

(19)



**Евразийское
патентное
ведомство**

(21) **202290863** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.07.28

(51) Int. Cl. **C12N 15/00** (2006.01)

(22) Дата подачи заявки
2020.09.12

**(54) ПОСЛЕДОВАТЕЛЬНОСТЬ N-КОНЦЕВОГО УДЛИНЕНИЯ ДЛЯ ЭКСПРЕССИИ
РЕКОМБИНАНТНЫХ ТЕРАПЕВТИЧЕСКИХ ПЕПТИДОВ**

(31) **201941009728**

(32) **2019.09.13**

(33) **IN**

(86) **PCT/IN2020/050790**

(87) **WO 2021/048878 2021.03.18**

(71) Заявитель:
БАЙОЛОДЖИКАЛ И ЛИМИТЕД (IN)

(72) Изобретатель:

Матур Рамеш Венкат, Срираман

Раджан, Регатти Паван Редди,

Мантена Нарендер Дев, Датла

Махима (IN)

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

(57) Изобретение относится к последовательностям N-концевого удлинения, которые используются для повышения экспрессии рекомбинантных терапевтических пептидов. Также изобретение относится к способу экспрессии на высоком уровне рекомбинантных терапевтических пептидов с использованием указанных последовательностей N-концевого удлинения. Также изобретение относится к нуклеиновым кислотам, векторам и рекомбинантным клеткам-хозяевам для эффективного продуцирования биологически активных белков, таких как лирапептид.

202290863

A1

A1

202290863

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420- 573952EA/061

ПОСЛЕДОВАТЕЛЬНОСТЬ N-КОНЦЕВОГО УДЛИНЕНИЯ ДЛЯ ЭКСПРЕССИИ РЕКОМБИНАНТНЫХ ТЕРАПЕВТИЧЕСКИХ ПЕПТИДОВ

ОБЛАСТЬ ИЗОБРЕТЕНИЯ

Изобретение относится к последовательности N-концевого удлинения для экспрессии на высоком уровне рекомбинантных терапевтических пептидов. Изобретение также относится к способу экспрессии на высоком уровне рекомбинантных терапевтических пептидов с использованием указанной последовательности N-концевого удлинения.

УРОВЕНЬ ТЕХНИКИ ИЗОБРЕТЕНИЯ

Пептидные терапевтические средства играют значительную роль в медицинской практике с появлением инсулиновой терапии в 1920-х. В настоящее время в продаже существует более 60 одобренных пептидных лекарственных средств, и ожидается, что их количества значительно вырастут.

Глюкагон-подобный пептид-1 (GLP-1) представляет собой пептидный гормон длиной 31 аминокислоту, образующийся в результате тканеспецифического посттрансляционного процессинга пептида проглюкагона. Он продуцируется и секретируется кишечными энтероэндокринными L-клетками и определенными нейронами в ядре солитарного тракта в стволе головного мозга в ходе приема пищи. Лираглутид является производным инкретина (метаболический гормон) человека, глюкагон-подобного пептида-1 (GLP-1), который используется в качестве агониста рецептора глюкагон-подобного пептида-1 длительного действия, связывающегося с теми же рецепторами, что и эндогенный метаболический гормон GLP-1, который стимулирует секрецию инсулина.

Терипаратид представляет собой рекомбинантную белковую форму паратиреоидного гормона, состоящую из первых (N-концевых) 34 аминокислот, которые являются биологически активной частью гормона. Он является эффективным анаболическим (способствующим формированию костей) средством, используемым для лечения некоторых форм остеопороза.

Экспрессионную плазмиду конструируют так, чтобы она содержала регуляторные последовательности, которые выступают в качестве энхансерных и промоторных областей и обеспечивают эффективную транскрипцию гена, содержащегося в экспрессионном векторе. Целью хорошо сконструированного экспрессионного вектора является эффективное продуцирование белка, и оно может достигаться посредством синтеза значительного количества стабильной матричной РНК.

Является возможным конструирование экспрессионных векторов, которые проявляют строгий контроль экспрессии и белок продуцируется в высоких количествах, только когда это необходимо, посредством использования подходящих условий экспрессии. В отсутствии строго контроля экспрессии гена белок также может

экспрессироваться конститутивно.

В патенте США № 4916212 описана последовательность ДНК, кодирующая предшественники биосинтеза инсулина, и способ получения предшественников инсулина и человеческого инсулина в дрожжевых клетках.

В патенте США № 7572884 описан способ получения рекомбинантного лирапептида, предшественника лираглутида, в *Saccharomyces cerevisiae*.

В 201741024763 А описан способ получения лираглутида посредством экспрессии синтетического олигонуклеотида, кодирующего лирапептид, который функционально связан с олигонуклеотидной последовательностью сигнального пептида в дрожжевой клетке.

В WO 1998/008871 A1 описаны производные GLP-1 и их аналоги, полученные с использованием способа рекомбинантных ДНК.

В WO 1998/008872 A1 описаны производные GLP-2, полученные с использованием способа рекомбинантных ДНК.

В WO 1999/043708 A1 описаны производные эксендина и GLP-1(7-C), полученные с использованием способа рекомбинантных ДНК.

В WO 2017/021819 A1 описан способ получения пептидов или белков, или их производных, посредством экспрессии синтетического олигонуклеотида, кодирующего желаемый белок или пептид, в прокариотической клетке в качестве слитой конструкции убиквитина.

В Avicenna J Med Biotech 2017; 9(1): 19-22 описана сверхэкспрессия терипаратида (1-34), рекомбинантной биологически активной части паратиреоидного гормона (PTH) человека, в *Escherichia coli*.

Авторы настоящего изобретения в попытках усилить экспрессию рекомбинантных терапевтических пептидов в несколько раз, предложили использовать короткую последовательность N-концевого удлинения, которая не описана в упомянутых выше документах уровня техники.

ЗАДАЧА ИЗОБРЕТЕНИЯ

Задачей настоящего изобретения является обеспечение повышенной в несколько раз экспрессии терапевтических пептидов.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

Настоящее изобретение относится к N-концевым удлинениям, нуклеиновым кислотам, векторам и рекомбинантным клеткам-хозяевам для эффективного продуцирования биологически активных пептидов, таких как лирапептид.

Изобретение охватывает многомерный подход для достижения высокого выхода пептидов, таких как лирапептид, в клетке-хозяине путем предоставления экспрессирующей конструкции, в которой нуклеиновая кислота, кодирующая лирапептид, функционально слита с модифицированной последовательностью гена, кодирующей участок расщепления TEV (вирус гравировки табака) и N-концевое удлинение (NE-3).

Настоящее изобретение относится к последовательности N-концевого удлинения,

как указано под SEQ ID NO: 1 (NE-3), для повышения экспрессии терапевтического пептида в бактериях или дрожжах.

Также настоящее изобретение относится к экспрессионным векторам и рекомбинантным клеткам-хозяевам для экспрессии на высоком уровне лирапептида, где экспрессионный вектор содержит модифицированную последовательность гена, кодирующую последовательность N-концевого удлинения NE-3, модифицированную последовательность гена, кодирующую участок расщепления TEV (вирус гравировки табака), и модифицированную последовательность гена, кодирующую лирапептид.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

Фиг.1А: Схематическая диаграмма экспрессирующих кассет без N-концевого удлинения.

Фиг.1В: Схематическая диаграмма экспрессирующих кассет с N-концевым удлинением (NE1).

Фиг.1 С: Схематическая диаграмма экспрессирующих кассет с N-концевым удлинением (NE3).

Фиг.2А: Экспрессионная плазмида без N-концевого удлинения.

