



(19) **RU** ⁽¹¹⁾ **2 069 694** ⁽¹³⁾ **C1**

(51) МПК⁶ **C 12 N 15/52**

РОССИЙСКОЕ АГЕНТСТВО
ПО ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ РОССИЙСКОЙ ФЕДЕРАЦИИ**

(21), (22) Заявка: 4028319/13, 17.10.1986

(30) Приоритет: 18.10.1985 US 789206
18.03.1986 US 841121

(46) Дата публикации: 27.11.1996

(56) Ссылки: Патент Великобритании N 2091271, кл.
C 12 N 15/00, 1982.

(71) Заявитель:
Пфайзер Инк. (US)

(72) Изобретатель: Лэнс Стивен Дэвидов[US],
Джон Роберт Дизивув[US], Артур Эрнст
Фрэнке[US]

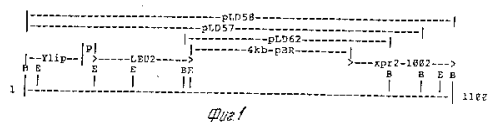
(73) Патентообладатель:
Пфайзер Инк. (US)

(54) СПОСОБ ПОЛУЧЕНИЯ ПРОРЕННИНА, СПОСОБ ПОЛУЧЕНИЯ ШТАММА ДРОЖЖЕЙ - ПРОДУЦЕНТА ПРОРЕННИНА, СПОСОБ ПОЛУЧЕНИЯ РЕКОМБИНАНТНОЙ ПЛАЗМИДНОЙ ДНК РХХ-33, СПОСОБ ПОЛУЧЕНИЯ РЕКОМБИНАНТНОЙ ПЛАЗМИДНОЙ ДНК РХХ-22, СПОСОБ ПОЛУЧЕНИЯ РЕКОМБИНАНТНОЙ ПЛАЗМИДНОЙ ДНК РХХ-11

(57) Реферат:

Использование: биотехнология, молекулярная биология, в частности получение проренина генно-инженерным путем. Сущность изобретения: проренин получают путем конструирования рекомбинантных плазмидных ДНК PLS/3 или РХХ-11, или РХХ-22, или РХХ-33, или РХХ-34, трансформации полученными рекомбинантными ДНК штаммов *J.lipolytica* с последующим культивированием

трансформантов, выделением и очисткой целевого продукта. При этом выделены штаммы *L.lipolytica* ATCC-206.88, или 20774, или 20794. Раскрыто конструирование плазмид РХХ 33, РХХ-22 и РХХ-11. 5 с. и 1 з.п. ф-лы, 17 ил., 8 табл.



RU 2 069 694 C1

RU 2 069 694 C1



(19) **RU** ⁽¹¹⁾ **2 069 694** ⁽¹³⁾ **C1**
 (51) Int. Cl.⁶ **C 12 N 15/52**

RUSSIAN AGENCY
 FOR PATENTS AND TRADEMARKS

(12) **ABSTRACT OF INVENTION**

(21), (22) Application: 4028319/13, 17.10.1986
 (30) Priority: 18.10.1985 US 789206
 18.03.1986 US 841121
 (46) Date of publication: 27.11.1996

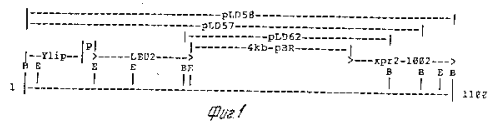
(71) Applicant:
 Pfajzer Ink. (US)
 (72) Inventor: Lehns Stiven Dehvidov[US],
 Dzhon Robert Diziuv[US], Artur Ehrnst
 Frehnke[US]
 (73) Proprietor:
 Pfajzer Ink. (US)

(54) METHOD OF PRORENNIN PREPARING, METHOD OF PREPARING YEAST STRAIN - A PRODUCER OF PRORENNIN, METHOD OF PREPARING RECOMBINANT PLASMID DNA PX-33, METHOD OF PREPARING RECOMBINANT PLASMIN DNA PX-22, METHOD OF PREPARING RECOMBINANT PLASMID DNA PX-11

(57) Abstract:

FIELD: biotechnology, molecular biology, genetic engineering. SUBSTANCE: prorennin is prepared by construction of recombinant plasmid DNA PLS/3 or PXX-11, or PX-22, or PX-33, or PX-34, transformation of the strains *L. lipolytica* with prepared recombinant DNAs followed by transformants cultivation, isolation and purification of the end product. Strains *L. lipolytica* ATCC

- 206.88, or 20774, or 20794 were isolated. Construction of plasmids PXX-33, PXX-22 and PXX-11 is described. EFFECT: improved method of preparing. 2 cl, 17 dwg



RU 2 0 6 9 6 9 4 C 1

RU 2 0 6 9 6 9 4 C 1

Изобретение относится к получению гетерологичных протеинов в клетках дрожжей с использованием рекомбинантных векторов клонирования *Yarrowia lipolytica*, содержащих чужеродную ДНК, кодирующую экспрессию и секрецию протеинов млекопитающих, векторов экспрессии, содержащих промотор гена *Y.lipolytica* (например XPR2 или LEU2), сигнальную (или предсигнальную) последовательность щелочной протеазы, область pro и участок терминатора XPR2, а также их вариантов и функциональных эквивалентов, учитывая вырожденность генетического кода, или с использованием других генетических компонентов *Y.lipolytica*.

S. cerevisiae имеет ряд ограничений как система секреции протеинов. Европейская патентная заявка N 0123544, опубликованная 31 октября 1984, описывает выделение генов α -фактора и использование их промотор и/или сигнальный пептид частей в сочетании с ДНК, кодирующей чужеродные для дрожжей протеины, в плазмиде для трансформации клеток дрожжей, способных продуцировать отдельный зрелый протеин в клеточной культуре. Европейская заявка N 0088632, опубликованная 14 сентября 1983, описывает способ экспрессии и выделения чужеродного протеина в *S.cerevisiae*. Однако размер протеинов, которые *S.cerevisiae* эффективно секретирует с использованием с данных и других систем секреции, ограничен примерно 20000 дальтон. Преодоление этой неэффективности *S.cerevisiae* как секретирующего организма потребовало многих мутационных изменений как описано Smith и др. Science 229, 1219-1224 (1985). Единственным исключением для этого направления является наблюдение, что ферменты *Aspergillus* величиной более 20000, очевидно могут секретироваться *S.cerevisiae*, но эти ферменты в значительной степени гликозилируются *S.cerevisiae*, что может влиять на эффективность секреции.

Особый интерес представляет *Yarrowia lipolytica*, промышленно важный вид дрожжей, используемый для производства лимонной кислоты и клеточных протеинов. Он также может быть использован для получения эритрита, таннина и изопропилаблочной кислоты. В противоположность *S.cerevisiae*, *Y.lipolytica* представляет особый интерес и ценность из-за своей способности эффективно выделять высокомолекулярные протеины (щелочную протеазу, кислую протеазу и ДНК-азу) в культуральную среду, представляя таким образом потенциальную возможность извлечения чужеродных протеинов в нативном состоянии без необходимости разрушения продуцирующих клеток. Кроме того, *Y.lipolytica* количественно выделяет очень мало протеинов, таким образом предоставляя возможность получения нужного чужеродного протеина в питательной среде в качестве доминирующей разновидности протеинов и облегчая извлечение указанного чужеродного протеина.

Y. lipolytica производит большое количество внеклеточной протеазы. Это доминирующий протеин, выделяемый *Y.lipolytica*. Вид протеазы (щелочная, кислая или нейтральная) зависит от используемого штамма *Y.lipolytica* (Ogrydziak и др. I.Gen. Microbiol (1982) 128, 1225-1234). Частичный

анализ N-концевой аминокислотной последовательности щелочной внеклеточной протеазы описан Ogrydziak и др. (местное цитирование).

Находящаяся на рассмотрении заявка N 634505, поданная 25 июля 1984, описывает методы трансформации *Y.lipolytica* и клонирования генов *Y.lipolytica* путем комплементации мутаций. В ней описано клонирование гена XPR2, который кодирует секретируемую щелочную протеазу, путем комплементации xpr2 мутаций *Y.lipolytica*. Метод включает трансформацию штамма-хозяина *Y.lipolytica* частичным BglII-переваром библиотеки генов *Y.lipolytica* в векторе pLD 40, описанный в Европейской заявке N 0138508, опубликованный 24 апреля 1985, являющейся частью вышеуказанной заявки США.

Предложенное изобретение раскрывает процесс экспрессии и секреции зрелых гетерологичных протеинов, в частности прореннина и анафилатоксина с 5 человек в культурах *Y.lipolytica*. Способ, раскрытый в настоящей заявке, может быть применен с соответствующими необходимыми изменениями для получения и выделения любых известных чужеродных протеинов, перечисленных в патенте США N 4532207, опубликованном 30 июля 1985 г. Кроме того, любой другой ген *Y.lipolytica* для секреции протеинов, такой как гены рибонуклеазы и кислой протеазы, может быть использован вместо гена XPR2, а также гибридные гены, сконструированные путем объединения фрагментов двух или более указанных генов, например сигнальной последовательности гена XPR2 и промоторной последовательности гена рибонуклеазы.

Экспрессия и секреция в *Y.lipolytica* прореннина и анафилатоксина C5a человека достигнута с использованием сигналов экспрессии и секреции генов *Y. lipolytica* XPR2 и/или LEU2. ДНК-последовательности для прореннина и человеческого анафилатоксина C5a были "сшиты" с помощью синтетических олигонуклеотидов с генной последовательностью XPR2 в сайтах, предположительно кодируемых сигнальный пептид щелочной протеазы, или в сайтах процессинга предшественника протеазы, обозначенных здесь как pro1 и pro2, и использованы для получения генных конструкторов, которые были затем введены в *Y.lipolytica* путем интегративной трансформации. Рекомбинантные культуры экспрессировали и секретируют в питательную среду чужеродные протеины, имеющие молекулярный вес и иммунореактивность прореннина и человеческого анафилатоксина C5a. Прореннин, полученный таким образом находится в нативной конфигурации, так как после удаления пропептида он обнаруживает полную ферментативную активность.

Использовались следующие материалы и методы

Материалы. Рестрикционные эндонуклеазы и T4 лигаза были получены из New England Biolabs, бактериальная щелочная фосфатаза из Bethesda Research Laboratories, T4 полинуклеотидкиназа из PL-Biochemicals и [гамма-³²p] АТФ из New England Nuclear. Все ферменты использовались при условиях,

рекомендованных поставщиком.

Среды. Среда ГПП (среда глицерин/протеаза-пептон) содержала (на литр): 6,7 г глицерина, 1,6 г Difco протеаза-пептона, 1,7 г Difco азотистого основания дрожжей без аминокислот и сульфата аммония, 30 мг урацила и 0,5 мл/л полипропиленгликоля мол.в. 2000 (Polysciences) в 40 мМ-фосфатном буфере (рН 6,8) (полипропиленгликоль не был включен, когда среда использовалась для выращивания культур для анализа ферментной активности реннина). Протеаза-пептон был отдельно обработан в автоклаве в фосфатном буфере.

Среда ДЭПД (среда дрожжи-экстракт/пептон/декстроза) содержала (на литр): 5 г экстракта дрожжей, 10 г пептона и 20 г декстрозы.

E. coli выращивалась в среде LB при 37 ° С. Среда LB содержала (на литр): 10 г Бактотриптона, 10 г Бактодрожжевого экстракта, 10 г хлористого натрия, доводилась до рН 7,5 гидроокисью натрия.

Анализ последовательности ДНК. Фрагменты ДНК из различных здесь описанных плазмид изолировались на геле из полиакриламида и секвенировались по методу Maxam и др. *Methods in Enzymology* 65, 499 (1980).

Процедуры лигирования. Фрагменты ДНК, включая расщепленные векторные плазмиды, лигировались путем смешения нужных компонентов (фрагменты ДНК с концевыми участками, специально сконструированными для обеспечения правильного спаривания) с Т4 ДНК-лигазой. Добавлялось примерно 10 единиц лигазы на мкг количества вектора и вводимых фрагментов. Реакционная лигирующая смесь трансформировалась в компетентные клетки штаммов MM294 (ATCC-33625) или HB101 (ATCC-33694) *E. coli* K12.

Получение химически синтезированной ДНК. Для конструирования гибридных генов для экспрессии и секреции прореннина восемь олигонуклеотидов были синтезированы по модифицированной фосфорамидатной методике (Sinha и др. *Tetrahedron letters* 24, 5843 (1983) на Genetic Design 6500 (Watertown, MA) автоматическом ДНК-синтезаторе и очищены на 6М мочевино-20% полиакриламидном геле. Аликвоты комплементарных олигонуклеотидов смешивались и сшивались друг с другом при 4°C в течение ночи в ТЭ (10 мМ трис-НСl, рН 8,0, 1 мМ натриевой соли этилендиаминтетрауксусной кислоты).

Аликвоты (около 2 мкг) двухцепочечных олигонуклеотидов фосфорилировались в 20 мкл реакционной смеси, содержащей 70 мМ трис (рН 7,6), 10 мМ MgCl₂, 5 мМ дитиотрейтола, 5 мМ АТФ, при 37° С, используя Т4 полинуклеотидкиназу.

Получение плазмидной ДНК. Плазмидная ДНК в больших количествах получалась по методике Holmes и др. *Anal biochem* 114, 195-197 (1981), с последующим центрифугированием в градиенте плотности в этидиум-бромид-СsCl. Миниколчества плазмидной ДНК получались по щелочной SDS методике Birnboim и др. *NAR* 1, 1513 (1979).

Конструирование векторов экспрессии/секреции для прореннина. Была

приготовлена серия различных конструкций для получения конечных векторов экспрессии. Все стадии представлены в виде схемы на фигурах. В общем случае, фрагменты ДНК изолировали путем гель-электрофореза и сшивали с другими фрагментами или расщепленной плазмидной ДНК в 20 мкл 50 мМ трис-НСl (рН 7,5), 10 мМ MgCl₂, 20 мМ дитиотрейтола, 1 мМ АТФ, и 200 единиц Т4 лигазы при 4°C. Если требовалось частичное расщепление ДНК рестрикционной эндонуклеазой, оптимальное время расщепления устанавливалось экспериментально.

Идентификация прореннина в культуральной жидкости

Трансформанты дрожжей, содержащие векторы экспрессии, выращивались в течение ночи в среде ГПП (см. выше). После центрифугирования для удаления клеток дрожжей, добавляли 1 мл 50% трихлоруксусной кислоты к каждой 5 мл аликвоте культуральной жидкости и выдерживали при 4°C 60 мин. Центрифугированием получали шарики, которые дважды промывали 2 мл холодного ацетона. Осажденный протеин растворяли в 100 мкл стандартного SDS буфера и аликвоты подвергали электрофорезу на 10% SDS-полиакриламидных гелях (Laemmli U.K. (1970) *Nature* 227, 680). Освобожденные из геля протеины электрофоретически переносили на нитроцеллюлозу (Schleicher and Schuell, 0,22 мкм) и прореннин определяли иммуноточечным анализом на пластинке геля (Hawkes R. и др. (1982) *Anal. Biochem.* 119, 142). На фильтр наносили кроличье антитело к прореннину с последующей инкубацией с конъюгированным с пероксидазой козьим антителом к антителу кролика (Cappel, Malvern, Pa). Связанные антитела определялись окрашиванием смесью 4-хлор-1-нафтол и перекись водорода.

Активность по свертыванию молока прореннина в культуральной жидкости. Культуральная жидкость различных трансформантов *Y. lipolytica* исследовалась на активность по свертыванию молока по модифицированному методу Erbstrom, I. *Dairy Sci* 41, 1664 (1958). Коротко говоря, исследование включало измерение времени, требуемого реннину в активированных супернатантах культуры для свертывания буферного снятого молока и соотношения значений со стандартными для очищенного реннина. Культуры дрожжей (25 мл) выращивались в течение ночи в среде ГПП. После центрифугирования для удаления клеток 5 мл аликвоты супернатантов культуры высушивались замораживанием под вакуумом. Каждый лиофилизированный супернатант был ресуспендирован в 300 мкл дистиллированной воды. Серия разбавленной очищенного прореннина теленка была также приготовлена в качестве контрольного стандарта. Прореннин в среде концентратов и контрольные образцы активировались добавлением примерно 5-10 мкл конц. НСl до получения рН приблизительно 2 и инкубировались один час при 22°C. Снятое молоко приготавливали путем прибавления 60 г сухого порошка снятого молока (Difco) к 500 мл 41,5 мМ ацетата натрия (рН 6,3) и 13,6 мМ CaCl₂ и перемешивания 20 мин при 4 °С.

Субстрат был использован для анализа немедленно после приготовления. Аликвоту 60 мкл (эквивалент 1 мл супернатанта культуры) каждого препарата фермента добавляли к 1 мл аликвоты снятого молока при 37°C и регистрировали время свертывания.

