

Область изобретения

Настоящее изобретение относится к синтетическим лидерным пептидным последовательностям для секреции полипептидов в дрожжах.

Предпосылки изобретения

Дрожжевые организмы продуцируют большое число белков, синтезируемых внутриклеточно, но функционирующих за пределами клетки. Такие внеклеточные белки называют секретруемыми белками. Эти секретруемые белки сначала экспрессируются внутри клетки в форме предшественника (или пре-белка), содержащего предпоследовательность, обеспечивающую эффективное направление экспрессированного продукта через мембрану эндоплазматического ретикулаума (ЭР). Предпоследовательность, которую обычно называют сигнальным пептидом, обычно отщепляется от интересующего нас продукта в ходе транслокации. Предназначенный для секреции белок транспортируется в аппарат Гольджи. Из аппарата Гольджи белок может быть различными путями доставлен в компартменты, такие как клеточная вакуоль или клеточная мембрана, или может быть выведен из клетки для секреции во внешнюю среду [1].

Было предложено несколько подходов к экспрессии и секреции в дрожжах гетерологичных для них белков. Известен [2] способ, при котором гетерологичные для дрожжей белки экспрессируются, процессируются и секретируются путем трансформации дрожжевого организма вектором экспрессии, несущим кодирующую интересующий нас белок ДНК и сигнальный пептид, получения культуры трансформированного организма, выращивания этой культуры и выделения указанного белка из такой культуральной среды. Сигнальный пептид может быть собственным сигнальным пептидом интересующего нас белка, гетерологичным сигнальным пептидом или гибридом нативного и гетерологичного сигнального пептида.

При использовании гетерологичных для дрожжей сигнальных пептидов может встать проблема не обеспечения гетерологичным сигнальным пептидом эффективной транслокации и/или последующего отщепления такого сигнального пептида.

Известно [3], что *Saccharomyces cerevisiae* MF α 1 (α -фактор) синтезируется в препро форме из 165 аминокислот, содержащей сигнальный или препептид из 19 аминокислот, за которым следует "лидер" или пропептид из 64 аминокислот, включающий в себя три N-связанных сайта гликозилирования, за которым следует (LysArg(Asp/Glu,Ala)₂₋₃ α -фактор)₄. Сигнально-лидерная часть препроMF α 1 широко используется для достижения синтеза и секреции гетерологичных белков в *S. cerevisiae*.

Также известно [4-10] использование сигнального/лидерного пептида гомологичного для дрожжей.

Описано как использование предшественника α -фактора *S. cerevisiae* [9], так и использование сигнального пептида инвертазы *Saccharomyces cerevisiae* PH05 [10] и использование сигнального пептида PH05 [11] для секреции чужеродных белков.

Описаны также [4-8] способы, при помощи которых сигнал-лидер из α -фактора *Saccharomyces cerevisiae* (MF α 1 и MF α 2) используют для секреции экспрессированных гетерологичных белков в дрожжах. Путем присоединения ДНК-последовательности, кодирующей сигнал/лидер последовательность *S. cerevisiae* MF α 1, к 5 концу гена, кодирующего интересующий нас белок, была продемонстрирована секреция и процессинг интересующего нас белка.

Известна [12] система секреции полипептида из *S. cerevisiae* с использованием лидерной последовательности α -фактора, усеченной для исключения четырех единиц α -фактора, присутствующих на нативной лидерной последовательности, с тем чтобы убрать сам лидерный пептид, присоединенный к гетерологичному полипептиду через процессинговый сайт α -фактора LysArgGluAlaGluAla. Указано, что такая конструкция приводит к эффективному процессингу меньших пептидов (менее 50 аминокислот). Для секреции и процессинга больших пептидов нативная лидерная последовательность α -фактора была усечена, чтобы убрать одну или две единицы α -фактора между лидерным пептидом и полипептидом.

Некоторое количество секретированных пептидов было подвергнуто воздействию протеолитической процессинговой системе, способной расщеплять пептидную связь на карбокси-конце двух последовательных основных аминокислот. Эта ферментативная активность кодируется в *S. cerevisiae* геном KEX 2 [13]. Процессинг продукта KEX 2 протеазой необходим для секреции активного *S. cerevisiae* скрепляющего фактора α 1 (MF α 1 или α -фактора), в то время как KEX 2 не вовлечен в секрецию активного *S. cerevisiae* скрепляющего фактора α .

Секреция и корректный процессинг полипептида, который должен быть секретирован, достигается в некоторых случаях при культивировании дрожжевых организмов, которые трансформированы вектором, сконструированным как указано в приведенных выше ссылках. Однако, во многих случаях уровень секреции очень низок или секреция отсутствует, или протеолитический процессинг осуществляется некорректно или неполностью. Таким образом, задачей данного изобретения является создание лидерных пептидов, которые обеспечивают более эффективную экспрессию и/или процессинг полипептидов.

Резюме

Неожиданно был обнаружен новый тип лидерного пептида, который позволяет достигать высокого уровня секреции полипептидов в дрожжах.

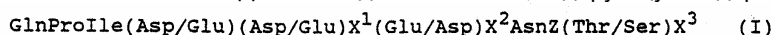
Согласно этому, данное изобретение относится к ДНК экспрессирующей кассете, включающей в себя следующую последовательность:

5'-P-SP-IS-PS-*ген*-(T)_i-3'

где P обозначает промоторную последовательность,

SP обозначает ДНК последовательность, кодирующую сигнальный пептид,

LS обозначает ДНК последовательность, кодирующую лидерный пептид общей формулы I:



где X^1 обозначает пептидную связь или кодируемую аминокислоту,

X^2 обозначает пептидную связь или кодируемую аминокислоту, или последовательность до 4 одинаковых или разных кодируемых аминокислот,

Z обозначает кодируемую аминокислоту кроме Pro, и

X^3 обозначает последовательность от 4 до 30 одинаковых или разных кодируемых аминокислот,

PS обозначает ДНК последовательность, кодирующую процессинговый сайт,

ген обозначает ДНК последовательность, кодирующую полипептид,

T обозначает терминаторную последовательность, и

i равно 0 или 1.

В данном контексте под термином "лидерный пептид" понимают пептид, чьей функцией является позволить экспрессированному полипептиду, быть направленным из эндоплазматического ретикулума в аппарат Гольджи и далее в секреторные везикулы для секреции в среду (то есть передвижение экспрессированного полипептида через клеточную стенку или по меньшей мере через клеточную мембрану в периплазматическое пространство клетки). Выражение "синтетический", используемое по отношению к лидерным пептидам, подразумевает, что такой лидерный пептид не обнаруживается в природе.

Выражение "сигнальный пептид" обозначает препоследовательность, главной особенностью которой является гидрофобная природа, и которая присутствует как N-концевая последовательность в предшественнике внеклеточного белка, экспрессируемого в дрожжах. Функция сигнального пептида заключается в том, чтобы позволить экспрессированному белку, который должен быть секретирован, попасть в эндоплазматический ретикулум. Сигнальный пептид в норме отщепляется в течение данного процесса. Сигнальный пептид может быть гетерологичным или гомологичным для дрожжевого организма, продуцирующего указанный белок.

Выражение "полипептид" обозначает как гетерологичный полипептид, то есть полипептид, который не продуцируется в природе хозяйским дрожжевым организмом, так и гомологичный полипептид, то есть полипептид, который продуцируется в природе хозяйским дрожжевым организмом, и любые их преформы. В предпочтительном воплощении экспрессирующая кассета по изобретению кодирует гетерологичный полипептид.

Выражение "кодируемая аминокислота" подразумевает аминокислоту, которая может быть закодирована при помощи триплета ("кодона") нуклеотидов.

Если в аминокислотной последовательности, приведенной в настоящем описании, трехбуквенные коды двух аминокислот, разделенных косой чертой, даны в скобках, например (Asp/Glu), то подразумевается, что данная последовательность содержит либо одну, либо другую из этих аминокислот в соответствующем положении.

Еще одним аспектом данного изобретения является способ продуцирования полипептидов в дрожжах, при котором в подходящей среде культивируют дрожжевую клетку, способную экспрессировать полипептид и которая трансформирована дрожжевым вектором экспрессии, как описано выше, содержащим лидерную пептидную последовательность по изобретению, с достижением экспрессии и секреции полипептида, после чего этот полипептид выделяют из среды.

Краткое описание рисунков

Настоящее изобретение иллюстрируется далее со ссылкой на прилагаемые рисунки, на которых

Фиг.1 схематично изображает плазмиду pAK492;

Фиг.2 показывает часть последовательности ДНК, кодирующей сигнальный пептид/лидер/M13 предшественник инсулина;

Фиг.3 показывает конструкцию плазмиды pAK546;

Фиг.4 показывает аминокислотную последовательность лидера SEQ ID No.4 и кодирующую ее последовательность ДНК;

Фиг.5 показывает ДНК последовательность экспрессионной плазмиды pAK546 *S.cerevisiae*, кодирующей YAP3 сигнальный пептид, лидера SEQ ID No.4 и предшественник инсулина M13, и закодированную аминокислотную последовательность;

Фиг.6 показывает аминокислотную последовательность лидера SEQ ID No.6 и кодирующую ее последовательность ДНК;

Фиг.7 показывает аминокислотную последовательность лидера SEQ ID No.8 и кодирующую ее последовательность ДНК;

Фиг.8 показывает аминокислотную последовательность лидера SEQ ID No.17 и кодирующую ее последовательность ДНК;

Фиг.9 показывает аминокислотную последовательность лидера SEQ ID No.16 и кодирующую ее последовательность ДНК;

Фиг.10 показывает аминокислотную последовательность лидера SEQ ID No.19 и кодирующую ее последовательность ДНК;

Фиг.11 показывает аминокислотную последовательность лидера SEQ ID No.20 и кодирующую ее последовательность ДНК;

Фиг.12 показывает аминокислотную последовательность лидера SEQ ID No.21 и кодирующую ее последовательность ДНК;

Фиг.13 показывает фрагмент ДНК pAK527, используемый как прямая матрица в конструкции SEQ ID No.4 и 6.

Фиг.14 показывает фрагмент ДНК pAK531, используемый как прямая матрица в конструкции SEQ ID No.8.

Фиг.15 показывает фрагмент ДНК pAK555, используемый как прямая матрица в конструкции SEQ ID No.16 и 17.

Фиг.16 показывает фрагмент ДНК pAK559, используемый как прямая матрица в конструкции SEQ ID No.19 и 20.

Фиг.17 показывает фрагмент ДНК pAK562, используемый как прямая матрица в конструкции SEQ ID No.21.

Фиг.18 показывает аминокислотную последовательность лидера SEQ ID No.27 и кодирующую ее

последовательность ДНК SEQ ID No.66;

Фиг.19 показывает аминокислотную последовательность SEQ ID No.70, удлинненного с N-конца предшественника инсулина MI3, и кодирующую ее последовательность ДНК SEQ ID No.71;

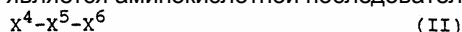
Фиг.20 показывает аминокислотную последовательность лидера SEQ ID No. 67 и кодирующую ее последовательность ДНК SEQ ID No.69;

Фиг.21 показывает фрагмент ДНК SEQ ID No.72 pAK614, используемый как прямая матрица в конструкции SEQ ID No.27.

Фиг.22 показывает фрагмент ДНК SEQ ID No.73 pAK625, используемый как прямая матрица в конструкции SEQ ID No.67.

Детальное описание изобретения

Когда X^1 в общей формуле I обозначает аминокислоту, она предпочтительно является Ser, Thr или Ala. Когда X^2 в общей формуле I обозначает аминокислоту, она предпочтительно является Ser, Thr или Ala. Когда X^2 в общей формуле I обозначает последовательность двух аминокислот, она предпочтительно представляет собой SerIle. Когда X^2 в общей формуле I обозначает последовательность трех аминокислот, она предпочтительно представляет собой SerAlaIle. Когда X^2 в общей формуле I обозначает последовательность четырех аминокислот, она предпочтительно представляет собой SerPheAlaThr. В предпочтительном воплощении X^3 является аминокислотной последовательностью общей формулы II



где X^4 обозначает последовательность от 1 до 21 одинаковой или разной кодируемой аминокислоты, X^5 обозначает Pro или одну из аминокислотных последовательностей ValAsnLeu или LeuAlaAsnVal-AlaMetAla, и X^6 обозначает последовательность от 1 до 8 одинаковых или разных кодируемых аминокислот.

