCCTGT
TCCAGGCCGGCCTGGGCAGAATGATGATGAAGTACAGGGACCAGAGAGCCGGCA
AGATCAGCGAGAGACTGGTGATCACCAGCGAGATGATCGAGAACATCCAGAGCG
TGAAGGCCTACTGCTGGGAGGAGGCCATGGAGAAGATGATCGAGAACCTGAGAC
AGACCGAGCTGAAGCTGACCAGAAAGGCCGCCTACGTGAGATACTTCAACAGCA
GCGCCTTCTTCTCAGCGGCTTCTTCGTGGTGTTTCCTGAGCGTGCTGCCCTACGCC
CTGATCAAGGGCATCATCCTGAGAAAGATCTTCACCACCATCAGCTTCTGCATCG
TGCTGAGAATGGCCGTGACCAGACAGTTCCTGAGGCGGTGCAGACCTGGTACG
ACAGCCTGGGCGCCATCAACAAGATCCAGGACTTCCTGCAGAAGCAGGAGTACA
AGACCCTGGAGTACAACCTGACCACCACCGAGGTGGTGATGGAGAACGTGACCG
CCTTCTGGGAGGAGGGCTTCGGCGAGCTGTTTCGAGAAGGCCAAGCAGAACAACA
ACAACAGAAAGACCAGCAACGGCGACGACAGCCTGTTCTTCAGCAACTTCAGCC
TGCTGGGCACCCCGTGCTGAAGGACATCAACTTCAAGATCGAGAGAGGCCAGC
TGCTGGCCGTGGCCGGCAGCACCGGCGCCGGCAAGACCAGCCTGCTGATGGTGA
TCATGGGCGAGCTGGAGCCCAGCGAGGGCAAGATCAAGCACAGCGGCAGAATC
AGCTTCTGCAGCCAGTTCAGCTGGATCATGCCCGGCACCATCAAGGAGAACATC
ATCTTCGGCGTGAGCTACGACGAGTACAGATACAGAAGCGTGATCAAGGCCTGC
CAGCTGGAGGAGGACATCAGCAAGTTCGCCGAGAAGGACAACATCGTGCTGGGC
GAGGGCGGCATCACCTGAGCGGCGGCCAGAGAGCCAGAATCAGCCTGGCCAG
AGCCGTGTACAAGGACGCCGACCTGTACCTGCTGGACAGCCCCTTCGGCTACCTG
GACGTGCTGACCGAGAAGGAGATCTTCGAGAGCTGCGTGTGCAAGCTGATGGCC
ACAAGACCAGAATCCTGGTGACCAGCAAGATGGAGCACCTGAAGAAGGCCGA
CAAGATCCTGATCCTGCACGAGGGCAGCAGCTACTTCTACGGCACCTTCAGCGA
GCTGCAGAACCTGCAGCCCGACTTCAGCAGCAAGCTGATGGGCTGCGACAGCTT

CGACCAGTTCAGCGCCGAGAGAAGAAACAGCATCCTGACCGAGACCCTGCACAG
ATTCAGCCTGGAGGGCGACGCCCCGTGAGCTGGACCGAGACCAAGAAGCAGAG
CTTCAAGCAGACCGGCGAGTTCGGCGAGAAGAGAAAGAACAGCATCCTGAACCC
CATCAACAGCATCAGAAAGTTCAGCATCGTGCAGAAGACCCCCCTGCAGATGAA
CGGCATCGAGGAGGACAGCGACGAGCCCCTGGAGAGAAGACTGAGCCTGGTGC
CCGACAGCGAGCAGGGCGAGGCCATCCTGCCCAGAATCAGCGTGATCAGCACCG
GCCCCACCCTGCAGGCCAGAAGAAGACAGAGCGTGCTGAACCTGATGACCCACA
GCGTGAACCAGGGCCAGAACATCCACAGAAAGACCACCGCCAGCACCCAGAAAG
GTGAGCCTGGCCCCCAGGCCAACCTGACCGAGCTGGACATCTACAGCAGAAGA
CTGAGCCAGGAGACCGGCCTGGAGATCAGCGAGGAGATCAACGAGGAGGACCT
GAAGGAGTGCTTCTTCGACGACATGGAGAGCATCCCCGCCGTGACCACCTGGAA
CACCTACCTGAGATACATCACCGTGCACAAGAGCCTGATCTTCGTGCTGATCTGG
TGCTGGTGATCTTCTGGCCGAGGTGGCCGCCAGCCTGGTGGTGCTGTGGCTGC
TGGGCAACACCCCCCTGCAGGACAAGGGCAACAGCACCCACAGCAGAAACAAC
AGCTACGCCGTGATCATCACCCAGCACCCAGCAGCTACTACGTGTTCTACATCTACG
TGGGCGTGGCCGACACCCTGCTGGCCATGGGCTTCTTCAGAGGCCTGCCCTGGT
GCACACCCTGATCACCGTGAGCAAGATCCTGCACCACAAGATGCTGCACAGCGT
GCTGCAGGCCCCCATGAGCACCCCTGAACACCCTGAAGGCCGGCGGCATCCTGAA
CAGATTCAGCAAGGACATCGCCATCCTGGACGACCTGCTGCCCTGACCATCTTC
GACTTCATCCAGCTGCTGCTGATCGTGATCGGCGCCATCGCCGTGGTGGCCGTGC
TGCAGCCCTACATCTTCGTGGCCACCGTGCCCGTGATCGTGGCCTTCATCATGCT
GAGAGCCTACTTCCTGCAGACCAGCCAGCAGCTGAAGCAGCTGGAGAGCGAGGG
CAGGAGCCCCATCTTCACCCACCTGGTGACCAGCCTGAAGGGCCTGTGGACCCTG
AGAGCCTTCGGCAGACAGCCCTACTTCGAGACCCTGTTCCACAAGGCCCTGAACC
TGCACACCGCCAACCTGGTTCTGTACCTGAGCACCCCTGAGATGGTTCCAGATGAG
AATCGAGATGATCTTCGTGATCTTCTTCATCGCCGTGACCTTCATCAGCATCCTGA
CCACCGGCGAGGGCGAGGGCAGAGTGGGCATCATCCTGACCCTGGCCATGAACA
TCATGAGCACCCCTGCAGTGGGCCGTGAACAGCAGCATCGACGTGGACAGCCTGA
TGAGAAGCGTGAGCAGAGTGTTCAAGTTCATCGACATGCCACCGAGGGCAAGC
CCACCAAGAGCACCAGCCCTACAAGAACGGCCAGCTGAGCAAGGTGATGATCA
TCGAGAACAGCCACGTGAAGAAGGACGACATCTGGCCCAGCGGCGGCCAGATG
ACCGTGAAGGACCTGACCGCCAAGTACACCGAGGGCGGCAACGCCATCCTGGAG
AACATCAGCTTCAGCATCAGCCCCGGCCAGAGAGTGGGCCTGCTGGGCAGAACC

