(19) Евразийское патентное ведомство

(11) 017712 (13) B1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2013.02.28

(21) Номер заявки
200970212

(22) Дата подачи заявки
2007.08.21

(51) Int. Cl. ***C12N 9/44*** (2006.01)
C12N 15/12 (2006.01)
C12N 5/10 (2006.01)

(54) КОМПОЗИЦИИ ДЛЯ РАЗЖИЖЕНИЯ И ОСАХАРИВАНИЯ КРАХМАЛА, ОСНОВАННЫЕ НА АКТИВНОСТИ ПУЛЛУЛАНАЗЫ, И СПОСОБЫ ПОЛУЧЕНИЯ ПЕПТИДОВ С АКТИВНОСТЬЮ ПУЛЛУЛАНАЗЫ

(31) **60/839,735; 60/903,247**
(32) **2006.08.23; 2007.02.23**
(33) **US**
(43) **2009.08.28**
(86) **PCT/US2007/018523**
(87) **WO 2008/024372 2008.02.28**

(71)(73) Заявитель и патентовладелец:
ДАНИСКО ЮЭс, ИНК., ДЖЕНЕНКОР ДИВИЖН (US)

(72) Изобретатель:
Инглэнд Джордж (US), Колкман Марк (NL), Миллер Брайан С. (US), Врумен Каспер (NL)

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

(56) DATABASE Geneseq [Online], 9 February 2006 (2006-02-09), "Bacterial pullulanase protein sequence", XP002469439, retrieved from EBI accession no. GSP:AEE27547, Database accession no. AEE27547, abstract

DATABASE Geneseq [Online], 24 September 2001 (2001-09-24), "Bacillus deramificans pullulanase", XP002469440, retrieved from EBI accession no. GSP:AAE05690, Database accession no. AAE05690, abstract

DATABASE Geneseq [Online], 24 September 2001 (2001-09-24), "Bacillus deramificans pullulanase DNA", XP002469441, retrieved from EBI accession no. GSN:AAD10891, Database accession no. AAD10891, abstract

WO 01/51620 A (NOVOZYMES AS [DK]; SVENDSEN ALLAN [DK]; ANDERSEN CARSTEN [DK]; VEDEL B), 19 July 2001 (2001-07-19), the whole document

US-B1-6429358

(57) Изобретение относится к новым вариантам ферментного пептида пуллуланазы, к генным последовательностям, кодирующим указанные новые пептиды, к векторам экспрессии, содержащим эти генные последовательности, а также к организмам, экспрессирующим новые варианты пуллуланазы. Новые варианты пуллуланазы, предлагаемые изобретением, были изготовлены эмпирически с применением процедур кодонной оптимизации, селективного усечения молекул "дикого типа", а также с замещением избранных аминокислотных остатков. Кроме того, изобретение относится к применению этих новых пептидов пуллуланазы в производстве текстиля, в процессах брожения, в пищевой промышленности и в других отраслях индустрии.

017712 B1

017712 B1

Перекрестная ссылка на родственные заявки

Эта заявка притязает на приоритет предварительных заявок Соединенных Штатов Америки US 60/903247, зарегистрированной 23 февраля 2007 г., и US 60/839735, зарегистрированной 23 августа 2006 г., каждая из которых включена сюда в качестве ссылки во всей полноте.

Область техники, к которой относится изобретение

Изобретение относится к новым вариантам ферментного пептида пуллуланазы, к генным последовательностям, кодирующим указанные новые пептиды, к векторам экспрессии, включающим эти генные последовательности, а также к организмам, экспрессирующим новые варианты пуллуланазы. Кроме того, изобретение относится к применению этих новых пептидов пуллуланазы в производстве текстиля, в процессах брожения, в пищевой промышленности и в других отраслях индустрии.

Уровень техники

Пуллуланазы - это ферменты, нашедшие полезное применение в многочисленных промышленных приложениях, особенно в пищевой индустрии и в производстве напитков. Пуллуланазы представляют собой ферменты, деветвящие молекулы крахмала: они эффективно устраняют ветвление продуктов гидролиза крахмала (полезных при изготовлении теста), ветвление β-лимитирующих декстранов (полезных в пивоварении), а также полезны при выработке сахарных сиропов, например, из кукурузы, картофеля, пшеницы, маниоки и риса. Пуллуланазы как ферменты классифицированы в документе ЕС 3.2.1.41, причем такие ферменты характеризуются по способности гидролизовать α-1,6-гликозидные связи, например, в амилопектине и пуллулане.

Пуллуланазы являются продуктом жизнедеятельности бактерий, особенно рода Bacillus. Выработка пуллуланаз для промышленного применения далеко не беспроблемна. Пуллуланазы быстро разрушаются под воздействием различных протеаз, также вырабатываемых бактериями, в связи с чем извлечение пуллуланаз в больших количествах становится дорогостоящим и неэффективным. Многие специалисты в данной области техники разрабатывали способы увеличения выработки пуллуланазы, основанные на ограничении ее разрушения в культуре. Например, ранее мы показали, что для экономичной экспрессии активной пуллуланазы необходимо удаление из штаммов, вырабатывающих пуллуланазу, генов AprL и Mpr, которые экспрессируют протеазы. Кроме того, чтобы уменьшить протеолитическое разрушение и потерю активности вырабатываемой пуллуланазы, время брожения необходимо ограничить 51-60 ч. Svendsen сконструировал варианты пуллуланазы, имеющие измененную трехмерную конформацию фермента, что увеличивает его термостабильность или изменяет характеристики расщепления ферментов его субстрата (см. патенты США 6350599 и 6838257, а также патентную заявку США № 2004/0082028).

Относительно недавно мы показали, что временной паттерн распада пуллуланазы был определен со следующим результатом: в промежутке времени от 30 до 50 ч наблюдалась частичная обрезка полноразмерной молекулы пуллуланазы с образованием усеченных молекул, утративших 98 и 102 N-концевые аминокислоты соответственно. Обрезку, которая происходит на N-конце в остатках глутаминовой кислоты E99 и E103, соответственно можно визуализировать посредством HPLC. Удивительно, что усеченные молекулы пуллуланазы 1-98 и 1-102 сохраняют активность пуллуланазы, и еще более удивительно, что эта активность усеченных молекул согласно полученным данным превышает активность полноразмерной молекулы пуллуланазы. Спустя 51 ч дальнейшее разрушение молекул пуллуланазы приводит к спаду активности, который, в конечном итоге, завершается полной утратой активности.

По-прежнему, имеется возможность конструктивного улучшения пептидов пуллуланазы и тех нуклеотидных последовательностей, которые кодируют эти пептиды. Таким образом, существует потребность в соединениях такого рода и в способах более эффективной выработки пуллуланазы, которые включают, например, уменьшение протеолитического разрушения, повышение титров сбраживания или увеличение активности пуллуланазы.

Сущность изобретения

Изобретение относится к новым и неочевидным формам пуллуланазы, пептида со свойствами фермента.

Пуллуланазы, предлагаемые изобретением, включают новые модификации, обладающие улучшенными рабочими характеристиками в отношении выработки титров и/или устойчивости к разрушению (например, ферментативному разрушению протеиназами) и/или более активные в расщеплении целевых субстратных материалов по сравнению с родительскими пептидами ("дикого типа").

В этом смысле изобретение относится к пуллуланазам с пептидными последовательностями SEQ ID NO: 2, SEQ ID NO: 4 и SEQ ID NO: 6, как это показано на фиг. 7B, 8B и 9B соответственно. Изобретение также относится к нуклеотидным последовательностям, кодирующим аминокислотные последовательности SEQ ID NO: 2, 4 и 6. Соответствующие нуклеотидные последовательности также показаны на фиг. 7A, 8A и 9A как SEQ ID NO: 1, 3 и 5 соответственно. Как хорошо известно в данной области знаний, генетический код в избыточном количестве содержит множественные нуклеотидные кодоны, кодирующие одну и ту же аминокислоту. Таким образом, изобретение дополнительно относится к любым чередующимся нуклеотидным последовательностям, которые кодируют полипептиды с SEQ ID NO: 2, 4 и 6. Компетентный специалист способен определить нуклеотидные последовательности, которые кодируют

пептидные последовательности, предлагаемые изобретением, основываясь на изучении описания этого изобретения и на общих знаниях в данной области техники.

Изобретение относится к экспрессирующим конструкциям, которые кодируют пептиды, предлагаемые изобретением. Изобретение не ограничивается какой-либо специфической экспрессирующей конструкцией при том условии, что она способна экспрессировать пептиды, предлагаемые изобретением. В этом отношении не ограничивающие примеры подходящих экспрессирующих конструкций схематически показаны в части (a) фиг. 4-6.

Как разъясняется более обстоятельно в подробном описании изобретения и в разделах примеров этого описания, в одном из вариантов реализации изобретения последовательность, кодирующая родительский пептид пуллуланазы (из Bacillus deramificans), была модифицирована с применением методик кодонной оптимизации для получения нуклеотидной последовательности с оптимизированным составом кодонов [SEQ ID NO: 1], кодирующей дубликат родительской (т.е. дикого типа) пуллуланазы с аминокислотной последовательностью [SEQ ID NO: 2]. Нуклеотидная последовательность, кодирующая аминокислотную последовательность [SEQ ID NO: 2], была клонирована в двух направлениях в сайте XhoI вектора интеграции pICatH B. licheniformis, создавая версии экспрессирующей конструкции Ori1 (pICatH-PUL-Ori1) и Ori2 (pICatH-PUL-Ori2).

В другом варианте реализации изобретения последовательность, кодирующая родительский пептид пуллуланазы, была модифицирована для получения пептида пуллуланазы с N-концевой делецией 104 аминокислот. Нуклеотидная последовательность, кодирующая эту новую пуллуланазу, в одном из вариантов реализации изобретения была также оптимизирована по кодонному составу [SEQ ID NO: 3]. Пептид, экспрессированный этой конструкцией, PULm 104, описан в SEQ ID NO: 4 (см. фиг. 8(b)). Кодирующая нуклеотидная последовательность [SEQ ID NO: 4] была клонирована в двух направлениях в сайте XhoI вектора интеграции pICatH B. licheniformis, создавая версии экспрессирующей конструкции Ori1 (pICatH-PUL-Ori1) и Ori2 (pICatH-PUL-Ori2).

Еще в одном варианте реализации изобретения последовательность, кодирующая родительский пептид пуллуланазы, была изменена для замещения аминокислотных остатков глутаминовой кислоты (E) в положениях 99 и 103 на глутамин (Q) с той целью, чтобы результирующий пептид стал более устойчивым к протеолитическому разрушению в этих положениях. Нуклеотидная последовательность, кодирующая эту новую пуллуланазу, в одном из вариантов реализации изобретения была также оптимизирована по кодонному составу [SEQ ID NO: 5]. Пептид, экспрессированный этой последовательностью нуклеиновой кислоты, PUL_E99Q_E103Q, описан в SEQ ID NO: 6. Кодирующая нуклеотидная последовательность [SEQ ID NO: 6] была клонирована в двух направлениях в сайте XhoI вектора интеграции pICatH B. licheniformis, создавая версии экспрессирующей конструкции Ori1 (pICatH-PUL_E99Q_E103Q-Ori1) и Ori2 (pICat-PUL_E99Q_E103Q-Ori2).

Изобретение также относится к трансфекции экспрессирующей конструкции, предлагаемой изобретением, в подходящие организмы-хозяева. Изобретение не ограничивается каким-либо конкретным организмом-хозяином. Организмом-хозяином может быть, например, микроорганизм, эукариотическая клетка или тканевая культура, растительная клетка или тканевая культура, либо грибковая клетка или тканевая культура. В предпочтительном варианте реализации изобретения организмом-хозяином является микроорганизм. Предпочтительные организмы-хозяева включают, но не ограничиваясь ими, разные виды Bacillus (особенно Bacillus subtilis, B. licheniformis и B. deramificicans), Escherichia coli, Trichoderma reesei, Saccharomyces cerevisiae или Aspergillus niger. В наиболее предпочтительном варианте реализации изобретения организмом-хозяином является B. licheniformis.

Изобретение относится к выделению и очистке пептидов, предлагаемых изобретением, из среды, в которой были культивированы организмы-хозяева, относящиеся к изобретению. В этом отношении изобретение не ограничивается какой-либо конкретной методикой выделения и очистки при том условии, что она позволяет достичь минимальной степени чистоты 10%. В более предпочтительном варианте реализации изобретения минимальная чистота выделенного и очищенного пептида составляет 25%, а еще в более предпочтительном варианте реализации изобретения минимальная чистота выделенного и очищенного пептида составляет 50%. Еще в более предпочтительном варианте реализации изобретения минимальная чистота выделенного и очищенного пептида составляет 75%. В наиболее предпочтительном варианте реализации изобретения минимальная чистота выделенного и очищенного пептида, предлагаемого изобретением, составляет 90%. Минимальная чистота может быть измерена в процентах общего сухого веса или выражена в других подходящих показателях, которые известны в данной области знаний.

Изобретение не ограничивается какими-либо конкретными способами выделения и очистки пептидов, предлагаемых изобретением. Подходящими считаются любые средства очистки пептидов, известные в данной области знаний. Не ограничивающие примеры пригодных средств очистки включают аффинную хроматографию, осаждение, эксклюзионную хроматографию, тонкослойную хроматографию, электрофорез, гель-фильтрацию и т.д.

Компетентный специалист признает, что биологически активный фрагмент пуллуланаз, предлагаемых изобретением, можно применять вместо полноразмерной последовательности или как ее эквивалент

в контексте изобретения. Понятие "биологически активный фрагмент" охватывает все аналоги, миметики, усечения, делеции и/или замещения в последовательностях, предлагаемых изобретением. Пептидомиметики пуллуланаз и активные домены пуллуланаз, предлагаемых изобретением, могут быть сконструированы с применением вычислительного подхода, опираясь на структурные данные, как это известно в данной области знаний. Дополнительно, в одном из вариантов реализации изобретения рассматривается возможность генерирования аналогов и мутаций нуклеотидных последовательностей пуллуланаз, предлагаемых изобретением, посредством направленной молекулярной эволюции. Методики направленной молекулярной эволюции известны в данной области знаний (см., например, патент США № 5605793, выданный Stemmer, et al., или патент США № 6537776, выданный Short, которые включены сюда в качестве ссылки). Белки, созданные посредством направленной молекулярной эволюции, будут иметь меньшую, большую или равную способность функционировать в качестве пуллуланазы по сравнению с пептидами, предлагаемыми изобретением.

Еще в одном варианте реализации изобретения пептиды, предлагаемые изобретением, применяются как гибридные белки, полученные в результате гибридизации, например, с другими структурными или функциональными доменами пептида. Такие домены способны, например, придавать гибридным белкам другие ферментативные свойства или привязывать пептид к поверхности.

Пептиды, предлагаемые изобретением, также включают пептиды, являющиеся результатом существования полимерных генов, альтернативных событий в ходе транскрипции, альтернативных событий в ходе сплайсинга РНК, а также альтернативных трансляционных и посттрансляционных событий. Полипептид может экспрессироваться в системах, например, в культивируемых клетках, что сопровождается, по существу, теми же посттрансляционными модификациями, которые наблюдаются при экспрессии пуллуланазы в нативных клетках, либо в таких системах, которые не приводят к посттрансляционным модификациям, наблюдаемым при экспрессии пуллуланазы в нативных клетках.

В сферу изобретения также входят композиция, которая включает один или более пептидов пуллуланазы (или кодирующую их нуклеиновую кислоту), и один или более добавочных компонентов, например носитель, разбавитель или растворитель. Добавочным компонентом может быть такой компонент, который делает композицию пригодной для применения in vitro, in vivo, а также фармацевтического или ветеринарного применения.

Еще в одном аспекте изобретение предлагает в достаточной степени чистую нуклеиновую кислоту, имеющую или включающую нуклеотидную последовательность, которая кодирует полипептид, аминокислотная последовательность которого включает или представляет собой последовательность пептида пуллуланазы, предлагаемого изобретением.

В предпочтительных вариантах реализации изобретения рассматриваемая нуклеиновая кислота пуллуланазы включает регуляторную последовательность транскрипции, например по меньшей мере одну из последовательностей промотора транскрипции или энхансера транскрипции, функционально связанных с последовательностью гена пуллуланазы, например, для того, чтобы сделать ген пуллуланазы пригодным для использования в качестве вектора экспрессии.

Еще в одном предпочтительном варианте реализации изобретения нуклеиновая кислота, кодирующая пептид пуллуланазы, предлагаемый изобретением, гибридизируется в жестких условиях с зондом нуклеиновой кислоты, соответствующим по меньшей мере 12 последовательным нуклеотидам из SEQ ID NO: 1, 3 или 5, предпочтительнее по меньшей мере с 20 последовательными нуклеотидами из SEQ ID NO: 1, 3 или 5.

Еще один предпочтительный вариант реализации изобретения предлагает применение описанной здесь пуллуланазы в различных промышленных приложениях. Например, изобретение относится к применению новых вариантов пуллуланазы, предлагаемых этим изобретением, в производстве продуктов питания (включая, но не ограничиваясь ими, различные виды теста и сиропы), напитки (включая, но не ограничиваясь ими, различные сорта пива и эля), биоэтанола и еще многих продуктов, хорошо известных специалистам, работающим в этой области.

Краткое описание чертежей

Фиг. 1 представляет схематическое изображение текущего процесса выработки пуллуланазы и вариантов пуллуланазы, сконструированных в этом проекте;

фиг. 2 демонстрирует молекулярную архитектуру конструкций, которые представлены в штаммах, экспрессирующих пуллуланазу в соответствии с изобретением, по сравнению с конструкцией, экспрессирующей пуллуланазу в штамме BMP 139;

фиг. 3 представляет схематическое изображение вектора pICatH. Вектор pICatH содержит температурочувствительную начальную точку репликации (ori pE194, для репликации в Bacillus), ori pBR322 (для амплификации в E. coli), ген устойчивости к неомицину для отбора, нативный ген B. licheniformis, детерминирующий устойчивость к хлорамфениколу (cat) с повторами для отбора, а также обладает способностью интегрирования с хромосомой и кассетной амплификации;

фиг. 4 представляет схематическое изображение конструкции pICatH-PUL Ori1;

фиг. 5 представляет схематическое изображение конструкции pICatH-PULm104Ori1;

фиг. 6 представляет схематическое изображение конструкции pICatH-PUL_E99Q_E103Q Ori;

Фиг. 7A, B отображают (a) последовательность нуклеиновых кислот [SEQ ID NO: 1] и (b) аминокислотную последовательность [SEQ ID NO: 2] пуллуланазы "дикого типа" (PUL), оптимизированной по составу кодонов;

фиг. 8A, B отображают (a) последовательность нуклеиновых кислот [SEQ ID NO: 3] и (b) аминокислотную последовательность [SEQ ID NO: 4] пуллуланазы PULm104;

фиг. 9A, B отображают (a) последовательность нуклеиновых кислот [SEQ ID NO: 5] и (b) аминокислотную последовательность [SEQ ID NO: 6] пуллуланазы PUL_E99Q_E103Q.

Раздел определений

Следует отметить, что при использовании в этом описании изобретения и прилагаемых пунктах формулы изобретения формы единственного числа подразумевают ссылку на формы множественного числа, если из контекста явно не следует иное. Таким образом, ссылка на композицию, содержащую "соединение", включает смесь двух или более соединений. Следует также отметить, что термин "или" обычно употребляется в том смысле, что он включает вариант "и/или", если из контекста явно не следует иное.