Фиг.2В: Экспрессионная плазмида с N-концевым удлинением-1 (NE1).

Фиг.2С: Экспрессионная плазмида с N-концевым удлинением-3 (NE3).

Фиг.3: Экспрессия лирапептида согласно ELISA.

Фиг.4: Сухая масса клеток с лирапептидом.

ОПИСАНИЕ СПИСКА ПОСЛЕДОВАТЕЛЬНОСТЕЙ

SEQ ID NO: 1 (аминокислотная последовательность последовательности N-концевого удлинения NE-3)

EEQAE

SEQ ID NO: 2 (аминокислотная последовательность модифицированного участка расщепления TEV)

ENLYFQ

SEQ ID NO: 3 (аминокислотная последовательность лирапептида)

HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG

SEQ ID NO: 4 (последовательность нуклеиновой кислоты, кодирующая последовательность N-концевого удлинения NE-3, для *Pichia pastoris*)

gaagaacaagccga

SEQ ID NO: 5 (последовательность нуклеиновой кислоты, кодирующая модифицированный участок расщепления TEV для *Pichia pastoris*)

gagaactgtacttccaa

SEQ ID NO: 6 (последовательность нуклеиновой кислоты, кодирующая лирапептид, для *Pichia pastoris*)

cacgctgagggtactttacctctgacgtgtcctcttacttgagggtcaagctgccaaagagttcattgcctgggtggttagaggt
agaggtag

SEQ ID NO: 7 (последовательность нуклеиновой кислоты, кодирующая

последовательность N-концевого удлинения NE-3, для Corynebacterium glutamicum)

gaagaacaggcagaa

SEQ ID NO: 8 (последовательность нуклеиновой кислоты, кодирующая модифицированный участок расщепления TEV, для Corynebacterium glutamicum)

gaaaacctgtactccag

SEQ ID NO: 9 (последовательность нуклеиновой кислоты, кодирующая лирапептид, для Corynebacterium glutamicum)cacgcagaaggcacctttacctccgatgtgtcctctacctggaaggccaggcagcaaaagaattcattgcatggctggt
tcgcggtcgcggttag**SEQ ID NO: 10 (последовательность нуклеиновой кислоты, кодирующая последовательность N-концевого удлинения NE-3, для Escherichia coli)**

gaagaacaggcagaa

SEQ ID NO: 11 (последовательность нуклеиновой кислоты, кодирующая модифицированный участок расщепления TEV, для Escherichia coli)

gaaaacctgtactccag

SEQ ID NO: 12 (последовательность нуклеиновой кислоты, кодирующая лирапептид, для Escherichia coli)catgcggaaggcaccttcaccagcgatgtagcagctacctggagggtcaggcggcgaaggaatttatcgctggctgggtcgt
ggccgtggttaa**SEQ ID NO: 13 (последовательность нуклеиновой кислоты, кодирующая последовательность N-концевого удлинения NE-3, для Bacillus subtilis)**

gaagaacaagccgaa

SEQ ID NO: 14 (последовательность нуклеиновой кислоты, кодирующая модифицированный участок расщепления TEV, для Bacillus subtilis)

gagaactgtacttcaa

SEQ ID NO: 15 (последовательность нуклеиновой кислоты, кодирующая лирапептид, для Bacillus subtilis)cacgctgagggtactttacctctgacgtgtcctcttacttgagggtcaagctgccaaagagttcattgcctggttggttagaggt
agaggtag**SEQ ID NO: 16 (аминокислотная последовательность терипаратида)**

SVSEIQLMHNLGKHLNSMERVEWLRKKLQDVHNF

SEQ ID NO: 17 (аминокислотная последовательность последовательности N-концевого удлинения NE-1)

EEA

SEQ ID NO: 18 (последовательность нуклеиновой кислоты, кодирующая последовательность N-концевого удлинения NE-1)

gaggaagcg

SEQ ID NO: 19 (слитый белок, содержащий лирапептид, функционально слитый с последовательностью N-концевого удлинения NE-3 и участком расщепления TEV)

EEQAEENLYFQHAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG

SEQ ID NO: 20 (слитый белок, содержащий лирапептид, функционально слитый с последовательностью N-концевого удлинения NE-1 и участком расщепления TEV)

EEAENLYFQHAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG

ОПРЕДЕЛЕНИЯ

Если не определено иначе, все технические и научные термины, используемые в настоящем описании, обладают тем же значением, которое обычно подразумевает специалист в области, к которой относятся способы. Хотя для применения на практике или тестирования векторов, клеток-хозяев, способов и композиций также можно использовать любые векторы, клетки-хозяева, способы и композиции, сходные или эквивалентные тем, которые описаны в настоящем описании, далее описаны репрезентативные примеры.

Когда предоставлен диапазон величин, понятно, что каждая встроенная величина между верхним и нижним пределом этого диапазона и любая другая указанная или промежуточная величина в этом указанном диапазоне охватывается способами и композициями. Верхние и нижние пределы этих меньших диапазонов могут быть независимо включены в меньшие диапазоны и также охватываются способами и композициями, за исключением какого-либо конкретно исключенного предела указанного диапазона. Когда указанный диапазон включает один или оба из пределов, диапазоны, не включающие один или оба из этих включенных пределов, также включены в способы и композиции.

Следует принять во внимание, что определенные признаки способов, которые для ясности описаны в контексте отдельных вариантов осуществления, также могут быть предоставлены в комбинации в одном варианте осуществления. Напротив, различные признаки способов и композиций, которые для краткости описаны в контексте одного варианта осуществления, также могут быть предоставлены отдельно или в любой подходящей субкомбинации. Следует отменить, что, как используют в рамках изобретения и в прилагаемой формуле изобретения, форма единственного числа включает множественное число упоминаемых объектов, если контекст явно не определяет иное. Кроме того, следует отметить, что формула изобретения может быть составлена для исключения какого-либо необязательного элемента. Таким образом, это утверждение предназначено для того, чтобы служить в качестве антецедентной основы для применения такой исключающей терминологии, как "единственно", "только" и т.п. применительно к перечислению элементов формулы изобретения или применению "негативного" ограничения.

Как будет понятно специалистам в данной области при прочтении настоящего описания, каждый из индивидуальных вариантов осуществления, описанных и проиллюстрированных в настоящем описании, имеет характерные компоненты и признаки, которые могут быть без труда отделены от или скомбинированы с признаками

любого из других вариантов осуществления без отклонения от объема или сущности настоящих способов. Любой приведенный способ можно проводить в указанном порядке событий или в любом другом порядке, который является логически возможным.

Термин "клетка-хозяин" включает индивидуальную клетку или клеточную культуру, которая может быть или является реципиентом рассматриваемых экспрессирующих конструкций. Клетки-хозяева включают потомство единичной клетки-хозяина. Клетка-хозяин для целей настоящего изобретения относится к любому штамму *Pichia pastoris*, *Saccharomyces cerevisiae*, *Corynebacterium glutamicum*, *Escherichia coli* и *Bacillus subtilis*, который может подходящим образом использоваться для целей настоящего изобретения.

Термин "рекомбинантный штамм" или "рекомбинантная клетка-хозяин" относится к клетке-хозяину, которая трансфицирована или трансформирована экспрессирующими конструкциями или векторами по настоящему изобретению.

Термин "экспрессионный вектор" или "экспрессирующая конструкция" относится к любому вектору, плазмиде или носителю, сконструированному для обеспечения экспрессии встроенной последовательности нуклеиновой кислоты после трансформации хозяина.