В общей формуле II X^4 предпочтительно обозначает аминокислотную последовательность, которая включает в себя один или несколько мотивов LeuValAsnLeu, SerValAsnLeu, MetAlaAsp, ThrGluSer, ArgPheAlaThr или ValAlaMetAla; или X^4 обозначает аминокислотную последовательность, которая включает в себя последовательность AsnSerThr или AsnThrThr; или X^4 обозначает аминокислотную последовательность, которая включает в себя последовательность

(Ser/Leu)ValAsnLeu,
(Ser/Leu)ValAsnLeuMetAlaAsp,
(Ser/Leu)ValAsnLeuMetAlaAspAsp,
(Ser/Leu)ValAsnLeuMetAlaAspAspThrGluSer,
(Ser/Leu)ValAsnLeuMetAlaAspAspThrGluSer Ile или
(Ser/Leu)ValAsnLeuMetAlaAspAspThrGluSerArgPheAlaThr; или

X^4 обозначает аминокислотную последовательность, которая включает в себя последовательность

Asn(Thr/Ser)ThrLeu,

Asn(Thr/Ser)ThrLeuAsnLeu или

Asn(Thr/Ser)ThrLeuValAsnLeu; или любую их комбинацию.

В общей формуле II X^5 предпочтительно обозначает Pro или аминокислотную последовательность, которая включает в себя последовательность ValAsnLeu, LeuAlaAsnValAlaMetAla, LeuAspValVal-AsnLeuProGly или LeuAspValValAsnLeuIleSerMet.

Когда X^6 в общей формуле II обозначает одну аминокислоту, она предпочтительно является Ala, Gly, Leu, Thr, Val или Ser. Когда X^6 в общей формуле II обозначает последовательность двух аминокислот, она предпочтительно представляет собой GlyAla или SerAla. Когда X^6 в общей формуле II обозначает последовательность трех аминокислот, она предпочтительно представляет собой AlaValAla. Когда X^6 в общей формуле II обозначает последовательность восьми аминокислот, она предпочтительно представляет собой GlyAlaAspSerLysThrValGlu.

Примерами предпочтительных лидерных пептидов, кодируемых последовательностью ДНК LS являются:

SEQ ID No. 1 GlnProIleAspGluAspAsnAspThrSerValAsnLeuProAla;

SEQ ID No. 2 GlnProIleAspAspGluAsnThrThrSerValAsnLeuProAla;

SEQ ID No. 3 GlnProIleAspAspGluSerAsnThrThrSerValAsnLeuProAla;

SEQ ID No. 4 GlnProIleAspAspGluAsnThrThrSerValAsnLeuProVal;

SEQ ID No. 5 GlnProIleAspAspThrGluAsnThrThrSerValAsnLeuProAla;

SEQ ID No. 6 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAla;

SEQ ID No. 7 GlnProIleAspAspGluAsnThrThrSerValAsnLeuMetAla;

SEQ ID No. 8 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProGlyAla;

SEQ ID No. 9 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAla;

SEQ ID No. 10 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnValProThr;

SEQ ID No. 11 GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnValPro-
Thr;

SEQ ID No. 12 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuPro-
Thr;

SEQ ID No. 13 GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnValPro-
GlyAla;

SEQ ID No. 14 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaProAlaValAla;

SEQ ID No. 15 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AspLeuAlaValGlyLeuProGlyAla;

SEQ ID No. 16 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGly-
Ala;

SEQ ID No. 17 GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-
ProGlyAla;

SEQ ID No. 18 GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnLeuPro-
GlyAla;

SEQ ID No. 19 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsn-
LeuProLeu;

SEQ ID No. 20 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuAlaAsn-
ValAlaMetAla;

SEQ ID No. 21 GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuValAsn-
LeuProGlyAla;

SEQ ID No. 22 GlnProIleAspAspThrGluSerPheAlaThrAsnThrThrLeuVal-
AsnLeuProGlyAla;

SEQ ID No. 23 GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-
MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuVal-
AsnLeuProLeu;

SEQ ID No. 24 GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-
MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAsp-
ValValAsnLeuProGlyAla;

SEQ ID No. 25 GlnProIleAspAspThrGluSerAlaAlaIleAsnThrThrLeuVal-
AsnLeuProGlyAla;

SEQ ID No. 26 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsn-
LeuAlaAsnValAlaMetAla;

SEQ ID No. 27 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAspVal-
ValAsnLeuIleSerMetAla;

SEQ ID No. 28 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAsnThrThrGluSerArgPheAlaThrAsnThrThrLeuAspVal-
ValAsnLeuIleSerMetAla; and

SEQ ID No. 67 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAlaLeu-
AspValValAsnLeuIleSerMetAla;

К конкретным предпочтительным лидерным пептидам, кодируемым последовательностью ДНК LS относятся:

SEQ ID No. 15 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AspLeuAlaValGlyLeuProGlyAla;

SEQ ID No. 16 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGly-
Ala;

SEQ ID No. 17 GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-
ProGlyAla;

SEQ ID No. 18 GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnLeuPro-
GlyAla;

SEQ ID No. 19 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsn-
LeuProLeu;

SEQ ID No. 20 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuAlaAsn-
ValAlaMetAla;

SEQ ID No. 21 GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuValAsn-
LeuProGlyAla;

SEQ ID No. 22 GlnProIleAspAspThrGluSerPheAlaThrAsnThrThrLeuVal-
AsnLeuProGlyAla;

SEQ ID No. 23 GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-
MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuVal-
AsnLeuProLeu;

SEQ ID No. 24 GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-
MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAsp-
ValValAsnLeuProGlyAla;

SEQ ID No. 25 GlnProIleAspAspThrGluSerAlaAlaIleAsnThrThrLeuVal-
AsnLeuProGlyAla;

SEQ ID No. 26 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsn-
LeuAlaAsnValAlaMetAla; and

SEQ ID No. 28 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAsnThrThrGluSerArgPheAlaThrAsnThrThrLeuAspVal-
ValAsnLeuIleSerMetAla.

SEQ ID No. 67 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAlaLeu-
AspValValAsnLeuIleSerMetAla.

Сигнальная последовательность (SP) может кодировать любой сигнальный пептид, который обеспечивает эффективное направление экспрессированного полипептида по секреторному пути в клетке. Этот сигнальный пептид может быть встречающимся в природе сигнальным пептидом или его функциональной частью либо синтетическим пептидом. Было обнаружено, что подходящими сигнальными пептидами являются сигнальный пептид α -фактора, сигнальный пептид амилазы из слюны мыши [14], модифицированный сигнальный пептид карбоксипептидазы [15] или дрожжевой BAR1 сигнальный пептид [16] или сигнальный пептид липазы из *Humicola lanuginosa* или их производные. Описан [17] сигнальный пептид дрожжевой аспартиновой протеазы 3.

Подходящим дрожжевым процессинговым сайтом, кодируемым последовательностью ДНК PS, может быть любая парная комбинация Lys и Arg, такая как LysArg, ArgLys, LysLys или ArgArg, которая позволяет осуществлять процессинг данного полипептида с помощью KEX2 протеазы *Saccharomyces cerevisiae* или эквивалентной протеазы в других фидах дрожжей [13]. Если KEX2 процессинг не подходит, например если он будет вести к расщеплению полипептидного продукта, например из-за присутствия двух последовательных основных аминокислот в самом интересующем нас продукте, то может быть выбран процессинговый сайт для другой протеазы, включающий в себя комбинацию аминокислот, не встречающуюся в полипептидном продукте, например процессинговый сайт для FXa, IleGluGlyArg [18].

Белком, продуцированным способом по изобретению, может быть любой белок, который успешно продуцируется в дрожжах. Примерами таких белков являются аprotинин, ингибитор тканевого фактора пути метаболизма или ингибиторы других протеаз, инсулин или предшественники инсулина, человеческий или бычий гормон роста, интерлейкин, глюкагон, глюкагон-подобный пептид-1 (GLP-1), IGF-I, IGF-II, тканевой активатор плазминогена, трансформирующий фактор роста α или β , происходящий из тромбоцитов фактор роста, ферменты или функциональный аналог любого из этих белков. В данном контексте термин "функциональный

аналог" обозначает белок со сходными с нативным белком функциями (подразумевается скорее природа, чем уровень биологической активности нативного белка). Белок может быть структурно схож с нативным белком и может быть получен из нативного белка путем добавления одной или нескольких аминокислот к C- и/или N-концу нативного белка, путем замещения одной или нескольких аминокислот в одном или нескольких различных сайтах нативной аминокислотной последовательности, путем делеции одной или нескольких аминокислот с одного или с обоих концов нативного белка или в одном или нескольких сайтах аминокислотной последовательности, или путем вставки одной или нескольких аминокислот в один или несколько сайтов нативной аминокислотной последовательности. Такие модификации хорошо известны для некоторых упомянутых выше белков. Также предшественники или промежуточные продукты для других белков могут быть получены способом по изобретению. Примером подходящего предшественника является предшественник инсулина MI3, содержащий аминокислотную последовательность B(1-29)-Ala-Ala-Lys-A(1-21), где A(1-21) обозначает A цепь человеческого инсулина и B(1-29) обозначает B цепь человеческого инсулина, в которой отсутствует Thr (B30).

Предпочтительны ДНК-конструкции, кодирующие лидерные последовательности, показанные на Фиг. 4-12 или их подходящие модификации. Примерами подходящих модификаций последовательности ДНК являются замены нуклеотидов, которые не приводят к получению другой аминокислотной последовательности белка, а соответствуют использованию кодонов дрожжевого организма, в который вводят эту последовательность ДНК, или замены нуклеотидов, которые не приводят к получению другой аминокислотной последовательности и, следовательно, возможно, другой структуре белка. Другими примерами возможных модификаций являются вставки одного или нескольких кодонов в последовательность или добавление одного или нескольких кодонов к любому концу последовательности, или делеция одного или нескольких кодонов на любом конце или внутри данной последовательности.

Рекомбинантным вектором экспрессии, несущим экспрессирующую кассету

5'-P-SP-LS-PS-*ген*-(T)_i-3'

где P, SP, LS, *ген*, T и i определены выше,

может быть любой вектор, способный реплицироваться в дрожжевых организмах. Промотором может быть любая последовательность ДНК, обнаруживающая транскрипционную активность в дрожжах, и которая может быть получена из генов, кодирующих либо гомологичные, либо гетерологичные для дрожжей белки. Такой промотор предпочтительно получают из генов, кодирующих гомологичные для дрожжей белки. Примерами подходящих промоторов являются промоторы *Saccharomyces cerevisiae* MF α 1, TPI, ADH или PGK.

Показанные выше последовательности должны также предпочтительно быть присоединены к соответствующему терминатору, например TPI терминатору [19].

Рекомбинантный вектор экспрессии по изобретению содержит далее последовательность ДНК, способствующую репликации данного вектора в дрожжах. Примерами таких последовательностей являются гены репликации REP 1-3 и сайт инициации репликации дрожжевой плазмиды 2 μ . Такой вектор также может содержать селективный маркер, например TPI ген *Schizosaccharomyces pombe* [20].

Методы, используемые для лигирования последовательности 5'-P-SP-LS-PS-*ген*-(T)_i-3' и вставки их в подходящие дрожжевые векторы, содержащие необходимую для репликации в дрожжах информацию, являются хорошо известными специалистам [18]. Очевидно, что такой вектор может быть сконструирован либо путем получения ДНК-конструкции, содержащей полную последовательность 5'-P-SP-LS-PS-*ген*-(T)_i-3' с последующей вставкой этого фрагмента в подходящий вектор экспрессии, либо путем последовательной вставки в подходящий вектор фрагментов ДНК, кодирующих информацию об индивидуальных элементах (таких как промоторная последовательность, сигнальный пептид, лидерная последовательность GlnProlle(Asp/Glu)(Asp/Glu)X¹(Asp/Glu)X²AsnZ(Thr/Ser)X³, процессинговый сайт, полипептид и, если имеется, терминаторная последовательность), с последующим лигированием.

Дрожжевым организмом, используемым в способе по изобретению, может быть любой подходящий дрожжевой организм, который при культивировании продуцирует большие количества интересующего нас полипептида. Примерами подходящих дрожжевых организмов служат штаммы дрожжей видов *Saccharomyces cerevisiae*, *Saccharomyces kluyveri*, *Saccharomyces pombe* или *Saccharomyces uvarum*. Трансформацию дрожжевых клеток можно осуществить путем создания протопластов с последующей трансформацией известными способами. Средой, используемой для культивирования клеток, может быть любая традиционная среда, подходящая для выращивания дрожжевых организмов. Секретированный полипептид, значительная часть которого должна присутствовать в среде в правильно процессированной форме, может быть выделен из среды традиционными способами, включая отделение дрожжевых клеток от среды путем центрифугирования или фильтрации, осаждение белковых компонентов супернатанта или фильтрата с помощью соли, например сульфата аммония, с последующей очисткой различными хроматографическими методами, например ион-обменной хроматографией, аффинной хроматографией и т.д.