Термин "пуллуланаза" относится к особой разновидности глюканазы, амилолитическому эндоферменту, расщепляющему пуллулан. Он вырабатывается грамотрицательными бактериями рода Klebsiella, например, как внеклеточный липопротеин, закрепленный на поверхности. Однако грамположительные бактерии вырабатывают пуллуланазы как секретируемые белки. Пуллуланазы I типа специфически атакуют α-1,6 связи, в то время как пуллуланазы II типа способны также гидролизировать α-1,4 связи. Пуллуланазу также вырабатывают некоторые другие бактерии и археи. В биотехнологии пуллуланазу используют как детергент. Пуллуланаза (EC 3.2.1.41) также известна под названием пуллулан-6-глюканогидролаза (деветвящий фермент). Пуллулан считается цепью блоков мальтотриозы, соединенных α-1,6-глюкозидными связями. Пуллуланаза способна гидролитически расщеплять пуллулан (α-глюкановые полисахариды).

Термин "кодонная оптимизация" относится к методикам, направленным на повышение уровня экспрессии посредством замены нуклеотидных кодонов в кодирующей последовательности другими кодонами, которые кодируют ту же аминокислоту, но обеспечивают более эффективный процессинг в организме-хозяине. Предпочтение кодонов может разительно отличаться у различных биологических видов. Для того чтобы повысить уровень экспрессии чужеродного белка в определенной экспрессирующей системе (клетках бактерий, грибков, дрожжей, насекомых, растений или млекопитающих), очень важно подобрать для чужеродного белка такую частоту кодонов, которая как можно лучше соответствовала бы экспрессирующей системе организма-хозяина. Классическим примером такого рода является GFP (зеленый флуоресцентный белок), который был оптимизирован для достижения высокого уровня экспрессии в клетках млекопитающих. Таким образом, кодонную оптимизацию можно использовать для экспрессии белков, предлагаемых изобретением, в широком спектре организмов-хозяев, где такие последовательности вообще не могли бы экспрессироваться или экспрессировались бы неэффективно.

Термин "вектор экспрессии" при использовании здесь относится к рекомбинантной молекуле ДНК, содержащей желательную кодирующую последовательность и соответствующие последовательности нуклеиновой кислоты, необходимые для экспрессии функционально связанной кодирующей последовательности в конкретном организме-хозяине. Последовательности нуклеиновых кислот, необходимые для экспрессии в прокариотах, включают промотор, по выбору - последовательность оператора, сайт связывания рибосом, а также, возможно, другие последовательности.

Термин "гетерологичный промотор" при использовании здесь означает промотор, который не имеет природной связи с геном или очищенной нуклеиновой кислотой. Термины "промотор", "промоторный элемент" или "промоторная последовательность" при использовании здесь относятся к последовательности ДНК, которая при сшивании с представляющей интерес нуклеотидной последовательностью способна управлять транскрипцией этой важной последовательности в мРНК. Промотор типично, хотя и не обязательно, локализован на 5'-конце (т.е. в верхней части) представляющей интерес нуклеотидной последовательности, транскрипцию которой в мРНК он контролирует, кроме того, он создает сайт для связывания с РНК-полимеразой и другими факторами транскрипции для инициации транскрипции.

Термин "специфичный по типу клетки" или "специфичный для организма-хозяина", а также эквивалентные термины, употребляемые применительно к регуляторному элементу, относятся к регуляторному элементу, который способен вызывать направленную селективную экспрессию представляющей интерес нуклеотидной последовательности в определенном типе клетки или в определенном организме при относительном отсутствии экспрессии той же представляющей интерес нуклеотидной последовательности в другом типе клеток или организмов в пределах той же ткани. Термин "специфичный по типу клетки" или "специфичный для организма-хозяина", употребляемый применительно к регуляторному элементу, также означает регуляторный элемент, способный стимулировать селективную экспрессию представляющей интерес нуклеотидной последовательности в регионе в пределах какой-либо ткани или организма соответственно.

Терминологические определения "выделенный", "очищенный препарат" или "по существу чистый

препарат" полипептида при использовании здесь означают полипептид, который был идентифицирован и отделен по меньшей мере от одного загрязнителя, с которым он обычно ассоциирован в природном состоянии или при получении из его фактического источника. По меньшей мере одним добавочным загрязнителем могут быть, например, другие белки, липиды и нуклеиновые кислоты, встречающиеся в комбинации с указанным полипептидом в природном состоянии. Предпочтительно, чтобы полипептид также подвергался отделению от таких веществ, как, например, антитела или желеобразная матрица, например полиакриламид, которые были использованы для его очистки. Предпочтительно, чтобы содержание полипептида составляло по меньшей мере 10, 20, 50, 70, 80 или 95% от сухого веса очищенного препарата. Предпочтительно, чтобы препарат содержал достаточное количество полипептида для секвенирования белка (по меньшей мере 1, 10 или 100 мг полипептида).

Терминологическое определение "очищенный препарат клеток" при использовании здесь относится (применительно к растительным или животным клеткам) к препарату таких клеток in vitro, а не ко всему интактному растению или животному. Применительно к культивируемым клеткам или микробным клеткам такой препарат состоит по меньшей мере из 10%, предпочтительнее из 50% целевых клеток.

Терминологическое определение "по существу чистая нуклеиновая кислота", например по существу чистая ДНК, означает нуклеиновую кислоту, которая имеет одно (любое) или оба следующих свойства: не граничит непосредственно с одной или обеими последовательностями, например кодирующими последовательностями (т.е. одна на 5'-конце, а другая на 3'-конце), с которыми она граничит в природном геноме организма, из которого извлечена эта нуклеиновая кислота, или, по существу, не содержит последовательности нуклеиновой кислоты, с которой она встречается в том организме, из которого извлечена эта нуклеиновая кислота. Этот термин включает, например, рекомбинантную ДНК, которая встроена в вектор, например, в автономно реплицирующиеся плазмиду или вирус, или в геномную ДНК прокариота либо эукариота, или которая существует как отдельная молекула (например, кДНК или фрагмент геномной ДНК, полученный в результате ПЦР или обработки рестрикционной эндонуклеазой) независимо от других последовательностей ДНК.

Кроме того, термин "выделенная" при использовании по отношению к нуклеиновой кислоте, например "выделенная последовательность нуклеиновой кислоты", относится к последовательности нуклеиновой кислоты, которая идентифицирована и отделена по меньшей мере от одной загрязняющей нуклеиновой кислоты, с которой она обычно соединена в природном состоянии или при получении из фактического источника. Выделенная нуклеиновая кислота представляет собой нуклеиновую кислоту, представленную в такой форме или в таком окружении, которые отличаются от природного состояния этой нуклеиновой кислоты. В отличие от этого невыделенные нуклеиновые кислоты представляют собой такие нуклеиновые кислоты как ДНК и РНК, которые представлены в том состоянии, в котором они существуют в природе. Например, данная последовательность ДНК (например, ген) находится в хромосоме клетки-хозяина вблизи от соседних генов; последовательности РНК, такие как специфическая последовательность мРНК, кодирующая специфический белок, находятся в клетке в виде смеси со многими другими мРНК, которые кодируют множество белков. Однако выделенная последовательность нуклеиновой кислоты, содержащая, например, SEQ ID NO: 1, включает в качестве примера такие последовательности нуклеиновой кислоты в клетках, которые обычно содержат SEQ ID NO: 1, причем последовательность нуклеиновой кислоты в хромосомном или внехромосомном расположении отличается от таковой в природных клетках или иным образом фланкируется другой последовательностью нуклеиновой кислоты по сравнению с той, что встречается в природе. Выделенная последовательность нуклеиновой кислоты может быть представлена в однонитевой или двунитевой форме. Если выделенная последовательность нуклеиновой кислоты должна быть использована для экспрессии белка, то указанная последовательность нуклеиновой кислоты будет содержать (как минимум) по меньшей мере часть смысловой или кодирующей нити (т.е. последовательность нуклеиновой кислоты может быть однонитевой). В альтернативном варианте она может содержать как смысловую, так и антисмысловую нити (т.е. последовательность нуклеиновой кислоты может быть двунитевой).

Термин "гомологичный (гомологичная)" при использовании здесь относится к сходству последовательностей между двумя молекулами полипептида или между двумя молекулами нуклеиновой кислоты. Если определенное положение в обеих сравниваемых последовательностях занято одним и тем же основанием или мономерной субъединицей аминокислоты, например если положение в каждой из двух молекул ДНК занято аденином, то молекулы считаются гомологичными в этом положении. Процентный показатель гомологии между двумя последовательностями представляет собой функцию, выраженную числом совпадающих или гомологичных положений в двух последовательностях, деленным на число проверенных положений и умноженным на 100. В качестве примера последовательности ДНК ATTGCC и TATGGC гомологичны на 50%. Обычно сравнение проводят тогда, когда две последовательности выровнены для достижения максимальной гомологии.

Термины "пептид (пептиды)", "белок (белки)" и "полипептид (полипептиды)" используются здесь взаимозаменяемо.

Термин "протеаза" означает белковый или полипептидный домен белка или полипептида, извлеченного из микроорганизма, например грибка, бактерии, или из растения, или из животного, который

обладает способностью катализировать расщепление пептидных связей в одном или более различных положений скелета белковой молекулы.

Предпочтительно, чтобы белки пуллуланазы, соответствующие изобретению, были выделены или очищены. Под очисткой или выделением подразумевается, что белок пуллуланазы изменяет свое природное состояние на основании отделения пуллуланазы от некоторых или всех природных составляющих, с которыми он соединен в природе. Такие выделение или очистка могут быть достигнуты посредством методик разделения, общепризнанных в данной области техники, например ионообменной хроматографии, аффинной хроматографии, гидрофобного разделения, диализа, обработки протеазой, осаждением сульфатом аммония или осаждением белка другими солями, центрифугированием, гель-проникающей хроматографией, фильтрацией, микрофильтрацией, гель-электрофорезом или разделением в градиенте для удаления целых клеток, обломков клеток, загрязнений, посторонних белков или ферментов, присутствие которых в конечном составе композиции нежелательно. Кроме того, затем можно добавлять в композицию, содержащую пуллуланазу, такие компоненты, которые придают ей дополнительные полезные свойства, например активирующие агенты, противоингибирующие агенты, желательные ионы, соединения для подгонки pH или другие ферменты. Предпочтительно, чтобы белки пуллуланазы, соответствующие изобретению, были получены рекомбинантными способами.

При использовании здесь термин "микроорганизм" относится к бактериям, грибкам, вирусам, простейшим (протозойным) животным, а также к другим микробам или микроскопическим организмам. В изобретении микроорганизмы применяются в качестве организмов-хозяев для экспрессии экзогенных белков.

При использовании здесь термины "производное", "вариант" или "модифицированный пептид, полипептид, белок" означают белок, полученный из белка-предшественника (например, природного белка) за счет добавления к нему одной или более аминокислот на C-конце и/или на N-конце, замены одной или более аминокислот в одном или нескольких разных сайтах аминокислотной последовательности, делеции одной или более аминокислот на одном или обоих концах белка или в одном или более сайтов аминокислотной последовательности либо инсерции одной или более аминокислот в одном или более сайтов аминокислотной последовательности. Приготовление производного пуллуланазы предпочтительно достигается за счет модификаций последовательности ДНК, кодирующей природный белок, трансформации этой последовательности ДНК в подходящего хозяина и экспрессии модифицированной последовательности ДНК с получением на выходе производного пуллуланазы. "Производные" в изобретении включают пептиды, в том числе с измененной аминокислотной последовательностью по сравнению с аминокислотной последовательностью предшественника (например, диким типом или природным состоянием пуллуланазы), причем производные пептиды сохраняют основную природную сущность пуллуланазы-предшественника, однако имеют измененные свойства в некоторых специфических аспектах. Например, производное пуллуланазы может иметь улучшенный оптимум pH, повышенную устойчивость к ферментативному разрушению или иному разрушению, более высокую ферментативную эффективность, повышенную термостабильность и устойчивость к окислению, но при этом будет сохранять свою характерную ферментативную модифицирующую активность. Сходным образом, производные, в соответствии с изобретением включают белок или другой домен, связывающий субстрат, который был добавлен или модифицирован для изменения его способности связывать субстрат.

Предполагается, что производные согласно изобретению должны быть получены из фрагмента ДНК, кодирующего производное пуллуланазы, причем функциональная активность экспрессированного производного пуллуланазы должна сохраняться. Далее, производные включают химические модификации, которые изменяют свойства пуллуланазы.

Обычно производное пуллуланазы будет иметь идентичность аминокислотной последовательности по меньшей мере приблизительно 50, 70 или 85%, предпочтительно по меньшей мере приблизительно 85%, предпочтительнее по меньшей мере приблизительно 90%, еще предпочтительнее по меньшей мере приблизительно 95% и еще более предпочтительно по меньшей мере 98%. Предпочтительно, чтобы любые замещения аминокислот представляли собой "консервативные замещения аминокислот" с использованием L-аминокислот, причем одна аминокислота замещается другой биологически сходной аминокислотой.

Консервативными замещениями аминокислот являются такие замещения, которые позволяют сохранить общий заряд, гидрофобность/гидрофильность и/или стерический объем замещенной аминокислоты. Примерами консервативных замещений являются указанные ниже замещения в следующих группах: Gly/Ala, Val/Ile/Leu, Lys/Arg, Asn/Gln, Glu/Asp, Ser/Cys/Thr и Phe/Trp/Tyr. Производное может, например, отличаться только по 1-10 аминокислотным остаткам, в частности по 6-10, только по 5, только по 4, 3 или даже по 1 аминокислотному остатку.

Приведенная здесь табл. 1 иллюстрирует типичные замещения аминокислот, общепризнанные в данной области техники.

При использовании здесь терминологическое определение "природная последовательность" пуллуланазы или последовательность пуллуланазы "дикого типа" включает полипептид, имеющий ту же аминокислотную последовательность, что и пуллуланаза, извлеченная из природного родительского штамма,

или ту же аминокислотную последовательность, что и пуллуланаза, из которой была изготовлена модифицированная или извлеченная пуллуланаза, например пуллуланаза, экспрессированная родительским штаммом B. deramificans (BMP 139) в соответствии с изобретением. Такая природная последовательность пуллуланазы может быть выделена из природного источника либо получена рекомбинантными или синтетическими способами. В одном из вариантов реализации изобретения термин "природная последовательность" или "дикий тип" пуллуланазы относится к пептиду пуллуланазы, из которого были получены варианты пуллуланазы, соответствующие изобретению (этот полипептид показан на фиг. 7b как SEQ ID NO: 2).

При использовании здесь понятие "процентный показатель (%) идентичности последовательности" по отношению к указанным здесь аминокислотным и нуклеотидным последовательностям определяется как процентная доля аминокислотных остатков или нуклеотидов в последовательности-кандидате, идентичных аминокислотным остаткам или нуклеотидам в последовательности пуллуланазы, после выравнивания этих последовательностей и, при необходимости, введения гэпов для достижения максимальной процентной идентичности последовательностей, причем любые консервативные замещения не рассматриваются как часть идентичности последовательности. Способы осуществления выравнивания последовательности и определения идентичности последовательности, известные компетентным специалистам, могут быть реализованы без излишнего экспериментирования, а вычисления показателей идентичности могут быть проведены с достаточной определенностью результатов; см., например, Ausubel, et al., eds. (1995) Current Protocols in Molecular Biology, Chapter 19 (Greene Publishing and Wiley-Interscience, New York) и программу ALIGN (Dayhoff, 1978) в ссылке Atlas of Protein Sequence and Structure 5:Suppl. 3 (National Biomedical Research Foundation, Washington, D.C.). Существует множество алгоритмов для выравнивания последовательностей и определения идентичности последовательности, которые включают, например, алгоритм выравнивания гомологии Needleman et al. (1970) J. Mol. Biol. 48:443, алгоритм локальной гомологии Smith, et al. (1981) Adv. Appl. Math. 2:482, метод поиска на сходство Pearson, et al. (1988) Proc. Natl. Acad. Sci. 85:2444, алгоритм Smith-Waterman (Meth. Mol. Biol. 70:173-187 (1997), а также алгоритмы BLASTP, BLASTN и BLASTX (см. Altschul, et al. (1990) J. Mol. Biol. 215:403-410). Доступны также компьютерные программы, основанные на этих алгоритмах, которые включают, но, не ограничиваясь ими: программные продукты ALIGN или Megalign (DNASTAR), или WU-BLAST-2 (Altschul, et al., Meth. Enzym., 266:460-480 (1996)), или GAP, BESTFIT, BLAST (Altschul, et al.), выше, FASTA и TFASTA, доступные в пакете программ Genetics Computing Group (GCG), версия 8, Мэдисон, штат Висконсин, США, а также CLUSTAL в пакете программ PC/Gene (Intelligenetics, Маунтин-Вью штат Калифорния). Компетентные специалисты могут определить соответствующие параметры для измерения выравнивания, включая алгоритмы, необходимые для достижения максимального выравнивания по всей длине сравниваемых последовательностей. Предпочтительно, чтобы идентичность последовательности была определена с применением параметров по умолчанию, заданных в программе. Более точно, идентичность последовательности можно определить по поисковому алгоритму гомологии Smith-Waterman (Meth. Mol. Biol. 70:173-187 (1997)), реализованному в программе MSPRCH (Oxford Molecular), с применением поиска по аффинным гэпам со следующими поисковыми параметрами: балльная оценка пени открытия гэпа - 12 и балльная оценка пени расширения гэпа - 1. Предпочтительно, чтобы сравнения парных аминокислот проводились с использованием программы GAP из пакета программ анализа последовательностей GCG (Genetics Computer Group, Inc., Мэдисон, штат Висконсин), а также с задействованием матрицы замещений аминокислот blosum62 при весе гэпа 12 и весе длины 2. Что касается оптимального выравнивания двух последовательностей аминокислот, то смежный сегмент вариантной последовательности аминокислот может иметь дополнительные аминокислотные остатки или делеции аминокислотных остатков по отношению к эталонной последовательности аминокислот. Смежный сегмент, используемый для сравнения с эталонно аминокислотной последовательностью, должен включать по меньшей мере 20 смежных аминокислотных остатков и может состоять из 30, 40, 50 или более аминокислотных остатков. Поправки на увеличенную идентичность последовательности, связанную с включением гэпов в аминокислотные последовательности производных, можно сделать за счет присвоения пени по гэпам.

При использовании здесь термин "экспрессирующая конструкция" (или "экспрессирующий вектор") означает конструкцию ДНК, включая последовательность ДНК, которая функционально связана с подходящей управляющей последовательностью, способной влиять на экспрессию ДНК в подходящем хозяине. Такие управляющие последовательности могут включать промотор для воздействия на транскрипцию, необязательную последовательность оператора для управления транскрипцией, последовательность, кодирующую соответствующие сайты связывания рибосом на мРНК, и последовательности, которые управляют окончанием транскрипции и трансляции. Изобретение не ограничивается применением какой-либо конкретной экспрессирующей конструкции. Для разных типов клеток предпочтительно использовать разные экспрессирующие векторы. Например, предпочтительным промотором для векторов, используемых в Bacillus subtilis, является промотор AprE, предпочтительным промотором для векторов, используемых в Bacillus deramificans, является промотор amyL, предпочтительным промотором, используемым в E. coli, является промотор Lac, предпочтительным промотором, используемым в Saccharomyces cerevisiae, является промотор PGK1, предпочтительным промотором, используемым в Asper-

gillus niger, является glaA, а предпочтительным промотором для Trichoderma reesei является cbhI. Вектором может быть плазмида, фаговая частица или просто потенциальный геномный инсерт (вставка).