Термин "промотор" относится к последовательностям ДНК, которые определяют, где начинается транскрипция гена. Промоторные последовательности, как правило, находятся непосредственно выше или на 5'-конце участка инициации транскрипции. РНК-полимераза и необходимые факторы транскрипции связываются с промоторной последовательностью и иницируют транскрипцию. Промоторы могут представлять собой либо конститутивные, либо индуцибельные промоторы. Конститутивные промоторы представляют собой промоторы, которые позволяют непрерывную транскрипцию ассоциированных с ними генов, поскольку их экспрессия обычно не обуславливается факторами внешней среды или факторами стадии развития. Конститутивные промоторы являются в высокой степени полезными инструментами генной инженерии, поскольку конститутивные промоторы обеспечивают экспрессию генов в свободных от индуктора условиях и часто демонстрируют лучшие характеристики, чем часто используемые индуцибельные промоторы. Индуцибельные промоторы представляют собой промоторы, которые индуцируются присутствием или отсутствием биотических или абиотических и химических или физических факторов. Индуцибельные промоторы являются очень мощным инструментом генной инженерии, поскольку экспрессия генов, функционально связанных с ними, может быть включена или выключена на определенных стадиях развития или роста организма или в конкретной ткани или клетках.

Термин "экспрессия" относится к биологическому продуцированию продукта, кодируемого кодирующей последовательностью. В большинстве случаев, последовательность ДНК, включающая кодирующую последовательность, транскрибируется с образованием матричной РНК (мРНК). Затем матричная РНК

транслируется с образованием полипептидного продукта, который имеет соответствующую биологическую активность. Также процесс экспрессии может вовлекать дальнейшие стадии процессинга в РНК-продукт транскрипции, такие как сплайсинг для удаления интронов и/или посттрансляционный процессинг полипептидного продукта.

Термин "модифицированная нуклеиновая кислота", как используют в рамках изобретения, используется для указания на нуклеиновую кислоту, кодирующую модифицированный лирапептид, соответствующий SEQ ID NO: 19 или 20, или его функционально эквивалентный вариант. Функциональный вариант включает любую нуклеиновую кислоту, имеющую существенную или значительную идентичность или сходство последовательности с SEQ ID NO: 19 или 20, и сохраняющую биологическую активность белка.

Термины "полипептид", "пептид" и "белок" используются в настоящем описании взаимозаменяемо для указания на два или более аминокислотных остатков, соединенных друг с другом пептидными связями или модифицированными пептидными связями. Эти термины применимы к полимерам аминокислот, в которых один или несколько аминокислотных остатков являются искусственными миметиками соответствующей встречающейся в природе аминокислоты, а также к встречающимся в природе полимерам аминокислот, полимерам аминокислот, содержащим модифицированные остатки, и к не встречающемуся в природе полимеру аминокислот. "Полипептид" относится как к коротким цепям, часто называемым пептидами, олигопептидами или олигомерами, и к более длинным цепям, обычно упоминаемым как белки. Полипептиды могут содержать аминокислоты, отличные от 20 кодируемых генами аминокислот. Аналогично, "белок" относится по меньшей мере к двум ковалентно связанным аминокислотам, которые включают белки, полипептиды, олигопептиды и пептиды. Белок может состоять из встречающихся в природе аминокислот и пептидных связей, или синтетических структур-пептидомиметиков. Таким образом, "аминокислота" или "пептидный остаток", как используют в рамках изобретения, означает как встречающиеся в природе, так и синтетические аминокислоты. "Аминокислота" включает остатки аминокислот, такие как пролин и гидроксипролин. Боковые цепи могут иметь конфигурацию либо (R), либо (S).

Термин "N-концевое удлинение" относится к пептидной или полипептидной последовательности, которая удаляемым образом связана с N-концевой аминокислотой желаемого полипептида. В предпочтительном варианте осуществления N-концевое удлинение содержит аминокислотную последовательность SEQ ID NO: 1.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

Настоящее изобретение относится к N-концевым удлинениям, нуклеиновым кислотам, векторам и рекомбинантным клеткам-хозяевам для эффективного продуцирования биологически активных пептидов, таких как лирапептид.

Изобретение относится к многомерному подходу для достижения высокого выхода рекомбинантного лирапептида в клетке-хозяина посредством предоставления

экспрессирующей конструкции, в которой нуклеиновая кислота, кодирующая лирапептид, функционально слита с модифицированной последовательностью гена, кодирующей участок расщепления TEV (вирус гравировки табака) и N-концевое удлинение (NE-3).

В одном варианте осуществления изобретение относится к последовательности N-концевого удлинения, указанной в SEQ ID NO: 1 (NE-3). Изобретение относится к способу повышения в несколько раз экспрессии рекомбинантных терапевтических пептидов с использованием короткой последовательности N-концевого удлинения, указанной в SEQ ID NO: 1, в бактериях или дрожжах.

В другом варианте осуществления также объемом изобретения охватываются нуклеиновые кислоты, кодирующие последовательность N-концевого удлинения, указанную в SEQ ID NO: 1 (NE-3).

Подходящая клетка-хозяин для экспрессии рекомбинантного терапевтического пептида выбрана из эукариотических хозяев, таких как, но не ограничиваясь ими, дрожжи, которые включают *Pichia pastoris* и *Saccharomyces cerevisiae*. Также могут использоваться бактериальные хозяева, такие как, но не ограничиваясь ими, *Corynebacterium glutamicum*, *Escherichia coli* и *Bacillus subtilis*.

Термин "терапевтический пептид" относится к пептидам, таким как, но не ограничиваясь ими, дирапептид, терипаратид, эксенатид и т.п.

В одном или нескольких вариантах осуществления настоящего изобретения в экспрессирующих кассетах могут использоваться конститутивные или индуцибельные промоторы, известные специалисту в данной области.

В другом варианте осуществления настоящее изобретение относится к экспрессирующим кассетам, содержащим промотор, сигнальную последовательность, N-концевое удлинение (NE-3), ген, кодирующий лирапептид или терипаратид, участок расщепления TEV и терминатор.

В одном варианте осуществления настоящее изобретение относится к последовательности N-концевого удлинения, как указано под SEQ ID NO: 1, для повышения экспрессии терапевтического пептида в дрожжах, где дрожжи представляют собой *Pichia pastoris*, *Saccharomyces cerevisiae*, *Corynebacterium glutamicum*, *Escherichia coli* и *Bacillus subtilis*.

Настоящее изобретение также относится к участку расщепления TEV, имеющему аминокислотную последовательность, указанную под SEQ ID NO: 2.

В одном варианте осуществления настоящее изобретение относится к повышенной экспрессии лирапептида, как указано под SEQ ID NO: 3, и терипаратида, как указано под SEQ ID NO: 16, в бактериях или дрожжах с использованием последовательности N-концевого удлинения, как указано под SEQ ID NO: 1.

Экспрессирующие конструкции, известные специалисту в данной области, для экспрессии прокариотических или эукариотических белков могут использоваться в одном или нескольких вариантах осуществления настоящего изобретения.

В одном варианте осуществления настоящее изобретение также относится к

экспрессирующей конструкции для экспрессии на высоком уровне лирапептида, которая содержит:

1. модифицированную последовательность гена, кодирующую последовательность N-концевого удлинения (NE3),
2. модифицированную последовательность гена, кодирующую участок расщепления TEV (вирус гравировки табака), и
3. модифицированную последовательность гена, кодирующую лирапептид.