Настоящее изобретение далее демонстрируется на следующих примерах, которые не ограничивают рамки заявленного изобретения.

Примеры

Плазмиды и ДНК

Все плазмиды для экспрессии имеют C-POT тип. Такие плазмиды описаны [21], и характеризуются наличием гена триозо фосфат изомеразы (POT) из *Schizosaccharomyces pombe* для селекции и стабилизации плазмид. Содержащие ген POT плазмиды могут быть получены из депонированного штамма *E. coli* (ATCC 39685). Далее эти плазмиды содержат промотор и терминатор триозо фосфат изомеразы из *S. cerevisiae* (P_{TPI} и T_{TPI}). Они идентичны pMT742 [22] (см. Фиг. 1), за исключением области, определенной рестрикционным сайтом EcoRI-XbaI, охватывающим область, кодирующую сигнал/лидер/продукт.

Плазмиды pAK527, pAK531, pAK555, pAK559, pAK562, pAK614 и pAK625 были использованы в качестве ДНК-матриц в ПЦР, примененной в конструкции лидеров, описанной в примерах. Синтетические фрагменты ДНК, служащие в качестве прямых матриц, показаны на Фиг. 13-17. За исключением показанных областей ДНК эти плазмиды идентичны pAK492, показанной на Фиг. 1.

Синтетические фрагменты ДНК были синтезированы на автоматическом ДНК-синтезаторе (Applied Biosystems model 380A) с использованием фосфорамидитной химии и коммерчески доступных реагентов [23].

Все остальные использованные материалы и методы являются широко известными [18].

Пример 1

Синтез лидера SEQ ID No. 4 для экспрессии предшественника инсулина MI3 в *S.cerevisiae* (штамм уAK546).

Лидер SEQ ID No 4. имеет следующую аминокислотную последовательность:

GlnProIleAspAspGluAsnThrThrSerValAsnLeuProVal.

Были синтезированы следующие олигонуклеотиды:

# 94	5'-TAAATCTATAACTACAAAAACACATA-3'	SEQ ID No. 29
# 333	5'-GACTCTCTTAACTGGCAAGTTGACA-3'	SEQ ID No. 30
# 312	5'-AAGTACAAAGCTTCAACCAAGTGAGAACACACAAGTGT	
	GGTTAACGAATCTCTT-3'	SEQ ID No. 31
# 1845	5'-CATACACAATATAAACGACGG-3'	SEQ ID No. 32

Следующие полимеразные цепные реакции (ПЦР) были осуществлены с использованием набора реагентов Gene Amp PCR (Perkin Elmer, 761 Main Avenue, CT 06859, USA), согласно инструкции производителя. В течение реакции на ПЦР-смесь наслаивали 100 мкл минерального масла (Sigma Chemical CO, St. Louis MO, USA):

Полимеразная цепная реакция N° 1

5 мкл олигонуклеотида # 94 (50 пмоль)

5 мкл олигонуклеотида # 333 (50 пмоль)

10 мкл 10x ПЦР-буфера

16 мкл смеси dNTP

0,5 мкл Taq фермента

0,5 мкл плазмиды рAK527 (Фиг.13) в качестве матрицы (0,2 мкг ДНК)

63 мкл воды

Было проведено в общем 12 циклов, один цикл составлял 94°C в течение 1 мин, 37°C в течение 2 мин, 72°C в течение 3 мин. ПЦР-смесь затем наносили на 2%-ный агарозный гель и стандартным образом проводили электрофорез [18]. Полученные фрагменты ДНК вырезали из агарозного геля и выделяли, используя Gene Clean kit (Bio 101 inc., PO BOX 2284, La Jolla, CA 92038, USA), согласно инструкции производителя.

Полимеразная цепная реакция N° 2

5 мкл олигонуклеотида # 312 (50 пмоль)

5 мкл олигонуклеотида # 94 (50 пмоль)

10 мкл 10x ПЦР-буфера

16 мкл смеси dNTP

0,5 мкл Taq фермента

10 мкл очищенного фрагмента ДНК из ПЦР N° 1

53,5 мкл воды

Было проведено в общем 12 циклов, один цикл составлял 94°C в течение 1 мин, 37°C в течение 2 мин, 72°C в течение 3 мин. Полученные после ПЦР N°2 фрагменты ДНК выделяли и очищали, используя Gene Clean kit (Bio 101 inc., PO BOX 2284, La Jolla, CA 92038, USA), согласно инструкции производителя.

Очищенные ПЦР фрагменты ДНК растворяли в 10 мкл воды и буфера для рестрикционных нуклеаз, и разрезали рестрикционными эндонуклеазами Asp 718 и Hind III в общем объеме 15 мкл, используя стандартные методики [18]. Фрагмент ДНК 167 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали, используя Gene Clean kit как описано. *S.cerevisiae* экспрессионная плазмида рAK492 (показано на Фиг.1) является производной описанной выше плазмиды рMT742, в которой фрагмент, кодирующий сигнал/лидер/предшественник инсулина был заменен на EcoRI-XbaI фрагмент, показанный на Фиг.2. Этот фрагмент был синтезирован на Applied Biosystem ДНК синтезаторе, согласно инструкциям производителя. Плазмиду рAK492 разрезали рестрикционными эндонуклеазами Asp 718 и Hind III и выделяли фрагмент ДНК 140 пн, кодирующий часть предшественника инсулина MI3. Три фрагмента ДНК лигировали вместе, используя T4 ДНК-лигазу в стандартных условиях [18]. Затем лигированной смесью трансформировали компетентный штамм *E.coli* (R-, M+) и определяли трансформантов по устойчивости к ампицилину. Выделяли плазмиды из полученных колоний *E.coli*, используя стандартные методики [18], осуществляя проверку подходящими рестрикционными эндонуклеазами, то есть EcoRI, XbaI, NcoI и Hind III. С помощью ДНК секвенс-анализа (Sequenase, U.S. Biochemical Corp.) с использованием праймера # 94 было показано, что отселектированная плазмида рAK546 содержит последовательность ДНК, кодирующую лидер SEQ ID No.4. Последовательность ДНК, кодирующая лидер SEQ ID No.4, показана на Фиг.4. Плазмидой рAK546 трансформировали штамм MT663 *S.cerevisiae* [24] и назвали полученный штамм уAK546. Последовательность ДНК области экспрессионной плазмиды, кодирующей белок, дана на Фиг.5.

Пример 2

Синтез лидера SEQ ID No. 6 для экспрессии предшественника инсулина MI3 в *S.cerevisiae* (штамм уAK531).

Лидер SEQ ID No 6. имеет следующую аминокислотную последовательность :

Были синтезированы следующие олигонуклеотиды:

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAla.

331 5'-GAATCTCTTAGCTGGCAAGTTGACAGAAGTAGTGTAG

TTTCAGAGTCGTCATT-3'

SEQ ID No. 33

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида #

333 использовали олигонуклеотид # 331.

Фрагмент ДНК 168 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1. Фрагмент ДНК Asp 718/Hind III субклонировали в *S.cerevisiae* экспрессионной плазмиде как описано в Примере 1. С помощью ДНК сиквенс-анализа, как описано в Примере 1, было показано, что отселектированная плазида рАК531 содержит последовательность ДНК, кодирующую лидер SEQ ID No.6. Последовательность ДНК, кодирующая лидер SEQ ID No.6, показана на Фиг.6. Плазмидой рАК531 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уАК531. Последовательности ДНК, кодирующие сигнальный пептид и предшественник инсулина М13, являются такими же, как на Фиг.5.

Пример 3

Синтез лидера SEQ ID No. 8 для экспрессии предшественника инсулина М13 в *S.cerevisiae* (штамм уАК547).

Лидер SEQ ID No 8. имеет следующую аминокислотную последовательность:

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAla.

Были синтезированы следующие олигонуклеотиды:

345 5'-AACGAATCTCTTAGCACCTGGCAAGTTGACAGAAGT-3' SEQ ID No. 34

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида # 333 использовали олигонуклеотид # 345 и в качестве матрицы - плазмиду рАК531 (Фиг.14).

Фрагмент ДНК 171 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1. Фрагмент ДНК Asp 718/Hind III субклонировали в *S.cerevisiae* экспрессионной плазмиде как описано в Примере 1. С помощью ДНК сиквенс-анализа, как описано в Примере 1, было показано, что отселектированная плазида рАК547 содержит последовательность ДНК, кодирующую лидер SEQ ID No.8. Последовательность ДНК, кодирующая лидер SEQ ID No.8, показана на Фиг.7. Плазмидой рАК547 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уАК547. Последовательности ДНК, кодирующие сигнальный пептид и предшественник инсулина М13, являются такими же, как на Фиг.5.

Пример 4

Синтез лидера SEQ ID No.17 для экспрессии предшественника инсулина М13 в *S.cerevisiae* (штамм уАК561).

Лидер SEQ ID No. 17 имеет следующую аминокислотную последовательность:

GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnLeu-

ProGlyAla.

Были синтезированы следующие олигонуклеотиды:

25 # 376 5'-AACGAATCTCTTAGCACCTGGCAAGTTGACCAAAGTAG
TGTTGATAGATTCAAGTGTCTGTC-3' SEQ ID No. 35

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида # 333 использовали олигонуклеотид # 376 и в качестве матрицы - плазмиду рАК555 (Фиг.15).

Фрагмент ДНК 180 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1. Фрагмент ДНК Asp 718/Hind III субклонировали в *S.cerevisiae* экспрессионной плазмиде как описано в Примере 1. С помощью ДНК сиквенс-анализа, как описано в Примере 1, было показано, что отселектированная плазида рАК561 содержит последовательность ДНК, кодирующую лидер SEQ ID No.17. Последовательность ДНК, кодирующая лидер SEQ ID No.17, показана на Фиг.8. Плазмидой рАК561 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уАК561. Последовательности ДНК, кодирующие сигнальный пептид и предшественник инсулина М13, являются такими же, как на Фиг.5.

Пример 5

Синтез лидера SEQ ID No. 16 для экспрессии предшественника инсулина М13 в *S.cerevisiae* (штамм уАК559).

Лидер SEQ ID No.16 имеет следующую аминокислотную последовательность:

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-

AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGlyAla.

Были синтезированы следующие олигонуклеотиды:

375 5'-AACGAATCTCTTAGCACCTGGCAAGTTAACCAAAGTAGT
GTTGATAGATTCAAGTGTCTGTCAGCCATCAAGTTGAC-3' SEQ ID No. 36

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида # 333 использовали олигонуклеотид # 375 и в качестве матрицы - плазмиду рАК555 (Фиг.15).

Фрагмент ДНК 222 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1. Фрагмент ДНК Asp 718/Hind III субклонировали в *S.cerevisiae* экспрессионной плазмиде как описано в Примере 1. С помощью ДНК сиквенс-анализа, как описано в Примере 1, было показано, что отселектированная плазида рАК559 содержит последовательность ДНК, кодирующую лидер SEQ ID No.16. Последовательность ДНК, кодирующая лидер SEQ ID No.16, показана на Фиг.9. Плазмидой рАК559 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уАК559. Последовательности ДНК, кодирующие сигнальный пептид и предшественник инсулина М13, являются такими же, как на Фиг.5.

Пример 6

Синтез лидера SEQ ID No. 19 для экспрессии предшественника инсулина М13 в *S.cerevisiae* (штамм уАК580).

Лидер SEQ ID No.19 имеет следующую аминокислотную последовательность:

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-

AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsnLeuProLeu.

Были синтезированы следующие олигонуклеотиды:

384 5'-AACGAATCTCTTCAATGGCAAGTTAACCAAAGTAGTGT

TAGTAGCGAATCTAGATTCAAGTGTCTGTCAGCCAT-3' SEQ ID No. 37

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида # 333 использовали олигонуклеотид # 384 и в качестве матрицы - плазмиду pAK559 (Фиг.16).

Фрагмент ДНК 228 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1. Фрагмент ДНК Asp 718/Hind III субклонировали в *S.cerevisiae* экспрессионной плазмиде как описано в Примере 1. С помощью ДНК сиквенс-анализа, как описано в Примере 1, было показано, что отселектированная плазида pAK580 содержит последовательность ДНК, кодирующую лидер SEQ ID No.19. Последовательность ДНК, кодирующая лидер SEQ ID No.19, показана на Фиг.10. Плазмидой pAK580 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уAK580. Последовательности ДНК, кодирующие сигнальный пептид и предшественник инсулина MI3, являются такими же, как на Фиг.5.