После трансформации (или трансфекции) в подходящего хозяина экспрессирующая конструкция может реплицироваться и функционировать независимо от генома хозяина или при соответствующих условиях встраиваться в геном хозяина. В описании изобретения термины "плазмида", "вектор" и "экспрессирующая конструкция (конструкции)" иногда употребляются взаимозаменяемо. Однако изобретение рассчитано на то, чтобы включать другие формы экспрессирующих векторов, реализующих те же функции, которые уже известны или будут известны в данной области техники. Полезные для изобретения экспрессирующие векторы, не упоминаемые в каких-либо других местах этого описания изобретения, могут состоять, например, из участков последовательностей хромосомной, нехромосомной и синтетической ДНК, таких как различные известные производные SV40 и известные бактериальные плазмиды, например плазмиды из E. coli, включая col E1, pCR1, pBR322, pMb9, pUC 19 и их производные, плазмиды с более широким кругом хозяев, например RP4, фаговые ДНК, например многочисленные производные фага λ, например NM989, другие ДНК-содержащие фаги, например M13 и волокнистые фаги с однонитевой ДНК, дрожжевые плазмиды, например плазмида 2υ или ее производные, векторы, полезные в эукариотических клетках, в частности векторы, полезные в клетках животных, и векторы, полученные из комбинаций плазмид и фаговой ДНК, например плазмиды, которые были модифицированы для применения фаговой ДНК или других последовательностей, управляющих экспрессией.

Методики экспрессии с применением экспрессирующих векторов в соответствии с изобретением известны в данной области техники и в целом описаны, например, в ссылке Sambrook, et al., MOLECULAR CLONING: A LABORATORY MANUAL, SECOND EDITION, Cold Spring Harbor Press (1989). Часто такие экспрессирующие векторы, включая последовательности ДНК, рассматриваемые в изобретении, трансформируются в одноклеточного хозяина посредством прямой вставки в геном определенных биологических видов, осуществляемой через событие интеграции (см., например, Bennett & Lasure, MORE GENE MANIPULATIONS IN FUNGI, Academic Press, San Diego, p. 70-76 (1991) и статьи, цитируемые в этой ссылке, где описана целенаправленная геномная вставка в грибковых хозяев).

Термины "функционально связанный", "в функциональной комбинации" и "в функциональном расположении" при использовании здесь относятся к сцеплению последовательностей нуклеиновой кислоты, которое позволяет им осуществлять свою целевую функцию. Например, функциональная связь промоторной последовательности с представляющей интерес нуклеотидной последовательностью означает сцепление промоторной последовательности и представляющей интерес нуклеотидной последовательности таким образом, что промоторная последовательность способна управлять транскрипцией представляющей интерес нуклеотидной последовательности и/или синтезом полипептида, кодируемого представляющей интерес нуклеотидной последовательностью. Сходным образом, функциональная связь последовательности нуклеиновой кислоты, обладающей повозрастной регуляторной активностью, с промоторной последовательностью и с представляющей интерес нуклеотидной последовательностью означает сцепление последовательности нуклеиновой кислоты, обладающей повозрастной регуляторной активностью, промоторной последовательности и представляющей интерес нуклеотидной последовательности таким образом, что последовательность нуклеиновой кислоты, обладающая повозрастной регуляторной активностью, способна за определенный период времени изменять уровень транскрипции в мРНК представляющей интерес нуклеотидной последовательности и/или синтез полипептида, кодируемого представляющей интерес нуклеотидной последовательностью.

При использовании здесь термины "организм-хозяин", "штамм-хозяин" или "клетка-хозяин" означают подходящего хозяина для экспрессирующего вектора, включая ДНК, соответствующую изобретению. Клетки-хозяева, полезные в изобретении, в целом, представляют собой прокариотических или эукариотических хозяев, включая любой трансформируемый микроорганизм, в котором можно осуществить экспрессию. Например, в контексте изобретения штаммами-хозяевами могут быть Bacillus subtilis, Bacillus deramificans (или другие виды Bacillus), Escherichia coli, Trichoderma reesei, Saccharomyces cerevisiae или Aspergillus niger. Клетки-хозяева трансформируются или трансфицируются векторами, сконструированными на основе методик рекомбинантной ДНК. Такие трансформированные клетки-хозяева могут быть способны реплицировать векторы, кодирующие пуллуланазу и ее производные (варианты или мутанты), и/или экспрессировать желательный пептидный продукт.

Термины "культура" или "условия культивирования" при использовании в контексте выращивания популяции организмов-хозяев относятся к сосуду для культивирования, к культуральной среде и к условиям культивирования, которые пригодны для выращивания организма-хозяина, а в том случае, если организм-хозяин трансфицирован нуклеотидными последовательностями, рассматриваемыми в изобретении, или их вариантами, эти термины относятся к выработке пуллуланаз, предлагаемых изобретением. Изобретение не ограничивается какой-либо конкретной культурой или условиями культивирования до тех пор, пока выполняется указанное выше.

При использовании здесь термины "функционально присоединенный" или "функционально связанный" означают, что регуляторный регион, например промотор, терминатор, сигнал секреции или энхан-

сер присоединен к структурному гену или сцеплен с ним и управляет экспрессией этого гена.

При использовании здесь вещество (например, полинуклеотид или белок) "извлеченное из" микроорганизма, означает, что вещество от природы присуще этому микроорганизму.

Термины "Trichoderma" или "Trichoderma sp." относятся к любым грибковым штаммам, которые ранее были классифицированы как Trichoderma или в настоящее время классифицируются как Trichoderma. Предпочтительными видами этого рода микроорганизмов являются Trichoderma longibrachiatum, Trichoderma reesei или Trichoderma viride. В изобретении различные виды Trichoderma могут быть использованы как организмы-хозяева.

Как описано здесь, в одном из аспектов отличительным признаком изобретения является "по существу чистая" (или рекомбинантная) нуклеиновая кислота, которая включает нуклеотидную последовательность, кодирующую полипептид пуллуланазы, и/или эквиваленты таких нуклеиновых кислот. Термин нуклеиновая кислота при использовании здесь может включать фрагменты и эквиваленты. Термин "эквивалент" относится к нуклеотидным последовательностям, кодирующим функционально равноценные полипептиды или функционально равноценные белки. Эквивалентные нуклеотидные последовательности включают последовательности, которые отличаются по одному или более замещений, вставок или делеций нуклеотидов (в частности, речь может идти об аллельных вариантах), а также включают последовательности, которые отличаются от нуклеотидной последовательности пуллуланазы, представленной в SEQ ID NO: 1, вследствие вырожденности генетического кода.

Практика изобретения использует, если не указано иное, традиционные методики клеточной биологии, культивирования клеток, молекулярной биологии, трансгенной биологии, микробиологии, рекомбинантной ДНК и иммунологии, которые известны компетентным специалистам. Эти методики описаны в литературе. Все цитируемые здесь публикации, патенты и патентные заявки включены сюда в качестве ссылок во всей полноте и с учетом всех целей; см., например, Molecular Cloning A Laboratory Manual, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); DNA Cloning, Volumes I and II (D.N. Glover ed., 1985); Oligonucleotide Synthesis (M.J. Gait ed., 1984); Mullis et al. патент США № 4683195; Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription And Translation (B.D. Hames & S.J. Higgins eds. 1984); Culture Of Animal Cells (R.I. Freshney, Alan R. Liss, Inc., 1987); Immobilized Cells And Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide To Molecular Cloning (1984); трактат Methods In Enzymology (Academic Press, Inc., N.Y.); Gene Transfer Vectors For Mammalian Cells (J.H. Miller and M.P. Calos eds., 1987, Cold Spring Harbor Laboratory); Methods In Enzymology, Vols. 154&155 (Wu, et al., eds.), Immunochemical Methods In Cell And Molecular Biology (Mayer and Walker, eds., Academic Press, London, 1987); Handbook Of Experimental Immunology, Volumes I-IV (D.M. Weir and C.C. Blackwell, eds., 1986); Manipulating the Mouse Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Кроме того, информацию о способах приготовления, экспрессии, выделения и применения протеаз можно получить из обзора патента США № 6768001, который включен сюда в качестве ссылки во всей полноте.

Другие отличительные признаки и полезные эффекты изобретения будут очевидны из приведенных ниже подробного описания и формулы изобретения.

Подробное описание изобретения

Теперь изобретение будет описано более подробно на основе ссылок с использованием только следующих определения и примеров. Все патенты и публикации, включая все последовательности, раскрытые в этих патентах и публикациях, имеющих отношение к данной заявке, определенно включены сюда в качестве ссылок.

Если в тексте этой заявки не дано другого определения, все упоминаемые здесь технические и научные термины имеют то же значение, которое считается общепринятым среди специалистов с обычным уровнем компетенции в той области техники, к которой относится изобретение. Такие публикации, как Singleton, et al., DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2D ED., John Wiley and Sons, New York (1994), and Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, NY (1991) дают для специалистов общий словарь многочисленных терминов, использованных в описании этого изобретения. Хотя в практике и при проверке изобретения могут быть использованы любые способы и материалы, аналогичные или равноценные описанным здесь, в данной заявке описаны предпочтительные способы и материалы. Числовые диапазоны включают значения, определяющие границы диапазона. Если в тексте не указано иное, нуклеиновые кислоты записаны слева направо в направлении от 5'-конца до 3'-конца, а аминокислотные последовательности записаны слева направо в направлении от амино- до карбокси- соответственно. Специалистов-практиков мы особенно отсылаем к Sambrook et al., 1989 и Ausubel F.M. et al., 1993 для определений и терминов, употребляемых в данной области техники. Необходимо понимать, что это изобретение не ограничивается какой-либо конкретной методологией, протоколами и описанными здесь реактивами, поскольку они могут варьироваться.

Числовые диапазоны включают значения, определяющие границы диапазона.

Если в тексте не указано иное, нуклеиновые кислоты записаны слева направо в направлении от 5'-конца до 3'-конца, а аминокислотные последовательности записаны слева направо в направлении от амино- до карбокси- соответственно.

Представленные в этой заявке рубрики не являются ограничениями различных аспектов или вариантов реализации изобретения, сферу действия которого можно оценить по описанию в целом. В соответствии с этим те термины, которые будут определены непосредственно ниже, более полно можно определить, рассматривая данное описание изобретения в целом.

Молекулярная биология.

В одном из вариантов реализации это изобретение предусматривает экспрессию гетерологичных генов под управлением промотора amyL. Следовательно, это изобретение основывается на общепринятых методиках в области рекомбинантной генетики. Основные руководства, раскрывающие общие методы, используемые в этом изобретении, включают Sambrook, et al., Molecular Cloning, A Laboratory Manual (2nd ed. 1989); Kriegler, Gene Transfer and Expression: A Laboratory Manual (1990) и Ausubel, et al., eds., Current Protocols in Molecular Biology (1994).

Гетерологичные гены, включающие промоторные последовательности гена целлюлазы мицелиальных грибов (гифомицетов), типично клонируются в промежуточные векторы перед трансформацией в клетки Trichoderma reesei для репликации и/или экспрессии. Типичными промежуточными векторами являются прокариотические векторы, например плазмиды или челночные векторы.

Для того чтобы получить высокий уровень экспрессии клонированного гена, предпочтительно расположить гетерологичный ген приблизительно на том же расстоянии от промотора, какое имеется в природном гене целлюлазы. Однако, как известно в данной области техники, вполне допустимы некоторые вариации этого расстояния без потери функции промотора.

Компетентные специалисты знают, что природный промотор можно модифицировать без изменения его функции перестановкой, замещением, добавлением или удалением одного или более нуклеотидов. Практика изобретения охватывает такие изменения промотора, не ограничивающие его функцию.

Типичный экспрессирующий вектор/конструкция содержит блок транскрипции или экспрессирующий кластер, который, в свою очередь, содержит все дополнительные элементы, необходимые для экспрессии гетерологичной последовательности. Таким образом, типичный экспрессирующий кластер содержит промотор, функционально связанный с гетерологичной последовательностью нуклеиновой кислоты, и сигналы, необходимые для эффективного полиаденилирования транскрипта, сайты связывания рибосом и терминатор трансляции. Дополнительные элементы кластера могут включать энхансеры и, если в качестве структурного гена используется геномная ДНК, интроны с функциональными донорскими и акцепторными сайтами сплайсинга.

Практика изобретения не имеет ограничений, связанных с выбором промотора в генетической конструкции. Однако образцовыми промоторами являются такие промоторы Trichoderma reesei, как cbh1, cbh2, eg1, eg2, eg3, eg5, xln1 и xln2.

В дополнение к промоторной последовательности экспрессирующий кластер также должен содержать регион терминации транскрипции, расположенный ниже структурного гена и обеспечивающий эффективное окончание считывания. Регион терминации можно получить из того же гена, откуда взята промоторная последовательность, или из других генов.

Хотя в изобретении функциональным может любой грибковый терминатор, предпочтительные терминаторы включают терминатор из гена trpC Aspergillus nidulans (Yelton, M. et al. (1984) PNAS USA 81:1470-1474, Mullaney, E.J. et al. (1985) MGG 199:37-45), из генов глюкоамилазы Aspergillus awamori или Aspergillus niger (Nunberg, J.H. et al. (1984) Mol. Cell Biol. 4:2306, Boel, E. et al. (1984) EMBO J. 3:1581-1585) и из гена карбоксилпротеазы Mucor miehei (EPO Publication No. 0215594).

Конкретный экспрессирующий вектор, используемый для транспортировки генетической информации в клетку, не имеет принципиального значения. Можно использовать любой общепринятый вектор из числа применяемых для экспрессии в эукариотических или прокариотических клетках. Стандартные бактериальные экспрессирующие векторы включают бактериофаги λ и M13, плазмиды, например, на основе pBR322, pSKF, pET23D, а также такие гибридные экспрессирующие системы, как MBP, GST и LacZ. К рекомбинантным белкам можно также добавить хвосты эпитопов, чтобы создать условия для удобных способов выделения, например, c-myc.

Типичные элементы, вставляемые в экспрессирующие векторы, также включают репликон, ген, кодирующий устойчивость к антибиотикам для осуществления отбора бактерий, содержащих рекомбинантные плазмиды, а также уникальные сайты рестрикции в несущественных регионах плазмид, позволяющие осуществлять вставки гетерологичных последовательностей. Выбор конкретного гена устойчивости к антибиотикам не имеет принципиального значения, и с этой целью можно применять любой из множества генов устойчивости, известных в данной области техники. Предпочтительно прокариотические последовательности выбирают таким образом, чтобы они не мешали репликации или встраиванию ДНК в Trichoderma reesei.

Способы трансформации, используемые в изобретении, могут приводить к стабильному встраиванию всего вектора трансформации или его части в геном гифомицетов. Однако также рассматривается возможное применение трансформации, приводящей к поддержанию самостоятельной реплицирующегося внехромосомного вектора трансформации.

Для продуцирования клеточных линий Trichoderma reesei, экспрессирующих гетерологичный белок

в больших количествах, можно использовать множество стандартных способов трансфекции. Некоторые из опубликованных в литературе способов введения конструкций ДНК в штаммы Trichoderma включают методики, описанные Lorito, Hayes, DiPietro and Harman, 1993, Curr. Genet. 24: 349-356; Goldman, Van-Montagu and Herrera-Estrella, 1990, Curr. Genet. 17:169-174; Penttila, Nevalainen, Ratto, Salminen and Knowles, 1987, Gene 6: 155-164, для Aspergillus, Yelton, Hamer and Timberlake, 1984, Proc. Natl. Acad. Sci. USA 81: 1470-1474, для Fusarium Bajar, Podila and Kolattukudy, 1991, Proc. Natl. Acad. Sci. USA 88: 8202-8212, для Streptomyces Hopwood et al., 1985, The John Innes Foundation, Norwich, UK и для Bacillus Brigidi, DeRossi, Bertarini, Riccardi and Matteuzzi, 1990, FEMS Microbiol. Lett. 55: 135-138.

Однако для введения чужеродной нуклеотидной последовательности в клетку-хозяина можно использовать любую из хорошо известных процедур. К их числу относятся применение трансфекции с фосфатом кальция, полибрен, слияние протопластов, электропорация, биолистика, липосомы, микроинъекция, плазматические векторы, вирусные векторы и любой из других хорошо известных способов введения клонированной геномной ДНК, кДНК, синтетической ДНК или другого чужеродного генетического материала в клетку-хозяина (см., например, Sambrook et al., выше). Кроме того, применим опосредованный Agrobacterium способ трансфекции, описанный в патенте США № 6255115. Необходимо только, чтобы конкретная генно-инженерная процедура, которая используется для переноса генетического материала, была способна успешно ввести по меньшей мере один ген в клетку-хозяина, которая, в свою очередь, способна экспрессировать гетерологичный ген.

Изобретение также относится к пуллуланазе, гетерологично выработанной микроорганизмом. Примерами подходящих бактерий являются такие грамположительные бактерии, как Bacillus subtilis, Bacillus licheniformis, Bacillus lentus, Bacillus brevis, Bacillus stearothermophilus, Bacillus alkalophilus, Bacillus amyloliquefaciens, Bacillus coagulans, Bacillus circulans, Bacillus lautus, Bacillus megaterium, Bacillus thuringiensis, или Streptomyces lividans, или Streptomyces murinus, а также такие грамотрицательные бактерии, как E. coli. Трансформацию бактерий можно осуществить, например, посредством трансформации протопластов или с применением компетентных клеток таким образом, который известен сам по себе.

В целом изобретение заключает в себе способ продуцирования пуллуланазы за счет экспрессии ДНК, встроенной в экспрессирующую систему, которая была трансформирована в клетку-хозяина. Для экспрессирования последовательностей ДНК в изобретении можно задействовать широкий ряд комбинаций хозяин/экспрессирующий вектор. Для приобретения доступны многие прокариотические и эукариотические экспрессирующие векторы. Выбор соответствующих экспрессирующих векторов находится в пределах знаний компетентных специалистов. Вектором может быть плазмида, фаговая частица или просто потенциальный геномной инсерт. После трансформации в подходящего хозяина вектор может реплицироваться и функционировать независимо от генома хозяина, но в некоторых случаях может встраиваться в геном хозяина. В описании изобретения термины плазмида и вектор иногда употребляются взаимозаменяемо, поскольку плазмида в настоящее время является наиболее частой используемой формой вектора. Однако изобретение направлено на то, чтобы включать и другие формы экспрессирующих векторов, которые способны выполнять равноценные функции и которые уже известны или будут известны в данной области техники. Например, полезные экспрессирующие векторы включают сегменты хромосомной, нехромосомной и синтетической ДНК, такие как различные известные плазмиды и фаги, полезные для применения в этих целях. В дополнение к этому, в указанных векторах может применяться любая последовательность из широкого ряда управляющих последовательностей экспрессии.

Клетки-хозяева, полезные в изобретении, обычно относятся к прокариотическим или эукариотическим хозяевам, включая любой трансформируемый микроорганизм, в котором можно вызвать экспрессию пуллуланазы, предлагаемую изобретением. Клетки-хозяева трансформируются или трансфицируются векторами, сконструированными на основе методик рекомбинантной ДНК. Такие трансформированные клетки-хозяева способны или реплицировать векторы, кодирующие пуллуланазу и ее варианты (мутанты), или экспрессировать желательную пуллуланазу. К числу таких хозяев относятся хорошо известные эукариотические и прокариотические хозяева, например штаммы E. coli, Pseudomonas, Bacillus, Streptomyces, различные грибки, дрожжи и животные клетки. Предпочтительно, чтобы в соответствии с изобретением хозяин экспрессировал пуллуланазу внеклеточно, что облегчает ее очистку и последующий процессинг.