В другом варианте осуществления настоящее изобретение также относится к экспрессирующей конструкции для экспрессии на высоком уровне лирапептида, которая содержит:

1. последовательность гена, кодирующую последовательность N-концевого удлинения (NE-3), как указано под SEQ ID NO: 4,
2. последовательность гена, кодирующую участок расщепления TEV (вирус гравировки табака), как указано под SEQ ID NO: 5, и
3. последовательность гена, кодирующую лирапептид, как указано под SEQ ID NO: 6.

В другом варианте осуществления настоящее изобретение также относится к способу экспрессии на высоком уровне лирапептида, как указано под SEQ ID NO: 3, в *Pichia Pastoris*, который включает:

1. конструирование рекомбинантного вектора (экспрессирующая конструкция), содержащего последовательности гена, как указано под SEQ ID NO: 4, 5 и 6,
2. трансформацию экспрессирующей конструкции в *Pichia Pastoris*,
3. оценку клона и его селекцию,
4. проведение процесса ферментации отобранных клонов,
5. выделение и очистку лирапептида, и
6. отщепление последовательности N-концевого удлинения от очищенного лирапептида.

В другом варианте осуществления настоящее изобретение также относится к экспрессирующей конструкции для экспрессии на высоком уровне лирапептида, которая содержит:

1. последовательность гена, кодирующую последовательность N-концевого удлинения, как указано под SEQ ID NO: 7,
2. последовательность гена, кодирующую участок расщепления TEV, как указано под SEQ ID NO: 8, и
3. последовательность гена, кодирующую лирапептид, как указано под SEQ ID NO: 9.

В одном из вариантов осуществления настоящее изобретение также относится к способу экспрессии на высоком уровне лирапептида, как указано под SEQ ID NO: 3, в *Corynebacterium glutamicum*, который включает:

1. конструирование рекомбинантного вектора (экспрессирующей конструкции),

содержащего последовательности генов, как указано под SEQ ID NO: 7, 8 и 9,

2. трансформацию экспрессирующей конструкции в *Corynebacterium glutamicum*,
3. оценку клона и его селекцию,
4. проведение процесса ферментации отобранных клонов,
5. выделение и очистку лирапептида, и
6. отщепление последовательности N-концевого удлинения от очищенного лирапептида.

Настоящее изобретение относится к экспрессирующей конструкции для экспрессии на высоком уровне лирапептида, которая содержит:

1. последовательность гена, кодирующую последовательность N-концевого удлинения, как указано под SEQ ID NO: 10,
2. последовательность гена, кодирующую участок расщепления TEV, как указано под SEQ ID NO: 11, и
3. последовательность гена, кодирующую лирапептид, как указано под SEQ ID NO: 12.

В одном из вариантов осуществления настоящее изобретение также относится к способу экспрессии на высоком уровне лирапептида, как указано под SEQ ID NO: 3, в *Escherichia coli*, который включает:

1. конструирование рекомбинантного вектора (экспрессирующей конструкции), содержащего последовательности генов, как указано под SEQ ID NO: 10, 11 и 12,
2. трансформацию экспрессирующей конструкции в *Escherichia coli*,
3. оценку клона и его селекцию,
4. проведение процесса ферментации отобранных клонов,
5. выделение и очистку лирапептида, и
6. отщепление последовательности N-концевого удлинения от очищенного лирапептида.

Настоящее изобретение относится к экспрессирующей конструкции для экспрессии на высоком уровне лирапептида, которая содержит:

1. последовательность гена, кодирующую N-концевое удлинение последовательность, как указано под SEQ ID NO: 13,
2. последовательность гена, кодирующую участок расщепления TEV, как указано под SEQ ID NO: 14, и
3. последовательность гена, кодирующую лирапептид, как указано под SEQ ID NO: 15.

В одном из вариантов осуществления настоящее изобретение также относится к способу экспрессии на высоком уровне лирапептида, как указано под SEQ ID NO: 3, в *Bacillus subtilis*, который включает:

1. конструирование рекомбинантного вектора (экспрессирующей конструкции), содержащего последовательности генов, как указано под SEQ ID NO: 13, 14 и 15,
2. трансформацию экспрессирующей конструкции в *Bacillus subtilis*,

3. оценку клона и его селекцию,
4. проведение процесса ферментации отобранных клонов,
5. выделение и очистку лирапептида, и
6. отщепление последовательности N-концевого удлинения от очищенного лирапептида.

Настоящее изобретение относится к экспрессии на высоком уровне терипаратида, как указано под SEQ ID NO: 16, в *Corynebacterium glutamicum* с использованием экспрессирующей конструкции, содержащей последовательность N-концевого удлинения.

Настоящее изобретение относится к экспрессии на высоком уровне терипаратида, как указано под SEQ ID NO: 16, в *Pichia pastoris* с использованием экспрессирующей конструкции, содержащей последовательность N-концевого удлинения.

Настоящее изобретение относится к экспрессии на высоком уровне терипаратида, как указано под SEQ ID NO: 16, в *Corynebacterium glutamicum*, которая включает следующие стадии:

1. конструирование рекомбинантного вектора (экспрессирующей конструкции)
2. трансформацию экспрессирующей конструкции в *Corynebacterium glutamicum*,
3. оценку клона и его селекцию,
4. проведение процесса ферментации отобранных клонов,
5. выделение и очистку терипаратида, и
6. отщепление последовательности N-концевого удлинения от очищенного терипаратида,

Настоящее изобретение также относится к последовательности N-концевого удлинения, как указано под SEQ ID NO: 1, для повышения экспрессии терипаратида в *Pichia pastoris*, которое включает следующие стадии:

1. конструирование рекомбинантного вектора (экспрессирующей конструкции),
2. трансформация сконструированного вектора в *Pichia pastoris*,
3. оценка клона и его селекция,
4. проведение процесса ферментации отобранных клонов,
5. выделение и очистка терипаратида, и
6. отщепление последовательности N-концевого удлинения от очищенного терипаратида.

В другом варианте осуществления изобретение относится к модифицированному лирапептиду, где лирапептид функционально слит с участком расщепления TEV (вирус гравировки табака) и последовательностью N-концевого удлинения (NE-3), и где модифицированный лирапептид является таким, как указано под SEQ ID NO: 19.

В другом варианте осуществления изобретение относится к способу экспрессии лирапептида с использованием рекомбинантных клеток-хозяев по настоящему изобретению, где процесс ферментации включает:

- а. культивирование рекомбинантных клеток-хозяев в среде BMGY в течение приблизительно 24 ч;

- b. сбор рекомбинантных клеток-хозяев посредством центрифугирования;
- c. ресуспендирование рекомбинантных клеток-хозяев до OD_{600nm} приблизительно 10 в среде BMMY;
- d. инкубацию клеток-хозяев в инкубаторе с перемешиванием в течение приблизительно 24 ч при 30°C;
- e. сбор и очистку культуральных супернатантов с получением лирапептида.

Лираглутид является аналогом GLP-1 человека и действует в качестве агониста рецептора GLP-1. Лираглутид получают путем присоединения C-16 жирной кислоты (пальмитиновая кислота) со спейсером из глутаминовой кислоты на оставшемся остатке лизина в положении 26 пептидного предшественника (лирапептид, как указано под SEQ ID NO: 3).

В другом варианте осуществления изобретение относится к получению лираглутида, которое включает конъюгацию лирапептида, полученного в соответствии с изобретением, с производным пальмитилглутамата, таким как 1-метилпальмитилглутаминовая кислота, с использованием способов, известных в данной области.