GGCAGCGGCAAGAGCACCTGCTGAGCGCCTTCCTGAGACTGCTGAACACCGAG
GGCGAGATCCAGATCGACGGCGTGAGCTGGGACAGCATCACCTGCAGCAGTGG
AGAAAGGCCTTCGGCGTGATCCCCAGAAGGTGTTTCATCTTCAGCGGCACCTTCA
GAAAGAACCTGGACCCCTACGAGCAGTGGAGCGACCAGGAGATCTGGAAGGTG
GCCGACGAGGTGGGCTGAGAAGCGTGATCGAGCAGTTCCCCGGCAAGCTGGAC
TTCGTGCTGGTGGACGGCGGCTGCGTGCTGAGCCACGGCCACAAGCAGCTGATG
TGCCTGGCCAGAAGCGTGCTGAGCAAGGCCAAGATCCTGCTGCTGGACGAGCCC
AGCGCCACCTGGACCCCGTGACCTACCAGATCATCAGAAGAACCCTGAAGCAG
GCCTTCGCCGACTGCACCGTGATCCTGTGCGAGCACAGAATCGAGGCCATGCTG
GAGTGCCAGCAGTTCCTGGTGATCGAGGAGAACAAGGTGAGACAGTACGACAGC
ATCCAGAAGCTGCTGAACGAGAGAAGCCTGTTTCAGACAGGCCATCAGCCCCAGC
GACAGAGTGAAGCTGTTCCCCACAGAAACAGCAGCAAGTGCAAGAGCAAGCCC
CAGATCGCCGCCCTGAAGGAGGAGACCGAGGAGGAGGTGCAGGACACCAGACT
GTGA (SEQ ID NO: 37)

ATGCAGAGGTCACCTCTGGAAAAGGCTAGCGTGGTCAGCAAGCTATTTTTTTCCT
GGACCCGCCCGATACTCAGGAAGGGCTACCGACAGCGGCTGGAGCTGAGTGACA
TTTATCAGATTCCCTCCGTCGATTCCGCTGACAACCTGTCTGAGAACTGGAGCG
GGAATGGGATAGGGAACTGGCGTCCAAAAAAAACCCCAAACCTCATCAATGCACT
CCGCAGATGCTTCTTCTGGCGGTTTATGTTTTATGGCATATTCCTGTATCTGGGGG
AGGTGACGAAAGCCGTGCAGCCGCTGCTGCTTGGTCGCATTATCGCGTCATACGA
TCCAGATAACAAGGAGGAAAGAAGTATCGCTATCTATCTCGGGATAGGGCTGTG
CCTGCTCTTCATTGTGCGGACTCTTCTCTTGCACCCCGCCATTTTCGGTCTGCATC
ATATAGGTATGCAGATGAGAATTGCGATGTTCTCATTGATTTACAAAAAACGCT
TAAGCTAAGTTCAAGGGTGCTAGATAAGATATCGATCGGCCAGCTGGTGTCTCTG
CTTAGCAACAACCTCAATAAATTCGACGAAGGCCTTGCCTGGCCACTTCGTGT
GGATCGCCCCTCTGCAGGTGGCTCTGCTGATGGGGTTAATATGGGAGCTGTTGCA
GGCCTCCGCTTTTTGTGGCCTGGGGTTTCTCATCGTGTTGGCCTTGTTTCAGGCAG
GGCTGGGACGTATGATGATGAAATATAGGGATCAGAGGGCTGGCAAATCTCTG
AGCGCCTGGTTATTACGAGTGAAATGATTGAGAACATCCAGTCAGTGAAGGCCT
ATTGCTGGGAGGAGGCCATGGAAAAAATGATTGAGAACCTACGCCAGACTGAGC
TGAAGTTAACCAGAAAAGCCGCCTATGTGCGCTACTTTAACAGTAGCGCATTTTT
CTTCTCCGGTTTTTTTCGTGGTGTTCCTTAGTGTGTTGCCGTATGCCTTAATCAAGG