В некоторых вариантах реализации изобретения клетка-хозяин принадлежит к роду Bacillus, тогда как в некоторых вариантах реализации изобретения представляющий интерес штамм Bacillus является промышленным штаммом Bacillus. Примеры промышленных штаммов Bacillus включают, но не ограничиваясь ими, B. licheniformis, B. subtilis, B. lentus, B. amyloliquefaciens. В дополнительных вариантах реализации изобретения штамм-хозяин Bacillus выбирают из группы, состоящей из B. lentus, B. brevis, B. stearothermophilus, B. alkalophilus, B. coagulans, B. cirulans, B. pumilus, B. thuringiensis, B. clausii и B. megaterium, а также из других микроорганизмов, принадлежащих к роду Bacillus, как это обсуждалось выше. В некоторых вариантах реализации изобретения используется B. subtilis. В других вариантах реализации изобретения используется B. licheniformis. Например, патенты США №№ 5264366 и 4760025 (RE34.606), а также документ US 2002/0182734 (международная публикация № WO 02/14490) описывают различные штаммы-хозяева рода Bacillus, которые могут найти применение в изобретении, хотя с этой

целью могут рассматриваться и другие подходящие штаммы. Предпочтительно использовать протеазонегативный штамм Bacillus (например, с делецией генов Δapr или Δnpr среди прочих).

Известны различные способы, применяемые для трансформации различных видов Bacillus. Действительно, хорошо известны способы изменения хромосомы Bacillus, связанные с плазмидными конструкциями и трансформацией плазмид в E. coli. В большинстве способов плазмиды впоследствии выделяют из E. coli и трансформируют в Bacillus. Однако использование такого промежуточного микроорганизма как E. coli не является принципиально необходимым, и в некоторых вариантах реализации изобретения конструкцию ДНК напрямую трансформируют в компетентного хозяина Bacillus через протопласты или трансформацию компетентной клетки. Экспрессию и очистку мутантной пуллуланазы в соответствии с изобретением можно осуществить на основе известных в данной области техники средств для выполнения таких процессов.

После введения экспрессирующего вектора в клетки эти трансфицированные клетки культивируют в таких условиях, которые создают преимущества для экспрессии генов под управлением промоторных последовательностей гена протеазы. Большие партии трансформированных клеток можно культивировать так, как это описано в примерах, см. ниже. Наконец, продукт восстанавливают из культуры, применяя стандартные методики.

Таким образом, описанное здесь изобретение предлагает экспрессию и усиленную секрецию желательных полипептидов, экспрессия которых происходит под управлением промоторных генных последовательностей, включая встречающиеся в природе гены амилазы, гибридных последовательностей ДНК и разных гетерологичных конструкций. Изобретение также предлагает высокопроизводительные процессы для осуществления экспрессии и секреции указанных желательных полипептидов.

Экспрессия белка.

Белки, предлагаемые изобретением, вырабатываются культивируемыми клетками, которые были трансформированы экспрессирующим вектором, содержащим гены, экспрессия которых находится под управлением промоторных последовательностей гена амилазы. Изобретение особенно полезно для улучшения внутриклеточной и/или внеклеточной выработки белков. Белок может быть гомологичным или гетерологичным. Белки, которые могут вырабатываться в соответствии с изобретением, включают, но не ограничиваясь ими, гормоны, ферменты, факторы роста, цитокины, антитела и т.п.

Ферменты включают, но не ограничиваясь ими, гидролазы, например протеазу, эстеразу, липазу, фенолоксидазу, пермеазу, амилазу, пуллуланазу, целлюлазу, глюкозоизомеразу, лактазу и белковую дисульфидизомеразу.

Условия, подходящие для экспрессии указанных генов, включают доставку в культуру индуцирующей питательной композиции (см., например, документ US-2004-0121446).

Оптимальные условия для выработки белков могут варьироваться в зависимости от выбора клетки-хозяина и от выбора белка, который должен экспрессироваться. Такие условия может без проблем определить компетентный специалист на основе обычных экспериментальных приемов или общепринятой оптимизации.

Представляющий интерес белок, например пуллуланазу, как это описано здесь, в типичном случае очищают или выделяют после экспрессии. Представляющий интерес белок можно выделить или очистить разными способами, известными компетентному специалисту, в зависимости от того, какие компоненты представлены в образце. Стандартные способы очистки включают электрофоретические, молекулярные, иммунологические и хроматографические методики, в том числе ионообменную, гидрофобную, аффинную и обращенно-фазовую хроматографию (типа HPLC), а также хроматофокусирование. Например, представляющий интерес белок можно очистить, применяя стандартную колонку с антителами против представляющего интерес белка. Также полезны методики ультрафильтрации и диафильтрации в сочетании с повышением концентрации белка. Общие руководящие указания по подходящим методикам очистки см. в ссылке Scopes, Protein Purification (1982). Необходимая степень очистки будет варьироваться в зависимости от представляющего интерес белка. В некоторых случаях очистка вообще не требуется.

Аналоги пуллуланазы, рассматриваемые в изобретении.

Аналоги могут отличаться от родительской пуллуланазы "дикого типа" или от природных образцов пуллуланазы по аминокислотной последовательности и/или по другим признакам, не связанным с аминокислотной последовательностью. Модификации не связанные с последовательностью, включают образование химических производных пуллуланазы in vivo или in vitro. Модификации, не связанные с последовательностью, включают изменения в ацетилировании, метилировании, фосфорилировании, карбоксилировании или гликозилировании.

Предпочтительные аналоги включают пуллуланазу (или ее биологически активные фрагменты), последовательности которых отличаются от последовательности "дикого типа" или от природной последовательности по одному или более консервативных замещений аминокислот или по одному или более неконсервативных замещений аминокислот, делеций или вставок, которые не аннулируют биологическую активность пуллуланазы. Типичные консервативные замещения включают замещение одной аминокислоты на другую аминокислоту, обладающую аналогичными характеристиками, например замены в

пределах следующих групп: валин-глицин, глицин-аланин, валин-изолейцин-лейцин, аспарагиновая кислота-глутаминовая кислота, аспарагин-глутамин, серин-треонин, лизин-аргинин, а также фенилаланин-тирозин.

Консервативные замещения могут быть осуществлены, например, в соответствии в приведенной ниже таблицей, которая описывает общепринятую группировку аминокислот в соответствии с диаграммой Венна.

Таблица 1

Группа		Подгруппа	
Гидрофобные	FWYHKMILVAGC	Ароматические	FWYH
		Алифатические	ILV
Полярные	WYHKREDCSTNQ	Заряженные	HKRED
		С положительным зарядом	HKR
		С отрицательным зарядом	ED
Мелкие	VCAGSPTND	Миниатюрные	AGS

Другие консервативные замещения могут быть определены по приведенной ниже таблице.

Консервативные замещения аминокислот

Аминокислота	Код	Замещается любой из следующих аминокислот
Аланин	A	D-Ala, Gly, бета-Ala, L-Cys, D-Cys
Аргинин	R	D-Arg, Lys, D-Lys, гомо-Arg, D-гомо-Arg, Met, He, D-Met, D-Ile, Orn, D-Ora
Аспарагин	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gin, D-Gln
Аспарагиновая кислота	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gin, D-Gln
Цистеин	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Глутамин	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Глутаминовая кислота	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gin, D-Gln
Глицин	G	Ala, D-Ala, Pro, D-Pro, b-Ala Acp
Изолейцин	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Лейцин	L	D-Leu, Val, D-Val, lie, D-Ile, Met, D-Met
Лизин	K	D-Lys, Arg, D-Arg, гомо-Arg, D-гомо-Arg, Met, D-Met, He, D-Ile, Om, D-Orn
Метионин	M	D-Met, S-Me-Cys, lie, D-Ие, Leu, D-Leu, Val, D-Val
Фенилаланин	F	D-Phe, Tyr, D-Tyr, L-Dopa, His, D-His, Trp, D-Trp, транс-3,4, или 5-фенилпролин, цис-3,4, или 5-фенилпролин
Пролин	P	D-Pro, L-I-тиазолидин-4-carboxylic acid, D- или L-1-оксазолидин-4-карбоксильная кислота
Серин	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Треонин	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Тирозин	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Валин	V	D-Val, Leu, D-Leu, He, D-Ile, Met, D-Met

Другие аналоги в рамках изобретения представляют собой модификации, которые увеличивают стабильность пептида. Такие аналоги могут, например, содержать в пептидной последовательности одну или более непептидных связей (замещающих пептидные связи). Сюда также включены аналоги, в составе которых имеются остатки, отличающиеся от природных L-аминокислот, например D-аминокислоты, или не встречающиеся в природе, или синтетические аминокислоты, например α или β-аминокислотные аналоги и циклические аналоги.

Другие варианты реализации изобретения

В изобретение включены аллельные вариации, природные мутанты, индуцированные мутанты, белки, кодируемые ДНК, которая гибридизируется в строгих или нестрогих условиях с нуклеиновой кислотой, кодирующей полипептид с SEQ ID NO: 1 (определение строгих и нестрогих условий см. в публикации Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1-6.3.6, которая включена сюда в качестве ссылки), а также полипептиды, специфически связывающиеся с пуллуланазой посредством иммунных сывороток, особенно иммунных сывороток к активному сайту или домену связывания пуллуланазы.

Нуклеотидные кислоты и полипептиды, рассматриваемые в изобретении, включают те разновидности, которые отличаются от раскрытых здесь последовательностей в силу ошибок секвенирования в раскрытых последовательностях.

Изобретение также включает фрагменты, предпочтительно биологически активные фрагменты или аналоги пуллуланазы. Биологически активный фрагмент или аналог проявляет in vivo или in vitro активность, характерную для пуллуланазы, описанной в SEQ ID NO: 2, 4 и 6, или для других природных пуллуланаз, например, один или более видов биологической активности, описанных здесь. Особенно пред-

почтительны фрагменты, которые существуют in vivo, например фрагменты, которые возникают в ходе посттранскрипционного процессинга или трансляции альтернативно сплайсированных РНК. Фрагменты включают те варианты, которые экспрессируются в нативных или эндогенных клетках, например, в результате посттрансляционного процессинга, например в результате удаления аминоконцевой сигнальной последовательности, а также те варианты, которые получены в экспрессирующих системах, например в клетках CHO. Особенно предпочтительны такие фрагменты, например активные фрагменты, которые порождаются в результате протеолитического расщепления или событий альтернативного сплайсинга. Поскольку такие пептиды, как пуллуланаза проявляют ряд физиологических свойств, и поскольку такие свойства могут быть приписаны разным частям молекулы, полезный фрагмент пуллуланазы или аналог пуллуланазы - это такой фрагмент или аналог, который проявляет биологическую активность в любом биологическом анализе на активность пуллуланазы. Наиболее предпочтительно, чтобы фрагмент или аналог обладал 10, 40, 60, 70, 80% или по меньшей мере 90% активности пуллуланазы (SEQ ID NO: 2, 4 и 6) в любом анализе in vivo или in vitro на активность пуллуланазы. Один из способов изготовления таких аналогов пуллуланазы включает синтез аналогов пуллуланазы через прямую молекулярную эволюцию, как это обсуждается ниже.

Фрагменты пуллуланазы можно генерировать способами, известными компетентному специалисту. Способность фрагмента-кандидата проявлять биологическую активность пуллуланазы можно оценивать способами, известными компетентному специалисту, как это описано здесь. В изобретение также включены пептиды пуллуланазы, содержащие остатки, которые не требуются для биологической активности пептида или появляются в результате событий альтернативного сплайсинга мРНК либо альтернативного процессинга белка.

Для того чтобы получить пептид пуллуланазы, ДНК, кодирующую пуллуланазу, вводят в экспрессирующий вектор, вектор - в клетку, пригодную для экспрессии желательного белка, а пептид восстанавливают и очищают способами, ранее известными в данной области техники. Антитела к пептидам и белкам можно получить посредством иммунизации животного, например кролика или мыши, причем восстановление антител против пуллуланазы также проводят с применением способов, ранее известных в данной области техники.

Возможности промышленного применения изобретения

Изобретение имеет множество практических применений в промышленности, как это рассматривается здесь, однако данное описание следует воспринимать как иллюстративное, но не исчерпывающее.

В некоторых вариантах реализации изобретение имеет предполагаемое применение в производстве этилового спирта, выпечке, производстве фруктовых соков, пивоварении, перегонке, виноделии, выделке кожи, производстве масел и жиров, бумаги и пульпы, а также кормов для животных.

В других вариантах реализации изобретение имеет предполагаемое применение в активном "биологическом" компоненте детергентов и материалов для очистки. В таких случаях для расщепления белка, крахмала и удаления жирных пятен применяют протеазы, амилазы и липазы. Варианты реализации изобретения включают проверку совместимости ферментов с ингредиентами детергента путем проведения исследований стабильности и тестирования указанных продуктов в разных составных рецептурах.

Еще в одном варианте реализации изобретение имеет предполагаемое применение в текстильной промышленности, главным образом, при завершающей обработке тканей и готовой одежды. Основные варианты применения включают расшлихтовку, устранение ограничивающих габаритов (то есть удаление жестких элементов волокон) из нитей в тканях после выработки текстиля. Процесс биополировки для уменьшения тенденции к пилингу (скатыванию) и придания тканям более гладкого и блестящего внешнего вида. Процесс биопритирки, когда небольшая доза фермента может заменить традиционную пемзу, применяемую при обработке джинсовой ткани для придания ей поношенного вида.

Еще в одном из вариантов реализации изобретение имеет предполагаемое ферментативное применение для разжижения и осахаривания крахмала с целью его превращения в глюкозу и изомеризации во фруктозу. Изобретение можно использовать для превращения больших объемов кукурузы и других зерновых в подсластители, например в кукурузный сироп с высоким содержанием фруктозы и в мальтозный сироп.

Примеры

В приведенном ниже раскрытии изобретения с экспериментальной стороны употребляются следующие сокращения: eq (эквиваленты), M (молярный), мкМ (микромолярный), N (нормальный), моль (моли), ммоль (миллимоли), мкмоль (микромоли), нмоль (наномоли), г (граммы), мг (миллиграммы), кг (килограммы), мкг (микрограммы), л (литры), мл (миллилитры), мкл (микролитры), см (сантиметры), мм (миллиметры), мкм (микрометры), нм (нанометры), °C (градусы Цельсия), ч (часы), мин (минуты), с (секунды), мс (миллисекуды), ТСХ (тонкослойная хроматография), нт (нуклеотиды), Q (глутамин), E (глутаминовая кислота), ХАФ (хлорамфеникол).

Изобретение более подробно описано в следующих примерах, которые ни в коем случае не направлены на ограничение сферы его применения в соответствии с пунктами формулы изобретения. Прилагаемые чертежи необходимо рассматривать как составную часть спецификации и описания изобретения. Все процитированные библиографические источники специфически включены сюда в качестве ссылок с

учетом всего их содержания. Следующие примеры приведены в иллюстративных целях, но не в качестве обозначения границ изобретения в соответствии с его формулой.

Пример 1. Конструирование вариантов пуллуланазы

Были сконструированы следующие варианты пуллуланазы (см. фиг. 1).

"PUL". Это "дикий тип" пуллуланазы из B. deramificans, идентичный молекуле, экспрессируемой BMP139. Ген был оптимизирован по составу кодонов, управляется промотором amyL (LAT) и имеет сигнальную последовательность amyL. Различия между этой конструкцией и конструкцией, представленной в BMP 139, таковы (см. фиг. 2). Во-первых, новая конструкция имеет кодонно-оптимизированный кодирующий регион по сравнению с природной кодирующей последовательностью в BMP 139. Во-вторых, новая конструкция имеет укороченный промоторный регион, состоящий из 100 нт по сравнению приблизительно с 800 нт в BMP 139. В-третьих, новая конструкция имеет терминатор amyL, тогда как BMP 139 имеет терминатор пуллуланазы B. deramificans. Как новая, так и старая конструкции имеют сигнальную последовательность amyL и экспрессируют идентичные молекулы пуллуланазы. Несмотря на то что молекула, экспрессируемая этой новой конструкцией, идентична имеющему хождение продукту, ее возможные выгоды связаны с производственными титрами в результате кодонной оптимизации.

"PULm104". Это пуллуланаза B. deramificans, из которой были удалены 104 N-концевых аминокислотных остатка. Конструкция заключает в себе промотор amyL, сигнальную последовательность amyL, кодонно-оптимизированный кодирующей регион пуллуланазы, в котором утрачена последовательность, кодирующая 104 N-концевые аминокислоты зрелой пуллуланазы, а также терминатор amyL. Усеченная пуллуланаза PULm 104 сходна с молекулами PULm 98 и PULm 102, вырабатываемыми при N-концевой обрезке полноразмерной пуллуланазы в положениях E99 и E103. В молекуле пуллуланазы была произведена делеция до 104-й аминокислоты, целью которой было получение идеальной целевой консенсусной последовательности для сигнальной пептидазы между сигнальной последовательностью amyL и последовательностью пуллуланазы: ASA-A. Логическое обоснование этого усеченного варианта пуллуланазы исходит из предыдущих неожиданных наблюдений, когда более высокая специфическая активность была отмечена для обрезанных вариантов пуллуланазы по сравнению с полноразмерной молекулой.

"PUL_E99Q_E103Q". Это пуллуланаза B. deramificans, в которой целевые мотивы протеазы в положениях E99 и E103 были модифицированы в Q99 и Q103 с той целью, чтобы сделать молекулу пуллуланазы устойчивой к обрезке в положениях E99 и E103. Кроме того, в том случае, если разрушение пуллуланазы после 51 ч зависит от начальной обрезки в положениях E99 и E103, можно ожидать, что такая модификация будет предупреждать разрушение и падение активности пуллуланазы после 51 ч.

Пример 2. Конструкция и трансформация плазмид.

Были синтезированы две кодонно-оптимизированные пуллуланазные конструкции: одна, кодирующая белок пуллуланазы "дикого типа", и другая, кодирующая вариант E99Q_E103Q.

Обе конструкции включали 57 нуклеотидов промотора amyL (который ранее продемонстрировал способность клонирования в E. coli, более длинные промоторные отрезки детальны), сигнальную последовательность amyL, кодонно-оптимизированную последовательность пуллуланазы (вариант) и терминатор amyL. Эти конструкции послужили шаблонами для ПЦР-генерирования трех описанных выше конструкций пуллуланазы.

"PUL". Для амплификации пуллуланазной конструкции "дикого типа" из синтетической пуллуланазной конструкции "дикого типа" были использованы следующие праймеры (сайт XhoI выделен жирным шрифтом):

Plat5-XhoI_FW: cccccgctcgaggcttttcttttggaagaaaatatagggaaaatggtacttgttaaaaattcggaatatttatacaatatcatatgtttacattgaaagggg [SEQ ID NO: 7].

Tlat-XhoI_RV: tggaatctcgaggttttatcctttaccttgtctcc [SEQ ID NO: 8].

"PULm104". Экспрессирующий кластер для усеченной пуллуланазы был генерирован посредством ПЦР со слиянием. Из синтетической пуллуланазной конструкции "дикого типа" были амплифицированы и затем слиты два следующих фрагмента:

фрагмент, охватывающий промотор amyL и сигнальную последовательность amyL;

фрагмент, охватывающий кодирующую последовательность усеченной пуллуланазы и терминатор amyL.

Дополнение A).

Для амплификации промотора и сигнальной последовательности amyL были использованы следующие праймеры:

Plat5-XhoI_FW [SEQ ID NO: 7] (see above).

ssLAT-PULm104_RV: gcgttgctgactgccggtttagcagctgctgaagctgcagaatgaggcagc [SEQ ID NO: 9] (праймер гибридизации, обратная последовательность пуллуланазы, начинающаяся с кодона 105 выделена жирным шрифтом).

Дополнение B).

Для амплификации кодирующей последовательности пуллуланазы и терминатора amyL были использованы следующие праймеры:

ssLAT-PULm104_FW: gctgcctcattctgcagcttcagcagctgctaaaccggcagtcagcaacgc [SEQ ID NO: 10] (праймер

гибридизации, последовательность пуллуланазы, начинающаяся с кодона 105 выделена жирным шрифтом).

Tlat-XhoI_RV: [SEQ ID NO: 8] (см. выше).

Каждый из праймеров гибридизации включает два отрезка последовательности, удаленных на 312 нт от последовательности-шаблона (они представляют 104 N-концевые аминокислоты). Два фрагмента ПЦР, описанные в пп.А) и В), были слиты в реакции ПЦР с использованием праймеров Plat5-XhoI_FW и Tlat-XhoI_RV.