В другом варианте осуществления изобретение относится к получению лираглутида, которое включает конъюгацию лирапептида, полученного в соответствии с изобретением, с производными пальмитилглутамата, где производные являются такими, как производные с метилом (1-метилпальмитилглутаминовая кислота), этилом, пропилом, проп-2-илом, бутилом, бут-2-илом, 2-метилпроп-1-илом, 2-метил-проп-2-илом (трет-бутилом), гексилом и т.п., с использованием способов, известных в данной области. Эту реакцию конъюгации проводят в присутствии конденсирующего реагента. Конденсирующий реагент может быть выбран из группы DIC/6-Cl-HOBt, DIC/HOBt, HBTU/HOBt/DIEA или DIC/Oxyma.

В другом варианте осуществления изобретение относится к способу получения лираглутида, причем указанный способ включает стадии:

- a. культивирование рекомбинантной клетки-хозяина по настоящему изобретению в подходящей культуральной среде с получением лирапептида;
- b. конвертирование лирапептида в лираглутид, где способ включает конъюгацию лирапептида, полученного на стадии (a), с производным пальмитилглутамата.

Приведенное выше описание описывает в общем настоящее изобретение. Более полное понимание может быть получено при помощи приведенных ниже конкретных примеров. Этот пример описан только для иллюстрации и не предназначен для ограничения объема изобретения. Хотя здесь используются конкретные термины, такие термины имеют описательное значение и не предназначены для целей ограничения.

ПРИМЕРЫ

Пример 1: Модифицированная нуклеиновая кислота для экспрессии лирапептида

Экспрессирующие кассеты, кодирующие пептид-предшественник лираглутида,

модифицировали для оптимальной экспрессии в *Pichia pastoris*, *Corynebacterium glutamicum*, *Escherichia coli* и *Bacillus subtilis*. Модифицированная открытая рамка считывания содержала нуклеотидную последовательность, кодирующую лирапептид, слитый с последовательностью, кодирующей участок расщепления TEV (вирус гравировки табака), и последовательность, кодирующую N-концевое удлинение (NE-3 или NE-1). Вместо редких кодонов использовались предпочтительные кодоны для экспрессии в *Pichia pastoris*, *Corynebacterium glutamicum*, *Escherichia coli* и *Bacillus subtilis*.

В качестве контроля получали открытую рамку считывания, содержащую нуклеотидную последовательность, кодирующую лирапептид, без какого-либо N-концевого удлинения.

Для экспрессии в *Pichia pastoris*

Для экспрессии в *Pichia Pastoris* проводили модификацию нуклеотидной последовательности, кодирующей лирапептид, нуклеотидной последовательности, кодирующей участок расщепления TEV, и нуклеотидной последовательности, кодирующей N-концевое удлинение.

Нуклеотидная последовательность, кодирующая лирапептид, представлена под SEQ ID NO: 6. Нуклеотидная последовательность участка расщепления TEV представлена под SEQ ID NO: 5. Нуклеотидная последовательность, кодирующая N-концевое удлинение (NE-3), представлена под SEQ ID NO: 4.

Эта модифицированная открытая рамка считывания была искусственно синтезирована с использованием последовательности лирапептида, последовательности участка расщепления TEV и последовательности N-концевого удлинения.

Модифицированная открытая рамка считывания, содержащая ДНК, кодирующую лирапептид без N-концевого удлинения (фиг.1A) и с N-концевым удлинением 1 (NE1) (GAGGAAGCG - фиг.1B), N-концевым удлинением 3 (NE3) (GAAGAACAAGCCGAA - фиг.1C), вместе с последовательностью расщепления TEV (вирус гравировки табака) (GAGAACTTGTA CT TCCAA) и кассетами сигнальных последовательностей представлены на фиг.1.

Модифицированную последовательность, кодирующую рекомбинантный лирапептид, клонировали в экспрессионный вектор pD912 (Atum, США). Рекомбинантная плазмида содержит открытую рамку считывания и промотор.

Карта вектора pD912 представлена на фиг.1.

Экспрессия в *Corynebacterium glutamicum*

Для экспрессии в *Corynebacterium glutamicum* проводили модификацию нуклеотидной последовательности, кодирующей лирапептид, нуклеотидной последовательности, кодирующей участок расщепления TEV, и нуклеотидной последовательности, кодирующей N-концевое удлинение.

Нуклеотидная последовательность, кодирующая лирапептид, представлена под SEQ ID NO: 9. Нуклеотидная последовательность участка расщепления TEV представлена под SEQ ID NO: 8. Нуклеотидная последовательность, кодирующая N-концевое

удлинение (NE-3), представлена под SEQ ID NO: 7.

Эта модифицированная открытая рамка считывания была искусственно синтезирована с использованием способа Thermo Fisher Scientific с использованием последовательности лирапептида, последовательности участка расщепления TEV и последовательности N-концевого удлинения.

Эту модифицированную последовательность, кодирующую рекомбинантный лирапептид, клонировали в экспрессионный вектор pD912 (Atum, США). Рекомбинантная плазмида содержит открытую рамку считывания и промотор.

Экспрессия в *Escherichia coli*

Для экспрессии в *Escherichia coli* проводили модификацию нуклеотидной последовательности, кодирующей лирапептид, нуклеотидной последовательности, кодирующей участок расщепления TEV, и нуклеотидной последовательности, кодирующей N-концевое удлинение.

Нуклеотидная последовательность, кодирующая лирапептид, представлена под SEQ ID NO: 12. Нуклеотидная последовательность участка расщепления TEV представлена под SEQ ID NO: 11. Нуклеотидная последовательность, кодирующая N-концевое удлинение (NE-3), представлена под SEQ ID NO: 10.

Эта модифицированная открытая рамка считывания была искусственно синтезирована с использованием последовательности лирапептида, последовательности участка расщепления TEV и последовательности N-концевого удлинения.

Модифицированную последовательность, кодирующую рекомбинантный лирапептид, клонировали в экспрессионный вектор pD912 (Atum, США). Рекомбинантная плазмида содержит открытую рамку считывания и промотор.

Экспрессия в *Bacillus subtilis*

Для экспрессии в *Bacillus subtilis* проводили модификацию нуклеотидной последовательности, кодирующей лирапептид, нуклеотидной последовательности, кодирующей участок расщепления TEV, и нуклеотидной последовательности, кодирующей N-концевое удлинение.

Нуклеотидная последовательность, кодирующая лирапептид, представлена под SEQ ID NO: 15. Нуклеотидная последовательность участка расщепления TEV представлена под SEQ ID NO: 14. Нуклеотидная последовательность, кодирующая N-концевое удлинение (NE-3), представлена под SEQ ID NO: 13.

Эту модифицированную открытую рамку считывания искусственно синтезировали с использованием последовательности лирапептида, последовательности участка расщепления TEV и последовательности N-концевого удлинения.

Модифицированную последовательность, кодирующую рекомбинантный лирапептид, клонировали в экспрессионный вектор pD912 (Atum, США). Рекомбинантная плазмида содержит открытую рамку считывания и промотор.

Подтверждение линеаризации плазмидной ДНК

Синтетическую ДНК, кодирующую лирапептид, без N-концевого удлинения, N-

концевого удлинения 1, N-концевого удлинения 3 и плазмиду pD912, расщепляли ферментами рестрикции EcoRI и BglII. Расщепленные фрагменты рестрикции лигировали и трансформировали в штамм *Escherichia coli*. Полученные плазмиды, содержащие экспрессирующие кассеты лирапептида без N-концевого удлинения (фиг.2А), N-концевого удлинения 1(фиг.2В), N-концевого удлинения 3 (фиг.2С) секвенировали для подтверждения последовательности лирапептида, N-концевого удлинения и последовательности расщепления TEV. Плазмидную ДНК с подтвержденной последовательностью линеаризовывали посредством фермента Sac I.