GAATAATACTCCGGAAGATTTTCACTACCATCAGCTTCTGTATCGTGTTGCGGAT
GGCCGTCACCCGGCAGTTTCCCTGGGCAGTACAGACTTGGTACGATTCTCTCGGA
GCAATTAACAAAATCCAAGACTTTCTACAAAAGCAGGAGTACAAGACCCTGGAG
TACAATCTGACCACCACAGAAGTCGTAATGGAGAATGTAAGTGCCTTCTGGGAA
GAGGGCTTTGGCGAACTCTTTGAAAAGGCCAAGCAGAACAATAACAACCGGAAG
ACCTCCAACGGGGACGACAGCTTATTTTTTCAGCAATTTTTCTTTGCTCGGGACCCC
TGTAAGTAAAGATATTAAGTTCGAGCGCGGACAACCTCTGGCTGTCGCC
GGCAGCACTGGAGCTGGAAAAACATCACTGCTTATGGTGATAATGGGAGAAGTCTC
GAACCAAGCGAGGGAAAAATAAAGCACTCTGGACGGATTAGTTTTTGCTCCCAG
TTCTCGTGGATAATGCCTGGCACCATTAAGGAGAATATCATCTTTGGAGTGAGTT
ACGACGAATACCGGTACCGGTCCGTTATCAAGGCTTGTCAACTCGAGGAGGACA
TTTCTAAATTGCGCGAAAAAGATAATATAGTGCTGGGCGAAGGAGGCATTACAC
TGAGCGGGGGTTCAGAGAGCTCGAATTAGCCTCGCCCGAGCAGTCTATAAAGACG
CCGATCTTTACCTGCTGGATTCCCCTTTTGGGTATTTGGATGTTCTGACAGAGAAG
GAAATCTTTGAATCATGTGTCTGTAAACTGATGGCCAATAAGACTAGGATTCTAG
TGACTTCGAAAATGGAGCACCTGAAAAAAGCGGACAAAATTCTGATACTCCATG
AAGGGTCTTCCTACTTCTACGGCACCTTCTCAGAGTTGCAGAACTTACAACCTGA
TTTTTCATCTAAGCTTATGGGGTGCAGTCTGTTGACCAGTTCTCCGCTGAAAGAC
GAAACAGCATCTTAACGGAACTCTTCACAGGTTCTCATTAGAGGGAGATGCGC
CGGTGTCCTGGACAGAGACAAAAAACAGTCTTTCAAACAGACAGGAGAGTTTG
GCGAGAAGAGAAAAAACTCAATCCTCAATCCCATCAATTCTATTAGAAAGTTTA
GCATCGTCCAAAAACACCATTGCAGATGAATGGGATTGAGGAGGACAGTGATG
AGCCTTTGGAACGAAGACTGTCCCTGGTACCCGATAGCGAACAGGGTGAGGCCA
TCCTTCCTAGGATCTCGGTCATAAGTACAGGGCCCACTGCAGGCCAGGCGAC
GTCAAAGTGTCTCAATCTTATGACGCACAGTGTGAATCAGGGGCAGAACATCC
ATCGTAAGACGACAGCTTCAACTCGAAAGGTCAGTCTAGCTCCACAAGCCAATCT
TACAGAGCTGGACATTTATTCCCGCCGCCTCAGTCAGGAGACCGGATTGGAAATA
TCAGAGGAAATTAATGAAGAGGATCTGAAGGAATGCTTCTTTGATGACATGGAA
TCGATCCCCGCTGTTACTACCTGGAACACATATCTGAGATATATTACCGTCCATA
AGAGCTTAATCTTTGTAAGTATGGTGCTTGGTGATTTTCTGGCAGAGGTTGC
GGCGAGTTTGGTTCGTGCTATGGCTCCTTGGAAACACTCCCCTGCAGGATAAGGGG
AACTCCACTCATAGCAGGAATAACAGCTATGCCGTGATCATCACCTCTACCTCCT
CTTATTACGTGTTTTACATATACGTCGGTGTGCGGATACCCTGTTGGCAATGGG

GTTCTTTAGAGGACTACCCCTAGTTCACACCCTGATCACCGTTTCGAAGATCTTGC
ACCACAAGATGCTTCATAGCGTTCTCCAAGCTCCTATGAGCACCCCTTAATACACT
GAAAGCAGGAGGTATCCTTAACCGCTTTTCCAAAGACATCGCTATACTCGACGAT
TTGCTCCCATTGACCATCTTCGACTTCATTCAGCTGCTCCTCATTGTGATCGGGCGC
CATTGCCGTGGTCGCAGTGTTACAGCCATATATTTTCGTAGCCACCGTGCCCGTC
ATCGTGGCATTATCATGCTGCGCGCATATTTCTTACAGACATCTCAGCAACTGA
AGCAGCTGGAATCTGAGGGCAGATCTCCTATTTTTACACACCTGGTTACCAGCCT
GAAGGGCCTGTGGACCCTGCGTGCTTTTCGGTCGCCAACCCTACTTTGAGACTCTC
TTCCATAAGGCTCTGAATTTACATACTGCCAATTGGTTCCTATACCTTAGTACCCT
TCGGTGGTTCAGATGCGGATAGAAATGATCTTCGTGATTTTCTTCATCGCAGTC
ACTTTCATCTCTATTTTGACGACCGGTGAGGGCGAGGGCAGGGTGGGCATCATT
TGACTTTGGCCATGAACATTATGTCAACACTCCAGTGGGCCGTTAATTCAAGCAT
TGATGTGGATTCTTGATGCGTTCGTCAGCAGGGTATTTAAATTCATAGACATG
CCCACCGAGGGCAAGCCAACAAAATCTACCAAGCCATACAAAATGGCCAACTA
AGCAAGGTCATGATTATCGAGAATTCTCATGTGAAAAGGACGACATTTGGCCTT
CCGGGGGTCAAATGACTGTAAAGGACCTGACGGCTAAATACACTGAGGGCGGTA
ATGCTATCTTGAGAACATCTCTTTCAGCATCTCCCCTGGCCAGAGAGTGGGACT
GCTCGGGCGGACAGGCTCCGGAAAGTCTACGCTCCTTTCAGCATTCCTTAGACTT
CTGAACACCGAAGGTGAGATTCAGATTGACGGGGTCTCTTGGGACTCCATCACAC
TTCAGCAATGGAGGAAGGCATTCGGTGTAAATCCCCAAAAGGTTTTTATCTTCTC
CGAACATTTTCGTAAGAATCTGGACCCGTACGAGCAGTGGTCAGATCAGGAGAT
CTGGAAAGTAGCAGACGAGGTCGGGCTACGGAGCGTTATTGAACAGTTTCCTGG
CAAACCTGGACTTCGTTTTGGTGGACGGAGGCTGTGTGCTGAGTCACGGCCATAAA
CAACTGATGTGCTTAGCTAGGTCTGTTCTCAGCAAGGCAAAGATTTTACTGCTGG
ATGAACCAAGCGCCCACCTTGATCCAGTGACATATCAAATCATCAGAAGAACTCT
TAAACAGGCGTTCGCCGACTGCACAGTGATCCTGTGTGAGCACAGAATAGAAGC
CATGCTGGAATGTCAACAGTTTCTCGTGATTGAGGAGAACAAGGTGCGCCAGTA
CGATAGCATCCAGAAGTTACTCAATGAAAGGTCACTCTTCAGGCAGGCCATCTCA
CCCAGCGACCGCGTTAAGCTGTTTCCACACCGAAACAGTTCCAAGTGCAAAAAGT
AAGCCACAGATTGCTGCACTGAAGGAAGAGACAGAAGAAGAAGTTCAGGACAC
TCGGCTCTGA (SEQ ID NO: 38)