"PUL_E99Q_E103Q". Конструкция варианта пуллуланазы.

E99Q_E103Q была идентична конструкции пуллуланазы "дикого типа" (см. 1 выше) с использованием синтетической конструкции E99Q_E103Q в качестве шаблона.

Генерированные фрагменты были клонированы в двух направлениях в сайте XhoI интегрирующего вектора pICatH B. licheniformis (фиг. 3). Это привело к созданию шести конструкций: pICatH-PUL-Ori1 (фиг. 4), pICatH-PUL-Ori2, pICatH-PULm104-Ori1 (фиг. 5), pICatH-PULm104-Ori2, pICatH-PUL_E99Q_E103Q-Ori1 (фиг. 6) и pICatH-PUL_E99Q_E103Q-Ori2. Конструкции Ori1 и Ori2 имеют противоположное направление гена пуллуланазы по сравнению с геном устойчивости к хлорамфениколу. Фиг. 4-6 отображают карты плазмиды трех конструкций Ori1.

Все шесть конструкций были трансформированы в B. subtilis и скринированы на образование ореола в покрывающих слоях AZCL-пуллулана (Megazyme) (0,1% в 100 mM NaAc pH 5, 1% агара). Трансформанты pICatH-PULm104 давали более выраженное образование ореола, чем трансформанты с полноразмерной пуллуланазой. Конструкции были верифицированы по последовательности и трансформированы в штаммы хозяева B. licheniformis BML612 и BML780 с применением трансформации протопластов.

Пример 3. Встраивание в геном B. licheniformis.

После трансформации проводили отбор трансформантов на планшетах минимальной регенерации, содержавших 5 мкг/мл хлорамфеникола и 10 мкг/мл неомицина. Трансформанты были точными копиями нанесены на два планшета с агаром, содержащим сердечный настой (известным компетентным специалистам), куда были добавлены те же антибиотики, что и в прокрывающий слой AZCL-пуллулана для отбора пуллуланазопозитивных трансформантов. Аналогично ситуации с B. subtilis трансформанты PULm104 демонстрировали более выраженные ореолы. Плазмиды были интегрированы в локус catH хромосомы B. licheniformis. Таким образом, для эксцизии/амплификации был использован следующий набор трансформантов, представленный в табл. 2.

Таблица 2

BML612 PUL Ori1
BML612 PUL Ori2
BML612 PULm104 Ori1
BML612 PULm104 Ori2
BML612 PUL_E99Q_E103Q Ori1
BML612 PUL_E99Q_E103Q Ori2
BML780 PUL Ori1
BML780 PUL Ori2
BML780 PULm1O4 Ori1
BML780 PULm104 Ori2
BML780 PUL_E99Q_E103Q Ori1
BML780 PUL_E99Q_E103Q Ori2

Эксцизию плазмид и кластерную амплификацию проводили следующим образом. Штаммы без чужеродной ДНК ("свободные штаммы") получали посредством эксцизии векторных последовательностей ("выпетливаний"), оставляя только catH - экспрессирующий кластер пуллуланазы, встроенный в хромосому. Затем экспрессирующий кластер амплифицировали, подвергая штаммы воздействия хлорамфеникола с пошаговым увеличением концентрации (5, 25, 50, 75 мкг/мл). Выработку пуллуланазы мониторировали, покрывая копии планшетов AZCL-пуллуланом после каждого этапа амплификации. Для каждого штамма были получены четыре уровня амплификации: CAP5, CAP25, CAP50 и CAP75.

Пример 4. Оценка штаммов B. licheniformis на пуллуланазу.

Штаммы были отобраны с дублированием на всех уровнях амплификации и внесены в один большой планшет с агаром, содержащим сердечный настой, с добавлением хлорамфеникола 5 мкг/мл и выращивались в течение ночи при 37°C. Штамм BMP 139, вырабатывающий пуллуланазу, был введен в качестве репера. Активность пуллуланазы визуализировали, покрывая планшеты слоем агара с AZCL-пуллуланом. Нанесенный покровный слой инкубировали 8 ч при 37°C с последующей 16-часовой инкубацией при комнатной температуре. Сводка результатов представлена ниже в табл. 3.

Таблица 3

Штамм	CAP5	CAP25	CAP50	CAP75	
BML612 PUL Ori1	+	+	+	++	
BML612 PUL Ori2	+	+	+	++	
BML612 PULml04 Ori1	+	++	++	+++	
BML612 PULml04 Ori2	+	+	+++	+++	
BML612 PUL_E99Q_E103Q Ori1	+	+	++	++	
BML612 PUL_E99Q_E103Q Ori2	+	+	++	++	
BML780 PUL Ori1	+	+	+	++	
BML780 PUL Ori2	+	+	++	++	
BML780 PULml04 Ori1	++	++	+++	+++++	
BML780 PULml04 Ori2	++	++	+++	+++++	
BML780 PUL_E99Q_E103Q Ori1	++	++	++	++	
BML780 PUL_E99Q_E103Q Ori2	++	++	++	+++	
BMP139					++

Легенда:

+ = диаметр ореола 7-9 мм

++ = диаметр ореола 10-12 мм

+++ = диаметр ореола 13-15 мм

++++ = диаметр ореола 16-18 мм

+++++ = диаметр ореола 19-21 мм

Из результатов проведенной оценки явно следует, что амплификация приводит к увеличению титров и/или производительности. Более специфические выводы.

1. Штаммы PUL с N-концевым усечением (PULm 104) имеют намного более высокую производительность по сравнению со штаммами, вырабатывающими полноразмерную пуллуланазу. Штаммы BML780 PULm 104 CAP75 продуцируют ореолы с площадью поверхности, увеличенной в 2,5 раза (увеличение диаметра более чем в 1,5 раза) по сравнению со штаммом BMP 139 и штаммами BML780 PUL CAP75. Это позволяет предположить, что укороченная молекула вырабатывается с более высокими титрами или ее активность увеличена по сравнению с полноразмерными молекулами пуллуланазы.

2. Вариант PUL_E99Q_E103Q, возможно, имеет небольшое преимущество перед "диким типом" PUL. Ореолы, продуцируемые штаммами BML780 PUL_E99Q_E103Q, несколько больше по сравнению со штаммами BML780 PUL.

3. Продуктивный штамм BMP139 по своей производительности примерно соответствует амплифицированным штаммам "дикого типа" PUL CAP75. Таким образом, основываясь на оценке планшетов, можно сделать вывод о том, что кодонная оптимизация не приводит к повышению производительности.

4. Штаммы BML780 обычно имеют лучшую производительность по сравнению со штаммами BML612. Это наблюдение согласуется с предыдущими данными о разрушении пуллуланазы в исходном материале BML612. Тот факт, что сконструированные здесь штаммы с пуллуланазой BML612 демонстрируют приемлемую очистку AZCL, позволяет предположить, что в планшетах пуллуланаза относительно стабильна, даже в исходном материале BML612.

5. Между штаммами Ori1 и Ori2, вырабатывающими пуллуланазу, не обнаружено явных различий по производительности.

Все публикации и патенты, упоминаемые в приведенном выше описании изобретения, включены сюда в качестве ссылок. Для компетентного специалиста будут очевидны различные модификации и варианты описанных здесь способов и систем изобретения, которые могут быть осуществлены без отклонений от сферы действия и духа этого изобретения. Хотя изобретение было описано в связи с особыми и наиболее предпочтительными вариантами его реализации, необходимо понимать, что изобретение в заявленном здесь виде не ограничено этими специфическим вариантами его реализации. Несомненно, что различные модификации описанных способов и методов реализации изобретения, которые очевидны компетентному специалисту, не должны выходить за пределы представленных ниже пунктов формулы изобретения.

Список последовательностей

<110> Danisco US, Inc., Genencor Division
England, George
Kolkman, Marc
Miller, Brian S.
Vroemen, Casper

<120> Варианты пуллуланазы с повышенной продуктивностью

<130> GC907-PCT

<140> PCT/US2007/018523
<141> 2007-08-21

<150> US 60/903,247
<151> 2007-02-23

<150> US 60/839,735
<151> 2006-08-23

<160> 10

<170> PatentIn version 3.4

<210> 1
<211> 2871
<212> ДНК
<213> Искусственная

<220>
<223> синтетическая кодон-оптимизированная последовательность из Bacillus deramificans

<400> 1
atgaaacaac aaaaacggct ttacgcccga ttgctgacgc tgttatttgc gctcatcttc 60
ttgctgcctc attctgcagc ttcagcagat ggcaatacga cgacgatcat cgtccattat 120
tttagaccgg cgggagatta tcaaccgtgg agcctttgga tgtggccgaa agatggagga 180
ggagcggaat atgattttaa ccagccggca gattcacttg gagcagtcgc ctcagcagat 240
attccgggaa atccgagcca agtcggcatc atcgtcagaa cacaggattg gacgaaagat 300
gtcagcgccg atcgctatat cgatctgagc aaaggcaatg aagtctggct ggtcgaaggc 360
aacagccaga tcttttatag cgaaaaagac gccgaagatg ctgctaaacc ggcagtcagc 420
aacgcttatc tggatgccag caaccaagtc ctggtcaaac tgagccaacc gctgacactt 480
ggagaaggag cgagcggatt tacggtccat gatgacacgg cgaacaaaga tatcccggtc 540
acgagcgtta aagatgctag cctgggccaa gatgtcacag cagttctggc gggcacgttt 600
caacatatct ttggcggatc agattgggca ccggataatc acagcacgct gctgaaaaaa 660
gtcacgaaca acctgtatca gtttagcgga gatctgccgg aaggcaacta tcaatataaa 720
gtcgccctga acgatagctg gaacaatccg agctatccga gcgataacat caatctgaca 780
gtcccggcag gcggagcaca tgtcacgttt agctatatcc cgagcacaca tgccgtctat 840
gacacgatca acaacccgaa cgccgatctt caagtcgaaa gcggcgtcaa aacggatctg 900
gtcacagtca cattgggaga agatccggat gtcagccata cactgagcat ccaaacggat 960
ggctatcaag cgaaacaagt catcccgaga aacgtcctga acagcagcca gtattattat 1020
agcggcgatg atctgggcaa cacgtataca caaaaagcga cgacgtttaa agtttgggcg 1080
ccgacaagca cacaagtcaa cgtcctgctg tatgattcag caacaggcag cgtcacaaaa 1140
atcgtcccga tgacagcatc aggacatgga gtctgggaag cgacggtcaa ccaaaacctg 1200
gaaaactggt attatatgta tgaagtcacg ggccaaggat caacaagaac agcggtcgat 1260
ccgtatgcta cagcaatcgc cccgaatgga acaagaggca tgatcgtcga tctggcaaaa 1320

```
acagacccgg caggctggaa tagcgataaa catatcacgc cgaaaaacat cgaagatgaa   1380
gtcatctatg aaatggacgt ccgggatttt agcatcgatc cgaacagcgg catgaaaaac   1440
aaaggcaaat atctggcgct gacggaaaaa ggaacaaaag gcccggataa cgtcaaaaca   1500
ggcatcgata gcctgaaaca actgggcatc acacatgtcc aactgatgcc ggtctttgct   1560
agcaatagcg tcgatgaaac ggacccgaca caagataact ggggctatga cccgagaaat   1620
tatgatgtcc cggaaggcca atatgccacg aacgccaatg gaaacgcccg gatcaaagaa   1680
tttaaagaaa tggtcctgag ccttcataga gaacatatcg gcgtcaacat ggacgtcgtc   1740
tataaccata cgtttgccac acagatcagc gactttgata aaatcgtgcc ggaatattat   1800
tatcggacgg atgacgccgg caattatacg aatggcagcg gcacaggaaa tgaaatcgcc   1860
gccgaaagac cgatggtcca gaaatttatc atcgacagcc ttaaatattg ggtcaacgaa   1920
tatcatatcg acggctttcg ctttgatctg atggcgctgc tgggcaaaga tacaatgagc   1980
aaagcggcga gcgaacttca tgctatcaat ccgggcatcg ctctttatgg agaaccgtgg   2040
acaggaggaa catcagcact gccggatgat caactgctga caaaaggcgc ccaaaaagga   2100
atgggagtcg ccgtctttaa cgacaacctg agaaatgccc tggatggcaa cgtttttgat   2160
agcagcgccc aaggatttgc tacaggagcg acaggactga cagatgccat caaaaatggc   2220
gtcgaaggca gcatcaacga ttttacaagc agcccgggag aaacgatcaa ttatgtcacg   2280
agccatgaca actatacgct gtgggacaaa atcgctctga gcaacccgaa tgatagcgaa   2340
gcggaccgga tcaaaatgga tgaactggca caagcagtcg tcatgacatc acaaggcgtc   2400
ccgtttatgc aaggcggaga agaaatgctg agaacgaaag gcggcaacga caacagctat   2460
aatgccggcg atgccgtcaa tgaatttgac tggagccgga aagcacaata tccggacgtc   2520
tttaactatt attcaggact tatccatctg agactggacc atccggcgtt tagaatgacg   2580
acggcgaacg aaatcaacag ccatcttcag tttctgaaca gcccggaaaa tacggtcgcc   2640
tatgaactga cggaccatgt gaacaaagac aaatggggca acatcatcgt cgtttataac   2700
ccgaacaaaa cggtcgccac aatcaatctt ccgagcggca aatgggcaat caatgccaca   2760
agcggcaaag ttggagaaag cacactggga caagcagaag gatcagtcca agtcccggga   2820
atcagcatga tgatccttca tcaagaagtc agcccggacc acggcaaaaa a            2871
```

<210> 2
<211> 957
<212> PRT
<213> Bacillus deramificans

<400> 2

```
Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe
1               5                  10                  15
Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ser Ala Asp Gly Asn
            20                  25                  30
Thr Thr Thr Ile Ile Val His Tyr Phe Arg Pro Ala Gly Asp Tyr Gln
        35                  40                  45
Pro Trp Ser Leu Trp Met Trp Pro Lys Asp Gly Gly Gly Ala Glu Tyr
    50                  55                  60
Asp Phe Asn Gln Pro Ala Asp Ser Leu Gly Ala Val Ala Ser Ala Asp
65                  70                  75                  80
Ile Pro Gly Asn Pro Ser Gln Val Gly Ile Ile Val Arg Thr Gln Asp
                85                  90                  95
Trp Thr Lys Asp Val Ser Ala Asp Arg Tyr Ile Asp Leu Ser Lys Gly
            100                 105                 110
Asn Glu Val Trp Leu Val Glu Gly Asn Ser Gln Ile Phe Tyr Ser Glu
        115                 120                 125
Lys Asp Ala Glu Asp Ala Ala Lys Pro Ala Val Ser Asn Ala Tyr Leu
    130                 135                 140
Asp Ala Ser Asn Gln Val Leu Val Lys Leu Ser Gln Pro Leu Thr Leu
145                 150                 155                 160
Gly Glu Gly Ala Ser Gly Phe Thr Val His Asp Asp Thr Ala Asn Lys
                165                 170                 175
Asp Ile Pro Val Thr Ser Val Lys Asp Ala Ser Leu Gly Gln Asp Val
            180                 185                 190
```

```
Thr Ala Val Leu Ala Gly Thr Phe Gln His Ile Phe Gly Gly Ser Asp
        195                 200                 205
Trp Ala Pro Asp Asn His Ser Thr Leu Leu Lys Lys Val Thr Asn Asn
    210                 215                 220
Leu Tyr Gln Phe Ser Gly Asp Leu Pro Glu Gly Asn Tyr Gln Tyr Lys
225                 230                 235                 240
Val Ala Leu Asn Asp Ser Trp Asn Asn Pro Ser Tyr Pro Ser Asp Asn
                245                 250                 255
Ile Asn Leu Thr Val Pro Ala Gly Gly Ala His Val Thr Phe Ser Tyr
            260                 265                 270
Ile Pro Ser Thr His Ala Val Tyr Asp Thr Ile Asn Asn Pro Asn Ala
        275                 280                 285
Asp Leu Gln Val Glu Ser Gly Val Lys Thr Asp Leu Val Thr Val Thr
    290                 295                 300
Leu Gly Glu Asp Pro Asp Val Ser His Thr Leu Ser Ile Gln Thr Asp
305                 310                 315                 320
Gly Tyr Gln Ala Lys Gln Val Ile Pro Arg Asn Val Leu Asn Ser Ser
                325                 330                 335
Gln Tyr Tyr Tyr Ser Gly Asp Asp Leu Gly Asn Thr Tyr Thr Gln Lys
            340                 345                 350
Ala Thr Thr Phe Lys Val Trp Ala Pro Thr Ser Thr Gln Val Asn Val
        355                 360                 365
Leu Leu Tyr Asp Ser Ala Thr Gly Ser Val Thr Lys Ile Val Pro Met
    370                 375                 380
Thr Ala Ser Gly His Gly Val Trp Glu Ala Thr Val Asn Gln Asn Leu
385                 390                 395                 400
Glu Asn Trp Tyr Tyr Met Tyr Glu Val Thr Gly Gln Gly Ser Thr Arg
                405                 410                 415
Thr Ala Val Asp Pro Tyr Ala Thr Ala Ile Ala Pro Asn Gly Thr Arg
            420                 425                 430
Gly Met Ile Val Asp Leu Ala Lys Thr Asp Pro Ala Gly Trp Asn Ser
        435                 440                 445
Asp Lys His Ile Thr Pro Lys Asn Ile Glu Asp Glu Val Ile Tyr Glu
    450                 455                 460
Met Asp Val Arg Asp Phe Ser Ile Asp Pro Asn Ser Gly Met Lys Asn
465                 470                 475                 480
Lys Gly Lys Tyr Leu Ala Leu Thr Glu Lys Gly Thr Lys Gly Pro Asp
                485                 490                 495
Asn Val Lys Thr Gly Ile Asp Ser Leu Lys Gln Leu Gly Ile Thr His
            500                 505                 510
Val Gln Leu Met Pro Val Phe Ala Ser Asn Ser Val Asp Glu Thr Asp
        515                 520                 525
Pro Thr Gln Asp Asn Trp Gly Tyr Asp Pro Arg Asn Tyr Asp Val Pro
    530                 535                 540
Glu Gly Gln Tyr Ala Thr Asn Ala Asn Gly Asn Ala Arg Ile Lys Glu
545                 550                 555                 560
Phe Lys Glu Met Val Leu Ser Leu His Arg Glu His Ile Gly Val Asn
                565                 570                 575
Met Asp Val Val Tyr Asn His Thr Phe Ala Thr Gln Ile Ser Asp Phe
            580                 585                 590
Asp Lys Ile Val Pro Glu Tyr Tyr Tyr Arg Thr Asp Asp Ala Gly Asn
        595                 600                 605
Tyr Thr Asn Gly Ser Gly Thr Gly Asn Glu Ile Ala Ala Glu Arg Pro
    610                 615                 620
Met Val Gln Lys Phe Ile Ile Asp Ser Leu Lys Tyr Trp Val Asn Glu
625                 630                 635                 640
Tyr His Ile Asp Gly Phe Arg Phe Asp Leu Met Ala Leu Leu Gly Lys
```

```
                645                 650                 655
Asp Thr Met Ser Lys Ala Ala Ser Glu Leu His Ala Ile Asn Pro Gly
            660                 665                 670
Ile Ala Leu Tyr Gly Glu Pro Trp Thr Gly Gly Thr Ser Ala Leu Pro
        675                 680                 685
Asp Asp Gln Leu Leu Thr Lys Gly Ala Gln Lys Gly Met Gly Val Ala
    690                 695                 700
Val Phe Asn Asp Asn Leu Arg Asn Ala Leu Asp Gly Asn Val Phe Asp
705                 710                 715                 720
Ser Ser Ala Gln Gly Phe Ala Thr Gly Ala Thr Gly Leu Thr Asp Ala
                725                 730                 735
Ile Lys Asn Gly Val Glu Gly Ser Ile Asn Asp Phe Thr Ser Ser Pro
            740                 745                 750
Gly Glu Thr Ile Asn Tyr Val Thr Ser His Asp Asn Tyr Thr Leu Trp
        755                 760                 765
Asp Lys Ile Ala Leu Ser Asn Pro Asn Asp Ser Glu Ala Asp Arg Ile
    770                 775                 780
Lys Met Asp Glu Leu Ala Gln Ala Val Val Met Thr Ser Gln Gly Val
785                 790                 795                 800
Pro Phe Met Gln Gly Gly Glu Glu Met Leu Arg Thr Lys Gly Gly Asn
                805                 810                 815
Asp Asn Ser Tyr Asn Ala Gly Asp Ala Val Asn Glu Phe Asp Trp Ser
            820                 825                 830
Arg Lys Ala Gln Tyr Pro Asp Val Phe Asn Tyr Tyr Ser Gly Leu Ile
        835                 840                 845
His Leu Arg Leu Asp His Pro Ala Phe Arg Met Thr Thr Ala Asn Glu
    850                 855                 860
Ile Asn Ser His Leu Gln Phe Leu Asn Ser Pro Glu Asn Thr Val Ala
865                 870                 875                 880
Tyr Glu Leu Thr Asp His Val Asn Lys Asp Lys Trp Gly Asn Ile Ile
                885                 890                 895
Val Val Tyr Asn Pro Asn Lys Thr Val Ala Thr Ile Asn Leu Pro Ser
            900                 905                 910
Gly Lys Trp Ala Ile Asn Ala Thr Ser Gly Lys Val Gly Glu Ser Thr
        915                 920                 925
Leu Gly Gln Ala Glu Gly Ser Val Gln Val Pro Gly Ile Ser Met Met
    930                 935                 940
Ile Leu His Gln Glu Val Ser Pro Asp His Gly Lys Lys
945                 950                 955
```