Получение синтетических олигонуклеотидов для гена С5а

Олигодезоксинуклеотиды, используемые в синтезе структурного гена С5а, получали по модифицированной фосфорамидатной методике (Sinha и др. ук.соч.) с использованием стеклянной подложки с контролируемым размером пор на Genetic Design 6500 (Watertown, MA) автоматическом ДНК-синтезаторе. При получении

использовали 3% (вес/объем) дихлоруксусную кислоту в дихлорметане для удаления тритильной группы, активирование фосфорамидатов насыщенным тетразолом в ацетонитриле, обработку ди-этоксифосфинтетразолидом и окисление водным раствором иод/ТТФ/Matteucci и др. 1981, I. Am. Chem. Soc. 105, 3183). Полное время на цикл добавления составляло 14 мин. Десять 47-мер сегментов А-I рисунка 9 получали с 98,8% средним выходом на стадию (по тритильному анализу), деблокировали по методике Matteucci и др. местное цитирование, осаждали этанолом из 0,3 М ацетата натрия и выделяли путем препаративного гель-электрофореза в 10% денатурированном геле полиакриламид-мочевина перед сшиванием.

Сборка, клонирование и секвенирование человеческого гена С5а

На фиг. 9 показана аминокислотная последовательность нужного протеина и расположение синтетических олигонуклеотидов, необходимых для получения гена, кодирующего человеческий протеин С5а. Все олигомеры, кроме А и F, фосфорилировались по их 5'-концам Т4 полинуклеотидкиназой. Сборка гена включала две первичных реакции отжиг/лигирование, включающие: олигомеры А, В, I и J и олигомеры С, D, E, F, G и H. Образующиеся В 94 п.о. и 141 п.о. двухцепочечные фрагменты ДНК изолировались после электрофореза в 10% полиакриламидном геле, сшивались вместе в их В 235 п.о. продукт изолировался гель-электрофорезом. В 235 п.о. фрагмент ДНК, содержащий структурный ген, кодирующий С5а, вводился между EcoRI и HindIII сайтами вектора рBP322 и трансформировался в компетентные клетки штамма RV101 E.coli K-12. Рестрикционный анализ плазмидной ДНК, изолированной из 6 трансформантов, показал, что 5 из 6 клонов содержали EcoRI/EcoRI/HindIII фрагмент правильного размера. Нуклеотидную последовательность участка гена С5а каждой из этих плазмид определяли по методу Maxam и др. Methods Enzymol. 65, 499 (1980).

Конструирование и характеристика плазмиды экспрессии С5а для E.coli.

Методика изолирования фрагментов ДНК и условия реакции сшивания были такими же, как описаны Maniatis и др. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor E.coli trp промотор-оператор был первоначально получен из ptr pL1 (Edman и др.) (1981) Nature 291, 503). 360 п.о. фрагмент EcoRI, содержащий trp

промотор-операторную последовательность, используемый в плазмиде экспрессии С5а (pC5A-48), изолировался из плазмиды экспрессии прореннина рPFZ-R2, описанной в Европейской заявке N 0147178, опубликованной 3 июля 1985 г.

Идентификация С5а в культуральной жидкости Y.lipolytica.

Методика была такой же, как описана выше для прореннина, за исключением того, что в иммуноанализе были использованы козы антитела к С5а и кроличьи антитела к козьему IgG (Cappel). Козье антитело к С5а человека получали методом Mandezino и др. J.Ymmunol. Methods 53, 41-50 (1982).

Векторы

pLD 40 описанный в Европейской заявке N 0128508, опубликованной 24 апреля 1985 г.

Микроорганизмы

ATCC 20774 Yarrowia lipolytica PC 30869

ATCC 20781 Yarrowia lipolytica DL 112-PC-30869, трансформант с XPR2

ATCC 20776 Yarrowia lipolytica DL-148 Трансформант Y.lipolytica ATCC 20688 с SnaBI дайджестед pLS-3

ATCC 20775 Yarrowia lipolytica DL-144 Трансформант Y.lipolytica ATCC 20688 с нерасщепленной pLs-3

ATCC 20777 Трансформант Y. lipolytica PC-30869 с SnaBI, расщепленной pC5aX-3

ATCC 20778 Трансформант Y. lipolytica PC-30869 с SnaBI, расщепленной pXX-II

ATCC 20779 Трансформант Y. lipolytica PC-30869 с SnaBI, расщепленной PXX-22

ATCC 20780 Трансформант Y. lipolytica PC-30869 с SnaBI, расщепленной pXX-33

ATCC 20794 Трансформант Y.lipolytica PC-30869 с pLD56

ATCC 20795 Трансформант Y.lipolytica ATCC 20794 с NruI, расщепленной pLX-34

Они были депонированы в соответствии с Будапештским договором в американской Коллекции типовых культур, Rockville, Maryland.

Таксономическое исследование Y.lipolytica ATCC 20774 (идентифицированного в коллекции культур Pfizer, Inc. как PC 30869) проведено д-ром J.R.Dezeew, который подготовил нижеследующее описание. Использовались методики, рекомендованные J.L.Lodder в "The Yeasts", второе издание, N.Holland Publishing Co. Амстердам, 1970.

CBS 599, типовая культура вида Candida lipolytica ("The Yeasts", второе издание, N.Holland Publishing Co. Амстердам, 1970) и CBS 6124, типовая культура Saccharomcopsis lipolytica в "The Yeasts", третье издание, были взяты для сравнения. Ранее вид был также сопоставлен с Endomycopsis lipolytica. Его состояние Candida lipolytica.

Таксономическое положение вида было установлено van der Walt и von Arx, Antonie van Leeuwenhoek 46, 517-521 (1980). Предпочитаемое название в настоящее время Yarrowia lipolytica.

Культуральные, морфологические и физиологические характеристики штамма PC 30869 согласуются со стандартным описанием вида, указанного как Saccharomycopsis lipolytica в "The Yeasts", третье издание /под ред. Kreger-Van Rij, сс. 406-408, Elsevier Science Publishers B.V. Амстердам, 1984.

Сравнение штаммов Yarrowia lipolytica представлено в табл.1.

PC-30869 был сконструирован с помощью генетической рекомбинации подходящих мутантов *Y.lipolytica* PC-22208, Pfizer почвенный изолят, и *Y.lipolytica* PC-30026, субкультура NRRL Y-1094. PC-30869 фенотипически отличается от своих родителей дикого типа следующим: (1) не производит активной внеклеточной щелочной протеазы, (2) требует биотина для роста и (3) требует источник L-лейцина.

В течение лог-фазы роста PC-30869 в среде дрожжевой экстракт-пептон-глюкоза (ДЭПГ) почкующиеся клетки имеют яйцевидную форму и средний размер 2,6x5,5 мкм. На ДЭПГ-агаре доминируют псевдо- и истинные мицелии. Бластоспоры присутствуют в большинстве как одиночки. Не обнаруживаются каротиноидных пигментов. Культура ведет себя как гаплоид В-типа спаривания в скрещивании с аутентичными тестовыми штаммами видов (табл. 5). Типичная аскоспоруляция наблюдается на агаре V8. Особенности ассимиляции углерода указаны в табл. 2. Ферментация отсутствует. Ион аммония и мочевины, но не нитрат, утилизируются как единственные источники азота (табл. 3). Штамм PC-30869 требует витаминов тиамин и D-биотин (табл. 4). Только тиамин требуется для родителей культуры дикого типа. Никакого роста не наблюдается при 37 °С. Другие характеристики представлены в табл.6.

PC-30869 отличается от других штаммов *Y.lipolytica*, описанных в патентной литературе, как это следует из сравнения их фенотипов (табл. 7 и 8).

ATCC 20228 (Nubel и др. Патент США N 4155811) обнаруживает вид питания дикого типа, подобно типовым штаммам для вида CBS 599 и CBS 6124. Конкретно, он не нуждается для роста в урациле, лейцине или биотине, и он разжижает желатин.

ATCC 20628 (Dezeew и др. Патент США N 4407953) в отличие от ATCC 20228 требует дополнительного лейцина для роста. Подобно ATCC 20228 он не нуждается в урациле или биотине. Он также разжижает желатин.

ATCC 20688 (Европейская заявка N 0138508) растет, только если в среду добавлены как урацил, так и лейцин. Эта необходимость в урациле отличает ATCC 20688 как от ATCC 20228, так и ATCC 20628. ATCC 20688 не требует биотин и разжижает желатин.

Культура PC-30869 отличается от всех вышеуказанных. Она нуждается для роста в биотине и лейцине, но не в урациле. Она не разжижает желатин.

Полная среда содержала 16,7 г/л бактодрожжевой основы, свободной от витаминов, плюс 100 мг/л урацила, 100 мг/л лейцина, 10 мкг/л D-биотин и 200 мкг/л тиамин, HCl.

Среда содержала 120 г/л желатина и 16,7 г/л бактодрожжевой основы, свободной от витаминов, плюс 100 мг/л урацила, 100 мг/л L-лейцина, 10 мкг/л D-биотин и 200 мкг/л тиамин, HCl.

На фиг.1 показана частичная линейная рестрикционная карта перекрывающихся плазмид pLD 57, pLD 58 и pLD 62, изолированных из штамма *Y.lipolytica* DL 112.

На фиг.2 пробы синтетических олигонуклеотидов для гена XPR2. Из описанной последовательности для

большинства первых 25 аминокислотных остатков активной протеазы (Ogrudziak и др. ук.соч.) два участка, отмеченные I и II, дают возможность конструирования 14-мер олигонуклеотидных проб с 32-кратной или меньшей вырожденностью. Эти два участка начинаются соответственно у аминокислот 7 и 18. Четыре различные восьмикратно вырожденные смешанные пробы были приготовлены для каждого участка, они получили номера от 170 до 186, как показано. В изображенной предсказанной последовательности нуклеиновой кислоты "X" означает все 4 основания, "U" означает оба пурина и "Y" означает оба пиримидина.

На фиг. 3-5 нуклеотидная последовательность гена XPR2 с указанием промотора, пре (от -157 до -136), про 1 (от -135 до -98), про 2 (от -97 до -1), щелочной внеклеточной протеазы и концевых последовательностей.

На фиг. 6 сконструированная последовательность для терминатора вектора pterm 4.

На фиг. 7 сконструированная последовательность и рестрикционная карта плазмиды pLS-3.

На фиг. 8 сконструированная последовательность и рестрикционная карта плазмиды pXX-33.

На фиг. 9 сконструированная последовательность и рестрикционная карта плазмиды pXX-22.

На фиг. 10 сконструированная последовательность и рестрикционная карта плазмиды pXX-11.

На фиг.11 аминокислотная последовательность человеческого анафилатоксина C5a.

На фиг.12 рестрикционная карта плазмиды pC5a-48.

На фиг. 13 сконструированная последовательность и рестрикционная карта плазмиды pC5aX-3.

На фиг.14-16 нуклеотидная последовательность гена LEU-2.

На фиг. 17 сконструированная последовательность и рестрикционная карта плазмиды pLX-34.

Анализ последовательности гена XPR2. Анализ ДНК-последовательности клонированного гена XPR2 проводился методом химической дегградации (Maxam и др. 1980, *Methods Enzymol* 65, 499) на перекрывающихся фрагментах рестрикции, полученных из плазмид pLD 57, pLD 58, pLD 62 (фиг.1) и pLD 84 и pLD 86 (см. ниже). Результаты показали, что клонированные геномные ДНК дрожжей действительно содержат ген для внеклеточной щелочной протеазы. Нуклеотидная последовательность гена XPR2 и аминокислотная последовательность предшественника щелочной протеазы с ее сигнальной последовательностью, выведенная из нуклеотидной последовательности, показаны на фиг.3-5. Значительная часть аминокислотной последовательности внеклеточной протеазы была неизвестна (Ogrudziak и др. ук. соч.) и представлена здесь впервые. Более того, последовательности, требуемые для экспрессии и выделения внеклеточной протеазы, описаны здесь в первый раз. Последовательность ДНК, кодирующая

щелочную протеазу, ее предшественник и сигнальные последовательности, состоит из 1362 пар оснований (фиг.3-5). Первичная структура этой полипептидной цепи, выводимая из нуклеотидной последовательности, должна содержать 454 аминокислотных остатка. Щелочная протеаза синтезируется в клетке в форме предшественника, который протеолизитически переводился в секретируемую или зрелую форму. Анализ N-концевой аминокислотной последовательности, выведенной из нуклеотидной последовательности, показал наличие сигнального пептида в молекуле предшественника. Этот сигнальный пептид содержит 22 аминокислотных остатка и его структурные особенности подобны особенностям сигнальных пептидов высших эукариотов и прокариотов (Perlman и др. 1983, J. Mol. Biol. 167, 391). Участку в выведенной аминокислотной последовательности в согласии с известными 25 N-концевыми аминокислотами активной щелочной протеазы (Ogrydziaк и др. 1982, J. Gen. Microbiol, 128, 1225) предшествуют 157 аминокислотных остатков, содержащих сигнальный пептид и два сайта расщепления трипсинового типа (Lys-Arg). Эти сайты расщепления использовались для разделения области rpo на rpo 1 (от -135 до -98) и rpo 2 (от -97 до -1), см. фиг.3-5. Активная щелочная протеаза имеет 297 аминокислот, как следует из нуклеотидной последовательности. Аминокислотные последовательности для различных форм протеазы, выведенные из нуклеотидной последовательности, находятся в сочетании с размерами очищенных форм фермента. В дополнение к структурной последовательности предшественника щелочной протеазы было определено примерно 700 п.о. 5'-боковой последовательности и 600 п.о. 3'-боковой последовательности. Анализ этих участков показал, что они содержат последовательности, аналогичные другим эукариотическим промоторам и терминаторам, и, вероятно, они играют существенную роль в экспрессии щелочной протеазы.

Как указано выше, методика трансформации *Y.lipolytica* и клонирования генов *Y. lipolytica* путем комплементации мутаций, включая клонирование гена XPR2, который кодирует выделяемую щелочную протеазу, путем комплектации мутации xpr², описаны в Европейской заявке N 0138508. Методика, описанная в ней, включает трансформацию клетки-хозяина штамма *Y.lipolytica* частичным Bg Ш-переваром библиотеки генов *Y.lipolytica* в векторе pLD 40. Данный вектор характеризуется тем, что он содержит в себе малый сегмент, содержащий LEU2 участок *Y.lipolytica* и сайты рестрикции эндонуклеазы 3EcoRI, 4EcoRV, 6A val, IBgIII, INCoI, IAPaI, 2XhoI и IBst XI. Один из трансформантов XPR2 *Y.lipolytica* были использованы для извлечения гена дико типа (pLD 84 и pLD 86) из *Y.lipolytica* NRRL Y-1094 для применения в конструировании вектора экспрессии/секреции, как описано в примере 1.

Анализ последовательности гена LEU2. Анализ последовательности ДНК

клонированного гена LEU2 в pLD 25 (Европейская заявка N 0138508) проводился методом химической деградации (Maxam и др. 1980, Methods, Enzymol. 65, 499) на перекрывающихся фрагментах рестрикции. Для локализации области, кодирующей бета-изопротилмалат (ИПМ) дегидрогеназу, и соответствующей рамки считывания была использована информация о аминокислотной последовательности, ранее определенной для гена LEU2 *S.cerevisiae* (Andreadis и др. 1984, J. Biol. Chem. 259, 8059). Участок гомогенной последовательности *Y.lipolytica* который кодирует аминокислотную последовательность, гомологичную участку протеиновой последовательности *S.cerevisiae* был идентифицирован. Нуклеотидная последовательность гена 2,8-kb LEU2 и аминокислотная последовательность бета ИПМ-дегидрогеназы, выведенная из нуклеотидной последовательности, показаны на фиг. 14-16. Кроме того, последовательности, требуемые для экспрессии бета-ИПМ-дегидрогеназы *Y. lipolytica* описаны здесь впервые. Последовательность ДНК, кодирующая этот протеин, состоящий из 405 аминокислот, содержит 1215 пар оснований (фиг. 14-16). В дополнение к последовательности, кодирующей бета-ИПМ-дегидрогеназу, были определены примерно 798 п.о. 5'-боковой последовательности и 797 п.о. 3'-боковой последовательности (включая кодон терминации трансляции TAA). Анализ этих участков показал, что они содержат последовательности, аналогичные другим эукариотическим промоторам и терминаторам, и что они имеют существенное значение для экспрессии.

5'-участок выше гена LEU2 *Y. lipolytica* содержит TATATA-последовательность в 78 п.о. перед сайтом начала трансляции и в 30 п.о. перед предлагаемым сайтом начала транскрипции мРНК. Вторая последовательность, важная для инициации транскрипции в эукариотах, представляет собой бокс CAAT, который в гене LEU2 расположен в 74 п.о. перед предлагаемым сайтом инициации транскрипции, расположенным в -48 п.о. от ATG (фиг.14-16).