Пример 2: Получение рекомбинантной клетки-хозяина посредством трансформации рекомбинантными плазмидами

Рекомбинантные плазмиды pD912, как описано в вышеуказанном примере, содержащие ген пептида-предшественника лираглутида, слитого с сигнальными пептидами, использовали для получения рекомбинантных хозяев.

Клетки-хозяева *Pichia pastoris* (полученные от Atum, США) трансформировали с использованием плазмид способом электропорации.

Трансформированные клетки высевали на агар YPD (дрожжевой пептон-декстроза), содержащий 100 мкг/мл зеоцина. Трансформированные клетки *Pichia pastoris* выращивали в 20 мл среды BMGY в течение 24 ч. Клетки собирали посредством центрифугирования и ресуспендировали до OD_{600nm} 10 в 20 мл среды BMMY. Клеточную суспензию инкубировали в инкубаторе с перемешиванием в течение 24 ч при 30°C.

Пример 3: Анализ и оценка экспрессии лирапептида

Через 24 часа культуральные супернатанты собирали, очищали и анализировали в отношении экспрессии лирапептида посредством ELISA (фиг.3) с использованием моноклонального антитела, специфичного к лирапептиду. Далее проводили измерение сухой клеточной массы с использованием анализатора влажности (фиг.4).

В таблице 1 и таблице 2 приведено сравнение, показывающее эффективность N-концевых удлинений в отношении повышения выхода лирапептида.

Таблица 1: Экспрессия различных N-концевых удлинений и лирапептида по сравнению с контролем без N-концевого удлинения

Удлинение	Клоны	Экспрессия LP (процент от контроля)	OD при 450 нм
Нет (контроль)		100	0,52
EEA	1	103	0,561
EEA	2	75	0,409
EEA	3	87	0,474
EEQAE	1	459	2,488
EEQAE	2	479	2,598
EEQAE	3	443	2,4

Таблица 2: Различные N-концевые удлинения и кратность изменения экспрессии лирапептида по сравнению с контролем без N-концевого удлинения

Удлинение	Клоны	Кратность отличий
EEA	1	1,0
EEA	2	0,8
EEA	3	0,9
EEQAE	1	4,6
EEQAE	2	4,8
EEQAE	3	4,4

Описанные выше данные отчетливо демонстрируют, что N-концевое удлинение способно повышать экспрессию лирапептида приблизительно в 5 раз по сравнению с контролем и известными N-концевыми удлинениями.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. N-концевое удлинение, содержащее аминокислотную последовательность SEQ ID NO: 1.
2. Нуклеиновая кислота, кодирующая N-концевое удлинение по п.1.
3. Нуклеиновая кислота по п.2, где нуклеиновая кислота выбрана из группы, включающей SEQ ID NO: 4, SEQ ID NO: 7, SEQ ID NO: 10 и SEQ ID NO: 13.
4. Вектор, содержащий модифицированную нуклеиновую кислоту по п.3.
5. Вектор по п.4, где вектор представляет собой pD912.
6. Вектор по п.4, где вектор содержит модифицированный участок расщепления TEV, выбранный из группы, включающей SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 и SEQ ID NO: 14.
7. Вектор для рекомбинантной экспрессии лирапептида, содержащий:
 - a. модифицированную последовательность гена, кодирующую последовательность N-концевого удлинения (NE-3), выбранную из группы, включающей SEQ ID NO: 4, SEQ ID NO: 7, SEQ ID NO: 10 и SEQ ID NO: 13;
 - b. модифицированную последовательность гена, кодирующую участок расщепления TEV (вирус гравировки табака), выбранную из группы, включающей SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 11 и SEQ ID NO: 14; и
 - c. модифицированную последовательность гена, кодирующую лирапептид, выбранную из группы, включающей SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 12 и SEQ ID NO: 15.
8. Вектор по п.7, где вектор модифицирован для экспрессии на высоком уровне лирапептида в *Pichia pastoris*, включающий:
 - a. модифицированную последовательность гена, кодирующую последовательность N-концевого удлинения (NE-3), включающую нуклеотидную последовательность SEQ ID NO: 4;
 - b. модифицированную последовательность гена, кодирующую участок расщепления TEV (вирус гравировки табака), включающую нуклеотидную последовательность SEQ ID NO: 5; и
 - c. модифицированную последовательность гена, кодирующую лирапептид, включающую нуклеотидную последовательность SEQ ID NO: 6.
9. Вектор по п.7, где вектор модифицирован для экспрессии на высоком уровне лирапептида в *Corynebacterium glutamicum*, включающий:
 - a. модифицированную последовательность гена, кодирующую последовательность N-концевого удлинения (NE-3), включающую нуклеотидную последовательность SEQ ID NO: 7;
 - b. модифицированную последовательность гена, кодирующую участок расщепления TEV (вирус гравировки табака), включающую нуклеотидную последовательность SEQ ID NO: 8; и
 - c. модифицированную последовательность гена, кодирующую лирапептид,

включающую нуклеотидную последовательность SEQ ID NO: 9.

10. Вектор по п.7, где вектор модифицирован для экспрессии на высоком уровне лирапептида в *Escherichia coli*, включающий:

а. модифицированную последовательность гена, кодирующую последовательность N-концевого удлинения (NE-3), включающую нуклеотидную последовательность SEQ ID NO:10;

б. модифицированную последовательность гена, кодирующую участок расщепления TEV (вирус гравировки табака), включающую нуклеотидную последовательность SEQ ID NO: 11; и

с. модифицированную последовательность гена, кодирующую лирапептид, включающую нуклеотидную последовательность SEQ ID NO: 12.

11. Вектор по п.7, где вектор модифицирован для экспрессии на высоком уровне лирапептида в *Bacillus subtilis*, включающий:

а. модифицированную последовательность гена, кодирующую последовательность N-концевого удлинения (NE-3), включающую нуклеотидную последовательность SEQ ID NO:13;

б. модифицированную последовательность гена, кодирующую участок расщепления TEV (вирус гравировки табака), включающую нуклеотидную последовательность SEQ ID NO: 14; и

с. модифицированную последовательность гена, кодирующую лирапептид, включающую нуклеотидную последовательность SEQ ID NO: 15.

12. Рекомбинантная клетка-хозяин, содержащая вектор по п.7.

13. Рекомбинантная клетка-хозяин по п.12, где рекомбинантная клетка-хозяин выбрана из группы, включающей *Pichia pastoris*, *Saccharomyces cerevisiae*, *Corynebacterium glutamicum*, *Escherichia coli* и *Bacillus subtilis*.

14. Модифицированный лирапептид, содержащий аминокислотную последовательность SEQ ID NO: 19, где лирапептид функционально слит с участком расщепления TEV (вирус гравировки табака) и последовательностью N-концевого удлинения (NE-3).

15. Способ экспрессии лирапептида с использованием рекомбинантных клеток-хозяев по п.12, где процесс ферментации включает:

а. культивирование рекомбинантных клеток-хозяев в среде BMGY в течение приблизительно 24 ч;

б. сбор рекомбинантных клеток-хозяев посредством центрифугирования;

с. ресуспендирование рекомбинантных клеток-хозяев до OD_{600nm} приблизительно 10 в среде BMMY;

д. инкубацию клеток-хозяев в инкубаторе с перемешиванием в течение приблизительно 24 ч при 30°C; и

е. сбор и очистку культуральных супернатантов с получением лирапептида.