ATGCAGAGGAGCCACTGGAGAAAGCCTCCGTGGTGAGTAAACTCTTTTTTAGTT
GGACCAGACCCATCCTGCGAAAAGGATACAGGCAGCGCCTCGAGTTGTCAGATA
TCTACCAGATTCTTCTGTGGACTCAGCTGACAATTTGAGTGAGAAGCTGGAGCG
GGAGTGGGATAGAGAGCTGGCGAGCAAAAAAACCCCAAGCTTATCAATGCTCT
GCGCCGCTGCTTTTTCTGGAGGTTTCATGTTTTATGGGATCTTCCTGTACCTGGGGG
AGGTCACCAAAGCTGTTTCAGCCGCTCCTTCTTGGCCGCATCATCGCCAGCTATGA
CCCTGATAATAAAGAAGAAAGGTCTATTGCTATTTATCTGGGAATTGGCCTCTGC
TTGCTCTTCATCGTCCGCACCCTTCTGCTGCACCCTGCCATTTTTGGCCTTCACCA
CATCGGCATGCAAATGAGAATTGCCATGTTCTCCCTCATTACAAAAAGACCCTG
AACTTTCCTCAAGAGTGTTAGATAAAATATCCATTGGTCAGCTGGTCAGCCTGC
TGTCACAACAATCTTAACAAATTTGATGAAGGCTTGGCGCTGGCCCACTTCGTGTG
GATTGCACCTCTGCAGGTGGCCCTGTTGATGGGACTTATATGGGAGCTGCTTCAA
GCCTCTGCTTTCTGTGGGCTGGGCTTTTTGATTGTACTGGCACTTTTTCAGGCTGG
GCTCGGAAGAATGATGATGAAATACAGAGATCAGCGGGCCGGGAAGATATCAG
AGCGACTTGTGATCACCAGTGAAATGATTGAAAATATTCAGAGCGTGAAAGCCT
ACTGCTGGGAAGAAGCCATGGAGAAGATGATTGAGAACCTGAGGCAGACAGAG
CTCAAGCTCACTCGGAAGGCTGCTTATGTTTCGCTATTTCAACAGCAGCGCCTTCTT
CTTCAGTGGCTTCTTTGTTGTCTTCCTGTCTGTTCTGCCATATGCACTGATAAAAG
GCATTATTTTACGAAAGATCTTCACCACCATCAGTTTTTGCATCGTTCTCAGGATG
GCCGTCACAAGACAGTTCCCCTGGGCTGTGCAGACCTGGTACGATTCCTTGGGGG
CCATCAACAAGATTCAAGATTTCTTGCAAAAACAAGAATATAAACTTTAGAAT
ACAACCTCACCACCACTGAAGTGGTCATGGAAAATGTGACAGCCTTTTGGGAGG
AGGGTTTTGGAGAATTGTTTCGAGAAGGCAAAGCAGAATAACAACAACAGGAAG
ACGAGCAATGGGGACGACTCTCTCTTCTTCAGCAACTTTTCACTGCTCGGGACCC
CTGTGTTGAAAGATATAAACTTCAAGATCGAGAGGGGCCAGCTCTTGGCTGTGGC
AGGCTCCACTGGAGCTGGTAAAACATCTCTTCTCATGGTGATCATGGGGGAACTG
GAGCCTTCCGAAGGAAAAATCAAGCACAGTGGGAGAATCTCATTCTGCAGCCAG
TTTTCTGGATCATGCCCGGCACCATTAAGGAAAACATCATATTTGGAGTGTCT
ATGATGAGTACCGCTACCGGTCAGTCATCAAAGCCTGTCAGTTGGAGGAGGACA
TCTCCAAGTTTGCAGAGAAAGACAACATTGTGCTTGGAGAGGGGGGTATCACTCT
TTCTGGAGGACAAAGAGCCAGGATCTCTTTGGCCCGGGCAGTCTACAAGGATGC
AGACCTCTACTTGTTGGACAGTCCCTTCGGCTACCTCGACGTGCTGACTGAAAAA
GAAATTTTTGAAAGCTGTGTGTGCAAACTGATGGCAAACAAGACCAGGATTCTTG