<210> 3
<211> 2559
<212> ДНК
<213> Искусственная

<220>
<223> синтетическая кодон-оптимизированная пуллуланаза

<400> 3

```
atgaaacaac aaaaacggct ttacgcccga ttgctgacgc tgttatttgc gctcatcttc     60
ttgctgcctc attctgcagc ttcagcagct gctaaaccgg cagtcagcaa cgcttatctg    120
gatgccagca accaagtcct ggtcaaactg agccaaccgc tgacacttgg agaaggagcg    180
agcggattta cggtccatga tgacacggcg aacaaagata tcccggtcac gagcgttaaa    240
gatgctagcc tgggccaaga tgtcacagca gttctggcgg gcacgtttca acatatcttt    300
ggcggatcag attgggcacc ggataatcac agcacgctgc tgaaaaaagt cacgaacaac    360
ctgtatcagt ttagcggaga tctgccggaa ggcaactatc aatataaagt cgccctgaac    420
gatagctgga acaatccgag ctatccgagc gataacatca atctgacagt cccggcaggc    480
```

```
ggagcacatg tcacgtttag ctatatcccg agcacacatg ccgtctatga cacgatcaac    540
aacccgaacg ccgatcttca agtcgaaagc ggcgtcaaaa cggatctggt cacagtcaca    600
ttgggagaag atccggatgt cagccataca ctgagcatcc aaacggatgg ctatcaagcg    660
aaacaagtca tcccgagaaa cgtcctgaac agcagccagt attattatag cggcgatgat    720
ctgggcaaca cgtatacaca aaaagcgacg acgtttaaag tttgggcgcc gacaagcaca    780
caagtcaacg tcctgctgta tgattcagca acaggcagcg tcacaaaaat cgtcccgatg    840
acagcatcag gacatggagt ctgggaagcg acggtcaacc aaaacctgga aaactggtat    900
tatatgtatg aagtcacggg ccaaggatca acaagaacag cggtcgatcc gtatgctaca    960
gcaatcgccc cgaatggaac aagaggcatg atcgtcgatc tggcaaaaac agacccggca   1020
ggctggaata gcgataaaca tatcacgccg aaaaacatcg aagatgaagt catctatgaa   1080
atggacgtcc gggattttag catcgatccg aacagcggca tgaaaaacaa aggcaaatat   1140
ctggcgctga cggaaaaagg aacaaaaggc ccggataacg tcaaaacagg catcgatagc   1200
ctgaaacaac tgggcatcac acatgtccaa ctgatgccgg tctttgctag caatagcgtc   1260
gatgaaacgg acccgacaca agataactgg ggctatgacc cgagaaatta tgatgtcccg   1320
gaaggccaat atgccacgaa cgccaatgga aacgcccgga tcaaagaatt taaagaaatg   1380
gtcctgagcc ttcatagaga acatatcggc gtcaacatgg acgtcgtcta taaccatacg   1440
tttgccacac agatcagcga ctttgataaa atcgtgccgg aatattatta tcggacggat   1500
gacgccggca attatacgaa tggcagcggc acaggaaatg aaatcgccgc cgaaagaccg   1560
atggtccaga aatttatcat cgacagcctt aaatattggg tcaacgaata tcatatcgac   1620
ggctttcgct ttgatctgat ggcgctgctg ggcaaagata caatgagcaa agcggcgagc   1680
gaacttcatg ctatcaatcc gggcatcgct ctttatggag aaccgtggac aggaggaaca   1740
tcagcactgc cggatgatca actgctgaca aaaggcgccc aaaaaggaat gggagtcgcc   1800
gtctttaacg acaacctgag aaatgccctg gatggcaacg tttttgatag cagcgcccaa   1860
ggatttgcta caggagcgac aggactgaca gatgccatca aaaatggcgt cgaaggcagc   1920
atcaacgatt ttacaagcag cccgggagag acgatcaatt atgtcacgag ccatgacaac   1980
tatacgctgt gggacaaaat cgctctgagc aacccgaatg atagcgaagc ggaccggatc   2040
aaaatggatg aactggcaca agcagtcgtc atgacatcac aaggcgtccc gtttatgcaa   2100
ggcggagaag aaatgctgag aacgaaaggc ggcaacgaca acagctataa tgccggcgat   2160
gccgtcaatg aatttgactg gagccggaaa gcacaatatc cggacgtctt taactattat   2220
tcaggactta tccatctgag actggaccat ccggcgttta gaatgacgac ggcgaacgaa   2280
atcaacagcc atcttcagtt tctgaacagc ccggaaaata cggtcgccta tgaactgacg   2340
gaccatgtga acaaagacaa atggggcaac atcatcgtcg tttataaccc gaacaaaacg   2400
gtcgccacaa tcaatcttcc gagcggcaaa tgggcaatca atgccacaag cggcaaagtt   2460
ggagaaagca cactgggaca agcagaagga tcagtccaag tcccgggaat cagcatgatg   2520
atccttcatc aagaagtcag cccggaccac ggcaaaaaa                          2559
```

<210> 4
<211> 853
<212> PRT
<213> Искусственная

<220>
<223> синтетическая кодон-оптимизированная пуллуланаза

<400> 4

```
Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe
1               5                  10                  15
Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ser Ala Ala Ala Lys
            20                  25                  30
Pro Ala Val Ser Asn Ala Tyr Leu Asp Ala Ser Asn Gln Val Leu Val
        35                  40                  45
Lys Leu Ser Gln Pro Leu Thr Leu Gly Glu Gly Ala Ser Gly Phe Thr
    50                  55                  60
Val His Asp Asp Thr Ala Asn Lys Asp Ile Pro Val Thr Ser Val Lys
65                  70                  75                  80
Asp Ala Ser Leu Gly Gln Asp Val Thr Ala Val Leu Ala Gly Thr Phe
                85                  90                  95
```

```
Gln His Ile Phe Gly Gly Ser Asp Trp Ala Pro Asp Asn His Ser Thr
            100                 105                 110
Leu Leu Lys Lys Val Thr Asn Asn Leu Tyr Gln Phe Ser Gly Asp Leu
        115                 120                 125
Pro Glu Gly Asn Tyr Gln Tyr Lys Val Ala Leu Asn Asp Ser Trp Asn
    130                 135                 140
Asn Pro Ser Tyr Pro Ser Asp Asn Ile Asn Leu Thr Val Pro Ala Gly
145                 150                 155                 160
Gly Ala His Val Thr Phe Ser Tyr Ile Pro Ser Thr His Ala Val Tyr
                165                 170                 175
Asp Thr Ile Asn Asn Pro Asn Ala Asp Leu Gln Val Glu Ser Gly Val
            180                 185                 190
Lys Thr Asp Leu Val Thr Val Thr Leu Gly Glu Asp Pro Asp Val Ser
        195                 200                 205
His Thr Leu Ser Ile Gln Thr Asp Gly Tyr Gln Ala Lys Gln Val Ile
    210                 215                 220
Pro Arg Asn Val Leu Asn Ser Ser Gln Tyr Tyr Tyr Ser Gly Asp Asp
225                 230                 235                 240
Leu Gly Asn Thr Tyr Thr Gln Lys Ala Thr Thr Phe Lys Val Trp Ala
                245                 250                 255
Pro Thr Ser Thr Gln Val Asn Val Leu Leu Tyr Asp Ser Ala Thr Gly
            260                 265                 270
Ser Val Thr Lys Ile Val Pro Met Thr Ala Ser Gly His Gly Val Trp
        275                 280                 285
Glu Ala Thr Val Asn Gln Asn Leu Glu Asn Trp Tyr Tyr Met Tyr Glu
    290                 295                 300
Val Thr Gly Gln Gly Ser Thr Arg Thr Ala Val Asp Pro Tyr Ala Thr
305                 310                 315                 320
Ala Ile Ala Pro Asn Gly Thr Arg Gly Met Ile Val Asp Leu Ala Lys
                325                 330                 335
Thr Asp Pro Ala Gly Trp Asn Ser Asp Lys His Ile Thr Pro Lys Asn
            340                 345                 350
Ile Glu Asp Glu Val Ile Tyr Glu Met Asp Val Arg Asp Phe Ser Ile
        355                 360                 365
Asp Pro Asn Ser Gly Met Lys Asn Lys Gly Lys Tyr Leu Ala Leu Thr
    370                 375                 380
Glu Lys Gly Thr Lys Gly Pro Asp Asn Val Lys Thr Gly Ile Asp Ser
385                 390                 395                 400
Leu Lys Gln Leu Gly Ile Thr His Val Gln Leu Met Pro Val Phe Ala
                405                 410                 415
Ser Asn Ser Val Asp Glu Thr Asp Pro Thr Gln Asp Asn Trp Gly Tyr
            420                 425                 430
Asp Pro Arg Asn Tyr Asp Val Pro Glu Gly Gln Tyr Ala Thr Asn Ala
        435                 440                 445
Asn Gly Asn Ala Arg Ile Lys Glu Phe Lys Glu Met Val Leu Ser Leu
    450                 455                 460
His Arg Glu His Ile Gly Val Asn Met Asp Val Val Tyr Asn His Thr
465                 470                 475                 480
Phe Ala Thr Gln Ile Ser Asp Phe Asp Lys Ile Val Pro Glu Tyr Tyr
                485                 490                 495
Tyr Arg Thr Asp Asp Ala Gly Asn Tyr Thr Asn Gly Ser Gly Thr Gly
            500                 505                 510
Asn Glu Ile Ala Ala Glu Arg Pro Met Val Gln Lys Phe Ile Ile Asp
        515                 520                 525
Ser Leu Lys Tyr Trp Val Asn Glu Tyr His Ile Asp Gly Phe Arg Phe
    530                 535                 540
Asp Leu Met Ala Leu Leu Gly Lys Asp Thr Met Ser Lys Ala Ala Ser
```

```
545                 550                 555                 560
Glu Leu His Ala Ile Asn Pro Gly Ile Ala Leu Tyr Gly Glu Pro Trp
                565                 570                 575
Thr Gly Gly Thr Ser Ala Leu Pro Asp Asp Gln Leu Leu Thr Lys Gly
            580                 585                 590
Ala Gln Lys Gly Met Gly Val Ala Val Phe Asn Asp Asn Leu Arg Asn
        595                 600                 605
Ala Leu Asp Gly Asn Val Phe Asp Ser Ser Ala Gln Gly Phe Ala Thr
    610                 615                 620
Gly Ala Thr Gly Leu Thr Asp Ala Ile Lys Asn Gly Val Glu Gly Ser
625                 630                 635                 640
Ile Asn Asp Phe Thr Ser Ser Pro Gly Glu Thr Ile Asn Tyr Val Thr
                645                 650                 655
Ser His Asp Asn Tyr Thr Leu Trp Asp Lys Ile Ala Leu Ser Asn Pro
            660                 665                 670
Asn Asp Ser Glu Ala Asp Arg Ile Lys Met Asp Glu Leu Ala Gln Ala
        675                 680                 685
Val Val Met Thr Ser Gln Gly Val Pro Phe Met Gln Gly Gly Glu Glu
    690                 695                 700
Met Leu Arg Thr Lys Gly Gly Asn Asp Asn Ser Tyr Asn Ala Gly Asp
705                 710                 715                 720
Ala Val Asn Glu Phe Asp Trp Ser Arg Lys Ala Gln Tyr Pro Asp Val
                725                 730                 735
Phe Asn Tyr Tyr Ser Gly Leu Ile His Leu Arg Leu Asp His Pro Ala
            740                 745                 750
Phe Arg Met Thr Thr Ala Asn Glu Ile Asn Ser His Leu Gln Phe Leu
        755                 760                 765
Asn Ser Pro Glu Asn Thr Val Ala Tyr Glu Leu Thr Asp His Val Asn
    770                 775                 780
Lys Asp Lys Trp Gly Asn Ile Ile Val Val Tyr Asn Pro Asn Lys Thr
785                 790                 795                 800
Val Ala Thr Ile Asn Leu Pro Ser Gly Lys Trp Ala Ile Asn Ala Thr
                805                 810                 815
Ser Gly Lys Val Gly Glu Ser Thr Leu Gly Gln Ala Glu Gly Ser Val
            820                 825                 830
Gln Val Pro Gly Ile Ser Met Met Ile Leu His Gln Glu Val Ser Pro
        835                 840                 845
Asp His Gly Lys Lys
    850
```

<210> 5
<211> 2871
<212> ДНК
<213> Искусственная

<220>
<223> синтетическая модифицированная пуллуланаза

<400> 5

```
atgaaacaac aaaaacggct ttacgcccga ttgctgacgc tgttatttgc gctcatcttc     60
ttgctgcctc attctgcagc ttcagcagat ggcaatacga cgacgatcat cgtccattat    120
tttagaccgg cgggagatta tcaaccgtgg agcctttgga tgtggccgaa agatggagga    180
ggagcggaat atgattttaa ccagccggca gattcacttg gagcagtcgc ctcagcagat    240
attccgggaa atccgagcca agtcggcatc atcgtcagaa cacaggattg gacgaaagat    300
gtcagcgccg atcgctatat cgatctgagc aaaggcaatg aagtctggct ggtcgaaggc    360
aacagccaga tcttttatag cgaaaaagac gccgaagatg ctgctaaacc ggcagtcagc    420
aacgcttatc tggatgccag caaccaagtc ctggtcaaac tgagccaacc gctgacactt    480
```

```
ggagaaggag cgagcggatt tacggtccat gatgacacgg cgaacaaaga tatcccggtc    540
acgagcgtta aagatgctag cctgggccaa gatgtcacag cagttctggc gggcacgttt    600
caacatatct ttggcggatc agattgggca ccggataatc acagcacgct gctgaaaaaa    660
gtcacgaaca acctgtatca gtttagcgga gatctgccgg aaggcaacta tcaatataaa    720
gtcgccctga acgatagctg gaacaatccg agctatccga gcgataacat caatctgaca    780
gtcccggcag gcggagcaca tgtcacgttt agctatatcc cgagcacaca tgccgtctat    840
gacacgatca acaacccgaa cgccgatctt caagtcgaaa gcggcgtcaa aacggatctg    900
gtcacagtca cattgggaga agatccggat gtcagccata cactgagcat ccaaacggat    960
ggctatcaag cgaaacaagt catcccgaga aacgtcctga acagcagcca gtattattat   1020
agcggcgatg atctgggcaa cacgtataca caaaaagcga cgacgtttaa agtttgggcg   1080
ccgacaagca cacaagtcaa cgtcctgctg tatgattcag caacaggcag cgtcacaaaa   1140
atcgtcccga tgacagcatc aggacatgga gtctgggaag cgacggtcaa ccaaaacctg   1200
gaaaactggt attatatgta tgaagtcacg ggccaaggat caacaagaac agcggtcgat   1260
ccgtatgcta cagcaatcgc cccgaatgga acaagaggca tgatcgtcga tctggcaaaa   1320
acagacccgg caggctggaa tagcgataaa catatcacgc cgaaaaacat cgaagatgaa   1380
gtcatctatg aaatggacgt ccgggatttt agcatcgatc cgaacagcgg catgaaaaac   1440
aaaggcaaat atctggcgct gacggaaaaa ggaacaaaag gcccggataa cgtcaaaaca   1500
ggcatcgata gcctgaaaca actgggcatc acacatgtcc aactgatgcc ggtctttgct   1560
agcaatagcg tcgatgaaac ggacccgaca caagataact ggggctatga cccgagaaat   1620
tatgatgtcc cggaaggcca atatgccacg aacgccaatg gaaacgcccg gatcaaagaa   1680
tttaaagaaa tggtcctgag ccttcataga gaacatatcg gcgtcaacat ggacgtcgtc   1740
tataaccata cgtttgccac acagatcagc gactttgata aaatcgtgcc ggaatattat   1800
tatcggacgg atgacgccgg caattatacg aatggcagcg gcacaggaaa tgaaatcgcc   1860
gccgaaagac cgatggtcca gaaatttatc atcgacagcc ttaaatattg ggtcaacgaa   1920
tatcatatcg acggctttcg ctttgatctg atggcgctgc tgggcaaaga tacaatgagc   1980
aaagcggcga gcgaacttca tgctatcaat ccgggcatcg ctctttatgg agaaccgtgg   2040
acaggaggaa catcagcact gccggatgat caactgctga caaaaggcgc ccaaaaagga   2100
atgggagtcg ccgtctttaa cgacaacctg agaaatgccc tggatggcaa cgtttttgat   2160
agcagcgccc aaggatttgc tacaggagcg acaggactga cagatgccat caaaaatggc   2220
gtcgaaggca gcatcaacga ttttacaagc agcccgggag aaacgatcaa ttatgtcacg   2280
agccatgaca actatacgct gtgggacaaa atcgctctga gcaacccgaa tgatagcgaa   2340
gcggaccgga tcaaaatgga tgaactggca caagcagtcg tcatgacatc acaaggcgtc   2400
ccgtttatgc aaggcggaga agaaatgctg agaacgaaag gcggcaacga caacagctat   2460
aatgccggcg atgccgtcaa tgaatttgac tggagccgga aagcacaata tccggacgtc   2520
tttaactatt attcaggact tatccatctg agactggacc atccggcgtt tagaatgacg   2580
acggcgaacg aaatcaacag ccatcttcag tttctgaaca gcccggaaaa tacggtcgcc   2640
tatgaactga cggaccatgt gaacaaagac aaatggggca acatcatcgt cgtttataac   2700
ccgaacaaaa cggtcgccac aatcaatctt ccgagcggca aatgggcaat caatgccaca   2760
agcggcaaag ttggagaaag cacactggga caagcagaag gatcagtcca agtcccggga   2820
atcagcatga tgatccttca tcaagaagtc agcccggacc acggcaaaaa a            2871
```