3'-участок ниже гена имеет последовательность от 72 до 120 нуклеотидов после стоп-кодона (TAA), гомологичную последовательности 5'-TAG.TA/T/GT. TTT-3', предложенной Zaret и др. Cell 28, 563 (1982) как имеющей важное значение для терминации транскрипции в *S.cerevisiae*.

Изобретение иллюстрируется следующими примерами.

Пример 1. Использованным штаммом-хозяином был ATCC 20774 (MATB Ieu 2-40 bio-6 xр 2-1002). Трансформант XPR2, *Y.lipolytica* ATCC 20781, определялся как колония, образующая зону на индикаторных чашках со снятым молоком после посева методом отпечатков на чашки с недостатком лейцина. Хромосомную ДНК получали из трансформанта по методу заявки EP N 0138508 и использовали для извлечения гена для выделяемой протеазы. Хромосомную ДНК частично расщепляли ферментом BgIII, сшивали для образования кольцевого фрагмента, содержащего репликон *E.coli* и ген устойчивости к ампициллину из вектора, и использовали для

трансформации *E.coli*. Хромосомную ДНК также расщепляли ферментом *Sall* и использовали в эксперименте по Southern, который показал, что нормальный участок LEU2 трансформанта не был нарушен. (520 п.о. *Sall*-Eco RI сегмент области LEU2, представляющий 5'-конец сегмента LEU2 pLD40, использовался в качестве пробы). Поэтому, так как гомология необходима для интеграции плазмидной библиотеки в *Y.lipolytica*, участок XPR2 должен быть сайтом интеграции. Три перекрывающиеся, но различные плазмиды, pLD 57- pLD 58 и pLD 62, были первоначально извлечены из *Y.lipolytica* ATCC 20781. Они изображены на фиг. 1. Гибридизация с синтетическими олигонуклеотидными пробами гена XPR2, основанными на известной последовательности первых 25 аминокислотных остатков выделенной активной протеазы (фиг. 2-5), показали, что ген секретирующей протеазы был клонирован. Для определения того, представляет ли извлеченный ген копию дикого типа или мутантную копию, штамм-реципиент *Y.lipolytica* был трансформирован pLD 58. Так как не возникло трансформантов, положительных по протеазе, из любых лейцин-независимых трансформантов, был сделан вывод, что pLD 58 содержит мутантный аллель гена.

Форма гена XPR2, присутствующая в штамме NRRL Y-1094 дикого типа, была получена с *E.coli* в эксперименте по гибридизации колонии. В качестве пробы использовался 2 тыс.п. фрагмент *PvuI*-EcoRI, который, как следует из данных по секвенированию, содержит полный структурный ген. Из оригинальной библиотеки частично расщепленных *Sau* 3A фрагментов NRRL Y-1094 ДНК в pLD 40, описанного в EP 0138508, было получено несколько колоний, гибридизующихся с пробой. Две из этих колоний содержали очень схожие плазмиды, обозначенные pLD84 и pLD86, которые были использованы для получения векторов экспрессии. Обе плазмиды содержат одинаковый 5'-конец участка XPR2-сайт *Sau* 3A (который был присоединен и восстанавливал сайт *Bam*HI вектора), от которого начинается последовательность на фиг.3-5. Каждый содержит целиком структурный ген протеазы и предполагаемый терминатор транскрипции и включает приблизительно 4-5 тыс. п. введенных из участка XPR2 штамма NRRL Y-1094. Включение в pLD 86 содержит дополнительно несколько сотен пар оснований на 3'-конце. Так как мы использовали 3'-цепь до сайта *Bg*III (2655 пар оснований) для конструирования вектора экспрессии, две плазмиды обеспечивают одинаковую ДНК, которая функционально идентична последовательности на фиг.3-5

Конструирование векторов экспрессии/секреции. План, разработанный для получения экспрессии и секреции прореннина в *Y.lipolytica*, предусматривает конструирование различных гибридных генов в интегративном векторе клонирования. Подобный подход требует нескольких различных плазмид, которые содержат обширные участки общих последовательностей ДНК. Фактически была использована модулированная

конструкционная схема для сборки векторов с геном прореннина, введенным в положение 3' предсказанного сайта XPR2 процессинга сигнального пептида, предлагаемым рго I сайтом процессинга и сайтом расщепления для получения зрелой щелочной протеазы. Желательно, чтобы чужеродный ген был введен между последовательностями дрожжей для промотора и терминатора экспрессии. Ясно, что N-концевая часть последовательностей гибридного гена будет изменяться в различных конструкциях плазмид, однако последовательность структурного гена прореннина, последовательность терминатора XPR2 и ДНК челночного вектора будут одинаковы в каждой конструкции плазмиды экспрессии. Предполагалось, что один и тот же фрагмент структурного гена прореннина и терминатор/векторная плаزمида будут использованы в каждой конструкции плазмиды экспрессии, как описано ниже. Различные конструкции плазмид экспрессии секреции прореннина различаются в участке сразу ниже от промоторной последовательности гена XPR2 по длине N-концевой последовательности предшественника щелочной протеазы, которая предшествует последовательности гена прореннина. Поэтому промоторный фрагмент каждой плазмиды экспрессии конструировался как изменяемая последовательность в области соединения XPR2 прореннин. Все векторы экспрессии секреции были собраны сходной реакцией лигирования, включающей три компонента.

Экспериментальные стадии, использованные для конструирования вектора-терминатора pterm 4, изображены на фиг.6. Сначала синтетический линкер присоединяли к фрагменту, содержащему 3'-конец гена XPR2, включающему сигналы окончания транскрипции и полиаденилирования. Плазмиду pLD 84 расщепляли эндонуклеазой *Kpn*I и лигировали с синтетической двухцепочечной линкерной ДНК, изображенной на фиг.6. Продукт лигирования расщепляли эндонуклеазами *Hind*III и *Bg*III, и фрагмент из 760 п.о. вводили в плазмиду pLD 41, линеализованную теми же двумя эндонуклеазами до образования pterm 4. Плазмиду pterm 4 идентифицировали по ее рестрикционной карте. Результаты серии рестрикционных расщеплений эндонуклеазой при использовании *Eco*Rv, *Eco*RI, *Kpn*I, *Bg*III-*Hind*III и *Bg*III-*Bcl*I анализировались. Расщепления дают пригодные фрагменты, подтверждающие присутствие синтетического линкера и "полного" 3'-конца гена XPR2 в челночной плазмиде pLD 41, как описано в EP N 0138508. Частичная карта этого 7,3 kb вектора-терминатора показана на фиг.6.

Конструирование плазмиды экспрессии/секреции pLS-3. На фиг. 7 показано конструирование первоначальной плазмиды, использованной для секреции прореннина в *Y.lipolytica*. Ее рестрикционная карта представлена на фиг. 7. Конструирование плазмиды секреции прореннина было инициировано приготовлением фрагмента, содержащего большую часть последовательности структурного гена прореннина. Фрагмент ДНК *Bcl*I-*Bam* HI (частичный) из 1080 п.о.

содержащий кодирующую последовательность для остатков прореннина 6-365, изолировали из плазмиды экспрессии прореннина *E.coli* pPFZ-84A. (Плаزمида pPFZ-84A является производной плазмиды экспрессии прореннина pPFZ-R2, конструкция которой описана в заявке EP N 0147178, опубликованной 3 июля 1985 г. и которая была получена сайт-направленным мутагенезом с использованием синтетических олигонуклеотидов с замещением фрагментов рестрикции. Конкретно, pPFZ-84A отличается от pPFZ-R2 только двумя парами оснований для аминокислотных остатков прореннина 214(Asn → Asp) и 286 (Asp → Gly) т.е. кодирует так называемый прореннин А, однако обе плазмиды содержат необходимую последовательность для прореннина и функционально эквивалентны в данном примере. Фрагмент промотора XPR2, содержащий кодирующую последовательность для предшественника щелочной протеазы 1-157 и прореннина 1-5, получали следующим образом. Фрагмент HindIII-AvaI ДНК из 870 п.о. содержащий участок промотора и 5'-конец гена щелочной протеазы, изолировали из субклона XPR2 плазмиды pLD 90. Этот фрагмент связывали с синтетическим фрагментом, имеющим структуру

5'CCGAGATTCGGCTTCTTCTAATGCCAAGCGAGCTGAGATCACTAG 3'

3'

СТАAGGACGAAGAAGATTACGGTTCGCTCGACTCTAGTGATCCTAG 5'

направление считывания _____→

Эта последовательность содержит липкий конец AvaI, за которым следует последовательность из последних девяти кодонов про-пептида щелочной протеазы, за которой следует последовательность, кодирующая первые четыре аминокислоты прореннина, и оканчивающаяся в сайте Bam HI. Фрагмент промотора получали по стандартной реакции связывания, используя синтетический фрагмент и фрагмент HindIII-AvaI из 870 п.о. и T4-лигазу с последующим расщеплением HindIII и BamHI. Образовавшиеся связанные последовательности очищали электрофорезом на геле полиакриламида с выделением соответствующего фрагмента ДНК HindIII-BamHI из 916 п.о. 3'-конец гибридного гена получали из терминатор/векторной плазмиды pterm 4, описанный выше. Плазмида pterm 4 расщеплялась HindIII и BclI, и, приблизительно 7,3 тыс.п. HindIII-BclI фрагмент ДНК терминатор/вектор, содержащий терминатор XPR2, селективный маркер LEU2 и pBR322, выделяли на геле агарозы.

Плазмиду экспрессии/секреции прореннина pLS-3 собирали путем инкубирования трех компонентов ДНК (HindIII-BclI фрагментов плазмиды pterm 4, промотора HindIII-BamHI из 916 п.о. и фрагментов BamHI-BclI из 1080 п.о.), содержащих ген прореннина, сконструированных, как описано выше в присутствии T4-лигазы (см. фиг.7). Лигирующую смесь использовали для трансформации штамма MM294 *E.coli* K12 по CaCl₂-методу Dagerl и др. Gene 6, 23-28 (1979). Плазмиды выделяли из отобранных

трансформантов, устойчивых к ампициллину, и плазмиду pLS-3 идентифицировали по ее рестрикционной карте (фиг.6,А). Участок XPR2-прореннин этой плазмиды секвенировали для подтверждения наличия нужной последовательности синтетической ДНК и правильного соединения нужных фрагментов.

Получение pLD090. Эта плазмиды содержит субклон из pLD 84. Участок ДНК от сайта PvuI в области промотора XPR2 до сайта EcoRI в области терминатора был субклонирован в сайт HindIII pBR322 по следующей методике. Несколько микрограмм pLD 84 расщепляли двумя вышеупомянутыми ферментами рестрикции. Затем в "липкие" концы расщепленных молекул ДНК вводился фрагмент кленова ДНК-полимеразы I. Затем обработанные киназой HindIII линкоры (CAAGCTTG из New England Biolabs) добавляли к концам с помощью T4-лигазы. Избыток линкеров удаляли и получали липкие концы HindIII последующим расщеплением ферментом HindIII. Смесь молекул ДНК разделяли на препаративном агарозном геле, отделяя нужную 2 тыс.п. полосу, очищали и вводили в реакцию лигирования с HindIII-расщепленным вектором pBR322, обработанным бактериальной щелочной фосфатазой. Лигирующую смесь использовали для трансформации компетентных *E. coli*. Ориентация сайта EcoRI терминатора XPR2, более близкая к сайту EcoRI pBR322, была названа pLD 90, а обратная ориентация pLD 91.

5'конец участка промотора XPR2, который был включен в pLS-3, представляет собой сайт PvuI, приблизительно в 280 п.о. от начала области, последовательность которой указана на фиг.3-5. Было обнаружено, что плазмиды, содержащие ген протеазы дикого типа под контролем только этой части промотора, при интеграции в геном сайте, удаленном от резидентного локуса xpr2, не придают трансформанту способность производить большие количества протеазы (определено по зонам просветления на чашках со снятым молоком).

Мы обнаружили, что если pLS-3 содержит укороченный и поэтому "дефектный" промотор, что интегрант, возникающий при рекомбинации между плазмидой и резидентным геном XPR2 дикого типа, будет образовывать полный промотор, направляющий экспрессию продукта слияния прохимозина, но дефектный по отношению к экспрессии протеазы. Аналогичный эксперимент с нарушением гена был выполнен с актин-геном *S.cerevisiae* Shortle и др. (Science 217:371-373 1982). В соответствии с нашими ожиданиями некоторые лейцин-независимые трансформанты с pLS-3 были действительно дефектны по отношению к протеазе. Трансформанты, дефектны по протеазе, представляли собой скорее нужные интегранты в локусе XPR2, чем нежелательные побочные продукты, такие как генные конверсии в leu2. Со штаммом-реципиентом ATCC 20688 мы обнаружили, что нерасщепленная pLS-3 производила 6,5% дефектных по протеазе трансформантов, тогда как SpaBI-расщепленная плазмиды давала приблизительно 70% трансформантов,

дефектных по протеазе. Аспект нарушения гена при этой трансформации был использован для того, чтобы избежать необходимости в большом количестве Саузерн-блот экспериментов для нахождения правильного интегранта среди всех трансформантов.

Плазмиды, содержащие структурный ген протеазы дикого типа под контролем промотора XPR2 (начинающегося как указано на фиг.3-5) дают возможность экспрессии значительного количества протеазы, если интегрированы в клетки *Y. lipolytica* в сайте, ином, чем локус *xpr2*. Однако для эффективной экспрессии чужеродных генов этим видом интегрантов может потребоваться дальнейшая модификация этого контролирующего участка ДНК.

Секреция прореннина. Штамм ATCC 20688 *Y. lipolytica* был трансформирован нерасщепленной pLS-3 ДНК и SnaBI-расщепленной pLS-3 ДНК для получения *xpr⁺leu⁺* трансформантов ATCC 20775 (DL 114) и ATCC 20776 (DL 148) соответственно. Эти трансформированные штаммы высевались в тест-пробирку, содержащую среду ДЭПД. Клетки выращивались в течение ночи при 28°C. Аликвоту (250 мкл) этих культур разбавляли 1: 100 в 25 мл среды ГПП. Клетки выращивались во встряхиваемой колбе при 28°C в течение 16-18 ч до возникающей абсорбции 5,0-7,0 при 600 нм и отбирались центрифугированием. Образующаяся культуральная жидкость исследовалась на присутствие прореннина путем концентрирования супернатанта и нанесения концентрата на SDS-PAGE. Гель электрофоретически переносился на нитроцеллюлозную бумагу в присутствии 20 мМ трис-основания, 150 мМ глицина, 20% метанола при 500m amp в течение 2 ч при 4 °C. Удаление протеина из геля проверялось окрашиванием с голубым Кумасси.

Нитроцеллюлозную бумагу высушивали при 37°C и помещали на 65°C в течение 1 ч, затем промывали в TBS (200 мМ NaCl, 50 мМ трис-HCl, pH 7,5). Бумагу затем инкубировали при комнатной температуре 30 мин в TBS, содержащем 10% лошадиной сыворотки (Gibco, Chagrin Falls, Ohio) с последующей инкубацией в TBS, содержащем 10% лошадиной сыворотки и антитела прореннина в соответствующем разведении в течение 16 ч при комнатной температуре. Бумагу затем три раза промывали в TBS с последующей инкубацией в TBS, содержащем 10% лошадиной сыворотки, с последующей инкубацией в течение 2 ч в TBS, содержащем 10% лошадиной сыворотки и козьего антитела к антителу кролика IgG, конъюгированного с пероксидазой. Бумагу затем промывали три раза в течение 10 мин в TBS и проявляли в присутствии 4-хлор-1-нафтола (3 мг/мл в метаноле), добавленного к концентрации 0,5 мг/мл в TBS, содержащем 0,01% перекиси водорода. Присутствие прореннина молекулярного веса 40000 было подтверждено в обоих супернатантах.

После кислотной активации концентрированных супернатантов культуры (см. выше) значительная активность по свертыванию молока присутствовала в образцах, приготовленных из трансформированных культур ATCC 20775 и

20776, содержащих pLS-3. Как ожидалось, активность по свертыванию молока не обнаружена в супернатантах контрольной культуры штамма реципиента *Y. lipolytica* ATCC 20688.

5 Конструирование плазмиды экспрессии/секреции pXX-33.

Модификация для превращения pLD-3 в усовершенствованную плазмиду экспрессии pXX-33 изображена на фиг.8. Такая модификация увеличивает участок промотора XPR2 на 280 п.о. Как в случае pLD-3, плаزمида экспрессии pXX-33 содержит гибридный ген, кодирующий весь препро-пептид (157 аминокислотных остатков) щелочной протеазы, присоединенный к полной структурной генной последовательности прореннина.

10 Перед конструированием плазмид экспрессии/секреции прореннина с последовательностью промотора XPR2 на 280 п.о. больше, чем в pLS-3, было необходимо субклонировать фрагмент рестрикции, содержащий полный ген щелочной протеазы, в сайт HindIII. Этот субклон был собран путем добавления синтетических линкеров к фрагменту рестрикции, выделенному из клона XPR2 геномной библиотеки pLD86.