Пример 7

Синтез лидера SEQ ID No.20 для экспрессии предшественника инсулина MI3 в *S.cerevisiae* (штамм уAK583).

Лидер SEQ ID No.20 имеет следующую аминокислотную последовательность:

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-

AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuAlaAsnValAlaMetAla.

Были синтезированы следующие олигонуклеотиды:

390 5'-AACGAATCTCTTAGCCATGGCAACGTTAGCCAAGTTAA

'CCAAAGT-3'

SEQ ID No. 38

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида # 333 использовали олигонуклеотид # 390 и в качестве матрицы - плазмиду pAK559 (Фиг.16).

Фрагмент ДНК 231 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1. Фрагмент ДНК Asp 718/Hind III субклонировали в *S.cerevisiae* экспрессионной плазмиде как описано в Примере 1. С помощью ДНК сиквенс-анализа, как описано в Примере 1, было показано, что отселектированная плазида pAK583 содержит последовательность ДНК, кодирующую лидер SEQ ID No.20. Последовательность ДНК, кодирующая лидер SEQ ID No.20, показана на Фиг.11. Плазмидой pAK583 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уAK583. Последовательности ДНК, кодирующие сигнальный пептид и предшественник инсулина MI3, являются такими же, как на Фиг.5.

Пример 8

Синтез лидера SEQ ID No. 21 для экспрессии предшественника инсулина MI3 в *S.cerevisiae* (штамм уAK586).

Лидер SEQ ID No.21 имеет следующую аминокислотную последовательность:

GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuValAsn-

LeuProGlyAla.

Были синтезированы следующие олигонуклеотиды:

401 5'-AACGAATCTCTTAGCACCTGGCAAGTTGACCAAAGTAG

TGTTGATAGCAGATTCAAGTGTCTG-3'

SEQ ID No. 39

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида # 333 использовали олигонуклеотид # 401 и в качестве матрицы - плазмиду pAK562 (Фиг.17).

Фрагмент ДНК 183 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1. Фрагмент ДНК Asp 718/Hind III субклонировали в *S.cerevisiae* экспрессионной плазмиде как описано в Примере 1. С помощью ДНК сиквенс-анализа, как описано в Примере 1, было показано, что отселектированная плазида pAK586 содержит последовательность ДНК, кодирующую лидер SEQ ID No.21. Последовательность ДНК, кодирующая лидер SEQ ID No.21, показана на Фиг.12. Плазмидой pAK586 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уAK586. Последовательности ДНК, кодирующие сигнальный пептид и предшественник инсулина MI3, являются такими же, как на Фиг.5.

Пример 9

Экспрессия предшественника инсулина с использованием выбранных лидерных последовательностей согласно настоящему изобретению.

Дрожжевой штамм, несущий плазмиды, как описано выше, выращивают на YPD среде [26]. Для каждого штамма помещают на качалку при 30°C на 72 ч 6 индивидуальных культур по 5 мл до финальной OD₆₀₀ около 15. После центрифугирования удаляют супернатант для HPLC анализа, с помощью которого измеряют концентрацию секретируемого предшественника инсулина [27].

В таблице 1 даны уровни экспрессии предшественника инсулина MI3, полученные с использованием выбранных лидерных последовательностей согласно изобретению, в виде процента от уровня, полученного у трансформантов pMT742 с использованием MF α (1) лидер *S.cerevisiae*.

Таблица 1

лидер	уровень экспрессии, %
MT748 α-лидер	100
SEQ ID No.15	87
SEQ ID No.16	215
SEQ ID No.17	157
SEQ ID No.19	166
SEQ ID No.20	86
SEQ ID No.21	145
SEQ ID No.22	137
SEQ ID No.23	121

Пример 10

Синтез лидера SEQ ID No.27 для экспрессии предшественника инсулина MI3 в *S.cerevisiae* (штамм уAK677).

Лидер SEQ ID No.27 имеет следующую аминокислотную последовательность :

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAspValValAsnLeuIle-
SerMetAla.

Были синтезированы следующие олигонуклеотиды:

440 5'-GGTTAACGAACCTTGGAGCTTCAGCTTCAGCTTCTCTCTTAGCCAT
GGAGATCAAGTTAACAACATCCAAAGTAGTGT-3' SEQ ID No. 64

и

441 5'-CAAGTACAAAGCTTCAACCAAGTGGGAACCGCACAAGTGTGGTTAACG
AACTT-3' SEQ ID No. 65

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида # 333 использовали олигонуклеотид # 440 и в качестве матрицы - плазмиду рAK614. Для второй ПЦР вместо олигонуклеотида # 312 использовали олигонуклеотид # 441.

Очищенные ПЦР фрагменты ДНК выделяли и разрезали рестрикционными эндонуклеазами Asp 718 и Hind III, как описано в примере 1. Фрагмент ДНК 268 пн Asp 718/Hind III подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1. Фрагмент ДНК Asp 718/Hind III субклонировали в *S.cerevisiae* экспрессионной плазмиде, как описано в Примере 1, за исключением того, что фрагмент 140 пн HindIII/XbaI был получен из рAK602 и кодирует Asp^{B28} человеческий инсулин. С помощью ДНК сиквенс-анализа, как описано в Примере 1, было показано, что отселектированная плаزمида рAK616 содержит последовательность ДНК, кодирующую лидер SEQ ID No.27. Последовательность ДНК SEQ ID No.66, кодирующая лидер SEQ ID No.27, показана на Фиг.11. Фрагмент ДНК 268 пн Asp 718/Hind III из рAK616 выделяли и лигировали с фрагментом ДНК 10986 пн Asp 718/XbaI из рAK464 (кодирующим удлинненный Asp^{B28} человеческий инсулин), обозначив рAK625. Фрагмент ДНК 180 пн Asp718/NcoI из рAK625 был выделен и лигирован с фрагментом ДНК 221 пн NcoI/XbaI из рJB146 (кодирующим удлинненный предшественник инсулина) и фрагментом ДНК 10824 пн Asp718/XbaI из рAK601, полученную плазмиду обозначили рAK677. Плазмидой рAK677 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уAK677. За исключением последовательности ДНК, кодирующей лидер, последовательность ДНК, кодирующая сигнальный пептид, показана на Фиг.5. Последовательность ДНК, кодирующая удлинненный предшественник инсулина MI3, показана на Фиг.19.

Пример 11

Синтез лидера SEQ ID No.67 для экспрессии предшественника инсулина MI3 в *S.cerevisiae* (штамм уAK680).

Лидер SEQ ID No.67 имеет следующую аминокислотную последовательность:

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAlaLeuAspValValAsnLeu-
IleSerMetAla.

Были синтезированы следующие олигонуклеотиды:

577 5'-TCTCTTAGCCATGGAGATCAAGTTAACAACATCCAAAG
CCAAAGTAGTGT-3' SEQ ID No. 68

Полимеразная цепная реакция была проведена как описано в Примере 1, но вместо олигонуклеотида # 333 использовали олигонуклеотид # 557 и в качестве матрицы - плазмиду рAK625, а вторую ПЦР не проводили. ПЦР фрагмент разрезали рестрикционными эндонуклеазами Asp 718 и NcoI, как описано в примере 1.

Фрагмент ДНК 190 пн Asp 718/NcoI подвергали электрофорезу на агарозном геле и очищали как описано в Примере 1, за исключением того, что векторный фрагмент ДНК 10824 пн Asp 718/XbaI был выделен из рAK601. Фрагмент ДНК 190 пн Asp 718/NcoI субклонировали в *S.cerevisiae* экспрессионной плазмиде, как описано в

Примере 1, за исключением того, что фрагмент 221 пн NcoI/XbaI (кодирующий удлиненную версию предшественника инсулина) был получен из рAK677 и использован вместо фрагмента HindIII/XbaI ДНК. С помощью ДНК секвенс-анализа, как описано в Примере 1, было показано, что отселектированная плазмида рAK680 содержит последовательность ДНК, кодирующую лидер SEQ ID No.67. Последовательность ДНК SEQ ID No.69, кодирующая лидер SEQ ID No.67, показана на Фиг.20. Плазмидой рAK680 трансформировали штамм MT663 *S.cerevisiae* [25] и назвали полученный штамм уAK680. За исключением последовательности ДНК, кодирующей лидер, последовательность ДНК, кодирующая сигнальный пептид, показана на Фиг.5. Последовательность ДНК, кодирующая удлиненный предшественник инсулина MI3, показана на Фиг.19.

Пример 12

Экспрессия удлиненного с N-конца предшественника инсулина MI3 с использованием лидерных последовательностей SEQ ID No.27 и SEQ ID No.67 согласно настоящему изобретению

Дрожжевой штамм, несущий плазмиды, как описано выше, выращивают на YPD среде [26]. Для каждого штамма помещают на качалку при 30°C на 72 ч 6 индивидуальных культур по 5 мл до финальной OD₆₀₀ около 15. После центрифугирования удаляют супернатант для HPLC анализа, с помощью которого измеряют концентрацию секретированного предшественника инсулина [27].

В таблице 2 даны уровни экспрессии некоторых удлиненных с N-конца предшественников инсулина MI3, полученных с использованием лидерных последовательностей SEQ ID No.27 и SEQ ID No.67 согласно изобретению, в виде процента от уровня, полученного у трансформантов рMT742 с использованием MF α (1) лидера *S.cerevisiae*.

Таблица 2

Штамм	Сигнальный пептид	Лидер	Удлиняющая последовательность	Относительно MT748
MT748	α	α		
уAK675	YAP3	SEQ ID No.27	EEAEAEAP K	251%
уAK677	YAP3	SEQ ID No.27	EEAEAEAE PK	224%
уAK681	YAP3	SEQ ID No.67	EEAEAEAP K	248%
уAK680	YAP3	SEQ ID No.67	EEAEAEAE PK	362%

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

(1) ОБЩАЯ ИНФОРМАЦИЯ:

(1) ЗАЯВИТЕЛЬ:

- (А) ИМЯ: Novo Nordisk A/S
- (Б) УЛИЦА: Novo Alle
- (Г) ГОРОД: DK-2880 Bagsvaerd
- (Д) СТРАНА: Denmark
- (Ж) ТЕЛЕФОН: +45 4444 8888
- (З) ТЕЛЕФАКС: +45 4449 3256
- (И) ТЕЛЕКС: 37173

(2) НАЗВАНИЕ ИЗОБРЕТЕНИЯ:

СИНТЕТИЧЕСКИЕ ЛИДЕРНЫЕ ПЕПТИДНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

(3) КОЛИЧЕСТВО ПОСЛЕДОВАТЕЛЬНОСТЕЙ: 73

(4) КОРРЕСПОНДЕНТСКИЙ АДРЕС

- (А) АДРЕС: Novo Nordisk A/S
Corporate Patents

- (Б) УЛИЦА: Novo Alle
- (Г) ГОРОД: DK-2880 Bagsvaerd
- (Д) СТРАНА: Denmark

(5) ФОРМА КОМПЬЮТЕРНОГО ПРЕДСТАВЛЕНИЯ:

- (А) ТИП НОСИТЕЛЯ: гибкий диск
- (Б) КОМПЬЮТЕР: IBM PC совместимый
- (В) ОПЕРАЦИОННАЯ СИСТЕМА: PC-DOS/MS-DOS
- (Г) ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ: Patentin Release #1.0, Version #1.25

(6) ДАННЫЕ О ТЕКУЩЕЙ ЗАЯВКЕ:

- (А) НОМЕР ЗАЯВКИ:
- (Б) ДАТА ПОДАЧИ:
- (В) КЛАССИФИКАЦИЯ:

(7) ДАННЫЕ О ПРИОРИТЕТНОЙ ЗАЯВКЕ:

- (А) НОМЕР ЗАЯВКИ: DK 0705/94 и US 08/282,852
- (Б) ДАТА ПОДАЧИ: 16.06.94 и 29.07.94

(8) ИНФОРМАЦИЯ О ПОВЕРЕННОМ/АГЕНТЕ:

- (А) ИМЯ: Jorgensen, Dan et al.
- (В) НОМЕР ССЫЛКИ/ДЕЛА: 4085-WO,DJ

(9) ТЕЛЕКОММУНИКАЦИОННАЯ ИНФОРМАЦИЯ:

- (А) ТЕЛЕФОН: +45 4444 8888
- (В) ТЕЛЕФАКС: +45 4449 3256

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 1

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 15 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.1:

Gln Pro Ile Asp Glu Asp Asn Asp Thr Ser Val Asn Leu Pro Ala
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 2

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 15 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.2:

Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn Leu Pro Ala
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 3

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 16 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.3:

Gln Pro Ile Asp Asp Glu Ser Asn Thr Thr Ser Val Asn Leu Pro Ala
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 4