TCACCAGCAAGATGGAACATCTGAAGAAAGCGGACAAAATTCTGATTCTGCATG
AAGGGAGCTCCTACTTCTATGGAACATTTAGCGAGCTTCAGAACCTACAGCCAGA
CTTCTCCTCCAAATTAATGGGCTGTGACTCCTTCGACCAGTTCTCTGCAGAAAGA
AGAAACTCTATACTCACAGAGACCTCCACCGCTTCTCCCTTGAGGGAGATGCCC
CAGTTTCTTGACAGAAACCAAGAAGCAGTCCTTTAAGCAGACTGGCGAGTTTG
GTGAAAAGAGGAAAAATTCAATTCTCAATCCAATTAACAGTATTCGCAAGTTCA
GCATTGTCCAGAAGACACCCCTCCAGATGAATGGCATCGAAGAAGATAGTGACG
AGCCGCTGGAGAGACGGCTGAGTCTGGTGCCAGATTCAGAACAGGGGGAGGCCA
TCCTGCCCCGGATCAGCGTCATTTCCACAGGCCCCACATTACAAGCACGGCGCCG
GCAGAGTGTTTTAAATCTCATGACCCATTCAGTGAACCAGGGCCAAAATATCCAC
AGGAAGACTACAGCTTCTACCCGGAAAGTGTCTCTGGCCCCCTCAGGCCAATCTGA
CCGAGCTGGACATCTACAGCAGGAGGCTCTCCAGGAAACAGGGCTGGAAATAT
CTGAAGAGATTAATGAAGAGGATCTTAAAGAGTGCTTCTTTGATGACATGGAGA
GCATCCCCGCGGTGACCACATGGAACACCTACCTTAGATATATTACTGTCCACAA
GAGCCTCATATTTGTCCTCATCTGGTGCCTGGTTATTTTCCTCGCTGAGGTGGCGG
CCAGTCTTGTTGTGCTCTGGCTGCTGGGCAACACTCCTCTCCAGGACAAGGGCAA
TAGTACTCACAGCAGAAATAATTCTTATGCCGTCATCATTACAAGCACCTCCAGC
TACTACGTGTTCTACATCTATGTGGGCGTGGCTGACACCCTCCTGGCCATGGGTTT
CTTCCGGGGCCTGCCTTTGGTGCACACCCTCATCACAGTGTCAAAAATTCTGCAC
CATAAAATGCTTCATTCTGTCCTGCAGGCACCCATGAGCACTTTGAACACATTGA
AGGCTGGCGGCATCCTCAACAGATTTTCTAAAGATATTGCTATCCTGGATGATCT
CCTCCCCCTGACAATCTTTGACTTTATCCAGCTTCTGCTGATCGTGATTGGAGCCA
TAGCAGTGGTTGCTGTCTGCAGCCCTACATTTTTGTGGCCACCGTGCCCGTGATT
GTTGCCTTTATTATGCTCAGAGCTTACTTCCTGCAAACCTTCTCAACAGCTCAAACA
GCTAGAATCTGAGGGCCGGAGCCCCATTTTTACCCACCTGGTGACTTCCCTGAAG
GGACTGTGGACTCTGAGAGCATTCGGGCGACAGCCTTACTTTGAGACACTGTTCC
ACAAGGCCCTGAACTTGACACTGCCAACTGGTTTCTTTACCTGAGCACACTCCG
CTGGTTCCAGATGCGGATAGAGATGATCTTCGTCATCTTTTTTATAGCTGTAACCT
TCATTTCTATCCTTACAACAGGAGAAGGAGAGGGCAGGGTGGAATCATCCTCA
CGCTGGCTATGAACATAATGTCCACCTTGCAGTGGGCCGTGAATTCAGTATAGA
TGTGGATTCTCTAATGAGGAGTGTCTCCCGGGTGTTTAAATTCATTGATATGCCTA
CTGAGGGGAAACCCACCAAGTCAACAAAACCTTATAAGAATGGACAGCTGAGCA
AGGTGATGATAATTGAGAACAGCCACGTGAAGAAGGATGACATTTGGCCCAGCG

GGGGCCAGATGACTGTGAAGGACCTGACGGCCAAGTACACCGAAGGTGGAAATG
CCATTTTGGAAAACATCAGCTTCTCAATCTCTCCTGGGCAGAGAGTTGGATTGCT
GGGTCGCACGGGCAGCGGCAAATCAACCCTGCTCAGTGCCTTCCTTCGGCTCCTG
AATACAGAAGGCGAAATCCAAATTGACGGGGTGAGCTGGGACAGCATCACCTG
CAGCAGTGGAGAAAAGCATTTGGGGTCATTCCACAGAAAGTTTTTCATCTTCTCTG
GCACTTTCAGAAAAGAACCTGGACCCCTATGAGCAGTGGAGCGACCAGGAGATCT
GGAAGGTTGCAGATGAAGTTGGCCTGCGGAGTGTGATAGAACAATTTCTGGCA
AGCTGGATTTTGTGCTGGTAGATGGAGGCTGCGTGCTGTCCCACGGCCACAAACA
GCTGATGTGCCTCGCCCGCTCCGTTCTTTCAAAGGCCAAAATCTTGCTTTTGGATG
AGCCCAGTGCTCACCTCGACCCAGTGACCTATCAGATAATCCGCAGGACCTTAAA
GCAAGCTTTTGGCCACTGCACCGTCATACTGTGTGAGCACCGGATTGAAGCAATG
CTGGAATGCCAGCAGTTTCTGGTGATCGAGGAGAATAAGGTCCGGCAGTACGAC
AGCATCCAGAAGTTGTTGAATGAGCGCAGCCTTTTCCGCCAGGCCATCTCCCCAT
CTGACAGAGTCAAGCTGTTTCCACATAGGAACTCCTCTAAGTGCAAGTCCAAGCC
CCAGATCGCTGCCCTCAAGGAGGAAACTGAGGAAGAGGTGCAGGATACCCGCCT
GTGA (SEQ ID NO: 39)

ATGCAACGGAGTCCTCTGGAAAAAGCCTCTGTCGTATCTAAGCTTTTCTTCAGTT
GGACACGCCCGATTTTGGAGAAAGGGTTATCGGCAACGCTTGGAACCTTAGTGACA
TCTACCAAATTCCAAGTGTAGACTCAGCCGATAACTTGAGCGAAAAGCTCGAAC
GAGAGTGGGATCGGAACTGGCTAGCAAAAAAATCCCAAACCTCATAAATGCC
TGCGACGCTGTTTCTTTTGGCGATTTATGTTTTACGGTATTTTCTTTATTTGGGTG
AGGTCACGAAGGCTGTACAGCCACTGCTGCTGGGTCGCATCATTGCCTCTTACGA
CCCTGACAACAAAGAGGAGCGGTCAATAGCTATCTACCTTGGTATAGGACTTTGC
TTGCTCTTCATAGTCCGCACGTTGCTTCTCCACCCTGCTATATTTGGTCTCCATCA
CATTGGGATGCAAATGCGGATCGCGATGTTTCAGTCTTATATATAAAAAGACTCTT
AAACTTTCAGCCGGGTTCTGGATAAGATCTCTATTGGTCAACTGGTATCTCTTTT
GTCTAACAACTGAATAAGTTCGACGAGGGCCTTGCATTGGCCCATTTTGTATGG
ATTGCCCTTTGCAAGTCGCCCTCCTGATGGGATTGATCTGGGAACTCCTGCAAG
CTAGTGCTTTTTGCGGATTGGGATTCCTCATAGTCCTTGCCTCTTTCAGGCGGGA
CTTGACGCATGATGATGAAGTATCGCGACCAACGAGCTGGCAAGATCAGTGAA
CGGCTTGTAATAACCAGTGAAATGATAGAGAACATCCAGAGCGTAAAAGCTTAC
TGTTGGGAAGAAGCGATGGAAAAGATGATTGAGAACCTTCGCCAGACAGAACTT