15 Конструирование этого субклона XPR2 с сайтом HindIII, локализованным слева, было иницировано приготовлением фрагмента ДНК, содержащего весь ген щелочной протеазы. 2,3 тыс.п. EcoRI-BamHI фрагмент (частичный) геномной области клона XPR2 pLD-86 был очищен электрофорезом на агарозном геле и связан с синтетическим фрагментом, имеющим последовательность

5' GATCGAAGCTTG 3'
3' TTCGAACCTAA 5'

20 Эта линкерная последовательность содержит липкий конец BamHI (но не восстанавливает сайт BamHI), за которым следует сайт HindIII и затем липкий конец EcoRI. Продукт лигирования расщеплен HindIII и встроен в HindIII сайт pBR322. Плазмида pXHP-24 была идентифицирована по ее рестрикционной карте и стала источником фрагментов промотора XPR2 для будущих конструкций экспрессии.

25 В плазмиде pXHP-24 субклонированный ген XPR2 содержит приблизительно на 280 п. л. больше 5' последовательности промотора XPR2, чем в pLS-3. Первый фрагмент промотора был получен стандартной реакцией лигирования при использовании синтетического фрагмента ДНК (описанного выше для pLS-3) фрагмента HindIII-AvaI с 1150 п.о. pXHP-24 и T4 лигазы с последующим расщеплением HindIII и BamHI. Образующиеся связанные

30 последовательности очищали гель-электрофорезом, выделяя фрагмент HindIII-BamHI приблизительно в 1196 п.о. Второй фрагмент, содержащий последовательности, кодирующие аминокислотные остатки прореннина 6-151 получали из pLS-3 расщеплением BamHI и XmaI и очисткой на геле образующегося фрагмента ДНК BamHI-XmaI в 440 п.о. Третий фрагмент, содержащий остальную часть гена прореннина, терминатор XPR2 и последовательности вектора получали из pLS-3 расщеплением HindIII и XmaI и очисткой на геле фрагмента вектора HindIII-XmaI

приблизительно из 8,0 тыс.п. Эти три фрагмента затем связывали, используя стандартную методику, описанную выше. Реакционную лигирующую использовали для трансформации штамма MM294 E.coli K12. Плазмиды изолировали из трансформантов, отобранных на основе устойчивости к ампициллину, и плазмиду рХХ-33 идентифицировали по ее рестрикционной карте (фиг. 6). Область протеаза-прореннина этой плазмиды секвенировали для подтверждения правильного соединения нужных фрагментов.

Y.lipolytica ATCC 20774 затем трансформировали рХХ-33, расщепленной SnaBI, для получения Y.lipolytica ATCC 20780 и прореннин, секретлируемый трансформированными культурами в культуральный бульон, анализировали, как описано выше в случае рLS-3. Было подтверждено присутствие прореннина в культуральном супернатанте.

После кислотной активации концентрированных культуральных супернатантов (см. выше) значительная активность по свертыванию молока наблюдалась в образцах, приготовленных из трансформированной культуры Y.lipolytica ATCC 20780.

Конструирование плазмиды экспрессии/экспрессии рХХ-22.

Экспериментальные стадии, использованные для конструирования плазмиды экспрессии/секреции рХХ-22, показаны на фиг.7. Вектор экспрессии отличается от рLS-3 в двух отношениях. Подобно рХХ-33, он содержит дополнительный сегмент последовательности промотора XPR2 из 280 п.о. Во-вторых, он содержит последовательность, кодирующую сигнальный пептид щелочной протеазы и только 38 аминокислотных остатков про-пептида (vgo I).

План конструирования рХХ-22 был аналогичен использованному для рХХ-33. Первый фрагмент промотора был получен по стандартной реакции связывания при использовании фрагмента HindIII-BglIII из 890 п.о. рХНР-24 и синтетического фрагмента с последовательностью

5' GATCTTGCTCAGATCACTAG 3'
3' AACGACT3ATGATCCTAG 5'

и T4 лигазы с последующим расщеплением HindIII и BamHI. Образующиеся связанные последовательности очищали гель-электрофорезом, выделяя фрагмент ДНК HindIII-BamHI из 920 п.о. Второй фрагмент, кодирующий остатки прореннина 6-151, изолировали из рLS-3 расщеплением BamHI и XmaI, и очисткой на геле образующегося фрагмента ДНК BamHI-XmaI из 440 п.о. Третий фрагмент, содержащий остальную часть гена прореннина, терминатор XPR2 и векторные последовательности, получали расщеплением рLS-3 HindIII и XmaI и очисткой на геле фрагмента вектора приблизительно в 8,0 тыс.п. Затем эти три фрагмента ДНК связывали, используя стандартную методику, описанную выше. Реакционную лигирующую смесь использовали для трансформации штамма MM294 E.coli K12. Плазмиды изолировали из отобранных трансформантов и плазмиду рХХ-11 идентифицировали по ее рестрикционной карте (фиг.7).

Y. lipolytica ATCC 20774 затем трансформировали рХХ-22, расщепленной SnaBI, для получения Y. lipolytica ATCC 20799, и прореннин, выделяемый в культуральный бульон трансформированными культурами, анализировали, как описано выше в случае рLS-3. Присутствие прореннина в культуральном супернатанте подтверждалось согласно методике, описанной выше. После кислотной активации концентрированных культуральных супернатантов (см. выше), значительная активность по свертыванию молока наблюдалась в образцах, приготовленных их трансформированной культуры Y.lipolytica ATCC 20799.

Конструирование плазмиды экспрессии/секреции рХХ-11.

Экспериментальные стадии конструирования плазмиды экспрессии/секреции прореннина рХХ-11 изображены на фиг.10. Эта плазида содержит последовательность промотора XPR2 и сигнальный пептид из 22 аминокислотных остатков, присоединенный к последовательности, кодирующей прореннин. План конструирования, использованный для рХХ-11, был подобен использованному для рХХ-22 и рХХ-33. Фрагмент промотора получали по стандартной реакции связывания при использовании фрагмента ДНК HindIII-BglIII рХНР-24, приблизительно, 750 п.о. и синтетического фрагмента с последовательностью

5'
TGGCCGCTCCCCTGCGCCGCCCTGCCGCTG
AGATCACTAG 3'

3'
AAGACCGGCGAGGGCACCGGCGGGGACGG
CGACTCTAGTGATCCTAG 5' →

и T4 лигазы с последующим расщеплением HindIII и BamHI. Образующиеся связанные последовательности очищали гель-электрофорезом, выделяя фрагменты ДНК HindIII-BamHI в 790 п.о. Второй фрагмент, кодирующий остатки прореннина 6-151, изолировали из рLS-3 расщеплением BamHI и XmaI и очисткой на геле образующегося фрагмента ДНК BamHI-XmaI в 440 п.л. Третий фрагмент, содержащий остальную часть структурного гена прореннина, терминатор XPR2 и последовательности челночного вектора получали расщеплением рLS-3 HindIII и XmaI очисткой на геле фрагмента вектора приблизительно в 8,0 тыс.п. Затем эти три фрагмента ДНК связывали, используя стандартную методику, описанную выше. Реакционную лигирующую смесь использовали для трансформации штамма MM294 E. coli K12. Плазмиды выделяли из отобранных трансформантов и плазмиду рХХ-11 идентифицировали по ее рестрикционной карте (фиг.10). Часть XPR2 прореннин этой плазмиды секвенировали для подтверждения нужной последовательности синтетической ДНК и правильного соединения нужных фрагментов.

Y.lipolytica ATCC 20774 затем трансформировали рХХ-11, расщепленной SnaBI, для образования Y.lipolytica 20778, и прореннин, выделенный в культуральную среду трансформированными культурами, анализировали, как описано выше в случае

pLS-3. Присутствие прореннина в культуральном супернатанте подтверждалось согласно вышеописанной методике.

Анализы на свертываемость молока (см. выше) показали, что имеется значительная активность по свертыванию молока в культуральном супернатанте трансформантов ATCC 20778, содержащих рХХ-11.

Пример 2. Конструирование "платформы".

Ген В10 дикого типа, соответствующий bio-6 аллелю в ATCC 20774, клонировали путем комплементации следующим образом. Библиотека генов частично Sau 3A-расщепленной хромосомной ДНК *Y.lipolytica*, введенной в BamHI сайт [LD 40 (которая представляет собой рBR322 плюс LEU2 в сайте EcoRI), была сконструирована и приготовлено большое количество библиотечной ДНК как плазмидного препарата смешанной культуры *E.coli* (это та же самая библиотека, что была использована для клонирования гена XPR2). Несколько микрограмм библиотечной ДНК расщепляли ферментом AраI (который расщепляет одну связь в участке LEU2). Затем эту ДНК использовали для трансформации ATCC 20774 (*leu2 xpr2 bio*), и трансформированные смеси вызывались на синтетической среде с недостатком лейцина. Были получены десятки тысяч лейцин-независимых трансформантов. Чтобы обнаружить колонии, если они есть, содержащие библиотечные плазмиды, включающие ген В10, лейцин-независимые трансформанты методом отпечатков высевались на агаровые чашки, содержащие биотин-селективную среду (рецептура на литр: 25 мг дестибиотина, 20 г глюкозы, 5 г сульфата аммония, 1 г KH_2PO_4 , 0,5 г $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,1 г CaCl_2 , 0,1 г NaCl , 500 мкг борной кислоты, 400 мкг тиамин. HCl, 400 мкг $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 400 мкг $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 200 мкг $\text{Na}_2\text{MgO}_4 \cdot 2\text{H}_2\text{O}$, 200 мкг $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, 100 мкг KI и 40 мкг $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$).

Один из нескольких *Y. lipolytica* В10-трансформантов, растущий на биотин-селективной среде, был назван DL 31. Затем приступили к извлечению плазмидной библиотеки генов, содержащей ген В10 штамма DL 31 *Y.lipolytica*. Была приготовлена хромосомная ДНК из культуры штамма DL 31. Несколько микрограммов этой хромосомной ДНК расщепляли ферментом рестрикции AраI для того, чтобы вырезать плазмиду библиотеки. Аликвоту расщепленной ДНК использовали в реакции связывания для циркуляризации неизвестной плазмиды библиотеки. Продукт связывания затем использовали для трансформации культуры *E.coli* к устойчивости к ампициллину путем введения неизвестной В10-содержащей плазмиды в *E. coli*. Было получено несколько трансформантов *E.coli*, устойчивых к ампициллину. Из трансформантов *E.coli* были получены препараты плазмид. Рестрикционные перевары таким образом полученной плазмидной ДНК показывают, что неизвестная В10-содержащая плаزمид эквивалентна рLD 40 со встройкой в сайт BamHI. Эта плазмид должна была первоначально происходить из нашей библиотеки генов и была названа рLD 51.

Плазмид рLD 56 была образована как субклон рLD 51 путем удаления гена LEU2 из

рLD 51, следующим образом. Аликвоту плазмиды рLD 51 расщепляли ферментом EcoRI для удаления участка LEU2. Расщепленную ДНК использовали в реакции связывания ДНК для рециркуляризации плазмиды. Затем была выполнена трансформация *E.coli* для клонирования меньшей В10-содержащей плазмиды. Один из ампициллин-устойчивых трансформантов *E.coli*, как показано, содержал ожидаемую меньшую плазмиду, которая была названа рLD56. Было выполнено несколько рестрикционных расщеплений рLD 56. В10-содержащий сегмент рLD56, который обнаруживается как вставка в BamHI сайт рBR322, имеет длину приблизительно 3,6 тыс.п.

Очень грубая рестрикционная карта вставки ДНК *Y.lipolytica* из 3,6 тыс.п. в BamHI сайт рBR322 (включая рLD 56) описана ниже с указанием в скобках примерного расстояния в парах оснований от начала вставки. Размеры были оценены на нескольких агарозных гелях и, возможно, содержат относительно большие количественные погрешности: PvuII (800), PvuII (1200), PstI (1800), MluI (2000), PstI (2300), EcoRV (2700), NcoI (3200) (для ориентации: SalI сайт рBR 322 будет предшествовать описанным сайтам и сайт HindIII будет следовать за ними).

Штамм ATCC 20774 (MATB *leu* 2-40 bio-6 xpr 2-1002) трансформировали интактной рLD 56 (рBR322 плюс приблизительно 3,6 тыс.п. хромосомной ДНК *Y.lipolytica*, содержащей ген В10). Три различных биотин-независимых трансформанта испытывали на высокочастотную трансформацию NruI-расщепленной (мишень для рKR322) рLD 40 (LEU2 на рBR322) относительно родительского штамма для определения, который из них содержал резидентную рBR322, интегрированную в участок В10. Все три показали способность к высокочастотной трансформации ввиду интеграции рLD40 в резидентную рBR322. Это было подтверждено Саузерн-блот анализом по гибридизации. Один из трех исходных В10-трансформантов *Y.lipolytica* был назван DL118 и использован далее как реципиент ДНК. Вышеупомянутая рестрикционная карта требовалась для установления: а) что использовать в качестве В10-специфичной пробы гибридизации (область NcoI-PvuII), б) какой фермент требуется, чтобы правильно вырезать плазмиду рLD 56 (Ml и I), в) какой фермент расщепляет только одну связь в области рBR322 (ClaI) и г) какой фермент не расщепляет плазмиду вообще (AраI). Гибридизация по Саузерну ClaI и AраI переваров ДНК из ATCC 20774 и DL 118 (испытанная с В10-фрагментом) показала, что, как ожидалось, биотин-участок DL118) при сравнении с В10-участком ATCC 20774) был разрушен добавлением ДНК размера приблизительно рLD 56. MluI перевары ДНК DL 118 (проба с рBR322) показали далее, что добавка была того же размера, как интактная рLD 56.

Конструирование плазмиды экспрессии/секреции рLX-34.

Была сконструирована плазмид экспрессии, которая содержит последовательность, кодирующую прореннин, с сигналами секреции XPR2

(препропоследовательности из 157 аминокислот) ниже от промотора LEU2. Эта плаزمид экспрессии демонстрирует, что может быть использован иной, чем XPR2 промотор для достижения секреции чужеродных протеинов. Кроме того, этот вектор экспрессии способен вызвать экспрессию/секрецию независимо от сайта интеграции в геноме *Y.lipolytica*. Успешное выделение прореннина с промотором, иным, чем промотор XPR, демонстрирует возможность конструкции вектора экспрессии для идентификации альтернативных новых сильных промоторов в *Y.lipolytica*. В дополнение к этому данный подход может быть использован для получения культуры экспрессии с двумя отдельными гибридными генами прореннина, один экспрессируемый промотором LEU2, а другой промотором XPR2, интегрированным в различных сайтах генома хозяина.

Экспериментальные стадии, использованные для конструирования вектора экспрессии, который содержит ген прореннина с сигналами секреции щелочной протеазы (предпоследовательность XPR2 157 аминокислот), экспрессируемые последовательностями промотора LEU2, изображены на фиг.17. Конструирование этой плазмиды было инициировано приготовлением фрагмента промотора LEU2, содержащего около 300 пар оснований 5'-нетранслированной последовательности, предшествующей кодону инициации трансляции ATG гена бета-изопротилмалатдегидрогеназы (фиг.14-16). Фрагмент ДНК HindIII-FokI из 300 п.о. содержащий участок из 270 п.о. последовательности промотора LEU2, был выделен из челночного вектора pLD 40. Этот фрагмент был связан с синтетическим линкером из 54 п.о. имеющим последовательность

```
-----Leu2-----
Fok I
5'-ATACAACCACFCFCFTCCACAATG
3'-TTGGTGTGTGTAGGTGTTAC
-----Xpr-----
AAGCTCGCTACCGCCTTTACTATTCTCAC
TGCCGTTTC-3'
TTCGAGCGATGGCGGAAATGATAAGAGT
GACGGC-5'
```

T4 лигазой с последующим расщеплением HindIII. Образующиеся линкерные последовательности очищали гель-электрофорезом, выделяя фрагмент ДНК HindIII-BglI приблизительно из 600 п.о. Второй фрагмент, кодирующий остальную часть препропоследовательности XPR2 и первые 152 аминокислоты прореннина, изолировали из плазмиды экспрессии pXX-33 (фиг.6) расщеплением BglI и XmaI и очисткой на геле образующегося фрагмента ДНК из 887 п.о. Третий фрагмент, содержащий остальную часть гена прореннина, терминатор XPR2 и последовательности вектора, готовили расщеплением pXX-33 HindIII и XmaI и очисткой на геле, выделяя приблизительно 8,0 тыс.п. фрагмента вектора. Три фрагмента ДНК связывались при использовании стандартной методики, описанной выше. Лигирующая смесь использовалась для трансформации штамма HB101 *E.coli* K12. Плазмиды изолировали из трансформантов, отобранных на основе устойчивости к

ампициллину, и плазмиду pL X-34 идентифицировали по ее рестрикционной карте (фиг.13).