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 15 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.4:

Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn Leu Pro Val
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 5

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 16 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.5:

Gln Pro Ile Asp Asp Thr Glu Asn Thr Thr Ser Val Asn Leu Pro Ala
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 6

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 17 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.6:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro
1 5 10 15

Ala

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 7

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 15 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.7:

Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn Leu Met Ala
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 8

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 18 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.8:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro
1 5 10 15

Gly Ala

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 9

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 17 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.9:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Ala

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 10

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 17 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.10:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Val Pro
1 5 10 15

Thr

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 11

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 17 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.11:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Leu Val Asn Val Pro
1 5 10 15

Thr

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 12

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 17 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.12:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro
1 5 10 15

Thr

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 13

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 18 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.13:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Leu Val Asn Val Pro
1 5 10 15

Gly Ala

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 14

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 21 аминокислота
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.14:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Ala Pro Ala Val Ala

20

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 15

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 25 аминокислот

(Б) ТИП: аминокислотная
(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.15:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Asp Leu Ala Val Gly Leu Pro Gly Ala
20 25

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 16

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 33 аминокислоты

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.16:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Pro Gly
20 25 30

Ala

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 17

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 19 аминокислот

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.17:

Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu
1 5 10 15

Pro Gly Ala

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 18

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 18 аминокислот

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.18:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Leu Val Asn Leu Pro
1 5 10 15

Gly Ala

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 19

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 35 аминокислот

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.19:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn
20 25 30

Leu Pro Leu
35

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 20

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 36 аминокислот

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.20:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Ala Asn
20 25 30

Val Ala Met Ala
35

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 21

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 20 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 21:

Gln Pro Ile Asp Asp Thr Glu Ser Ala Ile Asn Thr Thr Leu Val Asn
1 5 10 15
Leu Pro Gly Ala
20

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 22

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 21 аминокислота
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 22:

Gln Pro Ile Asp Asp Thr Glu Ser Phe Ala Thr Asn Thr Thr Leu Val
1 5 10 15
Asn Leu Pro Gly Ala
20

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 23

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 36 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 23:

Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu
1 5 10 15
Met Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val
20 25 30
Asn Leu Pro Leu
35

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 24

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 39 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 24:

Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu
1 5 10 15
Met Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Asp
20 25 30
Val Val Asn Leu Pro Gly Ala
35

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 25

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 21 аминокислота
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 25:

Gln Pro Ile Asp Asp Thr Glu Ser Ala Ala Ile Asn Thr Thr Leu Val
1 5 10 15
Asn Leu Pro Gly Ala
20

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 26

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 39 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 26:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
 1 5 10 15
 Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn
 20 25 30
 Leu Ala Asn Val Ala Met Ala
 35

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 27

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 39 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 27:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
 1 5 10 15
 Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Asp Val
 20 25 30
 Val Asn Leu Ile Ser Met Ala
 35

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 28

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 39 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 28:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
 1 5 10 15
 Ala Asn Thr Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Asp Val
 20 25 30
 Val Asn Leu Ile Ser Met Ala
 35

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 29

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 27 пар оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 29:

ТАААТСТАТА АСТАСААААА АСАСАТА

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 30

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 25 пар оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(3) ГИПОТЕТИЧЕСКАЯ: НЕТ

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 30:

ГАСТСТСТТА АСТГГСААГТ ТГАСА

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 31

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 56 пар оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 31:

ААГТАСААААГ СТТСААССАА ГТГАГААССА САСААГТГТТ ГГТТААСГАА ТСТСТТ

56

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 32

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 21 пара оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 32:

САТАСААТ АТАААСGAG G	21
(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 33	
(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:	
(А) ДЛИНА: 55 пар оснований	
(Б) ТИП: нуклеиновая кислота	
(В) СКРУЧЕННОСТЬ: однонитевая	
(Г) ТОПОЛОГИЯ: линейная	
(2) ТИП МОЛЕКУЛЫ: кДНК	
(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.33:	
GAATCTCTTA GCTGGCAAGT TGACAGAAGT AGTGTAGTT TCAGAGTCGT CAATT	55
(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 34	
(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:	
(А) ДЛИНА: 36 пар оснований	
(Б) ТИП: нуклеиновая кислота	
(С) СКРУЧЕННОСТЬ: однонитевая	
(Г) ТОПОЛОГИЯ: линейная	
(2) ТИП МОЛЕКУЛЫ: кДНК	
(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.34:	
AACGAATCTC TTAGCACCTG GCAAGTTGAC AGAAGT	36
(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 35	
(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:	
(А) ДЛИНА: 60 пар оснований	
(Б) ТИП: нуклеиновая кислота	
(В) СКРУЧЕННОСТЬ: однонитевая	
(Г) ТОПОЛОГИЯ: линейная	
(2) ТИП МОЛЕКУЛЫ: кДНК	
(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.35:	
AACGAATCTC TTAGCACCTG GCAAGTTGAC CAAAGTAGTG TTGATAGATT CAGTGTCTGC	60
(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 36	
(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:	
(А) ДЛИНА: 75 пар оснований	
(Б) ТИП: нуклеиновая кислота	
(В) СКРУЧЕННОСТЬ: однонитевая	
(Г) ТОПОЛОГИЯ: линейная	
(2) ТИП МОЛЕКУЛЫ: кДНК	
(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.36:	
AACGAATCTC TTAGCACCTG GCAAGTTAAC CAAAGTAGTG TTGATAGATT CAGTGTCTGC	60
AGCCATCAAG TTGAC	75
(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 37	
(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:	
(А) ДЛИНА: 72 пары оснований	
(Б) ТИП: нуклеиновая кислота	
(В) СКРУЧЕННОСТЬ: однонитевая	
(Г) ТОПОЛОГИЯ: линейная	
(2) ТИП МОЛЕКУЛЫ: кДНК	
(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.37:	
AACGAATCTC TTCAATGGCA AGTTAACCAA AGTAGTGTTA GTAGCGAATC TAGATTCAGT	60
GTCGTGAGCC AT	72
(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 38	
(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:	
(А) ДЛИНА: 45 пар оснований	
(Б) ТИП: нуклеиновая кислота	
(В) СКРУЧЕННОСТЬ: однонитевая	
(Г) ТОПОЛОГИЯ: линейная	
(2) ТИП МОЛЕКУЛЫ: кДНК	
(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.38:	
AACGAATCTC TTAGCCATGG CAACGTTAGC CAAGTTAACC AAAGT	45
(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 39	
(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:	
(А) ДЛИНА: 61 пара оснований	
(Б) ТИП: нуклеиновая кислота	
(С) СКРУЧЕННОСТЬ: однонитевая	
(Г) ТОПОЛОГИЯ: линейная	
(2) ТИП МОЛЕКУЛЫ: кДНК	
(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.39:	
AACGAATCTC TTAGCACCTG GCAAGTTGAC CAAAGTAGTG TTGATAGCAG ATTCAGTGTC	60
G	61

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 40

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 372 пары оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
- (Б) ПОЛОЖЕНИЕ: 82..351

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.40:

```
GAATTCATTC AAGAATAGTT CAAACAAGAA GATTACAAC TATCAATTC ATACACAATA      60
TAAACGACGG GTACCAAAAT A ATG AAA CTG AAA ACT GTA AGA TCT GCG GTC      111
      Met Lys Leu Lys Thr Val Arg Ser Ala Val
      1              5              10
CTT TCG TCA CTC TTT GCA TCT CAG GTC CTT GGC CAA CCA ATA GAC GAA      159
Leu Ser Ser Leu Phe Ala Ser Gln Val Leu Gly Gln Pro Ile Asp Glu
      15              20              25
GAC AAC GAC ACT TCT TCC ATG GCT AAG AGA TTC GTT AAC CAA CAC TTG      207
Asp Asn Asp Thr Ser Ser Met Ala Lys Arg Phe Val Asn Gln His Leu
      30              35              40
TGC GGT TCC CAC TTG GTT GAA GCT TTG TAC TTG GTT TGC GGT GAA AGA      255
Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg
      45              50              55
GGT TTC TTC TAC ACT CCT AAG GCT GCT AAG GGT ATT GTC GAG CAA TGC      303
Gly Phe Phe Tyr Thr Pro Lys Ala Ala Lys Gly Ile Val Glu Gln Cys
      60              65              70
TGT ACC TCC ATC TGC TCC TTG TAC CAA TTG GAA AAC TAC TGC AAC TAGACGCAGC
358
Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
75              80              85              90
CCGCAGGCTC TAGA      372
```

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 41

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 89 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 41:

```
Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
1              5              10              15
Ser Gln Val Leu Gly Gln Pro Ile Asp Glu Asp Asn Asp Thr Ser Ser
20              25              30
Met Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val
35              40              45
Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro
50              55              60
Lys Ala Ala Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser
65              70              75              80
Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
85
```

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 42

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 45 пар оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
- (Б) ПОЛОЖЕНИЕ: 82..351

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.42:

```
CAA CCA ATT GAC GAC GAA AAC ACT ACT TCT GTC AAC TTG CCA GTT
Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn Leu Pro Val
1              5              10              15
```

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 43

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 15 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(3) ГИПОТЕТИЧЕСКАЯ: НЕТ

(3) АНТИСМЫСЛОВАЯ: НЕТ

(6) ПРОИСХОЖДЕНИЕ:

(А) ОРГАНИЗМ: синтетическая

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 43:

Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn Leu Pro Val
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 44

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 297 пар оснований

(Б) ТИП: нуклеиновая кислота

(В) СКРУЧЕННОСТЬ: однонитевая

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

(А) ИМЯ/КЛЮЧ: CDS

(Б) ПОЛОЖЕНИЕ: 1..276

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.44:

ATG AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA 48
Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
1 5 10 15
TCT CAG GTC CTT GGC CAA CCA ATT GAC GAC GAA AAC ACT ACT TCT GTC 96
Ser Gln Val Leu Gly Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val
20 25 30
AAC TTG CCA GTT AAG AGA TTC GTT AAC CAA CAC TTG TGT GGT TCT CAC 144
Asn Leu Pro Val Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His
35 40 45
TTG GTT GAA GCT TTG TAC TTG GTT TGC GGT GAA AGA GGT TTC TTC TAC 192
Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr
50 55 60
ACT CCT AAG GCT GCT AAG GGT ATT GTC GAA CAA TGC TGT ACC TCC ATC 240
Thr Pro Lys Ala Ala Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile
65 70 75 80
TGC TCC TTG TAC CAA TTG GAA AAC TAC TGC AAC TAGACGCAGC CCGCAGGCTC 293
Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
85 90
TAGA 297

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 45

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 91 аминокислота

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 45:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
1 5 10 15
Ser Gln Val Leu Gly Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val
20 25 30
Asn Leu Pro Val Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His
35 40 45
Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr
50 55 60
Thr Pro Lys Ala Ala Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile
65 70 75 80
Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
85 90

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 46

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 51 пара оснований

(Б) ТИП: нуклеиновая кислота

(В) СКРУЧЕННОСТЬ: однонитевая

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК (9) ОСОБЕННОСТЬ:

(А) ИМЯ/КЛЮЧ: CDS

(Б) ПОЛОЖЕНИЕ: 1..51

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.46:

CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT GTC AAC TTG CCA	48
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro	
1 5 10 15	

GCT	51
Ala	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 47

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 17 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 47:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro
1 5 10 15

Ala

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 48

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 54 пары оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
- (Б) ПОЛОЖЕНИЕ: 1..54

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 48:

CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT GTC AAC TTG CCA	48
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro	
1 5 10 15	

GGT GCT	54
Gly Ala	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 49

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 57 пар оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
- (Б) ПОЛОЖЕНИЕ: 1..57

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.49:

CAA CCA ATT GAC GAC ACT GAA TCT ATC AAC ACT ACT TTG GTC AAC TTG	48
Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu	
1 5 10 15	

CCA GGT GCT	57
Pro Gly Ala	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 50

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 99 пар оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
- (Б) ПОЛОЖЕНИЕ: 1..99

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.50:

CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT GTC AAC TTG ATG	48
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met	
1 5 10 15	

GCT GAC GAC ACT GAA TCT ATC AAC ACT ACT TTG GTT AAC TTG CCA GGT	96
Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Pro Gly	
20 25 30	

GCT	99
Ala	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 51

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 105 пар оснований
- (Б) ТИП: нуклеиновая кислота

- (В) СКРУЧЕННОСТЬ: одонитевая
 (Г) ТОПОЛОГИЯ: линейная
 (2) ТИП МОЛЕКУЛЫ: кДНК
 (9) ОСОБЕННОСТЬ:
 (А) ИМЯ/КЛЮЧ: CDS
 (Б) ПОЛОЖЕНИЕ: 1..105
 (11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.51:

CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT GTC AAC TTG ATG	48
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met	
1 5 10 15	
GCT GAC GAC ACT GAA TCT AGA TTC GCT ACT AAC ACT ACT TTG GTT AAC	96
Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn	
20 25 30	
TTG CCA TTG	105
Leu Pro Leu	
35	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 52

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 108 пар оснований
 (Б) ТИП: нуклеиновая кислота
 (В) СКРУЧЕННОСТЬ: одонитевая
 (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
 (Б) ПОЛОЖЕНИЕ: 1..108

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.52

CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT GTC AAC TTG ATG	48
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met	
1 5 10 15	
GCT GAC GAC ACT GAA TCT ATC AAC ACT ACT TTG GTT AAC TTG GCT AAC	96
Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Ala Asn	
20 25 30	
GTT GCC ATG GCT	108
Val Ala Met Ala	
35	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 53

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 60 пар оснований
 (Б) ТИП: нуклеиновая кислота
 (В) СКРУЧЕННОСТЬ: одонитевая
 (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
 (Б) ПОЛОЖЕНИЕ: 1..60

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 53:

CAA CCA ATT GAC GAC ACT GAA TCT GCT ATC AAC ACT ACT TTG GTC AAC	48
Gln Pro Ile Asp Asp Thr Glu Ser Ala Ile Asn Thr Thr Leu Val Asn	
1 5 10 15	
TTG CCA GGT GCT	60
Leu Pro Gly Ala	
20	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 54

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 276 пар оснований
 (Б) ТИП: нуклеиновая кислота
 (В) СКРУЧЕННОСТЬ: одонитевая
 (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
 (Б) ПОЛОЖЕНИЕ: 113..274

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.54:

TTAAATCTAT AACTACAAAA AACACATACA GGAATTCATT CAAGAATAGT TCAAACAAGA	60
AGATTACAAA STATCAATTT CATACACAAT ATAAACGACG GGTACCAAAA TA ATG	115
Met 1	
AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA TCT	163
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser	
5 10 15	
CAG GTC CTT GGC CAA CCA ATT GAC GAC GAA AAC ACT ACT TCT GTT AAC	211
Gln Val Leu Gly Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn	
20 25 30	
TTG CCA GCT AAG AGA TTC GTT AAC CAA CAC TTG TGC GGT TCC CAC TTG	259
Leu Pro Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His Leu	
35 40 45	
GTT GAA GCT TTG TAC TT	276
Val Glu Ala Leu Tyr	
50	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 55

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 54 аминокислоты

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 55:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala	
1 5 10 15	
Ser Gln Val Leu Gly Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val	
20 25 30	
Asn Leu Pro Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His	
35 40 45	
Leu Val Glu Ala Leu Tyr	
50	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 56

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 282 пары оснований

(Б) ТИП: нуклеиновая кислота

(В) СКРУЧЕННОСТЬ: однонитевая

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

(А) ИМЯ/КЛЮЧ: CDS

(Б) ПОЛОЖЕНИЕ: 113..280

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.56:

TTAAATCTAT AACTACAAAA AACACATACA GGAATTCATT CAAGAATAGT TCAAACAAGA	60
AGATTACAAA STATCAATTT CATACACAAT ATAAACGACG GGTACCAAAA TA ATG	115
Met 1	
AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA TCT	163
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser	
5 10 15	
CAG GTC CTT GGC CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT	211
Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser	
20 25 30	
GTC AAC TTG CCA GCT AAG AGA TTC GTT AAC CAA CAC TTG TGC GGT TCC	259
Val Asn Leu Pro Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser	
35 40 45	
CAC TTG GTT GAA GCT TTG TAC TT	282
His Leu Val Glu Ala Leu Tyr	
50 55	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 57

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 56 аминокислот

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 57:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
 1 5 10 15
 Ser Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr
 20 25 30
 Ser Val Asn Leu Pro Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly
 35 40 45
 Ser His Leu Val Glu Ala Leu Tyr
 50 55

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 58

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 282 пары оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: одонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
- (Б) ПОЛОЖЕНИЕ: 113..280

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.58:

TTAAATCTAT AACTACAAAA AACACATACA GGAATTCATT CAAGAATAGT TCAACAAGA 60
 AGATTACAAA STATCAATTT CATACACAAT ATAAACGACG GGTACCAAAA TA ATG 115
 Met
 1
 AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA TCT 163
 Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser
 5 10 15
 CAG GTC CTT GGC CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT 211
 Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser
 20 25 30
 GTC AAC TTG ATG GCT AAG AGA TTC GTT AAC CAA CAC TTG TGC GGT TCC 259
 Val Asn Leu Met Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser
 35 40 45
 CAC TTG GTT GAA GCT TTG TAC TT 282
 His Leu Val Glu Ala Leu Tyr
 50 55

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 59

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 56 аминокислот
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 59:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
 1 5 10 15
 Ser Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr
 20 25 30
 Ser Val Asn Leu Met Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly
 35 40 45
 Ser His Leu Val Glu Ala Leu Tyr
 50 55

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 60

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 330 пар оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: одонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

- (А) ИМЯ/КЛЮЧ: CDS
- (Б) ПОЛОЖЕНИЕ: 113..328

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.60:

TTAAATCTAT AACTACAAAA AACACATACA GGAATTCATT CAAGAATAGT TCAAACAAGA 60
AGATTACAAA STATCAATTT CATACACAAT ATAAACGACG GGTACCAAAA TA ATG 115
Met
1

AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA TCT 163
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser
5 10 15

CAG GTC CTT GGC CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT 211
Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser
20 25 30

GTC AAC TTG ATG GCT GAC GAC ACT GAA TCT ATC AAC ACT ACT TTG GTT 259
Val Asn Leu Met Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val
35 40 45

AAC TTG CCA GGT GCT AAG AGA TTC GTT AAC CAA CAC TTG TGC GGT TCC 307
Asn Leu Pro Gly Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser
50 55 60 65

CAC TTG GTT GAA GCT TTG TAC TT 330
His Leu Val Glu Ala Leu Tyr
70

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 61

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 72 аминокислоты

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 61:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
1 5 10 15
Ser Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr
20 25 30
Ser Val Asn Leu Met Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu
35 40 45
Val Asn Leu Pro Gly Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly
50 55 60
Ser His Leu Val Glu Ala Leu Tyr
65 70

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 62

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 288 пар оснований

(Б) ТИП: нуклеиновая кислота

(В) СКРУЧЕННОСТЬ: одонитевая

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(9) ОСОБЕННОСТЬ:

(А) ИМЯ/КЛЮЧ: CDS

(Б) ПОЛОЖЕНИЕ: 113..286

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 62:

TTAAATCTAT AASTACAAAA AACACATACA GGAATTCATT CAAGAATAGT TCAAACAAGA 60
AGATTACAAA STATCAATTT CATACACAAT ATAAACGACG GGTACCAAAA TA ATG 115
Met
1

AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA TCT 163
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser
5 10 15

CAG GTC CTT GGC CAA CCA ATT GAC GAC ACT GAA TCT ATC AAC ACT ACT 211
Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr
20 25 30

TTG GTC AAC TTG CCA GGT GCT AAG AGA TTC GTT AAC CAA CAC TTG TGC 259
Leu Val Asn Leu Pro Gly Ala Lys Arg Phe Val Asn Gln His Leu Cys
35 40 45

GGT TCC CAC TTG GTT GAA GCT TTG TAC TT 288
Gly Ser His Leu Val Glu Ala Leu Tyr
50 55

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 63

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(А) ДЛИНА: 58 аминокислот

(Б) ТИП: аминокислотная

(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 63:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
1 5 10 15
Ser Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr
20 25 30
Thr Leu Val Asn Leu Pro Gly Ala Lys Arg Phe Val Asn Gln His Leu
35 40 45
Cys Gly Ser His Leu Val Glu Ala Leu Tyr
50 55

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 64

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 82 пары оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.64:

GGTTAACGAA CTTTGGAGCT TCAGCTTCAG CTTCTCTCT CTTAGCCATG GAGATCAAGT 60
TAACAACATC CAAAGTAGTG TT 82

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 65

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 54 пары оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.65:

CAAGTACAAA GCTTCAACCA AGTGGGAACC GCACAAGTGT TGGTTAACGA ACTT 54

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 66

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 117 пар оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.66:

CAACCAATTG ACGACACTGA ATCTAACACT ACTTCTGTCA ACTTGATGGC TGACGACACT 60
GAATCTAGAT TCGCTACTAA CACTACTTTG GATGTTGTTA ACTTGATCTC CATGGCT 117

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 67

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 41 аминокислота
- (Б) ТИП: аминокислотная
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 67:

Gln Pro Ile Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15
Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Ala Leu
20 25 30
Asp Val Val Asn Leu Ile Ser Met Ala
35 40

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 68

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 51 пара оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.68:

TCTCTTAGCC ATGGAGATCA AGTTAACAAC ATSCAAAGCC AAAGTAGTGT T 51

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 69

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 123 пары оснований
- (Б) ТИП: нуклеиновая кислота
- (В) СКРУЧЕННОСТЬ: однонитевая
- (Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.69:

CAACCAATTG ACGACACTGA ATCTAACACT ACTTCTGTCA ACTTGATGGC TGACGACACT	60
GAATCTAGAT TCGCTACTAA CACTACTTTG GCTTTGGATG TTGTTAACTT GATCTCCATG	120
GCT	123

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 70

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 65 аминокислот
(Б) ТИП: аминокислотная
(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: пептид

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ 70:

1 Lys Arg Glu Glu Ala Glu Ala Glu Ala Glu Pro Lys Phe Val Asn Gln
 5 10 15
 His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly
 20 25 30
 Glu Arg Gly Phe Phe Tyr Thr Pro Lys Ala Ala Lys Gly Ile Val Glu
 35 40 45
 Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys
 50 55 60
 Asn
 65

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 71

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 219 пар оснований
(Б) ТИП: нуклеиновая кислота
(В) СКРУЧЕННОСТЬ: однострелковая
(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(5) ТИП ФРАГМЕНТА: внутренний

(11) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No. 71:

AAGAGAGAAG	AAGCTGAAGC	TGAAGCTGAA	CCAAAGTTGC	TTAACCAACA	CTGTGTGGT	60
TCTCACTTGG	TGAAGCTTT	GTACTTGGTT	TGCGGTGAAA	GAGGTTTCTT	CTACACTCCT	120
AAGGCTGCTA	AGGGTATTGT	CGAACAAATC	TGTACCTCCA	TCTGCTCCTT	GTACCAATTG	180
GAAAACACT	GCAACTAGAC	GCAGCCCGCA	GGCTCTAGA			219

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 72

(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 348 пар оснований
(Б) ТИП: нуклеиновая кислота
(В) СКРУЧЕННОСТЬ: однонитевая
(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: кДНК

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.72:

TAAATCTAT AACTACAAA AACACATACA GGAATTCCAT TCAAGAA ¹ TAG TTCAAACAAG	60
AAGATTACAA ACTATCAATT TCATACACAA TATAAACGAC GGTACCAAAA TAATGAAACT	120
GAAAACTGTA AGATCTCGGG TCTTTTCGTC ACTCTTTCGA TCTCAGGTCC TTGGCCAACC	180
AATTGACGAC ACTGAATCTAACTACTCTCTGTCAACTTBTATGG-CTG-ACG-	
ACACTG-AATCTABATTGG-CTACTAACACTACTTTGGATGTTG-TTAACTTBTATCT	
CCTGGCTAAGAGAGAGAAGCTGAAAG-CTGAAG-CTGAACCAAA&TTCGTT	
AACCAACACTTGCTG TGCTTCTCACTTGGTTGAAGCTTTGTACTTG	

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ 73

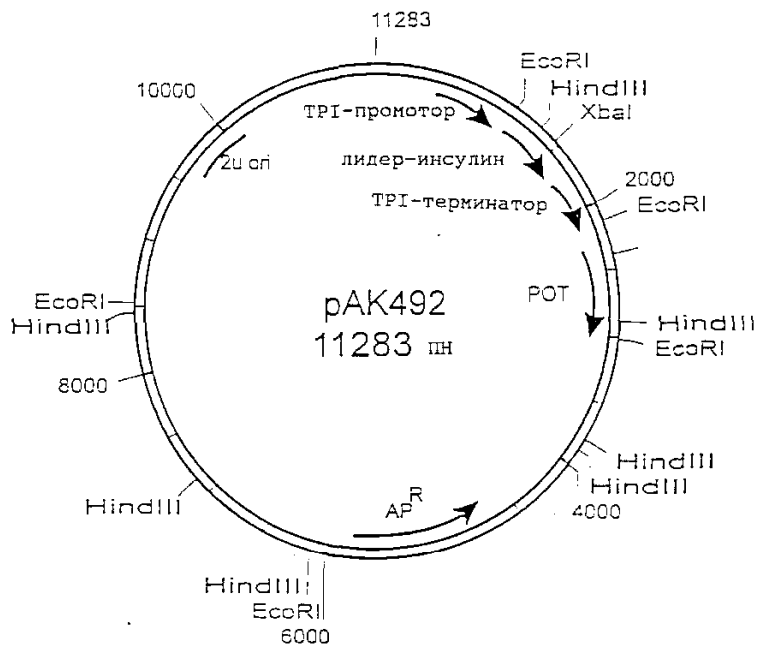
(1) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (А) ДЛИНА: 379 пар оснований
(Б) ТИП: нуклеиновая кислота
(В) СКРУЧЕННОСТЬ: однострелвая
(Г) ТОПОЛОГИЯ: линейная