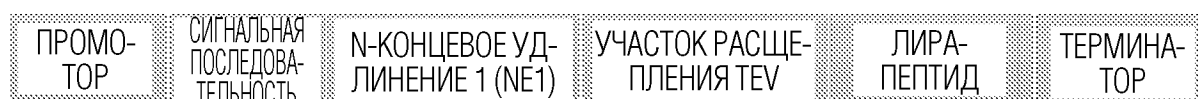
16. Способ получения лираглутида, причем указанный способ включает стадии:

с. культивирования рекомбинантной клетки-хозяина по п.12 в подходящей культуральной среде с получением лирапептида;

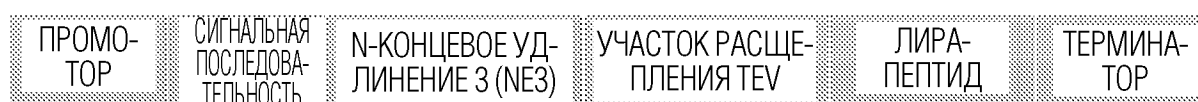
d. конвертирование лирапептида в лираглутид, где способ включает конъюгацию лирапептида, полученного на стадии (а), с производным пальмитилглутамата.



А

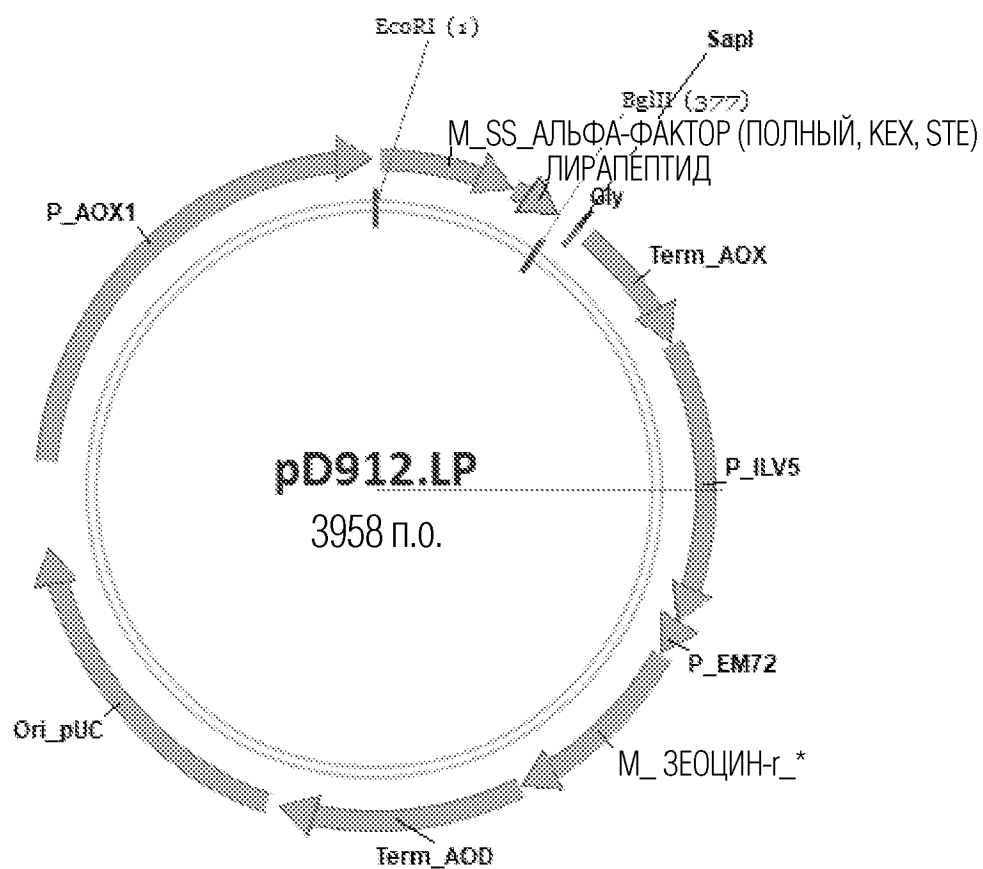