Y. lipolytica ATCC 20794 (DL 118) трансформировали NruI-расщепленной ДНК pL X-34 для получения *Y.lipolytica* ATCC 20795 (DL 251) и прореннин, выделенный в культуральную жидкость лейцин-независимыми трансформантами культуры, анализировали, как описано выше в случае pLS-3. Данная методика трансформации обеспечивала интеграцию pLX-34 в последовательность pBR322, предварительно введенную в bio локус в хромосоме хозяина (описано выше). Интеграция pL X-34 в этот сайт была подтверждена Southern анализом.

Использование DL 118 в качестве реципиента

Гибридизация по Саузерну выполнялась следующим образом: NruI перевары ДНК из трансформантов DL 118 (гибридирующихся с прохимозинового пробой, например, когда входящая плазмида представляла плазмиду экспрессии прохимозина) точно вырезали вошедшую плазмиду. Несколько наногрмм NruI расщепленной трансформирующей плазмиды использовались для проверки правильности размера гибридирующейся полосы. Также NruI перевары (MluI не выщепляет трансформирующую плазмиду) ДНК из этих трансформантов (анализированных с 32p-помеченной pBR322), что резидентная pBP322 последовательность DL 118 была разрушена при добавлении одной или более молекул трансформирующей плазмиды. Это показывает, что интеграция произошла в нужном сайте.

Трансформированную культуру *Y.lipolytica* ATCC 20795 (DL 251) выращивали в среде ДЭПД при 22°C, чтобы благоприятствовать экспрессии промотора LEU2. Присутствие прореннина в культуральных супернатантах было подтверждено анализом на свертываемость молока (см.выше) культуральных супернатантов, активированных кислотой, и проверялось иммуноблоттингом (см. выше). Эти результаты показывают, что гибридный ген представляет собой независимую единицу экспрессии, способную к экспрессии/секреции при интеграции в сайт, иной чем XPR2 или LEU2. Это позволяет сконструировать культуру экспрессии с множественными гибридными генами, потенциально способными увеличивать уровень внеклеточного прореннина.

Пример 3. Последовательность синтетического гена для человеческого C5a.

План, разработанный для достижения бактериальной продукции человеческого анафилатоксина C5a, был аналогичен предыдущим методам, использованным для синтеза и экспрессии EGF, как описано в EP заявке N 0147178. В нем применялось конструирование гена, в котором кодирующую последовательность для активированного дополнительного компонента C5a, получали синтетически. Зная аминокислотную последовательность человеческого C5a, сконструировали фрагмент ДНК, кодирующий 74 аминокислоты (фиг.9). Синтетическая генная последовательность была выбрана, чтобы максимизировать утилизацию

предпочтительного кодона E.coli и S.cerevisiae и обеспечить несколько сайтов рестрикции эндонуклеазой для облегчения характеристики. Данный подход обеспечивает прямую экспрессию анафилактоксина в E.coli путем введения кодона инициации ATC для синтеза протеина перед триплетом, кодирующим первую аминокислоту полипептида C5a. Для облегчения его введения в нужной ориентации в плазмиду pBR322, был сконструирован синтетический ген C5a, содержащий сайты EcoRI и HindIII на своих концах. Для получения конечной генной последовательности C5a, были синтезированы фосфорамидатным методом десять 47-меров и соединены в двухцепочечный фрагмент ДНК из 235 п.л. Фрагмент гена C5a вводили с соответствующим образом расщепленную pBR322 и клонированный ген идентифицировали путем рестрикционного анализа плазмидной ДНК из произвольно выбранных трансформантов. Несколько клонов C5a затем анализировали секвенированием ДНК для идентификации клона с правильной последовательностью. Нужная нуклеотидная последовательность для участка гена C5a была найдена в 2 из 5 исследованных клонов.

Бактериальная экспрессия человеческого C5a. Конструирование плазмиды экспрессии C5a было инициировано расщеплением субклона C5a эндонуклеазой рестрикции EcoRI с последующим дефосфорилированием путем обработки бактериальной щелочной фосфатазой. Используя фрагмент ДНК EcoRI из pPFZ-R2 в 360 п.о. содержащий trp промотор-оператор и последовательности сайта связывания с рибосомой, была сконструирована плаزمида экспрессии C5a. Компонентные клетки штамма HB101 E.coli трансформировали лигирующей смесью. Несколько устойчивых к лекарствам колоний от каждой трансформации очищали, и их плазмидные ДЕК подвергали рестрикционному анализу для идентификации ДНК с trp промотором в ориентации, которая будет проявляться в транскрипции гена C5a. Множественные изоляты из этой реакции связывания были идентифицированы с помощью плазмид, содержащих ген анафилактоксина, примыкающий к последовательности бактериального промотора в конфигурации, требуемой для прямой экспрессии C5a. Рестрикционная карта плазмиды экспрессии C5a p C5a-48 представлена на фиг.12. Экспрессия и выделение человеческого анафилактоксина в Y.lipolytica.

Вектор экспрессии/секреции pC5aX-3, обеспечивающий секрецию человеческого анафилактоксина C5a, был приготовлен по методике, изложенной в примере 1 для pXX-33. Y.lipolytica ATCC 20774 затем трансформировали этим вектором секреции, и человеческий C5a, продуцируемый трансформированными культурами, анализировали, как описано выше, за исключением того, что в иммуноблоттинге были использованы козы антитела к C5a и кроличьи антитела к козьему IgG. Для плазмиды, описанной в данном примере, было подтверждено наличие C5a в культуральном супернатанте.

Конструирование плазмиды

экспрессии/секреции pC5aX-3.

Конструирование плазмиды экспрессии/секреции анафилактоксина pC5aX-3 показаны на фиг. 11. Эта плазмида содержит последовательность для "полного" промотора XPR2 и сигнала из 157 аминокислотных остатка и про-пептид, присоединенный к синтетической последовательности, кодирующей 74 аминокислотных остатка C5a. План конструирования, использованный для pC5aX-3, был подобен использованному для pXX-33. Сначала плазмиду pXHP-24 (или другую плазмиду, содержащую нужную последовательность) расщепляли HindIII и Aval и фрагмент из 1150 п. о. содержащий промотор XPR2, очищали на геле. Второй фрагмент, содержащий 3' конец про-пептида XPR2 и последовательность структурного гена C5a, получали по стандартной реакции связывания при использовании фрагмента ДНК HinfI-HindIII с приблизительно 220 п.о. из плазмиды экспрессии E.coli pC5a-48 и синтетического фрагмента с последовательностью

5'
CCGAGATTCCTGCTTCTTCTAATGCCAAGCG
A 3'
3'
СТАAGGACGAAGAAGATTACGGTTCGCTTGA
5'

и T4 лигазы с последующим расщеплением Aval и HindIII. Образующиеся связанные последовательности очищали гель-электрофорезом, отбирая фрагмент Aval-HindIII из около 250 п.о. фрагмент HindIII-Aval, содержащий промотор, и фрагмент Aval-HindIII, кодирующий C5a, затем связывают T4 лигазой с последующим расщеплением HindIII. Фрагмент приблизительно в 1,4 тыс.п. очищали на геле и использовали для связывания с HindIII расщепленным pterm 4 (описанным выше). Продукт связывания использовали для трансформации штамма MM294 E.coli K12. Плазмиды, отобранные по устойчивости к ампициллину, изолировали из отобранных трансформантов, и плазмиду pC5aX-3 идентифицировали по ее рестрикционной карте. Штамм Y.lipolytica PC-30869, ATCC 20774 затем трансформировали SnaBI расщепленной pC5aX-3 и анафилактоксин, выделенный в культуральную среду, трансформированными культурами, анализировали, как описано выше. Присутствие C5a в культуральном супернатанте было подтверждено по вышеописанной методике.

Известно, что многие протеины, которые синтезируются рибосомами, связанными с эндоплазмической сетью, производятся в виде гликопротеинов. Действительно, гликозилирование может оказывать влияние на секрецию данного протеина. N-связанное гликозилирование эукариотических протеинов происходит в трипептидных последовательностях аспарагин-X-треонин и аспарагин-X-серин, где X может быть любой аминокислотой, кроме, возможно, аспарата (Hubbard S. и др. 1981, Ann. Rev. Biochem. 50, 555). Аминокислотная последовательность проренина включает два таких трипептида, однако гель-электрофоретический анализ проренина, выделяемого в культурах

Y.lipolytica, не показывает наличия гликозилирования. В других выделяемых эукариотических протеинах не все аспарагин-Х-треонин/серин сайты гликозилированы. Видимо, определенные аспарагины в трипептидных последовательностях не гликозилируются, так как они недоступны ферментам гликозилирования.

В случае человеческого С5а аминокислотная последовательность включает отдельный сайт гликозилирования или трипептидную последовательность (Asp-Ile-Ser), которая обычно содержит сложный олигосахарид, присоединенный к аспарагину (Fernandez Н. и др. 1976, *J.Immunol*, 117, 1668). Часть молекул С5а, выделяемых в культуральную среду *Y.lipolytica*, как представляется, гликозилирована, так как широкая область антигенной активности наблюдается в высокомолекулярной части при иммуноблоттинге. Данная гетерогенная электрофоретическая мобильность аналогична наблюдаемой для других секретируемых протеинов и, вероятно, является следствием различий в степени добавления углеводов. По настоящему изобретению очевидное гликозилирование определенных секретируемых чужеродных протеинов предполагает, что экспрессия и секреция *Y.lipolytica* будут полезны для получения многих обычно гликозилированных эукариотических протеинов.

Формула изобретения:

1. Способ получения прореннина, включающий культивирование штамма дрожжей, трансформированного рекомбинантной плазмидной ДНК, кодирующей протеин, и выделение протеина, отличающийся тем, что культивируют *J.lipolytica*, трансформированный рекомбинантной плазмидной ДНК рL5/3, или рХХ-11, или рХХ-22, или рХХ-33, или рLХ-34.

2. Способ по п. 1, отличающийся тем, что культивируют штамм дрожжей *J.lipolytica* ATCC 20775 и 20776, или 20778, или 20779, или 20780, или 20795.

3. Способ получения штамма дрожжей-продуцента прореннина, предусматривающий трансформацию штамма дрожжей рекомбинантной плазмидой ДНК, кодирующей протеин, и отбор трансформантов, отличающийся тем, что в качестве рекомбинантной плазмидной ДНК используют рL5-3 или рХХ-33, или рХХ-22, или рХХ-11, или рLХ-34, а трансформируют штамм дрожжей *J.lipolytica* ATCC-20688, или 20774, или 20794.

4. Способ получения рекомбинантной плазмидной ДНК рХХ-33, заключающийся в том, что проводят расщепление плазмиды рХНР-24 рестриктазами Hind III и Ava I до образования фрагмента размером 1150 п.о. изолируют указанный фрагмент, связывают его с синтетическим фрагментом ДНК со следующей нуклеотидной

последовательностью: 5'CCG AGA TTCCTGCTT CTT СТААТGCC AAGCGAGCTGA 3' CT AAGGACGAA GAA GATTACGGTTCGCTCGACT TCACTAG 3' AGTGATC 5', в присутствии лигазы полученную рекомбинантную молекулу расщепляют рестриктазами Hind III и Bam HI до образования фрагмента размером 1196 п.о. фрагмент выделяют, связывают его с фрагментом размером 440 п.о. полученным в результате расщепления плазмиды рL53 рестриктазами Bam HI и Xma I и Hind-III-Xma I фрагментом плазмиды рL53 размером 8,0 т.п. в присутствии T₄ лигазы, далее трансформируют штамм бактерий *E.coli*, полученный рекомбинантной ДНК, отбирают устойчивые к ампициллину трансформанты и выделяют плазмиду рХХ-33.

5. Способ получения рекомбинантной плазмидой ДНК рХХ-22, заключающийся в том, что расщепляют плазмиду рХНР24 рестриктазами Hind III и Bgl II до образования фрагмента размером 890 п.о. изолируют указанный фрагмент, связывают его фрагмент с синтетическим фрагментом ДНК со следующей нуклеотидной последовательностью:

5'GATCTTGCTGAGATCACTAC AACGACTCTAGTGATCCTAG 5' в присутствии T₄ лигазы, расщепляют полученную рекомбинантную ДНК рестриктазами Hind III и Bam HI до образования фрагмента размером 920 п.о. фрагмент отделяют, связывают его с фрагментом размером 440 п.о. полученным в результате расщепления плазмиды рL53 рестриктазами Bam HI и Xma I и фрагментом размером 8,0 т.п. в присутствии T₄ лигазы, далее трансформируют штамм *E.coli* полученной рекомбинантной ДНК, отбирают устойчивые к ампициллину трансформанты, из которых выделяют плазмиду рХХ-22.

6.Способ получения рекомбинантной плазмидной ДНК рХХ-11, заключающийся в том, что расщепляют плазмиду рХНР24 рестриктазами Hind III-BnlI до образования фрагмента размером 750 п.о. отделяют указанный фрагмент, связывают его с синтетическим фрагментом ДНК со следующей нуклеотидной

последовательностью: 5' TGGCCGCTCCCCTGGCGCCCCTGCCGCTGA G 3' AAG ACCGGCGAGGGGACCGGCGGGTTCGGCGAC TC ATCACTAG 3' TAGTGATCCTAG 5' в присутствии T₄ лигазы, расщепляют полученную рекомбинантную ДНК рестриктазой Hind III и Bam HI до образования фрагментов размером 790 п.о. связывают указанный фрагмент с фрагментом размером 440 п.о. полученным в результате расщепления плазмиды рL53 рестриктазами Bam HI и Xma I размером 8 т.п. в присутствии T₄ лигазы, далее транспортируют штамм *E.coli* полученной рекомбинантной ДНК, отбирают устойчивые к ампициллину трансформанты и изолируют плазмиду рХХ-11 из указанных трансформантов.

Таблица 1

Сравнение штаммов *Yarrowia lipolytica*

Регистрационный номер Pfizer	Источник	Генотип
PC-30265	NRRL YB-423 (также CBS 6124), типовая культура в <i>The Yeasts</i> , 3-е издание	Дикий тип диплоид
PC-30286	CBS 599, типовая культура в <i>The Yeasts</i> , 2-е издание	MATA дикий тип гаплоид
PC-30869	См. ниже	MATB bio-6 leu 2-40 xpr-1002

Таблица 2

Ассимиляция углерода ^{/a/}

Источник	Ссылка ^{/b/} в перечне описаний	Культура		
		30265	30286	30869
1. L-арабиноза	-	-	-	-
2. Целлобиоза	-	-	-	-
3. Эритрит	+	+++	+++	+++
4. D-галактоза	-	-	-	-
5. D-глюкоза	+	+++	+++	+++
6. Инозит	-	-	-	-
7. Лактоза	-	-	-	-
8. Мальтоза	-	-	-	-
9. D-маннит	+	+++	+++	+++
10. Рафиноза	-	-	-	-
11. Рибитол	-	-	-	-
12. D-рибоза	-/+	-	-	++
13. L-рамноза	-	-	-	-
14. Растворимый крахмал	-	-	-	-
15. Сакроза	-	-	-	-
16. Трихалоза	-	-	-	-
17. D-ксилоза	-	-	-	-
18. Янтарная кислота	+	+++	+++	+++
19. Лимонная кислота	+	+++	+++	+++

^{/a/} Основной средой была Бактодрожжевая азотная основа с добавками 10 мкг/л D-биотина и 149 мг/л L-лейцин-этилового эфира. HCl для образования 100 мг/л L-лейцина.

^{/b/} Kreger-van Rij (ук. соч.).

RU 2069694 C1

RU 2069694 C1

Таблица 3

Ассимиляция азота ^{/a/}

Источник	Ссылка ^{/b/} в перечне описа- ний	Культура		
		30265	30286	30869
1. /NH ₄ / ₂ SO ₄	+	+++	+++	+++
2. KNO ₃	-	-	-	-
3. Мочевина	+	+++	+++	+++

^{/a/} Основной средой была Бактодрожжевая углеродная основа с добавками 116 мг/л кетонизокапроата натрия, чтобы обеспечить эквивалент 100 мг/л L-лейцина, и 10 мкг/л D-биотина.

^{/b/} Kreger-van RIJ (местное цитирование).

Таблица 4

Требование витаминов ^{/a/}

Добавка к основной среде ^{/c/}	Заключение ^{/b/} источника	Культура		
		30265	30286	30869
1. Отсутствует	-	tr	tr	-
2. Тиамин. HCl	+	+++	+++	-
3. D-биотин	-	tr	tr	tr
4. Тиамин+биотин	+	+++	+++	+++

^{/a/} Основной средой была Бактодрожжевая основа, свободная от витаминов плюс 149 мг/л L-лейцин-этилового эфира. HCl для образования 100 мг/л L-лейцина.

^{/b/} Kreger-van RIJ (местное цитирование).