(2) ТИП МОЛЕКУЛЫ: КДНК

(11)ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID No.73:

TTAAATCTAT AACTACAAAA AACACATACA GGAATTCCAT TCAAGAATAG TTCAACAAG	60
AAGATTACAA ACTATCAATT TCATACACAA TATAAACGAC GGTACCAAAA TAATGAAACT	120
GAAAACTGTA AGATCTGCGG TCCTTTCGTC ACTCTTTGCA TCTCAGGTCC TTGGCCAACC	180
AATTGACGAC ACTGAATCTA ACACTACTTC TGTCAACTTG ATGGCTGACG ACACTGAATC	240
TAGATTGCTT ACTAACAATA CTTTGGATGT TGTAACTTG ATCTCCATGG CTAAGAGAGA	300
AGAAGCTGAA GCTGAAGCTG AACCAAGTT CGTTAACCA CACTTGTGTG GTTCTCACTT	360
GGTTGAAGCT TTGTACTTG	379



ФИГ. 1

EcoR I

GAATTCATTCAAGAATAGT

CTTAAGTAAGTTCTTATCA

TCAAACAAGAAGATTACAACTATCAATTTCATACACAATATAAACGACGGGTACCAAAA
AGTTTGTTCCTTCTAATGTTTGATAGTTAAAGTATGTGTTATATTGCTGCCCATGGTTTT

TAATGAAACTGAAAACTGTAAGATCTGCGGTCCTTTTCGTCACCTCTTGTCATCTCAGGTCCTT
ATTACTTTGACTTTTGACATTCTAGACGCCAGGAAAGCAGTGAGAAACGTAGAGTCCAGGAA
MetLysLeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeu

GGCCAACCAATAGACGAAGACAACGACACTTCTTCCATGGCTAAGAGATTTCGTTAACCAA
CCGGTTGGTTATCTGCTTCTGTTGCTGTGAAGAAGGTACCGATTCTCTAAGCAATTGGTT
GlyGlnProIleAspGluAspAsnAspThrSerSerMetAlaLysArgPheValAsnGln

CACTTGTGCGGTTCCCACTTGGTTGAAGCTTGTACTTGGTTTTCGCGTGAAAGAGGTTTC
GTGAACACGCCAAGGGTGAACCAACTTCGAAACATGAACCAAACGCCACTTTCTCCAAAG
HisLeuCysGlySerHisLeuValGluAlaLeuTyrLeuValCysGlyGluArgGlyPhe

TTCTACACTCCTAAGGCTGCTAAGGGTATTGTGCGAGCAATGCTGTACCTCCATCTGCTCC
AAGATGTGAGGATTCCGACGATTCCCATACAGCTCGTTACGACATGGAGGTAGACGAGG
PheTyrThrProLysAlaAlaLysGlyIleValGluGlnCysCysThrSerIleCysSer

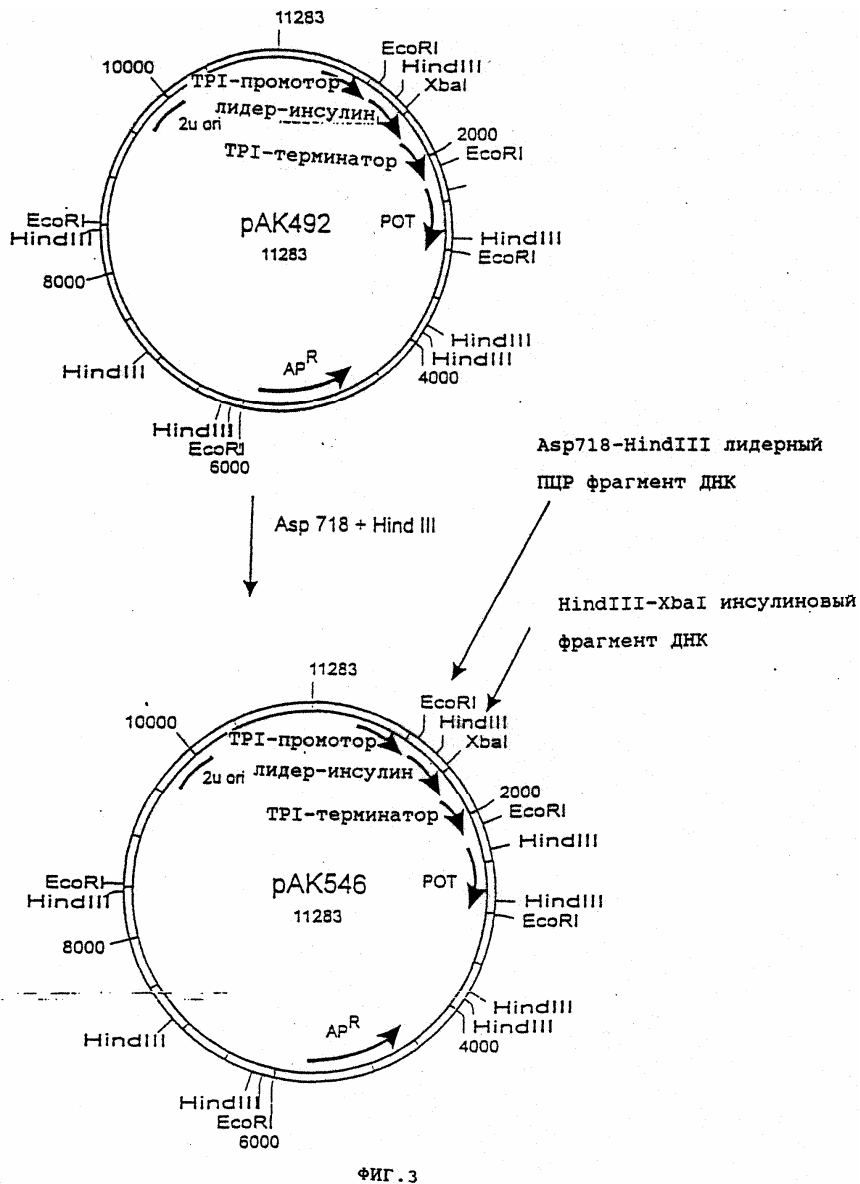
Xba I

TTGTACCAATTGGAAAACTACTGCAACTAGACGCAGCCCGCAGGCTCTAGA SEQ ID No.40

AACATGGTTAACCTTTTGATGACGTTGATCTGCGTCGGCGTCCGAGATCT

LeuTyrGlnLeuGluAsnTyrCysAsn*

SEQ ID No.41



ФИГ. 3

ФИГ. 4

CAACCAATTGACGACGAAAAACACTACTTCTGTCAACTTGCCAGTT
 GTTGGTTAACTGCTGCTTTTGTGATGAAGACAGTTGAACGGTCAA
 GlnProIleAspAspGluAsnThrThrSerValAsnLeuProVal

SEQ ID No. 42

SEQ ID No. 43

ФНГ.5

ATGAAACTGAAAACTGTAAGATCTGCGGTCCTTTCGTCACTCTTTGCA
TACTTTGACTTTTGTGACATTCTAGACGCCAGGAAAGCAGTGAGAAACGT
MetLysLeuLysThrValArgSerAlaValLeuSerSerLeuPheAla

TCTCAGGTCCTTGGCCAACCAATTGACGACGAAAACACTACTTCTGTG
AGACTCCAGGAACCGGTTGGTTAACTGCTGCTTTTGTGATGAAGACAG
SerGlnValLeuGlyGlnProIleAspAspGluAsnThrThrSerVal

AACTTGCCAGTTAAGAGATTCGTTAACCAACACTTGTGTGGTTCTCAC
TTGAACGGTCAATTCTCTAAGCAATTGGTTGTGAACACACCAAGAGTG
AsnLeuProValLysArgPheValAsnGlnHisLeuCysGlySerHis

TTGGTTGAAGCTTTGTACTTGGTTTTCGGTGAAAGAGGTTTCTTCTAC
AACCAACTTCGAAACATGAACCAACGCCACTTTCTCCAAAGAAGATG
LeuValGluAlaLeuTyrLeuValCysGlyGluArgGlyPhePheTyr

ACTCCTAAGGCTGCTAAGGGTATTGTGCAACAATGCTGTACCTCCATC
TGAGGATTCCGACGATTCCCATAACAGCTTGTACGACATGGAGGTAG
ThrProLysAlaAlaLysGlyIleValGluGlnCysCysThrSerIle

TGCTCCTTGTACCAATTGGAAAACTACTGCAACTAGACGCAGCCCGCA
ACGAGGAACATGGTTAACCTTTTGATGACGTTGATCTGCGTCGGGCGT

CysSerLeuTyrGlnLeuGluAsnTyrCysAsn*

SEQ ID No. 45

GGCTCTAGA
CCGAGATCT

SEQ ID No. 44

ФНГ.6

CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGCCAGCT SEQ ID No. 46

GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACGGTCGA

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAla SEQ ID No. 47

ФНГ.7

CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGCCA

GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACGGT

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuPro

GGTGCT

SEQ ID No. 48

CCACGA

GlyAla

SEQ ID No. 8

ФНГ.8

CAACCAATTGACGACACTGAATCTATCAACACTACTTTGGTCAACTTG

GTTGGTTAACTGCTGTGACTTAGATAGTTGTGATGAAACAGTTGAAC

GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu

CCAGGTGCT

SEQ ID No. 49

GGTCCACGA

ProGlyAla

SEQ ID No. 17

ФНГ.9

CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGATG
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACTAC
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet

GCTGACGACACTGAATCTATCAACACTACTTTGCTTAACTTGCCAGGTGCT SEQ ID No. 50
CGACTGCTGTGACTTAGATAGTTGTGATGAAACCAATTGAACGGTCCACGA
AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGlyAla SEQ ID No. 16

ФНГ.10

CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGATGGCTGACGACACT
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACTACCGACTGCTGTGA
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThr

GAATCTAGATTTCGCTACTAACACTACTTTGGTTAACTTGCCATTG SEQ ID No. 51
CTTAGATCTAAGCGATGATTGTGATGAAACCAATTGAACGGTAAC
GluSerArgPheAlaThrAsnThrThrLeuValAsnLeuProLeu SEQ ID No. 19

ФНГ.11

CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGATGGCTGACGACACT
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACTACCGACTGCTGTGA
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThr

GAATCTATCAACACTACTTTGGTTAACTTGGCTAACGTTGCCATGGCT SEQ ID No. 52
CTTAGATAGTTGTGATGAAACCAATTGAACCGATTGCAACGGTACCGA
GluSerIleAsnThrThrLeuValAsnLeuAlaAsnValAlaMetAla SEQ ID No. 20

ФНГ.12

CAACCAATTGACGACACTGAATCTGCTATCAACACTACTTTGGTCAAC
GTTGGTTAACTGCTGTGACTTAGACGATAGTTGTGATGAAACAGTTG
GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuValAsn

TTGCCAGGTGCT SEQ ID No. 53
AACGGTCCACGA
LeuProGlyAla SEQ ID No. 21

ФНГ.13

PAK527:

TTAAATCTATAACTACAAAAAACACATACAGGAATTCATTCAAGAATAGTTCAAACAA
AATTTAGATATTGATGTTTTTTGTGTATGTCCTTAAGTAAGTTCTTATCAAGTTTGTT

GAAGATTACAACTATCAATTTTCATACACAATATAAAACGACGGGTACCAAAATAATGAAA
CTTCTAATGTTTGATAGTTAAAGTATGTGTTATATTTGCTGCCCATGGTTTTATTACTTT
MetLys

CTGAAAACCTGTAAGATCTGCGGTCCTTTTCGTCACTCTTTGCATCTCAGGTCCTTGGCCAA
GACTTTTGACATTCTAGACGCCAGGAAAGCAGTGAGAAACGTAGAGTCCAGGAACCGGTT
LeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeuGlyGln

CCAATTGACGACGAAAACACTACTTCTGTAACTTGCCAGCTAAGAGATTCGTTAACCAA
GGTTAACTGCTGCTTTTGTGATGAAGACAATTGAACGGTCGATTCTCTAAGCAATTGGTT
ProIleAspAspGluAsnThrThrSerValAsnLeuProAlaLysArgPheValAsnGln

CACTTGTGCGGTTCCCACTTGTTGAAGCTTTGTACTT SEQ ID No. 54
GTGAACACGCCAAGGGTGAACCACTTCGAAACATGAA
HisLeuCysGlySerHisLeuValGluAlaLeuTyr SEQ ID No. 55

ФНГ.14
pAK531:

TTAAATCTATAACTACAAAAACACATACAGGAATTCATTCAAGAATAGTTCAAACAA
AATTTAGATATTGATGTTTTTTGTGTATGTCCTTAAGTAAGTTCTTATCAAGTTTGTT

GAAGATTACAACTATCAATTTTCATACACAATATAAAACGACGGGTACCAAAATAATGAAA
CTTCTAATGTTTGATAGTTAAAGTATGTGTTATATTTGCTGCCCATGGTTTTATTACTTT
MetLys

CTGAAAACCTGTAAGATCTGCGGTCCTTTTCGTCACTCTTTGCATCTCAGGTCCTTGGCCAA
GACTTTTGACATTCTAGACGCCAGGAAAGCAGTGAGAAACGTAGAGTCCAGGAACCGGTT
LeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeuGlyGln

CCAATTCACGACACTGAATCTAACACTACTTCTGTCAACTTGCCAGCTAAGAGATTCGTT
GGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACGGTCGATTCTCTAAGCAA
ProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAlaLysArgPheVal

AACCAACACTTGTGCGGTTCCCACTTGGTTGAAGCTTTGTACTT SEQ ID No. 56
TTGGTTGTGAACACGCCAAGGGTGAACCAACTTCGAAACATGAA
AsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyr SEQ ID No. 57

ФНГ.15

PAK555

TTAAATCTATAACTACAAAAACACATACAGGAATTCATTCAAGAATAGTTCAAACAA
AATTTAGATATTGATGTTTTTTGTGTATGTCCTTAAGTAAGTTCTTATCAAGTTTGTT

GAAGATTACAACTATCAATTCATACACAATATAAACGACGGGTACCAAAATAATGAAA
CTTCTAATGTTTGATAGTTAAAGTATGTGTTATATTTGCTGCCCATGGTTTTATTACTTT
MetLys

CTGAAAACTGTAAGATCTGCGGTCCTTTTCGTCACTCTTTGCATCTCAGGTCCTTGGCCAA
GACTTTTGACATTCTAGACGCCAGGAAAGCAGTGAGAAACGTAGAGTCCAGGAACCGGTT
LeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeuGlyGln

CCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGATGGCTAAGAGATTCGTT
GGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACTACCGATTCTCTAAGCAA
ProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaLysArgPheVal

AACCAACACTTGTGCGGTTCCCACTTGGTTGAAGCTTTGTACTT SEQ ID No. 58
TTGGTTGTGAACACGCCAAGGGTGAACCAACTTCGAAACATGAA
AsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyr SEQ ID No. 59

ФНГ.16

PAK559:

TTAAATCTATAACTACAAAAACACATACAGGAATTCATTCAAGAATAGTTCAAACAA
AATTTAGATATTGATGTTTTTTGTGTATGTCCTTAAGTAAGTTCTTATCAAGTTTGTT

GAAGATTACAACTATCAATTCATACACAATATAAACGACGGGTACCAAAATAATGAAA
CTTCTAATGTTTGATAGTTAAAGTATGTGTTATATTTGCTGCCCATGGTTTTATTACTTT
MetLys

CTGAAAACCTGTAAGATCTGCGGTCCTTTTCGTCACTCTTTGCATCTCAGGTCCTTGGCCAA
GACTTTTGACATTCTAGACGCCAGGAAAGCAGTGAGAAACGTAGAGTCCAGGAACCGGTT
LeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeuGlyGln

CCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGATGGCTGACGACACTGAA
GGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACTACCGACTGCTGTGACTT
ProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThrGlu

TCTATCAACACTACTTTGGTTAACTTGCCAGGTGCTAAGAGATTCGTTAACCAACACTTG
AGATAGTTGTGATGAAACCAATTGAACGGTCCACGATTCTCTAAGCAATTGGTTGTGAAC
SerIleAsnThrThrLeuValAsnLeuProGlyAlaLysArgPheValAsnGlnHisLeu

TGCGGTTCCCACTTGTTGAAGCTTTGTACTT SEQ ID No. 60
ACGCCAAGGGTGAACCAACTTCGAAACATGAA
CysGlySerHisLeuValGluAlaLeuTyr SEQ ID No. 61

ФНГ.17

ПАХ562:

TTAAATCTATAACTACAAAAACACATACAGGAATTCATTCAAGAATAGTTCAAACAA
AATTTAGATATTGATGTTTTTGTGTATGTCCTTAAGTAAGTTCTTATCAAGTTTGT

GAAGATTACAAACTATCAATTTCATACACAATATAAACGACGGGTACCAAAATAATGAAA
CTTCTAATGTTTGATAGTTAAAGTATGTGTTATATTGCTGCCCATGGTTTTATTACTTT
MetLys

CTGAAAACCTGTAAGATCTGCGGTCCTTTTCGTCACTCTTTGCATCTCAGGTCCTTGGCCAA
GACTTTTGACATTCTAGACGCCAGGAAAGCAGTGAGAAACGTAGAGTCCAGGAACCGGTT
LeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeuGlyGln

CCAATTGACGACACTGAATCTATCAACACTACTTTGGTCAACTTGCCAGGTGCTAAGAGA
GGTTAACTGCTGTGACTTAGATAGTTGTGATGAAACAGTTGAACGGTCCACGATTCTCT
ProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGlyAlaLysArg

TTCGTTAACCAACACTTGTGCGGTTCCCACTTGGTTGAAGCTTTGTACTT SEQ ID No. 62
AAGCAATTGGTTGTGAACACGCCAAGGGTGAACCAACTTCGAAACATGAA
PheValAsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyr SEQ ID No. 63

ФНГ.18

CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGATG
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACTAC
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet

GCTGACGACACTGAATCTAGATTGCTACTAACACTACTTTGGATGTT
CGACTGCTGTGACTTAGATCTAAGCGATGATTGTGATGAAACCTACAA
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAspVal

GTAACTTGATCTCCATGGCT
CAATTGAACTAGAGGTACCGA
ValAsnLeuIleSerMetAla

SEQ ID No. 66

SEQ ID No. 27

ФНГ.19

AAGAGAGAAGAAGCTGAAGCTGAAGCTGAACCAAAGTTCGTTAACCAA
TTCTCTCTTCTTCGACTTCGACTTCGACTTGGTTTCAAGCAATTGGTT
LysArgGluGluAlaGluAlaGluAlaGluProLysPheValAsnGln

CACTTGTGTGGTTCTCACTTGGTTGAAGCTTTGTACTTGGTTTGCGGT
GTGAACACACCAAGAGTGAACCAACTTCGAAACATGAACCAAACGCCA
HisLeuCysGlySerHisLeuValGluAlaLeuTyrLeuValCysGly

GAAAGAGGTTTCTTCTACACTCCTAAGGCTGCTAAGGGTATTGTGCAA
CTTTCTCCAAAGAAGATGTGAGGATTCCGACGATTCCCATAACAGCTT
GluArgGlyPhePheTyrThrProLysAlaAlaLysGlyIleValGlu

CAATGCTGTACCTCCATCTGCTCCTTGTACCAATTGGAAACTACTGC
GTTACGACATGGAGGTAGACGAGGAACATGGTTAACCTTTTGATGACG
GlnCysCysThrSerIleCysSerLeuTyrGlnLeuGluAsnTyrCys

Xba I

AACTAGACGCAGCCCGCAGGCTCTAGA
TTGATCTGCGTCGGGCGTCCGAGATCT
Asn***

SEQ ID No. 71

SEQ ID No. 70

ФНГ.20

CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTGATG
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAGACAGTTGAACTAC
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet

GCTGACGACACTGAATCTAGATTCGCTACTAACACTACTTTGGCTTTG
CGACTGCTGTGACTTAGATCTAAGCGATGATTGTGATGAAACCGAAAC
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAlaLeu

GATGTTGTAACTTGATCTCCATGGCT
CTACAACAATTGAACTAGAGGTACCGA
AspValValAsnLeuIleSerMetAla

SEQ ID No. 69

SEQ ID No. 67

ФНГ.21

pAX614

TTAAATCTATAACTACAAAAACACATACAGGAATTCATTCAAGA
AATTTAGATATTGATGTTTTTTGTGTATGTCCTTAAGGTAAGTTCT

ATAGTTCAAACAAGAAGATTACAACTATCAATTCATACACAATATA
TATCAAGTTTGTTCTTCTAATGTTTGATAGTTAAAGTATGTGTTATAT

AACGACGGTACCAAAATAATGAACTGAAACTGTAAGATCTGCGGTC
TTGCTGCCATGGTTTTATTACTTTGACTTTTGACATTCTAGACGCCAG
MetLysLeuLysThrValArgSerAlaVal

CTTCGTCACTCTTTGCATCTCAGGTCCTTGCCCAACCAATTGACGAC
GAAAGCAGTGAGAAACGTAGAGTCCAGGAACCGGTTGGTTAACTGCTG
LeuSerSerLeuPheAlaSerGlnValLeuGlyGlnProIleAspAsp

ACTGAATCTAACACTACTTCTGTCAACTTGATGGCTGACGACACTGAA
TGACTTAGATTGTGATGAAGACAGTTGAACTACCGACTGCTGTGACTT
ThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThrGlu

TCTAGATTGCTACTAACACTACTTTGGTTAACTTGGCTAACGTTGCC
AGATCTAAGCGATGATTGTGATGAAACCAATTGAACCGATTGCAACGG
SerArgPheAlaThrAsnThrThrLeuValAsnLeuAlaAsnValAla

AACCAACACTTGTGTGGTTCTCACTTGGTTGAAGCTTTGTACTTATGG
TACCGATTCTCTAAGCAATTGGTTGTGAACACACCAAGAGTGAACCAA
MetAlaLysArgPheValAsnGlnHisLeuCysGlySerHisLeuVal

CTAAGAGATTCGTT
CTTCGAAACATGAA
GluAlaLeuTyr

SEQ ID No. 72

ФНГ.22
pAK625

TTAAATCTATAACTACAAAAACACATACAGGAATTCATTCAAGA
AATTTAGATATTGATGTTTTTGTGTATGTCCTTAAGGTAAGTTCT

ATAGTTCAAAACAAGAAGATTACAACTATCAATTTCATACACAATATA
TATCAAGTTTGTTCCTTAATGTTTGATAGTTAAAGTATGTGTTATAT

AACGACGGTACCAAAATAATGAACTGAAACTGTAAGATCTGCGGTC
TTGCTGCCATGGTTTTATTACTTTGACTTTTGACATTCTAGACGCCAG
MetLysLeuLysThrValArgSerAlaVal

CTTTCGTCACCTTTTGCATCTCAGGTCCTTGGCCAACCAATTGACGAC
GAAAGCAGTGAGAAACGTAGAGTCCAGGAACCGGTTGGTTAACTGCTG
LeuSerSerLeuPheAlaSerGlnValLeuGlyGlnProIleAspAsp

ACTGAATCTAACACTACTTCTGTCAACTTGATGGCTGACGACACTGAA
TGACTTAGATTGTGATGAAGACAGTTGAACTACCGACTGCTGTGACTT
ThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThrGlu

TCTAGATTGCTACTAACACTACTTTGGATGTTGTTAACTTGATCTCC
AGATCTAAGCGATGATTGTGATGAAACCTACAACAATTGAACTAGAGG
SerArgPheAlaThrAsnThrThrLeuAspValValAsnLeuIleSer

ATGGCTAAGAGAGAAGAAGCTGAAGCTGAAGCTGAACCAAAGTTCGTT
TACCGATTCTCTCTTCTTCGACTTCGACTTCGACTTGGTTTCAAGCAA
MetAlaLysArgGluGluAlaGluAlaGluAlaGluProLysPheVal

AACCAACACTTGTGTGGTTCTCACTTGGTTGAAGCTTTGTACTTG
TTGGTTGTGAACACACCAAGAGTGAACCAACTTCGAAACATGAAC
AsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyr

SEQ ID No. 73