В

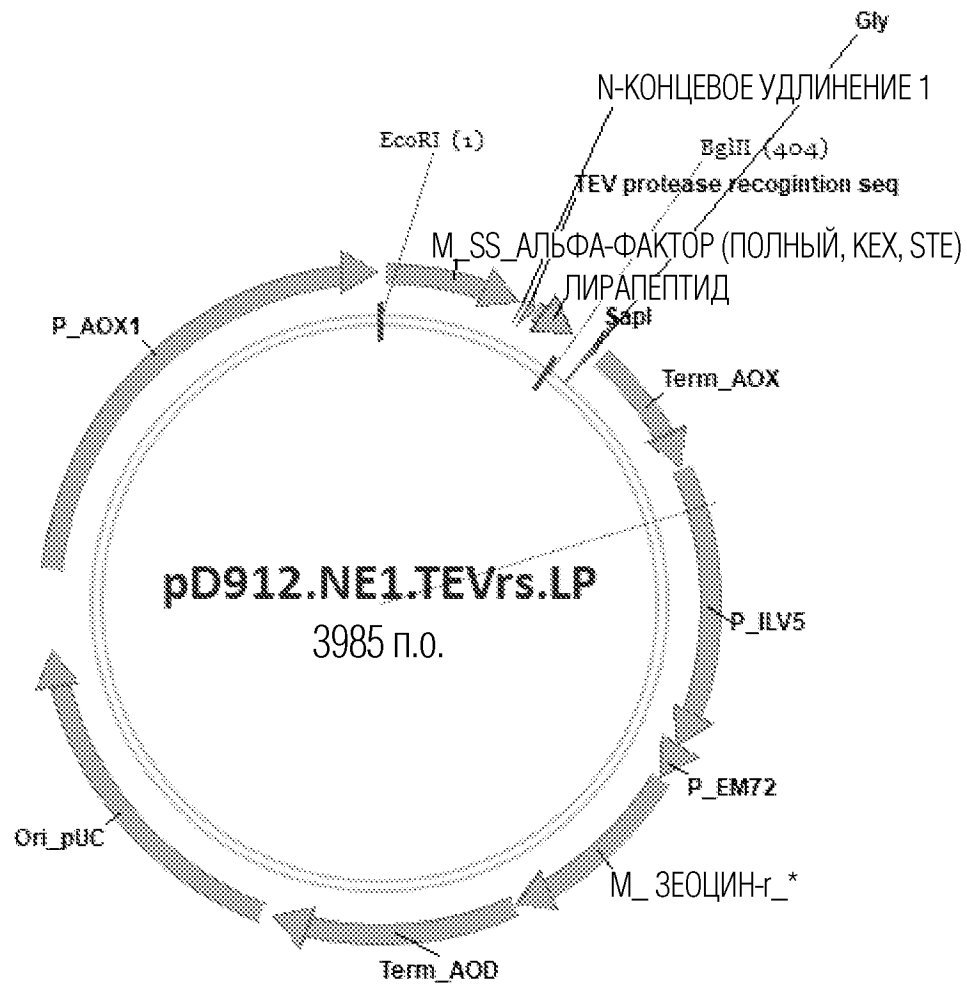


С

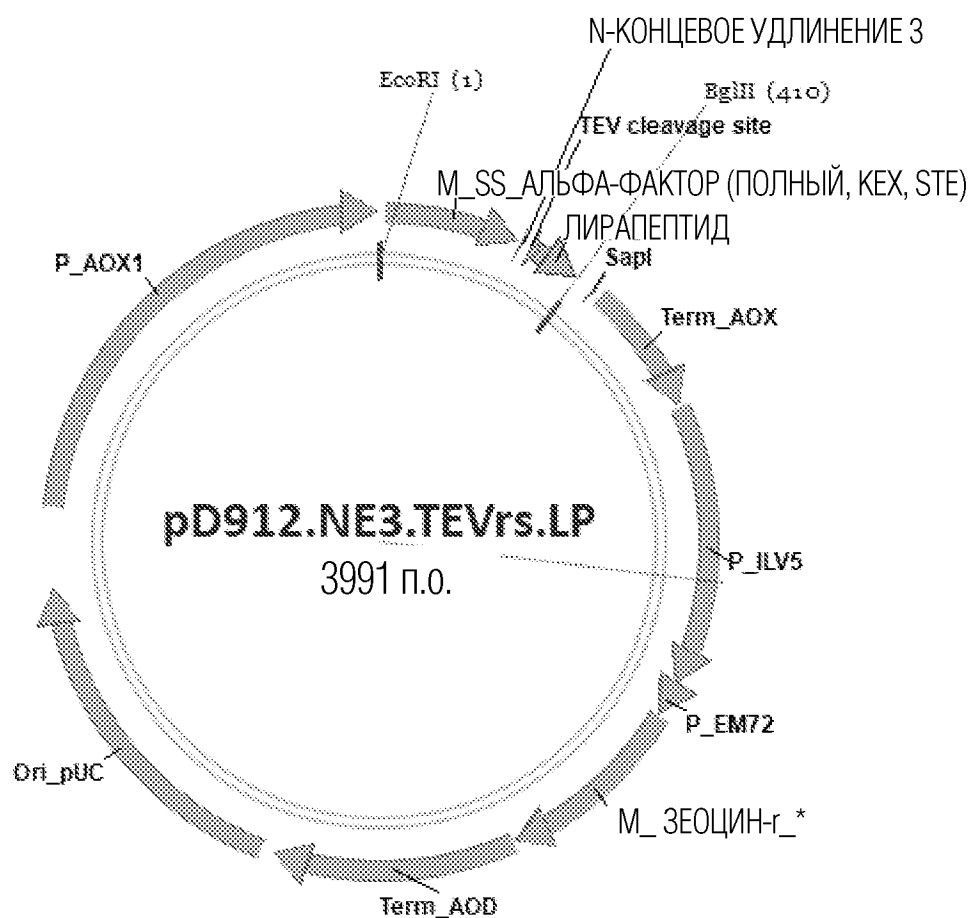
ФИГ. 1 Схематическое представление экспрессирующих кассет



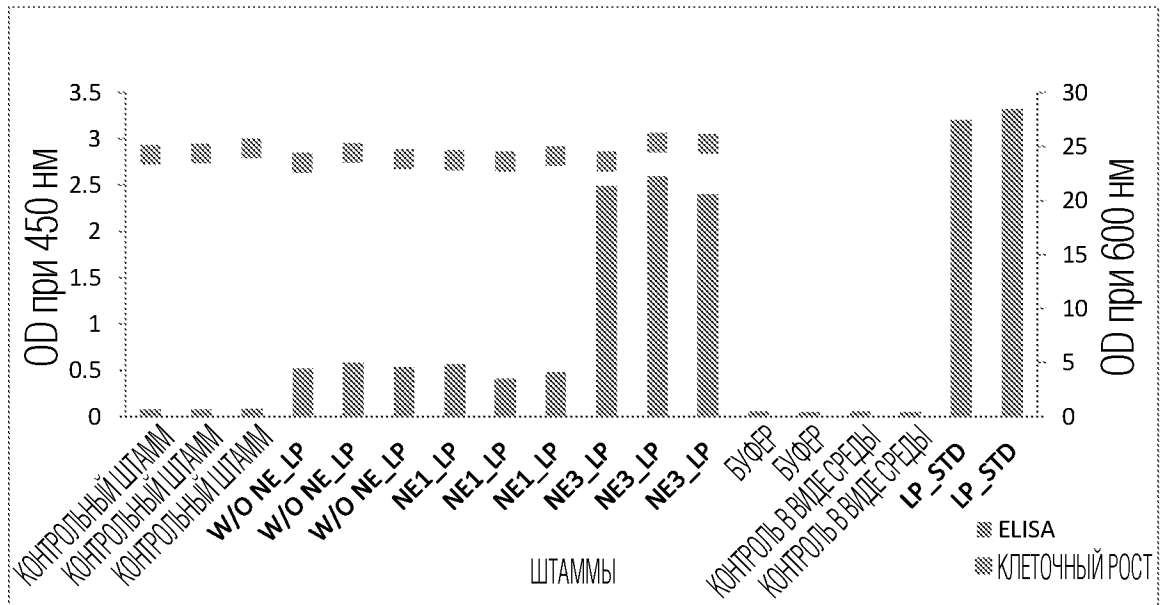
ФИГ. 2А Экспрессионная плазми-
да без N-концевого удлинения



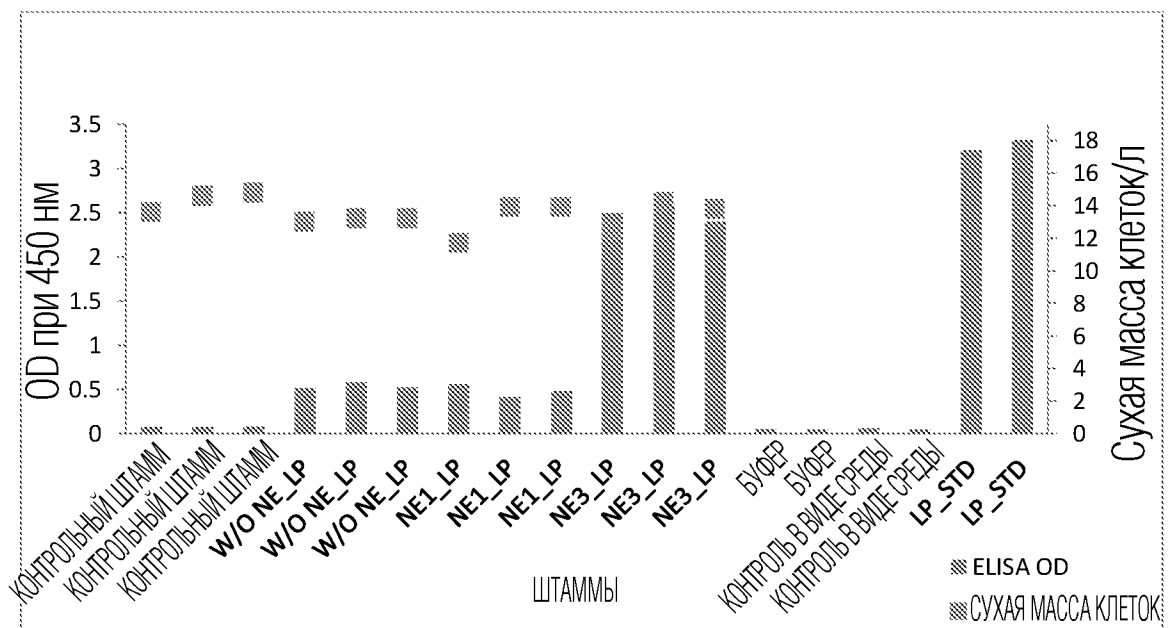
ФИГ. 2В Экспрессионная плазмида
с N-концевым удлинением 1



ФИГ. 2С Экспрессионная плаз-
мида с N-концевым удлинени-
ем 3



ФИГ. 3



ФИГ. 4