AAACTTACACGAAAGGCCGCTTATGTCCGGTACTTCAACTCTTCAGCATTTTTTTTT
TAGTGGCTTCTTTGTAGTGTTCCCTGTCCGTCCTTCCGTATGCACTTATCAAGGGTA
TAATACTTAGGAAAATCTTCACAACAATCAGTTTTTGCATAGTCCTTCGCATGGC
AGTAACTCGCCAATTTCCCTGGGCAGTTCAGACGTGGTACGACTCACTTGGCGCA
ATTAACAAAATTCAAGATTTCCCTCCAAAAGCAAGAGTATAAAAACCTTGGAAATAC
AACCTTACCACCACAGAAGTTGTAATGGAAAATGTCACAGCCTTCTGGGAGGAA
GGTTTCGGCGAACTTTTTGAGAAGGCCGAAGCAAATAACAATAATCGGAAAACA
TCAAACGGTGACGATTCCTGTTCTTTTCTAACTTTAGCCTTCTTGGGACGCCCGT
CCTGAAGGACATAAACTTTAAGATTGAACGGGGTCAACTTCTCGCGGTTCGCAGG
GAGTACTGGAGCGGGGAAAACGAGCCTGCTGATGGTGATAATGGGGGAGTTGGA
GCCCTCAGAAGGCAAGATCAAGCATAGTGGTAGAATTAGCTTCTGCAGTCAATTT
AGTTGGATTATGCCGGGCACGATCAAAGAAAATATAATCTTTGGGGTATCCTACG
ATGAATACAGGTACCGATCAGTGATAAAAGCGTGCCAGCTTGAAGAAGACATTT
CAAAGTTTGCTGAGAAGGATAATATCGTACTTGGAGAAGGAGGTATCACCTGT
CTGGGGGTCAACGAGCGAGGATCTCCCTGGCACGCGCCGTCTACAAGGACGCGG
ACCTCTATCTGTTGGATTCACCGTTCGGATATTTGGACGTGCTTACGGAGAAAGA
AATATTTGAGAGCTGTGTTTGCAAGCTCATGGCAAATAAAAACCAGAATATTGGTT
ACAAGCAAGATGGAGCATCTTAAGAAAGCAGATAAAATCCTGATATTGCACGAG
GGCTCTTCATACTTCTACGGGACGTTTTCTGAGTTGCAGAACCTCCAGCCGGATTT
CAGCTCTAAGCTGATGGGCTGTGATTCCTTTGATCAGTTTAGTGCGGAAAGACGA
AACAGTATACTCACCGAAACACTGCACAGGTTCTCTCTGGAGGGCGACGCCCCG
GTTTCCTGGACAGAGACGAAGAAGCAGTCCTTCAAACAGACAGGCGAGTTTGGG
GAGAAAAGGAAAAATAGCATACTCAACCCGATTAACAGCATTTCGCAAGTTCAGT
ATAGTACAAAAGACCCCGTTGCAGATGAACGGTATAGAGGAAGATTCTGATGAG
CCACTGGAAAGACGGCTTTCTCTCGTTCCGGACAGTGAACAGGGAGAGGCAATA
CTGCCTCGGATCAGCGTTATCTCTACAGGACCTACTTTGCAAGCTCGGCGCCGAC
AGTCAGTCTTGAATCTTATGACTCATAGTGTTAATCAAGGCCAGAATATCCATCG
CAAGACCACCGCAAGTACAAGGAAAGTGAGCTTGGCACCTCAAGCAAACCTTAC
TGAACCTTGATATCTACTCACGGCGACTTTCACAGGAGACCGGACTTGAAATTAGT
GAAGAAATTAACGAGGAGGACCTCAAGGAGTGCTTCTTCGATGACATGGAATCA
ATCCCCGCAGTCACAACCTGGAACACTTATCTGAGGTATATAACAGTTCACAAGA
GCCTCATTTTTGTACTTATTTGGTGTGGTAATTTTCCCTGGCGGAGGTTGCTGCT
TCTTTGGTCGTCCTTTGGCTCCTCGGGAATACACCGCTCCAAGACAAAGGCAACT