<210> 6
<211> 957
<212> PRT
<213> Искусственная

<220>
<223> синтетическая модифицированная пуллуланаза

<400> 6

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

```
    50                  55                  60
Asp Phe Asn Gln Pro Ala Asp Ser Leu Gly Ala Val Ala Ser Ala Asp
65                  70                  75                  80
Ile Pro Gly Asn Pro Ser Gln Val Gly Ile Ile Val Arg Thr Gln Asp
                85                  90                  95
Trp Thr Lys Asp Val Ser Ala Asp Arg Tyr Ile Asp Leu Ser Lys Gly
            100                 105                 110
Asn Glu Val Trp Leu Val Glu Gly Asn Ser Gln Ile Phe Tyr Ser Gln
        115                 120                 125
Lys Asp Ala Gln Asp Ala Ala Lys Pro Ala Val Ser Asn Ala Tyr Leu
    130                 135                 140
Asp Ala Ser Asn Gln Val Leu Val Lys Leu Ser Gln Pro Leu Thr Leu
145                 150                 155                 160
Gly Glu Gly Ala Ser Gly Phe Thr Val His Asp Asp Thr Ala Asn Lys
                165                 170                 175
Asp Ile Pro Val Thr Ser Val Lys Asp Ala Ser Leu Gly Gln Asp Val
            180                 185                 190
Thr Ala Val Leu Ala Gly Thr Phe Gln His Ile Phe Gly Gly Ser Asp
        195                 200                 205
Trp Ala Pro Asp Asn His Ser Thr Leu Leu Lys Lys Val Thr Asn Asn
    210                 215                 220
Leu Tyr Gln Phe Ser Gly Asp Leu Pro Glu Gly Asn Tyr Gln Tyr Lys
225                 230                 235                 240
Val Ala Leu Asn Asp Ser Trp Asn Asn Pro Ser Tyr Pro Ser Asp Asn
                245                 250                 255
Ile Asn Leu Thr Val Pro Ala Gly Gly Ala His Val Thr Phe Ser Tyr
            260                 265                 270
Ile Pro Ser Thr His Ala Val Tyr Asp Thr Ile Asn Asn Pro Asn Ala
        275                 280                 285
Asp Leu Gln Val Glu Ser Gly Val Lys Thr Asp Leu Val Thr Val Thr
    290                 295                 300
Leu Gly Glu Asp Pro Asp Val Ser His Thr Leu Ser Ile Gln Thr Asp
305                 310                 315                 320
Gly Tyr Gln Ala Lys Gln Val Ile Pro Arg Asn Val Leu Asn Ser Ser
                325                 330                 335
Gln Tyr Tyr Tyr Ser Gly Asp Asp Leu Gly Asn Thr Tyr Thr Gln Lys
            340                 345                 350
Ala Thr Thr Phe Lys Val Trp Ala Pro Thr Ser Thr Gln Val Asn Val
        355                 360                 365
Leu Leu Tyr Asp Ser Ala Thr Gly Ser Val Thr Lys Ile Val Pro Met
    370                 375                 380
Thr Ala Ser Gly His Gly Val Trp Glu Ala Thr Val Asn Gln Asn Leu
385                 390                 395                 400
Glu Asn Trp Tyr Tyr Met Tyr Glu Val Thr Gly Gln Gly Ser Thr Arg
                405                 410                 415
Thr Ala Val Asp Pro Tyr Ala Thr Ala Ile Ala Pro Asn Gly Thr Arg
            420                 425                 430
Gly Met Ile Val Asp Leu Ala Lys Thr Asp Pro Ala Gly Trp Asn Ser
        435                 440                 445
Asp Lys His Ile Thr Pro Lys Asn Ile Glu Asp Glu Val Ile Tyr Glu
    450                 455                 460
Met Asp Val Arg Asp Phe Ser Ile Asp Pro Asn Ser Gly Met Lys Asn
465                 470                 475                 480
Lys Gly Lys Tyr Leu Ala Leu Thr Glu Lys Gly Thr Lys Gly Pro Asp
                485                 490                 495
Asn Val Lys Thr Gly Ile Asp Ser Leu Lys Gln Leu Gly Ile Thr His
            500                 505                 510
```

```
Val Gln Leu Met Pro Val Phe Ala Ser Asn Ser Val Asp Glu Thr Asp
        515                 520                 525
Pro Thr Gln Asp Asn Trp Gly Tyr Asp Pro Arg Asn Tyr Asp Val Pro
    530                 535                 540
Glu Gly Gln Tyr Ala Thr Asn Ala Asn Gly Asn Ala Arg Ile Lys Glu
545                 550                 555                 560
Phe Lys Glu Met Val Leu Ser Leu His Arg Glu His Ile Gly Val Asn
                565                 570                 575
Met Asp Val Val Tyr Asn His Thr Phe Ala Thr Gln Ile Ser Asp Phe
            580                 585                 590
Asp Lys Ile Val Pro Glu Tyr Tyr Tyr Arg Thr Asp Asp Ala Gly Asn
        595                 600                 605
Tyr Thr Asn Gly Ser Gly Thr Gly Asn Glu Ile Ala Ala Glu Arg Pro
    610                 615                 620
Met Val Gln Lys Phe Ile Ile Asp Ser Leu Lys Tyr Trp Val Asn Glu
625                 630                 635                 640
Tyr His Ile Asp Gly Phe Arg Phe Asp Leu Met Ala Leu Leu Gly Lys
                645                 650                 655
Asp Thr Met Ser Lys Ala Ala Ser Glu Leu His Ala Ile Asn Pro Gly
            660                 665                 670
Ile Ala Leu Tyr Gly Glu Pro Trp Thr Gly Gly Thr Ser Ala Leu Pro
        675                 680                 685
Asp Asp Gln Leu Leu Thr Lys Gly Ala Gln Lys Gly Met Gly Val Ala
    690                 695                 700
Val Phe Asn Asp Asn Leu Arg Asn Ala Leu Asp Gly Asn Val Phe Asp
705                 710                 715                 720
Ser Ser Ala Gln Gly Phe Ala Thr Gly Ala Thr Gly Leu Thr Asp Ala
                725                 730                 735
Ile Lys Asn Gly Val Glu Gly Ser Ile Asn Asp Phe Thr Ser Ser Pro
            740                 745                 750
Gly Glu Thr Ile Asn Tyr Val Thr Ser His Asp Asn Tyr Thr Leu Trp
        755                 760                 765
Asp Lys Ile Ala Leu Ser Asn Pro Asn Asp Ser Glu Ala Asp Arg Ile
    770                 775                 780
Lys Met Asp Glu Leu Ala Gln Ala Val Val Met Thr Ser Gln Gly Val
785                 790                 795                 800
Pro Phe Met Gln Gly Gly Glu Glu Met Leu Arg Thr Lys Gly Gly Asn
                805                 810                 815
Asp Asn Ser Tyr Asn Ala Gly Asp Ala Val Asn Glu Phe Asp Trp Ser
            820                 825                 830
Arg Lys Ala Gln Tyr Pro Asp Val Phe Asn Tyr Tyr Ser Gly Leu Ile
        835                 840                 845
His Leu Arg Leu Asp His Pro Ala Phe Arg Met Thr Thr Ala Asn Glu
    850                 855                 860
Ile Asn Ser His Leu Gln Phe Leu Asn Ser Pro Glu Asn Thr Val Ala
865                 870                 875                 880
Tyr Glu Leu Thr Asp His Val Asn Lys Asp Lys Trp Gly Asn Ile Ile
                885                 890                 895
Val Val Tyr Asn Pro Asn Lys Thr Val Ala Thr Ile Asn Leu Pro Ser
            900                 905                 910
Gly Lys Trp Ala Ile Asn Ala Thr Ser Gly Lys Val Gly Glu Ser Thr
        915                 920                 925
Leu Gly Gln Ala Glu Gly Ser Val Gln Val Pro Gly Ile Ser Met Met
    930                 935                 940
Ile Leu His Gln Glu Val Ser Pro Asp His Gly Lys Lys
945                 950                 955
```

<210> 7
<211> 102
<212> ДНК
<213> Искусственная

<220>
<223> праймер

<400> 7
cccccgctcg aggcttttct tttggaagaa aatataggga aaatggtact tgttaaaaat 60
tcggaatatt tatacaatat catatgttta cattgaaagg gg 102

<210> 8
<211> 35
<212> ДНК
<213> Искусственная

<220>
<223> праймер

<400> 8
tggaatctcg aggttttatc ctttaccttg tctcc 35

<210> 9
<211> 51
<212> ДНК
<213> Искусственная

<220>
<223> праймер

<400> 9
gcgttgctga ctgccggttt agcagctgct gaagctgcag aatgaggcag c 51

<210> 10
<211> 51
<212> ДНК
<213> Искусственная

<220>
<223> праймер

<400> 10
gctgcctcat tctgcagctt cagcagctgc taaaccggca gtcagcaacg c 51

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Композиция для разжижения и осахаривания крахмала, содержащая:

(a) выделенную молекулу пептида, имеющую аминокислотную последовательность SEQ ID NO: 2 или ее фрагмент с удаленной аминоконцевой сигнальной последовательностью; или

(b) выделенную молекулу пептида, имеющую по меньшей мере 90% идентичность последовательности с аминокислотной последовательностью SEQ ID NO: 4 или ее фрагментом с удаленной аминоконцевой сигнальной последовательностью и содержащую делецию N-концевых аминокислотных остатков; или

(c) выделенную молекулу пептида, имеющую по меньшей мере 90% идентичность последовательности с аминокислотной последовательностью SEQ ID NO: 6 или ее фрагментом с удаленной аминоконцевой сигнальной последовательностью и содержащую аминокислотные замещения, соответствующие E99Q и E103Q,

причем указанная выделенная молекула пептида имеет активность пуллуланазы.

2. Композиция для разжижения и осахаривания крахмала, содержащая выделенную молекулу нуклеиновой кислоты, состоящую из нуклеотидной последовательности, выбранной из группы, состоящей из SEQ ID NO: 1, 3 и 5, причем указанная нуклеотидная последовательность кодирует пептид с активностью пуллуланазы, указанный в п.1.

3. Композиция для разжижения и осахаривания крахмала, содержащая выделенную ДНК, кодирующую пептид с аминокислотной последовательностью SEQ ID NO: 2, 4 или 6, имеющий активность пуллуланазы.

4. Композиция для разжижения и осахаривания крахмала, содержащая экспрессирующую конструкцию, включающую ДНК, состоящую из нуклеотидной последовательности, выбранной из группы, состоящей из SEQ ID NO: 1, 3 и 5, где указанная нуклеотидная последовательность кодирует пептид, имеющий активность пуллуланазы.

5. Экспрессирующая конструкция, охарактеризованная в п.4, функционально связанная с направляющими экспрессию последовательностями, распознаваемыми клеткой-хозяином, трансформированной указанной экспрессирующей конструкцией.

6. Экспрессирующая конструкция по п.5, трансфицированная в организм хозяина.

7. Экспрессирующая конструкция по п.6, где указанный организм хозяина выбран из группы, состоящей из грибов, бактерий и эукариотических клеток.

8. Экспрессирующая конструкция по п.7, где указанный организм хозяина выбран из группы, состоящей из Bacillus sp., Bacillus subtilis, Escherichia coli, Trichoderma reesei, Saccharomyces cerevisiae, Aspergillus niger и B. licheniformis.

9. Композиция для разжижения и осахаривания крахмала, содержащая экспрессирующую конструкцию, включающую ДНК, кодирующую пептид с аминокислотной последовательностью, выбранной из группы, состоящей из SEQ ID NO: 2, 4 и 6, имеющий активность пуллуланазы.

10. Способ продуцирования пептида с активностью пуллуланазы, предусматривающий:

a) получение i) экспрессирующей конструкции, содержащей ДНК, состоящую из нуклеотидной последовательности, кодирующей пептид с активностью пуллуланазы, причем указанная нуклеотидная последовательность выбрана из группы, состоящей из SEQ ID NO: 1, 3 и 5, ii) организма хозяина и iii) культуральной среды;

b) трансфекцию указанной экспрессирующей конструкции в указанный организм хозяина для создания трансфицированного организма хозяина;

c) культивирование указанного трансфицированного организма хозяина в указанной культуральной среде в течение промежутка времени и в условиях, достаточных для продуцирования пептида с активностью пуллуланазы.

11. Способ по п.10, в котором указанный пептид выделяют из указанной культуральной среды.

12. Способ по п.10, в котором указанный организм хозяина выбран из группы, состоящей из грибов, бактерий и эукариотических клеток.

13. Способ по п.12, в котором указанный организм хозяина выбран из группы, состоящей из Bacillus sp., Bacillus subtilis, Escherichia coli, Trichoderma reesei, Saccharomyces cerevisiae, Aspergillus niger и B. licheniformis.

14. Способ продуцирования пептида с активностью пуллуланазы, предусматривающий:

a) получение i) организма хозяина, трансфицированного экспрессирующей конструкцией, содержащей ДНК, состоящую из нуклеотидной последовательности, кодирующей пептид с активностью пуллуланазы с аминокислотной последовательностью, выбранной из группы, состоящей из SEQ ID NO: 2, 4 и 6, и ii) культуральной среды;

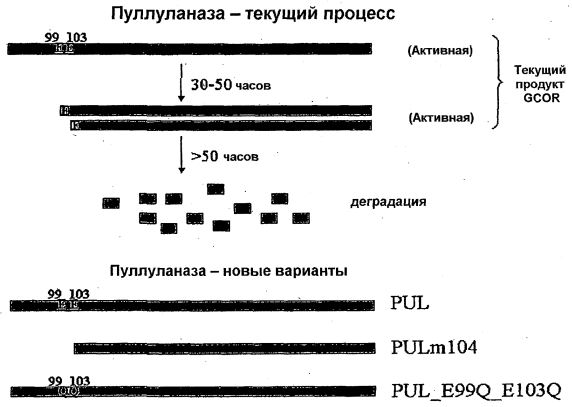
b) культивирование указанного трансфицированного организма хозяина в указанной культуральной среде в течение промежутка времени и в условиях, достаточных для продуцирования пептида с активностью пуллуланазы.

15. Способ по п.14, в котором указанный пептид выделяют из указанной культуральной среды.

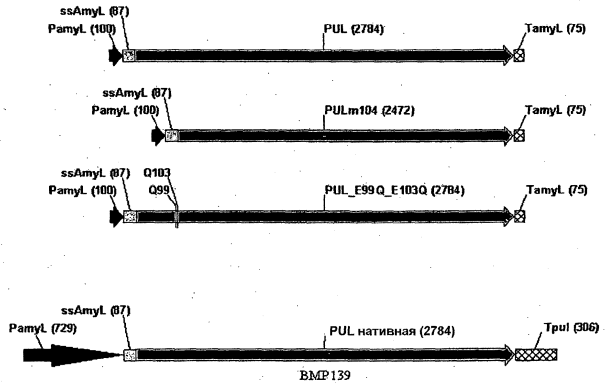
16. Способ по п.14, в котором указанный организм хозяина выбран из группы, состоящей из грибов, бактерий и эукариотических клеток.

17. Способ по п.16, в котором указанный организм хозяина выбран из группы, состоящей из Bacillus sp., Bacillus subtilis, Escherichia coli, Trichoderma reesei, Saccharomyces cerevisiae, Aspergillus niger и B. licheniformis.

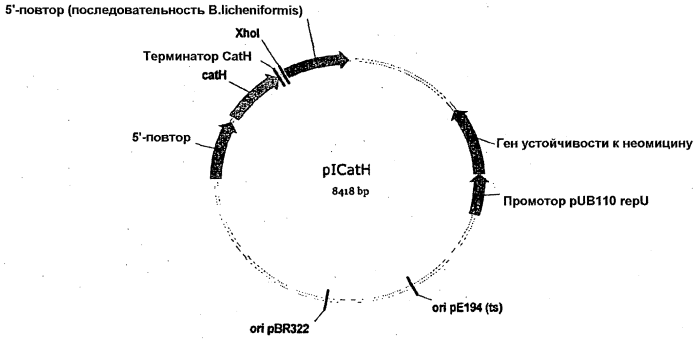
18. Комбинация штаммов для получения композиции по п.1, содержащая один или более штаммов B. licheniformis, приведенных в табл. 3 описания.



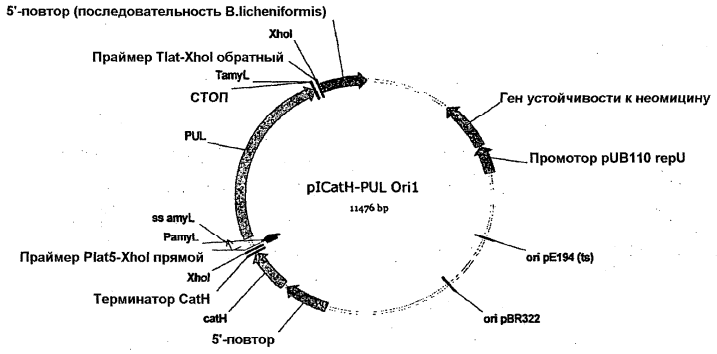
Фиг. 1



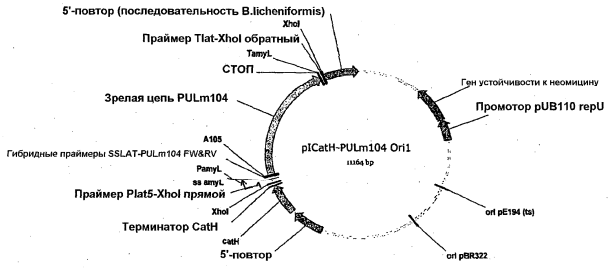
Фиг. 2



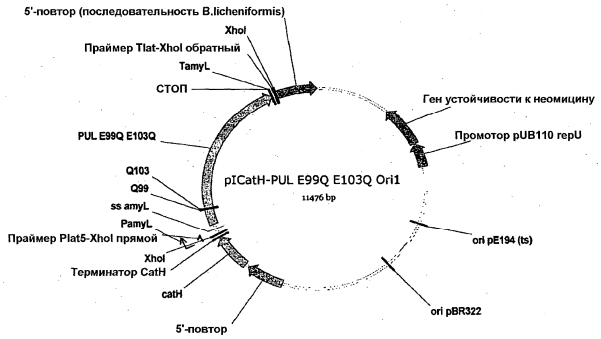
Фиг. 3



Фиг. 4



Фиг. 5



Фиг. 6

Нуклеотидная последовательность ДНК PUL дикого типа
ЖИРНЫЙ = *сигнальная последовательность amyL*
Черный = зрелая кодирующая последовательность PUL