^{/c/} 200 мкг/л тиамина. HCl и/или 10 мкг/л D-биотина, как указано.

Таблица 5

Аскоспоруляция

Тестовый штамм	Спаренная культура		
	30265	30286	30869
Отсутствует (сама спариваемая культура)	++	-	-
30264 (спариваемый тип А)	++	-	+++
30267 (спариваемый тип В)	++	+++	-

^{/a/} Культуры 30264 и 30267, гаплоидные штаммы противоположного спариваемого типа, любезно предоставлены д-ром L.J. Wickerham. Они формально описаны в Science **167**, 1141 (1970).

^{/b/} 30264 представляет собой Wickerham's C. Ilpolytica YB-421.

^{/c/} 30267 представляет собой Wickerham's C. Ilpolytica YB-423-12.

Другие характеристики

	Значение в /a/ источнике	Культура		
		30265	30286	30869
Форма клеток Средний размер клеток (микроны) Вегетативное размножение Ферментация Рост при 37 °С Рост колоний	яйцевидная (2-4,5) x (4-22) почкованием отсутствует нет	яйцевидная 3,3 x 9,1 почкованием отсутствует нет	яйцевидная 3,0 x 8,2 почкованием отсутствует нет	яйцевидная 2,6 x 5,5 почкованием отсутствует нет
Три культуры росли схожим образом в согласии с литературным описанием. Доминируют псевдо- и истинные мицелии. Присутствуют бластоспоры, преимущественно в виде одиночек. Каротеноидных пигментов не обнаруживается.				

/a/ Kreger-van Rij (местное цитирование).

Таблица 7

Требования к питанию

Культура	Питательное вещество, не включаемое в описанную среду			
	Никакое	Лейцин	Урацил	Биотин
CBS 599	+++	+++	+++	+++
CBS 6124	+++	+++	+++	+++
ATCC 20228	+++	+++	+++	+++
ATCC 20628	+++	-	+++	+++
ATCC 20688	+++	-	-	+++
PC-30869	+++	-	+++	-

Таблица 8

Разжижение желатина

Культура	Разжижение
CBS 599	+
CBS 6124	+
ATCC 20228	+
ATCC 20628	+
ATCC 20688	+
PC-30869	-

I последовательность белка val thr gln trp gly

mRNA 5'		GTX ACX CAU TGG GG	3'
probes 3'		CAX TGX GTY ACC CC	5'
170	3'	CAA TGX GTY ACC CC	5'
172	3'	CAT TGX GTY ACC CC	5'
174	3'	CAG TGX GTY ACC CC	5'
176	3'	CAC TGX GTY ACC CC	5'

II последовательность белка lys lys ala gln thr

mRNA 5'		AAU AAU GCX CAU AC	3'
probes 3'		TTY TTY CGX GTY TG	5'
180	3'	TTC TTY CGX GTC TG	5'
182	3'	TTT TTY CGX GTC TG	5'
184	3'	TTT TTY CGX GTT TG	5'
186	3'	TTC TTY CGX GTT TG	5'

Phi2.2

RU 2069694 C1

RU 2069694 C1

GATCCGGGGTTCTATGCATCCCTGAAACATTGATTGAAAATTAACATATGAGCTGCCGTG
 CTTTTTGCATTCAAGGGCCGACCTTATCTTGTATCCTTAATTACACATGACCTCTTGAG
 CGCCACGGTACATTCCCTGGCGTCAGTTCCGGTGGAGCGGACACTTTTCTCTCCTTTGTCT
 GACATGTTGGTTAAGTTGTAGTCCAGGGACACAAGGGGTTCCAACGGCAGTGGCAGCCT
 ACCCCACGCTACCCACCCTGGCCCTGGTCTAACTTCGACGATCGGCATCAGGGTTCAT
 GGATAGGGCGGTGTGATTTACGATGTGATGGACAATGTTAGAGAGATCCCCTACTTGTA
 GTCAGGCCATCTTTTACGTACGCACTGTACCATGATGTCAATGGAGTATGATGAACCGA
 CTTTGAGAGACTCACATCTGCACAACACCATGTTTCAGCGGAATCCGACTTCCAACCCA
 AACCCAAGCCCCTGTGAGATATCGTGAGAAGGCACGGCACCAACTAATGCACACACTCC
 ACCTGTATTGCACCAAGATAATGAGGGCATCGTCTTGGCGCGTCTTGGCGAGAGCCGCTG
 TTTTCGTGACGCAATCAGAGCAGTTTCTGGATAGTATCTTGTCCAGAAACACGATATAAA
 CCCCATCGACGGGGCCCGTTGAAGAGCACCAACCCACTATCCAATCCTCCAATCCAACA

-157 -150
 met lys leu ala thr ala phe thr ile leu thr ala val leu ala
 ATG AAG CTC GCT ACC GCC TTT ACT ATT CTC ACT GCC GTT CTG GCC

140 -130
 ala pro leu ala ala pro ala pro ala pro asp ala ala pro ala
 GCT CCC CTG GCC GCC CCT GCC CCT GCT CCT GAT GCT GCC CCT GCT

-120
 ala val pro glu gly pro ala ala ala ala tyr ser ser ile leu
 GCT GTG CCT GAG GGC CCT GCC GCC GCT GCC TAC TCA TCT ATT CTG

-110 -100
 ser val val ala lys gln ser lys lys phe lys his his lys arg
 TCC GTG GTC GCT AAG CAG TCC AAG AAG TTT AAG CAC CAC AAG CGA

-90
 asp leu asp glu lys asp gln phe ile val val phe asp ser ser
 GAT CTT GAT GAG AAG GAT CAG TTC ATC GTT GTC TTT GAC AGT AGG

-80 -70
 ala thr val asp gln ile ala ser glu ile gln lys leu asp ser
 GCT ACT GTT GAC CAG ATC GCC TCC GAA ATC CAG AAG CTG GAC TCT

-60
 leu val asp glu asp ser ser asn gly ile thr ser ala leu asp
 CTG GTC GAC GAG GAC TCG TCC AAC GGT ATC ACC TCT GCT CTT GAT

-50 -40
 leu pro val tyr thr asp gly ser gly phe leu gly phe val gly
 CTT CCT GTC TAC ACG GAT GGA TCT GGC TTT CTC GGA TTT GTT GGA

φu2.3

RU 2069694 C1

RU 2069694 C1

RU 2069694 C1

RU 2069694 C1

-30

lys phe asn ser thr ile val asp lys leu lys glu ser ser val
AAG TTC AAC TCC ACT ATC GTT GAC AAG CTC AAG GAG TCG TCT GTT

-20

leu thr val glu pro asp thr ile val ser leu pro glu ile pro
GTG ACG GTC GAG CCC GAT ACC ATT GTG TCT CTC CCC GAG ATT CCT

-10

1

ala ser ser asn ala lys arg ala ile gln thr thr pro val thr
GCT TCT TCT AAT GCC AAG CGA GCT ATC CAG ACT ACT CCC GTC ACT

10

gln trp gly leu ser arg ile ser his lys lys ala gln thr gly
CAA TGG GGC CTG TCT AGA ATC TCT CAT AAG AAG GCC CAG ACT GGA

20

30

asn tyr ala tyr val arg glu thr val gly lys his pro thr val
AAC TAC GCC TAC GTT CGA GAG ACA GTT GGC AAG CAC CCC ACC GTT

40

ser tyr val val asp ser gly ile arg thr thr his ser glu phe
TCT TAC GTT GTT GAC TCT GGT ATC CGA ACC ACC CAC TCC GAG TTC

50

60

gly gly arg ala val trp gly ala asn phe ala asp thr gln asn
GCA GGC CGA GCT GTC TGG GGA GCC AAC TTC GCT GAC ACA CAG AAC

70

ala asp leu leu gly his gly thr his val ala gly thr val gly
GCT GAT CTT CTC GGT CAC GCC ACT CAC GTT GCA GGT ACC GTG GGA

80

90

gly lys thr tyr gly val asp ala asn thr lys leu val ala val
GGA AAG ACA TAC GGA GTC GAC GCC AAC ACC AAG CTG GTG GCC GTC

100

lys val phe ala gly arg ser ala ala leu ser val ile asn gln
AAG GTG TTT GCA GGC CGA TCC GCA GCT CTC TCC GTC ATC AAC CAG

110

120

gly phe thr trp ala leu asn asp tyr ile ser lys arg asp thr
GGC TTC ACC TGG GCT CTC AAC GAC TAC ATC TCC AAG CGA GAC ACT

130

leu pro arg gly val leu asn phe ser gly gly gly pro lys ser
CTG CCT CGA GGA GTG CTG AAC TTC TCT GGA GGA GGA CCC AAG TCC

140

150

ala ser gln asp ala leu trp ser arg ala thr gln glu gly leu
GCT TCC CAG GAC GCC CTA TGG TCT CGA GCT ACC CAG GAG GGT CTG

160

leu val ala ile ala ala gly asn asp ala val asp ala cys asn
CTT GTC GCC ATC GCT GCG GGA AAC GAT GCC GTG GAC CCC TGT AAC

170

Fig. 4

180
 asp ser pro gly asn ile gly gly ser thr ser gly ile ile thr
 GAC TCT CCC GGT AAC ATT GGA GGC TCC ACC TCT GGT ATC ATC ACT

190
 val gly ser ile asp ser ser asp lys ile ser val trp ser gly
 GTG GGT TCC ATT GAC TCT AGC GAT AAG ATC TCC GTC TGG TCC GGT

200
 gly gln gly ser asn tyr gly thr cys val asp val phe ala pro
 GGA CAG GGA TCC AAC TAC GGA ACT TGT GTT GAT GTC TTT GCC CCC

210
 gly ser asp ile ile ser ala ser tyr gln ser asp ser gly thr
 GGC TCC GAT ATC ATC TCT GCC TCT TAC CAG TCC GAC TCT GGT ACT

220
 leu val tyr ser gly thr ser met ala cys pro his val ala gly
 TTG GTC TAC TCC GGT ACC TCC ATG GCC TGT CCC CAC GTT GCC GGT

230
 leu ala ser tyr tyr leu ser ile asn asp glu val leu thr pro
 CTT GCC TCC TAC TAC CTG TCC ATC AAT GAC GAG GTT CTC ACC CCT

240
 ala gln val glu ala leu ile thr glu ser asp thr gly val leu
 GCC CAG GTC GAG GCT CTT ATT ACT GAG TCC AAC ACC GGT GTT CTT

250
 pro thr thr asn leu lys gly ser pro asn ala val ala tyr asn
 CCC ACC ACC AAC CTC AAG GGC TCT CCC AAC GCT GTT GCC TAC AAC

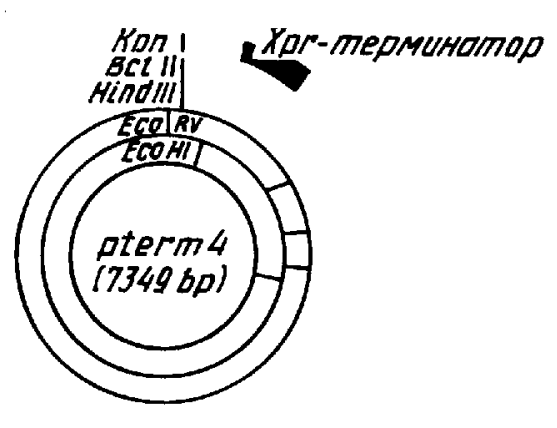
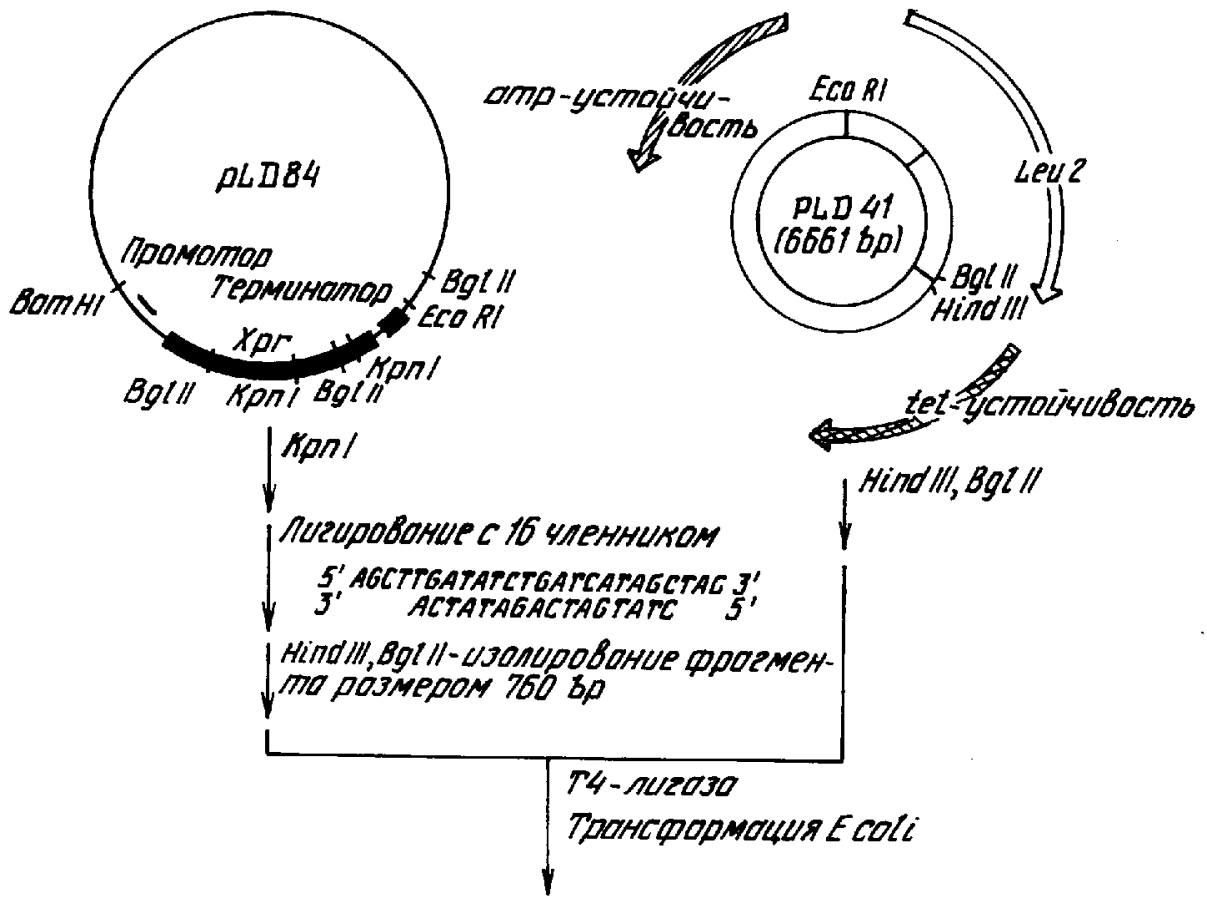
260
 gly val gly ile
 GGT GTT GGC ATT TAG GCAATTAACAGATAGTTTGCCGGTGATAATTCTCTTAACCTC

270
 CCACACTCCTTTGACATAACGATTTATGTAACGAAACTGAAATTTGACCAGATATTGTTG
 TAAATAGAAAATCTGGCTTGTAGGTGGCAAAATCCCGTCTTTGTTTCGTTCGGTTCCTCTG
 TGA CTGCTCGTCCCTTTGTGTTCCGACTGTCGTGTTTTGTTTTCCGTGGGTCGCAAG
 TGAGATGCCCGTGTTCGAATTCGGTAGTCCGACGGACCATCGGTTGCTCTGCAACACAC
 ACACCGGAGGCTGCAACCTACATCAGAGCACTACTTGCAGGGTTGATGCAACATTCACA
 AAAGCGCAAGCAGTGGGTGATGTATAGCAGCTAACAGCAACTACTGCTCAACATGAAAAA
 GGAGGGTGTAAAGACGGCCAAGACTGCTTTCTGTCTACGCCCTGAGCAACGTGCTCTGCAA
 CAGAGCAACAGATAATCGCCTACGGAGACAGAGACAGAGACAGAAACAGAAACAAAAGCA
 ACAGAAACTGCTGTAGTGTGTTGCAGTGAGGGCGAGATTTAACCGTATAATTCACGCTCA
 GATCT