CTACCCATAGTAGGAACAATTCATATGCAGTGATTATAACCAGTACATCATCTTA
TTACGTTTTCTATATTTATGTCGGGGTAGCTGACACGCTGTTGGCGATGGGCTTCT
TTAGGGGCCTCCCCTTGGTACACACCCTTATCACGGTGAGTAAAATCCTGCATCA
CAAAATGCTTCATTCTGTACTCCAAGCGCCGATGAGTACGCTTAATACGCTGAAA
GCAGGAGGGATACTGAATCGGTTTCAGCAAGGACATCGCCATTCTGGATGACCTG
CTTCCATTGACAATATTTGATTTTCATTTCAGCTCCTTCTCATAGTTATTGGAGCCAT
AGCGGTGGTGGCTGTGCTTCAGCCTTATATATTCGTTGCCACAGTTCCCGTTATAG
TGGCATTTATAATGCTCAGGGCCTACTTTCTCCAGACTTCCCAGCAGTTGAAGCA
ACTCGAATCAGAAGGAAGGTCACCTATTTTCACACATCTTGTGACTTCCTTGAAG
GGCTTGTGGACGCTGCGGGCCTTCGGAAGACAACCATAATTTTGAAACTCTCTTCC
ACAAAGCTTTGAATCTTCATACTGCGAACTGGTTCCTGTATTTGAGTACTTTGCGC
TGGTTCAGATGAGGATAGAAATGATATTCGTTATCTTCTTTATCGCGGTTACGTT
CATAAGTATCCTCACTACGGGGGAGGGTGAGGGTAGAGTGGGCATAAATACTGAC
CCTCGCCATGAACATTATGTCCACCCTGCAGTGGGCGGTAAACAGCAGCATAGAT
GTGGATTCTTTGATGCGCAGTGTGAGCAGGGTTTTTAAGTTTATCGATATGCCGA
CGGAAGGAAAGCCACTAAAAGCACGAAACCCTATAAAAATGGACAGCTTAGCA
AAGTAATGATAATCGAGAATAGCCATGTGAAAAAGGATGACATATGGCCTTCCG
GAGGCCAAATGACTGTTAAAGATCTGACCGCTAAATATACCGAGGGCGGCAACG
CAATACTCGAAAACATAAGCTTTTCCATAAGCCCCGGCCAACGCGTGGGTCTTCT
GGGGAGGACTGGCTCCGGAAAATCAACGTTGCTTAGCGCGTTTTTGGCGCTCCTT
AACACTGAAGGTGAGATCCAAATAGATGGCGTTAGTTGGGACTCTATAAFACTG
CAACAATGGCGGAAAGCTTTCGGCGTCATACCTCAGAAGGTGTTTCATCTTTAGCG
GAACGTTTCAGGAAGAAGCTTGGATCCCTACGAACAATGGAGTGATCAAGAAATAT
GGAAAGTGGCAGATGAGGTAGGCTTGCGCAGTGTTCATTGAACAATTCCCAGGGA
AACTCGACTTTGTACTGGTGGACGGCGGTTGCGTCTTGTACACGGGCACAAACA
GTTGATGTGTTTGGCCCAGTGTTTTGTCTAAGGCGAAGATTCTGTTGCTCGAC
GAACCGAGTGCTCATCTTGATCCCGTCACCTACCAAATCATCAGAAGGACGTTGA
AGCAAGCTTTCGCCGACTGCACTGTAATCCTTTGTGAGCATAGGATCGAAGCAAT
GCTCGAGTGCCAACAGTTCTTGGTTATAGAGGAGAATAAGGTTTCGGCAATACGA
CTCAATACAGAAACTGCTTAATGAGCGGTCACTCTTTCGACAAGCTATCTCTCCT
AGTGACAGGGTAAAGCTTTTTCTCATCGGAATTCAGCAAGTGTAAGAGTAAAC
CACAGATCGCCGCCCTTAAAGAGGAGACCGAAGAAGAGGTGCAGGATACGAGA
CTTTAG (SEQ ID NO: 40)

ЭКВИВАЛЕНТЫ

[0164] Специалисты в данной области смогут определять или с помощью лишь стандартных экспериментов смогут устанавливать множество эквивалентов конкретных вариантов осуществления изобретения, описанного в настоящем документе. Предполагается, что объем настоящего изобретения не ограничен приведенным выше описанием, а соответствует изложенному в следующих пунктах формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

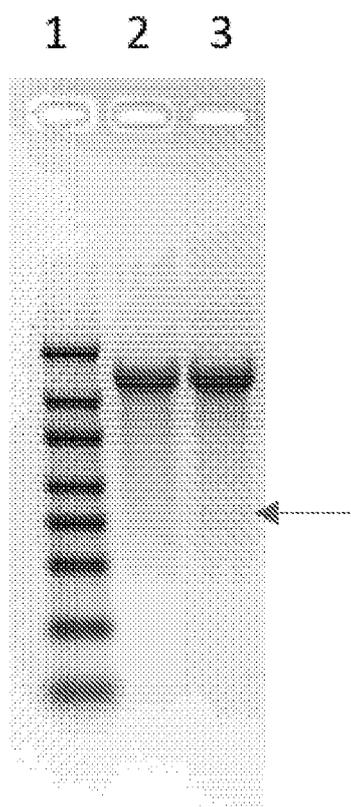
1. Фармацевтическая композиция для лечения муковисцидоза, содержащая кодон-оптимизированную мРНК, кодирующую муковисцидозный регулятор трансмембранной проводимости (CFTR), и при этом кодон-оптимизированная мРНК CFTR содержит полинуклеотидную последовательность, по меньшей мере на 85% идентичную SEQ ID NO: 1.
2. Фармацевтическая композиция по п. 1, в которой кодон-оптимизированная мРНК CFTR содержит SEQ ID NO: 1.
3. Фармацевтическая композиция по любому из пп. 1–2, в которой кодон-оптимизированная мРНК CFTR дополнительно содержит последовательность 5'-нетранслируемой области (UTR) с SEQ ID NO: 4.
4. Фармацевтическая композиция по любому из пп. 1–3, в которой кодон-оптимизированная мРНК CFTR дополнительно содержит последовательность 3'-нетранслируемой области (UTR) с SEQ ID NO: 5 или SEQ ID NO: 6.
5. Фармацевтическая композиция по любому из пп. 1–4, в которой кодон-оптимизированная мРНК CFTR, кодирующая белок CFTR, инкапсулирована внутри наночастицы.
6. Фармацевтическая композиция по п. 5, в которой наночастица представляет собой липосому.
7. Фармацевтическая композиция по п. 6, в которой липосома содержит один или более катионных липидов, один или более некаатионных липидов, один или более липидов на основе холестерина и один или более ПЭГ-модифицированных липидов.
8. Фармацевтическая композиция по п. 6, в которой липосома содержит не более трех отдельных липидных компонентов.
9. Фармацевтическая композиция по п. 6, в которой один отдельный липидный компонент представляет собой катионный липид на основе стерола.