atgaaacaacaaaaacggctttacgcccgattgctgacgctgttatttgcgctcatcttcttgctgcc
tcattctgcagcttcagcagatggcaatacgacgacgatcatcgtccattattttagaccggcgggag
attatcaaccgtggagcctttggatgtggccgaaagatggaggaggagcggaatatgattttaaccag
ccggcagattcacttggagcagtcgcctcagcagatattccgggaaatccgagccaagtcggcatcat
cgtcagaacacaggattggacgaaagatgtcagcgccgatcgctatatcgatctgagcaaaggcaatg
aagtctggctggtcgaaggcaacagccagatcttttatagcgaaaaagacgccgaagatgctgctaaa
ccggcagtcagcaaccgcttatctggatgccagcaaccaagtcctggtcaaactgagccaaccgctgac
acttggagaaggagcgagcggatttacggtccatgatgacacggcgaacaaagatatcccggtcacga
gcgttaaagatgctagcctgggccaagatgtcacagcagttctggcgggcacgtttcaacatatcttt
ggcggatcagattgggcaccggataatcacagcacgctgctgaaaaaagtcacgaacaacctgtatca
gtttagcggagatctgccggaaggcaactatcaatataaagtcgccctgaacgatagctggaacaatc
cgagctatccgagcgataacatcaatctgacagtcccggcaggcggagcacatgtcacgtttagctat
atcccgagcacacatgccgtctatgacacgatcaacaacccgaacgccgatcttcaagtcgaaagcgg
cgtcaaaacggatctggtcacagtcacattgggagaagatccggatgtcagccatacactgagcatcc
aaacggatggctatcaagcgaaacaagtcatcccgagaaacgtcctgaacagcagccagtattattat
agcggcgatgatctgggcaacacgtatacacaaaaagcgacgacgtttaaagtttgggcgccgacaag
cacacaagtcaacgtcctgctgtatgattcagcaacaggcagcgtcacaaaaatcgtcccgatgacag
catcaggacatggagtctgggaagcgacggtcaaccaaaacctggaaaactggtattatatgtatgaa
gtcacgggccaaggatcaacaagaacagcggtcgatccgtatgctacagcaatcgccccgaatggaac
aagaggcatgatcgtcgatctggcaaaaacagacccggcaggctggaatagcgataaacatatcacgc
cgaaaaacatcgaagatgaagtcatctatgaaatggacgtccgggattttagcatcgatccgaacagc
ggcatgaaaaacaaaggcaaatatctggcgctgacggaaaaaggaacaaaaggcccggataacgtcaa
aacaggcatcgatagcctgaaacaactgggcatcacacatgtccaactgatgccggtctttgctagca
atagcgtcgatgaaacggacccgacacaagataactggggctatgacccgagaaattatgatgtcccg
gaaggccaatatgccacgaacgccaatggaaacgcccggatcaaagaatttaaagaaatggtcctgag
ccttcatagagaacatatcggcgtcaacatggacgtcgtctataaccatacgtttgccacacagatca
gcgactttgataaaatcgtgccggaatattattatcggacggatgacgccggcaattatacgaatggc
agcggcacaggaaatgaaatcgccgccgaaagaccgatggtccagaaatttatcatcgacagccttaa
atatttgggtcaacgaatatcatatcgacggctttcgctttgatctgatggcgctgctgggcaaagata
caatgagcaaagcggcgagcgaacttcatgctatcaatccgggcatcgctctttatggagaaccgtgg
acaggaggaacatcagcactgccggatgatcaactgctgacaaaaggcgcccaaaaaggaatgggagt
cgccgtctttaacgacaacctgagaaatgccctggatggcaacgtttttgatagcagcgcccaaggat
ttgctacaggagcgacaggactgacagatgccatcaaaaatggcgtcgaaggcagcatcaacgatttt
acaagcagcccgggagaaacgatcaattatgtcacgagccatgacaactatacgctgtgggacaaaat
cgctctgagcaacccgaatgatagcgaagcggaccggatcaaaatggatgaactggcacaagcagtcg
tcatgacatcacaaggcgtcccgtttatgcaaggcggagaagaaatgctgagaacgaaaggcggcaac
gacaacagctataatgccggcgatgccgtcaatgaatttgactggagccggaaagcacaatatccgga
cgtctttaactattattcaggacttatccatctgagactggaccatccggcgtttagaatgacgacgg
cgaacgaaatcaacagccatcttcagtttctgaacagcccggaaaatacggtcgcctatgaactgacg
gaccatgtgaacaaagacaaatggggcaacatcatcgtcgtttataacccgaacaaaacggtcgccac
aatcaatcttccgagcggcaaatgggcaatcaatgccacaagcggcaaagttggagaaagcacactgg
gacaagcagaaggatcagtccaagtcccgggaatcagcatgatgatccttcatcaagaagtcagcccg
gaccacggcaaaaaa

Фиг. 7A

Белковая последовательность PUL дикого типа
ЖИРНЫЙ = сигнальная последовательность amyL
Черный = зрелая последовательность PUL

MKQQKRLYARLLTLLFALIFLLPHSAASADGNTTTIIVHYFRPAGDYQPWSLWMWPKDGGGAEYDFNQ
PADSLGAVASADIPGNPSQVGIIVRTQDWTKDVSADRYIDLSKGNEVWLVEGNSQIFYSEKDAEDAAK
PAVSNAYLDASNQVLVKLSQPLTLGEGASGFTVHDDTANKDIPVTSVKDASLGQDVTAVLAGTFQHIF
GGSDWAPDNHSTLLKKVTNNLYQFSGDLPEGNYQYKVALNDSWNNPSYPSDNINLTVPAGGAHVTFSY
IPSTHAVYDTINNPNADLQVESGVKTDLVTVTLGEDPDVSHTLSIQTDGYQAKQVIPRNVLNSSQYYY
SGDDLGNTYTQKATTFKVWAPTSTQVNVLLYDSATGSVTKIVPMTASGHGVWEATVNQNLENWYYMYE
VTGQGSTRTAVDPYATAIAPNGTRGMIVDLAKTDPAGWNSDKHITPKNIEDEVIYEMDVRDFSIDPNS
GMKNKGKYLALTEKGTKGPDNVKTGIDSLKQLGITHVQLMPVFASNSVDETDPTQDNWGYDPRNYDVP
EGQYATNANGNARIKEFKEMVLSLHREHIGVNMDVVYNHTFATQISDFDKIVPEYYYRTDDAGNYTNG
SGTGNEIAAERPMVQKFIIDSLKYWVNEYHIDGFRFDLMALLGKDTMSKAASELHAINPGIALYGEPW
TGGTSALPDDQLLTKGAQKGMGVAVFNDNLRNALDGNVFDSSAQGFATGATGLTDAIKNGVEGSINDF
TSSPGETINYVTSHDNYTLWDKIALSNPNDSEADRIKMDELAQAVVMTSQGVPFMQGGEEMLRTKGGN
DNSYNAGDAVNEFDWSRKAQYPDVFNYYSGLIHLRLDHPAFRMTTANEINSHLQFLNSPENTVAYELT
DHVNKDKWGNIIVVYNPNKTVATINLPSGKWAINATSGKVGESTLGQAEGSVQVPGISMMILHQEVSP
DHGKK

Фиг. 7B

Нуклеотидная последовательность PULm104
ЖИРНЫЙ = сигнальная последовательность amyL
Черный = зрелая кодирующая последовательность PULm104

atgaaacaacaaaaacggctttacgcccgattgctgacgctgttatttgcgctcatcttcttgctgcc
tcattctgcagcttcagcagctgctaaaccggcagtcagcaacgcttatctggatgccagcaaccaag
tcctggtcaaactgagccaaccgctgacacttggagaaggagcgagcggatttacggtccatgatgac
acggcgaacaaagatatcccggtcacgagcgttaaagatgctagcctgggccaagatgtcacagcagt
tctggcgggcacgtttcaacatatctttggcggatcagattgggcaccggataatcacagcacgctgc
tgaaaaaagtcacgaacaacctgtatcagtttagcggagatctgccggaaggcaactatcaatataaa
gtcgccctgaacgatagctggaacaatccgagctatccgagcgataacatcaatctgacagtcccggc
aggcggagcacatgtcacgtttagctatatcccgagcacacatgccgtctatgacacgatcaacaacc
cgaacgccgatcttcaagtcgaaagcggcgtcaaaacggatctggtcacagtcacattgggagaagat
ccggatgtcagccatacactgagcatccaaacggatggctatcaagcgaaacaagtcatcccgagaaa
cgtcctgaacagcagccagtattattatagcggcgatgatctgggcaacacgtatacacaaaaagcga
cgacgtttaaagtttgggcgccgacaagcacacaagtcaacgtcctgctgtatgattcagcaacaggc
agcgtcacaaaaatcgtcccgatgacagcatcaggacatggagtctgggaagcgacggtcaaccaaaa
cctggaaaactggtattatatgtatgaagtcacgggccaaggatcaacaagaacagcggtcgatccgt
atgctacagcaatcgccccgaatggaacaagaggcatgatcgtcgatctggcaaaaacagacccggca
ggctggaatagcgataaacatatcacgccgaaaaacatcgaagatgaagtcatctatgaaatggacgt
ccgggattttagcatcgatccgaacagcggcatgaaaaacaaaggcaaatatctggcgctgacggaaa
aaggaacaaaaggcccggataacgtcaaaacaggcatcgatagcctgaaacaactgggcatcacacat
gtccaactgatgccggtctttgctagcaatagcgtcgatgaaacggacccgacacaagataactgggg
ctatgacccgagaaattatgatgtcccggaaggccaatatgccacgaacgccaatggaaacgcccgga
tcaaagaatttaaagaaatggtcctgagccttcatagagaacatatcggcgtcaacatggacgtcgtc
tataaccatacgtttgccacacagatcagcgactttgataaaatcgtgccggaatattattatcggac
ggatgacgccggcaattatacgaatggcagcggcacaggaaatgaaatcgccgccgaaagaccgatgg
tccagaaatttatcatcgacagccttaaatattgggtcaacgaatatcatatcgacggctttcgcttt
gatctgatggcgctgctgggcaaagatacaatgagcaaagcggcgagcgaacttcatgctatcaatcc
gggcatcgctctttatggagaaccgtggacaggaggaacatcagcactgccggatgatcaactgctga
caaaaggcgcccaaaaaggaatgggagtcgccgtctttaacgacaacctgagaaatgccctggatggc
aacgtttttgatagcagcgcccaaggatttgctacaggagcgacaggactgacagatgccatcaaaaa
tggcgtcgaaggcagcatcaacgattttacaagcagcccgggagagacgatcaattatgtcacgagcc
atgacaactatacgctgtgggacaaaatcgctctgagcaacccgaatgatagcgaagcggaccggatc
aaaatggatgaactggcacaagcagtcgtcatgacatcacaaggcgtcccgtttatgcaaggcggaga
agaaatgctgagaacgaaaggcggcaacgacaacagctataatgccgggcgatgccgtcaatgaatttg
actggagccggaaagcacaatatccggacgtctttaactattattcaggacttatccatctgagactg
gaccatccggcgttagaatgacgacggcgaacgaaatcaacagccatcttcagtttctgaacagccc
ggaaaatacggtcgcctatgaactgacggaccatgtgaacaaagacaaatggggcaacatcatcgtcg
tttataacccgaacaaaacggtcgccacaatcaatcttccgagcggcaaatgggcaatcaatgccaca
agcggcaaagttggagaaagcacactgggacaagcagaaggatcagtccaagtcccgggaatcagcat
gatgatccttcatcaagaagtcagcccggaccacggcaaaaaa

Фиг. 8A

Белковая последовательность PULm104
ЖИРНЫЙ = сигнальная последовательность amyL
Черный = *зрелая последовательность* PULv104

MKQQKRLYARLLTLLFALIFLLPHSAASAAAKPAVSNAYLDASNQVLVKLSQPLTLGEGASGFTVHDD
TANKDIPVTSVKDASLGQDVTAVLAGTFQHIFGGSDWAPDNHSTLLKKVTNNLYQFSGDLPEGNYQYK
VALNDSWNNPSYPSDNINLTVPAGGAHVTFSYIPSTHAVYDTINNPNADLQVESGVKTDLVTVTLGED
PDVSHTLSIQTDGYQAKQVIPRNVLNSSQYYYSGDDLGNTYTQKATTFKVWAPTSTQVNVLLYDSATG
SVTKIVPMTASGHGVWEATVNQNLENWYYMYEVTGQGSTRTAVDPYATAIAPNGTRGMIVDLAKTDPA
GWNSDKHITPKNIEDEVIYEMDVRDFSIDPNSGMKNKGKYLALTEKGTKGPDNVKTGIDSLKQLGITH
VQLMPVFASNSVDETDPTQDNWGYDPRNYDVPEGQYATNANGNARIKEFKEMVLSLHREHIGVNMDVV
YNHTFATQISDFDKIVPEYYYRTDDAGNYTNGSGTGNEIAAERPMVQKFIIDSLKYWVNEYHIDGFRF
DLMALLGKDTMSKAASELHAINPGIALYGEPWTGGTSALPDDQLLTKGAQKGMGVAVFNDNLRNALDG
NVFDSSAQGFATGATGLTDAIKNGVEGSINDFTSSPGETINYVTSHDNYTLWDKIALSNPNDSEADRI
KMDELAQAVVMTSQGVPFMQGGEEMLRTKGGNDNSYNAGDAVNEFDWSRKAQYPDVFNYYSGLIHLRL
DHPAFRMTTANEINSHLQFLNSPENTVAYELTDHVNKDKWGNIIVVYNPNKTVATINLPSGKWAINAT
SGKVGESTLGQAEGSVQVPGISMMILHQEVSPDHGKK

Фиг. 8B

Нуклеотидная последовательность PUL E99Q_E103Q
ЖИРНЫЙ = сигнальная последовательность amyL
Черный = зрелая кодирующая последовательность PUL E99Q_E103Q

atgaaacaacaaaaacggctttacgcccgattgctgacgctgttatttgcgctcatcttcttgctgcc
tcattctgcagcttcagcagatggcaatacgacgacgatcatcgtccattattttagaccggcgggag
attatcaaccgtggagcctttggatgtggccgaaagatggaggaggagcggaatatgattttaaccag
ccggcagattcacttggagcagtcgcctcagcagatattccgggaaatccgagccaagtcggcatcat
cgtcagaacacaggattggacgaaagatgtcagcgccgatcgctatatcgatctgagcaaaggcaatg
aagtctggctggtcgaaggcaacagccagatcttttatagcgaaaaagacgccgaagatgctgctaaa
ccggcagtcagcaacgcttatctggatgccagcaaccaagtcctggtcaaactgagccaaccgctgac
acttggagaaggagcgagcggatttacggtccatgatgacacggcgaacaaagatatcccggtcacga
gcgttaaagatgctagcctgggccaagatgtcacagcagttctggcgggcacgtttcaacatatcttt
ggcggatcagattgggcaccggataatcacagcacgctgctgaaaaaagtcacgaacaacctgtatca
gtttagcggagatctgccggaaggcaactatcaatataaagtcgccctgaacgatagctggaacaatc
cgagctatccgagcgataacatcaatctgacagtcccggcaggcggagcacatgtcacgtttagctat
atcccgagcacacatgccgtctatgacacgatcaacaacccgaacgccgatcttcaagtcgaaagcgg
cgtcaaaacggatctggtcacagtcacattgggagaagatccggatgtcagccatacactgagcatcc
aaacggatggctatcaagcgaaacaagtcatcccgagaaacgtcctgaacagcagccagtattattat
agcggcgatgatctgggcaacacgtatacacaaaaagcgacgacgtttaaagtttgggcgccgacaag
cacacaagtcaacgtcctgctgtatgattcagcaacaggcagcgtcacaaaaatcgtcccgatgacag
catcaggacatggagtctgggaagcgacggtcaaccaaaacctggaaaactggtattatatgtatgaa
gtcacgggccaaggatcaacaagaacagcggtcgatccgtatgctacagcaatcgccccgaatggaac
aagaggcatgatcgtcgatctggcaaaaacagacccggcaggctggaatagcgataaacatatcacgc
cgaaaaacatcgaagatgaagtcatctatgaaatggacgtccgggatttagcatcgatccgaacagc
ggcatgaaaaacaaaggcaaatatctggcgctgacggaaaaaggaacaaaaggcccggataacgtcaa
aacaggcatcgatagcctgaaacaactgggcatcacacatgtccaactgatgccggtctttgctagca
atagcgtcgatgaaacggacccgacacaagataactggggctatgacccgagaaattatgatgtcccg
gaaggccaatatgccacgaacgccaatggaaacgcccggatcaaagaatttaaagaaatggtcctgag
ccttcatagagaacatatcggcgtcaacatggacgtcgtctataaccatacgtttgccacacagatca
gcgactttgataaaatcgtgccggaatattattatcggacggatgacgccggcaattatacgaatggc
agcggcacaggaaatgaaatcgccgccgaaagaccgatggtccagaaatttatcatcgacagccttaa
atattgggtcaacgaatatcatatcgacggctttcgctttgatctgatggcgctgctgggcaaagata
caatgagcaaagcggcgagcgaacttcatgctatcaatccgggcatcgctctttatggagaaccgtgg
acaggaggaacatcagcactgccggatgatcaactgctgacaaaaggcgcccaaaaaggaatgggagt
cgccgtctttaacgacaacctgagaaatgccctggatggcaacgtttttgatagcagcgcccaaggat
ttgctacaggagcgacaggactgacagatgccatcaaaaatggcgtcgaaggcagcatcaacgatttt
acaagcagcccgggagaaacgatcaattatgtcacgagccatgacaactatacgctgtgggacaaaat
cgctctgagcaacccgaatgatagcgaagcggaccggatcaaaatggatgaactggcacaagcagtcg
tcatgacatcacaaggcgtcccgtttatgcaaggcggagaagaaatgctgagaacgaaaggcggcaac
gacaacagctataatgccggcgatgccgtcaatgaatttgactggagccggaaagcacaatatccgga
cgtctttaactattattcaggacttatccatctgagactggaccatccggcgttagaatgacgacgg
cgaacgaaatcaacagccatcttcagtttctgaacagcccggaaaatacggtcgcctatgaactgacg
gaccatgtgaacaaagacaaatggggcaacatcatcgtcgtttataacccgaacaaaacggtcgccac
aatcaatcttccgagcggcaaatgggcaatcaatgccacaagcggcaaagttggagaaagcacactgg
gacaagcagaaggatcagtccaagtcccgggaatcagcatgatgatccttcatcaagaagtcagcccg
gaccacggcaaaaaa

Фиг. 9A

Белковая последовательность PUL E99Q_E103Q
ЖИРНЫЙ = сигнальная последовательность amyL
Черный = зрелая последовательность PUL E99Q_E103Q

MKQQKRLYARLLTLLFALIFLLPHSAASADGNTTTIIVHYFRPAGDYQPWSLWMWPKDGGGA
EYDFNQPADSLGAVASADIPGNPSQVGIIVRTQDWTKDVSADRYIDLSKGNEVWLVEGNSQI
FYSQKDAQDAAKPAVSNAYLDASNQVLVKLSQPLTLGEGASGFTVHDDTANKDIPVTSVKDA
SLGQDVTAVLAGTFQHIFGGSDWAPDNHSTLLKKVTNNLYQFSGDLPEGNYQYKVALNDSWN
NPSYPSDNINLTVPAGGAHVTFSYIPSTHAVYDTINNPNADLQVESGVKTDLVTVTLGEDPD
VSHTLSIQTDGYQAKQVIPRNVLNSSQYYYSGDDLGNTYTQKATTFKVWAPTSTQVNVLLYD
SATGSVTKIVPMTASGHGVWEATVNQNLENWYYMYEVTGQGSTRTAVDPYATAIAPNGTRGM
IVDLAKTDPAGWNSDKHITPKNIEDEVIYEMDVRDFSIDPNSGMKNKGKYLALTEKGTKGPD
NVKTGIDSLKQLGITHVQLMPVFASNSVDETDPTQDNWGYDPRNYDVPEGQYATNANGNARI
KEFKEMVLSLHREHIGVNMDVVYNHTFATQISDFDKIVPEYYYRTDDAGNYTNGSGTGNEIA
AERPMVQKFIIDSLKYWVNEYHIDGFRFDLMALLGKDTMSKAASELHAINPGIALYGEPWTG
GTSALPDDQLLTKGAQKGMGVAVFNDNLRNALDGNVFDSSAQGFATGATGLTDAIKNGVEGS
INDFTSSPGETINYVTSHDNYTLWDKIALSNPNDSEADRIKMDELAQAVVMTSQGVPFMQGG
EEMLRTKGGNDNSYNAGDAVNEFDWSRKAQYPDVFNYYSGLIHLRLDHPAFRMTTANEINSH
LQFLNSPENTVAYELTDHVNKDKWGNIIVVYNPNKTVATINLPSGKWAINATSGKVGESTLG
QAEGSVQVPGISMMILHQEVSPDHGKK

Фиг. 9B