7025

RU 2069694 C1

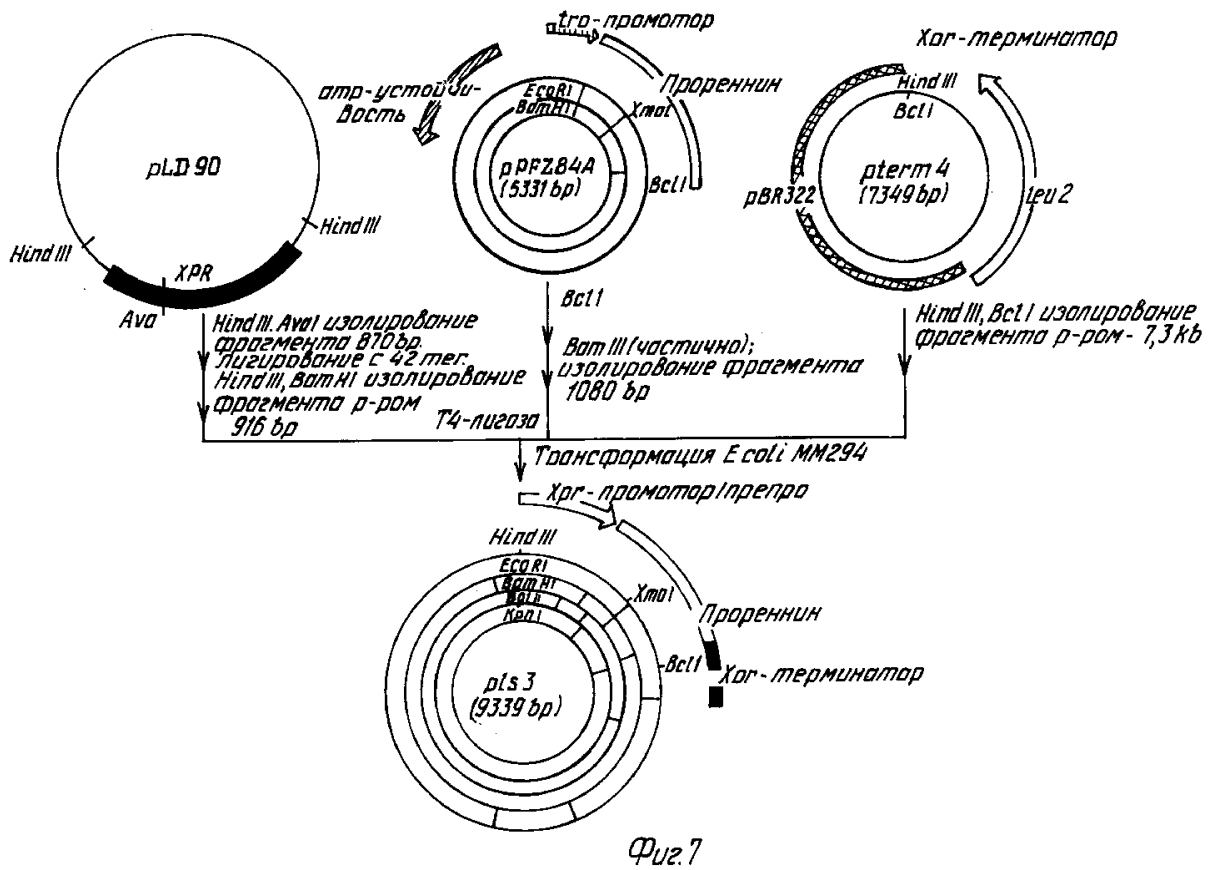
RU 2069694 C1

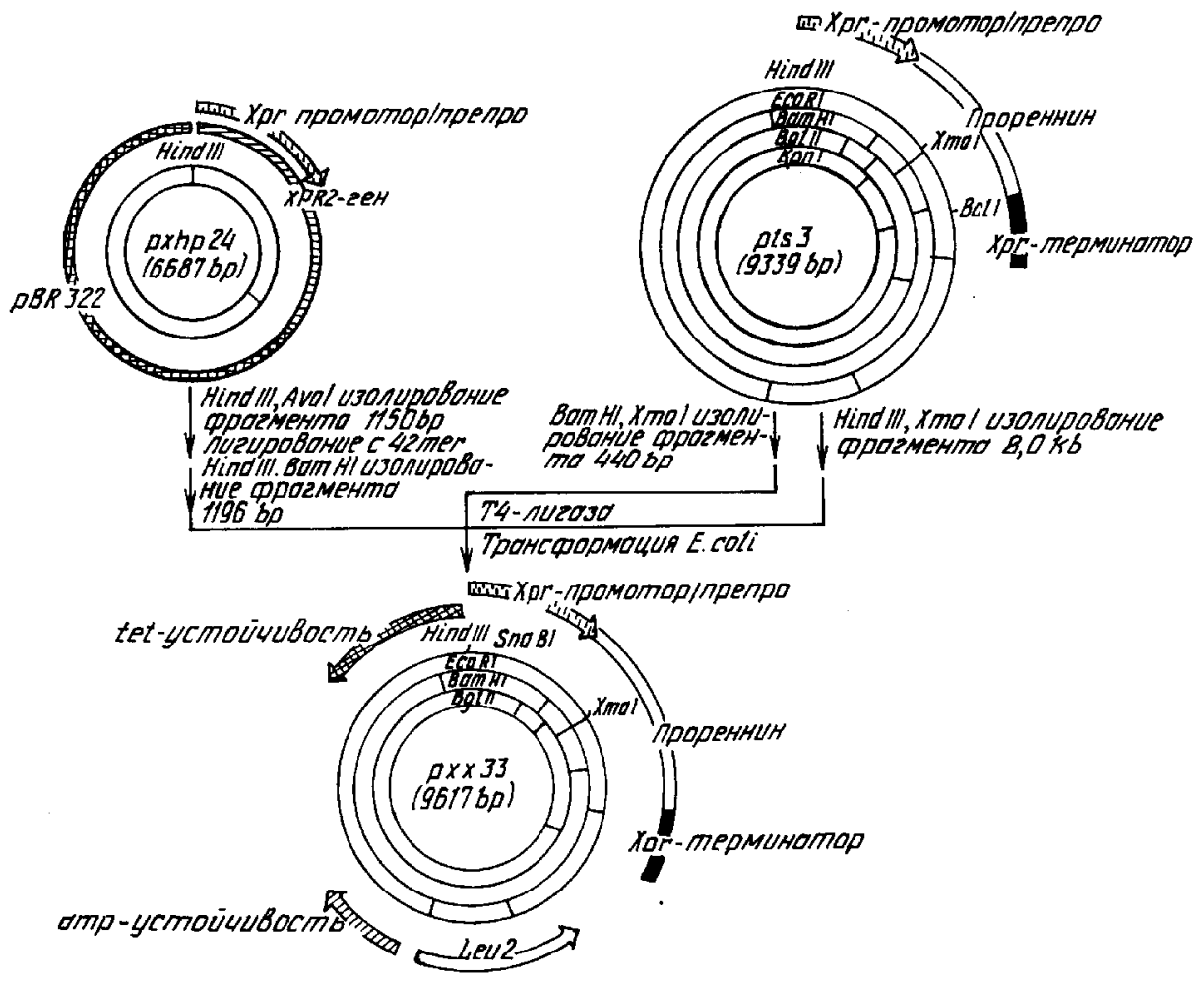


Фиг. 6

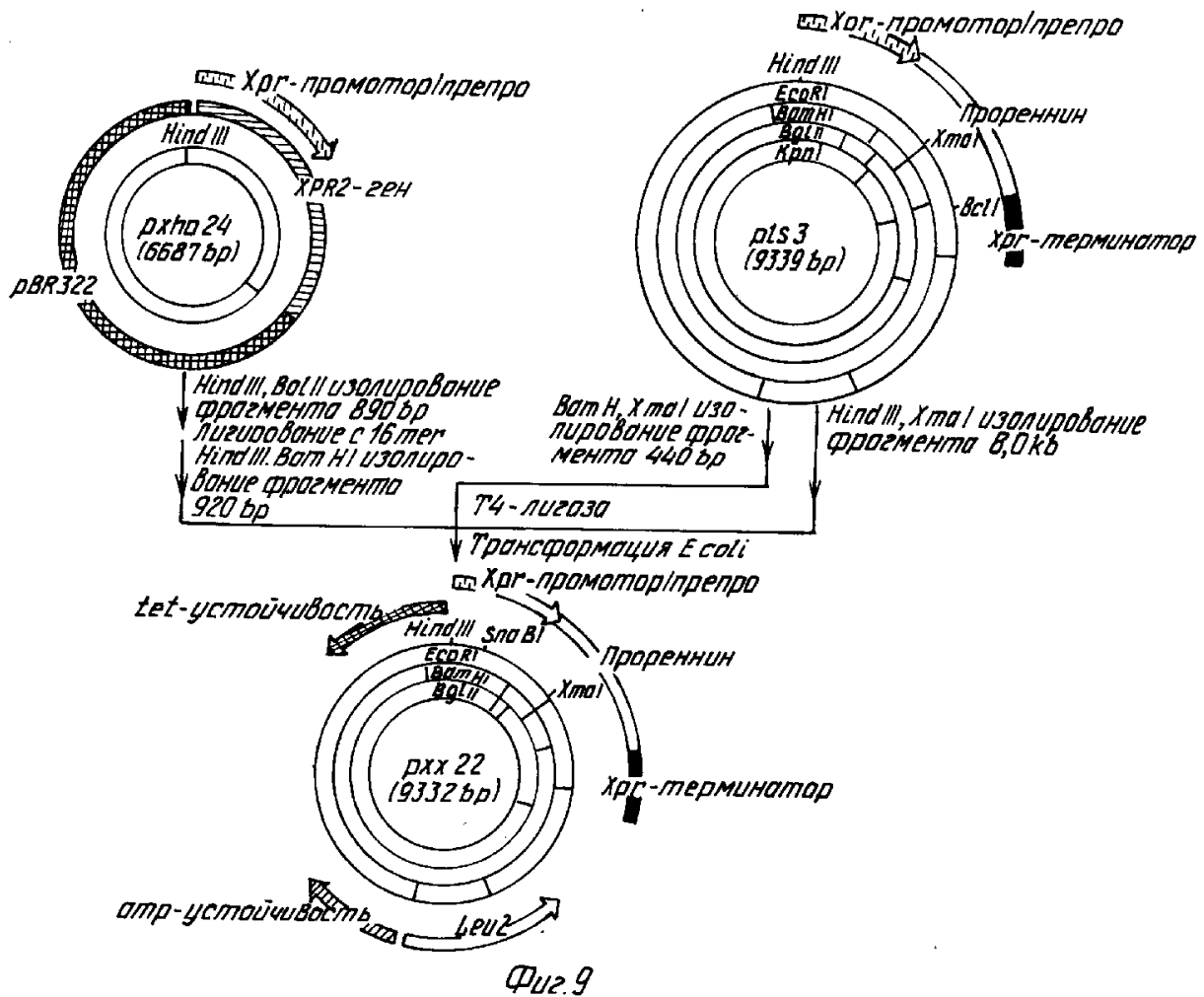
RU 2069694 C1

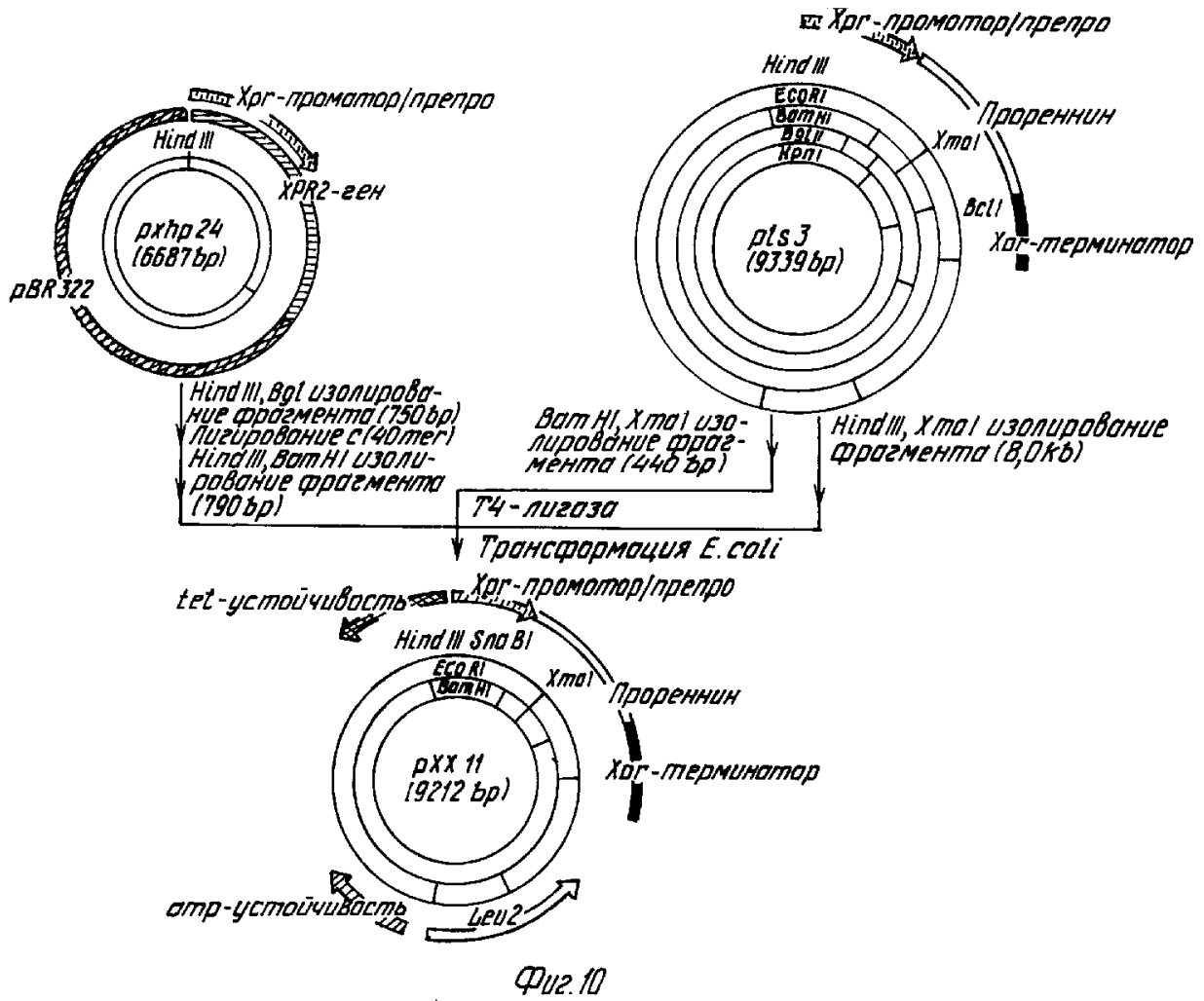
RU 2069694 C1





Фиг. 8





1 10 20
 thr leu gln lys lys ile glu glu ile ala ala lys tyr lys his ser val val lys lys cys cys tyr asp gly ala cys
 (A) (B)

TCT ATG ACT CTC CAA AAG AAG ATC GAA GAA ATC GCT GCT AAG TAC AAG CAC TCC CTC GTT AAG AAG TGT TGT TAC GAT GGT GCA TGC
 CA TAC TGA GAC GTT TTC TTC TAG CIT CTT TAG CGA CGA TTC ATG TTC GTG ACG CAG CAA TTC TTC ACA ACA ATG CTA CCA CGT ACG
 (J) (I)

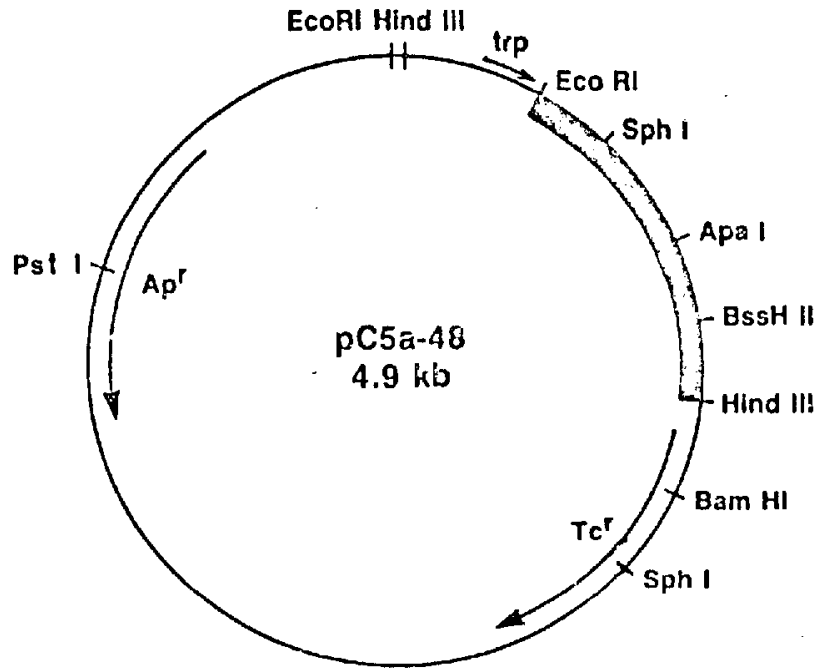
30 40 50
 asn asn asp glu thr cys glu gln arg ala ala arg ile ser leu gly pro arg cys ile lys ala phe thr glu cys cys val val
 (C) (D)