10. Фармацевтическая композиция по любому из пп. 6–9, в которой липосома имеет размер менее чем приблизительно 100 нм.
11. Способ крупномасштабного производства кодон-оптимизированной мРНК, кодирующей муковисцидозный регулятор трансмембранной проводимости (CFTR), включающий в себя синтез *in vitro* кодон-оптимизированной мРНК CFTR с использованием РНК-полимеразы SP6,
причем по меньшей мере 80% синтезированных кодон-оптимизированных молекул мРНК CFTR являются полноразмерными, и при этом по меньшей мере 100 мг кодон-оптимизированной мРНК синтезируют в одной партии.
12. Способ по п. 11, в котором синтезированная *in vitro* кодон-оптимизированная мРНК CFTR по существу не содержит вторичного полинуклеотидного фрагмента длиной приблизительно 1800 нуклеотидов.
13. Способ по п. 11, в котором синтез *in vitro* кодон-оптимизированной мРНК CFTR приводит к получению вторичного полинуклеотидного фрагмента, который составляет менее чем 10%, 5%, 4%, 3%, 2%, 1%, 0,5%, 0,4%, 0,3%, 0,2% или 0,1% от общей синтезированной мРНК.
14. Способ по п. 11, в котором по меньшей мере 85%, 90%, 95%, 96%, 97%, 98% или 99% синтезированных кодон-оптимизированных молекул мРНК CFTR являются полноразмерными.
15. Способ по п. 11, в котором синтезированные кодон-оптимизированные молекулы мРНК CFTR являются по существу полноразмерными.
16. Способ по любому из пп. 11–15, в котором по меньшей мере 200 мг, 300 мг, 400 мг, 500 мг, 600 мг, 700 мг, 800 мг, 900 мг, 1 г, 5 г, 10 г, 25 г, 50 г, 75 г, 100 г, 150 г, 200 г, 250 г, 500 г, 750 г, 1 кг, 5 кг, 10 кг, 50 кг, 100 кг, 1000 кг или более кодон-оптимизированной мРНК CFTR синтезируют в одной партии.
17. Способ по любому из пп. 11–16, в котором белок CFTR содержит аминокислотную последовательность SEQ ID NO: 3.

18. Способ по любому из пп. 11–17, в котором кодон-оптимизированная мРНК CFTR содержит полинуклеотидную последовательность, по меньшей мере на 85% идентичную SEQ ID NO: 1.
19. Способ по любому одному из пп. 11–18, в котором кодон-оптимизированная мРНК CFTR дополнительно содержит последовательность 5'-нетранслируемой области (UTR) с SEQ ID NO: 4.
20. Способ по любому из пп. 11–19, в котором кодон-оптимизированная мРНК CFTR дополнительно содержит последовательность 3'-нетранслируемой области (UTR) с SEQ ID NO: 5 или SEQ ID NO: 6.
21. Способ по любому из пп. 11–20, дополнительно включающий в себя стадию добавления кэпа и/или хвоста к синтезированной кодон-оптимизированной мРНК CFTR.
22. Композиция, содержащая кодон-оптимизированную мРНК, кодирующую муковисцидозный регулятор трансмембранной проводимости (CFTR), синтезированный способом по любому из пп. 11–21.
23. Способ лечения муковисцидоза, включающий в себя введение нуждающемуся в лечении субъекту композиции, содержащей кодон-оптимизированную мРНК, кодирующую муковисцидозный регулятор трансмембранной проводимости (CFTR), причем кодон-оптимизированная мРНК CFTR содержит полинуклеотидную последовательность, по меньшей мере на 85% идентичную SEQ ID NO: 1.
24. Способ по п. 23, в котором кодон-оптимизированная мРНК CFTR содержит SEQ ID NO: 1.
25. Способ по любому из пп. 23–24, в котором кодон-оптимизированная мРНК CFTR дополнительно содержит последовательность 5'-нетранслируемой области (UTR) с SEQ ID NO: 4.
26. Способ по любому из пп. 23–35, в котором кодон-оптимизированная мРНК CFTR дополнительно содержит последовательность 3'-нетранслируемой области (UTR) с SEQ ID NO: 5 или SEQ ID NO: 6.

27. Способ по любому из пп. 23–26, в котором кодон-оптимизированная мРНК CFTR инкапсулирована внутри наночастицы.
28. Способ по п. 27, в котором наночастица представляет собой липосому.
29. Способ по п. 28, в котором липосома содержит один или более катионных липидов, один или более некаатионных липидов, один или более липидов на основе холестерина и один или более ПЭГ-модифицированных липидов.
30. Способ по п. 28, в котором липосома содержит не более трех отдельных липидных компонентов.
31. Способ по п. 30, в котором один отдельный липидный компонент представляет собой катионный липид на основе стерола.
32. Способ по любому из пп. 28–31, в котором липосома имеет размер менее чем приблизительно 100 нм.
33. Способ по любому из пп. 23–32, в котором кодон-оптимизированную мРНК CFTR вводят субъекту посредством доставки через легкие.
34. Способ по п. 33, в котором доставка через легкие представляет собой распыление.
35. Фармацевтическая композиция для лечения муковисцидоза, содержащая мРНК, кодирующую муковисцидозный регулятор трансмембранной проводимости (CFTR), и при этом мРНК, кодирующая белок CFTR, содержит полинуклеотидную последовательность, по меньшей мере на 85% идентичную любой из SEQ ID NO: 21–40.
36. Фармацевтическая композиция по п. 35, в которой мРНК имеет полинуклеотидную последовательность, по меньшей мере на 85% идентичную SEQ ID NO: 21, SEQ ID NO: 24, SEQ ID NO: 28, SEQ ID NO: 30 или SEQ ID NO: 35.
37. Фармацевтическая композиция по п. 35 или 36, в которой мРНК содержит SEQ ID NO: 21, SEQ ID NO: 24, SEQ ID NO: 28, SEQ ID NO: 30 или SEQ ID NO: 35.

38. Фармацевтическая композиция по любому из пп. 35–37, в которой мРНК дополнительно содержит последовательность 5'-нетранслируемой области (UTR) с SEQ ID NO: 4.
39. Фармацевтическая композиция по любому из пп. 35–38, в которой мРНК дополнительно содержит последовательность 3'-нетранслируемой области (UTR) с SEQ ID NO: 5 или SEQ ID NO: 6.
40. Фармацевтическая композиция по любому из пп. 35–39, в которой мРНК, кодирующая белок CFTR, инкапсулирована внутри наночастицы.
41. Фармацевтическая композиция по п. 40, в которой наночастица представляет собой липосому.
42. Фармацевтическая композиция по п. 41, в которой липосома содержит один или более катионных липидов, один или более некаатионных липидов, один или более липидов на основе холестерина и один или более ПЭГ-модифицированных липидов.
43. Фармацевтическая композиция по п. 41, в которой липосома содержит не более трех отдельных липидных компонентов.
44. Фармацевтическая композиция по п. 41, в которой один отдельный липидный компонент представляет собой катионный липид на основе стерола.



ФИГ. 1