ATC AAC CAC GAA ACC TGT GAA CAA CGA GCT GCT GGT ATT TCT CTC GGC CCT CCC TCT ATC AAG CCT TTC ACT GAA TGT TGT GTT GTC
 TTG TTT CTG CTT TGG ACA CTT GTT GCT CGA CGA CCA TAA ACA GAC CCG GGA GCG ACA TAG TTC CGA AAG TGA CTT ACA ACA CAA CAG
 (H) (G)

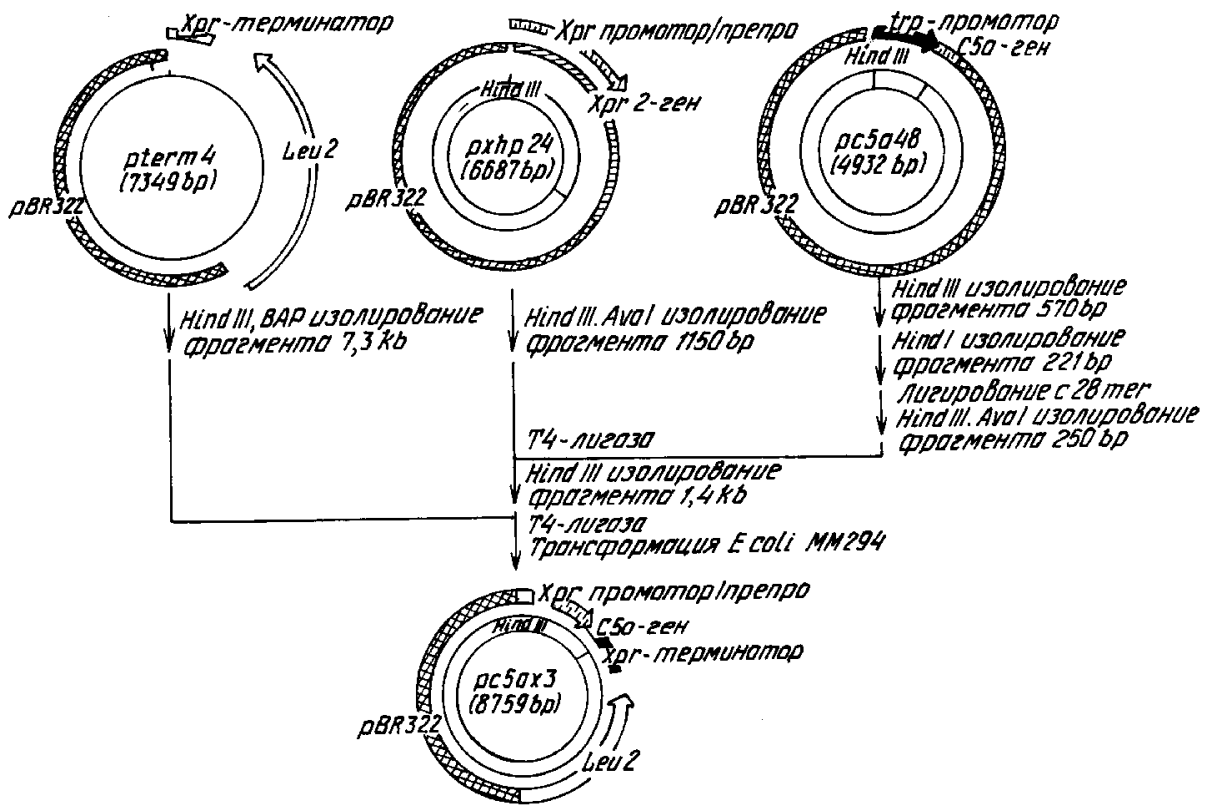
60 70 74
 ser gln leu arg ala asn ile ser his lys asp met gln leu gly arg stop
 (E) HindIII

TCC CAA CTG GGC GCT AAC ATT TCT CAC AAG GAC ATG CAA CTC GGC GCG TAA A
 AGC GTT CAC GCG CGA TTG TAA AGA TG TTC CTG TAC GTT GAG CCG GCG ATT TTCGA
 (F)

Фиг. 11



Фиг. 12



Фиг. 13

GTCGACACCATATCATATAAACTAACAATGCATTGCTTATTACGAAGACTACCCGTTGCTATCTCCAC
 ACCGTTATCTCCACGGTCCAAAGGCTGCTCAATGTGCTGCATACGTAAACGTGGGGTGCAACCTTGAGCA
 CATAGTACTTTTCCGAAAACCGGCGATAATTAAGTGTGCACTCCAACCTTTTCACTGAGCGTAAAATG
 TGGAGAAGAAATCGGCACTAAAAAGTCAGGTAGACTGGAAAATCGCCATGAAATGAATATCTCTTGCT
 ACAGTAATGCCCAGCATCGAGGGGTATTGTGTACCAACACTATAGTGGCAGCTGAAGCGCTCGTGATT
 GTAGTATGAGTCTTTATTGGTGATGGGAAGAGTTCACTCAATATTCTCGTTACTGCCAAAACACCACGG
 TAATCGGCCAGACACCATGGATGTAGATACCAAGCCTGTGAATGTTATTCCAGCTAAAATGCACATGG
 TTGGTGAAGGAGTAGTTGCTGTGCAATTCCGTCGTCGCTGAGTCATCATTATTACCAGTTGGCCA
 CAAACCCTTGACGATCTCGTATGTCCCCTCCGACATACTCCCGGCCGGCTGGGGTACGTTCCGATAGCGC
 TATCGGCATCGACAAGGTTTGGGTCCCTAGCCGATAACCGCACTACCTGAGTCACAATCTTCGGAGGTTT
 AGTCTTCCACATAGCACGGGCAAAAGTGCCTATATATACAAGAGCGTTTGCCAGCCACAGATTTTCACT

1 ----->
 met glu pro glu thr lys lys
 CCACACACCACATCACACATAACAACCACACACATCCACA ATG GAA CCC GAA ACT AAG AAG

10 20
 thr lys thr asp ser lys lys ile val leu leu gly gly asp phe
 ACC AAG ACT GAC TCC AAG AAG ATT GTT CTT CTC GGC GGC GAC TTC

30
 cys gly pro glu val ile ala glu ala val lys val leu lys ser
 TGT GGC CCC GAG GTG ATT GCC GAG GCC GTC AAG GTG CTC AAG TCT

40 50
 val ala glu ala ser gly thr glu phe val phe glu asp arg leu
 GTT GCT GAG GCC TCC GGC ACC GAG TTT GTG TTT GAG GAC CGA CTC

60
 ile gly gly ala ala ile glu lys glu gly glu pro ile thr asp
 ATT GGA GGA GCT GCC ATT GAG AAG GAG GGC GAG CCC ATC ACC GAC

70 80
 ala thr leu asp ile cys arg lys ala asp ser ile met leu gly
 GCT ACT CTC GAC ATC TGC CGA AAG GCT GAC TCT ATT ATG CTC GGT

90
 ala val gly gly ala ala asn thr val trp thr thr pro asp gly
 GCT GTC GGA GGC GCT GCC AAC ACC GTA TGG ACC ACT CCC GAC GGA

100 110
 arg thr asp val arg pro glu gln gly leu leu lys leu arg lys
 CGA ACC GAC GTG CGA CCC GAG CAG GGT CTC CTC AAG CTG CGA AAG

120
 asp leu asn leu tyr ala asn leu arg pro cys gln leu leu ser
 GAC CTC AAC CTG TAC GCC AAC CTG CGA CCC TGC CAG CTG CTG TCG

CPU 2.14

RU 2069694 C1

RU 2069694 C1

130 pro lys leu ala asp leu ser pro ile arg asn val glu gly thr
 CCC AAG CTC GCC GAT CTC TCC CCC ATC CGA AAC GTT GAG GGC ACC
 140
 asp phe ile ile val arg glu leu val gly gly ile tyr phe gly
 GAC TTC ATC ATT GTC CGA GAG CTC GTC GGA GGT ATC TAC TTT GCA
 150
 160 glu arg lys glu asp asp gly ser gly val ala ser asp thr glu
 GAG CGA AAG GAG GAT GAC GGA TCT GGC GTC GCT TCC GAC ACC GAG
 170
 180 thr tyr ser val pro glu val glu arg ile ala arg met ala ala
 ACC TAC TCC GTT CCT GAG GTT GAG CGA ATT GCC CGA ATG GCC GCC
 190
 phe leu ala leu gln his asn pro pro leu pro val trp ser leu
 TTC CTG GCC CTT CAG CAC AAC CCC CCT CTT CCC GTG TGG TCT CTT
 200
 210 asp lys ala asn val leu ala ser ser arg leu trp arg lys thr
 GAC AAG GCC AAC GTG CTG GCC TCC TCT CGA CTT TGG CGA AAG ACT
 220
 val thr arg val leu lys asp glu phe pro gln leu glu leu asn
 GTC ACT CGA GTC CTC AAG GAC GAA TTC CCC CAG CTC GAG CTC AAC
 230
 240 his gln leu ile asp ser ala ala met ile leu ile lys gln pro
 CAC CAG CTG ATC GAC TCG GCC GCC ATG ATC CTC ATC AAG CAG CCC
 250
 ser lys met asn gly ile ile ile thr thr asn met phe gly asp
 TCC AAG ATG AAT GGT ATC ATC ATC ACC ACC AAC ATG TTT GGC GAT
 260
 270 ile ile ser asp glu ala ser val ile pro gly ser leu gly leu
 ATC ATC TCC GAC GAG GCC TCC GTC ATC CCC GGT TCT CTG GGT CTG
 280
 leu pro ser ala ser leu ala ser leu pro asp thr asn glu ala
 CTG CCC TCC GCC TCT CTG GCT TCT CTG CCC GAC ACC AAC GAG GCG
 290
 300 phe gly leu tyr glu pro cys his gly ser ala pro asp leu gly
 TTC GGT CTG TAC GAG CCC TGT CAC GGA TCT GCC CCC GAT CTC GGC
 310
 lys gln lys val asn pro ile ala thr ile leu ser ala ala met
 AAG CAG AAG GTC AAC CCC AAT GCC ACC ATT CTG TCT GCC GCC ATG
 320
 330 met leu lys phe ser leu asn met lys pro ala gly asp ala val
 ATG CTC AAG TTC TCT CTT AAC ATG AAG CCC GCC GGT GAC GCT GTT

Фиг. 15

RU 2069694 C1

RU 2069694 C1

340 350
glu ala ala val lys glu ser val glu ala gly ile thr thr ala
GAG GCT GCC GTC AAG GAG TCC GTC GAG GCT GGT ATC ACT ACC GCC

360
asp ile gly gly ser ser ser thr ser glu val gly asp leu leu
GAT ATC GGA GGC TCT TCC TCC ACC TCC GAG GTC GGA GAC TTG TTG

370 380
pro thr arg ser arg ser cys ser arg arg ser lys ser phe leu
CCA ACA AGG TCA AGG AGC TGC TCA AGA AGG AGT AAG TCG TTT CTA

390
arg arg ile asp gly arg ser lys leu thr arg leu arg val gly
CGA CGC ATT GAT GGA AGG AGC AAA CTG ACG CGC CTG CGG GTT GGT

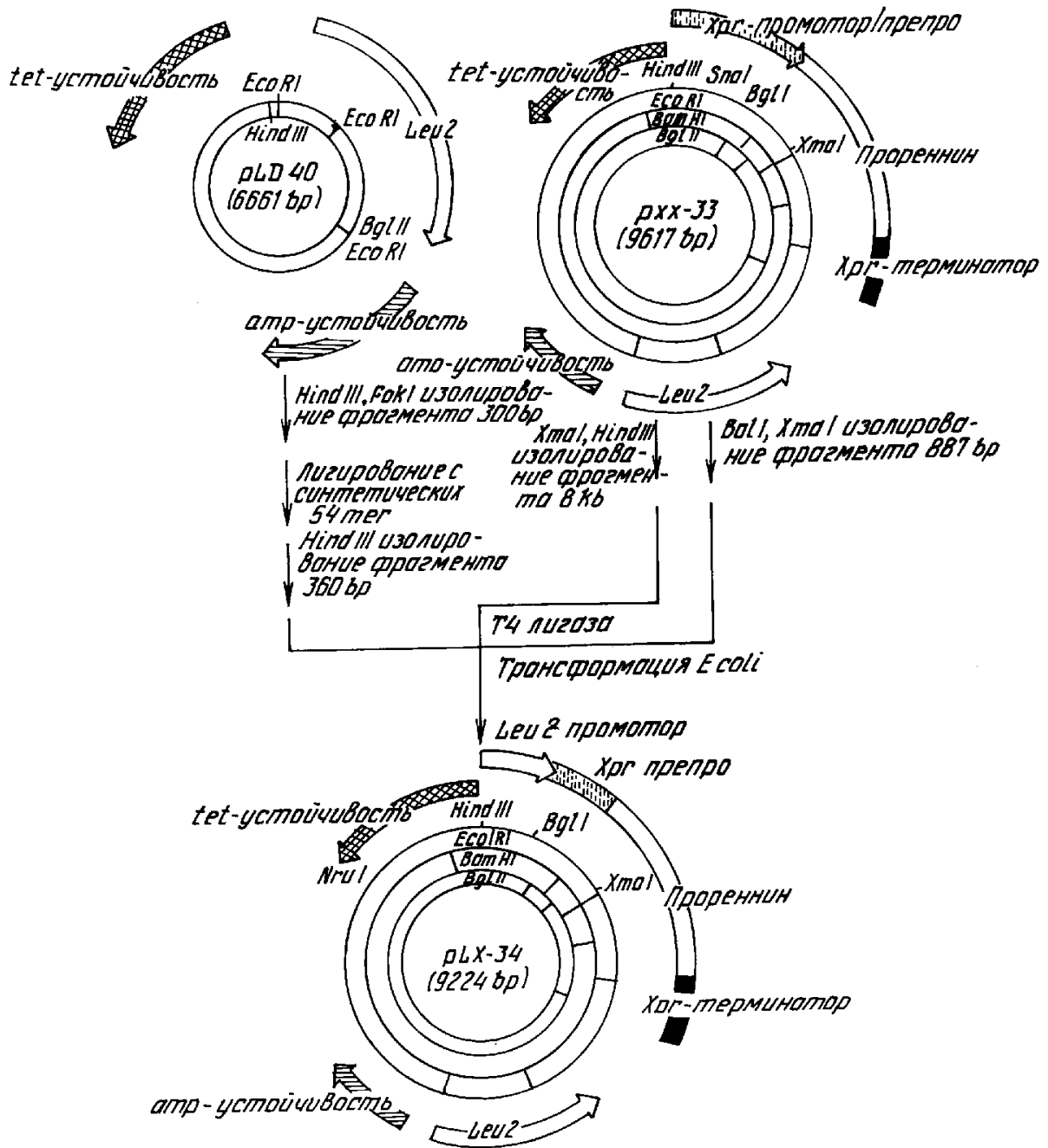
400 405
leu pro ala gly ser ala ser val OC
CTA CCG GCA GGG TCC GCT AGT GTA TAA GACTCTATAAAAAGGGCCCTGCCCTGCTAATGAA

ATGATGATTTATAATTTACCGGTGTAGCAACCTTGACTAGAGAAGCAGATTGGGTGTGTTTGTAGTGGA
GGACAGTGGTACGTTTTGGAAACAGTCTTCTTAAAAGTGTCTTGTCTACAGTATATTCACCTCATAACCTC
AATAGCCAAGGGTGTAGTCGGTTTATTAAAGGAAGGGAGTTGTGGCTGATGTGGATAGATATCTTTAAGC
TGGCGACTGCACCCAACGAGTGTGGTGGTAGCTTGTTACTGTATATTCCGTAAGATATATTTTGTGGGGT
TTTAGTGGTGTGTTGGTAGGTTAGTGCTTGGTATATGACTTGTAGGCATGACAATTCGAAAGGGGTGGAC
TTTGGGAATATTGTGGGATTCAATACCTTAGTTTGTACAGGGTAATTGTTACAAATGATACAAAGAACT
GTATTTCTTTTCATTTGTTTAAATTGTTGTATATCAAGTCCGTTAGACGAGCTCAGTGCCATGGCTTTT
GGCACTGTATTTTCAATTTTAGAGGTACACTACATTCACTGAGGTATGGTAAGGTTGAGGGCATAATGAAG
GCACCTTGTACTGACAGTCACAGACCTCTCACCGAGAATTTTATGAGATATACTCGGGTTCATTTTAGGC
TCCGATTCGATTCAAATTATTACTGTGCAAATCGGTTGAGCATCCGTTGATTTCCGAACAGATCTCGGCA
GTCTCTCGGATGTAGAATTAGGTTTCCTTGAGGGCGAAGATCCGTTTGTGTGACATGAATT

φuz.16

RU 2069694 C1

RU 2069694 C1



Фиг. 17